

## WormBase Newsletter May 2004

### Using WormBase for bioinformatics:

#### Version freezes:

We would like to remind our users that WormBase offers version freezes. Every tenth release is available, so that bioinformaticians can refer to and work with a particular version of the *C. elegans* genome (eg., <http://ws120.wormbase.org>). If you need to refer to WormBase in a publication, please refer to one of the freeze versions.



#### Linking and Mining WormBase:

A powerful aceserver ([aceserver.cshl.org](http://aceserver.cshl.org)) has been set up and can be remotely accessed via command-line and scripts. It is easy to link to WormBase or extract information for data-mining purposes. You can link to HTML pages, text-only dumps and XML pages. For documentation regarding linking and mining WormBase (eg., mining WormBase with AcePerl, mining WormBase with MySQL and PostgreSQL, etc.) please see: <http://aceserver.cshl.org/about/linking.html#aceperl>

#### An update on the sequencing status of the *Caenorhabditis* genomes:

The whole genome shotgun sequencing phase of *C. remanei* has been completed and the ~1.43 million reads can be searched using BLAST at <http://genome.wustl.edu/blast/client.pl>. Fosmid end reads are in progress to be followed by two rounds of automated finishing, which should result in a high-quality draft assembly by the end of the summer. *C. japonica* and CB5161 are being backcrossed and checked for heterozygosity. These genomes are scheduled for whole genome shotgun sequencing in August and December, respectively.

#### A second mirror site for WormBase:

A second mirror site for WormBase has been set up at the California Institute of Technology, Pasadena (<http://caltech.wormbase.org>). This site and the first mirror site at the Institute of Molecular Biology and Biotechnology in Crete, Greece (<http://worm.imbb.forth.gr>) act as back-up servers and will promote greater accessibility to WormBase.

#### Submitting your data to WormBase:

WormBase provides on-line data-submission forms for the easy and quick submission of data from individual users. We would like to encourage you to use these forms and provide us with feedback on their ease of use and information content. This would help us improve and further develop this resource. Forms are available to submit genetic map data (like allele data, deletion/duplication data, SNP map data, etc.), transgene data, new/altered locus data, etc. For a complete list of data submission forms and contacts please see <http://www.wormbase.org/db/curate/base>

For submission of large scale data sets, please contact WormBase directly.

#### Using the power of Textpresso to search the *C. elegans* literature via 'category' based queries:

Textpresso now has 3,810 full text worm papers. In addition to the commonly used 'key-word' search, Textpresso allows searching of the literature based on 'categories'. This is a powerful feature and returns 'facts' about the entity (eg., a gene) you are interested in. For example, from the simple retrieval page of Textpresso (<http://www.textpresso.org/>) you can use the keyword (query), 'let-23' and specify the categories 'regulation', 'cell or cellgroup' and 'molecular function'. Performing the search returns 245 matches in 113 publications (matches can be viewed in the sentences in which they occur along with the surrounding sentences).

