

Obsah

Předmluva.....	7
1 K čemu a o čem je tento text	9
1.1 Modelové problémy.....	11
2 Praktické předpoklady	13
3 Práce se sekvenčními daty	15
3.1 Zápis sekvencí a běžné formáty datových souborů	17
3.2 Mapování, restrikce, translace a klonování in silico.....	21
4 Veřejně dostupné zdroje dat	25
4.1 „Velká trojka“ primárních databází nukleotidových sekvencí – GenBank, EMBL, DDBJ.....	26
4.2 Identifikační čísla, verze a typy datových souborů.....	28
4.3 Jak se data dostanou do databází?	29
4.3.1 Anatomie databázového záznamu	30
4.4 Jednoduché vyhledávání v databázích Velké trojky.....	33
4.5 Specializované databáze včetně genomových	35
4.6 Databáze propojené se sbírkami biologických materiálů	37
5 Posuzování podobnosti sekvencí	38
5.1 Modelový postup stanovení podobnosti	38
5.2 Globální a lokální přiřazení.....	40
5.3 Empirické parametry – substituční matice a cena mezer.....	42
6 Prohledávání databází podle podobnosti se známou sekvencí.....	47
6.1 FASTA jakožto modelový heuristický algoritmus	48
6.2 BLAST.....	52
6.2.1. PSI-BLAST a pozičně specifické substituční matice (PSSM).....	56
6.3 Volba metody a parametrů pro konkrétní situace.....	58
6.4 Prohledávání strukturních databází sekvenčním dotazem	61
7 Hledání známých motivů v proteinových sekvencích.....	63
7.1 Mnohočetné přiřazení a způsoby zápisu sekvenčních motivů	63
7.2 Databáze známých motivů	66
7.3 SMART.....	68
7.4 Predikce buněčné lokalizace proteinů	69
7.5 Vyhledávání dalších motivů v proteinových sekvencích.....	70
8 Konstrukce a interpretace mnohočetného přiřazení.....	73
8.1 Zápis mnohočetného přiřazení	73
8.2 Ruční a automatizované postupy konstrukce mnohočetného přiřazení	77

8.2.1	CLUSTAL jako modelový automatizovaný postup	77
8.2.2	Poloautomatické a ruční metody	80
8.2.3	Konstrukce přiřazení pomocí programů na veřejných serverech.....	84
8.3	Přiřazování trojrozměrných struktur a základy modelování struktury proteinů.....	85
8.4	Interpretace mnohočetného přiřazení	88
8.5	Odvozování vlastních konzervativních motivů a jejich použití v prohledávání databází.....	90
9	Analýza struktury a funkce genů	92
9.1	Identifikace kódujících oblastí a predikce sestřihu.....	92
9.1.1	Jak kvalitní je anotace v databázích?.....	98
9.2	Analýza genové exprese.....	101
10	Studium příbuzenských vztahů biologických sekvencí	105
10.1	Základní pojmy	106
10.2	Metody konstrukce dendrogramů.....	110
10.3	Hodnocení kvality dendrogramu	113
10.4	Praktická konstrukce dendrogramu	115
10.4.1	Příprava výchozích dat	115
10.4.2	Metoda NJ (<i>neighbor-joining</i>).....	117
10.4.3	Další metody konstrukce dendrogramů	122
11	Tři příklady na závěr.....	126
	Slovníček a seznam zkratk	134
	Literatura	139
	Rejstřík	146