**Manual for AncFil**

Version 1.0

1. Introduction

The main function of the AncFil is to rescue the highly homologous-contaminated ancient samples based on the post-mortem damage patterns. There are several key parameters that will influence the filtering accuracy including the length of ancient DNA, the depurination characteristic and the deaminated C-to-T and G-to-A changes at ends of ancient DNA fragments. We strongly recommended to screen reads with at least one C-to-T or G-to-A changes within the first or last 15 bp at 3’ or 5’ ends for this script.

1. Requirements
2. Python >=3.7.7
3. Python package pyfaidx
4. Parameters

|  |  |  |
| --- | --- | --- |
| Parameters | Parameter Type | Description |
| -i/--input | string | Sam file generated by BWA |
| -o/--output | string | The output sam file after filtering. |
| -r/--reference | string | Reference genome with index built ( .fai format file). |
| -m/--mode | string | Filtering mode based on depurination or deamination  (-m: depurination/deamination) |
| -DeamNum | num | Screening reads with at least “-DeamNum” C-to-T or G-to-A changes at ends of DNA fragments |
| -DetectRange | num | Screening reads with C-to-T or G-to-A changes within the first or last “-DetectRange” base pair |
| -DoubleOrSingle | string | Screening reads with C-to-T or G-to-A changes at 3’ and/or 5’ ends. (-DoubleOrSingle: and/or) |
| -t | num | Number of threads [5] |
| -h | string | Show the help message |

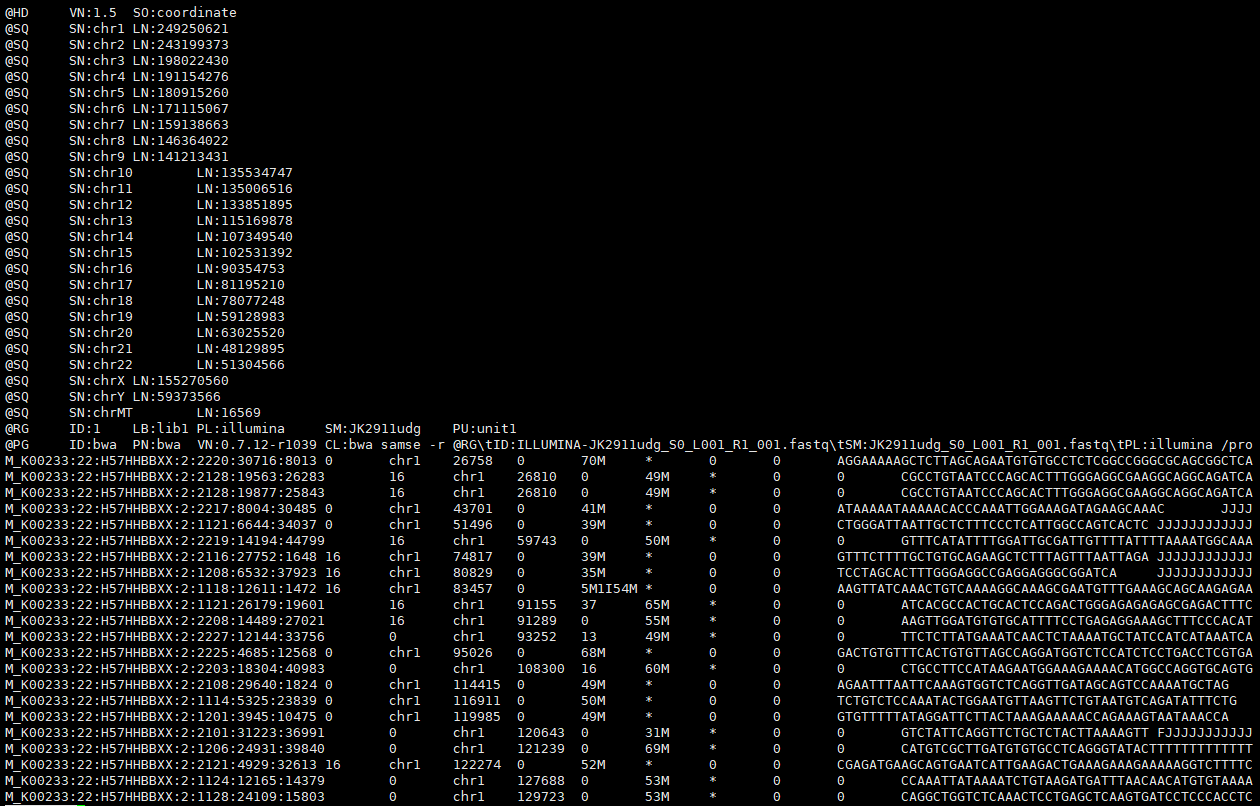
1. Example:
2. Screening based on deamination patterns:

python AncFil.py -i path/test.sam -o path/output.sam -r reference\_path/ref.fa -m deamination -DeamNum 1 -DetectRange 15 -DoubleOrSingle or

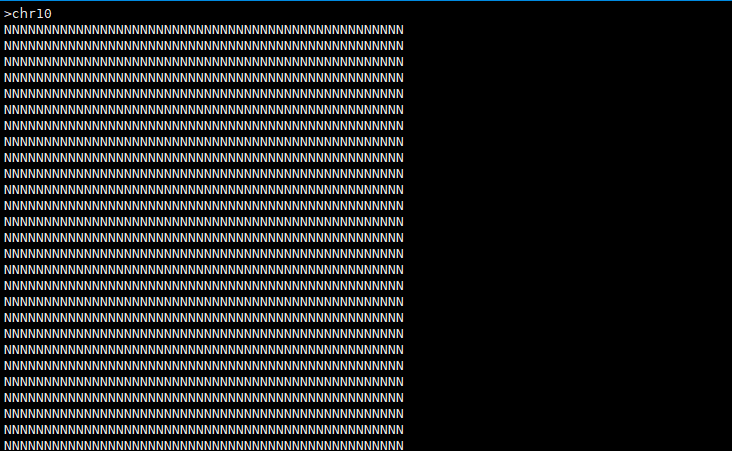
1. Screening based on depurination patterns:

python AncFil.py -i path/test.sam -o path/output.sam -r reference\_path/ref.fa -m depurination

1. The format of test.sam/output.sam: This is a normal Sam file with a format as following:



1. ref.fa: This is a reference genome file with the normal Fasta format:



1. Output files:

The AncFil will finally generate 1 output file:

1. Output.sam: this is a sam file without headers:

