



Mikromacierze DNA i ich zastosowanie w badaniach ekspresji genów u roślin

Agnieszka Ludwików, Krzysztof Krasowski, Lucyna H. Misztal,
Jan Sadowski

Zakład Biotechnologii, Instytut Biologii Molekularnej i Biotechnologii,
Wydział Biologii, Uniwersytet im. Adama Mickiewicza, Poznań

Application of microarray technology in plant genes expression

Summary

A major progress has been made in the last decade in the understanding of molecular mechanisms involved in plant responses to stresses. This has been achieved by application of a high-throughput genomics tools. Transcriptome analysis based on the DNA microarray technique has proved its effectiveness in studies of genome level responses to environmental changes and genetic manipulations. From the perspective of a decade, DNA microarrays serve as an invaluable tool for global gene expression analyses, unravelling new information about genes, pathways and their transcriptional regulation networks. In this review, we present applications of the DNA microarrays and a basic analysis of microarray data in relation to the transcriptional response of plants to ozone and drought stresses.

Key words:

DNA microarray technology, transcriptome profiling, environmental stresses.

Adres do korespondencji

Jan Sadowski,
Zakład Biotechnologii,
Instytut Biologii
Molekularnej
i Biotechnologii,
Wydział Biologii,
Uniwersytet
im. Adama Mickiewicza,
ul. Umultowska 89,
61-614 Poznań;
e-mail: jsad@amu.edu.pl

1. Wprowadzenie

W każdym żywym organizmie geny oraz kodowane przez nie białka funkcjonują w niezwykle złożony, ściśle zorganizowany sposób. Liczne sieci oddziaływań i ich regulacja będąca pod wpływem nawet nieznacznie zmieniających się czynników genetycznych, jak i środowiskowych, wymagają od badacza wprowadzenia właściwych metod badawczych pozwalających na efektywną

obserwację i jednocześnie zbieranie wielu danych eksperymentalnych. Niezbędnym etapem tych prac jest gromadzenie i szybkie przetwarzanie informacji, co prowadzi do szerokiej interpretacji wyników. Jednym z pierwszych narzędzi genomiki funkcjonalnej w masowej analizie ekspresji genów była mikromacierz DNA (1). Pozwala ona na analizę ekspresji tysięcy genów dla wielu prób biologicznych jednocześnie. W ostatniej dekadzie technologia mikromacierzy DNA stała się ważnym narzędziem badawczym ze względu na prostotę i masowość analizy. W wielu przypadkach, dotychczas skonstruowane mikromacierze DNA stwarzają możliwość analizy organizmów blisko spokrewnionych o nie znanych jeszcze sekwencjach genomowych. Obecnie, mikromacierze DNA wykorzystuje się w wielu dziedzinach biologii eksperymentalnej, począwszy od taksonomii, a skończywszy na diagnostyce medycznej. Ich znaczenie wzrosło wraz z dopracowaniem metodycznym kilku elementów np. amplifikacji DNA przeznaczonego do nadruku, warunków hybrydyzacji, znakowania DNA i skanowania obrazu mikromacierzowego, które zapewnia powtarzalność procesu obróbki danych z mikromacierzy. Obecnie w eksperymentach, w których wykorzystuje się technikę mikromacierzy DNA, generuje się powtarzalne dane o wysokiej jakości, stanowiące cenną informację biologiczną. Stosowane protokoły dotyczą dwóch głównych typów mikromacierzy DNA: mikromacierzy cDNA oraz tzw. chipów genowych produkowanych przez firmy Affymetrix i Agilent.

Alternatywą dla stosunkowo drogiej mikromacierzy są makromacierze na membranach nylonowych, przygotowywane bezpośrednio przez zainteresowane laboratoria. Ze względu jednak na nieperspektywiczność makromacierzy, wynikającą z niemożności zautomatyzowania konstrukcji i analizy, podejście to nie będzie omawiane w tym opracowaniu.

Rozwojowi techniki mikromacierzy DNA towarzyszy także udoskonalanie narzędzi wykorzystywanych w jej przygotowaniu i analizie wyników. Instrumenty używane do nadruku sond, modyfikacje podłoży i rodzaje opłaszczeń płytki, udoskonalanie skanerów to przykład „małej rewolucji” myśli technicznej. Tworzenie komercyjnych i publicznie dostępnych programów, skryptów oraz baz danych dla eksperymentów mikromacierzowych ma na celu gromadzenie szczegółowych informacji. Wreszcie, doskonalenie metod weryfikujących proces obróbki mikromacierzy DNA również przyczyniło się do prawidłowej analizy i interpretacji eksperymentu transkryptomicznego.

Uniwersalność mikromacierzy DNA odzwierciedla się w ich szerokim zastosowaniu do celów innych niż profilowanie ekspresji genów (2-4). Aktualnie, mikromacierze DNA wykorzystuje się do genotypowania, określania liczby kopii genów, mapowania, sekwencjonowania oraz badania oddziaływań między cząsteczkami (5-7). Technologia ta umożliwia nie tylko kontynuację szczegółowych badań nad strukturą genomu, ale przede wszystkim badanie funkcji genów (8,9). W Polsce dostępne są obecnie cztery platformy oferujące wykonanie analizy transkryptomicznej z wykorzystaniem mikromacierzy DNA (Instytut Onkologii w Gliwicach, Instytut Chemii Bioorganicznej PAN w Poznaniu, oraz Instytut Biochemii i Biofizyki PAN w Warsza-

wie) i masowego PCR w czasie rzeczywistym (Instytut Biologii Molekularnej i Biotechnologii UAM w Poznaniu).

Pomimo wielu możliwości jakie stwarza użycie mikromacierzy DNA, ich wykorzystanie do profilowania ekspresji genów wciąż przeważa w porównaniu do pozostałych zastosowań. W tej pracy przedstawimy wybrane aspekty projektowania doświadczenia mikromacierzowego i analizy danych, oraz podsumujemy postępy w transkryptomice stresu u roślin, by zademonstrować korzyści, jakie płyną z zastosowania masowej analizy ekspresji genów w badaniach dotyczących odpowiedzi roślin na stesy abiotyczne takie jak susza i podwyższony poziom ozonu w środowisku.

2. Rodzaje mikromacierzy DNA i ich zastosowanie

W rozwoju technologii mikromacierzowej wyróżnić można dwa podstawowe systemy: mikromacierze cDNA i mikromacierze oligonukleotydowe. Te ostatnie, ze względu na technikę nanoszenia sond DNA na płytkę, dzielą się na mikromacierze „drukowane” i syntezowane *in situ*. Przykładem tych mikromacierzy są chipy genowe firmy Affymetrix lub mikromacierze oligonukleotydowe firmy Agilent. Chociaż wszystkie wymienione typy mikromacierzy DNA są używane do analiz transkryptomicznych, tj. analizy ekspresji genów, przedstawione warianty są zasadniczo różne. W przypadku mikromacierzy cDNA, nadrukowywane są stosunkowo długie fragmenty DNA (EST). Z technicznego punktu widzenia, najistotniejszymi cechami odróżniającymi je od mikromacierzy oligonukleotydowych jest wspomniana długość sond i znacznie mniejsza ich liczba na jednej płytce. Do jej zalet należy zaliczyć możliwość dowolnego ukierunkowania badań, np. poprzez wybór mniejszych zestawów genów dla więcej niż jednego gatunku oraz ich zastosowanie w badaniach gatunków roślin o słabo poznanym genomie. Wybór sond genowych do tego typu mikromacierzy dostosowywany jest do celu badań i opiera się głównie na dostępnych danych literaturowych. Istotnym elementem projektowania takiej płytki jest uwzględnienie kontroli hybrydyzacji i czystości materiału. Wadą tego podejścia jest eliminacja genów, których funkcja nie została poznana.

Mikromacierze oligonukleotydowe są przygotowywane poprzez syntezę *in situ* (Affymetrix) lub nadruk konwencjonalnie zsyntetyzowanych oligonukleotydów. Ten typ mikromacierzy DNA, poza analizą ekspresji genów, znalazł zastosowanie przy wykrywaniu mutacji, mapowaniu oraz analizie ekspresji w obrębie rodzin genowych.

Projektowanie oligonukleotydów do analiz mikromacierzowych jest zadaniem wymagającym znajomości wszystkich sekwencji kodujących badanego organizmu. Produkcja różnych mikromacierzy zwiększa się wraz z postępem prac nad sekwencjonowaniem nowych genomów i/lub kolekcji klonów cDNA. Firma Affymetrix udostępnia obecnie w swojej ofercie chipy DNA z nadrukowanymi oligomerami o długości 25 nukleotydów dla 11 różnych gatunków roślin (tab.). Dostępne chipy firmy

Affymetrix dla cytrusów, lucerny oraz soi oprócz 25-nukleotydowych sekwencji sond genowych, zawierają również sekwencje patogenów specyficznych dla danego gatunku.

Tabela

Lista chipów DNA Affymetrix służących do analiz ekspresji genów

| Gatunek | Chip | Liczba transkryptów na chipie | Liczba dostępnych ESTów | Liczba poznanych genów |
|-------------|-------------------------------|-------------------------------|-------------------------|------------------------|
| | | | www.ncbi.nlm.nih.gov | |
| rzodkiewnik | Arabidopsis ATH1 Genome Array | 24k | 1 526 133 | ** 32 029 |
| ryż | Rice Genome Array | ~ 49k | 1 220 261 | ** 29 424 |
| cytrusy | Citrus Genome Array | ~ 33k | 19 294 | 142 |
| bawełna | Cotton Genome Array | ~ 21k | 265 833 | 131 |
| kukurydza | Maize Genome Array | ~ 14k | 1 462 612 | 2019 |
| lucerna | Medicago Genome Array | ~ 49k | 261 522 | 114 |
| topola | Poplar Genome Array | ~ 56k | 409 702 | 275 |
| soja | Soybean Genome Array | ~ 35k | 394 370 | 1245 |
| pomidor | Tomato Genome Array | ~ 9k | 258 943 | 1195 |
| winogrono | Vitis vinifera (Grape) Array | ~ 15,5k | 352 984 | 140 |
| pszenica | Wheat Genome Array | ~ 54k | 1 050 414 | 1492 |

**znana sekwencja genomu

Czułość oraz powtarzalność eksperymentów z udziałem mikromacierzy oligonukleotydowych można zwiększyć stosując oligomery dłuższe niż 25 nukleotydów. Zapewnia to specyficzność hybrydyzacji w obrębie rodzin genowych zmniejszając ryzyko wystąpienia hybrydyzacji krzyżowych. Wychodząc tym wymaganiom naprzeciw, firma Agilent zaproponowała rozwiązanie w postaci 60-nukleotydowych sond dla dwóch gatunków: *Arabidopsis thaliana* (na jednej mikromacierzy łącznie 40 000 sekwencji genów) i ryżu (łącznie ponad 42 000 sekwencji).

Oligonukleotydowe mikromacierze DNA stanowią cenną alternatywę metodyczną w identyfikacji i analizie ekspresji transkryptów powstających w procesie alternatywnego splicingu. Obecnie stosuje się mikromacierze DNA dla kilku różnych gatunków łącznie: Made-to-Order Arrays (Affymetrix) oraz Custom Gene Expression Microarrays (Agilent). Ponadto, międzynarodowe inicjatywy badawcze tworzą nowe rozwiązania w dziedzinie mikromacierzy. Dostępne są na przykład płytki mikromacierzowe zawierające dłuższe oligonukleotydy (70 pz), które konstruowane były w ramach projektów badawczych dotyczących genomu ryżu „Rice Oligonucleotide

Array Project” (www.ricearray.org) i kukurydzy „Maize Oligonucleotide Array Project” (www.maizearray.org).

Najnowszym osiągnięciem w dziedzinie mikromacierzy DNA są tzw. nanochipy (10). Matrycą może tu być zarówno twarde podłoże, do którego przyczepia się tysiące próbek DNA, jak również dwuniciowy DNA z bakteriofaga M13, który zawiera sekwencje komplementarne do RNA docelowego genu. Nanomacierz można wykorzystać do detekcji nawet pojedynczych cząsteczek RNA (www.biotechnology.net).

Zakres możliwości poznawczych mikromacierzy DNA nie musi ograniczać się do analiz ekspresji sekwencji kodujących białka. Mikromacierze z sekwencjami reprezentującymi miliony różnych fragmentów genomu, wraz z ich wariantami, zastosowano w metodzie zwanej Whole Genom Tiling Array. Tutaj sondy w postaci krótkich 25-70-nukleotydowych oligomerów reprezentują niepowtarzające się sekwencje DNA całego genomu. Tiling microarray pozwala nie tylko pełniej scharakteryzować transkryptom i identyfikować nowe geny, ale także badać ich różne izoformy splicingowe, rolę metylacji DNA w regulacji ekspresji genów, określać precyzyjnie miejsca wiązania czynników transkrypcyjnych (11) oraz różnorodność sekwencji regulatorowych w promotorach (12). Obecnie, dostępne są kolejne tego typu mikromacierze dla genomu *Arabidopsis thaliana*: GeneChip® *Arabidopsis* Tiling 1.0R Array (Affymetrix) oraz Custom Tiling Microarrays (Agilent). Niewątpliwie, technika „tiling” wnosi znaczny postęp do badań nad regulacją procesów biologicznych na poziomie RNA.

3. Planowanie eksperymentu z użyciem mikromacierzy DNA

Zasady, które powinny być przestrzegane podczas przygotowywania eksperymentu z użyciem mikromacierzy DNA, a także na etapie weryfikacji i interpretacji wyników, oceny powtarzalności, oraz sposoby dokumentowania zawarte są w formacie MIAME (ang. *Minimum Information About Microarray Experiment*) (13). Format ten wyznacza przejrzyste reguły planowania i przygotowywania eksperymentu mikromacierzowego, a rygorystyczne przestrzeganie tych wytycznych pomoże w wyjaśnieniu postawionych problemów badawczych i hipotez.

Jednym z czynników determinujących wybór platformy mikromacierzowej jest ilość materiału biologicznego do badań; z nim wiąże się także zagadnienie liczby i jakości stosowanych powtórzeń hybrydyzacji w eksperymencie. Ilość wyjściowego RNA używana do pojedynczej hybrydyzacji jest uwarunkowana m. in. sposobem jego izolacji i właściwościami badanej tkanki. Stosowanie powtórzeń hybrydyzacji, poza względami statystycznymi, ma na celu zwiększenie detekcji poziomu zmian – im więcej replikacji, tym niższe krotności wzrostu/spadku transkrypcji będziemy w stanie zweryfikować. Liczba powtórzeń musi spełniać wymagania statystyki, a ich charakter (powtórzenia biologiczne i techniczne) może wynikać ze specyfiki eksperymentu, np. charakteru zastosowanego czynnika stresowego, natężenia obserwo-

wanego efektu, zmienności osobniczej, a także założonego % wyników fałszywie pozytywnych. Ponieważ wpływ tych czynników na jakość wyników nie jest określony, sugeruje się przeprowadzać kilkukrotne powtórzenia hybrydyzacji dla każdego wariantu doświadczenia (14).

Z uwagi na matematyczne transformacje, którym są poddawane bezwzględne wartości sygnału uzyskane w wyniku hybrydyzacji uważa się, że mikromacierz jest techniką półilościową. Dlatego istotnym elementem analiz z użyciem mikromacieczy DNA jest weryfikacja poziomu sygnału otrzymanego w procesie normalizacji. Analiza ta wymaga zaadaptowania metod ilościowych określających poziom ekspresji genu, którą aktualnie jest metoda PCR w czasie rzeczywistym. Porównuje się wartości poziomu zmian ekspresji dla wybranych genów obserwowane w analizie danych pochodzących z mikromacieczy i otrzymanych alternatywną metodą (15). Zauważono, że poziom zmian uzyskany w wyniku analizy mikromacieczowej jest niższy w porównaniu do eksperymentów kontrolnych przeprowadzonych przy użyciu technik ilościowych. Prawidłowość taka jest uzasadniona wyłącznie dla genów o określonym natężeniu sygnału i jest związana ze stosowaną procedurą normalizacji. W analizie wykorzystuje się również założone minimalne wartości wzrostu lub spadku poziomu sygnału jako dodatkowy wyznacznik prawdopodobieństwa wykrytej zmiany w poziomie transkrypcji. Zazwyczaj za istotną i powtarzalną zmianę przyjmuje się 2-krotny wzrost lub zahamowanie ekspresji, bez względu na poziom sygnału. Jednocześnie jest to najniższa przyjmowana wartość wzrostu/spadku transkrypcji zmierzona w eksperymentach kontrolnych, wynikająca z rodzaju stosowanych narzędzi pomiarowych i algorytmów przetwarzających dane ilościowe.

4. Obróbka i analiza danych mikromacieczowych

Zmienność intensywności sygnału obserwowana po skanowaniu obrazu mikromacieczy DNA wynika z technicznej obróbki mikromacieczy (tzw. błąd systematyczny), a nie biologicznej różnorodności prób. Dlatego najważniejszym etapem doświadczenia z użyciem mikromacieczy DNA jest właściwa normalizacja poziomu sygnałów otrzymanych bezpośrednio po hybrydyzacji. Normalizacja danych stanowi główny etap poprzedzający właściwą analizę profilu ekspresji genów. Usuwa zmienność techniczną i umożliwia porównywanie ekspresji w relatywnej skali, podkreślając w ten sposób podobieństwo wzorów ekspresji. Dlatego postęp techniczny w przygotowaniu i obróbce mikromacieczy DNA, oraz weryfikacja zmian w poziomie transkrypcji dla genów o niskiej ekspresji, wyznacza kierunek rozwoju metod normalizacji. Od właściwie przeprowadzonej normalizacji zależy jakość gromadzonych danych. Wybór metody normalizacji jest konsekwencją zastosowanego typu mikromacieczy i rodzaju znakowania. Na początku wykonuje się normalizację danych w obrębie jednej mikromacieczy, następnie normalizację między wszystkimi płytkami wchodzącymi w skład eksperymentu. Każda płytka może charakteryzować

się swoistym rozkładem logarytmu współczynnika intensywności sygnału i , na tym etapie, poszczególne hybrydyzacje nie powinny być porównywane między sobą. Porównanie poziomu ekspresji między płytkami jest możliwe dopiero po przeprowadzeniu normalizacji, zwanej skalowaniem. Znanych jest wiele metod normalizacji, m. in. normalizacja globalna (16), normalizacja RMA (17) czy normalizacja Lowess (18,19). Przy wyborze algorytmu istotny jest również fakt, że stosowane algorytmy nie korygują jednoznacznie i porównywalnie sygnałów o bardzo wysokiej i niskiej intensywności. Dlatego każda procedura normalizacji generuje swoisty rozkład intensywności sygnałów, a tym samym inną liczbę genów o zróżnicowanej ekspresji. By zilustrować te zależności, warto przytoczyć porównanie wyników dwóch normalizacji stosowanych w analizach z użyciem chipów firmy Affymetrix – normalizacje dCHIP i RMA (20). Obie normalizacje generowały różny rozkład intensywności sygnałów na mikromacierzach, co bezpośrednio wpływało m. in. na liczbę genów uwzględnianych do dalszych analiz oraz na poziom wzrostu lub spadku ekspresji tych genów. Przeprowadzone porównanie wykazało również, że algorytm dCHIP, w przeciwieństwie do RMA, nie jest tak skuteczny w normalizacji sygnałów o niskim poziomie. Natomiast na podstawie wyników uzyskanych podczas weryfikacji eksperymentu mikromacierzowego wykazano, że sygnały poddane normalizacji RMA posiadały 96,6% zgodności ze zmianami (spadek/wzrost poziomu transkrypcji) odnotowanymi metodą PCR w czasie rzeczywistym. Na tej podstawie uznano, że RMA jest właściwszą metodą normalizacji dla chipów firmy Affymetrix. Chociaż analiza mikromacierzy DNA prowadzi do tworzenia list genów zmieniających poziom transkrypcji w zależności od zastosowanego czynnika, istotniejsze jest, które z tych genów mają decydujący wpływ na obserwowany profil. Do identyfikacji genów o zróżnicowanym poziomie transkrypcji wykorzystuje się obecnie metody statystyczne (15). Służą one do określenia zmiany w poziomie transkrypcji dla jak największej liczby genów przy jak najniższym odsetku wyników fałszywie pozytywnych (14). W związku z faktem, że konwencjonalnie stosowane testy statystyczne przypisują najniższe wartości prawdopodobieństwa genom o bardzo niskiej intensywności sygnału, obserwuje się tendencję do przeprowadzania analiz statystycznych na grupach wstępnie wyselekcjonowanych genów.

Ważną przesłanką do interpretacji wyników w badaniach mikromacierzowych jest określenie związku między poziomem transkryptu a sekwencją genu. To zadanie realizowane jest na kilka sposobów i służy jednemu celowi: poznaniu biologicznego znaczenia informacji, jaka kryje się za obszerną listą genów ulegających zróżnicowanej transkrypcji. Biologiczna interpretacja danych stwarza wiele trudności wynikających z faktu, że produktowi genu można przypisać więcej niż jedną funkcję. Takie uogólnienie informacji o funkcji genów podkreśla złożoność procesów biologicznych w komórce i problematyczność analiz mikromacierzowych. Dlatego analizę funkcjonalną genów w eksperymentach mikromacierzowych rozpoczyna się od klasyfikacji grup genów o zróżnicowanej transkrypcji lub określaniu kategorii funkcjonalnych, które są częściej reprezentowane w analizowanych da-

nych w odniesieniu do ich zestawów kontrolnych. Zaletą tej klasyfikacji jest łatwość budowania w dalszej analizie sieci interakcji (związków przyczynowo-skutkowych) pomiędzy genami (zależność typu gen-gen), kategoriami funkcjonalnymi (zależność typu termin-termin) czy genami a kategoriami funkcjonalnymi (zależność typu gen-termin). W późniejszym etapie prac, opisywane podejście może być pomocne także w tworzeniu list potencjalnych oddziaływań między białkami, określaniu związków między genami a obserwowanym fenotypem, oraz analizie kontroli transkrypcyjnej procesów metabolicznych. Narzędziem, które umożliwi interpretację wyników są analizy matematyczne m. in. analizy skupień (6,21,22), a także bardziej zaawansowane podejścia związane z eksploatacją danych (data mining) i korzystające z metod rozpoznawania wzorców (23). Podejścia te umożliwiają identyfikację grup genów ulegających podobnej ekspresji, tworzących regulony, czyli grupy genów regulowanych wspólnie. Określenie hipotetycznych regulonów umożliwi uporządkowanie obserwacji oraz eliminację nieistotnych kategorii funkcjonalnych. Ponadto, użycie tych metod powinno polepszać wykrywalność wspólnych domen funkcjonalnych (np. domen białkowych) i motywów (np. motywów w promotorach genów), co także podkreśla funkcjonalną współzależność części genów.

W przypadku analiz promotorów genów ulegających podobnej ekspresji, uporządkowanie genów we wspólnie regulowane zestawy sprzyja identyfikacji motywów promotorowych, które przypuszczalnie determinują obserwowane wzory ekspresji. Początkowo, nie uwzględniano złożoności procesu transkrypcji w analizie promotorów. Było to także związane z brakiem zaawansowanych metod i programów. Dlatego analizy polegały głównie na prostej identyfikacji motywów w rejonach promotorowych badanych genów. Aktualnie, metodykę analiz promotorowych dopasowano do istniejących wymagań, uwzględniając złożoną regulację procesu transkrypcji, która wynika ze współdziałania wielu czynników transkrypcyjnych. Porównywanie zidentyfikowanych sekwencji promotorowych pomiędzy zestawami genów o przeciwnym profilu ekspresji, w powiązaniu z ich modelowaniem (dodawanie lub wykluczanie motywów z kombinacji) pomaga określić, jaki zestaw motywów może być odpowiedzialny za obserwowany wzór ekspresji.

5. Informacje uzyskiwane z eksperymentu mikromacierzowego

Nawet w najprostszych systemach biologicznych dochodzi do skoordynowanej aktywacji wielu procesów biologicznych, które są regulowane na trzech poziomach: transkryptomycznym, proteomicznym i metabolicznym. Dlatego, by w pełni zrozumieć dane pochodzące z analizy mikromacierzy DNA należałoby rozpatrywać je w ich kontekście biologicznym. Pomimo faktu, że mikromacierze DNA ujawniają kompleksowość funkcjonowania organizmu, jego fenotyp zaledwie w części określany jest przez transkryptom. Zatem analizy mikromacierzowe, choć klarownie od-

dają stan fizjologiczny (kondycję) organizmu w danym momencie, nie pozwalają na określenie stopnia integracji uzyskanej informacji z pozostałymi poziomami ekspresji genów. Dotychczas, przeniesienie informacji z analiz transkryptomicznych na profil białkowy miał pewne merytoryczne uzasadnienie: regulacja ekspresji białek na poziomie transkrypcji jest o wiele bardziej wydajna w porównaniu z regulacją posttranskrypcyjną. Dlatego informatywność analiz profilu ekspresji genów jest wysoce pomocna w określaniu funkcji genów, identyfikacji mechanizmów regulatorowych i analizie szlaków biochemicznych. Należy pamiętać, że analiza sieci szlaków metabolicznych, bez znajomości ekspresji kontrolujących je genów może prowadzić do błędnych wniosków.

Do profilowania transkrypcji używa się wielu zaawansowanych metod i narzędzi bioinformatycznych. Otrzymany wynik zależy od przyjętych parametrów i wstępnej obróbki danych, w tym głównie zastosowanej procedury normalizacji. Należy zatem pamiętać, że wnioski wyciągane na podstawie badań z użyciem mikromacierzy DNA, szczególnie analiz *in silico*, mają charakter hipotetyczny i wymagają dalszej weryfikacji eksperymentalnej.

6. Transkryptomika stresu u roślin

Analiza profili ekspresji genów jest niezwykle pomocna w identyfikacji funkcji genów i badaniu mechanizmów regulujących ekspresję genów w odpowiedzi na czynniki zewnętrzne. Dlatego metoda mikromacierzy DNA jest szeroko stosowana w badaniach odpowiedzi roślin na stres. Chociaż dane literaturowe dostarczają informacji o profilowaniu transkrypcji dla wszystkich rodzajów stresów, w tym opracowaniu skupimy się na dwóch najważniejszych: stresie ozonowym i suszy, które są od dawna intensywnie badane w kontekście identyfikacji mechanizmów regulujących odpowiedź roślin na stres, także przez nasz zespół.

Stres ozonowy i susza należą do bardzo uciążliwych i szkodliwych czynników środowiska wpływających na wegetację roślin, obniżając produktywność upraw i kondycję obszarów leśnych (24-26). Identyfikacja mechanizmów molekularnych przyczyniających się do wzrostu odporności roślin na oba stresy stała się głównym wyzwaniem dla wielu laboratoriów. Niewątpliwie szybki rozwój technik mikromacierzowych dostarczył nowych informacji dotyczących odpowiedzi i adaptacji roślin do obu warunków stresowych. W tych eksperymentach wykorzystywano szeroki wachlarz dostępnych konstrukcji i platform mikromacierzowych (15,20,27-29). W większości przeprowadzone eksperymenty i analizy mikromacierzowe nie miały na celu prostej identyfikacji genów uruchamianych w odpowiedzi na stres (15,29), ale projektowano je, by poznać sieć zależności między genami i jej regulację (20,29,30). Należy również nadmienić, że w analizach tych pokazano użyteczność dostępnych narzędzi bioinformatycznych do pozyskiwania nowych informacji o zróżnicowaniu transkryptomu w warunkach obu stresów.

Analizy transkryptomyczne stresu ozonowego miały na celu: identyfikację różnic w poziomie transkrypcji genów należących do rodzin wielogenowych; identyfikację genów związanych ze stresem oksydacyjnym (15,28); identyfikację oddziaływań między sygnalizacją kwasu absycynowego (ABA) a sygnalizacją reaktywnych form tlenu (RFT)(20); dotyczyły także oszacowania specyficzności sygnalizacji RFT (31); oraz określenia zmian w profilu ekspresji genów u ekotypów *Arabidopsis* różniących się podatnością na stres ozonowy (32). Analizy te zaowocowały m. in. identyfikacją nowych genów uczestniczących w transkrypcyjnej regulacji metabolizmu roślin w tym stresie, a także ujawniły nowe szlaki regulujące poziom tolerancji roślin na stres ozonowy (20,28).

Analizy transkryptomyczne w stresie suszy są nastawione głównie na identyfikację mechanizmów molekularnych odpowiedzi na stres (20,29) i typowanie genów, których modyfikacje mogą potencjalnie zwiększyć poziom odporności roślin użytkowych (29). W pierwszym przypadku istotnym elementem badań jest analiza dynamiki odpowiedzi na stres i dotyczy ona głównie porównania szybkiej i ostrej odpowiedzi z łagodnym i stopniowym ubytkiem wody podobnie do warunków naturalnych (33). W drugim przypadku mają charakter ogólny i są wielowątkowym opracowaniem profilu transkrypcyjnego rośliny w stresie. Koncepcja ogólności tych analiz spowodowała, że uzyskane wyniki najczęściej dotyczą tej części genów, których zachowanie jest już znane i nie zawsze musi być związane z procesem nabywania tolerancji na stres u roślin uprawnych. Zaledwie mały odsetek tych badań prowadzi do identyfikacji nowych genów, które mogą aktywnie funkcjonować w stresie i były ważne w ewolucji odpowiedzi na stres. Ponadto brak szerokiej kolekcji mutantów roślin uprawnych przyczynia się do braku możliwości połączenia aktywności wytypowanych genów z fenotypem rośliny w stresie. Dlatego dość częstym podejściem w badaniach roślin uprawnych jest analiza porównawcza profilu transkrypcyjnego między odmianą wrażliwą i tolerującą stres suszy (34). Towarzyszące transkryptomice użycie innych technologii genomiki funkcjonalnej, tj. proteomiki i metabolomiki, nie ułatwia tych analiz. Nie pozwalają one również na identyfikację konkretnych produktów genowych czy metabolitów, które są istotne w procesie nabywania tolerancji na stres. Należy pamiętać, że i one służą zaledwie ogólnej charakterystyce poszczególnych etapów ekspresji genów.

Z pewnością poszukiwanie związków między genami a obserwowanym fenotypem, w warunkach jednego lub kilku stresów, wymusza profilowanie transkrypcji w układzie odmiennych, dobrze zdefiniowanych genotypów (np. mutantów). Badania tego typu umożliwiają porównawczą analizę złożonej odpowiedzi adaptacyjnej roślin na stres np. identyfikację nowych, nieznanych interakcji między szlakami sygnalizacyjnymi (20), tak jak to miało miejsce w przypadku stresu ozonowego; także identyfikację potencjalnych elementów kaskady sygnałowej funkcjonującej w warunkach stresu suszy (20), a nawet informację o zróżnicowaniu transkryptomu w stresie, wynikającym z procesu alternatywnego splicingu. Ta ostatnia obserwacja była dość niespodziewana z uwagi na fakt, że analizy wykonano z użyciem mikromacie-

rzy DNA nie projektowanych do tego celu. W przypadku pozostałych wymienionych zastosowań mikromacierzy do analiz fenotypu, na ich podstawie można identyfikować zarówno te geny, które determinują fenotyp, jak również i te, które go znoszą. W porównaniu odpowiedzi transkrypcyjnej mutantów w omawianych warunkach stresowych, uzyskuje się obraz konstytutywnych i specyficznych zmian transkrypcyjnych świadczących o dynamice i plastyczności transkryptomu w odpowiedzi na stres (20). Obserwacja ta odnosi się także do klasyfikacji funkcjonalnych genów, obserwowanego wzoru ich ekspresji, a także liczby uruchamianych genów w stresie. Choć te różnice są częściowo wynikiem stosowanych metod statystycznych i normalizacji (a powinny w całości być przez nie niwelowane), świadczą przede wszystkim o szkodliwości stosowanych warunków stresowych i kondycji rośliny w warunkach stresu. Obecność konstytutywnych i specyficznych wzorów ekspresji genów stwierdza się także w analizach ogólnego profilu transkrypcyjnego u roślin typu dzikiego (20,35,36).

Identyfikacja procesów biologicznych, które ulegają modyfikacji w warunkach stresu stała się priorytetem w analizie adaptacji do stresów ozonowego i suszy. Ukierunkowanie badań na identyfikację tych zmian powoduje, że w analizach transkryptomicznych stosuje się klasyfikacje funkcjonalne genów, by w konsekwencji znaleźć związki między genotypem, stresem a zaobserwowaną kondycją rośliny (tj. fenotypem wrażliwości lub tolerancji na stres). Na podstawie przeprowadzonych prac badawczych można stwierdzić, że odpowiedź transkrypcyjna roślin na oba te stropy w przeważającej części związana jest ze zmianą ekspresji genów kodujących enzymy szlaków metabolicznych, genów uczestniczących w procesach transdukcji sygnałów i transkrypcji. Zatem transkrypcyjna regulacja tych procesów jest istotnym elementem zachowania homeostazy roślin w stresie. Z punktu widzenia biologii systemu jest to punkt wyjścia do badań związanych z dostosowaniem mechanizmu transdukcji sygnału, metabolizmu, maszynierii transkrypcyjnej i sieci oddziaływań do zaistniałych szkodliwych warunków środowiska. Warto w tym miejscu nadmienić, że użycie tej informacji w powiązaniu z danymi eksperymentalnymi weryfikującymi te zależności, będzie w przyszłości stanowić cenne narzędzie w modelowaniu systemu biologicznego. Niewątpliwie wyniki analiz z użyciem mikromacierzy posłużą do selekcji odmian roślin uprawnych bądź konstruowania nowych linii i odmian hodowlanych odpornych na oba stropy.

7. Podsumowanie

Omówione w tym opracowaniu elementy metodyczne i analiza profilu transkrypcyjnego dają ogólny obraz eksperymentu z użyciem mikromacierzy DNA. Wprowadzenie tego podejścia zaowocowało zgłębieniem wiedzy o mechanizmach funkcjonowania organizmów żywych. Trzeba zaznaczyć, że ogólny profil transkrypcyjny nie dostarcza pełnych informacji o jakimkolwiek mechanizmie biologicznym. Wskazuje

jednak na aktywację lub hamowanie ekspresji genu w badanym procesie, czy korelację grup genów, a także na uruchamianie lub wyłączanie określonych szlaków metabolicznych. Wszystko to informuje o kierunku i tempie przebiegu procesów komórkowych i rozwojowych. Weryfikacja budowanych na podstawie ogólnego profilu transkrypcyjnego hipotez związanych z funkcjonowaniem genu, szlaku czy procesu wymaga na ogół dalszych ukierunkowanych badań, począwszy od potwierdzenia zmian w ekspresji na poziomie białka. Czy jednak budowanie nowych, oryginalnych hipotez nie jest najważniejszą częścią projektu naukowego?

Praca finansowana z funduszy projektu badawczego PBZ-MNiSW-2/3/2006/19.

Literatura

1. Schena M., (1996), *Bioassays*, 18, 427-431.
2. Aharoni A., Keizer L. C. P., Bouwmeester H. J., Sun Z., Alvarez-Huerta M., Verhoeven H. A., Blaas J., van Houwelingen A. M. M. L., de Vos R. C. H., van der Voet H., Jansen R. C., Guis M., Mol J., Davis R. W., Schena M., van Tunen A. J., O'Connell A. P., (2000), *Plant Cell*, 12, 647-662.
3. Reymond P., Weber H., Damond M., Farmer E., (2000), *Plant Cell*, 12, 707-719.
4. Renn S. C., Aubin-Horth N., Hofmann H. A., (2004), *Genomics*, 5, 42.
5. Barrett M. T., Scheffer A., Ben-Dor A., Sampas N., Lipson D., Kincaid R., Tsang P., Curry B., Baird K., Meltzer P. S., Yakhini Z., Bruhn L., Laderman S., (2004), *Proceedings of the National Academy of Sciences*, 101, 17765-17770.
6. Smid M., Dorssers L. C. J., Jenster G., (2003), *Bioinformatics*, 19, 2065-2071.
7. Rauch A., Ruschendorf F., Huang J., Trautmann U., Becker C., Thiel C., Jones K. W., Reis A., Nürnberg P., (2004), *Journal of Medical Genetics*, 41, 916-922.
8. Fadiel A., Naftolin F., (2003), *International Archives of Bioscience*, 1111-1121.
9. Yazaki J., Kishimoto N., Fujii F., Shimbo K., Otsuka Y., Oda T., Iba K., Yamamoto K., Sakata K., Sasaki T., Kikuchi S., (2001), *Plant Cell Physiology*, 42, 150.
10. Ke Y., Lindsay S., Chang Y., Liu Y., Yan H., (2008), *Science*, 180-183.
11. Gregory B. D., Yazaki J., Ecker J. R., (2008), *Plant Journal*, 53, 636-644.
12. Strausberg R. L., Levy S., (2007), *Genome Resources*, 17, 965-968.
13. Brazma A., Hingamp P., Quackenbush J., Sherlock G., Spellman P., Stoeckert Ch., Aach J., Ansorge W., Ball C. A., Causton H. C., Gaasterland T., Glenisson P., Holstege F. C. P., Kim I. F., Markowitz V., Matese J. C., Parkinson H., Robinson A., Sarkans U., Schulze-Kremer S., Stewart J., Taylor R., Vilo J., Vingron M., (2001), *Nature Genetics*, 29, 365-371.
14. Ludwików A., Misztal L. H., Sadowski J., (2005), *Genomika i bioinformatyka roślin*, red. Chelkowski J., Koczyk G., w: *Rozprawy i Monografie IGR PAN*, 16, 55-67.
15. Ludwików A., Gallois P., Sadowski J., (2004), *Cellular and Molecular Biology Letters*, 9, 829-842.
16. Yang Y. H., Speed T., (2002), *Nature Reviews Genetics*, 3, 579-588.
17. Bolstad B. M., Irizarry R. A., Astrand M., Speed T. P., (2003), *Bioinformatics*, 19, 185-193.
18. Berger J. A., Hautaniemi S., Järvinen A. K., Edgren H., Mitra S. K., Astola J., (2004), *Bioinformatics*, 5, 194.
19. Quackenbush J., (2002), *Nature Genetics*, 32 Suppl, 496-501.
20. Ludwików A., (2006), *Charakterystyka projektu transkrypcyjnego genomu *Arabidopsis thaliana* linii *Columbia* i mutanta *abi1* w odpowiedzi na stres ozonowy i suszę*, praca doktorska, Poznań.
21. Toronen P., Kolehmainen M., Wong G., Castren E., (1999), *FEBS Letters*, 451, 142-146.
22. Herrero J., Valencia A., Dopazo J., (2001), *Bioinformatics*, 17, 126-136.
23. Valafar F., (2002), *Annals of the New York Academy of Sciences*, 980, 41-64.
24. Ahlfors R., Lång S., Overmyer K., Jaspers P., Brosche M., Tauriainen A., Kollist H., Tuominen H., Belles-Boix E., Piippo M., Inze D., Palva E.T., Kangasjärvi J., (2004), *Plant Cell*, 16, 1925-1937.

25. Rao M. V., Koch J. R., Davis K. R., (2000), *Plant Molecular Biology*, 44, 345-358.
26. Evans N. H., McAinsh M. R., Hetherington A. M., Knight M. R., (2005), *Plant Journal*, 41, 615-626.
27. Mahalingam R., Shah N., Scrymgeour A., Fedoroff N., (2005), *Plant Molecular Biology*, 57, 709-730.
28. Tosti N., Pasqualini S., Borgogni A., Ederli L., Falistocco E., Crispi S., Paolucci F., (2006), *Plant Cell Environment*, 29, 1686-16702.
29. Shinozaki K., Yamaguchi-Shinozaki K., (2007), *Journal of Experimental Botany*, 58, 221-227.
30. Narusaka Y., Nakashima K., Shinwari Z. K., Sakuma Y., Furihata T., Abe H., Narusaka M., Shinozaki K., Yamaguchi-Shinozaki K., (2003), *Plant Journal*, 34, 137-148.
31. Gadjev I., Vanderauwera S., Gechev T. S., Laloi C., Minkov I. N., Shulaev V., Apel K., Inzé D., Mittler R., van Breusegem F., (2006), *Plant Physiology*, 141, 436-445.
32. Li P., Mane S. P., Sioson A. A., Heath L. S., Bohnert H. J., Grene R., (2006), *Plant Cell Environment*, 29, 854-868.
33. Talam V., Ozturk N. Z., Bohnert H. J., Tuberosa R., (2007), *Journal of Experimental Botany*, 58, 229-240.
34. Kathiresan A., Lafitte H. R., Chen J., Mansueto L., Bruskiwich R., Bennett J., (2006), *Field Crops Research*, 97, 101-110.
35. Seki M., Narusaka M., Ishida J., Nanjo T., Fujita M., Oono Y., Kamiya A., Nakajima M., Enju A., Sakurai T., Satou M., Akiyama K., Taji T., Yamaguchi-Shinozaki K., Carninci P., Kawai J., Hayashizaki Y., Shinozaki K., (2002), *Plant Journal*, 31, 279-292.
36. Tamaoki M., Nakajima N., Kubo A., Aono M., Matsuyama T., Saji H., (2003), *Plant Molecular Biology*, 53, 443-456.