

# Genomic Resources for Zebrafish

## Table of Contents

- [Chemical Phenomics Initiative](#)
  - [Genome and Map Resources](#)
  - [Primers](#)
  - [Microarrays](#)
  - [cDNA and EST Resources](#)
  - [Stock Centers](#)
  - [Mutagenesis Projects](#)
  - [Gene Trap, Enhancer Trap, Insertion Mutagenesis and CreloxP System Resources](#)
  - [TALENs/CRISPRs](#)
  - [Morpholinos](#)
  - [Pathways](#)
  - [Neurobehavioral / Physiological Models](#)
  - [Gene Expression Resources](#)
  - [White Papers](#)
- 

## Chemical Phenomics Initiative

<https://my.vanderbilt.edu/chemicalphenomics/>

As a way of introduction, I do chemical genetics of zebrafish early development. <https://medschool.vanderbilt.edu/chaz-hong-lab/>

Briefly, in a manner analogous to classic mutagenesis screens, we conduct high-throughput chemical screens using zebrafish to discover small molecules that specifically perturb embryonic pattern formation. Using the interdisciplinary chemical genetic approach, we have discovered exquisitely selective modulators of the Bone Morphogenetic Protein (BMP), Wnt and Hedgehog pathways, as well as important new signaling components that direct early vertebrate development. For example, I discovered dorsomorphin, the first small molecule inhibitor of BMP signaling. The technology from this effort is nearing an Investigational New Drug (IND) stage for several devastating human diseases, such as fibrodysplasia ossificans progressiva, and incurable pediatric cancers.

Importantly, from this effort, we also have discovered many compounds that have really interesting effects on zebrafish embryos, but we don't know the mechanism of action. I hope to initiate a "crowd sourcing" to accelerate discovery and translation.

Please do not hesitate to contact me with questions about this effort and how it might benefit the zebrafish community. I look forward to hearing from you.

## Genome and Map Resources

- **[The Danio rerio Sequencing Project at Sanger:](#)**

The [Sanger zebrafish homepage](#) hosts all of the efforts of the zebrafish genome sequencing project, including the [whole genome sequencing and assembly project](#) with automated annotation in [Ensembl](#) and the [clone mapping and sequencing project](#) with manual annotation in [Vega](#).

- **[Zebrafish Genome Annotation Browser \(Vega\):](#)**

This site allows browsing of zebrafish BAC and PAC clones that were sequenced, analysed and manually annotated within the Sanger zebrafish genome project. Vega includes

- manually annotated genes with poly A features
- protein feature and domain annotation
- mapping of unists markers
- ab initio gene predictions
- tRNA prediction
- CpG island prediction and repeat masking

For Blast searches against all finished and unfinished clones go [here](#).  
The manual annotation is a joint effort between the Sanger Institute and ZFIN.

- **[Zebrafish Genome Browser at Ensembl:](#)**

This site allows browsing of the latest zebrafish whole genome assembly [Zv9](#)

- automatically annotated Ensembl genes
- automatically annotated Ensembl EST genes
- protein feature and domain annotation
- Blast searches against all zebrafish ESTs, cDNAs and the WashU and IMCB EST clusters
- Blast searches against all vertebrate proteins, Uniprot, all vertebrate cDNAs and ESTs

- mapping of unists markers
- mapping of the Affymetrix zebrafish probes
- mapping of SNPs in dbSNP
- comparative analysis with other Ensembl species (human, mouse, rat, chicken, fugu, tetraodon)
- BAC end placement
- transcription start site predictions with Eponine
- ab initio gene predictions
- tRNA prediction
- CpG island prediction and repeat masking

Additional DAS tracks with SNP and indel analysis derived from the comparison of the whole genome shotgun reads are available. Ensembl also offers completely customizable downloads with [EnsmMart](#). You can search Ensembl with [BLAT](#).

#### Additional services available at the Sanger Institute

- The [WebFPC](#) database hosts the physical map and allows searching and browsing for clone information. Please note that this database is not updated anymore. For up to date information on the clone path please check [chromoview](#)
- The [RepeatMasker](#) service masks your sequence data with a choice of repeat libraries.
- The [miRBase](#) service is the new home for microRNA data, incorporating the database and gene naming roles previously provided by the miRNA Registry, and including the new miRBase Target database.

#### • [Zebrafish Genome Browser at UCSC:](#)

A joint collaboration between UCSC and the Zebrafish Genome Initiative at Children's Hospital Boston provides several useful tracks of genome data, including:

- Mapping & Sequencing: Position / Conigs / Scaffolds / RH Map / BAC End Pairs & Singles / Gap / GC Percent
- Genes & Gene Prediction: RefSeq Genes / Ensembl Genes
- mRNA & EST: ZF mRNAs & ESTs / spliced ESTs / WZ EST Clusters / nonZF mRNAs
- Expression & Regulation: Affy ZF
- Comparative Genomics: Fugu Chain & Net / Human Proteins / HG Blast
- Variation & Repeats: RepeatMasker / SimpleRepeats

Custom annotation tracks for the zebrafish research community include:

- RH Map (about 9K sequences)
- BAC End Clones (over 288K pairs, 23K singletons)
- WZ ESTs (43K sequences, 76K alignments, 39K EST alignments)
- Affymetrix Zebrafish Microarray Probe Alignments (15K probes)
- Human Protein (over 40K alignments)

#### • [Zebrafish Genome Resources at NCBI:](#)

This page provides information about genomic resources available from NCBI and the zebrafish research community, including sequence, mapping, clone information, and NCBI's independent gene annotation.

#### • [SNPFisher : Variant Report Compiler:](#)

Variation data for FLI, TL, and WIK. The SNPfisher website includes the SNPfisher Variant Reporter tool, which provides the genomic position, alternate allele read frequency, strain specificity, restriction enzyme recognition site changes and flanking primers for all SNPs and Indels in a user-defined gene or region of the zebrafish genome. The SNPfisher site also contains links to display our SNP data in the UCSC genome browser.

Butler MG, Iben JR, Marsden KC, Epstein JA, Granato M, Weinstein BM. 2015. SNPfisher: tools for probing genetic variation in laboratory-reared zebrafish. *Development*. 142: 1542-52.

#### • [FishMap Zebrafish Genomics Knowledge base:](#)

A community resource for Zebrafish Genomics.

Bhartiya D, Maini J, Sharma M, Joshi P, Laddha SV, Jalali S, Patowary A, Purkanti R, Lalwani M, Singh AR, Chauhan R, Singh N, Bhardwaj A, Scaria V, Sivasubbu S. FishMap Zv8 update--a genomic regulatory map of zebrafish. *Zebrafish*. 2010 Jun;7(2):179-80.

Meli R, Prasad A, Patowary A, Lalwani MK, Maini J, Sharma M, Singh AR, Kumar G, Jadhav V, Scaria V, Sivasubbu S. FishMap: a community resource for zebrafish genomics. *Zebrafish*. 2008 Summer;5(2):125-30.

#### • [BLAST hits of Human Genes from Sanger sequence maintained by Zon lab:](#)

Each human gene annotated on the Santa Cruz site was compared by BLAST against the Sanger Institute zebrafish genome sequence. These data can be used to help identify zebrafish orthologues of human or Fugu genes.

- **Zebrafish BLAST server at MGH:**

MGH site provides a BLAST search of a gene or marker of interest to the assembled sequence contigs of the zebrafish genomic shotgun sequence from the Sanger project.

- **Tübingen Map of the Zebrafish Genome**
- **RH, meiotic and integrated maps of the genome:**

Centralized site for all meiotic and radiation hybrid maps of the zebrafish genome.

## Primers

- **MGH Web-based Tool for Automating Primer Selection:**

MGH site offers this tool which searches for a supercontig, scaffold, or finished BAC clone by name or accession number, derives all di- and tri- nucleotide repeat sequences in the supercontig, masks the sequence around the repeats for larger zebrafish repetitive sequences and then designs unique sequence primer pairs to amplify the simple sequence repeats. These are potentially polymorphic simple sequence repeats (SSRs) that can be used in genetic mapping/positional cloning projects. This site also offers a tool to [search for marker primer pairs by chromosome coordinates \(3MB chromosome segments max.\)](#)

- **qPrimerDB - qPCR Primer Database**

The qPrimerDB database (real-time quantitative PCR Primer Database), <https://biodb.swu.edu.cn/qprimerdb> is the most comprehensive qPCR primer database available to date, with a web front-end providing gene-specific and pre-computed primer pairs across 516 important organisms, including human, mouse, zebrafish, yeast, thale cress, rice, and maize. The qPrimerDB provides an interactive and information-rich web graphical interface to display search and BLAST results as table-based descriptions and associated links. In this database, we provide 9,273,735 of the best primer pairs for each gene based on primer pair coverage (PPC), as well as 68,653,054 alternative gene-specific primer pairs, which can be conveniently batch downloaded. We validated the specificity and efficiency of the qPCR primer pairs for 66 randomly selected genes in six different organisms through qPCR assays and gel electrophoresis. The qPrimerDB database represents a valuable, timesaving resource for gene expression analysis. This resource, which will be updated routinely.

## Microarrays

- **Zebrafish Affy Chip Annotation Project:**

The Trans-NIH Zebrafish Genome Initiative at Children's Hospital Boston aims to annotate zebrafish genes with human known genes for functional and comparative studies. This website provides a web service for retrieving annotated zebrafish genes represented on Affymetrix Zebrafish Genome Arrays and Roche-NimbleGen Zebrafish Expression Arrays to the research community.

- **Microarray expression at ZF-Espresso**
- **Zebrafish Oligonucleotide Arrays :**

The Norwegian Microarray Consortium offers their array products for academic use to customers world wide. One limitation is oligonucleotide arrays, which cannot be provided to users in the US and Japan, due to licensing conditions. A moderate license fee will be added to the base price for oligonucleotide arrays sold to users in the UK, France, Germany, Italy, Austria, Belgium, Sweden, Switzerland, Luxembourg, and The Netherlands.

## cDNA and EST Resources

- **Zebrafish Gene Collection (ZGC):**

NIH initiative for production of full length sequences and cDNA clones.  
1819 ZGC clones recloned in pENTR223.1 In May 2008, ZGC has added 1819 full length ZGC clones recloned into a Gateway Entry vector (pENTR223.1;) for expression analysis and other applications.  
These clones can be found by searching Entrez nucleotide for "NIH\_ZGC\_36".

- **DFCI Zebrafish Gene Index (ZGI):**

The DFCI Zebrafish Gene Index (ZGI) integrates research data from international zebrafish gene research projects and represents a non-redundant view of all zebrafish genes and data on their expression patterns, cellular roles, functions, and evolutionary relationships.

- **Zebrafish Inner Ear ESTs Database:**

This database consists of 4694 contigs or singletons obtained from ~18000 reads of cDNAs that were isolated from the microdissected otocysts of zebrafish embryos at 20-30 hours postfertilization, following subtraction with a pool of liver cDNAs from adult fishes.

- **Zebrafish EST Database at Singapore**

## Stock Centers

- **Zebrafish International Resource Center (ZIRC)**
- **European Zebrafish Resource Center (EZRC)**
- **China Zebrafish Resource Center (CZRC)**

## Mutagenesis Projects

- Targeted mutations using zinc-finger nucleases (ZFNs):

The [Lawson](#) and [Wolfe](#) labs have compiled resources relating to their recent [publication](#) on targeted gene inactivation using ZFNs: A list of plasmids and strains (available at [addgene](#)): <http://lawsonlab.umassmed.edu/PDFs/B1Hrec.pdf>

Help on getting started with bacterial 1-hybrid screens: <http://lawsonlab.umassmed.edu/PDFs/B1HStart.pdf>

An on-line search tool to identify and design library oligos: <http://pgfe.umassmed.edu/ZFPsearch.html>

- **Zebrafish Mutation Resource (Sanger Institute):**

This project seeks to identify, phenotype and distribute a large number of mutant zebrafish lines. Mutations are identified from populations of chemically mutagenised animals using PCR of specific exons followed by re-sequencing. Mutants are analysed for morphological and molecular differences and distributed to the community. This is an open resource and is now open to requests from the community. The [Zebrafish Mutation Resource Tracking page](#) lists the mutants generated by the Sanger [Zebrafish Mutation Resource](#) and the [ZF-HEALTH](#) consortium.

- **Hopkins lab large insertional mutagenesis screen database :**

Detailed information on the mutants resulting from the large insertional mutagenesis screen described in Amsterdam, et al. (2004) Identification of 315 genes essential for early zebrafish development. [Proceeding of the National Academy of Sciences 101: 12792-12797](#)

## Gene Trap, Enhancer Trap, Insertion Mutagenesis and CreloxP System Resources

- [Zfishbook](#): real-time database of gene-break transposon mutagenesis projects in the zebrafish (Steve Ekker laboratory).
- [Digital Fish](#): flip trap database and digital atlas (Scott Fraser laboratory).
- [CreZoo](#): database for zebrafish Cre driver and responder lines (Michael Brand laboratory).
- [zTrap](#): Tol2-mediated gene and enhancer traps that express GFP or Gal4 in specific cells, tissues, and organs (Koichi Kawakami laboratory).
- [Zebrafish Enhancer Trap Database](#): Gal4FF enhancer trap lines enriched for brain selective expression maintained at NICHD (Harold Burgess laboratory).
- [ZInC](#): zebrafish insertion mutation collection (Shuo Lin/Shawn Burgess laboratories).
- [Engert lab neural enhancer trap line resource](#): This neural enhancer trap line resource is a collection of Gal4 expressing enhancer trap lines which have been registered to the ZBrain for expression comparison and annotation. Full resolution stacks can be downloaded from the website and lines can be ordered from the Engert lab.

## TALENs/CRISPRs

- **Feng Zhang lab's CRISPR design:**

The CRISPR design tool is a web tool crafted to simplify the process of CRISPR guide selection in an input DNA sequence by (i) discovering possible offtargets genome-wide, (ii) highlighting guides with high target specificity, and (iii) flagging guides with numerous or genic offtargets in target genomes.

- **Michael Boutros lab's E-CRISP:**

E-CRISP is a software tool to design and evaluate target sites for use with the Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) system. The web application uses fast algorithms to identify sgRNA target sequences in any nucleotide sequence for use in CRISPR/Cas mediated genome editing. E-CRISP analyzes target specificity of the putative designs and assesses their genomic context (e.g. exons, transcripts, CpG islands). The design process incorporates different options of how CRISPR constructs can be used in experimental applications

- **Michael Boutros lab's E-TALEN:**

E-TALEN is a web service to design TALENs for introducing knock-out mutations, for endogenous tagging and targeted excision repair. The tool can be used to design TALENs against a single target but also for up to 50 target genes in parallel. E-TALEN guides the user through end-to-end de-novo design process for specific sequences or genomic loci. In addition, it can also be used to evaluate existing TALEN designs.

- **RGEN Tools Cas-OFFender:**

Identifies gRNA target sequences from an input sequence and checks for off-target binding.

- **Mojo Hand:**

A TALEN and CRISPR design tool provided by the Mayo Clinic.

- **C RISPRz :**

CRISPRz is a curated database of validated CRISPR targets in zebrafish maintained by Shawn Burgess' lab at NHGRI/NIH. The database contains target sites from the Burgess lab's large-scale mutagenesis project, as well as those reported in the literature. A search form allows users to search by various identifiers in order to determine if a validated CRISPR target exists for a gene of interest.

## Morpholinos

- **Morpholino Database**
- **Quality control testing of Gene Tools morpholinos**

## Pathways

- **Zebrafish Pathways at WikiPathways:**

This site allows browsing, editing, downloading and publishing of zebrafish pathways that are annotated for computational analysis and illustrated for human consumption.

- includes metabolic, signaling, regulatory and disease pathways
- browse and search for pathways by title or gene content
- download pathways as images for presentations
- download pathway for data visualization and analysis
- edit pathways using a built-in drawing tool
- publish new pathways based on your knowledge and research
- add literature references to important pathway components or steps
- discuss, share and collaborate on pathways with colleagues globally
- compete with others to curate the highest ranking pathways

WikiPathways is also accessible programmatically through a [web service API](#), which supports search, download, and data mapping.

## Neurobehavioral / Physiological Models

- **Zebrafish Neurophenome Project (ZNP):**

ZNP is a comprehensive resource for neurobehavioral and physiological data of adult zebrafish models. ZNP incorporates validated and curated data from work published in this field, to improve the accessibility of current knowledge to researchers interested in using adult zebrafish models.

## Gene Expression Resources

- **Tomo-seq Genome-wide 3D Expression Atlas**

Tomo-seq is a high-resolution genome-wide 3D atlas of gene expression in the zebrafish embryo at three developmental stages, produced using RNA Tomography.

## White Papers

- [The Zebrafish & Disease Project: Zebrafish as a model system to study and cure human diseases](#)