Comparing the label accuracy of an expensive probiotic versus a cheaper probiotic: Are you getting your money's worth?

Team 2

By: Hannah Berg, Alea Confer, Erika Henderson, Sophia Ripley

LB 145 Cell & Molecular Biology
Section 010
Dr. Masani
Alexander Leonowicz, Maggie Leff
23 April 2020

Abstract:

Medical professionals have recommended probiotics due to their associated health benefits. However, it’s easy to become overwhelmed by the different types of probiotics and how they vary dramatically in price for what seems to be the same thing. Therefore, we chose to investigate the accuracy of probiotic labels regarding species viability and diversity with the research question: *How does the cheaper probiotic from Company 1 compare with the more expensive probiotic from Company 2, given that they have similar claims with respect to CFU count and species identity?* This study is important to decipher which of the two probiotics will be more beneficial to consumers. In order to test this, we compared species viability by performing and plating serial dilutions on MRS agar. PCR was performed to amplify samples. 5µL of the amplified fragment was run on agarose gel electrophoresis to identify the presence of the 16s rRNA gene. Successful samples were sent for sequencing and BLAST analysis determined the bacterial genus and species components of the probiotics. Results showed Probiotic A had a higher viability (6223%) compared to Probiotic B (23.02%) due to contamination of Ralstonia found in sequencing and probiotic B had a more accurate species diversity according to Company 2’s label while Probiotic A didn’t due to Ralsontia being present. In conclusion, due to the harmfulness associated with the Ralstonia bacteria in Probiotic A and the accuracy of Probiotic B regarding species diversity, it was concluded that the more expensive Probiotic B was the safer choice.

Introduction:

Probiotics contain living microorganisms that are essential to maintaining a healthy gut by enhancing the bioavailability of minerals and treating gastrointestinal disorders (Haffner *et al.* 2016, Prasad *et. al* 1998). Probiotics stimulate the pre-existing bacteria in the gut which assists in nutritional benefits like aiding travelers' diarrhea, and reducing lactose intolerance and food allergies. *Lactobacillus,* a common bacterial strain, is present in many probiotics and commonly found in the gut (Goldin 1998). Each strain of *Lactobacillus* has intestinal health benefits such as prolonging remission of many types of inflammatory bowel disease, maintaining or improving intestinal barrier function, and helping with intestinal mucosal immunity (Engen *et. al* 2015).

Regarding safety aspects relating to probiotic regulation, there is no common worldwide safety standard for the legal status of probiotics as a dietary supplement (Wassenaar *et al.* 2008). The U.S., European Union (EU) and countries of the EU vary within probiotic protocols (Wassenaar *et al.* 2008). The FDA lacks strict requirements regarding the accessibility of probiotics as a dietary supplement to consumers (Degnan 2008). While differing in protocol methods, various probiotics companies claim to have the same health benefits; this variation creates uncertainty for the average consumer when deciding what probiotic best fits their needs.

Due to the many options available, it is unclear how to determine a probiotic's quality. This research is imperative to the real world because it is important for consumers to know if a cheaper and expensive probiotic will reap the same benefits. Therefore, we are investigating the accuracy of probiotic labels in terms of how many and what type of bacteria is present, and if the assumed relationship between price and quality is valid. To conduct this research, our group studied viability and species diversity to distinguish the better probiotic by comparing those values to the label. To compare viability, serial dilutions were plated on MRS Agar, a selective growth medium for lactobacillus, and then CFUs were counted (Bujalance 2006). BLAST analysis was used to identify species by comparing sequences of the 16s rRNA gene: a gene that is often used for bacterial identification due to its nature of slow evolutionary changes (Woese and Fox 1977). This allowed us to compare a cheaper probiotic from Company 1 (Probiotic A) versus a more expensive probiotic from Company 2 (Probiotic B), given that their labels claim to contain similar strains and CFUs. Knowing which probiotic is better and what is more true to their label allows customers to make better informed decisions.

Results:

To compare the viability of the two probiotics.

Serial dilutions (10^{-1} through 10^{-9}) and spread plating of the dilutions were performed on MRS Agar plates. Eighteen dilutions were plated for Probiotic A; however, only six plates contained successful viability counts that had 30-300 colonies. Twenty dilutions were plated for Probiotic B; however, only five plates contained successful viability counts. The experiments performed on 1/30/20 of measuring the viability of Probiotic And Probiotic B were successful (Figure 1 A,C). The experiment performed on 1/30/20 of Probiotic A found that the viability for the 10^{-7} dilution was 5.15×10^9 CFUs/pill (Figure 1A). The experiment performed on 1/30/20 of Probiotic B found that the viability for the 10^{-5} dilution was 2.52×10^9 CFUs/pill (Figure 1C). The experiments performed on 3/14/20 of measuring the viability of Probiotic A and Probiotic B were unsuccessful since no isolated colonies were present. (Figure 1B, D).

Probiotic A had a significantly higher CFU count/mL than Probiotic B ($t = 1.5928 \times 10^{10}$, $p = 1.9984 \times 10^{-11}$) (Figure 2B); Probiotic A had a higher average CFU count (6.223×10^{11} CFUs/pill) than Probiotic B (2.302×10^9 CFUs/pill) (Table 1). In addition, Probiotic A had a CFU count that was significantly different ($X^2 = 357.41$, p-value $= 4.44 \times 10^{-75}$) compared to the label (Figure 2C). Probiotic B had a CFU count that was not significantly different ($X^2 = -3.849$, p-value = 1) compared to the label (Figure 2C); Probiotic B had a lower average accuracy (23.02%) (Table 1). Therefore, Probiotic A had a higher average accuracy than Probiotic B regarding viability (Table 1, Figure 2).

Results for the amplification of 16s rRNA gene

We ran 5µL of the amplified fragment of 16s rRNA gene on five separate gels for Probiotic A, resulting in a total of twenty five individual wells/colonies from Probiotic A. Fourteen wells out of all the gels ran showed successful amplification from PCR. We ran the 16s rRNA gene on five separate gels for Probiotic A, resulting in a total of twenty seven individual wells/colonies from Probiotic B. Thirteen wells out of all the gels ran showed successful amplification from PCR.

Successful amplification was indicated by the absence of a band in the negative control and the presence of a 1500bp band in the samples (Figure 3B, 4B). Not all PCR reactions were successful, as indicated by the lack of the 1500bp product in some wells (Figure 3A, 4A). All successfully amplified products were cleaned up and sent for Sanger sequencing.

Sanger sequences were obtained for 12 independent bacterial colonies from each probiotic. Successful sequences showed clear peaks and nucleotide base identification and were used for BLAST analysis (Figure 4A,C). Sequences with unclear peaks and nucleotide base identification were deemed unsuccessful and were not used for BLAST analysis (Figure 4B,D).

Successful identification for BLAST analysis was determined as any sequence that has a query cover of 100%, an E value of zero or as close to zero as possible, and a percent identity of 100% (nothing lower than 97%). For some samples, we could not distinguish between multiple species. For example, the BLAST sequence shows that *Lactobacillus plantarum* and *Lactobacillus pentosus* were not present, but were possible matches within Probiotic A from Sequence 2 (Figure 5A). The BLAST sequence shows that *Lactobacillus rhamnosus*, *Lactobacillus acidophilus*, *Lactobacillus paracasei*, *Lactobacillus casei*, *Lactobacillus zeae*, and *Lactobacillus fermentum* were present for Probiotic B from Sequence 12 (Figure 5B). For samples where an expected bacterial species was a possibility, we assumed that this was the bacterial identity for all further analysis, and is represented in bold (Table 2).

The bacterial species ratios that were expected within the samples for both Probiotics A and B were 60% *Lactobacillus plantarum*, 20% *Lactobacillus rhamnosus*, 10% *Bacillus coagulans*, and 10% *Lactobacillus salivarus*. The bacteria species ratios that were found within Probiotic A were 58.3% *Lactobacillus plantarum*, 16.7% *Lactobacillus rhamnosus*, 8.33% *Bacillus coagulans*, and 0% *Lactobacillus salivarus*. The bacteria species ratios that were found within Probiotic B were 58.3% *Lactobacillus plantarum*, 25% *Lactobacillus rhamnosus*, 8.33% *Bacillus coagulans*, and 0% *Lactobacillus salivarus.* In addition, Probiotic A contained an unexpected bacterial species, which was identified as a species of the genus *Ralstonia* bacteria. *Both probiotic A and B had samples whose identity could not be determined (A: Uncultured bacterium; B: Uncultured bacillus* and *Uncultured bacterium)* (Table 2).

From this data, it was reported that Probiotics A and B had identical accuracy values for *Lactobacillus plantarum* (97.2%), *Bacillus coagulans* (83.3%), and *Lactobacillus salivarius* (0%). The accuracy differed regarding *Lactobacillus rhamnosus* in which Probiotic B (125%) had greater accuracy than Probiotic A (83.5%). Thus, Probiotic B had a higher overall accuracy than Probiotic A regarding species diversity because Probiotic A was contaminated with bacteria from the Ralstonia genus (Table 2, Figure 6).

Discussion:

We compared two differently priced probiotics that had similar claims with respect to CFU count and species identity. When measuring viability, some experiments were unsuccessful due to not letting the metal spreader cool properly, limiting the number of successful replicates (Figure 2). In future, we could perform more replicates for each probiotic, however, we did have seven successful replicates for Probiotic A and five successful replications for Probiotic B, which allowed us to estimate viability. Despite both claiming to contain 10 billion live bacteria (CFUs) on their label, Probiotic A had a significantly higher CFU count/mL than Probiotic B ($t = 1.5928 \times 10^{10}$, $p = 1.9984\times 10^{-11}$). In addition, Probiotic A had a CFU count that was significantly different ($X^2 = 357.41$, p-value $=4.44 \times 10^{-75}$) compared to the label. The higher than expected CFU count could be due to contamination. Probiotic B had a CFU count that was not significantly different ($X^2 = -\ 3.849$, p-value = 1) compared to the label.

Probiotic A had a higher overall viability than probiotic B and in fact had a higher CFU count as compared to the claim on the label, but higher does not always mean better. In fact, studies have shown that too much probiotics can cause bloating, can be dangerous to severely ill or immunocompromised individuals, and can cause major infections to people who have central venous catheters (Williams 2010). Another concern is that we also don't know the exact number of excess probiotics that would cause these side effects (Haffner *et al.* 2016). Finally, the higher viability could be a reflection of contamination by other probiotic or non-probiotic bacterial species, that are not mentioned on the label. To identify the species present in each probiotic, we compared the 16S rRNA gene sequences of 12 different isolates from each pill.

In identifying bacterial species by comparing the 16S rRNA gene sequence, we often had multiple bacterial species that were identified as possible matches. This is because closely related bacterial species may have an identical 16S rRNA gene sequence, resulting in an identical query cover, E value, and percent identity (Table 2). Due to this similarity in sequence, it may not be possible for BLAST to provide an accurate species ID based solely on the sequence of a single gene. In future studies, we can extend upon this research by sequencing more genes than only the 16S rRNA to see greater differences in the sequences. To account for this implication, the BLAST database is updated daily in attempt to keep up with these evolutionary changes, but this means that query cover, E value, and percent identity values could change daily regarding the same bacterial sequence. Therefore, some discrepancy may exist in important research data values which causes an overarching limitation to the research. In future studies, we can extend upon this research by either getting all our sequencing results from BLAST in one day or if we get the results over multiple days, to take the average of the query cover, E value, and percent identity values.

Despite having similar species diversity claims on the label, Probiotic B had higher species diversity accuracy than Probiotic A (Table 2, Figure 6). Probiotic B has higher species accuracy based on Company 2's label and does not contain any bacterial strains not mentioned on the label. *Ralstonia*, which was found in Probiotic A can be harmful to the human body. *Ralstonia* species may cause infection when found within contaminated infusates or immunocompromised hosts (Steinberg and Burd 2015). This causes major concern regarding the safety of Probiotic A and calls into question the credibility of Company 1 (Figure 5). It also seems to support that the assumed positive relationship between price and quality is valid. Thus, we concluded from our research that Probiotic B, the more expensive probiotic, is the better probiotic based on viability and species diversity accuracy; likewise, Probiotic A of Company 1 is not the better probiotic because it did not include the additional bacterial strain Ralstonia on their label.

The presence of Ralstonia in Probiotic A may be attributed to possible contamination due to experimental error or industrial error within the manufacturing and transportation processes. Experimental human error could be from pipette tip transfer, supplies not being sterile, or improper plating techniques. To account for this, in future, we plan to have more replicates and plate a negative control for each sample to test if the contamination came from human error or from industrial error. Industrial errors could occur during the manufacturing/ transportation process in which protocols that are important to the viability and quality of probiotic supplements could be overlooked (Fenster *et al. 2019*). Manufacturing and storage of the probiotics are essential as the temperatures, pH, and presence of microorganisms can affect the viability before it reaches shelves in the stores (Collins 1998). Probiotic regulation includes reaching GRAS (Generally Regarded as Safe) status as well as being able to survive specific conditions within the body (Collins 1998). Also, probiotic labels do not need preapproval from the FDA, unless specific health claims are listed (Degnan 2008). Therefore, lack of strict regulations from the FDA regarding probiotics as a dietary supplement may also contribute to contamination.

We found that while Probiotic B had a lower viability than expected, it is the safer choice due to the dangers associated with the *Ralstonia* found in Probiotic A. We expected these results since some studies have shown how probiotic products have exceeded their label claim in terms of viable bacteria due to contamination (Weese and Martin 2011). Not only does our results speak volumes on the inconsistency of the FDA guidelines, but it also leaves the mystery of not knowing how much probiotic is too much in terms of looking at percent viability in a given pill. To date, the FDA has no exact therapeutic dosing recommendation (recommended doses range from 10^7-10^{12} CFU/day); this recommendation, however, would all depend on the probiotic genus, species, or even the strain (Haffner *et al.* 2016). Previous research conducted also found that the expensive probiotics showed a higher efficiency in clinical trials and mentioned incorrect information on some labels; some probiotics included strains not mentioned on their labels or had lower amounts of bacterial strains than expected (Ciclitira 2006). This study motivated our research to find if the more expensive probiotic had higher quality; also it guided our expectations to be that the more expensive probiotic would provide better results.

Overall, our results are important because they will better inform consumers on which probiotic is worth buying. Investing in a more expensive probiotic may increase the quality you are getting by limiting the contamination present in the product. There should be better regulation of all probiotic products to ensure that their labels are more accurate corresponding to the bacterial strain components and amounts it includes.

Appendix: Figures and Tables

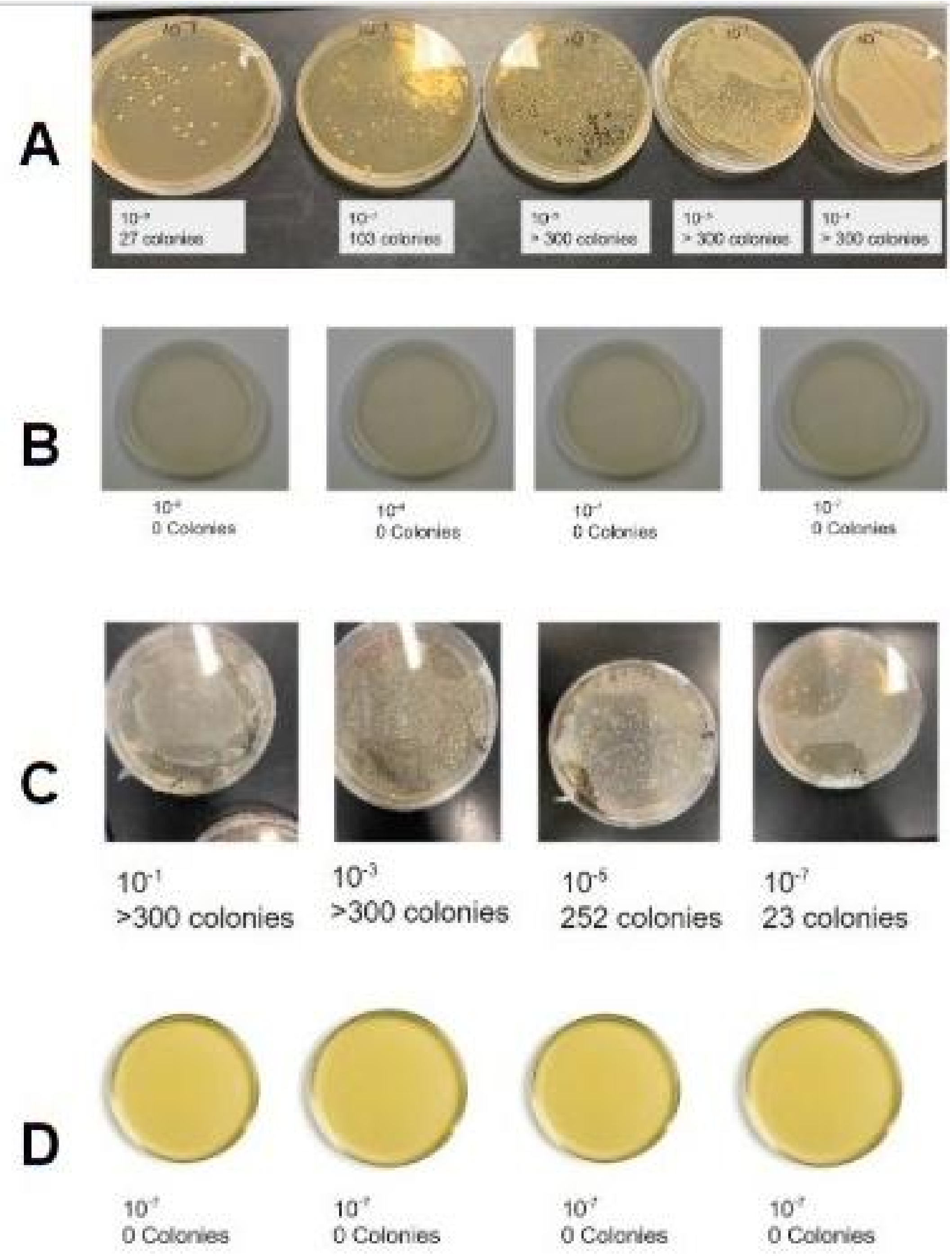


Figure 1: Measuring Viability of Probiotics A and B. The number of isolated colonies for plates with 30-300 colonies was counted because this was our threshold for a successful CFU count. **A.** Serial dilutions (10^{-1} through 10^{-9}) and spread plating (10^{-1}, 10^{-3}, 10^{-5},10^{-7}, and 10^{-9}) were performed on MRS Agar plates on 1/30. Viable CFU/pill counts were then calculated for the

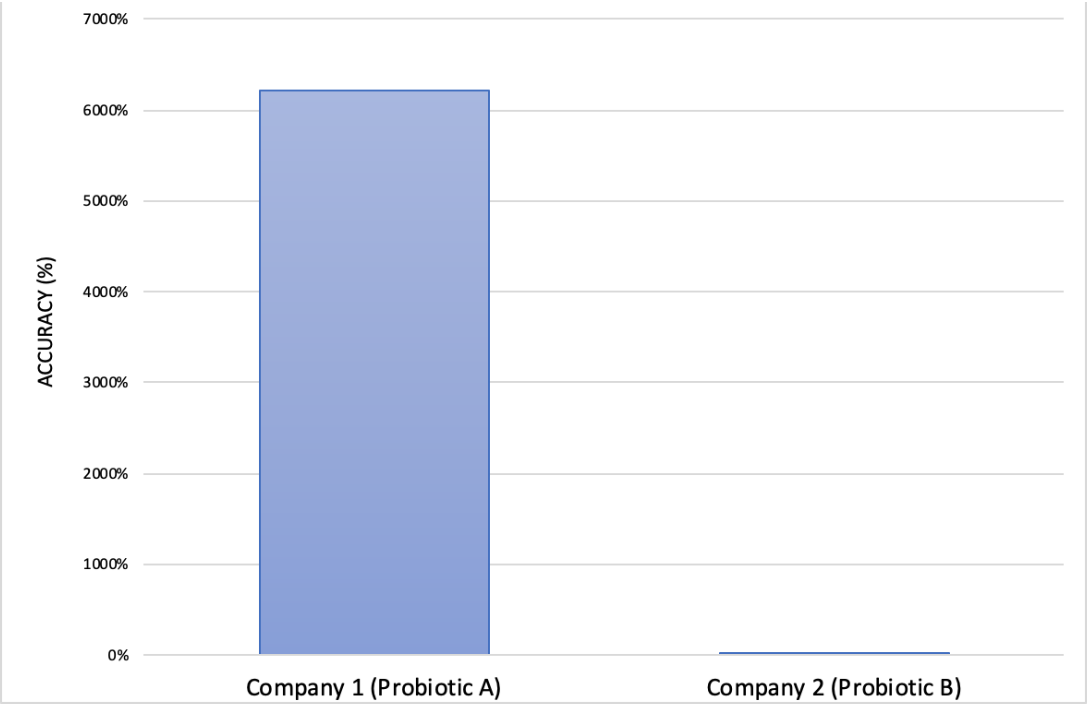
10^{-7} dilution (Table 1A) as 5.15 x 10^{9} CFU/pill. **B.** Serial dilutions (10^{-1} through 10^{-9}) and spread plating (all 10^{-7}) were performed on MRS Agar plates on 3/14. No colonies were seen so the number of isolated colonies for plates could not be counted. **C.** Serial dilutions (10^{-1} through 10^{-9}) and spread plating (10^{-1}, 10^{-3}, 10^{-5} and 10^{-7}) were performed on MRS Agar plates. Viable CFU/pill counts were then calculated for the 10^{-5} dilution (Table 1B) as 2.52 x 10^{9} CFU/pill. **D.** Serial dilutions (10^{-1} through 10^{-9}) and spread plating (10^{-7}) were performed on MRS Agar plates. No colonies were seen so the number of isolated colonies for plates could not be counted.

Table 1: Comparing Viability of Probiotic A and B.

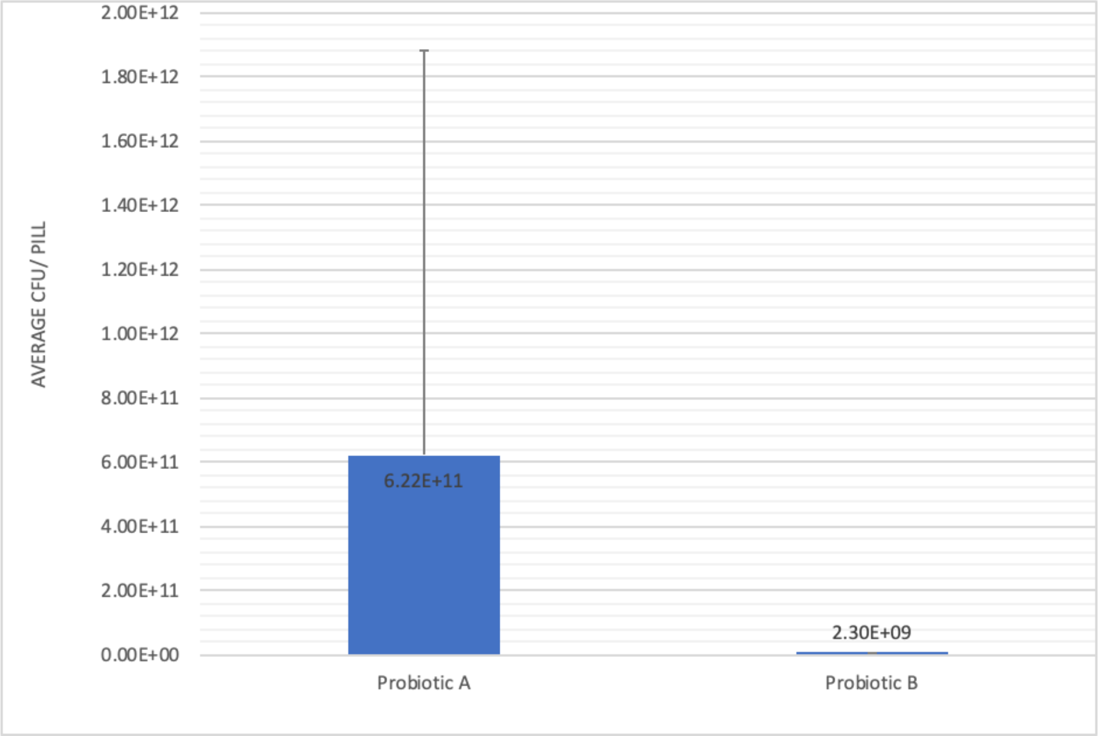
Serial dilutions (10^{-1} through 10^{-9}) and spread plating of the dilutions were performed on MRS Agar plates. Our threshold for successful plating was determined by the number of colonies observed on a given plate within the range 30-300 colonies. These successful plates were included in the calculations of the CFUs/pill and label accuracy values. We found that Probiotic A had a higher average CFU count (6.223×10^{11} CFUs/pill) than Probiotic B (2.302×10^{9} CFUs/pill). Probiotic A had a much higher average accuracy (6223%) than expected; Probiotic B had a lower average accuracy (23.02%) than expected. Therefore, Probiotic A had a higher average accuracy than Probiotic B regarding viability.

<table>
<tr><th colspan="6">Compiled Data for Probiotic A</th></tr>
<tr><th>Date</th><th>Dilution</th><th>Number of colonies observed on plate</th><th>CFUs/mL</th><th>CFUs/pill</th><th rowspan="7">Accuracy</th></tr>
<tr><td>01.30</td><td>10^{-7}</td><td>103</td><td>1.03×10^{10}</td><td>1.03×10^{11}</td></tr>
<tr><td>02.14</td><td>10^{-9}</td><td>32</td><td>3.2×10^{11}</td><td>3.2×10^{12}</td></tr>
<tr><td></td><td>10^{-7}</td><td>96</td><td>9.6×10^{9}</td><td>9.6×10^{10}</td></tr>
<tr><td>02.18</td><td>10^{-7}</td><td>111</td><td>1.11×10^{10}</td><td>1.11×10^{11}</td></tr>
<tr><td></td><td>10^{-7}</td><td>125</td><td>1.25×10^{10}</td><td>1.25×10^{11}</td></tr>
<tr><td></td><td>10^{-7}</td><td>99</td><td>9.9×10^{9}</td><td>9.9×10^{10}</td></tr>
<tr><td colspan="4">Average</td><td>6.223×10^{11}</td><td>6223%</td></tr>
<tr><th colspan="6">Compiled Data for Probiotic B</th></tr>
<tr><th>Date</th><th>Dilution</th><th>Number of colonies observed on plate</th><th>CFUs/mL</th><th>CFUs/pill</th><th rowspan="6">Accuracy</th></tr>
<tr><td>01.30</td><td>10^{-5}</td><td>252</td><td>2.52×10^{8}</td><td>2.52×10^{9}</td></tr>
<tr><td>02.14</td><td>10^{-5}</td><td>222</td><td>2.22×10^{8}</td><td>2.22×10^{9}</td></tr>
<tr><td>02.18</td><td>10^{-5}</td><td>212</td><td>2.12×10^{8}</td><td>2.12×10^{9}</td></tr>
<tr><td></td><td>10^{-5}</td><td>230</td><td>2.30×10^{8}</td><td>2.30×10^{9}</td></tr>
<tr><td></td><td>10^{-5}</td><td>235</td><td>2.35×10^{8}</td><td>2.35×10^{9}</td></tr>
</table>

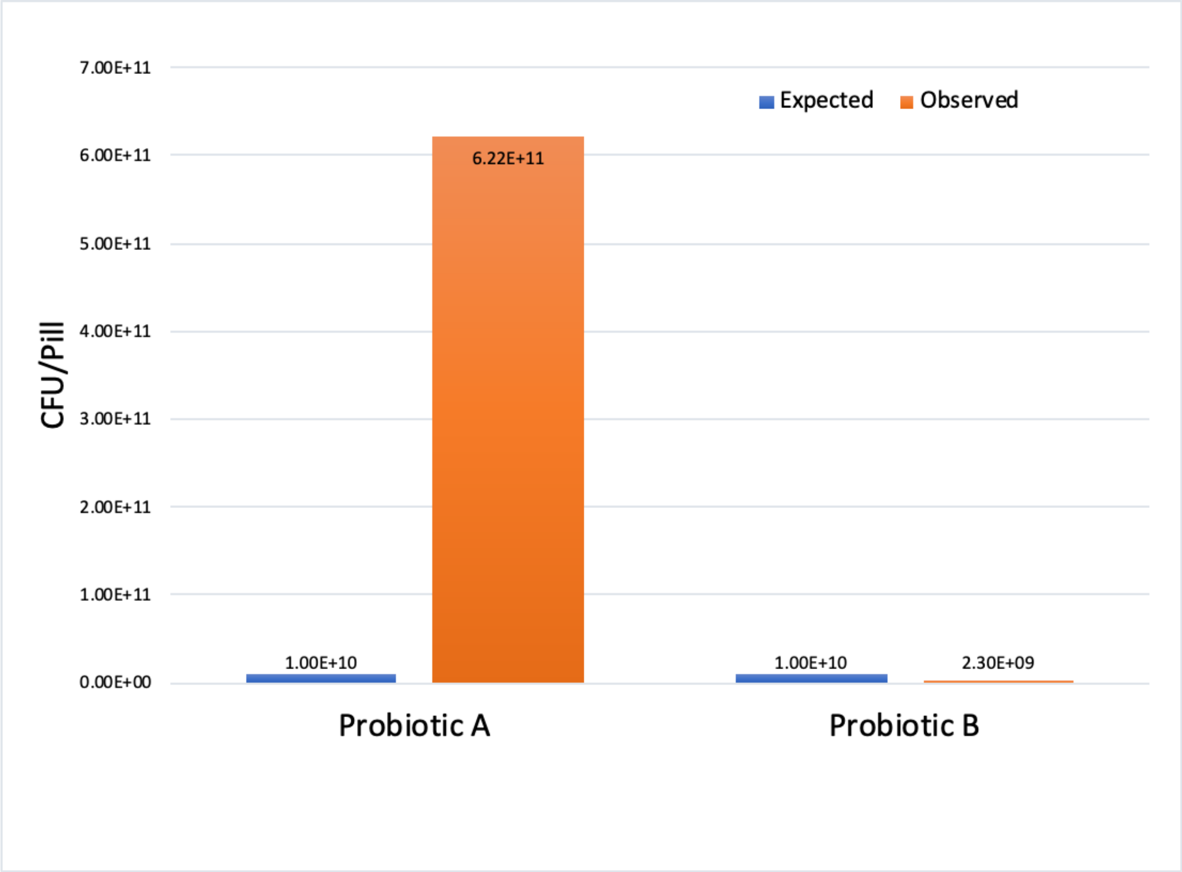
Average	2.302×10^9	23.02%



A.


AVERAGE CFU/ PILL
2.00E+12
1.80E+12
1.60E+12
1.40E+12
1.20E+12
1.00E+12
8.00E+11
6.00E+11
4.00E+11
2.00E+11
0.00E+00
6.22E+11
2.30E+09
Probiotic A
Probiotic B

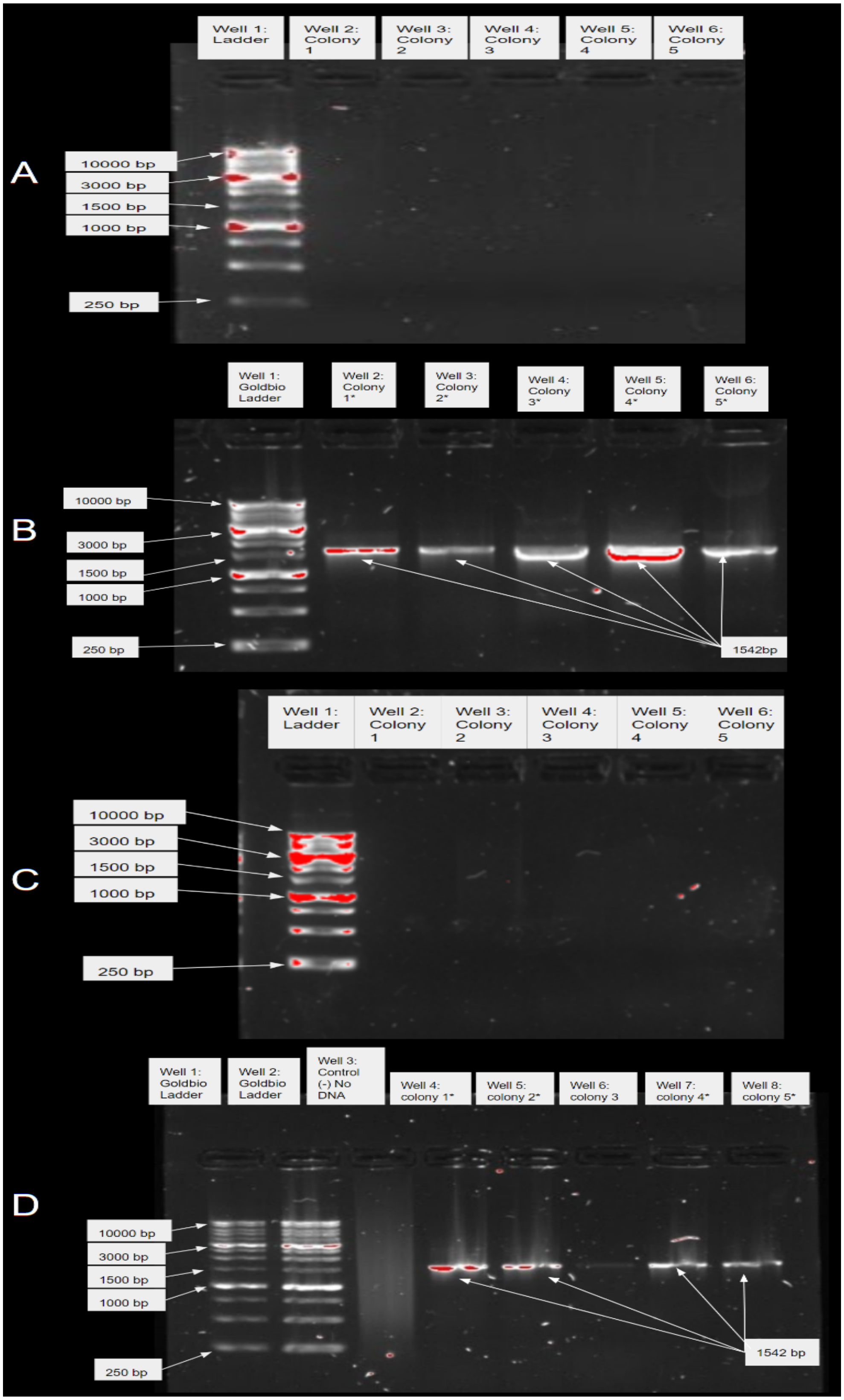
B.



C.

Figure 2: Comparing Viability of Company 1 and Company 2.

A. Plating and calculation of CFUs/pill were used to determine viability. Probiotic A showed to have a higher viability than probiotic B. The average CFU accuracy for both were calculated to be 6223% for Probiotic A and 23.02% for Probiotic B. **B.** A T-test was ran for Probiotics A and B (t-value = 1.5928×10^{10}, p-value = 1.9984×10^{-11}). **C.** A chi-squared goodness of fit was ran for Probiotic A ($X^2 = 357.41$, p-value =4.44×10^{-75}) and Probiotic B ($X^2 = -3.849$, p-value = 1).



*represents successful samples used for PCR cleanup to be sent to sequencing

Figure 3: Using gel electrophoresis to test for successful amplification of the 16s rRNA gene. PCR and gel electrophoresis techniques were performed on the 16s rRNA gene and shown and shown under UV light to differentiate between successful and unsuccessful samples to be used for Sanger Sequencing. Successful amplification of a 5 µL fragment of the 16s rRNA gene is indicated by the absence of a band in the negative control and the presence of a 1500bp band in the samples. The brighter the band, the more DNA is present. To be an unsuccessful sample, a band does not appear on gel at 1500 bp. The 1 kb goldbio ladder was shown as well to help with identification of the gene: the sample bands are matched to the ladder. **A.** Probiotic A colonies 1-5 were pipetted into the corresponding wells (2-6) and the negative control was pipetted into well 7. No visible bands were present on the gel. **B.** The negative control was pipetted into well 3, and Probiotic A colonies 1-5 were pipetted into the corresponding wells (4-8) and showed bands that lined up around the 1500 base pair mark on the kb ladders. Colony 3 showed to contain the least amount of this DNA as it showed the dullest band. **C.** Probiotic B colonies 1-5 were pipetted into the corresponding wells (2-6) and the negative control was pipetted into well 7. No visible bands were present on the gel. **D.** Probiotic B colonies 1-5 were pipetted into the corresponding wells (2-6) and the negative control was pipetted into well 7. Colonies 1-5 showed to be successful as the 16s rRNA gene is 1542 bp pairs and there was a band that lined up with kb ladders around 1500 base pairs for each sample.

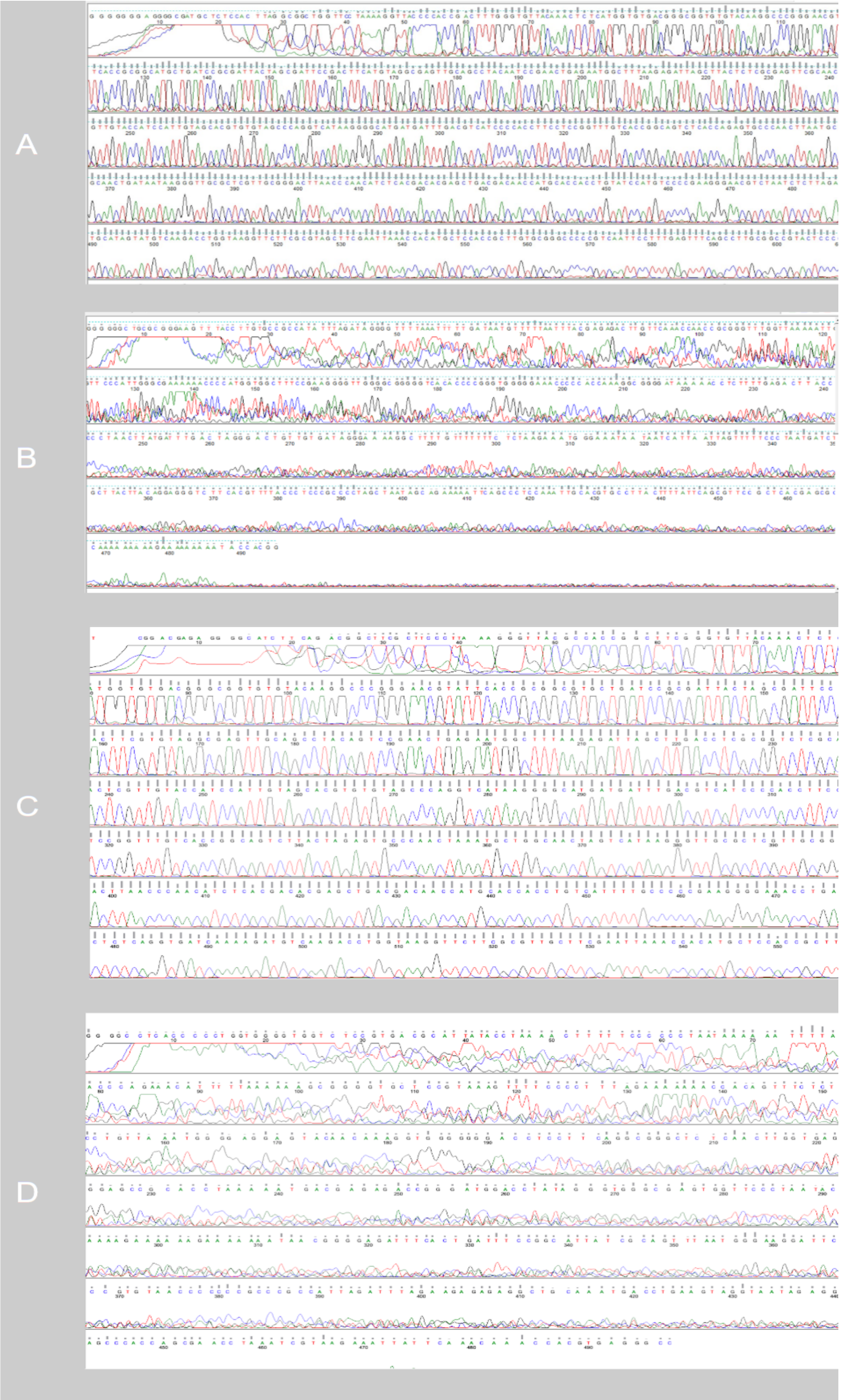

A
B
C
D

Figure 4: Determination of the 16s rRNA gene sequence for Probiotic A and B. Sanger sequencing was performed on samples of Probiotic A and B. The Sanger Sequencing shows light wavelength read outs in the form of different colored peaks. These different colored peaks represent a nucleotide in the DNA sequence: a blue peak is for Cytosine, a green peak is for Adenine, a black peak is for Guanine, and a red peak is for Thymine. **A, C.** These Sanger Sequences for Probiotic A and B respectively, are DNA fragments within the 16s rRNA gene that show clear data or peaks that don't overlap and each peak corresponds to a single nucleotide base. **B, D.** These Sanger Sequences for Probiotic A and B respectively, show the peaks that are overlapping and provide unclear data of individual nucleotide base identification.

```
Lactobacillus plantarum subsp. plantarum strain HBUAS56191 16S ribosomal RNA gene, partial sequence
Sequence ID: MT231813.1  Length: 1489  Number of Matches: 1

Range 1: 1078 to 1342  GenBank  Graphics          ▼ Next Match  ▲ Previous Match

Score            Expect     Identities        Gaps          Strand
490 bits(265)    1e-134     265/265(100%)     0/265(0%)     Plus/Minus

Query  1     AGCGATTCCGACTTCATGTAGGCGAGTTGCAGCCTACAATCCGAACTGAGAATGGCTTTA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1342  AGCGATTCCGACTTCATGTAGGCGAGTTGCAGCCTACAATCCGAACTGAGAATGGCTTTA  1283

Query  61    AGAGATTAGCTTACTCTCGCGAGTTCGCAACTCGTTGTACCATCCATTGTAGCACGTGTG  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1282  AGAGATTAGCTTACTCTCGCGAGTTCGCAACTCGTTGTACCATCCATTGTAGCACGTGTG  1223

Query  121   TAGCCCAGGTCATAAGGGGCATGATGATTTGACGTCATCCCCACCTTCCTCCGGTTTGTC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1222  TAGCCCAGGTCATAAGGGGCATGATGATTTGACGTCATCCCCACCTTCCTCCGGTTTGTC  1163

Query  181   ACCGGCAGTCTCACCAGAGTGCCCAACTTAATGCTGGCAACTGATAATAAGGGTTGCGCT  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1162  ACCGGCAGTCTCACCAGAGTGCCCAACTTAATGCTGGCAACTGATAATAAGGGTTGCGCT  1103

Query  241   CGTTGCGGGACTTAACCCAACATCT  265
             |||||||||||||||||||||||||
Sbjct  1102  CGTTGCGGGACTTAACCCAACATCT  1078
```

```
Lactobacillus rhamnosus strain IGM5-11 16S ribosomal RNA gene, partial sequence
Sequence ID: MT197267.1  Length: 1479  Number of Matches: 1

Range 1: 1039 to 1326  GenBank  Graphics          ▼ Next Match  ▲ Previous Match

Score            Expect     Identities        Gaps          Strand
532 bits(288)    3e-147     288/288(100%)     0/288(0%)     Plus/Minus

Query  1     TTCGTGTAGGCGAGTTGCAGCCTACAGTCCGAACTGAGAATGGCTTTAAGAGATTAGCTT  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1326  TTCGTGTAGGCGAGTTGCAGCCTACAGTCCGAACTGAGAATGGCTTTAAGAGATTAGCTT  1267

Query  61    GACCTCGCGGTCTCGCAACTCGTTGTACCATCCATTGTAGCACGTGTGTAGCCCAGGTCA  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1266  GACCTCGCGGTCTCGCAACTCGTTGTACCATCCATTGTAGCACGTGTGTAGCCCAGGTCA  1207

Query  121   TAAGGGGCATGATGATTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCTT  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1206  TAAGGGGCATGATGATTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGCAGTCTT  1147

Query  181   ACTAGAGTGCCCAACTAAATGCTGGCAACTAGTCATAAGGGTTGCGCTCGTTGCGGGACT  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1146  ACTAGAGTGCCCAACTAAATGCTGGCAACTAGTCATAAGGGTTGCGCTCGTTGCGGGACT  1087

Query  241   TAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTG  288
             ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1086  TAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTG  1039
```

```
Lactobacillus plantarum strain HBUAS56098 16S ribosomal RNA gene, partial sequence
Sequence ID: MT211326.1  Length: 1485  Number of Matches: 1

Range 1: 1031 to 1335  GenBank  Graphics          ▼ Next Match  ▲ Previous Match

Score            Expect     Identities        Gaps          Strand
564 bits(305)    1e-156     305/305(100%)     0/305(0%)     Plus/Minus

Query  1     GATTCCGACTTCATGTAGGCGAGTTGCAGCCTACAATCCGAACTGAGAATGGCTTTAAGA  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1335  GATTCCGACTTCATGTAGGCGAGTTGCAGCCTACAATCCGAACTGAGAATGGCTTTAAGA  1276

Query  61    GATTAGCTTACTCTCGCGAGTTCGCAACTCGTTGTACCATCCATTGTAGCACGTGTGTAG  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1275  GATTAGCTTACTCTCGCGAGTTCGCAACTCGTTGTACCATCCATTGTAGCACGTGTGTAG  1216

Query  121   CCCAGGTCATAAGGGGCATGATGATTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1215  CCCAGGTCATAAGGGGCATGATGATTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACC  1156

Query  181   GGCAGTCTCACCAGAGTGCCCAACTTAATGCTGGCAACTGATAATAAGGGTTGCGCTCGT  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1155  GGCAGTCTCACCAGAGTGCCCAACTTAATGCTGGCAACTGATAATAAGGGTTGCGCTCGT  1096

Query  241   TGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTAT  300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1095  TGCGGGACTTAACCCAACATCTCACGACACGAGCTGACGACAACCATGCACCACCTGTAT  1036

Query  301   CCATG  305
             |||||
Sbjct  1035  CCATG  1031
```

```
Lactobacillus rhamnosus strain IGM1-9 16S ribosomal RNA gene, partial sequence
Sequence ID: MT197200.1  Length: 1471  Number of Matches: 1

Range 1: 1057 to 1325  GenBank  Graphics          ▼ Next Match  ▲ Previous Match

Score            Expect     Identities        Gaps          Strand
497 bits(269)    8e-137     269/269(100%)     0/269(0%)     Plus/Minus

Query  1     TCCGACTTCGTGTAGGCGAGTTGCAGCCTACAGTCCGAACTGAGAATGGCTTTAAGAGAT  60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1325  TCCGACTTCGTGTAGGCGAGTTGCAGCCTACAGTCCGAACTGAGAATGGCTTTAAGAGAT  1266

Query  61    TAGCTTGACCTCGCGGTCTCGCAACTCGTTGTACCATCCATTGTAGCACGTGTGTAGCCC  120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1265  TAGCTTGACCTCGCGGTCTCGCAACTCGTTGTACCATCCATTGTAGCACGTGTGTAGCCC  1206

Query  121   AGGTCATAAGGGGCATGATGATTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGC  180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1205  AGGTCATAAGGGGCATGATGATTTGACGTCATCCCCACCTTCCTCCGGTTTGTCACCGGC  1146

Query  181   AGTCTTACTAGAGTGCCCAACTAAATGCTGGCAACTAGTCATAAGGGTTGCGCTCGTTGC  240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1145  AGTCTTACTAGAGTGCCCAACTAAATGCTGGCAACTAGTCATAAGGGTTGCGCTCGTTGC  1086

Query  241   GGGACTTAACCCAACATCTCACGACACGA  269
             |||||||||||||||||||||||||||||
Sbjct  1085  GGGACTTAACCCAACATCTCACGACACGA  1057
```

```
Ralstonia pickettii strain AJC02 16S ribosomal RNA gene, partial sequence
Sequence ID: MN069645.1  Length: 1462  Number of Matches: 1

Range 1: 34 to 423  GenBank  Graphics          ▼ Next Match  ▲ Previous Match

Score            Expect     Identities        Gaps          Strand
721 bits(390)    0.0        390/390(100%)     0/390(0%)     Plus/Minus

Query  1    AGCCATTTCTTTCCGGACAAAAGTGCTTTACAACCCGAAGGCCTTCTTCACACACGCGGC  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  423  AGCCATTTCTTTCCGGACAAAAGTGCTTTACAACCCGAAGGCCTTCTTCACACACGCGGC  364

Query  61   ATTGCTGGATCAGGCTTTCGCCCATTGTCCAAAATTCCCCACTGCTGCCTCCCGTAGGAG  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  363  ATTGCTGGATCAGGCTTTCGCCCATTGTCCAAAATTCCCCACTGCTGCCTCCCGTAGGAG  304

Query  121  TCTGGGCCGTGTCTCAGTCCCAGTGTGGCTGATCGTCCTCTCAGACCAGCTACTGATCGT  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  303  TCTGGGCCGTGTCTCAGTCCCAGTGTGGCTGATCGTCCTCTCAGACCAGCTACTGATCGT  244

Query  181  CGCCTTGGTGGGCCTTTACCCCACCAACTAGCTAATCAGACATCGGCCGCTCCTATAGCA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  243  CGCCTTGGTGGGCCTTTACCCCACCAACTAGCTAATCAGACATCGGCCGCTCCTATAGCA  184

Query  241  TGAGGCCTTGCGGTCCCCCACTTTCACCCTCAGGTCGTATGCGGTATTAGCTAATCTTTC  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  183  TGAGGCCTTGCGGTCCCCCACTTTCACCCTCAGGTCGTATGCGGTATTAGCTAATCTTTC  124

Query  301  GACTAGTTATCCCCCACTACAGGGCACGTTCCGATGTATTACTCACCCGTTCGCCACTCG  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  123  GACTAGTTATCCCCCACTACAGGGCACGTTCCGATGTATTACTCACCCGTTCGCCACTCG  64

Query  361  CCATCAATCAGCAGAGCTAGATCATGCTGC  390
            ||||||||||||||||||||||||||||||
Sbjct  63   CCATCAATCAGCAGAGCTAGATCATGCTGC  34
```

Figure 5. Successful species identification of sequences for Probiotic A and Probiotic B. Our threshold for successful identification for BLAST analysis is as any sequence that has a query cover of 100%, an E value of zero or as close to zero as possible, and a percent identity of 100%

(nothing lower than 97%). **A.** The BLAST sequence shows that *Lactobacillus plantarum* is present which is expected based on the label for Probiotic A. *Lactobacillus pentosus* and *Lactobacillus fermentum* were two other types of bacteria that the BLAST detected, but this was a species that was not expected. **B.** Probiotic A Sequence 11. The BLAST sequence shows that Lactobacillus plantarum is present which is expected based on the label for Probiotic A. *Lactobacillus rhamnosus, Lactobacillus paracasei, Lactobacillus fermentum, Lactobacillus casei, Lactobacillus zeae, and Lactobacillus acidophilus* were also present but these species were not expected based on Company 1's label. **C.** Probiotic Sequence 3. The BLAST sequence shows that *Lactobacillus plantarum* is present which is expected based on the label for Probiotic B. *Lactobacillus pentosus* was another type of bacteria that the BLAST detected, but this was a species that was not expected. **D.** Probiotic B Sequence 12. The BLAST sequence shows that *Lactobacillus rhamnosus* is present which is expected based on the label for Probiotic B. *Uncultured bacterium, Lactobacillus paracasei, Lactobacillus casei, Lactobacillus zeae, and Lactobacillus fermentum* are others type of bacteria that the BLAST detected, but these species were not expected based on Company 2's label. **E.** Probiotic A Sequence 10. The BLAST sequence shows that *Ralstonia bacteria, Uncultured ralstonia,* and *Uncultured burkholderiales* are present and these species were not expected based off of Company 1's label.

Table 2: Species Identification and Diversity.

BLAST was used for species identification; our threshold for successful identification for BLAST analysis is a query cover of 100%, an E value of zero or as close to zero as possible, and a percent identity of 100% (nothing lower than 97%). 12 samples for each probiotic were sequence and the results of the possible species identification are shown below. Due to consistent Query Cover, E Value, and Percent Identity values, the bacterial genus is known, but the species remains unknown. For samples where an expected bacterial species from the label was a possibility, we assumed that this was the bacterial identity for all further analysis, and is represented in bold. In addition, Probiotic A contained an unexpected bacterial species, which was identified as a species of the genus *Ralstonia* bacteria. *Both probiotic A and B had samples whose identity could not be determined (A: Uncultured bacterium; B: Uncultured bacillus* and *Uncultured bacterium).* From these observed species identifications, we then compared label accuracy for both Probiotic A and Probiotic B. Expected bacterial percentages from the probiotic labels were found by dividing each bacterial count by the total expected CFU/pill: 1.0 x 10^{10}. We then found the observed percentage for each bacterial genus by dividing the number of BLAST sequencing files containing that bacterial species for each probiotic. We calculated the label accuracy for each from dividing the observed percentage by the expected percentage and multiplying by 100.

Probiotic A				
Sequencing file	Species type	Query Cover	E Value	Percent Identity
1	***Lactobacillus plantarum****, pentosus, fermentum*	100%	0	100%
2	***Lactobacillus plantarum****, pentosus, fermentum*	100%	0	100%
3	*Bacillus bacteria*	100%	3e-121	100%
4	***Lactobacillus rhamnosus****,*	100%	9e-111	100%

	paracasei, fermentum, zeae casei			
5	***Lactobacillus plantarum****, pentosus, paraplantarum, fermentum*	100%	6e-123	100%
6	N/A	N/A	N/A	N/A
7	***Lactobacillus plantarum****, pentosus, fermentum*	100%	2e-128	100%
8	***Lactobacillus plantarum****, pentosus, paraplantarum, fermentum*	100%	4e-94	100%
9	***Lactobacillus plantarum****, pentosus*	100%	6e-123	100%
10	*Ralstonia bacteria, Uncultured ralstonia, Uncultured burkholderiales*	100%	3e-131	100%
11	***Lactobacillus rhamnosus****, paracasei, fermentum, casei, zeae, acidophilus*	100%	2e-158	100%
12	***Lactobacillus plantarum****, pentosus, paraplantarum, fermentum,*	100%	2e-102	100%
Probiotic B				
Sequencing file	Species type	Query Cover	E Value	Percent Identity
1	N/A	N/A	N/A	N/A
2	***Lactobacillus plantarum****, pentosus*	100%	0	100%

3	***Lactobacillus plantarum**, pentosus,*	100%	1e-150	100%
4	***Lactobacillus rhamnosus**, paracasei, fermentum, casei, zeae; Uncultured bacterium*	100%	0	100%
5	***Lactobacillus plantarum**, pentosus, fermentum*	100%	0	100%
6	***Bacillus coagulans**; Uncultured bacillus*	100%	5e-134	99.63%
7	***Lactobacillus plantarum**, pentosus, paraplantarum, fermentum*	100%	5e-160	100%
8	***Lactobacillus plantarum**, pentosus, fermentum*	100%	1e-175	100%
9	***Lactobacillus plantarum**, pentosus*	100%	2e-132	100%
10	***Lactobacillus rhamnosus**, paracasei, fermentum, casei, acidophilus, zeae; Uncultured bacterium*	100%	0	100%
11	***Lactobacillus plantarum**, pentosus, fermentum*	100%	4e-150	100%
12	***Lactobacillus rhamnosus**, paracasei, casei, fermentum, zeae; Uncultured bacterium*	100%	2e-164	100%
Quantified Bacterial Species				

Bacteria Species Types	Probiotic A/B: Expected	Probiotic A: Observed	Label Accuracy for Probiotic A	Probiotic B: Observed	Label Accuracy for Probiotic B
Lactobacillus plantarum	60 %	58.3%	97.2 %	58.3%	97.2%
Lactobacillus rhamnosus	20%	16.7%	83.5%	25%	125%
Bacillus coagulans	10%	8.33%	83.3%	8.33%	83.3%
Lactobacillus salivarus	10%	0%	0%	0%	0%
*Due to consistent Query Cover, E Value, and Percent Identity values, the bacterial genus is known, but the species remains unknown.					

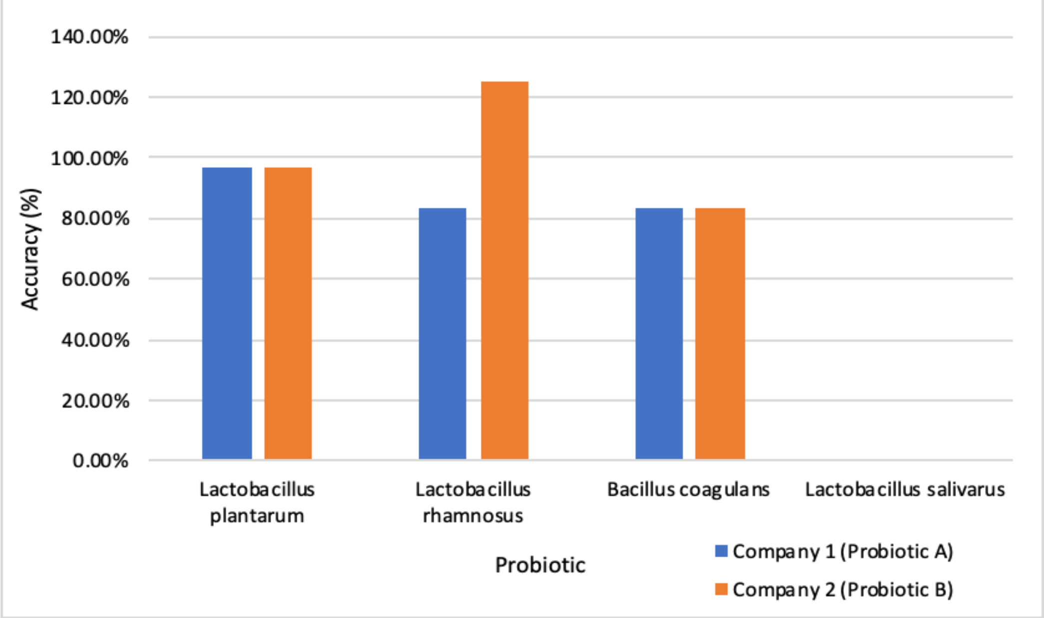


Figure 6: Comparing Species Diversity of Company 1 Versus Company 2. The BLAST database was used for bacterial identification. We plotted the expected bacterial genus and species from the label on the x axis and the accuracy percentage on the y axis for both probiotics based on data from Table 2. The label accuracies for Probiotic A and Probiotic B were similar; however, Probiotic B compared to Probiotic A had a higher overall accuracy regarding *Lactobacillus rhamnosus* determined by using BLAST and comparing expected versus observed values (Table 2). There were bacteria found in sequencing files that were not expected based on the label for both Company 1 and 2. Unexpected species that were not included on the label were not included in the graph because we were only concerned with focusing on bacterial species on each company's label. Probiotic A contained the following unexpected bacterial species: *Lactobacillus pentosus, Lactobacillus fermentum, Lactobacillus paracasei, Lactobacillus fermentum, Lactobacillus zeae casei, Lactobacillus paraplantarum, Ralstonia bacteria, Uncultured ralstonia, Uncultured burkholderiales,* and *Lactobacillus acidophilus.* Probiotic B contained the following unexpected bacterial species: *Lactobacillus pentosus, Lactobacillus paracasei, Lactobacillus fermentum, Lactobacillus casei, Lactobacillus zeae, Uncultured bacterium, Uncultured bacillus,* and *Lactobacillus acidophilus.*

References:

Bujalance C, Jimenez-Valera M, Moreno E, Ruiz-Brave A (2006). A selective differential medium for *Lactobacillus plantarum. Journal of Microbiological Methods* **66:**572-575.

Ciclitira PJ, Clark Sk, Nicholls RJ, Tekkis PP (2006). Effective probiotic treatment is rarely cheap. *BMJ* **333**:1006-8.

Collins J.K, Thornton G, Sullivan G.0 (1998). Selection of probiotic strains for human applications. *International Dairy Journal* **8**:487-490.

Degnan F (2008). The US food and drug administration and probiotics: regulatory categorization. *Clinical Infectious Disease* **46**:S133-6.

Engen PA, Green SJ, Voigt RM, Forsyth CB, and Keshavarzian A (2015). The gastrointestinal microbiome: alcohol effects on the composition of intestinal microbiota. *Alcohol Research* **37**:223-236.

Fenster K, Freeburg B, Hollard C, Wong C, Laursen RR, and Ouwehand AC (2019). The production and delivery of probiotics: a review of a practical approach. *Microorganisms* **10:**3025.

Goldin BR (1998). Health benefits of probiotics. *British Journal of Nutrition* **80**:S203-S207.

Haffner FB, Diab R, and Pasc A (2016). Encapsulation of probiotics: insights into academic and industrial approaches. *AIMS Materials Science* **3**:114-136.

Prasad J, Gill H, Smart J, and Gopal PK (1998). Selection and characterisation of *Lactobacillus* and *Bifidobacterium* strains for use as probiotics. *International Dairy Journal* **8**:993-1002.

Sohrabvandi S, Razavi SH, Mousavi SM, and Mortazavian AM (2010). Viability of probiotic bacteria in low-alcohol and non-alcoholic beer during refrigerated storage. *The Philippine Agricultural Scientist* **93**:24-28.

Steinberg JP and Burd EM (2015). Other gram-negative and gram-variable bacilli. pp. 2667-2683 in Bennett JE, Dolin R, and Blaser MJ, eds., Mandell, douglas, and bennett's principles and practices of infectious diseases. Elsevier Inc, Amsterdam.

Wassenazar TM and Klein G (2008). Safety aspects and implications of regulation of probiotic bacteria in food and food supplements. *J Food Prot* **71**:1734-1741.

Weese JS and Martin H (2011). Assessment of commercial probiotic bacterial contents and label accuracy *The Canadian Veterinary Journal.* **52**:43-46.

Williams NT (2010). Probiotics. *Am J Health Syst Pharm* **67**:449-458.

Woese CR and Fox GE (1977). Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proceedings of the National Academy of Science of the United States of America* **11**:5088-5090.