

# IV107 Bioinformatika I

## Přednáška 1

Katedra informačních technologií  
Masarykova Univerzita Brno

Jaro 2011

### Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

### Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

### Molekulární biologie v kostce

Centrální dogma

### Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

Struktura DNA

Transkripce a translace

Struktura proteinů

## Úvod do bioinformatiky

**Organizační záležitosti**

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA

Transkripce a translace

Struktura proteinů

- ▶ Dr. Matej Lexa, C506 (lexa@fi.muni.cz)
- ▶ Přednáška Po 08:00-09:50 (B410)
- ▶ Konzultace Čt 13:00-15:00
- ▶ <http://www.fi.muni.cz/~lexa/teaching.html>

## Úvod do bioinformatiky

### Organizační záležitosti

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

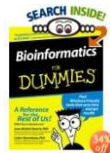
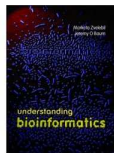
## Příště

Struktura DNA

Transkripce a translace

Struktura proteinů

1. Zvelebil and Baum (2007).  
Understanding bioinformatics, Garland Science,  
Oxford, 772 s. (ISBN: 0-8153-4024-9)
2. Krane and Raymer (2005).  
Fundamental concepts in bioinformatics,  
Benjamin Cummings, London, 320 s. (ISBN  
0-8053-4633-3)
3. Claverie (2005).  
Bioinformatics for dummies, Wiley Publishing,  
Hoboken, 452 s. (ISBN: 0-7645-1696-5)



## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

- ▶ Bioinformatics
- ▶ BMC Bioinformatics
- ▶ J. of Bioinformatics and Computational Biology
- ▶ Briefings in Bioinformatics
- ▶ Genome Informatics
- ▶ Theoretical Biology and Medical Modelling
- ▶ InSilico Biology
- ▶ Biosemiotics

## Úvod do bioinformatiky

### Organizační záležitosti

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA

Transkripce a translace

Struktura proteinů

- ▶ Bakalářská a magisterská úroveň
- ▶ Lze zvolit i v průběhu studia
- ▶ Základní sada předmětů Aplikované informatiky na FI a čtyři předměty na LF a PŘF.
- ▶ Povinnost vypracovat bioinformatickou závěrečnou práci
- ▶ <http://www.fi.muni.cz/~lexa/teaching.html.cz>

▶ [https://is.muni.cz/auth/setkavani/kruh.pl?kruh\\_id=7161](https://is.muni.cz/auth/setkavani/kruh.pl?kruh_id=7161)  
Bioinformatika@FI Muni

## Úvod do bioinformatiky

### Organizační záležitosti

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA

Transkripce a translace

Struktura proteinů

- ▶ IV108 - Bioinformatika II (podzim)
- ▶ IV105/IV106 - Seminář z bioinformatiky P/G (Út 8:00 B411)
- ▶ IV110/IV114 - Projekt z bioinformatiky (podzim)
- ▶ IV116 - Evolutionary Bioinformatics (podzim?)
- ▶ PB051 - Výpočetní metody v bioinformatice a systémové biologii

## Úvod do bioinformatiky

### Organizační záležitosti

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA

Transkripce a translace

Struktura proteinů

- ▶ IV109 - Modelování a simulace
- ▶ IV117/8 - Systémová biologie

## Úvod do bioinformatiky

### Organizační záležitosti

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA

Transkripce a translace

Struktura proteinů



- ▶ Rychlý úvod do molekulární biologie (do poloviny března)
- ▶ Semestrální test (březen/duben)

## Úvod do bioinformatiky

### Organizační záležitosti

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA

Transkripce a translace

Struktura proteinů

- ▶ Hodnotí se
  - ▶ Semestrální test 20 bodů
  - ▶ Zkouška 80 bodů
- ▶ Klasifikační stupnice
  - ▶ A 90 - 100
  - ▶ B 80 - 89
  - ▶ C 70 - 79
  - ▶ D 60 - 69
  - ▶ E 50 - 59
  - ▶ F méně než 50

## Úvod do bioinformatiky

### Organizační záležitosti

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA

Transkripce a translace

Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

Struktura DNA

Transkripce a translace

Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA

Transkripce a translace

Struktura proteinů

## Bioinformatika

Studuje metody shromáždění, přístupování a analýzy rozsáhlých souborů biologických dat, zejména molekulárně – biologických.

### Další disciplíny

- ▶ Výpočetní nebo matematická biologie  
matematické přístupy k reprezentaci a zkoumání biologických procesů, často simulace
- ▶ Lékařská informatika  
práce s medicínskými daty, převážně záznamy pacientů

### Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

### Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

### Molekulární biologie v kostce

Centrální dogma

### Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

# Předmětem zájmu nebo používanými metodami se bioinformatika prolíná s

1. molekulární biologii
2. genomikou a proteomikou
3. genetikou
4. výpočetní biologii
5. matematickou či teoretickou biologii
6. systémovou biologii
7. biomedicínskou informatikou
8. biomedicínským inženýrstvím
9. výpočetní chemií
10. informatikou
11. počítačovou lingvistikou

Převzato z <http://cz.wikipedia.org/wiki/Bioinformatics> 16.2.2008

## Úvod do bioinformatiky

Organizační záležitosti

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA

Transkripce a translace

Struktura proteinů

- ▶ Analýza sekvencí
- ▶ Anotace genomů
- ▶ Evoluční bioinformatika
- ▶ Studium biodiverzity
- ▶ Analýza exprese genů
- ▶ Analýza genové regulace
- ▶ Analýza proteomu
- ▶ Odhad struktury proteinů
- ▶ Srovnávací genomika
- ▶ Modelování biologických systémů
- ▶ Analýza obrazu
- ▶ Studium strukturních interakcí proteinů

## Úvod do bioinformatiky

Organizační záležitosti

Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

Bioinformatická data

Práce bioinformatika

Historie bioinformatiky

Zkoumání lidského genomu

Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA

Transkripce a translace

Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

**Objekty: geny, molekuly, buňky**  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma  
Struktura DNA  
Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

**Objekty: geny, molekuly, buňky**  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů





# Složitost biologických systémů na molekulární úrovni

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

**Člověk:** cca  $10^{14}$  buněk.

**Buňka:**  $3 \times 10^9$  párů nukleotidů DNA (A:T a C:G).

**Nukleotidy:** vytváří sřetěženými kombinacemi cca 20000 genů (a statisíce funkčních míst)

**Geny:** kódují (a aktivitou vytváří) statisíce molekul (proteinů a RNA)

**Buňka:** aktivuje v daném momentu určitou podmnožinu této sady

**Výsledek:** obrovské množství možných stavů buněk ( $2^{20000}$  je velmi podceňující odhad)

**Geny:** evolucí vybrané sady z cca  $4^{1000}$  možných sekvencí DNA (1000 nukl./gen)

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky

### Bioinformatická data

Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma  
Struktura DNA  
Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
**Bioinformatická data**  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

- ▶ Sekvence DNA a RNA
- ▶ Sekvence proteinů
- ▶ Struktura proteinů
- ▶ Údaje o aktivitě genů DNA čip, microarray, RNA-Seq
- ▶ Údaje o expresi proteinů 2-D gely + MS
- ▶ Mapy interakcí mezi proteiny a DNA - Chip-Seq
- ▶ Mapy interakcí mezi proteiny navzájem - Y2H
- ▶ Literatura

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
**Bioinformatická data**  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

AUGACAGUUGACGAGUGCA  
ATAGCAGTGCGCATGCAGT  
MASAQSFYLLMDDHLAVFM

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
**Bioinformatická data**  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

DNA ATAGCAGTGCGCATGCAGT  
RNA AUGACAGUUGACGAGUGCA  
Protein MASAQSFYLLMDDHLAVFM

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
**Bioinformatická data**  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů



# Spřístupnění dat uživatelům – NCBI Genome Viewer

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

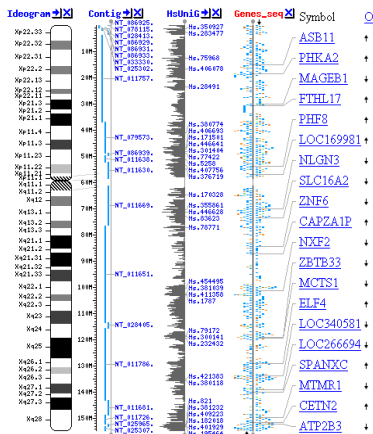
Objekty: geny, molekuly, buňky  
**Bioinformatická data**  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů



Zobrazení informací o genech na chromozomu

- ▶ Grafika je zbytečná. Prvořadá je rychlost a možnost automatizace manipulace s daty
- ▶ BioJava, BioPerl, BioPython, Bioconductor (R) a další knihovny pro většinu jazyků a prostředí
- ▶ servery poskytující syrová data (holý text, obrázky, XML a jiné struktury přes HTTP, SOAP, ODBC)

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
**Bioinformatická data**  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů



## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
**Práce bioinformatika**  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma  
Struktura DNA  
Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
**Práce bioinformatika**  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

výraz	Google (tis. výskytů)	
	2004	2011
<i>et tu brutus</i>	212	195
<i>in vino veritas</i>	162	1130
<i>veni vidi vici</i>	132	2340
<i>in vivo</i> (biolog)	19100	11400
<i>in vitro</i> (biochemik)	12900	18000
<i>in silico</i> (bioinformatik)	349	1790

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
**Práce bioinformatika**  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

- ▶ Umí pracovat s velkými datovými soubory
- ▶ Moudrými triky ovláda výkonné počítače
- ▶ V datech hledá zajímavé subsekvence
- ▶ Srovnává podobné sekvence
- ▶ Předpovídá strukturu a funkci genů a proteinů
- ▶ Studuje vývoj sekvencí a organismů
- ▶ Data a výsledky analýz zobrazuje graficky

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
**Práce bioinformatika**  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

- KLASIK** směs biologie, chemie, fyziky atd.
- MECHANIK** živé buňky jsou stroje, které chceme pochopit a ovládat
- HRA** sekvence jsou definiční soubory hráčů
- SEMIOTIK** život je signalizace a interpretace signálů
- JAZYK** sekvence se skládají z modulů (slov) s určitou funkcí vykazujících gramatické uspořádání

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
**Práce bioinformatika**  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
**Historie bioinformatiky**  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma  
Struktura DNA  
Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika

### **Historie bioinformatiky**

Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

1951	Pauling	struktura proteinů
1952	Turing	chem. základy vývoje
1953	Watson, Crick, Franklin	struktura DNA
1956	Gamow et al.	genetický kód
1959	Chomsky	gramatiky
1962	Shannon a Weaver	informační teorie
1966	Martin-Lof	náhodné řetězce
1966	Neumann	automata
1969	Britten a Davidson	génová regulace

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
**Historie bioinformatiky**  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

# Historie bioinformatiky do sformování disciplíny

- 1967 Fitch and Margoliash: sestrojení prvních fylogenetických stromů z biologické sekvence
- 1970 Needleman and Wunsch: zarovnání dvou sekvencí
- 1974 Chou and Fasman: predikce sekundární struktury proteinů
- 1978 Dayhoff: první sbírka sekvencí proteinů
- 1981 Kabsch and Sander: modelování struktury proteinů
- 1987 Feng and Doolittle: mnohonásobné zarovnání sekvencí
- 1990 Altschul et al.: efektivní hledání lokálních podobností
- 1998 The Journal Comp Appl Biosci se přejmenovává na Bioinformatics

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
**Historie bioinformatiky**  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

- Organizační záležitosti
- Zaměření bioinformatiky

## Bioinformatická data

- Objekty: geny, molekuly, buňky
- Bioinformatická data
- Práce bioinformatika
- Historie bioinformatiky
- Zkoumání lidského genomu**
- Aktuální problémy

## Molekulární biologie v kostce

- Centrální dogma
- Struktura DNA
- Transkripce a translace
- Struktura proteinů

## Úvod do bioinformatiky

- Organizační záležitosti
- Zaměření bioinformatiky

## Bioinformatická data

- Objekty: geny, molekuly, buňky
- Bioinformatická data
- Práce bioinformatika
- Historie bioinformatiky
- Zkoumání lidského genomu**
- Aktuální problémy

## Molekulární biologie v kostce

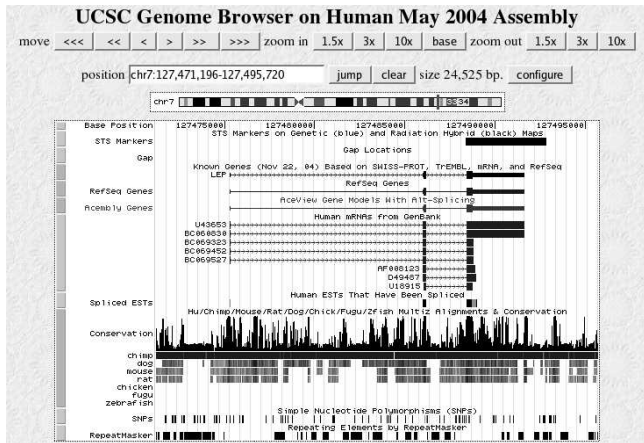
- Centrální dogma

## Příště

- Struktura DNA
- Transkripce a translace
- Struktura proteinů







Flexibilní nástroj určen k interaktivnímu prohlížení genomů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

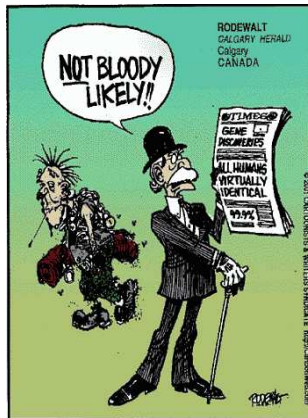
Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

# Homo/Homo

- ▶ rozdíl každých 1000 nukleotidů
- ▶ 90% variace je mezi africkými populacemi
- ▶ na Zemi je tolik lidí a četnost mutací je tak vysoká, že každý ze jmenovaných nukleotidů je v dané generaci mutován několikrát
- ▶ lidský genom obsahuje stovky nepříjemných mutací. Většina je recesivních, projeví se jenom ojediněle, pokud je mají oba rodiče



## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů



- ▶ rozdíl každých 100 nukleotidů
- ▶ transpozon každých 50000 nukleotidů
- ▶ dva chromozomy spojené, jinak podobná struktura

Podle Jim Kent: "The Genes, the Whole Genes, and Nothing But the Genes", BioCon 2003.

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

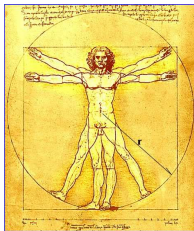
Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
**Zkoumání lidského genomu**  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů



- ▶ 40% nukleotidů byli od dob společného předka změněny
- ▶ Ve funkčních oblastech se změnilo jenom 15% nukleotidů
- ▶ úseky podobnosti mezi genomy člověka a myši jsou kandidáti na biologické funkce

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
**Zkoumání lidského genomu**  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů



## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
**Aktuální problémy**

## Molekulární biologie v kostce

Centrální dogma  
Struktura DNA  
Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
**Aktuální problémy**

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

- ▶ Základní výskum
- ▶ Medicína a jiné aplikace
- ▶ Bezpečnost na molekulární úrovni
- ▶ Komerční data

V současnosti např. nastupuje "osobní genomika"

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
**Aktuální problémy**

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů



# HT-Seq: objem dat z jednoho merení a cena za 1 Mbp

- ▶ Solexa pyrosequencing (Illumina) 18 Gbp \$2
- ▶ 454 (Roche) 0.5 Gbp \$60 (ale delší sekvence)
- ▶ SOLiD (Life Technologies) 24 Gbp \$2
- ▶ Heliscope (Helicos) 28 Gbp \$1
- ▶ Polonator (Danaher Motion) 8 Gbp \$1
- ▶ Zero-mode waveguide sequencing (Pacific Biosciences) 10 Gbp? \$10?
- ▶ Nanoball sequencing (CompleteGenomics) 70 Gbp \$1
- ▶ FRET sequencing (Visigen) ?
- ▶ Nanopore sequencing (Oxford Nanopore) ?

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
**Aktuální problémy**

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

```
>P11633 NONHISTONE CHROMOSOMAL PROTEIN 6B.  
Score = 54.8 bits (155), Expect = 1e-10 Identities =  
19/43 (46%), Positives = 24/43 (62%)  
Query: 2 TKKFKDPNRPPSAFFLFCSEYRKIKGEHPGLSIGDVAKKLGEM 52  
:      T : KDPNR  SA:  F :E R I  E:P :: G V : LGE  
Sbjct: 5 TTRKKDPNRGLSAYMFFANENRDIRSENPDVTFGQVGRILGER 55
```

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu

### Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

1. Mam z toho velkou radost.
2. Mam toho kocoura dost.

```
Mamztohovelk ouradost.  
::: :::: : :::::::::::  
Mam toho kocouradost.
```

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
**Aktuální problémy**

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

**Centrální dogma**  
Struktura DNA  
Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

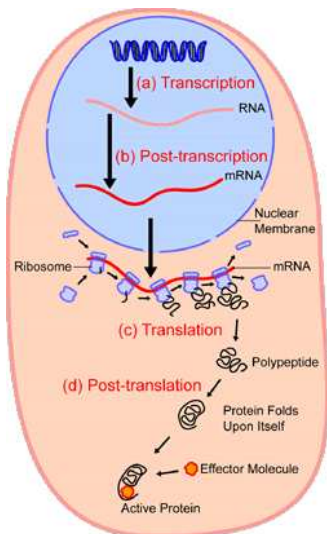
## Molekulární biologie v kostce

**Centrální dogma**

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

# Informace v DNA určuje existenci proteinů v buňce



## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

- ▶ Struktura DNA
- ▶ Struktura proteinů
- ▶ Přenos genetické informace

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma  
**Struktura DNA**  
Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

### **Struktura DNA**

Transkripce a translace  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma  
Struktura DNA  
**Transkripce a translace**  
Struktura proteinů

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
**Transkripce a translace**  
Struktura proteinů



## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma  
Struktura DNA  
Transkripce a translace  
**Struktura proteinů**

## Úvod do bioinformatiky

Organizační záležitosti  
Zaměření bioinformatiky

## Bioinformatická data

Objekty: geny, molekuly, buňky  
Bioinformatická data  
Práce bioinformatika  
Historie bioinformatiky  
Zkoumání lidského genomu  
Aktuální problémy

## Molekulární biologie v kostce

Centrální dogma

## Příště

Struktura DNA  
Transkripce a translace  
**Struktura proteinů**