

Určovanie genetických predispozícií pomocou regulárnych výrazov

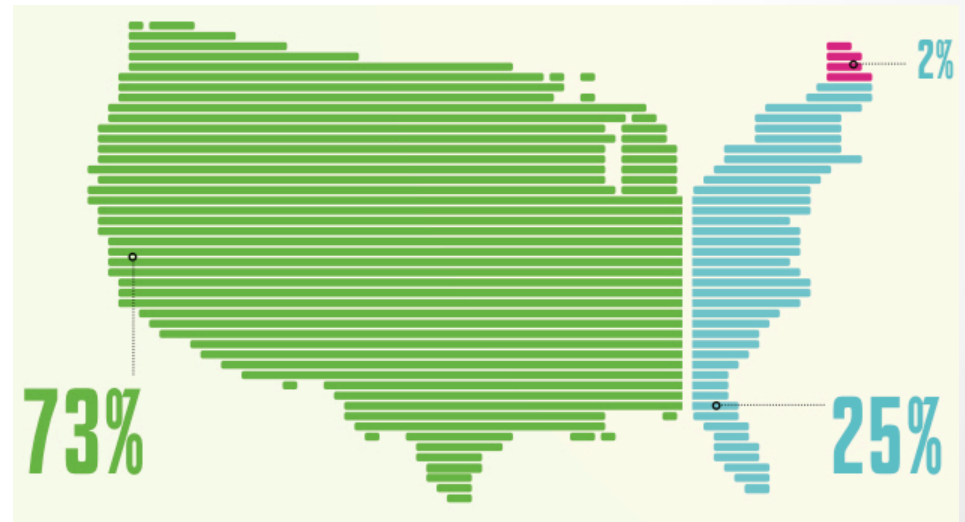
Diplomová práca

Autor: Bc. Jakub Kanitra

Vedúca práce: Mgr. Zuzana Ševčíková

Prečo genetika?

- 6000 genetických porúch
- 20% umrtí novorodencov
- 30% hospitalizovaných

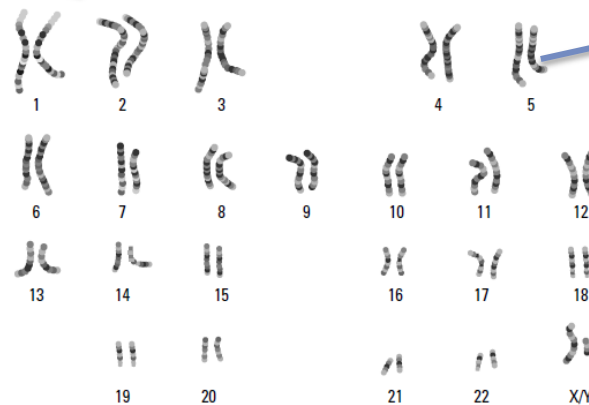
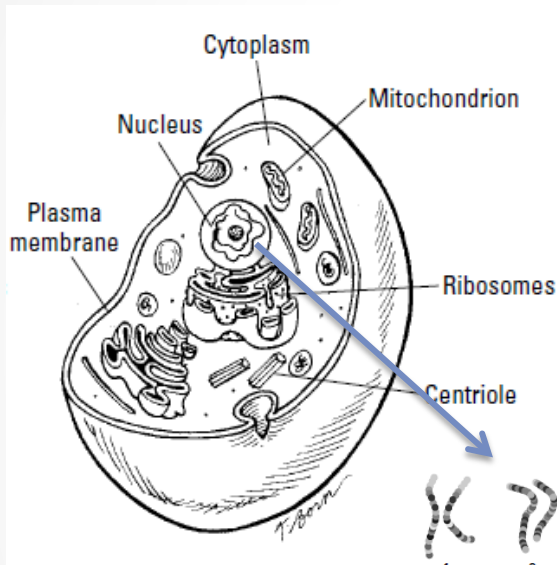


Výstup

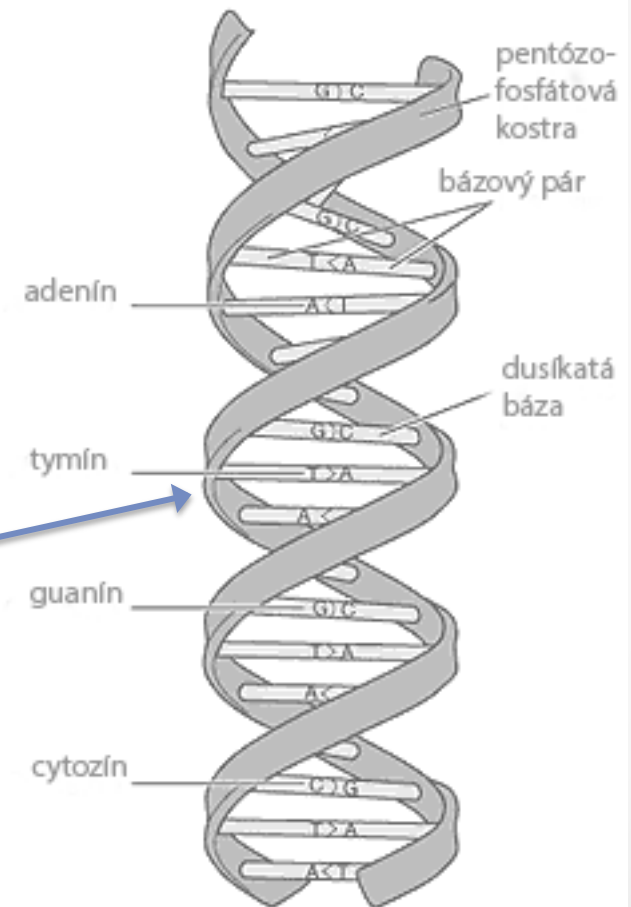
Sample #33456:

#	Gene	Result
1.	Breast cancer	Negative
2.	Huntington disease	Positive
3.	Hemophilia	Negative
4.	Neurofibromatosis	Negative

Genetické minimum

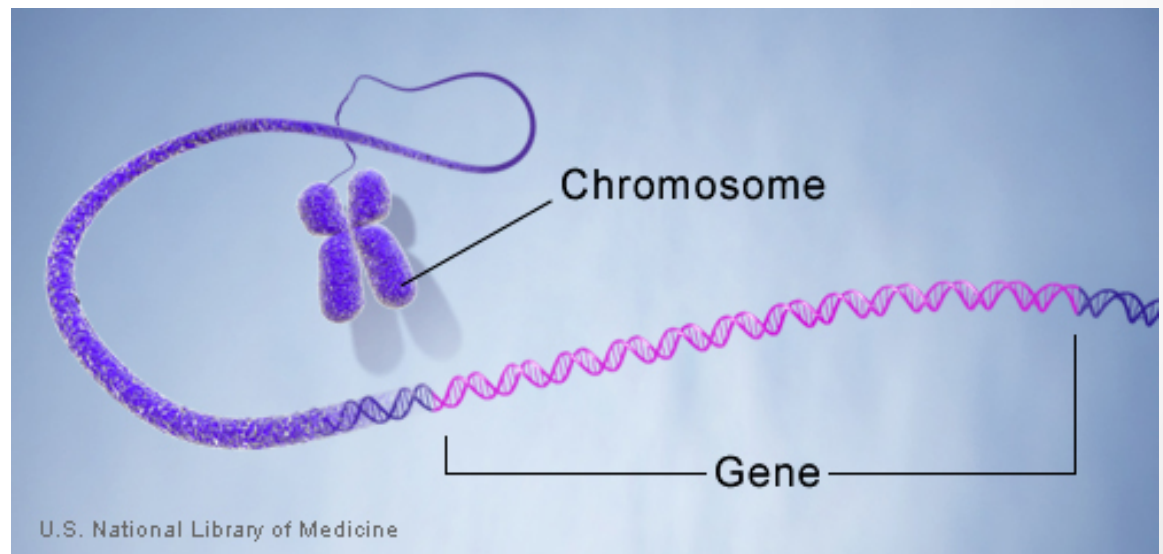


Normal Karyotype



Gén, mutácia

- Gén – vlastnosť jedinca (500bp – 2,3Mbp)
- Mutácia – permanentná zmena DNA



HUGO project

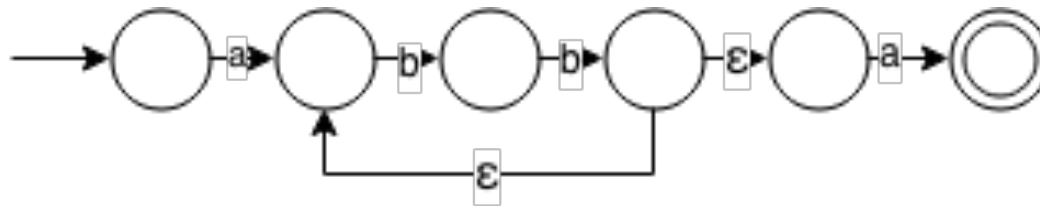
- 1993 – 2003, 3 miliardy \$
- Zdokumentovanie 21 000 génov

Gene: **BRCA2** ENSG00000139618

Description	breast cancer 2, early onset [Source:HGNC Symbol;Acc:HGNC:1101]
Synonyms	BRCC2, FACD, FAD, FAD1, FANCD, FANCD1
Location	Chromosome 13: 32,315,474-32,400,266 forward strand.
INSDC coordinates	chromosome:GRCh38:CM000675.2:32315474:32400266:1
Transcripts	This gene has 7 transcripts (splice variants) Show transcript table

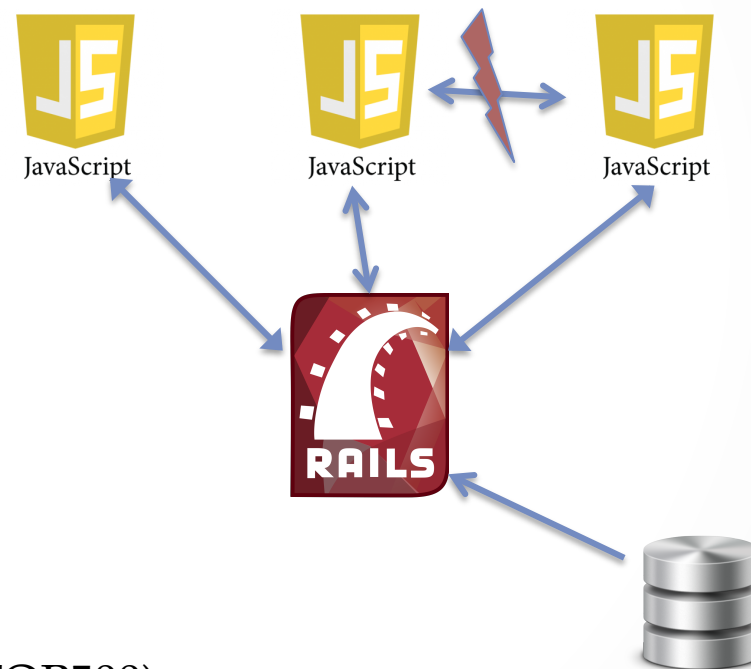
Regulárne výrazy

- Nie je čo riešiť...
 - **Konečný akceptor**
 - Thompsonov konštrukčný algoritmus



Distributívny systém

- Client – server
- Rozhrania



Einstein@home - 470 TFLOPS (TOP500)

Ensembl genome
browser

Plán práce

- Data mining z genome browser-a/-ov
- Komprimačný algoritmus
- Thompsonov konštrukčný algoritmus
- Definovanie a implementácia rozhraní distributívneho systému

Ďakujem za pozornosť

...

Priestor na otázky a diskusiu tu

A

@JakubKanitra