Determination of Puerto Rican Cacao’s Mucilage Microbial Population and their Role in Fermentation via Metagenomic Analysis

A collaborative project between
ISBiot, Inter American University of Puerto Rico, Barranquitas Campus
and
Jeanmarie Chocolat, Aguada

Cristian M. González Rivera
B00411406
Biot 4954
Research Methods in Biotechnology
Professor Melendez
I. Problem

The chocolate industry in Puerto Rico is not completely developed, but has a great potential for economic growth. However, little is known about Puerto Rican cacao strains and its fermentation process. A previous report in phylogenetic analysis of the community diversity of a single spontaneous cacao bean box fermentation sample through a metagenomics approach\(^1\); in this investigation it was found that the cocoa beans contain more bacterial and fungal diversity than found in previous investigations. This indicates a great advantage in using metagenomics techniques for these types of investigations. Due to weather conditions in Puerto Rico the fermentation process of cacao beans take six days and a timeline was made to determine which microbe species (bacterial or yeast) and find when are they active during the process. It is proposed a sampling method that includes taking the mucilage of cacao, knowing that it’s present in fermentation since the opening of the pod, isolating DNA, amplifying the 16S rRNA gene for bacteria and the 18S rRNA gene for yeast with a PCR. This will make possible to identify microbes after NGS Sequencing and a bioinformatics analysis. Finally, once identified it is possible to isolate these microorganisms and culture them. This could bring a better understanding of the process and the part it takes in the final quality of the product. The investigation can lead to further develop new mechanics to increase the product quality or make different variations from the same strains recalibrating the fermentation system by adding or subtracting microbes involved in the process.
II. Justification

A lot of investigations about cacao can be found with similar objectives but most of them are done in other countries. This could be starting point to understand how or if Puerto Rican cacao is unique in a way that it could be used as a major exportation product. Firstly, it is important to know how the fermentation works from a micro perspective and which species of microbes are involved to analyze and calibrate the process and improve it so a higher quality product can be obtained. Some industrial ways of fermentation include the use of starter microbial starter cultures and thus enhancing the efficiency. But even with the controlled fermentation, the process itself is still spontaneous and relies greatly in the indigenous microbial population\(^2\). From a medical perspective, it is a fact that cacao has medicinal properties with a high content of phenolic antioxidants. With such impacts like minimize risk for diabetes by influencing insulin resistance, protect human skin from UV radiation, cognitive function improvement and even affects our mood\(^3\). By determining and improving the quality of the final product with use of controlled fermentation the health of chocolate consumers can be improved and a successful business model can be created from the results. Furthermore, more tests would be necessary like isolation and culture of the microbial population and check if cacao from other regions contain the same microbiota as the cacao in Aguada.
III. Theoretical framework

Chocolate has been consumed historically as early as 460 AD as food and beverage but also because it contains healing properties. Cacao or cocoa has been used as a primary medicine or as secondary to deliver other medicines. As treatment, cacao has over 100 uses that have been documented but the most common are: (1) to help emaciated patients gain weight; (2) stimulate the nervous system and (3) to improve digestion and elimination. Indigenous people from South America that consume three 10-ounces cups of a cacao beverage have very low hypertension rates, blood pressure doesn’t increase with age, there is a low rate of diabetes, strokes and cancer.

The fine chocolates market has been growing parallel to the overall commercial mass chocolate market and it’s <5% the global chocolate consumption. In 2010/11, it was reported by the International Cocoa Organization that seven largest fine flavor chocolate producing countries were Ecuador, Dominican Republic, Colombia, Venezuela, Madagascar, Nicaragua and Bolivia. Because of the elevated prices for quality cocoa, small producers that account for the bulk of global cocoa supplies can receive a good income. There is a transition from bulk cacao to a higher-value cacao but because of the informal distribution of planting material without the guarantee of performance and identity and the lack of suitable genetic materials the transition is affected.

The fermentation of cacao beans is one of the few fermentation process that rely on spontaneous microbial processes in the modern food industry. In the fermentation of the cocoa pulp there is a variable and complex microbiota, which leads to inconsistent production efficiency and cocoa quality. Results in previous investigations show that yeasts, lactic acid
bacteria (LAB) and acetic acid bacteria (AAB) are the main microorganisms that participate in the fermentation, during which the mucilage that surrounds the beans is liquefied. Even when the process is well known, because it relies on spontaneous processes there are variations in the active microbiota and thus it’s not a consistent process. Because of this variability, most of the bulk food fermentation process, this includes beer, wine and bread, are nowadays inoculated with starter culture to obtain a consistent product quality with reproducible parameters\(^5\). The parameters for cacao beans fermentation should been taken into consideration as to improve it, the microbiota is a topic of interests, from there a starter culture can be developed to give aromas or flavors in the future.

The fermentation of cacao is divided in two main stages: (1) microorganisms consume the pulp and produce metabolites such as organic acids and ethanol; (2) inside the cotyledon, occur some hydrolytic reactions with the help of the metabolites produced in the first stage. The first microorganisms to start the fermentation are yeasts, they metabolize the carbohydrates in the pulp into ethanol to generate an anaerobic environment. Then, Lactic Acid Bacteria (LAB) convert citric acids into lactic acid, increasing the pH. Finally, acetic acid bacteria (AAB) convert the ethanol into acetic acid in aerobic conditions. Exothermic reactions of acetic acid and ethanol occur, inducing high temperatures and contributes to the death of the bean releasing endogenous enzymes\(^6\).

The dominants microorganisms in the spontaneous fermentation were \(S. \text{cerevisiae}, H. \text{uvarum}, H. \text{guilliermondii}, Lactobacillus fermentum, Pedicoccus \text{sp.}, \text{and} \ Acetobacter pasteurianus\)^6. In another test done with by Denaturing Gradient Gel Electrophoresis (DGGE), the microorganism assessed were: \(S. \text{cerevisiae}, P. \text{kluveri, Candida \text{sp.}, Pediococcus \text{sp.}, and A. pasteurianus}\). Also in the same investigation with gas chromatography/mass spectrometry
(GC/MS), sixty-seven volatile compounds were detected. The main group of volatile compounds was esters with 39%. Several papers has studied the diversity in microorganism in spontaneous cacao pulp fermentation. Mostly the investigations done rely on culture-dependent techniques like plating and characterization of the colonies but these methods have some notable drawbacks like media bias and some species cannot be cultured thus, remaining unidentified. Therefore, molecular techniques like DGGE, can be recommended because it can detect even uncultivable species. Recently, the development of microbial population profiling methods with direct sequencing of mixed communities allows high-throughput and analysis of the microbial population in the spontaneous fermentation. Metagenomics methods can help identify species that are not identifiable by culture-dependent and culture-independent methods. But it also has some disadvantages, this approach has a dependency on the extraction method and it can influence the structure of detected population and low discriminative power between species with the lack of appropriate primers.

The microbial ecosystem of the cacao beans pulp has species such as Hanseniaspora sp., Saccharomyces cerevisiae, Lactobacillus fermentum, Lactobacillus plantarum, and Acetobacter pasteurianus. But as mentioned, both culture-dependent and culture-independent have some drawbacks which undermines an accurate view on the microbial population of the ecosystem, and also implies that more, yet unidentified species, might play a role in the fermentation process which can lead to wrong results. To counter the drawbacks of culture-dependent techniques, culture-independent techniques are used like PCR-DGGE by amplifying small fragments of targeted genes or rRNA gene clone library sequencing have also been used. These methods have been used with metagenomics approach but because they rely on PCR the results might be biased and suffer from preferential DNA amplification.
Moreover, these techniques are based on the amplifications of small, variable regions of the 16S (bacteria) or 26S rRNA genes (yeast), but the resolution of these, in some genera is limited. The sequencing data of a whole-community obtained by high-throughput parallel sequencing of metagenomics DNA, overcome the drawbacks of the aforementioned methods\textsuperscript{1}.

It has been noted that cacao fermentation is one of the few processes that still relies on spontaneous fermentation even when other industries like beer and wine have a more controlled and consistent process by using a starter culture. However, the flavor of chocolate is influenced by many parameters, this includes the genotype and growing conditions of the cacao trees, also a poor fermentation process leads to a bitter tasting chocolate.
IV. Objectives

1. Obtain Cacao pulp samples for metagenomics analysis.
2. Isolate DNA from the microbial ecology for sequencing of the ribosomal subunits of yeasts and bacteria.
3. Identify the microbes, yeast and bacteria, via bioinformatics analysis involved in the fermentation process.
4. Calibrate the fermentation process for prime results in the product.

V. Methodology

Note: This project will be developed from the fermentation process to the calibration of DNA extraction.

A. Fermentation Process-Materials preparation
   a. To ferment the cacao correctly a good quantity of cacao pods are required.
      For this experiment a box with the following measures will be used:
      i. Width: 2ft.
      ii. Length: 2ft
      iii. Height: 1ft
   b. It is highly important to use untreated wood for the making of the box, given that the box microbial ecology could be somehow involved in the fermentation.
   c. In the floor of the box, holes were done with a drill to drain and oxygenate the seeds during the fermentation. They were done in columns and rows of eighteen cm between each hole and columns.

B. Fermentation Process- Seeds extraction
a. The fermentation process starts once all the seeds are taken from the inside of the pods. The following steps should be taken to break them:

i. With a blunt object strike the pods until they open, it shouldn’t be done too violently or the seeds could be damage or fall from the pod and get crossed contaminated.

ii. Inspect the inside of the pulps carefully, if the seeds seem damaged or dry they are not good for the fermentation and should be discarded. The healthy seeds have a sugar coat mucilage and those are the one that should be added to the box.

iii. The pulp is the most important material, is where the sugars are located and the ones that will be fermented first, therefore when putting the seeds on the box try to get all the pulp from the pod as possible.

iv. The seeds come attached to a vine inside the pod, be sure to not add any of it in the fermentation box.

v. A good quantity of seeds are needed for the fermentation to take place. Once all the seeds are added to the box add some plantain leaves covering them, this helps to maintain the heat in the exothermic reactions. When the seeds are well covered, close the box and make sure it has limited airflow.

vi. The fermentation process in Puerto Rico takes six days. For the fermentation to be even through all the seeds, some movements should be done in the following time periods:
1. After the first 48 hours

2. After every 24 hours until fermentation is over

vii. Once the fermentation process is finished the seeds can be taken to be dried.

C. Fermentation Process-Sample taking

Note: This table of the sampling strategy was prepared by Dr. Arun

<table>
<thead>
<tr>
<th>Sample number</th>
<th>Hours</th>
<th>Biological replicate</th>
<th>Biological replicate</th>
<th>Biological replicate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>36</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>6</td>
<td>48</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>56</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>72</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>80</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>90</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>120</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>12</td>
<td>144</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>192</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>Pod</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>Seed</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>Container</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td>leaf</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

a. The samples of the cacao pulp should be taken from the seeds inside the box in the time periods showed on the table.

b. Gloves should be worn all the time during the sampling to avoid crossed contamination. Recollection materials like the centrifuge tubes should be sterile. Good aseptic techniques must be used.
c. To take the pulp, hold a seed with forceps, then with a scissor take the pulp from the seed. This must be done carefully to avoid taking seeds samples instead of the pulp. All tools must be sprayed with 70% alcohol beforehand. When adding the pulp to the centrifuge tubes, open the cap carefully and put all the pulp inside, when a lot of pulp accumulates, close the tube well and do a fast movement to make it descend to the bottom.

d. When finished collecting a sample in the tube, close it and cover it with parafilm, then refrigerator immediately.

**Note:** All samples must be taken this way. To refrigerate the samples faster, the sample could be done with the centrifuge tube in a rack inside a little cooler with ice.

D. **Plant Genomic DNA Extraction using CTAB**

**Note:** All credit for this protocol goes to Nirmal Joshee and his laboratory.

a. **Materials**
   i. CTAB buffer
   ii. Microfuge tubes
   iii. Mortar and Pestle
   iv. Liquid Nitrogen
   v. Microfuge
   vi. Absolute Ethanol (ice cold)
   vii. 70% Ethanol (ice cold)
   viii. 7.5 M Ammonium Acetate
   ix. 55°C water bath
   x. Chloroform : Iso Amyl Alcohol (24:1)
   xi. Water (sterile)
   xii. Agarose
   xiii. 6x Loading Buffer
   xiv. 1x TBE solution
   xv. Agarose gel electrophoresis system
   xvi. Ethidium Bromide solution
b. **CTAB buffer (100ml) preparation**
   
i. 2.0 g CTAB (Hexadecyl trimethyl-ammonium bromide)

   ii. 10.0 ml 1 M Tris pH 8.0

   iii. 4.0 ml 0.5 M EDTA pH 8.0 (Ethylenediaminetetra Acetic acid disodium salt)

   iv. 28.0 ml 5 M NaCl

   v. 40.0 ml H2O

   vi. 1 g PVP 40 (polyvinyl pyrrolidone (vinylpyrrolidine homopolymer) MW: 40,000)

   vii. Adjust all to pH 5.0 with HCl and make up to 100 ml with H2O.

c. **1 M Tris (pH 8.0) preparation**
   
i. Dissolve 121.1 g of Tris base in 800 mL of H2O. Adjust pH to 8.0 by adding 42 ml of concentrated HCl. Allow the solution to cool to room temperature before making the final adjustments to the pH. Adjust the volume to 1 L with H2O. Sterilize using an autoclave.

d. **5x TBE buffer preparation**
   
i. 54 g Tris base

   ii. 27.5 g boric acid

   iii. 20 ml of 0.5M EDTA (pH 8.0)

   iv. Make up to 1 L with water.

   v. To make a 0.5x working solution, do a 1:10 dilution of the concentrated stock.
e. 1% Agarose gel preparation
   i. 1 g Agarose dissolved in 100 mL TBE.

f. Procedure for DNA extraction
   i. Grind 200 mg of plant tissue to a fine paste in approximately 500 μL of
      CTAB buffer.
   ii. Transfer CTAB/plant extract mixture to a microfuge tube.
   iii. Incubate the CTAB/plant extract mixture for about 15 min at 55°C in a
      recirculating water bath.
   iv. After incubation, spin the CTAB/plant extract mixture at 12000 g for 5
      min to spin down cell debris. Transfer the supernatant to clean
      microfuge tubes.
   v. To each tube add 250 μL of Chloroform:Iso Amyl Alcohol (24:1) and
      mix the solution by inversion. After mixing, spin the tubes at 13000 rpm
      for 1 min.
   vi. Transfer the upper aqueous phase only (contains the DNA) to a clean
      microfuge tube.
   vii. To each tube add 50 μL of 7.5 M Ammonium Acetate followed by 500
      μL of ice cold absolute ethanol.
   viii. Invert the tubes slowly several times to precipitate the DNA.
       Generally, the DNA can be seen to precipitate out of solution.
       Alternatively, the tubes can be placed for 1 hr at 20 °C after the addition
       of ethanol to precipitate the DNA.
ix. Following precipitation, the DNA can be pipetted off by slowly rotating/spinning a tip in the cold solution. The precipitated DNA sticks to the pipette and is visible as a clear thick precipitate. To wash the DNA, transfer the precipitate into a microfuge tube containing 500 μL of ice cold 70% ethanol and slowly invert the tube. Repeat. (Alternatively, the precipitate can be isolated by spinning the tube at 13000 rpm for a minute to form a pellet. Remove the supernatant and wash the DNA pellet by adding two changes of ice cold 70% ethanol).

x. After the wash, spin the DNA into a pellet by centrifuging at 13000 rpm for 1 min.

xi. Remove all the supernatant and allow the DNA pellet to dry (approximately 15 min).

xii. Do not allow the DNA to over dry or it will be hard to re-dissolve.

xiii. Resuspend the DNA in sterile DNase free water (approximately 50-400 μL H2O; the amount of water needed to dissolve the DNA can vary, depending on how much is isolated). RNaseA (10 μg/mL) can be added to the water prior to dissolving the DNA to remove any RNA in the preparation (10 μL RNaseA in 10 mL H2O).

xiv. After resuspension, the DNA is incubated at 65°C for 20 min to destroy any DNases that may be present and store at 4°C.

xv. Agarose gel electrophoresis of the DNA will show the integrity of the DNA, while spectrophotometry will give an indication of the concentration and cleanliness.
g. DNA quality confirmation procedure

i. Prepare a 1% solution of agarose by melting 1 g of agarose in 100 mL of 0.5x TBE buffer in a microwave for approximately 2 min. Allow to cool for a couple of minutes then add 2.5 μL of ethidium bromide, stir to mix.

ii. Cast a gel using a supplied tray and comb. Allow the gel to set for a minimum of 20 min at room temperature on a flat surface.

iii. Load the following into separate wells:

iv. 10 μL 1kb ladder.

v. 5 μL sample + 5 μL water + 2 μL 6x loading buffer.

vi. Run the gel for 30 min at 100 V.

vii. Expose the gel to UV light and photograph (demonstration).

viii. Confirm DNA quality, presence of a highly resolved high molecular weight band indicates good quality DNA, presence of a smeared band indicates DNA degradation.

Note: If the CTAB method doesn’t work the DNeasy PowerSoil Kit can be used
VI. Bibliography


### VII. Chronogram

<table>
<thead>
<tr>
<th>Sun</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>20</td>
</tr>
<tr>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
<td>25</td>
<td>26</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>CTAB Extraction</td>
<td>CTAB Extraction</td>
<td>CTAB Extraction</td>
<td>CTAB Extraction</td>
<td>CTAB Extraction</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTAB Extraction</td>
<td>CTAB Extraction</td>
<td>CTAB Extraction</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Determination of Puerto Rican Cacao’s Mucilage Microbial Population and their Role in Fermentation via Metagenomic Analysis

<table>
<thead>
<tr>
<th>Sun</th>
<th>Mon</th>
<th>Tue</th>
<th>Wed</th>
<th>Thu</th>
<th>Fri</th>
<th>Sat</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>New</td>
<td>Take</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>fermentation</td>
<td>samples</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Take</td>
<td>Take</td>
<td>Take</td>
<td>Take</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>samples</td>
<td>samples</td>
<td>samples</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
<td>15</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>19</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>29</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CTAB Extraction week long

June