

Zintegrowana genomika u podstaw nowych agrobiotechnologii

Andrzej B. Legocki

*Instytut Chemii Bioorganicznej PAN
ul. Noskowskiego 12/14, 61-704 Poznań*

Słowa kluczowe: genomika roślin, proteomika, wyłączenie genów, roślinne układy ekspresyjne

Zasadnicze cele nowych biotechnologii zbieżne są w swej istocie z celami tradycyjnych dziedzin rolniczych oraz praktyki hodowlanej. Nowe technologie i zdobycze biologii molekularnej skróciły czas wyprowadzania nowych linii hodowlanych i wpłynęły na niezawodność osiągania celów konwencjonalnych. Poprzez wprowadzenie nowatorskich podejść genetyki niemendlowskiej otworzone zostały nieznane dotychczas horyzonty poznawcze oraz perspektywy aplikacyjne [1, 2]. Niektóre z tych osiągnięć mają charakter przełomowy i mogą odegrać zasadniczą rolę w rozwiązaniu problemów żywienia ludzkości [3]. Główne cele poznawcze stawiane dla polepszenia i zwiększenia pierwotnej produkcji roślinnej bez względu na rodzaj podejść — tradycyjnych bądź nowatorskich — można zgrupować w następujące obszary:

- Podwyższenie wydajności plonowania poprzez zwiększenie odporności roślin uprawnych na chemiczne środki ochrony oraz niekorzystne warunki środowiska (np. zasolenie, susza, krańcowe pH gleby). Wprowadzenie do genomu roślin uprawnych pojedynczych genów odporności może poprawić plonowanie, jednakże większość cech użytkowych zależy od całych zespołów genowych, których przenoszenie jest bardzo trudne z uwagi na konieczność przezwyciężenia barier regulacyjnych.
- Zwiększenie wartości żywieniowych i technologicznych roślin uprawnych. Rozważane tutaj podejścia uwzględniają modyfikowanie endogennych genów strukturalnych, np. w kierunku podwyższenia zawartości żywieniowo korzystnych aminokwasów.
- Obniżenie energochłonności produkcji rolnej poprzez polepszenie wydajności fotosyntetycznej roślin i zwiększenie udziału wiązania biologicznego w bilansie azotu, co mogłoby obniżyć zależność plonowania od kosztownego i ekologicznie niekorzystnego nawożenia chemicznego.

W dłuższej perspektywie czasowej strategiczne cele agrobiotechnologii należy odnosić do globalnych problemów demograficznych, a także uwzględniać kierunki zmienności środowiska naturalnego. Przypuszcza się, że w XXI wieku liczba ludności na Ziemi osiągnie równowagę na poziomie 8–13 mld, przy czym 95% przyrostu odnosić się będzie do krajów rozwijających się w regionach tropikalnych i subtropikalnych Afryki, Ameryki i Azji [4]. Przy aktualnym areale uprawnym naszego globu — 1,5 mld hektarów i przy założeniu przeciętnego wzrostu produkcji roślinnej od $2 \text{ t} \cdot \text{ha}^{-1}$ do $5 \text{ t} \cdot \text{ha}^{-1}$, globalną wielkość produkcji roślinnej szacować można na ekwiwalent 8 mld t pszenicy, co wystarcza do zaspokojenia potrzeb żywieniowych 11 mld ludzi. Z pewnym uproszczeniem przyjąć można, iż taka właśnie jest globalna wydajność żywniowa rolnictwa tradycyjnego [5]. Zaspokojenie większego zapotrzebowania na produkty roślinne wymagałoby wprowadzenia do rolnictwa nowych technologii oraz wydajnych genotypów uprawnych, które umownie określa się nowymi agrobiotechnologiami. Najistotniejszy wkład w unowocześnianie tradycyjnego rolnictwa przypisać należy podejściom wywodzącym się z poznania budowy i funkcjonowania użytecznych genów, a ostatnio całych genomów roślin uprawnych [6].

Ostatnia dekada XX wieku minęła w naukach przyrodniczych pod znakiem przełomu wywołanego niespotykanym rozwojem metod sekwencjonowania genomów. Wyłoniła się nowa dziedzina empiryczna — genomika, łącząca możliwości metodyczne takich nauk podstawowych jak biologia molekularna, genetyka, chemia organiczna, chemia fizyczna, matematyka czy informatyka. W wyniku skupienia celów poznawczych każdej z tych dziedzin na opisanu materii dziedziczenia — DNA powstała nowa, zintegrowana dziedzina przyrodnicza — genomika, która przez najbliższe lata będzie odgrywała zapewne istotną rolę w rozwoju większości kierunków podstawowych i stosowanych [7, 8].

Genomika jest zespoleniem technologii umożliwiających analizowanie tysięcy genów jednocześnie dla wyczerpującego charakteryzowania stanów komórkowych żywych organizmów. Jest ona połączeniem empirycznych i teoretycznych koncepcji opisywania przyrody dla rozwiązania zasadniczych kwestii dotyczących budowy oraz funkcjonowania genów oraz całych genomów.

Jednym z podstawowych podejść stosowanych przez nową dyscyplinę jest analiza porównawcza pomiędzy genomami różnych, często ewolucyjnie oddalonych organizmów. Poprzez rozpoznanie analogii strukturalnych między homologicznymi genami różnego pochodzenia można wnioskować o funkcji tych jednostek dziedziczenia, których budowa nie została jeszcze w pełni poznana. Stąd też tak dużą rolę odgrywa zsekwencjonowanie genomów wybranych organizmów modelowych.

Pierwszym organizmem żywym, którego genom został poznany, była pasożytnicza bakteria *Haemophilus influenzae* [9]. Obecnie, w sześć lat później po opublikowaniu tej sekwencji, banki danych zawierają kompletne sekwencje nukleotydowe ponad 60 genomów, w tym szeregu organizmów wyższych, łącznie z genomem człowieka i rośliną wyższą. Rośliną tą jest niewielki chwast *Arabidopsis thaliana*, który podnie-

siony został przez genetyków do roli podstawowego modelu roślin okrytozalążkowych [10]. Genom *Arabidopsis*, którego sekwencja została opublikowana w grudniu 2000 r., należy do najmniejszych wśród roślin wyższych i jest zbudowany z 130 mln par zasad nukleotydowych (pz) rozmieszczonych na pięciu chromosomach [11]. Jest on 5-krotnie większy niż genom jednokomórkowych drożdży i 15-krotnie większy od genomu *E. coli*. Genom muszki owocowej *Drosophila melanogaster* jest dwa razy większy, zaś podobnej wielkości co *Arabidopsis* jest genom liczącego 3 tys. komórek nicienia *Caenorhabditis elegans*. Dla porównania — genom człowieka ma 3 mld (3×10^9) pz. Ogromną zaletą *Arabidopsis* jako rośliny modelowej jest jej duża podatność na transformację i krótki czas regeneracji wynoszący 5–6 tygodni. Umożliwiło to uzyskanie wielu tysięcy genetycznie zdefiniowanych mutantów punktowych, dzięki którym możliwe jest przypisywanie poszczególnym genom określonej funkcji [12].

Niewielki genom *Arabidopsis* jest wynikiem długotrwałej presji selekcyjnej faworyzującej ewolucyjnie „ekonomizowanie” zawartości jądra komórkowego: 80% jego puli DNA koduje białka występujące w jednej lub kilku kopiach. Jedynie 20% DNA stanowią sekwencje powtarzające się. Ocenia się, że 50–60% genów każdej rośliny można zidentyfikować poprzez porównanie i odnalezienie homologii z sekwencjami *Arabidopsis* lub z DNA innych organizmów zapisanymi w bazach danych. Skatalogowanie genów uczestniczących w procesach o podobnej funkcji metabolicznej u różnych organizmów może wyjaśnić ogólne prawa różnicowania i rozwoju.

Drugim po *Arabidopsis* genomem modelowym roślin wyższych jest zawierający 3,5 razy więcej DNA genom ryżu *Oryza sativa* [13], który z kolei jest 5-krotnie mniejszy od genomu kukurydzy i ponad 30-krotnie mniejszy od genomu pszenicy. Zsekwencjonowanie genomu ryżu, które jest już daleko zaawansowane, ma duże znaczenie, ponieważ ryż stanowi podstawowe źródło pożywienia dla 1/4 populacji ludzkiej, ale także dlatego, iż posiada on jeden z najmniejszych genomów roślin zbożowych. Jego poznanie umożliwi zidentyfikowanie drogą analiz porównawczych wiele ważnych genów u innych zbóż i jednoliściennych roślin użytkowych o ogromnych genomach, których analiza jest zaawansowana w mniejszym stopniu [14]. Podobna sytuacja dotyczy roślin motylkowatych, dla których wybrane zostały dwa modelowe gatunki: *Medicago truncatula* oraz *Lotus japonicus*. Prowadzone badania sekwencyjne na genomach tych roślin pozwolą na charakteryzowanie innych gatunków drogą analiz porównawczych [15, 16, 17].

Użytecznym narzędziem dla rozpoznawczego porównywania genomów oraz charakteryzowania genów indywidualnych okazały się zbiory (biblioteki) fragmentów sekwencji kodujących EST (Estimated Sequence Tags). Sekwencje te stanowią krótkie (200–600 pz) fragmenty regionów kodujących i niekodujących, z wyłączeniem fragmentów będących powtórzeniami najbardziej liczebnych genów. Wykorzystywane są one jako markery genetyczne dla sporządzania map fizycznych, a także dla uproszczonej klasyfikacji funkcjonalnych jednostek genomów [18, 19]. Obecnie ok.

55% wszystkich genów znajdujących się w genomie *Arabidopsis* posiada już rozpoznaną funkcję. Wśród nich zaskakująco duża wydaje się liczba sekwencji kodujących białka transkrypcyjne (8% wszystkich sekwencji), co pośrednio może wskazywać na pierwszorzędowe znaczenie aparatu regulacji ekspresji genów.

Dużego znaczenia nabrała w ostatnich latach analiza zbiorów sekwencji charakteryzujących określone tkanki i organy roślin. Organospecyficzna ekspresja genów odniesiona do omnipotentnego zestawu wszystkich genów w każdej komórce jest u roślin zjawiskiem bardzo znamienym. Na tym tle ważne wydaje się zrozumienie natury wyłączenia genów w trakcie cyklu życiowego rośliny. Jak wiadomo, ekspresja genów zachodzi w wyniku aktywacji szlaku informacji genetycznej. Większość wszakże genów ulega aktywacji jedynie na ściśle określony czas w zdefiniowanej fazie rozwoju. Regulacja tych procesów ma zasadnicze znaczenie dla rozwoju i różnicowania każdego organizmu [20]. Może ona obejmować zarówno mechanizmy transkrypcyjne, jak i posttranskrypcyjne, modulowanie aktywności polimeraz RNA, jak i reakcje metylacji elementów genomu czy deacylację białek histonowych [21]. Rozpoznanie mechanizmów wyłączających ekspresję genów może być ważnym elementem dla opanowania kontrolowanej ekspresji genów po ich wprowadzeniu do genomu rośliny użytkowej.

Jedną z doniosłych zdobyczy biologii molekularnej roślin dwu ostatnich dekad było opanowanie metodyki trwałego modyfikowania zasobów genowych roślin poprzez wprowadzenie do ich genomów wybranych cząsteczek DNA kodujących określone cechy użytkowe. W ten sposób rośliny transgeniczne można było wykorzystać jako bioreaktory dla wytwarzania użytecznych białek rekombinowanych [22]. Ten ważny i przyszłościowy kierunek agrobiotechnologii może odegrać znaczącą rolę dla produkcji białek i peptydów wykorzystywanych w diagnostyce i medycynie.

Obecnie istnieją dwie zasadnicze strategie pozyskiwania w roślinach produktów transgenicznych o znaczeniu praktycznym: i) technologia oparta na wektorach wirusowych typu RNA, ii) transformowanie bezpośrednio genomu roślinnego obcymi DNA przy wykorzystaniu plazmidów bakteryjnych *Agrobacterium tumefaciens*.

Wykorzystanie wirusów roślinnych (TMV, AMV) do produkowania w komórkach roślinnych peptydów i białek opiera się na spostrzeżeniu, iż subgenomowy RNA, kodujący białko płaszcz wirusa, można wykorzystać jako dogodny nośnik do produkcjiżądanego produktu peptydowego [23, 24]. Poprzez namnażanie zrekombinowanego wirusa uzyskuje się wiele kopii zrekombinowanego białka płaszcz wirusa wydłużonego o polipeptyd docelowy, który można następnie wyodrębnić z roślinnego bioreaktora. Tą drogą uzyskano już na skalę preparatywną szereg cennych produktów białkowych, które znalazły szerokie zastosowania praktyczne [25].

Wykorzystanie wirusów roślinnych dla wytwarzania użytecznych białek w roślinnych bioreaktorach obarczone jest jednak pewnymi ograniczeniami, do których zaliczyć trzeba niestabilność wektorów rekombinowanych oraz patogenny wpływ wektorowego wirusa na roślinę. Oczekuje się, że udoskonalenie wektorów wirusowych, a

zwłaszcza wykorzystanie mieszanych wektorów pochodzących od dwóch różnych wirusów, może ograniczenia te przewyciężyć.

Drugi typ strategii pozyskiwania rekombinowanych peptydów i białek z roślin transgenicznych polega na bezpośredniej transformacji genomu rośliny biorcy za pomocą wektora plazmidowego *A. tumefaciens*, do którego wklonowana zostaje obca sekwencja DNA, kodująca produkt docelowy. Sekwencja ta zostaje losowo wprowadzona w jedno lub więcej miejsc genomu jądrowego rośliny. Jest to obecnie najpowszechniej stosowany model transformacji roślin. Istotnym elementem wektora używanego do transformowania, zapewniającym wyrażanie nadproduktowanego białka w roślinie transgenicznej, jest sekwencja promotorowa, decydująca o poziomie i miejscu ekspresji danego produktu. Typem promotora zapewniającym ekspresję konstytutywną (tzn. we wszystkich organach rośliny) rekombinowanego białka jest promotor 35S — fragment genomu wirusa mozaiki kalafiora CAMV. Z kolei dla zapewnienia ekspresji organospecyficznej transgenów wykorzystywane są regulatorowe sekwencje DNA charakterystyczne na przykład dla białek zapasowych, które ulegają ekspresji w nasionach roślin. Wykorzystywane są także sekwencje „adresujące”, umożliwiające przeniesienie produktów transgenów do wakuoli, retikulum endoplazmatycznego czy plastydów.

Ostatnie lata przyniosły interesujący rozwój koncepcji szerokiego wykorzystywania roślin transgenicznych. Należą do nich próby wyrażania w komórkach roślinnych immunogennych antygenów wirusowego pochodzenia celem otrzymania tzw. szczepionek jadalnych [26, 27, 28]. Szczepionki takie podawane drogą pokarmową wprowadzałyby pewną ilość immunogenego antygeny dla aktywacji układu odpornościowego człowieka lub zwierzęcia. Już pierwsze próby przeprowadzone na początku lat dziewięćdziesiątych wykazały, że nowe formy szczepionek roślinnych mogą w przyszłości stać się niezwykle użyteczne w immunoprewencji i zwalczaniu groźnych chorób wywoływanych przez rozmaite patogeny. Do zalet szczepionek roślinnych nowej generacji zaliczyć można niski koszt, możliwość wytwarzania w dużej skali, eliminację możliwych odczynów alergicznych, a także łatwy sposób rozprowadzania i aplikacji.

Obecnie znanych jest szereg obiecujących przykładów ekspresji wirusowych antygenów w roślinach transgenicznych, które po podaniu drogą pokarmową zdolne są do indukowania odpowiedzi immunologicznej u ludzi i u zwierząt. Duże zainteresowanie budzą w tym względzie prace nad wytworzeniem szczepionki roślinnej przeciwko wirusowemu zapaleniu wątroby typu B. Prace te są prowadzone w kilku laboratoriach i z uwagi na duże zainteresowanie społeczne śledzone są z dużym zainteresowaniem. Według raportu WHO z roku 1998, na świecie żyje obecnie ponad 2 mld nosicieli wirusa HBV. Pionierskie prace nad uzyskaniem ekspresji jednego z antygenów tego wirusa — białka powierzchniowego HbsAg — pochodzą z pracowni Arntzena. Wykazał on, że w liściach tytoniu (*Nicotiana tabacum*) mogą tworzyć się wirusopodobne agregaty (VLP) z rekombinowanych podjednostek białka S otoczki wirusa

HBV, podobne do agregatów VLP z drożdży, wykorzystywanych w szczepionkach komercyjnych [27]. Następnym krokiem było wykazanie, że ekstrakt roślin transgeniczných, zawierający rekombinowany antygen HBsAg, indukował odpowiedź immunologiczną u myszy po podaniu domięśniowym. Ważnym etapem na drodze do otrzymania efektywnej szczepionki roślinnej było przeprowadzenie pierwszych testów na ochotnikach spożywających transgeniczną roślinę, wytwarzającą antygen HBsAg. Wykazały one możliwość indukcji tą drogą ogólnoustrojowej odpowiedzi immunologicznej u ludzi [30, 31]. Dalsze prace prowadzone są w kierunku wyznaczenia optymalnej dozy antygeny, a także określenia roli i doboru adiuwanty. Można oczekiwać, że kontynuowanie intensywnych badań przedklinicznych i testów klinicznych doprowadzi w ciągu 5–8 lat do pojawienia się na rynku medycznym szczepionek roślinnych nowej generacji, które będą szczególnie przydatne w tropikalnych krajach rozwijających się [32, 33].

Dotychczasowe osiągnięcia biotechnologii molekularnej roślin ukazały, że w roślinach transgeniczných i w hodowlach komórkowych możliwe jest wytwarzanie białek czynnych o znaczeniu terapeutycznym i diagnostycznym. „Zielone fabryki” mogą odegrać wielką rolę w przewyżnianiu ograniczeń wynikających z kosztocłonności i energocłonności biotechnologii klasycznych. Zależać to jednak będzie zarówno od wypracowania oczekiwanego postępu w udoskonalaniu badanych układów, jak i też od publicznej akceptacji nowych technologii opartych na rekombinowanych cząsteczkach DNA [34]. Panuje na ogół powszechne przekonanie, że nowe biotechnologie mogą odegrać istotną rolę w pozyskiwaniu dobrej jakościowo żywności w warunkach pogłębiających się globalnych zmian środowiska naturalnego. Mogą one także odegrać decydującą rolę w zachowaniu różnorodności biologicznej przyrody ożywionej, co jest jednym z podstawowych warunków zachowania równowagi ekosystemowej na Ziemi.

Literatura

- [1] Meyerowitz E.M. 1999. The logic of development. *Trends in Biol. Sci.* 24: 968–1004.
- [2] Phillips R.L., Freeling M. 1998. *Proc. Natl. Acad. Sci. USA* 95: 1969–1970.
- [3] Miflin B.J. 2000. *Plant Physiology* 123: 17–27.
- [4] Wambugu F. 1999. *Nature* 400: 15–16.
- [5] Miflin B.J. 2000. *J. Exp. Bot.* 51: 1–8.
- [6] Miklos G.L.G., Rubin G.M. 1966. *Cell* 86: 521–529.
- [7] Pandej A., Lewitter F. 1999. *Trends in Biol. Sci.* 24: 276–280.
- [8] Schuler G.D. et al. 1996. *Science* 274: 540–546.
- [9] Fleischmann R.D. et al. 1995. *Science* 269: 496–512.
- [10] Meinke D.W., Cherry J.M., Dean C., Rounsley S.D., Koornneef M. 1998. *Science* 282: 662–682.

- [11] Dennis C., Surridge C. 2000. *Nature* 408: 791.
- [12] Pruitt R. E., Meyerowitz E.M. 1986. *J. Mol. Biol.* 187: 169–183.
- [13] Sasaki T. et al. 1996. *Genome Res.* 6: 661–666.
- [14] Bevan M., Murphy G. 1999. *Trends in Genetics* 15: 211–214.
- [15] Gale M.D., Devos K.M. 1998. *Science* 282: 656–659.
- [16] Pennisi E. 2000. *Science* 290: 32–35.
- [17] Keller B., Feuillet C. 2000. *Trends in Plant Science* 5: 246–251.
- [18] Bennetzen J.L., SanMiguel P., Chen M., Tikhonov A., Francki M., Avramova Z. 1998. *Proc. Natl. Acad. Sci. USA* 95: 1975–1978.
- [19] Lottspeich F. 1999. *Angew. Chem. Int. Ed.* 38: 2476–2492.
- [20] Velculescu V.E., Vogelstein B., Kinzler K.W. 2000. *Trends in Genetics* 16: 423–425.
- [21] Iyer L.M., Kumpatla S.P., Chandrasekharan M.B., Hall T.C. 2000 *Plant Mol. Biol.* 43: 323–346.
- [22] Steplewski K.M. 1998. Ph.D. Dissertation „Plant viruses as vectors for the expression of antigens from rabies virus and HIV-1”. Philadelphia.
- [23] Moffat A.S. 1995. *Science* 265: 658–663.
- [24] Modelska A., Dietzchold B., Fleysh N., Fu Z., Steplewski K., Hooper D., Koprowski H., Yusibov V. 1997. *Proc. Natl. Acad. Sci. USA* 95: 2481–2485.
- [25] Dalsgaard K., Uttenthal A., Jones T.D., Xu F., Merryweather A., Hamilton W. D.O., Langeveld J.P.M., Boshuizen R.S., Kamstrup S., Lomonossoff G. P., Porta C., Vela C., Casal J.I., Meloen R.H., Rodgers P.B. 1997. *Nature Biotechnol.* 15: 248–252.
- [26] Verch T., Yusibov V., Koprowski H. 1998. *J. Immunological Methods* 220: 69–75.
- [27] Mason H.S., Lam D.M.-K., Arntzen C.J. 1992. *Proc. Natl. Acad. Sci. USA* 89: 11745–11749.
- [28] Yusibov V., Modelska A., Steplewski K., Agadjanyan M., Weiner D., Hooper D.C., Koprowski H. 1997. *Proc. Natl. Acad. Sci. USA* 94: 5784–5788.
- [29] Thanavala Y., Yang Y.F., Lyons P., Mason H.S., Arntzen C.J. 1995. *Proc. Natl. Acad. Sci. USA* 92: 3358–3361.
- [30] Kapusta J., Modelska A., Figlerowicz M., Pniewski T., Letellier M., Lisowa O., Yusibov V., Koprowski H., Plucienniczak A., Legocki A.B. 1999. *The FASEB Journal* 13: 1796–1799.
- [31] Kapusta J., Modelska A., Pniewski T., Figlerowicz M., Jankowski K., Lisowa O., Plucienniczak A., Koprowski H., Legocki A.B. 2001. *Adv. in Exp. Medicine and Biology. Progress in Basic and Clinical Immunology.* A. Mackiewicz Ed. in press.
- [32] Arakawa T., Chong D.K.X., Langridge W.H.R. 1997. *Nature Biotechnology* 16: 292–297.
- [33] Walmsley A. M., Arntzen C. J. 2000. *Current Opinion in Biotechnology* 11: 126–129.
- [34] Flavell R. B. 2000. *Current Opinion in Plant Biology* 3: 143–146.

Integrated genomics as a basis for novel agrobiotechnologies

Key words: plant genomics, proteomics, gene silencing, plant expression systems

Summary

Over last two decades new highly advanced genomic approaches have created unprecedented progress in understanding the structure-function relationship between nucleic acids and proteins across diverse organisms. The sequencing data combined with computation analysis have fundamentally changed the prospects of modern agrobiotechnology. This led to the development of novel strategies for plant protection, increasing crop yield and producing in plants some pharmaceutical compounds including those used for immunoprotection of human and animals. Plant comparative genomics is based on selected model plants such as *Arabidopsis thaliana*, rice or some legumes. The emerging fields such as functional genomics and proteomics address the questions of functional significance of individual genes for plant development under various environmental conditions.