

Database

Open Access

## A PATO-compliant zebrafish screening database (MODB): management of morpholino knockdown screen information

Michelle N Knowlton\*<sup>1,6</sup>, Tongbin Li<sup>2</sup>, Yongliang Ren<sup>2</sup>, Brent R Bill<sup>1</sup>, Lynda BM Ellis<sup>3,4</sup> and Stephen C Ekker<sup>1,4,5</sup>

Address: <sup>1</sup>Department of Genetics, Cell Biology and Development, University of Minnesota, Minneapolis, MN 55455, USA, <sup>2</sup>Department of Neuroscience, University of Minnesota, Minneapolis, MN 55455, USA, <sup>3</sup>Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, MN 55455, USA, <sup>4</sup>University of Minnesota Arnold and Mabel Beckman Center for Transposon Research, Minneapolis, MN 55455, USA, <sup>5</sup>Department of Biochemistry and Molecular Biology, Mayo Clinic, Rochester, MN 55905, USA and <sup>6</sup>Box 45, The Jackson Laboratory, 600 Main Street, Bar Harbor, Maine 04609, USA

Email: Michelle N Knowlton\* - [mnk@informatics.jax.org](mailto:mnk@informatics.jax.org); Tongbin Li - [toli@biocompute.umn.edu](mailto:toli@biocompute.umn.edu); Yongliang Ren - [yongliang@biocompute.umn.edu](mailto:yongliang@biocompute.umn.edu); Brent R Bill - [bill0115@umn.edu](mailto:bill0115@umn.edu); Lynda BM Ellis - [lynda@umn.edu](mailto:lynda@umn.edu); Stephen C Ekker - [ekker.stephen@mayo.edu](mailto:ekker.stephen@mayo.edu)

\* Corresponding author

Published: 7 January 2008

Received: 28 April 2007

BMC Bioinformatics 2008, 9:7 doi:10.1186/1471-2105-9-7

Accepted: 7 January 2008

This article is available from: <http://www.biomedcentral.com/1471-2105/9/7>

© 2008 Knowlton et al; licensee BioMed Central Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/2.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Abstract

**Background:** The zebrafish is a powerful model vertebrate amenable to high throughput *in vivo* genetic analyses. Examples include reverse genetic screens using morpholino knockdown, expression-based screening using enhancer trapping and forward genetic screening using transposon insertional mutagenesis. We have created a database to facilitate web-based distribution of data from such genetic studies.

**Description:** The MORpholino DataBase is a MySQL relational database with an online, PHP interface. Multiple quality control levels allow differential access to data in raw and finished formats. MODBv1 includes sequence information relating to almost 800 morpholinos and their targets and phenotypic data regarding the dose effect of each morpholino (mortality, toxicity and defects). To improve the searchability of this database, we have incorporated a fixed-vocabulary defect ontology that allows for the organization of morpholino affects based on anatomical structure affected and defect produced. This also allows comparison between species utilizing Phenotypic Attribute Trait Ontology (PATO) designated terminology. MODB is also cross-linked with ZFIN, allowing full searches between the two databases. MODB offers users the ability to retrieve morpholino data by sequence of morpholino or target, name of target, anatomical structure affected and defect produced.

**Conclusion:** MODB data can be used for functional genomic analysis of morpholino design to maximize efficacy and minimize toxicity. MODB also serves as a template for future sequence-based functional genetic screen databases, and it is currently being used as a model for the creation of a mutagenic insertional transposon database.

## Background

Advances in scientific technology and the availability of genomic sequence for an increasing number of species have led to a paradigm shift from identifying sequence to gene product function. Large functional genetic screens are becoming increasingly prevalent and require a new way of thinking about data to facilitate cross-species comparison and deciphering gene product function in different organisms. Zebrafish are particularly amenable to large screens as they have a tremendous ability to reproduce, and their embryos finish the development of most organ rudiments by 2 days post-fertilization [1]. Their semi-transparent embryos and the ready availability of transgenic lines facilitate functional annotation of the many genes required for numerous organ systems and pathways. While historically their most common use in genetic screenings has been primarily restricted to ENU mutagenesis screens [2-4] and retroviral mutagenesis [5], the availability of morpholinos [6], antisense knockdown oligonucleotides, and insertional mutagenic transposons [7] have opened new avenues for functional genetic research.

Recently, several morpholino screens have been published in collaboration with the Ekker laboratory [8,9]. The largest among them, the Secretome Screen [9,10], identified likely secreted genes from available transcriptome databases using a bioinformatics approach. The Vertebrate Secretome DataBase (VSDB) [4] was established to identify co-translationally translocated (CTT) proteins in several model organism species [10]. Using a combination of signal sequence, transmembrane domain, and initiation site predictors as well as homology to known secreted proteins, a subset of zebrafish CTT genes were identified (0.3× genome coverage based on zebrafish loaded and predicted protein sequences in the CTTome database and Genescan gene prediction estimates of the zebrafish genome [4,5]). 150 gene targets were selected at random for morpholino design, carried out using an Assisted MORpholino Design tool (AMOD) [11,12]. Each gene was knocked down in zebrafish via morpholino-injection then screened for defects in general morphology and several organ systems including ocular, renal, pigment, vascular, hematopoietic, cranio-facial, and motor [9]. A second morpholino knockdown screen, the Hematopoietic Stem Cell Screen, was conducted using zebrafish homologues of human genes differentially expressed in hematopoietic stem cell differentiation by microarray analysis [8].

Currently, the zebrafish community stores published and sequence-related results on several web-based databases. The principal database for zebrafish-related research is the Zebrafish Information Network (ZFIN) [13,14]. Ensembl [5], likewise, stores sequences targeted by published mor-

pholinos. Ensembl is limited in its ability to store and display experimental data, and while ZFIN contains a wealth of information on published results it lacks specific information such as dosage and searchable phenotype data for published morpholinos and phenotypic data for largely unpublished morpholinos.

Many model organism communities have developed online databases to access results from mutagenic, knock-down/out and over-expression studies. While mouse knockout mutant [15,16] and *C. elegans* RNAi [17,18] experimental result databases with phenotypic content exist, the only morpholino-specific database available is for a pilot *Xenopus* morpholino screen in which 202 genes were targeted and subsequent phenotypes displayed on a webpage [19,20]. The *Xenopus tropicalis* Morpholino Screen database is a flat format website that allows for records to be accessed and viewed by gene targeted, syn-phenotype group or specific defect. Notably this database does not include search functions and does not offer a means for further input by outside sources. However, the intent of this database as stated by the authors is to display the results from their specific screen [19,20].

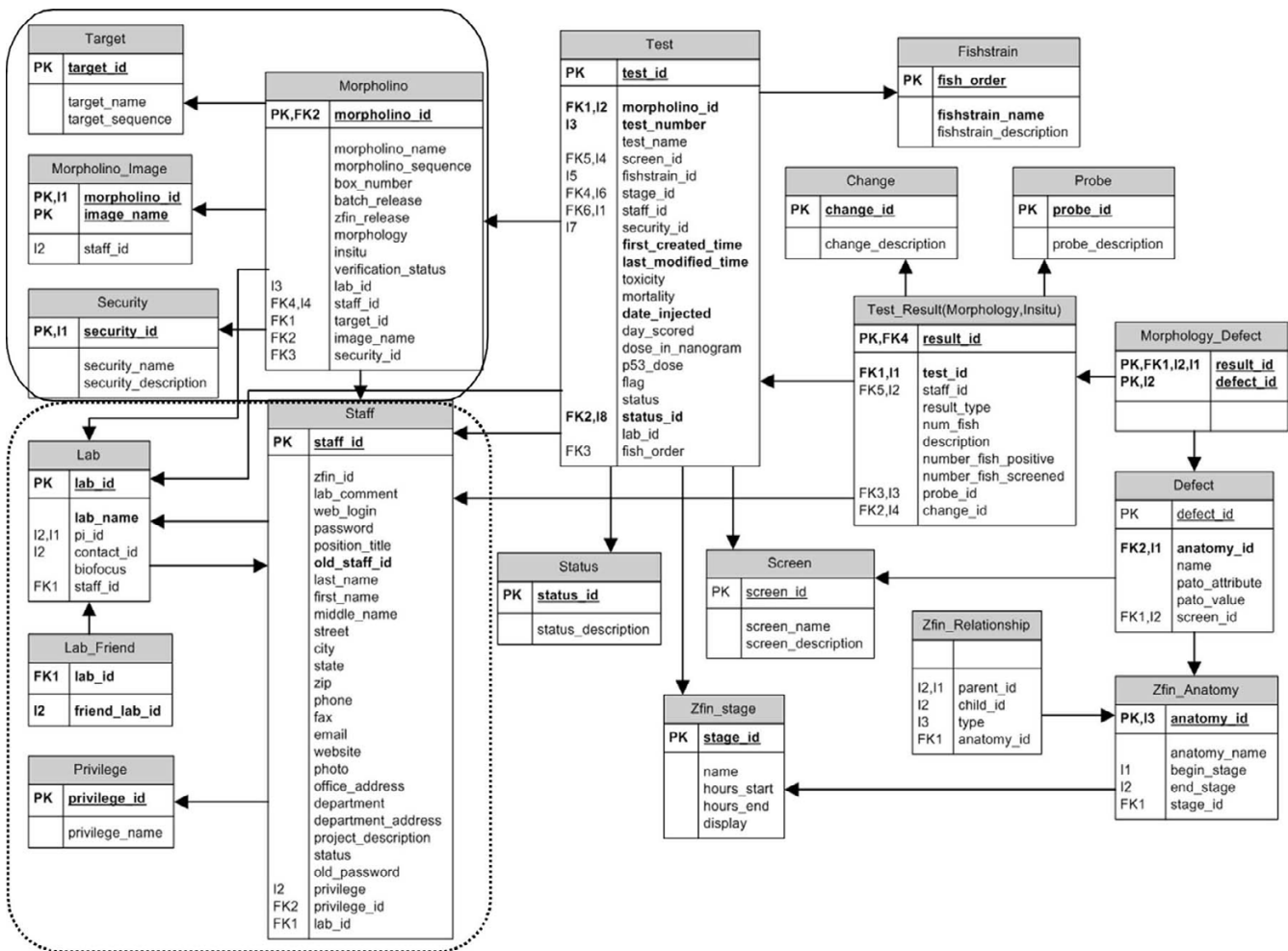
Given the intent of our database to display the large amount of data generated from our on-going collaborative screens, we have developed of a scalable, centralized yet easily accessible worldwide, MORpholino DataBase (MODB) [3] to store and facilitate retrieval of zebrafish morpholino screening results. Additional access has been provided by active links between MODB and ZFIN.

MODB is a scalable, searchable and readily adaptable online MySQL database with a PHP interface. MODB includes storage and retrieval of results generated by our screens and public access of published and unpublished information on morpholinos MODB not only allows dissemination of target-specific information such as effect of knockdown, but also facilitates analysis of morpholino-specific data such as toxicity, dosage, mortality and design. MODB is a new, online, publicly accessible database for data-mining zebrafish morpholino knockdown effects and morpholino non-specific effects.

## Construction and Content

### Construction

The MODB web project is currently running on an Apache web server [21] with PHP [1] on a Linux system (Intel Xeon, 2.8 GHz, CentOS release 4.4). The project architecture was developed according to the Model-View-Controller (MVC) design paradigm. The Model (database tables) is stored in a MySQL relational database [22]. Major tables include Morpholino, Staff, Lab, Test, and Test\_Result and store all information pertinent to each morpholino and users (Figure 1). The View was developed with the help of



**Figure 1**  
**Entity-Relationship diagram of MODB.** The organization of data contained within MODB is centered on the unique morpholino sequence. Tables pertaining to the morpholino-specific information are located in the upper left corner (solid box). The tables in the lower left corner establish the relations between laboratories, users and security level designation (dashed line box). The tables on the right half of the diagram represent the tests and test components such as anatomical structure, screen and defect observed (unboxed area).

Smarty [23], a template engine in PHP that helps the designer make modifications in the layout without breaking PHP logic. PHP scripts control the connections between the Model and the View. The GD graphics library (a PHP module) [24] is used for the graphic functions displayed in the summary page.

NCBI BLAST 2.2.13 [25] is invoked by PHP through system calls in the sequence search functions. The MODB BLAST target database is updated automatically when a new morpholino sequence is entered into the database. The ZFIN anatomy ontology [26] was downloaded (release date 08/16/2007) from the ZFIN web site in the OBO format [26] and transferred into the database. An

anatomy browser was written in PHP and incorporated in MODB.

**Content**

The unique identifier in MODB is the sequence of each morpholino. Many morpholino targets in our screens are derived from transcriptome sequence data from The Institute for Genomic Research (TIGR) [27] and GenBank [28] that are subject to change as new sequence data becomes available. For every unique morpholino entered into the database, there is an accompanying set of tables devoted to information pertaining to that morpholino such as name, designator, laboratory created in, contact person, target name, target sequence, gene targeted, reference

General Information [\[Add\]](#) [\[Del\]](#) [\[Edit\]](#)

Morpholino name: HS2012  
 Box number: 431  
 Morpholino sequence: ACTCCATCTGTAGGTCGTGTTCCCTC  
 Target name: SPRY1

Created in: No Lab  
 User name:  
 Verified?: No

Target sequence:

```
GTAGAAGACTGTGTCCTTACTGCAGTGCACATCGCACGGATAGACGTCGCTTTTACTCCGCGG
CTTCGT
TTGGATCTATCTGTCTCTCTTTTTGCCACTGGAATATATCGCAACGACCTGTTTCATCTGAGAAGG
```

Reference: [?](#) [Add Reference](#)

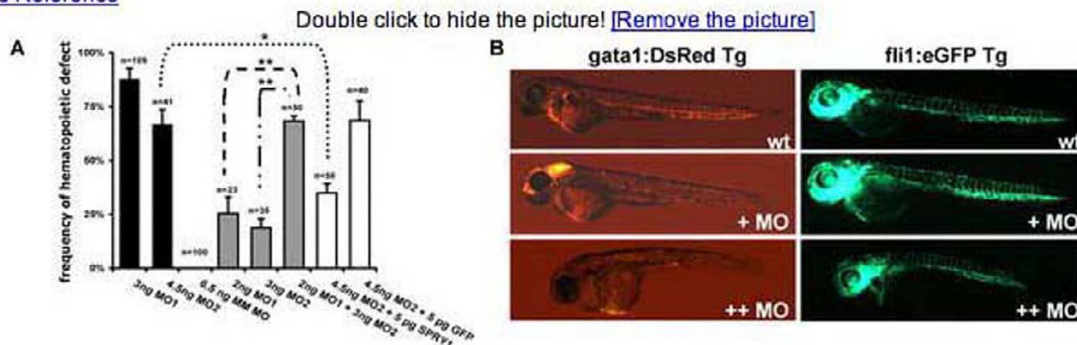


Figure 2

**Morpholino-specific information.** Information such as morpholino name and sequence, in laboratory morpholino designator (box number), creating laboratory, initiating user and any reference images associated with the morpholino are displayed on the initial morpholino page for each morpholino.

images and number of tests created (Figure 2). This concise format allows key information regarding the morpholino to be viewed at once.

Early in 2007, MODB contained over 700 unique morpholino sequences generated from several screens, individual projects and control morpholinos. The first generation Secretome Screen [9] is the largest overall contributor of morpholino sequences. From 150 candidate genes sequences, 209 morpholinos were designed. Each morpholino was subjected to a variety of morphological and molecular assays to determine the effect of down-regulating a particular gene product during the first 5-days of zebrafish development [9]. This Secretome Screen identified 26/150 gene targets that produced reproducible effects on specific aspects of early development in early morphology and a variety of organ systems including visual, renal, cardiovascular, cranio-facial, hematopoiesis, and pigment. The more focused Hematopoietic Stem Cells Screen identified 16/64 targets as being necessary for proper blood development [8]. The bulk of the data entered into MODB (124/150 target gene morpholinos in the Secretome Screen and 48/64 in the Hematopoietic Stem Cell Screen) consists of morpholinos that had no observed developmentally specific effect in our screening

tests, caused toxic effects such as a high amount of cell death and/or mortality, or produced pleiotropic effects we could not decipher. This information is of use to the scientific community for optimizing morpholino design and learning dosage data and overall effect of previously designed and tested morpholinos.

Since MODB is a living database, new data is regularly being submitted. In the coming year another 50 to 150 morpholinos will be added to the database for the continuing Hematopoietic Stem Cell and secretome screens. To allow the release of data in defined formats and quality, MODB has three data sharing quality control levels. Users are capable of accessing and modifying data related to their designated morpholinos. To facilitate collaborative efforts, a shared designation has been developed to allow members of consortium laboratories to access data related to morpholinos designed by any user within the laboratories that are connected by a relationship within MODB. The third designation ('public') allows morpholino data to be used to generate the summary reports available to the open scientific community via the MODB website. Nearly all morpholinos (779/784) in MODB, with the exception of non-zebrafish morpholinos, are designated as public in MODBv1.

<b>[Test 26] [Top] [Edit][Del]</b>	<b>Vasculature</b>	<b>Created By</b>	Michelle Knowlton (ZFIN) (No Lab)
<b>Mortality:</b>	0	<b>Toxicity:</b>	0
<b>Time Scored :</b>	(48-60) Hatching:Long-pec	<b>Flag for retest:</b>	No
<b>Call Suggestion:</b>	None	<b>Fish strain used:</b>	Fli-1/Gata-1
<b>Morpholino dose:</b>	3ng	<b>Security:</b>	shared
<b>P53 Dose:</b>	0	<b>File last modified:</b>	2007-01-12 14:47:49
<b>Date Injected:</b>	0000-00-00 00:00:00		

**Add Morphology Information (Morphology) For Test 26 of HS2012**

Positive Number

Screened Number

Description

**none**

<input type="checkbox"/> Edema	<input type="checkbox"/> Plieotropic; very messed up	<input type="checkbox"/> Undifferentiated; blobs etc	<input type="checkbox"/> Head size; smaller
<input type="checkbox"/> Head size; larger	<input type="checkbox"/> Delayed development		

**Toxicity**

<input type="checkbox"/> Cell death; head	<input type="checkbox"/> Cell death; trunk	<input type="checkbox"/> Cell death; telencephalon	<input type="checkbox"/> High death
---	--	--	-------------------------------------

**Morphology**

<input type="checkbox"/> Holoprosencephaly	<input type="checkbox"/> Midbrain/hindbrain boundary reduced	<input type="checkbox"/> Yolk extension size; shorter	<input type="checkbox"/> Head morphology; dysmorphic
<input type="checkbox"/> Notochord shape; undulating	<input type="checkbox"/> Notochord length; short/reduced	<input type="checkbox"/> Fin development; none	<input type="checkbox"/> Somites Shape; blocky, short
<input type="checkbox"/> Tail size; edema	<input type="checkbox"/> Tail Shape; curved etc	<input type="checkbox"/> Otolith; 0	<input type="checkbox"/> Otolith; 1
<input type="checkbox"/> Otolith; detached	<input type="checkbox"/> Pectoral fin; reduced	<input type="checkbox"/> Tail; shorter	<input type="checkbox"/> Yolk; edema
<input type="checkbox"/> Morphology; normal	<input type="checkbox"/> Trunk; curvature	<input type="checkbox"/> Notochord shape; kinked (one bend)	<input type="checkbox"/> Hindbrain; fluid retention
<input type="checkbox"/> Axis; dysmorphic	<input type="checkbox"/> Axis; shortened	<input type="checkbox"/> Notochord; Thinned	<input type="checkbox"/> Somite; dorsal curve
<input type="checkbox"/> Somite; ventral curve			

**Vasculature**

<input type="checkbox"/> Circulation; decreased	<input type="checkbox"/> CVP circulation; reduced	<input type="checkbox"/> CVP circulation; absent	<input type="checkbox"/> ISV; present
<input type="checkbox"/> ISV circulation; reduced	<input type="checkbox"/> ISV circulation; absent	<input type="checkbox"/> Heart beat; none	<input type="checkbox"/> Heart; edema
<input type="checkbox"/> Cardinal vein; enlarged	<input type="checkbox"/> Dorsal aorta; reduced circulation	<input type="checkbox"/> Heart; no looping	<input type="checkbox"/> Ventricle; enlarged
<input type="checkbox"/> Atria; enlarged	<input type="checkbox"/> Aorta; abnormal	<input type="checkbox"/> ISV; multiple banches	<input type="checkbox"/> ISV; uncoordinated
<input type="checkbox"/> CVP; decreased remodeling	<input type="checkbox"/> Dorsal aorta; shortened	<input type="checkbox"/> Dorsal aorta; premature return	<input type="checkbox"/> CVP; blood pooling
<input type="checkbox"/> Yolk; blood pooling	<input type="checkbox"/> Head; blood pooling	<input type="checkbox"/> Tail; blood pooling	<input type="checkbox"/> Trunk; blood pooling
<input type="checkbox"/> Dorsal aorta; gapped	<input type="checkbox"/> Circulation; absent	<input type="checkbox"/> ISV; short	<input type="checkbox"/> CVP; expanded
<input type="checkbox"/> ISV; absent	<input type="checkbox"/> Vasculature; normal	<input type="checkbox"/> Heart beat; slow	<input type="checkbox"/> Tail Vaculature; reduced
<input type="checkbox"/> Trunk vasculaure; reduced			

**Blood**

<input type="checkbox"/> Hemoglobin reduced	<input type="checkbox"/> Hemoglobin absent	<input type="checkbox"/> ICM expanded	<input type="checkbox"/> ICM reduced
<input type="checkbox"/> ICM pooling	<input type="checkbox"/> Blood reduced cell number	<input type="checkbox"/> Blood; none	<input type="checkbox"/> Blood; reduced
<input type="checkbox"/> Blood island patch	<input type="checkbox"/> Gata-1 expression; reduced	<input type="checkbox"/> Gata-1 expression; absent	<input type="checkbox"/> Blood pooling
<input type="checkbox"/> Blood; normal			

**Figure 3**  
**Common name defect check box.** Entry of observations following morpholino injection into zebrafish embryos is catalogued using a fixed vocabulary checkbox form. Penetrance of observed defect is also recorded for subsequent analysis.

Defect Name	Quality Value	Pato ID	Screen	
Unilateral Anophthalmia	count	PATO:0000053	Eye	<a href="#">Edit</a> <a href="#">Delete</a>
Bilateral Anophthalmia	count	PATO:0000053	Eye	<a href="#">Edit</a> <a href="#">Delete</a>
Optic fissure; open	open	PATO:0000610	Eye	<a href="#">Edit</a> <a href="#">Delete</a>
Eye Pigment, Reduced	absent	PATO:0000462	Eye	<a href="#">Edit</a> <a href="#">Delete</a>
Blood pooling	[proposed pooling]		Blood	<a href="#">Edit</a> <a href="#">Delete</a>
Gata-1 expression; absent	absent	PATO:0000462	Blood	<a href="#">Edit</a> <a href="#">Delete</a>
Gata-1 expression; reduced	decreased	PATO:0000468	Blood	<a href="#">Edit</a> <a href="#">Delete</a>
Blood island patch	arrested	PATO:0000297	Blood	<a href="#">Edit</a> <a href="#">Delete</a>
Blood; reduced	decreased	PATO:0000468	Blood	<a href="#">Edit</a> <a href="#">Delete</a>
Blood; none	absent	PATO:0000462	Blood	<a href="#">Edit</a> <a href="#">Delete</a>
Blood reduced cell number	decreased number	PATO:0000419	Blood	<a href="#">Edit</a> <a href="#">Delete</a>
ICM pooling	[proposed pooling]		Blood	<a href="#">Edit</a> <a href="#">Delete</a>
ICM reduced	small size	PATO:0000587	Blood	<a href="#">Edit</a> <a href="#">Delete</a>
ICM expanded	large size	PATO:0000586	Blood	<a href="#">Edit</a> <a href="#">Delete</a>
Hemoglobin absent	absent	PATO:0000462	Blood	<a href="#">Edit</a> <a href="#">Delete</a>
Hemoglobin reduced	decreased	PATO:0000468	Blood	<a href="#">Edit</a> <a href="#">Delete</a>

**Figure 4**  
**Correspondence between defect description and PATO designation.** As work is still ongoing on establishing a cross-species phenotypic vocabulary PATO, terminology is constantly evolving. Defect names are being used to maintain some consistency but are rooted in PATO designations via a translation table. When no PATO term exists to adequately describe the defect observed, a proposed term is submitted to NCBO via ZFIN.

**Utility**  
**Data Entry**

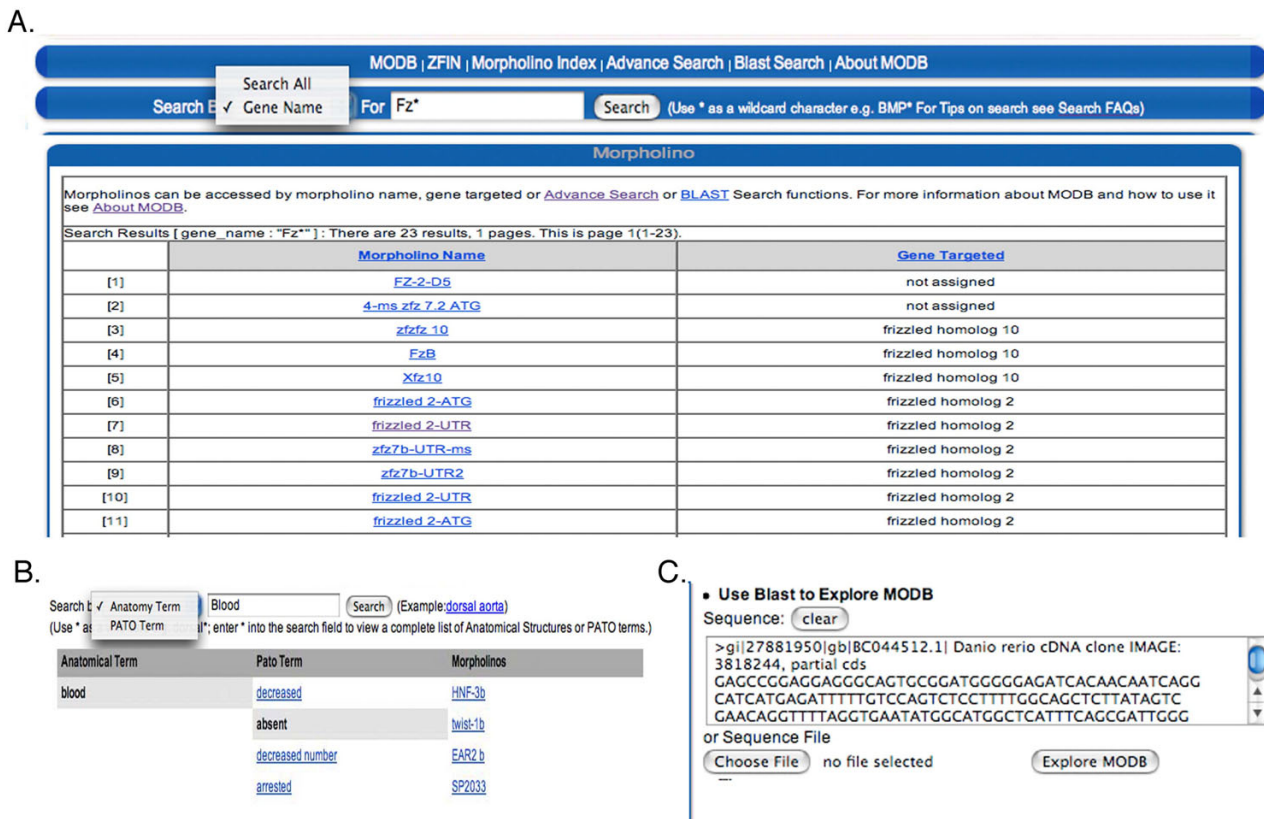
A major function of MODB is data compilation of morpholino-specific information generated by primary research. There are two main ways to add data. Currently,

administrators alone enter new listings that provide basic information on a morpholino. Authorized users can submit morpholino sequence and results from specific morphological or molecular experiments for a given observation point and dose. Researchers can apply for an

authorized-user identification and password (see author contact information).

These results are called 'tests'. In early 2007, MODB contained over 9000 individual test entries. A free text notes field permits detailed description defects observed in morpholino-injected embryos once a fixed vocabulary defect term has been selected. The defects ontology check box was established with common scientific terms used in our screening protocols (Figure 3). Specific in-laboratory scientific descriptions correspond to established defects ontology. The reduction of free text and the promotion of fixed vocabularies increases the searchability of MODB for

search by keyword, anatomical structure or fixed defects term and allows for future integration with other fixed-vocabulary phenotype databases such as WormBase and the Mouse Phenome database [15,29]. To utilize the data entered into MODB in a meaningful manner, we developed a variety of search functions. In addition to the parameters already mentioned (laboratory designator, morpholino name, morpholino target name, gene targeted, laboratory created in, or keyword), MODB can be queried by a controlled defects vocabulary. This defects ontology was developed in conjunction with ZFIN to model the current status of Phenotypic Attribute Trait Ontology (PATO) developed by the National Center for



**Figure 5**  
**MODB search functions.** **A.** Data retrieval from MODB is simplified by morpholino name and gene targeted. Results can be sorted morpholino name and gene targeted. **B.** An additional search in the public interface allows for searches based on PATO term and anatomical structure affected. **C.** An additional sequence-based search using morpholino or target sequence can be performed.

Biomedical Ontology (NCBO) [30]. To facilitate the use of this common language, MODB users assess phenotypes using defect descriptions linked to current PATO termi-

nology (Figure 4). Initially, PATO was designed to describe a defect observed in an anatomical structure (the entity, e.g. tail) using a trait (e.g. relative length) and

**General Information**

**Morpholino name:** SP2009 A  
**Morpholino sequence:** ACAGCAAAAACACACGCATCCTTCT  
[Find in ZFIN](#) [Find in Genbank](#)  
**Target name:** ambpl

**Target sequence:** (Morpholino Target and ATG)  
[Search in Genbank](#) [Search in Ensembl](#)

CTACTGGGACGATTGGTTTGAGTCAACAGGCAGATAAAAAGGACCTTTGCTCTGTTTGTGTCTGGCCTGGTTTTGGGGGT  
 GTTGTTCAGTAGGTCAGTGTCTCTGTGGACAGTCCGGCATCTCCTTCAGAGCTTGATTGGGACCTCAGACGAACAAGAGGA  
TGCGTGTGTTTTTGGCTGTTTCTGCCTCTGCTGGGGCTCCTGGAGGCTGGACCTCTGCCGGATGAGCCCGTCTGCAGACA  
 CAGGAGAACTTCAATGTCAATCGGTTTCATGGGGAAATGGTATGATATTGCGGTGGCGTCCACATGTCTCTGGATGAAGCG

**Gene targeted:**

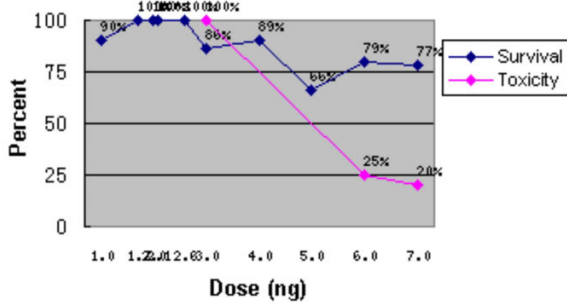
alpha-1-microglobulin/bikunin precursor, like  
[Find in ZFIN](#) [Find in Pubmed](#)  
[All the morpholinos targets to the gene](#)

**Synonyms:**

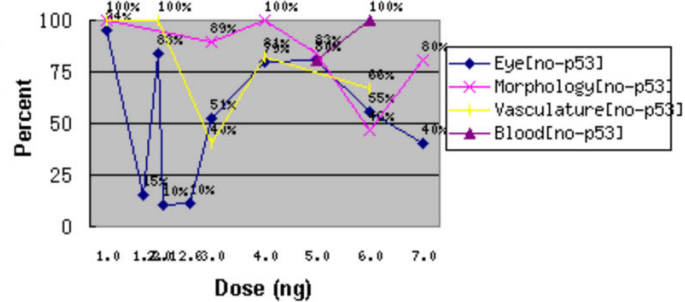
**Comments:**

**Reference:**

**Survival Curve:**



**Normality Curve:** ALL



Note: When more than one data point existed per dose an average was used. p53 morpholino co-injections were used to minimize non-specific toxicity effects (Robu et al. PLoS Genetics 2007)

**Screens Performed (number of tests):**

Eye: (28)  
 Vasculature: (9)  
 Morphology: (12)  
 Blood: (2)

**Defects Observed:**

- [head count](#)  
(10%, dose:1ng, p53:ng, days: Day: 3)
- [eye small size](#)  
(42%, dose:1.7ng, p53:0ng, days: -)
- [eye absent](#)  
(10%, dose:2ng, p53:ng, days: Day: 2)
- [eye dysplastic](#)  
(10%, dose:2ng, p53:ng, days: Day: 2)
- [eye small size](#)  
(38%, dose:2.1ng, p53:0ng, days: -)

**Images**



Figure 1. Eye defect.

**Figure 6**

**Public summary page.** A publicly accessible summary page is available for released morpholinos in MODB. This page contains morpholino and target information such as name, sequence etc. and a review of knockdown results such as dosing for mortality and normalcy. A list of tests performed, defects observed and images is also displayed.



attribute of that trait (e.g. short length). This system has subsequently been replaced with an entity and quality designation (e.g. tail, short length). MODB's defects ontology contains the latter version of PATO. When no PATO term exists to adequately describe the defect observed, a proposed term is submitted to NCBO via ZFIN to assist in the development of a cross-species phenotypic ontology. The importance of a common phenotypic vocabulary resides in the ability to perform cross-species analysis of gene function and conserved pathways.

#### **Data retrieval**

The usefulness of the data stored within any database is subject to the retrieval methods available. MODB has been designed with numerous search and sort functions that allow for retrieval by morpholino sequence, morpholino name and gene targeted (Figure 5A). An advanced search feature in the public interface allows users to search by anatomy and phenotype to identify morpholinos with a given defect (Figure 5B). A sequence-based search is also available, using either the target morpholino sequence or the sequence of the morpholino (Figure 5C). By offering numerous search and sorting options, results of nearly 800 morpholinos and more than 9000 tests can be accessed and analyzed for further scientific insight into the knockdown effect of a specific morpholino target, genes effecting the same developmental pathways, and non-specific morpholino effects.

#### **MODB v1 Public Interface**

To facilitate display and digestion of the large amount of data stored within MODB, a summary sheet for each morpholino was created (see construction and content) and is publicly available. Each morpholino summary contains sequence and target information as well as survival and normalcy (percent normal embryos observed overall or for specific screen types) curves based on dose (Figure 6). Several search functions can be used to access this information including by gene targeted, BLAST and an advanced search function that facilitates searching for morpholinos by PATO designation and anatomical structure affected (Figure 5B). MODB thus presents users with the valuable knowledge, such as mortality and dose curves, effectiveness and morphant phenotype. Currently, 101 morpholinos contain hyperlinks to specific morpholino entries in ZFIN and hyperlinks are provided for gene targets in MODBv1.

#### **Discussion**

With advent of whole-genome sequencing technology, a greater emphasis on gene function is emerging. Functional genomic screening tools available in zebrafish such as ENU mutagenesis, GripNAs, RNAi, insertional mutagenesis and morpholinos allow for large functional gene screens that produce an enormous amount of biological

information (reviewed in [2,6,7,31-33]). Efficient data storage is critical to allow for retrieval of results in a meaningful fashion and to facilitate data mining. MODB is a non-proprietary, relational database with online accessibility that acts as a storehouse for critical information relating to sequence-specific morpholino knockdown screens. MODB is, to our knowledge, unique among the available databases in that it offers searchable morpholino-specific data and fixed-vocabulary defects ontological cataloging of morphants that allows for cross-species analysis using a common, PATO-derived ontology. Cross-linking to ZFIN provides additional value of MODB for the research community.

Continued screening efforts in zebrafish using morpholinos and other sequence-specific knockdown or knockout technologies will expand our knowledge of gene function. MODB provides a central repository for morpholino information. This schema is potentially applicable to any sequence-based screening approach in zebrafish or related systems. For example, MODB is currently being adapted to the creation of a transposon database that includes insertional mutagenesis and expression analysis data.

#### **Conclusion**

By offering public access to our schema and the data contained within MODB, we hope to aid in the advancement of functional genetic studies and gene annotation as well as improvements in morpholino analysis and design. The dissemination of valuable knowledge gained via large screens such as the Secretome [9] and Hematopoietic Stem Cell [8] screens goes beyond published gene target characterization and provides a foundation for future research.

#### **Availability and requirements**

MODB is accessible via the web <http://www.secretomes.umn.edu/MODB/>. To be added as an authorized user, contact the corresponding author M.K. or S.C.E. Morpholinos are readily available for purchase through Gene Tools [34] or Open Biosystems [35]. The authors have no affiliation, financial or otherwise, with either corporate entity.

#### **Authors' contributions**

MK designed, tested and managed the database. TL and YR were responsible for the programming, debugging and informatics support. BB aided in the conception, design and testing of the database. EL and SE aided in the conception and design of the database and coordinated the project as a whole. All authors read and approved the final manuscript.

## Acknowledgements

The authors would like to acknowledge the numerous collaborations that have contributed to the development of MODB such as the laboratories of Drs. Farber, Hammerschmidt, Pickart, Schimmenti, Sivasubbu, and Verfaillie. We would also like to acknowledge the bioinformaticians who have worked on this project including Dr. Eric Klee and Kyong Jin Shim. Funding for this project was obtained from the NIH to SCE (GM63904 and CA65493). This work was supported in part by the University of Minnesota Supercomputing Institute.

## References

- Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF: **Stages of embryonic development of the zebrafish.** *Dev Dyn* 1995, **203(3)**:253-310.
- Geisler R, Rauch GJ, Geiger-Rudolph S, Albrecht A, van Bebber F, Berger A, Busch-Nentwich E, Dahm R, Dekens MP, Dooley C, et al.: **Large-scale mapping of mutations affecting zebrafish development.** *BMC Genomics* 2007, **8(1)**:11.
- Draper BV, McCallum CM, Stout JL, Slade AJ, Moens CB: **A high-throughput method for identifying N-ethyl-N-nitrosourea (ENU)-induced point mutations in zebrafish.** *Methods Cell Biol* 2004, **77**:91-112.
- Wienholds E, van Eeden F, Kusters M, Mudde J, Plasterk RH, Cuppen E: **Efficient target-selected mutagenesis in zebrafish.** *Genome Res* 2003, **13(12)**:2700-2707.
- Amsterdam A, Burgess S, Golling G, Chen W, Sun Z, Townsend K, Farrington S, Haldi M, Hopkins N: **A large-scale insertional mutagenesis screen in zebrafish.** *Genes Dev* 1999, **13(20)**:2713-2724.
- Pickart MA, Sivasubbu S, Nielsen AL, Shriram S, King RA, Ekker SC: **Functional genomics tools for the analysis of zebrafish pigment.** *Pigment Cell Res* 2004, **17(5)**:461-470.
- Taylor MR, Kikkawa S, Diez-Juan A, Ramamurthy V, Kawakami K, Carmeliet P, Brockerhoff SE: **The zebrafish pob gene encodes a novel protein required for survival of red cone photoreceptor cells.** *Genetics* 2005, **170(1)**:263-273.
- Eckfeldt CE, Mendenhall EM, Flynn CM, Wang TF, Pickart MA, Grindle SM, Ekker SC, Verfaillie CM: **Functional analysis of human hematopoietic stem cell gene expression using zebrafish.** *PLoS Biol* 2005, **3(8)**:e254.
- Michael A, Pickart EVK, Nielsen Aubrey L, Sivasubbu Sridhar, Mendenhall Eric M, Bill Brent R, Chen Eleanor, Eckfeldt Craig E, Knowlton Michelle, Robu Mara E, Larson Jon D, Deng Yun, Schimmenti Lisa A, Ellis Lynda BM, Verfaillie Catherine M, Hammerschmidt Matthias, Farber Steven A, Ekker Stephen C: **Genome-Wide Reverse Genetics Framework to Identify Novel Functions of the Vertebrate Secretome.** *PLoS One* 2006, **1(1)**:e104.
- Klee EW, Ellis L BM: **Evaluating eukaryotic secreted protein prediction.** *BMC Bioinformatics* 2005, **6**:256.
- Sadl VS, Sing A, Mar L, Jin F, Cordes SP: **Analysis of hindbrain patterning defects caused by the kreisler(enu) mutation reveals multiple roles of Kreisler in hindbrain segmentation.** *Dev Dyn* 2003, **227(1)**:134-142.
- Klee EW, Shim KJ, Pickart MA, Ekker SC, Ellis LB: **AMOD: a morpholino oligonucleotide selection tool.** *Nucleic Acids Res* 2005:W506-511.
- Miller-Bertoglio V, Carmany-Rampey A, Furthauer M, Gonzalez EM, Thisse C, Thisse B, Halpern ME, Solnica-Krezel L: **Maternal and zygotic activity of the zebrafish ogon locus antagonizes BMP signaling.** *Dev Biol* 1999, **214(1)**:72-86.
- Sprague J, Bayraktaroglu L, Clements D, Conlin T, Fashena D, Frazer K, Haendel M, Howe DG, Mani P, Ramachandran S, et al.: **The Zebrafish Information Network: the zebrafish model organism database.** *Nucleic Acids Res* 2006:D581-585.
- Mouse Phenome Database** [<http://phenome.jax.org/pub/cgi/phenome/mpd.cgi?rtn=docs/home>]
- Bogue MA, Grubb SC, Maddatu TP, Bult CJ: **Mouse Phenome Database (MPD).** *Nucleic Acids Res* 2007:D643-649.
- Horn T, Arziman Z, Berger J, Boutros M: **GenomeRNAi: a database for cell-based RNAi phenotypes.** *Nucleic Acids Res* 2007:D492-497.
- GenomeRNAi** [<http://www.dkfz.de/signaling2/rnai/ernai.html>]
- Xenopus tropicalis Morpholino Screen** [<http://www.gurdon.cam.ac.uk/~smithlab/screens/Xenopus-morpholino-pilot/morpholino-screen-lists.html#synphenotypes>]
- Rana AA, Collart C, Gilchrist MJ, Smith JC: **Defining synphenotype groups in Xenopus tropicalis by use of antisense morpholino oligonucleotides.** *PLoS Genet* 2006, **2(11)**:e193.
- Apache Web Server** [<http://www.apache.org>]
- Nasevicius A, Ekker SC: **Effective targeted gene 'knockdown' in zebrafish.** *Nat Genet* 2000, **26(2)**:216-220.
- Smarty** [<http://smarty.php.net>]
- GD Graphics** [<http://www.boutell.com/gd/>]
- Altschul SF, Gish W, Miller W, Myers EV, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215(3)**:403-410.
- ZFIN Anatomy Ontology** [<http://www.geneontology.org/GO.format.shtml>]
- The Institute for Genomic Research (TIGR)** [<http://www.tigr.org/>]
- GenBank** [<http://www.ncbi.nlm.nih.gov/Genbank/index.html>]
- WormBase** [<http://www.wormbase.org/>]
- National Center for Biomedical Ontology (NCBO)** [<http://www.bioontology.org/>]
- Amsterdam A: **Insertional mutagenesis in zebrafish: genes for development, genes for disease.** *Brief Funct Genomic Proteomic* 2006, **5(1)**:19-23.
- Amsterdam A, Hopkins N: **Mutagenesis strategies in zebrafish for identifying genes involved in development and disease.** *Trends Genet* 2006, **22(9)**:473-478.
- Dahm R, Geisler R: **Learning from small fry: the zebrafish as a genetic model organism for aquaculture fish species.** *Mar Biotechnol (NY)* 2006, **8(4)**:329-345.
- Gene Tools** [<http://www.gene-tools.com/>]
- Open Biosystems** [<http://www.openbiosystems.com/>]

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:  
[http://www.biomedcentral.com/info/publishing\\_adv.asp](http://www.biomedcentral.com/info/publishing_adv.asp)

