

ZFIN: enhancements and updates to the zebrafish model organism database

Yvonne Bradford*, Tom Conlin, Nathan Dunn, David Fashena, Ken Frazer, Douglas G. Howe, Jonathan Knight, Prita Mani, Ryan Martin, Sierra A. T. Moxon, Holly Paddock, Christian Pich, Sridhar Ramachandran, Barbara J. Ruef, Leyla Ruzicka, Holle Bauer Schaper, Kevin Schaper, Xiang Shao, Amy Singer, Judy Sprague, Brock Sprunger, Ceri Van Slyke and Monte Westerfield

ZFIN, 5291 University of Oregon, Eugene, OR 97403-5291, USA

Received September 13, 2010; Revised October 13, 2010; Accepted October 14, 2010

ABSTRACT

ZFIN, the Zebrafish Model Organism Database, <http://zfin.org>, serves as the central repository and web-based resource for zebrafish genetic, genomic, phenotypic and developmental data. ZFIN manually curates comprehensive data for zebrafish genes, phenotypes, genotypes, gene expression, antibodies, anatomical structures and publications. A wide-ranging collection of web-based search forms and tools facilitates access to integrated views of these data promoting analysis and scientific discovery. Data represented in ZFIN are derived from three primary sources: curation of zebrafish publications, individual research laboratories and collaborations with bioinformatics organizations. Data formats include text, images and graphical representations. ZFIN is a dynamic resource with data added daily as part of our ongoing curation process. Software updates are frequent. Here, we describe recent additions to ZFIN including (i) enhanced access to images, (ii) genomic features, (iii) genome browser, (iv) transcripts, (v) antibodies and (vi) a community wiki for protocols and antibodies.

INTRODUCTION

ZFIN is a curated resource for zebrafish biology comprised of the following primary data types: genes, phenotypes, genotypes, gene expression, functional and phenotypic annotations, anatomical structures, orthology, nucleotide and protein sequence associations and reagents such as morpholinos and antibodies. Table 1 lists ZFIN

data contents as of July 2010. A tabular presentation of ZFIN's growth over the years can be accessed from the database (http://zfin.org/zf_info/zfin_stats.html). ZFIN data can be accessed using any of the data-type specific search forms, site search, BLAST, or GBrowse. A comprehensive suite of download files provide a means of accessing large quantities of data for further analysis. Special requests for data reports can be requested from zfinadmin@zfin.org.

ZFIN participates in regularly scheduled data exchanges, ranging from daily to monthly, with major bioinformatics organizations, such as the Wellcome Trust Sanger Institute, Ensembl, NCBI and UniProt resulting in reciprocal links that provide valuable cross-site data integration. These exchanges enhance data accuracy and consistency because curators work continuously to resolve identified discrepancies. In addition, we provide links to many community resources on our home page.

ZFIN's curation process utilizes bioinformatics community-supported best practices to ensure data are described accurately and consistently. One such practice is the use of standardized nomenclature. ZFIN, in conjunction with the Zebrafish Nomenclature Committee, serves as the authoritative source of gene and allele nomenclature. Standardized nomenclature is essential to unambiguous communication. Zebrafish nomenclature guidelines are coordinated with guidelines used for human and mouse genes. Similarly, standardization of functional and phenotypic gene annotations promotes robust searching and comparisons within and among species. ZFIN's annotations are based on the structured vocabularies and relationships defined by biological ontologies. These ontologies are evolving resources that require community input to ensure completeness and accuracy. ZFIN collaborates with the bioinformatics community on the development of several ontologies including

*To whom correspondence should be addressed. Tel: 541 346 4115; Fax: 541 346 0322; Email: ybradford@zfin.org

Table 1. Summary of ZFIN data content (July 2010)

ZFIN data statistics	July 2010
Genes	30 783
Genes on assembly	15 341
Transcripts	25 916
EST/cDNAs	34 865
Full length cDNA clones (ZGC)	17 191
Genomic features	6603
Transgenic features	2002
Transgenic constructs	591
Genotypes	9565
Genes with GO annotations	15 502
Genes with IEA GO annotations	11 741
Genes with non-IEA GO annotations	8728
Total GO annotations	1 08 694
Morpholinos	3571
Antibodies	542
Gene expression patterns	49 811
Images (phenotypes, expression patterns)	76 911
Anatomical structures	2669
Genes with human orthology	8937
Genes with mouse orthology	5506
Publications	13 063
Researchers	5177
Laboratories	640
Companies	125

Gene Ontology [GO; (1)], Cell Ontology [CL; (2)] and Phenotype Quality Ontology [PATO; (3)]. ZFIN also develops and maintains the Zebrafish Anatomical Ontology [ZFA; (4)]. ZFIN is the authoritative source for zebrafish GO annotations. Standardized evidence codes are used to support GO and orthology annotations. All data are attributed to their original sources.

ZFIN encourages comments and suggestions from the community. A ‘Your Input Welcome’ button is provided on every ZFIN data page to facilitate communication. ZFIN curators address incoming questions and data submissions. Requests for new features and enhancements, combined with annual user surveys results, play a key role in determining future directions.

NEW TO ZFIN

Enhanced access to images

ZFIN maintains an extensive repository of annotated figures derived from current literature and data submitted directly to ZFIN by researchers. Recent enhancements, based mainly on user requests, provide increased access to these images and have quickly become ZFIN’s most popular feature.

Annotated figures of gene expression patterns are included in this repository. Annotations associate genes, fish, developmental stages and terms from the ZFA ontology to each figure. It is often desirable to browse these figures, using the gene expression search form, for a marker with a particular gene expression pattern. A search for ‘integument’ returns nearly 700 markers. Individually reviewing this large number of matching markers can be a daunting task. A figure gallery thumbnail strip (Figure 1), displaying each figure that matches the search criteria, has been added at the top of each gene

expression search results page to provide a quick means to find the desired pattern. Mousing over a thumbnail pops up a larger image with links to detailed information. Controls located above the strip provide navigation through multiple thumbnail strips.

Terms from the ZFA ontology play an integral role in annotations of phenotype and gene expression data in ZFIN. In addition, definitions and relationships assist in the identification of anatomical structures throughout the development cycle. Each ZFA term is represented on an anatomy page that details synonyms, a definition, developmental stages during which the structure is present, parent and child structures and links to expression and phenotype data for the structure. Images of a select group of approximately 300 anatomical structures are now also available on these pages. Initial images are from Wolfgang Driever’s developmental atlas (http://zfin.org/zf_info/anatomy.html) of the embryo spanning stages prim-5 to Day 5 and images of high quality, labeled mRNA in-situ hybridization studies (5). Plans are underway to add postembryonic images (6) and ZFIN curators will add images from publications as part of our ongoing curation process. Submissions from researchers are also welcome. The anatomy detail page can be accessed using the anatomy browser available from our home page. Extensive data integration produces links to anatomy pages from figure, gene, feature and genotype pages.

Genomic feature page

Researchers gain an understanding of how genes function by investigating mutations, or genomic features, of a particular locus. These genomic features include point mutations, insertions and other chromosomal aberrations like deficiencies and translocations. Understanding how a genomic feature affects gene function is key to elucidating biological pathways, disease, gene product interaction and gene regulation. To help researchers begin understanding genomic alterations, the Genomic Feature page (Figure 2) was developed as a resource page that details information about individual genomic features. This page provides information about affected genes, mutation type, protocol used to induce the mutation, lab of origin, mapping details and a table of all associated genotypes with links to phenotype and gene expression data. A camera icon indicates that genotype images are available on ZFIN. Links provide easy navigation to affected gene and genotype pages. The Feature page can also be accessed from links on the Mutant/Morphants/Transgenics Search Results page.

GBrowse genome browser

ZFIN has implemented a genome browser, GBrowse, developed by the Generic Model Organism Database (GMOD; <http://gmod.org>). GBrowse provides a graphical, interactive and customizable web interface to explore the zebrafish genome, including information about genome annotations and the positions of genes on the chromosomes. Tracks are aligned to the current version of the Wellcome Trust Sanger Institute Vertebrate

Modify Search
Your Input Welcome

Figure Gallery (164 images)

Search matched on current symbol: **fgf8a**

Publication (current status) Data

Dutton et al., 2009	Fig. 4
	Fig. 6
Hong et al., 2009	Fig. 6
Paridaen et al., 2009	Fig. 3
Regan et al., 2009	Fig. 2
Shima et al., 2009	Fig. 2
	Fig. 4

Plaster et al., 2007 **Fig. 6**

All Fig. 6 genes: [dlx2a](#), [fgf8a](#), [pax2a](#)

[hindbrain boundary neural keel, te](#)

Figure 1. Gene expression results page depicting the figure galley thumbnail strip with an enlarged thumbnail image showing Figure 6 from Plaster et al., 2007 (9). Thumbnail strip navigation controls have been used to advance to strip 14 of 17.

ZFIN ID: ZDB-ALT-960203-1091

Genomic Feature: **ti282a**

Previous Names:

Affected Genes: [fgf8a](#)

Type: Point Mutation (1)

Protocol: adult males treated with ENU

Lab Of Origin: [Nüsslein-Volhard Lab](#)

Map: LG: 13, [Details](#)

Notes:

GENOTYPES:

Genotype (Background)	Affected Genes	Phenotype	Gene Expression
Et(krt4-EGFP)sqet1.1;fgf8a^{ti282a/ti282a}	fgf8a	--	1 figure from Regan et al., 2009
fgf3^{124149/+};fgf8a^{ti282a/+}	fgf8a , fgf3	--	1 figure from Nechiporuk et al., 2007
fgf3^{124149/+};fgf8a^{ti282a/ti282a}	fgf8a , fgf3	1 figure from Nechiporuk et al., 2007	1 figure from Nechiporuk et al., 2007
fgf3^{124149/n24149};fgf8a^{ti282a/ti282a}	fgf8a , fgf3	1 figure from Nechiporuk et al., 2007	2 figures from 2 publications
fgf3^{124149/n24149};fgf8a^{ti282a/+}	fgf3 , fgf8a	1 figure from Nechiporuk et al., 2007	1 figure from Nechiporuk et al., 2007

[Show all 20 genotypes](#)

Figure 2. Screenshot depicting the ti282a feature page. Links are provided to all ti282a data in ZFIN.

Genome Annotation [Vega; (7)] genome assembly and link to data-specific pages at ZFIN or Vega. Individual tracks show ZFIN genes with phenotype, ZFIN genes with gene expression data, Vega genes, transcripts, BACs used as assembly components, morpholinos and antibodies. GBrowse can be launched from ZFIN's home page or from the GBrowse interactive graphic found on the gene, transcript and clone detail pages.

Transcript page

Transcript pages at ZFIN were created to integrate data from the zebrafish whole genome sequence and annotation project at the Wellcome Trust Sanger Institute and present users with the set of known transcripts for a gene. This is a necessary first step towards associating expression, function and phenotype data with specific transcripts.

ZFIN ID: ZDB-TSCRIPT-090929-6190

Your Input Welcome

Transcript Name: fgf8a-001
Transcript Type: mRNA

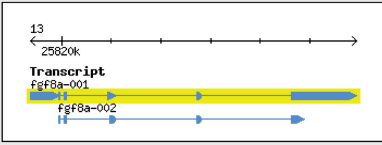
Sequence:
 OTTDART00000022966 (1) [Show] [Download] - Select Analysis Tool -

Associated with Genes: fgf8a (1)

SEGMENT (CLONE AND PROBE) RELATIONSHIPS:
 fgf8a-001 Contained in [BAC] CH211-19418 (1) (order this)

fgf8a TRANSCRIPTS:

Type	Name	Length (bp)	Analysis
mRNA	fgf8a-001	2238	
	fgf8a-002 (1)	600	- Select Tool -



OTHER fgf8a-001 TRANSCRIPT PAGES:
 • [Vega Trans:OTTDART00000022966 \(1\)](#)

PROTEIN PRODUCTS:
 • [VEGAPROT:OTTDARPO00000019154 \(1\)](#)

ZEBRAFISH SUPPORTING SEQUENCES:

Type	Accession #	Length (bp/aa)	Analysis
Polypeptide	GenPept:AAB82614 (1)	210 aa	- Select Tool -
	GenPept:AAC41302 (1)	210 aa	- Select Tool -
RNA	GenBank:BC086718 (1)	2146 bp	- Select Tool -
	GenBank:AF030560 (1)	1710 bp	- Select Tool -
	GenBank:AF034264 (1)	1460 bp	- Select Tool -

Figure 3. The transcript page summarizes transcript-specific data for fgf8a-001.

The transcript page (Figure 3) currently provides information about the transcript including name, type, associated genes and sequences and other transcripts produced by the gene. Supporting sequences, used as evidence for the transcript by the Sanger annotation team, are also provided. Transcript types are defined using commonly accepted terms from the Sequence Ontology [SO; (8)]. Links to Vega transcript and protein pages are also provided. Sequence searches can be performed at ZFIN, Vega, Ensembl, NCBI and UCSC by selecting the appropriate option from the sequence analysis tools menu. Additionally, the transcript page includes a GBrowse graphic depicting the relevant portion of the genome. Clicking the graphic launches GBrowse that provides access to data for transcripts at the genome level (Figure 4). All transcripts are searchable using the Genes/Markers/Clones search form. Transcript records can also be accessed from gene pages and by sequence similarity search using the BLAST tool at ZFIN.

Antibodies

Antibodies are widely used as probes for gene expression studies and as labels of anatomical structures, creating a need for a centralized resource that provides information about antibodies known to work in zebrafish. ZFIN now captures antibody data as part of our literature

curation process. An antibody nomenclature convention has been adopted using gene symbols or gene family stem symbols when possible. Antibody data can be retrieved using the new antibody search form. The search form provides options to find antibodies that label specific anatomical structures or that target products of a specific gene. Antibody specific information, such as name, type, host organism and assay may also be used to refine searches. A search results page provides links to matching antibody pages. The information provided on the antibody page (Figure 5) includes the name, aliases, host organism, isotype, type and assays. A summary of wild-type zebrafish labeling patterns includes labeled anatomical structures, developmental stage, assay, associated gene and links to published figures. Sources and usage notes are provided when available. A link is provided to the corresponding community wiki antibody page for community-contributed comments. Links to antibodies that label a specific structure are also available on ZFIN anatomy pages.

Community wiki

The zebrafish community has an untapped wealth of valuable unpublished knowledge about protocols and reagents. Recently, ZFIN implemented a Community Wiki (<http://wiki.zfin.org>) for experimental protocols

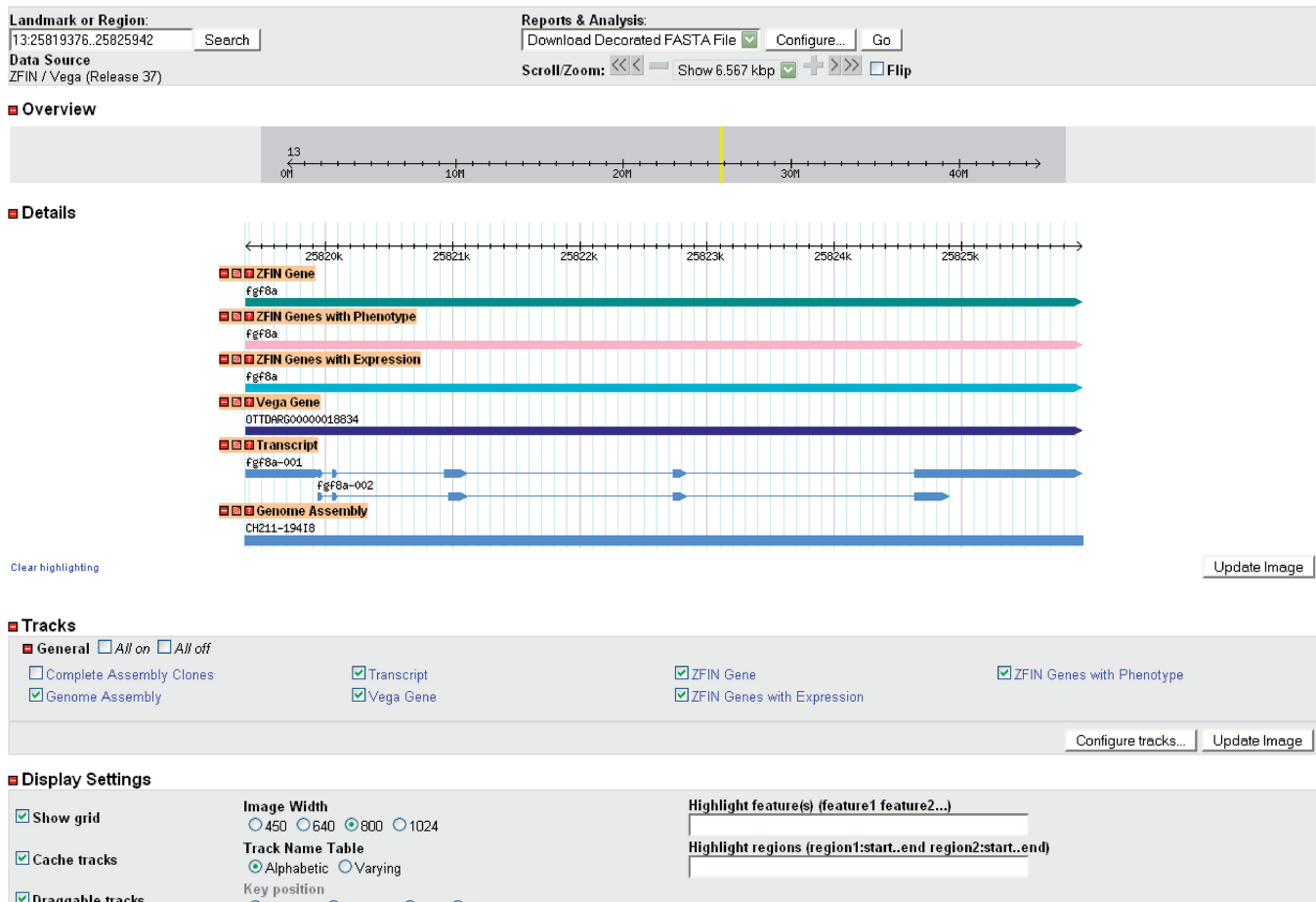


Figure 4. The GBrowse view of the genome, centered on *fgf8a-001* transcript.

and antibodies to foster community collaboration and sharing in a centralized location. Community provided information supplements, but is not directly integrated with, the published data curated at ZFIN. Contributions provide useful practical tips such as RNAi hairpin design for zebrafish transgenics or preferred fixatives for antibody staining. The wiki format supports easy searching, record creation and editing. A table of contents, the most recent comments and a catalog of the most recently updated pages are also displayed. Login is not required to view information in the wiki. A login link allows registered users to login and contribute new records or comments. The adjacent sign up link can be used by new users to create an account. Instructions for adding new records and comments are provided on the home page for each wiki section (Figure 6). All wiki records can be viewed online or exported as PDF or Word documents.

The Protocol Wiki contains protocols from *The Zebrafish Book* (5th edn), along with protocols shared by researchers through direct submission. Only the submitter can modify a protocol. Other registered users are encouraged to use the comments field to provide additional tips. The Antibody Wiki contains community submitted information, all ZFIN curated antibody records and all

antibodies available at the Zebrafish International Resource Center (ZIRC). Individual antibody pages (Figure 7) provide information about antibody names, aliases, catalog IDs, antibody details, structures the antibody labels, target molecules, recognized genes with links to the ZFIN gene page, suppliers, assays tested, notes, comments and links to corresponding ZFIN antibody pages. The community wiki can also be accessed from links provided on ZFIN's home page or by using the Site Search function provided on ZFIN's home page.

SUBMITTING DATA

ZFIN encourages researchers to share unpublished data using the Phenote software package. Phenote facilitates the annotation of gene expression patterns and mutant phenotypes with the same zebrafish ZFA, GO and PATO ontology terms used by ZFIN curators to annotate published data. The use of common terms provides easy integration into and searching of ZFIN. Phenote is available at <http://www.phenote.org/download.shtml>. All submitted data are directly attributed to their sources.

ZFIN ID: ZDB-ATB-081118-2

Antibody Name: **Ab3-dag1** [ab3-dag1 Wiki Page](#) Your Input Welcome
Alias: 43DAG1/8D5 (1), anti beta dystroglycan (1), anti-beta-dystroglycan (1), NCL-b-DG (1)

Host Organism: Mouse
Immunogen Organism: Human
Isotype: IgG2a
Type: monoclonal
Assays: Immunohistochemistry , Western blot
Antigen Genes: [dag1](#) (1)

NOTES: None Submitted

ANATOMICAL LABELING

Anatomy : Substructure	Stage	Assay	Gene	Data
brain	Prim-15	IHC	dag1	1 figure from Moore <i>et al.</i> , 2008
cerebellum	Prim-15	IHC	dag1	1 figure from Moore <i>et al.</i> , 2008
cloacal chamber	Prim-15	IHC	dag1	1 figure from Moore <i>et al.</i> , 2008
lens	Prim-15	IHC	dag1	1 figure from Moore <i>et al.</i> , 2008
myotome	Prim-15	IHC	dag1	1 figure from Moore <i>et al.</i> , 2008

[Show all 9 labeled structures](#)

SOURCE:
[Leica Microsystems Inc.](#)

CITATIONS (7)

Figure 5. ZFIN Antibody page for Ab3-dag1 antibody.

ZFIN Antibody Wiki Edit Add Tools

Added by [Jonathan Knight](#), last edited by [Nathan Dunn](#) on Jan 29, 2010 ([view change](#))

Welcome to the ZFIN Community Antibody Wiki. This is where zebrafish researchers can help each other by sharing antibody information - new antibodies, protocols, tips - anything that might be helpful.

Finding an antibody

Antibodies often have several different names. If you don't see an antibody in the list below, double-check by searching for it or any of its related information using the antibody search-field:

Adding information to an existing antibody

To add a short note or comment to an existing antibody, click the link "Add Comment". To add a new experimental protocol, create the protocol in the [Protocol Wiki](#) and use the comments to link to the correct protocol in the antibody (or update the notes if it is your antibody).

Entering a new antibody

To add a new antibody to the wiki, first make sure that the antibody is not already represented here, then enter the antibody name and click the create-new-antibody button below.

The new-antibody template has several fields. In addition to the antibody name, try to fill in as much information as possible. You can use this [antibody example](#) as a template.

New Antibody Name:

Community Submitted Antibodies

Antibodies represented here on the Wiki but not (yet) in the [ZFIN antibody database](#).

- [Anti - Cadherin E](#)
- [Anti-tyrosine hydroxylase antibody](#)
- [AP 2](#)
- [example-antibody](#)

[View All Community Antibodies](#)

ZFIN Antibodies

These antibodies have been imported from the [ZFIN antibody database](#).

- [ab-10e4](#)

Recent Comments

Recently Updated

Anti - Cadherin E	commented by Ping Yang	Mar 04
example-antibody	commented by Nathan Dunn	Jan 29
ab1-elav	commented by Dawn Westhelle	Jan 06
ab1-elav	commented by Adam Rich	Oct 25, 2009

Recently Updated Pages

Recently Updated

ab1-panx1	updated by Web Service User (view change)	Jun 09
ab2-army	created by Web Service User	Jun 06
znp-1	updated by Web Service User (view change)	Jun 03

Figure 6. Community Wiki Antibody section homepage.


[Edit](#) [Add](#) [Tools](#)

 Added by [Web Service User](#), last edited by [Web Service User](#) on May 30, 2010 ([view change](#))

[Add Comment](#)

Antibody Name:	Ab3-dag1 from the ZFIN antibody database .
Other names, clone ids, catalog ids etc.	43DAG1/8D5 , anti beta dystroglycan , anti-beta-dystroglycan , NCL-b-DG , dag1
Does it work on zebrafish?	yes
Host organism	Mouse
Antibody isotype	IgG2a
Antibody type	monoclonal
Anatomical structures recognized (use terms from the ZFIN Anatomical Ontology)	brain cerebellum cloacal chamber lens myotome retina vertical myoseptum whole organism
Recognized target molecules (gene names, domains, epitopes ...)	
Recognized ZFIN genes current ZFIN-gene symbols only ZFIN Gene Search	dag1
Supplier(s)	Leica Microsystems Inc.

Assays Tested

Assay	Prep	Worked	Notes
immunohistochemistry		yes	from ZFIN curation
western blot		yes	from ZFIN curation

Notes

- Imported from ZFIN Antibody page [ab3-dag1](#)
No notes imported.
- [Citations for Ab3-dag1 at ZFIN](#)

[Add Comment](#)

Figure 7. Antibody Wiki page for the Ab3-dag1 antibody.

FUTURE DIRECTIONS

ZFIN will continue detailed curation of current data types. Support will be expanded to include associations between human diseases and zebrafish genes and phenotypes. In addition, ZFIN is designing a new collection of search and browsing tools that will provide enhanced access to the rapidly expanding collection of data. We will soon integrate InterMine (<http://www.intermine.org>) into the ZFIN site.

CITING ZFIN

Please cite this article for a general reference to the ZFIN database. In addition, the following format is suggested for citing a specific entry in ZFIN. [Type of] data for this paper were retrieved from the Zebrafish Model Organism Database (ZFIN), University of Oregon, Eugene, OR 97403-5274; <http://zfin.org/>; [the date you retrieved the data cited].

IMPLEMENTATION

ZFIN is currently implemented with IBM/Informix relational database management software. Web-based HTML

forms combined with Java/JSP, GWT, JavaScript, Perl and CGI scripts provide access to the database. The community wiki is powered by Atlassian Confluence software (<http://www.atlassian.com/software/confluence/>).

FUNDING

National Institutes of Health (P41 HG002659, R01 HG004838 and R01 HG004834). Funding for open access charge: National Institutes of Health (HG002659).

Conflict of interest statement. None declared.

REFERENCES

- The Gene Ontology Consortium. (2010) The Gene Ontology in 2010: extensions and refinement. *Nucleic Acids Res.*, **38**, D331–D335.
- Bard,J., Rhee,S.Y. and Ashburner,M. (2005) An ontology for cell types. *Genome Biol.*, **6**, R21.
- Washington,N.L., Haendel,M.A., Mungall,C.J., Ashburner,M., Westerfield,M. and Lewis,S.E. (2009) Linking human diseases to animal models using ontology-based phenotype annotation. *PLoS Biol.*, **7**, e1000247.
- Sprague,J., Bayraktaroglu,L., Clements,D., Conlin,T., Fashena,D., Frazer,K., Haendel,M., Howe,D.G., Mani,P., Ramachandran,S.

- et al.* (2006) The zebrafish information network: the zebrafish model organism database. *Nucleic Acids Res.*, **34**, D581–D585.
5. Thisse, B., Pflumio, S., Furthauer, M., Loppin, B., Heyer, V., Degrave, A., Woehl, R., Lux, A., Steffan, T., Charbonnier, X.Q. *et al.* (2001) Expression of the zebrafish genome during embryogenesis (NIH R01 RR15402), ZFIN Direct Data Submission.
 6. Parichy, D.M., Elizondo, M.R., Mills, M.G., Gordon, T.N. and Engeszer, R.E. (2009) Normal table of postembryonic zebrafish development: staging by externally visible anatomy of the living fish. *Dev. Dyn.*, **238**, 2975–3015.
 7. Wilming, L.G., Gilbert, J.G., Howe, K., Trevanion, S., Hubbard, T. and Harrow, J.L. (2008) The vertebrate genome annotation (Vega) database. *Nucleic Acids Res.*, **36**, D753–D760.
 8. Eilbeck, K., Lewis, S., Mungall, C.J., Yandell, M., Stein, L., Durbin, R. and Ashburner, M. (2005) The Sequence Ontology: a tool for the unification of genome annotations. *Genome Biol.*, **6**, R44.
 9. Plaster, N., Sonntag, C., Schilling, T.F. and Hammerschmidt, M. (2007) REREa/Atrophin-2 interacts with histone deacetylase and Fgf8 signaling to regulate multiple processes of zebrafish development. *Dev. Dyn.*, **236**, 1891–1904.