

Bioinformatica
Programma d'esame
dal testo "Introduzione alla Bioinformatica" ed. Zanichelli

Capitolo 1 Introduzione alla bioinformatica 1

- 1.1 **Cenni introduttivi 1**
- 1.2 **Pietre miliari della bioinformatica 2**
- 1.3 **Infrastrutture bioinformatiche 4**
- 1.4 Indirizzi web delle più rilevanti infrastrutture bioinformatiche 5
- 1.5 Letture di approfondimento consigliate 6

Capitolo 2 Banche dati biologiche 7

- 2.1 **Introduzione 7**
- 2.2 **Sistemi di interrogazione delle banche dati biologiche 8**
 - 2.2.1 SRS 9
 - 2.2.2 Entrez 11
 - 2.2.3 ACNUC 12
 - 2.2.4 ACEDB 13
- 2.3 **Banche dati primarie e banche dati specializzate 13**
- 2.4 **Banche dati di sequenze nucleotidiche 16**
- 2.5 **Banche dati di sequenze proteiche 18**
- 2.6 **Banche dati di motivi e domini proteici 19**
- 2.7 **Banche dati di strutture proteiche 25**
- 2.8 **Banche dati biologiche per il sistema immunitario 26**
 - 2.8.1 IMGT 26
 - 2.8.2 Banche dati di peptidi correlati al complesso MHC 26
- 2.9 **Risorse genomiche 27**
- 2.10 **Banche dati di geni 28**
- 2.11 **Banche dati del trascrittoma 30**
- 2.12 **Banche dati di profili di espressione 30**
- 2.13 **Banche dati di polimorfismi e mutazioni 32**
- 2.14 **Banche dati di *pathways* metabolici 34**
- 2.15 **Banche dati mitocondriali 36**
- 2.16 **Indirizzi web per banche dati e risorse biologiche 39**

Capitolo 3 Allineamento di sequenze di acidi nucleici e proteine 41

- 3.1 **Introduzione al problema degli allineamenti di sequenze 41**
 - 3.1.1 **Similarità di sequenze e algoritmi di allineamento 41**
 - 3.1.2 **Allineamenti di sequenze biologiche con gap 43**
 - 3.1.3 **Visualizzazione di dot matrix 44**
- 3.2 **Le matrici di sostituzione 45**
 - 3.2.1 **Le matrici PAM 46**
 - 3.2.2 **Le matrici BLOSUM 48**
- 3.3 **Metodi di allineamento esatto 49**
 - 3.3.1 **Allineamenti globali e locali 49**
 - 3.3.2 **Algoritmi dinamici di allineamento 50**
 - 3.3.3 **L'algoritmo di Smith e Waterman per la ricerca di similarità locali 53**
- 3.4 **Metodi euristici di allineamento 54**
 - 3.4.1 **Ricerca di similarità in banche dati: FASTA 54**
 - 3.4.2 **BLAST 58**
 - 3.4.3 **Altri programmi di allineamento 62**
- 3.5 **Principali risorse disponibili in rete 63**
- 3.6 **Letture di approfondimento consigliate 63**

Capitolo 4 **Allineamento multiplo di sequenze** 65

4.1 Introduzione 65

4.2 Algoritmi per l'allineamento multiplo 66

4.3 Misura della qualità di un allineamento multiplo 70

4.4 Strumenti per la visualizzazione e manipolazione dei multiallineamenti 71

4.5 Generazione e applicazione dei profili dei multiallineamenti 73

4.6 Principali risorse disponibili in rete 78

4.7 Letture di approfondimento consigliate 79

Capitolo 5 **Ricerca di pattern e di motivi funzionali** 81

5.1 Introduzione 81

5.1.1 Definizione di motivo 81

5.1.2 Misure della sensibilità e della specificità di un motivo 82

5.2 Ricerca di pattern e di motivi funzionali in sequenze nucleotidiche 84

5.2.1 Generalità 84

5.2.2 Filtri per l'esclusione delle sequenze ripetitive 85

5.2.3 Ricerca e interpretazione dei promotori eucariotici 85

5.2.4 Siti di giunzione tra introni ed esoni 87

5.2.5 Siti di inizio della traduzione 88

5.2.6 Segnali di terminazione 89

5.2.7 Analisi della strategia dell'uso del codice genetico 89

5.2.8 Conclusioni 92

5.3 Ricerca di pattern e di motivi funzionali in sequenze proteiche 92

5.3.1 Generalità 92

5.3.2 PROSITE 93

5.3.3 Sequenze segnale 94

5.4 Metodi di apprendimento automatico: reti neurali, *Hidden Markov Models* e algoritmi genetici 95

5.4.1 Generalità 95

5.4.2 Le reti neurali 95

5.4.3 Catene di Markov e Hidden Markov Models (HMM) 98

5.4.4 Algoritmi genetici 101

5.5 Principali risorse disponibili in rete 102

5.6 Letture di approfondimento consigliate 103

Capitolo 7 **Analisi strutturale delle proteine** 135

7.1 Introduzione 135

7.1.1 Struttura delle proteine 136

7.2 Metodi classici per la determinazione della struttura proteica 137

7.2.1 Diffrazione ai raggi X 137

7.2.2 Spettroscopia a risonanza magnetica nucleare (NMR) 138

7.2.3 Conclusioni 138

7.3 Banche dati di strutture proteiche 140

7.3.1 PDB 140

7.3.2 MMDB 142

7.3.3 DSSP 142

7.3.4 HSSP 143

7.3.5 FSSP 144

7.3.6 SCOP 144

7.3.7 CATH 145

7.4 Programmi di visualizzazione 145

7.4.1 Metodi di rappresentazione delle strutture proteiche 145

7.4.2 Programmi di grafica molecolare 147

7.5 Ricerche di similarità strutturale 147

7.5.1 DALI 148

- 7.5.2 SSAP 148
- 7.5.3 CE 149
- 7.6 **Metodi per la predizione della struttura secondaria** 149
 - 7.6.1 PHD 151
 - 7.6.2 PSIPRED 152
 - 7.6.3 PREDATOR 152
 - 7.6.4 JPRED 153
- 7.7 **Metodi per la predizione della struttura terziaria** 153
 - 7.7.1 Homology modelling 154
 - 7.7.2 Threading 156
- 7.8 Principali risorse disponibili in rete 157
- 7.9 Letture di approfondimento consigliate 158

Capitolo 11 **Genomica funzionale** 223

- 11.1 **Metodi sperimentali per l'analisi dell'interazione proteina-proteina** 223
 - 11.1.1 Tecnica del doppio ibrido in lievito. 224
 - 11.1.2 Spettrometria di massa per l'identificazione di complessi proteici 225
 - 11.1.3 Interazioni genetiche 226
 - 11.1.4 FRET 227
 - 11.1.5 Conclusioni 227
- 11.2 **Predizioni bioinformatiche di interazioni proteina-proteina** 228
 - 11.2.1 Mutazioni correlate 228
 - 11.2.2 Docking 229
 - 11.2.3 SPOT: uso di frequenze di residui in posizioni di contatto tra proteine 230
- 11.3 **Associazioni funzionali indirette** 232
 - 11.3.1 Predizione dell'interazione tra proteine omologhe a proteine che interagiscono 232
 - 11.3.2 Profili filogenetici 232
 - 11.3.3 Stele di Rosetta 233
 - 11.3.4 Similarità di alberi filogenetici 234
 - 11.3.5 Prossimità genomica 235
- 11.4 **Banche dati di interazioni proteiche** 235
- 11.5 Principali risorse disponibili in rete 237
- 11.6 Letture di approfondimento consigliate 237

Appendice B **Elementi di informatica** 281

- B.1 **Struttura e funzionamento di un personal computer** 281
 - B.1.1 Calcolatori multiutente 281
 - B.1.2 Struttura di un calcolatore multiutente 282
 - B.1.3 Rete di computer 282
 - B.1.4 Cos'è un sistema operativo 282
 - B.1.5 Cos'è un «file» 283
 - B.1.6 Cos'è una «directory» 283
 - B.1.7 Cos'è un programma per calcolatore 284
- B.2 **Introduzione ai sistemi operativi *linux* e *unix*** 285