



ZFIN NEWS

The Zebrafish Information Network

<http://zfin.org>

Volume 1, Number 1
Summer 2004

In this issue:

- ZFIN - A valuable resource for the research community (**pg. 1**)
- Zebrafish Gene Collection now available in ZFIN (**pg. 1**)
- What's new on the ZFIN Gene Page? (**pg. 2**)
- Easier Access to Gene Expression & Anatomy (**pg. 3**)
- Phenotype Futures (**pg. 3**)
- Gene Ontology Annotation of Zebrafish Genes (**pg. 4**)

ZFIN - A valuable resource for the zebrafish research community

What is ZFIN?

The Zebrafish Information Network, ZFIN, serves as your zebrafish model organism database by providing a central location for the curation and integration of zebrafish genetic, genomic and phenotypic data. Web accessible query interfaces for mutant, gene, marker, clone, mapping and expression data facilitate integrated studies of zebrafish functional genomics.

ZFIN - Past & Present


ZFIN is a dynamic resource with new data added daily and new features added frequently. Scientific curators extract and attribute pertinent information from the literature. Collaborations with zebrafish laboratories, data exchanges with other bioinformatics organizations such as NCBI, the Sanger Institute, the Zebrafish Gene Collection (ZGC) and Swiss-Prot, and personal communications from you regularly provide data to ZFIN. Table 1 depicts the growth of ZFIN over the years.

We are currently expanding curation of zebrafish phenotypes and expression data, annotating ZFIN gene records using the structured vocabularies

developed by the Gene Ontology (GO) consortium to describe and classify genes and gene products, and increasing support for whole genome sequence data.

ZFIN - The Future

Communication from the zebrafish research community is invaluable in helping us maximize the utility and usability of ZFIN. We encourage you to send your comments and suggestions to zfinadm@zfin.org

We will publish ZFIN News periodically to help keep you abreast of new ZFIN resources. Watch for it on our website (<http://zfin.org>) in PDF format. 

Zebrafish Gene Collection now available in ZFIN

We have incorporated full length cDNA clones from the Zebrafish Gene Collection (ZGC) initiative into ZFIN. ZGC cDNA clones are available to the scientific community without restriction. Purchasing information is available at <http://zgc.nci.nih.gov/Info/Buy>.


Many cDNA sequences are curated and associated with their encoding genes, and we will add more as data become available. For sequences that represent genes novel to ZFIN, we make the association after the creation of a new gene record. These novel genes have nomenclature based upon the ZGC clone accession ID. ZFIN will continue to revise and update nomenclature for these genes as information from paralogous and orthologous genes becomes available. Electronic Gene Ontology (GO) annotations based upon Swiss-Prot keyword, enzyme commission number and InterPro domain translations are also available for some of these genes. 

Table 1: ZFIN growth - 1998 to 2004

	1998	1999	2000	2001	2002	2003	2004 (thru June)
Genes & Genetics:							
Genes	220	608	974	1384	1715	12205	14517
Alleles	1896	1951	2018	2088	2278	2343	2484
ESTs	0	97	5845	9169	10493	11847	12134
Genes with GO Annotations	0	0	0	0	0	1353	2790
Expression & Phenotypes							
Gene expression patterns	0	0	0	83	739	1295	1364
Images (phenotypes, expression patterns)	1851	1858	1866	2598	8473	16874	17650
Anatomical structures	0	0	0	747	1112	1132	1132
Genomics:							
Mapping panels	3	5	5	7	7	7	7
Marker mappings	3138	2942	10509	14891	25483	26858	31017
Links to other databases	0	8621	31589	46522	60417	99118	116931
Community Information:							
Publications	1913	2415	2856	3517	4246	4986	5385
Researchers	1403	1640	2092	2431	2683	2916	3033
Laboratories	190	228	280	324	329	372	391



Be sure to visit the ZFIN exhibit at the Zebrafish Meeting in Madison!

What's new on the ZFIN gene page?

In ZFIN, the gene page is a central hub from which a variety of information is accessible (Figure 1). This wealth of summarized information has made gene pages the most accessed in ZFIN. We update the gene pages frequently due to ongoing development of existing data types, as well as the addition of new data. Keeping abreast of these changes can present a challenge to even the most seasoned users. The most recent changes to the ZFIN gene pages include:

GENE PRODUCTS: This section contains a list of Gene Ontology terms (GO) as well as links to external databases storing information on protein families, domains, and sites. Gene Ontology information and its display are explained in more detail in a companion article in this issue (see page 4).

SEGMENT(CLONE and PROBE) RELATIONSHIPS: This section provides links to ZFIN cDNA and genomic DNA segment records. The relationship between the gene and the DNA segment is also indicated. ZFIN has recently incorporated clones from the Zebrafish Gene Collection (ZGC) and genome sequencing project. Links to DNA segment pages supporting ZFIN mapping and expression data can also be found in this section.

SEQUENCE INFORMATION: The redesigned sequence information section better summarizes sequences, provides more information, and increases functionality. Sequences are summarized by links to representative cDNA, genomic, and polypeptide sequences as well as the UniGene cluster on the gene page. All cDNA, genomic, polypeptide, and cluster sequences associated with the gene are found on a separate page accessed by selecting the "All Sequence Information" link.

ZFIN Publications People Labs Companies Acc#
Home Mutants / Transgenics Wild-Types Genes / Markers / Clones Expression Maps

ZFIN ID: ZDB-GENE-980526-436

Gene Name: *retinoid x receptor, beta*
Gene Symbol: *rxrb*
Previous Names: etlD309733.19 (1) ; rxre (1)
[Nomenclature History](#)
Description:

GENE PRODUCTS:

Gene Ontology

Ontology	Qualifier	GO Term
Molecular Function	Contributes to	transcriptional repressor activity (more)
Biological Process		transcription (more)
Cellular Component		nucleus

[All GO Annotations \(10\)](#)

Protein Families, Domains and Sites:

- [InterPro:IPR000003](#) (1)
- [InterPro:IPR000536](#) (1)
- [InterPro:IPR001628](#) (1)
- [InterPro:IPR001723](#) (1)
- [InterPro:IPR008946](#) (1)
- [PROSITE:PS00031](#) (1)
- [Pfam:PF00104](#) (1)
- [Pfam:PF00105](#) (1)

SEGMENT (CLONE AND PROBE) RELATIONSHIPS:

rxrb Contained in [PAC] [BUSM1-12F11](#) (1)
rxrb Encodes [cDNA] [MGC:66116](#) (1)

SEQUENCE INFORMATION:

Type	Accession #	Length	
cDNA:	RefSeq:NM_131275 (1)	2508 bp	(BLAST IT!)
Genomic:	Genbank:AL672176 (1)	100037 bp	(Mega BLAST IT!)
Polypeptide:	SWISS-PROT:Q7SYN5 (1)	471 aa	(BLAST IT!)
Sequence Clusters:	UniGene:351 (1)		

[All Sequence Information \(13\)](#)

OTHER *rxrb* GENE / MARKER PAGES:

- [LocusLink:30530](#) (1)
- [VEGA:OTTDARG0000001497](#) (1)

MAPPING INFORMATION:

LG: 19 [Details](#) View Map: [Merged](#) [Individual Panels](#)

ORTHOLOGUES:

Species	Symbol	AccID	Evidence	
			AA	CL
Zebrafish	<i>rxrb</i>		•	•
Human	<i>RXRB</i>	<ul style="list-style-type: none"> OMIM:180246 LocusLink:6257 	•	•
Mouse	<i>Rxrb</i>	<ul style="list-style-type: none"> MGI:98215 LocusLink:20182 	•	•

Number of References (1) (1) (2) (1)

[CITATIONS](#) (16)

Figure 1: ZFIN gene page example

Additional information such as length and sequence type has been added to each sequence. We have increased the functionality of the sequence information section by adding "BLAST IT!" or "MegaBLAST IT!" buttons that bring up the sequence in a NCBI BLAST query form.

OTHER ... GENE/MARKER PAGES: We now provide links directly to zebrafish gene pages in

LocusLink and in the Sanger Institute's Vertebrate Genome Annotation database (Vega). Researchers can also access marker pages from The Sanger Institute's fingerprinting map of the zebrafish genome (Fingerprint Contig or FPC) from this location.

ORTHOLOGUES: The redesigned orthology display has evidence codes to indicate the types of evidence

(continued from page 2)

used to support assertions of orthology. Evidence includes amino acid comparison, nucleotide sequence comparison, and conserved location. Multiple columns in a single evidence code section indicate that different publications use that particular type of evidence to support orthology between different sets of species. ↩

Phenotype Futures

ZFIN is improving its support of zebrafish phenotypes with the goal of providing comprehensive information about mutants and morpholinos. We offer annotation tools to laboratories that are generating phenotype data. These tools promote the use of approved nomenclature for anatomical structures, biological processes, and phenotypic attributes that are consistent with the structured vocabularies (ontologies) in ZFIN. We encourage investigators who are interested in these tools to contact us at zfinadmn@zfin.org.

ZFIN curators are beginning to curate phenotypes reported in the literature using these same vocabularies. In the meantime, we are developing new query forms and page displays that will fully integrate phenotypes with other ZFIN data. ↩



Keep your ZFIN personal and lab records up to date!

For assistance, or to request a new ZFIN record, contact:

zfinadmn@zfin.org

Easier Access to Gene Expression and Anatomy

Patterns of gene expression in developing zebrafish embryos contribute vitally to the understanding of vertebrate gene function. ZFIN is expanding its coverage of gene expression and has significantly changed its expression query page to make searching faster and easier.

Researchers may search ZFIN for expression patterns by gene name, mutant background, developmental stage, and other criteria, as well as by entering anatomical structures. Researchers may also use an anatomical hierarchy to browse for data of interest or to find descriptions of specific structures (Figure 2).

Currently ZFIN has a large database of directly-submitted, annotated expression patterns (Figure 3). In addition, ZFIN curators have indexed hundreds of recent publications reporting gene expression, which can be found with the new search interface. We are developing the means to add figures to ZFIN from the literature, to associate them with anatomical structures, and to make them fully searchable. With time, we will populate anatomical structure descriptions with definitions, related structures, images, links to expression and phenotypic data, and homologous structures of other model organisms. ↩

Segmentation:14-somite (16.00h-19.00h)

- [embryo](#) [1285 gene(s) expressed at this stage]
- [branchial arches](#) [332 gene(s) expressed at this stage]
- [ectoderm](#) [741 gene(s) expressed at this stage]
- [epidermis](#) [148 gene(s) expressed at this stage]
- [neural crest](#) [52 gene(s) expressed at this stage]
- [endoderm](#) [31 gene(s) expressed at this stage]
- [EVL](#) [86 gene(s) expressed at this stage]
- [hatching gland](#) [86 gene(s) expressed at this stage]
- [macrophages](#) [14 gene(s) expressed at this stage]
- [mesenchyme](#) [466 gene(s) expressed at this stage]
- [head mesenchyme](#) [45 gene(s) expressed at this stage]
- [mesenchyme derived from head mesoderm](#) [15 gene(s) expressed at this stage]
- [mesenchyme derived from head neural crest](#)

Figure 2: Hierarchic list of structures from the ZFIN anatomical browser.

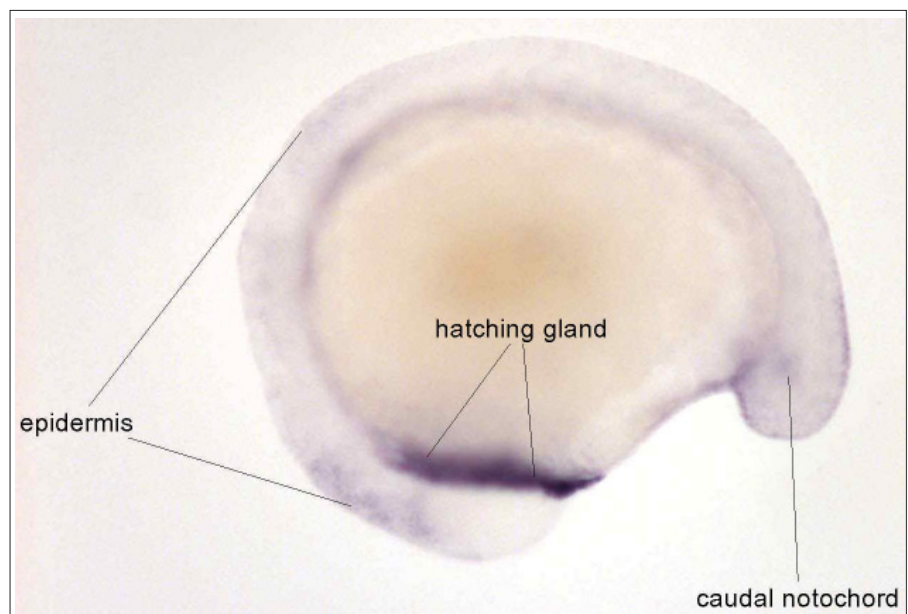


Figure 3: Expression pattern of clone cb918 annotated with structures from the anatomical dictionary. Source: Direct data submission, Thisse Laboratory, Strasbourg, FR

Gene Ontology Annotation of Zebrafish Genes

Biological ontologies are constrained and structured vocabularies used to describe a specific biological domain such as anatomy, nucleotide sequence features, phenotypes, or protein functions. In recent years there has been an increased effort to produce and employ such ontologies. The shared use of these biological ontologies promotes consistent usage of terms and their meaning across databases, supporting cross-species genomics studies and computational aspects of data storage and retrieval.

What is the Gene Ontology?

The Gene Ontology (GO) is a set of three orthogonal ontologies describing the Cellular Components, Biological Processes, and Molecular Functions of gene products (Table 2). These ontologies are produced and utilized by members of the Gene Ontology Consortium (<http://www.geneontology.org>), including ZFIN. Terms from these ontologies are associated with gene records in ZFIN as a routine part of literature curation, as well as by automated computational methods.

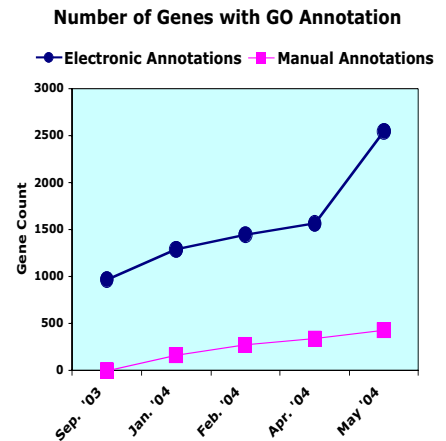
Where are the GO annotations in ZFIN?

GO annotations are located in the Gene Products section of gene pages in ZFIN (Figure 4A). The total number of GO annotations applied to a gene is listed below the GO categories. The "All GO Terms" hyperlink brings up the GO details page, where all the GO terms and their details can be reviewed (Figure 4B). The GO annotation details include evidence codes, inference data and publications. Evidence codes indicate what type of experimental support exists for an annotation. For example, the IEA (Inferred from Electronic Annotation) evidence code indicates that the annotation was produced through automated electronic methods, and has not been reviewed

by a curator. All other evidence codes indicate that an annotation has been reviewed by a curator and reflect the nature and quality of the data backing that annotation. To learn more about GO annotation, click on an evidence code from within ZFIN, or visit the Gene Ontology Consortium website.

How can I make GO-based gene queries?

Annotation of zebrafish gene records with GO terms is now an integral part of the ZFIN literature curation effort. Many uncharacterized genes, such as those solely represented by a cDNA sequence, also get GO annotation through our electronic annotation methods. Thus, GO annotation is an exciting method for describing both



characterized and novel genes. GO-based queries for zebrafish genes can be made using AmiGO (<http://www.godatabase.org/cgi-bin/amigo/go.cgi>), a web based query tool developed by the GO consortium. 🐟

Table 2: Orthogonal components of Gene Ontology (GO).

Ontology	Content	Example
Molecular Function	The tasks performed by individual gene products	<ul style="list-style-type: none"> carbohydrate binding ATPase activity protein kinase activity
Biological Process	Broad biological goals generally accomplished by ordered assemblies of Molecular Functions	<ul style="list-style-type: none"> mitosis retinoid metabolism gastrulation
Cellular Component	Subcellular structures, locations and macromolecular complexes	<ul style="list-style-type: none"> nucleus telomere cap complex connexon complex

A. Gene Ontology

Ontology	GO Term
Molecular Function	DNA binding (more)
Biological Process	positive regulation of transcription, DNA-dependent (more)
Cellular Component	nucleus
All GO Terms (9)	

B. GO Details

Gene Name: *SRY-box containing gene 32*
 Gene Symbol: [sox32](#) Your Input Welcome

Ontology	GO Term	Evidence	Inferred From	Reference(s)
Molecular Function	DNA binding	IDA		1
	DNA binding	IEA		1
	transcription factor activity	IDA		1
	transcription factor activity	IGI	pou5f1	1
Biological Process	determination of left/right symmetry	IMP	cas1a56	1
	embryonic heart tube development	IMP	cas1a56	1
	endoderm cell fate determination	IMP		1
	endoderm development	TAS		1
	positive regulation of transcription, DNA-dependent	IDA		1
	positive regulation of transcription, DNA-dependent	IGI	pou5f1	1
	regulation of transcription, DNA-dependent	IEA		1
Cellular Component	nucleus	IC	GO:0003677	1

Figure 4: Example of GO annotations in ZFIN gene pages.