Banana Disease-Protective Microbiome Enrichment and Analysis

Lahowetz, Rachel M\textsuperscript{1}, Abdollahi Aghdam, Shiva\textsuperscript{1}, Singh, Simrandeep\textsuperscript{1}, Brown, Amanda MV\textsuperscript{1}

\textsuperscript{1}Department Of Biological Sciences, Texas Tech University, Lubbock, TX.
Introduction

- Bananas are an important agricultural project threatened by variety of diseases

- **Microbiome**: collective microorganisms in an environment
  - **Endophytes**: organisms that live in a plant without causing harm; part of the plant microbiome

- Plants contain a complex microbiome that may aid in protection against disease\(^9,12\)
  - This includes a virome – some viruses may provide protection against certain pathogens\(^9,12\)

- This project: Bioinformatic analysis of plant microbiome
  - Determine microbes present, what they do, and how they protect against disease
  - Processing and analysis of DNA and RNA sequences
  - multi-omics approach: Project will incorporate genome (DNA), transcriptome (RNA), proteome (protein), etc.

*We hypothesize that cultivation and domestication have altered plant microbiomes, possibly reducing disease resistance.*
Significance

• Banana: major food staple and agricultural crop in many countries
• Threatened by variety of diseases
  • Mid-twentieth century: Gros Michel banana industry wiped out by *Fusarium oxysporum* f. sp. *cubense* (*Foc*)
  • Cavendish – current cultivar: susceptible to *Foc* Tropical Race 4 (Panama Disease)
• This project:
  • Address whether less-cultivated bananas have a more diverse microbiome that gives greater immunity
  • Aid in production of a disease-protective inoculum that could be used to protect vulnerable banana varieties such as the Cavendish
Methods: Part 1

1.) Surface-sterilize, extract above and below ground samples
   Leaf and corm/root, yielding 14 samples of 100 grams each

2.) Nycodenz density centrifugation to enrich microbial cells

3.) Extract microbiome DNA and RNA – Qiagen DNEasy and RNEasy
   14 samples (7 leaf, 7 root/corm) x 2 (DNA/RNA) = 28 Samples

4.) Prepare shotgun metagenomic Illumina libraries
   QIAseq FX (DNA), RiboZero (deplete rRNA), NEBNext Ultra (RNA)
   Shotgun metagenomics = sequence all genes

5.) Sequence – Illumia HiSeq

6.) Combine, trim, and assemble sequences
    Programs used: Pear\textsuperscript{11}, Trimmomatic\textsuperscript{3}, metaSPAdes\textsuperscript{7}/rnaSPAdes

7.) Part 2: Bioinformatic Processing and Analysis
Methods: Part 2

Bioinformatic Processing and Analysis of DNA and RNA Sequences

Reference Genome Mapping of Enriched Sample Reads –

- Determine effectiveness of microbial enrichment
- Use unassembled DNA sequences
- Enrichment appears to be successful

Shared potential antimicrobial biosynthesis-like genes

- Comparison of DNA samples to custom protein database of potential antimicrobial compound-producing genes
- Confirm hits by comparison to nr (protein) database

Virome Investigation

- Blast DNA and RNA samples to viral database
- Confirm hits by comparison to nt (nucleotide) database
- Gene annotation
- Find viral orthologs
- Ongoing
Reference Genome Mapping of Enriched Sample Reads

- bwa/samtools\textsuperscript{6} used to map DNA sequences to reference genome.
- Non-mapped genes are primarily microbial; gives approximate level of enrichment

**Biological Significance:**
- Approx: genome size of banana: 331 Mb and Average microbial genome: 6 Mbp
- Accounting for this, the microbe:banana cell ratio ranges from 2:1 (MBL) to 3007:1 (MBR)
  - Vast majority – ratio above 20:1
- **This means that overall, enrichment was successful**
- Enrichment may have worked better for below-ground samples, or these samples may contain more microbes.

### Reference Genome Mapping of Enriched Sample Reads

#### Percent Mapped Reads

<table>
<thead>
<tr>
<th>Leaf Sample</th>
<th>Root/Corm Sample</th>
<th>Banana Name</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCL</td>
<td>DCR</td>
<td>Dwarf Cavendish</td>
<td>AAA</td>
</tr>
<tr>
<td>WHL</td>
<td>WHC</td>
<td>William’s Hybrid</td>
<td>AAA</td>
</tr>
<tr>
<td>MBL</td>
<td>MBR</td>
<td>Musa balbisiana</td>
<td>BB</td>
</tr>
<tr>
<td>BBL</td>
<td>BBC</td>
<td>Black balbisiana</td>
<td>BB</td>
</tr>
<tr>
<td>FHL</td>
<td>FHC</td>
<td>FHIA-25</td>
<td>AAB</td>
</tr>
<tr>
<td>MTL</td>
<td>MTC</td>
<td>Musa textilis</td>
<td>--</td>
</tr>
<tr>
<td>MSL</td>
<td>MSR</td>
<td>Musa sikkimensis</td>
<td>--</td>
</tr>
</tbody>
</table>

*Microbiome percentage includes banana mitochondrial reads. However, preliminary data suggests this percentage is negligible.*
Shared potential antimicrobial biosynthesis-like genes

- Made custom protein database of potential antimicrobial biosynthesis-like genes
- Conduct BLAST of DNA samples to custom database
  - Find regions of similarity between our sequences and those on custom database
  - Second BLAST to protein (nr) database – confirm hits
- Results and Biological Significance:
  - Above and below-ground microbiome of each plant contains many of the same antimicrobial biosynthesis-like genes
  - More unique hits in below-ground samples – greater diversity

Database of potential antimicrobial biosynthesis-like genes include genes related to:

<table>
<thead>
<tr>
<th>Keyword</th>
<th>Lactone</th>
<th>Lactam</th>
<th>Protease inhibitor</th>
<th>Quorum sensing</th>
<th>Biofilm</th>
<th>Homoserine lactone (HSL)</th>
<th>Terpene</th>
<th>Bacteriocin</th>
<th>Ectoine</th>
<th>Phosphonate</th>
<th>Indole</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>67</td>
<td>23</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>340</td>
<td>27</td>
<td>4</td>
<td>20</td>
<td>13</td>
</tr>
</tbody>
</table>
Virome Investigation

• Analysis of DNA and RNA sequences to identify viruses that affect banana plant and/or its pathogens
  • Example: BBTV can cause severe crop losses, but also shown to increase plant resistance to \textit{Foc}^{12}
  • Goal: identify other viruses that affect plant heath and microbiome, positively or negatively

• Ongoing process
  • Blast DNA and RNA samples to a viral database, confirm hits via comparison to nucleotide (nt) database
  • Prokka: gene annotation\textsuperscript{8}
  • Orthofinder: identify viruses and their genes in samples\textsuperscript{4}
    • No marker gene for all viruses exists – use orthologs from known viruses to uncover other viral genes

• Database used:
  • All virus sequences on NCBI virus database at time of download (January 2020)
  • 3,250,606 entries
Future Directions

• Further refinement of antimicrobial biosynthesis-like gene database

• Gene Ontology (GO)
  • Gene function/products

• Transcriptomics (mRNA)
  • Confirm gene expression of antimicrobial biosynthesis genes

• Metabolomics (proteins)
  • Confirm presence of antimicrobial compounds

• Virome Investigation
  • Continue processing and analysis
  • Geneious (phylogenetic analysis)
  • Continue work on RNA samples

**Ultimate goal:** use this data to develop a disease-protective inoculum that can be used to protect susceptible banana strains against a variety of diseases.
Acknowledgements

Past students who worked in Dr. Brown’s lab

Funding support was provided by Texas Tech University and the Department of Biological Sciences. Funding to R.L. was provided by TTU Undergraduate Research Scholar program.

HPCC resources provided by High Performance Computing Center (HPCC) at Texas Tech University
References