

目次

はじめに	河野暢明	3
本書の使い方		10
プログラミングの基本事項	河野暢明	12

第 I 部 生命科学データ別にファーストアクションを学ぶ

第 1 章	ゲノム	河野暢明	22
-------	------------	------	-----------

はじめに	23
1 ゲノム統計値	24
1.1 ゲノム統計値の計算	24
1.2 assembly-stats による描画	29
2 コンタミネーション評価	33

第 2 章	トランスクリプトーム	河野暢明, 山本フィリップ	41
-------	-------------------	---------------	-----------

はじめに	42
1 遺伝子発現の定量, 発現変動遺伝子の抽出, 機能解析	43
1.1 遺伝子発現の定量	44
1.2 発現変動遺伝子の抽出 (DEG 解析)	48
1.3 機能解析 (GO エンリッチメント解析)	50
2 ドットプロット	57
3 バブルプロット	60
4 ネットワーク	63
5 ウェブアプリケーションの紹介	66

第3章

メタゲノム /16S rRNA アンプリコンシーケンスデータ

平岡聡史 69

はじめに	70
1 シーケンスデータのクオリティコントロール (QC)	72
2 全メタゲノムに対する配列データのカバレッジ推定	78
3 群集構造解析	81
4 NMDS (非計量多次元尺度構成法) 解析	85
5 遺伝子機能アノテーション	88
6 メタゲノムアセンブリ	92
7 MAGの構築と系統推定	95
8 16S rRNA アンプリコン配列解析	102

第4章

ChIP (クロマチン免疫沈降)-seq

大島 拓, 深町和貴 108

はじめに	109
1 ピークの検出	111
1.1 ChIP-seq データピーク検出手順	113
2 ピーク位置の推定とコンセンサス配列の決定	115
2.1 転写因子が関わる生理活性	115
2.2 転写因子が結合している DNA の推定	116
2.3 ピーク配列推定手順	117
2.4 ピーク推定とバックグラウンド	118
2.5 コンセンサス配列の推定	120
3 複数のChIP-seqデータの比較	121
まとめ	122

第5章

Hi-C データ

竹俣直道 124

はじめに	125
1 データのクオリティチェックのための <i>cis</i> 相互作用定量	126
2 ヒートマップによるHi-Cコンタクトマップの可視化	128
3 サンプル間での相互作用の変化の可視化	129
4 A/Bコンパートメントの形成パターンの可視化	130
5 サドルプロットを用いたA/Bコンパートメント強度の解析	132
6 insulation scoreのプロット1	133
7 insulation scoreのプロット2	137
8 contact probabilityのプロット	139
9 contact probabilityの一次導関数解析	140

第6章

シングルセルトランスクリプトーム

公開データとの統合・比較解析

佐伯憲和 143

はじめに	144
モジュールインポート・環境準備	145
1 取得データ (クエリデータ) の解析	146
2 リファレンスデータへのマッピング	153
2.1 リファレンスデータの準備 (解析例: 内胚葉オルガノイドアトラス)	153
2.2 scPoli を用いたリファレンスマッピングとラベル転移	156
3 統合・ラベル転移データを用いた発現プロファイル比較	160

第7章

プロテオーム

網羅的なタンパク質解析と図表作成・データ登録

岩崎未央 166

はじめに	167
1 Perseus を用いた PCA 解析とクラスタリング解析の可視化	168
1.1 テキストファイルの作成	168
1.2 Perseus の基本的な使い方	173
2 質量分析データのリポジトリ登録	180

第8章

顕微鏡データ デジタルステイニングによる細胞核の可視化
と形態解析

増田正人, 山中大介, 宮本崇史 189

はじめに	190
1 市販の画像解析ソフトウェアを用いたデジタルステイニング	191
2 Python を用いた核の画像の二値化	192
3 Python を用いた核の形態の解析	195
4 Python を用いた核の形態情報の UMAP によるクラスタリング	206

第9章

高速原子間力顕微鏡 (AFM) データ

西口茂孝 210

はじめに	211
1 高速 AFM 画像の見方と表示方法	211
2 高速 AFM 画像の質の評価	213
3 高速 AFM 画像の前処理	214
4 分子の大きさの解析	216
5 分子の高さの解析	218
6 分子の動きの解析	220

第Ⅱ部 LLM との対話を通じて理想のデータビジュアルをつくる

第1章 データから適切なグラフを選ぶ方法 中川 恒 224

はじめに	224
1 データ可視化手法を探索する	224
2 データ可視化に ChatGPT を利用する	228
3 ChatGPT を用いたデータ解析や可視化時の注意点	232

第2章 環状ゲノムマップ Circos を用いた作成方法 竹川瑠華 233

1 環状ゲノムマップの読み方と読み取れる情報	234
2 環状ゲノムマップが用いられるシーンと Circos の概要	236
3 Circos で描画を行う手順	236

第3章 ヒートマップ 藤本悠人 261

はじめに	262
1 ヒートマップ作成の準備	262
2 使用データの準備	264
3 ヒートマップの作成	267
Column 誰もが見やすい図を作ろう	269

第4章 主成分分析 (PCA) 山本フィリップ 271

はじめに～主成分分析 (PCA) とは	272
1 メタボロミクスの PCA	272
2 シングルセルプロテオミクスの PCA	281

第5章 注釈付き系統樹 荻野 馨, 河野暢明 290

はじめに	291
1 注釈付き系統樹の読み方	291
2 注釈付き系統樹の作成	291

第6章

ChIP (クロマチン免疫沈降)-seq ピーク解析

中川 恒 301

はじめに	302
1 ChIP-seq 解析の図の読み方	302
2 ChIP-seq 解析とは	303
3 実行コード	304
4 WashU Epigenome Browser での可視化	309

第7章

UMAP・Trajectory 解析

シングルセル RNA-seq データの描画

増井真那 311

はじめに	312
1 次元削減手法の概要	312
2 UMAP プロットの作成	313
3 次元削減後の解析 (シングルセルTrajectory 解析)	319

第8章

MA プロット・Volcano プロット

遺伝子発現変動を示す図の描き方

RStudio を使用した

松本美亜, 河野暢明 328

はじめに	328
1 DEG 解析とは	329
2 MA プロット	330
3 Volcano プロット	338

索引	345
----	-----

執筆者一覧	348
-------	-----