



Jakub Godlewski Ph.D.
Chair of the Department of Neurooncology
Mossakowski Medical Research Institute
Polish Academy of Sciences
02-106 Warsaw 5 Pawińskiego Street
jgodlewski@imdik.pan.pl

RE: Recenzja rozprawy doktorskiej

Warszawa 17/03/2025

Doktorant: mgr. inż. Paweł Segit

.....

Tytuł rozprawy doktorskiej:

Analiza profili transkryptomicznych pojedynczych komórek mieloidalnych w mikrośrodowisku glejaków.

.....

Promotorzy: Prof. dr hab. Bożena Kamińska-Kaczmarek

Dr hab. Jakub Mieczkowski

.....

Recenzent: Dr hab. n. med. Jakub Godlewski

.....

Ocena ogólna i zakres pracy

Praca doktoranta, mgr inż. Pawła Segita wyróżnia się szczegółowym podejściem do analizy danych scRNA-seq w badaniach nad glejakami, z naciskiem na opracowanie i optymalizację metodyki bio-informatycznej. W przeciwieństwie do tradycyjnych analiz przed-klinicznych, autor skupił się na technologicznym i analitycznym aspekcie badania, co stanowi duże wyzwanie ze względu na ogromną złożoność oraz wielowymiarowość generowanych danych.

Wyzwania analityczne i osiągnięcia

Realizacja projektu wymagała opanowania zaawansowanych narzędzi analitycznych oraz precyzyjnego dostrojenia parametrów, co jest niebagatelnym wyzwaniem przy analizie dużych zbiorów danych scRNA-seq. Doktorant wykazał się nie tylko umiejętnością obsługi nowoczesnych narzędzi bio-informatycznych, ale także zdolnością krytycznej oceny zastosowanych metod, co przejawia się w szczegółowej analizie wpływu różnych parametrów na ostateczną segmentację komórek.

Wartość naukowa pracy

Rozprawa wnosi istotny wkład do badań nad heterogenicznością komórek układu odpornościowego w mikrośrodowisku glejaków. Opracowanie zoptymalizowanej metodyki analizy i wizualizacji danych scRNA-seq



Jakub Godlewski Ph.D.
Chair of the Department of Neurooncology
Mossakowski Medical Research Institute
Polish Academy of Sciences
02-106 Warsaw 5 Pawińskiego Street
jgodlewski@imdik.pan.pl

stanowi nowatorskie podejście, umożliwiające precyzyjne rozróżnienie profili transkryptomicznych subpopulacji komórek mieloidalnych w mikrośrodowisku ludzkich guzów glejaków wielopostaciowych. Dzięki zastosowaniu nowoczesnych technologii oraz kompleksowej analizie danych zarówno z modeli doświadczalnych, jak i próbek ludzkich, praca ta otwiera nowe perspektywy w badaniach nad patogenezą glejaków oraz potencjalnymi celami terapeutycznymi.

Rozprawa została oparta na solidnych podstawach teoretycznych oraz aktualnej literaturze, co umożliwia logiczne i systematyczne przedstawienie zagadnienia. Doktorant wykazał się umiejętnością adaptacji zaawansowanych metod analitycznych do specyficznych problemów badawczych, takich jak analiza danych scRNA-seq z komórek mieloidalnych. Szczególnie cenne jest porównanie profili transkryptomicznych komórek układu odpornościowego w mózgach myszy (zarówno w stanie naiwnym, jak i po operacji pozorowanej) oraz w próbkach z ludzkich guzów, co znacząco przyczynia się do lepszego zrozumienia różnic pomiędzy modelami eksperymentalnymi a stanem klinicznym.

Uwagi metodologiczne i krytyczne

Praca wyróżnia się starannym doбором parametrów analizy, co gwarantuje optymalne wyniki wizualizacji i interpretacji danych. Niemniej jednak, cel pracy mógłby być sformułowany w sposób jeszcze bardziej precyzyjny, aby wyraźniej ukazać naukowy oraz potencjalny kliniczny sens badań. W niektórych partiach rozprawy przydałaby się głębsza dyskusja nad ograniczeniami zastosowanych metod oraz możliwością ich dalszej optymalizacji. Mimo to, przeprowadzone analizy oraz porównania między różnymi modelami badawczymi świadczą o wysokiej jakości podejściu bio-informatycznym autora. Szczegółowa analiza metodologii obejmuje następujące zagadnienia:

- Przetwarzanie danych:

Praca obejmuje kompleksowy „pipeline”, rozpoczynający się od konwersji plików BCL do formatu FASTQ, poprzez mapowanie odczytów do transkryptomu referencyjnego, aż po stworzenie macierzy zliczeń. Autor szczegółowo opisuje etapy kontroli jakości, normalizacji oraz integracji danych z różnych eksperymentów, co stanowi fundament dla wiarygodnej analizy dalszych etapów badania.

- Redukcja wymiarów i grupowanie komórek:

Kluczowym elementem pracy jest optymalizacja parametrów redukcji wymiarów (np. PCA, t-SNE czy UMAP) oraz precyzyjny dobór algorytmów grupowania. Autor dokładnie analizuje wpływ zmiany parametrów (takich jak rozdzielczość klastrowania) na ostateczny dobór grup komórkowych. Wykazana została umiejętność identyfikacji subtelnych różnic transkryptomicznych, co pozwala na wyodrębnienie biologicznie uzasadnionych subpopulacji komórek mieloidalnych.

- Analiza wariacji i interpretacja wyników:

Praca podejmuje trudny problem interpretacji zmienności ekspresji genów między różnymi grupami komórek. Autor przeprowadza dogłębną analizę różnicowej ekspresji, uwzględniając zarówno czynniki techniczne, jak i biologiczne,



Jakub Godlewski Ph.D.
Chair of the Department of Neurooncology
Mossakowski Medical Research Institute
Polish Academy of Sciences
02-106 Warsaw 5 Pawińskiego Street
jgodlewski@imdik.pan.pl

co umożliwia lepsze zrozumienie heterogeniczności mikrośrodowiska glejaka. Wyniki dotyczące doboru modelu klastrowania oraz analizy wariacji stanowią cenny wkład do dziedziny bio-informatyki, pokazując, jak drobne modyfikacje parametrów analitycznych mogą wpływać na interpretację danych.

Rekomendacje i potencjalny rozwój projektu

Wdrożenie interaktywnej aplikacji do eksploracji danych znacząco zwiększyłoby potencjał wykorzystania uzyskanych wyników, co znajduje potwierdzenie w najnowszych publikacjach w *Nature* i *Cell*. Praca opublikowana ostatnio w *Nature* (PMID: 40011771) wskazuje, że nawet subtelne różnice w programach transkryptomicznych komórek mieloidalnych mają kluczowe znaczenie dla funkcjonowania mikrośrodowiska glejaka, podkreślając tym samym wartość dokładnej analizy na poziomie pojedynczych komórek. Z kolei niedawna publikacja w *Immunity* (PMID: 37451265) ukazuje, jak złożone interakcje komórkowe w specyficznych, niedotlenionych niszach guza wpływają na re-programowanie komórek odpornościowych, co jeszcze bardziej uwypukla potrzebę zastosowania zaawansowanych narzędzi analitycznych. Interaktywna aplikacja umożliwiłaby dynamiczne wizualizowanie heterogenicznych klastrow, modyfikację parametrów analizy w czasie rzeczywistym oraz szybkie testowanie hipotez dotyczących mechanizmów immunosupresji. Taka platforma stałaby się nieocenionym narzędziem, które wspierałoby dalsze badania translacyjne, pozwalając naukowcom na głębsze zrozumienie złożoności patologii glejaka oraz przyczyniając się do szybszego wdrażania innowacyjnych strategii terapeutycznych.

Język i forma

Praca charakteryzuje się klarownym stylem oraz logiczną strukturą, co ułatwia odbiór przedstawionych zagadnień. Opisy metodologii i wyników są szczegółowe, a stosowane narzędzia analityczne zostały przedstawione w sposób przystępny, co świadczy o wysokim poziomie kompetencji autora w dziedzinie bio-informatyki.

Podsumowanie

Rozprawa doktorska mgr inż. Pawła Segita to wyróżniający się przykład połączenia zaawansowanej bio-informatyki z nowoczesnymi metodami analizy danych scRNA-seq. Autor z dużą precyzją i umiejętnością optymalizuje każdy etap przetwarzania danych, skupiając się na krytycznych aspektach, takich jak redukcja wymiarów, grupowanie komórek i analiza wariacji transkryptomicznych. Pomimo ogromnych wyzwań analitycznych, praca prezentuje solidne wyniki, które mogą stanowić podstawę do dalszych badań. Wdrożenie interaktywnej aplikacji do eksploracji danych znacząco zwiększyłoby potencjał wykorzystania uzyskanych wyników, co dodatkowo podkreśla wartość naukową projektu.

Reasumując, stwierdzam ze recenzowana rozprawa doktorska mgr inż. Pawła Segita pt. "Analiza profili transkryptomicznych pojedynczych komórek mieloidalnych w mikrośrodowisku glejaków" spełnia warunki określone w art. 187 Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2023 r. poz. 742



Jakub Godlewski Ph.D.
Chair of the Department of Neurooncology
Mossakowski Medical Research Institute
Polish Academy of Sciences
02-106 Warsaw 5 Pawińskiego Street
jgodlewski@imdik.pan.pl

z późn. zm.). W związku z powyższym, wnioskuję do Rady Naukowej Instytutu Biologii Doświadczalnej o dopuszczenie mgr. inż. Pawła Segita do dalszych etapów postępowania w sprawie nadania stopnia doktora.

Wnioski końcowe

Ponadto uważam, że całość analizy badań eksperymentalnych przeprowadzonych przez doktoranta stanowi pionierskie osiągnięcie na skalę krajową. Zastosowanie zaawansowanych technik scRNA-seq do detalicznego opisu mikrośrodowiska nowotworowego glejaka nie tylko wpisuje się w światowe trendy w badaniach nad nowotworami, ale również wyróżnia się na tle innych prac w tym obszarze. Aplikacyjność wyników tych badań może znacząco wpłynąć na rozwój nowych technologii badawczych oraz metod diagnostycznych i terapeutycznych, umożliwiając bardziej precyzyjne podejście do leczenia nowotworów mózgu. W mojej opinii, przedstawiona praca zasługuje na wyróżnienie ze względu na jej innowacyjność, głębię analizy oraz potencjalny wpływ na rozwój nowoczesnej onkologii.

Jakub Godlewski, Ph.D.



Wrocław, dn. 25 marca 2025 r.

Recenzja pracy doktorskiej mgr. inż. Pawła Segita pt. „Analiza profili transkryptomycznych pojedynczych komórek mieloidalnych w mikrośrodowisku glejaków”, zrealizowanej w Pracowni Neurobiologii Molekularnej Instytutu Biologii Doświadczalnej im. M. Nenckiego Polskiej Akademii Nauk pod kierunkiem Prof. dr hab. Bożeny Kamińskiej-Kaczmarek i dr. hab. Jakuba Mieczkowskiego

Recenzowana rozprawa doktorska liczy 96 stron i składa się z: wstępu (18 stron), poprzedzonego spisem treści, wykazem skrótów oraz streszczeniami w języku polskim i angielskim; następnie celu pracy (1 strona); materiałów i metod (11 stron); wyników (34 strony); dyskusji (11 stron); podsumowania i wniosków (1 strona); piśmiennictwa (7 stron) oraz wykazu publikacji (1 strona).

Wstęp stanowi przystępne i jednocześnie bardzo wartościowe merytorycznie wprowadzenie do tematyki analiz pojedynczych komórek za pomocą sekwencjonowania RNA i cytometrii przepływowej. Autor w sposób zwięzły i klarowny przedstawił zarówno metody nisko- i wysokoprzepustowe, jak i podejścia do wizualizacji danych, podejmując się trudnego zadania opisanie zastosowań matematycznych algorytmów w analizach próbek biologicznych. Kolejne podrozdziały wstępu w sposób wyczerpujący omawiają tematykę glejaków i ich mikrośrodowiska, uwzględniając aktualną klasyfikację oraz dane transkryptomiczne, co stanowi solidne uzasadnienie dla zaplanowanych badań. Dużym atutem jest szczegółowa charakterystyka populacji komórek układu odpornościowego obecnych w mikrośrodowisku glejaków, ze szczególnym uwzględnieniem komórek mieloidalnych oraz ich profilu transkryptomicznego i proteomicznego. Taki sposób ujęcia problematyki pozwala czytelnikowi dobrze zrozumieć znaczenie analizowanych danych oraz ich wpływ na dalsze wyniki pracy. Zabrakło mi tylko omówienia technik transkryptomiki przestrzennej, która zyskuje coraz większe znaczenie w badaniach tkanek i ich organizacji przestrzennej, sięgając już poziomu pojedynczych komórek.

Cel główny pracy został jasno sformułowany i dotyczył optymalizacji analizy oraz wizualizacji danych pochodzących z sekwencjonowania RNA z pojedynczych komórek mieloidalnych, wyizolowanych

Strona 1 z 6

z mikrośrodowiska glejaków w modelu mysim oraz z materiału klinicznego pozyskanego od pacjentów z glejakiem wielopostaciowym. W ramach czterech celów szczegółowych autor pracy skoncentrował się oddzielnie na materiale mysim i ludzkim, dążąc do oceny wpływu różnych parametrów analitycznych na wyniki oraz ich wizualną interpretację. Jeden z celów obejmował również analizę wpływu operacji neurochirurgicznej na profil transkryptomyczny mózgowych komórek mieloidalnych oraz ocenę zasadności stosowania nieoperowanych myszy jako grupy kontrolnej dla myszy z implantowanymi komórkami glejaka.

W rozdziale **Materiały i metody** Doktorant szczegółowo przedstawił kluczowe etapy analizy danych transkryptomicznych, co znacząco ułatwiło zrozumienie wyników zaprezentowanych w dalszej części pracy. Opis ten świadczy o bardzo dobrej znajomości stosowanych narzędzi bioinformatycznych i doskonałym przygotowaniu merytorycznym autora. Rozdział może również pełnić funkcję praktycznego przewodnika dla innych badaczy rozpoczynających pracę z analizą danych typu single-cell RNA-seq.

Wyniki zostały przedstawione w sposób systematyczny i przejrzysty. W treść rozdziału wkomponowano liczne wykresy, mapy cieplne, tabele i grafy, które ilustrują kolejne etapy przeprowadzonych analiz. Zaprezentowane dane dokumentują krok po kroku proces przetwarzania wyników sekwencjonowania RNA z pojedynczych komórek pochodzących z różnych próbek. Dzięki zawartym wizualizacjom czytelnik może łatwo prześledzić wpływ poszczególnych parametrów analizy na identyfikację klastrów oraz końcową charakterystykę populacji komórek mieloidalnych mózgu. Autor pracy wykazał m.in., że sam zabieg neurochirurgiczny nie wpływa istotnie na ekspresję genów w tej populacji komórek, co uzasadnia zastosowanie nieoperowanych myszy jako grupy kontrolnej wobec myszy z wszczepionymi komórkami nowotworowymi. W przypadku ludzkich tkanek glejaka, autor wyodrębnił 15 klastrów komórek, obejmujących mikroglej i makrofagi o zróżnicowanych profilach transkryptomicznych, co umożliwiło przypisanie im potencjalnych funkcji pełnionych w mikrośrodowisku guza. Należy podkreślić, że doktorant poświęcił należytą uwagę elementom kontroli jakości analizy, takim jak jakość próbek, parametry sekwencjonowania czy udział poszczególnych próbek w całkowitej puli analizowanych komórek, co sprzyja ograniczeniu liczby artefaktów w wynikach. Z edukacyjnego punktu widzenia pewien niedosyt budzi brak omówienia wpływu wybranych parametrów analitycznych,

Strona 2 z 6

zastosowanych w niektórych funkcjach przy przetwarzaniu danych mysich i ludzkich, np. funkcje FindNeighbors i FindAllMarkers, które miały różne wartości zależnie od źródła tkanek. Jaki wpływ na wyniki analizy mają te parametry i dlaczego zdecydowano się na ich zróżnicowanie pomiędzy tkankami mysimi i ludzkimi? Uzupełnienie tych informacji pozwoliłoby lepiej zrozumieć proces decyzyjny w zakresie doboru ustawień analitycznych.

W **Dyskusji** Doktorant podsumował uzyskane wyniki zarówno od strony metodologicznej, jak i pod względem interpretacji biologicznej. Poprawnie zwrócono uwagę na ograniczenia metodologiczne analizy transkryptomicznej pojedynczych komórek, co pokazuje dojrzałość naukową autora pracy. Zabrakło mi jednak odniesienia się do sposobu prowadzenia analiz i detekcji klastrów przez innych badaczy, których artykuły były cytowane w pracy.

Rozprawę kończy pięć wniosków, które odnoszą się do wcześniej sformułowanych celów pracy. Choć są one trafne, mogłyby zostać przedstawione w sposób bardziej szczegółowy, uwzględniający kluczowe przesłania oraz znaczenie przeprowadzonych badań dla rozwoju nauki.

Ostatni rozdział – **Piśmiennictwo** – obejmuje 79 pozycji literaturowych, stanowiących wartościowy i dobrze dobrany zbiór publikacji z zakresu tematyki rozprawy. Zgromadzone pozycje pozwalają na zapoznanie się z aktualnym stanem wiedzy w wybranej dziedzinie oraz umożliwiają porównanie wyników uzyskanych przez Doktoranta z dotychczasowymi osiągnięciami w omawianym obszarze badawczym.

W mojej opinii praca doktorska mgr inż. Pawła Segita świadczy o bardzo dobrej ogólnej wiedzy teoretycznej Doktoranta w dyscyplinie nauk biologicznych. Doktorant wykazał się umiejętnością samodzielnego prowadzenia pracy naukowej a sama rozprawa stanowi oryginalne osiągnięcie naukowe, prezentujące sposób analizy danych transkryptomicznych komórek mikrośrodowiska glejaka wielopostaciowego i dostarczające nową wiedzę na temat zaangażowania komórek mieloidalnych w procesy rozrostowe tego najbardziej agresywnego typu guzów mózgu. Oprócz dużej wartości naukowej pracy, rozprawa od strony metodologicznej jest doskonałym wprowadzeniem do analiz single-cell RNA-seq dla innych naukowców, planujących badania w tym obszarze.



Poniższe pytania i uwagi mają charakter redakcyjny i stymulujący dyskusję, i nie wpływają na ocenę końcową pracy:

1. Str. 14. Autor stwierdza, że „otrzymany materiał jest amplifikowany przy użyciu metody amplifikacji w czasie rzeczywistym - PCR (ang. polymerase chain reaction), aby po zakończeniu procesu uzyskać ilość materiału odpowiednią do wykonania sekwencjonowania NGS.” Czy na pewno chodzi o PCR w czasie rzeczywistym, czy raczej o klasyczną reakcję PCR?
2. Str. 22. Autor stwierdza, że „zidentyfikowano „sygnaturę genową” (ang. stemness signature), czyli zestaw genów o wysokiej i specyficznej ekspresji, która była najsilniejsza w komórkach podtypu proneuralnego i klasycznego GBM, a słabiej reprezentowana w podtypie mezenchymalnym”. Tłumaczenie „stemness signature” należałoby uściślić – sygnatura genowa komórek macierzystych.
3. Str. 25. Autor sugeruje, że naciek komórek do mózgu odbywa się poprzez barierę krew-mózg, która może ulec rozszczelnieniu w stanie patologicznym, np. w wyniku procesów zapalnych. Warto uściślić tę informację, że infiltracja komórek układu odpornościowego może zachodzić również przez układ limfatyczny mózgu.
4. W pracy przedstawiono opisy metod do wizualizacji danych za pomocą PCA, UMAP i t-SNE. Komórki z mikrośrodowiska mysich glejaków wizualizowano poprzez PCA/t-SNE/UMAP, ludzkie tkanki za pomocą PCA/UMAP. Jakie są zalety i ograniczenia t-SNE i UMAP? Z czego wynikają różnice w wyglądzie danych i innym rozkładzie populacji na Ryc. 6.11 (myszy kontrolne) i Ryc.12a (myszy naiwne)?
5. Ryc. 6.17 (prawy panel) ma prezentować dane dla komórek z wszystkich 5 guzów z kolorystycznym oznaczeniem informacji, z którego guza pochodziły analizowane komórki. Autor stwierdza, że „otrzymane wyniki wskazywały, że większość klastrów zawiera komórki pochodzące w zdecydowanej większości z jednej próbki”. Na przedstawionej rycinie widać różnorodność klastrów bez wyraźnej dominacji jeden z próbek, co z kolei jest do zauważenia przy podobnym „efekcie próbki” na Ryc. 6.18 (prawy panel). Czy Rycina 6.17 (prawy panel) jest prawidłowa?
6. Czy na podstawie uzyskanych danych można przypisać badanym ludzkim glejakom określony typ wg nowej klasyfikacji, opisanej w pracy (Park *et al.*, 2023)? Czy różnice pod względem płci obserwowane u myszy miały potwierdzenie w próbkach ludzkich?



7. W dyskusji zabrakło odniesień do danych przedstawionych w pracach Sankowski et al. 2019, Klemm et al., 2020 i Antunes et al. (2021). Czy można wyróżnić populacje CD49d⁺ i CD49d⁻ w uzyskanych przez autora danych? Czy podział na komórki CD11b⁺CD45^{low} i CD11b⁺CD45^{high} zaproponowany w pracy Gabrusiewicz et al. 2011 i pozwalający na rozróżnienie komórek mieloidalnych rezydentnych i naciekających znalazł potwierdzenie w pracy autora? Czy ekspresja genów dla CD49d i CD45 koreluje z poziomem białka i można identyfikować wymienione wcześniej populacje na poziomie mRNA?
8. W przypadku analiz ludzkich glejaków nie założono górnego limitu liczby genów a w mysich glejakach próg ten wynosił 3000 w celu odfiltrowania potencjalnych multipletów. Czy brak górnego progu w analizie ludzkich próbek może wpłynąć na wyniki analizy?
9. W analizach ludzkich glejaków autor wykazał obecność 15 klastrów w oparciu o ekspresję genów. Jaki był udział procentowy komórek z poszczególnych klastrów? Podobnie, jaki był udział komórek w fazach G1 i S/G2 w analizowanych próbkach, skoro faza cyklu komórkowego ma znaczenie przy analizie?
10. Czy procedura integracji (standardowej) danych w celu eliminacji „efektu próbki” tylko ulepsza analizę i powinna być stosowana jako procedura standardowa we wszystkich tego typu analizach?
11. Str. 69. Autor stwierdza: „W przypadku grupowania komórek w zbiorach scRNA-seq priorytetem jest właściwa interpretacja biologiczna uzyskanych populacji. Grupowanie, jako metoda nienadzorowana, podczas przypisywania komórek do klastrów nie uwzględnia zewnętrznych informacji (np. o znanych markerach komórek określonego typu) analizując jedynie profile transkryptomyczne pojedynczych komórek. Mając to na uwadze, zastosowano iteracyjne podejście do oceny wyników grupowania i weryfikacji uzyskanych klastrów, opierającej się na wiedzy biologicznej.” Czy widzi Pan tu jakiś potencjał do użycia sztucznej inteligencji jako pomoc w weryfikacji klastrów o znaczeniu biologicznym? Jak bardzo prawdopodobne jest to, że przy innej, może lepszej, metodzie analizy danych będziemy w stanie jednak wydzielić nowe klastry w jednorodnej grupie komórek i interpretować dane inaczej?

Podsumowując stwierdzam, że przedstawiona mi do recenzji rozprawa doktorska spełnia warunki określone w art.187 Ustawy z dnia

Strona 5 z 6



20 lipca 2018 r. - Prawo o szkolnictwie wyższym i nauce (Dz.U. z 2023 r. poz.742 z późn. zm.). W związku z powyższym, wnioskuję do Rady Naukowej Instytutu Biologii Doświadczalnej o dopuszczenie mgr. inż. Pawła Segita do dalszych etapów postępowania w sprawie nadania stopnia doktora.



Dr hab. Grzegorz Chodaczek

Grupa Badawcza Immunoterapii
Sieć Badawcza Łukasiewicz – PORT
Polski Ośrodek Rozwoju Technologii





UNIwersytet
Warszawski

Wydział Matematyki, Informatyki i Mechaniki

Instytut Informatyki

dr hab. Norbert Dojer, prof. UW

Warszawa, 25.04.2025.

Recenzja rozprawy doktorskiej *mgr inż. Pawła Segita*
zatytułowanej

*Analiza profili transkryptomycznych pojedynczych
komórek mieloidalnych w mikrośrodkowisku glejaków*

Rozprawa poświęcona jest analizie profili transkryptomycznych w subpopulacjach komórek w mikrośrodkowiskach glejaków. Przedmiotem badań są dane z sekwencjonowania RNA pojedynczych komórek w doświadczalnych glejakach mysich oraz w ludzkich guzach glejaka wielopostaciowego. W rozprawie postawiono cztery szczegółowe cele badawcze, spośród których dwa dotyczą przeprowadzenia analiz tych danych, zaś dwa pozostałe optymalizacji procesu analizy i wizualizacji.

Formalna strona rozprawy

Strukturę rozprawy można uznać za standardową. Składa się ona ze streszczeń w językach polskim i angielskim, listy skrótów, sześciu podstawowych rozdziałów oraz bibliografii i listy publikacji Autora rozprawy. Rozdział wstępny stanowi wprowadzenie w tematykę pracy i przybliża technologie badania transkryptomów pojedynczych komórek oraz zagadnienia związane z mikrośrodkowiskiem glejaków, w tym najistotniejsze wyniki uzyskane wysokoprzepustowymi metodami badań pojedynczych komórek. Na tym tle w następnym rozdziale zostały przedstawione wspomniane wcześniej cele rozprawy. Kolejny rozdział (Metody) poświęcony jest procedurze analizy danych z sekwencjonowania transkryptomów pojedynczych komórek. Przedstawione są kolejne etapy analizy, ich rola i zastosowane metody, a także wyjaśnione jest znaczenie poszczególnych parametrów. Następny rozdział (Wyniki) przedstawia dane eksperymentalne i przeprowadzone analizy. Podrozdział 6.1 poświęcony jest mysiemu glejakowi GL261, zaś podrozdział 6.2 ludzkiemu glejakowi

wielopostaciowemu. Dyskusja przedstawionych wyników zawarta jest w kolejnym rozdziale. Ostatni rozdział zawiera podsumowanie i wnioski, ujęte w pięciu podpunktach. Spis literatury zawiera 79 pozycji, w tym trzy współautorstwa Doktoranta. Około 80% przytoczonych publikacji pochodzi z ostatnich 10 lat, co potwierdza osadzenie pracy w nurcie współczesnych badań naukowych. Rozprawa jest napisana w sposób bardzo przejrzysty, jest też dobrze zredagowana i nie zawiera istotnych błędów językowych.

Główne wyniki rozprawy i wkład doktoranta

Rozprawa koncentruje się na bioinformatycznej analizie danych eksperymentalnych, która obejmuje następujące kroki:

1. wstępne przetwarzanie danych z sekwencjonowania: konwersja formatów, mapowanie i zliczanie odczytów,
2. przekształcanie macierzy ekspresji: kontrola jakości i filtrowanie, normalizacja i integracja danych z wielu próbek, redukcja wymiaru,
3. właściwa analiza ekspresji: grupowanie profilów ekspresji, identyfikacja genów o ekspresji różnicującej grupy, scharakteryzowanie grup.

Pierwszy z powyższych kroków został wykonany za pomocą narzędzia Cell Ranger opracowanego przez firmę 10x Genomics (producenta technologii sekwencjonowania pojedynczych komórek), do kolejnych kroków zastosowano środowisko Seurat v3. O ile ogólny schemat procedury analizy danych jest w zasadzie standardowy, jakość końcowego wyniku istotnie zależy od wyboru parametrów dla kolejnych kroków analizy. Należą do nich:

- progi odfiltrowania znaczników pochodzących od kropli nie zawierających komórek, zawierających multiplety bądź komórki obumierające,
- metoda integracji danych z poszczególnych próbek,
- poziom redukcji wymiaru,
- rozdzielczość grupowania,
- parametry wyznaczania genów markerowych dla grup.

Dla większości parametrów właściwy wybór wymaga umiejętnego połączenia biologicznej wiedzy eksperckiej ze zrozumieniem procedury analizy danych. W takich przypadkach istotna jest odpowiednia wizualizacja rezultatów poszczególnych etapów analizy i umiejętność zaaplikowania wniosków z oceny eksperckiej do kolejnych kroków procedury.

Mając powyższe na względzie, uważam, że opisany w rozprawie proces optymalizacji procedury analizy danych z sekwencjonowania transkryptomu z pojedynczych komórek stanowi oryginalne rozwiązanie problemu naukowego i dokumentuje umiejętność samodzielnego prowadzenia pracy naukowej przez Doktoranta. Z kolei zawartość rozdziałów Wstęp i Metody prezentuje ogólną wiedzę teoretyczną w dyscyplinie nauki biologiczne, w szczególności w obszarze badań nad gļejakami oraz technik badania transkryptomów pojedynczych komórek.

Uwagi krytyczne

1. Proces doboru parametrów został w rozprawie opisany rzetelnie, ale w niektórych przypadkach podjęte decyzje wydają się niedostatecznie poparte konkretnymi argumentami. Dotyczy to w szczególności kryteriów wyboru rozdzielczości grupowania. W rozdziale 6.1.3 pada ogólne stwierdzenie, że wybrano rozdzielczość 0.3 zapewniającą zgodność charakterystyki klastrów z wiedzą biologiczną. Z kolei w rozdziale 6.1.4 stwierdzono „wysoką stabilność przypisania komórek do klastrów”, po czym wybrano rozdzielczość 0.6 bez dodatkowego uzasadnienia, podobnie w rozdziale 6.2.3 wybrano rozdzielczość 0.4. Sprawia to wrażenie dużej arbitralności, zwłaszcza w kontekście postawionych w pracy szczegółowych celów badań – „opracowanie zoptymalizowanej metodyki analizy (...)”, „ocena wpływu wartości parametrów (...) dla uzyskania optymalnych wyników (...)”. Oczekiwałbym tu co najmniej podania uzasadnienia dla wybranych rozdzielczości oraz sformułowania kryteriów, które pozwalają uznać przypisanie komórek do klastrów za stabilne (być może dla określonego zakresu rozdzielczości).
2. W rozdziale 6.1.3 nie opisano, do jakich typów komórek zostały przypisane poszczególne klastry, co nie pozwala zweryfikować kończącego ten rozdział stwierdzenia, że „analizowane dane w większości przypadków są spójne pomiędzy biologicznymi powtórzeniami, w szczególności w obrębie populacji mikrogleju, makrofagów i BAM, które zostały wybrane do szczegółowej analizy”.
3. W kilku miejscach rozprawy podkreśla się znaczenie porównania w przeprowadzonych badaniach danych scRNA-seq pochodzących od mysich samic i samców. Tymczasem obserwacje związane z tym porównaniem ograniczają się do bardzo ogólnikowych stwierdzeń typu „występuje dysproporcja w liczbie komórek pochodzących od samic/samców w części klastrów zdominowanych przez komórki z próbek z grupy z nowotworem”, nie popartych konkretnymi danymi.

4. Rozdział Dyskusja w wyczerpujący sposób odnosi się celów pracy dotyczących przeprowadzenia analiz danych, natomiast cele związane z optymalizacją procesu analizy potraktowane są dość pobieżnie. Brakuje mi tu próby syntezy doświadczeń płynących z wykonanej pracy, np. w postaci odpowiedzi na następujące pytania:

- Jaki jest zakres rozsądnych wartości dla poszczególnych parametrów?
- Które parametry mają kluczowe znaczenie dla jakości wyników?
- Czy wybrane wartości parametrów odbiegały od rekomendowanych przez twórców pakietu Seurat? Jeśli tak, z czego to wynikało?

Konkluzja

Podsumowując, pomimo wymienionych powyżej uwag, stwierdzam, że mgr inż. Paweł Segit wykazał się wiedzą i umiejętnościami uprawniającymi do ubiegania się o stopień doktora nauk w dyscyplinie nauki biologiczne. Przedstawiona rozprawa doktorska spełnia warunki określone w art. 187 Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz.U. z 2023 r. poz. 742 z późn. zm.). W związku z powyższym, wnioskuję do Rady Naukowej Instytutu Biologii Doświadczalnej o dopuszczenie mgr inż. Pawła Segita do dalszych etapów postępowania w sprawie nadania stopnia doktora.

Norbert Dojer