Lab meeting Illumina pipeline demo instruction sheet

March 20, 2017

1. Git clone the repository to access the scripts.

cd ~/Desktop

git clone https://bitbucket.org/tcf-lab/illumina\_pipeline.git

2. Set add the folder “illumina\_pipeline” to your path. This will allow you to run the script without being in the folder that contains it.

cd ~

nano .bash\_profile

export PATH=“$PATH:~/Desktop/illumina\_pipeline”

press control + x to exit.

Type “y” to save changes.

3. Navigate back to your Desktop directory where you git cloned the illumina\_pipeline folder and cd into the illumina\_pipeline directory

cd ~/Desktop/illumina\_pipeline

**4. Run the commands that we will execute without the pipeline.**

**A. First, make a folder called “example” and copy (cp) the fastq files, reference sequence into it, and rename them.**

mkdir example

cp \*.fastq example

cp S301\_A\_CAMBODIA\_U0417030\_2010\_S301\_L001/ AJ4MBL718F512\_A\_CAMBODIA\_U0417030\_2010full\_genome.fasta ~/Desktop/lab\_meeting\_example\_170306/example/

cd example/

delete the S303 fastq files

mv S301\_A\_CAMBODIA\_U0417030\_2010\_S301\_L001\_R1\_001.fastq sample\_R1.fastq

mv S301\_A\_CAMBODIA\_U0417030\_2010\_S301\_L001\_R2\_001.fastq sample\_R2.fastq

mv AJ4MBL718F512\_A\_CAMBODIA\_U0417030\_2010full\_genome.fasta ref.fasta

**B. Trim reads using trimmomatic.**

java -jar /usr/local/bin/Trimmomatic-0.36/trimmomatic-0.36.jar SE sample\_R1.fastq sample\_R1.trimmed.fastq SLIDINGWINDOW:5:30 MINLEN:100

**C. Map using bowtie2. Mapping in bowtie2 requires 2 steps: building the reference and performing mapping:**

Run this to build your reference.

bowtie2-build ref.fasta ref

Perform mapping:

bowtie2 -x ref -U sample\_R1.trimmed.fastq,sample\_R2.trimmed.fastq -S sample.sam ---local

**D. Call variants**

Convert the sam file to a sorted bam file:

samtools view -bS sample.sam > sample.bam

samtools sort sample.bam > sample.sorted.bam

**Call variants with LoFreq:**

lofreq call -f ref.fasta –o sample.lofreq.vcf sample.sorted.bam

lofreq filter --cov-min 100 --snvqual-thresh 30 --af-min 0.01 -i sample.lofreq.vcf -o sample.lofreq.filtered.vcf

**Call variants with Varscan:**

First, generate a pileup file:

samtools mpileup -d 100000 sample.sorted.bam > sample.pileup -f ref.fasta

Type this to call variants/generate your vcf file:

java -jar /usr/local/bin/VarScan.v2.3.9.jar mpileup2snp sample.pileup --min-coverage 100 --min-avg-qual 30 --min-var-freq 0.01 --strand-filter 1 --output-vcf 1 > sample.varscan.vcf