Spring 2011

Microbiology of Bottled Water: A Molecular View

Reece Gesumaria

University of Colorado Boulder

Follow this and additional works at: http://scholar.colorado.edu/honr_theses

Recommended Citation

Microbiology of Bottled Water:  
A Molecular View

Reece Gesumaria

Molecular, Cellular, and Developmental Biology  
University of Colorado at Boulder

Thesis Advisor: Norm Pace (MCDB)

Thesis Committee Members:  
Norm Pace (MCD Biology)  
Diana Nemergut (Environmental Studies)  
Christy Fillman (MCD Biology)

Oral Defense: Monday, April 4, 2011
Abstract

Americans drink over 23 million gallons of bottled water every day, generating approximately 36 billion bottles annually. The false perception of the purity and cleanliness of expensive bottled water compared to cheap tap sources does not concur with scientific evidence. Many studies have demonstrated the presence of coliforms and heterotrophic bacteria in bottled water, and detected these organisms in counts greatly exceeding the contamination standards set for human consumption. While bacteria have been isolated from bottled water by classic microbiological culture-based methods, these techniques are capable of detecting only a subset of the true microbial constituents. This study analyzes the microbial assemblages and bacterial load of bottled water from two different sources (municipal and spring) using culture-independent molecular techniques. Data collected from 16S ribosomal RNA gene sequencing and DAPI-stained cell counts demonstrate a correlation between different water sources and the unique and reproducible bacterial quality and load among individual brands. The sequences generated from bottled water samples bear identity to bacteria found in freshwater aqueous environments and humans; some sequences correlate with known pathogens.
Introduction

Americans consume over 23 million gallons of bottled water every day, yet there are no standards for the quality of this product\textsuperscript{1,2}. The National Primary Drinking Water Regulations are the legally enforceable standards of the Environmental Protection Agency (EPA) that apply specifically to public water systems\textsuperscript{3}. These primary standards are intended to protect public health by limiting the levels of contaminants, including bacteria, in public drinking water from municipal systems. The Food and Drug Administration is responsible for protecting public health by supervising and regulating the goods consumed by the public. However, this regulatory agency has not set a limit for the heterotrophic bacteria counts in bottled drinking water. Unlike the stringent and enforceable regulations placed on U.S. municipal tap water by the EPA, the health standard for bottled water is merely voluntary\textsuperscript{4}. Distinct from municipal water regulations, there is no disinfection requirement, nor required testing for \textit{E.coli}, \textit{Cryptosporidium}, \textit{Giardia}, fecal coliform, or viruses. There is no requirement for contamination tests to be done in certified labs, and no obligation to report contamination violations. Most importantly, there is no requirement to inform the consumer of any form of contamination\textsuperscript{5}.

The United States Pharmacopeia (USP) established a maximum of 500 colony forming units (cfu) per milliliter in drinking water, yet this standard is not enforced for bottled sources\textsuperscript{6}. On top of the mounting evidence for the presence of radioactive isotopes, pharmaceuticals, disinfection byproducts, heavy metals, fertilizer residue, plasticizers, and other industrial chemicals\textsuperscript{4}, there have been many studies that have identified microbial presence in supposedly clean bottled water. Studies that have tested for the bacterial load in bottled water using the Heterotrophic Plate Count (HPC)\textsuperscript{4,7,8} or Plate Count Agar (PCA)\textsuperscript{9} approach found that the vast majority of the samples exceeded the legal limit by several orders of magnitude. While this number is quite high, studies have demonstrated that
HPC testing already greatly underestimates direct cell counts by a factor of 500- to 15000-fold\textsuperscript{10,11}. Thus, the false perception of the absolute purity and cleanliness of expensive bottled water compared to cheap tap sources does not concur with scientific evidence.

The current understanding of water microbiology, specifically regarding the ensemble of microbes in drinking water, has been largely informed from studies using culture-dependent techniques\textsuperscript{7,10,12-18}. Culturing methods used to assess microbial assemblages are now considered biased and provide an inadequate analysis of the microbiology. This is because the vast majority of bacteria in natural environments are as yet uncultured\textsuperscript{19}. The field of microbiology has undergone profound technological developments in the approach of microbial detection and compositional analysis. Culture-independent studies, including in situ hybridization\textsuperscript{20,21}, real time PCR\textsuperscript{22}, restriction length polymorphism analysis (RFLP)\textsuperscript{23} and microarrays\textsuperscript{24} have been performed and these shed light on the microbiology of environmental water samples.

Previous studies have seen Betaproteobacteria\textsuperscript{23,33} and Alphaproteobacteria\textsuperscript{21} as the dominant organisms in bottled water samples. The culture-independent techniques used in these studies do not explore the full microbial diversity of these samples because these technologies are targeted towards the detection of specific organisms. The analyses performed have not characterized the microbiology of bottled drinking water at the 16S rRNA gene sequencing level.

Studies which have examined the microbiology of bottled water from retail outlets have all found high total colony counts\textsuperscript{25}. The National Resources Defense Council tested over 100 brands of bottled water and found that one in three contained significant contamination levels of chemical or bacterial contaminants in at least one test\textsuperscript{5}. Seventeen percent of the brands tested contained more bacteria than permitted under microbiological-purity guidelines based on heterotrophic plate count bacteria levels. A similar study detected significant bacterial contamination in 40\% of the brands
tested, and isolated 38 chemical contaminants, including chloroform and arsenic, from 10 major brands of bottled water\(^4\). Although informative, these culture-dependent approaches do not provide significant insight into phylogenetic representation.

Another culture-dependent study found that both the microbial load and composition in bottled mineral water is inferior to direct tap sources\(^{16}\). Researchers demonstrated that 76.6% of bottled samples and 36.4% of municipal samples were contaminated by at least one coliform and/or at least one pathogenic bacterial strain\(^{26}\). The study also identified *Escherichia coli* and fecal streptococci in 6.4% and 9.0%, respectively, of bottled samples; these microbes were undetected in municipal samples. Median HPC data were notable in tap and bottled water samples at 95 and 2.3 x \(10^4\) cfu mL\(^{-1}\), respectively. These results are comparable to other studies where HPC and DAPI-staining cell counts of bottled mineral water were approximately \(10^4\) cfu mL\(^{-1}\) and \(1.7 \times 10^5\) cells per mL\(^{23,24}\).

Water treatment is considered the single most important and direct means of influencing the quality, growth rate, and composition of the microbial assemblage of drinking water\(^{12,14,16}\). Treatment should remove the majority of microbes. Additionally, investigators have demonstrated the potential for bacteria to grow in spring water for up to three weeks after bottling if the microbes are not adequately removed, resulting in microbial loads of up to \(10^4\) bacteria mL\(^{-1}\) at 37 °C\(^{12,14,16}\). This high microbial load can be explained by the utilization of nutrients naturally present in the water. Treatment to remove these nutrients would limit microbial growth, again supporting the principle of the significance of water treatment on microbiology.

There are numerous approaches used in the filtration and purification of bottled water, including filtration, UV, ozonation, and reverse osmosis. Filters vary significantly in size, resulting in the selective removal of sediments, particles, microorganisms, etc. For example, an Absolute Filter is
a filter capable of removing all solid particles above the specified micron size for that filter\textsuperscript{26}. Another type of filter commonly used is an activated carbon filter. These filters remove many dissolved solvents, including chlorine, sediment, and volatile organic compounds, but do not remove bacteria. The efficiency of activated carbon filtration is influenced by the amount of carbon in the unit and the amount of time the contaminant spends in contact with it\textsuperscript{27}.

Ultraviolet radiation treatment has been shown to be an effective mechanism of water disinfection with adequate exposure. However, it has been demonstrated that this mechanism’s effectiveness decreases as turbidity increases\textsuperscript{28}. Ozonation disinfects and purifies by infusing ozone into the water, which degrades into free oxygen radicals that destroy bacteria and oxidize metals, allowing them to be post-filtered\textsuperscript{10}. Reverse osmosis purifies water by using high pressure to force it through a semi-permeable molecular-level membrane. This membrane has the potential to be a source of contamination, possibly through the colonization of bacteria and the capability for biofilm formation. It is important to recognize the vast potential that different water processing and purification approaches and techniques offer for microbial contamination. For instance, a contaminated Absolute Filter could remove some bacteria while introducing others.

Keeping the potential for variation within these processing techniques in mind, I investigated the microbial approach used by four nationally distributed (Aquafina, Dasani, Arrowhead and 365 Spring) and one local (Eldorado) bottled water brands. As expected from the lack of an obligation of bottled water companies to disclose relevant consumer information, the details and complete procedures regarding water processing by these brands were largely unavailable. This presents the challenge of an adequate comparison between different brands, considering the potential differences in UV light frequency, filter size and type, and differences in storage, to name a few.
Dasani and Aquafina obtain water from public municipal water sources. Dasani filters through granular activated carbon filters, applies reverse osmosis, disinfects with UV light, adds minerals, and finally, ozonates the water. Taking a similar approach with a slightly different order, Aquafina applies its municipal waters through several filters prior to exposure to UV light, reverse osmosis, re-filtering through an activated carbon filter, and ozonation (Figure 1).

Figure 1. Illustration of the purification process of Aquafina (www.aquafina.com). This approach is highly comparable to Dasani, only varying in the number of filtration steps, and UV and reverse osmosis treatments are in opposite order in the Dasani system.

365 Spring, Arrowhead, and Eldorado obtain their water from a natural spring source. These brands are less comparable in their treatments relative to Dasani and Aquafina, as they all take different approaches to water treatment. 365 Spring processes its water using a 0.1 micron Absolute Filter. A filter of this size would, in theory, completely remove any bacteria from the water; no
additional treatment method was disclosed. Arrowhead collects and stores its spring water, applies micro-filtration, and disinfects with UV light (Figure 2). Arrowhead does not divulge its filter size, but claims that their micro-filters are designed to remove particles and microorganisms as small as 0.2 micrometers in diameter. Eldorado is a smaller company that uses spring water from a distinct location, the artesian springs located in Eldorado Canyon, Colorado. This brand claims that the water is naturally filtered through a layer of sandstone\(^2\). After collecting the spring water, it is processed through screening to remove debris, and undergoes UV light disinfection, making it the least processed of the bottled water tested.

![Figure 2. Illustration of the purification process of Arrowhead spring water.](http://www.nestle-watersna.com/pdf/AH_BWQR.pdf)
Many bottled water companies claim total purity (e.g. 365 Spring representative claims their Absolute Filter technology will remove 100% of all particles 0.1 micron or larger [personal communication], “[Arrowhead] filters are pharmaceutical grade and are designed to remove particles as small as 0.2 micron in diameter” [www.arrowhead.com]). However, scientific research has proven these claims to be predominantly false through the identification of bacterial contaminants. Yet the microbiology of these products has yet to be comprehensively explored on a molecular scale. Although HPC counts and other culture-dependent approaches provide useful information for specific culturable microbes, it is possible that 99% of the species present have been overlooked. Previous culture-independent studies have used techniques which target specific organisms but do not survey the full microbial diversity of a sample. For a more thorough analysis of the microbiology of bottled water, a different approach may prove more advantageous.

The DNA sequencing of the highly conserved 16S rRNA gene provides scientists with the ability to analyze the microbiology of complex environmental samples without the limiting factor of the capacity of cultivation. Rather than simply detecting or focusing on the properties of several species, a culture-independent 16S rRNA gene sequence-based approach allows for the analysis of whole microbial assemblages. DNA sequences can be compared to a database of characterized sequences, allowing for the phylogenetic identification of the organisms present. These advancements have significantly improved scientific insight and understanding of microbial diversity.

This study analyzed the microbiology of bottled water using 16S rRNA Sanger sequencing, a molecular phylogenetic approach, to characterize the bacteria present in bottled water. There were two major hypotheses tested in this project: 1) There is significant unrecognized microbiology in bottled water, and 2) water processing treatments influence the microbial load of bottled water. The first hypothesis was tested through DNA sequencing of 16S rDNA and analysis using phylogenetic
databases. This approach demonstrates the diversity of microbial species in each sample. The second hypothesis was tested through DAPI cell counts, a quantitative technique that provides relative comparison of the microbial load in each sample.

Materials and Methods

**Sample Collection:** Bottled water samples were purchased at local retail outlets.

**Epiflorescent Microscopy:** Water aliquots were taken from individual bottles (at different volumes depending on initial concentration estimates), and cells were fixed in paraformaldehyde and stained with 4′,6′-diamidino-2-phenylindole (DAPI) for 5 minutes. The samples were filtered onto a black 0.2 micron polycarbonate filter and each filter (25mm diameter) was mounted on a glass slide. The cells were visualized and enumerated at [400x] magnification using an epifluorescence microscope. Cell counts were converted to cells mL$^{-1}$ (Table 1). The standard deviations, which looks at the square root of the variance in a dataset, was calculated using Excel.

**Filtering for Bacterial Cells:** Cells were collected by filtration through a 0.2 micron polycarbonate membrane (Millipore Isopore) into a sterile filter holder (Nalgene). For spring water samples (365, Eldorado, Arrowhead), 1.5 L plastic bottles were filtered onto membranes. For purified samples (Aquafina, Dasani), two 1.5 L plastic bottles were filtered onto membranes (3.0 L each in total).

**DNA Extraction and Environmental PCR:** Immediately after filtering, the filter was removed and isolated under a laminar flow hood, and placed into a 2 mL tube containing a phenol-chloroform mix to dissolve the polycarbonate filter. The bacterial cells were mechanically lysed using mechanical disruption (“bead-beating”). After bead-beating for two minutes using a high
powered shaker, the DNA was isolated by phenol-chloroform extraction and isopropanol precipitation. The total genomic DNA was PCR-amplified (30 cycles) with universal rRNA primers 515F and 1391R to generate amplicons 800-nucleotides in length.

**Cloning and T3/T7 PCR:** The rRNA amplicons were gel-purified using a 1% agarose/TAE gel, run at ~100 Volts for approximately one hour, to separate the 16S rDNA amplicons from any remaining nucleic acids and other PCR byproducts. Once purified, the 16S sequences were ligated into TOPO vectors and cloned into TOPO-10 electrocompetent cells (Invitrogen Corp) by electroporation at 1.5 Volts. The cloned cells were allowed to grow for 75 minutes at 37°C, then were plated on LB-amp plates at several different dilutions and grown overnight at 37°C. Colonies from LB-amp plates with the appropriate dilution were picked with sterile pipet tips into 96-well plates and grown overnight in 37°C in 500 uL of 2xYT media with 1 µg/mL ampicillin to a concentration of ~5x10⁹ cells/mL. The overnight cultures were diluted with TE buffer and heated to 85°C to lyse the cells, then centrifuged at 2800 rcf to remove cellular debris. Supernatant was then transferred to PCR mix in 96-well plates containing T3 and T7 primers, specific for the region of TOPO vector flanking the rRNA gene insert. The reaction underwent 40 cycles of PCR to amplify the rRNA gene segment, and amplicons were visualized on a 1% agarose gel to verify the inserts.

**Sequencing:** The T3/T7 PCR product was purified using an enzyme mix containing exonuclease and shrimp alkaline phosphatase (Exo-SAPIT). Once purified, the DNA was fluorescently labeled to prepare the samples for sequencing. Ammonium acetate and isopropanol were used to precipitate out the DNA, which was then resuspended in formamide. The fluorescently labeled DNA product was then sequenced using the MegaBACE 96-well capillary per manufacturer’s instructions (GE Healthcare).
Analysis: Chromatograms generated during sequencing were imported into the open-source XplorSeq software program\textsuperscript{32}. The Basic Local Alignment Search Tool (BLAST) was used to compare the DNA sequences generated to those of cultured organisms in the Living Tree Project (LTP) database (http://www.arb-silva.de/projects/living-tree/). This program associates species level relatedness to a 97% sequence identity. In addition, sequences were aligned using the SINA alignment provided on the Silva website (http://www.arb-silva.de/aligner/) and inserted by parsimony with the ARB\textsuperscript{28} software package into a Silva Reference Database (version 104). The phylogenetic identity of each sequence was assigned based on its position in the guide tree. These two analytic methods combined helped verify identifications of each gene sequence in the samples. Cluster tables were exported from XplorSeq into Excel to compare the phylogenetic diversity among all sample libraries in the study. Biodiv in XplorSeq was used to calculate $S_{\text{Chao1}}$ diversity estimates.

Results

Epifluorescent Microscopy

Cells were DAPI-stained from 5 individual water brands, and viewed under an epifluorescence microscope at 40x magnification (see Materials and Methods). The cell counts corroborate previously described drinking water data from several past studies\textsuperscript{33} (Table 1). The data range from $4.5 \times 10^2$ to $6.5 \times 10^4$ cells mL$^{-1}$. 
Table 1. This table depicts cell count data of water samples using DAPI staining.

<table>
<thead>
<tr>
<th>Sample Name</th>
<th>Volume filtered, mL</th>
<th>Avg cells/field</th>
<th>cells/filter</th>
<th>cells/mL</th>
<th>Final Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>365 Spring</td>
<td>100</td>
<td>15.2</td>
<td>7.2E+04</td>
<td>7.2E+02</td>
<td>6.5E+02</td>
</tr>
<tr>
<td>365 Spring</td>
<td>100</td>
<td>12</td>
<td>5.7E+04</td>
<td>5.7E+02</td>
<td>5.1E+02</td>
</tr>
<tr>
<td>365 Spring</td>
<td>75</td>
<td>7.9</td>
<td>3.8E+04</td>
<td>5.0E+02</td>
<td>4.5E+02</td>
</tr>
<tr>
<td>365 Spring</td>
<td>75</td>
<td>11.5</td>
<td>5.5E+04</td>
<td>7.3E+02</td>
<td>6.6E+02</td>
</tr>
<tr>
<td>Aquafina</td>
<td>100</td>
<td>29.3</td>
<td>1.4E+05</td>
<td>1.4E+03</td>
<td>1.3E+03</td>
</tr>
<tr>
<td>Aquafina</td>
<td>100</td>
<td>48.9</td>
<td>2.3E+05</td>
<td>2.3E+03</td>
<td>2.1E+03</td>
</tr>
<tr>
<td>Arrowhead</td>
<td>3</td>
<td>44</td>
<td>2.1E+05</td>
<td>7.0E+04</td>
<td>6.3E+04</td>
</tr>
<tr>
<td>Arrowhead</td>
<td>3</td>
<td>45.7</td>
<td>2.2E+05</td>
<td>7.2E+04</td>
<td>6.5E+04</td>
</tr>
<tr>
<td>Dasani</td>
<td>100</td>
<td>16.4</td>
<td>7.8E+04</td>
<td>7.8E+02</td>
<td>7.0E+02</td>
</tr>
<tr>
<td>Dasani</td>
<td>100</td>
<td>15</td>
<td>7.1E+04</td>
<td>7.1E+02</td>
<td>6.4E+02</td>
</tr>
<tr>
<td>Eldorado</td>
<td>3</td>
<td>40</td>
<td>1.9E+05</td>
<td>6.3E+04</td>
<td>5.7E+04</td>
</tr>
<tr>
<td>Eldorado</td>
<td>5</td>
<td>24.8</td>
<td>1.2E+05</td>
<td>2.4E+04</td>
<td>2.1E+04</td>
</tr>
<tr>
<td>Eldorado</td>
<td>5</td>
<td>18</td>
<td>8.6E+04</td>
<td>1.7E+04</td>
<td>1.5E+04</td>
</tr>
<tr>
<td>Eldorado</td>
<td>15</td>
<td>112</td>
<td>5.3E+05</td>
<td>3.5E+04</td>
<td>3.2E+04</td>
</tr>
</tbody>
</table>

Figure 1 illustrates the mean average cell counts for each brand of bottled water, graphed against a log scale for visualization purposes. The positive error bars represent the standard deviation combined with the average cell count for a brand. The negative error bars represent the difference between the standard deviation and the lowest cell count for that brand. This measurement helps to test the statistical significance of the cell count data, demonstrating the repeatability of the data collected.
Figure 1. This bar graph illustrates the mean cell counts using DAPI-staining for all bottled water samples. The error bars represent the standard deviation.

BLAST Sequencing Data: Dasani and Aquafina

Once the 16S rRNA gene amplicons were cloned and sequenced, the sequences were phylogenetically analyzed using BLAST against an online NCBI database\(^4\). This database provides the most similar organism in the database based on sequence identity. Figures 2-5 illustrate the bacterial genus and species present in each sample by percent abundance of the organism within the total number of sequences compared in that sample.

Figures 2 and 3 depict the bacterial phylotypes found in Dasani and Aquafina. Dasani generated a library with sequences at a mean average of 98.3% BLAST identity, and a range of 92% - 100% sequence identity. Aquafina has an average BLAST identity of 97.8% with a range of 86% - 100%. These bottled water samples, both of tap origin, have the same three most common BLAST hits: *Sphingomonas yanoikuyae*, *Acinetobacter junii*, and *Acidovorax temperans* [Phylum:
Proteobacteria]. The level of diversity between these is also comparable in that 23 unique sequence types were found out of 71 total sequence types in the Dasani sample and 20 unique sequences were found in 73 total sequences from the Aquafina sample.

**Figure 2.** Best BLAST hits for Dasani library in LTP database.

The dominant bacterial species in these municipal bottled water libraries is the same. Sequences related to *Sphingomonas yanoikuyae* [Phylum: Proteobacteria] make up 17% of the microbiology found in Dasani and over 34% of the number of sequences found in Aquafina.
**Figure 3.** Best BLAST hits for Aquafina library in LTP database.

**BLAST Sequencing Data: Arrowhead and Eldorado**

Figures 4 and 5 illustrate the relative abundances of bacterial sequences in Arrowhead and Eldorado using BLAST. These libraries have a high sequence identity with BLAST. The Arrowhead library has an average of 98.0%, and a range of 99% - 100% identity. The Eldorado library’s mean is 97.8%, with a range between 89% and 100% sequence identity. These bottled water samples of spring water origin are comparable at a species level. Most notably, the dominant sequences in both spring bottled water sources were related to *Curvibacter gracilis* [Phylum: Proteobacteria]. The blue (bottle 1) and red (bottle 2) bars in each figure represent a library from a different bottle from the same brand of bottled water, providing a useful means of intra-comparison within the brand. The Arrowhead library had relatively low bacterial diversity, each replicate represented by only six (Arrowhead 1) and four (Arrowhead 2) different bacterial species according to NCBI’s database. The
dominant bacterial species in the Arrowhead library, *Curvibacter gracilis*, is present at 89.8% (149 out of 166 total sequences).

**Figure 4.** Best BLAST hits for two Arrowhead libraries in LTP database.

It is important to note that the Eldorado library contained approximately four times the number of sequences compared to Dasani and Aquafina, and twice the number of sequences of Arrowhead. This brand was phylogenetically analyzed using two 1.5 L bottles of water, creating a library of 340 sequences. The overall number of distinct species assigned using the NCBI database is 22 for 169 sequences in the first bottle of Eldorado, and 26 different species out of 171 total sequences in the second bottle. The more sequence data obtained from an individual sample, the greater the expected number of observed species. Eldorado (Figure 5) spring water shows a relatively high amount of biodiversity, which may be a result of deeper sequencing. Figure 5 compares duplicate libraries and depicts the consistency between the 2 different bottles of Eldorado tested. The dominant sequences in the library (*Curvibacter gracilis* 35.6%, *Hydrogenophaga atypica* 23.5%, *Acidovorax temperans* 5.6%, and *Variovorax boronicumulans* 5.0%) are all present around the same
abundances in replicate libraries. The level of species diversity in this library is also similar; the first bottle has a library of 169 sequences with 22 unique species identities, the second contains 171 sequences and 27 unique species.

**Figure 5.** Best BLAST hits for two Eldorado libraries in LTP database.

**Using $S_{\text{Chao1}}$ to Estimate Species Coverage**

$S_{\text{Chao1}}$ is a statistical tool used to evaluate the effectiveness of collecting further samples in species taxonomic surveys\textsuperscript{35}. It estimates the new species expected to be observed in a second
sequencing survey based on the data from initial sequences. The generation of \( S_{\text{Chao1}} \) provides a way to perceive relative species diversities between samples. Calculating \( S_{\text{Chao1}} \) involves dividing the lower confidence interval (CI) by the upper CI. CI is an interval estimate of a population parameter used to determine the reliability of an estimate. \( S_{\text{Chao1}} \) is expected to correlate positively with the number of sequences - the more number of sequences per sample, the greater the \( S_{\text{Chao1}} \) value.

As denoted by the number of BLAST hits per library (Figures 2-5), diversity varied between the libraries. The predicted diversity of the Arrowhead library by \( S_{\text{Chao1}} \) is significantly different from that of the other three bottled water libraries. This means that the microbial diversity of Arrowhead (11.05) is significantly less than Aquafina (47.82), Dasani (34.42) and Eldorado (37.75). Aquafina is demonstrated to have the greatest amount of predicted species diversity.

<table>
<thead>
<tr>
<th>Environment</th>
<th>#Seq/Library</th>
<th>Chao (lower/upper Confidence Intervals)</th>
<th>Number of Sample Libraries Included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aquafina</td>
<td>70</td>
<td>47.82 (38/59)</td>
<td>-</td>
</tr>
<tr>
<td>Arrowhead</td>
<td>154</td>
<td>11.05 (8/16)</td>
<td>2</td>
</tr>
<tr>
<td>Dasani</td>
<td>70</td>
<td>34.42 (32/40)</td>
<td>-</td>
</tr>
<tr>
<td>Eldorado</td>
<td>336</td>
<td>37.75 (23/63)</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2. \( S_{\text{Chao1}} \): Absolute measure for the amount of diversity seen in the samples. This Table also includes the number of sequences per library and the number of sample libraries per bottle water brand (Environment).

**Phylogenetic Abundance**

The microbial composition of all samples is shown in Figure 6. Dasani and Aquafina (municipal source), and Eldorado and Arrowhead (spring source) have comparable bacterial clades. *Sphingobium, Planomicrobium,* and *Acinetobacter* are genuses found in relatively high abundance in
Dasani and Aquafina samples. Similarly, sequences from Comamodaceae are common in Eldorado and Arrowhead libraries. One Comamonadaceae genus in particular, *Hydrophaga*, is abundant in Eldorado libraries only. There is a large occurrence of Alpha- and Betaproteobacteria found in all bottled water libraries.

**Figure 6.** Phylogenetic abundance of all bottled water samples using ARB Lineage information.
Relative Group Abundances

In ARB, the four libraries contained minimal diversity at the group level (Figure 7). Aquafina (Figure 7A) and Dasani (Figure 7B) are dominated by Firmicutes, Alpha-, Beta- and Gammaproteobacteria. These two libraries are both from heavily treated water of municipal origin. The number of groups represented are comparable between these samples as well. The Aquafina library contains 8 unique groups, and the Dasani library contains 7 distinct bacterial groups. The diversity at the group level of these libraries is relatively high compared to the less treated spring water, Arrowhead (Figure 7C) and Eldorado (Figure 7D), with 3 and 4 bacterial groups, respectively. The vast majority of the sequences in these bottled spring water samples are most closely identified as Betaproteobacteria using ARB. Although Betaproteobacteria is one of the four most abundant groups present in Dasani and Aquafina, there is a relatively higher abundance of Alphaproteobacteria per individual library.

Another notable comparison between spring and municipal sourced water is that although Aquafina and Dasani have approximately twice as many groups represented in their libraries, Arrowhead and Eldorado have sequences most closely related to Holophagae [Phylum: Acidobacteria] which is not found in the municipal-sourced water libraries. These libraries appear to have unique and comparable microbial signatures at the group level.
Figure 7. Pie charts comparing relative group abundances of bacteria using ARB database.

Discussion

Dominant Bacterial Groups  (Note: * = opportunistic pathogens; ** = primary pathogens)

The predominant species found in the phylogenetic libraries from bottled water samples are listed and described below. Some libraries have a high percentage of sequences belonging to a single NCBI species identity (most notably the Arrowhead samples.) The lineage information provided by
the best BLAST hit corresponds with the phylogenetic identification from ARB, unless otherwise stated. The most abundant species found in these data have been identified in more than one sample.

**Curvibacter gracilis** (Arrowhead 89.8%, Eldorado 35.6%, Dasani 4.2%, Aquafina 2.7%)

There were inconsistencies between the ARB and NCBI databases for this species. ARB describes five distinct clades for the sequences NCBI describes for one species: *Comomadaceae*, *Holophagaceae* marine group, *Comamonadaceae Limnohabitans*, *Comamonadaceae* uncultured, and *Comamonadaceae* *Pseudorhodoferax*. However, only one lineage is shared between all four brands (*Comomadaceae*) and the last three phylogenetic groups are found in the spring water samples only. These organisms have been found in environmental samples such as soil and freshwater.

*Sphingomonas yanoikuyae* (Aquafina 34.2%, Dasani 16.9%, Eldorado 2.4%)

Members of the *Sphingomonas* genus are known to be decomposers of aromatic compounds, often used to remove pollutants from the environment as bioremediators. This Biological Safety Level 2 (BSL2), indicates potential concern for opportunistic pathogenicity. *S. yanoikuyae* has been isolated from Finnish and Swedish public water municipal systems and from the Elbe River in central Europe.

**Hydrogenophaga atypica** (Eldorado 23.5%)

This species has been isolated from the activated sludge of wastewater in Munich, Germany. Closely related members of *Hydrogenophaga* have been proposed to be used to reduce the eutrophication process in lakes through the removal of phosphate.

**Acidovorax temperans** (Aquafina 15.1%, Dasani 9.9%, Eldorado 5.6%)
Although BLAST aligned sequences to the same species, ARB designated the sequences into five distinct clades: Comamonadaceae Acidovorax 5, 6 and 7, Comamonadaceae Simplicispira, Oxalobacteriae Janthinobacterium, Oxalobacteraceae Naxibacter, and Comamonadaceae Brachymona. The most important observation regarding this distinction is that only two out of these seven are shared between both municipal and spring bottled waters: Simplicispira and Acidovorax 6. Simplicispira has been described as a sewage-derived microorganism, isolated from wastewater treatment plants\textsuperscript{41}. Acidovorax and Janthinobacterium have been found in ground water used for a municipal system\textsuperscript{42}.

**Acinetobacter junii** (Dasani 12.7%, Aquafina 11.0%)

Best BLAST identity ranged from 98 – 100% with a mean bit score of 1136. Members of the genus Acinetobacter have been established as a cause of nosocomial infections. The species, A. junii, is documented as a rare cause of disease, associated with cases of septicemia in neonates, pediatric oncology patients, and in an adult oncologic patient\textsuperscript{43}. This microbe suggests a potential pathogen in the water, but more studies are needed to determine its viability.

**Planomicrobium okeanokoites** (Dasani 7.0%)

There are two specific clades in the genus using ARB, Planomicrobium and Planococcus. Members of the genus Planomicrobium have been isolated from fermented seafood, marine mud, Antarctic samples, intertidal sediments and glacier\textsuperscript{44}. Planococcus has been isolated from sea water\textsuperscript{45}.

**Flavisolibacter ginsengiterra** (Dasani 5.6%, Aquafina 5.5%)

ARB recognizes this sequence as Chitinophagaceae Sediminibacterium. Both clades have been isolated from soil\textsuperscript{46}. 


**Cloacibacterium normanense** (Dasani 5.6%, Aquafina 1.4%)

This species belongs to the Flavobacterium group, which has been shown to constitute a significant portion of activated sludge from wastewater plants. *C. normanense* has specifically been isolated from untreated wastewater from a water-treatment plant located in Norman, Oklahoma\(^47\). This may be an indicator of municipal origin.

**Paenibacillus alginolyticus** (Aquafina 5.5%)

*P. alginolyticus* is a xanthan-degrading species that have been found in soil\(^48\). Bacteria belonging to this genus have been detected in a variety of environments including soil, water, and humans\(^36\).

**Staphylococcus saccharolyticus** (Aquafina 5.5%, Dasani 4.2%)

*Staphylococci* are widespread in nature; however, they have primarily been isolated from the skin and other cutaneous tissue of mammals and birds. *S. saccharolyticus* is a species that has been found on humans and other primates\(^36\). The presence of this organism may indicate human contamination.

**Variovorax boronicumulans** (Eldorado 5.0%, Arrowhead 2.3%)

This species has been isolated from environmental samples, such as soil and water\(^49\). One study isolated *V. boronicumulans* from an aquifer in Cape Cod, Massachusetts, contaminated with the antimicrobial sulfamethoxazole\(^50\). This study describes the selection of microbes under these particular antimicrobial conditions to favor *V. boronicumulans*. This particular antimicrobial is commonly associated with wastewater or livestock contaminations, perhaps indicating the exposure to a contaminant that would select for this species. Further research is required to verify this speculation.
**Pseudomonas otitidis** (4.2% Aquafina, 2.7% Dasani)

Best BLAST identity ranged from 97 – 99% with a mean bit score of 809. Members of this species have been identified in humans from clinical specimens of infected ears\(^5^6\). *Pseudomonas* have also been shown as a dominant genus in water contaminated with plasticizers\(^5^7\). Although the presence of this species is important to note, the low representation and mediocre bit score of this sequence in Aquafina and Dasani libraries necessitates further investigation.

ARB identified a sequence in the Aquafina library as *Clostridium_1*. Although this singleton has a low occurrence in the Aquafina library, it is a genus of the primary pathogens *Clostridium difficile*, *perfringens*, *tetani* and *botulinum*. Generating another library would be useful to test for any re-occurrence.

It is important to note that bacteria from the genus *Mycobacterium* were not detected in the samples. This opportunistic pathogen is known for resistance to chlorine disinfection, and its absence in this dataset may be explained by the lack of chlorination treatment in all four bottled water brands.

**Cell Counts**

Previous studies have found on average 5.3 x 10\(^2\) in tap water and approximately 10\(^3\) - 10\(^4\) bacteria mL\(^{-1}\) in bottled water sample\(^1^2,1^4,1^6,2^3,2^4\). These numbers correlate with the samples in this study, which established cell counts ranging from to 4.5 x 10\(^2\) to 6.5 x 10\(^4\).

The cell counts from this study support the proposed hypothesis: water treatment affects microbial load. The cell counts of the majority of the spring water samples are similar to a previous study that tested for the quantity of bacteria in untreated bottled spring water. The cell counts for
Eldorado and Arrowhead spring water brands are one or two orders of magnitude greater than Aquafina and Dasani, respectively. These numbers seem generally proportionate to the amount of treatment the water received, specifically if the water received reverse osmosis and ozonation processing. However, more research needs to be done to reasonably elucidate this correlation.

365 Spring provides an interesting exception to the cell count data. This sample generated the smallest cell count, and the extraction of DNA from this sample yielded poor results, rendering sequence data unobtainable at this point. As described in the background, this brand uses the smallest filter of all five bottled water brands tested with a pore size of 0.1 micron. Both Arrowhead and 365 Spring use filters that, in theory, would remove the majority of the microorganisms from water. However, one study demonstrates that there may be a dramatic underestimation of the diversity of 0.2 µm-filterable bacteria\textsuperscript{35}. In correlation of this past study, the cell count data also dampen the absolute reliability of Absolute and micro-filters in the microbial decontamination of bottled water. Although 365 Spring demonstrated an extremely low cell count, it is only slightly lower than tap sources.

The lack of chemical processing in the Eldorado sample would suggest potentially higher cell count data compared to the other four brands. However, Eldorado cell counts are less than Arrowhead at \(3.1 \times 10^4\) and \(6.4 \times 10^4\) cells per mL, respectively. The higher cell count for Eldorado compared to Dasani and Aquafina is predictable due to this brand’s lack of filtration, but Arrowhead uses microfiltration prior to UV disinfection. Based on their micro-filtration process, it is unexpected that Arrowhead water was demonstrated to have over twice as many bacterial cells as Eldorado water. This may indicate greater cell growth after bottling. Many of the organic compounds present in raw water are potential nutrients for bacteria, so it is possible that distinct nutrient compositions contributed to this qualitative observation\textsuperscript{51}.

**Qualitative Observations**
There are many factors that could potentially influence the microbial load and membership found in bottled water. Varying nutrient compositions and different filtering techniques are discussed above. However, the downstream effects of nutrient composition and filtering stringency could also greatly influence the microbiology of these products.

It is widely known that turbidity in the water reduces the efficacy of ultraviolet light to pass through, and the microbes will attach to the turbidity, rendering them much harder to inactivated when bound\textsuperscript{52}. Water treatment can reduce turbidity, making UV disinfection ideally performed as a final purification step. Dasani is the only brand tested that uses both reverse osmosis and filtration prior to UV light disinfection. Aquafina filters its water, and then performs UV disinfection prior to reverse osmosis (Figure 1). Similar to Aquafina, Arrowhead filters, then performs UV disinfection treatment, but does not perform reverse osmosis on its water. Eldorado does not filter before UV disinfection, providing the least amount of treatment.

It is possible that the use of multiple filters, an approach employed by Dasani and Aquafina, decreases the turbidity enough to have a greater UV light exposure, killing the majority of the bacteria. This could explain the lower cell counts (Table 1) and greater microbial diversity, demonstrated by significant differences in $S_{\text{Chao1}}$ estimates compared to Arrowhead (Table 2). Because Eldorado does not filter its water, it is difficult to compare. Diverse approaches to water purification are likely to influence the microbial load and membership of the final product; although, more tests must be performed to thoroughly determine these relationships.

Similar to previous studies, the presence of Betaproteobacteria is considerably different between municipal and spring bottled water (Figure 7)\textsuperscript{33}. The notable smaller quantity of Betaproteobacteria sequence abundance in municipal waters compared to spring water indicates the sensitivity of this group to drinking water disinfection treatment.
Based on the phylogenetic comparisons, the qualitative microbiology of these samples is impressive. There is a noticeable pattern in bacterial groups between libraries of the same water source and similar processing techniques (Figure 7). This pattern could potentially be interpreted as a fingerprint of municipal vs spring water sources, or using reverse osmosis vs UV/UV and filtration. Further studies need to be performed to elucidate this potential.

These data document a high degree of reproducibility between replicate libraries. Both brands that were analyzed twice using distinct bottles to test for consistency between individual bottle units, Arrowhead and Eldorado, displayed clear similarities in both in assigned species and species abundance. The most abundant species for Arrowhead, *Curvibacter gracilis*, was found at an 89.1% and 90.5% abundance. As the dominant species for Eldorado libraries, *C. gracilis* abundances were 38.5% and 32.7% per library. These brands display fairly different abundances from one another, but are highly similar to themselves. These data also support the potential for source fingerprinting.

**Possible Contamination**

The bottled water libraries contained sequences most closely related to sequences from cultured pathogens deposited in NCBI and Silva databases. There were particularly notable strains shared in Aquafina and Dasani libraries: *Actinobacter junii*, a pathogen associated with nosocomial infections; *Staphylococcus saccharolyticus*, a possible indication of human contamination; and *Pseudomonas otitidis*, a pathogen associated with human ear infections. *Sphingomonas yanoikuyae* is a species that has been assigned to BSL2 as a potentially infectious bacterium. *Clostridium* was identified in one library (Aquafina) at the genus level, necessitating the collection of further data to associate this occurrence with a pathogenic species.
Currently, the majority of dose-response relationships between the pathogens found in water and healthy individuals are not fully understood, and the risk assessment of new pathogens is completely unknown\textsuperscript{52}. Many of the bacteria found in drinking water samples described in past studies are shown to be human secondary opportunistic pathogens, with immuno-compromised groups at a greater risk for adverse effects\textsuperscript{51}. This imprecise relationship between pathogenic and nonpathogenic bacteria needs to be further defined to determine what is of concern and what is harmless in products sold for human consumption. It is critical to both characterize the microbiology and implement stringent standards of the disinfection of bottled water.

Many of the bacterial clade classified in these libraries identify with organisms naturally found in environmental samples, such as water and soil, or municipal systems. However, there are dominant sequences in these samples that could be a major sign of wastewater contamination. These libraries contain sequences associated with species that have been found in activated sludge and raw sewage, such as *Hydrogenophaga atypical*, *Comamonadaceae Simplicispira*, and *Cloacibacterium normanense*.

**Conclusions**

Considering that a disconcerting 32\% of bottled water companies conceal treatment methods and purity testing and 13\% publish water quality reports that lack any tangible testing results, the investigation of what is actually in these products is crucial\textsuperscript{53}. The results described in this study begin to elucidate the microbiology of these products, both qualitatively and quantitatively. This study demonstrates that different brands of bottled water are highly inter- and intra-comparable to one another with regards to microbial composition and diversity. The inter-comparability of the dataset
lies in the distinct water source, municipal or spring, and the downstream processes applied. The intra-comparability is established in the notable similarity of libraries generated from different bottles of the same brand. The microbial species present in these samples are similar from bottle to bottle; these data suggest a unique and repeatable bacterial quality among distinct brands, different water sources, and the numerous water processing techniques.

Dasani and Aquafina both use municipal tap water, and are unique from the other brands tested in that they process their water using reverse osmosis and ozonation. The most representative bacterial sequences present in these brands are from *Sphingomonas yanoikuae*, *Acinetobacter junii*, and *Acidovorax temperans* (BLAST). In the spring waters, *Curvibacter gracilis* is the most abundant at 89.8% and 50.3% for Arrowhead and Eldorado, respectively.

These water types also have comparable relative group abundances; municipal libraries are represented predominantly by Alpha-, Beta- and Gammaproteobacteria, and Firmicutes. The vast majority of sequences in the Arrowhead and Eldorado libraries identify with Betaproteobacteria at 91.9% and 80.1% abundance, respectively. Alphaproteobacteria are also identified in Eldorado’s libraries at 18.5% abundance, and Holophagae in Arrowhead’s libraries at 7.5%.

Intra-comparability within the samples is also significant. The replicate datasets for both Eldorado and Arrowhead bottled water are similar in that libraries from both brands contain the same dominant species and in fairly high abundance similarities (Figures 4 and 5). The levels of diversity are also repeatable within these replicate libraries. There is a very high diversity of unique sequences per library in Eldorado and a particularly low diversity in Arrowhead. *S*<sub>chao1</sub> diversity estimates support this relative species diversity comparison (Table 2). Arrowhead had significantly less amount of predicted species diversity compared against all bottled water libraries.
Based on cell count data alone, 365 Spring bottled water demonstrates better bacterial water quality than the other four brands, with a bacterial load of $5.7 \times 10^2$. This suggests that the best purification approach is a 0.1 micron Absolute Filter, as it competes for bacterial removal with Dasani ($6.7 \times 10^2$) and Aquafina ($1.7 \times 10^3$), which have numerous processing steps. However, bacterial cell counts are significantly smaller relative to Arrowhead ($6.4 \times 10^4$) and Eldorado ($3.1 \times 10^4$). It is unclear whether these numbers correlate to water processing or original water sources.

Although cell counts are valuable indicators for bacterial load, it is crucial to investigate the state of these cells. For example, if Eldorado had the least amount of post-collection contamination, and low enough turbidity such that the UV light was highly effective in disinfection, the cellular load may be irrelevant because the microbial inhabitants are no longer living. Similarly, if the majority of the living organisms in Dasani samples are of the pathogenic species *Actinobacter junii*, this poses a major health concern.

In conclusion, the microbiology revealed in these samples through the cloning and sequencing of the 16S rRNA gene and the quantification of relative microbial loads present novel insight into these supposedly pure water samples. It is likely that the differences in the microbiology between different brands of bottled water are primarily due to the selective pressures posed by the filtration and purification processes and by the original water source. The known pathogens, human-skin associated microbes, and bacteria commonly isolated from raw sewage associated with these samples (the majority of these specifically found in Dasani and Aquafina libraries) render the current quality of this product of great concern. While this study contributes to further discrediting of the misconception of bottled water’s pristine and sterile quality, there is still additional research to be performed.
Future Work

The most meaningful supplemental experiment to be performed is a Live/dead test. It is crucial to prove the viability of the bacteria in each sample prior to presuming potential health implications. Because the Live/dead test was designed for monocultures and high-density cell samples, the protocol must be manipulated such that the cells can be concentrated and stained without affecting their viability. Possible approaches include 1) pelleting the cells via centrifuge, removing the supernatant/medium, resuspending in saline solution (repeat 2-3 times), and 2) staining the cells directly on a glass slide.

Additionally, the collection of more samples would be extremely useful to further compare the microbiology between and within varying brands. This would also provide a more comprehensive analysis of the microbial effects of different processing techniques.

In light of the growing concern of antibiotic resistant bacteria and the contamination of antibiotics in sewage systems, analyzing the presence of antibiotics and/or antibiotic metabolites in drinking water would be crucial, and the resulting microbial assemblage which these antibiotics select for. Comparing bottled water from a tap source to unprocessed tap water, and these samples to bottled water from a natural spring source would be particularly useful in serving our understanding of the effects of this pharmaceutical contamination and potential exposure to antibiotic resistant bacteria. One study demonstrated the presence of these antibiotic resistant bacteria in bottled water, finding 45% of the strains in 30 brands of bottled water to be resistant to two or more antibiotics. A similar study found 70% of the total isolates from dominant bacterial species were resistant against two or more antibiotics.
To both supplement the detection of antibiotic resistant microbes and general understanding of what is in bottled water, a broad chemical analysis would be invaluable. An ideal chemical analysis would test for antibiotics, plasticizers, oxygen, fluorine and chlorine levels, and other chemicals. This data could be explored such that the chemicals present in bottled water could indicate specific chemical signatures of different municipal and freshwater sources. These chemical signatures are likely to also contribute to microbial selection, along with specific processing techniques and water origin, rendering further potential for a distinct microbial signature associated with individual brands. There is a prospective for the unique microbial compositions of bottled water to provide a way of identifying the initial water source⁵⁴.

**Acknowledgments**

I’d like to thank Norm Pace for the incredible opportunity to perform this research in such a great environment. I’d also like to thank Kim Ross, Laura Baumgartner, and Alex Greenspan for their continual help with experimental procedures, data analysis, and developing this paper. Additionally, I’d like to thank Steven Mihaylo and Elizabeth Costello for training me in the methods used in this study. I would also like to thank Rob Knight for my time in his lab, providing me with the opportunity to pursue original research ideas and interact with brilliant computer scientists, affording me invaluable exposure to the computational analysis approach and development behind microbiology. Finally, I want to reverently acknowledge Claire Fraser-Liggett for inspiring me to pursue my love of microorganisms, and providing me with the encouragement and opportunity to begin my lab experiences at The Institute for Genomic Research in 2006, continuing at the J. Craig Venter Institute in 2007, and The Institute for Genome Sciences at the University of Maryland School of Medicine in 2008. Dr. Fraser-Liggett has been the dominant catalyst in my pursuit of microbial
knowledge, from acquainting me with this world at the age of 15, to introducing me to Dr. Knight and Dr. Pace as an undergraduate. Without any of these people, this research would not be possible!

References


