

Module 3: Genes and Sequences (NCBI)

iv - What Information Is Available For a ZGC Clone and its Protein?

Aims

- Introduce options to mine data related to a ZGC clone
- Suggest alternative sources of functional information
- Provide examples of orthologs providing functional information
- Show examples of links on nucleotide and protein records

Introduction

There are a number of alternative ways to determine what is known about a particular Zebrafish Gene Collection (ZGC) cDNA clone. To quickly determine if there are any related records in any of the Entrez databases, you can click on 'Links' in the upper right hand corner of the GenBank page for your cDNA of interest.

Following these links you could answer these questions:

- Is the cDNA clone in a UniGene cluster or Gene record ?
- Are there similar proteins in zebrafish or orthologs whose function is known ? (see also Module 3_ii)
- Where is this cDNA known to be expressed, based on expression of similar ESTs ? (see also Module4_i)
- Does this cDNA encode a conserved domain ? (see also Module3_ii)
- Has this cDNA been annotated on the genome ? (see Module2_iii)

Exercises

1. Examine links to related resources found on a Entrez Nucleotide record.
2. Examine links to related resources found on a Entrez Protein record.
3. Examine links to related resources found on a UniGene record to other similarly expressed UniGene records or homologous UniGene records.

1. Examine links to related resources found on an Entrez Nucleotide record.

To begin, go to the Entrez Nucleotide (GenBank) record for the cDNA. (for example, BC071422:

<http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?db=Nucleotide&dopt=GenBank&val=47937949>). From the Nucleotide record, click on the Links pull down menu:

to view the set of available related links.

NCBI Nucleotide

Search Nucleotide for [] Go Clear

Limits Preview/Index History Clipboard Details

Display GenBank Send all to file

Range: from begin to end ☐ Reverse complemented strand

Features: ☐ SNP ☐ CDD ☒ MGC ☐ HPRD ☐ STS

1: BC071422. Reports Danio rerio zgc:8...[gi:47937949]

LOCUS BC071422 714 bp mRNA linear VRT 08-MAR-2005

DEFINITION Danio rerio zgc:86750, mRNA (cDNA clone MGC:86750 IMAGE:6899054), complete cds.

ACCESSION BC071422

VERSION BC071422.1 GI:47937949

KEYWORDS MGC.

SOURCE Danio rerio (zebrafish)

ORGANISM Danio rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (bases 1 to 714)

AUTHORS Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Sherman CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Ustin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Kettelman M, Madan A, Rodrigues S, Sanchez A, Whiting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U, Smallos DE, Schnerch A, Schein JE, Jones SJ and Marra MA.

CONSRM Mammalian Gene Collection Program Team

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

PUBMED 12477932

REFERENCE 2 (bases 1 to 714)

AUTHORS Director MGC Project.

TITLE Direct Submission

JOURNAL Submitted (01-JUN-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer

Links

- Gene
- Full text in PMC
- Related Sequences
- Protein
- PubMed
- Taxonomy
- LinkOut

2. Examine links to related resources found on an Entrez Protein record.

You could also start at the Entrez Protein page (AAH71422:

<http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?val=AAH71422.1>)

to view the set of related databases in the Links pull-down menu.

Navigate to related proteins records such as BLink or the conserved domains database (CDD)

NCBI Entrez Protein

Search Protein for [] Go Clear

Limits Preview/Index History Clipboard Details

Display GenPept Send all to file

Range: from begin to end Features: ☐ SNP ☐ CDD ☒ MGC ☐ HPRD ☐ STS

1: [AAH71422](#). Reports Zgc:86750 [Danio ...[gi:47937950]

BLink, Domains, MGC cDNA clone, Links

Links

- Gene
- Full text in PMC
- Related Sequences
- Domain Relatives
- Nucleotide
- OMIM
- PubMed
- Taxonomy
- LinkOut

LOCUS AAH71422 174 aa linear VRT 08-MAR-2005

DEFINITION Zgc:86750 [Danio rerio].

ACCESSION AAH71422

VERSION AAH71422.1 GI:47937950

DBSOURCE accession [BC071422.1](#)

KEYWORDS MGC.

SOURCE Danio rerio (zebrafish)

ORGANISM [Danio rerio](#)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

REFERENCE 1 (residues 1 to 174)

AUTHORS Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD, Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg B, Buetow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T, Max SI, Wang J, Hsieh F, Diatchenko L, Marusina K, Farmer AA, Rubin GM, Hong L, Stapleton M, Soares MB, Bonaldo MF, Casavant TL, Scheetz TE, Brownstein MJ, Ussdin TB, Toshiyuki S, Carninci P, Prange C, Raha SS, Loquellano NA, Peters GJ, Abramson RD, Mullahy SJ, Bosak SA, McEwan PJ, McKernan KJ, Malek JA, Gunaratne PH, Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SW, Villalón DK, Muzny DM, Sodergren EJ, Lu X, Gibbs RA, Fahey J, Helton E, Kettelman M, Madan A, Rodrigues S, Sanchez A, Whiting M, Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakesley RW, Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J, Schmutz J, Myers RM, Butterfield YS, Krzywinski MI, Skalska U, Smailus DE, Schnerch A, Schein JE, Jones SJ and Marra MA.

CONSRMT Mammalian Gene Collection Program Team

3. Examine links to related resources found on a UniGene record to other similarly expressed UniGene records or homologous UniGene records.

Query the UniGene database:

Click on links found in the drop-down Links menu to go to other UniGene clusters with similar expression patterns (**Co-expressed UniGene**) or homologous UniGene clusters determined by Homologous comparisons (**Homologous UniGene**)

The screenshot shows the NCBI UniGene search results page. The search query is "Dr.25683 OR Dr.24802". The results are displayed in a table with two entries:

Item	Gene	Sequence(s)
1	Dr.25683 ctslb: Cathepsin L, b	<i>Danio rerio</i> , 1287 sequence(s)
2	Dr.24802 hsps8: Heat shock protein 8	<i>Danio rerio</i> , 1404 sequence(s)

Arrows point from the "Links" column of the table to two separate "Links" pop-up windows. The left window shows a list of links: Gene, HomoloGene, Nucleotide, Protein, PubMed, Taxonomy, Homologous UniGene, and UniSTS. The right window shows a list of links: Gene, HomoloGene, Nucleotide, Protein, PubMed, Taxonomy, Co-expressed UniGene, Homologous UniGene, and UniSTS.