```
# Download the reference genome and gene annotation files from NCBI using the provided GCF
(GenBank assembly) accession number.
# Assembly: GCF_000146045.2 (Saccharomyces cerevisiae S288C, R64)
# This sequence corresponds to the S. cerevisiae S288C (R64) reference genome.

# ------------------------------
#  1) create direcotry
# ------------------------------
mkdir ref_genome

# ------------------------------
# 2) Download the genomic DNA sequence in FASTA format.
# ------------------------------
# File extension:  .fna.gz  → compressed FASTA file containing the reference genome sequence

wget
https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/146/045/GCF_000146045.2_R64/GCF_000146045.2_
R64_genomic.fna.gz

# ------------------------------
# 3) Download the gene annotation in GTF format.
# ------------------------------
# File extension: .gtf.gz → compressed GTF file containing gene, transcript, exon coordinates

wget
https://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/146/045/GCF_000146045.2_R64/GCF_000146045.2_
R64_genomic.gtf.gz

# ------------------------------------
# 4) Decompress downloaded files
# ------------------------------------
gunzip *.gz

# Uncompress all .gz files in the current directory.
# Resulting files:
#  GCF_000146045.2_R64_genomic.fna
#  GCF_000146045.2_R64_genomic.gtf

# ---------------------------------------------------------
# Building an RSEM reference for yeast
# ---------------------------------------------------------
#   Building an RSEM reference for yeast using:
#    genome FASTA: GCF_000146045.2_R64_genomic.fna
#    annotation GTF: GCF_000146045.2_R64_genomic.gtf
#   and generate Bowtie2 indices for read alignment and quantification within RSEM.
#   Run this in the directory that contains the .fna and .gtf files.

# ---------------------------------------------------
# 5) Creating a reference genome
# ---------------------------------------------------

# ---------------------------------------------------
# Load Conda
# ---------------------------------------------------

module avail conda    # shows available conda/anaconda modules
module load anaconda3

# ---------------------------------------------------
#  Configure conda channels for bioinformatics installs
# ---------------------------------------------------

conda config --add channels defaults
conda config --add channels bioconda
```

```
conda config --add channels conda-forge


# -----------------------------------------------------------------------------
# Create and activate an environment containing RSEM
# -----------------------------------------------------------------------------

# Create and activate  new conda environment named rsem_env and install the package "rsem"
conda create -n rsem_env rsem
conda activate rsem_env

# -----------------------------------------------------------------------------
# Check that RSEM is installed and report its version
# -----------------------------------------------------------------------------
rsem-calculate-expression --version

# -----------------------------------------------------------------------------
# Load Bowtie2 (aligner used by RSEM for indexing/alignment
# -----------------------------------------------------------------------------

module avail bowtie2
module load bowtie2
#--------------------------------
# Check Bowtie2 version
# ------------------------------
bowtie2 --version

# -----------------------------------------------------------------------------
# Prepare the RSEM reference (build transcript reference + Bowtie2 indices)
# -----------------------------------------------------------------------------

rsem-prepare-reference \
  --bowtie2 \                             # tells RSEM to specifically build Bowtie2 indices
  --gtf GCF_000146045.2_R64_genomic.gtf \  # provides gene annotation in GTF format
  GCF_000146045.2_R64_genomic.fna \        # reference genome FASTA
  yeast_rsem                             # output reference prefix (creates multiple files)

# -------------------------------------------
# What this command produces
# -------------------------------------------
#   RSEM creates a set of reference/index files with prefix "yeast_rsem",
#   including:
#    transcript reference files derived from the genome + GTF
#    Bowtie2 index files needed for alignment-based quantification
```