

Sekwencjonowanie transkryptomu pojedynczych komórek: droga do odkrywania bioróżnorodności komórkowej

dr Anna Samelak-Czajka,

dr inż. Małgorzata Marszałek-
-Zeńczak,

dr hab. Paulina Jackowiak,
prof. ICHB PAN✉

Pracownia Analiz Pojedynczych Komórek, Instytut Chemii Bioorganicznej Polskiej Akademii Nauk

https://doi.org/10.18388/pb.2021_532

✉autor korespondujący: paulinaj@ibch.poznan.pl

Słowa kluczowe: transkryptomika, sekwencjonowanie wysokoprzepustowe, analiza pojedynczych komórek

Wykaz skrótów: RNA-Seq – wysokoprzepustowe sekwencjonowanie RNA; scRNA-Seq – wysokoprzepustowe sekwencjonowanie transkryptomu pojedynczych komórek; snRNA-Seq – wysokoprzepustowe sekwencjonowanie transkryptomu pojedynczych jąder komórkowych

STRESZCZENIE

Sekwencjonowanie transkryptomu pojedynczych komórek (scRNA-Seq) to przełomowa technologia, która otworzyła drogę do opisywania ekspresji genów z niespotykaną dotąd rozdzielczością. Umożliwia ona odkrywanie różnorodności komórkowej organizmów i śledzenie ich procesów rozwojowych. Opracowano szereg rozwiązań technologicznych umożliwiających analizy od setek do nawet miliona komórek w jednym eksperymencie, a także obszerny zestaw narzędzi do bioinformatycznej analizy uzyskanych danych. Bogactwo informacji dostarczanych przez scRNA-Seq oraz możliwość wykorzystania tej metody do badania komórek, organoidów, tkanek, a nawet całych organizmów, decydują o szerokim spektrum jej zastosowań. W niniejszej pracy przedstawiamy przebieg części eksperymentalnej i obliczeniowej procedury scRNA-Seq, a także prezentujemy najważniejsze zastosowania tej technologii w biomedycynie, biologii rozwoju i biologii roślin.

ROZWÓJ TECHNIK SEKWENCJONOWANIA TRANSKRYPTOMU

Transkryptom to zestaw wszystkich cząsteczek RNA powstałych w procesie transkrypcji genów, obecnych w danym układzie biologicznym (komórce, tkance, organizmie) w określonym czasie. Tworzą go zarówno cząsteczki matrycowego RNA (ang. *messenger RNA*, mRNA), niosącego informację o sekwencji białek, jak i RNA niekodujące (ang. *non-coding RNA*, ncRNA), które warunkują przebieg najważniejszych konstitutywnych procesów komórkowych i pełnią funkcje regulatorowe [1–3]. Cechą transkryptomu jest jego zmienność w zależności od czynników, takich jak stan komórki i etap jej rozwoju czy bodźce środowiskowe. Zmiany te przejawiają się w różnicach jakościowych i ilościowych poszczególnych transkryptów. Badanie transkryptomu tkanek, fragmentów narządów, czy całych organizmów w różnych warunkach i punktach czasowych pozwala zatem na poznanie niektórych mechanizmów regulacji ekspresji genów, identyfikację genów zaangażowanych w konkretne procesy biologiczne (takie jak rozwój, choroby, odpowiedź na bodźce zewnętrzne), a także umożliwia odkrywanie nowych markerów diagnostycznych, prognostycznych lub terapeutycznych [1].

Zainteresowanie badaniami transkryptomycznymi wzrosło wraz z opracowaniem technologii mikromacierzy, a następnie wysokoprzepustowego sekwencjonowania drugiej generacji (ang. *next generation sequencing*, NGS). Technologie te umożliwiły analizę całych transkryptomów i stworzenie baz danych transkryptomicznych dla wielu organizmów i ich etapów rozwojowych czy stanów chorobowych. Dzięki tym technikom możliwa była również identyfikacja nowych typów ncRNA [1]. Ostatecznie, metody sekwencjonowania RNA z wykorzystaniem technik NGS (ang. *RNA sequencing*, RNA-Seq) zyskały popularność i w dużej mierze wyparły mikromacierze, ze względu na wyższą czułość, umożliwienie analizy wszystkich transkryptów w próbce, zdolność do wykrywania nowych transkryptów i większy zakres dynamiczny pomiaru (zdolność do precyzyjnego wykrywania różnic w poziomach transkryptów, zarówno w przypadku genów o wysokiej, jak i niskiej ekspresji) [1].

RNA-Seq ma charakter sekwencjonowania masowego (ang. *bulk sequencing*), to znaczy, że biblioteka otrzymana w wyniku przepisania RNA na DNA, reprezentuje ogół cząsteczek RNA wyizolowanych z fragmentów tkanek/organów składających się z różnych typów komórek, o różnym poziomie ekspresji poszczególnych genów. Tym samym zbadany poziom ekspresji danego genu obrazuje jedynie wartość średnią dla wszystkich komórek w próbce. Tymczasem ekspresja genów ma charakter czasowo- i komórkowo-specyficzny. To, które geny podlegają ekspresji i na jakim poziomie, pozwala wnioskować nie tylko o rodzaju komórki, ale odzwierciedla również jej stan, funkcję i odpowiedź na czynniki zewnętrzne. Dlatego, aby dokładniej zbadać zmienność ekspresji genów, lepiej poznać bioróżnorodność komórkową i zrozumieć zależności między

komórkami funkcjonującymi w heterogenicznych, złożonych układach, należy skupić się na transkryptomie pojedynczych komórek [2-4].

Pierwsze próby badania transkryptomu pojedynczych komórek podjęto w latach dziewięćdziesiątych XX. i na początku XXI. wieku. Polegały one na umieszczeniu pojedynczej komórki w probówce, przeniesieniu informacji zawartej w mRNA na cDNA (odwrotna transkrypcja) i amplifikacji uzyskanego w ten sposób cDNA metodą ilościowej łańcuchowej reakcji polimerazy (ang. *polymerase chain reaction*, PCR) lub transkrypcji *in vitro* (ang. *in vitro transcription*, IVT) [3,4]. Z czasem zaadaptowano do profilowania ekspresji genów pojedynczych komórek popularne wówczas mikromacierze. Ograniczeniem tych podejść była ich niska przepustowość (niewielka liczba badanych komórek) i brak możliwości analizy całego transkryptomu. Przełomowa okazała się praca Tang i współpracowników [5], którzy po raz pierwszy zaadaptowali NGS do zbadania transkryptomu komórek myszy na wczesnym etapie rozwoju embrionalnego. Początkowo, sekwencjonowanie transkryptomu pojedynczych komórek (ang. *single cell RNA sequencing*, scRNA-Seq) było jednak ograniczone przez technologiczne i metodyczne bariery. Głównym wyzwaniem była konieczność fizycznego odseparowania poszczególnych komórek od siebie podczas całej wieloetapowej procedury, aby zapobiec zmieszaniu ze sobą transkryptów (i odpowiadających im cDNA) pochodzących z różnych komórek. Z czasem zaczęto wykorzystywać do tego celu sortowanie komórek aktywowane fluorescencją (ang. *fluorescence-activated cell sorting*, FACS), aby umieścić pojedyncze komórki w 384 dołkach mikropłytki, w których następnie wykonywano kolejne etapy procedury konstrukcji biblioteki do sekwencjonowania [6, 7]. Metody te nadal jednak pozwalały na równoczesną analizę jedynie od kilkudziesięciu do kilkuset komórek. Zdecydowanym usprawnieniem było pojawienie się na rynku urządzeń, które umożliwiły automatyzację całego procesu - od precyzyjnego wychwytywania pojedynczych komórek aż po uzyskanie gotowej do sekwencjonowania biblioteki [8,9]. Jednocześnie opracowane zostały metody indeksowania transkryptów pochodzących z tej samej komórki wspólnym oligonukleotydowym znacznikiem (ang. *barcoding*) dodawanym w procesie odwrotnej transkrypcji [10]. Dzięki temu możliwe stało się łączenie cDNA pochodzącego z różnych komórek i konstrukcja bibliotek w ramach jednej reakcji.

Kolejny przełom w transkryptomice pojedynczych komórek związany był z rozwojem technologii kroplowych, wykorzystujących układy mikroprzepływowe. Obecnie metody polegające na w pełni zautomatyzowanym zamykaniu pojedynczych komórek w obrębie nanolitrowej kropli cieszą się największą popularnością. Wysoka przepustowość, którą zapewniają pozwalająca na zbadanie transkryptomu do 20 000 komórek w pojedynczej reakcji, umożliwia wykrywanie i badanie rzadkich populacji [2,4]. Jeszcze większą przepustowością, sięgającą nawet miliona komórek w jednym eksperymencie, a do tego brakiem wymagań sprzętowych, charakteryzują się tak zwane metody płytkowe [11-14]. Choć pod pewnymi względami przypominają pierwsze metody badania transkryptomów pojedynczych komórek,

to wykorzystują one indeksowanie kombinatoryczne transkryptów znajdujących się w obrębie danej komórki.

W ostatnich latach naukowcy opracowali wiele protokołów bazujących na opisanych powyżej technologiach, powstało także kilka platform dostępnych komercyjnie. Ich porównanie można odnaleźć w licznych pracach przeglądowych [2,24].

Wraz z pojawianiem się kolejnych innowacji technologicznych i nowych metod scRNA-Seq, szybko rozwinął się także zasób narzędzi i algorytmów bioinformatycznych, nie tylko umożliwiających różnicową analizę ekspresji genów, ale też pozwalających na identyfikację różnych typów komórek, charakteryzowanie stanów komórkowych i rekonstrukcję trajektorii rozwojowych [4,15].

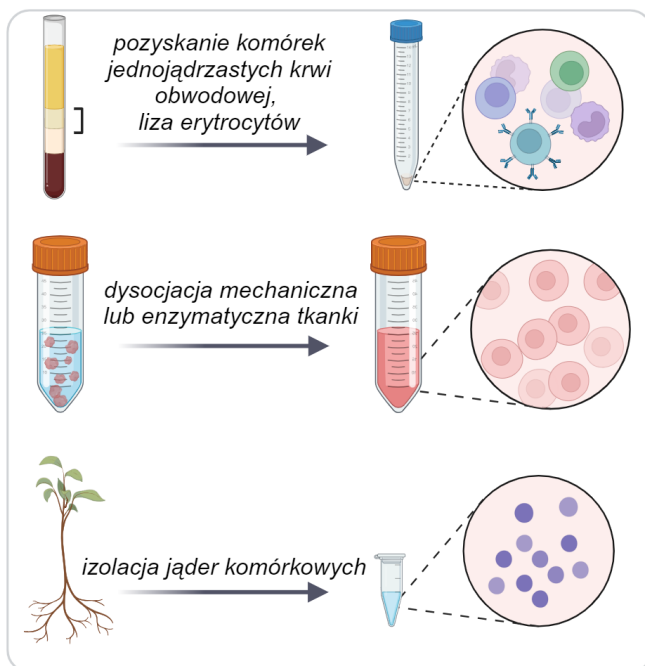
PRZEBIEG EKSPERYMENTU SEKWENCJONOWANIA TRANSKRYPTOMU POJEDYNCZYCH KOMÓREK - ETAPY I ZWIĄZANE Z NIMI WYZWANIA

Analiza transkryptomu pojedynczych komórek wymaga pokonania dwóch wyzwań, które nie występują w standardowych metodach sekwencjonowania masowego: izolacji i separacji pojedynczych komórek oraz amplifikacji niewielkich ilości mRNA. Protokoły przeznaczone do scRNA-Seq, choć w szczegółach różnią się między sobą, obejmują te same etapy: separację pojedynczych komórek, lizę komórek, odwrotną transkrypcję, amplifikację cDNA, konstrukcję biblioteki, sekwencjonowanie oraz bioinformatyczną analizę danych [16]. Kroki te poprzedza przygotowanie próbki będącej zawiesiną pojedynczych komórek lub jąder komórkowych (Ryc. 1). Etap ten jest kluczowy dla uzyskania danych o odpowiedniej jakości i w dużym stopniu decyduje o powodzeniu całego eksperymentu.

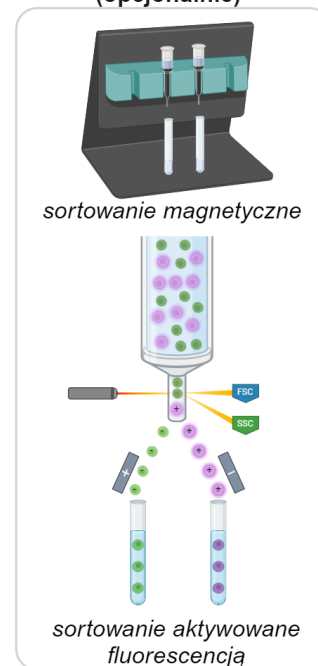
PRZYGOTOWANIE ZAWIESINY POJEDYNCZYCH KOMÓREK/JĄDER KOMÓRKOWYCH

Złożoność i stopień trudności procesu przygotowania zawiesiny pojedynczych komórek zależy od rodzaju i charakteru materiału wyjściowego oraz docelowej populacji, której transkryptom ma zostać zbadany [17]. W przypadku krwi czy szpiku kostnego, w których nie ma konieczności rozbijania połączeń międzykomórkowych, przygotowanie próbki zwykle ogranicza się do lizy erytrocytów i kilku-etapowego przemycia komórek buforem, choć w przypadku agregowania materiału może także wymagać sączenia przez filtry. Hodowle komórkowe również najczęściej wymagają jedynie łagodnej dysocjacji enzymatycznej np. roztworem trypsyny [17]. Z kolei fragmenty tkanek i organów, guzy nowotworowe, a nawet całe organizmy, niejednokrotnie wymagają połączenia mechanicznych i enzymatycznych metod dysocjacji. Dla niektórych tkanek, zwykle pochodzenia ludzkiego, mysiego lub szczurzego, dostępne są gotowe mieszanki enzymów i zoptymalizowane protokoły dysocjacji. Planując procedurę przygotowania materiału do badań należy jednak pamiętać, że żywotność komórek w docelowej próbce powinna wynosić co najmniej 80%. Tymczasem narażenie na nadmierne działanie enzymów może uszkadzać komórki i prowadzić do ich śmierci. Ponadto, wykazano, że dysocjacja wpływa na indukcję ekspresji genów odpowiedzi na stres, co w efekcie może prowadzić do błęd-

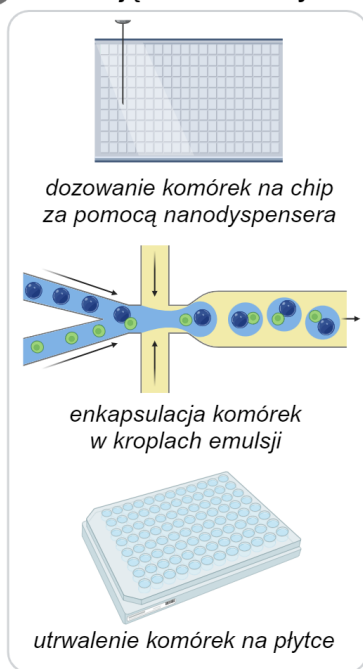
1 przygotowanie zawiesiny pojedynczych komórek lub jąder komórkowych



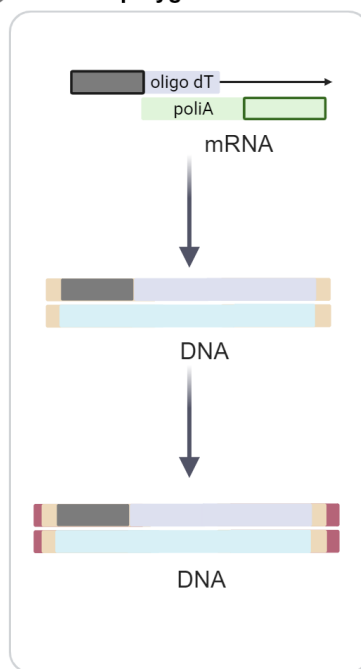
2 wyodrębnienie frakcji komórek lub jąder komórkowych (opcjonalnie)



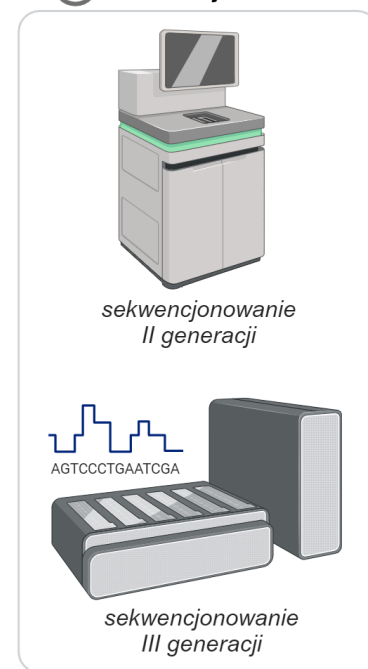
3 separacja pojedynczych komórek lub jąder komórkowych



4 odwrotna transkrypcja, amplifikacja DNA i przygotowanie bibliotek



5 wysokoprzepustowe sekwencjonowanie



Rycina 1. Etapy procedury sekwencjonowania transkryptomu pojedynczych komórek. Pierwszym etapem jest pozyskanie bądź przygotowanie zawiesiny pojedynczych komórek lub jąder komórkowych (1). Otrzymany w ten sposób materiał można następnie dodatkowo oczyścić lub wzbogacić w pożądaną frakcję (2). Kluczowym etapem jest rozdzielenie pojedynczych komórek lub jąder komórkowych do niezależnych przedziałów reakcyjnych (3). Następnie zachodzi odwrotna transkrypcja, amplifikacja DNA i przygotowanie bibliotek. Dzięki zastosowaniu znaczników, zachowana zostaje informacja o tym, z której komórki pochodził dany transkrypt (4). Ostatni etap obejmuje wysokoprzepustowe sekwencjonowanie, które może być wykonane z wykorzystaniem technologii krótkich lub długich odczytów (5).

nych wniosków dotyczących procesu biologicznego poddawanego badaniom [18–20]. Z tej przyczyny, niezwykle istotna jest optymalizacja dysocjacji pod kątem takich czynników jak stężenie enzymów czy czas i temperatura inkubacji. Alternatywnym podejściem, mającym zastosowanie

do materiału trudnego do dysocjacji lub tkanek uprzednio zamrożonych, jest izolacja jąder komórkowych [17,21,22].

Niezależnie od podejścia, należy zadbać o to, aby udział poszczególnych typów komórek w próbce odzwierciedlał

ich proporcje w materiale wyjściowym, chyba, że celem eksperymentu jest zbadanie konkretnej populacji komórek. Zastosowanie mają wówczas metody polegające na wzbogaceniu próbki w określony rodzaj komórek, takie jak: FACS czy sortowanie magnetyczne (ang. *magnetic-activated cell sorting*, MACS). FACS jest powszechnie stosowaną metodą polegającą na separacji komórek określonego typu na podstawie sygnału fluorescencyjnego. Sygnał ten może pochodzić między innymi od białka reporterowego powstającego w komórce, czy znacznika fluorescencyjnego połączonego z przeciwciałem rozpoznającym określony antygen na powierzchni lub wewnątrz komórki. Metoda ta pozwala na stosunkowo precyzyjne wyselekcjonowanie pożądaných komórek. W procedurze MACS stosuje się kulki paramagnetyczne sprzęgnięte z przeciwciałami skierowanymi przeciwko określonemu antygenowi na powierzchni komórki. Próbkę inkubuje się z kulkami, a następnie przepuszcza przez kolumnę separacyjną znajdującą się w polu magnetycznym. Komórki związane z kulkami paramagnetycznymi pozostają na kolumnie, a pozostałe swobodnie przez nią przepływają. Metoda ta służy głównie do eliminacji z próbki komórek niepożądanych, np. dominujących pod względem ilościowym, w przypadku, gdy celem eksperymentu jest zbadanie mniej licznej populacji. Powyższe metody stosuje się również w celu pozbycia się z próbki komórek martwych lub fragmentów komórkowych (ang. *cell debris*) trudnych do odseparowania standardowymi metodami oczyszczania próbki (filtracja, płukanie, wirowanie).

SEPARACJA POJEDYNCZYCH KOMÓREK

Istotą tego etapu jest odseparowanie od siebie poszczególnych komórek tak, aby umożliwić późniejsze wyznaczenie ich transkryptów. Stanowi to wyzwanie zwłaszcza w przypadku próbek zawierających liczne populacje komórek, zróżnicowane pod względem wielkości i kształtu. Dodatkowo, manipulacja komórkami powinna być na tyle łagodna i sprawna, aby nie wywierała wpływu na ekspresję genów i pozwalała zachować integralność błon komórkowych. Rozpad komórek powoduje bowiem uwolnienie transkryptów do buforu i prowadzi do artefaktów w analizie. Ponadto, dużą uwagę przywiązuje się do skalowalności metod wychwytywania pojedynczych komórek. Najbardziej pożądane są rozwiązania, które umożliwiają analizę zarówno próbek ubogokomórkowych, jak i tych o dużym zróżnicowaniu, liczących kilkadziesiąt tysięcy komórek, z zachowaniem wysokiej precyzji i niezawodności.

Zasadniczo stosuje się trzy podejścia. W pierwszym, historycznie najstarszym, komórki rozdzielone są fizycznie poprzez umieszczenie ich w oddzielnych dołkach na mikroplacie lub kompartmentach na specjalnym czipie. Kolejny sposób separacji pojedynczych komórek reprezentują technologie kroplowe, bazujące na mikroprzepływach i wytwarzaniu emulsji typu woda w oleju. Poszczególne komórki zamykane są w oddzielnych kroplach, które stanowią przestrzeń do zachodzenia kolejnych procesów aż do uzyskania pierwszej nici cDNA. Podejściem alternatywnym wobec powyższych są metody płytkowe stosujące indeksowanie kombinatoryczne, w których nie dochodzi do fizycznego odseparowania komórek. Komórki do analizy zostają specjalnie utrwalone i same w sobie stanowią odrębny prze-

dział, w ramach którego dochodzi do przyłączania znaczników oligonukleotydowych i odwrotnej transkrypcji.

LIZA KOMÓREK, ODWROTNA TRANSKRYPCJA I AMPLIFIKACJA cDNA

We wszystkich metodach, poza wspomnianym powyżej podejściem płytkowym, odseparowane od siebie komórki ulegają lizie celem uwolnienia znajdujących się w nich transkryptów. W klasycznej metodzie RNA-Seq wymagane jest 0,1–1,0 µg całkowitego RNA do stworzenia biblioteki. Tymczasem ilość całkowitego RNA obecnego w pojedynczej komórce jest znacznie mniejsza i wynosi od 1–50 pg, w zależności od typu komórki [23], a mRNA stanowi zwykle jego niewielki odsetek, ok. 1–5%. Aby przewyciężyć problem małej ilości materiału, wychwytuje się cząsteczki mRNA i selektywnie przepisuje się je na cDNA w procesie odwrotnej transkrypcji, a następnie amplifikuje przed utworzeniem biblioteki. Przeważająca większość technik do specyficznego wychwytywania cząsteczek mRNA wykorzystuje znajdujący się na ich końcach 3' ciąg reszt adeninowych (fragment poliadenylowy, ogon poli(A)). Pozwala to uniknąć amplifikacji obecnego w dużych ilościach rybosomalnego RNA. Analiza niepoliadenylowanych RNA jest trudniejsza i wymaga zastosowania specjalistycznych protokołów. Ogon poli(A) wylapywany jest przy użyciu oligonukleotydu złożonego z reszt tymidynowych (oligo(dT)). Odwrotna transkryptaza wykorzystuje oligo(dT) jako starter do syntezy pierwszej nici cDNA, komplementarnej do mRNA.

W zależności od protokołu scRNA-Seq, do oligo(dT) dołączone są również inne sekwencje nukleotydowe, takie jak: (i) znacznik (ang. *barcode*) jednakowy dla wszystkich cząsteczek mRNA z danej komórki; (ii) odcinki adaptorowe umożliwiające sekwencjonowanie na platformach NGS; (iii) unikatowe identyfikatory molekularne (ang. *unique molecular identifiers*, UMI), służące do jednoznacznego oznaczenia pojedynczej cząsteczki mRNA. Pozwalają one na identyfikację duplikatów wynikających z amplifikacji PCR i ich usunięcie podczas analizy bioinformatycznej.

Dodawanie znaczników umożliwia połączenie materiału pochodzącego z wielu komórek i przeprowadzanie kolejnych etapów eksperymentu (synteza drugiej nici cDNA, amplifikacja, konstrukcja bibliotek) w sposób zbiorczy, w jednej próbce. Aby zwiększyć przepustowość, czyli umożliwić łączenie materiału pochodzącego z większej liczby komórek, konieczne jest użycie większej liczby znaczników, co z kolei wymaga zastosowania dłuższych oligonukleotydów. Można jednak uniknąć tego problemu, stosując strategię kombinatorycznego przyłączania kilku krótszych fragmentów w następujących po sobie rundach ligacji – rozwiązanie to znalazło zastosowanie w metodach płytkowych typu SPLiT-seq (ang. *split pool ligation-based transcriptome sequencing*) czy sci-RNA-seq (ang. *single cell combinatorial indexing RNA sequencing*) [11–14].

Wyznakowane, jednoniciowe cDNA z pojedynczych komórek na tym etapie są łączone, a dalsze procesy przebiegają zbiorczo. Amplifikacja cDNA najczęściej odbywa się z wykorzystaniem PCR, rzadziej na drodze transkrypcji

in vitro, po której następuje kolejna runda odwrotnej transkrypcji [2,4,24].

Kluczowe dla etapu wychwytywania transkryptów i amplifikacji materiału, szczególnie w przypadku genów o niskiej ekspresji, jest zapewnienie wysokiej wydajności procesu odwrotnej transkrypcji i syntezy drugiej nici cDNA [17].

SEKWENCJONOWANIE

Wyznakowane cząsteczki cDNA podlegają standardowym etapom przygotowania bibliotek do NGS, zależnie od wybranej platformy. Końcowa biblioteka po raz kolejny podlega amplifikacji za pomocą PCR. W ostatnim czasie coraz większym zainteresowaniem cieszy się adaptacja do scRNA-Seq technologii sekwencjonowania trzeciej generacji, wykorzystujących długie odczyty [25–27]. Ich zaletą, w porównaniu z technikami drugiej generacji, jest wyższa skuteczność w wykrywaniu alternatywnego splicingu lub rearanżacji DNA (fuzji genowych) dzięki bezpośredniemu sekwencjonowaniu nienaruszonych cDNA o pełnej długości oraz ulepszone wykrywanie izoform, jak i nowych transkryptów, w tym także fuzyjnych [28,29].

BIOINFORMATYCZNA ANALIZA DANYCH

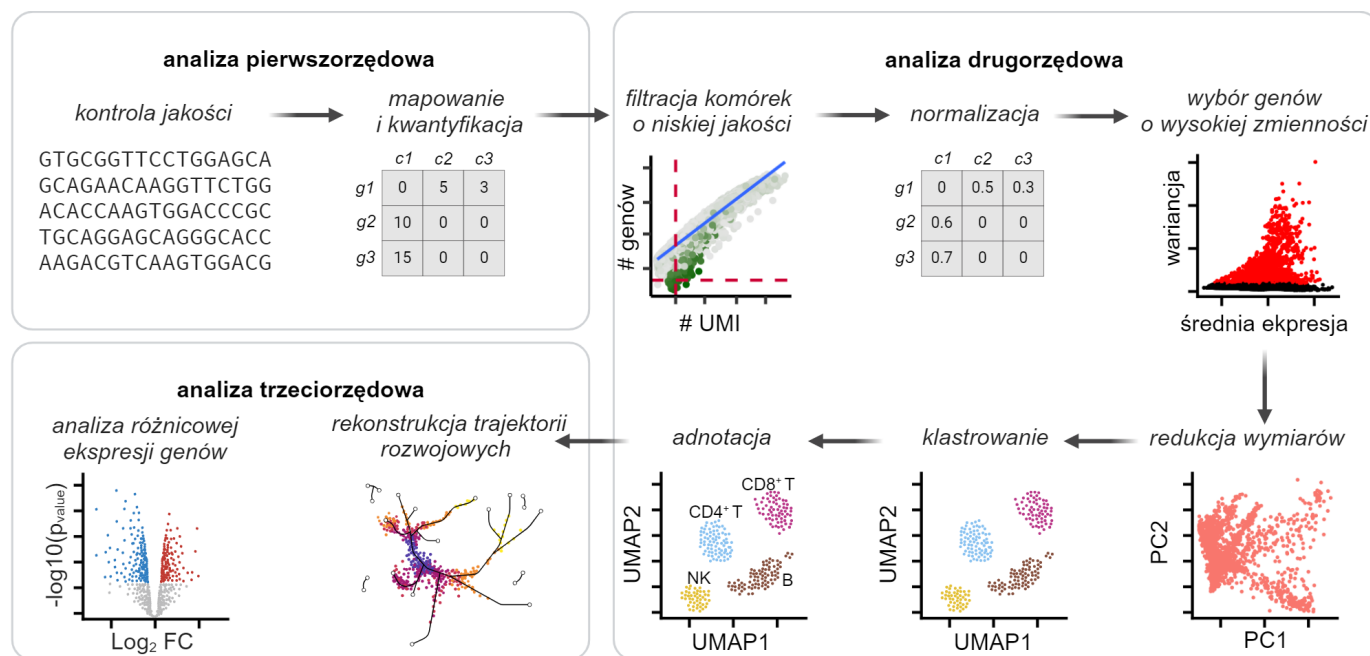
Coraz większa dostępność metod eksperymentalnych sprawia, że badania transkryptomów pojedynczych komórek cieszą się stale rosnącą popularnością. Bioinformatyczna analiza danych scRNA-Seq to wieloetapowy proces, który ze względu na złożoność i rozmiar uzyskanych danych często stanowi duże wyzwanie. Dobór odpowiednich narzędzi i metod zależy m.in. od problemu badawczego, wykorzystanej technologii oraz charakterystyki danych. Według bazy danych scRNAtools (<https://www.scrna-tools.org/>, [2,30]) obecnie dostępnych jest 1 720 narzędzi do analizy

danych scRNA-Seq (stan na 29 lutego 2024), co podkreśla kluczowe znaczenie metod analitycznych.

W ogólnym zarysie, proces analizy danych scRNA-Seq można podzielić na trzy główne etapy: analizę pierwszo-, drugo- i trzeciorzędową (Ryc. 2) [4].

Analiza pierwszorzędowa dotyczy przetwarzania surowych danych sekwencjonowania (FASTQ), w tym kontroli jakości odczytów oraz mapowania (przyrównania) odczytów do sekwencji referencyjnej. Filtrowanie danych o niskiej jakości umożliwia uzyskanie bardziej precyzyjnych i wiarygodnych wyników na późniejszych etapach analizy. Sekwencje znaczników komórkowych (ang. *barcode*) i molekularnych (UMI) zawarte w danych sekwencjonowania są wykorzystywane do stworzenia wielowymiarowej macierzy zliczeń, która przedstawia poziom ekspresji genów w poszczególnych komórkach [15,31].

Analiza drugorzędowa obejmuje szereg etapów dotyczących przetwarzania macierzy ekspresji, takich jak: kontrola jakości i filtrowanie komórek o niskiej jakości, normalizacja, selekcja genów o wysokiej zmienności (ang. *highly variable genes*, HVGs), redukcja wymiarowości danych, klastrowanie oraz adnotacja typów komórek [15,31]. Jej głównym celem jest wyselekcjonowanie komórek o wysokiej jakości oraz zniwelowanie szumu, który może przesłaniać rzeczywisty sygnał biologiczny. Analiza danych scRNA-Seq w większości opiera się na założeniu, że każda kropka zawiera RNA z nienaruszonej, pojedynczej komórki. Stąd kontrola jakości i filtrowanie macierzy ekspresji ma tu kluczowe znaczenie. Korekta macierzy uwzględnia filtrowanie pustych kropek (zawierających wyłącznie wolne, bezkomórkowe RNA), kropek zawierających więcej niż jedną komórkę (tzw. dubletów) oraz komórek o niskiej jakości, takich jak martwe lub uszkodzone. Identyfikacja tych ostatnich prowadzona jest w oparciu o podstawowe wskaźniki kontroli jakości



Rycina 2. Etapy bioinformatycznej analizy danych sekwencjonowania transkryptomu pojedynczych komórek.

m.in. liczbę genów i liczbę transkryptów (UMI) w komórce oraz procentowy udział genów mitochondrialnych. Aby skutecznie zidentyfikować komórki o niskiej jakości, konieczne jest, by te metryki były analizowane i oceniane łącznie, dla każdej próbki niezależnie. Mimo, że w literaturze podawane są przykładowe graniczne wartości liczbowe dla każdego ze wspomnianych parametrów, należy pamiętać, że stanowią one ogólne wytyczne i sprawdzają się tylko dla niektórych rodzajów badanego materiału biologicznego. Niska liczba genów i UMI przypisanych do danego znacznika komórki może świadczyć o uszkodzeniu tej komórki lub o tym, że w rzeczywistości do znacznika przypisane zostały przypadkowe transkrypty uwolnione do próbki z martwych komórek. W takim przypadku może wzrastać także procentowy udział genów mitochondrialnych (po rozpadzie komórek uwolnione z nich mitochondria zachowują integralność i to one stanowią materiał do procedury eksperymentalnej). Z kolei wysoka liczba genów i UMI może wskazywać (choć nie musi), że takie dane pochodzą z więcej niż jednej komórki. Dobór progów filtrowania zależy od rodzaju danych, jakości i rodzaju komórek oraz problemu badawczego. W ustaleniu kryteriów filtrowania pomocne jest wykreślenie rozkładu dla poszczególnych metryk, na przykład przy użyciu histogramu czy wykresu pudełkowego, a także analiza korelacji [15,31]. Normalizacja danych stanowi kolejny istotny etap analizy. Różnice w wielkości komórek oraz nierównomierność sekwencjonowania skutkują zróżnicowaną liczbą genów w komórkach badanej próbki. Dlatego, aby umożliwić porównanie ze sobą profili ekspresji całej puli komórek w badanej próbce, konieczna jest wcześniejsza normalizacja danych. Wybór genów o wysokiej zmienności opiera się na założeniu, że większość genów w każdej komórce odpowiada za jej metabolizm podstawowy, a ich poziom ekspresji jest stały. Skoncentrowanie analizy na genach wykazujących znaczne różnice w ekspresji między komórkami pozwala zatem zidentyfikować geny najbardziej istotne dla danego eksperymentu, jak również przyspiesza analizę tak złożonych danych. Stosowane metody selekcji powinny jako istotne traktować geny, które są kluczowe dla zmienności biologicznej w zestawie danych, skupiając się na tych, które różnią się między typami komórek w badanej próbce. Wielowymiarowe dane scRNA-Seq stanowią ogromne wyzwanie zarówno pod względem obliczeniowym jak i w interpretacji wyników. Dlatego kolejny krok to redukcja wymiarowości danych. W analizie scRNA-Seq popularną techniką redukcji wymiarowości danych jest analiza składowych głównych (ang. *principal component analysis*, PCA). Wykorzystuje ona fakt, że grupy genów mają zbliżone wzorce ekspresji w określonych warunkach, dzięki czemu nie ma potrzeby szczegółowego śledzenia każdego z nich z osobna. Możliwość zastąpienia skorelowanych cech przez nieskorelowane składowe główne pozwala na zmniejszenie wymiarowości zbioru danych, jednocześnie zachowując istotne różnice między danymi. Techniki redukcji wymiarowości można zastosować do wizualizacji lub podsumowania podstawowej topologii danych. Złożoną strukturę danych scRNA-Seq trudno jest przedstawić przy pomocy dwóch lub trzech składowych głównych. Stąd, aby uzyskać dwuwymiarową projekcję analizowanego zbioru danych wykorzystuje się wybrany wcześniej zestaw składowych głównych w algorytmach

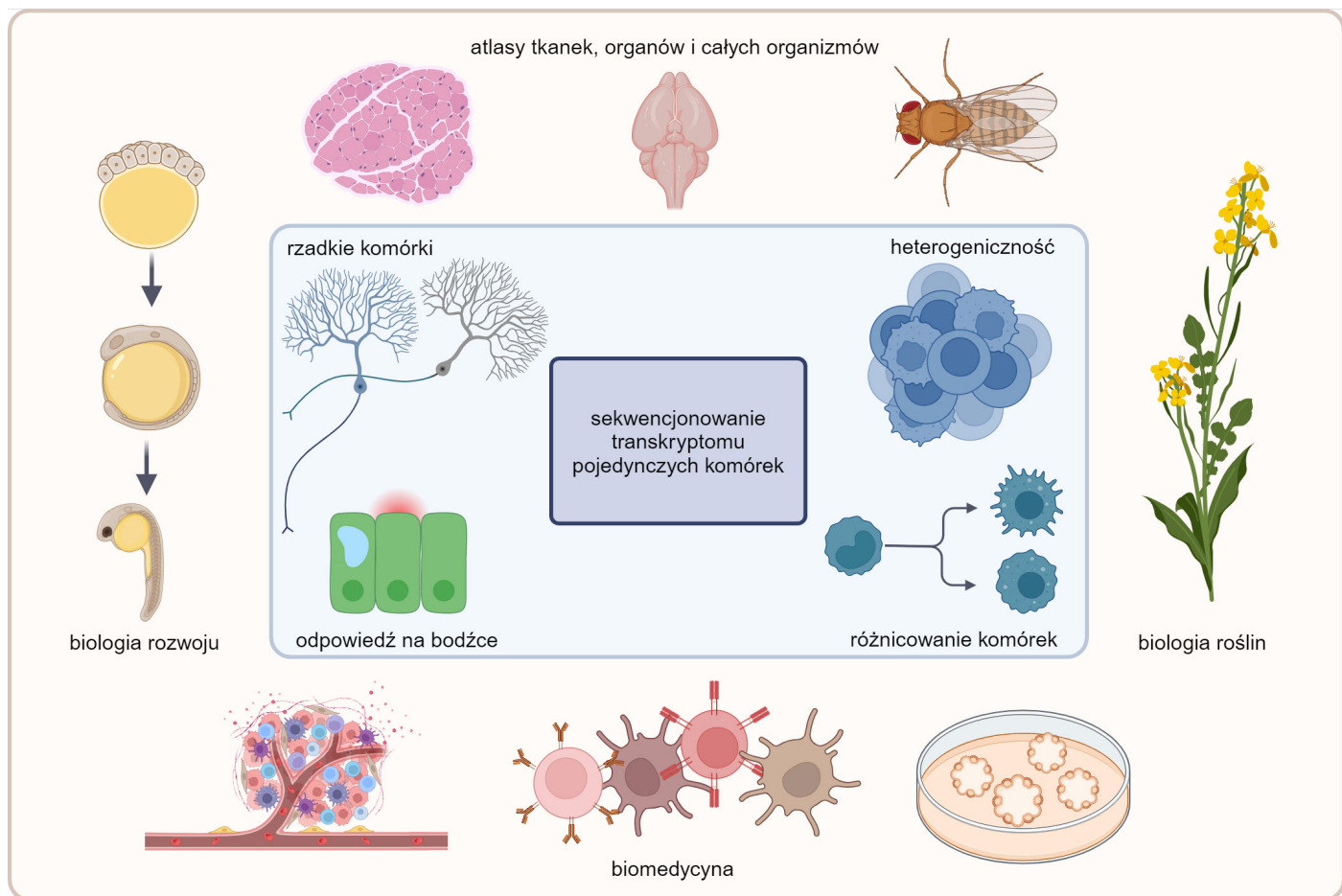
takich jak UMAP (ang. *uniform manifold approximation and projection*) lub t-SNE (ang. *t-distributed stochastic neighbor embedding*) [15,31]. Są to najpowszechniej używane narzędzia do wizualizacji danych w postaci subpopulacji składających się z pojedynczych komórek, które łączy podobny profil transkryptomyczny. Ich zastosowanie ułatwia dalsze etapy analizy i eksplorację danych.

Kolejny etap obejmuje klastrowanie i adnotację wyodrębnionych subpopulacji komórek. Klastrowanie ma kluczowe znaczenie dla większości analiz, ponieważ identyfikuje grupy komórek o podobnych profilach ekspresji. Niektóre z tych grup mogą reprezentować różne typy komórek, a inne można uznać za pośrednie stany komórkowe, w zależności od danego układu biologicznego. Adnotacja to proces przypisania zidentyfikowanym skupiskom komórek nazw biologicznych, np. typu komórki. Stosuje się tu podejścia ręczne lub automatyczne. Adnotacja automatyczna wykorzystuje publicznie dostępne zestawy danych, które zawierają informacje o znanych typach komórek, ich markerach genetycznych czy funkcjach. Proces ten polega na dopasowaniu profili ekspresji komórek w analizowanej próbce do znanych typów komórek z bazy danych. Najbardziej popularną bazą komórek ludzkich jest Human Cell Atlas (HCA, <https://www.humancellatlas.org/>). Adnotacja automatyczna zwykle nie jest wystarczająca i można ją traktować jako etap wstępny. Złoty standardem jest adnotacja ręczna. W tym przypadku poszczególne subpopulacje komórek opisuje się w oparciu o geny markerowe zidentyfikowane dla każdego klastra. Zaleca się, o ile to możliwe, wykorzystanie zarówno adnotacji automatycznej jak i ręcznej w celu uzyskania jak najbardziej rzetelnych i wiarygodnych wyników [15,31].

Analiza trzeciorzędowa to etap, który w największym stopniu uzależniony jest od konkretnego pytania biologicznego postawionego na początku eksperymentu. Obejmuje m.in. analizę różnicowej ekspresji genów, analizy funkcjonalne, rekonstrukcję trajektorii rozwojowych komórek, identyfikację oddziaływań międzykomórkowych i wiele innych [4,15,31].

PODSTAWOWE TECHNOLOGIE

Technologie transkryptomiki pojedynczej komórki różnią się między sobą przede wszystkim sposobem separacji pojedynczych komórek i znakowania transkryptów. Najbardziej popularne obecnie są metody kroplowe, wykorzystujące układy mikroprzepływowe do wytworzenia emulsji typu woda w oleju. Krople generowane są w taki sposób, aby znalazły się w nich hydrożelowa lub polistyrenowa kulka z przytwierdzonym do niej oligonukleotydowym znacznikiem połączonym z ogonem oligo(dT) oraz pojedyncza komórka. W kolejnych krokach w obrębie każdej kropli dochodzi do lizy komórki, uwolnienia transkryptów, ich wychwycenia i wyznakowania oraz odwrotnej transkrypcji. Zaletą tego podejścia jest jego wysoka przepustowość. Obecnie rozwiązania oferowane przez wiodącego dostawcę kroplowych technologii scRNA-Seq, umożliwiają enkapsulację (wychwycenie i zamknięcie w obrębie kropli) nawet do 20 000 komórek lub jąder komórkowych w jednej próbce. Ograniczenia metod wykorzystujących mikrokapilary wynikają z trudności z wychwyceniem komórek o dużych



Rycina 3. Zastosowania sekwencjonowania transkryptomu pojedynczych komórek.

rozmiarach i/lub nietypowym kształcie, które mogłyby spowodować zablokowanie układu przepływowego. Dla tego typu próbek, a także dla otoczonych ścianą komórkową komórek roślinnych, zaleca się podejście polegające na izolacji i enkapsulacji jąder komórkowych (ang. *single nuclei RNA sequencing*, snRNA-Seq). Metody kroplowe wykorzystują specjalne czipy i urządzenia wyposażone w systemy pomp. Niestety, z uwagi na ich koszt, dostęp do nich jest ograniczony. Laboratoria świadczące usługi przygotowania bibliotek do sekwencjonowania z wykorzystaniem metod kroplowych często zlokalizowane są w pewnej odległości od miejsca prowadzenia badań, co sprawia, że dostarczenie próbek o odpowiedniej jakości, czyli zawiesiny pojedynczych komórek charakteryzującej się wysokim odsetkiem komórek żywych, jest problematyczne. W takim przypadku rozwiązaniem może być zastosowanie metody płytkowej, wykorzystującej indeksowanie kombinatoryczne. Brak wymagań co do specjalistycznego sprzętu oraz ograniczeń w kwestii rozmiaru czy kształtu komórek sprawia, że metoda ta może być zastosowana praktycznie w każdym laboratorium. Dodatkową zaletą jest praca z komórkami utrwalonymi, dzięki czemu możliwe jest zebranie próbek w różnych punktach czasowych, utrwalenie ich i późniejsza analiza w ramach jednego eksperymentu. Utrwalone i permeabilizowane komórki rozdziela się do dołków mikroplątki w ilości do kilkunastu tysięcy komórek na dołek. Proces odwrotnej transkrypcji i przyłączenia znacznika komórkowego wraz

z oligonukleotydowym łącznikiem odbywa się *in situ* wewnątrz komórek. Po nim, wszystkie komórki są zbierane i mieszane, a następnie ponownie rozdzielane do dołków mikroplątki. W dwóch kolejnych rundach ligacji (po każdej następuje zebranie, zmieszanie i ponowne rozdzielanie komórek do dołków) dodawane są dwa dodatkowe znaczniki, UMI oraz adaptory umożliwiające przeprowadzenie NGS. Dopiero po tym etapie zebrane komórki poddawane są lizie, a wyznakowane cząsteczki cDNA podlegają amplifikacji i konstruowana jest właściwa biblioteka do sekwencjonowania. Z każdą rundą znakowania wykładniczo rośnie liczba kombinacji przyłączanych znaczników. Metody płytkowe cechują się wyjątkowo wysoką przepustowością, umożliwiając zbadanie od 10 000 do nawet 1 000 000 komórek. Niekiedy jednak materiał do badań jest ograniczony ilościowo. Wówczas najlepszym rozwiązaniem może okazać się technika, w której od kilkudziesięciu do kilkuset pojedynczych komórek jest umieszczanych za pomocą sortera w dołkach mikroplątki, albo przy użyciu precyzyjnych dyspenserów w zagłębieniach na specjalnym chipie. Nowoczesne nanodyspensery wyposażone w moduł do obrazowania umożliwiają wybranie jedynie tych dołków, w których znalazły się żywe komórki. Następnie automatycznie wykonują kolejne etapy lizy komórek, znakowania transkryptów i syntezy cDNA oraz konstrukcji i amplifikacji bibliotek, zużywając niewielkie objętości odczynników.

ZASTOSOWANIA

Transkryptomika pojedynczych komórek umożliwia eksplorację heterogeniczności komórkowej, zbadanie rzadkich populacji komórek w złożonych układach tkankowych i odkrywanie nowych populacji. Dostarcza kompleksowego obrazu obecnych w próbce typów i podtypów komórek wraz z informacją o ich profilach transkrypcyjnych, a także daje możliwość śledzenia trajektorii rozwojowych czy progresji choroby. Pozwala również na badanie dynamiki ekspresji genów, zmian stanu komórek i ich odpowiedzi na bodźce. Bogactwo informacji dostarczanych przez analizy scRNA-Seq oraz możliwość wykorzystania tych metod do badania komórek z hodowli *in vitro*, organoidów, fragmentów tkanek ludzkich, zwierzęcych i roślinnych, a nawet całych organizmów, decydują o szerokim spektrum zastosowań tej technologii (Ryc. 3).

TWORZENIE ATLASÓW TKANEK, ORGANÓW I CAŁYCH ORGANIZMÓW

Atlasy tkanek, organów, czy całych organizmów stanowią nieocenione źródło wiedzy, ukazując zarówno różnorodność, jak i hierarchię komórek w danym układzie biologicznym. Tworzenie i utrzymywanie takich atlasów to obecnie jeden z kluczowych aspektów skupiających międzynarodowe wysiłki różnych konsorcjów, takich jak Human Cell Atlas (HCA, <https://www.humancellatlas.org/>). Ten rozbudowany atlas komórek ludzkich stanowi

obecnie istotne narzędzie nie tylko w badaniach podstawowych procesów biologicznych, lecz także w diagnostyce i leczeniu chorób, zwłaszcza tych związanych z zaburzeniami regulacji ekspresji genów w konkretnych komórkach [32]. Oprócz niego dostępny jest bogaty zestaw innych atlasów komórek człowieka, uwzględniający różne etapy rozwojowe, w tym także embrionalny i płodowy. Szczegółowe zestawienie tych atlasów znajdzie Czytelnik w pracy Ye i wsp. [33]. Ponadto, w ostatnich kilkunastu latach powstało wiele atlasów, obejmujących różne zwierzęta i rośliny, m.in. organizmy modelowe i ważne z ekonomicznego punktu widzenia (Tabela 1) [33,34]. Atlasy referencyjne umożliwiają automatyzację adnotacji nowych zbiorów danych i prowadzenie analiz porównawczych różnych tkanek czy stanów chorobowych. Niemniej jednak, różnorodność protokołów eksperymentalnych i metod użytych do generowania danych przez poszczególne laboratoria wprowadza znaczącą zmienność techniczną. W celu jej zniwelowania konieczne jest zastosowanie zaawansowanych technik integracji danych. W erze wielkoskalowych, multiomicznych i integracyjnych metaanaliz, atlasy pojedynczych komórek dostarczają nowych informacji, umożliwiając identyfikację związków między genami, chorobami, a terapiami na poziomie komórkowym i tkankowym [33].

Tabela 1. Wybrane atlasy transkryptomiczne pojedynczych komórek zwierząt i roślin (na podstawie [33,34]).

Nazwa łacińska	Nazwa polska	Przynależność taksonomiczna	Rodzaj atlasu	Źródło
<i>Nematostella vectensis</i>	zawilec morski	parzydełkowce	cały organizm	[44]
<i>Hydra vulgaris</i>	stulbia pospolita	parzydełkowce	cały organizm	[45]
<i>Clytia hemisphaerica</i>		parzydełkowce	cały organizm	[46]
<i>Xenia sp.</i>		parzydełkowce	cały organizm	[47]
<i>Stylophora pistillata</i>		parzydełkowce	cały organizm	[48]
<i>Spongilla lacustris</i>	nadecznik stawowy	gąbki	cały organizm	[49]
<i>Schmidtea mediterranea</i>		robaki płaskie	cały organizm	[50]
<i>Caenorhabditis elegans</i>		nicienie	cały organizm	[51]
<i>Eisenia andrei</i>		pierścienice	cały organizm	[52]
<i>Drosophila melanogaster</i>	muszka owocowa	owady	cały organizm	[53]
<i>Ciona intestinalis</i>	przejrystka	żachwy	cały organizm	[54]
<i>Danio rerio</i>	danio pręgowany	ryby	cały organizm	[52]
<i>Xenopus laevis</i>	platana szponiasta	plazy	cały organizm (stadia larwalne), 17 tkanek (osobniki dorosłe)	[55]
<i>Gallus gallus domesticus</i>	kura domowa	ptaki	cały organizm (embrion)	[56]
<i>Sus scrofa domesticus</i>	świnia domowa	ssaki	20 tkanek	[57]
<i>Bos taurus</i>	bydło domowe	ssaki	10 tkanek	[58]
<i>Mus musculus</i>	mysz domowa	ssaki	51 tkanek, organów i kultur komórkowych	[59]
<i>Macaca fascicularis</i>	makak krabozerny	ssaki	45 tkanek	[60]
<i>Arabidopsis thaliana</i>	rzodkiewnik pospolity	kapustowate	korzeń	[61]
<i>Oryza sativa</i>	ryż siewny	wiechlinowate	siewka (fragmenty pędu i korzenia)	[62]
<i>Camellia sinensis var. sinensis</i>	herbata chińska	herbatowate	liść	[43]
<i>Arachis hypogaea</i>	orzech ziemny	bobowate	liść	[63]
<i>Populus alba</i>	topola biała	wierzbowate	łodyga	[64]

Transkryptomika pojedynczych komórek stała się rewolucyjnym narzędziem biomedycyny. Analizy scRNA-Seq są na szeroką skalę stosowane w onkologii, immunologii, neurobiologii i kardiologii. Wykorzystuje się je by poznać mechanizmy procesu starzenia oraz rozwoju chorób metabolicznych i zakaźnych. Ponadto, informacje zebrane dzięki wykonaniu scRNA-Seq, przyczyniają się do rozwoju inżynierii tkankowej, odkrywania i opracowywania nowych leków oraz rozwoju podejść medycyny spersonalizowanej [2,4,22,24,35]

scRNA-Seq w badaniu nowotworów

Sekwencjonowanie transkryptomu pojedynczych komórek pozwoliło naukowcom i klinicyście przyjrzeć się bliżej tajnikom inicjacji i progresji nowotworu, powstawaniu przerzutów oraz odpowiedzi na leczenie. Informacje zdobyte dzięki zastosowaniu scRNA-Seq mogą posłużyć do opracowania nowych strategii terapeutycznych.

W guzach nowotworowych komórki złośliwe, komórki niezłośliwe i czynniki pozakomórkowe oddziałują synergistycznie, tworząc skomplikowany „ekosystem guza”. Składa się on z dwóch części: komórek nowotworowych (złośliwych) i mikrośrodowiska guza (ang. *tumor microenvironment*, TME), tworzonego przez komórki niezłośliwe (m.in. limfocyty T, limfocyty B, komórki mieloidalne, fibroblasty, osteoblasty, komórki śródbłonna, adipocyty) i składniki macierzy zewnątrzkomórkowej [36]. Oprócz oczywistych badań ukierunkowanych na opisanie heterogeniczności komórkowej samego guza, cennych informacji dostarczają także analizy jego mikrośrodowiska [36]. Wykazano, że składniki TME mają wpływ na rozwój guza, między innymi przez bezpośrednie oddziaływania międzykomórkowe oraz wydzielanie cząsteczek będących ligandami dla receptorów powierzchniowych komórek nowotworowych, mogącymi zmieniać określone szlaki sygnałowe tych komórek. Transkryptomika pojedynczych komórek umożliwia nie tylko identyfikację typów komórek obecnych w ekosystemie guza, ale również rozróżnienie ich stanów. Dostarcza tym samym informacji o etapach cyklu komórkowego, procesach metabolicznych zachodzących w komórkach TME i uruchomionych szlakach sygnałowych, co ma zarówno istotne implikacje biologiczne, jak i kliniczne [24,36].

Analizy scRNA-Seq, ze względu na oferowaną możliwość wykrywania rzadkich populacji komórek, stanowią doskonałe narzędzie do badania krążących komórek nowotworowych (ang. *circulating tumor cells*, CTCs), czyli uwolnionych z guza pierwotnego komórek nowotworowych obecnych we krwi obwodowej. CTCs są istotnym czynnikiem powstawania przerzutów, a ich wczesne wykrycie we krwi jest ważną wskazówką dla rokowania, oceny skuteczności terapii i spersonalizowanego leczenia pacjentów [35].

scRNA-Seq w immunologii

Jednym z głównych zastosowań transkryptomiki pojedynczych komórek jest badanie różnorodności typów i stanów funkcjonalnych komórek odpornościowych oraz

ich oddziaływań między sobą i z innymi komórkami. Pozwala to lepiej zbadać procesy prawidłowej i patologicznej odpowiedzi immunologicznej [2,24]. ScRNA-Seq jest szczególnie potężnym narzędziem do zgłębiania różnorodności i cech receptorów limfocytów T i B. Te kompleksy białkowe są kluczowymi elementami odporności nabytej, odpowiedzialnymi za rozpoznawanie i reagowanie na konkretne antygeny. Wykazują zatem wysoką różnorodność, generowaną m.in. dzięki zjawisku rekombinacji somatycznej, podczas której następuje rearanżacja segmentów genów kodujących białka tworzące receptory (tzw. rekombinacja V(D)J). Dzięki wykorzystaniu scRNA-Seq do badania tych receptorów, naukowcy mogą lepiej scharakteryzować ogromny zbiór sekwencji kodujących receptory, a także wzorce ekspresji genów związane z ich aktywacją i funkcją, czy też śledzić klonalność limfocytów B i T. Na szczególną uwagę zasługują tutaj metody, dzięki którym możliwe jest przygotowanie dwóch rodzajów bibliotek dla tych samych komórek. Jedna biblioteka umożliwia wówczas określenie profilu transkryptomycznego, a druga zbadanie segmentów V(D)J. Informacje te można połączyć na etapie analizy danych i przypisać komórce o danym układzie V(D)J określony profil ekspresji genów.

W licznych badaniach naukowych analizuje się zmiany ekspresji genów w komórkach odpornościowych przed i po stymulacji lub w kontekście choroby, w celu opisanego mechanizmów molekularnych związanych z aktywacją immunologiczną, różnicowaniem i funkcją tych komórek. Te informacje pomagają w zrozumieniu odpowiedzi immunologicznej na patogeny, choroby autoimmunologiczne, nowotwory i inne zaburzenia odpornościowe [2,24]. Przykładowo, analizy scRNA-Seq jednojądrzastych komórek krwi obwodowej oraz komórek wyizolowanych z popłuczyn pęcherzykowo-oskrzelikowych wykazały, że SARS-CoV-2, wywołujący chorobę zakaźną COVID-19, indukuje unikatowe szlaki sygnałowe w komórkach układu odpornościowego. Wykazano również, że proporcje poszczególnych typów komórek odpornościowych są odmienne u rekonwalescentów i pacjentów w umiarkowanym oraz ciężkim stadium choroby. Dodatkowo, przeprowadzono badania jąder komórkowych wyizolowanych z tkanek płuc, nerek, wątroby oraz serca, pochodzących z autopsji pacjentów zmarłych wskutek ciężkiego przebiegu COVID-19. Ujawniły one zmiany strukturalne tkanki płucnej, zaburzenia różnicowania się komórek pęcherzykowych, ekspansję fibroblastów oraz zmiany transkrypcyjne w komórkach pozostałych organów [4].

scRNA-Seq w neurobiologii

Badania transkryptomyczne pojedynczych komórek zrewolucjonizowały odkrywanie złożoności budowy mózgu. Większość analiz w tym zakresie wykonano wykorzystując mózgi myszy na etapie rozwoju embrionalnego oraz dojrzałych. Skonstruowano kompleksowe atlasy komórek z różnych regionów mózgu oraz wykazano istnienie różnicowania transkryptomycznego między neuronami a komórkami spoza układu nerwowego [22]. Wyniki te służą jako istotny punkt odniesienia dla badań nad patologiami mózgu, w tym chorobami neurodegeneracyjnymi, zaburzeniami neurologicznymi, chorobami psychicznymi i guzami

mózgu, podkreślając przydatność transkryptomiki pojedynczych komórek RNA w poszerzaniu naszego zrozumienia zaburzeń mózgu [22].

scRNA-Seq przyczyniło się do poznania budowy i funkcjonowania mózgu, szczególnie w kontekście selektywnej wrażliwości neuronów. W większości chorób neurodegeneracyjnych istnieją określone grupy komórek, które są szczególnie podatne na degenerację, choć mechanizmy molekularne leżące u podstaw tej selektywnej wrażliwości pozostają słabo rozumiane i stanowią jedną z głównych przeszkód w leczeniu tych stanów [22]. Na przykład, w chorobie Alzheimera neurony z niektórych obszarów mózgu, takie jak neurony piramidowe warstwy II kory śródwężowej, komórki piramidowe CA1 hipokampu oraz neurony piramidowe w obszarach asocjacyjnych kory nowej, są szczególnie podatne na degenerację, co potwierdziły badania wykorzystujące snRNA-Seq. Podobne podejście zastosowano w badaniach choroby Parkinsona, gdzie zidentyfikowano subpopulację neuronów dopaminergicznych charakteryzującą się zwiększoną podatnością na degenerację. W przypadku zaburzeń neuropsychiatrycznych, takich jak schizofrenia czy zaburzenia depresyjne, zastosowano scRNA-Seq w połączeniu z informacjami o zestawach genów znanych z zaburzeń neuropsychiatrycznych, co pozwoliło na wyizolowanie komórek i podtypów neuronów zaangażowanych w te zaburzenia [22].

Biologia rozwoju

Podstawowym celem biologii rozwoju jest poznanie ścieżek różnicowania poszczególnych typów komórek. Transkryptomika pojedynczych komórek umożliwia uchwycenie komórek znajdujących się na różnych etapach rozwoju w jednym eksperymencie. Dzięki temu scRNA-Seq pozwala na eksplorację i lepsze zrozumienie procesów leżących u podstaw tworzenia tkanek i narządów. Jednym z jej głównych zastosowań w badaniu biologii rozwoju jest identyfikacja zdarzeń mających wpływ na różnicowanie poszczególnych linii komórkowych i śledzenie rozwoju tych linii. Analizując transkryptomy poszczególnych komórek na różnych etapach rozwoju, można śledzić ekspresję markerów specyficznych dla linii i identyfikować odrębne populacje komórek odpowiadających różnym liniom. Pomaga to opisać trajektorie rozwoju poszczególnych linii i relacje między nimi oraz poznać ich hierarchię [2]. Ponadto, porównując profile ekspresji genów komórek na różnych etapach rozwoju danej linii, można wyznaczyć potencjalne regulatory i szlaki sygnałowe procesów determinujących losy komórek. Z kolei profilowanie ekspresji genów podczas kształtowania i organizacji tkanek w procesie morfogenezy, pozwala zidentyfikować geny i szlaki zaangażowane w adhezję i migrację komórek.

W ostatnich latach wykonano wiele szeroko zakrojonych badań rozwoju organizmów modelowych, takich jak m.in. makak, mysz, danio pręgowany, muszka owocowa i nicien *Caenorhabditis elegans* [33,37,38]. Podjęto także szereg inicjatyw prowadzących do stworzenia kompleksowego atlasu rozwoju człowieka, obejmującego wszystkie etapy od zapłodnionego oocyta do urodzenia. Ponadto, scRNA-Seq wykorzystuje się w analizach przebiegu procesów rozwojo-

wych u roślin. Posłużono się nimi na przykład do opisanego rozwoju tkanki przewodzącej – ksylemu, u czterech gatunków roślin okrytozalążkowych w kontekście ewolucyjnym [39].

Wiedza płynąca z badań transkryptomicznych ukierunkowanych w dalszej perspektywie na zrozumienie biologii rozwoju organizmu znajduje zastosowania w biologii komórek macierzystych i medycynie regeneracyjnej. Przykładowo, technologia scRNA-Seq przyczyniła się do lepszego poznania rozwoju komórek β trzustki i patologii cukrzycy. Badania modeli mysich, a także ludzkich zarodkowych komórek macierzystych i indukowanych komórek pluripotencjalnych doprowadziły, między innymi, do odkrycia nowego markera specyficznego dla komórek α trzustki. Pozwoliły także na opisanie heterogeniczności komórek β i zmian transkrypcyjnych zachodzących w nich podczas rozwoju [4]. Ponadto, dzięki eksperymentom scRNA-Seq zidentyfikowano ponad 30 genów związanych z wrodzonymi wadami serca oraz wykazano kluczową rolę szlaków sygnałowych Notch i BMP w różnicowaniu i dojrzewaniu komórek serca [37].

BIOLOGIA ROŚLIN

Mimo, że większość eksperymentów angażujących transkryptomikę pojedynczych komórek prowadzona jest nadal z wykorzystaniem komórek pochodzenia zwierzęcego i ludzkiego, technologie te cieszą się coraz większym zainteresowaniem w badaniu roślin. W tym przypadku największym wyzwaniem nadal pozostaje przygotowanie materiału o odpowiedniej jakości – ze względu na obecność ściany komórkowej, stosuje się podejścia oparte na analizie protoplastów bądź izolowanych jąder komórkowych. Dotychczas metodą scRNA-Seq posługiwano się głównie w pracach badawczych poświęconych rzodkiewnikowi, jako najlepiej poznanej roślinie modelowej, niemniej jednak transkryptomika pojedynczych komórek coraz częściej znajduje zastosowanie w badaniach innych gatunków roślin, takich jak ryż, kukurydza, bawełna, pomidor, topola oraz pszenica [34,40]. Wykorzystanie scRNA-Seq pozwoliło na lepsze zrozumienie komórkowo-specyficznej odpowiedzi na stres abiotyczny i sygnalizacji międzykomórkowej [41]. Na przykład, wykazano, że w przypadku łagodnego stresu suszy znaczące obniżenie ekspresji genów ma miejsce głównie w komórkach mezofilu, podczas gdy w epidermie ekspresja większości genów jest aktywowana. Ponadto, analizy scRNA-Seq zastosowano do zbadania złożonych procesów rozwojowych zachodzących podczas dojrzewania pyłku, rozwoju tkanek przewodzących, różnicowania tkanek korzenia i liści [40,41]. Przykładowo, stworzono atlas różnicującego się drewna u topoli kanadyjskiej, obejmujący komórki naczyń, komórki włókienkowe, komórki miększu promieniowego oraz komórki prekursorowe drewna oraz zidentyfikowano potencjalne czynniki transkrypcyjne zaangażowane w kontrolę różnicowania komórek drewna [40]. Transkryptomika pojedynczych komórek dostarczyła również cennych informacji na temat szlaków biosyntezy metabolitów u roślin. Na przykład, scRNA-Seq umożliwiło powiązanie określonego typu komórek korony kwiatów dzikiego tytoniu z syntezą substancji zapachowej – benzylacetone [42]. Dodatkowo, analiza koekspresji doprowadziła

do wytypowania potencjalnych genów zaangażowanych w szlak biosyntezy tego związku. Z kolei badania herbaty chińskiej przyczyniły się do odkrycia nowego szlaku metabolicznego estrów katechin [43].

Technologia scRNA-Seq jest coraz częściej wykorzystywana w badaniach rozwoju i odpowiedzi na stres u roślin uprawnych, takich jak ryż i kukurydza, co w przyszłości może pomóc hodowcom zaprojektować lepsze strategie uprawy [41].

PODSUMOWANIE I PERSPEKTYWY

Technologie ukierunkowane na analizy pojedynczych komórek zrewolucjonizowały podejście do badań w naukach biologicznych. Dzięki ogromnemu postępowi technologicznemu, opracowaniu nowych metod wysokoprzepustowych oraz narzędzi bioinformatycznych, eksperymenty scRNA-Seq stały się bardziej dostępne niż jeszcze kilka lat temu. To z kolei sprawia, że znajdują one liczne zastosowania w odkrywaniu różnorodności komórkowej organizmów, badaniu procesów rozwojowych, śledzeniu odpowiedzi komórek na rozmaite czynniki i szeroko pojętej biomedycynie. W przeciwieństwie do tradycyjnej analizy typu masowego, która mierzy ekspresję genów w populacjach komórkowych jako całości, scRNA-Seq umożliwia identyfikację różnic w ekspresji genów między poszczególnymi komórkami w obrębie jednej próbki. Podejścia oparte na kroplach i znakowaniu kombinatorycznym umożliwiają analizy od dziesiątek tysięcy do nawet miliona komórek w jednym eksperymencie, co pozwala na dokładne zbadanie wysoce złożonych układów tkankowych. Wciąż jednak metody transkryptomiki pojedynczych komórek mają pewne ograniczenia. Jednym z nich jest utrata informacji o organizacji przestrzennej badanej tkanki. Dodatkowo, istnieją obawy, że procesy związane z trawieniem tkanek i izolacją komórek, mogą zmienić ekspresję genów i wpływać na obecność określonych typów komórek.

Z tego względu obecnie dużym zainteresowaniem cieszą się podejścia transkryptomiki przestrzennej umożliwiające profilowanie ekspresji genów na poziomie pojedynczych komórek z jednoczesnym zachowaniem informacji o lokalizacji w strukturze przestrzennej tkanki. Ponadto, mimo że koszt eksperymentu scRNA-Seq w przeliczeniu na komórkę został obniżony, koszt analizy w przeliczeniu na próbkę nadal jest stosunkowo wysoki. Ogranicza to możliwości wprowadzenia scRNA-Seq jako narzędzia diagnostycznego. Bardzo istotny jest więc dalszy rozwój tych technik, zarówno w aspekcie eksperymentalnym, jak i obliczeniowym, który umożliwi ich upowszechnienie. Można oczekiwać, że przyszłość analiz transkryptomu pojedynczych komórek leży w ich powiązaniu z genomiką, epigenomiką, proteomiką, metabolomiką czy bioobrazowaniem.

PODZIĘKOWANIA

Ryciny i streszczenie graficzne przygotowano z użyciem BioRender.com (*figures and graphical abstract were created with BioRender.com*).

PIŚMIENNICTWO

1. Lowe R, Shirley N, Bleackley M, Dolan S, Shafee T (2017) Transcriptomics technologies. *PLoS Comput Biol* 13(5): e1005457
2. Qu HQ, Kao C, Hakonarson H (2024) Single-Cell RNA Sequencing Technology Landscape in 2023. *Stem Cells* 42(1):1-12
3. Tang F, Lao K, Surani MA (2011) Development and applications of single-cell transcriptome analysis. *Nat Methods* 8(4 Suppl):S6-11
4. Jovic D, Liang X, Zeng H, Lin L, Xu F, Luo Y (2022) Single-cell RNA sequencing technologies and applications: A brief overview. *Clin Transl Med* 12(3):e694
5. Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, Wang X, Bodeau J, Touch BB, Siddiqui A, Lao K, Surani MA (2009) mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 6:377-382
6. Jaitin DA, Kenigsberg E, Keren-Shaul H, Elefant N, Paul F, Zaretzky I, Mildner A, Cohen N, Jung S, Tanay A, Amit I (2014) Massively parallel single cell RNA-Seq for marker-free decomposition of tissues into cell types. *Science* 343:776
7. Muraro MJ, Dharmadhikari G, Grün D, Groen N, Dielen T, Jansen E, van Gurp L, Engelse MA, Carlotti F, de Koning E, van Oudenaarden A (2016) A Single-Cell Transcriptome Atlas of the Human Pancreas. *Cell Syst* 3:385-394.e3
8. DeLaughter DM (2018) The use of the Fluidigm C1 for RNA expression analyses of single cells. *Curr Protoc Mol Biol* 122:e55
9. Hochgerner H, Lönnerberg P, Hodge R, Mikes J, Heskol A, Hubschle H, Lin P, Picelli S, La Manno G, Ratz M, Dunne J, Husain S, Lein E, Srinivasan M, Zeisel A, Linnarsson S (2017) STRT-seq-2i: dual-index 5-single cell and nucleus RNA-seq on an addressable microwell array. *Sci Rep* 7(1):16327
10. Islam S, Kjällquist U, Moliner A, Zajac P, Fan JB, Lönnerberg P, Linnarsson S (2011) Characterization of the single-cell transcriptional landscape by highly multiplex RNA-seq. *Genome Res* 21:1160-1167
11. Cao J, Packer JS, Ramani V, Cusanovich DA, Huynh C, Daza R, Qiu X, Lee C, Furlan SN, Steemers FJ, Adey A, Waterston RH, Trapnell C, Shendure J (2017) Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science* 357:661-667
12. Rosenberg AB, Roco CM, Muscat RA, Kuchina A, Sample P, Yao Z, Graybuck LT, Peeler DJ, Mukherjee S, Chen W, Pun SH, Sellers DL, Tasic B, Seelig G (2018) Single-cell profiling of the developing mouse brain and spinal cord with split-pool barcoding. *Science* 360:176-182
13. Salmen F, De Jonghe J, Kaminski TS, Alemany A, Parada GE, Verity-Legg J, Yanagida A, Kohler TN, Battich N, van den Brekel F, Ellermann AL, Arias AM, Nichols J, Hemberg M, Hollfelder F, van Oudenaarden A (2022) High-throughput total RNA sequencing in single cells using VASA-seq. *Nat Biotechnol* 40:1780-1793
14. Roco C, O'Connor E, Hills E, Tran V, Rosenberg A (2020) Scalable and Highly Sensitive Single Cell RNA-seq through Combinatorial Barcoding. *J Biomol Tech* 31:S16
15. Andrews TS, Kiselev VY, McCarthy D, Hemberg M (2021) Tutorial: guidelines for the computational analysis of single-cell RNA sequencing data. *Nat Protoc* 16(1):1-9
16. Ziegenhain C, Vieth B, Parekh S, Reinius B, Guillaumet-Adkins A, Smets M, Leonhardt H, Heyn H, Hellmann I, Enard W (2017) Comparative Analysis of Single-Cell RNA Sequencing Methods. *Mol Cell* 65:631-643.e4
17. Lafzi A, Moutinho C, Picelli S, Heyn H (2018) Tutorial: guidelines for the experimental design of single-cell RNA sequencing studies. *Nat Protoc* 13:2742-2757
18. Denisenko E, Guo BB, Jones M, Hou R, De Kock L, Lassmann T, Poppe D, Poppe D, Clément O, Simmons RK, Lister R, Forrest A (2020) Systematic assessment of tissue dissociation and storage biases in single-cell and single-nucleus RNA-seq workflows. *Genome Biol* 21:1-25
19. Van Den Brink SC, Sage F, Vértesy Á, Spanjaard B, Peterson-Maduro J, Baron CS, Robin C, Van Oudenaarden A (2017) Single-cell sequencing reveals dissociation-induced gene expression in tissue subpopulations. *Nat Methods* 14:935-936

20. Adam M, Potter AS, Potter SS (2017) Psychrophilic proteases dramatically reduce single-cell RNA-seq artifacts: a molecular atlas of kidney development. *Development* 144:3625–3632
21. Cervantes-Pérez SA, Thibivilliers S, Tennant S, Libault M (2022) Review: Challenges and perspectives in applying single nuclei RNA-seq technology in plant biology. *Plant Sci* 325:111486
22. Piwecka M, Rajewsky N, Rybak-Wolf A (2023) Single-cell and spatial transcriptomics: deciphering brain complexity in health and disease. *Nat Rev Neurol* 19(6):346–362
23. Boon WC, Petkovic-Duran K, Zhu Y, Manasseh R, Horne MK, Aumann T (2011) Increasing cDNA yields from single-cell quantities of mRNA in standard laboratory reverse transcriptase reactions using acoustic microstreaming. *J Vis Exp* (53):e3144
24. Wang S, Sun ST, Zhang XY, Ding HR, Yuan Y, He JJ, Wang MS, Yang B, Li Y (2023) The Evolution of Single-Cell RNA Sequencing Technology and Application: Progress and Perspectives. *Int J Mol Sci* 24(3):2943
25. Tian T, Zhang J, Lin X, Wei Z, Hakonarson H (2021) Model-based deep embedding for constrained clustering analysis of single cell RNA-seq data. *Nat Commun* 12:1–12
26. Shi T, Beaulieu MO, Saunders LM, Fabian P, Trapnell C, Segil N, Crump JG, Raible D (2023) Single-cell transcriptomic profiling of the zebrafish inner ear reveals molecularly distinct hair cell and supporting cell subtypes. *Elife* 12:e82978
27. Shiao CK, Lu L, Kieser R, Fukumura K, Pan T, Lin HY, Yang J, Tong EL, Lee GH, Yan Y, Huse JT, Gao R (2023) High throughput single cell long-read sequencing analyses of same-cell genotypes and phenotypes in human tumors. *Nat Commun* 14(1):4124
28. Singh A, Saint-Antoine M (2023) Probing transient memory of cellular states using single-cell lineages. *Front Microbiol* 13:1050516
29. Fan J, Slowikowski K, Zhang F (2020) Single-cell transcriptomics in cancer: computational challenges and opportunities. *Exp Mol Med* 52:1452–1465
30. Zappia L, Phipson B, Oshlack A (2018) Exploring the single-cell RNA-seq analysis landscape with the scRNA-tools database. *PLoS Comput Biol* 14(6):e1006245
31. Heumos L, Schaar AC, Lance C, Litnetskaya A, Drost F, Zappia L, Lücken MD, Strobl DC, Henao J, Curion F, Aliee H, Ansari M, Badia-i-Mompel P, Büttner M, Dann E, Dimitrov D, Dony L, Frishberg A, He D, Hediyezh-zadeh S, Hetzel L, Ibarra IL, Jones MG, Lotfollahi M, Martens LD, Müller CL, Nitzan M, Ostner J, Palla G, Patro R, Piran Z, Ramírez-Suástegui C, Saez-Rodriguez J, Sarkar H, Schubert B, Sikkeema L, Srivastava A, Tanevski J, Virshup I, Weiler P, Schiller HB, Theis FJ (2023) Best practices for single-cell analysis across modalities. *Nat Rev Genet* 24:1
32. Rood JE, Maartens A, Hupalowska A, Teichmann SA, Regev A (2022) Impact of the Human Cell Atlas on medicine. *Nat Med* 28:2486–2496
33. Ye F, Wang J, Li J, Mei Y, Guo G (2024) Mapping Cell Atlases at the Single-Cell Level. *Adv Sci* 11:2305449
34. Chen C, Ge Y, Lu L (2023) Opportunities and challenges in the application of single-cell and spatial transcriptomics in plants. *Front Plant Sci* 14:1185377
35. Chen S, Jiang W, Du Y, Yang M, Pan Y, Li H, Cui M (2023) Single-cell analysis technologies for cancer research: from tumor-specific single cell discovery to cancer therapy. *Front Genet* 14:1276959
36. Huang D, Ma N, Li X, Gou Y, Duan Y, Liu B, Xia J, Zhao X, Wang X, Li Q, Rao J, Zhang X (2023) Advances in single-cell RNA sequencing and its applications in cancer research. *J Hematol Oncol* 16(1):98
37. Shangguan Y, Li C, Lin H, Ou M, Tang D, Dai Y, Yan Q (2020) Application of single-cell RNA sequencing in embryonic development. *Genomics* 112(6):4547–4551
38. Jiang M, Xu X, Guo G (2021) Understanding embryonic development at single-cell resolution. *Cell Regen* 10(1):10
39. Tung CC, Kuo SC, Yang CL, Yu JH, Huang CE, Liou PC, Sun YH, Shuai P, Su JC, Ku C, Lin Y (2023) Single-cell transcriptomics unveils xylem cell development and evolution. *Genome Biol* 24(1):3
40. Kulak K, Wojciechowska N, Samelak-Czajka A, Jackowiak P, Ba-gniewska-Zadworna A (2023) How to explore what is hidden? A review of techniques for vascular tissue expression profile analysis. *Plant Methods* 19:1–18
41. Bawa G, Liu Z, Yu X, Qin A, Sun X (2022) Single-Cell RNA Sequencing for Plant Research: Insights and Possible Benefits. *Int J Mol Sci* 23(9):4497
42. Zheng D, Xu J, Lu Y, Chen H, Chu Q, Fan L (2023) Recent progresses in plant single-cell transcriptomics. *Crop Design* 2(2): 100041
43. Wang Q, Wu Y, Peng A, Cui J, Zhao M, Pan Y, Zhang M, Tian K, Schwab W, Song C (2022) Single-cell transcriptome atlas reveals developmental trajectories and a novel metabolic pathway of catechin esters in tea leaves. *Plant Biotechnol J* 20:2089–2106
44. Sebé-Pedrós A, Saudemont B, Chomsky E, Plessier F, Mailhé MP, Renno J, Loe-Mie Y, Lifshitz A, Mukamel Z, Schmutz S, Novault S, Steinmetz P, Spitz F, Tanay A, Marlow H (2018) Cnidarian Cell Type Diversity and Regulation Revealed by Whole-Organism Single-Cell RNA-Seq. *Cell* 173:1520–1534.e20
45. Siebert S, Farrell JA, Cazet JF, Abeykoon Y, Primack AS, Schnitzler CE, Juliano C (2019) Stem cell differentiation trajectories in Hydra resolved at single-cell resolution. *Science* 365:eaav9314
46. Chari T, Weissbourd B, Gehring J, Ferraioli A, Leclère L, Herl M, Gao F, Chevalier S, Copley RR, Houliston E, Anderson DJ, Pachter L (2021) Whole-animal multiplexed single-cell RNA-seq reveals transcriptional shifts across Clytia medusa cell types. *Sci Adv* 7:1683
47. Hu M, Zheng X, Fan CM, Zheng Y (2020) Lineage dynamics of the endosymbiotic cell type in the soft coral *Xenia*. *Nature* 582:534–538
48. Levy S, Elek A, Grau-Bové X, Menéndez-Bravo S, Iglesias M, Tanay A, Mass T, Sebé-Pedrós A (2021) A stony coral cell atlas illuminates the molecular and cellular basis of coral symbiosis, calcification, and immunity. *Cell* 184:2973–2987.e18
49. Musser JM, Schippers KJ, Nickel M, Mizzon G, Kohn AB, Pape C, Ronchi P, Papadopoulos N, Tarashansky AJ, Hammel JU, Wolf F, Liang C, Hernández-Plaza A, Cantalapedra CP, Achim K, Schieber NL, Pan L, Ruperti F, Francis WR, Vargas S, Kling S, Renkert M, Polikarpov M, Bourenkov G, Feuda R, Gaspar I, Burkhardt P, Wang B, Bork P, Beck M, Schneider TR, Kreshuk A, Wörheide G, Huerta-Cepas J, Schwab Y, Moroz LL, Arendt D (2021) Profiling cellular diversity in sponges informs animal cell type and nervous system evolution. *Science* 374:717–723
50. Fincher CT, Wurtzel O, de Hoog T, Kravarik KM, Reddien P (2018) Cell type transcriptome atlas for the planarian *Schmidtea mediterranea*. *Science* 360:eaq1736
51. Packer JS, Zhu Q, Huynh C, Sivaramakrishnan P, Preston E, Dueck H, Stefanik D, Tan K, Trapnell C, Kim J, Waterston RH, Murray J (2019) A lineage-resolved molecular atlas of *C. Elegans* embryogenesis at single-cell resolution. *Science* 365:eaax1971
52. Li J, Wang J, Zhang P, Wang R, Mei Y, Sun Z, Fei L, Jiang M, Ma L, Chen H, Wang X, Fu Y, Wu H, Liu D, Wang X, Li J, Guo Q, Liao Y, Yu C, Jia D, Wu J, He S, Liu H, Ma J, Lei K, Chen J, Han X, Guo G (2022) Deep learning of cross-species single-cell landscapes identifies conserved regulatory programs underlying cell types. *Nat Genet* 54:1711–1720
53. Li H, Janssens J, de Waegeneer M, Kolluru SS, Davie K, Gardeux V, Saelens W, David F, Brbić M, Spanier K, Leskovec J, McLaughlin CN, Xie Q, Jones R, Brueckner K, Shim J, Tattikota SG, Schnorrer F, Rust K, Nyström TG, Carvalho-Santos Z, Ribeiro C, Pal S, Mahadevaraju S, Przytycka TM, Allen A, Goodwin SF, Berry CW, Fuller MT, White-Coper H, Matunis EL, DiNardo S, Galenza A, O'Brien LE, Dow J, Jasper H, Oliver B, Perrimon N, Deplancke B, Quake S, Luo L, Aerts S (2022) Fly Cell Atlas: A single-nucleus transcriptomic atlas of the adult fruit fly. *Science* 375:eabk2432
54. Cao C, Lemaire LA, Wang W, Yoon PH, Choi YA, Parsons LR, Matese JC, Wang W, Levine M, Chen K (2019) Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature* 571:349–354
55. Liao Y, Ma L, Guo Q, et al (2022) Cell landscape of larval and adult *Xenopus laevis* at single-cell resolution. *Nat Commun* 13:1–15
56. Estermann MA, Williams S, Hirst CE, Roly ZY, Serralbo O, Adhikari D, Powell D, Major AT, Smith C (2020) Insights into Gonadal Sex Dif-

- ferentiation Provided by Single-Cell Transcriptomics in the Chicken Embryo. *Cell Rep* 31:107491
57. Wang F, Ding P, Liang X, Ding X, Brandt CB, Sjöstedt E, Zhu J, Bolund S, Zhang L, de Rooij L, Luo L, Wei Y, Zhao W, Lv Z, Haskó J, Li R, Qin Q, Jia Y, Wu W, Yuan Y, Pu M, Wang H, Wu A, Xie L, Liu P, Chen F, Herold J, Kalucka J, Karlsson M, Zhang X, Helmig RB, Fagerberg L, Lindskog C, Pontén F, Uhlen M, Bolund L, Jessen N, Jiang H, Xu X, Yang H, Carmeliet P, Mulder J, Chen D, Lin L, Luo Y (2022) Endothelial cell heterogeneity and microglia regulons revealed by a pig cell landscape at single-cell level. *Nat Commun* 13:1–18
 58. Wu JJ, Zhu S, Gu F, Valencak TG, Liu JX, Sun H (2022) Cross-tissue single-cell transcriptomic landscape reveals the key cell subtypes and their potential roles in the nutrient absorption and metabolism in dairy cattle. *J Adv Res* 37:1–18
 59. Han X, Wang R, Zhou Y, Fei L, Sun H, Lai S, Saadatpour A, Zhou Z, Chen H, Ye F, Huang D, Xu Y, Huang W, Jiang M, Jiang X, Mao J, Chen Y, Lu C, Xie J, Fang Q, Wang Y, Yue R, Li T, Huang H, Orkin SH, Yuan GC, Chen M, Guo G (2018) Mapping the Mouse Cell Atlas by Micro-well-Seq. *Cell* 172:1091–1107.e17
 60. Han L, Wei X, Liu C, Volpe G, Zhuang Z, Zou X, Wang Z, Pan T, Yuan Y, Zhang X, Fan P, Guo P, Lai Y, Lei Y, Liu X, Yu F, Shangguan S, Lai G, Deng Q, Liu Y, Wu L, Shi Q, Yu H, Huang Y, Cheng M, Xu J, Liu Y, Wang M, Wang C, Zhang Y, Xie D, Yang Y, Yu Y, Zheng H, Wei Y, Huang F, Lei J, Huang W, Zhu Z, Lu H, Wang B, Wei X, Chen F, Yang T, Du W, Chen J, Xu S, An J, Ward C, Wang Z, Pei Z, Wong C, Liu X, Zhang H, Liu M, Qin B, Schambach A, Isern J, Feng L, Liu Y, Guo X, Liu Z, Sun Q, Maxwell PH, Barker N, Muñoz-Cánoves P, Gu Y, Mulder J, Uhlen M, Tan T, Liu S, Yang H, Wang J, Hou Y, Xu X, Esteban MA, Liu L (2022) Cell transcriptomic atlas of the non-human primate *Macaca fascicularis*. *Nature* 604:723–731
 61. Jean-Baptiste K, McFaline-Figueroa JL, Alexandre CM, Dorrity MW, Saunders L, Bubb KL, Trapnell C, Fields S, Queitsch C, Cuperusa J (2019) Dynamics of Gene Expression in Single Root Cells of *Arabidopsis thaliana*. *Plant Cell* 31:993–1011
 62. Wang Y, Huan Q, Li K, Qian W (2021) Single-cell transcriptome atlas of the leaf and root of rice seedlings. *JGG* 48:881–898
 63. Liu H, Hu D, Du P, Wang L, Liang X, Li H, Lu Q, Li S, Liu H, Chen X, Varshney RK, Hong Y (2021) Single-cell RNA-seq describes the transcriptome landscape and identifies critical transcription factors in the leaf blade of the allotetraploid peanut (*Arachis hypogaea* L.). *Plant Biotechnol J* 19:2261–2276
 64. Chen Y, Tong S, Jiang Y, Ai F, Feng Y, Zhang J, Gong J, Qin J, Zhang Y, Zhu Y, Liu J, Ma T (2021) Transcriptional landscape of highly lignified poplar stems at single-cell resolution. *Genome Biol* 22:1–22

Sequencing the transcriptome of single cells: a path to discovering cellular biodiversity

Anna Samelak-Czajka, Małgorzata Marszałek-Zeńczak, Paulina Jackowiak✉

Laboratory of Single Cell Analyses, Institute of Bioorganic Chemistry Polish Academy of Sciences

✉corresponding author: paulinaj@ibch.poznan.pl

Keywords: transcriptomics, next generation sequencing, single cell analysis

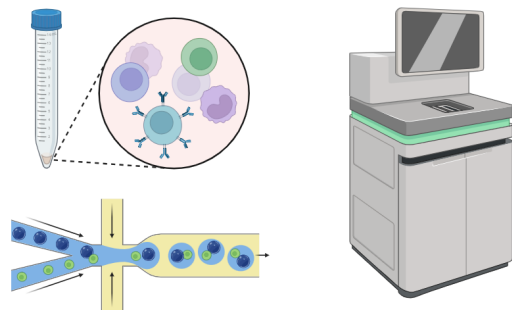
ABSTRACT

Single-cell transcriptomics (scRNA-Seq) is a breakthrough technology that has opened the way to characterizing gene expression with unprecedented resolution. It has enabled the discovery of the cellular diversity of organisms and tracing their developmental processes. A range of technological solutions have been developed to allow analysis of hundreds to even a million cells in a single experiment, as well as an extensive set of tools for bioinformatics analysis of the generated data. The wealth of information provided by scRNA-Seq and the possibility of using this method to study cells, organoids, tissues and even entire organisms determine its wide range of applications. In this paper, we present the experimental and computational parts of the scRNA-Seq procedure, as well as the most important applications of this technology in biomedicine, developmental biology and plant biology.

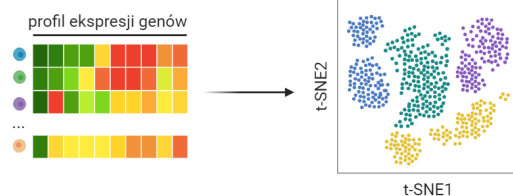
sekwencjonowanie transkryptomu pojedynczych komórek

etapy i wyzwania

procedura eksperymentalna



bioinformatyczna analiza danych



zastosowania

atlas tkanek, organów i całych organizmów

