HLA-HD download page

HLA-HD (HLA typing from High-quality Dictionary) can accurately determine HLA alleles with 6-digit precision from NGS data (fastq format).
HLA-HD is freely available for academic use and research purposes upon registration.

Released versions

#Version 1.7.0 February 9, 2023
Implement extraction step of read pairs overlap with HLA exons (increase mapping speed for WGS data).
Increase of the calculation speed in hla_est (same speed with version 1.5.0).
Reduction of the required memory usage (approximately 75% memory size down compared with version 1.6.1).

#Version 1.6.1 January 5, 2023
HLA-DQB2 were added in IMGT/HLA after release 3.50.0. Therefore, decoy sequences of HLA-DQB2 have been removed from update.dictionary.sh.
Deal with result of "not consistent" (output as "Not typed").

#Version 1.6.0 December 27, 2022
Improve read score calculation using intron mapped reads (this change increases memory and computational costs).
Prepare HLA_gene.split.txt file (HLA_gene.split.3.50.0.txt) suit to new IMGT/HLA database (release 3.50.0).
Decoy sequences (HLA-DQB2 and HLA-U) are added in update.dictionary.sh.

#Version 1.5.0 June 22, 2022
Bug fix (Correct the disconcordance of allele pair scores because of a rounding error).

#Version 1.4.0 March 10, 2021
Accelerate typing spped and memory reduction in hla_est.

#Version 1.3.0.2 February 10, 2021
Bug fix (memory allocation error in hla_est).

#Version 1.3.0 July 22, 2020
Accelerate typing spped of hla_est.
Correct bug in pm_extract.

#Version 1.2.1 June 26, 2019
Available to input gz compressed fastq files (zcat is needed).

#Version 1.2.0.1 July 11, 2018
Correct bug that hlahd outputs incorrect positions to read.txt in some genes (DRB6,DRB8,DRB9).

#Version 1.2.0 June 25, 2018
Modify default dictionary to type HLA-DRB5 and add some genes to HLA_gene.split.txt (HLA-DPA2, -T, -W, -Y were added to HLA_gene.split.3.32.0.txt for current release, see Ruuning)

#Version 1.1.0.1 November 15, 2017
Adapt the reference data of IPD-IMGT/HLA after the release 3.30.0.

#Version 1.1.0 October 02, 2017
The database update feature was implemented (see section Updating the HLA dictionary).

#Version 1.0.0 April 27, 2017

Download

[v1.7.0](#) (latest)

[v1.6.1](#)

[v1.6.0](#)

[v1.5.0](#)

[v1.4.0](#)

[v1.3.0.2](#)

[v1.2.1](#)

[v1.2.0.1](#)   (If you already installed v1.2.0, see [below](#))

[v1.1.0.1](#)

[v1.0.0](#)

Installation

HLA-HD requires [bowtie2](#) to map NGS reads.
Please install bowtie2 on your computer and set path to your environment variables.
For example, if you are using bash, add to your .bashrc the following command.
export PATH=$PATH:/path_to_bowtie2

Uncompress the downloaded tar.gz file by
> tar -zxvf hlahd.version.tar.gz
Then, move to the uncompressed directory and type
> sh install.sh
For the installation, the [g++ compiler by the GNU Compiler Collection](#) must be installed on your computer.

After the installation, add the current directory to your PATH
export PATH=$PATH:/path_to_HLA-HD_install_directory/bin

Update from v1.2.0 to v1.2.0.1 (July 11, 2018)
1. Uncompress the downloaded tar.gz file
2. Replace src folder in current hlahd folder to uncopressed src folder.
3. Move to hlahd directory
4. Type
> g++ ./src/hla_estimation.cpp -O3 -o ./bin/hla_est

If you want to correct read.txt file of already typed result, execute as follows:
1. Move to directory where hlahd was performed.
2. Type
>sh Output.directory/outputID/estimation.sh

Updating the HLA dictionary (after v.1.1.0)

You can update the HLA allele dictionary to the current release of the IPD-IMGT/HLA database by the command,
> sh update.dictionary.sh

Wget is required for the database update.

You can also use any release by getting hla.dat file from the github site.
Put hla.dat file on parent dicretory of hlahd and executing the update.dictionary.sh by deleting the line of the first wget command.

The latest release can adopt the newest rare alleles.
In contrast, the old release tend to yield conservative result.

Default dictionary of the installation is created from release 3.15.0.

Running

Before running the HLA-HD, check the value of open files on your computer by typing:
> ulimit -Sa
If open files are less than 1024, please type:
> ulimit -n 1024
or change /etc/security/limits.conf according to your system environment.

If you have fastq.gz file, unzip gz file in advance.

You can run the HLA-HD by typing the following commands:

> hlahd.sh -t [thread_num] -m [minimum length of reads] -c [trimming rate] -f [path_to freq_data directory] fastq_1 fastq_2 gene_split_filt path_to_dictionary_directory IDNAME[any name] output_directory

For example:

> hlahd.sh -t 2 -m 100 -c 0.95 -f freq_data/ data/sample_1.fastq data/sample_2.fastq HLA_gene.split.txt dictionary/ sampleID estimation

If you want to type HLA-DPA2, -T, -W, -Y, replace HLA_gene.split.txt
to HLA_gene.split.3.32.0.txt and update the dictionary to current. (after v.1.2.0)

Options

-m : A read whose length is shorter than this parameter is ignored. Default size is 100.

-t : Number of cores used to execute the program.

-c : Trimming option. If a match sequence is not found in the dictionary, trim the read until some sequence is matched to or reaches this ratio. Default is 1.0. We recommend 0.5 for 300×2 base-paired end-read data.

-f : Use information of allele frequencies. The default data exist in the installed directory (/hlahd.version/freq_data).

Reference

Kawaguchi, S. *et al*. "Comprehensive HLA Typing from a Current Allele Database Using Next-Generation Sequencing Data", *Methods Mol Biol*., 2018;1802:225-233, doi: 10.1007/978-1-4939-8546-3_16, 2018.

Kawaguchi, S. *et al*. "HLA-HD: An accurate HLA typing algorithm for next-generation sequencing data" *Hum Mutat*., Jul;**38**(7):788-797, doi: 10.1002/humu.23230, 2017.

Contact:

Shuji Kawaguchi: shuji@genome.med.kyoto-u.ac.jp