Salmonella typing directly from stool using long read metagenomics

@MetaRudder
@LangridgeLab

ICASE partner

Steven Rudder[1*], Bilal Djeghout[1], Georgios Patsos[2], Ngozi Elumogo[1,3], Nicol Janecko[1], and Gemma C. Langridge[1]

[1]Quadram Institute Bioscience, Norwich, UK, [2]RevoluGen, Hadfield, Glossop, UK, [3]Eastern Pathogen Alliance, Norwich, UK

The aim of this work was to test the bacterial typing and AMR gene identification capability of our HMW-DNA extraction pipeline in combination with long read sequencing

Background

- Culture-based methods are laborious with inherent biases towards specific species or subtypes and often generate results after an outbreak has become established [1]
- Although routine sequencing is conducted for salmonellosis cases, for other enteric pathogens rapid culture-independent testing (CIT) is the predominant choice for diagnostic laboratories, yielding no isolates for sequencing [2]
- As laboratories continue to adopt CIT-only testing strategies, a gap in pathogen genomic epidemiology is created.
- One potential solution is sequence-based metagenomic approaches to identify bacterial pathogens during an outbreak [3]
- To aid sequence-based long read metagenomic approaches improved high molecular weight (HMW) DNA extraction processes are needed

Methods

- A HMW DNA extraction method for stool is in development using RevoluGen's automated Fire Monkey HMW DNA extraction kit in multi-well filter plate format
- We have achieved recovery of DNA from clinical stools samples with an average length of 19kb which includes fragments of DNA >150kb (Fig. 1)
- 200ng of DNA was run on a Nanopore MinION sequencer using the LSK-109 library preparation kit

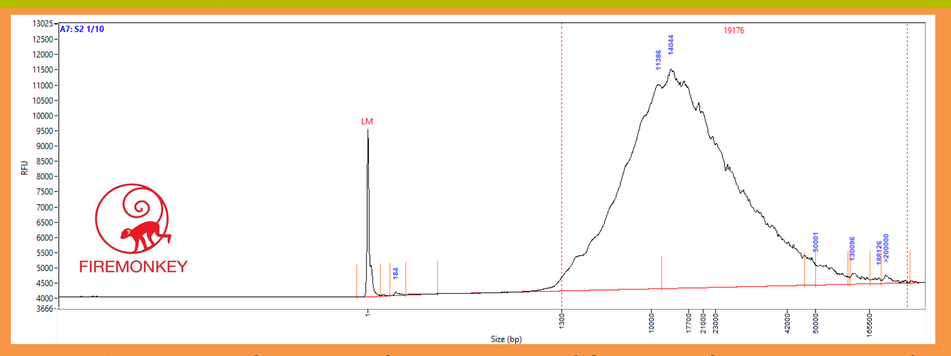

Figure 1: Femto Pulse trace of DNA extracted from stool using Fire Monkey

Results: Metagenome classification and *Salmonella* identification

- After base calling and QC, the yield was 884,500 reads with an N50 of 9,273bp (longest read was 154,201bp)
- Flye was used on meta mode to assemble contigs, N50 of contigs was 145,133bp (Table 1)
- Kraken2 was run on Flye assembled contigs and the data was visualised using Krona (Fig. 2)
- 2.14% of the overall sequence data was predicted to be *Salmonella enterica*, with 0.05% of the sequence data predicted to be *Salmonella enterica* subsp. *salamae*
- In addition to the identification of *Salmonella enterica* subsp. *salamae* a snapshot of the patients microbiome was captured. A large proportion of the data (71%) had no hits when using Kraken2. *Faecalibacterium prausnitzii* was identified as the most abundant bacterial species in the sample (present at 5.16% of overall sequencing)

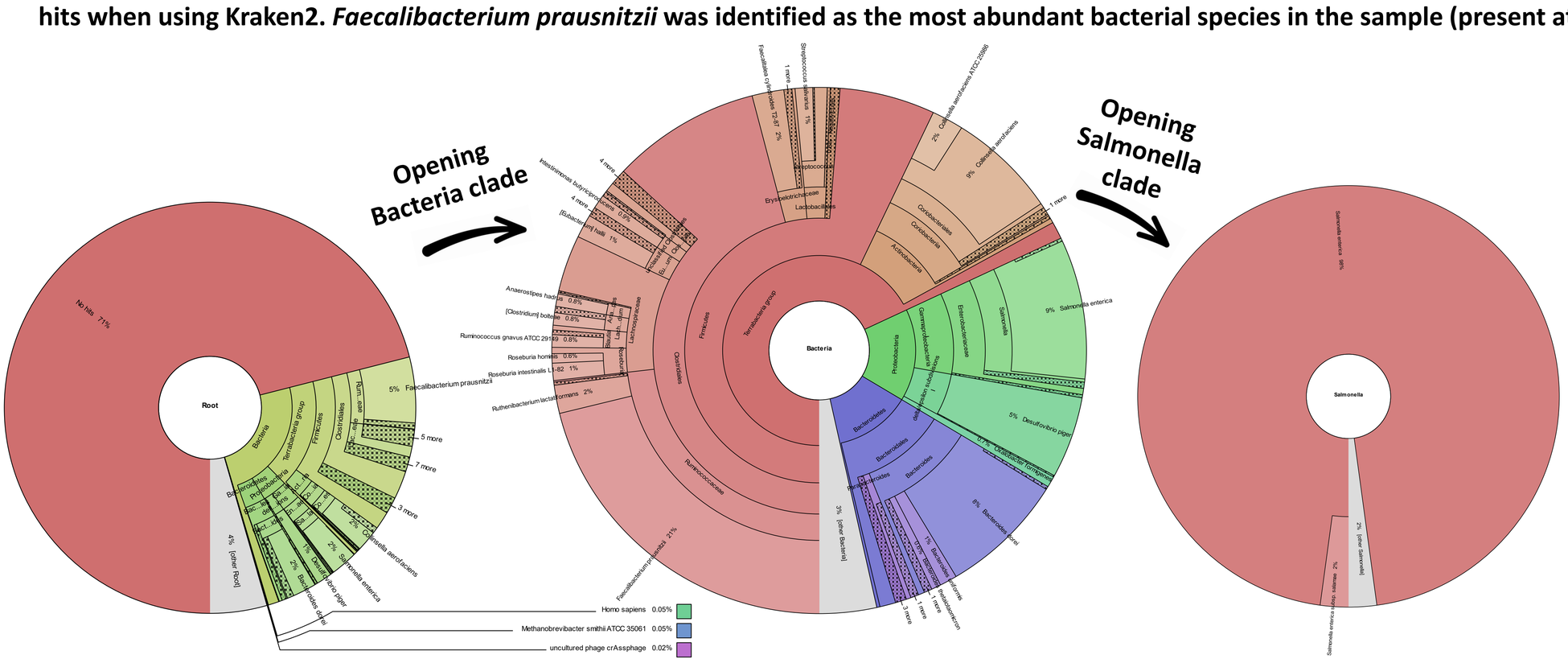

Figure 2: Krona visualisation of a Kraken2 report generated using contigs assembled by Flye

Table 1. displaying the number and size of contigs assembled by Flye

Contig size (bp)	Contigs (*n*)
2-4 Mbp	6
1-2 Mbp	15
500-999 Kbp	25
250-499 Kbp	74
100-249 Kbp	302
50-99 Kbp	453
25-49 Kbp	759
10-24 Kbp	1103
Below 10 Kbp	1525
Total	4262

Results: Comparing a pure culture isolate to a metagenome derived genome (MAG)

- From the same stool sample a single colony was isolated and sequenced with Nanopore and Illumina technologies to create a hybrid consensus sequence
- This pure culture isolate was used as a benchmark to compare the typing results from the metagenome data (Table 2 & 3)
- Metabat2 was used to bin the Flye assembled metagenome contigs. CheckM was used to identify the bin containing *Salmonella* reads

Table 2: *Salmonella In Silico* Typing Resource (SISTR) results

	Pure Culture Isolate	Metagenome derived genome
Subspecies	*Salamae*	*Salamae*
Serovar	II 6,7:g,[m],s,t:[z42]|II 6,7:m,t	II 6,7:g,[m],s,t:[z42]|II 6,7:m,t
Serogroup	C1	C1
O-antigen	6,7	6,7

Table 3: Resfinder results

	Pure Culture Isolate	Metagenome derived genome
AMR gene 1	*aac(6')-Iaa_1*	*aac(6')-Iaa_1*
AMR gene 2	*mdf(A)_1*	*mdf(A)_1*
Resistance profile	sensitive	sensitive

Conclusions

- We are developing a method to isolate HMW DNA from stool utilising RevoluGen's automated Fire Monkey HMW DNA extraction kit
- Using long read metagenomic sequencing we were able to identify the subspecies, serovar, and AMR profile of the targeted pathogen in line with results from a pure culture
- Our pipeline is showing promising signs that high-throughput stool metagenomics could offer a data rich alternative to CIT-only testing strategies in the near future

Thank you for the travel grant

References

1. Forbes, Jessica D., et al. "Metagenomics: the next culture-independent game changer." *Frontiers in microbiology* 8 (2017): 1069.
2. UK Standards for Microbiology Investigations: Gastroenteritis. (https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/930517/S_7i2_FINAL-UKSMI.pdf)
3. EFSA Panel on Biological Hazards (EFSA BIOHAZ Panel), et al. "Whole genome sequencing and metagenomics for outbreak investigation, source attribution and risk assessment of food-borne microorganisms." EFSA Journal 17.12 (2019): e05898.