

APARATURA

BADAWCZA I DYDAKTYCZNA

Metody sekwencjonowania nowej generacji oraz ich wykorzystanie w genetyce, hodowli i biotechnologii roślin

MARTA ORŁOWSKA, MARTYNA SOBCZYK

KATEDRA GENETYKI, HODOWLI I BIOTECHNOLOGII ROŚLIN, ZACHODNIOPOMORSKI UNIWERSYTET
TECHNOLOGICZNY W SZCZECINIE

Słowa kluczowe: NGS, pirosekwencjonowanie, sekwencjonowanie nowej generacji, sekwencjonowanie przez ligację, sekwencjonowanie przez syntezę

STRESZCZENIE:

W niniejszej pracy przedstawiono i scharakteryzowano metody sekwencjonowania nowej generacji – NGS (ang. *Next-Generation Sequencing*). W tym krótkim przeglądzie zaprezentowano najważniejsze techniki sekwencjonowania drugiej generacji. Metody te umożliwiają odczytanie kolejności nukleotydów w łańcuchu DNA. Szybki rozwój technik sekwencjonowania zaczął się na początku lat 70. XX wieku, kiedy to opublikowano dwie techniki pierwszej generacji – metodę Sangera, na której bazują techniki NGS, oraz metodę Maxama i Gilberta. Późniejsze ich modyfikacje (m.in. wprowadzenie elektroforezy kapilarnej, fluorescencyjne znakowanie nukleotydów, zastosowanie mikroprocesora) doprowadziły do pełnej automatyzacji procesu i przyczyniły się do jego udoskonalenia. Obecnie metody NGS są narzędziem wykorzystywanym w wielu dziedzinach nauki, m.in. w genetyce, biotechnologii, biologii molekularnej i w nowoczesnej hodowli roślin. Technologie prezentowane w artykule cechują się wieloma zaletami, ale nie są też pozbawione wad. Dokładność odczytu sekwencji oraz ogólne koszty eksploatacyjne wskazują na to, że najbardziej obiecująca metoda sekwencjonowania nowej generacji opiera się na systemie SOLiD. Z kolei zaś sekwencjonowanie poprzez ligację (platforma SOLiD) oprócz dużej dokładności charakteryzuje się dłuższym czasem analiz w porównaniu do innych omawianych metod.

Application of Next Generation Sequencing methods in genetics, plant breeding and biotechnology

Keywords: NGS, pyrosequencing, next generation sequencing, sequencing by ligation, sequencing by synthesis

ABSTRACT:

In this article next generation sequencing (NGS) methods were characterized. This short review focus on the most important methods of second and third generation sequencing techniques. NGS techniques give unique opportunity to read nucleotides order in DNA strand. Fast, rapid development and progress of genetic sequencing methods began in the seventies. Then two new sequencing techniques were established – enzymatic method (Sanger's sequencing), chemical method (Maxam's and Gilbert's sequencing). Later modifications such as use of capillary electrophoresis, fluorescent nucleotides and application of microprocessor led to total automation of the process and contributed to the methods improvements. Nowadays NGS methods are tools used in many scientific fields e.g. genetic, biotechnology, molecular biology and modern plant breeding programs. Technologies presented in this article have many advantages, but are also not without flaws. Reading accuracy and the overall costs indicate that method based on SOLiD system is the most promising next generation sequencing. Sequencing by ligation (SOLiD) characterized by high precision but also run time is longer compared to the other method described in this article.

1. WSTĘP

Analiza DNA stanowi podstawę dla wielu współczesnych nauk, m.in. genetyki, biotechnologii, a także medycyny i kryminologii. W ostatnich latach nastąpił ogromny postęp w dziedzinie badań genetycznych, związany głównie z opracowaniem nowych metod sekwencjonowania DNA (kwas deoksyrybonukleinowy, z ang. *deoxyribonucleic acid*) [2, 4, 20]. Sekwencjonowanie to proces mający na celu precyzyjne ustalenie sekwencji nukleotydów, wchodzących w skład łańcucha DNA [14].

Stosunkowo szybkie i wydajne procedury sekwencjonowania kwasów nukleinowych opracowano dopiero w drugiej połowie lat 70. XX wieku. Niemalże równocześnie opublikowane zostały dwie techniki, umożliwiające poznanie sekwencji zasad w łańcuchu DNA: pierwsza, zwana metodą chemicznej degradacji łańcucha DNA, zaproponowana przez Maxama i Gilberta, oraz druga – metoda terminacji łańcucha, stworzona przez Sangera i jego współpracowników. Początkowo obie te metody cieszyły się jednakową popularnością, jednakże po latach, to metoda Sangera stała się powszechniej stosowana, m.in. ze względu na możliwość jej automatyzacji [9, 18].

Impulsem do poszukiwania nowych, wysoko-wydajnych metod sekwencjonowania DNA był

przebieg projektu sekwencjonowania genomu człowieka, który pomimo wysiłków licznych zespołów badawczych pochłonął znacznie więcej funduszy oraz czasu, niż początkowo przewidywano. Pomimo dużej dokładności odczytów i wysokiej wiarygodności sekwencji stosowane wówczas metody sekwencjonowania okazały się zbyt mało wydajne dla szybkiego odczytu ogromnej, liczącej około 3 mld par zasad sekwencji genomowej [4]. Na podstawie osiągnięć Sangera oraz Maxama i Gilberta w zakresie technik sekwencjonowania wyodrębniono kolejne metody, określane mianem pierwszej generacji sekwencjonowania. Biorąc pod uwagę różnorodność stosowanych procedur, złożoność niezbędnej aparatury, zakres pojedynczego odczytu, a przede wszystkim wymaganą ilość próby wyjściowej do analizy, sklasyfikowano techniki sekwencjonowania według stopnia ich technologicznego zaawansowania, nadając im kolejne numery [2, 12].

Sekwencjonowanie pierwszej generacji obejmuje metody opracowane przez Sangera oraz Maxama i Gilberta. Sekwencjonowanie drugiej i trzeciej generacji określane jest wspólnym mianem sekwencjonowania nowej generacji – NGS (ang. *Next-Generation Sequencing*). Głównymi zaletami najnowszych metod sekwencjonowania są przede wszystkim wysoka przepustowość i niski koszt analizy jednej próbki w porównaniu do

starszych technik. Dodatkowo istnieje możliwość analizy ogromnej ilości danych w stosunkowo krótkim czasie. Technologie NGS generują dużą liczbę wyników, które następnie należy przetwarzać, wykorzystując innowacyjne narzędzie bioinformatyczne oraz komputery, charakteryzujące się dużą mocą obliczeniową. Ma to ogromne znaczenie w przypadku badań nad genomami roślin, charakteryzującymi się dużym rozmiarem, związanym z występowaniem poliploidalności oraz wielu sekwencji niekodujących. Obecnie wyzwaniem jest ich przetwarzanie oraz analiza danych, umożliwiająca ich aplikacyjne zastosowanie, m.in. w nowoczesnej hodowli i procesie efektywniejszego ulepszania odmian roślin uprawnych [22].

2. SEKWENCJONOWANIE PIERWSZEJ GENERACJI

2.1 Sekwencjonowanie z wykorzystaniem metod enzymatycznych – metoda Sangera

Sekwencjonowanie z wykorzystaniem metod enzymatycznych jest jedną z najpopularniejszych i najszerzej wykorzystywanych technik sekwencjonowania.

Po raz pierwszy metodę tę zastosowano do poznania sekwencji DNA faga X174 o długości 5,4 tys. nukleotydów. W późniejszych latach metoda Sangera ulegała licznym modyfikacjom, m.in. poprzez zastosowanie rekombinowanych polimeraz DNA czy też znaczników fluoroforowych (zamiast izotopowych) [6].

Zmiany standardowych protokołów umożliwiły poznanie najtrudniejszych do analizy rejonów genomu [5, 13]. Klasyczna technika opracowana przez Sanger'a i współpracowników opiera się na wykorzystaniu czterech dideoksynukleotydów (ddNTP). Cząsteczki te, oprócz typowych nukleotydów, uczestniczą w reakcji PCR, a w wyniku wbudowania się ich do polimeryzowanego łańcucha DNA dochodzi do terminacji zachodzącego procesu [18, 19].

Zasadniczą różnicą pomiędzy ddNTP a normalnym nukleotydem jest obecność wodoru na końcu 3' każdego dideoksynukleotydu zamiast standardowej grupy -OH. Ta modyfikacja sprawia, że niemożliwe staje się dołączenie kolejnego nukleotydu do cząsteczki ddNTP, w wyniku czego następuje koniec syntezy DNA. Proces syntezy DNA przy użyciu wzorca jest dokonywany niezależnie od siebie w 4 próbkach, różniących się typem ddNTP zawartego w roztworze.

W momencie, kiedy dochodzi do zakończenia reakcji we wszystkich czterech probówkach reakcyjnych, ich produkty przygotowywane są do elektroforezy żelowej. W celu rozdzielenia fragmentów o różnych długościach zawartość każdej z probówek jest umieszczana w kapilarze wypełnionej żelem. Użycie wysokorozdzielczej elektroforezy żelowej umożliwia rozdzielenie fragmentów różniących się długością z dokładnością do jednego nukleotydu. Następnie żel jest naświetlany promieniami UV lub Roentgena w zależności od metody znakowania DNA. W tej fazie możliwe jest już uzyskanie obrazu wizualizującego wyznakowane wcześniej nukleotydy ddNTP – terminatory reakcji syntezy DNA, w różnych odległościach od początku sekwencji (tj. startera). Dzięki temu uzyskiwany jest autoradiogram, czyli obraz występowania kolejnych nukleotydów w badanym DNA. Jego odczytanie jest końcowym etapem metody, dzięki któremu poznawana jest kolejność występowania nukleotydów w analizowanej sekwencji DNA [11, 18, 19].

2.2 Sekwencjonowanie z wykorzystaniem metod chemicznych – metoda Maxama i Gilberta

Metoda Maxama-Gilberta, która została stworzona w latach 1976-1977 przez Allana Maxama i Waltera Gilberta, nazywana jest również sekwencjonowaniem chemicznym [9]. Metoda ta wykorzystuje specjalnie dobrane związki chemiczne do przeprowadzania cięć w odpowiednim miejscu (G, A+G, T+C, C), dzięki czemu otrzymywane są fragmenty badanego DNA kończące się odpowiednim, znanym nukleotydem.

W ramach sekwencjonowania do podziałów wykorzystywanych jest wiele kopii tej samej sekwencji. Jeden koniec (5') każdej z nich jest radioaktywnie wyznakowany, zazwyczaj za pomocą radioaktywnego fosforu. Fragmenty DNA z czterech niezależnych reakcji otrzymanych w wyniku cięć chemicznych umieszcza się obok siebie na płytce z żelem i poddaje elektroforezie. Umożliwia to rozdzielenie fragmentów o tej samej liczbie nukleotydów, a dzięki radioaktywnemu oznakowaniu danego końca istnieje możliwość otrzymania obrazu, prezentującego poszczególne frakcje. Interpretując prążki w kolejności od najbardziej wysuniętych względem miejsca naniesienia próbek na płytkę, można określić sekwencję analizowanego DNA [9].

2.3 Pirosekwencjonowanie

Pirosekwencjonowanie jest technologią oznaczania sekwencji kwasów nukleinowych, którą opracował w 1996 roku w Królewskim Instytucie Technologicznym w Sztokholmie profesor Pal Nyren i jego uczeń Mostafa Ronaghi. Metoda ta nazywana jest też sekwencjonowaniem przez syntezę (DNA) w czasie rzeczywistym [16].

Bazuje ona na pomiarze ilości pirofosforanu (PP_i), uwalnianego w momencie włączenia komplementarnej zasady do nowo powstającej nici DNA. Detekcji dokonuje się pośrednio, poprzez rejestrację sygnału świetlnego.

Pierwszym krokiem do rozpoczęcia pirosekwencjonowania jest przygotowanie jednoniciowej matrycy za pomocą reakcji PCR (ang. *Polymerase Chain Reaction*) oraz jej oczyszczenie z nieprzyłączonych nukleotydów i starterów. Następnie przeprowadza się hybrydizację matrycy ze znakowanym biotyną primerem oraz inkubację z odpowiednimi enzymami (polimerazą DNA, sulfurylazą ATP, lucyferazą, apyrazą) i substratami reakcji (adenozyno-5' -fosfosiarczanem i lucyferyną).

W trakcie reakcji do mieszaniny pojedynczo, kolejno po sobie, dodawane są nukleotydy, które przyłączane są komplementarnie do matrycy DNA. Powoduje to uwolnienie pirofosforanu (PP_i) w ilości odpowiadającej liczbie przyłączonych nukleotydów.

PP_i jest substratem dla sulfurylasy ATP, która w obecności adenozy-no-5-fosfosiarczanu (APS) katalizuje przekształcenie pirofosforanu do ATP, napędzającego reakcję katalizowaną przez kolejny enzym – lucyferazę - i polegającą na konwersji lucyferyny do oksylucyferyny, czemu towarzyszy zjawisko chemiluminescencji, które rejestrowane jest przez kamerę ze światłoczułą matrycą CCD. Wyniki wyświetlane są w czasie rzeczywistym na ekranie komputera na wykresie zwanym pirogramem. Wysokość i ilość pików jest w nim proporcjonalna do liczby przyłączonych nukleotydów, a ich kolejność pozwala bezpośrednio odczytać sekwencję powstającej nici. Pozostałe nukleotydy, które nie zostały przyłączone do matrycy, są degradowane przez apyrazę [15, 16].

Zaletą pirosekwencjonowania jest wysoka czułość, wiarygodność, powtarzalność wyników (dokładność powyżej 99%) oraz specyficzność. Dzięki użyciu zdefiniowanych starterów powielony i poddany analizie zostaje jedynie wybrany fragment DNA. Dzięki pominięciu elektroforezy analiza przebiega znacznie szybciej (~10 minut)

w porównaniu z konwencjonalnym sekwencjonowaniem Sanger (24-55 h). Kolejną zaletą tej techniki jest także wysoki stopień automatyzacji – proces dodawania nukleotydów i analizy wyników wykonywany jest przez urządzenie. Jednocześnie analizowanie dużej liczby prób pozwala na obniżenie kosztów prowadzonych badań [15, 16].

3. SEKWENCJONOWANIE NOWEJ GENERACJI

3.1 Sekwencjonowanie techniką 454

Metoda określana mianem 454 stanowiła pierwszą z technik sekwencjonowania nowej generacji. Została ona wprowadzona w 2005 r. Wymaga przeprowadzenia kolejno zachodzących po sobie procesów:

- przygotowania biblioteki jednoniciowego DNA,
- powielenia składowych powyższej biblioteki w procesie amplifikacji klonalnej z zastosowaniem emPCR (reakcja łańcuchowa polimerazy w emulsji wodno-olejowej),
- sekwencjonowania przez syntezę z zastosowaniem pirosekwencjonowania.

W pierwszym etapie w celu wygenerowania bibliotek jednoniciowego DNA konieczna jest fragmentacja badanego DNA lub cDNA w procesie nebulizacji. Proces ten polega na przepychaniu roztworu DNA przez szczelinę stożkowatego cylindra, za pomocą wytworzonego nadciśnienia. Poprzez dobór odpowiedniej szczeliny można tą metodą uzyskać fragmenty żądanej długości. Do otrzymanych fragmentów, w procesie ligacji, przyłączane są następnie adaptory, które stanowią odcinki DNA, o znanej sekwencji nukleotydów, znakowane na końcu 5' biotyną. Umożliwia to selektywne związanie DNA do ziaren opłaszczonych streptawidyną. Immobilizowane fragmenty DNA poddaje się procesowi denaturacji w celu pozyskania pojedynczych nici DNA które w drugim etapie posłużą jako biblioteki DNA, do reakcji emPCR [3].

Uzyskane jednoniciowe fragmenty DNA są unieruchamiane poprzez hybrydizację do podłoża stałego, w postaci ziaren, zawierających na swojej powierzchni sekwencje oligonukleotydowe, komplementarne do sekwencji jednego z adapterów okalających analizowane fragmenty DNA. W celu otrzymania swoistych mikroreaktorów warunki reakcji są ściśle określone. W jednej kropli wody musi być obecna jedna mikrokulka z pojedynczą, unikalną sekwencją DNA, która posłuży jako matryca podczas procesu amplifikacji. emPCR pozwala na jednoczesne powielenie wszystkich

fragmentów DNA, co umożliwia uzyskanie miliona kopii każdej analizowanej sekwencji. Otrzymane dwuniciowe DNA jest denaturowane do postaci jednoniciowej formy, niezbędnej podczas ostatniego etapu – procesu pirosekwencjonowania. W tym celu dodawane są kulki magnetyczne, opłaszczone streptawidyną, do których wiążą się za pośrednictwem biotyny dwuniciowe amplifikowane DNA. Następnie w wyniku denaturacji pozyskuje się matryce do sekwencjonowania związane z ziarnami.

3.2 Metoda Illumina

Metoda sekwencjonowania zastosowana w sekwencjonatorze firmy Illumina, podobnie jak sekwencjonowanie Sangera i pirosekwencjonowanie, należy do metod sekwencjonowania przez syntezę. Technika ta została opracowana w 2006 r. i podobnie jak metoda 454 wymaga przeprowadzenia trzech etapów:

- przygotowania jednoniciowych bibliotek DNA,
- amplifikacji fragmentów DNA z zastosowaniem reakcji PCR typu „koci grzbiet”
- oraz sekwencjonowania z zastosowaniem znakowanych fluorescencyjnie nukleotydów, tzw. odwracalnych terminatorów (ang. *reversible terminators*).

Tworzenie bibliotek DNA przebiega podobnie jak w przypadku sekwencjonowania w systemie 454. W drugim etapie uzyskane jednoniciowe fragmenty DNA unieruchamiane są na mikroplątce pokrytej gęsto sekwencjami oligonukleotydowymi (sekwencje adaptorowe) poprzez hybrydyzację komplementarnych sekwencji, a następnie są one zaginane w mostek, tzw. „koci grzbiet”. Mostek ten tworzy się na skutek oddziaływania wolnego końca analizowanych fragmentów DNA z drugą unieruchomioną sekwencją oligonukleotydową znajdującą się w pobliżu. Następnie immobilizowane pojedyncze nici DNA poddawane są reakcji polegającej na przemiennej prowadzeniu syntezy i denaturacji, w wyniku czego płytka zawiera tzw. klastry zbudowane z wielu kopii danego wzorca [8].

Tak przygotowana płytka jest przemywana roztworem, zawierającym wszystkie cztery rodzaje odwracalnych terminatorów. W każdym cyklu dochodzi do dołączenia jednego terminatora komplementarnego do syntetyzowanych sekwencji we wszystkich klastrach, a pozostałe nieprzyłączone terminatory są wyflukiwane z płytki. Dzięki fluorescencyjnemu oznakowaniu terminatorów

istnieje możliwość uzyskania obrazu (fotografii), prezentującego, które terminatory zostały przyłączone w poszczególnych klastrach. W momencie, w którym następuje przyłączanie odpowiednich nukleotydów w klastrach, tworzone są przez kamerę sekwencjonatora następne fotografie, rejestrujące syntezę w kolejnych cyklach. Na ich podstawie można odtworzyć kolejność nukleotydów w sekwencji DNA. Przed ponownym przemyciem płytki roztworem odwracalnych terminatorów przeprowadza się reakcje chemiczne, prowadzące do usunięcia fluorescencyjnego znacznika i odblokowania możliwości przyłączenia kolejnego terminatora.

3.3 Sekwencjonowanie z zastosowaniem techniki SOLiD (ang. *Support Oligonucleotide Ligation and Detection*)

Metoda ta po raz pierwszy została wprowadzona w 2007 r. przez firmę Applied Biosystems (Foster City, USA). Opiera się ona na ligacji oligonukleotydów sekwencji znakowanych fluorescencyjnie do matryc immobilizowanych na mikroplątce. W metodzie tej można wyróżnić trzy etapy. Dwa pierwsze są podobne jak w przypadku sekwencjonowania 454, a ostatni etap to sekwencjonowanie z użyciem ligazy DNA.

Po przeprowadzeniu reakcji PCR koniec 3' jednej z nici DNA jest modyfikowany, co umożliwia jej kowalencyjne przyłączenie do stałego podłoża, którym jest szklana mikroplątka. W taki sposób, podobnie jak we wcześniejszej metodzie Solexa, otrzymuje się mikroplątkę z sektorami, reprezentującymi różne powielone fragmenty. Następnie zachodzi hybrydyzacja startera komplementarnego do sekwencji adaptorowej obecnej w każdym fragmencie DNA. W kolejnym etapie następuje przyłączenie się na zasadzie ligacji komplementarnych zmodyfikowanych sekwencji oligonukleotydów. Stanowią one kombinację 16 znanych dinukleotydów w danej sekwencji, znakowanych czterema specyficznymi fluoroforami. W następnej kolejności po usunięciu syntetyzowanej nici dodawany jest starter, o jeden nukleotyd krótszy od wcześniej zastosowanego. Tego rodzaju proces polegający na przemiennej ligacji kolejnych odcinków, a następnie denaturacji powstałej nici, powtarzany jest dla pięciu różnej długości starterów, określanych jako pięć rund. Użycie czterech znaczników oraz specyficznej metody analizy sekwencji, polegającej na zastosowaniu pięciu różnych starterów (długość startera: n, n-1, n-2, n-3,

n-4), powoduje uzyskanie podwójnego pokrycia dla każdej pozycji w sekwencji, natomiast identyfikacja nukleotydów oparta jest na detekcji specyficznego koloru znacznika [21].

3.4 Metoda IonTorrent

Metoda IonTorrent stanowi jedną z najnowocześniejszych nanotechnologii sekwencjonowania, którą przeprowadza się poprzez syntezę z zastosowaniem mikroprocesora, posiadającego na swej powierzchni gęstą sieć stworzoną, przez ponad milion studzienek reakcyjnych sprzężonych z opatentowanym sensorem.

Zasada postępowania w przypadku tego typu sekwencjonowania jest zbliżona do techniki 454, jednak zamiast pomiaru emisji światła dokonywana jest detekcja zmiany pH mieszaniny reakcyjnej, będącej skutkiem wbudowania nukleotydu. Do studzienek wprowadza się mikro-kulki opłaszczane jednoniciowym badaniem DNA, przy czym każda studzienka zawiera inny szablon DNA, co umożliwia masowe równoległe sekwencjonowanie. Proces syntezy nowo powstających nici katalizowany jest przez polimerazę DNA, która włączając kolejne komplementarne nukleotydy, powoduje uwolnienie kationu wodoru do mieszaniny reakcyjnej jako produktu ubocznego.

Na mikroplątkę w każdym cyklu dozowany jest określony rodzaj nukleotydów, a emitowany sygnał jest wprost proporcjonalny do liczby nukleotydów włączonych do nowo powstających nici w określonych studzienkach. Detekcja dokony-

wana jest jednocześnie na powierzchni całej platformy. Działanie takie jest możliwe dzięki zastosowaniu systemu sensorycznego PGM (ang. *Personal Genome Machine*), stanowiącego najmniejszy na świecie pehametr. Wcielenie nukleotydu przejawia się skokiem napięcia, czego konsekwencją jest możliwość dokonania odczytu i określenie rodzaju dołączonego nukleotydu [7, 10, 17, 20]. W Tabeli 1 porównano wydajności poszczególnych metod sekwencjonowania nowej generacji, na tle sekwencjonowania enzymatycznego.

4. WYKORZYSTANIE TECHNIK SEKWENCJONOWANIA NOWEJ GENERACJI W GENETYCE, HODOWLI I BIOTECHNOLOGII ROŚLIN

Wprowadzenie metod NGS umożliwiło poznanie sekwencji nukleotydowej roślin innych niż charakteryzujące się niewielkim genomem organizmy modelowe takie jak *Arabidopsis thaliana* [23]. Zainteresowanie wzbudzają głównie gatunki uprawne takie jak: zboża, kawa i trzcina cukrowa. Koszt takich badań zmniejszył się ponad 1000-krotnie w porównaniu do metod klasycznego sekwencjonowania. Dodatkowo skróceniu uległ czas analiz dzięki możliwości jednoczesnego przeprowadzania milionów reakcji. Od czasu zsekwencjonowania genomu pierwszej rośliny modelowej w 2000 r. poznano sekwencje ponad 100 innych gatunków roślin [24].

Inną strategią sekwencjonowania, wykorzystywaną głównie w celu zbadania interakcji roślin

Tabela 1 Porównanie wydajności metod sekwencjonowania nowej generacji oraz sekwencjonowania enzymatycznego

Platforma	Sanger 3730	454 Roche	SOLiD	Illumina	IonTorrent	HeliScope	PacBio
Obecna firma	Thermo Fisher Scientific	Roche	Life Technologies	Illumina	Life Technologies	Helicos	Pacific Bioscience
Model	Sanger 3730xl	454RocheFLX	SOLiDv4	HiSeq2000	Ion PGM	NA	PacBioRSII
Mechanizm sekwencjonowania	Terminacja łańcucha dideoxy	Pirosekwencjonowanie	Sekwencjonowanie przez ligację	Sekwencjonowanie przez syntezę	Synteza (detekcja H+)	Synteza	Synteza
Rodzaj amplifikacji	Dye-terminator PCR	emPCR	emPCR	BridgePCR	emPCR	Brak	Brak
Szybkość wykonania jednego cyklu pracy	2 h	10-24 h	7-14 dni	1-10 dni	2 h	8 dni	0,5-2 h
Długość czytanych fragmentów (pz)	400-900	450-700	~100	100-300	~100	~35	~3500
Koszt/milion pz (\$)	1500	10	0,12	0,10	1-50		11
Najczęstsze błędy	Substytucja	Indel	Błędy A-T	Substytucja	Indel	Indel	Delecje CG
Dokładność	99,999	99,9	99,9	98	99	99-99,9	85

i środowiska, jest wykorzystanie metod NGS do poznania transkryptomu roślin w poszczególnych stanach fizjologicznych. Analizując sekwencje cDNA, otrzymuje się informację o sekwencyjnych znacznikach ekspresji (ESTs), które ulegają transkrypcji w poszczególnych tkankach i organach; mimo pewnych ograniczeń, dane te są bardzo przydatne dla hodowców [25].

Techniki sekwencjonowania nowej generacji umożliwiają również badania jakościowe i ilościowe genów, ulegających ekspresji w różnych warunkach. Dane te, umieszczone w bazach, są podstawą i narzędziem dla przyszłych eksperymentów, a dzięki informacjom tam zamieszczonym można projektować markery molekularne oraz zajmować się transkryptomiką porównawczą. Znając sekwencję nukleotydową genomu bądź transkryptomu, można również znaleźć mutacje pojedynczych nukleotydów (SNP) lub proste powtórzenia sekwencji nukleotydów (SSR). Ze

względu na to, że różnice te występują w sekwencjach niekodujących, wykazują one dużą zmienność i polimorfizm. Dane te przydatne są również w tworzeniu markerów molekularnych typu SNP oraz In/Del, które wykazują charakter kodominujący, a tym samym są najlepszymi kandydatami na markery molekularne, wykorzystywanymi w hodowli i selekcji odpowiednich genotypów do dalszych krzyżowań. Dzięki powyższym cechom markery SNP i SSR mają też wiele innych zastosowań, do których należą m.in. tworzenie molekularnych map sprzężeń czy znajdowanie obszarów QTL (ang. *Quantitative Trait Loci*), które odpowiadają za dziedziczenie cech ilościowych. Dodatkowo wykorzystywane są do analiz pochodzenia, poszukiwania „odcisku palca” odmian hodowlanych, w badaniach różnicowania genetycznego w populacjach, przepływu genów i pracach nad genetyką ewolucyjną roślin [26].

LITERATURA

- [1] Buermans H. P., den Dunnen J. T., Next generation sequencing technology: Advances and applications. *Biochimica et Biophysica Acta*, 1842, (2014), 1932-1941.
- [2] Gurgul A., Ząbek T., Bugno-Poniewierska M., Wykorzystanie wysokowydajnych technik analiz genomu w badaniach naukowych i hodowli zwierząt gospodarskich. *Rocz. Nauk. Zoot.*, 42, (2015), 81-91.
- [3] Shendure J., Ji H., Next-generation DNA sequencing. *Nat. Biotech.*, 26, (2008), 1135-1145.
- [4] Raszka A., Ziemińska A., Wiechetek A., Metody i techniki biotechnologii molekularnej w biotechnologii środowiskowej. *Czasopismo Techniczne*, 2(106), (2009), 101-114.
- [5] Maxam A., Gilbert W., A new method for sequencing DNA. *Proceedings of National Academy of Sciences of the USA*, 74(2), (1977), 560-564.
- [6] Sanger F., Nicklen S., Coulson A. R., DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.*, 74(12), (1975), 5463-5467.
- [7] Piątkowski J., Skalniak A., Bodzioch M., Pach D., Hubalewska-Dydejczyk A., Wprowadzenie do sekwencjonowania ludzkiego genomu w diagnostyce. *Przegląd Lekarski*, 70(7), (2013), 458-462.
- [8] Brautigam A., Gowik U., What can next generation sequencing do for you? Next generation sequencing as a valuable tool in plant research. *Plant Biology*, 12, (2010), 831-841.
- [9] Kotowska M., Zakrzewska-Czerwińska J., Kurs szybkiego czytania DNA – nowoczesne techniki sekwencjonowania. *Biotechnologia*, 4(91), (2010), 24-38.
- [10] Kieleczawa J., Fundamentals of sequencing of difficult templates—an overview. *J. Biomol. Tech.*, 17(3), 2006, 207-217.
- [11] Prober J. M., Trainor G. L., Dam R. J., Hobbs F. W., Robertson C. W., Zagursky R. J., Cocuzza A. J., Jensen M. A., Baumeister K., A system for rapid DNA sequencing with fluorescent chain-terminating dideoxynucleotides. *Science*, 238(4825), (1987), 336-341.

- [12] Sanger, A., Coulson R., A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *Journal of Molecular Biology*, 94(3), (1975), 441-448.
- [13] Ronaghi M., Uhl n M., Nyr n P., A Sequencing Method Based on Real-Time Pyrophosphate. *Science*, 281(5375), (1998), 363-365.
- [14] Ronaghi M., Karamohamed S., Pettersson B., Uhl n M., Nyr n P., Real-time DNA sequencing using detection of pyrophosphate release. *Analytical Biochemistry*, 242(1), (1996), 84-89.
- [15] Dressman D., Yan H., Traverso G., Kinzler K. W., Vogelstein B., Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations, *Proc. Natl. Acad. Sci.*, 100, (2003), 8817-8822.
- [16] Nakano M., Komatsu J., Matsuura S., Takashima K., Katsura S., Mizuno A., Single-molecule PCR using water-in-oil emulsion, *J. Biotechnol.*, 102, (2003), 117-124.
- [17] Mardis R. E., Next-generation DNA sequencing methods. *Annual Review of Genomics and Human Genetics*, 9, (2008), 387-407.
- [18] Valouev A., Ichikawa J., Tonthat T., Stuart J., Ranade S., Peckham H., Zeng K., Malek J., Costa G., McKernan K., Sidow A., Fire A., Johnson S., A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res.*, 18, (2008), 1051-1063.
- [19] Loman N. J., Misra R. V., Dallman T. J., Constantinidou C., Gharbia S. E., Wain J., Pallen M. J., Performance comparison of benchtop high-throughput sequencing platforms. *Nat. Biotechnol.*, 30(5), (2012), 434-439.
- [20] Merriman B., Ion Torrent R&D Team, Rothberg J. M., Progress in Ion Torrent semiconductor chip based sequencing. *Electrophoresis*, 33(23), (2012), 3397-3417.
- [21] Rothberg J. M., Hinz W., Rearick T. M., Schultz J., Mileski W., Davey M., et al., An integrated semiconductor device enabling non-optical genome sequencing. *Nature*, 475(7356), (2011), 348-352.
- [22] Shendure J., Ji H., Next-generation DNA sequencing. *Nat. Biotech.*, 26, (2008), 1135-1145.
- [23] The Arabidopsis Genome Initiative, Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*, 408, (2000), 796-815.
- [24] Michael T. P., VanBuren R., Progress, challenges and the future of crop genomes. *Curr. Opin. Plant Biol.* 24, (2015), 71-81.
- [25] P rez-de-Castro A. M., Vilanova S., Ca izares J., Pascual L., Blanca J. M., D ez M. J., Prohens J., Pic  B., Application of Genomic Tools in Plant Breeding. *Curr. Genomics*, 13, (2012), 179-195.
- [26] Egan A. N., Schlueter J., Spooner D. M., Applications of next-generation sequencing in plant biology. *Am. J. Bot.*, 99(2), (2012), 175-185.