

Mikromacierze DNA i profil ekspresji genów raka brodawkowego tarczycy

Barbara Jarzab¹, Elżbieta Gubała¹, Dariusz Lange²

¹ Zakład Medycyny Nuklearnej i Endokrynologii Onkologicznej,

² Zakład Patologii Nowotworów Centrum Onkologii - Instytut im. Marii Skłodowskiej-Curie, Oddział w Gliwicach

Streszczenie

W pracy przedstawiamy zasady badania profilu ekspresji genów na mikromacierzach DNA i porównanie dwóch głównych platform technologicznych używanych do tego celu – mikromacierzy oligonukleotydowych wysokiej gęstości i mikromacierzy cDNA. Na tym tle przedstawimy dotychczasowe wyniki analiz raka brodawkowego tarczycy z użyciem mikromacierzy DNA.

Słowa kluczowe: mikromacierz DNA, profil ekspresji genów, rak brodawkowy tarczycy



Centrum Onkologii
Instytut im. Marii Skłodowskiej-Curie,
Oddział w Gliwicach
Wybrzeże Armii Krajowej 15,
44-101 Gliwice

Praca wykonana w ramach grantu PBZ-KBN-040-P04/2001

W niniejszym opracowaniu omówione zostaną prace dotyczące profilu ekspresji genów w raku brodawkowym tarczycy. Zanim jednak przejdziemy do tego tematu, konieczne jest podanie kilku ogólnych informacji o technikach badania z zastosowaniem mikromacierzy DNA.

Mikromacierze DNA: zasada metody, typy mikromacierzy

Projekt poznania ludzkiego genomu (Human Genome Project), zakończony w ubiegłym roku, przyniósł nie tylko wiedzę o sekwencji ludzkich genów. Ogrom zadania, jaki stanął przed jego wykonawcami, mobilizował do odkryć ułatwiających sekwencjonowanie z jednej strony, a wykorzystujących zdobytą w ten sposób wiedzę o ludzkich genach z drugiej strony. Jednym z najważniejszych kierunków badawczych, które zostały zainicjowane „wokół” tego projektu stało się wdrożenie badania profilu molekularnego komórek i tkanek przez równoczesną hybrydyzację izolowanego z nich RNA do wielu tysięcy sond genowych. W ten sposób powstały pierwsze mikromacierze DNA (DNA microarrays) [1]. Od początku lat dziewięćdziesiątych, czyli niemal przez 10 lat udoskonalano technikę syntezy *in situ* na małych płytkach szklanych krótkich oligonukleotydów o z góry zaplanowanych sekwencjach homologicznych do sekwencji poznawanych genów [2]. Przy tej syntezie, sterowanej komputerowo, wykorzystywano doświadczenia zdobyte wcześniej przy produkcji fotolitolgraficznej obwodów scalonych. Tym razem na płytkę nakładano jednak materiał biologiczny – odpowiednio uporządkowane nukleotydy syntetyzowane w krótkie łańcuchy jednoniciowego DNA.

W początkowym okresie taka technika była niezwykle droga, a ponieważ jednocześnie swoistość

wykrywania genów za pomocą krótkich sond wcale nie była wysoka, wdrożenie tej techniki do badań molekularnych następowało bardzo powoli. Przełomem stało się opublikowanie w 1999 roku w *Science* pracy poświęconej profilowi molekularnemu blastów białaczkowych [3]. Okazało się, że badając profil ekspresji genów metodą mikromacierzy DNA można odróżnić ostrą białaczkę szpikową od limfoblastycznej na podstawie danych dotyczących ekspresji zaledwie kilkudziesięciu genów. Zastosowane tu mikromacierze produkowane fotolitolgraficznie przez firmę Affymetrix®, która posiada patent na tę technologię, nazywane mikromacierzami oligonukleotydowymi wysokiej gęstości (ponieważ system został nazwany Genechip®, często spotyka się też określenie chipy DNA) [4, 5].

Równoległe do tej technologii, zakładającej kontrolowaną jednoczesną syntezę wielu tysięcy różnych sond na jednym szkiełku, rozwijała się druga platforma, w której wykorzystano biblioteki klonów DNA, automatycznie nakrapiane na szkiełko [6, 7, 8]. Te macierze nazwano mikromacierzami cDNA. Opracowane na Uniwersytecie Stanforda [9], wzbudziły wielkie zainteresowanie z kilku powodów. Po pierwsze, można było za ich pomocą badać ekspresję genów nawet nie znając ich pełnej sekwencji, o ile dysponowało się odpowiednimi klonami DNA namnożonymi przez odwrotną transkrypcję i reakcję łańcuchową polimerazy z mRNA izolowanego z danych komórek. Takie podejście ułatwiało odkrywanie nowych genów ulegających transkrypcji w poszczególnych typach komórek i tkanek. Drugą, niezwykle ważną zaletą był fakt, że w przeciwieństwie do metody poprzedniej, nie została ona natychmiast skomercjalizowana, autorzy szeroko ją udostępniali [7, 10]. Wystarczyło więc zakupić odpowiedni automat nakrapia-

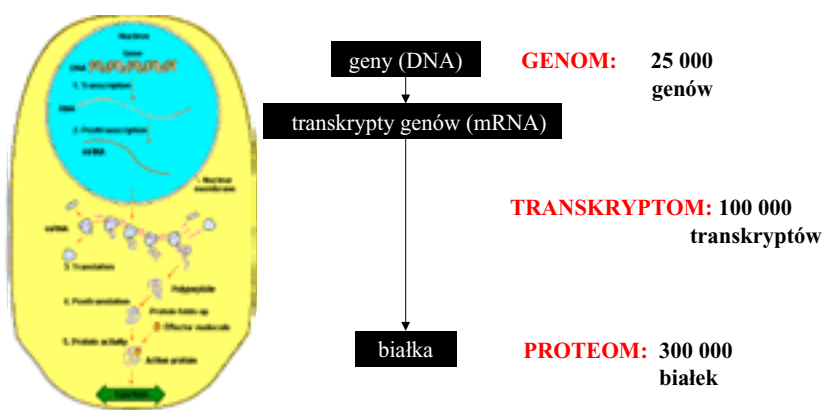
jący tysiące różnych sekwencji w sposób uporządkowany na szkiełko, trzeba było oczywiście też dysponować własną biblioteką klonów i umiejętnie je namnażać. W miarę postępu tej metody dostępne stały się biblioteki komercyjne (można więc kupić potrzebny zestaw klonów), a także produkowane seryjnie mikromacierze.

Trzecią, na początku chyba najważniejszą zaletą metody mikromacierzy cDNA była elastyczność tej technologii. W miarę zdobywania doświadczeń o ekspresji genów w interesujących nas komórkach można bowiem modyfikować skład mikromacierzy i nanosić na szkiełko nowe sekwencje. W mikromacierzach oligonukleotydowych wysokiej gęstości takie częste zmiany nie są możliwe, gdyż matryce używane w technice fotolitograficznej są bardzo drogie. Łatwiej też, dysponując klonami DNA, po wybraniu interesujących nas genów ułożyć macierz dedykowaną, składającą się z niewielkiej liczby genów używanych w celach diagnostycznych. Twórcy i pierwsi użytkownicy tych mikromacierzy spodziewali się też, że dłuższe sekwencje cDNA (na ogół używano klonów o długości do 200 nukleotydów) zapewnią większą swoistość hybrydyzacji w stosunku do macierzy oligonukleotydowych.

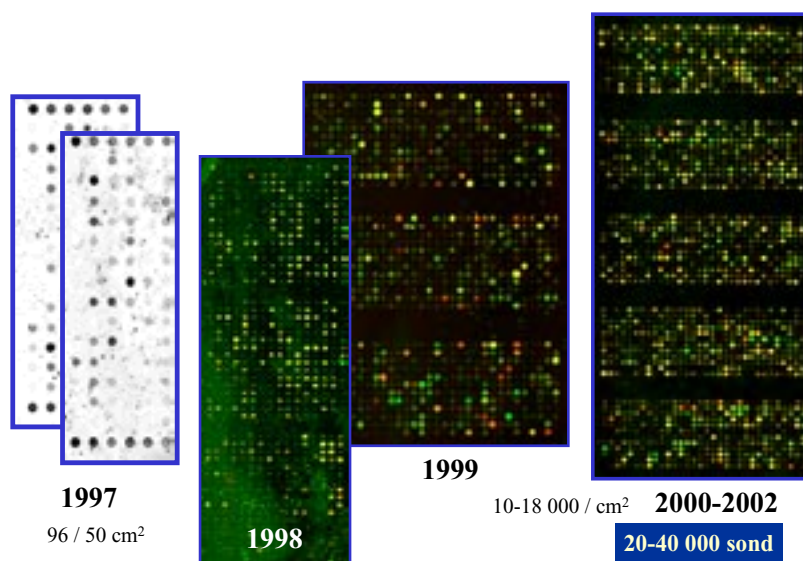
Zasada metody była jednak taka sama w obydwu przypadkach (ryc. 1): Badany jest cały transkryptom komórki, po izolacji całkowitego RNA, które następnie przepisywane jest w procesie odwrotnej transkrypcji na cDNA. W mikromacierzach cDNA ten właśnie materiał jest łączony z barwnikiem fluorescencyjnym i nanoszony na mikromacierz, a każde jednoniciowe cDNA w tej mieszaninie hybrydyzuje tylko z sondą o homologicznej sekwencji. Jeżeli więc wypłuczemy niezwiązane cDNA, to sygnał barwny – plamka odpowiadająca danemu transkryptowi – może zostać wykryty przez czytnik laserowy. W mikromacierzach oligonukleotydowych wysokiej gęstości na macierz nakłada się fragmenty cRNA, a znakowanie fluorescencyjne następuje już po hybrydyzacji. Ponieważ na mikromacierzach znajdują się tysiące takich sond (ryc. 2), czytnik musi bardzo dokładnie lokalizować położenie każdej plamki i odnosić je do zakodowanej informacji, jaką sekwencję ona reprezentuje.

Różnice między mikromacierzami cDNA i mikromacierzami oligonukleotydowymi wysokiej gęstości

Przedstawione powyżej różnice w wytwarzaniu obydwu typów mikromacierzy muszą być uzupełnione o dalsze informacje dotyczące różnic w hybrydyzacji. Najważniejsza różnica dotyczy porównawczego charakteru badania wykonanego na mikromacierzy cDNA w stosunku do bezwzględnego pomiaru na macierzy wysokiej gęstości (ryc. 3 i 4). W praktyce oznacza to, że na mikromacierz cDNA nakładamy zawsze cDNA uzyskane z dwóch różnych typów komórek wyjściowych – jedno cDNA jest znakowane barwnikiem zielonym a drugie czerwonym (ryc. 3). Obydwa konkurują w czasie hybrydyzacji o te same sondy. Jeżeli więc w jednym typie komórek ekspresja danego genu jest znacznie silniejsza niż w drugim, to z sondą zwiąże się przede wszystkim cDNA znakowane jednym barwnikiem – na przykład czerwonym. Jeżeli plamka po hybrydyzacji będzie miała kolor



Ryc. 1 Badanie profilu ekspresji genów dotyczą transkryptomu komórki

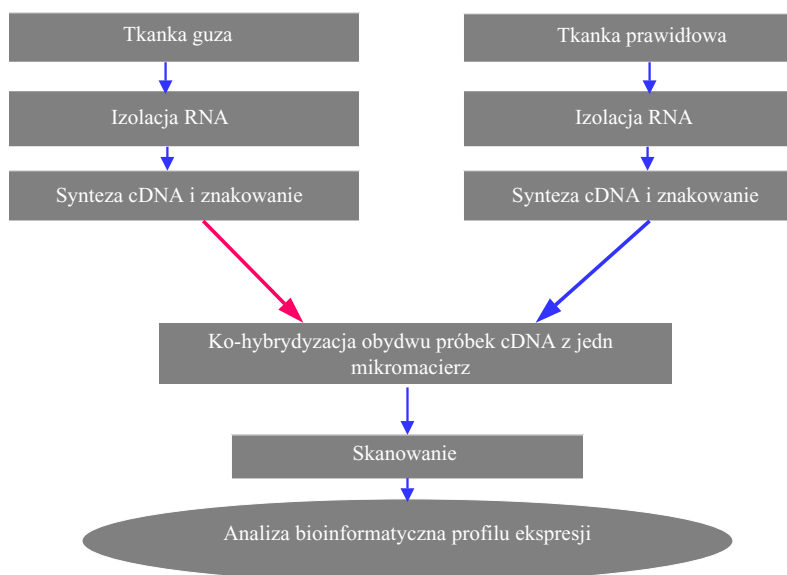


Ryc. 2 Rozwój techniki mikromacierzy DNA

zielony, to będzie to oznaczało, że transkryptu badanego w tej chwili genu było znacznie więcej w drugim ekstrakcie. Jeżeli wreszcie obydwa typy komórek nie będą się różniły zawartością tego transkryptu, to kolor będzie pośredni, podczas gdy brak zabarwienia będzie oznaczał, że w obydwu typach komórek badany gen nie ulegał ekspresji. W praktyce najczęściej stosuje się ekstrakt z komórek zdrowych i chorych lub też ekstrakt z komórek w stanie wyjściowym i z komórek poddanych działaniu badanego czynnika i mówi o nadekspresji lub hamowaniu ekspresji danego genu w badanym typie komórek w porównaniu do komórek kontrolnych. Ponieważ czynniki mikromacierzy musi być wówczas zdolny do wykrywania dwóch barwników jednocześnie, jest on *de facto* zmodyfikowanym (przystosowanym do bardzo dokładnego przesuwu nad plamkami) mikroskopem konfokalnym. Oceniane plamki mają na ogół średnicę około 100 μm i znajdują się od siebie w odległości 200-300 μm .

W mikromacierzach oligonukleotydowych wysokiej gęstości na jedną płytkę nakłada się tylko jeden materiał biologiczny – odczyt fluorescencji nie ma więc charakteru porównawczego tylko bezwzględny (ryc. 4). Jeżeli więc chcemy porównać dwa typy komórek, musimy wykonać dwa razy więcej hybrydyzacji, co zwiększa nakład pracy i koszty, ale jednocześnie umożliwia wykonywanie wielu porównań między różnymi typami komórek.

Druga poważna różnica dotyczy charakteru pomiaru. W mikromacierzach cDNA najczęściej jednemu genowi odpowiada tylko jedna „plamka” (*spot*) cDNA. Jeżeli więc z jakiegokolwiek powodu dojdzie do nierównej hybrydyzacji w różnych miejscach mikromacierzy, spowoduje to znaczne błędy w ocenie ilości transkryptu. Ponieważ typ barwnika i kolejność nakładania mają znaczenie, stosuje się 2 równoległe oznaczenia, raz znakując badaną tkankę na czerwono, a raz na zielono. W mikromacierzach wysokiej gęstości ekspresję jednego genu oceniamy na podstawie hybrydyzacji transkryptu z kilkunastoma sondami homologicznymi do różnych odcinków jego sekwencji (ale za to każda z tych sekwencji jest znacznie krótsza niż sekwencja wykrywana w czasie hybrydyzacji



Ryc. 3. Mikromacierze na cDNA: ko-hybrydyzacja dwóch próbek do jednej mikromacierzy

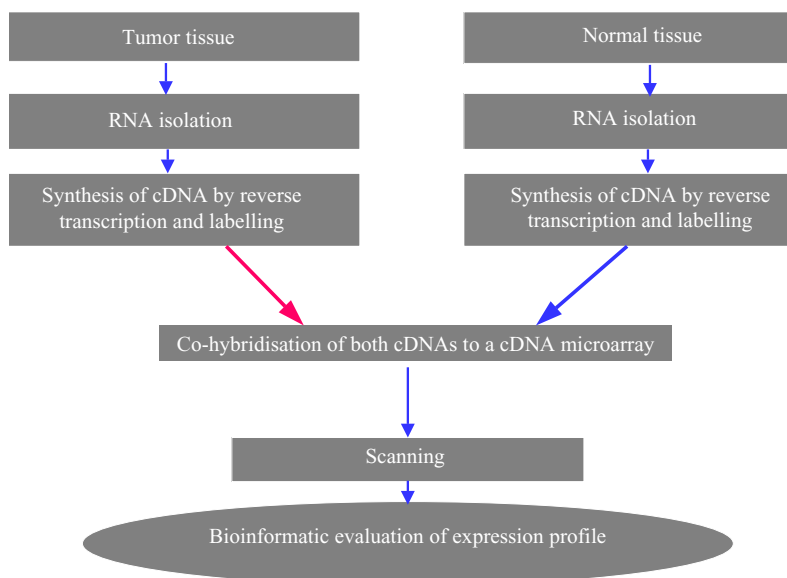
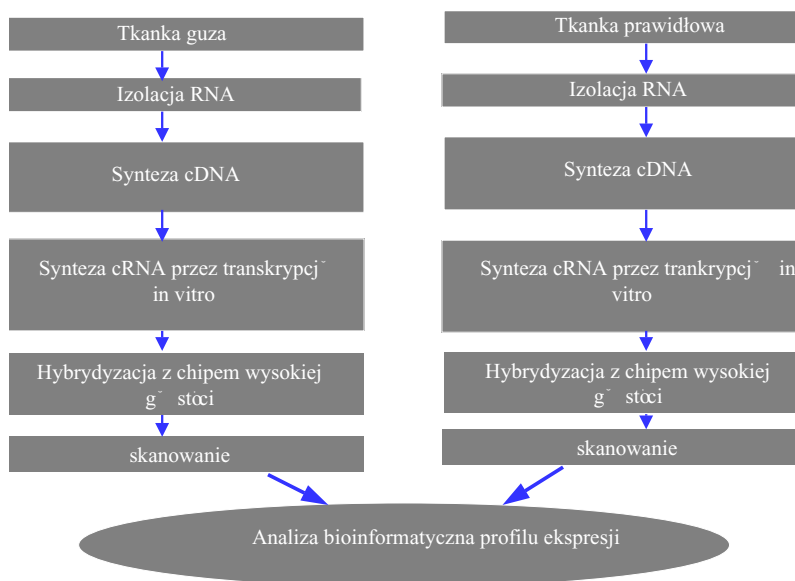


Fig. 3. cDNA microarrays: cohybridisation of two tissue samples to one microarray

z cDNA). Te sondy rozmieszczone są w różnych miejscach mikromacierzy, więc błąd związany z lokalizacją sondy na mikromacierzy jest w tym przypadku znacznie mniejszy. Ponieważ jednak sondy używane do detekcji są znacznie krótsze (składają się na ogół z 25 nukleotydów), ryzyko nieswoistej hybrydyzacji jest większe. Producent dodał więc dla każdej sondy osobną kontrolę, identyczną co do sekwencji z wyjątkiem podstawienia centralnej zasady. Założenie jest więc takie, że hybrydyzacja naszego badanego transkryptu do tej sondy kontrolnej musi być znacznie niższa niż hybrydyzacja do sondy właściwej, żeby uznać tę ostatnią za hybrydyzację swoistą.



Ryc. 4. Mikromacierze wysokiej gęstości: każda tkanka analizowana jest na osobnej płytce

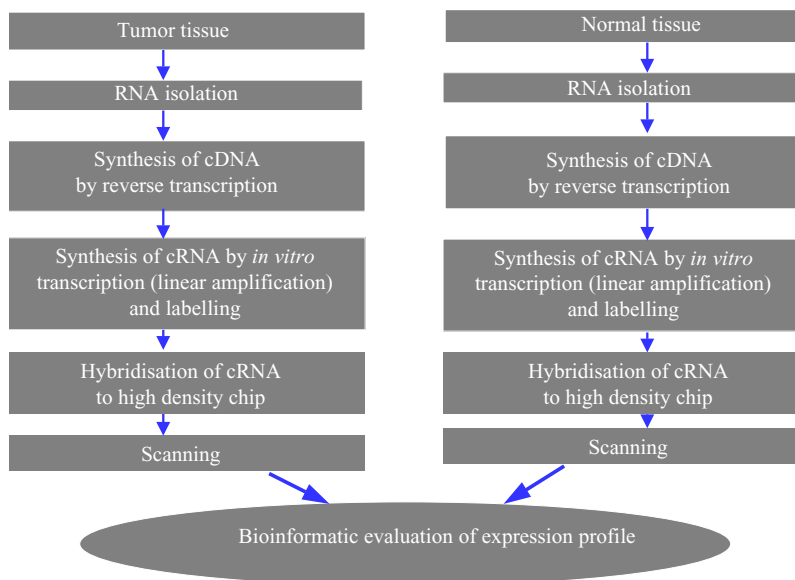


Fig. 4. High density oligonucleotide microarrays: hybridisation of each tissue sample to separate microarray

W sumie, jeżeli wziąć pod uwagę wszystkie te różnice, to okazuje się, że pomiar na mikromacierzach wysokiej gęstości znacznie lepiej spełnia wymagania pomiaru ilościowego (choć nie jest w pełni bezwzględny) niż pomiar na mikromacierzach cDNA – sygnał ekspresji pojedynczego genu jest oceniany na podstawie analizy statystycznej wielu pomiarów i hybrydyzacji jednego rodzaju transkryptu do jednej sondy, a nie dwóch próbek z transkryptu dodanych równocześnie. Ze względu na obudowanie tej mikromacierzy własną komorą hybrydyzacyjną i automatyzację procesu barwienia sama metoda jest też znacznie bardziej powtarzalna i łatwiejsza do wdrożenia niż metoda z użyciem mikromacierzy cDNA. Te zalety okupione są

wysoką ceną mikromacierzy oligonukleotydowych wysokiej gęstości, a także niewielką elastycznością badań. Wadą systemu firmy Affymetrix jest też, że mamy do czynienia z systemem zamkniętym – do odczytu można użyć tylko czytnika stanowiącego integralną część systemu, z kolei nie można na nim odczytywać innych mikromacierzy cDNA. Jeśli zdecydujemy się na mikromacierze cDNA, będziemy dysponować systemem otwartym – na czytniku będziemy mogli odczytywać różne rodzaje mikromacierzy, łatwiej też przyjdzie nam modyfikować zestawy genów na tych mikromacierzach. Z kolei musimy się liczyć z tym, że ustawienie metody, jej kontrola i standaryzowanie może się okazać zadaniem daleko trudniejszym niż w przypadku mikromacierzy wysokiej gęstości, znacznie trudniej będzie nam też porównywać nasze wyniki z wynikami innych laboratoriów. Pewną poprawę w tym zakresie mogą przynieść ostatnio rozwijające się modyfikacje, w których nakrapianie klonów DNA zastępuje się syntezą *in situ* dłuższych nukleotydów, które dalej opracowywane są tak jak macierze cDNA.

Jaka jest dostępność badań mikromacierzowych w Polsce?

Pierwsze w kraju laboratorium mikromacierzowe zostało zorganizowane w Instytucie Onkologii w Gliwicach w ramach grantu zamawianego KBN i działa od blisko dwu lat. W laboratorium tym używamy mikromacierzy oligonukleotydowych wysokiej gęstości w systemie Genechip Affymetrix.

Niedługo potem powstały podobne laboratoria w Poznaniu i Sosnowcu. Wszystkie trzy działają w systemie zamkniętym. Według danych dostępnych autorom tego opracowania uniwersalny czytnik umożliwiający odczytywanie różnych typów mikromacierzy cDNA znajduje się w Instytucie Biochemii i Biofizyki PAN w Warszawie, a system do drukowania własnych mikromacierzy i ich odczytu zostanie też zainstalowany w Instytucie Chemii Bioorganicznej PAN w Poznaniu, a prawdopodobnie także w Łodzi.

Jakie geny oceniamy na mikromacierzy?

Jeżeli w publikacji czytamy, że używano macierzy cDNA 10K, to znaczy, że znalazło się na niej 10

tysięcy różnych sond (ryc. 2). W pewnym uproszczeniu mówimy wówczas o 10 000 genów, musimy jednak pamiętać, że czasem kilka sond wykrywa *de facto* jeden transkrypt. Czasem różnice w sekwencji powodują, że wykrywamy dwa warianty składania jednego genu, czasem jest jednak na odwrót – sekwencja sondy jest tak dobrana, że mimo istniejących *in vivo* różnic w sekwencji alternatywnych wariantów składania wszystkie one hybrydują z jedną sondą, pomiar jest więc wypadkową ich obecności w badanym ekstrakcie komórkowym. Na ogół dostępne obecnie mikromacierze nie są zdolne do rozróżniania poszczególnych wariantów składania danego genu, niektóre wariantowe sekwencje mogą być w ten sposób w ogóle pominięte.

Jak więc z kolei interpretować fakt, że na niektórych mikromacierzach znajduje się 30-50 000 różnych sekwencji ludzkich genów, podczas gdy liczba genów człowieka oscyluje wokół 25 000? Najczęściej mamy do czynienia z obecnością więcej niż jednej sondy dla danego genu, czasem sekwencja wybrana dla danej sondy nie jest optymalna i tak naprawdę nie odpowiada znanemu genowi tylko sekwencji EST, której ostateczne znaczenie biologiczne nie jest zdefiniowane.

Określenie "gen" dla sekwencji wykrywanych na mikromacierzach nie powinno więc być traktowane absolutnie dosłownie, tym bardziej, że *de facto* wykrywamy za pomocą macierzy ekspresyjnych nie tyle same geny, co ich transkrypty.

Jak interpretujemy wyniki badań na mikromacierzach DNA?

Jak już wspomnieliśmy, przedmiotem naszej oceny jest obecność transkryptów genów w badanych komórkach, pomiar mikromacierzy daje nam więc informację o ich ekspresji – możemy na podstawie tego badania ocenić, czy transkrypcja danego genu przebiegała intensywnie (a więc był on aktywny) czy też przeciwnie, doszło do zahamowania ekspresji genu. Ponieważ badanie to wykonujemy równoległe dla wielu tysięcy genów, otrzymany w jego efekcie zestaw wyników nazywamy profilem ekspresji genu, a samą mikromacierz mikromacierzą ekspresyjną [11, 12]. Należy bowiem dodać, że na mikromacierzach możemy również wykonywać różne zadania związane z genotypowaniem, które nie są jednak przedmiotem niniejszego omówienia [13, 14].

Profil ekspresji genów to określenie, które może być rozumiane dwojako: Nazywamy tak zarówno pełen zestaw pomiarów ekspresji genów, jak i pewną jego wyselekcjonowaną część, charakterystyczną dla danego typu komórki czy tkanki lub jej stanu czynnościowego. Mniejsze zestawy genów o szczególnym znaczeniu nazywamy też klasyfikatorami molekularnymi (jeżeli selekcjonujemy je w celach klasyfikacji badanych tkanek)

lub podpisami czy sygnaturami molekularnymi (*molecular signature, expression signature*), jeżeli za ich pomocą chcemy opisywać wybrane cechy badanych komórek lub ich stan.

Analiza danych z mikromacierzy

Jakkolwiek jesteśmy przyzwyczajeni do konieczności analizy statystycznej uzyskanych przez nas danych pomiarowych, to na początku badań z użyciem mikromacierzy nie docenialiśmy stopnia trudności analizy danych uzyskanych tą techniką. Abstrahując od oceny jakości wykonanych badań (która w mikromacierzach cDNA może stanowić prawdziwe wyzwanie, w mikromacierzach wysokiej gęstości jest natomiast bardziej wystandaryzowana), oraz normalizacji i wstępnego opracowania wyników, które stanowią odrębne i bardzo ważne zagadnienie [15, 16], trzeba mieć świadomość, że klasyczne metody statystyczne nie nadają się do oceny tych wyników. Jeżeli na przykład oceniamy mikromacierze, które zawierają 22 000 genów (U133A) (a w Gliwicach wdramy też mikromacierze U133.2 PLUS, które zawierają 38000 genów), to *de facto*, wykonując badanie na 20 tkankach (2 grupy po 10), otrzymamy tabelę wyników składającą się z 22 000 testów przeprowadzonych na tych 20 powtórzeniach. Ryzyko wyniku fałszywie dodatniego jest wówczas olbrzymie, jeżeli nie zastosuje się bardzo silnych poprawek na wielokrotność powtórzeń. Co więcej, metody analizy, który każdy gen rozpatrują w oderwaniu od pomiarów ekspresji innych genów, są tutaj nieadekwatne. Dlatego na potrzeby analizy mikromacierzy opracowano wiele nowych testów statystycznych lub zaadaptowano wcześniej istniejące, ale rzadko w praktyce biologicznej wykorzystywane metody matematyczne [12, 17, 18, 19]. Do tych ostatnich należą analiza klasterowa, analiza składowych głównych (*Principal Component Analysis*) czy bardzo podobna do niej dekompozycja wartości szczególnych (*Singular Value Decomposition*) [20]. Ogromne uznanie zdobywają metody oparte na testach permutacyjnych, przede wszystkim przygotowana specjalnie dla badań mikromacierzowych analiza znaczenia mikromacierzy (*Significance Analysis of Microarrays*) [21]. Ta ostatnia technika wkracza już w pole zainteresowań bioinformatyki. Szczególnie przydatne dla analizy mikromacierzy okazały się techniki bioinformatyczne oparte o sieci neuronowe z grupy tzw. maszyn wektorów podpieranych (*Supported Vector Machines*), które w analizach profilu ekspresji genów traktowane są ostatnio jako techniki z wyboru i szczególnie dobrze nadają się do selekcji optymalnych sygnatur czy klasyfikatorów molekularnych i następnego weryfikowania ich przydatności [22].

Zadaniem niniejszego opracowania nie jest opis analizy bioinformatycznej badań z mikromacierzy, a jedynie podkreślenie konieczności ich wdrożenia.

Każda grupa, która chce podjąć badania profilu ekspresji, musi zadbać z góry o kwestię analizy wyników, najlepiej przez podjęcie współpracy z przygotowaną do tego zadania grupą bioinformatyków. W Gliwicach mamy szczęście współpracować z niezwykle zaawansowaną w realizacji tych zadań grupą z Instytutu Automatyki Politechniki Śląskiej, kierowaną przez prof. Andrzeja Świerniaka.

Profil ekspresji raka brodawkowatego tarczycy

Rak brodawkowaty tarczycy, najczęstszy histotyp raka tarczycy, został po raz pierwszy przeanalizowany techniką mikromacierzy w roku 2001, kiedy grupa z laboratorium prof. de la Chapelle z Columbus, Ohio, opublikowała analizę 8 guzów i 8 tkanek zdrowych pobranych od tych samych chorych [23]. W wyniku tej analizy wytypowano około 50 genów, z wyróżnieniem genów ulegających nadekspresji i wyciszeniu w tkankach raka (tab. I). Te ostatnie zostały przeanalizowane nieco szerzej w kolejnej publikacji pod kątem ich znaczenia w transformacji nowotworowej komórki pęcherzykowej tarczycy [24]. W odniesieniu do genów hamowanych autorzy największe znaczenie przypisali genowi CITED1, który był jednym z genów po raz pierwszy wiązanych z rakiem brodawkowatym tarczycy przez to właśnie badanie.

W ostatnim czasie ukazało się także kilka innych prac stosujących technikę mikromacierzy DNA w raku tarczycy [25-29]. W tym opracowaniu ustosunkujemy się wyłącznie do tych z nich, które poruszają kwestie związane z rakiem brodawkowatym tarczycy. Porównując te prace trzeba wziąć pod uwagę różnice metodyczne, bardzo wyraźne tak jeżeli chodzi o zastosowane mikromacierze jak i analizę uzyskanych wyników (tab. II). Musimy uwzględnić poważne różnice między sondami obecnymi na poszczególnych mikromacierzach, a trzeba podkreślić, że dostępne informacje nie pozwalają na razie na wiarygodne porównanie wszystkich genów chociażby z tego

powodu, że stosowane są różne nazwy genów, nie ma dokładnych danych o sekwencji stosowanych sond itp. Trudno więc gen po genie porównywać analizę wykonaną przez Waseniusa i wsp. na błonach nylonowych które zawierają około 1000 sond najlepiej poznanych genów z analizą wykonaną przez Huanga i wsp., gdzie tych sond było 20x więcej, ale dla nakładających się genów wykrywane były często nieco inne sekwencje, czy analizą Mazzanti i wsp., wykonaną na mikromacierzach cDNA, gdzie nie dysponujemy informacją o sekwencjach cDNA. Dlatego w naszej analizie będziemy raczej uwypuklać podobieństwa niż ustosunkowywać się do różnic, gdyż w odniesieniu do tych ostatnich brak nam kryteriów dla określenia, czy mają one charakter pozorny czy rzeczywisty i oceny, na ile związane są ze stosowaną metodyką.

Analizując podobieństwa, trzeba przede wszystkim podkreślić, że genem wspólnym dla większości przeprowadzonych analiz jest gen *MET*. Ten proto-onkogen, receptor dla HGF - czynnika wzrostowego wydzielanego między innymi przez fibroblasty i nieco myląco nazwanego wątrobowym czynnikiem wzrostu – był już wcześniej opisywany w raku brodawkowatym tarczycy jako gen ulegający wyraźnej nadekspresji [30, 31]. Przyjęto, że nadekspresja *MET* jest wynikiem aktywacji *RET*. Rozpatrywano także znaczenie prognostyczne nadekspresji *MET*, wiążąc ją na ogół z rakami brodawkowatymi o większej agresywności i mniej korzystnym przebiegu klinicznym. Z badań mikromacierzowych wynika jednak, że nadekspresja *MET* jest jedną z charakterystycznych cech raka brodawkowatego, odróżniających ten typ guza nowotworowego od zdrowej tarczycy. *MET* znalazł się w grupie genów wytypowanej zarówno przez Huanga i wsp., Waseniusa i wsp., jak i Finley'a i wsp. [23, 27, 28] Także nasze własne badania [32] wskazują na istotną rolę *MET*. Dwie sondy dla tego genu weszły w skład klasyfikatora molekularnego raka brodawkowatego tarczycy,

Tab. I. Geny wyselekcjonowane w raku brodawkowatym tarczycy w analizie Huanga i wsp. (2001). Strzałką w górę zaznaczono geny ulegające nadekspresji, strzałką w dół geny, których transkrypcja została zahamowana

Tab. I. Genes selected in papillary thyroid carcinoma by Huang et al. (2001). First arrow indicates up-regulated genes, whereas second arrow indicates down-regulated genes

N33	↑	SCEL	SCYA21	↓	CRABP1
ODZ1		KRT19	AKR1C2		FABP4
UP		DPP4	HBA		DIO2
SPUVE		SDC4	HGD		DIO1
KIA0937		PROS1	UBD		TPO
TSSC3		P4HA2	C11ORF8		CDC2L1
RIL		CST6	MT1G		FOSB
CITED1		SFTPB	TFF3		RGS16
DUSP6		CHI3L1	FGL2		DUSP1
EPS8		LGALS3	DPT		MEOX2
ADORA1		SERPINA1	STC1		GPC3
MET		FN1	ITPR1		BCL2
			CITED2		GAS1
			APOD		

Tab. II.

Autor	Rok (pozycja piśmiennictwa)	Typ mikromacierzy	Liczba genów na macierzy
Takano i wsp.	2000 (26)	SAGE*	29 000*
Huang i wsp.	2001 (23)	oligonukleotydowe wysokiej gęstości (Affymetrix U95)	12 000
Wasenius i wsp.	2003 (27)	makromacierze cDNA na błonach nylonowych (Atlas)	1176
Mazzanti i wsp.	2004 (25)	cDNA (NCI/NIH)	10 000
Finley i wsp.	2004 (28)	oligonukleotydowe wysokiej gęstości (Affymetrix U95)	12 000

*technika SAGE (seryjna analiza ekspresji genów; serial analysis of gene expression) nie jest technika opartą o mikromacierze DNA, ale jest także wysokowydajną techniką badania profilu ekspresji genów, która historycznie rzecz biorąc, jest nieco wcześniejsza od techniki mikromacierzy DNA

Tab. II.

Author	Year of publication (number of citation)	Microarray type	Number of genes/ array
Takano et al.	2000 (26)	SAGE*	29 000*
Huang et al.	2001 (23)	high density oligonucleotide (Affymetrix U95)	12 000
Wasenius et al.	2003 (27)	nylon membrane cDNA macroarray (Atlas)	1176
Mazzanti et al.	2004 (25)	cDNA (NCI/NIH)	10 000
Finley et al.	2004 (28)	high density oligonucleotide (Affymetrix U95)	12 000

* SAGE technique (serial analysis of gene expression) is another high-throughput method of gene expression analysis, but not based on DNA microarrays

złożonego z 20 sond, który zapewniał maksymalną jakość dyskryminacji między tkanką raka brodawkowatego tarczycy i tkanką pozostałego gruczołu, wybranego techniką SVM.

Należy podkreślić, że w klasyfikatorach wielogenowych wybranych przez różnych autorów szczególnie często powtarza się gen galektyny 3, a także gen fibronektyny (FN), gen inhibitora tkankowego metaloproteinaz typu 1 (TIMP-1) i inne geny związane z białkami przestrzeni pozakomórkowej. Nadekspresja galektyny-3 jest znana z wcześniejszych badań raka tarczycy [33, 34], wiadomo już jednak także, że badanie tego genu jako pojedynczego markera raka tarczycy związane jest ze znaczącym ryzykiem wyniku fałszywie dodatniego [37]. Na tym polega właśnie przewaga klasyfikatorów wielogenowych, wybieranych na podstawie analizy całego profilu ekspresji genów guza, a nie pojedynczych badań. Interakcja między genami, niezwykle istotna w oddziaływaniach genów *in vivo*, odzwierciedla się bowiem w składzie klasyfikatora, o ile jest on wybierany odpowiednio zaawansowanymi metodami bioinformatycznymi. Ale też, przydatność badań mikromacierzowych byłaby kwestionowana, gdyby wyselekcjonowane na ich podstawie zestawy genów zawierały wyłącznie geny, które były wcześniej znane z badań klasycznych. Takie ryzyko istnieje, jeżeli zastosujemy mikromacierz zawierającą wyłącznie znane geny [27] lub geny wyselekcjonowane w kierunku tych uczestniczących w proliferacji nowotworo-

wej. Wyniki przeprowadzonych dotychczas analiz [23, 28, 36], jak i nasze własne doświadczenie [32] wskazują, że rak brodawkowaty tarczycy charakteryzuje się bardzo wysoką ekspresją wielu genów, które nie były dotychczas zaliczane do klasycznych markerów transformacji nowotworowej, a które uczestniczą w procesach adhezji komórkowej, przekładając się zarówno na miejscową inwazyjność nowotworu, jak i jego skłonność do tworzenia przerzutów.

Podsumowanie

W podsumowaniu należy stwierdzić, że badania profilu ekspresji genów z użyciem mikromacierzy DNA stanowią niezwykle ciekawą i płodną drogę analizy mechanizmów transformacji i progresji nowotworów [9, 37-43], a jednocześnie za ich pomocą można uzyskać dane do budowy nowych testów diagnostycznych i prognostycznych o potencjalnym ogromnym znaczeniu klinicznym. Jeżeli

porównany aktualny stan badań profilu ekspresji raka brodawkowatego tarczycy z zaawansowaniem i wynikami badań prowadzonych w innych nowotworach, możemy stwierdzić, że wyniki uzyskane w raku brodawkowatym są dość powtarzalne, a jednocześnie bogate i mogą już dzisiaj stanowić punkt wyjścia dla wielu niezwykle interesujących analiz. Trzeba bowiem podkreślić, że badania z użyciem mikromacierzy DNA służą raczej do budowania nowych hipotez, które muszą być następnie sprawdzone klasycznymi technikami biologii molekularnej czy analizami klinicznymi.

Piśmiennictwo

1. Blohm DH, Guiseppi-Elie A. New developments in microarray technology. *Current Opinion in Biotechnology* 2001; 12: 41-47.
2. Ishii M, Hashimoto S, Tsutsumi S et al. Direct comparison of GeneChip and SAGE on the quantitative accuracy in transcript profiling analysis. *Genomics* 2000; 68: 136-143.
3. Golub TR, Slonim DK, Tamayo P et al. Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring. *Science* 1999; 286: 531-537.
4. Tillib SV, Mirzabekov AD. Advances in the analysis of DNA sequence variations using oligonucleot microchip technology. *Curr Opin Biotechnol* 2001; 12: 41-47.
5. Warrington JA, Dee S, Trulson M. Large-scale genome analysis using Affymetrix GeneChip® probe arrays. *Microarray Biochip Technology*, BioTechniques Books, Natick, MA, 2000:119-148.
6. Meldrum D. Automation for genomics, part two: sequencers, microarrays, and future trends. *Genome Research* 2000; 10: 1288-1303.

7. Mace ML Jr, Montagu J, Rose DS, McGuinness G. Novel microarray printing and detection technologies. *Microarray Biochip Technology*, BioTechniques Books, Natick, MA, 2000: 39-64.
8. Worley J, Bechtol K, Penn S et al. A system approach to fabricating and analyzing DNA microarrays. *Microarray Biochip Technology*, BioTechniques Books, Natick, MA, 2000:65-86.
9. Perou CM, Sorliet T, Eisen MB et al. Molecular portraits of human breast tumours. *Nature* 2000; 406: 747-752.
10. Zubritsky E. Spotting a microarray system. A broader array of commercial products edges out home-built systems. *Modern Drug Discovery* 2001: 59-71.
11. Khan J, Bittner ML, Chen Y et al. DNA microarray technology: the anticipated impact on the study of human disease. *Bioch Biophys Acta* 1999; 1423: M17-M28.
12. Greenfield A. Applications of DNA microarrays to the transcriptional analysis of mammalian genomes. *Mammalian Genome* 2000; 11: 609-613.
13. Kurg A, Tönisson N, Georgiou I et al. Arrayed primer extension: solid-phase four-color DNA resequencing and mutation detection technology. *Genetic Testing* 1999; 1: 1-7.
14. Shoemaker DD, Schadt EE, Armour CD et al. From gels to chips: "minisequencing" primer extension for analysis of point mutations and single nucleotide polymorphisms. *Human Mutation* 1999; 13: 1-10.
15. Li C, Wong WH. Model-based analysis of oligonucleotide arrays: Expression index computation and outlier detection. *Proc Nat Acad Sci USA* 2001; 98: 31-36.
16. Schadt EE, Li C, Su C, Wong WH. Analyzing high-density oligonucleotide gene expression array data. *J Cell Biochem* 2000; 80: 192-202.
17. Brazma A, Vilo J. Gene expression data analysis. *FEBS Letters* 2000: 17-24.
18. Burke HB. Discovering patterns in microarray data. *Mol Diagn* 2000; 5(4): 349-357.
19. Getz G, Levine E, Domany E. Coupled two-way clustering analysis of gene microarray data *Proc Nat Acad Sci USA* 2000; 97: 12079-12084.
20. Simek K, Kimmel M. A note on estimation of dynamics of multiple gene expression based on singular value decomposition. *Mathematical Biosciences* 2002; 182: 183-199.
21. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *PNAS USA* 2001; 98 (9): 5116-5121 Harbor Laboratory. *Genome Research* 2000; 10: 1-3.
22. Fujarewicz K, Kimmel M, Rzeszowska-Wolny J, Świerniak A. A note on classification of gene expression data using support vector machines. *J Biol Sys* 2003; 11: 43-56.
23. Huang Y, Prasad M, Lemon WJ et al. Gene expression in papillary thyroid carcinoma reveals highly consistent profiles. *Proc Natl Acad Sci U S A*. 2001 Dec 18;98(26):15044-9.
24. Prasad ML, Pellegata NS, Kloos RT et al. CITED1 protein expression suggests Papillary Thyroid Carcinoma in high throughput tissue microarray-based study. *Thyroid* 2004; 14(3):169-175.
25. Mazzanti C, Zeiger MA, Costourous N et al. Using gene expression profiling to differentiate benign versus malignant thyroid tumors. *Cancer Res* 2004;64(8):2898-2903.
26. Takano T, Hasegawa Y, Matsuzuka F et al. Gene expression profiles in thyroid carcinomas. *Br J Cancer* 2000; 83(11):1495-1502.
27. Wasenius VM, Hemmer S, Kettunen E et al. Hepatocyte growth factor receptor, matrix metalloproteinase-11, tissue inhibitor of metalloproteinase-1, and fibronectin are up-regulated in papillary thyroid carcinoma: a cDNA and tissue microarray study. *Clin Cancer Res* 2003; 9(1):68-75.
28. Finley DJ, Arora N, Zhu B et al. Molecular profiling distinguishes papillary carcinoma from benign thyroid nodules. *J Clin Endocrinol Metab* 2004; 89(7):3214-3223.
29. Yano Y, Uematsu N, Yashiro T et al. Gene expression profiling identifies platelet-derived growth factor as a diagnostic molecular marker for papillary thyroid carcinoma. *Clin Cancer Res* 2004;10:2035-2043.
30. Inaba M, Sato H, Abe Y et al. Expression and significance of c-met protein in papillary thyroid carcinoma. *Tokai J Exp Clin Med* 2002; 27: 43-49.
31. Ramirez R, Hsu D, Patel A et al. Over-expression of hepatocyte growth factor/scatter factor (HGF/SF) and the HGF/SF receptor (cMET) are associated with a high risk of metastasis and recurrence for children and young adults with papillary thyroid carcinoma. *Clin Endocrinol (Oxf)* 2000;53:635-644.
32. Jarzab B, Wiench M, Fujarewicz K et al. Gene expression profile of papillary thyroid cancer: sources of variability and diagnostic implications (in press).
33. Aratake Y, Umeki K, Kiyoyama K et al. Diagnostic utility of galectin-3 and CD26/DPPIV as preoperative diagnostic markers for thyroid nodules. *Diagn Cytopathol* 2002;26:366-372.
34. Nagy N, Decaestecker C, Dong X et al. Characterization of ligands for galectins, natural galactoside-binding immunoglobulin G subfractions and sarcolectin and also of the expression of calyculin in thyroid lesions. *Histol Histopathol* 2000;15:503-513.
35. Niedziela M, Maceluch J, Korman E. Galectin-3 is not an universal marker of malignancy in thyroid nodular disease in children and adolescents. *J Clin Endocrinol Metab*. 2002 Sep;87(9):4411-5.
36. Finley DJ, Zhu B, Barden CB, Fahey TJ 3rd. Discrimination of benign and malignant thyroid nodules by molecular profiling. *Ann Surg*. 2004 Sep;240(3):425-36; discussion 436-7.
37. Brooks JD. Microarray analysis in prostate cancer research. *Current Opinion in Urology* 2002; 12: 395-399.
38. Dhanasekaran SM, Barrette TR, Ghosh D et al. Delineation of prognostic biomarkers in prostate cancer. *Nature* 2001; 412: 822-826.
39. Kitahara O, Furukawa Y, Tanaka T et al. Alternation of gene expression during colorectal carcinogenesis revealed cDNA microarrays after laser-capture microdissection of tumor tissues a normal epithelia. *Cancer Res* 2001; 1: 3544-3549.
40. Lin YM, Furukawa Y, Tsunoda T et al. Molecular diagnosis of colorectal tumors by expression profiles of 50 genes expressed differentially in adenomas and carcinomas. *Oncogene* 2002; 21: 4120-4128.
41. Notterman DA, Alon U, Sierk AJ, Levine AJ. Transcriptional gene expression profiles of colorectal adenoma, adenocarcinoma, and normal tissue examined by oligonucleotide arrays. *Cancer Res* 2001; 1: 3124-3130.
42. Vant't Veer LJ, Dai H, van de Vijver MJ et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* 2002; 415(6871): 530-6.
43. West M, Blanchette C, Dressman H et al. Predicting the clinical status of human breast cancer by using gene expression profiles. *PNAS USA* 2001; 98(20): 11462-11467.