

DETEKCIA MIKROSATELITOVEJ INSTABILITY Z
GENOMICKÝCH ÚDAJOV ZÍSKANÝCH
SEKVENOVANÍM NOVEJ GENERÁCIE

Školiteľ: Mgr. Jaroslav Budiš, PhD.

Dária Čárska

MIKROSATELITY

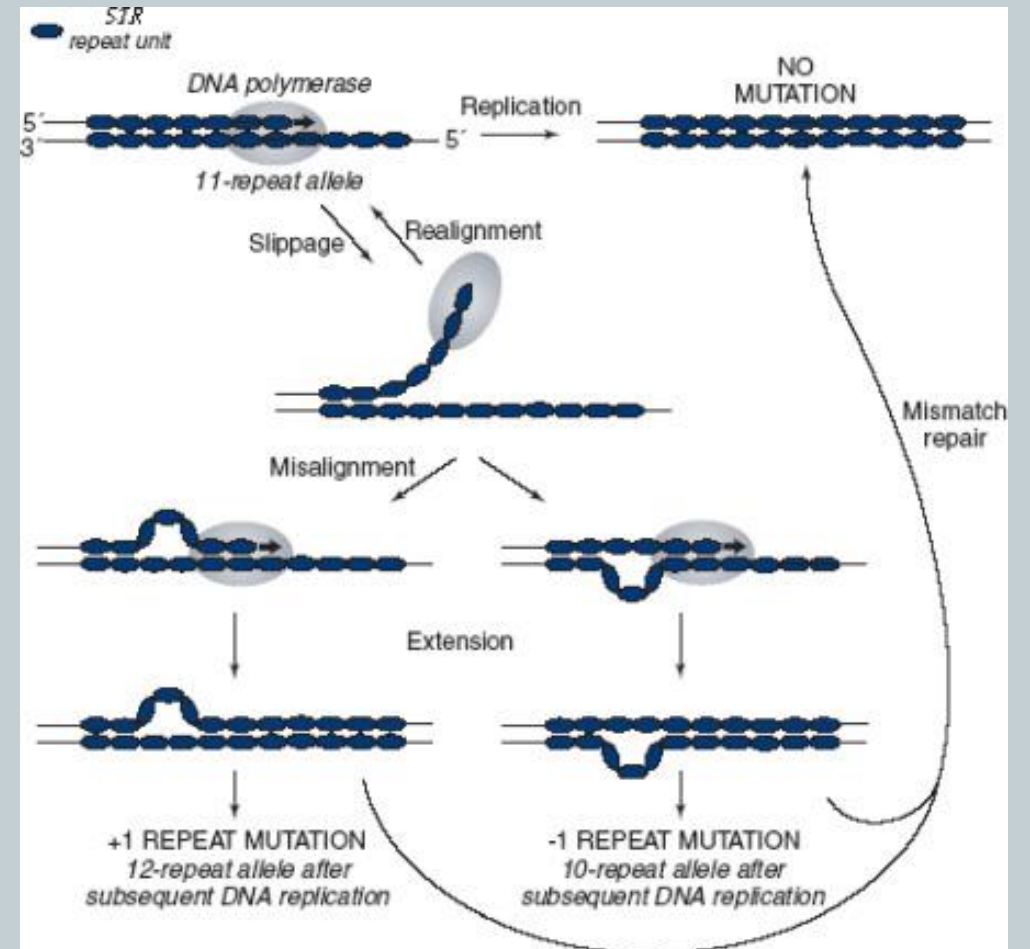
- Krátke tandemové repetície
- Krátke repetitívne DNA sekvencie
- Tvoria približne 3% ľudského genómu
- Zložené z opakujúceho sa motívu nukleotidov dĺžky 1-6
- Homopolymér – mononukleotidový mikrosatelit

GTCGA **CTA CTA CTA CTA CTA CTA** GGAT
trinukleotidový mikrosatelit

GACG**TTTTTTTT**TGCT
homopolymér

MIKROSATELITY

- Zdroj genomickej variability
- Vyššia mutačná rýchlosť
- Hlavným zdrojom ich vysokej premenlivosti v počte opakovaní je skĺznutie DNA-polymerázy



(Hao Fan and Jia-You Chu, 2007)

MIKROSATELITOVÁ INSTABILITA

- Stav genetickej hypermutability
- Výsledkom chybnjej funkcie mechanizmu opravy DNA (mismatch repair)
- Zmeny v dĺžke mikrosatelitovej sekvencie
- Niektoré druhy rakoviny vykazujú mikrosatelitovú instabilitu

POSTUP PRÁCE

Overenie hypotézy, že u pacientov s MSI sa vyskytujú vzácne formy homopolymérov, ktoré nie sú prítomné v populácii:

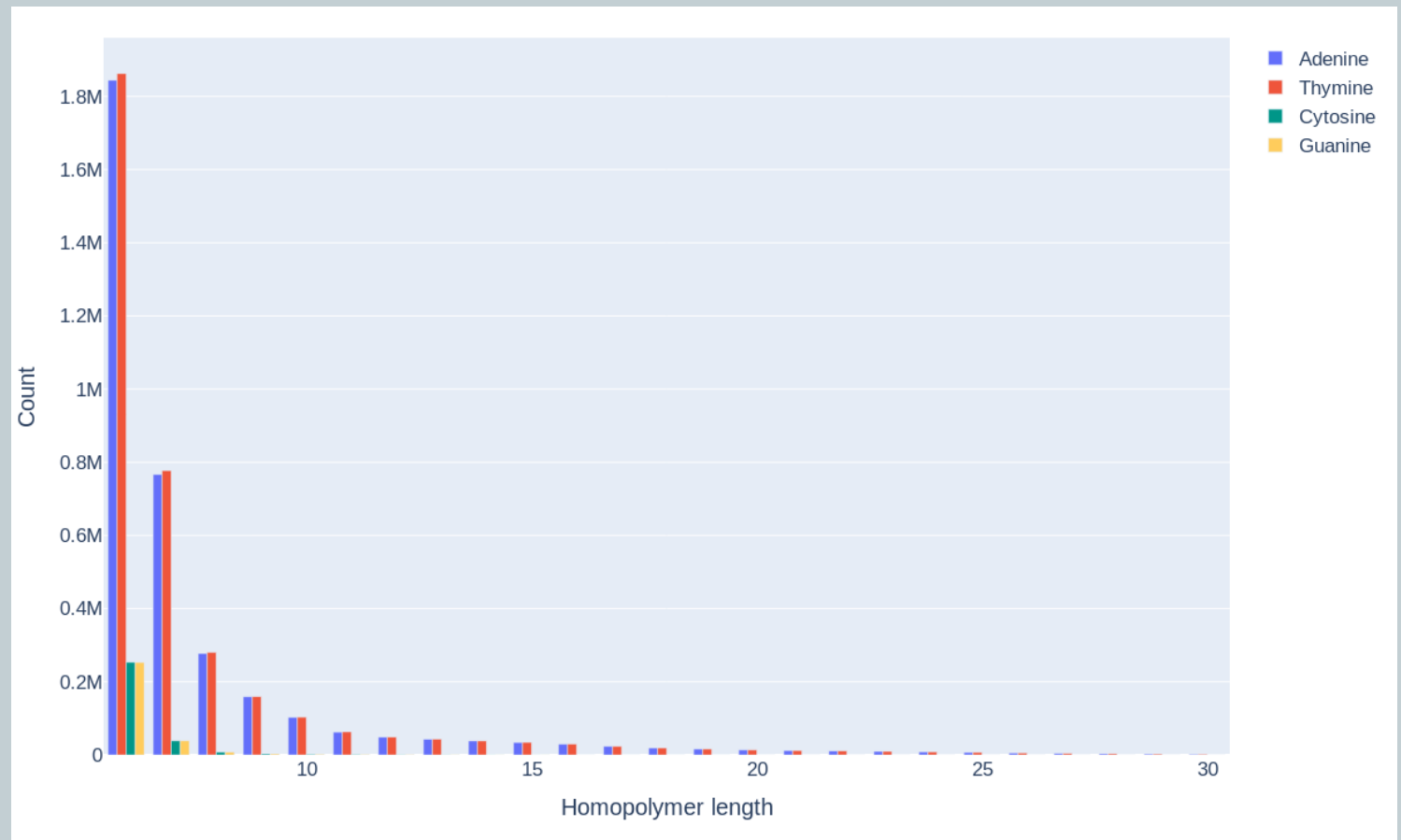
- Agregácia homopolymérov z 10,645 vzoriek slovenskej populácie
- Určenie populačných frekvencií pre ~7,000,000 lokusov
- Porovnanie vzoriek zdravých jedincov a onkologických pacientov
- Porovnanie geneticky vzdialených populácií (slovenské a indické vzorky)

ANALYZOVANÉ DÁTA

- Vzorok sekvenované novou generáciou:
 - Krátke čítania (35 báz)
 - Párové čítania
 - Vysoká kvalita čítaní (chybovosť < 1%)
 - Nízke pokrytie (< 1x)

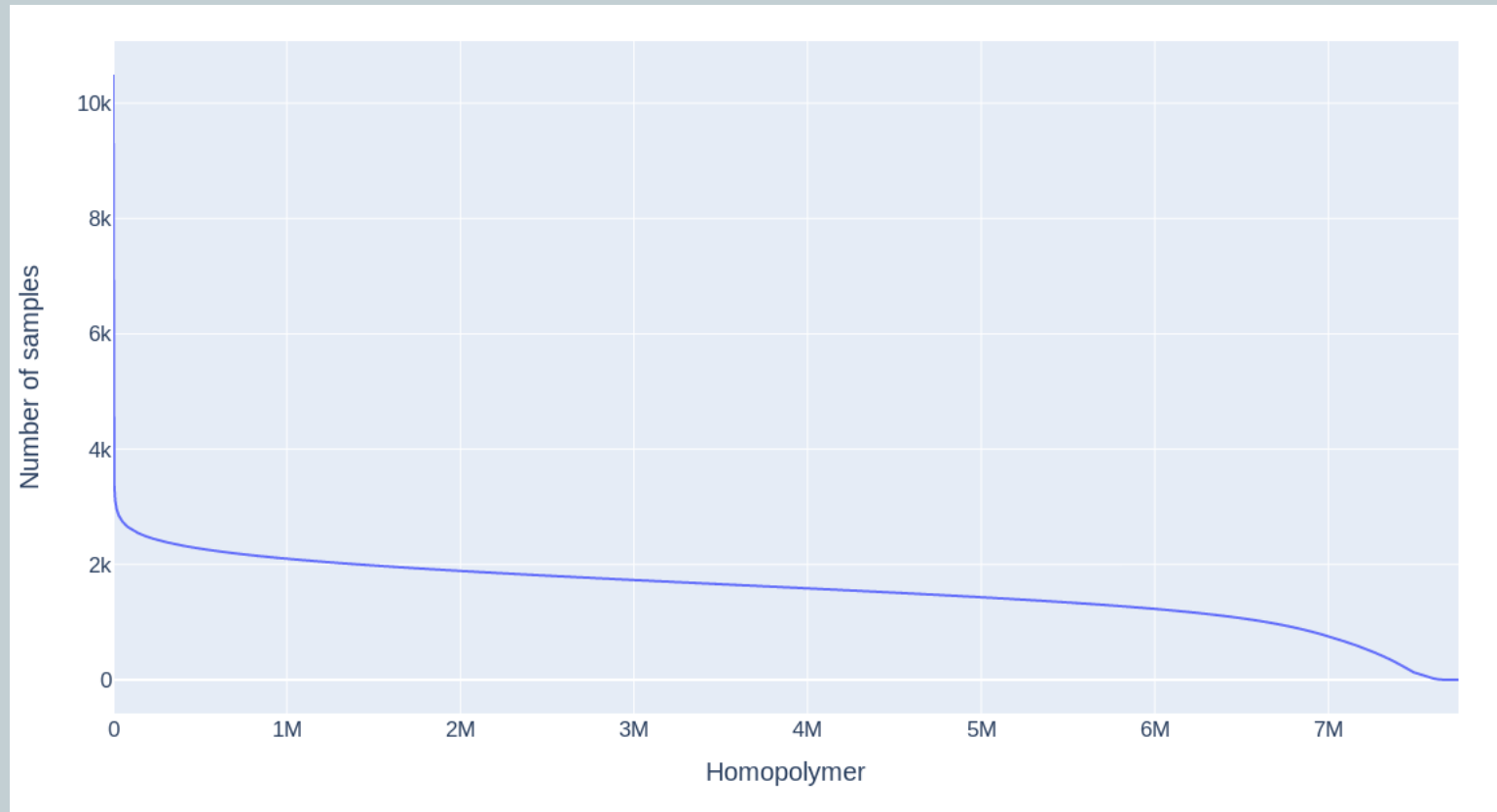
HOMOPOLYMÉRY V REFERENČNOM GENÓME

- Veľkosť ľudského genómu je ~ 3 Gbp
- 7,749,621 lokusov
- Vyššie zastúpenie adenínu a tymínu



POPULAČNÉ FREKVENCIE

- 10,645 vzoriek zdravých jedincov zo slovenskej populácie
- V priemere 14.46% homopolymérov plne pokrytých jednou vzorkou
- 7 miliónov homopolymérov pokrytých približne 2000 krát

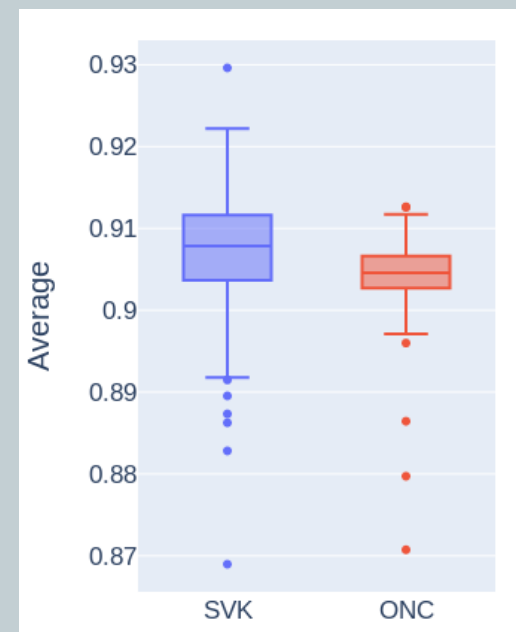


ZVOLENÉ METRIKY NA POROVNANIE RÔZNYCH SÚBOROV VZORIEK

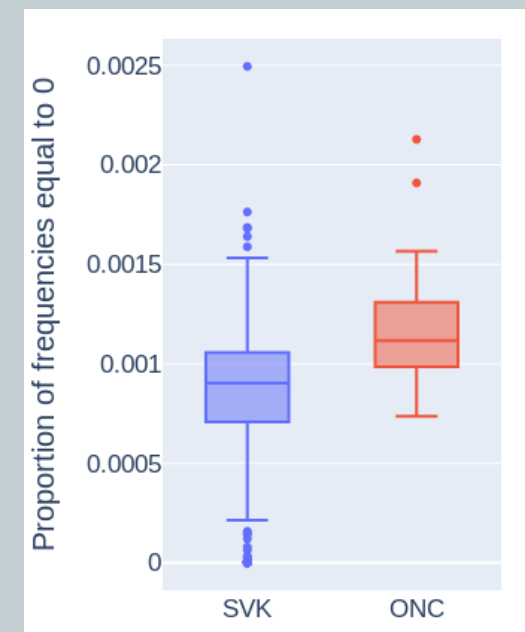
- Priemer relatívnych frekvencií
- Pomer výskytu homopolymérov vo vzorke s relatívnou frekvenciou rovnou 0

POROVNANIE ZDRAVÝCH A ONKOLOGICKÝCH VZORIEK

- Mann–Whitney U test
 - 200 vzoriek zdravých jedincov zo slovenskej populácie
 - 82 vzoriek onkologických pacientov zo slovenskej populácie (kolorektálne a gastrointestinálne nádory)
- Očakávaný trend pozorovaný v oboch metrikách



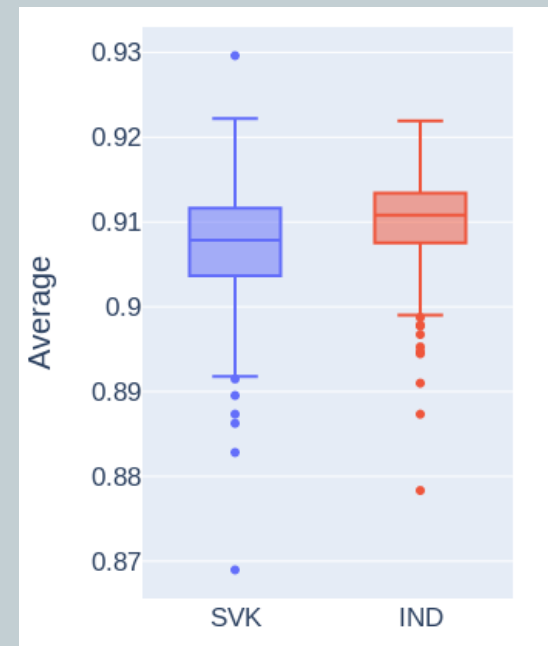
p-hodnota =
2.20002e-06



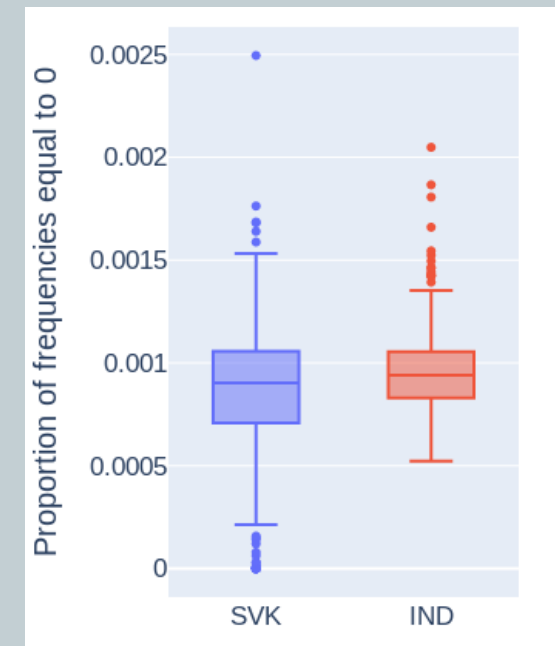
p-hodnota =
3.68497e-13

POROVNANIE SLOVENSKÝCH A INDICKÝCH VZORIEK

- Mann–Whitney U test
 - 200 vzoriek zdravých jedincov zo slovenskej populácie
 - 306 vzoriek zdravých jedincov z indickej populácie
- Rozdiely pozorované v oboch metrikách, sú menšie oproti porovnaniam s onkologickými vzorkami



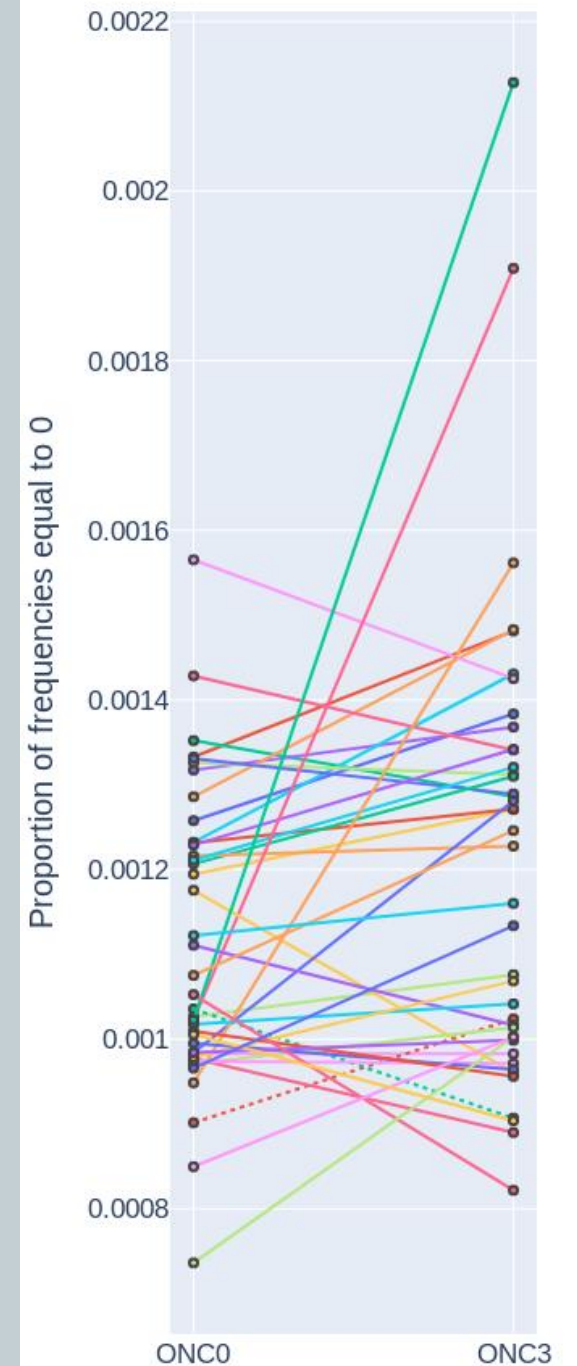
p-hodnota =
7.64507e-08



p-hodnota =
0.00341

POROVNANIE PÁROV ONKOLOGICKÝCH VZORIEK

- Wilcoxon signed-rank test:
 - 41 párov onkologických vzoriek (pred a po operácii)
- Nepodarilo sa ukázať pokles výskytu vzácnych foriem homopolymérov po operácii



VÝSLEDKY

- Vytvorenie populačných frekvencií
- Ukázali sme trend, že dĺžky homopolymérov u onkologických pacientov sa odlišujú od bežných foriem vyskytujúcich sa v populácii
- Homopolyméry majú potenciál na diagnostiku onkologických ochorení
- Výsledky ešte neboli postačujúce na spoľahlivú klasifikáciu

MOŽNOSTI ĎALŠEJ PRÁCE

- Porovnanie mikrosatelitovej instability u rôznych typov rakoviny
- Vytvorenie klasifikátora na rozlišovanie zdravých a onkologických vzoriek na základe mikrosatelitov

ĎAKUJEM ZA POZORNOST