Name: Ally Bonney

Title: Genome sequencing and annotation of *Pseudomonas veronii* isolated from Oregon State University soil and 16S rRNA characterization of Corvallis, OR soil microbial populations

Date: January 29, 2015 – February 24, 2015

Purpose

We used both culture-dependent and culture-independent approaches to purify genomic DNA, and PCR to amplify 16S rRNA genes from soil sample microbes. We then sequenced the 16S rRNA amplicons of the bacterial species as well as the genomic DNA of one chosen cultured organism to identify and describe the soil microbial ecology.

Methods

Each member of our group obtained a soil sample. We purified genomic DNA using both a culture-dependent and independent approach. In our culture-dependent approach, bacterial species were grown on LB agar and genomic DNA from a single colony of each sample was purified. In our culture-independent approach, genomic DNA from entire sample communities was purified. PCR was used to amplify 16S rRNA genes. The chosen genomic DNA sample and the 16S rRNA amplicons were sequenced via high-throughput (Illumina Mi-Seq), and 16S rRNA sequences were processed with BLAST and PRIMER. We assembled genomic DNA contigs with Velvet and annotated the genome with RAST. The procedure given on the "High-Throughput (Illumina Mi-Seq) DNA Sequence Analysis of 16S Ribosomal RNA Genes from Uncultured Bacterial Communities" handout was followed (2).

Results

In order to study the bacterial composition of our soil samples as well as isolate an organism for complete genome sequencing, we purified genomic DNA from a chosen cultured colony and amplified 16S rRNA genes via PCR from both cultured and non-cultured samples. We then performed sequence analyses on the amplicons and genomic DNA, giving us information on the species present, their abundance, and relevant genetic information of our chosen isolate.

My soil sample was taken from a paddock in Harmony Training Center in Corvallis, OR. After streaking this sample on an agar plate and incubating at 25°C for 5 days, three separate culturable organisms were evident: a fungal specimen, a bacterial species forming large, irregular off-white colonies, and a bacterial species forming small, round, yellow colonies. The large off-white colonies and the fungal specimen grew most rapidly, while only a few of the small yellow colonies appeared on the
The large off-white colonies were composed of gram-positive rods. PCR amplification of the 16S rRNA gene from a chosen large off-white colony was successful (Figure 1), yielding an amplicon 608 bp long (Figure 2). Both the purified and unpurified PCR amplicon bands contained about 80ng of DNA, giving the purification process a 100% yield. Since 5ul eluate was loaded onto each lane, there was 16ng/ul DNA in each sample (80ng/5ul = 16ng/ul).

Figure 1: Agarose gel electrophoresis of PCR amplicon fragments of 16S rRNA genes purified from isolated cultured soil bacterium. Lane 1 contains a molecular weight ladder. Lane 2 contains a control in which no DNA was added to the PCR reaction. As expected, there is no product present. Lane 3 contains unpurified 16S rRNA amplicon. Lane 4 contains purified 16S rRNA amplicon. Due to the comparable band concentrations, very little amplicon was lost during the purification process. Selected fragment lengths (in bp) of the ladder are given on the left.

Figure 2: Fragment size standard curve. Generated from a semilogarithmic plot of molecular weight marker fragment sizes (bp) vs. mobility on the gel (cm). A line of best fit and best fit equation are shown on the graph. Based on the standard curve, the PCR amplicon was 608 bp long.

Figure 3: Agarose gel electrophoresis of PCR 16S rRNA amplicon purified using culture independent method. Lane 1 contains a molecular weight ladder. Lane 19 contains purified amplicon from my soil sample obtained via culture independent methods. No product can be seen on the gel.
Sequence analysis of the 16S rRNA purified from the large off-white colony with BLAST indicated that it was *Bacillus pumilus*. The culture-independent approach did not yield a discernable amount of 16S rRNA amplicon on the agarose gel (Figure 3). Other members of my group isolated *Pseudomonas koreensis*, an *Arthrobacter* spp. (most likely *humicola* or *globiformis*), *Pseudomonas putida*, and *Bacillus simplex*. A comparison of sample ecology is shown in Figure 4.

The sequenced genome belonged to *Pseudomonas veronii*. Genome size was 6,058,927 bp. Using a k-mer of 95, n50 contig size was 85564 bp, and max contig size was 344433. Data for other k-mer values are given in Table 1. Genome annotation by RAST indicated the presence of 5396 coding sequences and 72 RNA genes. The depth of coverage was about 80. The recA base sequence is given on Attachment 1. The SSU rRNA gene contigs only included rRNA genes, and the depth of coverage was 1110. This is much greater than for the majority of the other genes, whose depths of coverage fell within the range of 60-100. No RepABC plasmid was detected by the RAST annotation. However, unique genes capable of metabolizing aromatics and taking up heavy metals were present.

<table>
<thead>
<tr>
<th>K-mer</th>
<th>75</th>
<th>79</th>
<th>85</th>
<th>87</th>
<th>95</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nodes</td>
<td>357</td>
<td>348</td>
<td>425</td>
<td>297</td>
<td>281</td>
</tr>
<tr>
<td>n50</td>
<td>137462</td>
<td>106670</td>
<td>137462</td>
<td>89282</td>
<td>85564</td>
</tr>
<tr>
<td>Max</td>
<td>500470</td>
<td>400613</td>
<td>500470</td>
<td>344433</td>
<td>344433</td>
</tr>
<tr>
<td>Total</td>
<td>6053775</td>
<td>6052631</td>
<td>6053554</td>
<td>6052255</td>
<td>6052189</td>
</tr>
</tbody>
</table>

Figure 4: Bray-Curtis comparison of microbial population resemblance.
Discussion

Soil is often prime habitat for a rich and diverse community of microorganisms. Many soil bacteria with novel genes have yet to be isolated and characterized. Following sampling from various locations in the Corvallis, OR area, *Pseudomonas veronii* was isolated and its genome sequenced. First isolated from spring water in France, the organism inhabits a variety of environments, and due to its unique metabolic abilities, can be isolated from sites contaminated with aromatic compounds (3). *P. veronii* has very promising applications in bioremediation. Genome assemblage via Velvet and annotation via RAST indicated a genome 6,052,544 bp long containing 5396 protein coding genes and 72 RNA genes. This length is comparable to other *P. veronii* isolates in the literature, which are around 6Mb long (3). It contained no RepABC plasmid, but a number of genes important for metabolizing aromatics and heavy metals such as iron. Sequencing and BLAST analysis of 16S rRNA of a large off-white colony cultured from my sample (following PCR amplification) identified the presence of *Bacillus pumilus*, a gram-positive, rod-shaped organism. The 16S rRNA amplicon was 608 bp long, which was slightly longer than the expected length of 530 bp. This could be due to error in measuring mobility on the gel. Since there was no band seen in lane 2 of Figure 1, little contaminating DNA was in the 16S rRNA amplicon of the culture-dependent purification technique. However, my sample did not yield a discernable amount of PCR amplicon following the culture-independent method (Figure 3). *Pseudomonas koreensis*, an *Arthrobacter* spp. (most likely *humicola* or *globiformis*), *Pseudomonas putida*, and *Bacillus simplex* were cultured from other soil samples. As seen in Figure 4, the park microbial communities are more closely related to each other than were communities from other sampling sites.

Conclusion

We isolated *Pseudomonas veronii* from soil and sequenced its genome, which was 6,058,927 bp long and included genes necessary for metabolizing aromatic compounds and iron. The 16S rRNA purified from a large off-white colony from my soil sample was identified as *Bacillus pumilus*, and other group members identified the presence of *Pseudomonas koreensis*, an *Arthrobacter* spp. (most likely *humicola* or *globiformis*), *Pseudomonas putida*, and *Bacillus simplex* in their samples.
References

