Studies on the acorn barnacle *Balanus amphitrite* and its associated bacteria

By Tristano Bacchetti De Gregoris

Submitted for the Degree of Doctor of Philosophy at

School of Marine Science and Technology

Submitted in Jan 2011

Abstract

Despite being a model organism to study settlement in marine invertebrates, little is known about the genetics of the barnacle *Balanus amphitrite*. To fill this gap, cDNA libraries representative of different developmental stages were generated and sequenced. Nearly 14,000 genes were annotated, which may represent 2/3 of the species' total protein coding regions. The database that was created to allow public access to this genetic information will profoundly benefit future research aiming to understand the molecular regulation of development and settlement in this species.

Furthermore, a quantitative real-time PCR assay to study gene expression in *B. amphitrite* was designed and validated. Eleven genes were studied for their ability to normalize qRT-PCR data. Total RNA extracted from seven developmental stages was reverse transcribed and the expression stability of the selected genes was compared. It was found that transcripts encoding cytochrome b and NADH dehydrogenase subunit 1 were expressed most stably, and their use to normalize gene expression data is recommended.

Conflicting evidence exists on the role of bacteria in *B. amphitrite* settlement. However, there is a paucity of information on the microbial community naturally associated with this barnacle. In order to reveal the existence of stable associations, a 16S rRNA-based, taxon-specific qPCR assay was developed to monitor the preponderance of 5 bacterial phyla and classes. Furthermore, attempts to profile these qPCR products by DGGE were made. This new method was applied to characterise the bacterial communities associated with different *B. amphitrite* developmental stages and body parts. It was found that the structure of these communities changed throughout the barnacle life cycle in a highly reproducible manner. Furthermore, bacteria isolated from the barnacle shell were capable of inducing settlement of conspecific larvae. The analysis of these communities at a lower taxonomic level should confirm if any of these ecologically important bacteria are vertically transmitted.

To the Star and Shadow community

Acknowledgements

I would like to thank the Natural Environment Research Council, UK, for awarding me the PhD fellowship NER/S/A/2005/13338, which supported me and most of my research. Additional financial support was devolved to this project by a number of organisations. These include The Boehringer Ingelheim Fonds; The European Community, through the Marine Genomics Network; The Office of Naval Research, USA; and The Prime Minister Initiative, UK. It would not have been possible to complete this thesis without their contributions.

I am grateful to my PhD supervisors Tony Clare and Grant Burgess. The combination of their qualities built up the perfect mentor, which has made my research at Newcastle enjoyable and interesting.

I am also indebted to the many researchers that have helped me acquiring fundamental technical competencies and a solid scientific approach. Among these, a special thank goes to Nick Aldred, Sheelagh Conlan, Andreia Cavaco, Jinwei Zhang, Reindert Nijland, Tahar Taybi, Jeremy Barnes, Michael Maguire, Ian Head, Fiona Read, Tom Curtis, Ehsan Mesbahi and Lisa Inganni, all at Newcastle University. Also, I am grateful to Bernardo da Gama from the Universidade Federal Fluminense, Brazil, to Dan Rittschof, Kevin Zelnio and Tom Schultz from the Duke Marine Laboratory, USA, to AC Anil and Lidita Khandeparker from the National Institute of Oceanography, India, to Oliver Rupp, Thomas Bekel and Alexander Goesman from Bielefeld University, Germany, to Sven Klages, Florian Knaust, Michael Kube and Reinhardt Richard from the Max-Plank Institute, Germany, and to Ina Arnone, Marco Borra and Elio Biffali from the Stazione Zoologica A. Dohrn, Italy.

Table of contents

Figure index

Supplementary figures

Table index

Supplementary tables

Chapter 1. Introduction

1.1 Biofouling: historical and environmental considerations

The biodeterioration of engineered structures by marine organisms is a problem that dates back to the early days of navigation. In their collection of Greek and Latin documents, Humphrey and colleagues (Humphrey*, et al.*, 1998) report that ship hulls were commonly pitched and coated with wax-based paints to retard fouling and rot. Since then, corrosion of submerged material and reduced performance of vessels due to the attachment of benthic organisms has encouraged the development of new technologies to protect marine infrastructures. The first patent for an antifouling coating was granted to William Beale in 1625 (Masseille, 1934) and the majority of early ship-bottom patents referred to some sort of metallic sheathing. Among the various materials tested, copper sheathing has been the most widely used throughout the 18^{th} and 19^{th} centuries (WHOI, 1952). Although the main purpose of this covering layer was to defend wooden hulls from burrowing by the teredo worm (Townsin, 2003), it soon became evident that copper was extremely effective in controlling biofouling development. The action of copper in keeping hulls clean is thus stated by the 'Engineer' of December 21st, 1866: "The beneficial action which continuously keeps a copper bottom from sea-weed is evidently due to a double process of exfoliation and of a generation of a poison (verdigris) inimical to the development of organic life" (reported in Young, 1867). However, this antifoulant technique was not to last.

With the advent of iron ships, the application of copper sheathing had a major drawback because of the accelerated corrosion of the iron and copper combination. The eminent chemist Sir Humphrey Davy helped to investigate this phenomenon and he demonstrated that this electrochemically driven corrosion could be stopped with the addition of zinc plaques (this process is nowadays known as “cathodic protection”). Despite successfully preserving the metals, the copper layer was quickly colonized by algae and barnacles, suggesting that the antifouling properties of this element were related with its gradual release into the water. This idea was developed further by Holzapfel, who in 1904 published the results of his tests on 100 different coatings in the port of Genoa (Holzapfel, 1904). He found that the more successful coatings were those with a high content of copper and mercury, but the most important feature of the varnish was its ability to slowly dissolve in sea water. This discovery became the milestone for the majority of innovations in antifouling research in the twentieth century, and the

whole idea of a matrix capable of gradually releasing bioactive compounds is still at the forefront of antifouling technologies (Yebra*, et al.*, 2004; Chambers*, et al.*, 2006).

Despite major expansions of both merchant and naval fleets in the 20th century, it was only in the 1970s that Holzapfel's findings were exploited practically (Milne, 1975). This happened with the introduction of self-polishing copolymer (SPC) based paint, where there is continual removal of the surface layer by polymer hydrolysis or ion exchange when the paint is in contact with sea water (Kiil*, et al.*, 2001; 2002). The first SPC commercialized was based on organotin compounds (particularly tributyltin, TBT) and the coating obtained was the unchallenged leader in the antifouling technology market for more than twenty years (Yebra*, et al.*, 2004). The gradual hydrolysis of the matrix allowed the slow release of toxic molecules in the water column. TBT is a bioactive compound, effective against a broad range of organisms, which inhibits the growth of unwanted biological material on ship hulls and other coated material (Champ & Seligman, 1996; Omae, 2003; Pellerito*, et al.*, 2006). Despite the initial promise, TBT degraded slowly in aquatic ecosystems (Seligman*, et al.*, 1996) and its dangerous accumulation in marine environments has been reported (Hall*, et al.*, 1985; Hall*, et al.*, 1987; Nogueira*, et al.*, 2003). Particularly, high levels of TBT have been found in mussel tissues (Oehlmann*, et al.*, 1998; Jacobsen & Asmund, 2000; Strand & Asmund, 2003; Meng*, et al.*, 2005) where it can cause morphological deformities in natural and cultivated populations of oysters (Alzieu*, et al.*, 1986; Pereira*, et al.*, 1999; Alzieu, 2000). Furthermore, it is feared that once organotins have entered the food web, they can become increasingly bioaccumulated in the higher trophic level (Kannan*, et al.*, 1996; Traas*, et al.*, 1996; Harino*, et al.*, 2005; Hu*, et al.*, 2006). Environmental problems posed by employing these substances have, therefore, led to legislation that limits their use on ship hulls, including a total ban from 2008 (Champ, 2000; Terlizzi*, et al.*, 2001; Roberts & Tsamenyi, 2008; IMO 2001). The industry has accepted that TBT has no future and has returned to the use of copper SPC paints boosted with biocides, usually herbicides, to control soft fouling that copper is less effective against (Voulvoulis*, et al.*, 1999). The EU Biocidal Product Directive will likely limit the number of biocides that will be permissible and as a result there is great interest in finding new non-toxic or at least environmentally benign solutions to fouling (Clare, 1998; Chambers*, et al.*, 2006). Although some alternatives to TBT-based coatings are already available (Yebra*, et al.*, 2004; Giriyan & Pangam, 2006), it is unlikely that a unique technology will be effective in both repelling marine organisms and being environmentally friendly. Nevertheless, an increased understanding of the biological

events leading to fouling establishment could help in developing techniques to fight the attachment of the most economically relevant species. With this in mind, this study was designed to extend our understanding of the life cycle of the acorn barnacle *Balanus amphitrit*e (syn. *Amphibalanus amphitrite*, Clare & Høeg, 2008) and, particularly, to test the ability of bacteria to influence development and settlement in this species.

1.2 Barnacle life cycle

Barnacles are crustaceans that live their adult stage enclosed in a calcareous shell. They are related to organisms like lobsters, shrimp and copepods, and belong to the phylum Arthopoda, which includes Hexapoda (terrestrial insects), Chelicerata (spiders), Myriapoda (millipedes), Xenocarida and Oligostraca. The evolutionary relationship between arthropods has long been debated, as some morphological characters do not fit with molecular data (Boxshall, 2007). Particularly, it was unknown whether insects were the sister group to a monophyletic Crustacea or a subset of crustaceans. The sequences of 62 single-copy nuclear protein-coding genes from 75 arthropod species were recently used to draw a reconciliatory tree (Regier*, et al.*, 2010), which suggested that the Hexapoda is a monophyletic group which diverged from 'true crustaceans' earlier than previously believed (Nardi*, et al.*, 2003). It also pointed out that the Maxillopoda, the class that includes barnacles, is a polyphyletic group. For this reason, it is expected that barnacle taxonomy may undergo some minor rearrangements following Regier *et al.* (2010). Nevertheless, the aim of this section is to discuss the biology of barnacles in order to highlight the possible mechanisms of interaction with bacteria, and barnacle taxonomy will be mentioned to pinpoint the potential evolution of these mechanisms.

Although all barnacles share a similar developmental fate, morphological and adaptive differences exist within this group. Therefore, unless otherwise stated, the model organism *Balanus amphitrite* (syn. *Amphibalanus amphitrite,* Clare & Høeg, 2008) will be used to draw the general life cycle of barnacles. *B. amphitrite* is a sessile, gregarious cirripede that predominantly inhabits intertidal rocky shores. It is an invasive species that has a worldwide distribution and is found in tropical/subtropical environments. Adult barnacles live enclosed in a calcareous shell (a cross-sectional diagram of an adult barnacle is presented in Figure 1.1) whose outermost structure is a thin cuticular membrane in continuity with the base, also known as cuticular slips, which seals off the internal parts of the animal from the external medium (Bourget,

1977). However, no research has been carried out on its possible function in interacting with the external environment (i.e. the biofilm). Interestingly, it was shown that *B. amphitrite* has one of the highest organics/weight of shell ratios among barnacles (Bourget, 1987). Little is known about the formation of the parietal and basal membranes, and more attention has been given to the formation of shell plates. In some barnacles, a thin layer of calcified shell is deposited at every tidal cycle (Crisp & Richardson, 1975). This is a consequence of calcium availability, as this element is not stored and becomes depleted in the animal during emersion. However, shell growth has also been shown to relate to the moulting cycle and some environmental stressors (Crisp & Richardson, 1975). *B. amphitrite* shell is composed of 8 calcified plates filled with hollow canals which are believed to reduce erosion and boring. The operculum is made of 4 extra plates whose movement is controlled by depressor and adductor muscles attached to the wall plates. Through the opening of these two pairs of plates, the adults can project parts of its body outside the shell for feeding and mating.

Generally, barnacle reproduction is characterised by simultaneous hermaphroditism, although other sexual patterns exist. It has been suggested that in some species, copulation occurs when a recently moulted adult, receptive as female, signals chemically 'her' status to the neighbours (Barnes*, et al.*, 1977). This induces adjacent individuals to start searching and when the receptive adult is found, the 'male' deposits sperm into the mantle cavity of the female. In *Semibalanus balanoides*, multiple copulations can occur with a single female and a single brood may therefore be fertilised by multiple individuals. Soon after copulation, the eggs are laid in two elastic sacs in the mantle cavity, where they remain until hatching (Walley*, et al.*, 1971), and it was suggested that a stimulus to oviposition is provided by the seminal fluid (Anderson, 1994). Certain species (e.g. *Semibalanus balanoides*) are capable of brooding the eggs for several months and generally reproduce once a year.

Regulation of hatching is best described in *S. balanoides*, possibly due to the synchronous release of the progeny by adults belonging to the same level on shore. The hatching is coincident with the spring algal bloom and it was shown that naupliar release occurs after feeding (Crisp, 1956). It was initially hypothesised that the adult accumulates a compound derived from diatoms which then acts as an excretory metabolite to promote egg hatching (Barnes, 1957). However, this compound is a product of the barnacle's own metabolism (Clare*, et al.*, 1984), and the presence of the algal bloom likely triggers its release as it is advantageous for the naupliar larvae to begin their development in a food-rich environment. The hatching substance possesses

comparable chemical proprieties to that of eicosanoids (Clare*, et al.*, 1982; Clare*, et al.*, 1984; Vogan*, et al.*, 2003). However, it appeared that this pheromone possesses an indirect effect on hatching, which is to cause the release of dopamine from within the embryo, which then stimulates strong muscular movements causing the eggs to hatch (Clare*, et al.*, 1984).

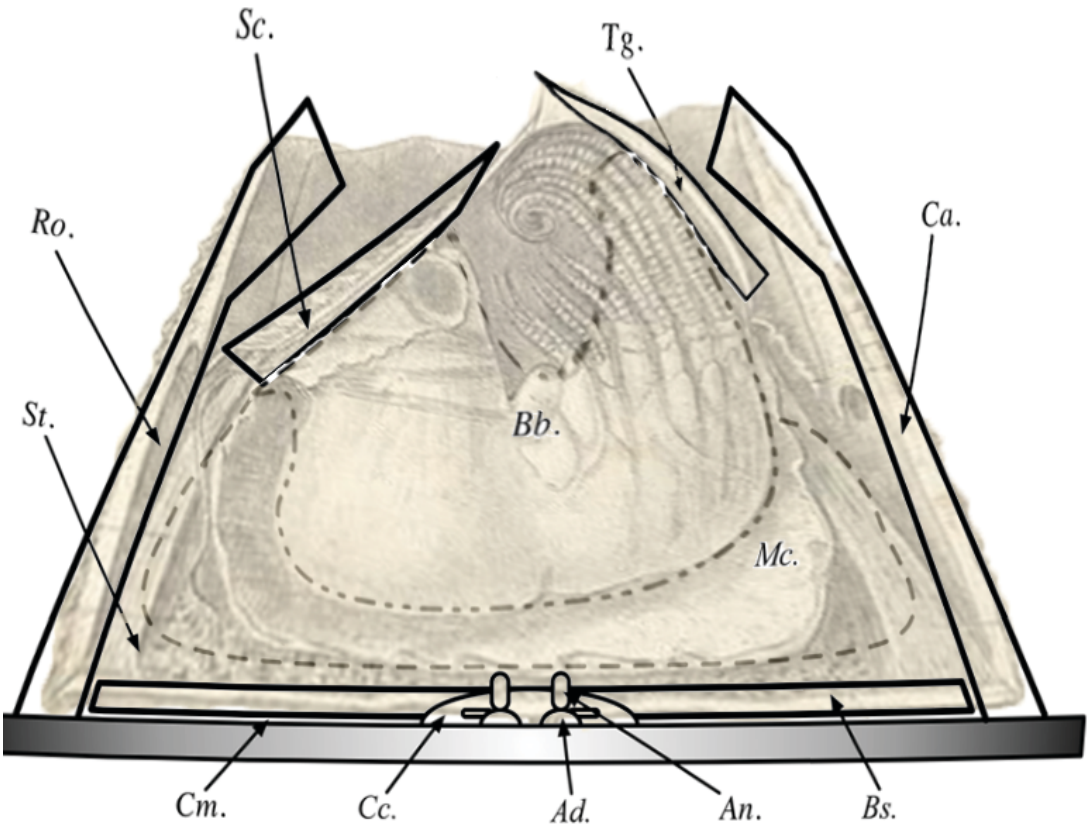


Figure 1.1 - **Cross-sectional diagram of an adult barnacle**

Drawing of a barnacle on a substratum (taken from Darwin, 1851), with important structural features highlighted (as in Aldred & Clare, 2008). Abbreviations are: *Ro*., the rostrum shell plate; *Ca*., the carina shell plate; *Sc*., the scutum opercular plate; *Tg*., the tergum opercular plate; *An*., the cyprid antennules embedded in *Cc*., cyprid cement with *Ad*., the adhesive discs viewable from beneath; *Cm*., the adult cement (not yet present in the juveniles) is secreted into the interface between *Bs*., the basis or base plate (calcified in a mature *B. amphitrite*, membraneous in a juvenile) and the glass surface; *St*., soft tissue, musculature and ovarian tissue; *Mc*., the mantle cavity containing the body of the barnacle and *Bb*., the barnacle body itself. *Ro*., *Sc*., *Tg*., *Ca*. and *Bs*. increase their calcification from juvenile toward mature adult.

The eggs hatch as stage I nauplii and are expelled from the mantle cavity to begin their planktonic phase. At this point in the life cycle, which is schematically represented in Figure 1.2, larvae need to feed and accumulate reserve energy which will then be

used by the non-feeding, settlement stage cyprid. There are 6 naupliar stages and, although morphological differences exist between instars, the feeding dynamics and behaviour appear to remain similar (Walker*, et al.*, 1987). Larval responses to light and pressure, combined with locomotor abilities, are used to remain in the upper water layer where most of the phytoplankton is located. However, swimming capabilities are relatively poor and larval dispersal has more to do with tidal- and wind-driven currents (De Wolf, 1973).

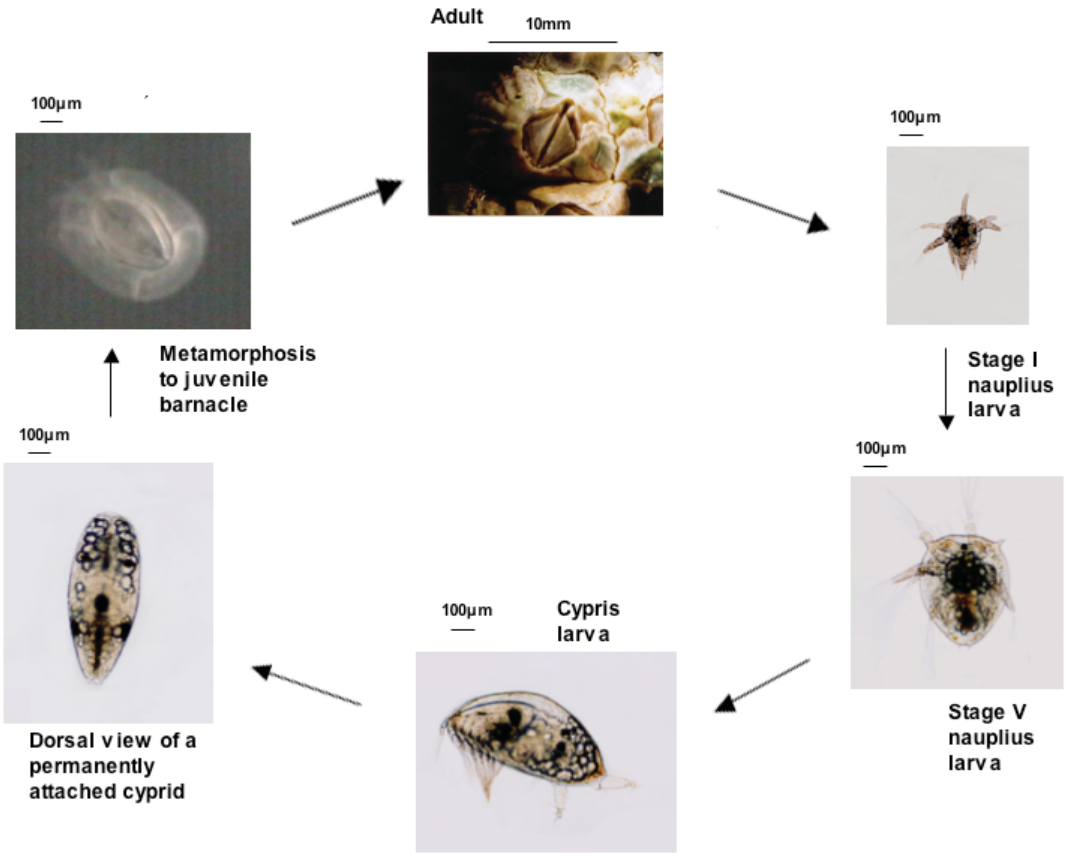


Figure 1.2 - ***Balanus amphitrite*'s life cycle**

Six well-distinct life stages are presented. This are arguebly the most important phases during barnacle development, but the egg stage and four more napliar instars (of a total of 6) exist.

It is interesting to mention the feeding strategy of nauplii in relation to the existence of chemoreceptors involved in microbial sensing. Larvae possess a ventral mouth between a pair of biramous antennae and a pair of biramous mandibles. In caudal position from the mouth, the ventral surface is covered by setae which are used to trap food particles and to move them towards the mouth (Lochhead, 1936). Interestingly, swimming speed, which is correlated with the capacity of capturing microalgae, positively correlates with the concentration of algal cells in the surrounding water.

Although the mechanism by which the nauplii detect the presence of food has not been completely elucidated, Yule (reported in Walker*, et al.*, 1987) showed that 10^{-4} M of glycine, glutamic acid, glucose and sucrose caused an increase in antennule swept volume. All these substances are known to be part of the excretory products of phytoplankton, and it was suggested that their detection in the water column may inform larvae about food content (Walker*, et al.*, 1987). Although chemosensitive organs have not been found around the mouth, it was shown that the dorsal carapace is likely equipped with a number of chemoreceptors (Walker & Lee, 1976), and the existence of sensory mechanisms utilised for the detection of compounds of microbial origin for larval feeding suggests that the cyprid may retain this ability in order to better monitor the environment during its quest for a settlement site. However, possible interactions between barnacles and bacteria in the process of settlement will be outlined in later sections of this thesis, and the attention is focused here on the development of morphological features that could be involved in such interactions.

After the naupliar stage VI, a single moult gives rise to the cyprid, the last planktonic stage, which is unique to barnacles. Whereas in other crustaceans the morphological transformation between the larva and the juvenile is carried out more gradually, in cirripedes this occurs through a rapid metamorphosis. Many features of nauplii, such as mouthparts, gut and segmented abdomen are missing, while the cyprid acquires specific structures involved in the settlement process (Anderson, 1994). Interestingly, the body organisation, size and behaviour during this stage are widely conserved in cirripedes, and generalisations can be made from single species investigations. The cyprid expresses the nearest approximation to a bivalved, free swimming maxillopodan, where most of the body becomes enclosed in a carapace that develops from the head shield of earlier stages (Glenner, 1999). Its fusiform shape allows a greater swimming agility in comparison to that of a nauplius, and the power generated by six pairs of thoracic limbs can propel the cyprids to a speed of up to 95 body lengths per second (Yule, 1984). The antennules, organs that are unique to this larval stage, protrude from the roof of the anterior mantle cavity and possess fundamental functions in settlement. They are composed of four segments and they enable ‘walking’, substrate discrimination and adhesion. The first two segments are believed to be mainly involved in movement, while the 3rd and 4th function in sensing and attaching to the substratum (Glenner, 1999). The third antennular segment ends with a bell-shaped attachment disc covered with a carpet of cuticular villi. Important receptors like the axial and postaxial sense organs are located on the disc. Furthermore,

an array of glands, including the cement gland, opens between the villi, allowing temporary and permanent attachment to take place (Walker & Yule, 1984; Walker*, et al.*, 1987). The fourth segment projects at the side of the attachment disc and is composed of a short mobile element that bears a variety of sensory setae, which in most species can be long and ramified. In spite of the growing morphological evidence that some antennular structures are devoted to sensing settlement cues, little is known about their ability to detect stimuli to which the cyprid responds. Nevertheless, the flicking movements of the fourth segment while the animal is anchored to a surface with the attachment disc, have supported the belief that this segment plays a fundamental role in substrate discrimination (Nott & Foster, 1969; Glenner, 1999).

Cyprids that are developmentally competent to settle display an adaptive ability to delay the pelagic-benthic transition, and settlement has been shown to be influenced by a variety of environmental stimuli, which are often species-specific. In *B. amphitrite,* factors such as larval history and cyprid age (Holm, 1990; Satuito*, et al.*, 1997), surface topology (Rittschof*, et al.*, 1984; Faimali*, et al.*, 2004), water flow (Mullineaux & Butman, 1991) and the presence of microbial film (Maki*, et al.*, 1988; Olivier*, et al.*, 2000) are all capable of affecting recruitment to a certain extent. However, in cross-fertilizing species such as *B. amphitrite*, the main force that drives settlement is the need of cypris larvae to attach close to individuals that will allow reproduction to take place (Crisp & Knight-Jones, 1953). For this reason, extensive research has been carried out in order to identify the chemical basis of conspecific recognition (Crisp & Meadows, 1962; Clare & Matsumura, 2000). As a result, a settlement-inducing protein complex (SIPC) was isolated from *B. amphitrite* (Matsumura*, et al.*, 1998a) and shown to be responsible for gregariousness (Dreanno*, et al.*, 2006b). The increased effect of SIPC when bound to a surface and the fact that its expression is localised to the cuticle (Dreanno*, et al.*, 2006a) suggests that it acts as a contact pheromone. SIPC is a high molecular weight protein composed of four subunits of 32, 76, 88 and 98 kDa respectively. The three larger subunits have been shown to link uncharacterised sugar chains (Matsumura*, et al.*, 1998a), making SIPC a large glycoprotein (Dreanno*, et al.*, 2006b). The addition of lectins to settlement-inducing adult extract was reported to block the promotory effect of SIPC, suggesting that the cue that cyprids respond to is located in its carbohydratic part (Matsumura*, et al.*, 1998b). Specifically, the effect was triggered by mannose-binding lectin (LCA and ConA) but not by galactose/N-acetylgalactosamine- or N-acetylglucosamine-binding lectin. Interestingly, a similar lectin-specific effect has been shown to block the settlement promotion induced by

certain microbial biofilms (Khandeparker*, et al.*, 2003). However, an in-depth analysis of bacteria-barnacle interaction is the focus of section 1.4.

Once the cyprid has located an optimal location for settlement, the release of proteinaceous substances allows the barnacle to strongly adhere to the substrate. These substances are produced in the cement glands whose formation has been studied in some detail (Nott & Foster, 1969; Walker, 1971). The mixture of proteins, phenol and phenoloxidase is first accumulated in the ducts and then completely secreted through the attachment disc when settlement is triggered. Over a period of a few hours the deposited cement hardens into a quinine-tanned protein that will trap the distal segments of the antennules. At this point, the cyprid initiates a complete rearrangement of the body in order to metamorphose into its juvenile stage. This complex transformation mostly takes place after the cyprid moults, when the freed hypodermis reshapes to the new organization before the juvenile cuticle hardens (Walley, 1969). At the end of this process, many characteristics of adult barnacles, such as the posture and the formation of a peduncle in pedunculate thoracicans or a basis in sessile forms, are already acquired, and the juvenile/adult growth pattern starts immediately. From the cyprid carapace are derived parts of the wall plates while others are laid down *in situ* (Anderson, 1994). Therefore, the plates that will form the external shell initially shape as chitinous rudiments and are calcified as development progresses. As can be expected, a lot of phagocytosis takes place in order to recycle disrupted cyprid structures and to permit the synthesis of a new musculature, of the missing past of the alimentary tract and new nervous and sensory systems (Walley, 1969). The cement glands of the larva are also reabsorbed and new glands develop from cells associated with the cement ducts (Yule & Walker, 1987). This will allow functional continuity as the barnacle uses the same glands to maintain adhesion throughout their life, although there is no evidence that cyprid and adult cement are related.

The last anatomical element of barnacles that need to be mentioned regarding the possible existence of intimate interactions with microbes is their digestive apparatus. *Balanus amphitrite* is a suspension feeder that protrudes its cirri outside of the shell to capture prey and other suspended particles. Like other crustaceans, its digestive tract is divided into a foregut (stomach), midgut (intestine) and hindgut (rectum), and undigested food is expelled in the form of faecal pellet enclosed in a membrane produced in the midgut (Anderson, 1994). The existence of intimate associations between barnacles and the microbial community inhabiting their guts has not been investigated. Although a comparative analysis on the functions that bacteria inhabiting

the gut of *B. amphitrite* may perform will be presented in section 1.4., it can be expected that a variety of microbial species contribute to barnacle health by participating in the degradation of food particles and by providing the host with fundamental nutritional requirements such as vitamins.

In conclusion, the existence of specific functions carried out by bacteria during *B. amphitrite* life cycle is unknown, as well as the presence of specific barnacle-bacteria associations, but many possible scenarios can be suggested from the study of barnacle anatomy and behaviour. In order to provide the reader with a solid background to understand the reasoning behind this study, the next important field to review is that concerning the bacterial counterpart. Specifically, the next section presents an overview on the system by which bacteria are classified and on the main variables that influence their distribution in the environment.

1.3 Microbial diversity

Since life most likely began in the ocean over 3.5 billion years ago (Nisbet & Sleep, 2001), for a long time this habitat was the only one to nurture the evolution of biochemical pathways. Photosynthesis for example, first accomplished by marine microbes, was responsible for the rise of atmospheric oxygen level, which consequently allowed the coming of organotrophic metabolism and the colonization of Earth. Nowadays, aquatic environments are still accountable for ~50% of global oxygen production (Field*, et al.*, 1998) and this can help us to understand the importance these environments have on the biogeochemical cycling that shapes the condition in which all other life forms have evolved. Therefore, marine microorganisms are not only considered the closest living descendents of the original form of life, but they also represent a major pillar that sustains the modern biosphere.

From the 'common ancestor', which rather than a single cell, was probably a population of very simple replicating organisms sharing genetic information (Woese, 1998), divergence crystallized as evolution favored more specific and efficient proteins and genes became less exchangeable. It is widely accepted that from this aquatic 'universal ancestor' independently arose the two domain Bacteria and Archaea (Pace, 1997), and, few hundred million years down the evolutionary path, symbiosis between members of this two domains is thought to have given birth to the domain Eucarya (Woese*, et al.*, 1990). As a result of this process, marine habitats are expected to harbour a major portion of global biodiversity. Furthermore, most of the many possible

ways different organisms interact with each other, from metabolically to behaviourally, have evolved among species inhabiting the oceans, beginning with microbial cell-to-cell communication and evolving into more complex inter-kingdom interactions. However, before attempting to investigate how invertebrates, and particularly barnacles, may be affected by the presence of microorganisms, it is important to address the concept of species in bacteriology, as researchers have yet to agree on its meaning.

First of all, the term microorganism is generally used to name unicellular or microscopic living entities, which includes members of groups as diverse as fungi, algae, viruses or amoebae. To avoid confusion, unless otherwise stated, the term microorganism and its synonyms are used in this text to refer only to the domain Bacteria. Interestingly, although members of this domain carry out unique and fundamental transformations in the biogeochemical cycles (Falkowski*, et al.*, 2008), and their total biomass (mostly localized in the oceanic subsurface) is comparable to that of terrestrial plants (Whitman*, et al.*, 1998), microbiology is a comparatively young discipline. Well into the 20th century bacteria were still classified as fungi, and only from 1930 did they start to have their own taxonomy. For the next 50 years they were catalogued according to their morphology or their metabolic activity, reaching some 40,000 type cultures by late 1970 (just a fraction of the known species of insects). However, the microbial classification used at that time was challenged by the pioneering work of Woese and colleagues (Woese*, et al.*, 1975; Woese & Fox, 1977), who began to reveal the correlation between nucleic acid sequence similarity, particularly that of 16S ribosomal RNA, and evolutionary relationships, opening the way for the utilization of molecular methods for bacterial systematics (Weisburg*, et al.*, 1991). Nowadays, 16S rRNA sequencing is a common tool to assess the taxonomy of newly isolated microbes (Vandamme*, et al.*, 1996) and over the last thirty years this technique has helped to redraw the phylogenetic tree of the Domain Bacteria (Brenner*, et al.*, 2000; Gevers*, et al.*, 2006). Furthermore, environmental surveys of 16S rRNA sequences have shown that the number of microorganisms we are able to grow in the laboratory represents a tiny proportion of the microbial diversity existing in any given ecosystem (Giovannoni*, et al.*, 1990; Curtis & Sloan, 2005; Gans*, et al.*, 2005). This latest discovery has renewed the interest in the field, opening the door for the establishment of microbial ecology as a discipline in its own right. At the end of 2008, the Ribosomal Database Project contained 81,607 full length 16S sequences from isolated bacteria and another 193,923 from uncultured strains. Less than 2 years later, the deposited sequences reached a total of 110,910 and 535;635 respectively, clearly

showing the exponential increase of 16S rRNA sequencing. However, prokaryotic diversity could exceed a million taxa (Dykhuizen, 1998; Curtis*, et al.*, 2002) and its real extent may well remain an estimation for a long time (Schloss & Handelsman, 2004; Pedrós-Alió, 2006). Particularly, this is because a clear definition of 'species' has yet to be agreed (Cohan, 2002; Bohannon, 2008).

In fact, one of the main challenges for the development of a robust bacterial systematics is that shared physiological or morphological traits do not always correspond to evolutionary proximity as determined by molecular approaches (Rossello-Mora, 2005). The main contributing factor to this phenomenon is the high level of genetic recombination occurring among natural bacterial populations (Feil*, et al.*, 2001; Whitaker*, et al.*, 2005; Fraser*, et al.*, 2007), also referred to as lateral gene transfer. Practically, genes are often exchanged between different taxa, making it impossible to estimate the metabolic repertoire of an isolate from its 16S rRNA gene sequence alone (Hanage*, et al.*, 2005; Hanage*, et al.*, 2006). For this reason, the whole idea of the existence of a phylogenetic tree capable of describing evolutionary relationships in bacteria has been challenged (Doolittle & Bapteste, 2007; Bapteste*, et al.*, 2009), and the discipline currently swing from a DNA-centric vision of taxonomy to a trait-centric system of classification; two faces of the same coin.

A graphical representation of how different rules to classify organisms can lead to an extensive rearrangement of the taxonomy is provided in Figure 1.3. Again, nucleic acid sequences often appear to be a good indicator of evolutionary relationships between different organisms, whether based on whole genome (Stackebrandt*, et al.*, 2002; Konstantinidis & Tiedje, 2005) or on fewer genetic markers (Konstantinidis*, et al.*, 2006), but the grouping of bacteria into operational unit based on the presence of characterizing traits may better serve the purpose of understanding ecological dynamics (Goodfellow*, et al.*, 1997; Green*, et al.*, 2008). Given that both strategies can result in fruitful classifications, it is expected that an approach capable of integrating information derived from DNA analysis together with phenotypic characteristics will facilitate the identification of ecologically distinct populations (Cohan, 2006; Cohan & Perry, 2007; Koeppel*, et al.*, 2008), and this could help to reveal processes regulating the structure/function of bacterial communities (Prosser*, et al.*, 2007). In order to expand this reasoning, it is interesting to look at the DNA vs. traits debate from another perspective; what causes bacterial distribution in natural ecosystems?

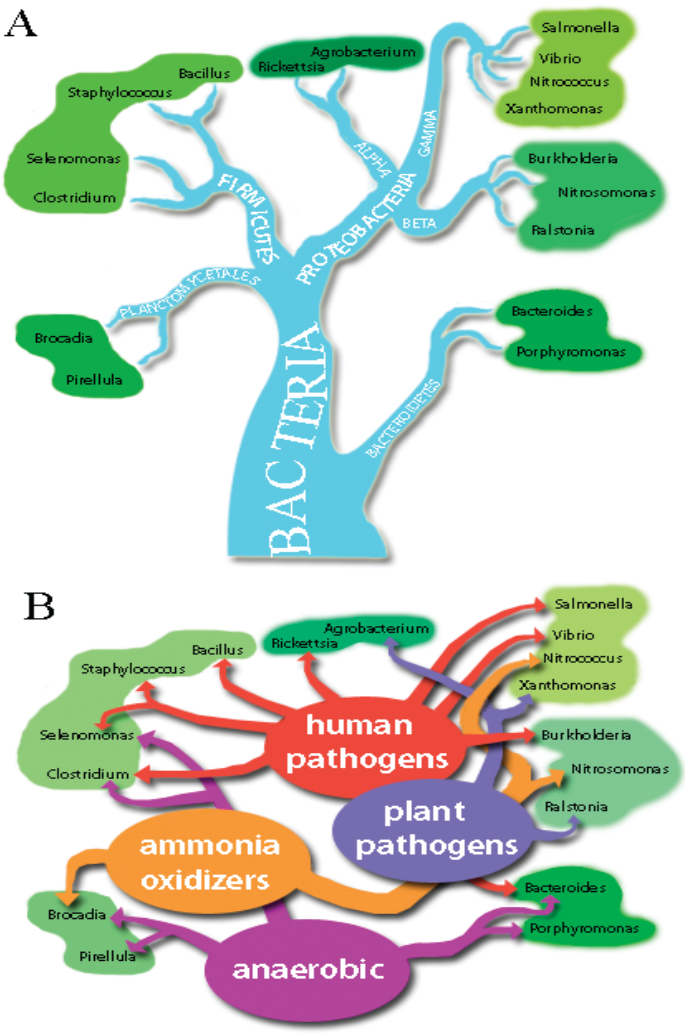


Figure 1.3 - **Which criteria for bacterial classification?**

The two pictures, A and B, show a simplified graphical representation of bacterial classification methods. Picture A highlights the evolutionary relationships between organisms as inferred from molecular data. Picture B shows a trait-based classification of the same bacteria "species" as in A.

The structure of bacterial communities appears to be shaped by a combination of niche and neutral processes (Hubbell, 2006). Although the capacity of niche and neutral models to predict community assembly is controversial (Harte, 2003; McGill, *et al.*, 2006), it seems that the latter is a better fit to ecosystems with limited immigration rate (Woodcock, *et al.*, 2007), while environmental variables play a major role in the distribution of bacteria when this is unconstrained. In fact, the composition of bacterial communities, or at least that of their dominant members, has been shown to respond to simple environmental parameters such as seasonal variation (Fuhrman, *et al.*, 2006), food source (Kent, *et al.*, 2006; Kent, *et al.*, 2007), pH (Fierer & Jackson, 2006; Palacios, *et al.*, 2008) and temperature (Lindstrom, *et al.*, 2005). Therefore, changes in

habitat conditions can determine clear patterns of bacterial distribution, but the taxonomic level at which these patterns emerge can vary. From an 'everything is everywhere, but the environment selects' point of view, it can be expected that the possession of a certain gene(s), rather than the belonging to a specific taxa, is responsible for the spatial distribution of bacteria. As a consequence, the fuzzy 'species' concept mentioned above would tend to blur the existence of a bacterial biogeography as determined by molecular tool. Nevertheless, habitat-dependant distributions have been reported from the sub-species level (Ward*, et al.*, 2006; Koeppel*, et al.*, 2008) to the phylum level (Ley*, et al.*, 2005; Philippot*, et al.*, 2009), clearly showing that taxa do possess a biological meaning, and attempting to draw a bacterial taxonomy is not a pointless exercise.

Concluding, despite the difficulties in linking a 16S rRNA gene sequence to the processes performed *in situ* by the bacterium owning such sequence, molecular surveys still play a fundamental part in our understanding of ecological dynamics. Furthermore, the determination of non-random spatial patterns of bacterial distribution is often considered the first step to revealing the role of certain taxa in any given ecosystem. For example, many specific animal-bacteria associations have been discovered through 16S rRNA sequencing. In the following paragraph, an overview of such associations is presented in order to suggest possible roles that bacteria may have in influencing *Balanus amphitrite*'s life cycle.

1.4 Microbe-invertebrate interactions

Most marine invertebrates possess a complex life cycle characterised by the occurrence of diverse larval stages before a pronounced rearrangement of the body tissues generates the juvenile morphology. The genome contains all the information required by an individual to develop functional organs and structures, but the understanding of how the actual timing and intensity of gene expression controls the developmental fate in natural environments is far from being achieved. This is because gene expression is influenced by a multitude of stimuli, of both internal and external origin, often with overlapping mechanisms of action or conflicting effects. Although it has been acknowledged for decades that the environment plays an important role in modulating such gene expression, the advent of molecular genetics led to a sort of deterministic view that unintentionally shadowed the importance of these environmental influences (Keller, 1995; Mead & Epel, 1995). The making of terms such as

"interactomics" (the study of molecular interactions in cells) clearly shows the tremendous impact that PCR-related technologies had on biology; but it also tells us how scientists are trying to embrace approaches capable of integrating information originating from different fields in order to understand biological phenomena. In this moving scenario, a multidisciplinary investigation is advocated to elucidate the events underlying invertebrate metamorphosis and settlement. In 2001; Gilbert argued that to understand how development takes place in the real world, the synergy of developmental biology and ecology could result in an extremely beneficial partnership (Gilbert, 2001). Environmental variables such as photoperiod, diet, temperature and the presence of promoters/inhibitors are known to influence the developmental programme at work in each individual. The aim of this section is to explore a subset of these factors that have so far received little attention; namely, the presence of bacteria cells, biofilm and secreted microbial compounds, and how they influence marine invertebrate development and settlement.

Invertebrate larvae can interact with microbes in three different ways:

1) by direct contact with bacterial cells, that can either attach on the external cuticle or colonize internal organs leading to symbiosis or pathogenesis;

2) by direct contact with biofilms, natural elements composed by microbial cells embedded in extracellular polymeric substances;

3) by chemical interaction, through the absorption/sensing of dissolved microbial compounds.

Considering the first set of possibilities, a typical example is the symbiosis between the squid *Euprymna scolopes* and the bacterium *Vibrio fischeri*, that has become a model system for the investigation of animal-bacteria interactions during development (McFall-Ngai, 1999). Hatching *E. scolopes* individuals are bacteria-free, but they are capable of acquiring *Vibrio* spp. from the environment within a few hours (McFall-Ngai & Ruby, 1998; Visick & McFall-Ngai, 2000). A highly regulated process allows the squid to recognize the specific symbionts and permit them to reach the light organs through pores present in the host's tissues (Nyholm*, et al.*, 2000; Nyholm*, et al.*, 2002). Once established, the bacteria induce a series of developmental changes in the squid that lead to the formation of a functional light organ, where the host provides a nutrient-rich and protected environment for *V. fischeri*, that in exchange exhibits the characteristic luminescence (Montgomery & McFall-Ngai, 1994; Kimbell & McFall-Ngai, 2003).

This overview of the squid-Vibrio interactions highlights the importance that bacteria play in the evolution and development of their eukaryotic counterpart. Particularly in the marine setting, where it is estimated that metazoan life began over 1 billion years ago, it is expected that multicellular organisms have evolved a vast array of strategies to take advantage of the microbial world surrounding them. New technological advances have contributed to shed light on the mechanisms that support such interactions, but researchers have mainly focused on the two more evident bacteria-animal partnerships: symbiosis and pathogenesis (Cirillo, 1999; Steinert*, et al.*, 2000; Goebel & Gross, 2001). However, recent understanding on how the vertebrate gut act together with the microbiota it hosts have suggested that bacteria influence animal development more strongly than previously expected (Nicholson*, et al.*, 2005; McFall-Ngai, 2006). It became apparent that individuals of a certain species are not alone, being instead composed of a complex network of eukaryotic and prokaryotic cells, and it has been suggested that community ecology might provide the correct context to understand the forces that shape animal evolution (Cirillo, 1999; Ruby*, et al.*, 2004). It is reasonable to hypothesize that tighter molecular interactions are established between bacteria and their eukaryotic counterpart when the bacteria reside within the host's body (e.g. endosymbionts, gut-associated microbes). However, 'external' contacts are also possible, and the following analysis refers to how microbial biofilms can influence invertebrate life cycle.

The sea bottom is one of the most microbe-rich environments known. A huge diversity of bacteria, microalgae and fungi contribute to shape the micro landscape of submerged surfaces (Battin*, et al.*, 2007), and it is expected that benthic organisms have evolved mechanisms to deal with such species richness. The recently published genome of the purple sea urchin *Strongylocentrotus purpuratus* (Sodergren*, et al.*, 2006) has shown that this organism possesses more than 600 genes involved in bacterial cell recognition (Hibino*, et al.*, 2006; Rast*, et al.*, 2006), nearly twice as many as humans do. The authors of this discovery suggested that these genes may engage in the defence against pathogenic bacteria and/or in the identification of the beneficial species that colonize the digestive tract augmenting the spectrum of catabolic activities performed by the host (Rast*, et al.*, 2006). However, more entire genome sequences are required to determine whether the incredibly high number of genes involved in non-self recognition found in *S. purpuratus* is an exception or a common feature in benthic invertebrates. In any case, all marine organisms that possess a larval stage which then metamorphoses to a sessile state encounter surfaces

which are covered by biofilm and it is generally agreed that this microbial layer provides precious information to the settling larvae (Bonar, *et al.*, 1986; Qian, *et al.*, 2007). As a model, a schematic representation of the connections between the biofilm and an exploring cyprid is provided in Figure 1.4.

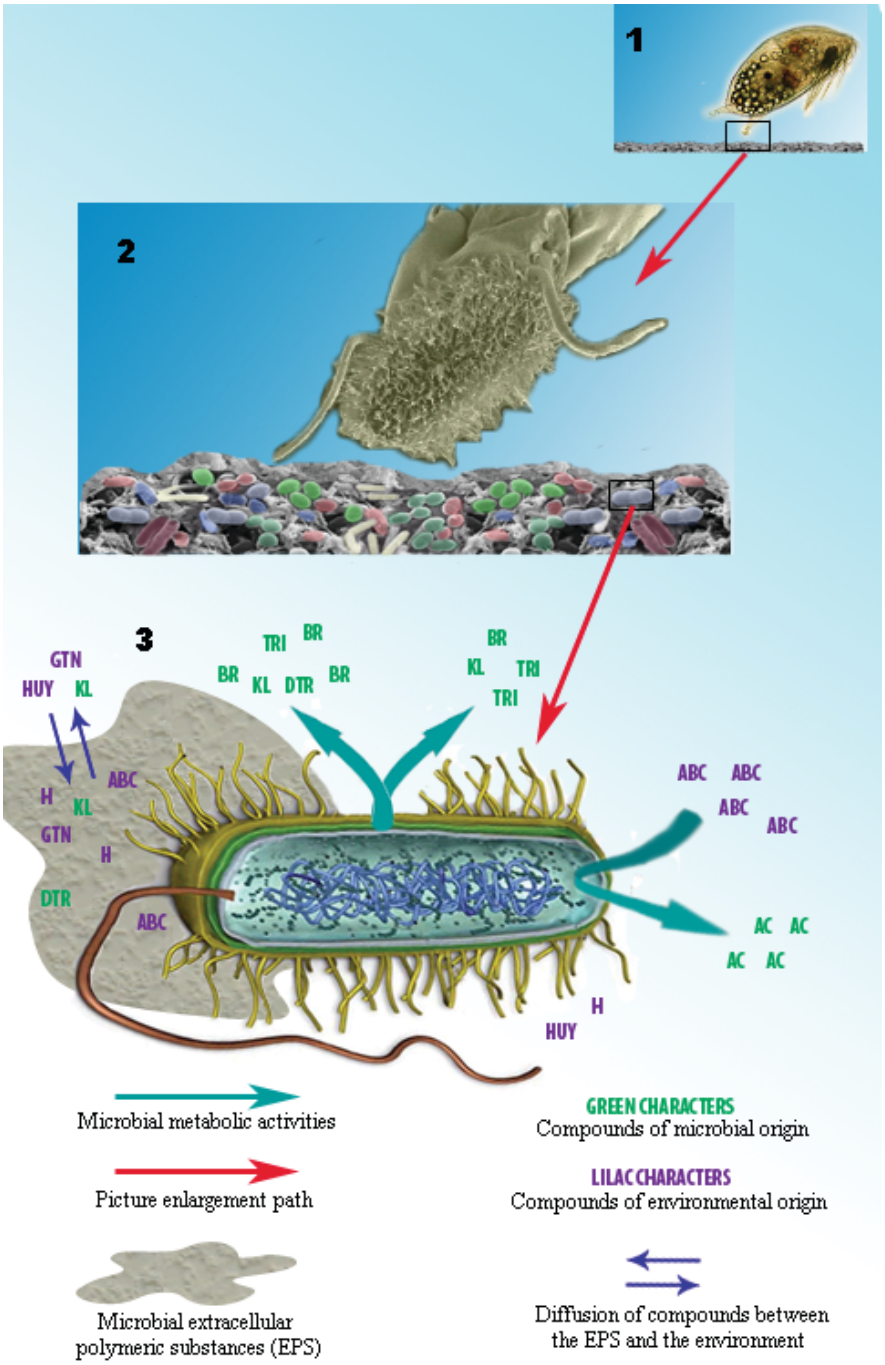


Figure 1.4 - **Interactions between bacteria and exploring barnacle larvae**

1 - Cypris larva; 2 - cyprid antennulae tip and the biofilm (approximately to scale); 3 - bacterial cell performing metabolic activities. The green/lilac characters shown do not refer to specific compounds but are just used to try to represent the molecular interactions between the cell and the environment.

The study of how surface-attached microbes influence invertebrate actions is a difficult task. This is mainly due to the fact that biofilms are dynamic environments which are continuously reshaped under the pressure of the physicochemical characteristics of the surroundings together with the biological activities triggered by commensal, antagonist and predatory behaviours (Watnick & Kolter, 2000; Parsek & Fuqua, 2004; Battin*, et al.*, 2007). Variables such as microbial species richness, characteristics of the extracellular polymeric substances (EPS), which constitute the biofilm matrix, and the presence of protozoa and other organisms which may disrupt the community architecture, all contribute to determine the properties of biofilms, making difficult to develop experimental procedure capable of assessing the relative importance of a single one of these variables on the output analysed. Nevertheless, ZoBell first suggested the importance of bacteria in influencing invertebrate settlement and metamorphosis more than 70 years ago (Zobell & Allen, 1935; ZoBell, 1938), and a vast amount of data has been produced on the topic ever since.

The importance of the biofilm characteristics such as the chemical composition, thickness and age in modulating larval settlement have been described, among others, for: the bivalve *Mytilus galloprovincialis* (Bao*, et al.*, 2007), the bryozoan *Bugula neritina* (Dahms*, et al.*, 2004), the polychaete *Janua brasiliensis* (Kirchman*, et al.*, 1981) and the sea urchins *Pseudocentrotus depressus* and *Anthocidaris crassispina* (Rahim*, et al.*, 2004). One of the best characterised effects shown by microbes is that of facilitating settlement in the polychaete worm *Hydroides elegans* (Hadfield*, et al.*, 1994; Harder*, et al.*, 2002; Lau*, et al.*, 2003). Qian and colleagues recently used the marine bacterium *Pseudoalteromonas spongiae* to present evidence in favour of the central role that biofilm formation plays in promoting *H. elegans* settlement (Huang*, et al.*, 2007). They showed that the bacterium was capable of inducing settlement when grown in media containing both peptone and yeast extract but not when one of the two ingredients was missing; introducing the idea that nutrient availability is a key factor influencing *P. spongiae*'s ability to produce a well established biofilm. They also demonstrated that blocking protein synthesis at the onset of biofilm formation significantly reduced the level of induction. Although their results suggest that a compound(s) involved in biofilm formation is responsible for *H. elegans* settlement stimulation, the procedure used did not allow the nature of the compound(s) to be clarified. A possible source of error in their procedure is that they started the experiment on protein synthesis inhibition with cells that had been washed and resuspended in sterile sea water, and did not consider that these steps may have strongly influenced

bacterial cell surface characteristics, hindering the possibility that settlement stimulation may depend on molecules attached to the cell wall rather than present in the EPS. Nevertheless, the results obtained by Huang *et al.* (2007) expand previous observations that *H. elegans* larvae were induced to settle by bacterial waterborne metabolites only when these compounds were bound to a surface (Harder*, et al.*, 2002), highlighting the importance that chemical signal-substratum association may have in determining the biological activity.

Biofilm architecture and EPS composition may be responsible for the effect that microbial films have on *Balanus amphitrite*'s behaviour (a review of the published work on *B. amphitrite*-microbes interactions is given in Table 1). For this organism, the literature presents a confusing picture, but it seems that biofilms, particularly old and thick ones, tend to inhibit settlement (Olivier*, et al.*, 2000; Faimali*, et al.*, 2004). The reason for this could be that EPS interferes with adhesive deposition and consequently the ability of the cypris larva to permanently attach to the surface. Although settling barnacles very rarely encounter microbe-free substrates in natural environments, they may have evolved mechanisms to sense the micro-patchiness of biofilms, favouring those areas where a direct contact between the surface and the attaching organ has became possible thanks to gaps in the EPS generated by the passage of a predator or by alternate dry-wet cycles. However, Hadfield and co-workers have recently provided evidence that advocates against the hypothesis that microbial films restrain barnacle attachment by simply getting in between larvae and the substratum (Zardus*, et al.*, 2008). On the contrary, they showed that the presence of biofilm has the ability to strengthen the tenacity of adhesion in larvae of *B. amphitrite* and other invertebrates, opening a new area of research that will focus on the physical architecture of the biofilm-invertebrate connection.

Related to the effect that EPS composition may have on settling larvae of *B. amphitrite*, some investigations seem to suggest that carbohydrates produced by bacteria can have promotory effects (Khandeparker*, et al.*, 2003), as it was shown that lectins were capable of turning the effect of *Pseudomonas aeruginosa* biofilm from promotion to inhibition. Lectins are a diverse group of proteins with carbohydrate-binding activity that is often involved in non-self recognition and innate immunity (Kilpatrick, 2002; Cambi & Figdor, 2003). Therefore, by adding lectins to a solution, it is possible to mask the presence of specific sugars in order to test their involvement in a biological process. Interestingly, the lectins used by Khandeparker *et al*. (2003) are the same ones capable of reducing the settlement-inducing activity of *B. amphitrite* adult extract (Matsumura*,*

et al., 1998b). As many sources of evidence have pointed out that the molecular signal that makes the settlement-inducing protein complex attractive for cyprids is resident in its sugar part (Matsumura*, et al.*, 1998a; 1998b; 1998c), the effect of lectins on biofilms allows a simple hypothesis to be advanced: carbohydrates, probably mannose and its polymers, forming the biofilm EPS are recognised by cypris larvae with the same chemical mechanism by which they recognise the SIPC. If this is the case, the settlement-inducing effect that certain bacteria have on *B. amphitrite* settlement could simply occur by chance, as they trigger the same signalling cascade activated by the SIPC. It would be of interest to test whether bacteria associated with adult *B. amphitrite* possess an EPS composed of proportionally more mannose-like residues in comparison to that of other marine ecotypes.

Finally, a last biofilm-specific mechanism that promotes larval settlement may exist: that the EPS acts as an adsorptive matrix for bioactive molecules. In fact, it was suggested that "exopolymers may retain bacterial metabolites on the cell surface and, thus, render the induction of larval settlement" (Harder*, et al.*, 2002). However, this retention mechanism may also involve waterborne cues of different origin that express their biological activity once absorbed in the biofilm. In fact, it is well known that the biofilm matrix is capable of binding and concentrating inorganic and organic matter dissolved in the sea (Wolfaardt*, et al.*, 1994; Buffle*, et al.*, 1998; Bhaskar & Bhosle, 2006). Although, much of the work on this topic has been done to elucidate the role of microorganisms in ecosystem dynamics and in the carbon trade-off in aquatic food-webs (Battin & Sengschmitt, 1999; Battin*, et al.*, 2003; Azam & Malfatti, 2007), it has been demonstrated that biofilm can accumulate antibiotics and other toxins (Wolfaardt*, et al.*, 1994; Labare*, et al.*, 1997; Dynes*, et al.*, 2006). Furthermore, evidence that biofilms are able to retain cues that induce settlement in larvae of the Serpulidae *Pomatoceros lamarkii* exists (Chan & Walker, 1998) and it will be interesting to test whether this is true also for molecules that can influence settlement in *B. amphitrite*. This hypothesis may also explain why natural biofilms show different effects on barnacle settlement if compared to laboratory-raised biofilm (O'Connor & Richardson, 1996; Hung*, et al.*, 2007). Although compound-specific dyes could be developed to test the adsorption capacity of the EPS *in situ*, a model system for multispecies biofilm development does not exist, making it difficult to test whether any effect found on larval settlement is due to the presence of absorbed molecular cues on the EPS or to the presence/proportion/activity of some bacterial ecotype. For these reasons, the study of how the biofilm, considered as a unique ecosystem, influences larval behaviour

represents a difficult task. Consequently, the best characterised of all possible bacteria-invertebrate interactions are those based on the biological activity of specific bacterial metabolites.

If a bioactive compound can be isolated, it becomes relatively easy to study its effect on larval behaviour and to reveal the biochemical events upon which the invertebrate responses are based. A typical case study is the effect that microbial biofilms have on settlement of zoospores of the marine green alga *Ulva* (formerly *Enteromorpha*) *intestinalis*. It has been demonstrated that natural biofilm is capable of attracting swimming zoospores and promoting settlement of *U. intestinalis* (Joint*, et al.*, 2000). Phylogenetic analysis of the biofilm community revealed that the zoospores responded to bacteria in a strain-specific manner (Patel*, et al.*, 2003). The compounds eliciting the algal behaviour were identified as *N*-acylhomoserine lactones (AHLs) (Joint*, et al.*, 2002; Tait*, et al.*, 2005), a class of molecules commonly employed in bacterial cell-to-cell communication within biofilms (Fuqua & Greenberg, 1998; Swift*, et al.*, 2001). Wheeler and colleagues showed that AHLs induce settlement by reducing the swimming speed of zoospores and that the rate of the chemoresponse was dependant on the AHL's acyl side chain (Wheeler*, et al.*, 2006). In less than ten years, the understanding of the mechanism by which bacteria attract settling *U. intestinalis* is well on its way and researchers are now ready to broaden their analysis to invertebrate species.

Few other examples of bacterial metabolites that influence invertebrate behaviour can be found in the literature (Kon-ya*, et al.*, 1995; Olguin-Uribe*, et al.*, 1997; Holmström & Kjelleberg, 1999; Dobretsov & Qian, 2004) and, due to the importance of chemical signalling in cross-kingdom interactions, it is expected that many other cases will be found. Although priorities have been given to the discovery of bioactive compounds that can be employed in antifouling technologies (Dobretsov*, et al.*, 2006; Fusetani & Clare, 2006), studies in the field of marine chemical ecology can also contribute to revealing processes by which bacteria influence eukaryotic evolution and development. This task requires the application of multiple disciplines such as genetics, developmental biology, biochemistry and ecology; an approach adopted here.

1.5 Aims of this project

In light of the economic and the environmental benefits that can result from an understanding of biological processes that lead to fouling establishment, this research

has focused on the interactions between bacteria and the barnacle *Balanus amphitrite*. For the reason that specific interactions were not known, this study was designed as an interdisciplinary investigation on two main themes. On the one hand, we intended to improve our knowledge of *B. amphitrite* genetics, to allow the exploration of settlement and metamorphosis at a molecular level. The use of ultra-high throughput sequencing techniques and the development of a reliable quantitative PCR assay were seen as key to this purpose. On the other hand, the existence of stable associations between the barnacle and bacteria was investigated through a biogeographical comparison of the microbial communities living within and around *B. amphitrite*. In theory, the combination of these two themes would have allowed the study of the way that bacteria found in association with *B. amphitrite* influence its gene expression.

Table 1.1

Effect of bacteria on settlement of barnacles (*Balanus amphitrite*)

INVERTEBRATE	MICROBES	ASSAYED ON	EFFECT	EXPLANATION	REFERENCE
B. amphitrite	*Pseudomonas aeruginosa*	Polystyrene dishes; laboratory	Settlement stimulation	Some not-identified sugar in the EPS promoted settlement	(Khandeparker, *et al.*, 2003)
B. amphitrite	*Citrobacter freundii, Bacillus pumilis*	Polystyrene dishes; laboratory	Settlement inhibition	Presence of ribose in the EPS	(Khandeparker, *et al.*, 2003)
B. amphitrite	Natural biofilms	Marble, quartz, glass, and cembonit	Young biofilm promoted while old inhibited settlement, with difference on the material tested	Not given	(Faimali, *et al.*, 2004)
B. amphitrite	*Deleya marina*	Polystyrene dishes and glass; laboratory	*D. marina* supernatant inhibits settlement on polystyrene, but results were not reproducible	Not given	(O'Connor & Richardson, 1998)
B. amphitrite	*Pseudomonas aeruginosa*	Polystyrene dishes; laboratory	Settlement stimulation, mainly when tested at 35‰ salinity and 26ºC temperature	Bacteria produce settlement promoters but dependent on the growth conditions	(Khandeparker, *et al.*, 2002)
B. amphitrite	*Bacillus pumilis, Citrobacter freundii, Pseudomonas aeruginosa*	Polystyrene dishes; laboratory	Different effects according to the experimental setting	Nutrient availability, salinity, interactions with barnacle extracts, fraction of the microbial metabolites tested (supernatant, leachates, cell extracts) all influenced the results	(Khandeparker, *et al.*, 2006)
B. amphitrite	Mixed biofilms developed in the field	PVC plates, assays in laboratory and in field	Biofilm from some environment recruited cyprids more than other	Particular bacterial species or group are responsible for the effects noted.	(Hung, *et al.*, 2007)

continues

Table 1.1 (continued)

INVERTEBRATE	MICROBES	ASSAYED ON	EFFECT	EXPLANATION	REFERENCE
B. amphitrite	Mixed natural biofilm	Polystyrene dishes; laboratory	Young biofilm (2-4d) inhibits, old biofilm (12-18d) promotes settlement	Suggested that the noted effect relates to the cyprid's need to evaluate tidal characteristics.	(Wieczorek, *et al.*, 1995)
B. amphitrite	Single species isolated from barnacles	Polystyrene dishes; laboratory	*Vibrio* spp. and other species inhibited settlement	Not given	(Avelin Mary, *et al.*, 1993)
B. amphitrite	Natural mixed population	Glass slides; laboratory	The presence of biofilm strengthens the adhesive tenacity	Film may stimulate/modulate production of adhesives; film may reduce shear stress on the surface	(Zardus, *et al.*, 2008)
B. amphitrite, Ciona intestinalis	A Gram negative anaerobic bacterium	Glass slides; laboratory	The bacterium inhibits larval settlement	The bacterium produced a compound toxic to larvae	(Holmström, *et al.*, 1992)
B. amphitrite	*Alteromonas* sp.	Glass slides; laboratory	The bacterium inhibits larval settlement	The bacterium produced ubiquinone, which was toxic	(Kon-ya, *et al.*, 1995)
B. amphitrite	Various single species biofilms	Polystyrene dishes; laboratory	Some species inhibited settlement, others had no effect	EPS might be involved in settlement inhibition	Maki et al 1998
B. amphitrite	*Deleya marina* sp.	Polystyrene dishes; laboratory	Some films inhibited settlement other had no effect	Bacteria did not interfere with the wettability of surfaces and may produce toxic compounds	(Maki, *et al.*, 1992)

Continues

Table 1.1 (continued)

INVERTEBRATE	MICROBES	ASSAYED ON	EFFECT	EXPLANATION	REFERENCE
B. amphitrite	Single diatoms species	Polystyrene dishes; laboratory	Diatoms films tend to promote settlement.	The characteristics of the EPS play a pivotal role in inducing larval recruitment	(Patil & Anil, 2005)
B. amphitrite	*Halomonas marina, Pseudomonas marina, Deleya marina.*	Polystyrene, glass; laboratory	Prevalent inhibition of settlement	Interactions between the bacterium and the surface influence the biofilm EPS that in turns influence the settlement behaviour.	(Maki*, et al.*, 2000)
B. amphitrite	Natural mixed species	Polystyrene; laboratory	The oldest the biofilm the less the larvae settled	Not given	(Olivier*, et al.*, 2000)
B. improvisus	*Deleya marina*	Glass and polystyrene in laboratory. Polystyrene in the field	Settlement inhibition on polystyrene and promotion on glass at certain salinity levels in the lab. Small promotory effect in the field.	Biofilm may act as a barrier for the detection of cues from the substratum or can take up cues from the medium. Different environment can affect the microbial metabolic activity.	(O'Connor & Richardson, 1996)
B. improvisus	*Alteromonas macleodii, Pseudomonas fluorescens*	Glass and Polystyrene; laboratory	Different results (salinity dependant) on polystyrene and small promotion on glass (*A. macleodii*)	Biofilm may act as a barrier for the detection of cues from the substratum.	(O'Connor & Richardson, 1996)
B. trigonus	Mixed natural population	Polystyrene, in the field	Biofilm form subtidal habitats favoured cyprids attachment.	Microbes provide cue to help in deciding vertical distribution. The type of cue has not been suggested	(Thiyagarajan*, et al.*, 2006)

Chapter 2. Developmental stage-specific transcriptome of *Balanus amphitrite*

2.1 Introduction

The barnacle *Balanus amphitrite* (syn. *Amphibalanus amphitrite*, Clare & Høeg, 2008) is one of the most widely used organisms to investigate settlement in marine invertebrates. It forms large colonies in temperate to tropical waters across the globe and it is considered a cosmopolitan hard fouler (Townsin, 2003). The life cycle of *B. amphitrite* is characterised by six successive larval stages, the naupliar instars I-VI, which are followed by the cyprid, a non-feeding planktonic stage with a modified morphology that is specialised for locating a suitable surface for permanent attachment (Glenner, 1999). Once settled, the cyprid undergoes an intense morphological rearrangement that leads to the juvenile stage, which acts as a suspension feeder using its cirri to capture plankton. *B. amphitrite* is a cross-fertilising hermaphrodite enclosed in six hard calcareous plates deposited to protect its body. Typically, recently moulted individuals are receptive as females and, when fertilised, deposit a pair of ovisacs that bear hundreds of eggs (Anderson, 1994).

Understanding the mechanism by which cypris larvae choose the settlement location is of economic importance. So far, however, studies in this field have been limited to the behavioural response of larvae to settlement cues. This has created a wealth of information on the environmental stimuli that *B. amphitrite* respond to during surface exploration (Rittschof*, et al.*, 1984; Holmström*, et al.*, 1992; O'Connor & Richardson, 1998; Olivier*, et al.*, 2000; Khandeparker*, et al.*, 2006), but the molecular mechanisms underlying these responses remain unclear. In recent years, Thiyagarajan and Qian (2008) have used a proteomics approach to investigate *B. amphitrite* development. Briefly, they have compared the phosphoprotein and total protein profiles of cyprids that were allowed to settle and metamorphose for 24h to that of cyprids in which metamorphosis was blocked at a molecular level by the addition of the compound genistein. The comparison of these profiles was performed by visualisation of 2-dimensional gels and differently expressed/phosphorylated proteins were excised from the gels and the aminoacid sequences determined by mass spectrometry (MS). This has led to the identification of a set of proteins possibly involved in larval attachment and metamorphosis. However, the authors claimed that the lack of genomic information for

this organism posed a limit to their MS-based protein identification approach (Thiyagarajan & Qian, 2008). The aim of this study was to provide the scientific community with an extensive dataset of gene sequences that can support studies on the molecular mechanisms regulating pelago-benthic transition.

Traditionally, genetic information derives from either direct sequencing of DNA fragments or from RNA transcripts that have been reverse-transcribed into cDNA. The first approach is fundamental to understanding the genetic potential of an organism and to identifying the regulatory elements that influence gene expression. On the other hand, cDNA or Expressed Sequence Tags (ESTs) sequencing surveys are fundamental for discovering new genes (Adams*, et al.*, 1991) and they often represent the initial step for the molecular characterization of the species of interest. In addition, EST-derived information supports genomic sequence annotation by suggesting intron/exon boundaries and the existence of previously undescribed transcription units; consequently, mRNA sequences are invaluable in comparative genomics (Marra*, et al.*, 1998). Interestingly, despite being a paradigm for the study of evolution (Darwin, 1851), both genomic and transcriptomic data for barnacles are scarce. For *B. amphitrite*, some *ad hoc* sequencing has been performed (Okazaki & Shizuri, 2000; Okazaki & Shizuri, 2001; Dreanno*, et al.*, 2006b) and a catalogue containing 609 unique ESTs exists (Bacchetti De Gregoris*, et al.*, 2009), which altogether may represent as little as 2% of its protein-coding sequences (Hou & Lin, 2009). Furthermore, in order to obtain biological information from cDNA sequences, a great use of comparative analysis is made with genetic data from related species. The current situation for barnacles is far from being optimal.

Phylogenetically, barnacles belong to the phylum Arthropoda, the largest animal lineage known, which includes insects, spiders, centipedes and other crustaceans. The phylogeny of this phylum has long been debated but a recent tree obtained with more than 40kb of concatenated gene sequences seems to suggest a reconciliatory view (Regier*, et al.*, 2010), which we will follow here on. For the whole subclass Thecostraca (barnacles), mainly mitochondrial genomes are available, with that of *Tetraclita japonica*, *Megabalanus volcano*, *Pollicipes polymerus* and *P. mitella* being fully sequenced (with accession numbers AB126701, AB167539, AY45618 and AY514042 respectively). To find the closest relative to *B. amphitrite* with an existing large collection of EST data we need to go back to the salmon louse *Lepeophtheirus salmonis*, which is classified in the same order with barnacles. Although modern sequencing techniques hold the promise to quickly resolve the issue, this clearly

highlights the gaps in genomic information that are currently available for many organisms that are important for both basic and applied research. Furthermore, few studies have investigated the regulation of the pelagobenthic life-cycle at a molecular level (Degnan & Morse, 1995; Woods*, et al.*, 2004), despite its broad distribution in marine invertebrates (Hadfield*, et al.*, 2001). The use of barnacles as model system for this purpose could answer fundamental biological questions.

In order to characterise the genetic repertoire of *B. amphitrite*, two alternative molecular approaches were undertaken. In the first, adult-extracted RNA was reverse transcribed into cDNA, which was then used to generate a clone library to facilitate sequencing. For clarity, the term 'EST library' will be used to refer to this library. Regarding the second molecular method implemented, RNA from 3 developmental stages was used as starting material. The 3 cDNA pools were then subtracted to reduce the preponderance of highly expressed transcripts and to facilitate gene discovery. The 3 'subtracted libraries' were then sequenced with 454 titanium technologies. The assemblies of 454 reads were aligned and annotated using SAMS 2.0 (Bekel*, et al.*, 2009), and a database was created to facilitate genetic analysis. Guidelines to access the database and a discussion of the biological significance of our findings are reported here.

2.2 Material and methods

2.2.1. Balanus amphitrite culture

Wild *B. amphitrite* adults were collected from Beaufort, North Carolina, USA (courtesy of Prof. D. Rittschof). Brood stocks were maintained in semi-static culture in UV-irradiated, 10 μm filtered natural seawater. The adults were fed on newly-hatched *Artemia* sp. nauplii (Artemia International LCC, U.S.A.). To obtain barnacle nauplii, the adults were placed in a tank of fresh seawater and released larvae were attracted to a point light source and collected by pipette over a 2 h interval. Nauplii were cultured at the density of ~1 larva ml^{-1} in an incubator at 28º C on a 12:12 light:dark cycle. The larvae were fed each day with 1 l of a *Skeletonema costatum* culture (~ 2 x 10^5 cells ml^{-1}) until they reached the cyprid stage (approx. 4-5 days). Cyprids were collected by filtering through a tier of filters (pore sizes of 350 and 250 μm) in order to discard undeveloped cyprids and microalgae, and stored at 6º C until use.

2.2.2. RNA extraction

In order to retrieve the largest possible collection of genes, we focused on the RNA transcripts expressed during three different developmental stages:

1) naupliar stage (a mix of instar I and II);

2) cyprid aged at 6^{o}C for 3 d (which is the standard age used in settlement assays);

3) adult.

Hatched nauplii were attracted to a light source, collected by pipette and checked under a microscope for the presence of contaminant organisms. As the first stage of naupliar development is of short duration, and the developmental stage was not assessed, the sample may have comprised a mix of stage 1 and 2 nauplii. Larvae were filtered though a 70 µm nylon membrane, rinsed with sterile sea water and resuspended in a 1.5-ml tube. The tube was briefly kept on ice to cause sinking of the nauplii and to allow removal of excess water. Finally, larvae were resuspended in 1 ml of TRIzol® reagent (Invitrogen) and sonicated mildly for 5 seconds twice. The cyprid sample comprised 100 larvae treated in the same way as nauplii. For the adult stage, 5 individuals were removed from their shells, snap-frozen in liquid nitrogen and ground with a mortar and pestle. The biological material in the mortar was resuspended in 5 ml of TRIzol® and collected in sterile 1- ml tubes. The particulates were left to settle to the bottom of the tube for 5 minutes on ice and 1-ml aliquots of the supernatant were loaded in 1.5-ml tubes.

For all samples, the RNA was isolated from TRIzol® following standard procedure. DNase treatment was performed with the RNA bound to RNeasy® spin columns (Qiagen) following the instruction provided with the kit. The RNA was then eluted using 80 µl of milliQ water, the purity evaluated using NanoDrop ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies) and the quality confirmed by running 2 µ l of each sample in a 1% agarose gel containing RiboRuler (Fermentas) as a reference RNA. Aliquots containing 2.5 µg of RNA were precipitated in 1/10 volume of 3M sodium acetate and 2.5 volumes of ice-cold 100% ethanol and stored at -20 ºC (or in dry ice) until used.

2.2.3. ESTs library creation and sequencing

In order to generate the first adult ESTs library, cDNAs were prepared by reverse transcription using the first strand synthesis primer 5'-GAGAGAGAGAGAGAGAGAGAACTAGT**CTCGAG**T^{17}-3' (complementary to the

poly-A mRNA tail), which contains an *XhoI* restriction site (in bold) to facilitate directional cloning of the 3' end of the ds-cDNA insert into the vector. The first strand synthesis used me5-dCTP rather than ordinary dCTP. Non-methylated dCTP was then used in the second strand reaction to make the complementary cDNA strand. This method prevents internal cleavage of the cDNA when the linker is digested subsequently with *XhoI*. Prior to cloning, a double stranded linker containing a 5-*EcoRI* overhang (5'-OH-**AATT**CGGCACGAGG-3', overhang given in bold) was blunt-end ligated onto the ds-cDNA. The lack of phosphate on the 5' overhang for the *EcoRI* linker prevented concatemerization during linker ligation (this was phosphorylated in a subsequent step). The linker-ligated cDNA was then digested with *XhoI* and cloned directionally into the multiple cloning site of the plasmid vector pBluescript II SK+, previously linearised by digestion with the restriction enzymes Eco-RI and *XhoI*. A total of 905, out of 960 positive clones that were sequenced provided high quality sequence data.

2.2.4. Clustering, assembly and functional annotation of the EST library

The bioinformatics analysis of the adult cDNA library BA23840 was performed using the sequence analysis and management system SAMS-2.0 (Bekel*, et al.*, 2008). We first applied a clustering step based on pair-wise comparison on the DNA level using the TIGR default parameters (Quackenbush*, et al.*, 2000) to avoid redundancies in the dataset. Individual ESTs fall in the same cluster if they show a similarity of at least 95% over a region of not less than 40 bp in a pair-wise alignment and unmatched flanking regions must not exceed a length of 20 bp. Each cluster was then assembled using CAP3 (Huang & Madan, 1999), to produce 79 tentative contiguous (TC) sequences and 530 singlets that were nearly free of redundancies and allowed the following functional analysis to be constructed within SAMS. After applying a modified GenDB (Meyer*, et al.*, 2003) annotation pipeline consisting of a collection of standard bioinformatics tools including BLAST (Altschul*, et al.*, 1997), HMMer (Eddy, 1998) and InterPro (Apweiler*, et al.*, 2000) on each sequence, we applied Metanor (Goesmann*, et al.*, 2005), the GenDB automatic function prediction programme. By interpreting all the tool results we obtained, we created consistent functional annotations and assigned gene products, EC numbers, GO terms and KOG functional categories (Tatusov*, et al.*, 2003). Finally, TCs and singlets were manually checked and gene names were given whenever possible. High quality ESTs were deposited in the EMBL database, accession numbers from FM882258 to FM883162.

2.2.5. Subtracted libraries creation and 454 sequencing

800ng of total RNA from the three developmental stages investigated were used for cDNA synthesis using the SMART approach (Mint-Universal cDNA synthesis kit, Evrogen). An aliquot of 100ng of cDNA was subsequently normalized using duplex-specific nuclease and re-amplified afterwards according to the manufacturers' instructions (Trimmer kit, Evrogen, Russia). Libraries for 454 sequencing were prepared from cDNA with the "GS FLX Titanium General Library Preparation Kit". The 454 library so obtained were immobilized on beads and clonally amplified using the "GS FLX Titanium LV emPCR Kit (Lib-L)". The library was then sequenced using the "GS FLX Titanium Sequencing Kit XLR70" and "GS FLX Titanium PicoTiterPlate Kit". All kits used were purchased from Roche and used according to the manufacturer's protocol.

2.2.6. Clustering, assembly and functional annotation of 454 reads

The assembly was performed using the GS De Novo Assembler software (version 2.3) with 'cDNA' mode (Roche). Thus the assembly results were given as 'isogroups', representing the genes, with 'isotigs', representing the isoforms/splice variants of the gene. The results of the assembly were uploaded into the modified version of the sequence analysis and management system SAMS-2.0 for further bioinformatic analysis (Bekel*, et al.*, 2009). SAMS was modified to enable storing and visualizing the isogroup/isotig data. After applying the SAMS annotation pipeline consisting of a collection of standard bioinformatics tools including BLAST (Altschul*, et al.*, 1997) and InterPro (Apweiler*, et al.*, 2000) on each sequence, we applied Metanor (Goesmann*, et al.*, 2005), the GenDB/SAMS automatic function prediction program. By interpreting all the tool results we obtained consistent functional annotations and assigned gene products, EC numbers, GO terms and KOG functional categories to the isotig (Tatusov*, et al.*, 2003).

2.3 Results and discussions

2.3.1. Annotation of sequences from the EST library

Balanus amphitrite is one of the most extensively studied barnacles and has been suggested as a candidate genetic model for larval settlement and metamorphosis

(Hadfield, 1998). Recently, a proteomics approach was used to investigate settlement regulation in this organism (Thiyagarajan & Qian, 2008). It was argued that the lack of deposited gene sequences hindered a full appreciation of the results. The creation of an adult *B. amphitrite* cDNA library, the sequence of 960 clones (of which 55 were excluded from our analysis as they showed insert length shorter than 50 bp) and the detection of 530 singlets and 79 TC sequences (a summary of the EST survey is given in table 2.1) is thus an important first-step towards understanding the molecular ecology of this barnacle. Among the 609 different genes we report, 107 appeared to be of mitochondrial origin and 75 showed similarities with previously published ribosomal sequences. Gene ontology entries (Ashburner, *et al.*, 2000) revealed that the main categories of the genes found in our library appear to be involved in electron transport, protein biosynthesis, catalytic activity, metal ion binding, metabolism and the biogenesis of structural elements such as muscle and cuticle (Figure 2.1).

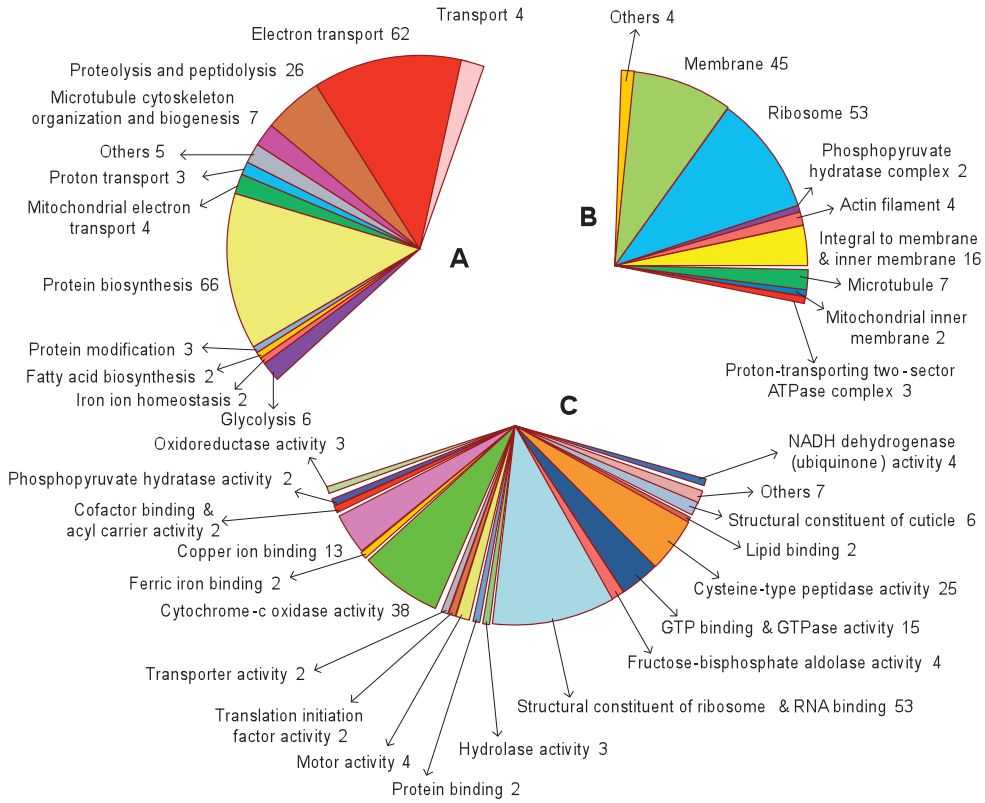


Figure 2.1 - **Gene Ontology mapping for *Balanus amphitrite*'s ESTs**

Relative distribution of ESTs within the three main subclasses existing in the GO classification: A) Biological process; B) cellular component; C) molecular function. Sequences obtained from the EST library were used to draw this figure.

Approximately 38% of the unique sequences we obtained have been functionally annotated and a corresponding gene name proposed, and informative annotation has

been given for an additional 18%. Of the remaining ESTs, 8% had a BLAST match with uncharacterized transcripts and 37% showed no appreciable similarity to previously published sequences. This distribution of ESTs among known/uncharacterized/unknown genes does not differ substantially from that found in recent EST surveys on other marine invertebrates (Venier*, et al.*, 2003; Tanguy*, et al.*, 2008). We also determined several transcripts that were highly similar to sequences derived from the deposited complete mitochondrial genomes of the two barnacles *Megabalanus volcano* and *Tetraclita japonica*.

Table 2.1 - **Classification of *Balanus amphitrite* ESTs**

Search method	**Putative source**	**E-value**	**Annotation**	**Nº of sequences**
Blast2n vs nt				457
InterPro				438
Blast2x vs SP				458
Blast2x vs KEGG				500
			Functionally annotated	449
			Unassigned ESTs	149
			Unknown ESTs	307
	Ribosomal RNA			123
	Mitochondrial DNA			202
	Genomic DNA			580
		< e-30		191
		> e-30		407

Vector clipped ESTs longer than 50 nucleotides and with an E-value > e-5 were functionally annotated whenever possible. Blast hits with E-values < e-5 were included in the unassigned proteins. Unknown ESTs in our dataset are those with no match in sequence databases. These ESTs were considered to be of genomic DNA origin as it is unlikely that the highly conserved mitochondrial gene would not result in a Blast hit.

Fragment assembly generated a total of 79 TCs comprising 375 ESTs. Sequences belonging to 8 different TCs were particularly frequent in our library, with the most common being an unassigned mitochondrial gene partly similar to 16S rRNA (with 52 entries), followed by cytochrome c oxidase subunit I (31), cysteine proteinase (17), cytochrome c oxidase subunit II (14), cytochrome b (12), a ribosomal RNA internal transcribed spacer (12), cytochrome c oxidase subunit III (11) and the elongation factor 1-α (11). The longest TC generated was 1703 nucleotides and

translated for the18S rRNA gene. Considering the 609 unique sequences we obtained, a total of 280 had a match in the NCBI nucleotide database. A taxonomic subdivision of the first hit produced by these 280 transcripts showed that 109 of them matched sequences from barnacles. The remaining sequences were represented among insects, vertebrates, arachnids, plants, fungi and various other groups (79, 53, 6, 8, 5 and 19 sequences, respectively). To annotate *B. amphitrite*'s genes, the proposed nomenclature for *Drosophila melanogaster* was used as a guide (http://flybase.org) and the corresponding gene symbol established in *D. melanogaster* was used when possible. However, in a slight departure, we decided to use the prefix mt- to identify mitochondrial genes.

2.3.2. Analysis of the subtracted libraries

The genome size of *Balanus amphitrite* is reportedly 1.4 pg (C-value), equivalent to 685 Mbp (Bachmann & Rheinsmith, 1973). By applying the formula presented by Hou and Lin (2009), the total number of protein-coding sequences in this organism is estimated to be around 28,500; nearly twice as many as are present in *Drosophila melanogaster* (fruit fly) and half the complement of *Oryza sativa* (rice). In order to reduce the preponderance of highly-expressed genes, a normalisation step, known to facilitate gene discovery (Shcheglov*, et al.*, 2007), was applied to the cDNA obtained from the three developmental stages investigated. This normalisation process is based on the denaturation and subsequent renaturation of double strand cDNA, followed by ds-cDNA digestion by duplex-specific nucleases, which allow to reduce the preponderance of those sequences that are highly expressed in the cDNA pool. The 3 cDNA pools obtained by this approach were then independently sequenced by 454 technologies and the total number of 185,564, 215,138 and 174,964 reads were generated for the nauplius, cyprid and adult libraries respectively. Altogether, ~191 Mb were sequenced, with an average read length of 331 bp.

Reads from single EST libraries were computed and annotated using SAMS (Bekel*, et al.*, 2009), which stores the sequences into three categories: 1) isotigs, which are sequences obtained with a classical alignment approach, basically all estimated ESTs (or part of); 2) contigs, which are the common sequences shared by two or more isotigs; and 3) isogroups, which are a list of isotigs that share at least one contig, in other words either isoforms of a gene or ESTs that share a conserved sequence. Therefore, the unique isotigs are defined by the assembly, and contigs and isogroups are

created in a second step using the isotig data. The results of the assembly are accessible from the following webpage: https://www.cebitec.uni-bielefeld.de/groups/brf/software/sams/cgi-bin/sams_login.cgi?project=SAMS_MGE_amphitrite_2010&chksum=FWJd9lIjRyV3MgXCUoHjiQ (Firefox supported). By accessing this link, a page containing the summary of the project is loaded on screen and by clicking on 'isotig', a webpage like that reproduced in Figure 2.2 appears.

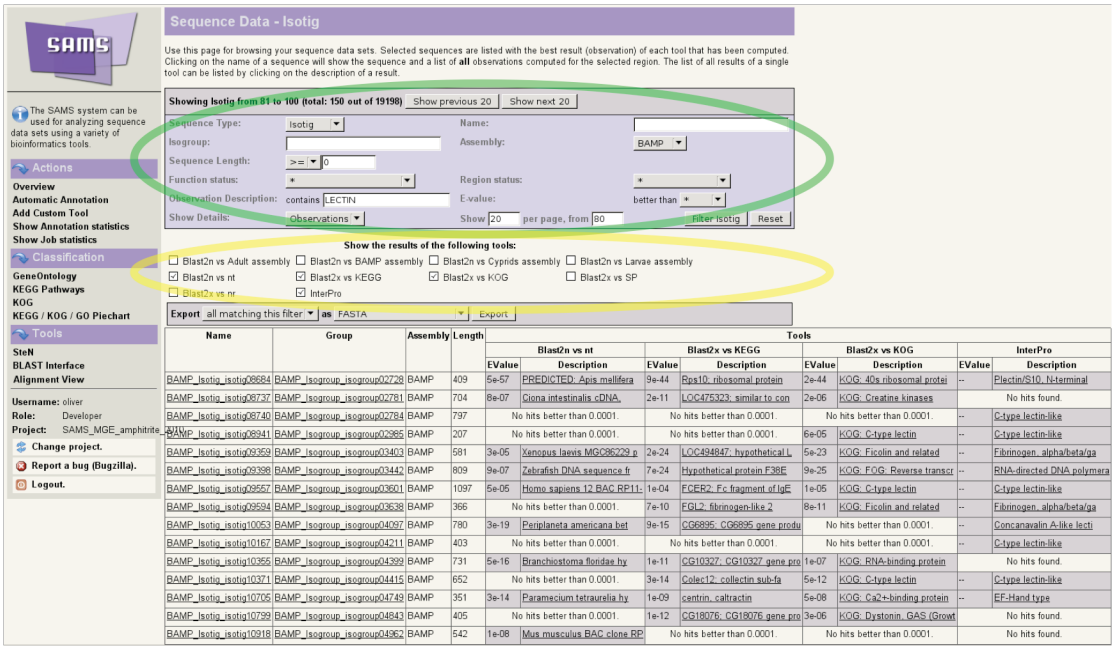


Figure 2.2 - **The web-interface of the SAMS database**

Guest users are allowed to search deposited *Balanus amphitrite*' sequences through the SAMS database. Isotigs can be browsed inserting a series of parameters in the green-circled area. It this example, BAMP sequences have been searched for isotigs that matched genes similar to 'lectin'. In the yellow-circled selectable area, it is possible to limit the results shown to specific databases (e.g. KOG, KEGG, Swiss Prot). On the left side of the SAMS interface, a series of features exist to help in analysing the sequences (e.g. specific KEGG pathways can be searched for matching isotigs).

Here the user is allowed to search the database by specifying a series of parameters and/or by searching for specific isotigs. In the SAMS database, the three assemblies generated were named 'Adult', 'Cyprids' and 'Larvae', with the latter created using reads from naupliar cDNA. By pooling the reads from all three libraries and re-running the assembly software, a fourth assembly (BAMP) was generated, to both increase the sequencing depth in respect to each assembly and to facilitate the analysis of genes that are uniquely expressed in one or more developmental stage. Therefore, BAMP represents an altogether assembly, in which, by using the SteN tool

(Fig. 2.2), it is possible to search for isotigs generated by the alignment of 454 reads derived from one 'or'/'and'/'and-not' another library.

The library-specific assembly of 454 reads identified 4,843 isotigs in the adult stage and 6,754 and 7,506 in the cyprid and naupliar stages respectively. Each isotig was BLAST searched against the other two assemblies in order to identify the developmental-specific genetic signature. Matching sequences with e value < $1e^{-50}$ and nucleotide identity > 95% were considered hits and the results were used to draw Figure 2.3 A and B with slightly different approaches. In part A, isotigs were first pooled into orthologous groups and then the developmental stage(s) in which each group was expressed was identified. In part B, isotigs from each assembly were BLASTed independently. As a consequence, for example, a gene that is not present in the naupliar assembly and that has 2 isoforms expressed in the cyprid and 5 in the adult stage, in Figure 2.3A is counted as 1 in the 'Adults + Cyprids' group, while in Figure 2.3B it is counted as 2 in the 'Adults + Cyprids' group of 'Cyprids Isotigs' and 5 in the 'Adults + Cyprids' group of 'Adults Isotigs'. From this analysis, it appeared that between 56 and 76% of isotigs in each assembly were uniquely found in that assembly (Fig 2.3B), suggesting that the genes expressed in each developmental stage are highly diverse. Furthermore, cyprid- and naupliar-specific isotigs were more similar to each other than to adults' genes, with only 4% that were retrieved from all 3 assemblies (Fig 2.3A). Altogether, 13,891 genes (19,198 isotigs) were identified, which may correspond to more than 2/3 of *B. amphitrite* protein-coding sequences. However, it has to be considered that the assembly of 454 data without a reference genome tends to overestimate the number of ESTs (Pop, 2009). Therefore, the full appreciation of the data presented here will be gained following the release of the complete genome sequence; work that is in progress.

Compared to the 4080 isotigs that this study found to be uniquely present in the cyprid stage, Okazaki *et al.* (2000) proposed the existence of six cypris larva-specific genes (*bcs*) in *B. amphitrite*, and their expression level during cyprid attachment and metamorphosis has been investigated recently (Li*, et al.*, 2010). It was shown that *bcs-1* to *bcs-5* were significantly down-regulated after attachment and toward the end of the metamorphic process, while *bcs-6* displayed the opposite trend.

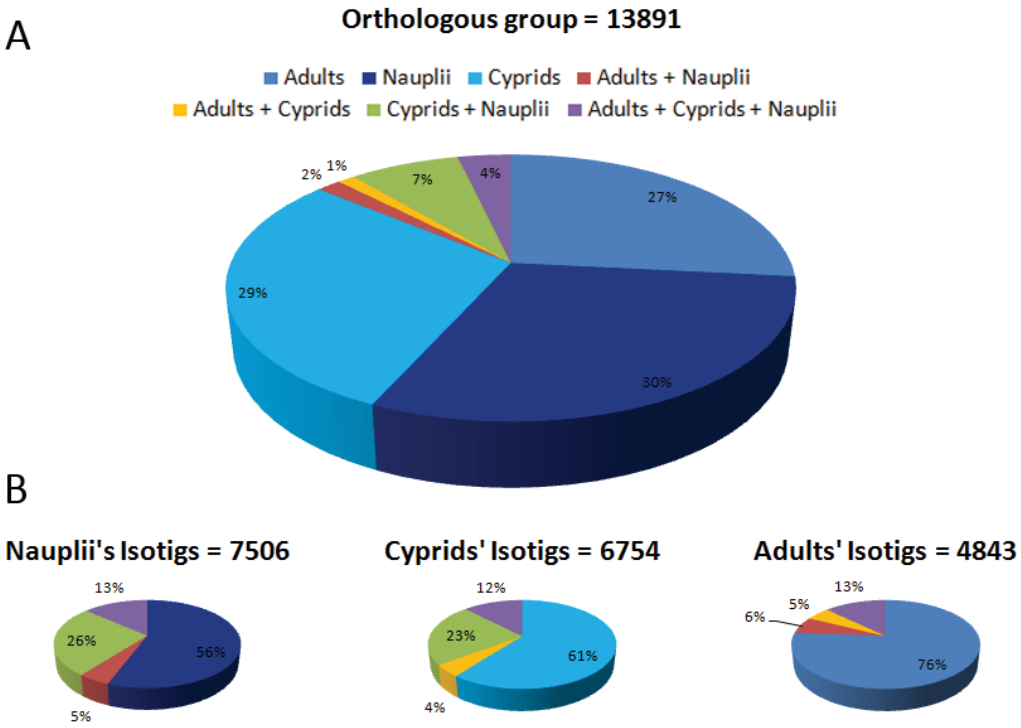


Figure 2.3 - **Proportion of isotigs expressed during the 3 developmental stages**

Isotigs contained in individually assembled libraries were BLASTed against each other and hits with e value < $1e^{-50}$ and identity > 95% were considered as matching sequences. In A, matching Isotigs were joined in orthologous groups before determining the developmental stage(s) in which every 'gene' was expressed. In B, single Isotigs in one library were BLASTed against Isotigs from the other two libraries and every hit counted.

Furthermore, the presence of settlement cues was capable of shifting the onset of *bcs-6* activation prior to attachment but it did not affect the other *bcs* genes. These results led the authors to conclude that *bcs-6* possesses a key regulatory role in larval attachment and metamorphosis in *B. amphitrite* (Li*, et al.*, 2010). In contrast, the present study found all 3 isoforms of *bcs-6* (*bcs-6a*, *b* and *c*) expressed in each library, suggesting that this gene is not cypris larva-specific. Although *bcs-6* may have multiple functions, its expression during naupliar and adult stages makes it unlikely to be a regulator of settlement, and its function during attachment should be reconsidered. Regarding other *bcs* genes, only *bcs-1*, *-3* and *-4* were found to be uniquely expressed during the cyprid stage. While *bcs-2* transcripts were present in both naupliar and cyprid ESTs, homologues to *bcs-5* were not found in this study.

Distribution of isotigs within Gene Ontology (GO) categories is commonly performed in order to assess the functions performed by RNA transcripts. In this study, having undertaken a cDNA normalisation process, nothing can be said about the relative

proportion of GO hits in the libraries, but important information can be obtained from the overall functional diversity of the genes expressed in each developmental stage. It was possible to annotate around 20% (966/4843) of isotigs of the adult with a GO number, with this proportion reduced to 11% (743/6754) and 8% (623/7506) for the cypris and naupliar stages respectively. These are relatively low percentages, reflecting the paucity of genetic information for the whole Vericrustacea (ex Maxillopoda) class, but yet not far from previous investigations (Fleury et al. 2009; Bettencourt et al. 2010).

Table 2.2 - ***Balanus amphitrite*** **gene with a significant match to deposited lectin-like proteins**

Isotig name*	Developmental Stage	Blast2n vs nt			Blast2n vs nr	
		Best hit	from species	identities	Best hit	identities
02490	A, C, N	Laminin alpha-4 chain	*Ixodes scapularis*	113/160	Extracellular matrix glycoprotein laminin	40/102
06444	A, C, N	Lectin 4 C-type	*Nasonia vitripennis*	111/136	C-type lectin	25/45
10053	C, N	Beta-1,3-glucanase	*Periplaneta Americana*	312/475	Beta-1,3-glucanase	97/158
13274	C, N	No hits			Hemolymph lipopoysaccharide-binding protein	48/126
13320	C, N	No hits			Predicted macrophage receptor 1	32/76
15095	C	No hits			Scavenger recptor class A like	44/103
15807	C, N	Lipoprotein receptor	*Bombix mori*	466/618	LDLa domain containing chitin binding protein	182/209
16152	C	No hits			Lectin BRA-3	41/132
17333	C				Hemolymph lipopoysaccharide-binding protein	42/111
17827	C	No hits			C-type lectin	no matching results given

As expected, most of the GO hits within the 'Biological Processes' category from each assembly matched translation-related transcripts (eg ribosomal proteins). Notable differences between the 3 developmental stages investigated were found in the following GO categories: 1) signal transduction (nauplii (N) = 4 hits; cyprids (C) = 2

hits and; adults (A) = 19 hits); 2) RNA-dependent DNA replication (N = 20; C = 15; A = 2); 3) proteolysis involved in cellular protein catabolic process (N = 1; C = 3; A = 14); 4) DNA integration (N = 2; C = 11; A = 1); 5) glycolysis (N = 1; C = 0; A = 10); 6) DNA replication (N = 1; C = 3; A = 10); and 7) vesicle-mediated transport (N = 0; C = 4; A = 10). A graphical representation of these results is given in Figure 2.4.

The molecular mechanism that allows kin recognition in *B. amphitrite* has been under intense study for decades (Crisp and Meadows 1962), and it is now clear that the settlement-inducing protein complex (SIPC) is fundamental to this process (Dreanno et al. 2006a). Our data show that the SIPC is expressed in all developmental stages investigated, confirming that this protein is not only produced by adults but that it may also possess a function in larval-larval interactions (Dreanno et al. 2006b). SIPC is a large glycoprotein present in all organs that are lined by adult cuticular tissues, and evidence seems to indicate that the carbohydrate component of the SIPC plays a fundamental role in kin recognition (Matsumura et al. 1998a; Matsumura et al. 1998b). Particularly, the addition of mannose-binding lectins to adult extract inhibited settlement induction by the SIPC (Matsumura et al. 1998a).

Lectins are a functionally diverse group of proteins possessing a carbohydrate-binding site (Kilpatrick 2002; Cambi and Figdor 2003), and their involvement in larval settlement has been hypothesised (Maki and Mitchell 1985). Therefore, a lectin-like compound(s) expressed by *B. amphitrite* may be responsible for sensing the SIPC-attached sugars. Accordingly, the 3 assemblies were searched for lectin genes. The initial 130 isotig (67 isogroups) hits obtained were filtered to remove random hits (eg plectin/S10), as well as ESTs that were similar to possible database biases (eg matching a ribosomal subunit with 98% identity as well as a lectin) and ESTs that were not expressed during the cyprid stage, and the remaining isotigs (10 isogroups) were used to draw Table 2.2.

Interestingly, the Blast2x vs KEGG tool showed that BAMP_Isotig_Isotig16152 matched a mannose receptor, C type 1, from *Xenopus laevis* with 49% of positively aligned amino acids ($3e^{-15}$). C-type lectins, which are Ca^{2+}-dependent, are a diverse group of sugar-binding proteins often involved in non-self recognition and innate immunity (Cambi and Figdor 2003). The TMHMM (Krogh, Larsson et al. 2001) and TMpred (Hofmann and Stoffel 1993) bioinformatic tools similarly predict that BAMP_Isotig_Isotig16152 is translated into a peptide possessing a transmembrane helix composed of 19 amino acids toward the N-terminus and with the C-terminus facing the external side of the membrane. Although *ad hoc* biochemical approaches are

required to confirm this experimentally, it suggests that the corresponding protein is located on the cell surface. Furthermore, BAMP_Isotig_Isotig16152 was uniquely expressed during the cyprid stage and it was determined by the perfect alignment of 8 reads. Its nucleotide sequence is 46% identical to that of BAMP_Isotig_Isotig15960, which is expressed uniquely during the adult stage and hits the same KEGG entry. These adult- and cyprid-specific predicted mannose-binding lectins are 33 and 31% identical to a lectin (BRA-3) isolated from *Megabalanus rosa* (Nobuhiko et al. 1994), and these 3 genes have likely evolved from a common ancestor. As many balanomorph and lepadomorph species contain SIPC-like molecules (Dreanno et al. 2006a), a comparative analysis of lectins in barnacles may suggest important structural features that allow kin recognition in different species. Furthermore, lectin-like proteins could also be involved in the perception of other environmental cues such as the bacterial biofilm (Khandeparker et al. 2003), and evidence of their involvement in settlement may suggest new biotechnological approaches to reduce its occurrence on submerged surfaces. The creation of the EST database presented here has the potential to significantly impact the design of studies aiming to understanding the molecular mechanisms underlying settlement.

2.4 Conclusions

The acorn barnacle *Balanus amphitrite* is commonly used as a model organism to investigate the regulation of settlement in marine invertebrate. However, the paucity of genomic information has hindered the understanding of the molecular mechanisms controlling the pelage-benthic transition. In order to resolve this issue, two independent sequencing surveys have been performed. For the first approach, a cDNA library was constructed by direct cloning of reverse transcribed adult RNA and 960 clones were sequenced. For the second approach, RNA from 3 developmental stages were normalised and directly sequenced by 454 titanium technology. All sequences were analysed and annotate using well-established bioinformatics tools. Altogether, these data suggest that around 2/3 of *B. amphitrite*'s gene content has been characterised. The creation of a publically accessible database containing this information will be of great support to further studies aiming at understanding the molecular regulation of development and settlement in this species.

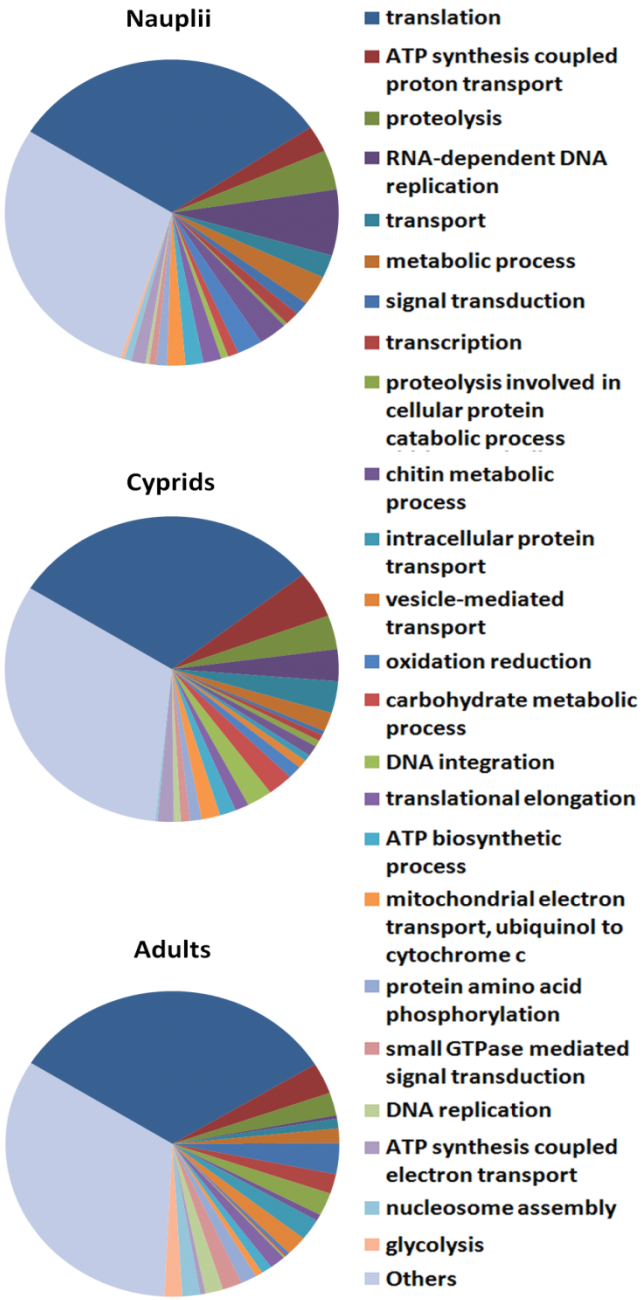


Figure 2.4 - **Distribution of GO hits within the 'Biological process' categories for the 3 normalised libraries**

GO subcategories for which matching sequences in none of the assemblies reached 1.7% (of the total hits) were placed in the 'Others' group.

Chapter 3. Best reference genes to normalise gene expression studies (qRT-PCR) in *Balanus amphitrite*

3.1 Introduction

Understanding the molecular mechanisms that regulate settlement in *Balanus amphitrite* is of great interest for both basic and applied research. Developing methods to assess gene expression in this organism is seen as central to gain this knowledge. Among the techniques used to monitor gene expression, quantitative real time PCR (qRT-PCR) is regarded as the most sensitive and reliable method to determine levels of mRNA transcription (Bustin, 2000; Nolan*, et al.*, 2006a). QRT-PCR is widely used in a number of biological disciplines including, but not limited to, medical research, developmental studies, forensic science and environmental topics (Deepak*, et al.*, 2007). This technique allows the kinetics of the PCR reaction to be monitored as it progresses (real time), and provides an indication on the initial amount of the transcript present in the sample under investigation (Nolan*, et al.*, 2006a). The detection of amplified products is allowed by the combined use of fluorescent dyes and a PCR thermocycler that can read the light emission at every progressing cycle, and two main reaction chemistries are exploited in this process (Bustin, 2000). The first relies on the use of fluorescent dyes, generally SYBR green, that bind specifically to double-stranded-DNA. The second method is based on the hybridisation of fluorescently-labelled probes to the correct amplicon, and the most commonly employed approach is the Taqman assay, which utilises the 5'-nuclease activity of the DNA polymerase to hydrolyse a hybridisation probe. Although the maths behind the calculation of gene expression level used in these two approaches do not substantially differ, a SYBR green-based assay has been developed in this study, and the following considerations focus on the use of this dye.

As SYBR green fluoresces when interlaced to double-stranded DNA, the increase in light emission recorded every PCR cycle follows the exponential progress of the reaction in which two specific primers are used to amplify a target cDNA sequence. Therefore, a typical fluorescence curve is logarithmic, characterised by an initial phase where the background noise masks the amplification, an exponential phase where PCR products double at every cycle, and a plateau phase where the fluorescence stabilises as a consequence of SYBR green or primer depletion, or polymerase inactivation. The qRT-PCR user commonly sets a fluorescence level (threshold) that is reached during the

exponential phase, and thermocycler data are recorded as Ct(s); the number of cycles required for the fluorescence curve to cross the threshold. Equal amounts of two genes can produce different Cts as a consequence of diverse amplification efficiencies. Therefore, it is fundamental to calculate the efficiency of each primer pair tested. One commonly employed method is to determine the slope of the line obtained by joining the Cts of 5 serial dilutions of template cDNA (Rasmussen, 2001). From the slope of the line, the efficiency (in percentage) is then calculated with the formula:

$$\text{Efficiency} = -1 + 10^{(-1/\text{slope})}$$

The application of qRT-PCR has proved particularly useful for comparative studies, where the expression of genes of interest (GOIs) in different samples is measured against the expression of endogenous reference genes (RGs). This normalization procedure is fundamental to minimizing inherent variability introduced during the RNA extraction or the reverse transcription steps (Bustin & Nolan, 2004; Huggett*, et al.*, 2005). Ideally, RGs should both maintain a stable transcription level in all cells, tissues or individuals under investigation and should not be influenced by the experimental conditions. House-keeping genes are often chosen to normalize qRT-PCR data, and those encoding for beta-actin, glyceraldehyde-3-phosphate dehydrogenase, elongation factors and ribosomal proteins have been widely used as internal standards (Spinsanti*, et al.*, 2006; Ahn*, et al.*, 2008; Infante*, et al.*, 2008). However, many studies have stressed that universal gene for data normalization do not exist and the utilization of an inopportune reference transcript can generate misleading results (Huggett*, et al.*, 2005). For this reason, the selection of the best RG(s) should be validated in every qRT-PCR assay designed (Bustin*, et al.*, 2005; Dheda*, et al.*, 2005).

A common way to determine the most stable genes is to comparing their expression level on a number of samples and modelling the Ct values obtained. There are a number of methods to do this analysis, including the use of software such as *GeNorm* (Vandesompele*, et al.*, 2002), *BestKeeper* (Pfaffl*, et al.*, 2004) and *NormFinder* (Andersen*, et al.*, 2004). *GeNorm* is a VBA applet for Microsoft Excel which was specifically developed to undertake this analysis (Vandesompele*, et al.*, 2002). The software determines the stability measure (M) of a gene within a pool, by analysing relative quantities obtained from qRT-PCR runs. The value M represents the mean pair-wise variation between a RG and all other tested candidates. The gene with

the highest M is then excluded from the analysis and the calculation is repeated in a stepwise fashion that allows gene ranking until the best two genes are found. *GeNorm* also generates a graph which indicates the pair-wise variation V between two sequential normalization factors containing an increasing number of genes. The authors of the applet suggest a 0.15 cut-off value of V, meaning that higher values indicate a big variation in the effectiveness of the normalization, and recommended that another RG is included in the assay (Vandesompele*, et al.*, 2002).

BestKeeper is an Excel-based tool determining the optimal RGs by using a pair-wise correlation analysis of all pairs of candidate genes and calculating the geometric mean of the best suited ones. After the descriptive statistics for the individual candidate, the first estimation of RG expression stabilities is given, based on the inspection of calculated standard deviation (SD) and coefficient of variance (CV) values. According to the variability observed, RGs are ordered from the most to the least stably expressed. Any studied gene with the SD higher than 1 can be considered inconsistent (Pfaffl*, et al.*, 2004) and is excluded for the calculation of the *BestKeeper* index. This index represents a measure of the variability of a specific sample, calculated as the geometric mean of the Ct values of all genes studied. Afterwards, the pair-wise correlation between genes and the correlation between each gene and the index are computed, describing the relation between the index and the contributing gene. In other words, the software attempts to normalise the Ct data by estimating the biases introduced by initial differences in the amount of cDNA used as template and suggests the genes that best represent the overall change in gene expression. Therefore, *BestKeeper* calculates the most stably expressed genes as well as the genes that best fit the index.

The third method described here in order to identify the optimal normalization gene among a set of candidates is *NormFinder*, another VBA applet. It provides a stability value for each gene, which is a direct measure for the estimated variability in expression, enabling the user to evaluate the systematic error introduced when using the gene for normalization. *NormFinder* also provides a measure of the stability of gene in different groups and estimates any bias in the expression of the genes between the groups based on two-way ANOVA. It calculates intra-group variability for the genes in each of the groups and the inter-group variability, or bias, between the groups. The optimum reference gene is the one that shows lowest total variability and the optimum pair shows the combined lowest variability (Andersen*, et al.*, 2004). Therefore, this software estimates the biases introduced by technical replicates as well as those determined by the different treatments (e.g. developmental stages, tissues etc).

The aim of this study was to develop a qRT-PCR assay to monitor gene expression in different developmental stages and in individuals exposed to morphogenetic cues. The ability of 11 *B. amphitrite* transcripts to normalize qRT-PCR data was determined by comparing relative quantities obtained from cDNAs representing 14 different samples and 7 developmental stages. The software *GeNorm* (Vandesompele*, et al.*, 2002), *BestKeeper* (Pfaffl*, et al.*, 2004) and *Normfinder* (Andersen*, et al.*, 2004) were used to obtain an estimation of the expression stability of each gene and, by comparing the results, to identify the most suitable genes for qRT-PCR data normalization in *B. amphitrite*.

3.2 Materials and methods

3.2.1. Balanus amphitrite, culturing and RNA extraction

Wild *B. amphitrite* adults were collected from Beaufort, North Carolina, USA (courtesy of Prof. D. Rittschof), and reared as described in chapter 2. In order to determine the best reference genes in this species, RNA was extracted from the following developmental stages:

N-1) naupliar instar I - just hatched;

N-6) naupliar instar VI - three-eyed stage;

C-0) young cyprids - recently metamorphosed;

C-3) mature cyprids - these are standard larvae for settlement assays and they are maintained for 72 h at 6º C after the C-0 stage;

C-I) mature cyprids (same as C-3) that have been exposed to sea water containing 10^{-5} M of 3-isobutyl-1-methylxanthine (IBMX) at 28º C for two hours (Clare*, et al.*, 1995);

J) juveniles collected ~24 hours after settlement onto glass slides;

A) adults.

All larvae were isolated under a dissecting microscope and placed in a 1.5 ml tube kept on ice. Cyprids were harvested using a 250 µm plankton mesh filter, washed thoroughly with sterile seawater and checked under the microscope for contamination by nauplii. The tubes containing the larvae were centrifuged briefly and the residual seawater was removed. The larvae were then resuspended in TRIzol (Invitrogen) prior

to storage at -20º C. Settled juveniles were collected by scraping them off the glass slides using a sterile scalpel. For the adult stage, the pooled soft tissues of ten individuals were dissected and ground under liquid nitrogen prior to RNA extraction.

After the larval stages were homogenized by pipetting and vigorous shaking the total RNA was extracted from each biological replicate using 1 ml TRIzol. The extracted RNA was then stored in 1 ml of isopropanol at -20º C. Prior to cDNA synthesis the stored RNA was precipitated by centrifugation at 12,000 g for 5 min at 4º, washed twice with 1 ml of 70% ethanol and then resuspended in milliQ water. The RNA purity and quality were evaluated using a NanoDrop ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies) and the quality was confirmed by gel electrophoresis.

3.2.2. Primer design

Tentative contiguous sequences (TCs) for RGs were analysed by Primer3 release 1.1.0 (Rozen & Skaletsky, 2000) (http://fokker.wi.mit.edu/primer3/input.htm) using the following parameters: a) product size range: 80-180; b) primer size: min 16, opt 19, max 22; c) primer Tm: minimum 55º, optimum 60º, maximum 65º; d) primer GC%: minimum 40, optimum 50, maximum 60; e) all other parameters were left as the default. Oligonucleotides (synthesized by Invitrogen) were resuspended as stock solutions containing 0.7 pmol/µl of both the forward and reverse primers. The cDNA obtained from adult RNA was initially used to visualize the melting curve of PCR products and to determine the possible formation of PCR artefacts. PCR products were also visualised by gel electrophoresis to check amplicon sizes.

3.2.3. cDNA synthesis, primer efficiencies and cycle parameter for qRT-PCR

Total RNA (1 µg) was reverse transcribed from each sample for 20 min at 42 ºC with the QuantiTect kit (Qiagen). After the genomic wipe-out step and prior to the reverse transcription, 1 µl was collected from each reaction to be later used as a negative-RT control to check for genomic contamination. Serial dilutions of 1:5, 1:10, 1:25, 1:50, 1:250, 1:500, 1:5000 and 1:50000 were then made from the initial 20 µl of adult cDNA. Each qRT-PCR experiments comprised 12.5 µl of Faststart SYBR green (Roche Diagnostics Ltd), 10.5 µl of stock primers (final concentration 0.3 µM each) and 2 µl of cDNA. Reactions were performed in sealed 96-well plates using a Chromo4 Research thermocyler and analyzed with the Opticon Monitor 3 software (BioRad).

The qRT-PCR thermal profile consisted of an initial step at 95 ºC for 5 min, followed by 40 cycles of 15 s at 95 ºC and 1 min at 60 ºC. A final elongation step at 72 ºC was included before the melting curve was determined by monitoring SYBR green fluorescence during the temperature ramp of 60 to 95 ºC with an increase of 0.5 ºC and a hold of 1 s. Primer efficiencies were determined using five cDNA dilution points for each primer pair that were chosen according to the expected expression level of the corresponding gene. Triplicates were tested for each dilution point and primer pair, together with a duplicate negative control that contained sterile water instead of cDNA. The resulting efficiency graphs are presented as Supplementary Figure S1. To determine the best RGs, 2 µl of the cDNA diluted 1:50 were used for all 14 samples and primers tested.

3.2.4. qRT-PCR data analysis

When required, raw Ct values were transformed to relative quantities by a comparative method based on the formula: $1/E^{(Ct\ value - lowest\ Ct)}$; where E is the primer efficiency and the lowest Ct refers to the smallest value obtained with each specific primer pair. The most stable RGs were then determined using software *geNorm* 3.5 (Vandesompele*, et al.*, 2002), *BestKeeper* (Pfaffl*, et al.*, 2004) and *NormFinder* (Andersen*, et al.*, 2004).

3.3 Results and Discussion

There is great interest in elucidating those genes involved in barnacle development and settlement. In this respect, qRT-PCR is particularly suitable to monitor how external cues, such as environmental variables, the presence of conspecific individuals or the occurrence of biofilm and/or of certain microorganisms, influence gene expression prior to and during settlement and metamorphosis. At the time of designing this study, the only genetic information available for *B. amphitrite* was that provided by the EST library (un-normalised) described in Bacchetti De Gregoris *et al.* (2009). Since most of the annotated ESTs were found to represent highly expressed housekeeping genes, this suggests that information from a few hundred clones is sufficient to provide a set of candidate genes among which the best RGs for subsequent qRT-PCR studies can be searched.

14 potential RGs were selected and designed PCR primer pairs to them (Table 3). These genes were chosen either because they were commonly used for other organisms or because they were often found in the EST library and their functional description indicated they might be useful candidate genes. Of the 14 primer pairs tested, three (gapdh, ald, mlc1) produced PCR artefacts and therefore were discarded from the analysis without attempting to design new primer pairs for them as a total of 11 genes were considered to be adequate for the analysis.

Table 3 - **List of primers and reference genes under investigation**

Gene	Gene's symbol	Forward primer	Reverse primer	Amplicon length	Primer efficiency
Ubiquitin c	*ubc*	GCGTCATAAGTTGCGGAGA	TCTTGGCCTTCACATTTTCA	106	100%
Fructose bisphosphate aldolase	*ald*	TATGTCCCAGCGTTGTGCT	TGGCACCAGACCATTCATT	166	Non-specific products
NADH dehydrogenase subunit 1	*mt-nd1*	CGGGCTGTTGCTCAAACTA	TTCGACAAAATCTTCCAATCT	102	100%
Tubulin alpha	*tuba*	CCTGCTGGGAGCTGTATTGT	ACAACAGTGGGCTCCAAATC	169	94%
NADH dehydrogenase subunit 4L	*mt-nd4L*	TTCTTGGTAGCTTCTGTGTGTG	TAGTCGGAACCATGTGATCG	80	100%
Tubulin beta	*tubb*	ACCTCAGCCTGGTCATCATC	GGCTTTCCTCCACTGGTACA	165	84%
Cysteine protease 1	*cp1*	GTTGAGCAGCACATGAAGGA	CGAACTCCTCAGAGGTCAGG	91	94%
Cytochrome b	*mt-cyb*	GGACACTGCATGCTAATGGA	AGGCAGCAGCCATAGTCAAG	144	91%
Acyl carrier	*mt-acp*	GATGTGGCGATTTGCTATCC	TTCTCCGGGTTGATCTTGTC	175	93%
Myosin 1-light chain	*mlc1*	AAGGATGAGGTTGACGCCTA	ACCCTGGTCCTTGTCCTTCT	174	Presence of primer dimers
60s ribosomal protein L15	*rplL5*	AAGCAGGGATACGCCATCTA	AGCTTCAGCTCGTTCACTCC	116	84%
Glyceraldehyde-3-phosphate dehydrogenase	*gapdh*	TCTGCGGCTTACTTGTCCTT	ACTCGCACTCGAGCATCTTT	154	Non-specific products
Elongation Factor 1 alpha	*ef1a*	GCCACAGGGATTTCATCAAG	TGGAGATACCAGCCTCGAAC	105	100%
Actin alpha	*act*	CAGTCCAAGCGTGGTATCCT	CGCACGCAGCTCGTTGTAGA	114	100%

Primer efficiencies of the remaining 11 RGs varied between 84% and 100% (primer efficiency graphs are provided in Supplementary Figure S1). To obtain reliable results it is important to achieve equivalent PCR efficiencies for the reference genes and the gene of interests. Therefore, pairs with a comparatively lower efficiency (84%) were kept in the analysis as they can be useful when investigating GOIs for which it is difficult to design primers with high efficiency.

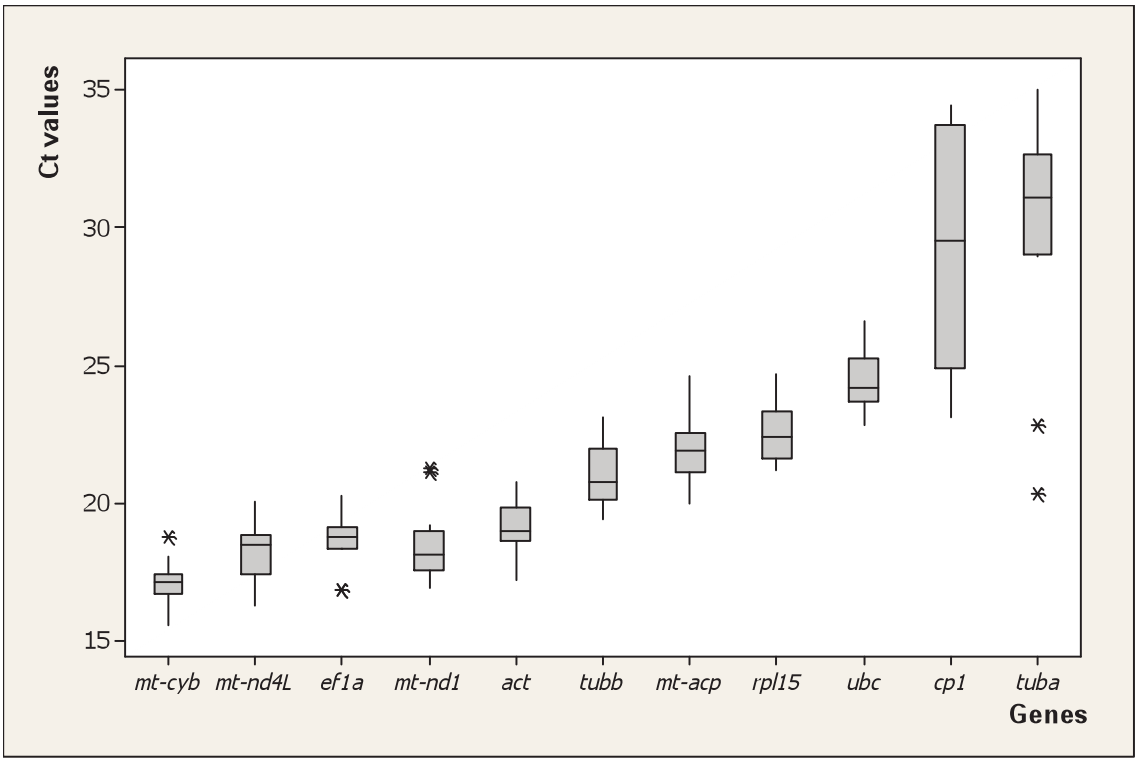


Figure 3.1 - **Comparative expression of the analysed genes**

Box-and-whisker plot representing the expression level (threshold cycles) of candidate reference genes in *B. amphitrite* (*n* = 14). The box plot, obtained using the software Minitab, shows the smallest observation, lower quartile, median, upper quartile, largest observation and indicates Ct value that might be considered outliers.

The expression levels of RGs were obtained from qRT-PCR reactions in the form of threshold cycle (Ct) values (Figure 3.1). The 14 Ct values collected for each primer pair were derived from the two biological replicates of the seven developmental stages under investigation. These samples were initially considered independent and the data they generated were analyzed with *geNorm* (Vandesompele*, et al.*, 2002), *BestKeeper* (Pfaffl*, et al.*, 2004) and *NormFinder* (Andersen*, et al.*, 2004) to determine the most steadily expressed genes. It can be argued that two biological replicates are not enough for statistical analysis, and this is the case for many biological systems (e.g. comparing different tissues, single individuals) where at least five replicates should be performed.

In our investigation, however, the RNA was extracted from ten (the adult stage) to hundreds of pooled individuals, so that the RNA could be considered an average sample of the developmental stage analysed. The high correlation found between the two biological replicates, at least for the most stably expressed genes (figure 3.2), confirmed the expectations.

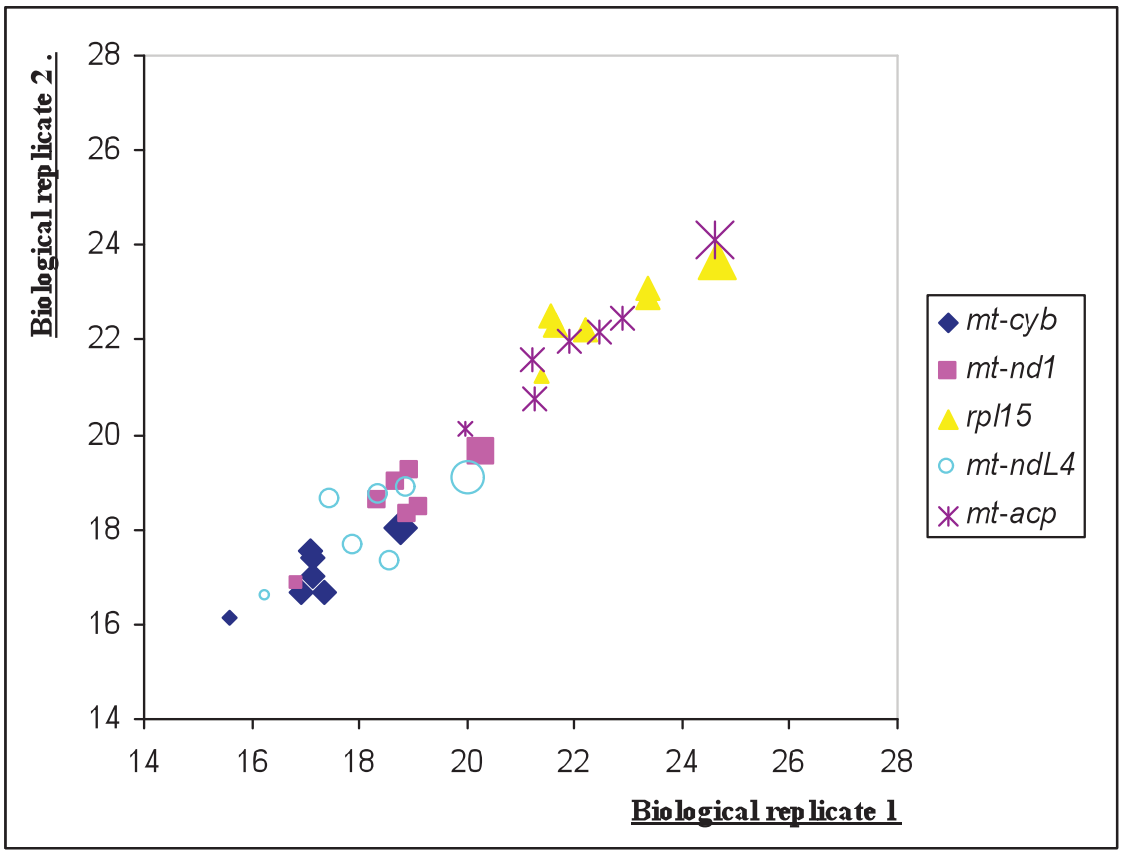


Figure 3.2 - **Correlation between biological replicates for the 5 best reference genes**

The Ct values (adjusted to primer efficiencies) obtained for the seven developmental stages we analysed were plotted for the five best reference genes. The size of the shape indicates the developmental stage: the smallest shapes represent values from just-released nauplii, whereas the largest represent values obtained from adult barnacles.

The software *geNorm* provides a ranking of the tested genes based on their stability measure (M), determining the two most stable RGs or a combination of multiple stable genes for normalization. The value M represents the mean pair-wise variation between a gene and all other tested candidates. The gene with the highest M is then excluded from the analysis and the calculation is repeated in a stepwise fashion that allows gene ranking until the best two genes are found. According to *geNorm*, the two most stable genes in the assay were *mt-nd1* and *mt-cyb* (Figure 3.3), with an M

value of 0.41. The threshold value M for considering a gene to be unsuitable for data normalization is suggested to be ≥ 1.5 (Vandesompele, *et al.*, 2002). Low values of the pair-wise variation V between two sequential normalization factors containing an increasing number of genes showed it was unnecessary to include another RG in the protocol (Figure 3.4). However, as *mt-nd1* and *mt-cyb* are both contained in the mitochondrial genome, it may be advisable to include a nuclear gene in the normalization strategy. In this case, *geNorm* suggested that either *act* or *ef1a* should be used.

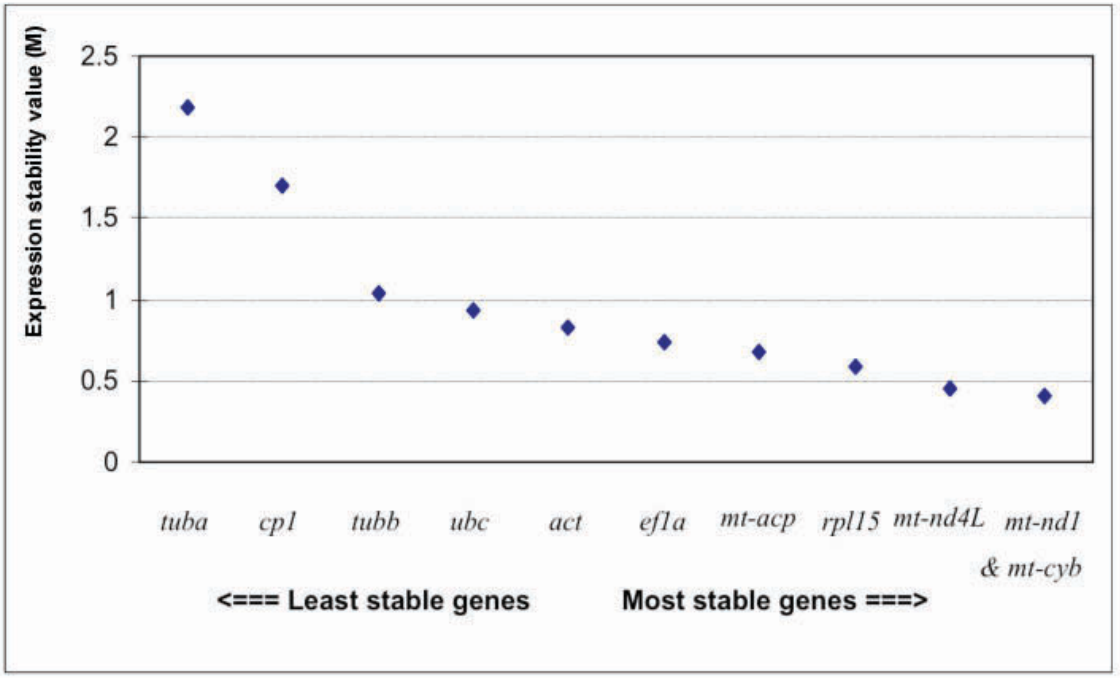


Figure 3.3 - **Gene expression stability M of candidate reference genes calculated by *geNorm***

The *geNorm* programme proceeds to the stepwise exclusion of the genes whose relative expression levels are more variable among samples. Data points represent the average expression stability values of remaining reference genes.

The highly variable *tuba* gene was excluded from the analysis using *BestKeeper* (Pfaffl, *et al.*, 2004), which can only consider 10 RGs. Since any gene calculated by *BestKeeper* to have a standard deviation >1 can be considered inconsistent, *cp1* (standard deviation = 3.69) was also excluded from the final calculation of the *BestKeeper* index. This index is a representation of the average over/under-regulation of all genes together in every developmental stage. The RG that best correlated with the *BestKeeper* index was *mt-acp*, followed by *rpl15, ef1a* and *mt-cyb* (figure 3.5). In other

words, *mt-acp* appeared to be the best candidate to represent the overall modulation of expression of the 9 RGs analysed. However, *mt-acp* was also the least stable gene and this was likely to influence the index to which it correlated. Nevertheless, the most stable gene, resulting from the *BestKeeper* analysis, was again *mt-cyb*, followed by *mt-nd1* and *rpl15* (figure 3.5).

NormFinder attempts to identify the optimal normalization gene among a set of candidates and provides a measure of the stability of genes' expression in different groups and at the same time estimates any bias in the expression of the genes between the groups based on two-way ANOVA (Andersen*, et al.*, 2004). When all data were analysed together, the most stable RG candidates in the assay using *NormFinder* were *mt-cyb* (stability value = 0.127), *rpl15* (0.137) and *mt-acp* (0.165), as shown in figure 3.6. The analysis was then repeated grouping samples by developmental stage to assess intergroup variability. In this case, the best genes that allow comparison of different developmental stages and/or treatments in *B. amphitrite*, which was the goal of this study, were *mt-cyb* (0.159), *mt-nd1* (0.167) and *mt-acp* (0.168), suggesting that these are the most suitable genes for data normalization.

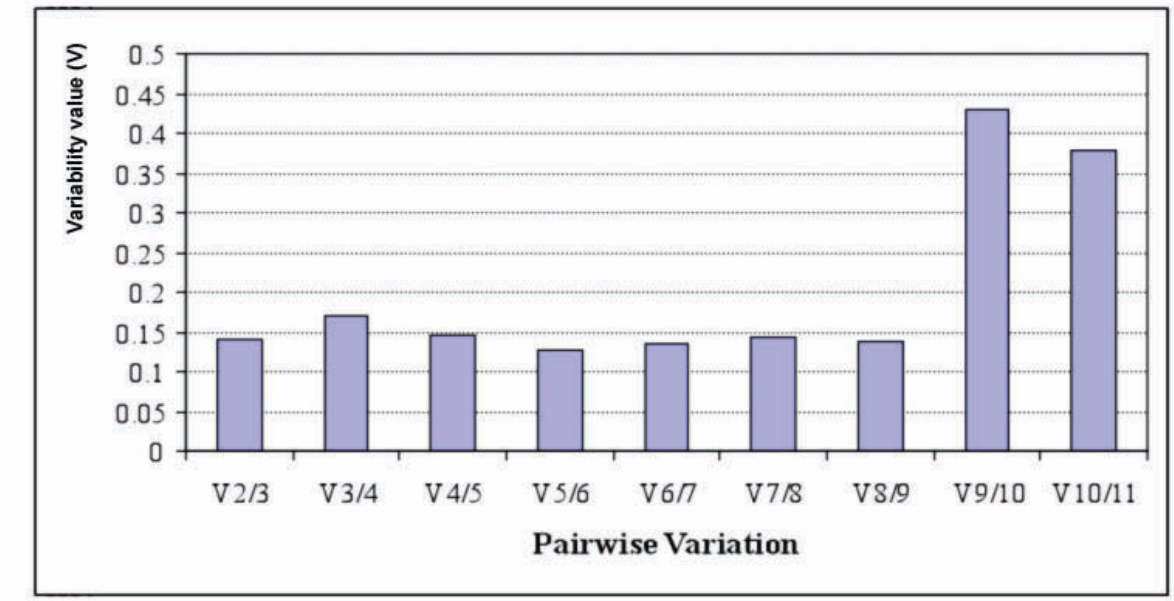


Figure 3.4 - **Determination of the optimal number of reference genes for data normalization**

Bar values indicate the magnitude of the change in the normalization factor after the inclusion of an additional reference gene. The authors of *geNorm* suggest that V > 0.15 should be considered as the threshold to include an extra RG into the assay.

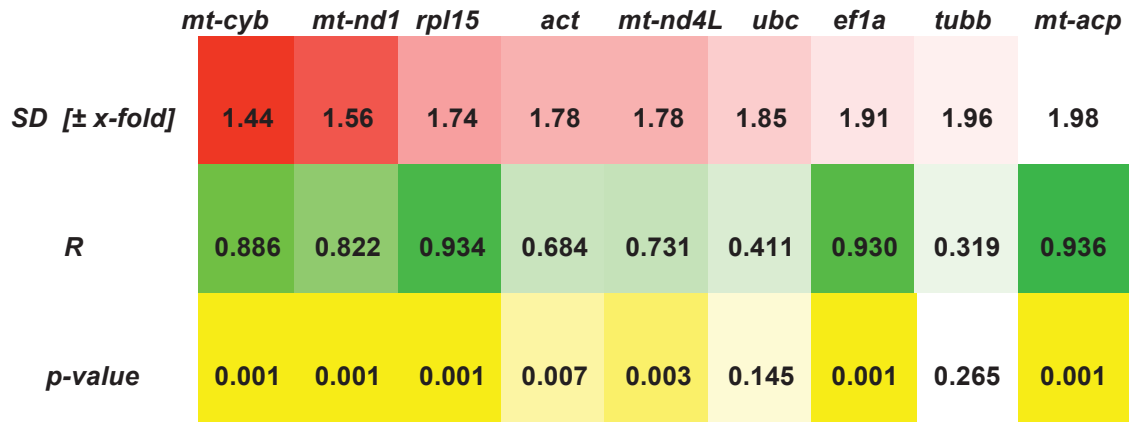


Figure 3.5 - **Results from *BestKeeper* correlation analysis.**

BestKeeper calculates the stability measure for each candidate gene and then ranks them from the most to the least stable (SD [± x-fold]). The coefficient of correlation (R) and the p-value measure the correlation between each gene and the *BestKeeper* index. For each variable presented in the figure (SD [± x-fold], R and p-value), genes that ranked comparatively better are highlighted with a more intense cell colour.

A final examination was performed adding the data to *NormFinder* as two subgroups (the two biological replicates for each developmental stage). As a result, the software produced the same gene ranking by their stability values, showing a very low variability between replicates, which was also confirmed by further statistical analysis; Pearson correlation coefficients of biological replicates for all RGs tested ranged between 0.711 and 0.979, with 9 genes out of 11 showing a significant correlation at the 0.01 level (2-tailed). Although the use of only 7 data points may affect the examination, the results suggested that the implemented protocol is effective in capturing meaningful differences in gene expression throughout *B. amphitrite* development.

As a general consideration, although *geNorm*, *BestKeeper* and *NormFinder* have the same aim, they employ different strategies to calculate the most stable genes and it is unlikely that they will give the same results. For example, looking at the absolute ranking of best genes, *mt-acp* scored 5th, 9th and 3rd. However, its stability values as determined by the tree software do not change substantially from that of the genes ranked closely (e.g. the value obtained by *BestKeeper* for *mt-acp* (9th) was 1.98 and that of *act* (4th) was 1.78). Finally, it was noted that Ct values for the best RGs tended to increase during the life cycle. This was particularly evident with the cDNA derived from adult barnacles, which required ~3 to 4 more cycles to reach the PCR exponential phase in comparison to the cDNA from larvae that had just hatched (Figure 3.2). While we cannot exclude the possibility that the genes analysed are down-regulated in the

adult stage, this trend could also be explained by the presence of reverse transcription inhibitors that concentrate or are synthesized in later stages of *B. amphitrite* development, as RT-inhibitors are known to be one of the main sources of variability in qRT-PCR experiments (Bustin & Nolan, 2004). Although the CT value shifts remained in an acceptable range, it may be advisable to include a reference assay to rule out the presence of inhibitors. This is commonly achieved by adding an aliquot of the RNA under investigation to a well characterised exogenous RNA and measuring the effect on the amplification of the cDNA derived from the latter (Smith*, et al.*, 2003; Nolan*, et al.*, 2006b).

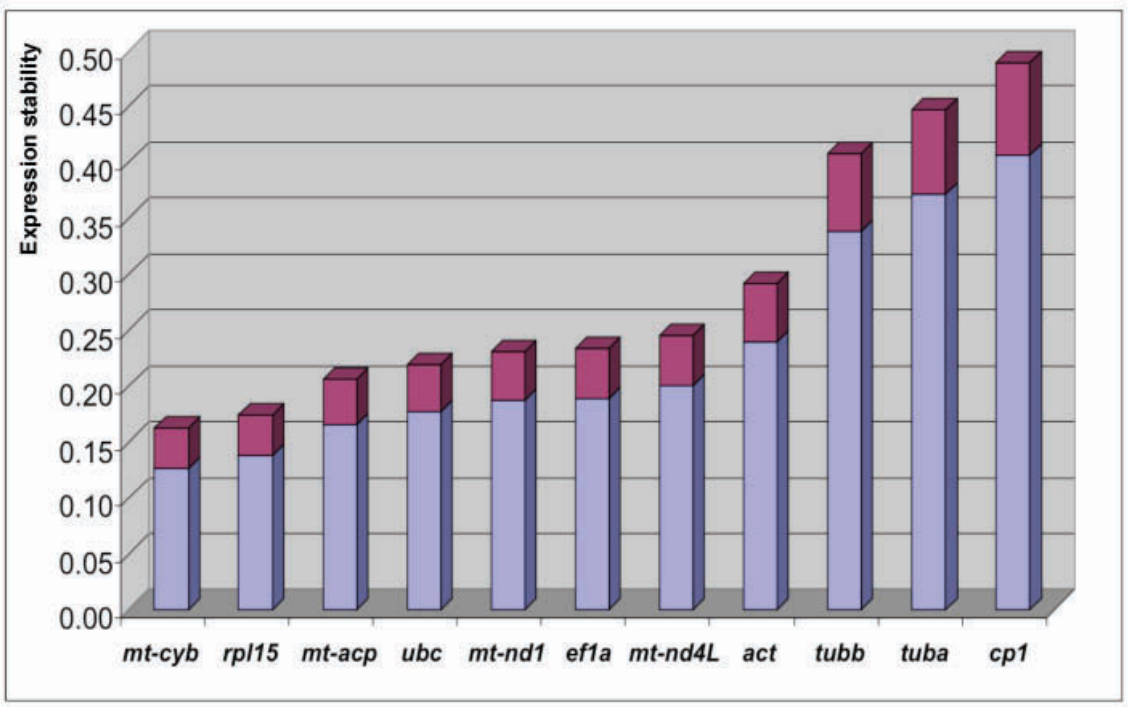


Figure 3.6 - **Determination of the most stable reference genes using *NormFinder***

The *NormFinder* algorithm ranks the set of candidate normalization genes according to their expression stability in a given experimental design. Blue bars represent the stability values of our candidate genes, while purple bars indicate their standard error.

3.4 Conclusions

Balanus amphitrite is already established as a model organism to study the pelagobenthic life cycle. Here, the development of a qRT-PCR assay is described to investigate gene expression in this organism. The three programmes implemented to analyse qRT-PCR results indicated that *tuba*, *tubb* and *cp1* are unsuitable genes for data normalization. They also showed that *mt-cyb* itself, and the pair *mt-cyb* - *mt-nd1*, were

the genes expressed most stably throughout life cycle of *B. amphitrite*, and so their use is recommend as reliable reference genes in future qRT-PCR experiments. Other genes that performed well in the analyses were *mt-acp*, *rpl15, mt-nd4L*, *ef1a*, *ubc* and *act*.

Chapter 4. A new taxon-specific method to investigate microbial communities

4.1 Introduction

Mapping the distribution of different species in natural ecosystems is fundamental to reveal ecological dynamics. For macro-ecologists, these data are often obtained by visual counting of sample areas, while the microscopic nature of bacteria and the inability to determine their taxonomic affiliation by eye have forced microbiologists to develop alternative strategies. Nowadays, a variety of methods exist to characterise microbial communities, and they can be divided into two groups: culture-dependent and culture-independent methods. The first approach relies on the isolation of bacteria on artificial media and their axenic growth under laboratory conditions. This was the only available method to study micro-organisms until the 1970s and much of our knowledge of the metabolic processes performed by bacteria was obtained thanks to their cultivation. Furthermore, the use of model organisms such as *Escherichia coli* and *Bacillus subtilis* in combination with molecular techniques has provided much of the information regarding the genetic regulation of the bacterial cell cycle, reinforcing the importance of such studies. However, the use of culture-dependent techniques to characterise microbial communities has serious limitations. Specifically, it has become clear that the vast majority of bacteria do not grow on artificial media (Whitman*, et al.*, 1998) and this unculturable proportion may account for more than 99% of the total diversity present in any given ecosystem (Amann*, et al.*, 1995). In fact, it is acknowledged that the use of high-nutrient content media in microbial surveys tends to allow the isolation of cosmopolitan heterotrophic bacteria (e.g. *Vibrio* or *Pseudoalteromonas* spp.), but misses many other ecotypes that may have important functions in the investigated habitat. For this reason, although the cultivation of bacteria is still seen as key to understand their ecological significance (Giovannoni & Stingl, 2007; Hunt*, et al.*, 2008), culture-dependent studies simply intended to describe the existing diversity of microbes are becoming rare. Instead, the vast majority of descriptive studies are based on molecular approaches.

Culture-independent techniques rely on the initial extraction of total genomic DNA (utilisation of RNA is also possible) from a sample, followed by the downstream characterisation of such genetic material. For many years, the most widely applied method used to characterise uncultured microbial communities was based on the amplification by PCR of 16S rRNA genes using universal primers and the subsequent

cloning and sequencing of the PCR products; an approach also known as metagenomics (Handelsman, 2004). The strength of sequencing-based techniques is the unmatched capability to identify bacteria at the lower taxonomic level, which is the reason why they have been considered the mainstay of characterising microbial assemblages for years. Although the recent development of ultra-high throughput sequencers now dwarf classical metagenomic approaches (Sogin*, et al.*, 2006), their utilisation is still costly, and microbial ecolosists often need to rely on other culture-independent strategies. The many existing methods used to compare bacterial communities can be divided into two groups: those aiming to quantify specific taxa across samples and those aiming to determine a microbial profile from each community under investigation (fingerprinting).

With regard to the first approach, the redundancy of 16S rRNA sequences in a clone library has often been used as a mean to estimate the abundance of the bacterium the sequence belongs to. However, the creation of clone libraries is affected by PCR biases (v. Wintzingerode*, et al.*, 1997; Sipos*, et al.*, 2007), where uneven primer affinities to target DNA result in a preferential amplification of certain sequences compared to others. Inference of population densities and dynamics from metagenomic data can, therefore, lead to serious misinterpretations (Morales & Holben, 2009). To overcome this problem, fluorescence *in situ* hybridisation (FISH) is often applied to quantify specific microbial populations and to gain insight into the spatial arrangement of bacteria in their habitats (Wagner*, et al.*, 2003). This technique has widespread applications in environmental microbiology and around 1,500 rRNA-targeted oligonucleotide probes have been characterised (Loy*, et al.*, 2006). However, the limited number of commercially available fluorescent dyes and the time-consuming microscopic counting required for quantitative data collection are the major limitations that hinder FISH-based investigations to encompass high sample numbers.

As an alternative, quantitative PCR (qPCR) has been used by microbial ecologists to determine the abundance of specific phylotypes for more than a decade (Becker*, et al.*, 2000; Suzuki*, et al.*, 2000). Since the kinetics of the qPCR reaction can be followed in real time, this technique has the potential to recognise and eliminate some of the biases characterising end-point PCR (Nolan*, et al.*, 2006a). Furthermore, compared to other methods, qPCR is relatively inexpensive and can be semi-automated, making it the option of choice for high throughput analysis. This technique has developed from gene expression studies, where the level of expression of the gene of interest is determined in a number of samples and normalised to an endogenous control (Bustin &

Nolan, 2004; Huggett, *et al.*, 2005). In order to avoid collecting misleading results, it is currently believed that the determination of primer efficiencies and the choice of a good reference gene are the most significant steps to validate qPCR data. Although these two steps are also considered fundamental for characterising microbial communities, the use of 16S rRNA as the gene of interest poses two serious limitations. The first problem is related to the variability of ribosomal RNA operon copy number in bacterial species (Klappenbach, *et al.*, 2001). Therefore, it is important to stress that 16S rRNA gene numbers from environmental samples cannot be directly converted to cell numbers if the operon copy number possessed by the resident species is unknown. On the other hand, the 16S rRNA gene offers limited possibility to design primers (for SYBR green detection, as opposed to probes used in TaqMan assays) that are specific to the desired bacterial taxa (Smith & Osborn, 2009), and the exponential increase in 16S rRNA sequences submitted to public databases makes it necessary to periodically re-evaluate existing primers. For these reasons, qPCR finds its best application in investigating how the abundance of specific taxa change over time and space, rather than in characterising single communities, in which the technical drawbacks of qPCR are likely to introduce serious biases.

Moving from quantitative to fingerprinting approaches, a number of methods have been developed to allow community profiles to be compared (e.g. T-RFLP, ARISA and DGGE). Among these, denaturing gradient gel electrophoresis (DGGE) is commonly used to visualise the presence of different genotypes in a sample (Muyzer, *et al.*, 1998). The strength of DGGE resides in the relatively inexpensive production of informative data on the general diversity of the community and the possible downstream identification of specific community members by single band sequencing. The use of so called 'universal' primers in the PCR step can result in uneven amplification efficiencies for different 16S rRNA genotypes, and the extrapolation of quantitative data from band intensities is open to criticism (Heuer, *et al.*, 2001). However, when samples are processed consistently, the same biases are likely to apply to each sample and it has been argued that for a reoccurring band, relative quantification can be inferred from comparable DGGE profiles (Woodcock, *et al.*, 2007). Nevertheless, because of the immense scale of microbial diversity (Curtis, *et al.*, 2002), a single DGGE profile obtained using 'universal' primers is likely to miss minor constituents of the microbial complex. To overcome this problem, a suggested approach to enhance the number of discernible bands is to fraction a DNA sample prior to DGGE analysis (Holben, *et al.*, 2004; Dar, *et al.*, 2005). As part of this strategy, the use of phylum- and class-specific

primers for DGGE was shown to improve the characterisation of microbial communities (Mühling, *et al.*, 2008). Compared to the ~18 and ~25 bands obtained from two different samples with "universal" primers alone, Mühling *et al.* (2008) managed to increase the total number of discernible bands to ~75 and ~65 respectively using a nested PCR-DGGE protocol with 7 taxon-specific primer pairs. Thus, the DNA fractioning approach based on developing higher taxa-specific DGGE profiles has the potential to provide a more detailed description of the microbial assemblages.

Apart from the impact that a cheap technique for the quantitative profiling of bacterial communities may have on the field of microbial ecology, a second rationale inspired the development of a phylum- and class-specific approach. In the introduction to this thesis we have seen how the high level of lateral gene transfer in bacteria makes it difficult to generate taxonomic groups that also make sense ecologically (Feil, *et al.*, 2001; Hanage, *et al.*, 2005; Whitaker, *et al.*, 2005; Fraser, *et al.*, 2007). For this reason, investigators have often focused on bacteria at the sub-species level, and it is generally believed that genes, more than taxa, are ultimately responsible for their distribution (i.e. their adaptability). However, the existence of genes that are unique to all members of high taxonomic groups such as phyla or classes (Gupta & Lorenzini, 2007; Gupta & Mok, 2007; Gao, *et al.*, 2009) suggests that ecological consistency may also be found in deep branches of the tree of life. For example, differences in the cell wall structure between Gram positive and Gram negative bacteria may provide them with unique advantages that are not detected when characterising the community at lower taxonomic levels. Interestingly, by estimating the effect of land management on bacterial communities, Acosta-Martínez and colleagues (2008) found that the total diversity at different sites changed significantly, but the relative proportion of various phyla and proteobacteria classes remained relatively constant. This phenomenon seems also to influence bacterial communities associated with different part of the human body (Ley, *et al.*, 2005; Ley, *et al.*, 2006; Grice, *et al.*, 2009), suggesting that the structure of bacterial communities may be controlled by undiscovered dynamics among higher taxonomic groups. Furthermore, it was recently shown that the distribution of higher taxa in soils responds to the organic content (Philippot, *et al.*, 2009), suggesting that phyla and classes can actually behave as uniform groups. Improving our capacity to characterise the distribution of higher bacterial taxa can help to reveal the factors underlying this coherence (Philippot, *et al.*, 2010), and the qPCR-DGGE technique presented here was designed for this purpose.

A protocol for SYBR green-based quantification of bacterial phyla and classes exists (Fierer*, et al.*, 2005). However, the technique developed by Fierer and colleagues (2005) relies on the comparison of threshold data (Ct) from 9 separate qPCR assays, which increases the cost as well as the probability of introducing biases. One aim of this study was to develop a single assay to quantify taxa that are dominant in marine ecosystems (Biers*, et al.*, 2009), namely, the Bacteroidetes, Firmicutes, Actinobacteria and the α and μ subdivision of Proteobacteria. A second aim was to develop the first combination of qPCR and DGGE that could allow the timely generation of data on both the diversity and abundance of microbial taxa. This was achieved through a stepwise process of single nucleotide modification of existing and newly-designed primers. The specificity of each primer pair was assessed *in silico* through the Ribosomal Database Project II (Maidak*, et al.*, 2001) and *in vitro* by amplifying a variety of target and non-target DNA templates. The unique qPCR protocol developed was validated on artificial mixtures of 16S rRNA genes similarly to previously described approaches (Becker*, et al.*, 2000; Suzuki*, et al.*, 2000). Following the successful establishment of the qPCR assay, a number of primers were tested for each taxon to allow a nested insertion of the GC-clamp required for DGGE analysis (Muyzer*, et al.*, 1998). Finally, the combination of qPCR and DGGE was used to investigate marine biofilms at the early onset of development on two alternative materials. As young marine biofilms have previously been characterised with a variety of methods, this latest test was included to assess the capacity of the qPCR-DGGE approach to produce comparable results. The advantages and drawbacks of the technique developed here as well as the biological significance of our findings are discussed.

4.2 Materials and methods

4.2.1. Primers design

The overall strategy used to develop the qPCR primers is presented in Figure 4.1. Briefly, in order to reveal phylum- and class-specific conserved regions, between 20 and 30 16S rRNA sequences for each taxonomic group were randomly downloaded from the Ribosomal Database Project II (Maidak*, et al.*, 2001) and grouped into fasta files. Sequences from each taxon were clustered using ClustalX (Thompson*, et al.*, 1997) and consensus sequences obtained using BioEdit (Hall, 1999). The alignment of these consensus sequences (Figure 4.2) was visually inspected to improve existing primers

and to design new ones. Primers were assessed *in silico* using the tool 'probe match' from the RDPII with data set options: strain 'both'; source 'both'; size >1200; quality 'good' (Maidak, *et al.*, 2001), and the result presented in Table 4. The annealing temperature that maximised primer specificity *in vitro* was determined by using 25 singles target and non-target DNA as template in parallel gradient PCR reactions with annealing ranging from 55º C and 68º C. The PCR thermocycle was composed of one initial denaturing step of 5 min at 95º C, thirty cycles of 95º C for 15 sec, gradient annealing for 15 sec and 72º C for 30 sec, and a final elongation step at 72º C for 5 min. Every PCR reaction contained 0.05 u/µl of Taq polymerase (Sigma Aldrich), 0.2 mM of each dNTP, 0.4 µM of each primer, 1X buffer, ~10 ng of DNA and water to 25 µl. Specificity was assessed by standard gel electrophoresis. Primers were chosen and further modified to allow the use of the same annealing temperature for every pair designed. All primers tested are reported in Supplementary Table S1.

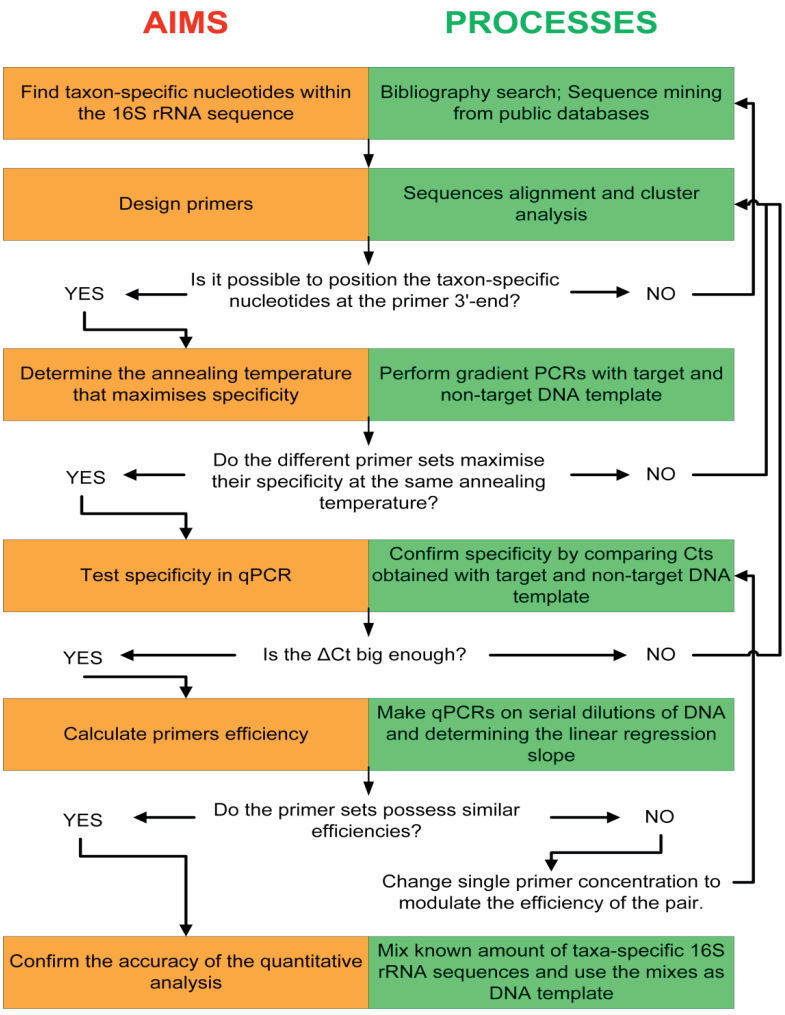
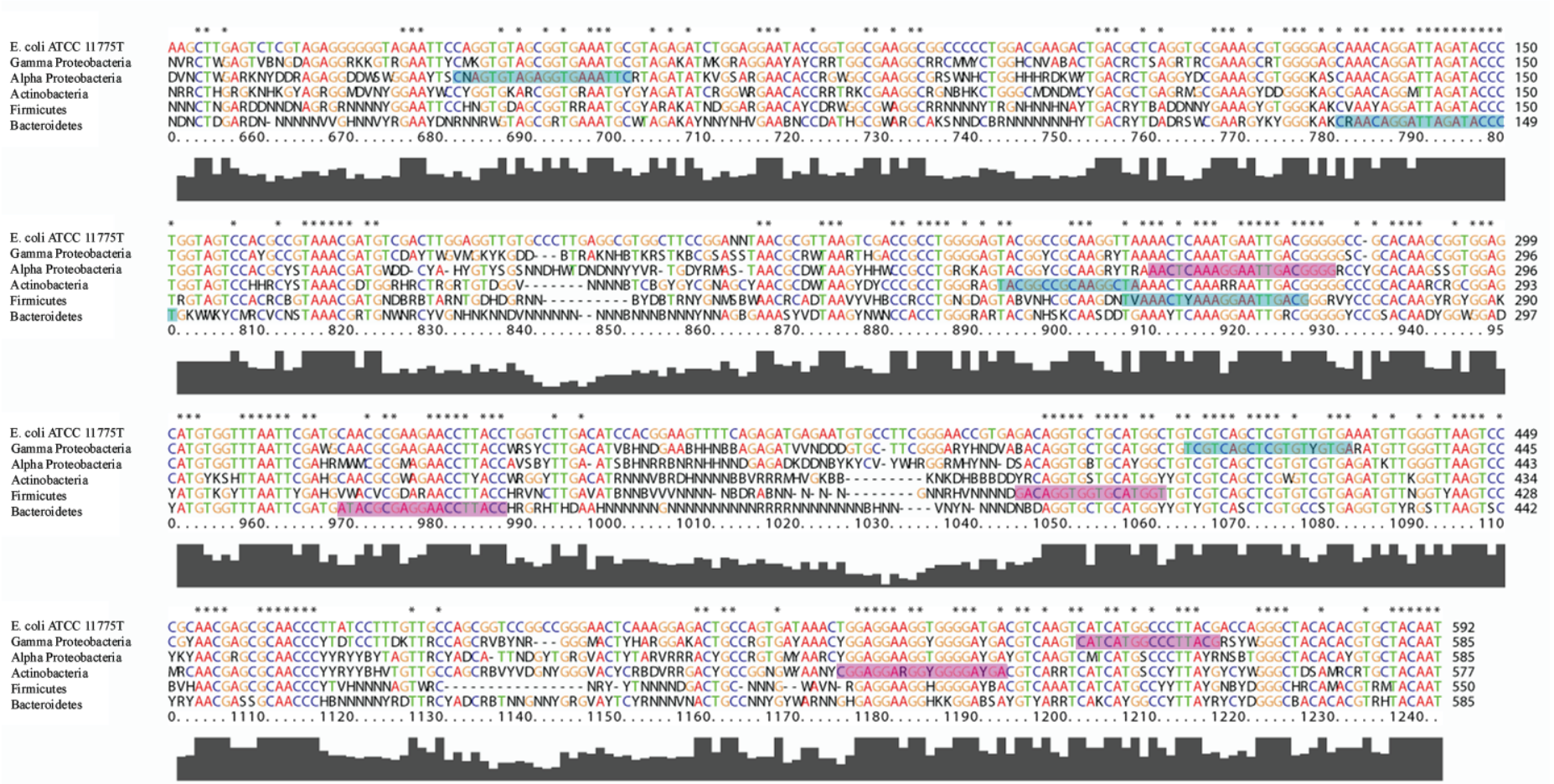


Figure 4.1 - **Flowchart of the process undertaken to develop the primers for the taxon-specific qPCR essay**

Figure 4.2 Alignment of taxon-specific consensus sequences for 16S rRNA



16S rRNA consensus sequences for five different classes/phyla of bacteria were aligned using ClustalX 2.0. The ribosomal sequences of Escherichia coli ATCC 11775T was included as a nucleotide position reference. The primer binding region for the five taxa investigated in this study are highlighted (in blue the forward primer and in pink the reverse primer).

TAXON-SPECIFIC PRIMER PAIRS USED

Target group	Name	Sequence
Universal	926F	AAACTCAAAKGAATTGACGG
	1062R	CTCACRRCACGAGCTGAC
α-Proteobacteria	α682F	CIAGTGTAGAGGTGAAATT
	908αR	CCCCGTCAATTCCTTTGAGTT
γ-Proteobacteria	1080γF	TCGTCAGCTCGTGTYGTGA
	γ1202R	CGTAAGGGCCATGATG
Bacteroidetes	798cfbF	CRAACAGGATTAGATACCCT
	cfb967R	GGTAAGGTTCCTCGCGTAT
Firmicutes	928F-Firm	TGAAACTYAAAGGAATTGACG
	1040FirmR	ACCATGCACCACCTGTC
Actinobacteria	Act920F3	TACGGCCGCAAGGCTA
	Act1200R	TCRTCCCCACCTTCCTCCG

PCR REACTION COMPONENTS

2x SYBR green buffer containing taq polymerase
0.3 µM of each primer
DNA template
Water to 25 µl

PCR THERMOCYCLE

initial denaturing step	95° C	for 5'
30 cycles	95° C	for 15''
	61.5° C	for15''
	72° C	for 20''
final elongation step	72° C	for 5'

Figure 4.3 - **The higher taxon-specific qPCR protocol**

4.2.2. Quantitative PCR optimization of thermocycle and reaction conditions

To investigate the effect that changing the polymerase and reaction buffer had on the optimal annealing temperature, a first round of gradient qPCR (annealing at 58.9 º, 59.3 º, 59.8 º, 60.5 º, 60.9 º, 61.2 º, 61.5 º and 62.5º C) was performed for all taxon-specific pairs. Reactions were performed in sealed 96-well plates using a Chromo4 MJ-Research thermocycler (BioRad) and analysed with the Opticon Monitor 3 software. All qPCR reactions contained 12.5 µl of 2X Faststart SYBR green (Roche Diagnostics Ltd), 10 µl of primers (final concentration 0.3µM), 2.5 µl of the DNA template (equilibrated to 5ng). Primer specificity was inferred from the shift of the threshold cycle (Ct) obtained amplifying target compared to non-target 16S rRNA sequences. It was found

that 61.5º C maximised specificity of all pairs and the ultimate qPCR protocol implemented is reported in Figure 4.3. In order to rule out the formation of PCR artefacts, a melting curve was generated at this point by monitoring SYBR green fluorescence in the temperature ramp 60º to 95º C with an increase of 0.5º C and a hold of 1 s.

4.2.3. Primer efficiencies

The amplification efficiency of each primer pair was determined by making dilution series of target DNA, calculating a linear regression based on the Ct data points, and inferring the efficiency from the slope of the line. Serial dilutions 1, 1:4, 1:16, 1:64 and 1:256 were utilized. Technical triplicates were tested for each dilution point and primer pair. A Non-Template Control (NTC) was included in each assay to confirm that the Ct value generated by the lowest concentrated DNA was not an artefact. Finally, Ct data were uploaded to an Excel spreadsheet and resulting efficiency graphs are given in Supplementary Figure S2.

4.2.4. Quantitative analysis of synthetic communities of 16S rRNA genes

In line with published approaches (Becker*, et al.*, 2000; Suzuki*, et al.*, 2000), the qPCR protocol was validated by analysing artificial mixtures of 16S rRNA genes. Briefly, for each taxon, equal amounts of genomic DNA from 5 bacteria chosen at random from 30 bacterial isolates that were previously genotyped by partial 16S rRNA sequencing (EMBL accession numbers from FN433050 to FN433079) were mixed and amplified with Taq polymerase and universal primers 27F and 1525R. PCR bands were excised from the gel and purified on-column (Qiaprep gel extraction kit, Qiagen), re-suspended in TE buffer (10 mM Tris and 1 mM EDTA, pH 8) and quantified using a NanoDrop ND-1000 UV-Vis spectrophotometer (NanoDrop Technologies). Serial dilutions (3) of equilibrated amounts of purified PCR products were then run on a gel and visualised with ethidium bromide to confirm the quantification.

Different proportions of these PCR products were mixed to obtain the artificial communities of 16S rRNA genes, which were then used as template for qPCR validation experiments. The proportion of taxon-specific qPCR products in each mixture is reported in Figure 4.4. A total of 4 mixes were created and amplified with all primer pairs in triplicate. The average Ct value obtained from each pair was transformed into a percentage with the formula:

$$X = \frac{(\text{Eff. Univ})^{Ct\ univ}}{(\text{Eff. Spec})^{Ct\ spec}} * 100$$

Where the Eff. Univ is the calculated efficiency of the universal primers (2 = 100% and 1 = 0%) and Eff. Spec refers to the efficiency of the taxon-specific primers. The Cts (univ and spec) are the threshold cycles registered by the thermocycler. Resolving this formula, X represents the percentage of 16S taxon-specific copy number existing in a sample. For each artificial community analysed, a chart containing the expected and the calculated percentages of 16S rRNA sequences belonging to each phylogenetic group was drawn.

4.2.5. Preparation of substrates for biofilm establishment

Glass microscope slides were cleaned using 10% nitric acid to produce an oxidised surface. Aminopropyltriethoxysilane (APS) and trimethylchlorosilane (TMS) were used for surface functionalisation following established procedures (Gerhart*, et al.*, 1992). Briefly, APS was diluted with ethanol to a 2% solution. The silanisation reaction was conducted for 30 mins under a fume hood and slides were subsequently rinsed twice in ethanol and once in methanol before baking at 100ºC for 15 minutes. This process yielded surfaces with an amine (NH_2) termination and an advancing water contact angle of 65º ± 5 at the time of the study. For methyl (CH_3) terminated surfaces, slides were incubated with a 1% solution of TMS in dichloromethane (CH_2Cl_2) for 60 minutes. Slides were then rinsed once in CH_2Cl_2 and baked for 30 mins at 100ºC yielding an advancing water contact angle of 71º ± 3. Contact angles were measured from the mean of three 10µl drops of distilled water on each surface type, using a ramé-hart contact angle goniometer.

Silanised slides were inserted in a PVC rack, which kept slides spaced 1.5 cm from each other. The rack was wrapped in a mesh net of average pore size 45 µm to prevent biofilm disruption by larger organisms. The slides were then submerged in a large tank containing free flowing, sand-filtered sea water, as well as a community of macroorganisms representative of the local diversity (experiment performed in Cullercoats, North-East coast of England). After 24 hours, three slides of each silane type were collected, the mesh was briefly rinsed to remove excessive clogging and the rack was repositioned. Biofilms were left to develop for another 3 weeks before a second sampling was performed.

4.2.6. Total DNA extraction from biofilm

Upon collection, slides were rinsed gently with seawater sterilised by 0.22 µm filtration (Millipore Express® PES membrane) to remove loosely attached bacteria. Biofilms were sampled from one whole side of the slides using sterile cotton swabs. The capacity of the swab to remove the biofilm was assessed by staining slides before and after swabbing with a solution 20 µl/ml of acridine orange in seawater for 10 minutes in the dark. After removal of the staining solution with sterile seawater, bacteria were viewed at 100X magnification (Nikon, eclipse 80i) under a fluorescent illumination system (Exfo, X-Cite 120). The area of the swabs with the cells embedded was cut off using a sterile scalpel and placed in a 1.5 ml tube for cell lysis and DNA extraction. Biofilms from three different slides for each surface type were collected after 1 and 21 days of immersion. Total DNA was extracted from all samples using the GeneElute bacterial genomic kit (Sigma-Aldrich), resuspended in 100 µl of TE buffer pH 8 and quantified with NanoDrop.

4.2.7. Analysis of community structure by qPCR

Equilibrated amounts of extracted DNA from marine biofilms were used as template for qPCR with the aim of measuring the proportion of 16S rRNA gene copy number belonging to the taxa Firmicutes, Actinobacteria, Bacteroidetes and the α and γ subdivisions of the Proteobacteria. All qPCR reactions were performed with the protocol described in Figure 4.3. Every DNA template was studied in triplicate with all six primer pairs. qPCRs were analysed with Opticon Monitor 3 software to determine the threshold cycle (Ct) for each run. The average Ct values from the three technical replicates were transformed into proportions of 16S rRNA copies applying the formula described earlier. A chart summarising the results obtained was drawn.

4.2.8. Evaluation of primers for taxon-specific DGGE

A new DGGE approach was developed in order to fingerprint the genotypes amplified by the taxon-specific qPCR. New primers that would allow a downstream insertion of the GC-clamp, required for DGGE analysis (Muyzer*, et al.*, 1993), into the qPCR products were tested. Firstly, the influence of 'carried over' SYBR green on the denaturing property of DNA was assessed by comparing three SYBR green-removal strategies with untreated qPCR and amplicons obtained without the addition of SYBR green. The qPCR products used as template in this test were obtained by amplifying

total bacterial DNA extracted from a marine biofilm unrelated to the samples described earlier using primer 2 and 3 without the GC-clamp (Muyzer*, et al.*, 1993). The three qPCR purification strategies were: 1) GenElute® PCR clean-up kit (Sigma-Aldrich); 2) ethanol precipitation performed with the qPCR products at room temperature; and 3) ethanol precipitation performed with the qPCR products at 95°C. Acrylamide gels (10% w/v) were prepared with a 30-55% denaturing gradients (100% denaturant is 7M urea plus 40% v/v formamide in 1X TAE buffer). The PCR products (15µl) were mixed with 2X loading buffer and loaded in the gels alongside the markers. The loaded material was run at 60º C in a stirred tank (Bio-Rad, DCode™ System) at 200V for 4.5h (900Vh). The gel was stained with SYBR gold (Roche) for 30 min in TAE buffer and viewed under ultraviolet light using a Bio-Rad Fluor-S imager. It appeared that the nucleic acid stain did not influence the formation of DGGE profiles. Therefore, 1µl of untreated qPCR products was used as template for each 60 µl PCR reaction thereafter. DGGE primers were designed internal to the qPCR amplicons whenever possible. A nested approach, and the capacity of 3 to 5 primer sets to separate distinct bands, was tested for each taxon (all primer pairs are reported in Supplementary Table S2). The same GC-clamp 5'-CGCCCGCCGCGCGCGGC-

GGGCGGGGCGGGGGCACGGGGGG-3' (Muyzer*, et al.*, 1998) was used and the optimal annealing temperature and DGGE gradient span were determined for all pairs. The primers and protocol that generated the best taxon-specific profiling of qPCR products (bolded in Supplementary Table S2) were used to investigate the qPCR from the young marine biofilms described earlier.

4.2.9. Sequencing of DGGE band and taxonomic analysis

In order to confirm the specificity of the qPCR-DGGE protocol developed here, a standard protocol was implemented to obtain the DNA sequences corresponding to the DGGE bands. Briefly, bands were excised from the gels with aseptic technique and placed in 50 µl of TAE buffer overnight. An aliquot of the buffer previously heated at 90° C was then used as template for a standard PCR with the DGGE primers but without the GC-clamp. PCR success was checked with gel electrophoresis and products cleaned with EXOsap-IT™ (USB). Sequencing primer and PCR product concentrations were equilibrated as requested by the company (GeneVision, Newcastle upon Tyne, UK). Sequencing was performed with a BigDye™ sequencer (Applied Biosystems). Both the quality-clipping and the subsequent base calling steps on the sequences were

performed using the Phred13 software (Ewing, *et al.*, 1998). Good quality sequences were taxonomically assigned by BLASTing them in the NCBI nucleotides database.

4.3 Results and Discussion

4.3.1. Specificity of newly designed primers targeting high bacterial taxa

The qPCR-DGGE approach presented in this chapter was developed to answer specific biological questions regarding temporal and spatial changes in microbial community structure and specifically in assemblages associated with the barnacle *Balanus amphitrite*. Quantitative PCR and DGGE are widely used techniques for investigating microbial consortia but this is the first study to perform DGGE analysis of qPCR products to gain information on both taxa abundance and species composition. However, this approach requires the sequential use of the two techniques and their results will be discussed separately.

New primers specific to Firmicutes, Actinobacteria, Bacteroidetes and the α and γ subdivisions of the Proteobacteria were developed. The intent was to place the taxon-specific region at the 3'end of the primers and to develop pairs that would amplify between 80 and 200bp. The taxon-specific regions on the 16S rRNA gene exploited to develop the primers are highlighted in Figure 4.2. Out of the 18 taxon-specific primers tested, which can be found in Supplementary Table S1, the 5 primers with the highest specificity were chosen for the qPCR protocol presented in Figure 4.3. Comparison of *in silico* specificity with previously published primers, performed as explained in the Materials and Methods section, is presented in Table 4.

Primer cfb967r have a single nucleotide change from CF967r (Chen, *et al.*, 2006), with which it shares a similarly high specificity. However, the validity of primer CF967r, to our knowledge, has never been tested for qPCR. The primer that has been used in qPCR to determining the abundance of Bacteroidetes is cfb319 (Fierer, *et al.*, 2005; Philippot, *et al.*, 2009). In comparison with cfb319, the primer developed here greatly expands the proportion of Bacteroidetes that can be amplified.

Regarding the Firmicutes, primer 1040firmR showed the greatest specificity of all primers tested here but ~17% of sequences deposited in the RDPII database do not possess the complementary region in this phylum. In particular, 1040firmR matched only 0.5% of Veillonellaceae, 38.8% of Lactobacillaceae, 4.6% of Incertae Sedis XI, 1% of Incertae Sedis XVIII and 35.4% of Erysipelotrichaceae, with all other families

showing >90% coverage. Nevertheless, they demonstrated better *in silico* coverage and specificity than published qPCR primers (Fierer*, et al.*, 2005; Guo*, et al.*, 2008).

For the Actinobacteria, primer act920F3 demonstrated a good specificity (>90%) with no sub-groups showing a coverage of <85%. Primer S-C-Act-235-a-S-20 (also called actino235) was previously validated for qPCR-based investigations but it was shown that only 60% of cloned sequences belonged to the target group (Fierer*, et al.*, 2005). Although primer act920F3 is slightly less specific than actino235, *in silico* analysis suggested that it possesses a much broader coverage of the Actinobacteria. Therefore, the use of act920F3 is recommended.

Finally, the primer targeting the α-proteobacteria demonstrated good specificity and group coverage. Primer α682F is a slight modification of primer Adf681F but in comparison with the latter, it showed a broader range of annealing temperatures at which it was highly specific *in vitro*. The reverse primer Alf685 was previously used for SYBR green-based qPCR assays but clone sequencing showed that it was less specific than expected (Fierer*, et al.*, 2005). The improvement of α682F in combination with 908αR suggested that this new pair should be preferred.

4.3.2. Taxon-specific qPCR protocol

The main scope of this study was to improve and validate a phylum- and class-specific qPCR assay. Compared with previously published taxon-specific quantitative analysis (Fierer*, et al.*, 2005; Philippot*, et al.*, 2009), the advantage of the protocol presented here lies in the ability to use all primer pairs with the same, highly-specific, annealing temperature. At 61.5ºC, all the primers had analogous efficiencies and they demonstrated high correlations, with R^2 values always > 0.99 (Supplementary Figure S2). Thus, the proportion of different taxa in multiple samples in the same PCR run can be investigated, reducing the biases generated by comparing Ct values from different runs. In particular, normalising with a Ct obtained from a different PCR should be avoided (Bustin & Nolan, 2004).

In order to validate the qPCR protocol, artificial mixes of 16S rRNA sequences were generated and tested. The validation departed slightly from the standard approaches in which plasmid constructs are used (Becker*, et al.*, 2000; Suzuki*, et al.*, 2000), as our artificial communities were created by mixing known amounts of taxon-specific PCR amplification of the 16S rRNA gene.

Table 4.1 Comparison of specificity between new and old taxon-specific primers

Target taxon	Primer name	Primer sequence	Distribution of matching sequences across taxa		Target taxon matched	Distance of specific nucleotide/s from the 3'-end	Reference
			Target	Non-target			
Bacteroidetes	cfb967R	GGTAAGGTTCCTCGCGTAT	96.7 %	3.3 %	95.3 %	0 and 1	This study
	CF967	GGTAAGGTTCCTCGCGTA	95.8 %	4.2 %	96.3 %	0	(Chen, *et al.*, 2006)
	cfb319	GTACTGAGACACGGACCA	98.1 %	1.9 %	44.1 %	**	(Manz, *et al.*, 1996)
	MIBF	GGCGACCGGCGCACGGG	99.9 %	0.1 %	45.5 %	5	(Nakanishi, *et al.*, 2006)
	MIBR	GRCCTTCCTCTCAGAACCC	99. %	0 %	55.2 %	**	(Nakanishi, *et al.*, 2006)
	Bact1060R	AGCTGACGACAACCATGCAG	95.2 %	4.8 %	87.3 %	8	(Guo, *et al.*, 2008)
γ-Proteobacteria	γ1202R	CGTAAGGGCCATGATG	85.8 %	14.2 %	87.2 %	**	This study
	Gamma395f	CMATGCCGCGTGTGTGAA	89.5 %	10.5 %	62.1 %	5	(Mühling, *et al.*, 2008)
	Gamma871r	ACTCCCCAGGCGGTCDACTTA	99.1 %	0.9 %	68.1 %	6 and 7	(Mühling, *et al.*, 2008)

(continues)

Table 4.1 (continued)

Target taxon	Primer name	Primer sequence	Distribution of matching sequences across taxa		Target taxon matched	Distance of specific nucleotide/s from the 3'-end	Reference
			Target	Non-target			
α-Proteobacteria	α682F	CIAGTGTAGAGGTGAAATTC	94.3 %	5.7 %	92.7 %	1 and 11	This study
	Alf28F	ARCGAACGCTGGCGGCA	96.2 %	3.8 %	50.5 %	**	(Ashelford, *et al.*, 2002)
	Alf968 (rev)	GGTAAGGTTCTGCGCGTT	84.6 %	15.4 %	80.0 %	6	(Amann & Fuchs, 2008)
	Alf685 (rev)	TCTACGRATTTCACCYCTAC	81.8 %	18.2 %	93.1 %	3 and 14	(Lane, 1991)
	Adf681F	AGTGTAGAGGTGAAATT	94.4 %	5.6 %	94.3 %	0 and 10	(Blackwood, *et al.*, 2005)
Firmicutes	1040firmR	ACCATGCACCACCTGTC	99.4 %	0.6 %	83.2 %	0 and 8	This study
	Firm350f	GGCAGCAGTRGGGAATCTTC	97.3 %	2.7 %	41.3 %	0 and 3	(Mühling, *et al.*, 2008)
	Firm814r	ACACYTAGYACTCATCGTTT	97.5 %	2.5 %	38.7 %	**	(Mühling, *et al.*, 2008)
	Lgc353	GCAGTAGGGAATCTTCCG	99.9 %	0.1 %	25.1 %	2 and 5	(Meier, *et al.*, 1999)
	Firm1060r	AGCTGACGACAACCATGCAC	75.5 %	24.5 %	90.3 %	0	(Guo, *et al.*, 2008)
	Firm934F	GGAGYATGTGGTTTAATTCGAAGCA	75.0 %	25.0 %	75.6 %	**	(Guo, *et al.*, 2008)

(continues)

Table 4.1 (continued)

Target taxon	Primer name	Primer sequence	Distribution of matching sequences across taxa		Target taxon matched	Distance of specific nucleotide/s from the 3'-end	Reference
			Target	Non-target			
	act920F3	TACGGCCGCAAGGCTA	91.8 %	8.2 %	95.3 %	0-4	This study
	Act1159r	TCCGAGTTRACCCCGGC	99.9 %	0.1 %	92.9 %	**	(Blackwood, *et al.*, 2005)
Actinobacteria	S-C-Act-235-a-S-20	CGCGGCCTATCAGCTTGTTG	99.0 %	1.0 %	60 %	10	(Stach, *et al.*, 2003)
	S-C-Act-878-a-A-19	CCGTACTCCCCAGGCGGGG	90.8 %	9.2 %	93.6 %	0-1	(Stach, *et al.*, 2003)
	AB1165r	ACCTTCCTCCGAGTTRAC	99.5 %	0.5 %	92.6 %	**	(Ludemann & Conrad, 2000)
	Act1360r	CTGATCTGCGATTACTAGCGACTCC	99.9 %	0.1 %	82.4 %	3 and 18	(McVeigh, *et al.*, 1996)

The symbol ** indicates that specificity is given by a combination rather than a single nucleotide

When investigating these mixtures by qPCR, the Ct ratio obtained between taxon-specific primers and universal primers (adjusted to the efficiencies) provided an estimate of the amount of 16S rRNA copies belonging to each phylum. By comparing the estimated and the known quantities existing in each mix it was possible to draw Figure 4.4. The results show that this technique successfully captured the proportion of phylum- and class-specific 16S rRNA sequences characterising the samples. In trying to understand the mechanisms that generate non-random distribution of higher bacterial taxa (Philippot, *et al.*, 2010), the qPCR protocol presented here facilitates data acquisition compared to existing approaches. It must be stressed that the qPCR assay provides information on the proportion of 16S rRNA gene copies and that, due to the variability in rRNA operon copy number in different bacteria, this proportion cannot be directly transformed into the number of cells (Smith & Osborn, 2009).

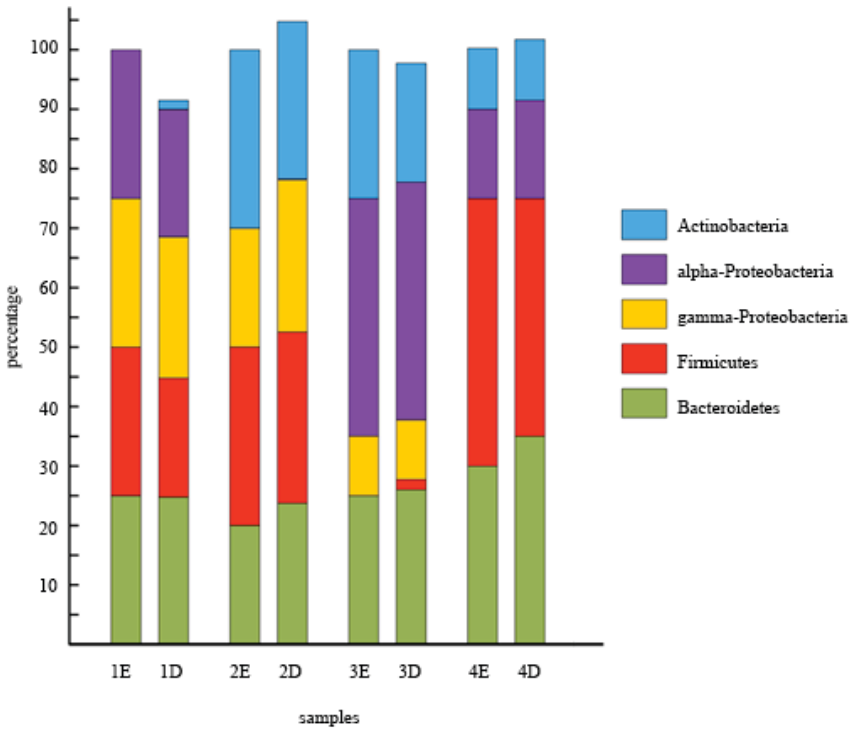


Figure 4.4 - **The use of the qPCR assay to determine the proportion of 16S rRNA sequences in artificial mixes**

Four artificial communities of 16S sequences (samples 1–4) were created with the following taxon-specific concentrations (%); for the Actinobacteria: (1) 0, (2) 30, (3) 25, and (4) 10. For the alpha-proteobacteria: (1) 25, (2) 0, (3) 40, and (4) 15. For the gamma-proteobacteria: (1) 25, (2) 20, (3) 10, and (4) 0. For the Firmicutes: (1) 25, (2) 30, (3) 0, and (4) 45. For the Bacteroidetes: (1) 25, (2) 20, (3) 25, and (4) 30. Samples E show the theoretical proportion of taxa as determined from the above combinations while the bars in samples D show the results obtained by qPCR.

Furthermore, the efficiency of the primers presented here needs to be re-calculated in individual laboratories and ideally in each environment studied. Nevertheless, the simplicity and the low cost of this method make it a suitable option for a wide range of investigations aiming to understand ecological processes affecting broad bacterial taxa.

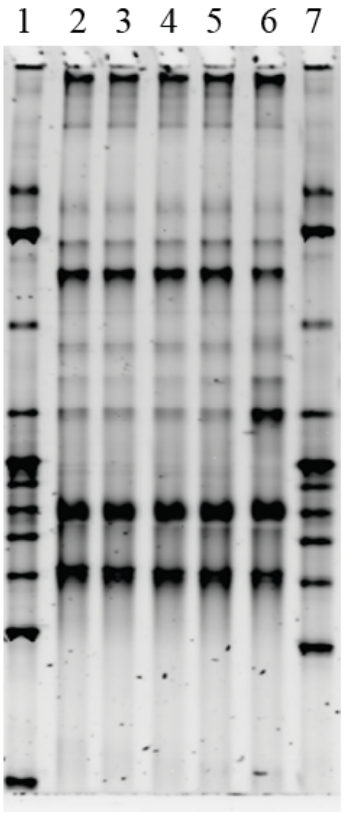


Figure 4.5 - **Comparison of the effect of SYBR green on the denaturing properties of PCR products.**

In the picture, from left to right: 1) marker; 2) spin column; 3) EtOH precipitation performed at room temperature; 4) EtOH precipitation performed at 95 degrees; 5) unpurified qPCR products; 6) regular PCR (without SYBR green addition); 7) marker.

4.3.3. DGGE profiling of qPCR products

Regarding DGGE analysis, a single profile obtained using 'universal' primers is likely to miss minor constituents of the microbial community and it is generally agreed that only phylotypes with abundance >0.1-1% can be revealed using this method (Fromin*, et al.*, 2002). To counteract this effect, it was suggested that fractioning a DNA sample prior to DGGE analysis could enhance the number of discernible bands (Holben*, et al.*, 2004; Dar*, et al.*, 2005). As part of this strategy, the use of phylum- and class-specific primers for DGGE was recently applied to the characterisation of microbial communities (Mühling*, et al.*, 2008). The taxon-specific profiling of qPCR products could represent an important improvement to the technique presented by Mühling and

colleagues, and its development was also considered as a substitute for the cloning and sequencing usually performed to confirm specificity of newly designed primers.

Before proceeding with the DGGE analysis of qPCR products, the influence of SYBR green on the denaturing properties of amplified DNA was investigated. It was found that dye carried over from the qPCR step did not change the DGGE banding patterns (Figure 4.5). Therefore, unpurified qPCR products, which contain SYBR green interlaced in the DNA helix, can be directly used as template for nested amplification and GC-clamp insertion. Once this was confirmed, a series of tests were run to determine the primers that would generate the best DGGE profiles for each taxon investigated. Initially, qPCR primers with the addition of a GC-clamp were tested using the DNA from young marine biofilms as template. This was done for Bacteroidetes and α- and μ-Proteobacteria but not for the two Gram positive taxa as their qPCR primers were developed in a later stage. The DGGE profiles obtained are presented in Figure 4.6. Sequencing of some bands from these gels suggested that the primers did not successfully separate the diverse genotypes amplified, as the sequence chromatograms often contained areas with clean single picks (generally associated with conserved nucleotides) alternate to areas with double picks. Although of poor quality, an attempt to clean these sequences and to determine their taxonomic affiliation was undertaken, and the results suggest that the primers amplified the expected taxa, further confirming the specificity of this approach.

At this point, nested PCR primers were designed for all taxa to target regions internal to the qPCR amplicons. For each group, 3 to 6 primer pairs were evaluated for their ability to generate DGGE profiles. DNA extracted from different environments was also compared to make sure that the improved spreading patterns were not merely the result of using specific DNA. Furthermore, PCR products were run in various acrylamide gels in order to determine the gradient span that would generate the finest band spread. The results obtained with α-Proteobacteria-specific primers are shown in Figures 4.7. For every taxon apart from the Firmicutes, the nested PCR approach generated both neater bands and better band spreads than the primers used in the qPCR step, and it was noted that the proximity of the probes without the GC-clamp to a variable region on the 16S rRNA gene would generally give rise to better DGGE profiles.

In order to design DGGE primers, Muyzer *et al.* (1993) performed time-lapsed gel runs of individually amplified bacterial DNA. The same approach could have been implemented in this study, but, similar to Müling *et al*. (2008), the comparison of DNA

from different microbial assemblages was considered as a sufficient step to demonstrate that the profiles did not arise from primers artefacts. Although certain taxon-specific profiles of different environments possessed re-occurring bands, suggesting that further optimisation of the DGGE primers should be considered, the primers in bold in Supplementary Table S2 were considered the most suitable and implemented for profiling qPCR products.

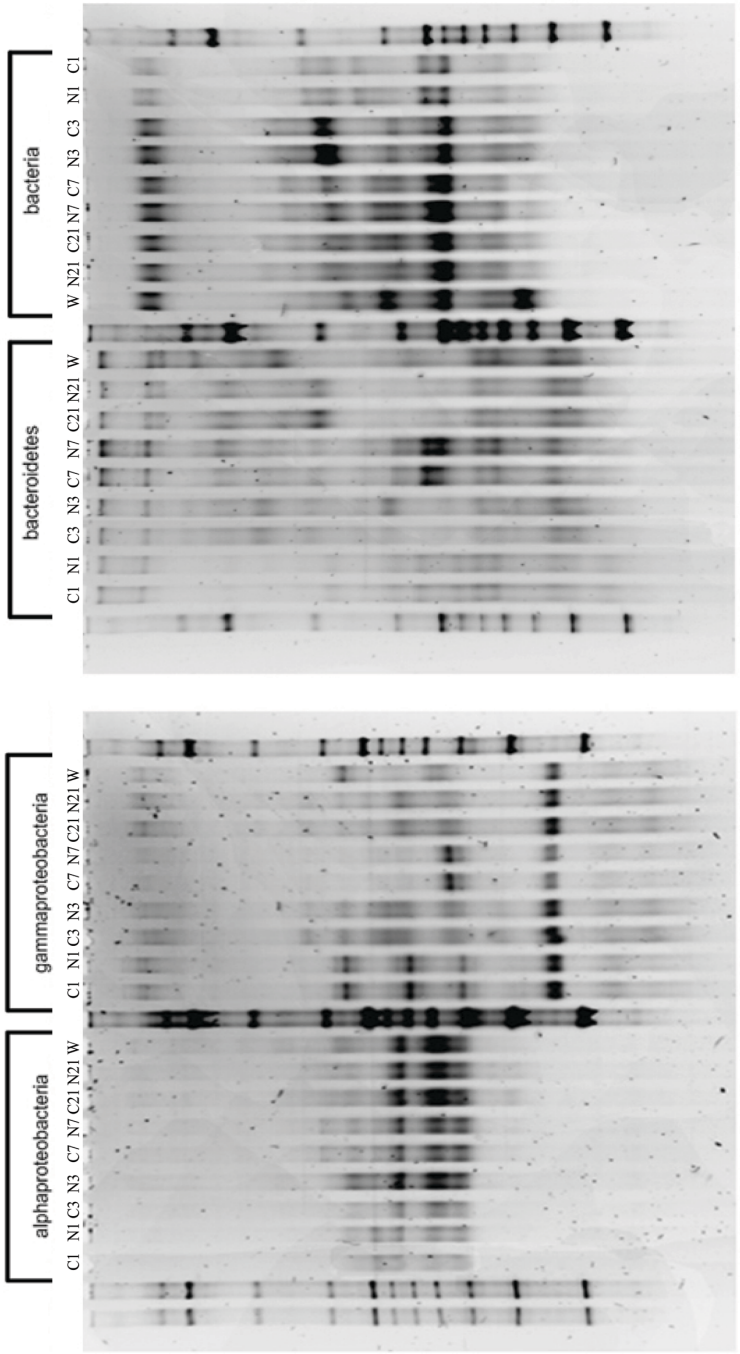


Figure 4.6 - **Initial assessment of DGGE primers.**

Primers tested here were designed by simply adding the GC-clamp (Muyzer et al., 2003) to the taxon specific primers used in the qPCR step.

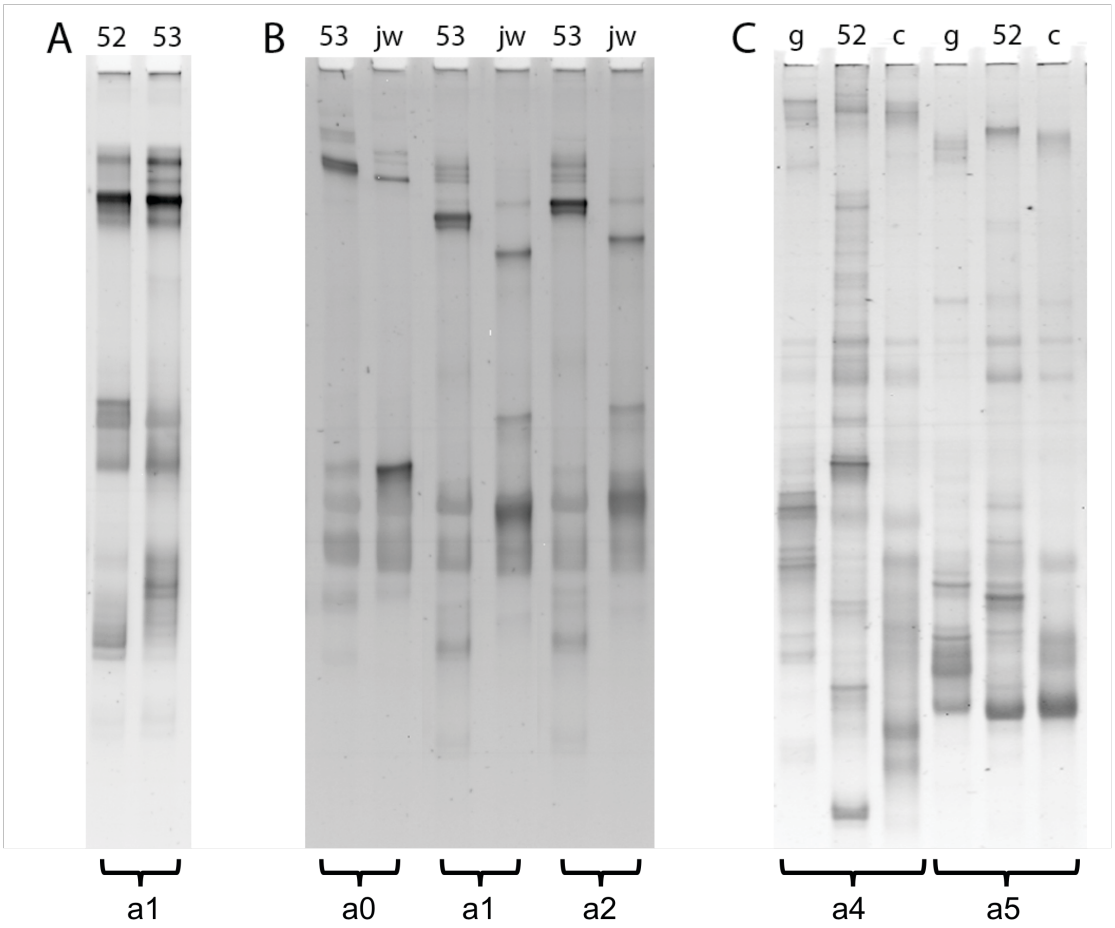


Figure 4.7 - **Improvement of DGGE primers specific to the α-Proteobacteria**

5 primer pairs were tested on α-Proteobacteria qPCR obtained from 5 DNA templates extracted from:

52 = sea water collected at 25m depth

53 = sea water collected at 100m depth

jw = mixed biofilm from fuel cell

g = soil

c = compost

The following primers were used (* = GC-clamp as in Muyzer *et* al, 1993):

a0 = (F) CIAGTGTAGAGGTGAAATT + (R) *CCCCGTCAATTCCTTTGAGTT

a1 = (F) *AGTGTAGAGGTGAAATTC + (R) CCCCGTCAATTCCTTTGAGTT

a2 = (F) *CIAGTGTAGAGGTGAAATTC + (R) CCCCGTCAATTCCTTTGAGTT

a4 = (F) TTAGATACCCTGGTAGTCCA + (R) *CCCCGTCAATTCCTTTGAGTT

a5 = (F) *AGTGTAGAGGTGAAATTC + (R) CTACCAGGGTCTCTAATCCT

While the taxon-specific DGGE profiles of young marine biofilms will be presented below, an overview of the results obtained by sequencing bands from a number of samples is given here to assess specificity as well as possible technical drawbacks.

Regarding the Bacteroidetes and the α-Proteobacteria, most sequences were of good quality and they all matched deposited sequences belonging to the target groups. Therefore, for these two taxa, it was confirmed that the qPCR-DGGE primers are 100% specific and this is a good improvement compared to previously published techniques (Blackwood*, et al.*, 2005), particularly for the α-Proteobacteria. For the μ-Proteobacteria, only 70% of the obtained sequences were affiliated with the corresponding taxon. Their qPCR primers showed the lowest *in silico* specificity (~85%) compared to those targeting the other taxa (Table 4). Nevertheless, the non-μ-Proteobacteria sequences retrieved from DGGE bands all hit bacteria that were predicted targets, mostly within the Sphingobacteria and the ε-Proteobacteria. On the contrary, primers specific to Firmicutes and Actinobacteria produced sequences of very poor quality. The PCR products were analysed by capillarity electrophoresis, which confirmed that single length amplicons were presented. Although PCR worked well for these 2 groups, what caused the low sequencing quality was not identified, which also prevent the final confirmation of specificity. For this reason, despite accurately determining the abundance of Firmicutes and Actinobacteria in artificial 16S rRNA communities by qPCR, caution is suggested when claiming specificity for these primers.

4.3.4. QPCR-DGGE analysis of developing marine biofilms

Microbial assemblages developing on microscope slides immersed in seawater were investigated to evaluate the capacity of the qPCR-DGGE method to resolve complex community structures. Glass slides were treated with either aminopropyltriethoxysilane (APS) or trimethylchlorosilane (TMS) to functionalise the surfaces. The different surface characteristics did not show a major effect on the abundance of the five taxa investigated. However, quantitative similarities clearly clustered samples according to the age of the biofilm (Figure 4.8). The γ-Proteobacteria dominated the communities in young biofilms but represented less than 10% of the community after 3 weeks, whereas the α-Proteobacteria and the Actinobacteria showed the opposite trend, increasing over time to become the most abundant taxa. The proportion of Bacteroidetes was stable over time, while Firmicutes, although present in young biofilms, were completely absent after three weeks. The relative proportions of the five taxonomic groups on independent slides followed reproducible temporal dynamics.

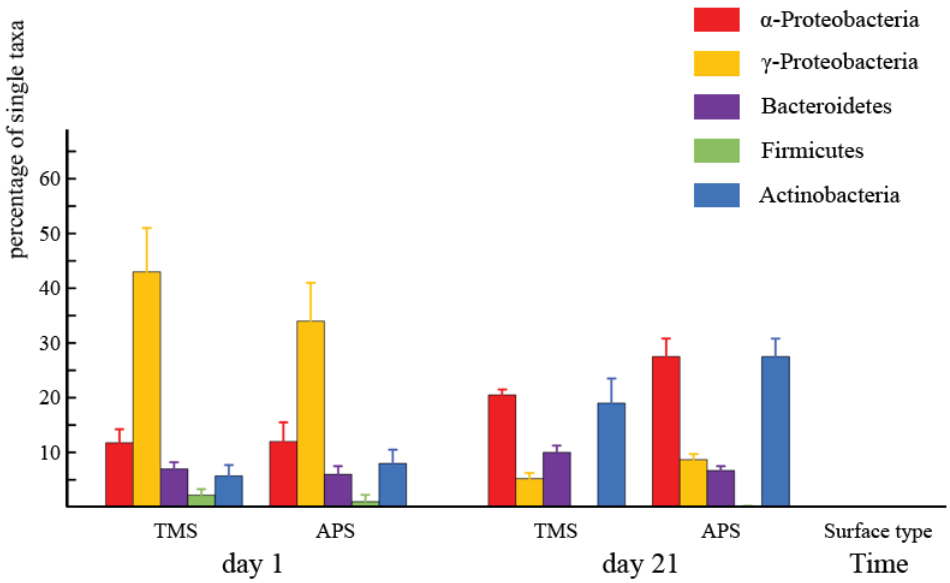


Figure 4.8 - **Taxon-specific quantitative analysis of biofilms developing on submerged glass slides**

Relative quantities of 16S rRNA copy number belonging to higher bacterial taxa are shown for each biofilm community as the mean of three replicates. Surface type TMS is methyl-terminated while type APS is amine-terminated glass.

Figure 4.9 shows the taxon-specific DGGE profiles. Multiple bands are present in every lane, confirming the success of the qPCR protocol in amplifying a variety of genotypes. Each primer pair had a specific gradient span that produced the best band spread (Supplementary Table S2). The diversity within the 5 taxa was comparable, with only the Bacteroidetes showing a relatively larger number of bands. The sum of all taxon-specific bands obtained from every sample was between 27 and 36. Community profiles of the biofilms from the two surface types were almost identical within time points but changed over time, with the only exception being the profiles of Actinobacteria at day 1, where 5 out of 10 bands present in total were only found from one surface type. The species abundance within α-Proteobacteria and Bacteroidetes appeared to increase over time, while proportions of Actinobacteria and γ-Proteobacteria remained relatively constant. However, within each taxon investigated, few genotypes appeared to be present at both day 1 and day 21. Finally, band intensities in DGGE profiles of the same taxon and at the same sampling time were comparable for α- and γ-Proteobacteria but changed for the Bacteroidetes and the Actinobacteria. The Firmicutes-specific profiles were characterised by a low number of bands. Control experiments with soil sample indicated the validity of the primers (Figure 4.9).

Similar changes in the proportions of α -Proteobacteria, Firmicutes, Actinobacteria, Bacteroidetes and γ -Proteobacteria occurred in independent biofilms. Different investigations have shown that members of α -Proteobacteria often comprise the most abundant group in marine biofilms (Jones*, et al.*, 2007; Huggett*, et al.*, 2009). The data show that the proportion of this taxon increased towards later stages of biofilm formation (Figure 4.9). This was also the case for the Actinobacteria, while the abundance of Bacteroidetes remained relatively constant. Members of the γ -Proteobacteria dominated the primary communities, and previous observations (Dang & Lovell, 2000; Jin-Woo*, et al.*, 2008) support these findings. Members of the Firmicutes were barely present in 1-day-old biofilms and disappeared after 3 weeks. Although a qPCR approach targeting higher taxonomic groups was previously used to demonstrate spatial patterning in soils (Fierer*, et al.*, 2005; Philippot*, et al.*, 2009), this is the first study to use this technique to monitor succession in bacterial communities. The reproducibility of community shifts at high taxa levels was striking.

Variations in phylum dominance may simply reflect patterns at the species level (Philippot*, et al.*, 2010). The majority of taxon-specific DGGE profiles characterised here support this hypothesis, with time-specific abundances that could be the result of reproducible shifts in species composition (Figure 4.9). However, Actinobacteria and Bacteroidetes profiles at day 1 indicated that phylum evenness was also conserved on different surfaces independent of species richness. Particularly for the Actinobacteria in young biofilms, diverse phylotype compositions resulted in similar phylum abundances on independent slides. This effect was also found in the Bacteroidetes, although here, this was due to different amounts of the same species within this taxon. However, this observation was made from relative DGGE band intensities and caution should be taken when interpreting this result (Fromin*, et al.*, 2002). This shows that even though there can be significant changes in species diversity, they may occur in the presence of stable phylum level abundances, confirming the importance of studying ecological dynamics at high taxonomic levels (Philippot*, et al.*, 2010).

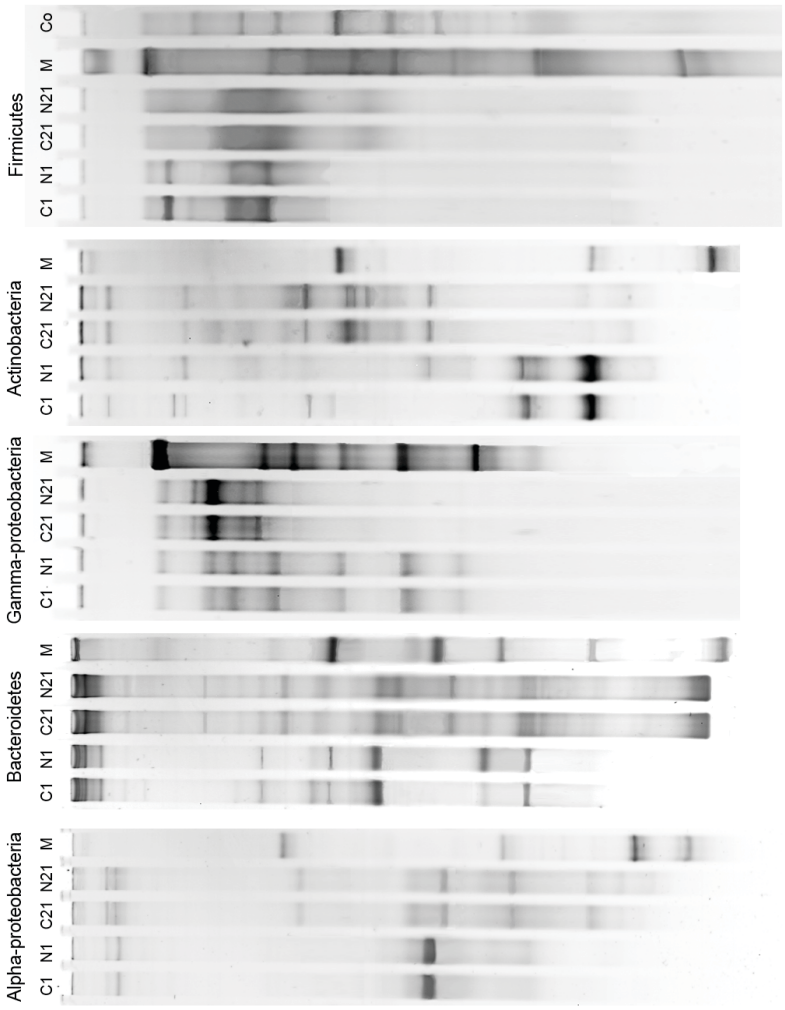


Figure 4.9 - **Taxon-specific DGGE profiles of developing biofilms**

Taxon-specific qPCR products were re-amplified with primers containing a GC-clamp, and the resulting amplicons analysed by DGGE. Lanes prefixed with C were extracted from communities attached to methyl-terminated glass slides, while samples prefixed with N were from amino-terminated surfaces. The numbers 1 and 21 refer to the collection day after slide immersion. Lanes labelled M are reference ladders and lane Co is a Firmicutes-specific profile obtained using soil DNA.

The study of initial biofilm formation in the marine environment has received particular attention because of its economic consequences (Qian*, et al.*, 2007) and interactions between the inert material and the cell surface are thought to play a fundamental role in bacterial adhesion (Neu, 1996). It was of interest to test whether higher taxonomic groups would demonstrate different adhesive strategies. For this reason, two widely implemented silanes were used (Gerhart*, et al.*, 1992; Huggett*, et al.*, 2009) in order to render glass slides with different values of hydrophobicity. However,

the results suggest that the surface treatment generally did not influence initial adhesion or development of bacteria in a phylum-specific manner. Only members of the Actinobacteria showed initial surface-specific differences at the species level, with 2 species present exclusively on APS and 3 occurring only on TMS-treated glass (Figure 4.9). As the biofilms developed independently, a new community appeared. Interestingly, the species composition of these two older biofilms was very similar, and this was also observed for all other high taxa investigated. Further applications of the technique presented here will contribute to our understanding of the processes that regulate succession.

4.4 Conclusions

Our level of understanding of the ecological mechanisms that regulate the formation of communities is based on the development of methods to capture species distribution. While for macro-ecologists counts of organisms are made by eye in sample areas, microbiologists need molecular tools to support their analysis. To this end, a new taxon-specific, SYBR green-based, quantitative PCR approach to characterise bacterial communities has been designed. The new primers generated were tested against both synthetic and natural microbial assemblages, confirming the validity of this new technique. Furthermore, an attempt to apply DGGE-sorting to qPCR products has been undertaken. Although sequencing of DGGE bands suggested that some taxon-specific primers require further optimisation, initial results indicated that the combination of these two techniques may provide rapid and relatively inexpensive data. By monitoring microbial communities developing in marine biofilms, an example of the ability of this method to help in elucidating the evolution of ecological traits is provided.

Chapter 5. Bacteria associated with the acorn barnacle Balanus amphitrite

5.1 Introduction

The vast majority of people perceive bacteria uniquely as etiological agents of diseases that menace human society and their surroundings (Postgate, 2000). Although this view is true for a few aggressive pathogens, the number of known beneficial actions performed by microbes is growing fast. For example, bacteria play a key role in the cycling of nutrients and a large proportion of the global oxygen production is performed by photosynthetic microorganisms living in the oceans (Field*, et al.*, 1998). Furthermore, new discoveries in the field of microbiology are often translated into biotechnological applications that can be exploited, including the production of vitamins or the removal of toxins from contaminated environments. However, in relation to the scope of this thesis, among the many new roles that have recently been attributed to bacteria, their importance in determining the evolution and development of eukaryotes is one of the most fascinating (McFall-Ngai, 1999; Rosenberg*, et al.*, 2007; Gilbert*, et al.*, 2010). The beneficial effects of commensal microbes inhabiting the gastro-intestinal tract of their host have been known for decades, but their influence on the mating behaviour of the host has only recently emerged (Sharon*, et al.*, 2010). Nevertheless, many of the methods used to investigate microbial metabolic diversity have only been available to researchers for the past two decades (Torsvik*, et al.*, 1990; Handelsman, 2004), and our understanding of the importance of bacteria in sustaining life as we know it may well be in its infancy.

Focusing on the marine environment, bacteria have been shown to play important roles in the life cycle of different invertebrates. One of the most famous is the effect of *Vibrio fischeri* in triggering the morphogenesis of the light organs in the sepiolid squid *Euprymna scolopes* (McFall-Ngai & Ruby, 1991; Montgomery & McFall-Ngai, 1994; Nyholm*, et al.*, 2000), which in turn protect the host from predators. Also, bacteria have been shown to be able to induce metamorphosis of various species, and the effect of *Alteromonas espejiana* on the hydroid *Hydractinia echinata* is a classic example (Leitz & Wagner, 1993). Furthermore, invertebrate-associated bacteria can provide protection against pathogens (Gil-Turnes*, et al.*, 1989; Gil-Turnes & Fenical, 1992) or defence against harmful fouling species (Armstrong*, et al.*, 2001), and growing evidence

suggests that bacteria and their hosts can sometime be considered as an unique evolutionary unit (Rosenberg*, et al.*, 2007; Gilbert*, et al.*, 2010). For this reason, the study of bacteria-invertebrate associations can answer fundamental biological questions on the evolution and development of marine species and on the mechanisms that regulate the pelago-benthic transition in larvae.

In biofouling research, particular focus has been given to the effect of microbially-produced compounds on invertebrate settlement (Dobretsov*, et al.*, 2006). It was shown that bacteria induce permanent attachment in a variety of invertebrate taxa, including the alga *Ulva* sp. (Joint*, et al.*, 2000), the coral *Acropora microphthalma* (Webster*, et al.*, 2004) and the polychaete *Hydroides elegans* (Lau*, et al.*, 2003). However, regarding *Balanus amphitrite* (syn. *Amphibalanus amphitrite*; Clare & Høeg, 2008) the literature looks more puzzling, as microbial biofilms stimulate contrasting settlement behaviours (Maki*, et al.*, 1992; O'Connor & Richardson, 1998; Khandeparker*, et al.*, 2006). In this species, a developmental stage – the cyprid – has evolved as a transitional phase between the pelagic larva and the benthic adult (Walker*, et al.*, 1987; Anderson, 1994; Glenner, 1999). Cyprids that are developmentally competent to settle display an adaptive ability to delay permanent attachment and discriminate between potential settlement sites. Although it was shown that the pelago-benthic transition is influenced by multiple factors (Rittschof*, et al.*, 1984; Maki*, et al.*, 1988; Mullineaux & Butman, 1991), cross-fertilisation is required for *B. amphitrite* to reproduce and it is clear that the fundamental driver of permanent attachment is the need for a cyprid to settle next to a mating partner (Crisp & Knight-Jones, 1953).

It was shown that kin-recognition in *B. amphitrite* is regulated at a molecular level (Crisp & Knight-Jones, 1953), and the settlement-inducing protein complex (SIPC) was identified as the molecule responsible and characterised in this species (Dreanno*, et al.*, 2006b). The increased effect of SIPC when bound to a surface and the fact that its expression is localised in the cuticle (Dreanno*, et al.*, 2006a) suggests that the SIPC acts as a contact pheromone. However, barnacle shells in their natural environment are covered by biofilm; a thick layer composed of microbial cells embedded in a sheet of polymeric substances (Branda*, et al.*, 2005). For this reason, the role of bacteria and their exudates in acting as intermediaries for barnacle kin recognition deserves further attention. Furthermore, it can be expected that the existence of bacteria promoting barnacle settlement would favour the adults they associate with, while the growth of cyprid-repelling bacteria around adult shells would not be favoured in an evolutionary sense. However, few attempts have been made to characterise bacteria associated

naturally with barnacles and to test their effect on larval behaviour (Khandeparker*, et al.*, 2002). A systematic survey of bacterial communities naturally associated with *B. amphitrite* has yet to be undertaken.

A common first step for studies on inter-kingdom interactions is that of characterising the microbial community associated with the organism of interest and showing the presence of reproducible patterns that can be then studied in the lab. The aim of the set of experiments presented in this chapter was to investigate whether some bacterial taxa are specifically associated with the acorn barnacle *Balanus amphitrite*, as the existence of such partnerships has not been previously described. Initially, the bacterial communities inhabiting the external shell of *B. amphitrite* as opposed to those colonising the surrounding rocks were compared. The effect of barnacle-associated bacteria on the behaviour of conspecific larvae was also assessed. Subsequently, the microbes associated with different developmental stages of *B. amphitrite* were analysed to assess whether reproducible communities existed and whether beneficial bacteria were transmitted vertically. The results suggest that specific bacterial assemblages colonised different developmental stages and *B. amphitrite* tissues, indicating that bacteria may have roles in barnacle development yet to be established.

5.2 Materials and methods

5.2.1. Biofilm sampling and DNA extraction

Samples were collected in the period between January and February 2010 in the intertidal area of Dona Paula Bay, Goa, India (15°27.5'N, 73°48'E). Bacteria growing attached to 4 different *B. amphitrite* shells were sampled as well as those adherent to the rock, both 2 and 4cm away from the basal plate of each barnacle shell sampled. Prior to collecting the bacterial communities, the sampling areas were washed with copious amount of sterile sea water to remove loosely attached organisms. Biofilms were collected from all surfaces using a sterile cotton swab with a method previously validated (chapter 4). The region embedding the microbial cells was sliced off the swabs with a sterile scalpel and total DNA extracted on column (GenElute, Sigma-Aldrich) directly from the cotton. DNA was precipitated and kept in ethanol until used.

5.2.2. Taxon-specific qPCR-DGGE

Microbial assemblages were characterised with the technique explained in chapter 4. Briefly, equilibrated amounts of total DNA were amplified in quantitative PCR using 16S rRNA-targeting primers specific to Firmicutes, Actinobacteria, Bacteroidetes and α-Proteobacteria. Universal primers were used as a reference to obtain the total amount of 16S ribosomal genes present in the extracted DNA. A delta-Ct method was used to determine the proportion of each taxa in the various samples. Quantitative sample similarities were investigated with the Kulczynski index using the relative proportion of higher bacterial taxa as variables (Clarke & Warwick, 2001). A non-metric multi-dimensional scaling (MDS) analysis was performed on the resemblance matrix obtained and 1000 restarts were carried out to draw a 2D scatter plot with Primer 6 (Clarke, 1993). For each bacterial group investigated, a one-way ANOVA was performed to assess the significance of the difference found between shell- and rock-associated communities.

Quantitative PCR products were re-amplified using a nested-PCR approach in order to insert the GC-clamp required for DGGE analysis (insert reference EM). GC-clamp-containing PCR products were run into acrylamide gels (10% w/v of acrylamide/bisacrylamide 37.5:1 in 1X TAE buffer) with various denaturing gradients: 35-55% for the Firmicutes, 25-45% for the Bacteroidetes, 25-50% for the Actinobacteria and 25-55% for the α -Proteobacteria. The 100% denaturing solution comprised 7M urea plus 40% v/v formamide in 1X TAE buffer. Gels were polymerised by adding 1:1000 vol/vol of TEMED and 1:100 vol/vol of ammonium persulphate to the solution. PCR products were run in DGGE tanks (Bio-Rad, DCode™ System) containing TAE buffer at 60°C for 4.5h at 200V (900V/h). At the end of the run, bands were stained for 30 min with 1X SYBR green, washed 5 min in water and visualised under UV light (Bio-Rad Fluor-S imager). A binary matrix (1=present and 0=absent) was determined from DGGE profiles by visual inspection of digitised gel images. The similarity matrix for the profiles was evaluated by determining Jaccard distances and clustering samples using the unweighted pair group method with arithmetic averages (UPGMA).

5.2.3. Balanus amphitrite culture

The barnacle breeding stock used in this study was obtained by collecting adults from the Duke Marine Laboratory, Beaufort, North Carolina, USA (34°40'N, 76°40'W).

Adults were maintained in the aquarium in Newcastle following established procedures (Bacchetti De Gregoris*, et al.*, 2009). Hatched larvae were released by the adults and attracted to a point light source for collection. Nauplii were cultured at a density of ~1 larva ml^{-1} in an incubator at 28°C on a 12:12 light:dark cycle. Larvae were provided with 1000ml a day of a *Skeletonema costatum* culture (~2×10^5 cells/ml) until they reached the cyprid stage (approx. for 4–5 days). Cyprids were collected by filtering through a tier of filters (pore sizes of 350 and 250 μm) in order to discard undeveloped cyprids and microalgae, and stored in filtered sea-water at 6°C until use. At this point, cyprids were ready to be used for settlement assays.

5.2.4. Media preparation, bacterial isolation and settlement assay

The effect of barnacle-associated bacteria on settlement of conspecific larvae was tested. To do this, a microbiological medium based on barnacle adult extract (AE medium) was prepared to increase the likelihood of growing naturally-associated bacteria in vitro. Basically, a mash of barnacles was obtained by grinding adults with a mortar and pestle. The resulting puree was boiled in seawater for 30 min (~100 individuals per 1 litre of sea-water) and then cooled down and decanted. The solution was centrifuged at 1,200g for 10 min to remove the particulates before sterilisation by autoclaving. The final protein content of the AE medium was, as determined by spectrophotometry with a previously described method (Lowry*, et al.*, 1951), ~170μg/ml. The capacity of two different microbial communities to induce cyprid settlement was investigated: barnacle-associated bacteria (BAB) and biofilms attached to rock (RB). For samples BAB, a living *B. amphitrite* adult was collected from our breeding stock, rinsed with copious amounts of sterile seawater and inoculated in 20ml of AE medium. For RB assemblages, inert material (a stone of roughly the same size as barnacles) was taken from another tank that contained mussels instead of barnacles. Inoculated AE media were shaken at 90rpm for 24h at 20°C to allow growth. Bacterial cells were then collected by centrifugation at 12,000g for 5 min and washed twice with sterile seawater and re-centrifuged. The microbial mixtures were finally re-suspended in seawater and equilibrated at the same concentration (cell/ml) by checking their turbidity at an optical density of 600nm. Aliquots of 2ml were used to inoculate 12 wells of a 24-well microtiter plate. The biofilm was allowed to develop for 20h before discarding the culture medium and rinsing the wells 3 times with 2ml of sterile seawater. At this point, ~15 cyprids (in 2ml) were added to each well and the plates incubated at 28°C. Settlement was checked after 24 and 72h. Also, biofilm growth was determined visually

by growing them on parallel Petri dishes treated as described above. The quantification of attached bacteria was done by staining cells with 20µl/ml of acridine orange and visualised under a 100X oil microscope (Nikon, eclipse 80i) equipped with a fluorescent illumination system (Exfo, X-Cite 120). The whole experiment was repeated 3 times to assess the reproducibility of the noted effect on settlement. Changes in settlement behaviour of larvae resulting from the treatments were normalised against the control and their significance was statistically assessed by ANOVA.

5.2.5. DGGE of settlement-inducing biofilms

After rinsing the wells and before adding the cyprids for the settlement assay, two replicates of biofilms formed by BAB and RB were collected. Bacterial cells were removed from the bottom of the wells using sterile cotton swabs and total DNA extracted as described above. Partial 16S ribosomal genes were amplified with universal primers 2 and 3 (Muyzer*, et al.*, 1998) with 35 PCR cycles of 95°C for 15 sec, 55°C for 15 sec and 72°C for 30 sec. A final elongation step of 20 min at 72°C was included. Amplification was performed using 0.025U/µl of Taq polymerase in a mixture containing 0.3µM of each primer, 0.2mM of each dNTP and ca. 10ng of template DNA. Successful amplification was confirmed by gel electrophoresis. PCR products were run in a 30-55% denaturing gradient gel as described earlier and bands were excised from the gel with aseptic technique and kept in 50µl of TE buffer at -20°C until used. For sequencing, bands were heated at 90°C for 5 min before a 5µl aliquot was taken and used as template for a new PCR reaction of a total volume of 50µl containing primers 2 and 3 without the GC clamp. Successful PCR products were purified on-column and sequenced with a BigDye™ sequencer (Applied Biosystems). Both the quality-clipping and the subsequent base calling steps on the sequences were performed using the Phred13 software (Ewing*, et al.*, 1998), resulting in a good quality, average read length of 161 nucleotides (trimmed of primers). A taxonomical representation of the BAB and RB bacteria was obtained by blasting the sequences in the NCBI nucleotides collection (Altschul*, et al.*, 1990).

5.2.6. Quantitative analysis of bacteria associated with laboratory-raised barnacles

Barnacles from Beaufort were cultured according to established procedures (Bacchetti De Gregoris*, et al.*, 2009). Total bacterial DNA was extracted with the GenElute bacterial genomic kit from the following developmental stages and body parts

of laboratory raised *B. amphitrite*: (a) just released nauplii; (b) cypris larvae; (c) whole adult body; (d) outer shell; (e) exuviae. For samples (a) and (b), larvae were collected and washed with sterile sea water through a 70μm mesh net and around 100 nauplii and 20 cyprids were respectively used for the DNA extractions. Adult bodies were removed from the shell and homogenised in a 1.5 ml tube with a sterile pestle. Triplicate DNA extractions were performed from 3 individuals. DNA from the outer shell was recovered with the swab method as described above. For sample (e), floating exuviae were collected from the tanks containing barnacle cultures and DNA extracted in triplicate directly from the cuticles after rinsing with sterile water. The bacterial community present in the water was used as control and DNA was extracted from cells collected by centrifuging 5 ml of water from the tanks at 12,000g for 5 min. All DNA samples were investigated with the qPCR method described earlier in order to determine the proportion of 16S rRNA copy numbers belonging to higher taxonomic groups.

5.2.7. Quantitative analysis of bacteria naturally associated with different barnacle developmental stages and tissues

For this last set of experiments, *B. amphitrite* specimens were collected from the intertidal zone around the Duke Marine Laboratory. Adult individuals were brought to the lab and processed under a dissecting microscope. To determine whether the bacteria associated with barnacles were vertically transmitted, four sample types were analysed: 1) testis; 2) eggs; 3) larvae and; 4) adult shell. For sample 1, seminal vesicles (testis) were dissected from the animal and rinsed 3 times with sterile sea-water. For sample 2, egg masses were removed from the adults bearing them, the egg sacs were opened in sterile water and further rinsed gently (by immersion) with sterile water. The stage of development of the eggs was assessed by visual inspection and egg's DNA classified accordingly (early, middle and late state, Figure 5.1). For sample 3, eggs from freshly-collected adults were allowed to spawn in sea-water and the larvae collected by attracting them to a point light source. Contaminant bacteria were eliminated by filtering the larvae out of the water and rinsing them further with sterile seawater. Bacteria growing on the outermost part of the adults' calcareous shells were removed with sterile swabs as previously described. The DNA was then extracted directly from all these samples with the GenElute kit (Sigma-Aldrich), as the tissues completely dissolved in the extraction buffer and no further disruption was required. In contrast to the DNAs extracted from laboratory-raised barnacles, qPCR for these samples was

performed in a Mastercycler ep realplex real-time PCR thermocycler (Eppendorf) using the SYBR Green JumpStart Taq 2X mastermix (Sigma-Aldrich).

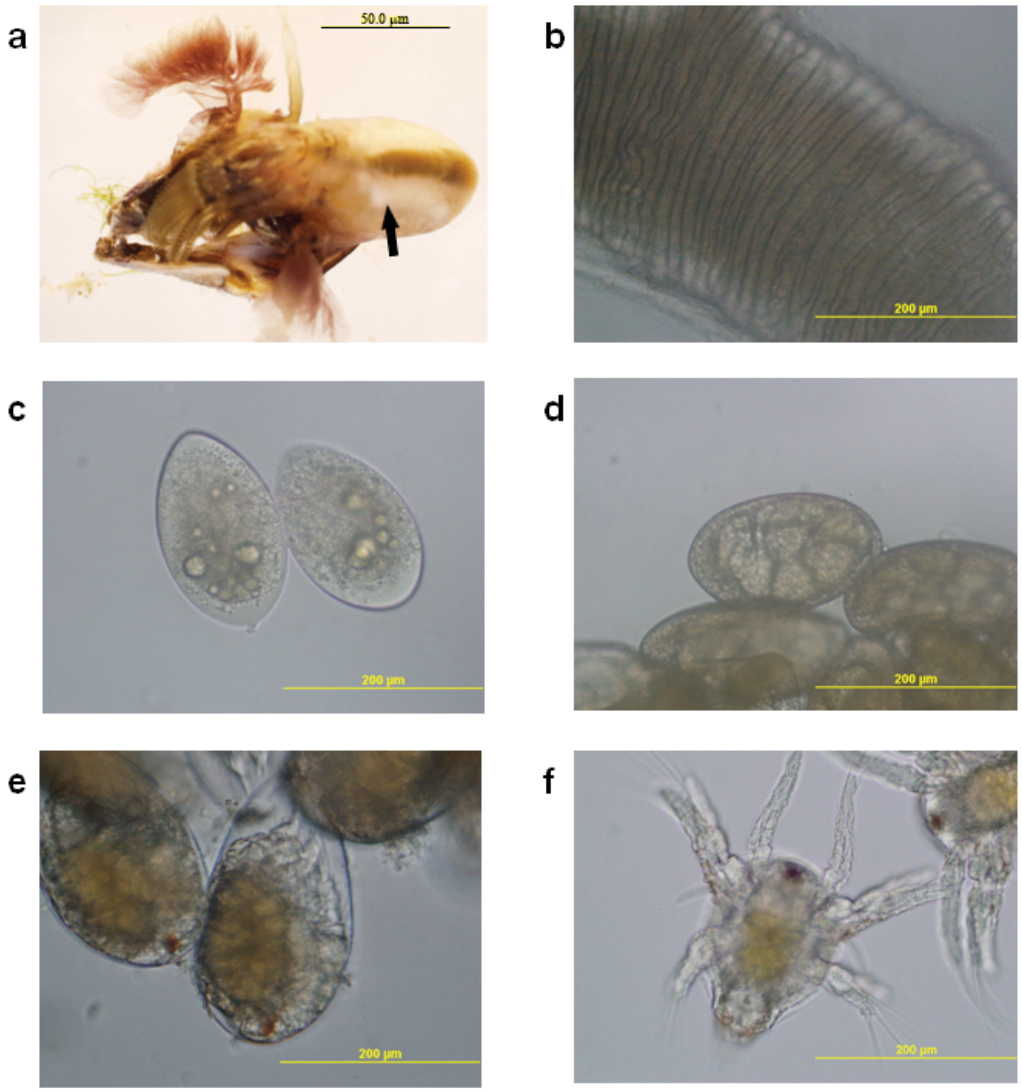


Figure 5.1 - **Extraction of 16S rDNA from barnacle specimens**

DNA from different *B. amphitrite* developmental stages was extracted. a) adult barnacle excised from the shell with arrow pointing at the seminal vesicle; b) seminal vesicle excised from adult body; c) eggs in early stage of development; d) eggs in middle stage of development; e) eggs in late stage of development; f) hatched nauplii.

All DNAs were analysed with the taxon-specific qPCR method. The quantitative sample similarities obtained were investigated with the Kulczynski index (Clarke & Warwick, 2001) using the relative proportion of Actinobacteria, Firmicutes, α - Proteobacteria and Bacteroidetes as variables. A non-metric multi-dimensional scaling (MDS) analysis was performed on the resemblance matrix obtained and 500 restarts were carried out to draw a 2D scatter plot (Clarke, 1993).

5.3 Results and Discussion

Stable associations between bacteria and animals are widely distributed in both marine and terrestrial ecosystems. Many of these partnerships involve mutualistic symbiosis between a single bacterial strain and its animal host (McFall-Ngai & Ruby, 1991; Polz*, et al.*, 1994). However, conserved associations with a more diverse microbial assemblage have also been reported, and bacteria living on sponges and corals are typical examples (Hentschel*, et al.*, 2002; Bourne*, et al.*, 2008). Furthermore, by studying coral microbiology, Rosenberg and colleagues came to realise that a dynamic relationship is established between microorganisms and corals, and that the best suited holobiont (the host species and its associated symbiotic microbiota) is selected according to the prevailing environmental conditions (Rosenberg*, et al.*, 2007). This led them to propose the 'hologenome theory of evolution' (Rosenberg*, et al.*, 2009), which posits that the holobiont is a unit of selection in evolution, and they elegantly demonstrated this theory by showing that gut microbiota were capable of influencing mating preference in *Drosophila melanogaster* (Sharon*, et al.*, 2010). As barnacles essentially decide their mating preference at settlement, the discovery of settlement-inducing bacteria stably associated with barnacle species could represent an interesting confirmation of Rosenberg's vision. The experiments presented here were designed to explore this possibility.

5.3.1. Quantitative fingerprinting of rock- vs. shell-associated biofilms

Regarding *B. amphitrite*, a number of studies have investigated the effect of bacteria on settlement, testing either natural biofilms (Qian*, et al.*, 2003; Hung*, et al.*, 2007) or isolated strains (Khandeparker*, et al.*, 2006; Zardus*, et al.*, 2008). However, this is the first attempt characterise bacterial communities associated with this barnacle. To test the hypothesis that barnacles harbour microbes that differ from those growing on adjacent rocks, a qPCR-DGGE method was deployed. The proportion of 16S rRNA copies belonging to the taxa Bacteroidetes, Actinobacteria, Firmicutes and α-Proteobacteria was determined. The non-metric multi-dimensional scaling of quantitative data showed a stronger clustering effect between biofilms collected from barnacle than from the surrounding rocks, suggesting that bacterial communities

attached to *B. amphitrite* shells were similar from the perspective of proportion of higher taxa (Figure 5.2A). Some rock biofilms clustered together with those from barnacles, particularly the sample collected around the shell number 1, but they generally tended to be more diverse between themselves. Therefore, to draw Figure 5.2B, results from rock biofilms (2 and 4 cm away from adult barnacles) were pooled together to add depth to the statistical analysis. Of the 4 taxa investigated, only the proportion of α-Proteobacteria was statistically influenced by the sampling location ($p < 0.01$). *B. amphitrite*-associated communities were dominated by α-Proteobacteria and Bacteroidetes, with the first group being particularly abundant. The proportions of these two taxa in rock-attached biofilms were more variable, with the α-Proteobacteria ranging between 1 and 22% and the Bacteroidetes between 3 and 33%. The proportion of both Actinobacteria and Firmicutes was very similar in all biofilms sampled, with the first group being always less than 4%. In conclusion, the results suggest that the microbial community structure in shell-associated biofilms is highly similar and, at least for the α-Proteobacteria, is different with respect to rock-associated biofilms.

To investigate whether differences could also be found at a lower taxonomic level, the phylotype diversity within each higher taxa was investigated by comparing DGGE profiles of 2 barnacle-associated communities and 2 rock-attached biofilms (Figure 5.3). It is good practice to support the comparison of DGGE by implementing software (e.g. BioNumerics, GelCompare) to standardise band sorting. However, these approaches require manual input of the parameters to define bands and are particularly useful to compare a high number of samples. For our duplicate analysis of the two environments sampled, the manual transformation of band patterns into binary matrices was considered sufficient to support the following analysis. For each sample, the number of bands found in the 4 bacterial groups (within taxa diversity) was similar, with only the Actinobacteria showing a proportionally lower diversity in certain profiles. This suggests that the dominance of α-Proteobacteria found by qPCR resulted from the higher abundance of specific bacterial strains rather than from an increased within-taxa diversity.

Resemblance analysis of the binary matrices obtained (band presence/absence) indicated that samples of the same type (shell or rock) always clustered together, with their branches on the UPGMA tree generated that departed between the 6% and 33% similarity level. In particular, of the 121 unique bands generated, 12% were found in all samples, 35% appeared in only 1 sample, and 22% and 23% were found in either both shell or both rock profiles respectively. The shell- and rock-specific bands appeared to

be evenly distributed across the 4 taxa investigated, suggesting that the capacity of bacteria to show a preference in regard to the surface they grow attached to is not a feature specific to certain higher taxa. However, the molecular mechanisms underlying this phenomenon remain unclear.

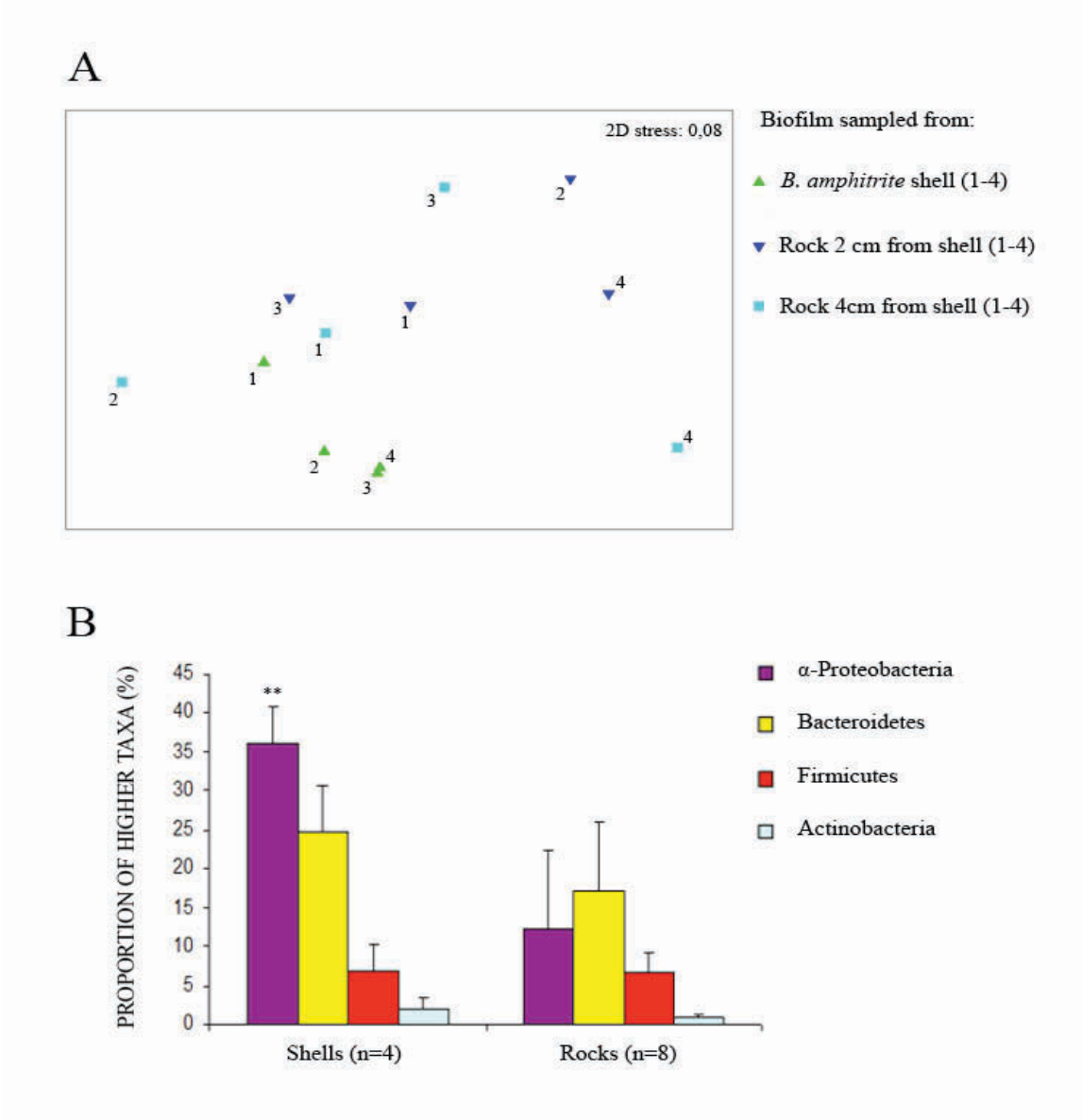


Figure 5.2 - **Comparison of biofilms growing on rocks and *Balanus amphitrite* shells.**

A: quantitative sample similarities were investigated with the Kulczynski index (Clarke & Warwick, 2001) using the relative proportion of higher bacterial taxa as variables. A non-metric multi-dimensional scaling (MDS) analysis was performed on the resemblance matrix obtained and 1000 restarts were carried out to draw a 2D scatter plot. B: means and standard deviations of the proportion of the four taxa investigated are represented for the 'shell' and 'rock' samples.

A number of factors influence the community structure of microbial biofilms, including environmental parameters (e.g. temperature, pH and nutrient levels), trophic

interactions between community members, phage and virus outbreaks, coordinated migratory events of specific populations and predation by other organisms (Battin*, et al.*, 2007). For this reason, understanding the mechanisms that lead to biofilm development on barnacles is inevitably difficult and falls outside the scope of this study. Nevertheless, what the results suggested is that the outer shell of *B. amphitrite* hosted bacteria that were not present on nearby rocks and favoured the dominance of α-Proteobacteria. The existence of specific association between barnacles and bacteria was not known before and these initial data provided a scientific base to support this. Although, the extent of this phenomenon needs to be confirmed with a wider sampling effort, it was hypothesised that these bacteria may play a role in barnacle kin recognition, and the experiments discussed below were designed to answer this biological question.

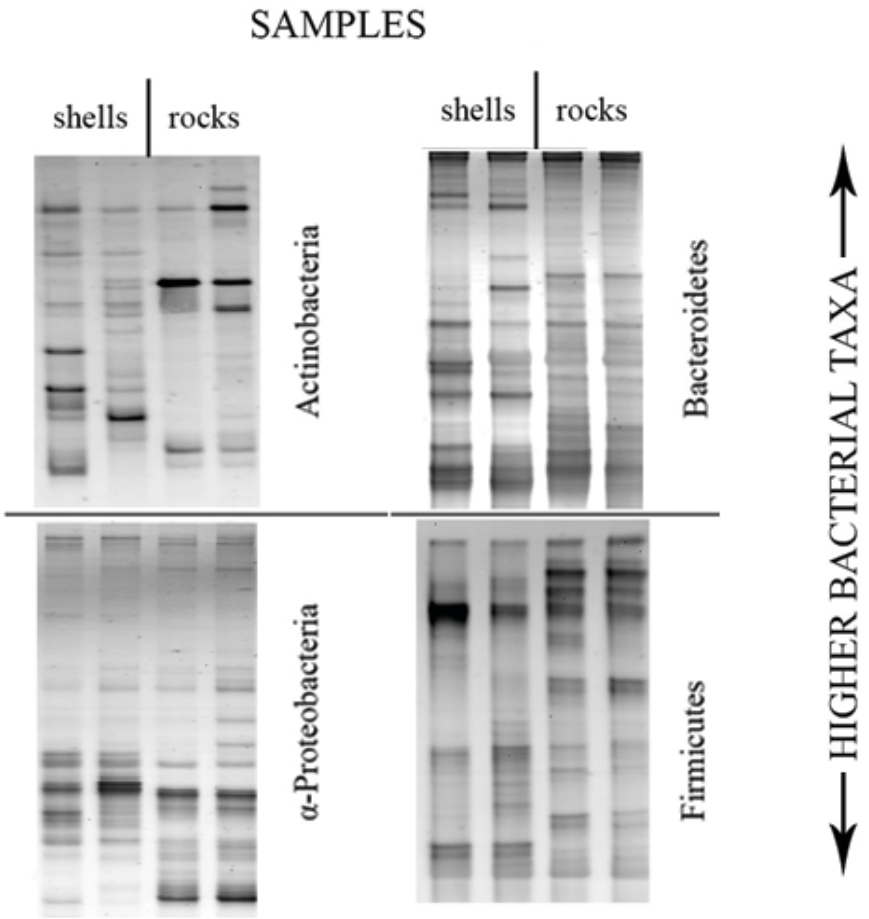


Figure 5.3 - **Taxon-specific DGGE profiles of rock- and shell-associated bacterial communities**

Two replicates for each sampling site were run for each taxon investigated.

5.3.2. Effect of bacteria on larval settlement

It is well known that the use of nutrient-rich media (e.g. marine broth) tends to favour the isolation of rapidly-growing, heterotrophic bacteria and often does not allow the identification of environment-specific, functionally important microbes. For this reason, a microbiological medium intended to mimic the barnacle-specific environment was designed for isolating bacteria associated with *B. amphitrite*. Medium AE was inoculated with either whole barnacles or with microbiologically fouled inert materials. Biofilms were allowed to develop on surfaces for 20h before their ability to influence larval settlement was assessed. It was found that bacteria originating from barnacles were capable of inducing permanent attachment of conspecifics (Figure 5.4). The experiment was repeated 3 times and the result was reproducible. In barnacles, the cyprid's capacity to settle often varies between batches of larvae. For this reason, settlement results were normalised as a percentage (+ or -) of the settlement that occurred in the control wells.

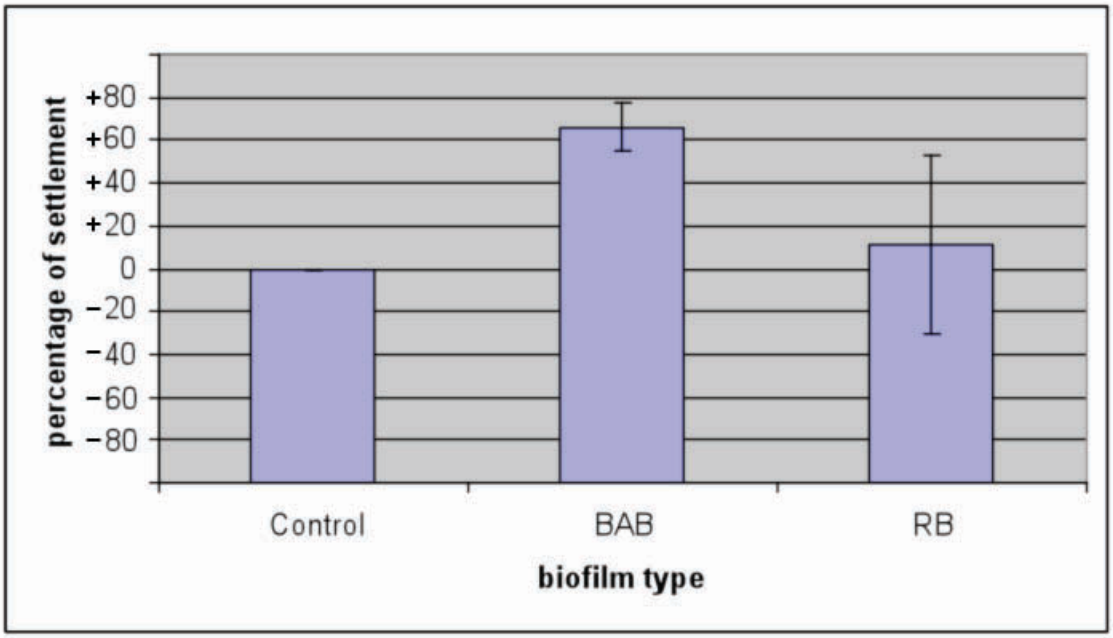


Figure 5.4 - **Effect of biofilm on larval settlement**

Comparison of *Balanus amphitrite* settlement between biofilm-free surface (Control), barnacle-originating biofilms (BAB) and rock-originating biofilms (RB). BAB and RB are expressed as +/- percentage of settlement on control surface.

Barnacle-associated bacteria (BAB) produced an average of 67% increase in settlement, with the promotion being statistically significant ($p<0.05$). For biofilm deriving from bacteria attached to inert materials (RB), the results were more variable, reflecting the data existing in the literature (Table 1), with both promotory and inhibitory effects. The same medium was used to grow all biofilms tested, suggesting that the noted effect is a consequence of the presence of specific bacteria rather than a different metabolic activity of similar communities. Biofilm structure was assessed with fluorescent microscopy and it was noted that 20h-old communities grew as multispecies microcolonies that created a patchiness on the surface (controls were biofilm-free), with a maximum calculated distance between colonies of 50µm. This biofilm architecture likely reflected the existence of bacterial aggregates growing in the original AE medium that had subsequently settled on the well bottom. The taxonomic composition of both BAB and RB communities was assessed by DGGE profiling and band sequencing.

Table 5 - **Sequenced bands from biofilms inducing and inhibiting settlement.**

Band	**Isolated from**	**Closest relative**	**Sequence length**	**Similarity**
34	RB	*Pseudoalteromonas mariniglutinosa*	163	99%
36	RB	*Vibrio rotiferianus*	162	100%
37	RB	*Pseudoalteromonas prydzensis*	162	99%
39	BAB	*Vibrio ichthyoenteri*	161	97%
40	BAB	*Pseudoalteromonas piscicida*	164	100%
41	BAB	*Vibrio ponticus*	164	97%
42	BAB	*Vibrio changasii*	158	100%
43	BAB	*Vibrio alginolyticus*	158	96%

Biofilms from one of the triplicate experiments, in which BAB had a average promotory effect on settlement (+73%) while RB was inhibitory (-37%), were removed from the surfaces by swabbing inoculated wells at the time of adding the cyprids (these wells were not used for settlement purposes) and DNA extracted. The PCR-DGGE

profiles obtained by amplifying 16S genes with universal primers (Figure 5.5) suggested that single biofilms were dominated by few bacterial strains, as profiles were characterised by a low number of bands. Replicate biofilms gave rise to similar banding patterns, but band intensity displayed major changes. This indicates that, even though the wells were inoculated with the same bacterial culture, well-specific microbial interactions led to the establishment of different community structures.

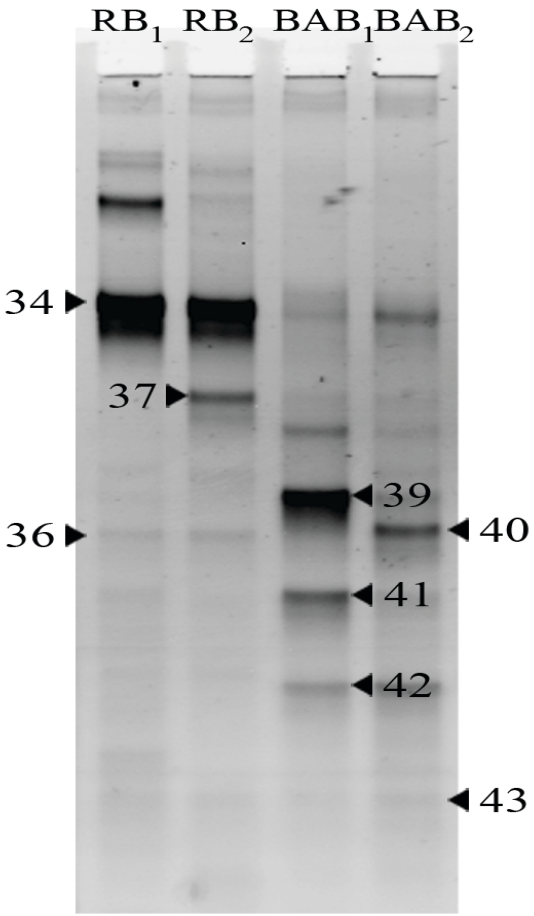


Figure 5.5 - **DGGE profiles of settlement-inducing or -inhibiting biofilms.**
RB = biofilms generated by rock-isolated bacteria (promontory); BAB = biofilms formed by barnacle-associated bacteria (inhibitory).

It is known that the existence of complex interspecies dynamics makes it difficult to generate a model biofilm that develops with a reproducible pattern. However, the settlement assay indicated that BAB communities possessed a conserved capacity to stimulate larval permanent attachment. For this reason, it was hoped to identify strains that were uniquely present in BAB communities as a possible indication of their

involvement in settlement facilitation. A number of bands were sequenced for this purpose, and the BLAST results of matching these sequences with those present in the NCBI database are reported in Table 5. All sequenced bands fell within the γ - Proteobacteria class, clustering in the *Pseudoaltermonas* and *Vibrio* groups; vast genera known to include strains that are functionally diverse. Unfortunately, the proportion of γ-Proteobacteria on natural shell-associated communities was not measured. Bands 39 and 42 (Table 5) appeared to be present in only BAB biofilms, and they were both tentatively identified as members of the *Vibrio*, a large taxonomic group for which short 16S rRNA sequences alone do not allow ecotype sorting (Zanetti*, et al.*, 1999; Le Roux*, et al.*, 2009). Interestingly, band 42 had a perfect match with a sequence from a bacterium that was recently isolated from *B. amphitrite* (unpublished results). Initial settlement assays with this isolated strain indicated that it was capable of inducing permanent attachment when growing as a single species biofilm, but the extent of this promotion was not as marked as that found for BAB (result not shown).

It is known that the influence of bacteria on *B. amphitrite* settlement is subjected to the complex interplay between taxonomic composition, nutrient media, secreted compounds and biofilm structure (Khandeparker*, et al.*, 2002; Khandeparker*, et al.*, 2003). Interestingly, Khandeparker and colleagues (2003) found that the effect of biofilms could turn to inhibitory by adding lectins to the media, with a similar mechanism to that found in lectin-treated settlement-inducing adult extract (Matsumura*, et al.*, 1998b). Lectins are a diverse group of proteins with carbohydrate-binding activity that is often involved in non-self recognition and innate immunity (Kilpatrick, 2002; Cambi & Figdor, 2003). As the settlement-inducing protein complex - SIPC - produced by *B amphitrite* is a large glycoprotein, it was suggested that the molecular signal recognised by exploring cyprids resides in the SIPC carbohydrates (Matsumura*, et al.*, 1998c). A similar effect on inducing biofilms indicates that their mechanism may also be based on the capacity of cyprids to recognise specific sugars present in the extracellular polymeric matrix (EPS) that keeps the biofilm bound together (Kachlany*, et al.*, 2001; Neu*, et al.*, 2001; Flemming & Wingender, 2010). In particular, settlement promotion was blocked by mannose-binding lectins in both adult extract and biofilms (Khandeparker*, et al.*, 2003). As mannose is the main carbohydrate present in the SIPC (Pagett et al., under review) and is also produced by many bacterial strains, it cannot be excluded that the promotion of permanent attachment by certain biofilms is due to random structural similarities between the SIPC and bacterial EPS. Nevertheless, the reproducible settlement-inducing effect of adult-associated communities suggests that

B. amphitrite may have evolved mechanisms to harbour bacteria that are more cyprid-friendly than those found in other environments. The development of model bacterial strains, which are both functionally important in the wild and cultivable, will help in elucidating the ecological mechanisms that lead to kin recognition in *B. amphitrite*.

5.3.3. Quantitative analysis of bacteria associated with laboratory-raised barnacles

As metazoan life likely arose in the oceans over 1 billion years ago, it is expected that multicellular organisms and bacteria have evolved a vast array of strategies to take advantage of each others' qualities in exploiting the surrounding environment. In fact, it is not uncommon to find that these strategies have developed into stable associations that increase the holobiont's adaptability. For example, the metabolism of many marine invertebrates is fuelled by chemoautotrophic bacteria (Dubilier, *et al.*, 2008), and symbiotic relationships exist both with the bacteria living inside the hosts' cells (Lösekann, *et al.*, 2008; Zielinski, *et al.*, 2009; Jensen, *et al.*, 2010; Kjeldsen, *et al.*, 2010) and extracellularly (Rinke, *et al.*, 2006; Duperron, *et al.*, 2008). Furthermore, the capacity of the α -Proteobacteria *Wolbachia* spp. to colonise the arthropod's reproductive system as a means to colonise new environments (i.e. individuals) has been known for decades (Charlat, *et al.*, 2003; Parlman, *et al.*, 2006) and many more bacteria symbiotic to invertebrates that are vertically transmitted are being discovered (Cary & Giovannoni, 1993; Sipe, *et al.*, 2000; Schmitt, *et al.*, 2007; Sharp, *et al.*, 2007; Bright & Bulgheresi, 2010). Therefore, it is not unreasonable to assume that barnacles have evolved mechanisms to interact with bacteria, but the existence of such relationships has never been documented.

It was possible to extract bacterial 16S rRNA genes from various *B. amphitrite* specimens, and the sorting of this gene within broad taxonomic lineages was performed to determine the predominant taxa associated with this barnacle. In all samples investigated, it was found that the sum of 16S rRNA copies belonging to the Bacteroidetes, Firmicutes, and α- and μ-Proteobacteria always constituted between 54 and 100% of the total ribosomal gene content (Figure 5.6). The microbial community existing in the tank water appeared to be dominated by Bacteroidetes, while α-Proteobacteria added up to a mere 11%, and the three biological replicates gave rise to highly similar community structures, suggesting that the qPCR technique presented in chapter 4 successfully captured the relative proportions of higher taxa. Whole adult bodies were found to be heavily colonised by α-Proteobacteria, while Firmicutes and μ-

Proteobacteria were almost absent. These latter groups are known to be predominant in the gastro-intestinal tract of many marine species (Harris, 1993), but this seems not to be the case for *B. amphitrite*. On the contrary, the dominance of α-Proteobacteria may indicate that this taxon serves important functions in maintaining the adult's health. Furthermore, the reduced proportion of α-Proteobacteria found associated with the exuviae (recently moulted cuticular vestige), suggests that members of this taxon inhabit some internal parts of the animal's body rather than simply attaching to its cuticle. As the mitochondria also belong to the α-Proteobacteria clade, it was investigated whether the strong signal for α-Proteobacteria was coming from the amplification of *B. amphitrite*'s mtDNA. The primers used in this study were checked against mitochondrial genes of this species (Bacchetti De Gregoris*, et al.*, 2009) and of the barnacle *Megabalanus volcano* (unpublished results, GenBank entry AB167539). The complementary sequences were not found, suggesting that the α-Proteobacteria are indeed predominant.

Interestingly, biofilms growing on barnacle shell were also dominated by α-Proteobacteria, as were biofilms growing on wild barnacles from Dona Paula, India (Figure 5.2). As the interplay of surface physicochemical proprieties and cell-wall characteristics are known to affect bacterial adhesion (Neu, 1996), it can be suggested that the specific shell chemistry favours the attachment of some α-Proteobacteria species. Although this mechanism may influence early colonisation, it is unlikely that it generates long lasting effects, as mature biofilms are often composed of very different species compared to early communities (Jackson*, et al.*, 2001; Jones*, et al.*, 2007). Therefore, the idea that a different mechanism determines the conserved dominance of certain α-Proteobacteria members is favoured, possibly involving the utilisation of barnacle-secreted compounds to sustain bacterial growth. However, the beneficial effects, if any, of these bacteria on the barnacle remain unknown. Finally, the relative proportions of bacteria forming the communities associated with each sample investigated appeared to be highly conserved. Each *B. amphitrite* body part or developmental stage housed a relatively stable community, which differed from those associated with other sample types (Figure 5.6). This is interesting, as it suggests that bacteria possess different roles during barnacle development and that interactions are established between *B. amphitrite* and various bacterial taxa. An investigation of the microbial communities at the lower taxonomic level in association with an imaging approach to revealing their spatial distribution could provide important clues to the activities of these microbiotas.

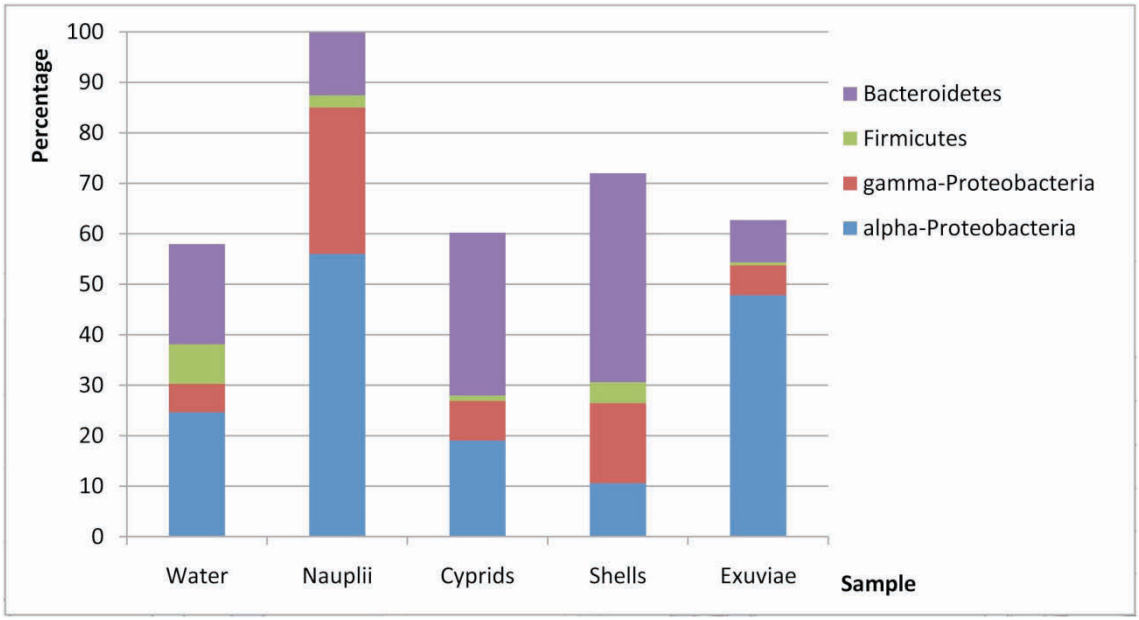


Figure 5.6 - **Proportion of higher bacterial taxa associated with laboratory-raised barnacles**

For each sample investigated with the qPCR protocol, DNA was extracted in triplicate. ΔCt data transformed in quantities (percentage) for individual samples are reported below.

	Water 1	Water 2	Water 3	Nauplii 1	Nauplii 2	Nauplii 3	Cyprids 1	Cyprids 2	Cyprids 3
α-Proteobacteria	7.66	14.42	9.73	17.53	28.64	27.78	49.40	58.05	69.99
µ-Proteobacteria	11.34	16.99	19.25	7.07	2.92	6.88	17.66	30.82	42.65
Firmicutes	3.54	5.04	3.87	15.10	5.36	3.14	3.58	2.05	1.76
Bacteroidetes	50.72	40.66	32.70	13.50	16.64	29.23	11.95	9.51	15.82

	Moults 1	Moults 2	Moults 3	Adults 1	Adults 2	Adults 3	Shells 1	Shells 2	Shells 3
α-Proteobacteria	15.06	27.83	14.26	71.08	35.50	77.29	29.04	39.30	50.87
µ-Proteobacteria	4.99	8.32	10.34	2.05	1.61	1.77	1.18	0.57	12.98
Firmicutes	1.74	0.78	0.53	0.30	0.31	0.57	0.27	0.29	0.96
Bacteroidetes	25.67	34.23	36.86	24.31	6.86	5.17	11.73	5.37	3.69

5.3.4. Quantitative analysis of bacteria naturally associated with Balanus amphitrite

Using laboratory-raised barnacles I showed that *B. amphitrite* associates with diverse microbial communities, some of which appear to be spatiotemporally stable, though how these consistent associations are assembled and maintained from one generation to the next was not investigated. To demonstrate that a number of bacteria are vertically transmitted in the wood-boring bivalve *Bankia setacea*, Sipe and colleagues (2000) isolated bacterial rRNA genes from the host gill, gonad tissue, and recently spawned eggs, and looked at the similarity between symbiotic communities through 16S sequencing and *in situ* hybridization. Similarly, in the experiment discussed here, DNA was extracted from the seminal vesicles (testis), egg masses, just spawned nauplii and adult shell surfaces of wild *B. amphitrite*. The proportion of 16S

rRNA copies belonging to the phyla Actinobacteria, Firmicutes, Bacteroidetes and the subdivision α of the Proteobacteria was determined. Furthermore, the characterisation of the bacterial communities at the lower taxonomic level was attempted by DGGE profiling of the taxon-specific qPCR products. Unfortunately, the quality of these gels was low, which prevented the inclusion of the DGGE analysis. Therefore, only the qPCR results are discussed here.

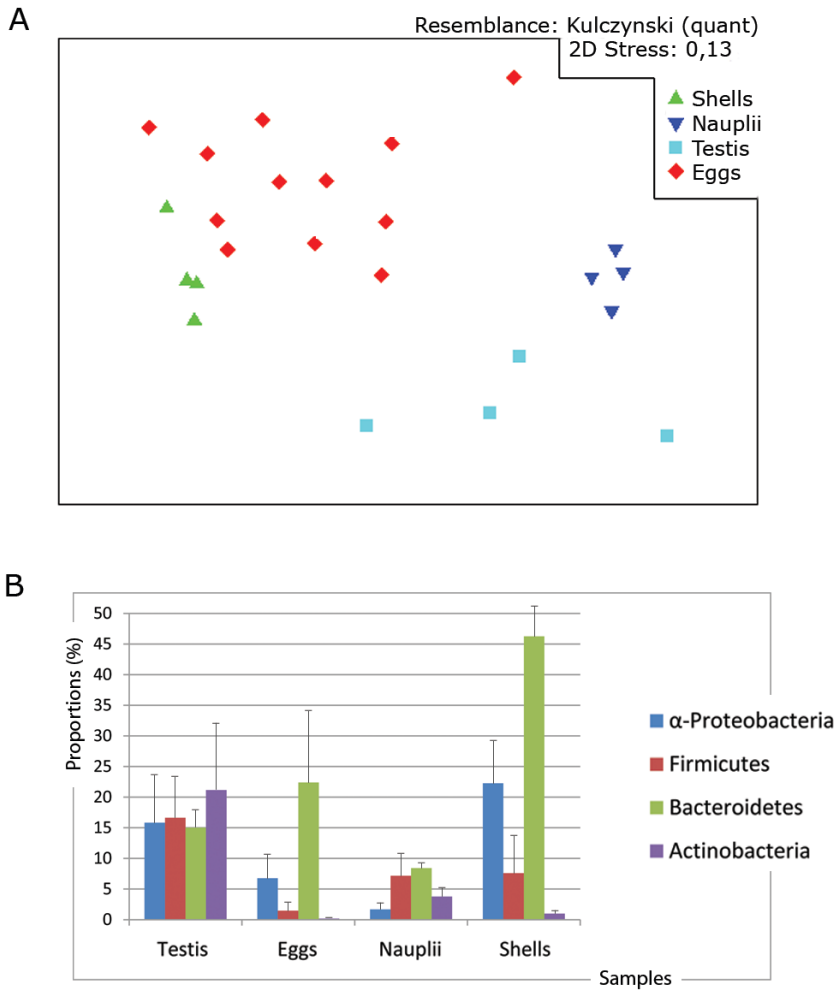


Figure 5.7 - **Proportion of higher bacterial taxa naturally associated with barnacles**

A = Non-metric multi-dimensional scaling (MDS) of samples resemblance (Kulczynski similarity) from the point of view of proportion of higher bacterial taxa as determined by 16S rRNA copy numbers; B = Average (colour bars), and standard deviation (black bars), proportion (calculated as percentages) of bacterial taxa associated with different body parts/developmental stage of *Balanus amphitrite*.

It was possible to extract bacterial 16S genes from all samples investigated, and the multi-dimensional scaling of the similarity between the structures of these associated microbiotas suggests that a relatively stable mixture of bacteria colonised the barnacles in a tissue/developmental stage-specific manner (Figure 5.7A). The

proportion of higher bacterial taxa in the testis appeared to be the most variable, even though the average concentration was between 15 and 21% for all 4 phylogenetic groups (Figure 5.7B). This was the only tissue in which the Actinobacteria were a significant fraction, constituting up to a third of the total 16S rRNA genes extracted. However, members of this group were not transmitted to the eggs, or if so, only in an almost undetectable way, suggesting that the Actinobacteria do not play a fundamental role during pre-hatching embryo development.

Of the 4 higher taxa quantified, the Bacteroidetes were the only conspicuous group associated with the egg masses, with the Gram positives almost completely absent. However, the major proportion of 16S rRNA genes isolated fell outside the 4 taxa, suggesting that other phyla are dominant in egg-associated communities. Furthermore, the bacteria associated with eggs at different stages of development (Figure 5.1) remained relatively stable, pointing out that the existing differences between the microbial communities associated with eggs and larvae are not achieved through a steady modification but rather through a sudden change. This may be explained by hypothesising that the dominant taxa colonising the eggs (i.e. Bacteroidetes) inhabit the external part of the egg membrane. As a consequence, the rupture of the membrane may lead to the hatching of larvae that have lost the majority of these bacteria. Although comparatively small, the same mechanism may explain the reduction in α-Proteobacteria as well as the proportional increase of the Gram positive phyla. However, microbial assemblages that colonised newly hatched larvae were dominated by taxa other than the 4 investigated here, and only the Firmicutes sporadically reached the 10% of total 16S copies present.

Interestingly, Bacteroidetes were strongly dominant in the biofilm growing on the outer shell of wild barnacles collected in Beaufort, USA. This is different from what was found attached to *B. amphitrite* from Dona Paula, India, where the α-Proteobacteria were the most abundant group. However, in both sampling sites, the sum of these two taxa always comprised around 2/3 of the total 16S rRNA genes present. Although these two taxa are often dominant in marine environments (Biers*, et al.*, 2009), their high proportion on barnacle shells raises questions on the influence of these bacteria on the life cycle of *B. amphitrite*. However, data on higher taxa abundance alone do not allow conclusions to be drawn on the biological significance of these partnerships, and more detailed investigations at the strain level are required to confirm the importance of these barnacle-bacteria associations.

5.4 Conclusions

Understanding the molecular cues that help barnacle larvae to locate surfaces suitable for settlement is of great importance in biofouling research. It was hypothesised that the bacteria naturally associated with adults may be partially responsible for attracting conspecific larvae. To begin answering this question, taxonomic differences between microbial communities attached to *B. amphitrite* shells and to the surrounding rocks were assessed. The quantitative comparison indicated that shell-associated biofilm hosted a larger proportion of α -Proteobacteria. Barnacle-specific communities also differed from rock biofilm when investigated at lower taxonomic level. It was possible to grow some members of this community on artificial media and the effect of multispecies biofilms on larval settlement was investigated. It was found that bacteria associated with adult barnacles promoted settlement in comparison to biofilm-free surfaces and to non-barnacle-specific microbial communities. Bacteria inducing settlement were tentatively identified as members of the *Vibrio*. Furthermore, bacteria associated with different developmental stages and body parts of *B. amphitrite* were investigated. It was found that microbial communities clustered according to barnacle development, indicating specificity. However, further analyses at the lower taxonomic level are required to confirm this finding.

Chapter 6. Summary and suggestions for future study

The aim of this study was to look at the diversity of bacteria associated with the barnacle *Balanus amphitrite* and to develop molecular tools that would allow the genes involved in inter-kingdom interactions to be studied.

6.1. Insights into *B. amphitrite* genetics

The recent development of high-throughput sequencing methods has allowed the rapid generation of a vast amount of data regarding the genes possessed by any species of interest, and obtaining these data is fundamental to achieving a thorough understanding of biological processes that are important for both basic and applied research. Regarding *B. amphitrite*, the literature presented only few genetic studies, mainly focusing on the molecular basis of gregariousness. In order to improve our knowledge on *B. amphitrite*'s transcriptome, a large sequencing effort was undertaken to characterise the genes expressed during the naupliar, cyprid and adult stages (in chapter 2). In total, more than 190 x 10^6 nucleotides were sequenced, which were clustered into 13,891 genes. The database created to store this information will be a valuable resource for future research aiming to understand the molecular regulation of the *B. amphitrite* life cycle. It was estimated that these genes may represent around 2/3 of the total protein coding sequences contained in the barnacle genome. Most of them were found to be expressed during a single developmental stage, with only 503 genes found in all 3 libraries.

Common bioinformatics tools were applied to extrapolate biological information from these sequences and the database was searched for genes previously suggested to be involved in settlement regulation. It was found that *bcs-6* was expressed in all 3 developmental stages, making it unlikely to be involved in permanent attachment. Furthermore, as lectins have been hypothesised to influence settlement cue recognition in barnacles, the database was searched for lectin-like isotigs. Out of the 10 best hits found, a proteins, uniquely expressed in the cyprid stage, matched a mannose receptor, and its nucleotide sequences was 33% identical to a lectin (BRA-3) isolated from *Megabalanus rosa*. As mannose has been suggested to be a molecular cue for conspecific recognition in *B. amphitrite*, further characterisation of this gene may suggest its involvement in settlement.

On the other hand, a qRT-PCR assay to monitor gene expression in this species was developed (in chapter 3). To this end, the capacity of 11 genes to normalise qRT-PCR data was investigated by determining their expression stability in 7 *B. amphitrite* developmental stages. It was found that the genes *mt-cyb* and *mt-nd1* were the most suitable normalisation factors. The opportunities that this new qRT-PCR approach opens to the study of gene functions in this species have been discussed.

6.2. Bacteria associated with *B. amphitrite*

All previous investigations on the interactions between *B. amphitrite* and bacteria have focused on the effect of microbes on larval settlement. However, these studies have been inconclusive, as they have generated contrasting results. Nevertheless, interactions between bacteria and invertebrates are very common in the marine environment, and it is likely that *B. amphitrite* has developed some too, not necessarily influencing settlement. To explore this territory, it was hypothesised that stable associations between the barnacle and specific bacteria should be found if their interactions are the basis of important biological mechanisms. For this reason, the first ever taxonomic survey of the bacteria associated with *B. amphitrite* was undertaken in this study.

Firstly, a new taxon-specific qPCR-DGGE was sought in order to generate a cheap and rapid alternative to commonly used molecular methods for the characterisation of microbial communities (in chapter 4). More than 60 primer pairs were tested to attain specificity and the final qPCR assay was validated on artificial mixes of bacterial 16S sequences. If specific association between barnacle and bacteria were known, widespread 16S rRNA cloning and sequencing may have served better the purpose of identifying functionally important bacterial strains than the qPCR assay presented here. However, as specific bacterial functions in the *B. amphitrite* life cycle have not been demonstrated in the literature, the inexpensive qPCR technique has allowed to investigate tens of samples and to obtain from their analysis an indication on the existence of specific barnacle-bacteria associations.

The data presented in chapter 5 indicate that bacteria are capable of colonising various organs during *B. amphitrite* development. It was possible to extract bacterial 16S genes from seminal vesicles, egg masses, nauplii and cypris larvae. Quantitative phylum- and class-specific sorting of these ribosomal sequences suggested that mixed microbial communities were associated with each sample. Interestingly, the structure of

these communities was found to be relatively stable in each sample type but different between types. These findings suggested that the nature of *B. amphitrite*-bacteria associations is subjected to profound changes during barnacle development. These changes in higher bacterial taxa preponderance likely reflected underlying alterations in strain composition. However, a characterisation of these communities at the lower taxonomic level was not achieved.

Finally, the biofilms-forming bacteria inhabiting the outer shell of barnacles were compared to those present on the surrounding rocks. The shells appeared to select for alpha-Proteobacteria, although other differences emerged from the taxon-specific DGGE profiles of these communities. Furthermore, multi-species biofilms isolated from *B. amphitrite*'s shell were capable of inducing settlement of conspecific larvae compared to other marine biofilms. However, no alpha-Proteobacteria were found in inductive biofilms, and member of the genus *Vibrio* or *Pseudoalteromonas* were more likely responsible for the noted effect. Member of the same taxa were also present in inhibitory microbial communities, indicating that the capacity of influencing *B. amphitrite* settlement behaviour is strain/ecotype-specific rather than conserved in broad taxonomic groups as in the alga *Ulva intestinalis*. The combination of these results with information existing in the literature led to the suggestion that specific carbohydrate composition of the biofilm's extracellular polymeric matrix may be responsible for promoting larval settlement in *B. amphitrite*.

6.3. Future work

As mentioned earlier in the text, the extensive gene database created in this study represents a valuable resource for future investigation aiming to understand *B. amphitrite* life cycle. Furthermore, despite settlement being an evolutionary conserved developmental mechanism that characterise many marine invertebrates, a model organism to study its regulation at a molecular level has yet to emerge. The commercial interests in understanding settlement in barnacles and the relative simplicity in growing them under laboratory conditions suggest that *B. amphitrite* should be a candidate species.

The possible influence of lectin-like protein in sensing settlement cues has emerged from the literature. In the *B. amphitrite* gene database presented here, a possible mannose-receptor was found. Future molecular and biochemical studies will confirm whether this gene is involved in settlement cue perception. For example, RNA

interference has proven to be a successful method to block the expression of specific genes in the copepod *Lepeophtheirus salmonis* (Frank Nilsen, University of Bergen, personal communication). A similar approach to functional genetics in *B. amphitrite* could be applied.

Regarding the bacteria interacting with barnacles, the existence of reproducible associations has emerged from this study. A biogeographical comparison of these partnerships at the lower taxonomical level should reveal whether some bacteria are stably associated with *B. amphitrite*. To this end, a full-length 16S rRNA survey should be undertaken. Furthermore, imaging of *ad hoc*-designed 16S probes (i.e. fluorescence *in situ* hybridisation) will indicate where specific microbes are located within the barnacle body and suggest a possible ecological explanation for their presence.

Finally, the taxon-specific qPCR-DGGE method presented here has been proven to be a powerful technique to characterise multispecies microbial communities. Although further optimisation in the DGGE step appeared to be necessary, this method possesses multiple applications in the fields of microbial ecology and evolution. In fact, a cheap two-steps approach capable of providing information on both the diversity and abundance of bacterial species can have a profound impact on our understanding of the mechanisms that influence microbial distribution, with particular reference on the formation and succession in natural biofilm communities.

References

Acosta-Martínez V, Dowd S, Sun Y & Allen V (2008) Tag-encoded pyrosequencing analysis of bacterial diversity in a single soil type as affected by management and land use. *Soil Biol and Biochem* **40**: 2762-2770.

Adams MD, Kelley JM, Gocayne JD*, et al.* (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**: 1651-1656.

Ahn K, Huh J-W, Park S-J*, et al.* (2008) Selection of internal reference genes for SYBR green qRT-PCR studies of rhesus monkey (*Macaca mulatta*) tissues. *BMC Mol Biol* **9**: 78.

Aldred N, Clare AS. (2008) The adhesive strategies of cyprids and development of barnacle-resistant marine coatings. *Biofouling*. **24**:351-63.

Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**: 403-410.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W & Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acids Res* **25**: 3389-3402.

Alzieu C (2000) Environmental impact of TBT: the French experience. *Sci Total Environ* **258**: 99-102.

Alzieu CL, Sanjuan J, Deltreil JP & Borel M (1986) Tin contamination in Arcachon Bay: Effects on oyster shell anomalies. *Mar Pollu Bull* **17**: 494-498.

Amann R & Fuchs BM (2008) Single-cell identification in microbial communities by improved fluorescence *in situ* hybridization techniques. *Nat Rev Micro* **6**: 339-348.

Amann R, Ludwig W & Schleifer KH (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol Rev* **59**: 143-169.

Andersen CL, Jensen JL & Orntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: A model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res* **64**: 5245-5250.

Anderson DT (1994) *Barnacles. Structure, function, development and evolution*. Chapman & Hall, London.

Apweiler R, Attwood TK, Bairoch A*, et al.* (2000) InterPro--an integrated documentation resource for protein families, domains and functional sites. *Bioinformatics* **16**: 1145-1150.

Armstrong E, Yan L, Boyd KG, Wright PC & Burgess JG (2001) The symbiotic role of marine microbes on living surfaces. *Hydrobiologia* **461**: 37-40.

Ashburner M, Ball CA, Blake JA*, et al.* (2000) Gene Ontology: tool for the unification of biology. *Nat Genet* **25**: 25-29.

Ashelford KE, Weightman AJ & Fry JC (2002) PRIMROSE: a computer program for generating and estimating the phylogenetic range of 16S rRNA oligonucleotide probes and primers in conjunction with the RDP-II database. *Nucl Acids Res* **30**: 3481-3489.

Avelin Mary S, Vitalina Mary S, Rittschof D & Nagabhushanam R (1993) Bacterial-barnacle interaction: Potential of using juncellins and antibiotics to alter structure of bacterial communities. *J Chem Ecol* **19**: 2155-2167.

Azam F & Malfatti F (2007) Microbial structuring of marine ecosystems. *Nat Rev Micro* **5**: 782-791.

Bacchetti De Gregoris T, Borra M, Biffali E, Bekel T, Burgess JG, Kirby R & Clare A (2009) Construction of an adult barnacle (*Balanus amphitrite*) cDNA library and selection of reference genes for quantitative RT-PCR studies. *BMC Mol Biol* **10**: 62.

Bachmann K & Rheinsmith EL (1973) Nuclear DNA amounts in Pacific Crustacea. *Chromosoma* **43**: 225-236.

Bao W-Y, Satuito C, Yang J-L & Kitamura H (2007) Larval settlement and metamorphosis of the mussel *Mytilus galloprovincialis* in response to biofilms. *Mar Biol* **150**: 565-574.

Bapteste E, O'Malley M, Beiko R*, et al.* (2009) Prokaryotic evolution and the tree of life are two different things. *Biology Direct* **4**: 34.

Barnes H (1957) Processes of restoration and synchronization in marine ecology. The spring diatom increase and the 'spawning' of the common barnacle, *Balanus balanoides* (L.) *Année Biol* **33**: 67-85.

Barnes H, Barnes M & Klepal W (1977) Studies on the reproduction of cirripedes I. Introduction: copulation, release of oocytes, and formation of the egg lamellae. *J Exp Mar Biol Ecol* **27**: 195-218.

Battin TJ & Sengschmitt D (1999) Linking sediment biofilms, hydrodynamics, and river bed clogging: Evidence from a large river. *Microb Ecol* **37**: 185-196.

Battin TJ, Kaplan LA, Denis Newbold J & Hansen CME (2003) Contributions of microbial biofilms to ecosystem processes in stream mesocosms. *Nature* **426**: 439-442.

Battin TJ, Sloan WT, Kjelleberg S, Daims H, Head IM, Curtis TP & Eberl L (2007) Microbial landscapes: new paths to biofilm research. *Nat Rev Microbiol* **5**: 76-81.

Becker S, Boger P, Oehlmann R & Ernst A (2000) PCR bias in ecological analysis: a case study for quantitative Taq nuclease assays in analyses of microbial communities. *Appl Environ Microbiol* **66**: 4945-4953.

Bekel T, Henckel K, Küster H*, et al.* (2009) The Sequence Analysis and Management System - SAMS-2.0: Data management and sequence analysis adapted to changing requirements from traditional sanger sequencing to ultrafast sequencing technologies. *J Biotech* **140**: 3-12.

Bettencourt R, Pinheiro M, Egas C, Gomes P, Afonso M, Shank T & Santos R (2010) High-throughput sequencing and analysis of the gill tissue transcriptome from the deep-sea hydrothermal vent mussel *Bathymodiolus azoricus*. *BMC Genomics* **11**: 559.

Bhaskar PV & Bhosle NB (2006) Bacterial extracellular polymeric substance (EPS): A carrier of heavy metals in the marine food-chain. *Environ Int* **32**: 191-198.

Biers EJ, Sun S & Howard EC (2009) Prokaryotic genomes and diversity in surface ocean waters: interrogating the global ocean sampling metagenome. *Appl Environ Microbiol* **75**: 2221-2229.

Blackwood CB, Oaks A & Buyer JS (2005) Phylum- and class-specific PCR primers for general microbial community analysis. *Appl Environ Microbiol* **71**: 6193-6198.

Bohannon J (2008) Confusing kinships. *Science* **320**: 1031-1033.

Bonar D, Weiner R & Colwell R (1986) Microbial-invertebrate interactions and potential for biotechnology. *Microb Ecol* **12**: 101-110.

Bourget E (1977) Shell structure in sessile barnacles. *Naturalists Can* **104**: 281-323.

Bourget E (1987) Barnacle shells: Composition, structure and growth. *Barnacle biology,*(Southward AJ, eds.), pp. 267-285. A.A.Balkema, Rotterdam.

Bourne DG, Iida Y, Uthicke S & Smith-Keune C (2008) Changes in coral-associated microbial communities during a bleaching event. *ISME J* **2**: 350-363.

Boxshall GA (2007) Crustacean classification: on-going controversies and unresolved problems. *Zootaxa* **1668**: 313-325.

Branda SS, Vik S, Friedman L & Kolter R (2005) Biofilms: the matrix revisited. *Trends Microbiol* **13**: 20-26.

Brenner D, Staley J & Krieg N (2000) Classification of procaryotic organisms and the concept of bacterial speciation. *Bergey's manual of systematic bacteriology,* Vol. 1 (Boone DR, Castenholz RW & Garrity GM, eds.), pp. 27-38. Springer Verlag, New York.

Bright M & Bulgheresi S (2010) A complex journey: transmission of microbial symbionts. *Nat Rev Micro* **8**: 218-230.

Buffle J, Wilkinson KJ, Stoll S, Filella M & Zhang J (1998) A generalized description of aquatic colloidal interactions: The three-colloidal component approach. *Environ Sci Technol* **32**: 2887-2899.

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

Bustin SA & Nolan T (2004) Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *J Biomol Tech* **3**: 155-166.

Bustin SA, Benes V, Nolan T & Pfaffl MW (2005) Quantitative real-time RT-PCR - a perspective. *J Mol Endocrinol* **34**: 597-601.

Cambi A & Figdor CG (2003) Dual function of C-type lectin-like receptors in the immune system. *Curr Opin Cell Biol* **15**: 539-546.

Cary SC & Giovannoni SJ (1993) Transovarial inheritance of endosymbiotic bacteria in clams inhabiting deep-sea hydrothermal vents and cold seeps. *Proc Natl Acad Sci USA* **90**: 5695-5699.

Chambers LD, Stokes KR, Walsh FC & Wood RJK (2006) Modern approaches to marine antifouling coatings. *Surf Coat Tech* **201**: 3642-3652.

Champ MA (2000) A review of organotin regulatory strategies, pending actions, related costs and benefits. *Sci Total Environ* **258**: 21-71.

Champ MA & Seligman PF (1996) An introduction to organotin compounds and their use in antifouling coatings. *Organotin: environmental fate and effects,*(Champ MA & Seligman PF, eds.), pp. Chapman and Hall, London.

Chan LCA & Walker G (1998) The settlement of *Pomatoceros lamarkii* larvae (Polychaeta: Sabellida: Serpulidae): a laboratory study. *Biofouling* **12**:71–80.

Charlat S, Hurst GDD & Merçot H (2003) Evolutionary consequences of *Wolbachia* infections. *Trends Genet* **19**: 217-223.

Chen X, Zeng Y & Jiao N (2006) Development and evaluation of specific 16S rDNA primers for marine Cytophaga-Flavobacteria cluster. *Mol Ecol Notes* **6**: 1278-1281.

Cirillo JD (1999) Exploring a novel perspective on pathogenic relationships. *Trends Microbiol* **7**: 96-98.

Clare AS (1998) Towards Nontoxic Antifouling. *J Mar Biotech* **6**: 3-6.

Clare AS & Matsumura K (2000) Nature and perception of barnacle settlement pheromones. *Biofouling* **15**: 57 - 71.

Clare AS & Høeg JT (2008) *Balanus amphitrite* or *Amphibalanus amphitrite*? A note on barnacle nomenclature. *Biofouling* **24**: 55 - 57.

Clare AS, Thomas R & Rittschof D (1995) Evidence for the involvement of cyclic AMP in the pheromonal modulation of barnacle settlement. *J Exp Biol* **198**: 655-664.

Clare AS, Walker G, Holland DL & Crisp DJ (1982) Barnacle egg hatching: a novel role for a prostaglandin-like compound *Mar Biol Lett* **3**: 113-120.

Clare AS, Walker G, Holland DL & Crisp DJ (1984) Nature of the barnacle hatching substance and the role of dopamine in the egg hatching process. *Advances in Invertebrate Reproduction,* Vol. 3 (Engels W, eds.), pp. 569. Elsevier, Amsterdam.

Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. *Aust J Ecol* **18**: 117-143.
Clarke KR & Warwick RM (2001) *Change in marine communities: an approach to statistical analysis and interpretation*. PRIMER-E, Plymouth.
Cohan FM (2002) What are bacterial species? *Annu Rev Microbiol* **56**: 457.
Cohan FM (2006) Towards a conceptual and operational union of bacterial systematics, ecology, and evolution. *Philos T Roy Soc B* **361**: 1985-1996.
Cohan FM & Perry EB (2007) A systematics for discovering the fundamental units of bacterial diversity. *Curr Biol* **17**: R373-R386.
Crisp DJ (1956) A Substance promoting hatching and liberation of young in cirripedes. *Nature* **178**: 263-263.
Crisp DJ & Knight-Jones EW (1953) The mechanism of aggregation in barnacle populations. *J Animal Ecol* **22**: 360-362.
Crisp DJ & Meadows PS (1962) The chemical basis of gregariousness in Cirripedes. *Proc R Soc Lond B* **156**: 500-520.
Crisp DJ & Richardson CA (1975) Tidally produced internal bands in the shell of *Elminius modestus*. *Mar Biol* **33**: 155-160.
Curtis TP & Sloan WT (2005) Exploring microbial diversity--A vast below. *Science* **309**: 1331-1333.
Curtis TP, Sloan WT & Scannell JW (2002) Estimating prokaryotic diversity and its limits. *Proc Natl Acad Sci USA* **99**: 10494-10499.
Dahms H-U, Dobretsov S & Qian P-Y (2004) The effect of bacterial and diatom biofilms on the settlement of the bryozoan *Bugula neritina*. *J Exp Mar Biol Ecol* **313**: 191-209.
Dang H & Lovell CR (2000) Bacterial primary colonization and early succession on surfaces in marine waters as determined by amplified rRNA gene restriction analysis and sequence analysis of 16S rRNA genes. *Appl Environ Microb* **66**: 467-475.
Dar SA, Kuenen JG & Muyzer G (2005) Nested PCR-denaturing gradient gel electrophoresis approach to determine the diversity of sulfate-reducing bacteria in complex microbial communities. *Appl Environ Microbiol* **71**: 2325-2330.
Darwin C (1851) *A monograph on the sub-class cirripedia with figures of all the species.* The Ray Society, London.
De Wolf P (1973) Ecological observations on the mechanisms of dispersal of barnacle larvae during planktonic life and settling. *Neth J Sea Res* **6**: 1-129.
Deepak SA, Kottapalli KR, Rakwal R*, et al.* (2007) Real-Time PCR: revolutionizing detection and expression analysis of genes. *Curr Genomics* **8**: 234-251.
Degnan BM & Morse DE (1995) Developmental and morphogenetic gene regulation in *Haliotis rufescens* larvae at metamorphosis. *Amer Zool* **35**: 391-398.
Dheda K, Huggett JF, Chang JS*, et al.* (2005) The implications of using an inappropriate reference gene for real-time reverse transcription PCR data normalization. *Analyt Biochem* **344**: 141-143.
Dobretsov S & Qian P-Y (2004) The role of epibotic bacteria from the surface of the soft coral *Dendronephthya* sp. in the inhibition of larval settlement. *J Exp Mar Biol Ecol* **299**: 35-50.
Dobretsov S, Dahms H-U & Qian P-Y (2006) Inhibition of biofouling by marine microorganisms and their metabolites. *Biofouling* **22**: 43-54.
Doolittle WF & Bapteste E (2007) Pattern pluralism and the Tree of Life hypothesis. *Proc Natl Acad Sci U S A* **104**: 2043-2049.
Dreanno C, Kirby RR & Clare AS (2006a) Locating the barnacle settlement pheromone: spatial and ontogenetic expression of the settlement-inducing protein complex of *Balanus amphitrite*. *Proc R Soc Lond B*.

Dreanno C, Matsumura K, Dohmae N, Takio K, Hirota H, Kirby RR & Clare AS (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**: 14396-14401.

Dubilier N, Bergin C & Lott C (2008) Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nat Rev Micro* **6**: 725-740.

Duperron S, Laurent MCZ, Gaill F & Gros O (2008) Sulphur-oxidizing extracellular bacteria in the gills of Mytilidae associated with wood falls. *FEMS Microbiol Ecol* **63**: 338-349.

Dykhuizen DE (1998) Santa Rosalia revisited: Why are there so many species of bacteria? *A van Leeuw J Microb* **73**: 25-33.

Dynes JJ, Lawrence JR, Korber DR, Swerhone GDW, Leppard GG & Hitchcock AP (2006) Quantitative mapping of chlorhexidine in natural river biofilms. *Sci Total Environ* **369**: 369-383.

Eddy SR (1998) Profile hidden Markov models. *Bioinformatics* **14**: 755-763.

Ewing B, Hillier L, Wendl MC & Green P (1998) Base-calling of automated sequencer traces using Phred. *Genome Res* **8**: 175-185.

Faimali M, Garaventa F, Terlizzi A, Chiantore M & Cattaneo-Vietti R (2004) The interplay of substrate nature and biofilm formation in regulating *Balanus amphitrite* Darwin, 1854 larval settlement. *J Exp Mar Biol Ecol* **306**: 37-50.

Falkowski PG, Fenchel T & Delong EF (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science* **320**: 1034-1039.

Feil EJ, Holmes EC, Bessen DE, Chan MS & Day NP (2001) Recombination within natural populations of pathogenic bacteria: short-term empirical estimates and long-term phylogenetic consequences. *Proc Natl Acad Sci USA* **98**: 182.

Field CB, Behrenfeld MJ, Randerson JT & Falkowski P (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* **281**: 237-240.

Fierer N & Jackson RB (2006) The diversity and biogeography of soil bacterial communities. *Proc Natl Acad Sci USA* **103**: 626-631.

Fierer N, Jackson JA, Vilgalys R & Jackson RB (2005) Assessment of soil microbial community structure by use of taxon-specific quantitative PCR assays. *Appl Environ Microbiol* **71**: 4117-4120.

Flemming H-C & Wingender J (2010) The biofilm matrix. *Nat Rev Micro* **8**: 623-633.

Fleury E, Huvet A, Lelong C*, et al.* (2009) Generation and analysis of a 29,745 unique Expressed Sequence Tags from the Pacific oyster (*Crassostrea gigas*) assembled into a publicly accessible database: the GigasDatabase. *BMC Genomics* **10**: 341.

Fraser C, Hanage WP & Spratt BG (2007) Recombination and the nature of bacterial speciation. *Science* **315**: 476-480.

Fromin N, Hamelin J, Tarnawski S*, et al.* (2002) Statistical analysis of denaturing gel electrophoresis (DGE) fingerprinting patterns. *Environ Microbiol* **4**: 634-643.

Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown MV & Naeem S (2006) Annually reoccurring bacterial communities are predictable from ocean conditions. *Proc Natl Acad Sci USA* **103**: 13104-13109.

Fuqua C & Greenberg EP (1998) Self perception in bacteria: quorum sensing with acylated homoserine lactones. *Curr Opin Cell Biol* **1**: 183-189.

Fusetani N & Clare AS (2006) *Antifouling compounds*. Springer, Berlin.

Gans J, Wolinsky M & Dunbar J (2005) Computational improvements reveal great bacterial diversity and high metal toxicity in soil. *Science* **309**: 1387-1390.

Gao B, Mohan R & Gupta RS (2009) Phylogenomics and protein signatures elucidating the evolutionary relationships among the Gammaproteobacteria. *Int J Syst Evol Microbiol* **59**: 234 - 247.

Gerhart DJ, Rittschof D, Hooper IR, Eisenman K, Meyer AE, Baier RE & Young C (1992) Rapid and inexpensive quantification of the combined polar components of surface wettability: Application to biofouling. *Biofouling* **5**: 251-259.

Gevers D, Dawyndt P, Vandamme P, Willems A, Vancanneyt M, Swings J & De Vos P (2006) Stepping stones towards a new prokaryotic taxonomy. *Philos T Roy Soc B* **361**: 1911-1916.

Gil-Turnes MS & Fenical W (1992) Embryos of *Homarus americanus* are protected by epibiotic bacteria. *Biol Bull* **182**: 105.

Gil-Turnes MS, Hay ME & Fenical W (1989) Symbiotic marine bacteria chemically defend crustacean embryos from a pathogenic fungus. *Science* **246**: 116.

Gilbert SF (2001) Ecological developmental biology: developmental biology meets the real world. *Develop Biol* **233**: 1-12.

Gilbert SF, McDonald E, Boyle N*, et al.* (2010) Symbiosis as a source of selectable epigenetic variation: taking the heat for the big guy. *Phil Trans R Soc B* **365**: 671-678.

Giovannoni S & Stingl U (2007) The importance of culturing bacterioplankton in the 'omics' age. *Nat Rev Micro* **5**: 820-826.

Giovannoni SJ, Britschgi TB, Moyer CL & Field KG (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**: 60-63.

Giriyan A & Pangam P (2006) *A Review of Antifouling Strategies: Alternatives to TBT*. TERI Press, New Delhi (India).

Glenner H (1999) Functional morpholohy of cirripede cypris: a comparative approach. *Barnacles. The biofoulers,*(Thompson M-F & Nagabhushanam R, eds.), pp. Regency publications, New Dheli.

Goebel W & Gross R (2001) Intracellular survival strategies of mutualistic and parasitic prokaryotes. *Trends Microbiol* **9**: 267-273.

Goesmann A, Linke B, Bartels D*, et al.* (2005) BRIGEP--the BRIDGE-based genome-transcriptome-proteome browser. *Nucl Acids Res* **33**: 710-716.

Goodfellow M, Manfio GP & Chun J (1997) Towards a practical species concept for cultivable bacteria. *The units of biodiversity,*(Claridge F, Dawah HA & Wilson MR, eds.), pp. Chapman and Hall, London.

Green JL, Bohannan BJM & Whitaker RJ (2008) Microbial biogeography: from taxonomy to traits. *Science* **320**: 1039-1043.

Grice EA, Kong HH, Conlan S*, et al.* (2009) Topographical and temporal diversity of the human skin microbiome. *Science* **324**: 1190-1192.

Guo X, Xia X, Tang R, Zhou J, Zhao H & Wang K (2008) Development of a real-time PCR method for *Firmicutes* and *Bacteroidetes* in faeces and its application to quantify intestinal population of obese and lean pigs. *Lett Appl Microbiol* **47**: 367-373.

Gupta RS & Mok A (2007) Phylogenomics and signature proteins for the alpha Proteobacteria and its main groups. *BMC Microbiol* **7**: 106.

Gupta RS & Lorenzini E (2007) Phylogeny and molecular signatures (conserved proteins and indels) that are specific for the Bacteroidetes and Chlorobi species. *BMC Evol Biol* **7**: 71.

Hadfield MG (1998) The D.P. Wilson Lecture: Research on settlement and metamorphosis of marine invertebrate larvae: past, present and future. *Biofouling* **12**: 9-29.

Hadfield MG, Unabia CRC, Smith CM & Michael TM (1994) Settlement preferences of the ubiquitous fouler Hydroides elegans. *Recent Developments in Biofouling Control,*(Thompson MF, Nagabhushanam R, Sarojini R & Fingerman M, eds.), pp. 65-74. Oxford University Press and IBH, New Delhi.

Hadfield MG, Carpizo-Ituarte EJ, del Carmen K & Nedved BT (2001) Metamorphic competence, a major adaptive convergence in marine invertebrate larvae. *Amer Zool* **41**: 1123-1131.

Hall LW, Pinkney AE & Laughlin R (1985) Acute and sublethal effects of organotin compounds on aquatic biota: An interpretative literature evaluation. *Crit Rev Toxicol* **14**: 159-209.

Hall LW, Lenkevich MJ, Scott Hall W, Pinkney AE & Bushong SJ (1987) Evaluation of butyltin compounds in Maryland waters of Chesapeake Bay. *Mar Pollu Bull* **18**: 78-83.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 95-98.

Hanage W, Fraser C & Spratt B (2005) Fuzzy species among recombinogenic bacteria. *BMC Biology* **3**: 6.

Hanage WP, Fraser C & Spratt BG (2006) Sequences, sequence clusters and bacterial species. *Philos T Roy Soc B* **361**: 1917-1927.

Handelsman J (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiol Mol Biol Rev* **68**: 669.

Harder T, Lau SCK, Dahms H-U & Qian P-Y (2002) Isolation of bacterial metabolites as natural inducers for larval settlement in the marine polychaete *Hydroides elegans* (Haswell). *J Chem Ecol* **28**: 2029-2043.

Harino H, O'Hara SCM, Burt GR, Chesman BS & Langston WJ (2005) Accumulation of butyltin compounds in benthic biota of the Mersey Estuary. *Mar Pollu Bull* **50**: 222-225.

Harris JM (1993) The presence, nature, and role of gut microflora in aquatic invertebrates: A synthesis. *Microb Ecol* **25**: 195-231.

Harte J (2003) Ecology: Tail of death and resurrection. *Nature* **424**: 1006-1007.

Hentschel U, Hopke J, Horn M, Friedrich AB, Wagner M, Hacker J & Moore BS (2002) Molecular evidence for a uniform microbial community in sponges from different oceans. *Appl Environ Microbiol* **68**: 4431-4440.

Heuer H, Wieland G, Schonfeld J, Schnwalder A, Gomes NCM & Smalla K (2001) Bacterial community profiling using DGGE or TGGE analysis. *Environmental Molecular Microbiology,*(Rochelle PA, eds.), pp. 177-190. Horizon Scientific Press, Wymondham, UK.

Hibino T, Loza-Coll M, Messier C*, et al.* (2006) The immune gene repertoire encoded in the purple sea urchin genome. *Develop Biol* **300**: 349-365.

Hofmann K & Stoffel W (1993) TMbase-a database of membrane spanning proteins segments. *Biol Chem* **374** 166-168.

Holben WE, Feris KP, Kettunen A & Apajalahti JHA (2004) GC fractionation enhances microbial community diversity assessment and detection of minority populations of bacteria by denaturing gradient gel electrophoresis. *Appl Environ Microbiol* **70**: 2263-2270.

Holm ER (1990) Attachment behavior in the barnacle *Balanus amphitrite amphitrite* (Darwin): genetic and environmental effects. *J Exp Mar Biol Ecol* **135**: 85-98.

Holmström C & Kjelleberg S (1999) Marine *Pseudoalteromonas* species are associated with higher organisms and produce biologically active extracellular agents. *FEMS Microbiol Ecol* **30**: 285-293.

Holmström C, Rittschof D & Kjelleberg S (1992) Inhibition of settlement by larvae of *Balanus amphitrite* and *Ciona intestinalis* by a surface-colonizing marine bacterium. *Appl Environ Microbiol* **58**: 2111-2115.

Holzapfel ACA (1904) Ships compositions. *Trans INA* 252-265.

Hou Y & Lin S (2009) Distinct gene number-genome size relationships for eukaryotes and non-eukaryotes: gene content estimation for dinoflagellate genomes. *PLoS ONE* **4**: e6978.

Hu J, Zhen H, Wan Y*, et al.* (2006) Trophic magnification of triphenyltin in a marine food web of Bohai Bay, North China: Comparison to tributyltin. *Environ Sci Technol* **40**: 3142-3147.

Huang X & Madan A (1999) CAP3: A DNA sequence assembly program. *Genome Res* **9**: 868-877.

Huang Y-L, Dobretsov S, Xiong H & Qian P-Y (2007) Effect of biofilm formation by *Pseudoalteromonas spongiae* on induction of larval settlement of the polychaete *Hydroides elegans*. *Appl Environ Microb* **73**: 6284-6288.

Hubbell SP (2006) Neutral theory and the evolution of ecological equivalence. *Ecology* **87**: 1387-1398.

Huggett J, Dheda K, Bustin S & Zumla A (2005) Real-time RT-PCR normalisation: strategies and considerations. *Genes Immun* **6**: 279-284.

Huggett MJ, Nedved BT & Hadfield MG (2009) Effects of initial surface wettability on biofilm formation and subsequent settlement of *Hydroides elegans*. *Biofouling* **25**: 387 - 399.

Humphrey JW, Oleson JP & N. SA (1998) *Greek and Roman technology: A sourcebook: Annotated translations of Greek and Latin texts and documents*. Routledge, Rome.

Hung OS, Thiyagarajan V, Zhang R, Wu RSS & Qian PY (2007) Attachment of *Balanus amphitrite* larvae to biofilms originating from contrasting environments. *Mar Ecol Progr Ser* **333**: 229-242.

Hunt DE, David LA, Gevers D, Preheim SP, Alm EJ & Polz MF (2008) Resource partitioning and sympatric differentiation among closely related bacterioplankton. *Science* **320**: 1081-1085.

Infante C, Matsuoka M, Asensio E, Canavate J, Reith M & Manchado M (2008) Selection of housekeeping genes for gene expression studies in larvae from flatfish using real-time PCR. *BMC Mol Biol* **9**: 28.

Jackson CR, Churchill PF & Roden EE (2001) Successional changes in bacterial assemblage structure during epilithic biofilm development. *Ecology* **82**: 555-566.

Jacobsen JA & Asmund G (2000) TBT in marine sediments and blue mussels (*Mytilus edulis*) from central-west Greenland. *Sci Total Environ* **245**: 131-136.

Jensen S, Duperron S, Birkeland N-K & Hovland M (2010) Intracellular *Oceanospirillales* bacteria inhabit gills of *Acesta bivalves*. *FEMS Microbiol Ecol* **74**: 523-533.

Jin-Woo L, Ji-Hyun N, Yang-Hoon K, Kyu-Ho L & Dong-Hun L (2008) Bacterial communities in the initial stage of marine biofilm formation on artificial surfaces. *J Microbiol* **46**: 174-182.

Joint I, Callow ME, Callow JA & Clarke KR (2000) The attachment of *Enteromorpha* zoospores to a bacterial biofilm assemblage. *Biofouling* **16**: 151-158.

Joint I, Tait K, Callow ME, Callow JA, Milton D, Williams P & Camara M (2002) Cell-to-cell communication across the prokaryote-eukaryote boundary. *Science* **298**: 1207.

Jones P, Cottrell M, Kirchman D & Dexter S (2007) Bacterial community structure of biofilms on artificial surfaces in an estuary. *Microb Ecol* **53**: 153-162.

Kachlany SC, Levery SB, Kim JS, Reuhs BL, Lion LW & Ghiorse WC (2001) Structure and carbohydrate analysis of the exopolysaccharide capsule of *Pseudomonas putida* G7. *Environ Microbiol* **3**: 774-784.

Kannan K, Corsolini S, Focardi S, Tanabe S & Tatsukawa R (1996) Accumulation patterns of butyltin compounds in dolphin, tuna and shark collected from Italian coastal waters. *Environ Contamin Toxicol* 19-23.

Keller EF (1995) *Refiguring life: Metaphors of twentieth century biology*. Columbia University Press, New York.

Kent AD, Yannarell AC, Rusak JA, Triplett EW & McMahon KD (2007) Synchrony in aquatic microbial community dynamics. *ISME J* **1**: 38-47.

Kent AD, Jones SE, Lauster GH, Graham JM, Newton RJ & McMahon KD (2006) Experimental manipulations of microbial food web interactions in a humic lake: shifting biological drivers of bacterial community structure. *Environ Microbiol* **8**: 1448.

Khandeparker L, Anil AC & Raghukumar S (2002) Factors regulating the production of different inducers in *Pseudomonas aeruginosa* with reference to larval metamorphosis in *Balanus amphitrite*. *Aquat Microb Ecol* **28**: 37-54.

Khandeparker L, Anil AC & Raghukumar S (2003) Barnacle larval destination: piloting possibilities by bacteria and lectin interaction. *J Exp Mar Biol Ecol* **289**: 1-13.

Khandeparker L, Anil AC & Raghukumar S (2006) Relevance of biofilm bacteria in modulating the larval metamorphosis of *Balanus amphitrite*. *FEMS Microbiol Ecol* **58**: 425-438.

Kiil S, Weinell CE, Pedersen MS & Dam-Johansen K (2001) Analysis of self-polishing antifouling paints using rotary experiments and mathematical modeling. *Ind Eng Chem Res* **40**: 3906-3920.

Kiil S, Dam-Johansen K, Weinell C, Pedersen M & Codolar S (2002) Dynamic simulations of a self-polishing antifouling paint exposed to seawater. *J Coat Technol* **74**: 45-54.

Kilpatrick DC (2002) Animal lectins: a historical introduction and overview. *Biochim Biophys Acta* **1572**: 187-197.

Kimbell JR & McFall-Ngai MJ (2003) The squid-Vibrio symbioses: From demes to genes. *Integr Comp Biol* **43**: 254-260.

Kirchman D, Graham S, Reish D & Mitchell R (1981) Bacteria induce settlement and metamorphosis of *Janua* (*Dexiospira*) *brasiliensis* Grube (Polychaeta:Spirpribidae). *J Exp Mar Biol Ecol* **56**: 153-163.

Kjeldsen KU, Obst M, Nakano H, Funch P & Schramm A (2010) Two types of endosymbiotic bacteria in the enigmatic marine worm *Xenoturbella bocki*. *Appl Environ Microb* **76**: 2657-2662.

Klappenbach JA, Saxman PR, Cole JR & Schmidt TM (2001) rrndb: the Ribosomal RNA Operon Copy Number Database. *Nucl Acids Res* **29**: 181-184.

Koeppel A, Perry EB, Sikorski J*, et al.* (2008) Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. *Proc Natl Acad Sci U S A* **105**: 2504-2509.

Kon-ya K, Shimidzu N, Otaki N, Yokoyama A, Adachi K & Miki W (1995) Inhibitory effect of bacterial ubiquinones on the settling of barnacle, *Balanus amphitrite*. *Cell Mol Life Sci* **51**: 153-155.

Konstantinidis KT & Tiedje JM (2005) Towards a genome-based taxonomy for prokaryotes. *J Bacteriol* **187**: 6258-6264.

Konstantinidis KT, Ramette A & Tiedje JM (2006) Toward a more robust assessment of intraspecies diversity, using fewer genetic markers. *Appl Environ Microb* **72**: 7286-7293.

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

Labare ML, Coon SL, Matthias C & Weiner RM (1997) Magnification of tributyl tin toxicity to oyster larvae by bioconcentration in biofilms of *Shewanella colwelliana*. *Appl Environ Microb* **63**: 4107-4110.

Lane DJ (1991) 16S/23S rRNA sequencing. *Nucleic acid techniques in bacterial systematics* (Stackebrandt E & goodfellow M, eds.), pp. 115-175. Wiley, Chichester.

Lau SC, Harder T & Qian PY (2003) Induction of larval settlement in the serpulid polychaete *Hydroides elegans* (Haswell): role of bacterial extracellular polymers. *Biofouling* **19**: 197-204.

Le Roux F, Zouine M, Chakroun N*, et al.* (2009) Genome sequence of *Vibrio splendidus*: an abundant planctonic marine species with a large genotypic diversity. *Environ Microbiol* **1**: 1.

Leitz T & Wagner T (1993) The marine bacterium *Alteromonas espejiana* induces metamorphosis of the hydroid *Hydractinia echinata*. *Mar Biol* **115**: 173-178.

Ley RE, Turnbaugh PJ, Klein S & Gordon JI (2006) Microbial ecology: Human gut microbes associated with obesity. *Nature* **444**: 1022-1023.

Ley RE, Bäckhed F, Turnbaugh P, Lozupone CA, Knight RD & Gordon JI (2005) Obesity alters gut microbial ecology. *Proc Natl Acad Sci U S A* **102**: 11070-11075.

Li H, Thiyagarajan V & Qian P-Y (2010) Response of cyprid specific genes to natural settlement cues in the barnacle *Balanus* (= *Amphibalanus*) *amphitrite*. *J Exp Mar Biol Ecol* **389**: 45-52.

Lindstrom ES, Kamst-Van Agterveld MP & Zwart G (2005) Distribution of typical freshwater bacterial groups is associated with pH, temperature, and lake water retention time. *Appl Environ Microb* **71**: 8201-8206.

Lochhead JH (1936) On the feeding mechanism of the nauplius of *Balanus perforatus* Bruguière. *J Linn Soc* **39**: 429-442.

Lösekann T, Robador A, Niemann H, Knittel K, Boetius A & Dubilier N (2008) Endosymbioses between bacteria and deep-sea siboglinid tubeworms from an Arctic Cold Seep (Haakon Mosby Mud Volcano, Barents Sea). *Environ Microbiol* **10**: 3237-3254.

Lowry OH, Rosenbrough NA, Farr AL & Randall RJ (1951) Protein measurement with the folin phenol reagent. *J Biol Chem* **193**: 265-275.

Loy A, Maixner F, Wagner M & Horn M (2006) ProbeBase--an online resource for rRNA-targeted oligonucleotide probes: new features 2007. *Nucl Acids Res* **35**: 800-804.

Ludemann H & Conrad R (2000) Molecular retrieval of large 16S rRNA gene fragments from an Italian rice paddy soil affiliated with the class *Actinobacteria*. *Syst Appl Microbiol* **23**: 582-584.

Maidak BL, Cole JR, Lilburn TG*, et al.* (2001) The RDP-II (Ribosomal Database Project). *Nucl Acids Res* **29**: 173-174.

Maki JS & Mitchell R (1985) Involvement of lectins in the settlement and metamorphosis of marine invertebrate larvae. *Bull Mar Sci* **37**: 675-683.

Maki JS, Rittschof D & Mitchell R (1992) Inhibition of larval barnacle attachment to bacterial films: An investigation of physical properties. *Microb Ecol* **23**: 97-106.

Maki JS, Rittschof D, Costlow JD & Mitchell R (1988) Inhibition of attachment of larval barnacles, *Balanus amphitrite*, by bacterial surface films. *Mar Biol* **97**: 199-206.

Maki JS, Ding L, Stokes J, Kavouras JH & Rittschof D (2000) Substratum/bacterial interactions and larval attachment: Films and exopolysaccharides of *Halomonas marina* (ATCC 25374) and their effect on barnacle cyprid larvae, *Balanus amphitrite* Darwin. *Biofouling* **16**: 159-170.

Manz W, Amann R, Ludwig W, Vancanneyt M & Schleifer K-H (1996) Application of a suite of 16S rRNA-specific oligonucleotide probes designed to investigate bacteria of the phylum cytophaga-flavobacter-bacteroides in the natural environment. *Microbiology* **142**: 1097-1106.

Marra MA, Hillier L & Waterston RH (1998) Expressed sequence tags - ESTablishing bridges between genomes. *Trends Genet* **14**: 4-7.

Masseille H (1934) Notes sur la salissure et les procédés d'entretien des carèens immergées. *Peintures-Pigments-Vernis* 54-57.

Matsumura K, Nagano M & Fusetani N (1998a) Purification of a larval settlement-inducing protein complex (SIPC) of the barnacle, *Balanus amphitrite*. *J Exp Biol* **281**: 12-20.

Matsumura K, Mori S, Nagano M & Fusetani N (1998b) Lentil lectin inhibits adult extract-induced settlement of the barnacle, *Balanus amphitrite*. *J Exp Biol* **280**: 213-219.

Matsumura K, Nagano M, Kato-Yoshinaga Y, Yamazaki M, Clare AS & Fusetani N (1998c) Immunological studies on the settlement-inducing protein complex (SIPC) of the barnacle *Balanus amphitrite* and its possible involvement in larva-larva interactions. *Proc R Soc Lond B* **265**: 1825-1830.

McFall-Ngai M (2006) Love the one you're with: Vertebrate guts shape their microbiota. *Cell* **127**: 247-249.

McFall-Ngai MJ (1999) Consequences of evolving with bacterial symbionts: Insights from the squid-vibrio associations. *Annu Rev Ecol Syst* **30**: 235-256.

McFall-Ngai MJ & Ruby EG (1991) Symbiont recognition and subsequent morphogenesis as early events in an animal-bacterial mutualism. *Science* **254**: 1491-1494.

McFall-Ngai MJ & Ruby EG (1998) Squids and vibrios: When they first meet. Reciprocal interactions between host and symbiont lead to the creation of a complex light-emitting organ. *Biosciences* 257-265.

McGill BJ, Maurer BA & Weiser MD (2006) Empirical evaluation of neutral theory. *Ecology* **87**: 1411-1423.

McVeigh HP, Munro J & Embley TM (1996) Molecular evidence for the presence of novel actinomycete lineages in a temperate forest soil. *J Ind Microbiol* **17**: 197-204.

Mead KS & Epel D (1995) Beakers versus breakers: How fertilisation in the laboratory differs from fertilisation in Nature. *Zygote* 95-99.

Meier H, Amann R, Ludwig W & Schleifer K-H (1999) Specific oligonucleotide probes for in situ detection of a major group of gram-positive bacteria with low DNA G+C content. *Syst Appl Microbiol* **22**: 186-196.

Meng P-J, Wang J-T, Liu L-L, Chen M-H & Hung T-C (2005) Toxicity and bioaccumulation of tributyltin and triphenyltin on oysters and rock shells collected from Taiwan maricuture area. *Sci Total Environ* **349**: 140-149.

Meyer F, Goesmann A, McHardy AC*, et al.* (2003) GenDB--an open source genome annotation system for prokaryote genomes. *Nucl Acids Res* **31**: 2187-2195.

Milne A (1975) Coated marine surfaces. eds.), pp. Internation Paint Co. Ltd.

Montgomery MK & McFall-Ngai M (1994) Bacterial symbionts induce host organ morphogenesis during early postembryonic development of the squid *Euprymna scolopes*. *Development* **120**: 1719-1729.

Morales SE & Holben WE (2009) Empirical testing of 16S PCR primer pairs reveals variance in target specificity and efficacy not suggested by *in silico* analysis. *Appl Environ Microbiol* 2677-2683.

Mühling M, Woolven-Allen J, Murrell JC & Joint I (2008) Improved group-specific PCR primers for denaturing gradient gel electrophoresis analysis of the genetic diversity of complex microbial communities. *ISME J* **2**: 379-392.

Mullineaux LS & Butman CA (1991) Initial contact, exploration and attachment of barnacle *(Balanus amphitrite*) cyprids settling in flow. *Mar Biol* **110**: 93-103.

Muyzer G, de Waal EC & Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* **59**: 695-700.

Muyzer G, Brinkhoff T, Nubel U, Santegoeds C, Schafer H & Waver C (1998) Denaturing gradient gel electrophoresis (DGGE) in microbial ecology. *Molecular Microbial Ecology Manual,*(Akkermans ADL, van Elsas JD & de Bruijn FJ, eds.), pp. 1-27. Kluwer Accademic Publishers, Dordrecht, The Netherland.

Nakanishi Y, Murashima K, Ohara H*, et al.* (2006) Increase in terminal restriction fragments of bacteroidetes-derived 16S rRNA genes after administration of short-chain fructooligosaccharides. *Appl Environ Microbiol* **72**: 6271-6276.

Nardi F, Spinsanti G, Boore JL, Carapelli A, Dallai R & Frati F (2003) Hexapod origins: monophyletic or paraphyletic? *Science* **299**: 1887-1889.

Neu T, Swerhone GD & Lawrence JR (2001) Assessment of lectin-binding analysis for *in situ* detection of glycoconjugates in biofilm systems. *Microbiology* **147**: 299-313.

Neu TR (1996) Significance of bacterial surface-active compounds in interaction of bacteria with interfaces. *Microbiol Rev* **60**: 151-166.

Nicholson JK, Holmes E & Wilson ID (2005) Gut microorganisms, mammalian metabolism and personalized health care. *Nat Rev Micro* **3**: 431-438.

Nisbet EG & Sleep NH (2001) The habitat and nature of early life. *Nature* **409**: 1083-1091.

Nobuhiko T, Masayuki H, Koji M, Hisao K & Tadayoshi S (1994) Cloning and analysis of the gene encoding lectin from the acorn barnacle *Megabalanus rosa*. *Gene* **150**: 407-408.

Nogueira J, Simplício B, Floréncio M & Bettencourt A (2003) Levels of tributyltin in sediments from Tagus estuary nature reserve. *Estuaries and Coasts* **26**: 798-802.

Nolan T, Hands RE & Bustin S (2006a) Quantification of mRNA using real-time RT-PCR. *Nat Protoc* **3**: 1559-1582.

Nolan T, Hands RE, Ogunkolade W & Bustin SA (2006b) SPUD: A quantitative PCR assay for the detection of inhibitors in nucleic acid preparations. *Analytic Biochem* **351**: 308-310.

Nott JA & Foster RB (1969) On the structure of the antennular attachment organ of the cypris larva of *Balanus balanoides*. *Philos T Roy Soc B* **256**: 115-134.

Nyholm SV, Stabb EV, Ruby EG & McFall-Ngai MJ (2000) Establishment of an animal-bacterial association: Recruiting symbiotic vibrios from the environment. *Proc Natl Acad Sci USA* **97**: 10231-10235.

Nyholm SV, Deplancke B, Gaskins HR, Apicella MA & McFall-Ngai MJ (2002) Roles of *Vibrio fischeri* and nonsymbiotic bacteria in the dynamics of mucus secretion during symbiont colonization of the *Euprymna scolopes* light organ. *Appl Environ Microb* **68**: 5113-5122.

O'Connor NJ & Richardson DL (1996) Effects of bacterial films on attachment of barnacle (*Balanus improvisus* Darwin) larvae: laboratory and field studies. *J Exp Mar Biol Ecol* **206**: 69-81.

O'Connor NJ & Richardson DL (1998) Attachment of barnacle (*Balanus amphitrite* Darwin) larvae: responses to bacterial films and extracellular materials. *J Exp Mar Biol Ecol* **226**: 115-129.

Oehlmann J, Stroben E, Schulte-Oehlmann U & Barbara B (1998) Imposex development in response to TBT pollution in *Hinia incrassata* (Strom, 1768) (Prosobranchia, Stenoglossa). *Aquatic Toxicol* **43**: 239-260.

Okazaki Y & Shizuri Y (2000) Structures of six cDNAs expressed specifically at cypris larvae of barnacles, *Balanus amphitrite*. *Gene* **250**: 127-135.

Okazaki Y & Shizuri Y (2001) Identification of the aromatic L-amino acid decarboxylase (AADC) gene and its expression in the attachment and metamorphosis of the barnacle, *Balanus amphitrite*. *Develop Growth Diff* **43**: 33-41.

Olguin-Uribe G, Abou-Mansour E, Boulander A, Débard H, Francisco C & Combaut G (1997) 6-Bromoindole-3-Carbaldehyde, from an *Acinetobacter* Sp. bacterium associated with the ascidian *Stomozoa murrayi*. *J Chem Ecol* **23**: 2507-2521.

Olivier F, Tremblay R, Bourget E & Rittschof D (2000) Barnacle settlement: field experiments on the influence of larval supply, tidal level, biofilm quality and age on *Balanus amphitrite* cyprids. *Mar Ecol Progr Ser* **199**: 185-204.

Omae I (2003) Organotin antifouling paints and their alternatives. *Appl Organometal Chem* **17**: 81-105.

Pace NR (1997) A molecular view of microbial diversity and the biosphere. *Science* **276**: 734-740.

Pagett EH, Abrahams LJ, Bones J, O'Donoghue N, Caldwell SG, Rudd MP & Clare AS (2011). Structural characterisation of the N-glycan moiety of the barnacle settlement-inducing protein complex (SIPC). Submitted to Glycobiology.

Palacios C, Zettler E, Amils R & Amaral-Zettler L (2008) Contrasting microbial community assembly hypotheses: a reconciling tale from the Rio Tinto. *PLoS ONE* **3**: e3853.

Parlman SJ, Hunter MS & Zchori-Fein E (2006) The emerging diversity of *Rickettsia*. *Proc R Soc Lond B* **273**: 2097-2106.

Parsek MR & Fuqua C (2004) Biofilms 2003: emerging themes and challenges in studies of surface-associated microbial life. *J Bacteriol* **186**: 4427-4440.

Patel P, Callow ME, Joint I & Callow JA (2003) Specificity in the settlement - modifying response of bacterial biofilms towards zoospores of the marine alga *Enteromorpha*. *Environ Microbiol* **5**: 338-349.

Patil JS & Anil AC (2005) Influence of diatom exopolymers and biofilms onmetamorphosis in the barnacle *Balanus amphitrite*. *Mar Ecol Progr Ser* **301**: 231-245.

Pedrós-Alió C (2006) Marine microbial diversity: can it be determined? *Trends Microbiol* **14**: 257-263.

Pellerito C, Nagy L, Pellerito L & Szorcsik A (2006) Biological activity studies on organotin(IV)n+ complexes and parent compounds. *J Organometal Chem* **691**: 1733-1747.

Pereira WE, Wade TL, Hostettler FD & Parchaso F (1999) Accumulation of butyltins in sediments and lipid tissues of the Asian clam, *Potamocorbula amurensis*, near Mare Island naval shipyard, San Francisco Bay. *Mar Pollu Bull* **38**: 1005-1010.

Pfaffl MW, Tichopad A, Prgomet C & Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. *Biotechnol Lett* **26**: 509-515.

Philippot L, Bru D, Saby PAN, Cuhel J, Arrouays D, Imek M & Hallin S (2009) Spatial patterns of bacterial taxa in nature reflect ecological traits of deep branches of the 16S rRNA bacterial tree. *Environ Microbiol* **11**: 3096-3104.

Philippot L, Andersson SGE, Battin TJ, Prosser JI, Schimel JP, Whitman WB & Hallin S (2010) The ecological coherence of high bacterial taxonomic ranks. *Nat Rev Micro* **8**: 523-529.

Pitombo FB (2004) Phylogenetic analysis of the *Balanidae* (Cirripedia, Balanomorpha). *Zool Scr* **33**: 261-276.

Polz MF, Distel DL, Zarda B, Amann R, Felbeck H, Ott JA & Cavanaugh CM (1994) Phylogenetic analysis of a highly specific association between ectosymbiotic, sulfur-oxidizing bacteria and a marine nematode. *Appl Environ Microb* **60**: 4461-4467.

Pop M (2009) Genome assembly reborn: recent computational challenges. *Brief Bioinform* **10**: 354-366.

Postgate JR (2000) *Microbes and Man*. Cambridge University Press, Cambridge.

Prosser JI, Bohannan BJM, Curtis TP*, et al.* (2007) The role of ecological theory in microbial ecology. *Nat Rev Micro* **5**: 384-392.

Qian P-Y, Thiyagarajan V, Lau SCK & Cheung SCK (2003) Relationship between bacterial community profile in biofilm and attachment of the acorn barnacle *Balanus amphitrite*. *Aquat Microb Ecol* **33**: 225-237.

Qian P-Y, Lau S, Dahms HU, Dobretsov S & Harder T (2007) Marine biofilms as mediators of colonization by marine macroorganisms: Implications for antifouling and aquaculture. *Mar Biotechnol*.

Quackenbush J, Liang F, Holt I, Pertea G & Upton J (2000) The TIGR Gene Indices: reconstruction and representation of expressed gene sequences. *Nucl Acids Res* **28**: 141-145.

Rahim SAKA, Li JY & Kitamura H (2004) Larval metamorphosis of the sea urchins, *Pseudocentrotus depressus* and *Anthocidaris crassispina* in response to microbial films. *Mar Biol* **144**: 71-78.

Rasmussen R (2001) Quantification on the LightCycler. *Rapid Cycle Real-time PCR, Methods and Applications,*(Meuer S, Wittwer C & Nakagawara K, eds.), pp. 21-34. Springer Press, Heidelberg.

Rast JP, Smith LC, Loza-Coll M, Hibino T & Litman GW (2006) Genomic insights into the immune system of the sea urchin. *Science* **314**: 952-956.

Regier JC, Shultz JW, Zwick A*, et al.* (2010) Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences. *Nature* **463**: 1079-1083.

Rinke C, Schmitz-Esser S, Stoecker K*, et al.* (2006) "Candidatus *Thiobios zoothamnicoli*," an ectosymbiotic bacterium covering the giant marine ciliate *Zoothamnium niveum*. *Appl Environ Microb* **72**: 2014-2021.

Rittschof D, Sandford Branscomb E & Costlow JD (1984) Settlement and behavior in relation to flow and surface in larval barnacles, *Balanus amphitrite* Darwin. *J Exp Mar Biol Ecol* **82**: 131-146.

Roberts J & Tsamenyi M (2008) International legal options for the control of biofouling on international vessels. *Marine Policy* **32**: 559-569

Rosenberg E, Sharon G & Zilber-Rosenberg I (2009) The hologenome theory of evolution contains Lamarckian aspects within a Darwinian framework. *Environ Microbiol* **11**: 2959-2962.

Rosenberg E, Koren O, Reshef L, Efrony R & Zilber-Rosenberg I (2007) The role of microorganisms in coral health, disease and evolution. *Nat Rev Micro* **5**: 355-362.

Rossello-Mora R (2005) Updating prokaryotic taxonomy. *J Bacteriol* **187**: 6255-6257.

Rozen S & Skaletsky HJ (2000) Primer3 on the WWW for general users and for biologist programmers. *Bioinformatics Methods and Protocols: Methods in Molecular Biology,*(Krawetz S & Misener S, eds.), pp. 365-386. Humana Press, Totowa, NJ.

Ruby E, Henderson B & McFall-Ngai M (2004) We get by with a little help from our (little) friends. *Science* **303**: 1305-1307.

Satuito CG, Shimizu K & Fusetani N (1997) Studies on the factors influencing larval settlement in *Balanus amphitrite* and *Mytilus galloprovincialis*. *Hydrobiologia* **V358**: 275-280.

Schloss PD & Handelsman J (2004) Status of the microbial census. *Microbiol Mol Biol Rev* **68**: 686-691.

Schmitt S, Weisz JB, Lindquist N & Hentschel U (2007) Vertical transmission of a phylogenetically complex microbial consortium in the viviparous sponge *Ircinia felix*. *Appl Environ Microb* **73**: 2067-2078.

Seligman PF, Maguire RJ, Lee RF, Hinga KR, Valkirs AO & Stang PM (1996) Persistance ond fate of tributyltin in aquatic ecosystems. *Organotin: environmental fate and effects* (Champ MA & Seligman PF, eds.), pp. Chapman & Hall, London.

Sharon G, Segal D, Ringo JM, Hefetz A, Zilber-Rosenberg I & Rosenberg E (2010) Commensal bacteria play a role in mating preference of *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **107**: 20051-20056.

Sharp KH, Eam B, Faulkner DJ & Haygood MG (2007) Vertical transmission of diverse microbes in the tropical sponge *Corticium* sp. *Appl Environ Microb* **73**: 622-629.

Shcheglov A, Zhulidov P, Bogdanova E & Shagin D (2007) Normalization of cDNA libraries. *Nucleic acids hybridization: modern applications* (Buzdin A & Lukyanov S, eds.), pp. 97-124. Springer, Dordrecht, The Netherlands.

Sipe AR, Wilbur AE & Cary SC (2000) Bacterial symbiont transmission in the wood-boring shipworm *Bankia setacea* (Bivalvia: Teredinidae). *Appl Environ Microb* **66**: 1685-1691.

Sipos R, Székely AJ, Palatinszky M, Révész S, Márialigeti K & Nikolausz M (2007) Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targetting bacterial community analysis. *FEMS Microbiol Ecol* **60**: 341-350.

Smith CJ & Osborn AM (2009) Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiol Ecol* **67**: 6-20.

Smith RD, Brown B, Ikonomi P & Schechter AN (2003) Exogenous reference RNA for normalization of real-time quantitative PC. *Biotechniques* **34**: 88-91.

Sodergren E & Weinstock GM & Davidson EH*, et al.* (2006) The Genome of the Sea Urchin Strongylocentrotus purpuratus. *Science* **314**: 941-952.

Sogin ML, Morrison HG, Huber JA, Mark Welch D, Huse SM & Neal PR (2006) Microbial diversity in the deep sea and the underexplored /`rare biosphere/'. *Proc Natl Acad Sci USA* **103**: 12115-12120.

Spinsanti G, Panti C, Lazzeri E, Marsili L, Casini S, Frati F & Fossi C (2006) Selection of reference genes for quantitative RT-PCR studies in striped dolphin (*Stenella coeruleoalba*) skin biopsies. *BMC Mol Biol* **7**: 32.

Stach JEM, Maldonado LA, Ward AC, Goodfellow M & Bull AT (2003) New primers for the class *Actinobacteria*: application to marine and terrestrial environments. *Environ Microbiol* **5**: 828-841.

Stackebrandt E, Frederiksen W, Garrity GM*, et al.* (2002) Report of the ad hoc committee for the re-evaluation of the species definition in bacteriology. *Int J Syst Evol Microbiol* **52**: 1043-1047.

Stanley D (2006) Prostaglandins and other eicosanoids in insects: biological significance. *Annu Rev Entomol* **51**: 25-44.

Steinert M, Hentschel U & Hacker J (2000) Symbiosis and pathogenesis: Evolution of the microbe-host interaction. *Naturwissenschaften* **87**: 1-11.

Strand J & Asmund G (2003) Tributyltin accumulation and effects in marine molluscs from West Greenland. *Environ Poll* **123**: 31-37.

Suzuki MT, Taylor LT & DeLong EF (2000) Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5'-nuclease assays. *Appl Environ Microbiol* **66**: 4605-4614.

Swift S, Downie JA, Whitehead NA, Barnard AML, Salmond GPC & Williams P (2001) Quorum sensing as a population-density-dependent determinant of bacterial physiology. *Advances in microbial physiology,* Vol. 45 eds.), pp. 199-270. Academic Press, London.

Tait K, Joint I, Daykin M, Milton DL, Williams P & Camara M (2005) Disruption of quorum sensing in seawater abolishes attraction of zoospores of the green alga *Ulva* to bacterial biofilms. *Environ Microbiol* **7**: 229-240.

Tanguy A, Bierne N, Saavedra C*, et al.* (2008) Increasing genomic information in bivalves through new EST collections in four species: Development of new genetic markers for environmental studies and genome evolution. *Gene* **408**: 27-36.

Tatusov R, Fedorova N, Jackson J*, et al.* (2003) The COG database: an updated version includes eukaryotes. *BMC Bioinformatics* **4**: 41.

Terlizzi A, Fraschetti S, Gianguzza P, Faimali M & Boero F (2001) Environmental impact of antifouling technologies: state of the art and perspectives. *Aquatic Cons: Mar Freshwater Ecosyst* **11**: 311-317.

Thiyagarajan V & Qian P-Y (2008) Proteomic analysis of larvae during development, attachment, and metamorphosis in the fouling barnacle, *Balanus amphitrite*. *Proteomics* **8**: 3164-3172.

Thiyagarajan V, Lau S, Cheung S & Qian P-Y (2006) Cypris habitat selection facilitated by microbial films influences the vertical distribution of subtidal barnacle *Balanus trigonus*. *Microb Ecol* **51**: 431-440.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F & Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucl Acids Res* **25**: 4876-4882.

Torsvik V, Goksoyr J & Daae FL (1990) High diversity in DNA of soil bacteria. *Appl Environ Microb* **56**: 782-787.

Townsin RL (2003) The ship hull fouling penalty. *Biofouling* **19**: 9-15.

Traas TP, Stab JA, Kramer PRG, Cofino WP & Aldenberg T (1996) Modeling and risk assessment of tributyltin accumulation in the food Web of a shallow freshwater lake. *Environ Sci Technol* **30**: 1227-1237.

v. Wintzingerode F, Gobel UB & Stackebrandt E (1997) Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. *FEMS Microbiol Rev* **21**: 213-229.

Vandamme P, Pot B, Gillis M, de Vos P, Kersters K & Swings J (1996) Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiol Rev* **60**: 407-438.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A & Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* **3**: 1-11.

Venier P, Pallavicini A, De Nardi B & Lanfranchi G (2003) Towards a catalogue of genes transcribed in multiple tissues of *Mytilus galloprovincialis*. *Gene* **314**: 29-40.

Visick KL & McFall-Ngai MJ (2000) An exclusive contract: Specificity in the *Vibrio fischeri-Euprymna scolopes* partnership. *J Bacteriol* **182**: 1779-1787.

Vogan CL, Maskrey BH, Taylor GW, Henry S, Pace-Asciak CR, Clare AS & Rowley AF (2003) Hepoxilins and trioxilins in barnacles: an analysis of their potential roles in egg hatching and larval settlement. *J Exp Biol* **206**: 3219-3226.

Voulvoulis N, Scrimshaw MD & Lester JN (1999) Review: alternative antifouling biocides. *Appl Organometal Chem* 135–143.

Wagner M, Horn M & Daims H (2003) Fluorescence *in situ* hybridisation for the identification and characterisation of prokaryotes. *Curr Opin Microbiol* **6**: 302-309.

Walker G (1971) A study of the cement apparatus of the cyprid larva of the barnacle *Balanus balanoides*. *Mar Biol* **9**: 205-212.

Walker G & Lee VE (1976) Surface structures and sense organs of the cypris larva of *Balanus balanoides* as seen by scanning and transmission electron microscopy. *J Zool* **178**: 161-172.

Walker G & Yule AB (1984) Temporary adhesion of the barnacle cyprid: The existance of an antennular adhesive secretion. *J Mar Biol Assoc U.K.* **64**: 679-686.

Walker G, Yule AB & Nott JA (1987) Structure and function in balanomorph larvae. *Barnacle biology,*(Southward AJ, eds.), pp. 307-328. A.A.Balkema, Rotterdam.

Walley LJ (1969) Studies on the larval structure and metamorphosis of *Balanus balanoides* (L.). *Phil Trans R Soc B* **256**: 237-280.

Walley LJ, White F & Brander KM (1971) Sperm activation and fertilization in *Balanus balanoides*. *J Mar Biol Assoc U.K.* **51**: 489-494.

Ward DM, Bateson MM, Ferris MJ, Kühl M, Wieland A, Koeppel A & Cohan FM (2006) Cyanobacterial ecotypes in the microbial mat community of Mushroom Spring (Yellowstone National Park, Wyoming) as species-like units linking microbial community composition, structure and function. *Phil Trans R Soc B* **361**: 1997-2008.

Watnick P & Kolter R (2000) Biofilm, city of microbes. *J Bacteriol* **182**: 2675-2679.

Webster NS, Smith LD, Heyward AJ, Watts JEM, Webb RI, Blackall LL & Negri AP (2004) Metamorphosis of a scleractinian coral in response to microbial biofilms. *Appl Environ Microbiol* **70**: 1213-1221.

Weisburg WG, Barns SM, Pelletier DA & Lane DJ (1991) 16S ribosomal DNA amplification for phylogenetic study. *J Bacteriol* **173**: 697-703.

Wheeler GL, Tait K, Taylor A, Brownlee C & Joint I (2006) Acyl-homoserine lactones modulate the settlement rate of zoospores of the marine alga *Ulva intestinalis* via a novel chemokinetic mechanism. *Plant Cell Environ* **29**: 608-618.

Whitaker RJ, Grogan DW & Taylor JW (2005) Recombination shapes the natural population structure of the hyperthermophilic archaeon *Sulfolobus islandicus*. *Mol Biol Evol* **22**: 2354-2361.

Whitman WB, Coleman DC & Wiebe WJ (1998) Prokaryotes: the unseen majority. *PNAS* **95**: 6578-6583.

WHOI (1952) The invention of protective devices. George Banta Publishing Co.

Wieczorek SK, Clare AS & Todd CD (1995) Inhibitory and facilitatory effects of microbial films on settlement of *Balanus amphitrite* larvae. *Mar Ecol Progr Ser* **119**: 221-228.

Woese C (1998) The universal ancestor. *Proc Natl Acad Sci U S A* **95**: 6854-6859.

Woese CR & Fox GE (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc Natl Acad Sci U S A* **74**: 5088-5090.

Woese CR, Kandler O & Wheelis ML (1990) Towards a natural system of organisms: Proposal for the domains Archaea, Bacteria, and Eucarya. *Proc Natl Acad Sci U S A* **87**: 4576-4579.

Woese CR, Fox GE, Zablen L*, et al.* (1975) Conservation of primary structure in 16S ribosomal RNA. *Nature* **254**: 83-86.

Wolfaardt GM, Lawrence JR, Headley JV, Robarts RD & Caldwell DE (1994) Microbial exopolymers provide a mechanism for bioaccumulation of contaminants. *Microb Ecol* **27**: 279-291.

Woodcock S, Van der Gast CJ, Bell T, Lunn M, Curtis TP, Head IM & Sloan WT (2007) Neutral assembly of bacterial communities. *FEMS Microbiol Ecol* **62**: 171-180.

Woods RG, Roper KE, Gauthier M, Bebell LM, Sung K, Degnan BM & Lavin MF (2004) Gene expression during early ascidian metamorphosis requires signalling by Hemps, an EGF-like protein. *Development* **131**: 2921-2933.

Yebra DM, Kiil S & Dam-Johansen K (2004) Antifouling technology-past, present and future steps towards efficient and environmentally friendly antifouling coatings. *Prog Organic Coatings* **50**: 75-104.

Young CFT (1867) *The fouling and corrosion of iron ship: their causes and means of prevention, with mode of application to the existing iron-clad*. The London Drawing Association, London.

Yule AB (1984) The effect of temperature on the swimming activity of barnacle nauplii. *Mar Biol Lett* **5**: 1-12.

Yule AB & Walker G (1987) Adhesion in barnacle. *Barnacle biology,*(Southward AJ, eds.), pp. 389-402. A.A. Balkema, Rotterdam.

Zanetti S, Deriu A, Dupre I, Sanguinetti M, Fadda G & Sechi LA (1999) Differentiation of *Vibrio alginolyticus* strains isolated from Sardinian waters by ribotyping and a new rapid PCR fingerprinting method. *Appl Environ Microb* **65**: 1871-1875.

Zardus JD, Nedved BT, Huang Y, Tran C & Hadfield MG (2008) Microbial biofilms facilitate adhesion in biofouling invertebrates. *Biol Bull* **214**: 91-98.

Zielinski FU, Pernthaler A, Duperron S, Raggi L, Giere O, Borowski C & Dubilier N (2009) Widespread occurrence of an intranuclear bacterial parasite in vent and seep bathymodiolin mussels. *Environ Microbiol* **11**: 1150-1167.

ZoBell CE (1938) The sequence of events in the fouling of submerged surfaces. *Official Digest Fed Paint Varnish Prod Clubs* 379-385.

Zobell CE & Allen EC (1935) The significance of marine bacteria in the fouling of submerged surfaces. *J Bacteriol* **29**: 239-251.

Supplementary Table S1 All primers used to develop the taxon-specific qPCR method

Primer name	Target taxon	Sequence (5'- 3')	Paired with primer	Maximum annealing Tº	Taxon-specific from Tº	Reference
ADF681F	α-proteobacteria	AGTGTAGAGGTGAAATT	907R 910R	61.6º 64º	aspecific 63º	Blackwood et al. 2005
α681F	α-proteobacteria	TNCIAGTGTAGAGGTGAAATT	910R	64º	62.5º	this study
α682F	α-proteobacteria	CIAGTGTAGAGGTGAAATT	908αR 910R 779R	64º 64º 60º	60.5º 62.5º 59º	this study
γR	γ-proteobacteria	CGTAAGGGCCATGAT	928F 926F	63.5º 63.5º	62.5º 63º	this study
γ1203R	γ-proteobacteria	CWISYCGTAAGGGCCATGAT	928F	n/a	aspecific	this study
γ1202R	γ-proteobacteria	CGTAAGGGCCATGATG	928F 1075F 1081F 1080Fγ	64º no band 63.5º 63º	aspecific no band aspecific 60.5º	this study
Firm1039R2	Firmicutes	CATGCACCACCTGTCW	926F	n/a	aspecific	this study
Firm1040R	Firmicutes	ACCATGCACCACCTGTC	798F 926F 928F-Firm 807F-Firm 807F2-Firm	62.5º 63º 63º 63º 64º	60.5º 60.5º 60.5º 61.6º 61.6º	this study
Firm1047R	Firmicutes	TGACGACAACCATGCAC	928F	n/a	aspecific	this study
CF967	Bacteroidetes	GGTAAGGTTCCTCGCGTA	798cfbF	63 º	60.5 º	Chen et al 2006
cfb967R	Bacteroidetes	GGTAAGGTTCCTCGCGTAT	798cfbF 799F2	64º 61.6º	60.5º n/a	this study
cfb967R2	Bacteroidetes	AGGTTCCTCGCGTAT	798F 799F2	61.6º 61.6º	60.5º n/a	this study
AB1165r	Actinobacteria	ACCTTCCTCCGAGTTRAC	926F	*	aspecific *	Ludemann et al, 2000

continues

Supplementary Table S1 (continued)

Primer name	Target taxon	Sequence (5'- 3')	Paired with primer	Maximum annealing Tº	Taxon-specific from Tº	Reference
Act1159r	Actinobacteria	TCCGAGTTRACCCCGGC	926F	*	aspecific *	Blackwood et al., 2005
Act920F	Actinobacteria	GAGTACGGCCGCAAGGCTA	1062R Act1200R Act1200R2	*	aspecific * aspecific * aspecific *	this study
Act920F2	Actinobacteria	GTACGGCCGCAAGGCTA	1062R Act1200R Act1200R2	*	aspecific * aspecific * aspecific *	this study
Act920F3	Actinobacteria	TACGGCCGCAAGGCTA	Act1200R Act1200R2	*	60.5 º * 60.5 º *	this study
Act920F4	Actinobacteria	GGCCGCAAGGCTA	Act1200R Act1200R2	*	no band no band	this study
U519F	General	CAGCMGCCGCGGTAATWC	779R 907R	61.6º 64º	general general	Suzuki et al 1996
798F	General	AAYAGGATTAGATACCCT	910R 1062R	61.6º n/a	general general	this study
926F	General	AAACTCAAAKGAATTGACGG	1062R	64º	general	Lane et al. 1991
798cfbF	General/Bacteroidetes	CRAACAGGATTAGATACCCT				this study
799F2	General/Bacteroidetes	AACAGGATTAGATACCCTG				Edwards et al 2007
928F-Firm	General/Firmicutes	TGAAACTYAAAGGAATTGACG				this study
807F-Firm	General/Firmicutes	TAGATACCCTRGTAGTCCA				this study
807F2-Firm	General/Firmicutes	GATTAGATACCCTRGTAGTCCA				this study
928F	General	TIAAACTCAAAKGAATTGACGGG				this study
1081F	General	GTCAGCTCGTGYYGTGAG				this study

continues

Supplementary Table S1 (continue)

Primer name	Target taxon	Sequence (5'- 3')	Paired with primer	Maximum annealing Tº	Taxon-specific from Tº	Reference
1080γF	Gen/γ-proteobacteria	TCGTCAGCTCGTGTYGTGA				this study
1075F	General	GGCTGTCGTCAGCTCGTGT				this study
779R	General	AGGGTATCTAATCCTRTT				this study
907R	General	CCGTCAATTCMTTTGAGTTT				Lane et al 1991
910R	General	GYCCCCGTCAATTCMTTTGAG				this study
908αR	General/α-proteobacteria	CCCCGTCAATTCCTTTGAGTT				this study
1062R	General	CTCACRRCACGAGCTGAC				this study
Act1200R	General/Actinobacteria	TCRTCCCCACCTTCCTCCG				this study
Act1200R2	General/Actinobacteria	CGTCRTCCCCACCTTCCTCCG				this study
27F	General	AGAGTTTGATCCTGGCTCAG				Lane et al 1991
1525R	General	AAGGAGGTGWTCCARCC				Lane et al 1991

Legend table 1

Nucleotides symbols: R = A or G; Y = C or T; N = any nucleotide; W = A or T; M = A or C; K = T or G; I = inosine; S = C or G. The PCR cycle used to obtain data regarding annealing temperatures was composed of one initial denaturating step of 5 min at 95º C, thirty cycles of 95º C for 15 sec, gradient annealing for 15 sec and 72º C for 30 sec, and a final elongation step at 72º C for 5 min. The appearance of a clear band in standard gel electrophoresis was used as an indicator of specificity. * Primers targeting Actinobacteria were purposely designed to be specific at 61.5º C as all other primer pairs worked best at this temperature.

Supplementary Table S2 All primers tested for DGGE profiling of taxon-specific qPCR products

Target taxon	Pair name	Forward primer	Position*	Reverse primer	Position*	Gradient span
α-Proteobacteria	αGC	CIAGTGTAGAGGTGAAATT	682	# CCCCGTCAATTCCTTTGAGTT	908	30-60
α-Proteobacteria	αGC1	# AGTGTAGAGGTGAAATTC	683	CCCCGTCAATTCCTTTGAGTT	908	30-60
α-Proteobacteria	αGC2	# CIAGTGTAGAGGTGAAATTC	683	CCCCGTCAATTCCTTTGAGTT	908	30-60
α-Proteobacteria	αGC3	CAAACAGGATTAGATAC	798	# CCCCGTCAATTCCTTTGAGTT	908	35-55
α-Proteobacteria	**αGC4**	TTAGATACCCTGGTAGTCCA	810	# CCCCGTCAATTCCTTTGAGTT	908	25-55
α-Proteobacteria	αGC5	# AGTGTAGAGGTGAAATTC	683	# CTACCAGGGTATCTAATCCT	787	25-55
μ-Proteobacteria	μGC	# TCGTCAGCTCGTGTYGTGA	1080	CGTAAGGGCCATGAT	1203	30-60
μ-Proteobacteria	μGC1	ATGTTGGGTTAAGTCCCG	1099	# CGTAAGGGCCATGATG	1202	35-60
μ-Proteobacteria	**μGC2**	ATGTTGGGTTAAGTCCCG	1099	# CCATGATGACTTGACGTC	1192	35-60

Supplementary Table S2 (continued)

Target taxon	Pair name	Forward primer	Position*	Reverse primer	Position*	Gradient span
Bacteroidetes	cfbGC	CRAACAGGATTAGATACCCT	798	GGTAAGGTTCCTCGCGTAT	967	25-50
Bacteroidetes	cfbGC1	AACAGGATTAGATACCCTG	799	# TAAGGTTCCTCGCGTAT	967	25-45
Bacteroidetes	**cfbGC2**	AACAGGATTAGATACCCTG	799	# TTCCTCGCGTATCATCGAAT	959	25-45
Firmicutes	**firmGC**	TGAAACTYAAAGGAATTGACG	928	# ACCATGCACCACCTGTC	1040	35-55
Firmicutes	firmGC1	AAAGGAATTGACGGG	930	# ACCATGCACCACCTGTC	1040	35-55
Firmicutes	firmGC2	# AAAGGAATTGACGGG	930	ACCATGCACCACCTGTC	1040	35-55
Actinobacteria	actGC	TACGGCCGCAAGGCTA	920	# CGTCRTCCCCACCTTCCTCCG	1200	35-55
Actinobacteria	actGC1	# TACGGCCGCAAGGCTA	920	TCRTCCCCACCTTCCTCCG	1200	35-55
Actinobacteria	actGC2	TACGGCCGCAAGGCTA	920	# TCCCCACCTTCCTCCG	1200	35-55

Supplementary Table S2 (continued)

Target taxon	Pair name	Forward primer	Position*	Reverse primer	Position*	Gradient span
Actinobacteria	actGC3	TTAATTCGATGCAACGCG	976	# TCCCCACCTTCCTCCG	1200	25-50
Actinobacteria	**actGC4**	TTAATTCGATGCAACGCG	976	# AACATCTCACGACACGAGCT	1067	25-50
Actinobacteria	actGC5	# AGCTCGTGTCGTGAGATGTT	1086	TCRTCCCCACCTTCCTCCG	1200	25-60
Actinobacteria	actGC6	TACGGCCGCAAGGCTA	920	# AACATCTCACGACACGAGCT	1067	25-55

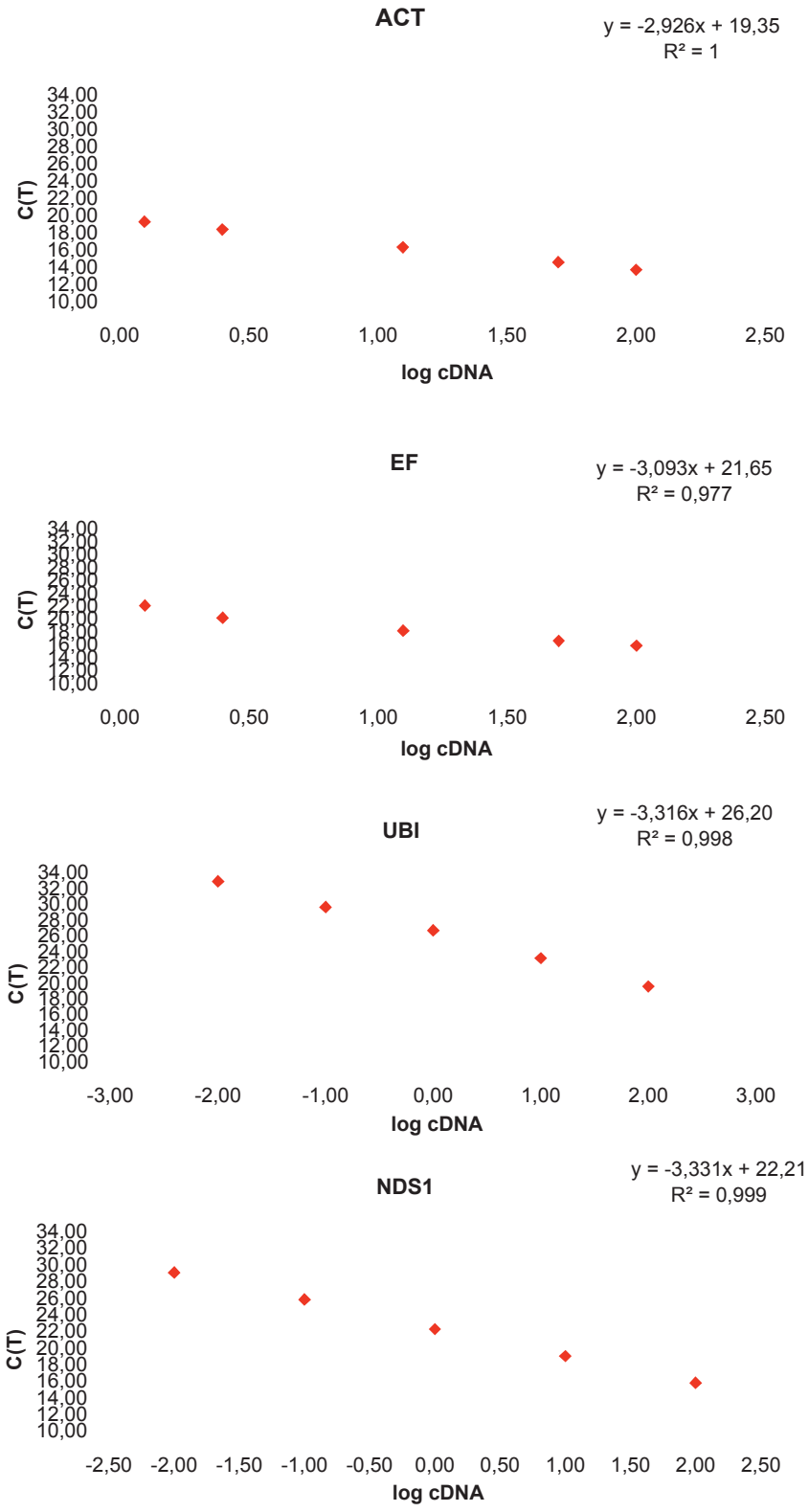
Primer pairs with bolded name were used to determine the DGGE profile of taxon-specific qPCR products. Gradient span: 100% denaturant is 7M urea plus 40% v/v formamide in 1X TAE buffer

* Position on the 16S rRNA gene as determined from alignment with the sequence from *Escherichia coli* ATCC 11775T

\# Primer containing the GC-clamp as in Muyzer *et al*. 1993

Supplementary Figure S1

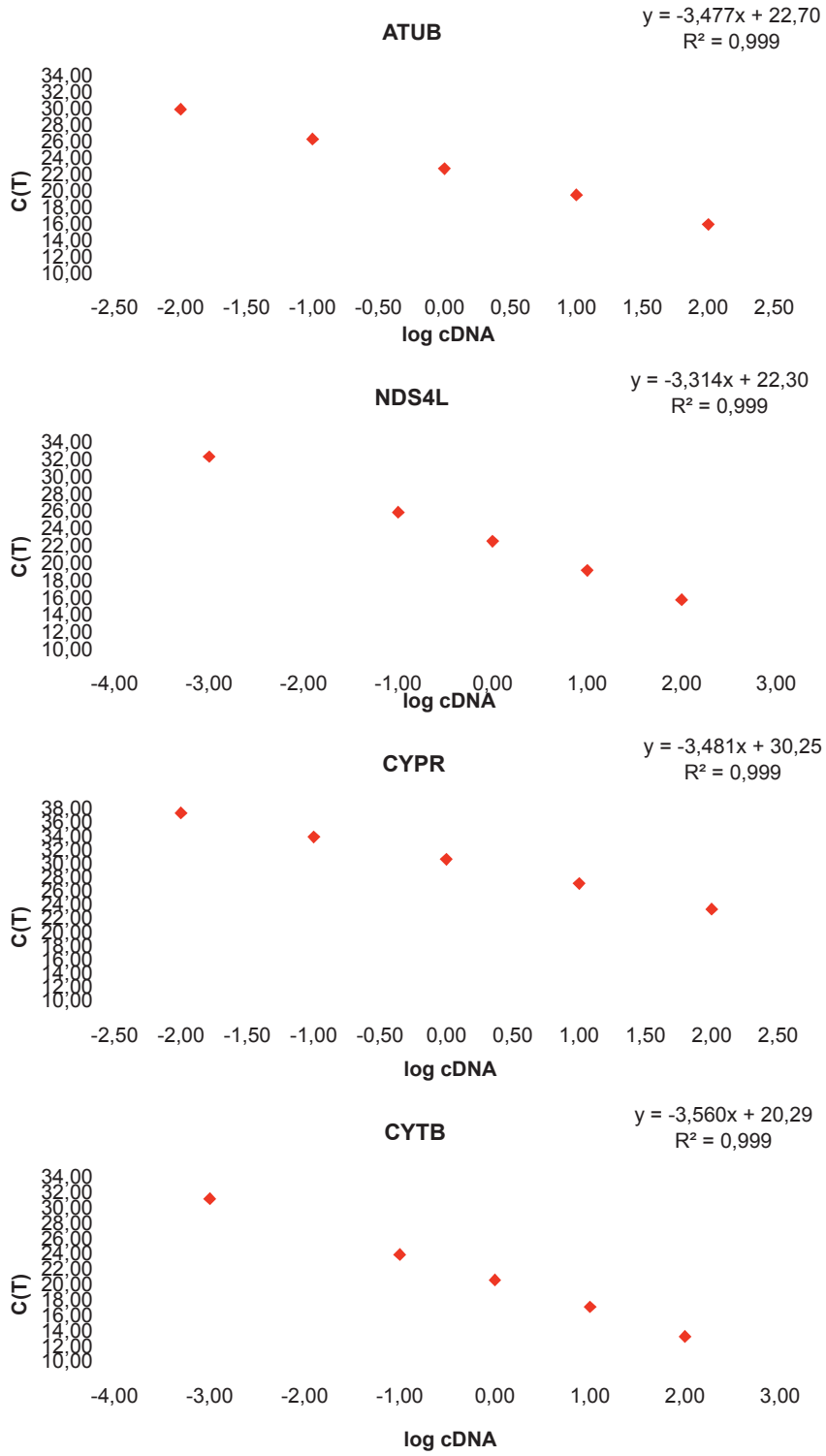
Efficiencies of primers amplifying candidate reference genes



continues

Supplementary Figure S1

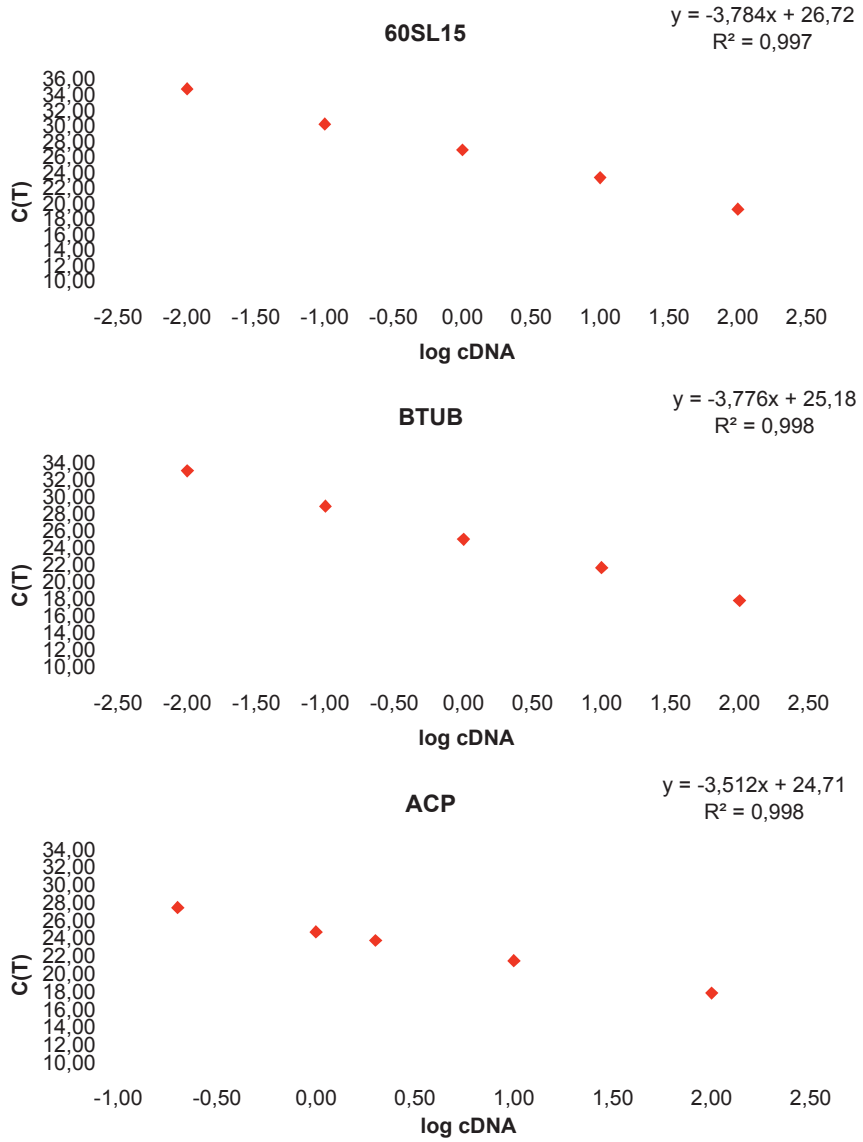
(continued)



continues

Supplementary Figure S1

(continued)

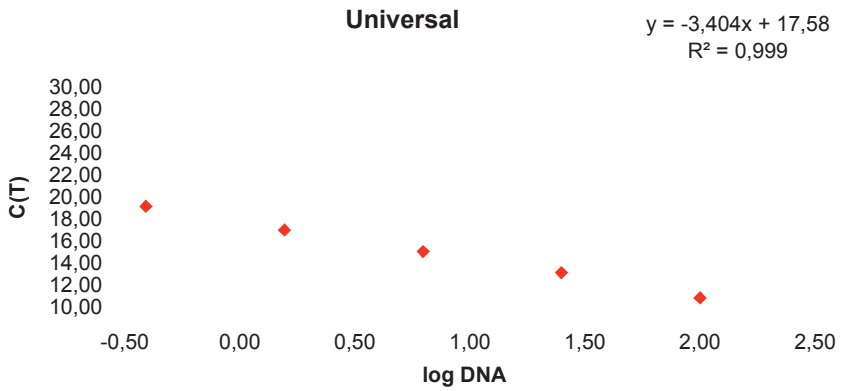


The amplification efficiency of each primer pair was determined by standard procedure by making dilution series of target DNA, calculating a linear regression based on the Ct data points, and inferring the efficiency from the slope of the line. R^2 = Correlation between data points. ACT = actin; EF = elongation factor; UBI = ubiquitin; NDS1 = NADH dehydrogenase subunit 1; ATUB = Tubulin A; NDS4L NADH dehydrogenase subunit 4L; CYPR = Cysteine protease 1; CYTB = cytochrome b; 60SL15 = 60s ribosomal protein L15, BTUB = tubulin B; ACP = acyl carrier.

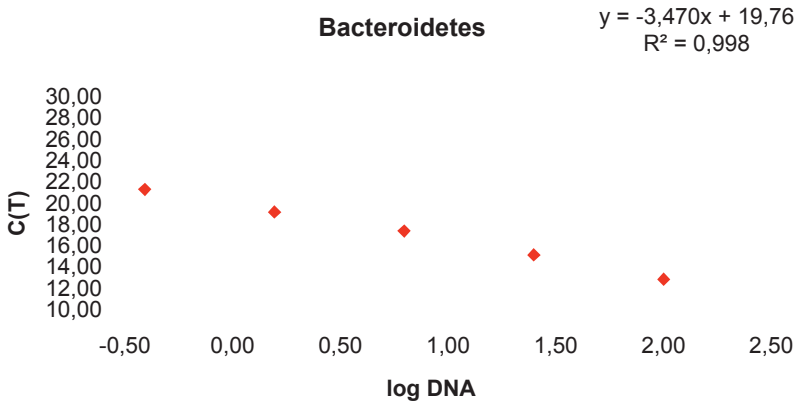
Supplementary Figure S2

Efficiencies of primers used in the taxon-specific qPCR assay

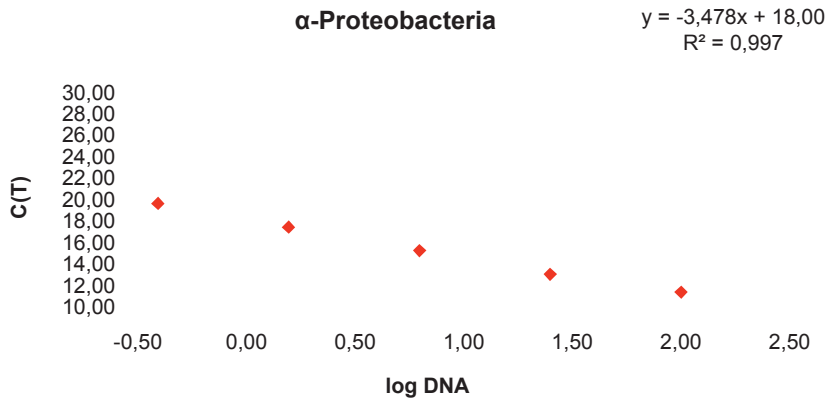
Universal: 907F + 1062R efficiency: 96.7%



Bacteroidetes: 798cfbF + cfb967R efficiency: 94.1%



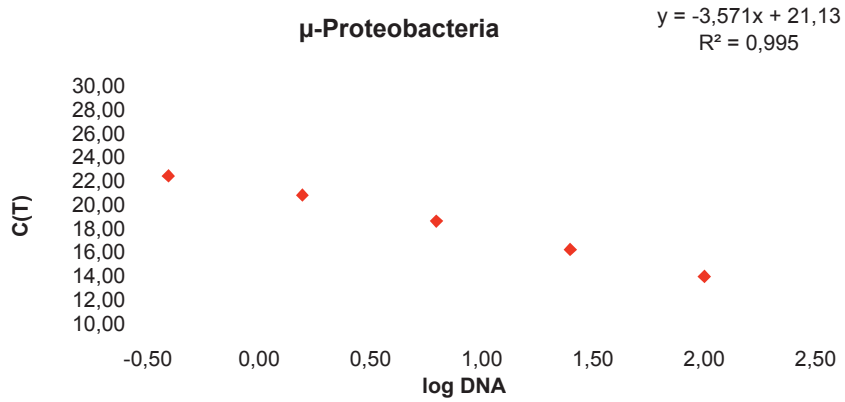
α-Proteobacteria: α682F + 908αR efficiency: 93.9%



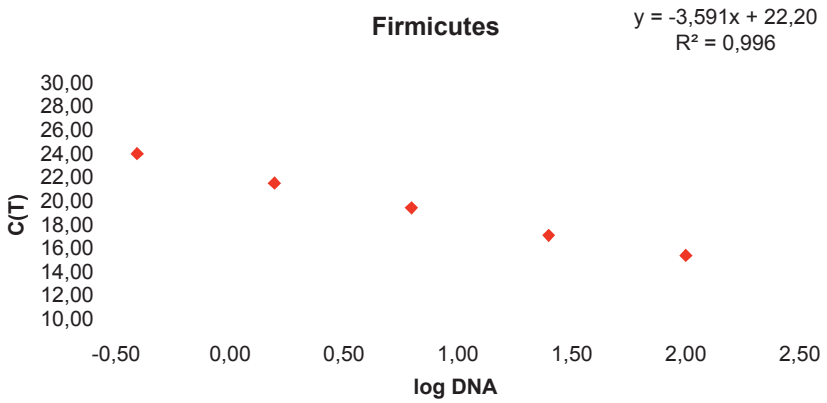
Supplementary Figure S2

(continued)

µ-Proteobacteria: 1080µF + µ1202R efficiency: 90.6 %



Firmicutes: 928firmF + 1040firmR efficiency: 89.9%



Actinobacteria: act920F3 + act1200R efficiency: 100%