WormBase Newsletter   January 2006

WormBase conducted an extensive survey for our users in Nov/Dec 2005. We thank the many people who carefully responded to it. There were 761 respondents, about a third each graduate students, postdoctoral fellows and PIs. Half the respondents used WormBase daily, the results are therefore, somewhat biased against the infrequent user.

We discussed the results of the survey at our Advisory Board Meeting in November 2005 (advisors include, Cori Bargmann, Marty Chalfie, Jonathan Hodgkin, Tim Schedl, Mike Cherry, Bill Gelbart and Jim Kent), and used the survey to help prioritize plans for the next year. We discuss below plans that will address the concerns/suggestions from the survey, some will be implemented immediately, and others will be worked on steadily in the coming months.

Many of our user's requests are for information we would like to provide but do not have available; the results of large-scale studies are often not available until after publication. We do try to have data in WormBase upon publication. Authors are encouraged to contact WormBase, any such communications will be treated as confidential information.

Website:

Two hundred respondents felt that it is sometimes confusing to find information in WormBase. There was a broad range of suggestions for improvements, and thus every aspect of the user interface and site needs to be examined and improved. Individual complaints or suggestions are taken seriously, so users are strongly encouraged to contact us when something seems less than optimal. Often there are simple fixes and these can be implemented quickly.

Speed of website. We note that the speed of WormBase remains a concern. We plan additional measures to decrease the response time of WormBase. A new architecture of the hardware serving WormBase allows us to add additional servers, when the need arises. All pages are stored in a cache ensuring that subsequent accesses are speedy (although complex pages may require some time to collate the first time they are accessed following a new release of the database). We will pre-cache the most often accessed gene pages to make even the initial response time more rapid.

Gene Structure and Maps:

Gene structures. The highest priority for data remains *C. elegans* gene structures. The majority of respondents considered the structures mostly accurate but some incorrect. In some cases the predictions will not change until there is experimental data or comparative genomic data, but we are taking a number of steps to ensure that existing data are incorporated. The sequence curators at Washington University and the Sanger Institute will systematically look at approximately a megabase of sequence to determine the most useful data for improving gene structures.

Regulatory region sequence download. We will add a simple link that will extract 5' sequence for features on gene and sequence pages.

Genetic Map. We will retire the dysfunctional genetic map and replace it with an enhanced Genome Browser style view that displays an integrated genome map. We think we can provide all the functionality of a genetic map and some nice features from having combined physical, functional and genetic information. The concept is to have separate tracks for classic and useful genetic markers, for deletions, for verified SNPs, etc. The Genome Browser (GBrowse) allows printing of high quality graphic images, and thus genetic maps of various types will be easily obtained. A demonstration version of this map will be

available by the end of January.

Data Types:

Functional descriptions. The functional descriptions were considered mostly accurate but some incorrect, and missing some critical details. Respondents felt that WormBase should have researchers provide input, and that the source of the data for the concise descriptions should be clearer. We would like to get closer to our goal of having all the relevant information on or accessible from the gene page. This is also a strong motivation for us to do a systematic curation of phenotypes.

Phenotypes. We have initiated a major effort to curate phenotypes associated with alleles, RNAi, transgenes, and cell ablations to an organized set of phenotype terms (Phenotype Ontology). This will allow searches of gene function and more robust data-mining. We will be contacting several people to help us refine the draft ontology that we have developed. We do not plan to assign three-letter phenotype names to all terms.

Allele information. We will curate alleles more intensely. You can help by providing information about alleles via submission of data through online forms. Our data submission forms are available from this page: http://www.wormbase.org/db/curate/base

Searches:

Gene expression search. Over the past year or so we have reannotated gene expression patterns to an organized set of cell and anatomy terms (Cell/Anatomy Ontology). In the coming weeks, there will be an enhanced search capability for groups of cells. For example, a query for genes expressed in neurons will return genes expressed in *all* neurons and in *any* neuron. Querying for genes expressed in ASHL will return all genes expressed in ASHL even if broadly expressed.

Literature search. Many respondents commented on the ease of the literature search at Leon Avery's site. We will institute a simple keyword search that will provide the simplicity of this search engine. In addition, Textpresso (http://www.textpresso.org) is undergoing a major overhaul, including it's user interface. In a few months, the updated search engine will be available. In the meantime, the existing Textpresso searches will be available within WormBase and at the textresso.org site, with new papers and abstracts added weekly.

Tools:

Interaction viewer. After a careful study, we have chosen 'nBrowse', developed by Kris Gunsalus and coworkers to visualize interactions among genes (physical interactions, genetic interactions, regulation, etc.). This will be implemented in a few months.

Query and analysis tools. BLAST searches still rank the highest priority. Half the respondents felt that a full-text literature search was high priority.

Use of *C. briggsae* sequence. As the sequences of other *Caenorhabditis* species become available, how WormBase enables their use will become a larger issue. Some of the requests include: synteny views, multiple alignments of nucleotide and protein, orthology assignments and promoter region comparisons.

WormMart. We will continue to expand the data available through the WormMart data mining interface. In several months we will have better methods to add other WormBase information to WormMart. WormMart is very powerful but apparently not very easy to use. We will remedy this in two ways. First, we will provide short tutorials with screen shots and an automated click-through of the steps of a few queries. Second, we will refine the user interface and flow of typical queries.

Data Contribution:

We appreciate the considerable interest in having researchers provide data for WormBase. We absolutely encourage you to do so; we provide many web-forms for this purpose. Please use the web forms at http://www.wormbase.org/db/curate/base and complain if they are annoying in any way or suggest how they might be improved. We will try to request specific data types during specific periods to see if this can be a more efficient method of incorporating data in WormBase.

Another way in which you can help WormBase is to make sure detailed information is available in your published papers (primer sequences, tables of strains, transgenes, end-sequences, sequences of constructs, etc). This information could be easily added to supplementary material. The more systematic your presentation, the easier it is for us to include it in WormBase. We would love to work with you to get your results into WormBase, to be released concomitant with publication.

Pictures. We are streamlining the process by which images can be added to WormBase, and we will work with you to obtain and display images.

Community Forum and Wiki:
More than half of the respondents said they would use a Wiki or forum. We will import bionet as a community forum. We will add a Wiki that will contain up-to-date information of using WormBase as well as user-contributed comments on genes, protcols, job-openings, etc. Details concerning the editorial policy are being discussed. Watch for an announcement introducing the WormBase Wiki very shortly.

Online Documentation:

While users had an interest in additional documentation for WormBase, this was not, in general, high priority among respondents compared to interest in the user interface and datasets. However, given the inherent bias in this survey towards the frequent user who is familiar with WormBase, we strongly believe additional documentation and tutorials are a high priority. We will provide documentation to specific pages and sections. Much of this exists in the official WormBase User's Guide and in various book chapters. We will make this documentation more readily accessible on the site as a part of the web-site software redesign.

General Comment from WormBase:

The WormBase staff was delighted by the extent of feedback via the survey. Whenever you notice an error, something confusing, an omission or if you are finding it difficult to find data, please let us know. We can provide help and try to fix the problem for the next person.