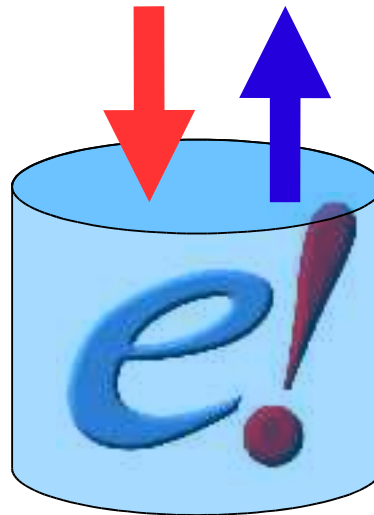


# Ensembl and DAS



Eugene Kulesha  
Wellcome Trust Sanger Institute  
April 2006



# DAS and Ensembl

- DAS Essentials
- Ensembl as DAS server
- Adding DAS sources to Ensembl views
- DAS configuration in Ensembl
- DAS code in Ensembl
- Future development

Ensembl



# DAS Essentials

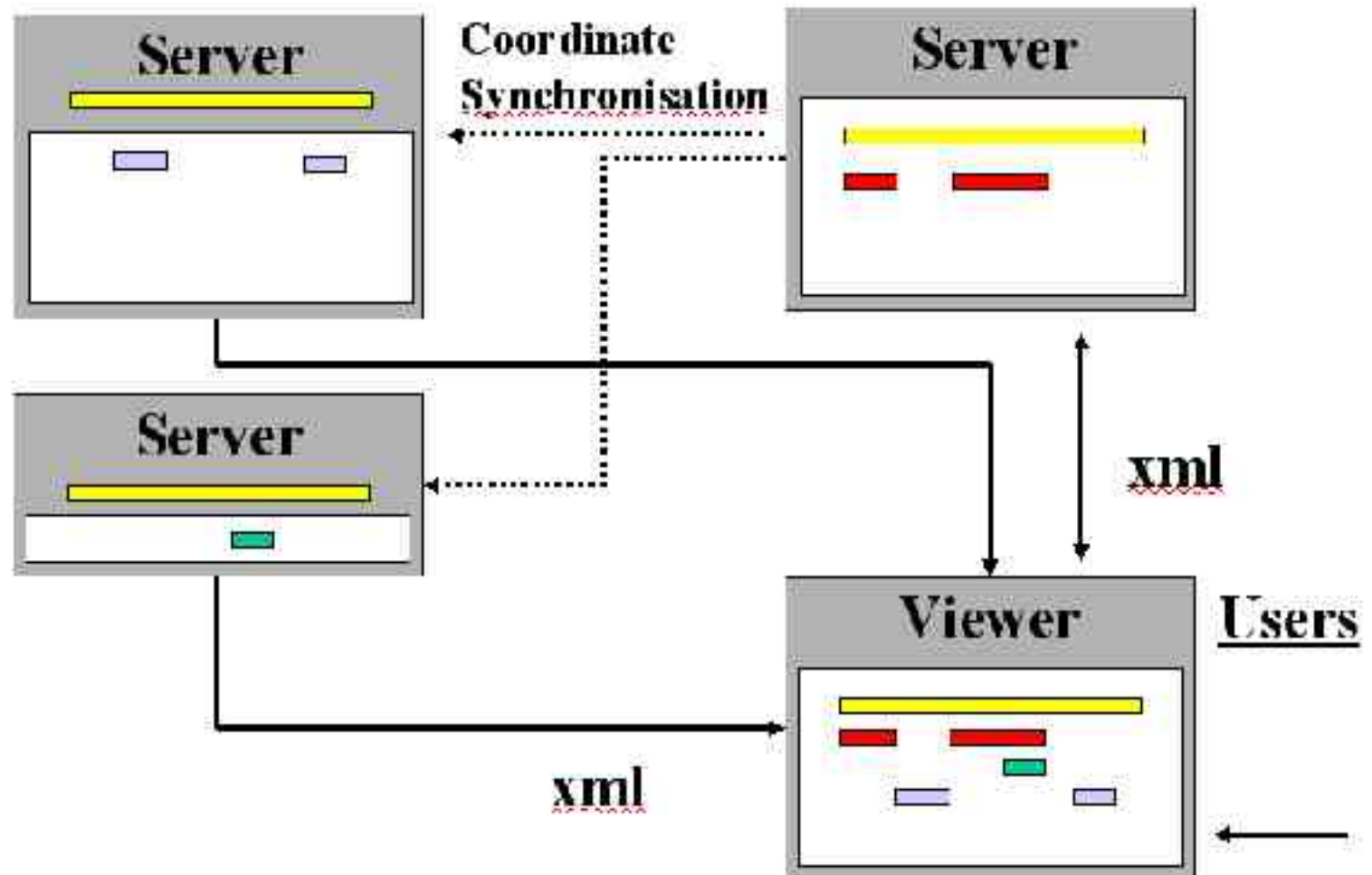
- Distributed Annotation System (DAS);
  - A server system for the sharing of Reference Sequences, and their annotations.
- DAS servers:
  - Reference server
  - Annotation server
- DAS requests:
  - dns
  - entry\_points
  - dna
  - features
- Ensembl as DAS server:

<http://www.biodas.org/documents/spec.html>

# Distributed Annotation System

## External Contributors

## Database providers



# DAS Essentials

``dsn`` - Get the list of available sources

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

[illegible]

# DAS Essentials

`features` - Get the annotations

[http://www.ebi.ac.uk/das-srv/asd/das/atd\\_human/features?segment=13:3770000,38100000](http://www.ebi.ac.uk/das-srv/asd/das/atd_human/features?segment=13:3770000,38100000)

```
<?xml version='1.0' standalone='yes'?>
<!DOCTYPE DASSET SYSTEM "http://www.biodas.org/dtd/dasgff.dtd">
<DASSET>
  <GFF version='1.0' href='http://www59-model.ebi.ac.uk:9100/das/asd_human/features'?>
    <SEGMENT id='13' version='1.0' start='3770000' stop='38100000'>
      <FEATURE id='5438:1.215740' label='6438:1.215740'>
        <TYPE id='exon' reference='no' subparts='no' superparts='no'>exon</TYPE>
        <START>3265804</START>
        <END>3258215</END>
        <ORIENTATION>--</ORIENTATION>
        <LINK href='http://www.ebi.ac.uk/asc/srv/atd.cgi?ret=cd-TRANSCRIPT;origin=AT;product=
        <GROUP id='5438' label='ENSG00000133111 AT-1' />
      </FEATURE>
      <FEATURE id='5438:1.215741' label='6438:1.215741'>
        <TYPE id='exon' reference='no' subparts='no' superparts='no'>exon</TYPE>
        <START>3265887</START>
        <END>32539780</END>
        <ORIENTATION>--</ORIENTATION>
        <LINK href='http://www.ebi.ac.uk/asc/srv/atd.cgi?ret=cd-TRANSCRIPT;origin=AT;product=
        <GROUP id='5438' label='ENSG00000133111 AT-1' />
      </FEATURE>
      <FEATURE id='5438:1.215742' label='6438:1.215742'>
        <TYPE id='exon' reference='no' subparts='no' superparts='no'>exon</TYPE>
        <START>32659475</START>
        <END>32537514</END>
        <ORIENTATION>--</ORIENTATION>
        <LINK href='http://www.ebi.ac.uk/asc/srv/atd.cgi?ret=cd-TRANSCRIPT;origin=AT;product=
        <GROUP id='5438' label='ENSG00000133111 AT-1' />
      </FEATURE>
```

# Ensembl as DAS server

- list of reference sources

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

```
- <DSN>
  <SOURCE id= 'Homo_sapiens' version= 37_35j">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.current.reference/entry\\_points](http://www.ensembl.org/das/Homo_sapiens.current.reference/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="http://www.ensembl.org/das" version="1.0">

<SEGMENT id="1" start="1" stop="245522847" orientation="+">1</SEGMENT>

<SEGMENT id="10" start="1" stop="135113628" orientation="1">10</SEGMENT>

<SEGMENT id="10\_NT\_079544" start="1" stop="115275" orientation="-">10\_NT\_079544</SEGMENT>

<SEGMENT id="11" start="1" stop="134152384" orientation="1">11</SEGMENT>

<SEGMENT id="12" start="1" stop="132449811" orientation="+">12</SEGMENT>

<SEGMENT id="12\_NT\_091701" start="1" stop="66818" orientation=" ">12\_NT\_091701</SEGMENT>

<SEGMENT id="13" start="1" stop="114142980" orientation="+">13</SEGMENT>

<SEGMENT id="13\_NT\_078092" start="1" stop="186858" orientation=" ">13\_NT\_078092</SEGMENT>

<SEGMENT id="14" start="1" stop="106368585" orientation="+">14</SEGMENT>

<SEGMENT id="15" start="1" stop="100338915" orientation="1">15</SEGMENT>

<SEGMENT id="15\_NT\_079546" start="1" stop="139280" orientation="-">15\_NT\_079546</SEGMENT>

<SEGMENT id="15\_NT\_079547" start="1" stop="168820" orientation=" ">15\_NT\_079547</SEGMENT>

<SEGMENT id="15\_NT\_079548" start="1" stop="119514" orientation="-">15\_NT\_079548</SEGMENT>

<SEGMENT id="15\_NT\_079549" start="1" stop="111861" orientation=" ">15\_NT\_079549</SEGMENT>

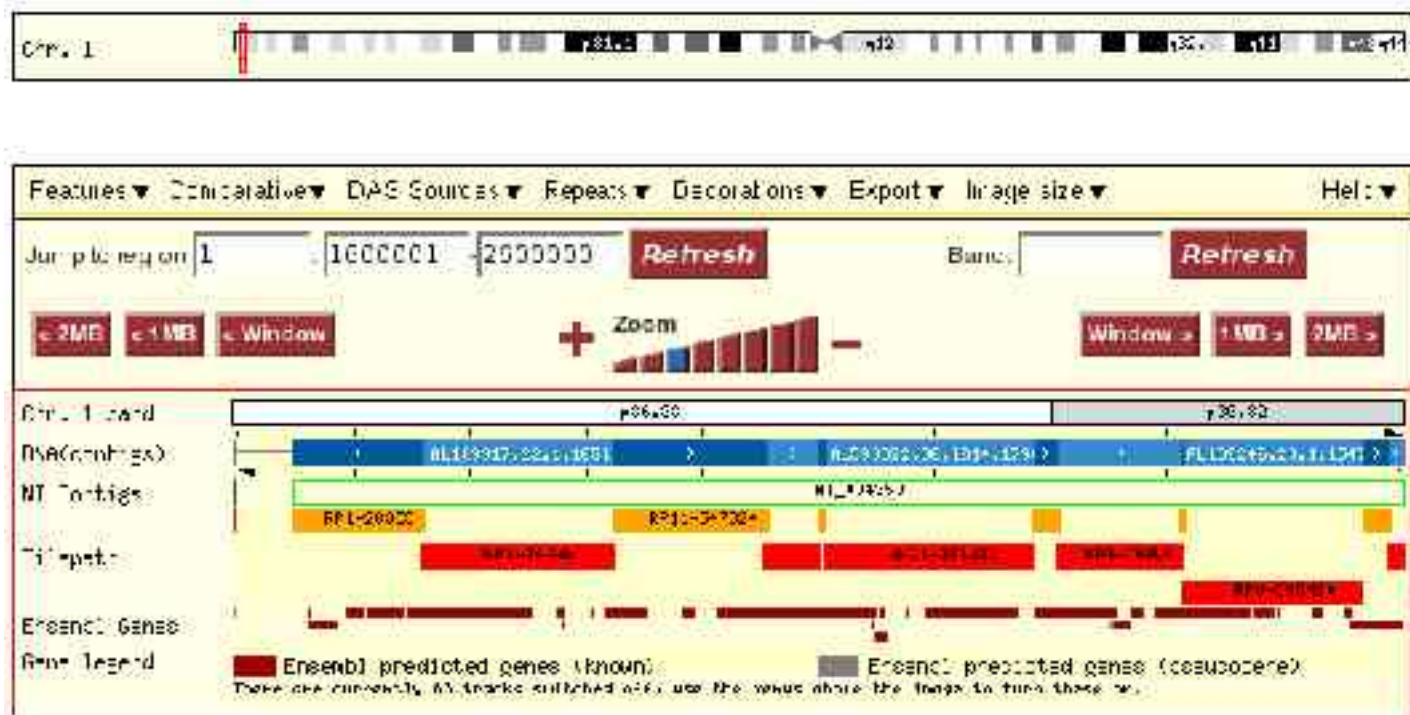
<SEGMENT id="15\_NT\_079553" start="1" stop="44888" orientation="-">15\_NT\_079553</SEGMENT>

<SEGMENT id="16" start="1" stop="88827254" orientation="1">16</SEGMENT>



# Ensembl as DAS server

- Exploring assembly



# Ensembl as DAS server

- Exploring assembly:

[http://www.ensembl.org/das/Homo\\_sapiens.current.reference/features?segment=1:1,1000000](http://www.ensembl.org/das/Homo_sapiens.current.reference/features?segment=1:1,1000000)

```
<DASGET>
  <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>557583</START>
        <STOP>511231</STOP>
        <TYPE id="supercontig" category="component" reference="yes" superparts="yes" subparts="yes">supercontig</TYPE>
        <TARGET id="NT_077912" start="1" stop="133619">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="435764">supercontig:NT_077913</TARGET>
      </FEATURE>
      - <FEATURE id="supercontig:NT_077402">
        <START>1</START>
        <STOP>167280</STOP>
```

# Ensembl as DAS server

- Exploring assembly:

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

```

-<DASGIT>
-<GET>
  <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="357383" stop="511231"/>
    </FEATURE>
  </SEGMENT>
</GET>
</DASGIT>

```

# Ensembl as DAS server

- Exploring assembly:  
[http://www.ensembl.org/das/Homo\\_sapiens.current.reference/features?segment=AL732372.15.1.153649](http://www.ensembl.org/das/Homo_sapiens.current.reference/features?segment=AL732372.15.1.153649)

```
- <DASGET>
- <GET>
- <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>
  <GET>
</DASGET>
```



# Ensembl as DAS server

- Exploring assembly:

[http://www.ensembl.org/das/Homo\\_sapiens.current.karyotype/features?segment=22](http://www.ensembl.org/das/Homo_sapiens.current.karyotype/features?segment=22)

```
- <SEGMENT id="22" start="1" stop="49591432">
  - <FEATURE id='p11.1'>
    <START>9607114</START>
    <STOP>11808744</STOP>
    <TYPE id='acen'>acen</TYPE>
    <METHOD id='ensembl'>ensembl</METHOD>
    <ORIENTATION/>
  </FEATURE>
  - <FEATURE id='p11.2'>
    <START>6604891</START>
    <STOP>9607113</STOP>
    <TYPE id='gvar'>gvar</TYPE>
    <METHOD id='ensembl'>ensembl</METHOD>
    <ORIENTATION/>
  </FEATURE>
  - <FEATURE id='p12'>
    <START>3002224</START>
    <STOP>6604890</STOP>
    <TYPE id='stalk'>stalk</TYPE>
    <METHOD id='ensembl'>ensembl</METHOD>
    <ORIENTATION/>
  </FEATURE>
```



Ensembl

## Adding DAS source to Ensembl





# Adding DAS source to Ensembl

- Switch on pre-configured server
- Set up your own server  
(ProServer, Dazzle, LDAS – see Ensembl docs)
- Configure an existing server
- Upload your own data
- View your data
- Send your source to a colleague
- Web-based sources (Contigview & Cytoview only)

# Switch on a pre-configured source in geneview, protview, transview

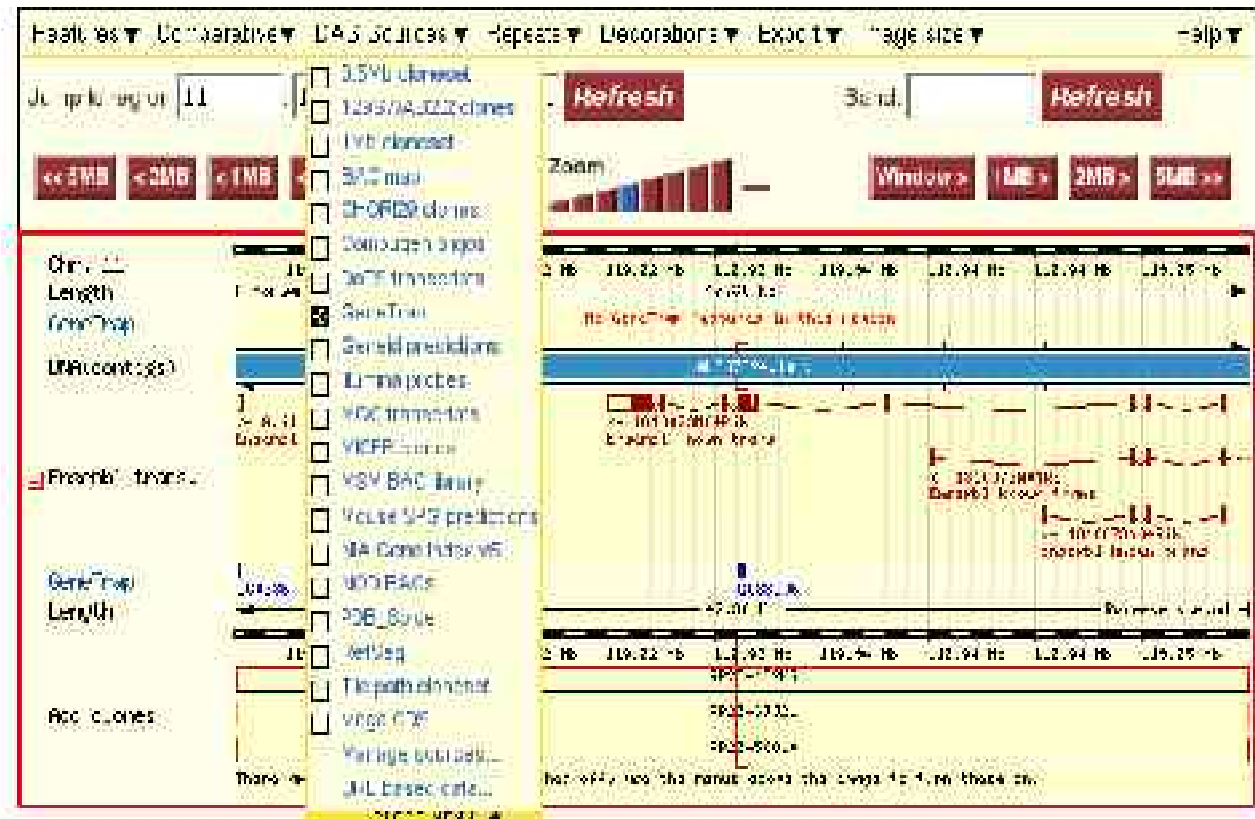
## Gene DAS Report

### DAS Sources

- ☐ [AltAtlas](#) (Alternative splice databases)
- ☐ [ALTtrans](#) (Alternative Transcript Diversity Database)
- ☐ [ArrayExpress](#) (Gene Expression Database)
- ☐ [GAD](#) (Genetic Association Database)
- ☐ [HGNC](#) (HUGO Gene Nomenclature Committee)
- ☐ [HUGO\\_text](#) (PubMed text-mining site HUGOsymbio)
- ☐ Phenotypes (Associated directly or via orthologues or protein families)
- ☐ [Protonet](#) (Global classification of proteins into hierarchical clusters)
- ☐ [ResFinder](#) (Knowledgebase of biological processes)
- ☐ [UniProt](#) (Protein knowledge base)

Manage SQL DBs

### E Detailed view



# Set up your own DAS source

- ProServer
  - ♦ <http://www.sanger.ac.uk/Software/analysis/proserver/>
- Dazzle
  - ♦ <http://www.biojava.org/dazzle/>
- LDAS
  - ♦ <http://biodas.org/servers/LDAS.html>

and register with DAS Registry !

<http://das.sanger.ac.uk/registry>

# DAS Registry

<http://www.ebi.ac.uk/das-srv/uniprot/das/erlistotlo/>

unique id: DS\_109

nickname: uniprot

display in DAS clients  
 Uniprot  
 EMBL  
 DAS

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 following kinds of protein ID / Accession number can be used: UniProt Accession numbers e.g. Q36502 Swiss-Prot IDs e.g. A/Human TrEMBL IDs e.g. Q12368 IPI IDs e.g. IPI00016171 UniProt IDs e.g. LP00001250561 for a full description of the service, visit: <http://www.ebi.ac.uk/uniprot-das/>

admin email: uniprot-das\_AT\_ebi.ac.uk

homepage: <http://www.ebi.ac.uk/uniprot-das/>

DAS capabilities: [sequence](#) [features](#) [entry](#) [points](#) [stylesheet](#)

Test Cases: PROPRC

Coordinate Systems: UniProt/Protein Sequence

Registered at: 2005.03.21

Last tested successfully at: 2006.03.17

label(s): UNISLMLL EriSapens



[send this das source to a friend](#)



[validate server](#)



[edit this server. \(das source admin only\)](#)



[remove this server. \(das source admin only\)](#)


# Configure an existing source

**e!Ensembl Human DatacrafView** Classic Interface




















Ensembl v57 Feb 2008 Help

**Manage Sources**

- Add Data Source
- Refresh your data

**DAS sources**

Name	DAS Server	Data Source	Coordinate System
 ARS data	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	ars_gene	Ensembl Gene ID
 ARS data	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	ars_gene	Ensembl Gene ID
 Arago Ensembl	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	arago_gene	Ensembl Gene ID
 ASD	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	asd	HUGO ID
 HUGO	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	HUGO	Ensembl Gene ID
 HUGO_L1	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l1	HUGO ID
 HUGO_L2	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l2	Ensembl Gene ID
 HUGO_L3	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l3	Ensembl Gene ID
 HUGO_L4	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l4	Ensembl Gene ID
 HUGO_L5	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l5	Ensembl Gene ID
 HUGO_L6	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l6	Ensembl Gene ID
 HUGO_L7	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l7	Ensembl Gene ID
 HUGO_L8	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l8	Ensembl Gene ID
 HUGO_L9	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l9	Ensembl Gene ID
 HUGO_L10	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l10	Ensembl Gene ID
 HUGO_L11	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l11	Ensembl Gene ID
 HUGO_L12	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l12	Ensembl Gene ID
 HUGO_L13	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l13	Ensembl Gene ID
 HUGO_L14	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l14	Ensembl Gene ID
 HUGO_L15	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l15	Ensembl Gene ID
 HUGO_L16	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l16	Ensembl Gene ID
 HUGO_L17	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l17	Ensembl Gene ID
 HUGO_L18	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l18	Ensembl Gene ID
 HUGO_L19	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l19	Ensembl Gene ID
 HUGO_L20	<a href="http://www.ncbi.nlm.nih.gov/arsdata/">http://www.ncbi.nlm.nih.gov/arsdata/</a>	hugo_l20	Ensembl Gene ID

© 2008 EBI. All rights reserved. [Download the Ensembl genome browser](#) [Download the Ensembl genome browser](#) [Download the Ensembl genome browser](#)



# Configure an existing source

## DAS Wizard Step 1 of 3: Data location

Please select DAS sources to which you wish to Next

**Next**

(for user uploaded sources enter DAS here)

DAS Server

(use the filter to narrow the list of sources)

☐ 3D-Genomics

<http://www.3dgenomics.org/>  
The 3D-Genomics server  
[details about 3D-Genomics](#)

☐ CSA - catalytic

<http://www.ebi.ac.uk/da/>  
Catalytic Site Atlas (CSA)  
[details about CSA - catalytic](#)

☐ CSA - literature

<http://www.ebi.ac.uk/da/>  
Catalytic Site Atlas (CSA)  
[details about CSA - literature](#)

☐ IMG - Genes

<http://img.jgi.doe.gov/IMG/Genes/>  
We provide a mirror of all the IMG/Genes data  
[details about IMG - Genes](#)

☐ MafPro

<http://mafpro.sanger.ac.uk/mafpro/>  
MafPro provides a list of mispredicted genes/proteins  
[details about MafPro](#)

DAS Registry  
DAS Registry  
<http://das.ensembl.org/das>  
<http://das.sanger.ac.uk/das>  
<http://das.sanger.ac.uk/0/das/das>  
<http://db.gensci.sanger.ac.uk/0/das>  
<http://genetics.mim.ac.uk/0/das>  
<http://img.jgi.doe.gov/0/das>  
<http://land.cmpg.sanger.ac.uk/0/das>  
<http://mafpro.sanger.ac.uk/0/das>  
<http://www.ebi.ac.uk/da/0/das>  
<http://www.genomics.sanger.ac.uk/0/das>  
<http://mafpro.sanger.ac.uk/0/das>

Next

1 of 1

# Configure an existing source

## DAS Wizard Step 2 of 3: Data appearance

**Coordinate System**

- ☒ Ensembl Gene ID
- ☐ Ensembl Location
- ☐ Chromosome
- ☐ Contig
- ☐ Contig
- ☐ Contig
- ☐ NT Contig
- ☐ Ensembl Peptide ID
- ☐ Ensembl Transcript ID
- ☐ Entrez Gene ID
- ☐ HUGO ID
- ☐ IPI Accession
- ☐ IPI ID
- ☐ UniprotSwiss-Prot Name
- ☐ UniprotSwiss-Prot Acc

**Enable on**

- ☒ genview
- ☐ trackview
- ☐ cytoview
- ☐ proteinview
- ☐ configview

**Back** **Next**

# Configure an existing source

## DAS Wizard Step 3 of 3: Display configuration

Name:	<input type="text" value="alignpubt"/>
Track label:	<input type="text" value="alignpubt"/>
Track URL:	<input type="text"/>
Link track:	<input type="text"/>
Link URL:	<input type="text"/>
Track colour:	<input type="text" value="Blue"/>
Group features:	<input type="text" value="No"/>
Display on:	<input type="text" value="Both strands"/>
Max rows to display:	<input type="text" value="20 rows"/>
Label features:	<input type="text" value="Inner feature"/>
Apply stylesheet:	<input type="text" value="No"/>
Use scores:	<input type="text" value="No"/>
<input type="button" value="Back"/> <input type="button" value="Finish"/>	

# Configure an existing source

## DAS sources

	Name	DAS Server	Data Source	Coordinate System
	AllSplice	<a href="http://www.ensembl.org/ensdb/was/das">http://www.ensembl.org/ensdb/was/das</a>	all_splice	Ensembl Gene ID
	AllTrans	<a href="http://www.ensembl.org/ensdb/was/das">http://www.ensembl.org/ensdb/was/das</a>	all_gene	Ensembl Gene ID
	Arise/Exon+	<a href="http://www.ensembl.org/ensdb/exonplus/das">http://www.ensembl.org/ensdb/exonplus/das</a>	arise/exon+	Uniprot/Swiss-Prot Acc
	gaf	<a href="http://www.ensembl.org/ensdb/wgencode/das">http://www.ensembl.org/ensdb/wgencode/das</a>	gaf	HUGO ID
	HOMO	<a href="http://hgdownload.soe.ucsc.edu/hg000/das">http://hgdownload.soe.ucsc.edu/hg000/das</a>	HOMO	Ensembl Gene ID
	HUGO_tax	<a href="http://www.ensembl.org/ensdb/wgencode/das">http://www.ensembl.org/ensdb/wgencode/das</a>	tax_naming	HUGO ID
	Phenotype	<a href="http://www.ensembl.org/ensdb/wgencode/das">http://www.ensembl.org/ensdb/wgencode/das</a>	phenotypes	Ensembl Gene ID
	Protein	<a href="http://www.protein.org.uk/das">http://www.protein.org.uk/das</a>	protein	Uniprot/Swiss-Prot Acc
	Proteome	<a href="http://www.proteome.org.uk/das">http://www.proteome.org.uk/das</a>	GH_human	Uniprot/Swiss-Prot Acc
	UniProt	<a href="http://www.uniprot.org/das">http://www.uniprot.org/das</a>	uniprot	Uniprot/Swiss-Prot Acc
	Hydraeul_0000128	<a href="http://www.ensembl.org/das">http://www.ensembl.org/das</a>	hydraeul_0000128	Ensembl Local ID
	FTDIL-DAS	<a href="http://www.ensembl.org/das">http://www.ensembl.org/das</a>	FTDIL	Uniprot/Swiss-Prot Acc
	Ensembl	<a href="http://www.ensembl.org/das">http://www.ensembl.org/das</a>	ensembl	Uniprot/Swiss-Prot Acc



# Upload your data

- File format:
  - annotation section

<http://www.sanger.ac.uk/Software/formats/GFF/>

- stylesheet section

[http://www.ensembl.org/info/data/external\\_data/das/CSS\\_support.pdf](http://www.ensembl.org/info/data/external_data/das/CSS_support.pdf)

- Examples:

[http://www.ensembl.org/info/data/external\\_data/das/das\\_upload\\_v1.euf](http://www.ensembl.org/info/data/external_data/das/das_upload_v1.euf)

Ensembl



# Upload your data

## DAS Wizard Step 1 of 3: Data location

Please provide your data location

Please [READ THE UPLOAD INSTRUCTIONS CAREFULLY](#) before uploading any data. Your data must be [formatted correctly](#) before uploading will work properly. The instructions page has detailed information about the data format.

Please read and understand the [Ensembl policy on uploaded data](#).

Email

Password   
your email and password to ensure that nobody else can modify your annotation

Paste your data

or choose a file to upload

Upload File:

If you want to update an existing annotation on Ensembl DAS Server enter its Data Source Name and select your action

Data source

Action

☐ Overwrite

☒ append

**Next**

Fields marked with \* are required

Ensembl



# Upload your data

## DAS Wizard Step 2 of 3: Data appearance

Successfully uploaded stylesheet  
Successfully uploaded 11 entries  
A new DAS source has been created at [http://www.ensembl.org/das/hydration/\\_00001000](http://www.ensembl.org/das/hydration/_00001000)

**Coordinate System** ☐ Ensembl location

**Enable on**

☒ geneview  
☐ transview  
☐ cytobview

☐ proteinview  
☐ contigview

**Back**

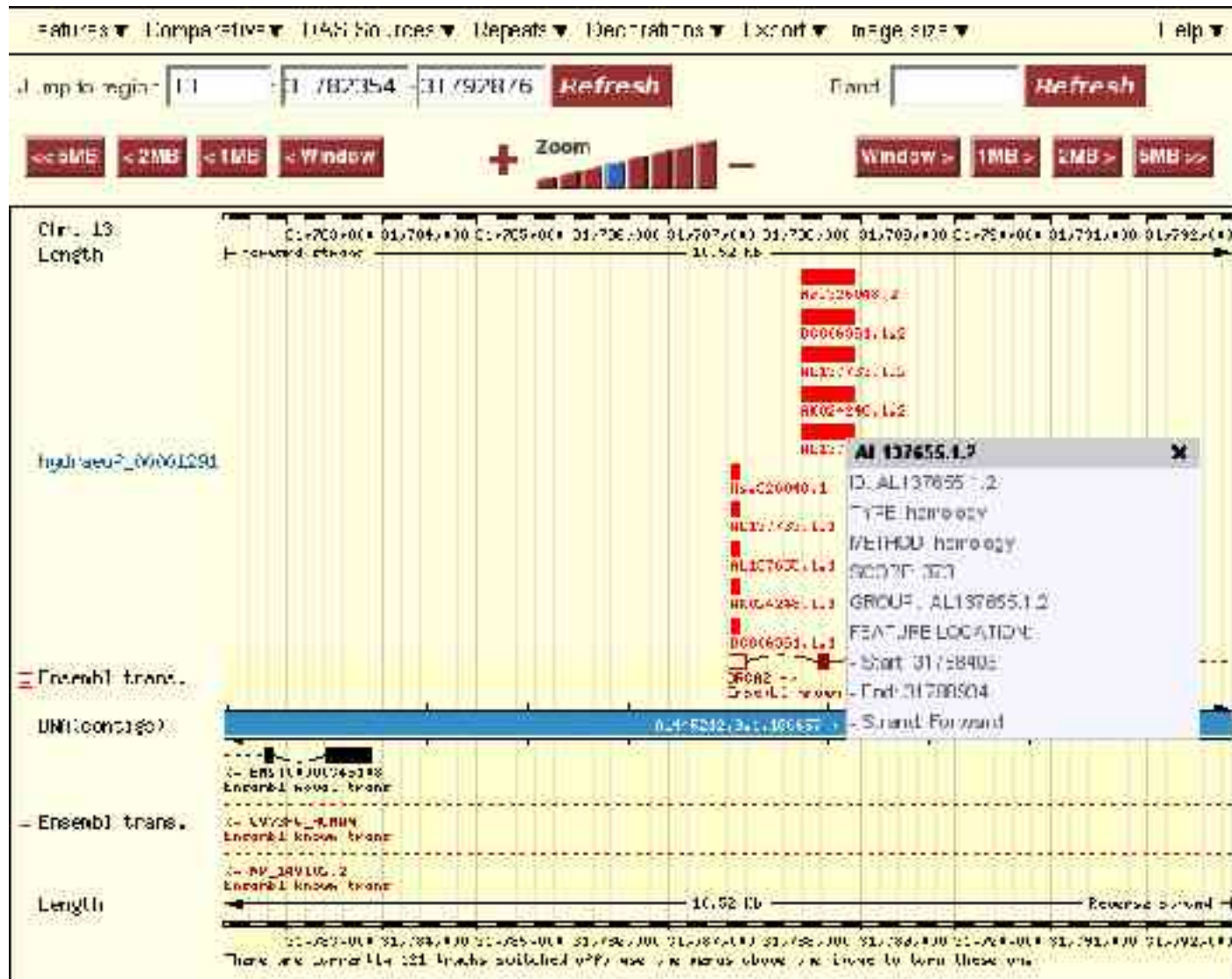
**Next**

# View your data in geneview

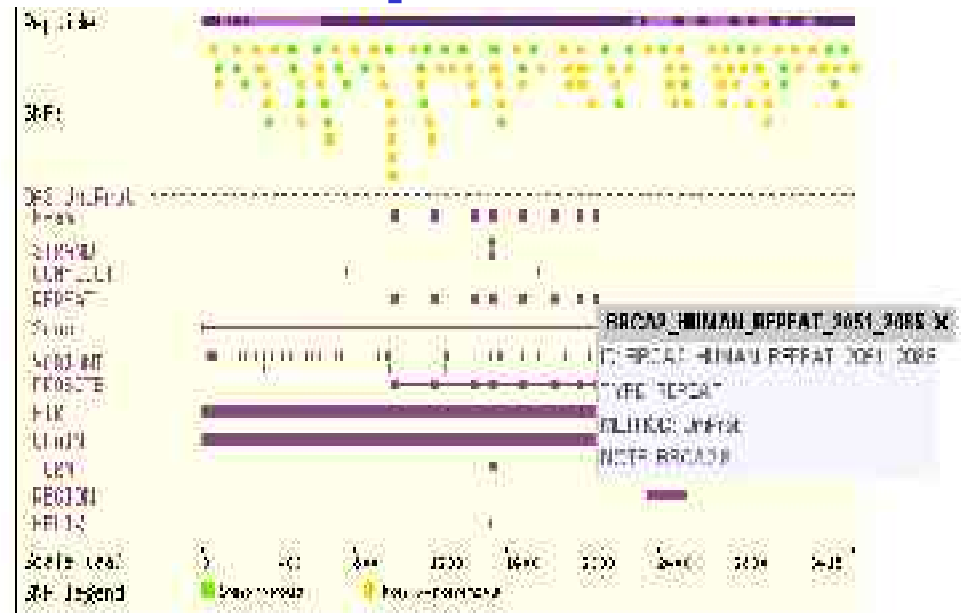
## Gene DAS Report

hydraeuf_00001291	homology AK024248.1.1	homology	Feature contained by gene	373.00
	homology AK024248.1.2	homology	Feature contained by gene	373.00
	homology AL137855.1.1	homology	Feature contained by gene	373.00
	homology AL137855.1.2	homology	Feature contained by gene	373.00
	homology AL137733.1.1	homology	Feature contained by gene	354.00
	homology AL137733.1.2	homology	Feature contained by gene	354.00
	homology BC006361.1.1	homology	Feature contained by gene	354.00
	homology BC006361.1.2	homology	Feature contained by gene	354.00
	homology Hs.326048.1	homology	Feature contained by gene	373.00
	homology Hs.326048.2	homology	Feature contained by gene	373.00
<div><div></div>DAS Sources</div>	<div><div></div><a href="#">AltSplice</a> (Alternative splice database)</div>			
	<div><div></div><a href="#">AltTrans</a> (Alternative Transcript Diversity Database)</div>			
	<div><div></div><a href="#">ArrayExpress</a> (Gene Expression Database)</div>			
	<div><div></div><a href="#">GAD</a> (Genetic Association Database)</div>			
	<div><div></div><a href="#">HGNC</a> (HUGO Gene Nomenclature Committee)</div>			
	<div><div></div><a href="#">HUGO_text</a> (PubMed text-mining via HUGO symbol)</div>			
	<div><div></div>PRIDE-DAS</div>			
	<div><div></div>Phenotypes (Associated directly or via orthologues or protein families)</div>			
	<div><div></div><a href="#">Protonet</a> (Global classification of proteins into hierarchical clusters)</div>			
	<div><div></div><a href="#">Reactome</a> (Knowledgebase of biological processes)</div>			
	<div><div></div><a href="#">UniProt</a> (Protein knowledgebase)</div>			
	<div><div><div></div></div>hydraeuf_00001291</div>			
<div>Manage Sources</div>				

## View your data in contigview



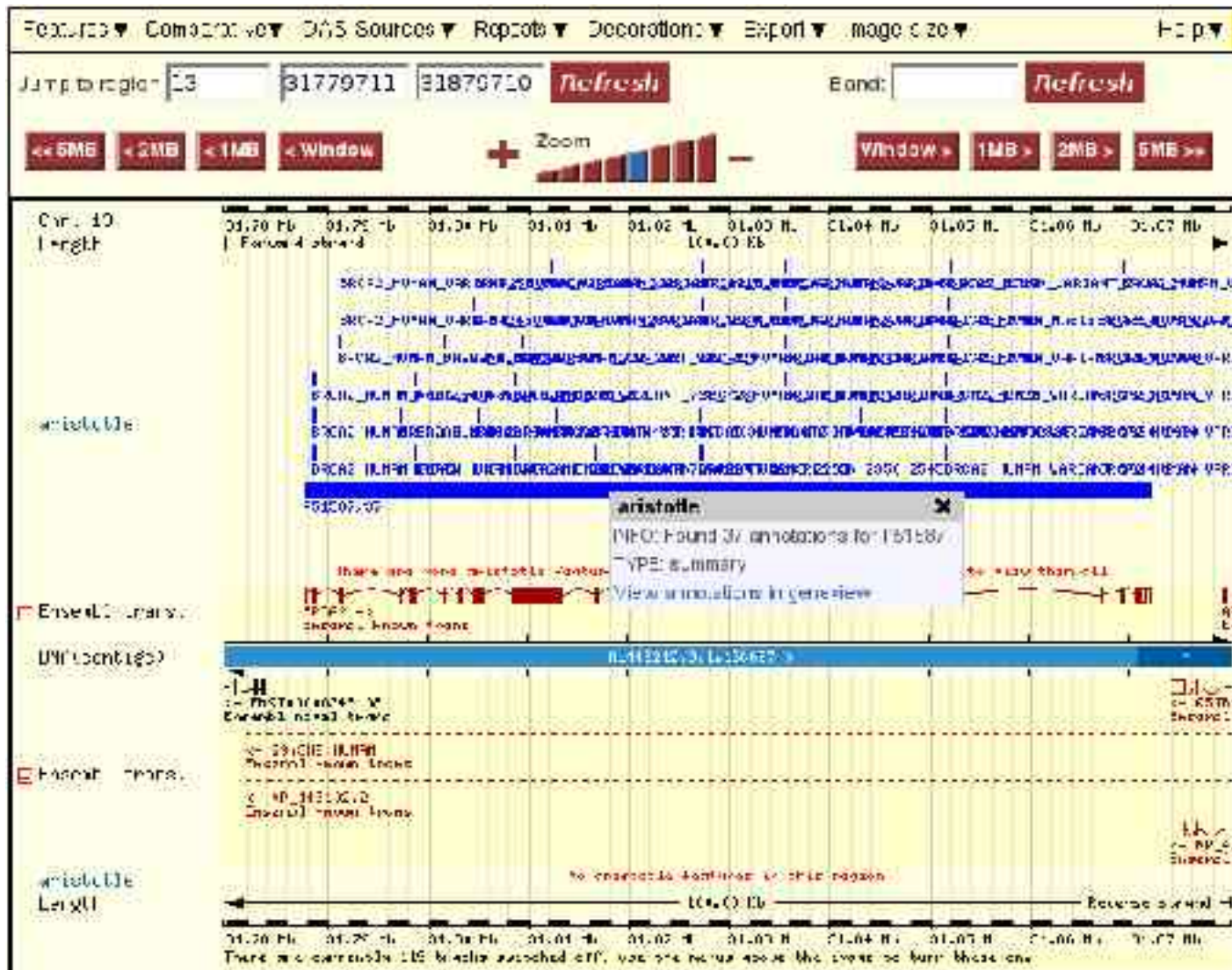
## Proven Features

[illegible]



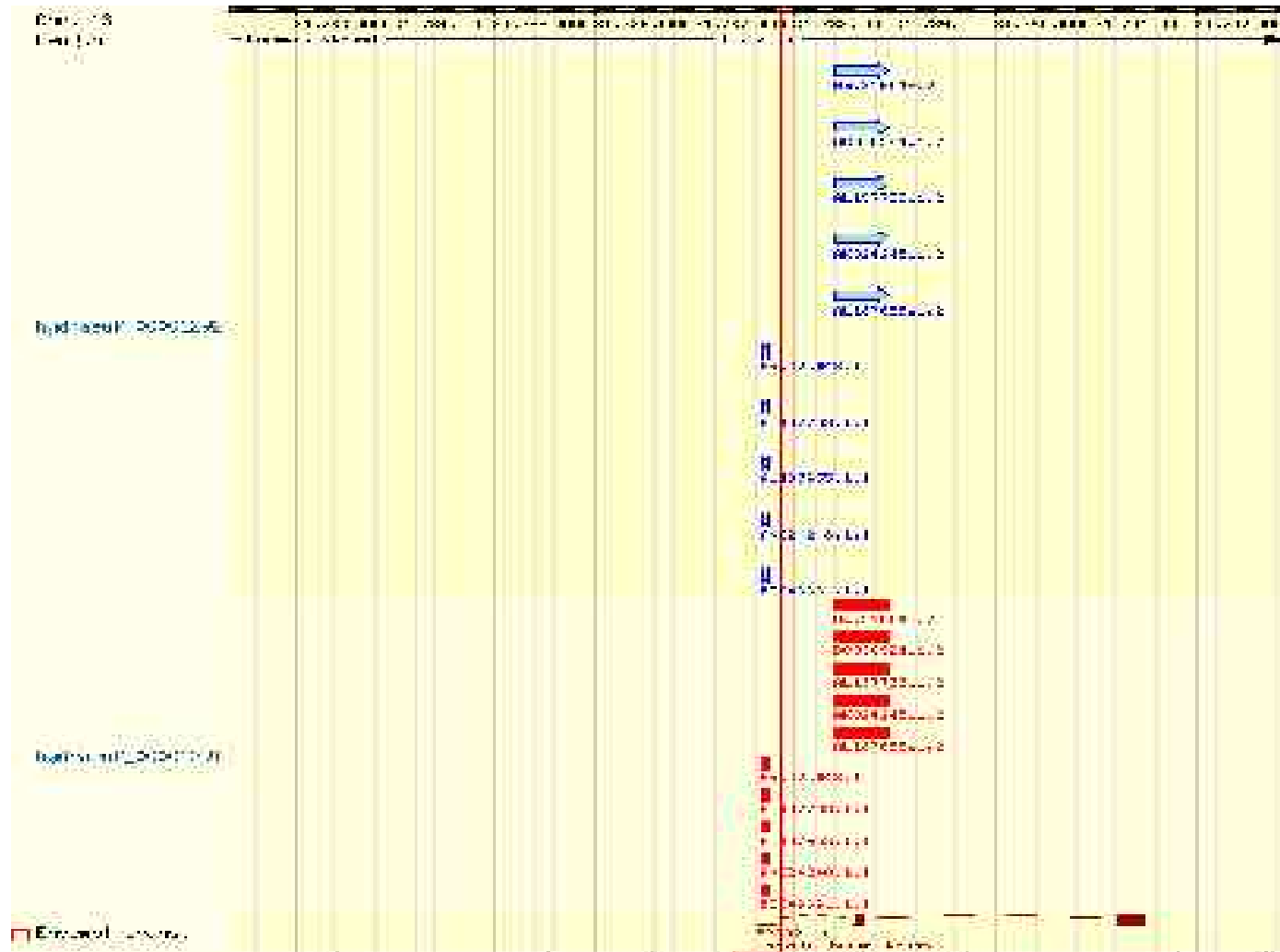


# Gene-based annotations in contigview



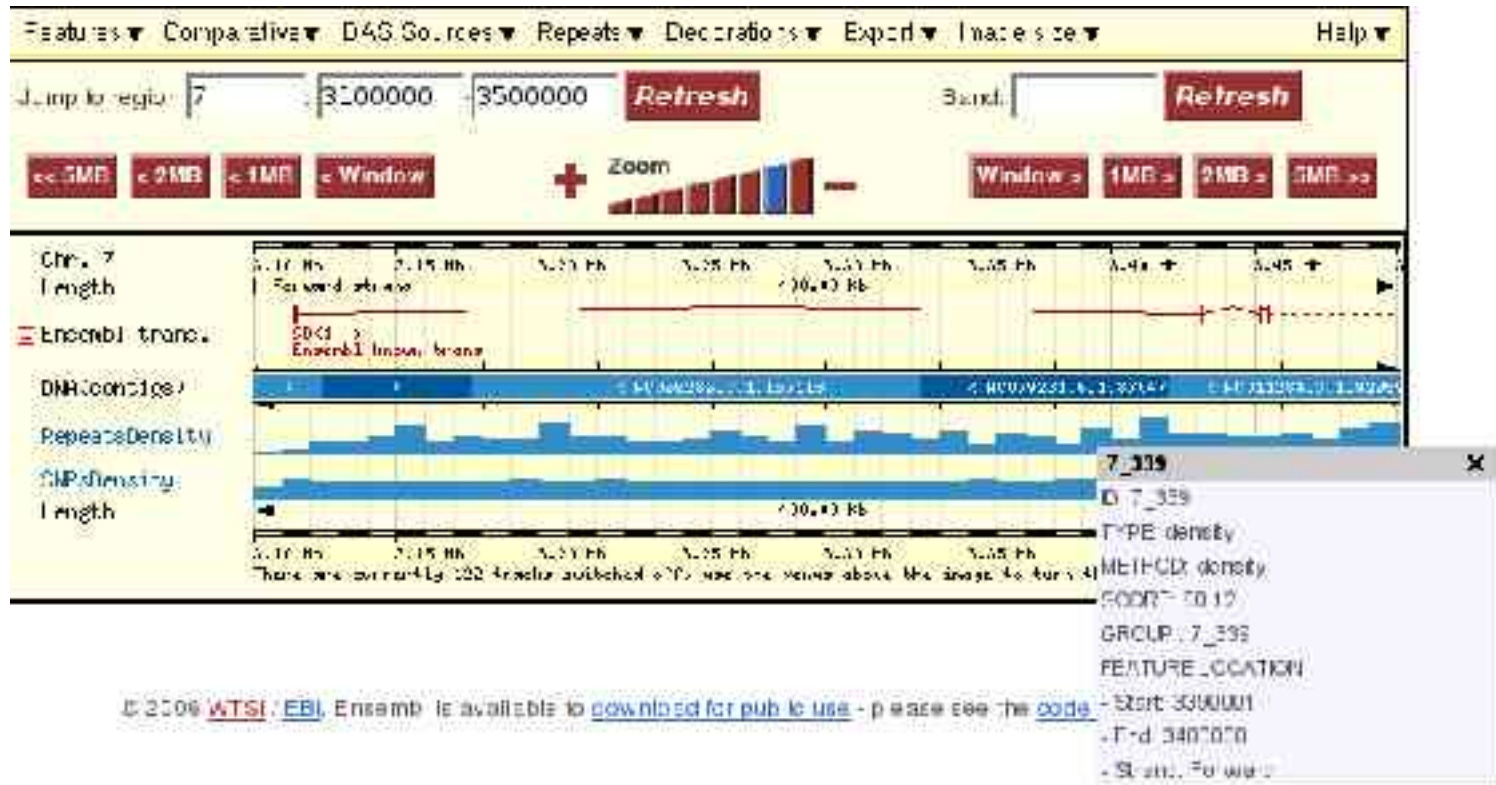


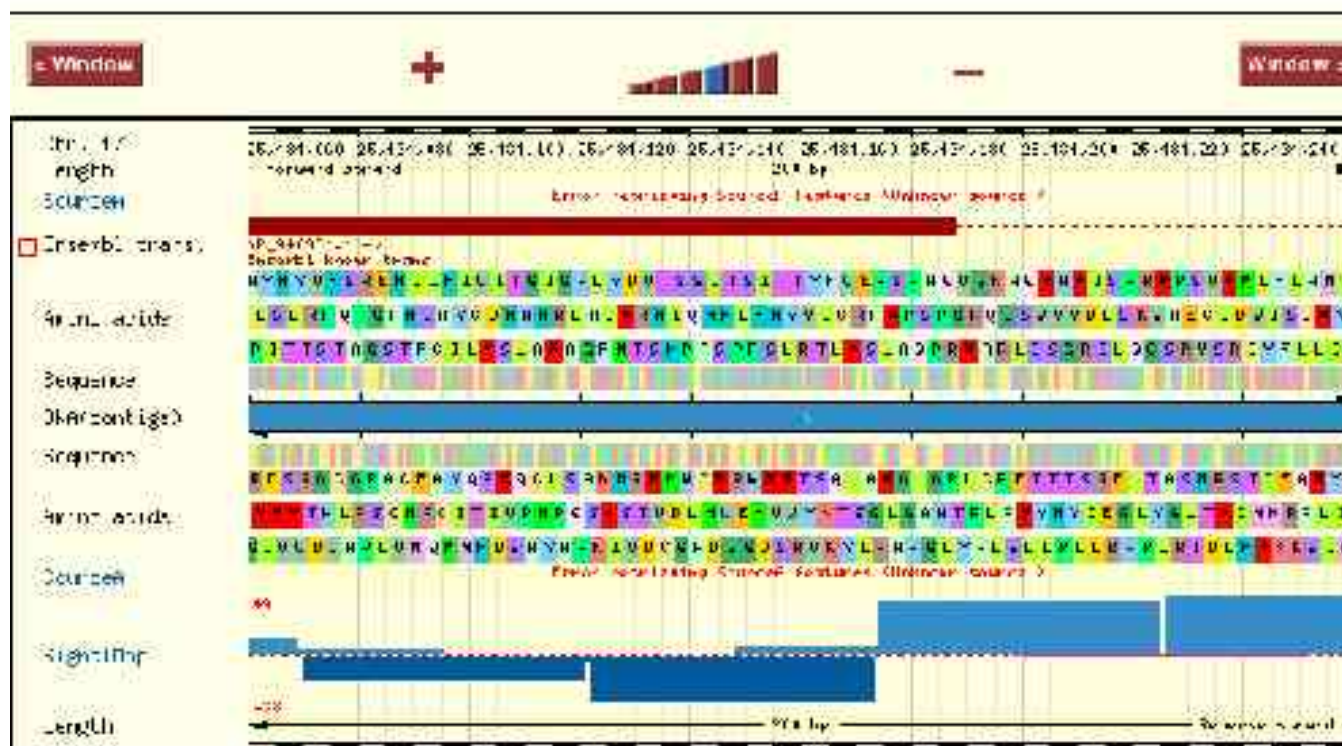
# Applying stylesheet



[http://www.ensembl.org/info/data/external\\_data/das/CSS\\_support.pdf](http://www.ensembl.org/info/data/external_data/das/CSS_support.pdf)

# Using score data







# Send the source

[http://www.ensembl.org/Homo\\_sapines/contigview?](http://www.ensembl.org/Homo_sapines/contigview?)

[conf\\_script=contigview;c=13:31829752.5:1;w=500000;h=;](#)

[add\\_das\\_source= \(name=hydraeuf\\_00001291+](#)

[url=http://das.ensembl.org/das+dsn=hydraeuf\\_00001291+](#)

[type=ensembl\\_location+color=red+strand=b+labelflag=u+stylesheet=n](#)

[+group=n+depth=10+score=n+active=1\)](#)



# Web-based source

```
browser position chr19:6900001-7400000
track name=BAC_end_pairs_bed description="BAC end pairs (BED)" useScore=1
color=ffcc66 url=http://www.ensembl.org/human/fastaview?faid=DNA_bacend_1533&id=$$
height=0.1
19      6503520 6641948 CTD-3084A17:AQ144413-AQ186293      1000      -      6503520
6641948 0      2      462,512 0,137917
19      6527567 6695619 RP11-357B22:AZ518790-AQ552859      1000      -      6527567
6695619 0      2      470,525 0,167528
19      6527640 6704480 RP11-134L9:AQ388324-AQ388334      1000      -      6527640
6704480 0      2      307,557 0,176284
```

[http://www.ensembl.org/info/data/external\\_data/examples/url\\_bed.txt](http://www.ensembl.org/info/data/external_data/examples/url_bed.txt)  
[http://www.ensembl.org/info/data/external\\_data/examples/url\\_pcl.txt](http://www.ensembl.org/info/data/external_data/examples/url_pcl.txt)  
[http://www.ensembl.org/info/data/external\\_data/examples/url\\_gff.txt](http://www.ensembl.org/info/data/external_data/examples/url_gff.txt)







# Exercise I

Create a file for upload :

- two features should belong to the group 'Probe A' and
- three features should belong to the group 'Probe B'
- stylesheet section that renders Probe A features in red and Probe B features in green

Upload the file and view the track in contigview

Attach the same source via auto-attach URL, but with stylesheet switched off and features grouped

- File format:

<http://www.sanger.ac.uk/Software/formats/GFF/>

[http://www.ensembl.org/info/data/external\\_data/das/CSS\\_support.pdf](http://www.ensembl.org/info/data/external_data/das/CSS_support.pdf)

- Examples:

[http://www.ensembl.org/info/data/external\\_data/das/das\\_upload\\_v1.euf](http://www.ensembl.org/info/data/external_data/das/das_upload_v1.euf)



Ensembl

## DAS setup





# DAS setup

- Resources
  - DasLite: <http://search.cpan.org/~rpettt/Bio-DasLite-0.14>
  - BioDas: <http://www.biodas.org/download/Bio::Das>
  - SOAP : <http://soaplite.com>
- Ini-files
  - DEFAULTS.ini
  - Preconfigure a source
- Data Upload
  - data journal
  - ProServer configuration



# DAS setup

DEFAULT.ini

```
;  
; Web services (Used for DAS etc)  
;  
ENSEMBL_WWW_PROXY      =  
DAS_REGISTRY_URL       = http://das.sanger.ac.uk/registry  
ENSEMBL_DAS_SERVERS     = [das.ensembl.org]  
ENSEMBL_DAS_UPLOAD_SERVER = das.mydomain.org  
ENSEMBL_DAS_UPLOAD_DB_HOST = das.mydomain.org  
ENSEMBL_DAS_UPLOAD_DB_PORT = 3306  
ENSEMBL_DAS_UPLOAD_DB_NAME = ens_upload  
ENSEMBL_DAS_UPLOAD_DB_USER = root  
ENSEMBL_DAS_UPLOAD_DB_PASS =
```



# DAS setup

## Preconfigure a source

Homo\_sapiens.ini

MULTI.ini

;

; DAS Config

;

[ENSEMBL\_INTERNAL\_DAS\_SOURCES]

SourceA = 1

SourceB = 1

SourceC = 1

[SourceA]

dsn = projectA

url = <http://das.ensembl.org/das>

on = [ contigview cytoview geneview ]



Ensembl

# DAS setup

[SourceA]

dsn	= projectA
url	= <a href="http://das.ensembl.org/das">http://das.ensembl.org/das</a>
on	= [ contigview cytoview geneview ]
type	= ensembl_location
label	= Source A probes
caption	= Source A probes
col	= black
labelflag	= U
homepage	= <a href="http://das.ensembl.org/das/projectA">http://das.ensembl.org/das/projectA</a>
helplink	= <a href="http://das.ensembl.org/das/projectA">http://das.ensembl.org/das/projectA</a>
strand	= r
depth	= 9999
group	= 0
stylesheet	= y
score	= n
select	= [ contigview ]
assembly	= NCBIM34
fasta	= [ DNA_A_probes ]

# DAS setup

## Data upload

```
mysql> describe hydra_journal;
```

Field	Type	Null	Key	Default	Extra
id	int(11)		PR	NULL	auto_increment
filetype	varchar(4)			ELF	
create_date	date	YES		NULL	
access_date	date	YES	PR	NULL	
email	varchar(64)	YES	PR	NULL	
passwd	varchar(128)	YES		NULL	
css	text	YES		NULL	

```
7 rows in set (0.03 sec)
```

```
mysql> select * from hydra_journal limit 1;
```

id	filetype	create_date	access_date	email	passwd	css
1	ELF	2005-02-24	2005-02-24	ek3@sanger.ac.uk		NULL



# DAS setup

## Data upload

Configure ProServer (eg/proserver.ini )

[hydraeuf]

state	= on
adaptor	= upload_euf
hydra	= dbi
transport	= dbi
basename	= euf
dbname	= ens_upload
host	= das.mydomain.org
port	= 3306
username	= rw_user
password	=



Ensembl

## Future development



# Future development



- User logins  
preserving Ensembl DAS configuration
- Ensembl as DAS server  
serving Ensembl data as DAS sources
- DAS for Functional genomics  
combining features and real value data
- Ensembl Upload Format, version 2  
uploading gene and protein based data
- `meta` request  
configure Ensembl display at the source
- Cross-assembly mapping  
display sources based on old assemblies on the current Ensembl
- DAS2 ?





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

[http://www.ensembl.org/info/data/external\\_data/das/DASTalk.ppt](http://www.ensembl.org/info/data/external_data/das/DASTalk.ppt)

[http://www.ensembl.org/info/data/external\\_data/das/DASTalk.pdf](http://www.ensembl.org/info/data/external_data/das/DASTalk.pdf)