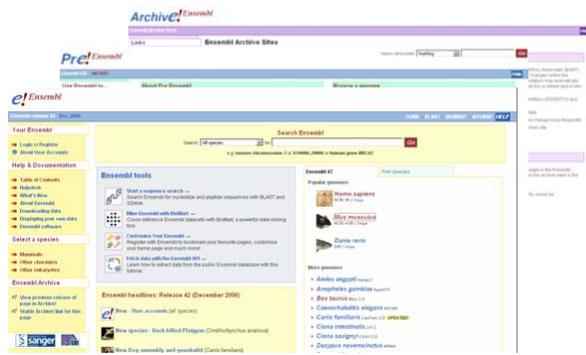


# DAS

## Integration of external annotation data



Xosé M<sup>a</sup> Fernández  
European Bioinformatics Institute

Feb 2007

## Outline of talk

- DAS
  - What is DAS?
  - DAS client DAS server
- DAS in Ensembl
  - Ensembl as DAS server
  - DAS essentials
  - Examples



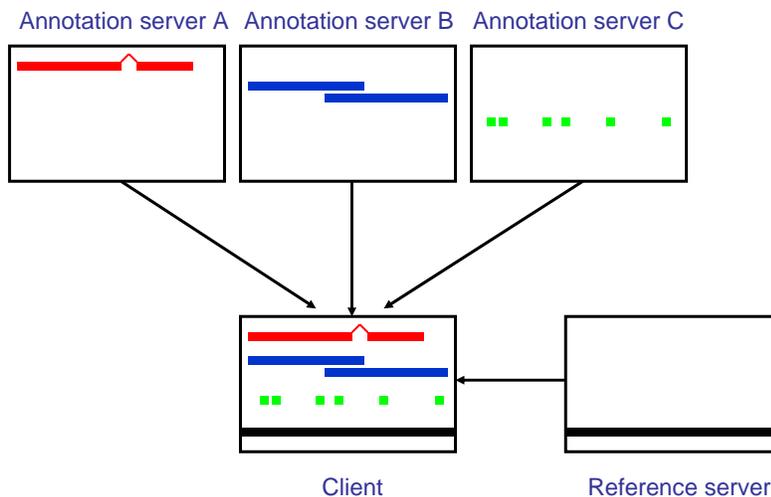
## What is DAS?

### DAS Distributed Annotation System

- Developed by Robin Dowell, Sean Eddy and Lincoln Stein (Dowell *et al.*, 2001)
- Specification of a protocol for requesting and returning sequence and annotation data
- Allows decentralised storage of sequence annotation, and integration on demand basis by client-side software
- Simple protocol that runs over HTTP

... see <http://biodas.org> for details

## DAS concept



## ***DAS concept***

The annotations may be shared more easily.

The amount of data which needs to be stored locally is decreased.

Responsibility for updating and maintaining the data is left with the original data provider.

Conflicting annotations are permitted, encouraging information dissension and dialogue.

Annotation can be stored in a variety of formats.

Other programs, even other types of programs (not annotation viewers), may use the data.

## ***DAS - Annotation***

- Annotation - An entity which:
  - Is anchored to the reference sequence via a stop and start value.
  - Possesses an ID unique to the server and a structured description of its types, methods, and categories.
- Global Annotation:
  - Annotation that applies to the entire reference sequence; start and stop are void

## ***DAS client DAS server***

### Definitions

- A **DAS client** is either an application (*Apollo*) or a web site (*Ensembl*) requesting and displaying annotations for a region of a sequence
- A **DAS server** is designated either as
  - a **reference server**, that provides essential structural information (e.g. entry points or sequence), or
  - an annotation server, that supplies specific annotation and has the knowledge about the sequence and/or annotation data stored in a flat file or database

## ***Ensembl, a DAS client***

### DAS clients in **Ensembl**

- **ContigView**  
Principal data visualisation tool for genome sequence annotation information
- **GeneView**  
Provides detailed information about gene model predictions
- **ProtView**  
Provides detailed information about protein model predictions

## Reference Sequences

- Applicable to **ContigView**
  - Chromosome
  - Contig
  - Clone
- Applicable to **ProtView** or **GeneView**
  - Protein
  - Gene (global annotation only)
  - cDNA (not yet implemented)

## DAS format

URL following a specific, standardised format:

`<prefix>/das/<dsn>/<cmd>?<args>`

- Site-specific prefix (server URL)
  - Keyword “das”
  - Data source name (dsn)
  - Command (cmd)
  - Arguments (args)

Example

`http://www.ebi.ac.uk/das-srv/uniprot/das/aristotle  
/features?segment=P51587`

## ***DAS format***

URL following a specific, standardised format:

`<prefix>/das/<dsn>/<cmd>?<args>`

- Site-specific prefix (server URL)
- **Keyword “das”**
- Data source name (dsn)
- Command (cmd)
- Arguments (args)

Example

`http://www.ebi.ac.uk/das-srv/uniprot/das/aristotle  
/features?segment=P51587`

## ***DAS format***

URL following a specific, standardised format:

`<prefix>/das/<dsn>/<cmd>?<args>`

- Site-specific prefix (server URL)
- Keyword “das”
- **Data source name (dsn)**
- Command (cmd)
- Arguments (args)

Example

`http://www.ebi.ac.uk/das-srv/uniprot/das/aristotle  
/features?segment=P51587`

## **DAS format**

URL following a specific, standardised format:

`<prefix>/das/<dsn>/<cmd>?<args>`

- Site-specific prefix (server URL)
- Keyword “das”
- Data source name (dsn)
- **Command (cmd)**
- Arguments (args)

Example

`http://www.ebi.ac.uk/das-srv/uniprot/das/aristotle  
/features?segment=P51587`

## **DAS format**

URL following a specific, standardised format:

`<prefix>/das/<dsn>/<cmd>?<args>`

- Site-specific prefix (server URL)
- Keyword “das”
- Data source name (dsn)
- Command (cmd)
- **Arguments (args)**

Example

`http://www.ebi.ac.uk/das-srv/uniprot/das/aristotle  
/features?segment=P51587`

## DAS essentials

<prefix>/das ...

- /dsn

Retrieves a list of data sources available from this server

- /<dsn>/types[?segment=<ref>]

Retrieves a list of types available from a data source or particular data source segments

- /<dsn>/features?segment=<ref>

Retrieves the annotations across a segment

Example

<http://www.ensembl.org/das/dsn>

'dsn' - Get the list of available sources

<http://www.ensembl.org/das/dsn>

<http://das.ensembl.org/das/dsn>

```
<?xml version="1.0" standalone="yes"?>
<!DOCTYPE DASDSN SYSTEM "http://www.biodas.org/dtd/dasdsn.dtd">
<DASDSN>
  <DSN>
    <SOURCE id="ena_34_affy10knp">ena_34_affy10knp</SOURCE>
    <MAPMASTER>http://das.ensembl.org/ensembl1834</MAPMASTER>
    <DESCRIPTION>Affymetryrix 10K SNP Array Mapping</DESCRIPTION>
  </DSN>
  <DSN>
    <SOURCE id="ena_35_affy10knp">ena_35_affy10knp</SOURCE>
    <MAPMASTER>http://das.ensembl.org/ensembl2635</MAPMASTER>
    <DESCRIPTION>Affymetryrix 10K SNP Array Mapping</DESCRIPTION>
  </DSN>
  <DSN>
    <SOURCE id="ena_anoph_dros_5_1">ena_anoph_dros_5_1</SOURCE>
    <MAPMASTER>http://servlet.sanger.ac.uk:8080/das/mosquito51</MAPMASTER>
    <DESCRIPTION>Ensembl-mapped Drosophila Genes</DESCRIPTION>
  </DSN>
  <DSN>
    <SOURCE id="ena_cbr25_fgenseh">ena_cbr25_fgenseh</SOURCE>
    <MAPMASTER>http://servlet.sanger.ac.uk:8080/das/cbriggsae1025</MAPMASTER>
    <DESCRIPTION>C. briggsae fgenseh predictions</DESCRIPTION>
  </DSN>
</DASDSN>
```

## ***DAS essentials***

<prefix>/das ...

- /dsn

Retrieves a list of data sources available from this server

- /<dsn>/types[?segment=<ref>]

•Retrieves a list of types available from a data source or particular data source segments

- /<dsn>/features?segment=<ref>

Retrieves the annotations across a segment

Example

[http://genome.dbs.dtu.dk:9000/das/cbs\\_ptm/tyes?segment=P51587](http://genome.dbs.dtu.dk:9000/das/cbs_ptm/tyes?segment=P51587)

## ***DAS essentials***

<prefix>/das ...

- /dsn

Retrieves a list of data sources available from this server

- /<dsn>/types[?segment=<ref>]

Retrieves a list of types available from a data source or particular data source segments

- /<dsn>/features?segment=<ref>

Retrieves the annotations across a segment

Example

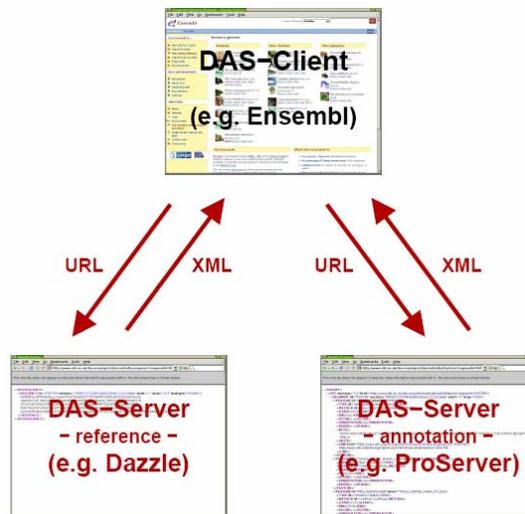
[http://das.ensembl.org/das/ens\\_36\\_refseq/features?segment=1:40000,60000](http://das.ensembl.org/das/ens_36_refseq/features?segment=1:40000,60000)

## 'features' - Get the annotations

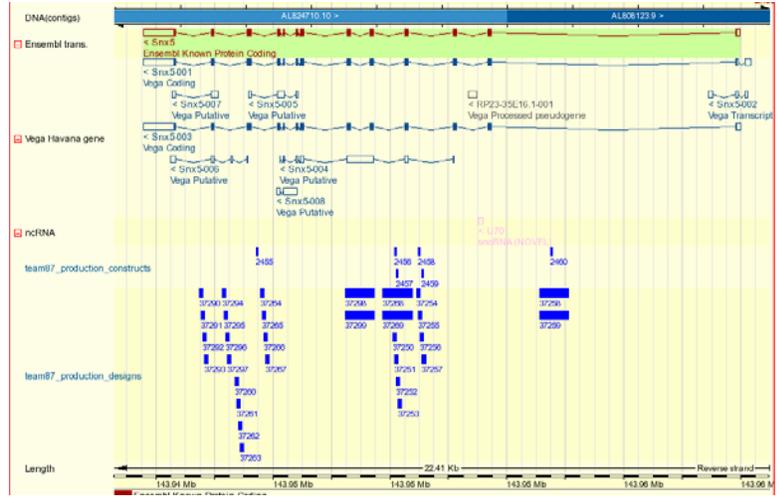
[http://www.ebi.ac.uk/das-srv/asd/das/atd\\_human/features?segment=13:3770000,3810000](http://www.ebi.ac.uk/das-srv/asd/das/atd_human/features?segment=13:3770000,3810000)

```
<?xml version="1.0" standalone="yes"?>
<!DOCTYPE DASGFF SYSTEM "http://www.biodas.org/dtd/dasgff.dtd">
<DASGFF>
<GFF version="1.01" href="http://web59-nodel.ebi.ac.uk:9100/das/atd_human/features">
<SEGMENT id="13" version="1.0" start="3770000" stop="3810000">
<FEATURE id="6438:1.215740" label="6438:1.215740">
<TYPE id="exon" reference="no" subparts="no" superparts="no">exon</TYPE>
<START>32658014</START>
<END>32658215</END>
<ORIENTATION>-</ORIENTATION>
<LINK href="http://www.ebi.ac.uk/asd-srv/Atd.cgi?method=TRANSCRIPT;origin=AT;product="
<GROUP id="6438" label="ENSG00000133121 AT-1" />
</FEATURE>
<FEATURE id="6438:1.215741" label="6438:1.215741">
<TYPE id="exon" reference="no" subparts="no" superparts="no">exon</TYPE>
<START>32639687</START>
<END>32639758</END>
<ORIENTATION>-</ORIENTATION>
<LINK href="http://www.ebi.ac.uk/asd-srv/Atd.cgi?method=TRANSCRIPT;origin=AT;product="
<GROUP id="6438" label="ENSG00000133121 AT-1" />
</FEATURE>
<FEATURE id="6438:1.215742" label="6438:1.215742">
<TYPE id="exon" reference="no" subparts="no" superparts="no">exon</TYPE>
<START>32637473</START>
<END>32637554</END>
<ORIENTATION>-</ORIENTATION>
<LINK href="http://www.ebi.ac.uk/asd-srv/Atd.cgi?method=TRANSCRIPT;origin=AT;product="
<GROUP id="6438" label="ENSG00000133121 AT-1" />
</FEATURE>
</GFF>
```

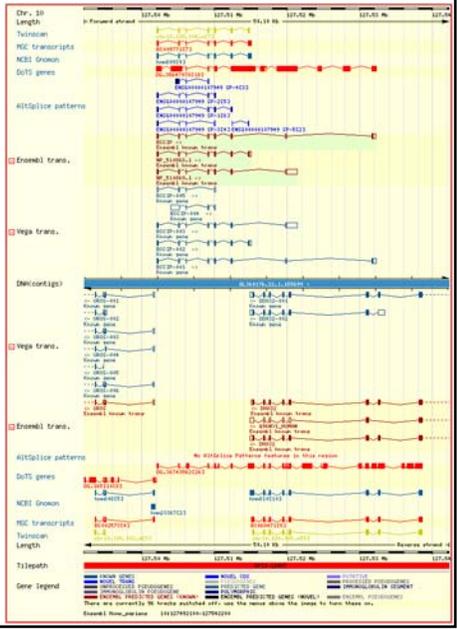
## Ensembl - Project



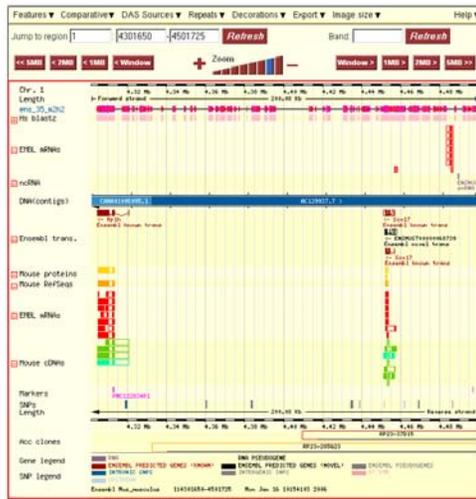
# EUCOMM



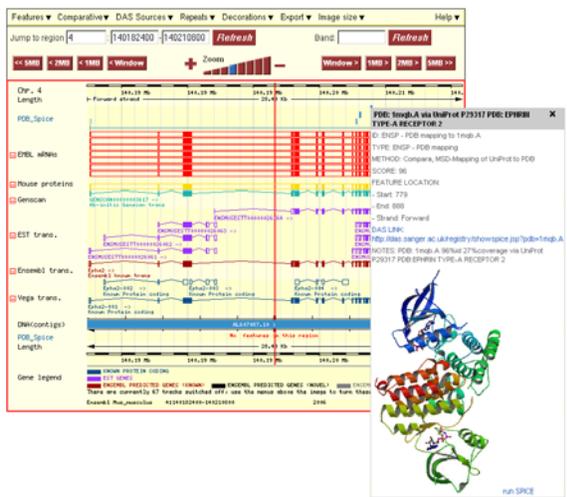
# Display 3rd party data



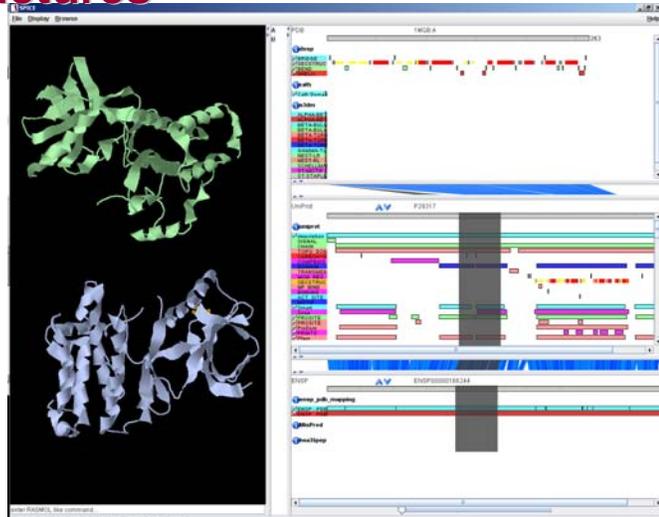
## Display Custom tracks



## Display 3D structures

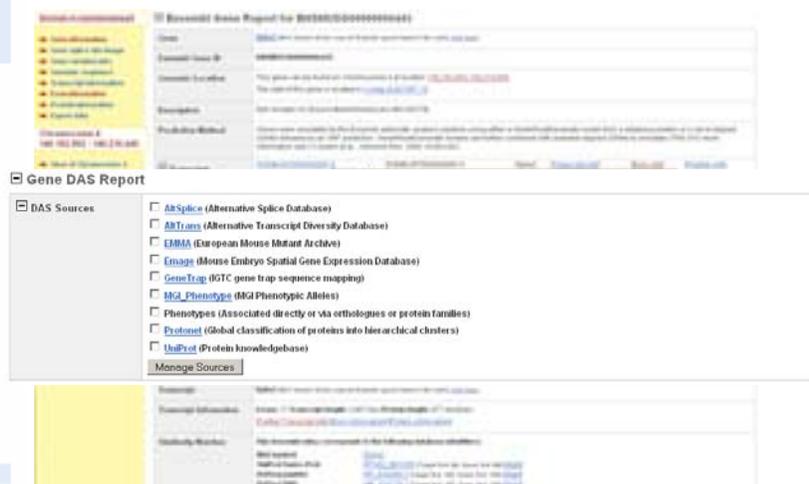


## Display 3D structures



The screenshot displays the Ensembl genome browser interface. On the left, two 3D protein structures are shown: a green ribbon structure and a grey ribbon structure. The main panel shows genomic tracks for the *TRIOA* gene, including the gene model, transcripts, and protein-coding regions. The interface includes a search bar and various navigation options.

## GeneDAS



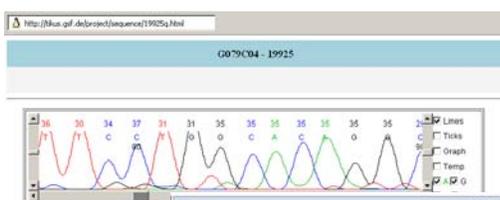
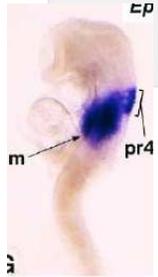
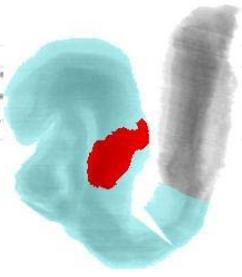
The screenshot shows the GeneDAS report interface. The top section displays a search result for the gene *TRIOA*. Below this, the 'Gene DAS Report' section is expanded, showing a list of DAS sources with checkboxes for selection. The sources include:

- [AltSplice](#) (Alternative Splice Database)
- [ALTtrans](#) (Alternative Transcript Diversity Database)
- [EMMA](#) (European Mouse Mutant Archive)
- [Emage](#) (Mouse Embryo Spatial Gene Expression Database)
- [GeneTrap](#) (IGTC gene trap sequence mapping)
- [MGD\\_Phenotype](#) (MGD Phenotypic Alleles)
- [Phenotypes](#) (Associated directly or via orthologues or protein families)
- [Prototax](#) (Global classification of proteins into hierarchical clusters)
- [UniProt](#) (Protein knowledgebase)

A 'Manage Sources' button is located at the bottom of the list. The interface also includes a search bar and various navigation options.

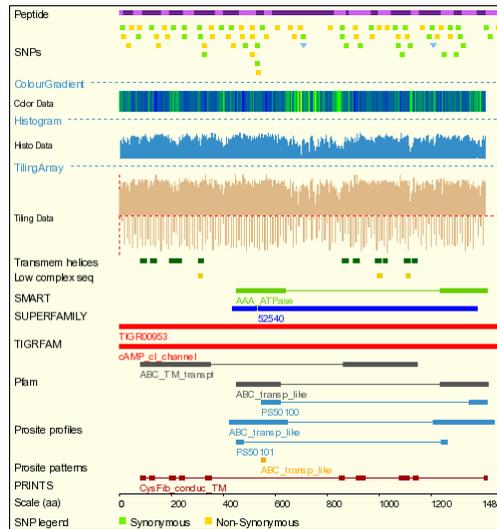
# GeneDAS

|                      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                          |
|----------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <b>EMage</b>         | description <a href="#">Epha2</a> description expression data in Their Stage <a href="#">L21</a> in <a href="#">20</a> Tissues                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                           |
| <b>MGI Phenotype</b> | Phenotypic Allele <a href="#">MG2248872</a> MGI SYMBOL: Epha2 <sup>206R05A(Betageno)CNI37m</sup> [NAME: g9 PUMBO: <a href="#">21127164</a> ]<br>Phenotypic Allele <a href="#">MG2268954</a> MGI SYMBOL: Epha2 <sup>306J(Betageno)J2A3Ru</sup> [NAME: gene <a href="#">1648815</a> ]<br>Phenotypic Allele <a href="#">MG2268815</a> MGI SYMBOL: Epha2 <sup>306J(Betageno)J2A3Ru</sup> [NAME: targeted mutation 1, ...]                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| <b>DAS Sources</b>   | <input type="checkbox"/> <a href="#">ASSplice</a> (Alternative Splice Database)<br><input type="checkbox"/> <a href="#">ALTomes</a> (Alternative Transcript Diversity Database)<br><input type="checkbox"/> <a href="#">EMMA</a> (European Mouse Mutant Archive)<br><input checked="" type="checkbox"/> <a href="#">EMage</a> (Mouse Embryo Spatial Gene Expression Database)<br><input type="checkbox"/> <a href="#">GeneTrap</a> (5'UTR gene trap sequence mapping)<br><input checked="" type="checkbox"/> <a href="#">MGI Phenotype</a> (MGI Phenotypic Alleles)<br><input type="checkbox"/> <a href="#">Phenotypes</a> (Associated directly or via orthologues or protein families)<br><input type="checkbox"/> <a href="#">Proteome</a> (Global classification of proteins into hierarchical clusters)<br><input type="checkbox"/> <a href="#">UniProt</a> (Protein knowledgebase)<br><input type="button" value="Manage Sources"/> |



| <p>Phenotypic Allele Detail</p> <p><b>Allele</b> <a href="#">Synbio: Epha2<sup>306J(Betageno)J2A3Ru</sup></a><br/> Name: gene trap (3A3), H Earl Raley<br/> MGI: 2668954</p> <p><b>Synopsis</b> <a href="#">eck<sup>1</sup></a></p> <p><b>Allele details</b><br/> Allele Type: Gene trapped<br/> Site of Origin: 1250'v<br/> ES Cell Line Not Specified<br/> ES Cell Line Strain Not Specified<br/> Mutation: Disruption caused by insertion of vector<br/> A protein-coding sequence, inserted into an exon 3.5 kb downstream of the alternatively spliced exon 5.2. Endogenous protein was undetectable in homozygous mutant mice by Western blot analysis. <a href="#">(L2202)</a><br/> International Mouse Strain Resource: <a href="#">Search for MGI strains</a> with Epha2 mutations)<br/> References and Additional Notes: <a href="#">[See below]</a></p> <p><b>Gene Information</b><br/> Name: <a href="#">Epha2</a><br/> Name: Eph receptor A2<br/> Chromosome: 4<br/> Genetic Position: 73.2 cM, cytoband D-E<br/> Genome Coordinates: 140102502-140210640 bp, + strand (from MGI annotation of MGI build 34)<br/> Mouse Ortholog: <a href="#">Epha2</a></p> <p><b>Phenotypes</b><br/> Phenotypic details for all genotypes that include at least one <a href="#">Epha2<sup>306J(Betageno)J2A3Ru</sup> allele</a></p> <table border="1"> <thead> <tr> <th>Allele Composition</th> <th>Genetic Background</th> </tr> </thead> <tbody> <tr> <td><a href="#">Epha2<sup>306J(Betageno)J2A3Ru</sup>/Epha2<sup>306J(Betageno)J2A3Ru</sup></a></td> <td>involves: L2N<sup>+/+</sup> * C57BL/6J</td> </tr> </tbody> </table> <p><b>mutual phenotype</b><br/> <a href="#">see Epha2 expression detected (L2202)</a></p> <p><b>References</b><br/> [Original] <a href="#">121972</a> Chen J et al., "Gene-line inactivation of the murine Eck receptor kinase by gene trap retroviral insertion." Oncogene 1996 Mar 7;11(5):979-88<br/> All references[1]</p> |                                         | Allele Composition | Genetic Background | <a href="#">Epha2<sup>306J(Betageno)J2A3Ru</sup>/Epha2<sup>306J(Betageno)J2A3Ru</sup></a> | involves: L2N <sup>+/+</sup> * C57BL/6J |
|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------|--------------------|--------------------|-------------------------------------------------------------------------------------------|-----------------------------------------|
| Allele Composition                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                              | Genetic Background                      |                    |                    |                                                                                           |                                         |
| <a href="#">Epha2<sup>306J(Betageno)J2A3Ru</sup>/Epha2<sup>306J(Betageno)J2A3Ru</sup></a>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       | involves: L2N <sup>+/+</sup> * C57BL/6J |                    |                    |                                                                                           |                                         |

## Style sheets



## Adding custom data to Ensembl Displays

- Add a new DAS server;
  - Switch on pre-configured server,
  - Configure a new server,
  - Set up your own server (see Ensembl docs),
  - ContigView, ProtView and GeneView.
  
- Use Ensembl DAS server to display your data, made accessible via;
  - file upload,
  - Web page (URL),
  - ContigView only at this time.

## URL-based Upload

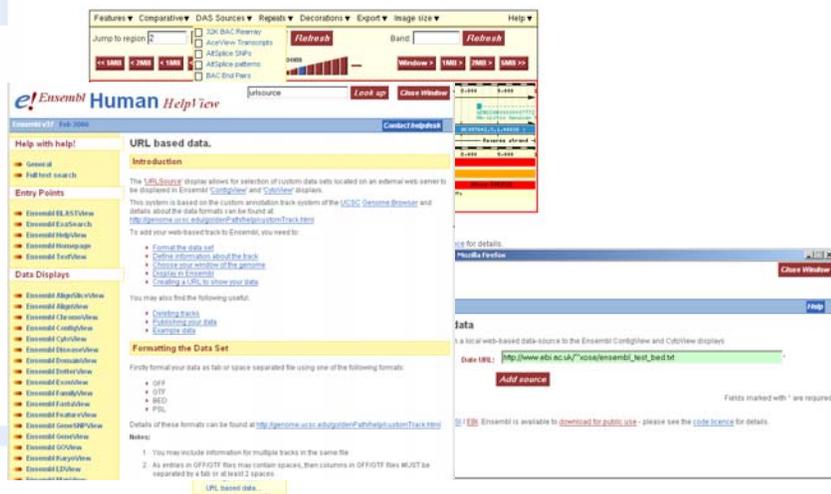
Based on the custom annotation track system of the UCSC browser.

Allowed formats: GFF, GTF, BED, PSL (see *also* <http://genome.ucsc.edu/goldenPath/help/customTrack.html>).

Display data by either entering the URL of the data file on the Urlsource page (ContigView: Data sources > URL based data ...) or by creating the following URL:

[http://www.ensembl.org/species/contigview?data\\_URL=URL](http://www.ensembl.org/species/contigview?data_URL=URL)

## Display of uploaded data



The screenshot shows the Ensembl Human Help page for 'URL based data'. The page is divided into several sections:

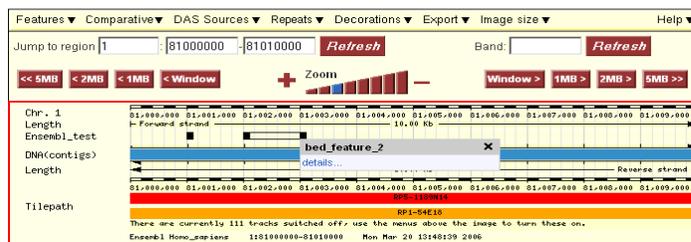
- Introduction:** Explains that the URL Source display allows for selection of custom data sets located on an external web server to be displayed in Ensembl's ContigView and Cytoside displays. It provides a link to the UCSC Genome Browser for more information on data formats.
- To add your web-based track to Ensembl, you need to:**
  - Format the data set
  - Define information about the track
  - Provide your website of the genome
  - Upload to Ensembl
  - Creating a URL to show your data
- You may also find the following useful:**
  - Contig Display
  - Submitting your data
  - Exporting data
- Formatting the Data Set:**
  - Firstly format your data as tab or space separated by using one of the following formats:
    - GFF
    - GTF
    - BED
    - PSL
  - Details of these formats can be found at <http://genome.ucsc.edu/goldenPath/help/customTrack.html>
  - Notes:**
    - You may include information for multiple tracks in the same file.
    - As entries in GFF/GTF files may contain spaces, then columns in GFF/GTF files MUST be separated by a tab or at least 2 spaces.
- URL Source data:** A section for adding a new URL source, with a text input field containing the URL: `http://www.ncbi.nlm.nih.gov/ensembl_test_data/` and an 'Add source' button.

## URL-based

```

browser position chr2:1-10000
track name=Ensembl_test description="Ensembl
workshop (BED)" color=000000
url=http://www.ebi.ac.uk/~xose/ensembl_test.html
2      1000   1100   bed_feature_1 1000   +
2      2000   2100   bed_feature_2 500    +
2      3000   3100   bed_feature_2 100    +
    
```

## Display of data via URL



The screenshot shows the Ensembl genome browser interface. At the top, there are navigation menus: Features, Comparative, DAS Sources, Repeats, Decorations, Export, Image size, and Help. Below these is a search bar with 'Jump to region' set to '1' and coordinates '81000000-81010000'. A 'Refresh' button is next to it. To the right, there is a 'Band:' field and another 'Refresh' button. Below the search bar is a zoom control with a '+' sign, a 'Zoom' label, and a '-' sign. Navigation buttons for '<< 5MB', '< 2MB', '< 1MB', '< Window', 'Window >', '1MB >', '2MB >', and '5MB >>' are also present. The main display area shows a genomic track for 'Chr. 1' with a length of '19.00 kb'. The 'Forward strand' is shown with a blue bar representing the 'bed\_feature\_2' track. A 'Reverse strand' is also indicated. Below the track, there is a 'Tilepath' section with a yellow background and the text 'RP1-54E18'. At the bottom, a message states: 'There are currently 111 tracks switched off; use the menus above the image to turn these on.' The footer shows 'Ensembl Home\_page' and '1081000000-81010000' with the date 'Mon Mar 20 15:49:19 2006'.

## Display of uploaded data

### DAS Wizard Step 2 of 3: Data appearance

Coordinate System: Provided by Registry

Enable on:  geneview  proteinview  transview  
 coreview  cybrowser

[Back](#) [Next](#)

### DAS Wizard Step 3 of 3: Display configuration

#### DAS sources

| Name             | DAS Server    |
|------------------|---------------|
| das_32K          | http://das1:  |
| das_ACEVIEW      | http://das.er |
| das_ALTSPLICE    | http://www.e  |
| das_ALTSPLICESNP | http://www.e  |
| das_ABas         | http://db.sys |
| das_ABasPlasma   | http://db.sys |
| das_BAC          | http://das1:  |
| das_CHORIS07     | http://das1:  |
| das_CHORI_243    | http://das.er |
| das_CISRED       | http://stand  |
| das_COMPUOEN     | http://das.er |
| das_CPG          | http://das.er |
| das_Decipher     | http://das1:  |
| das_DOTB         | http://das1:  |
| das_FIRSTEF      | http://das1:  |
| das_FOSMD        | http://das1:  |
| das_GCS_EnoFish  | http://www.e  |
| das_GCS_mRNA     | http://www.e  |

Name: As nickname in registry  
 Label: As nickname in registry

Help URL:   
 Link Text:   
 Link URL:

Track colour:

Group features:

Display on:

Label features:

Apply stylesheet:   
 Use score:

[Back](#) [Finish](#)

#### DAS sources

| Name             | DAS Server                            | Data Source                    | Coordinate System   |
|------------------|---------------------------------------|--------------------------------|---------------------|
| das_32K          | http://das1                           | sanger.ac.uk:7070/peridas      | ens_ncbi_35_22array |
| das_ACEVIEW      | http://das.ensembl.org/das            | ens_35_aceview                 | Ensembl Location    |
| das_ALTSPLICE    | http://www.ebi.ac.uk/das-srvs/das     | ssd_human                      | Ensembl Location    |
| das_ALTSPLICESNP | http://www.ebi.ac.uk/das-srvs/das     | ssd_stg_human                  | Ensembl Location    |
| das_ABas         | http://db.systemsbiology.net:8080/das | Human_Ena30_NCBI05_APD0405_P09 | Ensembl Location    |
| das_ABasPlasma   | http://db.systemsbiology.net:8080/das | HumanPlasma_ALL_Ena32_P09      | Ensembl Location    |

## File upload-based

```
#<col1>      <col2>      <col3>      <col4>
#<group>     <name>       <type>     <subtype>

Similarity   Fake_match_1  homology    wublastn
Transcription Fake_tscr_1   transcript   exon
Transcription Fake_tscr_1   transcript   exon

#<col5> <col6> <col7> <col8> <col9> <col10>
#<chr>  <start> <end>  <strand><phase> <score>

2       4000    4050   +       .       100
2       4200    4300   +       .       100
2       4400    4500   +       .       100
```

# Example: Epigenomic ChIP-chip data

e! Ensembl Mouse ContigView

Ensembl v85 - Apr 2006

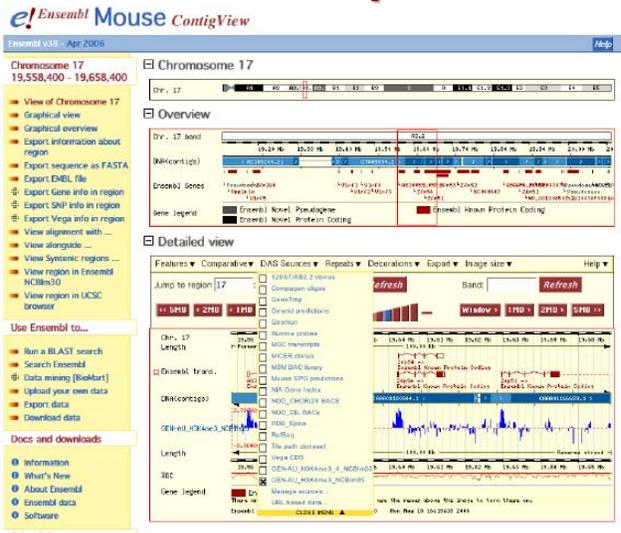
Chromosome 17  
19,558,400 - 19,658,400

View of Chromosome 17  
 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  
 View alignment with ...  
 View alongside ...  
 View Systemic regions ...  
 View regions in Ensembl NCBI-entrez  
 View region in UCSC browser

Use Ensembl to...  
 Run a BLAST search  
 Search Ensembl  
 Data mining (BioMart)  
 Upload your own data  
 Export data  
 Download data

Docs and downloads  
 Information  
 What's New  
 About Ensembl  
 Ensembl data  
 Software

Chromosome 17  
 Overview  
 Detailed view



# Example: Epigenomic ChIP-chip data

Detailed view

Features Comparative DAS Sources Repeats Decorations Export Image size Help

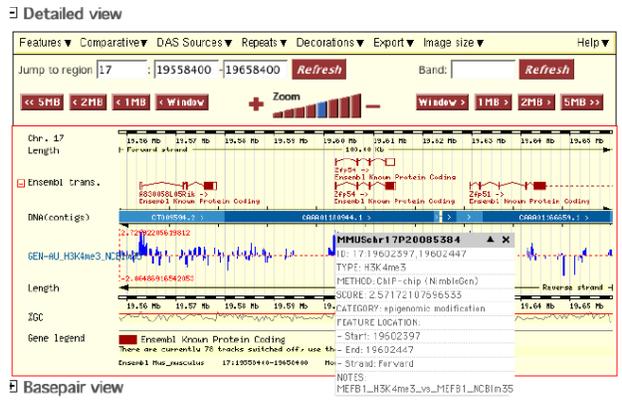
Jump to region 17 : 19558400 -19658400 Refresh Band: Refresh

5MB 2MB 1MB Window Zoom Window 1MB 2MB 5MB

Chr. 17  
 Length  
 Ensembl trans.  
 DNA(contigs)  
 GEN-HU\_H3K4me3\_NCBI  
 Length  
 ZIG  
 Gene legend

MMUSeqr17P20085384  
 ID: 1719602397,19602447  
 TYPE: H3K4me3  
 METHOD: ChIP-chip (NimbleGen)  
 SCORE: 2.51717210759535  
 CATEGORY: epigenomic modification  
 FEATURE LOCATION:  
 - Start: 19602397  
 - End: 19602447  
 - Strand: Forward

Basepair view  
 MEFB1\_H3K4me3\_vs\_H3F1\_NCBIin35





EMBL-EBI  

## DAS Registry

The Wellcome Trust  
Sanger Institute

available DAS services

Keyword search:

EMBL-EBI  
European Bioinformatics Institute

EBI Home About EBI Groups Services ToolBox Databases Downloads Submissions

UniProt DAS Services at the EBI

The distributed annotation system (DAS) is a client-server system in which a single client integrates information from multiple servers. The UniProt DAS server allows researchers to show their research results, for example identified peptides or signal sequences, on the UniProt reference sequence server, in the context of UniProt/Swiss-Prot annotation.

**The UniProt of Protein DAS Reference Server**

- Access the UniProt DAS Server

The UniProt DAS Reference Server serves both sequence and feature data from UniProt/Swiss-Prot and UniProt/TrEMBL. In addition to this, feature links to InterPro for all the proteins in UniProt are provided, including basic information (position, id and a brief description) and internet addresses to the relevant data source.

The server can be queried using any of the following:

- UniProt Accession numbers e.g. O35202 *sequence, features*
- Swiss-Prot ID's e.g. A4\_Human *sequence, features*
- TrEMBL ID's e.g. Q12368 *sequence, features*
- IPi ID's e.g. IPH0201571 *sequence*
- UniParc ID's e.g. UP000012665 *sequence*

All of these kinds of identifier will return sequence information. Identifier types marked with an \* will also return UniProt and InterPro feature hits.

For further details of the UniProt DAS server click on the link above that will take you to the UniProt DAS home page. This includes example requests using the DAS protocol for real proteins that can be found in UniProt.

<http://www.dasregistry.org>



EMBL-EBI  

## DAS Server Registration

Central DAS data source registration service at:  
<http://www.dasregistry.org/>

A registered data source

- Becomes publicly available
- can automatically be activated
  - in various DAS clients (Ensembl, SPICE, Dasty, etc.)
- gets “health”-checked

## DAS Server Implementations

Most commonly used DAS-Server implementations:

- **LDAS**  
The Lightweight DAS server using Perl, Apache and MySQL  
... see <http://www.biodas.org/servers/>
- **DAZZLE**  
DAS server written in Java  
... see <http://www.derkholm.net/thomas/dazzle>
- **ProServer**  
DAS server written in Perl  
... see <http://www.sanger.ac.uk/proserver>

## Ensembl as DAS server

- list of reference servers

<http://www.ensembl.org/das/dsn>

```
- <DSN>
  <SOURCE id="Homo_sapiens" version="37_35">Homo_sapiens</SOURCE>
  <MAPMASTER>http://www.ensembl.org/das/Homo_sapiens</MAPMASTER>
- <DESCRIPTION>
  Homo_sapiens Reference server based on NCBI35. Contains 111 entry points of the highest assembly tier.
</DESCRIPTION>
</DSN>
- <DSN>
  <SOURCE id="Macaca_mulatta" version="37_1a">Macaca_mulatta</SOURCE>
  <MAPMASTER>http://www.ensembl.org/das/Macaca_mulatta</MAPMASTER>
- <DESCRIPTION>
  Macaca_mulatta Reference server based on MMUL_0_1. Contains 192158 entry points of the highest assembly tier.
</DESCRIPTION>
</DSN>
- <DSN>
  <SOURCE id="Monodelphis_domestica" version="37_2a">Monodelphis_domestica</SOURCE>
  <MAPMASTER>http://www.ensembl.org/das/Monodelphis_domestica</MAPMASTER>
- <DESCRIPTION>
  Monodelphis_domestica Reference server based on BROADO2. Contains 5078 entry points of the highest assembly tier.
</DESCRIPTION>
</DSN>
```

# Ensembl as DAS server

- list of entry points

[http://www.ensembl.org/das/Homo\\_sapiens.NCBI36.reference/entry\\_points](http://www.ensembl.org/das/Homo_sapiens.NCBI36.reference/entry_points)

```

-<DASEP>
  -<ENTRY_POINTS href="Homo_sapiens.NCBI36.reference" version="1.0">
    <SEGMENT id="NT_113917" start="1" stop="19840" orientation="+">NT_113917</SEGMENT>
    <SEGMENT id="NT_113886" start="1" stop="96249" orientation="+">NT_113886</SEGMENT>
    <SEGMENT id="NT_113912" start="1" stop="185143" orientation="+">NT_113912</SEGMENT>
    <SEGMENT id="NT_113878" start="1" stop="106433" orientation="+">NT_113878</SEGMENT>
    <SEGMENT id="NT_113957" start="1" stop="166452" orientation="+">NT_113957</SEGMENT>
    <SEGMENT id="NT_113889" start="1" stop="161147" orientation="+">NT_113889</SEGMENT>
    <SEGMENT id="NT_113905" start="1" stop="183161" orientation="+">NT_113905</SEGMENT>
    <SEGMENT id="NT_113952" start="1" stop="184355" orientation="+">NT_113952</SEGMENT>
    <SEGMENT id="NT_113927" start="1" stop="111864" orientation="+">NT_113927</SEGMENT>
    <SEGMENT id="NT_113908" start="1" stop="13036" orientation="+">NT_113908</SEGMENT>
    <SEGMENT id="NT_113962" start="1" stop="217385" orientation="+">NT_113962</SEGMENT>
    <SEGMENT id="NT_113938" start="1" stop="44580" orientation="+">NT_113938</SEGMENT>
    <SEGMENT id="NT_113941" start="1" stop="37498" orientation="+">NT_113941</SEGMENT>
    <SEGMENT id="NT_113870" start="1" stop="145186" orientation="+">NT_113870</SEGMENT>
    <SEGMENT id="NT_113939" start="1" stop="147354" orientation="+">NT_113939</SEGMENT>
    <SEGMENT id="NT_113901" start="1" stop="182896" orientation="+">NT_113901</SEGMENT>
    <SEGMENT id="NT_113916" start="1" stop="173443" orientation="+">NT_113916</SEGMENT>
    <SEGMENT id="NT_113949" start="1" stop="159169" orientation="+">NT_113949</SEGMENT>
    <SEGMENT id="NT_113880" start="1" stop="185571" orientation="+">NT_113880</SEGMENT>
    <SEGMENT id="NT_113953" start="1" stop="131056" orientation="+">NT_113953</SEGMENT>
    <SEGMENT id="NT_113926" start="1" stop="119514" orientation="+">NT_113926</SEGMENT>
    <SEGMENT id="NT_113951" start="1" stop="152296" orientation="+">NT_113951</SEGMENT>
    <SEGMENT id="NT_113944" start="1" stop="182567" orientation="+">NT_113944</SEGMENT>
    <SEGMENT id="NT_113921" start="1" stop="39615" orientation="+">NT_113921</SEGMENT>
    <SEGMENT id="NT_113873" start="1" stop="51825" orientation="+">NT_113873</SEGMENT>
    <SEGMENT id="NT_113882" start="1" stop="172475" orientation="+">NT_113882</SEGMENT>
    <SEGMENT id="NT_113948" start="1" stop="92689" orientation="+">NT_113948</SEGMENT>
    <SEGMENT id="NT_113924" start="1" stop="139260" orientation="+">NT_113924</SEGMENT>
    <SEGMENT id="NT_113900" start="1" stop="112804" orientation="+">NT_113900</SEGMENT>
    <SEGMENT id="NT_113943" start="1" stop="81310" orientation="+">NT_113943</SEGMENT>
  
```

# Ensembl as DAS server

- Exploring assembly:

<http://www.ensembl.org/das/.../features?segment=1:1,1000000>

```

-<DASGFF>
  -<GFF>
    -<SEGMENT id="1" start="1" stop="1000000">
      -<FEATURE id="1">
        <START>1</START>
        <STOP>1000000</STOP>
        <TYPE id="chromosome" category="component" reference="yes" superparts="no" subparts="yes">chromosome</TYPE>
        <TARGET id="1" start="1" stop="1000000">1</TARGET>
      </FEATURE>
      -<FEATURE id="supercontig:NT_077912">
        <START>357583</START>
        <STOP>511231</STOP>
        <TYPE id="supercontig" category="component" reference="yes" superparts="yes" subparts="yes">supercontig</TYPE>
        <TARGET id="NT_077912" start="1" stop="153649">supercontig NT_077912</TARGET>
      </FEATURE>
      -<FEATURE id="supercontig:NT_077913">
        <START>561232</START>
        <STOP>1000000</STOP>
        <TYPE id="supercontig" category="component" reference="yes" superparts="yes" subparts="yes">supercontig</TYPE>
        <TARGET id="NT_077913" start="1" stop="438769">supercontig NT_077913</TARGET>
      </FEATURE>
      -<FEATURE id="supercontig:NT_077402">
        <START>1</START>
        <STOP>167280</STOP>
      </FEATURE>
    
```

## Ensembl as DAS server

### • Exploring assembly:

[http://www.ensembl.org/das/.../features?segment=NT\\_077912](http://www.ensembl.org/das/.../features?segment=NT_077912)

```
-<DASGFF>
-<GFF>
  -<SEGMENT id="NT_077912" start="" stop="">
    -<FEATURE id="NT_077912">
      <START/>
      <STOP/>
      <TYPE id="supercontig" category="component" reference="yes" superparts="yes" subparts="yes">supercontig</TYPE>
      <TARGET id="NT_077912" start="" stop="">NT_077912</TARGET>
    </FEATURE>
    -<FEATURE id="clone:AL732372.15">
      <START>1</START>
      <STOP>153649</STOP>
      <TYPE id="clone" category="component" reference="yes" superparts="yes" subparts="yes">clone</TYPE>
      <TARGET id="AL732372.15" start="1" stop="153649">clone AL732372.15</TARGET>
    </FEATURE>
    -<FEATURE id="chromosome:1">
      <START>1</START>
      <STOP>153649</STOP>
      <TYPE id="chromosome" category="supercomponent" reference="yes" superparts="no" subparts="yes"/>
      <TARGET id="1" start="357583" stop="511231"/>
    </FEATURE>
  </SEGMENT>
</GFF>
</DASGFF>
```

## Ensembl as DAS server

### • Exploring assembly:

<http://www.ensembl.org/das/.../features?segment=AL732372.15.1.153649>

```
-<DASGFF>
-<GFF>
  -<SEGMENT id="AL732372.15.1.153649" start="" stop="">
    -<FEATURE id="AL732372.15.1.153649">
      <START/>
      <STOP/>
      <TYPE id="contig" category="component" reference="yes" superparts="yes" subparts="no">contig</TYPE>
      <TARGET id="AL732372.15.1.153649" start="" stop="">AL732372.15.1.153649</TARGET>
    </FEATURE>
    -<FEATURE id="clone:AL732372.15">
      <START>1</START>
      <STOP>153649</STOP>
      <TYPE id="clone" category="supercomponent" reference="yes" superparts="yes" subparts="yes"/>
      <TARGET id="AL732372.15" start="1" stop="153649"/>
    </FEATURE>
  </SEGMENT>
</GFF>
</DASGFF>
```