

Komparatívna genomika (cvičenie pre biológov)

Broňa Brejová

14.11.2019

Objavenie génu HAR1 pomocou komparatívnej genomiky

Hľadáme úseky genómu, ktoré sa:

- dlho vyvíjali pomaly (purifikačná selekcia)
- v človeku sa vyvíjajú prekvapivo pomaly (pozitívna selekcia)

Postup: [Pollard et al. (2006) Nature]

- Všetky regióny dĺžky ≥ 100 s $> 96\%$ podobnosťou medzi šimpanzom a myšou/potkanom (35,000)
- Porovnali s ostatnými cicavcami, zistili, ktoré majú veľa mutácií v človeku, ale málo inde (pravdepodobnostný model)
- 49 štatisticky významných regiónov, 96% v nekódujúcich oblastiach
- Štatisticky najvýznamnejší HAR1 (Human Accelerated Region)

Human Accelerated Regions: HAR1

Oblasť dĺžky 118 báz

18 zmien medzi človekom a šimpanzom, 2 zmeny medzi šimpanzom a sliepkou

```
Clovek CTGAAATGATGGGCGTAGACGCACGTCAGCGGCGGAAATGGTTTCTATCA
Simpanz CTGAAAT TATAGGTGTAGACACATGTCAGCAGTGGAAATAGTTTCTATCA
Gorila CTGAAAT TATAGGTGTAGACACATGTCAGCAGTGGAAATAGTTTCTATCA
Rezus CTGAAAT TATAGGTGTAGACACATGTCAGCAGTGGAAATAGTTTCTATCA
Mys CTGAAAT TATAGGTGTAGACACATGTCAGCAGTGGAAATGGTTTCTATCA
Krava CTGAAAT TATAGGTGTAGACACATGTCAGCAGTGGAAACCGTTTCTATCA
Pes CTGAAAT TATAGGTGTAGACACATGTCAGCGGTGCAAAACAGTTTCTATCA
Sliepka CTGAAAT TATAGGTGTAGACACATGTCAGCAGTAGAAACAGTTTCTATCA
```

- Prekrývajúce sa RNA gény HAR1R a HAR1F
- HAR1F je exprimovaný v neokortexe u 7 a 9 týždenných embrií, neskôr aj v iných častiach mozgu (u človeka aj iných primátov)
- Všetky substitúcie v človeku A/T->C/G, stabilnejšia RNA štruktúra (ale tiež sú blízko k telomére, kde takéto mutácie časté)