#Commands for CloudBrush de novo genome assembler

#These are terminal shell commands

# https://github.com/ice91/CloudBrush

# failed on developer’s own sample data

#developers’ data

#convert fast to sfa

$ perl data/preprocessor.pl -renum data/Ec10k.sim.fastq > data/Ec10k.sim.sfa

#upload to HDFS

$ hadoop fs -put data/Ec10k/sim.sfa Ec10k

#CloudBrush alignment

$ hadoop jar CloudBrush.jar -reads Ec10k -asm Ec10k\_Brush -k 21 -readlen 36

########################################

#On teams’ Human micro biome data; also fails eventually##

# created .sfa:

$ perl data/preprocessor.pl -renum data/SRS019033/SRS019033.denovo\_duplicates\_marked.trimmed.singleton.fastq > data/SRS019033/SRS019033.denovo\_duplicates\_marked.trimmed.singleton.sfa

# put sfa in HDFS

# hadoop fs -put data/SRS019033/SRS019033.denovo\_duplicates\_marked.trimmed.singleton.sfa SRS019033.denovo\_duplicates\_marked.trimmed.singleton

# run command

$ hadoop jar CloudBrush\_ORIGINAL.jar -reads SRS019033.denovo\_duplicates\_marked.trimmed.singleton -asm SRS019033\_Brush -k 21 -readlen 36