

Tópico 02 - Navegação e Estrutura de Diretórios no Linux



1. Por que entender diretórios é essencial na bioinformática?

Trabalhar com Linux envolve lidar com **muitos arquivos ao mesmo tempo**: FASTQ, FASTA, GTF, VCF, tabelas, anotações, logs, etc.

Para análises como RNA-Seq, WGS, ChIP-Seq ou montagem genômica, você precisa:

- ✓ saber onde estão os dados
- ✓ organizar entradas e saídas
- ✓ apontar caminhos corretamente em scripts e pipelines

Um erro simples de caminho pode quebrar um pipeline inteiro.



2. Estrutura hierárquica do sistema Linux

O Linux organiza tudo como uma **árvore**:

```
/  
└── home/  
    └── usuario/  
        ├── documentos/  
        ├── downloads/  
        └── projetos/  
└── etc/  
└── usr/
```

```
└── bin/  
└── tmp/  
└── media/
```

Os diretórios mais usados na bioinformática:

Diretório	Função
/home/usuario	Onde ficam seus projetos e dados. ✓
/mnt / /media	Onde pendrives e HDs externos são montados.
/tmp	Arquivos temporários (cuidado ao salvar dados aqui).
/usr/bin , /usr/local/bin	Onde ferramentas ficam instaladas.

📌 **Regra geral:**

| Se é dado seu, coloque dentro de /home/seu_nome/.



3. Comandos essenciais de navegação

◆ Mostrar onde você está:

```
pwd
```

◆ Listar arquivos da pasta:

```
ls
```

◆ Entrar em uma pasta:

```
cd nome_da_pasta
```

◆ Voltar um nível:

```
cd ..
```

◆ Voltar para a pasta home:

```
cd ~
```

◆ Ir direto para um caminho específico:

```
cd /home/usuario/projetos/rnaseq
```

4. Caminho absoluto vs caminho relativo

Caminho absoluto

Sempre começa com `/`

É o endereço completo no sistema.

Exemplo:

```
cd /home/jean/projetos/rnaseq/raw_data
```

Caminho relativo

Começa da pasta onde você está.

Exemplo:

```
cd raw_data
```

 Em bioinformática, **use caminhos absolutos em scripts** para evitar erros, e **relativos ao trabalhar manualmente** no terminal.

5. Criando e organizando diretórios

Criar uma pasta:

```
mkdir minha_pasta
```

Criar várias:

```
mkdir raw results scripts qc logs
```

Criar com hierarquia:

```
mkdir -p projeto/data/fastq
```

6. Estrutura recomendada para um projeto de Bioinformática

Vamos criar a estrutura ideal de um pipeline RNA-Seq:

```
mkdir -p projeto_rnaseq/{raw_data,fastqc,multiqc,scripts,results,logs}
```

📌 Cada pasta terá um papel:

- **raw_data** → FASTQ originais
- **fastqc** → relatórios de qualidade
- **multiqc** → dashboard consolidado
- **results** → tabelas, figuras, saídas
- **scripts** → shell scripts e pipelines
- **logs** → arquivos de execução (debug)

Exercícios

Exercício 1 — Criar sua estrutura de projeto

Crie uma pasta chamada `meu_projeto_bioinfo` com subpastas:

- data
- results
- scripts
- qc

Exercício 2 — Navegar entre os diretórios criados

1 Entre em `data`

2 Volte um nível

3 Vá até scripts

4 Descubra onde está

Exercício 3 — Trabalhar com caminhos absolutos

1 Saia para a home:

2 Entre no projeto usando o caminho completo: