

PB051 Výpočetní metody v bioinformatice a systémové biologii

Týden 1

Katedra informačních technologií
Masarykova Univerzita Brno

Jaro 2015

Tento projekt je spolufinancován Evropským sociálním fondem a státním rozpočtem České republiky.



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

- ▶ Dr. Matej Lexa, C506 (lexa@fi.muni.cz)
- ▶ Kurz: Út 12:00-13:50 (B117)
- ▶ Konzultace: Čt 13:00-15:00 (C506)
- ▶ <http://www.fi.muni.cz/~lexa/teaching.html>

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

- ▶ Dr. David Šafránek, A408 (xsafran1@fi.muni.cz)

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

- ▶ Hodnotí se
 - ▶ Úkoly 4 x 5 bodů
 - ▶ Semestrální úkol 30 bodů
 - ▶ Zkouška 50 bodů
- ▶ Klasifikační stupnice
 - ▶ A 91 - 100
 - ▶ B 81 - 90
 - ▶ C 71 - 80
 - ▶ D 61 - 70
 - ▶ E 51 - 60
 - ▶ F méně než 51

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

- ▶ Genomové anotace
- ▶ Výpočty nad sekvencemi, konsenzus, repetice, mapování
- ▶ Aplikace Markovovských modelů v bioinformatice
- ▶ Statická analýza sítí
- ▶ Dynamická analýza sítí

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

- ▶ Prohlížeče genomů (UCSC, Ensembl, Argo, GBrowse)
- ▶ Pokročilé funkce UCSC Genome Browser a Table Browser
- ▶ Programovatelný přístup ke genomu přes rozhraní Biomart z prostředí R/Bioconductor
- ▶ Bioinformatika genové regulace (TRANSFAC, MatInspector))

Výpočty nad sekvencemi, konsenzus, repetice, mapování

Informace o kurzu

[Struktura genomu](#)

[Dynamika genomu](#)

[Genomové data prohlížeče](#)

- ▶ Mnohočetné zarovnání a konsenzuální sekvence
- ▶ Diagra "dot plot" a tandemové repetice
- ▶ Mapování sekvencí (BLAST, Bowtie-2)

Aplikace Markovovských modelů v bioinformatice

Informace o kurzu

[Struktura genomu](#)

[Dynamika genomu](#)

[Genomové data prohlížeče](#)

- ▶ Markovovy řetězce
- ▶ Markovovy řetězce proměnného řádu
- ▶ Skryté Markovovské modely (HMM)
- ▶ HMM profily
- ▶ HMM pro identifikaci genů

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

- ▶ metody a nástroje statické analýzy a integrace dat
 - ▶ integrace dat
 - ▶ rekonstrukce sítě genových interakcí z experimentálních dat
 - ▶ analýza interakční sítě jako obecného grafu

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

- ▶ statická analýza sítí a integrace dat
 - ▶ nástroje: Cytoscape s několika moduly
- ▶ rekonstrukce genových regulačních sítí z microarray dat
 - ▶ nástroje: GeneNetworks, GinSim
- ▶ dynamická analýza pravděpodobnostních modelů genových sítí
 - ▶ nástroje: Dizzy
- ▶ metabolické sítě a jejich analýza
 - ▶ nástroje: KEGG, metacyc, COPASI

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče



Informace o kurzu

Struktura genomu

Dynamika genomu

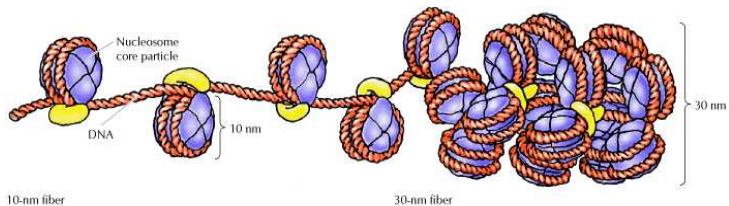
Genomové data prohlížeče

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

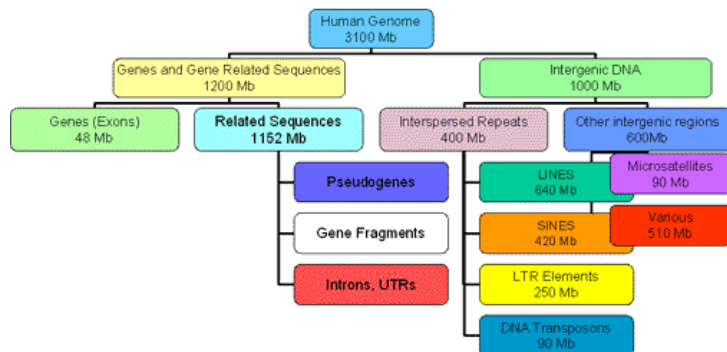


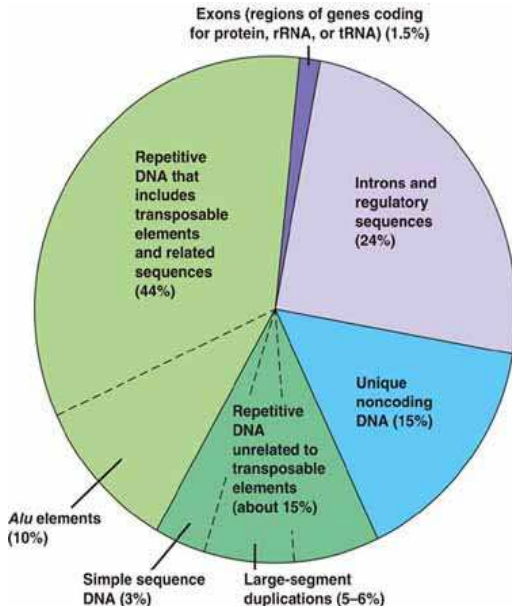
Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče





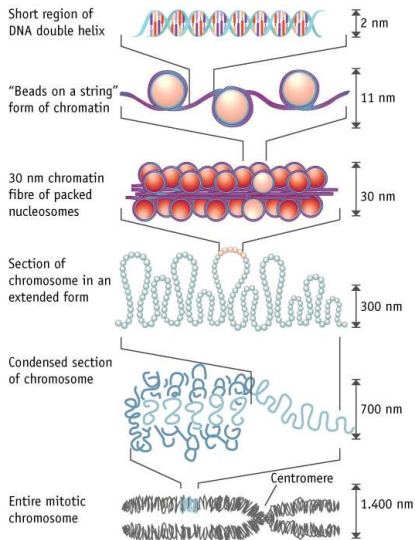
Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

Genome compacting



images/chromosomes

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

Eukaryotic transcription

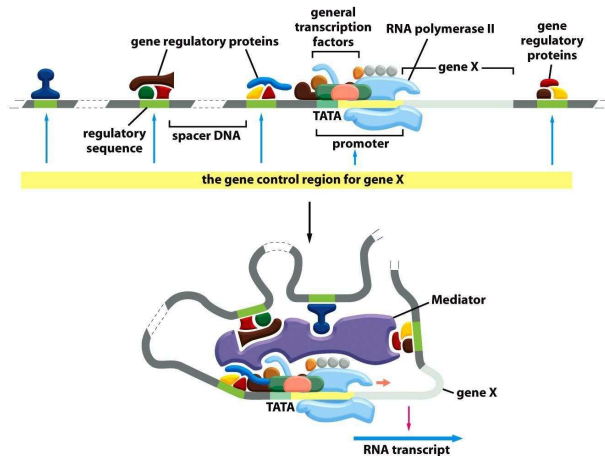


Figure 7-44 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

Selected organizational characteristics of genomes

- ▶ by topology
 - ▶ modular
 - ▶ hierarchical
 - ▶ discrete, but elements sometimes fuzzy
- ▶ by end-product
 - ▶ DNA (structure)
 - ▶ RNA (tRNA, ncRNA, rRNA, siRNA...)
 - ▶ protein
 - ▶ complexes and particles (membrane sensor, flagellum, viral particle)

- ▶ Geny
 - ▶ proteiny (kódující, exon, intron)
 - ▶ RNA
- ▶ Regulační sekvence
 - ▶ promotory
 - ▶ enhancery
 - ▶ jiné
- ▶ Repetitivní sekvence
 - ▶ mikrosatelity (STR)
 - ▶ minisatelity (VNTR)
 - ▶ satelity
 - ▶ DNA transpozony, helitrony
 - ▶ retrotranspozony (LINE, SINE, LTR)
- ▶ Cizí sekvence
 - ▶ viry
 - ▶ endo(retro)viry
- ▶ Oblasti (ne)podobnosti (homology)
 - ▶ SNP
 - ▶ delší strukturní variace
 - ▶ Genomické ostrovy

[Informace o kurzu](#)

[Struktura genomu](#)

[Dynamika genomu](#)

[Genomové data prohlížeče](#)

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

Genome changes (in sequence or number)

- ▶ topology unchanged
 - ▶ SNPs (point mutations)
 - ▶ tandem repeat expansion/contraction
- ▶ 1-point translocation
 - ▶ chromosome breakage
 - ▶ chromosome fusion
- ▶ 2-point translocation
 - ▶ deletions, conversions and exchanges (recombination effects)
 - ▶ cut-paste (DNA transposon)
 - ▶ copy-paste (retrotransposons)
 - ▶ rolling-circle (helitrons)
- ▶ 2-point translocation
 - ▶ DNA methylation
 - ▶ Histone methylation/acetylation

[Informace o kurzu](#)

[Struktura genomu](#)

[Dynamika genomu](#)

[Genomové data prohlížeče](#)

- ▶ at DNA level
 - ▶ exonization
 - ▶ exon shuffling
 - ▶ gene migration
 - ▶ genome duplication
- ▶ at RNA level
 - ▶ alternative splicing
 - ▶ transcriptional fusion
- ▶ at protein level
 - ▶ translational fusion

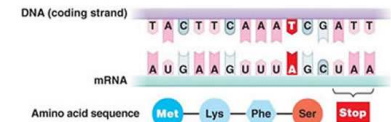
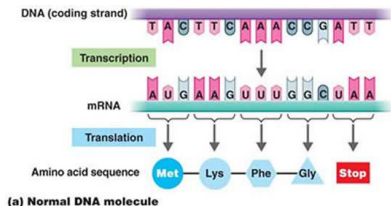
Mutations

Informace o kurzu

Struktura genomu

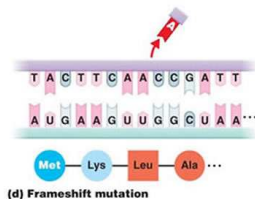
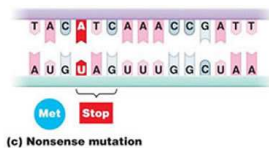
Dynamika genomu

Genomové data prohlížeče



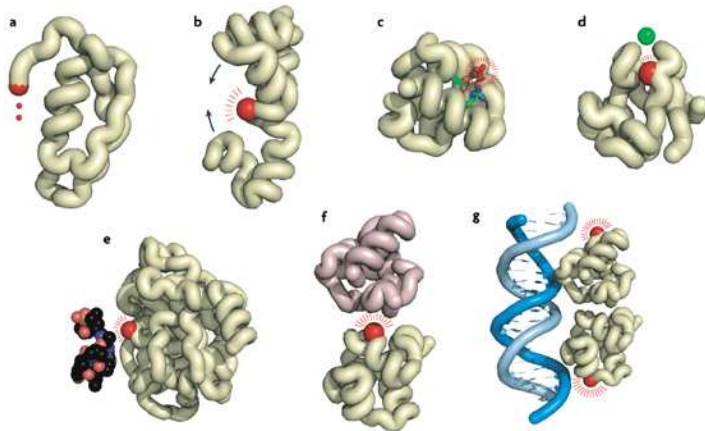
(b) Missense mutation

Copyright © 2010 Pearson Education, Inc.



(d) Frameshift mutation

Mutations



Tandem repeat expansion or contraction

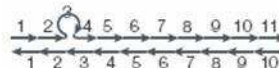
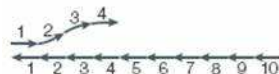
Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

Increase in repeat length



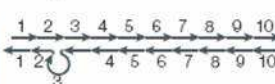
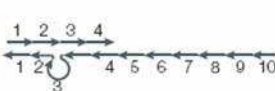
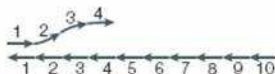
Initiation

Dissociation

Rehybridization
and
misalignment

The new strand
is a different
length to the
template

Decrease in repeat length



Chromosome breakage and repair

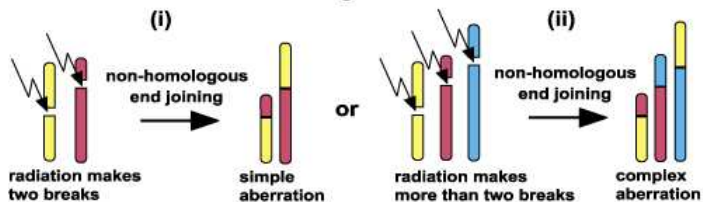
Informace o kurzu

Struktura genomu

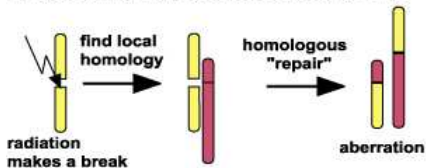
Dynamika genomu

Genomové data prohlížeče

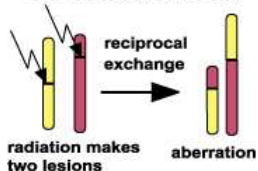
A. Breakage-and-reunion



B. Recombinational misrepair (1-hit)



C. Exchange theory



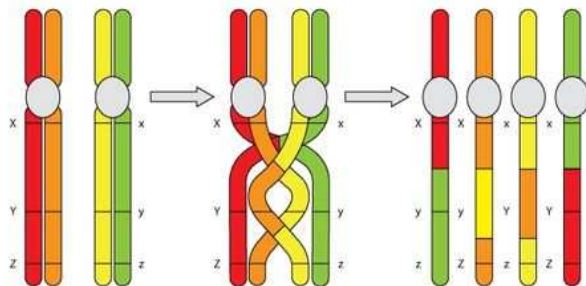
Chiasma and crossing-over during meiosis

Informace o kurzu

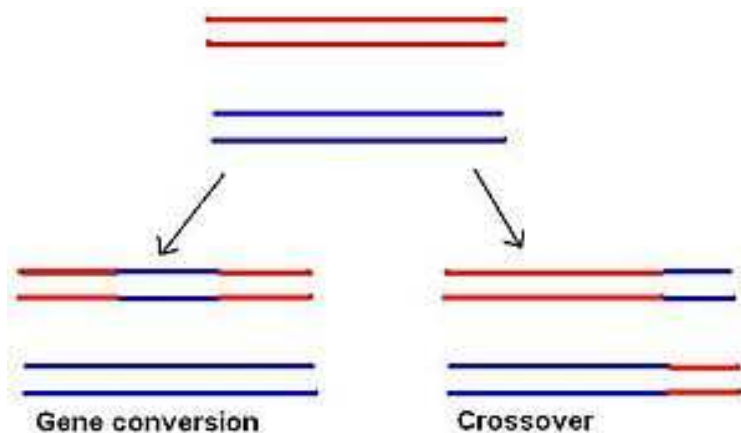
Struktura genomu

Dynamika genomu

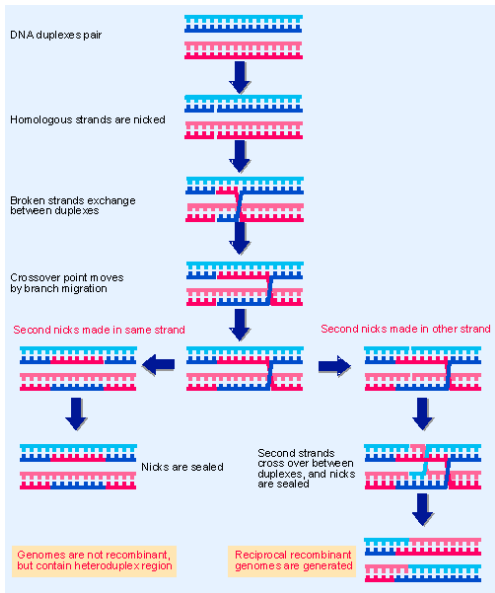
Genomové data prohlížeče



Gene cross-over or conversion



Gene cross-over or conversion



Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

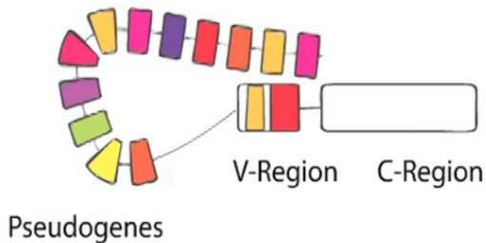
Gene conversion in immunoglobulin

[Informace o kurzu](#)

[Struktura genomu](#)

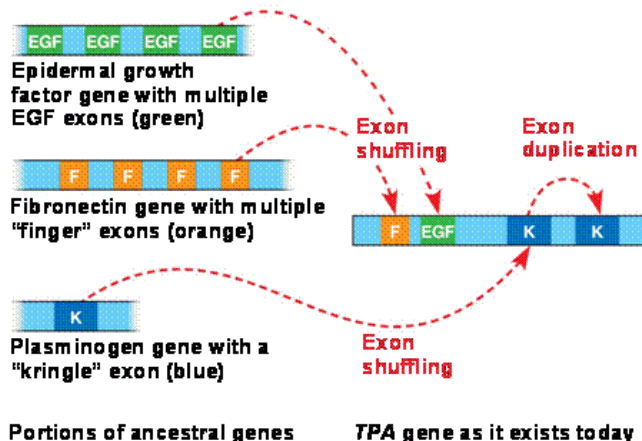
[Dynamika genomu](#)

[Genomové data prohlížeče](#)



Exon shuffling

Fig. 21-14



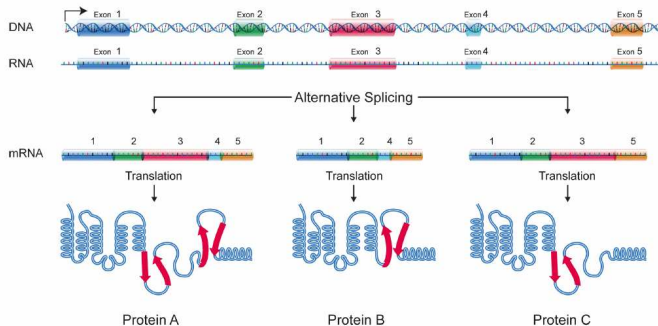
Alternative splicing

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

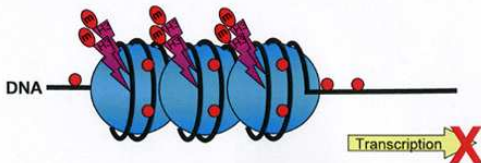
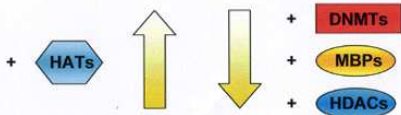
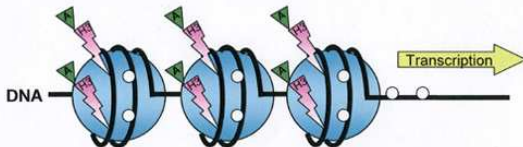


Epigenetic regulation of gene activity

Medscape®

www.medscape.com

A. Transcriptionally active chromatin



B. Transcriptionally inactive chromatin

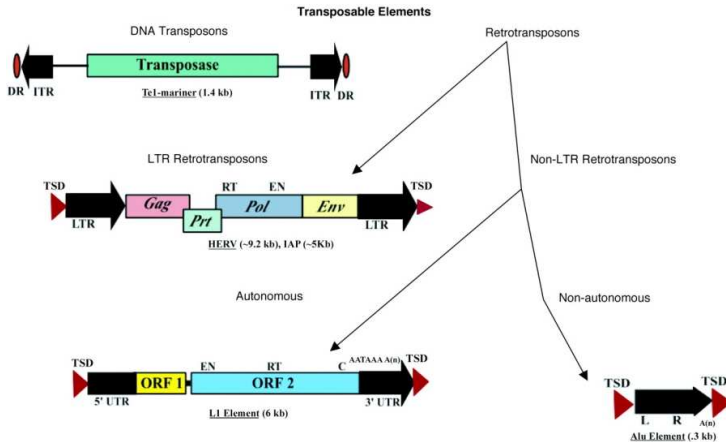
Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

Transposons



Informace o kurzu

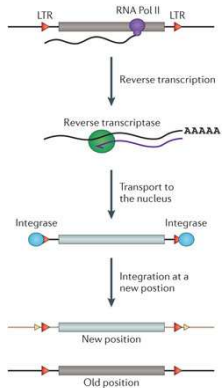
Struktura genomu

Dynamika genomu

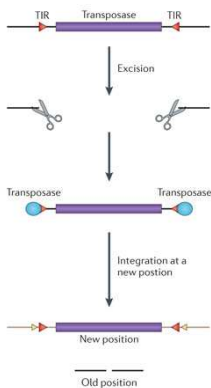
Genomové data prohlížeče

Transposons

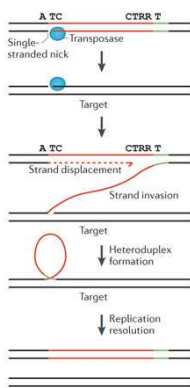
Class I element



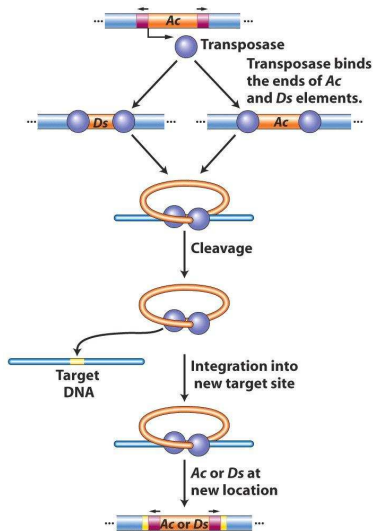
Class II element



Helitron



Transposase is the key protein in DNA transposons



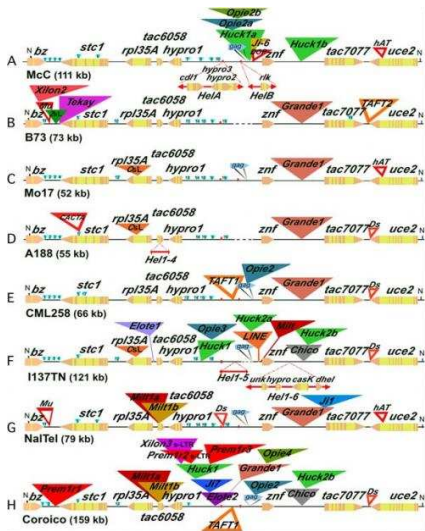
Transposons shape genomes of maize varieties

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče



Transposons shape phenotypes of maize kernels

PB051 Výpočetní metody v
bioinformatice a systémové
biologii - Týden 1

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče



Transposons shape phenotypes of grape varieties

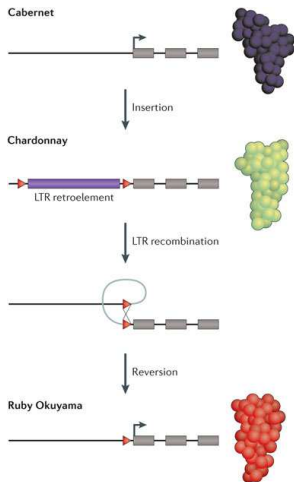
PB051 Výpočetní metody v
bioinformatice a systémové
biologii - Týden 1

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče



[Informace o kurzu](#)

[Struktura genomu](#)

[Dynamika genomu](#)

[Genomové data prohlížeče](#)

- ▶ organizace (konfokální a elektronová mikroskopie)
- ▶ sekvenace
- ▶ mapování metylace, nukleozomů
- ▶ měření transkripce (RNA-Seq, DNA čipy)
- ▶ identifikace regulačních sekvencí (Chip-Seq)
- ▶ funkce genů - podrobný výzkum

[Informace o kurzu](#)

[Struktura genomu](#)

[Dynamika genomu](#)

[Genomové data prohlížeče](#)

- ▶ predikce genů (např. GeneMark)
- ▶ homologie (zjišťování podobnosti sekvencí) (BLAT, MUMMER, BLAST)
- ▶ identifikace opakování (např. RepeatMasker, LTR Finder)

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

Ensembl Genome Browser

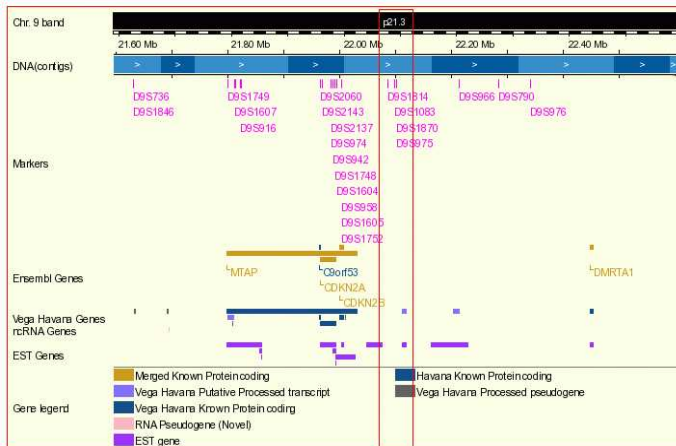
PB051 Výpočetní metody v
bioinformatice a systémové
biologii - Týden 1

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

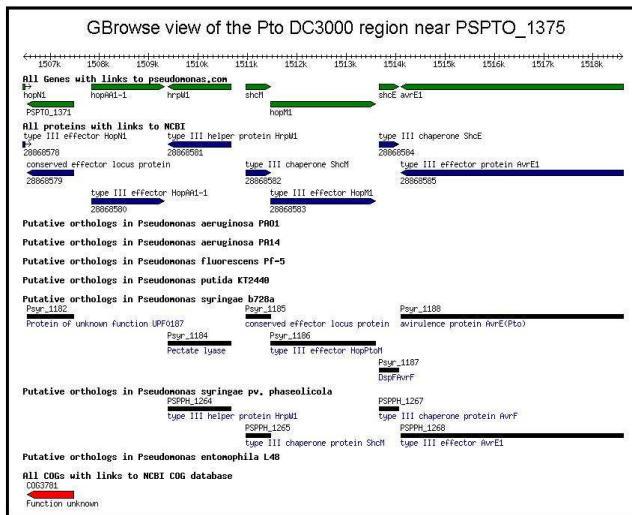


Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče



DecodeMe Browser

PB051 Výpočetní metody v
bioinformatice a systémové
biologii - Týden 1

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče



Golden Helix Genome Browser

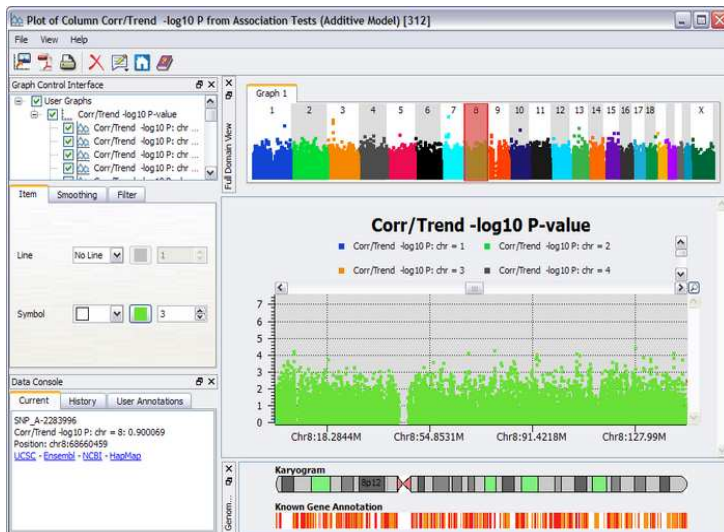
PB051 Výpočetní metody v
bioinformatice a systémové
biologii - Týden 1

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

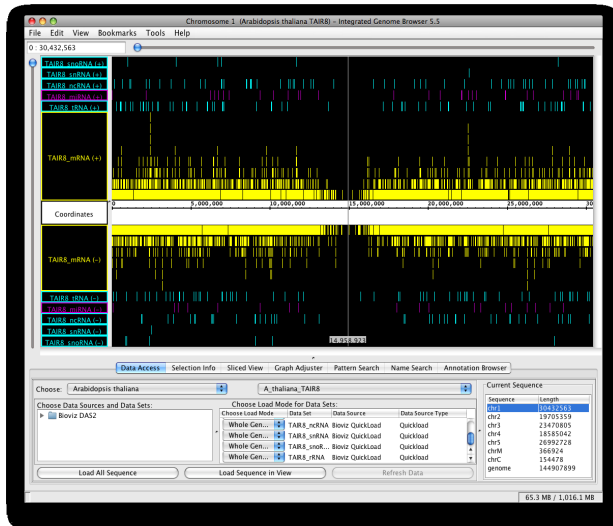


Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

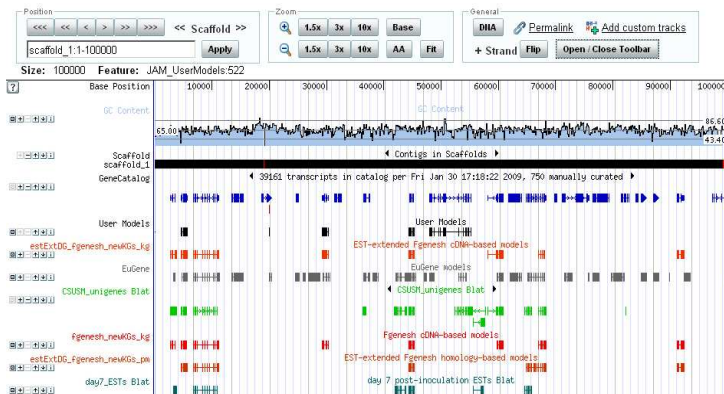


Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče



RIKEN Genome Browser

PB051 Výpočetní metody v
bioinformatické a systémové
biologii - Týden 1

The screenshot displays the RIKEN Genome Browser interface for the *Mus musculus* gene *Irf1*. The interface is divided into several numbered sections:

- 1**: Database list on the left, showing various annotations like Mouse mRBase version, Mouse FANTOM3, Mouse Fantom3 Mega Genes, Mouse CAGE, Mouse Mutant Resources, Mouse Ensembl Transcripts, Mouse Ensembl Vega Genes, Mouse GTOP, Mouse RefSeq Peptide, Mouse RefSeq DNA, Mouse UniProtKB SwissProt, Mouse UniProtKB TrEMBL, Mouse Riken Transcription, Mouse dbSNP (NCBI/Ensembl), Mouse Ensembl Gene 40, Mouse Ensembl Vega Gene, Mouse Entrez Gene NCBI, Mouse MGI Gene, and Mouse RefSeq (NCBI/Ensembl).
- 2**: Expert's set / User's set on the left, showing Genome View (Mouse, Human Homology, C elegans Homology, Other Homology, All, dbSNP) and Medline (PosMed (Positional Medline)), Transcriptome (FANTOM, CAGE).
- 3**: Search and registration options at the top, including "Go to Search page" and "Register current interval".
- 4**: Gene structure diagram for *Mus musculus : 1*, showing exons and introns with coordinates 82,100,466 bp, 81,637 bp, and 82,182,103 bp.
- 5**: Gene information and tracks, including:
 - Marker/Symbol/irs1/Uniprot/SWISSPROT/IRSI MOUSE/RefSeq_peptide/NP_034700.2/RefSeq_dnaNM_010570.2/Uniprot/SPTREMBL/GS43V3_MOUSE/Entrez
 - Mouse FANTOM3 tracks for cDNA(+) and cDNA(-)
 - Mouse CAGE (all tissues) Expression (TAG=249, TPM=6.62) Histogram(+/-)
 - Mouse Ensembl Transcript 43.36a tracks for transcript(+) and transcript(-)
 - Mouse Ensembl Vega Gene Transcript 43.36a tracks for transcript(+) and transcript(-)
 - Mouse GTOP tracks for gene(+) and gene(-)
 - Mouse RefSeq Peptide tracks for protein(+) and protein(-)
 - Mouse RefSeq DNA tracks for dna(+) and dna(-)
 - Mouse UniProtKB SwissProt tracks for protein(+/-)
- 6**: Navigation and utility icons at the bottom, including Config, Text, Print, Send mail, Select strand, Grid, and zoom controls.

Informace o kurzu

Struktura genomu

Dynamika genomu

Genomové data prohlížeče

