

Spis treści

Przedmowa do piątego wydania	XIII
Schemat książki	XIX
Wprowadzenie do bioinformatyki w sieci	XX
Podziękowania	XXI
1. Wprowadzenie	1
Życie w czasie i przestrzeni	4
Fenotyp = genotyp + środowisko + historia życia + epigenetyka	5
Ewolucja jako zmiana na przestrzeni czasu w świecie istot żywych	5
Klasyfikacja biologiczna i nazewnictwo	7
Dogmaty: główne i poboczne	10
Struktura DNA	10
Transkrypcja i translacja	13
Struktury białek	14
Statyka i dynamika	19
Biologia systemowa	19
Genom człowieka	21
Zmiany w sekwencjach genomu człowieka	22
Genom człowieka a medycyna	23
Bazy danych w biologii molekularnej	31
Archiwa rzeczy zauważalnych i danych	32
Baza danych bez skutecznych sposobów dostępu jako bezużyteczne narzędzie	33
Przepływ informacji w bioinformatyce	34
Organizacja, adnotacja i kontrola jakości	36
Baza danych WWW	37
Publikacje elektroniczne	38
Komputery i informatyka	38
Programowanie	39
Après moi, le déluge? Przepraszam, już za późno!	43
Jaka jest moc sekwencjonowania na świecie?	46
Jaki jest stosunek ilości danych w bioinformatyce do innych naukowych archiwów informacji?	46
Polecana literatura	47
Ćwiczenia i zagadnienia problematyczne	48
2. Od genetyki do genomów	53
Klasyczne tło genetyki	54
DNA reprezentuje geny	55
Mapy i przewodniki	56

Mapy połączeń genetycznych	56
Sprzężenia	56
Prążkowanie chromosomów	58
Mapy wysokiej rozdzielczości bazujące bezpośrednio na sekwencjach DNA	60
Mapy restrykcyjne	62
Sekwencjonowanie DNA	63
Frederick Sanger i rozwój sekwencjonowania DNA	63
Sekwencjonowanie DNA poprzez przerywanie replikacji łańcucha	64
Automatyzacja sekwencjonowania DNA	65
Sekwencjonowanie następnej generacji	66
Odczyty sparowane	72
Życie w pędzie	73
Zestawianie - aspekty obliczeniowe	74
Dopasowywanie wzorców	74
Drzewa sufiksowe	74
Łączenie fragmentów	76
Genomika w identyfikacji tożsamości	79
Genetyczne odciski palców	80
Identyfikacja tożsamości poprzez amplifikację określonych regionów w miejsce analizy polimorfizmu długości fragmentów restrykcyjnych (RFLP)	81
DNA mitochondrialne	83
Analiza sekwencji DNA nienależących do ludzi	84
Badania na ustalenie ojcostwa	85
Aspekty etyczne, prawne i społeczne	87
Bazy danych zawierające informacje o sekwencji DNA u ludzi	87
Wykorzystanie sekwencjonowania DNA w badaniach na organizmach ludzkich	90
Polecana literatura	90
Ćwiczenia i zagadnienia problematyczne	91
3. Panorama życia	95
Genomy, transkryptomy i proteomy	96
Geny	96
Proteomika i transkryptomika	99
Podsluchiwanie na transmisji informacji genetycznej	100
Projekty sekwencjonowania genomu	101
Genomy prokariotów	102
Genom bakterii <i>Escherichia coli</i>	103
Genom archeonu <i>Methanocaldococcus jannaschii</i>	105
Genom jednego z najprostszych organizmów: <i>Mycoplasma genitalium</i>	106
Metagenomika: pobranie genomów z próbek środowiskowych	107
Ludzki mikrobiom	109
Genomy eukariotów	110
Rodziny genów	111
Genom <i>Saccharomyces cerevisiae</i> (drożdży piekarniczych)	111
Genom <i>Caenorhabditis elegans</i>	113
Genom <i>Drosophila melanogaster</i>	115

Genom <i>Arabidopsis thaliana</i>	115
Genom <i>Homo sapiens</i> (genom człowieka)	117
Geny kodujące białka	118
Sekwencje powtarzalne	118
RNA	119
Polimorfizmy pojedynczego nukleotydu i haplotypy	119
Systematyczne pomiary i zbiory polimorfizmów pojedynczego nukleotydu	122
Różnorodność genetyczna w antropologii	123
Sekwencje DNA a języki	125
Ewolucja genomów	125
Proszę przekaż geny: poziomy transfer genów	128
Genomika porównawcza u eukariontów	129
Polecana literatura	130
Ćwiczenia i zagadnienia problematyczne	131
4. Dopasowania i drzewa filogenetyczne	133
Wprowadzenie do dopasowania sekwencji	134
Macierz punktowa a dopasowanie sekwencji	140
Miary podobieństwa sekwencji	142
Systemy punktacji	143
Pochodne od macierzy substytucji: macierze PAM	144
Obliczanie dopasowywania dwóch sekwencji	146
Odmiany i uogólnienia	146
Metody aproksymacji dla szybkiej identyfikacji baz danych	146
Algorytm programowania dynamicznego dla optymalnego dopasowania dwóch sekwencji	148
Znaczenie dopasowań	153
Dopasowanie wielu sekwencji	155
Zastosowanie dopasowania wielu sekwencji do przeszukiwania baz danych	156
Profile	158
PSI-BLAST	160
Całkowite dopasowanie par sekwencji dla ludzkiego białka PAX-6 i <i>Drosophila melanogaster</i> <i>yeless</i>	164
Ukryte modele Markowa (Hidden Markov Models, HMM)	165
Filogeneza	167
Określanie powiązań taksonomicznych na podstawie właściwości molekularnych	169
Wykorzystanie sekwencji przy określaniu związków filogenetycznych	172
Zastosowanie SINE oraz LINE do określania związków filogenetycznych	175
Drzewa filogenetyczne	177
Metody oparte na grupowaniu	179
Metoda największej wiarygodności	180
Rekonstrukcja sekwencji ancestralnych	180
Dekarboksylaza pirogronianu: synteza, aktywność i struktura krystaliczna przewidywanego przodka	181
Problem zmiennego tempa ewolucji	183
Metoda Bayesa	184

Czy drzewa są właściwym sposobem przedstawiania związków filogenetycznych?	184
Kwestie obliczeniowe	185
Podsumowanie	186
Polecana literatura	187
Ćwiczenia i problemy	187

5. Bioinformatyka strukturalna a opracowywanie leków 193

Wprowadzenie	194
Stabilność a fałdowanie białek	196
Wykres Ramachandrana jako opis dopuszczalnych konformacji łańcucha głównego	196
Łańcuchy boczne	198
Stabilność i denaturacja białek	198
Proces zwijania białek	202
Zastosowania hydrofobowości	204
Białka superhelikalne (<i>coiled-coil</i>)	204
Opis zróżnicowania struktur białek	208
Superpozycja struktur a dopasowanie strukturalne	209
Ewolucja struktur białek	214
Klasyfikacja struktur białek	216
SCOP	217
Przewidywanie i modelowanie struktur białek	219
Ocena krytyczna przewidywania struktur	221
Prognozowanie struktur drugorzędowych	222
Modelowanie homologiczne	223
Rozpoznawanie fałdu	223
Obliczanie energii konformacji i dynamiki molekularnej	226
ROSETTA	228
Przewidywanie struktury białek na podstawie map kontaktów pochodzących od skorelowanych mutacji w dopasowaniach wielu sekwencji	229
Tworzenie nowych białek	232
Opracowywanie i doskonalenie leków	235
Związek wiodący	236
Doskonalenie związku wiodącego: ilościowe modele zależności struktura-aktywność	237
Wykorzystanie bioinformatyki przy opracowywaniu i doskonaleniu leków	238
Modelowanie molekularne przy opracowywaniu leków	239
Polecana literatura	245
Ćwiczenia i problemy	247

6. Publikacje naukowe i archiwa: media, treść, dostęp i prezentacja 253

Literatura naukowa	254
Dostęp do publikacji naukowych	255
Otwarty dostęp	257
Public Library of Science	258
Biblioteki tradycyjne i cyfrowe	258

Jak zapełnić bibliotekę cyfrową	259
Eksploracja informacyjna	260
Sieć: wyższy wymiar	260
Nowe media: wideo, dźwięk	260
Przeszukiwanie literatury naukowej	261
Zarządzanie bibliografią	261
Bazy danych	263
Zawartość bazy danych	263
Kontrola jakości baz danych	264
Literatura jako baza danych	265
Organizacja bazy danych	265
Adnotacja	268
Języki znaczników	269
Dostęp do baz danych	271
Linki	271
Interoperacyjność baz danych	272
Eksploracja danych	276
Języki programowania i narzędzia do budowania i dostępu do baz danych	277
Tradycyjne języki programowania	277
Języki skryptowe	278
Biblioteki programów przeznaczone dla biologii molekularnej	278
Java - obliczenia w sieci	278
Przetwarzanie języka naturalnego	279
Przetwarzanie języka naturalnego przy przeszukiwaniu literatury biomedycznej	280
Zastosowania biomedyczne eksploracji tekstu	281
Tworzenie hipotez	286
Sieć związana z jaskrą uzyskana przez eksplorację danych	287
Polecana literatura	290
Ćwiczenia i problemy	291
7 Sztuczna inteligencja i uczenie maszynowe	293
Czym jest sztuczna inteligencja i uczenie maszynowe?	294
Klasyfikacja i grupowanie	295
Klasyfikator binarny	298
Krzywe ROC (Receiver Operating Characteristic)	300
Sztuczne sieci neuronowe	301
Mapy samoorganizujące	303
Drzewa decyzyjne	304
Maszyny wektorów nośnych (SVM)	308
Metody jądra	308
Grupowanie	310
Grupowanie za pomocą spektralnej teorii grafów	313
Polecana literatura	316
Ćwiczenia i problemy	316
8 Wprowadzenie do biologii systemów	319
Wprowadzenie	320

Sieci i grafy	321
Spójność w sieciach	322
Dynamika, stabilność i odporność na awarie	324
Niektóre źródła inspiracji dla biologii systemów	325
Złożoność sekwencji	325
Definicja entropii Shannona	326
Złożoność sekwencji	327
Zależność między złożonością, losowością i zdolnością do kompresji	328
Transformata Burrowsa-Wheelera	329
Odwracanie transformaty Burrowsa-Wheelera	329
Transformata Burrowsa-Wheelera grupuje powtórzenia, ułatwiając kompresję	330
Zastosowanie transformaty Burrowsa-Wheelera do poszukiwania wzorców w łańcuchach	330
Złożoność w innych typach danych biologicznych	331
Złożoność statyczna i dynamiczna	332
Przewidywalność i chaos	333
Analiza i porównywanie sieci	334
Analiza grafów za pomocą algebry macierzy	335
Izomorfizm grafów	335
Polecana literatura	338
Ćwiczenia i problemy	338
9 Szlaki metaboliczne	341
Wprowadzenie	342
Klasyfikacja funkcji białek	344
Komisja Enzymowa	344
Klasyfikacja funkcji białek przez Gene Ontology™ Consortium	344
Przewidywanie funkcji białka	345
Kataliza enzymatyczna	348
Miejsca aktywne	349
Kofaktory	349
Równowagi wiązania białko-ligand	350
Kinetyka reakcji enzymatycznych	351
Miary wydajności enzymów	353
Jak ewoluują nowe funkcje enzymów?	353
Kontrola aktywności enzymu	354
Mechanizmy strukturalne ewolucji zmienionych albo nowych funkcji białek	355
Szlaki i granice dywergencji sekwencji, struktury i funkcji	359
Ewolucja drogą duplikacji genów	359
Bazy szlaków metabolicznych	362
Kyoto Encyclopedia of Genes and Genomes (KEGG)	364
Ewolucja i filogeneza szlaków metabolicznych	366
Porównanie szlaków	366
Dopasowanie szlaków metabolicznych	368
Porównywanie liniowych szlaków metabolicznych	369
Porównywanie nieliniowych szlaków metabolicznych: szlak pentozofosforanowy i cykl Calvina-Bensona	371

Dynamika sieci metabolicznych	372
Odporność sieci metabolicznych	372
Modelowanie dynamiczne metabolizmu	373
Symulacja szlaków metabolicznych u <i>Plasmodium falciparum</i>	376
Human Metabolome Database wspiera zastosowania kliniczne w badaniach wrodzonych wad metabolicznych i raka	377
Polecana literatura	379
Ćwiczenia i problemy	379
10. Kontrola organizacji i organizacja kontroli	381
Transkryptomika	382
Projekt ENCODE	384
Wyznaczanie sekwencji RNA	385
Sekwencjonowanie RNA kontra mikromacierze	385
Mikromacierze DNA	385
Sekwencjonowanie RNA (RNAseq)	391
Projekt Genotype-Tissue Expression (GTEx)	392
Wzorce ekspresji w różnych stanach fizjologicznych	394
Zmiany wzorców ekspresji w cyklu życiowym <i>Drosophila melanogaster</i>	394
Różne stadia życiowe mają różne wymagania w stosunku do różnych genów	396
1. Zmiany we wzorcach ekspresji genów	398
2. Mutacje, które dają oporność na izoniazyd	399
3. Krystalografia	399
Kompleksy i agregaty białek	399
Właściwości kompleksów białko-białko	400
Sieci oddziaływań białek	402
Składniki kompleksu prymosomu u <i>Bacillus subtilis</i>	405
Sieci regulatorowe	406
Przekazywanie sygnału i kontrola transkrypcyjna	407
Biologia strukturalna sieci regulatorowych	407
Przykłady stosunkowo prostych regulatorowych sieci kontroli	410
Regulacja operonu laktozowego u <i>E. coli</i>	410
Przełącznik genowy bakteriofaga λ	412
Przeskok metaboliczny u <i>Saccharomyces cerevisiae</i>	416
Struktura logiczna sieci regulatorowych	418
Sieć regulatorowa transkrypcji u <i>E. coli</i>	418
Sieć regulatorowa transkrypcji u <i>Saccharomyces cerevisiae</i>	419
Zdolności adaptacyjne sieci regulatorowej drożdży	420
Polecana literatura	423
Ćwiczenia i problemy	423
Wnioski	426