What can we learn from sequencing mycetoma fungi?

Anastasia (Ana) Litvintseva, PhD

Mycotic Diseases Branch, Centers for Disease Control and Prevention, USA

11th ECTMIH, Liverpool 2019
The Whole Genome Sequencing (WGS) Process

WGS is a laboratory procedure that determines the order of bases in the genome of an organism in one process. WGS provides a very precise DNA fingerprint that can help link cases to one another allowing an outbreak to be detected and solved sooner.

1. DNA Extraction
   - Scientists take bacterial cells from an agar plate and treat them with chemicals that break them open, releasing the DNA. The DNA is then purified.

2. DNA Shearing
   - DNA is cut into short fragments of known length, either by using enzymes “molecular scissors” or mechanical disruption.

3. DNA Library Preparation
   - Scientists make many copies of each DNA fragment using a process called polymerase chain reaction (PCR). The pool of fragments generated in a PCR machine is called a “DNA library.”

4. DNA Library Sequencing
   - The DNA library is loaded onto a sequencer. The combination of nucleotides (A, T, C, and G) making up each individual fragment of DNA is determined, and each result is called a “DNA read.”

5. DNA Sequence Analysis
   - The sequencer produces millions of DNA reads and specialized computer programs are used to put them together in the correct order like pieces of a jigsaw puzzle. When completed, the genome sequence containing millions of nucleotides (in one or a few large pieces) is ready for further analysis.
Advantages of WGS for mycetoma community

• Better understanding of etiology of mycetoma

• Identification of novel targets for new diagnostics methods
**Understanding etiology of mycetoma: better species identification**

Molecular methods based on a single gene do not always provide enough resolution for identification of species. Rojas et al., 2016

### TABLE 2  Phenotypic and molecular data from eumycetoma agents

<table>
<thead>
<tr>
<th>Case</th>
<th>Morphological identification</th>
<th>Molecular identification</th>
<th>ITS GenBank accession number</th>
<th>GenBank accession number</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>This study/Reference</td>
<td></td>
<td>Identity</td>
</tr>
<tr>
<td>1</td>
<td>Madurella mycetomatis</td>
<td>Madurella pseudomykotomatis</td>
<td>KT834405/EU815933</td>
<td>596/597 (99%)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Exophiala jeanselmei</td>
<td>Cyphellophora oxyysora</td>
<td>KT323976/KM3968285</td>
<td>600/602 (99%)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Exophiala sp.</td>
<td>Exophiala oligosperma</td>
<td>KT323978/DQ836792</td>
<td>655/655 (100%)</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Exophiala dermatitidis</td>
<td>Exophiala dermatitidis</td>
<td>KT323977/AY213651</td>
<td>657/657 (100%)</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Scedosporium apiospermum</td>
<td>Scedosporium apiospermum</td>
<td>KT323975/A3849076</td>
<td>636/639 (99%)</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Aspergillus ustus</td>
<td>Aspergillus ustus</td>
<td>KT323974/EU326214</td>
<td>590/595 (99%)</td>
<td></td>
</tr>
</tbody>
</table>

Rojas et al., 2016

### TABLE 3  Phenotypic and molecular identification data from actinomycetoma agents

<table>
<thead>
<tr>
<th>16S rDNA</th>
<th>GenBank accession number</th>
<th>Identity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Require different genes for identification
Understanding etiology: diversity within species

Ahmed et al, 2014
Novel diagnostics

Ideal molecular target for DNA-based detection:

- Specific for mycetoma agents (does not cross-react with other soil fungi)
- Shared by different species/genera (*Madurella mycetomatis* and *Trematosphaeria grisea*)
- Present in multiple copies to increase sensitivity
This approach worked well for another fungus
Genome Sequence of *Madurella mycetomatis* mm55, Isolated from a Human Mycetoma Case in Sudan

Sandra Smit, Martijn F. L. Derks, Sander Bervoets, Ahmed Fahal, Willem van Leeuwen, Alex van Belkum, Wendy W. J. van de Sande

DOI: 10.1128/genomeA.00418-16

36.7 Mbp genome
804 scaffolds (N50 of 81.8 kb; G+C content of 54.9%).
Collaboration between Mycetoma Research Center, Sudan and CDC

Whole Genome Sequencing of fungal agents of Mycetoma
Study objectives

• Generate chromosomal quality annotated genomic assemblies of *M. mycetomatis* and *T. grisea* using long-read sequencing --- to provide a resource for community

• Generate WGS phylogeny of *M. mycetomatis* using clinical isolates from Sudan --- to understand the genetic diversity among isolates

• Use metagenomics to characterize “grains” from mycetoma patients--- to understand what pathogens actually are present in patients
Study Samples

• Received from Prof. Fahal’s group:
  • 128 DNA from grains
  • 50 cultures of *M. mycetomatis*

• Two isolates (one *M. mycetomatis* and one *T. grisea*) from CDC collection
Preliminary PCR analysis of grain samples (ITS and 16S)

<table>
<thead>
<tr>
<th>Organism</th>
<th>no 16S amplification</th>
<th>Actinomadura sp.</th>
<th>Uncultured/unsequenced</th>
<th>S. pyogenes</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>M. mycetomatis</em></td>
<td>92</td>
<td>0</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td><em>M. fahalii</em></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Falciformispora thompkinsii</em></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Falciformispora senegalensis</em></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Cladosporium sp.</em></td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Curvularia sp.</em></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><em>Fusarium solani</em></td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>no ITS amplification</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>

Of 126, 88 passed DNA quality control for WGS and good quality reads were obtained – analysis pending
Cultures

- Of 50, 29 cultures grew
- 26 were sent for WGS
- 3 are slow growing

M. mycetomatis

T. grisea
Preliminary WGS results

Madurellamyctomatis and Trematosphaeria grisea

Isolates from Sudan
CDC Collection
NCBI Reference
Preliminary WGS results

*M. mycetomatis* only

- ▲ Isolates from Sudan
- ◆ CDC Collection
- ■ NCBI Reference

GCA 001275765.2 ASM127576v2 genomic (Reference)
Next steps

- PacBio sequencing of 5 isolates, *T. grisea* and four *M. mycetomatis*
- Long-read assembly and annotation
- WGS phylogeny of *M. mycetomatis*
- Identification of potential PCR targets
- Collaboration of developing molecular tests
- WGS of isolates from other regions and other genera?
Acknowledgments

Mycetoma Research Center, Sudan
Prof. Ahmed Fahal
Sahar Bakhiet

Mycotic Diseases Branch, CDC
Lalitha Gade
Steven Hurst
Karlyn Beer
Tom Chiller
For more information, contact CDC
1-800-CDC-INFO (232-4636)

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.