IgE Antibody Assay-Based Allergen Database (CLSI-IL/A-37)

Abstract

There are 8 established allergen databases (IUIS/WHO ANDB, Allergome, AllFam, AllergenOnline, SDAP, COMPARE, IEDB and AllerBase). None of these provide a comprehensive listing of the internationally-accepted nomenclature codes and Linnean system descriptors of the ~1000 extract and molecular allergens used world-wide in diagnostic single-and multi-plex IgE antibody assays. The new Diagnostic Allergen Database (DADB) has been prepared by an international scientific committee of the Clinical and Laboratory Standards Institute to specifically provide these data together in a readily-searchable Excel database. Using peanut as an example, the DADB lists the allergens’ unique assay working codes (F13), general category (food), principal IgE antibody assay method(s) where the allergen specificity is available, common name (peanut), taxonomical name (Arachis hypogaea), and the NCBI, NPU, LOINC, Allergome and IUIS/WHO descriptor codes for allergens and their assays.

The DADB benefits manufacturers of allergen-specific IgE assays by providing unambiguous specificity descriptors. With the DADB, allergen extract manufacturers will have clearly defined specificity targets for allergenic product development. Government regulators will have clarity on the specificity of new allergen reagents that are being submitted for clearance. IgE antibody proficiency testing programs (e.g. CAP) will be able to clearly define the IgE antibody specificities being measured in their challenge sera. The clinician will precisely know the allergen specificity of IgE antibody that is measured in their reference laboratory. Finally, the allergic patient will be assured that their allergic sensitization is being assessed in an assay with a well-characterized allergen reagent.

Keywords: Allergen, Nomenclature, Databases, IgE Antibody, Diagnostic Assays, Extracts, Molecules, Molecular Allergology
IgE Antibody Assay-Based Allergen Database (CLSI-IL/A-37)

Review

Since the first recorded systemic allergic reaction thought to be due to a Hymenoptera sting in 5000 BC [1], the allergenic molecule has been and remains the quintessential reagent upon which all diagnostic allergy testing is based. By 1880, allergen extracts had become integral to diagnostic allergy testing in the skin. In 1967, allergen extracts were insubilized on allergosorbents and used in immunosassays for the measurement of specific IgE antibodies [2]. Physiological extracts of allergenic substances remained essential reagents for years in skin testing, IgE antibody assays and in in vivo provocation tests. By 1988-1991, the first allergens had been cloned, ushering in the era of molecular Allergyology [3]. Highly characterized molecular allergens, especially for foods such as peanut and hazelnut, have become essential not only in verifying sensitization (IgE antibody presence) but also for assessing relative risk for a severe allergic reaction [4]. Thus, the allergen remains the essential reagent in diagnostic tests of sensitization that are performed to support the clinical history-based diagnosis of human allergic disease. It is the clinician, however, who begins the diagnostic investigation of allergic sensitization with a careful clinical history that aims to identify an allergic symptom which associates temporally with exposure to a particular allergen. During this diagnostic investigation, a compendium of all the allergen specificities available diagnostically becomes invaluable to the clinician.

Established Allergen Databases

There have been a number of excellent reviews of allergen databases published from 2013 to 2019 [5-10]. Of these, Radauer and Breiteneder [9,10] provide the most comprehensive and current overview of the established allergen databases. Table 1 summarizes their names, general content, managing group, website and date of establishment. Each of these databases provides its own unique information about allergens for the clinician, researcher, government regulator, diagnostic reagent manufacturer and patient. The grandfather of all allergen databases, the Allergen Nomenclature Database (ANDB), was established in 2000 by the Allergen Nomenclature Sub-Committee (ANSC) of the World Health Organization and International Union of Immunological Societies (WHO/IUIS) Nomenclature Committee [11,12]. A novel allergenic molecule is first identified in a physiological extract using a variety of analytical and immunocchemical methods. Its associated structure, sequence data, genetic code and IgE binding properties are then submitted to the ANSC of the WHO which performs a peer review. If accepted as novel, the allergenic molecule is given a name based on the Linnean system and listed in the database. As of 2021, the ANDB contains 1036 allergens, 106 of which had been submitted in the last three years [12]. This database covers only molecular allergens and does not list all the available allergenic specificities that are provided as extracts and used in worldwide in IgE antibody assays.

While the ANDB is considered the official resource for molecular allergen designations, there are other established, intersecting and widely overlapping databases that are used to describe known allergens, their structure, genetic code and protein family relationships. These include the Allergome [13], Allergen Family Database (AllFam) [10], AllergenOnline [14], The Structural Database of Allergenic Proteins (SDAP) [15], Comprehensive Protein Allergen Resource (COMPARE) [10], Immune Epitope Database (IEDB) [16], AllerBase [17] Table 1. Interestingly, a number of these databases of allergenic molecules were originally established to provide researchers with the ability to search sequence data and identify potential new protein allergens in genetically-modified crops or to predict cross-reactivity across allergen groups (e.g. the PR10 protein family) as a result of structural homologies [18]. The relationships between these informative and overlapping allergen databases are extensively discussed elsewhere [9,10]. Like the ANDB, these other databases do not list all the available allergenic specificities used worldwide in IgE antibody assays.

New Diagnostic Allergen Database

To address this concern, a new allergen database has been created to serve the diagnostic IgE antibody testing community. The “Diagnostic Allergen Database” (DADB [ILA-37]) Table 1 serves a unique purpose by providing a comprehensive listing of all the allergen specificities as extracts and molecules currently used in diagnostic allergen-specific IgE antibody assays. An international committee composed of academic, industrial and government scientists convened under the auspices of the Clinical and Laboratory Standards Institute (CLSI) to establish this database. This was performed in parallel with another working group which is revising the CLSI ILA/20 reference guide that examines IgE antibody assay technology, analytical performance, calibration methods, degree of quantitation and quality control [19]. As a companion to the ILA/20 IgE antibody assay guidance document, a comprehensive Excel database has been prepared that contains a listing of all the currently available allergenic extracts and molecules that are used to manufacture allergosorbents for the principal allergen-specific IgE antibody serology assays that are used clinically throughout the world [20].

Figure 1 displays an illustrative subset of extract and component entries from the Diagnostic Allergen Database for peanut. The detailed descriptors in the DADB include the individual allergen’s current unique code (eg. F13), category (eg. food), method code (S, B, D, T, C [see descriptors below]), product common name (peanut), taxonomical name (Arachis hypogaea), IUIS Nomenclature code (only for allergenic molecules), and documentation level (“well documented” in the peer reviewed literature or undocumented).

The National Center for Biotechnology Information (NCBI) taxonomy database contains names that are determined to have a correct nomenclature and they are classified in an approximate phylogenetic hierarchy depending on the level of knowledge regarding phylogenetic relationships of a given group. The NCBI code for peanut is (3818). The Allergome code for peanut (1723) is also provided because the Allergome is one of the few databases in Table 1 that provides genus/species data and an extensive link to literature with references for both allergen extracts and allergenic molecules.

To allow the Diagnostic Allergen Database to be interconnected to laboratory descriptors for what is intended to be measured by the assays and by which units, the results are presented in healthcare sys-
IgE Antibody Assay-Based Allergen Database (CLSI-IL/A-37)

The IL/A 37 DADB is available in an Excel file download from the Clinical Laboratory Standards Institute at www.clsi.org

IgE Antibody Assay-Based Allergen Database (CLSI-IL/A-37) systems throughout the world using internationally recognized NPU [21] and LOINC [22,23] descriptor codes for the allergen specificities, and these are cross-referenced in the database. The NPU terminology (Nomenclature for Properties and Units) is a patient centered clinical laboratory terminology created by a joint effort of the IFCC and IUPAC, which enables clinical laboratory results to be used safely across technology, time and geography. The NPU code for peanut is NPU13208, respectively. The LOINC or Logical Observation Identifiers Names and Codes were established by a non-profit medical research organization associated with Indiana University with a similar purpose to the NPU terminology [23]. The LOINC code for peanut is 6206-7.

Table 1: Overview of Allergen Database.

<table>
<thead>
<tr>
<th>Listing</th>
<th>Name</th>
<th>Description</th>
<th>Website-Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allergome [13]</td>
<td>Allergome</td>
<td>comprehensive collection of allergen nomenclature data and references on allergens in extracts or as molecules</td>
<td><a href="http://www.allergome.org">www.allergome.org</a> Allergy Data Laboratories, Latina, Italy; est 2003</td>
</tr>
<tr>
<td>AllFam [10]</td>
<td>Allergen Family Database</td>
<td>classifies allergens into families of evolutionarily related proteins</td>
<td><a href="http://www.meduniwien.ac.at/">www.meduniwien.ac.at/</a> allfam/ Medical University of Vienna, Austria; est. 2007</td>
</tr>
<tr>
<td>AllergenOnline [14]</td>
<td>Food Allergen Research and Resource Program (FARRP) Allergen Protein Database</td>
<td>food allergen searchable DNA sequence data intended to evaluate new proteins in genetically modified crops or novel foods for prediction of allergenicity.</td>
<td><a href="http://www.allergenonline.org">http://www.allergenonline.org</a> University of Nebraska-Lincoln USA est. 2005</td>
</tr>
<tr>
<td>SDAP [15]</td>
<td>The Structural Database of Allergenic Proteins</td>
<td>allergen sequences, structures, and epitopes linked to bioinformatics tools to assist in studies of allergen structural biology, cross-reactivity and predicting the IgE-binding potential of genetically modified food proteins.</td>
<td><a href="https://fermi.utmb.edu/">https://fermi.utmb.edu/</a> SDAP University of Texas Medical Branch, the Department of Biochemistry and Molecular Biology, Galveston, TX USA est 2002</td>
</tr>
<tr>
<td>COMPARE [10]</td>
<td>COMprehensive Protein Allergen RESource</td>
<td>database containing protein sequences of known allergens initially built from the FARRP Allergen Online database v.16; IgE binding (allergenic) proteins are identified via “rule-based” text sorting algorithm; peer review committee based</td>
<td><a href="https://comparedatabase.org">https://comparedatabase.org</a> <a href="https://jifsan.umd.edu/">https://jifsan.umd.edu/</a> Joint Institute for Food Safety and Nutrition University of Maryland and Health and Environmental Sciences Institute (HESI) Washington DC USA est 2017</td>
</tr>
<tr>
<td>IEDB [16]</td>
<td>Immune Epitope Database</td>
<td>repository of T-cell, B-cell, and major histocompatibility complex protein epitopes including epitopes of allergens</td>
<td>iedb.org National Institute for Allergy and Infectious Diseases, Bethesda, MD, USA; est 2006</td>
</tr>
<tr>
<td>AllerBase [17]</td>
<td>Bioinformatics Centre, Savitribai Phule Pune University, India</td>
<td>Compilation of allergen data from various databases and the literature involving experimental data; does not specify inclusion criteria of non-WHO/IUIS listed allergens</td>
<td>bioinfo.net.in/AllerBase/Home.html Bioinformatics Centre, Savitribai Phule Pune University, India</td>
</tr>
</tbody>
</table>
Data extracted from references 9-17.

### Table 1: Diagnostic Allergen Database (DADB)

<table>
<thead>
<tr>
<th>Unique Code</th>
<th>Category</th>
<th>E</th>
<th>D</th>
<th>S</th>
<th>T</th>
<th>Product's common name</th>
<th>Taxoninal Name Used in Product Labeling</th>
<th>Biochemical or Protein Family Name</th>
<th>Current Taxoninal Name</th>
<th>NCII Taxoninal ID</th>
<th>Documentation Category</th>
<th>IUIS Nomenclature</th>
<th>NPU Code (quant)</th>
<th>LOINC Code</th>
<th>Allergen Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>f13</td>
<td>Foods</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Peanut</td>
<td>Arachis hypogaea</td>
<td>Arachis hypogaea</td>
<td>9818</td>
<td>Well documented in literature</td>
<td>NPU13208</td>
<td>6206-7</td>
<td>1728</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f422</td>
<td>Foods</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Ara h 1</td>
<td>Arachis hypogaea</td>
<td>75 globulin, violin</td>
<td>9818</td>
<td>Well documented in literature</td>
<td>Ara h 1</td>
<td>58779-0</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f423</td>
<td>Foods</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Ara h 2</td>
<td>Arachis hypogaea</td>
<td>25 albumin</td>
<td>9818</td>
<td>Well documented in literature</td>
<td>Ara h 2</td>
<td>58778-2</td>
<td>51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f424</td>
<td>Foods</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Ara h 3</td>
<td>Arachis hypogaea</td>
<td>115 globulin, glutenin, legumin</td>
<td>9818</td>
<td>Well documented in literature</td>
<td>Ara h 3</td>
<td>58777-4</td>
<td>52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f447</td>
<td>Foods</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Ara h 4</td>
<td>Arachis hypogaea</td>
<td>25 albumin</td>
<td>9818</td>
<td>Well documented in literature</td>
<td>Ara h 4</td>
<td>90880-6</td>
<td>3998</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f452</td>
<td>Foods</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Ara h 7</td>
<td>Arachis hypogaea</td>
<td>25 albumin</td>
<td>9818</td>
<td>Well documented in literature</td>
<td>Ara h 7</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>f453</td>
<td>Foods</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Ara h 8</td>
<td>Arachis hypogaea</td>
<td>Pathogenesis-related PF 10 (Pf-10)</td>
<td>9818</td>
<td>Well documented in literature</td>
<td>Ara h 8</td>
<td>65477-4</td>
<td>1215</td>
<td></td>
<td></td>
</tr>
<tr>
<td>f552</td>
<td>Foods</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>Ara h 9</td>
<td>Arachis hypogaea</td>
<td>Non-secreted lipid transfer protein type 1</td>
<td>9818</td>
<td>Well documented in literature</td>
<td>Ara h 9</td>
<td>64965-7</td>
<td>742</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 1:** Illustrative segment of the Diagnostic Allergen Database (ILA/37) using peanut and peanut components (Ara h 1, 2, 3, 6, 7, 8, 9) as an example of a food allergen extract and available components.

The allergen specificities presented in the Diagnostic Allergen Database are available in at least one immunoassay provided by the principal worldwide IgE antibody testing manufacturers. The methods represented (in order of worldwide market share) include the Thermo Fisher Scientific/Phadia ImmunoCAP (S = allergen attached to a cellulose sponge), the Siemens Healthcare IMMULITE (B = biotinylated allergen attached to an avidin coated bead), the Hycor Biomedical HyTech-88 (D = allergen attached to a cellulose disc), Quidel Corporation MAST (T = allergen attached to a cellulose thread), and the Thermo Fisher Scientific/Phadia ISAC (C = chip). Allergen specificities provided by other assays that are considered laboratory developed tests by regulators and are thus not regulatory laboratory government cleared are not listed, but can be added in the future as they are documented for clinical and research patient testing. The Hycor Biomedical HyTech-88 is currently being replaced by a newly US FDA cleared NovoNordisk Noveos system, but due to need to backlogged priority reviews, FDA's pace of clearance of allergens for this new assay has become glacial. The allergen categories in the DADB include the weed, grass, and tree pollen, mites, microorganisms, epidermals and animal proteins, foods, venom and insects, drugs, miscellaneous, parasites and occupational allergens. The descriptor nomenclature and cross-referenced codes in the DADB are provided for each of the approximately 1000 allergen (extracts and molecular components) that are used for clinical testing in IgE antibody immunoassays.

### Applications

The Diagnostic Allergen Database provides unique benefits to different users depending upon their needs. Manufacturers of allergen-specific IgE antibody assays will benefit greatly by having established international codes that simplify and provide unambiguous clarity for the specificity of their allergosorbent products. They can list their allergen specific IgE measurements with a variety of established nomenclature codes (CLSI and IUIS/WHO ANDB codes for molecules; CLSI, NPU, LOINC and Allergome codes for extracts and molecules) that are understood clearly across international laboratory boundaries. Allergen extract manufacturers will have clearly defined specificity targets for development of new allergenic in vivo (skin test and immunotherapy) products and source extracts for allergosorbent preparation. As new analytical methods such as mass spectrometry are applied to allergen extract characterization and cross-reactivity studies, researchers will be able to perform new clinical validation studies and this will lead to an increased number of molecular allergens being listed in the DADB. As a result, the number of molecular allergens utilized in both clinical and research IgE antibody testing will increase. Government regulators will benefit by ensuring clarity as they review clinical IgE antibody data for new allergen reagents that are being submitted for clearance. IgE antibody proficiency testing programs, such as the College of American Pathologists (CAP) [Diagnostic Allergy
IgE Antibody Assay-Based Allergen Database (CLSI-IL/A-37)

Informatics of Allergen Specificities related to the Database.

References