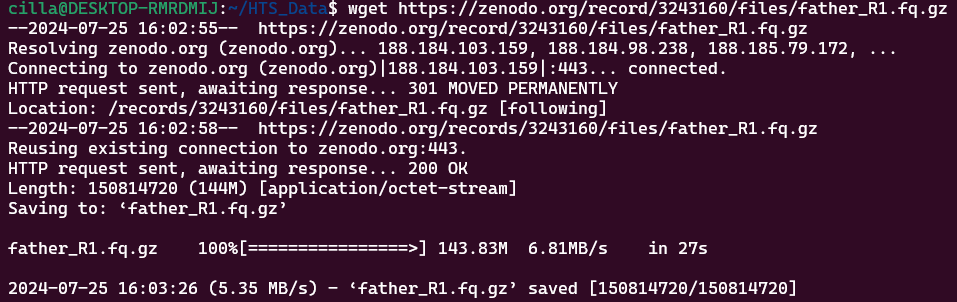
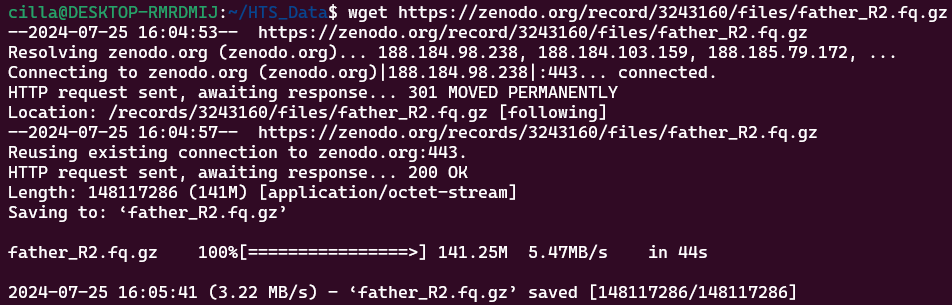
HIGHTHROUGHPUT PROJECT

1. Getting the father dataset

Code used: wget <https://zenodo.org/record/3243160/files/father_R1.fq.gz>

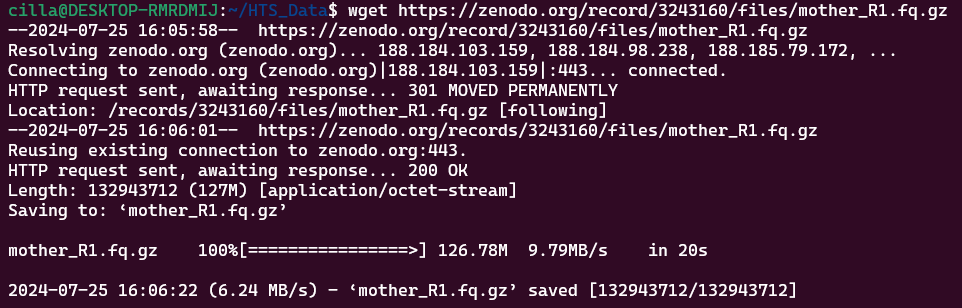


wget https://zenodo.org/record/3243160/files/father\_R2.fq.gz

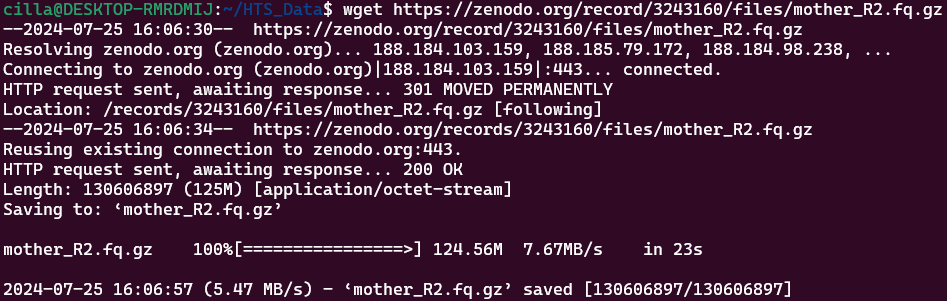


1. Getting the Mother dataset

Codes used: wget <https://zenodo.org/record/3243160/files/mother_R1.fq.gz>

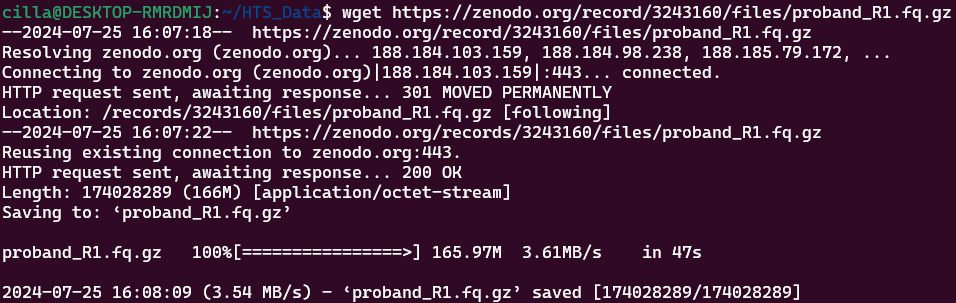


wget https://zenodo.org/record/3243160/files/mother\_R2.fq.gz

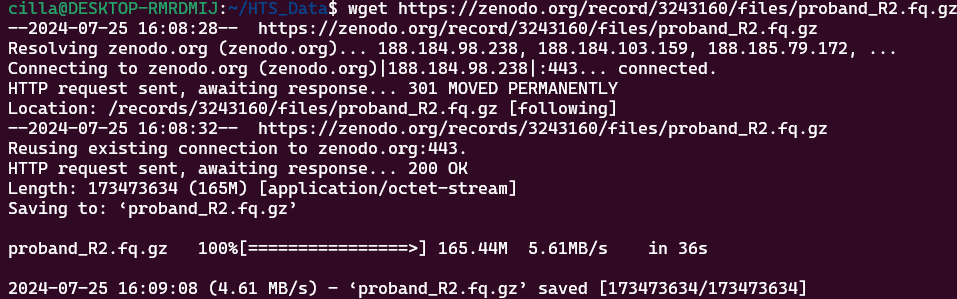


1. Child dataset

Code used: wget https://zenodo.org/record/3243160/files/proband\_R1.fq.gz

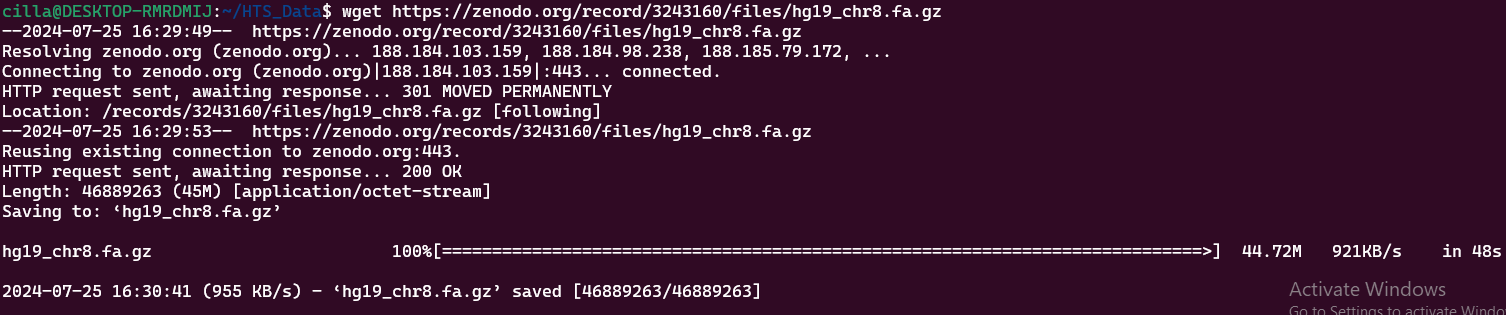


Code used: wget https://zenodo.org/record/3243160/files/proband\_R2.fq.gz



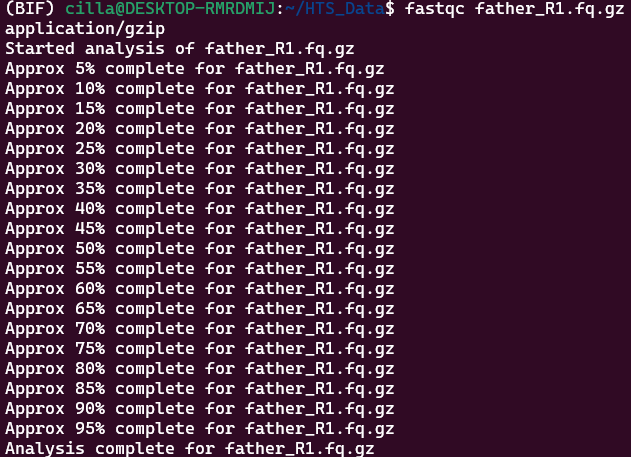
STEP2: GETTING THE hg19 version of the human dataset

Code used: wget <https://zenodo.org/record/3243160/files/hg19_chr8.fa.gz>

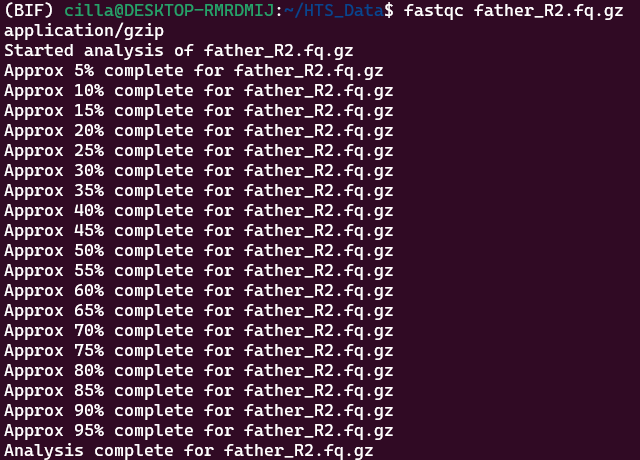


STEP 3

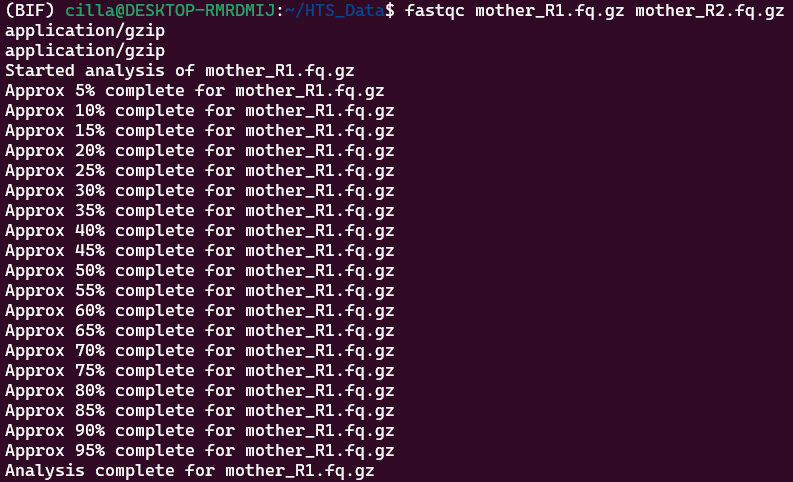
Running Fatsqc on the datasets

Code used: fastqc father\_R1.fq.gz

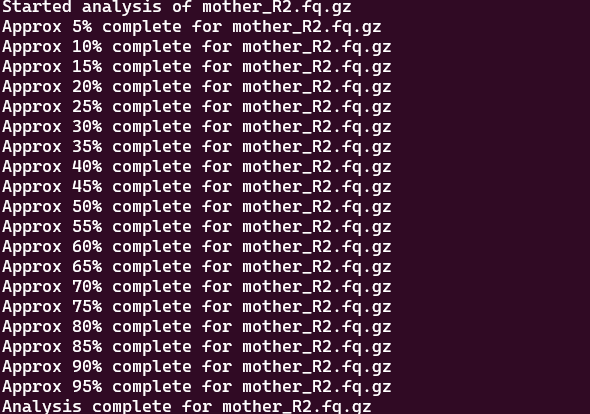
ii. fastqc father\_R2.fq.gz



iii. fastqc mother\_R1.fq.gz



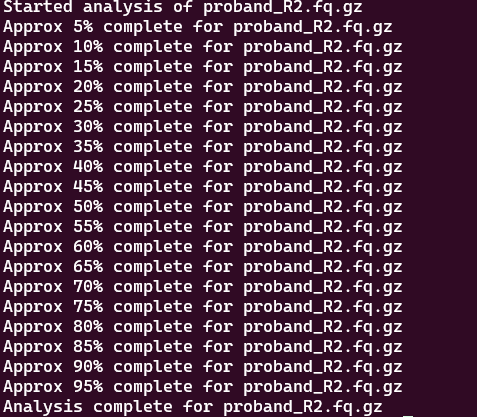
iv. fastqc mother\_R2.fq.gz



v. fastqc proband\_R1.fq.gz



vi. fastqc proband\_R2.fq.gz



STEP 4: INDEXING THE REFERENCE GENOME

Code used:

STEP 5: READ MAPPING

Code used:

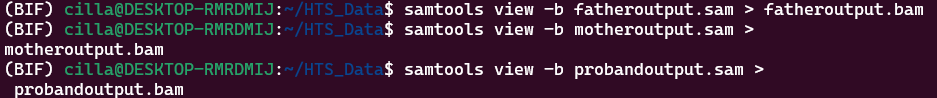
bwa mem -t 8 index/hg19\_chr8.fa.gz father\_R1.fq.gz father\_R2.fq.gz > fatheroutput.sam

bwa mem -t 8 index/hg19\_chr8.fa.gz mother\_R1.fq.gz mother\_R2.fq.gz > motheroutput.sam

bwa mem -t 8 index/hg19\_chr8.fa.gz proband\_R1.fq.gz proband\_R2.fq.gz > probandoutput.sam

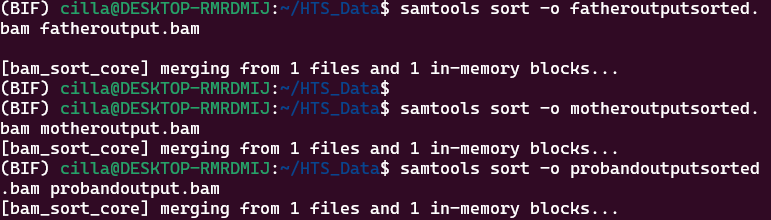
STEP 6: CONVERTING SAM FILES TO BAM FILES

Code Used

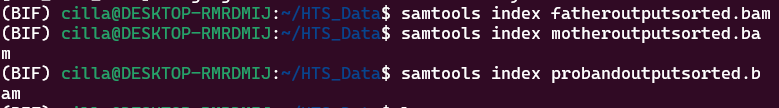


STEP 7: SORTING THE BAM FILES

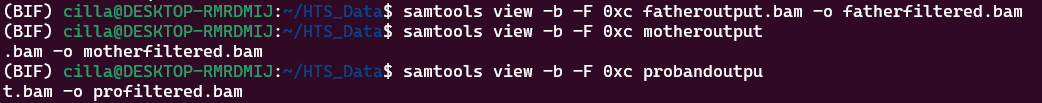
Code used:



STEP 8: INDEXING THE BAM FILES



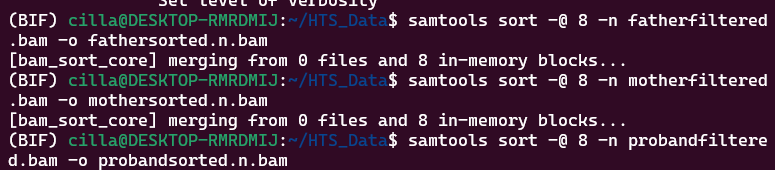
STEP 9: FILTERING THE BAM FILES



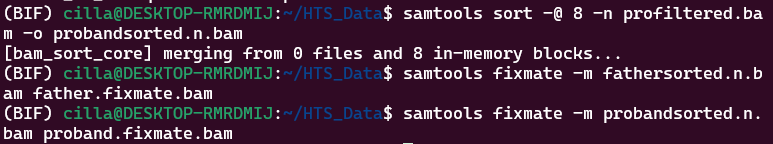
STEP 10: REMOVING DUPLICATE READS

First, we sorted the bam files based on names

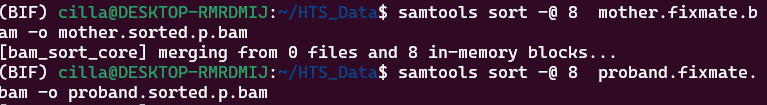
Code used:



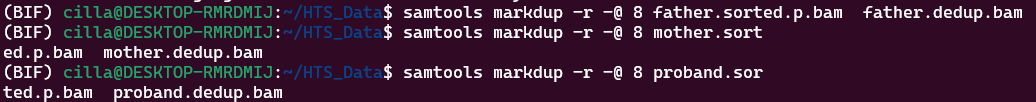
Then we fixmated the sorted files using these codes



Sorting the fixmated bam files according to coordinates



Removing duplicates



STEP 11

VARIANT CALLING

Before performing variant calling, we have to index our dedup bam file

Code Used

