



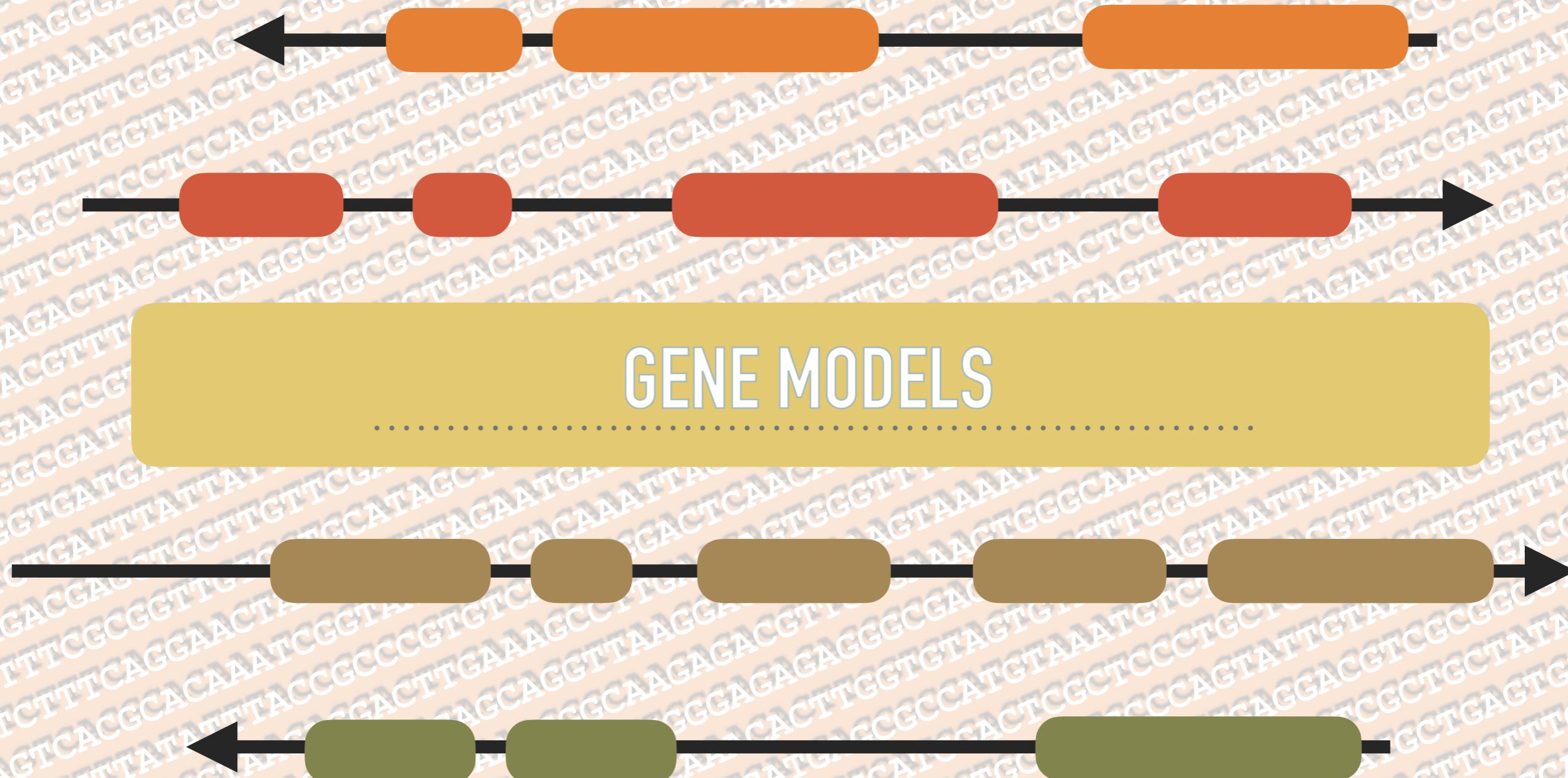
USING OPEN-SOURCE SOFTWARE TO SHARE GENOMIC DATA WITH THE COMMUNITY

Sofia Robb, Genomic Scientist, SIMR

ASSEMBLED GENOME

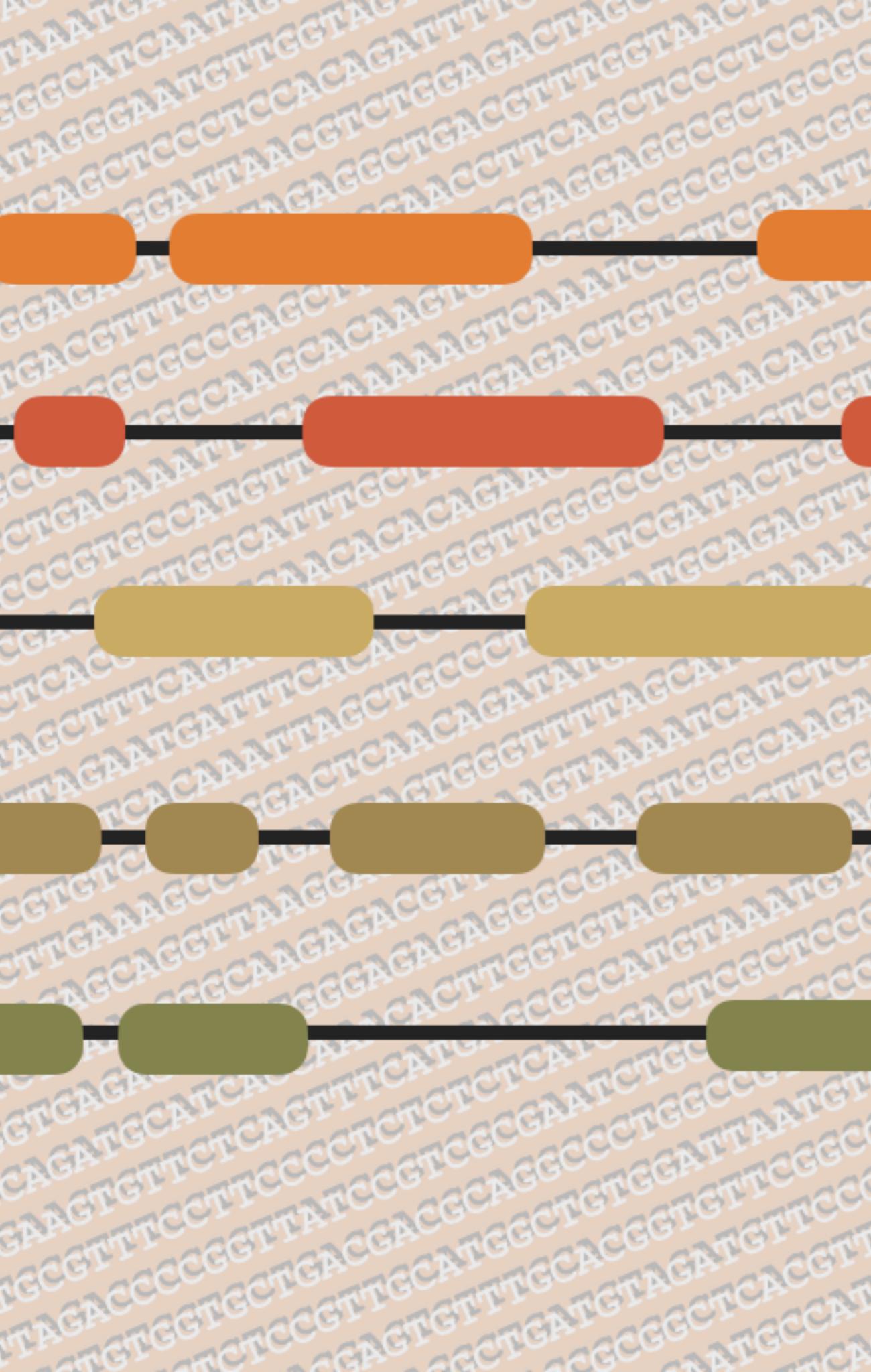
GENE MODELS

.....



AAGT
AAAAGTCAAATCC
CTGAGACTGTGGC
AAGCAAAGAAC
ATAACAGTCGAG
CACAGAAC
TTGGGTGGGG
CTGTAATCGATACT
GCAGAGTGG
AAAATGGCAG
ACCC
TCCC
AACAGATA
GTGGTTAGCA
AGTAAAATCATCT
AAGTGGCAAGAGAC
ACGT
AGAGAGGGGGA
CAGCTGGTAGTG
CAGCTAAATG
CAGCTCCCTCC
GAGCTGG
GAGCTGG
GAGCTGG
GAGCTGG





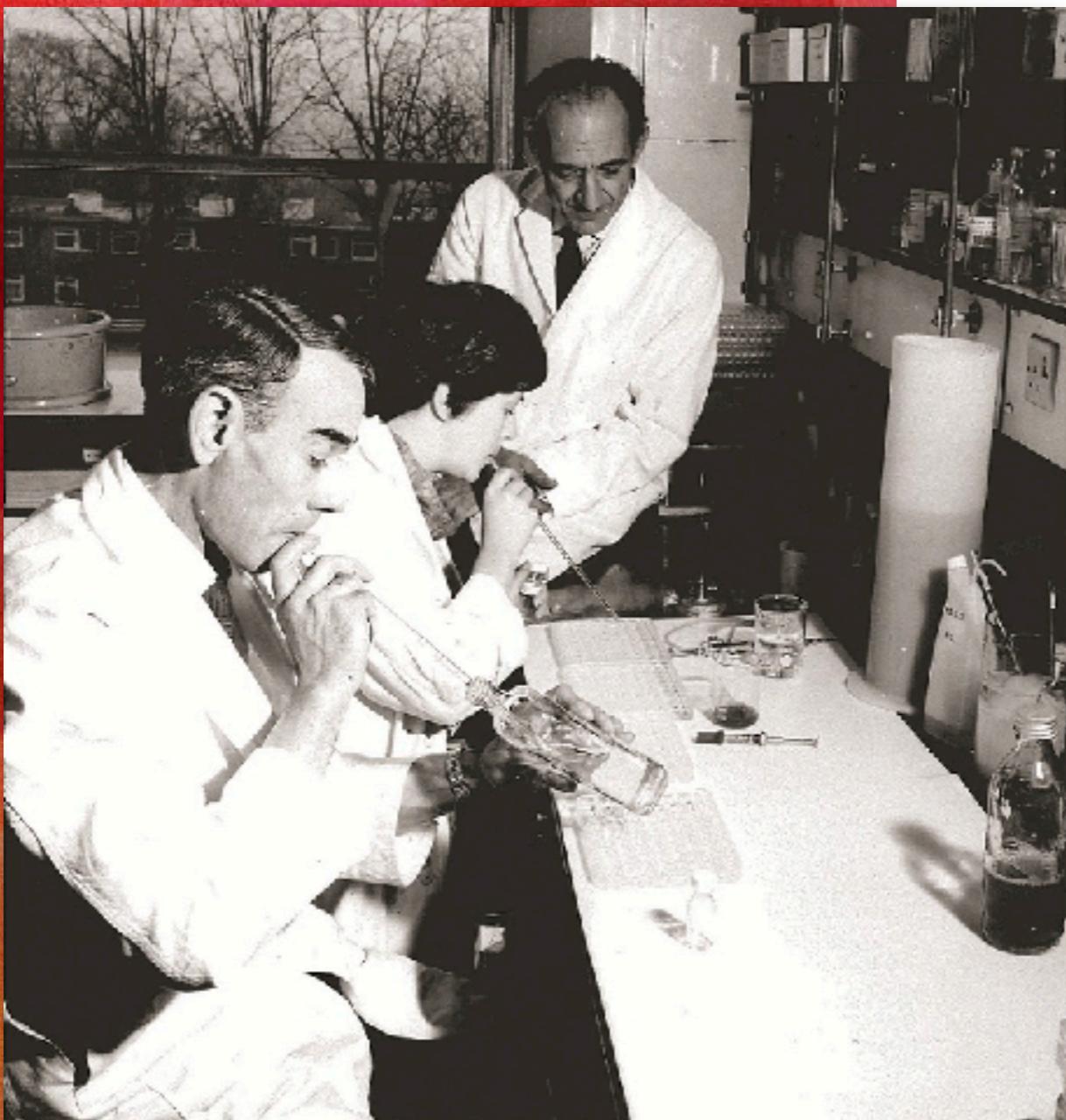
e!Ensembl



Phytozome

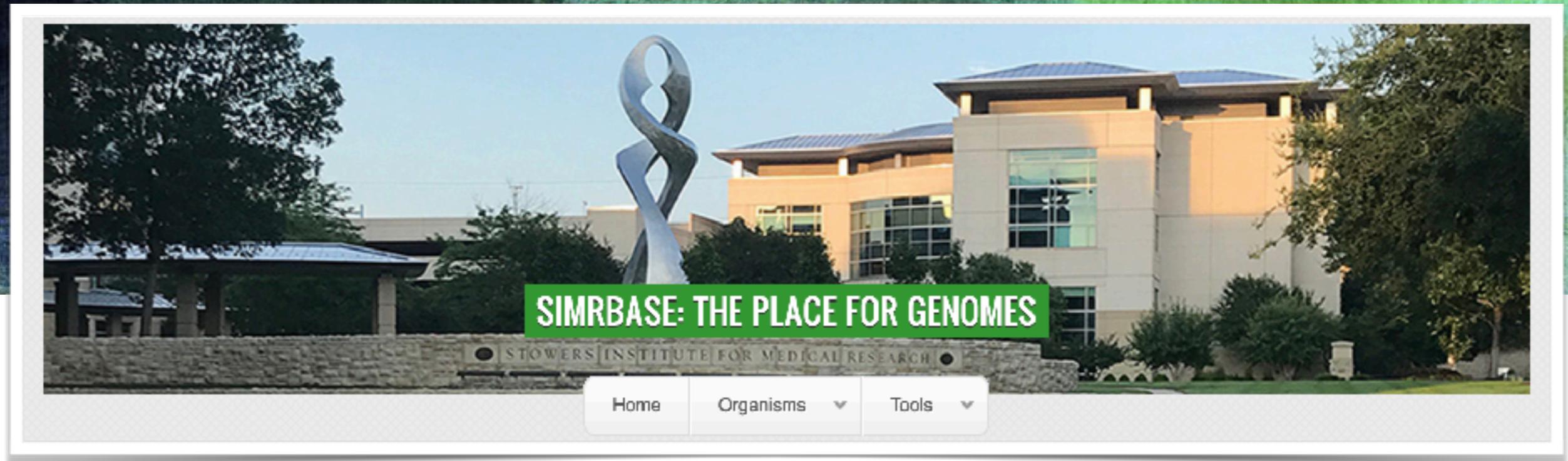
WHAT WE WANTED

.....



A platform to share
genomic data and
make it usable in a
useful way for bench
scientists

SIMRBASE



- Genomic Data Storage
- Genome Browser
- Gene Pages
- Gene Search Tools
- Sequence Similarity Search Tools
- Manual Gene Annotation

Stowers
Institute for
Medical
Research

SIMRB BASE ORGANISMS

Dugesia guanajuatiensis (Mexican planarian)



Our samples of *D. guanajuatensis* were collected in the state of Guanajuato in Mexico. They have unique motility and a "patchy" cilia pattern. Preliminary karyotyping suggests this planarian is primarily diploid, but some cells are multipliod.

Pomacea canaliculata (Apple Snail)





Pomacea canaliculata, common name the channeled apple snail, is a large freshwater snail with gills and an operculum, an aquatic gastropod mollusk. South American in origin, this species is considered an "Alien Species". It is also considered as about the alien species of gastropod in Europe -- [From Wikipedia](#).

alien species of gastropod in Europe -- From Wikipedia

Mus musculus (Mouse)





The house mouse (*Mus musculus*) is a small mammal with a pointed snout, small rounded ears, and a long tail. It is a common species of the genus *Mus*. Although it is often considered a pest, it has a symbiotic relationship with humans. The house mouse has been domesticated for thousands of years and is one of the most important model organisms in genetics. Its reference genome was sequenced in 2002.

Schmidtea mediterranea Sexual (Sexual Planaria)



Smed planarians are bilaterally symmetric, in differentiated cell types and organ systems. Smed animals that reproduce by fission, and obligate c (Newmark and Sánchez Alvarado, 2002; Newma newborn hatchlings grow and mature into adult worms without an i hatching, juveniles are sexually immature but otherwise possess a (Sánchez ... [more]

Astyanax mexicanus (Gavialfish)



 From Ensembl: The blind cave fish (*Astyanax* *Washington University*). The genome is 1Gb in length, which contains 19 chromosomes and 11,000 genes. The assembly represented here corresponds to GenBank accession number NC_007632. It was built using a mixed approach. Due to the lack of species specific sequence information, some scaffolds have been placed onto chromosomes.

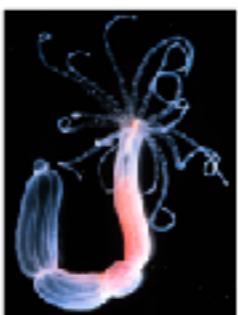
Danio rerio (Zebrafish)



The zebrafish (*Danio rerio*) is a tropical freshwater fish belonging to the minnow family (Cyprinidae) of the order Cypriniformes. Native to the Himalayan region, it is a popular aquarium fish, frequently sold under the trade name zebra danio. The zebrafish is also an important and widely used vertebrate model organism in scientific research, and was the first vertebrate to be cloned. It is particularly notable for its regenerative abilities, and has been modified by researchers to produce several transgenic strains.

The zebrafish is a derived member of the genus *Danio*, of the family Cyprinidae. It has a sister-group relationship with Da... [more]

Nematostella vectensis (Sea Anemone)



Cnidarians – the animal phylum that includes sea anemones, corals, jellyfish, and hydra – occupy a critical position in the tree of life. As the simplest Eumetazoans (animals with a tissue grade of organization) they are the most primitive animals with epithelial cells, neurons, stem cells, complex extracellular matrix, muscle fibers, and a fixed axis of symmetry. Ediacaran fossils with Cnidarian affinities are among the most ancient animal relics. As the principal architects and components of coral reefs, anthozoans (corals) are responsible for the maintenance of tremendous marine diversity. The sea anemone *Nematostella vectensis*, a mem... [more]

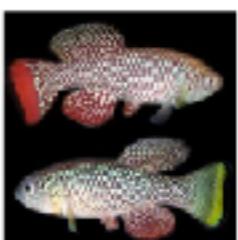
***Petromyzon marinus* (Sea Lamprey)**





The sea lamprey is a member of an ancient lineage that diverged from the vertebrate stem approximately 550 million years ago (MYA). By virtue of this deep evolutionary perspective, lamprey has served as a critical model for understanding the evolution of several conserved and derived features that are relevant to broad fields of biology and biomedicine. Studies have used lampreys to provide perspective on the evolution of developmental pathways that define vertebrate embryogenesis, vertebrate nervous and neuroendocrine systems, genome structure, immunity, clotting and others. These studies reveal aspects of vertebrate biology that have bee... [\[more\]](#)

Nothobranchius furzeri (Killifish)





A killifish is any of various oviparous (egg-laying) cyprinodontiform fish (including families Aplocheilidae, Cyprinodontidae, Fundulidae, Nothobranchiidae, Profundulidae, Rivulidae and Valenciidae). Altogether, there are some 1270 different species of killifish, the biggest family being Rivulidae, containing more than 320 species. Although killifish is sometimes used as an English equivalent to Cyprinodontidae, some species belonging to that family have their own common names, such as the pupfish and the mummichog. The origin of the name is not certain, but is likely to have come from the Dutch for a kill.

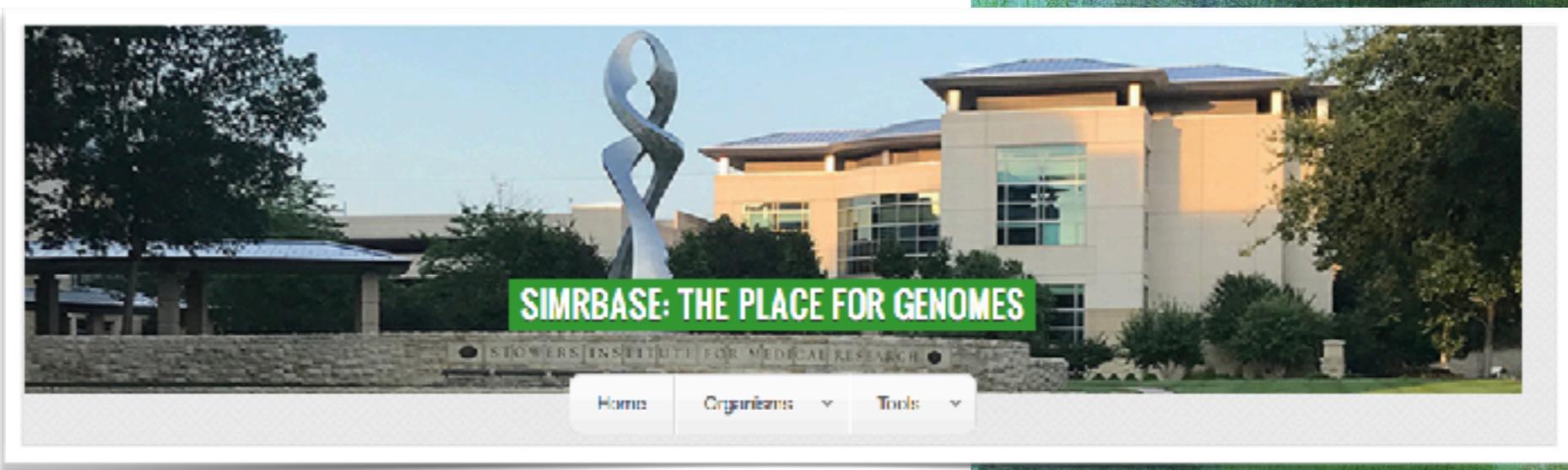
Because of living in ephemeral water... [more]

TOUR OF SIMRBASE

UNDER THE HOOD

TOUR OF SIMRBASE

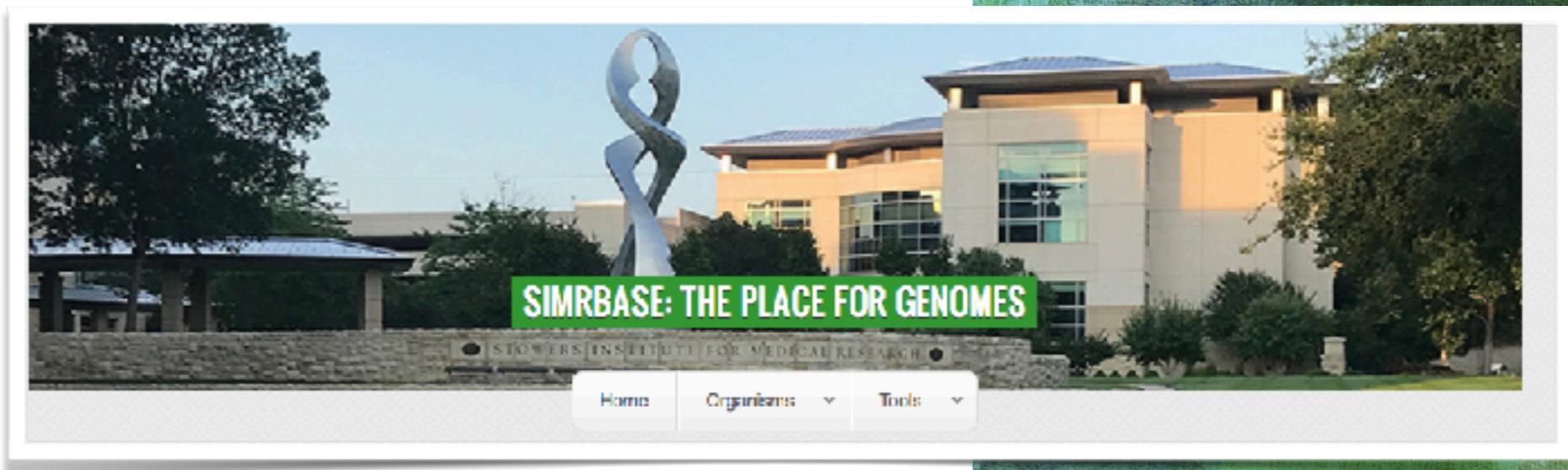
.....



- Genome Browser
- Gene Searches & Gene Pages
- User BLAST Interface
- Manual Gene Curation

TOUR OF SIMRBASE

.....

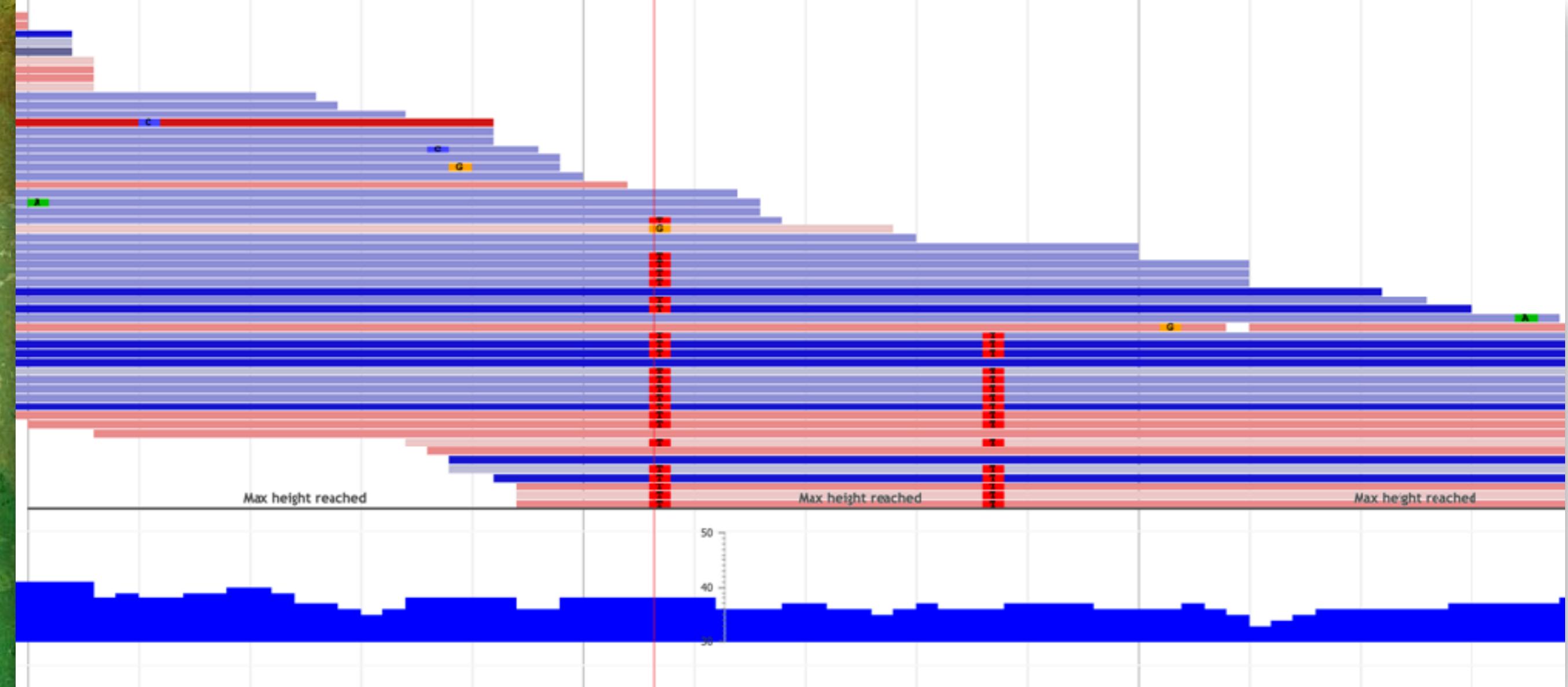


- **Genome Browser**
- **Gene Searches & Gene Pages**
- **User BLAST Interface**
- **Manual Gene Curation**

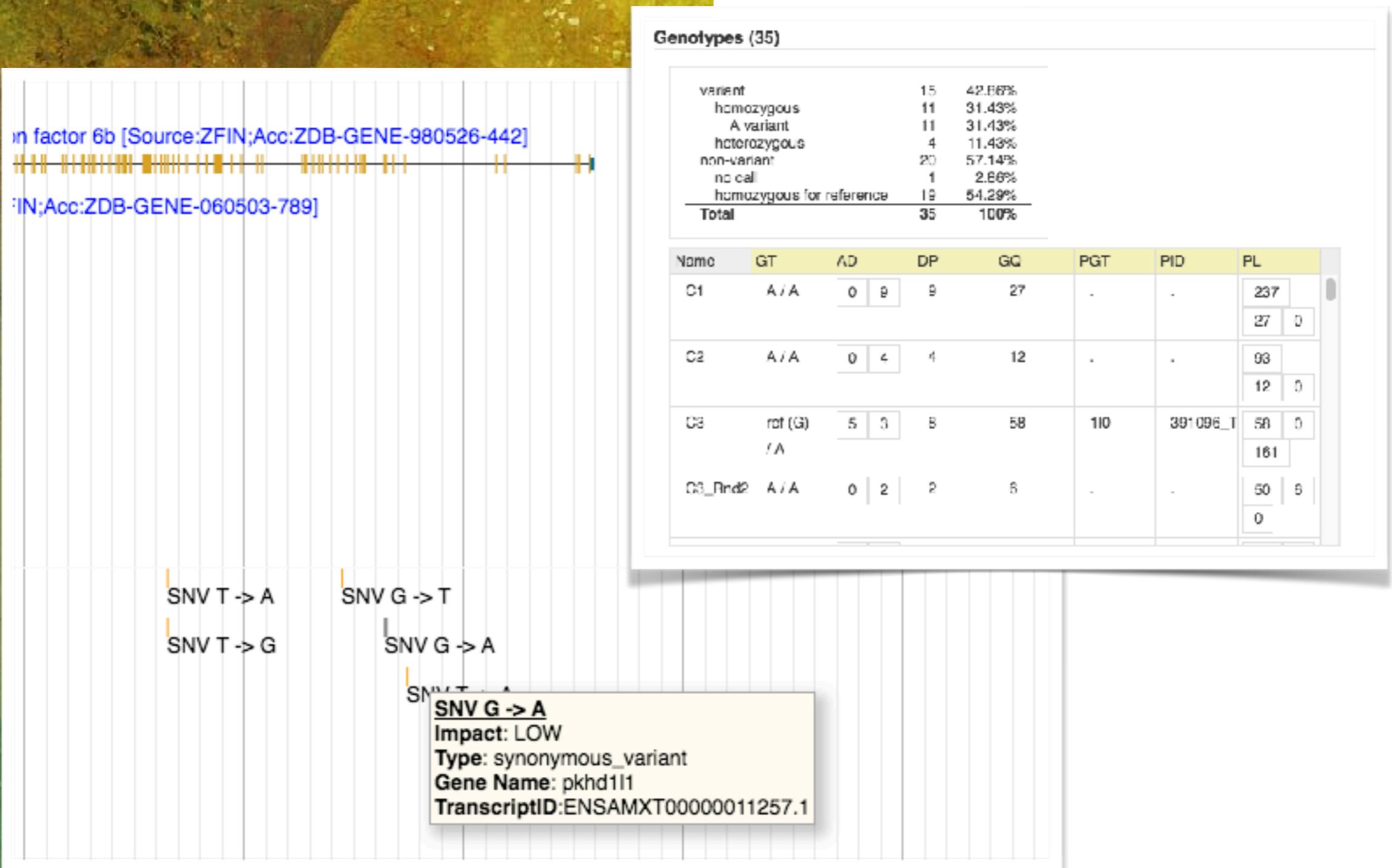
GENOME BROWSER



VISUALIZE NUCLEOTIDE DIFFERENCES



VISUALIZE SNPs



SEQUENCE RETRIEVAL

The screenshot shows a bioinformatics interface for sequence retrieval. On the left, a genome browser track displays two genes: Nv3_00000250-RA (Similar to TAF15: TATA-binding protein-as) and Nv3_00000251-RA (Similar to Armc5: Armadill). A context menu is open over the first gene, with the "View Feature Sequence" option highlighted in blue. On the right, a "FeatureSequence Viewer" window is open, showing options for highlighting and viewing different genomic features: CDSs, UTRs, Introns, Others, Three_prime_UTRs, Upstream, and Downstream. The "Others" section is currently selected. Below these controls is a large text area displaying the sequence for Nv3_00000250-RA.

Nv3_00000250-RA
Similar to TAF15: TATA-binding protein-as

Nv3_00000251-RA
Similar to Armc5: Armadill

View details
Highlight this mRNA
View Feature Sequence
View CDS Sequence

Nv3_00000250-RA
Similar to TAF15: TATA-binding protein-as

FeatureSequence Viewer

| Feature Type | Highlight: | Lowercase | Action |
|------------------|------------|-----------|---------------|
| CDSs | Highlight: | Lowercase | Hide |
| UTRs | Highlight: | Lowercase | Show |
| Introns | Highlight: | Lowercase | Show |
| Others | Highlight: | Lowercase | Show |
| Three_prime_UTRs | Highlight: | Lowercase | Show |
| Upstream | Highlight: | Lowercase | Select View ▾ |
| Downstream | Highlight: | Lowercase | Select View ▾ |

```
>Nv3_00000250-RA
ATGCCCAAGGACATCCAGCTTGGGAGGGAGAATTGGGATTCCTCTGGCTATAGCCAAGGAGGCTA
TCCCACTCTGGGGATATACCACTGTGGCTATTCCCACTACCCAGGGAGGGTATCAACAAAC
GATACAAAGCAAGACAGAAGGTACTTCATATGGCGGTACCAAACCAACACGTCTCATCTACTC
CCTCAAGATAACTCTGCCCTACCAGSCTACCAACASCTGGGACTTATCAGGCAGGTTCTGCCCTA
CAGCCTTCCTCAGGCTACCAGCAGTCACGGCTGTGACTCTAGGGGGCTACCACTGGGGGGGT
GGTGGCAGCTCTGACACGCTTCCGAAGTAGAGGTGTTAGTGGAGGCAGCTCCCATGGTAC
AGTGGGGCTCTATGGTAGCGACUGGGCAGCTCTATGGCAGGGACACAGGGGGCTCATATGC
GACAGGGGAGGTTCATATGGCGGAGACAGGGGGGGAGGTGACAGAGGTGGTGACCGTGGTGGCA
GGACGTGGCAGGGTACCCAGACCCAGCAGTGGATACTCAAATAGTGGTCAAGTTATGGCGC
GGACGTGGAGGGAGATGATCAAAGCAATACTCCCAGACACAATTGGTATCCAACATGAGCC
TCAACAGAGGAAGACATCAAGGGACTGTTGGCTCAATAGGAATTATAAAGATTGACAAAAAGCT
AAGCCAAGATCTGGATCTACAAGCATCCAGACGGTCATCAAAGGGAGAGTGTACGGTCACAT
GACCCCCAACAGCCTCTGCAGCTATTGAGTGGTTAATGGTAAAGACTCATGGACAAAGCAT
GTTGAACCTGCTGAAGCAAGAAACTCCCAAGTTGGTGTGCGGGGGAGGGCAGGGTGGAGGAGG
GGTGGCTTGAAACGTGGTGGTTCCGTGGTAGCAGAGGTGGTGGCGGCAGAGGAGGTGGTGGT
GCCGGAGGGAGGTGGAGATCGTTCTGGTACTGGACTGTCCATCATGTGTAATATGAACATTG
CGTGATCGCTGCAACCGTTGTCAAGGAGTCACGCCCTGGTGGTGGAGGCCGTGGTGGTGGCGC
GGTGGTGGGGGGGGAGGAGGAAGAAGGGGGGGTGGTGGGGGGTACAGAGGGAGGGGGTGGCGC
```

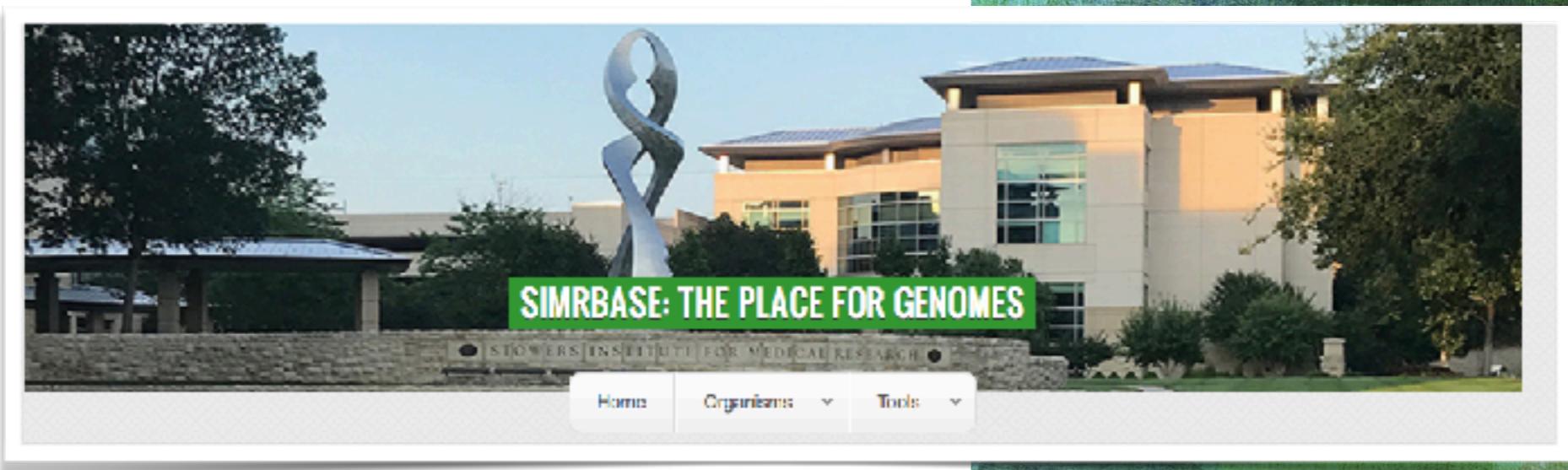
TRACK SELECTOR

Select Tracks

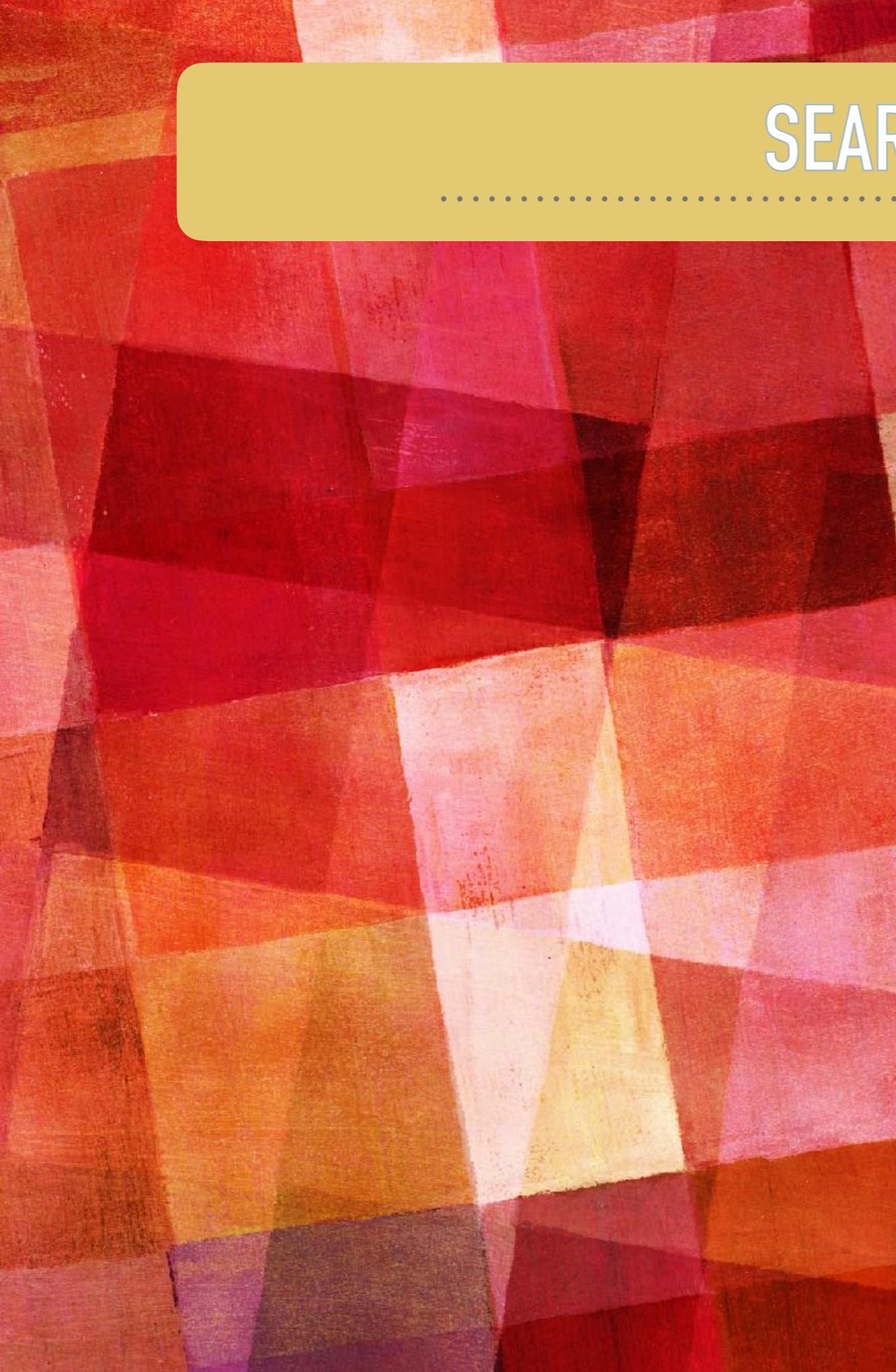
[Help](#)

| 44 tracks | | | | | | |
|--------------------------|-----------------|----------------------|--|---|---------------------|---|
| | Category | Technique | Name | Experiment | Developmental-stage | Summary |
| <input type="checkbox"/> | Alignments | Protein Alignment | Ipsilon Domains | MAKER + Interproscan | ... | MAKER run by Carson Holt 2016 |
| <input type="checkbox"/> | Alignments | Nucleotide Alignment | Repeats | MAKER + RepeatMasker | ... | MAKER run by Carson Holt 2016 |
| <input type="checkbox"/> | Alignments | Protein Alignment | Protein Alignments | MAKER | ... | MAKER run by Carson Holt 2016 |
| <input type="checkbox"/> | Alignments | Nucleotide Alignment | Transcript Alignments | MAKER + Exonerate/est2genome | ... | MAKER run by Carson Holt 2016 |
| <input type="checkbox"/> | Alignments | Nucleotide Alignment | Lethenteron camtschaticum HOX Alignments | Jeramiah Smith | ... | Lethenteron camtschaticum HOX Alignments |
| <input type="checkbox"/> | Alignments | DNAseq | Genotype-specific Regions | DNA from Blood and Sperm | germline | Relative coverage of DNA sequence from sperm and blood (standardized log2 ratio) calculated using DiffCover and DNACopy |
| <input type="checkbox"/> | Alignments | DNAseq | DNA Re-seq Sperm | DNA from Blood and Sperm; Illumina Sequencing of Blood DNA | ... | ... |
| <input type="checkbox"/> | Alignments | DNAseq | DNA Re-seq Blood | DNA from Blood and Sperm; Illumina Sequencing of sperm DNA | ... | ... |
| <input type="checkbox"/> | Alignments | DNAseq Reads | DNA Re-seq Sperm BAM | DNA from Blood and Sperm; Illumina Sequencing of Blood DNA | ... | ... |
| <input type="checkbox"/> | Alignments | DNAseq Reads | DNA Re-seq Blood BAM | DNA from Blood and Sperm; Illumina Sequencing of sperm DNA | ... | ... |
| <input type="checkbox"/> | Assembly | Sequence | Scaffolds | Dovetail | ... | Previous assembly placement in new Dovetail assembly |
| <input type="checkbox"/> | GC Content | ... | GCContentXY | ... | ... | ... |
| <input type="checkbox"/> | Gene Expression | RNAseq | C1 | Characterization of Somatically-Eliminated Genes During Development of the Sea Lamprey (<i>Petromyzon marinus</i>). | D1 (~Tahara 7) | Emeryonic Day 1 expression from Bryant et al. Molecular Biology and Evolution, 33:2337-2344. Mapped using TopHat v2.1.1 |
| <input type="checkbox"/> | Gene Expression | RNAseq | C2.5 | Characterization of Somatically-Eliminated Genes During Development of the Sea Lamprey (<i>Petromyzon marinus</i>). | D2.5 (~Tahara 11) | Emeryonic Day 2.5 expression from Bryant et al. Molecular Biology and Evolution, 33:2337-2344. Mapped using TopHat v2.1.1 |
| <input type="checkbox"/> | Gene Expression | RNAseq | C2 | Characterization of Somatically-Eliminated Genes During Development of the Sea Lamprey (<i>Petromyzon marinus</i>). | D2 (~Tahara 10) | Emeryonic Day 2 expression from Bryant et al. Molecular Biology and Evolution, 33:2337-2344. Mapped using TopHat v2.1.1 |
| <input type="checkbox"/> | Gene Expression | RNAseq | C3 | Characterization of Somatically-Eliminated Genes During Development of the Sea Lamprey (<i>Petromyzon marinus</i>). | D3 (~Tahara 12) | Emeryonic Day 3 expression from Bryant et al. Molecular Biology and Evolution, 33:2337-2344. Mapped using TopHat v2.1.1 |
| <input type="checkbox"/> | Gene Expression | RNAseq | ... | ... | ... | Emeryonic Day 4 expression from Bryant et al. Molecular Biology and Evolution, 33:2337-2344. Mapped using TopHat v2.1.1 |

TOUR OF SIMRBASE



- Genome Browser
- Gene Searches & Gene Pages
- User BLAST Interface
- Manual Gene Curation



SEARCHES

.....

- Gene Name/ID
- Keyword
- BLAST
- Genomic Location
- Expression Level
- Tissue or Region of Expression
- Gene Groups

GENE SEARCH

.....

GENE SEARCH

Search By: Gene ID/Name

Contains any word

Examples: MycD, SMED30002427. For multiple search terms use the "Contains any word" filter and separate with spaces or commas.

Search By: BLAST Best Hit

HDAC

Examples: HDAC

Search By: GO Term

Contains

Examples: protein binding, membrane. GO terms are displayed on the linked Gene page in the "Annotated Terms" Section.

Search By: Domains

Contains

Apply Reset

SEARCH RESULTS

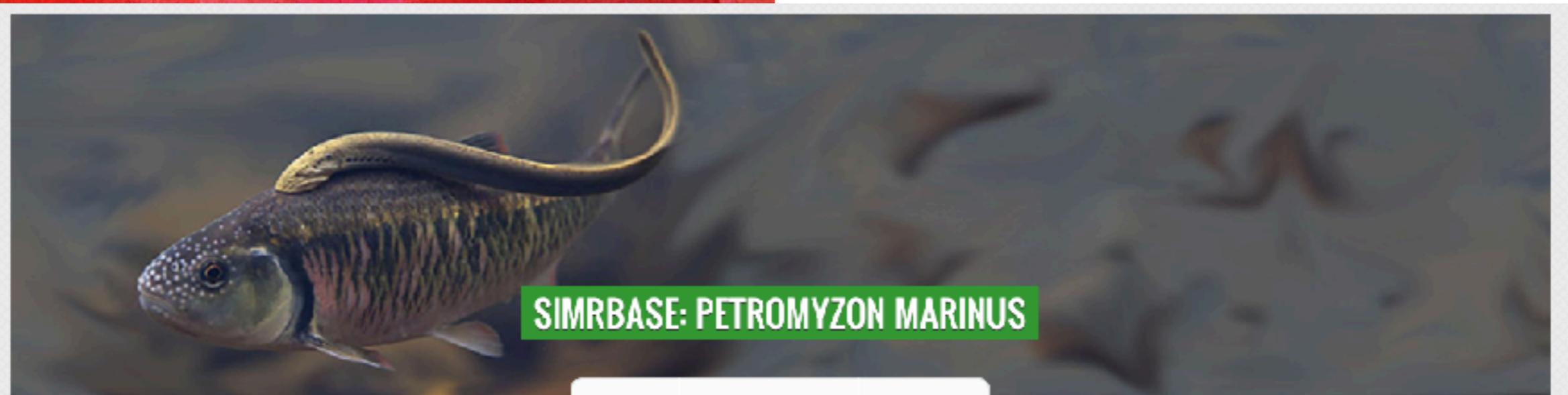
Download Results: [FASTA](#) | [CSV](#)

Results Total: 16. Displaying 1 - 16.

| Organism | Gene ID | Name | Alias | UniProt Hit | E-Value |
|-----------|---|---------|-------|---|--------------|
| Killifish | maker-sgr05-augustus-gene-169.2-mRNA-1 | hdac9 | | Histone deacetylase 9-B OS=Danio rerio GN=hdac9b PE=2 SV=1 | 2.716e-08 |
| Killifish | maker-sgr16-augustus-gene-360.11-mRNA-1 | hdac6;1 | | Histone deacetylase 6 OS=Homo sapiens GN=HDAC6 PE=1 SV=2 | 3.53262e-119 |
| Killifish | maker-sgr14-augustus-gene-21.5-mRNA-1 | hdac4;2 | | Histone deacetylase 4 OS=Mus musculus GN=Hdac4 PE=1 SV=1 | 1.43464e-176 |
| Killifish | augustus_masked-sgr15-processed-gene-355.0-mRNA-1 | hdac7;1 | | Histone deacetylase 7 OS=Mus musculus GN=Hdac7 PE=1 SV=2 | 0 |
| Killifish | maker-sgr16-augustus-gene-127.4-mRNA-1 | hdac5;2 | | Histone deacetylase 5 OS=Cricetulus griseus GN=HDAC5 PE=2 SV=1 | 1.22793e-131 |
| Killifish | maker-sgr14-augustus-gene-21.4-mRNA-1 | hdac4;3 | | Histone deacetylase 4 OS=Gallus gallus GN=HDAC4 PE=2 SV=1 | 0 |
| Killifish | maker-scaffold01460-augustus-gene-0.2-mRNA-1 | hdac3 | | Histone deacetylase 3 OS=Tetraodon nigroviridis GN=hdac3 PE=3 SV=1 | 0 |
| Killifish | augustus_masked-sgr06-processed-gene-498.2-mRNA-1 | hdac4;1 | | Histone deacetylase 4 OS=Rattus norvegicus GN=Hdac4 PE=1 SV=2 | 2.98052e-71 |
| Killifish | maker-scaffold00244-augustus-gene-0.5-mRNA-1 | hdac2;1 | | Probable histone deacetylase 1-B OS=Xenopus laevis GN=hdac1-b PE=1 SV=1 | 0 |
| Killifish | maker-sgr16-augustus-gene-127.3-mRNA-1 | hdac5;1 | | Histone deacetylase 5 OS=Cricetulus griseus GN=HDAC5 PE=2 SV=1 | 1.55978e-98 |
| Killifish | maker-sgr10-augustus-gene-416.0-mRNA-1 | hdac7;2 | | Histone deacetylase 7 OS=Homo sapiens GN=HDAC7 PE=1 SV=2 | 4.38746e-166 |
| Killifish | maker-sgr05-augustus-gene-169.1-mRNA-1 | hdac9b | | Histone deacetylase 9-B OS=Danio rerio GN=hdac9b PE=2 SV=1 | 7.44894e-108 |
| Killifish | maker-sgr13-augustus-gene-119.8-mRNA-1 | hdac10 | | Histone deacetylase 10 OS=Homo sapiens GN=HDAC10 PE=1 SV=1 | 1.91664e-148 |
| Killifish | maker-sgr16-augustus-gene-365.17-mRNA-1 | hdac2;2 | | Probable histone deacetylase 1-B OS=Xenopus laevis GN=hdac1-b PE=1 SV=1 | 0 |
| Killifish | maker-sgr06-augustus-gene-214.10-mRNA-1 | hdac12 | | Uncharacterized protein SYNPCC7002_A1628 OS=Synechococcus sp. (strain ATCC 27254 / PCC 7002 / PR-6) GN=SYNPCC7002_A1628 PE=8 SV=2 | 1.78533e-06 |

GENE PAGES

.....



SIMRBASE: PETROMYZON MARINUS

Home Organisms Tools

PXF-1, PMZ_0043432-RA (mRNA) PETROMYZON MARINUS

Overview

Alignments

Analyses

Annotated Terms

Cross References

Homology

InterPro

Properties

Relationships

Sequences

Overview

| | |
|-----------------|---|
| Name | pxf-1 |
| Unique Name | PMZ_0043432-RA |
| Type | mRNA |
| Organism | <i>Petromyzon marinus</i> (Sea Lamprey) |
| Sequence length | 255 |

ANNOTATED TERMS

Overview

Alignments

Analyses

Annotated Terms

Cross References

Homology

InterPro

Properties

Relationships

Sequences

Annotated Terms

The following terms have been associated with this mRNA:

Vocabulary: INTERPRO

| Term | Definition |
|------|------------|
|------|------------|

| | |
|-----------|--------------|
| IPR018490 | cNMP-bd-like |
|-----------|--------------|

| | |
|-----------|---------------------|
| IPR014710 | RmIC-like_jellyroll |
|-----------|---------------------|

Vocabulary: Molecular Function

| Term | Definition |
|------|------------|
|------|------------|

| | |
|------------|--|
| GO:0017034 | Rap guanyl-nucleotide exchange factor activity |
|------------|--|

Vocabulary: Biological Process

| Term | Definition |
|------|------------|
|------|------------|

| | |
|------------|---|
| GO:0042127 | regulation of cell population proliferation |
|------------|---|

| | |
|------------|--|
| GO:0043547 | positive regulation of GTPase activity |
|------------|--|

HOMOLOGY

Overview

Alignments

Analyses

Annotated Terms

Cross References

Homology

InterPro

Properties

Relationships

Sequences

Homology

The following BLAST results are available for this feature:

- BLAST of pxf-1 vs. Ensembl C.elegans
- BLAST of pxf-1 vs. Ensembl Fly
- BLAST of pxf-1 vs. Ensembl Mouse
- BLAST of pxf-1 vs. UniProt
- BLAST of pxf-1 vs. Ensembl Xenopus
- BLAST of pxf-1 vs. Ensembl Zebrafish
- BLAST of pxf-1 vs. Ensembl Human
- BLAST of pxf-1 vs. Ensembl Lamprey
- BLAST of pxf-1 vs. Ensembl Yeast
- BLAST of pxf-1 vs. Ensembl Cavefish
- BLAST of pxf-1 vs. Schmidtea mediterranea smes_v2_hcnnf_SMESG

BLAST of pxf-1 vs. Ensembl C.elegans

Analysis Date: 2019-10-09 (Petromyzon marinus PMZ_v3.1 BLASTX_Celegans_e198)

Total hits: 4

ZOOM | x 1 POSITION | 0

Sequence

pxf-1

Expect = 2.90e-11 / Id = 54.55

pxf-1

Expect = 2.90e-11 / Id = 54.55

pxf-1

Expect = 3.26e-11 / Id = 54.55

pxf-1

Expect = 3.26e-11 / Id = 54.55

20 40 60 80 100 120 140 160 180 200

Match Name E-value Identity Description

| | | | | |
|-------|-----------|-------|---|--------|
| pxf-1 | 2.905e-11 | 54.55 | Rap guanine nucleotide exchange factor [Source:UniProt] | [more] |
| pxf-1 | 2.905e-11 | 54.55 | Rap guanine nucleotide exchange factor [Source:UniProt] | [more] |

DOMAINS

Overview

Alignments

Analyses

Annotated Terms

Cross References

Homology

InterPro

Properties

Relationships

Sequences

InterPro

Analysis Name: **Petromyzon marinus PMZ_v3.1 Interproscan**

Date Performed: 2017-04-04



| IPR Term | IPR Description | Source | Source Term | Source Description | Alignment |
|-----------|--------------------------------|-------------|-------------------|--------------------------|---|
| IPR014710 | RmIC-like jelly roll fold | GENE3D | G3DSA:2.60.120.10 | | coord: 13..51 e-value: 3.5E-5 score: 22.6 |
| IPR018490 | Cyclic nucleotide-binding-like | SUPERFAMILY | SSF51206 | cAMP-binding domain-like | coord: 15..51 |

SEQUENCES

Overview

Alignments

Analyses

Annotated Terms

Cross References

Homology

InterPro

Properties

Relationships

Sequences

Sequences

The following sequences are available for this feature:

- mRNA sequence
- mRNA from alignment at scaf_00012:11464069..11476187+
- coding sequence from alignment at scaf_00012:11464069..11476187+

mRNA sequence

```
>PMZ_0043432-RA ID=PMZ_0043432-RA|Name=pxf-  
1|organism=Petromyzon marinus|type=mRNA|length=255bp  
ATGGTGAGGAGAGGGTCTGGTATGGTGAAGTGAGGGTCTGGTCTCTGTG  
CAAGCTGGTTCGCTACGAGCAGCACGAAGCCAATACGTCCATACTTCC  
CTGAGGATGTGAGCAGCTGCTGGTACATCCTGCTGTCGGGCTCAGTGTTC  
ATCGACGGCTCCATGTTCCCTGCCTCGCAGCAGGTCACTCGACCGCACCCC  
TCGGGGATCGCGGGCCGCGGTCAACGCCTTGCCCTTCGCGCGCACTCGA  
CGTGA
```

CUSTOM EXPRESSION LEVEL SEARCHES

SEARCH MOUSE CRANIAL FACIAL ATLAS

Search for genes by Ensembl Gene Name and/or ID or by expression values in tissues during development. For example, find all genes that have a FPKM of >=5 in E8.5 Mesoderm and E9.5 Epidermal Ectoderm. Download links for the results in CSV or FASTA format can be found below the results.

Search by Gene Name/ID

E8.5 Caudal Brain Neural Epithelium
Is greater than or equal to

Search by Gene Description

E8.5 Mesoderm
Is greater than or equal to

Search by Stage E8.5

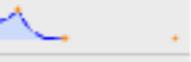
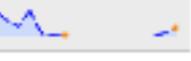
E8.5 Floor Plate
Is greater than or equal to

Search by Stage E9.5

E8.5 Non Floor Plate
Is greater than or equal to

Search by Stage E10.5

Apply **Reset**

| Ensembl ID | Gene Name | Description | FPKM |
|---------------------|-----------|---|---|
| ENSMUSG000000000805 | Car4 | carbonic anhydrase 4 [Source:MGI Symbol;Acc:MGI:1008674] |  |
| ENSMUSG00000061524 | Zic2 | zinc finger protein of the cerebellum 2 [Source:MGI Symbol;Acc:MGI:1066679] |  |

Results Total: 2. Displaying 1 - 2.

Download Results: [FASTA](#) | [CSV](#)

GENERAL TISSUE EXPRESSION SEARCHES

Search By: Tissue or Expression Pattern

Contains



See [Glossary](#) for Planarian Tissue Terms

Examples: muscle, neoblast.

Search By: BLAST Best Hit

Examples: HDAC

Search By: GO Term

Contains



Examples: protein binding, membrane. GO terms are displayed on the linked Gene page in the "Annotated Terms" Section.



PLANOSPHERE

<https://planosphere.stowers.org>

Embryonic origin of adult stem cells required
for tissue homeostasis and regeneration.
Davies et. al., 2017

TISSUE PAGE



DIGESTIVE SYSTEM

| Smed ID | Accession | Name | Alias | Enriched during stage(s) | Tissue/Pattern | Images |
|--------------|------------|------------------|----------|--|--|--------|
| SMED30027428 | AFJ24799.1 | forkhead box A-1 | foxA1 | Stage 3, Stage 4, Stage 5, Stage 6, Stage 7, Stage 8 | embryonic digestive system, definitive pharynx, definitive pharynx progenitor, digestive system, temporary embryonic pharynx | |
| SMED30007290 | AFF18488.1 | GATA456 | gata456a | Stage 2, Stage 5, Stage 6, Stage 7, Stage 8 | embryonic digestive system, digestive system, temporary embryonic pharynx, definitive gut, gamma neoblast | |
| SMED30005748 | AFF18489.1 | HNF4 | hnf-4 | Stage 2, Stage 5, Stage 6, Stage 7, Stage 8 | embryonic digestive system, digestive system, temporary embryonic pharynx, definitive gut, gamma neoblast | |

GENE GROUPS

INNEXIN, GSP_009731 (mRNA) GIRARDIA SP.

Overview

Analyses

Gene Groups

Homology

Sequences

Gene Groups

innixin is similar in sequence to the genes of this group: GG549

| Gene Name | Gene ID |
|-------------|-------------|
| SMU15000918 | SMU15000918 |
| SMU15003648 | SMU15003648 |
| SMU15003649 | SMU15003649 |
| Ddo_019521 | Ddo_019521 |
| innixin | Ddo_023720 |
| Gsp_005528 | Gsp_005528 |
| innixin | Gsp_009731 |
| Pgr_014738 | Pgr_014738 |
| Pgr_015079 | Pgr_015079 |



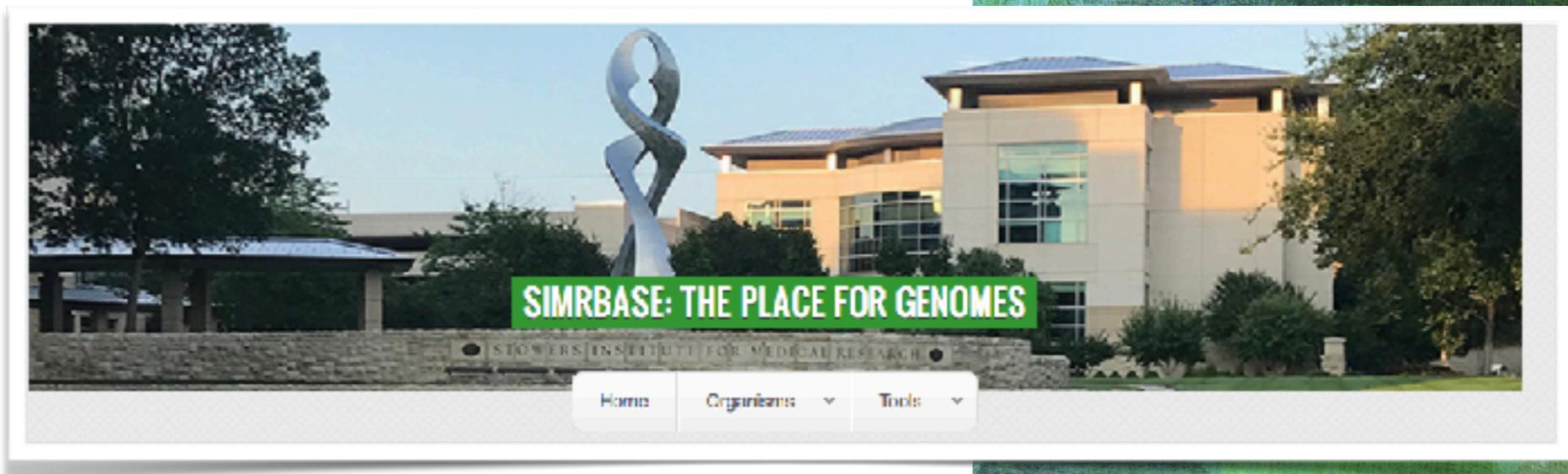
<https://cuttingclass.stowers.org>

Hands-On Classroom Activities for Exploring
Regeneration and Stem Cell Biology with Planarians.

Accorsi et. al., 2017

TOUR OF SIMRBASE

.....



- Genome Browser
- Gene Searches & Gene Pages
- User BLAST Interface
- Manual Gene Curation

BLAST DATABASES

PROTEIN TO NUCLEOTIDE BLAST (TBLASTN)

► See Results from a Recent BLAST

▼ Request a New BLAST

▼ Enter Protein Query Sequence

Enter one or more queries in the top text box or use the browse button to upload a file from your local disk. The file may contain a single sequence or a list of sequences. In both cases, the data must be in FASTA format.

Enter FASTA sequence(s)

Show an Example Sequence

Enter query sequence(s) in the text area.

Or upload your own query FASTA:

Choose File: No file chosen

Upload

The file should be a plain-text FASTA (.fasta, .fna, .fa, .fas) file. In other words, it cannot have formating as is the case with MS Word (.doc, .docs) or Rich Text Format (.rtf). It cannot be greater than

✓ Select a Dataset

- Astyanax mexicanus (Cleverfish) Genome Assembly: AstMex102
- Astyanax mexicanus (Cleverfish) cDNA (e!R67)
- Astyanax mexicanus (Surfacefish) Genome Assembly: AstMex2
- Astyanax mexicanus (Surfacefish) NCBI mRNA
- Astyanax_mexicanus (Surfacefish) cDNA (e!94)
- Danio rerio (Zebrafish) GRCz10 Genome Assembly
- Danio rerio (Zebrafish) cDNA (e!R86)
- Dugesia guanajuatensis (Mexican Planaria) Transcriptome
- Mus Musculus (Mouse) Genome Assembly GRCm38
- Mus musculus (mouse) cDNA (e!R88)
- Nematostella vectensis (Sea anemone) Genome Assembly: NV2
- Nematostella vectensis (Sea anemone) Genome Assembly: nemVec1
- Nematostella vectensis (Sea anemone) Nv3 Transcripts
- Nematostella vectensis (Sea anemone) nv2m-BETA-1 Transcripts
- Nematostella vectensis (Sea anemone) nv2m-BETA-2 Transcripts
- Nothobranchius furzeri (Killifish) Curated Gene Model Transcripts (RefSeq_r100_WEW_PPS_CKHU)
- Nothobranchius furzeri (Killifish) Gene Model Transcripts
- Nothobranchius furzeri (Killifish) Genome Assembly: Nfu_20160622
- Petromyzon marinus (Lamprey) Germline Genome Assembly gPmar100
- Petromyzon marinus (Lamprey) PMZ_v3.1 Transcripts
- Pomacea canaliculata (Apple Snail) Genome Assembly Pea_v1.1
- Pomacea canaliculata (Apple Snail) Genome Assembly RefSeq_v1
- Pomacea canaliculata (Apple Snail) MAKER Transcripts

Create

Save

BLAST REPORT

BLAST RESULTS

[Download: Alignment, Tab-Delimited, XML, GFF3](#)

Query Information: /tmp/2020Jan02_160824_query.fasta

Search Target: Petromyzon marinus (Lamprey) Germline Genome Assembly gPmar100

Submission Date: Thu, 01/02/2020 - 16:08

BLAST Command executed: blastn -max_target_seqs 500 -max_hsps 1 -evalue 0.001 -word_size 3 -gapopen -gapextend -culling_limit -matrix PAM30

Number of Results: 3

Number of Results: 3

The following table summarizes the results of your BLAST. Click on a triangle on the left to see the alignment and a visualization of the hit, and click the target name to get more information about the target hit.

| # | Query Name (Click for alignment & visualization) | Target Name | E-Value |
|-----|---|-------------|-------------|
| ▼ 1 | gi 49456395 emb CAG46518.1 HDAC1 [Homo sapiens] | scaf_00056 | 1.65021E-25 |
| ▼ 2 | gi 49456395 emb CAG46518.1 HDAC1 [Homo sapiens] | scaf_00026 | 3.15953E-8 |
| ▼ 3 | gi 49456395 emb CAC46518.1 HDAC1 [Homo sapiens] | scaf_00058 | 4.82913E-8 |

[Edit this query and re-submit](#)

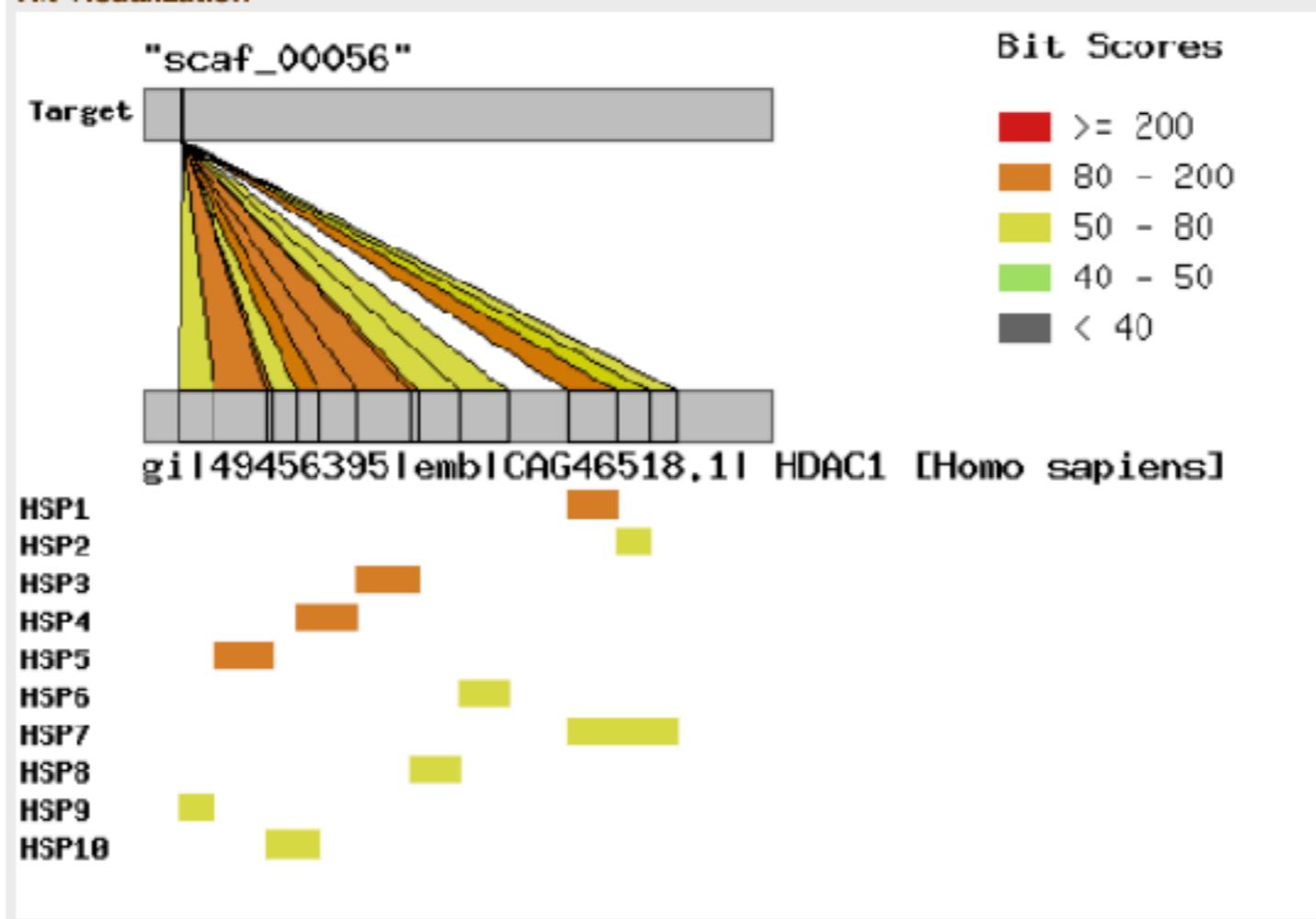
Recent Jobs

| Query Information | Search Target | Date Requested | See Results |
|---|--|-------------------------|-----------------------------|
| gi 49456395 emb CAG46518.1 HDAC1 [Homo sapiens] | Petromyzon marinus (Lamprey) Germline Genome Assembly gPmar100 | Thu, 01/02/2020 - 16:08 | See Results |

BLAST REPORT

| # | Query Name (Click for alignment & visualization) | Target Name | E-Value |
|-----|--|-------------|-------------|
| ▲ 1 | gi 49456395 emb CAG46518.1 HDAC1 [Homo sapiens] | scaf_00056 | 1.65021E-25 |

Hit Visualization



The image above shows the relationship between query and target for this particular BLAST hit.

Alignment

HSP 1

Identity= 36/38 (94.74%) , Positive= 36/38 (94.74%) Query Matches 326 to 363 Hit Matches = 375390 to 375503

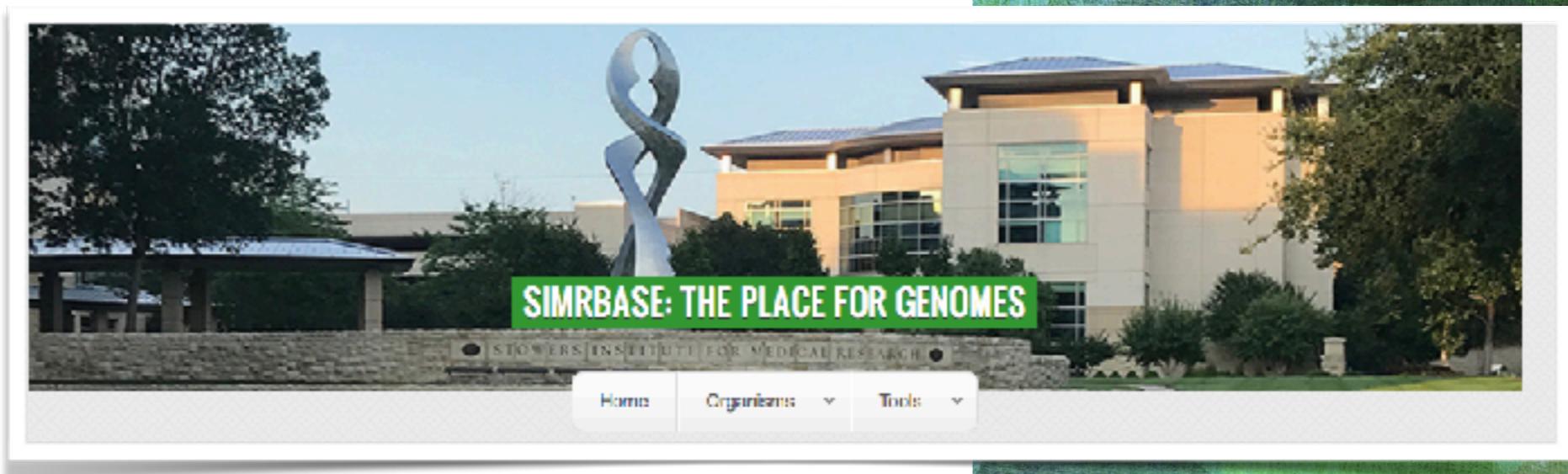
Query: 326 NELPYNDYFEYFCPDFKLHISPSNMTHQMTNEYLEKIK 363

ELPYNDYFEYFCPDFKLHISPSNMTHQMT EYLEKIK

Subject: 375390 SELPYNDYFEYFGPDFKLHISPSNMTHQNTPEYLEKIK 375449

TOUR OF SIMRBASE

.....



- Genome Browser
- Gene Searches & Gene Pages
- User BLAST Interface
- Manual Gene Curation

MANUAL GENE CURATION

The screenshot displays a manual gene curation interface for the Lamprey genome. On the left, a genomic browser shows a reference sequence (scaf_00028) from 11,067,500 to 11,112,500. It includes tracks for User-created Annotations (Drajc14-00001), Transcripts (Drajc14, GDF11), D1 (BAM), and D1 BAM. The Drajc14 transcript is shown with its coding regions and poly-A tail. On the right, an annotation table lists 143 entries for scaf_00028, with Drajc14 highlighted. A detailed view of Drajc14 shows its name, type (gene), sequence (34,104), length (577), and last update (Jan 31, 2017). The interface also includes sections for Annotations, Tracks, RefSequence, Organism, Users, Groups, and Admin.

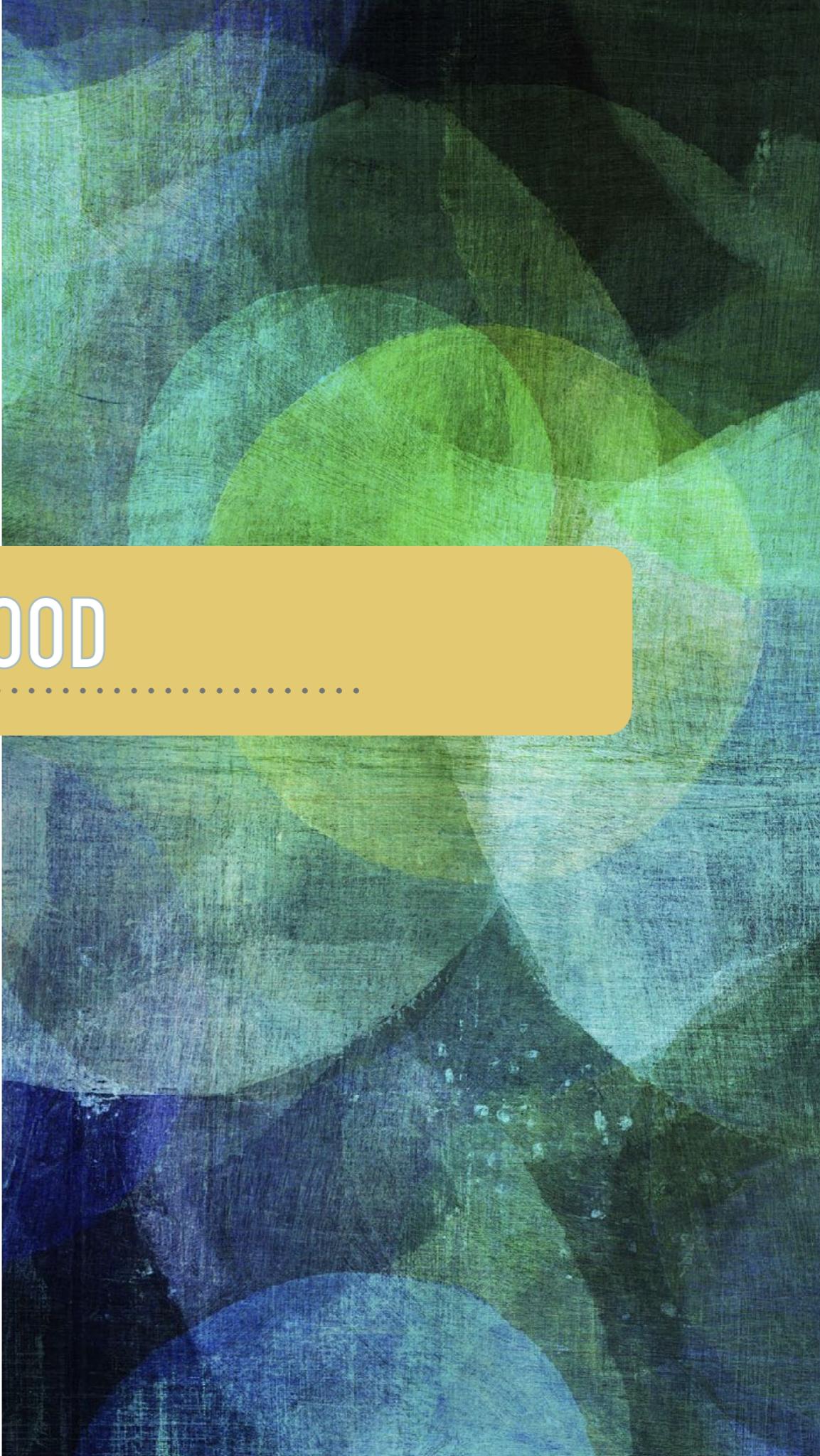
| Name | Seq | Type | Length | Updated |
|--------------------------|---------------|------|--------|--------------|
| hox-epsilon14(exon1) | scaf_00189 | gene | 577 | Jan 25, 2017 |
| hox-epsilon13 (partial?) | scaf_00006 | gene | 6,244 | Jan 26, 2017 |
| hox-zeta13 (partial) | scaf_00058 | gene | 306 | Jan 26, 2017 |
| hox-epsilon8(exon2) | scaf_00005 | gene | 309 | Jan 27, 2017 |
| HNRNPA3 | scaf_00005 | gene | 9,631 | Jan 31, 2017 |
| CHRNA1 | scaf_00004 | gene | 18,307 | Jan 31, 2017 |
| HNRNPA3b | scaf_00004 | gene | 6,567 | Jan 31, 2017 |
| Nfe2L2 | scaf_00004 | gene | 22,789 | Jan 31, 2017 |
| PRLH | scaf_00016 | gene | 4,671 | Jan 31, 2017 |
| CBX1c | scaf_00016 | gene | 3,225 | Jan 31, 2017 |
| rarga | scaf_00016 | gene | 60,841 | Jan 31, 2017 |
| ATF7 | scaf_00026 | gene | 18,689 | Jan 31, 2017 |
| Cdk5rap3 | scaf_00026 | gene | 11,629 | Jan 31, 2017 |
| AGPS | scaf_00026 | gene | 17,646 | Jan 31, 2017 |
| Drajc14 | scaf_00026 | gene | 34,104 | Jan 31, 2017 |
| | Drajc14-00001 | mRNA | 34,104 | Jan 31, 2017 |
| HOXB1 | scaf_00015 | gene | 3,063 | Jan 31, 2017 |
| HOXB13 | scaf_01238 | gene | 1,839 | Jan 31, 2017 |
| HOXC9 | scaf_01591 | gene | 3,441 | Jan 31, 2017 |
| HOXD8 | scaf_02915 | gene | 237 | Jan 31, 2017 |
| HOXB13a | scaf_05296 | gene | 2,077 | Jan 31, 2017 |

Details

| | |
|--------------|-------------------------------|
| Name | Drajc14 |
| Symbol | |
| Description | |
| Location | 11064382 - 11118788 strand(+) |
| Ref Sequence | scaf_00028 |
| Owner | hub@towers.org |

UNDER THE HOOD

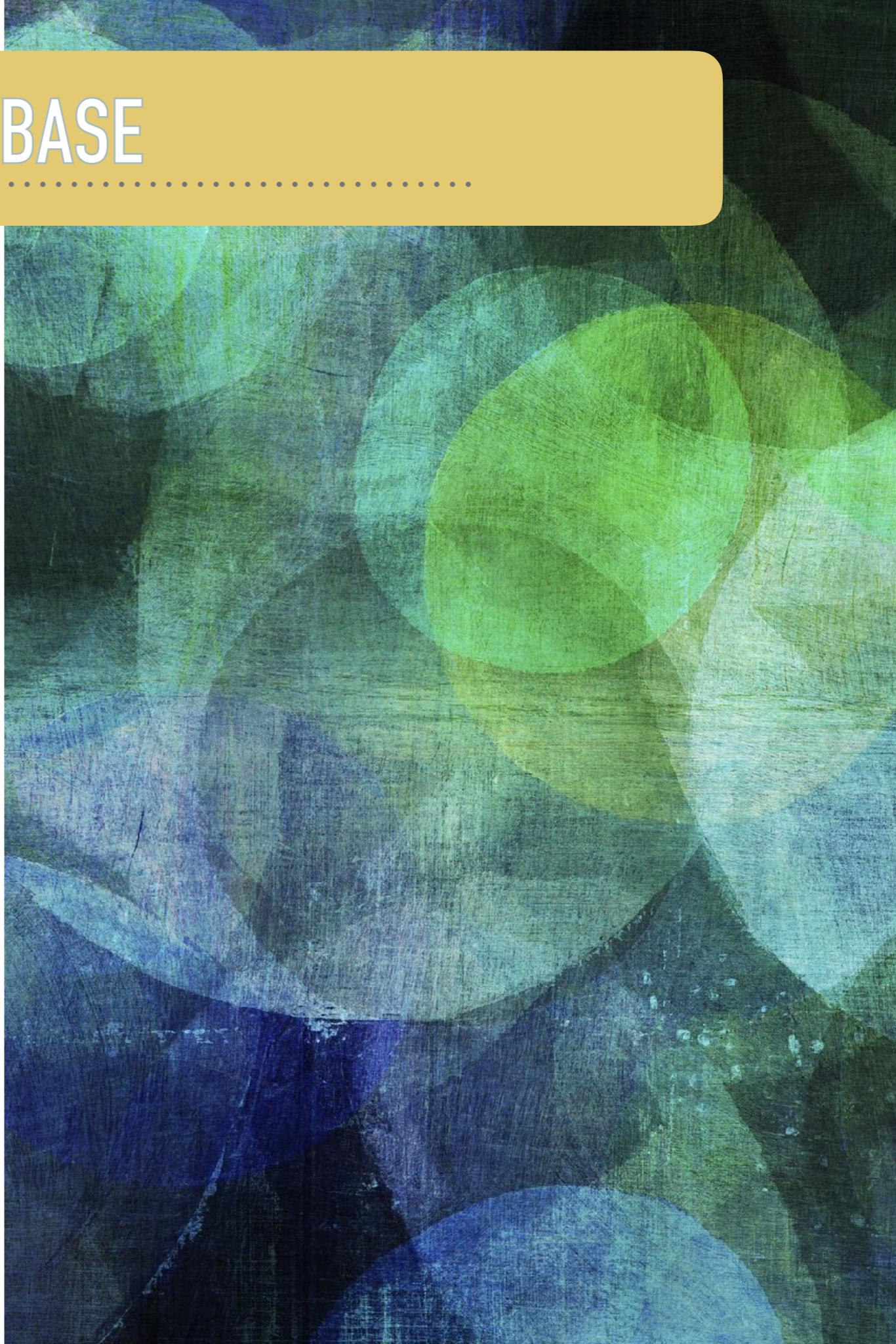
.....



SIMRBASE

.....

- Genome Annotation
- Data Storage
- Web Interface
- Sequence Searches
- Genome Visualization
- Manual Gene Curation



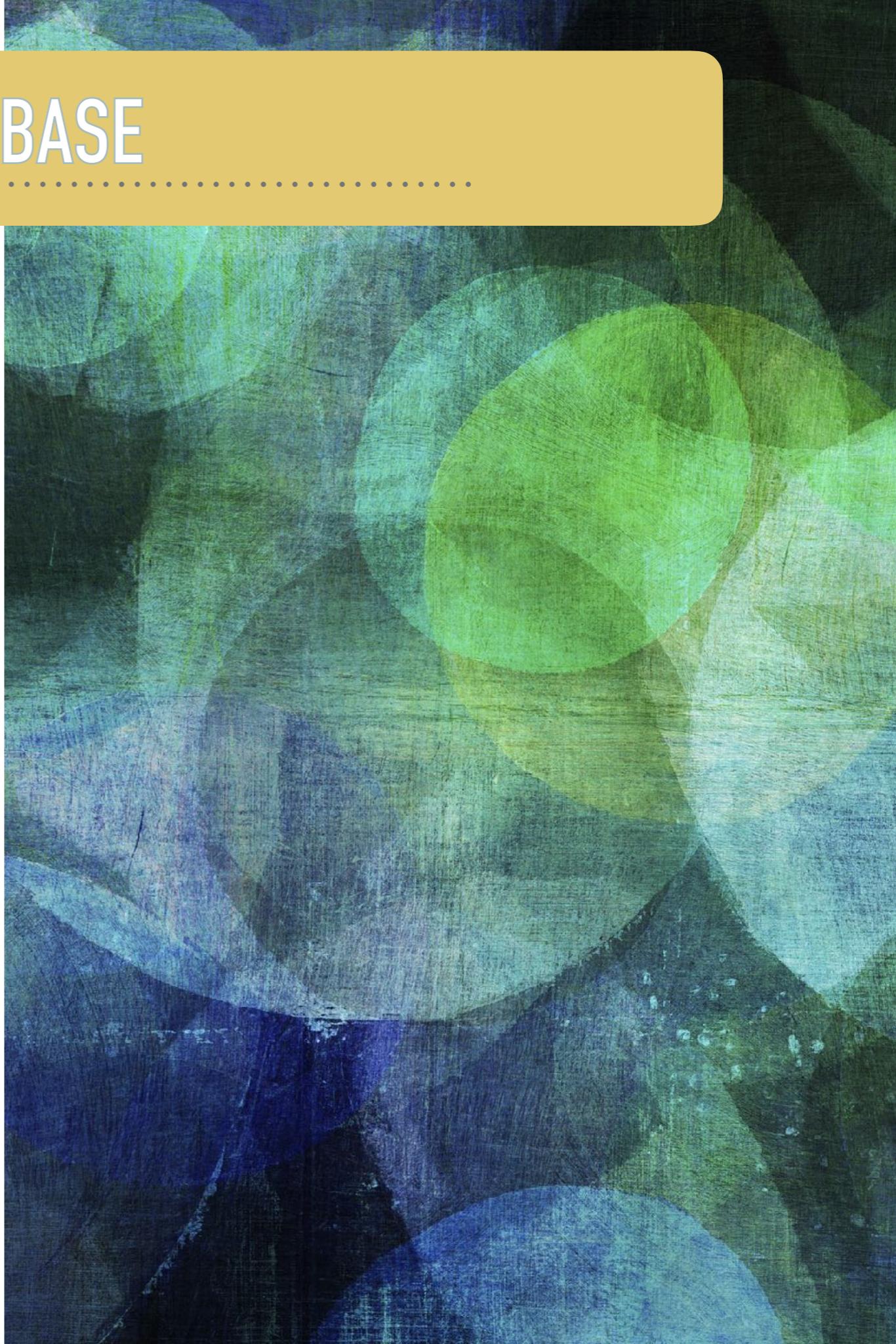


- Collection of open-source tools for managing, visualizing, storing, and disseminating genomic data.

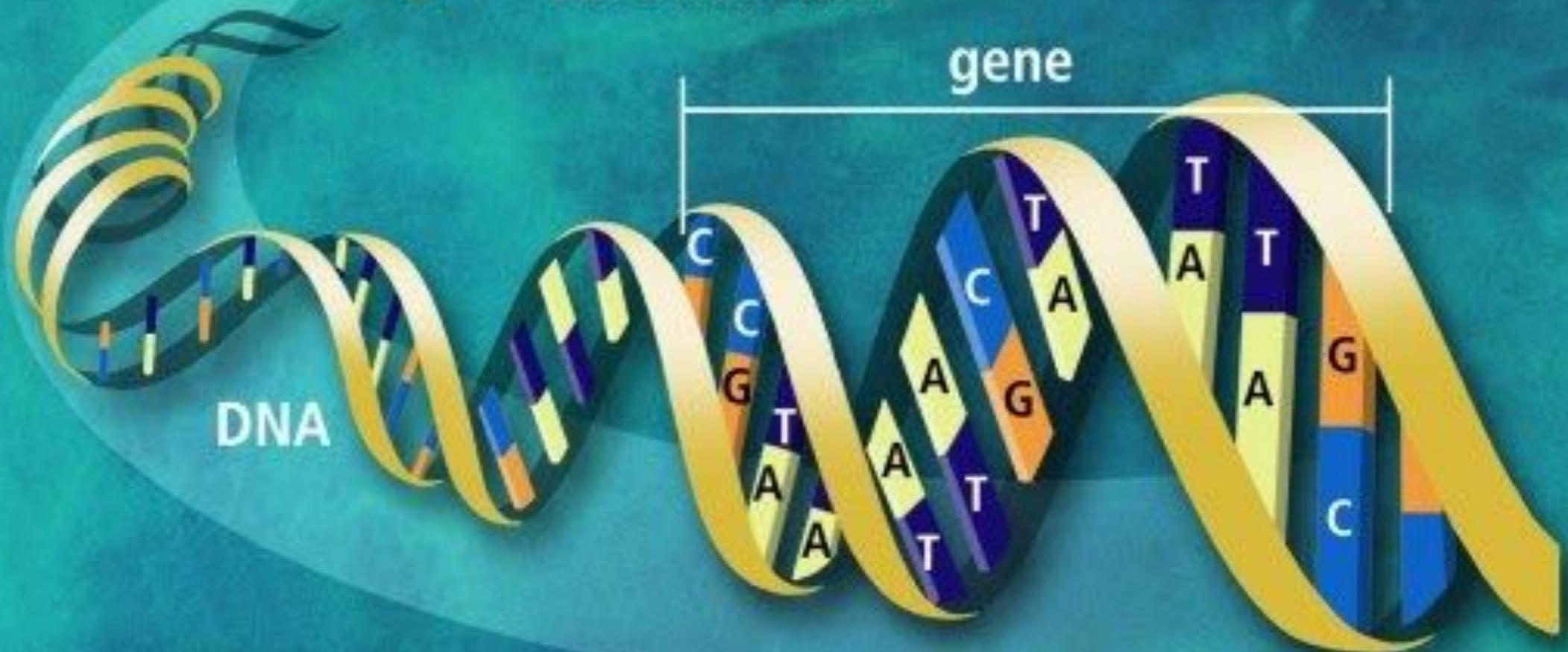
SIMRBASE

.....

- **Genome Annotation**
- **Data Storage**
- **Web Interface**
- **Sequence Searches**
- **Genome Visualization**
- **Manual Gene Curation**



GENOME ANNOTATION



- Confidently and accurately identify genes in your genome

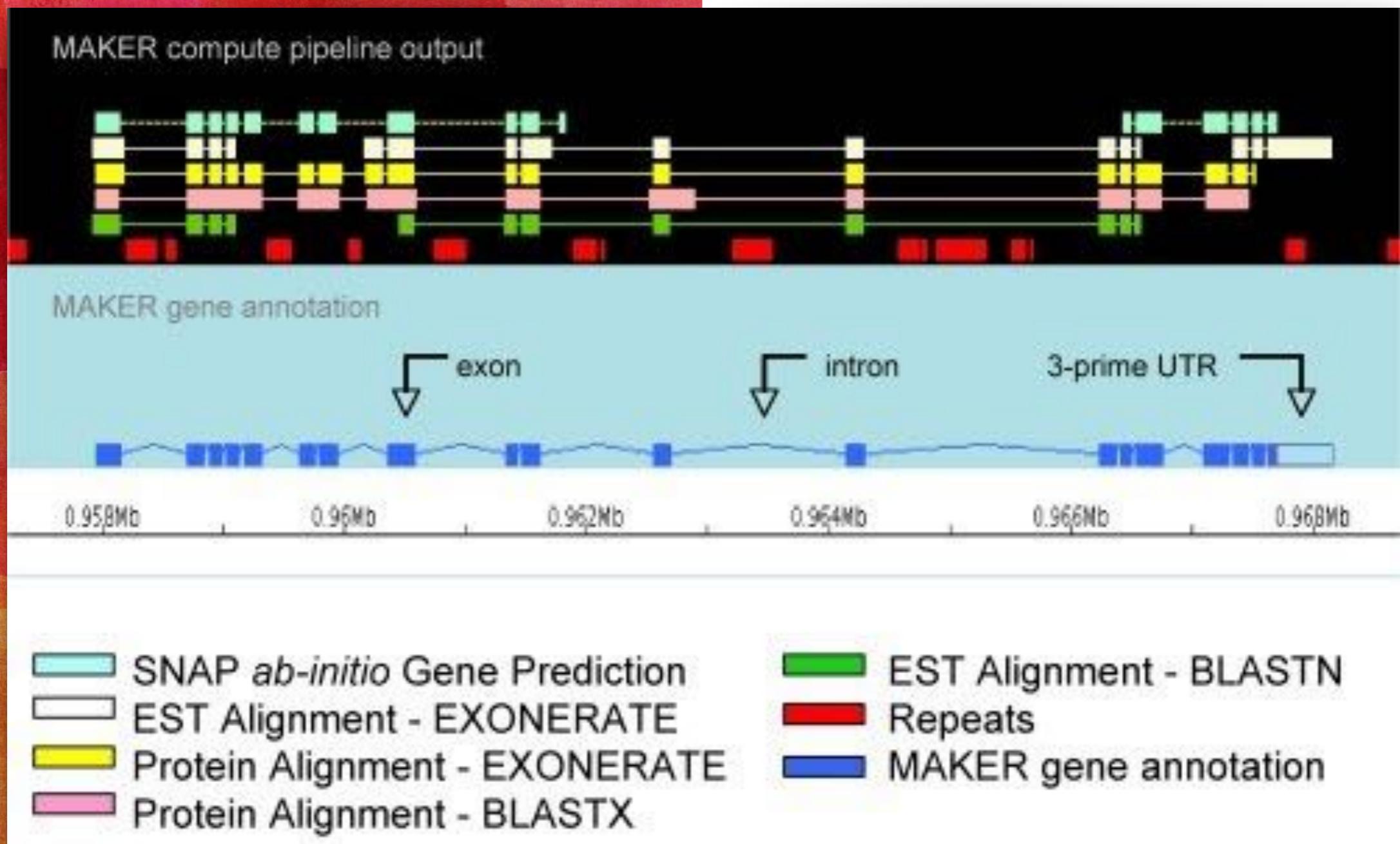


MAKER

Annotate this!

► MAKER is an easy-to-configure genome annotation pipeline.

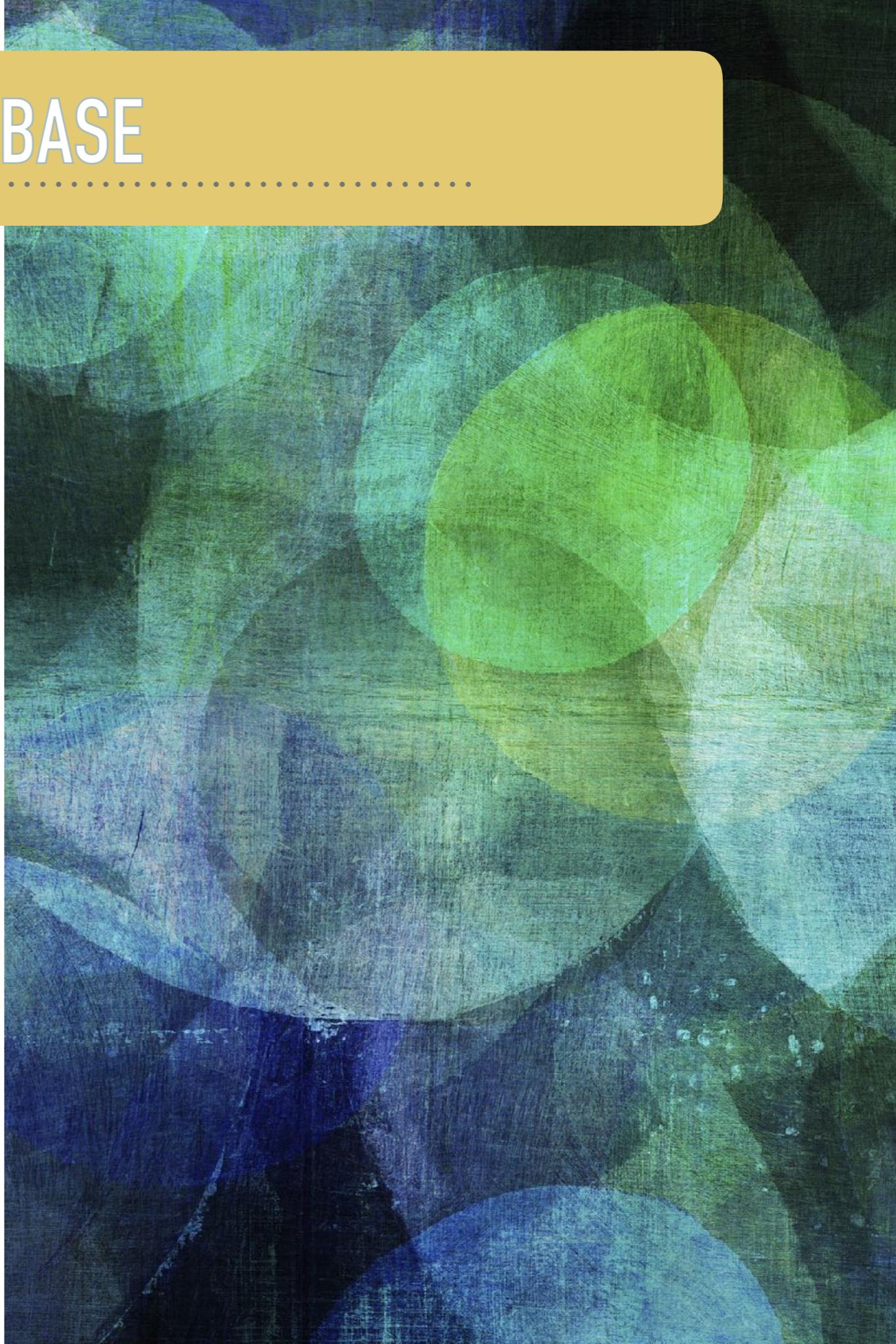
SYNTHEZIZES ALIGNED DATA INTO GENE ANNOTATIONS



SIMRBASE

.....

- Genome Annotation
- Data Storage
- Web Interface
- Sequence Searches
- Genome Visualization
- Manual Gene Curation

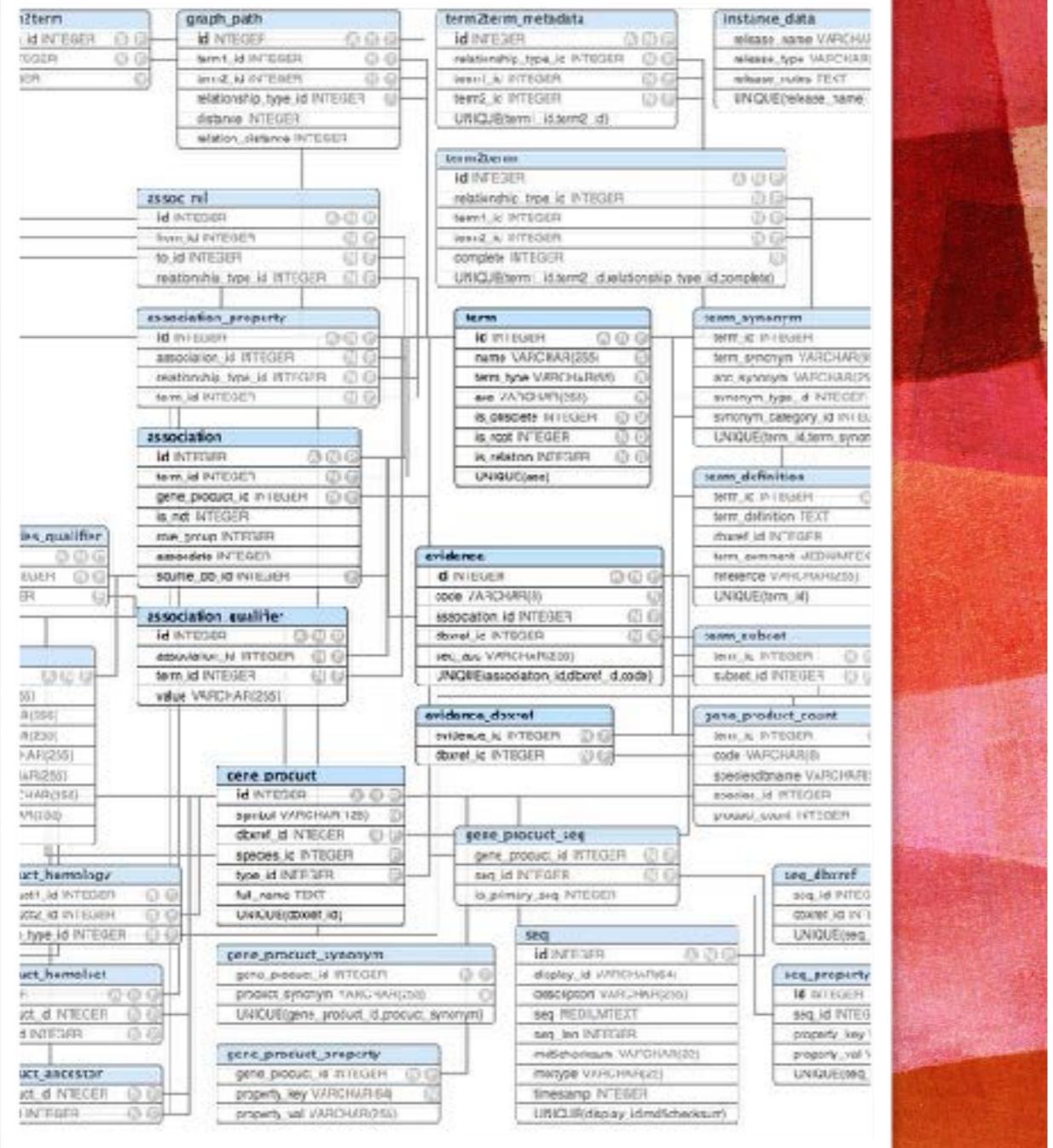


DATA STORAGE



- Organized
- Standard layout for easy
 - data input
 - data retrieval
 - data updating
 - tool integration
- Supported
 - bug fixes
 - new features
 - accommodate new data types

CHADO



- CHADO is a relational database schema that underlies many GMOD tools
 - Designed to handle complex representations of biological knowledge
 - <http://gmod.org/wiki/Chado>

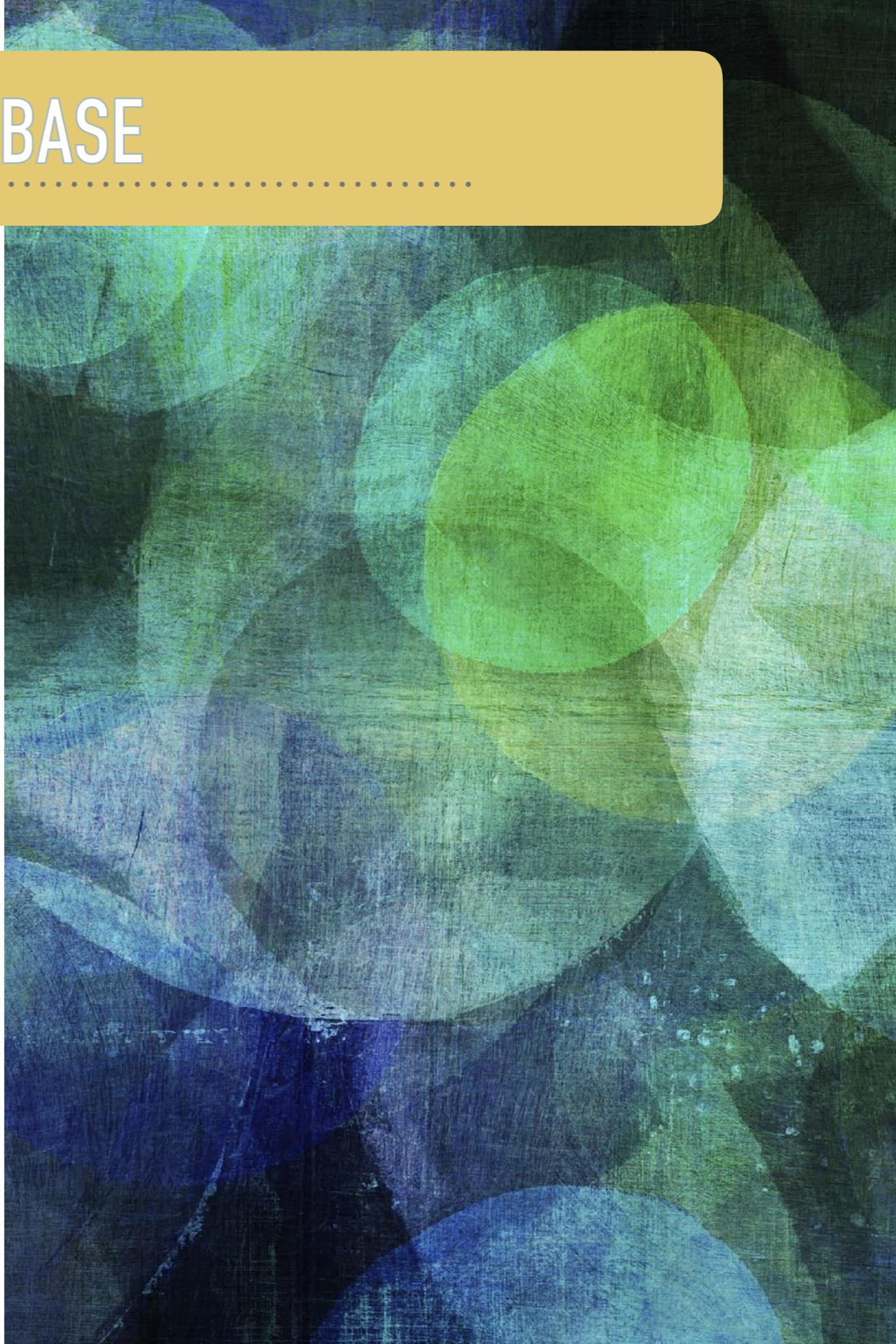
DATA FREQUENTLY USED IN BIOLOGY

- Organisms
- Genomes
- Transcripts
- Sequence alignments
- Sequence homology
- Phenotypes
- Ontology terms
- Publications
- Phylogeny
- more

SIMRBASE

.....

- Genome Annotation
- Data Storage
- Web Interface
- Sequence Searches
- Genome Visualization
- Manual Gene Curation



WEB INTERFACE

.....

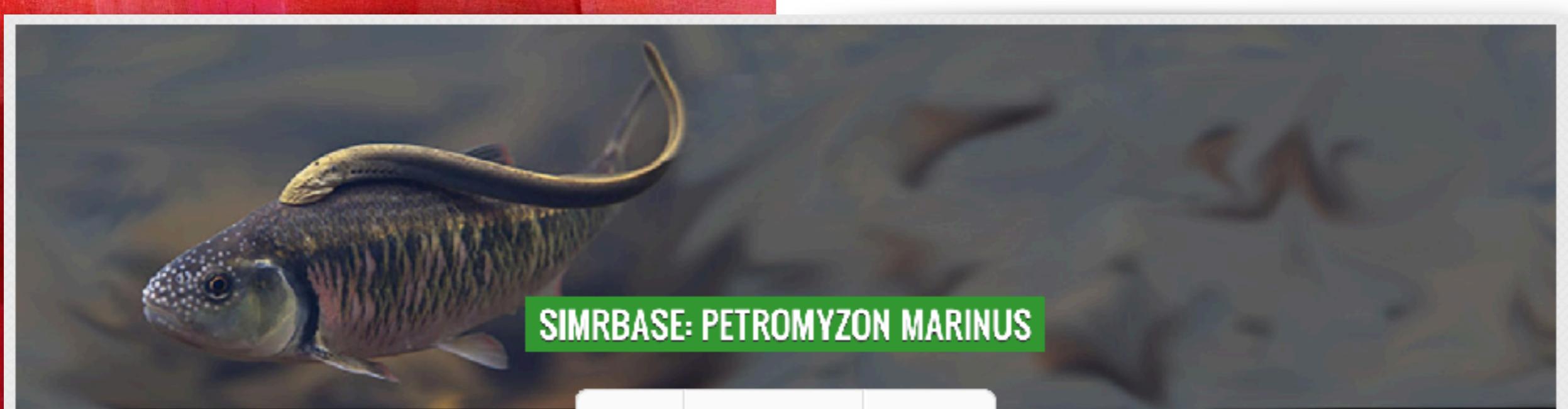


- Data should be easily accessible
- Clear interfaces
- Easy to use
- Easy to build and configure



- Tripal is a web front end for Chado database
- Drupal module
- creates gene pages and search interfaces
- <http://tripal.info/>

WEB ACCESSIBLE



SIMRBASE: PETROMYZON MARINUS

Home Organisms Tools

BLAST >

Search Genes

Retrieve Genes

Browse Genomes

WHAT'S NEW!

- LAMPREY PRECOMPUTED BLAST UPDATED
- SIMRBASE MAINTENANCE COMPLETE
- SIMRBASE MAINTENANCE 9/27/19
- GENOME BROWSER UPDATED
- NOW SEARCH BY GO IDS
- NOW SEARCH BY DOMAIN IDS
- SIMRBASE LAMPREY OPEN TO ALL
- OPEN TO COLLABORATORS
- NOW RETRIEVE FEATURES BY REFERENCE LOCATION
- NEW SITE LOOK!

SIMRBASE ORGANISMS

Petromyzon marinus (Sea Lamprey)



The sea lamprey is a member of an ancient lineage that diverged from the v million years ago (MYA). By virtue of this deep evolutionary perspective, model for understanding the evolution of several conserved and derived fe fields of biology and biomedicine. Studies have used lampreys to provide developmental pathways that define vertebrate embryogenesis, vertebrate nervous and ne structure, immunity, clotting and others. These studies reveal aspects of vertebrate biology that ha



LANGUAGES

.....

English

PORCUPINE-LIKE, GSP_005554 (mRNA) GIRARDIA S

Overview

Analyses

Gene Groups

Homology

Sequences

Overview

| | |
|-----------------|--|
| Name | porcupine-like |
| Unique Name | Gsp_005554 |
| Type | mRNA |
| Organism | <i>Girardia</i> sp. (<i>Girardia</i> sp.) |
| Sequence length | 1722 |

Spanish

PORCUPINE-LIKE, GSP_005554 (mRNA) GIRARDIA S

Resumen

Análisis

Grupos De Genes

Homología

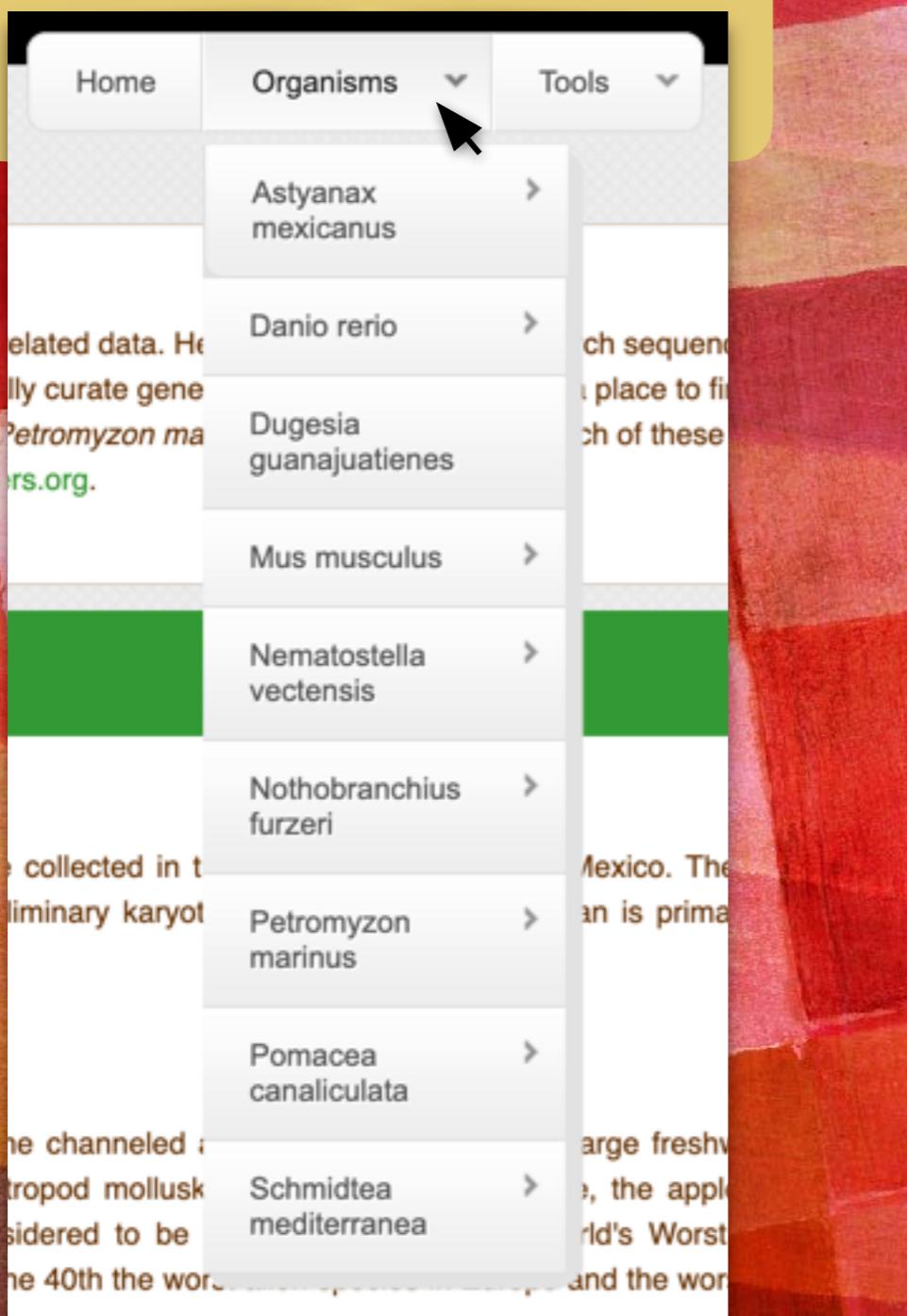
Secuencias

Resumen

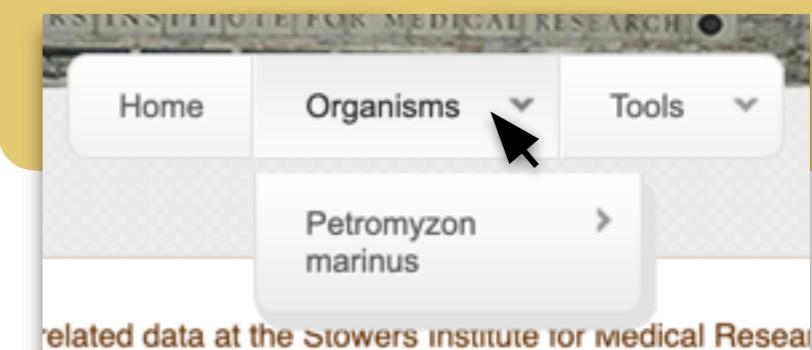
| | |
|-----------------------|--|
| Nombre | porcupine-like |
| Nombre único | Gsp_005554 |
| Tipo | mRNA |
| Organismo | <i>Girardia</i> sp. (<i>Girardia</i> sp.) |
| Secuencia de longitud | 1722 |

PERMISSIONS

Logged In



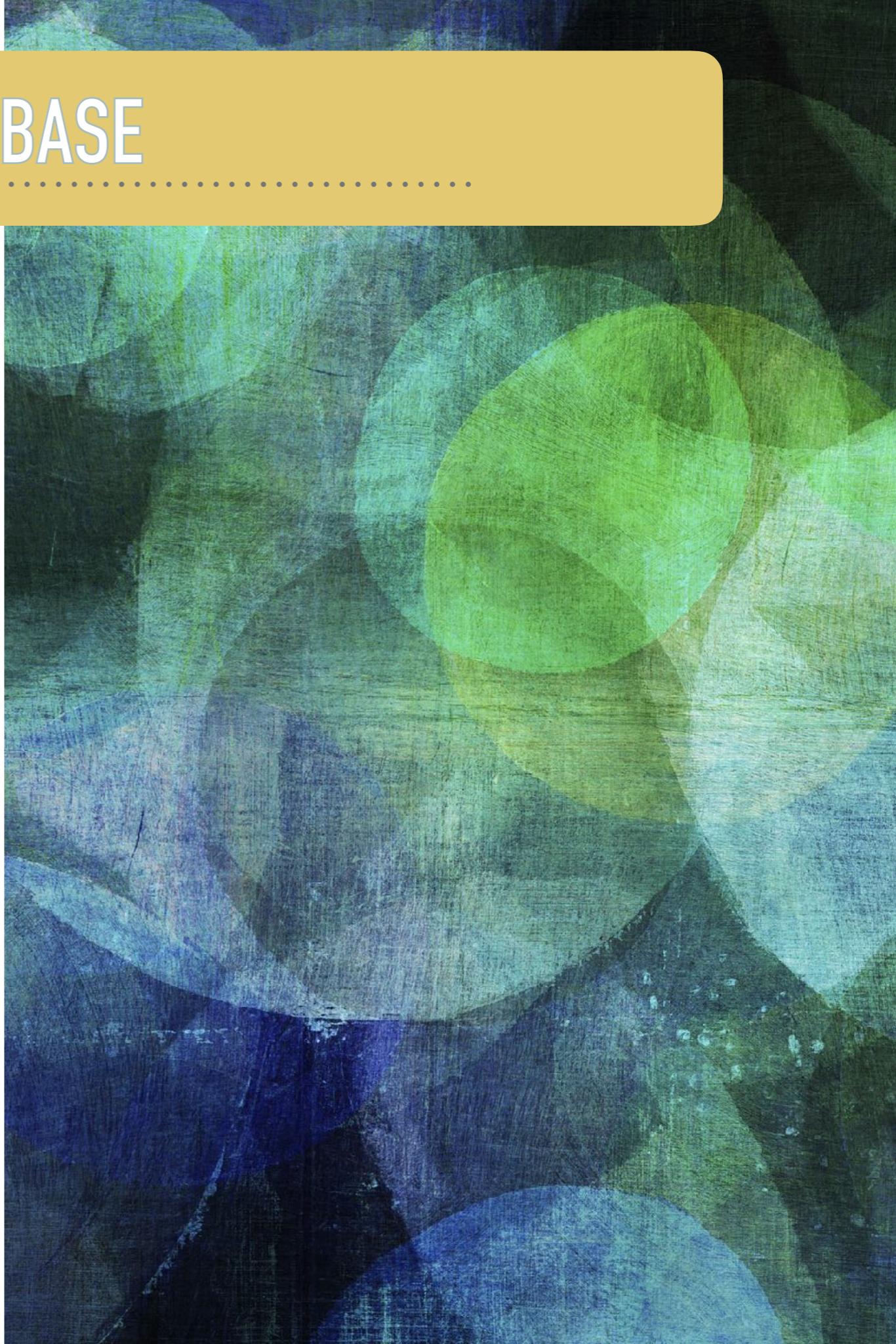
Not Logged in



SIMRBASE

.....

- Genome Annotation
- Data Storage
- Web Interface
- Sequence Searches
- Genome Visualization
- Manual Gene Curation



SEQUENCE SEARCHES

.....



- 1. Precomputed Homology**
- 2. BLAST: Sequence Similarity**

HOMOLOGY AND DOMAINS ON GENE PAGES

HDAC9, MAKER-SGR05-AUGUSTUS-GENE-189.2-MRNA-1 (mRNA) NOTHOBRANCHIUS FURZERI

Overview
Alignments
Analyses
Cross References
Homology
Properties
Relationships
Sequences

Homology

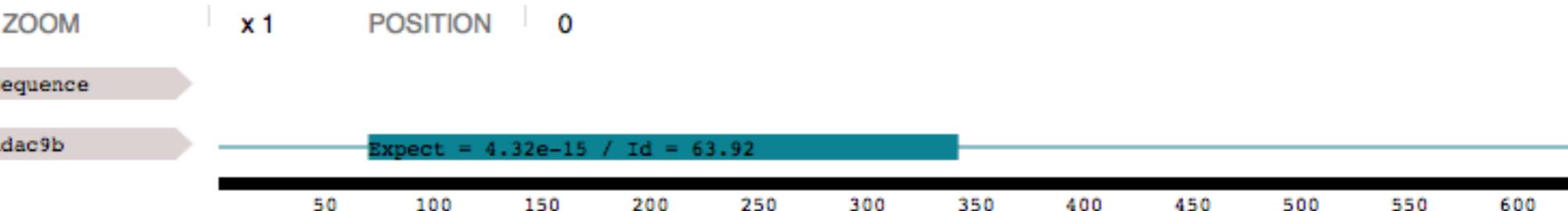
The following BLAST results are available for this feature:

- BLAST of hdac9 vs. Ensembl Medaka
- BLAST of hdac9 vs. Ensembl Zebrafish
- BLAST of hdac9 vs. UniProt
- BLAST of hdac9 vs. Ensembl Human

BLAST of hdac9 vs. Ensembl Medaka

Analysis Date: 2016-06-09 (Nothobranchius furzeri BLASTX Ensembl Medaka)

Total hits: 1



| Match Name | E-value | Identity | Description |
|------------|---------|----------|-------------|
|------------|---------|----------|-------------|

| | | | |
|--------|-----------|-------|--|
| hdac9b | 4.323e-15 | 63.92 | histone deacetylase 9b [Source:ZFIN;Acc:ZDB-GENE-0...] |
|--------|-----------|-------|--|

[more]

[back to top](#)

BLAST of hdac9 vs. Ensembl Zebrafish

Analysis Date: 2016-06-08 (Nothobranchius furzeri BLASTX Ensembl Danio)

Total hits: 2

GENE SEARCHES

GENE SEARCH

Search By: Gene ID/Name

Contains any word

Examples: MycD, SMED30002427. For multiple search terms use the "Contains any word" filter and separate with spaces or commas.

Search By: BLAST Best Hit

HDAC

Examples: HDAC

Search By: GO Term

Contains

Examples: protein binding, membrane. GO terms are displayed on the linked Gene page in the "Annotated Terms" Section.

Search By: Domains

Contains

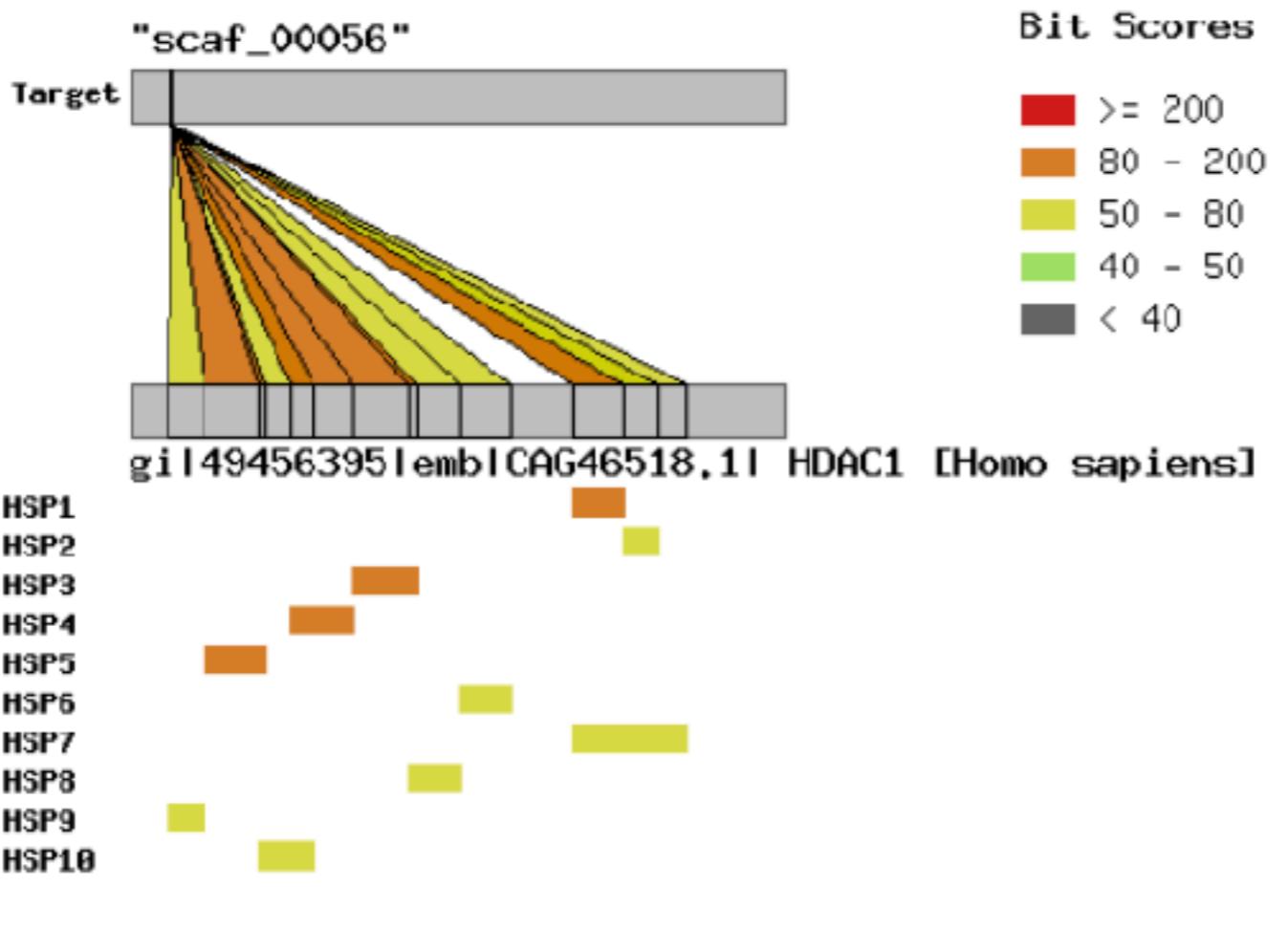
Apply

Reset

USER BLAST SEARCHES

| # | Query Name (Click for alignment & visualization) | Target Name | E-Value |
|-----|---|-------------|-------------|
| ▲ 1 | gi 49456395 emb CAG46518.1 HDAC1 [Homo sapiens] | scaf_00056 | 1.65021E-25 |

Hit Visualization



The image above shows the relationship between query and target for this particular BLAST hit.

Alignment

HSP 1

Identity= 36/38 (94.74%) , Positive= 36/38 (94.74%) Query Matches 326 to 363 Hit Matches = 375390 to 375503

Query: 326 NELPYNDYFEYFCPDFKLHISPSNMTWQNTNEYLERIK 363

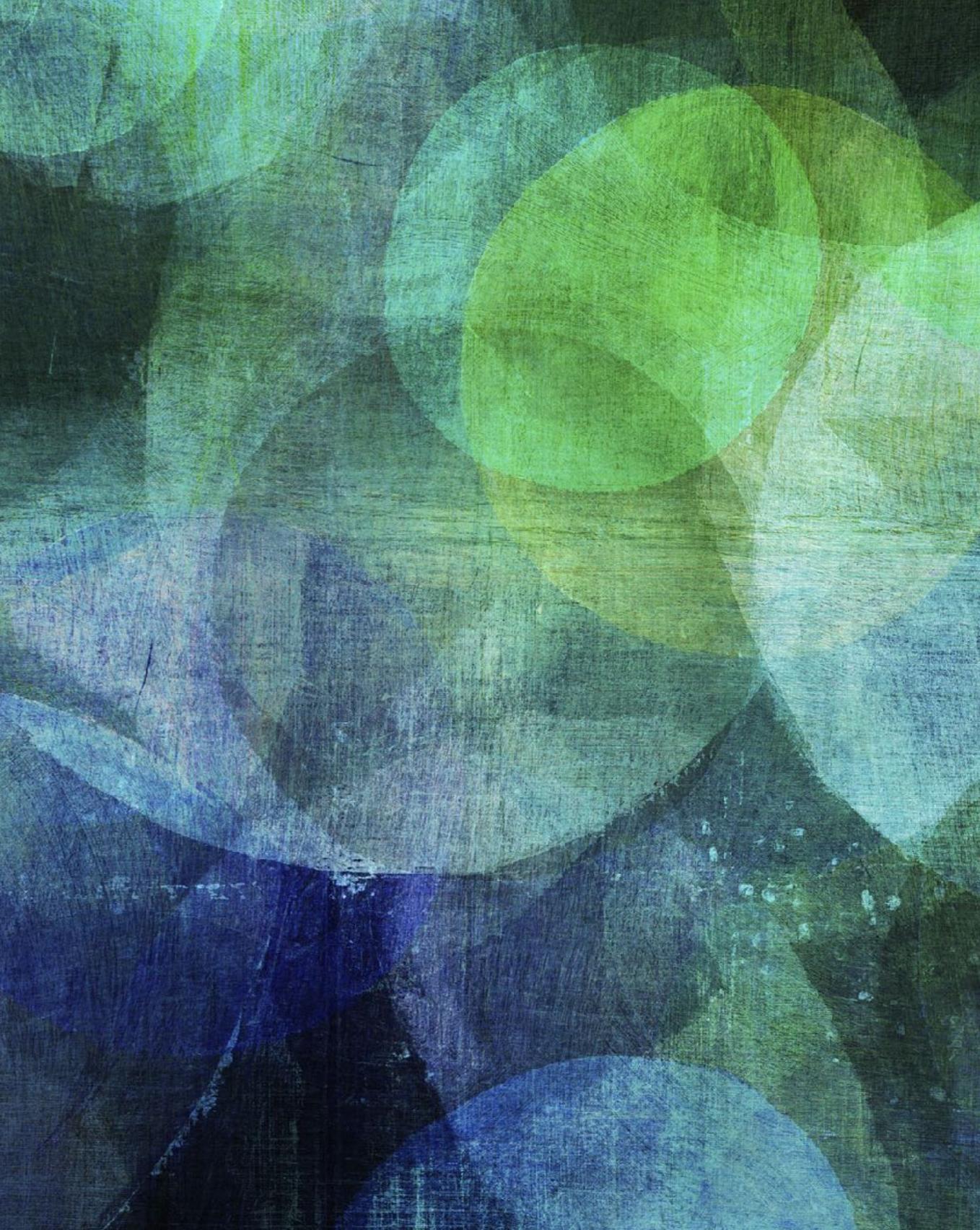
ELPYNDYFEYFCPDFKLHISPSNMTWQNT EYLERIK

Objct: 375390 SELPYNDYFEYFCPDFKLHISPSNMTWQNTPEYLERIK 375449

SIMRBASE

.....

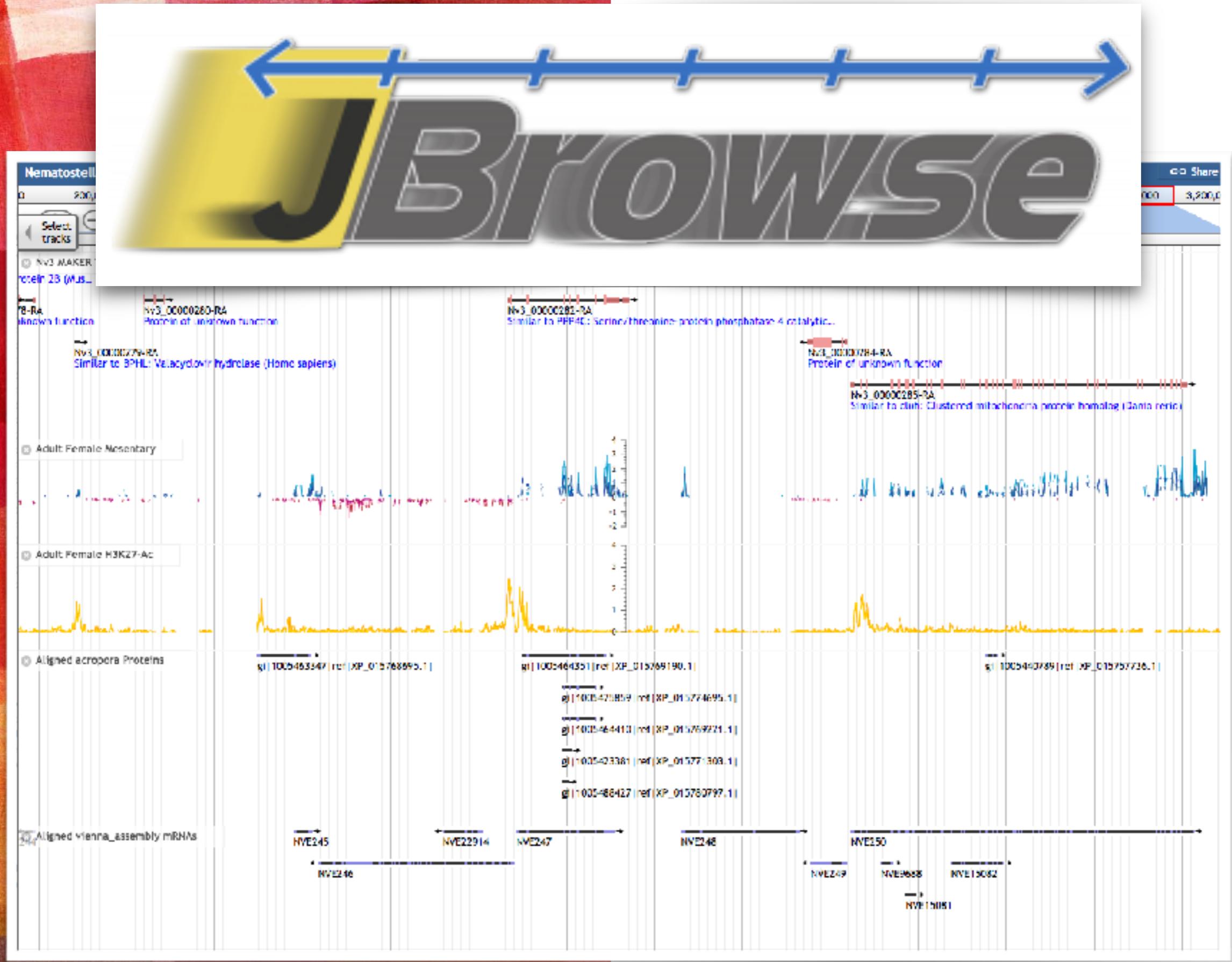
- Genome Annotation
- Data Storage
- Web Interface
- Sequence Searches
- **Genome Visualization**
- Manual Gene Curation



GENOME VISUALIZATION



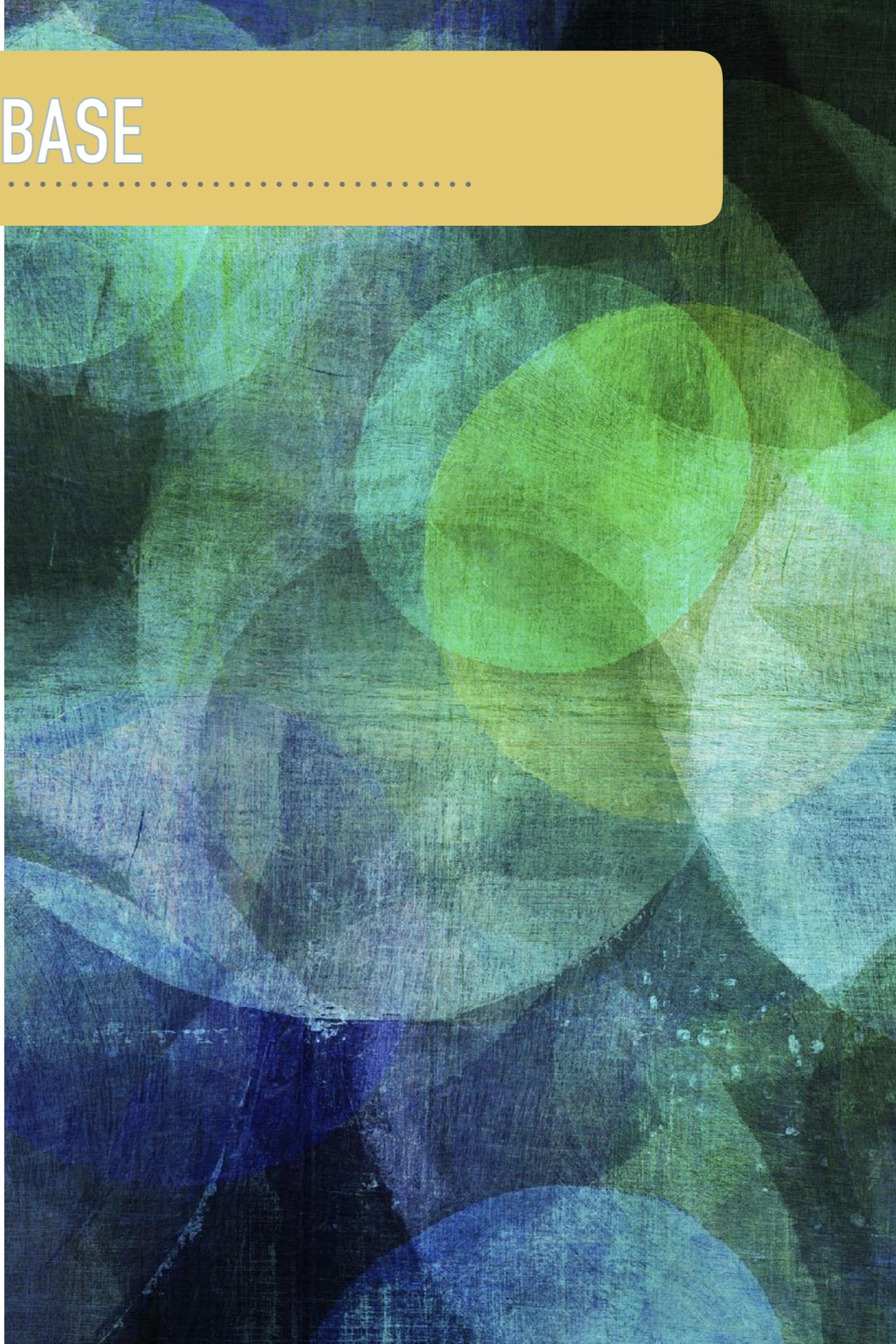
- Visualize all data that is localized to same region of the genome
- Search for features by name or genomic locations
- User upload of personal data



SIMRBASE

.....

- Genome Annotation
- Data Storage
- Web Interface
- Sequence Searches
- Genome Visualization
- Manual Gene Curation



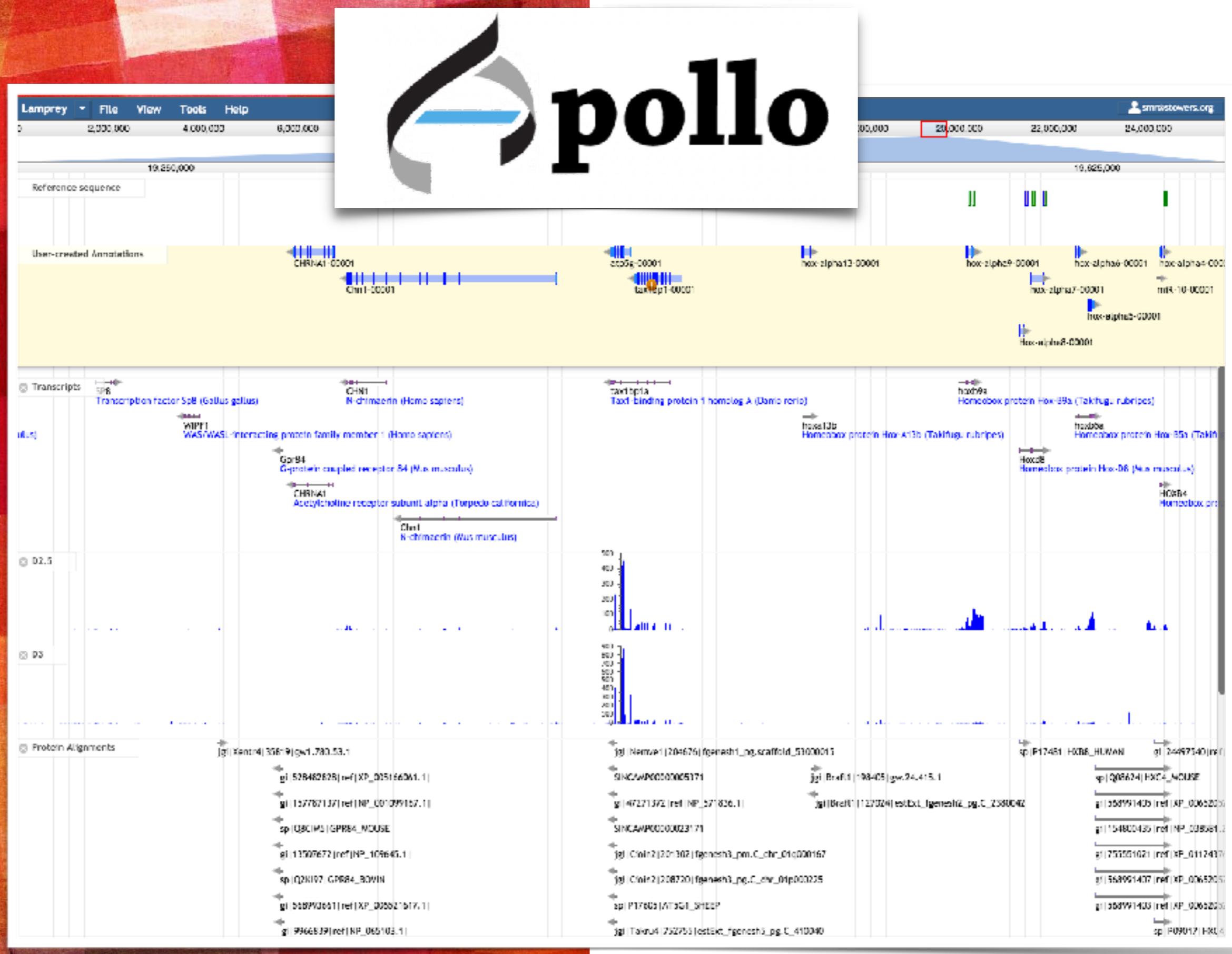
MANUAL GENE CURATION



- Gene function
- Gene name
- Split genes
- Merge genes
- Incorrect/Missing exons
- Incorrect/Missing UTR



- Web interface for manual gene annotation
- Password protected
- Integrates with the database for easy data management
- Approval required before addition to database
- <http://genomearchitect.github.io/>





SIMRBASE: PETROMYZON MARINUS

Home Organisms Tools

PETROMYZON MARINUS

Overview

BLAST

Genome Browser

Gene Search

Community Annotation

Data Analyses

Downloads

Feature Summary

Genome Consortium

Genomic Properties

Germline Genome Manuscript

Overview

| | |
|--------------|---------------------------|
| Full Name | <i>Petromyzon marinus</i> |
| Genus | <i>Petromyzon</i> |
| Species | <i>marinus</i> |
| Common Name | Sea Lamprey |
| Abbreviation | Pmar |



The sea lamprey is a member of the class Agnatha, which diverged from the vertebrates over 500 million years ago (MYA). By virtue of its unique perspective, lamprey has served as a model organism for understanding the evolution of many derived features that are relevant to medicine and biomedicine. Studies have shown that lampreys have a unique perspective on the evolution of the vertebrate genome that defines vertebrate embryology, neuroendocrine systems, genomic architecture, immunity, blood clotting and others. These studies reveal aspects of vertebrate biology that have been conserved over deep evolutionary time and reveal evolutionary modifications that gave rise to novel features that emerged within the jawed vertebrate lineage (gnathostomes). Lampreys also possess several features that are not observed in gnathostomes, which could represent either aspects of ancestral vertebrate biology that have not been conserved in the gnathostomes or features that arose since the divergence of the ancestral tissues that give rise to the



THANK YOU

.....

