

KATARZYNA PACHOTA
AGNIESZKA NIEDZIELA
RENATA ORŁOWSKA
PIOTR T. BEDNAREK

Zakład Biochemii i Fizjologii Roślin

Instytut Hodowli i Aklimatyzacji Roślin — Państwowy Instytut Badawczy, Radzików

Nowoczesne metody genotypowania DArT i GBS w hodowli gatunków roślin użytkowych

Modern methods of genotyping DArT and GBS in agriculture important crops

Rozwój nowoczesnych metod genotypowania sprawił, że konieczne stało się ich zaprezentowanie i pokazanie potencjału, jaki niosą generowane przez nie markery DNA dla badań nad roślinami. Niniejsza praca została poświęcona omówieniu markerów DArT i SNP, przy czym w ostatnim przypadku skupiono się na markerach wykorzystujących sekwencjonowanie nowej generacji. Przedstawiono metody uzyskania w/w markerów oraz ich zastosowania w genetyce roślin użytkowych. Nacisk położono na identyfikację markerów cech użytkowych w połączeniu z mapowaniem cech ilościowych, mapowaniem asocjacyjnym i selekcją genomową.

Słowa kluczowe: genotypowanie, DArT, Diversity Array Technology, GBS, Genotyping-by-sequencing, SNP

Advances in modern genotyping methods make it necessary to present them and their value for plant research. Thus, the review is dedicated to describing technologies allowing evaluation of high throughput markers such as DArTs or SNPs based on new generation sequencing approach. Various applications of these markers in plant research are discussed. Particular emphasis was put on the possibilities of using them for the identification of markers for agronomical traits using either QTL or association mapping methods. The opportunity of using them in genomic selection was also discussed.

Key words: genotyping, DArT; Diversity Array Technology, GBS, Genotyping-by-sequencing

WSTĘP

Markery genetyczne to różnego rodzaju znaczniki, za pomocą, których możliwe jest śledzenie określonego regionu DNA sprzężonego (czy asocjowanego) na przykład z pożądaną cechą. Wykorzystywane są do identyfikacji zmienności między osobnikami, odmianami i gatunkami (Semagn, 2006). Istnieje przyjęty podział markerów na:

Redaktor prowadzący: Katarzyna Mikołajczyk

morfolologiczne, białka strukturalne i funkcjonalne (izoenzymy) oraz markery DNA. Markery morfolologiczne w literaturze pojawiły się już w latach 30 (Sax, 1932) i były wykorzystywane do rozpoznawania cech zewnętrznych organizmu takich jak np. karłowatość czy zmieniona morfologia liścia. Ze względu na liczne wady i silny efekt epistatyczny znalazły one ograniczone zastosowanie. Markery białek funkcjonalnych, czyli izoenzymy (Markert i Møller, 1959), jako pierwsze wykorzystano do szacowania stopnia polimorfizmu oraz oceny struktury genetycznej populacji. Z oczywistych powodów liczba markerów tego typu była ograniczona, a analiza pracochłonna (Knopkiewicz i in., 2012), co ograniczało ich praktyczne wykorzystanie. Dopiero opracowanie markerów molekularnych DNA (Botstein i in., 1980; Welsh i McClelland, 1990; Akkaya i in., 1992; Vos i in., 1995; Jaccound i in., 2001) bazujących na wykorzystaniu enzymów (endonukleaz) restrykcyjnych i na reakcji łańcuchowej polimerazy (ang. Polymerase Chain Reaction, PCR) (Williams, 1989), sprawiło, że technologie markerowe zyskały szerokie zastosowanie w badaniach nad roślinami.

Idealny marker molekularny DNA powinien charakteryzować się łatwością identyfikacji, specyficnością dla danego gatunku, wysokim poziomem polimorfizmu oraz kodominującym charakterem dziedziczenia (Tomar i in., 2010). Zaletą markerów molekularnych, jest fakt, iż nie są one modyfikowane przez środowisko (Staub i in., 1997), co czyni je przydatnymi w programach hodowlanych (Düzyaman, 2005). Markery są istotne przy tworzeniu map genetycznych, pozwalają na lokalizowanie *loci* cech ilościowych (z ang. Quantitative Trait Loci, QTL) oraz mogą być przydatne do identyfikacji odmian (Hayward i in., 2014). Techniki oparte na markerach umożliwiają identyfikację polimorfizmu i struktury materiału genetycznego (Zagalska-Neubauer i Dubiec, 2007). Wykrywane zmiany w sekwencji DNA mogą być powodowane przez polimorfizm pojedynczych nukleotydów, insercje, delecje, czy też translokacje (Ziółkowski i in., 2010). Aby wykryć polimorfizm można zastosować analizę restrykcyjną, hybrydyzację lub amplifikację sekwencji DNA (Wenzl i in., 2008). Analizy te są łatwe do przeprowadzenia, a także nie wymagają zwykle dużych nakładów czasowych. Niezbędne staje się jednak badanie dużych populacji roślinnych, co powoduje konieczność wykorzystania nowocześniejszych narzędzi do genotypowania, umożliwiających szybkie analizowanie badanego materiału. W niniejszej pracy przedstawiono dwie wysokoprzepustowe technologie: markery DArT i wariant markerowania genetycznego z wykorzystaniem SNP bazujący na sekwencjonowaniu nowej generacji. Markery DArT należy zaklasyfikować do markerów mieszanych (Bednarek i Chwedorzewska, 2001) bazujących na trawieniu enzymami restrykcyjnymi amplifikacji PCR endonukleazach restrykcji, PCR i hybrydyzacji, natomiast technologię markerów identyfikującą polimorfizm pojedynczych nukleotydów należy zaliczyć do wysokowydajnych markerów sekwencyjnych typu SNP (ang. Single Nucleotide Polymorphism, polimorfizm pojedynczych nukleotydów).

WYBRANE METODY GENOTYPOWANIA

Pojęcie genotypowania odnosi się do technik opartych na markerach, które umożliwiają różnicowanie genetyczne badanych materiałów. Istnieje szereg podziałów metod genotypowania (Semagn i in., 2006; Mondini i in., 2009; Sztuba-Solińska, 2005). Według Sztuby-Solińskiej (2005), można wyróżnić techniki hybrydizacyjne DNA, techniki bazujące na amplifikacji PCR, (losowej lub specyficznej), oraz techniki mieszane bazujące zarówno na wykorzystaniu endonukleaz restrykcyjnych jak i PCR.

Do technik hybrydizacyjnych zalicza się metodę RFLP (ang. Restriction Fragments Length Polymorphism) (Botstein i in., 1980). Metoda ta bazuje na trawieniu genomowego DNA wykorzystując endonukleazy restrykcyjne. Produkty trawienia rozdzielane są elektroforetycznie a następnie przeprowadza się hybrydizację z sondą DNA stosując metodę Southern (Botstein i in., 1980). Zaletą omawianej metody jest kodominujący charakter markerów RFLP, możliwość rozpoznania dominujących i recesywnych alleli, natomiast wadą konieczność uzyskania dużej ilości wyjściowego DNA. Metoda uważana jest za mało wydajną, identyfikuje ograniczoną liczbę polimorfizmów i jest pracochłonna (Mondini i in., 2009). Markery RFLP wykorzystano do selekcji materiałów hodowlanych i opracowania map genetycznych (Chen i Gustafson, 1995), czy do kotwiczenia (ang. anchoring) map genetycznych (Bednarek i Chwedorzewska, 2001).

Przykładem metody w oparciu o amplifikację PCR, z wykorzystaniem arbitralnie wybranego, zwykle 10-nukleotydu startera jest RAPD (ang. Random Amplified Polymorphic DNA) (Williams i in., 1990; Welsh i McClelland, 1990). Produkty amplifikacji są rozdzielane w żelu agarozowym i uwidaczniane poprzez np. barwienie bromkiem etydy. Główną zaletą RAPD jest prosta metodyka i mała ilość matrycowego DNA niezbędna do przeprowadzenia powielania, wadą natomiast jest ich dominujący charakter (Yazdani i in., 1995). Markery RAPD okazały się przydatne między innymi do identyfikacji odmian uprawnych, w badaniach taksonomicznych, a także przy opracowaniu map genetycznych (Reiter i in., 1992). Obecnie technika RAPD ma ograniczone zastosowanie.

Przykładem metody opierającej się o amplifikację PCR z zastosowaniem pary starterów specyficznych, jest analiza polimorfizmu *loci* mikrosatelitarnych (ang. Simple Sequence Repeats, SSR) (Akkaya i in., 1992). Stosowane startery są specyficzne dla unikalnych sekwencji nukleotydujących oskrzydających poszczególne *loci* SSR. Sekwencje te są gatunkowo specyficzne, a polimorfizm *loci* SSR wynika z różnic długości amplifikowanych fragmentów spowodowanych różną ilością powtórzeń sekwencji tandemowych tworzących dane *locus* SSR. Istotne jest to, że *loci* SSR charakteryzują się wysokim polimorfizmem, dziedziczą się w sposób mendelowski, a markery SSR mają charakter kodominujący (Selkoe i Toonen, 2006). W związku z tym, stanowią dogodną, prostą i powtarzalną metodę analizy polimorfizmu genomów roślin użytkowych, znajdującą zastosowanie do mapowania QTL, asocjacyjnego, a także badania zróżnicowania genetycznego w obrębie populacji hodowlanych i kolekcji odmian. Produkty amplifikacji SSR można rozdzielać w 'tradycyjnych' żelach

poliakrylamidowych, (gdzie produkty mogą być znakowane radioaktywnie, bądź barwione srebrem lub innymi znacznikami) (Zane i in., 2002), a także coraz powszechniej, między innymi, metodą elektroforezy kapilarnej (ang. high performance capillary electrophoresis, HIPCE), przy czym znakowane fluorescencyjnie produkty analizowane są w systemie multipleks, co znacznie zwiększa efektywność. W przypadku genomów niezsekwencjonowanych, *loci* SSR dla danego gatunku identyfikuje się poprzez przeszukiwanie bibliotek genomowych z zastosowaniem sond specyficznych, a następnie sekwencjonowanie i projektowanie specyficznych par starterów dla sekwencji oskrzydających dane *locus* mikrosatelitarne. Natomiast, coraz powszechniej stosuje się metodę „in silico”, w związku z sekwencjonowaniem całych genomów roślin użytkowych. Markery SSR znalazły zastosowanie m. in., w charakterystyce zasobów genowych u jabłoni (Hawliczek i in., 2007) oraz hodowli roślin, między innymi w analizie różnorodności genetycznej pomiędzy liniami wsobnymi u kukurydzy (Shehata i in., 2009), a także w mapowaniu genów karłowatości u pszenicy (Ellis i in., 2005).

Do technik mieszanych należy zaliczyć metodę AFLP (ang. Amplified Fragment Length Polymorphism) (Vos i in., 1995). Technika ta polega na trawieniu genomowego DNA dwoma enzymami restrykcyjnymi najczęściej *Eco* RI/*Mse* I lub *Pst* I/*Mse* I w celu uzyskania fragmentów DNA z lepkimi końcami o znanej sekwencji. Po ligacji komplementarnych do tych sekwencji adaptorów następuje amplifikacja wstępna, a następnie amplifikacja z użyciem starterów selektywnych posiadających na końcu 3' od 2 do 4 różnicujących nukleotydów. Produkty PCR rozdzielane są w żelu poliakrylamidowym, a detekcja odbywa przez barwienie srebrem lub metodami izotopowymi przez znakowanie jednego z selektywnych starterów izotopem, np. ³²P (Vos i in., 1995). Obecnie powszechniej stosowana jest elektroforeza kapilarna w żelu poliakrylamidowym, wykorzystująca barwniki fluorescencyjne. Rozdział analizowanego materiału przebiega w długiej kapilarze kwarcowej wypełnionej żelem o określonej średnicy porów, w których szybkość migracji zależna jest od różnicy i wielkości elementów składowych badanego materiału. Zaletą kapilarnego wariantu AFLP jest możliwość automatyzacji analiz, wysoka czułość, powtarzalność wyników, krótki czas analizy a także niewielka ilość analizowanego materiału (Witkiewicz, 2005). Technikę AFLP cechuje także wysoka wydajność identyfikacji fragmentów oraz fakt, iż nie wymaga ona wcześniejszej znajomości sekwencji powielanego fragmentu DNA (Powel i in., 1996). Podstawową wadą technik mieszanych jest głównie dominujący charakter markerów. Metoda ta znalazła zastosowanie między innymi do badania odporności pszenicy na rdzę brunatną (Mago i in., 2005), do poszukiwania markerów sprzężonych z genami cech użytkowych, wzbogacania map genetycznych, a także do identyfikacji zmienności (epi)genetycznej (Bednarek i in., 2007; Machczyńska i in., 2014).

Omówione metody charakteryzuje ograniczona przepustowość, często relatywnie wysoki koszt analizy w przeliczeniu na pojedynczy marker oraz konieczność oceny polimorfizmu identyfikowanego w danym materiale badawczym, zależnego od wyboru starterów umożliwiających wykrycie zróżnicowania genetycznego (Kilian i in., 2005). Obecnie na uwagę zasługują metody genotypowania, które charakteryzują się możliwością generowania dużej liczby danych oraz szybkością wykrywania

polimorfizmu i mniejszymi kosztami w przeliczeniu na pojedynczy marker. Do takich metod należy zaliczyć technologię DArT, która opiera się o markery hybrydizacyjne oraz coraz powszechniej wykorzystywaną metodę genotyping-by-sequencing (GBS), bazującą na markerach sekwencyjnych.

TECHNOLOGIA DArT

Technologia DArT (ang. Diversity Arrays Technology) (Jaccound i in., 2001), została opracowana przez zespół dr. Andrzeja Kiliana w Centrum Biologii Molekularnej w Australii. Metoda DArT w dużej mierze jest podobna do AFLP, jednak zamiast frakcjonowania produktów reakcji PCR np. w żelu akrylamidowym, wykorzystuje hybrydizację badanego DNA z sondami umieszczonymi na mikromacierzach. Źródłem zmienności w DArT jest polimorfizm pojedynczych nukleotydów (ang. Single Nucleotide Polymorphism, SNP) w obszarze miejsca restrykcji, zmiany insercyjno- delecyjne (InDel), zmiany wzorów metylacji DNA oraz sekwencje powtórzone (Seroczyńska i Kilian, 2010). Analiza bioinformatyczna umożliwia badanie dużej ilości danych generowanych w pojedynczym doświadczeniu (Kilian i in., 2005). Metoda DArT nie wymaga znajomości sekwencji DNA w przeciwieństwie do np., metody SSR (Seroczyńska i Kilian, 2010). Koniecznym jest jednak opracowanie bibliotek genomowych celem uzyskania sond. To właśnie etap opracowania sond jest najbardziej pracochłonną częścią tej technologii. Wadą metody DArT jest potencjalna możliwość wykorzystania w analizie identycznych (redundantnych) sond. Ponieważ wybierane do analiz sondy nie są sekwencjonowane nie można wykluczyć, że sondy o identycznej sekwencji zostaną wykorzystane kilka razy. Może to powodować zmniejszenie liczby użytecznych markerów (Jaccound i in., 2001). Dużym ograniczeniem są również same mikromacierze, gdyż ich pojemność jest ograniczona oraz dominujący charakter markerów DArT. Zaletą markerów DArT jest natomiast możliwość analizy znacznej liczby prób DNA z wykorzystaniem 5000 do 10000 sond ukierunkowanych dla sekwencji niemetylowanych (głównie kodujących) oraz liczne kontrole wewnętrzne analizy, które zapewniają dokładność, powtarzalność i minimalizują występowanie brakujących danych (Huttner i in., 2005), co zwiększa wiarygodność wyników do ok. 99,8% (Xia i in., 2005).

Metoda DArT umożliwia analizę pojedynczych jak i zbiorczych próbek. Analiza pojedynczych preparatów (Jaccound i in., 2001) znalazła zastosowanie np., w genetyce populacyjnej (Sohail Q i in., 2015) czy przy identyfikacji markerów cech (Wenzl i in., 2006). Identyfikację markerów cech czasem korzystniej jest realizować poprzez próby zbiorcze (ang. Bulk Segregant Analysis, BSA) (Michelmore i in., 1991). Połączenie DArT i BSA stwarza taką możliwość i pozwala na typowanie markerów kandydatów cechy (Wenzl i in., 2007) z obniżeniem kosztów prac badawczych. Wykorzystanie tego wariantu metody stało się możliwe dzięki opracowaniu odpowiednich narzędzi bioinformatycznych, które ilościowo oceniają wkład sygnału danego markera w próbę zbiorczej (Magwene i in., 2011). Ten wariant techniki znalazł zastosowanie w przypadku poszukiwania markerów związanych z tolerancją na glin u jęczmienia (Wenzl i in., 2007).

PROCEDURA DArT

Procedura DArT składa się z przygotowania reprezentacji genomowego DNA, redukcji złożoności genomu, sporządzenia biblioteki genomowej, przygotowania sond i preparatów do analizy oraz etapu hybrydyzacji i odczytu uzyskanych wyników.

Przygotowanie reprezentacji genomowej polega na izolacji genomowego DNA. Zależnie od potrzeb DNA izoluje się z reprezentatywnej, zróżnicowanej puli genomowej danego gatunku by objąć możliwie duży poziom zmienności genetycznej. Redukcja złożoności genomu stosowana w technice DArT jest podobna do redukcji złożoności wykorzystywanej w technice AFLP. DNA jest trawiony endonukleazami restrykcyjnymi zwykle wrażliwymi na metylację cytozyny. Wykorzystuje się różne kombinacje enzymów, najczęściej są to endonukleazy *Pst* I, *Rag* I czy *Bst* NI. Dzięki zastosowaniu *Pst* I generowane są fragmenty DNA pochodzące głównie z obszarów niemetylowanych, czyli zwykle z rejonów kodujących genomu. Do uwolnionych fragmentów dołączane są syntetyczne dupleksy DNA (adaptory), które umożliwiają powielenie fragmentów (Wittenberg, 2007). Powielone fragmenty o zdefiniowanej masie cząsteczkowej są klonowane. Wektory zawierające klonowane fragmenty służą do powielenia sond, drukowanych na szklanych mikromacierzach. Powstałe produkty PCR zatężeją się i oczyszczą przed znakowaniem barwnikami fluorescencyjnymi Cy3-dUTP lub Cy5-dUTP. Tak oznaczone preparaty są hybrydyzowane do macierzy z sondami. Gotowe płytki z mikromacierzą po hybrydyzacji są skanowane za pomocą konfokalnego skanera laserowego (Raman i in., 2012). Analiza i identyfikacja polimorfizmów odbywa się komputerowo z wykorzystaniem odpowiedniego oprogramowania.

WYKORZYSTANIE TECHNOLOGII DArT

Pierwotnie metodę DArT opracowano dla diploidalnego genomu ryżu (Akbari i in., 2006). Metoda ta jest wykorzystywana dla dowolnych gatunków w tym również tych o mniejszym znaczeniu gospodarczym np. zioła czy gatunki chronione (Wittenberg, 2007). DArT znalazła zastosowanie w analizie genomów roślin poliploidalnych takich jak: pszenica, trzcina cukrowa i banan (Wenzl i in., 2008). Technologia DArT sprawdza się jako wydajne narzędzie diagnostyczne do badania różnorodności genotypowej (Wenzl i in., 2004 b). Markery DArT z powodzeniem zostały wykorzystane do badania różnorodności genetycznej i struktury populacji chińskiej pszenicy zwyczajnej (*Triticum aestivum* L). Zbadano łącznie 111 odmian i linii hodowlanych z północnych Chin. Wyniki dostarczyły informacji do dalszej selekcji form rodzicielskich i ustalenia heterozygotycznych badanych materiałów dla potrzeb chińskiego programu hodowli pszenicy (Zhang i in., 2011 b). Metoda DArT znalazła szerokie zastosowanie w analizie pokrewieństwa np. u owsa (*Avena sp.*), gdzie zbadano 134 odmiany i wyróżniono grupy odpowiadające formom ozimym i jarym (Tinker i in., 2009). Natomiast badania 232 form grochu bengalskiego (*Cajanus cajan*) wykazały niski stopień zróżnicowania materiałów. Spośród 696 markerów DArT tylko 64 okazało się polimorficznych przy czym wykazano, że formy dzikie są najbardziej zróżnicowane (Yang i in., 2006).

DArT doskonale sprawdza się do tworzenia średniej gęstości map dla gatunków posiadających poliploidalne genomy, na przykład u pszenicy (Wenzl i in., 2004 a). Dla gatunku tego skonstruowano mapę genetyczną zawierającą 749 markerów DArT (Akbari i in., 2006). Rozwój tej metody przyczynił się do opracowania map genetycznych dla wielu gatunków, między innymi dla ryżu, jęczmienia, owsa, trzciny cukrowej, żyta, sorga, (Seroczyńska i Kilian, 2010), oraz częściowo dla łubinu białego (*Lupinus albus* L.) (Vipin i in., 2013), chmielu zwyczajnego (*Humulus lupulus* L.) (Howard i in., 2011) i rzepaku (*Brassica napus* L.) (Raman i in., 2012). W oparciu o mapy genetyczne DArT zidentyfikowano np. QTL odporności na fuzariozę kłosa u jęczmienia (Rheault i in., 2007) czy aż 33 QTL porostania przedźniwnego u żyta (Myśków i in., 2012). W ostatnim przypadku wykorzystano mapę genetyczną zbudowaną na bazie rekombinowanych linii wsobnych (ang. recombinant inbred line, RIL), składającą się z 1285 markerów DArT. Technologia DArT może być wykorzystana między innymi do klonowania w oparciu o mapy genetyczne (Wenzl i in., 2004 b), czy w mapowaniu asocjacyjnym (ang. Association Mapping, AM) (Pritchard i in., 2000). Mapowanie asocjacyjne polega na identyfikacji markerów asocjowanych z określoną cechą w obrębie badanych odmian i/lub linii (Pritchard i in., 2000). Pierwsze mapowanie asocjacyjne opisano u pszenicy (Crossa i in., 2007), gdzie zidentyfikowano markery związane z odpornością na rdzę zbożową, rdzę żółtą, mączniakiem prawdziwym oraz także plonem ziarna. Przebadano 170 linii pszenicy ozimej (*Triticum aestivum* L.) wyprowadzonych z 5 odmian z CIMMYT (Centrum Rozwoju Odmian Pszenicy i Kukurydzy). W oparciu o mapę genetyczną zawierającą 1644 markery, z czego 813 stanowiły markery DArT, na podstawie AM wytypowano szereg markerów badanych cech, które umieszczono na odpowiednich chromosomach gatunku. Markery DArT okazały się również użyteczne w mapowaniu asocjacyjnym u pszenicy (Yu i in., 2014). Autorzy zidentyfikowali silnie zasocjowane markery związane z ważnymi cechami agrotechnicznymi użytecznymi w programach hodowlanych.

Innym zastosowaniem markerów DArT jest ich wykorzystanie w selekcji genomowej (ang. Genomic Selection, GS) (Meuwissen i in., 2001). GS jest podejściem, które pozwala prowadzić selekcję roślin w oparciu o całkowitą pulę markerów DNA dla wybranego modelu statystycznego. Selekcja jest prowadzona na bazie indeksów będących wypadkową powiązania wszystkich markerów z cechą. Selekcji podlega cały genom a nie jego wybrane obszary, uwzględniana jest całkowita złożoność interakcji genów a nie tylko tych obszarów, które odpowiadają za część wariacji cechy. GS pozwala ograniczyć konieczność fenotypowania i skrócenie cyklu hodowlanego (Desta i Ortiz, 2014). Jako pierwszy metodę opisał Meuwissen (Meuwissen i in., 2001), który zbadał dokładność selekcji genomowej przeprowadzonej z wykorzystaniem techniki DArT i porównał z selekcją fenotypową oraz z selekcją wspartą markerami molekularnymi MAS (ang. Marker Assisted Selection). Selekcja genomowa okazała się w 28% dokładniejsza od tradycyjnej selekcji wspomagananej markerami oraz niewiele mniej dokładna od selekcji fenotypowej. Wyniki badania dowodzą, że GS pozwala na zwiększenie rentowności hodowli (Heffner i in., 2011). Metoda z powodzeniem została wykorzystana u jęczmienia (Zhong i in., 2009), owsa (Asoro i in., 2011), a także

doskonale sprawdza się w poprawie efektywności hodowli gatunków długoletnich np. u eukaliptusa (*Eucalyptus L'Her*) (Grattapaglia i in., 2011).

MARKERY POLIMORFIZMU POJEDYNCZYCH NUKLEOTYDÓW Z ZASTOSOWANIEM SEKWENCJONOWANIA NOWEJ GENERACJI

Nowoczesne metody identyfikujące polimorfizm pojedynczych nukleotydów, SNP, wykorzystują technologię sekwencjonowania nowej generacji (ang. Next Generation Sequencing, NGS). Jako NGS określa się techniki sekwencjonowania opracowane w XXI wieku, zapewniające wyższą wydajność i przepustowość w stosunku do powszechnie stosowanej wcześniej techniki sekwencjonowania Sangera (Sanger i in., 1977). Do najpowszechniejszych technik NGS należą: pirosekwencjonowanie 454 (Roche, 2010), technika Solexa (Illumina), platforma SOLiD (Applied Biosystems), Polonator (Dover/Harvard) oraz HeliScope Single Molecule Sequencer (Helicos). Technologie te zapewniają niedrogie odczyty sekwencji całych genomów w wyniku zastosowania metod takich jak immunoprecypitacja chromatyn, mapowanie mutacji, wykrywanie polimorfizmów oraz wykrywanie niekodujących sekwencji RNA (Mardis i in., 2008). Metody sekwencjonowania takie jak: RAD (ang. Restriction site Associated DNA) (Baird i in., 2008), MSG (ang. Multiplexed Shotgun Genotyping) (Andolfatto i in., 2011), BSR-Seq (ang. Bulk segregant RNA-Seq) (Liu i in., 2012) umożliwiają identyfikację znacznej liczby markerów i pozwalają na dokładniejsze badanie wielu loci w małej ilości próbek. Szczegółowy opis technologii sekwencjonowania nowej generacji wykracza poza ramy niniejszej pracy. Obszerne dane dotyczące zagadnienia są dostępne w licznych pracach przeglądowych (Shendure i Ji., 2008; Mamanova i in., 2010; Zhang i in., 2011 a; Cronn i in., 2012; Buermans i Dunnen, 2014).

PROCEDURA METODY GBS

Metoda wykorzystująca podejście stosowane w Illumina dała podstawy do opracowania procedur GBS (Elshire i in., 2011 b), jak również DArTseq (Sansaloni i in., 2011) służących identyfikacji markerów pojedynczych polimorfizmów (SNP). Procedura GBS obejmuje kilka etapów, do których zalicza się przygotowanie prób DNA, trawienie genomowego DNA enzymami restrykcyjnymi, ligację adaptorów oraz niezależne utworzenie poszczególnych bibliotek i ich końcowe połączenie. Dalej odbywa się amplifikacja produktów oraz sekwencjonowanie i analiza wyników.

Przygotowanie prób DNA polega na izolacji DNA, a następnie ich normalizacji w celu zapewnienia identycznej ilości preparatu. Kolejnym etapem jest trawienie genomowego DNA enzymami restrykcyjnymi np. *Ape* KI, *Pst* I, *Msp* I w celu redukcji złożoności genomu. Wykorzystanie enzymów restrykcyjnych do kontrolowanej redukcji złożoności genomu w połączeniu z NGS po raz pierwszy opisał Baird i in. (2008). W oryginalnej metodzie GBS stosowano jeden enzym *Ape* KI (Elshire i in., 2011 b). Potem metoda została rozszerzona o dwa enzymy: jeden rzadko tnący *Pst* I w połączeniu z drugim często tnącym genomowy DNA *Msp* I (Poland i Rife, 2012). Zastosowanie

takiej wersji pozwala na tworzenie jednorodnej biblioteki, umożliwia wykrycie większości fragmentów powiązanych z enzymem rzadko tnącym. Stosowane enzymy restrykcyjne są wrażliwe na metylację, dzięki temu możliwe jest odfiltrowanie obszarów niekodujących oraz metylowanych sekwencji powtarzalnych, takich jak np. elementy mobilne. Fragmenty genomowego DNA pocięte przez enzymy restrykcyjne poddawane są ligacji z adaptorami. Ponieważ ostatnie zawierają w sobie identyfikatory, (zwane w języku angielskim „barcode”), pochodzenie każdej próby jest ściśle zdefiniowane, a identyfikatory muszą spełniać odpowiednie kryteria (Poland i in., 2012 a). Uzyskane produkty po PCR są analizowane pod względem rozmiaru i stanowią bibliotekę genomową, która jest następnie sekwencjonowana za pomocą najważniejszej platformy do NGS, Illuminy.

Niski koszt technologii GBS sprawia, że metoda staje się atrakcyjną i zarazem jest alternatywą dla innych platform genotypowania (Elshire i in., 2011 b). Koszt sekwencjonowania jednej próbki z użyciem markerów GBS nie przekracza 20 \$, a w niedalekiej przyszłości może zejść nawet poniżej 10 \$ (Poland i in., 2012 b). Biorąc pod uwagę, że metoda generuje setki tysięcy markerów SNP, koszt pojedynczego markera jest wielokrotnie niższy niż to ma miejsce w innych metodach genotypowania. Kluczowe zalety systemu GBS to: krótki czas przygotowania preparatów do analizy, zmniejszona liczba reakcji PCR i oczyszczania, brak frakcjonowania oraz zastosowanie identyfikatorów (krótkie sekwencje DNA umożliwiające przypisanie wyników do genotypów (Elshire i in., 2011 a). W większości metod genotypowania, w pierwszej kolejności następuje identyfikacja markerów, a potem genotypowanie, natomiast w metodyce GBS obie czynności są przeprowadzane równocześnie (Poland i Rife, 2012). Liczba SNP, którą można wygenerować za pomocą GBS zależna jest od rozmiaru genomu i różnorodności oraz tego, w jaki sposób próbki odzwierciedlają tę różnorodność (Elshire i in., 2011 b). Technologia GBS umożliwia identyfikację nawet kilkaset tysięcy kodominujących markerów DNA (Niedziela i in., 2015). Wykazano, że np. w przypadku pszenicy, przy użyciu techniki GBS można wykryć znacznie więcej markerów niż za pomocą techniki DaRT (Heslot i in., 2013).

WYKORZYSTANIE MARKERÓW SNP

Fu i Peterson (2011) wykorzystując 3980 SNP badali różnorodność genetyczną 16 lokalnych odmian jęczmienia. Dzięki technice Roche 454 GS FLX (Titanium, jedna z wersji NGS) (Roche 454 Sequencing), ujawniono grupowanie się materiałów zgodne ze znanym pochodzeniem geograficznym odmian (Fu i Peterson, 2011). Markery SNP wykorzystano w analizie różnorodności genetycznej 24 odmian gorczyca. Wykazano, że 26,1% całkowitej zmienności identyfikowanej na bazie 858 polimorficznych SNP dotyczyło odmian lokalnych i odmian hodowlanych, a 24,7% wykazywało zróżnicowanie pomiędzy materiałem o czarnych i żółtych ziarnach. Analiza hierarchiczna ujawniła grupowanie się odmian o czarnym zabarwieniu w obrębie jednego skupienia, natomiast linie hodowlane tworzyły skupienia zgodnie z ich znanym pochodzeniem (Fu i in., 2013). Markery SNP użyto w analizie genetycznej sałaty, prosa, kukurydzy, rzepaku oraz łubinu

(Yang i in., 2012; Lu i in., 2013; Sonah i in., 2013; Bus i in., 2012; Truong i in., 2012), w badaniach podstawowych w genetyce owsa (Huang i in., 2014), w charakterystyce genomu pszenicy i jęczmienia (Poland i in., 2012 b), w ocenie różnorodności genetycznej u kukurydzy (Romay i in., 2013), a także w analizie sierocych gatunków roślin bez znanej sekwencji genomu, czego przykładem jest tetraploidalna lucerna (Rocher i in., 2015).

Innym zastosowaniem markerów SNP jest ich wykorzystanie do tworzenia map genetycznych. I tak np. u pszenicy do opracowania mapy bazującej na 182 liniach podwojonych haploidów wykorzystano ok. 60000 markerów SNP. Ponadto integracja markerów SSR/STS i DArT z markerami SNP w GBS na mapie genetycznej umożliwiła określenie sprzężeń tych markerów z QTL wielu cech użytkowych (Saintenac i in., 2013). Markery SNP wykorzystano również do konstruowania dwóch map sprzężeń o wysokiej gęstości u drzewa kauczukowego (*Hevea brasiliensis*). Łącznie wykryto 1114 wspólnych markerów, które umożliwiły opracowanie zintegrowanej mapy składającej się z 2321 markerów o łącznej długości 2052 cM. Pokazano w ten sposób, że markery SNP nadają się do budowy zintegrowanych map u wysoce heterozygotycznych gatunków rolniczych (Pootakham i in., 2015). Wykorzystanie markerów SNP u ryżu (*Oryza sativa*) (Spindel i in., 2013) pozwoliło na opracowanie mapy genetycznej bazującej na 176 rekombinowanych liniach wsobnych (RIL) i identyfikację obszarów genomu odpowiedzialnych za sterility pyłku, identyfikację QTL szerokości liścia i tolerancji na glin (Spindel i in., 2013). Markery SNP użyto również do zbadania odporności na zgniliznę twardzikową (*Sclerotinia stem rot*) u soi. Przebadano 101 linii hodowlanych, wśród których znajdowały się odmiany uznane za odporne. Wykryto 8397 SNP, a następnie oznaczono trzy obszary genomu powiązane z badaną cechą (Iqaira i in., 2015). Ze względu na dostępność dużej liczby markerów SNP są one dobrym rozwiązaniem w połączeniu z mapowaniem asocjacyjnym np. u autotetraploidalnego ziemniaka (Uitdewilligen i in., 2013). Autorzy pracy identyfikowali silną asocjację markerów SNP z genami barwy bulw.

Technologia GBS może być efektywnym narzędziem wykorzystywanym w selekcji genomowej w programach hodowlanych, ze względu na możliwość identyfikacji dużej liczby markerów SNP w przypadku praktycznie dowolnego gatunku (Poland i in., 2012 a). Dla pszenicy, u 54 linii hodowlanych z CIMMYT (Centrum Rozwoju Odmian Pszenicy i Kukurydzy) sprawdzono przewidywanie selekcji genomowej na bazie 41371 polimorficznych SNP. Stwierdzono, że w jednym środowisku dokładność przewidywania selekcji genomowej jest wysoka. W przypadku różnych środowisk, odziedziczalność i dokładność przewidywania prawdopodobnie będzie mniejsza, z powodu wpływu interakcji genotypów ze środowiskiem. To samo ograniczenie dotyczy jednak selekcji fenotypowej (Poland i in., 2012a). Kolejnym przykładem użycia metody selekcji genomowej jest maniok (*Manihot esculenta* Crantz) gatunek uprawny z rodziny wilczomleczowatych, pochodzący z Brazylii. Na podstawie zaobserwowanych wyników wywnioskowano, że selekcja genomowa jest skuteczna w hodowli manioku (Ly i in., 2013).

PODSUMOWANIE

Rozwój nowych metod genotypowania bazujących na markerach hybrydacyjnych, czy też na sekwencjonowaniu nowej generacji sprawia, że są one coraz częściej stosowane w badaniach podstawowych. Dostępność dużej liczby markerów SNP czy powtarzalność technologii DArT oraz ich malejący koszt sprawiają, że w przypadku roślin ważnych gospodarczo, kiedy kryterium czasu jest bardziej istotne niż początkowe nakłady finansowe, nowoczesne metody zaczynają być wykorzystywane w badaniach stosowanych takich jak identyfikacja markerów cech czy też wręcz selekcja na poziomie całych genomów. Należy jednak zdawać sobie sprawę z wad i zalet omawianych markerów DArT czy SNP. Wydaje się, że markery DArT ze względu na ich dominujący charakter mogą czasem być wygodniejsze niż markery SNP (np. w przypadku gatunków poliploidalnych). Ich niewątpliwą zaletą jest dostępność sekwencji DNA sond a tym samym możliwość opracowania markerów specyficznych. Jednocześnie, technologia DArTseq (w przeciwieństwie do GBS) dostarcza dużej puli tak zwanych silicoDArTów, które również cechuje dominujący charakter (taki marker albo występuje, albo nie u danego genotypu, ale nie jest powiązany z różnicą w sekwencji DNA danego markera). Należy również zdawać sobie sprawę z faktu, iż w przypadku, gdy istotna jest lokalizacja chromosomowa, przypisanie grup sprzężeń do konkretnych chromosomów danego gatunku, markery DArT ze względu na ich znaną u wielu gatunków użytkowych lokalizację mogą okazać się lepszym rozwiązaniem niż markery SNP (Niedziela i in., 2014). Warto, więc rozważyć, który typ markerów będzie dawał większe korzyści w przypadku realizacji konkretnych badań.

LITERATURA

- Akbari M., Wenzl P., Caig V., Carling J., Xia L., Yang S., Uszynski G., Mohler V., Lehmensiek A., Kuchel H., Hayden M. J., Howes N., Sharp P., Vaughan P., Rathmell B., Huttner E., Kilian A. 2006. Diversity arrays technology (DArT) for high-throughput profiling of the hexaploid wheat genome. *Theor. Appl. Genet.* 113 (8): 1409 — 1420.
- Akkaya M. S., Bhagwat A. A., Cregan P. B. 1992. Length polymorphisms of simple sequence repeat DNA in soybean. *Genetics* 132 (4): 1131 — 1139.
- Adolfatto P., Davison D., Erezyilmaz D., Hu TT., Mast J., Sunayama-Morita T., Stern D. L. 2011. Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Res.* 21 (4): 610 — 617.
- Asoro F., Newell M., Beavis W., Jannink J. 2011. Accuracy and Training Population Design for Genomic Selection on Quantitative Traits in Elite North American Oats. *Plant Genome* 4 (2): 132 — 144.
- Baird N. A., Etter P. D., Atwood T. S., Currey M. C., Shiver A. L., Lewis Z. A., Selker E. U., Cresko W. A., Johnson E. A. 2008. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE* 3 (10).
- Bednarek P. T., Chwedorzewska K. 2001. Markery molekularne, ich charakterystyka genetyczna oraz wybrane zastosowania w analizie genetycznej roślin. *Biotechnologia* 1: 9 — 34.
- Bednarek P. T., Orłowska R., Koebner R. M. D., Zimny J. 2007. Quantification of the tissue-culture induced variation in barley (*Hordeum vulgare* L.). *BMC Plant Biology* 7 (10).
- Botstein D., White R. L., Skolnick M., Davis R. W. 1980. Construction of genetic linkage map in man using restriction fragment length polymorphism. *AJHG* 32 (3): 314 — 331.

- Buermans H. P., den Dunnen J. T. 2014. Next generation sequencing technology: Advances and applications. *Biochimica et Biophysica Acta* 1842 (10): 1932 — 1941.
- Bus A., Hecht J., Huettel B., Reinhardt R., Stich B. 2012. High-throughput polymorphism detection and genotyping in *Brassica napus* using next-generation RAD sequencing. *BMC Genomics* 13 (281).
- Chen J. M., Gustafson J. P. 1995. Physical mapping of restriction fragment length polymorphisms (RFLPs) in homoeologous group 7 chromosomes of wheat by in situ hybridization. *Heredity* 75 (3): 225 — 233.
- Cronn R., Knaus B. J., Liston A., Maughan P. J., Parks M., Syring J. V., Udall J. 2012. Targeted enrichment strategies for next-generation plant biology. *American Journal of Botany* 99 (2): 291 — 311.
- Crossa J., Burgueño J., Dreisgacker S., Vargas M., Herrera-Foessel S. A., Lillempo M., Singh R. P., Trethowan R., Warburton M., Franco J., Reynolds M., Crouch J. H., Ortiz R. 2007. Association Analysis of Historical Bread Wheat Germplasm Using Additive Genetic Covariance of Relatives and Population Structure. *Genetics* 177 (3): 1889 — 1913.
- Desta Z. A., Ortiz R. 2014. Genomic selection: genome-wide prediction in plant improvement. *Trends in Plant Science* 19 (9): 592 — 601.
- Düzyaman E. 2005. Phenotypic diversity within a collection of distinct okra (*Abelmoschus esculentus*) cultivars derived from Turkish land races. *Genetic Resources and Crop Evolution* 52 (8): 1019 — 1030.
- Ellis M. H., Rebetzke G. J., Azaña F., Richards R. A., Spielmeier W. 2005. Molecular mapping of gibberellin — responsive dwarfing genes in bread wheat. *Theor. Appl. Genet.* 111 (3): 423 — 430.
- Elshire R. J., Glaubitz J. C., Sun Q., Harriman J. V. 2011 a. Genotyping By Sequencing (GBS) Method Overview. http://cbsu.tc.cornell.edu/lab/doc/GBS_Method_Overview1.pdf.
- Elshire R. J., Glaubitz J. C., Sun Q., Poland J. A., Kawamoto K., Buckler E. S., Mitchell S. E. 2011b. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* 6 (5): e19379.
- Fu Y.-B., Cheng B., Peterson G. W. 2013. Genetic diversity analysis of yellow mustard (*Sinapis alba* L.) germplasm based on genotyping by sequencing. *Genetic Resources and Crop Evolution* 61 (3): 579 — 594.
- Fu Y.-B., Peterson G. W. 2011. Genetic diversity analysis with 454 pyrosequencing and genomic reduction confirmed the eastern and western division in the cultivated barley gene pool. *Plant Genome* 4: 226 — 237.
- Grattapaglia D., Resende M. D., Resende M. R., Sansaloni C., Petrolí C., Missiaggia A., Takahashi E., Zamprogno K., Kilian A. 2011. Genomic Selection for growth traits in Eucalyptus: accuracy within and across breeding populations. *BMC Proc.* 5 (7): 016.
- Hawliczek A., Stankiewicz-Kosyl M., Gawroński S. 2007. Wykorzystanie markerów SSR do molekularnej charakterystyki zasobów genowych jabłoni. *Roczniki Akademii Rolniczej w Poznaniu CCCLXXXIII* 41: 315 — 319.
- Hayward A. C., Tollenaere R., Dalton-Morgan J., Batley J. 2014. Molecular marker applications in plants. *Methods Mol. Biol.* 1245: 13 — 27.
- Heffner E. L., Jannink J.-L., Sorrells M. E. 2011. Genomic selection accuracy using multifamily prediction models in a wheat breeding program. *Plant Genome* 4 (1): 65 — 75.
- Heslot N., Rutkoski J., Poland J., Jannink J.-L., Sorrells M. E. 2013. Impact of marker ascertainment bias on genomic selection accuracy and estimates of genetic diversity. *PLoS ONE* 8 (9): e74612.
- Howard E. L., Whittock S. P., Jakše J., Carling J., Matthews P. D., Probasco G., Henning J. A., Darby P., Cerenak A., Javornik B., Kilian A. 2011. High-throughput genotyping of hop (*Humulus lupulus* L.) utilising diversity arrays technology (DArT). *Theor. Appl. Genet.* 122 (7): 1265 — 1280.
- Huang Y., Poland J. A., Wight C. P., Jackson E. W., Tinker N. A. 2014. Using genotyping-by-sequencing (GBS) for genomic discovery in cultivated oat. *PLoS ONE* 9 (7): e102448.
- Huttner E., Wenzl P., Akbari M., Caig V., Carling J., Cayla C., Evers M., Jaccoud D., Peng K., Patarapuwadol S. 2005. Diversity Arrays Technology: A novel tool for harnessing the genetic potential of orphan crops. *Conference of the World Biological Forum. UK* : 145 — 155.
- Iqura E., Humira S., François B. 2015. Association mapping of QTLs for sclerotinia stem rot resistance in a collection of soybean plant introductions using a genotyping by sequencing (GBS) approach. *BMC Plant Biology* 15 (1): 1 — 12.

- Jaccoud D., Peng K., Feinberg D., Kilian A. 2001. Diversity Arrays: a solid state technology for sequence information independent genotyping. *Nucl. Acids Res.* 29 (4): e25.
- Kilian A., Huttner E., Wenzl P., Jaccoud D., Carling J., Caig V., Evers M., Heller-Uszynska K., Uszynski G., Cayla C., Patarapuwadol S., Xia L., Yang S., Thomson B. R. 2005. The fast and the cheap: SNP and DArT-based whole genome profiling for crop improvement. *Proceedings of the International Congress "In the Wake of the Double Helix: From the Green Revolution to the Gene Revolution"* Avenue Media, Bolonia: 443 — 461.
- Knopkiewicz M., Gawłowska M., Świącicki W. 2012. Poszukiwanie polimorficznych markerów zdefiniowanych sekwencyjnie w populacji grochu Carneval x MP1401. *Fragm. Agron.* 29 (4): 87 — 94.
- Liu S., Yeh C.-T., Tang H. M., Nettleton D., Schnable P. S. 2012. Gene mapping via bulked segregant RNA-Seq (BSR-Seq). *PLoS One* 7 (5): e36406.
- Lu F., Lipka A. E., Glaubitz J., Elshire R., Cherney J. H., Casler M. D., Buckler E S, E C. D. 2013. Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network-based SNP discovery protocol. *PLoS Genet.* 9 (1): e1003215.
- Ly D., Hamblin M., Rabbi I., Melaku G., Bakare M., Gauch H. G., Okechukwu R., Dixon A., Kulakow P., Jannink J. 2013. Relatedness and Genotype x Environment Interaction Affect Prediction Accuracies in Genomic Selection: A Study in Cassava. *CS 53* (4): 1312 — 1325.
- Machczyńska J., Orłowska R., Zimny J., Bednarek P. T. 2014. Extended met AFLP approach in studies of tissue culture induced variation (TCIV) in triticale. *Mol. Breed.* 34 (3): 845 — 854.
- Mago R., Bariana H. S., Dundas I. S., Spielmeier W., Lawrence G. J., Pryor A. J., Ellis J. G. 2005. Development of PCR markers for the selection of wheat stem rust resistance genes Sr24 and Sr26 in diverse wheat germplasm. *Theor. Appl. Genet.* 111 (3): 496 — 504.
- Magwene P. M., Willis J. H., Kelly J. K. 2011. The Statistics of Bulk Segregant Analysis Using Next Generation Sequencing. *PLoS Comput. Biol.* 7 (11): e1002255.
- Mamanova L., Coffey A. J., Scott C. E., Kozarewa I., Turner E. H., Kumar A., Howard E., Shendure J., Turner D. J. 2010. Target-enrichment strategies for next-generation sequencing. *Nat. Meth.* 7 (2): 111 — 118.
- Mardis E. R. 2008. The impact of next-generation sequencing technology on genetics. *Trends in Genetics* 24 (3): 133 — 141.
- Markert C. L., Möller F. 1959. Multiple forms of enzymes: tissue, ontogenetic, and species specific patterns. *Proceedings of the National Academy of Sciences of the United States of America* 45 (5): 753 — 763.
- Meuwissen T. H. E., Hayes B. J., Goddard M. E. 2001. Prediction of Total Genetic Value using genome-wide dense marker maps. *Genetics* 157 (4): 1819 — 1829.
- Michelmore R. W., Paran I., Kessely R. V. 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci. U S A* 88 (21): 9828 — 9832.
- Mondini L., Noorani A., Pagnotta M. A. 2009. Assessing plant genetic diversity by molecular tools. *Diversity* 1 (1): 19 — 35.
- Myśków B., Łań A., Hanek M. 2012. Wpływ sposobu zapylenia żyta na mapowanie QTL dla porostania przedżniwnego. *Biul. IHAR* 264: 117 — 125.
- Niedziela A., Bednarek P.T., Labudda M., Mańkowski D.R., Anioł A. 2014. Genetic mapping of a 7R AL Tolerance QTL in triticale (*x Triticosecale Wittmack*). *J. Appl. Genet.* 55(1): 1 — 14.
- Niedziela A., Jarska W., Bednarek P. T. 2015. Wybrane elementy nowoczesnych rozwiązań wpływające na skuteczność programów. *Biul. IHAR* 275: 3 — 15.
- Poland J., Endelman J., Dawson J., Rutkoski J., Wu S., Manes Y., Dreisigacker S., Crossa J., Sanchez-Villeda H., Sorrells M. E., Jannink J. L. 2012 a. Genomic selection in wheat breeding using genotyping-by-sequencing. *Plant Genome* 5: 103 — 113.
- Poland J. A., Brown P. J., Sorrells M. E., Jannink J. L. 2012 b. Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE* 7 (2): e32253.
- Poland J. A., Rife T. W. 2012. Genotyping-by-Sequencing for Plant Breeding and Genetics. *Plant Genome* 5: 92 — 102.

- Pootakham W., Ruang-Areerate P., Jomchai N., Sonthirod C., Sangsrakru D., Yoocha T., Theerawattanasuk K., Nirapathpongporn K., Romruensukharom P., Tragoonrung S., Tangphatsornruang S. 2015. Construction of a high-density integrated genetic linkage map of rubber tree (*Hevea brasiliensis*) using genotyping-by-sequencing (GBS). *Front. Plant Sci.* 6 (367): 6:367.
- Powell W., Morgante M., Andre C., Hanafey M., Vogel J., Tingey S., Rafalski A. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* 2 (3): 225 — 238.
- Pritchard J. K., Stephens M., Rosenberg N. A., Donnelly P. 2000. Association mapping in structured populations. *The American Journal of Human Genetics* 67 (1): 170 — 181.
- Raman H., Raman R., Nelson M. N., Aslam M. N., Rajasekaran R., Wratten N., Cowling W. A., Kilian A., Sharpe A. G., Shondelmaier J. 2012. Diversity array technology markers: genetic diversity analyses and linkage map construction in rapeseed (*Brassica napus* L.). *DNA Res.* 19 (1): 51 — 65.
- Reiter R. S., Williams J., Feldman K. A., Rafalski J. A., Tingey S. V., Scolnik P. A. 1992. Global and local genome mapping in *Arabidopsis thaliana* by using recombinant inbred lines and random amplified polymorphic DNAs. *Proc. Natl. Acad. Sci. U S A* 89 (4): 1477 — 1481.
- Rheault M. E., Dallaire C., Marchand S., Zhang L., Lacroix M., Belzile F. 2007. Using DArT and SSR markers for QTL mapping of Fusarium head blight resistance in six-row barley. *Plant & Animal Genome XV conference San Diego, USA.*
- Roche 454 Sequencing. 2010. Multiplex Identifier (MID) Adaptors for Rapid Library Preparations 454 Life Sciences, Branford, CT.
- Rocher S., Jean M., Castonguay Y., Belzile F. 2015. Validation of Genotyping-By-Sequencing Analysis in Populations of Tetraploid Alfalfa by 454 Sequencing. *PLoS ONE* 10 (6): e0131918.
- Romay M. C., Millard M. J., Glaubitz J. C., Peiffer J. A., Swarts K. L., Casstevens T. M., Elshire R. J., Acharya C. B., Mitchell S. E., Flint-Garcia S. A., McMullen M. D., Holland J. B., Buckler E. S., Gardner C. A. 2013. Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biology* 14 (6): 1 — 18.
- Sanger F., Nicklen S., Coulson A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74 (12): 5463 — 5467.
- Saintenac C., Jiang D., Wang S., Akhunov E. 2013. Sequence-based mapping of the polyploid wheat genome. *G3 (Bethesda)* 3 (7): 1105 — 1114.
- Sansaloni C., Petrolu C., Jaccoud D., Carling J., Detering F., Grattapaglia D., Kilian A. 2011. Diversity Arrays Technology (DArT) and next-generation sequencing combined: genome-wide, high throughput, highly informative genotyping for molecular breeding of Eucalyptus. *BMC Proceedings* 5 (7): 1 — 2.
- Sax K. 1923. The Association of Size Differences with Seed-Coat Pattern and Pigmentation in PHASEOLUS VULGARIS. *Genetics* 8 (6): 552 — 560.
- Selkoe K. A., Toonen R. J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9 (5): 615 — 629.
- Semagn K., Bjørnstad Å., Ndjiondjop M. 2006. An overview of molecular marker methods for plants. *African Journal of Biotechnology* 5 (25).
- Seroczyńska A., Kilian A. 2010. Technologia Dart — nowe narzędzie do analizy zmienności genetycznej. *Zeszyty Problemowe Postępów Nauk Rolniczych* 555: 373 — 388.
- Shehata A. I., Al-Ghethar H. A., Al-Homaidan A. A. 2009. Application of simple sequence repeat (SSR) markers for molecular diversity and heterozygosity analysis in maize inbred lines. *Saudi Journal of Biological Sciences* 16 (2): 57 — 62.
- Shendure J., Ji H. 2008. Next-generation DNA sequencing. *Nat. Biotech.* 26 (10): 1135 — 1145.
- Sohail Q., Manickavelu A., Ban T. 2015. Genetic diversity analysis of Afghan wheat landraces (*Triticum aestivum*) using DArT markers. *Genet. Resour. Crop Evol.* 62 (8): 1147 — 1157.
- Sonah H., Bastien M., Iquira E., Tardivel A., Legare G., Boyle B., Normandeau E., Laroche J., Larose S., Jean M., Belzile F. 2013. An improved genotyping by sequencing (GBS) approach offering increased versatility and efficiency of SNP discovery and genotyping. *PLoS ONE* 8 (1): e54603.
- Spindel J., Wright M., Chen C., Cobb J., Gage J., Harrington S., Lorieux M., Ahmadi N., McCouch S. 2013. Bridging the genotyping gap: using genotyping by sequencing (GBS) to add high-density SNP markers

- and new value to traditional bi-parental mapping and breeding populations. *Theor. Appl. Genet.* 12: 2699 — 2716.
- Staub J. E., Serquen F. C., McCreight J. D. 1997. Genetic diversity in cucumber (*Cucumis sativus* L.): III. An evaluation of Indian germplasm. *Genetic Resour. Crop Evol.* 44 (4): 315 — 326.
- Sztuba-Solińska J. 2005. Systemy markerów molekularnych i ich zastosowanie w hodowli roślin. *Kosmos — Problemy Nauk Biologicznych* 54 (2–3): 227 — 239.
- Tinker N. A., Kilian A., Wight C. P., Heller-Uszynska K., Wenzl P., Rines H. W., Bjørnstad Å., Howarth C. J., Jannink J.-L., Anderson J. M., Rossnagel B. G., Stuthman D. D., Sorrells M. E., Jackson E. W., Tuvešson S., Kolb F. L., Olsson O., Federizzi L. C., Carson M. L., Ohm H. W., Molnar S. J., Scoles G. J., Eckstein P. E., Bonman J. M., Ceplitis A., Langdon T. 2009. New DArT markers for oat provide enhanced map coverage and global germplasm characterization. *BMC Genomics* 10 (1): 1 — 22.
- Tomar R. S., Parakhia M. V., Patel S. V., Golakiya B. A. 2010. *Molecular Markers and Plant Biotechnology*. New India Publishing Agency, New Delhi.
- Truong H. T., Ramos A. M., Yalcin F., de Ruiter M., van der Poel H. J. A., Huvenaars K. H. J., Hogers R. C. J., van Enckevort L. J. G., Janssen A., van Orsouw N. J., van Eijk M. J. T. 2012. Sequence-Based Genotyping for Marker Discovery and Co-Dominant Scoring in Germplasm and Populations. *PLoS ONE* 7 (5): e37565.
- Uitdewilligen J. G. A. M. L., Wolters A.-M. A., D'hoop B. B., Borm T. J. A., Visser R. G. F., van Eck H. J. 2013. A Next-Generation Sequencing Method for Genotyping-by-Sequencing of Highly Heterozygous Autotetraploid Potato. *PLoS ONE* 8 (5): e62355.
- Vipin C. A., Luckett D. J., Harper J. D. I., Ash G. J., Kilian A., Ellwood S. R., Phan H. T. T., Raman H. 2013. Construction of integrated linkage map of a recombinant inbred line population of white lupin (*Lupinus albus* L.). *Breed. Sci.* 63 (3): 292 — 300.
- Vos P., Hogers R., Bleeker M., Reijns M., van de Lee T., Hornes M., Frijters A., Pot J., Peleman J., Kuiper M., et al. 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* 23 (21): 4407 — 4414.
- Welsh J., McClelland M. 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Research* 18 (24): 7213 — 7218.
- Wenzl P., Caig V., Carling J., Cayla C., Evers M., Jaccound D., Patarapuwadol A., Uszyński G., Xia L., Yang S., Huttner E., Killian A. 2004a. Diversity Arrays Technology, a novel tool for harnessing crop genetic diversity. *Proceeding for the 4th International Crop Science Congress*. The Regional Institute Ltd, Brisbane, Australia.
- Wenzl P., Carling J., Kudrna D., Jaccound D., Huttner E., Kleinhofs A., Kilian A. 2004b. Diversity Arrays Technology (DArT) for whole-genome profiling of barley. *Proc. Natl. Acad. Sci. U S A* 101 (26): 9915 — 9920.
- Wenzl P., Huttner E., Carling J., Xia L., Blois H., Caig V., Heller-Uszyńska K., Jaccound D., Hopper C., Aschenbrenner-Kilian G., Evers M., Hok P., Duncan M., Miller K., Uszyński G., Kilian A. 2008. Diversity Arrays Technology (DArT): A generic high-density genotyping platform. *7th International Safflower Conference Wagga Wagga, Australia*: 1 — 7.
- Wenzl P., Li H., Carling J., Zhou M., Raman H., Paul E., Hearnden P., Maier C., Xia L., Caig V., Ovesná J., Cakir M., Poulsen D., Wang J., Raman R., Smith K. P., Muehlbauer G. J., Chalmers K. J., Kleinhofs A., Huttner E., Kilian A. 2006. A high-density consensus map of barley linking DArT markers to SSR, RFLP and STS loci and agricultural traits. *BMC Genomics* 7 (1): 1 — 22.
- Wenzl P., Raman H., Wang J., Zhou M., Huttner E., Kilian A. 2007. A DArT platform for quantitative bulked segregant analysis. *BMC Genomics* 8 (1): 1 — 10.
- Williams J. F. 1989. Optimization strategies for the polymerase chain reaction. *BioTechniques* 7 (7): 762 — 769.
- Williams J. G., Kubelik A. R., Livak K. J., Rafalski J. A., Tingey S. V. 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research* 18 (22): 6531 — 6535.
- Witkiewicz Z. 2005. *Podstawy chromatografii, kapilarne techniki elektromigracyjne*. Wydawnictwo Naukowo-Techniczne, Warszawa. 335 — 400.

- Wittenberg A. H. J. 2007. Genetic mapping using the Diversity Arrays Technology (DArT) — Application and validation using the whole-genome sequences of *Arabidopsis thaliana* and the fungal wheat pathogen *Mycosphaerella graminicola*. MS. Thesis. Wageningen University.
- Xia L., Peng K., Yang S., Wentzl P., de Vicente M. C., Fregene M., Kilian A. 2005. DArT for high-throughput genotyping of Cassava (*Manihot esculenta*) and its wild relatives. *Theor. Appl. Genet.* 110 (6): 1092 — 1098.
- Yang H., Tao Y., Zheng Z., Li C., Sweetingham M. W., Howieson J. G. 2012. Application of next-generation sequencing for rapid marker development in molecular plant breeding: a case study on anthracnose disease resistance in *Lupinus angustifolius* L. *BMC Genomics* 13 (1): 1 — 12.
- Yang S., Pang W., Ash G., Harper J., Carling J., Wenzl P., Huttner E., Zong X., Kilian A. 2006. Low level of genetic diversity in cultivated Pigeonpea compared to its wild relatives is revealed by diversity arrays technology. *Theor. Appl. Genet.* 113 (4): 585 — 595.
- Yazdani R., Yeh F. C., Rimsha J. 1995. Genomic mapping of *Pinus sylvestris* (L.) using random amplified polymorphic DNA markers. *For Genet* 2: 109 — 116.
- Yu H., Deng Z., Xiang C., Tian J. 2014. Analysis of diversity and linkage disequilibrium mapping of agronomic traits on B-genome of wheat. *J. Genomics* 2: 20 — 30.
- Zagalska-Neubauer M., Dubiec A. 2007. Techniki i markery molekularne w badaniach zmienności genetycznej ptaków. *Notatki Ornitologiczne* 48: 193 — 206.
- Zane L., Bargelloni L., Patarnello T. 2002. Strategies for microsatellite isolation: a review. *Molecular Ecology* 11 (1): 1 — 16.
- Zhang J., Chiodini R., Badr A., Zhang G. 2011a. The impact of next-generation sequencing on genomics. *Journal of Genetics and Genomics = Yi chuan xue bao* 38 (3): 95 — 109.
- Zhang L., Liu D., Guo X., Yang W., Sun J., Wang D., Sourdille P., Zhang A. 2011b. Investigation of genetic diversity and population structure of common wheat cultivars in northern China using DArT markers. *BMC Genetics* 12 (1): 1 — 11.
- Zhong S., Dekkers J. C., Fernando R. L., Jannink J. L. 2009. Factors affecting accuracy from genomic selection in populations derived from multiple inbred lines: a Barley case study. *Genetics* 182 (1): 355 — 364.
- Ziółkowski P., Babula-Skowrońska D., Kaczmarek M., Cieśla A., Sadowski J. 2010. Comparative sequencing of genomes: generating of INDEL and SNP genetic markers. *Biotechnologia* 4: 53 — 68.