

deMultiplexing

Nate Olowo

October 14, 2018

Demultiplexing Assignment

Code developed for paired end reads in which reads and indexes are in separate files. The following scripts will output write statistics and a graph of relative sample abundances for the demultiplexing.

Usage DeMultiAlgoArg.py [options]

Input flags

Required

-r1 [readfile1]

-i1 [indexfile1]

-r2 [readfile2]

-i2 [indexFile2]

Optional

-t [mean quality score threshold] *Default: 25*

Expected output from Test files

Write statistics

Records Read: 9

Matched Reads: 2

Below threshold: 1

unmatched indices: 2

bad reads(Ns): 3

Matched barcodes above threshold not found in indexes file: 1

Statistics defined

Records Read: Total number of reads processed in each file. All four files should have the same amount of lines, if not, the script will only read until your shortest file.

bad Reads: N found in barcode of either of the 2 index reads

Below Threshold: Average quality scores of each index read does not surpass threshold. No N's in barcode though

unmatched indices: Indexes don't match. Passes threshold and N tests stated above.

___Not found in indexes file:___ Index of forward read not found in indexes.txt input file. Passes all tests above.

Matched Reads: Passes all above test. Read written to respective sample with forward and reverse written separately.

```
## Parsed with column specification:
## cols(
##   Reads = col_double(),
##   Sample = col_character()
## )
```

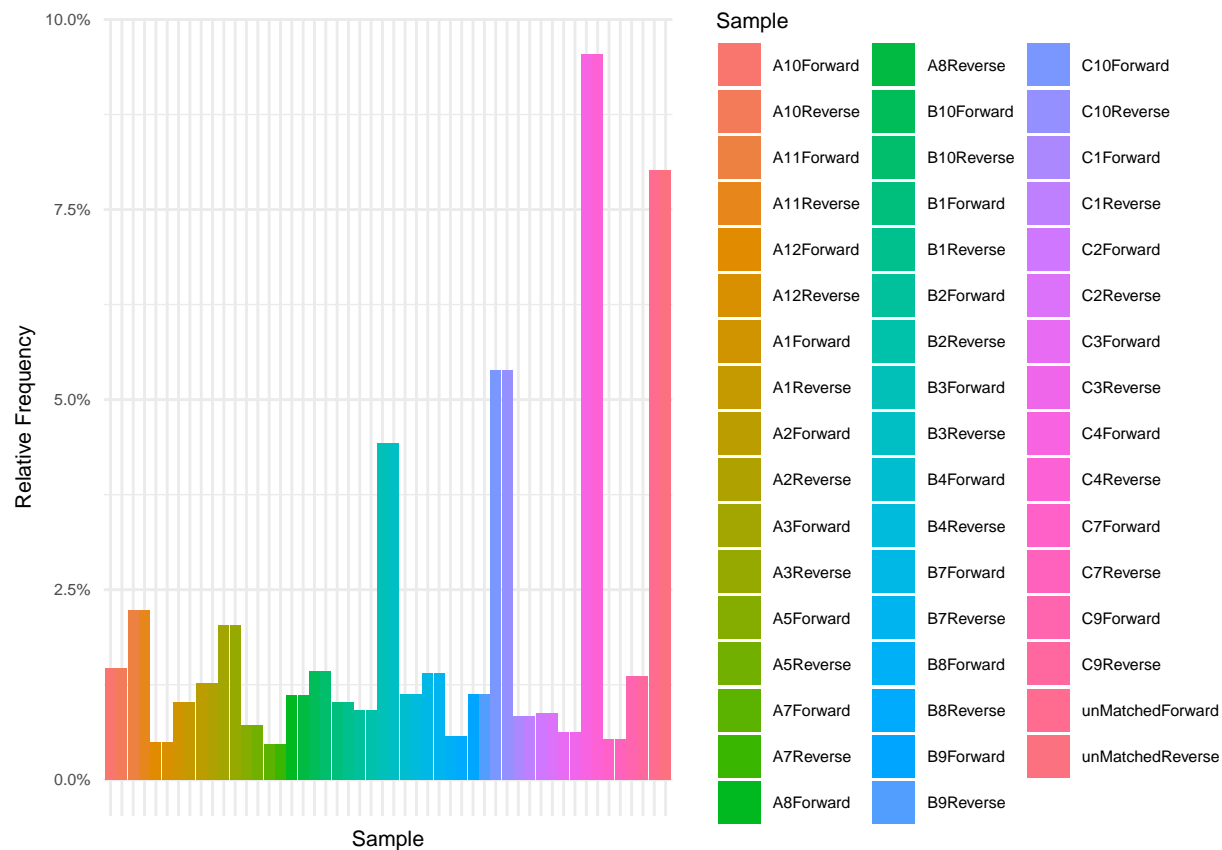


Figure 1: Relative abundance of reads mapping to each sample.