

# WEB SERVICES EN BIOINFORMÁTICA DAS, BIOMOBY Y TAVERNA PRÁCTICAS

José Manuel Rodríguez Carrasco

José María Fernández González

Biología Estructural y Bioinformática, CNIO  
GN2 (CNIO), INB





# Especificaciones DAS

- **dsn**  
Comando: `PREFIX/das/dsn`  
Función: Retorna los anotaciones que disponible el servidor DAS.
- **entry\_points**  
Comando: `PREFIX/das/DSN/entry_points`  
Función: Retorna la lista de las posibles entradas y sus respectivos tamaños para el servidor DAS.
- **sequence**  
Comando:  
`PREFIX/das/DSN/sequence?segment=RANGE[;segment=RANGE]`  
Función: Retorna la secuencia (nucleotidos o amino acidos) correspondiente al segmento dado.
- **dna**  
Comando: `PREFIX/das/DSN/dna?segment=RANGE[;segment=RANGE]`  
Función: Retorna la secuencia de "DNA" correspondiente al segmento dado.



# Especificaciones DAS (II)

- **types**

Comando: *PREFIX/das/DSN/types[?segment=RANGE][;segment=RANGE][;type=TYPE][;type=TYPE]*

Función: Retorna las anotaciones disponibles para un segmento dado.

- **features**

Comando: *PREFIX/das/DSN/features?segment=REF:start,stop[;segment=REF:start,stop][;type=TYPE][;type=TYPE][;category=CATEGORY][;category=CATEGORY]*

Función: Retorna las anotaciones cruzadas de una o más secuencias (nucleótidos o amino ácidos).

- **link**

Comando: *PREFIX/das/DSN/link?field=TAG;id=ID*

Función: Retorna una página web con descripciones sobre la anotación dada.

- **stylesheet**

Comando: *PREFIX/das/DSN/stylesheet*

Función: Retorna la hoja de estilos correspondiente.



# Respuesta de la lista de servidores DAS de anotaciones

```
<DASGFF>
  <DSN>
    <SOURCE id="elegans">
      elegans
    </SOURCE>
    <MAPMASTER>
      http://www.wormbase.org/db/das/elegans
    </MAPMASTER>
    <DESCRIPTION>
      C. elegans annotations from WormBase
    </DESCRIPTION>
  </DSN>
</DASGFF>
```



# Respuesta de un servidor DAS a sus puntos de entradas

```
<DASEP>
  <ENTRY_POINTS
    href="http://vab.wormbase.org:8080/db/das/elegans/entry_points"
    version="default">
    <SEGMENT id="00" size="100" subparts="yes" />
    <SEGMENT id="01" size="100" subparts="yes" />
    <SEGMENT id="02" size="100" subparts="yes" />
    ...
  </ENTRY_POINTS>
</DASEP>
```



# Respuesta de un servidor DAS a una secuencia dada (nucleotidos o aminoácidos)

<DASSEQUENCE>

<SEQUENCE id="P11766" version="X.XX" start="3" stop="300" moltype="Protein">

EVIKCKAAVAWEAGKPLSIEEIEVAPPKAHEVRIKIIATAVCHTDAYTLSGADPEGCFPV  
ILGHEGAGIVESVGEGVTKLKAGDTVIPLYIPQCGECKFCLNPKTNL CQKIRVTQ GKGLM  
PDGTSRFTCKGKTI LHYMG TSTFSEYTVVADISVAKIDPLAPLDKVCLLGCGISTGYGAA  
VNTAKLEPGSVCAVFLGGVGLAVIMGCKVAGASRIIGVDINKDKFARAKEFGATECINP  
QDFSKPIQEVLIEMTDGGVDYSFECIGNVKVMRAALEACHKGWGVSVVVGVAASGEEI

</SEQUENCE>

</DASSEQUENCE>



# Respuesta de un servidor DAS a las anotaciones de una proteína

```

<DASGFF>
  <GFF version="1.0" href="http://www.ebi.ac.uk/das-srv/uniprot/das/aristotle/features?segment=P11766">
    <SEGMENT id="P11766" start="1" stop="373">
      <FEATURE id="P11766" label="P11766">
        <TYPE id="description" category="">description</TYPE>
        <METHOD id="description">description</TYPE>
        <START>0</START>
        <END>0</END>
        <SCORE>-</SCORE>
        <ORIENTATION>0</ORIENTATION>
        <PHASE>-</PHASE>
        <NOTE>
          Alcohol dehydrogenase class 3 chi chain (EC 1.1.1.1) (Alcohol dehydrogenase class III
          chi chain) (S-(hydroxymethyl)glutathione dehydrogenase) (EC 1.1.1.284) (Glutathione-
          dependent formaldehyde dehydrogenase) (FDH).
        </NOTE>
        <LINK href="http://www.ebi.uniprot.org/uniprot-srv/uniProtView.do?proteinAc=P11766">
          http://www.ebi.uniprot.org/uniprot-srv/uniProtView.do?proteinAc=P11766
        </LINK>
      </FEATURE>
    </SEGMENT>
  </GFF>
</DASGFF>

```



# DAS v1

The response (DAS status codes):

200	OK, data follows
400	Bad command (command not recognized)
401	Bad data source (data source unknown)
402	Bad command arguments (arguments invalid)
403	Bad reference object (reference sequence unknown)
404	Bad stylesheet (requested stylesheet unknown)
405	Coordinate error (sequence coordinate is out of bounds/invalid)
500	Server error, not otherwise specified
501	Unimplemented feature



# Clientes DAS

- **Das in Ensembl**

[http://www.ensembl.org/info/data/external\\_data/das/ensembl\\_das.html](http://www.ensembl.org/info/data/external_data/das/ensembl_das.html)

- **Dasty**

<http://www.e-protein.org/e-proteindastypr.html>

- **SPICE**

<http://www.efamily.org.uk/software/dasclients/spice/spice.shtml>

- **MaDas**

<http://madas.bioinfo.cnio.es>

- **More Clients**

*Omnigene, Geodesic (Java), DasView (Perl), Gbrowse (perl)*



# DAS in Ensembl

The screenshot shows a Mozilla Firefox browser window displaying the Ensembl DAS reference server page. The browser's address bar shows the URL [http://www.ensembl.org/info/data/external\\_data/das/ensembl\\_das.html](http://www.ensembl.org/info/data/external_data/das/ensembl_das.html). The page title is "DAS in Ensembl - Mozilla Firefox". The browser's menu bar includes File, Edit, View, History, Bookmarks, Tools, and Help. The browser's toolbar shows various icons for navigation and utility. The page content is organized into several sections:

- Your Ensembl**: Includes links for "Login or Register" and "About User Accounts".
- Help & Documentation**: Includes links for "About Ensembl", "Genomic Data", "Comparative Genomics", "Custom Annotation", "Downloads", "FASTA Database Files", "FTP Directory Structure", "Genome Annotation", "Searching Ensembl", "Help & Information", and "Software".
- Ensembl Archive**: Includes links for "View previous release of page in Archive!" and "Stable Archive! link for this page".
- DAS in Ensembl**: The main content area, which includes:
  - The Ensembl DAS reference server**: A section explaining that Ensembl provides a DAS reference server for genomic sequences, gene predictions, karyotypes, and ditags. It lists sources served from the server as XML documents, with links to <http://www.ensembl.org/das/dsn> and <http://www.ensembl.org/das/sources>.
  - Example requests**: A section explaining the format of DAS request URLs. It provides the format `protocol://site-prefix/das/data-source/command?arguments` and gives an example: [http://www.ensembl.org/das/Homo\\_sapiens.NCBI36.transcript/features?segment=13:31787617,31871806](http://www.ensembl.org/das/Homo_sapiens.NCBI36.transcript/features?segment=13:31787617,31871806), which requests all transcripts in a specific region on human chromosome 13. Another example is [http://www.ensembl.org/das/Gallus\\_gallus.WASHUC1.reference/sequence?segment=1:1,1000](http://www.ensembl.org/das/Gallus_gallus.WASHUC1.reference/sequence?segment=1:1,1000), which requests the first 1000 bp of the first chicken chromosome.
  - DAS stylesheet support in Ensembl**: A section explaining that Ensembl supports a "stylesheet" request to make feature displays nicer, with a link to the "Ensembl Stylesheet Support" document.
  - DAS clients in Ensembl and semantic extensions to the DAS specification**: A section explaining that DAS clients are built into several Ensembl displays to incorporate third-party annotation data, and that some displays introduce semantic extensions to the DAS specification.
  - ContigView and CytoView**: A section explaining that a genomic DAS source serves data in the format specified in the DAS/1.5 specification, and that it can be attached to ContigView and CytoView. It also notes that for most species, Ensembl chromosomes are named like "1", "2", "X", "Y" rather than "chr1", "chr2", "chrX", "chrY".

At the bottom of the page, there is a banner for "Pufferfish Takifugu rubripes" and a footer with the URL [http://www.ensembl.org/Homo\\_sapiens/geneview?gene=P51587;db=core;das\\_sources=UniProt](http://www.ensembl.org/Homo_sapiens/geneview?gene=P51587;db=core;das_sources=UniProt).



# Ensembl ProtView

Ensembl release 45: Mus musculus Peptide Report for ENSMUSP00000026455 - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://www.ensembl.org/Mus\_musculus/protview?db=core;transcript=ENSMUST00000026455

Disables Cookies CSS Forms Images Information Miscellaneous Outline Resizer Tools View Source Options

DAS in Ensembl Ensembl release 45: Mu...

Ensembl release 45 - Jun 2007 HOME · BLAST · BIOMART · SITEMAP HELP

### Your Ensembl

- Login or Register
- About User Accounts

### ENSMUSP00000026455

- Gene Information
- Gene regulation info.
- Genomic sequence alignment
- Gene splice site image
- Gene tree info
- Gene variation info.
- ID history
- Genomic sequence
- Transcript information
- Exon information
- Protein information
- Export protein data

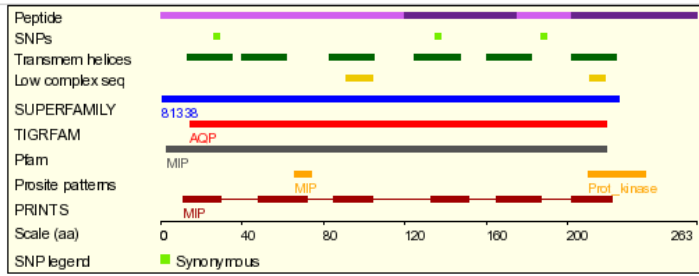
### Chromosome 10

127,628,831 - 127,633,453

- View of Chromosome 10
- Graphical view
- Graphical overview
- Export information about region
- Export sequence as FASTA
- Export EMBL file
- Export Gene info in region
- Export SNP info in region
- Export Vega info in region

### Ensembl Archive

## Ensembl Protein Report

<b>Peptide</b>	<a href="#">Mlp</a> (MGI Symbol) To view all Ensembl genes linked to the name <a href="#">click here</a> . This peptide is a member of the Mouse CCDS set: <a href="#">CCDS24265</a>
<b>Ensembl Peptide ID</b>	ENSMUSP00000026455
<b>Translation information</b>	This protein is a translation of transcript <a href="#">ENSMUST00000026455</a> , which is a product of gene <a href="#">ENSMUSG00000025389</a> .
<b>Genomic Location</b>	This peptide can be found on Chromosome 10 at location <a href="#">127,628,831-127,633,453</a> . The start of this peptide is located in <a href="#">Contig AC135859.4</a> .
<b>Description</b>	major intrinsic protein of eye lens fiber [Source:MarkerSymbol;Acc:MGI:96990]
<b>Prediction Method</b>	Genes were annotated by the Ensembl automatic analysis pipeline using either a GeneWise/Exonerate model from a database protein or a set of aligned cDNAs followed by an ORF prediction. GeneWise/Exonerate models are further combined with available aligned cDNAs to annotate UTRs (For more information see V.Curwen et al., Genome Res. 2004 14:942-50)
<b>InterPro</b>	<a href="#">IPR000719</a> Protein kinase - [ <a href="#">View other genes with this domain</a> ] <a href="#">IPR012269</a> Aquaporin - [ <a href="#">View other genes with this domain</a> ] <a href="#">IPR000425</a> Major intrinsic protein - [ <a href="#">View other genes with this domain</a> ]
<b>Protein Family</b>	<a href="#">ENSF00000000408</a> : AQUAPORIN AQP AQUAPORIN WATER CHANNEL This cluster contains 7 Ensembl gene member(s) in this species.
<b>Protein Features</b>	
<b>Protein Sequence</b>	<pre>MMEIERSASFWRAITFAEFFATLFFYVFFGLGASLRWADGPHLWQALAFGLALATLVQTVG HISGAHWNPVITFAFLVGSQMSLLRAFICYIAAQLLGAVAGAALVYSVTPPAVRGNLALNT LHAGISVGGATTVEIFLTLQFVLCIFATYDERRGRMG6SVALAVGFSLLTGLHFGMYTGG AGMNPARSFAPAILLTRFNSHMWYVWGPITGGGLGSLLYDFLLFRLKLSVSEKLSILKGA RPSDSNGQPEGTGEPVETKQAL</pre>

Show the following features:



# Ensembl GeneView

**Ensembl release 45: Homo sapiens Gene report for ENSG00000139618 - Mozilla Firefox**

File Edit View History Bookmarks Tools Help

http://www.ensembl.org/Homo\_sapiens/geneview?db=core:gene=ENSG00000139618

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options

DAS in Ensembl Ensembl release 45: Mus m... Ensembl release 45: Ho...

**e!Ensembl Human GeneView** Search: e! Human EBI Sanger -- Go

e.g. [ENSG00000139618](#), [ENSG00000138411](#)

Ensembl release 45 - Jun 2007 HOME · BLAST · BIOMART · SITEMAP HELP

**Your EnsEMBL**

- Login or Register
- About User Accounts

**ENSG00000139618**

- Gene Information
- Gene regulation Info.
- Genomic sequence
- Genomic sequence alignment
- Gene splice site Image
- Gene tree Info.
- Gene variation Info.
- LD Info
- ID history
- Compare SNPs in transcript
- Transcript Information
- Exon Information
- Protein Information
- Export gene data

**Chromosome 13**  
31,787,617 - 31,871,806

- View of Chromosome 13
- Graphical view
- Graphical overview
- Export Information about region
- Export sequence as FASTA
- Export EMBL file
- Export Gene Info In region
- Export SNP Info In region

**Ensembl Gene Report for ENSG00000139618**

<b>Gene</b>	<a href="#">BRCA2</a> (HGNC Symbol) To view all Ensembl genes linked to the name <a href="#">click here</a> . This gene is a member of the Human CCDS set: <a href="#">CCDS9344</a>
<b>Ensembl Gene ID</b>	ENSG00000139618
<b>Genomic Location</b>	This gene can be found on Chromosome 13 at location <a href="#">31,787,617-31,871,806</a> . The start of this gene is located in <a href="#">Contig AL445212.9.1.166657</a> .
<b>Description</b>	Breast cancer type 2 susceptibility protein (Fanconi anemia group D1 protein). <a href="#">Source: Uniprot/SWISSPROT P51587</a>
<b>Prediction Method</b>	gene containing both ensembl predicted transcripts and havana manual annotation

**Transcripts**

[ENST00000380152](#)    ENSP00000369497    [BRCA2\\_HUMAN](#)    [[Transcript info](#)]    [[Exon info](#)]    [[Peptide info](#)]

**Features**

Chr. 13  
Length 104,19 Kb  
Forward strand

Ensembl trans. [BRCA2\\_HUMAN](#) > Common Known Protein coding

DNA(contigs) [AL445212.9.1.166657](#) >

Ensembl trans. < NP\_001000001 > Common Known Protein Coding  
< NP\_001000001 > Common Known Protein Coding  
< NP\_431500 > Havana manual annotation  
< Q9Y270 > Ensembl  
< NP\_431500 > Ensembl  
< NM\_001000001 > Ensembl  
< Q5T000 > Ensembl

Length 104,19 Kb Reverse strand

**Alignments** This gene can be viewed in genomic alignment with other species



**e-PROTEIN**

### E-Protein Uniprot Search

Space separated words will be combined using '&' (AND)  
 Separate words with '|' (OR) to select entries containing either word  
 Wildcards (\*) will match any character(s)

Enter Query  
 ID/Accession  or Description   
 Select Organism  or enter names

Database: Uniprot Swissprot  Uniprot TrEMBL

595 hits found. (1)..2..3..4..5.....60 >

Links	Entry Name	Primary Accession	Description	Organism	Seq Length (AA)
<a href="#">D</a> <a href="#">U</a>	Q7WZ13_HAESO	Q7WZ13	High molecular weight-immunoglobulin binding proteins.	Haemophilus somnus	4095
<a href="#">D</a> <a href="#">U</a>	IGA1A_STRPN	Q97QP7	Immunoglobulin A1 protease precursor (EC 3.4.24.13) (IgA1 protease) (IgA-specific zinc metalloproteinase).	Streptococcus pneumoniae	2004
<a href="#">D</a> <a href="#">U</a>	IGA1_STRR6	Q59947	Immunoglobulin A1 protease precursor (EC 3.4.24.13) (IgA1 protease) (IgA-specific zinc metalloproteinase).	Streptococcus pneumoniae	1963
<a href="#">D</a> <a href="#">U</a>	IGA1B_STRPN	Q54875	Immunoglobulin A1 protease precursor (EC 3.4.24.13) (IgA1 protease) (IgA-specific zinc metalloproteinase).	Streptococcus pneumoniae	1927
<a href="#">D</a> <a href="#">U</a>	Q7X2A1_LEPIN	Q7X2A1	Immunoglobulin-like B protein.	Leptospira interrogans serovar Pomona	1889
<a href="#">D</a> <a href="#">U</a>	IGA1_STRSA	Q59986	Immunoglobulin A1 protease precursor (EC 3.4.24.13) (IgA1 protease) (IgA-specific zinc metalloproteinase).	Streptococcus sanguis	1854
<a href="#">D</a> <a href="#">U</a>	IGA4_HAEIN	P45386	Immunoglobulin A1 protease precursor (EC 3.4.21.72) (IGA1 protease).	Haemophilus influenzae	1849
<a href="#">D</a> <a href="#">U</a>	IGA2_HAEIN	P45384	Immunoglobulin A1 protease precursor (EC 3.4.21.72) (IGA1 protease).	Haemophilus influenzae	1702
<a href="#">D</a> <a href="#">U</a>	IGA0_HAEIN	P44969	Immunoglobulin A1 protease precursor (EC 3.4.21.72) (IGA1 protease).	Haemophilus influenzae	1694

Done



# Dasty

**Dasty - Mozilla Firefox**

File Edit View History Bookmarks Tools Help

http://www.ebi.ac.uk/das-srv/uniprot/dasty/content?ID=P51180:dis=DS\_168;DS\_174;DS\_199;DS\_200;DS\_216;

Disable Cookies CSS Forms Images Information Miscellaneous Outline Resize Tools View Source Options

DAS in Ensembl e-Protein: DASTY Protein DA... Dasty Dasty

**Dasty** Help

**NAME:** Lens fiber major intrinsic protein

**ID:** P51180

**LENGTH:** 263

**FEATURES** SORT

Positional  Default

Non Positions  Type

Both  Score

**VIEW**

Compressed  Expanded  Search again

**UNIPROT:**

**PDBSUM LIGANDS:**

**CHAIN**

Cath

MOTIF

MOTIF

PIR

TIGRFAMs

TOPO\_DOM

TOPO\_DOM

TOPO\_DOM

Roll over onto feature positions shows information in here

**pubmed:** 16141072 Carninci P., Kasukawa T., Katayama S., Gough J., Frith M.C., Maeda N., Oyama R., Ravas

**PDBSUM LIGANDS:**

BNG B-NONYLGLUCOSIDE

Done



# SPICE

The screenshot displays the SPICE web application interface. At the top, a Mozilla Firefox browser window shows the URL <http://www.efamily.org.uk/software/dasclients/spice/spice.shtml>. Below the browser, the application window is titled "SPICE - 1R3E - P00525" and includes a menu with "File", "Display", "Browse", and "Alignment".

The main interface is divided into two main sections:

- Left Panel:** A 3D molecular model of the protein structure, showing a complex fold with various colored spheres (pink, cyan, purple) and a stick representation of the backbone.
- Right Panel:** A series of horizontal tracks for protein annotations, including:
  - PDB:** Shows the PDB ID 1R3E.A and a sequence scale from 10 to 140.
  - dssp:** Displays secondary structure elements such as SECSTRUK, BEND, 3HELIX, and BRIDGE.
  - cath:** Shows the CATH domain classification s3dm, with sub-elements ALPHA-BE and ASX-MOT.
  - UniProt:** Shows the UniProt ID Q9WZWO and a sequence scale from 10 to 140. It includes annotations for CHAIN, SECSTRUK, ACT SITE, DOMAIN, TIGRFAMS, PROSITE, and PIR.
  - signalp:** Shows signal peptide predictions.
  - ENSP:** Shows ENSP annotations.
  - Other tracks:** superfam, ensp\_pdb, MisPred, hsa35pep, and Tanno are also visible.

At the bottom left, there is a text input field with the prompt "enter RASMOL like command..." and a status bar showing "Seq pos 101(E, GLU) PDB (110)". At the bottom right, a progress indicator shows "49 %".



MaDas - Mozilla Firefox

File Edit View History Bookmarks Tools Help

http://madas.bioinfo.cnio.es/MaDas/cgi-bin/MaDas

salud

Disable Cookies CSS

Type the protein name or the accession number

Q8IZT6 Submit Query clear

Overview

Sequence length

Details window position

Region from 1 to 3477 bp is shown in the detail window

Details

Window size 100 kbp <<< >>>

\* Click on the annotations to perform the view, edit or drop operations. Gray tracks correspond to MaDas annotations

CHAIN  
ASPM\_HUMAN\_CHAIN\_1\_3477  
Abnormal spindle-like microcephaly-associated protein/FTId=PR0\_0000191332

COILED  
ASPM\_HUMAN\_COILED\_1057\_1078

CONFLICT  
ASPM\_HUMAN\_CONFLICT\_1090\_1090 S -&#62; F (in Ref. 3; AAH34607)  
ASPM\_HUMAN\_CONFLICT\_2355\_2355 R -&#62; Q (in Ref. 4; BAA91660)  
ASPM\_HUMAN\_CONFLICT\_... I -&#62; V (in Ref. ...)  
ASPM\_HUMAN\_CONFLI... F -&#62; S (in Ref. ...)

DOMAIN  
ASPM\_HUMAN\_DOMAIN\_920\_1056 CH 1  
ASPM\_HUMAN\_DOMAIN\_1632\_1661 IQ 4  
ASPM\_HUMAN\_DOMAIN\_2384\_2415 IQ 24  
ASPM\_HUMAN\_DOMAIN\_1824\_1853 IQ 28  
ASPM\_HUMAN\_DOMAIN\_2624\_2653 IQ 38  
ASPM\_HUMAN\_DOMAIN\_2093\_2124 IQ 16  
ASPM\_HUMAN\_DOMAIN\_2859\_28... IQ 32  
ASPM\_HUMAN\_DOMAIN\_2116\_2147 IQ 33  
ASPM\_HUMAN\_DOMAIN\_2909... IQ 36  
ASPM\_HUMAN\_DOMAIN\_1582\_1613 IQ 3  
ASPM\_HUMAN\_DOMAIN\_2311\_2342 IQ 36  
ASPM\_HUMAN\_DOMAIN\_1655\_1684 IQ 5  
ASPM\_HUMAN\_DOMAIN\_2407\_2438 IQ 25  
ASPM\_HUMAN\_DOMAIN\_1728\_1757 IQ 6  
ASPM\_HUMAN\_DOMAIN\_2457\_2488 IQ 26  
ASPM\_HUMAN\_DOMAIN\_1751\_1782 IQ 7  
ASPM\_HUMAN\_DOMAIN\_2530\_2561 IQ 27  
ASPM\_HUMAN\_DOMAIN\_1801\_1830 IQ 8  
ASPM\_HUMAN\_DOMAIN\_2665\_2696 IQ 29  
ASPM\_HUMAN\_DOMAIN\_1874\_1903 IQ 8  
ASPM\_HUMAN\_DOMAIN\_2688\_2719 IQ 29

Annotate your sequence

Feature name\*

Type\* Pick one ...

Or enter some new type ...

Category

Method\* MaDas

Start\* 1 End\* 3477

Score

Orientation\* 0

Phase\* -

Note

Target name

Target Start Target End

Group Name Group Type

Group Note

Submit Annotation

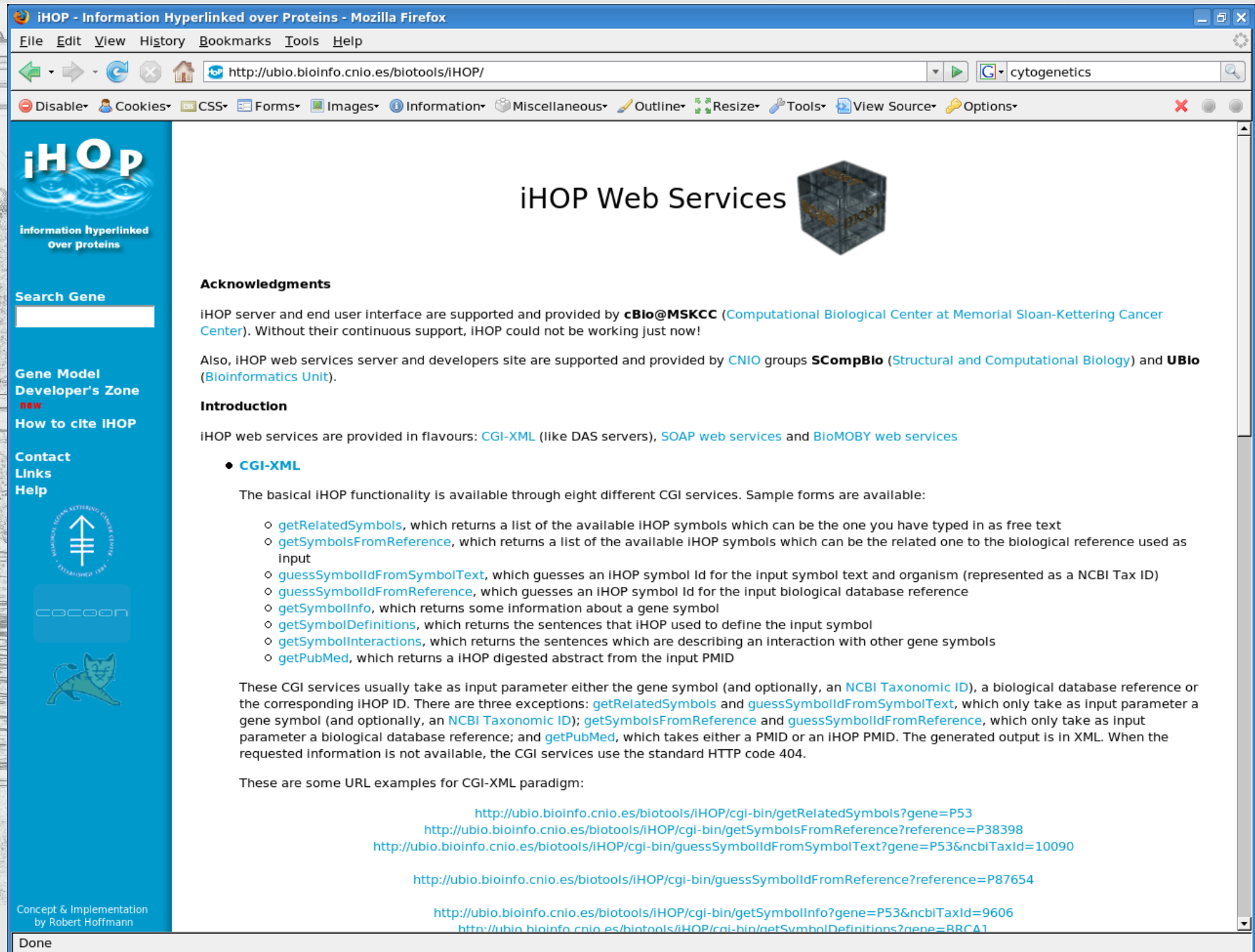
We are using as a model to upload the annotations the DAS feature response. Please consult the [DAS specification](#) and view the [example](#). The [HELP page](#) also provide information about it.

http://madas.bioinfo.cnio.es/MaDas/



# iHOP CGI-XML Services

17



The screenshot shows a Mozilla Firefox browser window displaying the iHOP website. The browser's address bar shows the URL <http://ubio.bioinfo.cnio.es/biotools/iHOP/>. The page title is "iHOP - Information Hyperlinked over Proteins". The browser's menu bar includes File, Edit, View, History, Bookmarks, Tools, and Help. The toolbar contains various icons for navigation and actions. The website content is displayed in a blue-themed layout. On the left, there is a sidebar with the iHOP logo, a search bar, and links for "Gene Model Developer's Zone", "How to cite iHOP", and "Contact Links Help". The main content area features the heading "iHOP Web Services" with a 3D cube icon. Below this, there are sections for "Acknowledgments", "Introduction", and "CGI-XML". The "CGI-XML" section lists eight services and provides several example URLs.

**iHOP**  
Information Hyperlinked  
Over Proteins

**Search Gene**

**Gene Model Developer's Zone**

**How to cite iHOP**

**Contact Links Help**

**iHOP Web Services**

**Acknowledgments**

iHOP server and end user interface are supported and provided by **cbio@MSKCC** (Computational Biological Center at Memorial Sloan-Kettering Cancer Center). Without their continuous support, iHOP could not be working just now!

Also, iHOP web services server and developers site are supported and provided by CNIO groups **SCompBio** (Structural and Computational Biology) and **UBIO** (Bioinformatics Unit).

**Introduction**

iHOP web services are provided in flavours: [CGI-XML](#) (like DAS servers), [SOAP web services](#) and [BioMOBY web services](#)

- **CGI-XML**

The basical iHOP functionality is available through eight different CGI services. Sample forms are available:

- [getRelatedSymbols](#), which returns a list of the available iHOP symbols which can be the one you have typed in as free text
- [getSymbolsFromReference](#), which returns a list of the available iHOP symbols which can be the related one to the biological reference used as input
- [guessSymbolIdFromSymbolText](#), which guesses an iHOP symbol Id for the input symbol text and organism (represented as a NCBI Tax ID)
- [guessSymbolIdFromReference](#), which guesses an iHOP symbol Id for the input biological database reference
- [getSymbolInfo](#), which returns some information about a gene symbol
- [getSymbolDefinitions](#), which returns the sentences that iHOP used to define the input symbol
- [getSymbolInteractions](#), which returns the sentences which are describing an interaction with other gene symbols
- [getPubMed](#), which returns a iHOP digested abstract from the input PMID

These CGI services usually take as input parameter either the gene symbol (and optionally, an [NCBI Taxonomic ID](#)), a biological database reference or the corresponding iHOP ID. There are three exceptions: [getRelatedSymbols](#) and [guessSymbolIdFromSymbolText](#), which only take as input parameter a gene symbol (and optionally, an [NCBI Taxonomic ID](#)); [getSymbolsFromReference](#) and [guessSymbolIdFromReference](#), which only take as input parameter a biological database reference; and [getPubMed](#), which takes either a PMID or an iHOP PMID. The generated output is in XML. When the requested information is not available, the CGI services use the standard HTTP code 404.

These are some URL examples for CGI-XML paradigm:

<http://ubio.bioinfo.cnio.es/biotools/iHOP/cgi-bin/getRelatedSymbols?gene=P53>  
<http://ubio.bioinfo.cnio.es/biotools/iHOP/cgi-bin/getSymbolsFromReference?reference=P38398>  
<http://ubio.bioinfo.cnio.es/biotools/iHOP/cgi-bin/guessSymbolIdFromSymbolText?gene=P53&ncbiTaxId=10090>  
<http://ubio.bioinfo.cnio.es/biotools/iHOP/cgi-bin/guessSymbolIdFromReference?reference=P87654>  
<http://ubio.bioinfo.cnio.es/biotools/iHOP/cgi-bin/getSymbolInfo?gene=P53&ncbiTaxId=9606>  
<http://ubio.bioinfo.cnio.es/biotools/iHOP/cgi-bin/getSymbolDefinitions?gene=BRC1>

Concept & Implementation  
by Robert Hoffmann



# Cómo usar los iHOP CGI-XML Services

**NAME**

- [NAME](#)
- [SYNOPSIS](#)
- [NOTES](#)
- [THE iHOPUserLibrary PACKAGE](#)
- [RELATED MODULES](#)
- [AUTHORS](#)

**NAME**

iHOPUserLibrary - A library which eases the use of iHOP CGI-XML web services

**SYNOPSIS**

```
use iHOPUserLibrary;
use HTTP::Response;
use XML::LibXML;

my($result)=iHOPUserLibrary::getSymbolInteractionsFromIHOP(32484);

if($result->is_success) {
    print $result->content;
}

my($definitions)=iHOPUserLibrary::getSymbolDefinitionsFromReference('P31947');

my($relatedSymbols)=iHOPUserLibrary::getRelatedSymbols('breast cancer');
if($relatedSymbols->is_success) {
    my($parser)=XML::LibXML->new();
    my($document)=$parser->parse_string($relatedSymbols->content);
}


```

**NOTES**

The information generated by this library was obtained from the iHOP server developed and created by Robert Hoffmann. iHOP - Information Hyperlinked over Proteins is licensed under a Creative Commons Attribution 2.5 License (<http://creativecommons.org/licenses/by/2.5/>).

The information generated by this library was obtained using the iHOP web services developed and created by José María Fernández González, CNIO.

Please fill-in either a gene symbol, a biological reference or iHOP symbol id

This is a demonstration HTML form about usage of [iHOP](#) [getSymbolInfo](#) [CGI-XML](#) service

Free text  NCBI TaxID\*  OR Bio Reference  OR iHOP symbol id

\* This is an optional parameter which constrains the search



# Taverna, un “banco de trabajo” para *web services*

- Taverna ( <http://taverna.sourceforge.net> ) es una aplicación Java que permite la construcción de flujos de trabajos a partir de *web services*.
- *Para más información:*

[http://www.mygrid.org.uk/usermanual1.5/workbench\\_windows.html](http://www.mygrid.org.uk/usermanual1.5/workbench_windows.html)





# Vista global de Taverna

The screenshot displays the Taverna Workbench v1.5.1.6 interface. The top menu bar includes 'File', 'Tools', 'Workflows', and 'Advanced'. Below the menu is a 'Design' tab and a 'Results' tab. The main workspace is divided into several panels:

- Available Processors:** A list of web services with their WSDL URLs, such as 'Local Services', 'WSDL @ http://soap.bind.ca/wsdll/bind.wsdl', and 'Biomart service @ http://www.biomart.org/biomart'.
- Advanced model explorer:** A tree view showing the workflow structure, including 'Workflow inputs' (gene input), 'Workflow outputs' (related symbols, Symbols, Genes, related PubMed), and 'Processors' (xpath, iHOPpmidXPath, interactorsXPath, extractPMID, extractSymbols, Remove\_duplicate\_strings, extractInteractors, getSymbolInteractionsFromIHO, getPubMedFromIHO, getRelatedSymbols).
- Workflow Diagram:** A central flowchart showing the execution of processors. It starts with 'gene input' leading to 'getRelatedSymbols', which outputs 'input' and 'result'. This is followed by 'extractSymbols' and 'extractPMID', which output 'nodelist' and 'ihopid'/'ihoppmid'. The final outputs are 'Med', 'Symbols', and 'related symbols'.

Menu de aplicaciones

Diagrama del flujo de trabajo

Procesos Web disponibles

Exploración detallada del flujo de trabajo

# Web Services (WSDL) en Taverna

- Servicios WSDL:

<http://ubio.bioinfo.cnio.es/biotools/iHOP/iHOP-SOAP.wsdl>



The screenshot shows the Taverna Workbench interface. On the left, a search pane lists various web services from the iHOP WSDL. The main workspace displays a workflow diagram with the following components:

- Workflow Inputs:** A box labeled "gene" with a red triangle icon.
- Processor:** A green box labeled "getSymbolInfoFromSymbol".
- Workflow Outputs:** A box labeled "symbol output" with a green inverted triangle icon.

Below the workflow diagram, the "Advanced model explorer" shows a tree view of the workflow object, including workflow inputs, outputs, processors, data links, and control links.



# Servicios moby en Taverna

- MOBY Central:

<http://mobycentral.icapture.ubc.ca/cgi-bin/MOBY05/mobycentral.pl>

The screenshot displays the Taverna Workbench v1.5.1.6 interface. The main window shows a workflow diagram with the following components:

- Workflow Inputs:** A dashed box containing a 'namespace' object (with a 'value' field) and a 'gene' object (with a red triangle).
- Processing:** An 'Object' (yellow box) with 'namespace id' and 'mobyData' fields. Below it is an 'Object(search)' (yellow box) with 'getSymbolInfo' and 'iHOPsymbol(result)' methods.
- Workflow Outputs:** A dashed box containing a 'symbol info' object (blue box) with a green triangle.

The 'Advanced model explorer' at the bottom left shows the workflow structure:

- Workflow: wf-MOBY-getSymbolInfo
- Workflow inputs: gene
- Workflow outputs: symbol info
- Processors: getSymbolInfo, Object, namespace: Global, value: 'text/plain'
- Data links: Object.mobyData-getSymbolInfo: Object(search), gene-Object.id, namespace:value-Object:namespace, getSymbolInfo:iHOPsymbol(result)-symbol info
- Control links

Rendering done.