

Warszawa, 14.03.2022

Prof. dr hab. Przemysław Juszczynski  
Z-ca Dyrektora ds. Nauki  
Instytut Hematologii i Transfuzjologii  
Ul. Indiry Gandhi 14  
Warszawa 02-776

## O C E N A

rozprawy doktorskiej mgr Michaliny Zajdel

### ***Analiza profili ekspresji mikroRNA i mRNA w chłoniaku podobnym do chłoniaka Burkitta z aberracją 11q (BLL,11q)***

W przedstawionej mi do recenzji pracy Doktorantka podjęła próbę analizy ekspresji mRNA i mikroRNA w w chłoniaku podobnym do chłoniaka Burkitta z aberracją 11q (BLL,11q) i identyfikacji różnic z chłoniakiem Burkitta i chłoniakami DLBCL.

Chłoniak podobny do chłoniaka Burkitta (BL) z aberracją w obrębie długiego ramienia chromosomu 11 (BLL,11q, ang. Burkitt-like lymphoma with 11q aberration) to bardzo rzadka nowa tymczasowa jednostka chorobowa wyodrębniona przez WHO w 2017 roku. Przypadki BLL,11q charakteryzują się szeregiem cech morfologicznych, immunofenotypowych, genetycznych i klinicznych bardzo zbliżonych do cech BL. Jednocześnie niektóre BLL,11q odróżnia od BL morfologia, odpowiadająca obrazowi

typowemu dla chłoniaków rozlanych z dużych komórek B, bliżej nieokreślonych (DLBCL,NOS, ang. diffuse large B-cell lymphoma, not otherwise specified) lub chłoniaków pośrednich pomiędzy BL a DLBCL. Chłoniaki BLL,11q nie ujawniają rearanżacji MYC, wykazują bardziej złożony kariotyp niż BL. Zmianą cytogenetyczną charakterystyczną dla BLL,11q jest obecność aberracji w obrębie długiego ramienia chromosomu 11 – jednoczesnej duplikacji fragmentu 11q23.2-23.3 oraz delekcji fragmentu telomerycznego 11q24.1-qter. Z uwagi na rzadkie występowanie tej tymczasowej jednostki, brak w literaturze kompleksowych opracowań dotyczących jej biologii, a jej diagnostyka sprawia wciąż trudności.

Zanim przejdę do oceny pracy, chciałbym nadmienić, iż pracę doktorską mgr M. Zajdel recenzuję po raz drugi. Poprzednio przedstawiona mi do oceny praca wykazywała szereg błędów, które nie pozwoliły mi jej ocenić pozytywnie. Wnioskowałem wówczas o uzupełnienie pracy o dodatkowe badania korygujące przedstawione wówczas niedociągnięcia i ponowne przedstawienie jej do recenzji.

Po pięciu latach otrzymałem pracę dotyczącą zbliżonej tematyki, z przerebadowanymi celami pracy i metodyką. Rozprawa tym razem przedstawiona do oceny posiada konstrukcję typową dla prac doktorskich i obejmuje wstęp, przedstawienie celów pracy, opis materiałów i metod, wyniki, dyskusję oraz wnioski. Rozprawę uzupełniają streszczenie w języku polskim i angielskim, wykaz użytych skrótów, spis tabel i rycin i piśmiennictwo. Podstawowym celem pracy, jaki Autorka przyjęła jest charakterystyka molekularna BLL,11q w porównaniu do chłoniaka Burkitta, a dodatkowo także do GCB DLBCL,NOS, w oparciu o badanie wzorów ekspresji mikroRNA i mRNA.

Uwagi recenzenta:

1. Autorka we Wstępie przedstawia aktualne aspekty klasyfikacji WHO chłoniaków agresywnych, ze szczególnym zwróceniem uwagi na kryteria diagnostyczne BL, BLL, 11q i DLBCL i ewolucję klasyfikacji WHO od 2008 roku. Niektóre stwierdzenia z wstępu wydają się sformułowane „na wyrost”: na przykład autorka twierdzi, że *„Bardzo skuteczną metodą diagnostyczną jest*

*ocena immunofenotypu komórek pobranych za pomocą biopsji cienkoigłowej metodą cytometrii przepływowej. Pozwala to na szybkie i precyzyjne postawienie diagnozy.*” Nie negując niekwestionowanej wartości cytometrii jako narzędzia diagnostycznego, nie jest ono narzędziem wystarczającym do „precyzyjnego postawienia diagnozy”, a stanowi jedną z komplementarnych służących temu technik, co z pełnym naciskiem podkreślają klasyfikacje WHO zarówno z 2008 jak i z 2017 roku, na które powołuje się autorka.

2. We wstępie brakuje podkreślenia ewolucji molekularnych klasyfikacji chłoniaków- począwszy od transkryptomicznych klasyfikacji (COO, CCC), po bardziej aktualne klasyfikacje grupy M. Shipp opublikowane w Nature Medicine w 2018 i grupy L Staudta (NEJM i Cancer Cell), oraz szerszej dyskusji o transkryptomicznych klasyfikatorach BL. Biologiczna charakterystyka tych chorób jest powierzchowna.
3. W kolejnej części wstępu dość autorka omawia zagadnienia związane z biogenezą i funkcją mikroRNA. Tu także brakuje mi uporządkowanego i głębokiego spojrzenia na tematkę : przedstawione dane, zwłaszcza w kontekście roli mikroRNA w patogenezie chłoniaków sa powierzchowne i hasłowe. Autorka pomija na przykład omówienie fundamentalnych publikacji dotyczących onkogennej funkcji miRNA 155 z modeli zwierzęcych (Proc Natl Acad Sci U S A. 2006;103(18):7024-9 , chociaż sama praca znalazła się w piśmiennictwie). W tej części brakło mi również choćby wzmianki, iż same mikroRNA są potencjalnymi celami terapeutycznymi.
4. W części „Materiały i metody autorka dość szeroko omawia proces diagnostyczny. W części 3.1.1. brakuje tabeli z charakterystyką wszystkich próbek, zwłaszcza w kontekście wyboru do dalszych badań NGS tylko części próbek, co budzi wątpliwości dotyczące selekcji danych, pozbawiając czytelnika możliwości samodzielnej oceny reprezentatywności wybranych próbek. **Poważnym przeoczeniem jest brak informacji o statusie EBV: może on istotnie zmieniać transkryptom zainfekowanych komórek, przez co wszelkie dalsze analizy, w kontekście przedstawionych celów, mogą być bezprzedmiotowe.** Proszę o przedstawienie tych informacji podczas obrony. Autorka natomiast zaznacza, że kryterium wykluczenia była obecność

wirusów takich jak: HIV, HBV, czy HCV. W opisie barwień IHC, autorka używa mało precyzyjnych i arbitralnych sformułowań „słabszy odczyn”, „wysoka ekspresja”.

5. W części dotyczącej metod molekularnych, autorka w wielu miejscach posługuje się objętościami używanych odczynników, zamiast ich ilościami molowymi (np. strona 62 - *Do każdej pojedynczej próby stosowano po 3  $\mu$ l specyficznych starterów*). Precyzja w tym zakresie jest niezbędna.
6. W części 3.11 (podstawowa ocena statystyczna) autorka wskazuje, iż rozkład zmiennych oceniała testem Shapiro – Wilka. Brak natomiast informacji o skutkach tej analizy – jakie kroki autorka podjęła, gdy rozkład był normalny, a jakie gdy nie ?
7. W części 3.12 przedstawione metody bioinformatyczne są mało precyzyjne i wydaje się nawet, że autorka nie wymienia wszystkich użytych pakietów środowiska R; na przykład, jak autorka generowała mapy ciepła? Jak skalowała dane (z-score?)

Ponieważ rozprawa w znacznej części bazuje na tej metodyce, warto było udostępnić napisane skrypty, np. za pośrednictwem portalu GitHub.

8. W części wyników stanowiących pierwszy etap pracy analizy statystyczne nie są wykonane poprawnie (Ryciny 1-4) - autorka powinna była użyć ANOVA (przy założeniu normalności rozkładu) lub testu Kruskala-Wallisa (rozkład zmiennych nieodpowiadający rozkładowi normalnemu), a następnie testów *post hoc* do oceny indywidualnych różnic między grupami.
9. Sposób obrazowania danych NGS jest, mówiąc najogólniej, suboptymalny. W przedstawionych mapach ciepła (na przykład rycina 17), skalowanie zmiennych wydaje się przeprowadzone „w kolumnach” – z map ciepła wynika, iż poziomy ekspresji genów lub miRNA różnią się między poszczególnymi transkryptami, ale pozostają bardzo zbliżone między próbkami. Ponieważ brakuje tu opisu metody, trudno te obserwacje zweryfikować obiektywnie, ale analiza map ciepła nie pozostawia raczej wątpliwości co do błędów w skalowaniu. Budzi to kolejne pytanie – czy wnioski płynące z wykonanych badań utrzymają się, jeśli skalowanie i obliczanie z-score w „wierszu” zostanie przeprowadzone poprawnie ? Jak będą wyglądać dendrogramy grupowania

hierarchicznego? Proszę Doktorantkę o szczegółowe odniesienie się do tej kwestii i zaprezentowanie prawidłowych analiz w trakcie obrony tej pracy, lub wyprowadzenie mnie z błędu, jeśli się mylę.

10. Analizy wyższego poziomu (GO enrichment) są pozbawione głębszych i logicznych analiz biologicznych. Autorka nie wykonała z niezrozumiałych przyczyn analiz GSEA z użyciem danych z bazy MSigDB, co jest niezrozumiałe choćby z uwagi na powszechnie używane grupy genów klasy „hallmark”, w tym genów regulowanych przez MYC - będących złotym standardem w tego typu analizach. Poproszę o uzupełnienie w trakcie obrony.
11. Analizy z użyciem genów „regulowanych przez MYC” bazują na bardzo uproszczonych kryteriach. Autorka jako geny regulowane przez MYC uznała geny wymienione w 2 wcześniejszych publikacjach sprzed ok. 15 lat, przy czym Autorka nie sprecyzowała sposobu skonstruowania tej listy (suma czy iloczyn zbiorów?). Tak czy owak – w świetle dzisiejszej wiedzy o biologii MYC, lista ta jest niekompletna i trudno w sposób biologicznie znaczący interpretować wyniki tych analiz. Istnieją liczne, nowsze publikacje, integrujące ChIP-Seq z RNA Seq, które w sposób daleko bardziej adekwatny mogłyby stanowić źródło danych do analiz. W mojej ocenie, byłoby bardzo pomocne zaprezentowanie danych z wykorzystaniem takich danych (na przykład z publikacji Seitz V, et al. *Deep sequencing of MYC DNA-binding sites in Burkitt lymphoma. PLoS One. 2011;6(11):e26837* lub Tesi A, et al. *An early Myc-dependent transcriptional program orchestrates cell growth during B-cell activation. EMBO Rep. 2019 Sep;20(9):e47987*. w trakcie obrony. Byłoby dobrze, gdyby autorka wykorzystwała różne publikacje jako źródła danych i zbadała na ile wnioski uzyskane przy różnych listach genów MYC-zależnych z różnych publikacji są ze sobą zbieżne.

Pomimo wskazanych powyżej wad i błędów, o których sprostowanie proszę w trakcie dyskusji, wydaje się że cel pracy został osiągnięty (o ile oczywiście wnioski się utrzymają i obronią). Autorka formułuje następujące wnioski płynące z pracy:

- Bank: PKO S.A. Oddział w Warszawie 16 1240 6003 1111 0000 4946 7956 • NIP: 525-000-94-24 • REGON: 000288484