ProteinDigester

Table of contents

1. Overview

This application digests fasta proteins and return peptide digests (sequence, molecular weight and isoelectric point).

2. Usage

```
usage: ProteinDigester <Fasta> [-a] [-d <arg>] [-e <arg>] [-f <arg>] [-h]
       [-i <arg>] [-l <arg>] [-m <arg>] [-o <arg>] [-p <arg>] [-q] [-v]
 -a,--average                      set the average mass accuracy for
                                   peptide mass calculation
                                   by default: monoisotopic.
 -d,--delimiter <arg>              define the field delimiter to display
                                   by default: \t.
 -e,--enzymes <arg>                define enzymes that digest proteins
                                   separately with:
                                   enzyme name among 'Caspase-1, Caspase-10,
Glu-C_bicarbonate, Caspase-3,
                                   Thermolysine, Lys-C, Pepsin_pH1.3,
Caspase-8, Glu-C_phosphate, Caspase-9,
                                   Caspase-5, CNBr, ChymoTrypsin_lowspec,
Trypsin, ChymoTrypsin_FYL,
                                   Proteinase-K, Caspase-7, Pepsin_pHgt2,
ChymoTrypsin_highspec,
                                   BNPS-Skatole, Enterokinase, Caspase-6,
Arg-C, Asp-N, ChymoTrypsin_FYLW,
                                   Caspase-4, Caspase-2'.
                                   or custom motifs respecting the following
grammar:
                                                                  <pre-cut>
<cut-token> <post-cut>
                                   <cut-token> := '|'
                                   <pre-cut> := (<AA> or <AA-class>)+
                                   <post-cut> := (<AA> or <AA-class>)+
                                   <AA> := [A-Z]
                                   <AA-class> := '[' AA+ ']'
                                   by default: Trypsin.
 -f,--fields <arg>                 define the fields to display
                                   (1:Sequence, 2:MW, 3:pI, 4:Enz, 5:MC)
                                   by default: [1, 2, 3].
 -h,--help                         print this message.
 -i,--setting-file <arg>           give a property file with all input
                                   settings.
 -l,--length <arg>                 define a filter over length of digested
                                   peptides
                                   by default: 5.
 -m,--max-missed-cleavage <arg>    define the number of maximum missed
                                   cleavages (for digestion)
                                   by default: 1.
 -o,--output <arg>                 set the output filename
```

```
-p,--precision <arg>                define the number of fractional digits
                                    for output
                                    by default: 6.
-q,--quiet                          quiet mode (verbose off)
                                    by default: false.
-v,--version                        print the version info.
```

3. Example

```
# a first way to execute the application with lots of options
$ java -jar ProteinDigester.jar -a -e Lys-C -f 1,2,4,5 -l 6 -p4 -q -o
/tmp/dig.tsv uniprot-human.fasta

# .. or the more compact way with all options defined in a setting file
$ java -jar ProteinDigester.jar -i settings.properties uniprot-human.fasta
```

4. Releases

The latest version is v1.22 - Download app.

4.1. Rel1.22

Type	Summary
Bug fixed	Fasta header manager bug fixed

4.2. Rel1.2

Type	Summary
Bug fixed	--max-missed-cleavage feature was not working properly

4.3. Rel1.1

Type	Summary
Bug fixed	parameter "output" defined in "your-params.properties" was not well set in ProteinDigester *-i "your-params.properties"* <fasta>.