

Posudek oponenta na diplomovou práci

oponentský posudek

Jméno posuzovatele:

Jaroslav Nunvář

Datum:

7.9.2021

Autor:

Bc. Kateřina Poláčková

Název práce:

Sekvenování nové generace v klinické virologii: optimalizace metody a použití na vzorcích s neznámým původcem infekce

Cíle práce

- na simulovaných infekčních vzorcích obsahující virus, nebo skupiny virů otestovat tři různé přípravy vzorků s použitím tří různých tří kitů od firmy Oxford Nanopore a vybrat nejvhodnější kit pro identifikaci patogenu v klinickém vzorku s neznámou infekcí
- na ředěných infekčních vzorcích otestovat sensitivitu použitých metod
- zhodnotit vliv přítomnosti kontaminujících nukleových kyselin na výsledky sekvenování
- kromě otestování a přípravy protokolu pro laboratorní část otestovat také různé nástroje, jejich výhody a nevýhody pro analýzu dat

Struktura (členění) práce, odpovídá požadovanému? ANO

Rozsah práce (počet stran):

- Úvod (1)
- Cíle práce (1)
- Přehled literatury (16)
- Materiál a metody (37)
- Výsledky (15)
- Diskuze (4)
- Souhrn (1)
- Seznam použité literatury (5)

Je uveden anglický abstrakt a klíčová slova? ANO

Je uveden seznam zkratek? ANO

Literární přehled:

Odpovídá tématu? ANO

Je napsán srozumitelně? ANO i NE

Použil(a) autor(ka) v rešerši relevantní údaje z literárních zdrojů? ANO

Jsou použité literární zdroje dostatečné a jsou v práci správně citovány? správně citovány ANO, dostatečné spíše NE

V úvodu literárního přehledu jsou popsány a porovnávány technologie sekvenování nové (2. a 3.) generace (NGS). Bohužel chybí pojednání o jejich chybovosti, což je nejzávažnější rozdíl a mj. faktor, díky němuž technologie 3. generace nenahradily 2. generaci. Rovněž kritéria finanční náročnosti NGS, zásadní pro aplikovatelnost, jsou pojednána jen okrajově. Velmi suboptimální je kapitola 3.4. Jedná se o zbytečně detailní a proto nepřehledné digesty ze šesti publikací, které se zabývaly reálným využitím sekvenace Oxford Nanopore v klinické virologické diagnostice. Výsledky těchto studií vůbec porovnávány (např. shrnující tabulka by zde byla velmi vhodná), čímž se ztrácí potenciální přidaná hodnota rešerše.

Materiál a metody:

Odpovídají použité metody experimentální kapitole? ANO

Kolik metod bylo použito?

zhruba 10 molekulárně-genetických metod (gelová elektroforéza, izolace NK 2x, kvantifikace NK 2x, deplece NK 2x, real-time PCR, příprava sekvenačních knihoven 3x) a metody bioinformatické analýzy (MG-RAST, sada analýz v programu Geneious)

Jsou metody srozumitelně popsány? ANO

Experimentální část:

Je vysvětlen cíl experimentů? ANO

Je dokumentace výsledků dostačující? ANO

Postačuje množství experimentů k získání odpovědi na zadané otázky? ANO

Hlavně díky dobrému experimentálnímu designu, danému dlouhodobými zkušenostmi Laboratoře molekulární genetiky, jsou design experimentální části a zpracování výsledků adekvátní (ojedinělé výjimky viz Otázky a připomínky oponenta níže).

Fundamentálním nedostatkem kapitoly Výsledky je způsob prezentace. Tato část DP sestává pouze z obrázků a textu v jejich legendách, bez „klasického“ pojícího textu. To je v nesouladu nejen se standardem diplomových prací, ale i výzkumných publikací obecně. Negativní vliv tohoto stylu prezentace na čtenářskou srozumitelnost je pochopitelně zásadní.

Diskuze:

Je opravdu diskuzí, nejde jen o konstatování vlastních výsledků? ANO

Jsou výsledky porovnávány s literaturou? NE

Jsou uvedeny nějaké hypotézy či návrhy na další řešení problematiky? NE

Diskuze diplomové práce je pojata jako „vnitřní“, tj. diskuze získaných výsledků vzhledem k mantinelům a proměnným daným zvoleným experimentálním nastavením (např. porovnání tří typů přípravy knihoven). Porovnání výsledků s literaturou v sekci Diskuze, až na dvě výjimky, absentuje, navzdory obšírnému popisu výsledků tematicky spřízněných studií v Literárním přehledu (viz výše), s nimiž se porovnání přímo nabízelo. Proto je v sekci Diskuze odkazováno na pouhých pět citací. Chybí zhodnocení vybraného způsobu testování biologických vzorků (Rapid kit) z hlediska možností jeho vylepšení/optimalizace, a předpokladů a možností jeho reálné použitelnosti v diagnostické praxi.

Závěry (Souhrn) :

Jsou výstižné? ANO

Závěry jsou maximálně stručné až lakonické.

Formální úroveň práce (obrazová dokumentace, grafika, text, jazyková úroveň):

Grafická úroveň obrázků je (v pdf verzi) dostatečná a obrázky vhodně vybrány vzhledem k textu.

Jazyková úroveň je na poměrně nízké úrovni, v práci se vyskytuje větší množství gramatických chyb (chybějící čárky v souvětí, shoda přísudku s podmětem atd.; zde nevypisují) a překlepů, vč. názvu práce. Seznam následujících překlepů, které jsem při letném pročitání textu zaregistroval, jistě není vyčerpávající:

... s neznámý původcem... (název práce), ... která se schopna získat... (s.7), ... pseky o délce... (s.9), ... konzervované a (hyper)variabilní oblasti... (s.11), ... po sekvenování jsou... (s.11), ... přespána do cDNA... (s.23), ... sekvovány na přístroji... (s.24), ... zachrnující kmeny... (s. 27), ... ústav lékařské mikrobiologie... (s.31), ... byly vzorky poolovány... (s.50), ... obsahem dlouhých četní... (s.56), ... artefaktů u jsme se zbavili... (s.59), ... na gelové elektroforéze... (s.78)

Splnění cílů práce a celkové hodnocení:

Cíle, které si autorka DP kladla, byly splněny, zásluhou vhodného designu, vlastního experimentálního provedení a analýzy výsledků. Splnění cílů je tedy hlavním kladem práce, zároveň však, přeneseně, i její hlavní slabinou, protože ostatním hodnoceným kritériím DP zdaleka nebyla věnována taková pozornost. Zpracování literárního přehledu a diskuze je nepoměrně slabší a v některých částech na hranici obhajitelnosti. V předložené formě se tak bohužel nepodařilo přenést práci z idiosynkratického prostředí klinicko-výzkumné virologické laboratoře širšímu akademickému publiku. I přes tyto výhrady práci doporučuji k obhajobě.

Otázky a připomínky oponenta:

Připomínky (faktické chyby a nesrovnalosti; reakce při obhajobě není nutná, ale je vítaná):

- s.15: „*Nanopory prochází ionty s konstantním napětím, které je narušeno průchodem vlákna DNA. Každá jeho báze je jinak veliká a charakteristicky naruší proud iontů a tím se změní i napětí. Každý nanopore je napojen na vlastní elektrodu, která je napojena na kanál a senzorový čip, který měří změnu napětí.*“ Měří se změna proudu, nikoliv napětí. Analogicky je chybně popsána osa y obr. 10.
- s.18: „... 33 % bylo identifikováno pomocí mNGS i konvenčních metod, 26 % pouze konvenčními metodami a 22 % pouze mNGS“ Obr. 12 udává ve stejném pořadí 19/26/13 vzorků, v prostřední hodnotě je rozpor.
- s.20: „*Při sekvenování technologií ON se používá buď sekvenování amplikonů, nebo sekvenování metagenomické. Sekvenování amplikonů má nespornou výhodu v tom, že pomocí specifických primerů lze ze vzorku tzv. „vytáhnout“ pouze virová čtení, a to pouze od viru, který chceme sekvenovat. To se velmi hodí při sekvenování viru přímo z biologického vzorku, který obsahuje i další kontaminující sekvence.*“ Tato pasáž je zbytečná a zavádějící. V tomto případě benefity dané samotným použitím patogen-specifické PCR (vysoká rychlost, nízká cena, možnost Sangerovsky sekvenovat PCR produkt) na primární vzorky kontraindikují následné použití ON na tyto amplikony.

Otázky (prosím o zodpovězení formou prezentace, vč. vypsání celého znění otázky):

- 1) s.33: „*Nejprve byla směs simulující infekční vzorek rozdělena na dvě poloviny, čímž tak vznikla RNA a DNA frakce. Tento krok byl zařazen proto, že virus v klinickém vzorku, i přes to že může mít vysokou kvantitu, je sekvenčních datech zastoupen většinou málo. Důvodem je to, že kontaminující DNA/RNA pocházející například z bakteriálních či lidských buněk je mnohem delší, což často úplně přehluší virový signál. Tím, že jsme vzorky rozdělili do dvou frakcí lze částečně odstranit tyto kontaminující sekvence pomocí enzymů RNázy a Dnázy.*“ Nerozumím logice tohoto zdůvodnění. Pokud rozdělíte vzorek na dvě poloviny, z každého odstraníte buď DNA, nebo RNA, a následně tyto vzorky znovu spojíte v poměru 1:1 (s.42), výsledkem bude původní vzorek s poloviční koncentrací DNA a poloviční koncentrací RNA. Prosím o vysvětlení. Dále uveďte, jakým způsobem („trikem“) se ve viromových studiích dosahuje (často výrazného) ochuzení podílu kontaminujících NK původem z hostitele a dalších přítomných buněk (např. bakterií).
- 2) Jak vysvětlíte skutečnost, že větvené produkty celogenomové amplifikace vykazují větší elektroforetickou mobilitu (obr. 22) než tytéž produkty linearizované („odvětvené“) působením T7 endoI (obr. 23), navzdory tomu, že jsou větší (obr. 21)?
- 3) Jaká konkrétní kritéria byla použita pro mapování readů na referenci v programu Geneious (maximální % neshodných bází mezi readem a referencí, maximální % gapů)? Toto jsou základní determinanty účinnosti mapování. Vzhledem k chybovosti sekvenací 3. generace bych očekával poměrně relaxovaná kritéria. Zkoušela jste, alespoň na jednom

vzorku, variovat mapovací nastavení a porovnat počty namapovaných readů?

- 4) s.58: „...existují mezi namapovanými sekvencemi také bloky velmi krátkých fragmentů do cca 50 nukleotidů (obrázek 24). Tyto sekvence se neshodovaly se sekvencí žádného barcodu, ani adaptéru a jejich původ se nám nepodařilo dohledat. Nicméně, ač se jednalo o cca stejnou sekvenci, mapovala se na více míst a vůbec se neshodovala s vybranou referencí.“ Navzdory tomu, že se nepodařilo vyhledat původ těchto „mysteriózních“ sekvencí (ani v GenBanku, s.77), lze podle mapovacího patternu na obr.29 jasně vydedukovat několik informací o jejich původu. Rozeberte obr.29 a uveďte nejdůležitější informace, které z něj lze zjistit.
- 5) Obrázek 26 je směsným vynesemím distribuce velikostí readů ze všech vzorků. Odpovídají jednotlivé peaky jednotlivým knihovnám, tj. je ve skutečnosti jeden velikostní peak na knihovnu?
- 6) Máte k dispozici informaci, který ze tří alternativních způsobů přípravy knihovny podchytil největší část diverzity/komplexity sekvencí přítomných ve vzorcích? Pokud ano, navrhnete zdůvodnění pro toto pozorování.
- 7) s.78, obr.23: „Při přípravě „WGS kitem“ po testování vzorků na gelové elektroforéze bylo vidět, že po izotermální amplifikaci a následné reakci s T7 endonukleázou I vznikly produkty o stejné délce 10 kbp.“ Jste si jistá, že produkty měly délku 10 kbp? Jakou metodu byste použila pro definitivní určení velikosti těchto fragmentů?

Návrh hodnocení oponenta (známka nebude součástí zveřejněných informací)

výborně velmi dobře dobře nevyhověl(a)

Podpis oponenta: