

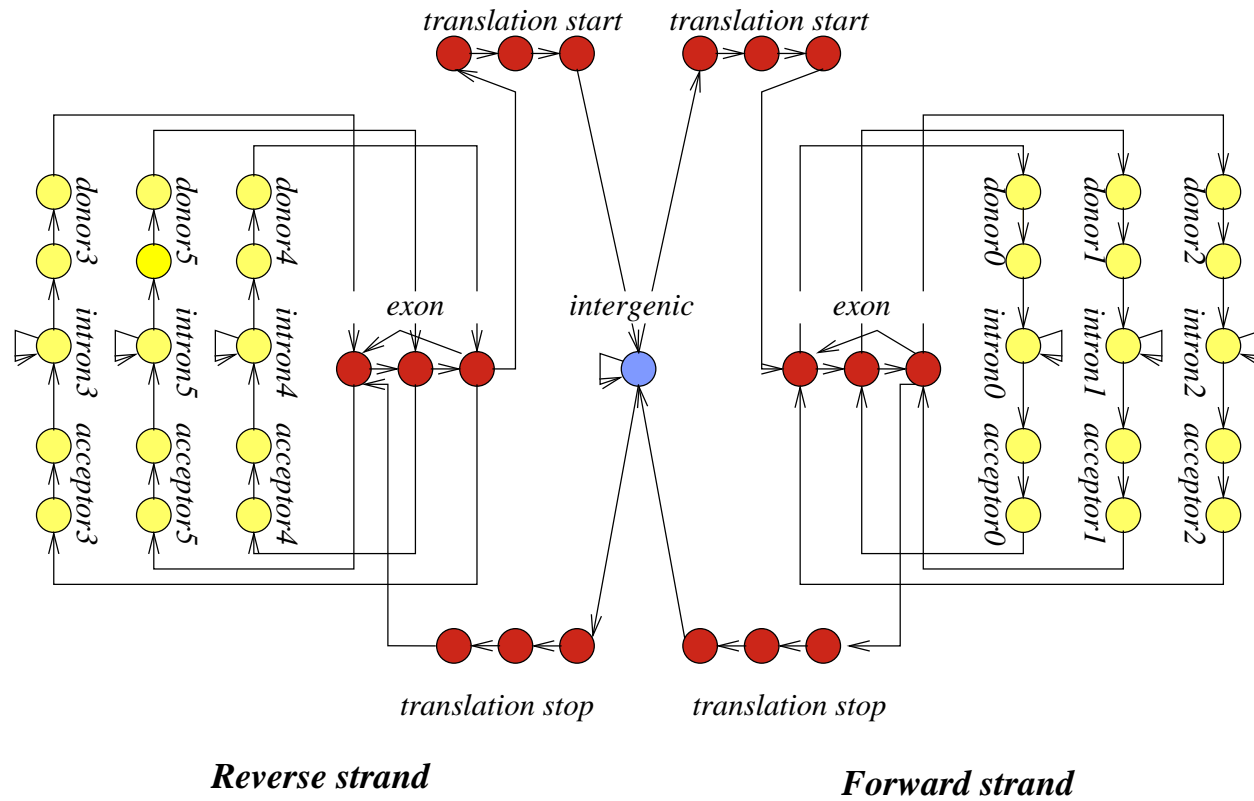
Hľadanie génov (cvičenie)

Broňa Brejová

15.11.2018

Hľadanie prokaryotických génov

HMM pre eukaryotické gény:



Prokaryotické gény jednoduchšie: zväčša nemajú intróny

Hľadanie prokaryotických génov

ORF: open reading frame, otvorený čítací rámec

- hľadanie ORFov je ľahké
- problémy:
 - ako nájsť začiatok,
 - ako rozlíšiť pseudogény a náhodné ORF-y

Anotácia prokaryotických génov nie je triviálny problém

- E.coli sekvenovaná a anotovaná 1997
- Porovnanie s verziou 2005 (4464 génov) [Riley et al NAR 2005]
(oprava sekvenovacích chýb aj chýb v anotácii)
682 zmien v štart kodóne
31 génov zrušených
48 nových génov

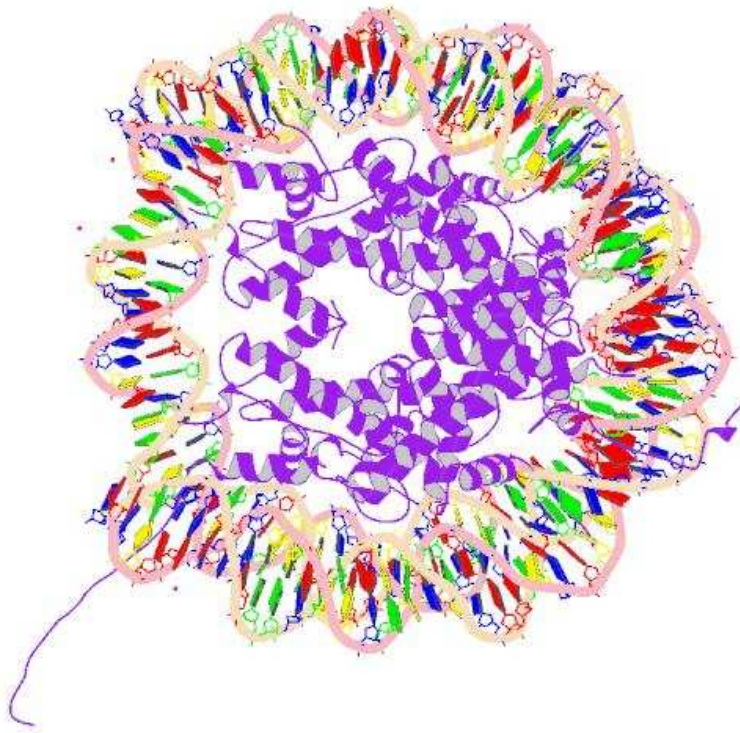
Hľadanie génov

Ideálne kombinácia výpočtových modelov a experimentálnej informácie

- Sekvenovanie RNA-seq; RT-PCR
- Metódy na detekciu proteínov
- Komparatívna genomika (celogenómové zarovnanie, zarovnanie proteínov z príbuzného organizmu)
- Stav chromatinu, histónové modifikácie

Históny a nukleozómy

- DNA v chromozómoch ovinutá okolo nukleozómov pozostávajúcich z histónov H2A, H2B, H3, H4
- 146 báz ovinutých okolo nukleozómu, cca 50 báz medzi nukleozómami

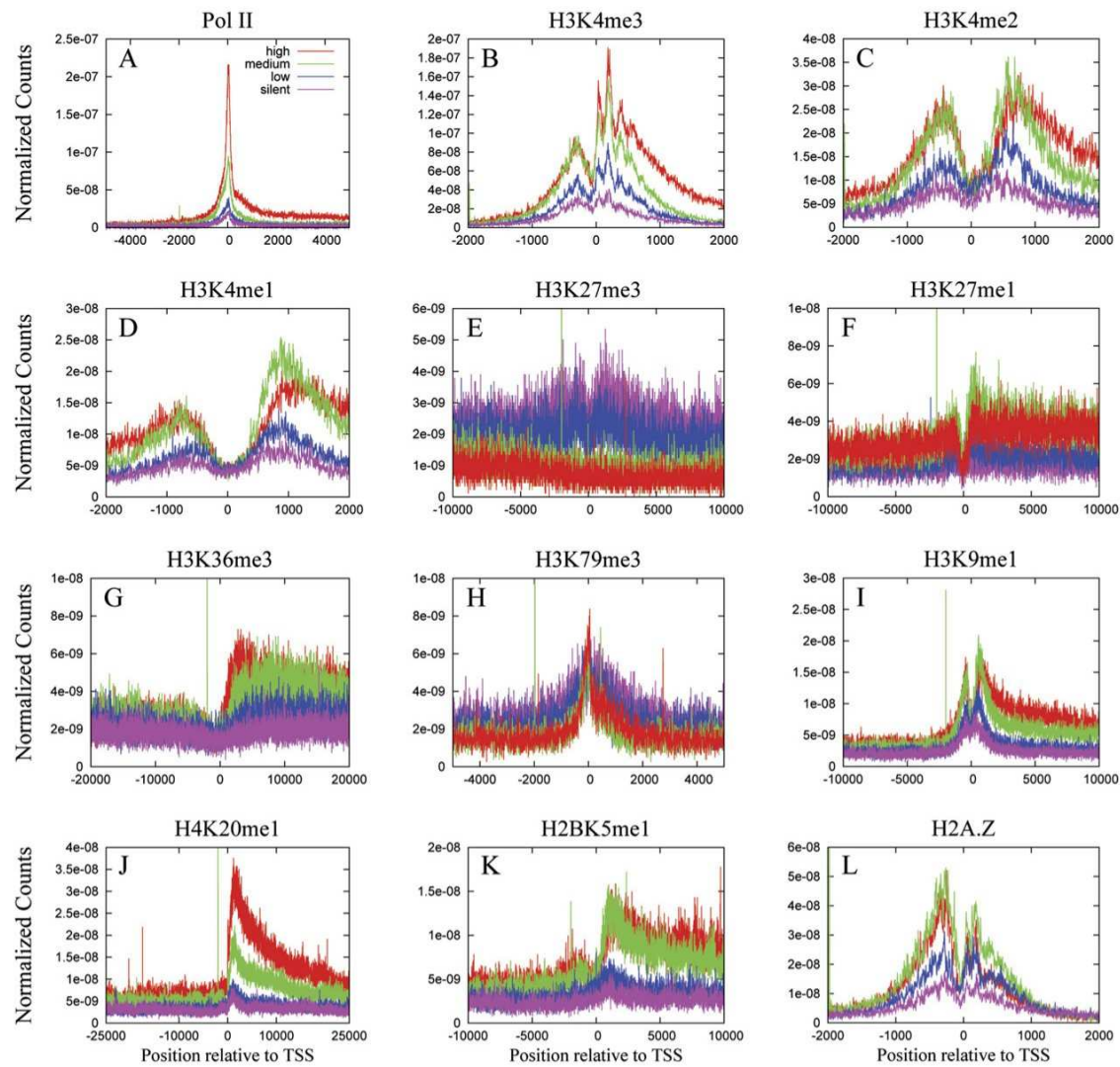


Histónové modifikácie

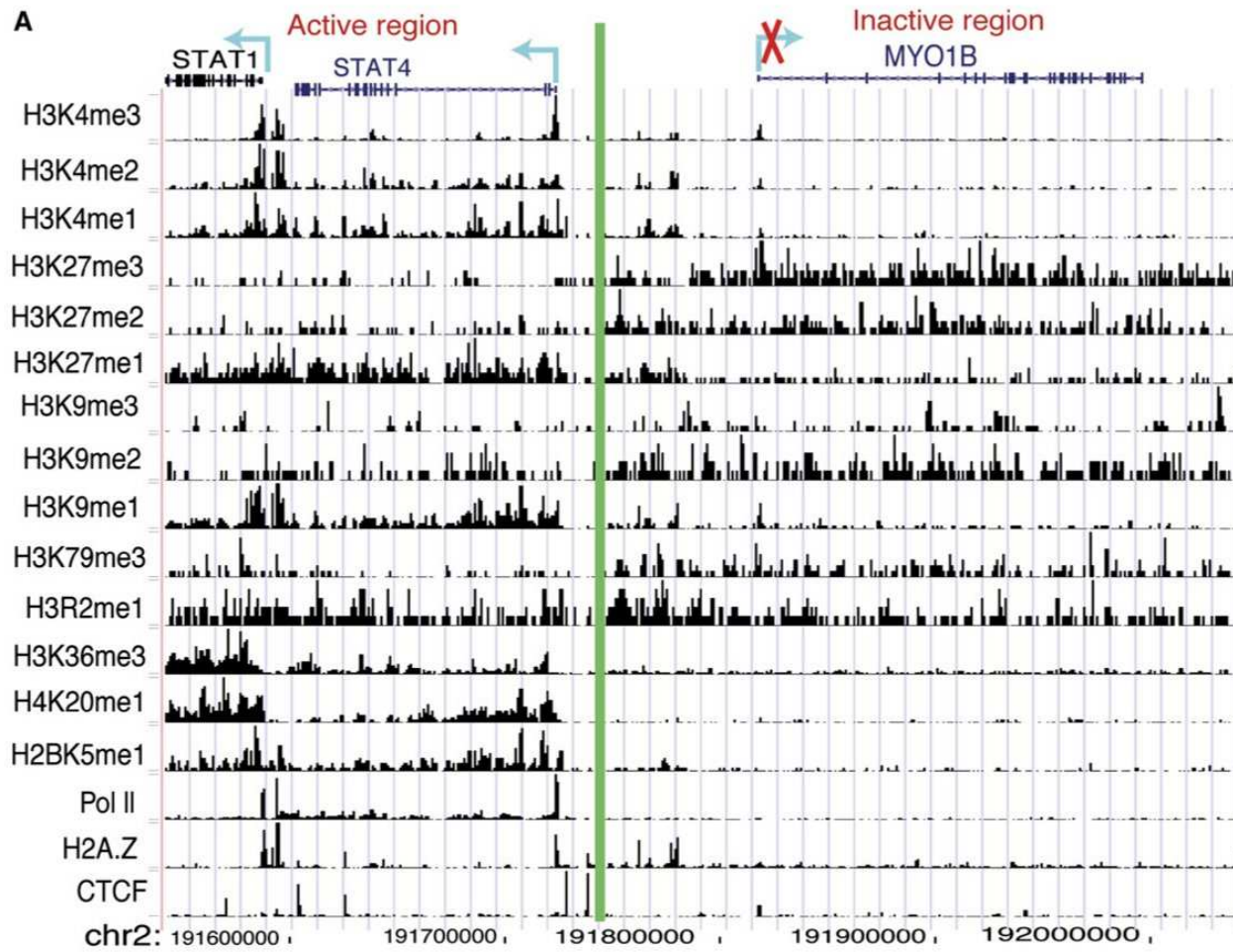
- Posttranslačné modifikácie, napr. metylácia
- Označenie napr. H3K4me1 je (mono-)metylácia štvrtej amino kyseliny (lyzínu) v proteíne H3

Zisťovanie v genóme

- Enzýmom nasekáme DNA medzi nukleozómami
- Nukleozómy s danou modifikáciou extrahujeme pomocou protilátky (chromatin immunoprecipitation, ChIP)
- Extrahovanú DNA identifikujeme pomocou microarray alebo sekvenovaním a mapovaním na genóm (ChIP-chip alebo ChIP-seq)



Priemerné profily okolo začiatku transkripcie, Barski et al 2007



Konkrétne gény, Barski et al, The Cell, 2007