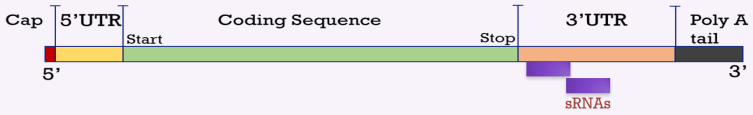
Using MinION sequencing to identify transcript variation between different life cycle stages of the parasitic nematode *Strongyloides ratti.*

Click for cool parasitic *S. ratti*

Dominika Lastik, Vicky Hunt
Department of Biology & Biochemistry, University of Bath
email: dl870@bath.ac.uk @DominikaLastik

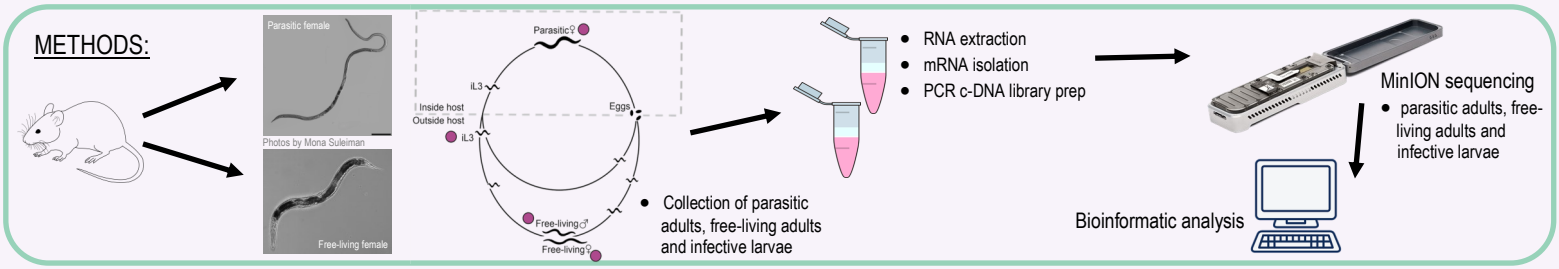
AIM: To identify full-length transcripts in *S. ratti* for study of alternatively spliced transcript and UTR variation

BACKGROUND:



- Our knowledge of parasitic nematode transcripts relies on short read sequencing, often missing information on poly A tails, UTR regions and alternatively spliced transcripts.
- This information is essential for studying gene expression and regulation and has to be estimated using bioinformatics, based on other organisms.
- Nanopore sequencing allows us to sequence full-length mRNA transcripts, including the 5' and 3' UTRs, which are important in gene regulation
- *Strongyloides ratti* have genetically identical parasitic and free-living generations; in which we know genes are differentially expressed and regulated by small RNAs - what we don't know is how?
- The study of transcripts in different life cycle stages will allow us to answer two main questions i) are transcripts spliced differently in different life cycle stages? and ii) are they important in gene regulation and parasitism?

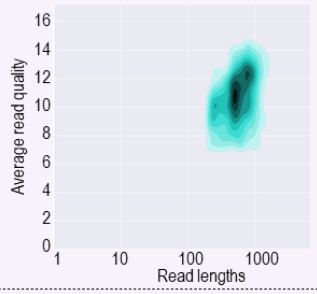
METHODS:



RESULTS & ANALYSIS:

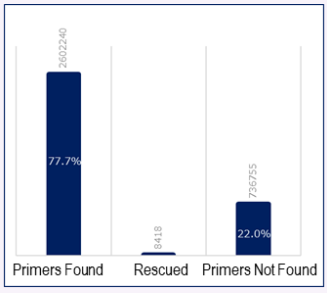
Nanopore produced up to 4M high quality reads per sample: Reads from all life cycle stages were combined and assessed for quality and size using Nanoplot. N50 read length is 720 bp with an average quality score of 10.5. These results are similar when looking at individual life cycle stage reads.

Fig 1. Read lengths vs Average read quality plot
Reads were combined to assess size and quality, darker colour represents higher proportion of reads.



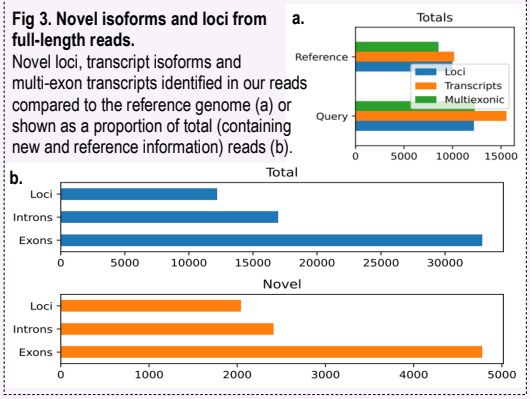
The majority of reads sequenced represented the full length of the mRNA transcript: At least 77% of total reads are full-length reads. Full-length reads are classified by recognition of primers present on both sides of the transcript. Reads with only 1 or no primers (22%) were not considered full-length and not used for further analysis.

Fig 2. Classification of output reads
Reads classified by Pychopper into full-length (primers found), rescued and without both primers.



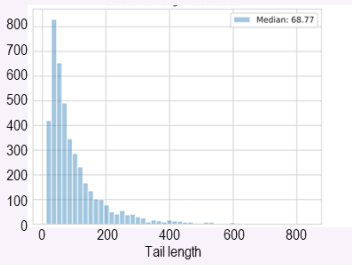
Over 5000 novel isoforms and 2000 new loci identified: Reads from all life cycle stages were combined prior to looking for alternatively spliced transcripts. 10000 isoforms available in *S. ratti* reference genome (a) to which we contributed 5000 novel isoforms. In addition, over 2000 new loci were identified which could potentially represent previously undiscovered genes.

Fig 3. Novel isoforms and loci from full-length reads.
Novel loci, transcript isoforms and multi-exon transcripts identified in our reads compared to the reference genome (a) or shown as a proportion of total (containing new and reference information) reads (b).



Poly A tails from full-length reads are around 70 nt long: Poly A tails predicted using Nanopolish from combined full-length reads, resulting in a median tail length of 69 nt. This is similar to information we have from *C. elegans* in which poly A tails vary from 70-90 nt. (Lima *et al.*, 2017)

Fig 4. Global poly A tail length distribution
Poly A tail length (in nt) showed in full-length reads of combined life cycle stages, with the median length just below 70 nt



CONCLUSION & FUTURE:

- We have identified 5000 new isoforms of *S. ratti* which will improve our current annotations of the genome for future analysis . In addition, the availability of full-length transcripts will allow us to calculate the 5' and 3' UTR regions which are important in gene regulation and small RNA binding.
- We have collected and sequenced parasitic adults, free-living adults and infective larvae separately and will compare isoforms, poly A tails and UTR regions to give an insight into whether alternative splicing between life cycle stages is important in parasitism.