

I. OSOBNÍ A STUDIJNÍ ÚDAJE

Příjmení: **Bůžková** Jméno: **Veronika** Osobní číslo: **500001**
 Fakulta: **Fakulta biomedicínského inženýrství**
 Studijní program: **Informatika a kybernetika ve zdravotnictví**
 Název práce: **Zpracování, uložení a agregace genetických variant**

II. HODNOCENÍ BAKALÁŘSKÉ PRÁCE

Kritéria hodnocení práce		Počet bodů
1.	<p>Splnění cíle a vhodnost struktury obsahu bakalářské práce z hlediska zadaného tématu (splnění zadání). (0 – 30)*</p> <p>Každá část či věta ze zadání musí mít jasný odraz ve zpracované práci. Excelentně splněné zadání může být ohodnoceno maximálním počtem bodů. V poměru rozsahu části v zadání, která není zcela vhodně či úplně zpracována, se hodnocení odpovídajícím způsobem snižuje.</p>	29
2.	<p>Teoretická úroveň a využití dostupné literatury v bakalářské práci. (0 – 30)*</p> <p>Oponent posuzuje relevantnost teoretické části k zadání, rozsah rešerší a systematické uspořádání zjištěných poznatků. Pokud převažuje doslovné převzetí textů, snižuje oponent hodnocení až o 15 bodů (přirozeně za předpokladu dodržení autorských práv). Důvodem pro snížení celkového hodnocení je dále nedostatečný výběr teoretických poznatků, literatury a zdrojů.</p>	25
3.	<p>Rozsah realizačních prací (SW, HW), aplikovaných vědomostí a znalostí, úroveň metodologického zpracování a závěrů práce. (0 – 30)*</p> <p>Maximální počet bodů lze udělit práci, která je vhodná k publikování. Tento aspekt se posuzuje zejména z hlediska významu pro obohacení teoretických poznatků a má praktický význam. Obzvláště pozitivně je hodnoceno vytvoření modelu, SW produktu a též technická realizace. Za drobné metodologické nedostatky se hodnocení snižuje až o 5 bodů. Nekonzistentnost zpracování s teoretickými východiskami a nejasný či ne zcela odborný metodologický přístup vede ke snížení minimálně o 15 bodů. Další snížení hodnocení lze udělit za nedostatečnou diskusi k závěrům. Celkem 30 bodů za velmi komplexní a bezchybnou práci včetně dalších aktivit jako je účast na vědecko-výzkumném projektu či grantu, aktivní účast na tvorbě publikací, patentů či užitečných vzorů.</p>	27
4.	<p>Formální náležitosti a úprava bakalářské práce (úroveň psaní, označení struktury textu, grafy, tabulky, citace v textu, seznam použité literatury apod.). (0 – 10)*</p> <p>Oponent hodnotí formální náležitosti z pohledu dodržení pravidel o psaní, atributů závěrečných prací, tj. formátování textu, struktury práce, seznamu použité literatury, vybavenosti bakalářské práce grafy a tabulkami, způsobu citování. Za nedodržení jednotlivých pravidel snižuje maximální hodnocení o 2 body za každý nerespektovaný atribut. Rovněž za výskyt gramatických chyb, překlepů a nevhodné stylistiky a terminologie se snižuje hodnocení o 2-4 body. V práci by se měla objevovat pouze standardní odborná terminologie a to zejména v českém jazyce (je třeba hodnotit schopnost vyjadřovat se technickým jazykem – 2 body), grafy jsou tvořeny podle zásad (viz tolerance a vliv statistického zpracování – 2 body), u grafů a tabulek jsou patřičné legendy a vše je čitelné (2 body), jsou dodržena citační pravidla podle ISO690 a ISO690-2 (2 body).</p>	4
5.	Celkový počet bodů	85

* Slovní hodnocení uveďte v komentáři.

III. NÁVRH OTÁZEK K OBHAJOBĚ

1. Cílem práce bylo vytvořit řešení umožňující v budoucnu tvorbu české populační databáze. Jaké plány máte v této oblasti a jaké vzorky a data bude tato databáze obsahovat?
2. Jaké české populační databáze už jsou dostupné a jaké další iniciativy v současnosti běží a jak se liší od té vaší? Máte v plánu agregovat data ze všech těchto iniciativ do jedné databáze?
3. Jaké vidíte technické možnosti implementace vaší databáze do webového prostředí, aby databáze byla dostupná veřejnosti?

IV. CELKOVÉ HODNOCENÍ ÚROVNĚ VYPRACOVÁNÍ BAKALÁŘSKÉ PRÁCE

Hodnocení**:	A (výborně)	B (velmi dobře)	C (dobře)	D (uspokojivě)	E (dostatečně)	F (nedostatečně)
Počet bodů:	100 - 90	89 - 80	79 - 70	69 - 60	59 - 50	< 50
	<input type="checkbox"/>	X	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

** v případě hodnocení F (nedostatečně) uveďte podrobný komentář

Bakalářskou práci hodnotím výše uvedeným klasifikačním stupněm a doporučuji/nedoporučuji k obhajobě.

V. KOMENTÁŘ

Cílem předkládané bakalářské práce bylo navrhnout metody, vybrat technologie a vytvořit skripty pro vytvoření databáze agregovaných anotovaných genetických variant pro daný panel vzorků. Jedná se o velmi potřebné, aktuální a zároveň složité, multioborové téma. Pro jeho zpracování bylo nutné pochopit a osvojit si základy molekulární biologie, nových metod sekvenace genomu, bioinformatiky a práce s big daty. Autorka tento úkol zvládla velmi dobře a splnila všechny zadané cíle. Předkládané řešení využívá moderních technologií (Hail, Apache Spark, VEP), které umožní jeho využití pro velké množství vzorků a je zároveň přenositelné a reprodukovatelné díky technologii Docker. Velmi oceňuji, že předkládané řešení je dostupné přes GitLab, úložiště umožňující verzování kódu. Možným vylepšením by bylo uvést odkaz na GitLab přímo v textu bakalářské práce a doplnit README.md o instrukce, jak kód spustit (viz. Kapitola 5 - Implementace).

Kvalitu práce bohužel snižuje velké množství gramatických chyb, překlepů, nevhodné stylistiky, terminologie (např. typ sekvenování není určitě fenotypem, místo "targetované sekvenování" se používá v češtině spíše termín "cílené sekvenování", místo "homozygot na referenci" se používá termín "referenční homozygot"). Dále práce obsahuje drobné nepřesnosti v uvedených citacích (např. u citace [11] chybí autor knihy, u citací [1] a [4] struktura neodpovídá požadované citační normě, na str. 27 je místo reference pouze nefunkční odkaz "Chyba! Nenalezen zdroj odkazů"). Diskuzi by bylo vhodné obohatit a doplnit například o porovnání předkládaného řešení s již existujícími populačními databázemi.

Celkově tuto bakalářskou práci hodnotím jako velmi zdařilou a doporučuji ji k obhajobě.

Jméno a příjmení: Mgr. et Mgr. Anna Přistoupilová, Ph.D.
Organizace: DNAnexus Czech Republic s.r.o.
Kontaktní adresa:

Podpis:

Datum: