



How Data are Retrieved, Entered, and Organized

www.iedb.org

Presented by: Randi Vita, M.D. Lead Ontology and Quality Manager

IEDB Scope

- Infectious diseases, Allergy, Autoimmunity, Transplant
 - HIV, cancer, etc only curated as structural data or when presented with above subjects as per NIH/NIAID
- Published literature and direct submissions
- Experimentally confirmed, no predictions, no reviews
- Negative data and supplemental data included
- Binding of an adaptive immune receptor to an epitope (T cell, B cell, MHC binding, MHC ligand elution)
- Must be epitope specific
 - All tested antigens are curated for epitope specific receptors

Minimal Criteria for Epitope Inclusion

- Linear peptide <50 amino acids in length
- Tested as an immunogen or an antigen
- Discontinuous residues shown to be important in recognition
- Non-peptidic epitopes <5000 Daltons
- Minimal information required (sequence, outcome, host, etc)

Literature Curation Process

PubMed / PDB

- Complex query
- Bi-weekly

230K retrieved

Classifier

- Content based categories
- Retrained annually

147K epitope related

Infectious Disease

46,898 refs

Allergy

9,042 refs

Autoimmunity

18,894 refs

Transplantation

13,968 refs

Other

58,971

Abstract Review

- Manual scan
- Confirmation of classification

36K likely curatable

Manual Curation

- Assigned to curators
- Peer review

20K curated

Infectious Disease

10,799 refs

Allergy

1,587 refs

Autoimmunity

5,017 refs

Transplantation

841 refs

Other

1,981 refs

Consistency and quality control measures

- Manually curated by a team of PhD level scientists with specific expertise
- Formal curation guidelines and peer review
- External immunological experts
- Built in validation in the curation application

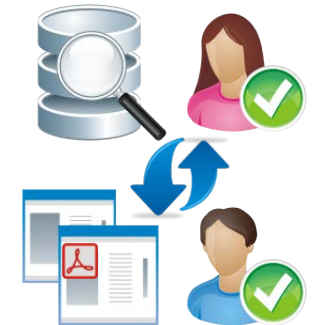
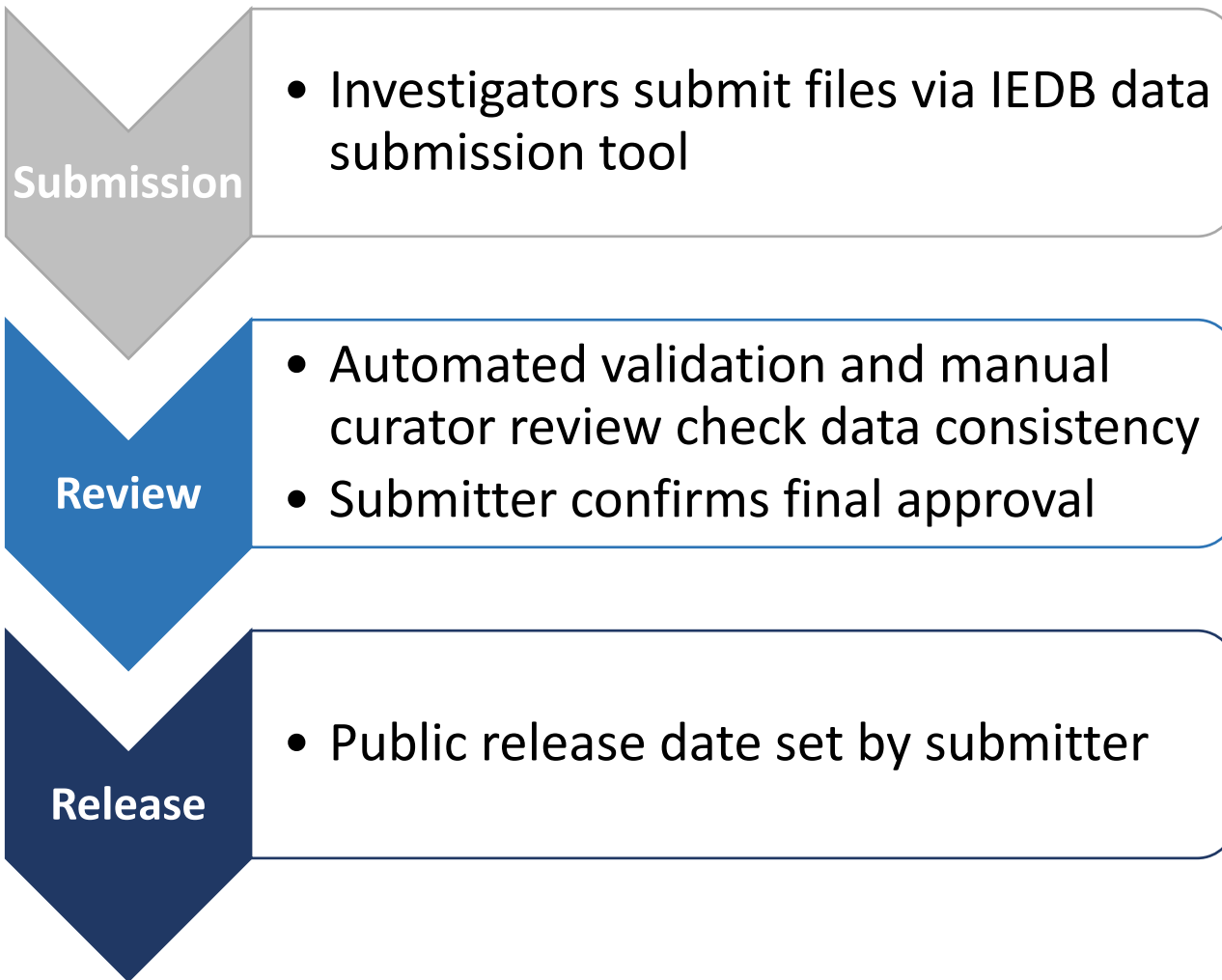
Submission Community

- Primarily Epitope Discovery Contract Holders

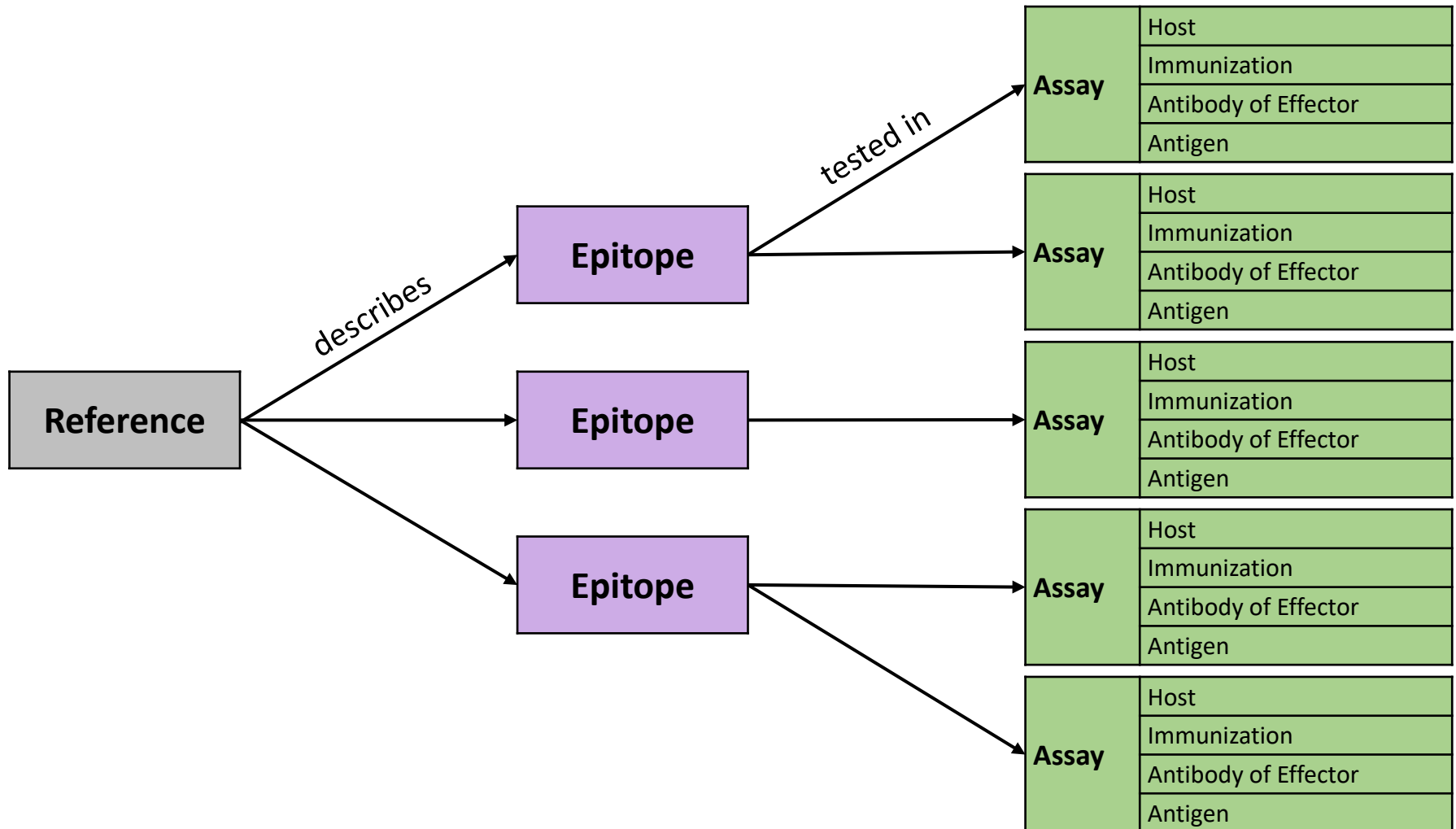
Past & Present:

- P. vivax, P. falciparum, Dengue, JEV, RVFV, Oropouche virus, Influenza, Ricin, Ebola, HCV, S. aureus (MRSA), HHV-6, Lassa, Vaccinia, VZV, common allergens, Chikungunya, Zika, Pertussis, Mycobacterium tuberculosis, Marburg, VEE, Rheumatoid Arthritis, HSV, Lyme
- Daved Fremont, Jim Crowe, Nicholas Mantis, Ben Doranz, Gregg Silverman, Stuart Levitz, Lawrence Stern, Michael Oldstone, David Koelle, William Kwok, Bjoern Peters, Alex Sette, Cecilia Lindestam Arlehamn, Daniela Weiskopf

Direct Submission Process



Data structure: a database of experiments

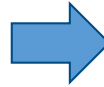


Collaborations with ontologies

- Provides standardized nomenclature, definitions, synonyms, and hierarchical relationships
- Ensures consistency and accuracy
- Finds errors
- Makes curation easier
- Enhances user experience
- Facilitates interoperability

External resources and ontologies

Peptidic Epitope	Amino acid sequence
	Protein source
	Organism source



Human herpesvirus 5 (HHV-5)
 NCBI
 taxon:10359

ORGANISM FINDER

Current Selection(s) Reset Apply

Search By

Name:

Organism ID:

Clear Search

Browse by Tree (Click to Select)

- Human herpesvirus 1
- Human herpesvirus 2
- Human herpesvirus 4 (Epstein Barr virus)
- Human herpesvirus 5 (Human cytomegalovirus)
- Human herpesvirus 5 (strain RV798)
- Human herpesvirus 5 TB40
- Human herpesvirus 5 strain AD169
- Human herpesvirus 5 strain Merlin
- Human herpesvirus 5 strain Toledo

Search Results (Click to Select)

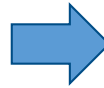
6 Records Found Page 1 of 2 5 Per Page

Organism Name	Synonyms	Organism ID
Cucumber mosaic virus (cucumber mosaic cucumovirus)	Cucumber mosaic virus, cucumber mosaic cucumovirus, cucumber mosaic virus CMV, cucumber mosaic cucumovirus CMV, cucumber mosaic virus, CMV, CMV	12305
Lymphocytic choriomeningitis mammarenavirus	Lymphocytic choriomeningitis mammarenavirus, Lymphocytic choriomeningitis virus, lymphocytic choriomeningitis virus LCMV, LCMV	11623
Murid betaherpesvirus 1 (Murine cytomegalovirus)	Murine cytomegalovirus, Murid betaherpesvirus 1, murine herpesvirus 1, murine cytomegalovirus MCMV, murine cytomegalovirus (MCMV), Murid herpesvirus 1, Mouse cytomegalovirus 1	10366
Human herpesvirus 5 (Human cytomegalovirus)	herpes virus 5,CMV,HHV5,HSV-5,HSV5, Human cytomegalovirus, Human betaherpesvirus 5, human herpesvirus type 5, Human herpesvirus 5, HHV-5	10359
Rhesus cytomegalovirus strain 68-1 (Rhesus cytomegalovirus (strain 68-1) (RhCMV))	Rhesus cytomegalovirus strain 68-1, Rhesus cytomegalovirus (strain 68-1) (RhCMV), Rhesus cytomegalovirus (strain 68-1)	103930

6 Records Found Page 1 of 2 5 Per Page

External resources and ontologies

Peptidic Epitope	Amino acid sequence
	Protein source
	Organism source



PKYVKQNTLKLAT
 hemagglutinin HA1
 GenBank/UniProt
 GI:AAL62329.1—UniProt:Q8V285

MOLECULE FINDER

Current Selection(s) Reset Apply

Search By

Name:

Molecule ID:

Source Organism: Finder

Clear Search

Browse by Tree (Click to Select)

- [-] Hepadnavirus protein
- [-] Herpesvirus protein
- [-] Influenza virus protein
 - [-] Influenza A virus protein ★★★
 - [-] Hemagglutinin ★★
 - [-] Matrix protein 1 ★★
 - [-] Matrix protein 2 ★★
 - [-] Neuraminidase ★★
 - [-] Non-structural protein 1 ★★

Search Results (Click to Select)

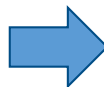
47 Records Found Page 1 of 10 5 Per Page

Molecule Name	Synonyms	Database ID	Organism Name
Influenza virus protein		IEDB [10002045]	
Influenza A virus protein ★★★		IEDB [11320]	Influenza A virus
Influenza B virus protein ★★★		IEDB [11520]	Influenza B virus
Influenza C virus protein ★★★		IEDB [11552]	Influenza C virus
Interferon-induced GTP-binding protein Mx1 ★★	MX1_MOUSE, myxoma resistance protein 1, Myxoma resistance protein 1, myxovirus (influenza) resistance 1 polypeptide, Interferon-induced GTP-binding protein Mx1, interferon-induced GTP-binding protein ...more...	UniProt [P09922]	Mus musculus (mouse)

47 Records Found Page 1 of 10 5 Per Page

External resources and ontologies

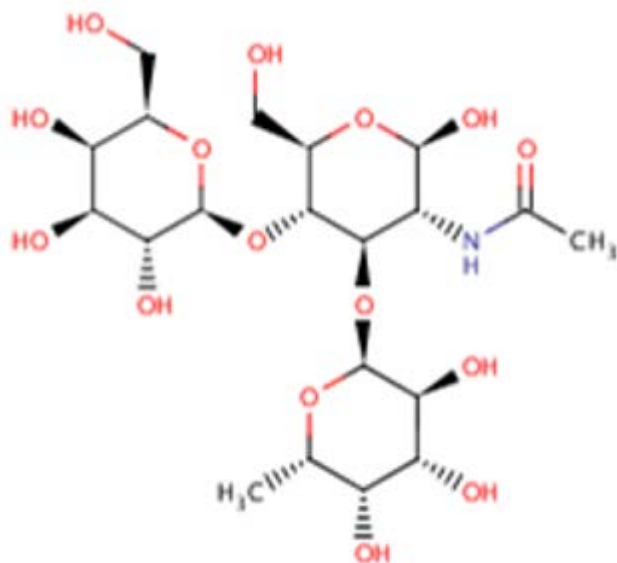
Non-peptidic Epitope	Structure name
	Source structure
	Organism source



α -L-Fucp-(1→3)-[β -D-Galp-(1→4)]- β -D-GlcpNAC
 CHEBI:59294
 Lipopolysaccharide
 CHEBI:16412



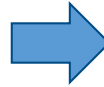
NCBI taxon:210
 Helicobacter pylori



ChEBI Name	α -L-Fucp-(1→3)-[β -D-Galp-(1→4)]- β -D-GlcpNAC
ChEBI ID	CHEBI:59294
ChEBI ASCII Name	alpha-L-Fucp-(1->3)-[beta-D-Galp-(1->4)]-beta-D-GlcpNAC
Definition	An α -L-Fucp-(1→3)-[β -D-Galp-(1→4)]-D-GlcpNAC where the glucosamine at the reducing end has β -configuration at its anomeric centre. Commonly known as Lewis x trisaccharide or Le ^x .
Stars	☆☆☆ This entity has been manually annotated by the ChEBI Team.

External resources and ontologies

Immunization Process(es)	Host
	Immunogen
	Administration
	Disease



Homo sapiens (human)
NCBI
taxon:9606



Peanut allergy
DO
DOID:4378



Metadata		Visualize
DOID	DOID:4378	
Name	peanut allergic reaction	
Definition	A food allergy that is an allergy or hypersensitivity to dietary substances from peanuts causing an overreaction of the immune system which in a small percentage of people may lead to severe physical symptoms. http://en.wikipedia.org/wiki/Allergy#Foods	
Synonyms	Allergy to peanuts (disorder) [EXACT] Peanut allergy [EXACT]	
Xrefs	MSH:D021183 SNOMEDCT_US_2015_03_01:213021008 SNOMEDCT_US_2015_03_01:91935009 UMLS_CUI:C0559470	
Relationships	is_a food allergy	

External resources and ontologies

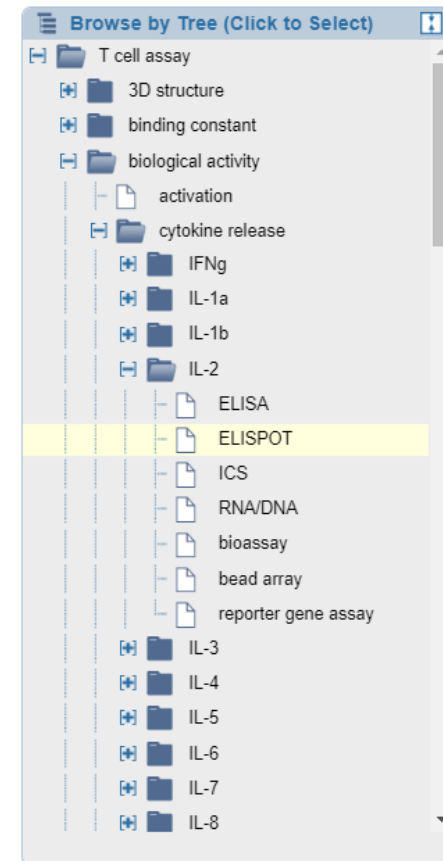
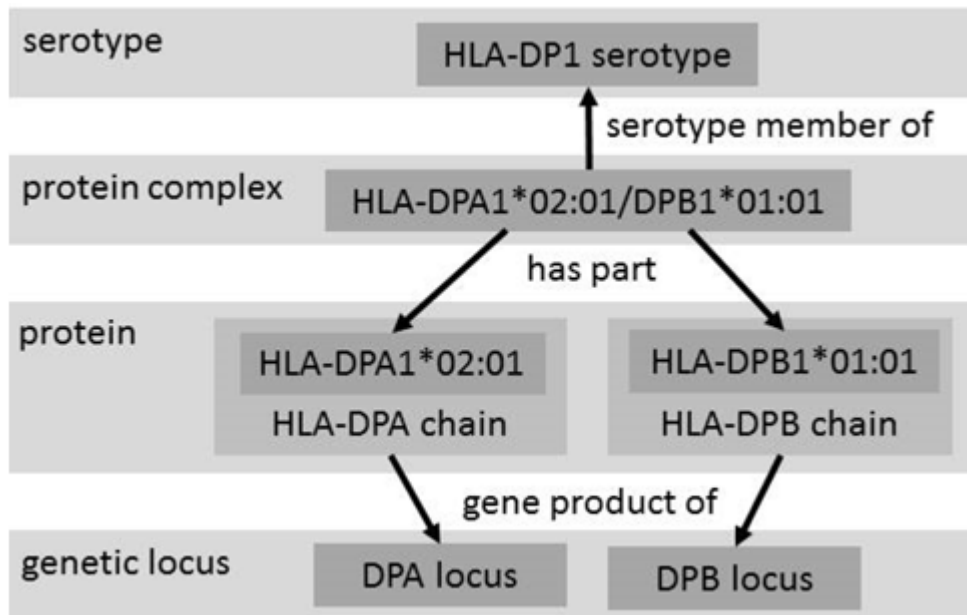
Immunization Process(es)	Host
	Immunogen
	Administration
	Disease



ELISA IL-2 release
 OBI
 OBI:1110152



MHC Restriction Ontology (MRO)



Data Aggregation

