

Tutorial Bioinformatica

Marcel Ferreira - Bolsista/CAPES

2023-12-05

Table of contents

Sobre esse curso	4
Realização	4
Apoio	4
Autores	4
Colaboradores	4
Introdução	5
Primeiros passos	6
Configurações de sistema	6
Softwares necessários	6
Usuários Windows	6
No Ubuntu	6
Instalação	7
Dados utilizados	7
Dia 1 - Sequenciamento de DNA	9
Arquivos	9
Atividades	9
Dia 2 - Alinhamento de sequências de DNA	11
Arquivos	11
Atividades	11
Dia 3 - Genotipagem	13
Arquivos	13
Atividades	13
Dia 4 - Análise de sequenciamento Oxford Nanopore	15
Atividades	15
Dia 5 - Genotipagem de STRs a partir de dados de NGS	19
Referências	20

ANEXO: Dicas para uso do Ubuntu/WSL	21
Navegação e Diretórios	21
Listar Conteúdo do Diretório	21
Mudar de Diretório	21
Diretório Atual	21
Criar Diretório	21
Manipulação de Arquivos	21
Copiar Arquivo	21
Mover/Renomear Arquivo	21
Remover Arquivo	22
Visualização de Conteúdo	22
Visualizar Conteúdo do Arquivo	22
Visualizar Conteúdo do Arquivo (página por página)	22
Pacotes e Atualizações	22
Atualizar Lista de Pacotes	22
Atualizar Pacotes Instalados	22
Instalar Novo Pacote	22
Gerenciamento de Usuários	22
Adicionar Usuário ao Grupo	22
Mudar Senha do Usuário	23
Processos	23
Listar Processos	23
Matar um Processo por ID	23
Monitorar Recursos do Sistema (htop)	23
Medir Tempo de Execução de um Comando (time)	23
Rede	23
Verificar Configurações de Rede	23
Testar Conexão com um Endereço IP	23
Entrada e Saída Padrão (stdin/stdout)	24
Outros Comandos Úteis	24
Ajuda sobre um Comando	24
Sair do Terminal	24

Sobre esse curso

Realização



Figure 1: Realização UnB, UNESP e USP.

Apoio



Figure 2: CAPES-PROCAD Edital nº 16/2020. Processos 88887.516236/2020-00 e 88881.516238/2020-01.

Autores

Celso Teixeira Mendes Junior
Erick da Cruz Castelli
Marcel Rodrigues Ferreira
Tamara Soledad Frontanilla Recalde

Colaboradores

Introdução

Bem-vindos ao Workshop de Bioinformática Aplicada à Genética Forense: Análise de Dados de Sequenciamento de Segunda e Terceira Geração. Este curso abrangente foi projetado para fornecer a vocês uma imersão prática nas técnicas de análise de dados genômicos, com foco especial na aplicação forense.

A genética forense tornou-se uma ferramenta essencial na resolução de casos criminais, identificação de indivíduos e estabelecimento de relações familiares. Neste workshop de cinco dias, exploraremos os fundamentos e as aplicações práticas do sequenciamento de DNA, abordando desde os conceitos básicos até as técnicas avançadas de genotipagem de STRs (*Short Tandem Repeats*) a partir de dados de *Next-Generation Sequencing* (NGS).

Dia 1 - Sequenciamento de DNA: Iniciaremos nossa jornada explorando os princípios fundamentais do sequenciamento de DNA de segunda geração. Compreenderemos as tecnologias por trás desses métodos e sua importância na geração de dados genômicos de alta qualidade. Também serão analisados dados brutos de sequenciamento e seu controle qualidade.

Dia 2 - Alinhamento de Sequências de DNA: No segundo dia, mergulharemos na etapa crucial de alinhamento de sequências de DNA. A precisão dessa fase é vital para extrair informações significativas dos dados brutos e identificar variações genéticas.

Dia 3 - Identificação de Variantes: Aprofundando-nos ainda mais, dedicaremos o terceiro dia à identificação de variantes genéticas. Exploraremos ferramentas e estratégias para detectar mutações, SNPs (*Single Nucleotide Polymorphisms*), InDels, e outras variantes que desempenham um papel crucial na individualidade genômica.

Dia 4 - Análise de Sequenciamento Oxford Nanopore: No quarto dia, abordaremos uma tecnologia revolucionária: o sequenciamento Oxford Nanopore. Compreenderemos suas vantagens, desafios e exploraremos casos de uso específicos na genética forense.

Dia 5 - Genotipagem de STRs a partir de dados de NGS: Encerraremos o workshop com uma exploração prática da genotipagem de STRs, uma ferramenta valiosa para estabelecer perfis genéticos únicos. Aprenderemos a interpretar e analisar esses dados, fornecendo *insights* fundamentais para investigações forenses.

Ao longo desta semana, vocês serão desafiados a aplicar os conhecimentos adquiridos em exercícios práticos e estudos de caso, preparando-os para enfrentar os desafios reais da genética forense na era da bioinformática. Esteja preparado para uma jornada intensiva de aprendizado e descoberta!

Primeiros passos

Configurações de sistema

Antes de iniciarmos o tutorial, é imperativo garantir que o sistema atenda às configurações mínimas para uma experiência estável. Utilizaremos sistema Linux. Recomenda-se que a máquina disponha de, no mínimo, 40 GB de armazenamento, 8 GB de memória RAM e um processador i5/i7 ou compatível. No entanto, para uma performance ideal e considerando o potencial de expansão das aplicações, encorajamos a utilização de um sistema com mais de 60 GB de armazenamento e, no mínimo, 16 GB de memória RAM. Essas configurações mais robustas assegurarão não apenas a instalação suave do software, mas também a capacidade de executar múltiplas aplicações de forma eficiente, proporcionando uma experiência mais fluida e responsiva ao usuário.

Softwares necessários

Usuários Windows

- [WSL](#) (Windows Subsystem for Linux)
- [IGV](#) (Robinson et al. 2011)
- [FASTQC](#)
- [notepad++](#)

 Tutorial para instalar o WSL

Siga o tutorial da microsoft para instalar o WSL.

<https://learn.microsoft.com/pt-br/windows/wsl/install>

No Ubuntu

- [IGV](#) (Robinson et al. 2011)
- [FASTQC](#)
- [Trimmomatic](#) (Bolger, Lohse, and Usadel 2014)
- [bwa](#) (Li 2013)

- [minimap2](#) (Li 2018, 2021)
- [samtools](#) (Danecek et al. 2021)
- [freebayes](#) (Garrison and Marth 2012)
- [vcftools](#) (Danecek et al. 2011)
- [bcftools](#) (Danecek et al. 2021)
- [NanoPlot](#) (De Coster and Rademakers 2023)
- [chopper](#) (De Coster and Rademakers 2023)
- [cramino](#) (De Coster and Rademakers 2023)
- [gzip](#)

Instalação

Usuários windows

Usuários windows precisam instalar o Subsistema Windows para Linux (WSL). Os softwares FASTQC e IGV precisam ser instalados no windows e **não no WSL**. Anote a **senha** que você configurou. Ela será fundamental durante o uso do WSL!!!!

Ao terminar a instalação do WSL e de configurar seu usuário no linux utilize os seguintes comandos:

```
sudo apt-get update
```

```
sudo apt-get upgrade
```

Estes comandos irão garantir que o seu sistema esteja atualizado.

Sobre o comando sudo

O comando **sudo** permite ao usuário executar comandos com permissão superior. Para isso você precisará da sua **senha** (ou do administrador)!

Para instalar softwares no linux (diretamente ou no WSL) utilize o comando **apt install** da seguinte forma:

```
sudo apt install [SOFTWARE]
```

Dados utilizados

Baixe os dados que serão utilizados neste workshop via [Google Drive](#);

! Utilize o email correto

Para ter acesso aos dados utilize o email que foi fornecido durante a inscrição no evento.

Em caso de erro, entre em contato com a organização.

Os dados totalizam ~20 GB. **Atente-se para isso!!!**

Os dados estão contidos nesta estrutura de pastas descritas a baixo:

WorkshopDados/

```
--genome/
|  |--chr17.fas1
|  °--hg38.fa
--fast5/
|  °--*arquivos.fast52
--guppy_installer/
|  °--ont-guppy-cpu_6.5.7_linux64.tar.gz3
--LongReadsFastq/
|  |--HG00096.hg38.fastq
|  °--HG00099.hg38.fastq
°--ShortReadsFastq/
|  |--HG00096_r1.fastq
|  |--HG00096_r2.fastq
|  |--HG00097_r1.fastq
|  |--HG00097_r2.fastq
|  |--HG00099_r1.fastq
|  |--HG00099_r2.fastq
|  |--HG00100_r1.fastq
|  |--HG00100_r2.fastq
|  |--NA18486_r1.fastq
|  |--NA18486_r2.fastq
|  |--NA18487_r1.fastq
|  |--NA18487_r2.fastq
|  |--NA18488_r1.fastq
|  |--NA18488_r2.fastq
|  |--NA19648_r1.fastq
|  |--NA19648_r2.fastq
|  |--NA19649_r1.fastq
|  |--NA19649_r2.fastq
|  |--NA19651_r1.fastq
|  °--NA19651_r2.fastq
```

¹Genoma do cromossomo 17 obtido em https://timkahlke.github.io/LongRead_tutorials/

²Arquivos **fast5** obtidos em https://timkahlke.github.io/LongRead_tutorials/

³Instalador obtido em <https://community.nanoporetech.com/>

Dia 1 - Sequenciamento de DNA

! Importante

Verifique se o **FASTQC** está instalado.

Arquivos

Serão utilizados os arquivos contidos na pasta **WorkshopDados/shortReads/**.

Estão dados vieram do projeto 1000 genomas (<https://www.internationalgenome.org/home>) e compreendem 10 amostras: dois trios e quatro amostras isoladas.

Atividades

O controle de qualidade (QC) dos dados é uma etapa crítica na análise de sequenciamento de nova geração (NGS) para garantir a confiabilidade dos resultados. Abaixo estão as etapas típicas do controle de qualidade:

1. Análise Inicial com FASTQC:

- Execute o **FASTQC** nas suas leituras brutas para avaliar a qualidade geral. Isso inclui gráficos e estatísticas que indicam a distribuição da qualidade das bases ao longo das reads, a presença de adaptadores, a presença de sequências *overrepresented*, entre outros.

2. Identificação de Adaptadores e Trimagem:

- Com base nos resultados do **FASTQC**, identifique a presença de adaptadores e sequências indesejadas nas extremidades das reads. Utilize o **Trimmomatic** para remover essas sequências, garantindo que apenas dados de alta qualidade sejam mantidos.
- [COMANDO]

3. Remoção de Leituras de Baixa Qualidade:

- Algumas leituras podem conter regiões de baixa qualidade. Considere a remoção dessas leituras ou a trimagem de regiões específicas usando ferramentas adequadas, dependendo da natureza do problema.
- [COMANDO]

4. Filtragem de Leituras Curtas ou Longas:

- Dependendo do seu experimento, você pode querer filtrar leituras muito curtas ou muito longas que possam representar artefatos ou problemas experimentais.

5. Avaliação de Qualidade Pós-Trimagem:

- Após a trimagem e filtragem, execute novamente o **FASTQC** para avaliar como essas etapas afetaram a qualidade dos dados. Isso ajudará a garantir que você atingiu os padrões de qualidade desejados.

Dia 2 - Alinhamento de sequências de DNA

Arquivos

Os arquivos utilizados para estas análises serão os `.fastq` analisados no primeiro dia.

Arquivos `.fastq`

Preste atenção para o caminho da pasta aonde estão os `.fastq`. Eles estão na pasta `WorkshopDados/shortReads/`

Atividades

1. Indexação do Genoma de Referência:

- Para o alinhamento de sequências utilizaremos o programa `bwa`. Há outros, como `bowtie2`, `minimap2`, que podem ser utilizados caso seja conveniente. Os comandos abaixo preparam o genoma de referência para o `bwa`.
- Antes de realizar o alinhamento, é necessário indexar o genoma de referência usando o comando `bwa index`. Isso cria arquivos que aceleram o processo de alinhamento.

```
bwa index hg38.fa
```

Indexação do genoma

O processo para criar o índice do genoma via `bwa index` demora bastante tempo para ser realizado. Mesmo em máquinas com grandes capacidades de memória. Devido a isso colocamos na pasta `WorkshopDados/genome/` os arquivos resultantes desta etapa, que são os arquivos com extensão `.amb`, `.ann`, `.bwt`, `.fai`, `.pac` e `.sa`. O genoma de referência humano esta no arquivo `hg38.fa`. Esse é a última versão do genoma de referência e a mais utilizada no mundo. Caso deseje testar em um genoma menor, utilize o `chr12.fasta`.

2. Alinhamento de Sequências:

- Use o **bwa mem** para alinhar suas sequências de DNA ao genoma de referência.
- `bwa mem -R "@RG\tID:{SAMPLE}\tSM:{SAMPLE}" hg38.fa {SAMPLE}_r1.fastq {SAMPLE}_r2.fastq -o {SAMPLE}.sam`

Substitua **reference_genome.fa**, **{SAMPLE}** pelos nomes dos arquivos correspondentes.

💡 Automatizando o processo (Opcional)

Caso tenha experiência em programação, você pode utilizar um **loop**, como **for** ou **while**, para rodar todas as amostras ao mesmo tempo. Você pode utilizar o próprio **bash** (ubuntu) ou sua linguagem de programação favorita, como **python**, **perl**, **R**, etc.

Você poderá incluir as etapas seguintes no mesmo **loop**.

3. Converter Formato SAM para BAM e ordenar os reads:

- O arquivo de saída do **bwa mem** é no formato SAM. Converta-o para o formato BAM, mais compacto e eficiente.
- `samtools sort {SAMPLE}.sam > {SAMPLE}.bam`

4. Indexar o Arquivo BAM:

- Ordene o arquivo BAM para facilitar a busca e indexe-o para melhorar o desempenho de ferramentas subsequentes.
- `samtools index {SAMPLE}.bam`

5. Visualização do Alinhamento:

- Use o **IGV** (Integrative Genomics Viewer) para visualizar o alinhamento (**{SAMPLE}.bam**) e verificar sua qualidade.

! Regiões com reads

Os arquivos **fastq** fornecidos para este curso foram preparados contendo apenas as regiões que estão nos arquivos **bed** fornecidos.

6. Use o IGV para observar várias amostras ao mesmo tempo

Dia 3 - Genotipagem

Arquivos

Atividades

1. Preparação do Ambiente:

- Certifique-se de que o **freebayes** está instalado no seu ambiente. Você pode instalar com `sudo apt install freebayes`.

2. Indexação do Genoma de Referência (se ainda não estiver indexado):

- Assim como na etapa de alinhamento, o genoma de referência deve ser indexado.
- `samtools faidx hg38.fa`
- Caso utilize um genoma menor, atualize o nome.

3. Crie um arquivo com os nomes das amostras BAM:

- Entre na pasta onde estão os arquivos BAM gerados e utilize o comando `ls` como abaixo
- `ls *.bam > bam_list.txt`

4. Chamada de Variantes com freebayes:

- Execute o **freebayes** para chamar variantes a partir do arquivo BAM gerado após o alinhamento.
- `freebayes -f reference_genome.fa -L bam_list.txt -t bed_file.bed > variantes.vcf`

Otimizando o uso do freebayes

Caso o computador que esteja utilizando aborte o processo por falta de memória, você pode optar por reduzir o número de amostras em `bam_list.txt`. É importante destacar que a análise correta de genotipagem via **freebayes**

requer que todas as amostras sejam analisadas simultaneamente, mas para fim de aprendizado esta é uma estratégia.

5. Filtragem de Variantes (opcional):

- Dependendo dos seus critérios e do tipo de análise, pode ser necessário filtrar as variantes chamadas pelo **freebayes** para reduzir o número de falsos positivos. Abaixo estão alguns exemplos
- `bcftools view --exclude 'QUAL>1' variantes.vcf > variantes_filtradas.vcf`
- `bcftools view --trim-alt-alleles variantes_filtradas.vcf > variantes_filtradas_trim.vcf`
- `bcftools view --min-ac 1 variantes_filtradas_trim.vcf > variantes_filtradas_trim_minac.vcf`
- `bcftools norm -f hg38.fa variantes_filtradas_trim_minac.vcf > variantes_filtradas_trim_minac_norm.vcf`
- Adapte os critérios de filtragem conforme necessário.

6. Análise e Interpretação de Variantes:

- Utilize ferramentas como **VCFtools** para realizar análises adicionais no arquivo VCF, como filtragem específica e anotações.

7. Visualização de Variantes:

- Use o **IGV** para visualizar as variantes em relação ao genoma de referência e avaliar sua qualidade.
- Você pode abrir ao mesmo tempo o VCF e o alinhamento (BAM) de uma mesma amostra. Caso deseje.

Dia 4 - Análise de sequenciamento Oxford Nanopore

! Importante

O software **guppy** só esta disponível para download via comunidade da Oxford Nanopore. Para este tutorial fornecemos um arquivo **.tar** para instalação em sua máquina. O arquivo esta na pasta **WorkshopDados/guppy_installer/**

Para instalar siga os seguintes passos:

Acesse a pasta que contem o arquivo **.tar** e descompacte;

```
tar -xf ont-guppy-cpu_6.5.7_linux64.tar.gz
```

Verifique o caminho completo para a pasta

```
pwd
```

Executando **guppy** via caminho completo (Exemplo pedindo ajuda)

```
./guppy_basecaller --help
```

Este tutorial foi inspirado no tutorial do Tim Kahlke⁴ e em nossas experiências durante os trabalhos com ONT.

Atividades

1. Realize uma chamada de base utilizando **guppy** (Opicional):

- Verifique os workflows disponíveis para esta versão de **guppy**;
- `./Downloads/ont-guppy-cpu/bin/guppy_basecaller --print_workflows`
- Sabendo que este sequenciamento foi realizado utilizando o kit SQK-LSK108 e a flowcell MIN106, qual a configuração a ser utilizada?
- Realize a chamada de base para todos os arquivos **fast5** contidos na pasta **WorkshopDados/fast5/**;

⁴ https://timkahlke.github.io/LongRead_tutorials/

- `./Downloads/ont-guppy-cpu/bin/guppy_basecaller -i [PASTAFAST5] -s /
./guppy_out -c [CONFIG].cfg --num_callers 2 --cpu_threads_per_caller
1`

Aviso

O tempo de execução da chama de base (para este conjunto de dados) é superior a **12 horas** em máquinas de uso pessoal, e pode acabar inutilizando seu uso. Caso deseje praticar, recomendamos que seja feito em um momento onde não precise da máquina para outras atividades. Para este tutorial utilize o resultado da chamada de base contido na pasta **WorkshopDados/LongReadsFastq/**

2. Avaliação da qualidade do arquivo resultado .fastq:

- Utilize o NanoPlot para gerar gráficos de qualidade da amostra;
- `NanoPlot --fastq [AMOSTRA].fastq -o [OUTDIR] --N50 --verbose`

3. Filtre as leituras baseado no tamanho e qualidade:

- Utilize `chopper` para isso.

Instalação do `chopper`

Para instalar o `chopper` pode ser necessário ter o `conda` em seu sistema. Para isso utilize os seguintes comandos:

```
mkdir -p ~/miniconda3
wget https://repo.anaconda.com/miniconda/Miniconda3-latest-Linux-x86_64.sh
-O ~/miniconda3/miniconda.sh
bash ~/miniconda3/miniconda.sh -b -u -p ~/miniconda3
rm -rf ~/miniconda3/miniconda.sh
~/miniconda3/bin/conda init bash
~/miniconda3/bin/conda init zsh
Será necessário reiniciar seu sistema (Ubuntu ou WSL) após isso.
conda install -c bioconda chopper
```

- `chopper < [AMOSTRA].fastq -q [QUALIDADE] -l [TAMANHO_MIN] > [AMOSTRA]_filtrada.fastq`

Por que usar `<` no comando `chopper`?

O comando `<` garante que o arquivo `fastq` da amostra seja direcionado para entrada (**stdin**) do comando no `chopper`. Em nossas experiências já tivemos a necessidade de utilizar algumas vezes e outras não.

4. Alinhamento das Sequências:

- Use o `minimap2` para alinhar suas sequências de DNA ao genoma de referência.
- `minimap2 -ax map-ont [REFERENCE_GENOME].fa [AMOSTRA].fastq /
-R "@RG\tID:{SAMPLE}\tSM:{SAMPLE}" -t [THREADS] > [AMOSTRA].sam`

Consumo de memória durante a etapa

O alinhamento via `minimap2` tem pico de consumo de memória de ~ 13 GB. Caso seu computador disponha desta capacidade utilize o arquivo resultante na pasta **WorkshopDados/pre_run/Dia4/**

- Realize novamente utilizando `bwa mem` desta vez (Opcional);

Tempo de execução do `bwa`

Atenção, o `bwa` demora quase 10x mais que o `minimap2`

- `bwa mem -x ont2d [REFERENCE_GENOME].fa [AMOSTRA].fastq /
-R "@RG\tID:{SAMPLE}\tSM:{SAMPLE}" -t [THREADS] > [AMOSTRA].sam`

Opcional

Realize uma montagem *de novo* utilizando o pipeline `minimap2-miniasm`. Este processo pode levar várias horas/dias. Tome cuidado!

5. Gerar o arquivo BAM indexado:

- Repita os passos 3 e 4 do dia 2 utilizando `samtools`;

6. Visualização do Alinhamento:

- Importe o BAM para o IGV e avalie sua qualidade;
- Compare as amostras com suas respectivas amostras de short reads;

7. Realize a Genotipagem das Amostras de long reads:

- Utilize o `freebayes` para realizar a genotipagem e modo similar ao passo 3 do dia 3;

Tempo de execução da genotipagem em Long Reads

O tempo de execução do `freebayes` para arquivos de ONT pode ser bem demorado! Você pode optar por modificar o arquivo `bed` fornecido para trabalhar com menos regiões.

Para os formulários utilize o **vcf** resultante em **WorkshopDados/pre_run/Dia4/**

- Repita as métricas de filtragem utilizadas no passo 4 do dia 3;
- Repita os passos 5 e 6 do dia 3, comparando aos resultados de short reads.

Dia 5 - Genotipagem de STRs a partir de dados de NGS

Referências

- Bolger, Anthony M., Marc Lohse, and Bjoern Usadel. 2014. “Trimmomatic: A Flexible Trimmer for Illumina Sequence Data.” *Bioinformatics* 30 (15): 2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
- Danecek, Petr, Adam Auton, Goncalo Abecasis, Cornelis A. Albers, Eric Banks, Mark A. DePristo, Robert E. Handsaker, et al. 2011. “The Variant Call Format and VCFtools.” *Bioinformatics* 27 (15): 2156–58. <https://doi.org/10.1093/bioinformatics/btr330>.
- Danecek, Petr, James K Bonfield, Jennifer Liddle, John Marshall, Valeriu Ohan, Martin O Pollard, Andrew Whitwham, et al. 2021. “Twelve Years of SAMtools and BCFtools.” *GigaScience* 10 (2). <https://doi.org/10.1093/gigascience/giab008>.
- De Coster, Wouter, and Rosa Rademakers. 2023. “NanoPack2: Population-Scale Evaluation of Long-Read Sequencing Data.” Edited by Can Alkan. *Bioinformatics* 39 (5). <https://doi.org/10.1093/bioinformatics/btad311>.
- Garrison, Erik, and Gabor Marth. 2012. “Haplotype-Based Variant Detection from Short-Read Sequencing.” <https://doi.org/10.48550/ARXIV.1207.3907>.
- Li, Heng. 2013. “Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM.” <https://doi.org/10.48550/ARXIV.1303.3997>.
- . 2018. “Minimap2: Pairwise Alignment for Nucleotide Sequences.” Edited by Inanc Birol. *Bioinformatics* 34 (18): 3094–3100. <https://doi.org/10.1093/bioinformatics/bty191>.
- . 2021. “New Strategies to Improve Minimap2 Alignment Accuracy.” Edited by Can Alkan. *Bioinformatics* 37 (23): 4572–74. <https://doi.org/10.1093/bioinformatics/btab705>.
- Robinson, James T, Helga Thorvaldsdóttir, Wendy Winckler, Mitchell Guttman, Eric S Lander, Gad Getz, and Jill P Mesirov. 2011. “Integrative Genomics Viewer.” *Nature Biotechnology* 29 (1): 24–26. <https://doi.org/10.1038/nbt.1754>.

ANEXO: Dicas para uso do Ubuntu/WSL

Navegação e Diretórios

Listar Conteúdo do Diretório

```
ls
```

Mudar de Diretório

```
cd nome_do_diretorio
```

Diretório Atual

```
pwd
```

Criar Diretório

```
mkdir nome_do_novo_diretorio
```

Manipulação de Arquivos

Copiar Arquivo

```
cp arquivo_origem destino
```

Mover/Renomear Arquivo

```
mv arquivo_origem novo_nome_ou_destino
```

Remover Arquivo

```
rm nome_do_arquivo
```

Visualização de Conteúdo

Visualizar Conteúdo do Arquivo

```
cat nome_do_arquivo
```

Visualizar Conteúdo do Arquivo (página por página)

```
less nome_do_arquivo
```

Pacotes e Atualizações

Atualizar Lista de Pacotes

```
sudo apt update
```

Atualizar Pacotes Instalados

```
sudo apt upgrade
```

Instalar Novo Pacote

```
sudo apt install nome_do_pacote
```

Gerenciamento de Usuários

Adicionar Usuário ao Grupo

```
sudo usermod -aG nome_do_grupo nome_do_usuario
```

Mudar Senha do Usuário

```
passwd nome_do_usuario
```

Processos

Listar Processos

```
ps aux
```

Matar um Processo por ID

```
kill -9 processo_id
```

Monitorar Recursos do Sistema (htop)

```
htop
```

Medir Tempo de Execução de um Comando (time)

```
time comando_a_ser_medido
```

Rede

Verificar Configurações de Rede

```
ifconfig
```

Testar Conexão com um Endereço IP

```
ping endereco_ip
```

Entrada e Saída Padrão (stdin/stdout)

- **stdin (Standard Input):** É a entrada padrão de dados. Um programa pode ler dados a partir do stdin. Exemplo:
 - `cat < nome_do_arquivo`
- **stdout (Standard Output):** É a saída padrão de dados. Um programa geralmente imprime resultados no stdout. Exemplo:
 - `ls > lista_de_arquivos.txt`
- **stderr (Standard Error):** É a saída padrão para mensagens de erro. Exemplo:
 - `comando_inexistente 2> erro.log`

Outros Comandos Úteis

Ajuda sobre um Comando

`man nome_do_comando`

Sair do Terminal

`exit`