

STEFAN MALEPSZY, ZBIGNIEW PRZYBECKI, CEZARY KOWALCZUK,
MARCIN FILIPECKI

*Katedra Genetyki
Hodowli i Biotechnologii Roślin
Wydział Ogrodnictwa i Architektury Krajobrazu
Szkoła Główna Gospodarstwa Wiejskiego w Warszawie
Nowoursynowska 159, 02-776 Warszawa
E-mail: stefan_malepszy@sggw.pl
zbigniew_przybecki@sggw.pl
cezary.kowalczuk@gmail.com
marcin_filipecki@sggw.pl*

SEKWENCJONOWANIE GENOMÓW STAJE SIĘ NOWYM SKŁADNIKIEM POSTĘPU W HODOWLI ROŚLIN

WPROWADZENIE

Minęło już ponad 10 lat od ogłoszenia sekwencji genomu ludzkiego, zapowiadane go jako wielki krok w dziejach ludzkości. Osiągnięcie to było efektem pracy wielkiego konsorcjum międzynarodowego (lista autorów publikacji liczy 130 nazwisk), które dysponowało budżetem ponad 3,5 mld dolarów. Od tego czasu jesteśmy coraz częściej informowani o kolejnym organizmie, dla którego odczytano kompletną sekwencję nukleotydową DNA. W elitarnym gronie organizmów o poznanym zapisie prawie wszystkich genów znajdują się ważne rośliny uprawne: ryż, topola, pomidor, melon, ziemniak, ogórek i wiele innych. W pracach nad genomami ziemniaka i ogórka uczestniczyli naukowcy pracujący w Polsce. W przypadku ziemniaka Polacy byli jednym spośród 12 narodowych

partnerów Konsorcjum (THE POTATO GENOME SEQUENCING CONSORTIUM 2011), które ustaliło między innymi, wielkość genomu ziemniaka na 844 mln nukleotydów i występowanie w nim 39 031 genów (więcej informacji także w GROMADKA i współaut. 2011). Obszerny opis tego odkrycia zamieszczono w numerze prestiżowego *Nature* z 14 lipca 2011r. Dwa tygodnie później ukazała się publikacja o genomie ogórka, w której lista autorów zawiera tylko 20 nazwisk z trzech jednostek naukowych, z tego 16 Polaków z jednej katedry.

Porównanie niektórych charakterystyk organizacyjnych wymienionych przedsięwzięć (Tabela 1) pokazuje, iż w ciągu ostatniej dekady nastąpił ogromny postęp metodyczny sekwencjonowania. Postęp ten sprawił, że poznanie sekwencji genomowej może być

Tabela 1. Porównanie wybranych parametrów organizacyjnych dotyczących trzech przedsięwzięć z okresu minionej dekady, w których zsekwencjonowano genomy organizmów eukariotycznych.

Nazwa przedsięwzięcia	Rok ogłoszenia	Czas trwania (lata)	Liczba osób	Liczba instytucji	Budżet (mln\$)
Genom człowieka	2001	15	134	120	3 500
Genom ziemniaka	2011	4,5	97	24	Brak danych
Genom ogórka	2009	2,5	20	4	0,80

Tabela 2. Grupy gatunków roślin uprawnych których genomy zostały zsekwencjonowane lub których sekwencjonowanie jest mocno zaawansowane i należy się spodziewać niebawem ostatecznego wyniku.

Grupa roślin/ gatunek	Nazwa zwyczajowa	Liczba sekwencjonowanych genotypów	Liczba skończonych wg NCBI	Liczba na ukończeniu
zbożowe		59	4	0
<i>Hordeum vulgare</i>	jęczmień	3	0	0
<i>Oryza sativa</i>	ryż	19	1	0
<i>Zea mays</i>	kukurydza	4	1	0
<i>Avena sativa</i>	owies	2	1	0
<i>Triticum</i>	pszenice	8	0	0
<i>Sorghum</i>	sorgo	20	1	0
<i>Secale cereale</i>	żyto	1	0	0
<i>Amaranthus tuberculatus</i>	amarantus	1	0	0
<i>Setaria italica</i>	czumiza; ber	1	0	0
oleiste		12	0	0
<i>Elaeis guineensis</i>	palma oleista	1	0	0
<i>Glycine max</i>	bawełna	3	0	0
<i>Helianthus annuus</i>	słonecznik	1	0	0
<i>Linum usitatissimum</i>	len	1	0	0
<i>Brassica napus</i>	rzepak	3	0	0
<i>Theobroma cacao</i>	kakaowiec	3	0	0
okopowe		5	0	0
<i>Solanum phureja</i>	ziemniak peruwiański	1	0	0
<i>Solanum tuberosum</i>	ziemniak	3	0	0
<i>Beta vulgaris</i>	burak	1	0	0
warzywne		17	0	0
<i>Brassica oleracea</i>	kapusta	1	0	0
<i>Brassica rapa</i>	rzodkiew	7	0	0
<i>Dioscorea alata</i>	słodki ziemniak	1	0	0
<i>Lactuca sativa</i>	sałata	1	0	0
<i>Manihot esculenta</i>	maniok	1	0	0
<i>Cucumis sativus</i>	ogórek	2	0	0
<i>Phaseolus vulgaris</i>	fasola	1	0	0
<i>Solanum lycopersicum</i>	pomidor	2	0	0
<i>Vigna radiata</i>	fasola złota; mungbeen	1	0	0
<i>Castanea mollissima</i>	chiński kasztan	1	0	0
ozdobne		16	0	0
<i>Aquilegia coerulea</i>	orlik	1	0	0
<i>Arundo donax</i>	trzcina ozdobna	1	0	0
<i>Millettia pinnata</i>	pangomia	1	0	0
<i>Mimulus guttatus</i>	kroplik żółty	1	0	0
<i>Miscanthus inensis</i>	Miskant	1	0	0
<i>Cenchrus americanus</i>	Rosplenica	1	0	0
<i>Eucalyptus grandis</i>	eukaliptus	1	0	0

<i>Panicum virgatum</i>	proso różgowe	1	0	0
<i>Phalaenopsis aphrodite</i>	storczyk	3	0	0
<i>Ricinus communis</i>	řącznik	2	0	0
<i>Solanum demissum</i>	dziki ziemniak	3	0	0
owocowe		12	0	0
<i>Citrus</i>	pomarańcza	2	0	0
<i>Fragaria vesca</i>	poziomka	1	0	0
<i>Musa acuminata</i>	banan	1	0	0
<i>Phoenix dactylifera</i>	palma daktylowa	1	0	0
<i>Prunus persica</i>	grusza	3	0	0
<i>Vaccinium macrocarpon</i>	zurawina	1	0	0
<i>Vitis vinifera</i>	winorośl	2	0	0
<i>Malus x domestica</i>	jabłń	1	0	0
pastewne		6	0	0
<i>Cenchrus americanus</i>	rosplenica	1	0	0
<i>Medicago truncatula</i>	lucerna	1	0	0
<i>Solanum phureja</i>	ziemniak peruwiański	1	0	0
<i>Solanum tuberosum</i>	ziemniak	3	0	0
przemysłowe		19	0	0
<i>Asclepias syriaca</i>	trojeść	1	0	0
<i>Ricinus communis</i>	řącznik	2	0	0
<i>Salix purpurea L.</i>	wierzba purpurowa	1	0	0
<i>Spirodela polyrhiza</i>	rzęsa	1	0	0
<i>Suaeda liaotungensis</i>	słonorośl	2	0	0
<i>Zostera marina</i>	trawa morska	1	0	0
<i>Nicotiana benthamia</i>	tytoń	2	0	0
<i>Beta vulgaris</i>	burak	1	0	0
<i>Millettia pinnata</i>	pangomia	1	0	0
<i>Jatropha curcas</i>	obrzydlenieć	7	0	0
sadownicze		10	0	0
<i>Phoenix dactylifera</i>	palma daktylowa	1	0	0
<i>Malus x domestica</i>	jabłń	1	0	0
<i>Citrus</i>	pomarańcza	2	0	0
<i>Musa acuminata</i>	banan	1	0	0
<i>Vitis vinifera</i>	winorośl	2	0	0
<i>Prunus persica</i>	grusza	3	0	0
drzewa krzewy		14	0	0
<i>Pinus taeda</i>	sosna taeda	1	0	0
<i>Populus trichocarpa</i>	topola kalifornijska	1	0	0
<i>Salix purpurea L.</i>	wierzba purpurowa	1	0	0
<i>Phoenix dactylifera</i>	palma daktylowa	1	0	0
<i>Malus x domestica</i>	jabłń	1	0	0
<i>Citrus</i>	pomarańcza	2	0	0
<i>Musa acuminata</i>	banan	1	0	0
<i>Vitis vinifera</i>	winorośl	2	0	0
<i>Prunus persica</i>	grusza	3	0	0

<i>Eucalyptus grandis</i>	eukaliptus	1	0	0
lecznicze		5	0	0
<i>Glycyrrhiza uralensis</i>	Chinese liquorice	1	0	0
<i>Panax ginseng</i>	żeńszeń	1	0	0
<i>Ricinus communis</i>	rącznik	2	0	0
<i>Salvia miltiorrhiza</i>	szałwia czerwona	1	0	0

zrealizowane w zasadzie przez jeden zespół badawczy, przy kosztach niewspółmiernych do sekwencjonowania genomu człowieka. W efekcie genomy można także z powodzeniem sekwencjonować w kraju takim jak Polska, czyli nie należącym do mocarstw naukowych, technologicznych i finansowych. To oznacza, że w najbliższym czasie możemy

oczekiwać dopływu ogromnej ilości nowych informacji o właściwościach genomów poszczególnych gatunków roślin uprawnych (Tabela 2). Wśród wielu skutków tej sytuacji jest także pytanie o wpływ jaki wiedza pochodząca z sekwencjonowania będzie miała na hodowlę roślin.

CZYM JEST SEKWENCJONOWANIE GENOMÓW

Sekwencjonowanie genomów jest odczytaniem zapisu genetycznego, jakim dany organizm dysponuje, wyrażonego w odpowiednim uporządkowaniu nukleotydów (otrzymujemy sekwencję kompletną). W jego wyniku uzyskujemy informację rzeczywistą o strukturze i położeniu genów i innych składowych genomu, a nie, jak dotychczas, informację względną, wynikającą z częstości crossing over i mapowania genetycznego. Sekwencjonowanie można nazwać remanentem genów, w wyniku którego można prognozować pewne szczególne cechy organizmu i wyjaśniać mechanizmy wielu reakcji. Remanent ten jest bardzo drobiazgowy i precyzyjny, gdyż pokazuje nie tylko obecność genów, ale ujawnia zarazem wiele szczegółowych ich właściwości, włącznie z elementami sterującymi aktywnością genów. Pokazuje także odległości fizyczne między genami i miejsca, w których występuje DNA zwany „śmieciwym”. Określenie to nie najlepiej oddaje jego rolę, albowiem powoli odkrywana jest dość niezwykła rola takiego niekodującego DNA w funkcjonowaniu i ewolucji genomu. Zapewne niebawem zaniechamy używania tego pejoratywnego określenia.

Znajomość sekwencji genomu jest dopiero początkiem poznawania mechanizmów funkcjonowania organizmu. Obecny stan wiedzy z zakresu genetyki molekularnej i zastosowanie specjalistycznych programów komputerowych pozwala na wykrycie obecności i przybliżone określenie funkcji większości genów – jednak nie wszystkich. Moż-

na powiedzieć że poznanie sekwencji genomu jest dopiero początkiem badań systemowych (w kontekście całego organizmu) nad funkcją i współdziałaniem genów. Ponadto, poza wciąż jeszcze dużym obszarem niewiedzy o sposobie działania genów zapisanym w dziedziczonej sekwencji DNA, istnieje trudniejszy do badania zapis epigenetyczny.

Zsekwencjonowanie genomu jest przedsięwzięciem złożonym, składającym się z trzech etapów: przygotowania genomu do sekwencjonowania; przeprowadzenia sekwencjonowania oraz złożenia uzyskanej sekwencji. Zakres prac na pierwszym etapie zależy od wielkości genomu i w przypadku nowosekwencjonowanych genomów roślin wyższych polega na przygotowaniu odpowiednich bibliotek zawierających duże fragmenty DNA, pokrywających cały genom. Dokładne ułożenie tych fragmentów wzdłuż chromosomów pozwala na stworzenie tzw. fizycznej mapy klonów, która, przy znajomości niewielkich odcinków sekwencji DNA, będzie stanowiła rusztowanie do układania sekwencji otrzymywanych w następnym etapie. Bardzo szybki postęp w sekwencjonowaniu genomów wynika z ogromnego przyspieszenia i uproszczenia drugiego etapu przedsięwzięcia, głównie dzięki zastosowaniu nowej generacji technologii samego sekwencjonowania, tzw. NGS (ang. next generation sequencing), i coraz doskonalszych programów bioinformatycznych pomagających w obróbce ogromnych ilości informacji. Ze względu na zaawansowanie i koszt aparatury

oraz konieczność utrzymania wysoce wykwalifikowanego personelu technicznego samo sekwencjonowanie przeprowadzają obecnie wyspecjalizowane firmy biotechnologiczne. Co więcej, pojawiają się wciąż nowe technologie i urządzenia do sekwencjonowania przyspieszające zarówno sam odczyt, jak i skracające, bądź eliminujące dosyć żmudny pierwszy etap. Aktualnie do odczytu sekwencji najczęściej stosuje się strategię losowego generowania zachodzących na siebie dosyć krótkich odcinków sekwencji z całego genomu, tzw. WGS (ang. whole genome shotgun). W całej procedurze najbardziej czasochłonne jest jednak złożenie uzyskanej sekwencji, uzupełnienie brakujących rejonów i wyjaśnienie zawilości w obszarach trudnych.

Przykładowo, genom człowieka złożono w pełni dopiero w 2010 r., a pierwsze publikowane wersje genomów są zwykle określane jako draft (ang. szkic, zarys) i stanowią 60-90% całkowitej wielkości genomu.

Należy podkreślić, że nowozsekwencjonowany genom staje się genomem referencyjnym, nieodzwoiercedlającym ani wewnątrzgatunkowej zmienności, ani stanu heterozygotyczności wewnątrz zsekwencjonowanego genomu. Dostęp do genomu referencyjnego radykalnie ułatwia i obniża koszty sekwencjonowania kolejnych przedstawicieli gatunku (tzw. resekwencjonowanie) umożliwiając zarówno badanie polimorfizmu, jak też użycie sekwencjonowania jako metody diagnostycznej.

WYMIERNA WARTOŚĆ SEKWENCJI

Wartość informacji wynikających z zsekwencjonowania mogą zilustrować dwa przykłady, z których jeden dotyczy człowieka i odnosi się zarazem do wartości społecznych, natomiast drugi pokazuje wymierną wartość pieniężną. Otóż Amerykanie zdecydowali się pod koniec lat 80. ubiegłego wieku na podjęcie programu zsekwencjonowania genomu człowieka, kierując się głównie szeroko pojętym aspektem praktycznym, u którego podstaw były dwa oczekiwania. Jedno dotyczyło zwielokrotnienia skuteczności opieki medycznej bezpośrednio, w wyniku rozwoju nowych metod diagnostyki medycznej i terapii. Natomiast oczekiwanie drugie zakładało, że dojdzie do ogromnego uproszczenia i obniżki kosztów sekwencjonowania, co uczyni tę metodę wręcz rutynową. Spodziewano się, że tak znaczne obniżenie kosztów spowoduje, że przysłowiowy Kowalski w USA będzie mógł w 2010 r. zamówić usługę sekwencjonowania swojego genomu za *ca.* 2000 \$ i uzyskać komentarz wskazujący na zakodowane w genach zagrożenia – między innymi większej od średniej w populacji podatności na choroby, wrażliwości na czynniki środowiska czy skłonności do „niezdrowych” zachowań. W ten sposób ów Kowalski bę-

dzie mógł w dużym stopniu sam chronić się przed skutkami wielu zagrożeń, prowadząc odpowiedni tryb życia. Komentarz do genomu indywidualnego mógłby zawierać nie tylko zalecenia profilaktyczne, lecz również informacje dla lekarzy np. o domniemanym tempie rozkładu podawanych leków, co może mieć ogromne znaczenie w leczeniu wielu poważnych schorzeń. Oczekiwanie to nie ziściło się w odniesieniu do czasu powszechnej dostępności sekwencjonowania, albowiem za 2000 \$ usługa ta będzie prawdopodobnie dostępna dopiero w 2015 r. Drugi z przykładów, pokazujący realną wartość informacji sekwencji genomu, dotyczy zagwarantowania wyłączności dostępu do sekwencji genomowej jednej z roślin uprawnych. Otóż jedna z firm hodowlano-nasiennych zapłaciła w 2009 r. kwotę 200 000 EUR za pół roku wyłącznego dostępu do takiej sekwencji u rośliny uprawnej znajdującej się na *ca.* 40 miejscu listy światowej gatunków o największym znaczeniu gospodarczym. Nie sposób wchodzić tutaj w detale skąd taka wartość, jednak w pewnym uproszczeniu jest to pochodna możliwości dokonania odpowiednich zabezpieczeń patentowych.

POZNANIE GENOMU OGÓRKA

Inicjatywa sekwencjonowania genomu ogórka powstała w Katedrze Genetyki, Hodowli i Biotechnologii Roślin SGGW w Warszawie, gdzie od lat prowadzono zarówno

badania podstawowe, jak i hodowlę tego gatunku. Już ponad 50 lat temu prof. Bogusław Kubicki opracował model genetycznej regulacji płci ogórka i wykorzystał tę wiedzę do

zaproprowania metodyki hodowli odmian heterozyjnych. Prace te kontynuowano po śmierci profesora w 1985 r., starając się wyjaśnić podłoże molekularne genów uczestniczących w kształtowaniu płci ogórka oraz podejmując tematy dotyczące poznania sposobu działania innych genów odpowiadających za kształtowanie cech ważnych z punktu widzenia hodowlanego. Prace molekularne są bardzo czasochłonne i kosztochłonne jeżeli nie zna się kompletnej sekwencji genów. Dlatego uznaliśmy, że dostęp do sekwencji genomowej ogórka mógłby radykalnie zwiększyć efektywność badań. Niestety pierwsze starania o realizację tego planu przez uczestnictwo w międzynarodowym konsorcjum (The Cucumber Genome Initiative; <http://www.icugi.org/>) kierowanym przez zespół z Chin nie powiodły się, ze względu na zbyt wysokie wymagania finansowe postawione poten-

cjalnym partnerom. Dokładna analiza finansowa i techniczno-organizacyjna pokazała, że można podjąć się tego zadania samodzielnie przy zaangażowaniu do części analiz bioinformatycznych zespołu z USA. W konsekwencji powstało Polskie Konsorcjum Sekwencjonowania Genomu Jądrowego Ogórka (ang. Polish Consortium of Cucumber Genome Sequencing; <http://csgenome.sggw.pl/>), którego uczestnicy doprowadzili przedsięwzięcie do końca. Do sekwencjonowania użyto jednopiennej linii wsobnej pochodzącej ze starej, już nie uprawianej odmiany polowej Borszczagowski, dla której istnieje duża kolekcja mutantów i prowadzone są wielokierunkowe badania molekularne. Na stronie internetowej konsorcjum zamieszczone są zbiory danych użytecznych z punktu widzenia możliwości korzystania z sekwencji.

CO WYNIKA Z ZSEKWENCJONOWANIA GENOMU OGÓRKA ?

We współczesnych badaniach genetycznych bardzo ważnym elementem są bazy danych (w większości publiczne) zawierające informacje o budowie, funkcji i wielu innych właściwościach genów i ich produktów. Bazy danych zamieszczają informacje odpowiednio przygotowane, a data zgłoszenia i autorstwo umożliwiają ocenę wkładu poszczególnych osób i grup badawczych do uzyskania zgromadzonych informacji. W przypadku genomów najbardziej znaną jest baza NCBI, w której zamieszczono sekwencję genomu ogórka 21 września 2009 r. i było to o miesiąc wcześniej aniżeli zgłoszenie przez konsorcjum kierowane przez Chińczyków. Okazało się, że genom ogórka jest zbudowany z 367 000 000 par nukleotydów, zorganizowanych w 26 500 genów, które stanowią niewiele ponad 37% genomu. Tak więc ogórek nie wyłamuje się z ogólnej prawidłowości obserwowanej wśród roślin wyższych, że większość genomu stanowi „śmieciowe” DNA, chociaż pod tym względem ogórek należy do gatunków średnio „zaśmieconych”. Wspomniany ziemniak ma ponad 90% takiej DNA o nieokreślonej funkcji.

Prawdziwą kopalnią informacji okazało się jednak dopiero porównanie odmiany Chinese Long z odmianą Borszczagowski (WOYCICKI i współaut. 2011), które wykonano dla sześciu grup genów odpowiedzialnych za podstawowe procesy fizjologiczne, związane ze zdolnością do adaptacji. Porówna-

nie to pokazało, że dla pewnych procesów fizjologicznych więcej genów ma odmiana Borszczagowski, natomiast dla innych - odmiana Chinese Long. W odmianie polskiej jest więcej genów biorących udział w takich procesach jak fotosynteza, metabolizm cukrów, oddychanie, regulacja ekspresji genów, przyswajanie jonów amonowych i degradacja chlorofilu, natomiast mniej genów odpowiadających za odporność na stres oksydacyjny i odporność na wysoką temperaturę. Różnice te są powiązane z odmiennością środowisk dla jakich odmiany te hodowano. Klimat umiarkowany Europy północnej charakteryzuje: chłód, niska intensywność promieniowania słonecznego, stała, wyższa, niż na terenie Azji południowo-wschodniej, emisja CO₂ (od początków ery industrialnej do lat 80. XX w.) i związana z tym, zmniejszona zdolność przyswajania jonów azotanowych. Z kolei dla strefy klimatu subtropikalnego południowo-wschodnich Chin charakterystyczne są: wysoka sezonowa intensywność promieniowania słonecznego w tym UV-B oraz wysoka temperatura, co wymaga zwiększonej odporności na stres oksydacyjny i termiczny. Wyniki te odnoszą się do zjawiska kształtowania się tła genetycznego w trakcie hodowli wskazując, że jest to proces postępujący samoistnie, jakby „przy okazji” prowadzonej przez hodowcę selekcji na inne cechy. Istnienie tła genetycznego nie jest niczym nowym, jednak do tej pory nie byliśmy go w stanie

skonkretyzować pod względem składu genów. Z tych danych dowiadujemy się, że w każdej z odmian są skumulowane określone geny dla wspomnianych procesów, a decydujące o takim lub innym ich zestawie są warunki w jakich jest prowadzona hodowla. Można więc powiedzieć, że genom odpowiedzialnym za cechy charakterystyczne dla Chinese Long (owoc sałatkowy, długi, białe kolce, brak brodawek, duże nasilenie żeńskości) towarzyszy tło genetyczne o większej liczbie genów odpowiedzialnych za procesy stresu oksydacyjnego i odporności na wysoką temperaturę oraz mniejszej liczbie genów odpowiadających za fotosyntezę, metabolizm cukrów, oddychanie i degradację chlorofilu. Z kolei genom odpowiadającym za cechy charakterystyczne dla odmiany Borszczagowski (owoc do kiszenia, ciemnokolcowy, brodawkowy, słabe nasilenie żeńskości) towarzyszy większa liczba genów odpowiadających za fotosyntezę, metabolizm cukrów, oddychanie i degradację chlorofilu oraz mniejsza liczba genów odporności na stres oksydacyjny i odporności na wysoką temperaturę.

Inna część uzyskanych wyników po raz pierwszy ujawniła jeszcze jeden intrygujący mechanizm. Okazało się mianowicie, że dynamicznym zmianom w zależności od warunków środowiska ulegają także promotory ge-

nów, a ściślej rzecz biorąc, krótkie sekwencje nukleotydowe wewnątrz promotorów, odpowiedzialne za wiele subtelnych reakcji na warunki zewnętrzne. W ten sposób dochodzi do zmian przystosowawczych w sieci oddziaływań między określonymi genami, co z kolei umożliwia efektywną adaptację do danych warunków środowiskowych, bez utraty zdolności do funkcjonowania poszczególnych genów.

Dane powyższe dotyczą małego fragmentu genomu ogórka, wskazują jednak, że genom jest strukturą bardzo dynamiczną, wbrew temu co do niedawna sądziliśmy. Świadczy o tym także wnikliwa analiza sekwencji genomowych obu odmian (Borszczagowski i Chinese Long), która wykazała dużą zmienność w występowaniu różnej wielkości insercji, delecji, rearanzacji oraz inwersji i translokacji wewnątrz chromosomowych oraz między chromosomami. Jest to dość zaskakujące, albowiem ogórek uchodził dotychczas za gatunek o bardzo słabym tak zwanym polimorfizmie, czyli zmienności na poziomie molekularnym, szczególnie w odniesieniu do białek. Konsekwencją tak dynamicznych zmian w genomie mogą być ograniczenia w uzyskiwaniu stabilnych markerów molekularnych opartych na DNA.

OD SEKWENCJI GENOMU DO KATALOGU POLIMORFIZMÓW

Sekwencjonowanie genomów dało początek nowemu rozdziałowi w badaniach nad funkcjonowaniem organizmów, który nazwano erą genomiki. Jest to dział genetyki, który ustala właściwości organizmów nie na podstawie różnic w pojedynczych genach, lecz na podstawie różnorodności całych genomów. Ta wiedza przynosi ogrom nowych informacji, z których na razie umiemy wykorzystać niewielką tylko część. Z punktu widzenia hodowcy, wraz ze znajomością sekwencji genomu otwierają się nowe możliwości markerowania molekularnego. Korzyści polegają tutaj głównie na dużo większej skuteczności i szybkości markerowania oraz znacznym obniżeniu kosztów (PRZYBECKI 2010). W praktyce postęp będzie polegał na wykorzystaniu sekwencji genomów czy ich części (powielone fragmenty jądrowego DNA lub transkryptomy) z setek lub nawet tysięcy spokrewnionych genomów (z określonych populacji) wykrywając tysiące polimorfizmów. W tym przypadku będziemy mówić o

resekwencjonowaniu genomów czy ich fragmentów, gdzie w odróżnieniu od sekwencjonowania *de novo*, przypisanie konkretnej sekwencji do *locus* jest znacznie uproszczone. Należy podkreślić, że nie jest to bynajmniej wizja futurystyczna. Jako przykład mogą posłużyć zrealizowane projekty wielkich firm nasiennych Dupont i Pioneer Hi-Bred, w których zsekwencjonowano znaczne obszary genomu 553 linii kukurydzy, a w konsekwencji zidentyfikowano gen odpowiedzialny za zawartość kwasu oleinowego w ziarniakach (BELÓ i współaut. 2008). Postęp w metodyce sekwencjonowania genomów spowodował, że w najbliższych latach nastąpi wręcz zalew informacji o genomach odmian i linii hodowlanych wielu gatunków roślin uprawnych (na liście zapowiedzi znajduje się 140 gatunków). Analiza tych informacji wymaga zastosowania odpowiednich programów komputerowych. Jako że polimorfizm pojedynczych nukleotydów jest najprostszym wykrywanym polimorfizmem w trakcie rese-

Tabela 3. Wybrane pakiety oprogramowania komputerowego służące do analiz porównawczych sekwencji genomowych, a w szczególności detekcji polimorfizmów.

Nazwa	URL	Zastosowanie
EagleView	http://bioinformatics.bc.edu/marthlab/EagleView	Składanie, wizualizacja i sprawdzanie jakości złożonych fragmentów genomu i weryfikacja potencjalnych polimorfizmów
Atlas-SNP2	http://www.hgsc.bcm.tmc.edu/cascade-tech-software_atlas_snp-ti.hgsc	Narzędzie do wykrywania SNP i indeli w toku resekwenjonowania na platformach Roche 454 i Illumina
SeqMap	http://www-personal.umich.edu/~jianghui/seqmap/index.html	Mapowanie krótkich sekwencji na genomie referencyjnym i detekcja substytucji i indeli.
Alpheus™	http://alpheus.ncgr.org/	Zintegrowana, sieciowa platforma do automatyzacji analiz, wizualizacji i interpretacji danych z projektów sekwencjonowania nowej generacji
CLCbio Genomics Workbench	http://www.clcbio.com/	Narzędzie do składania sekwencji <i>de novo</i> i do analiz resekwenjonowania zarówno w technologii Sangera jak i nowej generacji, zawierające możliwość detekcji i analiz SNP.

kwencjonowania narzędzia bioinformatyczne są głównie nakierowane na ich detekcję i analizę (VARSHNEY i współaut. 2009). Istnieje sporo programów do wykorzystania w takich analizach, a Tabela 3 pokazuje kilka przykładów.

Hodowla roślin w Polsce powinna umieć skorzystać z tych niezwykle cennych zasobów informacji. Więcej o znaczeniu sekwencjonowania można znaleźć w numerze 4 kwartalnika „Biotechnologia” z 2010 r., poświęconym temu zagadnieniu.

SEKWENCJONOWANIE GENOMÓW STAJE SIĘ NOWYM SKŁADNIKIEM POSTĘPU W HODOWLI ROŚLIN

Streszczenie

Sekwencjonowanie genomów jest odczytaniem zapisu genetycznego, jakim dany organizm dysponuje, wyrażonego w odpowiednim uporządkowaniu nukleotydów. W efekcie uzyskujemy informację rzeczywistą o strukturze i położeniu genów i innych składników genomu, co daje podstawy do szczegółowego prognozowania cech organizmu i pozwala

wyjaśniać mechanizmy wielu reakcji. Z tych powodów efekty sekwencjonowania mają wielorakie implikacje, między innymi oddziałują na nauki rolnicze w szczególności na postęp hodowlany. W artykule przedstawione zostały uwarunkowania ogólne sekwencjonowania i przybliżono znaczenie niektórych danych jakie uzyskano u ogórka.

GENOME SEQUENCING MAKES A SIGNIFICANT PROGRESS IN PLANT BREEDING

Summary

Genome sequencing is the reading of the genetic record of a given organism, expressed in the correct sequence of nucleotides. In effect, information about the real structure, location of genes and other elements of a genome is obtained. It gives a solid base for detailed prediction of the organism's traits and allows explain the mechanisms of many meta-

bolic reaction. This is why the sequencing has many implications, for example on agricultural science, especially in breeding programs. This article presents general sequencing strategies and gives a closer look at the meaning of experimental data obtained by cucumber.

LITERATURA

- BELÓ A., ZHENG P., LUCK S., SHEN B., MEYER D. J., LI B., TINGEY S., RAFALSKI A., 2008. *Whole genome scan detects an allelic variant of fad2 associated with increased oleic acid levels in maize*. Mol. Genet. Genomics 279, 1-10.
- GROMADKA R., GAWOR J., SZCZĘSNY P., ZAGÓRSKI W., 2011. *Kolejny wielki genom poznany przy udziale polskich laboratoriów. Genom ziemniaka zsekwencjonowany*. Kosmos 60, 491-497.
- PRZYBECKI Z., PAWEŁKOWICZ M., WÓYCICKI R., 2010. *Sekwencjonowanie genomów i rozwój biotechnologii*. Biotechnologia 4, 9-23.
- THE POTATO GENOME SEQUENCING CONSORTIUM, 2011. *Genome sequence and analysis of the tuber crop potato*. Nature 475, 189-195.
- VARSHNEY R. K., NAYAK S. N., MAY G. D., JACKSON S. A., 2009. *Next-generation sequencing technologies and their implications for crop genetics and breeding*. Trends Biotechnol. 27, 522-530.
- WOYCICKI R., WITKOWICZ J., GAWROŃSKI P., DĄBROWSKA J., LOMSADZE A. i współaut., 2011. *The genome sequence of the North-European cucumber (Cucumis sativus L.) unravels evolutionary adaptation mechanisms in plants*. PLOS ONE 6, e22728.

DODATKOWE ŹRÓDŁA

- <http://www.ncbi.nlm.nih.gov/genomes/leuks.cgi?taxgroup=11:|12:Land%20Plants&p3=12:Land%20Plants>
- <http://www.ncbi.nlm.nih.gov/genomes/PLANTS/PlantList.html>
- <http://www.ncbi.nlm.nih.gov/genome/browse/>
- http://www.genomesonline.org/cgi-bin/index.cgi?page_requested=Complete+Genome+Projects&subset_requested=EUKARYAL
- <http://www.genomesonline.org/cgi-bin/index.cgi>