

目次

第一部 生命データサイエンスの基礎体力づくり

第 1	章 Unix系環境の準備	中谷明弘	14
1	環境構築のための計算機の準備		15
2	プラットフォームごとの環境の準備		18
1	Microsoft Windows 11 または 10		18
2	Apple macOS		41
3	Linux		44
Column	次世代バイオビッグデータ時代に向けての人材育成	鈴木 穂	45
第 2	章 データ解析に向けた Unix系環境の使い方	中谷明弘	46
1	コマンドラインシェル (コマンドインタプリタ)		46
1	Bash (Bourne-again shell)		47
2	Unixのコマンド		47
3	ファイルとディレクトリ		48
4	スーパーユーザ (特権ユーザ) の管理権限		55
5	環境変数		56
6	リダイレクトとパイプ		58
2	パッケージマネージャ		64
1	APT (Ubuntuの場合)		64
2	MacPortsとHomebrew (macOSの場合)		66
3	テキストエディタ		70
1	vi		70
2	Emacs		71

4	Python 処理系の準備.....	73
1	Anaconda のインストール.....	74

第 3 章 Python によるデータ解析の基礎 中谷明弘 78

1	Python の起動と実行.....	78
1	対話型インタプリタを介した実行.....	78
2	スクリプトファイルとして実行.....	79
2	Python の基礎.....	81
1	データと変数.....	82
2	データ構造.....	85
3	文字列のパターンマッチング.....	92
4	制御構造.....	96
5	関数.....	99
6	ファイルの入出力.....	102
7	ソースコードの分割.....	106
8	起動時のコマンドライン引数.....	108

第 4 章 Python によるデータ解析の実践 中谷明弘 110

1	実践：matplotlib によるグラフの描画.....	110
2	実践：VCF 形式ファイル内の変異の集計.....	114

第 5 章 R 言語によるデータ解析の基礎と実践① 鹿島幸恵, 関 真秀, 鈴木絢子 122

1	解析環境と R のインストール.....	123
2	パッケージのインストール.....	125
3	R の基礎.....	127
1	起動と終了.....	127

2	基本操作.....	127
3	テーブル入力・出力.....	130
4	オブジェクトの保存・読み込み.....	131
5	Rスクリプトの実行.....	132
4	実践：遺伝子発現量の散布図を描画する.....	133
5	実践：特定の遺伝子における発現量の箱ひげ図を描画する.....	138
6	実践：発現レベルが>5 RPKMである遺伝子数を棒グラフで可視化する.....	140

第6章 R言語によるデータ解析の実践② 鹿島幸恵, 関 真秀, 鈴木絢子 144

1	実践：DEG解析.....	145
2	実践：階層的クラスタリング・ヒートマップ描画.....	149
3	実践：エンリッチメント解析.....	154
4	応用編：Rを用いたがんゲノムアトラス（TCGA）データの抽出.....	160

第7章 スパコン利用のための はじめの一步 鹿島幸恵, 関 真秀, 鈴木絢子 167

1	大規模計算機での解析.....	167
1	SSH接続.....	167
2	ファイルのやり取り.....	171
3	ジョブ管理システムとディスククォータ.....	173
2	国内におけるスパコンについて.....	174
1	バイオデータ解析向けのスーパーコンピュータ.....	174
Column	ヒトゲノム倫理とスパコン.....	鈴木 稯 175

第1章 オミクス解析の準備 鹿島幸恵, 関 真秀, 鈴木絢子 178

1	オミクス解析外観	178
2	オミクスシーケンス解析	180
3	解析環境について	182
4	主要なデータフォーマット	183
5	リファレンスゲノムについて	185
6	各種リソース・データベースについて	188
Column	ゲノム指針の改定	鈴木 穰 191

第2章 ゲノム解析 鹿島幸恵, 関 真秀, 鈴木絢子 192

1	実践：ゲノムシーケンスデータからの点変異の検出	193
2	その他のゲノム解析手法・ツールについて	208
Column	ゲノム配列を取り巻く国際情勢	鈴木 穰 211

第3章 トランスクリプトーム解析 鹿島幸恵, 関 真秀, 鈴木絢子 212

1	実践：RNA-seq データからの発現量算出とヒートマップ描画	215
Column	配列解析によるオミクス解析の未来	237

第4章 エピゲノム解析 鹿島幸恵, 関 真秀, 鈴木絢子 238

1	実践：ChIP-seq データからのヒストン修飾のパターン解析	240
2	その他のエピゲノム解析手法	249

第 **5** 章 シングルセル解析① 鹿島幸恵, 関 真秀, 鈴木絢子 251

- 1** 実践：scRNA-seq データの一次解析253
- 2** 実践：scATAC-seq データの一次解析259
- Column** 計測技術の進展とデータ量 鈴木 穰 267

第 **6** 章 シングルセル解析② 鹿島幸恵, 関 真秀, 鈴木絢子 268

- 1** 実践：scRNA-seq データからのクラスタリングと細胞種同定268
- 2** 実践：scATAC-seq データからのクラスタリングと coverage プロットの作成288
- 3** 実践：scRNA-seq と scATAC-seq のデータ統合300

第 **7** 章 ロングリード解析 鹿島幸恵, 関 真秀, 鈴木絢子 306

- 1** 実践：全長 cDNA-seq データからの発現量算出と短鎖 RNA-seq データとの比較311
- 2** 実践：長鎖ゲノムシーケンスデータからの構造変異の検出321
- 3** 実践：長鎖ゲノムシーケンスデータからのメチル化領域の同定328
- 4** その他ロングリードシーケンス技術の応用333

索引 336
