Bacterial community in the intertidal sediments populated by *Arenicola marina*,
a terminal restriction fragment length polymorphism study

DISSERTATION

Zur

Erlangung des Grades eines

Doktors der Naturwissenschaften

- Dr. rer. nat.-

Dem Fachbereich Biologie/Chemie der

Universität Bremen vorgelegt von

Maya Shovitri

aus

Indonesien

Bremen, 2008

Die Untersuchungen zur vorliegenden Doktorarbeit wurden am Max-Planck-Institut für Marine Mikrobiologie in Bremen durchgeführt

1. Gutachter : PD. Dr. Jens Harder, Universität Bremen

2. Gutachter : Prof. Dr. Rudolf Amann, Universität Bremen

Contents

Chapter 3: Bacterial community in the intertidal sediments populated by *Arenicola marina*

Acknowledgment

Summary

Sediments offer microorganisms an unexplored numbers of niches with an opportunity to evolve specialized microbial communities. The small size of microbial niches in biogeochemical gradients in sediments called for a high resolution study of the populations. We applied a genetic fingerprint method, the terminal restriction fragment polymorphism (T-RFLP), to characterize the diversity of the 16S rRNA gene present in thin sediment layers at which one TRF represents one operational taxonomic unit (OTU). A partial gene amplification and restriction enzyme digestion of the amplicon allows the detection of about 150 different fragments in an intensity range of 100 to 10000 relative fluorescence units as a picture of the richness and evenness of the bacterial community.

The T-RFLP method was established for the intertidal soft sediments from the Königshafen at the northern end of the Sylt Island, Germany. The variations in the results were correlated to variations in individual steps of the method protocol. Restriction enzyme digest and digest analysis on a capillary sequencer correlated with a dissimilarity of about 20% and 10% in the obtained replicate datasets describing one pooled amplicon from one DNA sample after binning with a fixed window size of 0.5 and 1 base pair, respectively. Biases in individual PCR reactions did not increase the dissimilarity after performing independent T-RFLP analyses from one DNA sample. The dissimilarity was partly caused by an imperfect binning. Working with a high resolution window size of 0.5 bp, no starting point (50.25, 50.20, 50.30 and 50.65 bp) gave a perfect binning result. Some of identical TRFs were always binned into two different TRFs, thus creating an additional OTU. A window size of 1 bp with starting point 50.50 bp gave similar dissimilarities. Although our results may require an improved binning technique to utilize the full biodiversity information in the profiles, the current T-RFLP technique clearly detected the biological variation in adjacent small sediment layers and can be used to characterize bacterial community in individual sediment layers.

Eukaryotes offer and create a number of niches. The lugworm *Arenicola marina* is a bioturbator in intertidal sediments. The T-RFLP method was applied to investigate bacterial community in the burrow of the lugworm *A. marina*. The U-shaped burrow is divided into three compartments: the vertical head shaft tube

through which the surface sediment is sinking down and ingested by the lugworm, the horizontal gallery tube at where the lugworm relatively stays permanent inside the sediment and the vertical tail shaft tube through which the lugworm does defecation by moving backward until the tail reaches sediment surface and ejects characteristic fecal cast on the sediment surface.

From the bulk sediment surrounding the U-shaped burrow, the sediment contained a number of different bacterial communities changing with depth. On the basis of an aerobic layer, a redox potential discontinuity (RDP) layer and an anoxic layer, the decreasing and the increasing TRFs over depth may represent surface and subsurface layer bacteria respectively at 0-2 cm and 2-10 cm depth. The T-RFLP data suggested that the RDP layer is at 3-5 cm sediment depth, because the unique TRFs of the surface layer and subsurface layers were not found at this depth and the change of abundance of TRFs was fast.

The T-RFLP analyses clearly grouped bacterial population in the head shaft tube with the sediment surface populations. The tail shaft tube was populated by different populations; close to the surface dominated by the surface bacteria and below 3 cm dominated by the subsurface bacteria. The populations in the gallery tube were similar to those in the head and tail shaft tube. The richness in the gallery tube was the lowest but had the highest evenness.

T-RFLP analyses of two mm-thick sediment layers from areas with *A. marina* and without *A. marina* also revealed a strong depth-dependence of the surface bacterial community composition. According to the T-RFLP analyses, the presence or absence of *A. marina* had no clear detectable influence on the bacterial populations in the top two centimeter of sediment. Most likely, the increase presence of other burrowing animals in the *A. marina* exclusion areas seems to form highly similar biogeochemical environments for the development of bacterial communities.

Chapter 1

Introduction, results and general discussion

1.1. Genetic diversity studies by DNA fingerprinting methods

Since cultivation methods have limitations, nowadays molecular methods provide several approaches for the investigation of the bacterial community and diversity. For instance, PCR (polymerase chain reaction) amplification and phylogenetic analysis of the 16S rRNA gene is widely used in the exploration of microbial environments and the identification of uncultured organisms (Amann *et al.*, 1995). DNA fingerprinting is one of several molecular methods. It uses DNA samples to describe the identity of a sample within a certainty. For complex and dynamic samples due to seasonal fluctuations or environmental perturbations, it had been well reported that DNA fingerprinting is a powerful tool for distinguishing bacterial communities (Muyzer, 1999).

Several DNA fingerprinting methods are summarized by Muyzer (1999) with the advantages and disadvantages:

1. Low molecular weight (LMW) RNA (5S rRNA and tRNA). The total target RNA is separated by high resolution polyacrylamide gel electrophoresis and visualized by silver staining or by autoradiography if the RNA is radioactively label. It is a direct approach without in vitro amplification. It is limited by rapid degradation of RNA and the phylogenetic information and length variation of the LMW RNA.

2. Denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE). The target PCR products are separated by polyacrylamide gel containing a linear gradient of DNA denaturants. Sequence variation among different DNA molecules influences the melting behavior and melting slows the migrating at different positions in the gel. TGGE works with a temperature gradient. It is applicable only for short fragments (ca. 550 bp). Double bands and heteroduplex molecules may interfere in the result.

3. Single stranded conformation polymorphisms (SSCP). The target PCR products are denatured and separated on a non-denaturing gel. Separation is based on differences in the folded conformation of single stranded DNA which influences the electrophoretic mobility. It is mostly applied for short fragments (ca. 150-400 bp) and has reproducibility problems.

4. Randomly amplified polymorphic DNA (RAPD). The PCR amplicons are amplified by short (5-10 nucleotides) and random primers which anneal at different positions on the genomic DNA. Various lengths of PCR products are separated on agarose or acrylamide gel and visualized by ethidium bromide or silver staining. It has reproducibility problems and phylogenetic information can not be obtained.

5. Automated ribosomal intergenic spacer analysis (ARISA). The region between 16S and 23S rRNA genes is amplified using a fluorescent primer, separated by automated electrophoresis and visualized by an electropherogram profile that consists of fragment length (in base pair) and intensity (fragment height or area in relative fluorescence unit). This region range is highly variable in length and nucleotide sequences that can be used to distinguish closely related species (Fisher and Triplett, 1999).

6. Amplified ribosomal DNA restriction analysis (ARDRA). The PCR product of single or multiple 16S rDNA genes is digested by enzyme restriction, separated by electrophoresis in agarose or acrylamide gel and visualized by ethidium bromide or silver staining. Number of bands are not directly related to number of community members and limited by a complex microbial community.

7. Terminal restriction fragment length polymorphism (T-RFLP). It is fully described in the following chapter.

1.2. Terminal restriction fragment length polymorphism (T-RFLP)

T-RFLP was first demonstrated by Liu *et al.* (1997) as a modified method derived from ARDRA. It is a more powerful method than ARDRA for assessing the 16S rRNA gene based bacterial community analysis as ARDRA is limited by the resolution in agarose gels, especially for complex communities which had a large number of restricted fragments. The initial steps of DNA isolation, PCR amplification and enzyme restriction in T-RFLP are similar to those used in ARDRA. In T-RFLP,

the primers are labeled with fluorescent dyes, so that only the fluorescent terminal restriction fragment (TRFs) are detected and quantified by a high resolution gel electrophoresis on an automated DNA sequencer (Fig. 1.1). The T-RFLP method relies on variations in the position of restriction sites among 16S rRNA gene sequences, thus the bacterial diversity of complex community is determined as a pattern composite of the number of fluorescently labeled TRFs with unique length sizes in base pairs and the intensity of each TRF in relative fluorescent unit (rfu) (Liu *et al.*, 1997, Dunbar *et al.*, 2001).
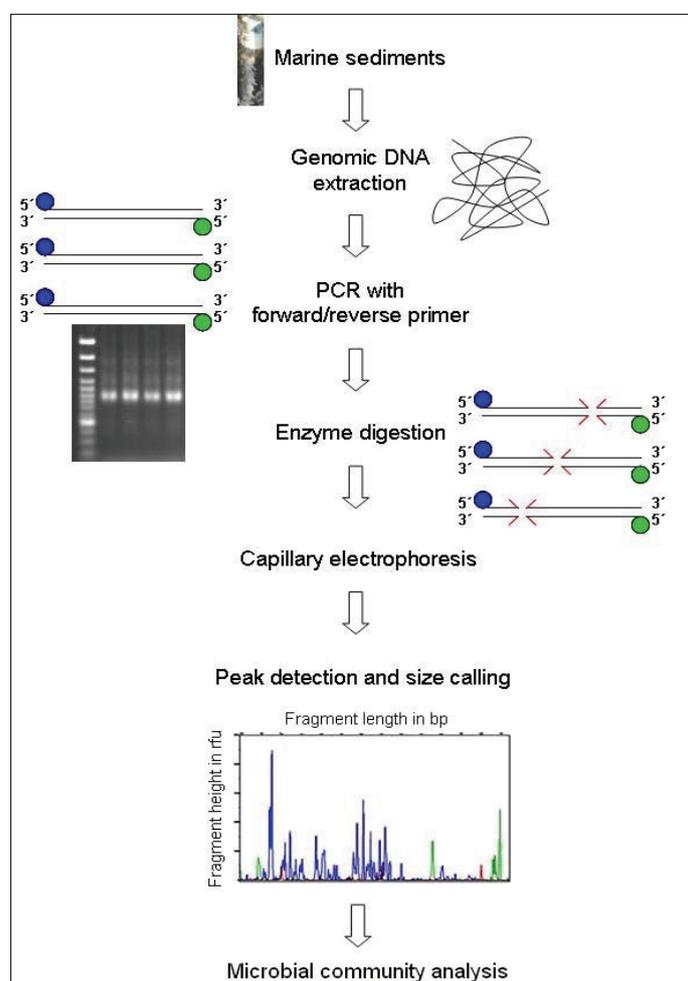

Fig. 1.1. T-RFLP scheme adapted from Application Note T-RFLP on the 3130/3730 (ABI, 2005).

T-RFLP had been reported as the most powerful fingerprinting method for a rapid comparison of the bacterial diversity from an environment based on the diversity of amplified 16S rRNA genes (Liu *et al.*, 1997; Marsh, 1999; Tiedje *et al.*, 1999) with a robust ability and reproducibility (Moeseneder *et al.*, 1999; Dunbar *et al.*, 2000; Osborn *et al.*, 2000). It successfully differentiates microbial communities when the optimal statistical approach is used in the study case, e.g. multivariate analysis (Blackwood *et al.*, 2003; Osborne *et al.*, 2006). This method had been applied for marine samples (Moeseneder *et al.*, 1999), soil samples (Clement *et al.*, 1998; Dunbar *et al.*, 2000; Osborn *et al.*, 2000, Blackwood *et al.*, 2003; Osborne *et al.* 2006), and in the activated sludge from the aeration tank, enrichment sludge from laboratory, aquifer sand from the groundwater and the gut of termite *Reticulitermes flavipes* (Liu *et al.*, 1997).

The quantitative data are sensitive to the technical variations which may arise from several causes as consequences applying molecular steps. Non-dominant microbial population can not be represented because the DNA template represented a small fraction of the total community DNA (Liu *et al.*, 1997). The low abundant species in nature may become the most abundance species after PCR or vice versa, due to the different copy number of 16S rRNA genes and preferences of PCR conditions, formation of PCR artifacts such as chimeric sequences and heteroduplex fragments, differential cell lyses and DNA extraction bias (Frey *et al.*, 2006; Dunbar *et al.*, 2001; von Wintzingerode *et al.*, 1997). The phylogenetic resolution is also limited since one TRF could be generated from multiple taxa. The analysis often goes to higher order group level than to species level, as phylogenetic related organisms could generate an identical size of TRF and it reduces the estimation of the diversity in complex community (Blackwood *et al.*, 2007; Engebretson and Moyer, 2003; Dunbar *et al.*, 2001; Liu *et al.*, 1997). Engebretson and Moyer (2003) presented data that a mean of 9.1 to 18.5 different sequences of a set 4600 16S rRNA gene sequences can generate one TRF. The conservation variability of restriction site position among 16S rRNA genes affects the resolution of T-RFLP. Incomplete or partial digestion may lead to an overestimation of the overall diversity within a community (Osborn *et al.*, 2000; Liu *et al.*, 1997; Farrelly *et al.*, 1995; Reysenbach *et al.*, 1992; Liesack *et al.*, 1991). TRFs are excluded if they are outside the determined size range. TRFs may not be detected because they are below the determined fluorescence threshold. Both processes clearly limit the true richness

(Blackwood *et al.*, 2007). Routinely occurring small pipetting errors, restriction enzyme digestion, TRFs separation, raw data analysis or statistical analysis are also potential causes of technical variation (Dunbar *et al.*, 2001).

A same fragment in different electrophoresis runs always shifts its position (Singh and Thomas, 2006) due to different running conditions in the capillary electrophoresis tube. The use of different fluorophores on the internal standard and sample fragments causes an absolute size difference. A fragment shifting of 0.5 bp was reported by Dunbar *et al.* (2001). A fluorophore is a molecule that is capable of fluorescing and used as a dyes, e.g. ethidium bromide and fluorescein. Different fluorophores have different electrophoretic mobility in the capillary electrophoresis (Abdo *et al.*, 2006; Hewson and Fuhrman, 2006; Dunbar *et al.*, 2001). This may introduce an error in accurate size determining of unknown fragments. Furthermore the analysis software also has a precision of the size calling algorithms of 0.01 bp. This precision may lead to an imprecision: a same fragment from different runs can be sized differently behind two decimal numbers. To count this fact, Abdo *et al.* (2006) and Hewson and Fuhrman (2006) suggested to do binning for minimizing total differences between replicate profiles within a sample as performed by Dunbar *et al.* (2001). Binning is combining the comparable fragment sizes found in the different runs into one length size within a defined window size. Defining a window size is based on the width range of fragment shifting (Hewson and Fuhrman, 2006).

Other related limitations are distinguishing true peak and noise. This decision is an important step in the T-RFLP method since true peaks, the unique fluorescently terminal restriction fragments (TRFs), correspond to operational taxonomic units (OTUs). Moreover, the peak height or peak area is incorporated as an additional quantitative parameter (evenness) in the profile comparison (Dunbar *et al.*, 2001). The optimal solution for this limitation is still under development. Several studies applied different statistical approaches, from simple steps to sophisticated steps. The simple step was normalization as conducted by Liu *et al.* (1997) and a sophisticated approach was an alterative and iterative filter with different kinds of threshold to exclude noise and to standardize the T-RFLP profiles as done by Dunbar *et al.* (2001), Saikaly *et al.* (2005), Osborne *et al.* (2006) and Abo *et al.* (2006). Other studies used a qualitative profile by converting T-RFLP profiles into binary profiles of presence and absence of TRFs for avoiding technical variations as applied by Fogarty and Voytek (2005); Clement *et al.* (1998). As different authors suggested

different kind of statistical approaches, the applied statistical approach needs a method establishment and a purpose dependent optimization.

1.3. Non-metric multi dimensional scaling (NMDS)

Multivariate method is a branch of statistics designed to reduce the complexity of high dimensional data by creating a low-dimensional data representation without ignoring the relationship among individual taxa. Concerning to the technical variation that potentially influences the profiles, multivariate method is the best choice for T-RFLP profiles to detect differences in community composition (Blackwood *et al.*, 2007). Multi dimensional scaling or ordination (MDS) and cluster analysis is one of several multivariate methods. Cluster analysis refers to a group of numerical techniques that attempt to classify individuals. This method takes the similarity matrix as the starting point and successively fuses the samples into groups in a hierarchical manner using group averages to link multiple samples (Clarke and Warwick, 2001).

MDS is a collection of data analysis techniques for embedding dissimilarity data in a space with a chosen dimensional Euclidean space and non-Euclidean space. The embedding is for data visualization and exploratory data analysis (van Wezel and Kosters, 2004; Young, 1985). Each object is represented by a point and the points are arranged in a space. The distances between pairs of points have the strongest possible relation to the similarities among the object pairs. Two similar objects are represented by two points that are close together, while two dissimilar objects are represented by two points that are far apart (Young, 1985). Classical MDS is subdivided into metric MDS (quantitative) and non-metric MDS (qualitative). Metric MDS assumes dissimilarities between objects are proportional to Euclidean distances while NMDS assumes they are related to Euclidean distances by some unknown monotone transformation (van Wezel and Kosters, 2004).

NMDS was first suggested by Shepard in 1962 *in* Kenkel and Orloci (1986). It is a simple statistical presentation of the differences in a profile. It is a parsimonious ordination of individuals in space dimensions, based on a rank order agreement between distances and similarities (Kenkel and Orloci, 1986). The iterative algorithm normally converges to an optimal ordination by successively refining positions of the points until they satisfy as closely as possible the similarity or dissimilarity relations between samples. But as it uses unknown transformation, a non-optimal ordination is

also possible, especially for a poorly structured data. Therefore a number of different starting configurations may have to be tested to get an optimal ordination with a low stress value (Shepard, 1974 *in* Kenkel and Orloci, 1986). The stress value *S* is a measure of deviation from monotonicity of observed dissimilarities and ordination distances (Kruskal *et al.*, 1964 *in* Kenkel and Orloci, 1986). It also refers to a statistic of goodness of fit. A stress value S>0.2 indicates that NMDS is close to random, *S*<0.2 indicates a useful 2 dimensional picture and *S*<0.1 indicates that NMDS corresponds to an ideal ordination with no real prospect of misinterpretation (Clarke and Gorley, 2001).

NMDS based on Euclidean coefficient had been proven as the best strategy for ordination non linear data structures such as typical ecological profiles. Recently this ordination method has been widespread applied in ecological studies that have complex environmental parameters. Because only rank order is used, NMDS has advantages: e.g. input could be a large variety of resemblance measures (Kenkel and Orloci, 1986).

Instead of Euclidean coefficient, Rees *et al.* (2004), Clark and Warwick (2001) and Faith *et al.* (1991) used Bray-Curtis coefficient for their T-RFLP datasets, as it is an appropriate coefficient to calculate datasets with a majority of blocks with zero numbers. Bray-Curtis similarity coefficient determines site similarities based on organism abundances. The distance obtained from Bray-Curtis coefficient reflects differences between two samples due to differing community composition and/or differing total abundance (Equation 1).

$$\delta_{jk} = 100 \left\{ 1 - \frac{\sum_{i=1}^{p} \left| y_{ij} - y_{ik} \right|}{\sum_{i=1}^{p} \left(y_{ij} + y_{ik} \right)} \right\} \qquad (1)$$

where :

    j and k = two compared samples

    y_{ij} = the abundance of the ith species in sample *j*

    y_{ik} = the abundance of the ith species in sample *k*.

NMDS can be complemented by the SIMPER (similarity percentage) and the ANOSIM (analysis of similarity) test. The SIMPER test assesses which species are primarily responsible for an observed difference between sample groups. The species will be listed in decreasing order of their importance in contributing to the average dissimilarity between two sample groups (Clarke and Gorley, 2001). Then the overall significance of the difference is assessed by the ANOSIM. The ANOSIM is non-parametric test. It uses the rank order of dissimilarity value between sample groups. The pairwise combination is divided into two types: between groups and within groups. If two sample groups are different in their species composition, then compositional dissimilarities between the sample groups are greater than those within the sample groups. The ANOSIM statistic R is based on the difference of mean ranks of all dissimilarities between sample groups (*rb*) and within sample groups (*rw*) (Equation 2). N is total number of replicates (Clarke and Gorley, 2001).

$$R = \frac{rb - rw}{N(N-1)/4} \tag{2}$$

The pairwise R values indicate the separation of the sample groups on a scale of -1 to 1. The large positive R (up to 1) signifies dissimilarity between sample groups: $R > 0.75$ = the samples groups are well separated; $R > 0.5$ = the sample groups are overlapping but clearly different; $R < 0.25$ = the sample groups are not separated at all. The significance level of the separation depends on replicate number in each sample group; if the replicate is few, the significance level is often low. Thus the significant level is not essential since the R value gives an absolute measurement for sample group separation (Clarke and Gorley, 2001).

In nature, negative R values were possible when the sampling area were very patchy so that replicates were variable, but each sample had similar amount of variability among replicates; when either or both samples contained an outlier; when sampling area had 2 different states and the replicates had sampled each of these states (Chapman and Underwood, 1999).

1.4. Ecosystem, diversity and bioturbation

A species is a natural group of actually or potentially interbreeding individuals and reproductively isolated from other groups. All individuals of a given species in an area constitute a population. Several different populations that occur together in an area constitute a community (Nybakken, 1997). The microbial populations within a community interact in an integrated manner. Each indigenous population has a specialized functional role called a niche. Populations compete to occupy the available niches and to use the same resources. Thus the successful population plays a contributable functional role in maintaining the community (Atlas and Bartha, 1997). Related to those, diversity refers to species richness (number of species within a community), species evenness (the total number of individuals among the species) and composition of living organisms (Nybakken, 1997).

An ecosystem is an assemblage of communities and abiotic environments (physical-chemical factors) which interact to each other in an area. Ecosystem is a natural system. Active interaction within this assemblage reflects to the ecosystem functioning which maintains natural processes and establishes complex ecological balances over time. The natural processes includes nutrient cycling (feeding, excreting, decomposing), breading, growing, adaptation and disturbance (Nybakken, 1997; Atlas and Bartha, 1997).

Species and community is not always in an equilibrium state. Diversity is maintained through continual or gradual environmental changes and periodic disturbances (Connell, 1978 *in* Nybakken, 1997). Disturbance means altering the physical and chemical condition of environments. It may influence the diversity, while it promotes a changing of species richness, evenness and species composition, although each species has a certain tolerances to all environmental factors (Nybakken, 1997).

Microorganisms have strategies to survive and maintain themselves in the environmental change. These strategies classify organisms along an r-K gradient (Equation 3) for a population growth in limited environmental conditions (Andrews, 1991 *in* Atlas and Bartha, 1997).

$$\frac{dX}{dt} \cdot \frac{1}{X} = r - \left[\frac{r}{K} \cdot X \right] \qquad (3)$$

where:

$\frac{dX}{dt} \cdot \frac{1}{X} =$ specific rate of population increase

r = per capita rate of increase of the population

K = carrying capacity of the environment

X = population density as either numbers or biomass

When X is low, the rate of population change is dominated by r. When X is high, the growth rate is limited by K. The r-strategist has a high rate of reproduction, while the K-strategist has an optimal utilization of environmental resources. Microorganism optimizes either reproductive capacity or resources conservation, but not both (Atlas and Bertha, 1997). A pioneer should be the r-strategist with the highest growth rate. The r-strategist has few other competitive adaptations and tends to prevail in limited resources; therefore they should have a high intrinsic growth rate. The population is extremely fluctuating. When the resources turn unfavorable, they experience rapid reduction. Cyanobacteria are an example for r-strategists that respond to nutrient enrichment with an explosive bloom. The K-strategist tend to compete successfully even in limited sources and has a slow reproductive rate. Their population is usually more stable and is a permanent member of a community. They prevail under condition of crowding and devote a smaller portion of their resources to reproduction. Soil *Streptomycetes* is a K-strategist which grows slowly on complex soil organic compounds (Atlas and Bertha, 1997).

Bioturbation is a disturbance caused by biological activities and strongly affects other organisms. All living organisms in some way affect their immediate abiotic environments. But only organisms that have with their presence or absence a disproportionately large impact on the ecosystem are bioturbators, e.g. dam-building beavers, earthworms and burrowing organisms. In sediment, a bioturbator may changes the physical habitat by feeding, reworking sediments, bioirrigation and biogenic structure building (e.g. burrow and tube construction) affecting flow of

resources and redox conditions for microorganisms (Kogure and Wada, 2005; Meysman *et al.*, 2006).

The reworking activity transports deposited and labile organic matter from surface to the deeper layer, enhance solute exchange between overlying and pore waters column to the deeper layer and removes the reduced compounds from deeper layer to the surface layer at which all of them stimulate redox rates relating to the bacterial population and bacterial processes over depth (Kristensen, 2001). The bioirrigation introduces oxygen and other solutes into formerly anoxic sediments (Hüttel, 1990; Kristensen *et al.*, 1985) and potentially inhibitor compounds as results of anaerobic metabolisms in deeper layer are removed to upper layer, e.g. sulfide (Kristensen, 2001; Hüttel, 1990; Kristensen *et al.*, 1985). The burrow and tube structures is considered as physical extension of the sediment-water interface and increases the surface area for solutes diffusion out or into the sediments via active bioirrigation (Kristensen *et al.*, 1985).

Not all bioturbators act in the same way in sediment. The mode of bioturbation (reworking sediments, bioirrigation and biogenic structure building) determines the impact on microbial activities and biogeochemistry processes. Characteristics of benthic systems can also have major impacts on sediment bioturbation and microbial activity as Covich *et al.* (2004) described two different benthic systems: (1) diffusion-dominated benthic systems characterized by fine-grained sediments and low hydrological connections between free water and interstitial water, and (2) advection-dominated benthic systems characterized by coarse sediment and strong hydrological connections between free water and interstitial water. In diffusion-dominated system, bioturbator can produce water fluxes at the water-sediment interface that may strongly influence microbial processes in sediments, whereas in advection-dominated system bioturbator can only modify the water circulation pattern in sediment and moderately affecting microbial processes. Marine water-sediment interface correspond to the diffusion-dominated system characterized by fine sediments (muddy sand) and negligible advective transport of water into sediment (Mermillod-Blondin and Rosenberg, 2006).

1.5. *Arenicola marina*

The lugworm *Arenicola marina* (Fig. 1.2.A), a burrowing polychaete, is a potential bioturbator that irrigates and reworks the deeper sediments by its feeding activity (Alyakrinskaya, 2003; Kristensen, 2001; Riisgard and Banta, 1998). It reaches about 20-30 individual/m^2 and plays an important role in the physical and chemical succession process on sandy flat intertidal zone of the Wadden Sea (Volkenborn et al., 2007; Volkenborn et al., 2007a). It lives relatively permanent in a burrow while eating subsurface and sunk down sediment and defecating at the sediment surface. It lives head down in a 20 – 40 cm deep J-shaped burrow in sediment with an adult body length about 15 to 25 cm. It ingests surface sediments in feeding pocket, and as a result sand above the lugworm head depresses downward forming a feeding funnel in the surface. The J-shaped burrow was then completed to a U-shaped burrow by a vertical head shaft through which the surface sediment is sinking down and ingested (Fig. 1.2.C). *A. marina* assimilates living bacteria, microphytobenthos, microfauna and meiofauna associated with sinking down sediment. During defecation, the lugworm moves backward through the tail shaft until the tail reaches sediment surface and ejects characteristic fecal cast (Fig. 1.2.B). Depression in the funnel feedings and fecal cast mounds were clearly seen in sediment surface of high populated area (Kristensen, 2001; Riisgard and Banta, 1998).

A. marina irrigates its burrow with oxygen-rich overlying water with a peristaltic movement in a posterior-anterior direction (Riisgard and Banta, 1998). The active irrigation period had a 5-10 minute duration and was interrupted by very short period of inactivity (Kristensen, 2001). Approximately 3 L $h^{-1}m^{-2}$ oxygen-rich overlying water could be pumped by a lugworm density of 30 ind. m^2 (Riisgard, *et al.,* 1996). The irrigation results in highly oxic and oxidized conditions in the burrow zone and surrounding sediments (Kristensen, 2001).
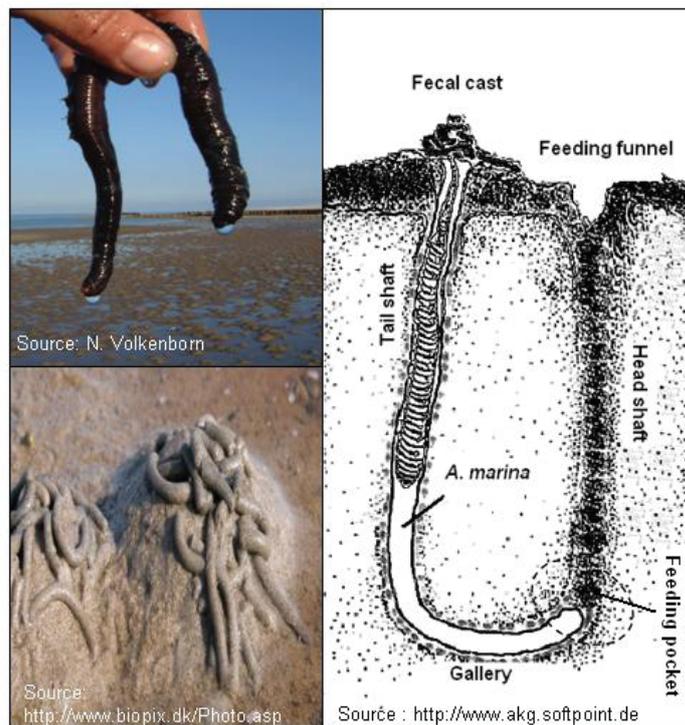
Fig. 1.2. (A). *Arenicola marina*. (B). Fecal cast on the surface indicated that *A. marina* lives in its U-shaped burrow beneath the marine sediment layers. (C). Scheme of the U-shaped lugworm burrow.

Ingestion and defecation is a cyclical pattern in phase with irrigation cycles and effects a sediment particle movement. The defecating interval is about 40 minute for large lugworm and about 15 minute for smaller one (Kristensen, 2001; Riisgard and Banta, 1998). A density of 30 ind/m^2 corresponded to 15 cm yr^{-1} sediment turnover (Cadee, 1976). This particle reworking affects sediment stability and composition. It developed successive sediment changes from muddy to sandy sediments that corresponded to higher sediment permeability and decreasing concentrations of phosphate, ammonium, nitrite, nitrate and sulfide (Volkenborn *et al.*, 2007). A distinct layer of gravel and bivalve shells was observed at a depth beneath the U-shaped burrow since the lugworm only ingests < 2mm particles and refuses larger particles (Riisgard and Banta, 1998).

1.6. Intertidal sediments

The intertidal or littoral zone is an area of a foreshore and seabed which is exposed to air at low tide and submerged at high tide. Organisms which are living in this area are tolerant to desiccation and well adapted to a changing environment. The adaptations may be behavioural (i.e. movements or actions), morphological (i.e. characteristics of external body structure), or physiological (i.e. internal functions of cells and organs). Salinity varies from fresh to highly saline due to rainwater or river inputs and tidal inundations. Temperature can be ranged widely from very hot with full sun to near freezing in cold climes and the waves can dislodge the residents (Nybakken, 1997).

Intertidal habitats may have hard or soft bottoms or substrates. Rocky shores tend to have higher wave action and allow inhabitants to adapt by attaching tightly to the rocks, e.g. headlands and cobble beaches. Soft sediment habitats include sandy beaches, mudflats, and salt marshes which are generally protected from large waves but tend to have more variable salinity levels. Many soft-sediment inhabitants are adapted for burrowing. Muddy sediments are more stable and conducive for establishing a permanent burrow. The intertidal zone is divided into four zones: the low tide zone is dry only at the lowest tides and contains the highest biodiversity. Organisms are not well adapted to periods of dryness and temperature extremes, e.g. tube worms. The middle tide zone is regularly covered and uncovered twice a day by tide sea water. The high tide zone is covered by water during high tide so it experiences dry periods daily and spray zone which survives on mist and spray water (Nybakken, 1997).

The sampling area in this study was a low tide zone in the Königshafen, at the northern end of the Sylt Island, Germany. The area was covered by sea water for 9-10 hours/tide with mean tidal about 1.8 m. The salinity was 27.5 $^o/_{oo}$ in spring and 31 $^o/_{oo}$ in summer with a negligible freshwater seepage. *A. marina* dominated this area about 20-30 individu/m^2 (Fig. 1.2.A). Its presence changed physical and chemical sediment properties (Table 1.1). Other burrowing inhabitants in this area which presence was significantly reduced by the presence of *A. marina* were *Nereis diversicolor, Pygospio elegans, Polydora cornuta, Tubificoides benedii, Capitella capitata* and *Scoloplos cf. armiger.* Other macrobenthose species mainly abundant in the absence of A. marina were mussel *Macoma balthica,* worm *Spio martinensis,*

snails *Hydrobia ulvae,* bivalve *Mya arenaria,* sand mason *Lanice conchilega* and bivalve *Cerastoderma edule* (Volkenborn, *et al..* 2007, 2007a; Volkenborn and Reise, 2007).

Table 1.1. The physical and chemical sediment properties influenced by the presence of *A. marina* summarized from Volkenborn, *et al.* 2007 and 2007a.

		Bioturbated plot	Non-bioturbated plot
Emersion period		9-10 h / tide	
Main hydrodynamic force		Tidal current	
Mean tidal height		≈ 1.8 m	
Salinity		27.5 $^o/_{oo}$ in spring	
		31 $^o/_{oo}$ in summer	
Average *A. marina* density		18-30 ind / m^2	-
Grain size of fine sand sediment		0 -1 cm : 204 µm	0 -1 cm : 190 µm
		1 - 5 cm : 218 µm	1 - 5 cm : 206 µm
Fine fraction (particle size <63 µm) at 0 - 5 cm depth		<1% dry wt	1% - 2.5% dry wt
Water content in 0 – 5 cm depth		16.20%	19.90%
Sediment permeability		2.6×10^{-12} m^2	$<1.0 \times 10^{-12}$ m^2
Approximately porewater profile through 0-10 cm depth	Ammonium	<100 µM	<150 µM
	Nitrite	<0.3 µM	<0.25 µM
	Nitrate	<2.5 µM	<2.5 µM
	Phosphate	<10 µM	<15 µM
	Sulphide	<25 µM	<150 µM
Average oxygen penetration		<4 cm depth	<1 cm depth

1.7. Vertical zonation in marine sediments related to bioturbation

The geochemical stratification in the marine sediments tends to develop a zonation of an aerobic layer, a redox potential discontinuity (RDP) layer and an anoxic layer. The zone dimension varies depending on the concentration of decomposable organic compounds in sediments, delivery of electron acceptor and

turnover rates in the sediments (Schulz, 2000; Kristensen, 2001; Nybakken, 1997). It is principally related to oxygen penetration and the sequence of available electron acceptors over depth which determine the kind of occurring carbon, nitrogen and sulfur cycles at the particular depth. This zonation directly effects the relative contribution of aerobic and anaerobic bacteria which are key player in those cycles (Nybakken, 1997). Organic compounds in marine sediments are degraded aerobically and anaerobically into precursors or inorganic compounds with a decaying rate depend on several factors: e.g. the kind of organic compound (cellulose, lignin, etc.), age (decomposition stage) and temperature (Kristensen, 2001). Generally the compounds formula is $(CH_2O)_x(NH_3)_y(H_3PO_4)_z$, where the C:N:P ratio of x:y:z is varying and depends on the origin and age of the compounds. Based on the Redfield composition, the x:y:z ratio is 106:16:1 (Schulz, 2000).

In an oxidized zone at which dissolved oxygen is available as an electron acceptor, the organic compounds can be completely oxidized to H_2O, CO_2 and inorganic compounds according to Equation (4). The RDP layer is the transition layer whereby the oxygen concentration drops sharply within a narrow depth and electron acceptors are shifting gradually from oxygen to NO_2^-, NO_3^-, Mn^{4+} and Fe^{3+}, thus denitrification, manganese and iron reduction take place and in parallel energy output decreases. Equations (5), (6) and (7) are examples of complete oxidation of the marine organic compounds with Mn^{4+}, NO_3^- and Fe^{3+} as an electron acceptor respectively. An anoxic zone characterized by black colour is the sulphidic layer and the methane layer. Sulfate reduction and methanogenesis dominantly occurs at this layer at where the electron acceptors SO_4^{2-} and CO_2 are present at which the methanogenesis layer is just beneath the sulfate reduction layer (Jorgensen, 2000; Kogure and Wada, 2005). Sulphate is reduced and sulphide is produced as presented in the Equation (8) (Froelich *et al.*, 1979 *in* Schulz, 2000). Fermentation, a redox process in the absence of exogenous electron acceptor, may also possibly occur in the anoxic layer and accomplishes the anaerobic degradation. As anaerobic organisms are more limited than aerobic organisms in the ability to degrade certain large complex compounds, anaerobic decomposition may occurs stepwise involving several different functional organisms (Kristensen, 2001).

$$(CH_2O)_{106}(NH_3)_{16}(H_3PO_4) + 138\ O_2 \rightarrow 106\ CO_2 + 16\ HNO_3 + H_3PO_4 + 122\ H_2O \qquad (-3190\ kJ/mol) \quad (4)$$

$$(CH_2O)_{106}(NH_3)_{16}(H_3PO_4) + 236\ MnO_2 \rightarrow 236\ Mn^{2+} + 106\ CO_2 + 8\ N_2 + H_3PO_4 + 366\ H_2O \qquad (-3090\ kJ/mol) \quad (5)$$

$$(CH_2O)_{106}(NH_3)_{16}(H_3PO_4) + 84.4\ HNO_3 \rightarrow 42.2\ N_2 + 16\ NH_3 + 106\ CO_2 + H_3PO_4 + 148.4\ H_2O \qquad (-2750\ kJ/mol) \quad (6)$$

$$(CH_2O)_{106}(NH_3)_{16}(H_3PO_4) + 212\ Fe_2O_3 + 848\ H^+ \rightarrow 424\ Fe^{2+} + 106\ CO_2 + 16\ NH_3 + H_3PO_4 + 530\ H_2O \quad (-1410\ kJ/mol) \quad (7)$$

$$(CH_2O)_{106}(NH_3)_{16}(H_3PO_4) + 53\ SO_4^{2-} \rightarrow 53\ S^{2-} + 106\ CO_2 + 16\ NH_3 + H_3PO_4 + 106\ H_2O \qquad (-380\ kJ/mol) \quad (8)$$

Bioturbation by burrowing organisms may destruct this vertical zonation and the associated biogeochemical processes within sediments. It significantly influences microbial activities and biogeochemical processes by modifying water and sediment fluxes at the water-sediment interface (Mermillod-Blondin and Rosenberg, 2006). Generally it leads to an expansion of the oxic zone in sediments and the exchange of reduced or oxidized compounds (Kogure and Wada, 2005; Kristensen, 2001). Reduced compounds or ions are accumulated in the suboxic zone or anoxic zone: e.g. NH_4^+, PO_4^{3-}, Fe^{2+}, Mn^{2+}, SH^- were transported upward, conversely the oxidized compounds or ions (e.g. NO_3^-, Fe^{3+}) were transported downward (Kogure and Wada, 2005). Fecal strings may create anaerobic microniches, at where anaerobic processes, e.g. denitrification and sulfate reduction, occur in apparently oxic surface sediments (Kristensen, 2001). This exchange enhances or suppresses the function of certain bacterial groups. Close coupling of the sulfur reduction and oxidation may occur on a small scale in the RDP zone: sulfate reduction is suppressed and sulfur oxidation is enhanced by oxic water. The nitrogen oxidation and reduction is strongly correlated to each other and stimulated by burrowing activity and probably coupled in the RDP zone. Under oxic condition, NO_2^- is oxidized to nitrate by nitrifying bacteria. Under anoxic condition, nitrate is reduced to gas nitrogen by denitrifying bacteria as reviewed by Kogure and Wada (2005).

1.8. The microbial community in the U-shaped burrow

The gardening phenomenon referred to a stimulation of microbial growth in the head shaft burrow due to high suspended nutritional compounds from lugworm's secretions, decaying plant matter, diffusion from the surrounding anoxic area and a supply of oxygenated overlying water through the burrow that is needed by bacteria

(Retraubun *et al.*, 1996; Riisgard and Banta, 1998; Kogure and Wada, 2005). The mucous wall lining described as sulphated or phosphate-rich mucopolysaccarides (Zola, 1967 and Muzii, 1968 *in* Kristensen *et al.*, 1985) should provide a potential degradable substrate for bacterial growth while microbial activity and biomass was higher in the wall of burrow than those found in the surrounding sediments (Aller and Yingst, 1978). This phenomenon in the U-shaped burrow was first suggested by Hylleberg in 1975, as he believed that feeding pause about 6 hours and was sufficient for a bacterial proliferation (Retraubun *et al.*, 1996) even Rijken (1979 *in* Retraubun *et al.*, 1996) argued against this phenomenon, because an ingestion rate in the head shaft was too fast for allowing a significant bacterial growth. A high potential of nitrification in the burrow walls correlated with the content of mucus (Kristensen *et al.*, 1985) but Nielsen *et al.* (2003) reported that mucus had no significant impact on the sulfate reduction rate in the tail shaft.

More studies supported the gardening phenomenon as below. The feeding funnel was assumed to form a trap for organic matter for bacterial growth by Retraubun *et al.* (1996), because the bacterial count in the feeding funnel (2.43 X 10^7 cell cm^{-3}) and in the head shaft (1.82 X 10^7 cell cm^{-3}) of the U-shaped burrow *A. marina* related to a detritus concentration (43.3±11.9 mg/g sediments) in the feeding funnel. Grossman and Reichardt (1991) reported similar counts: the feeding funnel and feeding pocket contained a bacterial density of 2 X 10^{10} cell cm^{-3}. Even though detritus was not directly consumed by *A. marina* due to no cellulose secretion (Longbottom, 1970 *in* Retraubun *et al.* 1996), detritus served evidently as a nutrient source for bacteria (Retraubun *et al.* 1996).

A higher bacterial number (2.5 - 7.8 X 10^{10} cell cm^{-3}) was counted in the foregut of *A.marina* than in the funnel feeding (about 2 X 10^{10} cell cm^{-3}), the feeding pocket (about 2 X 10^{10} cell cm^{-3}) and the feces (about 1 X 10^{10} cell cm^{-3}) while the lowest bacterial count number was in the hindgut (0.5 – 1.5 X 10^{10} cell cm^{-3}). Cytophaga-like bacteria were found about 0.5 to 3 X 10^5 colony forming units (CFU) cm^{-3} in the foregut and less than 0.5 X 10^5 CFU cm^{-3} in the hindgut (Grossman and Reichardt, 1991). This result showed that *A. marina* fed on bacteria, as supported by Plante and Mayer (1996) who reported a seasonal variation of bacteriolytic rates from *A. marina* digestive fluid. Grossman and Reichardt (1991) related this result to the gardening phenomenon, because the increase bacterial counts corresponded to high concentration of organic compounds in the foregut. Retraubun *et al.* (1996) provided

another evidence for the gardening phenomenon: they blocked the water current in the U-shaped burrow by inserting a plastic sheet. After 48 hour without water current, the bacterial number in the feeding funnel had statistically decreased, from 1.88 X 10^7 cell cm^{-3} to 1.35 X 10^7 cell cm^{-3}.

The direct effect of A. marina on the microbial population due to ingestion was reported by Goni-Urriza et al. (1999) as the counted that the bacterial density in non-bioturbated areas (5.9 to 8.1 X 10^9 cell cm^{-3}) was higher than those in the areas bioturbated by *A. marina* and the bivalve *Cerastoderma edule* (5.6 to 9.0 X 10^8 cell cm^{-3}). But the numbers were relatively constant from the surface to 5 cm depth. The indirect effect was also observed while the densities of colorless sulfur-oxidizing bacteria and anoxygenic phototrophic bacteria (85% affiliated to *Thiocapsa roseopersicina*) at the surface area (0 to 1 cm depth) were almost similar in the bioturbated and non-bioturbated area (Goni-Urriza et al., 1999). In subsurface layers (1 cm to 5 cm depth) they were higher in the bioturbated area; although they were assumed to occupy the small zone at the oxygen-sulfide interface in the non-bioturbated area (Goni-Urriza et al., 1999). This agreed with observations of Reichardt (1986). He detected a peak of CO_2 dark fixation within the RDP layer in about 2 cm depth in the bioturbated area. At this layer thiobacilli (sulfur-oxidizing bacteria) were assumed to live favorably. The electron donor $S_2O_3^{2-}$ stimulated the CO_2 dark fixation in the burrow wall of the polychaete *Nereis diversicolor* (Reichardt, 1986).

The bacterial community and activity in burrow depends on several factors: the physical and chemical properties of the burrow environments, habitat characteristics (e.g. organic compound content, grain size distribution, water column nutrient, phytoplankton concentration), ecology of burrow inhabitants (e.g. feeding type, irrigation pattern, mobility, type of secretion), burrow age and stability (Mermillod-Blondin and Rosenberg, 2006; Kristensen, 2001). Matsui et al. (2004) reported that the SRB assemblages within the Diopatra tubes were different between intertidal sandy flat and intertidal mud flat, as a greater diversity was found in the intertidal mud flat. Papaspyrou et al. (2006) investigated two different burrow walls of close related worms but behaviorally different: a facultative suspension feeder *Nereis diversicolor* and an obligate deposit feeder *N. virens*. Both burrow walls were different in content of particulate organic carbon, particulate organic nitrogen, ratio C:N and chlorophyll a; respectively it was 370 µmol.g^{-1}.dw^{-1} , 38 µmol.g^{-1}.dw^{-1} ,

9.8 (mol:mol) and 13 $\mu g.g^{-1}.ww^{-1}$ for the burrow wall of *N. diversicolor* and 225 $\mu mol.g^{-1}.dw^{-1}$, 22 $\mu mol.g^{-1}.dw^{-1}$, 10:1 (mol:mol) and 4.9 $\mu g.g^{-1}.ww^{-1}$ for the burrow wall of *N. virens*. Thus the rate of carbon decomposition, nitrogen decomposition and sulfate reduction was higher in the burrow wall of *N. diversicolor*, even though unexpectedly the bacterial abundance was lower than in the wall of *N. virens*. Furthermore the denaturating gradient gel electrophoresis (DGGE) profiles showed that bacterial communities in those two areas were different about 80% (Papaspyrou *et al.*, 2006).

1.9. Study aim

The effect of *A. marina* on the geochemical parameters that referred to the microbial processes aerobically and anaerobically whether directly in the U-shaped burrow or indirectly in the bulk sediment surrounding the burrow had been examined by many studies, e.g. recently published by Volkenborn *et al.* (2007 and 2007a). But it was limited for bacterial community especially for sample taken directly from the U-shaped burrow (Retraubun *et al.*, 1996; Grossmann and Reichardt, 1991; Reichardt, 1988). Grossmann and Reichardt (1991) suggested that the entire burrow system should be considered to evaluate the impact of burrowing bioturbator on the biogeochemical process in the sediment, while the burrow is a physically stable habitat on a day or week time scale and a chemically unstable habitat of oxic-anoxic change due to bioirrigation (Kristensen, 2001). After measuring the oxygen consumption, nitrification and denitrification directly in the U-shaped burrow, Kristensen (2001) considered that different microorganisms may grow in the burrow.

As a complement study, we investigated bacterial community change directly along the U-shaped burrow from the surface to 10 cm by applying the T-RFLP method with a resolution of 1 cm sediment compartment. Sediment samples were taken from the head shaft, the tail shafts and the gallery tube. The bacterial populations in those compartments were compared with those in the bulk sediment. The bulk sediment was from the middle part between the head shaft and the tail shaft. The change in the surface microbial populations due to the presence of the lugworm was also aimed in this study with a 2 mm resolution. As we applied 1 cm and 2 mm resolution, this study is the first T-RFLP application with a high resolution for tracking the bacterial change over depth in marine sediments.

The thesis is divided into 3 chapters:

− Chapter 1 gives a general introduction, an overview on the results and a general discussion.

− Chapter 2 describes a manuscript of the T-RFLP method establishment. Variations in the PCR reaction, the restriction enzyme digestion, the capillary electrophoresis, peak detection and peak sizing, and the binning strategy are evaluated by experimental approaches.

− Chapter 3 presents the application of the optimized T-RFLP method to determine the bacterial communities in the U-shaped burrow of the lugworm *A. marina* and in the populated versus the not populated area.

1.10. Results and general discussion

1.10.1. Method establishment

Enzyme digestion. Choosing an appropriate enzyme digestion is important to accurately reflect the microbial diversity, since T-RFLP diversity pattern relies on the unique restriction site variability in 16S rRNA gene sequences. Ideally the enzyme digestion should generate as many as possible unique TRF across the phylogenetic groups of bacteria. But as 16S rRNA gene sequences among bacteria consist of variable and conserved regions, the phylogenetic resolution of TRF is limited to this fact: either the unique restriction site is situated in the conserved region or in the variable region. Consequently one TRF size could be generated by several related or non-related bacteria (Liu *et al.*, 1997, Dunbar *et al.*, 2001). An enzyme digestion which potentially yields a set of predicted TRFs with a low redundancy can not provide information on the phylogenetic composition of a community (Dunbar *et al.*, 2001). Therefore the use of combinations of single enzyme digestion can be the best strategy for general profiling of bacterial community and phylogenetic interference (Liu *et al.*, 1997, Dunbar *et al.*, 2001; Osborne *et al.*, 2006).

The restriction enzymes *Alu*I, *Hha*I and *Msp*I were used in this study. They provide high unique numbers of 5´-TRFs and 3´-TRFs for the primer pair 27F-907R as revealed by an enzyme resolving power analysis and a virtual digestion with the free software MICA 3 (Table 1.2). But especially for the method establishment, we used only the restriction enzyme *Alu*I because in practice it generated a stable unique number of 5´-TRFs and 3´-TRFs in replicate profiles with a high similarity percentage of reproducible TRFs.

Table 1.2. The unique number of 5´-TRFs and 3´-TRFs amplified by primer pair 27F-907R predicted by using a free software MICA 3 (http://mica.ibest.uidaho.edu/).

Enzyme digestion	Enzyme resolving power analysis				A virtual digestion hit number (100 – 900 bp)	
	Total hits	Full length sequence	Unique 5´-TRF	Unique 3´-TRF	Unique 5´-TRF	Unique 3´-TRF
*Alu*I	29667	28968	541	506	455	434
*Msp*I	29667	29635	499	313	443	259
*Hha*I	29667	27224	717	607	610	499

Peak separation and peak detection. As a TRF is presented by a peak, distinguishing true peak and noise is a fundamental step. The first objective in this study was to optimize the parameter options for the peak detection which are available in the GeneMapper® Software v3.7. By applying default values, the capillary electrophoresis on the ABI Prism 3130XL genetic analyzer separates the fragments according to size. The GeneMapper® software v3.7 (ABI) performs peak analysis and peak sizing. The commercial available internal markers exhibited the peaks as described by the manufacturers with minor additional peaks. The main peaks were detected at stable retention times in all observed profiles.

We applied a peak amplitude threshold 100 rfu and light smoothing for analyzing the samples. This removed small true peaks in the digestion samples and false-positive peaks in the non-digested sample. Reducing small peaks can be critical if they are present frequently within replicates; they may be important for distinguishing samples, even if they resulted from PCR bias reactions (Blackwood *et al*., 2003). However, our analyses with a peak amplitude threshold of 100 rfu and light smoothing revealed a depth dependent change in community structure as we used depth dependent subsamples. The depth dependent change was not enhanced with a lower peak amplitude threshold. A constant peak threshold of 100 rfu also had been reported as a satisfactory threshold for standardization and presenting a similarity of the triplicate profiles by Osborne *et al.* (2006).

Binning and TRF profile reproducibility. The second objective in this study was to explore the reproducibility of peak number, peak size in bp and peak intensity in rfu. We performed two different strategies for generating replicate profiles of one subsample: nine replicate profiles and triplicate profiles strategy. Following the strategy of Dunbar *et al*. (2001), the nine replicate profiles were generated from pooled PCR products from 3 individual PCR mixtures. The pooled PCR product was divided into 3 individual enzyme digestion mixtures of which each of them was loaded into 3 different wheels of capillary electrophoresis. The triplicate profiles were generated from 3 individual PCR mixtures and continued to an individual enzyme digestion mixture and a single capillary run. These replicate strategies were purposed to address the technical variation on the level of PCR, restriction digestion and the capillary electrophoresis analysis.

To have a clear picture of replicate relationship we used ordination of the non-metric multi dimensional scaling (NMDS). NMDS is based on a rank order agreement between distances and similarities. An individual replicate profile of one subsample is represented by a point, thus the technical and biological variation is viewed as distance between the points. If the points of replicate profiles from one subsample are close together meant that these replicate profiles are similar; the biological variation between these replicate profiles is higher than the technical variation. The biological variation refers to the variation in bacterial composition consisting of TRFs and their abundance (relative area).

Binning can split a peak in independent samples into separate adjacent bins and attribute to the appearance or disappearance of a TRF (Hewson and Fuhrman, 2006). To optimize the binning, five starting points (50.25, 50.10, 50.20, 50.30 and 50.65 bp) for a fixed window size 0.5 bp were tested and compared with a man-made manual alignment. All automatically binned profiles were more diverse than the manual profile due to the introduction of false distributed peaks in the binned profile. Several identical 5´-TRFs in the nine replicate profiles appeared as two different fragments after binning. A binning window size of 1 bp with starting point 50.50 bp did not solve the problem.

The manual alignment yielded a high number of reproducible TRFs, over 90% of the observed peaks were reproducible. The automatic binning yielded on average a reproducibility of above 80%. Thus restriction enzyme digest and digest analysis on a capillary sequencer together with the data analysis may produce a dissimilarity of up to 20% in the obtained replicate profiles describing one pooled amplicon from one DNA sample (the nine replicate profiles). Due to that the 5 different starting points still yielded improper binning, for the triplicate profiles we applied a fixed window size 0.5 bp and 1 bp only respectively with starting point 50.25 and 50.50 bp. The biases in individual PCR reactions did not increase the dissimilarity among the triplicate profiles; the dissimilarity was same as indicated by the nine replicate profiles (10% - 20%).

The fixed window size 0.5 bp and 1 bp of the nine replicate and triplicate profiles generated a highly similar NMDS pattern: cluster of the replicate profiles and distances between the clusters representing the depth dependent differences in the bacterial communities (Fig. 1.3). The formation of depth dependent clusters indicated that the biological variation between the bacterial communities in the different layers was larger than the technical variations between the replicate profiles.
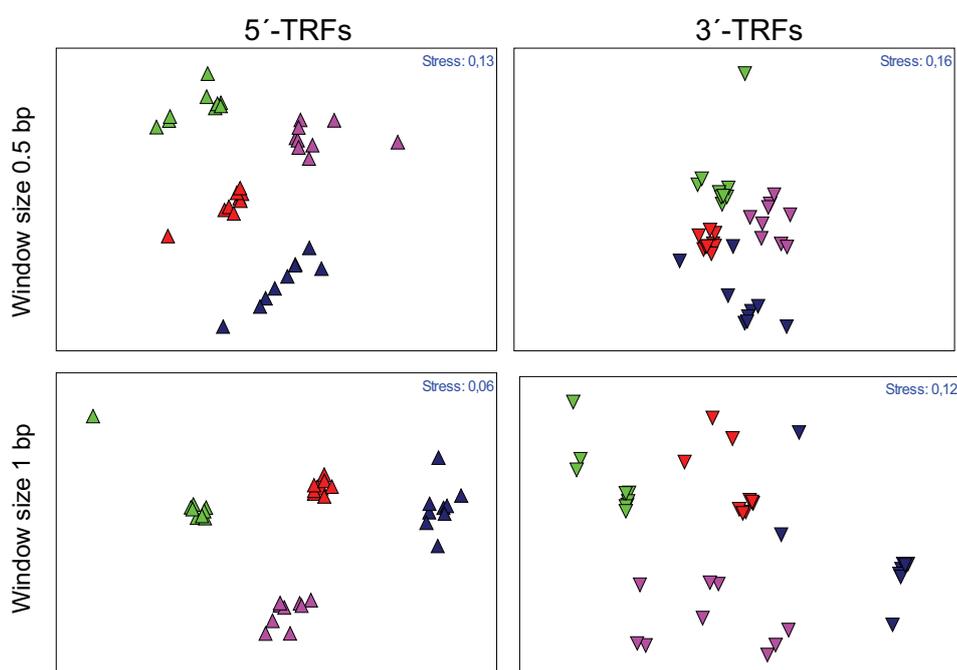
Fig. 1.3. The nine replicate profiles. NMDS ordination of quantitative 5´-TRF and 3´-TRF datasets after binning with a fixed window size 0.5 bp and 1 bp respectively with starting points 50.25 and 50.50 bp. The The green, pink, red and blue triangle is for sample 6.3 (1-1.5 cm depth), 6.4 (1.5-2 cm depth), 6.5 (2-2.5 cm depth) and 6.6 (2.5-3 cm depth), respectively.

The NMDS ordination (Fig. 1.3) incorporated richness (presence of a TRF) and evenness (area of the TRF) information. However, the utilization of evenness information has been questioned on the basis of eventually occurring PCR biases. Such quantitative T-RFLP studies should be interpreted as a reflection of differences in the bacterial community composition after the amplifying process rather than a true

difference in bacterial community diversity (Blackwood *et al.*, 2007; Saikaly *et al.*, 2005; Dunbar *et al.*, 2001; Liu *et al.*, 1997).

To clarify the influence of a PCR amplification bias, an NMDS ordination for qualitative datasets (binary datasets) applying the Jaccard coefficient was performed. In general the ordinations calculated by the free software package PAST 1.38 (Palaeontological Statistics, http://folk.uio.no/ohammer/past/) yielded a pattern similar to that of the quantitative datasets: the replicate profiles were reproducible and the depth dependent bacterial communities appeared in the presentations. This may indicate that a PCR bias is not detectable and that the quantitative datasets can be used for determining bacterial community changes.

The 3´-TRF datasets may generate less resolution than the 5´-TRF datasets (Saikaly *et al.*, 2005; Dunbar *et al.*, 2001). In this study, both 5´-TRF and 3´-TRF datasets yielded a similar presentation of the bacterial diversity in the sediment, even though in the quantitative and qualitative NMDS ordination the 5´-TRF datasets represented higher resolution. With the primer pair 6-FAM-27F and HEX–907R, the 5´- terminus may provide a greater discrimination due to the higher number of variable regions and less conserved region than the middle part of the gene. The application of forward and reverse labeled primers was intended to avoid an accidental low resolution, as related organisms may generate an identical size of 5´-TRF due to the similarity of conserved regions and enzyme recognition site, but different size of 3´-TRF (Abdo *et al.*, 2006).

1.10.2. Bacterial community in the intertidal sediments populated by *Arenicola marina*

The U-shaped burrow. In this study we used the triplicate profiles strategy and binning with a fixed window size 1 bp. The sediment samples were from the head shaft, the tail shafts and the gallery tube of the U-shaped burrow and from the bulk sediment. The 1 cm resolution sliced sediments of the T-RFLP profiling successfully tracked the bacterial community change in the burrow. Three enzymes digestion, *Alu*I, *Hha*I and *Msp*I, generated a highly similar reflection of local bacterial communities: a surface sediment signature in the whole head shaft tube and a bacterial community shifting with depth in the tail shaft tube (Fig. 1.4).
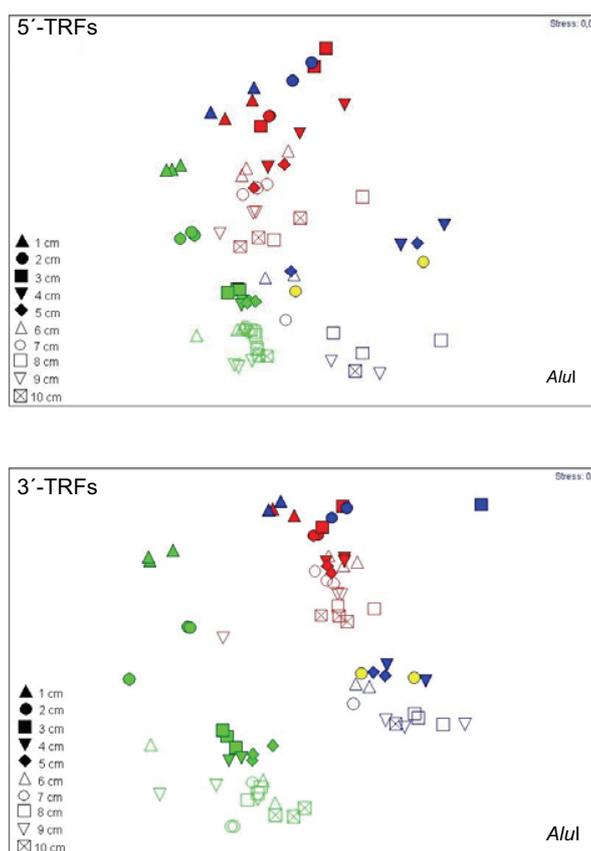
Fig. 1.4. The U-shaped burrow tubes and the bulk sediment. Bray-Curtis similarity based NMDS ordinations of 5´-TRF and 3´-TRF datasets after *Alu*I digestion represents a bacterial community shifting with depth in the tail shaft and the bulk sediments. The color code indicates green, blue, red and yellow respectively for the bulk sediment, tail shaft tube, head shaft tube and gallery tube.

The head shaft tube was a fragile and oxic habitat (Riisgard and Banta, 1998; Alyakrinskaya, 2003), thus T-RFLP result was coincident with the collected brownish sediment parts. The bacterial community over depth was detected as the surface layer bacterial communities similar with those at the surface layers in the tail shaft tube and the bulk sediment. In contrast, the tail shaft is a physically stable part due to mucous lining (Aller and Yingst, 1978; Papaspyrou, 2005; Kogure and Wada, 2005).

The bacterial community in the tail shaft represented distinctive surface and subsurface layer bacterial communities, even the collected samples were from the brownish sediment parts. This T-RFLP pattern may reflect to Jorgensen (1977) and Fenchel (1996). Jorgensen (1977) stated that the transported sediments, e.g. due to *A. marina* (Kristensen, 2001), even if it is a small particle could be a unique aerobic or anaerobic microhabitat for microbes. Fenchel (1996) observed that due to a brief period of anoxia between the continuous irrigation, the apparent vertical zonation of microbial processes and of microbial community in the burrow reflected a diminishing oxic fraction rather than an ideal vertical redox sequences (Fenchel, 1996) and the aerobic and anaerobic bacteria may present at the same depth from the surface layer to 10 cm depth (Fenchel, 1996; Kristensen *et al*., 1985) as the finding of aerobic ciliate (*Kentrophoros fasciolata* and *Euplotes sp.*) and anaerobic ciliate (*Myelostoma bipartititum* and *Parablepharisma pellitum*) at the same depth in the *A. diversicolor* burrow (Fenchel, 1996a).

The gallery tube is a transition microhabitat between the tail and head shaft tube, at where the physical and chemical sediments properties may change fast as the fact that the lugworm stays relative permanently in this part (Kristensen, 2001; Riisgard and Banta, 1998). Thus this condition was represented by the lowest richness and the highest evenness; only the highly adapted bacteria referred to the surface layer bacteria were growing.

In general, the T-RFLP profile is able to represent the reported effect of *A. marina* on the bacterial community either directly (Retraubun *et al*., 1996; Plantae and Mayer, 1994; Grossmann and Reichardt, 1991; Reichardt, 1988) or indirectly (Volkenborn et al., 2007, 2007a; Kristensen, 2001) as the U-shaped burrow had a lower richness and a higher evenness than the bulk sediment. Several TRFs referred to surface layer bacteria were more abundance in the U-shaped burrow probably a

reflection for a unique niche for them and the gardening phenomenon, as extremely found in the gallery tube.

The surface layers in the populated and not populated area. With a 2 mm resolution, a gradual bacterial community shifting at the surface layer was also successfully tracked by the method in both areas, but no clear differences between them (Fig. 1.5). This indicated that either *A. marina* has no influence on the surface microbial community or that other biologically or physically causes may influence the same effect on the surface bacterial community. The intertidal sediment surface is physically unstable due to e.g. waves and periodically tidal currents but chemically stable due to a continuous oxic condition in overlying water (Kristensen, 2001). In the absence of *A.marina*, other bioturbators from polychaetes were reported significantly inhabiting the non-bioturbated area; e.g. *Nereis diversicolor, Pygospio elegans, Polydora cornuta, Tubificoides benedii, Capitella capitata* and *Scoloplos cf. armiger* which are classified into surface-subsurface deposit feeding worms and make burrows into the sediments (Volkenborn and Reise, 2007).
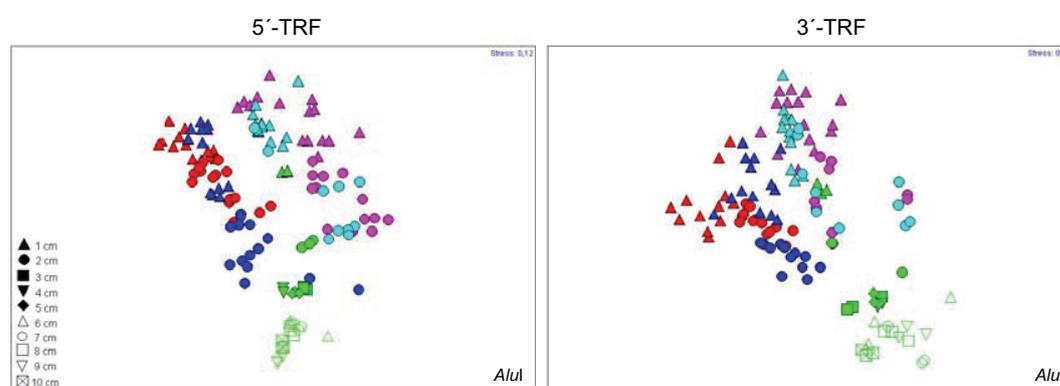
Fig. 1.5. The populated and not populated area. Bray-Curtis similarity based NMDS ordinations of 5´-TRF and 3´-TRF profile after *Alu*I digestion. They present a continuous bacterial community shifting with depth. The color code indicates green, red, pink, blue and light blue respectively for the bulk sediment, core I and core II of the not populated area, core I and core II of the populated area.

Our results may indicate that *A. marine* significantly influenced the subsurface bacterial community, especially direct in the U-shaped burrow tube. This agreed with Volkenborn *et al*. (2007 and 2007a) and Goni-Urriza *et al*. (1999), that the effect of the *A. marina* on the physical and chemical sediments properties (Volkenborn et al., 2007 and 2007) and on the bacterial counting number (Goni-Urriza *et al*., 1999) was significant at depth below 2 cm. Moreover Volkenborn *et al*. (2007 and 2007a) reported that the effect of *A. marina* was not restricted only to the burrow but to the entire area at which *A. marina* supported the intertidal habitat succession from muddy sediment to sandy sediment.

1. 11. Outlook

The U-shaped burrow is a unique habitat since it is considered as a physically stable habitat on a day or week time scale and a chemically unstable habitat of oxic-anoxic change in parallel with the lugworm activities (Kristensen, 2001). Therefore microbial populations and related microbial processes must be fluctuating following those conditions depend on several factors: the physical and chemical properties of the burrow environments, habitat characteristics (e.g. organic compound content, grain size distribution, water column nutrient, phytoplankton concentration), ecology of burrow inhabitants (e.g. feeding type, irrigation pattern, mobility, type of secretion), burrow age and stability (Mermillod-Blondin and Rosenberg, 2006; Kristensen, 2001). In this profiling study, it was detected that the U-shaped burrow had several microhabitats over depth which varied between the head shaft tube, the tail shaft tube and gallery tube. The head shaft was populated by bacteria of the oxic surface sediment. The gallery tube seemed to be populated by species of the oxic-anoxic transition zone. The tail shaft tube walls had distinctive surface layer bacterial communities at the top layers and subsurface layer bacterial communities in the lower part.

Further complementary studies may need to explore this unique habitat beside only over depth. For example with respect to the method, it is interesting to have a bacterial profile over the season at which in summer time the lugworm is more active than in winter time with *in situ* sampling; and over the age of the burrow at where the old structure might be different with the new structure in supporting the bacterial community and process, this probably can be performed with a microcosm study as

performed by Nielsen *et al*. (2003). To have a whole picture of direct and indirect effect of the lugworm on the bacterial aspects, those studies are combined with the environmental parameters (e.g. volume of oxic sediments in the taken sample, particle size) and product of aerobic and anaerobic metabolisms (e.g. carbon decomposition, nitrogen decomposition and sulfate reduction directly from the U-shaped burrow) with several available methods, e.g. radio isotopes study. Further molecular studies, e.g. fluorescence in situ hybridization with several chosen probes that relates to the measured bacterial products or the 16S rRNA gene clone libraries may provide more insight into the fine structure of the microbial communities.

Beside with molecular approaches, an attempt to isolate bacteria is also interesting since the U-shaped burrow is probably inhabited by highly adapted aerobic bacteria that may have unique metabolisms and mechanism to survive in the most fluctuating condition. Especially for aerobic bacteria found in the gallery tube, the obtained TRFs may represent that bacteria in there are able to proliferate faster than the ingestion speed of the lugworm. Finding out the physical and chemical factors, e.g. carbon source, might be critical in this attempt while the gardening phenomenon may play a role for the bacteria.

1.12. References

1. Abdo, Z., U.M.E. Schütte, S.J. Bent, C.J. Williams, L.J. Forney and P. Joyce. 2006. Statistical methods for characterizing diversity of microbial communities by analyzing of terminal restriction fragment length polymorphisms of 16S rRNA genes. Environ. Microbiol. 8:929-938.

2. Aller, R.C. and J.Y. Yingst. 1978. Biogeochemistry of tube-dwellings: a study of the sedentary polychaete *Amphitrite ornate* (Leidy). J. Mar. Res. 36:201-254.

3. Alyakrinskaya, I.O. 2003. Some ecological features of the lugworm *Arenicola marina* L. (Annelida, Polychaeta) and its morphological and biochemical adaptations to burrowing. Biology Bulletin. 30:411-418.

4. Amann, R.I., W. Ludwig and K-H. Scheifer, 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. 59:143-169.

5. Atlas, R.M and R. Bartha. 1997. Microbial Ecology, fundamentals and applications. 4th Edition. Benjamin/Cummings Publishing Company, Inc. California.

6. Blackwood, C.B., D. Hudleston, D.R. Zak and J.S. Buyer. 2007. Interpreting ecological diversity indices applied to terminal restriction fragment length polymorphism data: insights from simulated microbial communities. Appl. Environ. Microbiol. 73:5276-5283.

7. Blackwood, C.B., T. Marsh, S-H. Kim and E.A. Paul. 2003. Terminal restriction fragment length polymorphism data analysis for quantitative comparison of microbial communities. Appl. Environ. Microbiol. 69:926-932.

8. Cadee, G.C. 1976. Sediment reworking by *Arenicola marina* on tidal flats in the Dutch Wadden Sea. Neth. J. Sea. 1:440-460.

9. Chapman, M.G. and A. J. Underwood. 1999. Ecological pattern in multivariate assemblages: information and interpretation of negative values in ANOSIM test. Mar. Ecol. Prog. Ser. 180:257-265.

10. Clarke, K.R. and R.N. Gorley. 2001. PRIMER v5: User manual/tutorial, PRIMER-E, Plymouth UK.

11. Clarke, K.R. and R.M. Warwick. 2001. Change in Marine Communities: An Approach to Statistical Analysis and Interpretation. 2nd edition: PRIMER-E, Plymouth, UK.

12. Clement, B.G., L.E. Kehl, K.L. DeBord and C.L. Kitts. 1998. Terminal restriction fragment pattern (TRFPs), a rapid PCR-based method for the comparison of complex bacterial communities. J. Microbiol. Methods. 31:135-142.

13. Covich, A.P., M.C. Austen, F. Bärlocher, E. Chauvet, B.J. Cardinale, C.L. Bile, P. Inchausti, O. Dangles, M.Solan, M.O.Gessner, B. Statzner and B. Moss. 2004. The role of biodiversity in the functioning of freshwater and marine benthic ecosystem. Bioscience. 54:767-775.

14. DeFlaun, M.F. and L.M. Mayer, 1983. Relationship between bacteria and grain surfaces in the intertidal sediments. Limnol. Oceanogr. 28:873-881.

15. Dunbar, J., L.O. Ticknor and C.R. Kuske. 2001. Phylogenetic specificity and reproducibility and new method for analysing of terminal restriction fragment profile of 16S rRNA genes from bacterial communities. Appl. Environ. Microbiol. 67:190-197.

16. Dunbar, J., L.O. Ticknor and C.R. Kuske. 2000. Assessment of microbial diversity on two south western U.S. soils by terminal restriction fragment analysis. Appl. Environ. Microbiol. 66:2943-2950.

17. Engebretson, J.J and C. L. Moyer. 2003. Fidelity of select restriction endonucleases in determining microbial diversity by terminal restriction fragment length polymorphisms. Appl. Environ. Microbiol. 69:4823-4829.

18. Faith, D.P., C.L. Humphrey and P.L. Dostine. 1991. Statistical power and BACI designs in biological monitoring: comparative evaluation of measures of community dissimilarities based on benthic macroinvertebrate communities in Rockhole Mine Creek, Northern Territory, Australia. Aust. J. Mar. Freshwat. Res. 42:589-602.

19. Farrely, V., F.A. Rainey and E. Stackebrandt. 1995. Effect of genome size and rrn gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. Appl. Environ. Microbiol. 61:2798-2801.

20. Fenchel, T. 1996. Worm burrow and oxic microniches in marine sediments. 1. Spatial and temporal scales. Mar. Biol. 127: 289-295.

21. Fenchel, T. 1996a. Worm burrow and oxic microniches in marine sediments. 2. Distribution pattern of ciliated protozoa. Mar. Biol. 127:297-301.

22. Fisher, M.M. and E.W. Triplett. 1999. Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. Appl. Environ. Microbiol. 65:4630-4636.

23. Fogarty, L.R. and M.A. Voytek. 2005. Comparison of *Bacteroides-Prevotella* 16S rRNA genetic markers for fecal samples from different animal species. Appl. Environ. Microbiol. 71:5999-6007.

24. Frey, J.C., E.R. Angert and A.N. Pell. 2006. Assessment of biases associated with profiling simple, model communities using terminal-restriction fragment length polymorphism-based analysis. J. Microbiol. Met. 67:9-19.

25. Goni-Urriza, M., X. de Montaudouin, R. Guuyoneaud, G. Bachelet and R. de Wit. 1999. Effect of macrofaunal bioturbation on bacterial distribution in marine sandy sediments, with special reference to sulphur-oxidising bacteria. J. Sea Res. 41:269-279.

26. Grosmann, S. and W. Reichardt. 1991. Impact of *Arenicola marina* on bacteria in the intertidal sediments. Mar. Ecol. Prog. Ser. 77:85-93.

27. Hewson, I. and J. A. Fuhrman. 2006. Improved strategy for comparing microbial assemblage fingerprints. Microbiol. Ecol. 51:147-153.

28. Hylleberg, J. 1975. Selective feeding by *Abarenicola pacifica* with notes on *Abarenicola vagabunda* and a concept of gardening in lugworm. Ophelia. 14:113-137.

29. Hüttel, M. 1990. Influence of the lugworm *Arenicola marina* on porewater nutrient profiles of sand flat sediments. Mar. Ecol. Prog. Ser. 62:241-248.

30. Jorgensen, B.B. 2000. Bacteria and marine biogeochemistry, p:173-203. *In* Schulz, H.D. and M. Zabel (ed.), Marine geochemistry. Springer.

31. Jorgensen, B.B. 1977. Bacterial sulphate reduction within reduced microniches of oxidized marine sediments. Mar. Biology. 41:7-17.

32. Kogure, K and M. Wada. 2005. Minireview: Impact of macrobenthic bioturbation in marine sediments on bacterial metabolic activity. Microbes and Environments. 20:191-199.

33. Kenkel, N. C. and L. Orloci. 1986. Applying metric and nonmetric multidimensional scaling to ecological studies: some new results. Ecology. 67:919-928.

34. Kristensen, E. 2001. Impact of polychaetes (*Nereis* spp and *Arenicola marina*) on carbon biogeochemistry in coastal marine sediments. Geochem.Trans. 2:92-103

35. Kristensen, E., M.H. Jensen and T.K. Andersen. 1985. The impact of polychaete (*Nereis virens* Sars) burrows on nitrification and nitrate reduction in estuarine sediments. J. Exp. Mar. Biol. Ecol. 85:75-91.

36. Liesack, W., H. Weyland and E.Stackebrandt. 1991. Potential risks of gene amplification by PCR as determined by 16S rDNA analysis of a mixed culture of strick barophilic bacteria. Microbiol. Ecol. 21:191-198.

37. Liu, W., T.L. Marsh, Cheng H. and L.J. Forney. 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl. Environ. Microbiol. 63:4516-4522.

38. Marsh, T.L. 1999. Terminal restriction fragment length polymorphism (T-RFLP): an emerging method for characterizing diversity among homologous populations of amplification products. Current Opinion in Microbiol. 2:323-327.

39. Matsui,G.Y., D.B. Ringelberg and C.R. Lovell. 2004. Sulfate-reducing bacteria in tubes constructed by the marine infaunal polychaete *Diopatra cuprea*. Appl. Environ. Microbiol. 70:7053-7065.

40. Mermillod-Blondin, F. and R. Rosenberg. 2006. Ecosystem engineering: the impact of bioturbation of biogeochemical process in marine and freshwater benthic habitats. An overview article. Aquat. Sci. 68:434-442.

41. Meysmann, F.J.R, O.S. Galaktionov, B. Gribsholt, J.J. Middelbutg. 2006. Bioirrigation in permeable sediments: advective porewater transport induced by burrow ventilation. Limnol. Oceanogr. 51:142-156.

42. Moeseneder, M.M., J.M. Arrieta, G. Muyzer, C. Winter and G.J. Herndl. 1999. Optimization of terminal restriction fragment length polymorphism analysis for complex marine bacterioplankton communities and comparison with denaturing gradient gel electrophoresis. Appl. Environ. Microbiol. 65:3518-3525.

43. Muyzer, G. 1999. Genetic fingerprinting of microbial communities-present status and future perspectives. Proceeding of the 8th International Symposium on Microbial Ecology. Bell CR, Brylinsky M and Johnson-Green P. (Ed). Atlantic Canada Society for Microbial Ecology. Halifax. Canada.

44. Nielsen, O.I., E. Kristensen and M. Holmer. 2003. Impact of *Arenicola marina* (Polychaete) on the sediment sulfur dynamics. Aquatic Microbial Ecol. 33:95-105.

45. Nybakken, J. W. 1997. Marine Biology, an ecological approach. 4th Edition. Addition-Wesley Educational Publishers Inc. USA.

46. Osborn, A.M., E.R.B. Moore and K.N. Timmis. 2000. An evaluation of terminal restriction fragment length polymorphisms (T-RFLP) analysis for the study of microbial community structure and dynamics. Environ. Microbiol. 2:39-50.

47. Osborne, C.A., G.N. Rees, Y. Bernstein and P.H. Janssen. 2006. New threshold and confidence estimates for terminal restriction fragment length polymorphism analysis of complex bacterial communities. Appl. Environ. Microbiol. 72:1270-1278.

48. Papaspyrou, S., T. Gregersen, E. Kristensen, B. Christensen and R.P. Cox. 2006. Microbial reaction rates and bacterial communities in sediment surrounding burrows of two nereidid polychaetes (*Nereis diversicolor* and *N. virens*). Marine Biology. 148:541-550.

49. Plante, C.J and L.M. Mayer. 1994. Distribution and efficiency of bacteriolysis in the gut of *Arenicola marina* and three additional deposit feeders. Mar. Ecol. Prog. Ser. 109:183-194.

50. Ranjard, L., D.P.H. Lejon, C. Mougel, L. Schehrer, D. Merdinoglu and R. Chaussod. 2003. Sampling strategy in molecular microbial ecology: influence of soil sample size on DNA fingerprinting analysis of fungal and bacterial communities. Environ. Microbiol. 5:1111-1120.

51. Reichardt, W. 1988. Impact of bioturbation by *Arenicola marina* on microbiological parameters in intertidal sediments. Mar. Ecol. Prog. Ser. 44:149-158.

52. Reichardt, W.T. 1986. Polychaete tube walls as zonated microhabitats for marinebacteria. Deuxième Colloque International de Bactèriologie Marine. 415-425.

53. Rees, G.N., D.S. Baldwin, G.O. Watson, S. Perryman and D.L. Nielsen. 2004. Ordination and significance testing of microbial community composition derived from terminal restriction fragment length polymorphisms: application of multivariate statistics. Antonie van Leeuwenhoek. 86:339-347.

54. Retraubun, A.S.W., M. Dawson and S.M. Evans. 1996. The role of the burrow funnel in feeding processes in the lugworm *Arenicola marina* (L). J. Exp.Mar. Biol. and Ecol. 107-118

55. Reysenbach, A.L., L.J. Giver, G.S. Wickham and N.R. Pace. 1992. Differential amplification of rRNA gene by polymerase chain reaction. Appl. Environ. Microbiol. 58: 3417-3418.

56. Riisgard, H.U, I. Bernsten and B. Tarp. 1996. The lugworm (*Arenicola marina*) pump: characteristic, modelling and energy cost. Mar. Ecol. Prog. Ser. 138:149-156.

57. Riisgard, H.U and G.T. Banta. 1998. Irrigation and deposit feeding by the lugworm *Arenicola marina*, characteristics and secondary effects on the environment. A review of current knowledge. Vie Milieu. 48:243-257.

58. Saikaly, P.E., P.G. Stroot and D.B. Oether. 2005. Use of 16S rRNA gene terminal restriction fragment analysis to assess the impact of solids retention time on the bacterial diversity of activated sludge. Appl. Environ. Microbiol. 71:5814-5822.

59. Schulz, H.D. 2000. Quantification of early diagenesis: dissolved constituents in marine pore water, p:85-128. *In* Schulz, H.D. and M. Zabel (ed.), Marine geochemistry. Springer.

60. Singh, B.K. and N. Thomas. 2006. Multiplex-terminal restriction fragments length polymorphism. Nature Protocols. 5:2428-2433.

61. Tiedje, J.M., S.Asuming-Brempong, K.Nüsslein, T.L. Marsh amd S.J.Flynn. 1999. Opening the black box of soil microbial diversity. Appl. Soil Ecol. 13:109-122.

62. Van Wezel, M.C. and W.A. Kosters. 2004. Nonmetric multidimensional scaling: Neutral networks versus traditional techniques. Intelligent data analysis. IOS Press. 8:601-613.

63. Volkenborn, N., S.I.C. Hedtkamp, J.E.E. van Beusekom and K. Reise. 2007. Effects of bioturbation by lugworm (*Arenicola marina*) on physical and chemical sediment properties and implications for intertidal habitat succession. Estuatine, Coastal and Shelf Sci. 74:331-343.

64. Volkenborn, N., L. Polerecky, S.I.C. Hedtkamp, J.E.E. van Beusekom and D. de Beer. 2007a. Bioturbation and bioirrigation extend the open exchange regions in permeable sediments. Limnol. Oceanogr. 52:1898-1909.

65. Volkenborn, N. and K. Reise. 2007. Effects of *Arenicola marina* on polychaete functional diversity revealed by large-scale experimental lugworm exclusion. J. Sea Res. 57:78-88.

66. Von Wintzingerode, F., U.B. Göebel and E. Stackebrandt. 1997. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS. Microbial. Rev. 21:213-229.

67. Young, F.W. 1985. Multidimensional scaling. *In* Kotz-Johnson (ed.), Encyclopedia of statistical sciences, vol. 5. John Wiley & Sons, Inc. http://forrest.psych.unc.edu/teaching/p208a/mds/mds.html.

Chapter 2

Terminal restriction fragment length polymorphism,
a method evaluation for the intertidal sediments

2.1. Abstract

T-RFLP has gained recently a broad application as a genetic fingerprinting technique in biodiversity research at which one TRF represents one operational taxonomic unit (OTU). In this study, we tested the method on marine tidal soft sediments from the North Sea to evaluate the variation in the data with respect to variations in individual steps of the protocol. Restriction enzyme digest and digest analysis on a capillary sequencer showed a dissimilarity of about 20% and 10% in the obtained replicate profiles describing one pooled amplicon from one DNA sample after binning with a fixed window size of 0.5 and 1 base pair respectively with the starting point 50.25 and 50.50 bp. Independent T-RFLP analyses from one DNA sample indicated that the biases in individual PCR reactions did not increase the dissimilarity. The dissimilarity is partly caused by an imperfect binning. A high resolution window size of 0.5 bp, the starting point (50.25, 50.20, 50.30 and 50.65 bp) gave no perfect binning result. Some of identical TRFs were always binned into two different TRFs, thus creating an additional OTU. A window size of 1 bp with starting point 50.50 bp gave similar dissimilarities. Although our results indicate the requirement of an improved binning technique to utilize the full biodiversity information in the profiles, the current T-RFLP technique clearly detected the biological variation in adjacent small sediment layers and can be used to characterize the microbial community in individual sediment layers.

2.2. Introduction

Terminal restriction fragment length polymorphism (T-RFLP) is a PCR-based fingerprinting method. Modified from amplified rDNA restriction analysis (ARDRA), the initial steps of DNA isolation, PCR amplification and restriction enzyme digest are similar in both techniques. In T-RFLP, the primers are labeled with fluorescent dyes and only the fluorescent terminal restricted fragments (TRFs) are detected and quantified by a high resolution capillary electrophoresis on an automated DNA sequencer. As this method relies on variation in restriction sites among 16S rRNA gene sequences, the bacterial diversity of a community is determined as a pattern. The pattern consists of fluorescent TRFs with unique length sizes in base pairs (bp) and intensities in relative fluorescent units (rfu) (Liu *et al.*, 1997).

T-RFLP has been established as a powerful fingerprinting method for the rapid comparison of 16S rRNA genes based bacterial diversity (Liu *et al.*, 1997; Marsh, 1999; Tiedje *et al.*, 1999) with a robust ability and reproducibility (Moeseneder *et al.*, 1999; Dunbar *et al.*, 2000; Osborn, 2000). It successfully differentiates between microbial communities when the optimal statistically method is used in the study case, e.g. multivariate analysis (Blackwood *et al.*, 2003). This method had been applied for marine samples (Moeseneder *et al.*, 1999), soil samples (Clement *et al.*, 1998; Dunbar *et al.*, 2000; Osborn *et al.*, 2000, Blackwood *et al.*, 2003; Osborn *et al.* 2006), and in the activated sludge from the aeration tank, enrichment sludge from laboratory, aquifer sand from the groundwater and the gut of termite *Reticulitermes flavipes* (Liu *et al.*, 1997).

The quantitative data are sensitive to the technical variations which may arise from several causes. A minority microbial population may not be represented because the DNA template that is used in PCR represents a small fraction of the total community DNA (Liu *et al.*, 1997). A low abundant species in nature may become the most abundant species presented in the amplicon after PCR, due to the different copy number of 16S rRNA genes and preferences of PCR conditions, formation of PCR artifacts such as chimeric sequences and heteroduplex fragments, differential cell lyses and DNA extraction bias (Frey *et al.*, 2006; Dunbar *et al.*, 2001; von Wintzingerode *et al.*, 1997). The phylogenetic resolution is also limited since one TRF may be generated from multiple taxa (Blackwood *et al.*, 2007; Engebretson and Moyer, 2003; Dunbar *et al.*, 2001; Liu *et al.*, 1997). Engebretson and Moyer (2003)

showed that a mean of 9.1 to 18.5 different sequences of a set 4600 16S rRNA gene sequences could generate one TRF. Virtual digestions performed with MICA 3 also indicated that one TRF could represent about 1 to thousand uncultured bacterial species. TRFs are excluded if they are outside of the determined size range and TRFs are not being detected if their amount is below the determined fluorescence threshold, thus the detected richness has clearly limits (Blackwood *et al.*, 2007). Routinely occurring small pipetting errors, restriction enzyme digestion, TRFs separation, raw data analysis or statistical analysis are also potentially bias causes (Dunbar *et al.*, 2001). Due to those technical variations, the interpretation of the bacterial community can be underestimated or overestimated based on the observed TRFs of an environmental sample, whereby a TRF corresponds to an operational taxonomic unit (OTU). Therefore Blackwood *et al.* (2007), Saikaly *et al.* (2005), Dunbar *et al.* (2001) and Liu *et al.* (1997) suggested that the quantitative number in T-RFLP datasets should be interpreted as a reflection of differences in bacterial community composition after the amplification rather than a true difference in bacterial community diversity.

A crucial step influenced by humans is the raw data analysis. It can be seen as peak analysis. As a TRF is presented by a peak, the differentiation of a true peak and a noise peak is an important initial step. Concerning to that, the first objective in this study was to optimize the parameter options for peak detection which are available in the GeneMapper® Software v3.7: polynomial degree of peak, peak window size, peak amplitude threshold and smoothing. The quality of the capillary electrophoresis device for fragment separation was also tested.

Imprecision of the fragment size calling of peak is also a problem, since one identical peak can be assigned to different peak sizes in replicate raw profiles. This imprecision may arise from, e.g., the different mobility of internal size standards or different running conditions in the capillary electrophoresis tubes (Dunbar *et al.*, 2001; Hewson and Fuhrman, 2005). To minimize the total differences between replicate profiles within a sample due to this imprecision, combining the TRFs into a defined length size or window size is necessary (Hewson and Fuhrman, 2005). This process is called binning. It has two variables, the starting point and the window size. One problem related to binning is the distribution of the same TRF into two different range length sizes due to rounding up and down the decimal number of TRF length size. A single and wide peak could be split into separate adjacent bins by computer

software. Therefore the peak splitting due to binning need to be more considered, especially for studies which compare datasets across gradients or over time-series analysis. The appearance or disappearance of different fragments may be attributed by fragment movement into adjacent bin (Abdo *et al.*, 2006; Hewson and Fuhrman, 2005). To minimize the false distributed TFRs, the binning related to the starting point of the window sizes was examined as the second objective.

To relate both objectives to the reproducibility of peak number, peak size (in bp) and peak intensity, replicate profiles of a sample were generated with two different strategies: nine replicate profiles and triplicate profiles (see Fig. 1.2 in materials and methods). These replicate strategies were purposed to address the technical variation on the level of PCR, restriction digestion and the capillary electrophoresis analysis. The technical variations expected to be observed from these experimental strategies were compared with the biological variation. The biological variation referred to species richness (presence of 5´-TRFs and 3´-TRFs) and abundance (relative area). The intertidal sediments from the Königshafen, the Sylt Island in the North Sea are well known for a steep gradient from oxic to anoxic habitats in the upper sediment layers and were selected to evaluate the T-RFLP method.

2.3. Materials and methods

2.3.1. Sample area and sediment samples

Sediment samples were collected in October 2005 from populated and populated area developed by Volkenborn *et al.* (2007) on a low intertidal sandy flat in the Königshafen at the northern end of the island of Sylt in the North Sea, Germany ($55^o02´N$, $8^o26´E$). The not populated area of 20 m x 20 m was achieved by inserting a 1mm meshed net in 10 cm depths (Volkenborn *et al.*, 2007; Volkenborn *et al.*, 2007a). The populated area was approximately inhabited by 20-30 individuals/m^2 and characterized by fecal casts at the sediment surface. Vice versa, the not populated area had smooth surface indicating the absence of *A. marina* (Figure 2.1.A). The population was found to play an important role for physical and chemical processes and the benthic community (Volkenborn *et al.*, 2007; Volkenborn *et al.*, 2007a). The biochemical habitat description is presented in Table 2.1.

Photo and scheme : N. Volkenborn
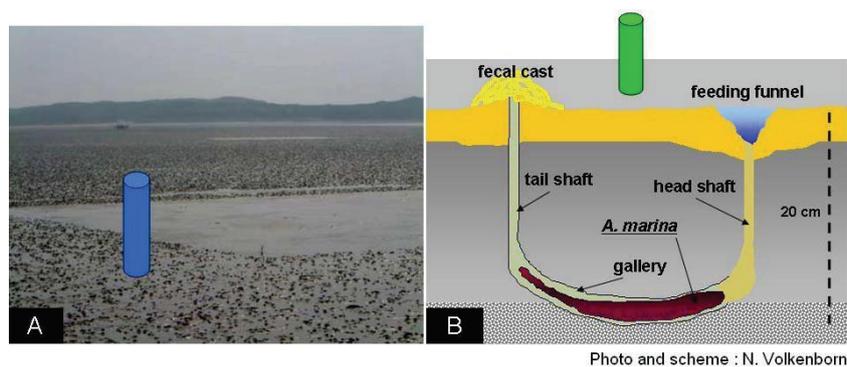
Fig. 2.1. (A) Smooth and non-smooth surface of the plot indicating the absence and presence of lugworm *A.marina*, respectively. One sample core was put into the populated area. (B) Scheme of the U-shaped lugworm burrow and one core collected the bulk sediments surrounding the burrow.

A piston core (2.5 cm x 20 cm) was taken from the populated area (Fig. 2.1.A) and coded as sample 6. A piston core collected the bulk sediments surrounding the U-shaped lugworm burrow (Fig. 2.1.B) and coded as sample 13. The sediments samples were immediately brought to the harbor laboratory and sliced within five hours into layers with a thickness of 0.5 cm for sample 6 and 1 cm for sample 13. Four subsamples were selected from sample 6 (6.3, 6.4, 6.5 and 6.6 for 1.5-2 cm, 2-2.5 cm, 2.5-3 cm and 3-3.5 cm depth, respectively) and five subsamples from sample 13 (13.1, 13.3, 13.5, 13.7 and 13.9 for 1-2 cm, 3-4 cm, 5-6 cm, 7-8 cm and 9-10 cm depth, respectively). The sliced sediment samples were frozen immediately and stored at -20°C until further analysis.

 Table 2.1. Sediment biogeochemistry in the sample area (from Volkenborn *et al*., 2007 and 2007a).

		Bioturbated plot	Non-bioturbated plot
Emersion period		9-10 h / tide	
Main hydrodynamic force		Tidal current	
Mean tidal height		$\approx$ 1.8 m	
Salinity		27.5 $^\circ/_{\circ\circ}$ in spring	
		31 $^\circ/_{\circ\circ}$ in summer	
Average *A. marina* density		18-30 ind / m^2	-
Grain size of fine sand sediment		0 -1 cm : 204 µm	0 -1 cm : 190 µm
		1 - 5 cm : 218 µm	1 - 5 cm : 206 µm
Fine fraction (particle size <63 µm) at 0 - 5 cm depth		<1% dry wt	1% - 2.5% dry wt
Water content in 0 – 5 cm depth		16.20%	19.90%
Sediment permeability		2.6 X 10^{-12} m^2	<1.0 X 10^{-12} m^2
Approximately porewater profile through 0-10 cm depth	Ammonium	<100 µM	<150 µM
	Nitrite	<0.3 µM	<0.25 µM
	Nitrate	<2.5 µM	<2.5 µM
	Phosphate	<10 µM	<15 µM
	Sulphide	<25 µM	<150 µM
Average oxygen penetration		<4 cm depth	<1 cm depth

2.3.2. Terminal restriction fragment length polymorphism (T-RFLP)

DNA extraction. Total genomic DNA from 0.5 g sediment samples was extracted using the FastDNA SPIN Kit for soil (Qbiogene, Carlsbad, USA) following the Manufacturer's protocols with a minor modification. The sample was centrifuged twice at 14.000 g for 30 seconds and the extracted genomic DNA was diluted twice in 25 µl PCR water. Extracted genomic DNA was measured qualitatively and quantitatively with agarose electrophoresis on 1.5% gels and a NanoDrop ND-1000 UV spectrophotometer (PEQLAB Biotechnology, Erlangen, Germany).

Two different strategies were applied for producing replicate profiles. Nine replicate profiles were generated from sample 6. The amplicons from 3 individual polymerase chain reaction (PCR) mixtures were pooled and digested in three individual reactions. Each of these was afterwards divided into three samples for capillary electrophoresis (Fig. 2.2.A). Three replicate profiles were generated from sample 13. Three individual PCR mixtures were then digested in three individual restriction enzyme digestions and fragment runs (Fig. 2.2.B).
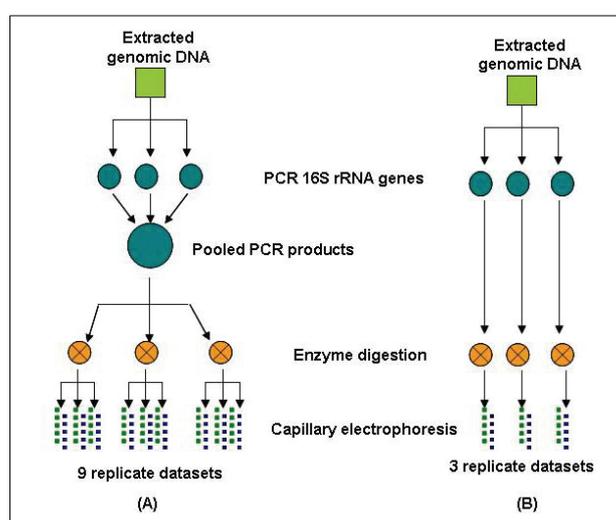
Fig. 2.2. Strategies for generating (A) the nine replicate profiles and (B) the triplicate profiles. The nine replicate profiles were generated with primer pair 6-FAM-27F/HEX-1492R and analyzed with internal standard marker Genescan500™-500-ROX™, while the triplicate profiles with 6-FAM-27F/HEX-907R and internal standard marker MapMarker1000®.

Polymerase chain reaction (PCR). The target bacterial 16S rRNA genes were amplified using two different primer pairs: 6-FAM-27F (AGAGTTTGATCCTGGCTCAG; Amann *et al*., 1995) and HEX-1492R (GGTTACCTTGTTACGACTT) for sample 6; 6-FAM-27F and HEX-907R (CCGTCAATTCCTTTRAGTTT; Muyzer *et al*., 1995) for sample 13. The forward primer was labeled at the 5´-terminus with 6-Carboxyfluorescein and the reverse

primer with 6-carboxy-2',4,4',5',7,7'-hexachloro-fluorescein. The 25µl-PCR reaction consisted of 12.5 µl PCR Master Mix (Promega), about 1 ng template of genomic DNA and 4 pmol/µl of each forward and reverse primer. PCR was performed in a Mastercycler Personal (Eppendorf) with the following thermal conditions: 5 minutes initial denaturation at 94°C and 35 cycles consisting of 4 minutes denaturation at 94°C, 1 minute annealing at 52°C for 6-FAM-27F/HEX-1492R or 50°C for 6-FAM-27F/HEX-907R and 1 minute elongation at 72°C. At the end, elongation was extended at 72°C for 10 minutes. The PCR products were passed through Sephadex™ G-50 Superfine columns according to a protocol suggested by Applied Biosystems (California, USA).

Enzyme digestion and capillary electrophoresis. Approximately 100 ng of PCR product was digested in 20 µl reaction volumes with 5 U of the restriction enzyme *Alu*I for 3 hours at 37°C and afterward the enzyme was inactivated at 65°C for 20 minutes. After desalting through Sephadex™G-50 Superfine, 5 µl digested product was mixed with 20 µl standard internal marker:deionized formamide (1:60 v/v) and denatured at 95°C for 10 minutes and placed immediately on ice. The marker Genescan500™-500-ROX™ (ABI, California, USA) was used for sample 6 and MapMarker1000® (BioVentures Inc. Murfreesboro, TN) for sample 13. The digested product was loaded on a capillary ABI Prism 3130XL genetic analyzer for TRFs separation. GeneMapper® Software v3.7 (ABI) was used for calculating the length size of observed 5´-TRFs and 3´-TRFs in bp with the local southern method and the amount as intensity in rfu.

2.3.3. Peak separation

The profile of two internal standard markers, Genescan500™-500-ROX™ Standard and MapMarker1000®, were analyzed using a default parameter as provided by the software as presented in Table 2.2. The peak and pattern distribution was reviewed based on the electropherogram and the position in the size match editor of the software.

Table 2.2. Default parameter values in the GeneMapper® software v3.7 (ABI)

Parameter	Value
Peak amplitude threshold	50 rfu
Min peak half width	2 data points
Polynomial degree	3
Peak window size	15 data points
Smoothing	None
Baseline window	51 data points
Size calling method	Local Southern Method

2.3.4. Peak detection

According to Table 2.2, the GeneMapper® Software v3.7 has options for defining a peak as a presentation of a TRF: polynomial degree, peak window size in data points, peak amplitude threshold in rfu, smoothing, minimal peak half width and baseline window in data points and size calling method. In this study we adjusted polynomial degree, peak window size in data points, peak amplitude threshold in rfu and smoothing. The adjustment was done stepwise and the effect on peak number and relative area (rfu x time) was analyzed using two-way ANOVA with replicates. Minimal peak half width, baseline window and size calling remained the default value.

Minimal peak half width defines what constitutes a peak. The range value is 2 to 99, at which a low number for narrow peaks and high number for ignoring noise. Baseline window adjusts the baselines of all detected dye colors to the same level for an improved comparison of relative signal intensity. Size calling method determines the molecular length of a known fragment (user guide of GeneMapper® software v3.7). The local southern method had been evaluated by Osborn *et al*. (2000) as the best calling algorithm for estimating the length size of unknown fragments. The algorithm counts the size of unknown fragments by using a reciprocal relationship between fragment length and mobility. The best-fit line value is determined by using four fragments closest in size to the unknown fragment (user guide of GeneMapper® software v3.7).

The sensitivity of peak detection depends on polynomial degree and peak window size. The peak detector calculates the first derivative of a polynomial curve which is fit to the data within a peak window size. The peak window size has a centre of data points of each peak in the analysis range. The sensitivity increases with

larger polynomial degree and smaller peak window size value. Larger polynomial degree allows the curve to more closely approximate the signal and smaller window size value allows the curve to better fit the underling data (user guide of Genemapper® software v3.7). To adjust the polynomial degree and peak window size, we used standard marker peaks as an effort to have a distinguishable peak from noise. Three polynomial degrees of 2, 3 and 4 was applied to 36 profiles of Genescan500™-500-ROX™ with other parameters as same as default (Table 2.2).

Adjustment of peak window size was performed using the known peak of MapMarker1000®. Peak window size was predicted from peak full width at a half maximum height. The peak full width at a half maximum height was calculated by dividing the range of observed data points between 50 and 900 bp with the base pair range between 50 bp and 900 bp for MapMarker1000®. The value was then halved to get the peak full width at a half maximum height. With a default parameter (Table 2.2), the peak full width at a half maximum height of 30 profiles of MapMarker1000® was 6.

As the peak window size is supposed around 1 to 2 times the full width at a half maximum height of the peaks (user guide of Genemapper® software v3.7), known fragments in MapMarker1000® were further analyzed by applying a serial peak window size 9, 13 and 15 data points while other parameters as same as default (Table 2.2) but with a chosen polynomial degree above. Number of observed peaks and average area in a particular profile was compared for determining the effect.

Smoothing optimizes peak size and reduces the number of false peaks detected (user guide of Genemapper® software v3.7). The amplitude (y-axis value) of a true peak changes over several data points rather smoothly as a function of the x-axis value, whereas noise is seen as rapid and random amplitude change from point to point within the signal. In smoothing, the signal data points are averaged so that an individual point that is higher than the adjacent point is reduced and vice versa. A peak amplitude threshold reports the peaks with heights that exceed the threshold value in the table data (user guide of Genemapper® software v3.7).

Smoothing and peak amplitude threshold were adjusted using profiles from internal standard marker and samples to determine whether peak amplitude threshold and smoothing could limit technical variation and to detect reproducibility of peak number and area within replicate profiles. Three smoothing options are

available in the software: none, light and heavy smoothing. None and light smoothing was chosen, because all observed samples produced half width peak number as same as internal standard peak. It indicated that the peak shape of the sample was as narrow as the internal standard in the electropherogram profile.
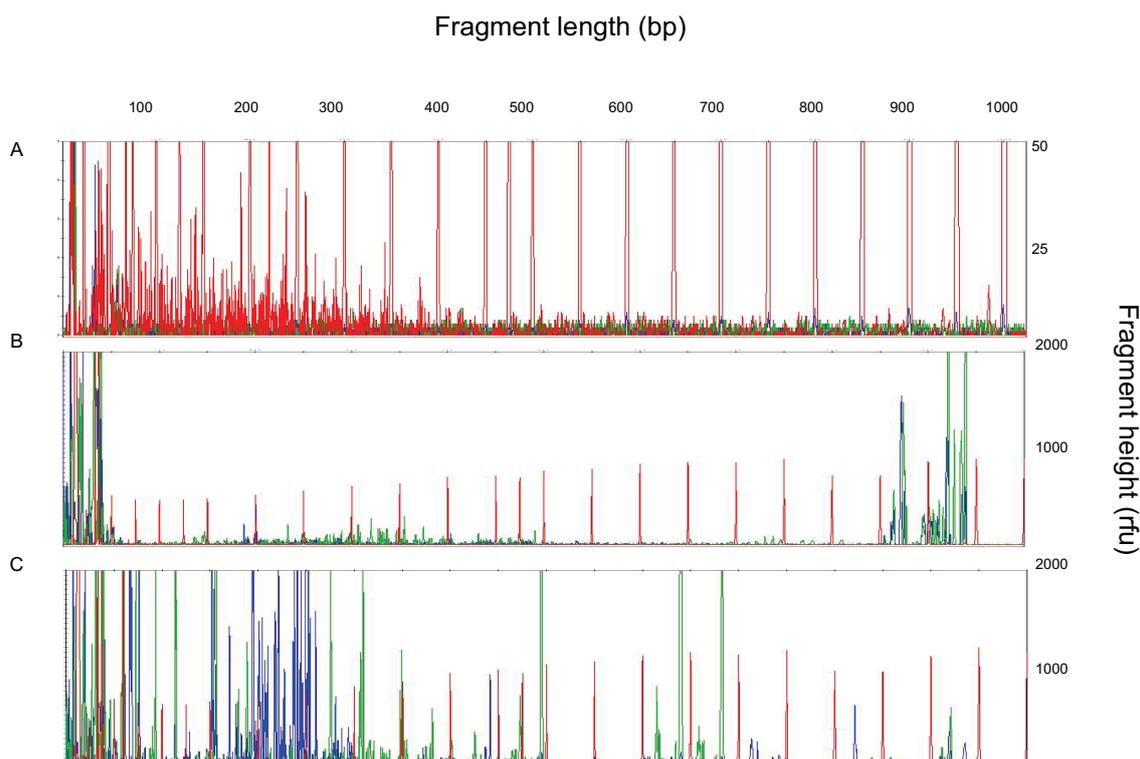
Fig. 2.3. Electropherogram of (A) MapMarker1000® with baseline noise in the range of 20 rfu. (B) Non-digested sample 13 at the depth 7-8 cm with baseline noise up to 20 rfu. Amplicons generated with primer pair 6-FAM-27F and HEX-907R are accumulated in size range around 900 bp. (C). Amplicons are digested with *Alu*I and distributed over the electropherogram.

The internal marker had a baseline noise in the range of 20 rfu and the non-digested sample (control) showed peaks of up to 200 rfu. Peaks in this non-digested sample were assumed as false-positive peaks (Fig. 2.3). In chromatography, the signal ratio of peak to noise should be at least 3:1. Therefore applied peak amplitude threshold for internal marker was only 50 rfu, whereas samples were 50 rfu and 100 rfu. A factorial analysis was performed for the peak amplitude threshold and the smoothing as below:

<u>Adjusted parameter :</u>
Peak amplitude threshold   50 and 100 rfu
Smoothing   none and light

<u>Fixed parameter :</u>
Minimum peak half width   2 data points
Polynomial degree   3
Peak window size   15 data points
Baseline window   51 data points
Size calling method   Local Southern Method

2.3.5. Binning

The T-RFLP profiles were analyzed by applying optimized values of peak detection parameters and exported in a text file containing 3 columns: sample name, size in bp and relative area (rfu x time). The T-RFLP datasets was defined by the labeled primer: 5´-TRF datasets and 3´-TRF datasets. These raw TRF datasets were binned using R software (www.r-project.org) with a language program written by A. Ramette (unpublished).
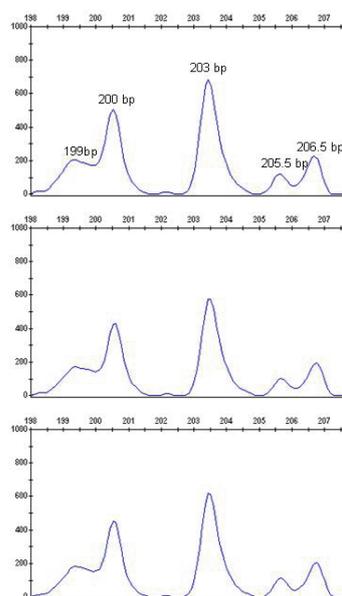
Fig. 2.4. Partial electropherograms from 3 out of 9 replicate profiles sample 6 at 3 cm depth after *Alu*l digestion.

Table 2.3. A partial 5´-TRF datasets from 9 replicate profiles sample 6 at 3 cm depth after *Alu*l digestion before binning. Numbers in the table were normalized area of 5´-TRFs in their particular profile.

TRF length (bp)	The 9 replicate datasets								
	F-6.6-A1	F-6.6-A2	F-6.6-A3	F-6.6-A4	F-6.6-A5	F-6.6-A6	F-6.6-A7	F-6.6-A8	F-6.6-A9
199.28	0	0	0	0	0	0	0	1.18	0
199.29	0.95	0	0	0	0	0	0	0	0
199.36	0	1.16	1.15	0	0	0	0	0	1.13
199.37	0	0	0	1.15	1.16	1.11	0	0	0
199.45	0	0	0	0	0	0	1.01	0	0
200.43	0	0	0	0	2.02	2.18	2.15	0	0
200.51	0	0	1.86	2.06	0	0	0	0	0
200.52	1.48	0	0	0	0	0	0	1.88	1.84
200.6	0	1.91	0	0	0	0	0	0	0
203.42	0	0	0	0	0	2.47	0	0	0
203.43	0	0	2.45	2.51	0	0	0	0	0
203.44	1.91	0	0	0	2.49	0	0	0	0
203.45	0	0	0	0	0	0	0	2.42	0
203.47	0	0	0	0	0	0	0	0	2.38
203.51	0	0	0	0	0	0	2.43	0	0
203.52	0	2.44	0	0	0	0	0	0	0
205.58	0.28	0	0	0	0	0	0	0	0
205.61	0	0	0	0	0	0	0	0.35	0
205.65	0	0	0.36	0	0	0	0.33	0	0
205.66	0	0.36	0	0	0	0	0	0	0
206.68	0	0	0	0	0	0	0	0	0.68
206.7	0.55	0	0	0	0	0	0	0	0
206.73	0	0	0	0	0	0	0	0.69	0
206.75	0	0	0.69	0	0	0	0	0	0
206.76	0	0	0	0.74	0	0	0	0	0
206.77	0	0.71	0	0	0	0	0	0	0
206.79	0	0	0	0	0.72	0	0	0	0
206.83	0	0	0	0	0	0.71	0	0	0
206.84	0	0	0	0	0	0	0.69	0	0

A detailed view on the electropherogram of one set of the nine replicate profiles revealed a high reproducibility of the method (Fig. 2.4), but the corresponding fragment sizes had an imprecision for size calling about 0.5 bp (Table 2.3). The same peaks in the nine replicate profiles were sized differently within 0.5 bp. This experimental data suggested binning with a fixed window size of 0.5 bp, because the window size should be as wide as the imprecision (Hewson and Fuhrman, 2005). It meant that TRFs which varied about 0.5 bp in length size were considered as an identical TRF (Fig. 2.5).
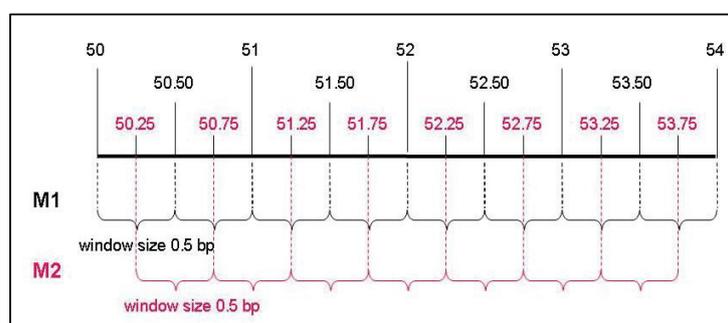
Fig. 2.5. Scheme of binning strategy I with a fixed window size 0.5 bp. Other binning strategies were generated from this scheme by shifting the starting point.

Considering that binning potentially split a peak into separate adjacent bins and attributes the appearance or disappearance of a TRF, five starting points of a fixed window size of 0.5 bp were tested to identify the optimal starting point for combining the same TRFs in the nine replicate profiles. The five starting points were 50.00, 50.10, 50.20, 50.30 and 50.40 bp and called respectively as binning strategy I, II, III, IV and V (Table 2.4). As a comparison a fixed window size of 1 bp was also performed with one starting point of 50.00 bp.

Table 2.4. Five different binning strategies which differ in the binning starting point. Each binning strategy produced 2 different output tables with a different binned starting point of TRFs length size.

Binning strategy	TRF lenght size (bp) as binning starting point	The starting point of output table (bp)			
		Table M1		Table M2	
		0.5 bp*	1 bp*	0.5 bp*	1 bp*
I	50	50	50	50.25	50.5
II	50.1	50.1		50.35	
III	50.2	50.2		50.45	
IV	50.3	50.3		50.55	
V	50.4	50.4		50.65	

* a fixed window size

The binning program yielded two binned output tables called M1 and M2 that differ in the starting point by one half of the window size (i.e. 0.25; Fig. 2.5 and Table 2.4). For example, the binning strategy I with a fixed window size 0.5 bp resulted in output table M1 with the starting point 50.00 bp for the first peak and table M2 with the starting point 50.25 bp. From each table a correlation matrix (c1 and c2, respectively) was calculated among samples. The means of the correlation matrices were compared to determine the one offering on average the highest similarity values among samples. Significance for the difference was tested by the parametric Student t test and by the non-parametric Wilcoxon test. The binning table associated with the correlation matrix offering the highest similarities among samples was then chosen for further statistical analyses. During the binning process, the raw profiles were normalized within a profile by diving the area of each peak by the total peak area in that particular profile. The 5´-TRF and 3´-TRF in the output table are presented by their size and their relative area. Dataset normalization counts the variations that may arise e.g. from DNA loading and sample sensitivity to the molecular steps (Osborne *et al.*, 2006; Abdo *et al.*, 2006; Saikaly *et al.*, 2005; Dunbar *et al.*, 2001).

2.3.6. Statistical analysis

The biological variations of the TRF datasets were analyzed by multivariate analysis: Non-metric Multi Dimensional Scaling (NMDS) with the statistical software packages Primer 5 for windows version 5.2.0 (Primer-E Ltd, Plymouth, UK) and PAST 1.38 (Palaeontological Statistics, http://folk.uio.no/ohammer/past/), a free software package for education and data analysis. Using Primer 5, the similarity matrix for NMDS and analysis of similarity (ANOSIM) was calculated based on the presence of 5´-TRFs or 3´-TRFs (the richness) and their abundance in relative area using Bray-Curtis coefficients without data transformation and standardization.

In order to get comparative ordinations, the similarity was counted from quantitative datasets (relative area) and binary datasets (absence/presence) with the Bray-Curtis coefficient using Primer 5 and the Jaccard coefficient using PAST respectively, since the Jaccard coefficient is not available in Primer 5. A hundred random restarts were used for calculating the iterative algorithm, because NMDS is sensitive to the initial configuration (Kenkel and Orloci, 1986; Rees *et al.*, 2004). The software Primer presents the ordination with the lowest stress value obtained in 100 calculations. Stress is a measure of deviation for ´a goodness of fit´ of the iterative algorithm (Kenkel and Orloci, 1986; Rees *et al.*, 2004). A stress value above 0.2 is a random ordination, less than 0.2 is a useful 2 dimensional ordination and less than 0.1 is an ideal ordination without potential misinterpretation (Clarke and Warwick, 2001). The random restart with PAST was done manually until obtaining a stress value as low as possible.

Contribution of 5´-TRF and 3´-TRF to the similarity within replicate profiles and to the dissimilarity between depth dependent sub samples was calculated with the Simper analysis. The list of 5´-TRFs or 3´-TRFs was cut off at 90% contribution; fragments will be listed in decreasing order of their importance in contributing to the average dissimilarity between two groups until 90% of the dissimilarity is explained.

2.4. Results

2.4.1. Peak separation

The internal markers were detected and the peaks were sized applying the GeneMapper® software v3.7 with the local southern method as prerequisite to determine the size of TRFs. The Genescan500™-500-ROX™ has 16 peaks of 35, 50, 75, 100, 139, 150, 160, 200, 250, 300, 340, 350, 400, 450, 490 and 500 bp. Only eight out of 120 profiles had the same peak numbers as described by the company (Fig. 2.6.A). The majority (118 profiles) had more than 16 peaks (Fig. 2.6.B). The peak variation was mainly present within size range 50 – 75 bp in 107 profiles (Fig. 2.6.C) and 50 – 100 bp in 5 profiles (Fig. 2.6.D). The peaks within size range 100 - 500 bp were separated and sized constantly in the 120 profiles (Fig. 2.6.A).

A
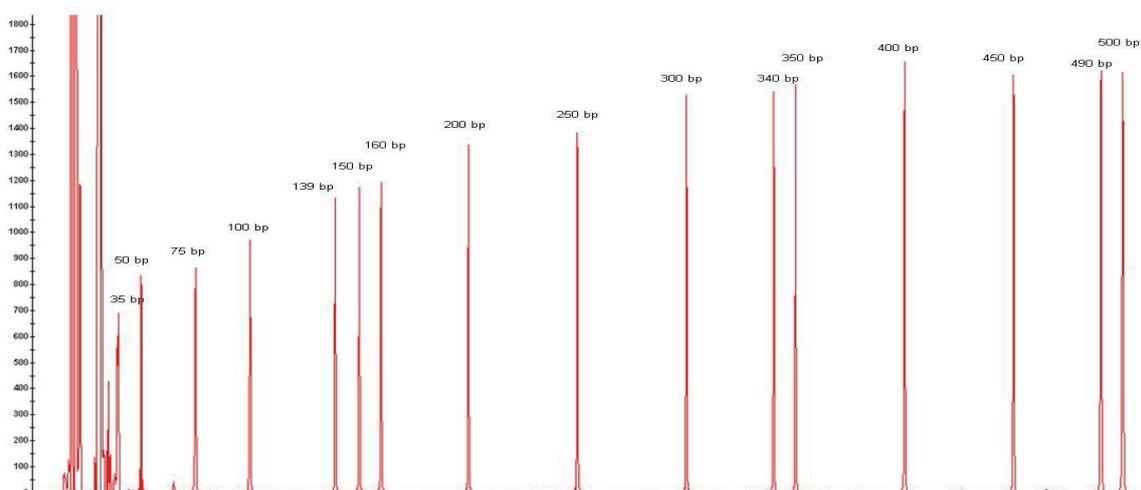
B

Peak number within 50 - 500 bp	Number of profiles
15	8
16	81
17	25
18	1
20	5
Total	120

Peak variation in range size	Number of profiles
50 - 75 bp	107
50 - 100 bp	5
same as recommended marker	8
Total	120

C
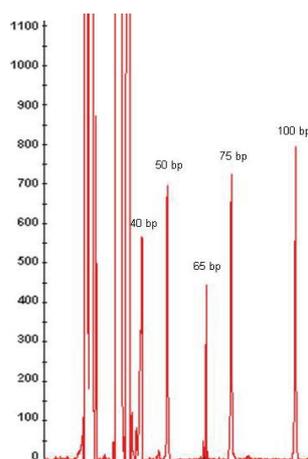
D
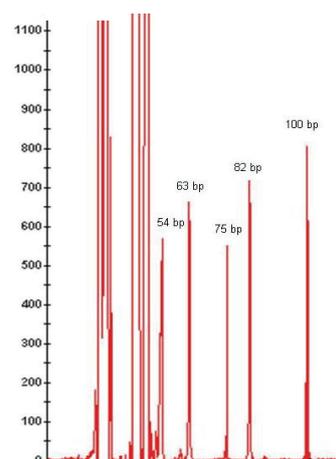
Fig. 2.6. (A). The electropherogram of the internal standard marker Genescan500™-500-ROX™. (B). Table shows the variations in peak numbers of 120 profiles. The electropherograms present the peak variation in region (C) 50-75 bp and (D) 50-100 bp.

Looking at to the size match editor of the GeneMapper® software v3.7 (ABI), the 8 perfect profiles (Fig. 2.6.A) and the 107 profiles with problems in the region 50-75 bp (Fig. 2.6.C) had the same pattern. The 50 bp and 75 bp peak was situated at the x-axis 1500-1600 (time) and 1800-1900 (time) respectively (Fig. 2.7.A), and one additional peak was constantly situated at the x-axis 1700-1800 (time) for the 107 profiles (Fig. 2.7.B).
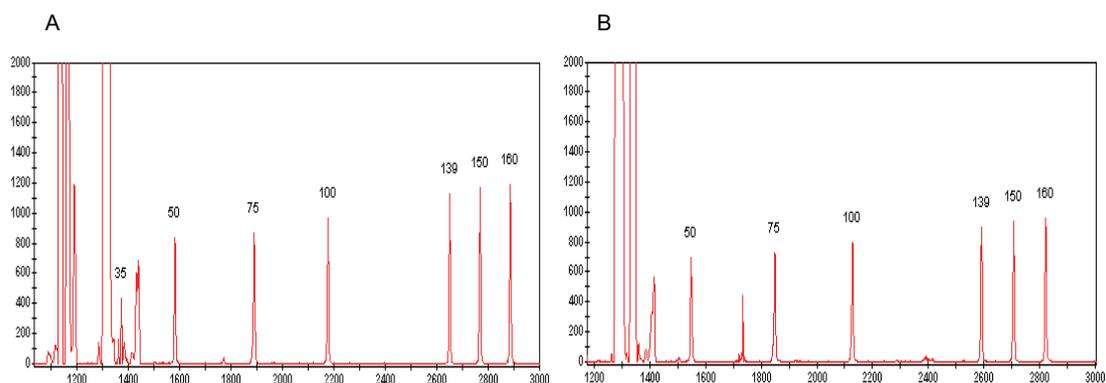
Fig. 2.7. Size match editor pattern of: (A) one profile of the 8 perfect profiles and (B) one profile of the 107 profiles which have peak variation within size range 50 – 75 bp. Height peak in rfu and x-axis in time.

The size match editor profile pattern of the 5 profiles which had peak variation within size range 50 – 100 bp (Fig. 2.6.D) differed with those above. The 50 bp and 75 bp peak was situated at x-axis 1300-1400 (time) and 1700-1800 (time), respectively (Fig. 2.8.A). By comparing to the perfect profile pattern (Fig. 2.7.A), the identification of standard peaks was possible to be edited. The 50 bp and 75 bp peak was shifted to x-axis 1500-1600 (time) and 1800-1900 (time) respectively and again one extra peak was identified between 50 and 70 bp (Fig. 2.8.B). As the consequence, the edited size calling curve was not valid anymore. The software automatically gave a warning. The edited size range under 100 bp was affected (Fig. 2.8.C and 2.8.D). It might relate to the size calling algorithm for counting the size of small unknown fragments less than 100 bp, as the local southern method has limitation of an inaccurate estimation if any of the standard peak run anomalously (user guide of Genemapper® software v3.7).
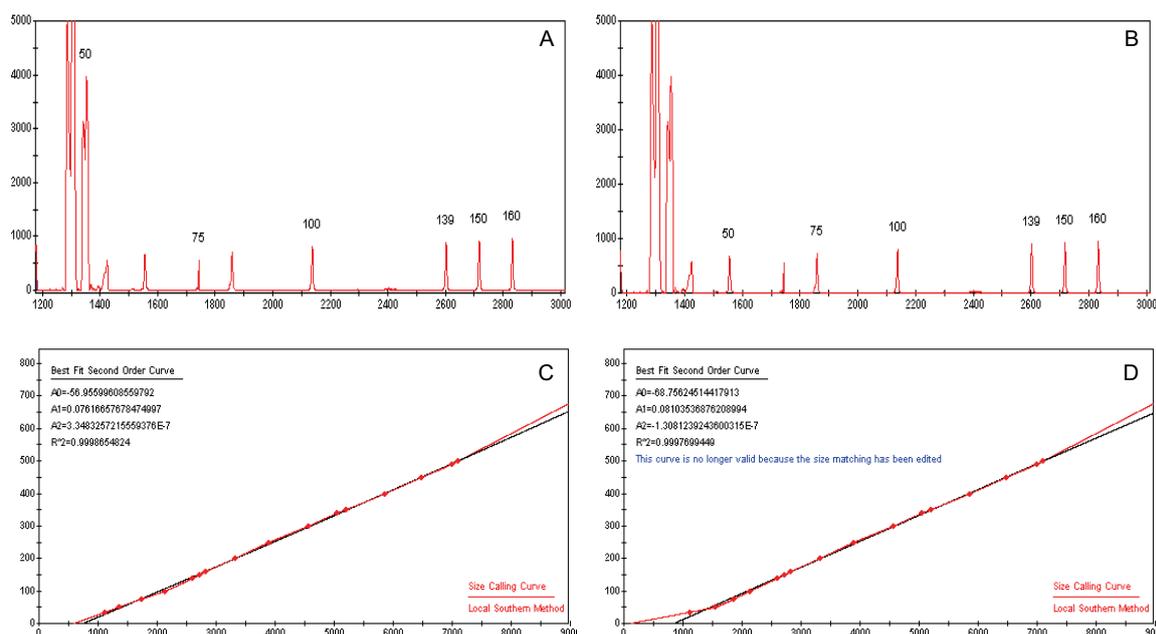
Fig. 2.8. Size match editor pattern of: (A) one profile of the 5 profiles which have peak variation within size range 50 – 100 bp before editing and (B) after editing of the peak identification. (C) Size calling curve after local southern method before editing and (D) after editing the standard peak identification. Height peak in rfu and x-axis in time.

The MapMarker1000® has 23 peaks of 50, 75, 100, 125, 150, 200, 250, 300, 350, 400, 450, 475, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950 and 1000 bp. With default parameters (Table 2.2), all of 60 profiles produced more than 23 TRFs. Mostly peak variation occurred within size range 50 – 475 bp by the presence of extra peaks and double peaks (Table 2.5). Based on their position in the size match editor, the main peaks in all profiles were situated constantly at the same range x-axis (Table 2.5).

Due to this observations, further on the TRFs of samples were analyzed within 3 different size ranges: 50 – 500 bp, 75 – 100 bp and 100 - 500 bp for Genescan500™-500-ROX™ and   50 – 900 bp, 75 - 900 bp and 100 - 900 bp for MapMarker1000®, especially for detecting the reproducibility of 5´-TRFs and 3´-TRFs within the replicate profiles.

Table 2.5. Six profiles from 60 profiles of MapMarker1000® that produce more than 23 TRFs of internal standard marker and average position of 23 TRFs in x-axis of the size match editor GeneMapper® software v3.7. Analyzes were performed with a default parameter provided by the software.

Identified internal marker peak (bp) in six datasets						X-axis position (scan line)
1	2	3	4	5	6	
50	**50**	**50**	**50**	**50**	**50**	**2900-2950**
59,56	59,64	59,61				
67,86	67,94	67,89	67,92	67,92	67,94	
75	**75**	**75**	**75**	**75**	**75**	**3200-3250**
91,53	91,53	91,55				
	93,16					
100	**100**	**100**	**100**	**100**	**100**	**3500-3550**
100,98	101,06		101,06	100,98	100,98	
			124,12			
125	**125**	**125**	**125**	**125**	**125**	**3800-3850**
125,87			125,87	125,95		
150	**150**	**150**	**150**	**150**	**150**	**4150-4200**
151,18	151,18	151,1	151,19	151,03	151,19	
156,14	156,15	156,15				
190,33						
195,05	195,04	195,04				
199	198,99	198,99	198,99	198,99	198,99	
200	**200**	**200**	**200**	**200**	**200**	**4750-4800**
220,33	220,41	220,33	220,37	220,37	220,4	
			238,85			
249,02	249,1	249,1	249,1	249,02	249,1	
250	**250**	**250**	**250**	**250**	**250**	**5450-5500**
299,02	298,95	299,02	299,02	299,02	298,94	
300	**300**	**300**	**300**	**300**	**300**	**6100-6150**
	305,88					
309,33	309,26	309,27				
348,95	348,95	349,02	348,95	348,95	349,02	
350	**350**	**350**	**350**	**350**	**350**	**6750-6800**
399,02	399,02	399,02	399,1	399,02	399,02	
400	**400**	**400**	**400**	**400**	**400**	**7400-7450**
	449	449	449	449	449,07	
450	**450**	**450**	**450**	**450**	**450**	**8050-8100**
	467,8					
473,88	473,96	473,96	473,96	473,88	473,96	
475	**475**	**475**	**475**	**475**	**475**	**8350-8400**
499,08					499,09	
500	**500**	**500**	**500**	**500**	**500**	**8700-8750**
550	**550**	**550**	**550**	**550**	**550**	**9350-9400**
600	**600**	**600**	**600**	**600**	**600**	**9950-10000**
				649,02	649,01	
650	**650**	**650**	**650**	**650**	**650**	**10500-10550**
700	**700**	**700**	**700**	**700**	**700**	**11000-11050**
750	**750**	**750**	**750**	**750**	**750**	**11600-11650**
800	**800**	**800**	**800**	**800**	**800**	**12150-12200**
850	**850**	**850**	**850**	**850**	**850**	**12600-12650**
			874,02	873,92	863,78	
					873,97	
900	**900**	**900**	**900**	**900**	**900**	**13000-13050**
			918,01	917,9	917,92	
					919,52	
950	**950**	**950**	**950**	**950**	**950**	**13500-13550**
1000	**1000**	**1000**	**1000**	**1000**	**1000**	**13900-13950**
41	42	38	39	38	40	Total peak

2.4.2. Peak detection

Applying default setting (Table 2.2) with three different polynomial degrees of 2, 3 and 4 did not statistically influence the peak number and the relative area in the profiles of Genescan500™-500-ROX™; moreover degree 3 and 4 generated exactly the same profile (Appendix 1). The peak window size of 9, 13 or 15 data points also did not statically affect the peak number and the average relative area of the MapMarker1000® profiles (Appendix 2). The different window sizes produced highly similar numbers: the average of peak number was 37.5, 36.8 and 36.3 and the average of relative area was 6846, 6965 and 7057 (rfu x time) respectively for peak window sizes of 9, 13 and 15 data points.

Either none or light smoothing did not influence the detected peaks of Genescan500™-500-ROX™; both gave the same peak number and relative area. MapMarker1000® significantly produced lower peak number with light smoothing, but no significant difference for total relative area between none and light smoothing (Appendix 3). Light smoothing might reduce small peaks as detected as extra peaks or double peaks which were situated in size range 50 – 475 bp.

The resultant of peak amplitude threshold and smoothing significantly affected the average peak number and had no significant effect on the average relative area for the nine replicate profiles (Fig. 2.9) and the triplicate profiles (Fig. 2.10). The effect of peak amplitude threshold and smoothing on the total peak number and total relative area of each replicate profiles can be seen in Appendix 4 and Appendix 5. A peak amplitude threshold 100 rfu and light smoothing might remove true peaks and false-positive noise peaks but kept the total relative area stable. It meant that the reduced peaks were small peaks in area. Anyhow the reduction was clearly observed in the non-digested sample. Approximately 10-30% 5´-TRFs in the non-digested sample 13 for each individual sub sample were reduced while 4%-8% in the digested sample 13. In order to have as little false-positive TRFs as possible; we applied a peak amplitude threshold 100 rfu and light smoothing. The difference of 5´-TRF datasets between none and light smoothing in the sample 13.3-A3 with peak amplitude threshold 100 rfu can be seen in Appendix 6.
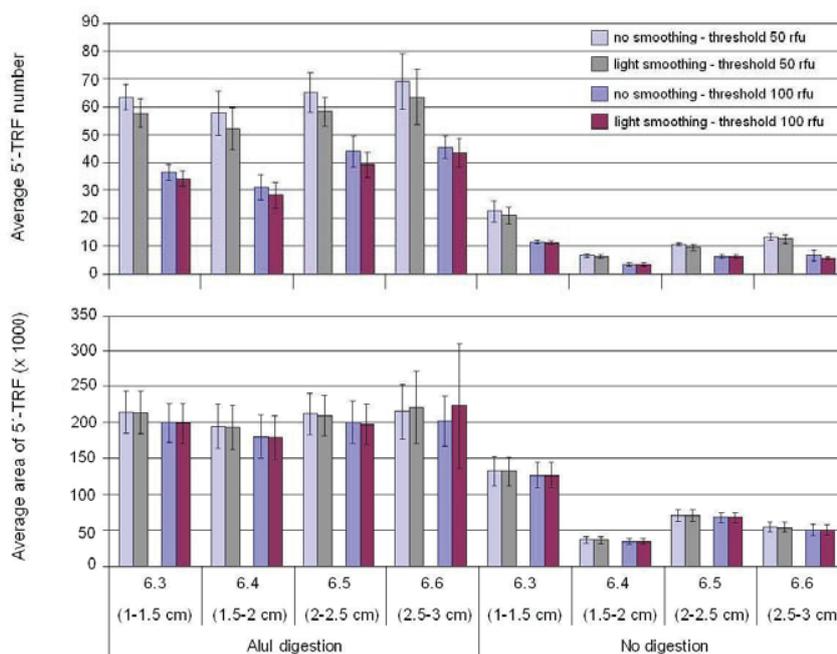
Fig. 2.9. The nine replicate profiles sized with Genescan500™-500-ROX™. Effect of peak amplitude threshold and smoothing on the average 5´-TRFs number and average area within size range 50-500 bp.
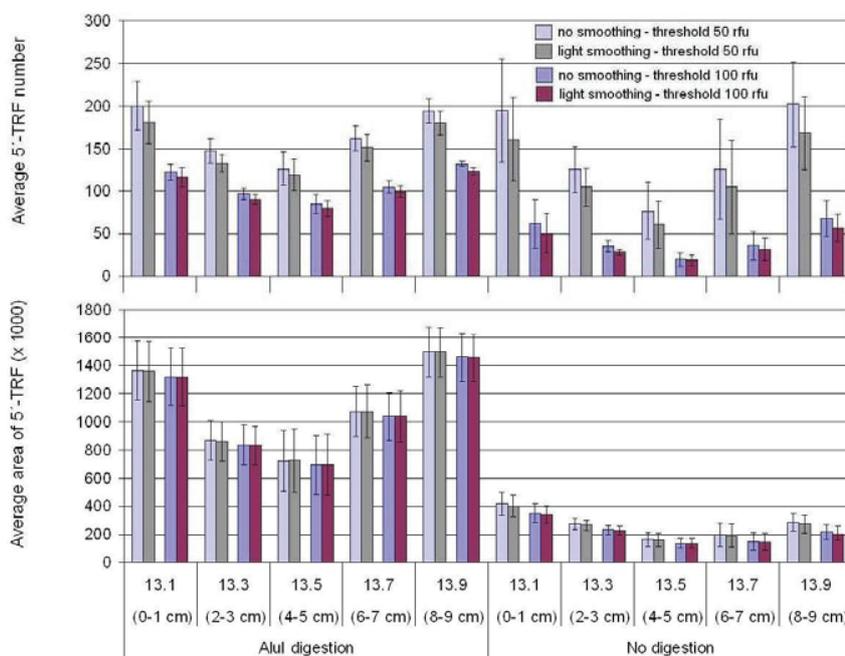
Fig. 2.10. The triplicate profiles sized with MapMarker1000®. Effect of peak amplitude threshold and smoothing on the average 5´-TRFs number and average area within size range 50-900 bp.

2.4.3. Binning strategy

The student t-test and the Wilcoxon test revealed that the binned output tables, M1 and M2 were significantly different (p-values were under 0.01). Their differences were also indicated by the large and positive value of the Wilcoxon test (W). Based on the mean of the correlation matrix (c1 or c2), the output table with the higher mean was considered to present the correct binning strategy (Table 2.6).

Table 2.6. The t-test and the Wilcoxon test results.

Binning strategy	Output table starting point M1	M2	t-test p-value	c1 mean	c2 mean	Wilcoxon test p-value	W	Chosen output table
A fixed window size 0.5 bp								
I	50	50,25	2,20E-16	0,22	0,26	2,20E-16	7,1E+08	M2
II	50,1	50,35	4,13E-05	0,26	0,25	2,20E-16	8,3E+08	M1
III	50,2	50,45	2,20E-16	0,26	0,21	2,20E-16	8,7E+08	M1
IV	50,3	50,55	2,20E-16	0,26	0,24	2,20E-16	7,8E+08	M1
V	50,4	50,65	2,20E-16	0,22	0,26	2,20E-16	6,6E+08	M2
A fixed window size 1 bp								
I	50	50,5	2,20E-16	0,39	0,46	2,20E-16	1,2E+07	M2

The five different binning strategies of a fixed window size 0.5 bp did not always properly combine the same TRF into a binned TRF. For example, the 5´-TRF range 206.68 to 206.84 bp in the electropherograms (Table 2.3) were not collected into one peak in the binning strategies I, III and IV (Table 2.7). This indicated that all the initially selected starting points 50.25, 50.10, 50.20, 50.30 and 50.65 bp introduced false distributed TRFs in the binning output table. A larger window size of 1 bp did not solve the problem, e.g. the same 5´-TRFs within range 199.5 – 200. 5 bp were not binned into one 5´-TRF. Always a few of the same TRFs were separated into two different TRFs during binning, creating a false-positive OTU.

In an additional experiment, the output table M2 of the binning strategy I (starting point of 50.25 bp) with a fixed window size 0.5 bp was realigned in a time-consuming procedure manually to obtain a profile with the peak variation as low as possible. The binned 5´-TRFs and 3´-TRFs which were distributed into different but close range size were checked back to the particular electropherogram profiles. Correlating peaks were then realigned manually (Table 2.8). This profile M2 was used to obtain a close to optimal binning in the statistical presentation.

Table 2.7. A partial output table for binning strategy I, II, III, IV and V with Table 8 as a partial 5´-TRF dataset. The TRFs within size range 205.58 – 206.84 bp and 199.5 – 200.5 bp are not successfully binned with a fixed window size 0.5 bp and 1 bp, respectively.

A fixed window size 0.5 bp

M2 Binning strategy I

Range length size (bp)		The 9 replicate profiles								
		F-6.6-A1	F-6.6-A2	F-6.6-A3	F-6.6-A4	F-6.6-A5	F-6.6-A6	F-6.6-A7	F-6.6-A8	F-6.6-A9
199,25	199,75	0,95	1,16	1,15	1,15	1,16	1,11	1,01	1,18	1,13
200,25	200,75	1,48	1,91	1,86	2,06	2,02	2,18	2,15	1,88	1,84
203,25	203,75	1,91	2,44	2,45	2,51	2,49	2,47	2,43	2,42	2,38
205,25	205,75	0,28	0,36	0,36	0	0	0	0,33	0,35	0
206,25	206,75	0,55	0	0	0	0	0	0	0,69	0,68
206,75	207,25	0	0,71	0,69	0,74	0,72	0,71	0,69	0	0

M1 Binning strategy II

Range length size (bp)		The 9 replicate profiles								
		F-6.6-A1	F-6.6-A2	F-6.6-A3	F-6.6-A4	F-6.6-A5	F-6.6-A6	F-6.6-A7	F-6.6-A8	F-6.6-A9
199,1	199,6	0,95	1,16	1,15	1,15	1,16	1,11	1,01	1,18	1,13
200,1	200,6	1,48	0	1,86	2,06	2,02	2,18	2,15	1,88	1,84
200,6	201,1	0	1,91	0	0	0	0	0	0	0
203,1	203,6	1,91	2,44	2,45	2,51	2,49	2,47	2,43	2,42	2,38
205,1	205,6	0,28	0	0	0	0	0	0	0	0
205,6	206,1	0	0,36	0,36	0	0	0	0,33	0,35	0
206,6	207,1	0,55	0,71	0,69	0,74	0,72	0,71	0,69	0,69	0,68

M1 Binning strategy III

Range length size (bp)		The 9 replicate profiles								
		F-6.6-A1	F-6.6-A2	F-6.6-A3	F-6.6-A4	F-6.6-A5	F-6.6-A6	F-6.6-A7	F-6.6-A8	F-6.6-A9
199,2	199,7	0,95	1,16	1,15	1,15	1,16	1,11	1,01	1,18	1,13
200,2	200,7	1,48	1,91	1,86	2,06	2,02	2,18	2,15	1,88	1,84
203,2	203,7	1,91	2,44	2,45	2,51	2,49	2,47	2,43	2,42	2,38
205,2	205,7	0,28	0,36	0,36	0	0	0	0,33	0,35	0
206,2	206,7	0	0	0	0	0	0	0	0	0,68
206,7	207,2	0,55	0,71	0,69	0,74	0,72	0,71	0,69	0,69	0

M1 Binning strategy IV

Range length size (bp)		The 9 replicate profiles								
		F-6.6-A1	F-6.6-A2	F-6.6-A3	F-6.6-A4	F-6.6-A5	F-6.6-A6	F-6.6-A7	F-6.6-A8	F-6.6-A9
199,3	199,8	0	1,16	1,15	1,15	1,16	1,11	1,01	0	1,13
200,3	200,8	1,48	1,91	1,86	2,06	2,02	2,18	2,15	1,88	1,84
203,3	203,8	1,91	2,44	2,45	2,51	2,49	2,47	2,43	2,42	2,38
205,3	205,8	0,28	0,36	0,36	0	0	0	0,33	0,35	0
206,3	206,8	0,55	0,71	0,69	0,74	0,72	0	0	0,69	0,68
206,8	207,3	0	0	0	0	0	0,71	0,69	0	0

M2 Binning strategy V

Range length size (bp)		The 9 replicate profiles								
		F-6.6-A1	F-6.6-A2	F-6.6-A3	F-6.6-A4	F-6.6-A5	F-6.6-A6	F-6.6-A7	F-6.6-A8	F-6.6-A9
199,15	199,65	0,95	1,16	1,15	1,15	1,16	1,11	1,01	1,18	1,13
200,15	200,65	1,48	1,91	1,86	2,06	2,02	2,18	2,15	1,88	1,84
203,15	203,65	1,91	2,44	2,45	2,51	2,49	2,47	2,43	2,42	2,38
205,15	205,65	0,28	0	0	0	0	0	0	0,35	0
205,65	206,15	0	0,36	0,36	0	0	0	0,33	0	0
206,65	207,15	0,55	0,71	0,69	0,74	0,72	0,71	0,69	0,69	0,68

A fixed window size 1 bp

M2 Binning strategy I

Range length size (bp)		The 9 replicate profiles								
		F-6.6-A1	F-6.6-A2	F-6.6-A3	F-6.6-A4	F-6.6-A5	F-6.6-A6	F-6.6-A7	F-6.6-A8	F-6.6-A9
199,5	200,5	0	0	0	0	2,02	2,18	2,15	0	0
200,5	201,5	1,48	1,91	1,86	2,06	0	0	0	1,88	1,84
202,5	203,5	1,91	0	2,45	2,51	2,49	2,47	0	2,42	2,38
203,5	204,5	0	2,44	0	0	0	0	2,43	0	0
205,5	206,5	0,28	0,36	0,36	0	0	0	0,33	0,35	0
206,5	207,5	0,55	0,71	0,69	0,74	0,72	0,71	0,69	0,69	0,68

Table 2.8. A partial output table M2 binning strategy I that had been manually realigned. Bold numbers of normalized areas of 5´-TRFs indicated that those peaks were removed to the range length size that had the majority 5´-TRFs.

M2 Binning strategy I after manual correction

Range length size (bp)		The 9 replicate profiles								
		F-6.6-A1	F-6.6-A2	F-6.6-A3	F-6.6-A4	F-6.6-A5	F-6.6-A6	F-6.6-A7	F-6.6-A8	F-6.6-A9
199,25	199,75	0,95	1,16	1,15	1,15	1,16	1,11	1,01	1,18	1,13
200,25	200,75	1,48	1,91	1,86	2,06	2,02	2,18	2,15	1,88	1,84
203,25	203,75	1,91	2,44	2,45	2,51	2,49	2,47	2,43	2,42	2,38
205,25	205,75	0,28	0,36	0,36	0	0	0	0,33	0,35	0
206,75	207,25	**0,55**	0,71	0,69	0,74	0,72	0,71	0,69	**0,69**	**0,68**

2.4.4. NMDS ordination of quantitative datasets presenting richness and evenness information

Each replicate profile of one subsample is represented by a point and the points are arranged in two dimensional NMDS ordination. The similar profiles are represented by points that are close together, while the dissimilar profiles are further apart. The manual realigned profile showed an ideal binning and a perfect ordination with stress values under 0.1, as expected for an ideal ordination without a potential misinterpretation. All nine replicate profiles of each subsample are placed closely and formed a cluster. Depth dependent cluster were clearly visible: one cluster represented one bacterial community at one particular depth (Fig. 2.11 and 2.12). It indicated that the biological variation between the bacterial communities at the different layers was larger than the technical variations between the replicate profiles, the uncertainty of the method procedures. The analyses of fragments in size range 50 – 500 bp produced 6 outliers out of 36 5´-TRF profiles (Fig. 2.11) and 3 outliers out of 36 3´-TRF profiles. One cause may be the small peaks below 75 bp of the internal marker Genescan500™-500-ROX™ were uncertainly determined or incorrectly recognized by the GeneMapper® Software v3.7. These outliers had different 5´-TRFs and 3´-TRFs in term of length size and number of peak and area.
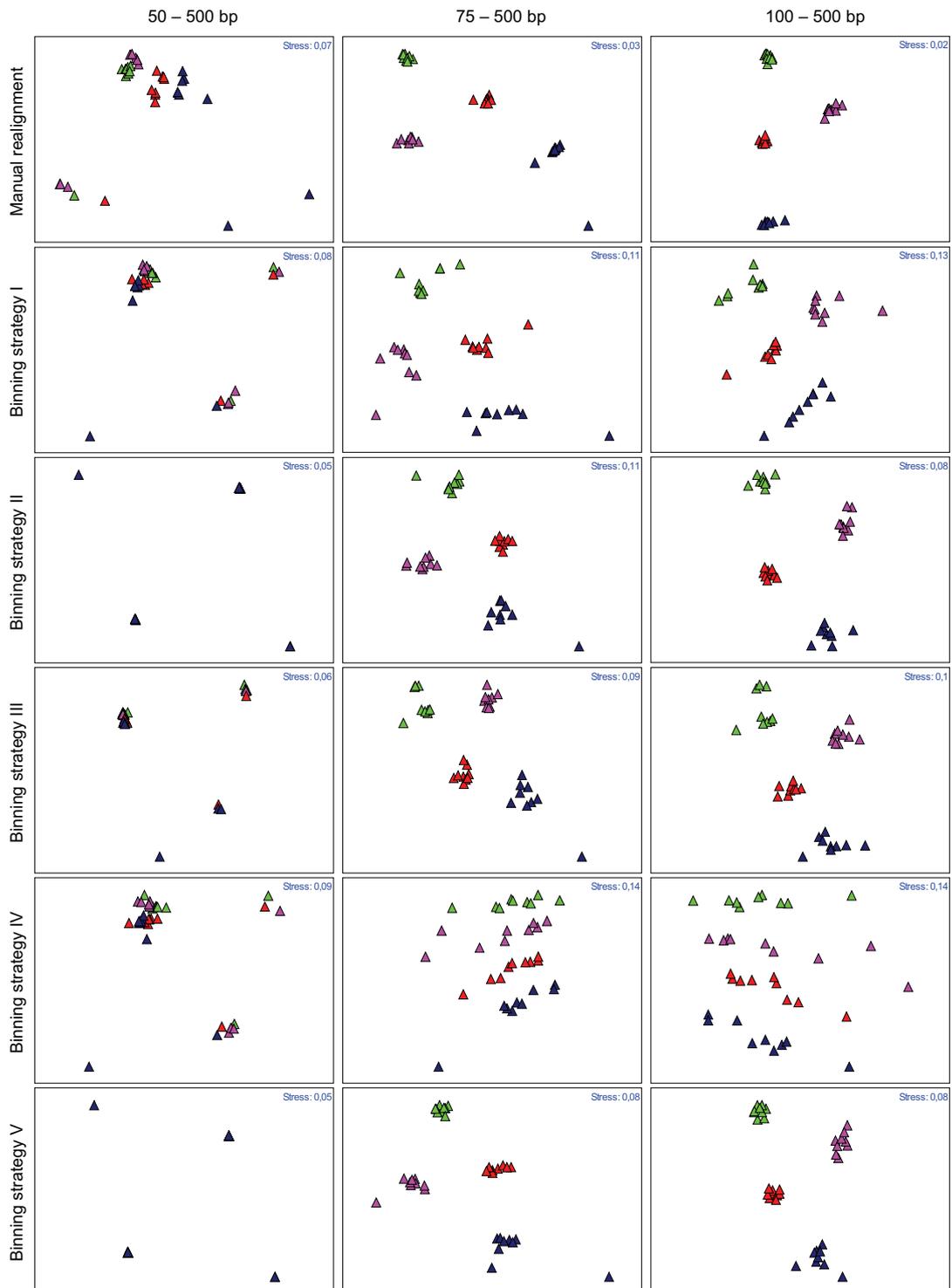
Fig. 2.11. The nine replicate profiles of quantitative 5´-TRF datasets with a fixed window size 0.5 bp. NMDS ordination within 3 different fragment size ranges. The starting point are 50.25, 50.10, 50.20, 50.30 and 50.65 bp respectively for binning strategy I, II, III, IV and V. The green, pink, red and blue triangle are for subsample 6.3 (1-1.5 cm depth), 6.4 (1.5-2 cm depth), 6.5 (2-2.5 cm depth) and 6.6 (2.5-3 cm depth), respectively.
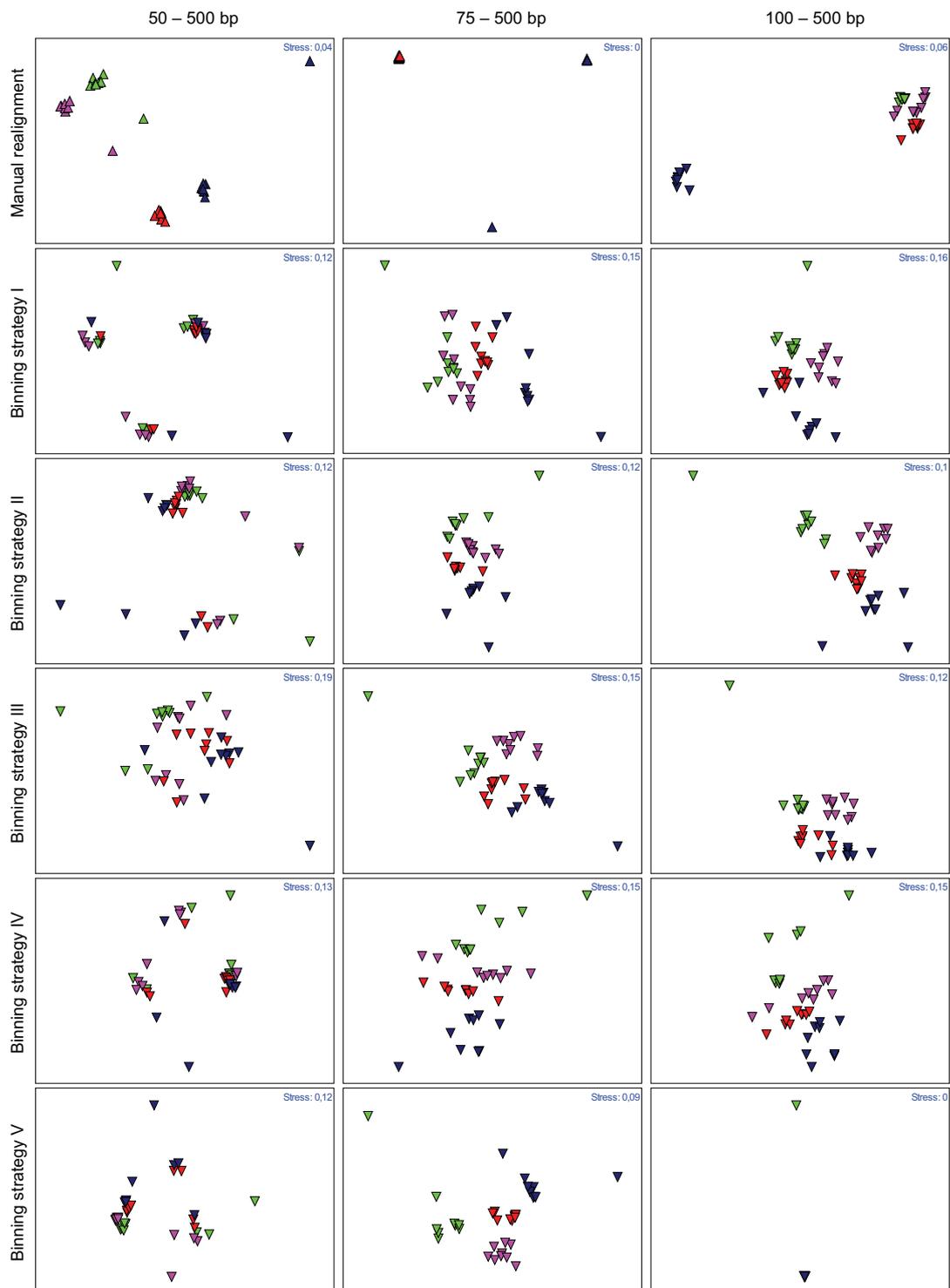
Fig. 2.12. The nine replicate profiles of quantitative 3′-TRF datasets with a fixed window size 0.5 bp. NMDS ordination within 3 different fragment size ranges. The starting points are 50.25, 50.10, 50.20, 50.30 and 50.65 bp respectively for binning strategy I, II, III, IV and V.The green, pink, red and blue triangle are for subsample 6.3 (1-1.5 cm depth), 6.4 (1.5-2 cm depth), 6.5 (2-2.5 cm depth) and 6.6 (2.5-3 cm depth), respectively.

Even binning still remained not an ideal process, all starting points of a fixed window size 0.5 bp generated NMDS ordinations representing bacterial community clusters over depth as same as the manual realignment (Fig. 2.11) with stress values 0.07 to 0.14 for analyses size range 75-500 bp and 100-500 bp. The ANOSIM results indicated that the bacterial clusters over depth were significantly different ($p<0.01$) in the bacterial composition within size range analysis of 75 – 500 bp (Appendix 7). The 3´-TRF datasets represented a similar NMDS pattern but with a relatively higher stress value 0.1 - 0.2. After excluding one outlier in the size range analysis 50 – 500 bp (the profile from R-6.3-A1 which was totally different with its 8 other replicate profiles), the depth dependent bacterial clusters were clearly seen as presented in Fig. 2.12.

Binning strategy IV produced a clear pattern of 5´-TRF dataset ordinations with stress value 0.14. But the nine replicate profiles representing points had distances like they were different. It might be the consequence of introducing many false distributed fragments within a particular replicate profiles by splitting the same fragment into separate adjacent bins that occurred in the binning process with starting point 50.30 (binning strategy IV).

Concerns that binning might introduce false distributed fragments in the nine replicate profiles, causes a study on the effects of binning on the reproducibility of 5´-TRFs and 3´-TRFs within the replicate profiles and their similarity based on the Bray-Curtis coefficient (Fig. 2.13). Manual realignment definitively yielded the highest number of reproducible 5´-TRFs and 3´-TRFs: about 80% of the average observed number. The similarity between the nine replicate profiles was up to 90% for the three different size ranges. The improperness binning of the 5 different binning strategies with a fixed window size 0.5 bp was indicated by the lower reproducibility number: approximately 50% - 60% of observed number were reproducible TFRs. The similarity was about 80% for fragment size range analysis 75 – 500 bp and 100 – 500 bp while the fragment range size 50 – 500 bp had the lowest similarity, meaning that the irreproducible fragments mainly had size less than 100 bp. Further more, binning strategy IV produced the lowest reproducible fragment number, about 36% for 5´-TRFs and 43% for 3´-TRFs with the similarity below 80%. This explains the scattered NMDS ordination of the datasets after binning strategy IV.
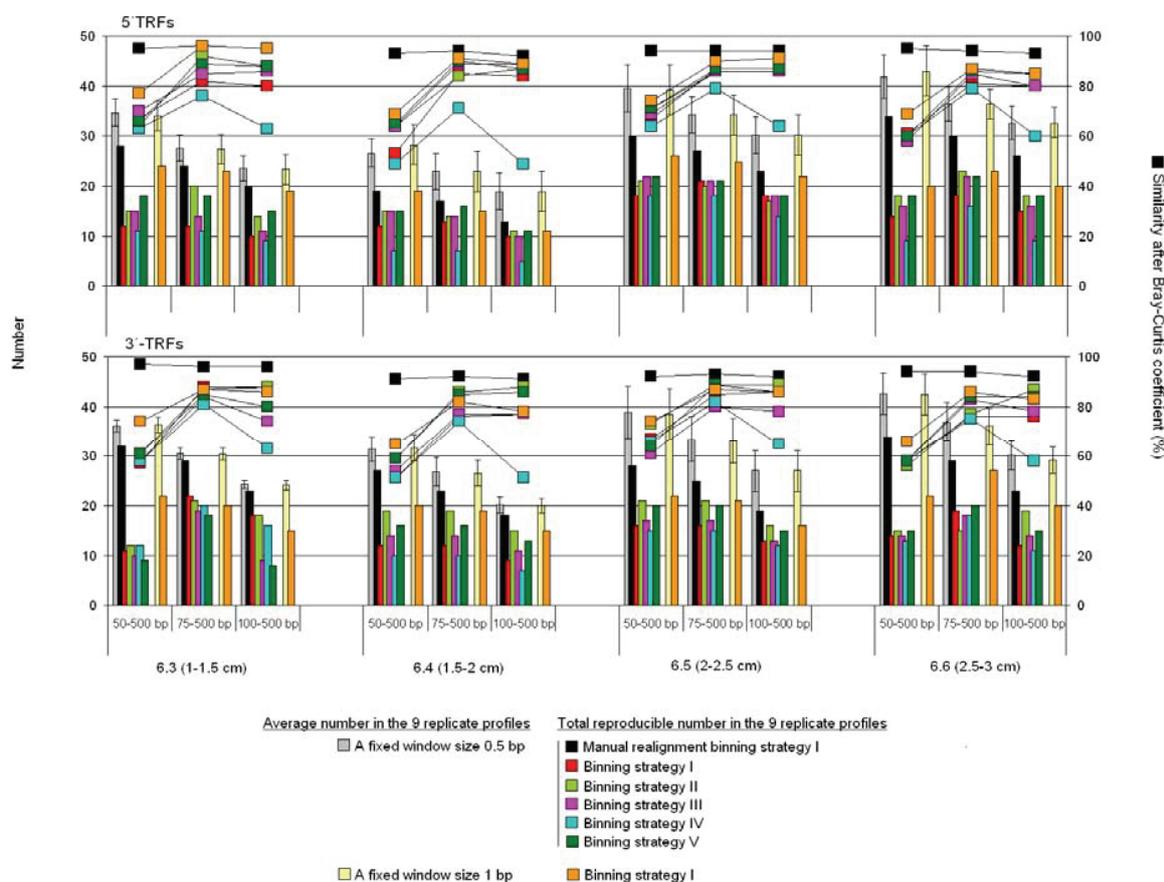
Fig. 2.13. The nine replicate profiles. Bar diagram shows the comparison of average observed number, total reproducible number and similarity percentage of 5´-TRF and 3´-TRF dataset with a fixed window size was 0.5 bp and starting point of 50.25, 50.10, 50.20, 50.30 and 50.65 bp respectively for binning strategy I, II, III, IV and V; and with a fixed window size 1 bp with starting point 50.50 bp. The outlier profile was excluded from the nine replicate profiles based on the NMDS ordination.

The average observed fragment number after binning with a fixed window size 1 bp was comparable with that after a binning with a fixed window size 0.5 bp. It meant that a fixed window size 1 bp did not drastically reduce the binned TRFs number even experimentally the closest distance between TRFs was about 0.80 bp. Approximately 60% - 80% of observed 5´-TRFs and 60% - 75% of 3´-TRFs were reproducible fragments within the nine replicate profiles with similarity percentage higher than 80% for fragment size range analysis of 75 – 500 bp and 100 – 500 bp. It seemed that irreproducible fragments less than 100 bp still could not be combined by

this window size. A fixed window size 1 bp with a starting point 50.50 bp also presented a similar NMDS pattern as presented by binning strategy I, II, III and V with a fixed window size 0.5 bp (Fig. 2.14). The dept dependent bacterial clusters were significantly different (p<0.01) indicated that they were different bacterial communities over depth (Appendix 7).
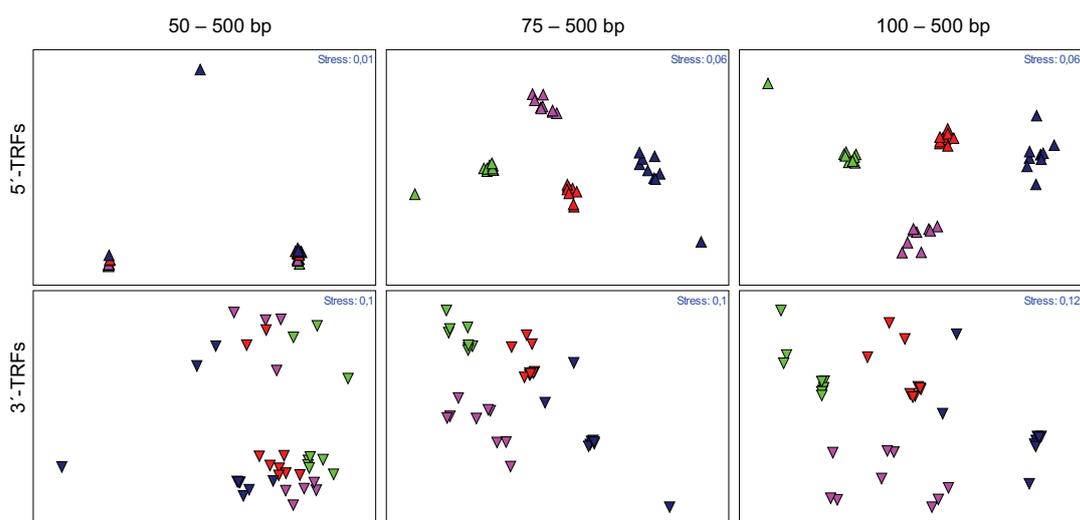
Fig. 2.14. The nine replicate profiles with a fixed window size 1 bp. NMDS ordination of quantitative 5´-TRF and 3´-TRF datasets within 3 different fragment range length size. The starting point is 50.50 bp. The green, pink, red and blue triangle are for 6.3 (1-1.5 cm depth), 6.4 (1.5-2 cm depth), 6.5 (2-2.5 cm depth) and 6.6 (2.5-3 cm depth), respectively.

The triplicate profiles strategy also presented reproducible 5´-TRF and 3´-TRF datasets. As previous results indicated that binning strategy I, II, III and IV produced similar ordinations, sample 13 was binned with binning strategy I for a fixed window size 0.5 bp and 1 bp. The chosen output tables for both window sizes are M2. Both fixed window sizes yielded a comparable number (Fig. 2.15). The similarity percentages between triplicate profiles were relatively stable within 3 different size range analyses. Fig. 2.15 also showed that the Simper analysis detected about 1% - 20% and 1% – 10% of reproducible 5´-TRFs were also present in the non-digested triplicate profiles respectively for a fixed window size 0.5 bp and 1 bp as well as 1% - 40% and 2 – 30% for 3´-TRF datasets.
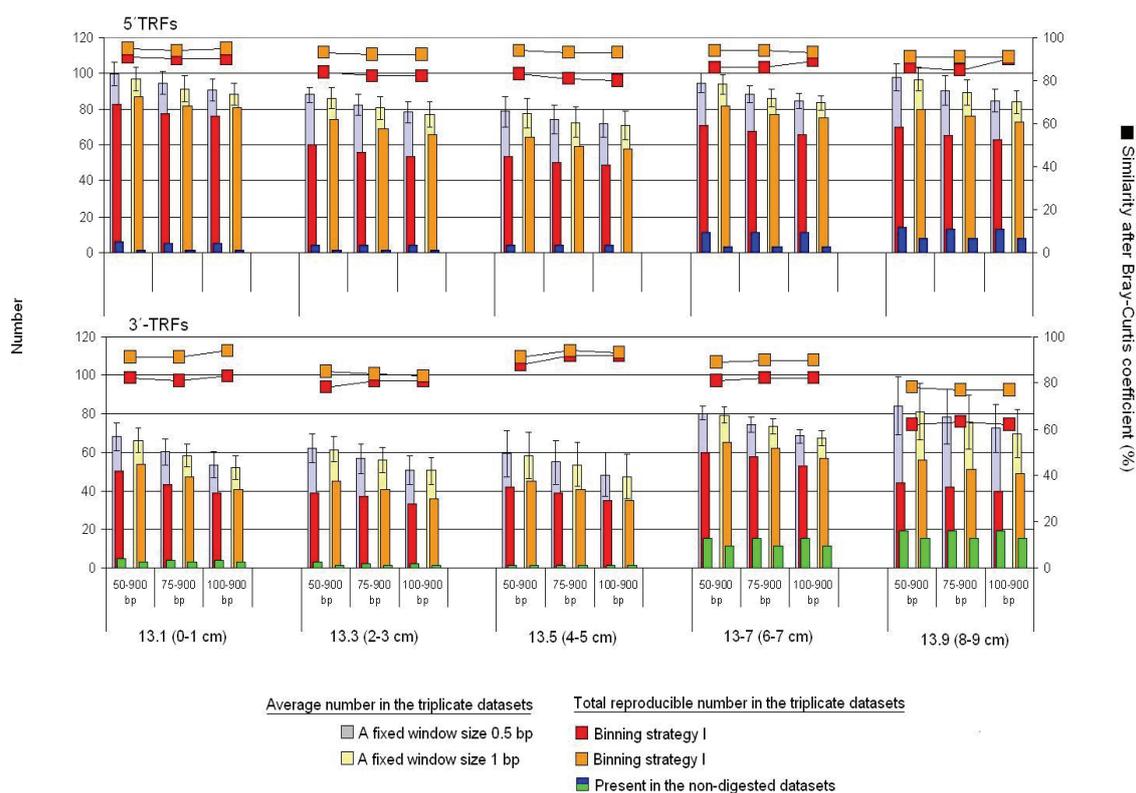
Fig. 2.15. The triplicate profiles. Bar diagram shows the comparison of average observed number, average reproducible number and similarity percentage of 5´-TRF and 3´-TRF datasets with starting point 50.25 and 50.50 bp respectively for a fixed window size 0.5 bp and 1 bp.

The NMDS ordination showed that the triplicate profiles of sample 13 were clustered together and different layer presented different bacterial cluster with stress values were less than 0.1 (Fig. 2.16). The ANOSIM results trended that the bacterial clusters were well separated to each other (R=1) related to bacterial composition differences. But because of the low number of replicates (one clusters had triplicate profiles), the significance value was above 0.05 (Appendix 7).

In contrast to the sample 6 (the nine replicate profiles), all size range analyses represented clear depth dependent bacterial clusters due to the reproducibility of TRFs in range size 50 – 900 bp. It might refer to the correctness of the GeneMapper® Software v3.7 detect the MapMarker1000® within range 50 – 900 bp. The size range analysis 50 – 900 bp with a fixed window size 1 bp more clearly differentiated the sediment surface layer from the deeper sediment layers. Due to

that a fixed window size 1 bp generated a higher similarity percentage within triplicate profiles; the triplicate profiles represented points were really overlapping to each other than those with a fixed window size 0.5 bp.
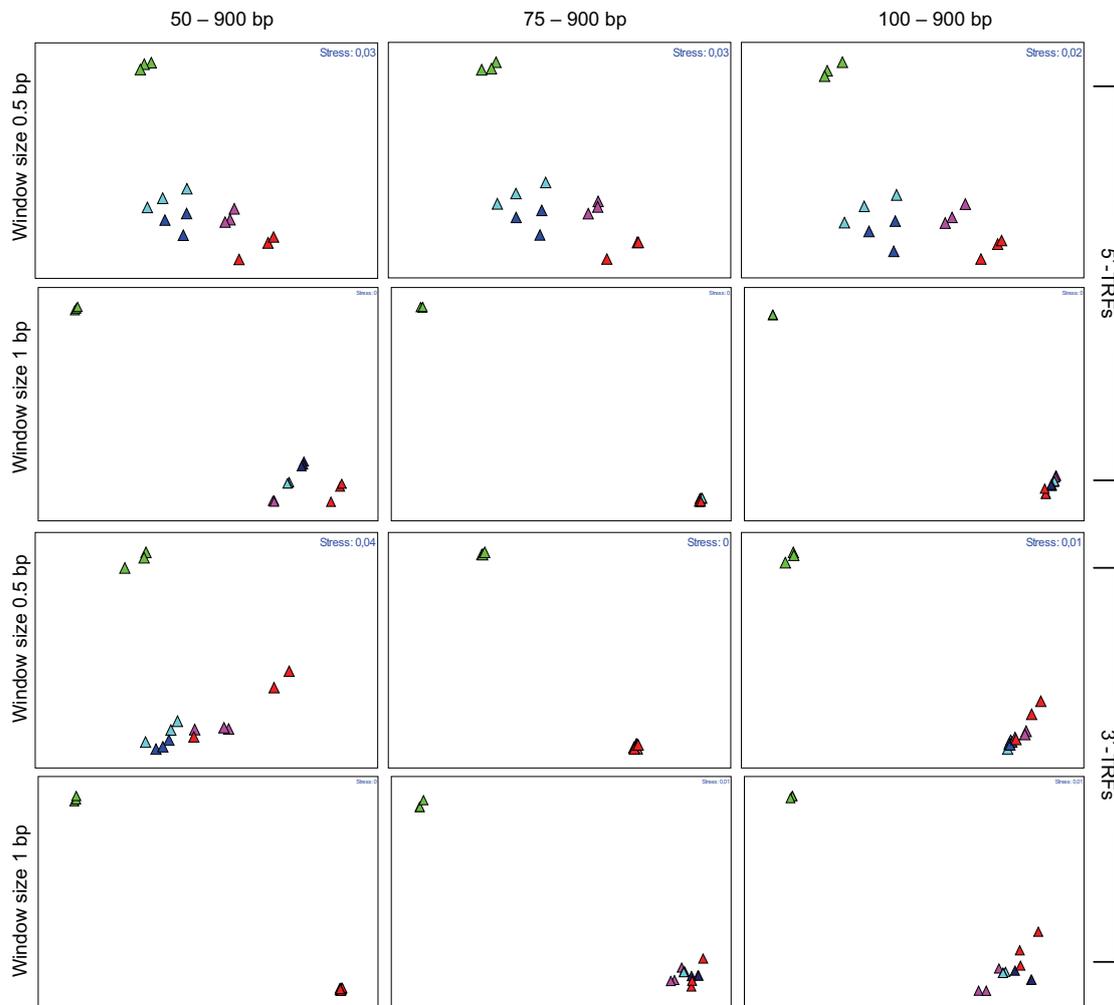
Fig. 2.16. The triplicate profiles of quantitative 5´-TRF and 3´-TRF datasets. NMDS ordination within 3 different fragment size ranges after binning with a fixed window size 0.5 bp and 1 bp and starting point respectively 50.25 and 50.50 bp. The green, light blue, dark blue, pink and red triangle are for sample 13.1 (0-1cm depth), 13.3 (2-3 cm depth), 13.5 (4-5 cm depth), 13.7 (6-7 cm depth) and 13.9 (8-9 cm depth), respectively.

2.4.5. Analysis of qualitative (binary) datasets presenting richness information

As the quantitative data presenting richness (the observed number of TRFs) and evenness (the relative area) are sensitive to the technical variations (Blackwood *et al.*, 2007; Engebretson and Moyer, 2003; Saikaly *et al.*, 2005; Dunbar *et al.*, 2001; Liu *et al.*, 1997), NMDS ordination was also performed with binary datasets (absence/presence). The Jaccard coefficient was used to account the similarity. It describes the similarity of each sample pair based on TRFs present in at least one sample. Thus, the absence of TRFs in the samples will not contribute to the similarity.

In general, the ordinations yielded similar patterns in presenting the reproducibility of 5´-TRF and 3´-TRF datasets and the depth dependent bacterial community as the analysis of quantitative profiles (Fig. 2.17). The ANOSIM also yielded a similar indication that the depth dependent clusters were well separated due to the deference of bacterial composition. In case of the nine replicate profiles, the outliers in sample 6 (F-6.6-A1, R-6.3-A1 and R-6.6-A1) which present in size range analysis 50 – 500 bp   were absent in size range 75 – 500 bp and 100 – 500 bp. It meant that TRFs above 75 bp were reproducible and increased the resolution of ordination. Manual realignment produced an ideal ordination with stress value 0.08 - 0.11 for 5´-TRF and 3´-TRF datasets. A fixed window size 1 bp represented useful two dimensional ordinations with stress value 0.1 – 0.2, while a fixed window size 0.5 bp generated random ordinations with stress value higher than 0.2 for both datasets and all starting points. In case of the triplicate profiles (Fig. 2.18), both fixed window sizes with starting point 50.00 bp generated an ideal binary ordination for 5´-TRF and 3´-TRF datasets. A fixed window size 1 bp generated a clearer reproducible datasets pattern as the three triplicate profiles represented points were close to each other. Because the observed reproducible TRFs number after a fixed window size 1 bp was higher than after a fixed window size 0.5 bp
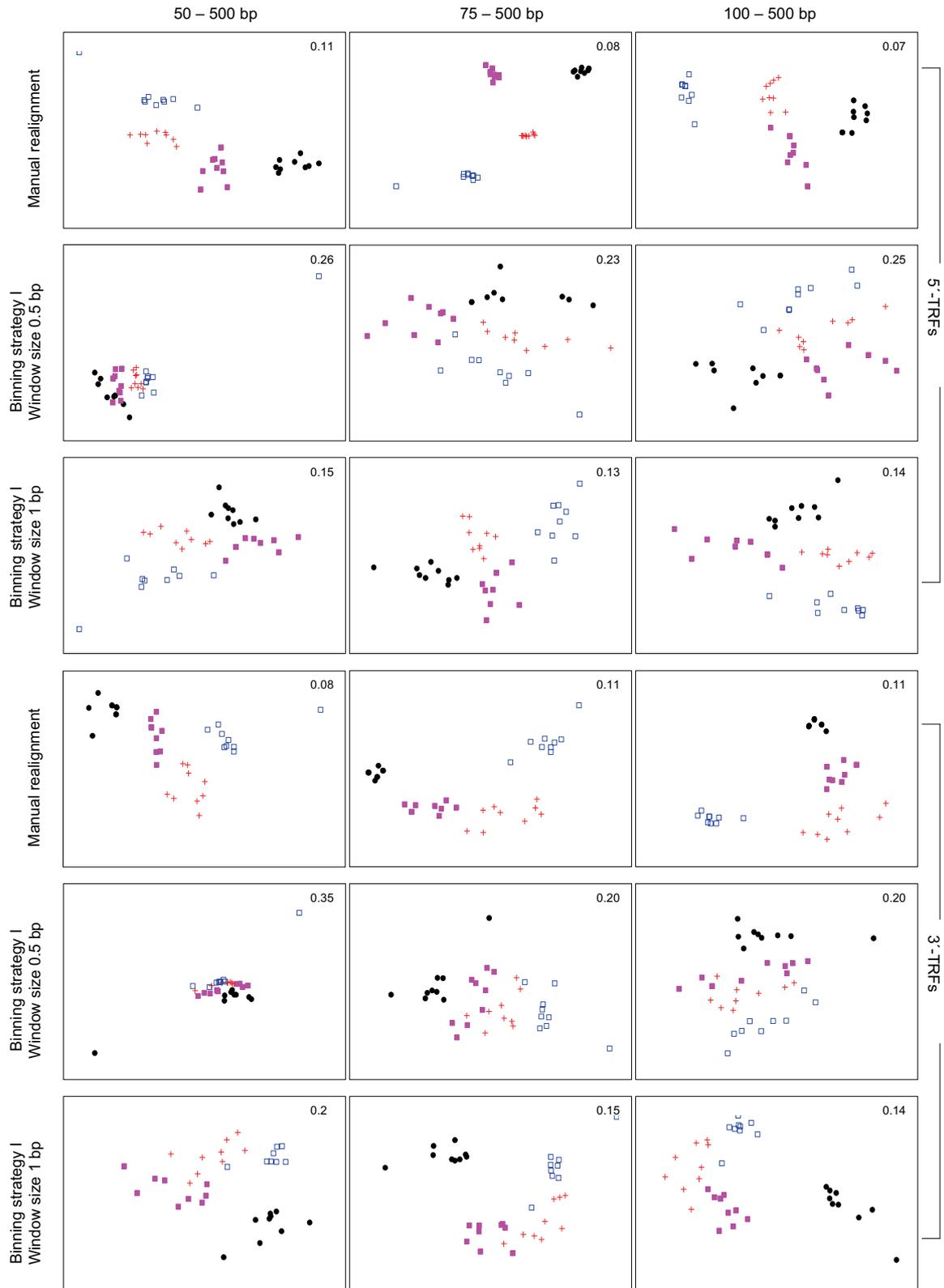
Fig. 2.17. The nine replicate profiles of binary 5´-TRF and 3´-TRF datasets with a fixed window size 0.5 bp and 1 bp respectively with starting point of 50.25 and 50.50 bp. NMDS ordination within 3 different fragment size ranges. The color code black, pink, red and blue are for sample 6.3 (1-1.5 cm depth), 6.4 (1.5-2 cm depth), 6.5 (2-2.5 cm depth) and 6.6 (2.5-3 cm depth).
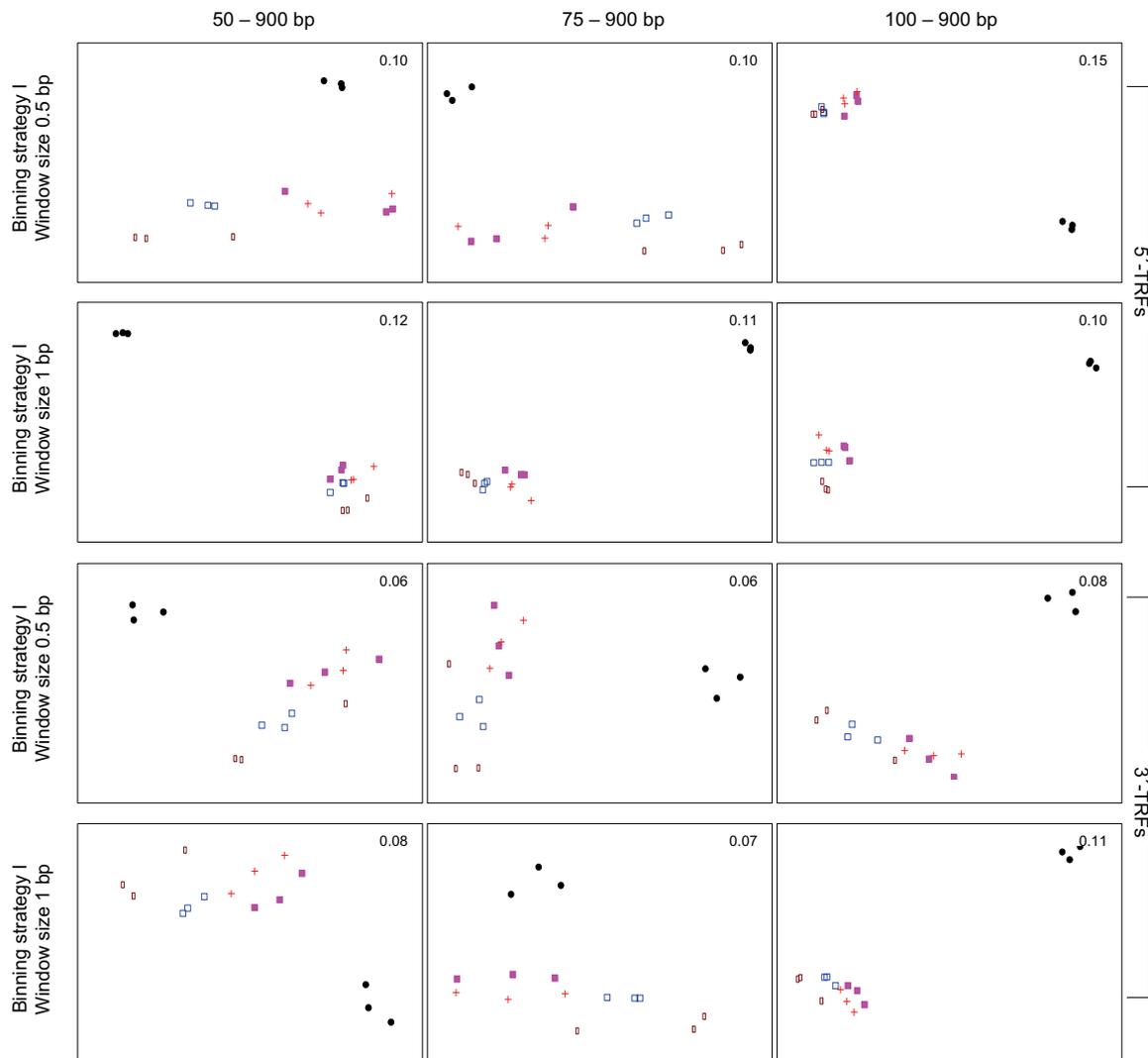
Fig. 2.18. The triplicate profiles of binary 5´-TRF and 3´-TRF datasets with a fixed window size 0.5 bp and 1 bp respectively with starting point 50.25 and 50.50 bp. NMDS ordination within 3 different fragment size ranges. The color code black, pink, red, blue and brown are for sample 13.1 (0-1cm depth), 13.3 (2-3 cm depth), 13.5 (4-5 cm depth), 13.7 (6-7 cm depth) and 13.9 (8-9 cm depth) respectively.

2.5. Discussion

2.5.1. Peak detection

Systematic and random error may occur in experiments. Systematic error refers to results either lower or larger than the real value by a fixed percentage or amount. Random error or noise is an unpredictable variation in the measurements. As a TRF is presented by a peak, distinguishing true peak and noise is a fundamental step. Our results indicated that the capillary electrophoresis on the used ABI Prism 3130XL genetic analyzer reproducibly does fragment separation according to the peak number of the internal standard markers with minor additional peaks. The source of these additional peaks is not known; they may be provided by the manufacturer or may have been generated during the handling. The presence of additional peaks may or may not influence the function of the internal marker do determination and sizing for unknown fragment. If this additional peak is not identified by the software as a peak with a defined length size (bp), this additional peak may not contribute in generating a size calling curve. Therefore manual editing affects the validity of size calling curve, as we introduced formerly an undefined peak to a defined peak.

Concerning that variation may come from different quantitative amount of taken sample influenced by automatically loading and different run (Liu *et al*., 1997; Dunbar *et al*., 2001; Osborn *et al*., 2000), we did not find a statically effect of the different loading and run on the peak number and peak area of the internal markers in parallel with the parameter adjustments of the GeneMapper® software v3.7 (ABI). The available default parameters were tested constantly performs peak analysis and peak sizing according to the peak number and peak area and also to the length size as described by the manufacturers.

The sensitivity of the capillary electrophoresis and the software to determine the small fragments bp is influenced by several factors, e.g. in our study it is the kind of internal marker. The small fragment of Genescan500™-500-ROX™ could not be identified correctly (112 out of 120 profiles had problem for fragment below 75 bp). Therefore samples analyzed with Genescan500™-500-ROX™ generated a scattered NMDS ordination with a lower similarity percentage within replicate profiles if small fragment less than 75 bp were included in the analysis. But in contrast with the

MapMarker1000®, including the small fragments in the analysis did not influence the reproducibility pattern of the replicate profiles of a sample. The less accuracy for fragment separation and identification at the beginning and at the close end of the run respectively for small and larger fragment was identified by Marsh (1999) and Liu *et al*. (1997), as they suggested confining analysis range to obtain a reproducibility replicate profiles. However defining the analysis range limits the obtained TRFs which may important for sample differentiation (Blackwood *et al.*, 2007).

As we assumed that peaks which were present in the non-digested sample were false positive peaks, we applied a peak amplitude threshold 100 rfu and light smoothing for determining the true peaks. This threshold and smoothing probably excluded the true peaks from our datasets, but the excluded peaks were mostly small peaks in area as the total area was not significantly different between high (100 rfu) and low (50 rfu) peak amplitude threshold. The presence or absence of small peak in low peak height (100 – 155 rfu) was assumed by Osborn *et al*. (2000) as variation occurred in either loading or the detection system. Reducing small peaks can be also critical if they are present frequently within replicates as they may be important for comparison between samples even if they are from PCR bias (Blackwood *et al.*, 2003). But the analyses with the peak amplitude threshold 100 rfu and light smoothing could represent the overall depth dependent bacterial community in this study. A constant peak threshold of 100 rfu also had been reported as a satisfactory threshold for standardization and presenting a similarity of the triplicate profiles by Osborne *et al.* (2006). A higher peak amplitude threshold than 100 rfu was not considered for sample 13, because it did not significantly reduce false positive peaks while peaks with big area were related to undigested amplicons. Blackwood *et al.* (2003) mentioned that peak amplitude threshold 50 and 100 rfu produced the same effect on the dendrograms error rate, while peak amplitude threshold 200 rfu produced the highest error rate. As a comparison, heavy smoothing was chosen by Clement *et al*. (1998) to differentiate bacterial communities between fecal pellet and sand and by Kaplan *et al*. (2001) for tracking bacterial change in the rat fecal.

2.5.2. Binning and dataset reproducibility

In principal binning process is rounding up and down the fragment length size. It is done either manually or automatically with available program or software, e.g. R-software which is used in this study. By applying automatic binning, the starting point for rounding up and down can be defined priory based on the window size and as a result two output tables can be generated consist of binned and rounded TRFs that differ in the starting point for rounding process; usually the difference is one half of the window size. As a parametric and non-parametric test are also available in the program, one of two generated output tables is   statistically option for the most optimal binned output combining the same TRFs. Therefore automatic binning provides advantage than to round the TRF length size up and down manually which may remain an erroneous (Abdo *et al*., 2006; Ruan *et al*., 2006; Hewson and Fuhrman, 2006).

All initially defined starting points of a high resolution fixed window size 0.5 bp did not allow to an ideal binning without introducing false distributed fragments. We applied this resolution due to the peak separation has resolution fragment shift of 0.5 bp and Hewson and Fuhrman (2006) suggested that a window size should be as wide as impression, while binning collects fragments which slightly differ in size into one area defined by a discrete size or called window size (Abdo *et al*., 2006; Hewson and Fuhrman, 2006). But manual realignment obviously combined all the same TRFs and yielded a high number of reproducible TRFs and a high similarity percentage between replicate profiles (about 90%). It meant that applying a high resolution fixed window size 0.5 bp could solve the fragment shift if followed by manual realignment. Even Clement *et al*. (1998) did manual alignment, but a manual realignment may not applicable if dealing with a huge dataset; it remains time consuming and potentially accompanied with human errors.

A comparable number of averages reproducible observed TRFs to those with manual realignment was obtained by applying a fixed window size of 1 bp, but this window size still exhibited the same binning problem. It showed that even experimentally the closest distance between TRFs in our datasets was about 0.80 bp, the fixed window size 1 bp did not dramatically reduce the binned TRFs number. Dunbar *et al*. (2001) reported the same resolution of 0.5 bp for the fragment shift and they binned the replicate profiles with window size 1 bp by using a clustering

algorithm. The bin size of 1 bp was also applied by many other T-RFLP studies, e.g. by Liu *et al* (1997), Clement *et al*. (1998), Osborn *et al*. (2000); Osborne *et al*. (2006).

No significant variation occurring from independent loading and detection system was detected for either internal markers or samples in this study. Fig. 2.9 and 2.10 showed that independent runs generated a similar peak number and peak area within the replicate profiles of sample 6 and sample 13 (the nine replicate strategy and the triplicate strategy) as indicated by low standard deviation. The variations occur between different PCRs from a single sample, between different digestions from a single PCR, between runs from a single digestion were examined with respect to the reproducibility of particular TRFs not only the peak number but the defined TRFs with length sizes. Both replicate strategies indicated the same pattern of reproducibility in the NMDS ordination with 90% - 80% similarity among the replicate profile by which 10% - 20% dissimilarity may from bias occurred in PCR, enzyme digestion, run and improperness binning. This may indicate that with and without pooling PCR product, the reproducibility pattern can be obtained in parallel with a single digestion mixture and run. The PCR reproducibility for T-RFLP study had been reported by Liu *et al*. (1997) while they introduced this method at the first time. But concerning to the potential bias, Liu *et al*. (1997); Clement *et al*. (1998); Dunbar *et al*. (2001) suggested to pool single PCRs which was then followed by other studies (e.g. Saikaly *et al*., 2005; Blackwood *et al*., 2003). Osborn *et al*. (2000) assumed that when individual PCRs results the similar profiles with those from pooled replicate PCRs then pooling PCRs is not essential. The variation between digestions also suggested not important, since they found that replicate digestions of the same PCR product had no significant variation in peak height (Osborn *et al*., 2000). But anyhow by applying molecular steps in the T-RFLP, the technical variation should be taken in consideration (Blackwood *et al*., 2007; Frey *et al*., 2006; Engebretson and Moyer, 2003; Dunbar *et al*., 2001; Osborn *et al*., 2000; Liu *et al*., 1997; von Wintzingerode *et al*., 1997; Farrelly *et al*., 1995; Reysenbach *et al*., 1992; Liesack *et al*., 1991).

The similarity for the NMDS ordination is counted based on the presence of reproducible TRF and abundance (relative area). The quantitative datasets generated in this study tended to reliable for representing the replicate reproducibility indicated by the stress value that goes to the interpretation range of useful 2 dimensional ordinations. A non optimal ordination is also possible due to the NMDS uses an iterative algorithm of a monotonic unknown transformation for refining

positions of sample points until they satisfy as closely as possible the similarity or dissimilarity relationship between samples; this happens for a poorly structured dataset (van Wezel and Kosters, 2004; Kenkel and Orloci, 1986; Young, 1985). Further more, in term of the biological variation, a high similarity percentage within the nine replicate profiles may indicate a similar bacterial community present among those profiles. The bacterial community at 1.5 - 2 cm, 2 – 2.5 cm, 2.5 – 3 cm and 3 – 3.5 cm depth was also probably different as sub sample represented clusters were clearly distinguishable to each other in the NMDS ordination, at which one cluster consisted of nine replicate profiles. This indication is also similar for the triplicate profiles: the bacterial community in the surface layer (1 cm depth) was different with those in the subsurface layers (3, 5, 7 and 9 cm depth).

The qualitative (binary) ordination was a comparative ordination for detecting technical variations effect on the presentation of biological variation. The ordination pattern was similar with that represented by the quantitative datasets for both replicate strategies. As the qualitative ordination was based on absent/present TRFs, a fixed window size 1 bp generated ordination with lower stress value than a fixed window size 0.5 bp, because a fixed window size 1 bp combined a higher number of reproducible TRFs.

The 3´-TRF dataset generated less resolution in the quantitative NMDS ordination than 5´-TRF dataset. It might be a consequence of the more conserved middle part of the 16S rRNA gene as amplicons were amplified with primer pair 27F – 907R. The 5´- terminus provided a greater discrimination as it has higher number of variable regions than 3´- terminus (Saikaly *et al.*, 2005; Dunbar *et al.*, 2001).

The qualitative NMDS ordination was more useful for 3´-TRFs dataset while abundance (relative area) does not influence the similarity within replicate profiles. The average observed 3`-TRF number was lower than 5´-TRFs, but the average 3´-TRF abundance was higher than 5´-TRFs. For example sample 13 after *Alu*I digestion yielded ratio 1:10000 (peak number : relative area) for the 3´-TRF datasets and 1:2000 for the 5´-TRF datasets. A higher abundance but not followed by proportional number of observed TRF might influence the percentage similarity within replicate profiles in the quantitative ordination of 3´-TRF dataset. Therefore the quantitative ordination of 3´-TRFs remained scattered than that of 5´-TRFs when representing the reproducibility of replicate profiles and the depth dependent bacterial community. The application of forward and reverse labeled primers was

intended to avoid an accidental low resolution, as related organisms may generate an identical size of 5´-TRF due to the similarity of conserved regions and enzyme recognition site, but different size of 3´-TRF (Abdo *et al.*, 2006). And in this study both 5´-TRF and 3´-TRF datasets represented the similar indication for the replicate profile reproducibility and the depth dependent bacterial community.

Conclusion

- The default parameters provided by the Genemapper® software v3.7 were provably applicable in this study for defining a peak. A peak amplitude threshold 100 rfu and light smoothing was technically the correct option for minimizing noise.

- Even though binning is not a perfect process in combining the same TRFs, but binning with a fixed window size 0.5 bp and 1 bp could minimize the replicate variation due to the fragment shift. Both window sizes gained the same similarity percentage in the replicate profiles for about 80% and 20% dissimilarity may arise from binning and molecular bias.

- T-RFLP is a reproducible finger printing method. The method can differentiate the biological variation between samples.

- A quantitative and qualitative profile represented the similar pattern in representing the replicate profiles reproducibility and the depth dependent bacterial community.

2.6. References

1. Abdo, Z., U.M.E. Schütte, S.J. Bent, C.J. Williams, L.J. Forney and P. Joyce. 2006. Statistical methods for characterizing diversity of microbial communities by analyzing of terminal restriction fragment length polymorphisms of 16S rRNA genes. Environ. Microbiol. 8:929-938.

2. Amann, R.I., W. Ludwig and K-H. Scheifer, 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. 59:143-169.

3. Blackwood, C.B., D. Hudleston, D.R. Zak and J.S. Buyer. 2007. Interpreting ecological diversity indices applied to terminal restriction fragment length polymorphism data: insights from simulated microbial communities. Appl. Environ. Microbiol. 73:5276-5283.

4. Blackwood, C.B., T. Marsh, S-H. Kim and E.A. Paul. 2003. Terminal restriction fragment length polymorphism data analysis for quantitative comparison of microbial communities. Appl. Environ. Microbiol. 69:926-932.

5. Clarke, K.R. and R.N. Gorley. 2001. PRIMER v5: User manual/tutorial, PRIMER-E, Plymouth UK.

6. Clarke, K.R. and R.M. Warwick. 2001. Change in marine community: an approach to statistical analysis and interpretation, 2nd edition. PRIMER-E. Plymouth. UK.

7. Clement, B.G., L.E. Kehl, K.L. DeBord and C.L. Kitts. 1998. Terminal restriction fragment pattern (TRFPs), a rapid PCR-based method for the comparison of complex bacterial communities. J. Microbiol. Methods. 31:135-142.

8. Dunbar, J., L.O. Ticknor and C.R. Kuske. 2001. Phylogenetic specificity and reproducibility and new method for analysing of terminal restriction fragment profile of 16S rRNA genes from bacterial communities. Appl. Environ. Microbiol. 67:190-197.

9. Dunbar, J., L.O. Ticknor and C.R. Kuske. 2000. Assessment of microbial diversity on two south western U.S. soils by terminal restriction fragment analysis. Appl. Environ. Microbiol. 66:2943-2950.

10. Engebretson, J.J and C. L. Moyer. 2003. Fidelity of select restriction endonucleases in determining microbial diversity by terminal restriction fragment length polymorphisms. Appl. Environ. Microbiol. 69:4823-4829.

11. Frey, J.C., E.R. Angert and A.N. Pell. 2006. Assessment of biases associated with profiling simple, model communities using terminal-restriction fragment length polymorphism-based analysis. J. Microbiol. Met. 67:9-19.

12. Hewson, I. and J. A. Fuhrman. 2006. Improved strategy for comparing microbial assemblage fingerprints. Microbiol. Ecol. 51:147-153.

13. Kaplan, C. W., J. C. Astaire, M. E. Sanders, B. S. Reddy and C. L. Kitts. 2001. 16S Ribosomal DNA terminal restriction fragment pattern analysis of bacterial communities in feces of rats fed Lactobacillus acidophilus NCFM. Appl. Environ. Microbiol. 67:1935-1939.

14. Kenkel, N. C. and L. Orloci. 1986. Applying metric and nonmetric multidimensional scaling to ecological studies: some new results. Ecology. 67:919-928.

15. Liu, W., T.L. Marsh, Cheng H. and L.J. Forney. 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl. Environ. Microbiol. 63:4516-4522.

16. Marsh, T.L. 1999. Terminal restriction fragment length polymorphism (T-RFLP): an emerging method for characterizing diversity among homologous populations of amplification products. Current Opinion in Microbiol. 2:323-327.

17. Moeseneder, M.M., J.M. Arrieta, G. Muyzer, C. Winter and G.J. Herndl. 1999. Optimization of terminal restriction fragment length polymorphism analysis for complex marine bacterioplankton communities and comparison with denaturing gradient gel electrophoresis. Appl. Environ. Microbiol. 65:3518-3525.

18. Muyzer, G., A. Teske and C.O. Wirsen. 1995. Phylogenic relation of Thiomicrospira sp. And their identification in deep sea hydrothermal vent samples by DGGE of 16S rRNA fragments, Arch. Microbiol. 59:695-700.

19. Muyzer, G. 1999. Genetic fingerprinting of microbial communities-present status and future perspectives. Proceeding of the 8th International Symposium on Microbial Ecology. Bell CR, Brylinsky M and Johnson-Green P. (Ed). Atlantic Canada Society for Microbial Ecology. Halifax. Canada.

20. Osborn, A.M., E.R.B. Moore and K.N. Timmis. 2000. An evelution of terminal restriction fragment length polymorphisms (T-RFLP) analysis for the study of microbial community structure and dynamics. Environ. Microbiol. 2:39-50.

21. Osborne, C.A., G.N. Rees, Y. Bernstein and P.H. Janssen. 2006. New threshold and confidence estimates for terminal restriction fragment length polymorphism analysis of complex bacterial communities. Appl. Environ. Microbiol. 72:1270-1278.

22. Rees, G.N., D.S. Baldwin, G.O. Watson, S. Perryman and D.L. Nielsen. 2004. Ordination and significance testing of microbial community composition derived from terminal restriction fragment length polymorphisms: application of multivariate statistics. Antonie van Leeuwenhoek. 86:339-347.

23. Saikaly, P.E., P.G. Stroot and D.B. Oether. 2005. Use of 16S rRNA gene terminal restriction fragment analysis to assess the impact of solids retention time on the bacterial diversity of activated sludge. Appl. Environ. Microbiol. 71:5814-5822.

24. Tiedje,J.M., S.Asuming-Brempong, K.Nüsslein, T.L. Marsh amd S.J.Flynn. 1999. Opening the black box of soil microbial diversity. Appl. Soil Ecol. 13:109-122.

25. Van Wezel, M.C. and W.A. Kosters. 2004. Nonmetric multidimensional scaling: Neutral networks versus traditional techniques. Intelligent data analysis. IOS Press. 8:601-613.

26. Volkenborn, N., S.I.C. Hedtkamp, J.E.E. van Beusekom and K. Reise. 2007. Effects of bioturbation by lugworm (*Arenicola marina*) on physical and chemical sediment properties and implications for intertidal habitat succession. Estuatine, Coastal and Shelf Sci. 74:331-343.

27. Volkenborn, N., L. Polerecky, S.I.C. Hedtkamp, J.E.E. van Beusekom and D. de Beer. 2007a. Bioturbation and bioirrigation extend the open exchange regions in permeable sediments. Limnol. Oceanogr. 52:1898-1909.

28. Von Wintzingerode, F., U.B. Göebel and E. Stackebrandt. 1997. Determination of microbial diversity in environmental samples: pitfalls of PCR-based rRNA analysis. FEMS. Microbial. Rev. 21:213-229.

Appendix 1:

Table 2.9. Effect polynomial degree 2, 3 and 4 on the total number of peak and area for known fragments Genescan500™-500-ROX™.

Sample	Repli-cate #	Polynomial degree 2		Polynomial degree 3 and 4	
		Σ peak	Σ area	Σ peak	Σ area
6.3	1	16	27465	17	27321
	2	16	32340	16	32030
	3	16	29968	16	29885
	4	19	22193	17	21940
	5	16	22290	16	22273
	6	16	23014	17	22813
	7	16	26157	17	25973
	8	16	22879	17	22765
	9	16	26081	17	25935
	10	16	32101	16	31878
	11	16	21364	16	21185
	12	16	23359	16	23208
6.4	1	16	22918	16	22760
	2	16	25550	17	25503
	3	16	28768	16	28600
	4	16	35604	16	35240
	5	16	33682	16	33317
	6	16	26637	16	26360
	7	16	25695	16	25456
	8	16	25873	16	25655
	9	16	33041	17	32814
	10	15	22053	15	21882
	11	15	25549	15	25369
	12	15	31220	15	30899
6.5	1	16	25659	16	25450
	2	16	21993	16	21841
	3	16	22321	16	22166
	4	16	24493	16	24484
	5	16	28248	16	28057
	6	16	32969	17	32665
	7	20	28545	17	28178
	8	16	22720	17	22473
	9	17	5128	18	5937
	10	16	23834	16	23639
	11	16	28045	16	27888
	12	16	25163	16	24992
6.6	1	16	22615	16	22328
	2	16	25128	16	25013
	3	16	28134	16	27806
	4	16	20477	17	20305
	5	16	21008	16	20804
	6	16	22194	16	22140
	7	16	29549	16	29362
	8	16	32805	16	32420
	9	16	29336	16	29024
	10	16	23748	16	23550
	11	16	23616	16	23399
	12	15	25953	15	25763

Table 2.10. Two-way ANOVA with replicate analysis for Table 2.11.

Total peak $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample	6.20833	3	2.06944	4.42375	0.00605	2.70819
Polynomial degree	0.375	1	0.375	0.80162	0.37305	3.94932
Interaction	0.20833	3	0.06944	0.14845	0.93043	2.70819
Within	41.1667	88	0.4678			
Total	47.9583	95				

Average area $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample	1061510	3	353837	3.82027	0.01266	2.70819
Polynomial degree	13326.5	1	13326.5	0.14388	0.70537	3.94932
Interaction	3783.3	3	1261.1	0.01362	0.99781	2.70819
Within	8150626	88	92620.7			
Total	9229245	95				

Appendix 2:

Table 2.11. Effect peak window size of 9, 13 and 15 data points on the total number of peak and area for known fragments MapMarker1000®

Sample	Repli-cate #	Peak window size 9		Peak window size 13		Peak window size 15	
		Σ peak	Σ area	Σ peak	Σ area	Σ peak	Σ area
13.1-a	1	43	363178	44	336496	41	361175
	2	43	334774	38	269624	42	334469
	3	40	270281	31	200456	38	269816
13.3-a	1	32	200667	33	221975	31	201326
	2	34	223259	39	315412	33	222015
	3	40	315702	33	215351	40	315857
13.5-a	1	34	215739	30	155292	33	216414
	2	31	156277	31	214618	30	156181
	3	31	214510	37	301268	31	214700
13.7-a	1	38	301372	34	250386	36	299544
	2	34	249179	38	303006	34	250627
	3	38	301387	46	347356	36	300841
13.9-a	1	48	349271	42	310871	46	349122
	2	42	309425	38	276653	41	309374
	3	38	276241	40	304182	36	276784
13.1-c	1	40	304021	38	277831	39	303891
	2	39	279453	40	265943	38	279211
	3	40	264698	31	177826	40	266124
13.3-c	1	32	179120	38	226534	31	177929
	2	39	226750	40	291377	37	225233
	3	41	291579	31	186903	40	292805
13.5-c	1	33	188472	33	194023	31	187001
	2	33	193904	32	204782	32	193811
	3	33	205237	38	268451	32	204921
13.7-c	1	38	266797	29	164557	37	266753
	2	30	164127	40	253676	29	165620
	3	40	252189	40	290135	39	252094
13.9-c	1	40	288366	38	289788	40	290206
	2	39	290131	41	292602	38	289853
	3	41	291022	42	362697	39	290208

Table 2.12. Two-way ANOVA with replicate analysis for Table 2.11.

MapMarker1000®

Total peak — $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample	745.8778	9	82.87531	5.89161	7.40E-06	2.040098
Peak window size	19.35556	2	9.677778	0.687994	0.506504	3.150411
Interaction	86.42222	18	4.801235	0.34132	0.993065	1.778446
Within	844	60	14.06667			
Total	1695.656	89				

Average area — $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample	29550826	9	3283425	6.846249	1.06E-06	2.040098
Peak window size	667419.1	2	333709.5	0.695816	0.502646	3.150411
Interaction	3564064	18	198003.6	0.412856	0.979974	1.778446
Within	28775684	60	479594.7			
Total	62557993	89				

Appendix 3:

Table 2.13. Effect smoothing with peak amplitude threshold 50 rfu on the total number of peak and total area of known fragments.

Genescan500™-500-ROX™

Sample	Repli-cate #	None smoothing Σ peak	None smoothing Σ area	Light smoothing Σ peak	Light smoothing Σ area
6.3	1	17	27321	16	27340
	2	16	32030	16	32189
	3	16	29885	16	29859
	4	17	21940	16	22026
	5	16	22273	16	22184
	6	17	22813	16	22867
	7	17	25973	16	26039
	8	17	22765	16	22767
	9	17	25935	16	25936
6.4	1	16	22760	16	22768
	2	17	25503	16	25413
	3	16	28600	16	28648
	4	16	35240	16	35402
	5	16	33317	16	33464
	6	16	26360	16	26495
	7	16	25456	16	25509
	8	16	25655	16	25670
	9	17	32814	16	32793
6.5	1	16	25450	16	25472
	2	16	21841	16	21791
	3	16	22166	16	22132
	4	16	24484	16	24390
	5	16	28057	16	28100
	6	17	32665	16	32691
	7	17	28178	17	28221
	8	17	22473	16	22537
	9	17	5118	17	5066
6.6	1	16	22328	20	22432
	2	16	25013	16	24997
	3	16	27806	16	27889
	4	17	20305	16	20226
	5	16	20804	16	20812
	6	16	22140	16	22087
	7	16	29362	16	29431
	8	16	32420	16	32556
	9	16	29024	16	29094

MapMarker1000®

Sample	Repli-cate #	None smoothing Σ peak	None smoothing Σ area	Light smoothing Σ peak	Light smoothing Σ area
13.1-a	1	41	361175	39	361008
	2	42	334469	41	334550
	3	38	269816	36	269406
13.3-a	1	31	201326	31	200734
	2	33	222015	31	221647
	3	40	315857	37	315677
13.5-a	1	33	216414	31	214953
	2	30	156181	29	155067
	3	31	214700	31	215022
13.7-a	1	36	299544	35	299884
	2	34	250627	32	250279
	3	36	300841	35	301069
13.9-a	1	46	349122	40	347345
	2	41	309374	39	308972
	3	36	276784	35	276019
13.1-c	1	39	303891	37	304047
	2	38	279211	36	277823
	3	40	266124	37	263955
13.3-c	1	31	177929	29	177397
	2	37	225233	35	224674
	3	40	292805	37	290924
13.5-c	1	31	187001	29	186503
	2	32	193811	31	193471
	3	32	204921	31	205160
13.7-c	1	37	266753	36	267060
	2	29	165620	29	164802
	3	39	252094	35	251056
13.9-c	1	40	290206	38	288520
	2	38	289853	35	288881
	3	39	290208	37	290219

Table 2.14. Two-way ANOVA with replicate analysis for Table 2.13.

Genescan500™-500-ROX™

Total peak α = 0.05

Source of Variation	SS	df	MS	F	P-value	F crit
Sample	0.597222	3	0.199074	0.567657	0.63833	2.748191
None/light smoothing	0.680556	1	0.680556	1.940594	0.168424	3.990924
Interaction	2.263889	3	0.75463	2.151815	0.102383	2.748191
Within	22.44444	64	0.350694			
Total	25.98611	71				

Total area α = 0.05

Source of Variation	SS	df	MS	F	P-value	F crit
Sample	2.32E+08	3	77490987	2.798864	0.047051	2.748191
None/light smoothing	14421.68	1	14421.68	0.000521	0.981862	3.990924
Interaction	7108.375	3	2369.458	8.56E-05	0.999999	2.748191
Within	1.77E+09	64	27686588			
Total	2.00E+09	71				

MapMarker1000

Total peak α = 0.05

Source of Variation	SS	df	MS	F	P-value	F crit
Sample	543.0667	9	60.34074	7.016365	5.27E-06	2.124029
None/light smoothing	52.26667	1	52.26667	6.077519	0.01808	4.084746
Interaction	5.066667	9	0.562963	0.065461	0.999916	2.124029
Within	344	40	8.6			
Total	944.4	59				

Total area α = 0.05

Source of Variation	SS	df	MS	F	P-value	F crit
Sample	1.13E+11	9	1.25E+10	7.806802	1.58E-06	2.124029
None/light smoothing	5269399	1	5269399	0.003286	0.954574	4.084746
Interaction	2314193	9	257132.6	0.00016	1	2.124029
Within	6.41E+10	40	1.60E+09			
Total	1.77E+11	59				

Appendix 4:

Table 2.15. Effect of peak amplitude threshold and smoothing on the 5´-TRFs number within size range 50-500 bp and total area in the nine replicate profiles of sample 6 with Genescan500™-500-ROX™.

Sample		Repli-cate #	Threshold 50 rfu				Threshold 100 rfu			
			None smoothing		Light smoothing		None smoothing		Light smoothing	
			ΣPeak	ΣArea	ΣPeak	ΣArea	ΣPeak	ΣArea	ΣPeak	ΣArea
AluI	6.3	1	70	258070	65	258022	38	239742	36	240531
	(1-1.5 cm)	2	66	234762	58	232965	37	217439	35	217904
		3	58	178216	52	176516	34	165642	30	163099
		4	65	226370	59	224990	38	210590	35	209136
		5	60	211933	58	210792	36	197500	34	196367
		6	59	179051	49	175145	33	164854	32	165322
		7	70	247649	62	246087	42	231873	39	230993
		8	63	206073	61	206093	38	192199	36	190802
		9	59	190071	55	189177	33	174974	31	174768
	6.4	1	55	179353	51	178225	30	166136	27	163609
	(1.5-2 cm)	2	44	156737	40	155212	24	143671	22	142184
		3	55	192380	51	191088	31	179629	28	177244
		4	68	218154	60	215957	34	201148	31	200104
		5	57	179270	50	177124	30	165368	25	161560
		6	53	175714	49	174766	30	163228	28	162113
		7	69	259303	67	258650	40	243551	37	242119
		8	64	209850	52	207058	33	193912	31	194180
		9	55	178053	49	176559	28	163018	25	161466
	6.5	1	74	242920	66	241635	49	230561	45	228940
	(2-2.5 cm)	2	63	200429	56	196855	43	188729	37	184531
		3	60	182644	55	181418	38	170411	34	168063
		4	62	224202	60	223761	47	215679	43	213541
		5	65	209785	56	206099	41	196056	37	193753
		6	69	228707	60	225624	49	218637	43	215399
		7	76	254464	66	249900	52	242570	44	237605
		8	63	195659	54	192924	41	182837	37	181550
		9	54	164809	52	163901	35	153767	33	151868
	6.6	1	77	247967	76	313594	49	232711	53	435177
	(2.5-3 cm)	2	69	212927	64	211148	47	200557	44	199677
		3	78	232013	67	228750	47	213975	45	213942
		4	61	185603	54	182878	44	177379	39	173992
		5	60	189534	54	187317	42	180368	40	178732
		6	51	154248	48	152717	37	146552	35	143957
		7	80	278411	77	277840	49	260555	48	260631
		8	77	237096	68	234097	50	222619	45	220056
		9	69	195654	63	195204	44	182973	42	183038
Control	6.3	1	27	155021	24	154774	12	146852	12	147694
	(1-1.5 cm)	2	21	122184	21	123339	11	117552	11	118535
		3	20	118281	18	117993	12	114376	11	114216
	6.4	1	7	41905	7	41828	4	39745	3	30408
	(1.5-2 cm)	2	7	33105	6	32235	3	30390	3	33937
		3	6	35968	6	35899	3	33941	4	39753
	6.5	1	11	67849	9	67076	6	64973	6	65047
	(2-2.5 cm)	2	11	79097	11	79090	7	76452	7	76542
		3	10	64809	9	64322	6	62312	6	62342
	6.6	1	14	60932	14	60914	9	58042	6	55840
	(2.5-3 cm)	2	14	56809	13	56085	6	51959	6	51991
		3	12	46502	11	46029	5	42637	5	42691

Table 2.16. Effect of peak amplitude threshold and smoothing on the 5´-TRFs number within size range 50-900 bp and total area in the triplicate profiles of sample 13 with MapMarker1000®.

Sample		Repli-cate #	Threshold 50 rfu				Threshold 100 rfu			
			None smoothing		Light smoothing		None smoothing		Light smoothing	
			Σ Peak	Σ Area	Σ Peak	Σ Area	Σ Peak	Σ Area	Σ Peak	Σ Area
AluI	13.1	1	200	1402603	180	1406675	122	1359397	117	1369302
	(0-1 cm)	2	229	1556917	206	1546122	132	1502128	127	1498731
		3	171	1135114	155	1124608	113	1104587	105	1096061
	13.3	1	141	821781	131	818876	97	793685	91	792187
	(2-3 cm)	2	136	752905	123	748078	90	723846	85	722382
		3	164	1027649	144	1018925	103	992690	95	987427
	13.5	1	144	932720	136	947916	94	900944	87	914370
	(4-5 cm)	2	106	506456	100	503468	73	485285	69	483200
		3	129	727188	121	722855	88	698374	82	696289
	13.7	1	174	1214046	161	1220257	111	1174188	105	1183079
	(6-7 cm)	2	145	877972	133	862049	97	849256	92	835044
		3	167	1133534	159	1141615	105	1092363	100	1100492
	13.9	1	182	1302203	169	1295203	128	1269392	120	1264501
	(8-9 cm)	2	210	1657256	196	1632630	134	1611945	128	1589189
		3	191	1536414	175	1555958	134	1504103	122	1521862
Control	13.1	1	161	367076	130	353128	45	310415	36	303699
	(0-1 cm)	2	159	372944	135	362756	45	318315	38	313066
		3	265	514829	218	492425	94	428485	77	412573
	13.3	1	95	229417	79	222791	27	197203	26	195991
	(2-3 cm)	2	137	302305	118	294361	40	255372	28	246815
		3	145	295079	117	282347	39	246076	31	238601
	13.5	1	49	124832	38	120384	12	106759	12	106529
	(4-5 cm)	2	114	218152	91	207973	28	176982	25	174624
		3	68	150583	53	144750	19	129074	19	128737
	13.7	1	166	255579	143	245936	47	194599	42	190924
	(6-7 cm)	2	59	102571	42	96443	17	84237	17	83908
		3	153	237103	130	227544	43	179989	35	174250
	13.9	1	156	223762	127	211233	47	164942	44	162533
	(8-9 cm)	2	255	356464	213	340618	90	269771	75	260716
		3	196	281671	164	267224	66	212130	52	200362

Appendix 5:

Table 2.17. Two-way ANOVA with replicate for Table 2.15

Total peak $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample : 6.3; 6.4; 6.5; 6.6	3346.799	3	1115.6	30.12872	8.36E-15	2.675387
Threshold and smoothing	20000.41	3	6666.803	180.0487	9.60E-46	2.675387
Interaction	194.1736	9	21.57485	0.582666	0.809386	1.953763
Within	4739.556	128	37.02778			
Total	28280.94	143				

Total area $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample : 6.3; 6.4; 6.5; 6.6	1.58E+10	3	5.27E+09	3.704833	0.013468	2.675387
Threshold and smoothing	5.07E+09	3	1.69E+09	1.187356	0.31727	2.675387
Interaction	2.53E+09	9	2.81E+08	0.197607	0.994077	1.953763
Within	1.82E+11	128	1.42E+09			
Total	2.06E+11	143				

Total peak $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Control : 6.3; 6.4; 6.5; 6.6	881.2292	3	293.7431	130.5525	5.00E-18	2.90112
Threshold and smoothing	443.7292	3	147.9097	65.73765	8.96E-14	2.90112
Interaction	99.02083	9	11.00231	4.889918	0.000373	2.188766
Within	72	32	2.25			
Total	1495.979	47				

Total area $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Control : 6.3; 6.4; 6.5; 6.6	5.97E+10	3	1.99E+10	159.7884	2.48E-19	2.90112
Threshold and smoothing	1.47E+08	3	48904868	0.392631	0.759122	2.90112
Interaction	22280091	9	2475566	0.019875	0.999999	2.188766
Within	3.99E+09	32	1.25E+08			
Total	6.39E+10	47				

Total peak with threshold 50 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample : 6.3; 6.4; 6.5; 6.6	1173.944	3	391.3148	7.126062	0.000333	2.748191
None and light smoothing	636.0556	1	636.0556	11.58293	0.001153	3.990924
Interaction	4.166667	3	1.388889	0.025292	0.994508	2.748191
Within	3514.444	64	54.91319			
Total	5328.611	71				

Total area with threshold 50 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample : 6.3; 6.4; 6.5; 6.6	6.12E+09	3	2.04E+09	1.777151	0.160416	2.748191
None and light smoothing	55945.13	1	55945.13	4.87E-05	0.994454	3.990924
Interaction	1.85E+08	3	61581973	0.053604	0.983498	2.748191
Within	7.35E+10	64	1.15E+09			
Total	7.98E+10	71				

Total peak with threshold 100 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample : 6.3; 6.4; 6.5; 6.6	2343.819	3	781.2731	40.81383	7.29E-15	2.748191
None and light smoothing	159.0139	1	159.0139	8.306911	0.005371	3.990924
Interaction	19.04167	3	6.347222	0.33158	0.802525	2.748191
Within	1225.111	64	19.14236			
Total	3746.986	71				

Total area with threshold 100 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample : 6.3; 6.4; 6.5; 6.6	1.03E+10	3	3.42E+09	2.01336	0.120899	2.748191
None and light smoothing	2.98E+08	1	2.98E+08	0.175754	0.676451	3.990924
Interaction	1.79E+09	3	5.95E+08	0.350791	0.788695	2.748191
Within	1.09E+11	64	1.70E+09			
Total	1.21E+11	71				

Total peak with threshold 50 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Control : 6.3; 6.4; 6.5; 6.6	769.4583	3	256.4861	69.95076	2.05E-09	3.238872
None and light smoothing	5.041667	1	5.041667	1.375	0.258118	4.493998
Interaction	1.458333	3	0.486111	0.132576	0.939248	3.238872
Within	58.66667	16	3.666667			
Total	834.625	23				

Total area with threshold 50 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Control : 6.3; 6.4; 6.5; 6.6	3.08E+10	3	1.03E+10	76.1921	1.09E-09	3.238872
None and light smoothing	345120.2	1	345120.2	0.002564	0.960245	4.493998
Interaction	404574.8	3	134858.3	0.001002	0.999954	3.238872
Within	2.15E+09	16	1.35E+08			
Total	3.29E+10	23				

Total peak with threshold 100 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Control : 6.3; 6.4; 6.5; 6.6	208.3333	3	69.44444	83.33333	5.57E-10	3.238872
None and light smoothing	0.666667	1	0.666667	0.8	0.384351	4.493998
Interaction	1	3	0.333333	0.4	0.754885	3.238872
Within	13.33333	16	0.833333			
Total	223.3333	23				

Total area with threshold 100 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Control : 6.3; 6.4; 6.5; 6.6	2.90E+10	3	9.65E+09	84.30731	5.10E-10	3.238872
None and light smoothing	2301.042	1	2301.042	2.01E-05	0.996478	4.493998
Interaction	1212332	3	404110.8	0.00353	0.999698	3.238872
Within	1.83E+09	16	1.14E+08			
Total	3.08E+10	23				

Table 2.18. Two-way ANOVA with replicate for Table 2.16

Total peak $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample : 13.1;13.3;13.5;13.7;13.9	27388.07	4	6847.017	34.21228	2.02E-12	2.605975
Threshold and smoothing	46030.93	3	15343.64	76.66711	1.24E-16	2.838745
Interaction	1799.4	12	149.95	0.74925	0.696033	2.003459
Within	8005.333	40	200.1333			
Total	83223.73	59				

Total area $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample : 13.1;13.3;13.5;13.7;13.9	5.01E+12	4	1.25E+12	36.33194	8.03E-13	2.605975
Threshold and smoothing	1.72E+10	3	5.73E+09	0.166171	0.918549	2.838745
Interaction	4.77E+08	12	39784353	0.001154	1	2.003459
Within	1.38E+12	40	3.45E+10			
Total	6.40E+12	59				

Total peak $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Control : 13.1;13.3;13.5;13.7;13.9	53183.77	4	13295.94	11.7547	2.09E-06	2.605975
threshold and smoothing	132120.1	3	44040.04	38.93502	6.09E-12	2.838745
Interaction	12004.37	12	1000.364	0.884404	0.569045	2.003459
Within	45244.67	40	1131.117			
Total	242552.9	59				

Total area $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Control : 13.1;13.3;13.5;13.7;13.9	3.87E+11	4	9.69E+10	28.87707	2.55E-11	2.605975
threshold and smoothing	3.54E+10	3	1.18E+10	3.517869	0.023547	2.838745
Interaction	4.01E+09	12	3.34E+08	0.099691	0.999934	2.003459
Within	1.34E+11	40	3.35E+09			
Total	5.61E+11	59				

Total peak with threshold 50 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample : 13.1;13.3;13.5;13.7;13.9	20748.53	4	5187.133	15.26376	7.14E-06	2.866081
None and light smoothing	1333.333	1	1333.333	3.923492	0.061536	4.351243
Interaction	125.3333	4	31.33333	0.092202	0.983825	2.866081
Within	6796.667	20	339.8333			
Total	29003.87	29				

Total area with threshold 50 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample : 13.1;13.3;13.5;13.7;13.9	2.54E+12	4	6.35E+11	17.81215	2.25E-06	2.866081
None and light smoothing	52068918	1	52068918	0.001459	0.969905	4.351243
Interaction	77645897	4	19411474	0.000544	0.999999	2.866081
Within	7.14E+11	20	3.57E+10			
Total	3.26E+12	29				

Total peak with threshold 100 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample : 13.1;13.3;13.5;13.7;13.9	8303.133	4	2075.783	34.34832	1.07E-08	2.866081
None and light smoothing	307.2	1	307.2	5.083287	0.035526	4.351243
Interaction	10.46667	4	2.616667	0.043298	0.99615	2.866081
Within	1208.667	20	60.43333			
Total	9829.467	29				

Total area with threshold 100 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Sample : 13.1;13.3;13.5;13.7;13.9	2.47E+12	4	6.17E+11	18.5483	1.65E-06	2.866081
None and light smoothing	2169216	1	2169216	6.52E-05	0.993636	4.351243
Interaction	41673099	4	10418275	0.000313	1	2.866081
Within	6.65E+11	20	3.33E+10			
Total	3.13E+12	29				

Total peak with threshold 50 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Control : 13.1;13.3;13.5;13.7;13.9	56921.47	4	14230.37	7.126824	0.000976	2.866081
None and light smoothing	4813.333	1	4813.333	2.410604	0.136197	4.351243
Interaction	412	4	103	0.051584	0.994606	2.866081
Within	39934.67	20	1996.733			
Total	102081.5	29				

Total area with threshold 50 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Control : 13.1;13.3;13.5;13.7;13.9	2.23E+11	4	5.59E+10	13.06369	2.17E-05	2.866081
None and light smoothing	8.80E+08	1	8.80E+08	0.205746	0.655005	4.351243
Interaction	87842427	4	21960607	0.005136	0.999942	2.866081
Within	8.55E+10	20	4.28E+09			
Total	3.10E+11	29				

Total peak with threshold 100 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Control : 13.1;13.3;13.5;13.7;13.9	7746.133	4	1936.533	7.293911	0.000858	2.866081
None and light smoothing	346.8	1	346.8	1.306215	0.266578	4.351243
Interaction	108.5333	4	27.13333	0.102197	0.980432	2.866081
Within	5310	20	265.5			
Total	13511.47	29				

Total area with threshold 100 rfu $\alpha = 0.05$

Source of Variation	SS	df	MS	F	P-value	F crit
Control : 13.1;13.3;13.5;13.7;13.9	1.68E+11	4	4.20E+10	17.25114	2.87E-06	2.866081
None and light smoothing	2.19E+08	1	2.19E+08	0.089939	0.767349	4.351243
Interaction	67468279	4	16867070	0.006933	0.999895	2.866081
Within	4.87E+10	20	2.43E+09			
Total	2.17E+11	29				

Appendix 6:

Table 2.19. Effect smoothing on the presence 5´-TRFs in sample 13 at 3 cm depth after *Alu*I digestion . The marker is MapMarker1000® and peak amplitude threshold is 100 rfu.

No.	None smoothing			Light smoothing			No.	None smoothing			Light smoothing		
	Size	Height	Area	Size	Height	Area		Size	Height	Area	Size	Height	Area
1	59.57	6286	62067	59.57	6231	62480	53	228.69	240	2231	228.69	228	2226
2	61.12	809	8783	61.12	787	8537	54	232.53	159	974	232.45	143	963
3	66.76	5516	39132	66.76	5147	40059	55	233.21	197	695	233.21	164	692
4	67.9	2274	28768	67.9	2136	27876	56	234.03	211	988	234.03	193	977
5	75.25	379	2837	75.25	359	3107	57	235.91	2364	25592	235.91	2240	25584
6	76.23	4770	31477	76.23	4406	31631	58	237.86	3504	36040	237.86	3296	36033
7	79.25	223	1648	79.25	206	1627	59	240.94	5088	35153	240.94	4808	35144
8	82.35	207	1648	82.35	195	1631	60	242.51	932	5674	242.44	860	5721
9	113.45	461	4660	113.45	447	4959	61	243.49	1287	14049	243.49	1279	14127
10	114.26	354	2990	114.26	330	2976	62	244.61	2882	17876	244.61	2668	17968
11	123.96	254	1810	123.96	229	1822	63	246.04	1384	15100	246.04	1349	15077
12	128.78	821	5204	128.78	752	5179	64	247.68	538	3691	247.76	477	4866
13	138.57	200	1202	138.57	182	1189	65	248.28	470	2294	-	-	-
14	143.87	106	1015	-	-	-	66	249.25	2659	24919	249.25	2521	25336
15	146.12	155	1495	146.12	150	1470	67	250.75	1153	8110	250.75	1090	7945
16	149.23	180	907	149.23	161	896	68	251.8	3391	38583	251.8	3218	38869
17	150.24	330	1535	150.24	296	1528	69	254.43	1481	14507	254.36	1427	14493
18	152.04	1926	23689	152.04	1853	23705	70	256.61	128	1056	256.54	120	1030
19	152.98	283	1286	152.98	244	1281	71	258.42	206	1103	258.42	186	1114
20	154.63	357	4758	154.63	345	4755	72	259.09	185	1284	259.09	175	1290
21	156.11	2085	13767	156.11	1952	13808	73	259.99	1182	8836	259.99	1113	8886
22	169.39	126	883	-	-	-	74	261.12	243	1642	261.12	228	1608
23	170.25	1456	11239	170.25	1346	11209	75	268.33	128	872	268.33	117	864
24	173.52	152	838	173.52	136	834	76	277.12	218	1799	277.12	207	1785
25	177.17	439	2797	177.17	400	2782	77	278.69	190	2000	278.69	183	2934
26	182.61	584	3809	182.61	541	3795	78	279.29	177	1235	-	-	-
27	186.64	107	1194	-	-	-	79	280.64	925	13141	280.64	908	13119
28	187.88	217	1225	187.88	199	1211	80	282.82	352	4363	282.82	330	4450
29	190.97	875	4935	190.97	778	4974	81	285.75	169	1870	285.75	164	1859
30	191.67	614	3775	191.67	571	3767	82	286.95	386	3460	286.95	371	3442
31	192.67	433	2272	192.67	389	2289	83	292.2	100	437	-	-	-
32	193.98	6419	49166	193.98	6080	49319	84	293.77	322	4507	293.77	309	4479
33	195.06	8652	92219	195.06	8550	92674	85	295.65	125	1028	295.58	120	1019
34	196.84	268	1145	196.84	236	1138	86	296.78	197	2128	296.78	189	2106
35	199.92	897	7137	-	-	-	87	309.31	178	1043	309.24	145	1025
36	200.61	1630	17227	200.68	1560	22845	88	371.57	152	2018	371.65	145	1970
37	201.75	2584	17344	201.75	2397	17404	89	436.2	211	1868	436.27	198	1856
38	202.97	128	898	202.97	120	884	90	437.42	130	1357	437.34	121	1342
39	204.18	4049	30332	204.18	3728	30304	91	441.39	1006	8814	441.39	950	8788
40	206.31	557	3946	206.31	517	3950	92	494.89	142	2380	494.89	137	2356
41	207.14	1278	8504	207.14	1182	8533	93	602.48	101	1126	-	-	-
42	209.27	1164	10527	209.27	1120	10524	94	607.44	114	1640	607.44	108	1611
43	210.33	2508	19631	210.33	2334	20018	95	610.41	125	1105	610.41	117	1084
44	211.16	723	4089	211.16	657	3874	96	639.91	161	3611	639.83	155	3553
45	217.22	3643	27407	217.22	3365	28473	97	682.49	134	3190	682.49	127	3131
46	217.98	1265	8251	217.98	1164	7639	98	713.22	778	13764	713.22	753	13713
47	219.26	310	1654	219.26	286	1667	99	719.37	376	3510	719.37	365	3553
48	221.38	2059	32027	221.38	1973	32030	100	735.29	105	1281	735.2	101	1260
49	223.49	697	4666	223.49	644	4712	101	741.97	118	2630	741.97	113	2577
50	224.32	369	2770	224.32	345	2770	102	820.64	1101	17092	820.59	1083	17064
51	226.28	994	6806	226.28	913	6793	103	822.55	295	3208	822.59	287	3192
52	227.56	888	6425	227.56	827	6418							

Appendix 7:

Table 2.20. The ANOSIM results for the nine replicate profiles.

Subsample		5'-TRF datasets											
		A fixed window size 0.5 bp										A fixed window size 0.5 bp	
		50.25 bp*		50.10 bp*		50.20 bp*		50.30 bp*		50.65 bp*		50.50 bp*	
		R	p	R	p	R	p	R	p	R	p	R	p
1.5cm-AluI	2cm-AluI	0.89	0.001	0.98	0.001	0.98	0.001	0.67	0.001	0.99	0.001	0.89	0.001
	2.5cm-AluI	0.96	0.001	0.98	0.001	0.99	0.001	0.85	0.001	1	0.001	0.93	0.001
	3cm-AluI	0.96	0.001	0.99	0.001	0.99	0.001	0.9	0.001	0.99	0.001	1	0.001
2cm-AluI	2.5cm-AluI	0.95	0.001	1	0.001	1	0.001	0.75	0.001	1	0.001	1	0.001
	3cm-AluI	0.92	0.001	0.97	0.001	0.95	0.001	0.76	0.001	0.96	0.001	0.99	0.001
2.5cm-AluI	3cm-AluI	0.77	0.001	0.87	0.001	0.85	0.001	0.66	0.001	0.85	0.001	0.92	0.001

Subsample		3'-TRF datasets											
		A fixed window size 0.5 bp										A fixed window size 0.5 bp	
		50.25 bp*		50.10 bp*		50.20 bp*		50.30 bp*		50.65 bp*		50.50 bp*	
		R	p	R	p	R	p	R	p	R	p	R	p
1.5cm-AluI	2cm-AluI	0.65	0.001	0.8	0.001	0.67	0.001	0.69	0.002	0.79	0.001	0.94	0.001
	2.5cm-AluI	0.78	0.001	0.81	0.001	0.73	0.001	0.74	0.001	0.79	0.001	0.99	0.002
	3cm-AluI	0.73	0.001	0.78	0.001	0.8	0.001	0.82	0.001	0.76	0.001	0.99	0.001
2cm-AluI	2.5cm-AluI	0.86	0.001	0.99	0.001	0.82	0.001	0.77	0.002	0.96	0.001	0.92	0.001
	3cm-AluI	0.84	0.001	0.81	0.001	0.9	0.001	0.8	0.001	0.82	0.001	0.94	0.001
2.5cm-AluI	3cm-AluI	0.76	0.001	0.73	0.001	0.83	0.001	0.63	0.001	0.71	0.001	0.9	0.001

* starting point

Table 2.21. The ANOSIM results for the triplicate profiles.

Subsample		A fixed window size 0.5 bp				A fixed window size 1 bp			
		5'-TRF		3'-TRF		5'-TRF		3'-TRF	
		50.25 bp*		50.25 bp*		50.50 bp*		50.50 bp*	
		R	p	R	p	R	p	R	p
1cm-AluI	3cm-AluI	1	0.1	1	0.1	1	0.1	1	0.1
	5cm-AluI	1	0.1	1	0.1	1	0.1	1	0.1
	7m-AluI	1	0.1	1	0.1	1	0.1	1	0.1
	9cm-AluI	1	0.1	1	0.1	1	0.1	1	0.1
3cm-AluI	5cm-AluI	0.59	0.2	0.33	0.2	1	0.1	0.56	0.1
	7m-AluI	1	0.1	0.7	0.1	1	0.1	0.89	0.1
	9cm-AluI	1	0.1	0.44	0.1	1	0.1	0.82	0.1
5cm-AluI	7m-AluI	1	0.1	0.85	0.1	1	0.1	0.78	0.1
	9cm-AluI	1	0.1	0.63	0.1	1	0.1	0.63	0.1
7m-AluI	9cm-AluI	1	0.1	0.41	0.1	1	0.1	0.44	0.1

* starting point

Chapter 3

Bacterial community in the intertidal sediments populated by *Arenicola marina*

3.1. Abstract

The lugworm *Arenicola marina* is a burrowing polychaete and lives semi-permanent in a U-shaped burrow consisting of head shaft, tail shaft and gallery tube inside marine sediments while digesting subsurface and sunk down surface sediment and defecating faeces at the sediment surface. It irrigates the burrow with oxygenated water (bioirrigation) and reworks the sediment during feeding and burrowing activity (bioturbation). The lugworm influences bacterial community, directly due to the feeding on microbes and indirectly through changes in the biogeochemical environment. In this study, the influence of the lugworm on bacterial populations was investigated with the terminal restriction fragment length polymorphism (T-RFLP). The T-RFLP analyses clearly grouped the bacterial population in the head shaft tube with sediment surface populations over depth, while the tail shaft tube was populated by different populations; the surface bacteria at 0-2 cm and the subsurface bacteria at 2-10 cm depth. The populations in the gallery tube were similar to those in the head and tail shaft tube. The absence of *A. marina* was reported to change the fine structure of the sediment. T-RFLP analyses showed that the presence of *A. marina* had no clear detectable influence on surface sediment populations, however, the T-RFLP analyses of two mm thick sediment layers revealed a strong depth-dependence of the community composition.

3.2. Introduction

The lugworm *Arenicola marina* is a marine burrowing polychaete that bioturbates and bioirrigates the sediments. It lives head down and relatively permanent in a 20 – 40 cm deep J-shaped burrow with an adult body length of about 15 to 25 cm. The burrow is completed to a U-shaped burrow by a vertical head shaft, though which surface sediment slides down into a feeding pocket and is ingested by the worm. As a result sand above the lugworm head depresses downward forming a feeding funnel in the surface. During defecation, the lugworm moves backwards through the tail shaft until the tail reaches the sediment surface and ejects a characteristic fecal cast. Ingestion and defecation is a cyclical pattern that effects a sediment particle movement (Kristensen, 2001; Riisgard and Banta, 1998). For respiration the lugworm actively irrigates its burrow with oxygen-rich overlying water with a peristaltic movement in a posterior-anterior direction (Riisgard and Banta, 1998). With a lugworm density of 30 individual/m^2 approximately 3 L h^{-1}m^{-2} oxygen-rich overlying water is pumped into the sediment (Riisgard *et al.*, 1996). The active irrigation period is 5-10 minutes, interrupted by short periods of inactivity (Kristensen, 2001). The continuous irrigation results in highly oxic and oxidized conditions in the burrow zone and surrounding sediments (Kristensen, 2001; Banta *et* al., 1999; Kristensen, 1985) and removes porewater nutrients and deeper microbial products, e.g. sulfide (Riisgard *et al.*, 1996; Volkenborn *et al.*, 2006).

A. marina may affect the microbial community directly by feeding on them. It assimilates bacteria, microphytobenthos, microfauna and meiofauna associated with the sediment. Its foregut had higher bacterial numbers than the feeding funnel, the head shaft, and the feces while the hindgut had the lowest number (Grossmann and Reichardt, 1991; Retraubun *et al.*, 1996). The bacteriolytic rate in its digestive fluid was also seasonally fluctuated (Plantae and Mayer, 1994). Indirectly *A. marina* affects the bacterial community by changing the biogeochemical environment through biotubative and bioirrigative activities (Volkenborn *et al.*, 2007; Volkenborn *et al.*, 2007a; Nielsen *et al.*, 2003; Banta *et al.*, 1999; Hüttel, 1990)

Many studies had examined the effect of *A. marina* on the sediment with the geochemical parameters that referred to the microbial processes aerobically and anaerobically whether directly in the U-shaped burrow, as performed by Nielsen *et al.* (2003) or indirectly in the bulk sediment surrounding the burrow, e.g. recently

published by Volkenborn *et al*. (2007 and 2007a). But studies on the bacterial community are rare, especially for sample taken directly from the U-shaped burrow (Retraubun *et al.*, 1996; Grossmann and Reichardt, 1991; Reichardt, 1988). While the burrow can be considered as a physically stable habitat on a day or week time scale and a chemically unstable habitat of oxic-anoxic change due to bioirrigation, it may support different microbial growths in the burrow (Kristensen, 2001). Therefore as a complement study and assuming that *A. marina* may provide unique niches for microbial populations, we investigated the bacterial community change directly along the U-shaped burrow from the surface to 10 cm by applying the terminal restriction fragment length polymorphism (T-RFLP) method.

T-RFLP is a fingerprinting method consisting of DNA isolation, PCR amplification enzyme restriction and capillary electrophoresis. Since the primers are labeled with fluorescent dyes, so that only the fluorescent terminal restriction fragment (TRFs) are detected and quantified by a high resolution gel electrophoresis on an automated DNA sequencer. The method relies on variations in the position of restriction sites among 16S rRNA gene sequences, thus the bacterial diversity of complex community is determined as a pattern composite of the number of fluorescently labeled TRFs with unique length sizes in base pairs and the intensity of each TRF in relative fluorescent unit (rfu) (Liu *et al.*, 1997, Dunbar *et al.*, 2001). We tested already the method in the previous study (Chapter 2) with sediments from the Wadden Sea and found a technically generated dissimilarity or artificial generated biodiversity of 10 to 20 percent in the TRF replicate dataset of a representative sample. But the study showed that the biological variation represented by bacterial communities in adjacent 0.5 cm and 1 cm thick sediment layers was larger than the technical variation which may potentially arise from the applied molecular steps, routinely small pipetting errors, raw data analysis or statistical analysis (Blackwood *et al.*, 2007; Frey *et al.*, 2006; Saikaly *et al.*, 2005: Engebretson and Moyer, 2003; Dunbar *et al.*, 2001; Liu *et al.*, 1997).

This study is a high resolution T-RFLP application for detecting the bacterial community change with depth. A resolution of 1 cm sliced sediment layer was applied for tracking the change along the U-shaped burrow of *A. marina*. Sediment samples were taken from the head shaft, the tail shafts and the gallery tube. The microbial populations in those compartments were compared with those in the bulk sediment. The bulk sediment was apart from the head shaft and the tail shaft and

assumed to be only indirectly influenced by the lugworm activities. The change in the surface bacterial community due to the presence of *A. marina* was also investigated with a 2 mm layer resolution.

3.3. Materials and methods

3.3.1. Sample area and sediment samples

Sediment samples were collected in October 2005 from populated and populated area developed by Volkenborn *et al.* (2007) on a low intertidal sandy flat in the Königshafen at the northern end of the island of Sylt in the North Sea, Germany (55°02´N, 8°26´E). The not populated area of 20 m x 20 m was achieved by inserting a 1mm meshed net in 10 cm depths (Volkenborn *et al*., 2007; Volkenborn *et al*., 2007a). The populated area was approximately inhabited by 20-30 individuals/m^2 and characterized by fecal casts at the sediment surface. Vice versa, the not populated area had smooth surface indicating the absence of *A. marina* (Figure 3.1.A). The population was found to play an important role for physical and chemical processes and the benthic community (Volkenborn *et al*., 2007; Volkenborn *et al*., 2007a). The biochemical habitat description is presented in Table 3.1.
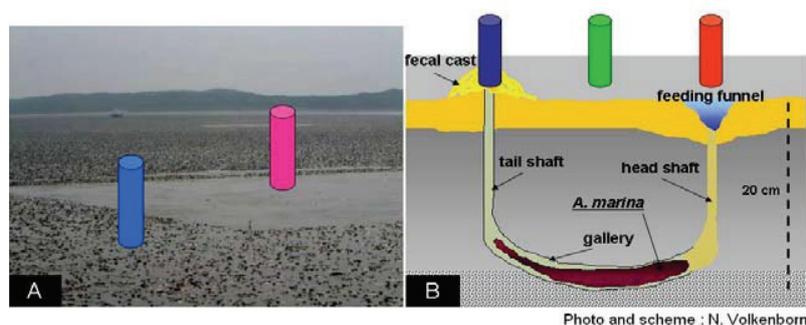
Fig. 3.1. (A) Smooth and non smooth sediments surface indicating the absence and presence of *A.marina*. Duplicate sample cores were put into both plots. (B) Scheme of the lugworm burrow and sediment collected cores put into feeding funnel, middle area and fecal cast. Sediments samples were sliced into two different resolutions of 0.2 and 1 cm over depth respectively for A and B.

Duplicate piston cores (2.5 cm x 20 cm) were taken from the populated and not populated areas (Fig. 3.1.A). A series of piston cores was placed into the U-shaped burrow directly (Fig. 3.1.B); in the feeding funnel (the head shaft tube), the bulk sediment and the fecal funnel (the tail shaft tube). The sediments samples were immediately brought to the harbor laboratory and sliced within five hours with a 0.2 cm sliced resolution for the populated and not populated samples and 1 cm for the U-shaped burrow samples. The populated and not populated samples each had 10 subsamples presenting 10 different layer depths (0-0.2 cm, 0.2-0.4 cm, 0.4-0.6 cm, 0.6-0.8 cm, 0.8-1 cm, 1-1.2 cm, 1.2-1.4 cm, 1.4-1.6 cm, 1.6-1.8 cm and 1.8-2 cm). For the U-shaped burrow cores, only the visible brownish sediment parts of the burrow tube were collected. The U-shaped burrow samples had also 10 subsamples presenting 10 different layer depths (0-1 cm, 1-2 cm, 2-3 cm, 3-4 cm, 4-5 cm, 5-6 cm, 6-7 cm, 7-8 cm, 8-9 cm and 9-10 cm). The gallery tube sediments were collected by digging the sediments vertically to open the U-shaped burrow and scarping the brownish sediments. The sliced and scarped sediment samples were frozen immediately and stored at -20°C until further analysis.

Table 3.1. Sediment biogeochemistry in the sample area from Volkenborn *et al.*, 2007 and 2007a.

		Bioturbated plot	Non-bioturbated plot
Emersion period		9-10 h / tide	
Main hydrodynamic force		Tidal current	
Mean tidal height		$\approx$ 1.8 m	
Salinity		27.5 $^{o}/_{oo}$ in spring	
		31 $^{o}/_{oo}$ in summer	
Average *A. marina* density		18-30 ind / m^2	-
Grain size of fine sand sediment		0 -1 cm : 204 μm	0 -1 cm : 190 μm
		1 - 5 cm : 218 μm	1 - 5 cm : 206 μm
Fine fraction (particle size <63 μm) at 0 - 5 cm depth		<1% dry wt	1% - 2.5% dry wt
Water content in 0 – 5 cm depth		16.20%	19.90%
Sediment permeability		2.6 X 10^{-12} m^2	<1.0 X 10^{-12} m^2
Approximately porewater profile through 0-10 cm depth	Ammonium	<100 μM	<150 μM
	Nitrite	<0.3 μM	<0.25 μM
	Nitrate	<2.5 μM	<2.5 μM
	Phosphate	<10 μM	<15 μM
	Sulphide	<25 μM	<150 μM
Average oxygen penetration		<4 cm depth	<1 cm depth

3.3.2. Terminal restriction fragment length polymorphism (T-RFLP)

Triplicate T-RFLP datasets of a depth subsample were generated from one genomic DNA preparation, followed by three parallel PCR and continued into three individual restriction enzyme digestions and three individual fragment runs (Fig. 3.2).
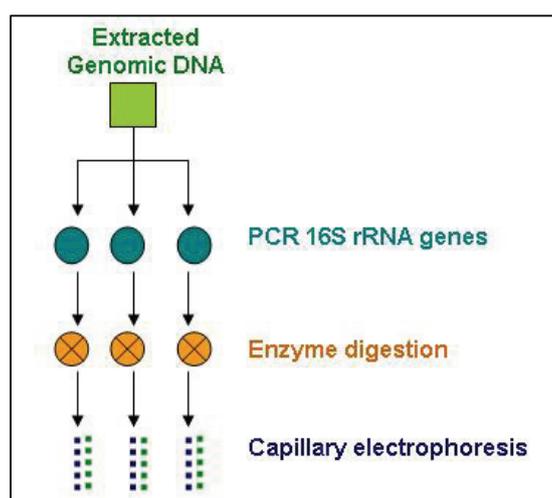
Fig. 3.2. A strategy for generating triplicate datasets.

DNA extraction. Total genomic DNA from 0.5 g sediment samples was extracted using the FastDNA SPIN Kit for soil (Qbiogene, Carlsbad, USA) following the Manufacturer's protocols with a minor modification. The sample was spinned twice at 14.000 g for 30 seconds and the extracted genomic DNA was diluted twice, each in 25 µl PCR water. Extracted genomic DNA was measured qualitatively and quantitatively with agarose electrophoresis on 1.5% gels and a NanoDrop ND-1000 UV spectrophotometer (PEQLAB Biotechnology, Erlangen, Germany).

Polymerase chain reaction (PCR). The target bacterial 16S rRNA genes were amplified with the primer pair 6-FAM-27F (AGAGTTTGATCCTGGCTCAG; Amann *et al*., 1995) and HEX-907R (CCGTCAATTCCTTTRAGTTT; Muyzer *et al*., 1995). The forward primer was labeled with 6-carboxyfluorescein and the reverse primer with 6-carboxy-2',4,4',5',7,7'-hexachloro-fluorescein at the 5´-terminus. The 25µl-PCR reaction consisted of 12.5 µl PCR Master Mix (Promega, USA), about 1 ng template of genomic DNA, 4 pmol/µl of each forward and reverse primer. PCR was performed in a Mastercycler Personal (Eppendorf, Germany) with the following

thermal conditions: 5 minutes initial denaturation at 94°C and 35 cycles consisting of 4 minutes denaturation at 94°C, 1 minute annealing at 50°C and 1 minute elongation at 72°C. At the end, elongation was extended for 10 minutes at 72°C. The PCR products were passed through Sephadex™ G-50 Superfine columns according to a protocol suggested by Applied Biosystems (California, USA).

Enzyme digestion and capillary electrophoresis. Approximately 100 ng of PCR product was digested in 20 µl reaction volumes independently with 5 U of the restriction enzyme *Alu*I, *Hha*I or *Msp*I for 3 hours at 37°C. Afterwards *Alu*I and *Msp*I were inactivated at 65°C for 20 minutes, while *Hha*I was inactivated at 80°C for 20 minutes. After desalting through Sephadex™G-50 Superfine columns, 5 µl digested product was mixed with 20 µl standard internal marker:deionized formamide (1:60 v/v) and denatured at 95°C for 10 minutes and placed immediately on ice. The marker was MapMarker1000® (BioVentures Inc. Murfreesboro, TN). The digested product was loaded onto a capillary ABI Prism 3130XL genetic analyzer for terminal restricted fragment (TRF) separation.

TRF detection and analyzing. The separated TRFs were detected and analyzed by using the GeneMapper® Software v3.7 (ABI) with the default parameter values provided by the software that previously had been evaluated the stability in detecting and analyzing the TRFs from our sediment samples (Chapter 2). The observed TRFs were identified by the length size (in base pairs) and the abundance represented as area intensity (rfu x time).

The TRF length size was determined with a precision of 0.01 bp. The T-RFLP datasets were priory modified before further statistical analysis, as below:

1. A TRF with a peak height of less than 100 rfu was excluded by applying a peak amplitude threshold 100 rfu.

2. A TRF with length size of less than 50 bp and more than 900 bp was removed.

3. Irreproducible electropherogram pattern within triplicate datasets were excluded; therefore not all depth dependent subsamples had 5´-TRF and 3´-TRF triplicate datasets.

4. The TRFs were exported in a text file containing 3 columns: sample name, length size in bp and area. These raw TRF datasets were the input and imported into R for binning.

5. The datasets were binned in a fixed window size 1 bp using free software R (www.r-project.org) with a language program written by A. Ramette (unpublished).

During the binning process, each dataset was normalized by proportioning the area of each TRF with the total areas in that particular dataset. The binned TRFs presented by their size in bp and their abundance in relative area (rfu x time) were the final TRFs dataset for statistical analysis.

3.3.3. Statistical analysis

The biological variations of the TRF datasets were represented by multivariate analysis: Non-metric Multi Dimensional Scaling (NMDS) and cluster analysis with the statistical software packages Primer 5 for windows version 5.2.0 (Primer-E Ltd, Plymouth, UK). The similarity matrix for NMDS, cluster analysis, analysis of similarity (ANOSIM) and SIMPER analysis (similarity percentages-species contributions) was calculated based on the presence of 5´-TRFs or 3´-TRFs (the richness) and their abundance in relative area (the evenness) after the Bray-Curtis coefficient without data transformation and standardization. Cluster mode of group average was used to construct the cluster analysis based on group average linkage.

Non-metric Multi Dimensional Scaling (NMDS) ordinates the relationship between the datasets in space dimension. A hundred random restarts were used for calculating the iterative algorithm as NMDS is sensitive to the initial configuration (Kenkel and Orloci, 1986; Rees *et al.*, 2004). The software Primer 5 presents the ordination with the lowest stress value obtained in 100 calculations. Stress value is a measure of deviation for ´a goodness of fit´ of the iterative algorithm (Kenkel and Orloci, 1986; Rees *et al.*, 2004). The NMDS ordination was interpreted following Clarke and Warwick (2001): stress value < 0.1 = ordination is an ideal ordination without potential misinterpretation, stress value < 0.2 = ordination is a useful 2 dimensional ordination and stress value > 0.2 = ordination was random. By using the same similarity matrix, a cluster analysis was performed to compare the similarity within triplicate datasets of one depth dependent subsample, between depth dependent subsamples and between the areas.

Due to the applied layer resolution (0.2 cm), the depth dependent subsample represented points are probably very close to each other in the NMDS ordination. The ANOSIM test was performed to analyze the separation degree of the surface

layer samples from the populated and not populated area. Clarke and Gorley (2001) interpretation of the R statistic for a pair wise group was used: R > 0.75 = groups well separated, R > 0.5 = groups overlapping but clearly different, R < 0.25 = groups barely separable.

Contribution of each 5´-TRF and 3´-TRF to the similarity within triplicate datasets of one depth dependent subsample and to the dissimilarity between depth dependent subsamples and area was calculated with the Simper analysis (similarity percentages-species contributions). The list of 5´-TRFs or 3´-TRFs was cut off at 90% contribution; fragments will be listed in decreasing order of their importance in contributing to the average dissimilarity between two groups until 90% of the dissimilarity is explained.

3.4. Results

3.4.1. The U-shaped burrow of *A. marina*

Observed TRF numbers and similarity percentage of triplicate datasets. Average numbers of observed 5´ and 3´-TRFs were presented in Fig. 3.3. The U-shaped burrow tubes gained a fluctuating TRF number with depth, but did not indicate depth dependence. Looking at to the enzyme digestion, the highest numbers was obtained after *Msp*I, then after *Hha*I digestion and *Alu*I digestion. Even if it produced the lowest TRFs number but the enzyme *Alu*I digestion generated the highest reproducible TRFs number within triplicate datasets with the lowest standard deviation. Approximately three fourth of the TRF numbers were reproducible TRFs after *Alu*I digestion with average similarity percentages 70% – 90%. In contrast, enzyme *Msp*I digestion generated lower reproducible TRF numbers (about two third of the observed TRFs number) with the most fluctuating similarity percentages (50% - 90%). Looking at to the tube part of the U-shaped burrow, the gallery tube had the highest standard deviation and the lowest reproducible TRF number with similarity percentage in range 60% - 80% while the numbers from the head and tail shaft tubes remained comparable.
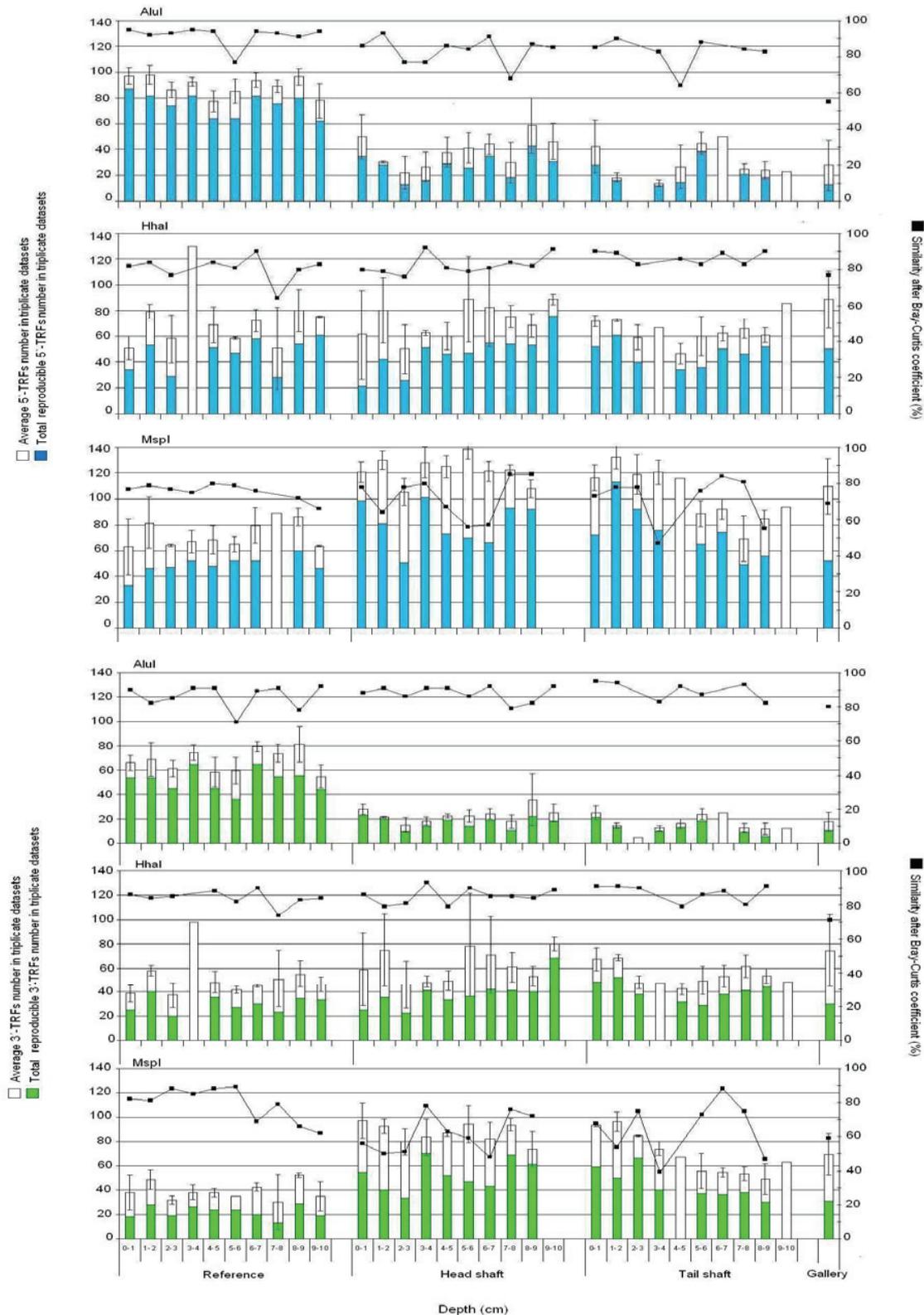
Fig. 3.3. The U-shaped burrow tubes and the bulk sediment. Average number and average similarity percentage of observed 5´and 3´-TRFs with depth after *Alu*I, *Hha*I and *Msp*I digestion. The similarity was counted after Bray-Curtis coefficient based on presence/absence and TRF abundance.

In the bulk sediment, the reproducible TRFs number within triplicate datasets was higher than those detected in the U-shaped burrow. The observed TRFs generated after *Alu*I digestion were mostly reproducible with similarity percentage about 90% within triplicate datasets. Enzyme *Hha*I and *Msp*I digestion also generated a high similarity of 80% but with a lower reproducible number than it after *Alu*I digestion.

The biological variation. The biological variation referred to the bacterial composition consisted of richness and evenness. This variation which may occur in the depth dependent subsamples was represented by the multivariate analysis Non-metric Multidimensional Scaling (NMDS) and the cluster analysis. The triplicate 5´-TRF or 3´-TRF datasets of one depth subsample are represented by three same points arranged in two dimensions space which usually form a cluster indicating a high similarity in bacterial composition. Distances between clusters relate to the dissimilarity among the depth dependent subsamples.
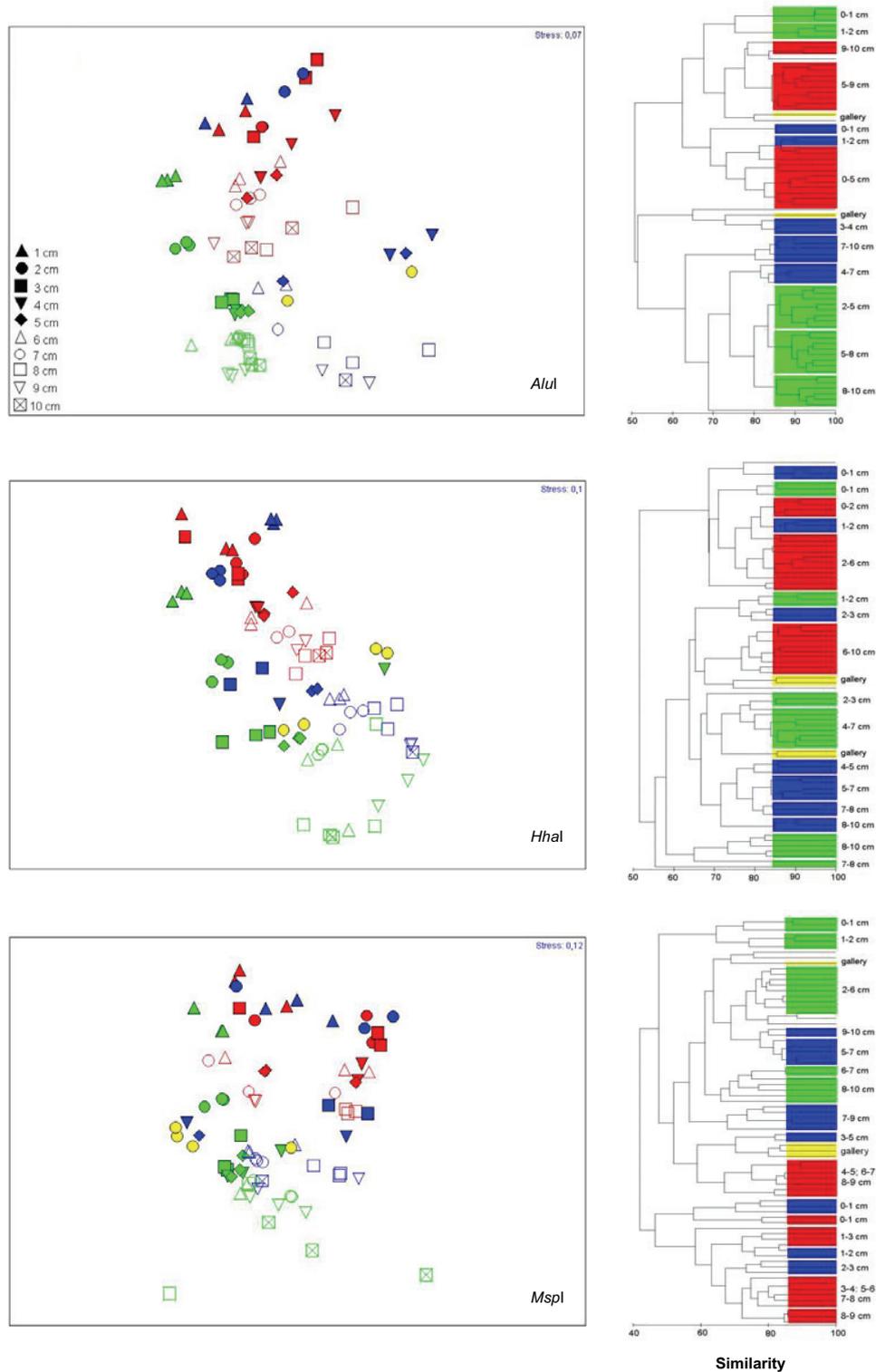
Fig. 3.4. The U-shaped burrow tubes and the bulk sediment. Bray-Curtis similarity based NMDS ordinations and similarity clusters of 5′-TRF datasets after *Alu*I, *Hha*I and *Msp*I digestion represents a bacterial community shifting with depth. The color code indicates green, blue, red and yellow respectively for the bulk sediment, tail shaft tube, head shaft tube and gallery tube.

From the U-shaped burrow, the 5´-TRF and 3´-TRF dataset after *Alu*I, *Hha*I and *Msp*I digestion generated a same NMDS pattern: a similar and likely mixed bacterial community in the head shaft tube over depth (red points in Fig. 3.4 and 3.5) and a bacterial community shifting with depth in the tail shaft tube from 0 to 10 cm (blue points in Fig. 3.4 and 3.5). The bacterial community shifting was more obviously detected in the bulk sediment than those in the U-shaped burrow, as the subsample represented clusters were distinguishable along the depth (green points in Fig. 3.4 and 5). From the tail shaft tube and the bulk sediment, the surface bacterial community at 0-1 cm depth was different with that at 1-2 cm depth. And the bacterial communities at the surface were different with those at subsurface (below 2 cm depth), as surface clusters were distant from the subsurface clusters.

The bacterial shifting in the U-shaped burrow was more clearly represented by the cluster analysis (Fig. 3.4 and 3.5). The cluster analysis may support the NMDS ordination in representing the relationship between the datasets (Clarke and Gorley, 2001). The bacterial communities in the head shaft tube performed big clusters and grouped with those found at the surface layers in the tail shaft tube and the bulk sediment. The 5´-TRF dataset showed that the bacterial clusters were at 0-5 cm and 5-10 cm after *Alu*I digestion or at 0-2 cm, 2-6 cm and 6-10 cm after *Hha*I digestion. While the 5´-TRF dataset after *Msp*I digestion revealed a different relationship; the head shaft bacterial populations clustered together with the surface bacterial community in the tail shaft tube but separated from those in the bulk sediment as also indicated by the 3´-TRF dataset after *Alu*I digestion. The 3´-TRF dataset after *Hha*I digestion showed the similar pattern of those indicated by the 5´-TRF dataset after *Alu*I and *Hha*I digestion. In the gallery tube, the bacterial communities were split into the surface and subsurface clusters even they were from one depth ($\pm$ 20 cm). The subsurface bacterial clusters in the tail shaft tube were at 3-4 cm, 4-7 cm and 7-10 cm based on the 5´-TRF dataset or at 3-6 cm and 6-10 cm based on the 3´-TRF dataset after *Alu*I digestion. The 5´-TRF and 3´-TRF dataset after *Hha*I and *Msp*I digestion represented the similar bacterial cluster.

The bacterial community shifting in the bulk sediment was similarly represented by 5´-TRF and 3´-TRF datasets generated after *Alu*I, *Hha*I and *Msp*I digestion. The surface bacterial community was distinct found at 0-1 cm and 1-2 cm. The subsurface bacterial communities were at 2-5 cm, 5-8 cm and 8-10 cm based on the 5´-TRF dataset or at 2-3 cm, 3-5 cm, 5-9 cm, 9-10 cm based on the 3´-TRF

dataset after *Alu*I digestion. Fig. 3.6 and Table 3.2 showed how the bacterial community shifting in the bulk sediment after *Alu*I digestion related to the presence and abundance of a particular TRF over depth. The 5´-TRFs and 3´-TRFs were mostly the same, but the abundance was different; they were stable (e.g. 5´-TRFs of 59.5, 67.5, 200.5, 237.5, 280.5 and 820.5 bp), increasing (5´-TRFs of 66.5, 75.5, 151.5, 216.5, 220.5, 244.5 and 251.5 bp) or decreasing (5´-TRFs of 194.5, 201.5, 203.5, 209.5, 245.5 and 248.5 bp) with depth. But several of them were unique for surface layers (5´-TRFs of 68.5, 190.5, 192.5, 712.5 and 821.5 bp) and subsurface layers (5´-TRF of 153.5, 169.5, 253.5 and 440.5 bp).
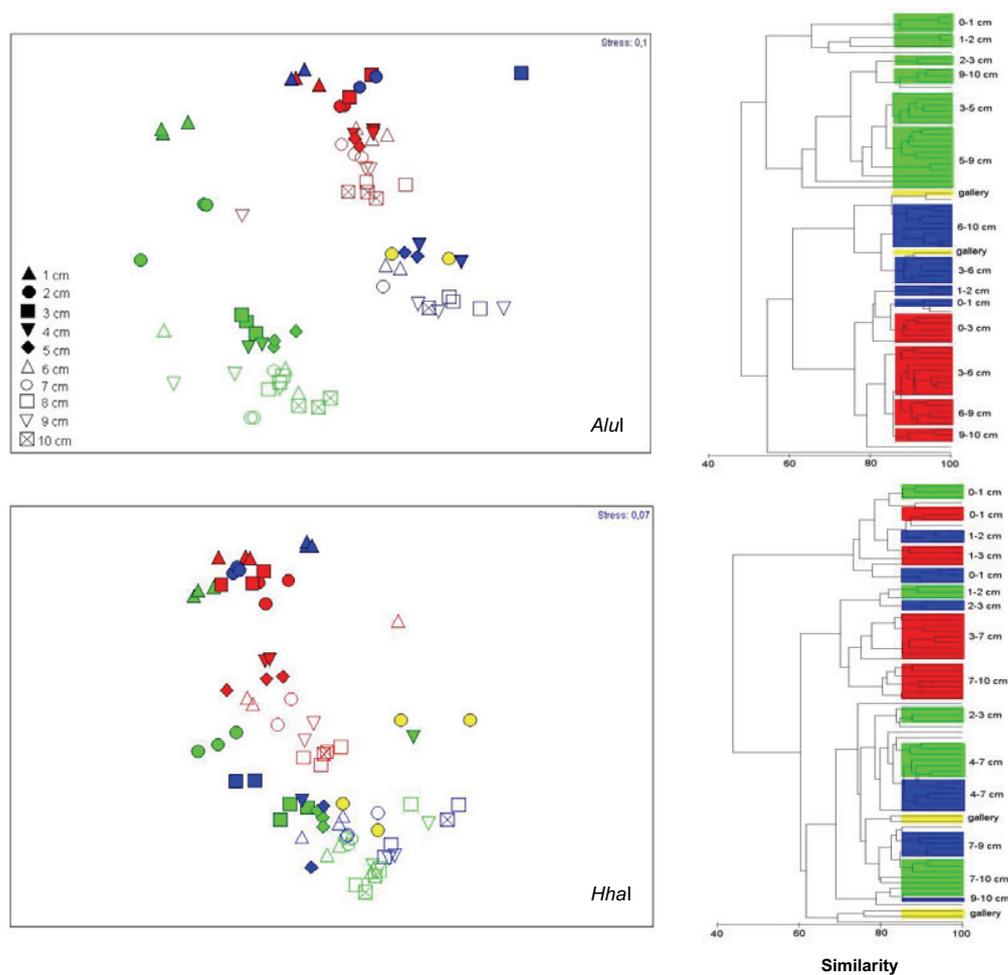
Fig. 3.5. The U-shaped burrow tubes and the bulk sediment. Bray-Curtis similarity based NMDS ordinations and similarity clusters of 3´-TRF datasets after *Alu*I and *Hha*I digestion represents a bacterial community shifting with depth. The color code indicates green, blue, red and yellow respectively for the bulk sediment, tail shaft tube, head shaft tube and gallery tube.

The richness and evenness of 5´-TRFs and 3´-TRFs differentiated the bacterial community between the tail shaft tube and the bulk sediment. The tail shaft tube was less richness than the bulk sediment (Table 3.2). But the total of relative average abundance indicated that the tail shaft tube had a higher evenness of a lower richness. Several 5´-TRFs that not or low detected in the tail shaft were abundant in the bulk sediment; e.g. 5´-TRFs of 59.5, 153.5, 192.5, 199.5, 235.5, 253.5, 440.5 and 712.5 bp or 3´-TRFs of 58.5, 72.5, 187.5, 349.5, 350.5 and 472.5 bp after *Alu*I digestion. Both areas had unique surface and subsurface TRFs. The 5´-TRF of 191.5 and 236.5 bp was only detected in the tail shaft area, while 192.5, 235.5 and 712.5 of 5´TRF was only in the bulk sediment.
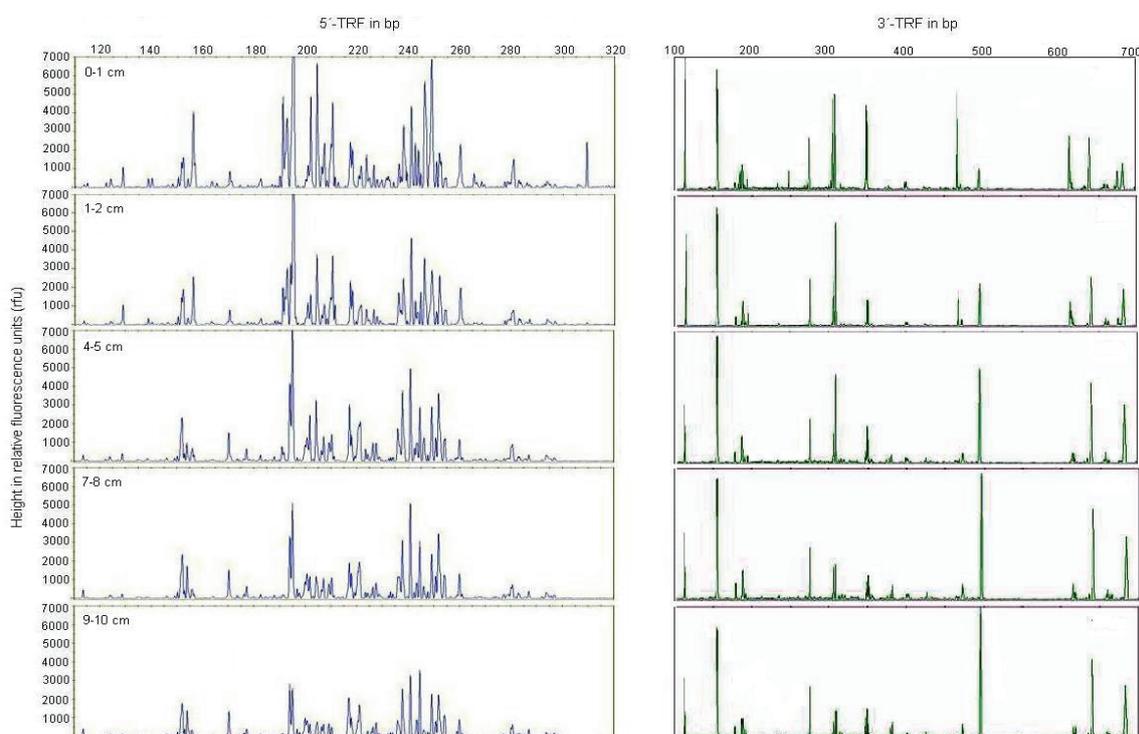
Fig. 3.6. The bulk sediments. A serial electropherogram of 5´-TRF and 3´-TRF dataset over depth after *Alu*I digestion.

Table 3.2 also showed that due to the unique 5´-TRFs surface layer of 68.5, 190.5, 191.5, 821.5 bp or 3´-TRFs of 89.5, 467.5, 612.5 bp, the bacterial community in the head shaft tube was then clustered together with the surface layer bacterial communities in the tail shaft tube and the bulk sediment. Those TFRs were detected

gradually decreasing in the head shaft tube with depth. And the gallery tube yielded the lowest richness with the highest evenness compared to the tail and head shaft tubes. The bacterial population in the gallery tubes was spread into the surface and subsurface TRFs without presenting unique TRFs of those layers. The presence of twelve 5´-TRFs and seven 3´-TRFs respectively contributed 70% and 77% to the average abundance in the gallery tube.

Table 3.2. The U-shaped burrow tubes and the bulk sediment. The 5´-TRFs and 3´-TRFs and their relative average abundance over depth after *Alu*I digestion.

3.4.2. The surface layers of the populated and not populated area

Observed TRF numbers and similarity percentage of triplicate datasets. The surface bacterial community in the populated and not populated area was investigated from 0 to 2 cm depth with a resolution of 2 mm thick layers. Duplicate sample cores were analyzed from each area. Fig. 3.7 showed the observed 5-TRF and 3´-TRF number after *Alu*I, *Hha*I and *Msp*I digestion. Enzyme *Alu*I digestion generated more than three-fourth reproducible TRFs from the observed TRFs number in the triplicate 5-TRF and 3´-TRF datasets with similarity percentages about 80% - 90%.

The biological variation. The 5´-TRF and 3´-TRF dataset after *Alu*I, *Hha*I and *Msp*I digestion from the populated and not populated area revealed similar pattern of bacteria communities: a gradual bacterial community shifting with depth at the surface layer sediments (Fig. 3.8). Range R values of 0.5 – 1 of the ANOSIM test statistically confirmed that the bacterial communities tended to be different communities along individual layers, even if the depth dependent subsample representing points were close to each other due to a high resolution sliced sediment layer (2 mm thick) (Appendix 1).

Fig. 3.7. The surface bacterial community in the not populated and populated area. Average number and average similarity percentage of observed 5´and 3´-TRFs with depth after *Alu*I, *Hha*I and *Msp*I digestion. The similarity was counted after Bray-Curtis coefficient based on presence/absence and TRF abundance.
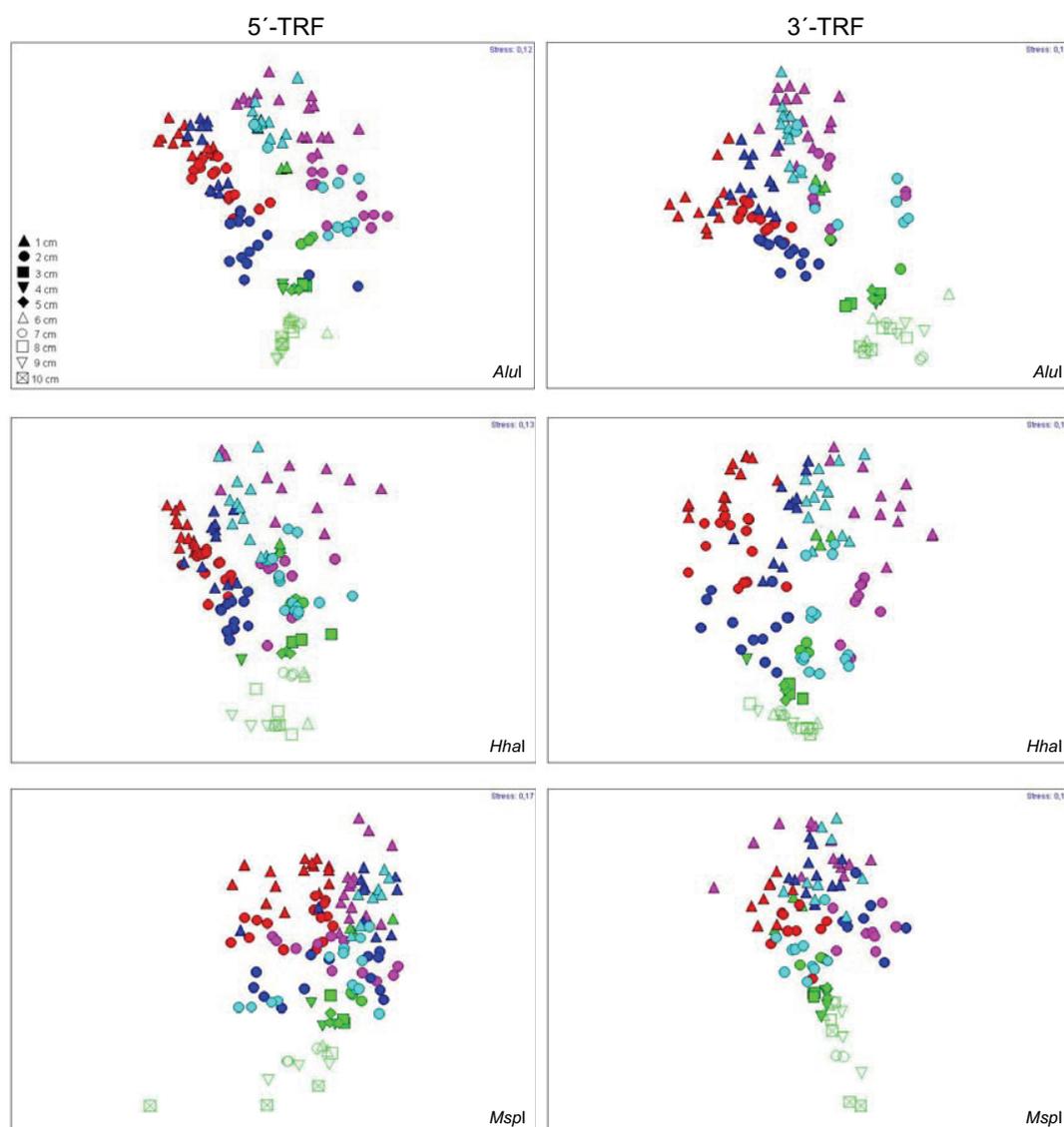
Fig. 3.8. The not populated and populated area. Bray-Curtis similarity based NMDS ordinations of 5´-TRF and 3´-TRF dataset after *Alu*I, *Hha*I and *Msp*I digestion. They presented a continuous bacterial community shifting with depth. The color code indicates green, red, pink, blue and light blue respectively for the bulk sediment, core I and core II of the not populated area, core I and core II of the populated area.

The bacterial communities in the not populated and populated area were similar. The core I represented points from both different areas were close to each other and likely made one cluster as well for core II, unless after *Msp*I digestion (Fig. 3.8). The dissimilarity percentage of bacterial composition between areas after *Alu*I digestion was 36% and 38% respectively for core I and core II (Table 3.3). The

dissimilarity between two areas was contributed by the presence and the abundance of about twelve 5´-TRFs or sixteen 3´-TRFs (Table 3.4), each TRF contributed above 1.5% to that dissimilarity percentage. There was no unique 5´-TRF and 3´-TRF found for each area. Mostly they were the same TRFs but different in abundances. The 5´-TRF of 200.5 bp after *Alu*I digestion was more abundance in the not populated area than it in the populated area and it gave the highest contribution (6%) to the dissimilarity percentage between those areas (Table 3.4).

Table 3.3. Dissimilarity percentage (%) after Bray-Curtis coefficient of the 5´-TRF and 3´-TRF datasets after *Alu*I, *Hha*I and *Msp*I digestion between duplicate cores (core I and core II) and areas (the populated and not populated area) with 2 mm sliced resolution.

Non-bioturbated		Bioturbated		5´-TRF datasets			3´-TRF datasets		
I	II	I	II	AluI	HhaI	MspI	AluI	HhaI	MspI
X	X			44	47	50	45	56	49
X		X		36	41	51	36	50	43
	X	X		47	51	53	45	57	47
X			X	42	47	52	46	55	43
	X		X	38	44	48	34	51	50
		X	X	42	44	50	42	48	46

Table 3.4. The average abundance of 5-TRFs and 3´-TRFs for the surface layers (0 to 2 cm depth) that contribute to the dissimilarity percentage between the duplicate sample cores and between two areas relating to the Table 3.4.

5'-TRF (bp)	Non-bioturbated area			Bioturbated area			Reference area		Non-bioturbated Core I / Bioturbated Core I			Non-bioturbated Core II / Bioturbated Core II			5'-TRF (bp)
	Core I Average abundance	Core II Average abundance	Contribution (%)	Core I Average abundance	Core II Average abundance	Contribution (%)	Average abundance	Contribution (%)	Core I Average abundance	Core I Average abundance	Contribution (%)	Core II Average abundance	Core II Average abundance	Contribution (%)	
58.5	1.9		2.18	2.28		2.72			1.9	2.28	3.07				58.5
59.5	1.49	3.93	3.53	0.71	2.54	2.83	4.96	5.65	1.49	0.71	2.22	3.93	2.54	3.86	59.5
66.5	0.51	1.8	2.21									1.8	0.29	2.5	66.5
67.5	3.78	3.54	1.54	3.06	2.82	1.81	2.9	3.13				3.54	2.82	2.51	67.5
68.5	6.21	4.25	2.92	5.45	5.22	4.62	3.58	3.28	6.21	5.45	4.44	4.25	5.22	4.59	68.5
191.5	1.35		1.46	1.64	3.05	2.7						0.2	3.05	3.91	191.5
192.5	0.03	2.29	2.61	0.12	0.87	1.09	3.85	4.37				2.29	0.87	2.71	192.5
194.5	5.83	14.56	10.33	4.58	12.67	9.68	2.9	3.13	5.83	4.58	3.97	14.56	12.67	7.36	194.5
200.5	5.24	0.52	5.6						5.24	0.93	6.41				200.5
201.5	0.78	3.91	4.22						0.78	2.03	3.13	3.91	2.91	2.38	201.5
216.5	5.1	4.44	1.3	3.99	2.66	2.2	1.69	1.99	5.1	3.99	2.15	4.44	2.66	2.81	216.5
237.5	4.94	3.14	2.49	3.63	3.29	1.22	2.7	3.21	4.94	3.63	2.41				237.5
240.5	1.07	1.29	0.89	3.3	2.1	1.72	2.85	3.05	1.07	3.3	3.13				240.5
242.5	1.87		2.15	1.42	0.36	1.39									242.5
245.5	5.74	4.3	1.83	6.56	6.45	2.66	4.45	4.92	5.74	6.56	3.05	4.3	6.45	3.14	245.5
248.5	6.11	5.52	2.29	5.93	7.96	4.35	5.72	6.15	6.11	5.93	4.15	5.52	7.96	4.31	248.5
251.5	1.19	1.03	1.19	2.15	0.83	2.01	2.27	2.21	1.19	2.15	1.79	1.03	0.83	1.51	251.5
821.5				2.06	3	2.34	2.03	1.76				1.4	3	2.81	821.5

3'-TRF (bp)	Non-bioturbated area			Bioturbated area			Reference area		Non-bioturbated Core I / Bioturbated Core I			Non-bioturbated Core II / Bioturbated Core II			3'-TRF (bp)
	Core I Average abundance	Core II Average abundance	Contribution (%)	Core I Average abundance	Core II Average abundance	Contribution (%)	Average abundance	Contribution (%)	Core I Average abundance	Core I Average abundance	Contribution (%)	Core II Average abundance	Core II Average abundance	Contribution (%)	
58.5	7.81	2.12	7.06	5.18	2.58	5.9	4.24	4.85	7.81	5.18	5.58	2.12	2.58	5.22	58.5
59.5	7.38	19.21	15.31	9.27	12.95	10.03	3.79	4.06	7.38	9.27	6.77	19.21	12.95	17.79	59.5
88.5	3.36		3.7						3.36		4.6				88.5
89.5	1.55	2.87	2.48	5.48	3.58	2.87	3.86	4.26	1.55	5.48	5.86				89.5
90.5	6.8	0.19	7.29	4.54	0.47	4.93	1.85	0.84	6.8	4.54	4.43				90.5
91.5		4.03	4.45		5.77	6.76						4.03	5.77	5.31	91.5
92.5	2.06	1.39	1.27	5.5	2.43	4.2	4.3	5.61	2.06	5.5	4.78	1.39	2.43	2.4	92.5
114.5	4.61	4.6	1.19	3.3	4.05	1.22	4.82	6.25	4.61	3.3	2.02				114.5
155.5	12.53	12.15	3.56	11.22	13.23	3.04	10.88	10.52	12.53	11.22	2.8	12.15	13.23	5.12	155.5
274.5	3.04	2.65	1.14	2.9	1.66	2.28						2.65	1.66	2.67	274.5
305.5	2.58	2.69	0.72	4.51	3.72	1.31	4.12	5.09	2.58	4.51	2.69				305.5
308.5	8.93	10.63	3.35	7.23	9.54	2.98	5.71	7.01	8.93	7.23	4.27	10.63	9.54	3.19	308.5
348.5	3.37	0.1	3.64	1.21	0.03	1.44			3.37	1.21	3.55				348.5
349.5	0.71	3.33	3.38	0.82	2.8	2.44	2.18	2.23							349.5
467.5	4.74	4.34	2.24	4.78	6.01	3.79	4.79	4.87	4.74	4.78	3.96	4.34	6.01	4.11	467.5
494.5	1.2	2.4	2.19	2.21	1.9	2.89	3.39	2.94	1.2	2.21	2.37	2.4	1.9	3.81	494.5
612.5	1.41	2.69	1.96	2.39	3.03	1.58	2.74	2.95	1.41	2.39	2.25	2.69	3.03	2.07	612.5
638.5		1.22	1.35	1.85	0.96	2.56				1.85	2.56	1.22	0.96	2.29	638.5
639.5	3.68	3.24	2.67	2.64	3.71	3.5	7.24	8.57	3.68	2.64	2.78	3.24	3.71	4.63	639.5
681.5		3.06	3.37	0.4	2.56	2.9	1.38	0.73				3.06	2.56	3.07	681.5
682.5	2.33		2.56	2.7		3.23	2.84	1.42	2.33	2.7	2.55				682.5

The bulk sediment in NMDS ordination (green points in Fig. 3.8) was the same bulk sediment used for the U-shaped burrow. With a 2 mm sliced resolution, the depth dependent subsamples representing points of the populated and not populated area were close to the surface layers representing points of the bulk sediment. The cluster analysis also indicated the same presentation: the surface bacterial communities from both areas with a 2 mm sliced resolution are clustered together with the surface bacterial communities in bulk sediments with a 1 cm sliced resolution (Fig. 3.9). This may indicate a high bacterial community relationship between two different sliced resolutions (2 mm and 1 cm sliced layer). The highest resolution (2 mm sliced layer) tracked the surface bacterial shifting smoothly and made a

connection between bacterial changes detected by the 1 cm layer resolution. Several undetected 5´-TRFs and 3´-TRFs in samples with a 1 cm sliced resolution were found contributing to the surface bacterial community in the populated sediments with a 2 mm sliced resolution; e.g. 5´-TRF of 191.5 bp after *Alu*I digestion (Table 3.4), as the bulk sediment was also from the populated area. Vice versa, the unique upper layers 5´-TRF of 192.5 bp could be detected (abundance in relative area) in 1 cm sliced layer, but was less intense in 2 mm sliced layer.
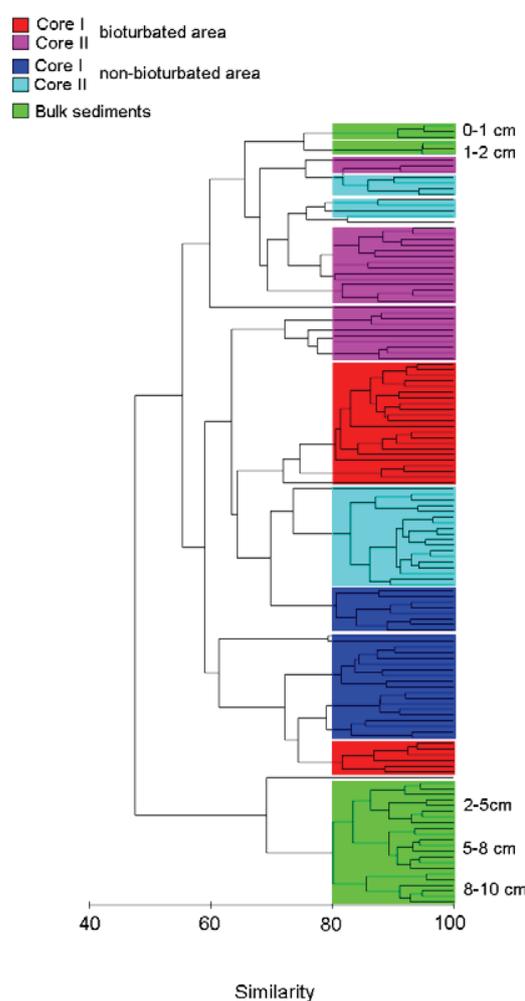
Fig. 3.9. The surface sediment layer. Bray-Curtis similarity based similarity clusters of 5´-TRF datasets after *Alu*I digestion represents a clustering of surface bacterial communities in the populated and not populated area.

After *Alu*I digestion, the bacterial communities between duplicates sample cores from the same area were likely different, as the represented points were part away and performing two obvious different clusters of core I and core II (Fig. 3.8), even though the TRF composition between the duplicate cores (Table 3.4) showed that they were mostly the same. The difference was due to different abundances, since the Bray-Curtis coefficient calculates the similarity/dissimilarity percentage between sample pairs based on the reproducible TRFs and their abundance. The NMDS ordination showed that the difference was decreasing slowly after *Hha*I and *Msp*I digestion; after *Hha*I digestion the points were getting closer and after *Msp*I digestion they were overlapping. The SIMPER analysis indicated contrary. Table 3.3 showed that the duplicate sample cores after *Alu*I digestion had the lowest dissimilarity percentage than those after *Hha*I and *Msp*I digestion. The duplicate sample core I and core II from the not populated area had dissimilarity percentage 44%, 47% and 50% respectively after *Alu*I*, Hha*I and *Msp*I digestion, as well as 42%, 44% and 50% for duplicate sample cores from the populated area. The dissimilarity after *Alu*I digestion was mostly contributed by 5′-TRFs of 194.5, 200.5, 201.5 bp and the 3′-TRFs of 59.5, 58.5, 90.5, 91.5 bp, because their abundances were highly different between the duplicate sample cores (Table 3.3). This contrary may relate to the difference of quality datasets generated from three different enzyme digestions. Enzyme digestion *Alu*I yielded better dataset quality than enzyme digestion *Hha*I and *Msp*I. Thus NMDS ordination generated from dataset after *Hha*I and *Msp*I digestion was more scattered and overlapping. A non-optimal ordination is possible due to NMDS uses unknown transformation (Shepard, 1974 *in* Kenkel and Orloci, 1986).

The dissimilarity percentage between the duplicate sample cores was mainly higher than it between the areas. Because TRFs that gave the highest contribution to the dissimilarity between duplicate cores (I and II) from the same area contributed less to the dissimilarity between the core I and core I or between core II and core II from different areas. For example 5′-TRF of 194.5 bp contributed 10% to the dissimilarity between duplicate sample cores from the not populated area (Table 3. 5), but only contributed about 4% to the dissimilarity between core I and core I or 7% between core II and core II from different areas. The difference between replicate cores from the same area was correlated with the differences between the bacterial populations in not populated and populated areas.

3.5. Discussion

3.5.1. The U-shaped burrow of *A. marina*

The geochemical stratification in the marine sediments develops zonations. This principally relates to oxygen penetration and the sequence of available electron acceptors over depth and determines the kind and intensity of occurring carbon, nitrogen and sulfur cycles at the particular depth. The zonation directly effects on the relative contribution of aerobic and anaerobic bacteria as the key player in those cycles (Jorgensen, 2000; Kristensen, 2001; Nybakken, 1997). Our results indicated that the zonation in the middle part of the U-shaped burrow had a gradient and was not physically and directly influenced by *A. marina*. An aerobic layer, a redox potential discontinuity (RDP) layer and an anoxic layer (Jorgensen, 2000; Kristensen, 2001; Nybakken, 1997) are expected to still exist through depth, whereby the decreasing TRFs represent surface bacteria and the increasing TRFs represent subsurface bacteria. The change of TRFs abundance may imply a vertical change of environmental conditions. The T-RFLP data suggest that the RDP layer is at 3-5 cm sediment depth. The unique TRFs of the surface layer and deep layers were not found in this depth. Below 3 cm depth, the abundance of surface and subsurface TRFs decreased or increased fast.

As *A. marina* interrupts this stratification and the associated biogeochemical process (Mermillod-Blondin and Rosenberg, 2006; Volkenborn *et al*., 2007 and 2007a; Nielsen *et al*., 2003; Kristensen, 2001), it significantly influences the microbial community composition (Goni-Urriza *et al.*, 1999; Retraubun *et al.*, 1996; Grossmann and Reichardt, 1991; Reichardt, 1988). In this study, the T-RFLP method enabled us to successfully detect a clear different bacterial community profile in the U-shaped burrow of *A. marina*. Three restriction enzymes, *Alu*I, *Hha*I and *Msp*I, generated a highly similar reflection of local bacterial communities: a surface sediment signature in the whole head shaft tube and a bacterial community shifting over depth in the tail shaft tube. In general, the richness (number of TRFs) was not different but the evenness (relative abundance) was decreasing or increasing with depth as obviously detected in the tail shaft tube. This T-RFLP profile may agree with Kristensen (2001); Banta *et al*. (1999) and Kristensen *et al*. (1985) who predicted that the same microbial community is probably present in the U-shaped burrow but with different

population sizes after he measured the nitrification and denitrification occurred in the burrow.

In the head shaft tube the bacterial community over depth was similar with those found at the surface layers in the tail shaft tube and the bulk sediment. Thus the brownish color of the sediment samples from the head shaft tube coincidences with the T-RFLP result representing the surface bacteria. A similar result was also reported by Retraubun *et al.* (1996) based on the identified bacterial taxa and bacterial counts: the same population density of aerobic bacteria were found in the feeding funnel and feeding pocket of the U-shaped burrow. This T-RFLP result may support the fragility and oxicity of the head shaft tube as reviewed by Riisgard and Banta (1998). This condition is due to several causes. *A. marine* continuously ingests the sediments in front of its head and makes a depression in the surface as a feeding funnel. From this oxidized surface sediments sink easily down to the feeding pocket passing the head shaft tube (Riisgard and Banta, 1998). It is not sealed off by mucous polysaccharides (Riisgard and Banta, 1998). The water pressure flows in a posterior-anterior direction and drops across the sediments in the head shaft and made this area labile; a normal pumping rate of 1.5 ml per minute could produce a head pressure about 20 cm H_2O (Riisgard and Banta, 1998). At low tide, the feeding funnel could be saturated by overlying water or dry. In dry conditions, the funnel and head shaft is exposed and open to air with the diameter of opening funnel about 37 - 80 mm and opening head shaft area about 12 - 25 mm depending on the lugworm size (Alyakrinskaya, 2003).

In the tail shaft tube a bacterial community shifting was detected with distinctive surface and subsurface layer bacterial communities. This T-RFLP result is probably in agreement with Fenchel (1996) who examined the role of the *Nereis diversicolor* (a subsurface deposit feeder) in aerobic microbial sediment processes by measuring the oxygen penetration and calculating the volume of the oxic sediment directly in the burrow. Due to a brief period of anoxia in the continuous irrigation, the apparent vertical zonation of microbial processes and of microbial community in the burrow over a centimeter scale was assumed as a reflection of diminishing oxic fraction rather than an ideal vertical redox sequence which is developed in the absence of the bioturbation (Fenchel, 1996). The aerobic and anaerobic bacteria may present at the same depth from the upper layer to 10 cm depth (Fenchel, 1996; Kristensen *et al.*, 1985) as the aerobic ciliate (*Kentrophoros fasciolata* and *Euplotes*

sp.) and the anaerobic ciliate (*Myelostoma bipartititum* and *Parablepharisma pellitum*) had been found at the same depth in the *A. diversicolor* burrow (Fenchel, 1996a). Jorgensen (1977) also concluded that a small transported sediment particle can be a unique aerobic or anaerobic microhabitat for microbes.

Comparing to the head shaft tube, physically the inside and outside of the tail shaft area is lined and rigid due to stabilizing of a lining matrix consisted of mucous polysaccharides secreted by lugworm and refractory detritus (Aller and Yingst, 1978; Papaspyrou, 2005; Kogure and Wada, 2005). The lining thickness was reported to be 2-8 μm and represents a dominant control on oxygen diffusion across the burrow-water boundary (Zorn *et al.*, 2006). It could reduce the solutes transportation between the sediment and burrow due to its low diffusivity (Aller and Yingst, 1978). Therefore, even if *A. marina* constantly pump oxygenated overlying water into its burrow in a posterior-anterior direction (Kristensen, 2001; Riisgard and Banta, 1998), oxygen only can penetrate to a distance of 0.95 mm in the tail shaft area (Zorn *et al.*, 2006). For comparison, in the populated area about 10 cm away from the feeding funnel or the cast deposit, the oxygen penetration could reach sediment 3.5 cm depth due to the higher permeability of surface sandy sediment (Volkenborn *et al.*, 2007a).

The gallery tube represented the lowest richness but with the highest evenness inhabited by surface and subsurface bacteria as found in the head and tail shaft tube. This may indicate that the gallery tube is a transition microhabitat as it is situated between the tail and head shaft tube. Consequently the physical and chemical sediments properties in this area may change fast in relation to the lugworm activities as the fact that the lugworm stays relative permanently between the interval times of the feeding-defecating cycle which is about 15-40 minutes or even more up to 5 days if the lugworm stays without feeding (Kristensen, 2001; Riisgard and Banta, 1998). Thus not all bacteria which were able growing in the head and tail shaft tube could grow in the gallery tube, unless only the highly adapted bacteria as represented by low number of observed TRFs. Even we did not find unique TRFs, but this area might be an unique niche for the adapted bacteria and well enriched by the lugworm which referred to the gardening phenomenon (Hylleberg, 1975) as indicated by the highest abundance of the observed TRFs that referred to the surface bacteria. The gardening phenomenon means that the lugworm stimulate the microbial growth by supplying oxygenated overlying water consisting of suspended nutritional compounds and lugworm's secretions that needed for the growth of

aerobic bacteria over the U-shaped burrow (Kogure and Wada. 2005; Riisgard and Banta, 1998; Retraubun *et al.*, 1996; Grossman and Reichardt, 1991; Hylleberg, 1975). The mucous polysaccharides lining the tail shaft and retaining its rigidity also provides a potential degradable substrate for bacterial growth (Papaspyrou, 2005; Kogure and Wada, 2005; Aller and Yingst, 1978).

Other side, the lowest richness and the highest evenness in the gallery tube might also due to the lugworm ingestion (Retraubun *et al.*, 1996; Plantae and Mayer, 1994; Grossmann and Reichardt, 1991; Reichardt, 1988) and to the particle size change from muddy to sandy (Volkenborn *et al.*, 2007), whereas if the grain size increase then the bacterial number also decrease (DeFlaun and Mayer, 1983). The feeding pocket is located in the gallery tube adjacent under the head shaft tube. This area is sandy since the lugworm ingests only small with <80 µm diameter subducted sediments and discards the bigger sediments accumulating below the feeding pocket and gallery area (Hylleberg, 1975; Riisgard and Banta, 1998).

In general we summarized that the reported effect of *A. marina* on the bacterial community whether directly (Retraubun *et al.*, 1996; Plantae and Mayer, 1994; Grossmann and Reichardt, 1991; Reichardt, 1988) or indirectly (Volkenborn et al., 2007, 2007a; Kristensen, 2001) could be also represented by the T-RFLP profile as the U-shaped burrow had a lower richness and a higher eveness than the bulk sediment. Following the Table 3.4 the missing TRFs were the subsurface bacteria (5´-TRFs of 153.5, 199.5, 253.5 and 440.5 bp). These bacteria could be obligate anaerobes that have no specific mechanism to survive, e.g. biofilms, aggregate or consortia formation and specific enzyme possession for neutralizing the effect of oxygen (Jorgensen, 1977; Van Niel and Gottschal, 1998; Voordouw and Voordouw, 1998) in the mosaic layer of oxic and anoxic microhabitat created by the lugworm (Fenchel, 2003; Kristensen 2001). And several TRFs which referred to the surface bacteria, e.g. 194.5 203.5, 245.5, 248.5 and 821.5 bp of 5´-TRFs, were more abundance in the U-shaped burrow probably indicating a reflection for a unique niche for them related to the gardening phenomenon, as extremely found in the gallery tube. Retraubun *et al.* (1996) further more added that the bacterial gardening in the head shaft provided a direct and indirect food for the lugworm whereby the increase of bacterial population might attract the meiofauna such as nematodes, flagellates and ciliates which were then also consumed by the lugworm.

3.5.2. The surface layers in the populated and not populated area

By applying a duplicate sample cores with a 2 mm resolution and three different enzyme digestions, a gradual bacterial community shifting at the surface layer was also obviously detected by the method in both populated and not populated areas. This result indicated a gradient of biological variation with depth, but no clear detectable difference between areas. This showed that either *A. marina* has no influence on the surface microbial community or that other biologically or physically causes may influence the same effect on the surface bacterial community. The intertidal sediment surface is physically unstable due to e.g. waves and periodically tidal currents but chemically stable due to a continuous oxic condition in overlying water (Kristensen, 2001). In the absence of *A.marina*, other bioturbators from polychaetes were reported significantly inhabiting the not populated area; e.g. *Nereis diversicolor, Pygospio elegans, Polydora cornuta, Tubificoides benedii, Capitella capitata* and *Scoloplos cf. armiger* which are classified into surface-subsurface deposit feeding worm and make burrow into the sediments (Volkenborn and Reise, 2007).

A. marine may only significantly influence the deeper bacterial community, especially direct in the U-shaped burrow tube as previously discussed. No detectable difference between both areas at the surface layer were also reported by Volkenborn *et al.* (2007 and 2007a) and Goni-Urriza *et al.* (1999). At 0 to 3 cm depth, the ammonium, nitrate and nitrite concentration was similar between both area respectively <50 µM, <2.5 µM and <0.25 µM (Volkenborn *et al.*, 2007a). This may indicate that biological processes relating to the nitrification were in the same rate at the sediment surface. Goni-Urriza *et al.* (1999) reported a similar indication; a similar number of the colorless sulphur bacteria (3.7 to 5.4 X 10^5 cells cm^{-3}) and the anoxygenic phototrophic bacteria were counted from the surface layer sediments of the populated and not-populated area by *A. marina* and *C. edule*. But after going deeper to depth, the effect of the *A. marina* on the physical and chemical sediments properties (Volkenborn et al., 2007 and 2007) and on the bacterial counting number (Goni-Urriza *et al.*, 1999) was significant. Furthermore Volkenborn *et al.* (2007 and 2007a) presented data that the effect of *A. marina* was not restricted only to the burrow but to the entire area at which *A. marina* influenced the physical and

chemical sediment properties and supported the intertidal habitat succession from muddy sediment to sandy sediment.

As we applied the method on the duplicate sample cores taken within a distance in the same sampling area, we found spatial and random effect in our T-RFLP results: bacterial community from the duplicate sample cores looked different with approximately dissimilarity percentage 47% and 45% respectively for the not populated and populated area. Blackwood *et al.* (2007) and Ranjard *et al.* (2003) mentioned that the relationship between diversity and sampling effort may or may not hold for T-RFLP profiles due to the dependence on the spatial scale of sampling. Our result from the surface samples may also agree to Osborne *et al.* (2006) and Dunbar *et al.* (2000) as we got similar to dissimilar microbial community pattern from different enzyme digestions applied to the same sample. Therefore using a combination of single enzyme digestions was suggested in a profiling study of bacterial community (Osborne *et al.*, 2006; Dunbar *et al.*, 2000).

3.6. References

1. Aller, R.C. and J.Y. Yingst. 1978. Biogeochemistry of tube-dwellings: a study of the sedentary polychaete *Amphitrite ornate* (Leidy). J. Mar. Res. 36:201-254.

2. Alyakrinskaya, I.O. 2003. Some ecological features of the lugworm Arenicola marina L. (Annelida, Polychaeta) and its morphological and biochemical adaptations to burrowing. Biology Bulletin. 30:411-418.

3. Amann, R.I., W. Ludwig and K-H. Scheifer, 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. 59:143-169.

4. Banta, G.T., M.Holmer, M. H. Jensen and E. Kristensen. 1999. Effect of two polychaete worms, *Nereis diversicolor* and *Arenicola marina*. On aerobic and anaerobic decomposition in a sandy marine sediment. Aquatic Microbial Ecol. 19:189-204.

5. Blackwood, C.B., D. Hudleston, D.R. Zak and J.S. Buyer. 2007. Interpreting ecological diversity indices applied to terminal restriction fragment length polymorphism data: insights from simulated microbial communities. Appl. Environ. Microbiol. 73:5276-5283.

6. Clarke, K.R. and R.N. Gorley. 2001. PRIMER v5: User manual/tutorial, PRIMER-E, Plymouth UK.

7. Clarke, K.R. and R.M. Warwick. 2001. Change in Marine Communities: An Approach to Statistical Analysis and Interpretation. 2nd edition: PRIMER-E, Plymouth, UK.

8. DeFlaun, M.F. and L.M. Mayer, 1983. Relationship between bacteria and grain surfaces in the intertidal sediments. Limnol. Oceanogr. 28:873-881.

9. Dunbar, J., L.O. Ticknor and C.R. Kuske. 2001. Phylogenetic specificity and reproducibility and new method for analysing of terminal restriction fragment profile of 16S rRNA genes from bacterial communities. Appl. Environ. Microbiol. 67:190-197.

10. Dunbar, J., L.O. Ticknor and C.R. Kuske. 2000. Assessment of microbial diversity on two south western U.S. soils by terminal restriction fragment analysis. Appl. Environ. Microbiol. 66:2943-2950.

11. Engebretson, J.J and C. L. Moyer. 2003. Fidelity of select restriction endonucleases in determining microbial diversity by terminal restriction fragment length polymorphisms. Appl. Environ. Microbiol. 69:4823-4829.

12. Fenchel, T. 1996. Worm burrow and oxic microniches in marine sediments. 1. Spatial and temporal scales. Mar. Biol. 127: 289-295.

13. Fenchel, T. 1996a. Worm burrow and oxic microniches in marine sediments. 2. Distribution pattern of ciliated protozoa. Mar. Biol. 127:297-301.

14. Frey, J.C., E.R. Angert and A.N. Pell. 2006. Assesment of biases associated with profiling simple, model communities using terminal-restriction fragment length polymorphism-based analysis. J. Microbiol. Met. 67:9-19.

15. Goni-Urriza, M., X. de Montaudouin, R. Guuyoneaud, G. Bachelet and R. de Wit. 1999. Effect of macrofaunal bioturbation on bacterial distribution in marine sandy sediments, with special reference to sulphur-oxidising bacteria. J. Sea Res. 41:269-279.

16. Grosmann, S. and W. Reichardt. 1991. Impact of *Arenicola marina* on bacteria in the intertidal sediments. Mar. Ecol. Prog. Ser. 77:85-93.

17. Hylleberg, J. 1975. Selective feeding by *Abarenicola pacifica* with notes on *Abarenicola vagabunda* and a concept of gardening in lugworm. Ophelia. 14:113-137.

18. Hüttel, M. 1990. Influence of the lugworm Arenicola marina on porewater nutrient profiles of sand flat sediments. Mar. Ecol. Prog. Ser. 62:241-248.

19. Jorgensen, B.B. 2000. Bacteria and marine biogeochemistry, p:173-203. *In* Schulz, H.D. and M. Zabel (ed.), Marine geochemistry. Springer.

20. Jorgensen, B.B. 1977. Bacterial sulphate reduction within reduced microniches of oxidized marine sediments. Mar. Biology. 41:7-17.

21. Kogure, K and M. Wada. 2005. Minireview : Impact of macrobenthic bioturbation in marine sediments on bacterial metabolic activity. Microbes and Environments. 20:191-199.

22. Kenkel, N. C. and L. Orloci. 1986. Applying metric and nonmetric multidimensional scaling to ecological studies: some new results. Ecology. 67:919-928.

23. Kristensen, E. 2001. Impact of polychaetes (*Nereis* spp and *Arenicola marina*) on carbon biogeochemistry in coastal marine sediments. Geochem.Trans. 2:92-103

24. Kristensen, E., M.H. Jensen and T.K. Andersen. 1985. The impact of polychaete (*Nereis virens* Sars) burrows on nitrification and nitrate reduction in estuarine sediments. J. Exp. Mar. Biol. Ecol. 85:75-91.

25. Liu, W., T.L. Marsh, Cheng H. and L.J. Forney. 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl. Environ. Microbiol. 63:4516-4522.

26. Mermillod-Blondin, F. and R. Rosenberg. 2006. Ecosystem engineering : the impact of bioturbation of biogeochemical process in marine and freshwater benthic habitats. An overview article. Aquat. Sci. 68:434-442.

27. Muyzer, G., A. Teske and C.O. Wirsen. 1995. Phylogenic relation of Thiomicrospira sp. And their identification in deep sea hydrothermal vent samples by DGGE of 16S rRNA fragments, Arch. Microbiol. 59:695-700.

28. Nielsen, O.I., E. Kristensen and M. Holmer. 2003. Impact of *Arenicola marina* (Polychaete) on the sediment sulfur dynamics. Aquatic Microbial Ecol. 33:95-105.

29. Nybakken, J. W. 1997. Marine Biology, an ecological approach. 4th Edition. Addition-Wesley Educational Publishers Inc. USA.

30. Osborne, C.A., G.N. Rees, Y. Bernstein and P.H. Janssen. 2006. New threshold and confidence estimates for terminal restriction fragment length polymorphism analysis of complex bacterial communities. Appl. Environ. Microbiol. 72:1270-1278.

31. Papaspyrou, S., T. Gregersen, E. Kristensen, B. Christensen and R.P. Cox. 2006. Microbial reaction rates and bacterial communities in sediment surrounding burrows of two nereidid polychaetes (*Nereis diversicolor* and *N. virens*). Marine Biology. 148:541-550.

32. Plante, C.J and L.M. Mayer. 1994. Distribution and efficiency of bacteriolysis in the gut of Arenicola marina and three additional deposit feeders. Mar. Ecol. Prog. Ser. 109:183-194.

33. Ranjard, L., D.P.H. Lejon, C. Mougel, L. Schehrer, D. Merdinoglu and R. Chaussod. 2003. Sampling strategy in molecular microbial ecology: influence of soil sample size on DNA fingerprinting analysis of fungal and bacterial communities. Environ. Microbiol. 5:1111-1120.

34. Reichardt, W. 1988. Impact of bioturbation by Arenicola marina on microbiological parameters in intertidal sediments. Mar. Ecol. Prog. Ser. 44:149-158.

35. Rees, G.N., D.S. Baldwin, G.O. Watson, S. Perryman and D.L. Nielsen. 2004. Ordination and significance testing of microbial community composition derived from terminal restriction fragment length polymorphisms: application of multivariate statistics. Antonie van Leeuwenhoek. 86:339-347.

36. Retraubun, A.S.W., M. Dawson and S.M. Evans. 1996. The role of the burrow funnel in feeding processes in the lugworm *Arenicola marina* (L). J. Exp.Mar. Biol. and Ecol. 107-118

37. Riisgard, H.U, I. Bernsten and B. Tarp. 1996. The lugworm (*Arenicola marina*) pump: characteristic, modelling and energy cost. Mar. Ecol. Prog. Ser. 138:149-156.

38. Riisgard, H.U and G.T. Banta. 1998. Irrigation and deposit feeding by the lugworm Arenicola marina, characteristics and secondary effects on the environment. A review of current knowledge. Vie Milieu. 48:243-257.

39. Saikaly, P.E., P.G. Stroot and D.B. Oether. 2005. Use of 16S rRNA gene terminal restriction fragment analysis to assess the impact of solids retention time on the bacterial diversity of activated sludge. Appl. Environ. Microbiol. 71:5814-5822.

40. Van Niel, E.W.J. and J.C. Gottschal. 1998. Oxygen consumption by *Desulfovibrio* strains with and without polyglucose. Appl. Environ. Mirobiol. 64:1034-1039.

41. Volkenborn, N., S.I.C. Hedtkamp, J.E.E. van Beusekom and K. Reise. 2007. Effects of bioturbation by lugworm (*Arenicola marina*) on physical and chemical sediment properties and implications for intertidal habitat succession. Estuatine, Coastal and Shelf Sci. 74:331-343.

42. Volkenborn, N., L. Polerecky, S.I.C. Hedtkamp, J.E.E. van Beusekom and D. de Beer. 2007a. Bioturbation and bioirrigation extend the open exchange regions in permeable sediments. Limnol. Oceanogr. 52:1898-1909.

43. Volkenborn, N. and K. Reise. 2007. Effects of Arenicola marina on polychaete functional diversity revealed by large-scale experimental lugworm exclusion. J. Sea Res. 57:78-88.

44. Voordouw, J.K. and G. Voordouw. 1998. Deletion of the *rho* gene increases the oxygen sensitivity of the sulfate reducing bacterium *Desulfovibrio vulgaris* Hildenborough. Environ. Mirobiol. 64:2882-2887.

45. Zorn, M.E., S.V. Lalonde, M.K. Gingras, S.G. Pemberton and K.O. Konhauser. 2006. Microscale oxygen distribution in various invertebrate burrow walls. Geobiology. 4:137-145.

Appendix 1:

Table 3.5. The ANOSIM results for the surface bacterial community profiles in the populated and not populated area.

5´-TRF datasets after Alul digestion

Pair wise sub sample		Bioturbated area I	II	Non-bioturbated area I	II
0 - 0.2 cm	0.2 - 0.4 cm	1	0.667	1	0.556
	0.4 - 0.6 cm	1	1	1	0.556
	0.6 - 0.8 cm	1	0.556	1	0.704
	0.8 - 1 cm	1	1	1	0.556
	1 - 1.2 cm	1	1	1	0.296
	1.2 - 1.4 cm	0.852	1	1	1
	1.4 - 1.6 cm	1	1	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	1	1	-
0.2 - 0.4 cm	0.4 - 0.6 cm	0.5	0.963	1	0.519
	0.6 - 0.8 cm	1	0.741	1	0.889
	0.8 - 1 cm	1	1	1	-0.111
	1 - 1.2 cm	1	1	1	0.185
	1.2 - 1.4 cm	0.333	1	1	1
	1.4 - 1.6 cm	1	1	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	1	1	-
0.4 - 0.6 cm	0.6 - 0.8 cm	0.75	0.556	1	0.926
	0.8 - 1 cm	0.833	1	1	0.222
	1 - 1.2 cm	0.917	1	1	-0.296
	1.2 - 1.4 cm	0.5	1	1	1
	1.4 - 1.6 cm	1	1	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	1	1	-
0.6 - 0.8 cm	0.8 - 1 cm	0.556	0.259	0.583	0.889
	1 - 1.2 cm	0.519	0.444	1	0.259
	1.2 - 1.4 cm	0.519	0.333	1	1
	1.4 - 1.6 cm	0.926	1	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	1	1	-
0.8 - 1 cm	1 - 1.2 cm	0.444	0.37	1	0.037
	1.2 - 1.4 cm	0.556	0.519	1	1
	1.4 - 1.6 cm	0.926	1	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	1	1	-
1 - 1.2 cm	1.2 - 1.4 cm	0.333	0.222	0.833	1
	1.4 - 1.6 cm	0.556	1	1	1
	1.6 - 1.8 cm	1	0.963	1	1
	1.8 - 2 cm	1	1	0.889	-
1.2 - 1.4 cm	1.4 - 1.6 cm	0.481	0.556	0.667	0.519
	1.6 - 1.8 cm	0.556	0.593	1	1
	1.8 - 2 cm	0.963	0.926	0.667	-
1.4 - 1.6 cm	1.6 - 1.8 cm	1	1	0.704	0.333
	1.8 - 2 cm	1	1	0.556	-
1.6 - 1.8 cm	1.8 - 2 cm	0.778	0.926	0.481	

5´-TRF datasets after Hhal digestion

Pair wise sub sample		Bioturbated area I	II	Non-bioturbated area I	II
0 - 0.2 cm	0.2 - 0.4 cm	1		0.741	0.148
	0.4 - 0.6 cm	1	0.222	1	0.222
	0.6 - 0.8 cm	1	0.333	1	0.333
	0.8 - 1 cm	1	0.778	1	1
	1 - 1.2 cm	1	1	1	1
	1.2 - 1.4 cm	1	1	1	1
	1.4 - 1.6 cm	1	-	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	-	1	1
0.2 - 0.4 cm	0.4 - 0.6 cm	0.852	-	0.815	0.667
	0.6 - 0.8 cm	1	-	1	1
	0.8 - 1 cm	1	-	1	1
	1 - 1.2 cm	1	-	1	1
	1.2 - 1.4 cm	1	-	1	1
	1.4 - 1.6 cm	1	-	1	1
	1.6 - 1.8 cm	1	-	1	1
	1.8 - 2 cm	1	-	1	1
0.4 - 0.6 cm	0.6 - 0.8 cm	1	0.037	1	0.556
	0.8 - 1 cm	0.917	0.481	1	0.889
	1 - 1.2 cm	1	1	1	0.852
	1.2 - 1.4 cm	1	0.926	1	1
	1.4 - 1.6 cm	1	-	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	-	1	1
0.6 - 0.8 cm	0.8 - 1 cm	0.75	0	-1	1
	1 - 1.2 cm	0.667	0.407	0.556	1
	1.2 - 1.4 cm	1	0.074	1	1
	1.4 - 1.6 cm	1	-	1	1
	1.6 - 1.8 cm	1	0.5	1	1
	1.8 - 2 cm	1	-	-	1
0.8 - 1 cm	1 - 1.2 cm	0	0.667	0.5	0.593
	1.2 - 1.4 cm	0.583	0.333	1	1
	1.4 - 1.6 cm	1	-	1	1
	1.6 - 1.8 cm	1	0.75	1	0.833
	1.8 - 2 cm	1	-	1	1
1 - 1.2 cm	1.2 - 1.4 cm	0.704	0.333	0.444	0.741
	1.4 - 1.6 cm	0.704	-	0.63	1
	1.6 - 1.8 cm	1	1	0.667	0.833
	1.8 - 2 cm	1	-	0.556	1
1.2 - 1.4 cm	1.4 - 1.6 cm	0.926	-	0.926	1
	1.6 - 1.8 cm	1	0.667	0.556	0.75
	1.8 - 2 cm	1	-	1	1
1.4 - 1.6 cm	1.6 - 1.8 cm	0.815	-	0.407	0.583
	1.8 - 2 cm	1	-	1	0.778
1.6 - 1.8 cm	1.8 - 2 cm	1	-	-0.111	0.667

5´-TRF datasets after Mspl digestion

Pair wise sub sample		Bioturbated area I	II	Non-bioturbated area I	II
0 - 0.2 cm	0.2 - 0.4 cm	0.259	0.593	1	0.444
	0.4 - 0.6 cm	0.407	1	0.583	0.167
	0.6 - 0.8 cm	0.481	1	0.5	
	0.8 - 1 cm	0.667	1	1	0.167
	1 - 1.2 cm	1	1	1	0.37
	1.2 - 1.4 cm	0.556	1	0.833	0.556
	1.4 - 1.6 cm	0.741	1	1	0.417
	1.6 - 1.8 cm	0.704	1	1	0.852
	1.8 - 2 cm	1		1	1
0.2 - 0.4 cm	0.4 - 0.6 cm	-0.111	0.815	0.333	0.583
	0.6 - 0.8 cm	-0.222	0.889	0.75	
	0.8 - 1 cm	0.111	0.917	1	1
	1 - 1.2 cm	0.519	1	1	0.778
	1.2 - 1.4 cm	0.074	1	0.704	1
	1.4 - 1.6 cm	0.407	1	0.926	1
	1.6 - 1.8 cm	0.407	1	1	1
	1.8 - 2 cm	0.852		1	1
0.4 - 0.6 cm	0.6 - 0.8 cm	-0.222	0.667	0.333	
	0.8 - 1 cm	0.148	0.917	0.417	1
	1 - 1.2 cm	0.407	0.852	0.704	0.667
	1.2 - 1.4 cm	0.037	1	0.481	1
	1.4 - 1.6 cm	0.259	1	0.556	1
	1.6 - 1.8 cm	0.37	1	1	1
	1.8 - 2 cm	0.481		1	1
0.6 - 0.8 cm	0.8 - 1 cm	-0.222	0.5	0.25	-
	1 - 1.2 cm	0.259	0.778	0.417	-
	1.2 - 1.4 cm	-0.074	0.778	0.25	-
	1.4 - 1.6 cm	0	1	0.333	-
	1.6 - 1.8 cm	0.148	1	1	-
	1.8 - 2 cm	0.593	0.917	1	-
0.8 - 1 cm	1 - 1.2 cm	0.37		0.667	-0.167
	1.2 - 1.4 cm	0	0.75	0.333	1
	1.4 - 1.6 cm	0.111	1	0.333	1
	1.6 - 1.8 cm	0.185	1	1	0.333
	1.8 - 2 cm	0.704		1	1
1 - 1.2 cm	1.2 - 1.4 cm	0.259	0.259	-0.074	0.37
	1.4 - 1.6 cm	0.593	1	0.111	0.583
	1.6 - 1.8 cm	0.481	0.778	0.333	0.407
	1.8 - 2 cm	0.815		0.917	1
1.2 - 1.4 cm	1.4 - 1.6 cm	0.074	0.75	0.074	0.917
	1.6 - 1.8 cm	0.148	0.667	-0.25	0.333
	1.8 - 2 cm	0.148		-0.083	1
1.4 - 1.6 cm	1.6 - 1.8 cm	-0.074	0.917	-0.167	-0.083
	1.8 - 2 cm	0.296		-0.25	1
1.6 - 1.8 cm	1.8 - 2 cm	0.074		-0.5	0.481

3´-TRF datasets after Alul digestion

Pair wise sub sample		Bioturbated area I	II	Non-bioturbated area I	II
0 - 0.2 cm	0.2 - 0.4 cm	-0.25	0.296	0.148	0.296
	0.4 - 0.6 cm	0.5	0.259	0.259	0.481
	0.6 - 0.8 cm	0.917	0.444	0.667	0.519
	0.8 - 1 cm	0.833	1	0.852	0.333
	1 - 1.2 cm	1	1	1	0.222
	1.2 - 1.4 cm	0.917	1	1	0.917
	1.4 - 1.6 cm	1	1	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	1	1	-
0.2 - 0.4 cm	0.4 - 0.6 cm	0	0.519	0.259	0.444
	0.6 - 0.8 cm	0.583	0.444	0.833	1
	0.8 - 1 cm	0.75	1	0.889	-0.111
	1 - 1.2 cm	0.25	1	1	0.111
	1.2 - 1.4 cm	0.667	1	1	1
	1.4 - 1.6 cm	0.833	1	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	1	1	-
0.4 - 0.6 cm	0.6 - 0.8 cm	1	0.185	0.5	1
	0.8 - 1 cm	0.333	0.63	0.519	-0.407
	1 - 1.2 cm	0.5	0.778	0.815	0.111
	1.2 - 1.4 cm	0.917	0.833	0.917	1
	1.4 - 1.6 cm	1	1	1	0.926
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	1	1	-
0.6 - 0.8 cm	0.8 - 1 cm	0.889	0	0.417	1
	1 - 1.2 cm	1	0.185	0.667	0.481
	1.2 - 1.4 cm	1	0.083	1	1
	1.4 - 1.6 cm	1	0.417	1	0.815
	1.6 - 1.8 cm	1	0.926	1	1
	1.8 - 2 cm	1	0.815	1	-
0.8 - 1 cm	1 - 1.2 cm	0.5	0.519	0.889	0
	1.2 - 1.4 cm	0.667	0.75	1	1
	1.4 - 1.6 cm	0.593	0.833	1	1
	1.6 - 1.8 cm	0.815	1	1	1
	1.8 - 2 cm	0.833	1	1	-
1 - 1.2 cm	1.2 - 1.4 cm	0.167	-0.083	-0.25	1
	1.4 - 1.6 cm	0.167	0.926	1	0.963
	1.6 - 1.8 cm	0.833	0.889	0.778	1
	1.8 - 2 cm	1	1	1	-
1.2 - 1.4 cm	1.4 - 1.6 cm	0.259	1	0.75	0.5
	1.6 - 1.8 cm	0.815	0.583	0.25	1
	1.8 - 2 cm	1	1	1	-
1.4 - 1.6 cm	1.6 - 1.8 cm	0.481	1	0.222	0.083
	1.8 - 2 cm	1	0.417	0.333	-
1.6 - 1.8 cm	1.8 - 2 cm	0.917	1	0.111	-

3´-TRF datasets after Hhal digestion

Pair wise sub sample		Bioturbated area I	II	Non-bioturbated area I	II
0 - 0.2 cm	0.2 - 0.4 cm	0.926	-	0.407	0.074
	0.4 - 0.6 cm	0.704	0.296	0.778	0.333
	0.6 - 0.8 cm	1	0.63	1	0.778
	0.8 - 1 cm	1	1	1	1
	1 - 1.2 cm	1	1	1	1
	1.2 - 1.4 cm	0.926	1	1	1
	1.4 - 1.6 cm	1	-	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	-	1	1
0.2 - 0.4 cm	0.4 - 0.6 cm	0.333	-	0.407	0.333
	0.6 - 0.8 cm	0.667	-	1	-0.333
	0.8 - 1 cm	0.917	-	1	1
	1 - 1.2 cm	1	-	1	1
	1.2 - 1.4 cm	0.963	-	1	1
	1.4 - 1.6 cm	0.963	-	1	1
	1.6 - 1.8 cm	1	-	1	1
	1.8 - 2 cm	1	-	1	1
0.4 - 0.6 cm	0.6 - 0.8 cm	0	0	0.778	0.556
	0.8 - 1 cm	0	0.444	0.75	1
	1 - 1.2 cm	0.333	0.852	1	0.963
	1.2 - 1.4 cm	0.778	0.815	1	1
	1.4 - 1.6 cm	0.852	-	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	-	1	1
0.6 - 0.8 cm	0.8 - 1 cm	0.5	0.481	0	0.556
	1 - 1.2 cm	0.667	0.815	0.778	1
	1.2 - 1.4 cm	0.5	0.63	1	1
	1.4 - 1.6 cm	0.667	-	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	-	-	1
0.8 - 1 cm	1 - 1.2 cm	0.25	0.963	0.833	0.556
	1.2 - 1.4 cm	0.5	0.556	0.667	1
	1.4 - 1.6 cm	0.833	-	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	-	1	1
1 - 1.2 cm	1.2 - 1.4 cm	0.444	0.296	0.333	1
	1.4 - 1.6 cm	0.556	-	0.778	1
	1.6 - 1.8 cm	1	0.917	0.63	1
	1.8 - 2 cm	1	-	0.556	1
1.2 - 1.4 cm	1.4 - 1.6 cm	0.37	0.583	0.148	1
	1.6 - 1.8 cm	1	-	0.407	1
	1.8 - 2 cm	1	-	0.556	1
1.4 - 1.6 cm	1.6 - 1.8 cm	0.593	-	0.407	1
	1.8 - 2 cm	0.5	-	0.556	1
1.6 - 1.8 cm	1.8 - 2 cm	0.917	-	0.111	1

3´-TRF datasets after Mspl digestion

Pair wise sub sample		Bioturbated area I	II	Non-bioturbated area I	II
0 - 0.2 cm	0.2 - 0.4 cm	1	0.444	0.583	1
	0.4 - 0.6 cm	0.75	0.778	0.667	0.5
	0.6 - 0.8 cm	-	0.889	1	-
	0.8 - 1 cm	1	1	1	0.75
	1 - 1.2 cm	0.667	1	1	0.833
	1.2 - 1.4 cm	1	1	1	1
	1.4 - 1.6 cm	1	1	1	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	0.833	1	-	-
0.2 - 0.4 cm	0.4 - 0.6 cm	0.5	0.593	0	0
	0.6 - 0.8 cm	-	0.519	0.111	-
	0.8 - 1 cm	0	0.25	0.667	0.167
	1 - 1.2 cm	0.333	0.556	1	0.519
	1.2 - 1.4 cm	1	1	1	1
	1.4 - 1.6 cm	1	1	0.917	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	1	1	-	-
0.4 - 0.6 cm	0.6 - 0.8 cm	-	0.296	-0.333	-
	0.8 - 1 cm	0	0.75	0.25	-0.5
	1 - 1.2 cm	-0.083	1	0.75	0.417
	1.2 - 1.4 cm	1	1	0.667	1
	1.4 - 1.6 cm	1	1	0.333	1
	1.6 - 1.8 cm	1	1	1	1
	1.8 - 2 cm	0.75	1	-	-
0.6 - 0.8 cm	0.8 - 1 cm	-	0.167	0	-
	1 - 1.2 cm	-	0.333	0	-
	1.2 - 1.4 cm	-	0.333	1	-
	1.4 - 1.6 cm	-	1	1	-
	1.6 - 1.8 cm	-	1	0.25	-
	1.8 - 2 cm	-	1	-	-
0.8 - 1 cm	1 - 1.2 cm	-0.778	1	1	-0.5
	1.2 - 1.4 cm	-1	1	1	0.583
	1.4 - 1.6 cm	-	1	1	1
	1.6 - 1.8 cm	-	1	0	1
	1.8 - 2 cm	0	1	-	-
1 - 1.2 cm	1.2 - 1.4 cm	-0.417	-	0.5	0
	1.4 - 1.6 cm	0.556	1	1	0.167
	1.6 - 1.8 cm	0.111	1	0.25	0.333
	1.8 - 2 cm	0.333	-	-	-
1.2 - 1.4 cm	1.4 - 1.6 cm	1	0	0	0.333
	1.6 - 1.8 cm	-	0.111	1	0.583
	1.8 - 2 cm	0.25	-	-	-
1.4 - 1.6 cm	1.6 - 1.8 cm	-	0.167	-	0
	1.8 - 2 cm	0	1	-	-
1.6 - 1.8 cm	1.8 - 2 cm	-1	0.111	-	-

Acknowledgements

Last but not least, I dedicated my lovely huge to my husband, my little girl who was born here during my study time, my parents and my sister.