

## Module 2: Maps and Genome Sequences

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

#### Aims

- Introduce ZFIN BLAST.

#### Introduction

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:

Program: Nucleotide - Nucleotide Database: ZFIN cDNA Sequences

Query sequence (maximum of 50,000 letters):

FASTA or free-text format:

Set subsequence: From  To

☐ Search for short, nearly exact matches

Sequence ID:  (one or multiple delimited by ",")

Sequence Type: Nucleotide

Upload a free-text file:  Browse

Clear sequence BLAST

**Options:**

Expect: 1e-25 Word Size: 11 Matrix: —

Filter options for DNA Queries: ☒ Low complexity ☒ Poly-A's filter

Filter options for Protein Queries:

☐ SEG - filter low compositional complexity regions

☐ XNU - filter short-periodicity repeats

**Format:**

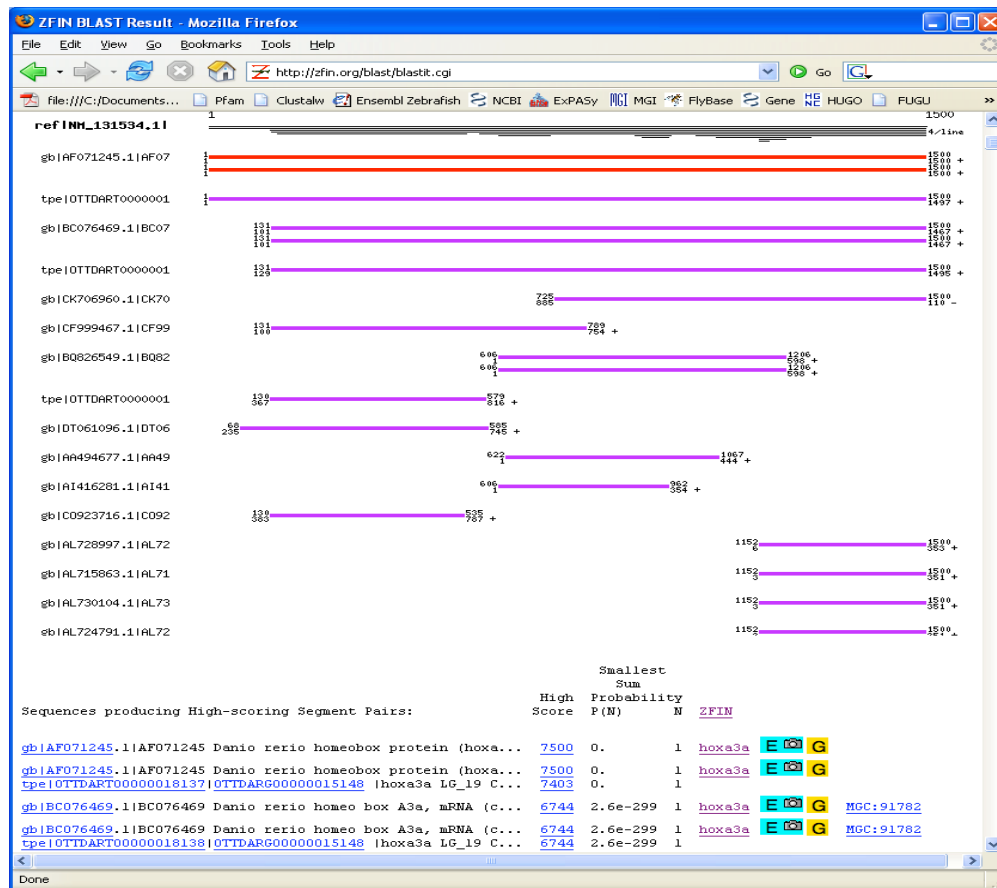
Show: ☒ Graphical Overview *limit of the first 50 alignments*

GenBank Zebrafish  
RefSeq Zebrafish mRNA  
ZFIN GenBank Sequences  
ZFIN cDNA Sequences  
ZFIN Genes with Expression  
ZFIN Morpholino Sequences  
ZFIN microRNA Sequences  
ZFIN Vega Transcripts  
Zebrafish mRNA  
EST Zebrafish  
Ensembl Zebrafish Transcripts  
TIGR Zebrafish Clusters  
Zebrafish DNA  
HTG Zebrafish  
GSS Zebrafish  
Zebrafish Trace Archive  
WGS Zebrafish  
RefSeq Zebrafish Protein  
UniProt / TrEMBL Zebrafish

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