Bioinformatics Tools and Resources for Identifying Gene Product Function

Emmanuel Dialynas

Bioinformatics Support Group @ IMBB
Proprietary -> Open Source

Command line  ->  Web
Web

Repositories/Databases

Data mining

Analysis

Sharing

Collaboration

Application Programming Interface

High Performance Computing
Paying tribute....
European Nucleotide Archive

The European Nucleotide Archive (ENA) provides a comprehensive record of the world’s nucleotide sequencing information, covering raw sequencing data, sequence assembly information and functional annotation. More about ENA

Access to ENA data is provided through the browser, through search tools, large scale file download and through the API.

Text Search

sodium channel
Examples: BN000065, histone

Search
Advanced search

Sequence Search

Enter or paste a nucleotide sequence or accession number

Search
Advanced search

Popular
- Submit and update
- Sequence submissions
- Genome assembly submissions
- Submitting environmental sequences
- Citing ENA data
- Rest URLs for data retrieval
- Rest URLs to search ENA

Latest ENA news

02 Nov 2015: Change to Globus endpoint for public ENA data
The Globus endpoint for public ENA data is changing from ebi#ena to ebi#public ('ena' subfolder).

23 Sep 2015: ENA Release 125
Release 125 of ENA’s assembled/annotated sequences now available
### Search results for **sodium channel**

#### Coding (Release) (6,349 results found)

<table>
<thead>
<tr>
<th>Download</th>
<th>1</th>
<th>6349</th>
<th>of 6,349 results in TEXT XML FASTA</th>
</tr>
</thead>
</table>

#### Showing results 11 - 20 of 6,349 results

<table>
<thead>
<tr>
<th>Accession</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAA16202</td>
<td>Doryteuthis opalescens sodium channel</td>
</tr>
<tr>
<td>AAA67105</td>
<td>Mus musculus (house mouse) sodium channel 21</td>
</tr>
<tr>
<td>AAA67106</td>
<td>Mus musculus (house mouse) sodium channel 25</td>
</tr>
<tr>
<td>AAP19643</td>
<td>Mus sp. partial sodium channel</td>
</tr>
<tr>
<td>AD028451</td>
<td>Ictalurus punctatus (channel catfish) sodium channel modifier 1</td>
</tr>
<tr>
<td>AGO33659</td>
<td>Culex pipiens pallens sodium channel</td>
</tr>
<tr>
<td>BAA78033</td>
<td>Homo sapiens (human) sodium channel</td>
</tr>
<tr>
<td>BAM84088</td>
<td>Thrips palmi partial sodium channel</td>
</tr>
<tr>
<td>BAO52751</td>
<td>Thrips palmi sodium channel</td>
</tr>
<tr>
<td>BA052752</td>
<td>Thrips palmi sodium channel</td>
</tr>
</tbody>
</table>
Culex pipiens pallens sodium channel

Organism: Culex pipiens pallens
Molecule type: transcribed RNA
Topology: linear
Data class: STD
Taxonomic Division: INV

Sequence length: 6,132
Sequence Version: 1

Lineage: Eukaryota, Metazoa, Ecdysozoa, Arthropoda, Hexapoda, Insecta, Pterygota, Neoptera, Endopterygota, Diptera, Nematocera, Culicoides, Culicidae, Culicinae, Culicini, Culex

Navigation | Overview | Source Feature(s) | Sequence | Publications | Submission Details
---|---|---|---|---|---
Overview
Features
Source: Culex pipiens pallens
precursor_RNA
CDS

Showing first 1 - 1000 of 6132

>ENA|AG033659|AG033659.1 Culex pipiens pallens sodium channel : Location:1..1000
ATGACGCAAAGACCTGCATCTGATCTGAGAAAGACGATGTGTTTGGTCGGCTCTTGCA
CGTAATCTATCTTGATGAAAGACGATGAAACATGAAACAGCAAAACAGCCGA
TTGGAAGAAGACGGCAAGCCAGGAGGATGAGTACGAGTAGAGACGAGGAGGATGAGT
CCCCACGGCACTCACCTCGACGACGAGGAGGATGCGACGTCCGTGCAATTGCAAGGACG
TTCCCTCCGAAATTGGCCTCACGCCCTCGAGAGATTAGCAGTCTTTTACTCCAAACATT
AAGAGATCTGATGTTAGTAAGGGAAGGATATTTTTTTTTCGCTTCGCCACCAATGCA
Want more?

Genome Browser

Gene Expression

Variance

Cross Species Comparison
Even more?
Gene: para AGAP004707

Description: voltage-gated sodium channel [Source: VB Community Annotation, Acc: AGAP004707]
Synonyms: VSC, kdr
Location: Chromosome 2L: 2356.158-2431.617 forward strand
About this gene: This gene has 3 transcripts (splice variants), 4 orthologues, and 9 paralogues

Genomic alignments

Alignment: Anopheles coluzzii - lastz

Download alignment

View an image of this alignment

A total of 7 alignment blocks have been found. Please select an alignment to view by selecting a Block from the Alignment column.
### GO: cellular component

The following terms describe the **cellular_component** of this entry in Gene Ontology:

<table>
<thead>
<tr>
<th>Chart</th>
<th>Accession</th>
<th>Term</th>
<th>Evidence</th>
<th>Annotation Source</th>
<th>GOSlim Accessions</th>
<th>GOSlim Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GO:0001518</td>
<td>voltage-gated sodium channel complex</td>
<td>IEA</td>
<td>UniProtKB/TrEMBL:Q7PJH0_ANOGA</td>
<td>GO:0005575, GO:0005623, GO:0043234, GO:0005886</td>
<td>cellular_component, cell, protein complex, plasma membrane</td>
</tr>
</tbody>
</table>

### GO: molecular function

The following terms describe the **molecular_function** of this entry in Gene Ontology:

<table>
<thead>
<tr>
<th>Chart</th>
<th>Accession</th>
<th>Term</th>
<th>Evidence</th>
<th>Annotation Source</th>
<th>GOSlim Accessions</th>
<th>GOSlim Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GO:0005248</td>
<td>voltage-gated sodium channel activity</td>
<td>IEA</td>
<td>UniProtKB/TrEMBL:Q7PJH0_ANOGA</td>
<td>GO:0003674, GO:0022857</td>
<td>molecular_function, transmembrane, transporter activity</td>
</tr>
<tr>
<td></td>
<td>GO:0005509</td>
<td>calcium ion binding</td>
<td>IEA</td>
<td>UniProtKB/TrEMBL:Q7PJH0_ANOGA</td>
<td>GO:0003674, GO:0043167</td>
<td>molecular_function, ion binding</td>
</tr>
</tbody>
</table>
The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.
Ensembl peptide - Interpro

Protein family membership
None predicted.

Domains and repeats
- Voltage-dependent channel, four helix bundle domain
  - IPR027359
  - G005A:1.10.12...
- Ion transport domain
  - IPR003582
  - IP00020 (ion_trans)
- Voltage-gated Na+ ion channel, cytoplasmic domain
  - IPR025483
  - IP11933 (ion_trans..)

Detailed signature matches
- no IPR

Integrating signatures
- CYTOPLECTIC_D...
- Cell
- NON_CYTOSKELETAL...

GO term prediction

Biological Process
- GO:0005961: Ion transport
- GO:0005959: membrane transport

Molecular Function
- GO:0098516: ion channel activity

Cellular Component
- GO:0016020: membrane
<table>
<thead>
<tr>
<th>Pfam</th>
<th>465</th>
<th>694</th>
<th>Domain of unknown function DUF3451</th>
<th>PF11933</th>
<th>IPR024383 [Display all genes with this domain]</th>
</tr>
</thead>
<tbody>
<tr>
<td>PROSITE profiles</td>
<td>1834</td>
<td>1869</td>
<td>EF-hand domain</td>
<td>PS50222</td>
<td>IPR002043 [Display all genes with this domain]</td>
</tr>
<tr>
<td>Pfam</td>
<td>164</td>
<td>409</td>
<td>Ion transport domain</td>
<td>PF00520</td>
<td>IPR005821 [Display all genes with this domain]</td>
</tr>
<tr>
<td>Pfam</td>
<td>620</td>
<td>1004</td>
<td>Ion transport domain</td>
<td>PF00520</td>
<td>IPR005821 [Display all genes with this domain]</td>
</tr>
<tr>
<td>Pfam</td>
<td>1201</td>
<td>1516</td>
<td>Ion transport domain</td>
<td>PF00520</td>
<td>IPR005821 [Display all genes with this domain]</td>
</tr>
<tr>
<td>Pfam</td>
<td>1610</td>
<td>1516</td>
<td>Ion transport domain</td>
<td>PF00520</td>
<td>IPR005821 [Display all genes with this domain]</td>
</tr>
<tr>
<td>Pfam</td>
<td>1020</td>
<td>1265</td>
<td>Sodium ion transport-associated</td>
<td>PF06512</td>
<td>IPR010528 [Display all genes with this domain]</td>
</tr>
<tr>
<td>Prints</td>
<td>1799</td>
<td>1516</td>
<td>Voltage gated sodium channel, alpha subunit</td>
<td>PR00170</td>
<td>IPR001696 [Display all genes with this domain]</td>
</tr>
<tr>
<td>Prints</td>
<td>1634</td>
<td>1647</td>
<td>Voltage gated sodium channel, alpha subunit</td>
<td>PR00170</td>
<td>IPR001696 [Display all genes with this domain]</td>
</tr>
<tr>
<td>PANTHER</td>
<td>38</td>
<td>450</td>
<td>Voltage gated sodium channel, alpha-4 subunit</td>
<td>PTHR10037:SF193</td>
<td>IPR028825 [Display all genes with this domain]</td>
</tr>
<tr>
<td>PANTHER</td>
<td>477</td>
<td>491</td>
<td>Voltage gated sodium channel, alpha-4 subunit</td>
<td>PTHR10037:SF193</td>
<td>IPR028825 [Display all genes with this domain]</td>
</tr>
</tbody>
</table>
PredictProtein Open

http://ppopen.informatik.tu-muenchen.de/visual_results?req_id=$1$04oxht0J$DgyLijzkqNhOOVVtxuHtR0#
## PredictProtein Open

### Molecular Function Ontology

<table>
<thead>
<tr>
<th>#</th>
<th>GO ID</th>
<th>GO Term</th>
<th>Reliability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GO:0022838</td>
<td>substrate-specific channel activity</td>
<td>54</td>
</tr>
<tr>
<td>2</td>
<td>GO:0022857</td>
<td>transmembrane transporter activity</td>
<td>54</td>
</tr>
<tr>
<td>3</td>
<td>GO:0022803</td>
<td>passive transmembrane transporter activity</td>
<td>54</td>
</tr>
<tr>
<td>4</td>
<td>GO:0005215</td>
<td>transporter activity</td>
<td>54</td>
</tr>
<tr>
<td>5</td>
<td>GO:0005216</td>
<td>ion channel activity</td>
<td>54</td>
</tr>
<tr>
<td>6</td>
<td>GO:0022861</td>
<td>substrate-specific transmembrane transporter activity</td>
<td>54</td>
</tr>
<tr>
<td>7</td>
<td>GO:0022862</td>
<td>substrate-specific transporter activity</td>
<td>54</td>
</tr>
<tr>
<td>8</td>
<td>GO:0015073</td>
<td>ion transmembrane transporter activity</td>
<td>54</td>
</tr>
<tr>
<td>9</td>
<td>GO:0015267</td>
<td>channel activity</td>
<td>54</td>
</tr>
<tr>
<td>10</td>
<td>GO:0022932</td>
<td>voltage-gated channel activity</td>
<td>51</td>
</tr>
</tbody>
</table>

### Biological Process Ontology

<table>
<thead>
<tr>
<th>#</th>
<th>GO ID</th>
<th>GO Term</th>
<th>Reliability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GO:0006811</td>
<td>ion transport</td>
<td>57</td>
</tr>
<tr>
<td>2</td>
<td>GO:0044699</td>
<td>single-organism process</td>
<td>57</td>
</tr>
<tr>
<td>3</td>
<td>GO:0044765</td>
<td>single-organism transport</td>
<td>57</td>
</tr>
<tr>
<td>4</td>
<td>GO:0006810</td>
<td>transport</td>
<td>46</td>
</tr>
<tr>
<td>5</td>
<td>GO:0001234</td>
<td>establishment of localization</td>
<td>40</td>
</tr>
<tr>
<td>6</td>
<td>GO:0001179</td>
<td>localization</td>
<td>46</td>
</tr>
<tr>
<td>7</td>
<td>GO:0000001</td>
<td>metal ion transport</td>
<td>42</td>
</tr>
<tr>
<td>8</td>
<td>GO:0006812</td>
<td>cation transport</td>
<td>42</td>
</tr>
<tr>
<td>9</td>
<td>GO:0011572</td>
<td>monovalent inorganic cation transport</td>
<td>42</td>
</tr>
<tr>
<td>10</td>
<td>GO:0006814</td>
<td>sodium ion transport</td>
<td>28</td>
</tr>
</tbody>
</table>

### Cellular Component Ontology

<table>
<thead>
<tr>
<th>#</th>
<th>GO ID</th>
<th>GO Term</th>
<th>Reliability (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>GO:0016006</td>
<td>membrane</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>GO:0005886</td>
<td>plasma membrane</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>GO:0044464</td>
<td>cell part</td>
<td>35</td>
</tr>
<tr>
<td>4</td>
<td>GO:000623</td>
<td>cell</td>
<td>35</td>
</tr>
<tr>
<td>5</td>
<td>GO:0071944</td>
<td>cell periphery</td>
<td>35</td>
</tr>
<tr>
<td>6</td>
<td>GO:0034706</td>
<td>sodium channel complex</td>
<td>33</td>
</tr>
<tr>
<td>7</td>
<td>GO:0034702</td>
<td>ion channel complex</td>
<td>33</td>
</tr>
<tr>
<td>8</td>
<td>GO:0034703</td>
<td>cation channel complex</td>
<td>33</td>
</tr>
<tr>
<td>9</td>
<td>GO:1902490</td>
<td>transmembrane transporter complex</td>
<td>33</td>
</tr>
<tr>
<td>10</td>
<td>GO:0044425</td>
<td>membrane part</td>
<td>33</td>
</tr>
</tbody>
</table>
PredictProtein Open
Gene Ontology and KEGG annotation

Welcome to DAVID 6.7

2003 - 2015

The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.7 is an update to the sixth version of our original web-accessible programs. DAVID now provides a comprehensive set of functional annotation tools for investigators to understand biological meaning behind large list of genes. For any given gene list, DAVID tools are able to:

- Identify enriched biological themes, particularly GO terms
- Discover enriched functional-related gene groups
- Cluster redundant annotation terms
- Visualize genes on BioCarta & KEGG pathway maps
- Display related many-genes-to-many-terms on 2-D view
- Search for other functionally related genes not in the list
- List interacting proteins
- Explore gene names in batch
- Link gene-disease associations
- Highlight protein functional domains and motifs
- Redirect to related literatures
- Convert gene identifiers from one type to another.
- And more

What's important in DAVID?

- Current (v 6.7) release note
- New requirement to cite DAVID
- Lists of Affy Exon and Gene arrays supported
- Novel Classification Algorithms
- Pre-built Affymetrix and Illumina backgrounds
- User-customized gene background
- Enhanced calculating speed

Statistics of DAVID

DAVID Bioinformatic Resources Citations

KEGG: Kyoto Encyclopedia of Genes and Genomes

KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism and the ecosystem, from molecular-level information, especially large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies. (See Release notes for new and updated features.)

New article - KEGG as a reference resource for gene and protein annotation

Main entry point to the KEGG web service

KEGG PATHWAY
KEGG BRITEx
KEGG MODULE
KEGG ORTHOLOGY
KEGG GENOME
KEGG GENES
KEGG COMPOUND
KEGG REACTION
KEGG DISEASE
KEGG DRUG
KEGG MEDICUS
KEGG Organism-specific entry points

Analysis tools

KEGG Mapper
KEGG Atlas
BlatKOALA
GhostKOALA
BLAST/FASTA
SIMCOMP

Screen Shot 1
Screen Shot 2
Screen Shot 3
Galaxy Server @ IMBB

NGS experiment analysis

Flexible workflow designing

Genome annotation

RNA structure prediction

Very active community

70+ public servers available
Chipster Analysis Platform @ IMBB

Next Generation Sequencing

- Genome browser
- Alignment
- Variants
- RNA-seq: Assembly & differential expression
- ChIP-seq: Peak detection motif scanning

Microarrays & Proteomics

- Normalization
- Filtering
- Plots (Venn, scattered, idiograms etc.)
- Clustering
- Pathway analysis (GO, KEGG, PFAM)
Chipster Analysis Platform
Thank you :-)  

bioinfo@imbb.forth.gr  

ENA - http://www.ebi.ac.uk/ena  

Ensembl - http://www.ensembl.org  

InterPro - http://www.ebi.ac.uk/interpro  

Predict Protein - http://ppopen.informatik.tu-muenchen.de