APPENDIX

# A

# Sample Laboratory Reports

The "Professor's Comments" on each laboratory report are shown in red in square brackets and refer to the preceding underlined words.

## REPORT 1

**Name:** Connie Lee
**Title** **[[Professor's Comment:** Good.**]]**
*Vibrio fischeri* Luciferase Operon on Plasmid pKN800 Contributes **[[Professor's Comment:** Confers**]]**Bioluminescence to *Escherichia coli*
**Date:** January 12, 2012–January 18, 2012

### Purpose

To study expression of the *Vibrio fischeri* luciferase (*lux*) operon in *Escherichia coli* (*E. coli*) by purifying a plasmid DNA from *E. coli*, determining the structure of the plasmid using restriction mapping, and introducing the recombinant plasmid into another *E. coli* strain using transformation. **[[Professor's Comment:** Nice, but too long. Break into two or three sentences.**]]**

### Methods

We received a strain of *E. coli* containing unspecified pKN800 plasmid DNA. We purified the plasmid DNA and used *Pst*I restriction enzyme to cut the plasmid DNA. We loaded *Pst*I-cut and uncut pKN800 plasmid DNA into an agarose gel for electrophoresis. We transformed *E.coli* strain DH5α with the pKN800 plasmid DNA.

We followed the procedures on p. 43–54 of the laboratory manual with the following exceptions (2): Step A-11 on p. 45: The tube was centrifuged again for five seconds after centrifuging the tube for 10 seconds. Step A-11 on p. 45: The cap of the tube was open to evaporate the ethanol. Step C-1 on p. 48: We used an 1.0% agarose gel to separate the restriction

fragments in the DNA samples. Step D-6 on p. 52: We incubated the samples at 225 rpm for 45 minutes at 37 °C. Steps E-1 to E-4 on p. 54: We examined the plates for two days.

## Results

We received an *E. coli* strain containing an unspecified pKN800 plasmid DNA. To determine whether the orientation of the *lux* operon in pKN800 affects luminescent ability in *E. coli*, we purified pKN800 plasmid DNA and determined the orientation of the *lux* operon. To confirm the function of pKN800 plasmid DNA in luciferase production, we transformed pKN800 plasmid DNA into *E. coli* DH5α and recorded the number of ampicillin-resistant and luminescent transformants. **[[Professor's Comment: Excellent.]]**

To determine the orientation of the purified pKN800, we digested the plasmid with *Pst*I, which cuts the plasmid cloning vector pBR322 once and the *lux* operon once. We separated the restriction fragments of the *Pst*I-cut and uncut pKN800 plasmid DNA by agarose gel electrophoresis. Lane one had distinctive bands, which were the 1000-bp molecular weight standard (Fig. A1.1). Lane two had two fragments, containing restriction fragments of the *Pst*I-cut pKN800 plasmid DNA; the top band was more intense than the bottom band. Lane three had only one band, containing the supercoiled uncut pKN800 plasmid DNA; the intensity of the band was low (Fig. A1.1).

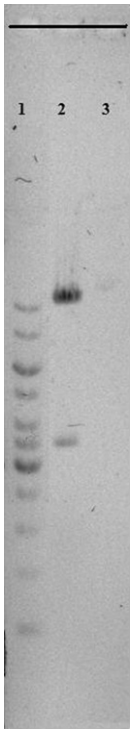


FIGURE A1.1 **Agarose gel electrophoresis of *Pst*I-cut and uncut pKN800 plasmid DNA:** lane 1, 1000-base pair (bp) molecular weight standard; lane 2, pKN800 plasmid DNA cut with *Pst*I; lane 3, pKN800 plasmid DNA not cut with *Pst*I.

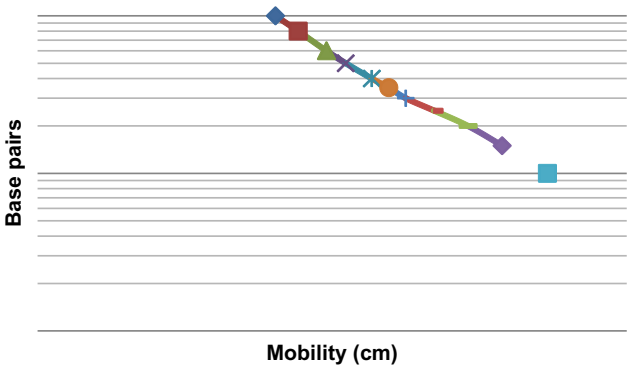


FIGURE A1.2 **Standard curve of 1000-bp DNA molecular weight standard.**

To estimate the size of the restriction fragments cut by *Pst*I, we constructed a standard curve of the 1000-bp DNA molecular weight standard and used it to estimate the size of the restriction fragments (Fig. A1.2). We compared the migration distances of the bands of lane two with the 1000-bp DNA molecular weight standards (Fig. A1.1). The migration distances of the restriction fragments of *Pst*I-cut pKN800 plasmid were 2.05 cm and 3.1 cm. *Pst*I digestion of pKN800 plasmid DNA produced two restriction fragments of 10,500 and 3,500 bp. We did not estimate the size of the uncut pKN800 plasmid DNA as we only used this sample to check whether *Pst*I digested the plasmid completely.

To confirm the function of pKN800 in the production of luciferase protein, we transformed pKN800 plasmid DNA into competent *E. coli* DH5α and observed luciferase reporter gene expression in ampicillin-resistant transformants (Table A1.1).

We calculated the efficiency of transformation to determine whether the transformation was efficient. To calculate the efficiency of transformation, we used the following formula:

$$\frac{\text{The number of ampicillin} - \text{resistant and luminescent transformants}}{\text{The amount of DNA used to transform}}$$

The number of ampicillin-resistant and luminescent transformants from $10^{-1}$ diluted culture transformed with the uncut pKN800 plasmid DNA was insignificant, so the efficiency of transformation only used the data from the undiluted culture transformed with the uncut pKN800 plasmid DNA. We estimated the amount of the uncut pKN800 plasmid DNA by comparing it with the intensity of the bands of the 1000-bp molecular weight standard (Fig. A1.3).

The amount of pKN800 plasmid DNA was 0.025 μg. The efficiency of transformation was

$$\frac{172 \text{ units}}{0.025\ \mu g} = 6880 \text{ transformants per microgram of plasmid DNA}$$

## Discussion

Plasmid pKN800 contains the *lux* operon, which contains α and β subunits that are responsible for luminescence (1). **[[Professor's Comment:** Nice! You may be the only student to cite more than the lab manual.**]]** Insertion of the *lux* operon into the plasmid

TABLE A1.1 The Average Number of Ampicillin-Resistant and Luminescent *E. coli* DH5α Transformants in Each Transformation

| Plate | Type of Medium | Number of Ampicillin-Resistant and Luminescent Transformants (units[b]) | Number of Colonies |
|---|---|---|---|
| Undiluted culture with uncut pKN800 plasmid DNA | | 172[c] | 178 |
| $10^{-1}$ diluted culture with uncut pKN800 plasmid DNA | LB + Amp agar[a] | Too few to count[d] | Too few to count[d] |
| Undiluted culture with *Pst*I-cut pKN800 plasmid DNA (positive control) | | Too few to count[e] | Too few to count[f] |
| Undiluted culture with no pKN800 plasmid DNA (negative control) | | 0 | 0 |
| $10^{-5}$ diluted culture with no pKN800 plasmid DNA | LB agar | 0 | Too numerous to count[g] |
| $10^{-6}$ diluted culture with no pKN800 plasmid DNA | | 0 | 221 |

[a]*This medium contained LB and ampicillin.*
[b]*One unit is one colony.*
[c]*The average was 171.5 units.*
[d]*The average numbers of transformants and colonies were both 10.5 units.*
[e]*The average number of transformants was 1.5 units.*
[f]*The average number of transformants was 4.5 units.*
[g]*The average number of transformants was 274 units.*

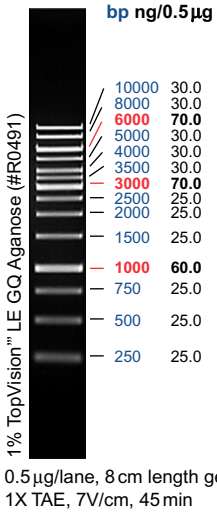


FIGURE A1.3 **The 1000-bp molecular weight standard containing the size and intensity of each band.** The bands with red numbers have twice as much DNA as the bands with blue numbers. The right column is the amount of DNA in each band.

cloning vector pBR322 confers bioluminescence and disrupts tetracycline resistance. A pKN800 plasmid DNA can have one of the two orientations. pKN800-A contains the *lux* operon in the same direction as the *tet* gene, and pKN800-B contains the *lux* operon and the *tet* gene in the opposite direction (2). To investigate whether the orientation of the *lux* operon in pKN800 plasmid affects light production, we purified the plasmid, determined the orientation of the plasmid, and transformed another *E. coli* strain to screen for ampicillin-resistant and luminescent transformants.

*Pst*I digestion of pKN800 plasmid DNA produces two fragments, and the size of each restriction fragment depends on the orientation of the *lux* operon on pKN800 plasmid DNA. To determine the orientation of the *lux* operon on the pKN800 plasmid, we compared the sizes of the restriction fragments with Fig. 1.2 on p. 35 of the laboratory manual (2). According to Fig. 1.2, *Pst*I-cut pKN800-A plasmid DNA produces restriction fragments of 11,540 bp and 1,400 bp; *Pst*I-cut pKN800-B plasmid DNA produces restriction fragments of 9,980 bp and 3,380 bp. Since the result matches closer to the theoretical sizes of the restriction fragments of *Pst*I-cut pKN800-B plasmid, we concluded that the orientation of the *lux* operon in the unspecified pKN800 plasmid is pKN800-B (antisense). *Pst*I digested all pKN800 plasmids completely because there were no unexpected bands on the agarose gel.

Transformation of *E. colii* **[[Professor's Comment:** Spelling.**]]** DH5α with pKN800 plasmid DNA helps us to confirm that pKN800 plasmid DNA encodes functional luciferase proteins. Transformation of pKN800 plasmid DNA into *E. coli* DH5α was successful as we observed growth of ampicillin-resistant and luminescent transformants on LB + Amp agar plates. The culture transformed with uncut pKN800 plasmid DNA contained genes coding for both ampicillin resistance and luminescent ability. The high efficiency **[[Professor's Comment:** In fact, the transformation efficiency you achieved is very low compared to that obtained with the very best cells (~10E9/μg).**]]** of the transformation indicates that the transformation of *E. colii* **[[Professor's Comment:** Spelling.**]]** DH5α with pKN800 DNA was efficient.

The positive control **[[Professor's Comment:** This is a type of negative control, not a positive one.**]]** indicates that *Pst*I successfully cut the pKN800 plasmid DNA as we observed no significant growth. (The few colonies in the positive control indicate that some uncut pKN800 plasmid DNA may be still present. However, the number of transformants is **[[Professor's Comment:** Wrong tense.**]]** too few to count and therefore insignificant.) The culture transformed with *Pst*I-cut pKN800 plasmid DNA cannot grow on LB + Amp agar plates because exonuclease V in *E. coli* destroys linear *Pst*I-cut pKN800 plasmid DNA when the linear DNA enters the bacteria. The culture did not contain cells transformed with the *Ap* gene, which is responsible for ampicillin resistance.

There were **[[Professor's Comment:** This is correct, but see Day's book, page 214, or example 2 in section A of Writing Tips.**]]** viable (ampicillin-sensitive) cells in the cultures as we observed growth on LB agar plates containing cultures that had no cells transformed with the pKN800 plasmid DNA. The negative control ensured that the competent cells are not resistant to ampicillin prior to transformation with pKN800 plasmid DNA as the culture without the pKN800 plasmid DNA did not grow on the LB + Amp plates. The cells in the culture did not contain the *Ap* gene, which is responsible for ampicillin resistance.

The orientation of the *lux* operon on pKN800 is antisense, and we transformed *E. coli* DH5α successfully and efficiently. However, we are not able to determine whether the orientation of the *lux* operon affects the luminescent ability in *E. coli* as all students received the pKN800-B plasmid DNA. We can only determine that the antisense orientation of the *lux* operon on pKN800 plasmid does not affect luminescent ability in *E. coli*.

## Conclusion

The orientation of the pKN800 plasmid DNA is antisense, and the pKN800-B plasmid DNA introduces ampicillin resistance and luminescent ability to *E. coli* DH5α efficiently.

## References

1. Engebrecht J, Nealson K, Silverman M. 1983. Bacterial bioluminescence: isolation and genetic analysis of functions from *Vibrio fischeri*. *Cell* **32**:773–81 **[[Professor's Comment:** Nice!**]]**.
2. Ream W, Geller B, Trempy J, Field K. 2003. *Molecular Microbiology Laboratory Manual*. San Diego, CA: Academic Press; pp. 32–54.

## Questions

**1.** The uncut pKN800 plasmid DNA only formed one band during agarose gel electrophoresis. The band represents the circular supercoiled pKN800 plasmid DNA.

**2.** My pKN800 plasmid DNA is pKN800-B. I calculated the size of the restriction fragments of a *Pst*I-cut pKN800 plasmid DNA based on the information provided on Fig. 1.2 on p. 35 of the laboratory manual (2). The two restriction fragments on pKN800-B in Fig. 1.2 were 9,980 bp and 3,380 bp, which match with the sizes of the restriction fragments of my *Pst*I-cut pKN800 plasmid DNA.

**3.** My *Pst*I digestion of pKN800 plasmid DNA went to completion. We observed no colonies[h] on undiluted *Pst*I-cut DNA on LB + Amp agar plates.

**4.** We got ampicillin-resistant transformants with restricted pKN800 plasmid DNA. We observed colonies on uncut pKN800 plasmid DNA on the LB + Amp plates. If we had no ampicillin-resistant transformants with restricted pKN800 plasmid DNA, we would have no growth on the LB + Amp plates.

**5.** I would *not* expect any of the transformations to yield ampicillin-resistant nonluminescent transformants. If the given pKN800 plasmid DNA contains recircularized cloning vector pBR322 **[[Professor's Comment:** I do not understand how this could happen.**]]**combined with the *lux* operon, the transformants would still be resistant to ampicillin (it still has the *Ap* gene) but could not glow in the dark. -9

**6.** I analyzed the uncut pKN800 plasmid DNA on the agarose gel because I wanted to confirm whether the *Pst*I-cut pKN800 plasmid was cut successfully. The uncut pKN800 plasmid DNA served as a comparison. Also, I wanted to estimate the amount of DNA used to transform *E. coli* by comparing the intensity of the band with the 1000-bp DNA molecular weight standard.

[h]A small number of colonies grew on the plates, but the average number was 4.5 colonies, which was too few to count and insignificant.

7. I plated bacteria that did not receive any plasmid DNA to set a negative control. The negative control on LB + Amp agar plate ensures that the competent cells were not resistant to ampicillin before the transformation with pKN800 DNA, and it also shows that the plates contained enough antibiotic to kill the untransformed cells. We also plated bacteria with no plasmid DNA on LB agar to ensure that the culture contained ampicillin-sensitive cells.
8. I would plate the uncut pKN800 plasmid transformants on LB agars. If we observe no growth on LB agar, then the problem is due to incorrect medium. If the medium is incorrect, the uncut plasmid transformants should still not be able to grow and glow in the dark as they cannot survive on the plates. If we observe colonies on LB agars, then the problem is due to poor transformation efficiency. If the transformation is inefficient, then some of the transformants can still grow on the plate and glow in the dark. Streak a **known ampicillin-resistant strain** on the plates; it should grow. For example, you could streak the strain you used to prepare pKN800 plasmid DNA because those cells were grown in ampicillin. -3

## REPORT 2

**Name:** Connie Lee
**Title**
Identification of probiotics, *L. delbrueckii* and *S. thermophilus*, **[[Professor's Comment:** Use the complete genus name the first time you mention it.**]]** in Yoplait® Original Vanilla using a culture-independent approach
**Date:** January 31, 2012–March 6, 2012

### Purpose

Probiotics in yogurt have many beneficial effects on human health, such as improving human's **[[Professor's Comment:** the**]]** immune system and lowering the risk of cancer and gastrointestinal disorders (1). Identification of probiotics in yogurt may help researchers to indicate which strain is the most efficient one in making yogurt while providing health benefits. Since many bacteria are unculturable, we used a culture-independent approach to identify and study the probiotic community in yogurt, and we compared the result we got **[[Professor's Comment:** Colloquial..**]]**from the culture-dependent approach. Using the culture-independent approach, we identified the probiotic community in yogurt by purifying genomic DNA directly from a sample collected for terminal restriction fragment length polymorphism (T-RFLP) analysis and DNA sequence analysis.

### Methods

We collected samples from Yoplait® Original Vanilla yogurt and extracted genomic DNA directly from the uncultured bacterial sample. We amplified 16S rRNA genes by polymerase chain reaction (PCR). We separated the PCR amplicons by agarose gel electrophoresis. We digested the purified PCR products with restriction enzymes and submitted

the digested DNA to the Central Services Lab for capillary electrophoresis. The data from capillary electrophoresis allowed us to perform terminal restriction fragment length polymorphism (T-RFLP) analysis on the PCR amplicons. We also transformed the PCR amplicons into *Escherichia coli* to create a molecular clone library of the 16S rRNA genes present in the PCR product. We sequenced the clones to identify the probiotics detected in T-RFLP analysis.

We followed the procedures on pp. 12–19 of the laboratory manual with the following exceptions (2): Step D-7 on p. 16: We added 50 μL of deionized water into the microcentrifuge tube. Step F-8: We did not re-streak the transformed *E. coli* on a fresh LB agar plate. Step H-1: We received 1 μL of restriction endonucleases +2 μL pf 10x buffer with BSA. Step I-1: For our group, we also sequenced additional clones from the agar plate.

## Results

Identification of probiotic community in yogurt may help researchers to indicate which strain is the most efficient one in making yogurt while providing health benefits. However, since many bacteria are currently unculturable in the laboratory, we used a culture-independent approach to identify the probiotic community in yogurt. **[[Professor's Comment:** You already said this.**]]** To collect DNA from the probiotic community in yogurt, we purified genomic DNA directly from the uncultured sample collected from yogurt. We observed the diversity of bacteria present in the yogurt by analyzing restriction-fragment-length polymorphisms (RFLPs) among the fluorescently labeled terminal fragments of the extracted **[[Professor's Comment:** amplified.**]]** DNA. To identify the bacteria detected in T-RFLP analysis, we created a molecular clone library and sequenced the clones.

To determine the most abundant probiotic in yogurt, we collected samples from Yoplait® Original Vanilla yogurt. Yoplait® Original Vanilla was a wet sample source. **[[Professor's Comment:** Unnecessary.**]]** We opened the seal of the yogurt bottle immediately before we collected the samples.

To determine the purity, size, and concentration of the purified PCR product, we separated the PCR products by agarose gel electrophoresis. Lane one had three bands, which were the low DNA mass ladder. Lane two had one band with smearing. Lane three was the no-template negative control, and the sample produced no band (Fig. A2.1).

To estimate the size of the purified PCR product, we constructed a standard curve of the low DNA mass ladder molecular weight standards (Figs. A2.2 and A2.3). To estimate the size of the PCR product, **[[Professor's Comment:** You just said this.**]]** we measured the migration distance of the band. We compared the migration distance of the band in lane two with the bands of the low DNA mass ladder. The band of lane two had a migration distance of 11.7 cm (Fig. A2.1). Extrapolating from the standard curve, we estimated that the band in lane two was 942 base pairs (Fig. A2.3).

To estimate the concentration of the purified PCR product, we estimated the intensity of the band in lane two relative to the molecular weight standards. The PCR amplicons had a concentration of 5 ng/μL. **[[Professor's Comment:** How did you calculate this number. The band on your gel is slightly stronger than the 80-ng standard, and your band is much weaker than the 120-ng standard. So, you have ~90 ng/5 ul loaded in the

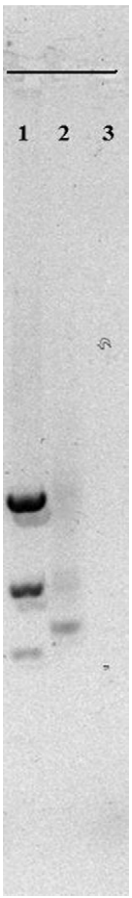


FIGURE A2.1 **Agarose gel electrophoresis of PCR products.** Lane one contains DNA mass ladder (Invitrogen). Lane two contains purified PCR products. Lane three contains a sample from the no-template PCR reaction.

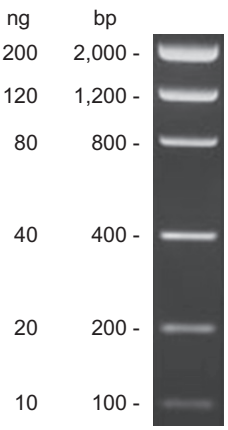


FIGURE A2.2 **Low DNA mass ladder (Invitrogen) molecular weight standards on a 2% 3:1 agarose gel.** Each lane contains 4 μL.

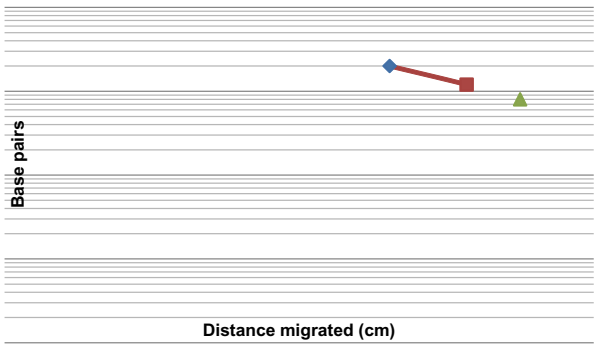


FIGURE A2.3 **Standard curve of the low DNA mass ladder molecular weight standards.**

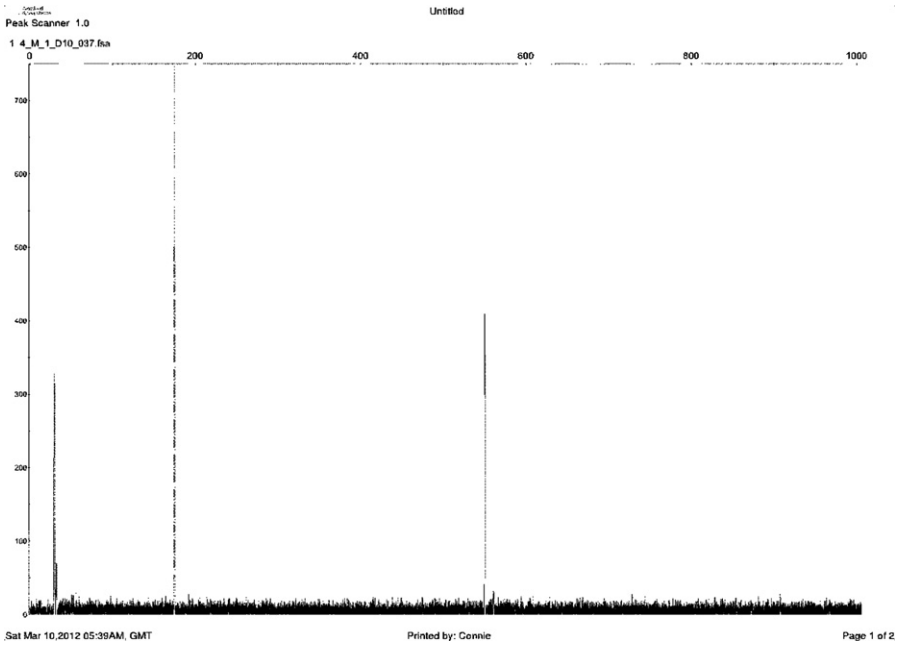


FIGURE A2.4 **T-RFLP profile of *Msp*I-cut PCR amplicons of the bacterial population from Yoplait® Original Vanilla yogurt.** We cut the purified PCR amplicons with *Msp*I restriction enzyme and the submitted sample **[[Professor's Comment:** Unnecessary.**]]** was in a dilution factor of **[[Professor's Comment:** Awkward.**]]** 20. The x-axis is the fragment length (nucleotides), and the y-axis is the fluorescence intensity (height of the peak).

lane = 18 ng/ul.**]]** The purified PCR product's concentration measured by the nanodrop spectrophotometer was 38 ng/μL.

To observe the complexity of the bacterial population in yogurt, we submitted our purified PCR products for T-RFLP analysis. The *Msp*I restriction endonuclease produced four **[[Professor's Comment:** I see only 3 peaks.**]]** major peaks, indicating that there were at least four different bacterial species in the sample (Fig. A2.4). The *Rsa*I endonuclease produced three major peaks, indicating there were at least three different bacterial species

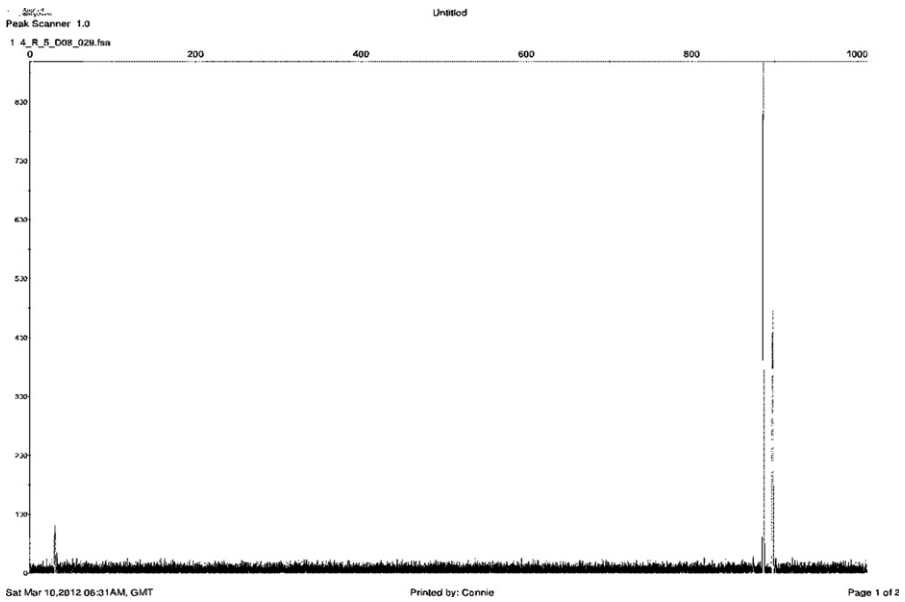


FIGURE A2.5 **T-RFLP profile of *Rsa*I-cut PCR amplicons of the bacterial population from Yoplait® Original Vanilla yogurt.** We cut the purified PCR amplicons with *Rsa*I restriction enzyme and the submitted sample **[[Professor's Comment:** Unnecessary.**]]** was in a dilution factor of **[[Professor's Comment:** Awkward.**]]** four. The x-axis is the fragment length (nucleotides), and the y-axis is the fluorescence intensity (height of the peak).

TABLE A2.1 Length and Intensity of the *Msp*I-Digested DNA Fragments of the Bacterial Population in Yoplait® Original Vanilla Yogurt

| Size of the Fragment (base pairs) | Intensity of the Fragment (height of the peak) |
|---|---|
| 30 | 329 |
| 32 | 68 |
| 174 | 750 |
| 550 | 410 |

The submitted sample was in a dilution factor of **[[Professor's Comment:** Awkward.**]]** 20.

in the sample (Fig. A2.5). The *Msp*I and *Rsa*I restriction endonucleases did not produce the same number of major peaks. **[[Professor's Comment:** I disagree.**]]** Analyzing the T-RFLP files, we obtained the length and relative quantity of each labeled terminal restriction fragment digested by either *Msp*I or *Rsa*I restriction endonucleases (Tables A2.1 and A2.2).

To identify the species associated with particular T-RFLP peaks, we constructed a molecular clone library and submitted the clones for DNA sequence analysis. To identify

TABLE A2.2 Length and Intensity of the *Rsa*I-Digested DNA Fragments of the Bacterial Population in Yoplait® Original Vanilla Yogurt

| Size of the Fragment (base pairs) | Intensity of the Fragment (height of the peak) |
|---|---|
| 30 | 78 |
| 886 | 869 |
| 897 | 445 |

The submitted sample was in a dilution factor of **[[Professor's Comment:** Awkward.**]]** four.

TABLE A2.3 Number of $Lac^-$ and $Lac^+$ Colonies in the Clone Library Made from T-RFLP Amplicons

| Concentration of the Transformants | Number of $Lac^-$ (white) Colonies | Number of $Lac^+$ (blue) Colonies |
|---|---|---|
| Original (not concentrated) transformed *E. coli* cells | Too numerous to count | 10 |
| Concentrated transformed *E. coli* cells | Too numerous to count | 20 |

The concentration of the transformed cells spread on the agar plate was not measurable. $Lac^-$ colonies were white because they did not produce active β-galactosidase to cleave ×-gal due to the inserted DNA. $Lac^+$ colonies were blue because they produced active β-galactosidase to cleave ×-gal.

the transformants with vectors that have DNA inserted, we observed the effect of the inserted DNA on *Lac*Z gene expression in ampicillin-resistant transformants (Table A2.3).

We sequenced the 16S rRNA gene to identify the bacteria corresponding to particular T-RFLP peaks. We indicated the portions of each sequence corresponding to the plasmid vector, PCR primers, and amplified 16S rRNA gene (Figs. A2.6 and A2.7). Appendix A shows the additional sequences of the molecular clone library. The bacteria were *Streptococcus thermophilus* and *Lactobacillus delbrueckii*. In a total of 22 DNA sequences, 16 sequences were *S. thermophilus*. Five out of the 22 sequences **[[Professor's Comment:** 16 + 5 = 21 What was #22?**]]** were *L. delbrueckii*. All sequences had 99–100% identity match in the BLAST search.

To determine the length of the corresponding T-RFLP fragments for each enzyme, we located the *Rsa*I (GTAC) and *Msp*I (CCGG) restriction sites in the 16S rRNA gene sequence (Figs. A2.6 and A2.7 and Appendix B). The *Msp*I-cut DNA fragment of *S. thermophilus* contained 555 nucleotides. The *Rsa*I-cut DNA fragments of *S. thermophilus* could contain either 32–34 nucleotides or 891 nucleotides. The *Msp*I-cut DNA fragment of *L. delbrueckii* had 179 nucleotides. The *Rsa*I-cut DNA fragments of *L. delbrueckii* had 905 nucleotides. Using the length of the fragments cut by each restriction enzymes, we identified the bacteria corresponding to particular T-RFLP peaks (Figs. A2.8 and A2.9). On the T-RFLP profile of *Rsa*I-cut PCR amplicons, both *S. thermophilus* and *L. delbrueckii* could correspond to the peak with the size of 897 base pairs (Fig. A2.9).

We used the Plymouth Routines in Multivariate Ecological Research (PRIMER) program to compare the bacterial population in the samples collected from yogurt and human saliva

>04A_4_M13R-27 TRIM QUALITY: 20
GTCCTGCAGGTttaACGAATTCGCCCTTCCgtaATTCCTTTGAGTTTCAACCTTGCGGTCGTAC
TCCCCAGGCGGAGTGCTTAATGCGTTTGCTGCGGCACTGAGGACCGGAAAGTCCCCAACACCTA
GCGCTCATCGTTTACGGCATGGACTACCAGGGTATCTAATCCTGTTCGCTCCCCACGCTTTCGA
GCCTCAGCGTCAGTTACAGACCAGAGAGCCGCTTTCGCCACCGGTGTTCCTCCATATATCTACG
CATTTCACCGCTACACATGGAATTCCACTCTCCCCTTCTGCACTCAAGTTTGACAGTTTCCAAA
GCGAACTATGGTTGAGCCACAGCCTTTAACTTCAGACTTATCAAACCGCCTGCGCTCGCTTTAC
GCCCAATAAATCCGGACAACGCTCGGGACCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGC
CGTCCCTTTCTGGTAAGCTACCGTCACAGTGTGAACTTTCCACTCTCACACCCGTTCTTGACTT
ACAACAGAGCTTTACGATCCGAAAACCTTCTTCACTCACGCGGCGTTGCTCGGTCAGGGTTGCC
CCCATTGCCGAAGATTCCCTACTGCTGCCTCCCGTAGGAGTCTGGGCCGTGTCTCAGTCCCAGT
GTGGCCGATCACCCTCTCAGGTCGGCTATGTATCGTCGCCTAGGTGAGCCATTACCTCACCTAC
TAGCTAATACAACGCAGGTCCATCTTGTAGTGGAGCAATTGCCCCTTTCAAATAAATGACATGT
GTCATCCATTGTTATGCGGTATTAGCTATCGTTTCCAATAGTTATCCCCCGCTACAAGGCAggT
TACCTACGCGTTACTCACCCGTTCCGCAACTCATCCAagaaCagCaaCGTCCTCTCTTCACCCTT
CTA

FIGURE A2.6 ***S. thermophilus*'s DNA sequence from the M13R primer.** Underlined nucleotides are part of the vector sequence. Blue nucleotides are 926-907R primer used in PCR amplification. Green nucleotides are the *Rsa*I restriction site. Pink nucleotides are the *Msp*I restriction site. Crossed-off nucleotides are ambiguities in the sequence, and we removed them before the BLAST search. Bold and italicized bases are the sequence we used for the BLAST search.

>04A_4_T7Prom TRIM QUALITY: 20
AGAGTTTGATCATGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTG
AAGAGAGGAGCTTGCTCTTCTTGGATGAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCC
TTGTAGCGGGGGATAACTATTGGAAACGATAGCTAATACCGCATAACAATGGATGACACATGTC
ATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTTGTATTAGCTAGTAGGTG
AGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACTGGGA
CTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACC
CTGACCGAGCAACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAA
GAACGGGTGTGAGAGTGGAAAGTTCACACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACT
ACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGAGCGTTGTCCGGATTTATTGGGCGTAAAGC
GAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACCATAGTTCGCTTTGGAA
ACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGAT
ATATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTGaggCTCGAAAG
CGTGGGGAGCGAACAGGATTAGATACCCTGGTAGTCCACGCCGTaaacgATGCACGCTAGGTGT
TCGGCACTTTCGGGTCGTCAgtgccgcagcAAACGCATTAAGCACT

FIGURE A2.7 ***S. thermophilus*'s DNA sequence from the T7 promoter primer.** In this sequence, we did not observe vector sequence. Yellow nucleotides are the 8-27F primer used in PCR amplification. Red nucleotides are the rest of the sequence of 8-27F primer that are not present in the sequence. There is no *Rsa*I restriction site. Pink nucleotides are the *Msp*I restriction site. Crossed-off nucleotides are ambiguities in the sequence, and we removed them before the BLAST search. Bold and italicized bases are the sequence we used for the BLAST search.

(Table A2.4).. **[[Professor's Comment:** One period will do.**]]** We transformed the data by finding the square root of the data. **[[Professor's Comment:** Redundant**]]** Samples with similar populations lie closer together than samples with dissimilar populations (Fig. A2.10). The bacterial populations from yogurt samples lied **[[Professor's Comment: lay.]]** close to

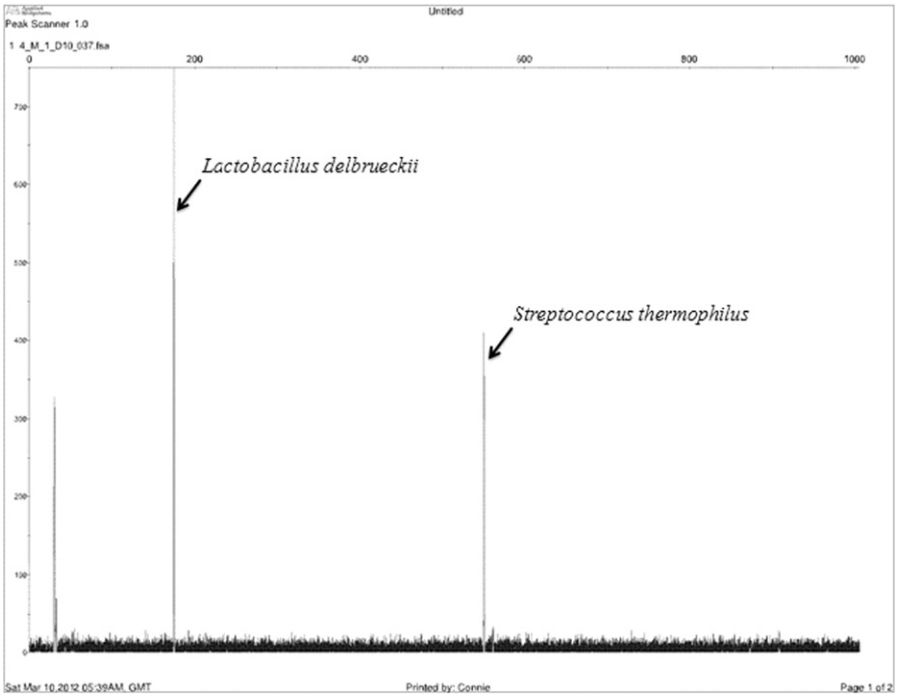


FIGURE A2.8 **T-RFLP profile of *Msp*I-cut PCR amplicons of identified bacterial population from Yoplait® Original Vanilla yogurt.** We cut the purified PCR amplicons with *Msp*I restriction enzyme and the submitted sample was in a dilution factor of 20. **[[Professor's Comment:** Unnecessary.**]]** The x-axis is the fragment length (nucleotides), and the y-axis is the fluorescence intensity (height of the peak).

each other. However, the bacterial populations sampled from human saliva lied **[[Professor's Comment:** lay or were.**]]** far away from each other.

## Discussion

Probiotics can help lower cholesterol level, reduce risk of cancer and gastrointestinal diseases, and improve the immune system (1). Probiotics are active ingredients that make yogurt a healthy food. Identification of probiotic populations in yogurt may help researchers to indicate which strain is the most efficient one in making yogurt while providing health benefits. Since many bacteria are unculturable in laboratory, we used a culture-independent approach to study the probiotic population in yogurt. To identify and study the probiotic population and determine the most abundant probiotic in yogurt, we collected samples from Yoplait® Original Vanilla, extracted genomic DNA directly from the sample for PCR amplification, purified PCR amplicons for agarose gel electrophoresis, digested PCR amplicons with restriction enzymes for T-RFLP analysis, and transformed competent *E. coli* cells with PCR amplicons to construct a clone library to identify T-RFLP peaks. **[[Professor's Comment:** This is the third time you said this.**]]**

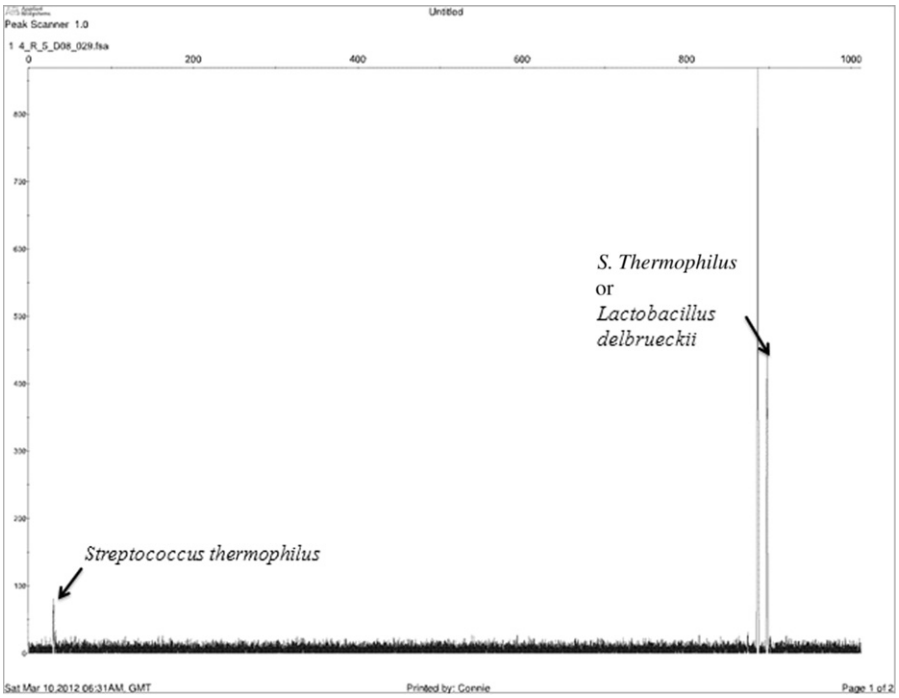


FIGURE A2.9 **T-RFLP profile of *Rsa*I-cut PCR amplicons of identified bacterial population from Yoplait® Original Vanilla yogurt.** We cut the purified PCR amplicons with *Rsa*I restriction enzyme and the submitted sample was in a dilution factor of four. **[[Professor's Comment:** Unnecessary.**]]** The x-axis is the fragment length (nucleotides), and the y-axis is the fluorescence intensity (height of the peak). **[[Professor's Comment:** You should attribute each peak to a particular species. Do not capitalize thermophilus.**]]**

The PCR amplification of the genomic DNA of the uncultured sample from Yoplait® Original Vanilla was successful. The agarose gel electrophoresis detected PCR product, indicating that the PCR reagents successfully reacted and amplified the DNA. The size of the band was 942 base pairs, confirming that the PCR amplicons had the anticipated length of 880 to 940 base pairs. The no-template negative control produced no band, indicating that the PCR reagents were pure.

The concentration of the PCR amplicons was inconsistent between two measurements. The concentration of PCR product estimated by agarose gel electrophoresis was lower than the concentration of PCR product measured by the spectrophotometer. Agarose gel electrophoresis can only give an estimation of the concentration of the PCR product. Instead of physically measuring the amount of product present in each band, we only estimated the intensities of the bands relative to the molecular weight standards. A nanodrop spectrophotometer **[[Professor's Comment:** This measured the signal in the smear as well as the band.**]]** gives a more accurate measurement **[[Professor's Comment:** I disagree.**]]** of the concentration of the PCR product.

To study the complexity of the bacterial population in yogurt, we submitted digested PCR amplicons for analysis by capillary electrophoresis and obtained T-RFLP data. The

TABLE A2.4 Bacterial Genera in Yogurt and Human Saliva

| Species Abundance Cut by Mspl | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **Genus** | **Seat 1** | **Seat 2** | **Seat 3** | **Seat 4** | **Seat 5** | **Seat 6** | **Seat 001** | **Seat 005** |
| Lactobacillus | 75 | 12.5 | 8.69 | 5 | 16.7 | 45.8 | 0 | 0 |
| Streptococcus | 25 | 87.5 | 91.3 | 16 | 79.2 | 54.2 | 34.43 | 46.5 |
| Anoxybacillus | 0 | 0 | 0 | 0 | 4.17 | 0 | 0 | 0 |
| Rothia | 0 | 0 | 0 | 0 | 0 | 0 | 4.34 | 0 |
| Veillonella | 0 | 0 | 0 | 0 | 0 | 0 | 17.39 | 0 |
| Lachnospiraceae | 0 | 0 | 0 | 0 | 0 | 0 | 4.35 | 0 |
| Campylobacter | 0 | 0 | 0 | 0 | 0 | 0 | 8.69 | 0 |
| Haemophilus | 0 | 0 | 0 | 0 | 0 | 0 | 2.17 | 25.6 |
| Capnocytophaga | 0 | 0 | 0 | 0 | 0 | 0 | 2.17 | 2.32 |
| Fusobacterium | 0 | 0 | 0 | 0 | 0 | 0 | 2.17 | 2.32 |
| Prevotella | 0 | 0 | 0 | 0 | 0 | 0 | 8.69 | 0 |
| Porphytomonas | 0 | 0 | 0 | 0 | 0 | 0 | 4.35 | 0 |
| Gemella | 0 | 0 | 0 | 0 | 0 | 0 | 2.17 | 0 |
| Catonella | 0 | 0 | 0 | 0 | 0 | 0 | 2.17 | 0 |
| Kingella | 0 | 0 | 0 | 0 | 0 | 0 | 2.17 | 0 |
| Lautropia | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 4.65 |
| Neisseria | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 16.3 |
| Actinomyces | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 2.32 |
| Source | Yogurt | Yogurt | Yogurt | Yogurt | Yogurt | Yogurt | Human Saliva | Human Saliva |

Numbers indicate the percent of the bacterial population comprised of each genus. The "source" row indicates the type of sources of each sample. The yogurt used by seats 1 and 2 was Tillamook® Vanilla. The yogurt used by seat 3 was Kroger® Plain. The yogurt used by seats 4, 5, and 6 was Yoplait® Original Vanilla. Seats 001 and 005 were sampled from human saliva.

T-RFLP data indicated that there are at least three or four bacteria present in the yogurt. However, more than one species can produce a particular fragment, so each T-RFLP peak may represent more than one species. Although T-RFLP can only give a minimal estimation of the number of species present in the sample, T-RFLP analysis is still useful in observing the complexity of the bacterial community in the yogurt.

The culture-independent approach in this experiment detected more species than the culture-dependent approach in experiment 3, which only detected one species. The

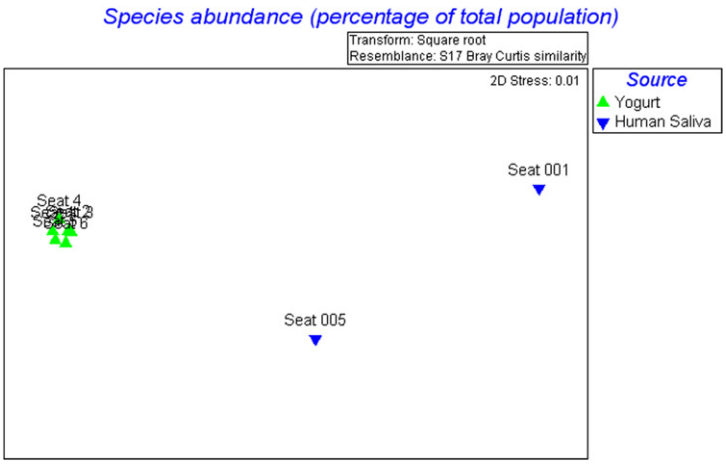


FIGURE A2.10 **Comparison of species abundance of bacterial genera in yogurts and human saliva.** The yogurt used by seats 1 and 2 was Tillamook® Vanilla. The yogurt used by seat 3 was Kroger® Plain. The yogurt used by seats 4, 5, and 6 was Yoplait® Original Vanilla. Seats 001 and 005 were sampled from human saliva.

difference in the number of species detected by the two approaches confirms that many bacteria are unculturable **[[Professor's Comment:** The species you found grow in the laboratory, but they do not grow on LB agar.**]]** in the laboratory. The culture-independent approach ensures that the analysis includes the genomic DNA of all bacteria present in the sample. The culture-independent approach is the more appropriate method for studying a bacterial community in an environment.

Transformation of *E. coli* with PCR amplicons helps us construct a molecular clone library of individual DNA molecules from the PCR product. Transformation of plasmid vector with DNA inserted into *E. coli* was successful as we observed growth of ampicillin-resistant transformants that had the $Lac^-$ phenotypes. The transformants with DNA inserted into the plasmid vector had the $Lac^-$ phenotypes because they lost the gene coding for β-galactosidase, so the $Lac^-$ transformants could not cleave X-gal and remained white.

DNA sequences of the $Lac^-$ transformants help us identify species associated with particular T-RFLP peaks. We cannot sequence the PCR product directly because we amplified the DNA from a mixed template. The BLAST search of the DNA sequences indicates that *S. thermophilus* and *L. delbrueckii* were present in the bacterial population sampled from the yogurt. We successfully indicated the species associated with particular T-RFLP peaks. However, the predicted length of the fragments did not perfectly correspond to the T-RFLP peaks. The predicted length of the fragments and T-RFLP peaks had a difference of three to six nucleotides. The differences may be due to the limitation of T-RFLP analysis. T-RFLP estimates the size of each labeled restriction fragment, and these estimations may differ from the true length by several nucleotides (2).

T-RFLP and DNA sequences help indicate the relative quantity of each species in the sample. However, we cannot determine the most abundant probiotic in the yogurt due to inconsistency in the data. The T-RFLP plot indicated that there were more *L. delbrueckii* in the yogurt, but the molecular clone library indicated that there were more *S. thermophilus* in the yogurt. The inconsistency may be due to incorrectly labeled PCR amplicon,

unsuccessfully digested amplicons, mutation in PCR or DNA sequencing primers, **[[Professor's Comment:** No. Unlikely.**]]** or the limitation of T-RFLP analysis. Also, since we only randomly chose colonies for DNA sequences, the inconsistency may be due to the probability of choosing a species more than the other species. Although we successfully observed the complexity of the bacterial population in the yogurt, we cannot determine the most abundant probiotic in the yogurt.

The presence of *L. delbrueckii* and *S. thermophilus* in the yogurt is logical as *L. delbrueckii* and *S. thermophilus* are the most commonly used probiotics in yogurt production (1). Compared to experiment 3, which only found one species in the sample, the culture-independent approach successfully included unculturable bacteria in the analysis. Also, since we did not detect *Bacillis pumilus* or *Bacillus safensis* in the culture-independent approach, it further confirms that the presence of *Bacillis pumilus* or *Bacillus safensis* in the sample in experiment 3 was due to contamination during the experiment.

The NMDS plot helped us compare the bacterial populations in two different environments. The bacterial populations sampled from yogurt were close to each other, indicating that the bacterial populations in yogurts are similar. However, the bacterial populations sampled from human saliva were far from each other, indicating that the bacterial populations in human saliva did not resemble each other. We attributed the differences of the bacterial populations in human saliva to different food consumed by the students or the health condition of the students. Similar environments had similar bacterial populations, and different environments had dissimilar bacterial populations.

## Conclusions

We collected and sampled bacteria from Yoplait® Original Vanilla. The PCR amplification was successful as the agarose gel electrophoresis detected PCR products. T-RFLP profile and DNA sequences indicated that *L. delbrueckii* and *S. thermophilus* were present in the yogurt. Since there were inconsistencies between T-RFLP data and DNA sequences, we could not determine the most abundant probiotic in the yogurt. The presence of *L. delbrueckii* and *S. thermophilus* in yogurt was logical because these bacteria are the most commonly used probiotics in yogurt production. The culture-independent approach successfully includes unculturable bacteria in the analysis.

## References

1. Ataie-Jafari A, Larijani B, Majdb H, Tahbaz F. 2009. Cholesterol-Lowering Effect of Probiotic Yogurt in Comparison with Ordinary Yogurt in Mildly to Moderately Hypercholesterolemic Subjects. *Ann Nutr Metab.* **54**:22–7.
2. Ream W, Geller B, Trempy J, Field K. 2003. *Experiment 4—Terminal Restriction Fragment Length Polymorphism Analysis of 16s Ribosomal RNA Genes from Uncultured Bacterial Communities*. San Diego, CA: Academic Press.

## Questions

**1.** I detected more bacterial species using a culture-independent approach. In the culture-independent approach, we detected at least three to four species present in the sample collected from the yogurt. However, in the culture-dependent approach, we only detected one species. The culture-independent approach allows us to extract

bacterial genomic DNA directly from uncultured sample. Therefore, we were able to do analysis on the mixed DNA sample that contains genomic DNA from both culturable and unculturable bacterial species.

2. Yes, the approach used in experiment 4 can detect dead bacteria. Since we extracted DNA directly from the sample rather than from a culture, we will get DNA from both live and dead bacteria as long as the bacteria are not lysed. Also, PCR amplification amplifies both dead and live bacteria. Since we used PCR amplicons for our T-RFLP analysis, the T-RFLP profile also cannot distinguish dead and live bacteria. It means that T-RFLP peaks on the plot can indicate bacteria that are either dead or alive.
3. No. A T-RFLP profile only gives minimal reflection of the bacterial population present in a sample. More than one species can produce a particular fragment, meaning that a single T-RFLP fragment detected can actually be fragments from more than one species. Without DNA sequences, a T-RFLP profile only indicates the complexity of the bacterial population but cannot accurately reflect the bacterial population in a sample.
4. I selected a $Lac^-$ (white) colony from my clone library for sequence analysis. $Lac^-$ (white) colonies were transformants that contained amplicon DNA inserted into the plasmid vector. We used the plasmid (pCR4-TOPO) to transform competent *E. coli* cells. The plasmid contains genes for ampicillin resistance, providing a selection for transformed cells. However, some transformed cells may contain empty vectors, and some may contain plasmid vectors with DNA inserted. Transformed cells that contained empty vectors still had intact *lacZ* $\alpha$ domain, so they can produce active $\beta$-galactosidase inside the transformed *E. coli* cells. Thus, transformed cells that contained empty vectors were blue because they can cleave X-gal. On the other hand, transformed cells with DNA inserted into the vector did not have an intact *lacZ* $\alpha$ domain, so they cannot product active $\beta$-galactosidase inside the transformed *E. coli* cells. Thus, transformed cells with DNA inserted into the vector remained white because they cannot cleave X-gal.
5. No. More than one species can produce a particular fragment, so T-RFLP fragment lengths alone cannot identify a species. We need to construct a molecular clone library and sequence the clones in order to identify the species that correspond to particular T-RFLP peaks.

## Appendix A.A

## The additional DNA sequences of the plasmid DNA from T7 promoter primer

Note: In all sequences, we did not observe vector sequence. Yellow nucleotides are 8-27F primer used in PCR amplification. Green nucleotides are *Rsa*I restriction site. Pink nucleotides are *Msp*I restriction site. Crossed-off nucleotides are ambiguities in the sequence, and we removed ambiguities before the BLAST search. Bold and italicized bases are the sequence we used for the BLAST search.

Sequence 1 — *L. delbrueckii*

TCCTGGCTCAG*GACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGAGCTGAATTCAAAGA~~tccCTTcgGG~~*
*~~GTGATTTGTTGGACGCTAGCGGCGGATGGGTGAGTAACacgTGGGCAATCTGCCCTAAAGACTGGGATACCACTTGG~~*
*~~AAACAGGTGCTAATA~~CCGC~~ATAACAACATGAATCGCATGATTCaAGTTTGAAAGGCGGCGCAagCTGTCACTTTaGG~~*
*~~ATGAGCCCGCGGCGCATTagcTAGTTGGTGGGGTAAAGGCCTACCAAGGCAATGATGCGtagccgagtTGAGAGACT~~*
*~~GATCGGccaCATTGGGACTGagaCACGGCCCAAACTCCTACGGGAGGCagcagTAGGGAATCTTCCACAATGGAcgc~~*
*~~AAGTCTGATGGAGCAACGCCGCGTGAGTGAAGAAGGTTTTcggATCGTAAAGCTCTGTTGTTGGTGAagaAGGATAG~~*
*~~AGGCAGTAACTGGTCTTTATTTGACGGTAATCAACcataaaGTCACGGCTAACTACGTGCCagcagCCGCGGTAATA~~*
*~~CGTaggTGGCAAGCGTTGT~~CCGG*~~ATTTATTGGGCGTAAAGCGAGCGCAggcggAATGATAAGTCTGATGTGAAAGCC~~*
*~~CACGGCTCAACCGTGGAACTGCATcggaAACTGTCATTCTTGagtGcaGaataggAGAGTGGAACTCCATGTGTAGc~~*
*~~ggtGGAATGCGTAcaTATATGGAAGaacACCAGtggCgaacgcgGCTCTCTGGTCTGCaACTGACGC~~

Sequence 2 — *S. thermophilus*

TCCTGGCTCAG*GACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTCTT*
*GGATGAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGCTA*
*ATACCGCATAACAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTTG*
*TATTAGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACTG*
*GGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAGC*
*AACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAAGT*
*TCACACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGAG*
*CGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACCA*
*TAGTTCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAG*
*ATATATGGAAGAACACCAGTGGCGAAGGCGGCTCTCTGGTCTGCAACTGACGCTG~~aggCTCGAAAGCGtgngGaGCG~~*
*~~AACAGGATTAGATACCCTGGTAGTCCACGCCGTaaaCGATGAGTGCTAGGTGTTGGATCCTTT~~CCGG~~GATTCAGTGC~~

Sequence 3 — *S. thermophilus*

*~~Gtca~~TTCCTTTGAGTTTCAACCTTGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTGCGGCACTGAATC*
*CCGGAAAGGATCCAACACCTAGCACTCATCGTTTACGGCGTGGACTACCAGGGTATCTAATCCTGTTCGCTCCCCAC*
*GCTTTCGAGCCTCAGCGTCAGTTACAGACCAGAGAGCCGCTTTCGCCACCGGTGTTCCTCCATATATCTACGCATTT*
*CACCGCTACACATGGAATTCCACTCTCCCCTTCTGCACTCAAGTTTGACAGTTTCCAAAGCGAACTATGGTTGAGCC*
*ACAGCCTTTAACTTCAGACTTATCAAACCGCCTGCGCTCGCTTTACGCCCAATAAATCCGGACAACGCTCGGGACCT*
*ACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTCCCTTTCTGGTAAGCTACCGTCACAGTGTGAACTTTCCACT*
*CTCACACCCGTTCTTGACTTACAACAGAGCTTTACGATCCGAAAACCTTCTTCACTCACGCGGCGTTGCTCGGTCAG*
*GGTTGCCCCCATTGCCGAAGATTCCCTACTGCTGCCTCCCGTAGGAGTCTGGGCCGTGTCTCAGTCCCAGTGTGGCC*
*GATCACCCTCTCAGGTCGGCTATGTATCGTCGCCTAGGTGAGCCATTACCTCACCTACTAGCTAATACAACGCAggT*
*CCATCTTGTAGTGGAGCAATTGCCCCTTTCAAATAAATGACATGTGTCATCCATTGTTATGCGGTATTAGCTATCGT*
*TTCCAATAGTTATCCCCCGCTACAAGGCAGGTTACCTACGCGTTACTCACCCGTTCGCAACTCATCCA~~agaagagca~~*
*~~agctCCTCTGCttcagcgTtcTACTTGCATGTATTaggcaCgccgccaGCGTTCGTCCTGagcCaTGA~~

Sequence 4 — *S. thermophilus*

GatatGGCTCAG*GACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTCT*
*TGGATGAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGCT*
*AATACCGCATAACAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTT*
*GTATTAGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACT*
*GGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAG*
*CAACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAAG*
*TTCACACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGA*
*GCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACC*
*ATAGTTCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTA*
*GATATATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTG~~aggCTCGAAAGCgtggggaGC~~*
*~~GAACAGGATTAGATACCCTggTAGTCCACGCcGTAAACGATGAGTGCTaggtGTTGGATCCTTT~~CCGG~~GATTcAGTG~~*
*~~CC~~*

Sequence 5 — *S. thermophilus*

ATCATGGCTCAG*GACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTCT*
*TGGATGAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGCT*
*AATACCGCATAACAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTT*

GTATTAGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACT
GGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAG
CAACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAAG
TTCACACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGA
GCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACC
ATAGTTCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTA
GATATATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTGaggCTCGAAAGCGTggngagCG
AACAGGATTAGATACCCTGGTAGTCCACGCCgtaaacgATGAGTGCTAGGTGTTGGATCCTTTCCGGGATTCAgtgC

Sequence 6 — *L. delbrueckii*

GATCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGAGCTGAATTCAAAGATCCCTTCG
GGGTGATTTGTTGGACGCTAGCGGCGGATGGGTGAGTAACACGTGGGCAATCTGCCCTAAAGACTGGGATACCACTT
GGAAACAGGTGCTAATACCGGATAACAACATGAATCGCATGATTCAAGTTTGAAAGGCGGCGTAAGCTGTCACTTTA
GGATGAGCCCGCGGCGCATTAGCTAGTTGGTGGGGTAAAGGCCTACCAAGGCAATGATGCGTAGCCGAGTTGAGAGA
CTGATCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCACAATGGAC
GCAAGTCTGATGGAGCAACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTTGGTGAAGAAGGAT
AGAGGCAGTAACTGGTCTTTATTTGACGGTAATCAACCAGAAAGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAA
TACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGAATGATAAGTCTGATGTGAAAG
CCCACGGCTCAACCGTGGAACTGCATCGGAAACTGTCATTCTTGAGTGCAGAAGAGGAGAGTGGAATTCCATGTGTA
GCGGTGGAATGCGTAGATATATGGAAGAACACCAGTGGCGAaggCGGCTCTCTGGTCTGCAACTGACGCtgaggCTC
GAAAGCATGGGTAGCGAACAGGATTAGaTACCCTGGTAGTCCATGCCGTaaaCGATGAGCGCT

Sequence 7 — *S. thermophilus*

TCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTCTT
GGATGAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGCTA
ATACCGCATAACAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTTG
TATTAGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACTG
GGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAGC
AACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAAGT
TCACACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGAG
CGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACCA
TAGTTCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAG
ATATATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTGAggCTCGAAAGCGTggnGaGCG
AACAGGATTAGATACCCTGGTAGTCCACGCCgtanacgATGAGTGCTaggTGTTGGATCCT

Sequence 8 — *S. thermophilus*

TCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTCTT
GGATGAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGCTA
ATACCGCATAACAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTTG
TATTAGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACTG
GGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAGC
AACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAAGT
TCACACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGAG
CGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACCA
TAGTTCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAG
ATATATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTGaggCTCGAAAGCGTGGGGAGCG
AACAGGATTAGATACCCTGGTAGTCCACGCCgtaaacgATGagtGCTAGGTGTTGGATCCTTTCCGGGATCAGTGCC
GAAGCTAACGC

Sequence 9 — *L. delbrueckii*

GATCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGAGCTGAATTCAAAGATCCCTTCG
GGGTGATTTGTTGGACGCTAGCGGCGGATGGGTGAGTAACACGTGGGCAATCTGCCCTAAAGACTGGGATACCACTT
GGAAACAGGTGCTAATACCGGATAACAACATGAATCGCATGATTCAAGTTTGAAAGGCGGCGCAAGCTGTCACTTTA
GGATGAGCCCGCGGCGCATTAGCTAGTTGGTGGGGTAAAGGCCTACCAAGGCAATGATGCGTAGCCGAGTTGAGAGA
CTGATCGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCACAATGGAC
GCAAGTCTGATGGAGCAACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTTGGTGAAGAAGGAT
AGAGGCAGTAACTGGTCTTTATTTGACGGTAATCAACCAGAAAGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAA

TACGTAGGTGGCAAGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGAATGATAAGTCTGATGTGAAAG
CCCACGGCTCAACCGTGGAACTGCATCGGAAACTGTCATTCTTGAGTGCAGAAGAGGAGAGTGGAATTCCATGTGTA
GCGGTGGAATGCGTAGATATATGGAAGAACACCAGTGGCGAAggCGGCTCTCTGGTCTGCAACTGACGCTgaggcTC
GAAAGCATGGGTAGCGAACAGGATTAGATACCCTGGTAGTCCATGCCgtaaacgATGagCGCTAGGTGTCGGGGACT
TTCCGGTTCTCagtGCCGC

Sequence 10 — *L. delbrueckii*

TTCGTCATTCCTTTGAGTttaACCTTGCGGTCGTACTCCCCAGGCGGAGCGCTTAATGCGTTTGCTGCGGCACTGAG
GACCGGAAAGTCCCCAACACCTAGCGCTCATCGTTTACGGCATGGACTACCAGGGTATCTAATCCTGTTCGCTACCC
ATGCTTTCGAGCCTCAGCGTCAGTTGCAGACCAGAGAGCCGCCTTCGCCACTGGTGTTCTTCCATATATCTACGCAT
TCCACCGCTACACATGGAATTCCACTCTCCTCTTCTGCGCTCAAGAATGACAGTTTCCGATGCAGTTCCACGGTTGA
GCCGTGGGCTTTCACATCAGACTTATCATTCCGCCTGCGCTCGCTTTACGCCCAATAAATCCGGACAACGCTTGCCA
CCTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGACTTTCTGGTTGATTACCGTCAAATAAAGACCAGTTA
CTGCCTCTATCCTTCTTCACCAACAACAGAGCTTTACGATCCGAAAACCTTCTTCACTCACGCGGCGTTGCTCCATC
AGACTTGCGTCCATTGTGGAAGATTCCCTACTGCTGCCTCCCGTAGGAGTTTGGGCCGTGTCTCAGTCCCAATGTGG
CCGATCAGTCTCTCAACTCGGCTACGCATCATTGCCTTGGTAGGCCTTTACCCCACCAACTAGCTAATGCGCCGCGG
GCTCATCCTAAAGTGACAGCTTACGCCGCCTTTCAAACTTGAATCATGCGATTCATGTTGTTATCCGGTATTAGCAC
CTGTTTCCAAGTGGTATCCCAGTCTTTAGGGCAGATTGCCCACGTGTTACTCACCCATCCGCCGCTAGCGTCCAaca
aatcaccccgaagGGATCTTTgaatTcAGCTCGCTCGACTTGCATGTA

Sequence 11 — *S. thermophilus*

GGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTCTTGGAT
GAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGCTAATAC
CGCATAACAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTTGTATT
AGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACTGGGAC
TGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAGCAACG
CCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAAGTTCAC
ACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGAGCGTT
GTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACCATAGT
TCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGATAT
ATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTGaggCTCGAAAGCGTgnggAGCGAACA
GGATTAGATACCCTGGTAGTCCACGCCGtaaacgATGAGTGCtaggtgTTGGATCCTTTCCggattCAGTGCCGcA
GCTaACGCATTAAGcACTCCG

Sequence 12 — *S. thermophilus*

ATCATGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTCT
TGGATGAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGCT
AATACCGCATAACAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTT
GTATTAGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACT
GGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAG
CAACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCACTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAAG
TTCACACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGA
GCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACC
ATAGTTCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTA
GATATATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTGaggCTCGAAAGCGTgngGaGC
GAACAGGATTAGaTACCCTGGTAGTCCACGCCGTanaCGATGAGTGCTaggTGTTGGATCCTTTCCGGGATTCAGTG
CCGCAGCTAACGCATTAAGcACT

Sequence 13 — *S. thermophilus*

TCCTGGCtcgGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTCTTG
GATGAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGCTAA
TACCGCATAGCAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTTGT
ATTAGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACTGG
GACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAGCA
ACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAAGTT
CACACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGAGC
GTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACCAT

AGTTCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGA
TATATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTGaggCTCGAAAGCGTgnggaGCGA
ACAGGATTAGATACCCTGGTAGTCCACGCCGTanacGATGAGTGCTAGgTGTTGGATCCTTTCCGGGATTCAG

Sequence 14 — *S. thermophilus*

GATCATGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTC
TTGGATGAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGC
TAATACCGCATAACAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGT
TGTATTAGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACAC
TGGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGA
GCAACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAA
GTTCACACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCG
AGCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAAC
CATAGTTCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGT
AGATATATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTGaggCTCGAAAGCGTGGGAGC
GAACAGGATTAGATACCCTGGTAGTCCACGCCgTanaCGATGAGTGCTAGGTGTTGGATC

Sequence 15 — *S. thermophilus*

TTCGTCATTCCTTTGAGTttaACCTTGCGGTCGTACTCCCCAGGCGGAGTGCTTAATGCGTTAGCTTCGGCACTGAA
TCCCGGAAAGGATCCAACACCTAGCACTCATCGTTTACGGCGTGGACTACCAGGGTATCTAATCCTGTTCGCTCCCC
ACGCTTTCGAGCCTCAGCGTCAGTTACAGACCAGAGAGCCGCTTTCGCCACCGGTGTTCCTCCATATATCTACGCAT
TTCACCGCTACACATGGAATTCCACTCTCCCCTTCTGCACTCAAGTTTGACAGTTTCCAAAGCGAACTATGGTTGAG
CCACAGCCTTTAACTTCAGACTTATCAAACCGCCTGCGCTCGCTTTACGCCCAATAAATCCGGACAACGCTCGGGAC
CTACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTCCCTTTCTGGTAAGCTACCGTCACAGTGTGAACTTTCCA
CTCTCACACCCGTTCTTGACTTACAACAGAGCTTTACGATCCGAAAACCTTCTTCACTCACGCGGCGTTGCTCGGTC
AGGGTTGCCCCCATTGCCGAAGATTCCCTACTGCTGCCTCCCGTAGGAGTCTGGGCCGTGTCTCAGTCCCAGTGTGG
CCGATCACCCTCTCAGGTCGGCTATGTATCGTCGCCTAGGTGAGCCATTACCTCACCTACTAGCTAATACAACGCAG
GTCCATCTTGTAGTGGAGCAATTGCCCCTTTCAAATAAATGACATGTGTCATCCATTGTTATGCGGTATTAGCTATC
GTTTCCAATAGTTATCCCCCGCTACAAGGCAGGTTACCTACGCGTTACTCACCCGTTCGCAACTCATCCAagaagag
cAAGCTCCTCTCTTCAGCGTTCTACTTGCATGTA

Sequence 16 — *S. thermophilus*

TCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTCTT
GGATGAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGCTA
ATACCGCATAACAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTTG
TATTAGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACTG
GGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAGC
AACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAAGT
TCACACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGAG
CGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACCA
TAGTTCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAG
ATATATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTGaggCTCGAAAGCGTgngGaGCG
AACAGGATTAGATACCCTGGTAGCCCACGCCGTanaCGATGAGTGCTAGGTGTTGGATCCTTTCCGGGATTCAGTGC
CGCAGCTAACGCATTAAGcACTCcgcCTGG

Sequence 17 — *S. thermophilus*

GGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTCTTGGAT
GAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCCTGTGGCGGGGGATAACTATTGGAAACGATAGCTAATAC
CGCATAACAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTTGTATT
AGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACTGGGAC
TGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAGCAACG
CCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAAGTTCAC
ACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGAGCGTT
GTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACCATAGT
TCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAGATAT
ATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTgaggcTCGAAAGCGTgngGAGCGAACA
GGATTAGATACCCTGGTAGTCCACGCCgtaaacGATGagcGCTAGGTGTTGGGgACtTT

Sequence 18 — *S. thermophilus*

GTCATTCCTTTGAGTTTCAACCTTGCGGTCGTACTCCCCAGGCGGAGCGCTTAATGCGTTTGCTGCGGCACTGAGGA
CCGGAAAGTCCCCAACACCTAGCACTCATCGTTTACGGCGTGGACTACCAGGGTATCTAATCCTGTTCGCTCCCCAC
GCTTTCGAGCCTCAGCGTCAGTTACAGACCAGAGAGCCGCTTTCGCCACCGGTGTTCCTCCATATATCTACGCATTT
CACCGCTACACATGGAATTCCACTCTCCCCTTCTGCACTCAAGTTTGACAGTTTCCAAAGCGAACTATGGTTGAGCC
ACAGCCTTTAACTTCAGACTTATCAAACCGCCTGCGCTCGCTTTACGCCCAATAAATCCGGACAACGCTCGGGACCT
ACGTATTACCGCGGCTGCTGGCACGTAGTTAGCCGTGACTTTCTGGTTGATTACCGTCAAATAAAGACCAGTTACTG
CCTCTATCCTTCTTGACTTACAACAGAGCTTTACGATCCGAAAACCTTCTTCACTCACGCGGCGTTGCTCGGTCAGG
GTTGCCCCCATTGCCGAAGATTCCCTACTGCTGCCTCCCGTAGGAGTCTGGGCCGTGTCTCAGTCCCAGTGTGGCCG
ATCACCCTCTCAGGTCGGCTATGTATCGTCGCCTAGGTGAGCCATTACCTCACCTACTAGCTAATACAACGCAGGTC
CATCTTGTAGTGGAGCAATTGCCCCTTTCAAATAAATGACATGTGTCATCCATTGTTATGCGGTATTAGCTATCGTT
TCCAATAGTTATCCCCCGCTACAAGGCAGGTTACCTACGCGTTACTCACCCGTTCGCAACTCATCCAagaagagcAA
GCTCCTCTCTTCAGCgTtcT

Sequence 19 — *L. delbrueckii*

TGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTCGAGCGAGCTGAATTCAAAGATTCCTTCGGGATG
ATTTGTTGGACGCTAGCGGCGGATGGGTGAGTAACACGTGGGCAATCTGCCCTAAAGACTGGGATACCACTTGGAAA
CAGGTGCTAATACCGGATAACAACATGAATCGCATGATTCAAGTTTGAAAGGCGGCGTAAGCTGTCACTTTAGGATG
AGCCCGCGGCGCATTAGCTAGTTGGTGGGGTAAAGGCCTACCAAGGCAATGATGCGTAGCCGAGTTGAGAGACTGAT
CGGCCACATTGGGACTGAGACACGGCCCAAACTCCTACGGGAGGCAGCAGTAGGGAATCTTCCACAATGGACGCAAG
TCTGATGGAGCAACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTTGGTGAAGAAGGATAGAGG
CAGTAACTGGTCTTTATTTGACGGTAATCAACCAGAAAGTCACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGT
AGGTGGCAAGCGTTGTCCGGAATTTATTGGGCGTAAAGCGAGCGCAGGCGGAATGATAAGTCTGATGTGAAAGCCCAC
GGCTCAACCGTGGAACTGCATCGGAAACTGTCATTCTTGAGTGCAGAAGAGGAGAGTGGAATTCCATGTGTAGCGGT
GGAATGCGTAGATATATGGAAGAACACCAGTGGCGaaggcGGCTCTCTGGTCTGCAACTGACGCTGaggCTCGAAAG
CATGGGTAGCGAACAGGATTAGATACCCTGGTAGTCCATGCCGtaaacGATGagcGCTAGGTGTtggggactTTCCC
CTCCTCagtGCCGC

Sequence 20 — *S. thermophilus*

ATCATGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTCT
TGGATGAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGCT
AATACCGCATAACAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTT
GTATTAGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACT
GGGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAG
CAACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAAG
TTCACACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGA
GCGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACC
ATAGTTCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTA
GATATATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTGaggCTCGAAAGCGTgngGAGC
GAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGATGAGTGCTaggTGTTGGATCCTTtcggggatcagTGC
CGCAgcTaACGCATTAAGCACTC

Sequence 21 — *S. thermophilus*

TCCTGGCTCAGGACGAACGCTGGCGGCGTGCCTAATACATGCAAGTAGAACGCTGAAGAGAGGAGCTTGCTCTTCTT
GGATGAGTTGCGAACGGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGCTA
ATACCGCATAACAATGGATGACACATGTCATTTATTTGAAAGGGGCAATTGCTCCACTACAAGATGGACCTGCGTTG
TATTAGCTAGTAGGTGAGGTAATGGCTCACCTAGGCGACGATACATAGCCGACCTGAGAGGGTGATCGGCCACACTG
GGACTGAGACACGGCCCAGACTCCTACGGGAGGCAGCAGTAGGGAATCTTCGGCAATGGGGGCAACCCTGACCGAGC
AACGCCGCGTGAGTGAAGAAGGTTTTCGGATCGTAAAGCTCTGTTGTAAGTCAAGAACGGGTGTGAGAGTGGAAAGT
TCACACTGTGACGGTAGCTTACCAGAAAGGGACGGCTAACTACGTGCCAGCAGCCGCGGTAATACGTAGGTCCCGAG
CGTTGTCCGGATTTATTGGGCGTAAAGCGAGCGCAGGCGGTTTGATAAGTCTGAAGTTAAAGGCTGTGGCTCAACCA
TAGTTCGCTTTGGAAACTGTCAAACTTGAGTGCAGAAGGGGAGAGTGGAATTCCATGTGTAGCGGTGAAATGCGTAG
ATATATGGAGGAACACCGGTGGCGAAAGCGGCTCTCTGGTCTGTAACTGACGCTGaggCTCGAAAGCGTGGGGAGCG
AACAggATTAGATACCCTGGTAGTCCACGCCgtaaacgATGagtGCTAGgTgttggATCCTTTCcgggattCAGTGC
CGCAGCTAACGCAT

## Appendix A.B

## Lengths of the corresponding T-RFLP fragments for each enzyme (*MspI* and *RsaI* restriction enzymes) in the additional sequences (*S. thermophilus* and *L. delbrueckii*) of the molecular clone library

| Sequence | Identification | Length of *Msp*I-cut Fragment (base pairs) | Length of *Rsa*I-cut Fragment (base pairs) |
|---|---|---|---|
| 1 | *L. delbrueckii* | 179 | 714 |
| 2 | *S. thermophilus* | 555 | Not applicable (NA) |
| 3 | *S. thermophilus* | 78 | 32 |
| 4 | *S. thermophilus* | 555 | NA |
| 5 | *S. thermophilus* | 555 | NA |
| 6 | *L. delbrueckii* | 179 | NA |
| 7 | *S. thermophilus* | 555 | NA |
| 8 | *S. thermophilus* | 555 | NA |
| 9 | *L. delbrueckii* | 179 | NA |
| 10 | *L. delbrueckii* | 80 | 34 |
| 11 | *S. thermophilus* | 555 | NA |
| 12 | *S. thermophilus* | 555 | NA |
| 13 | *S. thermophilus* | 545 | NA |
| 14 | *S. thermophilus* | 555 | NA |
| 15 | *S. thermophilus* | 80 | 34 |
| 16 | *S. thermophilus* | 555 | NA |
| 17 | *S. thermophilus* | 555 | NA |
| 18 | *S. thermophilus* | 78 | 32 |
| 19 | *L. delbrueckii* | 179 | NA |
| 20 | *S. thermophilus* | 555 | NA |
| 21 | *S. thermophilus* | 555 | NA |