

Integrating T-cell epitope annotations with sequence and structural information using DAS

Carmen M. Diez-Rivero¹, María García-Boronat¹ and Pedro A Reche^{1, *}

¹ImmuneMedicine Group, Department of Microbiology, Division of Immunology, Facultad de Medicina, Universidad Complutense de Madrid, Ave Complutense, s/n. Madrid 28040, Spain;
Pedro A Reche* - Email: parecheg@med.ucm.es; Phone: 34 91 394 7229; Fax: 34 91 394 1641; * Corresponding author

received November 09, 2008; accepted November 23, 2008; published December 06, 2008

Abstract:

Immunoinformatics is an emerging new field that benefits from computational analyses and tools that facilitate the understanding of the immune system. A large number of immunoinformatics resources such as immune-related databases and analysis software are available through the World Wide Web for the benefit of the research community. However, immunoinformatics developments have sometimes remained isolated from mainstream bioinformatics. Therefore, there is clearly a need for integration, which will empower the exchange of data and annotations within the scientific community in a quick and efficient fashion. Here, we have chosen the Distributed Annotation System (DAS), for integrating in house annotations on experimental and predicted HLA I-restriction elements of CD8 T-cell epitopes with sequence and structural information.

Keywords: DAS; annotation; epitope; HLA I

Abbreviations: CMV - Cumulative Phenotypic Frequency; DAS - Distributed Annotation System; HLA I - Human Leukocyte Antigen class I; PSSM - Position Specific Scoring Matrix

Background:

Recent years have witnessed the birth of Immunoinformatics, an emerging subdiscipline of Bioinformatics. With the burgeoning explosion of immunological data, computational analysis has become an essential element of immunology research, facilitating the understanding of the immune function by modeling the interactions among immunological components [1]. Another major role in Immunoinformatics is the efficient management, storage, and annotation of such data. Following those principles, a large number of immunoinformatics resources including immune-related databases and sophisticated analysis software, are available through the World Wide Web. Collectively, these resources contribute to the advances made in immunological research. Yet, there is still a major step to be taken towards the integration of all these resources, as ideally, multiple research groups should be able to exchange and compare their data, in a quick and efficient fashion.

The distributed annotation system (DAS) defines a communication protocol used to exchange biological annotations from a number of heterogeneous distributed databases [2]. The key idea behind the DAS concept is that annotations should not be provided by single centralized databases but instead be spread over multiple sites. DAS follows a simple http-based client-server protocol, where clients make requests in the form of a URL to the servers,

and receive simple XML responses. The basic system is composed of a reference server, one or more annotation servers, and an annotation viewer. The reference server is responsible for serving genome maps, sequences and information related to the sequencing process. Annotation servers are responsible for returning the annotations on a defined region (given a start and stop position coordinates) of the genome or proteome. The annotation viewer can either be a simple web browser, which will visualize the raw XML data provided by the server, or a graphical client which translates the XML annotations such as the Center for Biological Sequence Analysis (CBS) DAS viewer [3] accessible at <http://www.cbs.dtu.dk/cgi-gin/das>.

In this article, we will show how an epitope database can be integrated to other database resources using DAS. For that we will describe TEPIDAS, a DAS Annotation Server of HLA I-restricted CD8 T-cell epitopes specific of human pathogenic organisms. TEPIDAS falls into the category of annotation servers and is registered at the DAS registry since February of 2008, and has the unique id DS_545.

Description:

Overview

TEPIDAS is a DAS annotation server that follows the UniProt coordinates system to annotate the experimental and potential HLA I-restriction elements of a set of CD8 T-cell epitopes. TEPIDAS is implemented using ProServer [4], a lightweight Perl-based DAS server. When a client makes a query to the

Conclusion:

DAS is an important, simple and yet powerful system for exchanging and viewing biological data that is already being used in real-world bioinformatics applications. The TEPIDAS annotation server described in this chapter is a clear example of how epitope data can be integrated and shared by the research community using the DAS architecture. The complexity of immune interactions and the data intensive nature of immune research make Immunoinformatics a suitable area that could greatly benefit from the advantages of using such a powerful integration and annotation system, allowing to gain a more insightful understanding of the complexities of the immune system.

Acknowledgment:

We would like to thank Alfonso Valencia, Osvaldo Graña, and Jaime Fernandez Vera from the Spanish National Cancer Research Center (CNIO) for their helpful advice on DAS and ProServer. Work and authors were supported by grant SAF2006-07879 from the “Ministerio de Educación y Ciencia” of Spain, granted to PAR.

References:

- [01] M. N. Davies and D. R. Flower, *Drug Discov. Today*, 12: 389 (2007) [PMID: 17467575]
- [02] R. D. Dowell, R.M. *et al.*, *BMC Bioinformatics*, 2: 7 (2001) [PMID: 11667947]
- [03] P. I. Olason, *Nucleic Acids Res.*, 33: W468 (2005) [PMID: 15980514]
- [04] R. D. Finn *et al.*, *Bioinformatics* 23: 1568 (2007) [PMID: 17850653]
- [05] C. H. Wu *et al.*, *Nucleic Acids Res.*, 34: D187 (2006) [PMID: 16381842]
- [06] P. A Reche *et al.*, *Bioinformatics*, 21: 2140 (2005) [PMID: 15657103]
- [07] P. A Reche and E. L. Reinherz, *Methods Mol. Biol.*, 409:185 (2007) [PMID: 18450001]
- [08] P. A Reche *et al.*, *Med. Immunol.*, 5: 1 (2006) [PMID: 16674822]
- [09] Prlic *et al.*, *Bioinformatics*, 21: ii40 (2005) [PMID: 16204122]

Edited by P. Kanguane

Citation: Diez-Rivero *et al.*, *Bioinformatics* 3(4): 156-158 (2008)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited.