IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE

How Data are Retrieved, Entered, and Organized

www.iedb.org

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

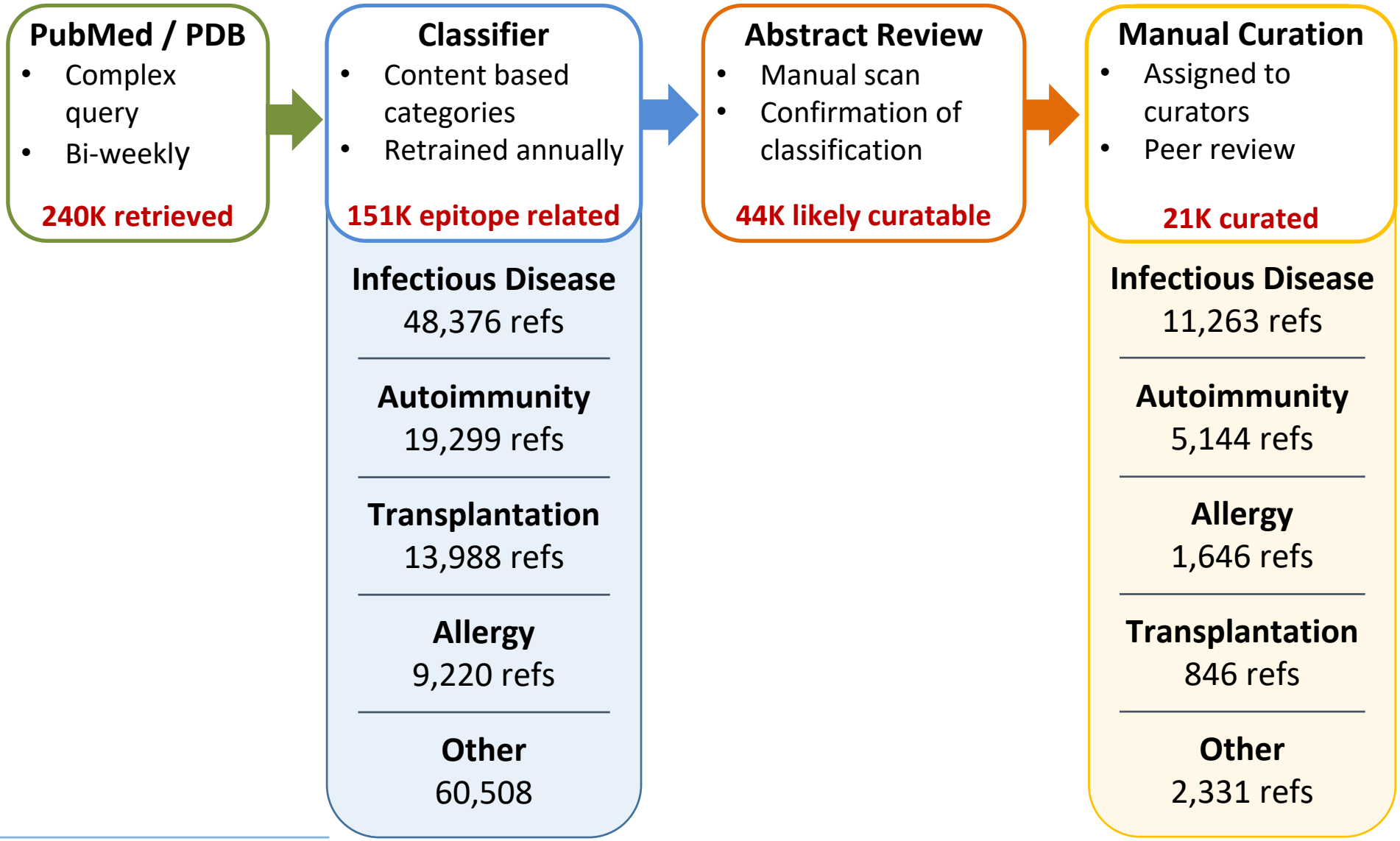
IEDB Scope

- 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

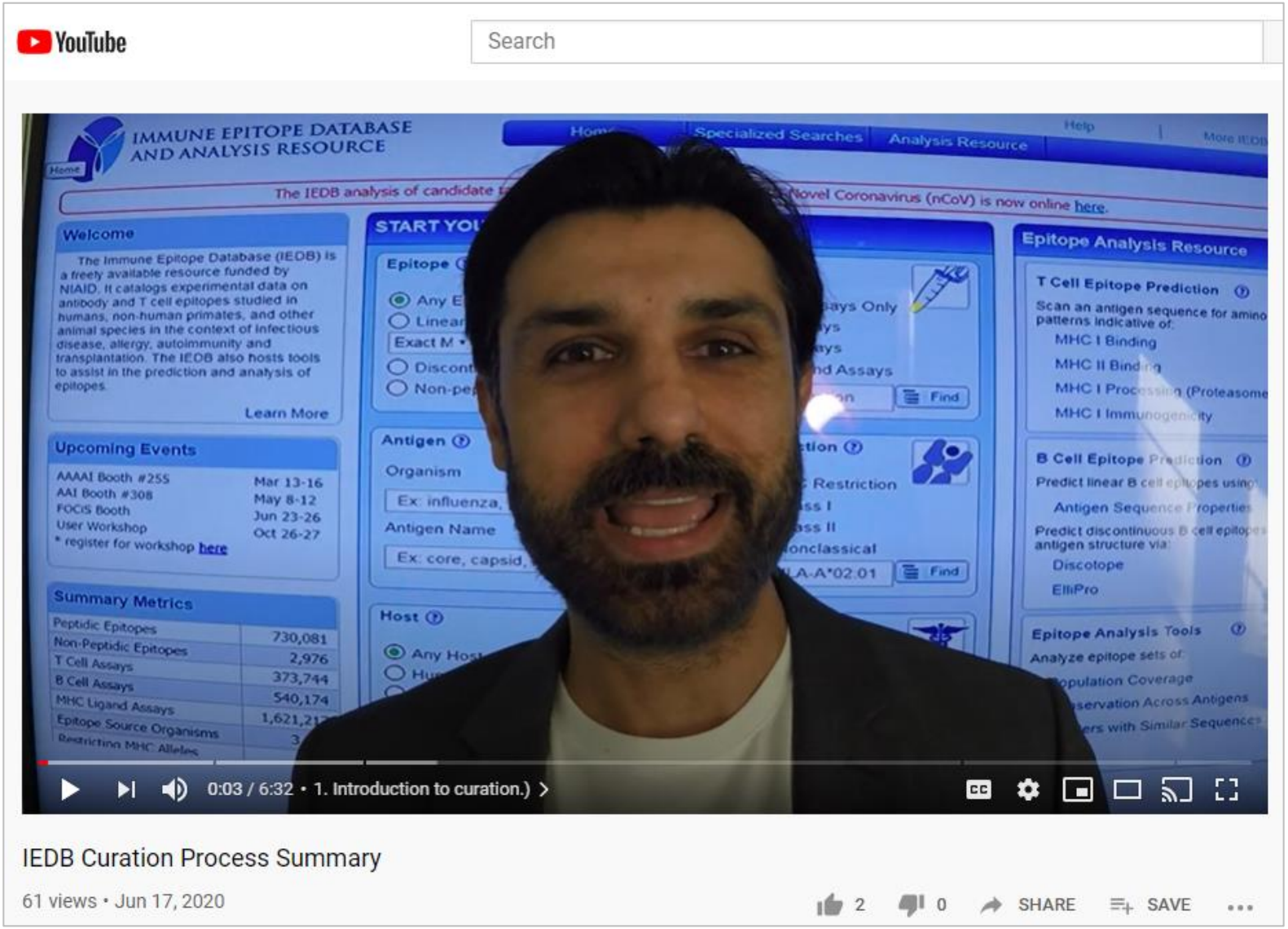
Minimal Criteria for Epitope Inclusion

- 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 <5,000 Daltons
- Minimal information required (sequence, outcome, host, etc.)

Literature Curation Process



Literature Curation Process: Video



https://youtu.be/GG6r1Vx8xBg

Consistency and Quality Control Measures

- Manually curated by a team of PhD-level scientists with specific expertise
- Formal curation guidelines and peer review
  Curation Manual: http://curationwiki.iedb.org/wiki/index.php/Main_Page
- External immunological experts
- Built in validation in the curation application

Data Submissions

- Primarily sourced through NIAID Epitope Discovery contracts
- Data deposition is open to the general research community on a case-by-case basis

As of October, 2020:

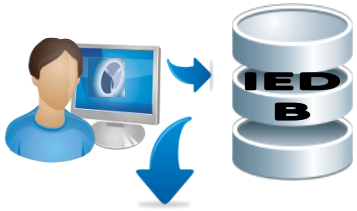
- Data from **306 submissions** is publicly available & **96 submissions** are in process or on-hold
- Submitted data comprises **28% of epitopes**

Inquire by contacting: nblazeska@lji.org or support@iedb.org!

Direct Submission Process

Submission

- Investigators submit files via IEDB data submission tool



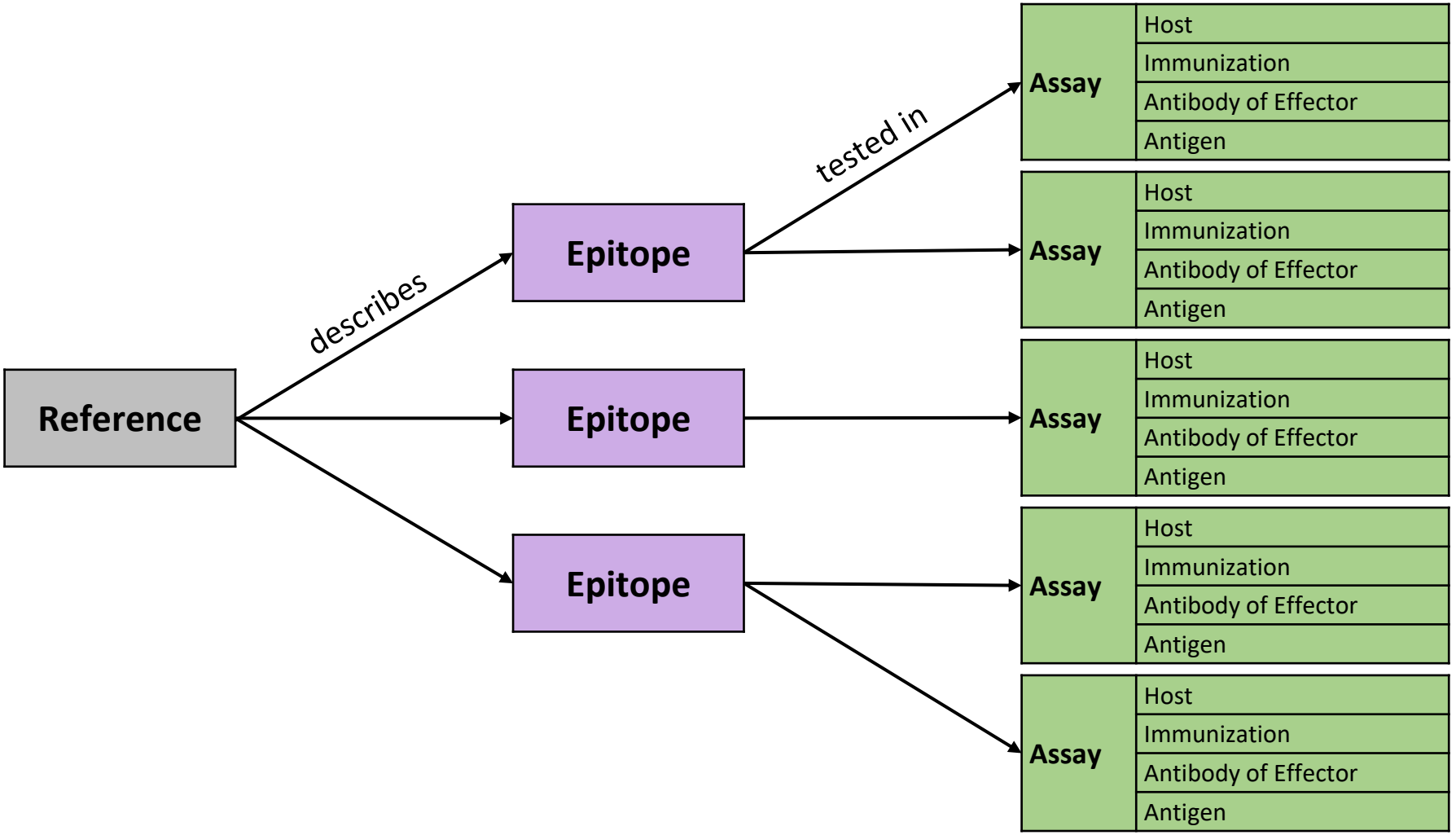
Review

- Automated validation and manual curator review check data consistency
- Submitter confirms final approval

Release

- Public release date set by submitter

Data Structure: A Database of Experiments



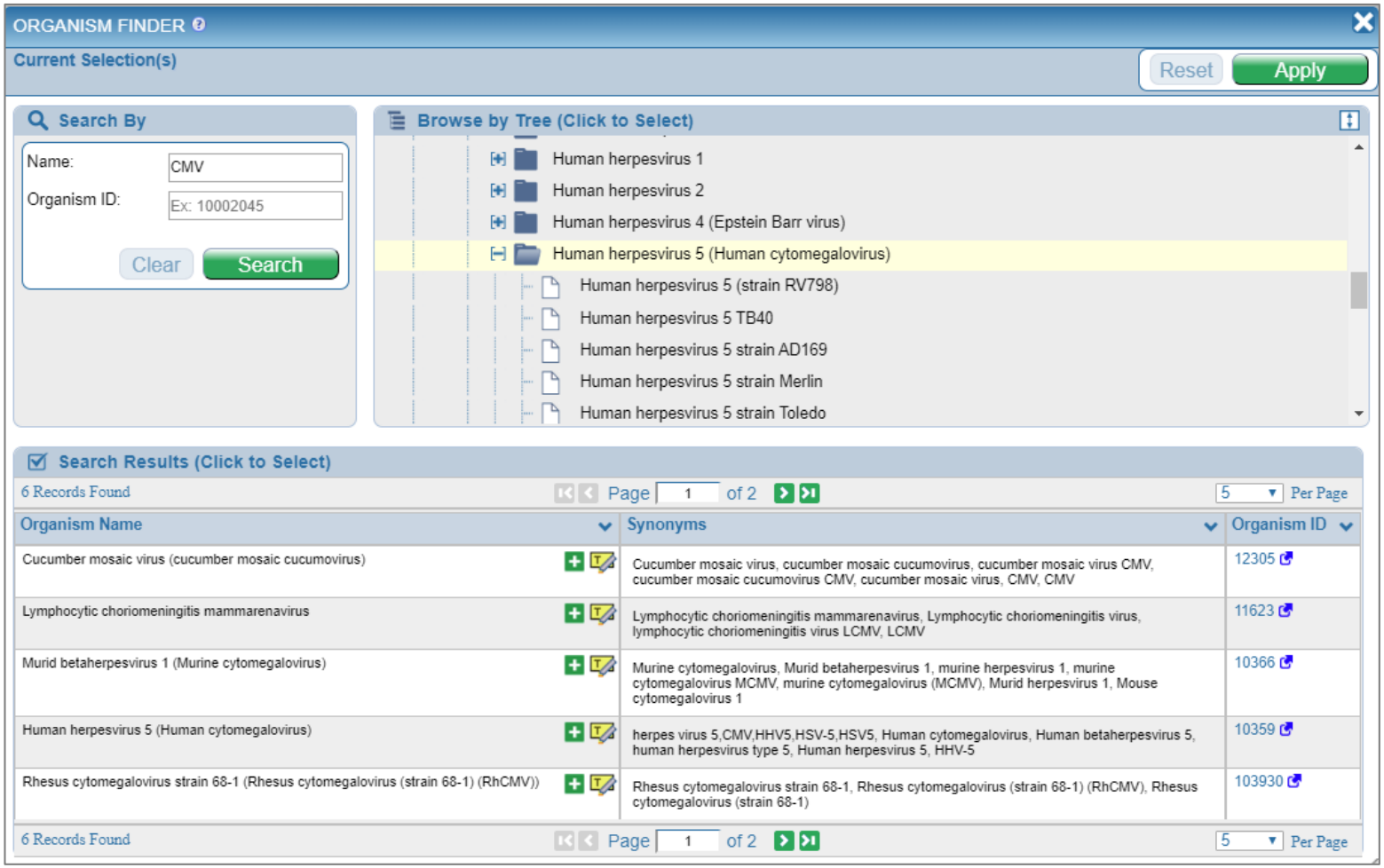
Collaborations with Ontologies

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

External Resources and Ontologies

Peptidic Epitope	Amino acid sequence
	Protein source
	Organism source

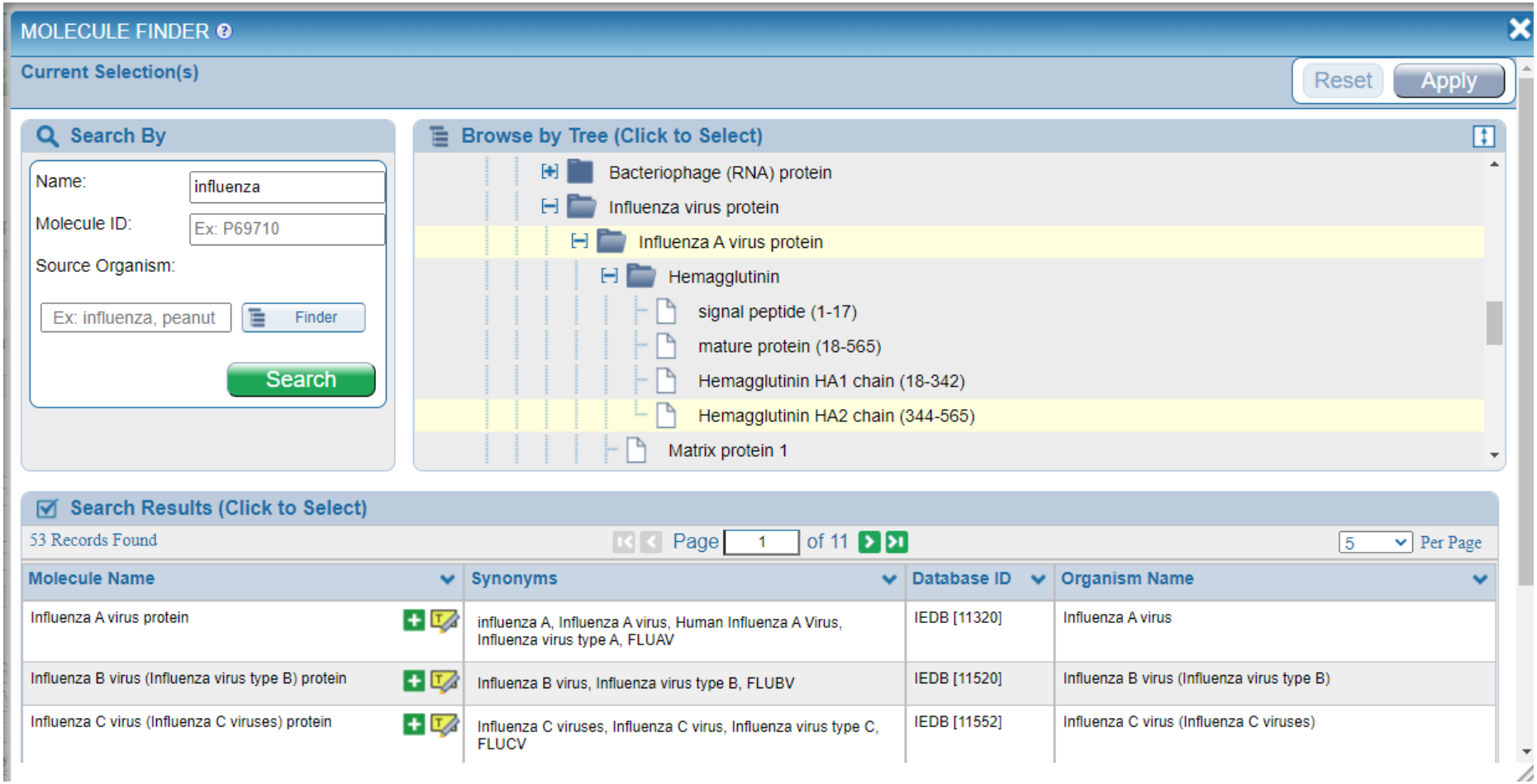
Human herpesvirus 5 (HHV-5)
NCBI
taxon:10359



External Resources and Ontologies

Peptidic Epitope	Amino acid sequence
	Protein source
	Organism source

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

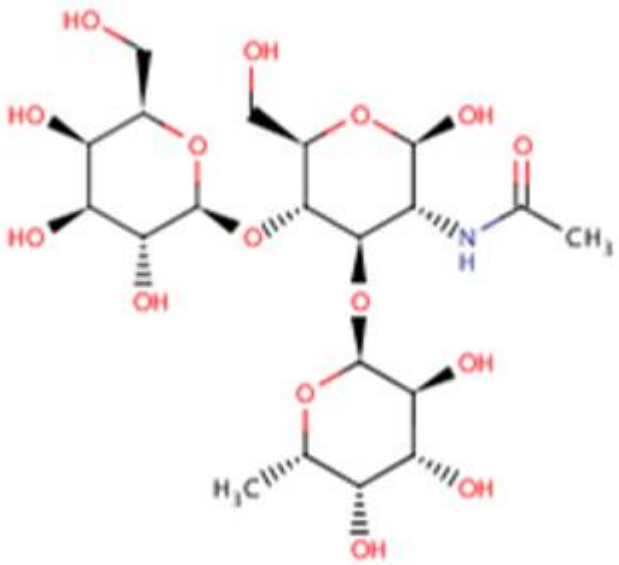


External Resources and Ontologies

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-Fuc*p*-(1→3)-[β-D-Gal*p*-(1→4)]-β-D-Glc*p*NAc
ChEBI ID	**CHEBI:59294**
ChEBI ASCII Name	alpha-L-Fucp-(1->3)-[beta-D-Galp-(1->4)]-beta-D-GlcpNAc
Definition	An α-L-Fuc*p*-(1→3)-[β-D-Gal*p*-(1→4)]-D-Glc*p*NAc 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

DISEASE ONTOLOGY

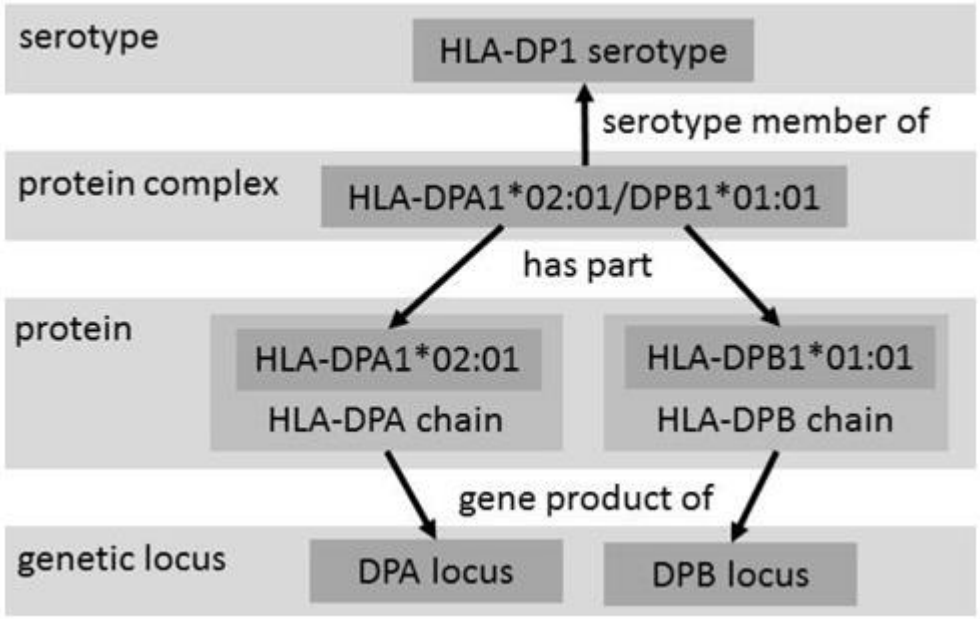
Metadata	Visualize
DOID	DOID:4378
Name	peanut allergic reaction
Definition	A food allergy that is an allergy or hypersensitivity to dietary substances from peanuts causing an overreaction of the immune system which in a small percentage of people may lead to severe physical symptoms. http://en.wikipedia.org/wiki/Allergy#Foods
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

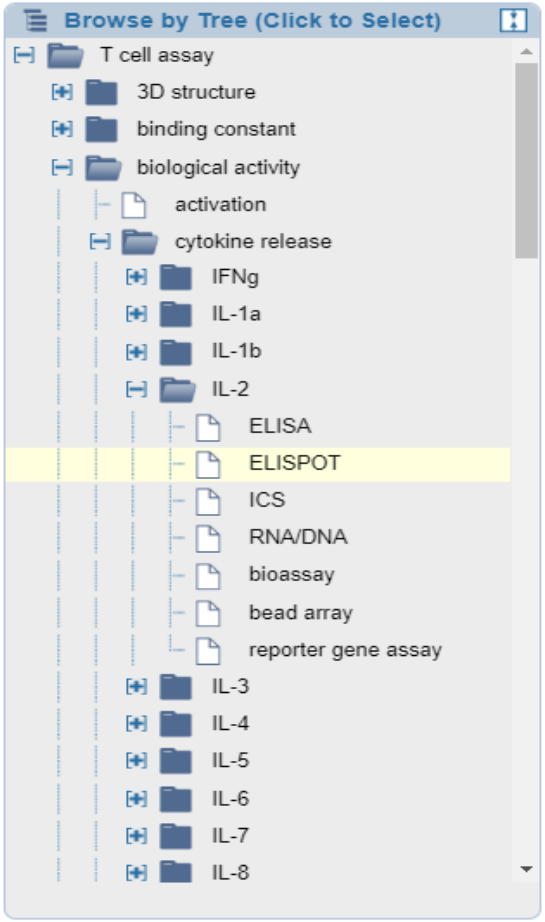
External Resources and Ontologies

Immunization Process(es)	Host
	Immunogen
	Administration
	Disease

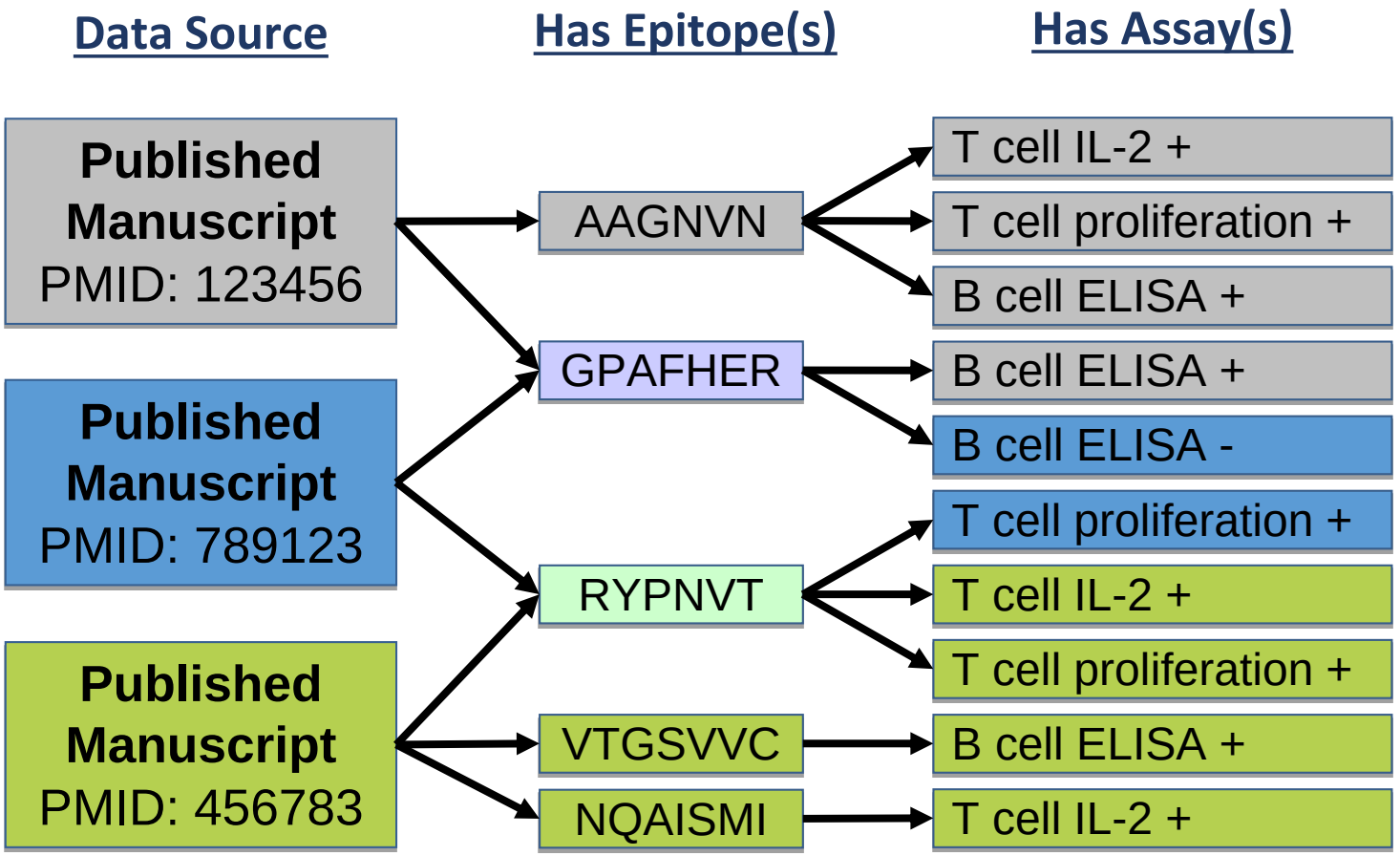
ELISA IL-2 release
OBI
OBI:1110152

MHC Restriction Ontology (MRO)





Data Aggregation

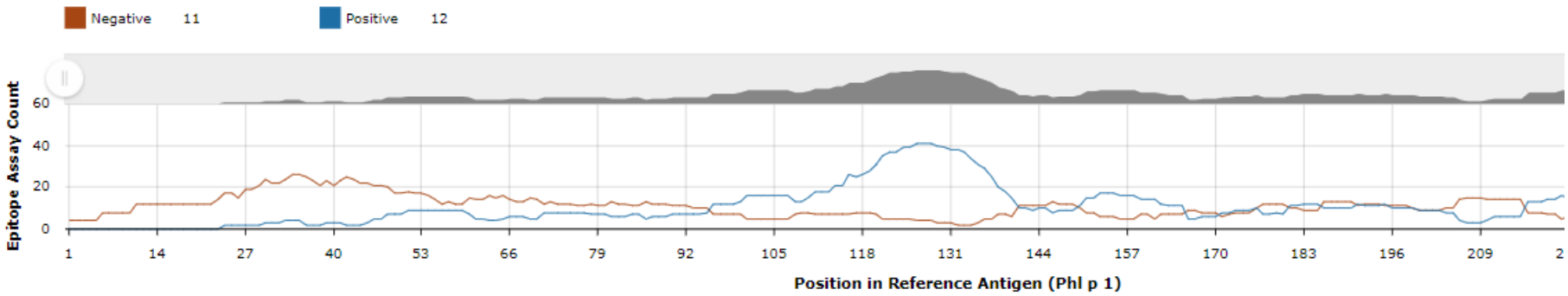


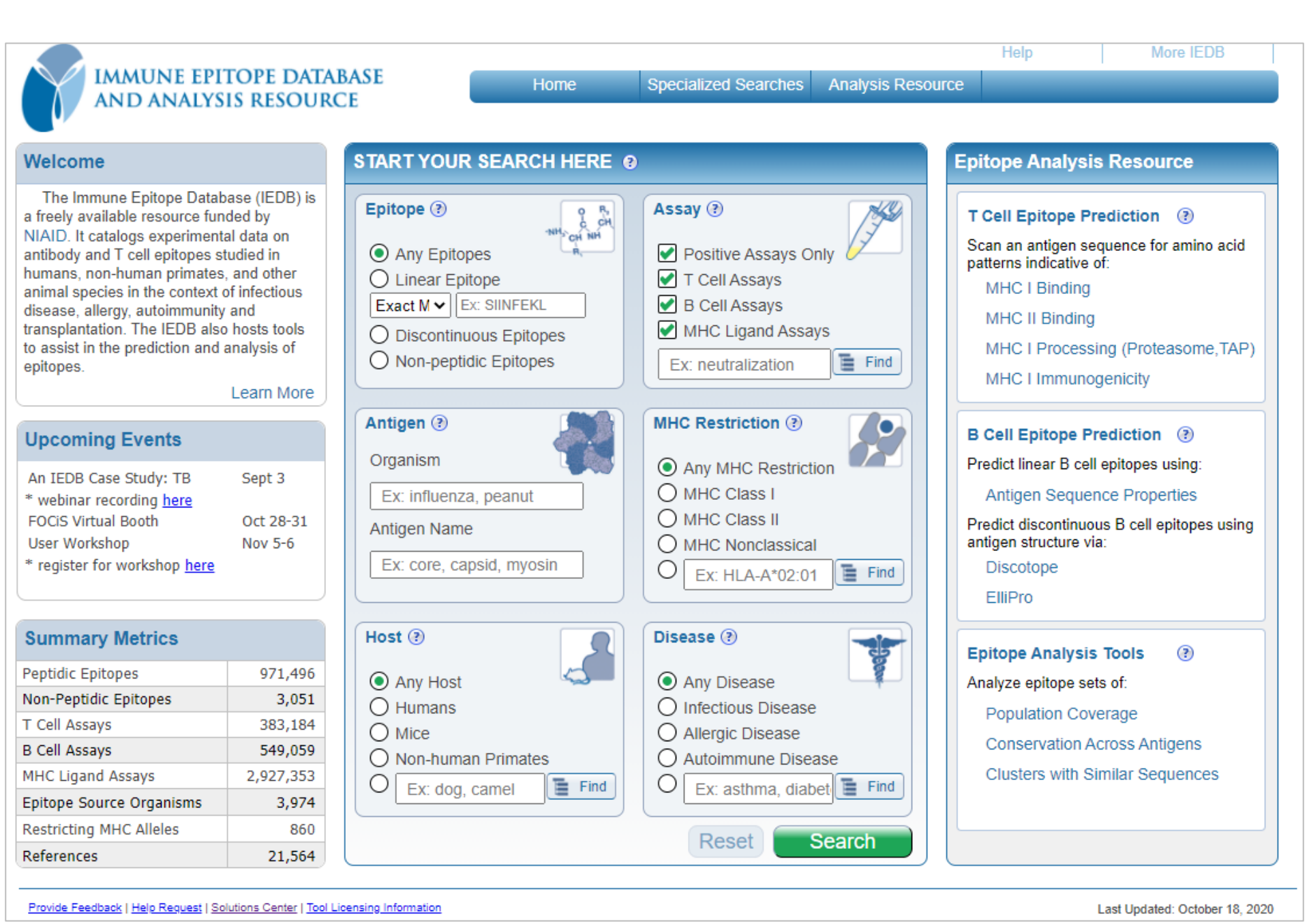
Data Aggregation

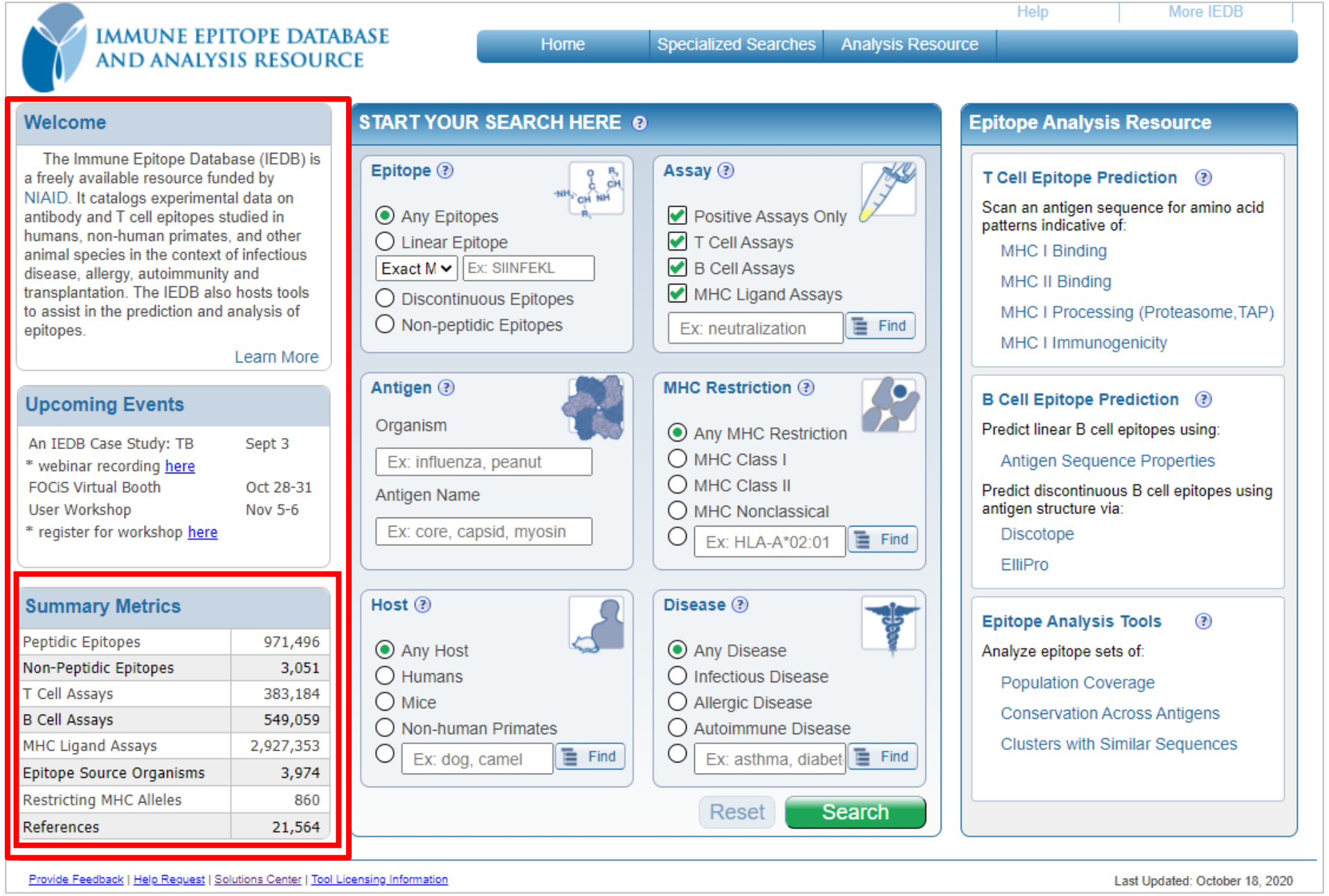
Phleum pratense (timothy grass) - Phl p 1

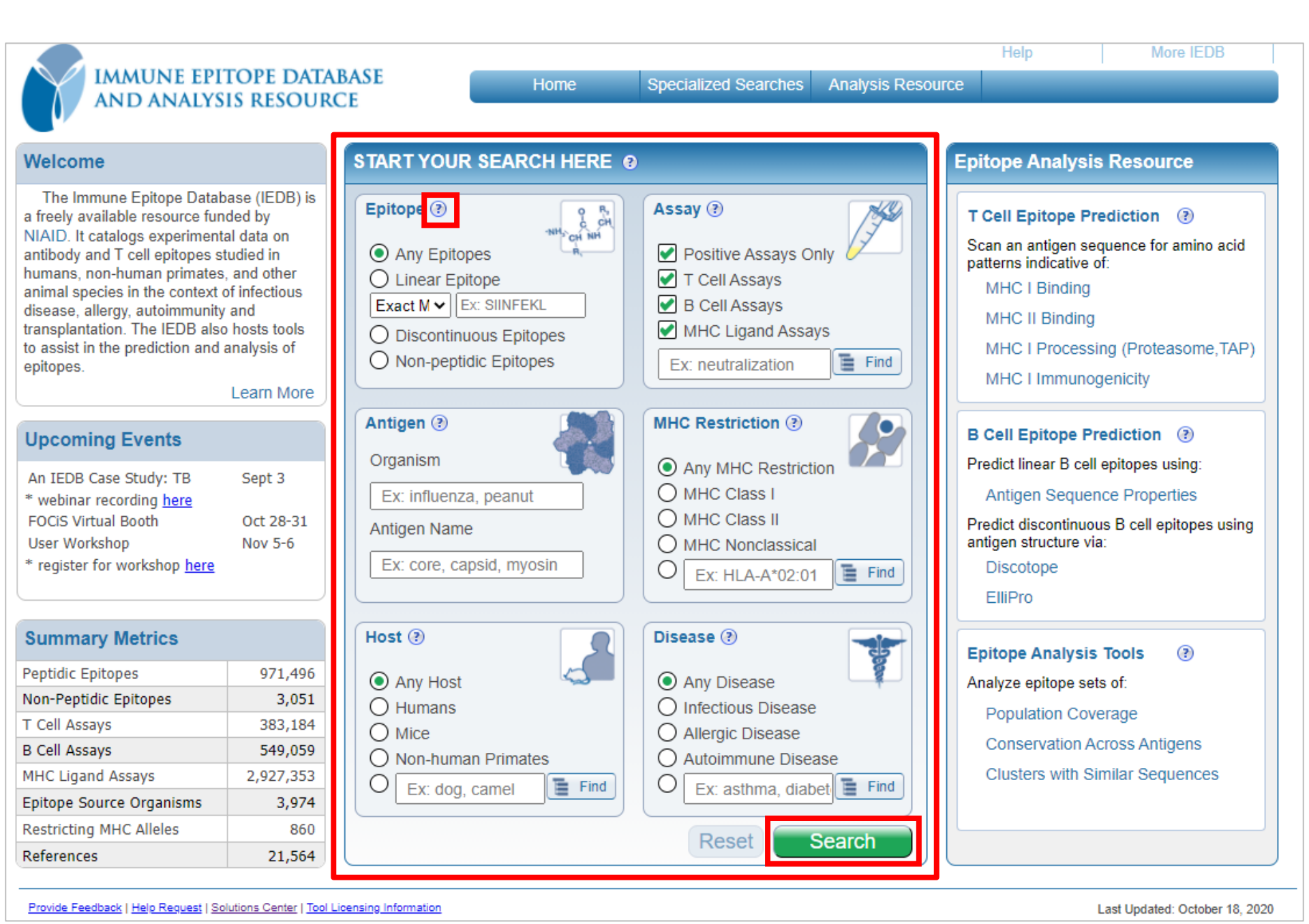
Pending Filters: Positive Assays Only | Organism: Phleum pratense (timothy grass) (ID:15957, timothy) | Host: Homo sapiens (human) | Disease Data: Allergic Disease

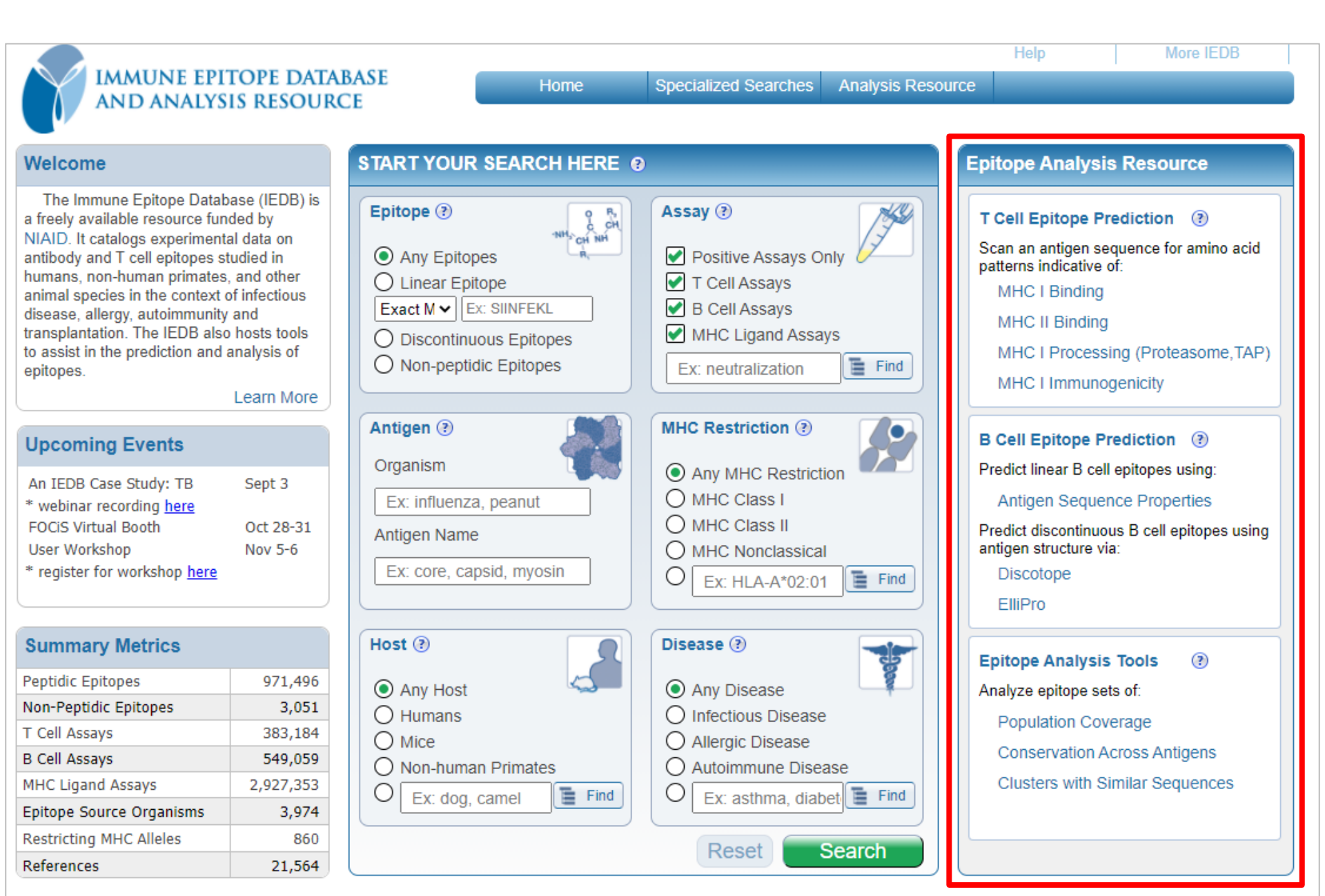
Epitope Assay Counts

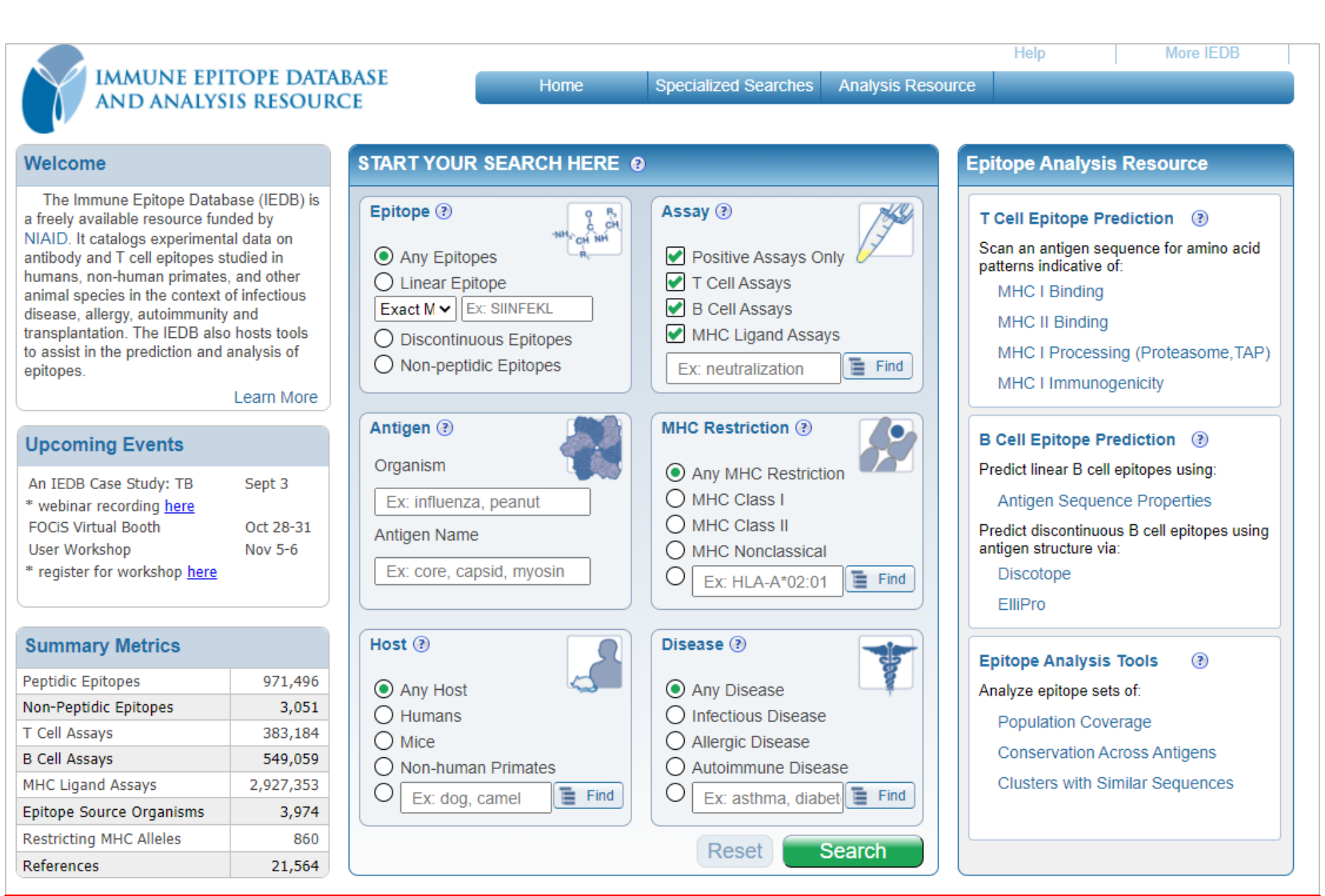



IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE
Help
More IEDB
Home
Specialized Searches
Analysis Resource
Welcome
The Immune Epitope Database (IEDB) is a freely available resource funded by NIAID. It catalogs experimental data on antibody and T cell epitopes studied in humans, non-human primates, and other animal species in the context of infectious disease, allergy, autoimmunity and transplantation. The IEDB also hosts tools to assist in the prediction and analysis of epitopes.
Learn More
Upcoming Events
An IEDB Case Study: TB Sept 3
* webinar recording here
FOCiS Virtual Booth Oct 28-31
User Workshop Nov 5-6
* register for workshop here
Summary Metrics
Peptidic Epitopes 971,496
Non-Peptidic Epitopes 3,051
T Cell Assays 383,184
B Cell Assays 549,059
MHC Ligand Assays 2,927,353
Epitope Source Organisms 3,974
Restricting MHC Alleles 860
References 21,564
START YOUR SEARCH HERE
Epitope
Any Epitopes
Linear Epitope
Exact M
Ex: SIINFEKL
Discontinuous Epitopes
Non-peptidic Epitopes
Assay
Positive Assays Only
T Cell Assays
B Cell Assays
MHC Ligand Assays
Ex: neutralization
Find
Antigen
Organism
Ex: influenza, peanut
Antigen Name
Ex: core, capsid, myosin
MHC Restriction
Any MHC Restriction
MHC Class I
MHC Class II
MHC Nonclassical
Ex: HLA-A*02:01
Find
Host
Any Host
Humans
Mice
Non-human Primates
Ex: dog, camel
Find
Disease
Any Disease
Infectious Disease
Allergic Disease
Autoimmune Disease
Ex: asthma, diabet
Find
Reset
Search
Epitope Analysis Resource
T Cell Epitope Prediction
Scan an antigen sequence for amino acid patterns indicative of:
MHC I Binding
MHC II Binding
MHC I Processing (Proteasome,TAP)
MHC I Immunogenicity
B Cell Epitope Prediction
Predict linear B cell epitopes using:
Antigen Sequence Properties
Predict discontinuous B cell epitopes using antigen structure via:
Discotope
ElliPro
Epitope Analysis Tools
Analyze epitope sets of:
Population Coverage
Conservation Across Antigens
Clusters with Similar Sequences
Provide Feedback | Help Request | Solutions Center | Tool Licensing Information
Last Updated: October 18, 2020

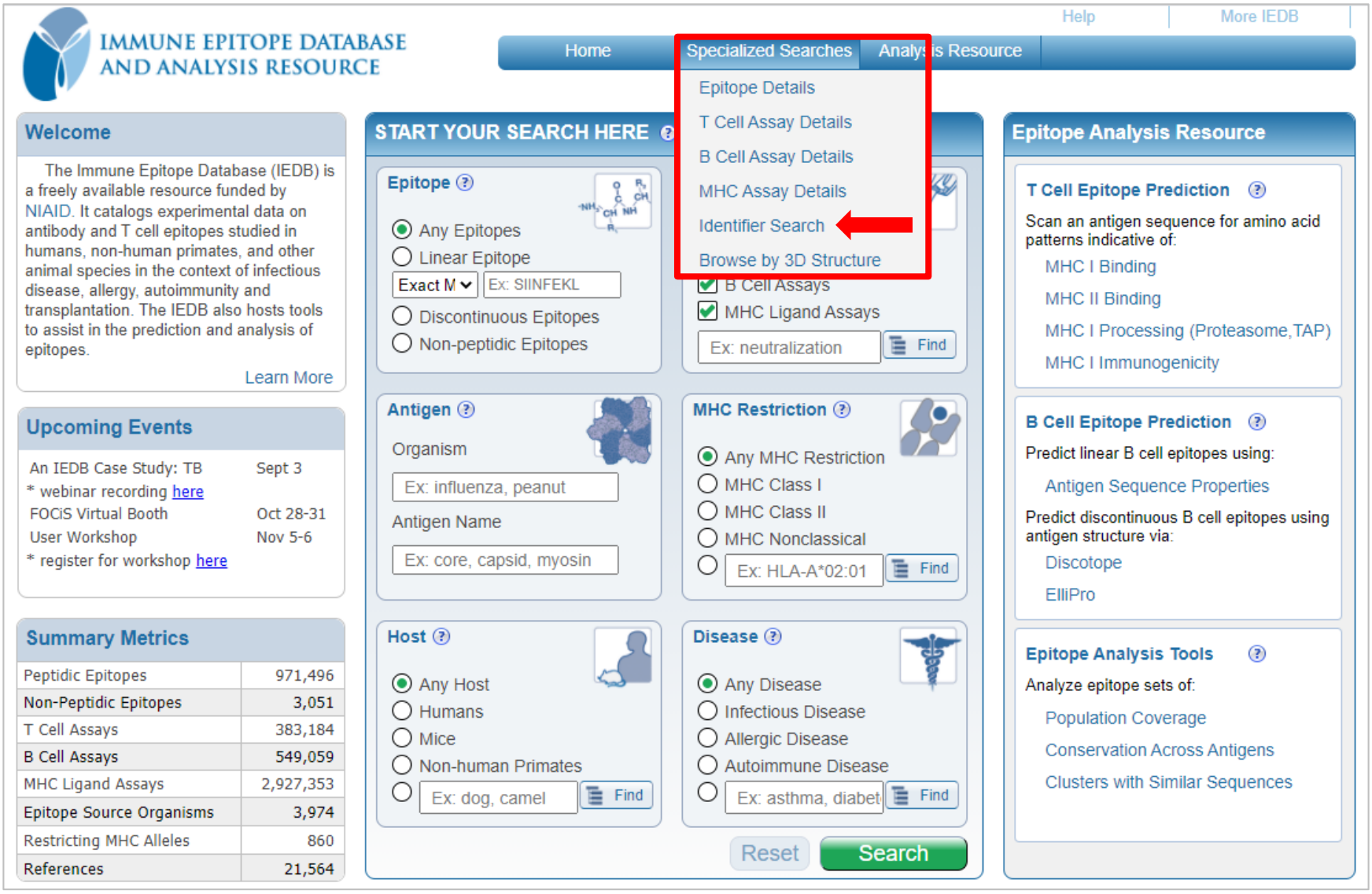

IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE
Help
More IEDB
Home
Specialized Searches
Analysis Resource
Welcome
The Immune Epitope Database (IEDB) is a freely available resource funded by NIAID. It catalogs experimental data on antibody and T cell epitopes studied in humans, non-human primates, and other animal species in the context of infectious disease, allergy, autoimmunity and transplantation. The IEDB also hosts tools to assist in the prediction and analysis of epitopes.
Learn More
Upcoming Events
An IEDB Case Study: TB Sept 3
* webinar recording here
FOCiS Virtual Booth Oct 28-31
User Workshop Nov 5-6
* register for workshop here
Summary Metrics
Peptidic Epitopes 971,496
Non-Peptidic Epitopes 3,051
T Cell Assays 383,184
B Cell Assays 549,059
MHC Ligand Assays 2,927,353
Epitope Source Organisms 3,974
Restricting MHC Alleles 860
References 21,564
START YOUR SEARCH HERE
Epitope
Any Epitopes
Linear Epitope
Exact M
Ex: SIINFEKL
Discontinuous Epitopes
Non-peptidic Epitopes
Assay
Positive Assays Only
T Cell Assays
B Cell Assays
MHC Ligand Assays
Ex: neutralization
Find
Antigen
Organism
Ex: influenza, peanut
Antigen Name
Ex: core, capsid, myosin
MHC Restriction
Any MHC Restriction
MHC Class I
MHC Class II
MHC Nonclassical
Ex: HLA-A*02:01
Find
Host
Any Host
Humans
Mice
Non-human Primates
Ex: dog, camel
Find
Disease
Any Disease
Infectious Disease
Allergic Disease
Autoimmune Disease
Ex: asthma, diabet
Find
Reset
Search
Epitope Analysis Resource
T Cell Epitope Prediction
Scan an antigen sequence for amino acid patterns indicative of:
MHC I Binding
MHC II Binding
MHC I Processing (Proteasome,TAP)
MHC I Immunogenicity
B Cell Epitope Prediction
Predict linear B cell epitopes using:
Antigen Sequence Properties
Predict discontinuous B cell epitopes using antigen structure via:
Discotope
ElliPro
Epitope Analysis Tools
Analyze epitope sets of:
Population Coverage
Conservation Across Antigens
Clusters with Similar Sequences
Provide Feedback | Help Request | Solutions Center | Tool Licensing Information
Last Updated: October 18, 2020


IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE
Help
More IEDB
Home
Specialized Searches
Analysis Resource
Welcome
The Immune Epitope Database (IEDB) is a freely available resource funded by NIAID. It catalogs experimental data on antibody and T cell epitopes studied in humans, non-human primates, and other animal species in the context of infectious disease, allergy, autoimmunity and transplantation. The IEDB also hosts tools to assist in the prediction and analysis of epitopes.
Learn More
Upcoming Events
An IEDB Case Study: TB Sept 3
* webinar recording here
FOCiS Virtual Booth Oct 28-31
User Workshop Nov 5-6
* register for workshop here
Summary Metrics
Peptidic Epitopes 971,496
Non-Peptidic Epitopes 3,051
T Cell Assays 383,184
B Cell Assays 549,059
MHC Ligand Assays 2,927,353
Epitope Source Organisms 3,974
Restricting MHC Alleles 860
References 21,564
START YOUR SEARCH HERE
Epitope
Any Epitopes
Linear Epitope
Exact M
Ex: SIINFEKL
Discontinuous Epitopes
Non-peptidic Epitopes
Assay
Positive Assays Only
T Cell Assays
B Cell Assays
MHC Ligand Assays
Ex: neutralization
Find
Antigen
Organism
Ex: influenza, peanut
Antigen Name
Ex: core, capsid, myosin
MHC Restriction
Any MHC Restriction
MHC Class I
MHC Class II
MHC Nonclassical
Ex: HLA-A*02:01
Find
Host
Any Host
Humans
Mice
Non-human Primates
Ex: dog, camel
Find
Disease
Any Disease
Infectious Disease
Allergic Disease
Autoimmune Disease
Ex: asthma, diabet
Find
Reset
Search
Epitope Analysis Resource
T Cell Epitope Prediction
Scan an antigen sequence for amino acid patterns indicative of:
MHC I Binding
MHC II Binding
MHC I Processing (Proteasome,TAP)
MHC I Immunogenicity
B Cell Epitope Prediction
Predict linear B cell epitopes using:
Antigen Sequence Properties
Predict discontinuous B cell epitopes using antigen structure via:
Discotope
ElliPro
Epitope Analysis Tools
Analyze epitope sets of:
Population Coverage
Conservation Across Antigens
Clusters with Similar Sequences
Provide Feedback | Help Request | Solutions Center | Tool Licensing Information
Last Updated: October 18, 2020

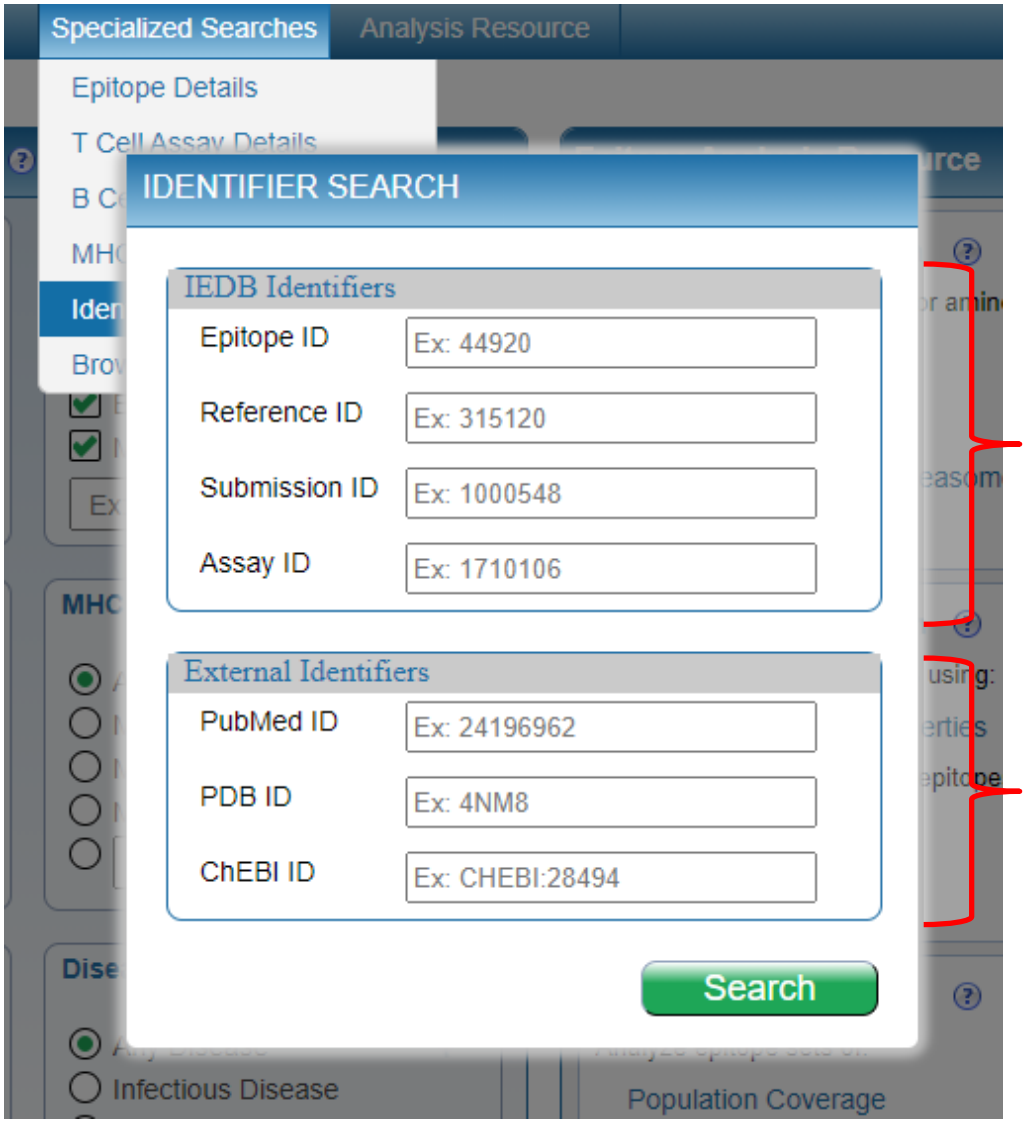

IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE
Help
More IEDB
Home
Specialized Searches
Analysis Resource
Welcome
The Immune Epitope Database (IEDB) is a freely available resource funded by NIAID. It catalogs experimental data on antibody and T cell epitopes studied in humans, non-human primates, and other animal species in the context of infectious disease, allergy, autoimmunity and transplantation. The IEDB also hosts tools to assist in the prediction and analysis of epitopes.
Learn More
Upcoming Events
An IEDB Case Study: TB Sept 3
* webinar recording here
FOCiS Virtual Booth Oct 28-31
User Workshop Nov 5-6
* register for workshop here
Summary Metrics
Peptidic Epitopes 971,496
Non-Peptidic Epitopes 3,051
T Cell Assays 383,184
B Cell Assays 549,059
MHC Ligand Assays 2,927,353
Epitope Source Organisms 3,974
Restricting MHC Alleles 860
References 21,564
START YOUR SEARCH HERE
Epitope
Any Epitopes
Linear Epitope
Exact M
Ex: SIINFEKL
Discontinuous Epitopes
Non-peptidic Epitopes
Assay
Positive Assays Only
T Cell Assays
B Cell Assays
MHC Ligand Assays
Ex: neutralization
Find
Antigen
Organism
Ex: influenza, peanut
Antigen Name
Ex: core, capsid, myosin
MHC Restriction
Any MHC Restriction
MHC Class I
MHC Class II
MHC Nonclassical
Ex: HLA-A*02:01
Find
Host
Any Host
Humans
Mice
Non-human Primates
Ex: dog, camel
Find
Disease
Any Disease
Infectious Disease
Allergic Disease
Autoimmune Disease
Ex: asthma, diabet
Find
Reset
Search
Epitope Analysis Resource
T Cell Epitope Prediction
Scan an antigen sequence for amino acid patterns indicative of:
MHC I Binding
MHC II Binding
MHC I Processing (Proteasome,TAP)
MHC I Immunogenicity
B Cell Epitope Prediction
Predict linear B cell epitopes using:
Antigen Sequence Properties
Predict discontinuous B cell epitopes using antigen structure via:
Discotope
ElliPro
Epitope Analysis Tools
Analyze epitope sets of:
Population Coverage
Conservation Across Antigens
Clusters with Similar Sequences
Provide Feedback | Help Request | Solutions Center | Tool Licensing Information
Last Updated: October 18, 2020


IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE
Help
More IEDB
Home
Specialized Searches
Analysis Resource
Welcome
The Immune Epitope Database (IEDB) is a freely available resource funded by NIAID. It catalogs experimental data on antibody and T cell epitopes studied in humans, non-human primates, and other animal species in the context of infectious disease, allergy, autoimmunity and transplantation. The IEDB also hosts tools to assist in the prediction and analysis of epitopes.
Learn More
Upcoming Events
An IEDB Case Study: TB Sept 3
* webinar recording here
FOCiS Virtual Booth Oct 28-31
User Workshop Nov 5-6
* register for workshop here
Summary Metrics
Peptidic Epitopes 971,496
Non-Peptidic Epitopes 3,051
T Cell Assays 383,184
B Cell Assays 549,059
MHC Ligand Assays 2,927,353
Epitope Source Organisms 3,974
Restricting MHC Alleles 860
References 21,564
START YOUR SEARCH HERE
Epitope
Any Epitopes
Linear Epitope
Exact M
Ex: SIINFEKL
Discontinuous Epitopes
Non-peptidic Epitopes
Assay
Positive Assays Only
T Cell Assays
B Cell Assays
MHC Ligand Assays
Ex: neutralization
Find
Antigen
Organism
Ex: influenza, peanut
Antigen Name
Ex: core, capsid, myosin
MHC Restriction
Any MHC Restriction
MHC Class I
MHC Class II
MHC Nonclassical
Ex: HLA-A*02:01
Find
Host
Any Host
Humans
Mice
Non-human Primates
Ex: dog, camel
Find
Disease
Any Disease
Infectious Disease
Allergic Disease
Autoimmune Disease
Ex: asthma, diabet
Find
Reset
Search
Epitope Analysis Resource
T Cell Epitope Prediction
Scan an antigen sequence for amino acid patterns indicative of:
MHC I Binding
MHC II Binding
MHC I Processing (Proteasome,TAP)
MHC I Immunogenicity
B Cell Epitope Prediction
Predict linear B cell epitopes using:
Antigen Sequence Properties
Predict discontinuous B cell epitopes using antigen structure via:
Discotope
ElliPro
Epitope Analysis Tools
Analyze epitope sets of:
Population Coverage
Conservation Across Antigens
Clusters with Similar Sequences
Provide Feedback | Help Request | Solutions Center | Tool Licensing Information
Last Updated: October 18, 2020

Identifier Search



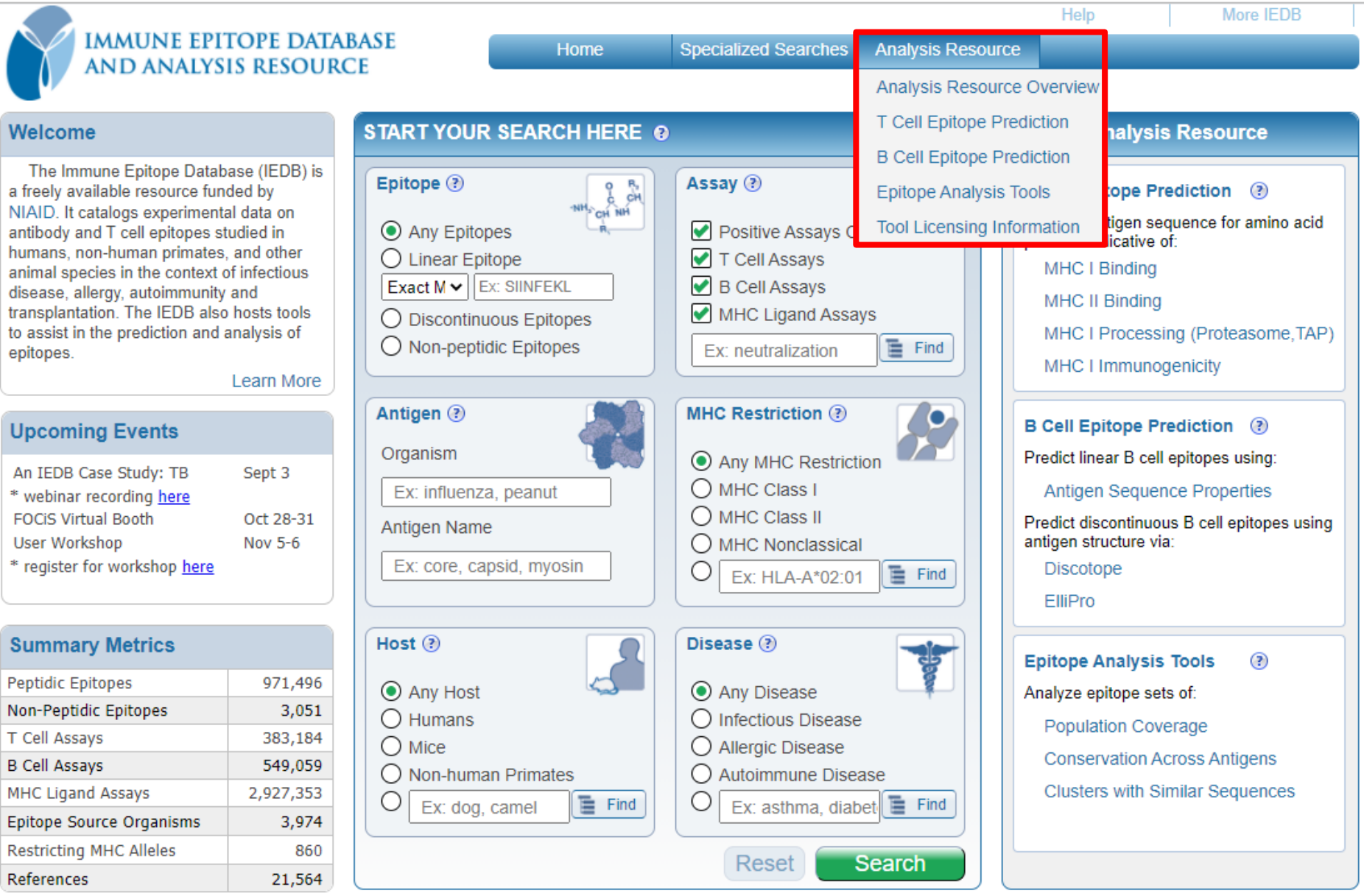
Identifier Search

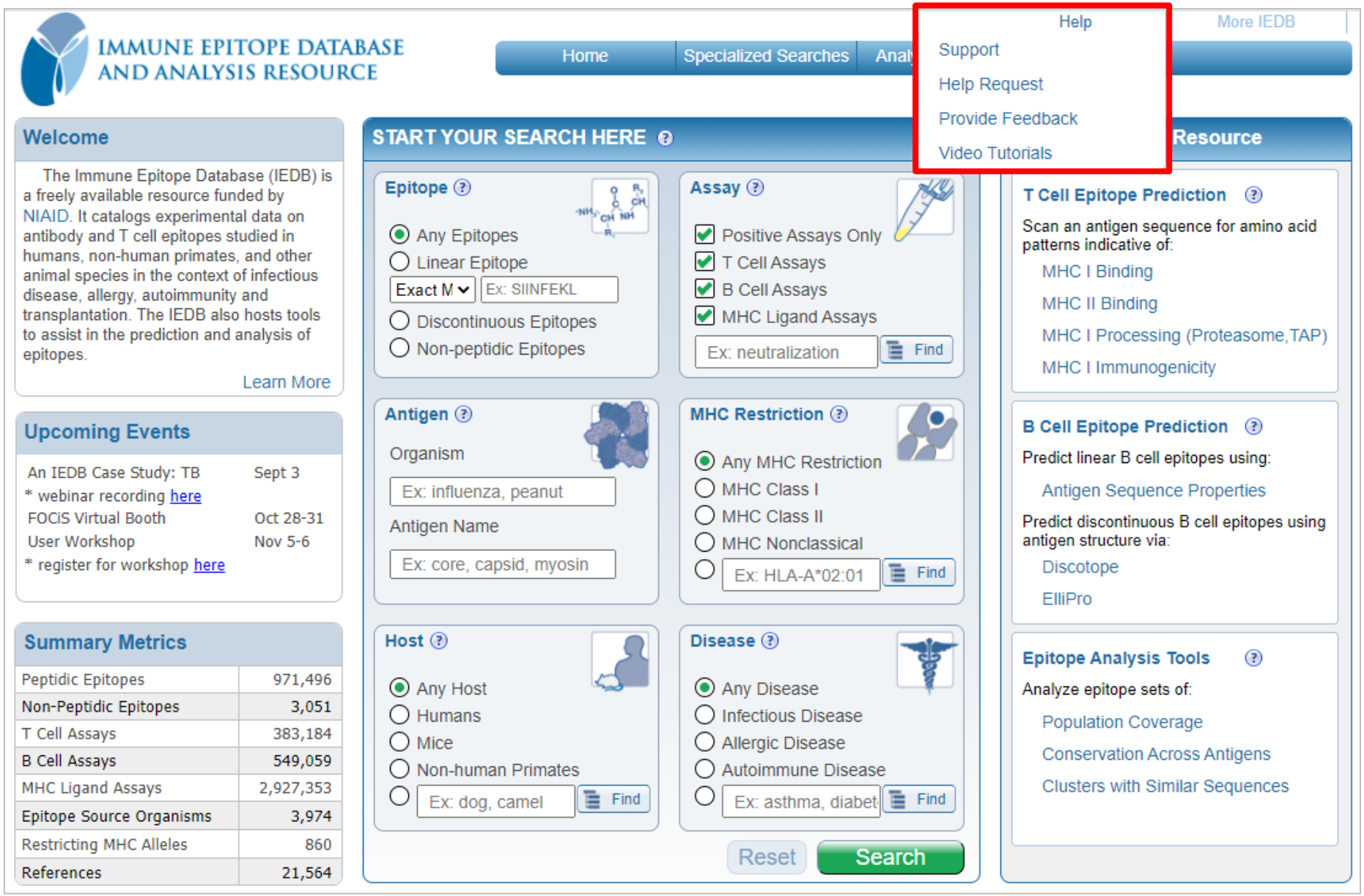


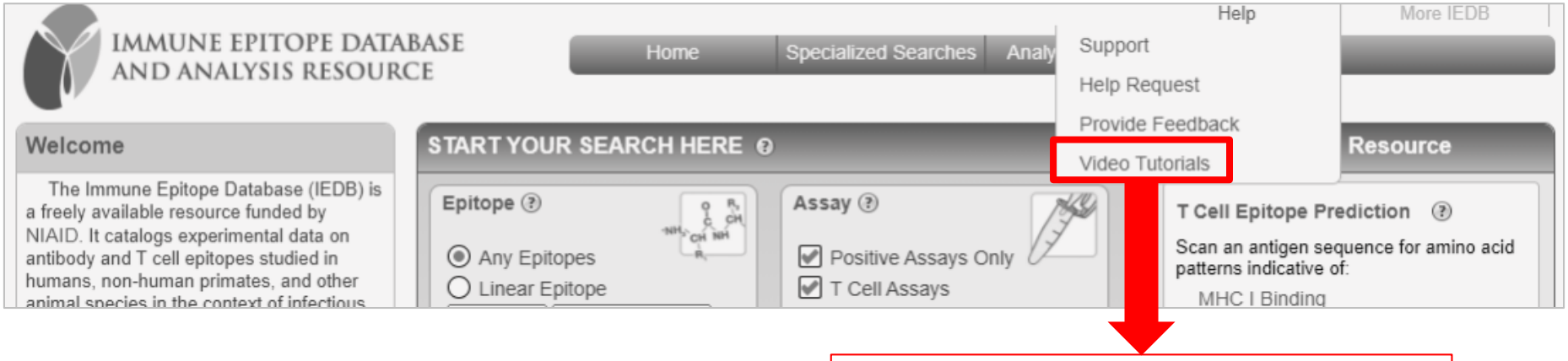
A quick way to retrieve data by using an identifier

Unique identifiers created by IEDB

Unique identifiers from other resources


IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE
Help
More IEDB
Home
Specialized Searches
Analysis Resource
Analysis Resource Overview
T Cell Epitope Prediction
B Cell Epitope Prediction
Epitope Analysis Tools
Tool Licensing Information
Welcome
The Immune Epitope Database (IEDB) is a freely available resource funded by NIAID. It catalogs experimental data on antibody and T cell epitopes studied in humans, non-human primates, and other animal species in the context of infectious disease, allergy, autoimmunity and transplantation. The IEDB also hosts tools to assist in the prediction and analysis of epitopes.
Learn More
Upcoming Events
An IEDB Case Study: TB Sept 3
* webinar recording here
FOCiS Virtual Booth Oct 28-31
User Workshop Nov 5-6
* register for workshop here
Summary Metrics
Peptidic Epitopes 971,496
Non-Peptidic Epitopes 3,051
T Cell Assays 383,184
B Cell Assays 549,059
MHC Ligand Assays 2,927,353
Epitope Source Organisms 3,974
Restricting MHC Alleles 860
References 21,564
START YOUR SEARCH HERE
Epitope
Any Epitopes
Linear Epitope
Exact M
Ex: SIINFEKL
Discontinuous Epitopes
Non-peptidic Epitopes
Assay
Positive Assays C
T Cell Assays
B Cell Assays
MHC Ligand Assays
Ex: neutralization
Find
Antigen
Organism
Ex: influenza, peanut
Antigen Name
Ex: core, capsid, myosin
MHC Restriction
Any MHC Restriction
MHC Class I
MHC Class II
MHC Nonclassical
Ex: HLA-A*02:01
Find
Host
Any Host
Humans
Mice
Non-human Primates
Ex: dog, camel
Find
Disease
Any Disease
Infectious Disease
Allergic Disease
Autoimmune Disease
Ex: asthma, diabet
Find
Reset
Search
nalysis Resource
tope Prediction
tigen sequence for amino acid
icative of:
MHC I Binding
MHC II Binding
MHC I Processing (Proteasome,TAP)
MHC I Immunogenicity
B Cell Epitope Prediction
Predict linear B cell epitopes using:
Antigen Sequence Properties
Predict discontinuous B cell epitopes using antigen structure via:
Discotope
ElliPro
Epitope Analysis Tools
Analyze epitope sets of:
Population Coverage
Conservation Across Antigens
Clusters with Similar Sequences


IMMUNE EPITOPE DATABASE AND ANALYSIS RESOURCE
Home
Specialized Searches
Anal
Help
Support
Help Request
Provide Feedback
Video Tutorials
More IEDB
Welcome
The Immune Epitope Database (IEDB) is a freely available resource funded by NIAID. It catalogs experimental data on antibody and T cell epitopes studied in humans, non-human primates, and other animal species in the context of infectious disease, allergy, autoimmunity and transplantation. The IEDB also hosts tools to assist in the prediction and analysis of epitopes.
Learn More
Upcoming Events
An IEDB Case Study: TB Sept 3
* webinar recording here
FOCiS Virtual Booth Oct 28-31
User Workshop Nov 5-6
* register for workshop here
Summary Metrics
Peptidic Epitopes 971,496
Non-Peptidic Epitopes 3,051
T Cell Assays 383,184
B Cell Assays 549,059
MHC Ligand Assays 2,927,353
Epitope Source Organisms 3,974
Restricting MHC Alleles 860
References 21,564
START YOUR SEARCH HERE
Epitope
Any Epitopes
Linear Epitope
Exact M
Ex: SIINFEKL
Discontinuous Epitopes
Non-peptidic Epitopes
Assay
Positive Assays Only
T Cell Assays
B Cell Assays
MHC Ligand Assays
Ex: neutralization
Find
Antigen
Organism
Ex: influenza, peanut
Antigen Name
Ex: core, capsid, myosin
MHC Restriction
Any MHC Restriction
MHC Class I
MHC Class II
MHC Nonclassical
Ex: HLA-A*02:01
Find
Host
Any Host
Humans
Mice
Non-human Primates
Ex: dog, camel
Find
Disease
Any Disease
Infectious Disease
Allergic Disease
Autoimmune Disease
Ex: asthma, diabet
Find
Reset
Search
Resource
T Cell Epitope Prediction
Scan an antigen sequence for amino acid patterns indicative of:
MHC I Binding
MHC II Binding
MHC I Processing (Proteasome,TAP)
MHC I Immunogenicity
B Cell Epitope Prediction
Predict linear B cell epitopes using:
Antigen Sequence Properties
Predict discontinuous B cell epitopes using antigen structure via:
Discotope
ElliPro
Epitope Analysis Tools
Analyze epitope sets of:
Population Coverage
Conservation Across Antigens
Clusters with Similar Sequences

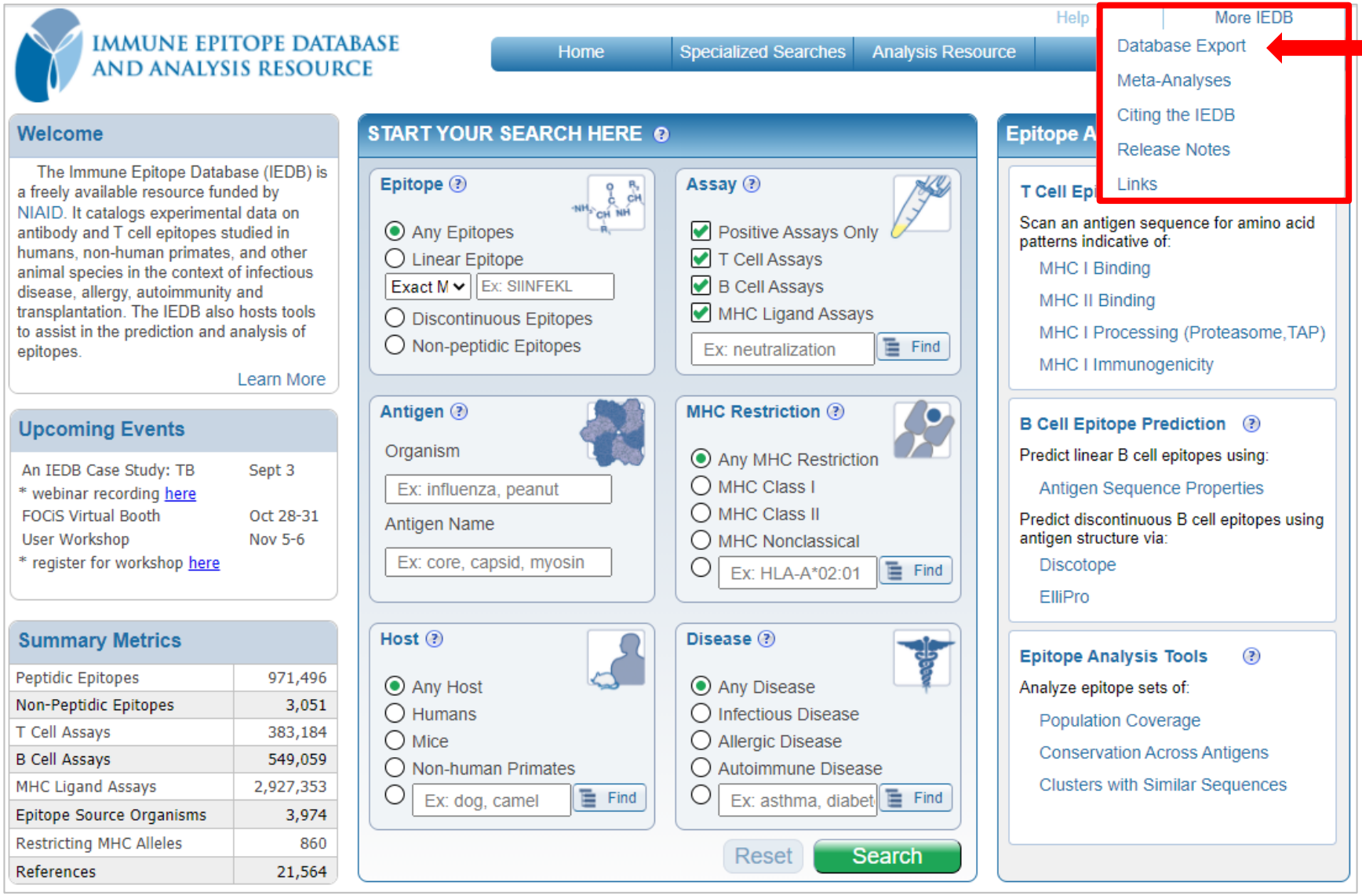


Tutorials and Reference Materials

Includes links to video tutorials,

- ★ 2019 IEDB User Workshop Presentations
- ★ IEDB User Documentation Release 3
- 2018 IEDB User Workshop Presentations
- Browse by Source Organism (IEDB 3.0)
- Browse by MHC Allele (IEDB 3.0)
- Molecule Finder (IEDB 3.0)
- Allele Finder (IEDB 3.0)
- Home Page Search (IEDB 3.0)
- Understanding Search Results (IEDB 3.0)
- Diseases in IEDB 3.0

More Exports



More Exports

http://www.iedb.org/database_export_v3

XML Database Export	
Complete Database Export	323MB
IedbAccessionList.zip	49kB
MhcAlleleNameList.zip	34kB
OrganismList.zip	35MB
AssayTypeList.zip	5kB
GeoLocList.zip	3kB

IEDB Schema	
Curation.xsd (Primary IEDB schema)	48kB
CurationSimpleTypes.xsd	303kB
IedbAccessionList.xsd	909B
MhcAlleleNameList.xsd	1kB
OrganismList.xsd	751B
AssayTypeList.xsd	771B
GeoLocList.xsd	642B

MySQL Database Export	
SQL Statement Export	379MB
MyISAM Binary Export	819MB

Physical Entity Relationship Diagram	
iedb_public_erd.pdf	31kB

CSV Metric Exports	
epitope_full_v3.zip	65MB
antigen_full_v3.zip	2MB
tcell_full_v3.zip	28MB
bcell_full_v3.zip	36MB
mhc_ligand_full (single_file.zip) (multi_file.zip)	168MB
reference_full_v3.zip	12MB
receptor_full_v3.zip	5MB
iedb_3d_full.zip	3MB

Stable ID Exports	
Linkout IDs Compact	21MB
Linkout IDs Full	40MB
ChEBI to Epitope ID Maps	158kB
PDB to Epitope ID Maps	51kB
Epitope ID to Linear Seq Maps	11MB

Receptor_full Export

- Receptor_full contains exports of all antibody and T cell sequences in the IEDB
- Provides nucleotide and protein full length sequences
- Includes CDR1, 2, and 3 sequences
- Includes gene usage
- Listed by each epitope that they were shown to recognize

Iedb_3d_full Export

- iedb_3d_full contains exports of all 3D structures in the IEDB
- Antibody, MHC, and T cell are in separate spreadsheets
- Provides structural details on receptor-Antigen interactions (inter-molecular contacts)
- Include PDB IDs and resolutions of complexes, PDB chain IDs of antibody, TCR, MHC, and antigen chains
- Includes full length receptor and antigen sequences
- Can be combined with receptor export to get information on CDRs and VDJ gene usage

After the Break

- How to query the data in the IEDB
- Your submitted questions

But first…