

Bioinformatics for Disease

Andrew Lynch

Table of contents

Preface	3
1 Molecular Sequencing Concepts and Technologies	4
I Data skills	5
2 Common data visualization techniques	7
II DNA	8
3 Aligning DNA sequences	9
4 DNA quality control	10
5 Variant calling	11
III RNA	12
6 Aligning RNA sequences	13
7 RNA quality control	14
8 Differential expression analysis	15
9 Gene set enrichment analysis (GSEA)	16
IV Cancer Analyses	17
10 Cancer drivers	18
11 Driver enrichment	19
12 Driver enrichment	20

V	Appendices	21
13	Reference versions	23
13.1	Downloading genome reference files	23
13.2	Additional Resources	23
14	Useful databases	24
14.1	Downloading genome reference files	24
14.2	Additional Resources	24
	References	25

Preface

This is a Quarto book.

To learn more about Quarto books visit <https://quarto.org/docs/books>.

1 + 1

[1] 2

1 Molecular Sequencing Concepts and Technologies

Part I

Data skills

In summary, this book has no content whatsoever.

1 + 1

[1] 2

2 Common data visualization techniques

Part II

DNA

3 Aligning DNA sequences

4 DNA quality control

Here I'm listing some useful reference material.

Chapter [14](#)

5 Variant calling

Part III

RNA

6 Aligning RNA sequences

7 RNA quality control

8 Differential expression analysis

9 Gene set enrichment analysis (GSEA)

Part IV

Cancer Analyses

10 Cancer drivers

11 Driver enrichment

12 Driver enrichment

Part V

Appendices

Here I'm listing some useful reference material.

Chapter [14](#)

13 Reference versions

13.1 Downloading genome reference files

Species	Version	Command
human	GRCh38/hg38	wget https://hgdownload.soe.ucsc.edu/goldenPath/hg38/bigZips/hg38.fa.gz
human	GRCh37/hg19	wget https://hgdownload.soe.ucsc.edu/goldenPath/hg19/bigZips/hg19.fa.gz
mouse	GRCm39/mm39	wget https://hgdownload.soe.ucsc.edu/goldenPath/mm39/bigZips/mm39.fa.gz
mouse	GRCm38/mm10	wget https://hgdownload.soe.ucsc.edu/goldenPath/mm10/bigZips/mm10.fa.gz

13.2 Additional Resources

1. [Broad Institute](#)

14 Useful databases

14.1 Downloading genome reference files

Species	Version	Command
human	GRCh38/hg38	wget https://hgdownload.soe.ucsc.edu/goldenP
human	GRCh37/hg19	wget https://hgdownload.soe.ucsc.edu/goldenP
mouse	GRCm39/mm39	wget https://hgdownload.soe.ucsc.edu/goldenP
mouse	GRCm38/mm10	wget https://hgdownload.soe.ucsc.edu/goldenP

14.2 Additional Resources

1. [Broad Institute](#)

References