

# The WormBase Newsletter

November 2006



---

## WormBase and WormAtlas begin a Worm Community Forum

The [Worm Community Forum](#) is an on-line forum for discussion of any issue related to worms. People may post job openings, meeting announcements, discuss experimental results, protocols etc. If you are interested in acting as a moderator for the WCF please contact Thomas Burglin ([thomas.burglin@biosci.ke.org](mailto:thomas.burglin@biosci.ke.org)) or Todd Harris ([harris@cshl.org](mailto:harris@cshl.org)).

---

## Phenotype annotation in WormBase

To provide more detailed phenotype curation and to facilitate searching for objects associated (or not associated) with particular phenotypes, WormBase has implemented a variety of changes with respect to phenotype annotation. WS160 marked the introduction of the phenotype ontology, a structured vocabulary for phenotypes that is organized hierarchically, based on relationships between phenotype terms. The phenotype ontology currently consists of over 1,200 terms. These phenotype terms replace the previously existing vocabulary of approximately 120 terms. Whereas the previous structured vocabulary had been used for annotation of phenotypes attached solely to RNAi objects, the new ontology is being used for the annotation of phenotypes attached to RNAi, allele and transgene objects. The ontology is also being used to annotate phenotypes communicated as *not* being associated with a particular object (allele, RNAi experiment, or transgene). Concurrent with the advent of the new ontology, RNAi experiments derived from large-scale screens were updated with such “not” phenotype annotations. As a result of these changes, phenotype information attached to RNAi experiments has increased several-fold. WS160 featured over 200,000 such “not” RNAi-phenotype connections. Additionally, some phenotype annotations were converted to more granular terms, and RNAi objects annotated with

an “Unclassified” phenotype term are being converted to more appropriate terms.

WormBase is currently working on a tool that provides advanced searching and browsing capabilities for the phenotype ontology and other ontologies in WormBase. In the interim a temporary [search tool](#) can be accessed.

---

## Genetic interactions data set updated

We have updated genetic and regulatory interactions in WS168. Redundant interactions were removed and ~700 new interactions added. In total, there are 4,920 interactions (3,650 genetic and 1,270 regulatory). The interactions were extracted in two steps. During the first phase, ~5,800 *C. elegans* papers were processed by Textpresso to obtain sentences enriched for interactions. During the second phase, a curator manually extracted the interactions from the sentences. Please note that due to this semi-automated data extraction process, an interaction is attributed to the paper from which it was extracted, and not necessarily to the original paper. We also plan to begin manual curation of genetic interactions from data tables in published papers since these are likely to be underrepresented in our data set.

---

## An update on the sequencing of nematode genomes

These 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/cbriggsae/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?GE>

[NOME=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-see news item below). The preliminary *C. remanei* assembly will be made available on a genome browser. Please visit the Worm Base [wiki](#) for detailed information about the sequencing status of other *Caenorhabditis* species (*C. remanei*, *C. species PB2801*, *C. japonica*), *Pristionchus* species, and the

following parasitic nematode genomes: *Haemonchus contortus*, *Heterorhabditis bacteriophora*, *Meloidogyne hapla*, *Meloidogyne incognita*, *Heterodera glycines*, *Brugia malayi*, *Trichinella spiralis*, *Ascaris suum* and *Ancylostoma caninum*.

## **The Nematode Genome Annotation Assesment Project (nGASP)**

---

nGASP will be evaluating gene prediction software using the *C. elegans* genome. Interested parties are invited to read more about this on the [wiki](#).

## **The Gene Ontology Newsletter**

---

The November issue of the Gene Ontology Newsletter has been released and can be found at: <http://www.geneontology.org/newsletter/current-newsletter.shtml>.