Review

International Archives of Allergy and Immunology

Int Arch Allergy Immunol 2013;160:334–345
DOI: 10.1159/000343880

Published online: November 21, 2012

Strategies to Query and Display Allergy-Derived Epitope Data from the Immune Epitope Database

Kerrie Vaughan[c] Bjoern Peters[c] Mark Larche[d] Anna Pomes[a] David Broide[b] Alessandro Sette[c]

[a]Indoor Biotechnologies, Charlottesville, Va., and [b]Allergy and Immunology, Department of Medicine, University of California, San Diego, and [c]Immune Epitope Database and Analysis Resource (IEDB), La Jolla Institute for Allergy and Immunology, La Jolla, Calif., USA; [d]Division of Clinical Immunology and Allergy, Department of Medicine, McMaster University, Hamilton, Ont., Canada

Key Words
Allergy · Allergen · Epitope · Database

Abstract
The recognition of specific epitopes on allergens by antibodies and T cells is a key element in allergic processes. Analysis of epitope data may be of interest for basic immunopathology or for potential application in diagnostics or immunotherapy. The Immune Epitope Database (IEDB) is a freely available repository of epitope data from infectious disease agents, as well as epitopes defined for allergy, autoimmunity, and transplantation. The IEDB curates the experiments associated with each epitope and thus provides a variety of different ways to search the data. This review aims to demonstrate the utility of the IEDB and its query strategies, including searching by epitope structure (peptidic/nonpeptidic), by assay methodology, by host, by the allergen itself, or by the organism from which the allergen was derived. Links to tools for visualization of 3-D structures, epitope prediction, and analyses of B and T cell reactivity by host response frequency score are also highlighted.

Introduction

The body of allergen-related immune epitope data has expanded greatly in the past few decades, providing a wealth of data potentially useful for basic and clinical applications. However, given the complexity of these data it can be a daunting task for the individual investigator to pinpoint the research results of greatest interest/relevancy. There are now hundreds of allergy epitope-related references describing thousands of determinants defined for antibodies and T cells in human subjects and animal models. The Immune Epitope Database (IEDB; www.iedb.org) was created by the National Institute of Allergy and Infectious Disease (NIAID) to provide a freely accessible repository of immune epitope data encompassing infectious disease, allergy, autoimmunity, and transplant-related reactivity.

The recognition of specific epitopes on allergens by either T cells or antibodies is a key element of allergic processes [1]. Recognition of epitopes by IgE leads to degranulation of mast cells and basophils and thus directly contributes to immunopathology. Likewise, recognition of specific epitopes by T cells leads to secretion of soluble mediators directly contributing to immunopathology.

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1018–2438/13/1604–0334$38.00/0

Accessible online at:
www.karger.com/iaa

Correspondence to: Dr. Kerrie Vaughan
La Jolla Institute for Allergy and Immunology
Vaccine Discovery
La Jolla, CA 92037 (USA)
E-Mail kvaughan@liai.org

Direct activation of T cells following peptide challenge of asthmatic subjects has been shown to induce symptoms of asthma [2, 3], and indirect contributions by T cells include modulation of IgE responses [1]. Furthermore, induction of regulatory T cells [4–9], immunodeviation [10, 11], and induction of competing IgG responses [12–14] have also been implicated in positive clinical effects following antigen-specific immunotherapy [4, 15–17].

Epitope-specific immunotherapy is an attractive alternative to conventional immunotherapy, since the peptide fragments recognized by T cells would not be likely to bind IgE [18, 19]. Information about specific IgE epitopes can still be applied to generate hypoallergenic molecules by modifying only specific amino acids involved in antibody binding. Mechanisms of allergen-antibody interaction can be evaluated in detail by site-directed mutagenesis followed by antibody binding analysis, especially if mutant design is based on X-ray crystallographic studies of the interactions [20–28].

In addition, the identification of cross-reacting IgG or IgE epitopes in homologous molecules may be of interest. While there are numerous papers describing cross-reacting linear epitopes, especially among food allergens [29–32], a small number of conformational IgE epitopes has been described [33]. The structural basis of antigenic cross-reactivity can help to predict the potential allergenicity of allergen homologs in different sources, which is of special interest for risk assessment in novel foods [34, 35] to elucidate more fundamental questions about the IgE immune response by using the epitopes as probes of the IgE repertoire [36].

In this light, the study of the specific epitopes associated with the phenomena described above is of significant interest, allowing the tracking, measurement, and characterization of allergic responses to a degree of accuracy and precision unthinkable until a few years ago [37–48]. The IEDB provides a platform upon which immunologists and informaticians alike can access all allergy-related epitope data. This review provides an introduction to the database, presenting strategies for accessing allergy-related data, and highlights several unique features that may be of interest to the community of allergists.

The IEDB Is a Source of Epitope-Related Data for the Scientific Community

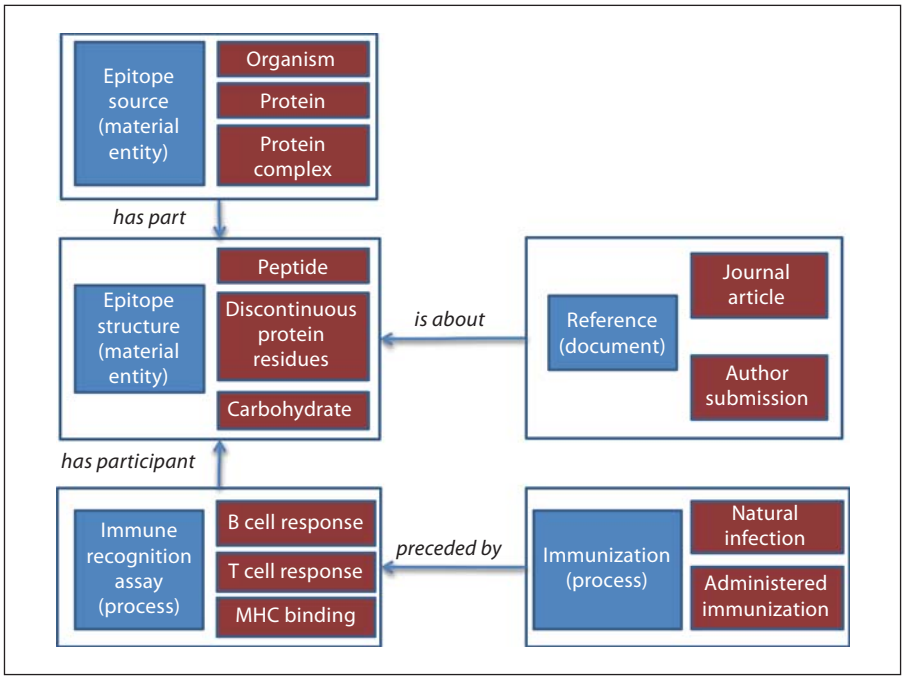
The IEDB is a repository of information relating to all immune epitope data reported in the literature representing infectious diseases (excluding HIV), transplantation, autoimmunity, and allergies. The resource is freely accessible and contains epitopes recognized by humans, nonhuman primates, rodents, and all animal species in which defined epitopes have been reported in the literature. In addition to literature reports, the IEDB also houses direct submissions from various NIH contracts and resources, which represent about a third of the data curated in the IEDB. Data in the IEDB are captured by MD and PhD level scientists experts in the fields of immunology, microbiology, and biochemistry. The IEDB is updated by quarterly queries to PubMed and the date of the last update is shown on the home page.

The IEDB captures both positive and negative data. The only caveat is that negative data are not curated when the effector receptor (T cell or B cell/antibody) tested is monoclonal and there is positive data for the same receptor available. This is due to the implication that monoclonal receptors are highly specific. Negative data generated from pools of peptides are also curated to help define nonreactive regions.

A fundamental issue at the very root of the establishment and curation (or populating) of an epitope database is how to define or present the most interesting, relevant, or prominent features associated with a given epitope. We find that opinions on this topic differ considerably as a function of different users and scientists in general and as it relates to allergy in particular. Some authors might only consider relevant epitopes related to recognition of IgE, while others might be focused on T cell responses, and different users might have very specific views and needs in terms of what assays are most informative, relevant, or even acceptable.

To resolve this issue the IEDB developed an assay-centric approach, where no value judgment is made on which epitopes and epitope features are interesting. Rather, the database captures the associated experimental data in its entirety for each identified epitope. This approach preserves the richness of the data and empowers the individual users and the scientific communities at large to choose which details are relevant. In this way, if the author feels that records defining IgE epitopes in humans with clinical allergy are of greatest significance, then only those records can be queried. Similarly, the author can specify that the search return only those epitopes defined using a specific technique (e.g. X-ray crystallography or the passive cutaneous anaphylaxis test).

The high level structure of the database is shown in figure 1. The database captures for each epitope its chemical structure, whether peptidic or nonpeptidic (carbohydrate, hapten, etc.), associated with a linear sequence or



Color version available online

Fig. 1. High-level ontology schema. Data within the IEDB are organized according to the epitope structure (sequence), type (peptide, carbohydrate, etc.), and source (antigen and organism), the immunological context in which it was defined (antibody response, T cell response, or MHC binding), and the immunization process described for the host (natural or administered), and all of this is linked to the reference of origin.

with a discontinuous (conformational) structure. In addition, the database fields capture the source organism from which the epitope originated (e.g. peanuts, *Arachis hypogaea*), the specific allergen (e.g. Ara h 1) from which it was derived, and an NCBI accession number link representing that protein.

A separate series of fields captures the host organism in which the epitope-specific immune responses were observed and details the specific immunization or in vivo disease processes that have led to the elicitation of the responses. Finally, the database captures, in each instance, the specific assays that were utilized by the investigators to detect the responses (e.g. ELISA, neutralization, ELISPOT) and on which indeed the existence of the epitope is predicated.

Different Ways to Query the IEDB for Allergy Epitopes

As of March 2012, the IEDB contains data related to recognition of 5,238 allergy-related immune epitopes recognized in 68 different hosts (humans, nonhuman primates, numerous mouse strains, rabbits, etc.) and derived from nearly 700 different peptidic and nonpeptidic allergens representing 146 organisms. To allow access to this large volume of detailed data and to match different specific user needs, multiple query strategies have been implemented and new strategies are constantly being developed. As mentioned above, users can currently query the IEDB by the structure of specific epitopes (using amino acid sequences), as a function of specific assays of interest (e.g. histamine release), as a function of the host organism (e.g. humans or Tg mice) or the allergen organism (e.g. birch), or as a function of a specific allergen molecule (e.g. Bet v 1). The user can also focus broadly on all immune response types (B cell, T cell, MHC binding and ligand elution), or alternatively specify that only T cell or antibody data be retrieved. The effector cell phenotype is searchable down to the allele for CD4+ and CD8+ (including nonclassical molecules like CD1d), and down to the specific isotype for immunoglobulins (IgG1, IgG2a, IgG2b, IgM, etc.), if reported by the authors. It is also possible using the 'advanced search' from the pull-down menu on the home page to search by 'effector cell type', and this includes 'T cell γδ'. The IEDB does not capture data related to innate immune receptors (e.g. NK cells).

Searching by Epitope Sequence or Structure

The user can specify any particular continuous or discontinuous peptide sequence and the query will return all records matching the specified sequence. For example, if the user were to enter the sequence 'YDTYKCIPSLEAAVK' from the Phl p 5 allergen, in the results sum-

ID	Reference	Epitope	Host	Immunization	Assay Antigen	Antigen Epitope Relation	MHC Restriction	Assay Description
1460448	W D Müller; Clin Exp Allergy 1998	YDTYKCIPSLEAAVK Pollen allergen Phl p 5b precursor (200-214) Phleum pratense	Homo sapiens	Allergy to Pollen from Phleum pratense (Derivative of Source Organism) followed by restimulation in vitro	YDTYKCIPSLEAAVK Pollen allergen Phl p 5b precursor (200-214) Phleum pratense	Epitope	HLA-DR	3H-thymidine cell proliferation **Positive**
1460449	W D Müller; Clin Exp Allergy 1998	YDTYKCIPSLEAAVK Pollen allergen Phl p 5b precursor (200-214) Phleum pratense	Homo sapiens	Allergy to Pollen from Phleum pratense (Derivative of Source Organism) followed by restimulation in vitro	Pollen allergen Phl p 5b precursor Pollen allergen Phl p 5b precursor Phleum pratense	Source Antigen	HLA-DR	3H-thymidine cell proliferation **Positive**
1460450	W D Müller; Clin Exp Allergy 1998	YDTYKCIPSLEAAVK Pollen allergen Phl p 5b precursor (200-214) Phleum pratense	Homo sapiens	Allergy to Pollen from Phleum pratense (Derivative of Source Organism) followed by restimulation in vitro	Pollen allergen Phl p 5a Pollen allergen Phl p 5a Phleum pratense	Other Structure from Source Organism	HLA-DR	3H-thymidine cell proliferation **Positive**
1460451	W D Müller; Clin Exp Allergy 1998	YDTYKCIPSLEAAVK Pollen allergen Phl p 5b precursor (200-214) Phleum pratense	Homo sapiens	Allergy to Pollen from Phleum pratense (Derivative of Source Organism) followed by restimulation in vitro	Pollen from Phleum pratense Phleum pratense	Derivative of Source Organism	HLA-DR	ELISA cytokine release IL-4 **Positive**
1460452	W D Müller; Clin Exp Allergy 1998	YDTYKCIPSLEAAVK Pollen allergen Phl p 5b precursor (200-214) Phleum pratense	Homo sapiens	Allergy to Pollen from Phleum pratense (Derivative of Source Organism) followed by restimulation in vitro	Pollen from Phleum pratense Phleum pratense	Derivative of Source Organism	HLA-DR	ELISA cytokine release IFNg **Positive**
1460453	W D Müller; Clin Exp Allergy 1998	YDTYKCIPSLEAAVK Pollen allergen Phl p 5b precursor (200-214) Phleum pratense	Homo sapiens	Allergy to Pollen from Phleum pratense (Derivative of Source Organism) followed by restimulation in vitro	Pollen from Phleum pratense Phleum pratense	Derivative of Source Organism	HLA-DR	ELISA cytokine release IL-13 **Positive**

2a

Color version available online

(For legend see next page.)

mary table we would see a total of 25 different curated T cell responses (some associated with positive results, some with negative results). Clicking on the '11' positive responses would display the results shown in figure 2a (the figure only shows a portion of the results table).

However, if at the level of epitope structure the user were to revise this search by selecting the option to return sequences with 70% homology with the same sequence, a different result would be obtained. Indeed, the results summary table would show a total of 26 different positive peptides retrieved by the search. Clicking on the '26' positive peptides shows the results listed in figure 2b. Many of the sequences are derived from the same allergen sequence, and overlap with the input sequence, but have different starting and end points. Interestingly, however, the search also reveals several additional epitopes derived from isoallergen sequences (Phl p 5a vs. 5b), and even highly homologous sequences independently described as epitopes in other allergens such as *Lolium perenne* (Lol p VA). In some cases (not shown here), the search can also reveal additional similar epitopes encoded in other antigens not related to any allergens, such as those derived from microbes, which may be of interest as potential targets for cross-reactive responses.

As stated above, the IEDB is capable of capturing discontinuous (nonlinear) epitopes defined for both peptidic and nonpeptidic allergens. In the first sections of the home page search there is a radio button, 'discontinuous

Epitope ID	Structure	Source Antigen	Source Organism
4792	ASTGGAYQSYKFIPALEAAV	Pollen allergen KBG 60 precursor	Poa pratensis
14609	ETYKFIPSLEAA	Pollen allergen Lol p VA precursor	Lolium perenne
18839	GAYETYKFIPSLEAAVKQAY	Pollen allergen Lol p VA precursor	Lolium perenne
28065	IPSLEAAVKQAYAATVAAAPQ	Pollen allergen Phl p 5b precursor	Phleum pratense
30748	KFIPALEAAVKQSYAATVAT	Pollen allergen KBG 60 precursor	Poa pratensis
30868	KFTVFEGAFNKAIKESTGGAYEAYKFIPSLETAVK	group V allergen	Holcus lanatus
44461	NKALNECTGGAYETYKFIPS	Pollen allergen Lol p VA precursor	Lolium perenne
49417	PSLEAAVKQAYAATVAAAPE	Pollen allergen Lol p VA precursor	Lolium perenne
73596	YDTYKCIPSLEAAVK	Group V allergen Phl p 5.0203 precursor (1 more)	Phleum pratense
73597	YDTYKCIPSLEAAVKQAYAATVAAAPQ	Pollen allergen Phl p 5b precursor	Phleum pratense
74450	YKFIPALEAAVK	Pollen allergen Phl p 5a	Phleum pratense
113139	QSYKFIPALE	Pollen allergen KBG 60	Poa pratensis
121968	VFTPRPPDNYKVIPS		
125764	AYESYKFIPALEAAV	Group V allergen Phl p 5.0103 precursor	Phleum pratense
125778	CIPSLEAAVKQAYAA	Group V allergen Phl p 5.0203 precursor	Phleum pratense
126392	KFIPALEAAVKQAYA	Group V allergen Phl p 5.0103 precursor	Phleum pratense
126535	LEAAVKQAYAATVAT	Group V allergen Phl p 5.0103 precursor	Phleum pratense
127142	STGGAYDTYKCIPSL	Group V allergen Phl p 5.0203 precursor	Phleum pratense
168246	AYDTYKSIPSLEAAV	pollen allergen Poa p 5	Poa pratensis
168380	DSYKFIPTLVAAVKQ	Major pollen allergen Lol p 5b	Lolium perenne
168503	ESTGGAYDTYKSIPS	pollen allergen Poa p 5	Poa pratensis
168623	GAYETYKFIPSLEAA	major allergen Pha a 5 isoform	Phalaris aquatica
169082	LEAAVKQAYAATIAA	pollen allergen Poa p 5	Poa pratensis
170098	YKFIPSLEAAVKQAY	major allergen Pha a 5 isoform	Phalaris aquatica
174415	CIPSLEAAVKQAYAATVAAA	pollen allergen PhlpVb	Phleum pratense

Color version available online

2b

Fig. 2. a Searching by epitope sequence or structure. Presented is the T cell response assay table generated following a query using a specified peptide sequence from Phl p 5. **b** Search by 70% sequence homology. List of additional epitopes returned in the results summary table following a query using the same Phl p 5 peptide sequence and a BLAST stringency of 70%.

peptide', that can be selected. The user can then specify the allergen of interest and perform a search (or add any number of other criteria). If the authors provide the data (not all papers do) we can and do capture residues reported from two different chains. There are currently 44 epitopes in the IEDB with multichain epitopes, e.g. epitope ID 98668 from Bla g 2, *Blattella germanica*.

Querying the IEDB for Nonpeptidic Allergen Epitopes

Recognition of nonpeptidic epitopes may play an important part in allergic disease, particularly in the case of food allergy, drug hypersensitivity, and contact dermatitis. Perhaps the best example of this is IgE antibodies specific for galactose-alpha-1,3-galactose [49]. The IEDB curates (or captures) all nonpeptidic epitopes and the specific strategies to curate and display nonpeptidic epitopes have been recently described [50]. The novel approach entails collaboration with Chemical Entities of Biological Interest (ChEBI) [www.ebi.ac.uk/chebi] to enable the curation of nonpeptidic epitopes, in the context of a rigorous yet flexible and information-rich ontology.

Several options are available to the user in terms of searching for nonpeptidic allergen epitopes. These include entering the exact name of the molecule and utilizing the molecule finders, which will retrieve the molecule name and common synonyms in an automatic auto-fill fashion. For example, we can use the nonpeptidic molecule finder to search for penicillin. Penicillins, like many other nonpeptidic allergens, are thought to act as haptens and bind to proteins in vivo, which results in the immune response. In the clinical setting, skin testing against penicillin is performed using pre-pen which is penicillin, linked to polylysine (simulating hapten + carrier protein). Executing this query will return 15 different molecules, including amoxicillin as one of these structures, as it is in the same chemical family. Selecting records relating to amoxicillin alone returns a total of 87 experimental (10 T and 77 B cell) assays derived from 17 different references. The query results include the chemical structure, drawn according to chemical convention, and other useful data such as detailed reference information and access to all experimental data (B and T cell assays), as well as links to ChEBI through the accession number.

Perhaps the most useful and flexible query opportunity is represented by the option to utilize the nonpeptidic structure tree, which reflects the ChEBI classification of chemical structures. Selecting the 'highlight in tree' option allows visualization of the relation of amoxicillin to the other penicillins, and the user can select the level of generality of the query (fig. 3). Clicking on the

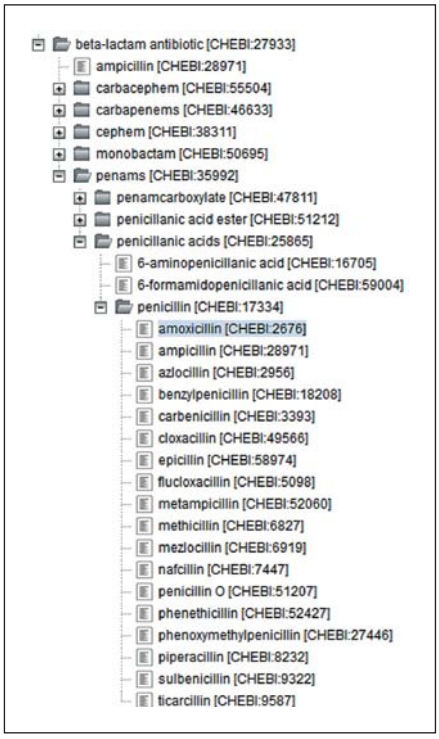


Fig. 3. Nonpeptidic structure tree. The nonpeptidic structure tree is organized according to chemical type and allows the user to visualize the big picture of chemical compound groupings and select higher nodes for inclusiveness.

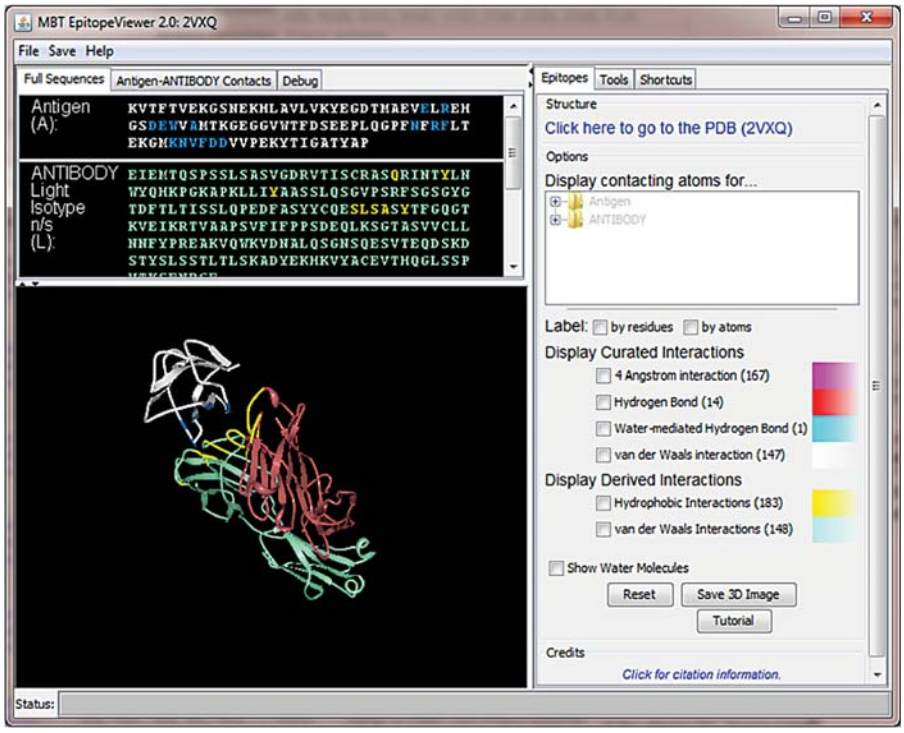


Fig. 4. Conformational Phl p 2 epitope in complex with IgE. 3-D representation of a conformational Phl p 2 epitope in complex with the human IgE monoclonal antibody (HuMab2) using EpitopeViewer.

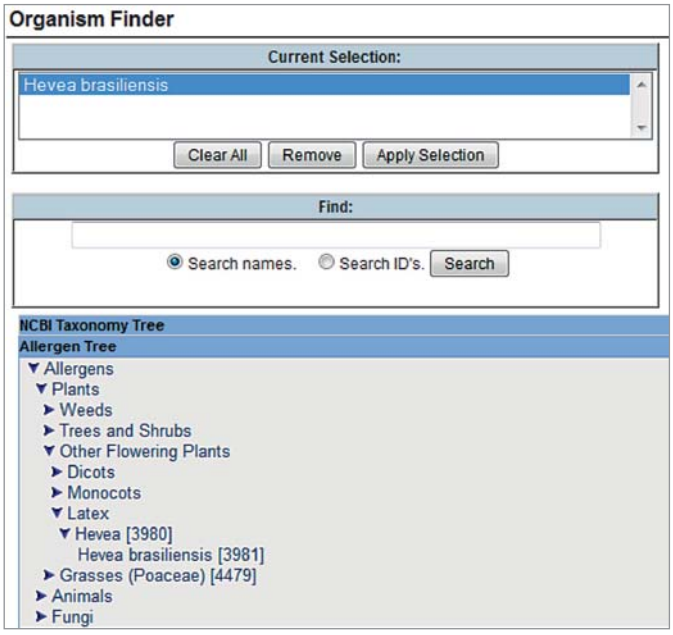
higher node of 'penicillin' will simultaneously return results relating to 19 different structures, such as flucloxacillin, ampicillin, and nafcillin, for a total of over 379 T and B cell response assays derived from 55 different references.

Searching for Results Associated with a Particular Assay Methodology

The user can also search the IEDB so that the results will contain only the assays or method(s) specified. We believe that this is one of the most powerful features of the overall design of the database, which takes advantage of the assay-centric structure of the IEDB. This option addresses one of the most common concerns of investigators, namely the desire to search and include only assays that are performed using the experimental methodology deemed most relevant for that investigator's purposes.

For example, the user can select and specify 'IgE-mediated histamine release' as an assay system. In this case the search will return only those epitopes assayed with that specific methodology. To further enhance this search option, specific assay finders and assay ontologies have been developed. This allows the user to select from an assay tree hierarchically organized as a function of their relatedness to each other, enabling searches of the desired level of stringency in terms of assay methodology.

Furthermore, the assay methodology specified as a search criteria can also be combined with other search criteria, such as selecting a specific allergen source or a given host organism. The combination of different search parameters is illustrated by an example in which the IEDB was queried for X-ray crystallography data related to IgE. This can be done by choosing the advanced B cell response search option and selecting both IgE as the assay antibody isotype and X-ray crystallography as the method/technique. Among the results reported by such a query is a record relating to a conformational Phl p 2 epitope [51]. The structure shown in figure 4 is this epitope in complex with the human IgE Fab (HuMab2) and is visualized using [52] the IEDB's Epitope Viewer. This tool al-



Color version available online

Fig. 5. Allergen tree. The tree is based on higher order categories such as plants, animals and fungi representing the biological entity from which the allergen was derived. Each high level category has multiple subcategories. In plants, for example, subcategories include weeds, trees and shrubs, other flowering plants, and grasses. These are then further organized by taxonomic families (e.g. Urticaceae or the nettle family).

lows the users to map conformational epitopes to the 3-D structure of the protein taking advantage of the linkage between the IEDB and the Protein Data Bank (PDB) (www.rcsb.org/pdb/home/home.do). Specific features within the tool allow the user to highlight epitope residues engaged by the immunoglobulin and, conversely, the specific residues on the antibody that contact the epitope (paratope). Further options include the ability to manipulate, label, and export the image.

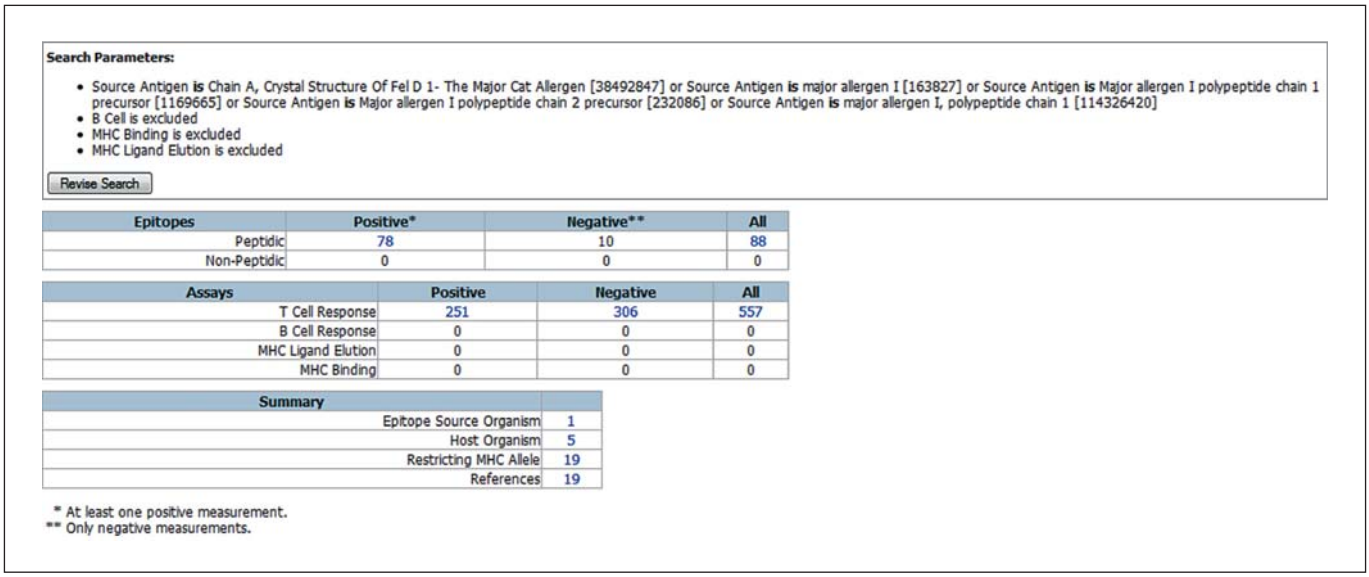
The Search by Organism: Taxonomy or Allergen Classification?

When the IEDB project was initiated, curation of epitopes derived from class A-C priority pathogens and microbes of biodefense concern, and then all other infectious agents in general, was assigned first priority. The most immediate and intuitive strategy to search for epitopes derived from infectious agents was to take advantage of the preexisting and independently maintained NCBI taxonomy. Accordingly, the user can enter a specific name in the organism finder and the system will provide the exact taxonomical denomination and NCBI taxonomy synonyms. As described above, a taxonomy tree allows the user to select more or less general searches (such as strain-, species-, genus-, or family-specific searches). Indeed the NCBI taxonomy is utilized as the default setting for organism searches in the IEDB.

However, this arrangement is not optimal in the case of allergens, as the most commonly utilized allergen classification is only partially overlapping with the NCBI taxonomy and utilizes concepts such as 'weeds' or 'grasses' which are not utilized in the context of the broader NCBI taxonomy, whose use would also be too detailed and appear unfamiliar to the general allergist. Furthermore the generally accepted nomenclature utilized by allergist is the World Health Organization/International Union of Immunological Societies (WHO/IUIS) allergen nomenclature (www.allergen.org) which also only partially overlaps with the National Center for Biotechnology Information (NCBI) nomenclature. For this reason we have developed an alternative way to search for allergen organisms, which is termed 'allergen tree' (fig. 5), and is accessible in the Epitope Source organism finder as an alternative to the NCBI tree. The allergen tree is designed to mirror the taxonomic classification most used by allergists, based on higher order categories such as plants, animals, and fungi representing the biological entity from which the allergen was derived. Each high level category has multiple subcategories. In plants, for example, subcategories include weeds, trees and shrubs, other flowering plants, and grasses. These are then further organized by taxonomic families (e.g. Urticaceae or the nettle family).

The Many Names of a Protein Allergen: Searching and Finding

A common search for allergen epitopes might not necessarily be as narrow as being centered on a given epitope sequence or structure, or as broad as encompassing the totality of all determinants for a given allergen. Nonetheless, indeed, many searches entail querying for data relating to a given protein. However, the issue is less trivial than it might at first appear, as different authors might utilize different accession numbers which sometimes represent the same protein (as the same exact sequence can be represented by multiple GenBank accession numbers). In addition there may be minor allelic variants of the same protein, sequences that have variations because of sequencing errors but also indeed other times represent true isoforms or related highly homologous proteins. To make matters worse the same exact protein can be referred to by several different synonyms. The published nomenclature does not always use WHO/IUIS conven-



Search Parameters:

- Source Antigen **is** Chain A, Crystal Structure Of Fel D 1- The Major Cat Allergen [38492847] or Source Antigen **is** major allergen I [163827] or Source Antigen **is** Major allergen I polypeptide chain 1 precursor [1169665] or Source Antigen **is** Major allergen I polypeptide chain 2 precursor [232086] or Source Antigen **is** major allergen I, polypeptide chain 1 [114326420]
- B Cell is excluded
- MHC Binding is excluded
- MHC Ligand Elution is excluded

Revise Search

Epitopes	Positive*	Negative**	All
Peptidic	78	10	88
Non-Peptidic	0	0	0

Assays	Positive	Negative	All
T Cell Response	251	306	557
B Cell Response	0	0	0
MHC Ligand Elution	0	0	0
MHC Binding	0	0	0

Summary	
Epitope Source Organism	1
Host Organism	5
Restricting MHC Allele	19
References	19

* At least one positive measurement.
** Only negative measurements.

Fig. 6. Fel d 1 from *F. domesticus* using all synonyms. Presented is the results summary table generated following a search using all listed synonyms of the major cat dander allergen Fel d 1.

tion, and the WHO/IUIS itself, while constantly revised, does not cover 100% of the proteins identified in the literature as being associated with allergic immune reactions [www.allergen.org; 53].

By way of example, let us assume a user is interested in retrieving epitope data relating to the common cat allergen Fel d 1. Entering in the molecular finder field the molecule accession number 163827 returns the 'major allergen I' of *Felis catus* (the common domestic cat), according to the GenBank name associated with that particular accession number. The scientific name for this species used by the WHO/IUIS allergen nomenclature is *Felis domesticus,* from which the allergen names are derived (Fel d 1 to Fel d 8). A search of the associated epitope data reveals 5 different epitopes and a total of 9 positive T cell responses. If, however, the 'highlight in tree' option is selected, the tree (according to NCBI/ GenBank taxonomy) of protein records derived from *F. catus* is shown. This reveals that a number of different records relating to the Fel d 1 antigen exist, in which the authors have utilized different synonyms, namely chain A, crystal structure of Fel D 1 – the major cat allergen, major allergen I, major allergen I polypeptide chain 1 precursor, or major allergen I polypeptide chain 2 precursor. The user can now select all of these synonyms, and in this case the search will return a total of 78 epitopes, described in 251 assays and 19 references (fig. 6). While the IEDB is currently compelled to utilize the nomenclature presented by the NCBI protein databank, future plans are to integrate the WHO/IUIS nomenclature into our allergy tree and provide a link to the IUIS allergen nomenclature tree view.

Alternatively, the user can search using the 'browse' feature. Browse by source organism is a quick and convenient way to hunt for specific data related to a certain allergy-associated plant, animal, or fungus using common names. In the search field, one can simply type in 'peanut' or 'house dust mite' or 'cow' to quickly retrieve tab delineated results starting with a list of all peptides and protein names derived from that organism. Clicking on the epitope ID enables access to reference and assay details. In this way, the user can get a preview of related data and/or drill down into the details of any given epitope derivation. Using the above examples, choosing 'peanut' will list epitopes from Ara h 1 and seed storage protein SSP2, 'house dust mite' will list peptides from Der f 2 or Der p 2, and 'cow' will show epitopes derived from cow's milk allergens such as casein and β-lactoglobulin.

Looking into the Future: New Features in Development

As alluded to above, new ways to search for epitopes and new strategies for displaying the results are constantly being considered. Two of these features are currently in development and are expected to be deployed

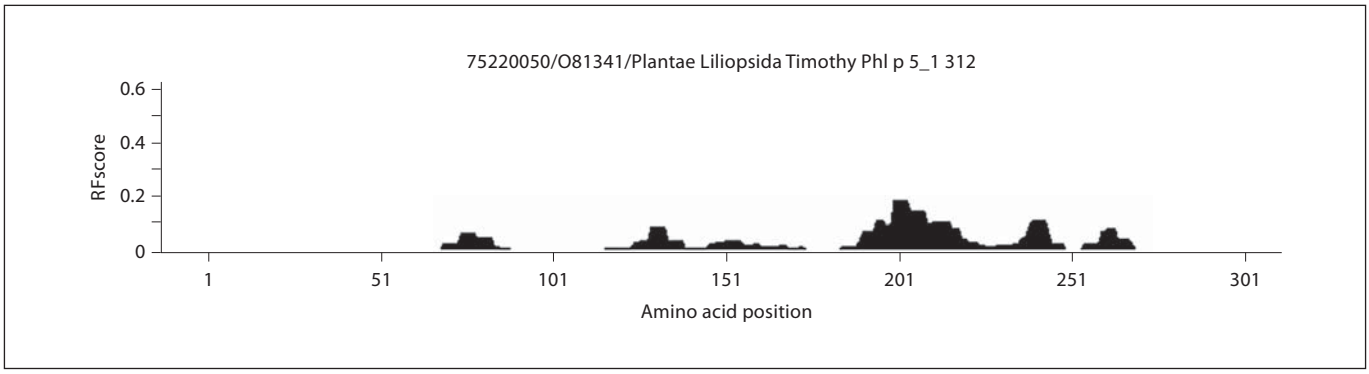


Fig. 7. RFscore plot for the common allergen Phl p 5. RFscores are displayed along the entire length of the Phl p 5 protein from Timothy grass *(Phleum pretense).*

within the 2012 calendar year. These include the expansion of the disease tree and the implementation of a protein browser.

The disease tree will allow results to be searchable and displayed by virtue of association of a particular sequence with a specific clinically diagnosed allergic condition. This feature complements the allergen tree, which currently allows a user to search for allergic epitopes as a function of the allergen source, or in this context the disease trigger. Specifically, in the disease tree the allergic disease is classified by anatomical location of the main symptoms, such as skin and connective tissue (contact dermatitis), the gastrointestinal tract (food allergy), and the respiratory tract (e.g. rhinitis, asthma). This feature will discriminate between records associated with a specific physiological diagnosis (reported in the patient history) and records associated with a specific trigger (e.g. drug hypersensitivity). A trial version of the disease tree and finder is available on the IEDB and allows the user to select the broader category of 'allergy'.

The second feature will allow visualization of all records relating to a given protein in a plot that summarizes the frequency or prevalence of immune reactivity per residue. These response frequency scores (RFscores) are calculated on the basis of the number of individuals tested and the number of individuals in which this test was positive weighted by the sample size. This feature enables the user to combine, in a single plot, all data related to a given allergen, visualizing the composite knowledge (all RFscores for a specified region) relating to epitope reactivity. This approach coalesces all data, including partially overlapping peptides spanning a sequence or to different truncated versions of the same epitope region, and also factors in how extensively each region has been studied and therefore how reliable the combined response frequency data is. A similar approach has been recently reported for HCV epitope records, where it revealed unexpected correlations between the structure and sequence variability of epitopes [54]. An example of this type of plot for the common allergen Phl p 5 is shown in figure 7. These results show three regions of CD4+ T cell reactivity in humans. The plot shows three regions of overall activity: a larger region towards the C-terminal end between aa180 and aa260, which includes a dominant domain at residues 190–210 (peak RFscores of 0.2 or 20%), and two less prominent regions at aa60–90 and aa120–170.

Discussion

In the setting of allergic diseases, mapping epitopes recognized by IgE provides insight into the nature of allergen-IgE interactions and into what makes a given protein allergenic. Furthermore, definition of IgE epitopes might be of therapeutic interest in the context of the generation of hypoallergenic antigens to be used in desensitization protocols and may help to predict the allergenicity of cross-reactive molecules and contribute to a better understanding of the IgE immune response [33, 36, 51, 55–60]. Likewise, T cell epitopes can be used to measure responses and gain information on the mechanisms involved in allergic reactions, since T cell responses directly contribute to immunopathology [2, 3, 39–44], and are thought to indirectly modulate IgE responses [1]. T cell epitopes also and have potential immunotherapeutic applications because the short peptides recognized by T

cells are less likely to cross-link allergen-specific IgE [45, 47, 61].

The IEDB is a freely available repository of published and user-submitted information relating to immune epitopes from allergens, microbes, transplantation, and auto-antigens. The database captures the experiments associated with each epitope and thus provides a variety of different ways to search the data, including querying by epitope structure (both peptidic and nonpeptidic), by assay methodology, by host, by the allergen itself, or by the organism from which the allergen was derived. In many cases, strategies and query tools specifically designed to facilitate analysis of allergen-derived epitopes have been designed. The IEDB is one of several tools available to the scientific community in terms of analysis and inventory of allergen-related information (like Allergome, AllergenOnline.org, and WHO/IUSI), and the information is complementary to that contained in the databases in that the IEDB provides an immunological context-based format in which to search the known allergen-specific literature. As an example, a recent study emphasized a strong association of HLA class II with asthma and IgE in genome-wide association studies [62]. It may be possible to use the IEDB to investigate the molecular mechanisms associated with this effect using search criteria specifying class II alleles, plus the occurrence of asthma, and then comparing the results with known targets of IgE responses. A previous meta-analysis performed by our group on all allergy-related data highlights the use of the IEDB in deriving disease-specific data from the cumulative data [63].

The IEDB can be used to quickly and conveniently survey the existing knowledge relating to allergen-derived epitopes. These results are relevant for the comparison and characterization of different types of allergen responses, to retrieve information relating to antibody epitopes for generation of hypoallergenic antigens and B cell-targeted vaccines. Data related to T cell epitopes is relevant for the selection of reagents to monitor and quantify allergen-specific T cell responses, and also have potential direct immunotherapeutic applications.

Future plans involve further integration of the IEDB with the official WHO/IUIS allergen nomenclature, and development of intuitive and biologically relevant query and display strategies, such as the disease tree and the protein browser. In this context, users' critique, feedback, and suggestions are key for the continued improvement of the IEDB resource.

Conclusion

The IEDB is a unique resource for accessing cumulative allergy-related epitope data. The assay-centric nature of the IEDB provides tremendous granularity, making it possible to search the data at every level, including epitope sequence, allergen source (peanut), or the organism from which the allergen was derived (e.g. Timothy grass). It is also possible to search for references or authors of interest. The most unique feature of the database is inclusion of details related to the immunization process and experimentation. In this way, the user can access data related to a particular host or a specific clinical condition, select an immunogen/antigen of interest, and then select only those assay types deemed relevant. Moreover, the database is a resource for B and T cell epitope prediction, as well as other analysis tools, including 3-D modeling. The IEDB therefore represents an important tool of potential relevance for researchers working in the allergy field.

Acknowledgments

The La Jolla Institute of Allergy and Immunology is supported by the National Institutes of Health NIAID, contract No. HHSN26620040006C, under the IEDB and Analysis Program and the project for the identification of potentially pathogenic and therapeutic epitopes from common human allergens under contract No. HHSN272200700048C as well as LIAI Epitope Validation Center: Characterization of Allergen-Specific T Cells grant No. 1U19AI100275-01. A.P. was funded by grant R01AI77653-1A2S1 (PI: A.P.) from the National Institute of Allergy and Infectious Diseases.

References

1 Akdis CA: Allergy and hypersensitivity: mechanisms of allergic disease. Curr Opin Immunol 2006;18:718–726.
2 Haselden BM, Kay AB, Larche M: Immunoglobulin E-independent major histocompatibility complex-restricted T cell peptide epitope-induced late asthmatic reactions. J Exp Med 1999;189:1885–1894.
3 Ali FR, Oldfield WL, Higashi N, Larche M, Kay AB: Late asthmatic reactions induced by inhalation of allergen-derived T cell peptides. Am J Respir Crit Care Med 2004;169:20–26.
4 Akdis CA, Akdis M: Mechanisms of allergen-specific immunotherapy. J Allergy Clin Immunol 2011;127:18–27.

5 Bohle B, Kinaciyan T, Gerstmayr M, Radakovics A, Jahn-Schmid B, Ebner C: Sublingual immunotherapy induces IL-10-producing T regulatory cells, allergen-specific T-cell tolerance, and immune deviation. J Allergy Clin Immunol 2007;120:707–713.
6 Verhoef A, Alexander C, Kay AB, Larche M: T cell epitope immunotherapy induces a CD4(+) T cell population with regulatory Activity. PLoS Med 2005;2:e78.
7 Francis JN, Till SJ, Durham SR: Induction of IL-10+CD4+CD25+ T cells by grass pollen immunotherapy. J Allergy Clin Immunol 2003;111:1255–1261.
8 Akdis CA, Blesken T, Akdis M, Wuthrich B, Blaser K: Role of interleukin 10 in specific immunotherapy. J Clin Invest 1998;102:98–106.
9 Bellinghausen I, Metz G, Enk AH, Christmann S, Knop J, Saloga J: Insect venom immunotherapy induces interleukin-10 production and a Th2-to-Th1 shift, and changes surface marker expression in venom-allergic subjects. Eur J Immunol 1997;27:1131–1139.
10 Ebner C, Siemann U, Bohle B, Willheim M, Wiedermann U, Schenk S et al: Immunological changes during specific immunotherapy of grass pollen allergy: reduced lymphoproliferative responses to allergen and shift from TH2 to TH1 in T-cell clones specific for Phl p 1, a major grass pollen allergen. Clin Exp Allergy 1997;27:1007–1015.
11 Varney VA, Hamid QA, Gaga M, Ying S, Jacobson M, Frew AJ, et al: Influence of grass pollen immunotherapy on cellular infiltration and cytokine mRNA expression during allergen-induced late-phase cutaneous responses. J Clin Invest 1993;92:644–651.
12 James LK, Shamji MH, Walker SM, Wilson DR, Wachholz PA, Francis JN et al: Long-term tolerance after allergen immunotherapy is accompanied by selective persistence of blocking antibodies. J Allergy Clin Immunol 2011;127:509–516.
13 Wachholz PA, Soni NK, Till SJ, Durham SR: Inhibition of allergen-IgE binding to B cells by IgG antibodies after grass pollen immunotherapy. J Allergy Clin Immunol 2003; 112:915–922.
14 Levy DA, Lichtenstein LM, Goldstein EO, Ishizaka K: Immunologic and cellular changes accompanying the therapy of pollen allergy. J Clin Invest 1971;50:360–369.
15 Durham SR: Allergen immunotherapy: 100 years on. Clin Exp Allergy 2011;41:1171.
16 Ring J, Gutermuth J: 100 years of hyposensitization: history of allergen-specific immunotherapy (ASIT). Allergy 2011;66:713–724.
17 Larche M: T cell epitope-based allergy vaccines. Curr Top Microbiol Immunol 2011; 352:107–119.
18 Tanabe S: Epitope peptides and immunotherapy. Curr Protein Peptide Sci 2007;8: 109–118.
19 Larche M, Wraith DC: Peptide-based therapeutic vaccines for allergic and autoimmune diseases. Nat Med 2005;11:S69–S76.
20 Glesner J, Wünschmann S, Li M, Gustchina A, Wlodawer A, Himly M, Chapman MD, Pomés A: Mechanisms of allergen-antibody interaction of cockroach allergen Bla g 2 with monoclonal antibodies that inhibit IgE antibody binding. PLoS One 2011;6:e22223.
21 Neudecker P, Lehmann K, Nerkamp J, Haase T, Wangorsch A, Fotisch K et al: Mutational epitope analysis of Pru av 1 and Api g 1, the major allergens of cherry *(Prunus avium)* and celery *(Apium graveolens):* correlating IgE reactivity with three-dimensional structure. Biochem J 2003;376:97–107.
22 Gafvelin G, Parmley S, Neimert-Andersson T, Blank U, Eriksson TL, van HM et al: Hypoallergens for allergen-specific immunotherapy by directed molecular evolution of mite group 2 allergens. J Biol Chem 2007;282: 3778–3787.
23 Flicker S, Steinberger P, Norderhaug L, Sperr WR, Majlesi Y, Valent P, et al: Conversion of grass pollen allergen-specific human IgE into a protective IgG(1) antibody. Eur J Immunol 2002;32:2156–2162.
24 Edlmayr J, Niespodziana K, Linhart B, Focke-Tejkl M, Westritschnig K, Scheiblhofer S, et al: A combination vaccine for allergy and rhinovirus infections based on rhinovirus-derived surface protein VP1 and a non-allergenic peptide of the major timothy grass pollen allergen Phl p 1. J Immunol 2009;182: 6298–6306.
25 Niespodziana K, Focke-Tejkl M, Linhart B, Civaj V, Blatt K, Valent P, et al: A hypoallergenic cat vaccine based on Fel d 1-derived peptides fused to hepatitis B PreS. J Allergy Clin Immunol 2011;127:1562–1570.
26 van Milligen FJ, van 't HW, van den BM, Aalberse RC: IgE epitopes on the cat *(Felis domesticus)* major allergen Fel d I: a study with overlapping synthetic peptides. J Allergy Clin Immunol 1994;93:34–43.
27 Wallmann J, Proell M, Stepanoska T, Hantusch B, Pali-Scholl I, Thalhamer T et al: A mimotope gene encoding the major IgE epitope of allergen Phl p 5 for epitope-specific immunization. Immunol Lett 2009;122:68–75.
28 Holmgren J, Czerkinsky C: Mucosal immunity and vaccines. Nat Med 2005;11:S45–S53.
29 Jenkins JA, Griffiths-Jones S, Shewry PR, Breiteneder H, Mills EN: Structural relatedness of plant food allergens with specific reference to cross-reactive allergens: an in silico analysis. J Allergy Clin Immunol 2005;115: 163–170.
30 Ferreira F, Hawranek T, Gruber P, Wopfner N, Mari A: Allergic cross-reactivity: from gene to the clinic. Allergy 2004;59:243–267.
31 Ayuso R, Reese G, Leong-Kee S, Plante M, Lehrer SB: Molecular basis of arthropod cross-reactivity: IgE-binding cross-reactive epitopes of shrimp, house dust mite and cockroach tropomyosins. Int Arch Allergy Immunol 2002;129:38–48.
32 Scheurer S, Son DY, Boehm M, Karamloo F, Franke S, Hoffmann A, Haustein D, Vieths S: Cross-reactivity and epitope analysis of Pru a 1, the major cherry allergen. Mol Immunol 1999;36:155–167.
33 Pomés A: Relevant B cell epitopes in allergic disease. Int Arch Allergy Immunol 2010;152: 1–11.
34 Goodman RE, Vieths S, Sampson HA, Hill D, Ebisawa M, Taylor SL, van Ree R: Allergenicity assessment of genetically modified crops – what makes sense? Nature Biotech 2008;26:73–81.
35 Aalberse RC: Assessment of allergen cross-reactivity, Clin Mol Allergy 2007;5:2.
36 Aalberse RC, Crameri R: IgE-binding epitopes: a reappraisal. Allergy 2011;66:1261–1274.
37 Van OL, Wambre E, Maillere B, Von HE, Louise A, Balazuc AM, et al: Assessment of Bet v 1-specific CD4+ T cell responses in allergic and nonallergic individuals using MHC class II peptide tetramers. J Immunol 2008;180:4514–4522.
38 Wambre E, DeLong JH, James EA, LaFond RE, Robinson D, Kwok WW: Differentiation stage determines pathologic and protective allergen specific CD4+ T-cell outcomes during specific immunotherapy. J Allergy Clin Immunol 2012;129:544–551.
39 James EA, Kwok WW: Autoreactive CD4(+) T cells in patients with atopic dermatitis. J Allergy Clin Immunol 2011;128:100–101.
40 DeLong JH, Simpson KH, Wambre E, James EA, Robinson D, Kwok WW: Ara h1-reactive T cells in individuals with peanut allergy. J Allergy Clin Immunol 2011;127:1211–1218.
41 Bonvalet M, Wambre E, Moussu H, Horiot S, Kwok WW, Louise A, Ebo D, Hoarau C, Van Overtvelt L, Baron-Bodo V, Moingeon P: Comparison between major histocompatibility complex class II tetramer staining and surface expression of activation markers for the detection of allergen-specific CD4+ T cells. Clin Exp Allergy 2011;41:821–829.
42 Kwok WW, Roti M, Delong JH, Tan V, Wambre E, James EA, Robinson D: Direct ex vivo analysis of allergen-specific CD4+ T cells. J Allergy Clin Immunol 2010;125:1407–1409.
43 Oseroff C, Sidney J, Kotturi MF, Kolla R, Alam R, Broide DH, Wasserman SI, Weiskopf D, McKinney DM, Chung JL, Petersen A, Grey H, Peters B, Sette A: Molecular determinants of T cell epitope recognition to the common Timothy grass allergen. J Immunol 2010;185:943–955.

44 Kinnunen T, Nieminen A, Kwok WW, Närvänen A, Rytkönen-Nissinen M, Saarelainen S, Taivainen A, Virtanen T: Allergen-specific naive and memory CD4+ T cells exhibit functional and phenotypic differences between individuals with or without allergy. Eur J Immunol 2010;40:2460–2469.
45 Kinnunen T, Jutila K, Kwok WW, Rytkönen-Nissinen M, Immonen A, Saarelainen S, Närvänen A, Taivainen A, Virtanen T: Potential of an altered peptide ligand of lipocalin allergen Bos d 2 for peptide immunotherapy. J Allergy Clin Immunol 2007;119:965–972.
46 Macaubas C, Wahlstrom J, Galvão da Silva AP, Forsthuber TG, Sønderstrup G, Kwok WW, DeKruyff RH, Umetsu DT: Allergen-specific MHC class II tetramer+ cells are detectable in allergic, but not in non-allergic, individuals. J Immunol 2006;176:5069–5077.
47 Kinnunen T, Kwok WW, Närvänen A, Rytkönen-Nissinen M, Immonen A, Saarelainen S, Taivainen A, Virtanen T: Immunomodulatory potential of heteroclitic analogs of the dominant T-cell epitope of lipocalin allergen Bos d 2 on specific T cells. Int Immunol 2005;17:1573–1581.
48 Campbell JD, Buckland KF, McMillan SJ, Kearley J, Oldfield WL, Stern LJ et al: Peptide immunotherapy in allergic asthma generates IL-10-dependent immunological tolerance associated with linked epitope suppression. J Exp Med 2009;206:1535–1547.
49 Commins SP, Platts-Mills TA: Anaphylaxis syndromes related to a new mammalian cross-reactive carbohydrate determinant. J Allergy Clin Immunol 2009;124:652–657.
50 Vita R, Peters B, Josephs Z, de Matos P, Ennis M, Turner S, Steinbeck C, Seymour E, Zarebski L, Sette A: A Model for collaborative curation, The IEDB and ChEBI curation of non-peptidic epitopes. Immunome Res 2011;7:1–8.
51 Padavattan et al. Padavattan S, Flicker S, Schirmer T, Madritsch C, Randow S, Reese G, Vieths S, Lupinek C, Ebner C, Valenta R, Markovic-Housley Z: High-affinity IgE recognition of a conformational epitope of the major respiratory allergen Phl p 2 as revealed by X-ray crystallography. J Immunol 2009;182:2141–2151.
52 Beaver JE, Bourne PE, Ponomarenko JV: EpitopeViewer: a Java application for the visualization and analysis of immune epitopes in the Immune Epitope Database and Analysis Resource (IEDB). Immunome Res 2007;3:3.
53 King TP, Hoffman D, Lowenstein H, Marsh DG, Platts-Mills TA, Thomas W: Allergen nomenclature: WHO/IUIS Allergen Nomenclature Subcommittee. Int Arch Allergy Immunol 1994;105:224–233.
54 Kim Y, Vaughan K, Greenbaum J, Peters B, Law M, Sette A: A meta-analysis of the existing knowledge of immunoreactivity against hepatitis C virus (HCV). PLoS One 2012;7:e38028.
55 Niemi M, Jylhä S, Laukkanen ML, Söderlund H, Mäkinen-Kiljunen S, Kallio JM, Hakulinen N, Haahtela T, Takkinen K, Rouvinen J: Molecular interactions between a recombinant IgE antibody and the beta-lactoglobulin allergen. Structure 2007;15:1413–1421.
56 Mirza O, Henriksen A, Ipsen H, Larsen JN, Wissenbach M, Spangfort MD, Gajhede M: Dominant epitopes and allergic cross-reactivity: complex formation between a Fab fragment of a monoclonal murine IgG antibody and the major allergen from birch pollen Bet v 1. J Immunol 2000;165:331–338.
57 Padavattan S, Schirmer T, Schmidt M, Akdis C, Valenta R, Mittermann I, Soldatova L, Slater J, Mueller U, Markovic-Housley Z: Identification of a B-cell epitope of hyaluronidase, a major bee venom allergen, from its crystal structure in complex with a specific Fab. J Mol Biol 2007;368:742–752.
58 Li M, Gustchina A, Alexandratos J, Wlodawer A, Wünschmann S, Kepley CL, Chapman MD, Pomés A: Crystal structure of a dimerized cockroach allergen Bla g 2 complexed with a monoclonal antibody. J Biol Chem 2008;283:22806–22814.
59 Li M, Gustchina A, Glesner J, Wünschmann S, Vailes LD, Chapman MD, Pomés A, Wlodawer A: Carbohydrates contribute to the interactions between cockroach allergen Bla g 2 and a monoclonal antibody. J Immunol 2011;186:333–340.
60 Chruszcz M, Chapman MD, Vailes LD, Stura EA, Saint-Remy JM, Minor W, Pomés A: Crystal structures of mite allergens Der f 1 and Der p 1 reveal differences in surface-exposed residues that may influence antibody binding. J Mol Biol 2009;386:520–530.
61 Worm M, Lee HH, Kleine-Tebbe J, Hafner RP, Laidler P, Healey D, et al: Development and preliminary clinical evaluation of a peptide immunotherapy vaccine for cat allergy. J Allergy Clin Immunol 2011;127:89–97.
62 Vercelli D: Remembrance of things past: HLA genes come back on the allergy stage. Allergy Clin Immunol 2012;129:846–847.
63 Vaughan K, Greenbaum J, Kim Y, Vita R, Chung J, Peters B, Broide D, Goodman R, Grey H, Sette A: Towards defining molecular determinants recognized by adaptive immunity in allergic disease: an inventory of the available data. J Allergy (Cairo) 2010;2010:628026.