Abundance and function of bacteria associated with the sponge *Crambe crambe*

Bart Verwaaijen

October 2010- May 2011

WAGENINGEN UR
For quality of life

qPCR of a *β-proteobacterium*, working towards the bacterial metagenome of *Crambe crambe*

Bart Verwaaijen
Supervisors: Dr. ir. D. Sipkema
Prof. Dr. H. Smidt

June 2011- December 2011
Laboratory of Microbiology
Wageningen University
The Netherlands

Front-page illustration: *C. crambe*,

Courtesy ; Miquel Pontes, marenostrum.org

Abstract

In previous research conducted by De Jaeger (2010) a highly abundant symbiotic *β-proteobacterium* was found within the sponge *Crambe crambe*. During this research we aim to retrieve the genomic sequence of this sponge-specific bacterium. The chosen method for sequencing is Illumina next generation sequencing of sponge samples enriched for bacteria by differential centrifugation.

Based upon the qPCR primers previously demonstrated by Aanstoot (2010) and Valev (2010) the most ideal sample for sequencing, containing a relative high concentration of bacterial DNA versus sponge DNA, was selected. This sample has been amplified by means of full genomic amplification and was submitted to Baseclear for sequencing.

The qPCR reactions have been standardized with a calibration plasmid for each of the primer sets and fine tuning of the thermocycler programs.

Furthermore it was discovered that the Rickettsia-like bacterium previously found by de Jaeger (2010) is present in all the presently available *C. crambe* samples.

In future research further analysis of the metagenomic data supported by additional 'wet' lab work is needed.

Preface

During my Msc. Biotechnology I decided to specialize myself in the Marine direction. This discussion was driven by my idea that this specialization would provide an ideal balance between engineering and 'wet lab' laboratory work. Within this specialization there were two compulsory courses that stood out for me, Marine biotechnology and advanced biotechnology. In the latter, two lectures particularity interested me. The lecture of Detmer Sipkema about microbial diversity in marine environments and a lecture of Klaske Schippers about her research on sponge cultivation. What appealed to me was the fundamental nature of this research in contrast to for example algae research that is more directed towards applications. There are still a lot of gaps in our fundamental knowledge about sponges and they are very reluctant to give up their secrets.

For my first thesis I approached Klaske Schippers. Within that thesis I tried to optimize some of the cultivation parameters of sponge tissue cultivation and to determine the effects of these parameters on contamination. Furthermore I tried to improve the flowcytometry protocol that was used to measure the viability of sponge cells. Because of these experiments I was in constant need of fresh sponge material. Every two weeks a diving trip was scheduled and on one of these trips Detmer participated. During the long car drive he passionately described his ongoing research on sponge symbionts.

As this was going to be my fifth internship/thesis I did not wanted to start from scratch on a whole new subject again. The fact that I already knew some about sponge research combined with Detmers enthusiasm made me decide for a thesis under his supervision.

I really enjoyed my time at Microbiology. The friendly environment both in the labs and the canteens is very comforting and allows for comfortable working. Of course I would like to thank Detmer for supervising me, apart from him I would like to specially mention some people; Hauke, Phillipe, Hans, Naim and Kyle for their help, Jasper and Matthijn for the many many coffee breakes.

With regards,

Bart Verwaaijen

Contents

1 Introduction

Sponges belong to the phylum Porifera and can be classified as sessile benthic filter-feeding invertebrate animals (Koopmans, Martens et al. 2009). Sponges are found in benthic habitats all over the globe in both sea and inland waters. People have been using sponges throughout history mainly for the absorbent properties of its skeleton. Sponges are the oldest animals still present to date. Li *et al* (1998) described the discovery of a 580 million year old sponge fossil.

Most of the sponge species, with the exception of some carnivorous species (Vacelet and Duport 2004), are sessile for the duration of their adult life. As a result of this sessile lifestyle they lack the ability to move away from predation or changing environmental conditions and therefore sponges developed a wide array of bioactive compounds during evolution. These make them responsible for producing up till 30% of all the known marine natural products (Faulkner 2000; Sipkema, Franssen et al. 2005). Some of which have antibiotic, antitumor, anti-inflammatory, antiviral, immune suppressive, antifouling or antimalarial properties (Taylor, Radax et al. 2007; Wright, McCluskey et al. 2011). Being such an untapped resource for novel medicines, sponges gained the interest of the pharmaceutical industry.

To date the application of sponge natural bioactive compounds is mainly hampered by the unavailability of sponge biomass. Sponges in general have very low growth rates and are notoriously difficult to culture. Therefore several researchers are aiming for *ex situ* cultivation of sponge tissue or sponge cell cultivation. One of the difficulties encountered during cultivation experiments is the high abundance of (symbioctic-) microorganisms within the sponge body. Amounts as high as 60% relative tot the total sponge biomass have been reported (Wilkinson 1978). Microorganisms found within sponges include; archea, heterotrophic bacteria, cyanobacteria, red and green algae(Taylor, Radax et al. 2007). These symbionts make it increasingly difficult to obtain a primary sponge cell line (Schippers, Martens et al. 2011). Furthermore in most cases it is not known if the sponge can survive without these 'commensal' inhabitants.

During the last few decades researchers have become more and more aware that a lot of formerly sponge associated bioactive compounds are actually being produced by members of the microbial community hosted within sponge (Taylor, Radax et al. 2007). An example is the neuroactive compound that was found in *Halicondria panicea* (Perovic, Wichels et al. 1998), for which further research pointed out it is produced by the symbiotic bacterial species *Psychroserpens burtonensis*. As a result of these discoveries, a shift of focus can be observed from sponge research towards the sponge associated bacteria (Lee, Lee et al. 2001).

In many cases tough it was discovered that only 0.1 % of the total diversity of microbes within sponges could be cultivated *in vitro* (Webster and Hill 2001). Several reasons for this inability to cultivate sponge-associated microbes have been debated. First of all sponges present their symbionts

with a microenvironment sheltered from the surrounding sea. Within this environment nutrient concentrations are higher, the mesohyl provides an substrate for surface attachment and so helps retention and protects from predation. It is therefore likely that these symbionts are dependent on these specific conditions, which are difficult to mimic *in vitro*. Furthermore it is likely that the sponge associated microbes partake in the nutritional process and thus benefit the sponge. Examples could be; intracellular digestion, nitrogen fixation, nitrification and photosynthesis (Wilkinson, Garrone et al. 1980; Wilkinson, Nowak et al. 1981). As long as it is unclear who produces the different metabolites in the mesohyl it is difficult to provide a symbiotic microorganism with the medium it requires. In addition it must be taken into account that a sponge usually maintains a very high flow trough of seawater and by doing so removes waste products from associated microorganisms very fast from its system. Therefore it is possible that these microorganisms do not grow under standard laboratory conditions because toxic metabolites are not dispersed fast enough, although this is purely hypothetical.

Of particular interest are the different symbiotic relations between the sponge-associated microbes and their hosts (Thomas, Rusch et al. 2010). As discussed previously, sponges are the oldest animals still in existence. Based upon that, one can suggests that these were one of the first multicellular organisms to establish symbiotic relations with prokaryotes. Even older examples of symbiosis between multicellular animals and prokaryotes may have existed but did not leave physical traces. Future discoveries on the mechanisms behind these relations might be extrapolated to give insight in other symbiotic relations and their development during evolution. A highly investigated field at this moment for example is the symbiosis between humans and gut micro biota. Although a lot of knowledge is being gained, sponges and their symbiotic relationships might provide a much less elaborate model to study the evolution of symbiosis as compared to the highly complex system of human gut.

1.1 Goal

Within this project the focus lies on the sponge *C. crambe* and its bacterial symbionts. This sponge was first described by Smidt in 1862 and can be readily found in the Mediterranean Sea. It is an orange to red encrusting sponge and sports a low microbial diversity. This low diversity makes this sponge an ideal candidate for different omics approaches. Specifically this research will target the most abundant bacterium, a sponge-specific *β-proteobacterium*. The main goal of this project is to obtain the genomic sequence of this *β-proteobacterium* in order to obtain insight in its genomic potential. Based upon this genetic makeup a list of putative enzymes with the corresponding substrates and products can be created. These will help to understand the mechanism behind its symbiosis with *C. crambe* and could be extrapolated to other sponges and their symbionts. In section 1.2 a summary of the preceding research conducted by Stefan Aanstoot and Lenny de Jaeger can be found

1.2 Preceding research

De Jaeger, 2010 started the research by attempting to cultivate the bacterial fraction found in the *C. crambe* sponge. He succeeded to obtain 16s rRNA sequences 10 different isolates. A 16S rRNA gene clone library experiment showed the targeted *β-proteobacterium* to be the dominant species and the known 16S rRNA gene sequence was extended to a total of 1508 base pairs. Furthermore it was shown that the nearest neighbor found within sponge related bacteria has only 92% sequence identity based on the 16S rRNA gene compared to this *β-proteobacterium*.

Aanstoot, 2010 used differential centrifugation in order to enrich the sponge cell fractions for bacterial cells. This was done in two steps starting with a 1500g centrifugation followed by a 16000g fraction, the former was called bacterial fraction and the latter 16K. The original fraction was called the sponge fraction. For a more comprehensive description of the procedure see, Aanstoot, 2010. Furthermore Aanstoot demonstrated the design and practical application of the *β-proteobacterium* specific 16S set of (q)PCR primers. These primers were used to determine the relative sponge and bacterial abundances within the different fractions. It was found that the bacterial versus eukaryotic DNA ratio increased during the two enrichment stages although the actual *β-proteobacterium* DNA concentration decreased due to the procedure.

2 Materials and Methods

2.1 Media and solutions

LB-agar medium; 10 g bacto tryptone, 5 g Bacto yeast, 5 g NaCL and 15 g agar dissolved in 1000 g tap water. Adjust to pH 7.0 with NaOH and autoclave.
LB liquid medium; 10 g bacto tryptone, 5 g Bacto yeast, 5 g NaCL and dissolved in 1000 g tap water. Adjust to pH 7.0 with NaOH and autoclave.
IPTG stock; Dissolve 24 mg in 1ml dH2O, sterilize through a 0.45 µm filter and store at -20 ºC
X-gal stock Dissolve 50 mg in 1ml N,N'-dimethylformamide and store at -20 ºC
Ampiciline stock; Dissolve 20 mg in 1ml dH2O sterilize through a 0.45 µm filter and store at -20 ºC
LB-agar Amp, X-gal, IPTG plates; Add per 100 ml LB medium, 200 µl X-gal, 160 µl IPTG and 200 µl Ampiciline stock solution.
SOB medium; 2 g tryptone, 0.5 g yeast extract, 0.05 g NaCl dissolved in 100g tap water. Autoclave and subsequently add 1 ml filter sterilized 1M $MgCl_2$ and 1 ml filter sterilized 1M $MgSO_4$.
SOC medium; Directly before use supplement 1.5 ml of SOB medium with 30 µl 20% w/v filter sterilized glucose.

2.2 Transformation

50 µl of XL-1 blue supercompetent cells (Stratagene) were placed in pre-chilled eppendorf tubes on ice and allowed to thaw. After thawing 0.85 µl β-mercaptoethanol was added, the tubes were gently swirled and incubated for 10 minutes on ice again. 3 µl of ligation mixture (mixture of template DNA and Promega pGEM-T easy vector, stored at 5 °C over night) was added (a higher volume might be applied if difficult constructs are anticipated), the tubes were gently swirled and incubated for another 30 minutes on ice. The cells were shocked for 45 seconds in a 42 °C water bath and incubated for 2 minutes on ice. The cells were supplemented with 450 µl of SOC medium preheated to 37 °C and then incubated 1 hour at 37 °C while shaking at 200 rpm. 10 µl and 100 µl of the culture suspension was plated on separate LB-Agar+Amp+X-gal+IPTG plates and incubated over night at 37 °C. Five separate colonies were marked on the plates and transferred with a sterile toothpick to a PCR tube, the toothpick was dropped in a 25 ml screw cap tube afterwards. 7 ml of LB+Amp medium was added to the tubes and these were incubated over night at 37 °C while shaking. Colony PCR was performed on the previously 'inoculated' PCR tubes with the corresponding primers and product sizes were checked by means of gel electrophoresis. A strain possessing the right product size was selected and 5 ml of suspension was pelleted in 2 ml tubes by centrifuging 5 minutes at 11000 RCF. Plasmids were isolated according to the Gene Jet Plasmid MiniPrep (Fermentas) kit protocol. Glycerol stocks were made by

adding 600 µl of the cloned culture and 400 µl 100% sterile glycerol into cryo-vials. After mixing these were immediately frozen at -80 °C.

2.3 qPCR calibration plasmids

Partial DNA sequences of *C. crambe* 18S, *β-proteobacterium* 16S and *Rickettsia*16S were ligated into the pGEM-t-easy plasmid and cloned into E.coli XL-1 Blue. The different inserts were checked by means of PCR and gel electrophoresis, not depicted in this report. The clones were cultivated and plasmids were isolated. These plasmids were then diluted to a concentration of 10 ng/ µl and alloquited. For each qPCR reaction that was done from this point on a fresh alloquit of one of the calibration plasmids was used to make the calibration curve. The different plasmids were named pGEM-t-easy-Cr18S, pGEM-t-easy-Betaproteo16S and pGEM-t-easy-*Rickettsia*16S. Glycerol stocks of these clones were cryopreserved at -80˚ C for future research.

2.4 qPCR

For each qPCR primer set a specific plasmid was constructed for calibration purposes, see section 2.2. The calibration curve ranged from 10 ng/ µl to 10 fg/ µl and was applied in triplicate to each reaction plate. The calibration curve was also used a a positive control, nuclease free water (Promega) was used as a negative control. A No Template Control (NTC) has not been used. 2 µl of each sample was applied to the reaction in a triplicates of 10, 100 and 1000 fold dilutions. The plates filled with template not containing reaction mixture or calibration curves yet were sealed and stored at -20 ˚C.
qPCR reactions were performed with a Bio-rad CFX-96 machine controlled by Bio-rad CFX Manager 2.1 software. Each qPCR mix contained 5 µl of SYBRgreen mastermix (Qiagen), forward primer and a reversed primer(primer specific concentrations), the reactions were topped up with Nuclease free water (Promega) to a total volume of 8 µl per reaction.
Each program was started with a 5 minutes step at 95 ˚C to denaturate the template and activate the polymerase. This step was followed by 40 cycles consisting of; 30 seconds at 94 ˚C, 40 seconds at the primer specific annealing temperature and 90 seconds at 72 ˚C. Subsequently a step of 5 minutes at 72 ˚C was used. The program was finalized with a melting curve ranging from 65 ˚C to 95 ˚C in steps of 0.5 ˚C (Bustin, Beaulieu et al. 2010).

2.5 PCR Primers

Table 2.1: List of oligonucleotide primers and their optimal annealing temperatures as used in this Research for PCR.

Primer	Sequence	Annealing temperature	Reference
Bet809-F	5'AACGATGCCCGCTAGCTGTCG'3	58 °C	Aanstoot (2010)
Bet1010-R	5'TCTCCGACTCCGCGACAGG'3	58 °C	Aanstoot (2010)
EUK345-F	5'AAGGAAGGCAGCAGGCG'3	55 °C	(Zhu et al. 2005)
EUK499-R	5'CACCAGACTTGCCCTCYAAT'3	55 °C	(Zhu et al. 2005)
T7	5'-TAATACGACTCACTATAGGG-'3	52 °C	Promega
SP6	5'-ATTTAGGTGACACTCAAGC-'3	52 °C	Promega
1369F	5'-CGGTGAATACGTTCYCGG-'3	52 °C	(Suzuki et al. 2000)
1492R	5'-GGWTACCTTGTTACGACT-'3	52 °C	(Suzuki et al. 2000)
27F	5'-AGRGTTYGATYMTGGCTCAG-'3	52 °C	(Takai and Horikoshi 2000)
EUKF	5'- AACCTGGTTGATCCTGCC-'3	55 °C	(Medlin et al. 1988)
EUKR	5'-TGATCCTTCTGCAGGTTCACCTAC-'3	55 °C	(Medlin et al. 1988)

Table 2.2: List of oligonucleotide primers and their optimal annealing temperatures and mastermix concentrations as used in this research for qPCR.

Primer	Sequence	Primer C (mmol/ µl)	Annealing temperature	Reference
Bet809-F	5'AACGATGCCCGCTAGCTGTCG'3	0.4	58 °C	Aanstoot (2010)
Bet1010-R	5'TCTCCGACTCCGCGACAGG'3	0.4	58 °C	Aanstoot (2010)
EUK345-F	5'AAGGAAGGCAGCAGGCG'3	0.25	59.2 °C	(Zhu et al. 2005)
EUK499-R	5'CACCAGACTTGCCCTCYAAT'3	0.25	59.2 °C	(Zhu et al. 2005)
1369F	5'-CGGTGAATACGTTCYCGG-'3	0.25	53.8 °C	(Suzuki et al. 2000)
1492R	5'-GGWTACCTTGTTACGACT-'3	0.25	53.8 °C	(Suzuki et al. 2000)
Rick F	GGC GAC CCG GAG CAA ATC CTG AG	0.4	65 °C	Dimitar Valev
Rick R	TCT CCG ACT CCG CGA CAG G	0.4	65 °C	Dimitar Valev

2.6 Repli-g

The Qiagen REPLI-g Mini kit was used to synthesize the required amount of genomic DNA for Illumina paired end sequencing. 100x Sigma TE buffer was diluted with nuclease free water

(Promega) to a 1x and 1.67x concentration. 6 µl of template DNA was buffered with 4 µl 1.67x TE buffer. In the final reaction 1.25 µl, 2.5 µl, 5 µl and 0 µl of TE buffered DNA was used topped up to a final volume of 5 µl with 1x TE.

2.7 Sponge and bacterial cell fractionation

During this research sponge samples enriched for their bacterial content by Stefan Aanstoot were used. The following protocol was taken from his report and gives a description of the enrichment method that was used.

Samples containing C. crambe and bacterial cells were diluted 10 times in CMFASW containing 10 mM Ethylenediaminetetraacetic acid (EDTA). After gentle shaking for 1 hour to disassociate cell clusters the cell suspension was filtered through a 70 µm cell strainer. The suspension was then centrifuged at 1500 g for 15 minutes. The pellet was collected as the Sponge Fraction (SF). The supernatant was again centrifuged 15 minutes at 1500 g °C, the resulting cell pellet was discarded. Two more pellets were collected from the remaining supernatant. One pellet was collected at 4700 g for a total of 35 minutes and the other pellet was collected from the supernatant of this fraction at 16000 g for 1 min. These fractions are respectively called the Bacterial Fraction (BF) and 16000 G Fraction (16K). All centrifugation steps were performed at 4 °C.

DNA isolations were performed with the qBiogene fastDNA spin kit for Soil, DNA isolation kit.

The C.crambe samples Cr9 and Cr10 were collected separately by Oriol Sacristan in Blanes, Spain. The samples were crushed and diluted with Calcium magnesium free seawater (CMFSW) + EDTA and shaken for one hour. The suspension was filtered over a 70 µm filter prior to a 15 minutes centrifugation step at 1500g. The pellet was discarded and the supernatant was crypresered with a concentration of 17% glycerol. These steps were done at Blannes, Spain further processing of the samples took place at Wageningen, the Netherlands according to the previously described protocol.

3 Results

3.1 DNA isolation

DNA was isolated from the SF fractions of the *C. crambe* samples number 2, 4, 5, 6 and 7-8. The resulting DNA samples were placed on agarose gel and measured on Nanodrop (Thermo scientific) for validation of the quality and concentration. In figure 3.1 it can be seen that there are DNA smears for each of the samples indicating the presence of different sizes of DNA including high molecular weight DNA. In table 3.1 the Nanodrop measurements are depicted. The DNA concentrations are not very high with values between 10 and 30 ng/ µl respectively but are sufficient for the use in PCR and qPCR. Furthermore the samples were tested for the presence of *C. crambe* and *β-proteobacterium* DNA with the Bet809-F and Bet1010-R primers.

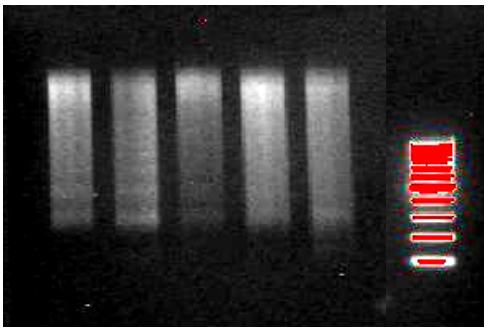

Figure 3.1: DNA isolation. 10 µl of total isolated DNA was placed on a 1% agarose gel. Cr2, Cr4, Cr5, Cr6, Cr7-8 and 5 µl of 1kb GeneRuler (Fermentas).

Concentrations,

Table 3.1: Nanodrop measured concentrations of DNA isolations from the samples Cr2, Cr4, Cr5, Cr6 and Cr7-8.

Sample	DNA concentration (ng/ µl)
Cr2	27.51
Cr4	23.40
Cr5	12.79
Cr6	20.88
Cr7-8	15.74

3.2 qPCR

qPCR reactions were performed on the *C. crambe* SF samples 2, 4, 5, 6, 7-8 with the universal 18S primer set and the universal 16S primer set. These series of reactions were done for several reasons. First of all to measure the relative 16S and 18S ratios within these samples. For the future

metagenomic sequencing it is important to have as much as possible bacterial DNA and as less as possible eukaryotic DNA. Secondly to get acquainted with the qPCR machine and the protocol.

The reactions were fine tuned for primer concentration and annealing temperature first. For each primer set three concentrations were used; 1, 2 and 4 mM respectively. A gradient of eight steps starting at 2˚ C under the calculated annealing temperature up to 2˚ C above this temperature was used, figure 3.2 depicts an example. For the universal 16S rRNA gene (abbreviated, 16S) primer set it was found that a primer concentration of 4 mM and an annealing temperature of 53.8˚ C is optimal. For the universal 18S rRNA gene (abbreviated, 18S) primer set it was found that a primer concentration of 2.5 mM and an annealing temperature of 59.2˚ C is optimal.

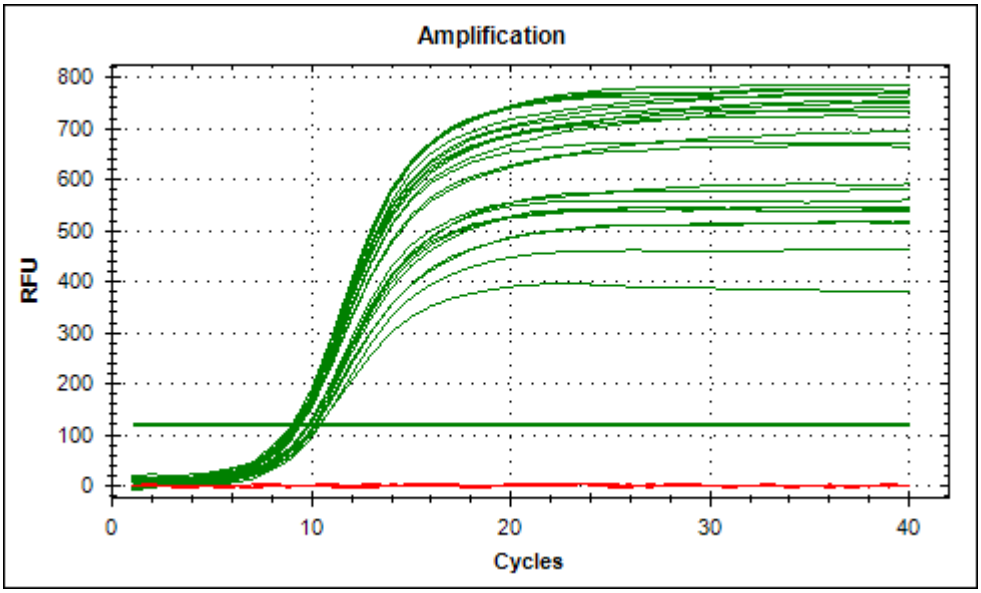

Figure 3.2: an example of a qPCR optimization reaction. Baseline subtracted relative fluorescentce units (RFU) plotted against the reaction cycles. With in green the sample curves and in red the Negative controle. The green line represends the RFU threshold at wich the thershold cycle is determined (Ct).

Figure 3.3 depicts an example of a calibration curve for the universal 18S primer set. Based on these calibration curves the efficiency of each corresponding qPCR reaction could be determined. Combining the reaction efficiency with the threshold cycle (Ct) of the samples the starting quantities (SQ) were calculated. Each reaction was ended with a melting curve as depicted in figure 3.5. From the melting curves it can be seen that no secondary or non specific products are formed that reach the threshold. Because the calibration curves for these reactions are based on *A. niger* and *Bacillus* genomic DNA respectively the calculated SQ's are only relative and are therefore not depicted in this report.

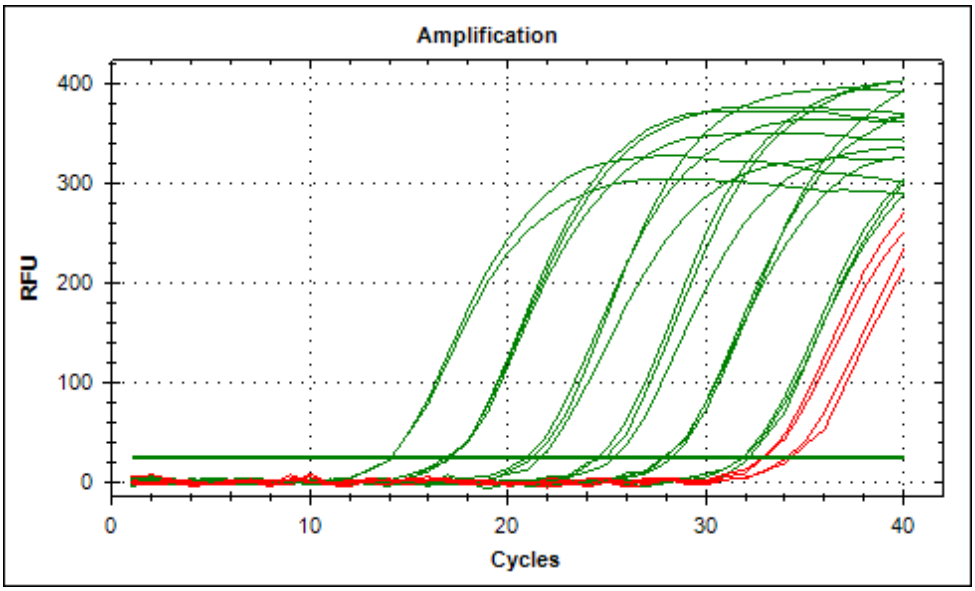

Figure 3.3: 18S qPCR calibration reaction based on *A. niger* genomic DNA. Baseline subtracted relative fluorescentce units (RFU) plotted against the reaction cycles. With in green the sample curves and in red the Negative controle. The green line represends the RFU threshold at wich the thershold cycle is determined (Ct).

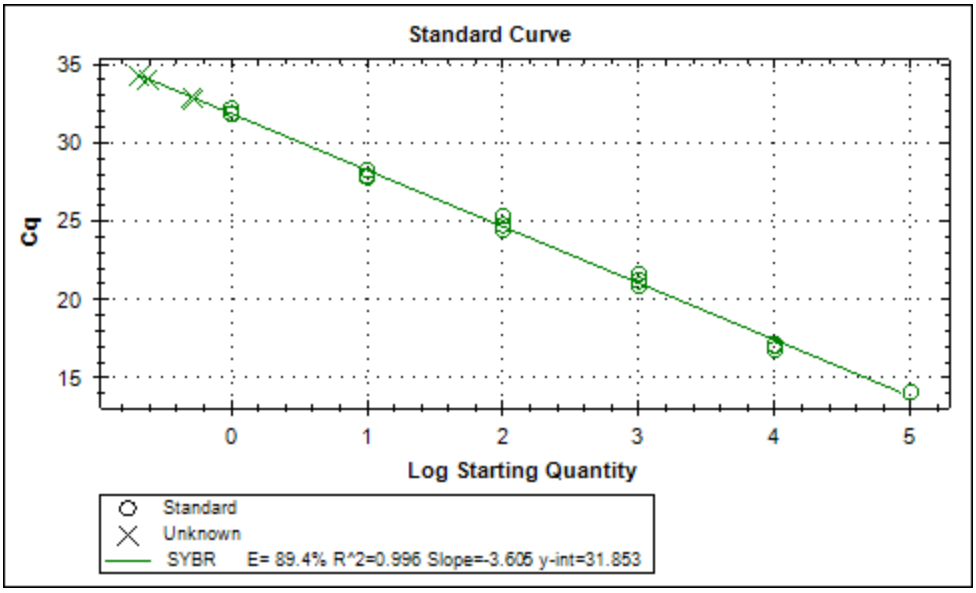

Figure 3.4: 18S qPCR calibration curve.

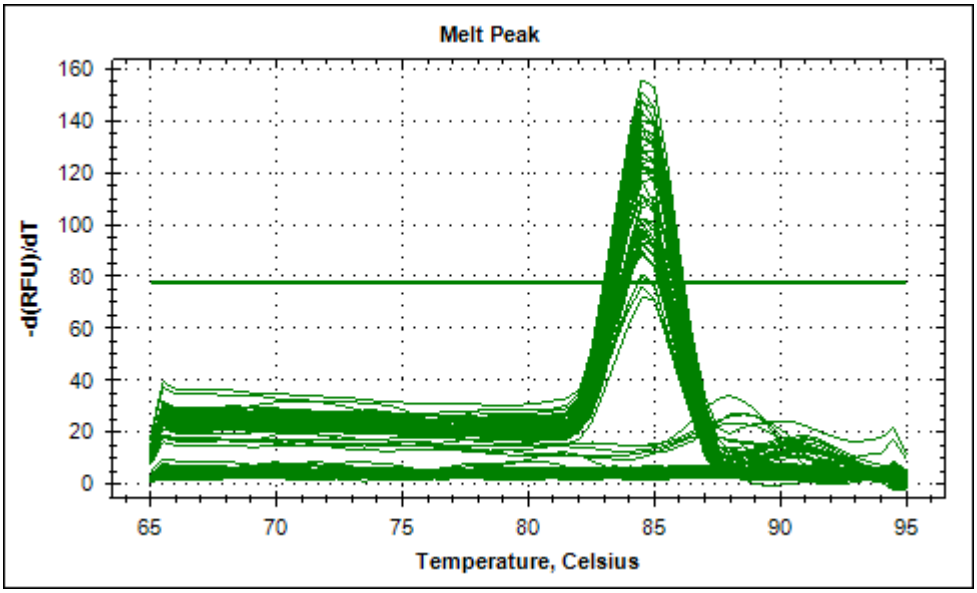

Figure 3.5: An example of a meltcurve.

By dividing the measured 18S concentrations trough the 16S concentrations, a ratio is calculated that can be used to indicate the size of the bacterial DNA portion within these specific samples, figure 3.6. The closer this value approaches zero the higher the bacterial DNA portion, which is favourable for metagenome sequencing.

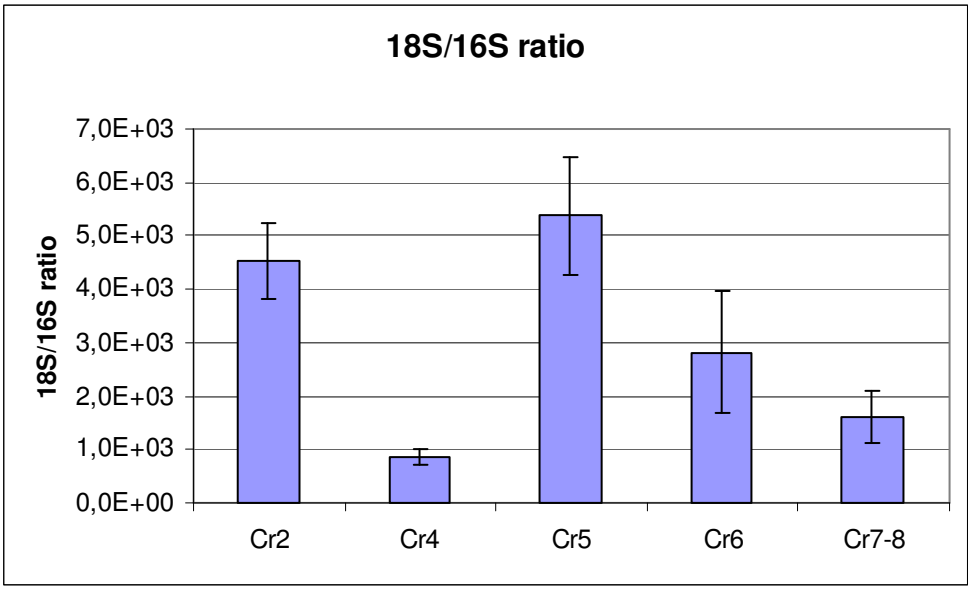

Figure 3.6: relative 18S/ 16S ratios as found in the the *C. crambe* SF samples 2, 4, 5, 6 and 7-8 by means of qPCR. Calibration curves were based on 10ng/ µl stock samples of *A. Niger* and *Bacillus* genomic DNA respectively. The error bars depict the STDEV over technical replicates.

In addition to the ratio between the bacterial and eukaryotic DNA the total amount of 16S and 18S DNA per sample was calculated, figure 3.7. These can be used to compare the total DNA concentrations between the samples within this data set.

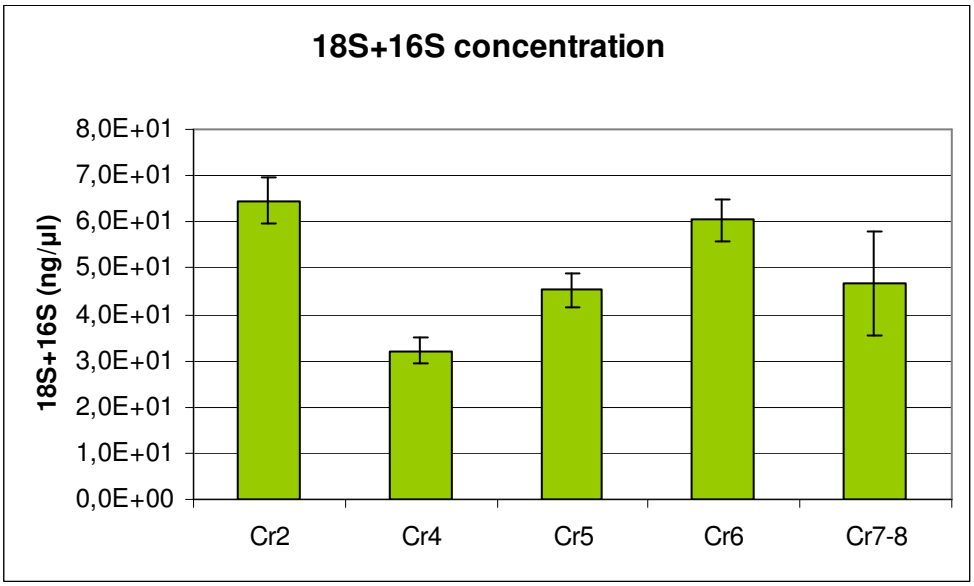

Figure 3.7: relative total amount of 18S + 16S as found in the the *C. crambe* SF samples 2, 4, 5, 6 and 7-8 by means of qPCR. Calibration curves were based on 10ng/ µl stock samples of *A. niger* and *Bacillus* genomic DNA respectively. The error bars depict the STDEV over technical replicates.

3.3 Sample selection (Illumina sequencing)

For the Illumina metagenome sequencing assay four samples were nominated. The original Cr1 BF and Cr1 16K made and described by Aanstoot 2010. Furthermore there were two fresh samples named Cr9 BF and Cr10 BF that were freshly imported from Spain by Oriol Sacristan and processed by dr. Detmer Sipkema. The 16S and 18S fractions were measured with the same method as previously described for the other samples. Figure 3.8 and figure 3.9 depict the 18S/ 16S ratio and the total 16S plus 18S respectively.

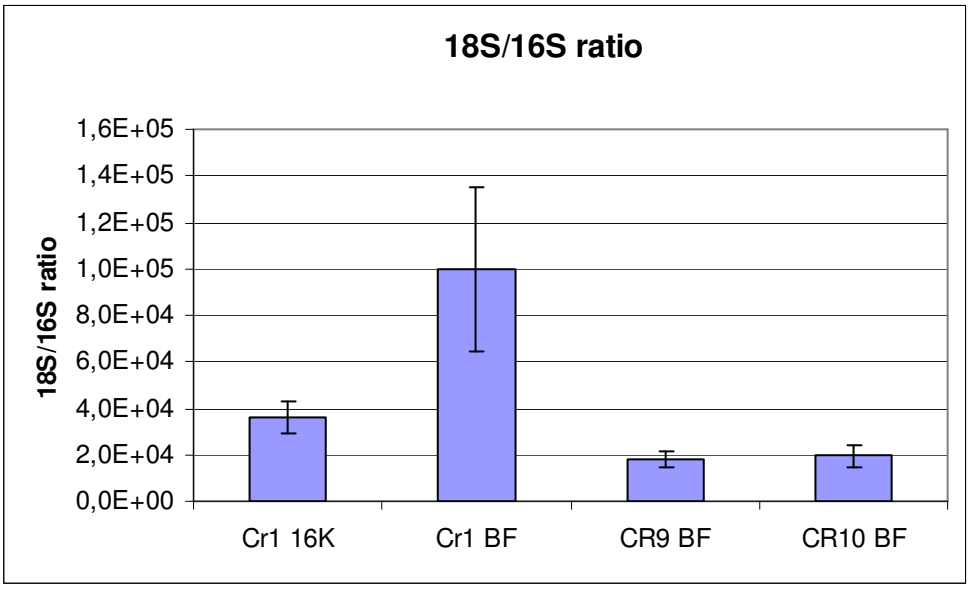

Figure 3.8: relative 18S/ 16S ratios as found in the the *C. crambe* samples Cr1 16K, Cr1 BF, Cr9 BF and Cr10 BF by means of qPCR. Calibration curves were based on 10ng/ µl stock samples of *A .Niger* and *Bacillus* genomic DNA respectively. The error bars depict the STDEV over technical replicates.

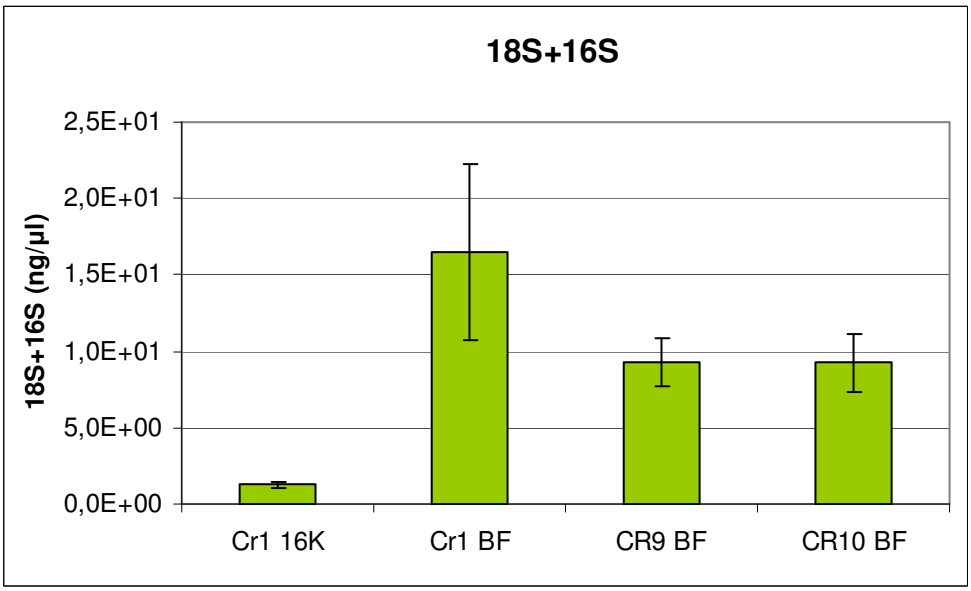

Figure 3.9: relative total amount of 18S + 16S as found in the the *C. crambe* samples Cr1 16K, Cr1 BF, Cr9 BF and Cr10 BF by means of qPCR. Calibration curves were based on 10ng/ µl stock samples of *A. Niger* and *Bacillus* genomic DNA respectively. The error bars depict the STDEV over technical replicates.

3.4 Repli-g amplification

Sample Cr9 BF was chosen for Illumina sequencing, but because of the low amount of biomass of this sample an amplification step of the DNA was necessary. A whole genome amplification was done based on the Qiagen Repli-g mini kit. Four different Cr9 BF template input amounts were used; 30 ng, 60 ng, 120 ng and a no template control. These were dubbed Cr9.1, Cr9.2, Cr9.3 and B respectively.

To test the yield of the Repli-g reactions 10 µl of each reaction mixture was placed on gel, depicted in figure 3.10. As can be seen in this picture there is a big DNA band exceeding the size of the 1 kb GeneRuler (Fermentas) plus some DNA still left in the well of each lane. Also note that the no template control of the Repli-g reaction, B also contains a similarly sized product.

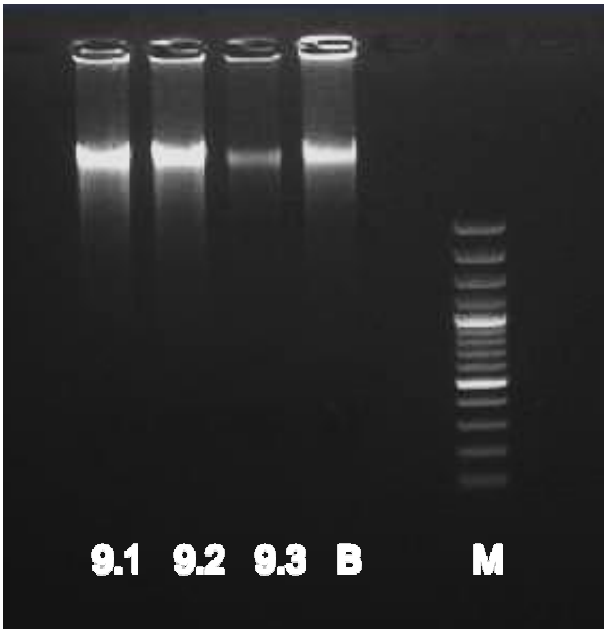

Figure 3.10: Gel electrophoresis of the Repli-g amplified DNA. From left to right 10 µL of Cr9.1, Cr9.2, Cr9.3 and B. As a marker 5 µl of 1kb GeneRuler (Fermentas).

The Repli-g products were tested for the presence of *β-proteobacterium* bacterium DNA by means of PCR with the Bet809-F and Bet1010-R primer set. The expected product size for this reaction was approximately 220 bp, figure 3.11.

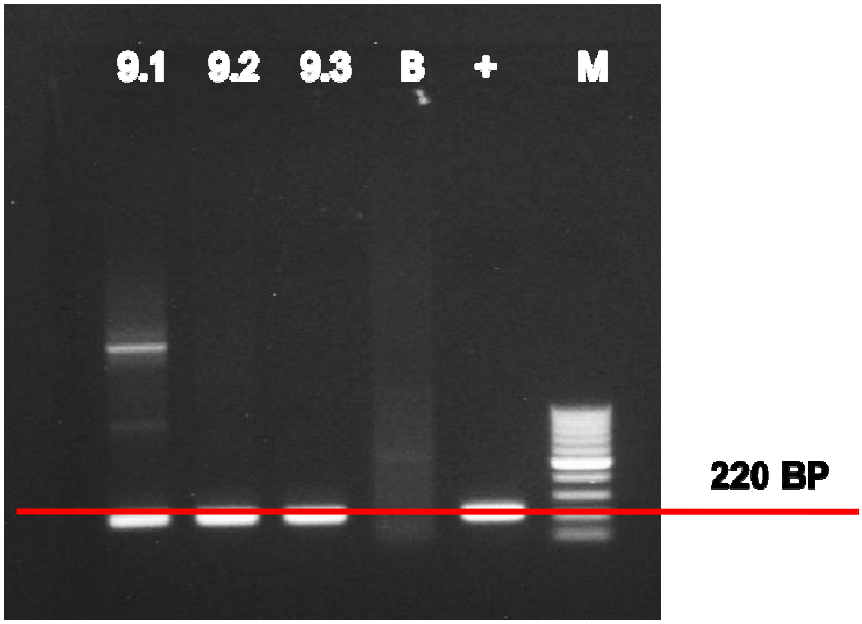

Figure 3.11: Gel electrophoresis of PCR products based on the *β-proteobacterium* 16S primer set with Repli-g amplified DNA as template. From left to right 10 µL product of Cr9.1, Cr9.2, Cr9.3, B and Cr9 (+). As a marker 5 µl of 1kb GeneRuler (Fermentas).

Based on qPCR of the 16S and 18S genes it was decided that Cr9.1 and Cr9.2 had the lowest and thus most favourable 18S/ 16S ratio. These data are not presented within this report because of the low quality, the calibration curves failed during the reaction.

For Illumina next generation sequencing Baseclear (Leiden, NL) demanded at least 10 µg of good quality DNA (260/ 280nm between 1.8 and 2.0). Because they deemed fluorescent and spectroscopy based measurements inaccurate a 10x dilution of Cr9.1 and Cr9.2 product was ran on gel. The weight of the different bands in the 1kb+ GeneRuler (Fermentas) were used to estimate the DNA concentrations, depicted in figure 3.12. Both the Cr9.1 and Cr9.2 were estimated to have at least 160 ng/ µl. The 260/ 280nm ratios were measured on a Nanodrop machine to be 1.73 and 1.77 for Cr9.1 and Cr9.2 respectively and thus approached the 1.8 ratio for pure DNA. The Cr9.1 and Cr9.2 samples were pooled and send off to Baseclear for Illumina paired end metagenomic sequencing.

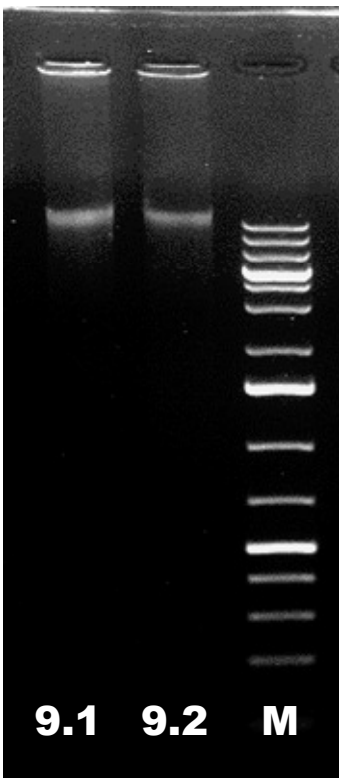

Figure 3.12: 10µL of 10x diluted Cr9.1 supplemented wit 3 µL loading dye of wich 10 µL was put on gel. Lane2; 10µL of 10x diluted Cr9.2 supplemented wit 3 µL loading dye of wich 10 µL was put on gel. Lane3; 5µL of Fermentas gene ruler 1kb plus @ 0.1 µg/ µL

3.5 Optimized qPCR reactions

For the samples Cr9 SF, Cr1 16K, Cr1 BF, Cr1 SF, Cr9 BF, Cr9.1/9.2 and B a new series of qPCR reactions was performed, calibrated with the previously described plasmids. In figure 3.13 the 18S/ 16S ratio is depicted and in figure 3.14 the total amount of 18S + 16S. Special attention in this measurement goes out to the samples Cr9 SF, Cr9 BF and Cr9.1/9.2 as these depict the different enrichment and amplification steps that are used for the original Cr9 sponge sample.

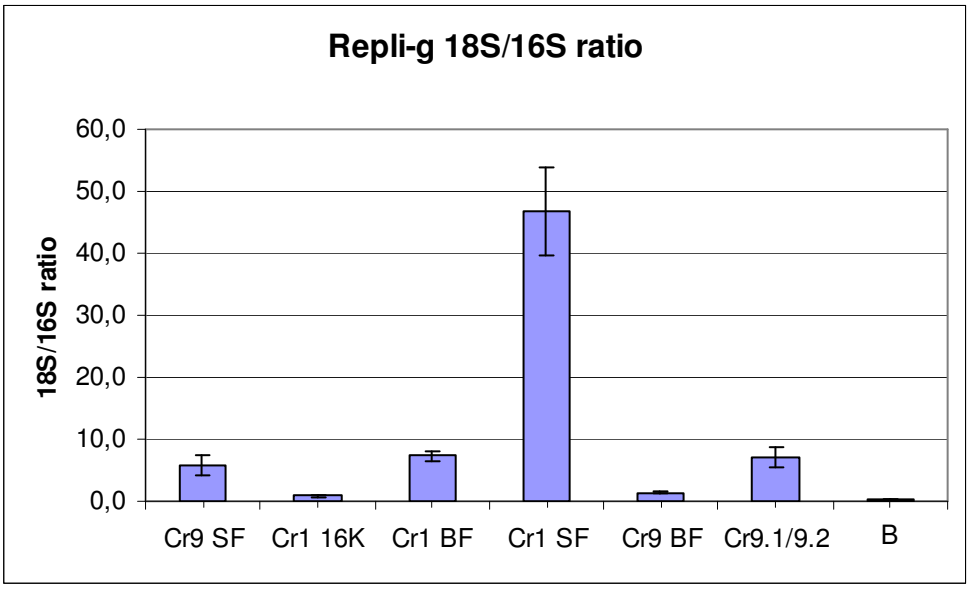

Figure 3.13: relative 18S/ 16S ratios as found in the the *C. crambe* samples Cr9 SF, Cr1 16K, Cr1 BF, Cr1 SF, Cr9 BF, Cr9.1/9.2 and Cr9.4 by means of qPCR. Calibration curves were based on 10ng/ µl stock samples of pGEM-t-easy-Cr18S and pGEM-t-easy-16S. The error bars depict the STDEV over technical replicates.

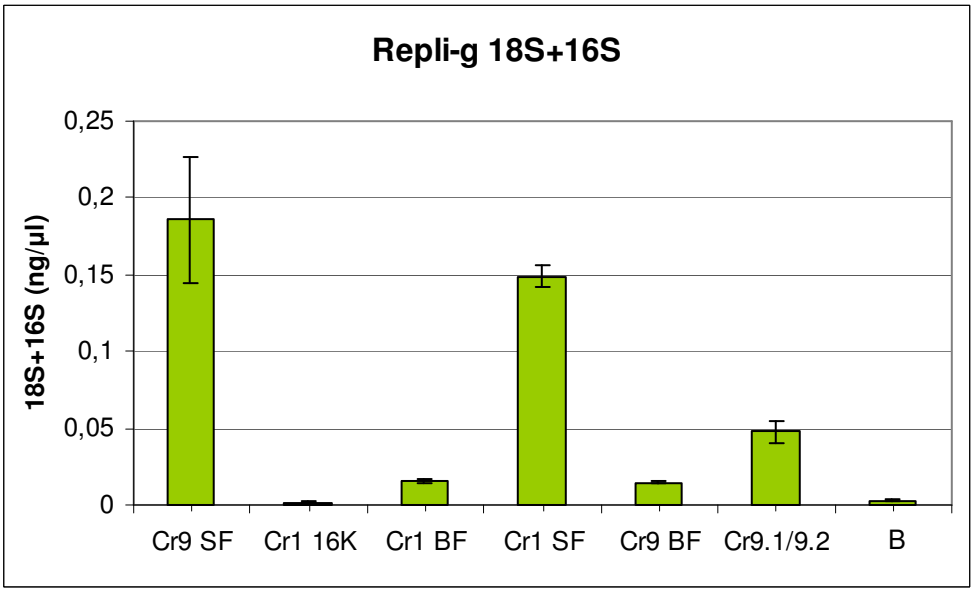

Figure 3.14: relative total amount of 18S + 16S as found in the the *C. crambe* samples Cr9 SF, Cr1 16K, Cr1 BF, Cr1 SF, Cr9 BF, Cr9.1/9.2 and Cr9.4 by means of qPCR. Calibration curves were based on 10ng/ µl stock samples of pGEM-t-easy-Cr18S and pGEM-t-easy-16S. The error bars depict the STDEV over technical replicates.

In addition to these also the specific *β-proteobacterium* 16S rRNA gene and *Rickettsia* 16S rRNA gene DNA fractions of these samples were measured. A complete overview of all measured fractions

of these samples is shown in table 3.1. The 16S *Rickettsia* DNA concentrations are depicted in red because their concentrations exceeded the used calibration curve.

Table 3.2: relative total amount of 18S, 16S, 16S *β-proteobacterium* and 16S *Rickettsia* DNA as found in the *C. crambe* samples Cr9 SF, Cr1 16K, Cr1 BF, Cr1 SF, Cr9 BF, Cr9.1/9.2 and Cr9.4 by means of qPCR. Calibration curves were based on 10ng/ µl stock samples of pGEM-t-easy-Cr18S, pGEM-t-easy-Betaproteo16S and pGEM-t-easy-*Rickettsia*16S. The given values are averages of triplicate measurements.

Sample	18S (ng/ µl)	16S (ng/ µl)	16S beta (ng/ µl)	16S Rick (ng/ µl)
Cr9 SF	1,6E-01	2,7E-02	2,4E-02	3E-06
Cr1 16K	9E-04	1,0E-03	2,3E-04	4,1E-07
Cr1 BF	1,4E-02	1,9E-03	1,7E-03	6E-06
Cr1 SF	1,5E-01	3,1E-03	3,2E-03	4E-06
Cr9 BF	9E-03	6E-03	8E-03	2,6E-06
Cr9.1/9.2	4,2E-02	5,8E-03	7,2E-03	7,82E-06
B	6,6E-04	2,5E-03	3E-06	4E-07

Table 3.3: relative total amount of 16S *Rickettsia* DNA as found in the *C. crambe* samples Cr2 SF, Cr4 SF, Cr5 SF, Cr6 SF and Cr7+8 SF by means of qPCR. Calibration curves were based on 10ng/ µl stock samples of pGEM-t-easy-*Rickettsia*16S. The given values are averages of triplicate measurements.

Sample	16S Rick (ng/ µl)
Cr2 SF	1,5E-06
Cr4 SF	1E-06
Cr5 SF	7,8E-07
Cr 6 SF	4,1E-06
Cr7+8 SF	9,7E-07

4 Discussion

4.1.1 qPCR

These samples were used to get acquainted with the qPCR equipment and the required protocols. During these sessions various problems were encountered. First of all it was found that the range of sample dilutions originally used by Aanstoot was too narrow. Instead of using pure sample, 20x and 40x dilutions it was decided to extend the range and use 10x, 100x and 1000x dilutions. Furthermore it was observed that even within the 10x diluted sample some reduced PCR efficiency could be observed, resulting in an artificially lower measured concentration. This is caused by either non DNA contamination within the sample or chemical residues from the extraction process. Based on this knowledge the concentrations within this report were calculated based on the 100x dilutions, except for the samples that fell outside the calibration curve in this dilution. Another problem encountered during these experiments were the relatively low Ct values observed for some of the negative controls especially with the universal 16S primers, figure 3.3. A similar problem was encountered by Aanstoot and although every kind of precaution was taken into account neither one of us succeeded to circumvent this problem. Later on in this research various samples were measured with both the universal 16S as the *β-proteobacterium* 16S specific primers, as can be seen in table 3.2 the concentrations measured with both primer sets are always in the same range. The universal 16S measurement appears not be influenced by the low Ct value of the negative control as long as samples exceeding this Ct are discarded.

4.1.2 Sample selection (Illumina sequencing)

For the Illumina metagenome sequencing assay four samples were nominated. The original Cr1 BF and Cr1 16K fractions plus the new Cr9 BF and Cr10 BF fractions. The qPCR was executed based on the same protocol used for the 'practice' samples. In figure 3.8 and figure 3.9 it can be seen that the Cr 16K fraction has a very much improved 18S/ 16S ratio as compared to the original Cr1 SF fraction. Proving the efficacy of this method to decrease the 18S content. During the differential centrifugation procedure tough a lot of template DNA was lost, therefore the total DNA concentration within this fraction was very low, as was depicted in figure 3.9. When comparing the Cr9 SF fraction with the Cr1 SF fraction one can see that the Cr9 has a much better 18S/ 16S ratio to start with. After the first differential centrifugation step that was used to create the Cr9 BF fraction the ratio became slightly

better as compared to the Cr1 16k fraction. Therefore it was chosen to use the Cr9 BF fraction with an approximately 2 fold lower 18S/ 16S ratio combined with an almost 5 fold higher total DNA concentration.

4.1.3 Repli-g amplification

Cr9 BF was amplified with the Repli-g kit, a form of multiple displacement amplification (Bayer, Schmitt et al. 2008; Binga, Lasken et al. 2008). The resulting product is depicted on gel in figure 3.10. Note that both the Cr9 BF sample as the negative control yield product. Based on the manufacturers literature it was found that this kit can form high molecular weight products if not presented with a template. The length of these products of approximately 25 kbp, see figure 3.12, are a characteristic of the polymerase that is used.

Because the DNA that was used for the calibration curve in previous experiments had ran out an alternative was tried. This resulted in no amplification at all within the calibration curves. At this stage we were only interested in the ratio's between 18S and 16S DNA and it was decided to determine these directly based on Ct values. Although these data are not scientifically significant they allowed for a swift selection of Cr9 BF as the best sample for Illumina sequencing. The Cr9.1 and Cr9.2 samples were pooled and send off to Baseclear for Illumina paired end metagenomic sequencing.

To obtain scientifically relevant qPCR data a series of calibration plasmids was made. With these plasmids it will be possible to standardize the measurement and compare the results of separate qPCR runs. When looking at figure 3.13 it can be seen that the 18S/ 16S ratio shifted slightly in favour of the Cr1 16K fraction as compared to the Cr9 BF fraction. Furthermore the negative control of the Repli-g reaction (Cr9.4) appears to have the best ratio of all measurements. In the melt curve graphs of the 16S and 18S reactions it was seen that these products have a different melting temperature and are therefore none specific. Although the Repli-g kit is claimed by the manufacturer to provide unbiased amplification we see, if we focus on the values measured for the Cr9 BF and Cr9.1/Cr9.2 fractions, that the 18S/ 16S ratio increases after the Repli-g amplification. This shift in ratio is probably caused by the logarithmic nature of PCR amplification techniques. For sequencing it would have been better if we had 10 µl of template DNA of the Cr9 BF fraction available. Since this was not the case we settled for second best.

Within previous research de Jaeger discovered that the abundance of the targeted *β-proteobacterium* was about 60 % of the total bacterial content of *C. crambe*. From the measurements depicted in table 3.2 it can be concluded tough, that this abundance might be much higher. In future research it would be interesting to measure some more SF fractions in order to get a real idea of the abundance of this bacterium within *C. crambe* and other sponges.

4.1.4 *Rickettsia*

During this study it was discovered that the *Rickettsia* like bacterium, previously described by de Jaeger, could be detected in all the samples that were used within this research, see table 3.2 and 3.3. Based on previous research it had been expected that this bacterium was quite unique and thus it was surprising to find it in every single sample. Therefore it is highly likely that this is a real symbiont of the sponge *C. crambe* and not a accidental discovery unique to the original *C. crambe* isolate.

Because the Ct values of the *Rickettsia* samples exceeded those of the used calibration curves it is impossible to give significant concentrations. What can be said is that the concentrations are very low with values at least below 10^{-5} ng/ µl. One footnote has to be added tough, based on literature *Rickettsia* is expected to be a intracellular symbiont and this might negatively influence the DNA extraction protocol, recovering a lower concentration than is actually available(Ilan, Micha and Abelson Avigdor. 1995; Sipkema, Detmer et al. 2009; Vacelet, Jean and Donadey Claude, 1977). On the other hand we do not observe higher concentrations of *Rickettsia* 16S DNA in the SF fractions as compared to the BF fractions. This would have been expected since the goal of the differential centrifugation is to remove sponge cells from the isolate and therefore would also remove intracellular symbionts. So either full recovery of the *Rickettsia* DNA present within the SF samples is not achieved or *Rickettsia* is not truly intracellular. Based on these data tough neither of these hypothesis can be validated.

4.1.5 Bioinformatics

Although the data from the Illumina run has been received from Baseclear not much has been learned from them yet. At this moment the process of assembling contigs from the +/-500.000 reads is arrested at the very first stage of binning the reads. Binning is a necessary step preceding the actual contig building whereby the reads are separated from each other on for example species level. In case of this research we wanted to separate the eukaryotic sponge reads from the prokaryotic reads. At first it was tried to bin the sequences based upon the BLAST (Basic Local Alignment Search Tool) similarity with partial *β-proteobacterium* 16S sequences. It was then attempted to build contigs based on the sequences that were successfully binned and subsequently blast the remaining reads against those contigs. Since this did not obtain the desired results a second approach is being tried at the moment. Based upon the article from Taylor et al. sponge associated *β-proteobacteria* were selected and the KEGG (Kyoto Encyclopedia of Genes and Genomes) data base was used to retrieve the gene sequences of enzymes involved in the ammonia pathways (Taylor, Radax et al. 2007). It is expected

that these gene's are unique to the bacteria within this metagenomic data set and that *C. crambe* does not posses this pathway. This tough is all still a work in progress.

4.2 Conclusion

The main goal of this research was to obtain the genome of the *β-proteobacterium* that was found in *C. crambe* by de Jaeger. Although the data output of the Illumina sequencing has been received no actual results were extracted yet. This was caused by a lack of facilities, expertise and foremost time.
What has been achieved is an optimization and standardisation of the qPCR reactions for the universal 16S, universal 18S, *Rickettsia* 16S and *β-proteobacterium* 16S genes. The plasmids that were produced can be useful for future research Secondly a metagenomic DNA sample from *C. crambe* was successfully amplified and validated. Finally the presence of the *Rickettsia* like bacterium previously found by de Jaeger was detected in multiple *C. crambe* samples

4.3 Future research

As mentioned previously, the Illumina sequencing data has been received but is still in the process of being analysed. The obvious first thing to do is to work further on this data set. These results will give a lot of new information that can not only be applied to this research in particular but also to follow up experiments and similar experiments on different sponges and/ or different symbionts. To assist the *in silico* contig assembly a BAC library can be constructed using the products from the Repli-g reaction(Ouyang, Dai et al. 2010). With the PCR primers used in this research the BAC's containing the *β-proteobacterium* and *C. crambe* genes can be selected and sequencing of these BAC can result in an important foothold for bining of the Illumina reads. If the genome or partial genome is successfully assembled it can be used to look into the genomic potential of this specific bacterium, which on it self can be used for the transcriptomics research that is in the pipeline.

In this research it was found that the *Rickettsia* like bacterium was found in every *C. crambe* sample that was tested. The concentrations were unanimously very low but still *Rickettsia* was present in every sample. This could suggest that this bacterium is essential to *C. crambe* and produces some low concentration metabolite or presents the sponge with a trace element that is not biologically available in seawater.
The concentrations that have been found for this bacterium suggest that the current approach that is being used for *β-proteobacterium* will not suffice to obtain enough bacterial template DNA. Since *Rickettsia* is supposed to be an intra cellular symbiont it might be possible to use a FACS methodology by double staining the *C. crambe* cells, containing *Rickettsia* bacteria, with two separate fluorophores.

Finally it would be interesting to look for *β-proteobacteria* and *Rickettsia* within more sponge and or sea water samples now that the qPCR reactions for these measurements are more standardized.

5 References

Bayer, K., S. Schmitt, et al. (2008). "Physiology, phylogeny and in situ evidence for bacterial and archaeal nitrifiers in the marine sponge Aplysina aerophoba." DOI. **10**: 2942-2955.

Binga, E. K., R. S. Lasken, et al. (2008). "Something from (almost) nothing: the impact of multiple displacement amplification on microbial ecology." ISME **2**(3): 233-241.

Bustin, S., J.-F. Beaulieu, et al. (2010). "MIQE precis: Practical implementation of minimum standard guidelines for fluorescence-based quantitative real-time PCR experiments." BMC Molecular Biology **11**(1): 74.

Faulkner, D. J. (2000). "Marine natural products." Nat.Prod.Rep. **17**: 7-55.

Ilan, M., A. Abelson. (1995). "The life of a sponge in a sandy lagoon." Biol.Bull. **189**: 363-369.

Koopmans, M., D. Martens, et al. (2009). "Towards commercial production of sponge medicines." Marine drugs **7**(4): 787-802.

Lee, K. L., J. H. Lee, et al. (2001). "Microbial symbiosis in marine sponges." J.Microbiol. **39**: 254-264.

Li, C.-W. J.-Y. T.-E. (1998). "Precambrian sponges with cellular structures." Science **279**(5352): 879.

Ouyang, Y., S. Dai, et al. (2010). "Isolation of High Molecular Weight DNA from Marine Sponge Bacteria for BAC Library Construction." Marine Biotechnology **12**(3): 318-325.

Perovic, S., A. Wichels, et al. (1998). "Neuroactive compounds produced by bacteria from the marine sponge Halichondria panicea: activation of the neuronal NMDA receptor." Environmental Toxicology and Pharmacology **6**(2): 125-133.

Schippers, K., D. Martens, et al. (2011). "Cell cycle analysis of primary sponge cell cultures." In Vitro Cellular & Developmental Biology - Animal **47**(4): 302-311.

Sipkema, D., M. C. F. Franssen, et al. (2005). "Marine sponges as pharmacy." Mar.Biotechnol. **7**: 142-162.

Sipkema, D., B. Holmes, et al. (2009). "Biological characterisation of Haliclona (?gellius) sp.: sponge and associated microorganisms." Microb.Ecol. **58**(4): 903-20.

Taylor, M. W., R. Radax, et al. (2007). "Sponge-associated microorganisms: evolution, ecology and biotechnological potential." Microbiol.Mol.Biol.Rev. **71**(2): 295-347.

Thomas, T., D. Rusch, et al. (2010). "Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis." **4**(12): 1557-1567.

Vacelet, J. and C. Donadey (1977). "Electron microscope study of the association between some sponges and bacetria." Mar.Biol. **30**: 301-314.

Vacelet, J. and E. Duport (2004). "Prey capture and digestion in the carnivorous sponge Asbestopluma (Porifera: Demospongiae)." Zoomorphology **123**(4): 179-190.

Webster, N. S. and R. T. Hill (2001). "The culturable microbial community of the Great Barrier Reef sponge Rhodopaloeides odorabile is dominated by an α-proteobacterium." Mar.Biol. **138**: 843-851.

Wilkinson, C. R. (1978). "Microbial associations in sponges. I. Ecology, physiology and microbial populations of coral reef sponges." Mar.Biol. **49**: 161-167.

Wilkinson, C. R., R. Garrone, et al. (1980). Nutrition of marine sponges. Involvement of symbiotic bacteria in the uptake of dissolved carbon. Nutrition in the lower metazoa, Pergamon**:** 157-161.

Wilkinson, C. R., M. Nowak, et al. (1981). "Specificity of bacterial symbionts in Mediterranean and Great Barrier Reef sponges." Microb.Ecol. **7**: 13-21.

Wright, A. D., A. McCluskey, et al. (2011). "Anti-malarial, anti-algal, anti-tubercular, anti-bacterial, anti-photosynthetic, and anti-fouling activity of diterpene and diterpene isonitriles from the tropical marine sponge Cymbastela hooperi." Organic & Biomolecular Chemistry **9**(2): 400-407.