

Analiza rakowego środowiska immunologicznego

Anna Szymik

21 maja 2024

1 Założenia projektu

1.1 Wprowadzenie

Ten projekt skupia się na analizie danych przestrzennych z wielokrotnej immunofluorescencji (mIF), pochodzących z biopsji nowotworów różnych pacjentów. Dane dostarczają informacji o lokalizacji i poziomach ekspresji markerów białkowych dla każdej komórki, dzięki czemu możliwe jest badanie przestrzennej organizacji i interakcji komórek w obrębie tkanki.

W przypadku tkanki nowotworowej, szczególnym obszarem zainteresowania w tym kontekście są trzeciorzędowe struktury limfatyczne (TLS) – formacje powstające w narządach niebędących częścią układu limfatycznego, takich jak nowotwory, w odpowiedzi na długotrwałą stymulację immunologiczną. Składają się one z różnych typów komórek odpornościowych, przede wszystkim limfocytów B i T, i odgrywają kluczową rolę w miejscowej odpowiedzi immunologicznej. Badania wykazały, że obecność TLS może być związana z lepszymi wynikami leczenia niektórych nowotworów [Fridman et al., 2022]. Analiza tych struktur jest zatem istotna dla zrozumienia mechanizmów odpowiedzi immunologicznej na nowotwory i może przyczynić się do rozwoju skuteczniejszych strategii leczenia.

1.2 Charakterystyka danych

Dane są w formacie tabelarycznym i zawierają informacje o aktywności białek, uzyskane z trzech paneli mIF dla różnych pacjentów. Każdy wiersz reprezentuje pojedyncze jądro komórkowe, wraz z jego współrzędnymi (mierzonymi w mikrometrach od początku płytki), znormalizowanymi wynikami aktywności określonych białek oraz odpowiadającym mu fenotypem, który został wywnioskowany z wyników aktywności.

1.3 Cel projektu

Głównym zadaniem projektu było umożliwienie wizualizacji danych z immunofluorescencji, identyfikacja potencjalnych trzeciorzędowych struktur limfatycznych (TLS) oraz zgrupowanie ich w klastry pod względem składu komórkowego. Końcowy cel stanowiło zapakowanie wszystkich tych elementów w formę intuicyjnej w obsłudze aplikacji.

2 Implementacja

2.1 Identyfikacja kandydatów na TLS

Identyfikacja potencjalnych struktur TLS w tym projekcie opiera się na implementacji grafowej. Z danych mIF, korzystając z bibliotek `sklearn` i `networkx` tworzę graf sąsiedztwa, w którym za sąsiadów uznaję komórki znajdujące się nie dalej niż $30\mu\text{m}$ od siebie. Następnie, jako potencjalny „trzon” struktur TLS, wybieram tylko te komórki, które oznaczone zostały jako *Bcell*, *Tcell* lub *BnTcell*. Na nich również buduję graf sąsiedztwa i znajduję w nim spójne składowe o 20 lub więcej wierzchołkach – wybór minimalnej liczby komórek jest oczywiście arbitralny. W kolejnym kroku dodaję do kandydatów sąsiadów każdej komórki w spójnej składowej z pełnego grafu, aby uzyskać informacje na temat rodzajów komórek sąsiadujących z komórkami odpornościowymi.

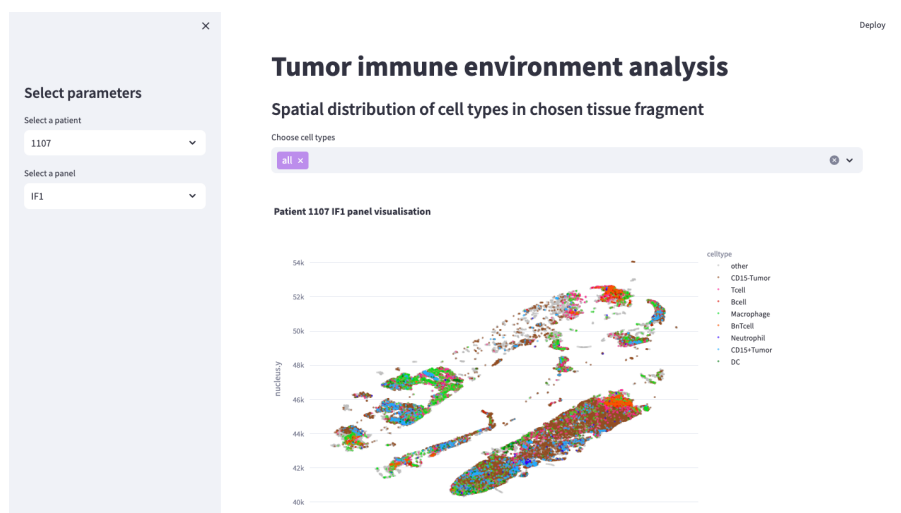
2.2 Analiza składu TLS

Dla każdego z pacjentów, po identyfikacji kandydatów na TLS, obliczam wektor procentowego udziału każdego typu komórek w tej strukturze. Następnie, na wszystkich powstałych w ten sposób wektorach przeprowadzam klastrowanie hierarchiczne, by móc zobaczyć grupy TLS podobne pod względem składu komórkowego.

Poza tym, postanowiłam przeanalizować, jak zmienia się sąsiedztwo limfocytów B w całym skrawku tkanki, jako że są to szczególnie istotne komórki w strukturach TLS. Dla każdej komórki B, przechodząc przez nie od lewego dolnego do prawego górnego rogu wykresu – rosnąco względem pozycji x, y – obliczam procentowy udział komórek, będących jej sąsiadami.

2.3 Wizualizacja danych

Do graficznego przedstawienia danych korzystam z biblioteki `plotly`, a całość postanowiłam opakować w formie aplikacji `streamlit`.



Rysunek 1: Interfejs graficzny aplikacji.

Dla wybranego pacjenta można znaleźć w niej:

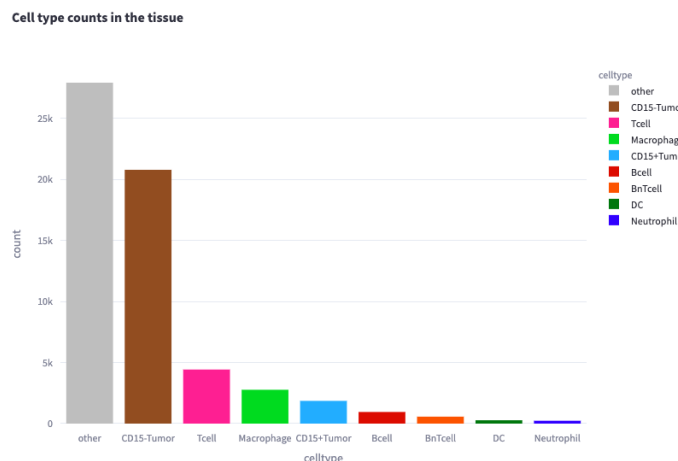
- wizualizację fragmentu tkanki reprezentowaną jako wykres punktowy, z komórkami pokolorowanymi według ich typu;
- statystyki dotyczące liczby typów komórek w próbce;
- wykresy ilustrujące zmiany składu komórkowego sąsiadów limfocytów B w tkance;
- przestrzenne rozmieszczenie kandydatów na TLS, z wyróżnionymi typami komórkowymi;
- procentowy udział poszczególnych typów komórek w każdym TLS.

Dodatkowo, można w niej zobaczyć, jak, pod względem występujących typów komórek, klastrują się TLS znalezione u wszystkich badanych pacjentów. Oprócz procentowego składu komórkowego w klastrach, przedstawiony jest również dendrogram odpowiadający przeprowadzonemu klastrowaniu hierarchicznemu.

W celu przyspieszenia renderowania aplikacji, dane dotyczące składu TLS są przechowywane w osobnym pliku, co eliminuje konieczność ich ponownego obliczania za każdym razem.

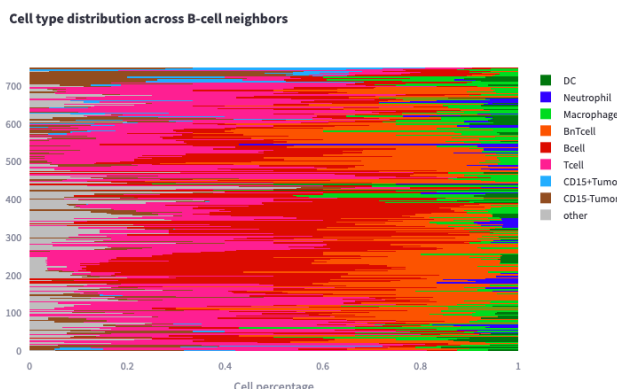
3 Wyniki

Niestety, z powodu szumu w danych, spowodowanych z obciążeniami metod eksperymentalnych oraz przypisywaniem markerów do jądra komórkowego, wiele komórek oznaczonych zostało bardzo ogólnie jako *other*, co utrudnia wyciąganie biologicznie istotnych wniosków (patrz: Rys. 2). Mimo to, z przeprowadzonych przeze mnie analiz można wysunąć kilka ciekawych obserwacji.



Rysunek 2: Przykładowy skład komórkowy w próbce, pochodzący od pacjenta 1107. Większość stanowią komórki *other*.

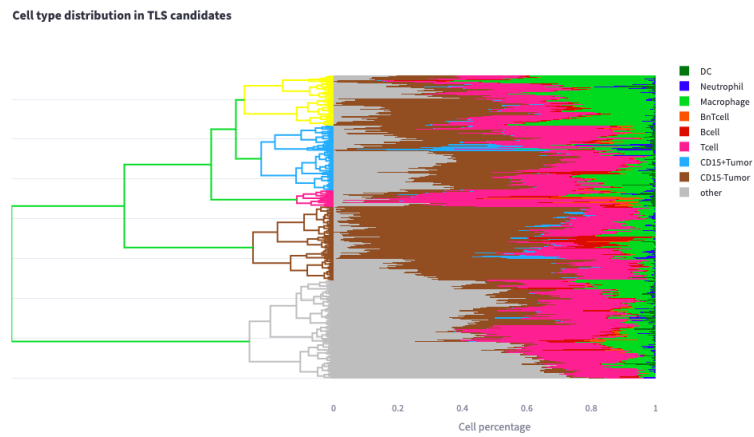
Z wykresu zmian sąsiedztwa limfocytów B, możemy niejednokrotnie wnioskować, w którą stronę zachodzi infiltracja. Na Rys. 3 możemy zauważyć, że u pacjenta 0374 wśród sąsiadów komórek w „prawym górnym rogu” tkanki przeważają komórki rakowe, zaś później skład sąsiadów ulega zmianie i zaczynają przeważać limfocyty T, co może sugerować, że odpowiedź immunologiczna zachodzi od tej właśnie strony. W przyszłości, tę funkcjonalność aplikacji można byłoby ulepszyć, ilustrując jak zmienia się sąsiedztwo limfocytów B w obrębie TLS, a nie całej tkanki.



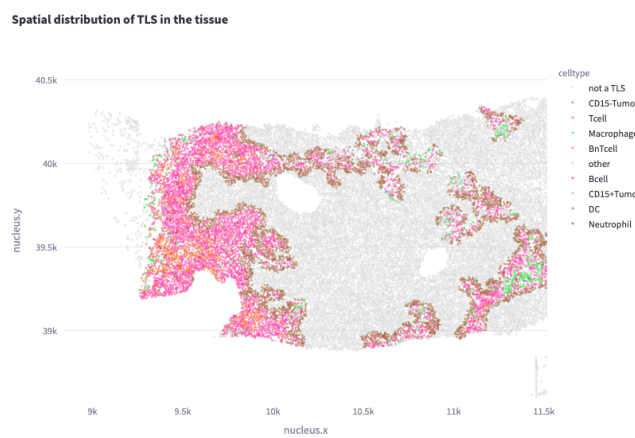
Rysunek 3: Wykres zmian składu komórkowego sąsiadów limfocytów B w próbce pacjenta 0374.

Jeśli chodzi o skład komórkowy zidentyfikowanych potencjalnych TLS, w tym przypadku również przeszkodą jest duża liczba komórek oznaczonych jako *other* – widoczne jest to szczególnie w przypadku ostatniego klastra (Rys. 4), w którym większość stanowią takie komórki, dlatego trudno wyciągnąć z niego biologicznie informatywne wnioski. Pozostałe zidentyfikowane klastry niosą już jednak ze sobą więcej informacji. Możemy znaleźć wśród nich chociażby klastry z przeważającą liczbą komórek nowotworowych CD15– czy klastry, w którym znaczącą większość stanowią limfocyty T. Ten drugi jest szczególnie ciekawy, jako że wskazuje na silną odpowiedź immunologiczną. Tego typu TLS występują na przykład u pacjenta 0316, co widoczne jest na Rys. 5.

Podsumowując, mimo pewnych ograniczeń związanych z szumem w danych, udało się zidentyfikować kilka istotnych obserwacji. Analiza zmian sąsiedztwa limfocytów B pozwoliła na wnioskowanie o kierunku infiltracji – choć można byłoby uszczegółowić te analizy – a identyfikacja klastrów w potencjalnych TLS dostarczyła cennych informacji o wzorcach w składzie komórkowym tych struktur. Te wyniki mogą stanowić podstawę do dalszych analiz złożonego zagadnienia, jakim jest powstawanie i funkcjonowanie trzeciorzędowych struktur limfatycznych.



Rysunek 4: Klastrowanie hierarchiczne zidentyfikowanych TLS.



Rysunek 5: Przykład TLS, w którym przeważają limfocyty T, pochodzący od pacjenta 0316.

4 Referencje

Fridman, W.H., Meylan, M., Petitprez, F. et al. B cells and tertiary lymphoid structures as determinants of tumour immune contexture and clinical outcome. *Nat Rev Clin Oncol* 19, 441–457 (2022).