Análise da Qualidade do Sequenciamento

Pablo Rodrigo Sanches

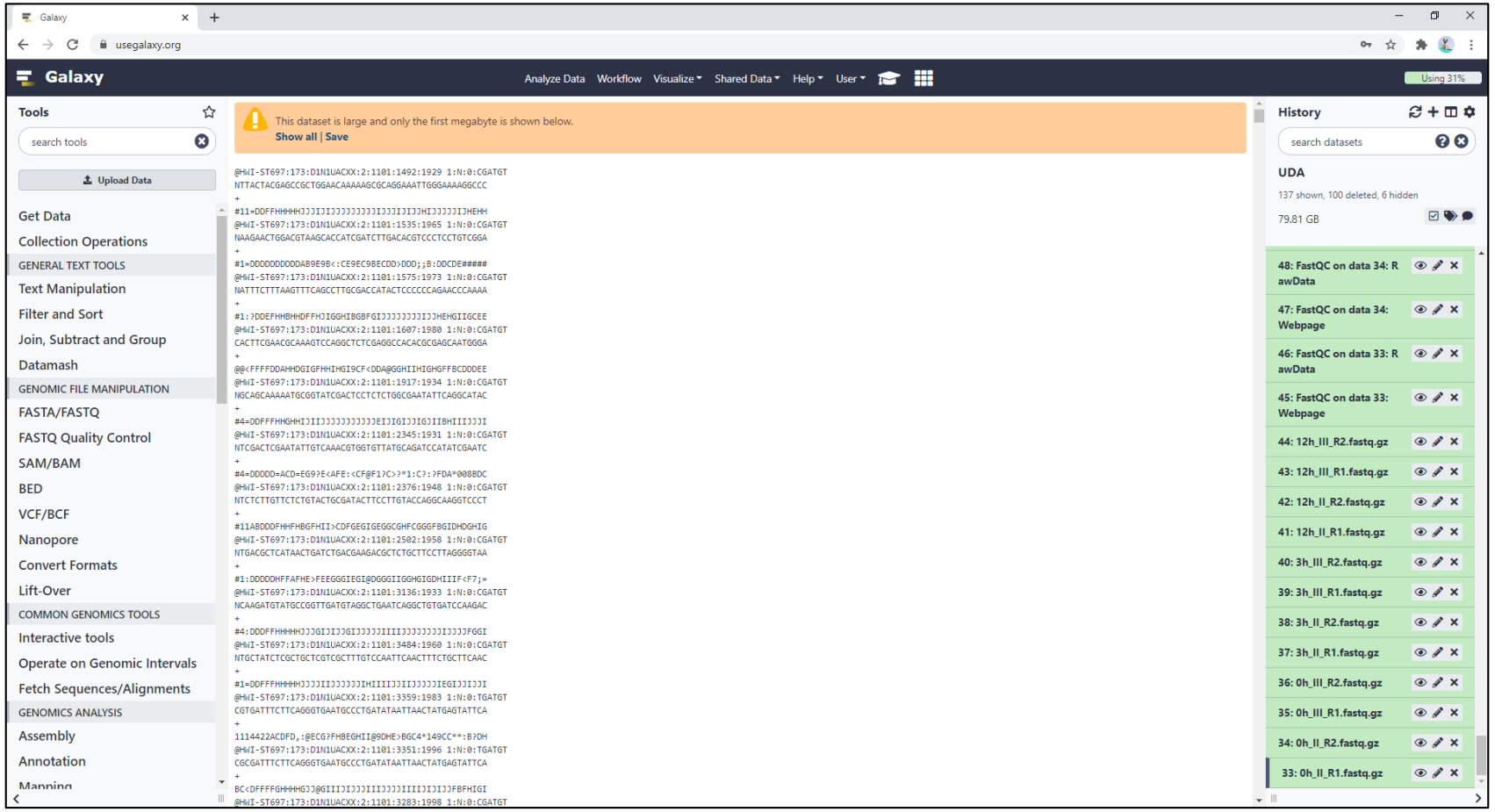
Departamento de Genética – FMRP/USP

psanches@usp.br

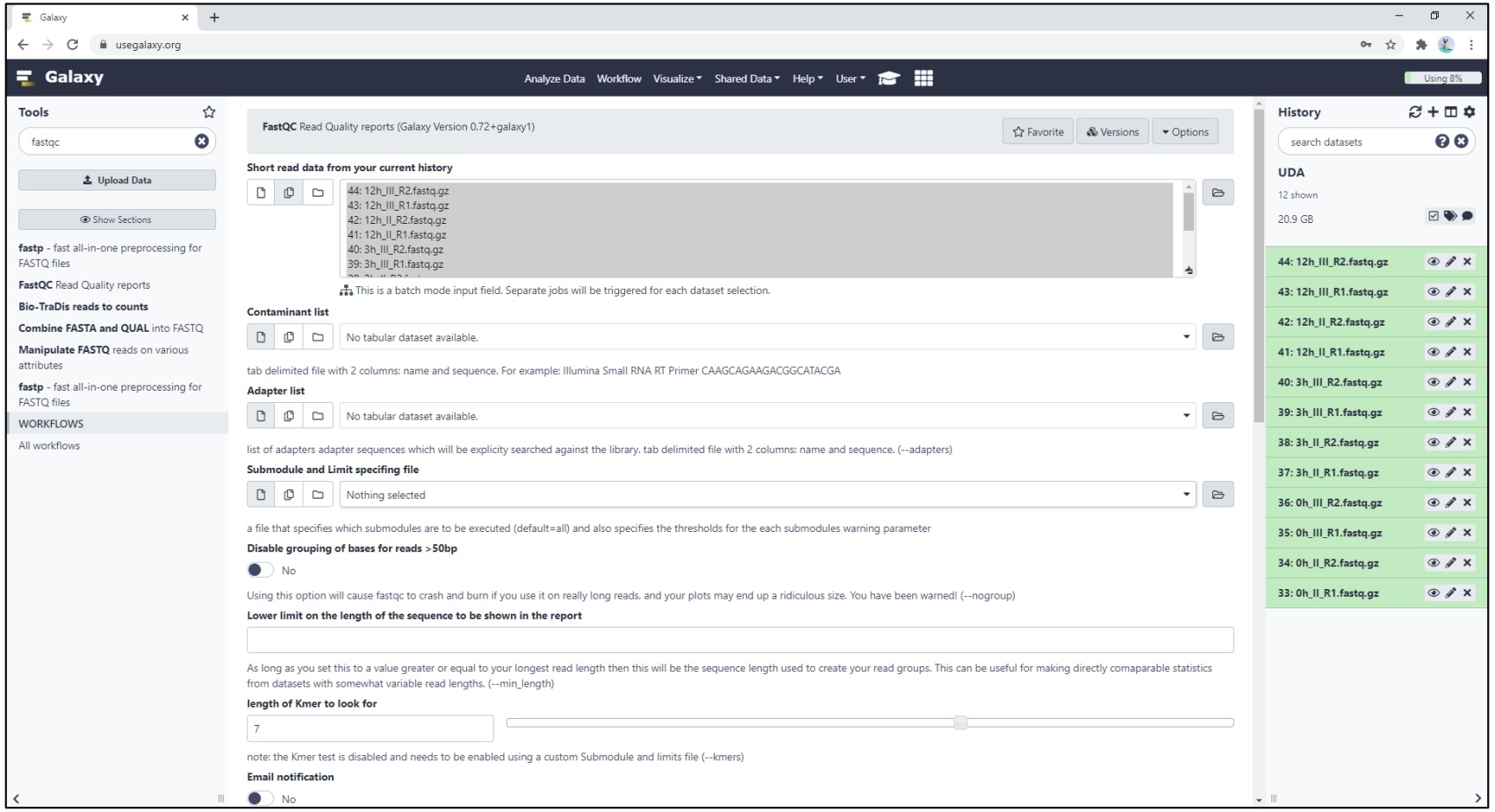
Roteiro de análise

1. FASTQC – 1ª rodada (análise de qualidade “pré-trimagem”)
2. Trimmomatic (“trimagem” de adaptadores e regiões de má qualidade)
3. FASTQC – 2ª rodada (análise de qualidade “pós-trimagem”)
4. Comparação dos resultados obtidos nas 1ª e 2ª rodadas
5. Desenvolvimento da tabela de resultados

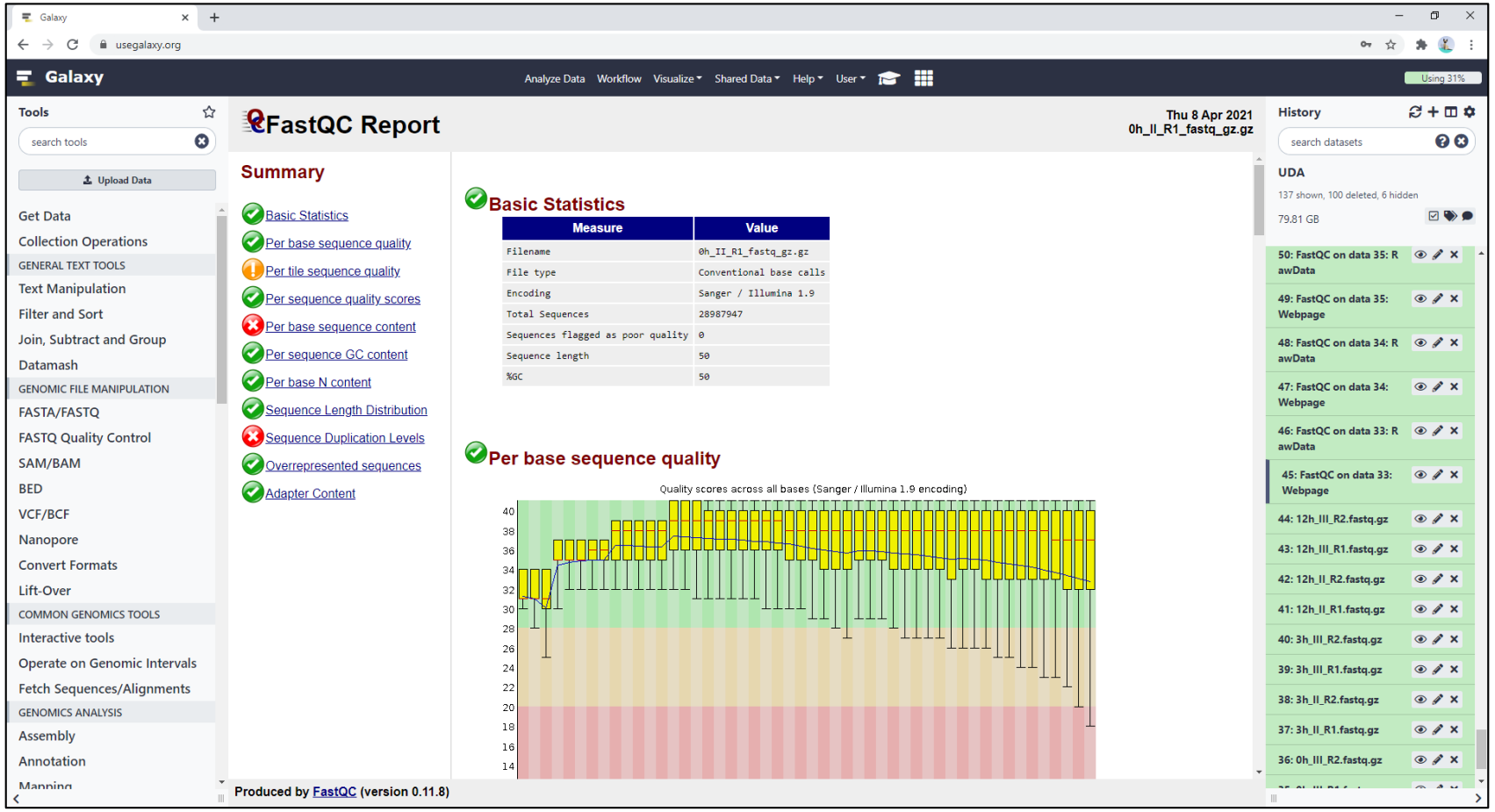
Arquivos Fastq (Dados Brutos)



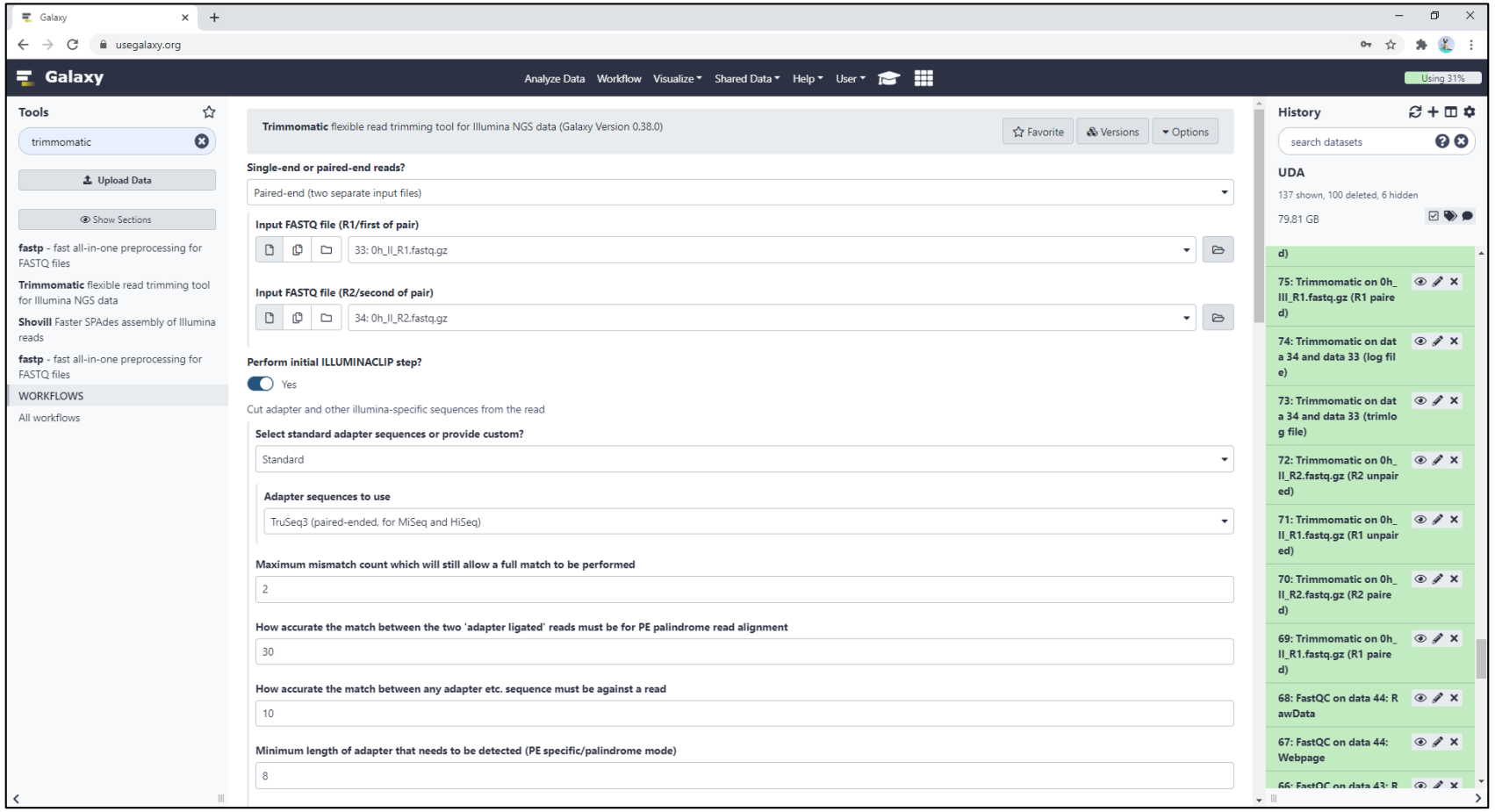
FASTQC - Via Galaxy – 1ª rodada



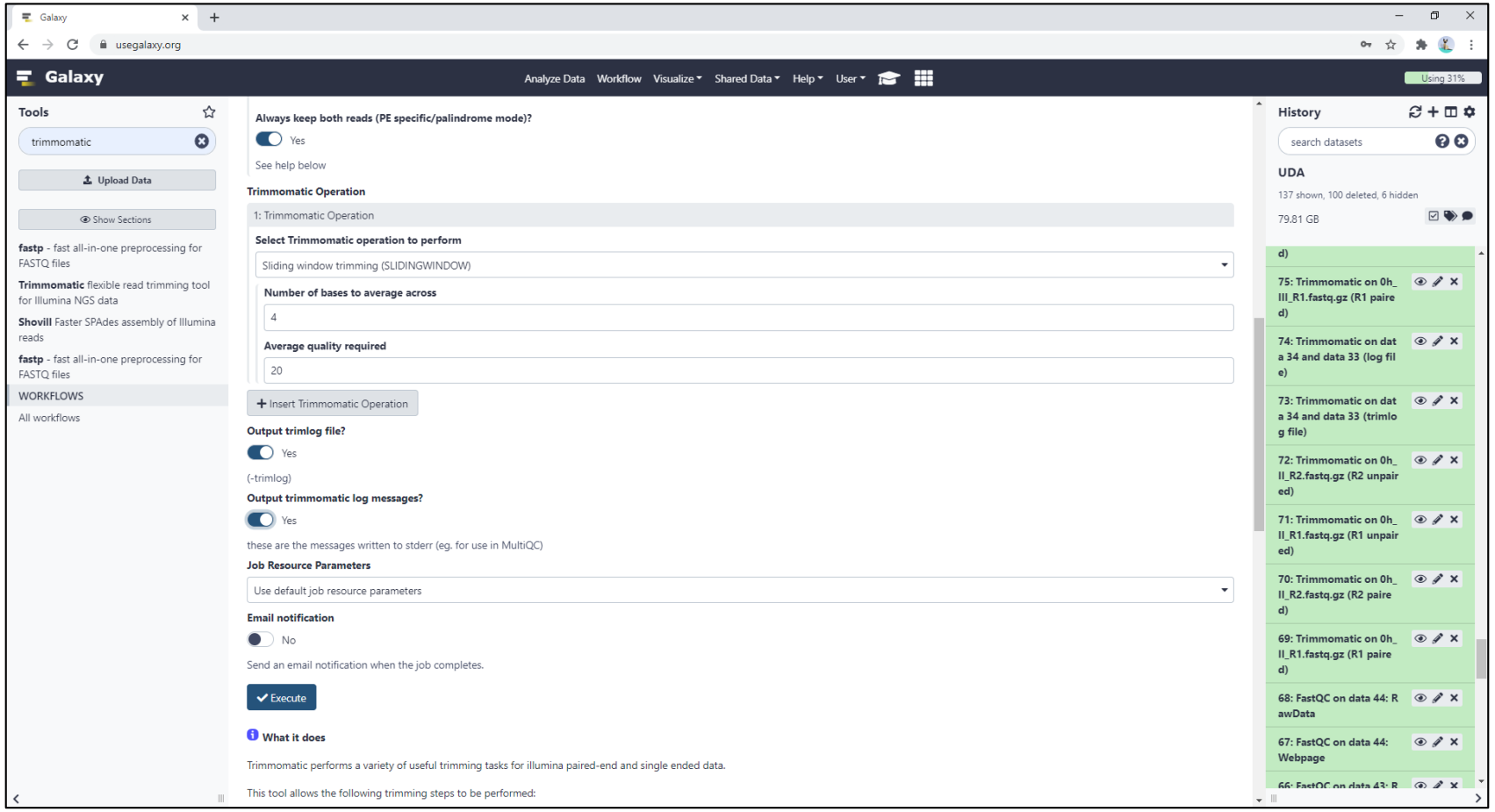
Exemplo de Resultado FASTQC



Trimmomatic – Via Galaxy



Trimmomatic – Via Galaxy (cont.)



Exemplo de Resultado Trimmomatic

```
Picked up _JAVA_OPTIONS: -Djava.io.tmpdir=/jetstream/scratch0/main/jobs/34439361/_job_tmp
TrimmomaticPE: Started with arguments:
 -threads 10 fastq_r1.fastqsanger.gz fastq_r2.fastqsanger.gz fastq_out_r1_paired.fastqsanger.gz fastq_out_r1_unpaired.fastqsanger.gz fastq_out_r2_paired.fastqsanger.gz
fastq_out_r2_unpaired.fastqsanger.gz ILLUMINACLIP:/cvmfs/main.galaxyproject.org/deps/_conda/envs/__trimmomatic@0.38/share/trimmomatic-0.38-1/adapters/TruSeq3-PE.fa:2:30:10:8:true
SLIDINGWINDOW:4:20 -trimlog trimlog
Using PrefixPair: 'TACACTCTTTCCCTACACGACGCTCTTCCGATCT' and 'GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT'
ILLUMINACLIP: Using 1 prefix pairs, 0 forward/reverse sequences, 0 forward only sequences, 0 reverse only sequences
Quality encoding detected as phred33
Input Read Pairs: 28987947 Both Surviving: 26800883 (92.46%) Forward Only Surviving: 1836438 (6.34%) Reverse Only Surviving: 274016 (0.95%) Dropped: 76610 (0.26%)
TrimmomaticPE: Completed successfully
```

Nro. Reads totais

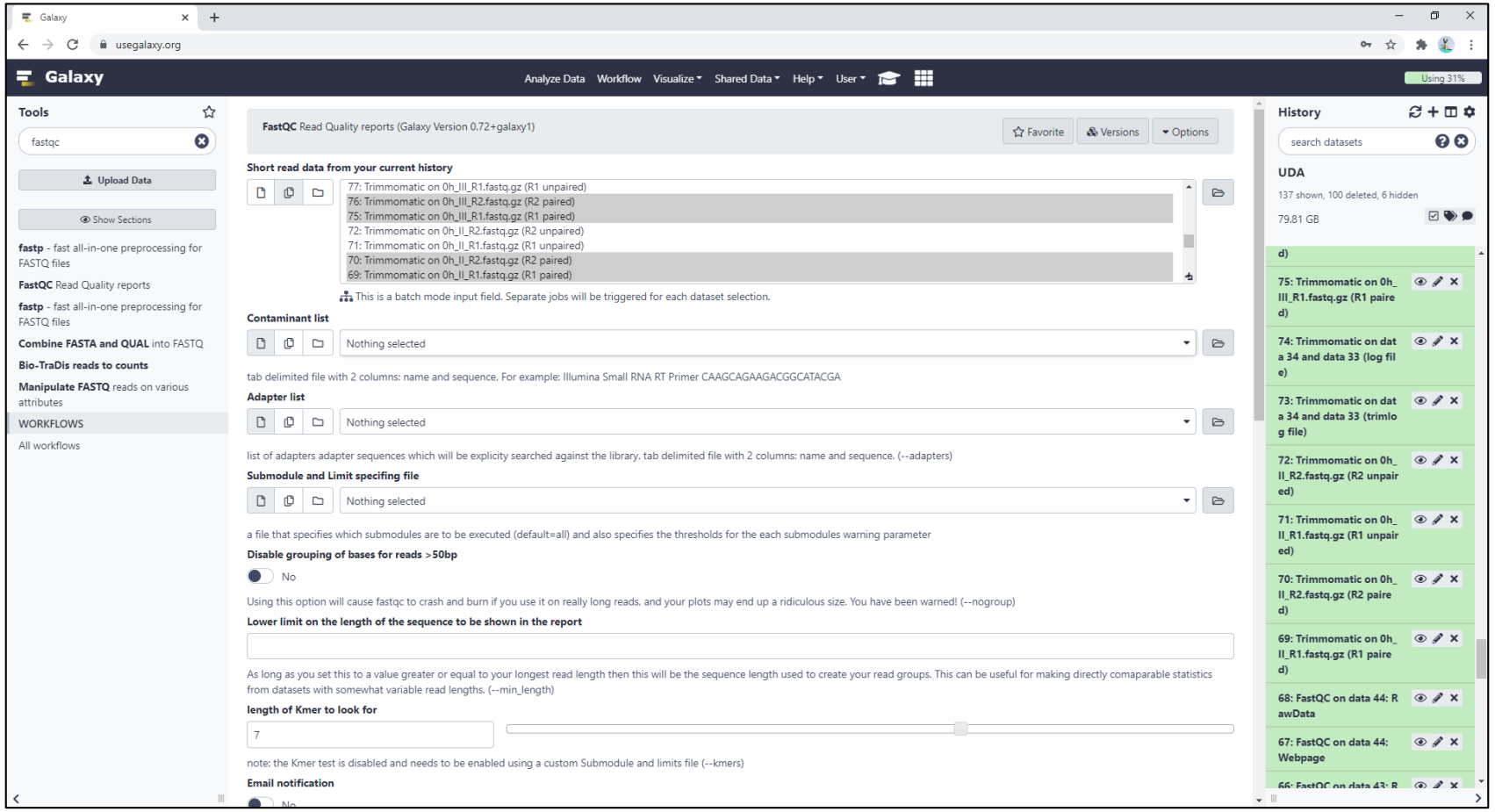
Nro. de pares de Reads após trimagem

Reads apenas em R1 após trimagem

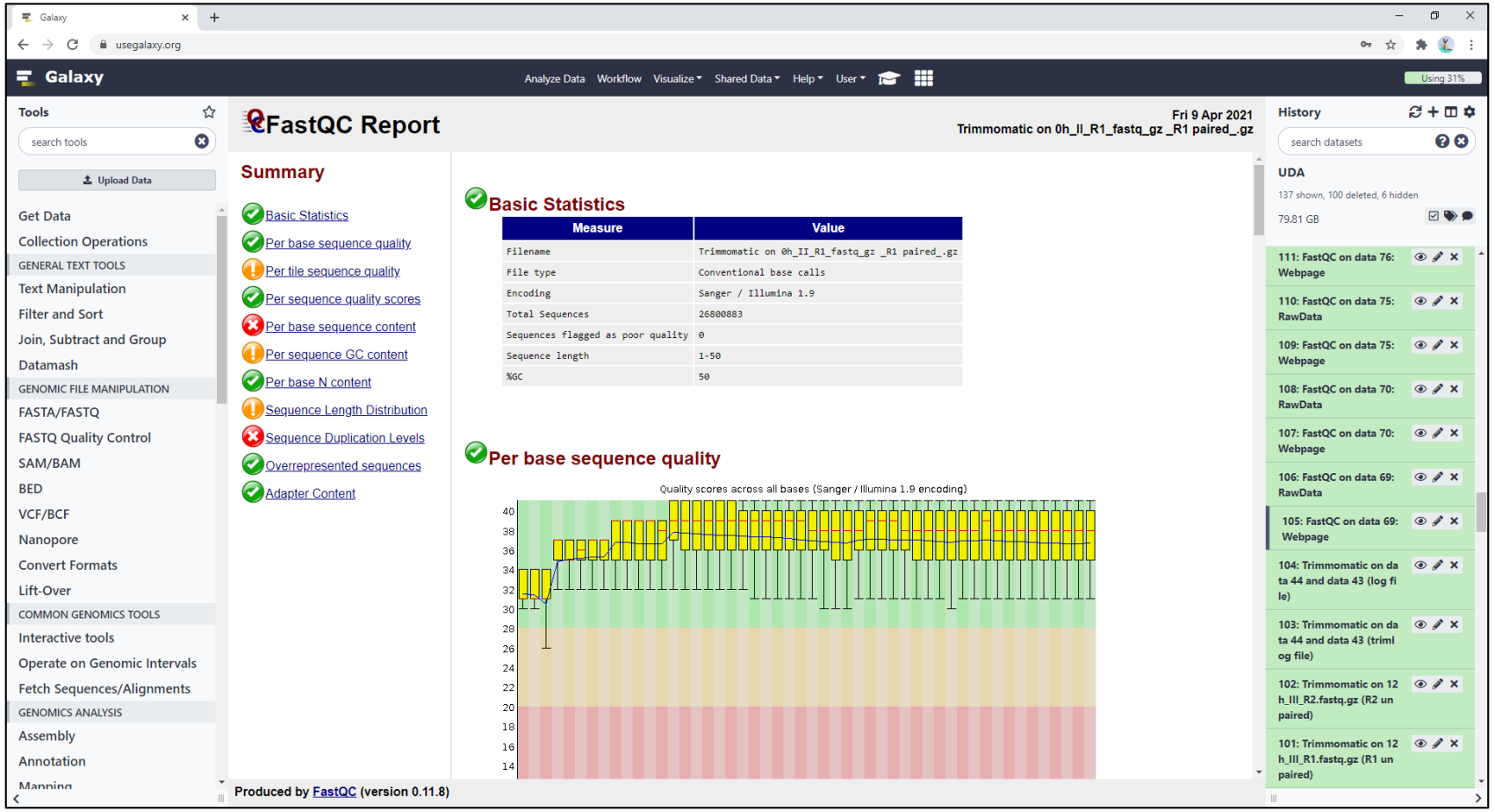
Reads apenas em R2 após trimagem

Reads retirados após trimagem

FASTQC - Via Galaxy – 2ª rodada

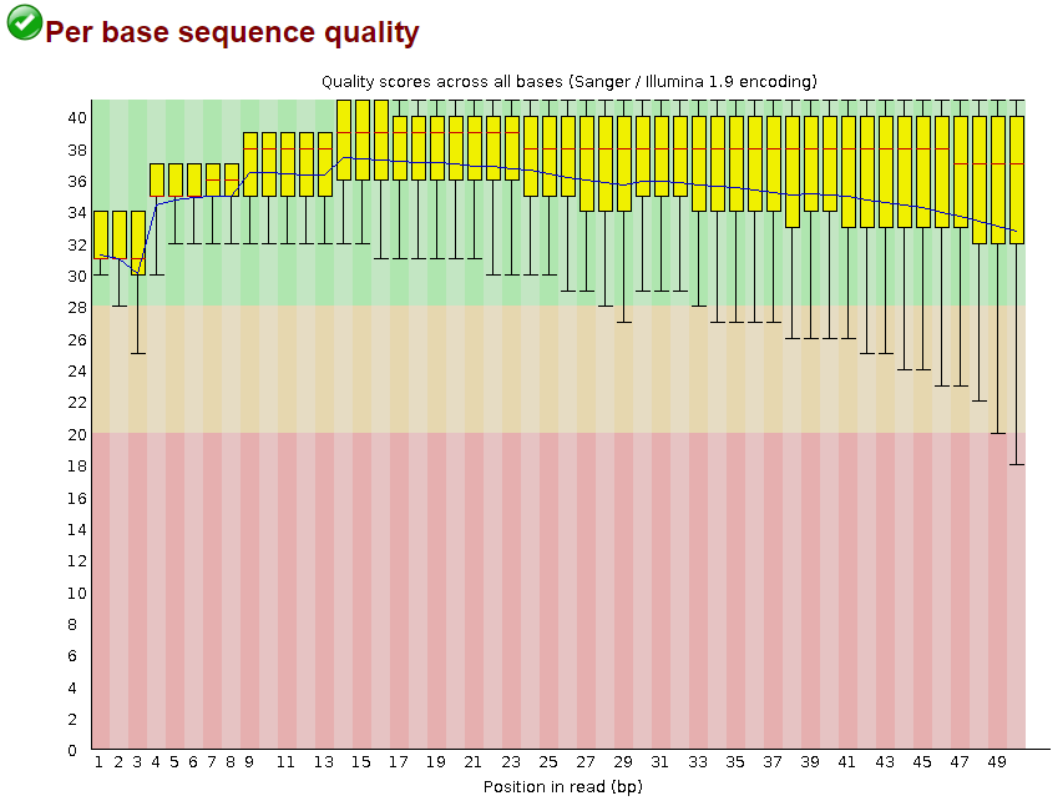


Exemplo de Resultado FASTQC

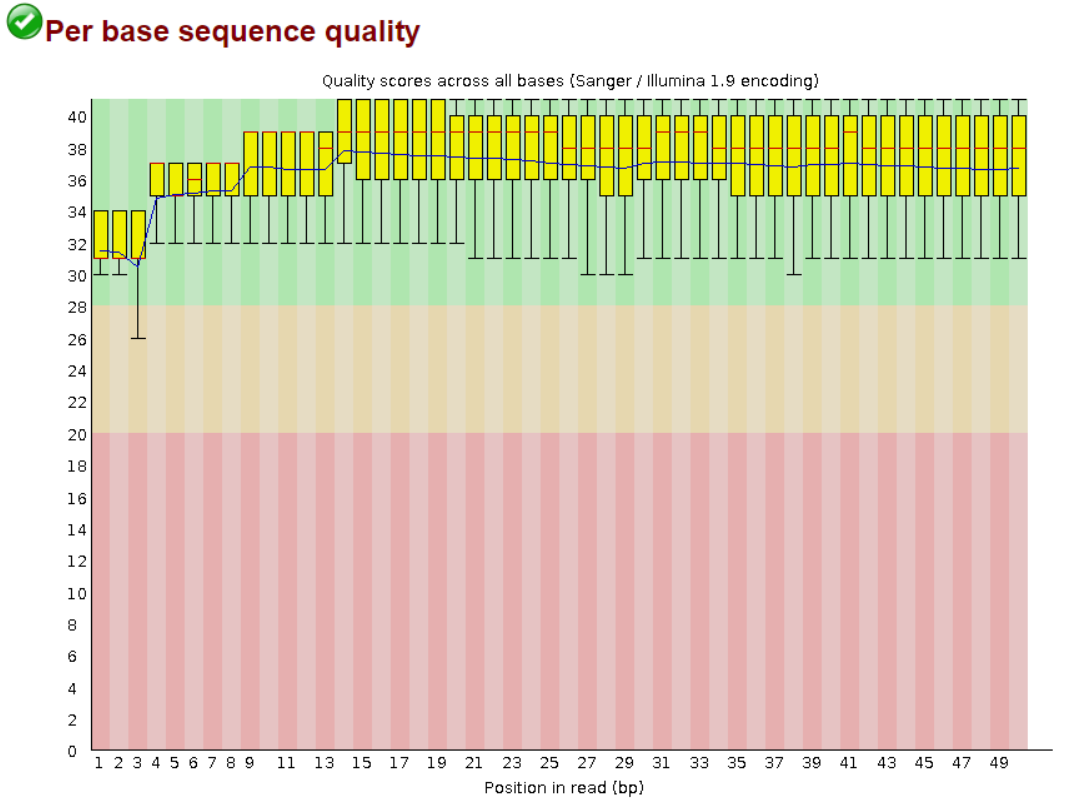


Comparação dos resultados obtidos

Antes (1ª rodada)



Depois (2ª rodada)



Desenvolver tabela de resultados

Table 1. General features of RNA-seq reads mapped to the T. rubrum reference genome

Sample	Raw reads	High-quality reads
0h II		
0h III		
3h II		
3h III		
12h II		
12h III		