La Jolla Institute FOR ALLERGY AND IMMUNOLOGY | Life Without Disease.

The Immune Epitope Database Analysis Resource:

Tool downloads
Epitope prediction benchmark references and datasets
Contributing tools to the Analysis Resource

Sinu Paul

spaul@lji.org

IEDB user workshop 2018

Oct 22-23, 2018

Tool downloads

IEDB Analysis Resource

Overview | T Cell Tools | B Cell Tools | Analysis Tools | Tools-API | Download | Datasets | Contribute Tools | References

IEDB Tools Downloads

Complete Download: IEDB Analysis Resource Virtual Machine Image

For users that would like to run the entire analysis resource locally, a virtual machine image file is available with a paid commercial license. The image is kept in sync with the current version of the IEDB Analysis Resource and is updated on a six month cycle. Please contact us for details on licensing options.

Standalone Downloads

Many of the tools hosted on the IEDB-AR are available as command-line tools. They are freely available to academic users through an open source license. Please contact us to inquire about a commercial license or if you have any questions in general. For a complete list of standalone tools, including previous versions, please click here.

Linear B cell epitope predictor
This allows for scoring of amino acid residues using the 6 scale-based methods of the linear B cell epitope prediction tool.
- Linear B cell epitope predictor

MHC class I & II epitope predictors
For users with batch processing needs, the MHC class I and II binding prediction tools are available as standalone scripts for download. These command line tools are kept in sync with the web tools and should therefore produce the same results as clicking through the web interface.
- MHC Class I
- MHC Class II

ElliPro
This will produce the same output as the web version of the tool but does not include the initial template finding functionality. It accepts either a PDB ID or a path to a PDB file as input. Usage instructions are printed by calling the program without any parameters.
- ElliPro

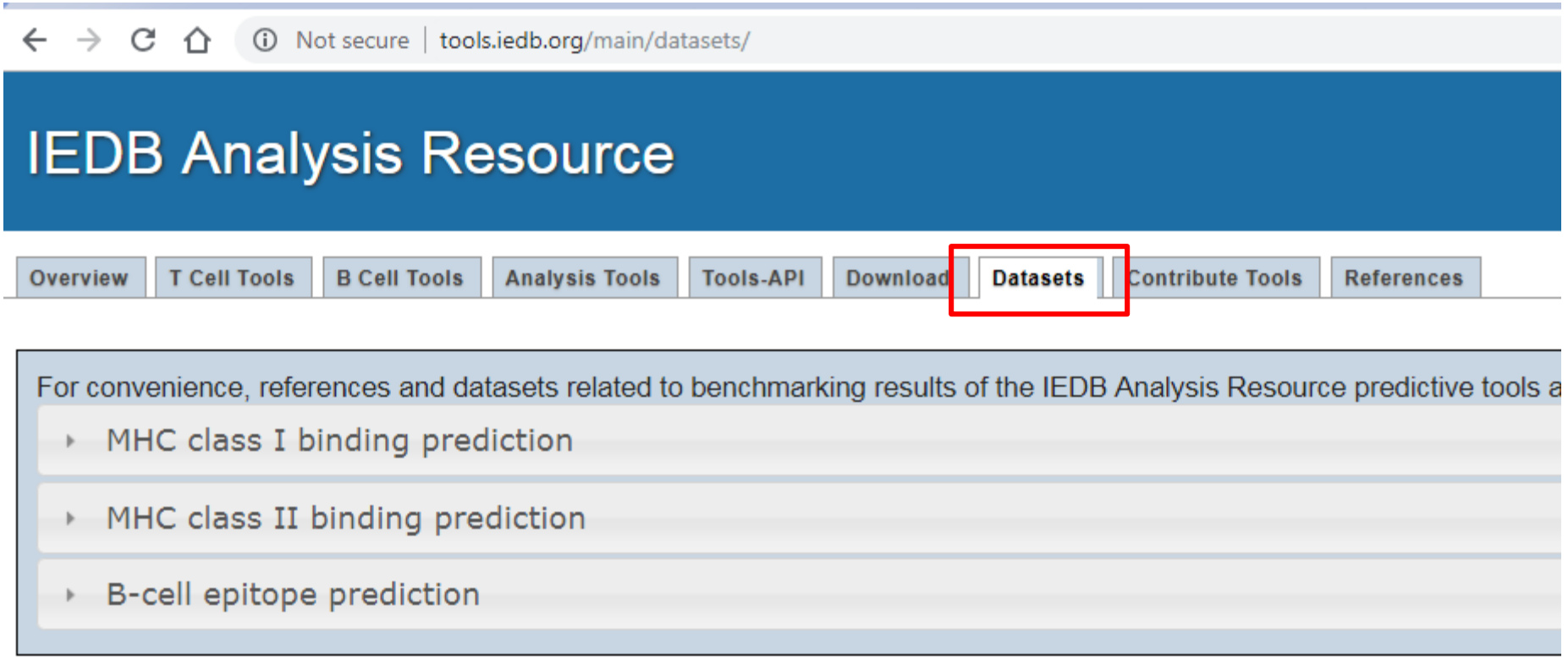
Epitope Cluster Analysis
This tool groups epitopes into clusters based on sequence identity. A cluster is defined as a group of sequences which have a sequence similarity greater than the minimum sequence identity threshold specified.
- Epitope Cluster Analysis

Class I Immunogenicity
This tool uses amino acid properties as well as their position within the peptide to predict the immunogenicity of a peptide MHC (pMHC) complex.
- Class I Immunogenicity

Proteasomal Cleavage Prediction
The distributions 'IEDB_NetChop-1.0.tar.gz' contains methods for NetChop, NetCTL and NetCTLpan.
NetChop is a predictor of proteasomal processing based upon a neural network. NetCTL is a predictor of T cell epitopes along a protein sequence. It also employs a neural network architecture. NetCTLpan is an update to the original NetCTL server that allows for prediction of CTL epitope with restriction to any MHC molecules of known protein sequence.
- Proteasomal Cleavage Prediction

Population Coverage
This tool calculates the fraction of individuals predicted to respond to a given epitope set on the basis of HLA genotypic frequencies and on the basis of MHC binding and/or T cell restriction data. HLA allele genotypic frequencies were obtained from Allele Frequency database.
- Population Coverage

Benchmark Datasets


Not secure | tools.iedb.org/main/datasets/
IEDB Analysis Resource
Overview
T Cell Tools
B Cell Tools
Analysis Tools
Tools-API
Download
Datasets
Contribute Tools
References
For convenience, references and datasets related to benchmarking results of the IEDB Analysis Resource predictive tools a
MHC class I binding prediction
MHC class II binding prediction
B-cell epitope prediction

IEDB Analysis Resource

Overview | T Cell Tools | B Cell Tools | Analysis Tools | Tools-API | Download | **Datasets** | Contribute Tools | References

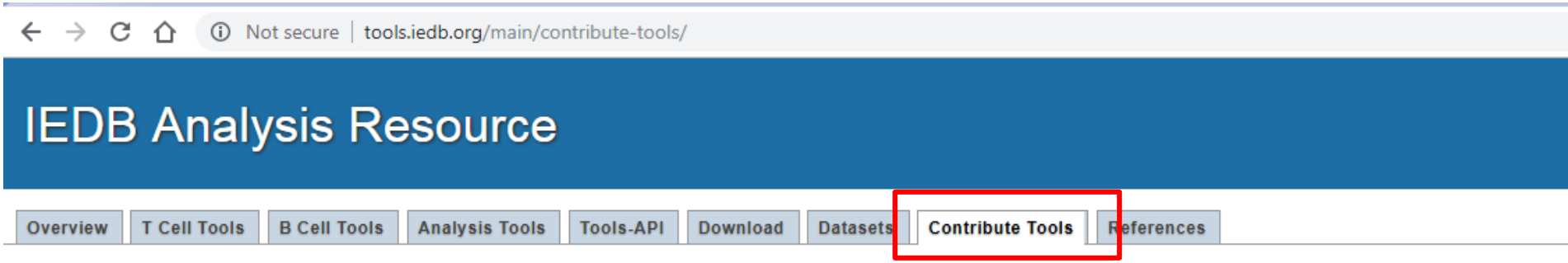
For convenience, references and datasets related to benchmarking results of the IEDB Analysis Resource predictive tools are collected here.

MHC class I binding prediction

- Automated benchmarking of peptide-MHC class I binding predictions
  Trolle T, Metushi IG, Greenbaum JA, Kim Y, Sidney J, Lund O, Sette A, Peters B, Nielsen M.
  Bioinformatics
  - **Description:**
    Numerous in silico methods predicting peptide binding to major histocompatibility complex (MHC) class I molecules have been developed over the last decades. However, the multitude of available prediction tools makes it non-trivial for the end-user to select which tool to use for a given task. To provide a solid basis on which to compare different prediction tools, we here describe a framework for the automated benchmarking of peptide-MHC class I binding prediction tools. The framework runs weekly benchmarks on data that are newly entered into the Immune Epitope Database (IEDB), giving the public access to frequent, up-to-date performance evaluations of all participating tools. To overcome potential selection bias in the data included in the IEDB, a strategy was implemented that suggests a set of peptides for which different prediction methods give divergent predictions as to their binding capability. Upon experimental binding validation, these peptides entered the benchmark study.
  - **Links:**
    Weekly results
    Participate
- Dataset size and composition impact the reliability of performance benchmarks for peptide-MHC binding predictions.
  Kim Y, Sidney J, Buus S, Sette A, Nielsen M, Peters B.
  BMC Bioinformatics
  - **Description of the dataset:**
    1) All binding data used in the paper: BD2009, BD2013, and Blind.

- Dataset size and composition impact the reliability of performance benchmarks for peptide-MHC binding predictions.
  Kim Y, Sidney J, Buus S, Sette A, Nielsen M, Peters B.
  BMC Bioinformatics
  - **Description of the dataset:**
    1) All binding data used in the paper: BD2009, BD2013, and Blind.
    2) For BD2009 data set, three cross-validation data partitions were generated: cv_rnd, cv_sr, and cv_gs.
    3) FILE_S1: Prediction performances for SMMPMBEC, NetMHC, and NetMHCpan against cv_rnd, cv_sr, cv_gs, and Blind data sets. An R script that constructs logistic regression models of deviations (i.e. |cv - blind|) is also included.
  - **Date of the dataset generation:** 2014
  - **Details on the dataset generation:** BD2009 and BD2013 refer to MHC-I binding data files compiled in 2009 and 2013. Blind data sets refer to data resulting after subtracting BD2009 from BD2013. In the paper, different cross-validation strategies (i.e. cv_rnd, cv_sr, and cv_gs) were tested. Please see the Methods section for details of the cross-validation strategies.
  - **Data format:** Text file format.
  - **Dataset availability:** benchmark_reliability.tar.gz
- Dataset used for retraining the IEDB class I binding prediction tools.
  - **Description of the dataset:**The dataset is largely identical to that of Kim et al (2014), described above, but includes additional data that was not publicly available at the time.
  - **Date of the dataset generation:**2013
  - **Details on the dataset generation:**The dataset was compiled from three sources: the IEDB, the Sette lab, and the Buus lab. If a peptide/allele combination had more than 1 measurement among the three sources, its geometric mean was taken.
  - **Data format:** Compressed text file containing binding data.
  - **Dataset availability:** binding_data_2013.zip
- Derivation of an amino acid similarity matrix for peptide: MHC binding and its application as a Bayesian prior.
  Kim Y, Sidney J, Pinilla C, Sette A, Peters B.
  BMC Bioinformatics, 2009.
  - **Description of the dataset:** Cross-validated predictive performances for SMMPMBEC using the same binding data set as in [Peters et al. PLOS Comput Biol 2006].
  - **Date of the dataset generation:** 2009
  - **Details on the dataset generation:** Using the same cross-valiation data partitions as was done for ANN and ARB in 2006, cross-validated predictions using SMMPMBEC were made.
  - **Data format:** A table in Excel file format.
  - **Dataset availability:** http://www.biomedcentral.com/1471-2105/10/394/additional
- A Community Resource Benchmarking Predictions of Peptide Binding to MHC-I Molecules.
  Peters B, Bui HH, Frankild S, Nielsen M, Lundegaard C, Kostem E, Basch D, Lamberth K, Harndahl M, Fleri W, Wilson SS, Sidney J, Lund O, Buus S, Sette A.
  PLOS Computational Biology, 2006.

Contributing tools to the IEDB



Contribute tools to the IEDB-AR

One of the overarching goals of the IEDB is to be the central repository for tools that are of general use to the Immunology and Immunoinformatics communities. As such, we encourage developers of such tools to contact us to inquire about hosting your tool at the IEDB. The IEDB team would work with the developers to create a web portal and keep it up and running indefinitely. We believe this arrangement benefits all parties involved and the Immunology community as a whole. The process for submitting your tool for inclusion at the IEDB-AR is outlined below

Tool contribution process

1. Send an email to help@iedb.org and include the following information:
   - A summary of the problem that is addressed by your tool and why it is of general interest.
   - The publication status of your tool.
   - If there is a web server that currently hosts your tool, please provide the URL.
   - The time frame in which you will be ready to hand off your tool to IEDB developers.
2. Submissions will be evaluated by IEDB staff to determine whether the tool fits within the scope of the IEDB and we have the capability (hardware, personnel, etc.) to implement it.
3. You will receive a reply within 2 weeks with either a decision or a request for further information.
4. Once your tool is approved for inclusion, you will work with IEDB developers to hand off code and create a web portal at the IEDB.
5. The tool will be thoroughly tested for bugs and the load it exerts on the IEDB servers.
6. After you give the go-ahead, links will be made public and it will be officially announced in the IEDB Newsletter as well as the IEDB-AR release notes. It will also be referenced in any future publication on the general capabilities of the IEDB-AR, (e.g., the annual NAR webserver issue).
7. Finally, any updates you make to the tool can be applied, tested, and released in our 6-month development cycle.

Thanks!