Why Hymenolepis?

- Hymenolepis species (e.g., H. diminuta, H. nana, and H. microstoma) have been used as laboratory models since the 1950s, and thus much of our basic understanding of tapeworm biology stems from work on these species.
- The entire life cycle can be kept in the lab, making them more practical models than groups having medical or veterinary importance, such as Taenia and Echinococcus.
- The highly inbred ‘Nottingham’ strain is expected to show reduced variability, resulting in fewer assembly problems.

Materials & Methods

Data derive from a laboratory strain of H. microstoma (ie. ‘Nottingham’) maintained in vivo at the NHM using outbred conventional mice and flour beetles (Tribolium confusum). Specimens were removed from the bile ducts and intestines of mice and genomic DNA extracted from somatic tissues (ie. anterior parts of worm). RNA was extracted from whole adult worms as well as from larvae during mid-metamorphosis in flour beetles and use hooks and glands to enter the haemocoel.

The genome was assembled from 5 full Roche 454 Titanium runs (3 unpaired, 2 paired with ~3 Kbp inserts) and 3 Illumina GAII lanes, with a read-length of 76 bp. Illumina insert sizes for 2 lanes range from 300-400 bp, and for the third ~3000 bp. A de novo assembly of the genome was made using the software Newbler 2.3 (for Roche/454) and AbySS 1.2.0 (for Illumina), and contigs then merged using minimus2 from the AMOS Pipeline. The transcriptome was sequenced using separate lanes of Illumina data (ie. adult vs. larva), and mapped to the genome using BWA after screening for contaminants such as host DNA/RNA. Gene models are currently being constructed using Augustus, SNAP and Jig-saw.

The finished project will result in:

- A high-quality whole genome-assembly
- Transcriptomic data from multiple life-stages
- Annotated gene-predictions for a majority of genes, based on transcriptomic data and gene-prediction algorithms.

These data will provide a platform for further research in all areas of tapeworm biology, including development and drug targets.

Discussion

After the first year of the project we have generated 42x coverage of the genome and comparisons with similar amounts of data from Echinococcus show that the Hymenolepis assembly is highly efficient. These data will provide a platform for further research in all areas of tapeworm biology, including development and drug targets.

- A high-quality whole genome-assembly
- Transcriptomic data from multiple life-stages
- Annotated gene-predictions for a majority of genes, based on transcriptomic data and gene-prediction algorithms.

The latest genome assembly is available from: http://www.sanger.ac.uk/Projects/Pathogens/