

# Characterization of Invasive Streptococci Isolated in a Nosocomial Outbreak with Illumina and Nanopore Whole Genome Sequencing

Mark M. Li <sup>1,2</sup>, Xiaojun Liu <sup>1</sup>, Peter Hu <sup>1</sup>, and Awdhesh Kalia <sup>1</sup>

<sup>1</sup> Program in Diagnostic Genetics & Genomics, School of Health Professions, UT MD Anderson Cancer Center

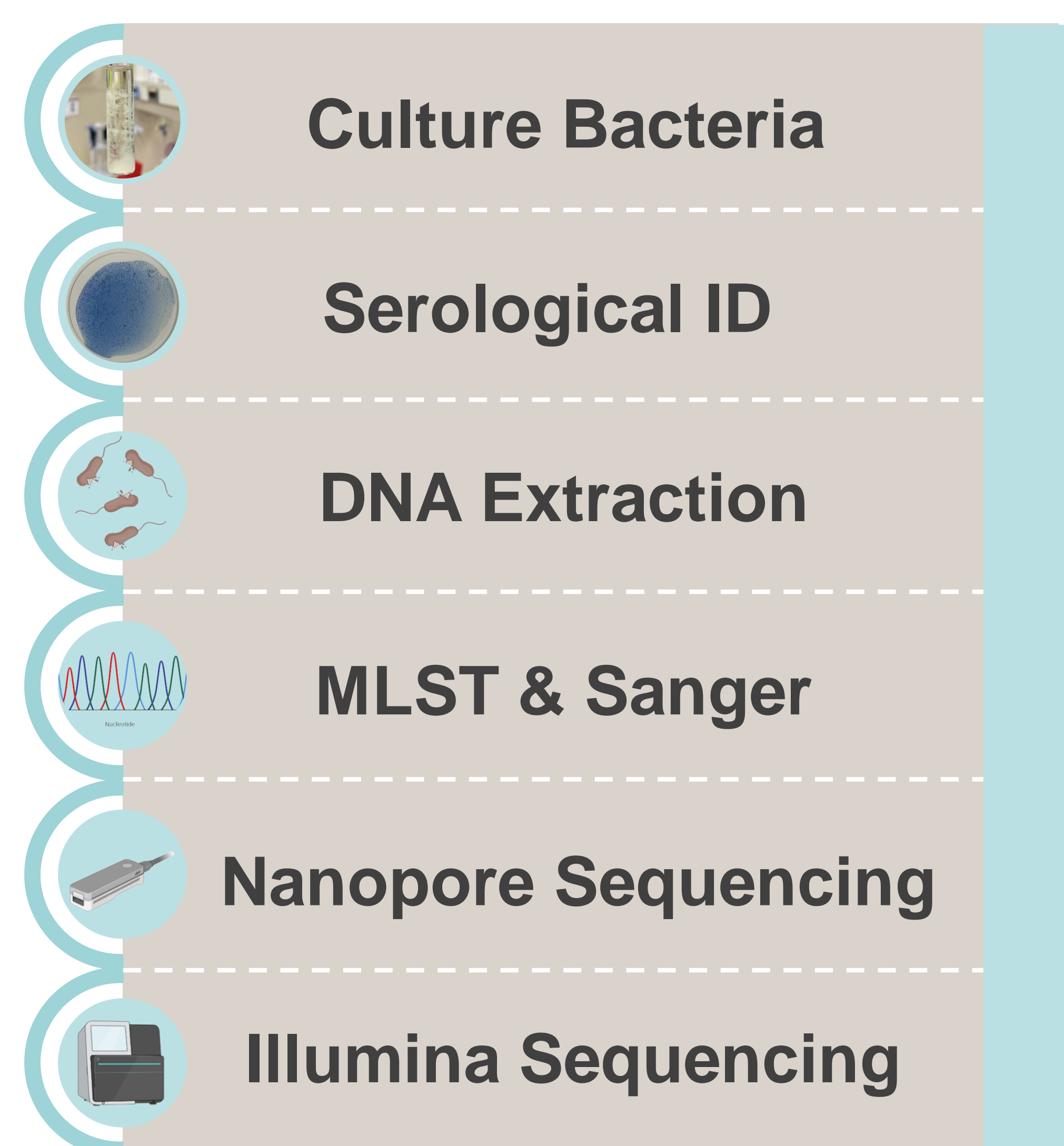
<sup>2</sup> University of Texas, Austin

## Introduction

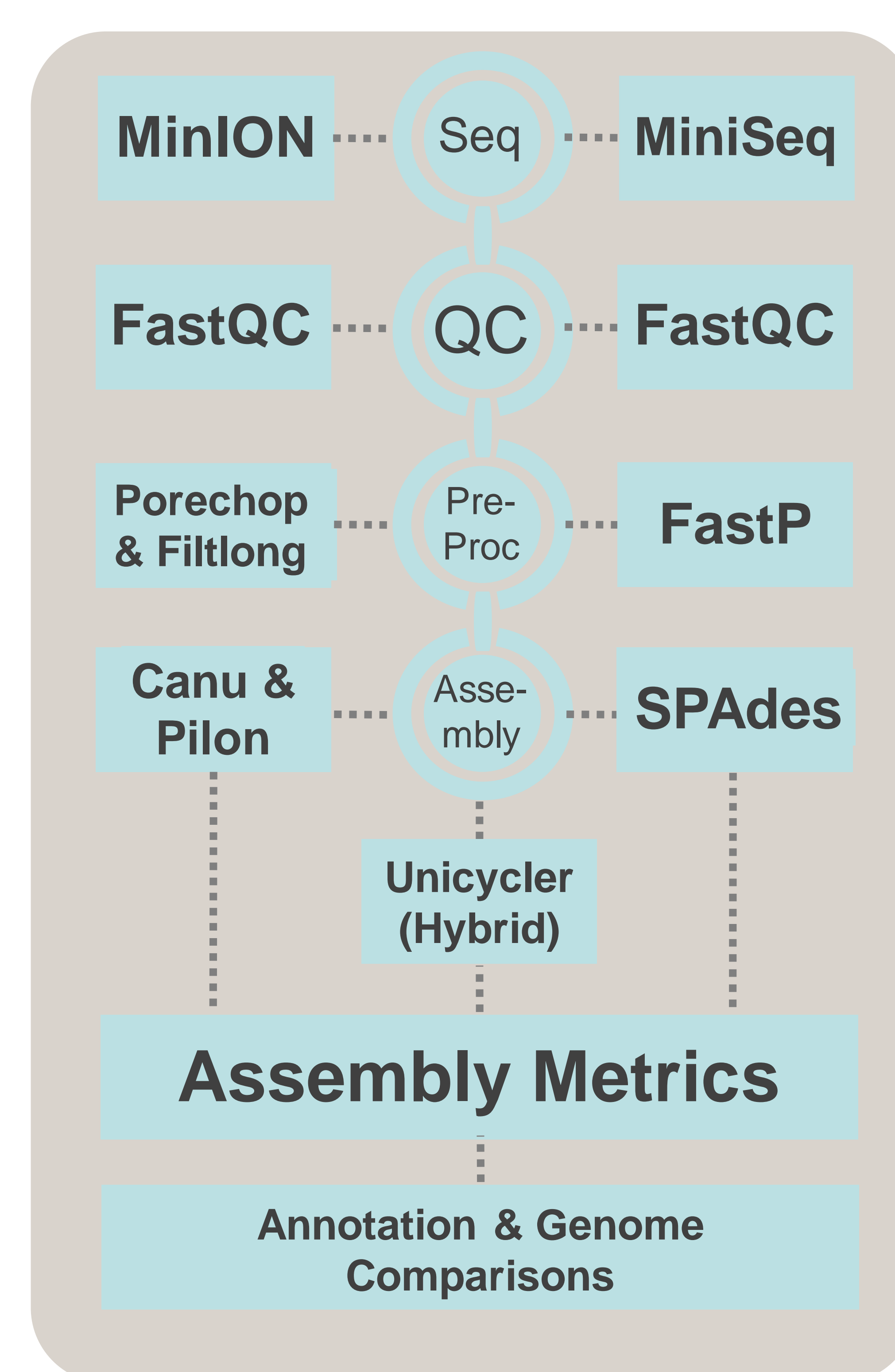
Nosocomial outbreaks caused by MDR pathogens threaten patient safety and treatment management. Genomics is providing unprecedented insights into pathogen evolution; however, difficulties in producing high quality complete genomes limits the information that can be gained.

## Methods

### Overall Workflow

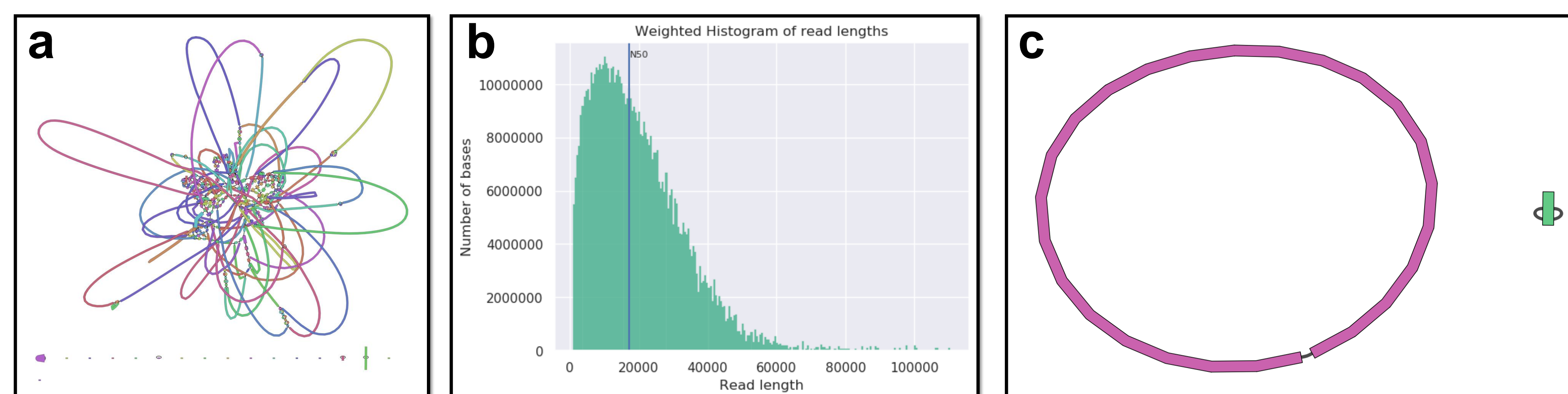


### Bioinformatics Workflow



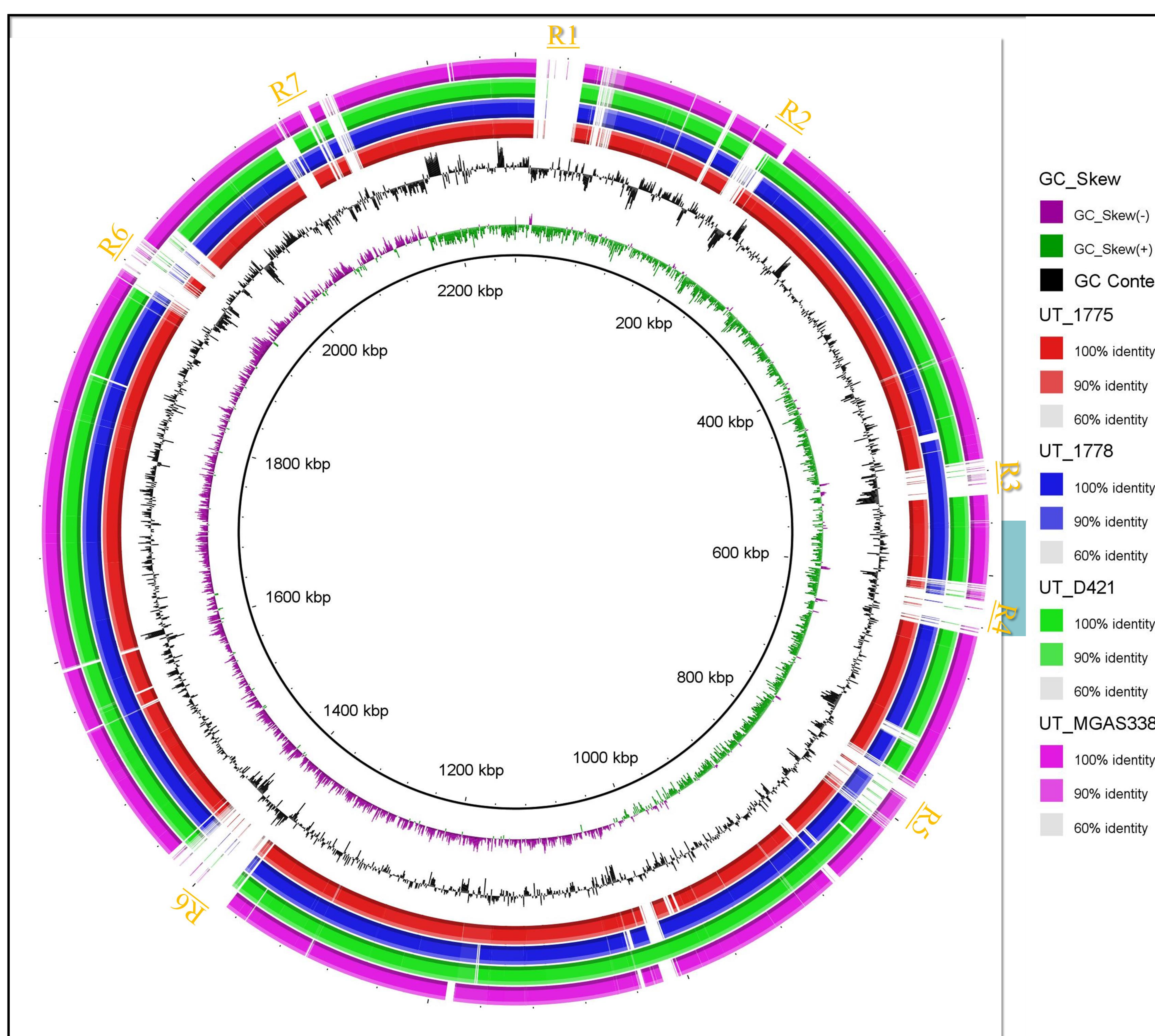
## Major Findings

1. Joint use of Illumina short-read and Nanopore long-read sequencing data produces high quality complete genome assemblies.



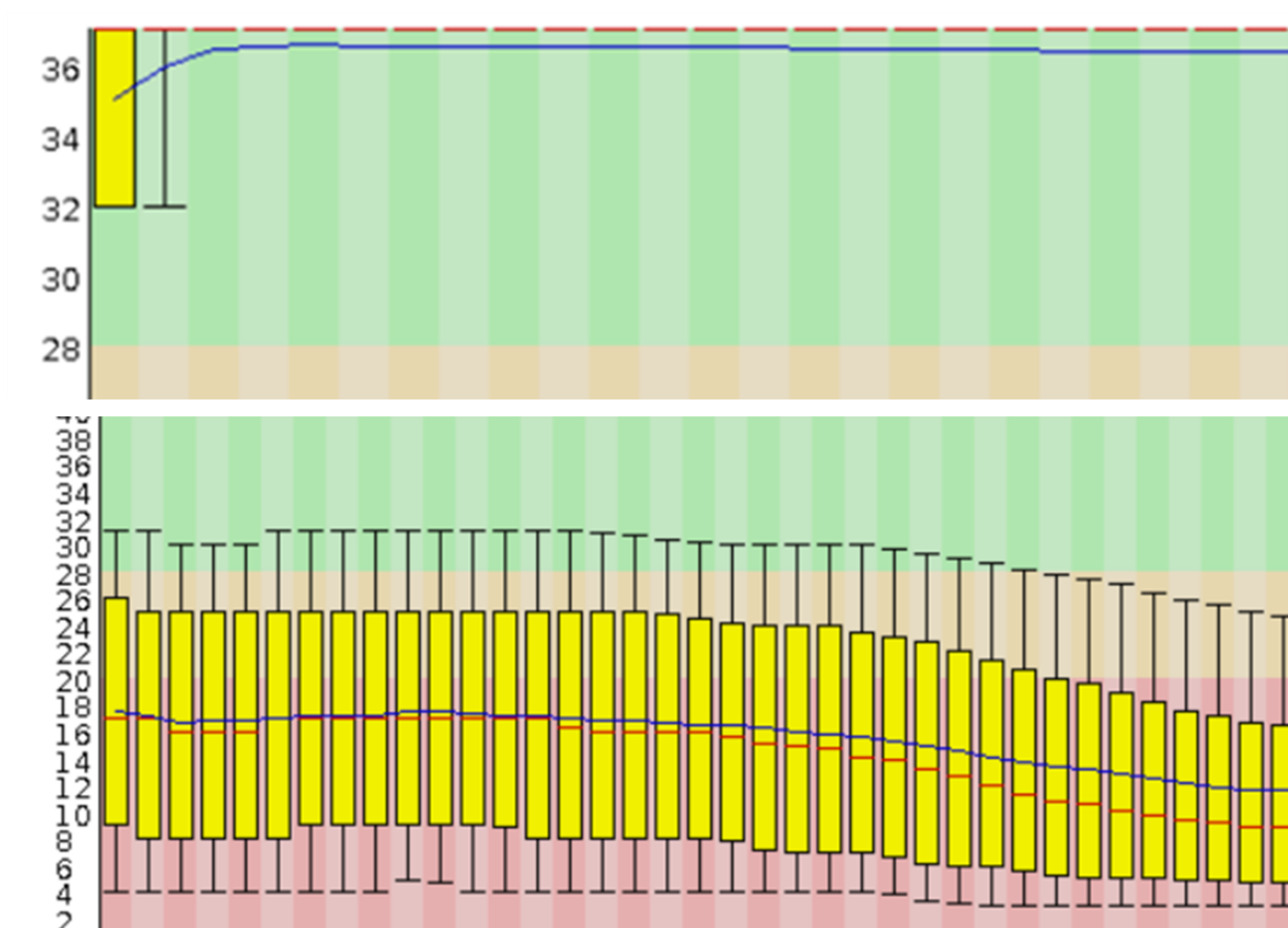
**Figure 1.** (a) deBruijn graph indicating fragmented genome assembly resulting from SPAdes genome assembly from Illumina reads; b) Weighted histogram of Nanopore reads indicating N50 of 17 Kb (c) deBruijn graph indicating a completely resolved circular bacterial chromosome along with an accessory plasmid.

2. All four isolates involved in the outbreak are genomically diverse yet harbor identical determinants of antimicrobial resistance to tetracycline, fluoroquinolones, and macrolides.



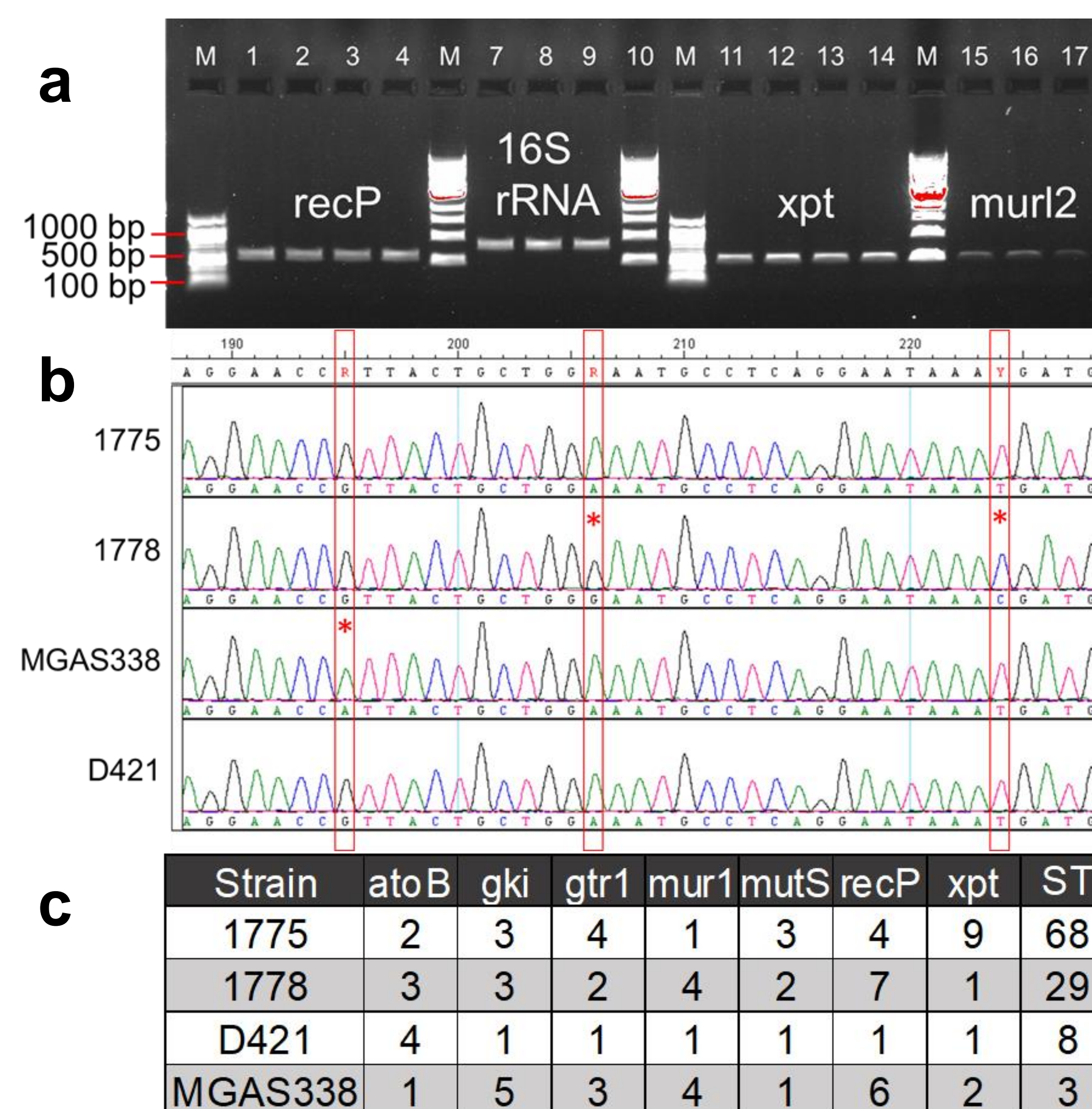
**Figure 2.** All four isolates were compared to the reference SDSE strain FDAARGOS\_1016 using blastn. Concentric circles from inside to outside show GC-skew, GC content and similarities or divergence among the outbreak strains relative to the reference genome. R1 to R7 indicate regions of plasticity wherein isolates show divergent sequence content (e.g., phage insertions, pathogenicity islands, etc.)

## Long-Read Sequencing Sacrifices Base Quality Score



**Figure 3.** Comparison of quality scores across bases between Illumina sequences (top) and Nanopore sequences (bottom).

## Multi-Locus Sequence Typing Identified Four Distinct Genotypes



**Figure 4.** (a) Agarose gel electrophoresis showing successful isolation of HK genes; (b) Sequenced MLST products in a chromatogram showing SNPs between strains; (c) Strains have different sequence types.

## Acknowledgements

This study was funded by the Reece Summer Research Program, School of Health Professions. Mark Li was supported by the CPRIT CURE Summer Undergraduate Research Program. We thank Vijay Varghese and Dr. Hongli Tang for their continued support and Daniel Domogala and Yasmine Hoballah for their help with bioinformatics analysis and poster preparation.