

Posudek oponenta na diplomovou práci	
<input checked="" type="checkbox"/> oponentský posudek	Jméno posuzovatele: RNDr. Dana Holá, Ph.D.
	Datum: 28.5.2015
Autor: Bc. Michaela Daňková	
Název práce: Cílené sekvenování nové generace kandidátních genů zodpovědných za poruchu spermatogeneze a neplodnost mužů	
Cíle práce: Stanovit sekvence dvanácti kandidátních genů u neplodných mužů s poruchou spermatogeneze a prokázat nálezný, případně absenci patogenních mutací v těchto genech. Nalezená mutace by byla vlastní etiologií neplodnosti. (přesná citace z kapitoly Cíle práce). Práce byla součástí širšího projektu řešeného na pracovišti školitele a podporovaného grantem MZ ČR „Screening genetických determinant oligo/teratozoospermie u neplodných mužů, s využitím informací získaných pomocí zvířecích modelů“.	
Struktura (členění) práce, odpovídá požadovanému? ANO <i>Rozsah práce (počet stran):</i> 62 <i>Je uveden anglický abstrakt a klíčová slova?</i> ANO <i>Je uveden seznam zkratek?</i> ANO	
Literární přehled: <i>Odpovídá tématu?</i> spíše ANO, očekávala bych nicméně kapitolu o známých mutacích v konkrétních genech, u nichž se studovala možná souvislost s neplodností/poruchami spermatogeneze u člověka (4 takové práce jsou pak zmíněny na začátku kapitoly Diskuze, existuje jich však samozřejmě mnohem víc, viz např. původní asociační studie uvedené v přehledovém článku Ferlin et al. 2007, který autorka četla, nebo v přehledových článcích Kovac et al. 2013, Hotaling and Carrell 2014, a mnohé další). Zařazení takové kapitoly je pochopitelně zcela na rozhodnutí školitele nebo diplomantky, domnívám se nicméně, že by výrazně přispěla k lepšímu seznámení se s celou problematikou a byla by pak vhodným základem Diskuze. Naopak kapitola „Sekvenování nové generace – platforma GS Junior (Roche)“ je v podstatě zbytečná, popisuje pouze základní princip pyrosekvenování na úrovni, kterou by v dnešní době kterýkoli čtenář orientující se v genetice měl znát. Tato kapitola by byla podle mého názoru opodstatněná tehdy, když by autorka např. ve vlastní diplomové práci porovnávala různé sekvenační metody; v každém případě bych při jejím zařazení do práce očekávala např. podrobnější přehled o tom, jaké a proč má tato metoda výhody, nevýhody a využití, podložený primárními citacemi a ne pouze internetovým odkazem na příslušný sekvenační systém a citací jednoho přehledového článku (str. 25). <i>Je napsán srozumitelně?</i> ANO, některé informace se místy zbytečně opakují, ale není jich mnoho. <i>Použil(a) autor(ka) v rešerši relevantní údaje z literárních zdrojů?</i> ANO, ale mohlo jich být rozhodně více. Tato část obsahuje cca 55 citací, z nichž cca 20 jsou buď citace lékařských učebnic/podobné literatury, nebo citace přehledových článků/kapitol v knihách či sbornících. Nic proti, tyto citace jsou na svých místech opodstatněné, ale očekávala bych, že autorka diplomové práce prostuduje větší počet primárních studií. Většina prací citovaných v Literárním přehledu pochází z let 2009-2011, mladší práce jsem našla pouze dvě, přičemž velmi pochybuji, že by se výzkum v této oblasti mezitím zastavil. U kapitol věnovaných jednotlivým kandidátním genům v některých případech chybí konkrétní informace o tom, jak/proč jsou jejich mutace u modelových savců spojovány s poruchami plodnosti (např. geny <i>CDC4B</i> , <i>CDC42</i>). <i>Jsou použité literární zdroje dostatečné a jsou v práci správně citovány?</i> K dostatečnosti literárních zdrojů viz komentář výše, ke správnému citování odpovídám ANO v případě odborných článků aj. publikací, NE v případě internetových odkazů. Všechny práce uvedené v textu jsou obsaženy i v Seznamu použité literatury a vice versa, způsob citování v Seznamu použité literatury je v pořádku až na několik drobných chyb (nesprávné zkratky časopisů u	

prací Field et al. 2011, Ivanov et al. 2000, Kara et al. 2010, Singh et al. 2014, nesprávně uvedené stránky u práce Nieschlag et al. 2010). Naproti tomu internetové odkazy jsou vždy uvedeny pouze jako odkaz na úvodní stránku příslušného webu. To znamená, že např. v kapitolách věnovaných jednotlivým kandidátním genům se jedná o úplně úvodní stránku National Center for Biotechnology Information (NCBI), přičemž odkazy v textu práce se týkají informací, které člověk najde spíše v Online Mendelian Inheritance in Man (OMIM) databázi. Na tu se samozřejmě přes NCBI lze dostat, ale v textu by měl u příslušného tvrzení vždy být uveden přesný internetový odkaz týkající se konkrétního genu, nebo – ještě lépe – přímo příslušná primární citace. Nestáčí, že tyto citace jsou uvedeny v Tab. 2.2. – z té není jasné, kterých konkrétních tvrzení v následujícím textu se týkají.

Materiál a metody:

Odpovídají použité metody experimentální kapitole? ANO

Kolik metod bylo použito? Kvalita diplomové práce nezáleží na tom, kolik metod během jejího zpracování bylo použito, ale zda byly metody zvoleny vhodně s ohledem na stanovené cíle a naplánované experimenty. V „laboratorní“ části práce měla autorka za úkol osekvenovat vybrané geny, přičemž měla podle svého vyjádření k dispozici vzorky RNA a DNA z periferní krve. Použila reverzní transkripci, PCR (primery navrhl její školitel), elektroforetickou separaci k ověření PCR produktů a přípravu vzorků pro sekvenování postupem doporučeným pro použitou pyrosekvenační platformu (sekvenátor GS Junior, Roche). V tomto ohledu lze metody považovat za zcela adekvátní. Dalo by se sice polemizovat o vhodnosti zvolené metody NGS (od pyrosekvenování se v současnosti upouští z mnoha důvodů; některé z nich sama autorka zmiňuje na závěr kapitoly Diskuse), ale to nepochybně záviselo na technickém vybavení pracoviště a rozhodně to nelze považovat za chybu. Ke zpracování získaných dat autorka použila dva bioinformatické programy (PROVEAN Human Genome Variants, PolyPhen-2). Získaná data však adekvátně statisticky nezhodnotila. Programy sice používají určité statistické metody, závěry založené pouze na nich však nemusí být správné (viz dále).

Jsou metody srozumitelně popsány? ANO v případě „laboratorních“ metod (i když v některých částech popis připomíná spíše „laboratorní kuchařku“ než vědecký text), NE u bioinformatických programů. V tomto případě bych přivítala podrobnější popis principů, podle kterých příslušné programy posuzují význam nalezených polymorfismů/alel (které parametry program počítá, podle čeho predikuje funkční význam jednotlivých alel, jaké statistické metody přitom používá ...), a v každém případě musí být uvedeny konkrétní parametry nastavení použité při příslušných analýzách (např. v případě PolyPhen-2 je – podle popisu na internetu – možné zvolit ze dvou výchozích modelů, případně rozhodnout se mezi dvěma kritickými hladinami, u PROVEAN je možné různě nastavit kritickou hranici apod.). Popis pouhým odkazem na příslušný internetový zdroj je nedostatečný – jednak odkaz za pár dnů již nemusí platit, jednak – a to hlavně – autoři programů mohou mezitím vyvinout nové verze, pracující na odlišných principech a se zcela odlišným nastavením parametrů. Budoucí čtenář práce pak, i když si vyhledá příslušný program na internetu, dostane zcela chybné informace o metodách použitých při hodnocení výsledků.

Experimentální část:

Je vysvětlen cíl experimentů? ANO v samostatné kapitole Cíle práce. Chybí však výchozí hypotéza, která by měla být na základě výsledků vhodně postavených experimentů posléze zamítnuta nebo nezamítnuta.

Je dokumentace výsledků dostačující? NE. Kapitola Výsledky má 6 stran, přičemž první 3 a půl strany patří ve skutečnosti do kapitoly Materiál a metody, neboť se týkají popisu hodnoceného souboru pacientů (stanovení karyotypu, mikrodeleci chromozómu Y, mutací v *CFTR* genu a diagnózy kvality spermií). Tuto rutinní diagnostiku autorka neprováděla (jak sama v práci uvádí). To samozřejmě nijak nevádí a dokonce by nevádílo ani to, že takový

popis zahrnula do části Výsledky, pokud by to ale mělo své opodstatnění. Tj. pokud by např. tuto tabulku spojila s vlastními výsledky (především výstupy z bioinformatických programů apod.) a použila by ji např. pro stanovení korelací nebo jinou statistickou analýzu. Na konci kapitoly Materiál a metody autorka píše, že „byla vytvořena tabulka v programu Microsoft Excel ...“, která zřejmě právě tento typ údajů obsahuje. Proč tedy tato tabulka není součástí diplomové práce (alespoň jako příloha)? Tak, jak jsou výsledky prezentovány v diplomové práci, nemá čtenář absolutně žádnou možnost si ověřit, že tvrzení v textu týkající se „škodlivosti“ jednotlivých variant jsou skutečně experimentálně podložena.

Postačuje množství experimentů k získání odpovědi na zadané otázky? NE. Hlavní problém je v tomto případě v tom, že byl hodnocen vlastně jen soubor pacientů (73 vzorků, z čehož bylo posléze ještě 12 vzorků vyloučeno) a k tomu neexistovaly adekvátní kontroly (hodnoceno bylo jen 7 kontrolních vzorků, a to, jak autorka uvádí na str. 46, jen z toho důvodu, že sloužily pro sledování kvality sekvenování). Proč autorka nehodnotila více kontrol (vzorky přitom k dispozici zřejmě byly – viz tvrzení autorky na str. 55, že na pracovišti je k dispozici cca 100 kontrolních vzorků) a hlavně – proč oba soubory spolu statisticky neporovnála? Veškeré závěry o nepatogenitě různých alelických variant nalezených u hodnocených pacientů jsou totiž v práci činěny pouze na základě toho, že použité bioinformatické programy nevyhodnotily příslušné alely jako potenciálně škodlivé. To samozřejmě závisí na nastavení daného programu a dostupných datech v bioinformatických databázích, ale – co je důležitější – to, že daná mutace genu neovlivní nějakou důležitou doménu proteinu (což je hlavní princip, na kterém oba programy pracují – alespoň pokud jsem měla možnost to z jejich popisu na internetu vyčíst), nemusí nic říci o případné asociaci dotyčné alely s onemocněním. Klasickým případem, kdy i např. synonymní mutace mohou být škodlivé, jsou choroby spojené s upřednostňováním určitých kodónů, tzv. „codon usage bias“. Práce, které hledají souvislost mezi určitou alelickou variantou nějakého genu a nějakou chorobou/stupněm jejího projevu, by měly být založeny na statisticky podloženém nálezu průkazné asociace např. na základě populačních frekvencí příslušných alel/genotypů u reprezentativních souborů postižených a zdravých osob apod. V předložené diplomové práci žádná taková statistická analýza, resp. porovnání pacientů a kontrol, provedena nebyla a těžko tedy lze hovořit o prokázané nebo neprokázané patogenitě jednotlivých alelických variant. Pokud by diplomová práce měla být schopna obstát jako samostatná vědecká jednotka (malý vědecký úkol, který má diplomant řešit), experimenty byly od počátku chybně postaveny. I pokud by experimenty v diplomové práci byly pojímány pouze jako rutinní analýza – část výzkumného úkolu řešeného na pracovišti školitele (přičemž by analýzu kontrolních vzorků prováděl někdo jiný – předpokládám, že tak tomu pravděpodobně bylo nebo bude), mělo by být v práci alespoň uvedeno, že získaná data budou posléze určitým způsobem porovnána s kontrolními daty, aby bylo zřejmé, že autorka ví, jak se tento typ výzkumu správně dělá. Nic takového ovšem v práci nikde zmíněno není.

Diskuze:

Je opravdu diskuzí, nejde jen o konstatování vlastních výsledků? ANO, je však velmi slabá.

Jsou výsledky porovnávány s literaturou? částečně ANO, ale očekávala bych širší pojetí (Diskuze má celkem 3 strany, některé její části by přitom spíše patřily do Úvodu nebo Literárního přehledu). Studií zabývajících se monogenními mutacemi zodpovědnými za mužskou neplodnost (poruchy spermatogeneze) je celá řada, autorka však v této části práce cituje pouze 4. Dalo by se také např. diskutovat o tom, proč u modelových savců asociace mutací v hodnocených genech s neplodností prokázána byla a u člověka podle autorky ne, apod. Rovněž zcela chybí diskuse o dalším plánovaném osudu získaných dat a způsobu jejich zpracování a hodnocení, aby mohly být vyvozeny rozumné závěry (viz výše).

Jsou uvedeny nějaké hypotézy či návrhy na další řešení problematiky? ANO, jedná se ovšem jen o návrh rozšířit hodnocený soubor pacientů (opět bez jakékoli zmínky o zahrnutí regulérních kontrol a srovnání s nimi), rozšířit skupinu kandidátních genů a přejít na jinou platformu sekvenování.

Závěry (Souhrn):

Jsou výstižné? NE (jsou vlastně jakýmsi druhým abstraktem, kde je znovu stručně popisován cíl, metody a materiál, výsledky – nepřesně, resp. nepodloženě – a návrh dalšího postupu)

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

Práce má dobrou formální úroveň, použité obrázky a tabulky jsou přehledné, obsahuje přijatelný počet překlepů a jazykových/stylistických chyb (týkají se většinou interpunkce), celková jazyková úroveň je dobrá (pouze v některých místech je práce psána až moc úsporným jazykem). Některé nevhodné formulace jsou spíše úsměvného charakteru, např. „„platforma GS Junior ... se hodí pro sekvenování genomu virů, bakterií a dalších mikrobů, plísní, zvířat a rostlin“ (přesná citace ze str. 25) – ničeho dalšího?; „sekvence různých podobných délek“ (str. 25); „normální genom“ (str. 27, 46) – lépe by bylo použít termín „referenční genom“; „pacienti, kteří se sekvenovali najednou v jednom běhu“ (str. 27); „malá/velká agaróza“ (str. 33). Také formulace názvu práce „Cílené sekvenování nové generace kandidátních genů zodpovědných za poruchu spermatogeneze a neplodnost mužů“ přímo vybízí k otázce, zda jde o „novou generaci kandidátních genů“ nebo (jak je to nepochybně míněno) „sekvenování nové generace“. Navíc použití termínu „zodpovědných“ není v tomto případě na místě (u mužů, resp. člověka, se neprokázalo, že by tyto geny byly skutečně zodpovědné za poruchu spermatogeneze a neplodnost).

Splnění cílů práce a celkové hodnocení:

Cíle práce, tak, jak byly vytyčeny v úvodu, nemohu považovat za splněné – především z důvodu nevhodného experimentálního designu (viz výše). Je samozřejmě otázkou, nakolik je toto chybou diplomantky a nakolik jde spíše o selhání zodpovědných pracovišť. Nepochybuji o tom, že data, které autorka získala, budou v kombinaci s dalšími výsledky školitelského pracoviště v budoucnu podkladem pro solidní vědeckou publikaci. Mé vyjádření o celkovém hodnocení práce v tomto ohledu závisí na tom, jaké jsou v oboru Genetika, molekulární biologie a virologie, akreditovaném v rámci navazujícího magisterského programu Biologie na Přírodovědecké fakultě UK v Praze nároky na celkové pojetí diplomové práce (viz výše uvedený komentář k Experimentální části práce). I když však od tohoto odhlédnu, nedostatečná dokumentace výsledků je podle mého názoru velmi závažným problémem a spolu s dalšími výše zmíněnými nedostatky výrazně snižuje kvalitu práce. O tom, zda tato diplomová práce v podobě, v jaké byla předložena k obhajobě, má takovou úroveň, aby mohla být podkladem pro získání magisterského titulu, musí samozřejmě rozhodnout příslušná komise.

Otázky a připomínky oponenta:

V této části uvádím již jen drobnější připomínky a dotazy, u nichž nepochybuji, že se k nim autorka při obhajobě práce uspokojivě vyjádří.

Dotazy:

- 1) jaká modelová zvířata se dnes používají při studiu poruch reprodukce kromě myši (str. 16)?
- 2) na jakém mechanismu je založena asociace mutací v genu *CDC42* s poruchami reprodukce? Jde přímo o poruchy spermatogeneze (dtto u genu *CDC14B*)?
- 3) co je to globozoospermie (str. 19) – v Tab. 2.1. není uvedeno. Jde rovněž o klasifikaci dle norem WHO?
- 4) na jakých modelových zvířatech byla studována transkripce a translace genu *ODF1*, nezbytnost genu *RIMBP3* pro normální morfologii spermií a plodnost samců a vliv genu *SPATA16* na spermatogenezi (str. 21)?

- 5) jaké bylo složení TBE pufu (str. 33)?
- 6) formulace „Přečištění probíhalo na kolonce s membránou, na kterou se váže DNA/cDNA a po přečištění se uvolnila elučním pufrem“ ve mně navozuje dojem, že se uvolnila membrána. Je tomu skutečně tak?
- 7) jestliže byly u 12 pacientů, kteří byli posléze z hodnoceného souboru vyřazeni, výsledky rutinního genetického vyšetření (které byly důvodem jejich vyřazení) známy již předem, proč vůbec byli do sekvenování zařazeni?
- 8) z čeho autorka usuzuje, že „ztrátová alela je úplně recesivní“ (a že je vůbec ztrátová) (str. 51)?

Připomínky:

- 1) na obr. 2.2. (str. 15) jsou do češtiny přeloženy jen některé popisky
- 2) dATP, dGTP atp. nejsou nukleotidy, ale nukleosidtrifosfáty, nukleotid také dost dobře nemůže uvolnit pyrofosfát (str. 22, 23, 24)
- 3) zkratka DTT (Tab. 28) chybí v seznamu zkratek
- 4) u některých centrifugačních údajů nejsou uvedeny časy / g (str. 36, 37), jinde je uváděn údaj RCF (mělo by to být sjednoceno, tedy místo RCF dát \times g) (str. 42, 43, 44)
- 5) Tab. 5.3. a Obr. 5.2 v kapitole Výsledky znázorňují totéž, jedná se tedy o duplikaci výsledků
- 6) Tab. 5.5 (str. 51) uvádí jen heterozygotní pacienty, vysvětlivky pod čarou jsou tedy zbytečné
- 7) Tvzení, že 4 studie citované v Diskusi na str. 53 jsou „podobné jako v našem případě“ neplatí – zmínění autoři na rozdíl od diplomantky správně porovnávali soubory pacientů s kontrolními soubory.

Návrh hodnocení oponenta (známka nebude součástí zveřejněných informací)

výborně velmi dobře dobře nevyhověl(a)

Podpis oponenta:

Dana Holá, v Praze 29.5.2015