

NGS dla początkujących: Illumina, Nanopore i PacBio w praktyce

Kompletny przewodnik po technologiach sekwencjonowania nowej generacji - zasady, koszty i zastosowania

Agnieszka Wiśniewska · 08.09.2026

W 2001 roku zsekwencjonowanie jednego ludzkiego genomu kosztowało około 3 miliardów dolarów i zajęło 13 lat pracy międzynarodowego konsorcjum. Dzisiaj, korzystając z platformy Illumina NovaSeq X Plus, można to zrobić w ciągu 24 godzin za mniej niż 200 dolarów. Ta zmiana nie jest tylko imponująca statystyka - to fundament, na którym zbudowano całą nowoczesną genomikę, ekologię molekularną i diagnostykę kliniczną. Jako osoba, która na co dzień pracuje z danymi sekwencjonowania w projektach bioremediacji, widzę, jak NGS stało się narzędziem tak powszechnym jak PCR - a mimo to wciąż budzi wiele pytań wśród osób, które dopiero zaczynają przygodę z biologią molekularną.

Czym jest NGS i dlaczego wyparło sekwencjonowanie Sangera?

Sekwencjonowanie nowej generacji (NGS, ang. Next-Generation Sequencing) to zbiorcza nazwa dla technologii umożliwiających równoległe odczytywanie milionów lub miliardów fragmentów DNA jednocześnie. W odróżnieniu od klasycznego sekwencjonowania Sangera, gdzie analizuje się jeden fragment na raz, NGS przetwarza masowo - stąd alternatywna nazwa: sekwencjonowanie masowe równoległe (ang. massively parallel sequencing).

Sekwencjonowanie Sangera nie zniknęło - wciąż się go stosuje do walidacji pojedynczych mutacji, klonowania czy sekwencjonowania krótkich fragmentów, gdzie wymaga się stuprocentowej dokładności odczytu. Ale gdy potrzebujemy przeskanować cały egzom, zbadać mikrobiom gleby liczący tysiące gatunków bakterii albo wykryć warianty somatyczne w guzie nowotworowym obecne w mniej niż 1% komórek, Sanger po prostu nie nadaje się do tego zadania. NGS zmienia skalę możliwości o kilka rzędów wielkości.

Trzy platformy, które definiują rynek

Rynek NGS jest zdominowany przez trzy technologie, radykalnie różniące się zasadą działania, długością odczytów i zastosowaniami. Warto je znać, zanim zaplanuje się eksperyment lub zinterpretuje dane z zewnętrznego laboratorium.

Illumina - złoty standard krótkich odczytów

Illumina przez ponad dekadę kontrolowała ponad 80% rynku sekwencjonowania na świecie. Jej technologia opiera się na metodzie sekwencjonowania przez syntezę z odwracalnymi terminatorami (ang. sequencing by synthesis, SBS). Praktycznie: fragmenty DNA są amplifikowane na powierzchni flow cella, tworząc klastry identycznych kopii. Następnie polimeraza DNA wbudowuje znakowane fluorescencyjnie nukleotydy jeden po drugim, a kamera rejestruje sygnał po każdym cyklu.

Odczyty Illuminy mają typową długość 150 par zasad (PE150) w trybie paired-end, choć dostępne są też konfiguracje 100 bp i 250 bp. Dokładność na poziomie Q30 (czyli 99,9% poprawnych nukleotydów) jest standardem. Platformy takie jak NovaSeq X Plus generują do

16 terapar zasad na jeden przebieg - wystarczająco, aby zsekwencjonować ponad 6000 ludzkich genomów z 30-krotnym pokryciem.

Słabością są krótkie odczyty, które utrudniają montaż genomów z dużą ilością powtórzeń oraz wykrywanie strukturalnych wariantów chromosomowych. Dlatego Illumina świetnie sprawdza się w resekwencjonowaniu (resequencing), analizie RNA-seq, ChIP-seq, ATAC-seq i amplikon-sekwencjonowaniu mikrobiomów.

Oxford Nanopore - sekwencjonowanie w czasie rzeczywistym

Oxford Nanopore Technologies (ONT) wprowadzili coś, co jeszcze 10 lat temu brzmiało jak science fiction: urządzenie wielkości pendrive'a zdolne do sekwencjonowania DNA w terenie. MinION podłącza się do laptopa przez USB i może pracować w szpitalu polowym, na łodzi badawczej czy - co widziałam na własne oczy - na skraju torfowiska podczas monitoringu mikrobiologicznego.

Zasada działania jest całkowicie inna niż w Illumina. Cząsteczka DNA przepływa przez nanopory białkowe osadzone w syntetycznej membranie. Każda kombinacja nukleotydów przesuwających się przez por zaburza przepływ prądu elektrycznego w charakterystyczny sposób. Algorytm (tzw. basecaller, np. Dorado) przekształca zapis zmian prądu w sekwencje nukleotydów.

Kluczową zaletą: długość odczytów jest praktycznie nieograniczona. Typowe odczyty mają 10-50 kb, ale zdarzają się odczyty przekraczające 2 Mb (megapary zasad) - to cały chromosom odczytany jako jeden ciągły fragment. Takie dane są bezcenne przy montażu genomów, wykrywaniu strukturalnych wariantów i analizie modyfikacji epigenetycznych (metylacji DNA) bezpośrednio z surowego sygnału.

Historyczną słabością Nanopore była wysoka stopa błędów - pierwsze wersje chemii sekwencjonowania osiągały dokładność 85-90%. Najnowsza chemia R10.4.1 w trybie duplex osiąga dokładność powyżej 99,5%, zbliżając się do standardów Illuminy przy zachowaniu zalet długich odczytów.

PacBio - długie odczyty z wysoką dokładnością

Pacific Biosciences (PacBio) oferuje technologię SMRT (ang. Single Molecule Real-Time sequencing). DNA polimeraza sekwencjonuje pojedynczą cząsteczkę DNA w czasie rzeczywistym w specjalnych studniach ZMW (ang. zero-mode waveguides). Fluorescencyjne sygnały emitowane przy wbudowywaniu każdego nukleotydu są rejestrowane z rozdzielczością nanosekundową.

Najnowszy tryb pracy, HiFi (CCS - Circular Consensus Sequencing), polega na wielokrotnym przejściu polimerazy po cząsteczce kolistej DNA. Efektem są odczyty o długości 15-25 kb z dokładnością powyżej 99,9% (Q30+). To unikalne połączenie długości i precyzji, które czyni PacBio niezastąpionym w de novo montażu genomów referencyjnych, analizie izoform RNA (Iso-Seq) i wykrywaniu tandemowych powtórzeń.

Wadą jest koszt - sekwencjonowanie PacBio jest droższe niż Illumina przy porównywalnej liczbie par zasad. Platforma Revio z 2023 roku znacznie obniżyła koszty (do około 10-15 dolarów za gigapary zasad), ale wciąż nie dorównuje ekonomii Illuminy w zastosowaniach wymagających dużej głębokości pokrycia.

Przygotowanie biblioteki DNA - co dzieje się przed sekwencjonowaniem?

Zanim próbka trafi do sekwenatora, musi przejść przez wieloetapowy proces przygotowania biblioteki. To właśnie tutaj popełnia się najwięcej błędów, które bezpośrednio przekładają się na złą jakość danych - i żaden algorytm bioinformatyczny nie potrafi tego naprawić.

Ekstrakcja i ocena jakości DNA

Punktem wyjścia jest izolacja DNA o odpowiedniej czystości i integralności. Dla sekwencjonowania Illumina wystarczy DNA o średniej długości fragmentów powyżej 20 kb i stosunku absorbancji A260/A280 w zakresie 1,8-2,0. Dla PacBio i Nanopore (długie odczyty) potrzebujemy DNA o długości fragmentów powyżej 50-100 kb, co wymaga delikatniejszych protokołów - pipetowanie z szeroką końcówką, unikanie vortexowania.

Do oceny integralności DNA stosuje się elektroforezę na żelu agarozowym, system Agilent Femto Pulse lub Genomic DNA ScreenTape (TapeStation). Steżenie mierzy się fluorometrycznie (Qubit), a nie spektrofotometrycznie (NanoDrop) - NanoDrop jest mniej czuły na zanieczyszczenia RNA i nie odróżnia jednocieczowego DNA od dwucieczowego.

Fragmentacja i naprawianie końców

W przypadku Illuminy DNA fragmentuje się do pożądanego długości - najczęściej 300-500 bp dla PE150. Fragmentację przeprowadza się enzymatycznie (np. zestawy Kapa HyperPlus) lub mechanicznie (sonikacja, np. Covaris M220). Po fragmentacji końce cząsteczek są naprawiane (end-repair), adenylowane (A-tailing) i ligowane z adapterami.

Adaptery to krótkie dwucieczowe oligonukleotydy zawierające sekwencje komplementarne do powierzchni flow cella oraz sekwencje indeksujące (barcodes), które pozwalają na multipleksowanie - łączenie wielu próbek w jednym przebiegu sekwenatora.

Multipleksowanie jest kluczem do ekonomiczności NGS: na jednym flow cella NovaSeq można jednocześnie sekwencjonować setki próbek.

Amplifikacja i kontrola jakości biblioteki

Gotową bibliotekę amplifikuje się PCR (najczęściej 4-8 cykli), a następnie ocenia jej jakość - steżenie (Qubit), rozkład wielkości fragmentów (Bioanalyzer lub TapeStation) i brak dimerów adapterów. Dimery adapterów to zmora każdego laboratorium: krótkie artefakty, które konkurują z prawdziwymi fragmentami podczas klastrowania i marnują przepustowość sekwenatora.