

Bacillus anthracis comparative genome analysis in support of the amerithrax investigation

Dreycey Albin (da39)

February 19, 2019

Scientific question

The scientific question and goal for this paper focuses on *the ability for microbiological analysis, whole genome sequencing, and bioinformatic analysis to aid in microbial forensics*. The research dives further into the details of how different strains involved in the 2001 Amerithrax attacks were characterized. Four different morphologically distinct strains were identified using samples obtained from the letters sent to the NY Post, Senator Leahy, and Senator Daschle. These strains were designated A, B, C/D, and E.

Hypothesis tested

The hypothesis tested in this research is that *forensics can benefit from a combination of microbiological testing, whole-genome sequencing, and comparative genomic analysis*. They convincingly justify this hypothesis by first retrieving information about the high-level morphological differences between the different strains, then by outlining the more detailed genomic characteristics that may yield the morphological results observed.

Methods/Results

There are two major ways that the researchers used to split the methods: (1) Microbiological Analysis; and (2) Whole-genome sequencing and comparative genomics.

1. **Microbiological Analysis** was used to characterize the observable characteristics of the different strains isolated from each of the envelopes. When plated out onto sheep blood agar (SBA), the colonies for the four different morphotypes were yellow in color, as compared to the normal gray-white, and exhibited spreading (Note that morphotype E had small colonies). In addition, congo red (A dye used in pH indication and biological staining) binding indicated a lower ability for these morphotypes to sporulate.
2. **Whole-genome sequencing and comparative genomics** was used to identify the genomic differences between the morphotypes and the Ames strain. In morphotype A, it was seen that there was a repeat region containing an unknown protein sequence and a 16S rRNA gene. This was first seen on the comparative bioinformatics analysis, and was thereafter validated using PCR. There was also a noticeable SNP found within the TrkA family potassium-uptake protein. Morphotypes B, C/D, and E all had genetic modification in areas that affected sporulation. Specifically, B had a SNP occurring between SpoOF and a hypothetical protein. Morphotypes C and D both had modifications that made an effect on a sensor histidine kinase, the difference being C had a SNP causing a stop codon to arrive prematurely, and D had a deletion in a large portion of the gene. This histidine Kinase is thought to play a role in the phosphorylation of SpoOF. Morphotype E contained an in-frame deletion, disrupting the response regulator involved in the dephosphorylation of SpoOF (i.e. a phosphatase).

Key implications of the results

The key implication from this research is the ability for a team to use a combination of techniques to assist in the analysis of biological sample for microbial forensics. They show this by identifying morphological differences between isolated *B. Anthracis* from the letters of the Amerithrax investigation. They then expand upon these morphological differences by identifying exact genomic explanations using whole-genome sequencing and comparative genomics.