Immune Epitope Database NEWSLETTER

Volume 1, Issue 2 http://www.iedb.org August 2012

The IEDB at AAI 2012

The IEDB exhibit booth was present at the Annual Meeting of the American Association of Immunologists (AAI) held May 4 – 8 in Boston. The booth was staffed for three days by Senior Curators Nima Salimi and Kerrie Vaughan, and by Bioinformatics postdoctoral fellow Yohan Kim. There were approximately 50 visitors per day, with a noticeable increase in visitors already familiar with the IEDB compared to previous years. Visitor interactions ranged from brief explanations of the IEDB and its capabilities, after which visitors were offered literature assembled by the IEDB team providing further details, to extended live demonstrations of the IEDB. The booth included two computers where visitors were invited to use the IEDB with guidance from IEDB staff. Users were generally successful in finding answers to their queries of interest, and were impressed by the quantity and quality of the data available.

In addition to the exhibit booth, the IEDB presented a workshop and a poster. The IEDB workshop was held on Sunday morning and was attended by approximately 20 people. The presentation was given by Nima Salimi and consisted of two parts. The first half of the workshop covered the basic features of the database website, the available tools, and new features introduced over the past year. It included a demonstration of three sample queries (infectious disease, autoimmunity, and allergy) and one sample T cell epitope prediction example. The second half of the workshop was allotted for audience questions. The poster, titled "A meta-analysis of the existing knowledge of immunoreactivity against hepatitis C virus (HCV)," was presented by Kerrie Vaughan.

The IEDB at FOCIS 2012

Senior Curators Kerrie Vaughan and Randi Vita staffed the IEDB exhibit booth at the Federation of Clinical Immunology Societies (FOCIS) conference in Vancouver, BC Canada, June 20-23. Additionally, a poster titled "The IEDB: Making Immune Epitope Data Freely Accessible" was presented by Randi Vita and attracted end-users with specific questions. The booth was visited by approximately 40 scientists with generally more

Continued on Page 2

Inside this Issue

Continued from Page 1

epitope specialists than in years past. The majority of the booth and poster visitors had specialized uses of the IEDB and brought specific questions. In the past, many visitors were seeking more general information about what the IEDB was, while this year it seemed more were wanting to know if and how the IEDB could benefit their particular needs.

The IEDB at ISMB 2012

The IEDB exhibit booth made its final appearance for the year at the Intelligent Systems for Molecular Biology (ISMB) 2012 Conference in Long Beach, CA, July 14 - 17. The booth was staffed by Senior Curators Nima Salimi and Debbie Shackelford, and Bioinformatics postdoctoral fellow Sinu Paul. ISMB is one of the largest bioinformatics conferences and is an official conference of the International Society for Computational Biology. This was the first time that the IEDB had exhibited at this meeting, and it was a wonderful opportunity to introduce a whole new community to the IEDB. Nima Salimi also presented a 25-minute technology talk on the general features of the IEDB and tools on the last day of the conference.

Figure 1. Senior Curator Debbie Shackelford, Ph.D. and Bioinformatics postdoctoral fellow Sinu Paul, Ph.D. staff the IEDB exhibit booth at ISMB 2012.

IEDB v2.9 Released in June 2012

Two noteworthy new features in the IEDB website were introduced in the latest release that was deployed on June 5, 2012. The first was revising the Epitope Type field on the Advanced Search pages to use radio buttons instead of pull-down menus. These now match the selections available on the home page search. The second was the introduction of a completely revised B cell assay finder, which is available on the Advanced B Cell Search page. Assays are now displayed in a tree structure, as seen in Figure 2, and a specific assay may appear in more than one location in the tree. Users can browse the tree structure and select their assays of interest. They can also search for assays in the Find box in the lower left corner of the Assay Finder. In the figure shown, a search for "cross blocking" resulted in one item found, as displayed in the top of the panel. By clicking on "Highlight in Tree", the user can expand the tree to reveal the assay of interest.

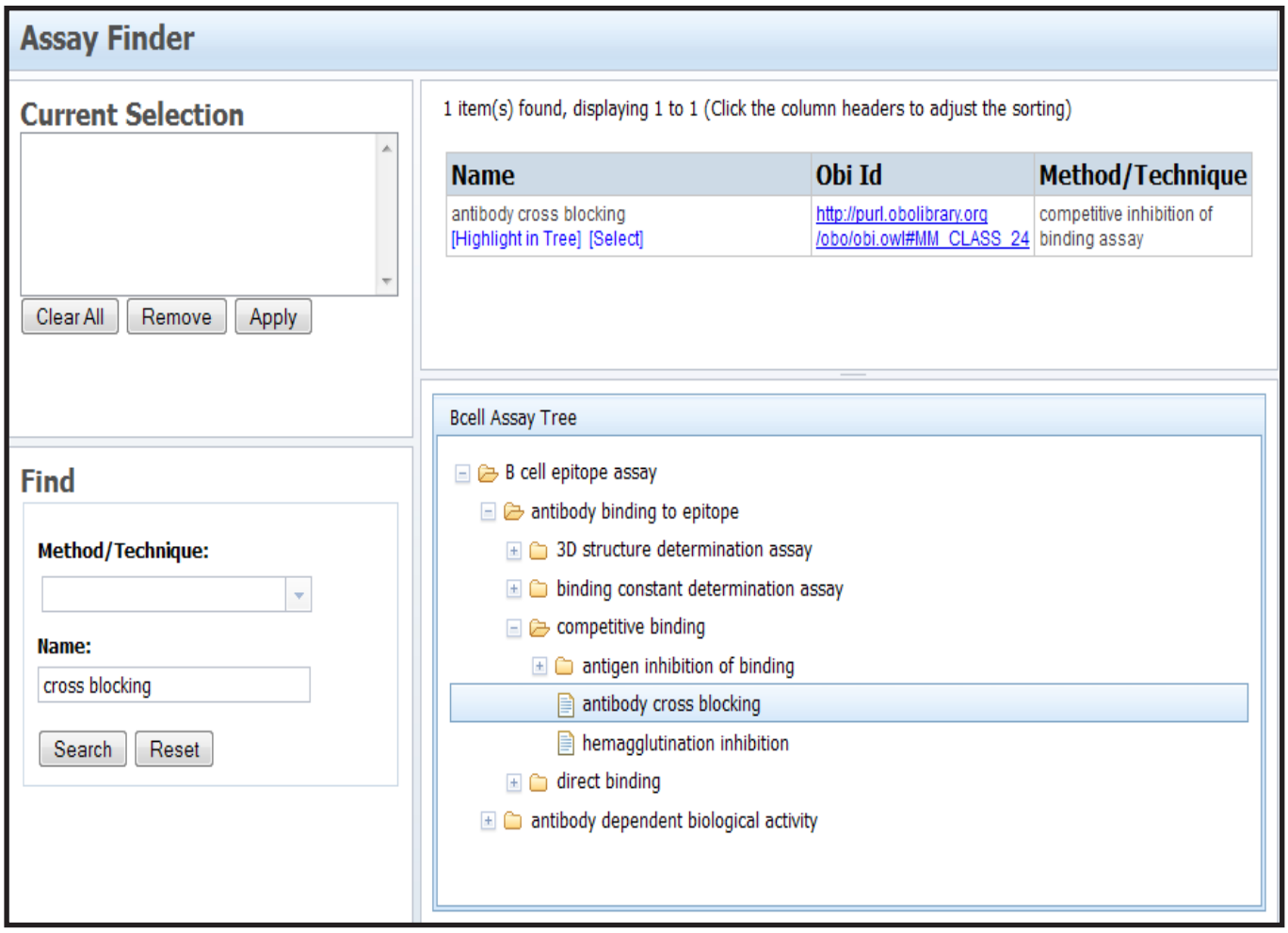

Figure 2. The latest release of the IEDB website features a new Assay Finder for the advanced B Cell Search, located on the Search pull-down menu. Assays are displayed in a tree structure. A search for a "cross blocking" assay in the lower left Find box found one result, "antibody cross blocking". Selecting "Highlight in Tree" expands the tree structure to show the item.

Recent Publications

The Immune Epitope Database: A Historical Retrospective of the First Decade

Salimi N, Fleri W, Peters B, Sette A.

Immunology. 2012 Jun 8. doi: 10.1111/j.1365-2567.2012.03611.x. [Epub ahead of print]

PMID: 22681406

As the amount of biomedical information available in the literature continues to increase, databases that aggregate this information continue to grow in importance and scope. The population of databases can occur either through fully automated text mining approaches or through manual curation by human subject experts. We here report our experiences in populating the NIAID sponsored Immune Epitope Database and Analysis Resource (IEDB, http://iedb.org) which was created in 2003, and as of 2012 captures the epitope information from approximately 99% of all papers published to date that describe immune epitopes (with the exception of cancer and HIV data). This was achieved using a hybrid model based on automated document categorization and extensive human expert involvement. This task required automated scanning of over 22 million PubMed abstracts followed by classification and curation of over 13,000 references, including over 7,000 infectious disease-related manuscripts, over 1,000 allergy-related manuscripts, roughly 4,000 related to autoimmunity, and 1,000 transplant/alloantigen-related manuscripts. The IEDB curation involves an unprecedented level of detail, capturing for each paper the actual experiments performed for each different epitope structure. Key to enabling this process was the extensive use of ontologies to ensure rigorous and consistent data representation as well as interoperability with other bioinformatics resources, including PDB, ChEBI, and NIAID Bioinformatics Resource Centers. A growing fraction of the IEDB data derives from direct submissions by research groups engaged in epitope discovery, and is being facilitated by the implementation of novel data submission tools. The present explosion of information contained in biological databases demands effective query and display capabilities in order to optimize the user experience. Accordingly, the development of original ways to query the database, on the basis of ontologically-driven hierarchical trees, and display of epitope data in aggregate in a biologically intuitive yet rigorous fashion is now at the forefront of the IEDB efforts. We also highlight advances made in the realm of epitope analysis and predictive tools available in the IEDB.

A Meta-Analysis of the Existing Knowledge of Immunoreactivity against Hepatitis C Virus (HCV)

Kim Y, Vaughan K, Greenbaum J, Peters B, Law M, Sette A.

PLoS One. 2012;7(5):e38028. Epub 2012 May 31.

PMCID: PMC3364976; PMID: 22675428 [PubMed - in process]

Approximately 3% of the world population is infected by HCV, which represents a major global health challenge. Almost 400 different scientific reports present immunological data related to T cell and antibody epitopes derived from HCV literature. Analysis of all HCV-related epitope hosted in the Im-

Continued on Page 5

Continued from Page 4

mune Epitope Database (IEDB), a repository of freely accessible immune epitope data, revealed more than 1500 and 1900 distinct T cell and antibody epitopes, respectively. The inventory of all data revealed specific trends in terms of the host and the HCV genotypes from which sequences were derived. Upon further analysis we found that this large number of epitopes reflects overlapping structures, and homologous sequences derived from different HCV isolates. To access and visualize this information we developed a novel strategy that assembles large sets of epitope data, maps them onto reference genomes and displays the frequency of positive responses. Compilation of the HCV immune reactivity from hundreds of different studies, revealed a complex and thorough picture of HCV immune epitope data to date. The results pinpoint areas of more intense reactivity or research activities at the level of antibody, CD4 and CD8 responses for each of the individual HCV proteins. In general, the areas targeted by the different effector immune functions were distinct and antibody reactivity was positively correlated with hydrophilicity, while T cell reactivity correlated with hydrophobicity. At the sequence level, epitopes frequently recognized by both T cell and B cell correlated with low variability, and our analysis thus highlighted areas of potential interest for practical applications. The human reactivity was further analyzed to pinpoint differential patterns of reactivity associated with acute versus chronic infection, to reveal the apparent impact of glycosylation on T cell, but not antibody responses, and to highlight a paucity of studies involved antibody epitopes associated with virus neutralization.

Immune epitope database analysis resource.

Kim Y, Ponomarenko J, Zhu Z, Tamang D, Wang P, Greenbaum J, Lundegaard C, Sette A, Lund O, Bourne PE, Nielsen M, Peters B.

Nucleic Acids Res. 2012 Jul;40(Web Server issue):W525-30. Epub 2012 May 18.

PMID: 22610854 [PubMed - in process]

The immune epitope database analysis resource (IEDB-AR: http://tools.iedb.org) is a collection of tools for prediction and analysis of molecular targets of T- and B-cell immune responses (i.e. epitopes). Since its last publication in the NAR webserver issue in 2008, a new generation of peptide:MHC binding and T-cell epitope predictive tools have been added. As validated by different labs and in the first international competition for predicting peptide:MHC-I binding, their predictive performances have improved considerably. In addition, a new B-cell epitope prediction tool was added, and the homology mapping tool was updated to enable mapping of discontinuous epitopes onto 3D structures. Furthermore, to serve a wider range of users, the number of ways in which IEDB-AR can be accessed has been expanded. Specifically, the predictive tools can be programmatically accessed using a web interface and can also be downloaded as software packages.

Curation Status

Curation of data relating to peptidic and non-peptidic epitopes for all infectious diseases, allergens, autoimmune diseases, and transplant/alloantigens is current for references appearing in PubMed. A query for new potentially relevant epitope references is run biweekly to update the database since all reference categories are in maintenance mode for both peptidic and non-peptidic epitopes. As of July 2012, data from approximately 14,200 literature references and direct data submissions have been incorporated into the IEDB. The IEDB contains data for over 94,000 epitopes, 2,932 epitope source organisms, and 615 restricting MHC alleles.

Contact Information

The Immune Epitope Database and Analysis Resource is suported by a contract from the National Institute of Allergy and Infectious Disease, NIH, DHHS (Contract HHSN272201200010C). The newsletter is distributed four times a year. We welcome communication from the users of the IEDB database and invite suggestions for articles in future issues. To subscribe to the IEDB newsletter or to contact project staff, send your email information to the email address below.

Email: contact@iedb.org
Web: http://www.iedb.org

Immune Epitope Database and Analysis Resource
c/o La Jolla Institute for Allergy & Immunology
9420 Athena Circle
La Jolla, CA 92037
(858) 752-6500

Principal Investigator:
Alessandro Sette, Ph.D.
alex@liai.org

Co-Principal Investigator:
Bjoern Peters, Ph.D.
bpeters@liai.org

Project Director:
Stephen Wilson, Ph.D.
swilson@liai.org

Production:
Emily Seymour
Ward Fleri, Ph.D.