# Module 1: Where do I start?

# ii. ZFIN, the zebrafish model organism database

#### <u>Aims</u>

- Introduce ZFIN web site
- Suggest starting points for different types of queries
- Describe methods for contacting ZFIN

#### Introduction

ZFIN is the zebrafish model organism database and as such we work towards integrating zebrafish biology with its genome. Data are updated daily by the ZFIN curatorial staff who extract relevant information from your publications. Large datasets submitted by zebrafish labs, routine data exchanges with organizations such as NCBI, the Sanger Institute and Swiss-Prot, as well as data submissions from individual investigators provide additional data for ZFIN. ZFIN also participates with the Sanger Institute in annotation of the whole genome sequence.

#### **ZFIN Home Page**

Query forms for mutants, genes, markers, clones, mapping and gene expression data facilitate integrated analysis of these data. These forms are available from the ZFIN home page, <a href="http://zfin.org">http://zfin.org</a>,

As you scroll down the ZFIN home page, take note of the links to query forms for:

- Mutants/Transgenes
   Search by name, map location, mutation type or phenotype
- Genes/Markers/Clones
   Search by name, accession number, LG, vector or sequence type
- Gene Expression
  Search by gene name, mutant background, author, developmental stage, anatomical structure or knockdown reagent gene name.
- BLAST
  Search for sequence alignment against ZFIN and zebrafish datasets
- Genetic Map
   Generate an integrated view of mapping panels or a consolidated map.
- Anatomy

Search the zebrafish anatomical dictionary.

#### Genes

- **Function**
- **Protein Families and** Domains
- **Probes**
- **Expression Patterns**
- **Orthologs**
- **Sequence Data**
- **Mapping Data**

General Information

**Mutants and Transgenes** 

Positions at ZFIN About ZFIN Citing ZFIN Help Resources Contact Us Site News Glossary **Download Data** 

**Transgenics** Wild-Type Stocks

utants /

Search for mutations / transgenic lines by gene name, map location or phenotype

rafish wild-type lines.

Genes / Markers / Clones

Gene Expression

Search for genes, markers and clon type or sequence type.

on pattern Search for gene structure, developmenta

Search for sequence alignment

Genomics

Zebrafish Ger Resource

BLAST

**Publications** 

Anatomy

People

Search for sequence alignment against ZFIN datasets and Zebrafish datasets.

**Genetic Maps** Generate graphical views of genetic, radiation hybrid or consolidated maps.

mary listing of zebrafish mapping panels.

**Gene Expression Patterns** 

> Meetings / Jobs Grants The Zebrafish Book The Zebrafish Science Monitor Zebrafish Newsgr Zebrafish for K-ZFIN Newslet

**Mapping Panels** Accession #

Search ZFIN by data accession nur

Search for zebrafish research public

Search the zebrafish anatomical on

**Graphical representation of** mapping panels

Search for zebrafish researchers by name or address.

Laboratories arch for laboratories by name, address or research interests.

Companies 4 es supplying zebrafish reagents. Search for com

ratory llele Designations Nomenclature Conventions

Login:

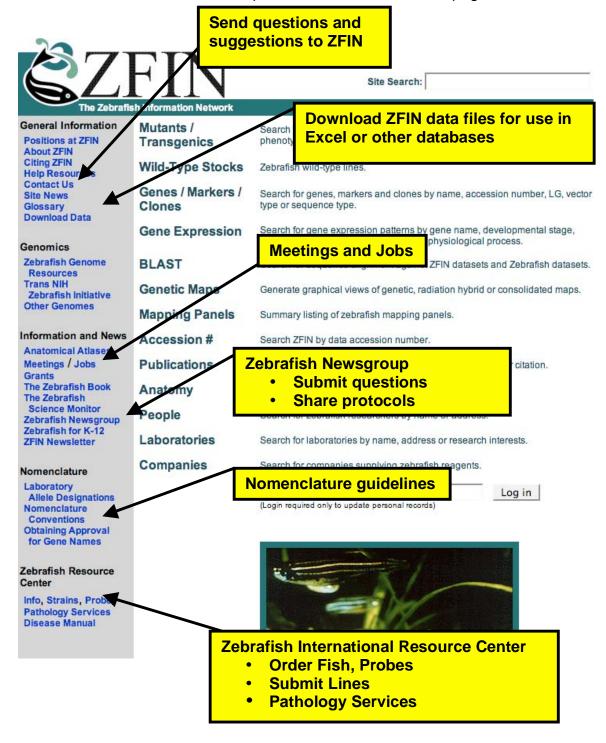
Passwor (Login required only to update personal records) Log in

**Anatomical Dictionary with** links to associated gene expression patterns

Pathology Services Disease Manual

**Zebrafish Research Community Contact Information** 

Links to additional resources are provided on the ZFIN home page side bar.



#### **Site Search**

Site Search is a quick way to search the entire ZFIN website. It's easy – you'll find the search box in the upper right hand corner of every ZFIN page - just enter words in the search box and press return.

The features of site search are shown in this example search for **cartilage**. The top of the results page displays a box where results are sorted into several different data categories. Categories help you narrow your search. The currently-selected category is marked with an arrowhead (). The number beside each category is the number of pages found. Look inside the categories box. Next to "All", you can see that there are 353 ZFIN pages containing the word "cartilage". 28 of these are Mutant/Transgenic pages.



Click the Mutants/Transgenics category to find mutant fish pages containing the word "cartilage". The *dirty south* locus page is in this group because the word "cartilage" is contained in the phenotype description.

#### Locus: dirty south

... some ventral **cartilages** reduced or absent), underdeveloped liver/gut. CITATIONS (1) Home Email ZFIN About ZFIN Helpful Hints ...

/cgi-bin\_almost/webdriver?Mlval=aa-locusview.apg&OID=ZDB-LOCUS-040927-5

The Gene Ontology (GO) category is a good place to look for genes related to cartilage. Here you see that sox9b has been annotated with the GO term 'cartilage development'. Follow the GO details link for sox9b to learn that this annotation was inferred from deficiency mutant phenotypes and morpholino experiments.

#### GO Details: sox9b

... s) Molecular Function DNA binding IEA 1 Biological Process cartilage development IMP MO3-sox9b MO2-sox9b 1 cartilage development IMP Df(LG03)sox8,sox9b b971 1 embryonic pectoral fin morphogenesis ...

**Cartilage** appears many times in the zebrafish anatomical ontology. Follow the links in the Anatomy category to view definitions, possible probes and expressed genes.

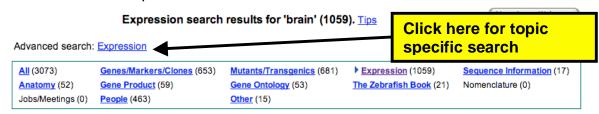
#### Anatomical Structure: ceratobranchial

... Definition: Ceratobranchials are bilaterally paired **cartilage** bones that form part of the ventral ... with the epibranchials. Ceratobranchials 1-5 ossify in the ceratobranchial **cartilages**. Appears ... //gi-bin\_almost/webdriver?Mlval=aa-anatomy\_item.apg&OID=ZDB-ANAT-011113-411

You might even find a cartilaginous job with Site Search – check the Jobs/Meetings category.

#### Site Search vs. Topic Specific Search

Site Search complements ZFIN's topic specific searches but it is important to understand the differences. Site search looks for word matches on ZFIN web pages. It's fast and it covers the entire database, but it's not very smart. The domain-specific searches *are* smart – they are designed take advantage of domain-specific associations between terms. For example, if you enter "brain" in the anatomy field of the domain-specific Gene Expression search page, your results will include all figures with cerebellar gene expression. Gene Expression search "knows" that the cerebellum is part of the brain. In contrast, the expression category of a Site Search for "brain" will return figures only if the word "brain" appears in figure captions, gene names or anatomical structures, as shown in the example below:



Exact Match: brain

#### Figure: Chou et al., 2006, Fig. 2

... Expression and characterization of a **brain-specific** protein kinase BSK146 from zebrafish. Biochem ... Anatomical Terms: **brain**, embryo, unspecified Stage Range: 5-9 somites to Adult Fig. 2 ZFIN is incorporating ... Not detected embryo RTPCR Prim-5 to Day 6 unspecified RTPCR Adult **brain** RTPCR Home Email ZFIN About ZFIN Helpful Hints ...

/cgi-bin\_almost/webdriver?Mlval=aa-fxfigureview.apg&OID=ZDB-FIG-060313-4

For a more detailed search, follow the 'Advanced search: Expression' link to the ZFIN topic specific gene expression search. As a convenience, "brain" will be automatically entered into the anatomy field of the expression search form.

#### **Synonyms**

When a synonym or previous name is entered into Site Search, you are provided with an Alternative search suggestion link. Click this link to start a fresh Site Search using the primary name of the term. This is worth a try, because it will typically provide you with overlapping and complementary results. Examples below:

*pou2* is a previous name for *pou5f1*. Site search for *pou2* produces *pou2* results and an optional alternative search for *pou5f1*:



Follow the alternate search link to find matches to *pou5f1*.

#### Search results for 'pou5f1' (38). Tips

Your Input Welcome

▶ <u>All</u> (38)	Genes/Markers/Clones (5)	Mutants/Transgenics (11)	Expression (13)	Sequence Information (1)
Anatomy (0)	Gene Product (1)	Gene Ontology (2)	The Zebrafish Book (0)	Nomenclature (0)
Jobs/Meetings (0)	People (5)	Other (0)		

#### Exact Match: pou5f1

#### Gene: pou5f1

... factor 1 Gene Symbol: pou5f1 Previous Names: chunp6868; pou2 ... KNOCKDOWNS: Mutant locus: spiel ohne grenzen (spg) has been shown to correspond to gene pou5f1 ... SEGMENT (CLONE AND PROBE) RELATIONSHIPS: pou5f1 Encodes [EST] cb197 (order this), fd18d06 ... /ogi-bin\_almost/webdriver?Mlval=aa-markerview.apg&OID=ZDB-GENE-980528-485

Similarly, entering a previous/alternate name for an anatomy term ("isthmus") will provide an alternative search using the primary anatomical structure name.

Search results for 'isthmus' (30). Tips

Your Input Welcome

Alternative search: midbrain hindbrain boundary (isthmus)

▶ <u>All</u> (30)	Genes/Markers/Clones (1)	Mutants/Transgenics (10)	Expression (9)	Sequence Information (0)
Anatomy (1)	Gene Product (0)	Gene Ontology (0)	The Zebrafish Book (0)	Nomenclature (0)
Jobs/Meetings (0)	People (9)	Other (0)		

#### Gene: pax2a

... Nomenclature History MUTANTS AND TARGETED KNOCKDOWNS: Mutant locus: no isthmus ... /ogi-bin\_almost/webdriver?Mlval=aa-markerview.apg&OID=ZDB-GENE-990415-8

#### Allele: b593 (no isthmus)

... ZDB-FISH-011017-10 Name: no isthmus Abbreviation: noi b593 Previous names: AFFECTED GENE: paired box gene 2a (pax2a) has been shown to correspond to locus no isthmus. (2) Images ... /cgi-bin\_almost/webdriver?Mival=aa-fishview.apg&0ID=ZDB-FISH-011017-10

#### Contacting ZFIN

ZFIN is your database. We welcome your comments and suggestions. Please use the **Contact Us** on our home page or the **Your Input Welcome** button provided on every data page to contact us.

# Module 2: Maps and Genome Sequence iii. ZFIN Map Viewer

#### <u>Aims</u>

- Introduce ZFIN Map Viewer
- Suggest starting points for various queries

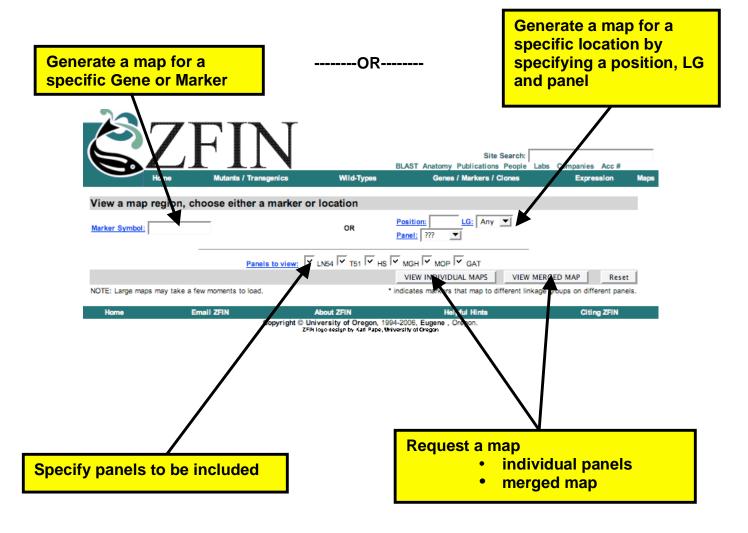
#### <u>Introduction</u>

ZFIN's map viewer provides an integrated view of the zebrafish mapping panels (HS, MGH, T51, LN54, GAT, MOP) to assist with identification of candidate loci and conserved chromosome segments. An integrated map, designed by John Postlethwait et al. is also available. <a href="http://zfin.org/cgi-bin/webdriver?Mlval=aa-crossview.apg&OID=ZDB-REFCROSS-010114-1">http://zfin.org/cgi-bin/webdriver?Mlval=aa-crossview.apg&OID=ZDB-REFCROSS-010114-1</a>.

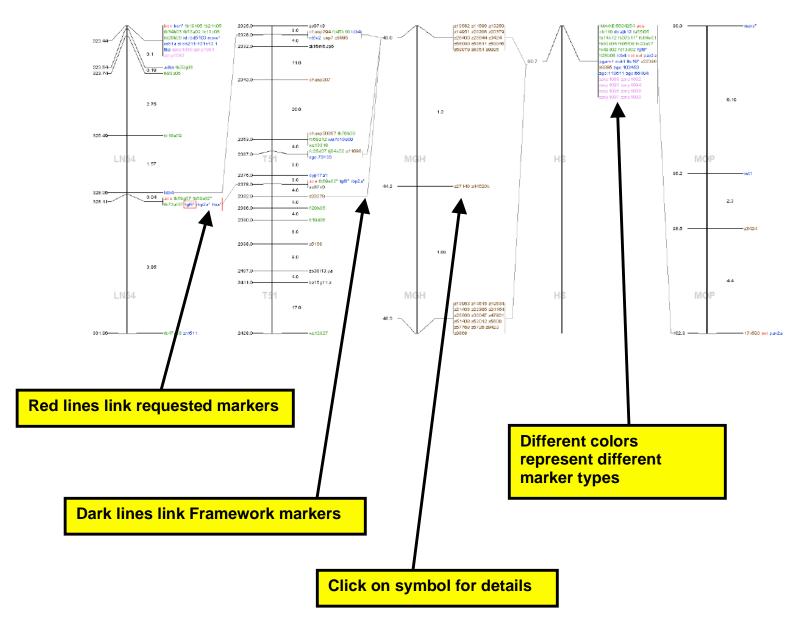
Framework markers are used to link panels. Genes, ESTs, anonymous markers and SNPs are displayed. Detailed information for a gene or marker may be found by clicking on the symbols displayed on the map.

Maps may be generated by specifying a gene or marker name or by specifying a location on a linkage group.

To generate an integrated view of mapping panel data, click on the 'Genetic Maps' link present on ZFIN's home page.



# Comparison of LN54, HS and T51 panels for marker fgf8



#### View of Merged (Consolidated) Map for fgf8



#### **Exercises**

- 1. Use the ZFIN Map Viewer to generate a map for a specified position on a linkage group.
- 2. Use the ZFIN Map Viewer to generate a merged (consolidated) map for a gene of your choice.

# Module 2: Maps and Genome Sequences

i. How do I find a zebrafish gene in the genome?

#### <u>Aims</u>

Introduce ZFIN BLAST.

#### <u>Introduction</u>

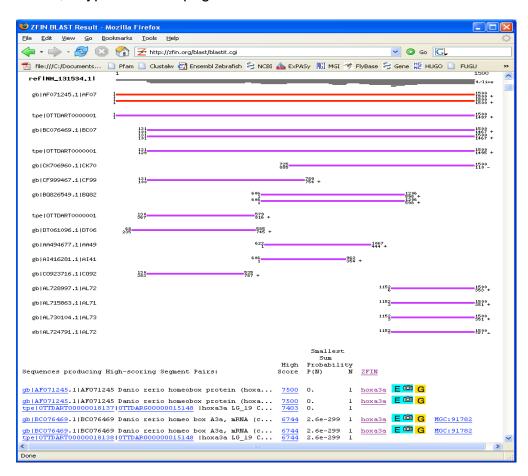
Sequence comparison, BLAST, may be used to locate the position of a gene on the genome. The ZFIN BLAST can be accessed from the BLAST link on the home page or from the navigation bar. The BLAST search page is shown below:

BLAST	Your Input Welcome		
Choose program and database:  GenBank Zebrafish GenBank Human GenBank Mouse RefSeq Zebrafish mRNA ZFIN GenBank Sequences			
Program: Nucleotide - Nucleotide  Database: ZFIN cDNA Sequences ZFIN Genes with Expressio ZFIN Morpholino Sequence ZFIN MicroRNA Sequence ZFIN Vega Transcripts Zebrafish mRNA	98 8		
Query sequence (maximum of 50,000 letters):	— GenBank Zebrafish		
FASTA or free-text format	RefSeq Zebrafish mRNA		
Set subsequence: From To To Search for short, nearly exact matches  Sequence ID: (one or multiple delimited by ",")	ZFIN GenBank Sequences ZFIN cDNA Sequences ZFIN Genes with Expression ZFIN Morpholino Sequences ZFIN microRNA Sequences ZFIN Vega Transcripts Zebrafish mRNA EST Zebrafish		
Sequence Type: Nucleotide 🗑			
Upload a free-text file: Browse  Clear sequence BLAST	Ensembl Zebrafish Transcripts TIGR Zebrafish Clusters Zebrafish DNA HTG Zebrafish		
Options:  Expect: 1e-25	GSS Zebrafish Zebrafish Trace Archive WGS Zebrafish RefSeq Zebrafish Protein UniProt / TrEMBL Zebrafish		
□ XNU - filter short-periodicity repeats  Format:  Show:    ☑ Graphical Overview limit of the first 50 alignments	Omi 10t/ HEMDE Zebiansii		

As shown in the inset box above, ZFIN BLAST provides a variety of sequence datasets to compare with your search sequence. These include GenBank zebrafish, curated Vega transcripts identified by the zebrafish genome project at Sanger, Ensembl transcripts and TIGR zebrafish clusters. Multiple databases can be selected. Simply hold down the shift key while selecting the databases. You can readily optimize the BLAST search parameters for short, nearly exact matches. This is particularly useful when searching for morpholinos.

The main advantage of ZFIN BLAST is tight integration with the ZFIN database via the five ZFIN specific data sets: ZFIN GenBank sequences, ZFIN cDNA sequences, ZFIN morpholino sequences, ZFIN microRNA sequences and ZFIN Vega transcripts.

Below, a typical results page:



Because ZFIN BLAST is integrated with the ZFIN database, you can navigate directly from BLAST results to ZFIN gene and clone pages. Genes with related expression and Gene Ontology data are indicated through the use of E and G icons respectively. The camera icons are used for instances where ZFIN has been able to provide a figure containing expression data. Similarly, you can view

gene and cluster records at Vega, Ensembl or TIGR for queries involving these resources.

The ZFIN BLAST server is a shared resource. To optimize performance for all, the following strategies have been implemented:

- WU BLAST program running on a dedicated multi-processor server
- query length up to 50,000 letters
- graphical display available for the first 50 alignments
- single queries for searching the zebrafish trace archive
- batch queries of up to 100 sequences for small to medium zebrafish datasets

\_\_\_\_\_

# Module 3: Genes and Sequences ii. How do I find information about a gene?

#### <u>Aims</u>

- Introduce the ZFIN Genes/Markers/Clones query form
- Suggest starting points for various queries
- Describe information that can be found pertaining to genes

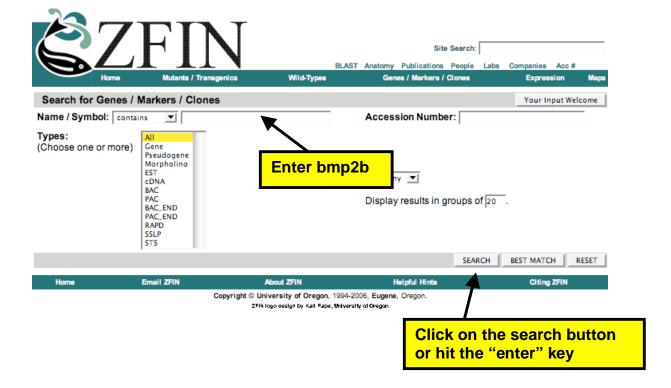
#### **Introduction**

ZFIN integrates data for named, published genes as well as unnamed genes identified by the Sanger Institute Vega project, by the zebrafish gene collection project and by various gene expression screens.

These data may be accessed using the Genes/Markers/Clones query form. A link to this form is available on ZFIN's home page. You may search using the approved name or symbol, other names that have been used for a gene or approved names of orthologs. GenBank accession number searches and LG searches are also possible.

The search results summary will display links to corresponding gene page(s) as well as links to associated BACs, PACs, ESTs and morpholinos.

### **Genes/Markers/Clones Query Form**



Enter bmp2b in the name field and click on the 'search' button. The following results summary is displayed.



Sile	Searc	III .				
		_	_	_	_	

BLAST Anatomy Publications People Labs Companies Acc# Wild-Types Genes / Markers / Clones Expression Markers

Search Results for: name contains 'bmp2b'

Your Input Welcome



- 1 Gene (1 with known mutant loci)
  - 2 Morpholinos
  - 1 BAC

Modify you	r searc	h.			
Name / Symbol:	contains bmp2b	•		Accession Numb	per:
Types: (Choose one o	or more)	All Gene Pseudogene Morpholino EST cDNA BAC PAC BAC_END PAC_END RAPD SSLP STS		LG: any J Display results in	
Home	Email	ZFIN	About ZFIN	Helpful Hints	Citing ZFIN

Copyright © University of Oregon, 1994-2006, Eugene, Oregon. ZFIN logo design by Kari Pape, University of Oregon

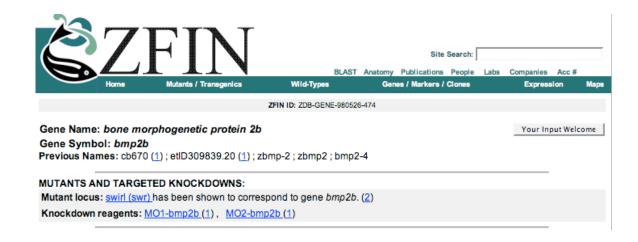
Click on the '1 Gene' link and note the following information as you scroll down the *bmp2b* gene page:

#### **Nomenclature**

ZFIN curators administer established nomenclature guidelines, <a href="http://zfin.org/zf\_info/nomen.html">http://zfin.org/zf\_info/nomen.html</a>, to ensure unambiguous communication and to support comparisons between species. Names and symbols are consistent with mammalian orthologs whenever possible. The 'nomenclature history' link provides a chronology of name changes, splits and merges.

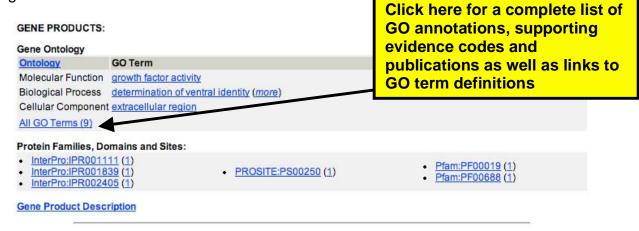
#### **Mutants and Targeted Knockdowns**

Links are provided to data for associated mutants and knockdown reagents and to the contributing publications.



#### **Gene Products**

Gene product information is displayed as you continue down the *bmp2b* gene page.



Gene ontology or GO terms provide insight into gene products and functions. As members of the gene ontology consortium, <a href="http://www.geneontology.org/">http://www.geneontology.org/</a>, we help to define terms used to describe the three GO ontologies - molecular function, cellular components and biological processes. These terms apply to all organisms and are valuable for direct comparisons across organisms. The <a href="https://www.geneontology.org/">bmp2b</a> gene page shows a representative term for each of the three ontologies. To view all terms click the 'All GO terms' link. The <a href="https://www.geneontology.org/">bmp2b</a> GO details page displays all annotations for <a href="https://www.geneontology.org/">bmp2b</a>.

Manual curation and a collaboration with Uni-Prot allow us to provide links to protein family and domain sites and to provide a gene product description.

#### **Gene Expression**

As you continue to scroll down *the bmp2b* gene page you will find links to gene expression data. Follow these links to view the expression data.

GENE EXPRESSION:(current status)

All expression data: 42 figure(s) from 33 publications

Directly submitted expression data: 6 figure(s) (47 images) from Thisse et al., 2001 [cb670]

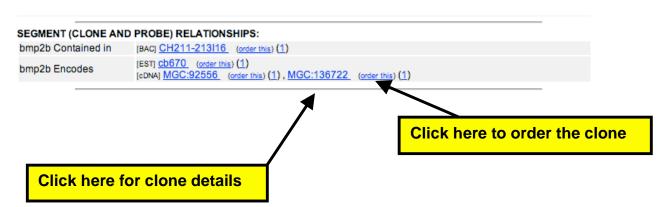
ZFIN displays three kinds of gene expression data:

- annotated images that have been directly submitted to ZFIN by researchers
- annotated data from the current literature, figures and legends are included when copyright permissions are available.
- an index of gene expression data from older publications

ZFIN began to include published figures in 2004. We are currently able to add figures from older publications only on an ad hoc basis. A more complete incorporation of figures from the older literature is a long-term goal.

#### Segment (clone and probe) Relationships

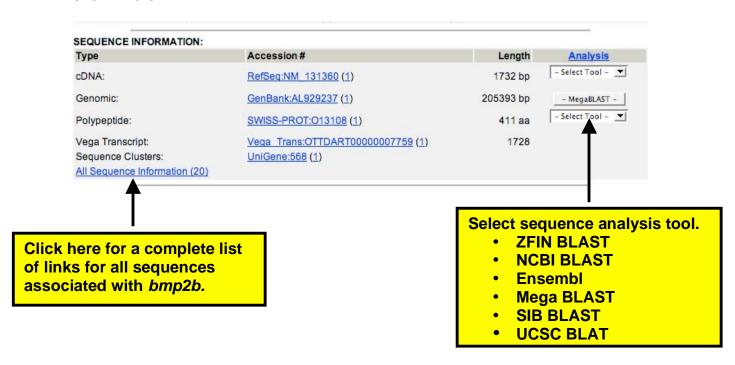
The next section you will find on the *bmp2b* page describes related molecular segments.



Here we see the BAC used by the Sanger Institute for the annotation of *bmp2b*, an EST for the Thisse expression screen (cb670) encoded by *bmp2b* and two full length cDNAs from the zebrafish gene collection (MGC:92556, MGC:136722) that is encoded by *bmp2b*. The 'Order this' link takes you to a site where you may order this probe. Links to the reference or publication that defines these relationships are provided (number in parentheses).

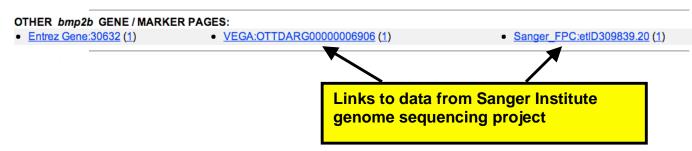
#### **Sequence Information**

Sequence information appears next on the *bmp2b* page. Manual curation and collaborations with NCBI, the Sanger Institute and Uni-Prot allow us to maintain an extensive set of links to sequence information. Representative links are shown here.



#### Other bmp2b gene pages

As you continue down the *bmp2b* page you see links to *bmp2b* gene pages at NCBI Entrez Gene, the Sanger Institute's Vertebrate Genome Annotation database (VEGA) and the marker report from the Sanger Institute's fingerprinting map (Fingerprint Contig or FPC) of the zebrafish genome.



#### Mapping Information

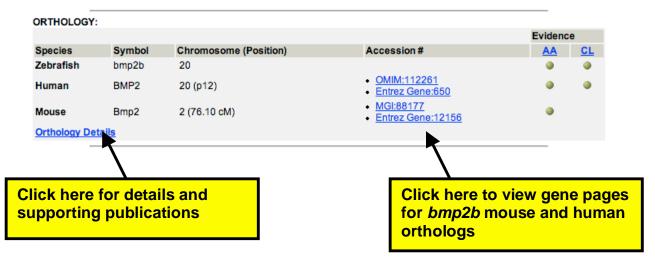
The next item on the *bmp2b* gene page is mapping information.



bmp2b has been mapped to LG 20. The 'Details' link will take you to detailed mapping information such as mapping panel, location and scoring data <a href="http://zfin.org/cgi-bin/webdriver?Mlval=aa-mappingdetail.apg&OID=ZDB-GENE-980526-474">http://zfin.org/cgi-bin/webdriver?Mlval=aa-mappingdetail.apg&OID=ZDB-GENE-980526-474</a>. Mapping information for related markers and clones is also available on the details page. Links to graphical map views are provided.

#### <u>Orthology</u>

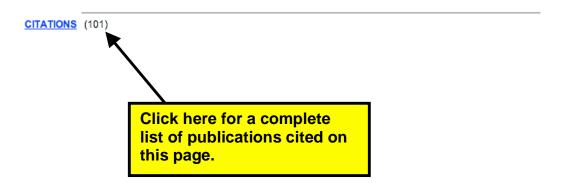
ZFIN curators capture and record orthology data for human, mouse and Drosophila genes. Approved nomenclature symbols and links to the gene page for other species are provided. Continuing down the bmp2b page you find orthology information. Ortholog chromosome (position) provides a quick glance at synteny.



Please see Module 3 section ii for additional details pertaining to orthologs.

#### **Citations**

The final item we see on the *bmp2b* gene page is a link to citations. Clicking on this link we find a listing of all publications that contributed data to the ZFIN *bmp2b* gene page. Publications discussing *bmp2b* are listed here. Data that has been entered through collaborations are cited here.



# **Exercises**

- Have any fgf8 knockdowns been studied?
  How many published studies of fgf8 gene expression have been curated at ZFIN?
- Have any large scale in situ screens provided images for fgf8 gene expression patterns?
- Are mouse or human orthologs known for fgf8?

# Module 3: Genes and Sequences

iii. Does my gene have known homologues/orthologs?

#### Aims

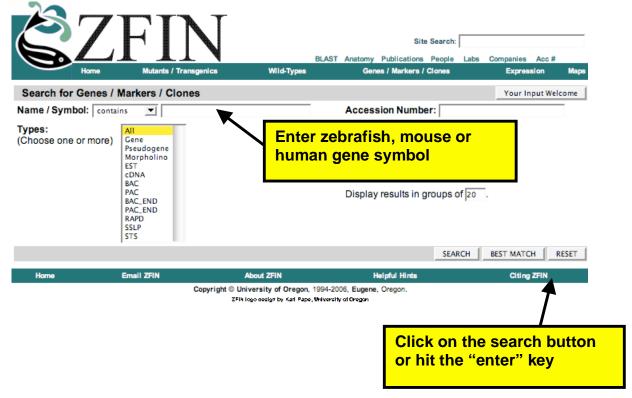
- Provide overview of ortholog curation at ZFIN
- Suggest starting points for finding ortholog data

#### Introduction

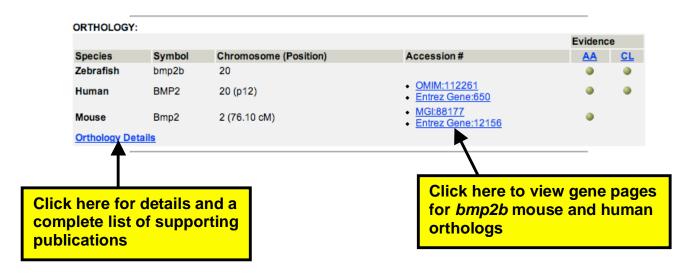
To facilitate an understanding of relationships between gene and gene functions in zebrafish and other organisms, ZFIN curators capture orthology data for human, mouse and Drosophila genes from zebrafish literature. Approved nomenclature symbols, chromosome(position) and links to the gene page for other species are provided from ZFIN gene pages. The evidence supporting the assertion as well as the source is provided. Definitions of the various evidence codes can be found by clicking on the code itself. <a href="http://zfin.org/zf\_info/oev.html">http://zfin.org/zf\_info/oev.html</a>.

Our curators also identify orthologs through their extensive analysis. These data are entered and attributed accordingly.

Use the ZFIN Genes/Markers/Clones query form, <a href="http://zfin.org/cgi-bin/webdriver?Mlval=aa-newmrkrselect.apg">http://zfin.org/cgi-bin/webdriver?Mlval=aa-newmrkrselect.apg</a> to locate orthology data for a gene. This form supports the use of approved symbols and names from other organisms.



Scroll to the **Orthologs** section of the resulting gene page.



In this case links to Entrez Gene, OMIM and the mouse model organism database, MGI, are provided. The chromosome (position) is provided for each ortholog. The evidence supporting the assertion and the source for the assertion are provided. Definitions of the various evidence codes can be found by clicking on the code itself. <a href="http://zfin.org/zf\_info/oev.html">http://zfin.org/zf\_info/oev.html</a> This example shows that amino acid sequence homology and conserved location have identified a human

*bmp2b* ortholog. A mouse ortholog has been identified via amino acid sequence homology and synteny.

## **Exercises**

• What type of evidence exists to support human and mouse ortholgues of fqf8?

# Module 3: Genes and Sequences

# iv. How can I find cDNA clones for my gene?

#### Aims

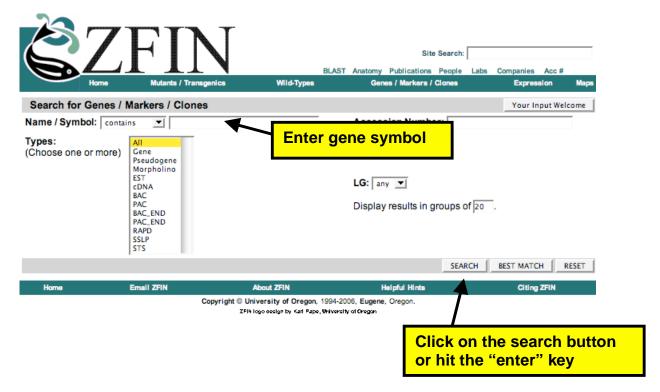
- · Describe curation of clones and probes at ZFIN
- Suggest starting points for finding cDNA clones and probes
- Describe how to find a source for the reagent.

#### Introduction

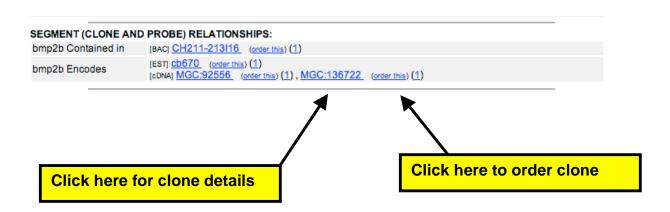
ZFIN incorporates information about full length cDNA clones from the Zebrafish Gene Collection (ZGC, <a href="http://zgc.nci.nih.gov/">http://zgc.nci.nih.gov/</a>), cDNAs used in large scale in situ screens conducted by the Dawid, Talbot and Thisse laboratories as well as from cDNAs cited in the literature. These cDNA sequences are curated and associated with their encoding genes. In some cases, it is not possible to associate the cDNA with a published gene. In these cases, we create novel gene records. In addition, BAC clones used by the Sanger Institute genome sequencing project are captured in ZFIN.

#### Finding cDNA clones

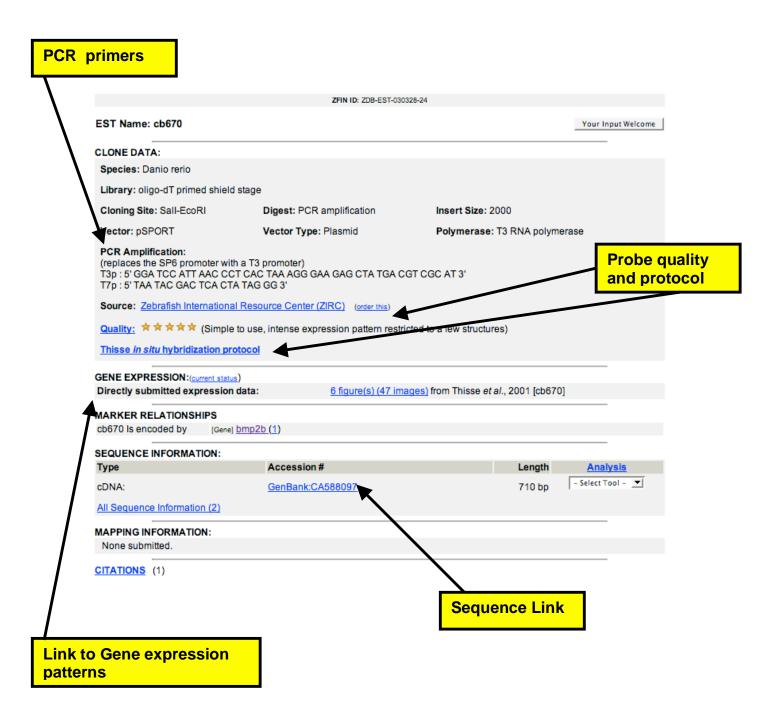
Links to cDNA clones and probe data can be found on ZFIN gene pages. Locate this information using our Genes/Markers/Clones query form, <a href="http://zfin.org/cgi-bin/webdriver?Mlval=aa-newmrkrselect.apg">http://zfin.org/cgi-bin/webdriver?Mlval=aa-newmrkrselect.apg</a>. Search by specifying your gene of interest or accession number.



Scroll to the **Segment (Clone and Probe) Relationship** section of the gene page.



Click on the cb670 link to view details about this clone.



The Christine and Bernard Thisse have ranked probes from their large scale insitu screens according to expression pattern quality. Probes with the highest ranking have the potential to be used as specific probes for an anatomical structure. These are also listed on the appropriate ZFIN anatomy page as High Quality Probes. Descriptions of the star ratings follow, from 5 start to 1 star.

- Simple to use, intense expression pattern restricted to a few structures.
- Nice strong expression pattern.
- Moderate expression pattern.
- Weak expression pattern.
- Probe is difficult to use. Generally basal level of expression with more intense labeling in particular structure.

## Exercises

- Locate probes or clones that could be useful for studies of *fgf8*.Can you find a way to order these reagents?

# Module 4: Function and Expression i. How do I determine gene function?

#### <u>Aims</u>

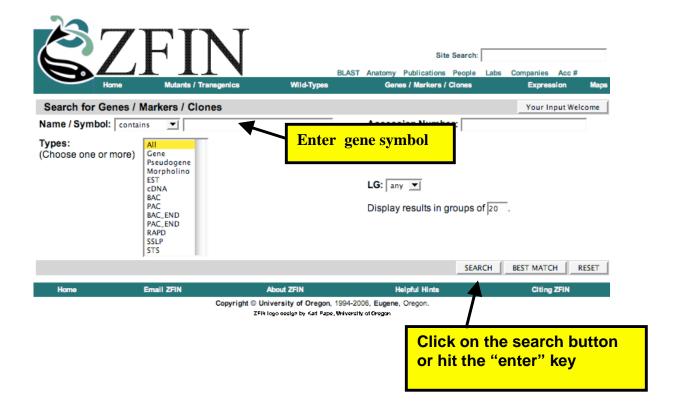
- Introduce gene ontology and gene expression data available in ZFIN
- · Describe how to find these data

#### Introduction

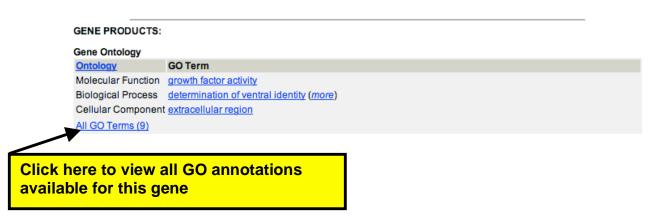
Gene Ontology (GO, <a href="http://www.geneontology.org/">http://www.geneontology.org/</a>) annotations and gene expression data offer powerful insights for understanding of biological processes and gene function. GO terms are associated with genes by literature curation and by automated computational means. These terms are defined by members of the GO consortium, including ZFIN, to describe biological processes, cellular components and molecular functions of gene products. These terms can be used to describe gene products in any organism thus promoting cross-species studies.

#### **Finding GO annotations**

GO annotations are located in the **Gene Products** section of a ZFIN gene page. Locate this information using our Genes/Markers/Probes query form, <a href="http://zfin.org/cgi-bin/webdriver?Mlval=aa-newmrkrselect.apg">http://zfin.org/cgi-bin/webdriver?Mlval=aa-newmrkrselect.apg</a>. Search by specifying your gene of interest.



Scroll down to the **Gene Products** section of the gene page.



The *bmp2b* gene page shows a representative term for each of the three ontologies. To view all terms click on 'All GO terms' link. The *bmp2b* GO details page displays all annotations for *bmp2b*.

	ZFIN ID: ZDB-GENE-9	80526-474		
GO Details Gene Name: bone mo Gene Symbol: bmp2b	orphogenetic protein 2b		_ Y	our Input Welcome
Ontology	GO Term	Evidence	Inferred From	Reference(s)
Molecular Function	growth factor activity	<u>IEA</u>		<u>2</u>
Biological Process	cell migration during gastrulation	<u>IGI</u>	SWrtc300a MO6-wnt8a MO5-wnt8a MO4-wnt8a MO3-wnt8a	1
	cell-cell signaling during cell fate commitme	ent IMP		<u>1</u>
	determination of ventral identity	<u>IDA</u>		1
	determination of ventral identity	<u>IGI</u>	twsg1b	1
	determination of ventral identity	<u>IGI</u>	MO2-bmp7 MO2-bmp2b	1
	determination of ventral identity	<u>IMP</u>	MO2-bmp2b	1
	determination of ventral identity	TAS		1
	dorsal/ventral pattern formation	<u>IEP</u>		1
	dorsal/ventral pattern formation	<u>IGI</u>	MO6-wnt8a MO5-wnt8a MO4-wnt8a MO3-wnt8a	1
	dorsal/ventral pattern formation	<u>IMP</u>	swrtc300a	1
	growth	<u>IEA</u>		1
	mesodermal cell fate commitment	<u>IGI</u>	swrtc300a MO6-wnt8a MO5-wnt8a MO4-wnt8a MO3-wnt8a	1
	notochord development	IGI	MO6-wnt8a MO5-wnt8a MO4-wnt8a MO3-wnt8a	1
Cellular Component	extracellular region	<u>IEA</u>		1
ere for GO te	Click here for Code definition	O evidence		nere for pu ling suppo

You may get additional information regarding a term by following the term link. Evidence codes supporting the annotation are provided. Evidence codes are standardized by the GO consortium and allow you to determine the confidence you may want to have in each GO term association. Again you may click on the evidence code for a description. The reference supporting the annotation is also provided.

## Making a GO based gene query

ZFIN's Site Search may be used for a GO based gene query can. See Module 1 for an example query.

As members of the GO consortium, we routinely make our annotations available to the centralized database maintained by the consortium. This allows you to use the GO term search engine, AmiGO <a href="http://www.godatabase.org/cgi-bin/amigo/go.cgi">http://www.godatabase.org/cgi-bin/amigo/go.cgi</a>, made available by the GO consortium, to search for zebrafish genes and genes of other organisms that are annotated with a specified GO term.

#### Gene expression data

ZFIN also incorporates large datasets of high quality annotated images from laboratories performing large scale *in situ* hybridization screens, gene expression data submitted by individual investigators and gene expression data from the literature. See Module 4 ii for a complete discussion.

#### **Morpholinos**

Morpholinos, antisense oligonucleotides, have become an important method for evaluating gene function in zebrafish. ZFIN curates morpholino data from published literature and is working with the Stephen Ekker laboratory to include data from their morpholino screen. See Module 4 ii for a detailed discussion.

#### **Exercises**

- What are possible ways to infer function for my gene?
- What molecular function has been attributed to fqf8?
- In what biological processes is *fqf8* involved?
- With what cellular components is fqf8 associated?
- What supporting evidence is available?

# Module 4: Function and Expression ii. How can I find gene expression data?

#### <u>Aims</u>

- Introduce gene expression data at ZFIN
- Suggest starting points for various queries
- Discuss morpholino curation at ZFIN

#### Introduction

ZFIN displays three kinds of gene expression data:

- annotated images that have been directly submitted to ZFIN by researchers
- annotated data from the current literature, figures and legends are included when copyright permissions are available.
- an index of gene expression data from older publications

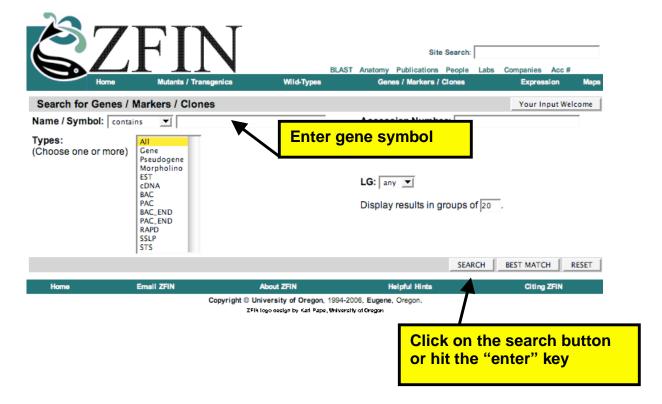
ZFIN began to include published figures in 2004. We are currently able to add figures from older publications only on an ad hoc basis. A more complete incorporation of figures from the older literature is a long-term goal. The zebrafish anatomical dictionary, <a href="http://zfin.org/cgi-bin/webdriver?Mlval=aa-anatdict.apg&mode=search">http://zfin.org/cgi-bin/webdriver?Mlval=aa-anatdict.apg&mode=search</a>, plays a central role in our curation of gene expression data.

#### Finding gene expression data

There are three methods for finding gene expression data in ZFIN.

1. Gene Expression data may be found by following the links provided on a gene page.

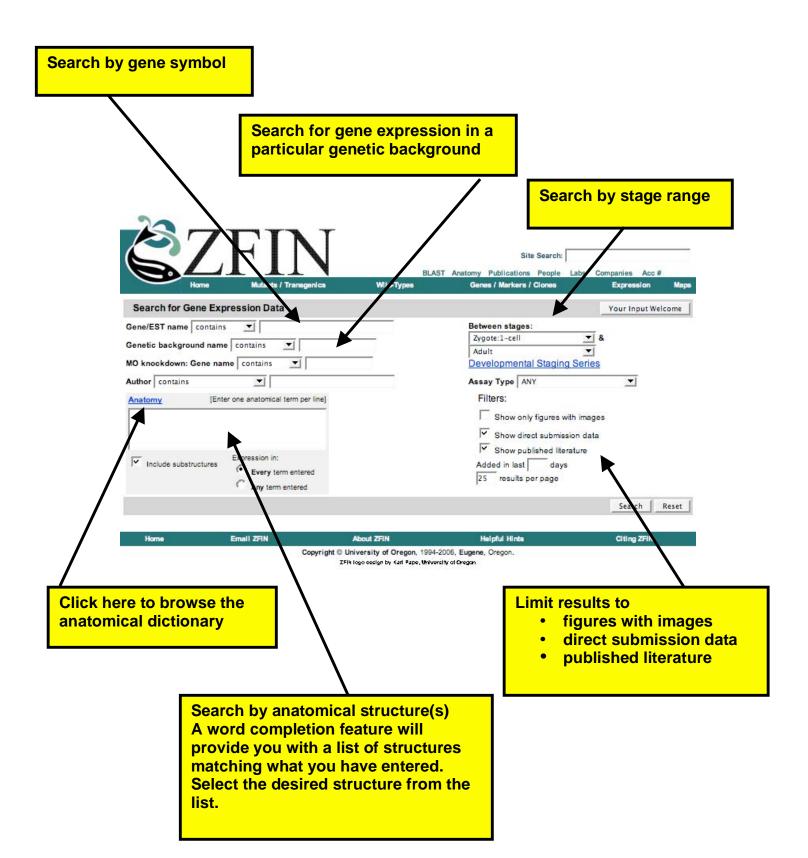
Locate this information using our Genes/Markers/Clones query form, <a href="http://zfin.org/cgi-bin/webdriver?Mlval=aa-newmrkrselect.apg">http://zfin.org/cgi-bin/webdriver?Mlval=aa-newmrkrselect.apg</a>. Search by specifying your gene of interest.



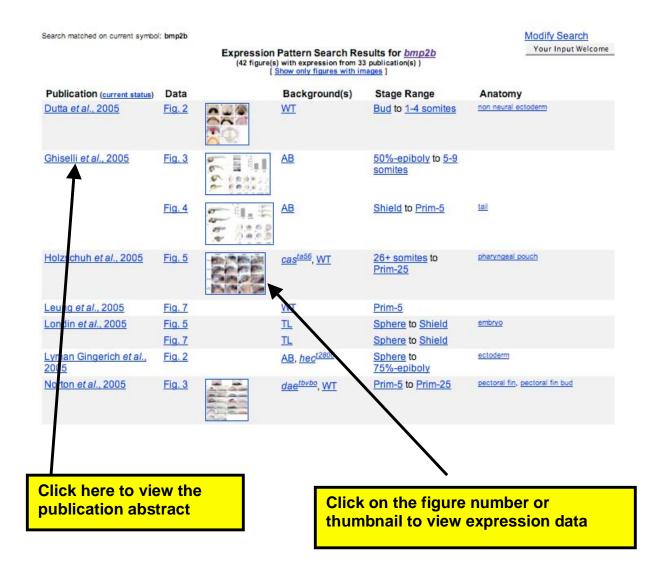
Scroll to the **Gene Expression** section of the gene page.



2. Gene expression data may also be found using the ZFIN expression query form.



A portion of the return results for a *bmp2b* gene expression search is shown below.



Click on the Fig.5 link of the Eivers et al. publication for figure image, legend and annotations.

Click here to view all expression data from this publication

### ZFIN ID: ZDB-FIG-050509-2

Dutta et al., 2005 - pitx3 defines an equivalence domain for lens and anterior pituitary placode. Sevelopment 132(7):1579-1590 Your Input Welcome

ADDITIONAL FIGURES

Genes: bmp2b ▼, ctslb ▼, dlx3b ▼, eya1 ▼, foxb1.2 ▼, pax6a ▼, pitx3 ▼, shha ▼, zic1 ▼

Genetic Background : WT

Anatomical Terms: ectoderm, neural plate, neuroectoderm, non neural ectoderm, polster, prechordal plate

Stage Range: Bud to 1-4 somites

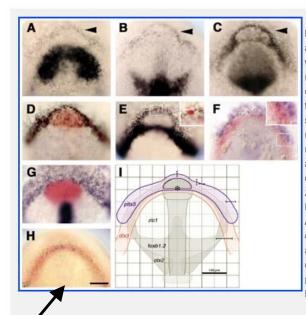


Fig. 2 pitx3 and dlx3b form overlapping expression domains at the anterior neural plate border. (A-D) Embryos labeled with probes for pitx3 (arrowheads) and (A) zic1, (B) foxb1.2; (C) pax6a or (D) hgg1 (red). (E) Anterior neural plate border cells express dlx3b. (E, inset) Single cell injected with fluorescein-dextran, labeled for lineage tracer (red). (F-H) A subset of cells co-expresses: (F) pitx3 (blue) and dlx3b (red); (G) bmp2b (blue, non-neural ectoderm), shh (blue, neural plate midline) and hgg1 (red, polster); (H) dlx3b (red) and eya1 (black). (I) Sketch of anterior neural plate gene expression domains (pitx3, blue, n=12; dlx3b, red, n=15; zic1, foxb1.2, otx2, grey) (Varga et al., 1999) to predict location of pitx3-expressing cells in live unlabeled embryos. Averaged gene expression domains; bars indicate standard deviations from the mean, relative to neural plate midline and center of polster (asterisk). All panels at bud stage, except C, which is at the one-somite stage. Dorsal views of prospective head region, ventral towards the top. Scale bars: 100 µm.

Click here to view a larger image

Continue to scroll down this page. You will find a tabular summary of the expression data discussed in this figure. This table is always available, even when we do not have privileges to display the figure's image and legend.

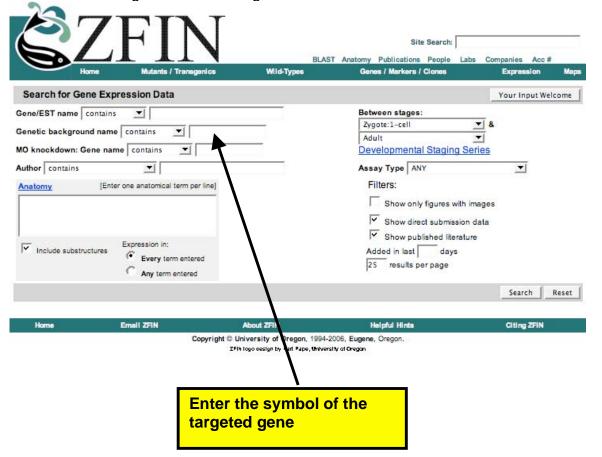
## Gene expression details

Gene		Fish	Stage	Anatomy	<u>Assay</u>
bmp2b	•	<u>WT</u>	<u>Bud</u>	non neural ectoderm	ISH
cts/b	•	WT	Bud	polster	ISH
<u>dlx3b</u>	*	WT	Bud	non neural ectoderm	ISH
eya1	*	WT	Bud	ectoderm	ISH
foxb1.2	*	WT	Bud	neuroectoderm	ISH
pax6a	•	<u>WT</u>	1-4 somites	neuroectoderm	ISH
pitx3	•	<u>WT</u>	Bud	ectoderm	ISH
			1-4 somites	ectoderm	ISH
<u>shha</u>	•	<u>WT</u>	Bud	neural plate	ISH
			Bud	prechordal plate	ISH
zic1	•	<u>WT</u>	Bud	neural plate	ISH
zic1	*	<u>WT</u>			

## Acknowledgments:

ZFIN wishes to thank the journal <u>Development</u> for permission to reproduce figures from this article. Please note that this material may be protected by copyright.

The gene expression query form can also be used to locate expression patterns from studies using knockdown reagents.



# A search for *gata1* targeted knockdowns returns the following:

Gene	Expression Data (current status)	Stage Range	Matching Text
alas2	1 figure(s) from Galloway et al., 2005	20-25 somites	MO symbol: gata1
cahz	1 figure(s) from Galloway et al., 2005	20-25 somites	MO symbol: gata1
<u>cmyb</u>	1 figure(s) from Galloway et al., 2005	20-25 somites	MO symbol: gata1
gata1	2 figure(s) from 2 publications	20-25 somites	MO symbol: gata1
glcci1	1 figure(s) from Galloway et al., 2005	10-13 somites to 26+ somites	MO symbol: gata1
hbae1	1 figure(s) from Rhodes et al., 2005	Prim-15	MO symbol: gata1
hbbe1	3 figure(s) from Galloway et al., 2005	20-25 somites to Prim-5	MO symbol: gata1
kiaa0650 <u>l</u>	1 figure(s) from Galloway et al., 2005	10-13 somites to 26+ somites	MO symbol: gata1
klf4	1 figure(s) from Galloway et al., 2005	10-13 somites to 20-25 somites	MO symbol: gata1
krcp	1 figure(s) from Galloway et al., 2005	10-13 somites	MO symbol: gata1
lcp1	3 figure(s) from 2 publications	Prim-5 to Day 4	MO symbol: gata1
lmo2	1 figure(s) from Rhodes et al., 2005	26+ somites	MO symbol: gata1
<u>трх</u>	3 figure(s) from 2 publications	26+ somites to Day 4	MO symbol: gata1
runx1	2 figure(s) from 2 publications	20-25 somites to 26+ somites	MO symbol: gata1
spi1	3 figure(s) from 2 publications	14-19 somites to Prim-5	MO symbol: gata1
tal1	1 figure(s) from Rhodes et al., 2005	26+ somites	MO symbol: gata1
znfn1a1	1 figure(s) from Galloway et al., 2005	20-25 somites	MO symbol: gata1
CHIN POLIT	N Sulfation Sulf	<u>Lo Lo domino</u>	me dymbon galar

Copyright privileges have not been granted to ZFIN by this journal, however, a detailed table describes the gene expression data presented in this figure. A portion of this table is depicted below.

image not available Fig. 1 ZFIN is incorporating published figure images and captions as part of an ongoing project. Figures from some publications have not yet been curated, or are not available for display because of copyright restrictions.

## Gene expression details

Gene		Fish	Stage	Qualifier	Anatomy	<u>Assay</u>
<u>cmyb</u>		TU, MO:gata2	20-25 somites		intermediate cell mass of mesoderm	ISH
		TU, MO:gata1	20-25 somites		intermediate cell mass of mesoderm	ISH
		TU, MO:gata1,gata2	20-25 somites		intermediate cell mass of mesoderm	ISH
		<u>TU</u>	20-25 somites		intermediate cell mass of mesoderm	ISH
hbbe1	_	TU, MO:gata1	Prim-5		intermediate cell mass of mesoderm	ISH
		<u>TU</u>	Prim-5		intermediate cell mass of mesoderm	ISH
lcp1		TU, MO:gata2	Prim-5		intermediate cell mass of mesoderm	ISH
			Prim-5		unspecified	ISH
		TU, MO:gata1	Prim-5		intermediate cell mass of mesoderm	ISH
			Prim-5		unspecified	ISH
		TU, MO:gata1,gata2	Prim-5		intermediate cell mass of mesoderm	ISH
			Prim-5		unspecified	ISH
		<u>TU</u>	Prim-5		intermediate cell mass of mesoderm	ISH
			Prim-5		unspecified	ISH
<u>mpx</u>		TU, MO:gata2	26+ somites	Not detected	intermediate cell mass of mesoderm	ISH
			26+ somites		unspecified	ISH
		TU, MO:gata1	26+ somites		intermediate cell mass of mesoderm	ISH
			26+ somites		unspecified	ISH
		TU, MO:gata1,gata2	26+ somites		intermediate cell mass of mesoderm	ISH
			26+ somites		unspecified	ISH
		<u>TU</u>	26+ somites	Not detected	intermediate cell mass of mesoderm	ISH
		N N	26+ somites		unspecified	ISH

Site Search:

BLAST Anatomy Publications People Labs Companies Acc#

ZFIN ID: ZDB-MRPHLNO-050208-10

Morpholino Name: MO1-gata1

Genes / Markers / Clones

Your Input Welcome

Previous Names: MO(T)-gata1 (1); gata1 MO (1); Gata1 morpholino (1)

Click here to view morpholino details

Sequence: 5' - CTGCAAGTGTAGTATTGAAGATGTC - 3' BLAST IT |

(Although ZFIN verifies reagent sequence data, we recommend that you conduct independent sequence analysis before ordering any reagent.)

Target Gene: gata1 (1)

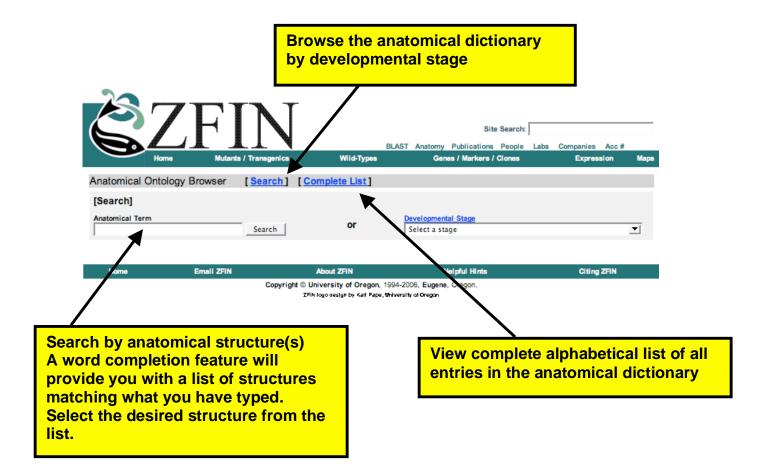
Note: A translation blocking morpholino targeting gata1.

This morpholino sequence was reported with an additional nucleotide in Rhodes et al. 2005 and is correct as displayed here

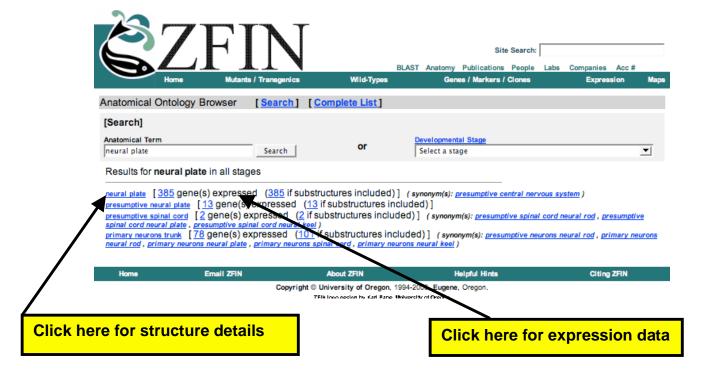
confirmed by author.

CITATIONS (3)

3. Gene expression data may also be found by browsing the anatomical dictionary, <a href="http://zfin.org/cgi-bin/webdriver?Mlval=aa-anatdict.apg&mode=search">http://zfin.org/cgi-bin/webdriver?Mlval=aa-anatdict.apg&mode=search</a>



A search for **neural plate** returns the following list:



Looking at the **neural plate** anatomy page, <a href="http://zfin.org/cgibin/webdriver?Mlval=aa-anatomy\_item.apg&OID=ZDB-ANAT-010921-5060">http://zfin.org/cgibin/webdriver?Mlval=aa-anatomy\_item.apg&OID=ZDB-ANAT-010921-5060</a>, notice synonyms and a definition for neural plate, the stages in which the neural plate is present, links to genes expressed in the neural plate, related structures and a list of publications with neural plate in their abstract.



## **Exercises**

- How can you find expression patterns for your marker in early wildtype development to determine if the observed change is due to misexpression or a delay in development?
- Can you find any knockdown studies for a gene you are studying?
- Use the gene expression query form to find expression data for an anatomical structure during a specified developmental stage range.

# Module 4: Function and Expression iii. How can I find possible molecular markers for an anatomical structure?

# <u>Aims</u>

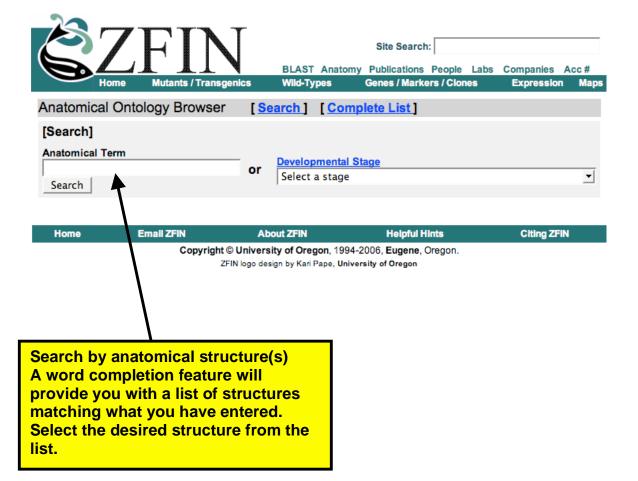
- Describe the ZFIN gene expression search form
- Suggest ways to customize a search for molecular markers for an anatomical structure

# **Introduction**

The zebrafish anatomical dictionary, <a href="http://zfin.org/cgi-bin/webdriver?Mlval=aa-anatdict.apg&mode=search">http://zfin.org/cgi-bin/webdriver?Mlval=aa-anatdict.apg&mode=search</a>, plays a central role in our curation of gene expression data. Associating anatomical structure with gene expression patterns supports queries that can locate possible molecular markers for specific anatomical structures.

# Finding a possible molecular marker

ZFIN integrates a large number of expression patterns from large scale *in situ* screens. These data provide the best source for molecular markers for anatomical structures. In many cases high quality probes have been identified and are listed on the representative anatomy page. This is a good place to begin.



Specify the anatomical structure of interest in the **Anatomy Term** box. A word completion feature will provide you with a list of structures matching your input text as you type. Select the desired structure from the list.

A search for **floor plate diencephalon** lead you to the following anatomy page.



Home Email ZFIN About ZFIN Helpful Hints Citing ZFIN

Copyright © University of Oregon, 1994-2006, Eugene, Oregon.
ZFIN logo design by Karl Pape, University of Oregon

In this case, a high quality probe, cb110, has been identified by Thisse et al.. This does not mean that this is the best probe for the structure merely that it is a good probe that possibly can be used as a marker for the structure. High quality probes are not yet identified for all structures. In those cases, you may follow the All Expressed Genes links to look for a clone that will meet your needs. Large scale in situ screens offer the best opportunity for identifying a probe. In all cases, check the images available on ZFIN before using the probe to ensure that it will meet your needs.

## **Exercises**

Find a possible molecular marker for a structure of interest.