bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Using eDNA methodologies to identify the presence/absence of *S. thermophilus* in yogurt products

Vladimir Bejdo[a,*], Claire Gao[a,]

[a] *University of Washington, Seattle, WA*

Abstract

The advent of low-cost nuclear acid extraction allows for the creation of low-cost assays which can specifically be used to determine the presence or absence of bacteria in a variety of environments. Commercially sold dietary yogurt claims to contain bacteria forming a microbiome which has been previously linked to improved health outcomes in incidence rates of type 2 diabetes in vulnerable populations. The predicted outcome was that a main bacterial culture used in yogurt production *(S. thermophilus)* would be present and would be detectable using eDNA methodologies since it is fundamental to the making of yogurt. We gathered DNA from yogurt using yogurt dilutions and filter paper; we then extracted the DNA and also sequenced and used PCR to amplify the DNA. We sequenced PCR products to verify their identity through processing with publicly available BLAST tools which reference already accessioned bacterial genomes. Yogurt from four different commercially available brands (in the U.S.) was tested; not all

*Principal corresponding author.
Email addresses: bejdo@uw.edu (Vladimir Bejdo), claireg5@uw.edu (Claire Gao)

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

yogurts tested positive for the bacteria, with higher concentrations of the bacteria in imported Greek-style yogurts, lower concentrations in domestically produced Greek-style yogurts, and no bacterial DNA detected in domestically produced 'regular' yogurt products. This research suggests that not all yogurts are equal, putting into question the whole-sale claims made on dietary yogurt's probiotic preventative health effects and calling for a more detailed analysis to determine firm causal links between the microbiota of yogurts and preventative health effects.

Keywords: environmental DNA, bacterial fingerprinting, *S. thermophilus*, probiotics, yogurt cultures

---

1 **1. Introduction**

2 Previous studies have shown (and advertisers often promote) the pos-
3 itive health impacts which seem to be correlated to eating yogurt and
4 other fermented foods. These effects have been attributed to the probiotics
5 which promote the fermentative processes that create these products; in
6 dietary yogurt, these two bacteria are *L. delbrueckii ssp. bulgaricus* and *S.*
7 *thermophilus* [1].

8 Consuming yogurt on a regular basis has been shown to reduce the
9 risk of T2D in elderly populations at a high risk of cardiovascular disease
10 in Spain [2]. The same effect has been found in general adult populations
11 in the United States, where over 194 thousand adults were examined on
12 eating habits and in particular on the types of dairy that they ate in a study
13 which conducted meta-analysis to find similar effects caused from yogurt
14 consumption [3]. In order for these yogurt products to have any positive

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

15 effect, they need to have the proper bacteria which facilitate fermentation;
16 this study aims to introduce a methodology to prove that yogurt, a po-
17 tential preventative against widespread disease like type 2 diabetes, as
18 commercially sold, truly has the effects attributed to these bacteria that
19 have been studied and publicized domestically and internationally by con-
20 firming their presence in the yogurt microbiome.

21 No recorded efforts have been made to confirm manufacturer claims on
22 the presence of live cultures in yogurt products. Most studies in the past
23 that have analyzed the bacterial content of yogurt have utilized phenotypic
24 and fermentative profiling, which is highly variable for each strain of the
25 bacteria potentially present [4, 5]. An emerging method of bacterial identi-
26 fication in environmental sources involves the use of DNA sequencing, or at
27 the very least PCR/amplification. Some of the most prevalent methods for
28 using DNA in bacterial identification involve DNA fingerprinting by test-
29 ing through methods like random amplified polymorphic DNA (RAPD)
30 analysis and restriction fragment-length polymorphism (RFLP) analysis
31 [6, 7, 8, 9]–both methods that can produce results even when the specifics
32 of the sample tested are not well known [10]. However, these methods may
33 produce wildly variable PCR products, are at times resource-heavy, and
34 are not always reproducible due to protocols that are extremely dependent
35 on individual testing and laboratory conditions [11]. Due to these issues,
36 these methods are sometimes not seen as stable enough to be used as
37 stand-alone methods [12], and are often accompanied with existing profil-
38 ing methods. DNA sequencing has, however, advanced to the point where
39 we can create species specific primers for the particular and replicable am-

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

40 plification of bacterial DNA samples, potentially negating the need for a
41 hyper-general testing method. Methods recently used for the isolation of
42 environmental DNA have become increasingly cost effective, costing a few
43 cents per sample identified, and can provide quick and certain results [13].

44 To make use of this technology, we have conducted preliminary test-
45 ing of commercially available dietary yogurts in order to begin to form
46 a consistent, cost-effective methodology for the use of eDNA methods in
47 the identification of probiotics in yogurt. This paper comes to preliminary
48 conclusions and suggestions for further research to help protect trade and
49 consumer interests by pointing out potential false advertising, and sug-
50 gests that this methodology provides an opportunity to more definitively
51 prove causation between these bacteria and the effects of yogurt on vulner-
52 able groups by proving whether or not these bacteria are actually present
53 in the disease-preventative yogurt noted in literature.

54 **2. Methodology**

55 This section outlines the methodology used in yogurt sample collection,
56 DNA extraction, amplification, sequencing, and sequence processing.

57 In order to conduct this experiment, four different brands of commer-
58 cially available yogurt products (available at time of publication in Seattle,
59 WA) were used. The four yogurts tested were Fage's *Total* (an imported
60 Greek yogurt product), Chobani Plain Yogurt, Greek Gods Greek Yogurt
61 (domestic Greek yogurt products), and Yoplait Original Strawberry (a do-
62 mestic yogurt product). All the brands claim to have been produced using
63 starter cultures containing both tested bacteria, and claim to have 'live and

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

active' cultures.

2.1. *Sample collection*

Samples were prepared for collection by preparing a 1:25 dilution in a test tube; this was done by taking 0.5 mL of yogurt and adding it to 12 mL of distilled water. This mixture is then capped and agitated for about 30 seconds. Precipitation may be present; this does not affect the overall sample quality. For the actual collection process, a piece of standard Whatman Grade 1 (15mm diameter) qualitative filter paper is placed into the mixture for 30 seconds with forceps, retrieved, and stored in an appropriately sized (2 mL) screw capped microcentrifuge tube with 1 mL of NP-40 lysis buffer for storage. This process allows for the DNA to be isolated and quickly prepared for extraction shortly after sampling. Alternatively, samples from the solution can be pipetted onto the filter paper for storage; the 'dipping' method allows for less specialized tools to be used to achieve the same effect, and is the one used in this experiment.

2.2. *DNA extraction/purification*

DNA extraction/purification was performed with a method similar to the one found in the article *Nucleic acid purification from plants, animals and microbes in under 30 seconds* which calls for directly dipping a dipstick exposed to the sample to an amplification reaction after a few dips in a wash buffer [13]. Instead, to extract as much DNA as possible from samples and additionally keep material costs down, a small (roughly 1 cm x 1 cm large) piece of the original filter paper was placed into 500 µL Tris buffer and then incubated in 200 µL of distilled water in order to extract DNA off the filter

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

88 paper. This allows for sufficient time to have a solution that can be used
89 for amplification.

90 *2.3. PCR amplification*

91 The particular PCR beads used to perform PCR amplification were
92 General Electric's illustra™ PuReTaq™ Ready-To-Go™ PCR beads. The
93 procedure is performed as standard–adding a reaction bead, 1 µL of each
94 forward and reverse primer for the desired bacteria, 18 µL of water, and
95 5 µL of the DNA template prepared in the previous stage into a 200 µL
96 Eppendorf tube.

97 *2.4. Primers*

98 Primers were designed using publicly available accessioned bacterial
99 genomes for *S. thermophilus* using NCBI's Primer-BLAST. The forward
100 primer is GCT TTA GGG CTA GCG TCG AT, while the reverse primer
101 is TAG GTC CCG ACT AAC CCA GG, listed in $5' \rightarrow 3'$ order. The ex-
102 pected product length was listed as 524 base pairs.

103 The following table lists additional information regarding the primers
104 designed, including length, melting and annealing temperatures (T_m and
105 T_a), GC content, and folding T_m.

Table 1: **Primer Data.** This table details pertinent primer design data for the primers used in this experiment, both forward and reverse.

Primer	T_m,°C	T_a, °C	GC content	Folding T_m, °C
S. thermophilus, forward	57.1	52.1	55.00%	25.3
S. thermophilus, reverse	57.9	52.9	60.00%	29.2

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

106 Primers were resuspended from the lyophilized form in which they
107 were delivered by the addition of molecular grade water to create a working
108 stock; 275 µL of water was added for the forward primer, and 283 µL of
109 water was added for the reverse primer.

110 *2.5. PCR purification, gel electrophoresis, sequencing*

111 All of these steps were performed by an external genomics service/CRO,
112 Genewiz, LLC in South Plainfield, New Jersey. PCR products were puri-
113 fied by magnetic bead-based clean up, and were sequenced with standard
114 Sanger sequencing methodologies.

115 *2.6. Sequence processing*

116 In order to produce FASTA files from .ab1 files (files containing elec-
117 tropherograms and base sequences), the program SnapGene Viewer was
118 used. The FASTA files produced from the electropherograms were then
119 processed through both the publicly available National Center for Biotech-
120 nology Information (NCBI) and Kyoto Encyclopedia of Genes and Genomes
121 (KEGG) databases through their respective BLAST tools in order to identify
122 and verify the identity of the sequenced DNA.

123 **3. Results**

124 *3.1. Gel electrophoresis results*

125 The results for gel electrophoresis are located in Figure 1. The figures
126 are labeled with a short-hand labelling convention; this is in the form of
127 [brand name], [first letter of bacterial genus tested]. Each brand name is
128 also abbreviated; Chobani is CHOB, Greek Gods Greek Yogurt is GGGY,

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Yoplait is YPLT, and Fage remains FAGE. Unlabeled wells indicate results which do not belong to this experiment. For *S. thermophilus*, Chobani, Greek Gods Yogurt, and Fage give positive gel results, while Yoplait gives a negative result.

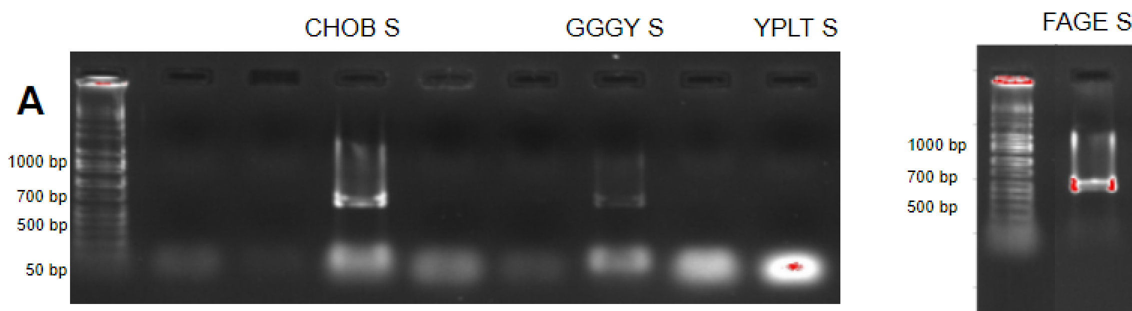

Figure 1: **Annotated electrophoresis results**. Row **A** contains all the electrophoresis results for *S. thermophilus*. CHOB S, GGGY S, and FAGE S were positive and are marked by red boxes. The e-gel DNA ladder from each gel is labeled with rough indicators of amplicon size in number of base pairs (bp). The results for FAGE S are shown seperately as they were processed on a separate gel.

The samples provided varying degrees of primer hybridization, especially in samples that provided visually weaker results. Stronger results are indicated by both stronger (whiter) bands in the gel result, and red marks indicate results with greater concentrations of DNA. This suggests greater concentrations of DNA product in the FAGE S sample, followed by the CHOB S and GGGY S samples respectively.

3.2. BLAST results (sequence identification)

After processing with publicly available BLAST tools from both NCBI and KEGG resources, all successful amplifications were identified as match-

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

142 ing with other existing accessioned *S. thermophilus* samples. Table **??** lists
143 a compiled record of accession numbers and match information for the
144 primary matches of each sample sequenced and referenced against NCBI
145 and KEGG databases. This data is accurate as of March 3rd, 2018.

146 The products of the Fage and Chobani samples all fully match to the
147 same genome sample for *S. thermophilus* in the NCBI database, as indicated
148 by a 100 percent identity percentage match to accession CP025400.1. The
149 products of the Greek Gods Greek Yogurt sample indicate a 99 percent
150 identity percentage match in the NCBI database to accession MG825731.1 (a
151 partial sequence of *S. thermophilus* from the 16S-rRNA region of a bacterial
152 sample).

153 4. Discussion of results

154 *4.1. Result accuracy*

155 *Yoplait sample.* The negative result for the Yoplait sample may be inter-
156 preted as no bacteria present in the sample; it may also be interpreted
157 as being a false-negative due to a lack of bacterial DNA in the particular
158 sample taken, and not in the yogurt itself. There is also a potential for
159 false-negatives due to poor reaction performance; however, this might not
160 be the case as the protocol followed is the same for all samples taken (3 of
161 which were successful) and is based on a proven protocol for the detection
162 of eDNA. There may very well have been no detectable bacterial DNA in
163 the yogurt; evidently enough, the primers simply reproduced themselves
164 (resulting in a strong band near the bottom of the gel image). While the
165 yogurt claimed to include live and active cultures, no specifics were listed

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

166 in the ingredients list on the yogurt tested itself.

167 *Primers and their role in testing.* The primers seem to have been non-specific
168 enough to reproduce part of the 16S rRNA gene in one of the samples,
169 which was not necessarily the original target sequence of the primers when
170 they were created. At the very least, this could mean that the primers are
171 not specific to a particular commercial strain of *S. thermophilus*, meaning
172 that they can be used to amplify more strains and can be used on a wider
173 range of commercially available commercial yogurt. However, they are are
174 non-specific enough to reproduce at least two differing regions of the same
175 bacterial genome.

176 *4.2. Contaminants and culture amounts from the point of production*

177 *Yogurt contents and detection.* It seems that all yogurts are simply not made
178 equal. Commercially sold dietary yogurt products could variably contain
179 substances that make it harder to determine the presence/absence of their
180 claimed bacterial cultures.

181 The strongest bands and consistent results come from yogurts that
182 only list two sets of ingredients; the Fage and Chobani-branded plain
183 yogurt products tested sampled only list milk and yogurt cultures as their
184 ingredients. Greek Gods Greek Yogurt, the one yogurt with a smaller
185 match percentage to a more general gene of the bacteria, contained cream
186 and pectin in addition to milk and yogurt cultures. These were most likely
187 placed in the yogurt as thickeners.

188 The one negative result comes from Yoplait yogurt, which most promi-
189 nently of various thickeners (such as pectin, gelatin, starches, corn syrups,

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

190 and other oils), flavorings, and colorants. The yogurt tested was Yoplait
191 Original Strawberry; while it claimed to contain live cultures, these were
192 not explicitly listed on the ingredient list.

193 It is very likely that the presence of these other substances may cause
194 issues in the detection of bacterial DNA, whether or not those bacteria are
195 actually dead or living. Issues of contamination could also be the cause
196 of the weaker bands in the Yoplait and Greek Gods samples; both of these
197 samples have more ingredients that could dilute the presence of these
198 bacterial cultures in yogurt.

199 5. Conclusions

200 *5.1. Re-evaluating links between yogurt, probiotics, and health impacts*

201 The uneven set of results raises concerns over the importance of bacte-
202 rial concentration and the actual composition of yogurt biota on the effects
203 of dietary yogurt. If yogurt with almost undetectable amounts of DNA
204 present for probiotic strains of bacteria is credited for providing positive
205 health impacts, are the bacteria themselves the beneficial agents making
206 yogurt a preventative for disease? Or, are bacterial byproducts (the fer-
207 mented yogurt itself) the actual source of these positive health impacts?

208 DNA testing through this cost-effective method could provide a gold-
209 standard for bacterial identification for the biota present in dietary prod-
210 ucts. With DNA quantification methods, one could further determine the
211 actual concentration and quality of DNA present in each sample mixture.

212 However, it might be more useful in certain cases to pursue further
213 research through cell culturing of yogurt products to both provide a more

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

214 complete picture of a commercial yogurt's biome and make more general-
215 ized statements about the realities of the probiotic value of commercially
216 sold yogurt. The contaminants which we speculated over may affect DNA
217 testing for bacteria, but not culturing; we can attempt to determine the
218 actual rate at which probiotics are present in dietary yogurt in order to de-
219 termine whether it is the probiotics being live in the yogurt sold or simply
220 the product of fermentative processing (yogurt itself, with or without live
221 cultures) that provides these speculated benefits.

222 Specialized testing through DNA methods is most useful to come to
223 important research leads about the causal link between bacteria and these
224 positive health effects. More extensive testing on larger batches of yogurt
225 could provide a better idea of overall product quality and consistency, lead-
226 ing to a baseline for consideration by nutritional and health researchers to
227 further specify proven causal relationships between probiotics and human
228 health benefits. The low cost and time required to prepare and process
229 samples from the point of collection could also promote more consumer-
230 oriented research into the quality of products sold; and, if this technology
231 indeed is used to prove that there is a strong causal link between *S. ther-*
232 *mophilus* and human health effects, could inform the creation of actionable
233 points to reinforce recommendations on issues of public and preventative
234 health, and could also lead to better decisions for consumers who try to
235 take advantage of probiotics in products sold to them.

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

236 ## Competing financial interests

237 There are no competing interests to declare.

238 ## Data availability

239 All the data related to this set of experiments is included in this article as
240 part of the text, as figures, or in the supporting information provided in the
241 text. Additional supplementary files analyzed in this set of experiments
242 have also been released.

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supporting Information

S1: FASTA file, FAGE S

This is a copyable version of the FASTA file for FAGE S which was interpreted from a chromatogram and used for BLAST.

```
>FAGE SAMPLE STREPTOCOCCUS THERMOPHILUS
AACGCATTAAGCACTCCGCCTGGGGAGTACGACCGCAAGGTTGAAACTCAAAGGAATTGAC
GGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAATTCGAAGCAACGCGAAGAACCTTACC
AGGTCTTGACATCCCGATGCTATTTCTAAAGATAGAAAGTTACTTCGGTACATCGGTGACA
GGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGATGTTGGGTTAAGTCCCGCAACGAGC
GCAACCCCTATTGTTAGTTGCCATCATTCAGTTGGGCACTCTAGCGAGACTGCCGGTAATA
AACCGGAGGAAGGTGGGGATGACGTCAAATCATCATGCCCCTTATGACCTGGGCTACACAC
GTGCTACAATGGTTGGTACAACGAGTTGCGAGTCGGTGACGGCGAGCTAATCTCTTAAAGC
CAATCTCAGTTCGGATTGTAGGCTGCAACTCGCCTACATGAAGTCGGAATCGCTAGTAATC
GCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGCCTTGTACACACCGCCCGTCACACCA
CNAGAGTTTGTAACACCCGAAGTCGGTGAGGTAACCTTTTGGAGCCAGCCGCCTAAGGTGG
GACANATGATTGGGGTGAAGTCGTAACAAGG
```

S2: FASTA file, CHOB S

This is a copyable version of the FASTA file for CHOB S which was interpreted from a chromatogram and used for BLAST.

```
>CHOB SAMPLE STREPTOCOCCUS THERMOPHILUS
AAGGTTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGTGGTTTAAT
TCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCCGATGCTATTTCTAGAGATAGA
AAGTTACTTCGGTACATCGGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTGTCGTGAGA
```

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

```
TGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATTGTTAGTTGCCATCATTCAGTTGGG
CACTCTAGCGAGACTGCCGGTAATAAACCGGAGGAAGGTGGGGATGACGTCAAATCATCAT
GCCCCTTATGACCTGGGCTACACACGTGCTACAATGGTTGGTACAACGAGTTGCGAGTCGG
TGACGGCGAGCTAATCTCTTAAAGCCAATCTCAGTTCGGATTGTAGGCTGCAACTCGCCTA
CATGAAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGTTCCCGGGC
CTTGTACACACCGCCCGTCACACCACGAGAGTTTGTAACACCCGAAGTCGGTGAGGTAACC
TTTTGGAGCCAGCCGCCTAAGGTGGGACAGATGATTGGGGTGAAGTCGTAACAAG
```

S3: FASTA file, GGGY S

This is a copyable version of the FASTA file for GGGY S which was interpreted from a chromatogram and used for BLAST.

```
>GGGY SAMPLE STREPTOCOCCUS THERMOPHILUS
ACGACCGCAAGGTTGAAACTCAAAGGAATTGACGGGGGCCCGCACAAGCGGTGGAGCATGT
GGTTTAATTCGAAGCAACGCGAAGAACCTTACCAGGTCTTGACATCCCGATGCTATTTCTA
AAAATAGAAAGTTACTTCGGTACNTCGGTGACAGGTGGTGCATGGTTGTCGTCAGCTCGTG
TCGTGAGATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCCTATTGTTAGTTGCCATCATT
CAGTTGGGCACTCTAACGAGACTGCCGGTAATAAACCGGAGGAAGGTGGGGATGACGTCAA
ATCATCATGCCCCTTATGACCTGGGCTACACACGTGCTACAATGGTTGGTACAACGAGTTG
CGAGTCGGTGACGGCGAGCTAATCTCTTAAAGCCNATCTCAGTTCGGATTGTAGGCTGCAA
CTCGCCTACATGAAGTCGGAATCGCTAGTAATCGCGGATCAGCACGCCGCGGTGAATACGT
TCCCGGGCCTTGTACACACCGCCCGTCACACCNNNAGAGTTTGTANNNNNCGAAGTCGGTG
AGGTAACCTTTTGGAGCCAGCCGCCTAAGGTGGGACAGATGATTGGGGTGAAGTCGTAACA
```

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

287 **S4: .zip archive, chroma_STherm.zip**

288 This archive contains the chromatogram files for each sample tested.
289 The following is a list of the zip-file's contents.

290 *FAGE_S-Stherm.ab1.* A file containing a full chromatogram for the Fage
291 sample, which can be viewed using many different programs dedicated to
292 viewing chromatograms and can also be converted to FASTA files as was
293 done in this paper.

294 *CHOB_S-Stherm.ab1.* A file containing a full chromatogram for the Chobani
295 sample.

296 *GGGY_S-Stherm-ab1.* A file containing a full chromatogram for the Greek
297 Gods Greek Yogurt sample.

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

[1] M. L. Marco, D. Heeney, S. Binda, C. J. Cifelli, P. D. Cotter, B. Foligné, M. Gänzle, R. Kort, G. Pasin, A. Pihlanto, E. J. Smid, R. Hutkins, Health benefits of fermented foods: microbiota and beyond, Current Opinion in Biotechnology 44 (2017) 94 – 102, food biotechnology • Plant biotechnology. doi:https://doi.org/10.1016/j.copbio.2016.11.010. URL http://www.sciencedirect.com/science/article/pii/S095816691630266X

[2] A. Díaz-López, M. Bulló, M. Martínez-González, D. Corella, R. Estruch, M. Fitó, E. Gómez-Gracia, M. Fiol, F. J. De la Corte, E. Ros, N. Babio, L. Serra-Majem, X. Pinto, M. Muñoz, F. Francés, P. Buil-Cosiales, J. Salas-Salvadó, Dairy product consumption and risk of type 2 diabetes in an elderly spanish mediterranean population at high cardiovascular risk, European journal of nutrition 55.

[3] M. Chen, Q. Sun, E. Giovannucci, D. Mozaffarian, J. E. Manson, W. C. Willett, F. B. Hu, Dairy consumption and risk of type 2 diabetes: 3 cohorts of us adults and an updated meta-analysis, BMC Med 12 (2014) 215, 215[PII]. doi:10.1186/s12916-014-0215-1. URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4243376/

[4] V. Xanthopoulos, D. Petridis, N. Tzanetakis, Characterization and classification of streptococcus thermophilus and lactobacillus delbrueckii subsp. bulgaricus strains isolated from traditional greek yogurts, Journal of Food Science 66 (5) 747–752. doi:10.1111/j.1365-2621.2001.tb04632.x.

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

322 URL https://onlinelibrary.wiley.com/doi/abs/10.1111
323 /j.1365-2621.2001.tb04632.x

324 [5] F. de Vin, P. Rådström, L. Herman, L. De Vuyst, Molecular and
325 biochemical analysis of the galactose phenotype of dairy strep-
326 tococcus thermophilus strains reveals four different fermentation
327 profiles, Applied and Environmental Microbiology 71 (7) (2005)
328 3659–3667. arXiv:https://aem.asm.org/content/71/7/3659.full.pdf,
329 doi:10.1128/AEM.71.7.3659-3667.2005.
330 URL https://aem.asm.org/content/71/7/3659

331 [6] M. C. Collado, M. Hernández, Identification and differentiation of
332 lactobacillus, streptococcus and bifidobacterium species in fermented
333 milk products with bifidobacteria, Microbiological Research 162 (1)
334 (2007) 86 – 92. doi:https://doi.org/10.1016/j.micres.2006.09.007.
335 URL http://www.sciencedirect.com/science/article/pii
336 /S0944501306001133

337 [7] G. Moschetti, G. Blaiotta, M. Aponte, P. Catzeddu, F. Villani, P. Deiana,
338 S. Coppola, Random amplified polymorphic dna and amplified ribo-
339 somal dna spacer polymorphism: powerful methods to differentiate
340 streptococcus thermophilus strains., Journal of Applied Microbiology
341 85 (1) (1998) 25–36. doi:10.1046/j.1365-2672.1998.00461.x.

342 [8] G. Moschetti, G. Blaiotta, F. Villani, S. Coppola, E. Parente, Com-
343 parison of statistical methods for identification of streptococcus ther-
344 mophilus, enterococcus faecalis, and enterococcus faecium from ran-
345 domly amplified polymorphic dna patterns, Appl Environ Micro-

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

346 biol 67 (5) (2001) 2156–2166, 1606[PII]. doi:10.1128/AEM.67.5.2156-
347 2166.2001.
348 URL http://www.ncbi.nlm.nih.gov/pmc/articles/PMC92850/

349 [9] P. Ivanova, S. Peykov, A. Dimitrova, S. Dimov, Molecular typ-
350 ing by genus-specific pcr and rapd profiling of diverse lactobacil-
351 lus delbrueckii strains isolated from cow, sheep and buffalo yo-
352 ghurts, Biotechnology & Biotechnological Equipment 22 (2) (2008)
353 748–753. arXiv:https://doi.org/10.1080/13102818.2008.10817545,
354 doi:10.1080/13102818.2008.10817545.
355 URL https://doi.org/10.1080/13102818.2008.10817545

356 [10] J. Welsh, M. McClelland, Fingerprinting genomes using pcr with ar-
357 bitrary primers, Nucleic acids research 18 (24) (1990) 7213–7218.

358 [11] G. A. Penner, A. Bush, R. Wise, W. Kim, L. Domier, K. Kasha,
359 A. Laroche, G. Scoles, S. J. Molnar, G. Fedak, Reproducibility of ran-
360 dom amplified polymorphic dna (rapd) analysis among laboratories.,
361 Genome Research 2 (4) (1993) 341–345. doi:10.1101/gr.2.4.341.
362 URL http://genome.cshlp.org/content/2/4/341.abstract

363 [12] V. Plengvidhya, F. Breidt Jr, H. Fleming, Use of RAPD-PCR as a
364 method to follow the progress of starter cultures in sauerkraut fer-
365 mentation, International journal of food microbiology 93 (3) (2004)
366 287–296.

367 [13] Y. Zou, M. G. Mason, Y. Wang, E. Wee, C. Turni, P. J. Blackall,
368 M. Trau, J. R. Botella, Nucleic acid purification from plants, animals

bioRxiv preprint doi: https://doi.org/10.1101/414060; this version posted November 27, 2018. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

369 and microbes in under 30 seconds, PLOS Biology 15 (11) (2017) 1–22.
370 doi:10.1371/journal.pbio.2003916.
371 URL https://doi.org/10.1371/journal.pbio.2003916