DNA Barcoding the Billions of Bacteria in Probiotics

Michelle Chen[1], Yumna Elhassa[2], Jeanmaire Molina (mentor)[3]

[1]Stuyvesant High School, [2]Brooklyn Technical High School, [3]Long Island University-Brooklyn

ABSTRACT

Probiotics are dietary supplements containing live bacterial strains supposedly beneficial to humans. The probiotic market has been booming in recent years. Currently, the US Food and Drug Administration (FDA) does not regulate most of these products. The lack of FDA regulation has led to some products not containing their declared species. This presented an opportunity for students, like us, to study and validate the bacterial species included in such products. As part of the Urban Barcode Research Program (UBRP), which is an education initiative to increase scientific research among New York city youth, we have authenticated bacterial species in probiotic supplements sold in drug stores in New York City. We used next-generation sequencing technology called metagenomic sequencing to identify, or barcode, bacterial species in probiotic products. After analyzing our data we found that 4 of the 5 probiotic brands had bacterial species that were not on the labels of their products.

INTRODUCTION

- The probiotics market is part of the US dietary supplement industry, which is expanding rapidly into a $30-billion global industry (Patro et al., 2016). Millions of Americans have been using probiotics (NCCIH, 2017) partly due to manufacturer claims that probiotics could enhance the well-being of an individual (de Simone, 2018). Bacterial species that are considered "probiotic" are believed to boost immunity, prevent and treat gastrointestinal infections, alleviate allergies, relieve constipation, etc. (NCCIH, 2018).
- Consequently, as scientists have begun to learn and publish more about probiotics and their potential benefits, the production of probiotic supplements, often available over-the-counter, has similarly increased.
- Many probiotics are sold as dietary supplements, which do not require FDA approval before they are marketed (NCCIH, 2018). This is an issue because probiotics that are not carefully regulated can possibly contain additives, substitutions, or microbial quantity lower than claimed quantity, unknown to the average consumer.
- To combat this, the FDA has developed the "GutProbe" DNA microarray (Patro et al. 2015), which is able to detect lot-to-lot differences in bacterial species within the same probiotic brand. Though it has been developed in 2015, probiotic manufacturers have yet to implement the technology. Thus, unscrupulous practices may still prevail, which is why we are interested in the authentication of probiotics, since we, and our families, personally consume them.
- As part of the Urban Barcode Research Program (UBRP), which seeks to promote scientific research among New York City youth, we aimed to identify, or barcode, bacterial species in probiotic supplements using a next-generation sequencing technique called metagenomic sequencing

MATERIALS & METHODS

- Five over-the-counter probiotics were purchased from various drugstores in New York City.
- We used the QIAamp PowerFecal DNA Kit (Qiagen Cat No./ID: 12830-50) workflow for DNA extraction. PCR amplification, library and template preparation, and sequencing were performed following the manufacturer's protocol for the Ion 16S metagenomics kit (catalog no. A26216). The kit uses two sets of primers that amplify the hypervariable regions of the 16S rDNA gene in bacteria and allows for bacterial identification at the genus or species level.
- Libraries were made from the amplified fragments and quantified using qPCR. The libraries were diluted for template preparation to allow only one type of template DNA to attach to Ion Sphere Particles (ISPs) for further amplification using emulsion PCR. The templated ISPs were then sequenced using the Ion PGM Sequencing 400 Kit (catalog no. A30044) on the Ion PGM platform housed at the Biology Department of Long Island University-Brooklyn. After the run, PGM generated a "Run Summary" file (Fig. 1).
- Genomic sequencing reads were downloaded and assembled into contigs using Geneious R11 (Biomatters Ltd.). Consensus sequences (>50 bp) of resulting contigs were then blasted against Genbank using Megablast (Morgulis et al. 2008), limiting hit results to 10 each per consensus. If species hits were unanimous or majority were identical, the match was taken to be the identity of the probiotic species. If hits were equivocal (=multiple species matching) then phylogenetic analysis was conducted, first by aligning the sequences of the hits with the query sequence using MAFFT (Katoh et al. 2017) then reconstructing the phylogeny using FastTree (Price et al. 2010). The most closely related species to the query (>90%) was assumed to be the query's taxonomic identity.

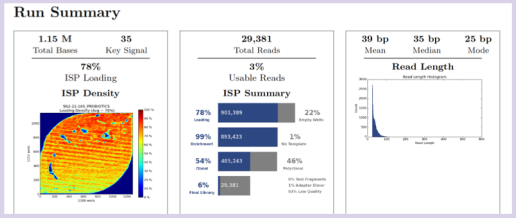

Fig. 1. Summary report from the PGM run. ISP (ion sphere particles) loading density was excellent with 78% wells loaded. Enrichment was superb with 99% of ISPs with template attached. However, 46% of the ISPs were polyclonal (i.e. different templates attached to same ISP which ideally should only be one type of template; acceptable <30%). Library preparation was poor at 6% (should be at least 90%) with mean read length at 39bp.

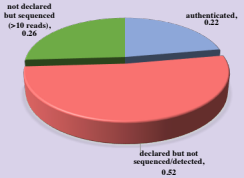

Fig. 2. Proportion of bacterial species in the 5 probiotic brands. Majority of the declared bacterial species were not detected perhaps due to technical errors. However, surprisingly, there were also undeclared species that were also sequenced.

Brand Name	Sequenced Species	Is sequenced species declared on label?	Other declared species but not sequenced/detected
TUU (4 spp. expected)	*Bifidobacterium longum*	*Yes*	*B. bifidum, L. gasseri*
	Lactobacillus acidophilus	*Yes*	
	Lactobacillus helveticus	*No*	
CTA (12 spp. expected)	*Lactobacillus plantarum*	*Yes*	*L. rhamnosus GG, L. paracasei, L. acidophilus, B. lactis*
	Lactobacillus reuteri	*Yes*	*L. salivarius, B. infantis, L. casei, B. bifidum, B. breve, B. longum*
RP (2 spp. expected)	*Lactobacillus paracasei*	*No*	*L. reuteri*
	Lactobacillus rhamnosus	*Yes*	
	Lactobacillus plantarum	*No*	
CTL (1 sp. expected)	*Lactobacillus paracasei*	*No*	*L. rhamnosus GG*
	Lactobacillus casei	*No*	
	Lactobacillus plantarum	*No*	
SD (1 sp. expected)	*Bacillus coagulans*	*Yes*	
	Staphylococcus aureus	*No*	
	Bacillus sp.	*Yes*	

Table 1. Probiotic products sampled and the bacterial species detected in each. Actual brand names have been withheld to prevent legal consequences from manufacturers.

RESULTS & DISCUSSION

- We attempted to barcode 5 different brands of probiotic supplements using next generation sequencing. Unfortunately, we did not generate enough sequencing reads. An ideal sequencing run with an Ion 314 chip produces 400,000 sequence reads. However, we were only able to retrieve 3941 reads which is ~1% of the ideal number of reads. This may be due to errors during our library preparation.
- Regardless, we were able to obtain results and ascertain some of the probiotic species contained in the supplements (Table 1; Fig. 1). Out of 27 bacterial species we detected in 5 different brands, 6 species were authentic (22%), while 14 expected species were undetected (52%, Fig. 2). However, we only had 1% of the expected sequence reads, which may be why these were undetected even if they were present in the sample. Repeating the PGM experiment to generate the ideal number of sequence reads would allow us to determine if indeed manufacturers did not include some bacterial species.
- There were 7 foreign bacterial species (26%) that were detected and may be attributed to contamination, perhaps during our lab work and/or in the probiotic manufacturing facility. In TUU probiotic, *L. helveticus* was detected but is not a declared species. However, bacterial species names often have synonyms, thus we used http://www.bacterio.net/, an archive of bacterial species and their synonyms, to check. *L. helveticus* was not a synonym of any of the declared species. We discovered that it is commonly used in the fermentation of cheeses (Taverniti and Guglielmetti, 2012), which leads us to the conclusion that a human food contamination could have occurred either from one of us, or the workers at the manufacturing facility who did not follow sanitary protocols. *L. helveticus* has also been demonstrated to be a probiotic for gastrointestinal health (Taverniti and Guglielmetti, 2012), which could also possibly mean that the manufacturing lab may have mislabeled/substituted the bacterial species.
- For CTL and RP, the expected species is *L. rhamnosus*, but we identified other similar species, namely *L. casei, L. paracasei, L. plantarum*. Interestingly *L. rhamnosus* was formerly known as *L. casei*, but it is now considered a separate species (http://www.bacterio.net/). It could be that the names have not been updated in Genbank, or that the manufacturer had actually used other species within the Lactobacillus casei group (LCG). *L. rhamnosus, L. casei*, and *L. paracasei* all fall under LCG—very closely related species used for similar purposes in commercial food production and medicine (Hill et al. 2018). DNA barcoding of the 16S gene is not the most reliable way to distinguish between species of LCG due to their high similarity. Since we used 16S in the experiment, this may explain why other LCG species matched. However, detecting *L. plantarum* instead of an LCG species, is surprising and leads us to think that this was a contamination or substitution.
- The results for SD probiotic showed detection of *Staphylococcus aureus* which is a bacteria commonly found on human skin (US Department of Health and Human Services, 2009), indicating there may have been contamination, which may have occurred during our lab work. It is also possible that *S. aureus* contamination occurred during the manufacturing process, which is disconcerting as this bacterial species can produce a toxin that can cause food poisoning (CDC, 2018).
- Though we were unable to authenticate 14 other declared bacterial species in the supplements, we, as UBRP participants have benefited in several ways—we learned the process of next generation sequencing using the Ion Torrent platform and learned to collaborate with one another. We also learned how to evaluate and analyze data, and critically think of possible explanations for why we have obtained the results that we have. Though we committed technical mistakes during the experiment, the fact that we were still able to detect undeclared bacterial species make us wary of consuming probiotic supplements, After this experience, we came to realize that we have to be more careful in selecting and consuming these over-the-counter supplements, inspiring us to be more pharmacovigilant (Molina et al. 2018) in the absence of stringent FDA regulation for such products.

ACKNOWLEDGMENTS

We thank UBRP of CSHL, Christine Marizzi, the Pinkerton Foundation, and the Biology Dept. of LIU Brooklyn for support and sponsorship. We would also like to thank Noor Mohammad of Midwood HS and LIU graduate students, Malini and Urdhva, for their help in lab work.

REFERENCES

- Morgulis, A., et al. (2008). *Bioinformatics (Oxford, England)*, *24*(16), 1757–1764. doi:10.1093/bioinformatics/btn322
- Center for Disease Control/CDC 2018. https://www.cdc.gov/foodsafety/diseases/staphylococcal.html
- Hill, D., et al. 2018. *Frontiers in Microbiology* 9: 2107. doi:10.3389/fmicb.2018.02107
- Katoh, K., et al. *Briefings in Bioinformatics*. https://doi.org/10.1093/bib/bbx108
- Molina J., et al. 2018. DNA barcoding of online herbal supplements: crowd-sourcing pharmacovigilance in high school. *Open Life Sciences* 13:48-55.
- NCCIH. 2017. nccih.nih.gov/research/statistics/NHIS/2012/natural-products/biotics
- NCCIH. 2018. nccih.nih.gov/health/probiotics/introduction.htm#hed4.
- Patro, J.N., et al. 2015. *Journal of Applied Microbiology* 118: 1478-1488.
- Patro, J.N., et al. 2016. *mSphere* 1(2) e00057-16.
- Price, M.N., et al. 2010. PLoS ONE, 5(3):e9490.
- Taverniti, V. & Guglielmetti, S. (2012). *Frontiers in microbiology*, *3*, 392. doi:10.3389/fmicb.2012.00392
- US Department of Health and Human Services. 2009. https://www.foodsafety.gov/poisoning/causes/bacteriaviruses/staphylococcus/index.html.