

Análise da Qualidade do Sequenciamento

(continuação)

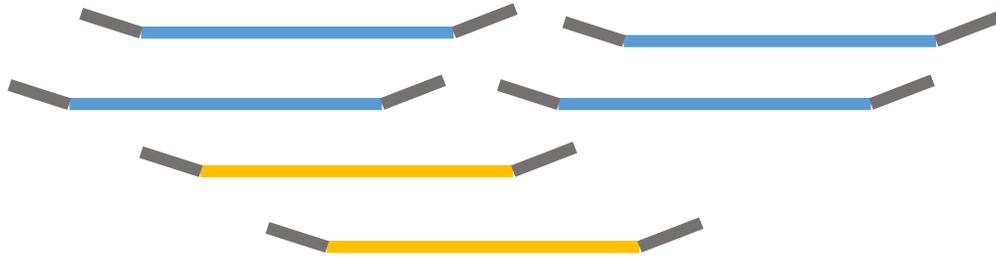
Pablo Rodrigo Sanches

Departamento de Genética – FMRP/USP

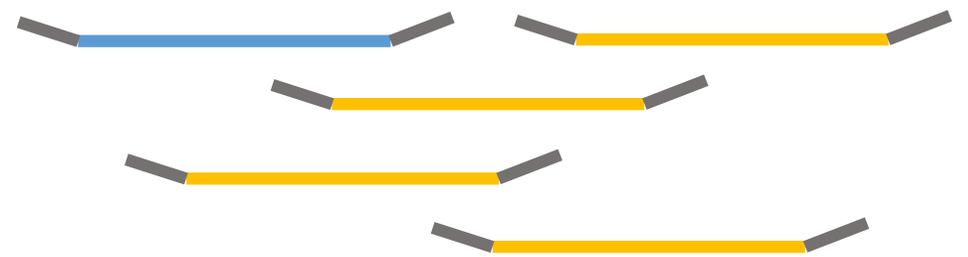
psanches@usp.br

Recordando...

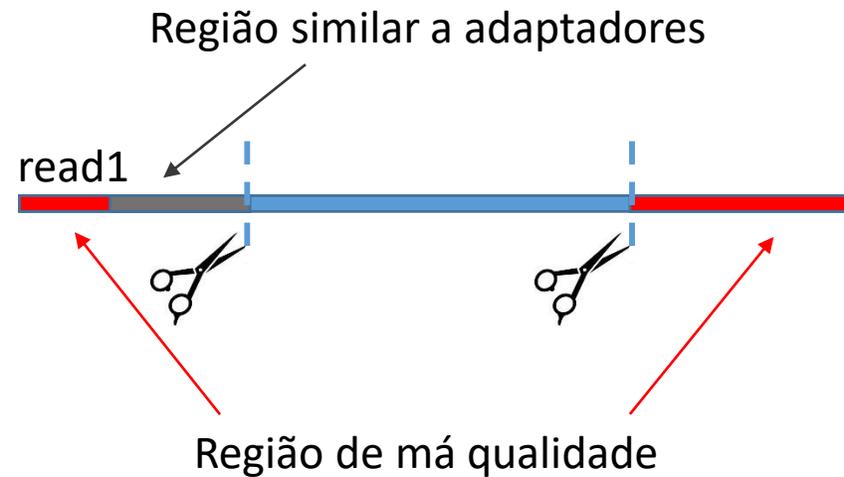
Controle (ex. não tratado)



Desafio (ex. tratado com droga)



Reads (leituras)
do RNA-Seq



Recordando...



Retirar regiões com score mínimo de qualidade

Read bruta

| Base | A | C | T | G | G | T | C | A | C | A | C | G | T | A | A | T | G | C | A | C |
|-----------|---|----|----|---|----|----|----|----|----|----|----|----|----|----|----|----|----|----|---|---|
| Qualidade | 5 | 10 | 12 | 8 | 20 | 22 | 25 | 15 | 30 | 30 | 30 | 35 | 35 | 32 | 20 | 18 | 10 | 10 | 5 | 2 |

Ex. corte de qualidade:
Tamanho da Janela=4
Qualidade média=20

Read “trimada” quanto a qualidade das bases

| Base | G | T | C | A | C | A | C | G | T | A | A |
|-----------|----|----|----|----|----|----|----|----|----|----|----|
| Qualidade | 20 | 22 | 25 | 15 | 30 | 30 | 30 | 35 | 35 | 32 | 20 |

Retirar regiões de adaptadores

Read “trimada” quanto a qualidade das bases

| Base | G | T | C | A | C | A | C | G | T | A | A |
|-----------|----|----|----|----|----|----|----|----|----|----|----|
| Qualidade | 20 | 22 | 25 | 15 | 30 | 30 | 30 | 35 | 35 | 32 | 20 |

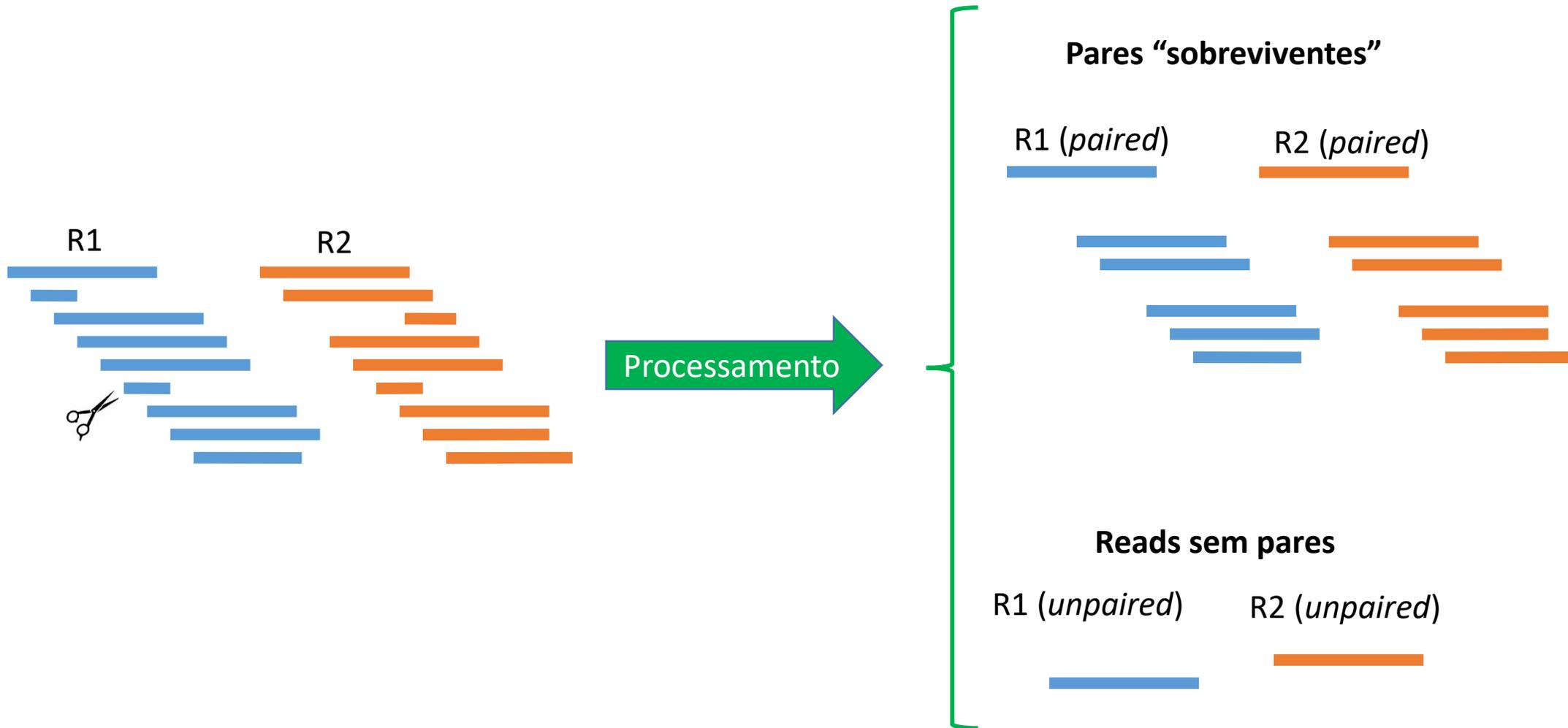
Sequências apresentam similaridade?

```
Read:      G T C A C A C G T A A
           |   | |
Adaptador: G A C A
```

Read “trimada” quanto a similaridade com adaptador

| Base | C | A | C | G | T | A | A |
|-----------|----|----|----|----|----|----|----|
| Qualidade | 30 | 30 | 30 | 35 | 35 | 32 | 20 |

Retirar reads de tamanho mínimo (*paired-end*)



Log de Resultado Trimmomatic

Input Read Pairs: 28987947

Both Surviving: 27213240 (93.88%)

Forward Only Surviving: 784150 (2.71%)

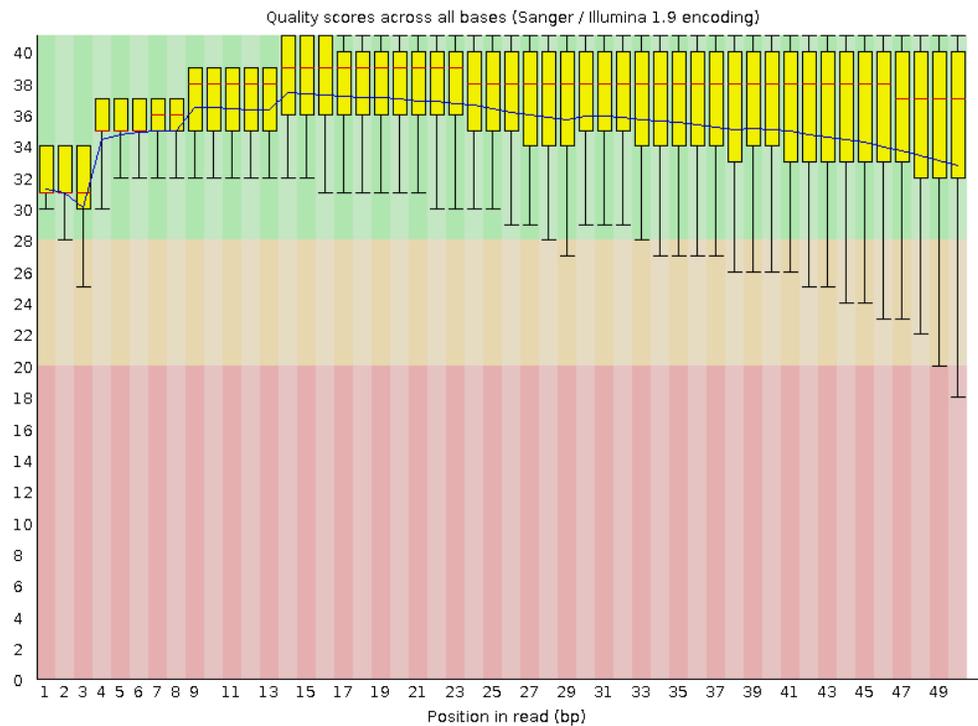
Reverse Only Surviving: 863068 (2.98%)

Dropped: 127489 (0.44%)

Comparação dos resultados obtidos

Pré-trimagem

✔ Per base sequence quality



Pós-trimagem

✔ Per base sequence quality

