

Výpočtová analýza genómu kvasinky *Magnusiomyces clavatus*

Dominika Békesová
doc. Mgr. Bronislava Brejová, PhD.

Úvod do práce

Magnusiomyces clavatus

- patogénna kvasinka
- nájdená najmä u pacientov s leukémiou
- podobná *M. capitatus*

Magnusiomyces spicifer

- nájdená v Arizone v zhnitom kaktuse

Výsledky

- poskladanie genómu *M. clavatus*
- nájdenie génov a ich proteínových domén
- porovnanie s *M. capitatus*
- čiastočné poskladanie *M. spicifer*

Charakter dostupných dát

- Illumina - krátke čítania
 - nízka chybovosť
 - fluorescenčné značenie
- Oxford Nanopore Technologies - dlhé čítania (rádovo tisíce)
 - väčšia chybovosť
 - kanál nanopóru - elektrický prúd
- PacBio - dlhé čítania
 - real-time
 - fluorescenčné značenie

dáta:

Illumina

Nanopore

PacBio

program:

SPAdes

Canu

Miniasm

+

M. clavatus:

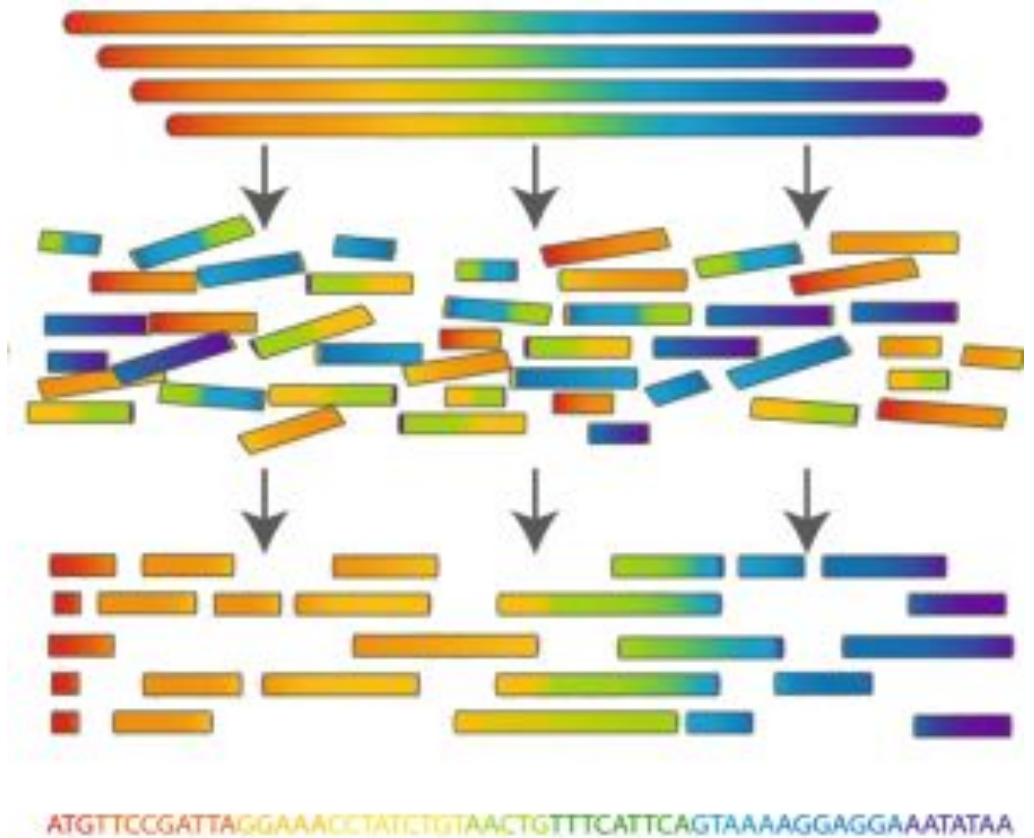
- skladanie - Miniasm, SPAdes, Canu
- čistenie - Racon, Pilon
- anotácia (RNA-seq) - Trinity, Augustus, InterProScan

M. spicifer:

- skladanie - Miniasm, SPAdes
- čistenie - Racon

Skladanie genómu

- vstup
 - čítania
- cieľ
 - zostaviť pôvodnú DNA sekvenciu
 - pomocou prekryvov v čítaniach



Čistenie genómu

- namapovanie čítaní na referenčnú sekvenciu
- princíp konsenzusu

sekvencie:

A A T **T** C C G G

A **T** T A C C G G

A A T A **G** C G G

A A T A C **T** G G

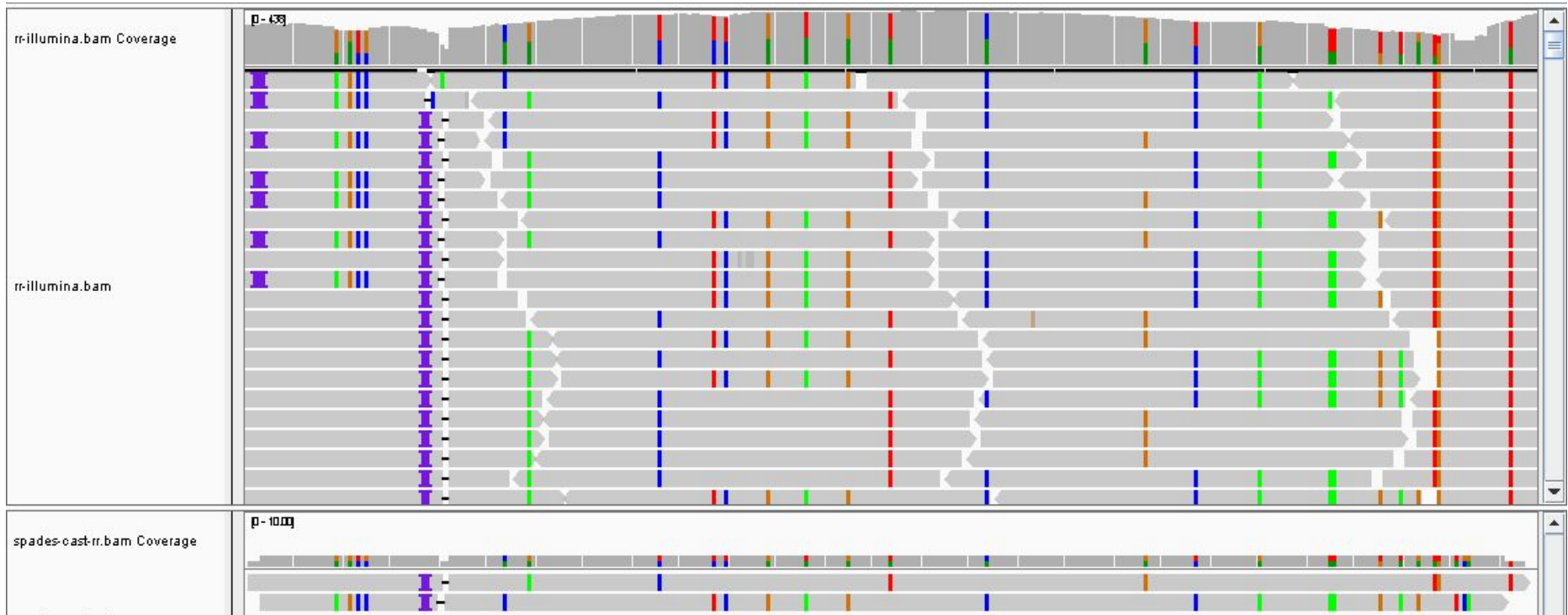
A **T** T **T** **G** **T** **C** G

výsledná
sekvencia:

A A T A C C G G

Magnusiomyces spicifer

- skoro dvojnásobná celková délka genómu → hypotéza dvou verzí



Magnusiomyces clavatus

- poskladaný genóm + transkripty z RNA-seq - anotácia
- bližšia analýza proteínovej domény Hsp70

	počet kontigov	počet kontigov ($\geq 50\,000\text{bp}$)	celková dĺžka	GC(%)	N50
illumina-SPAdes	874	83	17.55M	33.73	40 283
nanopore-canu	44	14	18.42M	34.22	3 721 805
nanopore2-canu	21	20	17.3M	34.29	1 121 059
nanopore-miniasm	230	227	42.15M	34.23	213 993
nanopore2-miniasm	21	16	18.26M	33.87	3 304 945

- chýbajúca SSB subfamily
- SSB nájdená v Canu a transkriptoch

	M. cla	M. cap	M. ing	Description
PF00012	11	47	27	Hsp70 protein

Možnosti ďalšej práce

- dokončiť genóm *M. spicifer*
- doplniť genóm *M. clavatus* o chýbajúce sekvencie proteínov
- bližšie preskúmať ostatné proteínové domény
- hotový genóm vložiť do databázy

Ďakujem za pozornosť