

NOWE OSIĄGNIĘCIA GENETYKI I FIZJOLOGII ROŚLIN W TWORZENIU I WYKORZYSTANIU POSTĘPU BIOLOGICZNEGO

Różnorodność genetyczna z perspektywy poznanych sekwencji genomów roślinnych

Jan Sadowski^{1,2}, Małgorzata Kaczmarek², Piotr A. Ziółkowski², Danuta Babula², Małgorzata Jakubowicz¹, Agnieszka Ludwików¹, Agata Stańko¹

¹*Zakład Biotechnologii, Uniwersytet im Adama Mickiewicza,
ul. Międzychodzka 5, 60-371 Poznań*

²*Instytut Genetyki Roślin, Polska Akademia Nauk,
ul. Strzeszyńska 34, 60-479 Poznań*

Słowa kluczowe: genom, rodziny genowe, duplikacje, *Arabidopsis thaliana*,
Brassica

Wstęp

Opublikowana w grudniu 2000 roku sekwencja nukleotydowa genomu rzodkiewnika *Arabidopsis thaliana* (L.) HEYNH [29], a w kwietniu 2002 roku – ryżu *Oryza sativa* L. [14, 34] daje specjalistom wielu dziedzin biologii nowe możliwości poznawania mechanizmów funkcjonowania organizmów roślinnych. Genetykom umożliwi wykorzystanie pełnego zestawu genów w analizie procesów komórkowych, co będzie pomocne w opracowywaniu i doskonaleniu materiałów hodowlanych, a także przybliży wyjaśnienie mechanizmów formowania nowych homologów genowych. W tym ostatnim wypadku genom *A. thaliana* jest niezwykle interesującym obiektem, bowiem stwierdzono na podstawie analizy jego sekwencji powstawanie i różnicowanie

się wielu tzw. rodzin genowych. Analiza i porównanie składu genów pomiędzy różnymi gatunkami pozwoli na określenie, które szlaki metaboliczne są wspólne, a które gatunkowo specyficzne i jakim modyfikacjom były one poddane. Zastosowanie nowych metod biologii molekularnej daje możliwość poznania ekspresji genów całego genomu w wybranych warunkach doświadczenia. Dostępność sekwencji genomowych to źródło markerów DNA zarówno do badań z dziedziny genetyki molekularnej (w tym porównawczej), jak i populacyjnej.

Przeprowadzona w okresie 2000–2003 dalsza analiza sekwencji nukleotydowej genomu *A. thaliana* rzuciła nowe światło na podstawy budowy, organizacji i funkcjonowania genomów licznych gatunków uprawnych, które mają podobne do tej rośliny modelowej pochodzenie poliploidalne (nazywane są one oddalonymi poliploidami lub paleopoliploidami). Cechą charakterystyczną genomów paleopoliploidów jest obecność zduplikowanych segmentów chromosomowych w skali całego genomu. Ten wysoki stopień zduplikowania wyraża się wzrostem liczby homologów tego samego genu. Homologi genowe są z reguły rozproszone w genomie, ulegając mniejszej lub większej dywergencji strukturalnej i funkcjonalnej. Intensywnie prowadzona analiza genomów gatunków spokrewnionych z *A. thaliana* pozwala wstępnie wyjaśnić, w jaki sposób dochodzi do różnicowania, ekspansji bądź redukcji rodzin genowych, które to zjawiska miały miejsce podczas ewolucji genomu modelowego gatunku [6, 29, 35].

Krótką charakterystyka strukturalna genomu *A. thaliana*

Genom *A. thaliana* obejmuje pięć par chromosomów pokrywających 125 Mbp (milionów par zasad). Dotychczas łącznie zsekwencjonowano 115,4 Mbp z pominięciem obszarów centromerowych i regionów rDNA, co odpowiada 92% długości całego genomu *A. thaliana* [29]. Liczba genów tworzących 11 000 rodzin genowych została wstępnie oszacowana na ponad 25 000, choć ostatnio opublikowane dane wskazują na istnienie ponad 27 000 genów w genomie *A. thaliana* (<http://www.arabidopsis.org/>). W obrębie puli zidentyfikowanych genów, 17% ułożonych jest w tandemowo powtórzonych segmentach. Sekwencje kodujące *A. thaliana* obejmują prawie 85% genomu, pozostałe obszary zajmują sekwencje powtarzające [5]. Szczegółowa analiza wykazała, że wartości takie, jak średnia gęstość genów i ich średnia długość są raczej wyrównane w całym genomie i wynoszą odpowiednio 4,5 kbp/gen oraz 2 kbp. Wstępna identyfikacja funkcji kodowanych przez geny produktów została przeprowadzona poprzez porównanie ich sekwencji nukleotydowej z sekwencjami genomów innych eukariontów zdeponowanymi w banku danych. W efekcie tej analizy 69% genów zostało zaszeregowanych do 12 kategorii funkcjonalnych, w tym dla 9% funkcja została potwierdzona eksperymentalnie. Zidentyfikowano 24 duże segmenty zduplikowane w genomie *A. thaliana*, obejmujące 65,6 Mbp, co stanowi 58% genomu [6,

29]. Udział homologicznych genów w regionach zduplikowanych waha się od 20% do 47%. Obie wartości wskazują, że organizacja genomu *A. thaliana* różni się od oczekiwanej prostej organizacji charakterystycznej dla diploida. Ponowna, bardziej wnikliwa analiza sekwencji całego genomu *A. thaliana* pozwoliła na zidentyfikowanie segmentów zduplikowanych, pochodzących z dwóch odległych w czasie zdarzeń duplikacyjnych, dotyczących najprawdopodobniej całego genomu [7]. Doniesienia te wskazują na możliwe trudności w interpretacji mechanizmów ewolucyjnych, które mogły doprowadzić genom *A. thaliana* do obecnej złożonej organizacji, która powinna być uwzględniana w analizie funkcjonalnej genów.

A. thaliana* – modelem dla gatunków uprawnych z rodzaju *Brassica

Ponieważ *A. thaliana* należy do rodziny *Brassicaceae*, dostępność baz danych i opracowanych „narzędzi” biotechnologicznych wraz z rozwiniętymi metodami biologii molekularnej umożliwiającymi masową analizę struktury i funkcji genów stawia innych przedstawicieli tej rodziny w pozycji beneficjentów, dzięki możliwej analizie porównawczej ich genomów. Możliwe stanie się wyjaśnienie genetycznych mechanizmów leżących u podstaw zdolności *Brassicaceae* do adaptacji do prawie wszystkich zmian środowiskowych, jak również dróg ewolucji specyficznych genów i szlaków biochemicznych.

Rodzina *Brassicaceae* obejmuje 360 rodzajów zorganizowanych w obrębie 13 plemion [2]. Gatunki należące do tej rodziny cechują się znacznym zróżnicowaniem liczby genomowej – od $n = 5$ u *A. thaliana* do $n = 19$ u *Brassica napus*, i zawartością DNA – od 38 Mpz u *Cardamine amara* do 1235 Mpz u *B. napus* [3]. Najintensywniej badanym obecnie taksonem, z wyjątkiem rośliny modelowej – *A. thaliana* – jest rodzaj *Brassica*, obejmujący około 300 gatunków diploidalnych i amfidiploidalnych, zarówno dzikich, jak i uprawnych. Najbardziej urozmaiconym pod względem morfotypów jest gatunek *B. oleracea* L., w obrębie którego można wyróżnić sześć odmian uprawnych i jedną dziką (*Brassica silvestris* L. – kapusta dzika). Poza sekwencją genomu istnieje obecnie liczna kolekcja roślin transgenicznych i mutantów *A. thaliana* dla większości etapów wzrostu i różnicowania oraz rozwoju w odmiennych warunkach, stanowiących źródło informacji dla tego typu prac nad gatunkami *Brassica*.

Analiza porównawcza diploidalnych gatunków z rodzaju *Brassica* i gatunku spokrewnionego – *A. thaliana* – ujawnia ich poliploidalne pochodzenie. Wyniki tych badań sugerują nie tylko pochodzenie diploidalnych *Brassica* od heksaploidalnego przodka, ale również wskazują na ewolucję ich genomów na drodze fuzji chromosomów i częstych rearanżacji [4, 19]. Genetyczne i ewolucyjne konsekwencje wielokrotnej poliploidyacji i następującej po nich restrukturyzacji genomu (zwanej diploidyacją) są widoczne w obecnej strukturze genomu *A. thaliana*.

Bazy danych sekwencji genów (EST)

Dostępność rosnących w szybkim tempie baz danych sekwencji DNA dla różnych gatunków modelowych pozwala na identyfikację anonimowych sekwencji wyizolowanych z genomów mało poznanych poprzez podobieństwo sekwencji. W ten sposób możliwe jest przypisanie domniemanej funkcji znacznej liczbie sekwencji zawartych w bibliotekach klonów cDNA. Obserwuje się także wzbogacenie zasobów baz danych o zestawy sekwencji EST (ang. expressed sequence tag) pochodzących z genomów roślinnych, stanowiące swoiste etykiety genów. Dotyczy to zwłaszcza gatunków o dużym znaczeniu rolniczym (<http://www.ncbi.nlm.nih.gov>) [8], takich jak jęczmień *Hordeum vulgare* – 346 108 ESTów czy ryż – 201 747 ESTów; w wypadku *A. thaliana* – 178 538 ESTów (dane z 6 czerwca 2003). W bazie danych (National Center for Biotechnology Information) zdeponowanych jest 8406 EST dla *Brassica* (stan z 30 maja 2003). Wysoki stopień podobieństwa na poziomie sekwencji kodujących pomiędzy *Brassica* i *A. thaliana*, szacowany na 87%, stwarza możliwość analizy zmian ekspresji dużej liczby genów *Brassica* przy użyciu mikromacierzy DNA (ang. DNA microarray) utworzonych na podstawie sekwencji genów *A. thaliana*. W tym kontekście ważna jest perspektywa określenia funkcji wszystkich genów w genomie *A. thaliana*, co zaplanowano do 2010 roku [28].

Konstrukcja map chromosomowych

Bazy danych sekwencji DNA wykorzystuje się pod różnym kątem: klony EST użyto jako markerów do budowy np. nasyconych map genetycznych ryżu i kukurydzy [11], mapy fizycznej ryżu [18], mapowania porównawczego *B. oleracea* i *A. thaliana* [4, 20] oraz *Solanum tuberosum* i *A. thaliana* [13]. Ponadto, dane pochodzące z sekwencjonowania są wykorzystywane w analizie rodzin genowych [10] i mogą stanowić podstawę obiecującego systemu markerowego SNP (polimorfizm pojedynczych nukleotydów) [9]. Poza zastosowaniem w mapowaniu genetycznym i fizycznym, sekwencje EST są także głównym źródłem informacji w analizie ekspresji genów z udziałem mikromacierzy DNA, jak to zostało pokazane na przykładzie *A. thaliana* [26]. Zgromadzone w bazach danych sekwencje EST dla wielu organizmów cechują się powtarzalnością odpowiadającą poziomowi ekspresji korespondujących do nich genów. W celu skatalogowania ich w zestawy reprezentujące każdy gen danego organizmu opracowano projekt Unigen [21].

Porównawcze mapowanie chromosomów

W ramach projektów genomowych identyfikowane są segmenty chromosomowe o konserwatywnej budowie w obrębie wielu taksonomicznie pokrewnych grup organizmów. Pozwala to na konstruowanie wspólnej mapy genetycznej dla wielu takich grup, np. traw [12]. Zastosowanie takiej mapy w badaniach genetycznych i biotechnologicznych powinno przyspieszyć postęp w hodowli. Przykładowo, szczegółowe informacje dotyczące struktury genomu uzyskane z analizy map i sekwencji rośliny modelowej, takiej jak *A. thaliana*, mogłyby wspomagać identyfikację ważnych z użytkowego punktu widzenia genów u gatunków uprawnych, bądź opartą na mapowaniu introgresję genów z dzikich gatunków pokrewnych, selekcję z wykorzystaniem markerów zlokalizowanych na mapach genetycznych i klonowania pozycyjnego. Znając pozycję interesującego nas genu w stosunku do genów markerowych u gatunku modelowego, można obliczyć prawdopodobieństwo tego sprzężenia u gatunku docelowego. Natomiast kompletne informacje o rearanżacjach chromosomowych różniących dwa gatunki są pomocne w wykrywaniu konserwatywnych bloków genów. Poza tym oszacowanie zakresu i typu rearanżacji chromosomowych może być pomocne w rekonstrukcji historii ewolucyjnej i testowaniu hipotez ewolucyjnych dotyczących tempa ewolucji chromosomowej [27].

Jednym ze sposobów wykorzystania informacji pochodzącej z analizy genomu *A. thaliana* jest identyfikacja genów interesujących z rolniczego punktu widzenia. Mapowanie porównawcze może przyspieszyć identyfikację prawdopodobnych ortologów *A. thaliana* w genomach *Brassica* kodujących cechy ilościowe. W celu ustalenia korelacji genów kodujących cechy ilościowe u *Brassica* z genami *A. thaliana*, wstępnie scharakteryzowanymi za pomocą technik genetyki molekularnej, można wykorzystać wiele strategii opierających się na mapowaniu genetycznym. Po pierwsze, kandydujące do tego geny z *A. thaliana* mogą zostać użyte jako markery do mapowania w populacji mapującej *Brassica* dla sprawdzenia, czy segregujące fragmenty restrykcyjne DNA (analiza typu RFLP) są skorelowane z interesującym nas QTL. Po drugie, markery *Brassica* leżące w sąsiedztwie QTL można wykorzystać do mapowania genetycznego u *A. thaliana* w celu zidentyfikowania korespondującego z nimi regionu genomowego. Sekwencje markerów molekularnych flankujących interesujący nas locus w genomie *Brassica* można określić i porównać z sekwencją *A. thaliana*, co – w związku z wysokim stopniem zachowawczości sekwencji genowej pomiędzy *A. thaliana* i *Brassica* – pozwoli z dużym prawdopodobieństwem na zlokalizowanie ich na chromosomach *A. thaliana*. W wypadku zidentyfikowania w genomie *A. thaliana* regionu zbiegającego się z QTL w gatunku uprawnym dostępność scharakteryzowanej sekwencji genomu *A. thaliana* stwarza możliwość szczegółowej analizy i identyfikacji locus kodującego np. enzym uczestniczący w interesującym nas szlaku metabolicznym.

Analiza cech użytkowych z pomocą modelu *Brassica/A. thaliana*

Naturalnie występujące geny kontrolujące niektóre cechy użytkowe w obrębie gatunków rodzaju *Brassica* mogą okazać się bardziej przydatne w badaniach nad przebiegiem rozwoju roślin, szczególnie przy braku odpowiedniego zróżnicowania genetycznego w obrębie kolekcyjnych linii *A. thaliana*. Wiele fenotypowych cech znajduje się pod bardziej złożoną genetyczną kontrolą u *Brassica*, niż podobne cechy u *A. thaliana*. Przykładem tej różnicy jest wykształcenie kwiatostanu podobnego do kalafiora, otrzymanego u *A. thaliana* poprzez mutacje w dwóch genach. Cecha ta u kalafiora kodowana jest przez większą liczbę genów [16]. Wysoki poziom duplikacji u *Brassica*, związany z dwoma dodatkowymi cyklami poliploidyzacyjnymi, które miały miejsce po rozejściu się dróg ewolucyjnych *Brassica* i *Arabidopsis*, umożliwia występowanie nowych mutacji w loci, które w genomie *A. thaliana* podlegają ściślejszej kontroli. Wstępne badania wskazują, że wiele genów u diploidalnych gatunków *Brassica* występuje w wyższej liczbie kopii niż u *A. thaliana*. Badanie ekspresji i funkcji wielu kopii genów u *Brassica* może stwarzać nowe możliwości w interpretacji roli intensywnego zduplikowania genomu w szerokiej zmienności fenotypowej, obserwowanej u gatunków tego rodzaju [24]. Gatunki z rodzaju *Brassica* pozwolą śledzić szybkie zmiany w genomie, skorelowane z poliploidalnością.

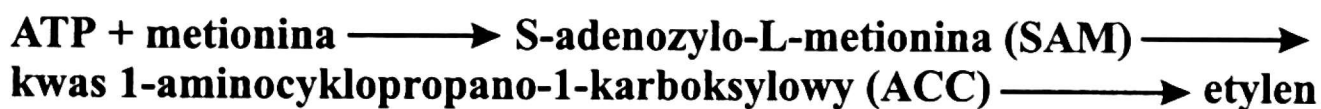
Różnorodność genetyczna u podstaw rozwoju i odpowiedzi roślin na stresy: rola genów uczestniczących w biosyntezie etylenu

Jak już wspomniano powyżej, dostęp do całkowitej sekwencji genomu umożliwia identyfikację wszystkich kopii danego genu. Niektóre z rodzin genowych były przedmiotem intensywnych badań ze względu na ich rolę w przebiegu ważnych procesów życiowych czy w kształtowaniu się cech użytkowych. Do nich należą syntazy i oksydazy ACC uczestniczące w biosyntezie etylenu, fitohormonu wpływającego na formowanie się wielu cech użytkowych u roślin. Jednak dopiero obecnie, dzięki identyfikacji w genomie *A. thaliana* wszystkich kopii kodujących je genów, można poznawać zróżnicowane funkcje, jakie pełnią w procesach rozwojowych czy obronnych u roślin. Przebieg wielu często niezależnie przebiegających procesów indukowanych przez etylen zapewniają mechanizmy precyzyjnej regulacji zarówno syntezy, jak i percepcji tego hormonu. Oddziałuje on na liczne procesy fizjologiczne, m.in.: zjawisko geotropizmu, inicjację kiełkowania nasion, rozwój pąków i bulw ziemniaka, dojrzewanie owoców, tworzenie kwiatów (u ananasa), a także ruchy płatków okwiatu. Etylen ma także wpływ na wzrost i rozwój pędów, korzeni (przybyszowych i włośni-

ków), stymuluje elongację międzywęzła u ryżu i uczestniczy w rozwoju organów w kulturach *in vitro*. Ponadto etylen jest cząsteczką sygnałową w procesie starzenia i odpowiedzi rośliny na czynniki stresowe, takie jak: zranienie, zasolenie, ozon, atak patogenów czy promieniowanie UV.

Geny związane z biosyntezą etylenu

Biosynteza etylenu podlega autokatalitycznej regulacji – negatywnej w warunkach wegetatywnego wzrostu, lub regulacji pozytywnej podczas procesów dojrzewania owoców i starzenia. Kluczowymi enzymami kontrolującymi dwa ostatnie etapy biosyntezy etylenu w roślinach wyższych jest syntaza ACC, katalizująca przekształcenie SAM w ACC, oraz oksydaza ACC, katalizująca przekształcanie ACC w etylen (rys. 1).



Rysunek 1. Schemat biosyntezy etylenu

Enzymy te występują w postaci spektrum izoenzymów. Kodujące je geny stanowią rodziny wielogenowe, a ich ekspresja u różnych gatunków regulowana jest przez wiele czynników hormonalnych i środowiskowych. Na przykład auksyny indukują ekspresję At-ACS4, At-ACS6 u *A. thaliana*, Le-ACS5 i Le-ACS3 u pomidora, St-ACS1A i St-ACS1b u ziemniaka. Cytokiny indukują At-ACS5 i At-ACS9 u *A. thaliana*. ABA indukuje ekspresję At-ACO u *A. thaliana*. Natomiast Le-ACS2-6 i Le-ACO1 u pomidora, St-ACS1A-B u ziemniaka, At-ACS2, At-ACS6, BO-ACS2, At-ACO1 u *A. thaliana* są regulowane przez zranienie. W wypadku genów At-ACS6 u *A. thaliana*, Le-ACS6 i Le-ACS1A u pomidora indukcja ekspresji powodowana jest przez dotyk. Jony miedzi, litu i żelaza indukują ekspresję Le-ACS3 u pomidora, St-ACS4 i St-ACS5 u ziemniaka, At-ACS2, At-ACS4, At-ACS6 i At-ACO1 u *A. thaliana*. Ozon stymuluje transkrypcję Le-ACS1b, Le-ACS2, Le-ACS6 u pomidora, At-ACS6 u *A. thaliana*, Nt-ACS6 u tytoniu, St-ACS4, St-ACS5 u ziemniaka. Podtapianie stymuluje transkrypcję Le-ACS2, Le-ACS3, Le-ACS7 u pomidora. ABA podnosi ekspresję Bo-ACO u kapust [25].

Podobnie w wypadku oksydazy ACC, jej różne izoformy są indukowane przez zranienie, jony srebra, infekcje wirusem, kwas salicylowy, kwas jasmonowy, etylen i auksyny. Geny kodujące ww. enzymy zostały scharakteryzowane m.in. u pomidora [1], tytoniu [31], *A. thaliana* [30] i brokuła [33].

Wykorzystanie genów biosyntezy etylenu w biotechnologii

Etylen przyspiesza dojrzewanie owoców wielu roślin klimakterycznych (pomidor, jabłko, banan, śliwka, melon, brzoskwinie) i powoduje starzenie kwiatów (etylen wpływa na opadanie płatków okwiatu). Po zbiorze warzyw i owoców wzrost produkcji etylenu jest przyczyną niekorzystnych zmian, takich jak żółknięcie brokułu czy czernienie kalafiora. W doskonaleniu roślin ozdobnych położono nacisk na poprawę trwałości kwiatów (storczyki, geranium, pelargonie, petunie, goździki). Dla osiągnięcia tego celu zastosowano także kilka dostępnych strategii manipulacji genetycznych, które obniżają lub znoszą produkcję etylenu, prowadząc do opóźnienia starzenia kwiatów i warzyw lub przedłużenia dojrzewania owoców, a nawet do całkowitego zahamowania dojrzewania owoców. Obniżenie produkcji etylenu można osiągnąć poprzez ingerencję w biosyntezę, percepcję etylenu lub przez obniżenie poziomu metabolitów biorących udział w biosyntezie etylenu. Pierwszym celem dla biotechnologii roślin był pomidor. W pracach tych wprowadzano do roślin antysensowe RNA syntaz ACC jak i oksydaz ACC. Hamilton i in. [15] zredukowali produkcję etylenu u pomidora przez wprowadzenie antysensowego konstruktów oksydazy ACC. Oeller i in. [23] obniżyli poziom etylenu u transgenicznego pomidora z nadekspresją syntazy ACC w orientacji antysensowej. W 1991 roku Klee i in. [17] do tego celu używali genu bakteryjnej deaminazy ACC i transformacji za pomocą *Agrobacterium tumefaciens*. Stwierdzono, że Etr1 z *A. thaliana* ulega ekspresji u petunii, przedłużając trwałość kwiatów; u pomidora natomiast opóźnia dojrzewanie owoców [32]. Antysensowe konstrukty do wyciszenia genów oksydazy ACC zostały przygotowane także dla papai [22].

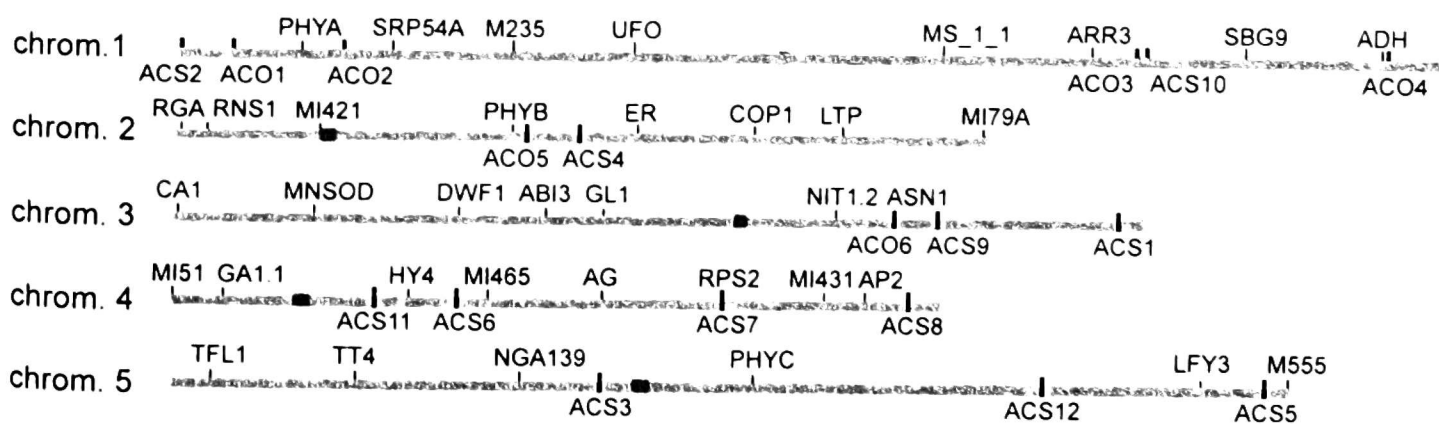
Wyjątkowe możliwości badania modelu ekspresji enzymów biosyntezy etylenu oraz wzajemnych zależności pojawiły się w momencie poznania całego genomu *A. thaliana*. Dotychczas w genomie *A. thaliana* zidentyfikowano geny 12 izoenzymów syntaz ACC, spośród których izoenzym AT-ACS1 reprezentuje najprawdopodobniej enzym nieaktywny katalitycznie, AT-ACS3 jest pseudogenem, natomiast model ekspresji pozostałych izoenzymów jest intensywnie badany. W rodzinie *Brassicaceae* dotychczas zidentyfikowano geny kilku izoenzymów z *Sinapis arvensis*, *Brassica oleracea* i *B. juncea*. Geny kodujące syntazy ACC z reguły zbudowane są z czterech eksonów i trzech intronów występujących zawsze w stałych pozycjach. Ponadto wśród genów syntaz ACC występują również takie, które utraciły intron pierwszy lub intron drugi.

Homologia aminokwasowa poszczególnych izoenzymów u *A. thaliana* może wahać się od 49% do 99%. Z tego względu można wśród nich wydzielić kilka klas grupujących te enzymy zgodnie z ich wzajemnym podobieństwem: 1) AT-ACS4, AT-ACS5, AT-ACS8, AT-ACS9 oraz AT-ACS11; 2) AT-ACS1, AT-ACS2 oraz AT-ACS3; 3) AT-ACS10 i AT-ACS12; 4) AT-ACS7; 5) AT-ACS6.

Można oczekiwać, że możliwość modyfikacji genetycznej aktywności poszczególnych izoenzymów wpłynie w zdecydowany sposób na cechy użytkowe, dlatego też syntazy ACC są przedmiotem intensywnych badań. Nasze zainteresowania są ukierunkowane m. in. na badanie organizacji genomowej jak i ekspresji syntaz ACC w roślinach kapustnych. Do tego celu wykorzystujemy dane pochodzące z programu genomowego *A. thaliana*.

Organizacja genów dla syntaz i oksydaz ACC w genomie *A. thaliana*

W genomie *A. thaliana* zidentyfikowano 12 genów kodujących syntazy ACC (rys. 2). Cztery z nich zlokalizowane są na chromosomie 4, natomiast trzy na chromosomie 5. Na chromosomach 1 i 3 znajdują się po dwa geny; jeden gen zlokalizowany jest na chromosomie 2. Organizacja genów oksydaz ACC jest zasadniczo podobna do organizacji syntaz ACC. Również w tym wypadku nie można wyróżnić żadnego miejsca w genomie, w którym akumulowałyby się te geny. Łącznie zidentyfikowano sześć kopii genów kodujących oksydazy ACC (*ACO1-ACO6**). Cztery z nich zlokalizowane są na chromosomie 1, w znacznych odstępach od siebie; natomiast po jednej kopii występuje na chromosomie 2 i 3 (rys. 2).



Rysunek 2. Organizacja genów syntaz ACC i oksydaz ACC w genomie *A. thaliana*. Poziome linie reprezentują poszczególne chromosomy. Ciemne prostokąty w obrębie chromosomów oznaczają położenie centromerów. Pod liniami reprezentującymi pięć chromosomów *A. thaliana* zaznaczono położenie genów syntaz ACC (*ACS1-ACS12*) i oksydaz ACC (*ACO1-ACO6*), natomiast nad liniami zaznaczono położenie wybranych markerów genetycznych stosowanych w programie sekwencjonowania

Chociaż w genomie *A. thaliana* geny syntaz ACC i oksydaz ACC są rozproszone, to jednak informacje dotyczące ich występowania w zduplikowanych blokach chromosomowych wydają się bardzo interesujące: okazuje się na przykład, że dziewięć z dwunastu genów kodujących syntazy ACC występuje w zduplikowanych segmentach chromosomowych. Oznacza to, że ich rozmieszczenie w genomie nie jest losowe, lecz ma podłoże ewolucyjne – ich położenie w poszczególnych regionach chromosomowych było zachowane w kolejnych rundach duplikacyjnych genomu *A. thaliana*. A zatem, w przeciwieństwie do większości genów, syntazy ACC „oparły się” intensywnemu procesowi utraty genów (ang. gene loss). Utrata genów jest typowym elementem zjawiska wtórnej diploidyzacji, które miało prawdopodobnie miejsce po każdym cyklu poliploidyzacyjnym u gatunków z rodziny *Brassicaceae*.

* W wypadku oksydaz ACC symbol i numeracja genów zostały utworzone na potrzeby niniejszego artykułu.

Podsumowanie i perspektywy

Bioinformatyczna analiza puli genowej poznanej w zsekwencjonowanych genomach *A. thaliana* i ryżu jest obecnie punktem odniesienia do złożonych genomów gatunków uprawnych, spokrewnionych z tymi modelowymi gatunkami. W wypadku genomu *A. thaliana* stanowi on punkt wyjścia dla poznania struktury i organizacji genomów licznych gatunków uprawnych, które okazują się być odległymi poliploidami. Są to, poza blisko spokrewnionymi z nim gatunkami z rodziny *Brassicaceae*, także kukurydza, soja, łubin, len, konopie i inne. Szczególnie interesująca będzie analiza różnych wariantów tego samego genu – zarówno allelicznych, jak i nieallelicznych – i ich znaczenie dla adaptacji roślin do warunków środowiska. Ostatnio ukazały się pierwsze doniesienia dotyczące tych zagadnień u *A. thaliana*, będące cennymi przykładami podjęcia tego typu prac u gatunków uprawnych.

Literatura

- [1] Alexander L., Grierson D. 2002. Ethylene biosynthesis and action in tomato: a model for climacteric fruit ripening. *J. Exp. Bot.* 53(377): 2039–2055.
- [2] Al-Shehbaz I.A. 1973. The biosystematics of the genus *Thelypodium* (*Cruciferae*). *Contrib. Gray. Herb. Harv. Univ.* 204: 3–148.
- [3] Arumuganathan K., Earle E.D. 1991. Estimation of nuclear DNA content of plants by flow cytometry. *Plant Mol. Biol. Rep.* 9: 208–218.
- [4] Babula D., Kaczmarek M., Barakat A., Delseny M., Quiros C.F., Sadowski J. 2003. Chromosomal mapping of *Brassica oleracea* based on ESTs from *Arabidopsis thaliana*: complexity of the comparative map. *Mol. Genet Genomics* 268: 656–665.
- [5] Barakat A., Matassi G., Bernardi G. 1998. Distribution of genes in the genome of *Arabidopsis thaliana* and its implications for the genome organization of plants. *Proc. Natl. Acad. Sci. USA* 95: 10044–10049.
- [6] Blanc G., Barakat A., Guyot R., Cooke R., Delseny M. 2000. Extensive duplication and reshuffling in the *Arabidopsis* genome. *Plant Cell* 12: 1093–1101.
- [7] Blanc G., Hokamp K., Wolfe K.H. 2003. A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis* genome. *Genome Res.* 13(2): 137–144.
- [8] Boguski M.S., Lowe T.M., Tolstoshev C.M. 1993. dbEST – database for „expressed sequence tags”. *Nat. Genet.* 4: 332–333.
- [9] Ching A., Rafalski A. 2002. Rapid genetic mapping of ESTs using SNP pyrosequencing and indel analysis. *Cell Mol. Biol. Lett.* 7(2B): 803–810.
- [10] Cooke R., Raynal M., Laudie M., Delseny M. 1997. Identification of members of gene families in *Arabidopsis thaliana* by contig construction from partial cDNA sequences: 106 genes encoding 50 cytoplasmic ribosomal proteins. *Plant J.* 11: 1127–1140.
- [11] Davis G.L., McMullen M.D., Baysdorfer C., Musket T., Grant D., Houchins K., Chao S., Coe E.H. 1999. A maize map standard with sequenced core markers, grass genome refe-

- rence points and 932 expressed sequence tagged site ESTs in a 1,736 – locus map. *Genetics* 152: 1137–1172.
- [12] Gale M.D., Devos K.M. 1998. Comparative genetics in the grasses. *Proc. Natl. Acad. Sci. USA* 95: 1971–1974.
- [13] Gebhardt C., Walkemeier B., Henselewski H., Barakat A., Delseny M., Stuber K. 2003. Comparative mapping between potato (*Solanum tuberosum*) and *Arabidopsis thaliana* reveals structurally conserved domains and ancient duplications in the potato genome. *Plant J.* 34(4): 529–541.
- [14] Goff S.A., Ricke D., Lan T.H., Presting G., Wang R., Dunn M., Glazebrook J., Sessions A., Oeller P., Varma H. 2002. A draft sequence of the rice genome (*Oryza sativa* L. ssp. *japonica*). *Science* 296: 92–100.
- [15] Hamilton A.J., Lycett G.W., Grierson D. 1990. Antisense gene that inhibits synthesis of the hormone ethylene in transgenic plants. *Nature* 346: 284–287.
- [16] Kempin S.A., Savidge B., Yanofsky M.F. 1995. Molecular basis of the cauliflower phenotype in *Arabidopsis*. *Science* 267(5197): 522–525.
- [17] Klee H., Hayford M., Kretzmer K., Barry G., Kishore G. 1991. Control of ethylene synthesis by expression of a bacterial ACC deaminase in transgenic tomato plants. *Plant Cell* 3: 1187–1193.
- [18] Kurata N., Umehera Y., Tanouse H., Sasaki T. 1997. Physical mapping of the rice genome with YAC clones. *Plant Mol. Biol.* 35: 101–113.
- [19] Lagercrantz U. 1998. Comparative mapping between *Arabidopsis thaliana* and *Brassica nigra* indicates that *Brassica* genomes have evolved through extensive genome replication accompanied by chromosome fusions and frequent rearrangements. *Genetics* 150: 1217–1228.
- [20] Lan T.H., DelMonte T.A., Reischmann K.P., Hyman J., Kowalski S.P., McFerson J., Kresovich S., Paterson A.H. 2000. An EST-enriched comparative map of *Brassica oleracea* and *Arabidopsis thaliana*. *Genome Res.* 10: 776–788.
- [21] Michalek W., Weschke W., Pleissner K.P., Graner A. 2002. EST analysis in barley defines a unigene set comprising 4,000 genes. *Theor. Appl. Genet.* 104(1): 97–103.
- [22] Neupane K.R., Mukatira U.T., Kato C., Stiles J.I. 1998. Cloning and characterization of fruit expressed ACC synthase and ACC oxidase from papaya (*Carica papaya* L.). *Acta Horticulturae* 461: 329–337.
- [23] Oeller P.W., Lu M.W., Taylor L.P., Pike D.A., Theologis A. 1991. Reversible inhibition of tomato fruit senescence by antisense RNA. *Science* 18; 254(5030): 437–439.
- [24] Paterson A.H., Lan T.-H., Amasino R., Osborn T.C., Quiros C. 2001. *Brassica* genomics: a complement to, and early beneficiary of, the *Arabidopsis* sequence. *Genome Biology* 3: 1011.1–1011.4.
- [25] Pogson B.J., Downs Ch.G., Davies K.M. 1995. Differential expression of two ACC oxidase genes in broccoli may indicate a role for reproductive structures in senescence. *Plant Physiology* 108: 651–657.
- [26] Schenk P.M., Kazan K., Wilson I., Anderson J.P., Richmond T., Somerville S.C., Manners J.M. 2000. Coordinated plant defense responses in *Arabidopsis* revealed by microarray analysis. *Proc. Natl. Acad. Sci. USA* 97: 11655–11660.
- [27] Schoen D.J. 2000. Comparative genomics, marker density and statistical analysis of chromosome rearrangements. *Genetics* 154: 943–952.

- [28] Somerville C., Dangl L. 2000. Genomics. Plant biology in 2010. *Science* 290(5499): 2077–2078.
- [29] The *Arabidopsis* Genome Initiative 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408: 796–815.
- [30] Wang K.L.C., Li H., Ecker J.R. 2002. Ethylene biosynthesis and signaling network. *The Plant Cell* S131–S151.
- [31] Weterings K., Pezzotti M., Cornelissen M., Mariani C. 2002. Dynamic 1-aminocyclopropane-1-carboxylate-synthase and-oxidase transcript accumulation patterns during pollen tube growth in tobacco styles. *Plant Physiology* 130(3): 1190–1200.
- [32] Wilkinson J.Q., Lanahan M.B., Clark D.G., Bleecker A.B., Chang C., Meyerowitz E.M., Klee H.J. 1997. A dominant mutant receptor from *Arabidopsis* confers ethylene insensitivity in heterologous plants. *Nature Biotech.* 15: 444–447.
- [33] Yang C.Y., Chu F.H., Wang Y.T., Chen Y.T., Yang S.F., Shaw J.F. 2003. Novel Broccoli 1-Aminocyclopropane-1-carboxylate Oxidase Gene (Bo-ACO3) Associated with the Late Stage of Postharvest Floret Senescence. *J. Agric. Food Chem.* 51(9): 2569–2575.
- [34] Yu J., Hu S., Wang J., Wong G.K., Li S., Liu B., Deng Y., Dai L., Zhou Y., Zhang X., Cao M., Liu J., Sun J., Tang J., Chen Y., Huang X., Lin W. 2002. Draft sequence of the rice genome (*Oryza sativa* L. ssp. *indica*). *Science* 296(5565): 79–92.
- [35] Ziółkowski P. A., Blanc G., Sadowski J. 2003. Structural divergence of chromosomal segments that arose from successive duplication events in the *Arabidopsis* genome. *Nucleic Acids Res.* 31: 1339–1350.

Genetic diversity in the light of available nucleotide sequences for plant genomes

Key words: genome, gene families, duplications, *Arabidopsis thaliana*, *Brassica*

Summary

Genome analyses in model species have provided value in biological research. As it has become clear that coding gene sequences are well conserved among genera and even tribes, comparative mapping makes it possible to transfer the information on chromosome structure and gene organization from species with well-developed genetic maps to species where such information is scarce. Furthermore, genome comparisons on the basis of a well characterized model genome in relation to more complex genomes in crop plants, offer new information about the mechanisms responsible for the evolution of chromosomal structure. Among dicotyledonous plants, *A. thaliana* plays the role of a model plant, especially for closely related crop species of the genus *Brassica* (cruciferous oilseeds and a variety of leafy vegetables). The *A. thaliana* genome project provides an opportunity of systematic and large-scale identification and isolation of *Brassica* genes, and to a better understanding various

aspects of the *Brassica* biology and, eventually, of its breeding problems. This aspects is demonstrated on an example of the *A. thaliana* gene families coding for ACC syntases and oxidases, key enzymes in ethylene biosynthesis. Further progress in the analysis of *A. thaliana* and rice genomes should establish the synteny with many related crop species by means of detailed comparative mapping. This may greatly facilitate gene identification and studies on their organization within the genome.