

Instituto Federal de Educação, Ciência e Tecnologia de Brasília.
Nome: Danyelle da Silva Oliveira Angelo.

1. a

a. Dados de sequenciamento: 59

DNA: 5

RNA: 54

b. Illumina HiSeq 2000

c.

i. 3750513

ii. RNA-Seq: é uma tecnologia de perfil de transcriptoma, essa tecnologia não depende dos dados existentes do genoma e pode rastrear novas transcrições e analisar a estrutura das transcrições, incluindo resolução de pares de bases simples o que torna-a útil para análise de genotipagem e ligação

iii. O transcriptoma é o conjunto completo de transcritos em uma célula e sua quantidade, para um estágio de desenvolvimento específico ou condição fisiológica.

iv. PAIRED-end.

O PAIRED-end é um sequenciamento de extremidade, ele fornece um alinhamento de alta qualidade entre regiões do DNA que contêm sequências repetitivas, além disso ele detecta rearranjos comuns de DNA, como inserções, deleções e inversões. A sua qualidade é melhor em relação ao SINGLE-end, porém o custo também é mais alto.

O SINGLE-end: é um sequenciamento de DNA de apenas uma extremidade, fornece grandes volumes de dados de alta qualidade, de forma rápida e econômica. Sua qualidade e seu custo é inferior ao PAIRED-end.

v. Runs: 62, tamanho dos arquivos: 211.51 Gb

d.

Código de Acesso	Estratégia de Sequenciamento	Layout	Runs	Tamanho dos Arquivos(Mb)
4482795	WGS	PAIRED	2	32718.65
4482794	WGS	PAIRED	1	25968.79
459029	WGS	PAIRED	1	21268.43
459028	WGS	PAIRED	2	36667.14
459027	WGS	PAIRED	2	34789.52

2.

a. 797.757

b. 469 resultados. A consulta ficou na forma: "toxin organism:spider AND reviewed:no"

- c. https://github.com/DanyelleAngelo/Bioinform-tica/raw/master/Trabalho1/uniprot-toxin%2Borganism_spider%2BAND%2BReviewed_no.fasta
- d. https://github.com/DanyelleAngelo/Bioinform-tica/blob/master/Trabalho1/uniprot-toxin+organism_spider+AND+reviewed_no.txt?raw=true
- e. Programa para coletar informações:
https://raw.githubusercontent.com/DanyelleAngelo/Bioinform-tica/master/Trabalho1/2-ac_namefunction.py?token=AB6SZSK5JKCHZJ7I43YIROK6OF674
 Informações coletadas:
<https://github.com/DanyelleAngelo/Bioinform-tica/raw/master/Trabalho1/saida2.txt>
 Tabela com as informações:
https://drive.google.com/open?id=1Fs5yBGwYyhnoafm_0kaf757VtlaVOapT

3.

Dados manualmente curados:
https://raw.githubusercontent.com/DanyelleAngelo/Bioinform-tica/master/Trabalho1/uniprot-toxin%2Borganism_spider.fasta?token=AB6SZSJ75SDFJ6D2NIK4FPK6OF6ZC