

## CONTROLE DE QUALIDADE

Foi possível identificar no *Fastq* a qualidade dos sequenciamentos e seus pontos de melhoria. Portanto, o primeiro passo foi usar o *Trimmomatic* remover adaptadores e filtrar os reads por qualidade dentro de uma janela passada por parâmetros.

Parâmetros do Trimmomatic:

ILLUMINACLIP: Remoção adaptadores.

LEADING: Remoção bases do início com baixa qualidade ou Ns.

TRAILING: Remoção bases do fim com baixa qualidade ou Ns.

SLIDINGWINDOW: Percorre o read com uma janela de valor n, removendo quando a qualidade média por base é menor do que um valor q.

Parâmetros usados para todos:

LEADING:10 TRAILING:10 SLIDINGWINDOW:5:20 ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10.

### **BcaribensisP30**

Input Read Pairs: 931205 Both Surviving: 922122 (99,02%) Forward Only Surviving: 6667 (0,72%) Reverse Only Surviving: 450 (0,05%) Dropped: 1966 (0,21%).

### **CupriavidusP15**

Input Read Pairs: 1545430 Both Surviving: 1503233 (97,27%) Forward Only Surviving: 38385 (2,48%) Reverse Only Surviving: 818 (0,05%) Dropped: 2994 (0,19%).

### **HerbaspirillumP19**

Input Read Pairs: 407873 Both Surviving: 397794 (97,53%) Forward Only Surviving: 7073 (1,73%) Reverse Only Surviving: 2641 (0,65%) Dropped: 365 (0,09%).