Project Title: *Salmonella enterica* **Strain Typing Using Next-Generation Sequencing Technology –**
Intrepid Bioinformatics Bacterial Typing Portal

Presenter: **Ted Kalbfleisch**, James Bono, Rong Wang, Tim Smith, Gregory Harhay,
Tommy Wheeler and Dayna Harhay
Presenters email address: Ted.Kalbfleisch@intrepidbio.com
Mailing address: Intrepid Bioinformatics
                 c/o MetaCyte Business Labs
                 301 East Market Street
                 Louisville, KY 40202

Category: Post-harvest research

Published: Unpublished to date

Objective: To create a web accessible *Salmonella enterica* genome data resource for assessing the relatedness of any given *Salmonella* strain via comparison with reference *Salmonella* genome sequences present in our database. The resultant platform allows whole genome next-generation sequence (NGS) data to be uploaded and mapped to extant reference genome sequences, and facilitates subsequent discovery of single nucleotide polymorphisms (SNPs) which may be used for strain identification.

Experimental Design & Analysis:

A comprehensive genome sequence comparison was conducted of 48 *Salmonella enterica* isolates (12 each of serotypes Anatum, Montevideo, Typhimurium and Newport) that had been fully sequenced, assembled and uploaded to the Intrepid Bioinformatics Bacterial Typing Portal.

Using data from the 48 strains sequenced by Dr. Harhay, we performed 48x48 comparisons to create a comprehensive catalog of variants (single nucleotide variant, and small insertions and deletions) that describe exactly how each strain varies from all the others. When an unknown strain is submitted, it will be mapped to each of the 48 strains, and it is possible to unambiguously identify to which of these strains the unknown strain is most closely related. Additional software is being developed currently that will produce the nucleotide level information required to design molecular assays that will differentiate either a submitted strain (known, or novel) from any other strain, or group of strains known to us. The web front end to this system is still under development, and we hope to be able to make it available for use to the general public in the near future.

Key Results:

Here, we present a web based system that will allow both researchers and producers to quickly, inexpensively, and unequivocally type strains of Salmonella enterica that are relevant to their work.

Industry Application:

Once available, producers, and researchers will be able to create secure accounts on our system, and, with drag and drop ease, upload their data for analysis. Within hours, they will receive a phylogenetic tree that will place their submitted strain with the know strain to which it is most closely related. They will also be able to use our algorithms to identify variants that can be used to distinguish their strain of interest from any other known to our system.
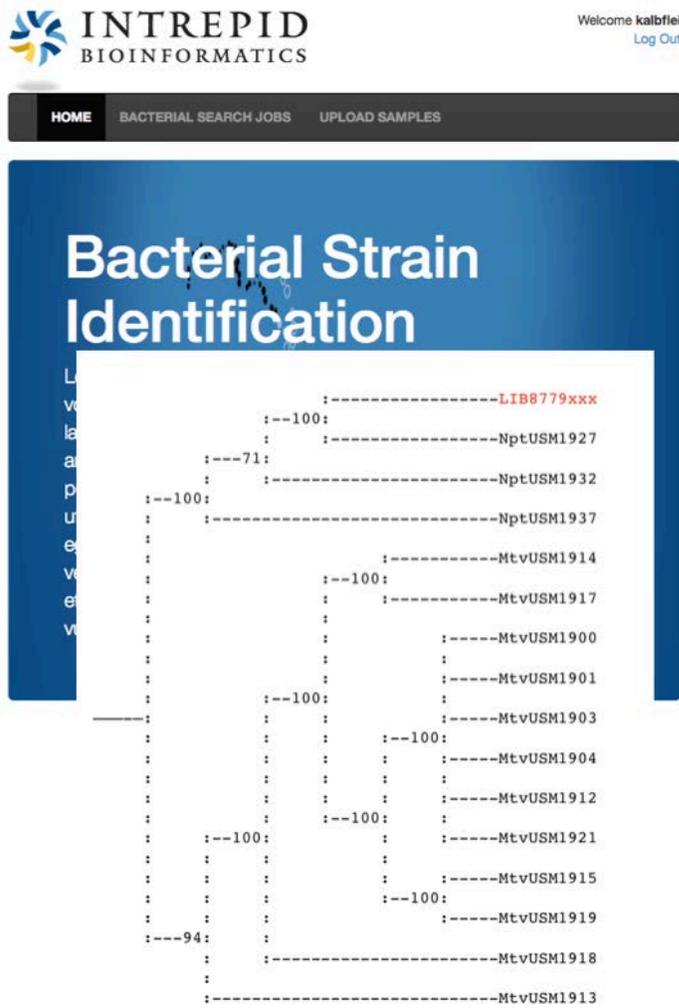
Figure 1: A screen shot of our web portal as well as an example of output for the submission of whole genome shotgun sequence. This example is from a library created with a Salmonella enterica strain closely related to Newport. Our process placed phylogenetically with its nearest Newport neighbor.