

Danuta Babula<sup>1</sup>, Małgorzata Kaczmarek<sup>1</sup>, Piotr Ziółkowski<sup>1</sup>, Jan Sadowski<sup>1,2</sup>

<sup>1</sup>Instytut Genetyki Roślin PAN w Poznaniu,

<sup>2</sup>Uniwersytet im. Adama Mickiewicza w Poznaniu, Zakład Biotechnologii Instytutu Biologii Molekularnej i Biotechnologii,

## Genetyczne i fizyczne podstawy organizacji chromosomów u wybranych gatunków z rodziny *Brassicaceae*

### Genetic and physical basis of chromosomal organization in selected species from *Brassicaceae* family

Słowa kluczowe: *Brassica*, *Arabidopsis thaliana*, genom, mapy chromosomowe

Key words: *Brassica*, *Arabidopsis thaliana*, genome, chromosomal maps

Postęp w poznawaniu struktury i funkcji genomów roślinnych następował wraz z rozwojem map genetycznych, cytogenetycznych i fizycznych, a ostatnio z zakończonym programem sekwencjonowania genomu *Arabidopsis thaliana*. Szybko rozwijająca się bioinformatyka i techniki masowej analizy genomu umożliwiają określenie podstawowych mechanizmów decydujących o kierunkach różnicowania się gatunków oraz stają się fundamentem dla funkcjonalnego scharakteryzowania genów związanych z wieloma procesami biologicznymi. W pracy dyskutowane są aspekty genomiki porównawczej z uwzględnieniem modelowego układu spokrewnionych gatunków: *Arabidopsis thaliana* i rodzaju *Brassica*.

Comprehension of plant genome structure and function has developed as an effect of advances in construction of genetic, cytogenetic and physical maps followed by recent completion of *Arabidopsis thaliana* genome sequencing program. Fast developing bioinformatics and techniques of genome mass analysis facilitate the determination of genetic basis of differentiation among higher plant species and characterization of genes involved in many biological processes. Aspects of comparative genomics in relation to model system of related species: *Arabidopsis thaliana* and genus *Brassica* are discussed.

## Wstęp

Badanie struktury i wewnętrznej organizacji genomów dostarcza istotnych informacji m.in. o liczbie genów i ich liniowym porządku na chromosomach, odległościach między nimi, stopniu skupienia bądź rozproszenia genów na chromosomach oraz różnego rodzaju aberracjach chromosomowych, które miały miejsce w trakcie ewolucji gatunku. W tym zakresie prowadzone są obecnie prace

na poziomie pierwszorzędowej struktury chromosomowej. Prace te mają na celu poznanie sekwencji nukleotydowej całych genomów. Pozwoli to, poza uzyskaniem ww. informacji, na określenie podstaw genetycznego zróżnicowania między różnymi gatunkami roślin wyższych i funkcjonalnego scharakteryzowania genów związanych z wieloma procesami biologicznymi, takimi jak obrona przed nieprzyjającymi czynnikami środowiska, fotomorfogeneza, regulacja ekspresji genów, przebieg rozwoju, fundamentalne szlaki metaboliczne, transport i naprawa DNA. Informacje dotyczące sekwencji nukleotydowej całych genomów, czy choćby wybranych segmentów chromosomowych, stają się obecnie ważnym elementem w dalszych badaniach nad procesami wspólnymi dla wielu organizmów.

Wielkość genomów roślinnych waha się od około 38 Mpz (Mega par zasad; 1 Mpz = 1 000 000 par zasad) dla *Cardamine amara* (*Brassicaceae*) do ponad 87 000 Mpz dla *Fritillaria assyriaca* (*Liliaceae*) (Flavell i in. 1974, Bennett i Leitch 1995). Pomimo tak wielkiego zróżnicowania w fizycznych rozmiarach genomów, liczba genów, która waha się prawdopodobnie pomiędzy 20 000 a 50 000 oraz ich liniowy porządek, szczególnie w krótkich odcinkach chromosomowych, są podobne. Różnice w wielkości genomów wynikają z różnic w poziomie poliploidalności, a także w zawartości sekwencji powtarzających się.

### **Charakterystyka genomów w obrębie rodziny *Brassicaceae***

---

Rodzina *Brassicaceae* obejmuje 360 rodzajów należących do 13 plemion (Schultz 1936, Al-Shehbaz 1973). Gatunki w obrębie tej rodziny cechują się znacznym zróżnicowaniem haploidalnej liczby chromosomów od  $n = 5$  (*Arabidopsis thaliana*) do  $n = 19$  (*Brassica napus*) i zawartością DNA od 38 Mpz u *Cardamine amara* do 1235 Mpz u *B. napus*. Rodzaj *Brassica* L. obejmuje kilkaset diploidalnych i amfidiploidalnych gatunków, zarówno dzikich jak i uprawnych. Wiele z nich jest uprawianych jako warzywa (np. kalafior, brukselka, kapusty, brokuły), pasza dla zwierząt (np. kapusty pastewne), źródło oleju (rzepak, rzepa, gorczyca sarepska, gorczyca czarna), składniki przypraw (np. gorczyca czarna, gorczyca sarepska), wykazują również ważną aktywność antyrakową (Beecher 1994). Na podstawie badań cytologicznych i analizy mieszańców ujawniono ciągły zakres haploidalnej liczby chromosomów od 7 do 12 (Mizushima 1980). Istotna praca U (1935), oparta na międzygatunkowej hybrydyzacji i metodach cytogenetycznych pozwoliła ustalić pokrewieństwo pomiędzy sześcioma gatunkami *Brassica*, które do dzisiaj znane jest jako „trójkąt U”. Ta i późniejsze prace wskazują na dużą złożoność genomów w rodzaju *Brassica*, wynikającą z występowania genomów o różnej liczbie chromosomów i wewnątrzgenomowych duplikacji. Oczekuje się, że zrozumienie genomowych relacji pomiędzy gatunkami *Brassica* przyniesie postęp w dalszych pracach hodowlanych i umożliwi międzygenomowy transfer genów.

## Roślina modelowa i jej genom

---

Przedstawicielem rodziny *Brassicaceae* jest również *Arabidopsis thaliana* (L.) Heynh. Sprawia to, że rodzina *Brassicaceae*, poza *Poaceae*, jest obecnie najintensywniej badaną wśród roślin. *Arabidopsis* jest blisko spokrewniony z uprawnymi gatunkami rodzaju *Brassica*, mimo przynależności do różnych plemion: odpowiednio *Sisymbrieae* i *Brassicaceae* (Meyerowitz i Pruitt 1985). Z tego powodu rośliny uprawne z rodzaju *Brassica* będą pierwszymi beneficjentami programu genomowego *A. thaliana*. Cechami *Arabidopsis* decydującymi o jego wybraniu na organizm modelowy dla roślin dwuliściennych były: krótki cykl rozwoju, mała wielkość, wielka liczba potomstwa, łatwość transformowania, stosunkowo mały genom jądrowy i niska zawartość sekwencji powtarzających się. Ta mała roślina cechuje się szerokim naturalnym rozmieszczeniem w Europie, Azji i Ameryce Północnej. Cały cykl życiowy wynosi 6 tygodni. Jest rośliną samopylną. Z naturalnych populacji wyselekcjonowano wiele ekotypów, a ekotypy Columbia i Landsberg zostały wybrane jako standardy dla badań genetycznych i molekularnych. W 1990 r. sporządzono zarys długoterminowych celów badań, które objęły stworzenie kolekcji mutantów, pozwalających na identyfikację poszczególnych genów. Przełom w badaniach nastąpił w 1996 r. wraz z powstaniem The *Arabidopsis* Genome Initiative (The AGI) — koordynatora projektu analizy genomu. Obejmowały one zadania m.in. sekwencjonowanie wielkiej liczby EST-ów (expressed sequence tag), konstrukcję kolekcji klonów YAC (yeast artificial chromosome) i BAC (bacterial artificial chromosome), budowanie szczegółowych map genetycznych i fizycznych dla wybranych regionów genomu i molekularną analizę wielu pojedynczych genów. Poznanie pełnej sekwencji genomu *Arabidopsis* pozwoliło precyzyjnie określić jego strukturę (The AGI 2000). Genom *Arabidopsis* obejmuje 5 par chromosomów o łącznej długości 125 Mpz; dla 115,4 Mpz (92%) określono sekwencję nukleotydową (dane z 2000 r.) z pominięciem obszarów centromerowych i regionów powtórzeń rDNA. Zidentyfikowano 25 498 genów kodujących białka, zebranych w 11 000 rodzin (The AGI 2000). Szczegółowa analiza sekwencji pozwoliła stwierdzić, że gęstość genów, poziomy ekspresji i rozmieszczenie sekwencji powtarzających się są stałe wzdłuż 5 chromosomów. Geny są rozmieszczone gęsto i pokrywają prawie 85% genomu, zaś regiony pozbawione genów są mocno zredukowane. Sekwencje powtarzające się obejmują natomiast jedynie około 20–30% genomu. Dane oparte na analizie EST-ów potwierdziły fakt, że pomimo niewielkiej wielkości genomu, *A. thaliana* zawiera dużą liczbę małych rodzin genów (Cooke i in. 1996).

Ważnym doniesieniem opartym na analizie prawie całej sekwencji genomu jest to, że większość genomu *Arabidopsis* występuje w postaci zduplikowanych dłuższych rejonów chromosomowych (The AGI 2000). Odkrycie to potwierdza wcześniejsze sugestie o segmentalnej duplikacji, oparte na mapowaniu gene-

tycznym (Kowalski i in. 1994). W genomie *A. thaliana* ujawniono 24 duże zduplikowane segmenty o długości od 100 kbp do 4 Mbp, pokrywające 58% (65,6 Mbp) genomu. To rzuca nowe światło na ewolucję genomu *A. thaliana*, wskazując na nagromadzenie się licznych rearanżacji obejmujących duplikacje, translokacje, insercje, inwersje i delecje. Trzeba zaznaczyć, że duplikacje genomu *A. thaliana* znacznie komplikowały fizyczne mapowanie i sekwencjonowanie (The AGI 2000).

Znajomość pełnej sekwencji *A. thaliana* jest ważna dla wszystkich biologów. Ostatnie odkrycia, które ukazują nieoczekiwany poziom syntenii w genomach różnych roślin sugerują, że nie tylko informacja o specyficznych genach i procesach będzie użyteczna w analizie genomu innych roślin, ale również wiele genomowych materiałów, narzędzi biologii molekularnych, technik opracowanych i zgromadzonych dla *A. thaliana* będzie wykorzystanych w pracach nad innymi gatunkami (Schmidt 2000). Analiza długości rejonów konserwatywnych w genomach *A. thaliana* i gatunków *Brassica* wskazuje, że obejmują one odcinki o długości 5–10 cm. Wydaje się więc, że odcinki chromosomowe *B. oleracea* o fizycznej długości 2,5–5,0 Mbp, zawierające około 300 genów, mają podobny porządek liniowy genów jak ich odcinki homologiczne u *A. thaliana*. Informacje te mają duże znaczenie dla dalszych prac genetycznych w rodzaju *Brassica*. Ostateczna integracja genetycznych (w tym genomowych), biochemicznych, fizjologicznych i morfologicznych danych pozwoli na poznanie wielu nowych aspektów biologii roślin.

Widoczne jest duże zainteresowanie wykorzystaniem sekwencji genomu *Arabidopsis* w badaniach innych roślin. Informacje wypływające z poznania całej sekwencji genomu *Arabidopsis*, dostępność wielu mutantów, licznych kolekcji sklonowanych sekwencji, wydajnych systemów transformacyjnych, szczegółowych map genetycznych i fizycznych oraz ponad 37 000 EST-ów są bardzo przydatne dla dalszych badań prowadzonych w różnych dziedzinach biologii i nauk rolniczych. Dostępna sekwencja nukleotydowa genomu *A. thaliana* jest ważnym, kolejnym krokiem w zrozumieniu organizacji genomów wszystkich roślin dwuliściennych. Poznanie sekwencji całego genomu organizmu modelowego znacznie przyspieszy i ułatwi identyfikację genów. Identyfikacja genów uczestniczących w rozwoju i fundamentalnych procesach metabolicznych otwiera nowe możliwości w manipulowaniu nimi przy wykorzystaniu metod inżynierii genetycznej celem ich modyfikacji.

## Mapy genetyczne

---

Większość informacji o strukturze i organizacji genomów pochodzi z badań genetycznych z zastosowaniem markerów molekularnych i analizy map genetycznych. Podobnie jak globalne sekwencjonowanie, mapy genetyczne typu np. RFLP (restriction fragment length polymorphism) pozwalają na lokalizację

poszczególnych genów na chromosomach. Ich dokładność zależy w dużej mierze od liczby zmapowanych loci. Należy tu podkreślić, że genomy gatunków z rodzaju *Brassica* cechują się znaczną złożonością, wynikającą z licznych duplikacji i innych wewnętrznych rearanżacji (Quiros i in. 1991, Teutonico i Osborn 1994). Rozwój szczegółowych map genetycznych i rozwijające się intensywnie fizyczne mapowanie genomów roślin (patrz niżej) pozwolą na skorelowanie fizycznych i genetycznych odległości. Pierwsze mapy sprzężeń dla gatunków *Brassica* oparte były na dziedziczeniu cech morfologicznych (Yarnell 1956, Sampson 1966). Analiza map sprzężeń ujawniła, że w obrębie tego rodzaju gatunki zachowują wysoki stopień kolinearności (Lydiate i in. 1993). Równolegle stwierdzono złożoność genomów w rodzaju *Brassica*, wynikającą z licznych duplikacji (Kianian i Quiros 1992, Lan i in. 2000, Babula i in. 2000). Stwierdzono liczne aberracje, związane głównie z duplikacjami, zarówno pomiędzy gatunkami jak i wśród morfotypów, np. u *B. oleracea*. Takson, taki jak rodzaj *Brassica*, który odznacza się wysokim tempem strukturalnych rearanżacji chromosomów może być doskonałym systemem dla badań procesów mutacji strukturalnych.

Mapy genetyczne uprawnych gatunków *Brassica* będą wykorzystywane w genetyce stosowanej i hodowli (Matuszczak i in. 1999). Pozwalają na określenie lokalizacji genów determinujących cechy ekonomicznie interesujące, badanie struktury oraz pochodzenia i ewolucji genomów. Szczególnie przydatne stają się tu mapy powstałe w oparciu o analizę RFLP, mimo, że budowa takiej mapy jest znacznie wolniejsza niż konstruowanie map opartych o markery DNA typu RAPD (random amplified polymorphic DNA) czy AFLP (amplified fragment length polymorphism). Zaletą tych map jest to, że pozwalają na identyfikację zestawów spokrewnionych loci występujących na homeologicznych chromosomach różnych genomów. Wysoka zachowawczość sekwencji genów podczas ewolucji otwiera możliwości wykorzystania markerów genetycznych z jednego gatunku jako sondy przy tworzeniu mapy dla innego gatunku. Zastosowanie wspólnych zestawów sond w mapowaniu genomów różnych gatunków staje się użytecznym narzędziem dla porównawczego mapowania, pozwalając na bezpośrednią identyfikację homeologicznych loci i kolinearnych segmentów chromosomowych. Przez ostatnich 10 lat przedstawiono wiele prac nad pokrewieństwem między genomami różnych gatunków. Do porównań wybierano gatunki stosunkowo blisko ze sobą spokrewnione. Odkrycie wielkich zachowawczych segmentów chromosomów wśród różnych taksonów sugeruje możliwość skonstruowania ujednoczonych map dla licznych grup organizmów, co może mieć znaczne konsekwencje dla genetycznych i biotechnologicznych zastosowań. Znajomość lokalizacji sekwencji kodujących jest jednak ograniczona, bowiem większość map zawiera tysiące anonimowych markerów, a tylko 5–10% markerów dotyczy bezpośrednio genów. W rodzinie *Brassicaceae* wykonano porównania pomiędzy gatunkami blisko z sobą spokrewnionymi, jak *B. oleracea* i *B. rapa* (Slocum 1989, McGrath i Quiros 1991),

*B. oleracea* i *B. napus* (Cheung i in. 1997), *B. oleracea* i *A. thaliana* (Kowalski i in. 1994, Lan i in. 2000, Babula i in. 2000), *B. rapa* i *A. thaliana* (Jackson i in. 2000), *B. napus* i *A. thaliana* (Lagercrantz i in. 1996), *B. rapa*, *B. napus* i *A. thaliana* (Osborn i in. 1997), *B. napus*, *B. oleracea* i *A. thaliana* (Teutonico i in. 1994), porównanie trzech podstawowych genomów (A, B i C) w rodzaju *Brassica* (Truco i in. 1996, Lagercrantz i Lydiate 1996). Prace te nie przedstawiają pełnego obrazu struktury genomów ze względu na ograniczoną liczbę wspólnych sond i poziom wykrytego polimorfizmu, rzucają jednak nowe światło na ewolucję gatunków w obrębie *Brassicaceae*. Wykrycie syntenii chromosomowej pomiędzy spokrewnionymi gatunkami umożliwi identyfikację i wyizolowanie genów z roślin uprawnych o wielkich genomach, wykorzystując informacje o homologicznych genach u pokrewnych roślin z mniejszymi genomami. Porównawcze mapowanie pomiędzy różnymi gatunkami umożliwi rekonstrukcję tempa i rodzajów wymiany segmentów chromosomów oraz dostarcza użytecznych informacji dla określenia filogenetycznych stosunków wśród odległych taksonów. Ostatecznie pozwoli zapewne sklonować geny, które mogły odegrać ważniejszą rolę w specjacji i adaptacji.

Wiele współczesnych gatunków cechuje się poliploidalnym pochodzeniem, m.in.: kukurydza, diploidalne gatunki w rodzinie *Brassicaceae*, pszenica, soja, łubin, diploidalna bawełna, sorgo i inne. Powstawanie poliploidów wiąże się z rozległą rekonstrukcją chromosomów, spowodowaną dodaniem chromosomów. Stan ten prowadzi najczęściej do dywergencji genów i dalszej rekonstrukcji chromosomów. Struktura populacji i niedawna poliploidyzacja są prawdopodobnie ważnymi czynnikami przyczyniającymi się do zachodzenia szybkich rearanzacji chromosomów w rodzinie *Brassicaceae*. Szczegółowe analizy map genetycznych poparte analizą sekwencji wybranych rejonów wskazują, że nawet proste genomy zawierają wielką liczbę segmentalnych duplikacji.

Rodzaje *Arabidopsis* i *Brassica* rozeszły się około 12,2–19,2 mln lat temu (Cavell i in. 1998). Mimo to, geny u *Brassica* wykazują wysoki stopień zachowawczości na poziomie sekwencji nukleotydowej z ich ortologami u *A. thaliana*, wynoszący średnio 86% w sekwencjach kodujących (Lydiate i in. 1993, Cavell i in. 1998). Genom *A. thaliana* jest w przybliżeniu czterokrotnie mniejszy niż genomy *Brassica* (Arumuganathan i Earle 1991) (tab. 1). Pomimo tej różnicy badania porównawcze u obu gatunków wskazują, że struktura genomów gatunków rodzaju *Brassica* jest podobna do *A. thaliana*, z tym, że cechuje się rozległą triplikacją i intensywną dywergencją (O'Neill i in. 2000). To nadzwyczajne tempo zmian w powtórzonych segmentach genomu *Brassica* wskazuje, że paleopoliploidy podlegały szybkiej ewolucji chromosomalnej. Analiza kolinearności w rodzinie *Brassicaceae* była także oparta na mapowaniu fizycznym (Sadowski i in. 1996, Sadowski i Quiros 1998, Conner i in. 1998, Cavell i in. 1998, Sadowski i in. 1998). Badania te umożliwiły identyfikację kolinearności chromosomowej na podstawie wybranych grup genów. Ponieważ oba rodzaje *Arabidopsis*

i *Brassica* oznaczają się wysokim poziomem zduplikowania genomów, porównawcze mapowanie pozwoli określić czy wykryte duplikacje zaszły przed czy po rozejściu się obu gatunków. Jednak liczne segmentalne inwersje i translokacje genów u *B. oleracea*, które nie były obserwowane w genomie *Arabidopsis* wskazują, że brały tu udział również inne mechanizmy. Określenie długości kolinearnych odcinków w genomach różnych gatunków pozwoli wydajnie stosować fizycznie zmapowane klony np. typu BAC i dane sekwencji genomowej *A. thaliana* w genetycznym mapowaniu i izolacji genów u uprawnych gatunków *Brassica*.

Tabela 1

Porównanie cech strukturalnych genomów *A. thaliana* i *B. oleracea*  
*Comparison of structural features in A. thaliana and B. oleracea genomes*

|  | <i>A. thaliana</i> | <i>B. oleracea</i> |
|--|--------------------|--------------------|
| Liczba genomowa — <i>Genomic number</i>  | n = 5              | n = 9              |
| Wielkość fizyczna (Mpz) — <i>Physical size</i>   | 125–130            | 599–662            |
| Duplikacje chromosomowe<br><i>Chromosome duplications</i>  | >75%               | >75% ?             |
| Liczba genów — <i>Gene number</i>  | 25 500             | 30 000–40 000      |
| Skupienie genów — <i>Gene</i>  | 1 gen/ok. 5 kpz    | 1 gen/15–20 kpz    |
| Podobieństwo sekwencji nukleotydowej genów<br><i>Similarity of nucleotide sequences of genes</i> |                    | ~ 86%              |

## Cytogenetyczne mapy chromosomowe

Trzeba zaznaczyć, że genetyczne mapy sprzężeń, ze względu na ograniczoną liczebność populacji mapujących i zakres polimorfizmu, zawsze obciążone są pewnym błędem; w miarę rozwoju mapy i przybywania nowych markerów następuje ciągła jej modyfikacja i uściślanie. Ograniczenia wynikają z samej istoty mapowania genetycznego, tj. wykrycia polimorfizmu. Aby dany marker dało się nanieść na mapę musi on segregować w pokoleniu F<sub>2</sub> — tylko pewna część sond pozwala na wykrycie absolutnego polimorfizmu, zwykle nie udaje się zmapować wszystkich kopii danego genu (wzgl. danej sekwencji niekodującej). W takiej sytuacji zachodzi potrzeba rozwijania nowych metod umożliwiających precyzyjne mapowanie całych genomów i wszystkich kopii danego genu. Jedną z nich jest technika fluorescencyjnej hybrydyzacji *in situ* (FISH), która u zwierząt i człowieka jest już z powodzeniem stosowana od stosunkowo dawna do mapowania unikalnych sekwencji (Landegent i in. 1985). Chociaż u roślin do tej pory próby mapowania pojedynczych genów i innych sekwencji unikalnych nie są zbyt częste (Fransz i in. 1996, Ohmido i in. 1998, Hoopen i in. 1999, Zhong i in. 1999),

to jednak rozwinięcie takiej metodyki pozwoliłoby na bardzo szybki postęp w poznawaniu roślinnych genomów.

Należy podkreślić, że informacje o ewolucji chromosomów u roślin wyższych oparte były do niedawna przede wszystkim na analizie mejotycznej i porównaniu kariotypów. Chromosomy u gatunków z rodzaju *Brassica* są stosunkowo małe i nie wykazują wyraźnych różnic w morfologii. Znacznie to utrudnia prowadzenie badań w oparciu o metody cytogenetyczne. Pierwsze prace nad rozróżnieniem poszczególnych chromosomów u gatunków diploidalnych oparte były na obserwacji koniugacji chromosomów (Robbelen 1960). Od kilku lat z powodzeniem stosuje się nowoczesne techniki oparte na hybrydyzacji *in situ* przy pomocy techniki FISH (fluorescent *in situ* hybridization) (Małuszyńska i Heslop-Harrison 1993, Fukui i in. 1998, Jackson i in. 2000, Ziółkowski i Sadowski 2001) i sond molekularnych z *Arabidopsis thaliana*. Obecnie interesującym narzędziem w mapowaniu mogą być sondy typu BAC noszące fragmenty chromosomowe *A. thaliana* o średniej długości 100 kpz i średniej liczbie 20 genów. Zastosowanie sond DNA z *A. thaliana* dla hybrydyzacji *in situ* do chromosomów *Brassica* jest możliwe ze względu na wysoką homologię sekwencji kodujących pomiędzy tymi dwoma rodzajami roślin. Liczne wcześniejsze badania, w szczególności prace nad porównawczym mapowaniem genomów *Brassica* przy zastosowaniu sond molekularnych z *A. thaliana* w oparciu o markery RFLP wykazały, iż w przypadku *Brassica* i *Arabidopsis* mamy do czynienia z podobną organizacją genową wielu segmentów chromosomowych (Sadowski i in. 1996). Włączenie do tych badań również techniki FISH usprawni i przyspieszy proces porównawczej analizy genomów. Wprowadzenie techniki heterologicznego mapowania FISH (sondy *A. thaliana* / chromosomy *B. oleracea*; Ziółkowski i Sadowski 2001) może mieć ogromne znaczenie dla poznawania genomów roślin uprawnych, strukturalnie bardziej złożonych, w oparciu o poznane na poziomie pierwszorzędowym, proste genomy roślin modelowych (*A. thaliana* i *Oryza sativa* — ryżu). Niewielkie rozmiary chromosomów metafazalnych u *Brassica* (1–5  $\mu\text{m}$ ) uniemożliwiają ich zastosowanie do liniowego mapowania kilkudziesięciu czy nawet więcej loci wzdłuż pojedynczego chromosomu. Ostatnio obiektem doświadczeń stają się biwalenty chromosomowe z profazy mejozy I w stadium pachytenu (Ziółkowski i Sadowski 2001), które będąc znacznie większe (70–100  $\mu\text{m}$ ) pozwalają na wykrywanie i mapowanie sekwencji fizycznie zlokalizowanych nawet bardzo blisko siebie, tzn. na odległości 120 kpz (średnia długość chromosomu np. u *B. oleracea* wynosi około 70 Mpz = 70 mln par zasad).

Ponieważ przynajmniej 60% genomu *Arabidopsis* jest zlokalizowane w zduplikowanych segmentach, może to sugerować, że *Arabidopsis*, podobnie jak kukurydza ma tetraploidalnego przodka (Blanc i in. 2000). Stosunkowo niewiele wiadomo o pochodzeniu tych duplikacji. Mogą one wynikać z co najmniej czterech wielkich zdarzeń duplikacji, które zdarzyły się 100–200 mln lat temu, w okresie różnicowania się nasiennych (Vision i in. 2000). Z innej strony sugeruje się,



że powstały one w pojedynczym zdarzeniu poliploidyzacji, a następnie uległy przemieszaniu przez chromosomalne rearanżacje (Blanc i in. 2000). Ta hipoteza przewiduje, że każdy region w genomie *Arabidopsis* reprezentowany jest w dwóch kopiach. Jednak ostatnie wyniki porównawczego mapowania pomiędzy *Arabidopsis* a pomidorem oraz między *Arabidopsis* a sorgo wskazują, że kilka regionów genomu obecne jest w trzech lub więcej kopiach, lecz w chwili obecnej nie można określić ich udziału w całej strukturze genomu (Grant i in. 2000, Ku i in. 2000). Jeśli genom obejmował wielokrotne duplikacje, wtedy obecność tylko 5 chromosomów wskazuje na wystąpienie fuzji chromosomów. Porównawcza analiza sekwencji *A. thaliana* i pomidora wykazała, że te duplikacje zdarzyły się około 112 mln lat temu, podczas formowania tetraploida. Poziomy zachowawczości zduplikowanych segmentów mogą wynikać z dywergencji od formy starożytnego autotetraploida lub mogły rzutować na różnice obecne w allotetraploidalnym przodku. Ale jest także możliwe, że kilka niezależnych segmentalnych duplikacji miało miejsce poza formowaniem i stabilizacją tetraploida. Rozległe duplikacje obserwowane w genomie *Arabidopsis* popierają hipotezę, że większość z ostatnich tych duplikacji było wynikiem zdarzeń poliploidyzacji. Jest to zgodne z sugestiami Kowalskiego i in. (1994), który przyjął, że *A. thaliana* można uważać za paleopoliploida. Częstość występowania tandemowych duplikacji i widocznych delecji pojedynczych genów lub mniejszych grup genów ze zduplikowanych regionów sugeruje, że kluczowym mechanizmem był nierówny crossing over wpływający na ewolucję genomu roślin. U *A. thaliana* pomimo, że 75% genomu jest zduplikowana, tylko 19% z przebadanych EST-ów weszło do sprężonych grup. Dywergencja tych zduplikowanych segmentów znacznie maskuje ich ewolucyjną historię. Nadal ważnym pytaniem pozostaje, co się dzieje z tymi genami po duplikacji? Wydaje się, że odpowiedzialne za to są mechanizmy prowadzące do utraty genów lub przemieszczania ich do nowych miejsc oraz zmiany funkcjonalne uniemożliwiające ich identyfikację.

## Podsumowanie

---

Ze względów metodycznych i finansowych nie jest możliwe sekwencjonowanie całych genomów wielu roślin wyższych, dlatego dostępne obecnie informacje o ich strukturze pochodzą często z analizy map genetycznych czy cytogenetycznych. Postęp w poznaniu struktury i funkcji genomów roślinnych następował z udziałem map genetycznych dla wielu różnych gatunków. Początkowo kilka roślin (kukurydza, pomidor, groch, ryż, jęczmień, tytoń i petunia) uznano za genetyczne systemy modelowe. W latach 80. biologowie rozpoczęli jednak poszukiwania innych gatunków modelowych, odpowiedniejszych dla złożonych analiz z wykorzystaniem narzędzi genetyki i biologii molekularnej.

Wybrano gatunki cechujące się małymi genomami: najpierw *Arabidopsis thaliana* (rzodkiewnik pospolity), jako model dla roślin dwuliściennych, a w późniejszym okresie *Oryza sativa* (ryż), będący modelem dla roślin jednoliściennych, głównie zbożowych. W drugiej połowie lat 90. nastąpił szybki i dynamiczny rozwój genomiki — dziedziny genetyki molekularnej zajmującej się strukturą, organizacją i funkcjonowaniem genomów. Obecnie znane są sekwencje kodujące wszystkich pięciu chromosomów *A. thaliana*. Pomimo, że w przypadku około 40% potencjalnych genów ich funkcja nie została poznana, to dostęp do uzyskanych informacji daje bardzo szerokie możliwości dalszych prac badawczych. Mogą być one wykorzystane m.in. dla analizy budowy genomów roślin użytkowych pokrewnych z *A. thaliana*, co w radykalny sposób może wpłynąć na postęp w pracach hodowlanych.

## Literatura

---

- Al-Shehbaz I.A. 1973. The biosystematics of the genus *Thelypodium* (*Cruciferae*). *Contrib. Gray. Herb. Harv. Univ.* 204: 3-148.
- Arumuganathan K., Earle E.D. 1991. Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.* 9: 208-218.
- Babula, D., Kaczmarek M., Delseny M., Quiros C.F., Sadowski J. 2000. Construction of a genetic map for *Brassica oleracea* based on ESTs from the *Arabidopsis thaliana* genome. *Acta Hort.* 539: 95-99.
- 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.
- Beecher C.W.W. 1994. Cancer preventative properties of *Brassica oleracea*. *Am. J. Clin. Nutr.* 56: 1166-1170.
- Bennett M.D., Leitch I.J. 1995. Nuclear DNA amounts in angiosperms. *Ann. Bot.* 76: 113-176.
- Blanc G., Barakat A., Guyot R., Cooke R., Delseny M. 2000. Extensive duplication and reshuffling in the *Arabidopsis* genome. *Plant Cell* 12: 1093-1102.
- Cavell A., Lydiate D., Parkin I., Dean C., Trick M. 1998. A 30 centimorgan segment of *Arabidopsis thaliana* chromosome 4 has six collinear homologues within the *Brassica napus* genome. *Genome* 41: 62-69.
- Cheung W.Y., Champagne G., Hulbert N., Landry B.S. 1997. Comparison of the genetic maps of *Brassica napus* and *Brassica oleracea*. *Theor. Appl. Genet.* 94: 569-582.
- Conner J.A., Conner P., Nasrallah M.E., Nashrallah J.B. 1998. Comparative mapping of the *Brassica* S locus region and its homeolog in *Arabidopsis*: implications for the evolution of mating systems in the *Brassicaceae*. *Plant Cell* 10: 801-812.
- Cooke R., Raynal M., Laudie M., Grellet F., Delseny M., Morris P.C., Guerrier D. i in 1996. Further progress towards a catalogue of all *Arabidopsis* genes: analysis of a set of 5000 non-redundant ESTs. *Plant J.* 9(1): 101-124.
- Flavell R.B., Bennett M.D., Smith J.B., Smith D.B. 1974. Genome size and the proportion of repeated nucleotide sequence DNA in plants. *Biochem. Genet.* 12: 257-269.

- Fransz P.F., Stam M., Montijn B., Hoopen R.T., Wiegant J., Kooter J.M., Oud O., Nanninga N. 1996. Detection of single-copy genes and chromosome rearrangement in *Petunia hybrida* by fluorescence *in situ* hybridization. *The Plant Journal* 9: 767-774.
- Fukui K., Ohmido N., Kamisugi Y., Mathias R.J., Kuginuki Y., Yamabe M. 1993. Analysis and utility of chromosome information. 54. Complete identification of *Brassica* A, B and C genome chromosomes. *Jpn. J. Breed.* 43 (suppl. 2): 360.
- Grant D., Cregan P., Shoemaker R.C. 2000. Genome organization in dicots. I. Genome duplication in *Arabidopsis* and synteny between soybean and *Arabidopsis*. *Proc. Natl. Acad. Sci. USA.* 97: 4168-4173.
- Hoopen R., Montijn B.M., Veuskens J.T., Oud O.J., Nanninga N. 1999. The spatial localization of T-DNA insertions in petunia interphase nuclei: consequences for chromosome organization and transgene insertion sites. *Chromosome Res.* 7: 611-23.
- Hu J., Sadowski J., Osborn T.C., Landry B.S., Quiros C.F. 1998. Linkage group alignment from four independent *Brassica oleracea* RFLP maps. *Genome* 241: 226-235.
- Jackson S.A., Cheng Z.K., Wang M.L., Goodman H.M., Jiang J.M. 2000. Comparative fluorescence *in situ* hybridization mapping of a 431-kb *Arabidopsis thaliana* bacterial artificial chromosome contig reveals the role of chromosomal duplications in the expansion of the *Brassica rapa* genome. 156: 833-838.
- Kianian S.F., Quiros C.F. 1992. Generation of a *Brassica oleracea* composite and RFLP map: linkage arrangements among various populations and evolutionary implications. *Theor. Appl. Genet.* 84: 544-554.
- Kowalski S.P., Lan T.-H., Feldmann K.A., Paterson A.H. 1994. Comparative mapping of *Arabidopsis thaliana* and *Brassica oleracea* chromosomes reveals islands of conserved organization. *Genetics* 138: 499-510.
- Ku H.M., Vision T., Liu J., Tanksley S.D. 2000. Comparing sequenced segments of the tomato and *Arabidopsis* genomes: large-scale duplication followed by selective gene loss creates a network of synteny. *PNAS USA* 97(16): 9121-9126.
- Lagercrantz U., Lydiate D.J. 1996. Comparative genome mapping in *Brassica*. *Genetics* 144: 1903-1910.
- 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.
- 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.
- Landegent J.E., Jansen in de Wal N., van Ommen G.-J., Baas F., de Vijlder J.J.M., van Duijn P., van der Ploeg M. 1985. Chromosomal localization of a unique gene by non-autoradiographic *in situ* hybridization. *Nature* 317: 175-177.
- Lydiate D., Sharpe A., Lagercrantz U., Parkin I. 1993. Mapping the *Brassica* genome. *Outlook on Agriculture* 22: 85-92.
- Małuszyńska J., Heslop-Harrison J.S. 1993. Physical mapping of rDNA loci in *Brassica* species. *Genome* 36: 774-781.
- Matuszczak M., Krzymański J. 1999. Poszukiwanie markerów RAPD różnicujących linie rzepaku ozimego o różnych cechach chemicznych. *Rośliny Oleiste XX* (2): 395-414.
- Meyerowitz E.M., Pruitt R.E. 1985. *Arabidopsis thaliana* and plant molecular genetics. *Science* (Washington, D.C.) 229: 1214-1218.

- Mizushima U. 1980. Genome analysis in *Brassica* and allied genera. In: Tsunoda S., Hinata K., Gomez-Campo C. (eds) *Brassica* crops and wild allies. Japan Scientific Societies Press, Tokyo, 89-105.
- Ohmido N., Akiyama Y., Fukui K. 1998. Physical mapping of unique nucleotide sequences on identified rice chromosomes. *Plant Mol. Biol.* 38: 1043-1052.
- O'Neill C.M., Bancroft I. 2000. Comparative physical mapping of segments of the genome of *Brassica oleracea* var. *alboglabra* that are homoeologous to sequenced regions of chromosomes 4 and 5 of *Arabidopsis thaliana*. *Plant J.* 23: 233-243.
- Osborn T.C., Kole C., Parkin I.A.P., Sharpe A.G., Kuiper M., Lydiate D.J., Trick M. 1997. Comparison of flowering time genes in *Brassica rapa*, *B.napus* and *Arabidopsis thaliana*. *Genetics* 146: 1123-1129.
- Quiros C.F., Hu J., This P., Chevre A.M., Delseny M. 1991. Development and chromosomal localization of genome-specific markers by polymerase chain reaction in *Brassica*. *Theor. Appl. Genet.* 82: 627-632.
- Sadowski J., Gaubier P., Delseny M., Quiros C.F. 1996. Genetic and physical mapping in *Brassica* diploid species of a gene cluster defined in *Arabidopsis thaliana*. *Mol. Gen. Genet.* 251: 298-306.
- Sadowski J., Quiros C.F. 1996. Comparative mapping of *Brassica* and *Arabidopsis*. *J. Appl. Genet.* 37A: 150-152.
- Sadowski J., Quiros C.F. 1998. Organization of an *Arabidopsis thaliana* gene cluster on chromosome 4 including the RPS2 gene, in the *Brassica nigra* genome. *Theor. Appl. Genet.* 96: 468-474.
- Sampson D.R. 1966. Genetic analysis of *Brassica oleracea* using nine genes from sprouting broccoli. *Can. J. Gen. Cytol.* 8: 404-413.
- Schmidt R. 2000. Synteny: recent advances and future prospects. *Curr. Opin. Plant Biol.* 3: 97-102
- Schultz O.E. 1936. Cruciferae. In: Engler A., Prantl K. (eds) *Die natürlichen Pflanzenfamilien*, 2nd edn. Engelmann, Leipzig, 17: 227-658.
- Slocum M.K. 1989. Analyzing the genomic structure of *Brassica* species using RFLP analysis. In: T. Helentjaris & B. Burr (eds.), *Development and application of molecular markers to problems in plant genetics*. Cold Spring Harbor Lab. Press, Cold Spring Harbor, NY.
- Teutonico R.A., Osborn T.C. 1994. Mapping of RFLP and qualitative trait loci in *Brassica rapa* and comparison to the linkage maps of *B. napus*, *B. oleracea*, and *Arabidopsis thaliana*. *Theor. Appl. Genet.* 89: 885-894.
- The *Arabidopsis* Genome Initiative 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. 408: 796-815.
- Truco M.J., Hu J., Sadowski J., Quiros C. 1996. Structure and intra-genomic homology of the *Brassica* genomes: implications for their origin and evolution. *Theor. Appl. Genet.* 93: 1225-1233.
- U N. 1935. Genome analysis in *Brassica* with special reference to the experimental formation of *B. napus* and peculiar mode of fertilization. *Jpn. J. Genet.* 7: 784-794.
- Vision J. 2000. The origins of genomic duplications in *Arabidopsis*. *Science* 290: 2114-2117.
- Yarnell S.H. 1956. Cytogenetics of the vegetable crops: II. Crucifers. *Bot. Rev.* 22: 81-166.
- Zhong X-B., Bodeau J., Fransz P.F., Williamson V.N., van Kammen A., de Jong J.H., Zabel P. 1999. FISH to meiotic pachytene chromosomes of tomato locates the root-knot nematode resistance gene *Mi-1* and the acid phosphatase gene *Aps-1* near the junction of euchromatin and pericentromeric heterochromatin of chromosome arms 6S and 6L, respectively. *Theor. Appl. Genet.* 98: 365-370.
- Ziółkowski P., Sadowski J. 2001. Mapping of *Brassica* pachytene chromosomes by fluorescent *in situ* hybridization. *Genome* (przyjęto do druku).