

Podstawy sekwencjonowania wysokoprzepustowego

ALEKSANDRA GIZA¹, EWELINA IWAN¹, ARKADIUSZ BOMBA¹, DARIUSZ WASYL^{1,2}

¹Zakład Analiz Omicznych, ²Zakład Mikrobiologii,
Państwowy Instytut Weterynaryjny – Państwowy Instytut Badawczy w Puławach, Al. Partyzantów 57, 24-100 Puławy

Otrzymano 30.06.2021

Zaakceptowano 10.09.2021

Giza A., Iwan E., Bomba A., Wasyl D.
Basics of high throughput sequencing

Summary

Sequencing can provide genomic characterisation of a specific organism, as well as of a whole environmental or clinical sample. High Throughput Sequencing (HTS) makes it possible to generate an enormous amount of genomic data at gradually decreasing costs and almost in real-time. HTS is used, among others, in medicine, veterinary medicine, microbiology, virology and epidemiology. The paper presents practical aspects of the HTS technology. It describes generations of sequencing, which vary in throughput, read length, accuracy and costs—and thus are used for different applications. The stages of HTS, as well as their purposes and pitfalls, are presented: extraction of the genetic material, library preparation, sequencing and data processing. For success of the whole process, all stages need to follow strict quality control measurements. Choosing the right sequencing platform, proper sample and library preparation procedures, as well as adequate bioinformatic tools are crucial for high quality results.

Keywords: high throughput sequencing, DNA extraction, library preparation, bioinformatic analysis

Mianem sekwencjonowania określa się wygenerowanie cyfrowego zapisu informacji genetycznej znajdującej się w DNA. Sekwencjonowanie stwarza możliwość scharakteryzowania pojedynczego lub wielu organizmów obecnych w próbce środowiskowej, lub w próbkach materiału klinicznego. Aktualnie dostępne technologie następnej generacji określane są wspólnie mianem sekwencjonowania wysokoprzepustowego (High throughput sequencing; HTS). Umożliwiają generowanie olbrzymiej ilości danych genomowych coraz niższym kosztem. HTS jest narzędziem o dużym potencjale aplikacyjnym. Znajduje zastosowanie w obszarach zdrowia publicznego, zdrowia zwierząt czy bezpieczeństwa żywności, szczególnie w dziedzinach: medycyny, weterynarii, mikrobiologii, wirusologii czy epidemiologii.

Praca ma na celu przedstawienie podstawowych i praktycznych informacji dotyczących sekwencjonowania wysokoprzepustowego. Technologia ta jest metodą uniwersalną, jednakże dla uzyskania wartościowych wyników wymaga stosownego sekwenatora, właściwego przygotowania materiału oraz doboru odpowiednich narzędzi bioinformatycznych.

Technologie sekwencjonowania

Sekwencjonowanie można przeprowadzić z wykorzystaniem różnych technologii. Tradycyjne sekwencjonowanie Sangera zaliczane jest do I generacji.

Generacje II i III są wspólnie określane mianem sekwencjonowania wysokoprzepustowego (High Throughput Sequencing; HTS). Wykorzystywane do analiz urządzenia – tzw. sekwenatory, w zależności od systemu mogą różnić się m.in. rodzajem zastosowanej technologii odczytu, długością odczytywanych fragmentów nici DNA czy przepustowością (tab. 1), jednak podstawową różnicą między tradycyjnym sekwencjonowaniem Sangera a HTS jest ilość generowanych danych w pojedynczym cyklu pracy urządzenia (run) (tab. 1). Dla przykładu: metoda Sangera umożliwia jednoczesne sekwencjonowanie np. kilkudziesięciu fragmentów genów, co dostarcza bardzo ograniczoną ilość informacji na temat badanego organizmu, przy czym wykorzystując HTS można równocześnie sekwencjonować w wielu powtórzeniach całe genomy o różnym stopniu złożoności. Ze względu na wielkoformatowy charakter dane uzyskiwane w tym procesie są przetwarzane przy użyciu skryptów bioinformatycznych i serwerów o znacznych mocach obliczeniowych (1, 4).

Sekwencjonowanie II generacji. Zasada działania systemów II generacji opiera się na masowym, wielokrotnym sekwencjonowaniu krótkich fragmentów DNA (tab. 1). Takie odczyty w toku dalszej analizy bioinformatycznej zostają połączone w dłuższe elementy, tzw. kontigi. Cały proces wiąże się z generowaniem bardzo dużej ilości danych i wymaga zastosowania specjalistycznego oprogramowania opartego na in-

Tab. 1. Porównanie systemów do sekwencjonowania

Systemy	I generacja (Sanger)	Sekwencjonowanie wysokoprzepustowe			
	ABI	II generacja	III generacja		
Wielkość sekwencjonowanych fragmentów	500-900 pz (max. 1500)	300-600 pz	200-400 pz	15 Kpz ÷ < 4,2 Mpz	< 90 Kpz
Maksymalna przepustowość/**	80 Kpz	6 Tpz	25 Gpz	14 Tpz	160 Gpz

Objaśnienia: pz – pary zasad azotowych; Kpz – 10^3 pz; Mpz – 10^6 pz; Gpz – 10^9 pz; Tpz – 10^{12} pz; * – ilość danych, które można uzyskać w pojedynczym cyklu pracy urządzenia; ** – dotyczy urządzeń o najwyższej przepustowości z danej generacji.

<https://emea.illumina.com/>; <https://www.pacb.com/>; <https://www.thermofisher.com/>; <https://nanoporetech.com/>

terfejsie „command-line” i sprzęcie o wysokiej mocy komputacyjnej. Do urządzeń II-generacji zalicza się m.in. sekwenatory Illumina oraz IonTorrent (1, 4).

Głównymi zaletami systemów II generacji są: stosunkowo niski koszt i krótki czas badania. Ponadto wielokrotne sekwencjonowanie tego samego fragmentu minimalizuje ryzyko potencjalnego błędu, zwiększając wiarygodność uzyskanych wyników. Krotność, z którą dana zasada azotowa została odczytana, określa się pojęciem pokrycia (coverage). W kontekście analizy bioinformatycznej stanowi ono jeden z elementów kontroli jakości uzyskanych danych, bowiem jakość otrzymanych wyników jest powiązana z uzyskaniem odpowiednio wysokiego pokrycia, a samo wyznaczenie minimalnego pokrycia zależy od rodzaju planowanych analiz (12).

Sekwencjonowanie III generacji. Charakterystyczną cechą sekwencjonowania III generacji jest generowanie długich fragmentów (tzw. długie odczyty; long reads). Tego rodzaju systemy znajdują zastosowanie w analizie obszarów bardzo problematycznych dla sekwenatorów II generacji, tj. rozległe delecje lub insercje, sekwencje powtórzone lub bogate w pary GC. Są przydatne do generowania ciągłych struktur genomicznych, np. całych chromosomów czy plazmidów, tworzenia wysokiej jakości sekwencji referencyjnych czy uzyskiwania informacji o przestrzennym położeniu genów. Znajdują także zastosowanie w sekwencjonowaniu *de novo* genomów nowych, nieznanymi gatunków mikroorganizmów (4, 13).

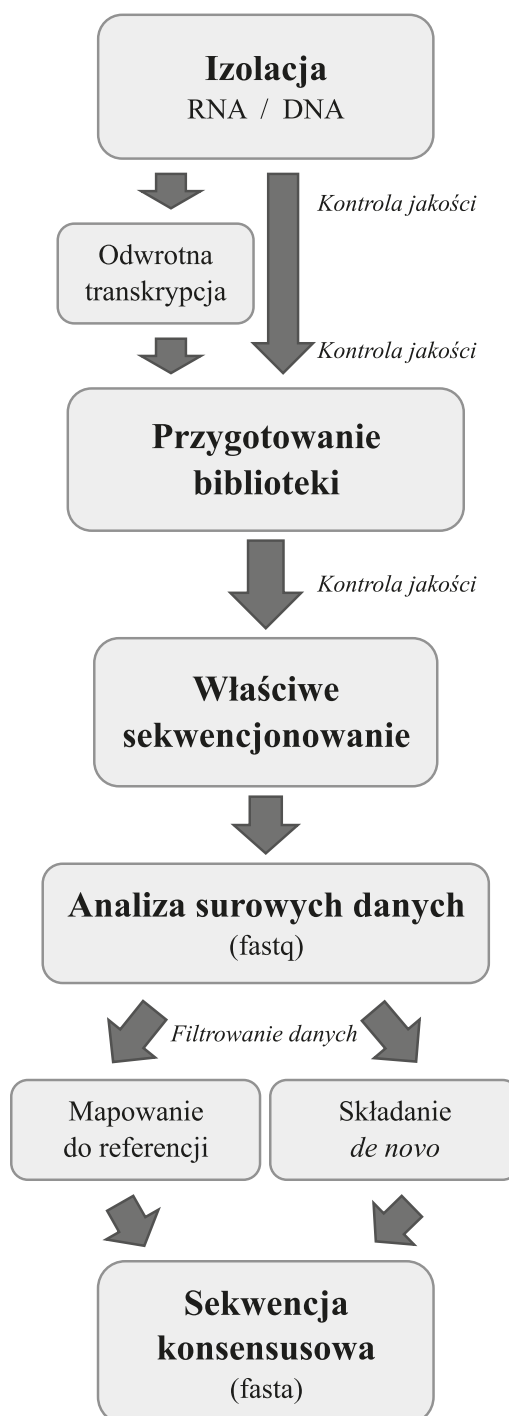
Wśród sekwenatorów III generacji wyróżnić można dwa główne systemy: PacBio oraz Oxford Nanopore. Urządzenia te mają możliwość odczytu bardzo długich fragmentów, zależnie od systemu może to być 30-90 Kpz dla PacBio (Sequel) czy nawet 4,2 Mpz w przypadku Oxford Nanopore.

Sekwencjonowanie III generacji stanowi przyszłość genetyki molekularnej, jednakże ze względu na wysoką cenę bądź wciąż niewystarczającą wiarygodność wyników, to urządzenia II generacji są aktualnie znacznie częściej stosowane.

Przebieg procesu sekwencjonowania HTS

Każde sekwencjonowanie HTS, niezależnie od generacji i systemu, składa się z kilku etapów: izolacji materiału genetycznego, przygotowania biblioteki, właściwego sekwencjonowania oraz bioinforma-

tycznej analizy danych (6). Poszczególne etapy HTS przedstawia rycina 1.



Ryc. 1. Etapy sekwencjonowania wysokoprzepustowego

Przygotowanie próbki do sekwencjonowania.

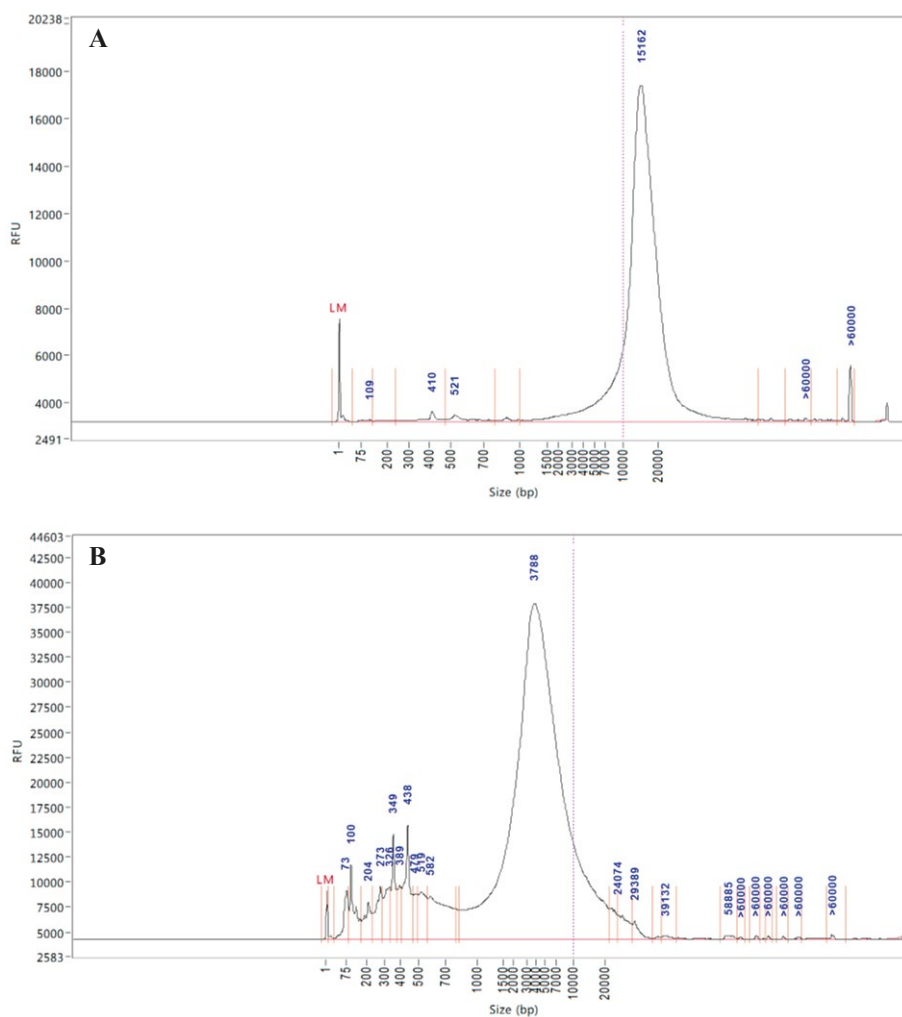
Kluczowym warunkiem prawidłowego przebiegu sekwencjonowania i uzyskania oczekiwanych wyników jest wykorzystanie materiału genetycznego spełniającego określone parametry jakościowe – czystość, stężenie oraz wysoki poziom integralności. Mierzona spektrofotometrycznie czystość DNA pozwala ocenić zanieczyszczenie próbki przez białka (λ 260/280) oraz odczynniki użyte do ekstrakcji (λ 260/230). Stężenie DNA mierzone fluorymetrycznie pozwala z wysoką precyzją określić ilość natywnego (dwuniciowego) DNA. Wartość stężenia mierzona przy użyciu spektrofotometru jest z reguły zawyżona ze względu na fakt, iż pomiar absorpcji przy długości fali λ 260 nm jest specyficzny dla kwasów nukleinowych i odnosi się do obecności w próbce, poza dwuniciowym DNA, również zdegradowanego DNA oraz RNA. Ocena integralności DNA odbywa się przy pomocy elektroforezy kapilarnej, dostarczając informacji na temat stopnia pofragmentowania materiału genetycznego. Nadmierna fragmentacja DNA może być skutkiem działania czynników fizykochemicznych, takich jak: wielokrotne rozmrażanie i zamrażanie próbki, nieodpowiednie warunki przechowywania, degradacja na skutek pozostałości odczynników do ekstrakcji czy sama metodyka ekstrakcji. Odpowiednia integralność DNA jest szczególnie ważna w przypadku sekwencjonowania III generacji ze względu na potrzebę uzyskiwania długich odczytów (5). Różnice pomiędzy integralnym i zdegradowanym DNA przedstawiono w rycinie 2.

Uzyskanie DNA dobrej jakości z materiału, który jest jednorodny genetycznie (np. czyste kultury bakteryjne) z reguły nie stanowi problemu. Dużo większym wyzwaniem są próbki ze środowiska (np. woda, kał, ścieki, żywność), które pod względem genetycznym odzwierciedlają skład genetyczny metaprobki. Uzyskanie wiarygodnych wyników wymaga wówczas głębokiego sekwencjonowania, co zwiększa koszt badań i wymagania komputacyjne analizy. Wyzwaniem jest również izolacja kwasów nukleinowych z bakterii wewnątrzkomórkowych oraz wirusów pochodzących bezpośrednio z materiału diagnostycznego, ale także namnażanych w hodowlach komórkowych lub zarodkach. Wirusy są zróżnicowane pod względem budowy, cyklu replikacyjnego czy powinowactwa do komórek gospodarza, dlatego każdy przypadek wyma-

ga podejścia indywidualnego. W znacznej większości urządzeń DNA jest jedynym rodzajem cząsteczek, które mogą być sekwencjonowane, dlatego w przypadku konieczności poznania informacji zapisanej w RNA, konieczne jest wykonanie odwrotnej transkrypcji w celu uzyskania cDNA (6).

Materiał genetyczny wirusów i bakterii wewnątrzkomórkowych po izolacji zawiera w większości DNA gospodarza. W wyniku sekwencjonowania takiej próbki uzyskiwane są głównie odczyty inne niż docelowe, dlatego większość sekwencji będzie pochodzić od gospodarza i zostanie odrzucona na etapie bioinformatycznej obróbki danych. W efekcie sekwencji docelowych może być zbyt mało (tzw. zbyt niskie pokrycie) do przeprowadzenia zamierzonych analiz. Problem ten można ograniczyć przez zastosowanie różnych technik wzbogacania materiału docelowego.

W zależności od rodzaju patogenu stosuje się różne podejścia, np. ultrawirowanie, trawienie nukleazami (DNA-za, RNA-za). Można także usunąć genom gospodarza z próbki przy pomocy komercyjnie dostępnych zestawów bazujących np. na różnicach w metylacji pomiędzy prokariotami i eukariotami (16). Inne



Ryc. 2. Bakteryjny DNA poddany rozdzielowi przy użyciu elektroforezy kapilarnej na aparacie Fragment Analyzer (Agilent Technologies), zestaw: HS Genomic DNA Kit; A – DNA o wysokim stopniu integralności – *Enterobacter cloacae*, B – DNA zdegradowany – *Staphylococcus aureus*

stosowane podejście, to selekcja pozytywna, polegająca na wzbogaceniu przez amplifikację fragmentów genomu wirusa/bakterii (17) lub zaprojektowaniu sond hybrydujących do specyficznych regionów DNA (hybrid capture), np. materiału genetycznego organizmu docelowego (9). Metoda oczyszczania powinna być dobrana zależnie od rodzaju i ilości patogenu w próbce oraz rodzaju materiału, z którego pochodzi, np. określonej tkanki czy organizmu.

Wzbogacenie przez sondy bądź amplifikację można wykorzystać do analizy próbek metagenomicznych, np. taksonomicznej identyfikacji bakterii (15), grzybów (14) lub roślin (3) obecnych w badanym materiale. Przy takim rodzaju próbek świetnie sprawdzają się również panele genowe, które umożliwiają analizę jedynie wybranych regionów. Panele takie pozwalają np. na identyfikację genów oporności na antybiotyki, sekwencjonowanie genomów wirusowych (np. rozróżnienie SARS-Cov2 od grypy i innych wirusów oddechowych) lub identyfikację mutacji w genach odpowiedzialnych za rozwój różnych chorób o podłożu genetycznym, co znajduje zastosowanie w ich diagnostyce i celowanym leczeniu (2).

Przygotowanie biblioteki DNA. Biblioteka to zbiór odpowiedniej długości, unikalnie oznakowanych i reprezentatywnych dla badanej próbki fragmentów DNA, które poddawane są sekwencjonowaniu HTS. W trakcie przygotowania bibliotek przeznaczonych do sekwencjonowania materiał genetyczny jest amplifikowany lub fragmentowany na cząstki DNA o określonej wielkości. W przypadku preferującego długie odczyty sekwencjonowania III generacji ten etap jest opcjonalny. Cząsteczka DNA może zostać „pocięta” na fragmenty w sposób mechaniczny (sonikacja, nebulizacja), przy użyciu enzymów restrykcyjnych lub w procesie tzw. tagmentacji (tagmentation), czyli losowej fragmentacji z równoczesnym przyłączeniem sekwencji adaptorowych (6, 10).

Do pociętych końców DNA są doczepiane tzw. adaptory – krótkie, oligonukleotydowe znaczniki o znanej sekwencji, składające się z kilku- kilkunastu par zasad. Jednym z elementów adaptora są tzw. indeksy, które służą do przypisania danego fragmentu DNA do określonej próbki w trakcie analizy bioinformatycznej. Metoda HTS umożliwia bowiem równoczesne sekwencjonowanie nawet kilkuset próbek w „jednej próbce” w trakcie pojedynczej rundy sekwencjonowania HTS – tzw. run. Adaptory umożliwiają również inkorporację bibliotek HTS do powierzchni stałej, na której zachodzi sam proces sekwencjonowania (6).

W zależności od protokołu oraz sekwenatora wykonuje się także m.in. naprawę pociętych końców DNA, selekcję długości uzyskanych fragmentów oraz amplifikację bibliotek. Istotna jest również kontrola jakości przygotowanych bibliotek (np. przy użyciu elektroforezy kapilarnej lub mikrokapilarnej), w celu analizy rozkładu wielkości uzyskanych fragmentów DNA. Kolejny etap – normalizacja – ma na celu do-

prowadzenie wszystkich próbek do odpowiedniego stężenia. W końcowym etapie przygotowane biblioteki są łączone w jednej próbce zawierającej materiał do sekwencjonowania w jednej rundzie, w której każda z próbek powinna być właściwie reprezentowana (6).

Procedury przygotowania bibliotek są bardzo różne w zależności od zastosowanej aplikacji, sekwenatora, ilości materiału genetycznego oraz celu sekwencjonowania.

Proces sekwencjonowania. W sekwenatorach II generacji właściwe sekwencjonowanie zachodzi na powierzchni stałej (płytki flow cell – Illumina; kulki beads immobilizowane w dołkach na płytce – Ion Torrent), w obrębie obszarów, na których każdy fragment przygotowanej biblioteki jest klonalnie reprezentowany. Sekwencjonowanie odbywa się z reguły przez syntezę (SBS – sequencing by synthesis – Illumina, Ion Torrent) i oparte jest o reakcję amplifikacji, podczas której przyłączane są komplementarne sekwencje nukleotydów. Metody detekcji mogą być różne, np.: nukleotydy znakowane fluorochromami, które są rejestrowane przy pomocy zaawansowanej aparatury optycznej sekwenatora (Illumina), lub rejestracja zmian pH w trakcie przyłączania zasad i detekcja protonów na układzie scalonym (Ion Torrent) (4).

W odróżnieniu od II generacji, sekwenatory III generacji odczytują sygnał bezpośrednio z nici DNA. Podczas sekwencjonowania na urządzeniach PacBio pojedyncze cząsteczki DNA ulegają inkorporacji w mikrostudzienkach, w których odbywa się proces syntezy nici. Sygnał fluorescencyjny jest następnie odczytywany w czasie rzeczywistym przez układ optyczny. Z kolei sekwenatory Oxford Nanopore wykorzystują technologię nanoporów, w której urządzenie rejestruje sygnał elektryczny generowany w trakcie przedostawania się nici DNA przez pory (4).

Sekwencjonowanie HTS może być przeprowadzone w trybie single-end lub pair-end. Sekwencjonowanie single-end polega na odczytywaniu nici DNA tylko z jednej strony i umożliwia uzyskiwanie dużej ilości danych w sposób szybki i ekonomiczny. Z kolei sekwencjonowanie z dwóch stron (pair-end) polega na sekwencjonowaniu każdego fragmentu DNA od obu końców. Ułatwia to składanie sekwencji oraz umożliwia lepsze zobrazowanie rearanżacji genomowych, wykrycie sekwencji powtarzających się, fuzji genowych oraz nowych transkryptów (11).

Po sekwencjonowaniu przy użyciu każdego z systemów uzyskuje się dane w postaci sekwencji nukleotydów, które wymagają analizy bioinformatycznej.

Bioinformatyczna analiza danych HTS. Ostatni element sekwencjonowania HTS stanowi bioinformatyczna analiza danych. Generowanie cyfrowych zapisów odczytywanych sekwencji określane jest mianem „basecalling” i odbywa się bezpośrednio w sekwenatorze. Wyjściowym formatem plików uzyskiwanych w HTS są tzw. pliki surowe – FASTQ. Jest to jeden ze standardowych formatów przechowywania

danych genomowych. Pliki FASTQ zawierają odczytane sekwencje nukleotydowe (reads) wraz z oceną prawdopodobieństwa błędnego odczytu każdej z zasad.

W dalszym toku wytwarzania sekwencji konsensusowej dokonywana jest kontrola jakości otrzymanych odczytów. W jej wyniku fragmenty sekwencji o niskiej jakości i adaptory są usuwane (trimming). Sekwencje są następnie składane w większe elementy tzw. kontigi – efektem czego jest powstanie plików FASTA. Samo składanie odbywać się może na dwa sposoby: przez mapowanie do sekwencji referencyjnej (mapping), o ile taka jest znana, lub składanie odczytów *de novo* przez ich porównanie, dopasowanie i orientowanie względem samych siebie (7).

Zarówno pliki FASTQ, jak i FASTA mogą być dowolnie przetwarzane przez użytkownika, a schemat takiego procesu analitycznego i interpretacja merytoryczna zależą od celu badania (np. analiza filogenetyczna i określenie stopnia pokrewieństwa, określanie profilu wirulencji czy oporności) i potrzeb użytkownika.

Podsumowanie

Sekwencjonowanie wysokoprzepustowe to uniwersalna metoda analityczna o szerokim zastosowaniu w bezpieczeństwie żywności oraz medycynie i medycynie weterynaryjnej. Jest procesem wieloetapowym. Ekstrakcja materiału genetycznego stanowi element bezpośrednio rzutujący na wynik badania, gdyż od niej zależy jakość uzyskanych danych. Próbkę, w której DNA/RNA docelowo stanowi niewielki procent całkowitego materiału, często wymagają wzbogacenia poprzez kombinację metod oczyszczania. Zdegradowany i zanieczyszczony (np. białkami lub odczynnikami użytymi do ekstrakcji) materiał genetyczny utrudnia, a czasem nawet uniemożliwia uzyskanie biblioteki o dobrej jakości. Proces przygotowywania bibliotek powinien być zoptymalizowany pod kątem różnorodności badanych próbek. Dostępne obecnie systemy HTS różnią się przepustowością, wiarygodnością, długością odczytów oraz ceną (zarówno urządzeń, jak i odczynników). Sekwenatory III generacji umożliwiają uzyskiwanie długich odczytów – przydatnych do analizy rozległych rearanżacji w genomie lub składania ciągłych cyrkularnych genomów. Sekwenatory II generacji dostarczają krótszych, lecz bardziej wiarygodnych odczytów. Wybór odpowiedniej metodyki sekwencjonowania powinien być uzależniony m.in. od rodzaju próbek i celu badań, możliwości sprzętowych i finansowych oraz posiadania odpowiedniej infrastruktury laboratoryjnej.

Piśmiennictwo

- Behjati S., Tarpey P. S.: What is next generation sequencing? Arch. Dis. Child. Educ. Pract. Ed. 2013, 98, 236-238.
- Belilovski Rozenblum A., Ilouze M., Dudnik E., Dvir A., Soussan-Gutman L., Geva S., Peled N.: Clinical Impact of Hybrid Capture-Based Next-Generation Sequencing on Changes in Treatment Decisions in Lung Cancer. J. Thorac. Oncol. 2017, 12, 258-268.

- Cheng T., Xu C., Lei L., Li C., Zhang Y., Zhou S.: Barcoding the kingdom Plantae: new PCR primers for ITS regions of plants with improved universality and specificity. Mol. Ecol. Resour. 2016, 16, 138-149.
- Goodwin S., McPherson J. D., McCombie W. R.: Coming of age: ten years of next-generation sequencing technologies. Nat. Rev. Genet. 2016, 17, 333-351.
- He H. J., Stein E. V., DeRose P., Cole K. D.: Limitations of methods for measuring the concentration of human genomic DNA and oligonucleotide samples. Biotechniques 2018, 64, 59-68.
- Head S. R., Komori H. K., LaMere S. A., Whisenant T., Van Nieuwerburgh F., Salomon D. R., Ordoukhanian P.: Library construction for next-generation sequencing: Overviews and challenges. Biotechniques 2018, 56, 61-77.
- Kämpf C., Specht M., Scholz A., Puppel S. H., Doose G., Reiche K., Schor J., Hackermüller J.: uap: reproducible and robust HTS data analysis. BMC Bioinformatics 2019, 20, 664.
- Kchouk M., Jean-François Gibrat J. F., Elloumi M.: Generations of Sequencing Technologies: From First to Next Generation. Biol. Med. 2017, 9, 3.
- Kim K. W., Deveson I. W., Pang C. N. I., Yeang M., Naing Z., Adikari T., Hammond J. M., Stevanovski I., Beukers A. G., Verich A., Yin S., McFarlane D., Wilkins M. R., Stelzer-Braid S., Bull R. A., Craig M. E., van Hal S. J., Rawlinson W. D.: Respiratory viral co-infections among SARS-CoV-2 cases confirmed by virome capture sequencing. Sci. Rep. 2021, 11, 3934.
- Knierim E., Lucke B., Schwarz J. M., Schuelke M., Dominik Seelow D.: Systematic Comparison of Three Methods for Fragmentation of Long-Range PCR Products for Next Generation Sequencing. PLoS One 2011, 6, 28240.
- Liu T., Chen C. Y., Chen-Deng A., Chen Y. L., Wang J. Y., Hou Y. I., Lin M. C.: Joining Illumina paired-end reads for classifying phylogenetic marker sequences. BMC Bioinform. 2020, 21, 105.
- Pedersen B. S., Quinlan A. R.: Mosdepth: quick coverage calculation for genomes and exomes. Bioinformatics 2018, 34, 867-868.
- Rhoads A., Au K. F.: PacBio Sequencing and Its Applications. Genomics Proteomics Bioinformatics 2015, 13, 278-289.
- Stielow J. B., Lévesque C. A., Seifert K. A., Meyer W., Iriny L., Smits D., Renfurm R., Verkley G. J. M., Groenewald M., Chaduli D., Lomascolo A., Welti S., Lesage-Meessen L., Favel A., Al-Hatmi A. M. S., Damm U., Yilmaz N., Houbraken J., Lombard L., Quedvlieg W., Binder M., Vaas L. A. I., Vu D., Yurkov A., Begerow D., Roehl O., Guerreiro M., Fonseca A., Samerpitak K., van Diepeningen A. D., Dolatabadi S., Moreno L. F., Casaregola S., Mallet S., Jacques N., Roscini L., Egidi E., Bizet C., Garcia-Hermoso D., Martin M. P., Deng S., Groenewald J. Z., Boekhout T., de Beer Z. W., Barnes I., Duong T. A., Wingfield M. J., de Hoog G. S., Crous P. W., Lewis C. T., Hambleton S., Moussa T. A. A., Al-Zahrani H. S., Almaghrabi O. A., Louis-Seize G., Assabgui R., McCormick W., Omer G., Dukik K., Cardinali G., Eberhardt U., de Vries M., Robert V.: One fungus, which genes? Development and assessment of universal primers for potential secondary fungal DNA barcodes. Persoonia 2015, 35, 242-63.
- Sturgeon A., Stull J. W., Costa J. M. C., Weese S.: Metagenomic analysis of the canine oral cavity as revealed by high-throughput pyrosequencing of the 16S rRNA gene. Vet. Microbiol. 2013, 162, 891-898.
- Thoendel M., Jeraldo P. R., Greenwood-Quaintance K. E., Yao J. Z., Chia N., Hanssen A. D., Abdel M. P., Patel R.: Comparison of microbial DNA enrichment tools for metagenomic whole genome sequencing. J. Microbiol. Methods 2016, 127, 141-145.
- Watson S. J., Welkers M. R. A., Depledge D. P., Coulter E., Breuer J. M., de Jong M. D., Kellam P.: Viral population analysis and minority-variant detection using short read next-generation sequencing. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 2013, 368.

Autor korespondencyjny: mgr Aleksandra Giza, Zakład Analiz Omicznych, Państwowy Instytut Weterynaryjny – Państwowy Instytut Badawczy, Al. Partyzantów 57, 24-100 Puławy; e-mail: aleksandra.giza@piwet.pulawy.pl