Databases and ontologies

Dasty and UniProt DAS: a perfect pair for protein feature visualization

Philip Jones¹,* Nisha Vinod¹, Thomas Down², Andre Hackmann¹, Andreas Kahari¹, Ernst Kretschmann¹, Antony Quinn¹, Daniela Wieser¹, Henning Hermjakob¹ and Rolf Apweiler¹

¹EMBL Outstation—European Bioinformatics Institute (EBI), Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK and ²Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

Received on April 4, 2005; accepted on May 17, 2005
Advance Access publication May 19, 2005

ABSTRACT

Summary: In this study, we present two freely available and complementary Distributed Annotation System (DAS) resources: a DAS reference server that provides up-to-date sequence and annotation from UniProt, with additional feature links and database cross-references from InterPro and a DAS client implemented using Java and Macromedia Flash that is optimized for the display of protein features.

Availability: The Dasty DAS client is freely available at the following URLs: Dasty DAS Client (http://www.ebi.ac.uk/das-srv/uniprot/dasty/), Dasty Example Query (http://www.ebi.ac.uk/das-srv/uniprot/dasty/index.jsp?id=Q24488), UniProt DAS Service (http://www.ebi.ac.uk/das-srv/uniprot/das), Dasty Documentation (http://www.ebi.ac.uk/das-srv/uniprot/das), and UniProt DAS Documentation (http://www.ebi.ac.uk/uniprot-das/). It is intended that the Dasty client package and source code will be freely accessed via the EBI site at the given link.

Contact: For further details of the Dasty client, please contact N.V. (nvinod@ebi.ac.uk). For further details of the UniProt DAS server, please contact P.J. (pjones@ebi.ac.uk).

1 INTRODUCTION

The Distributed Annotation System (DAS) (Dowell et al., 2001) specifies a lightweight protocol to allow the positional feature data for biological molecules to be requested using HTTP requests (i.e. a URL), with the response being returned as XML. The DAS specification was originally designed to allow the feature information to be served in relation to a genomic sequence, however, the DAS protocol can be leveraged to serve the protein sequence and feature information. DAS servers are separated into two kinds: reference servers provide sequence data and where appropriate scaffolding information and annotation servers that provide feature information only. An annotation server is required to provide coordinates for each feature that it serves and indicate a suitable DAS reference server from which the corresponding sequence can be obtained. A DAS client is an application that is able to connect to at least one reference server and one annotation server and merge the information from these servers in a unified display. The true power of this mechanism is that a DAS client is able to connect to any number of annotation servers, located anywhere in the world, serving annotation data from different laboratories (hence ‘distributed’).

2 DASTY: A DAS CLIENT OPTIMIZED FOR PROTEIN VISUALIZATION

Dasty has been developed with protein DAS services in mind. It offers an attractive and powerful graphical user interface (Fig. 1). A user can request any protein identification from the available DAS servers. The resulting view is furnished with tools that allow the user to zoom by sequence coordinates and DAS tracks; sort the order of DAS tracks by feature type or score; view details of each features type, description, start and end coordinates, and score and follow any hyperlinks to additional information or external data resources. Depending on the type of the feature, Dasty allows features to be colour coded. At present such settings can be set by editing a configuration file, however, a user interface to this configuration is under development. Dasty has been developed using Java and Macromedia Flash™. Dasty can be used in any modern Internet browser with the freely available Flash plugin installed.

The ability to view non-positional features available from UniProt DAS, such as protein synonyms and literature references, is currently being built into the client.

3 UniProt DAS PROTEIN REFERENCE SERVER

UniProt DAS has been made available from the EBI since July 2004. The server is able to access sequence and feature data from the UniProt data warehouse via the Aristotle API (Apweiler et al., 2004; Kretschmann et al., 2004), hence, it acts as both a reference and annotation server in one. Feature data available from UniProt DAS include all human annotations and feature predictions from SwissProt and TrEMBL. In both cases, this information is anchored to the latest sequence data. UniProt DAS is synchronized with the fortnightly release of UniProt, providing up-to-date sequence and annotation.

*To whom correspondence should be addressed.
In addition, the UniProt data warehouse provides secondary database cross-references from the InterPro (Mulder et al., 2005). UniProt DAS is able to serve these cross references as hyperlinks to the secondary database entry. The DAS response includes a brief description of the InterPro entry and where applicable, a note about the status of the annotation. UniProt DAS has been developed using the Java programming language and is based upon the Dazzle general purpose DAS server (Table 1).

Table 1. DAS resources on the Internet

<table>
<thead>
<tr>
<th>Resource</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>biodas.org website</td>
<td><a href="http://biodas.org/">http://biodas.org/</a></td>
</tr>
<tr>
<td>DAS specification</td>
<td><a href="http://biodas.org/documents/spec.html">http://biodas.org/documents/spec.html</a></td>
</tr>
<tr>
<td>Dazzle DAS server homepage</td>
<td><a href="http://www.biojava.org/dazzle/">http://www.biojava.org/dazzle/</a></td>
</tr>
</tbody>
</table>

4 CONCLUSIONS

Dasty and UniProt DAS together provide a freely available and powerful resource for distributed protein sequence and annotation services. UniProt DAS opens all UniProt, IPI and related InterPro protein sequence and feature data to query via the DAS protocol.

ACKNOWLEDGEMENTS

This work has been supported by the European FP6 Programme, ‘BioSapiens’ grant, contract number LHSG-CT-2003–503265 and by the BBSRC ‘ISPIDER’ grant.

REFERENCES


