```
################################
## ChIP-seq Analysis Workflow ##
################################
## Mireia Ramos-Rodríguez     ##
## 14th of March 2019         ##
################################

###############################################################################
# Set up your working environment and necessary files
###############################################################################
# Open the virtual machine, log in to the VPN and connect to `ironwomen`.
# There, you will need to create the following directory structure.

#------------------------------------------------------------------
# 1) Generate directory structure:
#------------------------------------------------------------------
# yourpersonalfolder
# |-- ChIP-seq
#     |-- fastq
#     |-- BAM
#     |-- peaks

cd yourpersonalfolder
mkdir ChIP-seq
cd ChIP-seq
mkdir fastq
mkdir BAM
mkdir peaks

#------------------------------------------------------------------
# 2) Copy sample file: NL1_h3k27ac.fastq.gz
#------------------------------------------------------------------
cp ~/workshop_ChIPseq/ChIP-seq_sample/fastq/NL1_h3k27ac_sample.fastq.gz
~/yourpersonalfolder/ChIP-seq/fastq/.

###############################################################################
## Quality control with FastQC
###############################################################################
# First, we are going to check the quality of the sequenced reads using FastQC.
# The otuput is an html file with a summary of the tests performed and if your
# reads are passing or not the quality control thresholds.

#------------------------------------------------------------------
# 1) Run FastQC on the sample file
#------------------------------------------------------------------
cd ~/yourpersonalfolder/ChIP-seq # Make sure you are here
mkdir FastQC # Make directory for output
fastqc fastq/NL1_h3k27ac_sample.fastq.gz -o FastQC # 15 seconds

#------------------------------------------------------------------
# 2) To see the results, copy the output html to the virtual machine
#------------------------------------------------------------------
# Open a terminal in the virtual machine
scp msuser@ironwomen:/home/labs/mslab/msuser/yourpersonalfolder/ChIP-
seq/FastQC/NL1_h3k27ac_sample_fastqc.html .
# Now open the hmtl from your virtual machine and explore the results

###############################################################################
## Alignment with Bowtie2
###############################################################################
#------------------------------------------------------------------
# 1) Align the sample file to the reference genome
#------------------------------------------------------------------
bowtie2 -t -x ~/workshop_ChIPseq/reference_genome/hg19 -U fastq/NL1_h3k27ac_sample.fastq.gz
-S BAM/NL1_h3K27ac_sample.sam

###############################################################################
## Post-processing with Samtools
###############################################################################
#------------------------------------------------------------------
# 1) Convert to BAM and sort
```

```
#-----------------------------------------------------------------
samtools view BAM/NL1_h3K27ac_sample.sam -b | samtools sort - -o
BAM/NL1_h3K27ac_sample.srtd.bam

#-----------------------------------------------------------------
# 2) Remove duplicates
#-----------------------------------------------------------------
markdup BAM/NL1_h3K27ac_sample.srtd.bam BAM/NL1_h3K27ac_sample.rmdup.bam -r -s
#-----------------------------------------------------------------
# 3) Index bam files
#-----------------------------------------------------------------
samtools index BAM/NL1_h3K27ac_sample.srtd.bam
samtools index BAM/NL1_h3K27ac_sample.rmdup.bam

#################################################################################
## Peak calling with MACS2
#################################################################################
activate-macs-git-2017.5.15 # Activate MACS2

mkdir tmp/ # Create folder for temporary files

#-----------------------------------------------------------------
# Run peak calling with MACS2
#-----------------------------------------------------------------
macs2 callpeak -f BAM -t BAM/NL1_h3K27ac_sample.rmdup.bam -c ~/workshop_ChIPseq/ChIP-
seq/BAM/NL1_input.rmdup.bam -g hs -n peaks/NL1_h3K27ac --tempdir tmp/ --broad --nomodel
```