

WormBase Newsletter February 2003

WormBase User's Survey

In preparation for our annual Advisory Board Meeting, we are conducting an on-line survey. Please access it from the WormBase home page. Your opinion in setting priorities and feedback on our quality of service would be of great help in planning the next year's activities. The survey is anonymous, but if you include your email, you will have a chance to win a WormBase mug with Thomas Boulin's award-winning WormBase Logo.



Release Schedule

A new version of the WormBase database is released approximately every two weeks. The new database can usually be downloaded every other Friday from the Sanger FTP site: <ftp://ftp.sanger.ac.uk/pub/wormbase> and this information is usually uploaded to the website over the subsequent weekend.

Proteome/Incyte

As many of you know, since the summer, access to Incyte/Proteome's useful wormPD database is now by subscription only. Over the past few years, we had purposely avoided focusing on information already being provided by wormPD. We have reset our priorities to provide users with the information most desired, based upon feedback at regional meetings this past summer. We have obtained additional resources from the NHGRI for this year to provide more information as soon as possible. Providing a brief description of each gene, now seen at the top of the newly reorganized Gene page, is currently one of our high priority tasks. Curating RNAi data from individual papers, and adding microarray data are other top priorities.

A complete *C. elegans* Genome Sequence!

WormBase now includes a complete *C. elegans* genome sequence. 100,264,081 base pairs are present in six contiguous segments corresponding to the six *C. elegans* nuclear chromosomes. On behalf of the *C. elegans* community we thank the *C. elegans* Genome Sequencing Consortium at the Wellcome Trust Sanger Institute and Washington University's Genome Sequencing Center for this impressive accomplishment. In spite of the high quality of the sequence, there are undoubtedly remaining errors in the genome sequence, and thus the Sequencing Consortium will remain active in responding to and resolving any potential errors that arise. The Sequencing Consortium wants to know of any potential problems in the now contiguous six *C. elegans* chromosomes, for example, any evidence of deleted YACs, etc. Please inform Alan Coulson (alan@sanger.ac.uk) or WormBase.

Genomic Clone Requests

Please send all requests for *C. elegans* genomic clones to Audrey Fraser (aef@sanger.ac.uk) with a cc to Ratna Shownkeen (rs2@sanger.ac.uk); please include a FedEx or DHL account number.

Genome Browser improvements

The Genome Browser now puts the track labels in the main image rather than in a key at the bottom. Some new tracks include: Knockout alleles and 3-frame translations.

Expression patterns: Request for data and corrections

We are up-to-date with extracting gene expression patterns from published papers. We will do our best to keep up with this information as new papers come out.

Now that WormBase has a very useful collection of expression patterns, we ask for help in three ways (in order of priority):

First, please check WormBase for your expression data. We apologize for any omissions or mistakes that we have made, but please tell us if you notice any! Please email any reports of errors, bugs or omissions to wormbase-help@wormbase.org.

Second, please submit expression pattern data. If you have old data that are in danger of being forgotten, submit it to WormBase via the homepage, click on Submit data; click on GeneExpression/Regulation or go directly to http://minerva.caltech.edu/~azurebrd/cgi-bin/expr_pattern.cgi. If you are willing to share data before publication, please submit it and it will be attributed to you.

Third, if you have a paper about to come out, feel free to send us a pre-print and we can get a head start on extracting the information; text files are useful since we can scan them for information about transgenes, etc.

RNAi

Not surprisingly, there is a remarkable increase in RNAi experiments in papers. We appreciate that these often have more thorough phenotypic analysis than the large scale screens. For technical reasons, the large-scale data gets into WormBase faster even if published later. We are working to catch up with the considerable backlog of the individual RNAi experiments from papers. If any omission bothers you, please let us know and we will get it in as fast as possible.

RNAi search improvements

The RNAi search now allows you to combine phenotypes with either AND or OR criteria.

The Textpresso search engine

WormBase has developed a search engine for *C. elegans* literature that allows queries with one or more key words as well as categories such as allele, gene, nucleic acid, pathway, etc. The search engine can be reached from the home page or directly at www.textpresso.org. Individual sentences in almost 2000 full text papers are included; more will be added in the coming months.

Persons and Authors

As part of an effort to better organize and correct mistakes in the connections between people and their papers, WormBase distinguishes Persons from Authors: Persons are unique individuals with contact information. We will be asking your help in making accurate connections between you and your publications. An extremely useful practice would be to use full names whenever possible in papers and abstracts. To check or update your contact information, please go to: <http://minerva.caltech.edu/~azurebrd/cgi-bin/forms/author.cgi>.

Coming Soon

More microarray data. A pipeline of data entry for Kim lab/Stanford microarrays has been established, and we will add data from published papers over the next few months.

People of WormBase

New Curators/Programmers at the Sanger Institute: **Anthony Rogers** and **Chao-Kung Chen** have joined the WormBase group. Chao-Kung is half-time with the CGC, and is the key liason between WormBase and the CGC.

New Programmers at Cold Spring Harbor Laboratories: **Jack Chen** and **Fiona Cunningham** have joined the WormBase group. Jack has a Ph.D. in neuroscience and an M.S. in computer science. Fiona has a background in bioinformatics and worked with decode Genetics before joining WormBase.

New Curators at Caltech: **Andrei Petcherski** (Ph.D. at the University of Wisconsin, Madison with Judith Kimble), **Carol Bastiani** (postdoc with Melvin Simon and Paul Sternberg at Caltech) and **Igor Anteshechkin** (former postdoc with Min Han at the University of Colorado, Boulder) have joined WormBase as biological curators. **Ranjana Kishore** (former postdoc with Meera Sundaram at the University of Pennsylvania) has joined WormBase as a Gene Ontology (www.geneontology.org) curator.

WormBase Advisory Board: Our advisory board helps set priorities and goals (and generally acts like a thesis committee). They ensure that the user community is best served, and that WormBase cooperates with other Model Organism Databases.

Thomas Blumenthal, *C. elegans* molecular biologist

Leon Avery, *C. elegans* neurobiologist; *Caenorhabditis elegans* WWW Server

Martin Chalfie, *C. elegans* developmental and neuro geneticist

Jonathan Hodgkin, *C. elegans* geneticist; curator of the CGC genetic map

J. Michael Cherry, PI, Saccharomyces Genome Database (SGD)

Janan Eppig, PI, Mouse Genome Database (MGD)

William Gelbart, PI, FlyBase

Stanley Letovsky, Bioinformatician.

Funding sources

WormBase is in its third year of funding from the National Human Genome Research Institute (# P41 HG02223), and receives additional support from the British Medical Research Council. Applications to these agencies are pending.