

Posudek na dizertační práci

Mgr. et Mgr. Anny Přistoupilové

Využití nových metod analýzy genomu ve studiu molekulární podstaty vzácných geneticky podmíněných onemocnění

Předkládaná práce má 59 stran plus seznam použité literatury a jako přílohu osm publikací v recenzovaných časopisech, z toho dvě se sdíleným spoluautorstvím a jednu se díleným spoluautorstvím v procesu publikace. Autorka uvádí dalších 16 publikací, na kterých se podílela, ale které nejsou součástí dizertační práce. Odborné publikace, na kterých se podílela již nasbíraly takřka 500 citací. To vše je úctyhodný přehled odvedené práce a svědčí o tom, že Anna je pro svůj tým platným spolupracovníkem.

Práce má klasické rozdělení kapitol, kde samozřejmě hlavní část tvoří publikace v recenzovaných časopisech. Přehled sekvenčních technik a jejich vývoje je velmi dobře napsaný, přitom přehledný a čtivý - doporučil bych ho i jako studijní materiál studentům biomedicínským oborů. Výsledky shrnují všechny publikace a v komentářích doplňují, na co v článcích nebylo místo. Celkově má tato dizertační práce vysokou úroveň.

Prezentovaná práce je dokladem autorčiny schopnosti pracovat v týmu na komplexních problémech současné medicíny. Použitá metodika zahrnuje nejenom rutinní používání nejnovějších bioinformatických postupů a softwaru, ale i tvorbu softwaru vlastního, připraveného na míru jednotlivým projektům. Výsledků je hodně a mají jednoznačně mezinárodní úroveň. Definované cíle dizertační práce byly splněny. Během čtení jsem nenarazil na žádné známky plagiátorství.

Předloženou dizertační práci s potěšením doporučuji k obhajobě.

V Praze 20. 5. 2020

Jan Pačes, Ph.D.

Ústav molekulární genetiky AV ČR

Vídeňská 1083, Praha

Otázky k obhajobě:

V kapitole 7 je jednou větou poměrně obecně popsán autorčin podíl na výsledcích - v podstatě se podílela na všech aspektech práce od zavádění metod, bioinformatické analýze, interpretaci dat i psaní publikací. U čtyř z prací jsem našel popis autorčina konkrétního příspěvku, u ostatních chybí. Protože se jedná vesměs o práce s velkým množstvím autorů a v případě prvoautorství jde o prvoautorství sdílené, bylo by žádoucí popsat podíl na výsledcích podrobněji. Jako oponent mám tu výhodu, že práci Anny znám a nepochybuji nebo přímo vím, že její přínos ke všem publikacím byl významný, nicméně rád bych si při obhajobě udělal detailnější obrázek. Jedná se o obecný problém současného biomedicínského výzkumu, na kterém se podílí velké týmy a při množství autorů je jejich podíl na práci nejasný.

Quality score: během vývoje jednotlivých technik docházelo ke změnám systému kódování kvalit jednotlivých sekvencí uvnitř formátu FASTQ a bohužel i jejich výpočtu. Z vlastní praxe jsem narazil na špatně udělanou kontrolu kvality ze starších dat, protože nebyl vzat do úvahy vývoj onoho kódování. Zajímalo by mě, jaký software na kontrolu kvality dat používáte, zejména jeho schopnost vypořádat se se zmíněným problémem.

PacBio vs Oxford Nanopore: cílem práce č. 3 bylo mimo jiné zavedení technik sekvenování třetí generace pro identifikaci variant. Nenalezl jsem nikde výsledky vycházející ze sekvenování pomocí technologie Oxford Nanopore, ačkoliv je v souhrnu výsledků tato technologie zmíněna. Zajímalo by mě totiž srovnání těchto dvou technologií právě pro identifikaci variant v souvislosti s jejich velkou chybovostí. Jaký software na čištění a opravu těchto sekvencí používáte a s jakým úspěchem?

Součástí analýz je i vlastní vyvinutá sada programů pro zpracování sekvencí třetí generace. Jsou programy open source a někde k dispozici? Jestliže ne, plánujete to?

