DAS - Putting it all together!

Rob Finn
DAS Workshop 2009
Review of DAS - SOA

- http://www.dasregistry.org

Registry

Publish
- ProServer (CPAN)
- Dazzle

Bind

Find
- Bio::Das::Lite (CPAN)
- Dasobert

Clients

Servers
**EnsEMBL Client**

- Sequence and Features With Ranges

---

### EnsEMBL Transcript View

- Chromosome bands
- Contigs
- Ensembl/Havana genes

- TP53-203: Known protein coding Ensembl gene
- TP53-202: Known protein coding Ensembl gene
- TP53-201: Known protein coding Ensembl gene

- CCDS set
- Human cDNA

- All variations
- Reg. Features

---

**Gene Legend**
- Known protein coding
- Unclassified

**Reg. Features Legend**
- Intronic
- Splice site SNP

**Variation Legend**
- Non-synonymous coding
- Essential splice site
- Synonymous coding

There are currently 177 tracks turned off.
Ensembl Homo sapiens version 53.36o (NCBI36) Chromosome 17: 7,517,000 - 7,519,057
Feature Aggregation

- **Distributed Annotation System**
  - Simple Exchange Protocol

```plaintext
das/sequence?segment=ID

das/features?segment=ID
```
Feature Aggregation

- **Distributed Annotation System**
  - Simple Exchange Protocol

Diagram:
- **DAS Client**
- **DAS Server**
  - HTTP request
  - XML response
Feature Aggregation

• **Distributed Annotation System**
  – Simple Exchange Protocol

- Key Features
  - Few Reference Servers, Many Annotation/Feature Servers
  - Dumb Servers, Smart Clients
Pfam DAS client

Features
Request
List of sources

Bio::DAS::Lite and Pfam drawing code lib
A Different View

http://www.ebi.ac.uk/dasty/client/ebi.php?q=P00784&label=any&t=3
Using Ontologies

- DAS Feature Servers Ontology compliant
- EBI - ontology lookup
  - http://www.ebi.ac.uk/ontology-lookup/browse.do?ontName=BS
- Aggregation of feature ‘types’
Spice

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

RUN SPICE

Press here to start SPICE and automatically load the structure of Rhodanese (PDB accession code 1boi).

Run Spice

Or: Start SPICE with accession code:

<table>
<thead>
<tr>
<th>PDB</th>
<th>Launch</th>
</tr>
</thead>
<tbody>
<tr>
<td>1ycs</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UniProt</th>
<th>Launch</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
If the demo does not work!
Structure

- Fetched Using PDB request
- Rendered using Jmol libraries
- Dasobert Libraries
DAS Alignments

• Can be considered as an elaborate feature
  – Pairwise
    – *PREFIX*/das/alignment?query=ID
DAS Alignments

• Can be considered as an elaborate feature
  – Pairwise
    – \texttt{PREFIX/das/alignment?query=ID}
DAS Alignments

- Can be considered as an elaborate feature
  - Pairwise
    - `PREFIX/das/alignment?query=ID`
DAS Alignments

• Can be considered as an elaborate feature
  – Pairwise
  – PREFIX/das/alignment?query=ID
More Complex Alignments

Structure
Multiple Sequence
Multiple Structure Alignment

Sisyphus Database - Structural alignments for proteins with non-trivial relationships
DAS Alignment

• First structure is reference
  – Subsequent structures aligned with respect to this structure
  – Bold regions represent spatial equivalent positions

• Same DAS command as pairwise
  – Same co-ordinate system
DAS Alignments

- The Next Step....
  - Multiple Sequence Alignments
  - PREFIX/das/alignment?query=ID
DAS Alignments

- **Dealing with large alignments**
  - `PREFIX/das/alignment?query=ID[&subject=ID[RANGE]]` or/and `&rows=START-END`
DAS Alignments

- Dealing with large alignments
  - `PREFIX/das/alignment?query=ID[&rows=START-END]`
In Practice

- Pfam alignments vary in size
  - 2 - 75,000+ sequences
  - Paging Essential
- Simple DAS alignment client
  - HTML, AJAX
Any Questions?