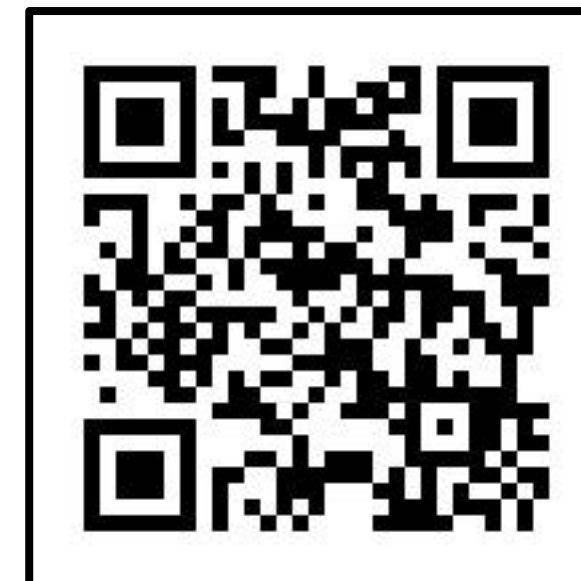


GENOMICS AND METAGENOMICS TECHNIQUES FOR CHARACTERIZING BACTERIA STRAINS AND COMMUNITIES



Emily Longo '21, Minjun Choung '21 and Funmilola A. Ayeni; Biology Department

INTRODUCTION

The global importance of microbes cannot be overemphasized in health and disease, thus, there is recent emergence of new techniques for rapid and in-depth investigation of these organisms. Whole genome sequencing, genome assembly, and downstream analysis can be used to genomically characterize the pathogenicity of individual microbes, while the taxonomic and functional profile of entire microbial communities can be explored using the metagenomics QIIME2 pipeline. The aim of the current study is to learn these two techniques for future application to novel studies.

METHODOLOGY

1. GENOMICS STUDY

- Whole genome assembly and analysis were learned and applied through a case study of *Salmonella*.
- 138 fresh bovine fecal samples were collected from healthy Sokoto Gudali cattle.
- Genus-specific PCR amplification of the *Salmonella invA* gene identified eighteen bacterial isolates as *Salmonella* spp..
- The isolates were whole-genome sequenced using Illumina MiSeq (read length 2x150, depth 50s)
- All reads were assembled *de novo* and subjected to downstream analysis (Figure 1).

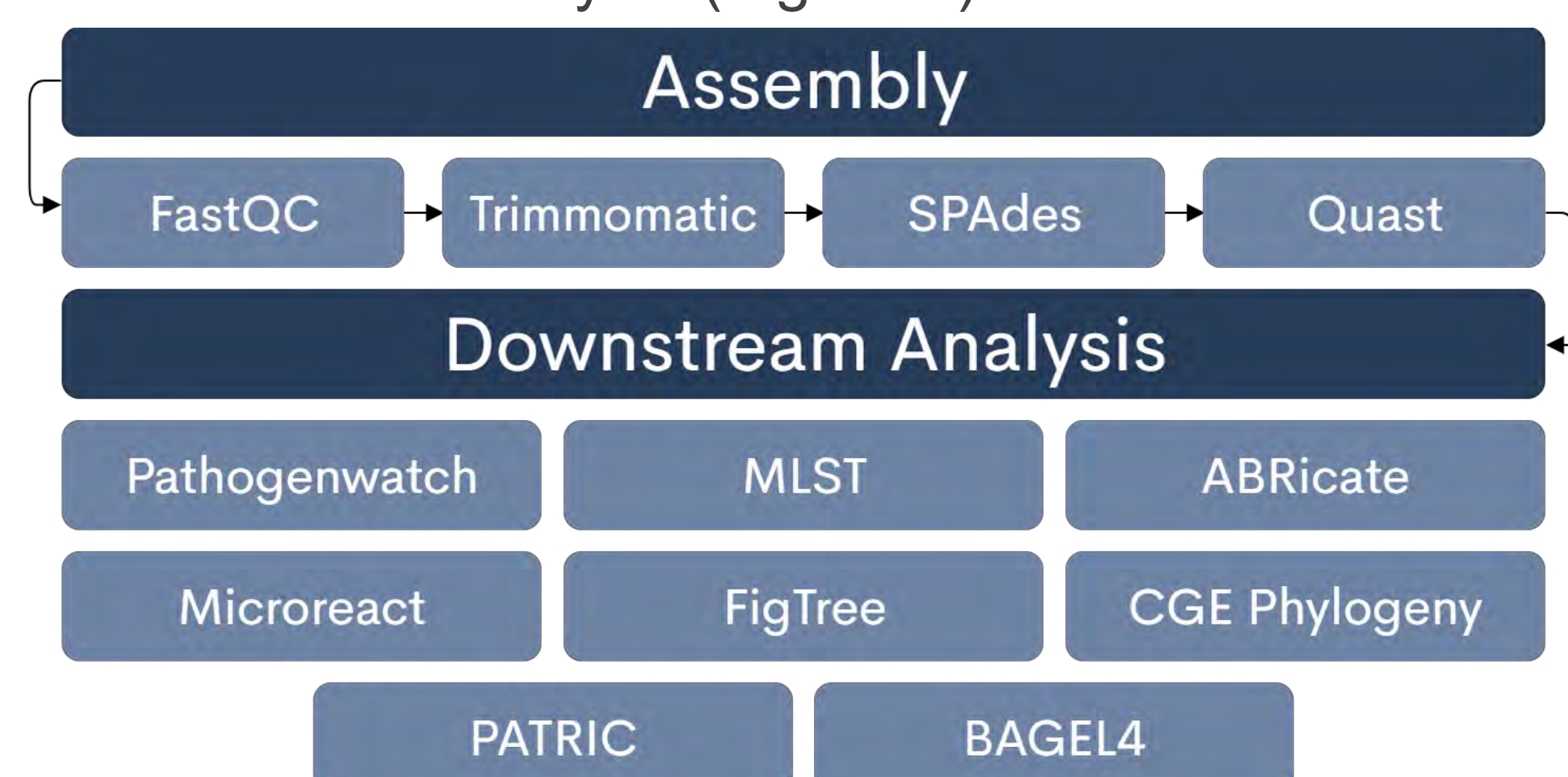


Figure 1: Workflow for bacterial genome assembly and downstream analysis.

2. METAGENOMICS STUDY

- The metagenomics QIIME2 pipeline (v. 2020.6) was applied to a subset of the raw data from Afolayan et al. (2020)
- 42 fecal samples were collected from 22 healthy and 20 diabetic elderly individuals.
- DNA extraction and sequencing of the total microbial community was performed, and the resulting reads were imported into QIIME2 to be analyzed (Figure 2).

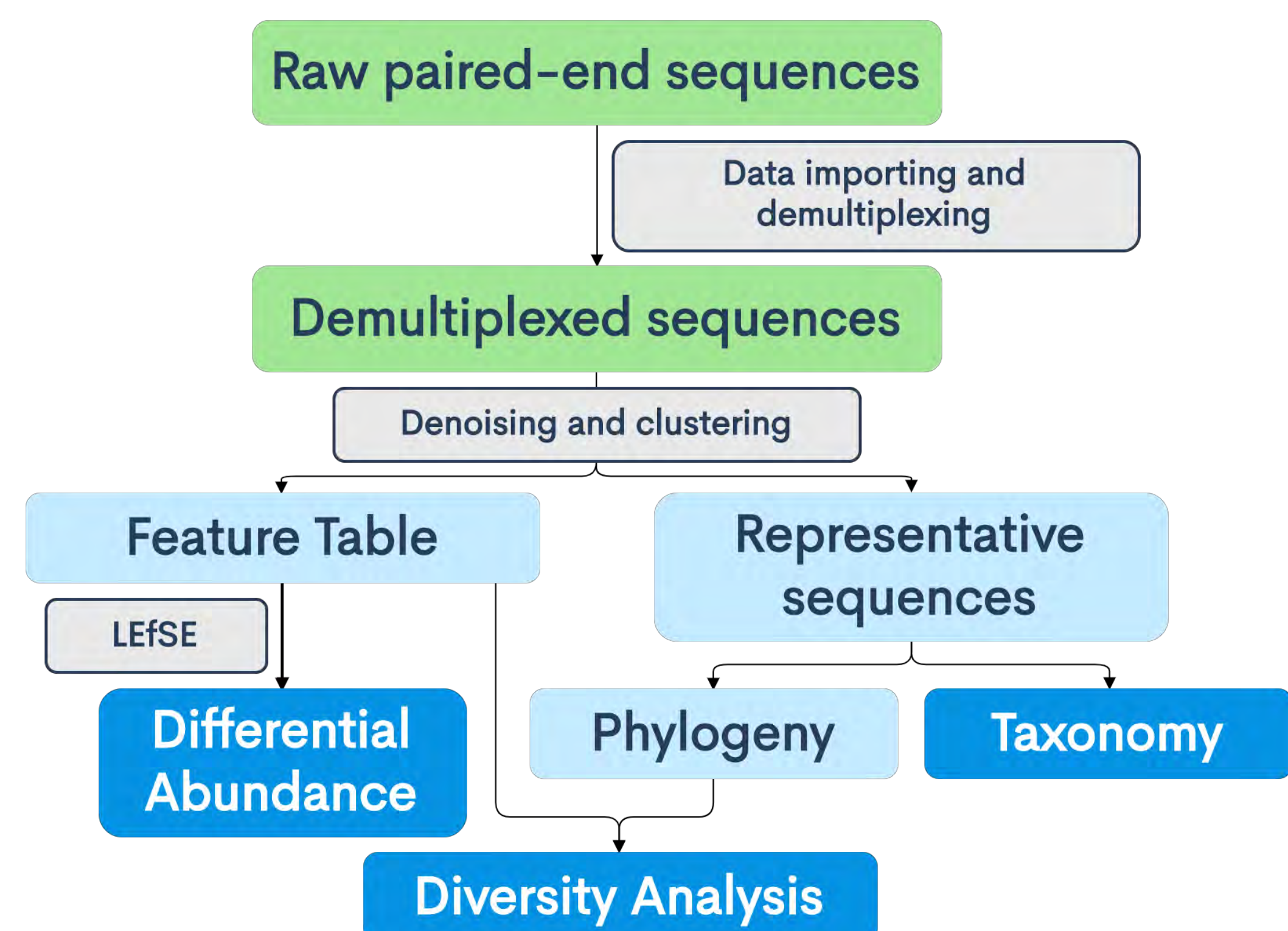


Figure 2: Workflow for the metagenomics QIIME2 pipeline.

RESULTS

1. GENOMICS STUDY

- All eighteen *Salmonella* strains were of the species *Salmonella enterica* ssp. *enterica*
- Five different sequence types and corresponding serovars were identified (Figure 3).
- 15 strains only contained one antimicrobial resistance gene (*aac(6')*-*laa*).
- Additional AMR genes were identified in ST519 and ST3961 (Table 1).
- There was an absence of virulence genes.

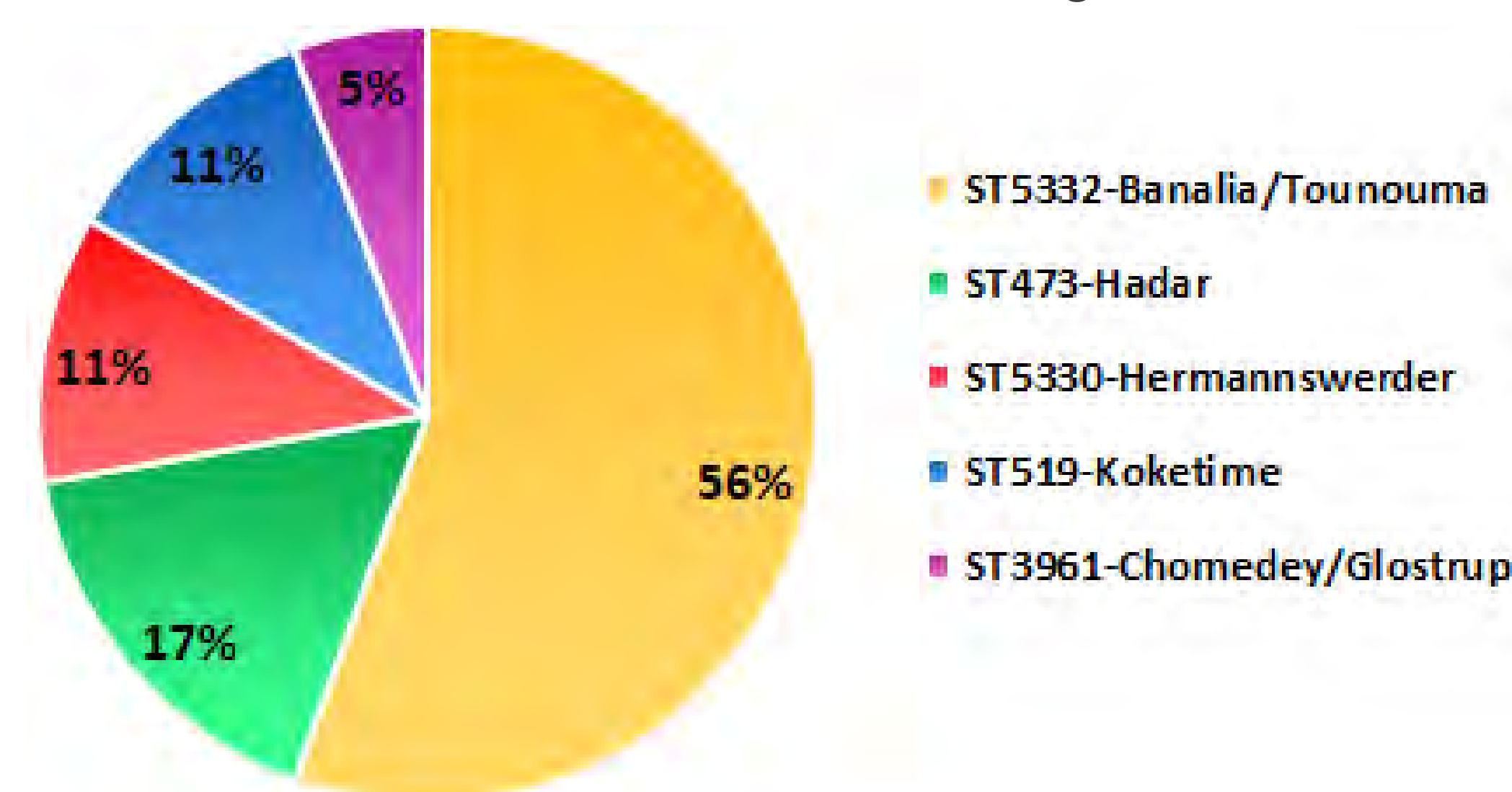


Figure 3: Sequence types and corresponding serovars of *S. enterica*.

| <i>aac(6')</i> | <i>fosA7</i> | <i>tet(A)</i> | <i>aph(6)</i> | <i>aph(3'')</i> | <i>sul2</i> | <i>qnrB19</i> |
|----------------|--------------|---------------|---------------|-----------------|-------------|---------------|
| X | | | | | | |
| X | | | | | | |
| X | | | | | | |
| X | X | | | | | |
| X | | X | X | X | X | X |

Table 1: Antimicrobial resistance (AMR) genes found in *S. enterica* sequence types, colored to match the legend in Figure 2. The presence of a genes is indicated by an X.

2. METAGENOMICS STUDY

- There were significant differences in alpha diversity ($p=0.03$) and beta diversity ($p=0.012$) between the gut microbiota of healthy and diabetic individuals
- 14 different phyla were observed in the gut microbial communities, with Firmicutes being most dominant (~72%) (Figure 4).
- 18 taxonomic biomarkers are associated with healthy individuals, while 8 are linked to diabetic individuals (Figure 5).

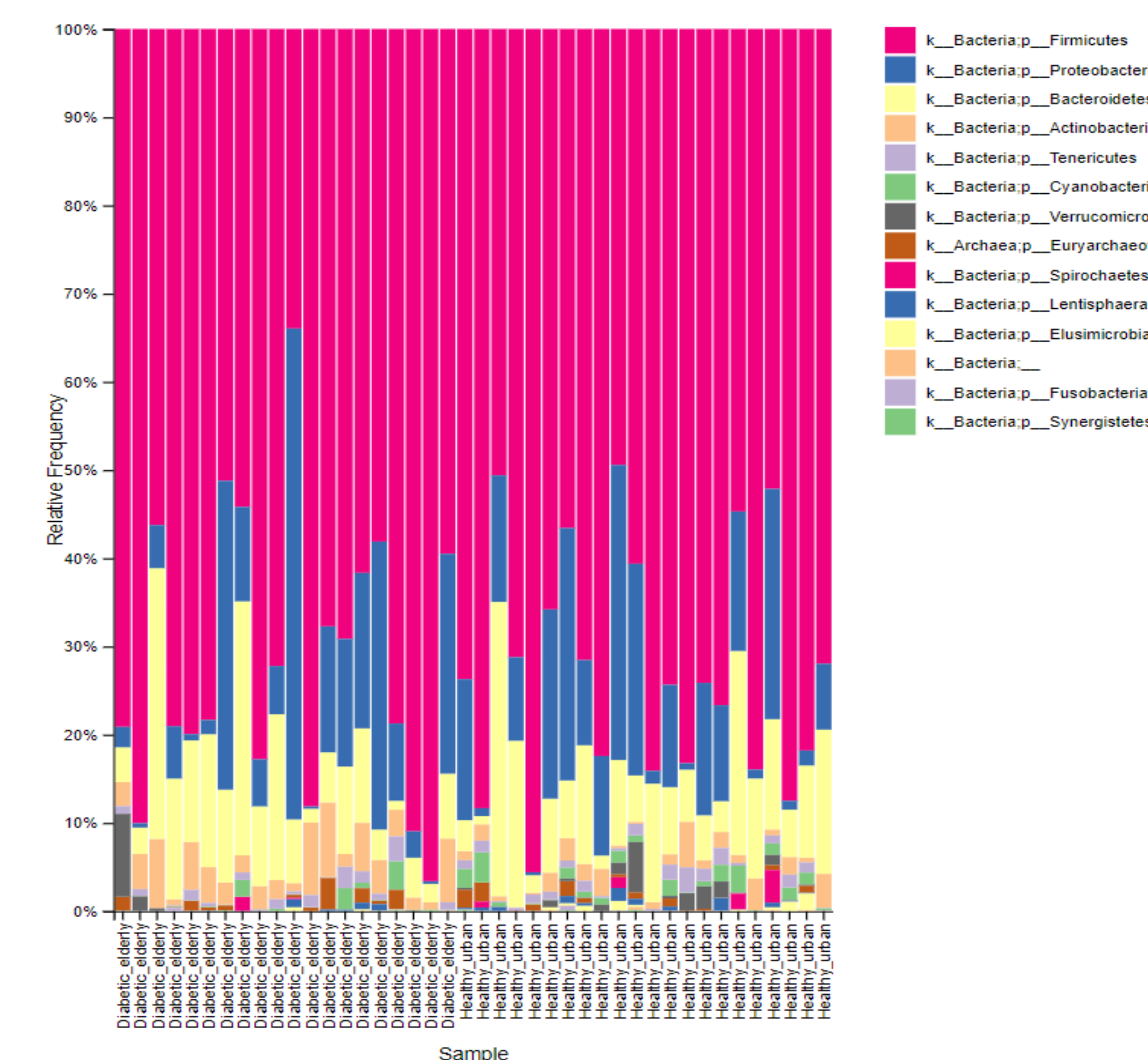


Figure 4: Barplot of taxonomic variation of the gut microbial communities of healthy and diabetic individuals at the phylum level.

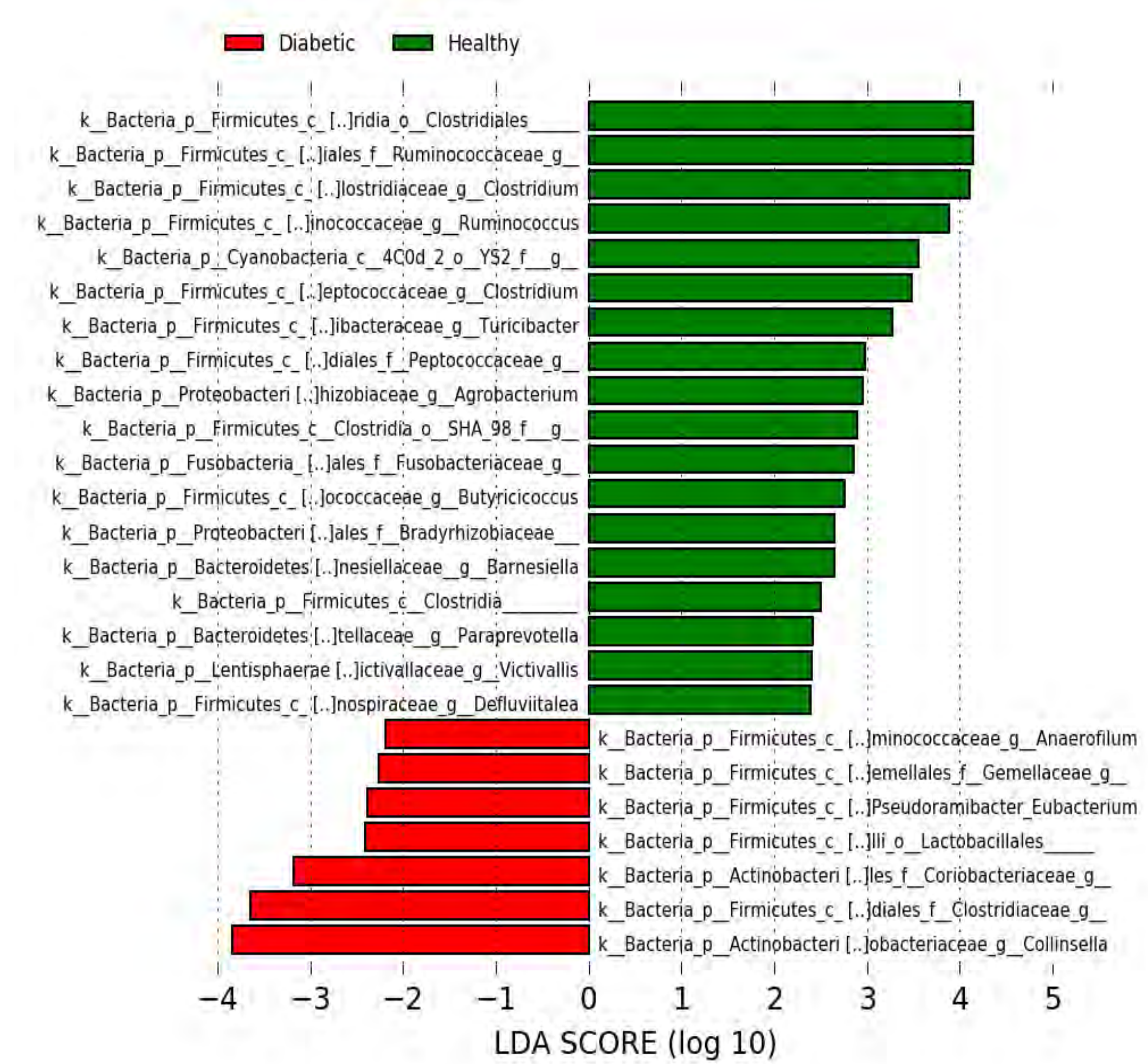


Figure 5: Differential abundance of discriminant features based on microbial taxa in healthy (green) and diabetic (red) individuals.

DISCUSSION

The results of the *Salmonella* case study suggest that:

- There were more *Salmonella* strains than expected in the cattle population.
- The *Salmonella* strains found in the gut microbiomes of the cattle population are relatively benign, which indicates minimal antibiotic exposure and healthy living conditions.

The conclusions from the diabetes case study can be summarized in the following:

- There are many bacterial taxa common to all human gut microbiota.
- Proportions of different microbes within gut microbial communities differ significantly between healthy individuals and diabetic individuals.

The genomic and metagenomic techniques learned in this study are extremely applicable to future research in characterizing infectious microorganisms, as well as determining the relationship between health and the human gut microbiome. Future studies that will benefit from this training include:

- The whole genome analysis of select *Lactobacillus* species to explore the genetic basis for their reported antimicrobial activity.
- A metagenomic analysis of the effect of *Capsicum spp.* consumption on the human gut microbiota.

ACKNOWLEDGEMENTS + REFERENCES

Special thanks to Ayorinde Afolayan, our bioinformatics advisor, as well as Matt Tarantino and Asprey CCAS for granting us access to Vassar's Hopper computer cluster. Also thank you to all of the URSI faculty and the contributing donors for making this research opportunity possible.

1. Afolayan, A. O., Adebuseye, L. A., Cadmus, E. O., & Ayeni, F. A. (2020). Insights into the gut microbiota of nigerian elderly with type 2 diabetes and non-diabetic elderly persons. *Heliyon*, 6(5).
2. Australian Galaxy Team. (2020). Galaxy Australia. Retrieved from <https://usegalaxy.org.au/>
3. QIIME2 Development Team. (2020). "Moving pictures" tutorial. Retrieved from <https://docs.qiime2.org/2020.8/tutorials/moving-pictures/>

