Nematode genome sequencing update November 2006

Caenorhaditis genomes

These certainly are exciting times for anyone interested in comparative nematode genomics. Of course there is the complete *C. elegans* genome, and a *C. briggsae* draft assembly has been available for several years, but thanks to work of Ray Miller and his group there is now a high density, genetic map for *C. briggsae* (http://snp.wustl.edu/snp-research/c-briggsae/index.html), that has been used by LaDeana Hillier to organize the Phusion, cb25 assembly into chromosomal AGP files (http://genome.wustl.edu/genome.cgi?GENOME=Caenorhabditis%20briggsae&SECTION=assemblies). The *C. briggsae* chromosomes and annotation will be available soon on a Genome browser at WormBase.

C. remanei was the third *Caenorhabditis* genome to be sequenced when a preliminary, 6x, draft assembly became available not long ago (GenBank accession AAGD01000000). It is now undergoing additional, whole-genome sequencing and automated, primer-directed, sequence improvement at the Washington University Genome Sequencing Center (GSC) with a final draft assembly expected in early 2007. When completed, WormBase will help generate a gene set using methods derived from the nematode genome annotation assessment project (nGASP). The preliminary *C. remanei* assembly is available on a Genome browser (http://wormbase.org/db/seq/gbrowse/remanei/).

The fourth *Caenorhabditis* genome to be sequence will be *C.* sp. 4 (PB2801). A preliminary, 9x WGS, draft assembly has recently been completed at the GSC. It is undergoing checks for contamination and quality control, and should be available on a Genome browser at WormBase in the near future. Keep an eye on the 'News' on WormBase's homepage for an announcement. PB2801 will also undergo primer-directed sequence improvement, with a final assembly expected next year.

The final *Caenorhabditis* species scheduled for a draft genome sequence is *C. japonica*. Due to difficulties growing worms this genome project is moving slower than expected. A small plasmid library is currently being sequenced to check for possible fungal infection or contamination, but with a bit of hard work and a little luck, preliminary sequence and an assembly should be forthcoming in the next year.

Pristionchus genomes

A non-*Caenorhabditis* worm familiar to most *C. elegans* researchers, *Pristionchus pacificus*, var california has recently been sequenced to 9x

coverage and assembled at the GSC. It is currently undergoing checks for contamination and quality control and will be submitted to GenBank soon. Low-coverage (2-3x) draft assemblies of another strain, *P. pacificus*, var washington, and two other species, *P. maupasi* and *P. entomophagus* have been funded through a grant from the Max-Planck Society to Ralf Sommer and are currently being sequenced.

Parasitic nematode genomes

Outside the *Caenorhabditis* group a number of parasitic nematode genomes are being sequenced.

Haemonchus contortus, a Strongylid parasite of goats and sheep, is being sequenced at the Wellcome Trust Sanger Institute. Sequencing to 5x WGS coverage is done and an assembly is available (http://www.sanger.ac.uk/Projects/H_contortus/). Additional reads from fosmids are being added and finishing is planned.

Heterorhabditis bacteriophora is a small parasitic nematode used widely in the control of soil-dwelling insects. It has been funded by NHGRI for sequencing at the GSC (http://genome.wustl.edu/genome.cgi?GENOME=Heterorhabditis%20bacteriophora). Small-scale survey sequencing and heterozygosity testing is currently in progress. The results of these will be used to determine the sequencing plan and approach.

Meloidogyne hapla is a plant parasite and is also known as the root-knot nematode. It has been funded by NSF and USDA for 6-8x, WGS coverage with the sequencing being done at JGI.

Meloidogyne incognita, another root-knot nematode, has been funded for 10x, WGS coverage and BAC sequencing at GENOSCOPE.

Heterodera glycines, the soybean cyst nematode, has been sequenced to 3x coverage using 454 sequencing technology by Kris Lambert at the University of Illinois.

Brugia malayi is a filarial nematode that causes filariasis, or commonly known as elephantiasis, in humans. A ~9x WGS, high-quality, draft assembly and annotation is available at TIGR (http://www.tigr.org/tdb/e2k1/bma1/). A publication is in preparation.

Trichinella spiralis is a parasite of swine, and the most distantly-related species to *C. elegans* currently being sequenced. There is a 34.6x WGS draft assembly available at the GSC (http://genome.wustl.edu/genome.cgi?GENOME=Trichinella%20spiralis&SECTION=assemblies) with primer-directed sequence improvement and a BAC-based physical map underway.

Ascaris suum and *Ancylostoma caninum* are parasites of pigs and dogs respectively. Low-coverage (1x) WGS sequencing of both have been done at the GSC and are searchable on the NemaBLAST website (http://nematode.net/BLAST/). *A. caninum* is scheduled for a high-quality WGS draft assembly (see below).

Future genomes

We can expect sequence from an abundance of additional parasitic nematode genomes in the next 2-5 years. A five year plan at the Welcome Trust Sanger Institute has Helminth sequencing as one of it major programs, including 14 nematode species (http://www.sanger.ac.uk/Projects/Helminths/). And on the US side of the pond, a white paper was submitted to NHGRI in January proposing high-quality draft sequencing of up to 20 parasitic, Strongylida genomes. Ten of these have been approved for sequencing (http://www.genome.gov/page.cfm?pageID=10002154&sortcol=1,3). They are listed below followed by the organism they most commonly infect:

Ancylostoma caninum - dogs
Cooperia oncophora - cattle
Dictyocaulus viviparous - cattle
Necator americanus - humans
Nematodirus battus - sheep and goats
Nippostrongylus brasiliensis - rats
Oesophagostomum dentatum - pigs
Ostertagia ostertagi - cattle
Teladorsagia circumcincta - sheep
Trichostrongylus vitrinus - sheep and goats

Clearly much will be learned about the biology of all these interesting nematodes from their high-quality draft genomes, but *C. elegans* biology can also expect to profit from the comparative genomics studies that are sure to follow.