Immune Epitope Database NEWSLETTER

Volume 7, Issue 4 http://www.iedb.org November 2010

Allergy Meta-Analysis

Scientists at LIAI performed a meta-analysis on immune epitope data related to allergy as of May 2010. The universe of allergens represents an incredible array of antigens from a broad spectrum of plants and animals. Moreover, defining allergic sequelae is complex in the human host and in their respective animal models. It is therefore challenging to analyze the collective immunological data related to human allergy. Analysis of epitope data provides an inventory of current immunological data and can highlight knowledge gaps and areas for future work. The meta-analysis includes all epitope structures (peptidic and non-peptidic), their source and associated immune reactivity, including assay details and disease-related data. To date, more than 4,500 allergy-related epitopes derived from 270 different allergens have been captured. This includes both peptidic and non-peptidic allergens defined in humans as well as animal models. It is noteworthy that of all the meta-analyses conducted by LIAI so far, this is the first one to include non-peptidic epitopes. Protein allergens were categorized according to their source organism which included plants, animals, insects, parasites, and fungi. Non-peptidic allergens were categorized into four groups including drugs and biologicals, industrial compounds or those related to occupational exposure, metals, model haptens, and carbohydrates from plants. Global analysis of the data provides, for the first time, an inventory of allergy data and serves to identify critical knowledge gaps.

The vast majority of allergy epitopes, both peptidic and non-peptidic, were defined for B cells/antibodies in humans. In these records, IgE-mediated reactivity figured prominently; however IgG, IgG1, IgG4, IgA, IgM, Ig2b (mouse), IgG2a (mouse), and IgG3 were also reported. Other hosts included monkeys, pigs, dogs, rabbits, guinea pigs, rats, and mice. The majority of peptidic epitopes were defined for foods (cow's milk, wheat, and peanuts) and plants (tree and grass pollens), while the majority of non-peptidic epitopes were defined for drugs and biologicals (antibiotics).

Continued on Page 2

Inside this Issue

Continued from Page 1

The majority of T cell epitopes were defined as CD4+/class II, with very few being defined for CD8+/class I. MHC-restriction was also captured for T cell reactivity when reported. Here, the vast majority of human epitopes were restricted by HLA-DQ and HLA-DR. Other reported alleles included HLA-A2, HLA-DP, HLA-DRB1, HLA-DRA, HLA-DRB3, HLA-DRB4, and HLA-DRB5.

Interesting findings included the ratio of B to T cell epitopes between food allergies and airborne/respiratory allergies. The vast majority of food allergen-related epitopes were described for B cells, whereas a fairly even number of B and T cell epitopes were defined for airborne allergens. It is not clear why this is the case, but may have to do with historical analysis of allergies to foods such as milk, peanuts, and eggs, which represent a large portion of that data. We also found that the distribution of epitopes varied greatly between allergen and species. For certain species, the majority (if not all) of the known allergens had epitope-related data (e.g.. timothy grass allergens), while other species had epitope data from only a small number of known allergens (e.g. apple). Likewise, of those allergens representing the epitope-related data, very few had an equal distribution of both B and T cell epitopes. Instead, the total number of epitopes reported per allergen tended to be of one response type (B cell) or the other (T cell), and not both. Nevertheless, the overall completeness of the epitope-specific allergy data with respect to known allergens on a species basis was relatively good (40%); however the epitope data represented only about 17% of all allergens listed by the International Union of Immunological Societies (IUIS). Finally, the ability to capture non-peptidics was unique and allowed a first-time inventory and assessment of important drug and contact allergens.

The allergy-related data contained within the IEDB to date comprise mainly type I (e.g. rhinitis and asthma) and type IV hypersensitivity (e.g. contact dermatitis). Indeed, to date, allergy specific-IgE and CD4+ T cell responses make up 83% of the data. However, type II (e.g. drugs reactions) and type III hypersensitivities (food and airborne allergens) were also described, and we saw this in the form of non-IgE antibody reactivities and CD8+ T cell reactivity. While it was surprising that so little of the non-human antibody responses were allergy-specific IgE, this may have simply reflected some of the barriers associated with these models. Whether this finding will be of significance going forward as it pertains to the use of mice and other small mammals for immune epitope identification targeting human allergens remains a point for discussion.

Lacking from both peptidic and non-peptidic data were allergy epitopes defined for class I/CD8+ T cells. Also lacking from the human data were epitopes identified as being therapeutic in vivo. It is anticipated that this would be of significant interest for the field going forward.

IEDB "How-To" Videos

A series of "How-To" video tutorials is available on the IEDB Solutions Center. They can be found in the *Knowledgebase and Forums* menu under *Tutorials and Reference Materials*. The two videos are divided into four categories: *Overview of the IEDB*, *Searching the IEDB*, *Understanding Query Results*, and *Overview of Support*. The short videos feature actual screen animations and voice-overs from IEDB curators. The collection of videos will grow and evolve as the interfaces and capabilities of the IEDB continually improve. Five

Continued on Page 3

Continued from Page 2

new video tutorials are now available in the IEDB Solutions Center under the *Searching the IEDB* category. They can be accessed directly at http://iedb.zendesk.com/entries/140865-how-to-videos. Their titles are *Search by Disease Type*, *Simple Search: Finding Epitopes from a Protein*, *Advanced Search Overview*, *Advanced B cell search*: *Example of finding antibody cross-reaction*, and *Advanced T cell search: Example of designating assay type*. Other videos for near-term production include *Introduction to the IEDB site and Search Options*, *Using Finders*, *Browse by Organism*, *Results: Epitope Listing*, *Details of Individual Entries*, and *Using the Export Function*, *Getting Help*, and *Posting a Question or Topic in User Forums*.

Data Submissions Tool Enhancements

Efforts are currently underway to further simplify and expand direct data submission to the IEDB via the Data Submission Tool. The original method of submitting data via an XML format had previously been augmented by the Data Submission Tool that allowed users to utilize a spreadsheet style system using templates for specific T cell and antibody epitope submissions. While not addressing all possible experiments, they fully addressed the needs of the Epitope Discovery contracts. These templates are currently being expanded to accommodate the submission requirements of the new Epitope Discovery contracts. In addition, a data submission wizard is being developed that will guide and prompt the user through the process of submitting data. A wizard for T cell epitope submissions is near completion, and the capability to handle MHC binding, B cell, and MHC ligand elution assays will be added in the next six months. The ability to export data from the wizard to a spreadsheet for further editing will also be added.

IEDB at 2011 Conferences

The IEDB exhibit booth will be present at two conferences in 2011. The first is the AAI Annual Meeting, March 18-22, in San Francisco. Plans are being made to hold a user workshop during the conference. The second is the Federation of Clinical Immunology Societies (FOCIS) conference in Washington, DC, June 23-26.

Recent Publications

Modeling biomedical experimental processes with OBI.

Brinkman RR, Courtot M, Derom D, Fostel JM, He Y, Lord P, Malone J, Parkinson H, Peters B, Rocca-Serra P, Ruttenberg A, Sansone SA, Soldatova LN, Stoeckert CJ Jr, Turner JA, Zheng J; OBI consortium.

J Biomed Semantics. 2010 Jun 22;1 Suppl 1:S7.

PMID: 20626927

BACKGROUND: Experimental descriptions are typically stored as free text without using standardized terminology, creating challenges in comparison, reproduction and analysis. These difficulties impose limitations on data exchange and information retrieval. RESULTS: The Ontology for Biomedical Investigations (OBI), developed as a global, cross-community effort, provides a resource that represents biomedical investigations in an explicit and integrative framework. Here we detail three real-world applications of OBI, provide detailed modeling information and explain how to use OBI. CONCLUSION: We demonstrate how OBI can be applied to different biomedical investigations to both facilitate interpretation of the experimental process and increase the computational processing and integration within the Semantic Web. The logical definitions of the entities involved allow computers to unambiguously understand and integrate different biological experimental processes and their relevant components. AVAILABILITY: OBI is available at http://purl.obolibrary.org/obo/obi/2009-11-02/obi.owl.

MHC class II epitope predictive algorithms.

Nielsen M, Lund O, Buus S, Lundegaard C.

Immunology. 2010 Jul;130(3):319-28. Epub 2010 Apr 12.

PMID: 20408898

SUMMARY: Major histocompatibility complex class II (MHC-II) molecules sample peptides from the extracellular space, allowing the immune system to detect the presence of foreign microbes from this compartment. To be able to predict the immune response to given pathogens, a number of methods have been developed to predict peptide-MHC binding. However, few methods other than the pioneering TEPITOPE/ProPred method have been developed for MHC-II. Despite recent progress in method development, the predictive performance for MHC-II remains significantly lower than what can be obtained for MHC-I. One reason for this is that the MHC-II molecule is open at both ends allowing binding of peptides extending out of the groove. The binding core of MHC-II-bound peptides is therefore not known a priori and the binding motif is hence not readily discernible. Recent progress has been obtained by including the flanking residues in the predictions. All attempts to make ab initio predictions based on protein structure have failed to reach predictive performances similar to those that can be obtained by data-driven methods. Thousands of different MHC-II alleles exist in humans. Recently developed pan-specific methods have been able to make reasonably accurate predictions for alleles that were not included in the training data. These methods can be used to define supertypes (clusters) of MHC-II alleles where alleles within each supertype have similar binding specificities. Furthermore, the pan-specific methods have been used to make a graphical atlas such as the MHCMotifviewer, which allows for visual comparison of specificities of different alleles.

Curation Update

Curation of data relating to peptidic epitopes for all infectious diseases and peptidic and non-peptidic epitopes for allergens is current for references appearing in PubMed as of the end of June 2010. A query for new potentially relevant epitope references is run quarterly to update the database. Curation of peptidic epitopes for all autoimmune diseases is essentially complete. Curation of non-peptidic epitopes for all autoimmune and infectious diseases is in progress and will be completed in early 2011. As of October 2010, data from approximately 11,200 references have been incorporated into the IEDB. The IEDB contains data for over 77,000 epitopes, 2,548 epitope source organisms, and 571 restricting MHC alleles. Users are invited to bring references to our attention that are potentially relevant to the IEDB but do not appear in the database. References that are deemed to meet the IEDB criteria for curation will be queued for processing in accordance to our NIAID-directed priorities (Category A-C priority pathogens, emerging and re-emerging infectious diseases, other infectious diseases, allergies, autoimmune diseases, and transplantation). Citations should be sent to help@iedb.org.

Contact Information

The Immune Epitope Database and Analysis Resource is supported by a contract from the National Institute of Allergy and Infectious Disease, NIH, DHHS (Contract HHSN266200400006C). 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

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

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

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

Production:
Emily Seymour
Ward Fleri, Ph.D.