

Bioinformatik

Sequenz - Struktur - Funktion

von

Reinhard Rauhut

WILEY-VCH

Weinheim - New York - Chichester - Brisbane - Singapore - Toronto

Inhaltsverzeichnis

| | |
|---|-----|
| Vorwort | V |
| Einleitung | 1 |
| Bioinformatik - Biologische Wissenschaften im 21. Jahrhundert | 1 |
| Empfohlene Literatur | 7 |
| 1 Sequenzen | 9 |
| 1.1 Der Evolutionsverlauf des Planeten Erde, die molekulare Evolution biologischer Systeme und die Suche nach Ähnlichkeiten | 9 |
| 1.2 Sequenzdatenbanken | 14 |
| 1.2.1 Das Beispiel GenBank | 24 |
| 1.2.2 Der NCBI Datenverbund und ENTREZ | 28 |
| 1.2.3 LocusLink und RefSeq | 28 |
| 1.2.4 UniGene | 30 |
| 1.3 Proteindatenbanken | 30 |
| 1.4 Alignments - Ähnlichkeiten zwischen Sequenzen | 38 |
| 1.4.1 Wie definiert und wie mißt man Ähnlichkeiten? | 38 |
| 1.4.2 Ein Wahrscheinlichkeitsmodell für Alignments - Algorithmen, gaps, Matrizen | 42 |
| 1.4.3 Die mathematische Entwicklung globaler und lokaler Alignments | 50 |
| IAA Was ist signifikant? | 59 |
| 1.4.5 Homologiesuche mit BLAST | 60 |
| 1.5 Das Identifizieren von ORFs in genomischer DNA | 80 |
| 1.5.1 Eukaryontische Gene | 80 |
| 1.5.2 Prokaryontische Gene | 83 |
| 1.6 Markov Modelle | 86 |
| 1.6.1 Beispiel CpG Inseln | 86 |
| 1.6.2 HMMs als Sequenzemitter oder Sequenzgenerator | 90 |
| 1.6.3 Hidden Markov Models und multiple Alignments | 94 |
| 1.6.4 Motive und Domänen: Prosite, Blocks, Pfam, Prodom | 103 |

VIII Inhaltsverzeichnis

| | | |
|----------|--|-----|
| 2 | Strukturen | 109 |
| 2.1 | Wie falten sich Proteine? | 109 |
| 2.1.1 | Grundlegende Konzepte | 109 |
| 2.1.2 | Strukturvorhersage | 114 |
| 2.1.3 | Ansätze zur <i>de novo</i> Faltungsvorhersage | 225 |
| 2.1.4 | Sekundärstrukturvorhersage in Proteinen | 120 |
| 2.1.5 | Threading (fold recognition) Methoden | 121 |
| 2.1.6 | Homology Modeling mit SWISS-Model | 124 |
| 2.2 | Strukturdatenbanken | 127 |
| 2.2.1 | Protein Database Files | 129 |
| 2.2.2 | Molecular Modeling Database des NCBI | 133 |
| 2.3 | Vorhersage von RNA-Strukturen | 237 |
| 2.4 | Pattern-Suche | 242 |
| 2.5 | Die Klassifizierung von Proteinstrukturen | 246 |
| 2.5.1 | Die hierarchische SCOP Klassifizierung | 248 |
| 2.5.2 | Die Beziehung zwischen Sequenz, Struktur und Funktion | 254 |
| 2.5.3 | Structural Genomics - Strukturelle Klassifizierung von Genomen | 267 |
| 3 | Genomics | 277 |
| 3.1 | Orthologe, Paraloge und globaler Aufbau von Genomen | 277 |
| 3.2 | Cluster von orthologen Gruppen | 284 |
| 3.3 | Wie sequenziert man Genome? | 292 |
| 4 | Functional Genomics | 297 |
| 4.1 | DNA Chiptechnologie und Expressionsarrays | 297 |
| 4.1.1 | Die Chipherstellung | 298 |
| 4.1.2 | Das experimentelle Prinzip und die Einsatzgebiete | 299 |
| 4.2 | Das Modell <i>Saccharomyces cerevisiae</i> | 201 |
| 4.2.1 | Expressionsanalyse mit Hefe-Chip | 204 |
| 4.2.2 | Mutanten und Chiptechnologie | 220 |
| 4.2.3 | Genomweite Mutantensammlungen | 225 |
| 4.3 | Anwendungsgebiete für Chiptechnologie | 226 |
| 4.4 | Chiptechnologie in der Pharmaforschung | 227 |
| 4.4.1 | Das Beispiel Tumorzelllinien | 228 |
| 4.5 | Pharmakogenetik | 226 |

| | | |
|----------|--|------------|
| 5 | Proteomics | 235 |
| 5.1 | Datenbankgestützte high-tech Sequenzierung von Proteinen | 235 |
| 5.2 | Genomweite Two-Hybrid Analyse in Hefe | 242 |
| 5.2.1 | Das Proteinnetzwerk der Hefe | 242 |
| 5.3 | Proteomarray mit exprimierten Hefe Proteinen - Die Suche nach enzymatischen Aktivitäten | 246 |
| 5.4 | Datenbanken für nonhomology Funktionsvorhersagen | 250 |
| 5.5 | Pathway-Datenbanken | 255 |
| 6 | Phylogenetik | 265 |
| 6.1 | Grundlagen | 265 |
| 6.1.1 | Methoden zur Konstruktion phylogenetischer Trees | 267 |
| 6.1.2 | Gen-trees | 268 |
| 6.2 | Gen-trees versus Spezies-trees | 269 |
| 7 | DNA-Computing - Ein Exot mit Potential | 277 |
| | Index | 282 |