

## Posudek oponenta bakalářské práce

Jméno a příjmení uchazeče/ky: Lucie Korená

Název práce: Tvorba a hodnocení kvality genomových assembly

### A. Bodové hodnocení jednotlivých aspektů práce (označte právě jednu z možností)

1. Rozsah BP a její členění	
	A - přiměřené, odpovídají charakteru BP a významu jednotlivých částí
<input checked="" type="checkbox"/>	B - nevyrovnané, členění není logické n. rozsah jednotlivých částí nekoresponduje s jejich významem
	C - uspokojivé, rozsah některých částí nedostačuje
	<b>N - nedostatečné</b>

2. Odborná správnost	
	A - výborná, bez závažnějších připomínek
<input checked="" type="checkbox"/>	B - velmi dobrá, s ojedinělými drobnými závadami (nejasnost výkladu, chyby ve vzorcích nebo chemických názvech, nedokonalý popis metod nebo výsledků)
	C - uspokojivá, s četnějšími drobnými závadami
	<b>N - nevyhovující, s hrubými chybami</b>

3. Uvedení použitých literárních a j. zdrojů	
<input checked="" type="checkbox"/>	A - bez připomínek, všechny převzaté údaje s citací zdroje, celkový počet citací odpovídá charakteru práce
	B - uspokojivé, s občasnými neobratnostmi zejm. v umístění odkazů, nebo s celkově nižším počtem citací
	C - s vážnějšími závadami, např. převažují "nestandardní" odkazy na učebnice, přednášky, webové stránky, nebo se ojediněle vyskytuje opominutí odkazu na zdroj převzatých dat
	<b>N - nevyhovující, velmi málo citací, ev. rysy plagiátu (časté opomíjení odkazu na zdroj převzatých dat, popř. opsání velkých částí textu)</b>

4. Jazyk práce	
	A - výborný, práce je napsána čtivě a srozumitelně, bez závažnějších gramatických n. pravopisných chyb
<input checked="" type="checkbox"/>	B - velmi dobrý, ojedinělé stylistické neobratnosti, gramatické n. pravopisné chyby
	C - uspokojivý, četnější slohové neobratnosti, gramatické n. pravopisné chyby, ojediněle se vyskytují obtížně srozumitelné n. nejednoznačné formulace
	<b>N - nevyhovující, s četnými hrubými chybami</b>

5. Formální a grafická úroveň práce	
	A - výborná, bez překlepů a chyb ve formátování
<input checked="" type="checkbox"/>	B - velmi dobrá, ojedinělé chyby formátu citací, překlepy, chybějící zkratky apod.
	C - uspokojivá, s ojedinělými většími (např. vynechání stránky) nebo četnějšími drobnými chybami
	<b>N - nevyhovující, s četnými hrubými chybami</b>

Případný slovní komentář k bodům 1. až 5.:

Téma předložené bakalářské práce se zabývá velmi zajímavým bioinformatickým tématem, kterým je snaha o sestavení co nejlepšího genomu (genomové assembly). Jedná se o problematiku, o které již bylo popsáno mnoho stran a jejíž shrnutí do jedné krátké bakalářské práce se může stát výzvou. Studentka se s tím poprala poměrně úspěšně, kdy na 30 stranách popisuje vybrané metody vedoucí ke genomovému assembly a následnému statistickému vyhodnocení, včetně vlastní praktické práce. Bohužel se čtenář místy neubrání pocitu, že některé kapitoly jsou popsány velmi povrchně a některé stěžejní myšlenky by mohly být rozvedeny více do hloubky, aby to celé dávalo větší smysl. Dále zamrzí totální absence diskuze, která by vzhledem k popisovaným algoritmům mohla být velmi zajímavá.

Práce samotná je čtivá, logicky členěná, s minimem gramatických a typografických chyb, občas s výskytem ne zcela jasných, zavádějících a chybných formulací a špatných překladů ustálených slovních spojení z angličtiny. Avšak jako celek práci hodnotím kladně.

Obsahuje-li práce i vlastní výsledky uchazeče/ky (nejsou povinnou součástí práce), pak prosíme o Vaše stanovisko k následujícímu:

Jsou řádně stanoveny a vysvětleny cíle experiment? **ANO**

Je množství experiment adekvátní k cílům? **ANO**, ale programů vybraných k porovnání by přeci jen mohlo být více (použity pouze Abyss a Masurca)

Je dokumentace výsledků dostačující? **ANO**, vzhledem k definovaným cílům

Jsou výsledky diskutovány a zasazeny do kontextu existující literatury? **NE**

Případný další slovní komentář k výsledkům autora:

Cílem praktické části bylo porovnat dvě genomové assembly vyprodukované dvěma programy (Abyss vs Masurca). Závěrem praktické části je víceméně pouze konstatování, že Masurca dovedla sestavit kvalitnější assembly, které i tak, na základně vstupních dat, není dost kvalitní. Diskuze zcela chybí.

## B. Obhajoba

### **Dotazy k obhajobě (povinná část posudku)**

Poprosil bych o vysvětlení pojmu (str. 2) “Nanooptická” detekce signálu při sekvenování platformou Oxford Nanopore.

Studentka popisuje programy Abyss (pro krátká Illumina čtení) a Masurca (zvládá i hybridní assembly). Mohla by studentka vybrat jeden assembler čistě pro dlouhá čtení a více do hloubky jej popsat? V podobném duchu, jak popisuje výše zmíněné dva programy v kapitole 3.3.

V praktické části studentka provedla assembly genomu *T. szidati* s raw daty deponovanými v databázi SRA. Jak by studentka vysvětlila, že ani pro jeden program vstupní data nijak neupravila? Bylo by vhodné vůbec používat upravená vstupní data (ready) např. pro program Masurca? Dá se očekávat, že úprava dat bude mít pozitivní vliv na výsledné assembly?

Stanovisko k opravě chyb v práci:

opravný lístek/oprava v textu **JE** / **NENÍ** (zakroužkujte) podmínkou přijetí práce

**C. Celkový návrh**

Práci **doporučuji** k přijetí k dalšímu řízení: **ANO** / **NE**

Navrhovaná celková klasifikace **B**

Datum vypracování posudku: **24. 6. 2020**

Jméno a příjmení, podpis oponenta (SIS): **Mgr. Jiří Vorel**

