



Oponentský posudek na diplomovou práci Bc. Lucie Korené - Analýza transkriptomu tasemnice *Mesocestoides corti*

Práce Lucie Korené se věnuje biologicky velmi zajímavé oblasti výzkumu – zda a jakým mechanismem může parazit pozitivně ovlivňovat svého hostitele, konkrétně potlačovat vznik a/nebo rozvoj nádorů v hostiteli. Autorka analyzovala transkriptomy tasemnic *Mesocestoides corti* kultivovaných buď *in vitro*, nebo v dvou kmenech myši (inbrední, outbrední) a její práce zahrnovala veškeré kroky od kultivace a zpracování biologického materiálu, přes přípravu sekvenačních knihoven pro bulk RNA-seq, až po bioinformatické zpracování a vyhodnocení transkriptomických dat.

Spis diplomové práce má klasické členění a rozsah 44 stran (bez citací). Součástí práce jsou navíc 4 skripty pro zpracování výsledků (Python) a 3 elektronické přílohy (sady rozsáhlých tabulek výsledků). Oceňuji, že autorka skripty a přílohy dala k dispozici na GitHubu. Autorka v práci cituje 102 literárních pramenů a 12 online zdrojů. Formální stránka práce je vcelku dobrá, i když mám jisté výhrady (moje drobné poznámky jsou uvedeny na konci posudku) – text podle mého soudu místy trpí poněkud nešikovnými či nepřesnými překlady anglických vět a/nebo termínů. Práce na začátku obsahuje velmi stručný slovníček používaných pojmů. To považuji za velmi dobrý nápad a je škoda, že jsou v něm uvedena jen 4 hesla, přičemž v textu je používána široká škála dalších vhodných adeptů. Seznam zkratk v práci chybí. Cíle práce jsou jasně definované a korespondují s Výsledky. Literární přehled obsahuje velmi stručný úvod do biologické problematiky, a dále pak užitečný přehled kroků a nástrojů pro zpracování a analýzu RNA-seq dat. Kapitola Materiál a metody je psána velmi úsporně (necelé 4 strany textu), ale možná zde ze mne jen mluví potrefená biologická husa zvyklá na podrobnější dokumentaci. Každopádně je škoda, že na GitHubu nejsou k dispozici i kompletní skripty pro preprocesing sekvenačních dat, přípravu count tables a detekci diferenciálně exprimovaných genů (DEGs). Významně by to napomohlo reproducibilitě práce, aneb ďábel bývá v detailech (nebo byly použité nástroje „klikací“?). Výsledková sekce popisuje technické aspekty přípravy sekvenačních knihoven a sekvenování, přípravu anotovaných transkriptomů *M. corti*, kvantifikaci genové exprese, detekci DEGs, detekci funkčního nabohacení u DEGs. Výsledky jsou opět popsány velmi úsporně – zde nic proti tomu, ale pak bych čekal, že tento poněkud suchý výčet bude oživen v rámci Diskuze. Ta je ale opravdu velmi stručná (4 strany), do značné míry jen sumarizuje/opakuje výsledky, popř. řeší technické aspekty práce. Biologický význam získaných výsledků dostal jen minimum prostoru. A to je velká škoda, protože zpracovávané téma je fascinující, s reálným praktickým výstupem pro humánní medicínu.

Přírodovědecká fakulta UK

RNDr. Martin Převorovský, Ph.D.



V souhrnu hodnotím kladně, že se autorka úspěšně popasovala s celým projektem od začátku (práce v biologické laboratoři) až do konce (bioinformatické analýzy) a odvedla nemalý kus práce. Myslím si ale, že získané výsledky by si zasloužily hlubší zhodnocení a biologickou diskuzi. Práci doporučuji k obhajobě a navrhuji hodnotit stupněm „velmi dobře“.

OTÁZKY

- 1) Autorka uvádí (str. 3 a Obr. 1), že jedním z mechanismů helminty vyvolané suprese nádorů je indukce angiogeneze u hostitele. Mohla by to autorka blíže vysvětlit? Očekával bych totiž naopak supresi angiogeneze.
- 2) Proč je mapovací nástroj HISAT2 vhodný spíše pro genomy s nízkou komplexitou a STAR pro genomy složitější?
- 3) Jaký je rozdíl či vztah mezi "velikostí sekvenační knihovny" a "hloubkou sekvenování" (str. 13)?
- 4) Co znamená, že nástroj má "dobrou časovou složitost" (str. 22)?
- 5) Jaký vliv mělo použití Rcorrector na mapping rate a výsledky analýzy DEGs?
- 6) Čím si vysvětlujete nízké procento namapovaných readů (55-73%)? Jaký je původ/zdroj nenamapovaných readů?
- 7) Tab. 8 a 9 - řada nejvíce regulovaných DEGs je ve skutečnosti málo exprimovaná (popř. má 0 counts v některých replikátech) - nebojíte se zavádějících výsledků/závěrů (velké chyby malých čísel)?
- 8) Tab. 12 – Zdůrazňujete, že homology (potenciálních) savčích tumorových supresorů jsou diferencielně exprimované v tasemnicích. To by podle mne mohlo suprimovat nádory u tasemnic... Jak by tím ale mohly být ovlivněny nádory savčího hostitele?
- 9) Kap. 5.4 - Chápu, že Bonferroniho korekce je příliš konzervativní. Mohla byste místo toho použít nějakou jinou metodu korekce na mnohonásobné testování (namísto "opatrného" používání nekorigovaných výsledků)?
- 10) Detekovala jste změny v expresi genů pro produkci sekundárních metabolitů tasemnicemi - není to nakonec ten nejzajímavější výsledek z hlediska ovlivnění nádorů hostitele? Proč ano / ne?

Přírodovědecká fakulta UK

RNDr. Martin Převorovský, Ph.D.



DROBNÉ POZNÁMKY

- překlepy (hlavně malá/velká písmena) už na titulní straně DP..., i v komentářích skriptů
- Tab. 2 v textu odkazovaná dříve než Tab. 1 = divné
- str. 9 - hromadění readů u 3' konců transkriptů neznamená nízkou kvalitu readů, ale vstupní RNA
- Tab. 3 - "Reference" je zavádějící (každý z programů lze použít s genomem i transkriptomem, navíc u prokaryot se rozdíl stírá). Mnohem lepší by bylo rozdělení SW na splice-un/aware
- seznam citací obsahuje všude zbytečný text "[online]"
- str. 15 - "mohou být terčem terapeutických cílů"... ??? "míra exprese je vyjádřena pomocí p-hodnoty" ???
- který SW je "nejlepší" se těžko stanovuje objektivně a asi bych se úplně nespolehal na 6 let starý benchmark (DESeq2 vs edgeR str. 16)
- obr. 6: "konzervativní úseky" -> konzervované úseky
- str. 17 - BLAST = basic local alignment search tool
- str. 21 - Bonferroniho korekce spočívá ve vynásobení (nikoli vydělení) p-value počtem testů
- str. 23 - není uvedena délka readů pro použité sekvenování na přístroji NovaSeq
- str. 26 - DEGs byly patrně vybrány jako $p < 0.05$ a nikoli 0.5
- Obr. 9 - osa y není popsána v grafu (pouze slovně v legendě)
- Obr. 10 - horní panel - ne všechny barvy jsou popsány v legendě

V Bystřici 4. 9. 2022

RNDr. Martin Převorovský, Ph.D.

Přírodovědecká fakulta UK

RNDr. Martin Převorovský, Ph.D.