Galaxy hands-on for NGS data analysis

Andreas Gisel, IITA - Ibadan, Nigeria

Introduction to Bioinformatics using the eBioKit Platform

Commission for Biotechnology & Genetic Engineering, National Center for Research
December 1 - 6, 2014

Exercise 1: Map cleaned Illumina reads against a reference sequence

1. upload the read file 'mac-18_noad.fastq'
2. upload the reference file 'peach_chloro.fa'
3. convert the score system
4. check the quality of the read sequences (statistics and graphics)
5. map the reads onto the reference sequence
6. convert from bam to sam file
7. export sam file
8. visualize result with 'table'

Exercise 2: Clean and map Illumina reads to the host and pathogen reference sequence.

1. upload the read file 'mac18red.fastq'
2. upload the reference file 'peach_chloro.fa'
3. upload the reference file 'PLMV_PC-C40.fasta'
4. convert the score system
5. check the quality of the read sequences (statistics and graphics)
6. trim adaptor sequence
7. map the reads onto the reference sequences
8. convert from bam to sam file
9. export sam file
10. visualize result with 'table'