



BACHELOR THESIS  
SIK- THE SWEDISH INSTITUTE FOR FOOD AND BIOTECHNOLOGY

# Microbial Diversity in Raw and Pasteurized Milk with Terminal Restriction Fragment Length Polymorphism (T-RFLP)

SOFIA SVENNINGSSON

# Microbial Diversity in Raw and Pasteurized Milk with Terminal Restriction Fragment Length Polymorphism

Sofia Svenningsson

*Faculty of Natural Science  
Department of Biology  
Lund University*

Supervisor  
Klara Båth

Examiner  
Nora Ausmees

Bachelor Thesis, 30hp  
SIK- The Swedish Institute for Food and Biotechnology  
Gothenburg Nov 2011/ March 2012



**LUND**  
UNIVERSITY

## Microbial Diversity in Raw and Pasteurized Milk with Terminal Restriction Fragment Length Polymorphism

### Abstract

In this thesis, a molecular PCR-based method was used to study the bacterial diversity in milk. The aim was to compare the microbiota of traditionally pasteurized milk with milk treated with a novel pasteurization technique, using terminal restriction fragment length polymorphism (T-RFLP). A second aim was to analyze the microbial composition of cheese produced from the two milk variants. Results of this molecular approach were compared with outcomes from traditional culturing on non-selective media, followed by 16S rRNA sequencing in order to evaluate the usefulness of the methods. Overall, the results demonstrated that T-RFLP is a powerful tool for analyzing microbiota in foods. In conclusion, both pasteurization techniques proved to be effective in reducing the number of bacteria. The initial hypothesis, that the two pasteurization techniques affect different parts of the microbiota of the raw milk, was confirmed. As expected, the molecular approach of DNA extraction direct from the milk detected a more differentiated microflora, compared to DNA extracted from cultivated bacteria from the milk. The results also indicated that the molecular approach was more reproducible between the sampling occasions. The cultivation and sequencing showed that the microbiota mainly consisted of *Firmicutes* and *Actinobacteria*, as well as *Bacteroidetes*.

# Contents

1. Introduction.....	5
1.1. Molecular approaches to describe the microflora.....	5
1.1.1 T-RFLP analysis .....	6
1.1.2 Clone library and identification.....	8
1.1.3 Method considerations.....	9
1.2 Microbial community in foods.....	10
1.2.1 Bacterial flora in milk and cheese .....	10
1.2.2 Pasteurization.....	10
2. Material and methods.....	12
2.1. Food sampling and DNA preparation .....	12
2.2 Determination of CFU.....	12
2.3. T-RFLP analysis.....	12
2.4. Clone library construction .....	12
2.5. Assembly of data for identification.....	13
3. Results .....	14
3.1 Determination of CFU.....	14
3.2 T-RFLP analysis.....	14
3.2.1 Bacterial community profiles .....	14
3.2.2. Method of detection.....	18
3.2.3 Clone library identification .....	18
4. Discussion .....	20
6. Acknowledgment.....	23
7. References.....	24
8. Appendix A- Complementary studies.....	26
8.1. Material and methods.....	26
Food sampling and DNA preparation .....	26
Clone library construct .....	26
8.2. Results.....	27

## **1. Introduction**

The number of species within a community (species richness) and the size of the species population (species evenness) are two fundamental parameters for assessing the microbial community diversity. Traditional methods are based on visual morphology (limited by the small number of morphologies), biochemical tests (limited to cultured species), while newer molecular genetics methods are based on differences in the DNA sequence. Most of the traditional methods include culturing on plates, which have proven to give a deceptive reflection of the original community structure (27). Previous studies have demonstrated that large fractions of the organisms are difficult to cultivate (36 and 38). It is e.g. estimated that only 0.1-10% of the bacteria from soil can be cultivated (23). Furthermore, culture-based identification is labor-intensive, and the identification of ten colonies will not be representative for the entire community but rather only the dominating flora. Thus, standard cultivating methods may not be a good enough tool for food manufactures to achieve the desired risk management.

### **1.1. Molecular approaches to describe the microflora**

Molecular biology methods based on variations in the DNA sequences allow for a high-throughput and reproducible characterization of microbial communities. The most commonly analyzed prokaryote gene is the gene encoding for 16S ribosomal RNA sequence, which is part of the small subunit of the ribosome. This universal gene has sufficiently conserved sequences and at the same time enough sequence variability to differentiate between prokaryotes. The 16S rRNA gene has been widely used as a powerful tool for microbial fingerprinting. These include denaturing gradient gel electrophoresis (20), ribosomal intergenic spacer analysis (4), single-strand conformation polymorphism (32), and terminal restriction fragment length polymorphism (T-RFLP) (Table 1.). These approaches do not require cultivation and therefore visualize a wider range of organisms compared to biochemical culture-based methods. T-RFLPs application in assessing microbial diversity and community structure was first described by Liu (16) and has since then been widely applied in the analysis of bacterial community characterization in various environments, human colonic microbiota (11,19), oral microbial profiles (29) as well as soil communities (33). In addition to this, T-RFLP has also been used to analyze lactic acid bacteria (21), fungal genes (35), nitrifying bacteria (39) as well as methanotrophs (17). It has proven to be a sensitive and reproducible method to isolate and amplify total community DNA (1).

Although T-RFLP is a widespread analysis tool for a quick assessment of the microbiota, the publications on the method employed to analysis of foods are still relatively few. In this study, the utility of T-RFLP as a method for analyzing bacterial genetic diversity in milk was evaluated. Samples were collected from three different sources; raw milk, traditional pasteurized milk and milk treated with a novel pasteurization technique. The novel technique

TABLE. 1. Summary of some main genomic-based methods involved in describing the microflora of dairy products

<b>Method</b>	<b>Principle</b>
<i>Culture dependent methods</i>	
Random Amplified Polymorphic DNA (RAPD)	Arbitrary short primers randomly amplify a large genomic template by low stringent hybridization. Separated amplicons results in a fingerprint pattern.
Restriction Fragment Length Polymorphism (RFLP), also named Amplified Ribosomal DNA Restriction Analysis (ARDA)	Specific restriction endonucleases digest DNA. A labeled RFLP probe hybridizes with the digested DNA fragments, giving a unique semi-quantative blotting pattern.
<i>Culture independent methods</i>	
Denaturing or Temporal Temperature Gradient Gel Electrophoresis (DGGE or TGGE)	Oligonucleotide primers PCR amplifies equally long fragments, which are separated electrophoretically based on their GC-content. DGGE utilizes a chemical gradient and TGGE a temperature gradient.
Single Stranded Conformation Polymorphism (SSCP)	Takes advantage of the unique conformational 3D structure of a single stranded DNA. The diversity of loops and folds are visualized on a gel electrophoresis.
Real-Time PCR (qPCR)	Specific primers target a gene. Quantification of the gene is achieved using a fluorescent probe to monitor the amplification in real time.
Intergenic Transcribed Spacer Analysis (ITS)	Visualizes the differences in ITS regions, located between the 16S and 23S ribosomal genes, between closely related species.
Terminal Restriction Fragment Length Polymorphism (T-RFLP)	Fluorescently marked primers are used to amplify the target gene. Restriction enzymes are used to digest the amplicons into T-RF, which are separated and visualized on a capillary gel electrophoresis platform.
Fluorescence In Situ Hybridization (FISH)	Fluorescently labeled oligonucleotide probes marks bacteria cells , which are detected by epifluorescent microscopy, confocal laser microscopy or flow cytometry.
Denaturing High Performance Liquid Chromatography (DHPLC)	Uses a reverse-phase high performance liquid chromatography (HPLC) to interrogate PCR amplicons.
Length Heterogeneity PCR (LH PCR)	A universal fluorescently marked primer labels the variable regions within the 16S rDNA. The fragment profile is detected by fluorescence.

is expected to be reduce a different set of bacteria in comparison to traditional pasteurization. This new microbiota may contribute to beneficial flavors and aromas when the milk later is utilized for cheese production. The T-RFLP profiles of the milk samples were compared to observe differences in their microbiota as well as potential contamination sources. Furthermore, the effectiveness of two different analysis approaches was compared by analyzing the total bacterial population, using DNA directly extracted from the foods, versus the cultivatable population, using DNA from cultivated bacteria.

### 1.1.1 T-RFLP analysis

T-RFLP analysis typically involve four steps: DNA isolation and purification, PCR amplification and restriction enzyme digestion, separation of digested products via capillary gel electrophoresis and finally analysis and clustering of data to generate a fragment profile for each sample (Fig. 1). The PCR reaction involving a fluorescently labeled primer, tags one end of the PCR product with the fluorescent dye phosphoramidite fluorochrome 6-carboxy fluorescein (FAM). The amplicon is subsequently cut by restriction enzymes and separated by electrophoresis, followed by visualization of the terminal restriction fragments (TRFs) by excitation of the FAM. The obtained data provide information about the size in number of

base pairs (bp) and intensity of fluorescence (peak height), representing the amount of DNA. Relative abundance of a peak can be calculated by summing the total area of the TRFs and dividing each peak by the sum of all. It results in a normalization of the peak areas, which are no longer restricted to the amount of DNA loaded onto the electrophoresis. False peaks accounted for by run-to-run variability, PCR biases etc. can be alerted for by carrying out replicates. The final results give a representative profile of the TRFs and their relative abundance in the community sample.

T-RFLP is a powerful fingerprinting tool for a rapid comparison of microbial community compositions. As with all molecular approaches, it is however repeatedly subjected to the difficulties of sequence amplification. The homology between the primer and its target sequence is thus greatly significant. Studies have shown that a universal primer such as f8 only amplifies 76-98 % of the total bacterial 16S rRNA sequences in the RDP database (37). One must also consider the lack of completion of the database. In addition to bacteria 16S rRNA, the f8 primer also matches archeal 16S rRNA (37). Hence the description of f8 as a “universal” primer is debatable. Although amplification of archeal genes is possible, it is not frequently occurring. Further on, the presence of fungi and molds will not be indicated. Another consideration of the T-RFLP method is the fact that the 16S rRNA

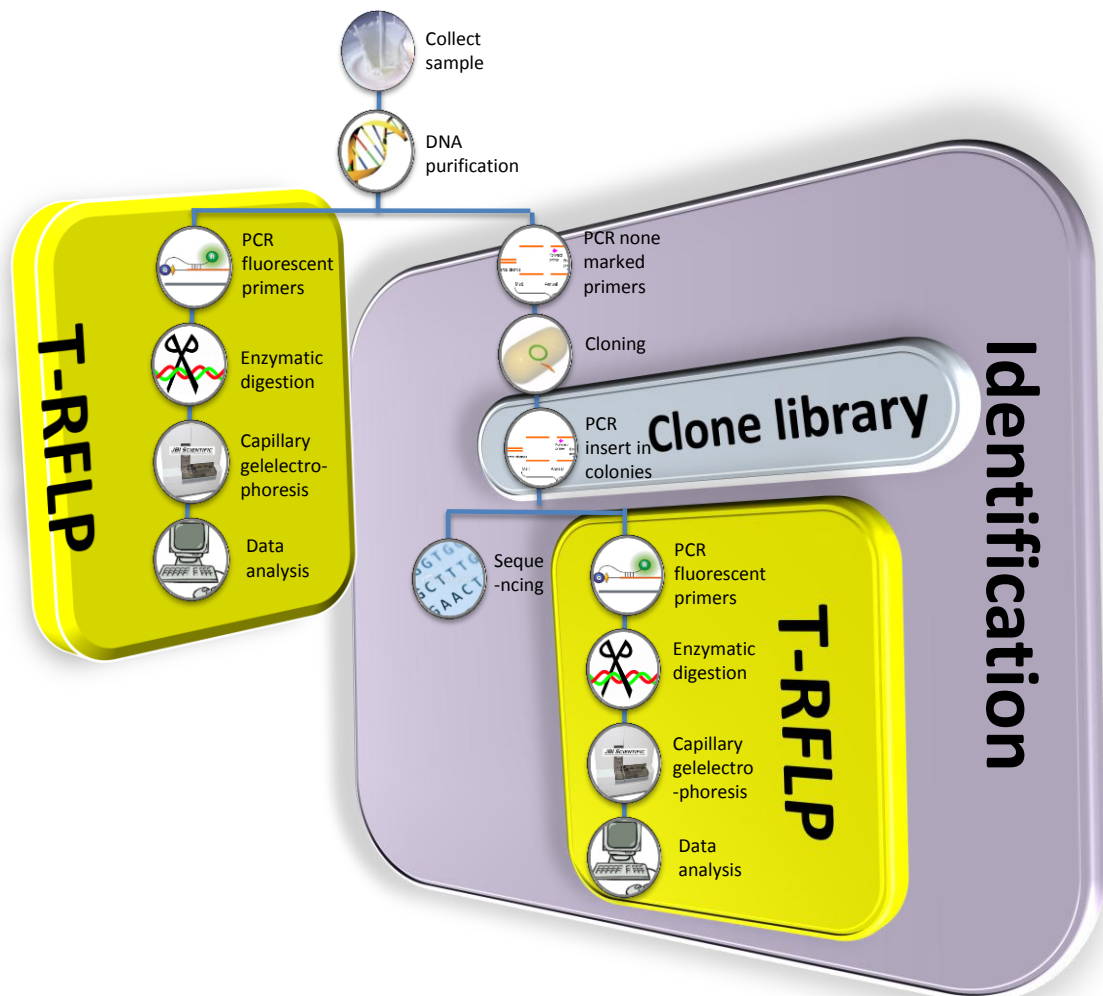


FIG. 1. Flowchart of the T-RFLP analysis and identification by construction of a clone library.

can exist in several copies within a single genome. Basing the bacterial abundance on the amount of amplicon of this gene, not considering the number of copies, can therefore give a deceptive view of the abundance.

Moreover, the number of PCR cycles influence the bacterial diversity found in a sample. More cycles have proven to give a higher bacterial diversity in clone libraries (2). The choice of PCR cycles also influences the product ratio (24). Polz with colleagues showed that differences in GC-content at the target site of the primer may contribute to PCR biases. The triple hydrogen bonds between G and C contribute to a higher melting temperature than for an AT-rich fragment. The GC-rich fragment will thus be more resistant to dissociation into single-stranded molecules. This outcome is an overrepresentation of molecules with a low GC-content in the final PCR product, due to its facilitated binding to the primer. Secondly, regions directly adjacent to the primer binding site may influence the hybridization efficiency of the primer. As concluded previously, T-RFLP is a rapid and high resolution method of analysis of communities, but due to the potential unequal amounts of amplicons it is thus not suitable for determining absolute abundance (34).

Another weakness of T-RFLP is the fact that one single TRF can originate from several phylogenetic separate species. To solve this bias there are two options; use several restriction enzymes to create diverse TRF lengths and later match the profiles with a database (8) or create a clone library. There are numerous web-based tools with T-RFLP profiles for different primer-restriction enzyme combinations assisting in the identification of the sample microbes. These include the phylogenetic assignment tool PAT (14), TRUFFLER (41) and APLAUS (31). Although they have been useful in gaining insight into the microbial sample community, there are still some aspects that need to be noted. Different fluorescent labeled dyes differ in electrophoretic mobility, contributing to differences between the empirically determined and the actual fragment size (Shyu, C., et al. Unpublished Data). Secondly, mismatches could occur due to the small percentage of 16S rRNA gene sequence in the database. The TRF sizing may therefore be imprecise and might not even give an identity to genus-specific taxonomy level. A more detailed taxonomy description is achieved through the use of a clone library. A library of inserts is constructed and its sequences are matched with the TRF profile. The great disadvantage of this approach is that it's a time consuming process. In addition, numerous clones may have to be screened to detect the TRFs with low abundance.

### **1.1.2 Clone library and identification**

A clone library is achieved by amplification of the 16S rRNA gene, followed by ligation of the PCR products into a vector. The vector is transformed into *Escherichia coli* cells that are cultured on selective plates. The plasmid inserts are amplified using universal primers and the insert is sequenced. In parallel to sequencing, the plasmid inserts are analyzed in a T-RFLP analysis. The sample profile from this analysis is combined with its sequence to give an identity to the observed peaks. This data is compared to the first environmental T-RFLP analysis to obtain the microbial identities of the sample (Fig. 1).



### **1.1.3 Method considerations**

This work involves combination of data from several different T-RFLP analyses; data from the initial T-RFLP analysis itself and data from the clone library. As previously mentioned, the method has 1 bp resolution and the boundaries of each TRF length are assessed manually. Thus, the combination of data between two T-RFLP analyses may contribute to misconceptions. For example, the same TRF could be assessed to have two closely spaced fragment lengths, e.g. 308 bp in lengths vs. 309 bp, in the initial T-RFLP analysis compared to the clone library. Combination of data between two T-RFLP analyses may therefore be deceptive. Further on, a source of error with molecular techniques is the DNA extraction. Bead beating, the method of choice for this work, is suitable for efficient lysis of Gram-positive bacteria, including spore-formers (42). It is critical that the bead beating is performed with a sufficient time to break the cell wall, but avoiding demolishing the DNA. Limitations in DNA extraction technique may contribute to a deceiving reflection of the original abundance.

## 1.2 Microbial community in foods

### 1.2.1 Bacterial flora in milk and cheese

The flavor and aromas of milk originates in its microbial content. The manufacturer faces a choice on how to treat the milk to develop the finest characteristics of the final product. Two diverse raw materials can be used for this purpose, raw milk or pasteurized milk. Utilization of both pasteurized and raw milk often involves addition of a lactic acid bacteria (LAB) starter culture which has been produced under controlled conditions in a laboratory. This method enables a regular product quality and a low health hazard. An alternative to pasteurized milk is raw milk. Raw milk with its natural LAB is among other things useful as a contribution to the flavor of the cheese. LAB produces lactic acid, degrades protein and produces carbon dioxide necessary for creating the cavities in cheese. Apart from contributing to flavor, lactic acid also lowers the pH for the coagulum to concentrate so that the whey can be separated from the milk. In addition to the starter culture, adjunct cultures are used to provide or enhance the characteristic flavors and textures of cheese.

The draw-back with using raw milk is the potential presence of pathogens. In the United States, cheese from raw milk must be ripened for at least 60 days at +4 °C (U.S. Food and Drug administration standard 21 CFR 133.182) to gain the pH, salt concentration, water activity ( $a_w$ ) and other parameters capable of inhibiting pathogens. By these means most pathogens are thought to be inhibited to reduce the health risk. The awareness and control of the bacteria diversity in the milk is therefore a very important aspect in production of cheese and other dairy product.

### 1. 2.2 Pasteurization

Pasteurization is used to minimize the pathogens and extend the shelf-life for food. It decreases the number of microorganisms in the product by heating, followed by an immediate cooling. The treatment does not sterilize, but do reduce the number of viable pathogens. Psychrophilic, gram-negative bacteria play an essential role in spoilage of dairy products (Table 2). The most critical group of bacteria contributing to milk spoilage through its production of lipolytic and proteolytic enzymes is *Pseudomonas* (40). At their optimal growth temperature of 0-10 °C they will degrade proteins, thereby contributing to spoilage of the milk by giving a disagreeable taste. *Pseudomonas* can among other things reduce the diacetyl-acetaldehyde ratio, giving a yogurt-like flavor to buttermilk. Presences of other pathogens such as *Mycobacterium paratuberculosis* and *Listeria monocytogenes* have also been reported in raw milk (6, 10).

TABLE 2. Bacterial activity in untreated milk at different temperatures. The star denotes spore-forming bacteria.

Classification	Growth range (°C)	Class	Genus
Psychrophile	-5 to + 22	<i>Actinobacteria</i>	<i>Corynebacterium</i>
		<i>Bacilli</i>	<i>Bacillus*</i>
		<i>Beta Proteobacteria</i>	<i>Alcaligenes</i>
		<i>Flavobacteria</i>	<i>Flavobacterium</i>
			<i>Acinetobacter</i>
		<i>Gammaproteobacteria</i>	<i>Proteus</i>
			<i>Pseudomonas</i>
Mesophile	+ 10 to + 47	<i>Cocci</i>	<i>Enterococcus</i>
		<i>Bacilli</i>	<i>Lactobacillus</i>
		<i>Actinobacteria</i>	<i>Microbacterium</i>
			<i>Micrococcus</i>
		<i>Bacilli</i>	<i>Streptococcus</i>
			<i>Bacillus*</i>
	<i>Clostridia</i>	<i>Clostridium*</i>	

Traditional pasteurization has proven successful in killing both gram-negative and gram-positive bacteria, as well as molds and yeasts (30). The technique is for instance effective in killing gram-negative Coliforms, which otherwise would cause a disagreeable taste and a destroyed texture of the cheese. Nevertheless, there are groups of organisms in milk that can survive pasteurization. Mesophilic bacteria thrive at 25-40 °C but some are capable of surviving the high temperatures of pasteurization (such as low temperature pasteurization at 72-73 °C). Typical examples of mesophiles found in pasteurized milk are found in the genera *Streptococcus* and *Lactobacillus*. Other relevant bacteria within this group for the food industry are the spore-forming bacteria *Bacillus* and *Clostridium*. For instance, *Clostridium tyrobutyricum* ferments lactic acid forming hydrogen gas and butyric acid, causing vast problems in both unwanted taste and health standards.

Further on, it is known that bacterial cells survive thermal heating with different success depending on its growth phase (i.e. lag-, logarithmic- or stationary phase, 13). They tend to be more stable at stationary growth phase and less resistant during the logarithmic phase. Older bacterial spores also tend to be more heat resistant than younger spores. Moreover, aggregation of spore-forming bacteria cells is favored during heating and cooling. Due to the many stages of growth in the raw milk, this provides a possible explanation to why some pathogens and other bacteria survive pasteurization.

## 2. Material and methods

### 2.1. Food sampling and DNA preparation

Raw milk, traditionally pasteurized milk (TP) and milk pasteurized with a novel pasteurization (NP) technique were collected from the cheese manufacturer. The milk samples with a potentially low bacterial load were concentrated by centrifuging  $3.000 \times g$  for 10 min in  $4^\circ \text{C}$  and the pellet was re-suspended in 1.8 ml peptone water. DNA isolation was performed according to MoBio PowerFood Microbial DNA Isolation Kit (Mobio, Carlsbad, CA, USA), with modified isolation by additional heating of tubes at  $70^\circ \text{C}$  for 10 min before the solution is transferred to Microbead tubes. Lysis was performed two times 5000 rpm for 50 sec in a BeadBeater (BioSpec Products, Inc., Bartlesville, OK, US), cooling the samples on ice between the repetition. The protocol was modified additionally for the cheese samples. DNA was isolated directly from 0.5 g cheese suspended in 450  $\mu\text{l}$  solution PF1. Further on, additional non-DNA organic and inorganic material such as cell debris and proteins were removed using solution PF2 incubated at an extended incubation time of 15 min. The concentration of purified DNA was controlled using gel electrophoresis and spectrophotometer (Eppendorf AG, Cambridge, UK). DNA was stored at  $-20^\circ \text{C}$  for further down-stream modification.

### 2.2 Determination of CFU

Plate count was used to determine the colony forming unit. The milk was diluted 10 fold and 1 ml was spread on a BAP (blood agar plate). 10 g of the cheese sample was mixed with peptone water and melted in  $45^\circ \text{C}$ , ca.60 min. The solution was diluted 10 fold and 1 ml was spread on BAP and TSA (tryptic soy agar). The plates were incubated aerobically at  $30^\circ \text{C}$  for 48 h. For DNA extraction, bacterial matter from the plates was pooled and dissolved in 1 ml peptone water. DNA isolation was performed as above.

### 2.3. T-RFLP analysis

The V3 region of the bacterial 16S rRNA was amplified using primer F8 (5'-AGA GTT TGA TCC TGG CTC AG- 3') (9) and 926R (5' -CCG TCA ATT CCT TT R AGT TT -3') labeled at the 5'-end with phosphoramidite fluorochrome 6-carboxy fluorescein (6-FAM) (15) to generate an FAM labeled amplicon of 918 bp. The PCR mixture (25  $\mu\text{L}$  per reaction) contained 2x iProof High-Fidelity Master Mix (Bio-rad Laboratories, Hercules, CA), 0.5  $\mu\text{molL}^{-1}$  of primer and 5  $\mu\text{L}$  of DNA template. A negative control (reaction without DNA template) was included in all experiments. The PCR amplification was carried out with a thermal cycler machine (Bio-rad Laboratories, Hercules, CA) and the cycle parameters were 3 min at  $98^\circ \text{C}$  for initial denaturation followed by 35 cycles with denaturation 10 s at  $98^\circ \text{C}$ , annealing for 30 s at  $55^\circ \text{C}$ , and extension for 1 min at  $72^\circ \text{C}$ ; and with a final elongation at  $72^\circ \text{C}$  for 10 min. The amplified products were visualized on a 1.5 % (w/v) agarose gel containing 0.4  $\mu\text{g.mL}^{-1}$  ethidium bromide along with a 100 bp DNA mass ladder (Bio-rad Laboratories, Hercules, CA). PCR products for each sample were digested in separate reactions according to the instructions provided by the manufacturer (New England BioLabs, Beverly, MA, USA) with 5 U of the restriction enzymes HaeIII (recognition site 5' GG/CC 3': 3' CC/GG 5') to a final volume of 25  $\mu\text{L}$ . The fluorescently labeled DNA fragments were separated and detected in capillary electrophoresis using an ABI3730XZ DNA Analyzer (Life technologies). TRF sizing was estimated using algorithms available in the software GeneScan (Applied Biosystems, Fosters City, CA, USA). The abundance of each TRF was expressed based on fluorescent intensity. Triplicates of the samples were aligned, and peaks that were not present in at least two of the triplicates were considered as background noise and removed to compensate for run-to-run variability in the T-RFLP analysis. The average relative abundance was calculated and the peaks with an average relative abundance lower than 0.005 were removed as background noise.

### 2.4. Clone library construction

To generate a clone library, the bacterial 16S rRNA gene was amplified from the extracted DNA using unlabeled primer set f8 and 926R. The PCR mixture (25  $\mu\text{l}$  per reaction) contained 3x iTaq DNA

Polymerase (Bio-rad Laboratories, Hercules, CA), 0.5nM of primer and 2  $\mu$ L of DNA template. A negative control (reaction without DNA template) was included in all experiments. The PCR amplification was carried out with a thermal cycler machine (Bio-rad Laboratories, Hercules, CA), and the cycle parameters were 3 min at 95 °C for initial denaturation followed by 35 cycles with denaturation 30 s at 95 °C, annealing for 30 s at 55 °C, and extension for 1 min at 72 °C; and with a final elongation at 72 °C for 30 min. Triplicate PCR products for each sample were pooled and the products were purified on a 0.7 % (w/v) agarose gel containing 0.4  $\mu$ g.mL<sup>-1</sup> ethidium bromide. PCR products were gel purified with QIAquick Gel Extraction Kit (Qiagen, Mississauga, ON, Canada). MilliQ water was used as elution buffer and was let standing in the column for 5 min before centrifugation. The purified DNA was pooled and its concentration visualized on a 1.5 % (W/V) agarose gel. The PCR product was ligated into a pCR4-TOPO vector (with ampicillin and kanamycin resistance genes for selection of positive clones) and transformed into *Escherichia coli* TOP10 OneShot chemical competent cells as specified by the manufacturer (Invitrogen, Carlsbad, CA). The competent cells were incubated overnight on selective lysogeny broth plates (50  $\mu$ g/ml ampicillin). Plasmid inserts from successfully transformed bacterial clones were amplified by PCR with the universal plasmid primers M13F (5'-GTA AAA CGA CGG CCA G-3') and M13R (5' -CAG GAA ACA GCT ATG AC -3'). The PCR product was diluted 50 folds and used in a F8 PCR reaction (25  $\mu$ l per reaction) as described above. The amplification product was visualized on a 1.5 % (W/V) agarose gel, The V3 region of the 16S rRNA of each plasmid vector was analyzed through T-RFLP, as described above.

## **2.5. Assembly of data for identification**

T-RFLP data were analyzed and peaks assumed to be background noise were removed as above. Data was assembled to create a sample profile of TRF length for each clone. The T-RFLP peaks from the clone library were identified using the sequenced clones (RDP and BLAST search). Spectrograms from the initial T-RFLP analysis and the identified clone library T-RFLP peaks were combined for a comprehensive visualization of the community structure.

### 3. Results

#### 3.1 Determination of CFU

To assess the number of bacteria in the food samples, the bacteria were cultivated on BAP. As expected, the raw milk exhibited a greater number of colony-forming units/ml (CFU/ml) for both sampling occasions compared to the pasteurized milk samples (Table 3). The CFU/ml value was in general higher for the first sampling of both the novel and the traditional technique. The second sampling showed more growth for the novel pasteurization compared to the traditional technique. In general, the CFU values of the pasteurized milks were relatively high in comparison to the CFU value of the raw milk. A reproduction of the experiment would therefore be desirable.

Raw milk		Novel Past.		Traditional Past.	
Sampl. 1	Sampl. 2	Sampl. 1	Sampl. 2	Sampl. 1	Sampl. 2
$3.8 \times 10^4$	$1.5 \times 10^4$	$9.0 \times 10^3$	$5.0 \times 10^3$	$6.5 \times 10^3$	$1.7 \times 10^3$

TABLE 3. Aerobic plate count (log CFU/ml) of cultured milk samples from raw milk, as well as milk pasteurized with traditional pasteurization and the novel pasteurization technique, from sampling 1 and 2.

#### 3.2 T-RFLP analysis

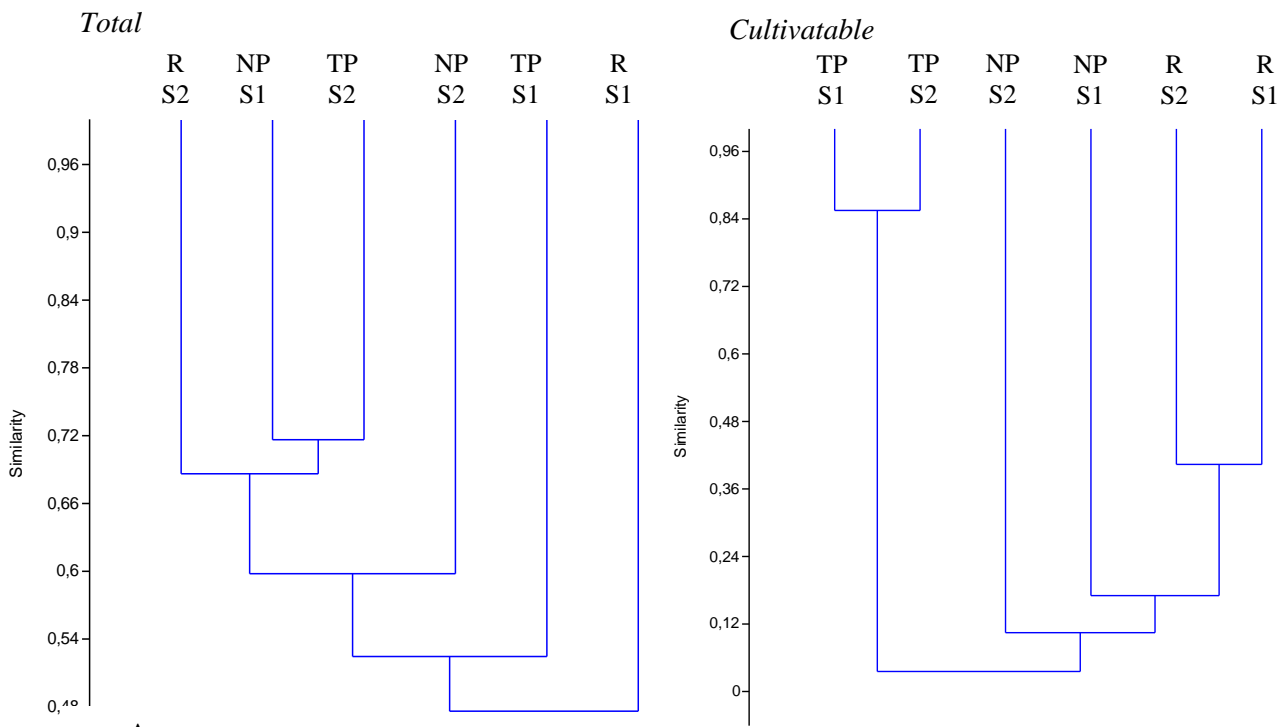
##### 3.2.1 Bacterial community profiles

The effectiveness of pasteurization was assessed by performing T-RFLP analyses on three different types of milks; raw milk, traditionally pasteurized milk (TP) and pasteurized milk with a novel technique (NP). Two separate T-RFLP analyses were performed on each milk type; one representing the cultivatable population using DNA from cultivated bacteria, e.g. cultivated bacteria from traditionally pasteurized milk (CTP), and one representing the total population using bacterial DNA directly extracted from the foods, e.g. DNA directly extracted from the traditionally pasteurized milk (DTP). The milk was analyzed from two manufacturing periods, sampling one and two.

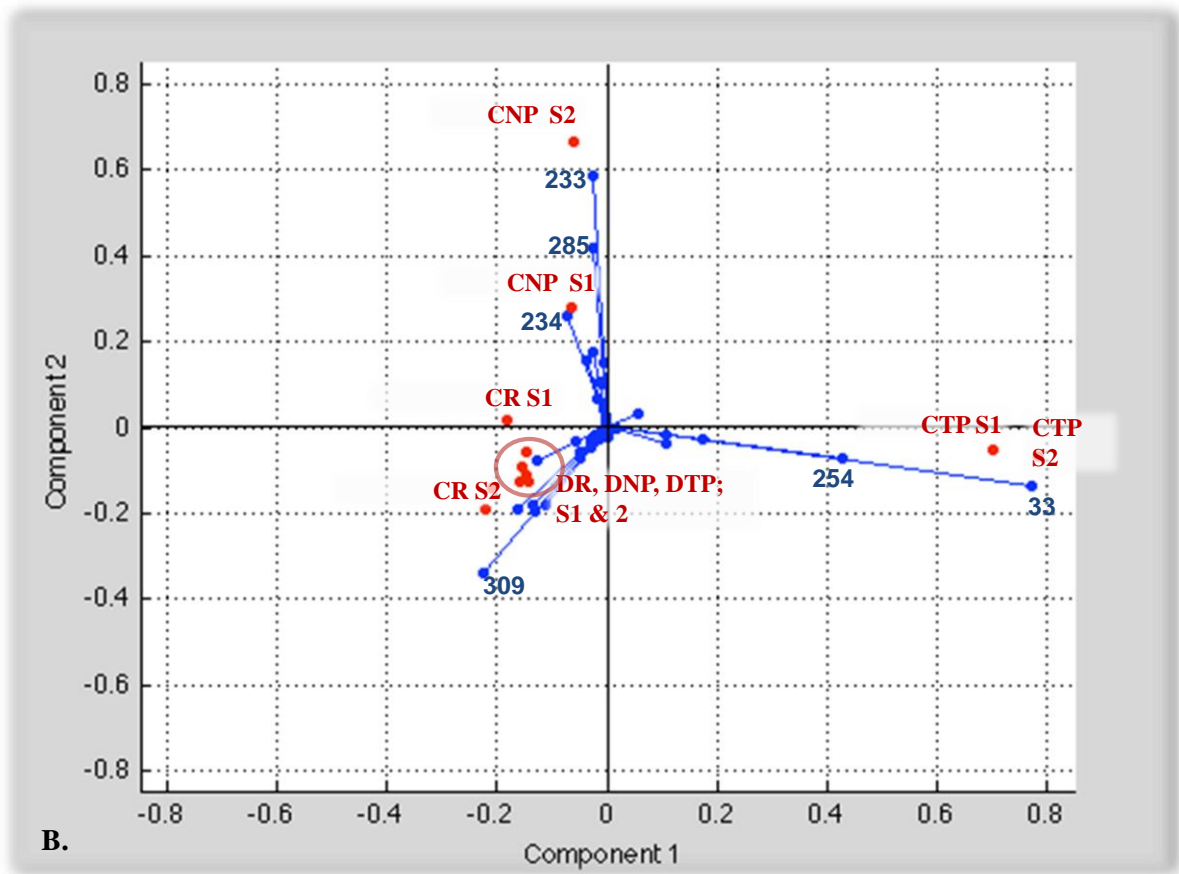
A total of 15 different TRFs, representing different genera, were found in the cultured raw milk (CR) from both sampling occasion one and two (Table 4.). Analysis of the cultivatable population of CNP (cultivated novel pasteurized) milk showed 13 and 11 TRFs for the first and second sampling respectively. Corresponding TRFs for the CTP (cultivated traditional pasteurized) milk were seven and nine. The numbers of TRFs found from the total population of raw milk (DR) from both

Sample	<i>Cultivable</i>			Total	<i>Total</i>			
	Sampling	Reduced	New		Sampling	Reduced	New	Total
Raw milk	1			15	1			17
	2			15	2			17
Nov past.	1	14	6	13	1	10	7	22
	2	15	5	11	2	13	4	19
Trad past.	1	20	6	7	1	16	6	18
	2	20	7	9	2	15	7	20
			Sum:	70			Sum:	113

TABLE 4. Number of TRFs found in the cultivatable and total population of the raw milk, novel pasteurized milk and the novel pasteurized milk from sample occasions one and two. The TRFs denoted *reduced* were detected in at least one out of the two samplings from the raw milk, but were not found in the denoted sample. TRFs denoted *new* were detected in at least one out of the two denoted samplings, but not in the raw milk.



A.



B.

FIG.1. Panel A shows a cluster analysis displaying the degree of similarity of the T-RFLP analysis of the cultivatable population and the total population of the raw milk (R), novel pasteurized milk (NP) and traditional pasteurized milk (TP), using Bray-Curtis parameters. Please note the differences in scale of similarity in the two analyses. Panel B shows a principal component analysis plot of the TRFLP analyses of the same samples from sampling one and two (S1 and S2). The blue color denotes the TRF length that motivates the position of the red dots (i.e. the samples).

samplings were 17. DNA from bacteria in DNP milk in the first sampling showed 22 TRFs and 19 at the second sampling. Finally, DNA extracted from bacteria in DTP milk from the first sampling had 18 TRFs and 20 at the second sampling. Each sample had a unique TRF profile, confirmed by principal component analysis (Fig. 1). A distinct outlier of the TRF profiles was the DNA extracted from CTP sampling one and two. In general, the cultivatable population showed more differences in the community profile between the sampling occasions than the total population.

TP was most effective in reducing the number of TRF from the raw milk, compared to NP milk. This pattern was reflected in the cultivatable population as well as the total population. Compared to the TRF in the raw milk, 14 different TRFs were successfully reduced in the CNP milk from first sampling (Table 4). Corresponding number of reduced TRFs from sampling two were 15. Similarly, the CTP milk reduced the number of TRFs by 20 from both. DNP milk showed that the treatment reduced 10 and 13 TRFs from sampling one and two, respectively. DTP milk reduced 16 and 15 TRF from the sampling one and two. The appearance of new TRFs, which were not found in the raw milk, was comparable for the cultivatable population (six and five for CNP milk sampling one and two, six and seven for the CTP milk). Corresponding numbers for the total population were seven from the first sampling and four from the second sampling from DNP milk. DTP milk had similar values, six and seven TRFs for sampling one and two. In conclusion, CTP had a smaller number of TRFs compared to CNP. Results show that T-RFLP analysis of the cultivatable population versus analysis of the total population visualizes different TRFs.

In contrast to the milk samples, the T-RFLP profile of the cheese samples showed one dominant characteristic TRF for each isolate (Fig.2). The TRF with the greatest relative abundance from the cultivatable population was 308 bp long. For isolate from the total population the corresponding TRF was 309 bp. Due to the limitations of one bp resolution in the T-RFLP analysis these peaks could however belong to the same bacteria or genera. In detail, DNA from CNP milk had five TRFs and CTP milk had seven. Corresponding numbers for DNP milk was two and for the DTP one.

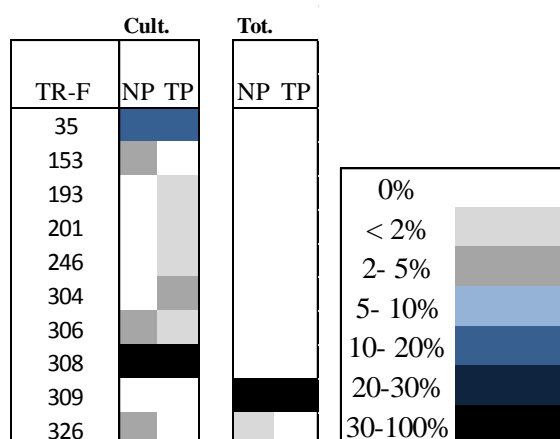


FIG 2. Relative abundance of TRFs from the cultivatable population (cult.) and the total population (pop.) The colors represent different relative abundance intervals of the TRFs found in the cheese produced from, the milk treated with the novel pasteurization technique (NP) as well as the traditional pasteurized method (C).



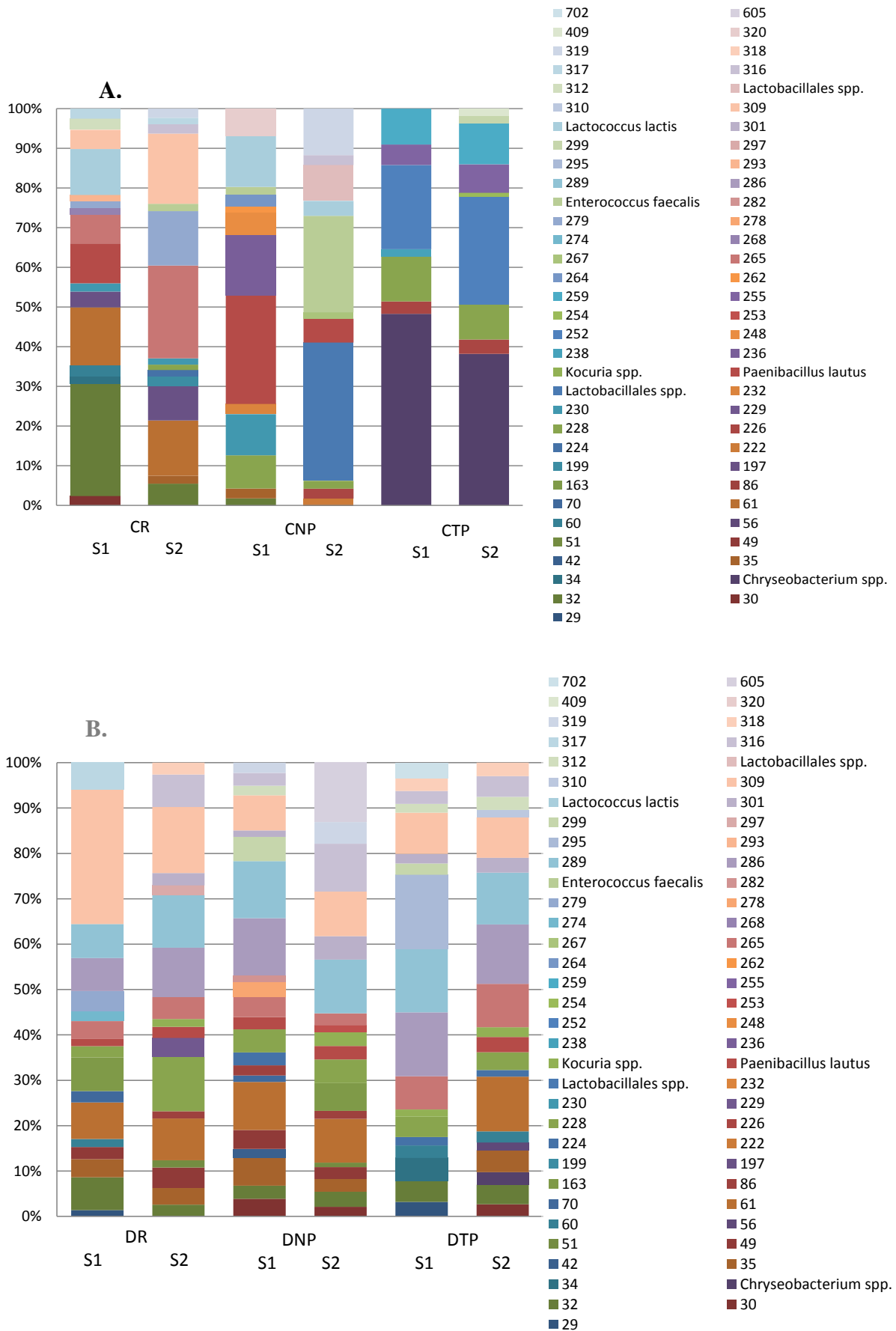


FIG 3. TRF profiles of the cultivatable population (A.) and the total population (B.). C denotes the cultivatable population and D the total population. Diagrams presents the relative abundance of the TRF (%) of the different samples of ram milk (R), novel pasteurized milk (NP) and traditionally pasteurized milk (NP) from sampling occasions one and 7

### 3.2.2. Method of detection

T-RFLP analyses were performed on bacterial DNA from two different approaches, the cultivatable microflora (C) and identification of the total flora (D). Their T-RF profiles were compared to assess the effectiveness of the different approaches in regards to the number and identity of TRFs they are able to detect.

Results showed that the TFR profiles between the cultivatable and the total population showed great differences in which TRF they visualized (Fig.3). T-RFLP analysis of the total population revealed 25 bacteria, from both sampling occasions, that could not be detected using the cultivation approach. Similarly, 23 bacteria groups were detected in the cultivable population that could not be detected in the total population. The sample profile demonstrated vast differences in the T-RFLP profile between the two different sampling occasions (Fig. 4). DNA extraction from the total population detects more TRFs, in total 70 from both samplings, compared to the DNA from the cultivatable population, 113 for both sampling occasions (Table 4). In conclusion, T-RFLP analysis from total population reveals a larger number of TRFs than the cultivatable approach. In addition to this, the profiles from the total population are more even between the sampling occasions. The TFRs that can be identified by the different approaches are different. The cheese TRF profiles were however contradictory to this (Fig 2). The cultivatable population showed a TRF profile with a total of 12 TRFs for TP and NP, while the T-RFLP analysis of the total population found two TRFs for the same samples. This result is most likely due to technical limitations in the technique so that not all the bacteria that are present in the samples are visualized.

### 3.2.3 Clone library identification

To give an identity to the TRF found in the T-RFLP analysis a clone library with a collection of diverse TRF lengths were created. Results from the clone library of the CNP from the second sampling successfully help to identify a total of 24 clones. The majority of these classified into the same bp range of TRF size, finally identifying seven diverse TRFs (Table 5). Most of these belonged to the lactic acid bacteria within the phylum *Firmicutes*. The TRF with the size of 33 bp were not detected in the T-RFLP analysis of the sample chosen for clone library construction. This TRF was however detected with a high relative abundance in the traditional pasteurized milk. A similar unexpected result is the TRF with 235 bp of length. This fragment was not found in any of the DNA from the cultivated samples, but was found in the DNA directly extracted from the milk. It is however desirable to expand the clone library with more identities to yield a more identification.

TRF (bp)	Phylum	Class	Family	Genus	Species
33	Bacteroidetes	Flavobacteria	Flavobacteriaceae	Chryseobacterium	<i>Chryseobacterium spp.</i>
233	Firmicutes	Lactobacillales	-	-	-
234	Firmicutes	Bacillales	Paenibacillaceae	Paenibacillus	<i>Paenibacillus lautus</i>
235	Actinobacteria	Actinobacteridae	Micrococcaceae	Kocuria	<i>Kocuria spp.</i>
285	Firmicutes	Lactobacillales	Enterococcaceae	Enterococcus	<i>Enterococcus faecalis</i>
308	Firmicutes	Lactobacillales	Streptococcaceae	Lactococcus	<i>Lactococcus lactis</i>
311	Firmicutes	Lactobacillales	-	-	-

TABLE 5. Phylogenetic distribution of clones identity and length (bp) recovered from the clone library combined with the T-RFLP analysis.

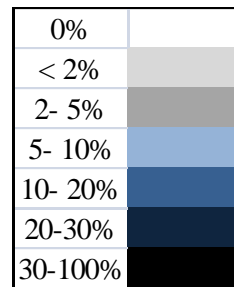
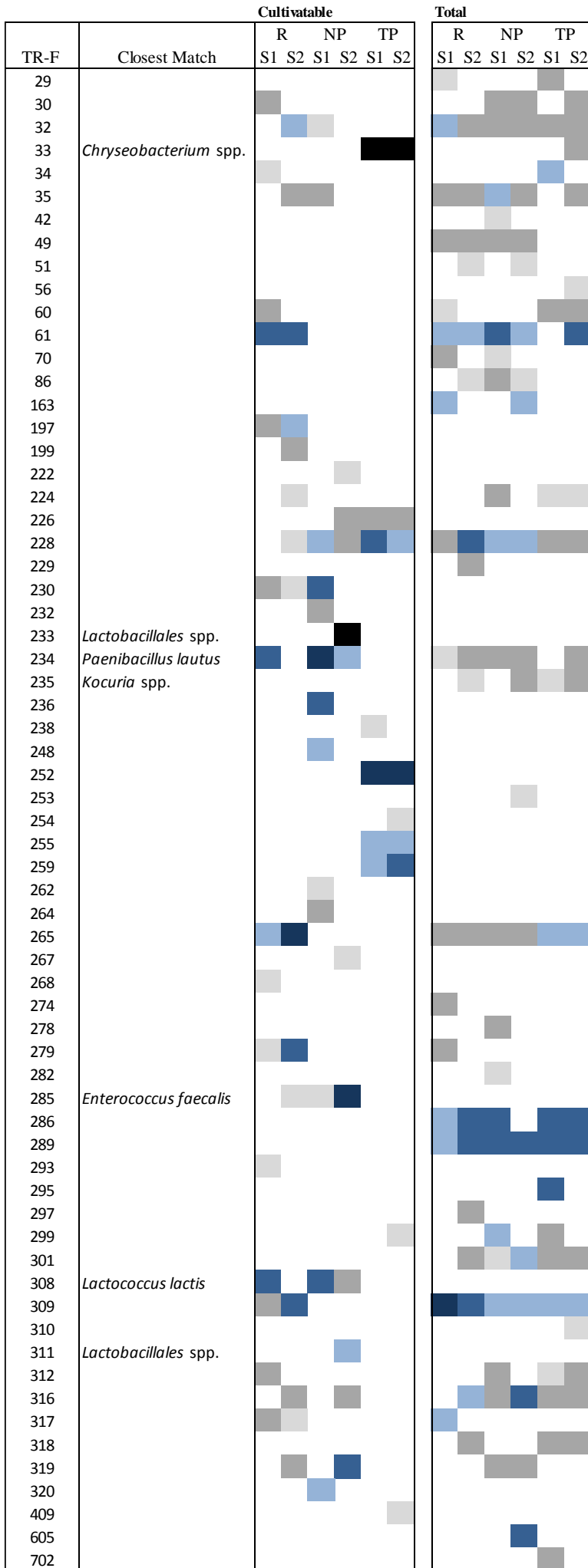


FIG 4. TRF profile of the cultivatable and the total population from sampling 1 and 2 (S1 and S2). The colors represent different relative abundance intervals of the TRFs found in the raw milk (R), the novel pasteurized milk (NP) as well as the traditional pasteurized milk (TP).

## 4. Discussion

In this work we analyzed and compared the microbial flora of raw milk, traditionally pasteurized milk and milk pasteurized with a novel technique. The microbial flora of cheese made from the traditionally pasteurized milk as well as cheese made from milk pasteurized with the novel technique was also investigated. All work was performed using molecular analysis. A conserved region of the bacterial 16S rRNA gene was analyzed using T-RFLP analysis (22, 7). The identities represented by the acquired TRFs were specified by constructing a clone library. In recent years, numerous studies have confirmed the advantages of molecular DNA analysis for microbial identification over traditional, culture-dependent approaches (18). Foremost, culture-independent methods do not apply any selective conditions of growth and therefore visualizes a wider range of bacteria. In combination with a clone library, T-RFLP also has the great advantage of identifying a large number of bacteria as close as down to species level; in comparison to culture-dependent approaches were this would be considerably more labor-intensive.

Pasteurization is used to control the microbiological quality of milk. In this study, traditional pasteurization showed a greater decrease in number of bacterial species than the novel pasteurization technique. The TRF profile of the bacteria in the milk samples treated with the two different pasteurization techniques were different, suggesting that the techniques indeed differentiate in which bacteria they reduce.

Interestingly, there are TRFs appearing after treatment that were not identified in the raw milk. These were different from each other between the pasteurization techniques and also between the two sampling occasions. This indicates a recontamination after pasteurization. The results are however somewhat confusing considering the setup of the production line, where the same equipment was used for both pasteurization techniques without cleaning between them. Another possible reason for why new TRFs appear after pasteurization could be that these fragments were indeed present in the raw milk, but were not visualized in the T-RFLP analysis due their small abundance. As some of the fragments with higher abundance are reduced through the pasteurization process, TRFs with smaller abundance can be visualized due to growth or technical limitations. Finally, a possible explanation to the diverse microbiota is the bias that the milk has different origin between production dates. As established by Bonizzi and his colleagues, the feeding area of the cows plays a critical role in determining the cheese features (3). The milk used for this study did however originate from a one single farm and from solely two batches of milk. This factor is therefore assumed to be negligible. The source of contamination remains unknown.

In contrast to the variable microbial population found in milk, the cheese microbiota was considerably less diverse. Similar to previous studies which characterized the microbiota in cheese, there was one dominant TRF length in each of the cheese isolates (26). In this case, these two most abundant bacteria are likely lactic acid starter culture. Due to lack of completion of the clone library, the identity of these TRFs does however remain unknown. In contrast to these results, other studies show a more diverse flora in cheeses depending on the

type of cheese examined; hard, semi-hard or soft cheese etc. (25). The key sources for such a diverse microbiota is partly due to the environmental conditions of the ruminant producing the milk yielding different whey cultures, the ripening conditions as well as alternative addition of a starter culture. Moreover, the time of sampling during the maturation/fermentation process is obviously crucial for what flora will be found.

As expected, the number of bacteria found from the total population was greater than from the cultivatable population. This observation is most likely due to the selective conditions applied during culturing, enabling a limited part of the microbiota to grow and be characterized. These conditions will enable some bacteria to grow more extensively than others, resulting in an overrepresentation in the total microbiota. As stated previously, the culture-based versus the molecular identification approach detects different parts of the microbiota. Surprisingly, these findings were contradictory for the cheese samples, where cultivated DNA had a greater number of bacteria than directly extracted DNA. It is likely due to favorable selectable medium, contributing enhanced growth for bacteria other than the dominating flora and therefore an overrepresentation of the bacterial abundance.

The similarity between the manufacturing dates of the total population was the most similar to each other, compared to the results from the cultivatable population. The culture-independent technique might therefore be advantageous when applying T-RFLP to visualize microbial flora in the food industry. By characterizing the normal flora of the foods and routinely look at TRF profiles of the product, a contamination creating a differentiated TRF profile could easily be detected. In addition to contributing to microbial safety, T-RFLP is also a powerful tool to gain knowledge about the microbiota in foods. In the food industry, these understandings may be used to developing novel aromas of the foods or create more sustainable foods with an increased shelf-life.

T-RFLP visualizes all DNA, living as well as dead. In this work, it has been a drawback. Owing to representation of dead bacterial DNA in the total characterization of microbiota, the efficiency of the different pasteurization techniques in killing off bacteria is difficult to assess from directly extracted DNA. This effect can only be evaluated through culture-based DNA, however with the drawback of selection. Previous molecular studies have been successful in distinguishing living from dead DNA, using ethidium monazide (EMA) or propidium monazide (PMA) (12, 27). These compounds penetrate the membrane of dead cells and prevent PCR amplification of dead DNA, thereby only visualizing the DNA from living cells. In recent years EMA has successfully been used in combination with T-RFLP to investigate the effect of antibiotics on the microbiota in the intestinal loop in ruminants (12). Discriminating compounds like these might therefore be an option for future similar analysis. In this study, it is however assumed that the abundance of dead DNA was relatively small in relationship to the living DNA (particularly in the cheese, where the bacteria had the ability to grow) and that this bias to some extent can be disregarded. Finally, it is also important to emphasize that the primers used in this research did not amplify the yeast population of milk, which are also present in raw milk and affects the milks quality and sustainability (5).

As expected, a majority of the identified bacteria belonged to the *Firmicutes* phylum and were lactic acid bacteria (LAB). These results correlates to parallel studies at SIK where sequencing of cultured bacterial DNA from the same milk samples showed an overrepresentation of LAB in the novel pasteurized milk, in comparison to the traditional pasteurized milk that after ripening only contained the starter culture. To the manufacturer, this opens up new possibilities of producing cheeses with a novel microbiota and flavor. The main reason for pasteurization is nevertheless to reduce the number of viable pathogens. The clone library proved to be successful in identifying seven of the TRFs. In comparison to all the clones screened and analyzed, this is however a fairly low success rate. The majority of the sequenced clones turned out to belong to the same TRF sizes, therefore limiting the number of identified TRFs. It is however unclear why some fragments were successfully ligated into the vector and others not. It seems to be no correlation between the relative abundance of the TRFs and their success ratio of ligation into the vector. A complete identification of the TRFs would therefore involve many clone libraries per sample and would result in a time consuming work.

The T-RFLP analysis revealed many similar peaks with less than two bp of difference in size. Without identification of these fragments using a clone library it is very hard to be completely confident of the separation of these fragments into individual bacteria during data analysis. To complete this study, it would thus be desirable to create additional clone libraries to complete the identification process. Moreover, some of the identities found in the sample selected for the clone library construction could not be found in the T-RFLP analysis of the same sample. This is likely due to the loss of data as the T-RFLP data is processed, neglecting the fragments with low abundance. Thus, some of the bacteria might be present in the physical sample but not represented in the T-RFLP profile.

## **5. Conclusion**

The microbial structure of differently pasteurized milk and cheese was revealed using T-RFLP combined with clone libraries. The different pasteurization techniques proved to be effective in reducing different bacteria in relation to each other. The key question of the identity of these bacteria does however still remain unanswered. The eligibility of using clone library for the identification process may be questioned. Even newer molecular techniques, such as pyrosequencing, could provide a more convenient approach for a rapid and complete identification of the microbiota, without the use of clone library (18). The results confirmed the advantages of using molecular opposed to culture-based approaches to characterize microbiota. The study also verified the applicability of T-RFLP for analyzing microbiota in foods.

## **6. Acknowledgment**

I would like to thank all my colleagues at SIK Microbiology and Process Hygiene department for making my work there a great and stimulating process. It has been a pleasure! In particular I want to thank my supervisor Klara Båth. Thank you for giving me the opportunity to familiarize with the fascinations and frustrations of molecular biology research. Your guidance and support have been crucial for this project.

## 7. References

1. **Blackwood, C.B., Marsh, T., Kim, S-H and Paul, E.A.** 2003. Terminal restriction fragment length polymorphism data analysis for quantitative comparison of microbial communities. *Appl. Environ. Microbiol.* **69**:926-932.
2. **Bonnet, R., et al.** 2002. Differences in rDNA libraries of faecal bacteria derived from 10-and 25-cycle PCRs. *Int. J. Syst. Evol. Microbiol.* **52**:757-763.
3. **Bonizzi, Buffoni, J.N., Feligini, M. and Enne, G.** 2009. Investigating the relationship between raw milk bacterial composition, as described by intergenic transcribed spacer-PCR fingerprinting, and pasture altitude. *J. Appl. Microbiol.* **107**:1319-1329
4. **Borneman, J. and Triplett, E. W.** 1997. Molecular microbial diversity in soils from eastern Amazonia: Evidence for unusual microorganisms and microbial population shifts associated with deforestation. *Appl. Environ. Microbiol.* **63**:2647-2653.
5. **Cocolin, L., et al.** 2002. An application of PCR-DGGE analysis to profile the yeast populations in raw milk. *Int. Dairy J.* **12**: 407-411.
6. **Doyle, M. P., et al.** 1987. Survival of *Listeria monocytogenes* in milk during high-temperature, short-time pasteurization. *Appl Environ Microbiol.* **53**: 1433-1438.
7. **Ercolini, D., Hill, J. P. and Dodd, E. R. C.** 2003. Bacterial community structure and location in Stilton cheese. *Appl. Environ. Microbiol.* **69**:3540-3548
8. **Edlund, A., Soule, T., Sjöling, S. and Jansson, J.K.** 2005. Microbial community structure in polluted Baltic Sea sediments. *Environ. Microbiol.* **8**:223-232.
9. **Edwards, U., et al.** 1989. Isolation and direct complete nucleotide determination of entire genes. Characterization of a gene coding for 16S ribosomal RNA. *Nucleic Acids Res.* **17**:7843-7853.
10. **Grant, I. R., H. J. Ball, Neill, S. D. and Michael, T. R.** 1996. Inactivation of *Mycobacterium paratuberculosis* in cows' milk at pasteurization temperatures. *Appl. Environ. Microbiol.* **62**:631-636.
11. **Hayashi, H., Sakamoto, M., Kitahara, M. and Benno, Y.** 2003. Molecular analysis of fecal microbiota in elderly individuals using 16S rDNA library and T-RFLP. *Microbiol. Immunol.* **47**:557-70.
12. **Inglis, G. D., Kastelic, J. P. and Uwiera, R. E.** 2010. Catheterization of intestinal loops in ruminants does not adversely affect loop function. *J. Am. Assoc. Lab. Anim. Sci.* **60**:469-478.
13. **Jay, J. M., Loessner, M. J. and Golden, D. A.** 2005. Food Protection with high temperatures, and Characteristics of thermophilic microorganisms, p. 415-421. *In: Heldman D.R (ed.), Modern Food Biology, 7nd ed. Springer Science+ Business Inc. New York, NY.*
14. **Kent, A.D., Smith D.J., Benson, B.J. and Triplett, E.W.** 2003. Web-based phylogenetic assignment tool for analysis of terminal restriction fragment length polymorphism profiles of microbial communities. *Appl. Environ. Microbiol.* **69**:6768-6776.
15. **Law, B. A. and Tamime, A. Y.** 2010. *Technology of cheesemaking, (2nd ed.), Wiley-Blackwell, Oxford, UK.* doi: 10.1002/9781444323740.
16. **Liu, W-T., Marsh, T. L., Cheng, H. and Forney, L. J.** 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl. Environ. Microbiol.* **63**:4516-4522.
17. **Lüke, C., Bodrossy, L., Lupotto, E. and Frenzel, P.** 2011. Methanotrophic bacteria associated to rice roots: the cultivar effect assessed by T-RFLP and microarray analysis. *Environ. Microbiol. Rep.* **3**:518-525.
18. **Masoud, W. et al.** 2011. Characterization of bacterial populations in Danish raw milk cheeses made with different starter cultures by denaturing gradient gel electrophoresis and pyrosequencing. *Int. Dairy J.* **21**:142-148.
19. **Matsumoto, M., Sakamoto, M., Hayashi, H. and Benno, Y.** 2005. Novel phylogenetic assignment database for terminal-restriction fragment length polymorphism analysis of human colonic microbiota. *J. Microbiol. Methods* **61**:305-319.
20. **Muyzer, G., de Waal, E. C. and Uitterlinden, A.G.** 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* **59**:695-700.
21. **Nieminen, .T., et. al.** 2011. Characterization of psychrotrophic bacterial communities in modified atmosphere-packed meat with terminal restriction fragment length polymorphism. *Int. J. Food Microbiol.* **144**: 360-366.
22. **Ogier, J-C., et al.** 2002. Identification of the bacterial microflora in dairy products by temporal temperature gradient gel electrophoresis. *Appl. Environ. Microbiol.* **68**:3691-3701.



23. **Øverås, L. and Torsvik, V.** 1998. Microbial diversity and community structure in two different agricultural soil communities. *Microbiol. Ecol.* **36**:303-315
24. **Polz., M. F. and Cavanaugh, C. M.** 1998. Bias in Template-to-Product Ratios in Multitemplate PCR. *Appl. Environ. Microbiol.* **64**:3724–3730.
25. **Quigley, L., et al.** 2011. Molecular approaches to analyzing the microbial composition of raw milk and raw milk cheese. *Int. J. Food Microbiol.* **150**:81-94.
26. **Rademaker, et al.** 2004. The surface microflora dynamics of bacterial smear-ripened Tilsit cheese determined by T-RFLP DNA population fingerprint analysis. *Int. Dairy J.* **15**:758-794.
27. **Rapp'e, M. S. and Giovannoni, S. J.** 2003. The uncultured microbial majority. *Annu. Rev. Microbiol.* **57**:369–94.
28. **Rogers, G. B., et al.** Assessing the diagnostic importance of nonviable bacterial cells in respiratory infections. *Diagn. Microbiol. Infect. Dis.* **62**:133–141.
29. **Sakamoto, M., et al.** 2004. Changes in oral microbial profiles after periodontal treatment as determined by molecular analysis of 16S rRNA genes. *J. Med. Microbiol.* **53**:563-571.
30. **Shah, N. P.** 1994. Psychrotrophs in milk: a review. *Milchwissenschaft* **49**:432–437.
31. **Shyu C., et al.** 2007. MiCA: a webbased tool for the analysis of microbial communities based on terminal-restriction fragment length polymorphisms of 16S and 18S rRNA genes. *Microb. Ecol.* **53**:562–570.
32. **Simon, L., Levesque R.C, and Lalonde, M.** 1993. Identification of endomycorrhizal fungi colonizing roots by fluorescent single-strand conformation polymorphism-polymerase chain reaction. *Appl. Environ. Microbiol.* **59**:4211–4215.
33. **Stres, B., et al.** 2008. Influence of temperature and soil water content on bacteria, archeal and denitrifying microbial communities in drained fen grassland soil microorganism. *FEMS Microbiol. Ecol.* **66**:110-122.
34. **Suzuki, M.T, and Giovannoni, S.J.** 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Appl. Environ. Microbiol.* **62**:625-630.
35. **Tonin, C., et al.** 2001. Assessment of arbuscular mycorrhizal fungi diversity in the rhizosphere of *Viola calaminaria* and effect of these fungi on heavy metal uptake by clover. *Mycorrhiza*, **10**:161-168.
36. **Torsvik, V., Gokosyr, J., and Daae, F.L.** 1990. High diversity of DNA of soil bacteria. *Appl. Environ. Microbiol.* **56**:782-787.
37. **Schütte, U. M. E., et al.** 2008. Advances in the use of terminal restriction fragment length polymorphism (T-RFLP) analysis of 16S rRNA genes to characterize microbial communities. *Appl. Microbiol. Biotechnol.* **80**:354-380
38. **Ward, D. M., R. Weller, and M. M. Bateson.** 1990. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* **345**: 63-65
39. **When, Q., Chen, Z. and Shi, H.** 2008. T-RFLP detection of nitrifying bacteria in a fluidized bed reactor of achieving simultaneous nitrification-denitrification. *Chemosphere.* **71**:1683-92.
40. **Wiedmann, M., et al.** 2000. Molecular and phenotypic characterization of *Pseudomonas* spp. Isolated from milk. *Appl. Environ. Microbiol.* **66**:2085–2095.
41. **Wise, M. J. and Osborn A. M.** 2001. TRUFFLER: programs to study microbial community composition and flux from fluorescent DNA fingerprinting data. *Proc IEEE Int Symp Bioinformatics Bioeng.* 129–135.
42. **Zoetendal, E. G., et al.** 2001. DNA isolation protocols affect the detection limit of PCR approaches of bacteria in samples from the human gastrointestinal tract. *System. Appl. Microbiol.* **24**:405–410.

## 8. Appendix A- Complementary studies

To better establish and evaluate the method for constructing the clone library, initial attempts were made to clone salmon DNA into *E.coli*. This section of the project is part of a previous study made at SIK, where T-RFLP was applied to analyze the microflora of salmon. Due to lack of completion and vagueness in the results of this clone library, attempts were made to re-construct the clone library.

### 8.1. Material and methods

#### Food sampling and DNA preparation

Bacteria were released from the salmon sample by running 25 g of foods with 225 ml peptone water in a stomacher. DNA from milk sample was extracted as above. 1.8 ml of the salmon solution was used for downstream applications. DNA isolation was performed according to MoBio PowerFood Microbial DNA Isolation Kit, with modified isolation by additional heating of tubes at 70°C for 10 min before the solution is transferred to Microbead tubes. Lysis was performed two times 5000 rpm for 50 sec in a BeadBeater (BioSpec), cooling the samples on ice between the repetition. The tubes were then centrifuged 10,000 x g for 1 min. DNA purification using solution PF2 was incubated at an extended incubation time of 15 min. The concentration of purified DNA was determined with gel electrophoresis. DNA was stored at - 20 °C for further down-stream modification.

#### Clone library construct

To generate a clone library, the bacterial 16S rRNA gene was amplified from the extracted DNA as described above using the unmarked primer set F8 and 926R. 4 numbers of tests were made to achieve a successful clone library using the salmon DNA. 1) Initial attempts were made with the PCR mixture (25 µl per reaction) contained 3x iProof DNA Polymerase (iProof High-Fidelity Master Mix, BIO-RAD), 0.5 µmol.l<sup>-1</sup> of primer and 2 µl of DNA template. An additional 2 µl BSA (10mg/ml) was added. A negative control (reaction without DNA template) was included in all experiments. The PCR amplification was carried out with a thermal cycler machine (BIO-RAD), and the cycle parameters were 3 min at 95 °C for initial denaturation followed by 35 cycles with denaturation 30 s at 95 °C, annealing for 30 s at 55 °C, and extension for 1 min at 72 °C; and with a final elongation at 72 °C for 7 min. Triplicate PCR products for each sample were pooled and the products were purified on a 1.5 % (w/v) agarose gel containing 0.4 µg.mL<sup>-1</sup> ethidium bromide. PCR products were purified with QIAquick PCR Purification Kit (Qiagen, Mississauga, ON, Canada). MilliQ water was used as elution buffer and was let standing in the column for 5 min before centrifugation. An initial attempt to purify the DNA was also made using QIAquick PCR Purification kit (Qiagen, Mississauga, ON, Canada). The purified DNA was pooled and its concentration visualized on a 1.5 % (W/V) agarose gel. The PCR product was ligated into a pCR4-TOPO vector and transformed into *Escherichia coli* TOP10 OneShot chemical competent cells as specified by the manufacturer (Invitrogen, Carlsbad, CA), using 2µ PCR product and cultivation on kanamycin plates. 2) A second attempt was performed using the same parameters as above, with disregard to the BSA in the PCR mixture, which was removed. A newer cloning set was also used. 3) In test number 3, 3'A-overhangs were added to the gel-purified PCR product from attempt 2 above according to the cloning kit manufacturer (Invitrogen). Cloned cells were cultured on Lysogeny broth (LB) plates containing 50µg/ml ampicillin. 4) In test number 4, the PCR mixture was exchanged to iProof High-Fidelity Master Mix (BIO-RAD), 0.5 µmol.l<sup>-1</sup> of primer and 5 µl of DNA template. The dNTP's used were taken from the TOPO cloning kit. A

negative control (reaction without DNA template) was included in all experiments. The PCR amplification was carried out with a thermal cycler machine (BIO-RAD), and the cycle parameters were 3 min at 95 °C for initial denaturation followed by 35 cycles with denaturation 30 s at 95 °C, annealing for 30 s at 55 °C, and extension for 1 min at 72 °C; and with a final elongation at 72 °C for 30 min. Cloning was performed. The following procedure was as described above. Cloned cells were cultured on Lysogeny broth (LB) plates containing 50µg/ml kanamycin, as well as 50µg/ml ampicillin LB plates. To get an indication of the character of the colonies, gram-staining was performed. After incubation 48h, the colonies were re-stroked onto a new LB plate. After another 48h incubation, plasmid inserts were amplified by PCR with the universal plasmid primers M13F (5'-GTAAAACGACGGCCAG-3') and M13R (5' -CAGGAAACAGCTA TGAC -3'). The PCR mixture (25 µl per reaction) iProof High-Fidelity Master Mix was PCR amplified as described above. The amplification product was visualized on a 1.5 % (W/V) agarose gel.

## **8.2. Results**

No colonies could be observed after incubation 48 h for tests 1 to 3. After incubation 24 h during test 4, very few colonies were observed. The number increased slightly after 48 h of incubation. Subsequent PCR reaction to amplify the PCR insert revealed a diverse set of fragment inserts, indicating that the cloning was unsuccessful. Eventually, a clone library was achieved as described above using DNA from milk. The key to success was an optimized gel purification procedure where a lower percentage of agarose (0.7 %) was used in combination with a lower voltage (85 V). Initially, PCR amplification of the 16S rRNA gene within the insert showed a larger fragment than 920 bp. This problem was resolved by diluting the DNA in the PCR product, from M13 primer reaction, 50-fold before it was used in the 16S rRNA PCR reaction.

