Ceviche is a traditional dish made from raw fish meat marinated in lime juice without any heat cooking step throughout its preparation process. Although the use of organic acids as antibacterial agents is well known, recent research indicates that lime juice actually can reduce the risk of V. parahaemolyticus infections but it is ineffective against other potential pathogens. Despite the fact that fresh fish meat is safe; exposed organs including skin, gills and guts represent a potential source of bacterial contamination. In Mexico, diarrheal diseases are caused mainly by contaminated food; it is estimated that almost 67% of infections are due the presence of bacterial agents mainly in frozen and fresh fish. The main objective of this study was to estimate the taxonomic diversity of microbial species present in ready-to-eat ceviche using a metagenomic approach. Six samples of commercially available ceviche were subjected to DNA high throughput sequencing and bioinformatics analyses, we identified between 65,000 and 131,000 reads per sample. The predominant phyla identified through the samples were Proteobacteria, Bacteroidetes and Firmicutes. We discuss the factors involved in the microbiological quality of this kind of raw foods and how they influence the bacterial diversity within the analyzed samples.

**Keywords:** Ceviche; bacteriome; food safety

**MATERIAL AND METHODS**

**Sample collection**
Ceviche samples, identified as S12, S13, S14 and S15 were purchased from restaurants and street vendors in Reynosa, Mexico, with every ingredient used in its preparation included (table S1). Each sample was transported on ice to the laboratory and preparation included (table S1). Each sample was transported on ice to the laboratory and sequencing of nucleic acids purified directly from the environmental sample, could improve significantly the limitations exhibited by the classic microbiological methods (De Filippis et al. 2018). In the present study, we analyzed six ceviche samples through massive sequencing of the 16S V3-V4 rRNA amplicon to estimate the bacterial diversity present in this kind of raw seafood. We obtained between 65,000 and 131,000 reads within the sequenced samples, the reads were quality filtered and assembled in contigs. The resulting contigs were compared and identified by homology using the small ribosomal subunit (SSU) SILVA database v. 132. The predominant phyla found within the samples were Bacteroidetes, Proteobacteria, Firmicutes and Actinobacteria while at genus level, Psychrobacter, Shewanella, Prevotella, Pseudomonas and Pantoea were found the most abundant.

**DNA purification and sequencing**
We obtained random and representative samples containing every ingredient within each collected ceviche. Samples were mechanically homogenized in 0.9% sterile saline solution before DNA purification, the total genomic DNA was recovered from 0.25 grams of sample using a modified chloroform-ethanol procedure using 1 mL of lysis buffer (0.5M EDTA, 10% SDS) and precipitated in two steps with isopropanol and ethanol in triplicate. Finally, it was diluted in 50
µL of AE buffer (Qiagen, Valencia, CA). Prior the library preparation, DNA quality control was performed with Bioanalyzer (Agilent, Santa Clara, CA) to determine the size of DNA fragments and its abundance throughout the samples. DNA pair-ended library was performed according to Illumina’s Multiplexing Sample Preparation Guide (Illumina, San Diego, CA). The DNA was fragmented, end repaired, A’ tagged, ligated to 357-F and CD-R adaptors, size-selected, and enriched with 25 cycles of PCR during which an index was incorporated to the samples (Table S2). The higher quality resulting library for each sample measured in DNA integrity and concentration was subjected to Illumina MiSeq (2x300) pair-ended sequencing reads were assembled in contigs and quality filtered following the quality control was performed with Bioanalyzer (Agilent, Santa Clara, CA) to determine the size of DNA fragments and its abundance throughout the samples.

Sequencing data analyses

Taxonomic labels were assigned to the raw reads in each sample by homology with a local 16S rRNA database using Kraken 2 (Wood and Salzberg 2014) in the Illumina BaseSpace Sequence Hub (https://basespace.illumina.com). Additionally, sequencing reads were assembled in contigs and quality filtered following the Mothur pipeline (Schloss et al. 2009) to assess species richness and alpha diversity. Contigs were compared and identified by homology with the ribosomal small subunit (SSU) SILVA Database v. 132 (Quast et al. 2013), alignment of good quality contigs was performed using a ksize of eight kmer and the Needleman Wunsch pairwise alignment method. Aligned sequences were screened for chimeric and non-bacterial sequences that ultimately were removed from the dataset. The alpha diversity for each sample was reported as observed calculated using the Chao1, ACE and Simpson’s diversity metrics and visualized throughout the Microbiome Analyst Portal (https://www.microbiomeanalyst.ca) (Dhariwal et al. 2017). Finally, metagenomeSeq (Paulson et al. 2013) was used to find differential abundance of microorganisms at different taxonomic levels among commercial and control samples.

RESULTS

Sequenced reads were deposited and are available in the NCBI sequence depository under the Raw Fish Ceviche Bacterial Diversity BioProject (ID: PRJNA512226). We identified between 65,000 and 131,000 reads among the sequenced samples. According to the taxonomic labels given by both pipelines, Kraken and Mothur, the predominant phyla across the ceviche samples were Proteobacteria, Actinobacteria, Firmicutes and Bacteroidetes. Additionally, we found several genera with low count of reads, Chloroflexi, Cyanobacteria, Fusobacteria and Gemmatimonadetes. At genus level, Kraken analysis showed high predominance of Bacteroidetes followed by Firmicutes in S12 but S13 had high presence of Firmicutes, whereas S14, S15, S16 and S17 showed predominance of bacteria belonging to the phylum Proteobacteria (Table 1).

At genus level, in S12 we have found a high prevalence of Bacteroidetes, mainly represented by Prevotella species (58%) such as P. oris, P. dentalis, P. bucca, P. denticola; the second largest group of bacteria in S12 belongs to the phylum Firmicutes and had homologous sequences to species such as Faecalibacterium prausnitzii, Eubacterium rectale, Ruminococcus bicirculans, Selenomonas ruminantium and S. sputigena. Megaplastea eldenii and in less frequency, we found S. aureus and Streptococcus sp. In S13 we observed a high relative abundance (RA) of not assigned bacterial genus (72.2%) but also we found Psychrobacter (9.34%), Rothia (5.08%) and Paracoccus (3.7%) additionally represented in less than 2% of RA, Enterobacteriaceae, Pantoaea, Propionibacterium and Shewanella. Sample S15 had 0.95% of RA of Empedobacter and Rheinheimera. S14 had a high presence of Pantoaea (50%) and genera such as Rosenbergiella (16.95%), unclassified Enterobacteriaceae (11.96%) and Weisella (5.46%). In less percentage, we found Myroides, Pseudomonas, Chryseobacterium, Rhizobium, Buttiauxella, Sphingobacterium, Sponnouza, Rothia and Pseudochrobacterium. Non-commercial samples, S16 and S17 were only fish meat marinated with lime juice. Whereas S16 exhibited Psychrobacter (16.45%), Pseudomonas (3.5%), Propionibacterium (1.9%), Shewanella, Roseflexus, Flavobacterium, Acinetobacter, Rothia, Bradirhizobium and Aeromonas on the other hand, S17 showed Aeromonas (19.13%), Shewanella (8.77%), Propionibacterium (8.09%), Psychrobacter (3.8%) and less represented, Corynebacterium, Acinetobacter, Roseflexus, Pseudomonas, Flavobacterium and Buttiauxella.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Total Reads</th>
<th>S12</th>
<th>S13</th>
<th>S14</th>
<th>S15</th>
<th>S16</th>
<th>S17</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phylum</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>1.1</td>
<td>702</td>
<td>0.9</td>
<td>194</td>
<td>2.5</td>
<td>90</td>
<td>0.2</td>
</tr>
<tr>
<td>Bacteroidetes</td>
<td>57.2</td>
<td>33725</td>
<td>2.4</td>
<td>489</td>
<td>12.4</td>
<td>440</td>
<td>2.7</td>
</tr>
<tr>
<td>Chloroflexi</td>
<td>0.02</td>
<td>12</td>
<td>0.1</td>
<td>27</td>
<td>0.8</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>Cyanobacteria</td>
<td>0.02</td>
<td>14</td>
<td>0.1</td>
<td>33</td>
<td>1.3</td>
<td>49</td>
<td>0.0</td>
</tr>
<tr>
<td>Fimbicutes</td>
<td>36.8</td>
<td>21718</td>
<td>86.7</td>
<td>17604</td>
<td>32.5</td>
<td>1151</td>
<td>32.7</td>
</tr>
<tr>
<td>Fusobacteria</td>
<td>0</td>
<td>1</td>
<td>0.02</td>
<td>6</td>
<td>0.3</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>Gemmatimonadetes</td>
<td>0</td>
<td>1</td>
<td>0.02</td>
<td>5</td>
<td>0.0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Proteobacteria</td>
<td>4.2</td>
<td>2510</td>
<td>8.5</td>
<td>1738</td>
<td>46.5</td>
<td>1646</td>
<td>63.9</td>
</tr>
</tbody>
</table>

Figure 1 Relative abundance of bacterial genera within the ceviche samples.

Table 1 Total reads identified to phylum for each sample tested in Kraken (numbers in bold represent the most representative phylum within the sample).
Accordingly, to the Kraken analysis, Mothur analyses revealed a high abundance of Bacteroidetes in the sample S12 but we found a relatively low abundance of Firmicutes, and the most abundant phyla in S13, S14, S15, S16 and S17 were Proteobacteria, Actinobacteria, Firmicutes, Chloroflexi and Armatimonadetes (Table S3). At genus level, we found mostly Prevotella in S12 (0.995) while in S13-S15 a higher diversity was observed (Figure 1). The rarefaction richness plot observed in samples was not saturated in the ceviche samples but it clearly indicated that sample S12 had the lowest richness while S17 showed the highest number of OTU’s. (Figure 2 and Table S4). Alpha diversity was assessed through observed and estimated analyses using the Mothur data. We did not found significant differences in observed diversity among samples ($P=0.29724$), however, according to the rarefaction analysis, S12 and S17 had the lowest and highest diversity with 60 and 197 OTU’s respectively (Table 2). In addition to the observed count of species we calculated the alpha diversity trough ACE ($P=0.71295$), Chao1 ($P=0.98386$) and Simpson’s diversity indexes ($P=0.2646$) (Table 2).

Correlation and clustering visualization analysis showed that based on the microbial profiles, laboratory-made samples S16 and S17 were clustered together whereas commercial samples formed the other cluster. It is observed that S16 and S17 samples were strongly positively correlated to the presence of bacteria belonging to the Psychrobacter, Pseudomonas, Bradyrhizobium, Shewanella, Propionibacterium, Corynebacterium, Aeromonas, Flavobacterium, Acinetobacter and Vagococcus genera which do not suggest evidence of human contamination (Figure 3). On the other hand, S12, S13 and S14 had a strong correlation with bacterial classes typically associated to human contamination such as Enterobacterales, Lactobacillales, Bacteroidales and Rhodobacterales. We found that Enterobacteriaceae ($P=2.63\times10^{-4}$), Prevotellaceae ($P=0.005$) and Leuconostocaceae ($P=0.005$) family members had a significant higher abundance in ceviche samples from restaurants than in the laboratory made samples whereas Moraxellaceae was more slightly abundant in the control samples ($P=0.02$) (Table S5).

Figure 2 Richness rarefaction plot of analyzed ceviche samples

Table 2 Alpha diversity observed and calculated (Chao1, Ace, Simpson) for analyzed samples

<table>
<thead>
<tr>
<th>Sample type</th>
<th>Sample ID</th>
<th>Chao1</th>
<th>ACE</th>
<th>Simpson</th>
<th>Observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Commercial</td>
<td>S12</td>
<td>8.33</td>
<td>0.91</td>
<td>11.27</td>
<td>1.79</td>
</tr>
<tr>
<td>Commercial</td>
<td>S13</td>
<td>24.75</td>
<td>4.19</td>
<td>27.82</td>
<td>2.71</td>
</tr>
<tr>
<td>Commercial</td>
<td>S14</td>
<td>17.2</td>
<td>0.61</td>
<td>17.94</td>
<td>1.85</td>
</tr>
<tr>
<td>Commercial</td>
<td>S15</td>
<td>18.5</td>
<td>2.57</td>
<td>19.6</td>
<td>1.9</td>
</tr>
<tr>
<td>Control</td>
<td>S16</td>
<td>23</td>
<td>0.48</td>
<td>23.26</td>
<td>2.31</td>
</tr>
<tr>
<td>Control</td>
<td>S17</td>
<td>22</td>
<td>0.48</td>
<td>22.38</td>
<td>1.92</td>
</tr>
</tbody>
</table>
DISCUSSION

Bacterial agents found in the tested ceviche samples showed high diversity of taxa across the samples, as stated previously. In general, bacterial in seafood belong to two different groups, native marine or freshwater bacteria present in naturally in the aquatic environment and those bacteria present due exogenous contamination of human or animal sources.

Among the native bacteria identified in this study we found *Soonwooa*, a member of the *Flavobacteriaceae* family is a previously isolated microorganism from the yellow sea waters from Korea (Joung et al. 2010), it was found mainly in S14 and S15. Nowadays, two species of *Soonwooa*, S. sp. and S. buanensis have been described in tissues from tilapia (*O. niloticus*) reared in dams in Mexico (Soto-Rodriguez et al. 2013) and in duck meat products from Korea (Kim et al. 2016) but were not identified as pathogenic agents. Within the *Flavobacteriaceae* family we also identified reads hitting to the gram-negative *Chryseobacterium* in S13 and S15, although we were not able to identify the bacteria at species level, its close relatives, previous reports mention that *C. meningosepticum* has been associated to foodborne septicaemia due consumption of raw seafood (Kim et al. 2016) and *C. indologenes* which is widely spread in soil and plants rarely was associated to human disease (Imataki and Uemura 2017). This last example could be present in ceviche since the samples involved raw vegetables in its preparation, however, further analyses are required to elucidate the origin of *Chryseobacterium* species present in raw seafood.

The genus *Rheinheimera*, found in S14 is widely distributed in freshwater, saltwater and soil environments. Is a Gram-negative bacterium that presents antimicrobial activity in some strains (Chen et al. 2010). Additionally, other bacteria naturally associated to aquatic environments such as *Shewanella*, *Psychrobacter*, *Vagococcus*, *Buttiauxella* and *Novosphingobium* were found, although they do not represent a potential source of foodborne illness.

Vegetables in ceviche, including cilantro, onion, tomato, pepper and avocado are usually consumed raw, in ceviche, they represent a major source of pathogenic bacteria that is associated to foodborne illness. In this study the ceviche samples tested had a variety of produce ingredients that constituted a factor of variability in the results. Among exogenous bacteria, we identified several members of the *Enterobacteriaceae* genus such as *Pantoea*, *Rosenbergiella* and *Morganella*. *Pantoea*, primarily found in S15 but also in S13, these genera are known to be associated as a pathogen or epiphytic bacteria in edible vegetables, e.g. *P. agglomerans* in tomato and onion (Nadarasah and Stavrinides 2014; Oie et al. 2008) which are main ingredients of Mexican made ceviche. *P. agglomerans*, has also been associated to soft bone infections (Cruz et al. 2007), although this opportunistic bacterium is not known to be a foodborne pathogen it has been associated as a contaminant in powered infant formula milk (Mardaneh and Dallal 2013). We hypothesized that *Pantoea* sp. found in ceviche have its origin in the vegetables used in the dish preparation since it was almost undetectable in the laboratory made samples S16 and S17 respectively.

*Rothia* sp., found in a highest frequency in S13 may cause a significant number of infections to immunocompromised hosts despite they are part of the normal microbiome of human oropharynx and respiratory tract (Ramanan et al. 2014), it is possible that *Rothia* in ceviche have its origin from anthropogenic contamination during its preparation.

The genus *Weissella*, found primarily in S15 but also in S13 is a gram-positive coco-bacilli, often misidentified by traditional microbiological methods but commonly found in fermented foods, including vegetables, meat, fish and raw milk. Although, this genus is rarely a known cause of human infections. Conversely, few *Weissella* strains may act as probiotics due its antimicrobial activity e.g. *W. hellenica* isolated from flounder intestine act as a defensive agent against host pathogens such as *Edwardsiella*, *Pasteurella*, *Aeromonas* and *Vibrio* (Abriouel et al. 2015; Cai et al. 1998).

*Aeromonas*, a genus found in high presence in S17, but present in all the tested samples is one of the etiological agents associated to disease-causing pathogen of fish and responsible of infectious complications in humans. This genus can be isolated from almost every ecological niche in earth, including aquatic habitats, fish, food and soils and it can be present in ceviche due multiple factors and represented by one or more species. In food, *Aeromonas* strains have been widely reported in dairy products, specifically unpasteurized milk, yogurt and cheese from diverse geographical origins.

*Prevotella*, was detected in all samples except S15 and S17, this agent is known to be a part of human gut, oral and vaginal microbiota, this bacterium is often an
indicator of fecal contamination of water (Fogarty and Voytek 2005; Koskey et al. 2014). However, in ceviche the high presence of this agent could be associated to low hygienic conditions in produce handling. Pseudomonas, a naturally widespread genus is a potential human pathogen previously reported in spinach, lettuce and red cabbage, it was found in every sample and probably its presence is due either contaminated vegetables and contamination of fish meat. 

In the present study, we did not found evidence of etiological agents associated to freshwater fish species outbreaks of foodborne illness such as Salmonella, Escherichia coli, Clostridium, Staphylococcus or Bacillus. In our knowledge, this study presents the first description of bacterial diversity in raw fish food using a metagenomic approach, and our study is focused in ceviche samples made in Reynosa, Mexico. However, the highly diverse conditions of preparation, side ingredients, origin of fish meat, and hygienic handling of this kind of foods in different regions of Mexico and other countries makes necessary further analyses in order to compare the dynamics of bacterial communities in ready-to-eat raw seafood.

Acknowledgments: We thank the Secretaría de Educación Pública de Mexico for their support given to finance this project through the PRODEP-NPTC fellowship granted to Humberto Martínez-Montoya and Myriam Elizabeth Ramírez-Martínez

REFERENCES


J Microbiol Biotech Food Sci / Ramirez-Martinez et al. 20xx : x (x) e3321


