
DamageProfiler Documentation

Release 1.0

Judith Neukamm, Alexander Peltzer

Aug 04, 2020

Contents:

1 Prerequisites for the installation of DamageProfiler	3
1.1 Operating System Support	3
1.2 Software Requirements for DamageProfiler	3
2 Installation Instructions for DamageProfiler	5
3 General Usage	7
3.1 How to run	7
3.2 Log file	8
4 Indices and tables	9

This is the main DamageProfiler documentation, where you can find information about the prerequisites, the installation and the usage of this tool.

Prerequisites for the installation of DamageProfiler

1.1 Operating System Support

DamageProfiler has been implemented as a platform-independent tool and can thus be installed and run on Linux, Windows and MacOS. A Java 8+ platform has to be installed on the workstation used for running the tool.

1.1.1 Linux

It has been successfully tested on different flavors of Linux based operating systems, including ArchLinux and Ubuntu 17.04.

1.1.2 Windows

The application can be run on Windows 10.

1.2 Software Requirements for DamageProfiler

Install a suitable Java 8 runtime environment. We tested both Oracle Java 8 SE and OpenJRE 8 on Linux. The latter requires to install JavaFX in addition, whether the Oracle JRE already ships with the required JavaFX libraries.

CHAPTER 2

Installation Instructions for DamageProfiler

The tool can be downloaded from [DamageProfiler's GitHub page](#) . After downloading the JAR file, you can start the application via double click on most operating systems (OSX, Windows and Linux). If not, please [install Java](#) on your workstation.

DamageProfiler can be used to calculate damage profiles of mapped ancient DNA reads given as SAM/BAM file and provides a graphical as well as text based representation.

It creates

- damage plots
- a fragment length distribution
- a read identity distribution
- a base frequency table of reference
- a table of different base misincorporations and their occurrences

3.1 How to run

```
java -jar DamageProfiler-VERSION.jar <arguments> [options]
```

Arguments:

- **-h,--help** : Shows this help page.
- **-i,--input <INPUT>** : The input sam/bam file.
- **-o,--output <OUTPUT>** : The output folder

Options:

- **-l,--length <LENGTH>** : Number of bases which are computations.
- **-mapped,--all_mapped_reads** : Use all mapped reads to calculate damage plots. Default: false.
- **-r,--reference <REFERENCE>** : The reference file. The reference is only needed if the mapping file doesn't have MD tags.

- **-s,--specie <SPECIE>** : RefSeq ID of the reference genome. This will just use the reads mapping to the specific reference for creating the damage profile. It can for example be used for MALT output files. In this case, please make sure that you run MALT without *-sparseSAM* option. This would create a sam file with is not readable. Example: `-s NC_022116.1`
- **-sf,--specieslist file <SPECIES LIST>** : List with species (RefSeq IDs) for which damage profile has to be calculated. One species per line, given as text file (.txt). Species must have the same format like *-s* parameter.
- **-t,--threshold <THRESHOLD>** : Number of bases which are considered for plotting nucleotide misincorporations.
- **-title,--title <TITLE>** : Title used for all plots (Default: file name of input SAM/BAM file).
- **-yaxis,--yaxis <YAXIS>** : Maximal value on y axis (Default: flexible, adapts to the calculated damage).

DamageProfiler can be used in offline mode. Running the jar file without any parameter starts a `org.damageprofiler.GUI` to configure the run.

3.2 Log file

DamageProfiler documents the configuration in a separate log file, which helps you to reconstruct your analysis at a later date. The file is saved in the specified result folder.

CHAPTER 4

Indices and tables

- `genindex`
- `modindex`
- `search`