

Instytut Chemii Bioorganicznej PAN
Zakład Biologii Molekularnej i Systemowej
Pracownia Mikromacierzy i Głębokiego Sekwencjonowania

*Zastosowanie mikromacierzowej hybrydyzacji międzygatunkowej
w badaniach transkryptomu *Nicotiana tabacum**

Rozprawa doktorska

Michał Góralski

Promotor: prof. dr hab. Marek Figlerowicz

Poznań 2014

*Pragnę serdecznie podziękować promotorowi,
Panu Profesorowi Markowi Figlerowiczowi za
opiekę naukową i cenne uwagi.*

*Gorące podziękowania kieruję do
Dr Agnieszki Żmieńko, na której wszechstronną
pomoc zawsze mogłem liczyć.*

SPIS TREŚCI

I. WSTĘP	7
I.1. MIKROMACIERZE DNA	8
I.1.1 PRZEBIEG EKSPERYMENTU MIKROMACIERZOWEGO	8
<i>I.1.1.2. Analiza bioinformatyczna</i>	12
I.1.2 TYPY MIKROMACIERZY DNA	13
<i>I.1.2.1. Mikromacierze cDNA</i>	14
<i>I.1.2.2. Mikromacierze L-oligo</i>	15
<i>I.1.2.3. Mikromacierze S-oligo</i>	15
I.2. MIKROMACIERZOWA HYBRYDYZACJA MIĘDZYGATUNKOWA	16
I.2.1 ZASTOSOWANIE CSH	19
<i>I.2.2.1. Wykorzystanie CSH w obrębie rodziny Solanaceae</i>	20
<i>I.2.2.2. Perspektywy wykorzystania CSH w kontekście intensywnego rozwoju genomiki</i>	21
I.2.2. SPECYFICZNOŚĆ HYBRYDYZACJI SONDA-PRÓBA	22
I.2.3. STRATEGIE EKSPERYMENTALNE STOSOWANE W CSH	25
<i>I.2.3.1. Dobór gatunku reprezentowanego na mikromacierzy</i>	26
<i>I.2.3.2. Wybór rodzaju mikromacierzy</i>	28
<i>I.2.3.3. Mikromacierzowe eksperymenty porównawcze</i>	29
I.2.4. WARUNKI EKSPERYMENTU CSH	30
I.3. ANALIZA DANYCH CSH	31
I.3.1. FILTRACJA NA PODSTAWIE PODOBIEŃSTWA SEKWENCJI NUKLEOTYDOWYCH	32
I.3.2. FILTRACJA NA PODSTAWIE PARAMETRÓW MORFOLOGICZNYCH PLAMEK	33
<i>I.3.2.1. Intensywność sygnału</i>	33
<i>I.3.2.2 Intensywność sygnału pochodzącego z hybrydyzacji z genomowym DNA</i>	34
<i>I.3.2.3 Parametry jakości plamek</i>	35
I.4. ODPOWIEDŹ ROŚLIN NA STRES SOLNY I KADMOWY - STAN BADAŃ TRANSKRYPTOMICZNYCH	37
I.5. PRZYSZŁOŚĆ TECHNIKI MIKROMACIERZOWEJ	39
II. CEL PRACY	40
III. MATERIAŁY I METODY	42
III.1. MATERIAŁY	43
III.1.1. MATERIAŁ ROŚLINNY	43
III.1.2. ZESTAWY DO PREPARATYKI KWASÓW NUKLEINOWYCH	43
III.1.3. ODCZYNNIKI CHEMICZNE	43
III.1.4. MIKROMACIERZE	44
III.1.5. BUFORY	44
III.1.6. APARATURA	44
III.1.7. OPROGRAMOWANIE	45
III.2. METODY	46
III. 2.1. IZOLACJA RNA	46
<i>III.2.1.1. Trawienie DNA przy użyciu zestawu TURBO DNA-free™ Kit</i>	46
<i>III.2.1.2. Wytrącanie RNA</i>	46
<i>III.2.1.3. Analiza ilościowa i jakościowa RNA</i>	47

<i>III.2.1.3.1 Pomiar stężenia RNA za pomocą aparatu NanoDrop ND-1000 (Thermo Scientific).</i>	47
<i>III.2.1.3.2. Elektroforeza kapilarna przy użyciu aparatu Bioanalyzer 2100 (Agilent)</i>	47
III.2.2. PRZYGOTOWANIE PRÓB DO HYBRYDYZACJI	47
<i>III.2.2.1. Metodyka stosowana w przypadku mikromacierzy TOM1, TOM2 i POT</i>	47
<i>III.2.2.1.1. Odwrotna transkrypcja</i>	47
<i>III.2.2.1.2. Znakowanie barwnikami fluorescencyjnymi</i>	47
<i>III.2.2.1.3. Oczyszczanie wyznakowanej próby</i>	48
<i>III.2.2.1.4. Przygotowanie mieszaniny hybrydizacyjnej</i>	48
<i>III.2.2.1.5. Dodanie modyfikowanych nukleotydów poli-A (tylko mikromacierze cDNA: POT i TOM1)</i>	49
<i>III.2.2.2. Metodyka stosowana w przypadku mikromacierzy TOB</i>	49
<i>III.2.2.2.1. Znakowanie barwnikami fluorescencyjnymi</i>	49
<i>III.2.2.2.2. Oczyszczanie wyznakowanej próby</i>	49
<i>III.2.2.2.3. Przygotowanie mieszaniny hybrydizacyjnej</i>	49
III.2.3. PRZYGOTOWANIE MIKROMACIERZY DO HYBRYDYZACJI-TYLKO MIKROMACIERZE TOM1, TOM2, POT	50
<i>III.2.3.1. Rehydratacja- tylko mikromacierze TOM2</i>	50
<i>III.2.3.2. Prehybrydizacja</i>	50
III.2.4. HYBRYDYZACJA	50
<i>III.2.4.1. Metodyka stosowana w przypadku mikromacierzy TOM1, TOM2 i POT</i>	50
<i>III.2.4.2. Metodyka stosowana w przypadku mikromacierzy TOB</i>	51
III.2.5. SKANOWANIE	51
<i>III.2.5.1. Skanowanie za pomocą skanera Scan Array Express (Perkin Elmer)</i>	51
<i>III.2.5.2. Skanowanie za pomocą skanera GenePix 4200 AL (Axon)</i>	51
III.2.6. METODY BIOINFORMATYCZNE	52
<i>III.2.6.1. Analiza obrazów mikromacierzy</i>	52
<i>III.2.6.1.1. Analiza obrazów mikromacierzy przy użyciu programu GenePixPro</i>	52
<i>III.2.6.1.2. Analiza obrazów mikromacierzy w programie Maia 2.75</i>	52
<i>III.2.6.2. Analiza homologii sond znajdujących się na mikromacierzach i ich docelowych transkryptów</i>	53
<i>III.2.6.3. Analizy bioinformatyczne danych mikromacierzowych</i>	54
IV. WYNIKI	55

IV.1. EKSPERYMENTY MIKROMACIERZOWE	56
IV.1.1. WYBÓR MODELU I MIKROMACIERZY DO BADAŃ PORÓWNAWCZYCH	56
IV.1.2. MAPOWANIE SOND POMIĘDZY RÓŻNYMI TYPAMI MIKROMACIERZY	59
IV.1.3. PRZYGOTOWANIE MATERIAŁU BIOLOGICZNEGO	61
V.1.4. ZNAKOWANIE I HYBRYDYZACJA MIKROMACIERZY	64
IV.1.5. OBRÓBKA OBRAZÓW I ANALIZA JAKOŚCIOWA MIKROMACIERZY.	68
IV.1.6. ANALIZA RÓŻNICOWA	76
IV.1.7. ANALIZA FUNKCJONALNA	79
IV.2. OCENA ZMIAN EKSPRESJI GENÓW W KORZENIACH TYTONIU PODDANYCH DZIAŁANIU STRESU SOLNEGO I METALI CIĘŻKICH	81
IV.2.1.1. OGÓLNY PROFIL EKSPRESJI GENÓW W KORZENIACH TYTONIU POD WPLYWEM STRESU ABIOTYCZNEGO	82
<i>IV.2.1.2. Profil ekspresji genów w korzeniach tytoniu poddanych działaniu stresu solnego</i>	91
<i>IV.2.1.3. Profil ekspresji genów w korzeniach tytoniu poddanych działaniu kadmu</i>	96

IV.3. MIKROMACIERZOWA HYBRYDYZACJA MIĘDZYGATUNKOWA- BADANIA PORÓWNAWCZE PROFILÓW EKSPRESJI GENÓW	102
IV.3.1. ANALIZA PORÓWNAWCZA EKSPERYMENTÓW CSH WYKONANYCH Z ZASTOSOWANIEM MIKROMACIERZY CDNA ORAZ L-OLIGO	118
IV.3.2. BADANIE WPLYWU POKREWIEŃSTWA GATUNKU BADANEGO I REPREZENTOWANEGO NA MIKROMACIERZY NA WYNIK ANALIZY EKSPRESJI GENÓW METODĄ CSH	122
IV.3.3. WPLYW DODATKOWYCH POWTÓRZEŃ TECHNICZNYCH NA WYNIKI CSH	126
IV.3.4. OPRACOWANIE ŚCIEŻKI FILTRACJI DANYCH	129
IV.3.4.1. Filtracja danych na podstawie kryteriów podobieństwa sekwencji nukleotydowej (filtracja HOM)	129
IV.3.4.2. Filtracja danych na podstawie kryteriów morfologii plamek (filtracja MORF)	132
V.DYSKUSJA	142
V.1. JAKOŚĆ UZYSKANYCH DANYCH	144
V.2. PROFIL EKSPRESJI GENÓW W KORZENIACH TYTONIU PODDANYCH DZIAŁANIU STRESU	145
V.2.1. OGÓLNY PROFIL EKSPRESJI GENÓW W KORZENIACH TYTONIU POD WPLYWEM STRESU ABIOTYCZNEGO	145
V.2.2. PROFIL EKSPRESJI GENÓW W KORZENIACH TYTONIU PODDANYCH DZIAŁANIU STRESU SOLNEGO.	147
V.2.3. PROFIL EKSPRESJI GENÓW W KORZENIACH TYTONIU PODDANYCH DZIAŁANIU KADMU.	149
V.3. ZASTOSOWANIE CSH W ANALIZIE TRANSKRYPTOMÓW	150
V.3.1. WPLYW DŁUGOŚCI SONDE NA JAKOŚĆ EKSPERYMENTÓW CSH	150
V.3.2. WPLYW POKREWIEŃSTWA GATUNKU BADANEGO I REPREZENTOWANEGO NA MIKROMACIERZY NA JAKOŚĆ CSH	152
V.3.3. WPLYW LICZBY ZASTOSOWANYCH POWTÓRZEŃ NA JAKOŚĆ CSH	154
V.3.4. WYBÓR OPTYMALNEJ ŚCIEŻKI FILTRACJI DANYCH	155
V.3.4.1. Filtracja HOM	155
V.3.4.2. Filtracja MORF	156
V.4. SKUTECZNOŚĆ CSH JAKO METODY BADANIA TRANSKRYPTOMU	157
V.5. OPTYMALNA ŚCIEŻKA ANALIZ CSH	161
V.6. ALTERNATYWNE ZASTOSOWANIA CSH	163
VI.WNIOSKI	165
VII.LITERATURA	167
STRESZCZENIE	176
ABSTRACT	178

Wykaz ważniejszych skrótów stosowanych w pracy

ABA- ang. *abscisic acid*- kwas abscysynowy

BLAST- ang. *Basic Local Alignment Search Tool* - narzędzie do poszukiwania lokalnych podobieństw pomiędzy sekwencjami

BSA- ang. *bovine serum albumin*- surowicza albumina wołowa

CSH- ang. *cross-species hybridization*- mikromacierzowa hybrydyzacja międzygatunkowa

EST- ang. *expressed sequence tag*- znaczniki sekwencji podlegających ekspresji

FDR- ang. *false discovery ratio*- odsetek wyników fałszywie pozytywnych

GO- ang. *Gene Ontology*- ontologia genów

L-oligo- długie oligonukleotydy

JA- ang. *jasmonic acid*- kwas jasmonowy

PCA- ang. *Principal Component Analysis*- Analiza głównych składowych

PCR- ang. *polymerase chain reaction*- reakcja łańcuchowa polimerazy

PMT- ang. *photomultiplier*- fotopowielacz

RIN- ang. *RNA Integrity Number*- miara integralności RNA o

RNAseq-ang. *RNA Sequencing*- sekwencjonowanie RNA

SNR- ang. *signal-to-noise ratio*- stosunek sygnału do szumu

SDS- ang. *sodium dodecyl sulfate*- dodecylosiarczan sodu

S-oligo- ang. *short oligonucleotides*- krótkie oligonukleotydy

SSC- ang. *sodium saline citrate buffer*- bufor cytrynianowy

SSH- ang. *single-species hybridization*- mikromacierzowa hybrydyzacja homologiczna

SEA-ang. *Singular Enrichment Analysis*- analiza pojedynczego wzbogacenia

I. WSTĘP

I.1. MIKROMACIERZE DNA

Mikromacierz DNA to zbiór wielu cząsteczek DNA (sond) związanych ze stałym podłożem i rozmieszczonych w ściśle określonym porządku. Doniesienia dotyczące pierwszych prób użycia mikromacierzy DNA pojawiły się na początku lat dziewięćdziesiątych XX wieku, kiedy to powstała potrzeba opracowania metody umożliwiającej jednoczesną analizę ekspresji dużej liczby genów [1,2]. Mikromacierze DNA okazały się być obiecującym narzędziem do badania transkryptomu i przez długi czas stanowiły technikę z wyboru w tego typu analizach. Pierwszy artykuł naukowy opisujący zastosowanie mikromacierzy w badaniach ekspresji genów opublikowano w roku 1995 [3]. Od tego czasu jesteśmy świadkami niesłabnącej popularności mikromacierzy DNA, wykorzystywanych nie tylko w analizie ekspresji genów, lecz także, dzięki postępowi technologicznemu i rosnącej liczbie modyfikacji i odmian tejże techniki, w innych badaniach [4-6], dotyczących np. alternatywnego składowania transkryptów [7,8], struktury genomu [9,10] czy oddziaływań DNA-białko [11,12].

I.1.1 Przebieg eksperymentu mikromacierzowego

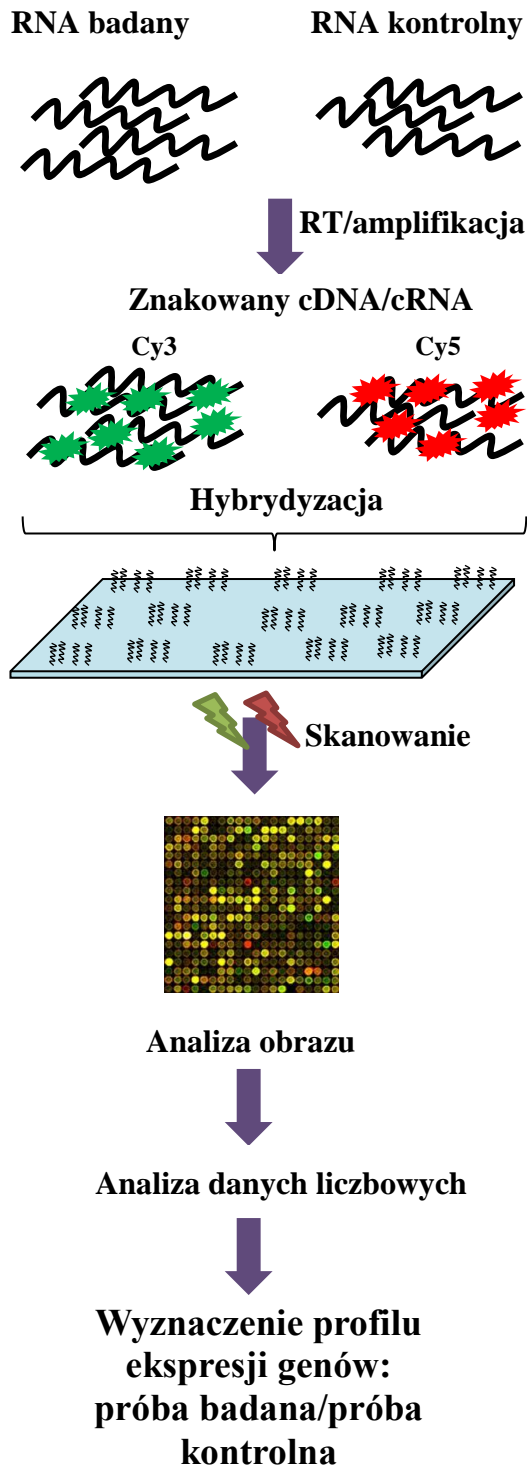
Mikromacierze DNA są techniką bazująca na zdolności komplementarnych cząsteczek kwasów nukleinowych do tworzenia struktur dwuniciowych. Proces ich powstawania nazywamy hybrydyzacją. W przypadku starszych metod wykorzystujących hybrydyzację, takich jak metoda Southerna (ang. Southern blot) lub metoda northern (ang. northern blot), unikalną cząsteczkę kwasu nukleinowego (sondę), wyznakowaną najczęściej za pomocą radioaktywnych izotopów, dodaje się do badanego kwasu nukleinowego (odpowiednio, DNA lub RNA), utrwalonego na nylonowej membranie (próbna). Wynik eksperymentu pozwala ocenić występowanie w badanej próbnie interesujących nas cząsteczek.

W przypadku mikromacierzy stosowana jest odmienna strategia (Rycina 1) - to cząsteczki kwasu nukleinowego zawarte w badanej próbnie (materiale pochodzącym np. z tkanki w określonym stanie fizjologicznym) są znakowane fluorescencyjnie. Próba poddawana jest hybrydyzacji z sondami, umieszczonymi na mikromacierzy. Po odmyciu niezwiązanej części próbny dokonuje się laserowego wzbudzenia sygnału fluorescencji tych cząsteczek, które uległy hybrydyzacji z elementami mikromacierzy. Informacja o sile sygnału jest utrwalona w postaci elektronicznego obrazu. Ten z kolei przetwarzany jest na

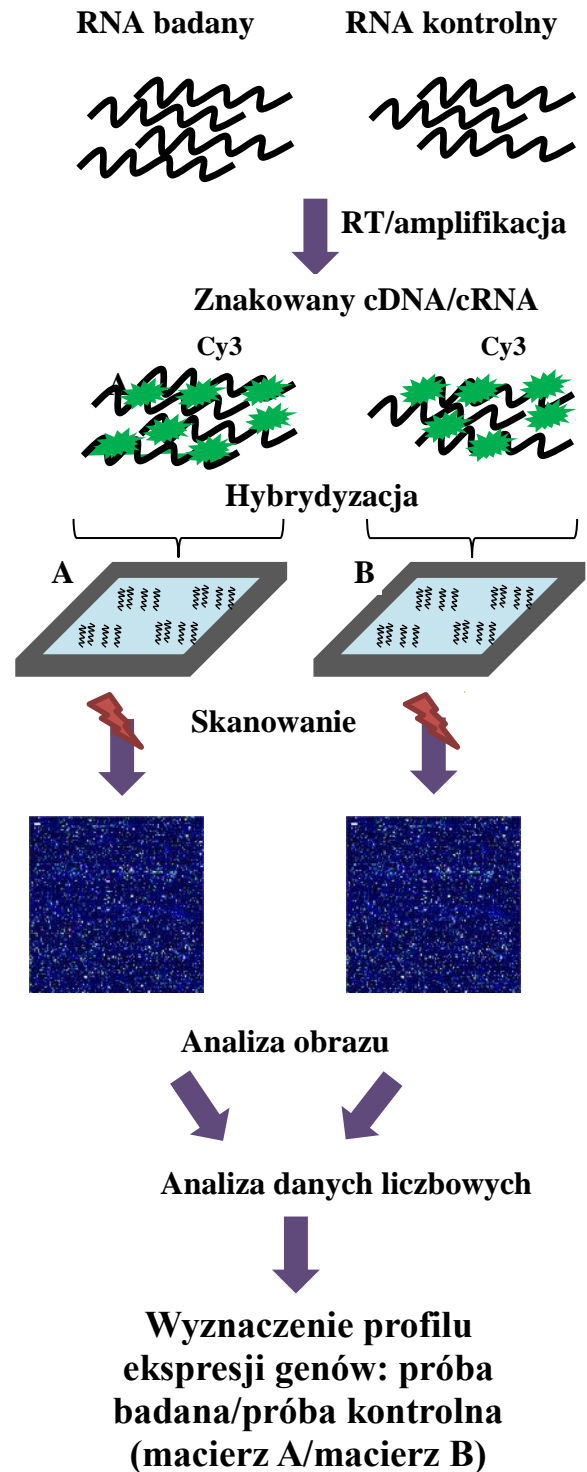
dane liczbowe określające względną zawartość kwasu nukleinowego o danej sekwencji w badanych próbach. W ten sposób możliwe jest jednoczesne badanie wielu rodzajów cząsteczek (np. transkryptów), reprezentowanych przez poszczególne sondy na mikromacierzy.

Mikromacierze najpowszechniej stosowane są do oceny ekspresji genów. Przeprowadzenie takiej analizy wymaga odpowiedniej ilości i jakości RNA, który posłuży do przygotowania próby [15,16]. W przypadku, gdy badacz dysponuje niewielką ilością materiału, skutecznym rozwiązaniem jest amplifikacja RNA, najczęściej za pomocą transkrypcji *in vitro* z zastosowaniem polimerazy T7, którą charakteryzuje wystarczająco duża dokładność i powtarzalność [16,18]. Znakowanie barwnikami fluorescencyjnymi można wykonać z zastosowaniem kilku metod (Rycina 2). W przypadku metody bezpośredniej, nukleotydy związane z fluoroforami wprowadzane są do cDNA powstającego w reakcji odwrotnej transkrypcji, bądź do cRNA amplifikowanego podczas transkrypcji *in vitro*. Metoda pośrednia polega na wprowadzeniu do syntetyzowanego cDNA lub cRNA modyfikowanych nukleotydów (np. amino-allilowych), do których w wyniku reakcji estryfikacji dołączane są fluorochromy. Możliwe jest również zastosowanie znakowania chemicznego (nieenzymatycznego), w którym barwniki (np. biotyna) przyłączają się bezpośrednio do kwasu nukleinowego. Podczas wyboru metody znakowania należy zwrócić uwagę na rodzaj stosowanych sond: np. w przypadku mikromacierzy oligonukleotydowych zawierających DNA sensowne, konieczne jest uzyskanie znakowanego produktu antysensownego [86]. Stosowane są dwa schematy eksperymentalne: z wykorzystaniem jednego lub dwóch barwników fluorescencyjnych (Rycina 1). W przypadku eksperymentu jednokolorowego, profil ekspresji dla każdej próby jest generowany na odrębnej mikromacierzy, po znakowaniu za pomocą pojedynczego barwnika fluorescencyjnego (np. fikoerytryny lub cyjaniny) [18]. Z kolei eksperyment dwukolorowy polega na połączeniu dwóch próbek znakowanych różnymi barwnikami (próba referencyjna i badana) w celu oceny ekspresji różnicowej.

I. Eksperyment dwukolorowy



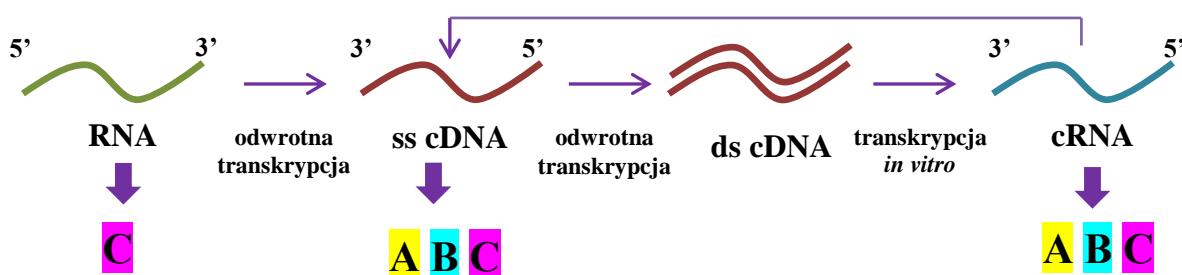
II. Eksperyment jednokolorowy



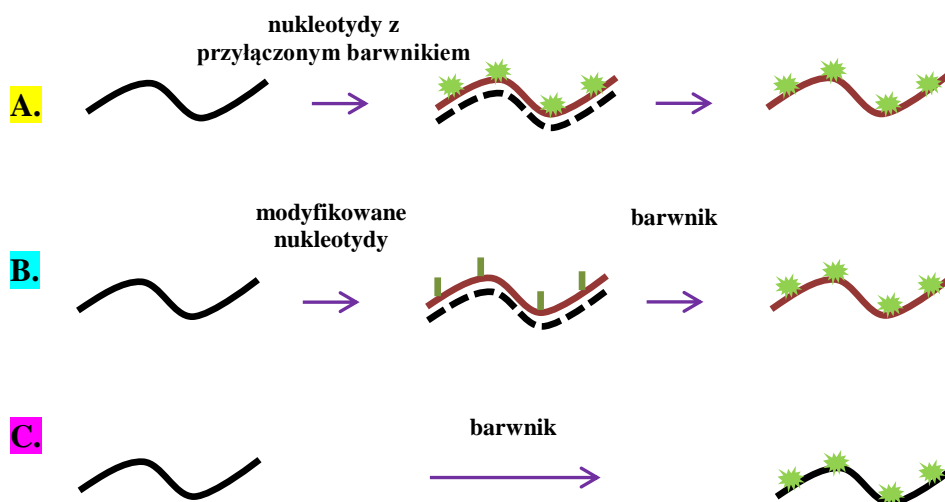
Rycina 1. Schemat przebiegu eksperymentu mikromacierzowego z wykorzystaniem dwóch (I) lub jednego (II) barwnika fluorescencyjnego.

Znakowanie za pomocą barwników fluorescencyjnych (np. cyjanin Cy3 i Cy5) o różnych parametrach wzbudzenia i emisji sygnału, umożliwia równoczesną hybryzację dwóch prób do pojedynczej mikromacierzy. Zakładając, że wpływ obu barwników na charakterystykę hybrydacji jest jednakowy, stosunki ilościowe poszczególnych transkryptów pierwotnie istniejących w próbce badanej i referencyjnej powinny być zachowane podczas hybrydacji do mikromacierzy [13,14,18].

I.



II.



Rycina 2. Możliwości i metody znakowania kwasów nukleinowych barwnikami fluorescencyjnymi podczas eksperymentów mikromacierzowych.

I- ścieżka postępowania podczas znakowania kwasów nukleinowych, za pomocą liter oznaczono możliwości zastosowania poszczególnych metod; II- metody znakowania: A- metoda enzymatyczna bezpośrednia, B- enzymatyczna pośrednia, C- chemiczna.

1.1.1.2. Analiza bioinformatyczna

Uzyskane w wyniku hybrydyzacji i skanowania „surowe dane” poddawane są analizie komputerowej, której celem jest wytypowanie w danej próbie zestawu genów cechujących się zmienionym poziomem ekspresji w stosunku do próby kontrolnej. Ta część analizy składa się z dwóch etapów: analizy wstępnej (ang. *preprocessing*) oraz analizy wyższego rzędu [27]. Podstawowym celem analizy wstępnej jest eliminacja zmian, których źródłem są przede wszystkim ograniczenia techniczne, wynikające z wieloetapowości i złożoności eksperymentu mikromacierzowego. Czynniki wprowadzającymi zmienność mogą być: procedura drukowania mikromacierzy, nierównomierna dystrybucja barwników, zaburzenia przebiegu hybrydyzacji oraz płukania, zanieczyszczenia szkiełka i wiele innych. Niektóre czynniki mogą wprowadzać systematyczne zaburzenia wielu prób [28]. Do przydatnych metod analizy wstępnej należą korekcja tła oraz normalizacja, po zastosowaniu których badacz powinien uzyskać jak najbardziej jednorodny zestaw danych, w którym maksymalnie ograniczona jest zmienność pochodząca z innych źródeł niż różnice biologiczne pomiędzy próbami. Weryfikację danych umożliwia szereg narzędzi diagnostycznych pozwalających przedstawić graficznie różne aspekty zmienności prób. Na tym etapie surowe dane dotyczące intensywności fluorescencji: dla kanału czerwonego (R, Cy5) oraz dla kanału zielonego (G, Cy3) (w przypadku mikromacierzy dwukolorowych) przekształcane są na wartości M i A, wyrażone poniższymi wzorami:

$$M = \log_2(R) - \log_2(G)$$

$$A = (\log_2(R) + \log_2(G))/2,$$

gdzie wartość M jest różnicą pomiędzy logarytmami intensywnościami fluorescencji na kanałach czerwonym i zielonym, obliczoną dla każdej sondy na danej mikromacierzy, a wartość A- średnią wartością logarytmów fluorescencji na obu kanałach.

Celem analizy wyższego rzędu jest wytypowanie genów charakteryzujących się ekspresją różnicową, a także pomoc w określeniu znaczenia biologicznego wykrytych różnic. Etap ten obejmuje ocenę ekspresji różnicowej z zastosowaniem odpowiedniego modelu statystycznego oraz porównań wielokrotnych. Efektem jest uzyskanie danych dotyczących różnic w ekspresji genów pomiędzy próbami oraz szeregu informacji określających prawdopodobieństwo, znaczenie statystyczne i wiarygodność wyników, np. dane na temat udziału wyników fałszywie pozytywnych (FDR, ang. *false discovery ratio*) [28]. Do najważniejszych parametrów obliczanych podczas analizy różnicowej należą: logFC oraz zmodyfikowana wartość p. Wartość logFC jest logarytmem o podstawie 2 ze

stosunku ekspresji genów próby badanej i referencyjnej. Należy pamiętać, że logFC jest wartością poddaną logarytmowaniu, zatem uzyskanie wartości 1 oznacza dwukrotny wzrost, a wartości -1- dwukrotne zmniejszenie ekspresji genu. Z kolei zmodyfikowana wartość p (ang. *adjusted P value*) jest miarą istotności statystycznej obliczeń i określa prawdopodobieństwo przypadkowego pojawienia się danej zmiany poziomu ekspresji. Po filtracji genów na podstawie kryteriów wyznaczonych przez badacza możliwe są dalsze etapy analizy, takie jak analiza skupień czy konstruowanie modeli interakcji bądź procesów biologicznych.

Warto przy tym zauważyć, że ze względu na dużą liczbę czynników zaburzających przebieg eksperymentu mikromacierzowego, uzyskiwane tą drogą wyniki zwykle wymagają walidacji za pomocą metod o większej czułości. Narzędziem standardowo stosowanym do walidacji wyników uzyskanych za pomocą mikromacierzy jest ilościowy PCR (real-time PCR, qPCR) [41,42].

I.1.2 Typy mikromacierzy DNA

Rozróżnienia rodzajów mikromacierzy DNA stosowanych w badaniach ekspresji genów możemy dokonać na podstawie długości i typu zastosowanych sond. Biorąc pod uwagę ten parametr możemy wyróżnić: (i) mikromacierze o krótkich sondach oligonukleotydowych (nazywane dalej mikromacierzami typu S-oligo; sondy 15-25 nukleotydowe); (ii) mikromacierze o długich sondach oligonukleotydowych (określane dalej jako L-oligo; sondy 50-120 nukleotydowe) oraz (iii) mikromacierze o sondach, które stanowi dwuniciowy cDNA (mikromacierze cDNA; sondy od stu do kilku tysięcy par zasad) [19]. Rodzaj zastosowanych sond determinuje szereg cech danej mikromacierzy, które należy wziąć pod uwagę przed rozpoczęciem eksperymentu (Tabela 1).

Tabela 1. Charakterystyka poszczególnych typów mikromacierzy DNA.

Mikromacierz	cDNA	L-oligo	S-oligo
Źródło sond	Klony cDNA	Synteza chemiczna	Synteza chemiczna
Długość sond	Zróznicowana (kilkaset-kilka tysięcy pz)	50-120 nt	15-25 nt
Pokrycie genów	Średnie	Duże	Duże
Specyficzność sond	Średnia	Duża	Duża
Dane dot. sekwencji sondy na etapie produkcji mikromacierzy	Niekonieczne	Wymagane	Wymagane
Koszt uzyskania sond	Mały	Średni	Duży
Schemat detekcji	Dwukolorowy	Dwukolorowy	Jednokolorowy

1.1.2.1. Mikromacierze cDNA

Jest to najwcześniej opracowany typ mikromacierzy, do produkcji której wykorzystuje się dwuniciowe sondy będące klonami cDNA amplifikowanymi w reakcji PCR. Klony stanowiące matrycę w reakcji zazwyczaj pochodzą z biblioteki klonów cDNA o sekwencji poznanej całkowicie lub częściowo (tzw. sekwencje EST, ang. *Expressed Sequence Tag*). Znajomość pełnej sekwencji sond nie jest warunkiem koniecznym do produkcji mikromacierzy cDNA, ponieważ są one zwykle amplifikowane przy pomocy uniwersalnych starterów. Sondy nanoszone są na podłoże stałe (płytkę) w sposób kontaktowy - za pomocą igieł drukarki, lub bezkontaktowy – przy użyciu techniki podobnej do stosowanej w drukarkach atramentowych. Mikromacierze cDNA zyskały znaczną popularność ze względu na możliwość uzyskania dużej ilości sond w standardowo wyposażonym laboratorium, a tym samym ich niską cenę i dużą dostępność. Dlatego są one szczególnie chętnie stosowane w przypadku małych projektów. Rodzaj sond determinuje wady i zalety tej technologii: z jednej strony, długie sondy cDNA generują sygnał o dużej intensywności, z drugiej - są bardziej podatne na hybrydyzację krzyżową, czyli wiązanie się z sondą cząsteczek cDNA niestanowiących sekwencji docelowych dla sondy, co obniża ich specyficzność [22]. Ten ostatni czynnik utrudnia również analizę rodzin wielogenowych. Przykładowo, w pracy dotyczącej analizy ekspresji genów należących do rodziny cytochromów P450 u *Arabidopsis thaliana* wykazano, że w

przypadku sekwencji wykazujących ponad 80% identyczności odsetek hybrydyzacji krzyżowej wynosi powyżej 20% [87]. Innym problemem może być niemożliwość rozróżnienia transkrypcji z nici plus i minus i generowanie wspólnego sygnału fluorescencyjnego dla transkryptów sensownych i antysensownych. Istotną kwestią jest również duża liczba zanieczyszczonych lub błędnie scharakteryzowanych sond znajdujących się na mikromacierzy (które można zidentyfikować wyłącznie poprzez powtórne sekwencjonowanie matrycowych klonów cDNA) [21,22,23].

1.1.2.2. Mikromacierze L-oligo

Mikromacierze zawierające długie sondy oligonukleotydowe łączą szereg zalet mikromacierzy cDNA i opisanych poniżej S-oligo. Zastosowanie sond oligonukleotydowych zapewnia specyficzną hybrydyzację i eliminuje problemy związane z potencjalną detekcją transkryptów antysensownych. Z kolei długość sond gwarantuje silny sygnał hybrydyzacji oraz możliwość wyprodukowania mikromacierzy metodą „drukowania” przy pomocy takiej samej aparatury, jak w przypadku mikromacierzy cDNA. Do zaprojektowania sond L-oligo, podobnie jak w przypadku S-oligo, niezbędna jest informacja dotycząca docelowej sekwencji nukleotydowej. Spośród licznych dostępnych mikromacierzy oligonukleotydowych warto wymienić mikromacierze firmy Agilent, w których 60-nukleotydowe sondy syntetyzowane są *in situ* z wykorzystaniem technologii bezkontaktowego nanoszenia na płytkę [19,21,22]. Ciekawa jest również oferta firmy Illumina- BeadChip® to technologia, w której oligonukleotydy o długości 50 nt połączone są z ziarnami umieszczonymi w dołkach na płytce silikonowej. Każda z sond występuje w 30 powtórzeniach, co umożliwia dokładniejszą analizę statystyczną wyników [25].

1.1.2.3. Mikromacierze S-oligo

Tego typu mikromacierze kojarzone są głównie z firmą Affymetrix i znane również pod nazwą chipów DNA (ang. *Gene Chip*). Wykorzystywana w chipach DNA technologia syntezy oligonukleotydów *in situ* z zastosowaniem fotolitografii umożliwia gęste upakowanie sond na płytce kwarcowej. Danemu transkryptowi odpowiada zestaw sond (ang. *probe set*) składający się z 11-22 różnych oligonukleotydów o długości 25 nt [21,22]. W przypadku mikromacierzy do badania ekspresji genów, przez długi czas sondy obejmowały eksony położone na końcu 3' danego genu. Od kilku lat stosuje się również

mikromacierze eksonowe, w których zaprojektowano po 4 sondy przypadające na każdy ekson genu [88]. W celu detekcji i kontrolowania hybrydyzacji krzyżowej, każda sonda o pełnej komplementarności do danego transkryptu (sonda PM, ang. *perfect match*) posiada swój odpowiednik ze zmienionym pojedynczym nukleotydem - taka sonda wykazuje niepełną komplementarność do transkryptu (sonda MM, ang. *mismatch*) [24]. Eksperymenty wykonywane są metodą jednokolorową, zatem nie dają możliwości bezpośredniego porównania prób podczas hybrydyzacji. Mikromacierze S-oligo wymagają zastosowania aparatury specyficznej dla tej technologii na każdym etapie eksperymentu.

I.2. MIKROMACIERZOWA HYBRYDYZACJA MIĘDZYGATUNKOWA

Szybki rozwój technik wielkoskalowego sekwencjonowania umożliwił w ostatnich latach poznanie wielu sekwencji genomowych, jednak liczba ta nadal jest ograniczona do organizmów modelowych i gatunków ważnych z punktu widzenia medycznego i komercyjnego [34]. Do tej pory opublikowano pełne sekwencje genomowe ponad 50 roślin wyższych (Tabela 2). Komercyjne mikromacierze stworzono do tej pory dla 19 gatunków roślin (Tabela 3). Firmy zajmujące się produkcją mikromacierzy (np. Affymetrix, Agilent), świadczą usługi obejmujące umieszczenie na mikromacierzy sond o sekwencjach wskazanych przez klienta, należy jednak pamiętać, iż zaprojektowanie mikromacierzy związane jest z olbrzymim nakładem pracy i bardzo często ograniczone brakiem dostępności odpowiedniej liczby sekwencji [31,33]. Biorąc pod uwagę powyższe względy wielu badaczy zainteresowało się wykorzystaniem istniejących już mikromacierzy do badania innych transkryptomów niż te, dla których zostały one stworzone. Wiadomo bowiem, że wiele organizmów wykazuje znaczny stopień homologii pomiędzy sekwencjami kodującymi, np. dla genomu człowieka i szympansa wynosi on 98% [89]. Z podobną sytuacją spotykamy się u roślin. Podobieństwa na poziomie sekwencji genomowych, umożliwiające identyfikację genów na podstawie homologii, wykazano w rodzinach *Poaceae*, *Solanaceae*, *Brassicaceae*, *Fabaceae* oraz *Pinaceae* [53]. Stosunkowo wysoki jest również poziom podobieństwa sekwencji kodujących geny pomiędzy rodzinami. Przykładowo, szacuje się, że 90% genów rzodkiewnika posiada homologię w genomie ryżu [32]. Na podstawie tych przesłanek wielokrotnie podejmowano próby zastosowania mikromacierzy DNA w badaniach organizmów innych niż reprezentowany na mikromacierzy [37-40, 45-59]. Technika taka nazwana została mikromacierzową hybrydyzacją międzygatunkową (CSH, ang. *cross-species hybridization*) lub hybrydyzacją heterologiczną (ang. *heterologous hybridization*), w odróżnieniu od hybrydyzacji

homologicznej (SSH, ang. *single-species hybridization*).

Tabela 2. Wybrane gatunki roślin wyższych o znanych sekwencjach genomowych, na podstawie: <http://synten.cnr.berkeley.edu> (University of California, Berkeley).

Gatunek	Sekwencja	Publikacja
Banan	Niepełna	D'Hont, A. i wsp. (2012) Nature, 9;488:213-7.
Brzoskwinia	Pełna	Ahmad, R i wsp. (2011) BMC Genomics, 22;12:569
Daktylowiec	Niepełna	-
Eukaliptus	Pełna	-
Jabłoń	Pełna	Velasco, R. i wsp. (2010) Nat Genet. 42(10):833-9
Kakaowiec	Pełna	Argout, X. i wsp. (2011) Nat Genet 43, 2, 101-108
Kłosownica	Pełna	Vogel, J. i wsp.(2010) Nature, 463, 763-768
Kroplik	Niepełna	-
Kukurydza	Pełna	Schnable, P. i wsp. (2009), Science, 326, 1112-1115
Maniok	Niepełna	-
<i>Medicago truncatula</i>	Pełna	Young, N. i wsp. (2011) Nature 16, 480, 520-4
Ogórek	Pełna	Huang, S. i wsp. (2009) Nat Genet, 41(12):1275-81
Orlik	Pełna	-
Papaja	Pełna	Ming, R. I wsp. (2008) Nature, 452 (7190), 991-996
Pomidor	Niepełna	Tomato Genome Consortium (2012) Nature 30;485 :635-41
Poziomka	Pełna	Shulaev, V. i wsp. (2011), Nat Genet 43, 2 ,109-116
Rącznik	Pełna	Chan, P. i wsp.(2010) Nat Biotechnol.;28(9):951-6.
Ryż	Pełna	Goff, S. (2002) Science, 296 (5565), 92-100
Rzodkiewnik	Pełna	The Arabidopsis Genome Initiative (2000), Nature, 408, 796-815
Rzodkiewnik	Pełna	Hu, T. i wsp. (2011) Nat Genet 43,476-481
Soja	Pełna	Schmutz, J. i wsp. (2010) Nature, 463 (7278), 178-183
Sorgo	Pełna	Paterson A. i wsp. (2009) Nature, 457 (7229), 551-556
Topola	Pełna	Tuskan, G. i wsp. (2006), Science, 313 (5793), 1596-1604
Winogrono	Pełna	Jaillon, O. i wsp. (2007), Nature, 449 (7161), 463-467
Włośnica ber	Pełna	-
Ziemniak	Niepełna	Potato Genome Sequencing Consortium (2011),Nature,10;475:189-95

Tabela 3. Dostępne komercyjnie mikromacierze dla roślin, wg: [30], www.affymetrix.com, www.nimblegen.com, www.agilent.com

Gatunek	Rodzaj mikromacierzy	Producent
Bawełna	S-oligo	Affymetrix
	L-oligo	Agilent
	S-oligo	Affymetrix
	L-oligo	Agilent
Brzoskwinia	L-oligo	Microarrays Inc
Groszek	L-oligo	Microarrays Inc
Jęczmień	S-oligo	Affymetrix
	L-oligo	Agilent
Kapusta	L-oligo	Agilent
Kukurydza	S-oligo	Affymetrix
	L-oligo	University of Arizona
<i>Medicago truncatula</i>	L-oligo	Agilent
	S-oligo	Affymetrix
	L-oligo	Microarrays Inc
Pomidor	S-oligo	Affymetrix
	L-oligo	Agilent
	L-oligo	Boyce Thompson Institute for Plant Research
	cDNA	Boyce Thompson Institute for Plant Research
Pszenica	S-oligo	Affymetrix
	L-oligo	Agilent
Ryż	L-oligo	Agilent
	L-oligo	University of Arizona
Rzodkiewnik	S-oligo	Affymetrix
	L-oligo	Agilent
	L-oligo	Nimblegen
	L-oligo	University of Arizona
Soja	S-oligo	Affymetrix
	cDNA	University of Illinois
Topola	S-oligo	Affymetrix
	L-oligo	Nimblegen
Trzcina	S-oligo	Affymetrix
Tytoń	L-oligo	Agilent
Winogrono	S-oligo	Affymetrix
	L-oligo	The Institute for Genomic Research (TIGR)
Ziemiak	cDNA	The Institute for Genomic Research (TIGR)

I.2.1 Zastosowanie CSH

Technika CSH wykorzystywana jest najczęściej w badaniach ekspresji genów organizmów, dla których informacja dotycząca sekwencji jest niepełna, bądź w badaniach porównawczych, oceniających ewolucyjnie konserwatywne mechanizmy kontroli ekspresji genów u pokrewnych organizmów [54]. Opublikowane do tej pory prace opisują zastosowanie CSH z wykorzystaniem wszystkich typów mikromacierzy- S-oligo [37,62,63,66], L-oligo [58,85,86] lub cDNA [40,51,53,55]. Przykładem zastosowania CSH w badaniach ekspresji genów mogą być analizy profilów ekspresji na różnych stopniach rozwoju gatunku much *Drosophila simulans*, wykorzystujące mikromacierze dla *D. melanogaster*. W badaniach tych wykazano, że różnice w profilu ekspresji pomiędzy dwoma gatunkami sięgały 27% [48]. W innych badaniach oceniano zmiany ekspresji genów u szympanсів zakażonych wirusem zapalenia wątroby typu C, z wykorzystaniem mikromacierzy dla człowieka. Stwierdzono m.in. znaczny wzrost ekspresji genów związanych z odpowiedzią na interferon już w ciągu 2 dni od infekcji [49]. Spośród badań wykonanych na roślinach, oprócz prac w obrębie rodziny *Solanaceae* (omówionych szerzej w rozdziale I.2.2.1), na uwagę zasługuje wykorzystanie mikromacierzy cDNA dla rzodkiewnika do badań gatunków pokrewnych. Przeprowadzono m.in. ocenę ekspresji genów u roślin rzepaku zakażonych patogenem *Alternaria brassicola*. Wśród zidentyfikowanych transkryptów charakteryzujących się ekspresją różnicową znalazły się takie, które wykazywały podobieństwo do znanych genów związanych z obroną, usuwaniem wolnych rodników czy odpornością na choroby [32]. Innym przykładem jest ocena profilu ekspresji genów związanych z odpowiedzią na wzrost stężenia metali w podłożu u *Arabidopsis halleri*, roślinie o dużo większej odporności na te czynniki niż *A. thaliana*. Ustalono między innymi, że strategia tolerancji u *A. halleri* związana jest z dużym konstytutywnym poziomem ekspresji genów regulujących homeostazę jonów metali [90]. Mimo, iż wielu autorów twierdzi, że za pomocą CSH można uzyskać istotne biologicznie informacje [40,48,50,56,59], a część przyznaje, że jest to możliwe po zastosowaniu odpowiedniej filtracji [37,62], istnieją prace sygnalizujące możliwość generowania fałszywie pozytywnych wyników nawet u blisko spokrewnionych gatunków [50,55] bądź występowania rozbieżności pomiędzy wynikami hybrydyzacji tych samych prób do różnych mikromacierzy CSH [54].

1.2.2.1. Wykorzystanie CSH w obrębie rodziny Solanaceae

U roślin z rodziny *Solanaceae* (psiankowate) wykonano szczególnie dużo eksperymentów CSH. Jest to związane z dostępnością kilku mikromacierzy opracowanych dla tej rodziny, istotnym znaczeniem gospodarczym wielu gatunków *Solanaceae*, a także ich dużym pokrewieństwem filogenetycznym, usprawiedliwiającym zastosowanie tego typu podejścia badawczego. Psiankowate to duża, zróżnicowana rodzina roślin, licząca powyżej 3000 gatunków. W jej skład wchodzi co najmniej 18 roślin o dużym znaczeniu gospodarczym (m.in. bakłażan, ziemniak, pomidor, papryka, tytoń) [53]. Moore i wsp. porównywali profile ekspresji genów związanych z dojrzewaniem owoców u trzech gatunków *Solanaceae* na mikromacierzy cDNA dla pomidora, w układzie CSH (papryka, bakłażan) i SSH (pomidor). Przed hybrydyzacjami przeprowadzono porównanie sekwencji badanych gatunków, stwierdzając, że >75% dostępnych sekwencji dla papryki i bakłażana posiada swoje homologi u pomidora. Podczas hybrydyzacji wykrywalny sygnał stwierdzono dla pomidora w 61% sond, dla bakłażana w 52% sond, i dla papryki w 45% sond. Badania pozwoliły wykryć wiele genów homologicznych dla 3 gatunków, między innymi związanych z fotosyntezą i fotooddychaniem, jak również wytypować procesy przebiegające w odmienny sposób u badanych gatunków [53]. Z kolei mikromacierz S-oligo dla pomidora wykorzystano między innymi do badania zjawiska indukowanej niską temperaturą degradacji skrobi w bulwach ziemniaków na rzecz syntezy sacharozy. Badanie umożliwiło wytypowanie szeregu genów związanych z odpowiedzią na obniżoną temperaturę [38]. Mikromacierz cDNA dla ziemniaka wykorzystano do analizy ekspresji genów u kilku gatunków *Solanaceae* (ziemniak, pomidor, tytoń, papryka, bakłażan, *Nicotiana benthamiana*, petunia), poddanych działaniu różnych czynników stresowych (zasolenie, wysoka i niska temperatura). Wyniki grupowania genów różnicujących wskazują na rozbieżności międzygatunkowe w profilach ekspresji genów związanych z odpowiedzią na stres niskiej temperatury, natomiast profile dotyczące stresu solnego były podobne u wszystkich badanych gatunków. Głównym celem tego badania była jednak ocena samej istoty hybrydyzacji międzygatunkowej, co zostało szczegółowo omówione w dalszej części niniejszej pracy (rozdział 1.2.3) [52]. W innym eksperymencie, tę samą mikromacierz cDNA dla ziemniaka wykorzystano do oceny profilu ekspresji genów *N. benthamiana* podczas zakażenia dwoma rodzajami wirusa: SYNV (ang. *Sonchus yellow net virus*) i INSV (ang. *Impatiens necrotic spot virus*). Stwierdzono istotne różnice w ekspresji genów w przebiegu zakażenia tymi wirusami [82].

CSH stosuje się również do oceny ekspresji genów u mniej znanych gatunków *Solanaceae*. Przykładem jest analiza transkryptomyczna szlaków metabolicznych w płatkach brunfelsji (*Brunfelsia calycina*) po otwarciu kwiatów, z wykorzystaniem mikromacierzy cDNA dla ziemniaka [84].

1.2.2.2. Perspektywy wykorzystania CSH w kontekście intensywnego rozwoju genomiki

Pomimo braku precyzyjnych wytycznych dotyczących stosowania CSH w badaniach genomicznych oraz opracowania nowych metod analizy genomu (por. rozdział I.5), istnieją obszary nauk biologicznych, w których technika ta może być z powodzeniem stosowana. Dotyczy to w szczególności badań porównawczych, oceniających podobieństwa i różnice pomiędzy gatunkami pokrewnymi. Mimo rozwoju i większej dostępności techniki, sekwencjonowanie każdego gatunku interesującego badacza jest trudne do wykonania z ekonomicznego punktu widzenia, szczególnie w przypadku, gdy badanie dotyczy gatunku o bardzo zbliżonej sekwencji genomowej do organizmu o poznanym genomie czy transkryptomie. W szczególności może to dotyczyć organizmów nieuznawanych za modelowe i niebędących gatunkami o istotnym znaczeniu gospodarczym. W przypadku stosowania CSH dla gatunku o bardzo bliskim pokrewieństwie z organizmem reprezentowanym na mikromacierzy możliwe jest wykonanie badań wstępnych mających na celu wytypowanie, które szlaki przemian zachodzą w podobny sposób w obu przypadkach oraz gdzie należy szukać potencjalnych różnic w późniejszych, dokładniejszych badaniach przy użyciu nowocześniejszych, ale droższych technik. Tego typu badania mogą być pomocne w uzyskiwaniu odpowiedzi na pytania z zakresu ewolucji, filogenezy czy ekologii. Takie podejście, poszukujące podobieństw profilów ekspresji genów pomiędzy różnymi gatunkami jest popularne w badaniach genomicznych, a rozwój baz danych i technik informatycznych i statystycznych sprzyja tego typu porównaniom [202,203]. Przykładowo, w pracy porównującej dane o ekspresji genów, uzyskane za pomocą ponad 3000 mikromacierzy dla człowieka, muszki owocowej (*D.melanogaster*), nicienia (*C.elegans*) oraz drożdży (*S. cerevisiae*), wyselekcjonowano 3400 genów wykazujących koekspresję, zaangażowanych w ponad 22 tys. interakcji. Uzyskane dane umożliwiły m.in. identyfikację funkcji niektórych genów związanych z proliferacją komórek [202]. W przypadku CSH, uzasadnione wydaje się użycie mikromacierzy heterologicznych do analizy genomu/transkryptomu pozwalającej zasygnalizować podobieństwa i wskazać procesy wymagające dokładniejszych badań. Tego typu badania, wykorzystujące istniejące dane SSH lub wykonywane równolegle z

takimi hybrydyzacjami pozwalają zminimalizować błędy eksperymentalne dzięki zastosowaniu tych samych warunków (mikromacierzy, hybrydyzacji, skanowania) dla obu badanych gatunków. Dodatkowym aspektem takiego podejścia może być poszukiwanie nowych gatunków modelowych, o odmiennych profilach ekspresji. Wydaje się, że takie zastosowanie, mające na celu poszukiwanie zmian raczej w grupach funkcjonalnych niż dotyczących selekcji pojedynczych genów ulegających zmianom ekspresji, może stanowić najbardziej odpowiedni obszar potencjalnego wykorzystania mikromacierzowej hybrydyzacji międzygatunkowej.

I.2.2. Specyficzność hybrydyzacji sonda-próba

Specyficzność hybrydyzacji, czyli zdolność sondy do wiązania tylko jednego rodzaju transkryptu, jest kwestią często poruszaną w kontekście mikromacierzy. Wyróżnić możemy cztery rodzaje relacji pomiędzy sondami a transkryptami (Rycina 3) [43]:

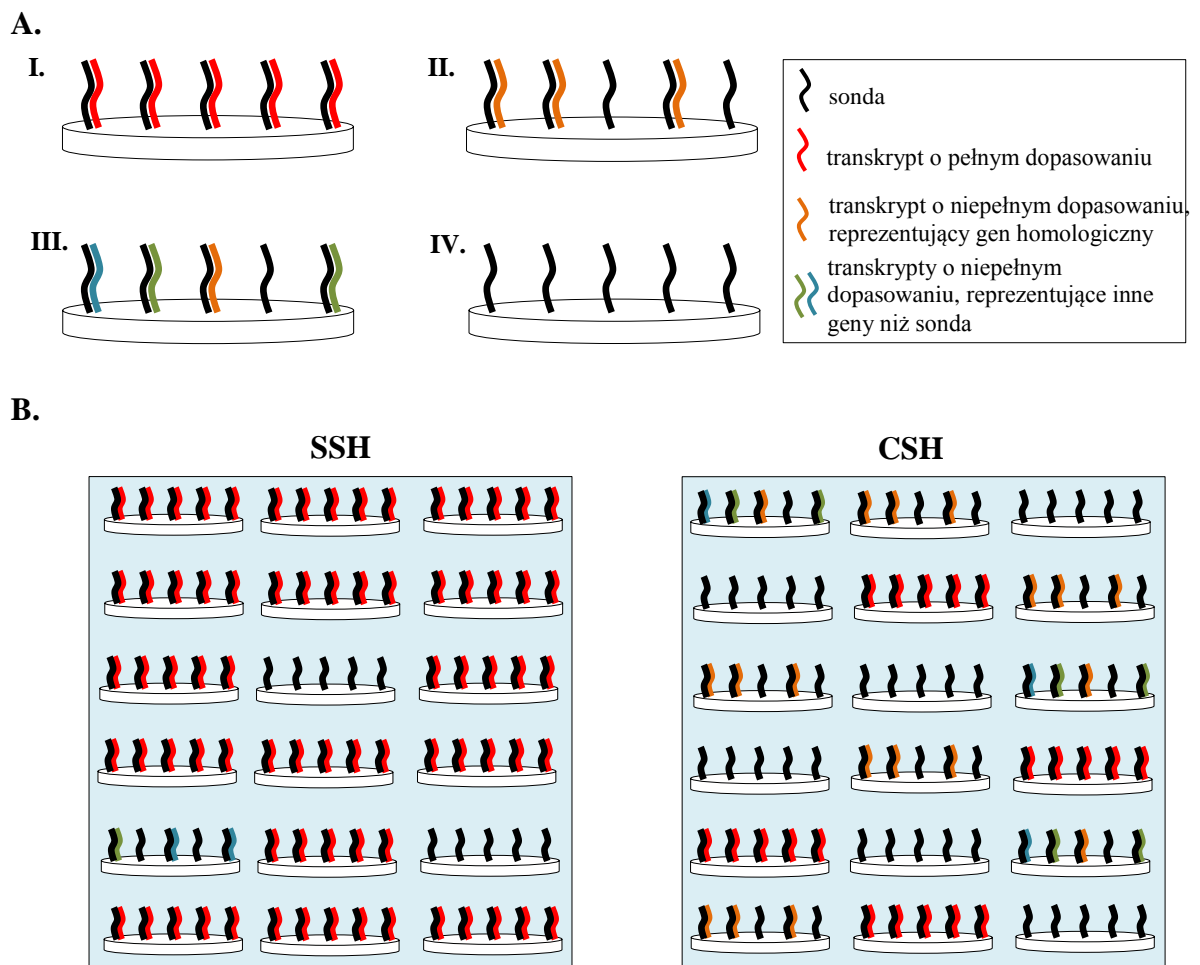
- I. Pełne dopasowanie- sonda wiąże się z transkrytem na całej długości
- II. Częściowe dopasowanie- wiązanie się fragmentu sondy z transkrytem
- III. Dopasowanie kilku transkryptów do sondy- wiązanie się sondy z więcej niż jednym transkrytem
- IV. Brak dopasowania- brak wiązania transkryptu z sondą

W wielu pracach zwrócono uwagę na zmniejszoną średnią intensywność sygnału podczas hybrydyzacji CSH w porównaniu z SSH [31,40,54]. Mimo to zaledwie w jednej pracy podjęto próbę wyjaśnienia mechanizmu tych różnic w obrazie mikromacierzy [52]. Zaproponowano, że podczas następującego po hybrydyzacji płukania mikromacierzy, transkrypty o słabym dopasowaniu oddysocjują od sond, czego efektem jest obecność pustych sond (Rycina 4.). Tego rodzaju transkrypty występują w dużo większej liczbie w CSH, dlatego w tego rodzaju eksperymentach powyższy proces występuje znacznie częściej, niż w przypadku SSH. Daną plamkę po skanowaniu tworzy wiele pikseli, z których każdy będzie miarą intensywności sygnału hybrydyzacyjnego dla wielu identycznych sond. Dlatego odpłukiwanie transkryptów może skutkować powstaniem piksela o zmniejszonym sygnale, który jest fałszywie zakwalifikowany jako tło. Wpływ takich zmian na pojedynczą plamkę może być przyczyną:

- zmniejszenia intensywności sygnału
- nierównomiernego rozkładu sygnału w obrębie plamki
- zaburzeń kształtu plamki

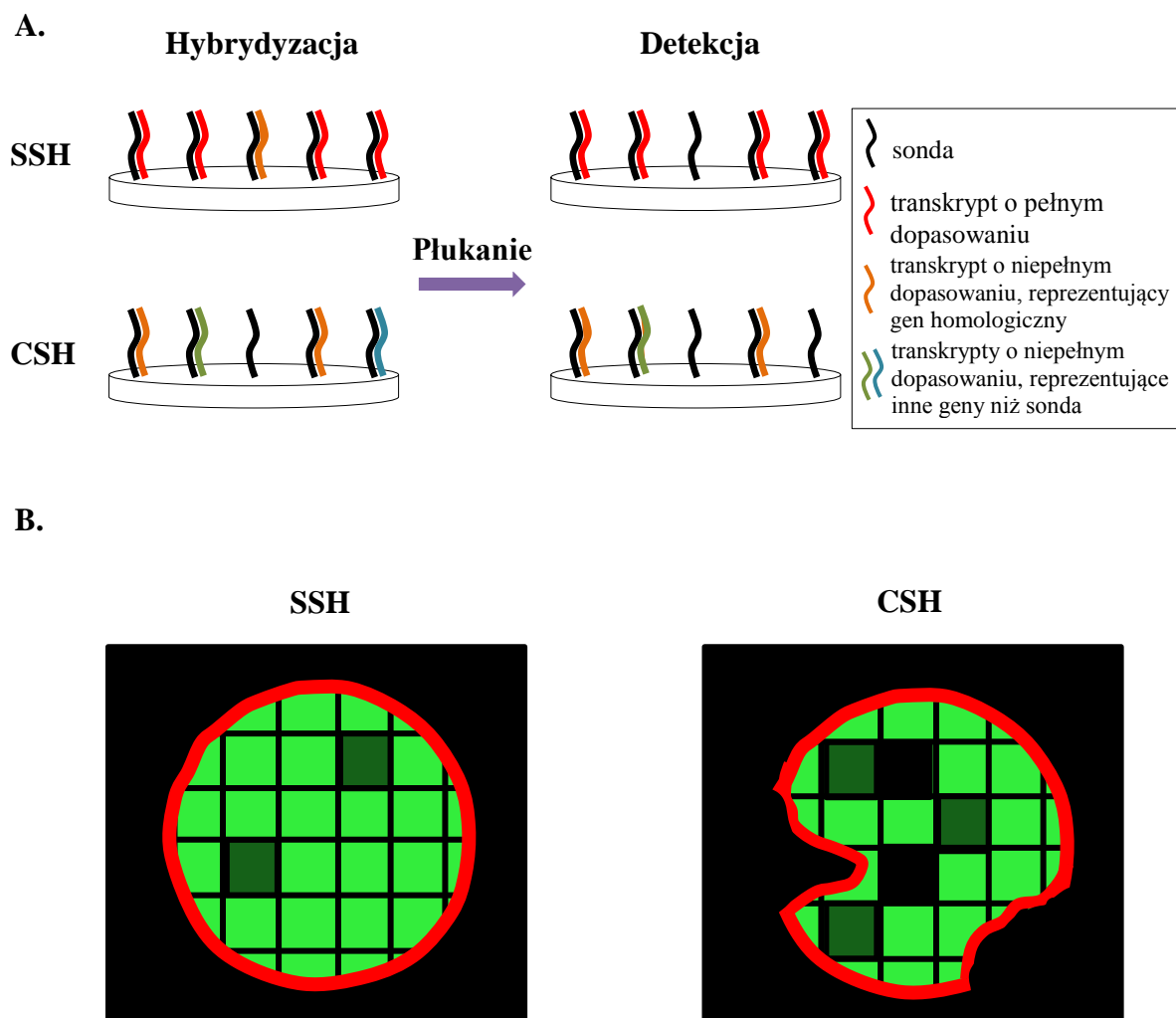
Wymienione czynniki wpływają na uzyskiwany obraz mikromacierzy. W przypadku gdy zaburzenia występują na obszarze całej płytki, mogą prowadzić do następujących niekorzystnych efektów, sygnalizowanych przez niektórych autorów:

- dodatkowy czynnik zaburzający/utrudniający analizę danych [53,56]
- zmniejszenie liczby statystycznie istotnych genów różnicujących [53,55,57,58]
- słaba powtarzalność wyników [52]



Rycina 3. Dopasowanie sond i transkryptów podczas hybrydyzacji:

A. Dopasowania na poziomie pojedynczej plamki: I. Pełne dopasowanie; II. Niepełne dopasowanie transkryptu homologicznego do genu docelowego dla sondy; III. Niepełne dopasowanie różnych transkryptów, dla genu homologicznego dla sondy oraz innych genów; IV. Brak dopasowania; B. Dopasowanie na poziomie całej mikromacierzy CSH i SSH. Widoczne zaburzenie specyficzności w przypadku CSH (Wg [46], zmodyfikowano) - w przypadku SSH większość plamek będzie prezentować pełne dopasowanie lub jego brak, niektóre cechują się hybrydyzacją krzyżową; Dla CSH zwiększa się liczba plamek o niepełnym dopasowaniu, w tym takich, które wiążą się z kilkoma transkryptami.



Rycina 4. Hipotetyczny mechanizm powstawania zaburzeń w obrazie mikromacierzy CSH w porównaniu z SSH. A. Odplukiwanie transkryptów o słabszym dopasowaniu. B. Obraz plamek po detekcji sygnału: zaciemnione piksele- słabszy sygnał, czarne piksele- brak sygnału, czerwona linia- kontur plamki (wg [52], zmodyfikowano)

Zaburzenia specyficzności mogą w istotny sposób wpływać na uzyskane profile ekspresji. Widocznym efektem są różnice w wynikach otrzymywanych dla różnych sond reprezentujących ten sam transkrypt, bądź też uzyskiwanych w ramach porównań różnych mikromacierzy [44]. Specyficzność jest kwestią szeroko dyskutowaną nawet w odniesieniu do mikromacierzy homologicznych, a efekty zaburzeń w przypadku CSH mogą być znacznie trudniejsze do przewidzenia. W przypadku SSH, niepełne dopasowanie sond występuje zdecydowanie rzadziej i może być efektem np. ekspresji genu należącego do tej samej rodziny, co sekwencja docelowa. Z kolei w CSH taki rodzaj oddziaływań zwykle związany jest z hybrydyzacją do sekwencji o mniejszym dopasowaniu, a co za tym idzie, wzrostem prawdopodobieństwa hybrydyzacji z kilkoma różnymi transkryptami [46]. Przykładowo, w pracy dotyczącej hybrydyzacji transkryptów ziemniaka do mikromacierzy

pomidorowej stwierdzono, że wśród genów wykazujących częściowe dopasowanie, ok. 16% charakteryzowało się homologią do więcej niż jednej sekwencji sond [54]. Brak dopasowania w przypadku SSH oznacza brak transkryptu docelowego, natomiast w CSH może dotyczyć braku tego transkryptu lub też braku genów o podobnej sekwencji. Ze względu na fakt, iż w przypadku hybrydyzacji międzygatunkowej znacznie większa liczba sond będzie charakteryzować się częściowym dopasowaniem lub jego brakiem, dane na wszystkich etapach ich analizy będą znacznie bardziej zaburzone, niż w przypadku SSH, co zwiększa prawdopodobieństwo uzyskania fałszywych wyników [46].

W tym kontekście na uwagę zasługują również wyniki badań CSH, które przeprowadzone zostały w naszej Pracowni z wykorzystaniem mikromacierzy do analizy ekspresji genów chloroplastowych ogórka. W przypadku hybrydyzacji międzygatunkowej RNA plastydowego rzodkiewnika, szpinaku i pomidora stwierdzono istotne zwiększenie udziału sond o niskim stosunku sygnału do szumu (SNR, ang. *signal-to-noise ratio*), który stanowi jeden z parametrów branych pod uwagę podczas oceny jakościowej sygnału fluorescencyjnego sond. Znajomość sekwencji genów chloroplastowych pozwoliła zbadać homologię pomiędzy sondami i transkryptami dla poszczególnych gatunków. Po odfiltrowaniu sond o zbyt małej komplementarności do analizowanych transkryptów w każdym przypadku stwierdzono zwiększenie udziału sond o prawidłowym SNR [91].

I.2.3. Strategie eksperymentalne stosowane w CSH

Podstawowym warunkiem uzyskania wiarygodnych wyników hybrydyzacji heterologicznej jest właściwa metodyka eksperymentu i dobrane skutecznych sposobów analizy danych. Należy wziąć pod uwagę istotne ograniczenia CSH, w szczególności różnice w stopniu dopasowania pomiędzy badanymi transkryptami a sondami znajdującymi się na mikromacierzy. Niestety, jak dotąd jedynie w części opublikowanych prac podczas analizy danych wzięto pod uwagę specyfikę eksperymentów CSH. Do tej pory nie ustalono również jednolitej metodyki postępowania w przypadku planowania eksperymentów CSH.

1.2.3.1. Dobór gatunku reprezentowanego na mikromacierzy

Pierwszym, podstawowym problemem, jaki należy rozważyć przed przystąpieniem do eksperymentu CSH jest wybór gatunku, dla którego mikromacierz została skonstruowana [54]. W przypadku, gdy znana jest sekwencja choćby niewielkiej liczby genów badanego organizmu, istotne jest określenie stopnia homologii pomiędzy nimi, a ich odpowiednikami występującymi u gatunku reprezentowanego na mikromacierzy. Stwierdzono, iż nawet niewielka różnica w dopasowaniu sekwencji (ok.1%) może w widoczny sposób wpływać na wyniki hybrydyzacji. Gilad i wsp. porównywali wpływ różnic pomiędzy sekwencjami DNA na jakość hybrydyzacji u 4 gatunków o różnej odległości filogenetycznej (człowiek, szympan, orangutan, rezus). W badaniach użyto mikromacierz cDNA zawierającą cztery wersje każdej sondy – każda z wersji była specyficzna dla jednego z badanych gatunków naczelnych. RNA poszczególnych małych hybrydyzowano w układzie dwukolorowym z RNA człowieka, a następnie porównywano wyniki (w postaci stosunku intensywności Cy3/Cy5) uzyskane dla obu gatunków. Oceniano przy tym wpływ średniej różnicy pomiędzy sekwencjami DNA obu porównywanych gatunków na intensywność sygnału uzyskaną dla danej sondy w układzie homo- i heterologicznym. W przypadku różnicy 1% (człowiek-szympan) efekt w postaci zmniejszenia intensywności podczas hybrydyzacji heterologicznej był widoczny dla 63% genów, a przy 5% (człowiek-rezus) aż dla 89% genów. [55]. Podobne obserwacje odnotowano nawet w przypadku dwóch gatunków należących do tego samego rodzaju: hybrydyzacja genomowego DNA *D. simulans* do mikromacierzy dla *D. melanogaster* skutkowałą zmniejszeniem sygnału hybrydyzacyjnego o ok. 4% w porównaniu z hybrydyzacją homologiczną *D. melanogaster*. Co ciekawe, różnica na poziomie sekwencji nukleotydowej pomiędzy tymi gatunkami szacowana jest również na około 4% [76]. W innej pracy opisującej hybrydyzację genomowego DNA trzech gatunków *Drosophila* do mikromacierzy cDNA dla *D. melanogaster* również stwierdzono znaczne zmniejszenie sygnału dla większości (średnio 84%) sond o stopniu identyczności z sekwencjami docelowymi poniżej 92%, w przypadku dopasowania sekwencji rzędu 92-97% spadek intensywności odnotowano w ok. 50% sond, a przy dopasowaniu >97%- jedynie w 13% sond [80]. Biorąc pod uwagę powyższe obserwacje, optymalnym rozwiązaniem w przypadku znajomości większej liczby sekwencji organizmu badanego, wydaje się wstępne zbadanie homologii sond znajdujących się na mikromacierzy i ich docelowych transkryptów, w celu oceny możliwości zastosowania danego układu CSH. Takie podejście zastosowano w badaniach, w których RNA psa hybrydyzowano do mikromacierzy cDNA

człowieka. Zakres podobieństwa sekwencji sond do cDNA wynosił tu 78%-100% (średnio 88%), co uznano za wystarczające [57].

W przypadku braku odpowiednich danych genomowych, przydatna może okazać się ocena pokrewieństwa filogenetycznego wybranych gatunków [46]. W kilku pracach zwrócono uwagę na wpływ pokrewieństwa filogenetycznego na jakość danych uzyskiwanych w eksperymentach CSH, sugerując wybór mikromacierzy gatunku o jak największym stopniu pokrewieństwa [40,53]. We wspomnianym już wcześniej badaniu opisującym hybrydyzację genomowego DNA kilku gatunków *Drosophila* stwierdzono, że zmniejszenie sygnału fluorescencyjnego jest proporcjonalne zarówno do odległości filogenetycznej pomiędzy gatunkami badanymi a reprezentowanym na mikromacierzy, jak i do stopnia zróżnicowania sekwencji [80]. W roku 2004 Renn i wsp. opisali pierwsze zastosowanie hybrydyzacji heterologicznej u kilku gatunków kręgowców. Stworzono mikromacierz cDNA zawierającą 4570 sond dla endemicznej ryby *Astatotilapia burtoni*. Za pomocą tego narzędzia badano ekspresję genów w mózgu *A. burtoni* oraz w mózgach 7 innych gatunków ryb o różnym stopniu pokrewieństwa (Tabela 4).

Tabela 4. Zbieżność wyników hybrydyzacji dla wybranych gatunków ryb z wynikami dla *A. burtoni* w kontekście pokrewieństwa filogenetycznego. Gatunki cechujące się krótszym czasem dywergencji w stosunku do referencyjnego charakteryzowały się większym podobieństwem wyników badań ekspresji genów [40].

Gatunek	Czas dywergencji (lata)	Zbieżność sond cechujących się wzrostem ekspresji w eksperymentach CSH, w odniesieniu do SSH
<i>Enantiopus melanogenys</i>	<10 mln	68%
<i>Neolamprologus brichardi</i>	<10 mln	68%
<i>Oreochromis niloticus</i>	<10 mln	62%
<i>Poecilia reticulata</i>	<65 mln	42%
<i>Xiphophorus sp</i>	<65 mln	42%
<i>Danio rerio</i>	<200 mln	29%
<i>Salmo salar</i>	<200 mln	27%

Uzyskane wyniki wskazują, że hybrydyzacja międzygatunkowa może być przydatnym narzędziem podczas analizy ekspresji genów, ale jakość i wiarygodność wyników maleje wraz ze zmniejszeniem pokrewieństwa filogenetycznego. W związku z tym sugeruje się, że czas dywergencji pomiędzy organizmami nie powinien przekraczać 10 mln lat [40]. Niemniej jednak, bliskość filogenetyczna nie zawsze wiąże się z

podobieństwem profilu ekspresji genów, na co wskazują wspomniane wyżej badania różnych gatunków *Solanaceae* z wykorzystaniem mikromacierzy dla pomidora. Dane filogenetyczne wskazują na bliższe pokrewieństwo pomidora z bakłażanem (ten sam rodzaj *Solanum*), niż z papryką. Niemniej jednak, profil genów różnicujących zarówno w niedojrzałych, jak i dojrzałych owocach cechuje się większym podobieństwem w przypadku pary pomidor-papryka, niż pomidor-bakłażan, co związane jest z większą zbieżnością fenotypową badanych owoców w przypadku pierwszej z par. Obserwacja ta wskazuje, że w przypadku CSH, nie zawsze podobieństwo filogenetyczne gatunków badanego i reprezentowanego na mikromacierzy będzie miało odzwierciedlenie w podobnej ekspresji genów [53].

1.2.3.2. Wybór rodzaju mikromacierzy

Jak już wspomniano, rodzaj stosowanych sond ma olbrzymi wpływ na specyficzność hybrydyzacji. Istotną kwestią w eksperymentach CSH jest zatem wybór najodpowiedniejszych spośród dostępnych sond. W wielu badaniach wykorzystywano mikromacierze cDNA z uwagi na fakt, iż dłuższe sondy powinny być mniej wrażliwe na brak pełnej komplementarności [40,56,77,78]. Mikromacierze cDNA są narzędziem z wyboru zwłaszcza w przypadku badań porównawczych, polegających na kompetytywnej hybrydyzacji dwóch prób różnych gatunków do tej samej mikromacierzy [61]. Oczywiście obecność długich sond zwiększy też znacznie prawdopodobieństwo wystąpienia hybrydyzacji krzyżowej. Na uwagę zasługuje też fakt, że w przypadku sond cDNA mamy do czynienia z utrudnioną identyfikacją transkryptu hybrydującego z sondą jak również oceną stopnia homologii sonda - transkrypt, ponieważ zazwyczaj znany jest jedynie fragment sekwencji EST sondy.

Alternatywnym rozwiązaniem jest zastosowanie mikromacierzy oligonukleotydowych. Z uwagi na swoją naturę, sondy oligo zapewniają większą specyficzność hybrydyzacji, a znajomość ich sekwencji daje większe możliwości precyzyjnego zbadania stopnia homologii sonda-transkrypt, jeśli dysponuje się odpowiednią ilością danych sekwencyjnych dla gatunku badanego. Należy jednak wziąć pod uwagę, że sondy oligo do mikromacierzy ekspresyjnych są najczęściej ulokowane w regionie 3' sekwencji docelowych – bardzo często w regionie 3' niekodującym, który wykazuje najwyższą specyficzność dla danego gatunku. Może to doprowadzić do wyeliminowania z eksperymentów CSH dużo większej liczby sond, z uwagi na ich niewystarczające podobieństwo sekwencji.

W przypadku występowania kilku sond hybrydujących w różnych miejscach tego samego transkryptu można wyselekcjonować te najbardziej odpowiednie do CSH z zastosowaniem filtracji bazującej na znajomości sekwencji bądź intensywności fluorescencji [31]. W praktyce takie podejście stosowano kilkakrotnie w przypadku mikromacierzy S-oligo, gdzie każdy zestaw sond ukierunkowany na dany transkrypt obejmuje średnio 11 sond [31,38]. Z drugiej strony, w takim przypadku następuje utrata korzyści płynących z posiadania zestawu sond pokrywających dany gen (założenia będącego przecież podstawą koncepcji mikromacierzy S-oligo) [74].

Dotąd przeprowadzono niewiele badań porównujących hybrydyzację różnych rodzajów mikromacierzy uzyskanych dla tego samego gatunku z transkryptami należącymi do innego gatunku. W pracy opisującej zastosowanie różnej długości sond oligonukleotydowych do zbadania metodą CSH ekspresji genów makaka, wykorzystano ludzkie mikromacierze zbudowanych z dłuższych (60-merów) lub krótszych (25-merów) oligonukleotydów. Zasugerowano, że większą liczbę genów różnicujących i mniejszą ilość fałszywie pozytywnych wyników uzyskuje się stosując dłuższe sondy [58].

1.2.3.3. Mikromacierzowe eksperymenty porównawcze

W przypadku CSH częstą praktyką jest wykonywanie dodatkowych eksperymentów porównawczych, pozwalających na dokładniejszą weryfikację badań. Jedno z podejść polega na porównaniu wyników CSH z wynikami uzyskanymi dla próby pochodzącej z organizmu, dla którego macierz została stworzona (hybrydyzacja homologiczna) [40,53]. Przykładem mogą tu być badania, w których porównywano wyniki hybrydyzacji RNA psa i człowieka do mikromacierzy S-oligo dla człowieka. Hybrydyzacja homologiczna posłużyła w tym przypadku do weryfikacji wykrywanego sygnału uzyskanego podczas hybrydyzacji heterologicznej [62]. Niemniej jednak strategia polegająca na weryfikacji danych CSH za pomocą SSH wymaga ostrożnego podejścia, ponieważ niektóre procesy biologiczne mogą przebiegać w odmienny sposób nawet u blisko spokrewnionych gatunków [90]. Należy wziąć pod uwagę również fakt, że to właśnie różnice pomiędzy gatunkami dotyczące ekspresji genów są często celem poszukiwań [46]. Inną możliwością jest umieszczenie na jednej mikromacierzy sond homologicznych i heterologicznych. W badaniu, w którym poszukiwano homologów genów człowieka i kałamarnicy, hybrydyzowano materiał genetyczny do mikromacierzy oligonukleotydowej zawierającej oba rodzaje sond. Sondy zaprojektowano na podstawie sekwencji EST obu gatunków znajdujących się w bazach danych. Po hybrydyzacji RNA

kałamarnicy typowano geny wykazujące homologię na podstawie ludzkich sond wykazujących odpowiednią intensywność fluorescencji. Sondy kałamarnicy posłużyły tu jako kontrola intensywności sygnału. Tego typu podejście wymagało jednak produkcji niestandardowej mikromacierzy, zaprojektowanej specjalnie pod kątem konkretnego eksperymentu i nie znajduje powszechnego zastosowania w badaniach ekspresji genów z wykorzystaniem metody CSH [34]. Najbardziej bezpośrednim, a przez to poprawnym sposobem wykorzystania hybrydyzacji SSH do weryfikacji wyników CSH może być użycie mikromacierzy homologicznej i heterologicznej do hybrydyzacji z próbą pochodzącą z tego samego materiału biologicznego. Tego typu strategię zastosowano podczas eksperymentów oceniających ekspresję genów ziemniaka na mikromacierzach cDNA dla pomidora (CSH) i ziemniaka (SSH). Eksperyment ten miał na celu ustalenie, czy uzyskane dane CSH będą odzwierciedlać wyniki otrzymane podczas hybrydyzacji tych samych prób do mikromacierzy homologicznej. Obserwowano tu znaczne rozbieżności, których liczba uległa wyraźnemu zmniejszeniu po wyodrębnieniu z mikromacierzy pomidorowej zestawu sond o najlepszym dopasowaniu sekwencji do cDNA ziemniaka [54]. To niezwykle istotne spostrzeżenie podnosi wagę odpowiedniego zaplanowania przebiegu i analizy eksperymentu CSH w celu uzyskania wyników poprawnych z biologicznego punktu widzenia. Jest to jedno z nielicznych doniesień, faktycznie porównujących zbieżność wyników ekspresji genów uzyskanych w dwóch typach hybrydyzacji (SSH i CSH). W oczywisty sposób, wobec dysponowania mikromacierzą homologiczną, przeznaczoną dla danego gatunku, przeprowadzanie dodatkowych eksperymentów w układzie CSH może służyć wyłącznie badaniu aspektów samej metody, stąd niewielka liczba publikacji tego typu.

I.2.4. Warunki eksperymentu CSH

Ustalenie optymalnych warunków eksperymentu ma kolosalne znaczenie nie tylko w przypadku hybrydyzacji heterologicznej. Ze względu na częste problemy z powtarzalnością wyników [52], pożądane jest zastosowanie możliwie dużej liczby powtórzeń biologicznych oraz technicznych. Zmieniając niektóre parametry hybrydyzacji (temperatura, płukanie) można wpływać na specyficzność hybrydyzacji, jednak w niewielu pracach związanych z CSH podjęto ten temat. Przykładowo, w badaniu transkryptomów tkanek mózgu psa, wilka i kojota z użyciem mikromacierzy cDNA człowieka, zastosowano dwie temperatury hybrydyzacji. Bardziej restrykcyjne warunki (temperatura 60°C) stosowano w celu poszukiwania szlaków specyficznych dla poszczególnych

regionów w mózgu. Z kolei hybrydyzacja w temperaturze 55°C miała służyć porównaniu wyników pomiędzy poszczególnymi gatunkami. Wyniki ilościowego PCR potwierdziły, że detekcja homologicznych transkryptów możliwa była również przy niższej specyficzności [57].

I.3. ANALIZA DANYCH CSH

Niezwykle istotnym elementem eksperymentów mikromacierzowych jest analiza danych. W przypadku CSH, ze względu na opisywany wcześniej szereg dodatkowych czynników zaburzających wiązanie próby i transkryptu, właściwa analiza i interpretacja danych jest sporym wyzwaniem [74]. W wielu badaniach opracowywano dane z wykorzystaniem standardowej ścieżki analizy, typowej dla eksperymentów SSH. Być może, ze względu na specyfikę CSH, istotną innowacją mogłoby być opracowanie nowej metody normalizacji, biorącej pod uwagę wpływ niepełnego dopasowania transkryptów i sond. Takie próby opisano dotychczas w dwóch pracach, jednak obie modyfikacje metody normalizacji dotyczą hybrydyzacji genomowego DNA do mikromacierzy i mogą stanowić potencjalną metodę walidacji wyników badań ekspresji genów [78,80]. Pierwsza z metod sugeruje normalizację na podstawie zestawu sond kontrolnych, komplementarnych do sekwencji gatunku innego niż reprezentowany na mikromacierzy (w tym przypadku *A. thaliana* i *Bacillus subtilis* dla mikromacierzy *Caenorhabditis elegans*). Zakłada się, że takie sondy wykazują równie niskie dopasowanie (brak dopasowania) względem docelowego DNA genomowego i intensywność ich fluorescencji mogłaby posłużyć jako punkt odniesienia w normalizacji [78]. Drugie podejście zakłada z kolei normalizację z wykorzystaniem zestawu sond o największym dopasowaniu do sekwencji genomowego DNA i przyjęcie ich intensywności za miarę dobrego dopasowania [80]. Powyższych metod nie stosowano jednak do tej pory w badaniach ekspresji genów. W związku z tym, obecnie jedyną możliwością uwzględnienia zaburzeń towarzyszących hybrydyzacji heterologicznej jest filtracja danych. W kilku pracach potwierdzono skuteczność filtracji jako procedury znacznie poprawiającej jakość wyników [55,54,62].

Właściwa filtracja powinna zminimalizować prawdopodobieństwo wygenerowania danych fałszywie pozytywnych (np. powstałych w wyniku hybrydyzacji krzyżowej) przy równoczesnym zachowaniu jak największej liczby informacji. Biorąc pod uwagę stosowane kryteria, możemy wyróżnić dwa zasadnicze typy filtracji danych CSH: opartą na analizie dopasowania sekwencji nukleotydowej transkryptów i sond oraz opartą na jakości uzyskanych sygnałów fluorescencyjnych.

I.3.1. Filtracja na podstawie podobieństwa sekwencji nukleotydowych

W niektórych pracach opisano skuteczność filtracji opartej na analizie homologii pomiędzy sekwencjami gatunku badanego i reprezentowanego na mikromacierzy. Przykładem mogą być wspomniane już badania polegające na hybrydyzacji RNA ziemniaka do mikromacierzy S-oligo dla pomidora, w ramach których utworzono bazę danych zawierającą wykaz sekwencji ziemniaka homologicznych do sekwencji docelowych dla sond obecnych na mikromacierzy. Dzięki temu wytypowano zestawy sond zawierające konserwatywne sekwencje, które były wykorzystane podczas późniejszej analizy danych [38]. W badaniu porównującym hybrydyzację cDNA ziemniaka do mikromacierzy homologicznej oraz heterologicznej, zawierającej sondy dla pomidora, zastosowano filtrację opartą na wytypowaniu homologicznych sekwencji genów pomidora i ziemniaka, dla których sondy znajdowały się na obu mikromacierzach. Ustalono próg podobieństwa sekwencji wynoszący $e < 0,001$ (wartość e – wartość oczekiwana, opisująca liczbę dopasowań, które mogą zostać znalezione „przez przypadek” podczas przeszukiwania określonej bazy danych): sekwencje uzyskujące wartość równą lub mniejszą od progowej uznawano za przydatne podczas analizy danych. W przypadku analizy bez użycia filtracji wykazano rozbieżności pomiędzy obiema grupami danych (m.in. niezgodność wyników analizy głównych składowych (PCA, ang. *Principal Component Analysis*). Dotyczyły one profilów ekspresji oraz liczby genów różnicujących uzyskiwanych z wykorzystaniem mikromacierzy SSH i CSH. Ze względu na wspomnianą powyżej weryfikację danych CSH za pomocą hybrydyzacji SSH, jest to jedna z prac w sposób wiarygodny potwierdzających istotność filtracji danych w eksperymentach z wykorzystaniem mikromacierzy heterologicznych [54]. W innej pracy hybrydyzowano próbki cRNA z komórek linii CHO (komórki jajnika chomika chińskiego) do komercyjnej mikromacierzy S-oligo dla myszy oraz mikromacierzy dla chomika (zaprojektowana przez badaczy). Oceniano różnice w ekspresji genów w tkance poddanej działaniu maślanu sodu, w porównaniu z tkanką kontrolną. W badaniu wyselekcjonowano pary sond odpowiadające genom ortologicznym u myszy i chomika. Zaobserwowano korelację pomiędzy intensywnością sygnału fluorescencyjnego i stopniem homologii w przypadku hybrydyzacji heterologicznej. Podobnie jak w poprzednim przykładzie obserwowano rozbieżności pomiędzy wynikami CSH i SSH w zakresie powtarzalności wykrytych zmian w ekspresji genów, jednak stwierdzono, iż właściwa selekcja sond na podstawie homologii istotnie zwiększa wiarygodność uzyskanych wyników [39]. Biorąc pod uwagę kryteria filtracji oparte na podobieństwie sekwencji należy pamiętać, że w miarę uzupełniania

genomicznych baz danych, informacje dotyczące homologii poszczególnych sekwencji mogą ulegać zmianie wskutek dodawania nowych sekwencji, aktualizacji już istniejących, istnienia różnych wariantów transkrypcyjnych lub rewidowania błędów. Ponadto, obydwie gatunki musi charakteryzować znaczna liczba sekwencji dostępnych w bazach danych, co w przypadku słabo poznanych genetycznie organizmów nie jest możliwe [54].

I.3.2. Filtracja na podstawie parametrów morfologicznych plamek

I.3.2.1. Intensywność sygnału

W niektórych pracach stosowano prostą filtrację polegającą na zliczaniu sond, które uległy hybrydyzacji (charakteryzowały się obecnością sygnału), niemniej jednak tego typu strategia wydaje się być mało wiarygodna, gdyż nie odróżnia braku ekspresji genu od braku właściwej sondy reprezentującej ten gen, jak również nie pozwala na jakąkolwiek weryfikację prawdziwości danych uzyskanych dla wyselekcjonowanych sond [63].

Metoda filtracji często stosowana podczas analizy mikromacierzy S-oligo polega na pominięciu podczas analizy sond charakteryzujących się sygnałem mniejszym niż ustalona wartość progowa [33,62,37,81]. Użyteczność tego typu filtracji potwierdzono w opisywanym już eksperymencie z zastosowaniem ludzkiej mikromacierzy s-oligo i RNA psa [62]. W przypadku tradycyjnej analizy jedynie 14% zestawów sond pozwalało na uzyskanie wyników. Po zastosowaniu filtracji na podstawie intensywności sygnału, weryfikowanej za pomocą hybrydyzacji homologicznej (człowiek-człowiek) okazało się, że po wzięciu pod uwagę pojedynczych sond, a nie całych zestawów, liczba transkryptów możliwych do identyfikacji wzrosła o 60%. Po weryfikacji za pomocą mikromacierzy homologicznej (pies-pies), odnotowano zwiększenie o 46% liczby prawidłowo wykrywanych transkryptów oraz zmniejszenie liczby wyników fałszywie negatywnych o 53% [53]. Degletagne i wsp. w eksperymencie polegającym na użyciu mikromacierzy dla kurczaka do badania ekspresji genów pingwina wykazał, że 96% zestawów sond posiadało co najmniej 1 sondę o stosunkowo intensywnym sygnale [31]. Na tej podstawie filtrowano najbardziej odpowiednie sondy reprezentujące dany gen. Ekspresję genów różnicujących potwierdzono za pomocą qPCR, a wyniki okazały się zbieżne w 90% przypadków. Niemniej jednak, istnieją badania negujące zastosowanie intensywności sygnału jako jedyne kryterium filtracji, mającego odzwierciedlać podobieństwo sekwencji, wskazujące na możliwość generowania fałszywych wyników. We wspomnianej wyżej pracy dotyczącej hybrydyzacji genomowego DNA różnych gatunków nicieni do

mikromacierzy dla *C. elegans* stwierdzono, że znaczny odsetek umieszczonych na mikromacierzy sond kontrolnych, specyficznych dla innych gatunków (*A. thaliana* i *B. subtilis*) i służących do normalizacji, wykazywał znaczną intensywność fluorescencji. Autorzy sugerują, że w przypadku tak krótkich sond, wpływ nawet jednonukleotydowych różnic w sekwencji może być znacznie bardziej nieprzewidywalny niż w przypadku sond cDNA. Różnice w intensywności sygnałów mogą w przypadku mikromacierzy S-oligo w znacznym stopniu zależeć nawet od samego składu sondy (zawartość par GC vs AT) [78].

1.3.2.2 Intensywność sygnału pochodzącego z hybrydyzacji z genomowym DNA

Interesującym podejściem jest zastosowanie hybrydyzacji genomowego DNA do mikromacierzy i wykorzystanie uzyskanych informacji do ustalenia progu filtracji. Dzięki temu możliwa jest identyfikacja i filtracja sond, które wykazują niewystarczające podobieństwo sekwencji do transkryptów, przed rozpoczęciem właściwych eksperymentów [37]. Metoda ta wykorzystywana jest w przypadku braku informacji na temat sekwencji poszczególnych genów i jest szczególnie użyteczna podczas analizy porównawczej transkryptomów kilku gatunków [61]. Procedurę taką zastosowano m.in. w badaniach transkryptomu kapusty (*Brassica oleracea*) (poddanej działaniu stresu fosforowego), w których użyto chipów DNA dla rzodkiewnika [37]. W tym eksperymencie, przeprowadzono hybrydyzację genomowego DNA kapusty, wyznakowanego biotyną, do mikromacierzy heterologicznej, a następnie odrzucano sondy o słabszej intensywności sygnału hybrydyzacyjnego. Uzyskane informacje wykorzystano do filtracji sond podczas badania ekspresji genów. Pomimo odrzucenia łącznie ponad 68% indywidualnych sond, uzyskano informację biologiczną dla aż 96% zestawów sond obecnych na chipie. Ponadto w tym przypadku zastosowanie filtracji ponad 10-krotnie zwiększyło liczbę genów różnicujących, w porównaniu z analizą zestawu bez filtracji. Hybrydyzację genomowego DNA do mikromacierzy można również wykorzystać do walidacji danych uzyskanych w eksperymentach polegających na bezpośrednim porównaniu transkryptomów dwóch gatunków na jednej mikromacierzy, stosowanych m.in. w badaniach nad ewolucją gatunków. Podejście to zastosowano m.in. w badaniach ryb pielęgnicowatych [61], z wykorzystaniem mikromacierzy cDNA stworzonej przez Renn i wsp [40]. Celem badania było porównanie profilów ekspresji genów związanych z płcią u 2 gatunków spokrewnionych z rybą *A. burtoni*, dla której stworzono mikromacierz. Przeprowadzono serię bezpośrednich porównań RNA samic i samców obu gatunków, zestawiając uzyskane dane z wynikami hybrydyzacji genomowego DNA. W przypadku,

gdy stosunek sygnałów hybrydacyjnych DNA genomowego obu gatunków był znacząco różny od 1, można było oczekiwać wpływu różnic w sekwencji na proces hybrydacji, co mogło zaburzać wyniki analizy ekspresji genów. W tego rodzaju eksperymentach, zastosowanie hybrydacji z genomowym DNA pozwala wyeliminować fałszywie pozytywne wyniki [61].

1.3.2.3 Parametry jakości plamek

Interesującym podejściem jest wykorzystanie do filtracji parametrów opisujących jakość plamek, jednak nie będących wyłącznie miarą intensywności sygnału hybrydacyjnego informującego o zmianach ekspresji genów. Należy jednak przy tym zaznaczyć, że dostępność informacji o poszczególnych parametrach zależy od rodzaju oprogramowania użytego do analizy obrazów mikromacierzy. Różnego rodzaju parametry dotyczące charakterystyki jakościowej plamek można automatycznie określić między innymi za pomocą komercyjnego programu GenePix Pro bądź darmowego programu Maia [71]. Bar-Or i wsp. wykorzystując parametry jakości plamek uzyskane za pomocą programu Maia wykazali, że niektóre z nich charakteryzuje korelacja ze stopniem homologii sekwencji pomiędzy sondą a transkryptem, wyrażonym jako wartość Bit Score obliczona za pomocą programu BLAST. Wartość Bit Score jest znormalizowaną miarą podobieństwa pomiędzy dwiema sekwencjami, zależną od długości sekwencji i rozmiaru biblioteki. Stwierdzono również, że korelacja między parametrami jakości a Bit Score jest mniejsza w przypadku dalszego pokrewieństwa pomiędzy organizmami badanym i reprezentowanym na mikromacierzy. Wspomniana korelacja była niepełna, jednak sugeruje się, że jest to skutek niekompletnych informacji dotyczących sekwencji w bazach danych. Parametry wytypowane jako odzwierciedlające stopień dopasowania wymieniono w Tabeli 5. Dzięki odrzuceniu sond, które nie osiągnęły wartości progowych ustalonych dla wszystkich 5 parametrów, uzyskano znaczną poprawę wyników analizy skupień w 2 z 3 eksperymentów, co zdaniem autorów świadczy o możliwości wykorzystania charakterystyki morfologicznej plamek do filtracji danych w eksperymentach CSH [52]. Wnioski z opisanej pracy (w ograniczonej formie) wykorzystali inni badacze, np. stosując średnicę plamek jako kryterium filtracji [84].

Tabela 5. Dostępne w programie Maia parametry jakości plamek, które mogą odzwierciedlać stopień dopasowania sond i transkryptów, [52,71].

Parametr	Opis	Dodatkowe informacje
Det	Współczynnik determinacji regresji liniowej pomiędzy intensywnością kanałów Cy3 i Cy5	Miara równomierności rozkładu sygnałów Cy3 i Cy5. Niższe wartości świadczą o dużym wpływie zmienności statystycznej, zwykle charakteryzującej plamki o małej intensywności, bądź też o obecności stosunkowo jasnych, lecz nie korelujących z ekspresją genu zanieczyszczeń.
Dia	Średnica plamki	Zależna od średniej wartości średnicy plamek na całej mikromacierzy.
GSy	Symetria geometryczna	Miara odchylenia plamki od kształtu koła.
CVR	Współczynnik wariancji pomiędzy wartościami współczynników RR (współczynnik regresji) i RS (współczynnik segmentacji) opisujących zależności pomiędzy kanałami Cy3 i Cy5.	Wysoka wartość oznacza obecność dodatkowych pikseli w plamce lub dodatkowych pikseli na zewnątrz plamki.
CV	Współczynnik wariancji pomiędzy wartościami współczynników ekspresji genów dla replik sond	Opisuje powtarzalność hybrydyzacji dla tych samych sond (możliwość stosowania jedynie dla replik sond)

I.4. ODPOWIEDŹ ROŚLIN NA STRES SOLNY I KADMOWY - STAN BADAŃ TRANSKRYPTOMICZNYCH

Rośliny jako organizmy nie posiadające możliwości przemieszczania się, są narażone na stałe oddziaływanie czynników stresowych ze strony środowiska. Stres abiotyczny w istotny sposób wpływa na plony roślin uprawnych, zmniejszając je nawet dwukrotnie. Przykładami stresowych czynników abiotycznych działających na rośliny są: susza, zalanie, nadmierne zasolenie, chłód lub wysoka temperatura, nadmiar lub niedobór światła oraz metale ciężkie [94]. W związku z ich działaniem, organizmy rozwinęły szereg mechanizmów obronnych, w ramach których dochodzi do zmian we wzorze ekspresji wielu genów.

Stres solny jest jednym z podstawowych czynników wpływających na produkcję rolną. Duży stopień zasolenia gleby zmniejsza światowe zbiory roślin uprawnych o około 20% [95]. Istniejąca wiedza na temat reakcji roślin na stres solny jest dość obszerna. Wiadomo, że wpływa on na gospodarkę wodną oraz homeostazę jonową rośliny. Akumulacja soli może modyfikować skład błon komórkowych i powodować stres hiperosmotyczny, zakłócając wzrost i rozwój rośliny. Znane są dwa rodzaje odpowiedzi rośliny, odmienne pod względem czasu występowania: krótkotrwała odpowiedź ochronna i długotrwała adaptacyjna. Geny odpowiedzi na stres możemy podzielić właśnie na takie dwie klasy [94]. Zaangażowanych w reakcję rośliny jest wiele genów tworzących zawiłą sieć zależności obejmującą szereg szlaków biochemicznych związanych z transdukcją sygnału, metabolizmem związków węgla, wytwarzaniem energii, ochroną przed stresem oksydacyjnym, gospodarką jonami sodu, modyfikacją struktury ścian i błon komórkowych, syntezą osmolitów i transporterów jonów, usuwaniem wolnych rodników i innych. W odpowiedzi biorą udział takie substancje jak: enzymy detoksyfikacji (np. peroksydaza glutationowa), osmolity (np. glicyna-betaina, LEA), flawoproteiny (np. HAL3). Odpowiedź na stres regulowana jest przy pomocy wielu czynników transkrypcyjnych, takich jak AP2/ERBP, DREB, SOS i innych [94]. W sygnalizację komórkową związaną ze stresem zaangażowane są również hormony, takie jak kwas jasmonowy, kwas abscysynowy czy etylen [94].

Istnieją liczne prace oceniające zmiany transkryptomiczne u roślin poddanych działaniu stresu solnego, jednak tylko nieliczne z nich dotyczyły specyficznej tkanki, jaką stanowią korzenie. Ponieważ są one odpowiedzialne za zaopatrzenie w wodę, stanowią pierwsze miejsce odbioru bodźców dla stresu osmotycznego [95]. W dotychczasowych badaniach stwierdzono odmienne profile ekspresji genów dla korzeni i części naziemnej

[96]. Z istotnych prac genomicznych z zastosowaniem mikromacierzy, badających działanie stresu solnego na korzenie roślin warto zwrócić uwagę na eksperymenty dla bawełny [100] oraz rzodkiewnika [95, 100]. Szczególnie w przypadku rzodkiewnika, eksperymenty pozwoliły nakreślić szczegółowy profil ekspresji genów w stresie solnym, obejmujący m.in. zmiany w zakresie aktywności genów osmoprotekcyjnych, sygnalizacji komórkowej i regulacji hormonalnej, budowie ściany komórkowej oraz transportu (por. V.4.1.1.) [95,100]. W przypadku rodziny *Solanaceae*, na uwagę zasługuje praca opisująca badanie stresu solnego w korzeniach pomidorów, która umożliwiła wytypowanie genów związanych z regulacją transkrypcji, metabolizmem ścian komórkowych oraz mechanizmami regulacji równowagi osmotycznej (geny SOS) [142].

Drugim czynnikiem stresowym, którego wpływ oceniano w niniejszej pracy jest chlorek kadmu. Kadm jest metalem ciężkim o działaniu toksycznym i jednym z pierwiastków zanieczyszczających środowisko, na działanie którego narażone są nie tylko rośliny, ale również ludzie i zwierzęta tworzące łańcuchy pokarmowe [97]. Związki kadmu absorbowane są przez korzenie i transportowane do organów wegetatywnych i generatywnych, wpływając negatywnie na szereg procesów życiowych rośliny, takich jak fotosynteza, synteza białek, aktywność enzymów, transport jonów itp. [97]. Rośliny wykształciły szereg mechanizmów obronnych związanych z odpowiedzią na działanie metali ciężkich, prowadzących do detoksyfikacji. Mechanizm odpowiedzi na działanie kadmu nie jest jeszcze tak szczegółowo poznany, jak w przypadku stresu solnego. Wiadomo, że związki kadmu w istotny sposób wpływają na ekspresję genów rośliny, która uruchamia procesy detoksyfikacji, polegające na wychwycie szkodliwych jonów, ich transporcie i sekwestracji w wakuoli. Z badań dotyczących działania kadmu na transkryptom korzeni roślin warto wskazać prace dotyczące rzodkiewnika [100], oraz ryżu [97]. Badania z wykorzystaniem mikromacierzy wskazują na tkankową specyficzność procesów ekspresji genów podczas odpowiedzi na stres wywołany obecnością jonów metali ciężkich. W korzeniach poddanych ich działaniu stwierdzono indukcję genów związanych z asymilacją/redukcją siarki oraz metabolizmem glutationu, prowadzącym do syntezy fitochelatyn [94]. Z kolei badanie korzeni rzodkiewnika poddanych działaniu kadmu pozwoliło zidentyfikować kategorie genów biorących udział w odpowiedzi: 1) geny wspólnej, konserwatywnej odpowiedzi na stres; 2) geny związane z odpowiedzią charakterystyczną dla metalofitów; 3) geny specyficznej odpowiedzi na kadm [98]. Odpowiedź niespecyficzna może obejmować akumulację osmolitów, antyoksydantów, aminokwasów oraz zmiany w równowadze hormonalnej. W przypadku rodziny

Solanaceae jak dotąd wykonano niewiele badań transkryptomicznych dotyczących wpływu kadmu.

I.5. PRZYSZŁOŚĆ TECHNIKI MIKROMACIERZOWEJ

Rozwój techniki RNAseq, umożliwiającej wysokowydajne sekwencjonowanie transkryptomu, stworzył nowe możliwości badania ekspresji genów. Technika ta posiada kilka atrybutów dających jej przewagę nad mikromacierzami, m.in. daje możliwość bezpośredniej oceny ilościowej transkryptów reprezentowanych przez odczyt (w przeciwieństwie do mikromacierzy, w której oceniamy ilość transkryptów na podstawie intensywności sygnału fluorescencyjnego), pozwala na ocenę zmian w znacznie większym zakresie dynamicznym dotyczącym różnic w ekspresji genów oraz nie jest ograniczona wyłącznie do sekwencji reprezentowanych przez sondy, dając również możliwość wykrywania i oceny ilościowej nowych transkryptów. Technika RNAseq umożliwia także rozróżnianie izoform transkryptów i wykrywanie nowych wariantów splicingowych.[169] Niemniej jednak, w chwili obecnej cena eksperymentu mikromacierzowego jest nadal kilkakrotnie niższa niż sekwencjonowanie transkryptomu, tym bardziej że podobnie jak w przypadku mikromacierzy, również technika RNAseq wymaga zastosowania powtórzeń, w celu uzyskania statystycznie istotnych wyników. Warto również podkreślić konieczność magazynowania i obróbki olbrzymiej ilości danych, co wymaga posiadania zaawansowanej bazy informatycznej. Ponadto, w ciągu kilkunastu lat stosowania szczegółowo opisano czynniki zakłócające przebieg eksperymentów mikromacierzowych i dokładnie opracowano metody analizy danych uwzględniające zaburzenia związane ze specyfiką takiego eksperymentu [35]. W przypadku sekwencjonowania, analiza danych jest w dalszym ciągu w trakcie optymalizacji, a źródła zakłóceń nie są tak dobrze poznane, jak w przypadku mikromacierzy [36]. Szczególnie analiza wyników sekwencjonowania *de novo*, bez genomowej sekwencji referencyjnej, obejmująca rekonstrukcję transkryptów ze zbioru krótkich odczytanych sekwencji (zazwyczaj o długości ok. 36-150 nt), jest trudnym wyzwaniem. Problemem jest tu między innymi zmienny stopień pokrycia związany z różnym poziomem ekspresji poszczególnych transkryptów oraz ze specyfiką metod przygotowania bibliotek sekwencyjnych (odmienną dla różnych platform technologicznych) [72]. Można zatem przypuszczać, że pomimo znacznie większych możliwości, jakie niesie ze sobą sekwencjonowanie nowej generacji, mikromacierze będą nadal stanowić atrakcyjne narzędzie globalnej analizy ekspresji genów.

II. CEL PRACY

Podstawowym celem niniejszej pracy jest kompleksowa ocena możliwości wykorzystania technik CSH w badaniach transkryptomów roślinnych, oraz wyselekcjonowanie czynników wykazujących najistotniejszy wpływ na wynik tego typu analiz.

Zgodnie z przyjętymi założeniami osiągnięcie tak zarysowanego celu wymagało realizacji następujących zadań:

- ustalenia w jaki sposób homologia pomiędzy genami gatunku badanego i reprezentowanego na mikromacierzy wpływa na wyniki CSH;
- określenia wpływu długości i rodzaju stosowanych sond na rezultaty badań
- opracowania metodyki analizy danych, ze szczególnym uwzględnieniem wyboru kryteriów wstępnej filtracji danych.

Równocześnie przyjęto, iż punktem odniesienia, pozwalającym na ocenę znaczenia poszczególnych czynników w hybrydyzacji CSH, powinny być wyniki uzyskane przy pomocy hybrydyzacji gatunkowo-specyficznej.

III. MATERIAŁY I METODY

III.1. MATERIAŁY

III.1.1. Materiał roślinny

W badaniach wykorzystywano RNA izolowany z korzeni tytoniu szlachetnego (*Nicotiana tabacum*). Rośliny uprawiano metodą hydroponiczną w Laboratorium Biochemii Roślin Instytutu Biochemii i Biofizyki Polskiej Akademii Nauk. Rośliny poddawano działaniu czynników stresowych: CdCl₂ w stężeniu 100 μM lub NaCl w stężeniu 150 mM w roztworze do uprawy hydroponicznej. Po upływie 6 godzin rośliny zbierano i natychmiast zamrażano w ciekłym azocie. Część roślin nie była poddana działaniu stresu i stanowiła grupę kontrolną.

III.1.2. Zestawy do preparatyki kwasów nukleinowych

Tabela 6. Stosowane zestawy do preparatyki kwasów nukleinowych

Nazwa	Producent	Zastosowanie
RNeasy Plant Mini Kit	Qiagen	Izolacja RNA, Oczyszczanie znakowanego cRNA
TURBO DNA-free™ Kit	Ambion	Hydroliza DNA
SuperScript Indirect cDNA Labeling System	Invitrogen	Odwrotna transkrypcja, znakowanie cDNA
MinElute Purification Kit	Qiagen	Oczyszczenie znakowanego cDNA
Quick Amp Labeling Kit, two-color	Agilent	Znakowanie RNA
Gene Expression Hybridization Kit	Agilent	Hybrydyzacja
RNA 6000 Nano kit	Agilent	Analiza ilościowa, jakościowa RNA
Agilent RNA Spike-in Kit	Agilent	Kontrola jakości mikromacierzy
Gene Expression Wash Buffer Kit	Agilent	Płukanie mikromacierzy

III.1.3. Odczynniki chemiczne

4M octan sodu (Invitrogen)

BSA (Sigma)

Chlorowodorek hydroksylaminy (Sigma)

DMSO (Invitrogen)

DTT (Invitrogen)

EDTA (Invitrogen)

Etanol (POCH Gliwice)

HCl (Sigma)
Izopropanol (POCH Gliwice)
NaOH (POCH Gliwice)
woda wolna od nukleaz (Ambion)
poli(A) (Sigma)
SDS (Sigma)
 β -merkaptoetanol (Sigma)
CyTM3 Mono-Reactive Dye (Amersham)
CyTM5 Mono-Reactive Dye (Amersham)

III.1.4. Mikromacierze

TOM1 (Boyce Thompson Institute for Plant Research)
TOM2 (Boyce Thompson Institute for Plant Research)
Tobacco Gene Expression Microarray, 4x44K- Agilent (TOB)
TIGR Potato cDNA Microarray (10Kv4)- The Institute for Genomic Research (TIGR)
(POT)

III.1.5. Bufory

Bufor prehybrydacyjny

- 5x SSC
- 1% BSA
- 0,1% SDS

Bufor sprzęgający (ang. *coupling buffer*)

- 0,1M NaHCO₃, pH \geq 9,0., nie starszy niż 30 dni

III.1.6. Aparatura

Bioanalyzer 2100 (Agilent)
Crosslinker UVilink CL508 (Uvitec)
Hybrydizator HybArray 12 (Perkin Elmer)
Hybrydizator HS 4800 Pro (Tecan)
Łaźnia wodna 503843/503844 (Julabo)
Koncentrator SpeedVac SVC 100H (Savant)
Mikrowirówka Vortex FVL-2400N (Biosan)
Skaner mikromacierzy Scan Array Express (Perkin Elmer)

Skaner mikromacierzy GenePix 4200 AL (Axon)

Spektrofotometr NanoDrop 1000 (Thermo Scientific)

Termoblok Bio TDB-100 (Biosan)

Termoblok z wytrząsarką Thermomixer® Comfort (Eppendorf)

Wirówka Universal 320 R (Hettich)

III.1.7. Oprogramowanie

Excel 2007 (Microsoft)

GenePix Pro 6.1 (Axon)

MAIA 2.75 (Curie)

R/Bioconductor 2.13.1 pakiety limma, marray, ABarray, matrixStats, pvclust

Agrigo (China Agricultural University)

MapMan (Max Planck Institute of Molecular Plant Physiology)

III.2. METODY

III. 2.1. Izolacja RNA

RNA izolowano z korzeni tytoniu przy użyciu zestawu RNeasy Plant Mini Kit firmy Qiagen. Tkankę rozcierano na proszek w moździerzu, w ciekłym azocie. Procedurę izolacji przeprowadzano wg zaleceń producenta, używając 100 mg materiału roślinnego na każdą kolumnkę. Procedurę zmodyfikowano, stosując dwukrotną elucję za pomocą 75µl wody wolnej od rybonukleaz, uzyskując końcową objętość roztworu RNA 150 µl.

Po izolacji wykonywano pomiar spektrofotometryczny za pomocą aparatu NanoDrop w celu określenia stężenia RNA w próbce przed trawieniem DNA.

III.2.1.1. Trawienie DNA przy użyciu zestawu TURBO DNA-free™ Kit

Trawienie wykonywano według procedury zalecanej przez producenta, z modyfikacjami polegającymi na:

1. Inkubacji mieszaniny reakcyjnej z 2 µl enzymu przez 20 min w 37°C.
2. Dodaniu 1 µl enzymu i dalszej inkubacji przez 15 min.

III.2.1.2. Wytrącanie RNA

1. Do próbki z roztworem RNA dodawano 2,5 objętości schłodzonego 96% etanolu oraz 0,1 objętości 3M octanu sodu, mieszano.
2. Inkubowano przez noc w temperaturze -20°C.
3. Wirowano: 18600 g, przez 30 min, w 4°C.
4. Bez naruszania osadu zbierano supernatant.
5. Osad przemywano za pomocą 1ml etanolu [70%].
6. Wirowano; 18600 g, przez 5min, w 4°C
7. Zbierano dokładnie supernatant.
8. Resztę etanolu odparowywano pozostawiając otwarte próbki w temperaturze pokojowej.
9. Osad rozpuszczano w wodzie wolnej od nukleaz, w takiej objętości, by uzyskać stężenie RNA nie mniejsze niż 1500 µg/µl.

III.2.1.3. Analiza ilościowa i jakościowa RNA

III.2.1.3.1 Pomiar stężenia RNA za pomocą aparatu NanoDrop ND-1000 (Thermo Scientific).

Przeprowadzano pomiar absorpcji 1 µl próby przy długości fali 230, 260 i 280 nm. Jako referencję stosowano wodę wolną od nukleaz. Miarą jakości RNA były:

- stosunek absorpcji A260/A280- za czyste uznawano próby o wartości 2,0-2,2
- stosunek absorpcji A260/A230- za czyste uznawano próby o wartości >2,0.

Stężenie RNA obliczane była automatycznie na podstawie wartości absorpcji przy długości fali 280 nm, według wzoru: A260 x rozcieńczenie x 40 mg/ml

III.2.1.3.2. Elektroforeza kapilarna przy użyciu aparatu Bioanalyzer 2100 (Agilent)

Jakość RNA analizowano także przy użyciu urządzenia Bioanalyzer 2100 zgodnie z protokołem Agilent RNA 6000 Nano. Za nadające się do dalszych analiz uznawano próby o wartości RIN (ang. *RNA Integrity Number*) >8.

III.2.2. Przygotowanie prób do hybrydizacji

III.2.2.1. Metodyka stosowana w przypadku mikromacierzy TOM1, TOM2 i POT

Każdorazowo przygotowywano jednocześnie próbę badaną i kontrolną, które były poddane hybrydizacji do tej samej mikromacierzy.

III.2.2.1.1. Odwrotna transkrypcja

Do każdej reakcji używano 15 µg RNA. Stosowano zestaw SuperScript Indirect cDNA Labeling System, postępując zgodnie z procedurą zalecaną przez producenta, z modyfikacją polegającą na dodaniu do mieszaniny reakcyjnej 1,5 µl enzymu SuperScript III RT [400 U/ µl), inkubacji przez 1,5 godziny w temperaturze 46°C, a następnie dodaniu tej samej ilości enzymu i powtórzeniu inkubacji. Druga modyfikacja polegała na dodaniu do uzyskanego roztworu cDNA 40 µl H₂O i wytrąceniu cDNA zgodnie z metodyką opisaną punkcie III.2.1.3.

III.2.2.1.2. Znakowanie barwnikami fluorescencyjnymi

Podczas przeprowadzania reakcji znakowania zwracano uwagę na ograniczenie ekspozycji barwników fluorescencyjnych na światło.

1. Osad cDNA uzyskany po wirowaniu rozpuszczano w 4,5 µl 2x stężonego buforu sprzęgającego.
2. Inkubowano w temperaturze pokojowej przez 15 min.

3. Dodawano 4,5 μ l barwnika (Cy3 i Cy5 do każdej pary prób, zgodnie ze schematem znakowania)
4. Inkubowano w temperaturze pokojowej przez 1,5 godzin z wytrząsaniem
5. Dodawano 6 μ l 4M chlorowodoru hydroksyloaminy (w celu zablokowania grup estrowych cyjanin).
6. Inkubowano w temperaturze pokojowej przez 15 minut z wytrząsaniem.

III.2.2.1.3. Oczyszczanie wyznakowanej próby

Po znakowaniu konieczne było usunięcie niezwiązanego barwnika przy pomocy zestawu MinElute Purification Kit (Qiagen).

1. Do probówki z wyznakowanym cDNA dodawano 1,16 μ l 3M octanu sodu (pH 5,2) w celu obniżenia pH, co umożliwiała optymalne wiązanie cDNA z membraną kolumny.
2. Dodawano 500 μ l buforu PB.
3. Wirowano przy 18600 g, przez 1 minutę w temperaturze pokojowej. Usuwano przesącz.
4. Dodawano 700 μ l buforu PE.
5. Wirowano przy 18600 g, przez 1 minutę w temperaturze pokojowej. Wymieniano probówkę zbierającą.
6. Wirowano przy 18600 g, przez 1 minutę w temperaturze pokojowej, bez dodawania buforu.
7. Przeprowadzano elucję, nakładając na membranę kolumny 20 μ l buforu EB, podgrzanego do 50°C, pozostawiano na 2 min.
8. Wirowano przy 18600 g, przez 1 minutę w temperaturze pokojowej.
9. Powtarzano elucję, nakładając na membranę kolumny 20 μ l buforu EB, podgrzanego do 50°C
10. Przeprowadzano pomiar stężenia znakowanego cDNA i związanego z nim barwnika na spektrofotometrze NanoDrop (tryb Microarray).

III.2.2.1.4. Przygotowanie mieszaniny hybrydyzacyjnej

1. Łączono wyznakowane preparaty próby kontrolnej i badanej (ok. 100 pmol każdej próby).
2. Odparowywano za pomocą koncentratora.
3. Zawieszano w 8 μ l 1mM EDTA.
4. Denaturowano w 95°C, 10 minut
5. Dodawano 118 μ l buforu hybrydyzacyjnego SlideHyb™ Hybridization Buffer#3

(podgrzanego do 68°C).

III.2.2.1.5. Dodanie modyfikowanych nukleotydów poli-A (tylko mikromacierze cDNA: POT i TOM1)

1. Do mieszaniny hybrydyzacyjnej dodawano 4 µl Poly(A)
2. Inkubowano przez 60 min. w temp. 42°C.

III.2.2.2. Metodyka stosowana w przypadku mikromacierzy TOB

III.2.2.2.1. Znakowanie barwnikami fluorescencyjnymi

Próbkę znakowano fluorescencyjnie z wykorzystaniem zestawu Quick Amp Labeling Kit firmy Agilent. Do reakcji używano 5 µg RNA, do którego dodawano RNA kontrolne z zestawu Agilent RNA Spike-in Kit: Spike A mix dodawano do prób znakowanych Cy3, Spike B mix- do prób znakowanych Cy 5.

III.2.2.2.2. Oczyszczanie wyznakowanej próby

Znakowane RNA oczyszczano za pomocą zestawu RNEasy mini kit (Qiagen), zgodnie z procedurą zalecaną przez producenta zestawu Agilent.

Przeprowadzano analizę ilościową i jakościową wyznakowanego cRNA z wykorzystaniem spektrofotometru NanoDrop. Warunkiem zakwalifikowania próby do hybrydyzacji była ilość cRNA >825 ng oraz aktywność specyficzna >8 pmol Cy3/Cy5 na 1 µg cRNA.

III.2.2.2.3. Przygotowanie mieszaniny hybrydyzacyjnej

1. Przygotowywano mieszaninę hybrydyzacyjną wg schematu:

cRNA znakowany Cy3	825 ng
cRNA znakowany Cy5	825 ng
10X Blocking Agent	6 µL
Woda wolna od nukleaz	do 28,8 µl
25X Fragmentation Buffer	1,2 µl
Łącznie	30 µl

2. Mieszaninę inkubowano w 60°C przez 30 minut w celu fragmentacji RNA.
3. Dodawano 30 µl buforu hybrydyzacyjnego 2x GEx Hybridization Buffer HI-RPM z zestawu Gene Expression Hybridization Kit. Delikatnie mieszano, unikając wprowadzenia

pęcherzyków powietrza.

4. Wirowano 18600 g, przez 1 min., w temperaturze pokojowej.

III.2.3. Przygotowanie mikromacierzy do hybrydyzacji-tylko mikromacierze TOM1, TOM2, POT

III.2.3.1. Rehydratacja- tylko mikromacierze TOM2

1. Sondy rehydratowano poprzez umieszczenie mikromacierzy nad łaźnią wodną o temp. 65°C na 5 sekund
2. Szkiełka odparowywano umieszczając na bloku grzejmym o temperaturze 65°C.
3. Mikromacierze poddawano działaniu promieniowania UV o natężeniu 125 mJ (crosslinker UVI LINK (Uvitec))
4. Szkiełka płukano w 2% SDS przez 2 min., następnie zanurzano trzykrotnie w wodzie i na 10 sek. w etanolu (96%).
5. Suszono przez wirowanie.

III.2.3.2. Prehybrydyzacja

1. Mikromacierze inkubowano w naczyniu z buforem prehybrydyzacyjnym, 42°C, 45 minut
2. Płukano w wodzie 2x 5 minut.
3. Płukano 2 minuty w izopropanolu.
4. Suszono przez wirowanie.

III.2.4. Hybrydyzacja

III.2.4.1. Metodyka stosowana w przypadku mikromacierzy TOM1, TOM2 i POT

Próby hybrydyzowano do przygotowanych wcześniej mikromacierzy w hybrydyzatorze HybArray12. Stosowano następujący protokół:

1. Gotowe komory z mikromacierzami inkubowano w temperaturze 70°C przez 2 minuty w celu uszczelnienia komór hybrydyzacyjnych.
2. Podawano mieszaninę hybrydyzacyjnej w objętości 115 µl w temperaturze 43°C.
3. Hybrydyzowano w temperaturze 43°C przez 18 godzin z mieszaniem roztworu.
4. Płukano roztworem 2xSSC, 0,1% SDS (5 cykli, przepływ roztworu przez komorę- 20 sekund, pozostawienie roztworu w komorze- 40 sekund, temperatura 43°C).
5. Płukano roztworem 0,5xSSC (5 cykli, przepływ roztworu przez komorę- 20 sekund,

pozostawienie roztworu w komorze- 40 sekund, temp. 30°C).

6. Płukano roztworem 0,05xSSC (5 cykli, przepływ roztworu przez komorę- 20 sekund, pozostawienie roztworu w komorze- 40 sekund, temp. 25°C).

Po zakończeniu hybrydyzacji szkiełka płukano w 0,05x SSC i suszono przez wirowanie.

III.2.4.2. Metodyka stosowana w przypadku mikromacierzy TOB

1. Hybrydyzację przeprowadzono w hybrydyzatorze Tecan HS 4800 Pro. Hybrydyzacja przebiegała wg następującego protokołu:
2. Mikromacierz płukano w buforze prehybrydyzacyjnym w temperaturze 65°C przez 1 minutę
3. Nakładano próbę o objętości 55 µl w temperaturze 65°C.
4. Hybrydyzowano w temperaturze 65°C przez 17 godz.
5. Płukano roztworem Agilent Wash Buffer 1 w temperaturze 23°C (2 cykle, przepływ roztworu przez komorę- 1 minuta, pozostawienie roztworu w komorze- 1 minuta)
6. Płukano roztworem Agilent Wash Buffer 2 w temperaturze 37°C (2 cykle, przepływ roztworu przez komorę- 1 minuta, pozostawienie roztworu w komorze- 1 minuta)
7. Suszono w temperaturze 30°C przez 2 minuty

III.2.5. Skanowanie

III.2.5.1. Skanowanie za pomocą skanera Scan Array Express (Perkin Elmer)

Za pomocą skanera Scan Array Express skanowano mikromacierze POT, TOM1 i TOM2. Wartość wzmocnienia fotopowielacza (PMT) dostosowywano ręcznie w taki sposób, by uzyskać podobną intensywność sygnału w obu kanałach oraz maksymalną intensywność przy minimalnej saturacji plamek. Wyniki skanowania zapisywano w formacie .tif, uzyskując pojedynczy obraz dla każdego barwnika fluorescencyjnego.

III.2.5.2. Skanowanie za pomocą skanera GenePix 4200 AL (Axon)

Za pomocą skanera GenePix skanowano mikromacierze TOB. Skanowanie przeprowadzono z automatycznym dostosowaniem wartości wzmocnienia fotopowielacza (PMT) na obu kanałach z progiem saturacji 0,2%, który ustalono na podstawie wstępnych eksperymentów. Wyniki skanowania zapisywano w formacie .tif, uzyskując pojedynczy obraz dla każdego barwnika fluorescencyjnego.

III.2.6. Metody bioinformatyczne

III.2.6.1. Analiza obrazów mikromacierzy

III.2.6.1.1. Analiza obrazów mikromacierzy przy użyciu programu GenePixPro

Analizę obrazów wykonywano przy użyciu programu GenePixPro. Analiza wymagała przypisania obrazów poszczególnych plamek do odpowiednich nazw i lokalizacji sond. W tym celu na obraz mikromacierzy nakładano wirtualną „siatkę”, tworzoną przez program na podstawie informacji zawartych w pliku *.gal dostarczonym wraz z dokumentacją mikromacierzy. Plik *.gal zawierał informacje o położeniu sond, ich nazwy i adnotacje funkcjonalne. Następnie dopasowywano kształt oczek siatki do obrazu plamek za pomocą funkcji programu GenePixPro i w razie potrzeby korygowano ręcznie ustawienie w taki sposób, by siatka dokładnie obejmowała kontury plamek. Po potwierdzeniu właściwego nałożenia siatki dokonano analizy jakościowej plamek, przyporządkowując plamkom wartości liczbowe (tzw. „flagi”), kategoryzując plamki o prawidłowej i gorszej jakości.

Stosowano dwa rodzaje „flagowania”:

- automatyczne: odrzucano plamki o saturacji $\geq 10\%$ na obu kanałach, sondy, których sekwencja nie została zwalidowana, a także 4 sektory mikromacierzy POT, które nie mieściły się w komorze hybrydyzacyjnej
- ręczne: na podstawie wyglądu mikromacierzy odrzucano plamki z widocznym artefaktem lub innym zaburzeniem spowodowanym czynnikami zewnętrznymi.

Na podstawie ostatecznego położenia siatki uzyskano wartości liczbowe odpowiadające intensywności fluoprescencji obu barwników, a także inne dane statystyczne dla poszczególnych plamek, np. intensywność tła, stopień saturacji, stosunek sygnału do szumu. Dane te zostały zapisane w plikach *.gpr, gdzie przyporządkowane były parametrom lokalizującym plamkę. Pliki te były używane w dalszych etapach analiz.

III.2.6.1.2. Analiza obrazów mikromacierzy w programie Maia 2.75

Podczas korzystania z programu Maia, podobnie jak w przypadku GenePixPro, korzystano również z plików *.gal, na podstawie których program tworzył siatkę służącą do lokalizacji plamek, która, w przeciwieństwie do programu GenePixPro, nie odzwierciedlała konturów plamek ale umiejscawiała każdą plamkę w pojedynczej, prostokątnej „komórce”. Po automatycznej i ręcznej korekcji ustawienia siatki, generowano pliki *.txt, zawierające informacje dotyczące położenia na mikromacierzy, intensywności fluorescencji, jak również szereg parametrów jakościowych plamek.

Kryteria flagowania były identyczne jak w przypadku programu GenePixPro

III.2.6.2. Analiza homologii sond znajdujących się na mikromacierzach i ich docelowych transkryptów

Analizie poddano sondy znajdujące się na 3 mikromacierzach oraz dostępne sekwencje dla tytoniu. Zbiorcze przeszukania zostały wykonane we współpracy z dr Aleksandrą Świercz z Pracowni Mikromacierzy i Głębokiego Sekwencjonowania Instytutu Chemii Bioorganicznej PAN. Przeszukania wykonano korzystając z algorytmu BLASTn. Skrypty przetwarzające i formatujące pliki zostały napisane w programie C++. Korzystano z sekwencji dla *N. tabacum* oraz *N. benthamiana* znajdujących się w bazie danych TIGR Gene Indices wersja 5.0 (<http://compbio.dfci.harvard.edu/tgi/plant.html>). Założono dwie bazy danych dla tytoniu z wykorzystaniem programu „formatdb” (dostępny w ramach programu BLAST). Jedna z nich zawierała wszystkie dostępne sekwencje *Nicotiana tabacum* i *Nicotiana benthamiana* (Źródło: TIGR Gene Indices wersja 5.0 (<http://compbio.dfci.harvard.edu/tgi/plant.html>)), a druga obejmowała wyłącznie sekwencje docelowe dla sond z mikromacierzy TOB (Źródło: TIGR Gene Indices wersja 7.0 (<http://compbio.dfci.harvard.edu/tgi/plant.html>), GenBank, Unigene).

Jako zapytanie stosowano sekwencje odpowiednich sond mikromacierzowych (lub ich fragmenty). Do przeszukiwań stosowano grupy sekwencji wymienione w Tabeli 7.

Tabela 7. Sekwencje sond mikromacierzowych używane podczas przeszukiwań BLAST

Mikromacierz	Źródło sekwencji	Sekwencje użyte do przeszukań
TOM1	Tomato functional genomics database (http://ted.bti.cornell.edu/)	Sekwencje z końców 3' i 5' sond
TOM2	Tomato functional genomics database (http://ted.bti.cornell.edu/)	Całe sondy
POT	TIGR Solanaceae Genomics Resource (http://jcvl.org/potato/)	Sekwencje z końców 3' i 5' sond

Sekwencje sond mikromacierzowych zostały dopasowane do sekwencji tytoniu z bazy danych przy użyciu programu blastall z opcją blastn. Warunkiem zakwalifikowania sondy jako wykazującej podobieństwo była wartość $e < 0,001$ (por. I.3.2.3). Wybrano najlepsze dopasowanie dla każdej sondy.

III.2.6.3. Analizy bioinformatyczne danych mikromacierzowych

Analizy ekspresji genów wykonano z zastosowaniem pakietu R/Bioconductor oraz programu Excel (Microsoft Office). W Tabeli 8 wymieniono funkcje i pakiety używane podczas wszystkich analiz wykonanych w środowisku R. Analizy ontologii genów i wzbogacania zestawu genów różnicujących wykonano z wykorzystaniem metody PAGE za pośrednictwem serwisu AgriGO (<http://bioinfo.cau.edu.cn/agriGO/>) oraz programu MapMan (Max Planck Institute of Molecular Plant Physiology, Potsdam-Golm).

Tabela 8. Wykaz ważniejszych pakietów i funkcji projektu Bioconductor stosowanych w tworzeniu scenariusza analizy danych pochodzących z mikromacierzy DNA.

Etap analizy	Nazwa funkcji	Nazwa pakietu
Ustalanie ścieżki filtracji danych	rowMeans doVennDiagram	base ABarray
Wczytywanie danych	read.GenePix read.maimages	marray limma
Wykresy diagnostyczne	maQualityPlots boxplot, plot plotMA, plotDensities	arrayQuality graphics limma
Korekcja tła	backgroundCorrect	limma
Normalizacja wewnętrzna	normalizeWithinArrays	vsn limma
Normalizacja zewnętrzna	normalizeBetweenArrays	limma
Uśrednianie powtórzeń technicznych	duplicateCorrelation	limma
Ekstrakcja genów	modelMatrix avedups lmFit makeContrasts contrasts.fit eBayes topTable	limma
Analiza genów różnicujących	decideTests	limma
Analiza skupień	pvclust hclust kmeans heatmap	pvclust clust kmeans stats

IV. WYNIKI

IV.1. EKSPERYMENTY MIKROMACIERZOWE

IV.1.1. Wybór modelu i mikromacierzy do badań porównawczych

Jako organizm modelowy do planowanych analiz wybrano tytoń. Należy on do rodziny *Solanaceae*, która jest jedną z najliczniej reprezentowanych rodzin w zakresie liczby i rodzajów dostępnych mikromacierzy u roślin (Tabela 9). Do zaplanowanych analiz porównawczych wytypowano łącznie 4 mikromacierze, najbardziej odpowiednie z punktu widzenia celów pracy oraz posiadanego sprzętu laboratoryjnego. Reprezentują one 3 gatunki: pomidora, ziemniaka i tytoń. Wybrano: (i) 2 mikromacierze dla pomidora: TOM1 (cDNA) i TOM2 (L-oligo), zaprojektowane i wytworzone przez jedną instytucję (Boyce Thompson Institute) oraz reprezentujące podobny zestaw transkryptów; (ii) mikromacierz cDNA dla ziemniaka –POT (The Institute for Genomic Research-TIGR); oraz (iii) mikromacierz L-oligo dla tytoniu (TOB) wykonaną przez firmę Agilent.

Tabela 9. Dostępne mikromacierze dla rodziny *Solanaceae* - wybrane do eksperymentów oznaczono kolorem niebieskim.

Gatunek	Nazwa mikromacierzy	Rodzaj sond	Producent	Liczba sond
Pomidor (<i>Solanum lycopersicum</i>)	GeneChip Tomato Genome Array	S-oligo	Affymetrix	10000 zestawów
	Tomato Gene Expression Microarray, 4x44K	L-oligo	Agilent	44000
	TOM1	cDNA	Boyce Thompson Institute for Plant Research	8951
	TOM2	L-oligo	Boyce Thompson Institute for Plant Research	12020
Tytoń (<i>Nicotiana tabacum</i>)	Tobacco Gene Expression Microarray, 4x44K (TOB)	L-oligo	Agilent	43730
Ziemniak (<i>Solanum tuberosum</i>)	TIGR Potato cDNA Microarray (10Kv4) (POT)	cDNA	The Institute for Genomic Research (TIGR)	10426

Aby określić przydatność poszczególnych mikromacierzy w eksperymentach CSH, przeprowadzono analizy podobieństwa sekwencji poszczególnych sond do zestawu

wszystkich sekwencji cDNA i EST tytoniu, dostępnych w bazie Plant Gene Indices, stosując algorytm BLAST. W momencie, gdy wykonywano porównania, w bazie danych dostępnych było 83000 unikalnych sekwencji dla *Nicotiana tabacum*, zdecydowano się powiększyć ten zestaw dodatkowo o 16000 sekwencji dla *Nicotiana benthamiana*, z uwagi na bliskość filogenetyczną obu gatunków [92]. W przypadku mikromacierzy cDNA (POT i TOM1) do przeszukań wykorzystano sekwencje UniGene oraz TC (ang. *tentative consensus*), będące złoženiami sekwencji EST obejmującymi znany fragment sondy cDNA (sondy na mikromacierzach POT i TOM1 identyfikowane są wyłącznie na podstawie sekwencji ich końców 3' i 5'). W przypadku mikromacierzy L-oligo (TOM2) możliwe było użycie do przeszukań zarówno sekwencji UniGene, jak i samych sond. Po zastosowaniu UniGene, wyniki analiz homologii (oznaczone jako TOM2_UG) były w zasadzie identyczne z wynikami dla TOM1, z uwagi na podobną reprezentację cDNA na tych mikromacierzach (Rycina 5). Znajomość pełnej sekwencji sond na mikromacierzy TOM2 dawała dodatkowo możliwość użycia ich do bezpośrednich analiz homologii i identyfikacji transkryptów, które mogą się wiązać z daną sondą, z pominięciem adnotacji (wyniki oznaczono jako TOM2_S).

Każdej sondzie z danej mikromacierzy przypisano sekwencję cDNA tytoniu lub *N. benthamiana* o najwyższym podobieństwie, wraz z parametrami określającymi stopień podobieństwa (wartość e oraz Bit Score, por. I.3.2.3). Sumaryczną charakterystykę sond z poszczególnych mikromacierzy pod kątem homologii do sekwencji *Nicotiana* przedstawiono w Tabeli 10. Przeprowadzone porównania potwierdziły, że wszystkie trzy wybrane mikromacierze będą przydatne podczas analizy, ponieważ znaczna część sond na każdej z nich (przynajmniej 44%) wykazywała podobieństwo do sekwencji tytoniu na poziomie $e < 0,001$. Najniższe wartości Bit Score sekwencji spełniających to kryterium wynosiły ok. 40, co w przypadku badanych sond odpowiadało pełnemu podobieństwu sekwencji 20 kolejnych nukleotydów. Zakresy wartości Bit Score uzyskiwanych dla wszystkich 3 rodzajów mikromacierzy przedstawiono na Rycinie 5. W przypadku porównań przeprowadzonych z użyciem sekwencji UniGene reprezentowanych na mikromacierzach POT i TOM1, a także TOM2_UG obserwujemy podobny rozkład wartości Bit Score dla ziemniaka i pomidora, co – biorąc pod uwagę podobny rozkład długości badanych sekwencji – świadczy o zbliżonym podobieństwie sekwencji pomiędzy tymi gatunkami, a tytoniem. W przypadku analiz TOM2_S zdecydowanie mniejsze wartości Bit Score wynikają z mniejszej długości sekwencji użytych do porównań (70-mera). W tym wypadku nawet przy dużym podobieństwie sekwencji wartości Bit Score są

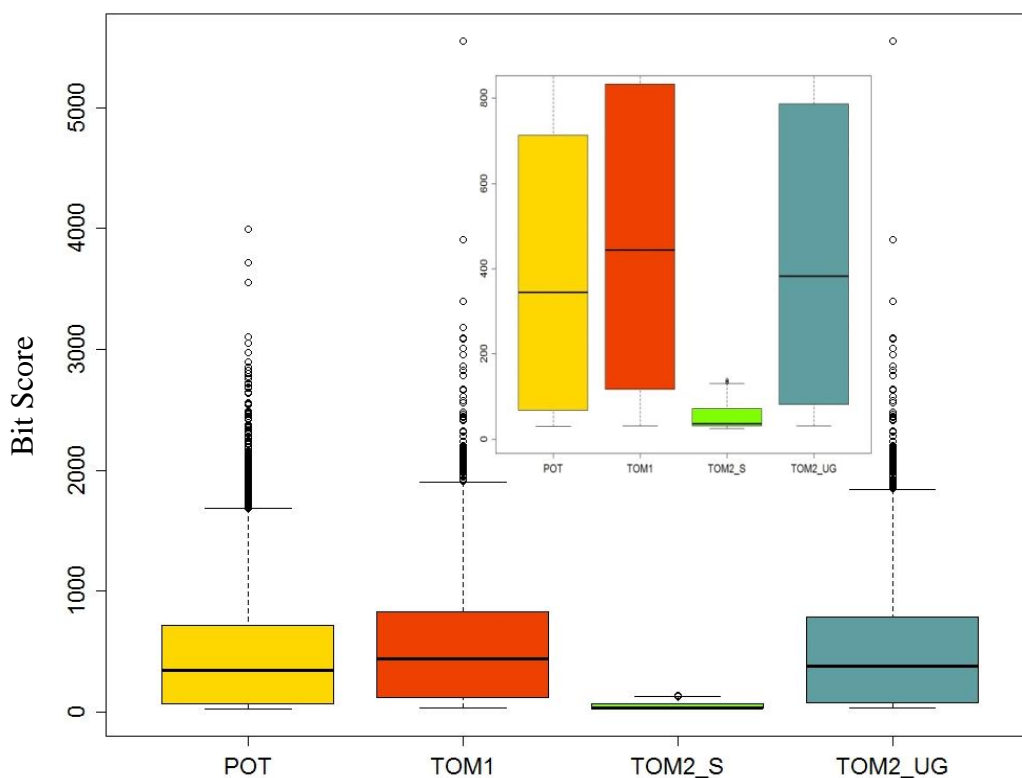
niższe - przykładowo, przy pełnym podobieństwie 70 nukleotydów uzyskuje się wartość Bit Score 139. Dla tej mikromacierzy odnotowano również zdecydowanie mniejszy odsetek sond o wartości $e < 0,001$ niż dla pozostałych. Niemniej, ponieważ wyniki porównań TOM2_S lepiej odzwierciedlają faktyczną zdolność sond oligonukleotydowych do oddziaływania z potencjalnymi docelowymi transkryptami tytoniu, w dalszych rozważaniach zdecydowano się wykorzystać te właśnie dane do charakterystyki mikromacierzy TOM2. Uzyskane dane o stopniu podobieństwa sekwencji sond CSH do tytoniu wykorzystano w dalszym etapie do filtracji danych (patrz punkt IV.3.5).

Tabela 10. Zestawienie wytypowanych do badań mikromacierzy pod kątem podobieństwa do sekwencji tytoniowych.

	POT*	TOM1	TOM2_S	TOM2_UG
Liczba sond	10426	8951	12020	7229
Średnia Bit Score	467	562	53	517
Mediana Bit Score	333	448	36	383
Maks. Bit Score	3994	5551	139	5551
Liczba sond o podobieństwie sekwencji $e < 0,001$ #	8668	8054	5944	6004
W tym o podobieństwie $e < 0,001$ do <i>N. tabacum</i>	6689	6142	4067	4543
W tym o podobieństwie $e < 0,001$ do <i>N. benthamiana</i>	1979	1576	1229	1461
Odsetek sond o podobieństwie sekwencji $e < 0,001$	83,1%	90,0%	49,0%	83,0%

*Sondy występujące na mikromacierzach dla ziemniaka są zduplikowane - podane wartości dotyczą pojedynczego powtórzenia.

Sondy zakwalifikowane jako odpowiednie do analizy CSH na podstawie homologii (por. sekcja IV.3.5).

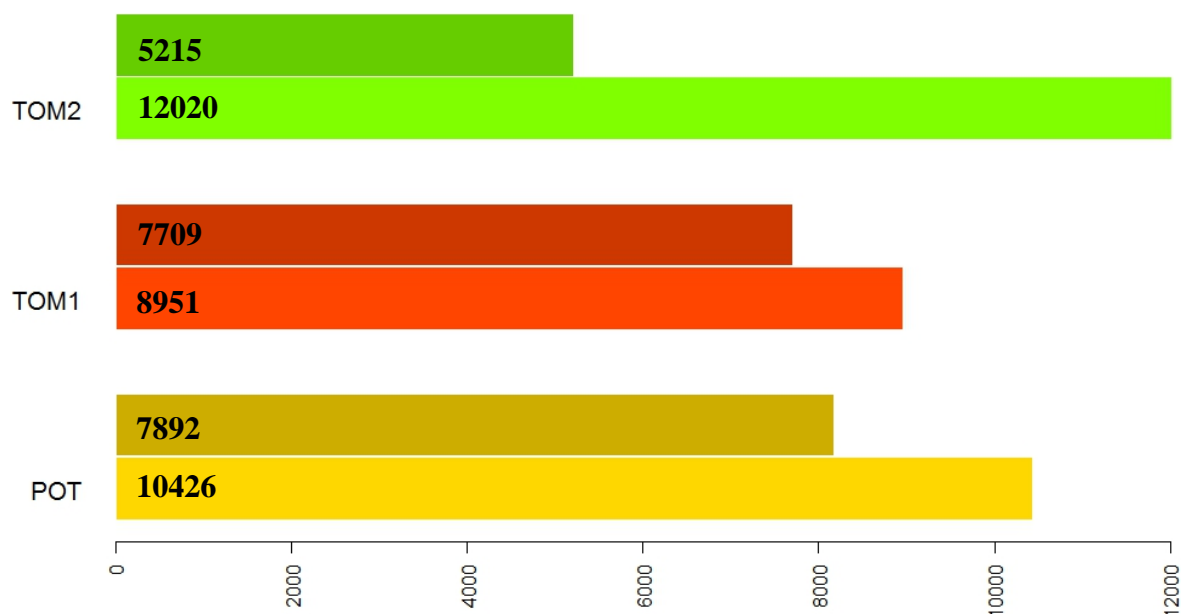


Rycina 5. Rozkład wartości Bit Score określającej podobieństwo sond do sekwencji tytoniu na poszczególnych mikromacierzach. Mniejszy rysunek przedstawia powiększony fragment wykresu. POT, TOM1- wyniki porównania z użyciem sekwencji UniGene i TC obejmujących sekwencje sond cDNA; TOM2_S – wynik porównania z użyciem właściwych sekwencji sond TOM2; TOM2_UG - wynik porównania z użyciem sekwencji UniGene reprezentowanych przez sondy TOM2.

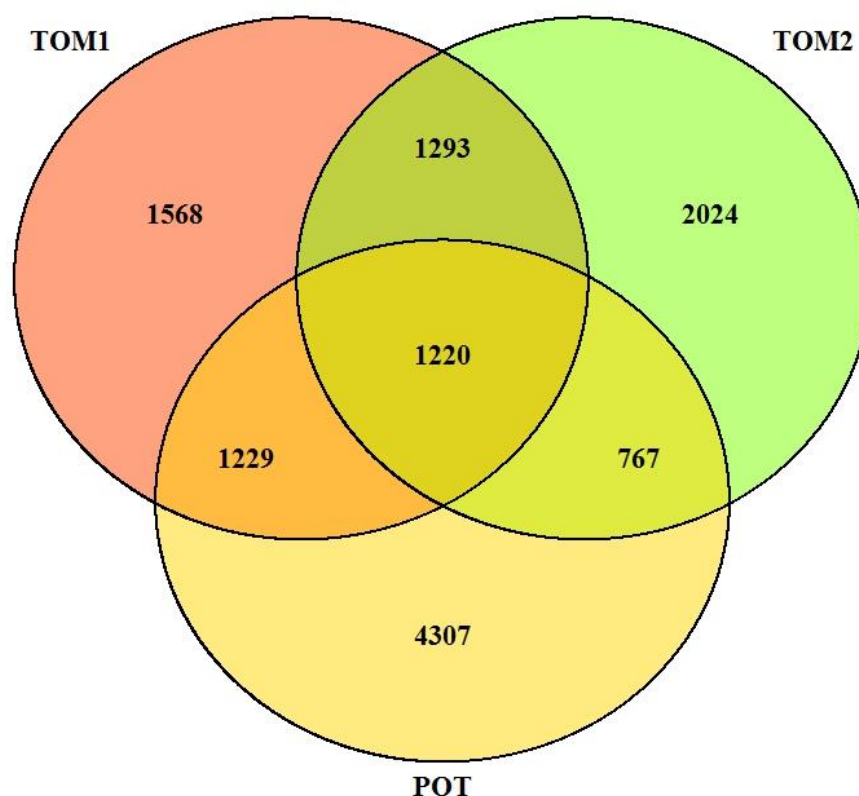
IV.1.2 Mapowanie sond pomiędzy różnymi typami mikromacierzy

Aby umożliwić porównania wyników uzyskanych przy pomocy różnych układów oraz odniesienie wyników hybrydyzacji w układzie CSH do hybrydyzacji SSH, przeprowadzono mapowanie sond pomiędzy mikromacierzami. Każdej sondzie na poszczególnych mikromacierzach CSH przypisano (jeśli to było możliwe) odpowiednik na mikromacierzy referencyjnej TOB, wykorzystując dane z przeprowadzonych wcześniej analiz homologii. W przypadku podobieństwa danej sekwencji tytoniowej do kilku sekwencji sond dla ziemniaka lub pomidora, notowano sondę o najwyższej wartości Bit Score. Na tej podstawie wyselekcjonowano zestawy odpowiadających sobie sond pomiędzy mikromacierzami CSH. Dopasowanie sond na poszczególnych mikromacierzach z sondami TOB przedstawiono na Rycinie 6. Odsetek sond z odpowiednikami na TOB wynosił 78%, 86% i 43%, odpowiednio dla mikromacierzy POT, TOM1 i TOM2. Zatem zdecydowana większość sond, wykazujących istotne podobieństwo do sekwencji

tytoniowych, posiadała swoje odpowiedniki na mikromacierzy TOB (por. Tabela 10) Biorąc pod uwagę mikromacierz TOB, łącznie 12408 sond mogło posłużyć do weryfikacji co najmniej jednej sondy z mikromacierzy heterologicznych (Rycina 7). Podsumowanie dotyczące liczby sond posiadających odpowiedniki na mikromacierzy TOB przedstawiono na Rycinie 6. W zestawieniu uwzględniono takie sondy TOB, które mogą posłużyć do weryfikacji co najmniej jednej sondy z mikromacierzy heterologicznych. 1220 sond TOB miało swoje odpowiedniki na wszystkich trzech porównywanych typach mikromacierzy.



Rycina 6. Porównanie łącznej liczby sond na poszczególnych mikromacierzach CSH (jaśniejszy kolor) i liczby sond posiadających odpowiedniki na mikromacierzy TOB (ciemniejszy kolor).



Rycina 7. Liczba sond na mikromacierzy TOB przyporządkowanych na podstawie podobieństwa ($e < 0,001$) do sond na mikromacierzach heterologicznych.

IV.1.3. Przygotowanie materiału biologicznego

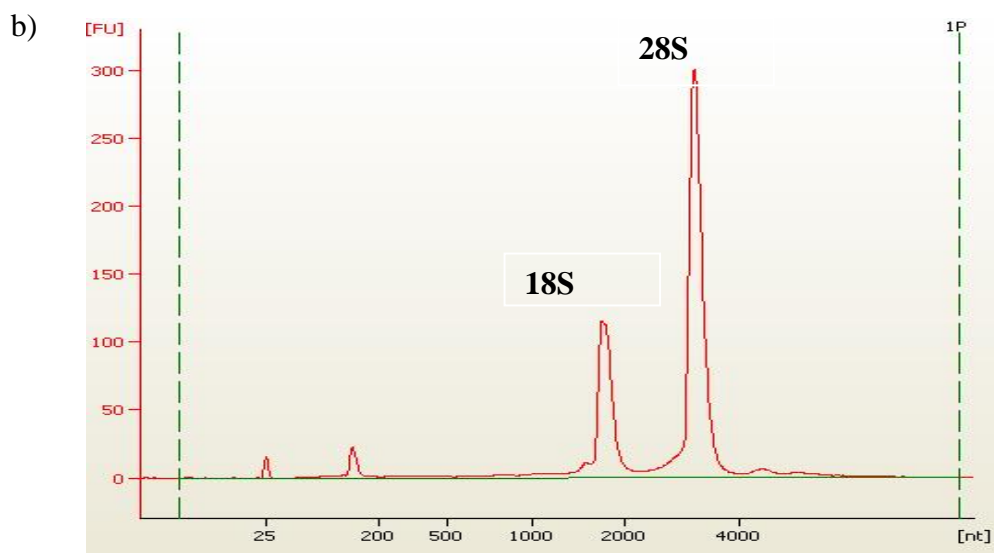
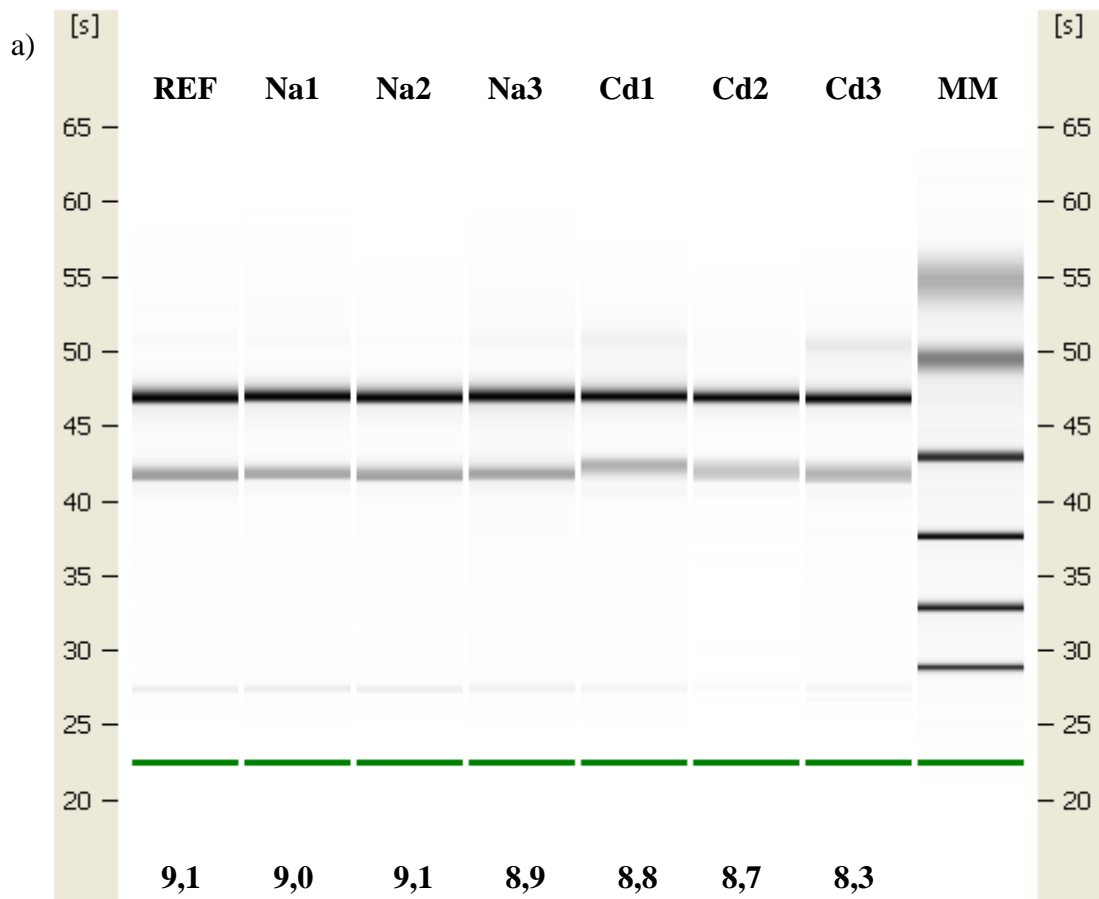
Wykorzystano dwa niezależne biologiczne układy eksperymentalne. Pierwszy z nich był ukierunkowany na analizę zmian transkryptomu korzeni tytoniu w odpowiedzi na zasolenie, a drugi - na obecność metali ciężkich. Odpowiednio, materiał do badań stanowiły korzenie tytoniu traktowane 150 mM chlorkiem sodu (NaCl) lub 100 μ M chlorkiem kadmu (CdCl_2), przez 6 godzin. Jako materiał kontrolny posłużyły korzenie z roślin nie poddanych żadnemu z tych stresów. Rośliny uzyskano w 3 niezależnych hodowlach (powtórzeniach biologicznych, oznaczonych jako 1-3), po kilkadziesiąt roślin na powtórzenie, dla każdego rodzaju warunków.

Na potrzeby hybrydyzacji przygotowano preparaty RNA z korzeni oddzielnie dla każdego typu eksperymentu i powtórzenia biologicznego (Na1 – Na3 dla stresu zasolenia, Cd1 – Cd3 dla obecności metali ciężkich). Przygotowano również jeden preparat RNA łącząc w równych ilościach wagowych rośliny niepoddane stresowi z wszystkich powtórzeń biologicznych (pula kontrolna, REF). Preparaty przygotowano z zastosowaniem

metody kolumnkowej. Stosunki absorbancji wskazane podczas pomiaru spektrofotometrycznego, a także wygląd wykresów i wartości parametru RIN obliczone za pomocą urządzenia Bioanalyzer wskazywały na wysoką czystość i integralność uzyskanego RNA (Tabela 11, Rycina 8). Wszystkie próbki spełniały zatem ogólnie przyjęte kryteria kwalifikacji do znakowania i analizy na mikromacierzach [101, 102].

Tabela 11. Stężenie i jakość preparatów RNA do analiz mikromacierzowych.

Nazwa próby	Stężenie ng/ μ l	A260	A280	260/280	260/230
REF	1612,01	40,3	19,649	2,05	2,46
Na1	1279,61	31,99	15,815	2,02	2,43
Na2	1790,56	44,764	22,298	2,01	2,41
Na3	1363,73	34,093	16,921	2,01	2,39
Cd1	1344,88	33,622	16,498	2,04	2,39
Cd2	1611,91	40,298	19,844	2,03	2,41
Cd3	1564,72	39,118	19,369	2,02	2,4



Rycina 8. Integralność preparatów RNA do analiz mikromacierzowych.

a) Obraz wirtualnego żelu z uwidocznionymi 28S i 18 S rRNA. U dołu każdej linii podano wartości parametru RIN uzyskane dla poszczególnych próbek.(MM - marker mas cząsteczkowych);

b) Wykres przedstawiający zawartość poszczególnych RNA w przykładowej próbce izolowanej z korzeni tytoniu.

V.1.4 Znakowanie i hybrydyzacja mikromacierzy

Aby zminimalizować wpływ warunków technicznych na wyniki porównań między mikromacierzami, starano się maksymalnie ujednolicić warunki przygotowania próbek oraz hybrydyzacji, jeśli pozostawało to w zgodzie z zaleceniami producentów mikromacierzy. W eksperymencie zaplanowano jednakową ilość powtórzeń i jednakowy układ hybrydyzacji dla wszystkich badanych mikromacierzy (Rycina 9a). Był to układ ze wspólną referencją, w którym zakładano użycie identycznej próbki (REF) do każdej hybrydyzacji, co miało umożliwić normalizację i analizę porównawczą wyników. We wszystkich eksperymentach CSH próbę referencyjną REF znakowano barwnikiem Cy5, a próbę badaną - barwnikiem Cy3. Dla mikromacierzy tytoniowej, która miała służyć jako układ referencyjny, przewidziano dodatkowo wykonanie powtórzeń technicznych w układzie z odwróceniem barwników (ang. *dye-swap*) (Ryc 9b). Oprócz tego, po wstępnej analizie wyników wykonano podobne, dodatkowe powtórzenia dla mikromacierzy POT. Znakowanie preparatów wykonywano stosując dwie różne metody, zależnie od zaleceń producenta danej mikromacierzy. Dla mikromacierzy TOM1, TOM2 i POT otrzymywano znakowany cDNA, a dla mikromacierzy TOB - znakowany cRNA. Następnie przeprowadzono ocenę uzyskanych próbek pod kątem ilości i jakości kwasu nukleinowego oraz stopnia wyznakowania. Istotnym parametrem oceny znakowanego kwasu nukleinowego jest specyficzna aktywność (SA) próbki, oznaczająca liczbę pmoli barwnika przypadającą na 1 µg kwasu nukleinowego. Obliczono ją według wzoru:

$$SA = \frac{\text{stężenie Cy3/Cy5 (pmol/}\mu\text{l)} * 1000}{\text{stężenie kwasu nukleinowego (ng/}\mu\text{l)}}$$

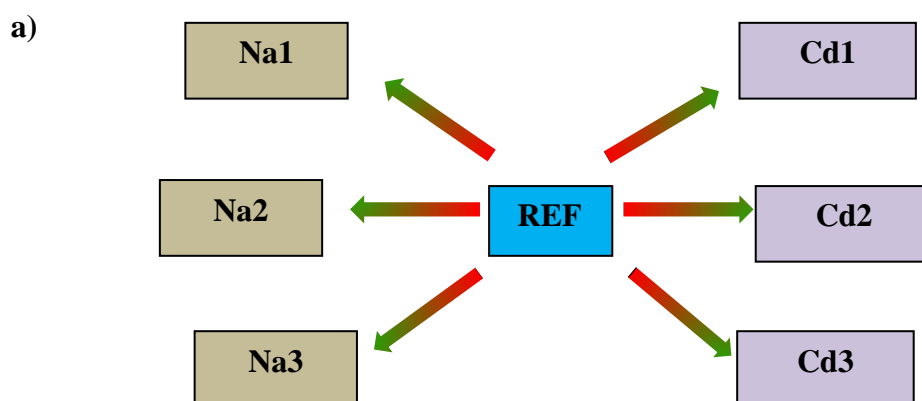
Warunkiem zakwalifikowania próbki do eksperymentu mikromacierzowego była:

- w przypadku cDNA - zawartość co najmniej 80 pmoli barwnika w każdej z próbek z pary nakładanej na mikromacierz.
- w przypadku cRNA - zawartość co najmniej 825 ng RNA w próbce oraz aktywność specyficzna >8 pmol barwnika/µg RNA.

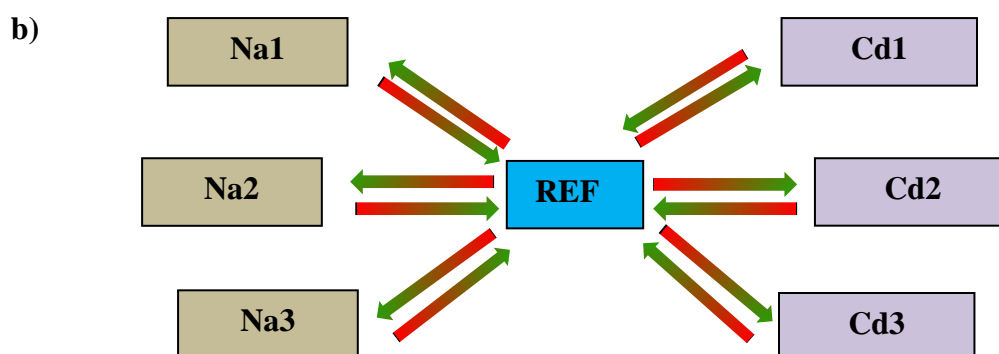
Pomiary znakowanych kwasów nukleinowych wskazywały na zgodną z oczekiwaniami, wysoką jakość i wydajność znakowania w przypadku obu metod. (Patrz Tabela 12,13).

Hybrydyzację prowadzono wg ustalonego schematu (Rycina 9). Mikromacierze TOM1, TOM2 oraz POT hybrydyzowano w stacji automatycznej HybArray 12 (Perkin Elmer) i skanowano na aparacie Scan Array Express (Perkin Elmer). Mikromacierze TOB, które charakteryzują się większą gęstością upakowania sond (4 mikromacierze na płytce o formacie szkiełka mikroskopowego), hybrydyzowano w automatycznej stacji HS 4800 Pro (Tecan), posiadającej adekwatne, czterodziałowe komory hybrydyzacyjne. Te mikromacierze były skanowane na aparacie GenePix 4200 AL (Axon). Dla wszystkich mikromacierzy uzyskano dane o tej samej rozdzielczości (5 μm) (Rycina 10).

Hybrydyzacje CSH (POT, TOM1, TOM2)



Hybrydyzacje SSH (TOB)



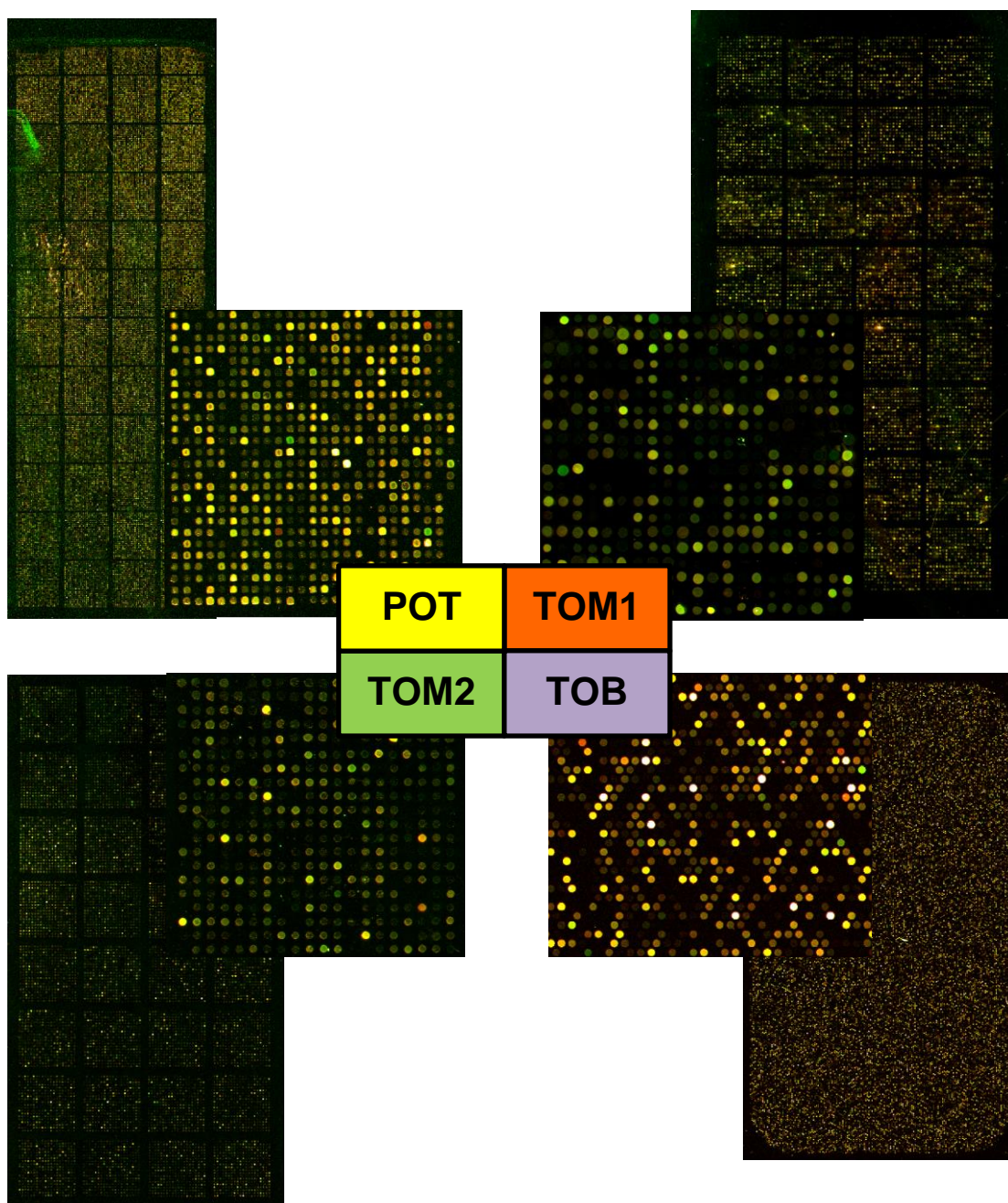
Rycina 9. Schemat hybrydyzacji transkryptów tytoniu na poszczególnych mikromacierzach. Każda strzałka oznacza jedną mikromacierz i określa sposób znakowania. Kolor czerwony - znakowanie Cy5, zielony - znakowanie Cy3; REF – rośliny nietraktowane; Na - traktowanie NaCl; Cd - traktowanie CdCl₂; 1,2,3 - numery powtórzeń biologicznych.

Tabela 12. Wyniki oceny ilościowej i jakościowej znakowanego cDNA z korzeni tytoniu. W tabeli podano parametry uzyskane dla prób poddanych hybrydyzacji do mikromacierzy POT - w przypadku TOM1 i TOM2 uzyskano podobne wartości.

Nazwa próby	Stężenie cDNA (ng/μl)	260/280	Ilość barwnika w próbce (pmol)	Aktywność specyficzna (pmol/μg cDNA)
REF	84,0	2	151,6	45,1
REF	120,4	2,02	171,6	35,6
REF	100,9	2,04	161,6	40,0
REF	113,5	2,04	173,2	38,2
REF	127,8	2,03	215,6	42,2
REF	129,8	2,06	227,6	43,8
Na1	111,1	2,01	134,8	30,3
Na2	92,3	1,95	159,2	43,1
Na3	123,6	2,03	136,8	27,7
Cd1	90,2	1,99	160	44,3
Cd2	110,9	1,97	186	41,9
Cd3	71,2	1,95	125,2	44,1

Tabela 13. Wyniki oceny ilościowej i jakościowej znakowanego cRNA.

Nazwa próby	Stężenie cRNA (ng/μl)	260/280	Ilość barwnika w próbce (pmol)	Aktywność specyficzna (pmol/μg cRNA)
REF	879,7	1,74	669	25,3
REF	868,3	1,88	780	29,9
REF	844,6	1,95	570	22,5
REF	809,6	1,78	738	30,3
REF	822,4	1,85	999	40,4
REF	827,8	1,79	780	31,4
REF	879,1	1,77	543	20,5
REF	824,5	2,04	483	19,5
Na1	841,2	1,80	750	28,9
Na1	840,0	1,94	714	28,3
Na2	871,6	1,84	678	25,9
Na2	869,4	1,81	591	22,6
Na3	868,8	1,98	837	32,1
Na3	889,5	1,98	624	23,3
Cd1	866,4	1,87	474	18,2
Cd1	780,7	2,10	348	14,8
Cd2	720,8	2,20	384	17,7
Cd2	755,8	2,13	309	13,6
Cd3	860,4	1,88	510	19,7
Cd3	818,1	2,02	399	16,2



Rycina 10. Obrazy całych mikromacierzy oraz wygląd plamek dla czterech analizowanych układów. Obrazy dla TOM1, TOM2 i POT uzyskano z wykorzystaniem skanera Scan Array Express (Perkin Elmer), obrazy dla TOB - za pomocą skanera GenePix 4200 AL (Axon). Wszystkie mikromacierze wizualizowano z wykorzystaniem programu GenePix Pro.

IV.1.5 Obróbka obrazów i analiza jakościowa mikromacierzy.

Obróbkę obrazów mikromacierzy po skanowaniu wykonano używając programu GenePix Pro. Zastosowano przy tym automatyczne oraz manualne flagowanie plamek o nieodpowiedniej jakości (przesyconych oraz zawierających artefakty), które odrzucano i nie brano pod uwagę w dalszych analizach. Po konwersji obrazów na dane liczbowe wykonano analizę wstępną oraz analizę różnicowej ekspresji genów dla każdej badanej mikromacierzy z użyciem programu R/Bioconductor, przestrzegając ogólnie przyjętych zasad analizy danych mikromacierzowych [19,27,29]. Wstępną analizę jakościową danych wykonano w celu: (i) weryfikacji technicznej poprawności wykonania eksperymentów i potwierdzenia dobrej jakości danych oraz (ii) przygotowania danych do analizy różnicowej, z uwzględnieniem specyfiki każdej mikromacierzy. W ramach analizy jakościowej wykonano szereg różnorodnych wykresów diagnostycznych pozwalających na wizualizację wyników eksperymentów. Na Rycinach 11-14 zaprezentowano niektóre z nich, a w Tabeli 14 zebrano informacje dotyczące metod analizy wstępnej zastosowanych dla poszczególnych mikromacierzy. Ponieważ stosowano jednakowe próby referencyjne dla obu typów eksperymentu (NaCl i CdCl₂), wszystkie mikromacierze danego typu analizowano wspólnie, co zwiększało moc statystyczną obliczeń. W przypadku każdej mikromacierzy oceniano kilka ścieżek analizy danych i wybierano optymalną metodę na podstawie wykresów diagnostycznych (Tabela 15), parametrów korelacji pomiędzy powtórzeniami i ostatecznej listy genów różnicujących. W efekcie, dla mikromacierzy TOB wybrano metodę obliczenia tła „morphological opening” połączoną z nieznacznym odcięciem najniższych wartości tła, z uwagi na najlepszy wygląd wykresów dla małej intensywności fluorescencji. Dla pozostałych mikromacierzy tło obliczano metodą „local feature”. Korekcję tła stosowano dla POT i TOM2, prezentujących podobny trend na wykresach pudełkowych intensywności plamek oraz tła na obu kanałach. W przypadku TOM1 trend nie był wyraźny, dlatego zrezygnowano z korekcji tła. Najlepsze wyniki normalizacji (wyrównanie wykresów pudełkowych) uzyskiwano stosując 2 metody oparte na regresji loess: „loess” (TOB i POT) lub „printtip-loess” (TOM1, TOM2) (Rycina 11-14) Przedstawione wykresy diagnostyczne świadczą o dobrej jakości wykonanych eksperymentów. Zbliżona rozpiętość wykresów pudełkowych oraz zbieżność wykresów gęstości wskazują na wybór prawidłowej metody normalizacji i jej właściwe funkcjonowanie. O powtarzalności eksperymentów świadczy z kolei wysoka korelacja sygnałów uzyskanych dla próby referencyjnej na poszczególnych mikromacierzach w obrębie danej mikromacierzy (Tabela 16).

Tabela 14. Metody obróbki danych zastosowane dla poszczególnych mikromacierzy.

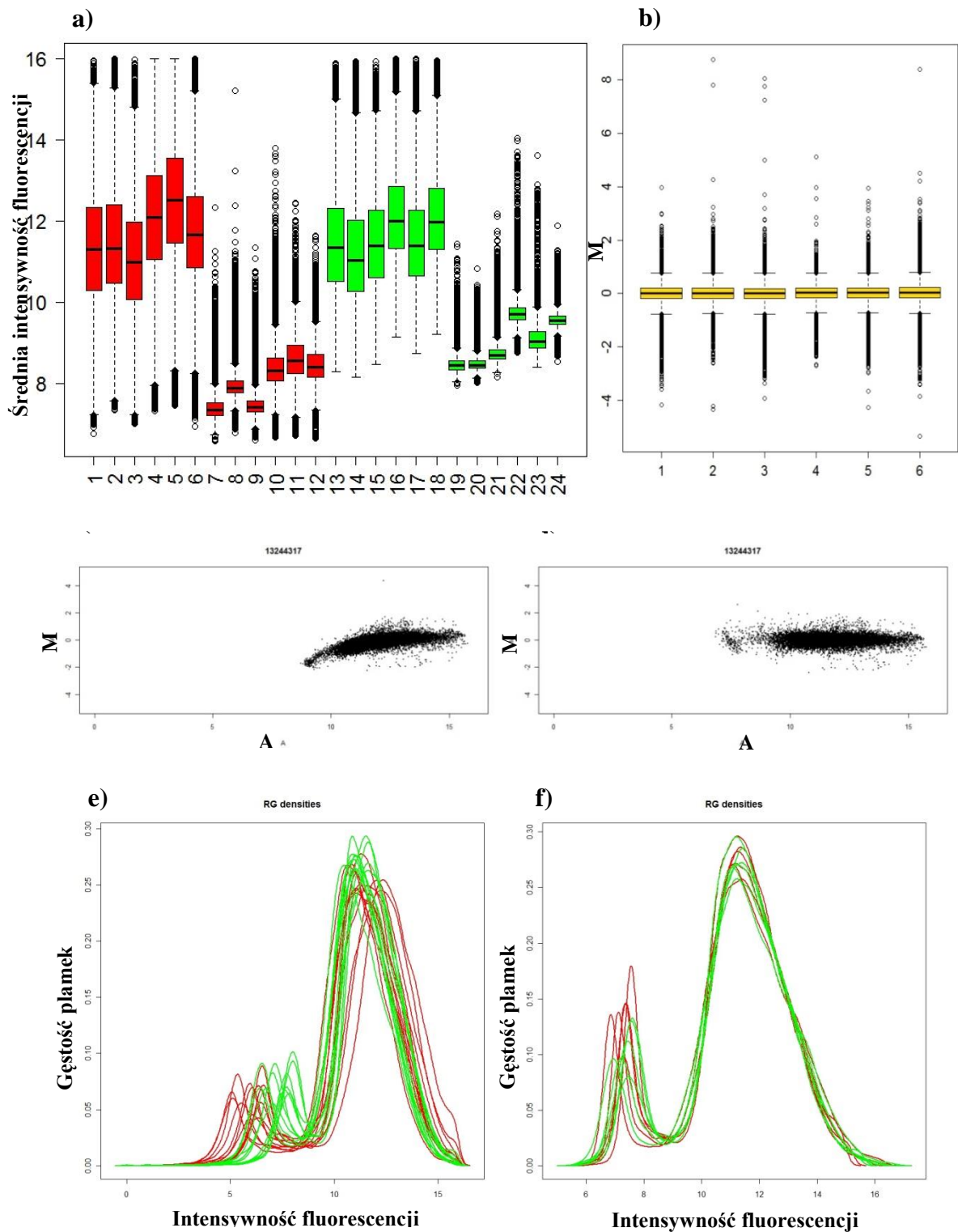
Mikromacierz	POT	TOM1	TOM2	TOB
Metoda obliczenia tła	Local feature	Local feature	Local feature	Morphological opening
Korekcja tła	Normexp, offset=50	-	Normexp, offset=50	Subtract, offset=10
Normalizacja wewnętrzna	Loess	Printtip loess	Printtip loess	Loess
Normalizacja zewnętrzna	Scale	Scale	Scale	Scale

Tabela 15. Wybrane wykresy diagnostyczne stosowane podczas analizy mikromacierzy.

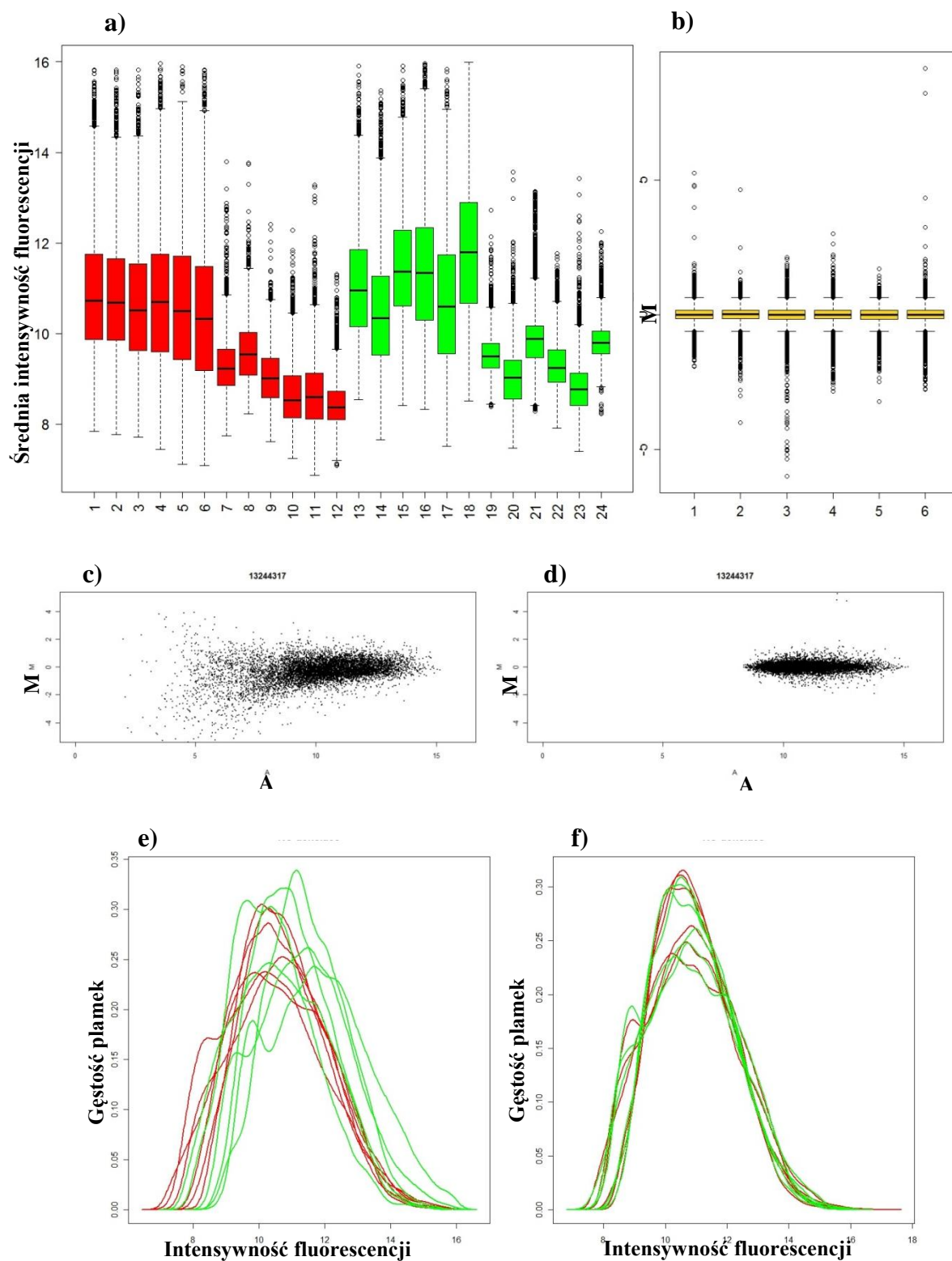
Wykres MA	Zależność pomiędzy różnicą intensywności fluorescencji pomiędzy kanałem Cy5 i Cy3 ($M=(\log_2(\text{cy5})-\log_2(\text{cy3}))$) oraz średnią intensywnością fluorescencji ($A=((\log_2(\text{cy5})+\log_2(\text{cy3}))/2)$) dla poszczególnych sond na mikromacierzy
Wykres pudełkowy	Zbiorczy wykres rozkładu danego parametru (np. różnicy intensywności fluorescencji M) na mikromacierzy, z uwzględnieniem obserwacji odstających, kwartyli i mediany.
Wykres gęstości	Rozkład intensywności fluorescencji na obu kanałach mikromacierzy
Wykres typu imageplot	Przestrzenny rozkład intensywności tła na całej powierzchni mikromacierzy

Tabela 16. Średnie wartości współczynników korelacji Pearsona prezentujących powtarzalność intensywności sygnału dla próby referencyjnej na poszczególnych mikromacierzach w obrębie każdej mikromacierzy; Wartość p wynosiła każdorazowo $<2,2e-16$

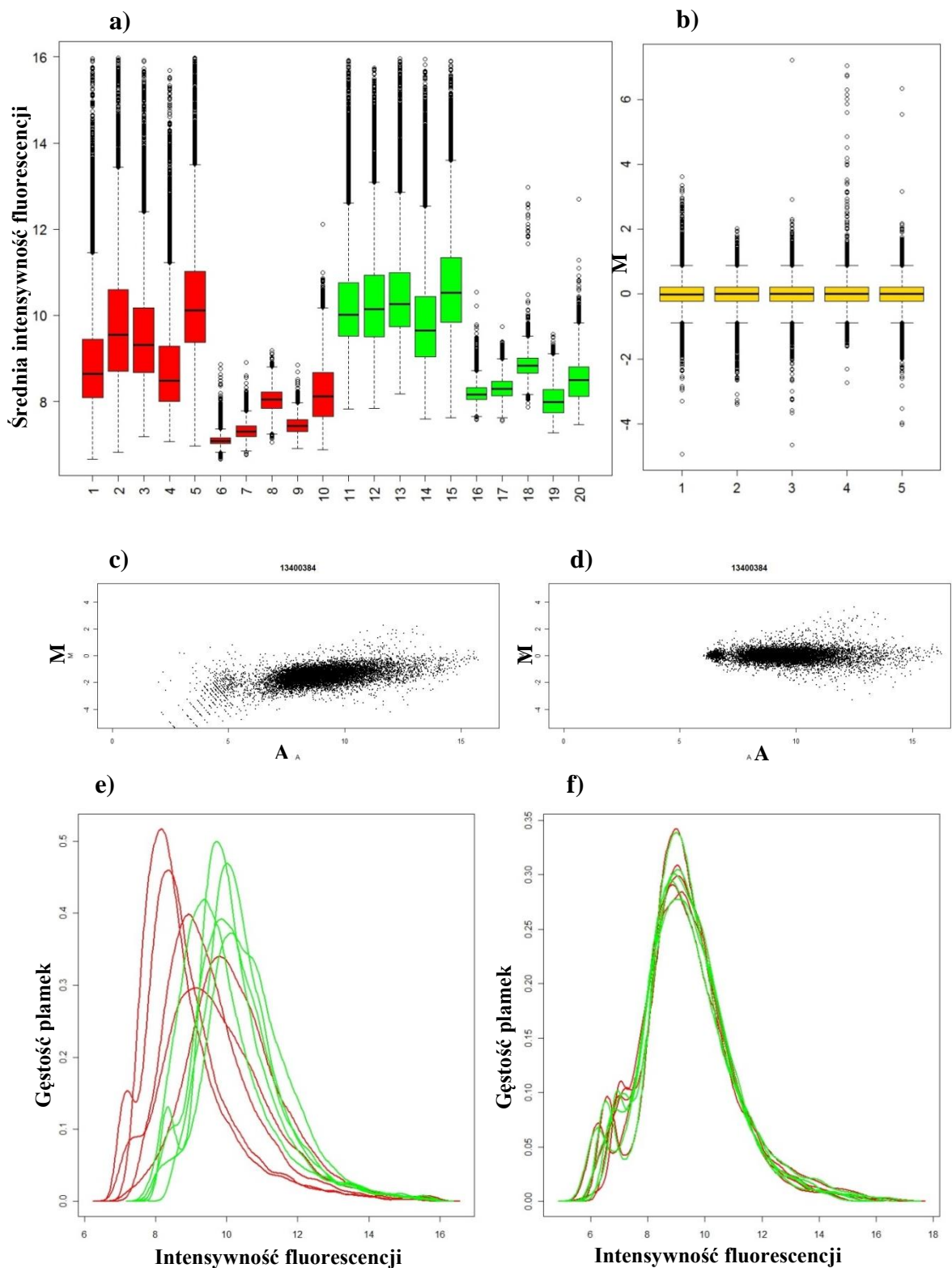
TOB	TOM1	TOM2	POT
0,97	0,78	0,88	0,90



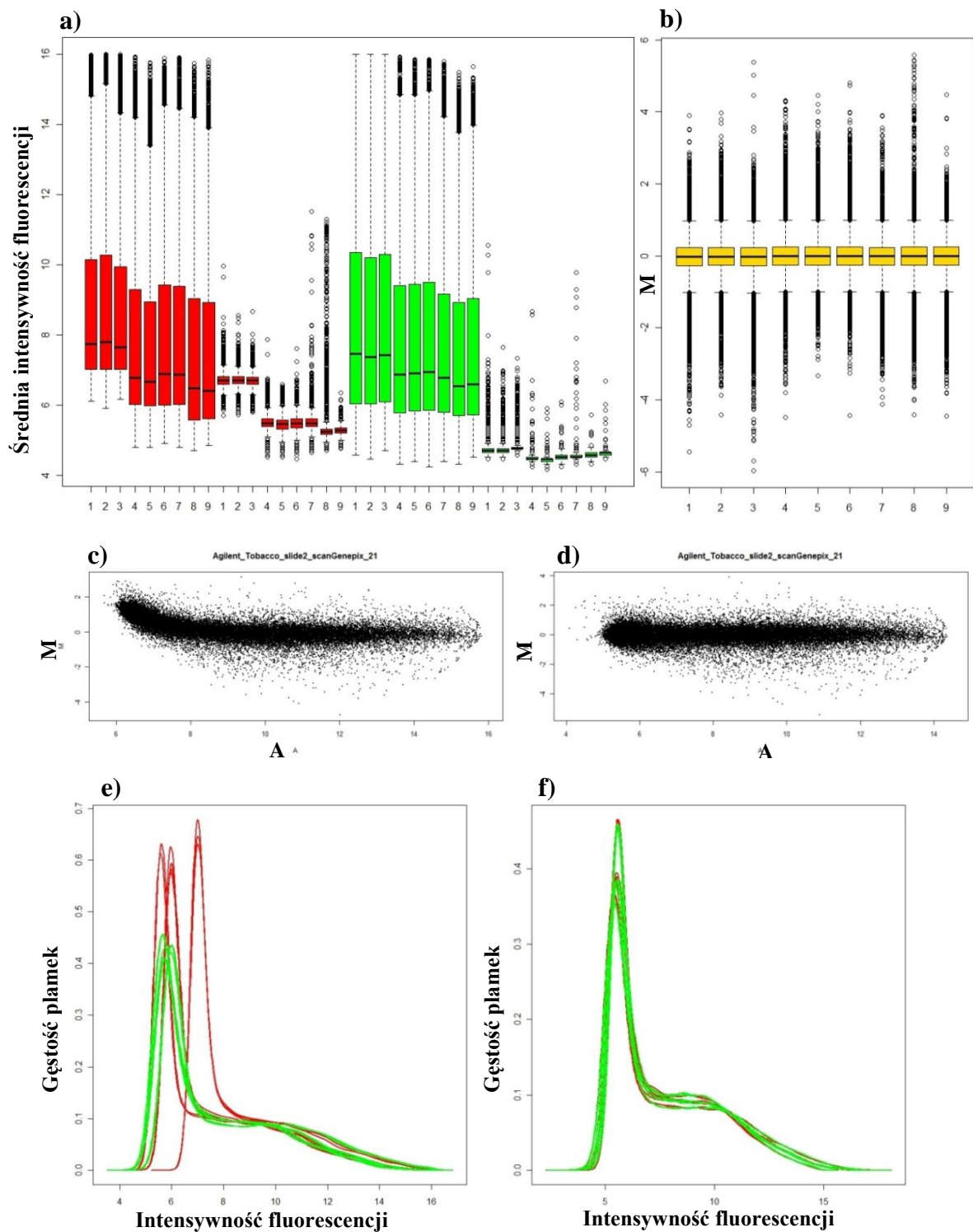
Rycina 11. Analiza jakościowa mikromacierzy POT. a) wykres pudełkowy intensywności fluorescencji surowych danych (kolor czerwony-Cy5, kolor zielony-Cy3; 1-6 i 13-18: średnia intensywność plamek, 17-12 i 19-24: średnia intensywność tła); b) wykres pudełkowy różnicy intensywności fluorescencji M po normalizacji; c) wykres MA surowych danych; d) wykres MA po normalizacji; e) wykres gęstości surowych danych; f) wykres gęstości danych po normalizacji (patrz Tabela 15)



Rycina 12. Analiza jakościowa mikromacierzy TOM1. a) wykres pudełkowy intensywności fluorescencji surowych danych (kolor czerwony-Cy5, kolor zielony-Cy3; 1-6 i 13-18: średnia intensywność plamki, 17-12 i 19-24: średnia intensywność tła); b) wykres pudełkowy różnicy intensywności fluorescencji M po normalizacji; c) wykres MA surowych danych; d) wykres MA po normalizacji; e) wykres gęstości surowych danych; f) wykres gęstości danych po normalizacji (patrz Tabela 15)



Rycina 13. Analiza jakościowa mikromacierzy TOM2. a) wykres pudełkowy intensywności fluorescencji surowych danych (kolor czerwony-Cy5, kolor zielony-Cy3; 1-5 i 11-15: średnia intensywność plamki, 6-10 i 16-20: średnia intensywność tła); b) wykres pudełkowy różnicy intensywności fluorescencji M po normalizacji; c) wykres MA surowych danych; d) wykres MA po normalizacji; e) wykres gęstości surowych danych; f) wykres gęstości danych po normalizacji (patrz Tabela 15). Z uwagi na niezadawalającą jakość hybrydyzacji odrzucono 1 z powtórzeń biologicznych TOM2

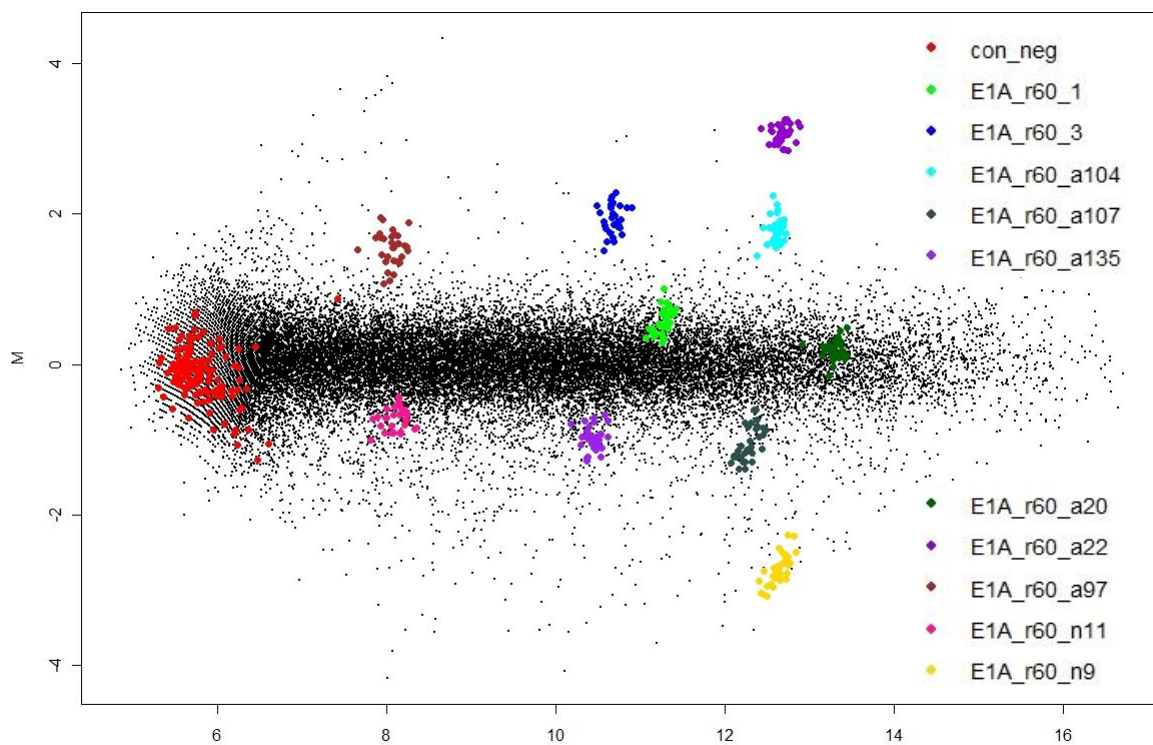


Rycina 14. Analiza jakościowa mikromacierzy TOB. Analiza jakościowa mikromacierzy TOM1. a) wykres pudełkowy intensywności fluorescencji surowych danych (kolor czerwony-Cy5, kolor zielony-Cy3; 1-6 i 13-18: średnia intensywność plamki, 17-12 i 19-24: średnia intensywność tła); b) wykres pudełkowy różnicy intensywności fluorescencji M po normalizacji; c) wykres MA surowych danych; d) wykres MA po normalizacji; e) wykres gęstości surowych danych; f) wykres gęstości danych po normalizacji (patrz Tabela 15) Z uwagi na niemożność użycia skanera GenePix do skanowania jednego z powtórzeń typu dye-swap dla NaCl, uzyskane dane odbiegały od pozostałych i zrezygnowano z ich analizy (nie pokazano).

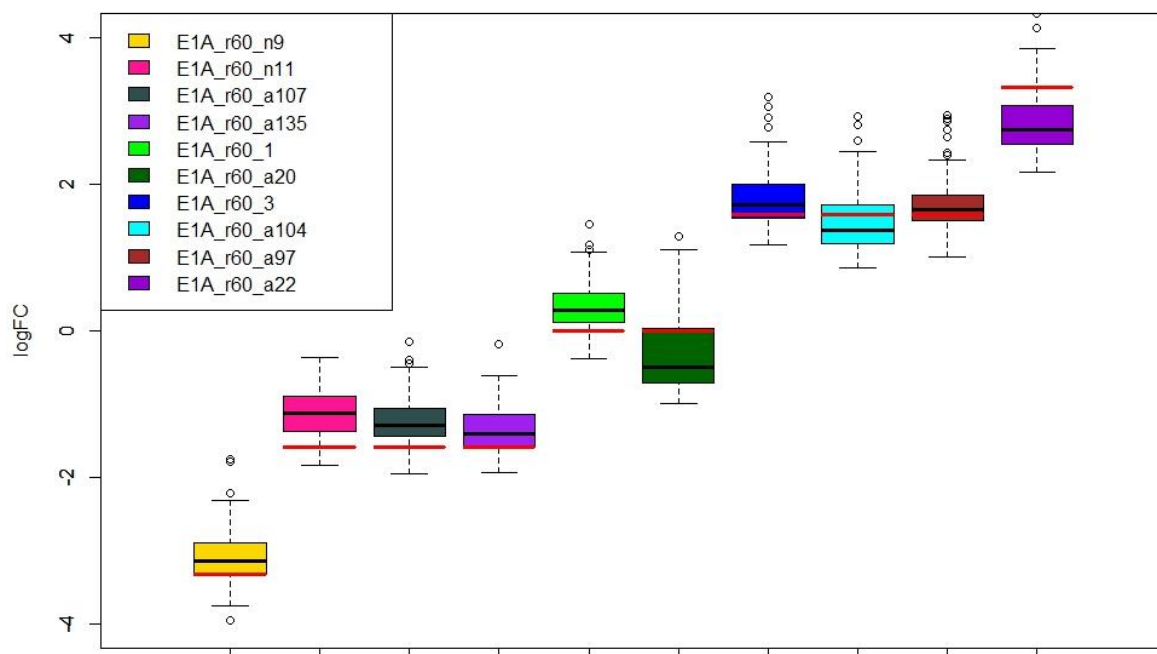
W przypadku mikromacierzy TOB zastosowano dodatkową kontrolę jakości każdego eksperymentu, zgodnie z zaleceniami producenta mikromacierzy. Do każdego RNA dodano po 10 syntetycznych 55-nukleotydowych RNA o sekwencji odpowiadającej RNA ludzkiego adenowirusa. Wszystkie mikromacierze TOB posiadały sondy pozwalające na detekcję sygnałów pochodzących od tych syntetycznych RNA, stanowiących kontrolę zewnętrzną. Znaną ilość każdego z kontrolnych RNA podano do każdej pary prób hybrydyzowanych do tej samej mikromacierzy przed znakowaniem Cy3 i Cy5 (Tabela 17). Na wykresach diagnostycznych zobrazowano dużą zbieżność stosunków stężeń obliczonych po analizie danych mikromacierzowych z wartościami rzeczywistymi, co potwierdziło dobrą jakość eksperymentów z wykorzystaniem mikromacierzy TOB. (Rycina 15,16) Korelacja Pearsona pomiędzy uzyskanymi w eksperymentach średnimi wartościami M (różnice w ekspresji pomiędzy Cy3 a Cy5) oraz rzeczywistymi stosunkami stężeń wynosiła 0,99 ($p= 1,194e-07$). Również znajdujące się na mikromacierzy sondy stanowiące kontrolę negatywną (tj. nie posiadające komplementarnych RNA w puli znakowanych cząsteczek) charakteryzuje prawidłowy, bardzo niski poziom ekspresji na obu kanałach (Rycina 15). Za pomocą powyższej analizy potwierdzono bardzo dobrą jakość wyników uzyskanych na mikromacierzy referencyjnej.

Tabela 17. Teoretyczne i uzyskane w eksperymencie różnice w intensywności kontroli zewnętrznych, wyrażone w postaci wartości M.

Nazwa kontroli spike-in	Wartość M zakładana na podst. różnicy stężeń RNA kontrolnych.	Średnia wartość M uzyskana w eksperymencie (po normalizacji)
(+)E1A_r60_n9	-3,32	-3,08
(+)E1A_r60_n11	-1,58	-1,04
(+)E1A_r60_a107	-1,58	-1,26
(+)E1A_r60_a135	-1,58	-1,35
(+)E1A_r60_1	0	0,31
(+)E1A_r60_a20	0	-0,37
(+)E1A_r60_3	1,58	1,76
(+)E1A_r60_a97	1,58	1,52
(+)E1A_r60_a104	1,58	1,45
(+)E1A_r60_a22	3,32	2,79



Rycina 15. Wykres MA przykładowej mikromacierzy TOB po normalizacji. Kolory obrazują intensywność poszczególnych kontroli zewnętrznych na tle pozostałych genów. Kolorem czerwonym oznaczono sondy stanowiące kontrolę negatywną (con_neg).



Rycina 16. Rozkłady wartości M uzyskiwanych dla poszczególnych kontroli zewnętrznych na mikromacierzach TOB. Czerwonymi kreskami oznaczono wartości M dla danych sond odzwierciedlające teoretyczne stosunki stężeń poszczególnych RNA.

IV.1.6 Analiza różnicowa

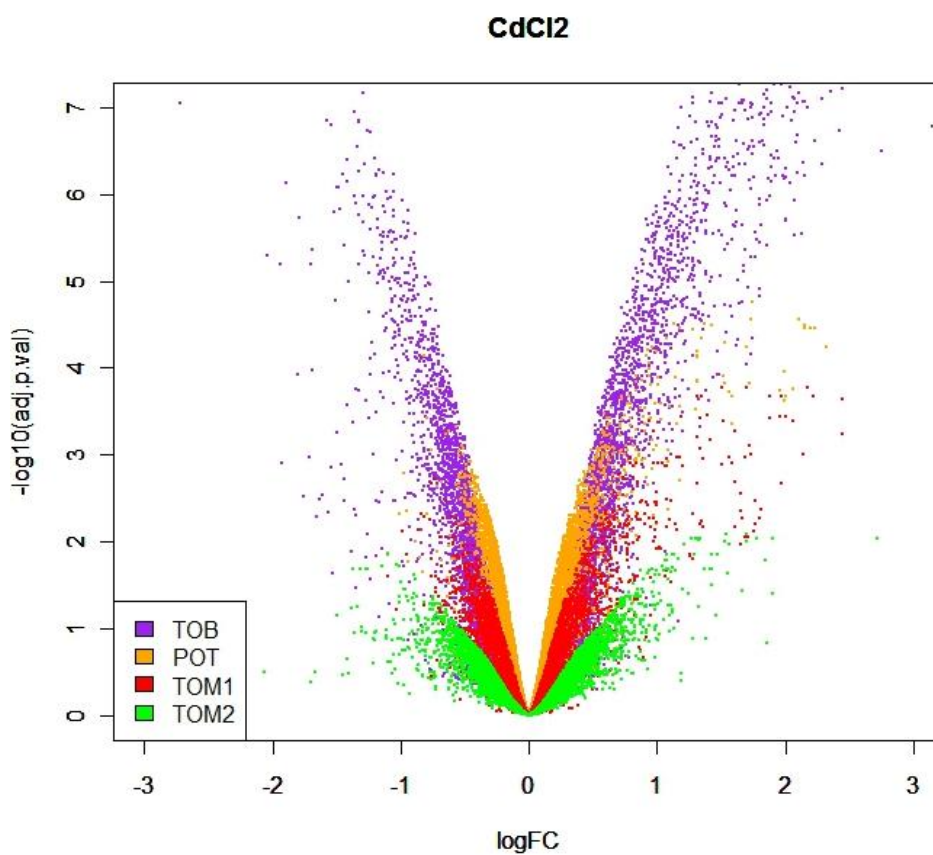
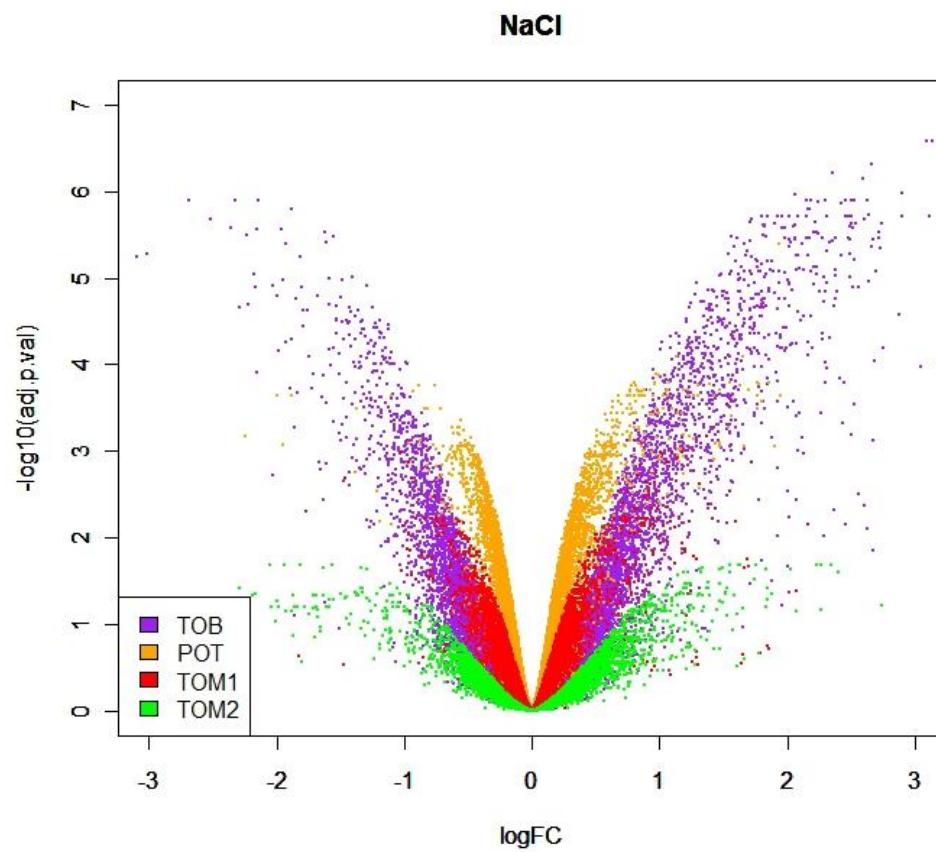
Znormalizowane dane poddano analizie różnicowej w celu uzyskania dla każdej mikromacierzy listy sond (sondy różnicujące) reprezentujących transkrypty o znacząco zmienionej ekspresji pod wpływem zastosowanego czynnika, w porównaniu z pulą kontrolną. Ze względu na słabo poznaną strukturę transkryptomów tytoniu, ziemniaka i pomidora oraz związaną z tym możliwość reprezentacji jednego genu przez więcej niż jedną sondę, podczas omawiania wyników hybrydyzacji unikano odnoszenia ilości sond do ilości genów ulegających ekspresji różnicowej, pozostając przy pojęciu „sondy różnicującej”. Zestawy takich sond wyselekcjonowano po zastosowaniu wieloetapowych procedur testowania wykorzystujących m.in. wnioskowanie bayesowskie (podejście oparte na twierdzeniu Bayesa, wykorzystujące do szacowania wartości parametrów, oprócz danych pochodzących z próby, także pewne z góry ustalone rozkłady) [93]. Spośród uzyskanych danych opisujących każdą sondę brano pod uwagę wartość logFC, będącą miarą zmiany ekspresji transkryptu oraz zmodyfikowaną wartość p , po uwzględnieniu poprawki dla testowań wielokrotnych (por. I.1.1.2.), obrazującą istotność statystyczną uzyskanego wyniku logFC. Zależność pomiędzy tymi wartościami demonstrują wykresy wulkaniczne (ang. *volcano plots*) (Rycina 17), których uzupełnienie stanowi Tabela 18. Mikromacierz TOB cechuje największy rozrzut obydwu wartości; w szczególności zakres zmodyfikowanych wartości p jest niewspółmiernie duży w stosunku do pozostałych mikromacierzy. Taki stan rzeczy był spodziewany z uwagi na użycie mikromacierzy SSH, zastosowanie powtórzeń technicznych, a także fakt, iż mikromacierz TOB została wykonana z zastosowaniem nowoczesnej techniki polegającej na syntezie sond in situ. Wspomniane aspekty mogą w znaczący sposób poprawić jakość eksperymentu i wpływać na liczbę uzyskanych genów różnicujących. Z kolei TOM2 wykazuje zakres wartości logFC podobny do pozostałych mikromacierzy, ale wartości te charakteryzują się najslabszą istotnością statystyczną - w każdym traktowaniu zaledwie ok.100 sond osiągnęło wartość $p < 0,05$. Mikromacierze TOM1 oraz POT osiągają pośrednie wyniki pod względem obu wartości, a ich wykresy mają podobny pokrój, choć zarówno w przypadku traktowania NaCl jak i CdCl₂ lepsze wartości p uzyskano dla mikromacierzy POT. Charakterystyczną cechą mikromacierzy TOB, POT i TOM1 jest duża asymetria rozkładu skrajnych wartości logFC, z wyraźną przewagą wartości dodatnich. W przypadku TOM2 dysproporcja ta jest mniej wyraźna (Patrz Rycina 17 i Tabela 18). Zauważalny jest również odmienny pokrój wykresów wulkanicznych dla eksperymentów z zastosowaniem NaCl i CdCl₂. Szczególnie silne różnice obserwujemy dla mikromacierzy oligonukletydowych -

TOB i TOM2, dla których różnice w ekspresji genów, dotyczące zarówno jej wzrostu, jak i spadku (rozpiętość wartości logFC, Rycina 17), są znacznie większe w eksperymencie NaCl, niż w przypadku traktowania CdCl₂.

Tabela 18. Zestawienie wyników analizy różnicowej dla korzeni tytoniu traktowanych NaCl oraz CdCl₂ uzyskanych na poszczególnych mikromacierzach.

	NaCl				CdCl ₂			
	TOB	POT	TOM1	TOM2	TOB	POT	TOM1	TOM2
Liczba sond	43730	10426	8951	12020	43730	10426	8951	12020
Liczba sond p<0,05 o podwyższonej ekspresji	2478	1583	295	54	3211	1342	412	77
Liczba sond p<0,05 o obniżonej ekspresji	1625	1521	235	37	2706	1244	215	23
Odsetek sond o wartości p<0,05 (%)	9,4	29,8	5,9	0,8	13,5	24,8	7,0	0,8
Odsetek sond o podwyższonej ekspresji (%)	60,4	51,0	55,7	59,3	54,3	51,9	65,7	77,0
Liczba sond o logFC > 1	929	57	52	128	613	46	72	63
Liczba sond o logFC < -1	305	10	8	101	178	2	4	46

Często przyjmowanym kryterium ekspresji różnicowej genów w badaniach mikromacierzowych jest osiągnięcie pewnej granicznej wartości p (zazwyczaj p<0,05), poniżej której obserwowane zmiany uważamy za znaczące. Analizowane listy sond wykazywały pod tym względem znaczne rozbieżności (Tabela 18), świadczące o zróżnicowanej skuteczności wykrywania istotnych sond przez poszczególne typy mikromacierzy. Najwięcej, bo od 4 do prawie 6 tysięcy takich sond otrzymano na mikromacierzy TOB - stanowiły one, w zależności od eksperymentu, 10-16% wszystkich sond obecnych na mikromacierzy. W eksperymentach CSH, największą liczbę sond różnicujących stwierdzono na mikromacierzy POT. Najmniejszą liczbę zmian wykryto za pomocą mikromacierzy TOM2 – dla obu rodzajów stresu nie przekraczała ona 100, co stanowi mniej niż 1% całkowitej liczby sond. We wszystkich porównaniach liczba sond o istotnie zwiększonej ekspresji w stosunku do próby referencyjnej była większa od liczby sond o znacząco obniżonej ekspresji. W przypadku traktowania NaCl odsetek sond o zwiększonej ekspresji wynosił od 51% do 60%, natomiast w przypadku traktowania CdCl₂ – od 52% do 77%. Największą rozbieżność pomiędzy wynikami dla obu traktowań zaobserwowano na mikromacierzy TOM2. Miała na to wpływ ogólnie niska liczba sond wykazujących statystycznie znaczącą zmianę ekspresji.



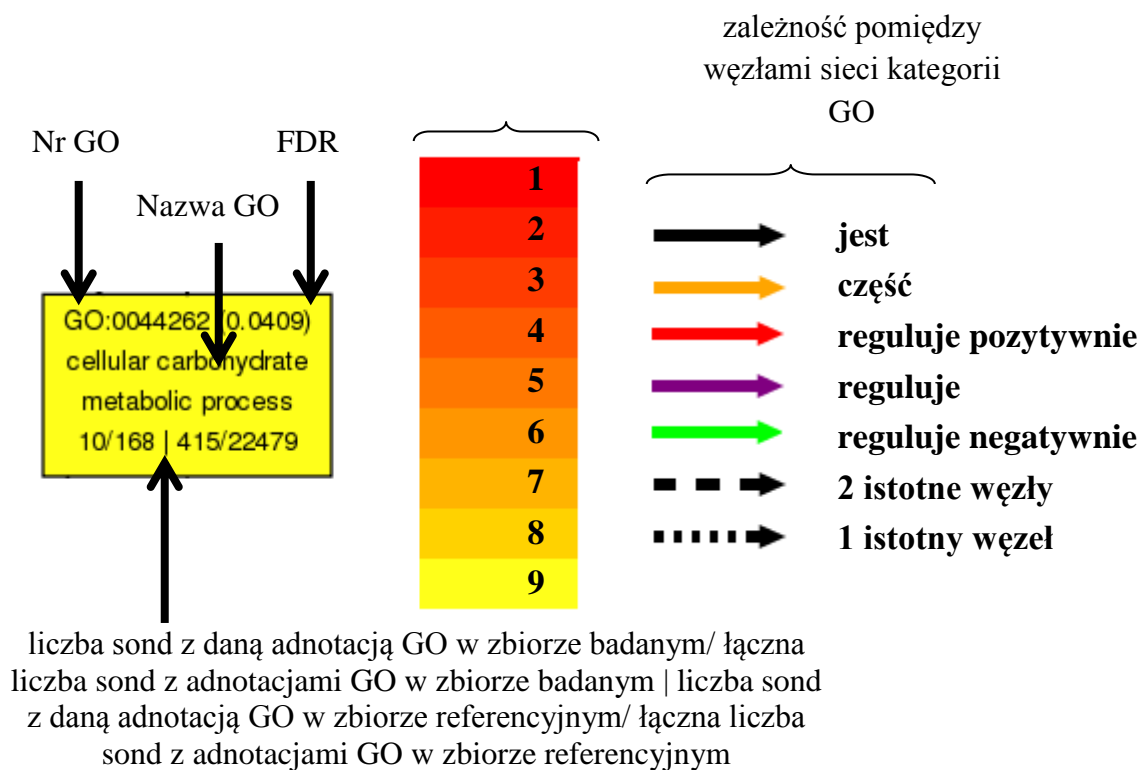
Rycina 17. Wykresy wulkaniczne prezentujące zależność między logFC a przekształconą logarytmicznie wartością p dla wszystkich mikromacierzy w dwóch eksperymentach (stres solny i kadmowy).

Ponieważ tak różne liczby sond w puli ograniczały możliwość przeprowadzenia porównań profilów ekspresji pomiędzy mikromacierzami, co było głównym celem niniejszej pracy, zmieniono kryterium ekspresji różnicowej dla wyników CSH, generując dla każdej mikromacierzy pulę zawierającą identyczną liczbę 600 sond najbardziej wyróżniających się zarówno pod względem wartości logFC jak i zmodyfikowanej wartości p. Uzyskano je sortując malejąco wszystkie sondy według bezwzględnej wartości log FC, a następnie sortując rosnąco 2000 elementów z czoła listy pod względem zmodyfikowanej wartości p. Listy „top600” sporządzono w celu zidentyfikowania sond o najbardziej istotnych zmianach, świadczących o różnicach w ekspresji reprezentowanych przez nie genów. Listy „top600” ułatwiły wykonywanie porównań szczególnie w przypadku mikromacierzy TOM2, umożliwiając statystyczną analizę wyników porównania.

IV.1.7 Analiza funkcjonalna

Uzyskane za pomocą CSH i SSH zestawy sond różnicujących dla korzeni tytoniu poddanych działaniu stresu liczyły od kilkudziesięciu do kilka tysięcy sond (patrz Tabela 18). Z uwagi na często dużą liczbę odnotowanych zmian, a także biorąc pod uwagę fakt, że nie wszystkie sondy posiadały adnotacje dotyczące potencjalnej funkcji danego transkryptu (np. 23499 spośród 43760 sond TOB), w celu opisanego profilu ekspresji konieczne było usystematyzowanie wyselekcjonowanych sond i stworzenie powiązanych funkcjonalnie zbiorów. Z tego powodu skorzystano z analizy opartej na danych dotyczących ontologii genów (ang. *Gene Ontology*, GO), umożliwiającej wyróżnienie kategorii GO reprezentujących grupy genów, które mogą być różnicujące dla danego procesu. Grupy GO podzielone są na trzy typy ontologii: „funkcja molekularna”, „proces biologiczny” i „lokalizacja komórkowa”, umożliwiającą przyporządkowanie danego genu do określonego procesu, funkcji lub miejsca w komórce. Stosowano metodę SEA (ang. *Singular Enrichment Analysis*) dostępną w ramach serwisu AgriGO [104]- umożliwiającą wskazanie różnicujących kategorii GO w grupie sond wyselekcjonowanych na podstawie wybranych kryteriów (w tym przypadku wartości p oraz kierunek zmian ekspresji (indukcja lub hamowanie). Metoda umożliwia wyróżnienie takich grup GO, których częstość występowania w danym zestawie różni się w sposób istotny (wartość $p < 0,05$) od częstości uzyskanej przypadkowo, przy uwzględnieniu referencyjnego zestawu grup GO- odpowiadającego sondom znajdującym się na mikromacierzy TOB, bądź też na mikromacierzach CSH. Analiza SEA pozwoliła na wyselekcjonowanie istotnie zmienionych kategorii funkcjonalnych GO wspólnych dla obu stresów oraz specyficznych dla każdego z nich. Kategorie te charakteryzują uzyskane zestawy sond różnicujących

ulegających zwiększonej lub zmniejszonej ekspresji. Serwis AgriGO daje możliwość wizualizacji wyników w postaci wykresów i schematów blokowych, których przykłady przedstawiono w dalszej części pracy. Na Rycinie 18 wyjaśniono znaczenie barw, kształtów i informacji przedstawionych na schematach.



Rycina 18. Wyjaśnienie schematów GO. Poziom 1 w skali kolorów oznacza największą istotność statystyczną, poziom 9- najmniejszą. FDR (ang. *false discovery ratio*)- odsetek wyników fałszywie pozytywnych

IV.2. OCENA ZMIAN EKSPRESJI GENÓW W KORZENIACH TYTONIU PODDANYCH DZIAŁANIU STRESU SOLNEGO I METALI CIĘŻKICH

Mimo, iż podstawowym założeniem niniejszej pracy była ocena zastosowania CSH w badaniach transkryptomów, w ramach wykonanych doświadczeń uzyskano również profile ekspresji genów stworzone na podstawie wyników hybrydyzacji SSH, zgodnie z powszechnie przyjętymi standardami pracy z mikromacierzami. Uzyskane dane posłużyły jako referencja, do której odnoszono wszelkie obserwacje odnotowane w eksperymentach CSH, co uzasadnia zamieszczenie w pracy profilów ekspresji uzyskanych z wykorzystaniem standardowych eksperymentów mikromacierzowych. Oprócz tego opisane w tej części wyniki same w sobie posiadają dużą wartość biologiczną z uwagi na niedobór prac opisujących globalne zmiany ekspresji genów pod wpływem stresu w korzeniach tytoniu. W opisanym profilu ekspresji zwracano uwagę przede wszystkim na funkcje zmienionych grup genów, uwzględniając wszystkie podstawowe kategorie funkcjonalne, w ramach których obserwowano istotne zmiany. Liczne różnicujące GO (Tabela 19) również klasyfikowano w bardziej ogólne grupy, formułując podstawowe obszary zmian ekspresji. W Tabelach 20-22 zebrano najważniejsze dane dotyczące zmian ekspresji w korzeniach tytoniu poddanych działaniu wspomnianych stresów, wyszczególniając najistotniejsze geny podlegające zmianom ekspresji w poszczególnych grupach funkcjonalnych. Wyselekcjonowano kategorie GO i pojedyncze geny charakteryzujące:

- uniwersalną odpowiedź na stres abiotyczny- geny o podobnej ekspresji w przypadku obu stresów
- odpowiedź specyficzną dla stresu solnego
- odpowiedź specyficzną dla stresu kadmowego.

Tabela 19. Liczba różnicujących grup GO wyselekcjonowanych za pomocą analizy SEA dla korzeni tytoniu traktowanych NaCl i CdCl₂. „+” oznacza podwyższenie ekspresji, „-” obniżenie ekspresji.

		NaCl	CdCl ₂	Wspólne	Specyficzne dla NaCl	Specyficzne dla CdCl ₂	
Liczba różnicujących grup GO	+	126	172	73	53	99	
	-	22	39	14	8	25	
Podział z uwzględnieniem typów GO	Proces biologiczny	+	114	146	68	46	78
		-	2	4	2	0	2
	Komponent komórkowy	+	0	3	0	0	3
		-	3	13	1	2	12
	Funkcja molekularna	+	12	23	5	7	18
		-	17	22	11	6	11

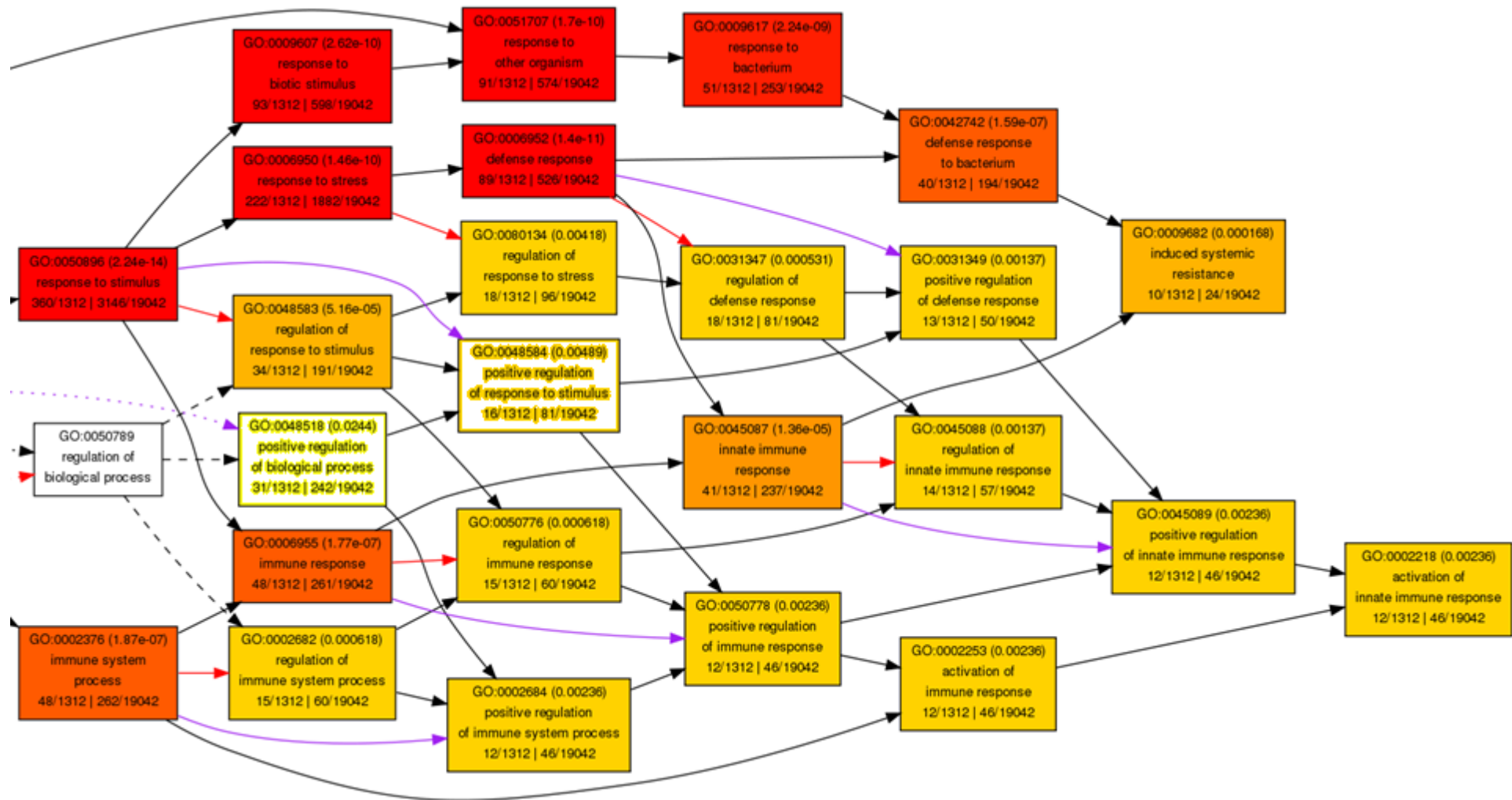
IV.2.1.1. Ogólny profil ekspresji genów w korzeniach tytoniu pod wpływem stresu abiotycznego

Wyselekcjonowano grupy GO ulegające podobnym zmianom w wyniku działania obu stresów, które potencjalnie mogą charakteryzować ogólny profil odpowiedzi na stres abiotyczny w korzeniach tytoniu. Łącznie wykryto 73 takie grupy wśród sond cechujących się wzrostem ekspresji oraz 14 w zestawie sond, których ekspresja ulegała zmniejszeniu. Na podstawie różnicujących grup GO wyselekcjonowano geny związane ze wspólną ścieżką odpowiedzi na stres solny i kadmowy w korzeniach tytoniu, które zamieszczono w Tabeli 20. Zestaw kategorii genów cechujących się podwyższeniem ekspresji obejmował w znacznej większości grupy GO należące do typu „proces biologiczny”. Aż 50 kategorii związanych było z odpowiedzią roślin na bodźce i działaniem układu odpornościowego. Podstawowe kategorie w tej grupie stanowiły: „proces układu odpornościowego” (GO:0002376) oraz „odpowiedź na bodziec” (GO:0050896). Pierwsza z nich zawierała szereg coraz bardziej szczegółowych kategorii dotyczących procesów odpornościowych, z których końcowe stanowiły „aktywacja odpowiedzi odpornościowej wrodzonej” (GO:0002218) oraz „indukcja odporności układowej” (GO:0009682), łącząca kaskady związane z odpornością oraz odpowiedzią na bodziec (Rycina 19). Z kolei w ramach rozlicznych kategorii GO dotyczących reakcji na bodziec możemy wyróżnić grupy związane z odpowiedzią na bodźce biotyczne („odpowiedź obronna na grzyby” (GO:0050832) oraz „- bakterie” (GO:0042742)), abiotyczne (UV (GO:0009411), intensywność światła (GO:0009642), ciepło (GO:0009408), brak wody (GO:0009414), związki chemiczne (leki (GO:0042493), nadtlenek wodoru (GO:0009411), chitynę (GO:0010200)), a także związane ze ścieżkami sygnałowymi hormonów (kwas

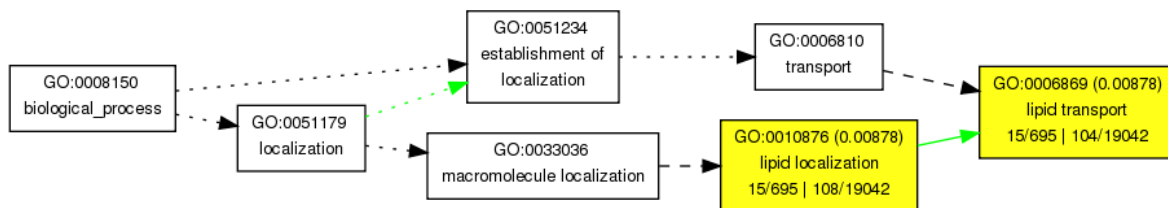
abscysynowy (GO:0009737), kwas salicylowy (GO:0071446), kwas jasmonowy (GO:0071395)). Powiązane z reakcjami odpornościowymi były GO opisujące regulację hormonalną, w tym kaskady sygnałów zależne od kwasu jasmonowego (GO:0009867) i salicylowego (GO:0009863)). Biorąc pod uwagę poszczególne sondy przyporządkowane do powyższych GO, część z nich została przypisana do genów, których zmiany ekspresji pod wpływem stresu abiotycznego zostały dość dokładnie zbadane i opisane. Należą do nich czynniki transkrypcyjne (ABF2, CBF1, DREB1A, ERF (2,4), HY5, RHL41, WRKY70) [121,119,120,107,113,114,106]. enzymy (GAPC1, F3H, MEKK1, CDPK1, UGT73C1) [130,134] i inne geny (ELIP2, HSP (101, 70, 18.2)) [131,132]. Obserwowano różnorodne zmiany dotyczące odpowiedzi na stres, m.in. indukcję genów związanych z sygnalizacją zależną od kwasu abscysynowego (ABA) (ABF2, HY5, ABR1, TCH2, MYB) oraz niezależną od ABA (CBF1, DREB1, ERF) [159-161]. Odnotowano również wzrost ekspresji genów świadczących o odpowiedzi związanej z działaniem kwasu jasmonowego, salicylowego i etylenu (WRKY70, ERF2, ERF4, JAZ3) [106, 107]. Wyróżnić można także grupę genów biorących udział w syntezie hormonów, jak OPR 1 i 2, LOX- reduktaz i lipooksygenazy wchodzących w skład ścieżki syntazy kwasu jasmonowego, o znanej roli w odpowiedzi na stres [135]. Spośród enzymów wyróżniają się geny kinaz powszechnie kojarzonych z odpowiedzią na różnorodne warunki stresowe: MAPK i CDPK [160,165]. Na uwagę zasługuje również zwiększenie ekspresji genów zwykle zaangażowanych w odpowiedź na stres biotyczny (MYB30, ELI3-2, EP3, NDR1) [152,153]. Istotnej zmianie ekspresji ulegały również kategorie związane z metabolizmem toksyn (3, w tym „katabolizm toksyn” GO:0009407), w ramach których wyselekcjonowano powszechnie znane geny: LEA oraz GST (Patrz Tabela 20), odpowiedzialne za tego rodzaju odpowiedź w różnorodnych warunkach stresowych [98,126,150]. W przypadku genów odpowiedzialnych za transport (m.in. „transport wielolekowy” (GO:0009863)), odnotowano wzrost ekspresji genu PDR9, którego indukcję pod wpływem kadmu, stresu solnego i innych czynników stresowych stwierdzono w korzeniach ryżu [129], a także MATE, którego zwiększoną ekspresję odnotowano u rzodkiewnika [151]. Ostatnim genem z tej grupy był MPR4 należący podobnie jak PDR9 do rodziny transporterów ABC. Wszystkie geny z tej grupy o podwyższonej ekspresji kodują transportery, które mogą być odpowiedzialne za usuwanie toksycznych substancji [129]. Cechująca się wzrostem ekspresji grupa GO:0034050 „programowana śmierć komórki indukowana przez symbiont” obejmuje geny ulegające ekspresji zarówno w wyniku stresu abiotycznego, jak i biotycznego jak EP3 (chitynaza) [151], ELI3-2 (dehydrogenaza kwasu cyjanonowego

[152] czy NDR1 [154]. Kolejną grupą genów tworzących uniwersalny profil odpowiedzi na stres są geny związane z sygnalizacją za pomocą wapnia („wiązanie jonów wapnia” (GO:0005509)). Sygnalizacja związana ze zmianami stężenia jonów wapnia w komórce w wyniku działania różnych czynników jest bardzo dokładnie opisanym zjawiskiem [155]. Wśród genów związanych z tego typu procesem, wykazujących wzrost ekspresji w obu rodzajach stresu znalazły się aneksyna, kalmodulina (CAM7) i inne białka wiążące wapń (CML37, TCH2). W ramach „aktywności syntazy alfa-humulenu i (-)-E-beta-kariofilenu” (GO:0080016, GO:0080017), odnotowano indukcję TPS21- genu syntetazy terpenów (beta-kardiofilenu i alfa-humulenu), słabo poznanego w kontekście stresu, choć wzrost ekspresji niektórych genów z tej rodziny w warunkach stresowych odnotowano u pomidora [156].

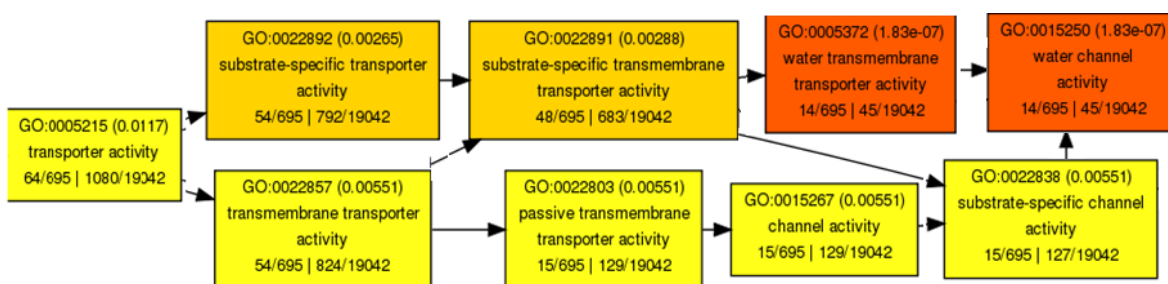
Geny, których ekspresja ulegała zahamowaniu w przypadku obu rodzajów stresu występowały w mniejszej liczbie i zostały podzielone na trzy grupy czynnościowe: transport tłuszczów, ściana komórkowa, aktywność kanałów wodnych. Wśród genów kodujących białka odpowiedzialne za transport lipidów (Rycina 20) („transport lipidów” GO:0006869, „wiązanie lipidów” (GO: 0008289)) wyróżniono białka błonowe o potwierdzonym spadku ekspresji w stresie (LTP) [95], oraz inne (DIR1, ELP). Białka LTP i ELP [157], są zaangażowane w proces rozluźnienia ściany komórkowej związanej z wzrostem roślin, co warto zaznaczyć w kontekście kolejnej grupy genów, związanych właśnie ze ścianą komórkową. Są to enzymy, takie jak celulaza (CEL5), a także komponenty ściany (5 białek arabinogalaktanowych). Obniżenie ekspresji powyższych genów sugeruje zahamowanie procesu rozluźnienia ściany związanego z inhibicją wzrostu wywołaną stresem, co jest zjawiskiem omawianym w literaturze z tej dziedziny [141,158]. Dziewięć kategorii GO związanych było z aktywnością transporterów (w tym „aktywność kanałów wodnych” GO:0015250) (Rycina 21). W ich skład wchodziły geny kodujące białka błonowe odpowiedzialne za przepływ wody przez tonoplast (TIP) i błonę komórkową (PIP)- akwaporyny („zakotwiczone w błonie” GO:0031225). Ich inhibicja w wyniku stresu abiotycznego została również odnotowana w innych eksperymentach i może być związana ze spowodowanymi warunkami stresowymi koniecznością utrzymania wody w komórce [143].



Rycina 19 Przykładowe wspólne kategorie GO związane z odpowiedzią na bodźce i odpornością w korzeniach tytoniu traktowanych NaCl oraz poddanych działaniu CdCl₂. Intensywność koloru oznacza większą istotność statystyczną. Kolorem białym oznaczono kategorie wykazujące istotną zmianę ekspresji tylko w jednym rodzaju stresu.



Rycina 20. Przykłady kategorii GO ulegających zmniejszonej ekspresji pod wpływem obu stresów: solnego i kadmowego: GO związane z transportem i lokalizacją lipidów.



Rycina 21. Przykłady kategorii GO ulegających zmniejszonej ekspresji pod wpływem obu stresów: solnego i kadmowego: GO związane z transportem przez błonowy

Tabela 20. Ważniejsze geny ulegające ekspresji różnicowej w korzeniach tytoniu w obu stresach (NaCl i CdCl₂). Kolor niebieski oznacza zwiększenie ekspresji, kolor czerwony- jej zmniejszenie.

Kategoria	Gen/grupa genów	Funkcja	Istotne GO	Doniesienia w literaturze
Odpowiedź na bodźce, procesy odpornościowe				
Czynniki transkrypcyjne	ABF2	Udział w odpowiedzi na ABA	GO:0009414	Indukcja w warunkach braku wody u rzodkiewnika [121]
	CBF1	Indukcja genów reakcji na stres chłodu	GO:0009414	Indukcja w warunkach chłodu u rzodkiewnika [120]
	ATAF1	Regulacja biosyntezy ABA [161]	GO:0009737	
	DREB1A	Regulator odpowiedzi na stres	GO:0009414	Indukcja w stresie suszy, solnym, zimna u rzodkiewnika [119]
	ERF2	Aktywacja genów indukowanych kwasem jasmonowym	GO:0009864,GO:0009682,GO:0002218,GO:0071395,GO:0009867,GO:0071395,GO:0010105,GO:0070298,	Indukcja w stresie zranienia u rzodkiewnika [107]
	ERF4	Represor genów indukowanych kwasem jasmonowym, etylenem, kwasem abscysynowym	GO:0009864,GO:0009682,GO:0002218,GO:0071395,GO:0009867,GO:0071395,GO:0010105,GO:0070298,GO:0016564	Indukcja w stresie solnym u rzodkiewnika [107]
	HY5	Wydłużenie hypocotyłu, odpowiedź na światło, hormonalna [112]	GO:0009411	Indukcja pod wpływem światła u rzodkiewnika [113]
	MYB30, MYB6	Odpowiedź na auksyny	GO:0034050	
	RHL41	Odpowiedź na światło	GO:0009411	Indukcja pod wpływem światła u rzodkiewnika [114]
	WRKY41, 50, 6	Odpowiedź na bodźce	GO:0006355	
	WRKY70	Aktywator genów obronnych zależnych od kwasy salicylowego, supresja sygnalizacji kwasu jasmonowego	GO:0009864,GO:0009682,GO:0002218,GO:0071446,GO:0071395,GO:0009867,GO:0009863,GO:0071395,GO:0016564	Indukcja w stresie osmotycznym (rzodkiewnik) [106]
	WRKY33	Odpowiedź na stres [124]	GO:0009414 GO:0052317	Indukcja w stresie cieplnym u rzodkiewnika [124]
Enzymy	ACLA-1,2,3	Syntetaza ATP cytrynianowa	GO:0009411	
	ATUBC2	Ligaza ubikwityny, aktywacja FLC (floral repressor), supresja	GO:0009411	

		kwitnienia[117]		
	CDPK1	Kinaza zależna od wapnia	GO:0009737	Udział w odpowiedzi na stres biotyczny i abiotyczny [165]
	GAPC1	Dehydrogenaza aldehydu 3-fosfoglicerynowego	GO:0042542	Indukcja pod wpływem stresu solnego w korzeniach rzodkiewnika [134]
	GST	S- Transferaza glutationowa, rola w detoksyfikacji [123]	GO:0009414,GO:0009407 GO:0004364,GO:0009407	Indukcja w warunkach stresu oksydacyjnego w tytoniu [122]
	F3H	3- Hydrolaza flawanononowa, biosynteza flawonoidów [110]	GO:0009411	Indukcja w stresie zranienia u bawełny [111]
	FMO1	Monooksygenaza zależna od flawin, odpowiedź odpornościowa [109]	GO:0002218, GO:0034050	
	MEKK1	Kinaza MAP, udział w szlaku MAPK	GO:0009414 GO:0042542	Indukcja pod wpływem kadmu i miedzi u rzodkiewnika [125]
	SIP3	Aktywność kinazy serynowej/treoninowej	GO:0009414	
	SGT1B	Fosfataza biorąca udział w odpowiedzi obronnej	GO:0071395	
	TT7	Monooksygenaza - 3'-flawonoidowa	GO:0009411 GO:0019825	
	UGT73C1	UDP-glukozylotransferaza, odpowiedź na stres [114]	GO:0009411	Indukcja w stresie oksydacyjnym, silnego światła, solnym, odwodnienia u rzodkiewnika [114]
Inne	ELIP2	Białko indukowane światłem	GO:0009411	Indukcja w stresie świetlnym u rzodkiewnika [116]
	BAG6	Związany z BCL2, wiązanie kalmoduliny	GO:0042542	
	HSP101	Białko szoku cieplnego	GO:0042542	Indukcja pod wpływem ciepła u rzodkiewnika[136]
	HSP70, HSP18.2	Białko szoku cieplnego, odpowiedź na stres	GO:0042542	Indukcja pod wpływem różnych rodzajów stresu, w tym cieplnego i oksydacyjnego u rzodkiewnika[131,132]
	JAZ3	Białko biorące udział w szlaku odpowiedzi związanej z kwasem	GO:0071395 GO:0009867	

		jasmonowym		
	SYP121	Syntaksyna	GO:0009737 GO:0071395 GO:0034050 GO:0009867 GO:0009863	
	LEA14	Białko odpowiedzi na stres	GO:0009414	Indukcja pod wpływem chłodu, stresu solnego, ABA u rzodkiewnika [126]
Geny odpowiedzi na stres biotyczny				
	ELI3-2	Dehydrogenaza alkoholu cynamonowego	GO:0034050	Indukcja w stresie biotycznym i abiotycznym [153]
	EP3	Chitynaza	GO:0034050	Indukcja w tytoniu pod wpływem stresu solnego i metali ciężkich [152]
	NDR1	Odpowiedź na patogeny	GO:0034050	
Transport				
Enzymy	PDR9	Atpaza zaangażowana w transport wielolekowy	GO:0042493	Indukcja przez metale ciężkie i niedobór tlenu w korzeniach ryżu [129]
Transportery	MRP4	Transporter leków	GO:0042493	
	MATE	Aktywność transportera leków	GO:0042493 GO:0006855	Indukcja w stresie solnym u rzodkiewnika [151]
Wiązanie wapnia				
Inne	ANNEXIN	Aneksyna, wiązanie jonów wapnia	GO:0005509	Indukcja pod wpływem stresu solnego, suszy u rzodkiewnika [138]
	CAM7	Kalmodulina, wiązanie wapnia	GO:0005509	
	CML37	Białko podobne do kalmoduliny	GO:0005509	Indukcja w stresie solnym i cieplnym [139]
	TCH2	Białko wiążące wapń	GO:0042542, GO:0005509	Indukcja pod wpływem różnych rodzajów stresu, w tym metali i soli u rzodkiewnika [137]
Biosynteza hormonów i innych substancji				
	LOX3	Lipoksygenaza	GO:0009695	Indukcja pod wpływem zranienia u rzodkiewnika [140]
	OPR1, OPR2	Reduktaza biorąca udział w biosyntezie kwasu jasmonowego	GO:0009695	Indukcja pod wpływem kadmu u rzodkiewnika [135]
	TPS21	Syntaza alfa-humulenu i (-)-E-beta-kariofilenu	GO:0080017 GO:0080016	
Transport tłuszczów				
Białka błonowe	DIR1	Potencjalny transfer tłuszczów	GO:0006869 GO:0008289	
	LTP	Białko transferu tłuszczów	GO:0006869 GO:0008289	Spadek ekspresji w korzeniach po wpływie NaCl u rzodkiewnika [141]

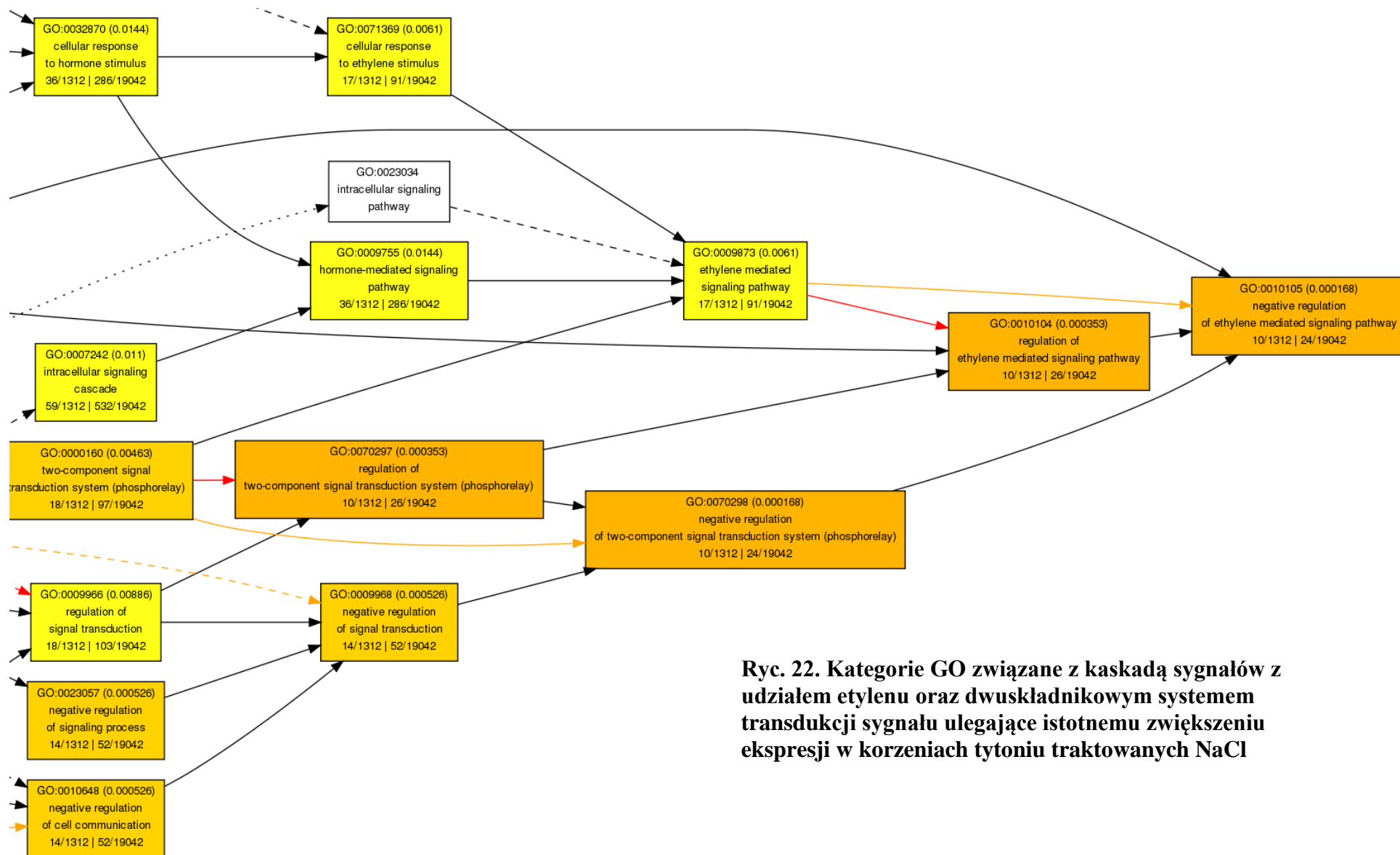
	ELP	Białko podobne do ekstensyny	GO:0006869 GO:0008289	
Ściana komórkowa				
Enzymy	BGAL1	Beta-galaktozydaza	GO:0016798	
	BGLU47	Beta-glukozydaza	GO:0016798	
	CEL5	Celulaza	GO:0016798	Spadek ekspresji w korzeniach rzodkiewnika pod wpływem NaCl [126]
	EXGT-A3,4, XTH9, XTR9	Transferaza endoksyloglukanowa	GO:0016798	Spadek ekspresji w korzeniach pomidora pod wpływem NaCl [142]
Białka strukturalne	AGP4, AGP9, AGP10, AGP18, AGP2,	Białka arabinogalaktanowe	GO:0031225	Spadek ekspresji w korzeniach pomidora pod wpływem NaCl [142]
Białka strukturalne	ATEXLA2, ATEXLB1	Geny podobne do ekspansyny	GO:0009828	ATEXLA2, ATEXLB1
	EXPA8	Ekspansyna	GO:0009828	EXPA8
Aktywność kanałów wodnych				
Białka błonowe	PIP1, PIP2	Integralne białka błony komórkowej, akwaporyny	GO:0015250	Spadek ekspresji w korzeniach rzodkiewnika pod wpływem NaCl [143]
	TIP	Integralne białka tonoplastu, akwaporyny	GO:0015250	Spadek ekspresji w korzeniach rzodkiewnika pod wpływem NaCl [143]

IV.2.1.2. Profil ekspresji genów w korzeniach tytoniu poddanych działaniu stresu solnego

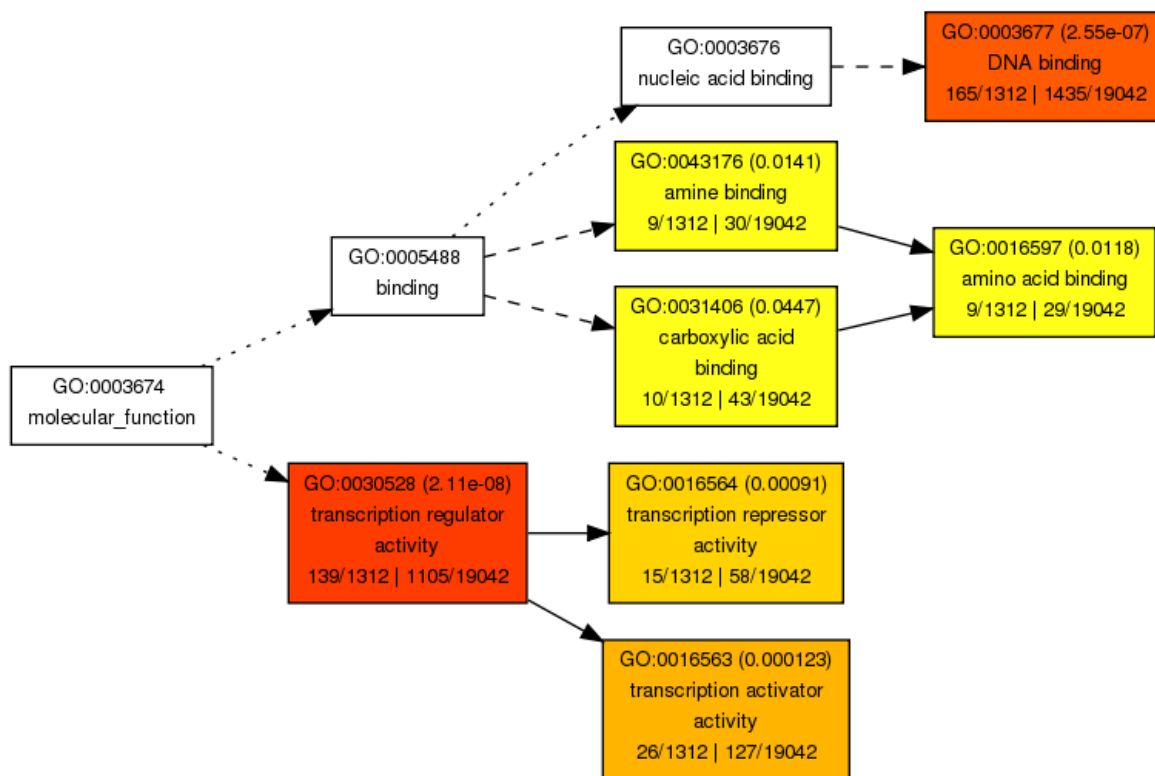
W przypadku oceny profilu ekspresji genów specyficznych dla korzeni tytoniu traktowanych NaCl wzięto pod uwagę 61 grup GO wykazujących istotne zmiany. Wyselekcjonowane grupy GO pozwoliły na wyróżnienie genów biorących udział w odpowiedzi korzeni tytoniu na stres solny. Zamieszczono je w Tabeli 21. W przypadku grup genów ulegających indukcji i należących do kategorii „proces biologiczny”, sześć kategorii związanych było z odpowiedzią na bodźce. Jedną z istotniejszych kategorii GO dla stresu solnego wydawała się „odpowiedź na brak wody”. Należące do niej geny ulegały ekspresji różnicowej w obu grupach (NaCl i CdCl₂). Dodatkowo, specyficzna dla stresu solnego była kategoria pokrewna „odpowiedź na działanie wody” (GO:0009415). Wśród ważniejszych genów z tej grupy specyficznie ulegających ekspresji tylko w przypadku stresu solnego, wyróżnić możemy m.in.: DREB2A, HIS4, ERD15 i ABA1. Na uwagę zasługuje tutaj gen DREB2A, który jest typowym genem stresu osmotycznego w ścieżce niezależnej od ABA [160]. Z drugiej strony, zmiana ekspresji wymienionych uprzednio genów świadczy o występowaniu ścieżki zależnej od ABA, co potwierdza również obecność genu ABA1, biorącego udział w jego syntezie [169]. Kategoria „odpowiedź komórkowa na etylen” (GO:0071369), stanowi jedną z czterech kategorii dotyczących regulacji hormonalnej i związana jest z kaskadą sygnałów z udziałem etylenu oraz jej negatywną regulacją (GO:0010105). Ta kaskada jest z kolei powiązana z dwuskładnikowym systemem transdukcji sygnału poprzez transfer grupy fosforanowej (ang. *phosphorelay*) (GO:0070298). Zależności te przedstawione są na Rycinie 22 i łącznie obejmują 14 istotnie zmienionych grup GO. Geny związane z odpowiedzią na etylen stwierdzono w przypadku obu traktowań (ERF2, ERF4), jednak w przypadku stresu solnego w tą ścieżkę sygnałową zaangażowana jest większa liczba genów i obserwuje się ekspresję różnicową kilku kategorii GO jej dotyczących. Wśród takich genów, potencjalnie specyficznych dla stresu solnego możemy wyróżnić EIN4 - przypuszczalny receptor etylenu [170], oraz EBF1- białko wiążące receptor EIN3 [170]. Wyselekcjonowane zostały również 2 kategorie genów związanych ze starzeniem (GO:0007568, GO:0010149). W ich skład wchodziły geny kodujące enzymy (ASP3, SAG101, ATG18a) oraz czynnik transkrypcyjny indukowany działaniem auksyn (ARF1). Istotnym zmianom uległa również kategoria „rozluźnienie ściany komórkowej typu roślinnego” (GO:0009828), w ramach której wzrostowi ekspresji ulegały geny ekspansyn. W ramach typu „funkcja molekularna” możemy wyróżnić dwa zasadnicze zespoły: wiązanie (amin, kwasów karboksylowych,

DNA, aminokwasów) oraz aktywność regulatorów (aktywatorów i represorów transkrypcji) (Rycina 23) , wśród których warto zwrócić uwagę na geny kodujące białko związane z tolerancją soli (STZ), czy regulator H⁺- ATPazy błon wakuoli i endosomu o wzroście ekspresji potwierdzonym w innych eksperymentach [142, 171].

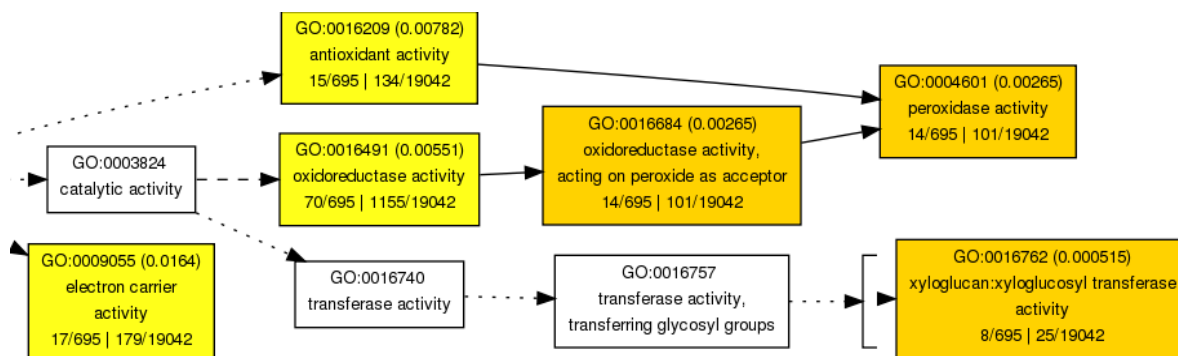
Z kolei GO ulegające istotnemu zahamowaniu, specyficzne dla traktowania NaCl, obejmowały 8 grup. Dwie dotyczyły "lokalizacji komórkowej" i wskazywały na procesy odbywające się w błonie i środowisku zewnątrzkomórkowym. Pozostałe 6 wchodziło w skład typu: „funkcja molekularna”: cztery tworzyły układ związany z aktywnością peroksydaz, pozostałe 2 dotyczyły aktywności przenośników elektronów i transferazy ksyloglukanowej (Rycina 24). Stwierdzono inhibicję transkrypcji kilku peroksydaz (PER12,64,42), katalazy (CAT3) i dysmutazy ponadtlenkowej miedź/cynk (CSD2). Peroksydazy, w tym wymienione enzymy, biorą udział m.in. w reakcjach obronnych roślin na powstawanie wolnych rodników a ich aktywność może być związana z odpowiedzią na stres. Tego typu aktywność wykazuje peroksydaza glutationowa, dla której stwierdzono wzrost ekspresji w obu traktowaniach i której sondy nie wchodziły w skład ulegającej zahamowaniu grupy GO.



Ryc. 22. Kategorie GO związane z kaskadą sygnałów z udziałem etylenu oraz dwuskładnikowym systemem transdukcji sygnału ulegające istotnemu zwiększeniu ekspresji w korzeniach tytoniu traktowanych NaCl



Rycina 23. Kategorie GO ulegające istotnemu zwiększeniu ekspresji w korzeniach tytoniu traktowanych NaCl, związane z wiązaniem substancji i aktywnością regulatorów transkrypcji.



Rycina 24. Przykładowe kategorie GO ulegające istotnemu zahamowaniu w korzeniach tytoniu traktowanych NaCl.

Tabela 21. Wybrane geny ulegające ekspresji różnicowej specyficznej dla korzeni tytoniu poddanych działaniu stresu solnego. Kolor niebieski oznacza indukcję ekspresji, kolor czerwony- jej zahamowanie.

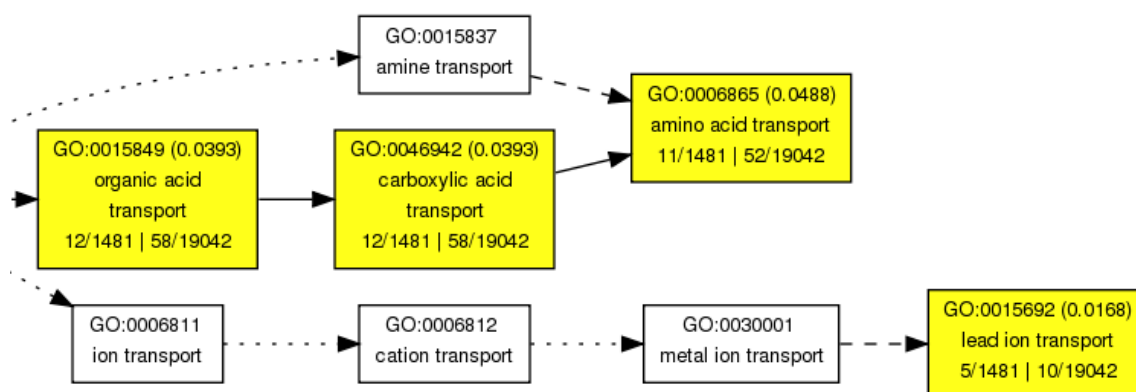
Kategoria	Gen/grupa genów	Funkcja	Istotne GO	Doniesienia w literaturze- stres solny
Niedobór wody				
Czynniki transkrypcyjne	DREB2A	Regulator odpowiedzi na stres	GO:0009411, GO:0042542, GO:0042542,	Indukcja pod wpływem stresu suszy i wysokiej temperatury[118] u rzodkiewnika
Enzymy	ABA1	epoksydaza zeaksantyny, udział w syntezie ABA	GO:0009414	
Inne	HIS4	Histon 4, odpowiedź na brak wody	GO:0009414	
	ERD15	Białko wczesnej odpowiedzi na brak wody, negatywna regulacja ABA [127]	GO:0009414	Indukcja w stresie niedoboru wody u rzodkiewnika [128]
Transdukcja sygnału za pośrednictwem etylenu				
Receptory	EIN4	Receptor etylenu [170]	GO:0009873	
Inne	EBF1	Białko wiążące receptor [170]	GO:0009873	
Starzenie				
Enzymy	AtATG18a	Transducyna/WD40 , udział w formowaniu autofagosomu	GO:0010149 ,GO:0007568	Indukcja w korzeniach rzodkiewnika pod wpływem stresu solnego [144]
	ASP3	Aminotransferaza asparaginianowa	GO:0010149, GO:0007568	
	SAG101	Hydrolaza acylowa	GO:0010149, GO:0007568	
inne	ARF1	Czynnik odpowiedzi na auksyny	GO:0010149, GO:0007568	
Regulacja transkrypcji				
Czynniki transkrypcyjne	AT-HSFB2B	Czynnik transkrypcyjny szoku cieplnego	GO:0016564	
	AHBP-1B	Czynnik transkrypcyjny z rodziny bZIP	GO:0016564	
	RAV2	Regulator H ⁺ - ATPazy błon wakuoli i endosomu	GO:0016564	
	STZ	Białko typu palec cynkowy związane z tolerancją soli	GO:0016564	Indukcja u rzodkiewnika pod wpływem stresu solnego[145]

Aktywność peroksydaz				
Enzymy	RCI3	Peroksydaza	GO:0004601 GO:0016209	Indukcja u rzodkiewnika pod wpływem chłodu [146]
	PER12,PER64, Per42	Peroksydazy	GO:0004601 GO:0016209 GO:0009055	
	CAT2,3	Katalaza	GO:0004601 GO:0016209	
	CSD2	Dysmutaza ponadtlenkowa miedź/cynk	GO:0004601 GO:0016209	Spadek ekspresji u rzodkiewnika w stresie solnym [147]

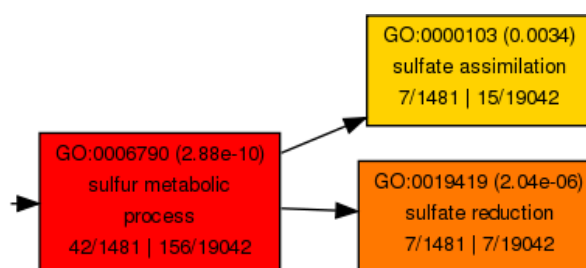
IV.2.1.3. Profil ekspresji genów w korzeniach tytoniu poddanych działaniu kadmu

Grupy genów o zmienionym poziomie ekspresji w korzeniach tytoniu traktowanych chlorkiem kadmu reprezentowane były przez 124 kategorie GO, z czego 99 ulegało indukcji. Spośród 78 należących do typu „proces biologiczny”, 10 dotyczyło odpowiedzi na bodźce. Na uwagę zasługują grupy genów związanych z odpowiedzią na chłód (GO:0009631), silne światło (GO:0009644), jony cynku (GO:0010043). W przypadku tej ostatniej grupy, zwraca uwagę wzrost ekspresji trzech genów biorących udział w odpowiedzi na działanie jonów metali. YSL i ZIF biorą udział w syntezie i transporcie nikotianaminy, chelatującej jony m.in. cynku i kadmu. Białko NRAMP3 jest transporterem jonów metali (w tym kadmu) zlokalizowanym w tonoplaście [173]. Pozostałe różnicujące kategorie GO związane z odpowiedzią na bodźce nie zawierały genów specyficznych dla odpowiedzi na kadm. Wśród innych GO, zdecydowaną większość (58) stanowiły kategorie dotyczące metabolizmu i biosyntezy różnorodnych substancji. Na uwagę zasługuje grupa genów biorących udział w syntezie kwasu jasmonowego (GO:0009694) na różnych etapach: lipooksygenazy, syntetazy JA oraz syntazy i cyklazy tlenu allenowego [173,174]. Regulacja odpowiedzi na działanie jonów metali ciężkich za pomocą kwasu jasmonowego była potwierdzona u kilku roślin [175]. Innym hormonem, w przypadku którego obserwujemy widoczną indukcję metabolizmu, jest etylen (GO:0009693). W tym przypadku również znane są doniesienia literaturowe potwierdzające indukcję syntezy hormonu w wyniku działania jonów metali, w tym kadmu [176]. Procesy z udziałem etylenu stwierdzono dla obu traktowań, jednak w przypadku stosowania kadmu wyraźnie widoczna jest indukcja genów odpowiedzialnych za jego syntezę: syntazy i oksydazy ACC, o potwierdzonym wzroście ekspresji przez innych autorów [176], oraz syntetazy adenozylometioniny (1 i 2). Kolejnym związkiem, którego biosynteza okazała się znaczącą kategorią GO był acetylokoenzym A (GO:0006084). W tym przypadku oprócz indukcji

genów kodujących enzymy biorące udział w jego syntezie (PANK1, MVA) stwierdzono również wzrost ekspresji genu kodującego enzym MVD1, uczestniczący w syntezie terpenów, w którą zaangażowany jest acetylokoenzym A [177]. Kolejną godną uwagi kategorią GO związaną z metabolizmem jest „Metabolizm L-seryny” (GO:0006563). Odnotowano indukcję genów kodujących enzymy biorące udział w biosyntezie tego aminokwasu: (PDGH, SHM, PSAT). Z powyższym zjawiskiem może być związany również wzrost ekspresji trzech grup GO związanych z asymilacją i redukcją siarczanów (Rycina 26) i genów odpowiedzialnych za asymilację siarki: reduktazy APS i siarczynowej. Asymilacja siarki jest zjawiskiem często obserwowanym w przypadku stresu kadmowego i solnego, a reduktaza APS jest enzymem charakterystycznym dla tego zjawiska. Zarówno seryna, jak i asymilowane siarczany, a także wspomniany wcześniej acetylokoenzym A biorą udział w syntezie glutationu [135, 178]. Wspomniane grupy GO nie wykazywały różnicowania w przypadku działania stresu solnego na korzenie tytoniu, jednak widoczna u roślin traktowanych NaCl indukcja niektórych pojedynczych sond (np. reduktazy APS) i zwiększona indukcja genów syntezy glutationu świadczy o obecności tego zjawiska również w przypadku stresu solnego, być może w mniejszym nasileniu. W tym kontekście warto wspomnieć również o ekspresji licznych genów z rodziny cytochromu P450, które również biorą udział w usuwaniu wolnych rodników. Ich zwiększoną ekspresję pod wpływem kadmu wykazano już w korzeniach żeńszenia i ryżu [149,179]. Cztery kategorie GO związane były z transportem substancji: kwasów karboksylowych, aminokwasów (gen permeazy aminokwasowej) oraz jonów ołowiu (wspomniany gen NRAMP3 oraz gen transportera PDR12) (Rycina 25). W ramach „lokalizacji komórkowej” znaczący wzrost ekspresji wykazały sondy związane z procesami zachodzącymi w retikulum endoplazmatycznym (3 kategorie). Z kolei „funkcje molekularne” opisywało 17 kategorii GO, z czego 14 związane było z aktywnością enzymów: reduktazy adenylilosiarczanowej (GO:0009973), adenozylotransferazy metioninowej (GO:0004478), transferazy grup acylowych innych niż aminoacyłowe (GO:01016747), glukozylotransferazy kwercetyny (GO:0080043), hydrolazy epoksydowej (GO:0004301), liazy węgiel-tlen (GO:0016838), cyklazy (GO:0009975) (Rycina 27). Pozostałe kategorie związane były z aktywnością transporterów błonowych (GO:0015238, GO:0022804) oraz wiązaniem tlenu (GO:0019825).



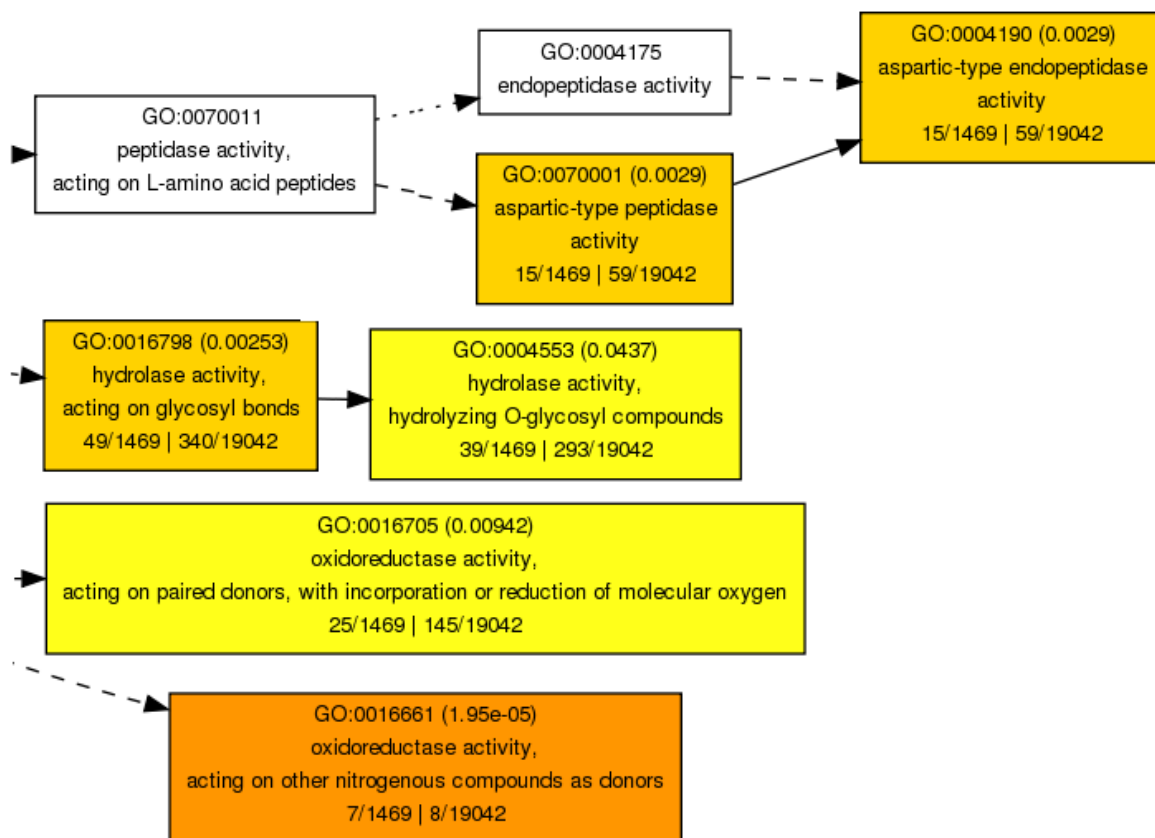
Rycina 25. Kategorie GO związane z transportem, których ekspresja uległa zwiększeniu w korzeniach tytoniu poddanych działaniu chlorku kadmu.



Rycina 26. Kategorie GO związane z metabolizmem siarki, których ekspresja uległa zwiększeniu w korzeniach tytoniu poddanych działaniu chlorku kadmu.

Zmniejszonej ekspresji uległo wiele szereg grup GO z zakresu „lokalizacja komórkowa”. Były to m.in. geny kodujące białka tworzące cytoszkielet: tubulinę, wilinę i aktynę (kompleks tubuliny (GO:0045298)), a także związane z wakuolą (błona wakuoli (GO:0009705)). To organelum odgrywa bardzo istotną rolę w detoksyfikacji komórek roślinnych poddanych działaniu kadmu. W transporcie kadmu lub jego chelatów biorą udział liczne transportery błony wakuolarniej, jak CAX [181]. W przypadku genów ulegających zahamowaniu ekspresji w niniejszym eksperymencie mamy do czynienia z transporterami usuwającymi do wakuoli inne jony, jak CAX3 (wapń) [181], VHA-A2 (wodór, stwierdzone eksperymentalnie hamowanie aktywności przez jony kadmu) oraz aktywująca transporter kinaza SOS2 [182], KCO6 (potas). Warto dodać, że inhibicji w warunkach stresu kadmowego ulegały również inne transportery np. anionów (siarczanów –SULTR, AST, azotanów-NRT, fosforanów- ATPT) lub kationów (potasu-ATKEA4, żelaza-IRT3). Dwie kategorie GO wyselekcjonowano w ramach typu „proces biologiczny”. Były to: „organizacja cytoszkieletu” (GO:0007010) oraz “wygaszanie niefotochemiczne” (GO:0010196). Z kolei w ramach „funkcji molekularnej” istotnemu zahamowaniu uległy grupy genów związanych z transportem („aktywność transporterów

śródbłonowych anionów nieorganicznych” (GO:0015103), „aktywność transporterów śródbłonowych kationów” (GO:0008324)), aktywnością enzymów (5 kategorii, Rycina 27), wiązaniem tetrapiroli (GO:0046906) i chlorofilu (GO:0016168) oraz „strukturalnymi składnikami cytoszkieletu” (GO:0005200). W Tabeli 22 zamieszczono geny związane z odpowiedzią na działanie kadmu, wyselekcjonowane na podstawie wyników GO.



Rycina 27. Kategorie GO dotyczące aktywności enzymów, których ekspresja uległa zahamowaniu w korzeniach tytoniu poddanych działaniu chlorku kadmu

Tabela 22. Wybrane geny ulegające ekspresji różnicowej w korzeniach tytoniu poddanych działaniu kadmu. Kolor niebieski oznacza zwiększenie ekspresji, kolor czerwony- jej zmniejszenie.

Kategoria	Gen/grupa genów	Funkcja	Istotne GO	Doniesienia w literaturze-stres metali ciężkich
Odpowiedź na jony metali				
Transportery	ZIF1	Białko indukowane jonami cynku	GO:0010043	Indukcja jonami cynku [173]
	NRAMP2, 3	Transporter błonowy jonów metali, w tym kadmu	GO:0010043	
	YSL2	Transporter metali-nikotianaminy	GO:0010043	
Metabolizm kwasu jasmonowego				
Enzymy	JAR1	Syntetaza kwasu jasmonowego	GO:0009694	
	LOX1, LOX5	lipoksygenaza	GO:0031407,GO:0009694	Indukcja jonami kadmu [174]
	AOS	Syntaza tlenu allenowego	GO:0031407,GO:0009694,GO:0019825	Indukcja jonami ołowiu [175]
	AOC3	Cyklaza tlenu allenowego	GO:0031407	
Metabolizm etylenu				
	EFE	Oksydaza ACC	GO:0009693	
	ACS6	Syntaza ACC	GO:0009693	Indukcja jonami kadmu [176]
	SAM1, SAM2	Syntetaza adenozylometioniny	GO:0009693 GO:0006556 GO:0004478	
Metabolizm acetylokoenzymu A				
Enzymy	MVD1	dekarboksylaza difosfomewalonianowa	GO:0006084	
	MVA1	Syntetaza hydroksymetyloglutarylo-CoA	GO:0006084	
	ATPANK1	Kinaza pantotenianowa	GO:0006084	
	PPC3	karboksylaza fosfoenolopirogronianowa	GO:0006084	Indukcja w korzeniach pomidora pod wpływem kadmu [148]
Metabolizm L-seryny				
Enzymy	PGDH	Dehydrogenaza fosfoglicerynianowa	GO:0006563	
	SHM6, SHM7	Hydroksymetylotransferaza serynowa	GO:0006563	Indukcja w komórkach rzodkiewnika pod wpływem kadmu [135]
	PSAT	Aminotransferaza fosfoserynowa	GO:0006563	
Asymilacja siarczanów				
Enzymy	SIR	Reduktaza siarczynowa	GO:0019419	
	APR1, APR3	Reduktaza APS	GO:0019419, GO:0009973	Indukcja w korzeniach rzodkiewnika pod wpływem kadmu [178]
Wiązanie tlenu				
Inne	CYP (wiele genów)	Geny cytochromu P450	GO:0019825	Indukcja pod wpływem kadmu w ryżu [149]

Cytoszkieleł				
Białka strukturalne	TUA, TUB	Tubulina	GO:0045298, GO:0007010	
	VLN	Wilina	GO:0007010	
	ACT	Aktyna	GO:0007010	
Procesy w wakuoli				
Enzymy	SOS2	Kinaza białkowa serynowo-treoninowa	GO:0000325, GO:0009705, GO:0044437	
Transportery	CAX3	Białko błonowe biorące udział w wymianie protonów i wapnia	GO:0000325, GO:0009705, GO:0044437	
	VHA-A2	Wakuolarna ATPaza protonowa	GO:0000325, GO:0009705, GO:0044437	
	KCO6	Kanał potasowy	GO:0000325, GO:0009705, GO:0044437	
Transport				
Transportery	SULTR3;2, AST	Transporter siarczanów	GO:0015103	
	NRT	Transporter azotanów	GO:0015103	
	ATPT2	Transporter fosforanów	GO:0015103, GO:0008324	
	ATKEA4	Antyporter wypływu jonów potasu	GO:0008324	
	KCO6	Kanał potasowy	GO:0008324	
	IRT3	Transporter błonowy jonów żelaza	GO:0008324	

IV.3. MIKROMACIERZOWA HYBRYDYZACJA MIĘDZYGATUNKOWA- BADANIA PORÓWNAWCZE PROFILÓW EKSPRESJI GENÓW

W kolejnych etapach analizy porównywano wyniki uzyskane na poszczególnych mikromacierzach CSH, biorąc pod uwagę aspekty wyszczególnione jako cele niniejszej pracy. Punkt odniesienia stanowiły każdorazowo wyniki uzyskane podczas eksperymentów SSH z zastosowaniem mikromacierzy TOB (por. IV.2). Na podstawie podobieństwa sekwencji (por. IV.1.2) zestawiono dane dla odpowiadających sobie sond z poszczególnych mikromacierzy CSH i mikromacierzy TOB i dokonano porównań, których wyniki zebrano w przedstawionych poniżej tabelach (Tabela 23-25). Dane z tabel zostały omówione w kolejnych rozdziałach, z odsyłaczami do numerów tabel i kolumn dotyczących analizowanego zagadnienia. W Tabeli 23 zebrano ogólne dane dotyczące analizy różnicowej, w Tabeli 24 dane dotyczące sond różnicujących („top600”) (por. IV.1.6), a w Tabeli 25 wyniki uzyskane podczas badań korelacji profilów ekspresji z danymi referencyjnymi.

Oprócz porównania na poziomie pojedynczych sond, wykonano również analizy funkcjonalne z wykorzystaniem kategorii GO. Uzyskane w ten sposób dane porównano i przedstawiono graficznie dzięki programowi SEACOMPARE (Rycina 28-31). Analizę funkcjonalną wykonywano dwoma sposobami:

- z wykorzystaniem danych GO dla mikromacierzy TOB dostępnych w programie agriGO: po uzyskaniu list genów różnicujących dla mikromacierzy CSH, selekcjonowano ich odpowiedniki w tytoniu (na podstawie podobieństwa $<0,001$) i dla zbioru odpowiadających im sond wykonywano analizę SEA dla tytoniu

- z wykorzystaniem danych GO dla gatunków reprezentowanych na mikromacierzach TOM1, TOM2 i POT znajdujących się w bazach danych (TOM1, TOM2: Tomato Functional Genomics Database (<http://ted.bti.cornell.edu/>), POT:TIGR Solanaceae Genomics Resource (http://jcvl.org/potato/sol_ma_microarrays.html)), co umożliwiło stworzenie profilu ekspresji pod kątem gatunków reprezentowanych na poszczególnych mikromacierzach.

Statystyczne podsumowanie analiz GO dla mikromacierzy CSH zamieszczono w Tabelach 26 i 27.

Tabela 23. Zestawienie wyników eksperymentów mikromacierzowych SSH i CSH analizowanych różnymi metodami. Podstawowe wyniki analizy różnicowej.

A)		NaCl	0	1	2	3	4	5	6	7	8	9	10	11	12
			ref	brak filtracji				filtracja HOM				filtracja MORF			
<i>Mikromacierz</i>		TOB	TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt	
<i>Liczba przeanalizowanych sond</i>		43730	8951	12020	10426	10426	8054	5944	8668	8668	-	-	7047	7047	
<i>Liczba sond o $p < 0,05$</i>		2478 (9,4%)	530 (6,0%)	91 (0,7%)	3104 (29,8%)	991 (9,5%)	481 (5,9%)	80 (1,3%)	2797 (32,2%)	942 (10,9%)	-	-	3377 (47,9%)	934 (13,3%)	
<i>Sondy wskazujące na co najmniej dwukrotną zmianę ekspresji</i>	<i>Ogółem</i>	1234	60	229	67	64	53	117	56	54	-	-	62	60	
	<i>W tym o ekspresji podwyższonej</i>	75,0%	86,7%	55,9%	85,1%	78,1%	86,8%	58,9%	89,2%	83,3%	-	-	88,7%	78%	

B)		CdCl₂	0	1	2	3	4	5	6	7	8	9	10	11	12
			ref	brak filtracji				filtracja HOM				filtracja MORF			
<i>Mikromacierz</i>		TOB	TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt	
<i>Liczba przeanalizowanych sond</i>		43730	8951	12020	10426	10426	8054	5944	8668	8668	-	-	7047	7047	
<i>Liczba sond o $p < 0,05$</i>		5917 (13,5%)	627 (7,0%)	100 (0,8%)	2586 (24,8%)	733 (7,0%)	600 (7,5%)	100 (1,7%)	2304 (26,5%)	704 (8,1%)	-	-	2984 (42,3%)	745 (10,5%)	
<i>Sondy wskazujące na co najmniej dwukrotną zmianę ekspresji</i>	<i>Ogółem</i>	791	76	109	48	47	72	63	43	41	-	-	38	44	
	<i>W tym o ekspresji podwyższonej</i>	77,5%	94,7%	57,7%	95,8%	82,9%	95,8%	63,5%	100%	87,8%	-	-	100%	79,6%	

Tabela 24. Zestawienie wyników eksperymentów mikromacierzowych SSH i CSH analizowanych różnymi metodami. Porównanie zgodności uzyskanych list sond różnicujących CSH („top600”) z listami sond oznaczonych jako różnicujące na mikromacierzy TOB

A)	<u>NaCl</u>	1	2	3	4	5	6	7	8	9	10	11	12
		brak filtracji				filtracja HOM				filtracja MORF			
<i>Mikromacierz</i>		TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt
<i>Liczba sond zmapowanych na mikromacierz TOB</i>		529	282	501	514	600	592	600	600	-	-	512	533
<i>Udział sond o podwyższonej ekspresji</i>		58%	73%	69%	63%	55%	62%	62%	61%	-	-	64%	61%
<i>Udział sond oznaczonych jako różnicowe na mikromacierzy TOB</i>	<i>Ogółem</i>	33%	24%	38%	42%	30%	20%	31%	41%	-	-	36%	45%
	<i>W tym o zgodnym kierunku ekspresji</i>	79%	73%	94%	98%	77%	73%	92%	98%	-	-	95%	99%
<i>Średnia wartość p</i>		0,028	0,113	0,003	0,001	0,031	0,166	0,002	0,001	-	-	0,001	0,001

B)	<u>CdCl₂</u>	1	2	3	4	5	6	7	8	9	10	11	12
		brak filtracji				filtracja HOM				filtracja MORF			
<i>Mikromacierz</i>		TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt
<i>Liczba sond zmapowanych na mikromacierz TOB</i>		535	295	527	529	600	591	600	600	-	-	525	539
<i>Udział sond o podwyższonej ekspresji</i>		66%	67%	73%	63%	65%	56%	67%	67%	-	-	68%	62%
<i>Udział sond oznaczonych jako różnicowe na mikromacierzy TOB</i>	<i>Ogółem</i>	44%	35%	42%	58%	40%	27%	42%	55%	-	-	41%	58%
	<i>W tym o zgodnym kierunku ekspresji</i>	89%	76%	97%	98%	89%	73%	96%	98%	-	-	98%	98%
<i>Średnia wartość p</i>		0,023	0,091	0,0028	0,004	0,023	0,099	0,027	0,005	-	-	0,003	0,004

Tabela 25. Korelacja wartości logFC uzyskanych dla odpowiadających sobie sond na mikromacierzach CSH i SSH. Dotyczy sond posiadających odpowiedniki na mikromacierzy TOB. * wartość p określa istotność statystyczną korelacji.

A)		<u>NaCl</u>		1	2	3	4	5	6	7	8	9	10	11	12
				brak filtracji				filtracja HOM				filtracja MORF			
Mikromacierz		TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt		
<i>Dla wszystkich sond na mikromacierzy CSH</i>	<i>Wsp. kor.*</i>	0,21	0,12	0,35	0,42	0,21	0,12	0,35	0,43	-	-	0,29	0,39		
	<i>Wartość p[#]</i>	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	-	-	p<2,2e-16	p<2,2e-16	
<i>Dla sond top600</i>	<i>Wsp. kor.*</i>	0,41	0,29	0,53	0,63	0,4	0,22	0,51	0,59	-	-	0,57	0,65		
	<i>Wartość p[#]</i>	p<2,2e-16	p=7,648e-07	p<2,2e-16	p<2,2e-16	p<2,2e-16	p=6,834e-08	p<2,2e-16	p<2,2e-17	-	-	p<2,2e-16	p<2,2e-16		
<i>Dla sond top600, których odpowiedniki są różnicujące na TOB</i>	<i>Wsp. kor.*</i>	0,5	0,39	0,79	0,86	0,52	0,34	0,78	0,84	-	-	0,81	0,86		
	<i>Wartość p[#]</i>	p<2,2e-16	p=0,000842	p<2,2e-16	p<2,2e-16	p<2,2e-16	p=0,000169	p<2,2e-16	p<2,2e-16	-	-	p<2,2e-16	p<2,2e-16		

B)		<u>CdCl₂</u>		1	2	3	4	5	6	7	8	9	10	11	12
				brak filtracji				filtracja HOM				filtracja MORF			
Mikromacierz		TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt		
<i>Dla wszystkich sond na mikromacierzy CSH</i>	<i>Wsp. kor.*</i>	0,34	0,19	0,42	0,55	0,34	0,16	0,42	0,5	-	-	0,51	0,46		
	<i>Wartość p[#]</i>	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	-	-	p<2,2e-16	p<2,2e-16	
<i>Dla sond top600</i>	<i>Wsp. kor.*</i>	0,55	0,41	0,61	0,72	0,55	0,34	0,61	0,7	-	-	0,64	0,71		
	<i>Wartość p[#]</i>	p<2,2e-16	p=5,862e-14	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,2e-16	-	-	p<2,2e-16	p<2,2e-16		
<i>Dla sond top600, których odpowiedniki są różnicujące na TOB</i>	<i>Wsp. kor.*</i>	0,64	0,54	0,76	0,85	0,65	0,44	0,77	0,84	-	-	0,78	0,85		
	<i>Wartość p[#]</i>	p= 8,988e-13	p= 2,159e-09	p<2,2e-16	p<2,2e-16	p<2,2e-16	p<2,266e-09	p<2,2e-16	p<2,2e-16	-	-	p<2,2e-16	p<2,2e-16		

Tabela 26. Zgodność różnicujących kategorii GO uzyskanych dla mikromacierzy CSH i SSH. Analiza GO wykonana metodą wykorzystującą adnotacje funkcjonalne dla TOB

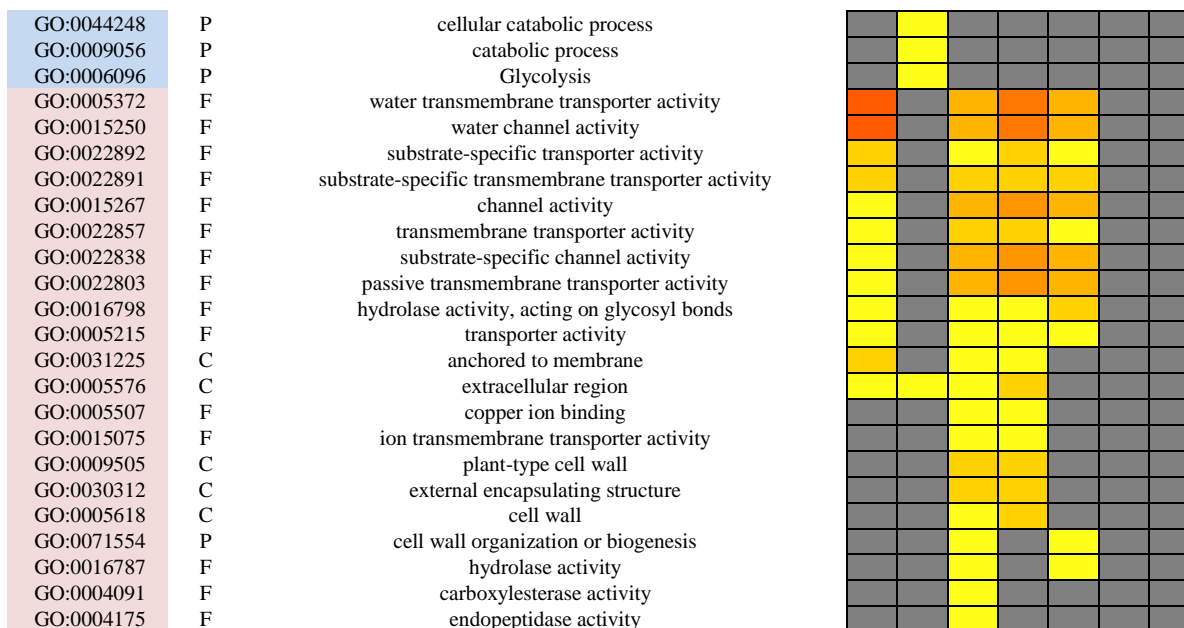
A)		<u>NaCl</u>	1	2	3	4	5	6	7	8	9	10	11	12
			brak filtracji				filtracja HOM				filtracja MORF			
Mikromacierz			TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt
<i>l. różnicujących grup GO</i>	+		23	0	30	48	7	0	35	52	-	-	18	27
	-		0	0	1	21	0	2	1	17	-	-	0	12
<i>l. różnicujących GO zbieżnych z TOB (% wykrytych)</i>	+		23 (18%)	0 (0%)	22 (17%)	38 (30%)	7 (5%)	0 (0%)	27 (21%)	40 (32%)	-	-	10 (8%)	19 (15%)
	-		0 (0%)	0 (0%)	1 (5%)	12 (55%)	0 (0%)	0 (0%)	0(0%)	12 (55%)	-	-	0 (0%)	10 (45%)
<i>l. różnicujących GO niezgodnych z TOB</i>	+		0	0	8	10	0	0	8	12	-	-	8	12
	-		0	0	0	9	0	2	1	5	-	-	0	2
B)		<u>CdCl₂</u>	1	2	3	4	5	6	7	8	9	10	11	12
			brak filtracji				filtracja HOM				filtracja MORF			
Mikromacierz			TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt
<i>l. różnicujących grup GO</i>	+		12	1	35	65	14	0	32	67	-	-	13	46
	-		2	0	0	19	3	0	0	21	-	-	0	27
<i>l. różnicujących GO zbieżnych z TOB (% wykrytych)</i>	+		11 (6%)	1 (2%)	32 (19%)	47 (27%)	13 (8%)	0 (0%)	25 (15%)	50 (29%)	-	-	11 (6%)	34 (19%)
	-		0 (0%)	0 (0%)	0 (0%)	9 (23%)	0 (0%)	0 (0%)	0 (0%)	13 (34%)	-	-	0 (0%)	14 (36%)
<i>l. różnicujących GO niezgodnych z TOB</i>	+		1	0	3	18	1	0	7	17	-	-	2	12
	-		2	0	0	10	3	0	0	8	-	-	0	13

Tabela 27. Zgodność różnicujących kategorii GO uzyskanych dla mikromacierzy CSH i SSH. Analiza GO wykonana metodą wykorzystującą adnotacje funkcjonalne dla gatunku reprezentowanego na mikromacierzy (TOM1, TOM2- pomidor, POT-ziemniak).

A)		<u>NaCl</u>	1	2	3	4	5	6	7	8	9	10	11	12
			brak filtracji				filtracja HOM				filtracja MORF			
Mikromacierz			TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt
<i>l. różnicujących grup GO</i>	+		52	5	42	53	44	19	47	65	-	-	30	39
	-		2	0	3	13	1	10	3	16	-	-	0	10
<i>l. różnicujących GO zbieżnych z TOB/ % wykrytych</i>	+		30 (24%)	0 (0%)	20 (18%)	26 (21%)	30 (24%)	0 (0%)	20 (18%)	27 (21%)	-	-	1 (0%)	9 (7%)
	-		0 (0%)	0 (0%)	3 (8%)	3 (8%)	0 (0%)	0 (0%)	3 (8%)	1 (5%)	-	-	0 (0%)	3 (8%)
<i>l. różnicujących GO niezgodnych z TOB</i>	+		22	5	22	27	14	19	27	38	-	-	29	30
	-		2	0	0	10	1	10	0	15	-	-	0	7
B)		<u>CdCl₂</u>	1	2	3	4	5	6	7	8	9	10	11	12
			brak filtracji				filtracja HOM				filtracja MORF			
Mikromacierz			TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt	TOM1	TOM2	POT-1powt	POT-2powt
<i>l. różnicujących grup GO</i>	+		60	2	48	85	52	2	53	81	-	-	44	69
	-		0	0	1	23	0	0	1	20	-	-	1	20
<i>l. różnicujących GO zbieżnych z TOB/ % wykrytych</i>	+		27 (16%)	2 (1%)	28 (16%)	47 (27%)	25 (15%)	2 (1%)	31 (18%)	46 (26%)	-	-	32 (18%)	42 (24%)
	-		0 (0%)	0 (0%)	0 (0%)	6 (15%)	0 (0%)	0 (0%)	0 (0%)	5 (13%)	-	-	0 (0%)	5 (13%)
<i>l. różnicujących GO niezgodnych z TOB</i>	+		33	0	20	38	27	0	22	38	-	-	12	27
	-		0	0	1	17	0	0	1	15	-	-	1	15

GO referencyjne: TOB, stres:NaCl

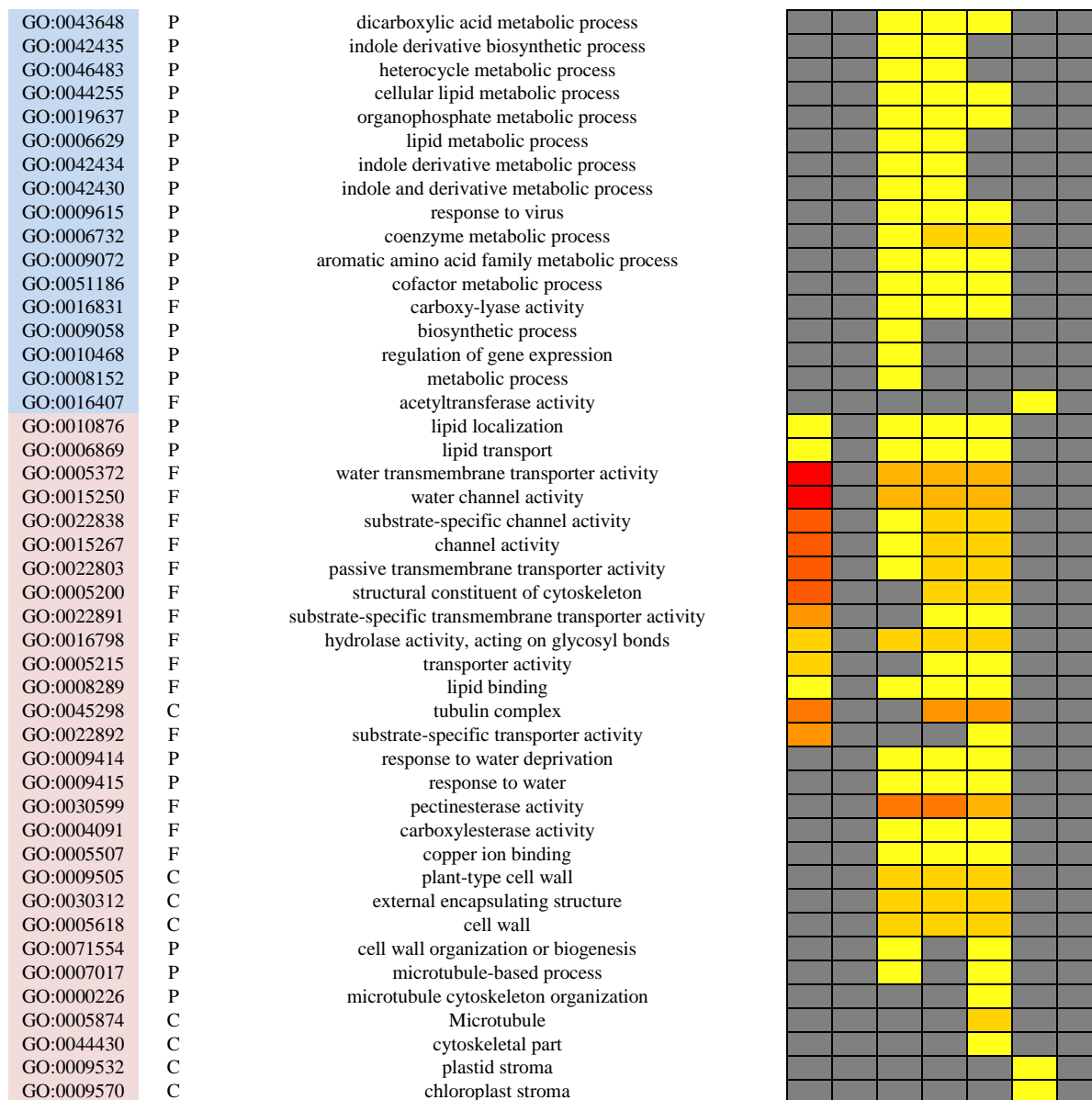
Nr GO	Typ	Opis	TOB	POT 1P	POT 2P	POT H 2P	POT SC 2P	TOM1	TOM2
GO:0010200	P	response to chitin	■		■				
GO:0050896	P	response to stimulus	■						
GO:0042221	P	response to chemical stimulus	■						
GO:0006952	P	defense response	■		■	■	■		
GO:0006950	P	response to stress	■		■				
GO:0009415	P	response to water	■						
GO:0009611	P	response to wounding	■						
GO:0009605	P	response to external stimulus	■						
GO:0009414	P	response to water deprivation	■						
GO:0003700	F	transcription factor activity	■		■	■			
GO:0030528	F	transcription regulator activity	■		■	■			
GO:0003677	F	DNA binding	■		■	■			
GO:0051707	P	response to other organism	■						
GO:0009743	P	response to carbohydrate stimulus	■		■				
GO:0009607	P	response to biotic stimulus	■		■	■			
GO:0009617	P	response to bacterium	■		■				
GO:0042742	P	defense response to bacterium	■		■				
GO:0051704	P	multi-organism process	■						
GO:0010556	P	regulation of macromolecule biosynthetic process	■						
GO:0050832	P	defense response to fungus	■						
GO:0009889	P	regulation of biosynthetic process	■						
GO:0031326	P	regulation of cellular biosynthetic process	■						
GO:0045449	P	regulation of transcription	■						
GO:0006355	P	regulation of transcription, DNA-dependent	■						
GO:0009620	P	response to fungus	■						
GO:0051171	P	regulation of nitrogen compound metabolic process	■						
GO:0051252	P	regulation of RNA metabolic process	■						
GO:0080090	P	regulation of primary metabolic process	■						
		regulation of nucleobase, nucleoside, nucleotide and nucleic acid	■						
		metabolic process	■						
GO:0019219	P	metabolic process	■						
GO:0009404	P	toxin metabolic process	■		■	■	■		
GO:0009407	P	toxin catabolic process	■		■	■	■		
GO:0006350	P	Transcription	■						
GO:0031323	P	regulation of cellular metabolic process	■						
GO:0010468	P	regulation of gene expression	■						
GO:0019748	P	secondary metabolic process	■						
GO:0008219	P	cell death	■						
GO:0016265	P	Death	■						
GO:0009408	P	response to heat	■						
GO:0012501	P	programmed cell death	■						
GO:0004364	F	glutathione transferase activity	■		■	■	■		
GO:0010033	P	response to organic substance	■						
GO:0009719	P	response to endogenous stimulus	■						
GO:0071495	P	cellular response to endogenous stimulus	■						
GO:0009867	P	jasmonic acid mediated signaling pathway	■						
GO:0071395	P	cellular response to jasmonic acid stimulus	■						
GO:0071310	P	cellular response to organic substance	■						
GO:0051716	P	cellular response to stimulus	■						
GO:0070887	P	cellular response to chemical stimulus	■						
GO:0009628	P	response to abiotic stimulus	■						
GO:0032870	P	cellular response to hormone stimulus	■						
GO:0009755	P	hormone-mediated signaling pathway	■						
GO:0060255	P	regulation of macromolecule metabolic process	■						
GO:0019222	P	regulation of metabolic process	■						
GO:0009312	P	oligosaccharide biosynthetic process	■		■	■	■		
GO:0009311	P	oligosaccharide metabolic process	■		■	■	■		
GO:0008483	F	transaminase activity	■						
GO:0016769	F	transferase activity, transferring nitrogenous groups	■						
		transferase activity, transferring alkyl or aryl (other than methyl)	■						
		groups	■						
GO:0016765	F	groups	■						
GO:0046351	P	disaccharide biosynthetic process	■		■	■	■		
GO:0005984	P	disaccharide metabolic process	■		■	■	■		
GO:0016138	P	glycoside biosynthetic process	■		■	■	■		
GO:0016137	P	glycoside metabolic process	■		■	■	■		
GO:0009615	P	response to virus	■						
GO:0006457	P	protein folding	■						
GO:0006915	P	Apoptosis	■						



Rycina 28. Porównanie profiliów GO (wyznaczonych z zastosowaniem referencyjnego zestawu danych GO dla mikromacierzy TOB) uzyskanych za pomocą mikromacierzy SSH oraz CSH dla korzeni tytoniu poddanych stresowi solnemu.

Kolorem niebieskim w kolumnie „nr GO” oznaczono kategorie GO charakteryzujące się wzrostem ekspresji, różowym- spadkiem ekspresji. Kolorami w skali od żółtego do czerwonego oznaczono GO cechujące się istotną ekspresją różnicową (patrz Rycina 18), kolor szary oznacza brak istotnej zmiany ekspresji. Nie pokazano kategorii GO dla TOB, które nie miały różnicujących odpowiedników na mikromacierzach CSH. 1P- jedno powtórzenie, 2P- dwa powtórzenia, H- filtracja HOM, SC- filtracja MORF.

GO referencyjne: TOB, stres:CdCl ₂			TOB	POT 1P	POT 2P	POT H 2P	POT SC 2P	TOM1	TOM2
Nr GO	Typ	Opis							
GO:0009404	P	toxin metabolic process	■	■	■	■	■		
GO:0009407	P	toxin catabolic process	■	■	■	■	■		
GO:0019748	P	secondary metabolic process	■	■	■	■	■	■	
GO:0009607	P	response to biotic stimulus	■	■	■	■	■		
GO:0051707	P	response to other organism	■	■	■	■	■		
GO:0009620	P	response to fungus	■	■	■	■	■		
GO:0006952	P	defense response	■	■	■	■	■		
GO:0051704	P	multi-organism process	■	■	■	■	■		
GO:0010200	P	response to chitin	■	■	■	■	■		
GO:0042221	P	response to chemical stimulus	■	■	■	■	■		
GO:0050832	P	defense response to fungus	■	■	■	■	■		
GO:0006790	P	sulfur metabolic process	■	■	■	■	■		
GO:0006950	P	response to stress	■	■	■	■	■		
GO:0050896	P	response to stimulus	■	■	■	■	■		
GO:0009408	P	response to heat	■	■	■	■	■		
GO:0045087	P	innate immune response	■	■	■	■	■		
GO:0006955	P	immune response	■	■	■	■	■		
GO:0002376	P	immune system process	■	■	■	■	■		
GO:0044283	P	small molecule biosynthetic process	■	■	■	■	■		
GO:0009617	P	response to bacterium	■	■	■	■	■		
GO:0042398	P	cellular amino acid derivative biosynthetic process	■	■	■	■	■		
GO:0006575	P	cellular amino acid derivative metabolic process	■	■	■	■	■		
GO:0009743	P	response to carbohydrate stimulus	■	■	■	■	■		
GO:0006519	P	cellular amino acid and derivative metabolic process	■	■	■	■	■		
GO:0006725	P	cellular aromatic compound metabolic process	■	■	■	■	■		
GO:0009266	P	response to temperature stimulus	■	■	■	■	■		
GO:0044272	P	sulfur compound biosynthetic process	■	■	■	■	■		
GO:0042742	P	defense response to bacterium	■	■	■	■	■		
GO:0009611	P	response to wounding	■	■	■	■	■		
GO:0019438	P	aromatic compound biosynthetic process	■	■	■	■	■		
GO:0009753	P	response to jasmonic acid stimulus	■	■	■	■	■		
GO:0034641	P	cellular nitrogen compound metabolic process	■	■	■	■	■		
GO:0044271	P	cellular nitrogen compound biosynthetic process	■	■	■	■	■		
GO:0009605	P	response to external stimulus	■	■	■	■	■		
GO:0009628	P	response to abiotic stimulus	■	■	■	■	■		
GO:0008299	P	isoprenoid biosynthetic process	■	■	■	■	■		
GO:0006720	P	isoprenoid metabolic process	■	■	■	■	■		
GO:0009700	P	indole phytoalexin biosynthetic process	■	■	■	■	■		
GO:0052315	P	phytoalexin biosynthetic process	■	■	■	■	■		
GO:0052314	P	phytoalexin metabolic process	■	■	■	■	■		
GO:0042431	P	indole metabolic process	■	■	■	■	■		
GO:0046217	P	indole phytoalexin metabolic process	■	■	■	■	■		
GO:0044249	P	cellular biosynthetic process	■	■	■	■	■		
GO:0006355	P	regulation of transcription, DNA-dependent	■	■	■	■	■		
GO:0051252	P	regulation of RNA metabolic process	■	■	■	■	■		
GO:0006084	P	acetyl-CoA metabolic process	■	■	■	■	■		
GO:0004364	F	glutathione transferase activity	■	■	■	■	■		
GO:0016765	F	transferase activity, transferring alkyl or aryl (other than methyl) groups	■	■	■	■	■	■	■
GO:0016746	F	transferase activity, transferring acyl groups	■	■	■	■	■		
GO:0019825	F	oxygen binding	■	■	■	■	■		
GO:0000096	P	sulfur amino acid metabolic process	■	■	■	■	■		
GO:0044106	P	cellular amine metabolic process	■	■	■	■	■		
GO:0009698	P	phenylpropanoid metabolic process	■	■	■	■	■		
GO:0009308	P	amine metabolic process	■	■	■	■	■		
GO:0044281	P	small molecule metabolic process	■	■	■	■	■		
GO:0006520	P	cellular amino acid metabolic process	■	■	■	■	■		
GO:0009699	P	phenylpropanoid biosynthetic process	■	■	■	■	■		
GO:0006979	P	response to oxidative stress	■	■	■	■	■		
GO:0006351	P	transcription, DNA-dependent	■	■	■	■	■		
GO:0032774	P	RNA biosynthetic process	■	■	■	■	■		
GO:0008654	P	phospholipid biosynthetic process	■	■	■	■	■		
GO:0006644	P	phospholipid metabolic process	■	■	■	■	■		
GO:0046912	F	transferase activity, transferring acyl groups, acyl groups converted into alkyl on transfer	■	■	■	■	■		
GO:0008610	P	lipid biosynthetic process	■	■	■	■	■		



Rycina 29. Porównanie profiliw GO (wyznaczonych z zastosowaniem referencyjnego zestawu danych GO dla mikromacierzy TOB) uzyskanych za pomocą mikromacierzy SSH oraz CSH dla korzeni tytoniu poddanych stresowi kadmowemu.

Kolorem niebieskim w kolumnie „nr GO” oznaczono kategorie GO charakteryzujące się wzrostem ekspresji, różowym- spadkiem ekspresji. Kolorami w skali od żółtego do czerwonego oznaczono GO cechujące się istotną ekspresją różnicową (patrz Rycina 18), kolor szary oznacza brak istotnej zmiany ekspresji. Nie pokazano kategorii GO dla TOB, które nie miały różnicujących odpowiedników na mikromacierzach CSH. 1P- jedno powtórzenie, 2P- dwa powtórzenia, H- filtracja HOM, SC- filtracja MORF.

**GO referencyjne: specyficzne dla danej mikromacierzy;
stres:NaCl**

Nr GO	Typ	Opis	TOB	POT 1P	POT 2P	POT H 2P	POT SC 2P	TOM1	TOM2
GO:0010200	P	response to chitin	Red					Yellow	
GO:0042221	P	response to chemical stimulus	Red					Yellow	
GO:0010033	P	response to organic substance	Red					Yellow	
GO:0009743	P	response to carbohydrate stimulus	Red					Yellow	
GO:0048583	P	regulation of response to stimulus	Yellow					Yellow	
GO:0009864	P	induced systemic resistance, jasmonic acid mediated signaling pathway	Yellow					Orange	
GO:0009719	P	response to endogenous stimulus	Yellow					Yellow	
GO:0009682	P	induced systemic resistance	Yellow					Orange	
GO:0071495	P	cellular response to endogenous stimulus	Yellow					Orange	
GO:0009867	P	jasmonic acid mediated signaling pathway	Yellow					Orange	
GO:0031347	P	regulation of defense response	Yellow					Orange	
GO:0071395	P	cellular response to jasmonic acid stimulus	Yellow					Orange	
GO:0002682	P	regulation of immune system process	Yellow					Orange	
GO:0050776	P	regulation of immune response	Yellow					Orange	
GO:0009725	P	response to hormone stimulus	Yellow					Yellow	
GO:0045088	P	regulation of innate immune response	Yellow					Orange	
GO:0031349	P	positive regulation of defense response	Yellow					Orange	
GO:0071310	P	cellular response to organic substance	Yellow					Orange	
GO:0002684	P	positive regulation of immune system process	Yellow					Orange	
GO:0045089	P	positive regulation of innate immune response	Yellow					Orange	
GO:0050778	P	positive regulation of immune response	Yellow					Orange	
GO:0002218	P	activation of innate immune response	Yellow					Orange	
GO:0002253	P	activation of immune response	Yellow					Orange	
GO:0051716	P	cellular response to stimulus	Yellow					Yellow	
GO:0070887	P	cellular response to chemical stimulus	Yellow					Yellow	
GO:0080134	P	regulation of response to stress	Yellow					Orange	
GO:0048584	P	positive regulation of response to stimulus	Yellow					Orange	
GO:0007242	P	intracellular signaling cascade	Yellow					Orange	
GO:0032870	P	cellular response to hormone stimulus	Yellow					Orange	
GO:0009755	P	hormone-mediated signaling pathway	Yellow					Orange	
GO:0050896	P	response to stimulus	Red		Yellow	Yellow	Yellow		
GO:0006952	P	defense response	Red			Yellow	Yellow		
GO:0006950	P	response to stress	Red	Yellow	Orange	Orange	Orange		
GO:0010556	P	regulation of macromolecule biosynthetic process	Orange						
GO:0009889	P	regulation of biosynthetic process	Orange						
GO:0031326	P	regulation of cellular biosynthetic process	Orange						
GO:0006351	P	transcription, DNA-dependent	Orange	Yellow					
GO:0032774	P	RNA biosynthetic process	Orange	Yellow					
GO:0045449	P	regulation of transcription	Orange						
GO:0006355	P	regulation of transcription, DNA-dependent	Orange						
GO:0051171	P	regulation of nitrogen compound metabolic process	Orange						
GO:0051252	P	regulation of RNA metabolic process	Orange						
GO:0080090	P	regulation of primary metabolic process	Orange						
GO:0019219	P	regulation of nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	Orange						
GO:0006350	P	Transcription	Orange	Yellow					
GO:0031323	P	regulation of cellular metabolic process	Orange	Yellow					
GO:0010468	P	regulation of gene expression	Orange						
GO:0009415	P	response to water	Orange	Grey	Yellow	Yellow	Yellow		
GO:0060255	P	regulation of macromolecule metabolic process	Orange						
GO:0019222	P	regulation of metabolic process	Orange						
GO:0009414	P	response to water deprivation	Orange	Grey	Yellow	Yellow	Yellow		
GO:0016070	P	RNA metabolic process	Orange						
GO:0012501	P	programmed cell death	Orange	Grey		Yellow	Yellow		
GO:0003700	F	transcription factor activity	Red	Orange	Orange	Orange	Yellow		
GO:0030528	F	transcription regulator activity	Red	Grey			Yellow		
GO:0003677	F	DNA binding	Red	Yellow	Yellow	Yellow	Yellow		
GO:0004364	F	glutathione transferase activity	Red	Red	Red	Red	Red		
GO:0050794	P	regulation of cellular process	Yellow		Yellow				
GO:0044275	P	cellular carbohydrate catabolic process	Grey	Orange	Orange	Orange	Orange		
GO:0046164	P	alcohol catabolic process	Grey	Orange	Orange	Orange	Orange	Yellow	
GO:0019320	P	hexose catabolic process	Grey	Yellow			Yellow	Yellow	
GO:0006007	P	glucose catabolic process	Grey	Yellow			Yellow	Yellow	
GO:0046365	P	monosaccharide catabolic process	Grey	Yellow			Yellow	Yellow	
GO:0006096	P	Glycolysis	Grey	Yellow			Yellow	Yellow	
GO:0071365	P	cellular response to auxin stimulus	Grey					Orange	

GO:0051179	P	Localization							
GO:0007167	P	enzyme linked receptor protein signaling pathway							
GO:0023034	P	intracellular signaling pathway							
GO:0007169	P	transmembrane receptor protein tyrosine kinase signaling pathway							
GO:0006468	P	protein amino acid phosphorylation							
GO:0016310	P	Phosphorylation							
GO:0006796	P	phosphate metabolic process							
GO:0006793	P	phosphorus metabolic process							
GO:0004857	F	enzyme inhibitor activity							
GO:0030234	F	enzyme regulator activity							

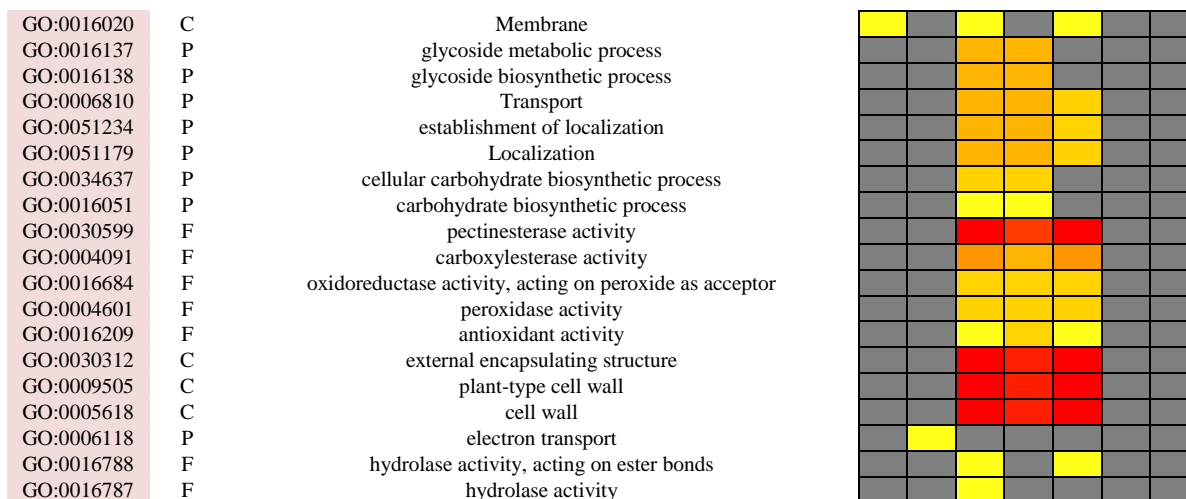
Rycina 30. Porównanie profiliw GO (wyznaczonych z zastosowaniem referencyjnych zestawów danych GO specyficznych dla poszczególnych mikromacierzy) uzyskanych za pomocą mikromacierzy SSH oraz CSH dla korzeni tytoniu poddanych stresowi solnemu.

Kolorem niebieskim w kolumnie „nr GO” oznaczono kategorie GO charakteryzujące się wzrostem ekspresji, różowym- spadkiem ekspresji. Kolorami w skali od żółtego do czerwonego oznaczono GO cechujące się istotną ekspresją różnicową (patrz Rycina 18), kolor szary oznacza brak istotnej zmiany ekspresji. Nie pokazano kategorii GO dla TOB, które nie miały różnicujących odpowiedników na mikromacierzach CSH. 1P- jedno powtórzenie, 2P- dwa powtórzenia, H- filtracja HOM, SC- filtracja MORF.

GO: specyficzne dla danej mikromacierzy; stres: CdCl₂

Nr GO	Typ	Opis	TOB	POT 1P	POT 2P	POT H 2P	POT SC 2P	TOM1	TOM2
GO:0019748	P	secondary metabolic process	Red	Yellow	Orange	Orange	Orange	Yellow	Grey
GO:0006790	P	sulfur metabolic process	Red	Yellow	Orange	Orange	Orange	Yellow	Grey
GO:0044283	P	small molecule biosynthetic process	Red	Yellow	Orange	Orange	Orange	Yellow	Grey
GO:0000096	P	sulfur amino acid metabolic process	Red	Yellow	Orange	Orange	Red	Yellow	Grey
GO:0042398	P	cellular amino acid derivative biosynthetic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0006575	P	cellular amino acid derivative metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0006519	P	cellular amino acid and derivative metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0006725	P	cellular aromatic compound metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0046500	P	S-adenosylmethionine metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0006556	P	S-adenosylmethionine biosynthetic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0000097	P	sulfur amino acid biosynthetic process	Orange	Yellow	Orange	Orange	Red	Yellow	Grey
GO:0044272	P	sulfur compound biosynthetic process	Orange	Yellow	Orange	Orange	Red	Yellow	Grey
GO:0019438	P	aromatic compound biosynthetic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0009309	P	amine biosynthetic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0034641	P	cellular nitrogen compound metabolic process	Orange	Yellow	Orange	Orange	Orange	Orange	Grey
GO:0044271	P	cellular nitrogen compound biosynthetic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0044106	P	cellular amine metabolic process	Orange	Yellow	Orange	Orange	Orange	Orange	Grey
GO:0009119	P	ribonucleoside metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0008652	P	cellular amino acid biosynthetic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0009698	P	phenylpropanoid metabolic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0046128	P	purine ribonucleoside metabolic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0042278	P	purine nucleoside metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0046394	P	carboxylic acid biosynthetic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0016053	P	organic acid biosynthetic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0009308	P	amine metabolic process	Orange	Yellow	Orange	Orange	Orange	Orange	Grey
GO:0044281	P	small molecule metabolic process	Orange	Yellow	Orange	Orange	Orange	Orange	Grey
GO:0006520	P	cellular amino acid metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0009116	P	nucleoside metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0009699	P	phenylpropanoid biosynthetic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0009066	P	aspartate family amino acid metabolic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0006555	P	methionine metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0008299	P	isoprenoid biosynthetic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0006720	P	isoprenoid metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0043436	P	oxoacid metabolic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0019752	P	carboxylic acid metabolic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0006082	P	organic acid metabolic process	Orange	Grey	Yellow	Grey	Grey	Yellow	Grey
GO:0042180	P	cellular ketone metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0044249	P	cellular biosynthetic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0055086	P	nucleobase, nucleoside and nucleotide metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Grey
GO:0004364	F	glutathione transferase activity	Red	Red	Red	Red	Red	Red	Grey
GO:0016765	F	transferase activity, transferring alkyl or aryl (other than methyl) groups	Red	Red	Red	Red	Red	Red	Orange
GO:0004478	F	methionine adenosyltransferase activity	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0016740	F	transferase activity	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0016801	F	hydrolase activity, acting on ether bonds	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0016667	F	oxidoreductase activity, acting on sulfur group of donors	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0019825	F	oxygen binding	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0016491	F	oxidoreductase activity	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0009693	P	ethylene biosynthetic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0009692	P	ethylene metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0043449	P	cellular alkene metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0043450	P	alkene biosynthetic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0016838	F	carbon-oxygen lyase activity, acting on phosphates	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0044255	P	cellular lipid metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0008610	P	lipid biosynthetic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0016114	P	terpenoid biosynthetic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0009809	P	lignin biosynthetic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0009808	P	lignin metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0006629	P	lipid metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0006721	P	terpenoid metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0009058	P	biosynthetic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0044237	P	cellular metabolic process	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0008171	F	O-methyltransferase activity	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0016616	F	oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0008168	F	methyltransferase activity	Orange	Orange	Orange	Orange	Orange	Orange	Orange
GO:0016741	F	transferase activity, transferring one-carbon groups	Orange	Orange	Orange	Orange	Orange	Orange	Orange

GO:0005829	C	Cytosol							
GO:0006059	P	hexitol metabolic process							
GO:0019405	P	alditol catabolic process							
GO:0019407	P	hexitol catabolic process							
GO:0046174	P	polyol catabolic process							
GO:0019594	P	mannitol metabolic process							
GO:0019592	P	mannitol catabolic process							
GO:0019400	P	alditol metabolic process							
GO:0019751	P	polyol metabolic process							
GO:0009067	P	aspartate family amino acid biosynthetic process							
GO:0032501	P	multicellular organismal process							
GO:0046148	P	pigment biosynthetic process							
GO:0042440	P	pigment metabolic process							
GO:0044282	P	small molecule catabolic process							
GO:0007275	P	multicellular organismal development							
GO:0006807	P	nitrogen compound metabolic process							
GO:0032502	P	developmental process							
GO:0004014	F	adenosylmethionine decarboxylase activity							
GO:0046029	F	mannitol dehydrogenase activity							
GO:0031320	F	hexitol dehydrogenase activity							
GO:0003773	F	heat shock protein activity							
GO:0016831	F	carboxy-lyase activity							
GO:0016829	F	lyase activity							
GO:0016830	F	carbon-carbon lyase activity							
GO:0016614	F	oxidoreductase activity, acting on CH-OH group of donors							
GO:0016853	F	isomerase activity							
GO:0016859	F	cis-trans isomerase activity							
GO:0003755	F	peptidyl-prolyl cis-trans isomerase activity							
GO:0005528	F	FK506 binding							
GO:0005527	F	macrolide binding							
GO:0008144	F	drug binding							
GO:0051186	P	cofactor metabolic process							
GO:0009791	P	post-embryonic development							
GO:0003824	F	catalytic activity							
GO:0071365	P	cellular response to auxin stimulus							
GO:0009734	P	auxin mediated signaling pathway							
GO:0023033	P	signaling pathway							
GO:0009733	P	response to auxin stimulus							
GO:0006595	P	polyamine metabolic process							
GO:0009755	P	hormone-mediated signaling pathway							
GO:0032870	P	cellular response to hormone stimulus							
GO:0008203	P	cholesterol metabolic process							
GO:0071495	P	cellular response to endogenous stimulus							
GO:0008216	P	spermidine metabolic process							
GO:0006695	P	cholesterol biosynthetic process							
GO:0008295	P	spermidine biosynthetic process							
GO:0009835	P	Ripening							
GO:0006769	P	nicotinamide metabolic process							
GO:0042401	P	cellular biogenic amine biosynthetic process							
GO:0046496	P	nicotinamide nucleotide metabolic process							
GO:0006596	P	polyamine biosynthetic process							
GO:0006730	P	one-carbon metabolic process							
GO:0006740	P	NADPH regeneration							
GO:0006098	P	pentose-phosphate shunt							
GO:0008202	P	steroid metabolic process							
GO:0019362	P	pyridine nucleotide metabolic process							
GO:0071310	P	cellular response to organic substance							
GO:0006576	P	cellular biogenic amine metabolic process							
GO:0016126	P	sterol biosynthetic process							
GO:0006739	P	NADP metabolic process							
GO:0009815	F	1-aminocyclopropane-1-carboxylate oxidase activity							
GO:0016641	F	oxidoreductase activity, acting on the CH-NH2 group of donors, oxygen as acceptor							
GO:0016638	F	oxidoreductase activity, acting on the CH-NH2 group of donors							
GO:0016407	F	acetyltransferase activity							
GO:0031418	F	L-ascorbic acid binding							
GO:0008080	F	N-acetyltransferase activity							
GO:0016410	F	N-acyltransferase activity							
GO:0010876	P	lipid localization							
GO:0006869	P	lipid transport							
GO:0016798	F	hydrolase activity, acting on glycosyl bonds							
GO:0008289	F	lipid binding							
GO:0004553	F	hydrolase activity, hydrolyzing O-glycosyl compounds							



Rycina 31. Porównanie profiliw GO (wyznaczonych z zastosowaniem referencyjnych zestawów danych GO specyficznych dla poszczególnych mikromacierzy) uzyskanych za pomocą mikromacierzy SSH oraz CSH dla korzeni tytoniu poddanych stresowi kadmowemu.

Kolorem niebieskim w kolumnie „nr GO” oznaczono kategorie GO charakteryzujące się wzrostem ekspresji, różowym- spadkiem ekspresji. Kolorami w skali od żółtego do czerwonego oznaczono GO cechujące się istotną ekspresją różnicową (patrz Rycina 18), kolor szary oznacza brak istotnej zmiany ekspresji. Nie pokazano kategorii GO dla TOB, które nie miały różnicujących odpowiedników na mikromacierzach CSH. 1P- jedno powtórzenie, 2P- dwa powtórzenia, H- filtracja HOM, SC- filtracja MORF.

IV.3.1. Analiza porównawcza eksperymentów CSH wykonanych z zastosowaniem mikromacierzy cDNA oraz L-oligo

Aby określić wpływ długości i rodzaju stosowanych sond na wyniki hybrydyzacji CSH, dokonano porównania danych dotyczących ekspresji różnicowej uzyskanych z zastosowaniem mikromacierzy TOM1 (sondy cDNA) i TOM2 (sondy L-oligo). Dane z mikromacierzy TOM1 i TOM2 zbadano pod kątem zgodności z wynikami uzyskanymi za pomocą mikromacierzy TOB, przy czym liczba sond posiadających odpowiedniki na mikromacierzy TOB była większa w przypadku mikromacierzy cDNA (7709-TOM1, 5215-TOM2, Rycina 6). Za pomocą mikromacierzy TOM1 uzyskano znacznie większą (ponad pięciokrotnie dla stresu solnego, ponad sześciokrotnie dla kadmowego) liczbę genów różnicujących po filtracji na podstawie wartości $p < 0,05$ (Tabela 23, kolumny 1,2). Po ujednoczeniu liczby genów różnicujących (listy „top600”) i ocenie, w jakim stopniu wyniki są zgodne z uzyskanymi na mikromacierzy referencyjnej, przewaga mikromacierzy cDNA jest również wyraźnie widoczna. Świadczy o niej większy odsetek sond, których odpowiedniki na mikromacierzy TOB również wykazywały różnicową ekspresję odpowiadających im genów, większa zgodność z danymi referencyjnymi dotycząca kierunku zmiany ekspresji (zmniejszenie lub zwiększenie), znacznie niższa średnia wartość p (Tabela 24, kolumna 1 i 2), a także wyższa korelacja uzyskanych wartości $\log_{2}FC$ (Tabela 25, kolumna 1 i 2). Mimo iż dla mikromacierzy cDNA uzyskano wyniki korzystniejsze, badania korelacji z danymi referencyjnymi wskazują na niską zbieżność w przypadku użycia wszystkich sond dla danej mikromacierzy posiadających odpowiedniki na mikromacierzy TOB. Korelacja profili ekspresji wzrastała w przypadku obu typów sond po zawężeniu badanej puli tylko do genów różnicujących, jednak uzyskane wartości były nadal stosunkowo niskie (Tabela 25, kolumna 1 i 2). Kontrolne analizy korelacji z zastosowaniem losowej puli zawierającej taką samą liczbę sond wskazywały na korelację zbliżoną do uzyskanej podczas analizy wszystkich sond, co dowodzi, że obserwowany wzrost korelacji wynika z wyselekcjonowania bardziej wartościowych danych, a nie z samego zawężenia puli. Oprócz danych prezentowanych w Tabeli 25, porównanie wartości $\log_{2}FC$ dla pojedynczych, odpowiadających sobie sond SSH i CSH zobrazowano na Rycinie 32.

Wykorzystując uzyskane dane dotyczące genów różnicujących, przeprowadzono analizę funkcjonalną GO i porównano otrzymane profile (Tabela 26,27, kolumna 1 i 2). W przypadku obu typów mikromacierzy: cDNA i oligo, wygenerowano profile GO dla tytoniu (z użyciem sond tytoniowych będących odpowiednikami różnicujących sond

TOM1 i TOM2) i pomidora (z użyciem różnicujących sond TOM1 i TOM2). W przypadku mikromacierzy L-oligo wykryto bardzo małą liczbę genów różnicujących, które wykazywały zgodność z danymi referencyjnymi (geny SAM, GST, MAT, ERD9 dla stresu kadmowego, patrz Tabela 20-22), dlatego liczba różnicujących kategorii GO uzyskanych podczas analizy funkcjonalnej była również bardzo znikoma, niezależnie od stosowanej metody badawczej. Zatem także w przypadku analizy funkcjonalnej, dużo bardziej wartościowe dane uzyskano korzystając z mikromacierzy cDNA. Otrzymane dla tego typu sond wyniki różniły się w zależności od zastosowanej metody analitycznej- tworzącej profil dla tytoniu bądź pomidora.

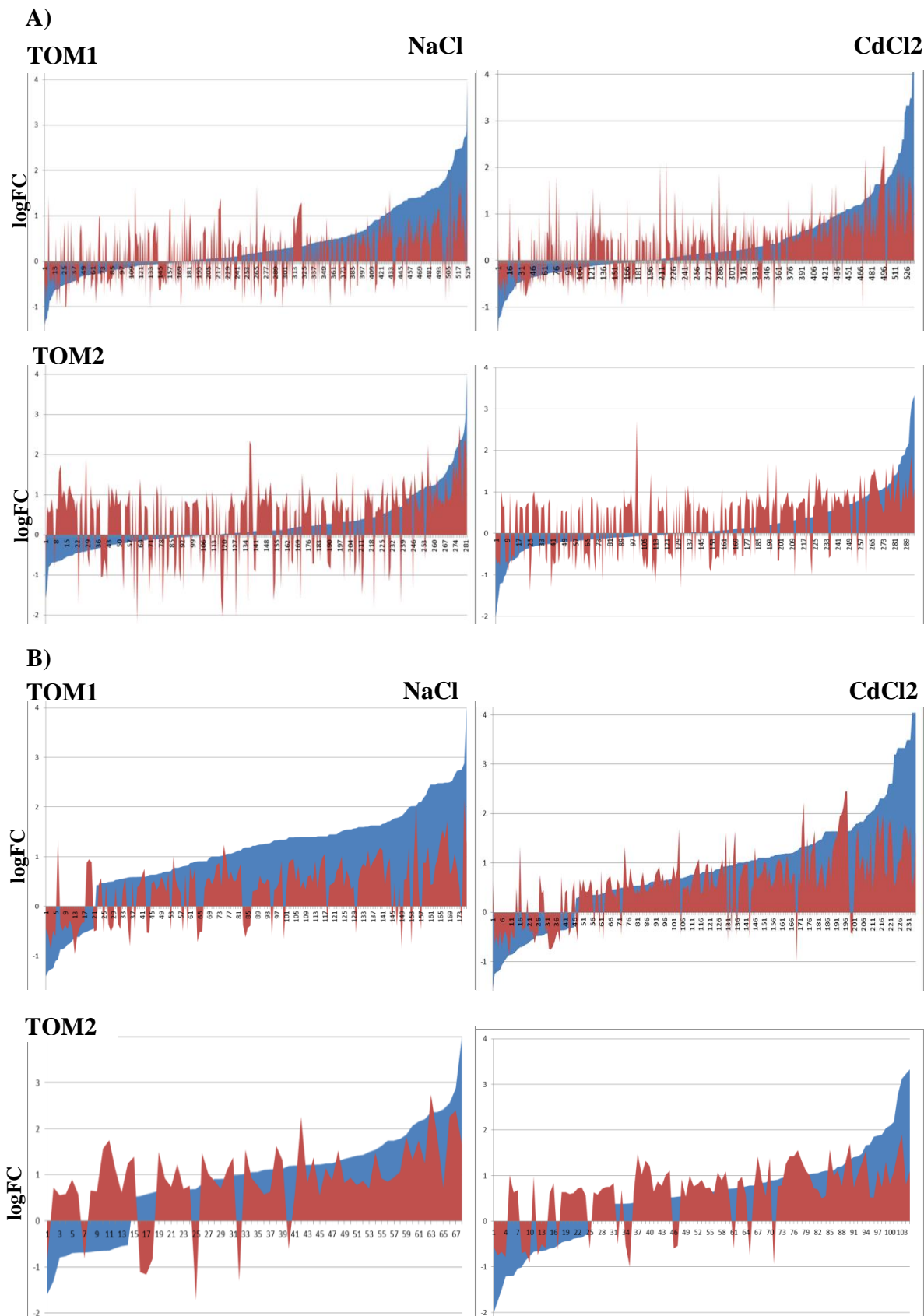
W przypadku oceny ontologii genów na podstawie sekwencji tytoniowych (Tabela 26, kolumna 1; Rycina 28,29) na uwagę zasługuje niewielka liczba grup GO wykazujących fałszywie pozytywne różnicowanie, (żadnej w stresie solnym, 3 w kadmowym). Z drugiej strony za pomocą tej metody wykryto mniejszą ilość zmian: w przypadku traktowania NaCl różnicujące kategorie GO potwierdzone na mikromacierzy referencyjnej ograniczały się do sygnalizacji komórkowej i odpowiedzi na stres (np. geny WRKY, ERF, RTF1, ATAF1, LOX3, HSF, ELI3), z kolei w ramach stresu kadmowego wyselekcjonowano kategorie związane ze szlakami biosyntezy (np. seryny, acetylokoenzymu A- geny PSAT, MAT, CYP). Należy zwrócić także uwagę na brak wspólnych kategorii GO wykrytych dla obu stresów, chociaż w obu grupach stwierdzono podobną ekspresję różnicową pojedynczych genów charakterystycznych dla stresu abiotycznego (np. MAPK, GST).

W przypadku analizy GO z uwzględnieniem sekwencji dla pomidora (Tabela 27, kolumna 1; Rycina 30,31), dla obu traktowań stwierdzono ponad dwukrotnie większą liczbę kategorii różnicujących, niż w przypadku drugiego podejścia, jednak znaczny odsetek stanowiły kategorie genów o fałszywie pozytywnej ekspresji. Stwierdzono osiem kategorii różnicujących GO wspólnych dla obu stresów. Cztery z nich wykazywały również istotną indukcję na mikromacierzy TOB i związane były ze ścieżką sygnałową odpowiedzi na stres indukowaną hormonami (geny ERF, GST, AP2, WRKY), a pozostałe cztery co prawda nie cechowały się istotną zmianą ekspresji w tytoniu ale dotyczyły pokrewnych kategorii (związanych z odpowiedzią na auksyny) i zawierały podobne geny (np. GST). Z kolei w przypadku analiz specyficznych dla danego stresu, podobnie jak w przypadku poprzedniego podejścia, w korzeniach traktowanych NaCl wyselekcjonowano znaczną liczbę kategorii GO związanych z odpowiedzią na stres i sygnalizacją za pomocą kwasu jasmonowego i auksyn. Oprócz tego stwierdzono wzrost ekspresji grup związanych z glikolizą i katabolizmem glukozy (np. enzymy glikolizy-dehydrogenaza pirogronianowa,

dehydrogenaza aldehydu-3-fosfoglicerynowego, enolaza, fosfofruktokinaza, których indukcję obserwowano również w przypadku oceny na mikromacierzach TOB, jednak bez odnotowania zmiany ekspresji w ramach grupy GO), transportem hormonów (pirofosfataza nieorganiczna) oraz aktywnością transporterów śródłonowych (pirofosfataza nieorganiczna, WRKY6, transporter ATM1 (mitochondrialny transporter ABC zaangażowany w odpowiedź na stres abiotyczny, m.in. kadmowy [183])).

W przypadku drugiego podejścia analitycznego stwierdzono większą liczbę grup GO wytypowanych jako różnicujące w CSH, a nie na mikromacierzy referencyjnej („fałszywie pozytywne”): można je podzielić na kilka zespołów. Jeden z nich tworzyły wspomniane już kategorie związane z sygnalizacją za pomocą auksyn i odpowiedzią komórkową (GST, CDPK, WRKY). Kolejny zespół tworzyły kategorie związane z syntezą poliamin, zawierające szereg sond dla genu dekarboksylazy argininy. Wzrost aktywności tego enzymu był obserwowany w korzeniach rzodkiewnika poddanych działaniu kadmu i związany był z antyoksydacyjnym działaniem poliamin [184]. Mimo braku indukcji dla tej kategorii GO, na mikromacierzy TOB odnotowano istotne zwiększenie ekspresji dla wszystkich sond tego genu. Indukcji ulegały także geny związane z syntezą cholesterolu (np. MVD, syntaza HMG-CoA). W przypadku części z nich wzrost ekspresji obserwowano również na mikromacierzy referencyjnej w ramach kategorii GO dotyczącej syntezy acetyloCoA. Kategoria „dojrzewanie” zawierała opisywane już geny związane z syntezą etylenu, jak EFE, ACS. Pozostałe geny należały do typu GO „funkcja molekularna” i dotyczyły głównie opisywanych już enzymów.

Wykonana analiza wskazuje na przydatność mikromacierzy cDNA w analizie CSH ,i wyodrębnienie części genów lub procesów zaangażowanych w stres abiotyczny przy stosunkowo niewielkiej liczbie obserwacji fałszywie pozytywnych.



Rycina 32. Wykresy warstwowe obrazujące porównanie wartości logFC TOM1 lub TOM2 z odpowiadającymi im sondami referencyjnymi (TOB), na poziomie: A) listy top600, B) wspólnych genów wykazujących różnicowanie zarówno dla mikromacierzy badanej, jak i referencyjnej. Wartości referencyjne posortowano rosnąco.

IV.3.2. Badanie wpływu pokrewieństwa gatunku badanego i reprezentowanego na mikromacierzy na wynik analizy ekspresji genów metodą CSH

Kolejnym celem badań była ocena wyników CSH uzyskanych z zastosowaniem mikromacierzy reprezentujących dwa różne gatunki z rodziny *Solanaceae*. Analiza homologii sekwencji oraz dane literaturowe dotyczące filogenezy wskazują, że obydwie gatunki tj. pomidor i ziemniak charakteryzują się zbliżonym podobieństwem do tytoniu [92]. Korelacja pomiędzy wartościami Bit Score uzyskanymi dla odpowiadających sobie sond POT i TOM1 wynosiła 0,61 ($p < 2,2e-16$). Wyniki porównań zaprezentowane są w Tabelach 23-27 i na Rycinach 28-31. Obie mikromacierze charakteryzowały się zbliżoną liczbą sond posiadających odpowiedniki na mikromacierzy TOB (Rycina 6.). Odsetek sond różnicujących, posiadających odpowiedniki na mikromacierzy TOB, także był podobny w przypadku obu mikromacierzy (od 33% do 44%), natomiast mikromacierz POT wykazywała w tej grupie większą zgodność zmian z platformą referencyjną (94 i 97% vs 79 i 89%) (Tabela 24. kolumna 1, 3). Współczynnik korelacji wartości logFC na wszystkich badanych poziomach był zdecydowanie korzystniejszy dla mikromacierzy POT (Tabela 25, kolumna 1, 3). Korelacja ta osiągała wartość do 0,79 dla wspólnych genów różnicujących POT i TOB. Ta duża zbieżność danych widoczna jest również na wykresach zaprezentowanych na Rycinie 33.

W przypadku analizy GO, podobnie jak w poprzednim punkcie, wyniki uzyskano z wykorzystaniem dwóch podejść: na podstawie GO tytoniowych oraz GO dla mikromacierzy TOM1 i POT (pomidora i ziemniaka). Dla pierwszego podejścia, w przypadku traktowania NaCl odnotowano podobną liczbę grup GO zbieżnych z profilem ekspresji genów tytoniu dla obu mikromacierzy, jednak mikromacierz TOM1 charakteryzowała się brakiem GO fałszywie pozytywnych. Z kolei podczas analizy korzeni tytoniu traktowanych chlorkiem kadmu stwierdzono niemal trzykrotnie większą liczbę różnicujących grup GO wyselekcjonowanych na mikromacierzy POT (Tabela 26. kol.1 i 3). Charakterystykę zmian wykrytych w wyniku eksperymentów na mikromacierzach TOM1 opisano dokładniej w punkcie poprzednim (IV.3.1). W przeciwieństwie do mikromacierzy TOM1 (brak grup GO wspólnych dla obu traktowań), dla POT odnotowano 11 grup GO charakteryzujących się wzrostem ekspresji w przypadku obu stresów. Były to kategorie związane z odpowiedzią na stres (geny ATAF, CDPK, MAPK, JAZ, ERD) i metabolizmem toksyn (GST). Obie te kategorie wykazywały indukcję także na mikromacierzy referencyjnej i to w obu stresach.

Wśród grup o ekspresji zwiększonej wyłącznie pod wpływem stresu NaCl stwierdzono m.in. grupy genów odpowiedzialnych za regulację transkrypcji. Odnotowano także kategorie GO, które nie figurowały na liście GO różnicujących dla mikromacierzy TOB. Były one związane z biosyntezą oligosacharydów (fosfataza trehalozowa, syntaza sacharozowa), glikolizą (fosfofruktokinaza, kinaza fosfoglicerynianowa) oraz aktywnością transaminaz (aminotransferaza asparaginianowa i alaninowa). Indukcję ekspresji aminotransferaz stwierdzono również na mikromacierzy TOB, a fosfataza trehalozowa jest enzymem biorącym udział w syntezie trehalozy: cukru, którego akumulację obserwowano w stresie abiotycznym u roślin, wykazującego działanie osmoprotekcyjne [186]. Jedyną kategorią wykazującą zmniejszenie ekspresji była GO „region pozakomórkowy”, obejmująca geny kodujące enzymy takie jak EXGT (inhibicja na mikromacierzy TOB), peptydazy serynowe, germiny i RuBisCO, o różnorodnej specyfice działania.

Z kolei GO specyficzne dla stresu kadmowego w większości przypadków pokrywały się z obserwowanymi na mikromacierzy TOB. Odnotowano tu indukcję genów związanych z metabolizmem siarki (SIR), acetylokoenzymu A (MVA), etylenu (EFE, ACS6), biosyntezą terpenów (TPS21), wiązaniem tlenu (CYP). Jedynie trzy kategorie GO ulegające indukcji nie były odnotowane wśród grup różnicujących na mikromacierzy TOB. Zawierały one geny związane z syntezą fosfolipidów (MVA, MVD) i transferaz (GST), obserwowane już w innych grupach.

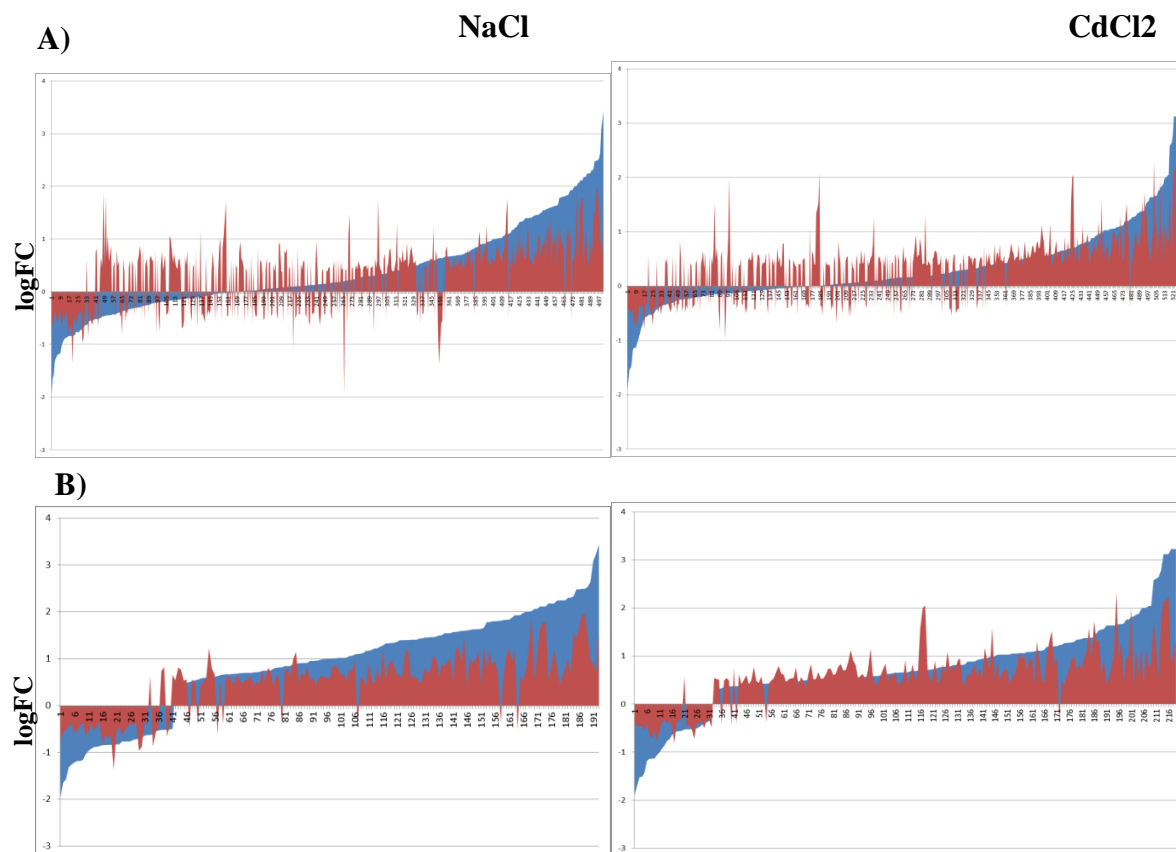
Podsumowując, w przypadku analizy bazującej na sekwencjach homologicznych do tytoniowych występujących na mikromacierzy POT można było zaobserwować większą różnorodność grup genów ulegających ekspresji różnicowej oraz wierniejsze odwzorowanie profilu referencyjnego (wspólny profil stresu abiotycznego). Odnotowano 8 kategorii GO wykazujących istotne zmiany na obu mikromacierzach w stresie solnym oraz 10 w stresie kadmowym.

Analiza profili GO charakterystycznych dla gatunków reprezentowanych na mikromacierzy (pomidora i ziemniaka) w roślinach poddanych działaniu stresu solnego wskazywała na nieco lepsze wyniki uzyskane za pomocą mikromacierzy TOM1 (większa liczba GO zbieżnych, podobna liczba GO fałszywie pozytywnych), z kolei w przypadku stresu kadmowego wykryto bardzo zbliżoną liczbę grup GO zbieżnych z mikromacierzą TOB dla obu mikromacierzy, jednak mniej zmian fałszywie pozytywnych odnotowano dla mikromacierzy POT (Tabela 27. kol.1 i 3).

W przypadku analizy profili ekspresji genów z wykorzystaniem kategorii GO dla ziemniaka zaobserwowano zdecydowanie większe rozbieżności pomiędzy wynikami dla

poszczególnych stresów (zaledwie 2 kategorie GO wspólne dla obu stresów, obie związane z aktywnością GST). Dla stresu NaCl odnotowano 15 kategorii GO wspólnych dla obu sposobów analizy oraz 20 wspólnych z mikromacierzą referencyjną. Wśród tych ostatnich stwierdzono indukcję genów związanych z odpowiedzią na stres (np. HSP, LEA, ER6-białko regulowane etylenem, MAPK) oraz szeregu grup związanych z regulacją transkrypcji (WRKY, EREBP6, ERF5, ARF1). Wśród kategorii GO fałszywie pozytywnych wyróżnić można grupy związane z biosyntezą cukrów (oligosacharydów, glikozydów (syntaza sacharozowa, fosfataza trehalozowa), glikolizą (GAPDH, kinaza pirogronianowa, fosfofruktokinaza, kinaza fosfoglicerynianowa), a także kategorie należące do typu „funkcja molekularna” i związane z aktywnością enzymów, zawierające geny peroksydaz (w tym glutationowej) oraz GST. Podobnie jak w przypadku mikromacierzy referencyjnej, grupy GO dotyczące aktywności peroksydaz wykazywały istotne różnicowanie również w przypadku sond charakteryzujących się spadkiem ekspresji. W korzeniach tytoniu poddanych działaniu stresu kadmowego obserwowano zwiększenie ekspresji grup GO związanych z syntezą izoprenoidów i terpenów (syntaza seskwiterpenów), oraz syntezą S-adenozylometioniny (SAM), które ulegały indukcji również w przypadku stosowania mikromacierzy TOB. Do grup genów nieulegających ekspresji na mikromacierzy referencyjnej należały kategorie związane z syntezą ligniny (katecholo-O-metylotransferaza) budującej ściany komórkowe i indukowanej w wyniku stresu [187], pozostałe grupy typu „proces biologiczny” tylko uzupełniały opisywane powyżej procesy. W przypadku typu „funkcja molekularna” wzrost ekspresji wykazywały GO opisujące aktywność enzymów: izomeraż, GST, O-metylotransferaz, oksydoreduktaz oraz wiązanie leków. Jediną kategorią GO ulegającą inhibicji był „transport elektronów” (geny cytochromu c6, reduktaza chelatów żelaza).

Podsumowując powyższe spostrzeżenia można uznać, że wyniki badań profilu ekspresji genów uzyskane z uwzględnieniem pojedynczych sond wskazują, iż lepsze odzwierciedlenie wyników SSH uzyskujemy na mikromacierzy POT, z kolei na podstawie profilowania GO trudno jest jednoznacznie wskazać, która z mikromacierzy prezentuje dane o większej wartości biologicznej.



Rycina 33. Profil ekspresji sond na mikromacierzy POT na tle próby referencyjnej na poziomie A) całej puli top 600 B) frakcji wspólnych genów różnicujących dla mikromacierzy badanej i referencyjnej. Wartości referencyjne posortowano rosnąco.

IV.3.3. Wpływ dodatkowych powtórzeń technicznych na wyniki CSH

W kolejnym etapie badań zdecydowano się sprawdzić, czy wykonanie dodatkowych powtórzeń eksperymentów mikromacierzowych będzie wpływać na jakość uzyskiwanych danych. Przeprowadzono zatem dodatkowy eksperyment (wykorzystując mikromacierz POT) polegający na uzyskaniu powtórzenia technicznego dla każdej mikromacierzy (łącznie 6 dodatkowych mikromacierzy) z zamianą barwników (ang. *dye swap*). W kolumnach 3 i 4 Tgtabel 26-30 zamieszczono porównanie danych POT uzyskanych dla jednego i dwóch powtórzeń technicznych w odniesieniu do wyników uzyskanych dla mikromacierzy TOB. Wyniki wskazują, że zastosowanie dodatkowego powtórzenia technicznego powoduje zmniejszenie liczby sond różnicujących (Tabela 23, kolumna 3 i 4). Jednocześnie poprawie uległy wyniki dotyczące udziału sond oznaczonych jako różnicujące na mikromacierzy referencyjnej, a także zgodności kierunku ekspresji (do 98%) (Tabela 24, kolumna 3 i 4), co świadczy o odrzuceniu sond prezentujących mniej wiarygodne dane. Ponadto widoczne są korzystne różnice w korelacji danych z danymi uzyskanymi dla TOB we wszystkich analizowanych zbiorach - współczynnik korelacji wzrastał średnio o ok. 0,1 w porównaniu z danymi uzyskanymi przy pojedynczym powtórzeniu technicznym (Tabela 25, kolumna 3 i 4). Współczynnik korelacji dla wspólnych sond różnicujących osiąga wartość do 0,86.

Również analizy GO wskazują w każdym przypadku na możliwość uzyskania bardziej wartościowych wyników po dodaniu jednego powtórzenia (Tabela 26,27, kolumna 3 i 4). Oprócz profilu GO specyficznego dla mikromacierzy POT w stresie solnym (Tabela 27 kolumna 1 i 4), gdzie uzyskano wyniki zbliżone do TOM1, w pozostałych przypadkach wyniki otrzymane dla dwóch powtórzeń na mikromacierzy POT wykazywały największą zbieżność z uzyskanymi na mikromacierzy referencyjnej ze wszystkich wykonanych eksperymentów.

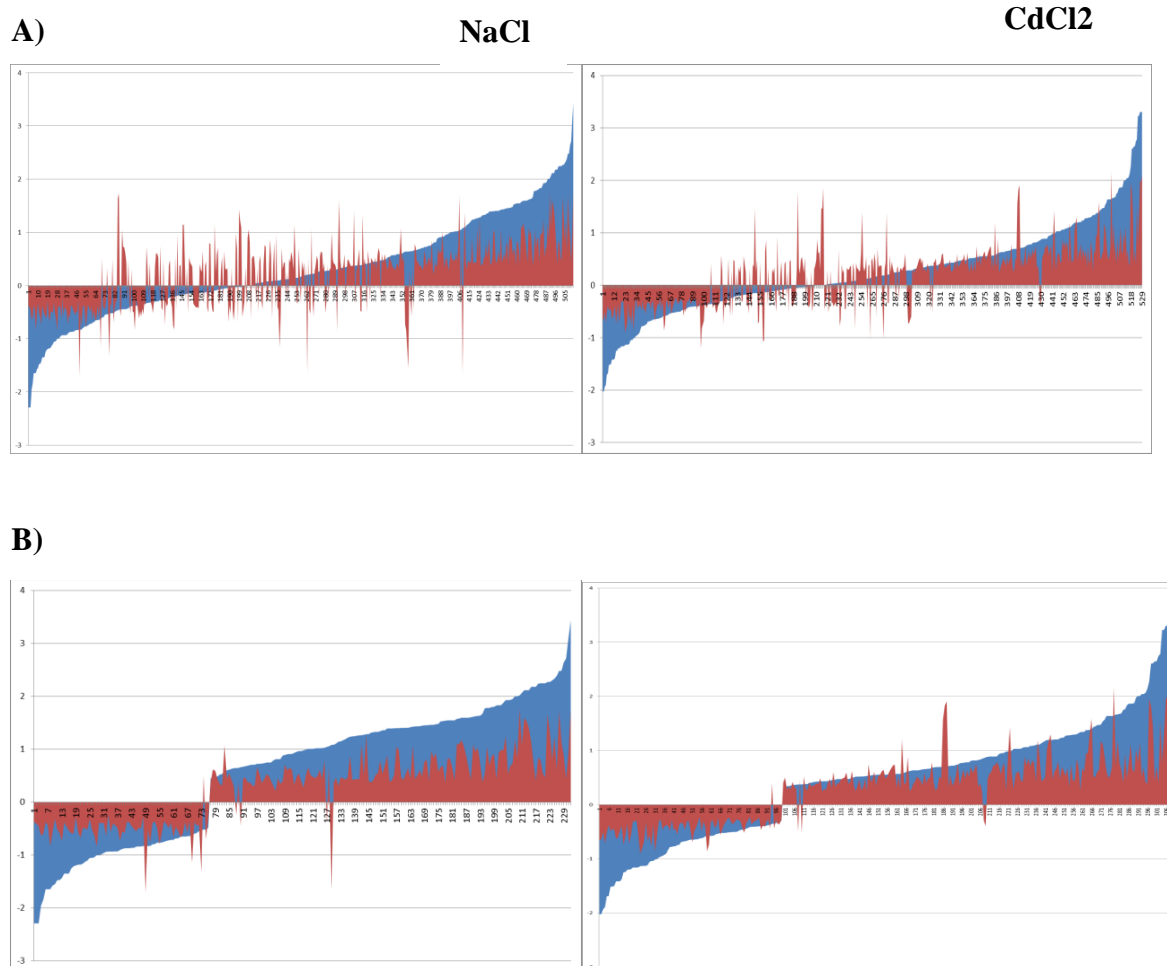
Biorąc pod uwagę analizę GO wykonaną pod kątem genów tytoniu, uzyskano odpowiednio o 38% i 46% więcej różnicujących kategorii GO, odpowiednio dla stresu solnego i kadmowego, wykrywając w ten sposób ok. 30% różnicujących GO uzyskanych na mikromacierzy referencyjnej. Odnotowano 20 wspólnych kategorii GO dla obu profili, chociaż wszystkie związane były z odpowiedzią na stres i aktywnością GST, podobnie jak dla pojedynczego powtórzenia. W przypadku genów specyficznych dla stresu solnego, dodatkowe kategorie dotyczyły głównie procesów regulacyjnych. Oprócz tego odnotowano 5 dodatkowych kategorii nie wykazujących istotnych zmian na mikromacierzy TOB, powiązanych z syntezą cukrów i odpowiedzią na stres. Korzystny

wpływ dodatkowego powtórzenia widoczny jest szczególnie w przypadku GO różnicujących cechujących się spadkiem ekspresji, gdzie uwidoczniła została inhibicja transporterów (TIP, PIP, ZIP). Dodatkowo, niewykazujące istotnych zmian na mikromacierzy TOB kategorie dotyczyły wiązania jonów miedzi (metalotioneiny), aktywności niektórych enzymów (XTH) i organizacji ściany komórkowej. Spośród tych kategorii wątpliwości w kontekście zbieżności z mikromacierzą referencyjną budzi jedynie wiązanie jonów miedzi. W przypadku stresu kadmowego dodatkowe powtórzenie pozwoliło na wyselekcjonowanie dodatkowych kategorii z zakresu metabolizmu (np. synteza fitoaleksyn) lub reakcji na stres (np. odpowiedź na kwas jasmonowy) dla genów o wzroście ekspresji, a także, podobnie jak w przypadku stresu solnego, na wyselekcjonowanie grup ujemnie różnicujących związanych z transportem przez błonowy i wiązaniem tłuszczów (zgodność z TOB). W przypadku genów „fałszywie pozytywnych” obserwowano obniżenie ekspresji GO dotyczących organizacji ściany komórkowej (np. XTH, pektynoesteraza – enzym biorący udział w procesie rozluźniania ściany komórkowej) i cytoszkieletu (tubulina). Są to zmiany, które można dostrzec na mikromacierzy referencyjnej. W tej grupie niepokojące może być jedynie zmniejszenie ekspresji GO związanych z odpowiedzią na brak wody. W ich skład weszły geny kodujące transportery błonowe (np. PIP), jak również geny odpowiedzi na odwodnienie (RD22, SAG21), których spadek ekspresji trudno jest wyjaśnić.

Analiza GO wykonana pod kątem genów dla ziemniaka również wskazywała na korzystny wpływ stosowania dwóch powtórzeń technicznych, choć w tym przypadku liczba kategorii „fałszywie pozytywnych” była znaczna (ok. dwukrotnie większa w przypadku stresu kadmowego). W stresie solnym stwierdzono jedynie 6 dodatkowych grup GO, o podobnym profilu genów do wykrytych za pomocą pojedynczego powtórzenia. Spośród grup GO wykazujących ekspresję różnicową pomimo braku na mikromacierzy TOB zwrócono uwagę na cztery grupy GO dotyczące rozwoju, które zawierały geny kodujące białka LEA, bardzo charakterystyczne dla stresu solnego. Pozostałe kategorie uzupełniały profil opisywany przy zastosowaniu jednego powtórzenia. Grupy GO o spadku ekspresji związane były, podobnie jak w poprzednim podejściu, z transportem przez błonowy i aktywnością peroksydaz.

W przypadku stresu kadmowego, wśród nowych kategorii funkcjonalnych niewyselekcjonowanych na mikromacierzy referencyjnej wyróżnił się zespół kategorii GO związanych z metabolizmem mannitolu, w których dominowały geny ELI3-2, o znanej roli w stresie. Geny charakteryzujące się spadkiem ekspresji, tworzyły, podobnie jak w

przypadku wcześniejszej analizy, grupy GO związane z transportem (w tym transportem tłuszczów), aktywnością peroksydaz, organizacją ściany komórkowej. Ciekawa jest inhibicja kategorii funkcjonalnych opisujących syntezę glikozydów, przede wszystkim syntazy sacharozowej, która ulegała wyraźnej indukcji na mikromacierzy POT w stresie solnym. Doniesienia literaturowe dotyczące ekspresji tego genu w stresie są sprzeczne [198,199].



Rycina 34. Profil ekspresji genów na mikromacierzy POT (2 powtórzenia techniczne) na tle próby referencyjnej na poziomie: A) listy top600, B) wspólnych genów różnicujących dla mikromacierzy badanej i referencyjnej. Wartości referencyjne posortowano rosnąco.

IV.3.4. Opracowanie ścieżki filtracji danych

IV.3.4.1. Filtracja danych na podstawie kryteriów podobieństwa sekwencji nukleotydowej (filtracja HOM)

Filtracja danych polegała na odrzuceniu z analizy sond o niskim podobieństwie do sekwencji *Nicotiana tabacum* i *Nicotiana benthamiana*, ocenionym za pomocą programu BLAST. Progiem odcięcia była wartość $e < 0,001$, będąca ogólnie przyjętą wartością, pozwalającą uznać podobieństwo za statystycznie istotne. Wartość e wyraża się wzorem:

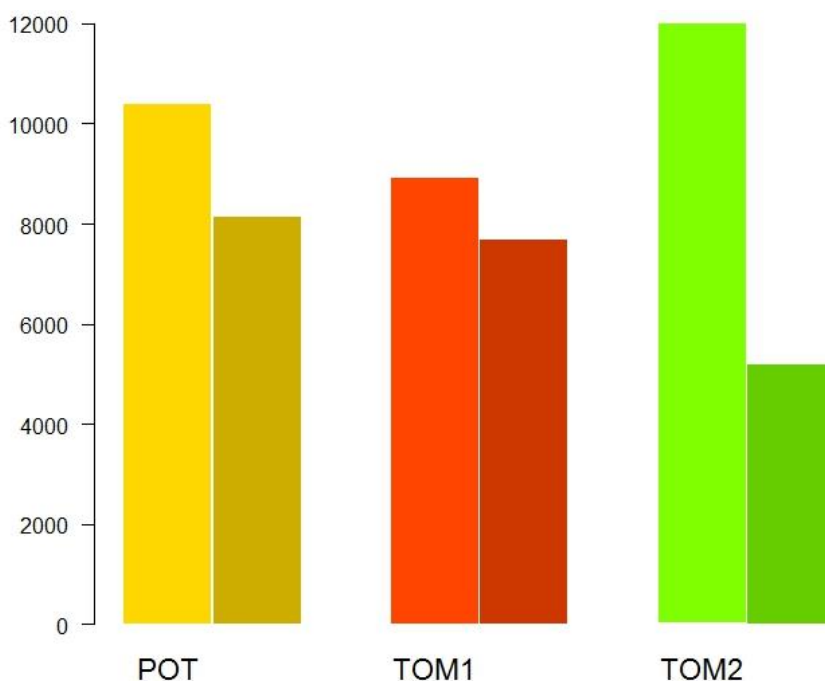
$$e = mn2^{-S}$$

gdzie m – długość analizowanej sekwencji, n – wielkość bazy, S – Bit Score).

Wartość e jest skorelowana z wartością Bit Score (por. I.3.2.3), określającą stopień dopasowania sekwencji, która jest stosowana w dalszych obliczeniach.

Sondy nie spełniające kryterium włączenia do analizy oznaczano w taki sposób, by móc swobodnie usunąć je w dowolnym etapie analizy danych mikromacierzowych przy użyciu pakietu funkcji R. Ostatecznym efektem było więc uzyskanie dla każdej mikromacierzy pliku gal zawierającego dodatkową kolumnę informującą o filtracji sond – sondy pozostawione w analizie oznaczano cyfrą „1”, usuwane z analizy – cyfrą „0”. Informacje dotyczące liczby sond spełniających kryteria filtracji HOM na poszczególnych mikromacierzach zamieszczono w Tabeli 10.

Analiza danych z zastosowaniem filtracji HOM polegała na wyeliminowaniu sond o niewystarczającej homologii przed etapem normalizacji. Inne etapy analizy były wykonane standardowymi metodami. Ponieważ uzyskane wyniki uległy zmianie, na każdym etapie przeanalizowano ponownie wykresy diagnostyczne. Potwierdziły one dobrą jakość danych analizowanych z zastosowaniem filtracji i prawidłowość zastosowanej metodyki.



Rycina 35. Filtracja HOM. Jaśniejsze słupki odzwierciedlają łączną liczbę sond na mikromacierzy, ciemniejsze przedstawiają liczbę sond pozostałych po filtracji HOM.

Po wykonaniu analizy uzyskano listy genów różnicujących, których podsumowanie zamieszczono w Tabelach 23-27 (kol.5-8), tylko nieliczne sondy z listy „top600” wyselekcjonowane bez zastosowania filtracji, a wykazujące podobieństwo do sond tytoniowych nie znalazły się na liście wytypowanej przy użyciu filtracji HOM. W przypadku mikromacierzy TOM1 i POT odnotowano podobną liczbę genów „top600” wytypowanych bez filtracji, które z uwagi na zbyt małe podobieństwo sekwencyjne nie przeszły filtracji HOM - liczba ta wynosiła od 100 do 73 genów w zależności od eksperymentu. Największe rozbieżności pomiędzy grupami genów różnicujących analizowanych z i bez użycia filtracji HOM obserwowano w przypadku mikromacierzy TOM2. Właśnie na tej mikromacierzy, z uwagi na dość znaczną liczbę sond o niewystarczającym podobieństwie sekwencyjnym spodziewano się największego wpływu filtracji. Liczba sond unikatowych dla zastosowanej filtracji wynosiła w przypadku tejże mikromacierzy ok. 300, co stanowiło połowę genów wytypowanych jako różnicujące (Tabela 28). W przypadku mikromacierzy TOM1 i POT zgodność z wynikami uzyskanymi dla TOB nie uległa istotnym zmianom po zastosowaniu filtracji w porównaniu z jej brakiem, co świadczy o stabilności grupy sond wyselekcjonowanych metodą „top600” nawet w przypadku analizy wykonanej z zastosowaniem metodyki typowej dla mikromacierzy homologicznych. W przypadku TOM2 zmniejszeniu uległ odsetek sond, które różnicowały również na mikromacierzy referencyjnej (Tabela 24. kol. 2 i 6) oraz

wartości współczynników korelacji dla sond różnicujących (Tabela 25. kol. 2 i 6), jednak zmiana ta może wynikać raczej z dwukrotnego zwiększenia liczby sond poddanych ocenie (posiadających odpowiedniki na TOB). Powyższe porównania nie wskazują zatem na poprawę wyników spowodowaną zastosowaniem filtracji HOM. Z kolei analiza GO na podstawie sond tytoniowych wskazuje na nieznacznie większą liczbę różnicujących GO zbieżnych z uzyskanymi na mikromacierzy TOB (Tabela 26. kol.5-8) dla mikromacierzy POT w przypadku stosowania 2 powtórzeń technicznych. Na pozostałych mikromacierzach nie odnotowano takiego zjawiska w obu stresach. Wspomniane wyniki dla mikromacierzy POT, w przypadku stresu solnego uwiarygodniły indukcję kilku kategorii GO związanych ze starzeniem. W stresie kadmowym filtracja pozwoliła uzyskać dodatkowe różnicujące grupy genów, związane m.in. z odpowiedzią na bodziec i syntezą acetylokoenzymu A (indukcja) oraz aktywnością transporterów i białek cytoszkieletu (inhibicja).

Analiza GO specyficznych dla danej mikromacierzy mogła w sposób bardziej obiektywny wskazać na znaczenie filtracji HOM w kontekście zgodności z mikromacierzą TOB. Uzyskane w tym przypadku wyniki były dość niejednoznaczne. Spośród ośmiu analizowanych przypadków (mikromacierze TOM1, TOM2 i POT (w 1 i 2 powtórzeniach), w czterech odnotowano zwiększenie liczby różnicujących GO, jednak tylko w jednym przypadku (POT 1 powt. CdCl₂) wzrost dotyczył liczby GO zbieżnych z TOB. W trzech przypadkach stwierdzono spadek liczby GO, a w jednym- brak zmian. Liczba GO fałszywie pozytywnych wzrosła w pięciu przypadkach, w dwóch zmalała a w jednym pozostała niezmienną. W przypadku mikromacierzy, dla których uzyskane dane wydają się być stabilniejsze (POT i TOM1), uzyskano zbliżone wyniki po zastosowaniu filtracji lub jej braku. Z kolei mikromacierz TOM2 wykazywała duże zróżnicowanie wyników analizy GO, co idzie w parze z niską jakością plamek i małą liczbą danych o wysokiej jakości. Trudno jest zatem wskazać, czy filtracja HOM wykazuje pozytywny bądź negatywny wpływ na uzyskane wyniki.

Tabela 28. Zgodność sond różnicujących („top600”) wyselekcjonowanych podczas analizy bez filtracji i z filtracją HOM.

Mikromacierz	POT 1 powt.		POT 2 powt.		TOM1		TOM2	
	NaCl	CdCl ₂	NaCl	CdCl ₂	NaCl	CdCl ₂	NaCl	CdCl ₂
Wspólna pula sond „top600” dla analiz bez filtracji i z filtracją H	500	525	513	529	523	533	281	295

IV.3.4.2.. Filtracja danych na podstawie kryteriów morfologii plamek (filtracja MORF)

W niniejszej pracy podjęto próbę opracowania alternatywnej ścieżki filtracji danych, niezależnej od informacji dotyczącej sekwencji i opierającej się na parametrach morfologicznych plamek. Początkowo wzięto pod uwagę 13 parametrów opisujących morfologię, które były obliczone za pomocą programów Maia i GenePix Pro (opisane w Tabeli 29). W ramach każdej mikromacierzy uśredniono parametry uzyskane dla poszczególnych powtórzeń, otrzymując zbiorczą tabelę zawierającą średnie wartości wszystkich parametrów dla każdej sondy na mikromacierzy. Uzyskane wartości charakterystyki morfologicznej plamek wykorzystano w dalszych obliczeniach.

Celem ustalenia, czy którekolwiek z parametrów wykazuje zależność od wartości Bit Score i ewentualnie może zastąpić ją w ramach filtracji danych przeprowadzono badania korelacji liniowej (Pearsona) tych parametrów z wartościami Bit Score. Uzyskane dane zamieszczono w Tabeli 30. Wartości współczynnika korelacji wskazywały na prawie całkowity brak zależności charakterystyki punktów od wartości Bit Score na mikromacierzach TOM1 i TOM2 (największa wartość współczynnika 0,12) oraz jedynie nieznaczną korelację ($>0,2$) w przypadku mikromacierzy POT dla 7 parametrów (Tabela 31). Wśród nich znalazło się 5 parametrów wyznaczonych za pomocą programu Maia (CVR, Det, Dia, GSy, Sig) oraz dwa obliczone z wykorzystaniem GenePix Pro (F(R) i RgnR2). Wielkości opisujące współczynnik determinacji regresji liniowej pomiędzy intensywnością fluorescencji na kanałach Cy3 i Cy5 (Det i RgnR2) oraz intensywności sygnału (Sig i R) były odpowiadającymi sobie wielkościami obliczanymi przy wykorzystaniu dwu różnych programów. W przypadku czterech parametrów (CVR, Det, Dia, GSy) istniały dane literaturowe sugerujące zależność wobec Bit Score, nie była to jednak ścisła korelacja liniowa [52]. Badanie wzajemnych zależności pomiędzy wszystkimi wyróżnionymi parametrami (Tabela 31) wskazuje na dużą wzajemną korelację (od 0,62 do 0,94) wartości 4 parametrów wyróżnionych przez Bar-Or i wsp. [52], a także RgnR2, powtarzalną na wszystkich mikromacierzach. Duża korelacja pomiędzy tymi parametrami sugeruje możliwość wykorzystania nawet jednego, wybranego parametru do wyselekcjonowania danych o lepszej jakości hybrydyzacji. Z kolei oba parametry opisujące intensywność sygnału nie wykazują tak znacznej korelacji z wymienionymi wyżej pięcioma, dlatego nie wzięto ich pod uwagę w dalszych obliczeniach.

Tabela 29. Parametry jakości plamek wzięte pod uwagę podczas ustalania alternatywnej ścieżki filtracji danych.

Skrót	Nazwa parametru	Opis	P/ LK*	Zakres i interpretacja wartości
ABac	Bezwzględna wartość tła	Porównanie fluorescencji tła ze średnią wartością tła dla całej mikromacierzy	M/2	
Con	Zanieczyszczenie plamki	Liczba nieprawidłowych pikseli w plamce	M/2	
CVR	Współczynnik zmienności dwóch stosunków intensywności	Współczynnik zmienności pomiędzy wartościami współczynników RR (współczynnik regresji) i RS (współczynnik segmentacji) opisujących zależności pomiędzy kanałami Cy3 i Cy5.	M/2	
Det	Współczynnik determinacji (2)	Współczynnik determinacji regresji liniowej pomiędzy intensywnością kanałów Cy3 i Cy5	M/2	
Dia	Średnica plamki	Średnica plamki: $D=2(S/\pi)^{1/2}$, S oznacza liczbę pikseli w obrębie plamki	M/1	
Dur	Statystyka Durбина-Watsona	Miara korelacji wartości odstających w rozkładzie liniowym intensywności plamki	M/2	
GSy	Symetria geometryczna	Odchylenie geometrii plamki od idealnego okręgu	M/1	
ISy	Symetria rozkładu intensywności	Określa symetrię intensywności fluorescencji w obrębie plamki	M/2	
F (R,G)	Intensywność sygnału	Średnia intensywność sygnału fluorescencji dla wszystkich pikseli danej plamki	GP/1	
B (Rb, Gb)	Intensywność sygnału tła	Intensywność sygnału tła otaczającego dany punkt-dla próbek nietraktowanych tytoniu	GP/1	
RBac	Jednorodność tła	Jednorodność tła w obrębie otoczenia plamki	M/2	
RgnR2	Współczynnik determinacji (1)	Współczynnik determinacji regresji liniowej pomiędzy intensywnością kanałów Cy3 i Cy5	GP/2	
Sig	Intensywność sygnału	$S=\min(S_{Cy5}-B_{Cy5}, S_{Cy3}-B_{Cy3})$, gdzie $S_{Cy5}(S_{Cy3})$ jest średnią intensywności w obrębie konturu danego punktu dla kanału Cy5(Cy3), $B_{Cy5}(B_{Cy3})$ jest średnią szacunkową tła dla kanału Cy5(Cy3).	M/2	
SNR	Stosunek sygnału do szumu	Stosunek średniej intensywności fluorescencji plamki do średniej intensywności fluorescencji tła (ang. "Signal-to-noise ratio")	GP/1	

*P=program (GP-GenePix Pro; M-Maia); LK=liczba kanałów wymaganych dla odczytania parametru; intensywniejszy kolor gwiazdek oznacza lepszą jakość plamki

Tabela 30. Korelacja (obliczona metodą Pearsona) pomiędzy średnimi wartościami wybranych parametrów jakościowych plamek a parametrem Bit Score uzyskana w eksperymentach CSH.

	POT	p*	TOM1	p*	TOM2	p*
ABac	0,21	#	0,06	-	0,03	##
Con	0,08	##	0,03	-	-0,02	-
CVR	-0,22	#	-0,08	##	-0,13	#
Det	0,26	#	0,11	##	0,14	#
Dia	0,23	#	0,09	##	0,15	#
Dur	0,00	##	-0,06	##	-0,02	##
GSy	-0,23	#	-0,07	##	-0,14	#
ISy	-0,07	#	0,07	##	-0,02	-
F (R)	0,23	#	0,10	#	0,02	-
B (Rb)	0,03	##	0,03	##	0,02	-
RBac	0,13	##	0,05	#	0,02	-
RgnR2	0,28	#	0,10	#	0,12	#
Sig	0,26	#	0,10	#	0,03	##
SNR_	0,02	##	0,10	#	0,04	##

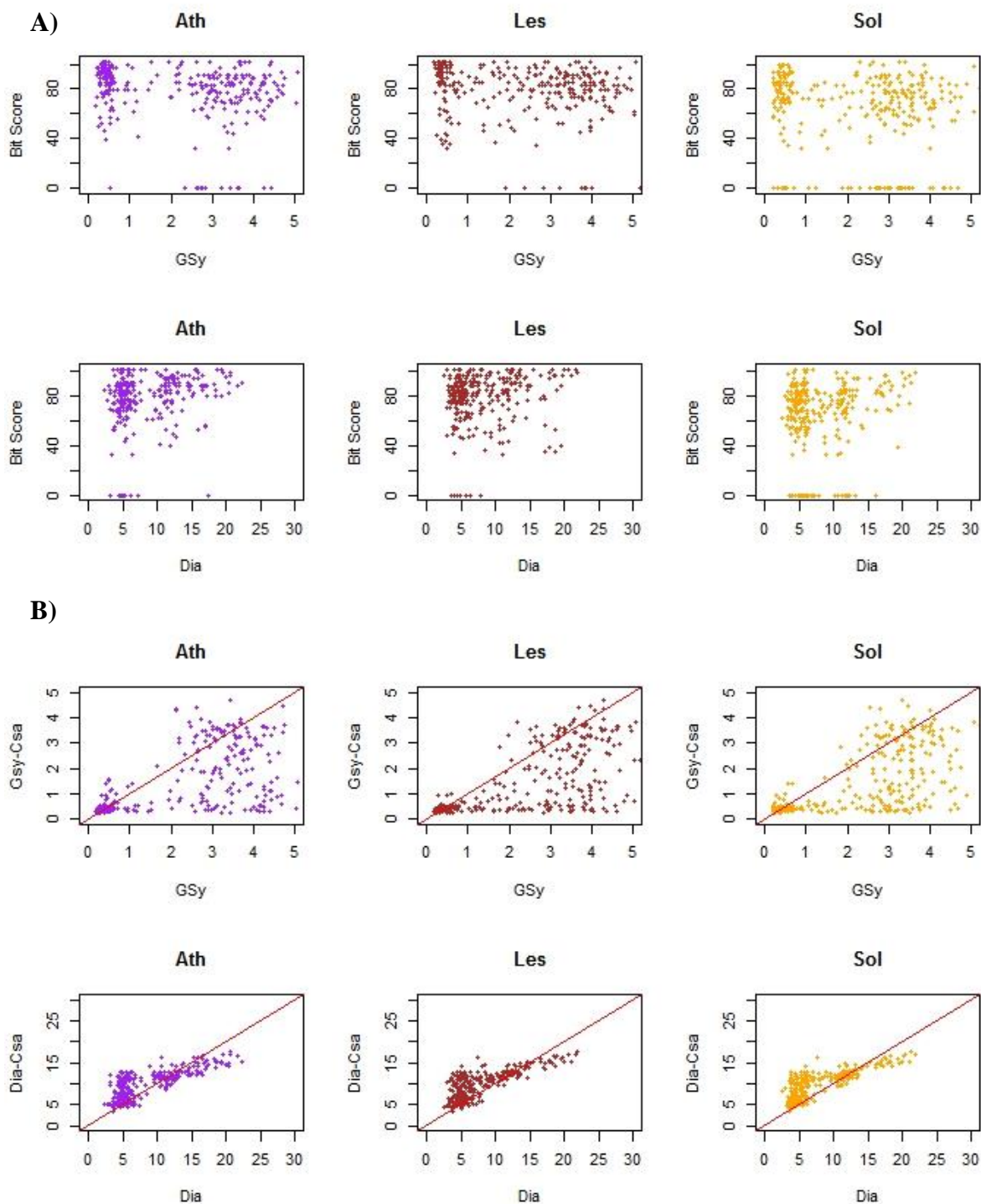
*wartości p mieściły się w wymienionych zakresach oznaczonych symbolami:

$p < 2,2e-16$
 ## $2,2e-16 < p < 0,05$;
 - $p > 0,05$

Tabela 31. Macierz korelacji (Pearsona) pomiędzy parametrami jakości plamek wytypowanymi jako potencjalnie użyteczne do filtracji MORF. Kolorem czerwonym oznaczono współczynniki korelacji o wartości $> 0,5$.

POT								TOM1							
	CVR	Det	Dia	GSy	Sig	R	Rgn R2		CVR	Det	Dia	GSy	Sig	R	Rgn R2
CVR		-0,9	-0,9	0,9	-0,5	-0,5	-0,8	CVR		-0,9	-0,9	0,8	-0,5	-0,5	-0,8
Det	-0,9		0,7	-0,8	0,7	0,6	1,0	Det	-0,9		0,9	-0,8	0,7	0,7	1,0
Dia	-0,9	0,7		-0,9	0,4	0,4	0,6	Dia	-0,9	0,9		-0,8	0,6	0,6	0,9
GSy	0,9	-0,8	-0,9		-0,4	-0,4	-0,6	GSy	0,8	-0,8	-0,8		-0,5	-0,5	-0,8
Sig	-0,5	0,7	0,4	-0,4		1,0	0,8	Sig	-0,5	0,7	0,6	-0,5		1,0	0,8
F (R)	-0,5	0,6	0,4	-0,4	1,0		0,7	F (R)	-0,5	0,7	0,6	-0,5	1,0		0,8
Rgn R2	-0,8	1,0	0,6	-0,6	0,8	0,7		Rgn R2	-0,8	1,0	0,9	-0,8	0,8	0,8	
TOM2															
	CVR	Det	Dia	GSy	Sig	R	Rgn R2								
CVR		-0,9	-0,8	0,6	-0,3	-0,3	-0,6								
Det	-0,9		0,9	-0,8	0,5	0,4	0,8								
Dia	-0,8	0,9		-0,9	0,4	0,4	0,8								
GSy	0,6	-0,8	-0,9		-0,4	-0,3	-0,7								
Sig	-0,3	0,5	0,4	-0,4		0,9	0,7								
F (R)	-0,3	0,4	0,4	-0,3	0,9		0,6								
Rgn R2	-0,6	0,8	0,8	-0,7	0,7	0,6									

Dotychczas analizowany zestaw danych nie dawał możliwości bezpośredniego skonfrontowania wyników CSH z wynikami SSH na tej samej mikromacierzy, co mogłoby ostatecznie zweryfikować istnienie zależności pomiędzy podobieństwem sekwencji DNA, a jakością plamek. W tym przypadku pomocne okazały się dane uzyskane w naszej pracowni przy pomocy wspomnianej wcześniej (I.2.3) oligonukleotydowej mikromacierzy chloroplastowej dla ogórka, PlasTi-array [91]. W badaniu stosowano hybrydyzację CSH trzech gatunków roślin: szpinaku, rzodkiewnika i pomidora do mikromacierzy dla ogórka, przy czym strategia stosowana w doświadczeniu polegała na znakowaniu jednym z barwników próby dla obcego gatunku, drugim zaś- dla ogórka. Każdy eksperyment wykonano w 2 powtórzeniach ze zamianą barwników, dodatkowo każda sonda była dwukrotnie powtórzona na mikromacierzy. Dane te stanowiły cenny materiał w kontekście oceny wpływu homologii sekwencji na wyniki CSH, ponieważ dostępne były pełne informacje o sekwencjach genomów plastydowych wszystkich badanych gatunków. Dzięki temu istniała możliwość porównania parametrów jakościowych plamek uzyskanych dla transkryptów CSH i SSH na tej samej mikromacierzy i ostatecznej weryfikacji wpływu podobieństwa sekwencji na ich wartości. Spośród stosowanych sond wybrano takie, które odpowiadały kodującym sekwencjom chloroplastowym, a także we wszystkich wykonanych eksperymentach emitowały sygnał fluorescencyjny po hybrydyzacji z chloroplastowym RNA ogórka, świadczący o ekspresji transkryptu docelowego. W ten sposób wyselekcjonowano pulę 280 sond, które były komplementarne do genów ulegających ekspresji w ogórku (na podstawie stosunku sygnału do szumu ($SNR > 3$)). Dane uzyskane dla tych sond stanowiły wartości referencyjne dla danych CSH. Uzyskano średnie wartości parametrów jakości, które obejmowały dwa zestawy (CSH i SSH) dla każdej sondy w ramach poszczególnych gatunków i zebrano dane dotyczące stopnia komplementarności każdej sondy do sekwencji genomu chloroplastowego danego gatunku. Z uwagi na sposób przeprowadzenia eksperymentu mikromacierzowego, ograniczeniem dla tego modelu była możliwość wykorzystania tylko takich parametrów, które mogły być obliczone dla eksperymentów jednokolorowych (patrz Tabela 29). Spośród wyselekcjonowanych parametrów były to Dia i GSy. Porównanie tych wielkości z Bit Score na wykresach (Rycina 36A) wskazuje na brak zależności liniowej pomiędzy danymi dotyczącymi homologii, a parametrami morfologicznymi. Niemniej jednak porównanie wartości Dia i GSy uzyskanych jednocześnie dla tych samych sond podczas hybrydyzacji CSH i SSH (Rycina 36B) wskazuje ewidentnie na korzystniejsze wyniki uzyskiwane dla SSH (mniejsze GSy, mniejsza rozpiętość Dia).



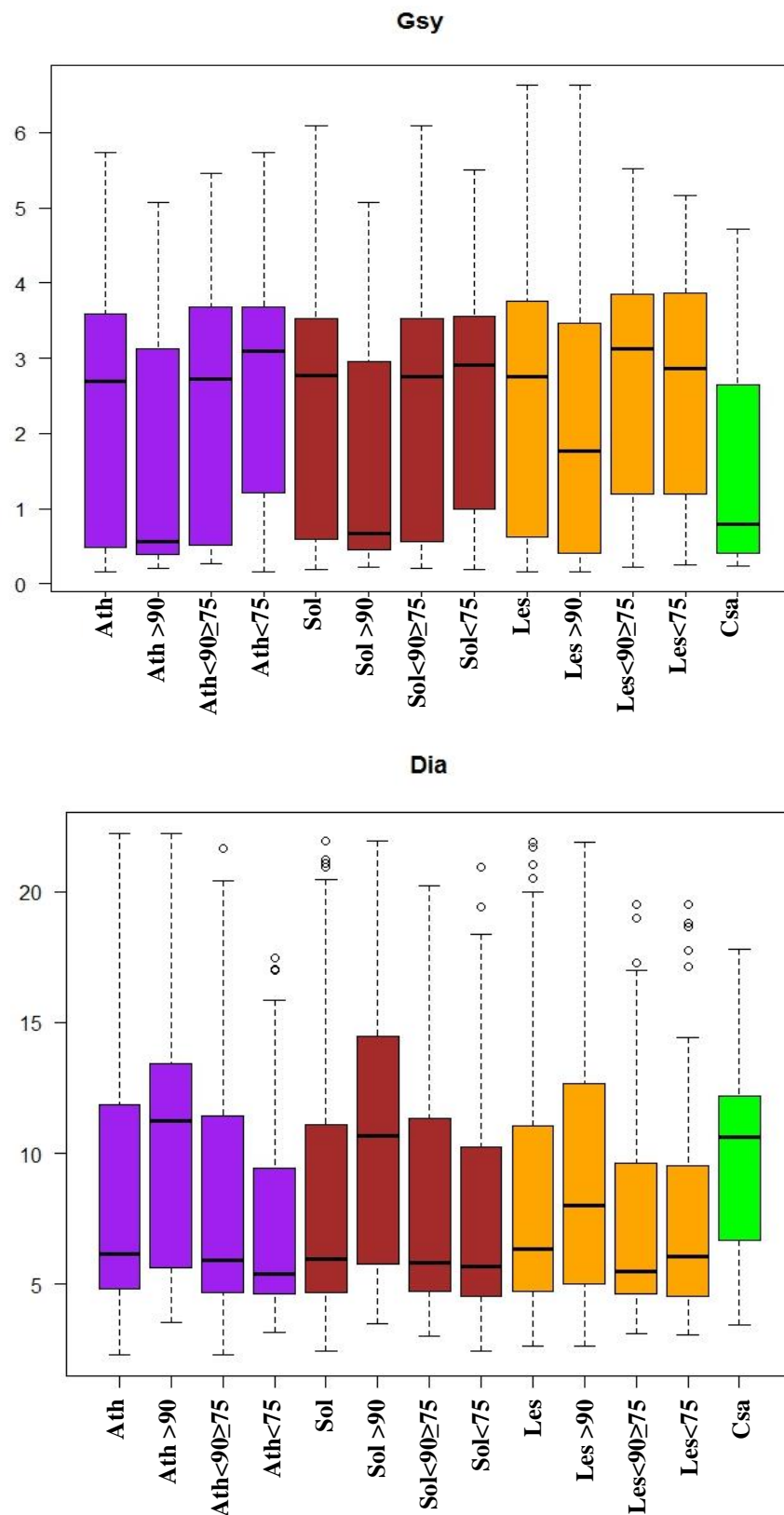
Rycina 36. Rozkład parametrów morfologicznych uzyskanych podczas eksperymentów CSH na mikromacierzy dla ogórka (Ath-rzodkiewnik, Sol-szpinak, Les-pomidor, Csa-ogórek)

A) w zależności od wartości Bit Score

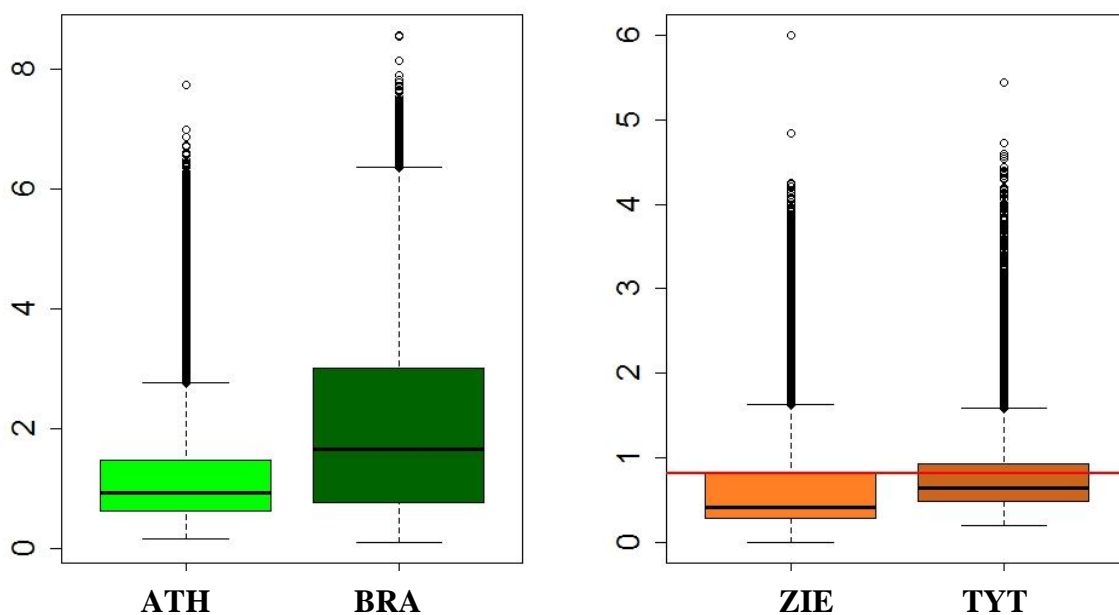
B) w porównaniu z wartościami uzyskanymi dla tych samych sond podczas hybrydyzacji SSH

Na wykresach pudełkowych prezentujących rozkład parametrów GSy i Dia dla plamek o różnych zakresach Bit Score, we wszystkich eksperymentach CSH widoczna jest gorsza jakość plamek dla zbiorów o mniejszym podobieństwie sekwencji (Rycina 37.). Na wykresach rozkładu wartości GSy i Dia dla danych CSH widać wyraźnie lepszą jakość danych wyłącznie dla sekwencji o Bit-Score > 90 i tylko w tym zakresie rozkład jest bliższy danym, uzyskanym w przypadku SSH. W przypadku wyselekcjonowania sekwencji o największych wartościach Bit-Score (>90) obserwujemy większe wartości parametrów jakościowych plamek. Tendencja ta jest widoczna dla obu parametrów jakości w przypadku wszystkich trzech gatunków CSH, przy czym nieco mniejszą, choć nadal wyraźną poprawę wyników odnotowano dla pomidora. Ponadto, efekt ten jest lepiej widoczny dla GSy, ponieważ posiada on zdefiniowaną wartość optymalną (0), w przypadku parametru Dia można jedynie wytypować orientacyjny zakres prawidłowych wartości. Wyniki eksperymentu na mikromacierzy dla ogórka charakteryzują się lepszymi parametrami jakościowymi dla sond o większym podobieństwie sekwencji, przy czym różnice mają charakter skokowy. Wskazuje to na możliwość wykorzystania parametrów morfologicznych do filtracji danych pomimo wspomnianego braku zależności liniowej. Punktem odniesienia dla ustalenia progu odcięcia pozwalającego przeprowadzić filtrację może być rozkład danych uzyskanych w eksperymentach SSH dla tej samej mikromacierzy. W tym przypadku dla wartości GSy próg odcięcia mogłaby stanowić wartość graniczna trzeciego kwartylu rozkładu uzyskanego dla hybrydyzacji SSH ogórka, prezentowana na Rycinie 37 (2,64).

Powyższe wyniki wskazywały na możliwość wykorzystania parametrów jakości uzyskiwanych dla danej mikromacierzy w ramach hybrydyzacji SSH do ustalenia progu odcięcia dla filtracji MORF podczas analizy eksperymentów CSH. Aby zbadać, czy takie założenia mają zastosowanie dla innych mikromacierzy, porównano rozkłady parametrów SSH i CSH dla innego zestawu danych wykonywanych w naszej pracowni: badań transkryptomu rzodkiewnika (SSH) i kapusty (*Brassica oleracea*) (CSH) wykonywanych na tej samej mikromacierzy oligonukleotydowej zawierającej sondy dla rzodkiewnika. W przypadku parametru GSy parametry jakościowe dla hybrydyzacji CSH są również znacznie gorsze. (Rycina 38). Wartość trzeciego kwartylu danych SSH, mogąca stanowić potencjalny próg odcięcia wynosiła ok. 1,48, co w tym wypadku oznaczałoby odrzucenie ok. 50% sond w przypadku badań transkryptomu kapusty.



Rycina 37. Rozkład parametrów jakości GSy i Dia podczas hybrydizacji CSH (Ath-rzodkiewnik, Sol- szpinak, Les-pomidor) i SSH (Csa-ogórek) uzyskanych dla sond w różnych przedziałach podobieństwa do sekwencji docelowej (wyrażonego w postaci Bit Score) - kolejno dla każdej rośliny CSH: wszystkie badane sondy, sondy o Bit-score >90; sondy o Bit-score <90 i ≥75, sondy o Bit-score <75..



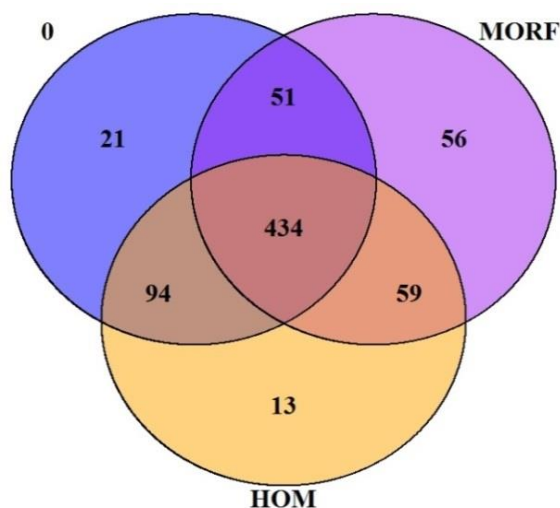
Rycina 38. Rozkład wartości GSy dla hybrydyzacji SSH i CSH a) na mikromacierzy oligo dla rzodkiewnika (ATH-SSH rzodkiewnik; BRA-CSH kapusta); b) na mikromacierzy POT (ZIE-SSH ziemniak, TYT- CSH tytoń), czerwoną linią oznaczono wytypowany próg odcięcia.

Po potwierdzeniu opisanej zależności w dwóch niezależnych układach zdecydowano się na ustalenie progu filtracji MORF w analogiczny sposób dla niniejszych badań transkryptomu tytoniu. W tym celu niezbędne było uzyskanie danych dotyczących charakterystyki jakości plamek w układzie SSH dla badanych mikromacierzy, analizowanych za pomocą programu Maia. Takie dane dla mikromacierzy POT dostępne były w załączniku do omawianej wcześniej publikacji Bar-Or i wsp.[52] w której hybrydyzacji SSH poddawano cDNA z roślin ziemniaka poddanych działaniu stresu solnego (6 powtórzeń). Porównanie rozkładu wartości GSy z hybrydyzacji SSH i CSH na mikromacierzy POT przedstawiono na wykresach pudełkowych (Rycina 38). Wyznaczono próg filtracji MORF na mikromacierzy POT korzystając z wartości trzeciego kwartylu dla hybrydyzacji SSH ziemniaka: 0,82. Używając tej wartości odcięcia przeprowadzono filtrację MORF na mikromacierzy POT i analizę danych, analogicznie do wcześniej wykonywanych analiz. W przypadku mikromacierzy TOM1 i TOM2 nie posiadano danych referencyjnych SSH, dlatego nie wykonano filtracji MORF na tych mikromacierzach.

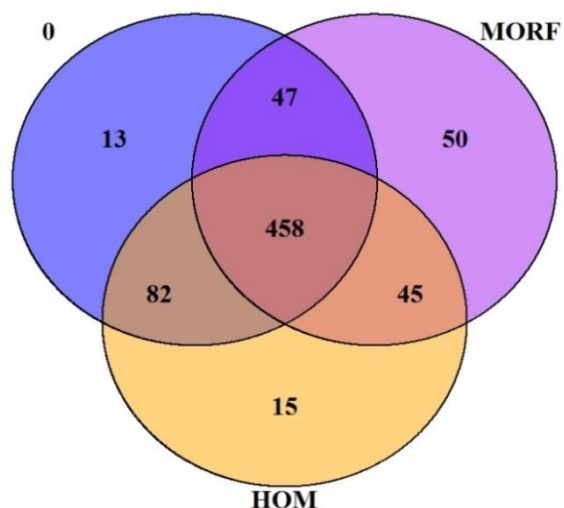
W wyniku zastosowania filtracji MORF z mikromacierzy POT usunięto ok. 32% sond (Tabela 23), a więc ok. dwukrotnie więcej niż w przypadku filtracji HOM. Podsumowanie statystyczne wyników zamieszczono w Tabelach 23-27. Mimo odrzucenia

większej liczby danych, ilość sond uzyskujących wartość $p < 0,05$ była podobna do otrzymanej przy braku filtracji oraz filtracji HOM. Wartości korelacji obrazujące zbieżność profilu ekspresji były na wszystkich ocenianych poziomach zbliżone do uzyskiwanych przy braku filtracji i filtracji HOM. Na Rycinie 39 zamieszczono porównanie zbieżności list „top600” uzyskiwanych przy braku filtracji, filtracji HOM i MORF na mikromacierzy POT w obu eksperymentach, przy zastosowaniu 1 i 2 powtórzeń technicznych. Wskazują one na znaczną zgodność wyników uzyskiwanych przy różnych filtracjach. Największą zbieżność (co najmniej 80 % sond powtarzających się we wszystkich 3 sposobach filtracji) uzyskano w przypadku zwiększenia liczby powtórzeń technicznych. Mimo tak niewielkich różnic w puli sond i podobnych wyników zbieżności dotyczących logFC pojedynczych sond jak w przypadku stosowania filtracji HOM lub jej braku, (Tabela 23-25 kol.11 i 12), wyniki analiz GO (Tabela 26,27, kol 11 i 12) wskazują na słabszą jakość danych filtrowanych przy użyciu morfologii plamek, niż danych analizowanych bez użycia filtracji lub z filtracją HOM. Wyjątek stanowi analiza z wykorzystaniem GO ziemniaka w przypadku korzeni traktowanych $CdCl_2$. W tym przypadku dane tylko nieznacznie odbiegały od uzyskanych przy pomocy innych filtracji, a nawet stwierdzono zmniejszenie liczby GO wykazujących brak zgodności z mikromacierzą referencyjną. Niemniej jednak, trudno wskazać na skuteczność tej metody filtracji, a tym bardziej na możliwość zastosowania jej jako alternatywnej ścieżki analizy danych w sytuacji braku możliwości skorzystania z badań podobieństwa sekwencji.

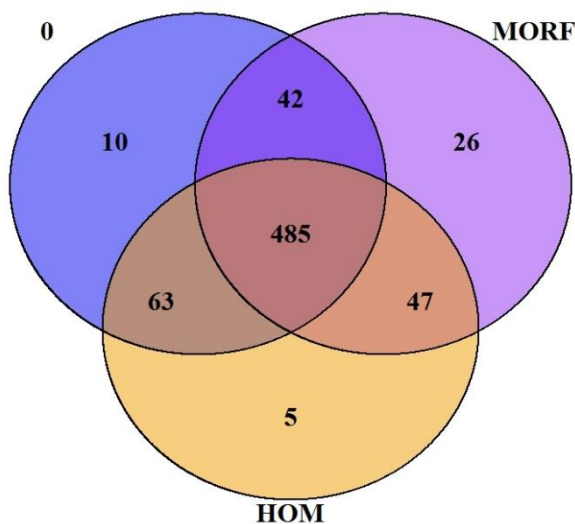
A) POT, NaCl, 1 powt.



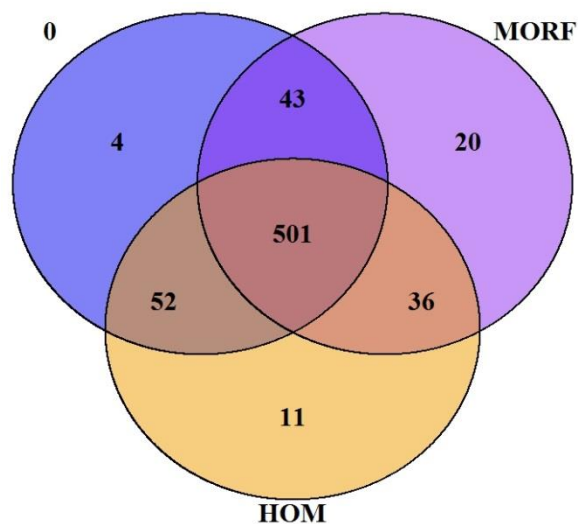
B) POT, CdCl₂, 1 powt.



C) POT, NaCl, 2 powt.



D) POT, CdCl₂, 2 powt.



Rycina 39. Zgodność sond różnicujących podczas hybrydyzacji CSH tytoniu na mikromacierzy POT przy braku filtracji (0) oraz zastosowaniu filtracji HOM i MORF.

V.DYSKUSJA

Mimo szybkiego rozwoju technik sekwencjonowania, mikromacierze w dalszym ciągu pozostają istotnym narzędziem analizy ekspresji genów, ze względu na dobrze opracowaną metodykę analizy, dokładnie określone czynniki zakłócające, dobrą jakość eksperymentów i niższą cenę. Olbrzymie ilości danych uzyskanych w badaniach mikromacierzowych i zdeponowanych w bazach, są nieustannie eksplorowanym źródłem cennych informacji biologicznych. Jednocześnie, CSH jest eksperymentem mikromacierzowym, którego analiza w dalszym ciągu budzi wiele wątpliwości. Dotychczasowe, bardzo nieliczne badania na temat CSH nie dawały przekonującej odpowiedzi na pytania o zasadność stosowania tej metody. Część danych sugerowała, że zarówno stopień podobieństwa sekwencji gatunku badanego i reprezentowanego na mikromacierzy, jak i długość zastosowanych sond mają znaczenie dla uzyskania wyników istotnych biologicznie. Do tej pory nie przeprowadzono jednak kompleksowych badań oceniających, w jakim stopniu czynniki te wpływają na jakość wyników analiz mikromacierzowych. Z drugiej strony, w literaturze można znaleźć szereg przykładów zastosowania metody CSH w badaniach podstawowych, bez uwzględnienia w jakimkolwiek stopniu ich specyfiki [57,90]. Opracowanie właściwej metodyki eksperymentów mogłoby się zatem przyczynić do popularyzacji i rozwinięcia nowych zastosowań metody CSH oraz umożliwić weryfikację poprawności zgromadzonych już danych eksperymentalnych.

W niniejszej pracy zastosowano szereg układów badawczych, których porównanie umożliwiło określenie zbioru reguł dotyczących wykonywania eksperymentów SCH. Jako model badawczy wybrano rodzinę *Solanaceae*, z uwagi na:

- a) wysoki stopień podobieństwa pomiędzy genami roślin należących do tej rodziny
- b) dostępność różnych typów mikromacierzy dla kilku gatunków roślin
- c) znaczenie gospodarcze wielu roślin psiankowatych
- d) znaczną liczbę publikacji dotyczących zastosowania CSH w tej grupie roślin
- e) dostępność stosunkowo dużej liczby sekwencji cDNA tytoniu, ziemniaka i pomidora w bazach danych, co umożliwiło przeprowadzenie badań homologii poszczególnych sond mikromacierzowych i sekwencji transkryptów, które potencjalnie mogły wiązać się z tymi sondami.

Dzięki zastosowaniu różnych typów mikromacierzy (oligonukleotydowe, cDNA) oraz układów hybrydyzacji międzygatunkowej (tytoń: ziemniak; tytoń: pomidor), możliwe było systematyczne zbadanie wpływu poszczególnych czynników związanych z hybrydyzacją CSH na wynik i biologiczną interpretację wyników. Ponadto, w celu

dokładniejszej weryfikacji zgodności wyników CSH z wynikami referencyjnymi (SSH), zdecydowano się na zastosowanie dwóch biologicznie niezależnych układów eksperymentalnych (traktowanie NaCl i CdCl₂). Taki obiekt badań dawał możliwość uzyskania nieopisanych do tej pory w literaturze danych biologicznych, związanych z aktualnymi zainteresowaniami badawczymi naszej pracowni oraz prowadzoną współpracą z Laboratorium Biochemii Roślin Instytutu Biochemii i Biofizyki Polskiej Akademii Nauk. Badania wpływu stresu solnego oraz metali ciężkich były dość często wykonywane na roślinach modelowych, jak rzodkiewnik i ryż [95-98,149,189], natomiast tylko nieliczne prace genomyczne dotyczyły rodziny *Solanaceae* [142], zaś mechanizmy odpowiedzi na stres w samym tytoniu, uważanym przecież za roślinę modelową, były do tej pory dość słabo poznane. Czynnikiem podnoszącym potencjalne znaczenie uzyskanych wyników jest to, że dotyczą one transkryptomu korzeni, a więc części rośliny, która często jako pierwsza kontaktuje się z czynnikami stresogennymi [95].

V.1. JAKOŚĆ UZYSKANYCH DANYCH

Analizy diagnostyczne uzyskanych wyników pokazały, że dla każdej z mikromacierzy w układzie CSH możliwe było zoptymalizowanie warunków hybrydyzacji i analizy tak, aby uzyskać dane o dobrej jakości. Poszczególne mikromacierze różniły się jednak zdecydowanie pod względem liczby genów różnicujących przy wartości $p < 0,05$. Z pewnością wpływ na taki stan rzeczy przynajmniej w części mogły mieć różnice techniczne w przebiegu eksperymentów. Te starano się jednak zminimalizować poprzez ujednolicenie, tam gdzie to było możliwe, procedur eksperymentalnych oraz toku analizy bioinformatycznej. Można zatem założyć, że obserwowane różnice odzwierciedlają również różnice w przydatności poszczególnych mikromacierzy, wynikające z ich rodzaju i/lub różnic pomiędzy gatunkami reprezentowanymi na tych mikromacierzach. Tezę tę może wspierać fakt, że dane o najwyższej jakości uzyskano przy wykorzystaniu mikromacierzy referencyjnej. O ich wysokiej jakości, oprócz samych wykresów diagnostycznych, niezbiemie świadczą uzyskane wartości ekspresji różnicowej dla sond kontrolnych. Przy użyciu tej mikromacierzy uzyskano zdecydowanie największą liczbę genów różnicujących (wartość $p < 0,05$). Taki stan rzeczy był spodziewany nie tylko z uwagi na użycie mikromacierzy SSH, ale także ze względu na zastosowanie powtórzeń technicznych oraz fakt, iż mikromacierz TOB została wykonana z zastosowaniem nowoczesnej techniki polegającej na syntezie sond *in situ*.

V.2. PROFIL EKSPRESJI GENÓW W KORZENIACH TYTONIU PODDANYCH DZIAŁANIU STRESU

V.2.1. Ogólny profil ekspresji genów w korzeniach tytoniu pod wpływem stresu abiotycznego

Uzyskane wyniki pozwoliły na opisanie profili ekspresji genów w korzeniach tytoniu w stresie abiotycznym. Na podstawie wyników eksperymentów mikromacierzowych wyselekcjonowano liczną grupę genów, które zachowywały się w podobny sposób w przypadku obu wybranych stresów i mogą być zaangażowane w ogólną odpowiedź roślin na stres abiotyczny, wspólną co najmniej dla obu czynników. Wielu autorów potwierdza obecność wspólnych profili transkryptomicznych w ramach odpowiedzi na różnego rodzaju bodźce [100, 159-161, 188]. Mimo iż wspólna odpowiedź na stres abiotyczny jest dość częstym tematem badań, do tej pory opublikowano tylko jedną pracę z zakresu transkryptomiki, w której bezpośrednio porównywano działanie stresu solnego i kadmowego na korzenie rzodkiewnika [100].

Na podstawie profili GO i genów różnicujących zdefiniowano najważniejsze kategorie ogólnej odpowiedzi na stres abiotyczny: odpowiedź na bodźce i procesy immunologiczne, katabolizm toksyn, transport, śmierć komórki, wiązanie wapnia, biosynteza hormonów (geny ulegające indukcji) oraz transport tłuszczów, procesy zachodzące w ścianie komórkowej, aktywność kanałów wodnych (geny ulegające zahamowaniu) (patrz Tabela 17).

Wśród genów ulegających indukcji, najliczniej reprezentowana jest grupa tych związanych z szeroko ujętą odpowiedzią na bodźce. Aktywność genów typowych dla reakcji na różnorodne bodźce świadczy o złożoności odpowiedzi na oba rodzaje traktowania, w której przeplata się wiele ścieżek sygnałowych. Wskazuje to również na istnienie wspólnych etapów cechujących odpowiedź na różne rodzaje stresu abiotycznego, a nawet biotycznego, co potwierdzają inni autorzy [159-161]. Uruchomienie ścieżki sygnalizacji za pośrednictwem ABA w przypadku korzeni różnych roślin poddanych działaniu stresu solnego lub kadmowego zaobserwowano również w innych pracach transkryptomicznych [95,99,100]. Podobne spostrzeżenia dotyczą działania etylenu [99, 178], chociaż w tym przypadku w stresie solnym odnotowano bardziej znaczącą indukcję genów i GO związanych z sygnalizacją za pomocą tego hormonu, a w stresie kadmowym wyraźniejszy był wzrost ekspresji genów odpowiedzialnych za jego syntezę. W przypadku

kaskady sygnałów zależnych od kwasu jasmonowego i salicylowego, napotkano doniesienia dotyczące przede wszystkim indukcji pod wpływem stresu solnego [95,99], chociaż w pracy oceniającej działanie obu stresów stwierdzono indukcję niektórych genów biorących udział w transdukcji sygnału za pośrednictwem JA (np. JAZ5).

W przypadku obu stresów, do genów o największym istotnym wzroście ekspresji należał gen S-transferazy glutationowej odgrywający rolę w detoksyfikacji i będący markerem stresu oksydacyjnego [122]. Wykazywał on wzrost ekspresji również w badaniu bezpośrednio porównującym oba stresy, wraz z peroksydazą glutationową, dla której podobne wyniki odnotowano w ramach niniejszej pracy [100]. Te obserwacje są zgodne z wynikami wcześniejszych badań transkryptomów korzeni roślin poddanych działaniu obu badanych czynników [95,100,189]. Wyraźnie zatem widać, że odpowiedź na oba rodzaje stresu w korzeniach tytoniu związana jest z uruchomieniem podobnych mechanizmów detoksyfikacji, opartych na działaniu glutationu. We wspomnianym doświadczeniu na korzeniach rzodkiewnika stwierdzono znacznie mniejszą liczbę genów różnicujących wspólnych dla obu stresów (kilkadziesiąt), niemniej jednak można wśród nich wyróżnić szereg takich, których indukcję odnotowano również w niniejszej pracy dla obu traktowań: ERF/AP2, WRKY33, HSP, ATAF, CYP, wspomniane GST i JAZ. Są to geny typowe dla odpowiedzi na stres u roślin [100]. W doświadczeniu dla rzodkiewnika odnotowano podobnie jak w przypadku naszego eksperymentu wspólną dla różnych stresów abiotycznych indukcję genów związanych z wiązaniem wapnia (białka CML –podobne do kalmoduliny). W korzeniach tytoniu mamy więc do czynienia z typową dla stresu sygnalizacją za pośrednictwem jonów wapnia.

Geny różnicujące związane z transportem (transportery ABC, MATE) ulegały również indukcji w niezależnych doświadczeniach na korzeniach rzodkiewnika poddanych działaniu stresu solnego [95] i kadmowego [149].

Analizując spadek ekspresji genów pod wpływem obu warunków stresowych, warto zwrócić uwagę na inne grupy genów związanych z transportem. Były to białka błonowe odpowiedzialne za transport tłuszczów, których zahamowanie ekspresji (geny LTP) odnotowano również w korzeniach rzodkiewnika pod wpływem NaCl [95], a także akwaporyny wchodzące w skład tonoplastu i błony komórkowej, których inhibicja jest zjawiskiem typowym dla przeciwdziałania odwodnieniu w wyniku stresu solnego [95, 143], jednak była spotykana również pod wpływem działania kadmu [190]. Ostatnią grupę cechującą się zmniejszoną ekspresją stanowią geny związane ze ścianą komórkową. Inhibicję ekspresji genów zaangażowanych w rozluźnienie ściany komórkowej (XTH,

arabinogalaktany) obserwowano w korzeniach pomidora poddanych działaniu NaCl [142], z kolei w korzeniach rzodkiewnika obserwowano inhibicję enzymów związanych z lignifikacją, indukcję genów XTH oraz ekspansyn. Wzrost ekspresji ekspansyn odnotowano w naszym doświadczeniu w obu grupach, a w przypadku stresu solnego widoczna była nawet indukcja kategorii GO „rozluźnienie ściany komórkowej”. Obserwujemy tu zatem sprzeczność z uprzednio dokonanymi spostrzeżeniami dotyczącymi inhibicji procesu rozluźnienia ściany komórkowej. Ekspansyny biorą udział w procesach rozluźnienia ściany komórkowej i ich ekspresja ulegała zahamowaniu m.in. pod wpływem stresu solnego w pędach kukurydzy [165]. W przypadku korzeni ekspansyny mogą stymulować ich wzrost w wyniku braku wody [166], a także brać udział w wykształcaniu odporności [167]. Jest to zróżnicowana grupa białek o różnych funkcjach, która może wykazywać odmienną ekspresję w zależności od lokalizacji w organie i warunków [141,167]. W przypadku pojedynczej sondy wykazującej podobieństwo do genu ekspansyny 11 odnotowano inhibicję ekspresji dla obu traktowań. Ciekawe jest podobne zachowanie genów odpowiedzialnych za właściwości ściany komórkowej w obu układach badawczych, jako że opisywane zjawiska obserwowano do tej pory jedynie w przypadku działania stresu solnego [95,141,142].

V.2.2. Profil ekspresji genów w korzeniach tytoniu poddanych działaniu stresu solnego.

Dyskutując profil ekspresji genów specyficznych dla stresu solnego warto odwołać się do pracy transkryptomicznej, kompleksowo opisującej działanie NaCl na korzenie rzodkiewnika. Chlorek sodu podawano tu w identycznym stężeniu, a próby zbierano m.in. po 6 godz., podobnie jak w niniejszym eksperymencie [95]. Na uwagę zasługiwały też prace dotyczące korzeni pomidora [142] i bawełny [99], jednak traktowanie przeprowadzono tu w nieco innych warunkach.

Wśród wielu genów ulegających indukcji w korzeniach poddanych działaniu stresu solnego, warto zwrócić uwagę na geny wchodzące w skład kategorii GO związanej z odpowiedzią na niedobór wody, co świadczy o cechach wspólnych odpowiedzi na stres solny i suszy. W innych doniesieniach również obserwowano wzrost ekspresji genów DREB2A, ABA1 i ERD15 w stresie solnym i osmotycznym, natomiast niewiele wiadomo na temat indukcji genów białek histonowych w tych warunkach.

Geny wchodzące w skład ścieżki zależnej od ABA, zarówno specyficzne dla odpowiedzi na NaCl, jak i ulegające indukcji w obu eksperymentach, świadczą o obecności w korzeniach tytoniu typowych ścieżek transdukcji sygnału dla stresu solnego i

pokrywają się z danymi uzyskiwanymi do tej pory dla rzodkiewnika i innych gatunków [160,161,191]. W eksperymencie na korzeniach rzodkiewnika, dominującymi hormonami regulującymi odpowiedź były kwas jasmonowy i etylen, drugi z hormonów, którego indukcję obserwowano nie tylko w niniejszym eksperymencie, ale też w innych pracach [95, 160,161].

Starzenie jest dobrze znanym skutkiem stresu solnego związanym z zahamowaniem wzrostu i objawiającym się rozpadem chlorofilu, żółknięciem liści, a w przypadku korzeni zmniejszeniem ich długości [164], regulowanym przez hormony, m.in. ABA i etylen [163]. W badaniach siewek rzodkiewnika [194] stwierdzono indukcję wielu genów zaangażowanych w proces starzenia wywołanych stresem solnym, przede wszystkim genów zawierających domenę NAC. Ekspresja tego typu genów w korzeniu nie była do tej pory opisana w pracach dotyczących transkryptomiki.

Kwestią dyskusyjną jest inhibicja kategorii GO dotyczącej aktywności peroksydaz, które zwykle kojarzone są z usuwaniem wolnych rodników. Prace dotyczące skutków rozmaitych czynników stresowych wskazują na różne zachowania katalazy i dysmutazy ponadtlenkowej w zależności od organu i rodzaju traktowania, prezentując sprzeczne wyniki nie pozwalające na ostateczne określenie roli tych genów [172]. W przypadku oceny aktywności w korzeniach fasolnika chińskiego stwierdzono zmniejszenie aktywności dysmutazy, podczas gdy poziom katalazy nie zmieniał się [172]. Być może zatem obserwowana inhibicja niektórych katalaz ma związek z innego rodzaju aktywnością niż usuwanie wolnych rodników. Podobną tendencję stwierdzono w stresie solnym w korzeniach rzodkiewnika [95], gdzie również stwierdzono zmniejszenie ekspresji niektórych peroksydaz.

Porównując uzyskane wyniki z innymi pracami transkryptomicznymi oceniającymi stres solny w korzeniach, można wyróżnić znaczną liczbę genów lub grup genów, których ekspresja w tytoniu ulega podobnym zmianom co u innych gatunków. Należą do nich geny charakterystyczne dla ścieżek sygnałowych stresu (MAPK, ERF, ATAF, EREBP, WRKY, HSP, CIPK), świadczące o sygnalizacji zależnej od hormonów (ABA, ET, JA), związane ze zmianami w ścianie komórkowej (XTH, arabinogalaktany) lub transportem (MATE, ABC, akwaporyny). Wśród genów, których występowanie w stresie solnym jest opisane w literaturze, a nie jest wyraźnie zaznaczone w niniejszych badaniach, można wymienić geny ścieżki SOS, dobrze opisanego w literaturze szlaku odpowiedzi na stres solny [195], geny osmoprotekcyjne (w niniejszym eksperymencie widoczna była jedynie indukcja genów fosfatazy trehalozy), geny związane z lignifikacją ściany komórkowej. Mimo to, jeśli

chodzi o liczbę widocznych zmian, jest ona większa niż np. w przypadku wspomnianych prac dla pomidora i bawełny.

V.2.3. Profil ekspresji genów w korzeniach tytoniu poddanych działaniu kadmu.

W przypadku ekspresji genów w korzeniach poddanych działaniu kadmu, dostępnych jest znacznie mniej badań transkryptomicznych w porównaniu z działaniem stresu solnego. Na uwagę zasługują właściwie 3 prace: 2 dotyczące korzeni ryżu [149, 189], a jedna rzodkiewnika [97]. Odnotowana w niniejszym badaniu indukcja genów charakterystycznych dla odpowiedzi na jony metali nie była odnotowana w wymienionych eksperymentach, niemniej jednak rola genów ZIF, NRAMP i YSL w chelatowaniu i transporcie jonów metali ciężkich została opisana w literaturze [173,196]. Spośród hormonów biorących udział w odpowiedzi na kadm, najbardziej wyraźna była indukcja ścieżki syntezy i sygnalizacji za pomocą etylenu i kwasu jasmonowego. Co ważne, w obu przypadkach wyraźnie zaznaczony jest wzrost ekspresji kilku genów biorących udział w szlaku syntezy obu związków, które są wymienione w literaturze jako hormony o istotnym znaczeniu dla odpowiedzi na stres metali ciężkich [175,176]. Indukcję genów odpowiedzialnych za sygnalizację za pośrednictwem kwasu jasmonowego wykryto za pomocą mikromacierzy w korzeniach ryżu [149].

Odnotowane zwiększenie ekspresji genów związanych z asymilacją siarki i syntezą acetylokoenzymu A może być związane z indukcją syntezy glutationu lub fitochelatyn [135,178,189]. Wzrost ekspresji takich genów odnotowano również w korzeniach ryżu traktowanych kadmem i miedzią, a udział powyższych zjawisk w odpowiedzi na stres kadmowy w korzeniach tytoniu jest ewidentny [189].

Bardziej skomplikowana jest kwestia ekspresji genów odpowiedzialnych za transport. Indukcji uległy geny związane m.in. z usuwaniem jonów metali (PDR i NRAMP), jednak zdecydowana większość genów związanych z transportem uległa inhibicji. W korzeniach ryżu obserwowano wzrost ekspresji transporterów PDR i innych z grupy ABC, a także MATE i CDF [149,189], natomiast zmniejszenia ekspresji tak wielu transporterów nie notowano w innych pracach transkryptomicznych. Być może podczas stresu kadmowego niektóre transportery mogą być hamowane w celu zapobieżenia wypływu jonów np. z korzeni do pędu, tak jak w przypadku niektórych transporterów SULTR (siarczanów) ulegających inhibicji w korzeniach rzepaku pod wpływem kadmu [183]. W przypadku innych genów, których ekspresja ulega zmniejszeniu, zwrócono uwagę na geny białek cytoszkieletu. Wpływ jonów kadmu na cytoszkielet objawiający się rozpadem mikrotubul i włókien aktyny jest zjawiskiem obserwowanym w komórkach

roślinnych i zwierzęcych, jednak niewiele wiadomo na temat zmian ekspresji genów kodujących białka cytoszkieletu [180]. Z całą pewnością inhibicja tych genów w warunkach stresu musi być wzięta pod uwagę w przypadku typowania genów referencyjnych o stałej ekspresji używanych w reakcji PCR w czasie rzeczywistym.

W przypadku innych prac oceniających wpływ kadmu na zmiany ekspresji genów w korzeniach roślin, do dość często powtarzających się genów należały geny związane z usuwaniem wolnych rodników (GST), sygnalizacją (MAPK, CDPK, DREB, WRKY, AP2), chelatowaniem jonów (glutation, metylotioneiny), transportem (PDR, MATE, CDF, ZIP), metabolizmem siarki czy modyfikacją ścian komórkowych [149,189]. Duża część tych procesów była w większym lub mniejszym stopniu odnotowywana w niniejszym eksperymencie, w ramach ogólnej odpowiedzi na stres abiotyczny lub reakcji specyficznej na jony kadmu. Obserwacje dotyczące wspólnych oraz zależnych od bodźca zmian w ekspresji genów korzeni tytoniu traktowanych CdCl_2 i NaCl zebrano na Rycinie 41. Ze względu na poboczny charakter tej części pracy, przybliżono jedynie globalny obraz zmian transkryptomu. Dla dokładnego poznania procesów stresowych w tym układzie badawczym niezbędne byłyby dodatkowe badania z wykorzystaniem np. różnych punktów czasowych i stężeń czynników stresowych. Niemniej, należy podkreślić, że uzyskane mapy aktywności transkryptomu w ww. stresach stanowią cenne i bogate źródło danych do ukierunkowanych badań fizjologii korzeni tytoniu w stresie. W szczególności znajomość sekwencji sond różnicujących, może okazać się przydatna do wyekstrahowania z baz danych informacji o pełnych sekwencjach cDNA konkretnych genów związanych z reakcją na stres w tytoniu.

V.3. ZASTOSOWANIE CSH W ANALIZIE TRANSKRYPTOMÓW

V.3.1. Wpływ długości sond na jakość eksperymentów CSH

Wpływ długości sond oceniano porównując wyniki uzyskane przy zastosowaniu mikromacierzy TOM1 i TOM2, ponieważ pochodziły one z tego samego laboratorium i reprezentowały podobny zestaw transkryptów. Gwarantowało to uzyskanie maksymalnie dużej liczby danych do późniejszych porównań między mikromacierzami. Biorąc pod uwagę liczbę genów różnicujących oraz zbieżność z danymi referencyjnymi, zdecydowanie lepsze wyniki, wykazujące większe podobieństwo do uzyskanych w eksperymencie SSH, otrzymano stosując mikromacierze z sondami cDNA. W przypadku mikromacierzy TOM2 uzyskano niewiele danych o znaczeniu biologicznym, a geny różnicujące, pokrywające się z uzyskanymi na mikromacierzy referencyjnej, są nieliczne,

co w porównaniu z wynikami otrzymanymi z wykorzystaniem mikromacierzy TOM1, wskazuje jednoznacznie na przewagę mikromacierzy cDNA nad oligo w eksperymentach CSH. Wynik ten jest zgodny ze spostrzeżeniem wskazującym na większą liczbę genów różnicujących i lepsze parametry FDR uzyskiwane przy pomocy dłuższych sond [58]. Nawet w przypadku wyselekcjonowania puli 600 sond o najlepszych parametrach statystycznych (top600), skuteczność mikromacierzy cDNA była zdecydowanie wyższa. Mimo iż takiego stanu rzeczy można było oczekiwać na podstawie wcześniejszych prac [40,56,77,78], spodziewano się jednak nieco większej specyficzności sond oligo. Co ciekawe, w zbiorze prac porównujących hybrydyzację mikromacierzy cDNA i oligo, można znaleźć doświadczenia wskazujące na niższą specyficzność mikromacierzy oligo w porównaniu z mikromacierzami cDNA (320-390 pz.) wynikającą z większej podatności na hybrydyzację krzyżową [197]. Stwierdzono, że hybrydyzacji krzyżowej sprzyja wyższe od 75% podobieństwo niespecyficznego transkryptu do sondy oraz obecność w pełni komplementarnych do transkryptu odcinków sondy o długości powyżej 15 zasad. Zatem w pracy tej podważono jeden z wyżej przytoczonych argumentów przemawiających za stosowaniem mikromacierzy oligo, mówiących o większej specyficzności wiązania transkryptów przez sondy. Równocześnie badane sondy wykazywały podobną czułość, określoną jako zdolność do wykrycia 3-krotnej różnicy w stężeniu dodawanych do próby transkryptów o znanej ilości. Autorzy nie negują przydatności sond oligo w zastosowaniach mikromacierzowych, a jedynie wskazują na konieczność starannego ich projektowania i znaczenie dostępności sekwencji DNA. Zatem, zgodnie z tymi spostrzeżeniami, w przypadku CSH, w której prawdopodobieństwo przyłączenia do sondy kilku podobnych transkryptów jest większe niż w tradycyjnej hybrydyzacji, stosowanie mikromacierzy oligo wymagałoby znacznie bardziej szczegółowej filtracji danych pod kątem podobieństwa sekwencji, uwzględniającej powyższe warunki [197].

Z kolei, w przypadku sond cDNA problemem mogą być trudności w identyfikacji rzeczywistych transkryptów reprezentowanych na mikromacierzach, nawet w przypadku SSH – sekwencje sond są tylko częściowo znane, dodatkowo często transkrypty roślin obecne w bazach danych mają postać kontigów (ciągów częściowo zachodzących na siebie sekwencji), które mogą zawierać błędy. Ponadto w przypadku większości roślin sekwencje te najczęściej są adnotowane na podstawie analiz funkcjonalnych, np. z wykorzystaniem genów rzodkiewnika. W takim przypadku dysponujemy zatem sondą „prawdopodobnie” reprezentującą dany kontig, ten z kolei tylko „prawdopodobnie” jest funkcjonalnym odpowiednikiem genu w roślinie modelowej. Wynik uzyskany przy pomocy takiej sondy

odnosimy do cDNA z organizmu badanego, który „prawdopodobnie” jest homologiem genu „prawdopodobnie” reprezentowanego na mikromacierzy. Biorąc pod uwagę tak dużą niepewność, nawet pomimo uzyskanych lepszych wyników dla sond cDNA niż w przypadku sond oligonukleotydowych, należy zachować dużą dozę ostrożności podczas wnioskowań dotyczących ekspresji poszczególnych genów. Dodatkowo warto tutaj wspomnieć o kwestii ogromnej zmienności genetycznej organizmów już na poziomie wewnątrzgatunkowym, na co wskazuje szereg badań. Przykład mogą stanowić wyniki eksperymentów Rna-seq dla kukurydzy - wskazujące na obecność lub brak niektórych transkryptów w zależności od odmiany [185]. Biorąc to pod uwagę, należy liczyć się z faktem, że każda analiza oparta o hybrydyzację jest ograniczona tylko do zestawu sond reprezentowanych na mikromacierzach, a istnienie genomu referencyjnego nie gwarantuje pełnej reprezentacji transkryptomu nawet w przypadku pracy na poziomie innej odmiany danego gatunku, nie wspominając już o badaniach międzygatunkowych.

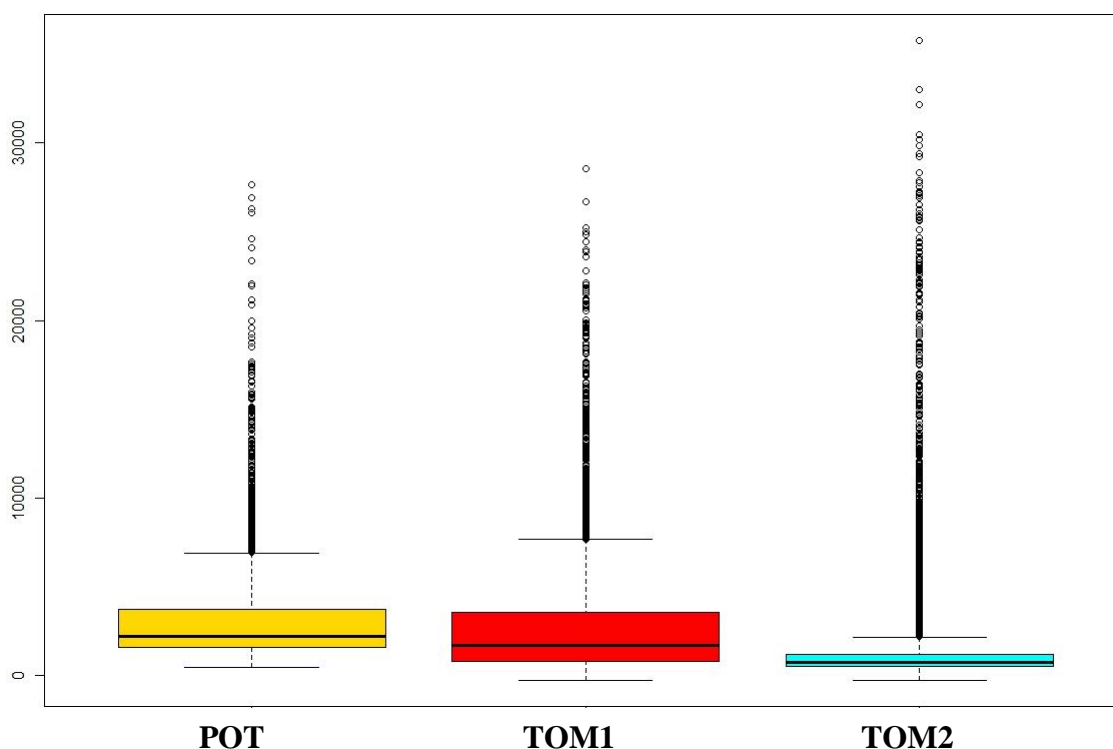
V.3.2. Wpływ pokrewieństwa gatunku badanego i reprezentowanego na mikromacierzy na jakość CSH

Biorąc pod uwagę stosunkowo niewielką liczbę mikromacierzy w kontekście istniejących gatunków, istnieje duże prawdopodobieństwo, że wybór badacza będzie oczywisty i ograniczony tylko do jednej mikromacierzy, lub też może wystąpić sytuacja, w której zbyt duża odległość filogenetyczna uniemożliwi wykonanie eksperymentów. Niemniej, jednak, w przypadku istnienia kilku mikromacierzy w obrębie jednej rodziny (tak jak dla *Solanaceae*, *Brassicaceae*) (por. Tabela 2), może zaistnieć konieczność wyboru najlepszej mikromacierzy spośród kilku opcji. W przypadku tego porównania analizowano eksperymenty wykonane na mikromacierzach cDNA dla pomidora i ziemniaka zawierających sondy o zbliżonym, choć niedokładnie określonym stopniu podobieństwa sekwencji do tytoniu i odznaczających się podobną liczbą sond użytkowych. Jeśli założyć, że elementem warunkującym jakość eksperymentów CSH jest podobieństwo sekwencji gatunków [55,57,80] nie powinno się uzyskiwać dużej rozbieżności wyników pomiędzy dwiema mikromacierzami w tych samych warunkach eksperymentalnych. Z punktu widzenia statystyki korelacji pojedynczych genów różnicujących, uzyskane wyniki wskazują na lepszą zgodność z danymi referencyjnymi (TOB) mikromacierzy POT. Z kolei ilościowe spojrzenie na dane uzyskane w wyniku analiz GO nie wskazuje tak jednoznacznie na przewagę jednej z mikromacierzy. Gorsze wyniki analizy na podstawie kategorii GO specyficznych gatunkowo mogą być uzasadnione ponad trzykrotnie większą

liczbą kategorii GO przyporządkowanych sondom TOM1, czego skutkiem może być brak niektórych kategorii w analizie mikromacierzy POT.

Tabela 32. Liczba kategorii GO dostępnych dla mikromacierzy POT i TOM1

	POT	TOM1
Kombinacje GO-sonda	16242	66851
Liczba sond posiadających GO	8837	6811
Liczba unikatowych GO	1169	3342



Rycina 40. Średnia intensywność sygnału (Sig) dla wszystkich sond na poszczególnych mikromacierzach CSH.

Większość prac poruszających kwestię pokrewieństwa pomiędzy gatunkiem badanym i reprezentowanym na mikromacierzy wskazywało na intensywność sygnału jako wyznacznik jakości hybrydyzacji, sugerując że nawet minimalne różnice (1%) sekwencyjne znajdują istotne odzwierciedlenie w intensywności sygnału sond [55,76,80]. W niniejszym eksperymencie mikromacierze POT i TOM1 cechowała podobna intensywność sygnału (Rycina 40), a wyniki badań homologii wskazywały na podobne, aczkolwiek nieznacznie większe podobieństwo sond do transkryptów na mikromacierzy TOM1 (Tabela 10, Rycina 15). Na podstawie niniejszego eksperymentu można zatem wyciągnąć wniosek, że o wyborze spośród kilku mikromacierzy o zbliżonym

podobieństwie sekwencji DNA do gatunku badanego powinny decydować inne aspekty niż niewielkie różnice w homologii sekwencji. Jednym z takich czynników może być podobieństwo gatunków dotyczące niektórych cech w specyficznych warunkach biologicznych, o którym wspomniano już w kontekście analizy unikalnych genów związanych z dojrzewaniem owoców papryki i bakłażana, podczas której stwierdzono większy wpływ na uzyskany profil ekspresji podobieństwa fenotypowego, niż bliskości filogenetycznej [53]. We wspomnianej pracy oceniano jednak kilka gatunków przy pomocy tej samej mikromacierzy, w niniejszym eksperymencie mieliśmy do czynienia z sytuacją odwrotną- ten sam materiał badano przy pomocy mikromacierzy dla różnych gatunków. W tej sytuacji warto również zwrócić uwagę na różnice w profilu sond TOM1 i POT (Rycina 7, punkt IV.1.2)- liczba sond POT i TOM1 posiadających wspólne odpowiedniki na mikromacierzy TOB wynosiła zaledwie ok.2500, co świadczy o tym, że przynajmniej część analizowanych sond na obu mikromacierzach nie pokrywała się, generując odmienny profil. Zatem skład sond w kontekście badanego stanu biologicznego jest również kwestią godną uwagi podczas wyboru mikromacierzy. Ostatnim czynnikiem, który należy przeanalizować jest liczba dostępnych danych genetycznych i adnotacji funkcjonalnych dla danego gatunku i sond, o którym wspomniano wyżej w kontekście kategorii GO. W przypadku mikromacierzy dla gatunków o zbliżonym podobieństwie do badanego warto rozważyć wybór tego, który daje większe możliwości analizy dzięki bogatszemu zasobowi danych.

V.3.3. Wpływ liczby zastosowanych powtórzeń na jakość CSH

Zastosowanie dodatkowych powtórzeń z zamianą barwników (*dye swap*) zdecydowanie wpłynęło na jakość uzyskanych danych dla mikromacierzy POT i ich zgodność z wynikami referencyjnymi. Przy zachowaniu proporcji sond wykazujących wzrost lub spadek ekspresji uzyskano poprawę parametrów zgodności z mikromacierzą referencyjną w każdym badanym aspekcie. Różnice pomiędzy wynikami uzyskanymi dla jednego lub dwóch powtórzeń były znacznie większe niż w przypadku omawianego wcześniej porównania mikromacierzy dwóch odmiennych gatunków. Zatem szczególnie w przypadku CSH rozważenie dodatkowych powtórzeń wydaje się być kwestią niezwykle istotną, a czynnik ten może mieć większe znaczenie dla powodzenia eksperymentu CSH niż nieznaczące różnice w podobieństwie sekwencji DNA. Paradoksalnie wydaje się zatem, że jednym z kluczowych elementów w przypadku hybrydyzacji CSH jest spełnienie ogólnie przyjętych warunków dobrej pracy z mikromacierzami, jakim jest wykonanie jak największej liczby powtórzeń, nawet w przypadku gdy odbywa się to kosztem mniejszej

liczby warunków eksperymentalnych. Podobnie jak w przypadku SSH, istotne jest również zastosowanie zamiany barwników (*dye swap*) pozwalającej wyeliminować efekt wyciszania jednego z barwników [200, 201]. W niniejszym eksperymencie dysponowano trzema powtórzeniami biologicznymi, dlatego zastosowano powtórzenia techniczne, które z punktu widzenia ekonomiki eksperymentu są dość kosztownym rozwiązaniem, choć korzyści z takiego modelu doświadczenia są wyraźnie widoczne. W przypadku posiadania większej liczby powtórzeń biologicznych, wskazane jest ich użycie zamiast powtórzeń technicznych, wraz z zastosowaniem zamiany barwników, gdyż takie rozwiązanie będzie skuteczniejsze z punktu widzenia eliminacji czynników zaburzających [201].

V.3.4. Wybór optymalnej ścieżki filtracji danych

V.3.4.1. Filtracja HOM

W przypadku badań Bar-Or i wsp. [54] zauważono poprawę jakości wynikającą z zastosowania filtracji, objawiającą się większym podobieństwem pomiędzy powtórzeniami biologicznymi, niemniej jednak nie przedstawiono porównania uzyskanych profilów w kontekście genów różnicujących. W niniejszej pracy porównywano ostateczne wyniki, które są przecież najbardziej interesujące dla biologa. Trudno jest jednak wskazać jednoznaczne korzyści z zastosowania filtracji bazującej na ocenie podobieństwa sekwencji genetycznej między sondami, a transkryptami na podstawie uzyskanych w niniejszej pracy wyników.

Dla mikromacierzy POT i TOM1, odsetek sond usuniętych w wyniku filtracji HOM był mniejszy (17% i 10 %) niż dla TOM2 (51%), być może dzięki temu filtracja ta nie wykazywała szczególnego wpływu na zbieżność wyników CSH z SSH. Największe różnice po zastosowaniu tego typu filtracji odnotowano dla mikromacierzy oligonukleotydowej, wykazującej większą podatność na rozbieżność sekwencji pomiędzy sondami, a transkryptami, niemniej jednak nie idą one w parze z poprawą jakości danych pod kątem zbieżności z SSH. W przypadku najdokładniej ocenianych w niniejszym eksperymencie mikromacierzy POT, zbieżność genów różnicujących dla analiz bez filtracji i z filtracją HOM wynosiła co najmniej 88% (Rycina 39), zatem być może samo właściwe przeprowadzenie analizy danych i wyselekcjonowanie genów o najlepszych parametrach statystycznych może wystarczyć do uzyskania wiarygodnych biologicznie wyników CSH, bez potrzeby stosowania dodatkowej filtracji.

Z całą pewnością, sam fakt dostępności sekwencji transkryptów daje możliwość oceny przydatności danej mikromacierzy już przed przystąpieniem do eksperymentów.

Ważnym aspektem, który należy poddać dyskusji jest nie tyle zastosowanie filtracji, co możliwość wykorzystania informacji na temat podobieństwa sond podczas analizy funkcjonalnej bazującej na danych dla mikromacierzy dla gatunku badanego. W przypadku analizy GO z wykorzystaniem adnotacji funkcjonalnych dla mikromacierzy TOB, uzyskano mniejszą liczbę danych niż w przypadku analizy uwzględniającej adnotacje funkcjonalne dla ziemniaka, jednak wykazywały one większy odsetek zgodności z danymi referencyjnymi. Z kolei analiza z wykorzystaniem adnotacji sond dla gatunku reprezentowanego na mikromacierzy tworzy profil ekspresji charakterystyczny dla genów tego gatunku (w przypadku POT-dla ziemniaka), a nie homologów występujących w genomie badanym. W tym drugim przypadku odnotowano większą liczbę różnicujących kategorii GO, ale też więcej niezgodności w stosunku do mikromacierzy TOB.

V.3.4.2. Filtracja MORF

Jak wspomniano w poprzednim punkcie, sama filtracja danych na podstawie homologii nie wykazywała dużego wpływu na wyniki wykonanych eksperymentów CSH, dlatego w kontekście takich obserwacji, trudno było oczekiwać znaczącej poprawy jakości danych po zastosowaniu filtracji MORF. Niemniej jednak wykonano badania mające na celu sprawdzenie, czy istnieją przesłanki do stworzenia alternatywnego sposobu filtrowania danych CSH, których podstawą były doniesienia Bar-Or i wsp. [52] wskazujące na wyraźną zależność niektórych parametrów jakości plamek od homologii sond i transkryptów, sugerujące możliwość ich wykorzystania do filtracji danych w przypadku braku dostępnych sekwencji. Stwierdzono, że korelacja między wartością Bit Score a tymi parametrami była zbyt słaba, by można było mówić o liniowej zależności tych parametrów. Niemniej jednak, dzięki zastosowanemu w niniejszej pracy podejściu wykorzystującym dane SSH zauważono, że z punktu widzenia całej mikromacierzy, lepsze parametry jakości uzyskuje się dla sond o wyższej homologii, co potwierdza obserwacje autorów wspomnianej pracy, dotyczące istniejącej zależności pomiędzy tymi wartościami. Obserwowana nieznaczna poprawa charakterystyki jakościowej plamek przy rosnącej homologii ma charakter skokowy, a wyniki uzyskane na mikromacierzach PLasTi i POT wskazują na konieczność indywidualnego ustalania progu filtracji, bazującego na wynikach SSH. Uzyskanie wyników gorszych niż w przypadku filtracji na podstawie homologii, a nawet braku jakiegokolwiek filtracji, podważa wnioski sformułowane przez Bar-Or i wsp. [52] a także innych autorów [37,40,61] sugerujących możliwość wykorzystania parametrów morfologicznych do skutecznej filtracji danych. Z całą pewnością wykonanie filtracji danych na podstawie niektórych parametrów

morfologicznych (SNR, intensywność sygnału) jest w niektórych laboratoriach standardową praktyką stosowaną nawet dla mikromacierzy homologicznych i może być pomocne dla odrzucenia mniej wartościowych danych [204], dlatego zaleca się używanie umiarkowanie restrykcyjnych metod filtracji na podstawie jakości plamek przy każdym eksperymencie mikromacierzowym. Należy również pamiętać, że pierwszy stopień filtracji odbywa się już na poziomie nakładania siatki przez program do obróbki danych, który nie będzie wychwytywał plamek o zbyt słabym sygnale. Użycie odpowiedniego, sprawdzonego programu i właściwe wykonanie etapu konwersji danych może być ważnym etapem dla uzyskania jak najlepszych wyników hybrydyzacji.

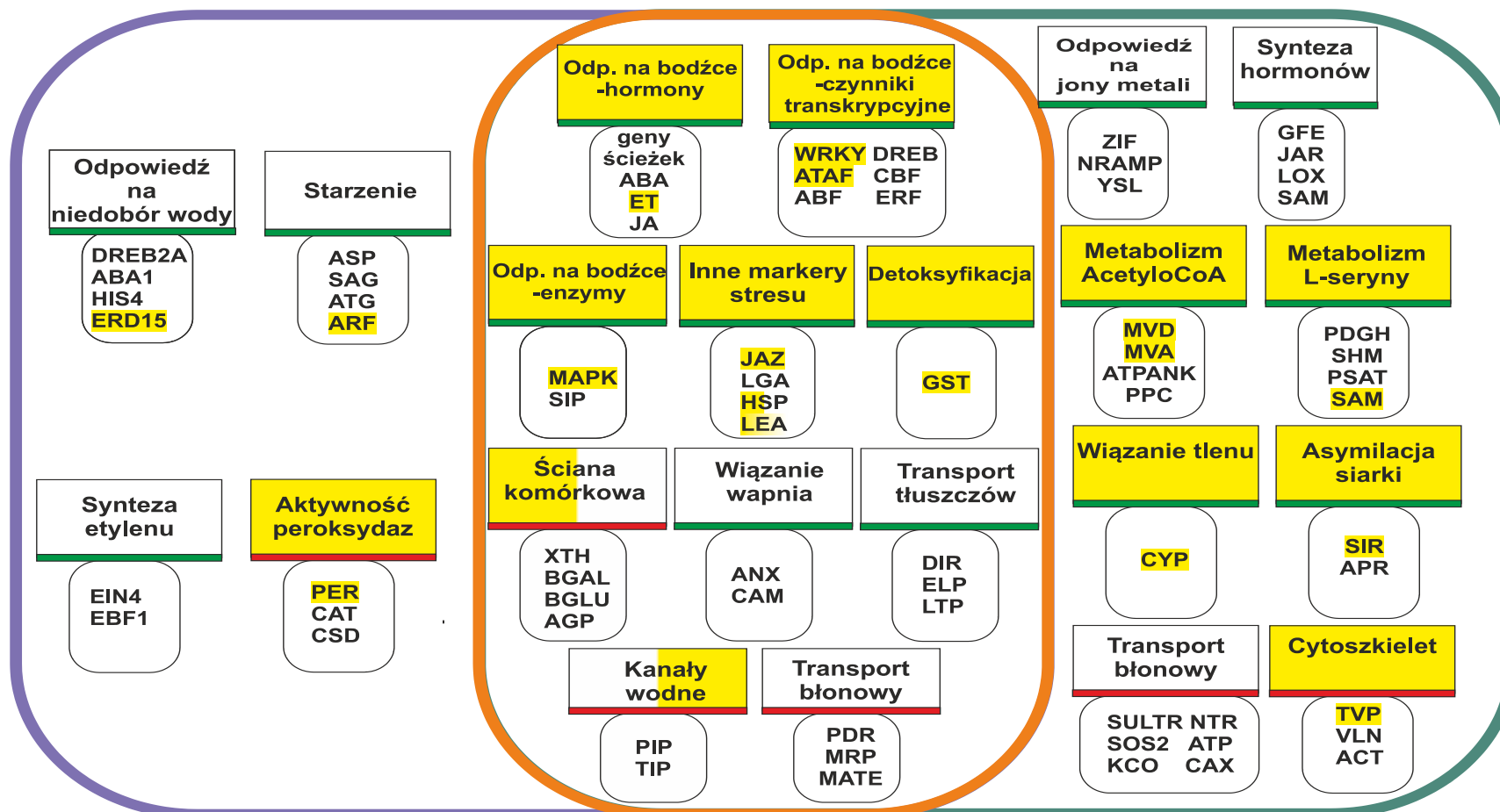
V.4. SKUTECZNOŚĆ CSH JAKO METODY BADANIA TRANSKRYPTOMU

Na podstawie wykonanych doświadczeń i porównań różnych aspektów hybrydyzacji CSH, można dokonać oceny przydatności CSH w analizie danych biologicznych. Porównania na poziomie pojedynczych sond oraz analiza funkcjonalna wskazują, że przy użyciu tej metody można uzyskać wiarygodne dane biologiczne. Najbardziej podobne do SSH wyniki uzyskano w przypadku stosowania mikromacierzy POT i dwóch powtórzeń technicznych każdej hybrydyzacji, z zamianą barwników. Na Rycinie 41 przedstawiono ogólny schemat przedstawiający zgodność wyników SSH i CSH. Oczywiście, za pomocą tej metody nie otrzymano wyników pozwalających wykryć wszystkie procesy. Niemniej jednak, analiza funkcjonalna pozwoliła na wyselekcjonowanie zmian w większości kategorii wytypowanych jako istotnie zmienione w tytoniu pod wpływem stresu. Warto zauważyć także dużą zgodność kierunku zmiany ekspresji grup i genów różnicujących, a także niewielką liczbę zmian, które nie występowały na mikromacierzy referencyjnej i można je określić jako wątpliwe. Więcej takich zmian stwierdzono w przypadku analizy GO wykorzystującej podejście korzystające z danych GO dla ziemniaka (wiążanie jonów miedzi w stresie solnym, inhibicja genów odpowiedzi na brak wody w stresie kadmowym), dlatego należy zachować ostrożność w przypadku tego typu podejścia i raczej traktować je jako uzupełniające wyniki analizy funkcjonalnej wykonane za pomocą danych dla gatunku reprezentowanego na mikromacierzy. W dwóch kategoriach funkcjonalnych (przedstawionych na Rycinie 41) stwierdzono różnicowanie tylko w przypadku pojedynczego stresu, podczas gdy mikromacierz referencyjna wykazała istotne zmiany w obu stresach. Oczywiście, ze względu na wspomniane ograniczenia (np. dotyczące danych GO) przedstawione w pracy wyniki raczej nie mogą posłużyć do wytyczenia dokładnych szlaków zmian ekspresji

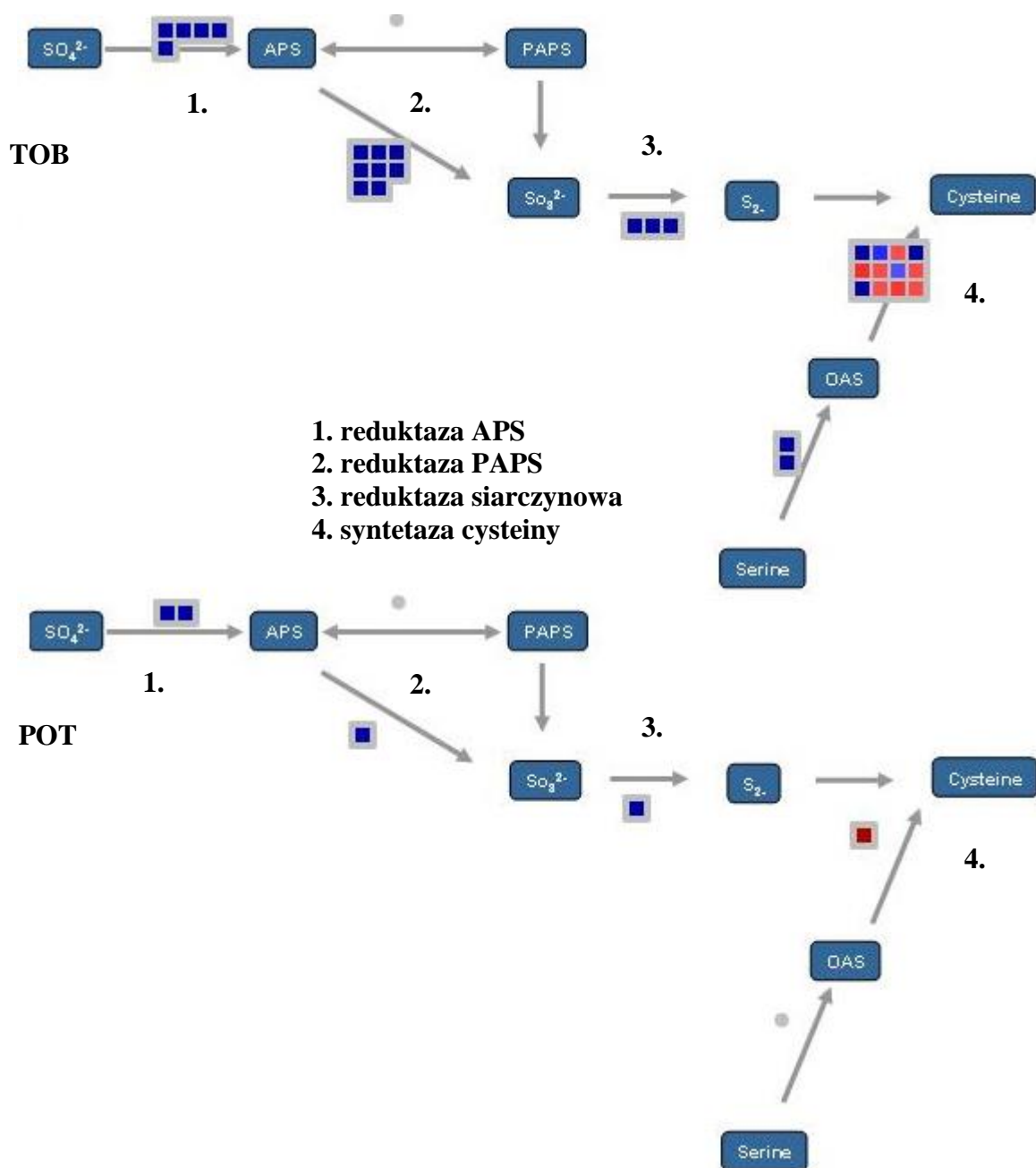
genów pod wpływem stresu abiotycznego, jednak są wystarczające do nakreślenia orientacyjnego profilu zmian i wyznaczenia ewentualnych obszarów zainteresowania podczas badań. Przykład wykorzystania uzyskanych danych do wyznaczenia profilu ekspresji genów związanych z asymilacją siarki w stresie kadmowym przedstawiono na Rycinie 42, na której można zauważyć zgodność wyników CSH z mikromacierzą referencyjną w przypadku bardziej szczegółowej analizy procesu biologicznego, mimo znacznie mniejszej liczby sond reprezentujących to zjawisko.

STRES SOLNY

STRES KADMOWY



Rycina 41. Skuteczność wykrywania zmian w transkryptomie za pomocą CSH (POT) w porównaniu z SSH (TOB). Przedstawiono wybrane procesy i geny tworzące profile ekspresji wyznaczone za pomocą mikromacierzy TOB. Kolorem żółtym oznaczono wybrane procesy/geny występujące na mikromacierzy POT i wykazujące zgodność z mikromacierzą referencyjną. Pola oznaczone w połowie oznaczają wykrycie zmian procesu/genu tylko w przypadku jednego traktowania.



Rycina 42. Wspólne cechy profilów ekspresji genów związanych z asymilacją siarki w korzeniach tytoniu traktowanych chlorkiem kadmu, badanych za pomocą mikromacierzy SSH (TOB) i CSH (POT). Schematy wykonano z wykorzystaniem programu Mapman. Każdy kwadrat oznacza pojedynczą sondę. Kolor czerwony oznacza spadek ekspresji genu, niebieski- wzrost ekspresji genu.

V.5. OPTYMALNA ŚCIEŻKA ANALIZ CSH

Jednym z ważniejszych celów niniejszej pracy było opracowanie optymalnej ścieżki postępowania podczas wykonywania eksperymentów CSH i analizy danych. Kluczowym elementem warunkującym powodzenie eksperymentu jest ocena dostępnych zasobów laboratoryjnych (mikromacierze) i analitycznych (dostępność sekwencji, danych GO, narzędzi analizy danych). Niniejsze wyniki wskazują, że znajomość stopnia podobieństwa sond do transkryptów jest niezwykle istotnym czynnikiem wpływającym na powodzenie eksperymentu CSH, biorąc pod uwagę filtrację danych oraz możliwości późniejszej analizy funkcjonalnej. Wykazano, że mikromacierze o dłuższych sondach cDNA sprawdzały się w tych eksperymentach znacznie lepiej niż oligonukleotydowe. Projekt eksperymentu powinien również zakładać maksymalizację liczby powtórzeń i zastosowanie zamiany barwników *dye-swap*, co udowodniono eksperymentalnie. Z kolei w przypadku analizy danych warto zwrócić uwagę na ocenę jakości plamek i w razie potrzeby zastosować nieco bardziej restrykcyjne progi odcięcia niż w przypadku SSH, ponieważ wykazano, że lepsze parametry morfologiczne uzyskują sondy o większej homologii. Innym pomocnym zabiegiem może być bardziej restrykcyjna niż w przypadku SSH filtracja na podstawie uzyskanych logFC i wartości P. W niniejszym eksperymencie selekcja 600 sond o najlepszych parametrach pozwoliła na stworzenie stabilnych grup genów różnicujących, pozwalających na uzyskanie w przypadku mikromacierzy TOM1 i POT wiarygodnych danych transkryptomocnych. Tego typu filtracja wydaje się mieć największy wpływ na jakość uzyskanych danych niż inne testowane tu metody. Ostatnim elementem jest analiza funkcjonalna, która wymaga dostępności odpowiednich danych (np. ontologicznych). Najbardziej wskazane wydaje się korzystanie z danych dla organizmu badanego i ewentualnie uzupełnienie braków adnotacjami dla gatunku reprezentowanego na mikromacierzy (Rycina 43).

PROJEKT EKSPERYMENTU

1. Ocena podobieństwa genetycznego gatunków

Dostępność danych o podobieństwie sekwencji sonda-transkrypt

2. Wybór mikromacierzy

cDNA

WYKONANIE EKSPERYMENTU

3. Układ hybrydyzacji

P Powtórzenia biologiczne
Dye-swap
Powtórzenia techniczne

ANALIZA DANYCH

4. Analiza obrazów

Ocena jakości plamek

5. Analiza statystyczna

Filtracja na podstawie logFC i wartości P

6. Analiza funkcjonalna

Dostępność adnotacji funkcjonalnych dla gatunku badanego
Adnotacje funkcjonalne dla gatunku reprezentowanego na mikromacierzy-jako uzupełnienie

Rycina 43. Optymalny przebieg eksperymentu CSH na podstawie wyników uzyskanych w niniejszej pracy

V.6. ALTERNATYWNE ZASTOSOWANIA CSH

W tym miejscu w oczywisty sposób nasuwa się kwestia potencjalnej rezygnacji ze stosowania mikromacierzy CSH na rzecz wysokowydajnego sekwencjonowania transkryptomu *de novo*. Mimo iż cena eksperymentu mikromacierzowego jest nadal niższa niż sekwencjonowania, zwłaszcza jeżeli badacz chce zastosować podobną liczbę powtórzeń, ta ostatnia technika staje się coraz atrakcyjniejsza i dostępniejsza w badaniach transkryptomicznych, a koszty mogą być zmniejszone np. poprzez zastosowanie „indeksów” umożliwiających analizę kilku próbek w jednej linii eksperymentu (platforma Illumina). Tego typu eksperyment teoretycznie umożliwiłby ominięcie jednego etapu wnioskowania na podstawie homologii i dotarcie do sekwencji właściwego transkryptu. Niemniej jednak, nawet pomijając ograniczenia jakie niesie za sobą asemblacja *de novo*, mapowanie badanego transkryptu wymaga użycia sekwencji innego organizmu, możliwie zbliżonego. Zatem eksperymentator nie jest w stanie całkowicie pominąć etapu, na którym wymagana będzie sekwencja DNA, przy pomocy której możliwe będzie wnioskowanie na podstawie homologii. Ponadto należy zwrócić uwagę na fakt, iż tego typu procedura stanowi ogromne wyzwanie bioinformatyczne, znacznie większe niż w przypadku analizy mikromacierzy, a nawet wysokowydajnego sekwencjonowania organizmów z wykorzystaniem genomu referencyjnego, co znacznie ogranicza jej dostępność dla przeciętnego badacza. Mimo powyższych ograniczeń, wydaje się, że ich przezwyciężenie jest jedynie kwestią czasu. W związku z tym można przypuszczać, że atrakcyjność CSH jako metody badania transkryptomu będzie malała. Również wyniki niniejszych badań wskazują, że mimo potencjału do generowania istotnych biologicznie i wiarygodnych wyników, uzyskane tą metodą dane wymagają dalszych, szczegółowych badań. Niemniej jednak, można wskazać inne możliwości wykorzystania techniki CSH w praktyce, świadczące o jej przydatności. W związku z intensywnym rozwojem genomicznych baz danych, coraz lepszym poznaniem funkcji poszczególnych genów oraz dużą liczbą danych mikromacierzowych zdeponowanych w bazach, dość atrakcyjne wydaje się użycie CSH w badaniach porównawczych z wynikami badań SSH, w kontekście poszukiwania podobieństw i różnic biologicznych pomiędzy gatunkami o pewnym stopniu pokrewieństwa. Wydaje się, że w najbliższych latach CSH powinno być stosowane jako badanie wstępne lub przesiewowe, pozwalające nakreślić ogólny profil ekspresji i wskazać potencjalnie zmienione miejsca w transkryptomie w porównaniu z istniejącymi gatunkami modelowymi. Dzięki takim badaniom możliwe będzie wskazanie gatunków, których profil ekspresji genów wykazuje różnice w ramach rodziny i które warto jest zbadać dokładniej

przy wykorzystaniu technik sekwencjonowania. W tym kontekście obecność różnic w sekwencji pomiędzy gatunkiem badanym a modelowym może być wręcz pożądana. Optymalne byłoby otrzymanie danych uzyskanych na tej samej mikromacierzy i w tych samych warunkach dla obu technik, co znacznie zmniejsza podatność metody na błędy eksperymentalne. Tego typu zastosowanie CSH w genomice, pod warunkiem prawidłowego przeprowadzenia badań i spełnienia nakreślonych wyżej warunków, wydaje się właściwym sposobem wykorzystania sporego potencjału tej techniki mikromacierzowej.

VI. WNIOSKI

Przeprowadzone badania porównawcze wpływu zastosowania metody CSH i SSH na efektywność detekcji zmian w obrazie transkryptomu *Nicotiana tabacum* pod wpływem stresu abiotycznego, pozwoliły na sformułowanie następujących wniosków:

1. CSH jest metodą umożliwiającą uzyskanie wiarygodnych danych transkryptomicznych, a stosowanie się do ustalonej w niniejszej pracy ścieżki postępowania zmniejsza ryzyko fałszywego wnioskowania.
2. Wyniki CSH o lepszej jakości i wyższym podobieństwie do danych referencyjnych uzyskano stosując mikromacierze cDNA, a gorsze gdy użyto mikromacierze oligonukleotydowe.
3. Niewielkie różnice w odległości filogenetycznej pomiędzy gatunkami nie mają dużego wpływu na jakość danych CSH. Istotniejsze znaczenie mają inne czynniki, takie jak liczba zastosowanych powtórzeń lub dostępność adnotacji funkcjonalnych.
4. Filtracja danych na podstawie homologii sekwencji ma istotny wpływ na wyniki jedynie w przypadku mikromacierzy o mniejszej liczbie sond wykazujących istotną homologię do sekwencji docelowych. W przypadku dużego podobieństwa pomiędzy sondami a transkryptami, wystarczające może być odfiltrowanie danych na podstawie parametrów statystycznych (wartości p i logFC).
5. Parametry jakości plamek wykazują związek ze stopniem homologii transkryptów wobec sond, jednak zależność ta nie jest liniowa. Ustalenie progu filtracji jest możliwe w przypadku zastosowania danych SSH z tej samej mikromacierzy jako referencji. Filtracja danych wyłącznie na podstawie charakterystyki morfologicznej nie wystarcza do uzyskania wyników o jakości porównywalnej z danymi filtrowanymi na podstawie homologii lub wyselekcjonowanymi wyłącznie na podstawie parametrów statystycznych.
6. Uzyskane wyniki eksperymentów na mikromacierzy tytoniowej pozwoliły na opisanie profilów ekspresji genów w korzeniach *Nicotiana tabacum* w stresie abiotycznym. Między innymi stwierdzono że:
 - a) Pod wpływem stresu solnego i kadmowego indukowane są liczne zmiany w obrębie transkryptomu korzeni tytoniu. Geny o zmienionej ekspresji należą do wielu różnych grup funkcjonalnych.
 - b) Większość wykrytych zmian jest zgodna z obserwowanymi wcześniej w korzeniach innych gatunków roślin (np. rzodkiewnika) pod wpływem stresu abiotycznego.
 - c) Nowo zidentyfikowane zjawiska (nieobserwowane do tej pory podczas stresu w tkance korzeniowej), takie jak zmiany w ścianie komórkowej w stresie kadmowym czy zmiany związane ze starzeniem w stresie solnym mogą przyczynić się do poszerzenia wiedzy na temat wpływu stresu na funkcjonowanie roślin.

VII.LITERATURA

1. Fodor SP, Read JL, Pirrung MC, Stryer L, Lu AT, Solas D. Light-directed, spatially addressable parallel chemical synthesis. *Science*. 1991 Feb 15;251(4995):767-73.
2. Fodor S. P. A., Read J. L., Pirrung M. C., Stryer L., Lu A., Solas D., Schena M., Shalon D., Davis RW., Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*. 1995 Oct 20;270(5235):368-371.
3. Schena M., Shalon D., Davis RW., Brown PO. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*. 1995 Oct 20;270(5235):368-371.
4. Lockhart D. J., Dong H., Byrne M. C., Follettie M. T., Gallo M. V., Chee M. S., Mittmann M., Wang C., Kobayashi M., Horton H., Brown E. L., Expression monitoring by hybridization to high-density oligonucleotide arrays. (1996), *Nat. Biotech.*, 14(13), 1675-1680.
5. Wodicka L., Dong H., Mittmann M., Ho M. H., Lockhart D. J., Genome-wide expression monitoring in *Saccharomyces cerevisiae* (1997), *Nat. Biotech.*, 15(13), 1359-1367
6. Lipshutz R. J., Fodor S. P., Gingeras T. R., Lockhart D. J., (1999), *Nat. Biotech.*, 21(1 Suppl), 20-24.
7. Moore MJ, Silver PA. Global analysis of mRNA splicing. *RNA*. 2008 Feb;14(2):197-203.
8. Ben-Dov C, Hartmann B, Lundgren J, Valcárcel J. Genome-wide analysis of alternative pre-mRNA splicing. *J Biol Chem*. 2008 Jan 18;283(3):1229-33.
9. Rauch A, Rüschenhoff F, Huang J, Trautmann U, Becker C, Thiel C, Jones KW, Reis A, Nürnberg P. Molecular karyotyping using an SNP array for genomewide genotyping. *J Med Genet*. 2004 Dec;41(12):916-22.
10. Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science*. 1992 Oct 30;258(5083):818-21.
11. B. Ren, F. Robert, J.J. Wyrick, O. Aparicio, E.G. Jennings, I. Simon, J. Zeitlinger, J. Schreiber, N. Hannett and E. Kanin et al., Genome-wide location and function of DNA binding proteins, *Science* 290 (2000), pp. 2306–2309
12. Hanlon SE, Lieb JD. Progress and challenges in profiling the dynamics of chromatin and transcription factor binding with DNA microarrays. *Curr Opin Genet Dev*. 2004 Dec;14(6):697-705.
13. Aharoni A, Vorst O. DNA microarrays for functional plant genomics. *Plant Mol Biol*. 2002 Jan;48(1-2):99-118.
14. Vorst, O., van Hal, N.L.W., Keijer, J. and Hall, R.D. 2001. cDNA microarrays as a tool for gene expression studies – theoretical and practical considerations. In: *Focus on Biotechnology 5: Plant Biotechnology*, Kluwer Academic Publishers, Dordrecht, Netherlands
15. Zhu B, Xu F, Baba Y. An evaluation of linear RNA amplification in cDNA microarray gene expression analysis. *Mol Genet Metab*. 2006 Jan;87(1):71-9. Epub 2005 Oct 7. PubMed PMID: 16214384.
16. Kaposi-Novak P, Lee JS, Mikaelyan A, Patel V, Thorgeirsson SS. Oligonucleotide microarray analysis of aminoallyl-labeled cDNA targets from linear RNA amplification. *Biotechniques*. 2004 Oct;37(4):580, 582-6, 588
17. Kiewe P, Gueller S, Komor M, Stroux A, Thiel E, Hofmann WK. Prediction of qualitative outcome of oligonucleotide microarray hybridization by measurement of RNA integrity using the 2100 Bioanalyzer capillary electrophoresis system. *Ann Hematol*. 2009 Dec;88(12):1177-83.
18. Gomes LI, Silva RL, Stolf BS, Cristo EB, Hirata R, Soares FA, Reis LF, Neves EJ, Carvalho AF. Comparative analysis of amplified and nonamplified RNA for hybridization in cDNA microarray. *Anal Biochem*. 2003 Oct 15;321(2):244-51.
19. Stears RL, Martinsky T, Schena M. Trends in microarray analysis. *Nat Med*. 2003 Jan;9(1):140-5
20. Mario Ouellet, Paul D Adams, Jay D Keasling, Aindrila Mukhopadhyay A rapid and inexpensive labeling method for microarray gene expression analysis *BMC Biotechnology* 2009, 9:97
21. Cowell JK, Hawthorn L. The application of microarray technology to the analysis of the cancer genome. *Curr Mol Med*. 2007 Feb;7(1):103-20.
22. Gupta V, Oliver B. *Drosophila* microarray mikromacierzys. *Brief Funct Genomic Proteomic*. 2003 Jul;2(2):97-105.
23. Draghici, S., Khatri, P., Eklund, A.C. and Szallasi, Z. (2006) *Trends Gene.*, 22, 101-109
24. Rouchka EC, Phatak AW, Singh AV. Effect of single nucleotide polymorphisms on Affymetrix match-mismatch probe pairs. *Bioinformatics*. 2008 Jul 14;24(9):405-11.
25. Kuhn K, Baker SC, Chudin E, Lieu MH, Oeser S, Bennett H, Rigault P, Barker D, McDaniel TK, Chee MS. A novel, high-performance random array mikromacierzy for quantitative gene expression profiling. *Genome Res*. 2004 Nov;14(11):2347-56.
26. Ku WC, Lau WK, Tseng YT, Tzeng CM, Chiu SK. Dextran sulfate provides a quantitative and quick microarray hybridization reaction. *Biochem Biophys Res Commun*. 2004 Feb 27;315(1):30-7
27. Stępniać P., Handschuh L., Figlerowicz M. Mikromacierze DNA – analiza danych, *Biotechnologia*. 2008, 4(83)

28. Sontrop HM, Moerland PD, van den Ham R, Reinders MJ, Verhaegh WF. A comprehensive sensitivity analysis of microarray breast cancer classification under feature variability. *BMC Bioinformatics*. 2009 Nov 26;10:389
29. Verducci JS, Melfi VF, Lin S, Wang Z, Roy S, Sen CK. Microarray analysis of gene expression: considerations in data mining and statistical treatment. *Physiol Genomics*. 2006 May 16;25(3):355-63.
30. W Rensink, CR Buell (2005) Microarray expression profiling resources for plant genomics. *TRENDS in Plant Science* 10, 12.
31. Degletagne C, Keime C, Rey B, de Dinechin M, Forcheron F, Chuchana P, Jouventin P, Gautier C, Duchamp C. Transcriptome analysis in non-model species: a new method for the analysis of heterologous hybridization on microarrays. *BMC Genomics*. 2010 May 31;11:344.
32. PM Schenk, SR Thomas-Hall, AV Nguyen, JM Manners, K Kazan, G Spangenberg (2007) Identification of plant defence genes in canola using *Arabidopsis* cDNA microarrays. *Plant Biology* 10, 539-547.
33. Chen Y, Wu R, Felton J, Rocke DM, Chakicherla A. A Method to Detect Differential Gene expression in Cross-Species Hybridization Experiments at Gene and Probe Level. *Biomed Inform Insights*. 2010 Mar 5;2010(3):1-10.
34. Ogura A, Yoshida MA, Fukuzaki M, Sese J. In vitro homology search array comprehensively reveals highly conserved genes and their functional characteristics in non-sequenced species. *BMC Genomics*. 2010 Dec 2;11 Suppl 4:S9.
35. J Malone, B Oliver Microarrays, deep sequencing and the true measure of the transcriptome, *BMC Biology* 2011,9:34
36. Auer PL, Doerge RW. Statistical design and analysis of RNA sequencing data. *Genetics*. 2010 Jun;185(2):405-16.
37. JP Hammond, MR Broadley, DJ Craigon, J Higgins, ZF Emmerson et al. (2005) Using genomic DNA-based probe –selection to improve the sensitivity of high –density oligonucleotide arrays when applied to heterologous species. *Plant Methods* 1,10
38. A Bagnaresi, A Moschella, O Beretta, F Vitulli, P Ranalli (2008) Heterologous microarray experiments allow the identification of the early events associated with potato tuber cold sweetening. *BMC Genomics* 9, 176.
- 39 Yee JC, Wlaschin KF, Chuah SH, Nissom PM, Hu WS. Quality assessment of cross-species hybridization of CHO transcriptome on a mouse DNA oligo microarray. *Biotechnol Bioeng*. 2008 Dec 15;101(6):1359-65.
- 40 S.C. Renn, N Aubin-Horth, HA Hoffmann (2004) Biologically meaningful expression profiling across species using heterologous hybridization to a cDNA microarray. *BMC Genomics* 5, 42.
41. Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR Jeanine S. Morey, 1 James C. Ryan,1 and Frances M. Van Dolah1 *Biol Proced Online*. 2006; 8: 175–193
42. Dallas PB, Gottardo NG, Firth MJ, Beesley AH, Hoffmann K, Terry PA, Freitas JR, Boag JM, Cummings AJ, Kees UR. Gene expression levels assessed by oligonucleotide microarray analysis and quantitative real-time RT-PCR -- how well do they correlate? *BMC Genomics*. 2005 Apr 27;6:59
43. Koltai H, Weingarten-Baror C. Specificity of DNA microarray hybridization: characterization, effectors and approaches for data correction. *Nucleic Acids Res*. 2008 Apr;36(7):2395-405.
44. Draghici S, Khatri P, Eklund AC, Szallasi Z. Reliability and reproducibility issues in DNA microarray measurements. *Trends Genet*. 2006 Feb;22(2):101-9. Epub 2005 Dec 27
45. W Rensink, CR Buell (2005) Microarray expression profiling resources for plant genomics. *TRENDS in Plant Science* 10, 12.
46. C Bar-Or, H Czosnek, H Koltai (2007) Cross-species microarray hybridizations: a developing tool for studying species diversity. *TRENDS in Genetics* 23, 4.
47. PM Schenk, SR Thomas-Hall, AV Nguyen, JM Manners, K Kazan, G Spangenberg (2007) Identification of plant defence genes in canola using *Arabidopsis* cDNA microarrays. *Plant Biology* 10, 539-547.
48. S.A. Rifkin (2003) Evolution in gene expression in the *Drosophila melanogaster* subgroup. *Nat. Genet*. 33, 138-144
49. CB Bigger (2001) DNA microarray analysis of chimpanzee liver during acute resolving hepatitis C virus infection. *J. Virol*. 75, 7059-7066.
50. J Adjaye, R Herwig, D Herrmann, W Wruck, A BenKahla et al. (2004) Cross-species hybridization of human and bovine orthologous genes on high density cDNA microarrays. *BMC Genomics* 5:83.
51. Y Gilad, A Oshlack, GK Smyth, TP Speed, KP White (2006) Expression profiling in primates reveals a rapid evolution of human transcription factors. *Nature* 440, 242-245.
52. C Bar-Or, E Novikov, A Reiner, H Czosnek, H Koltai (2007) Utilizing microarray spot characteristics to improve cross-species hybridization results. *Genomics* 90, 636-645.
53. S Moore, P Payton, M Wright, S Tanksley, J Giovannoni (2005) Utilization of tomato microarrays for comparative gene expression analysis in the Solanaceae. *Journal of Experimental Botany*, 56, 421, 2885-2895.

54. C Bar-Or, M Bar-Eyal, TZ Gal, Y Kapulnik, H Czosnek et al. (2006) Derivation of species-specific hybridization-like knowledge out of cross-species hybridization results. *BMC Genomics* 7, 110.
55. Y Gilad, SA Rifkin, P Bertone, M Gerstein, KP White (2005) Multi-species microarray reveal the effect of sequence divergence on gene expression profiles. *Genome Res.* 15, 674-680.
56. M Nowrousian, C Ringelberg, JC Dunlap, JJ Loros, U Kück (2005) Cross-species microarray hybridization to identify developmentally regulated genes in the filamentous fungus *Sordaria makrospora*. *Mol. Genet. Genomics* 273, 137-149.
57. P Saetre, J Lindberg, JA Leonard, K Olsson, U Pettersson et al. (2004) From wild wolf to domestic dog: gene expression changes the brain. *Brain Res. Mol. Brain Res.* 126, 198-206.
58. SJ Walker, Y Wang, KA Grant, F Chan, GM Hellmann (2006) Long versus short oligonucleotide microarrays for the study of gene expression in nonhuman primates. *J. Neurosci. Methods* 152, 179-189.
59. A Oshlack, A Chabot, G Smyth, Y Gilad (2007) Using DNA microarrays to study gene expression in closely related species. *Bioinformatics* 23, 10, 1235-1242.
60. F. Rutschmann, Molecular dating of phylogenetic trees: a brief review of current methods that estimate divergence times, *Divers. Distrib.* 12 (2006), pp. 35-48
61. Machado HE, Pollen AA, Hofmann HA, Renn SC. Interspecific profiling of gene expression informed by comparative genomic hybridization: A review and a novel approach in African cichlid fishes. *Integr Comp Biol.* 2009 Dec;49(6):644-59
62. Grigoryev DN, Ma SF, Simon BA, Irizarry RA, Ye SQ, Garcia JG. In vitro identification and in silico utilization of interspecies sequence similarities using GeneChip technology. *BMC Genomics.* 2005 May 4;6:62.
63. Tsoi SC, Cale JM, Bird IM, Ewart V, Brown LL, Douglas S. Use of human cDNA microarrays for identification of differentially expressed genes in Atlantic salmon liver during *Aeromonas salmonicida* infection. *Mar Biotechnol (NY).* 2003 Nov-Dec;5(6):545-54.
64. Jayaraman A, Hall CK, Genzer J. Computer simulation study of molecular recognition in model DNA microarrays. *Biophys. J.* 2006;91:2227-2236
65. Chen YA, Chou CC, Lu X, Slate EH, Peck K, Xu W, Voit EO, Almeida JS. A multivariate prediction model for microarray cross-hybridization. *BMC Bioinformatics.* 2006;7:101
66. Wu C, Carta R, Zhang L. Sequence dependence of cross-hybridization on short oligo microarrays. *Nucleic Acids Res.* 2005;33:e84.
67. Wick LM, Rouillard JM, Whittam TS, Gulari E, Tiedje JM, Hashsham SA. On-chip non-equilibrium dissociation curves and dissociation rate constants as methods to assess specificity of oligonucleotide probes. *Nucleic Acids Res.* 2006;34:e26.
68. Dai H, Meyer M, Stepaniants S, Ziman M, Stoughton R. Use of hybridization kinetics for differentiating specific from non-specific binding to oligonucleotide microarrays. *Nucleic Acids Res.* 2002;30:e86
69. Zhang Y, Hammer DA, Graves DJ. Competitive hybridization kinetics reveals unexpected behavior patterns. *Biophys. J.* 2005;89:2950-2959
70. Schaupp CJ, Jiang G, Myers TG, Wilson MA. Active mixing during hybridization improves the accuracy and reproducibility of microarray results. *Biotechniques.* 2005;38:117-119.
71. Novikov E, Barillot E. Software package for automatic microarray image analysis (MAIA). *Bioinformatics.* 2007 Mar 1;23(5):639-40. Epub 2007 Jan 19.
72. Pepke S, Wold B, Mortazavi A. Computation for ChIP-seq and RNA-seq studies. *Nat Methods.* 2009 Nov;6(11 Suppl):S22-32.
73. Lu Y, Huggins P, Bar-Joseph Z. Cross species analysis of microarray expression data. *Bioinformatics.* 2009 Jun 15;25(12):1476-83
74. Darby BJ, Jones KL, Wheeler D, Herman MA. Normalization and centering of array-based heterologous genome hybridization based on divergent control probes. *BMC Bioinformatics.* 2011 May 21;12:183
75. Mecham BH, Klus GT, Strovel J, Augustus M, Byrne D, Bozso P, Wetmore DZ, Mariani TJ, Kohane IS, Szallasi Z. Sequence-matched probes produce increased cross-mikromacierzy consistency and more reproducible biological results in microarray-based gene expression measurements. *Nucleic Acids Res.* 2004 May 25;32(9):e74
76. Ranz JM, Namgyal K, Gibson G, Hartl DL. Anomalies in the expression profile of interspecific hybrids of *Drosophila melanogaster* and *Drosophila simulans*. *Genome Res.* 2004 Mar;14(3):373-9
77. Enard W, Khaitovich P, Klose J, Zöllner S, Heissig F, Giavalisco P, Nieselt-Struwe K, Muchmore E, Varki A, Ravid R, Doxiadis GM, Bontrop RE, Pääbo S. Intra- and interspecific variation in primate gene expression patterns. *Science.* 2002 Apr 12;296(5566):340-3
78. Darby BJ, Jones KL, Wheeler D, Herman MA. Normalization and centering of array-based heterologous genome hybridization based on divergent control probes. *BMC Bioinformatics.* 2011 May 21;12:183
79. Machado HE, Renn SC. A critical assessment of cross-species detection of gene duplicates using comparative genomic hybridization. *BMC Genomics.* 2010 May 13;11:304.

80. Renn SC, Machado HE, Jones A, Soneji K, Kulathinal RJ, Hofmann HA. Using comparative genomic hybridization to survey genomic sequence divergence across species: a proof-of-concept from *Drosophila*. *BMC Genomics*. 2010 Apr 29;11:271.
81. Davey MW, Graham NS, Vanholme B, Swennen R, May ST, Keulemans J. Heterologous oligonucleotide microarrays for transcriptomics in a non-model species; a proof-of-concept study of drought stress in *Musa*. *BMC Genomics*. 2009 Sep 16;10:436
82. Senthil G, Liu H, Puram VG, Clark A, Stromberg A, Goodin MM. Specific and common changes in *Nicotiana benthamiana* gene expression in response to infection by enveloped viruses. *J Gen Virol*. 2005 Sep;86(Pt 9):2615-25.
83. Tomiuk S, Hofmann K. Microarray probe selection strategies. *Brief Bioinform*. 2001 Dec;2(4):329-40.
84. Bar-Akiva A, Ovadia R, Rogachev I, Bar-Or C, Bar E, Freiman Z, Nissim-Levi A, Gollop N, Lewinsohn E, Aharoni A, Weiss D, Koltai H, Oren-Shamir M. Metabolic networking in *Brunfelsia calycina* petals after flower opening. *J Exp Bot*. 2010 Mar;61(5):1393-403.
85. Chalmers AD, Goldstone K, Smith JC, Gilchrist M, Amaya E, Papalopulu N. A *Xenopus tropicalis* oligonucleotide microarray works across species using RNA from *Xenopus laevis*. *Mech Dev*. 2005 Mar;0122(3):355-63. Epub 2004 Oct 26
86. Kaposi-Novak P, Lee JS, Mikaelyan A, Patel V, Thorgeirsson SS. Oligonucleotide microarray analysis of aminoallyl-labeled cDNA targets from linear RNA amplification. *Biotechniques*. 2004 Oct;37[4]:580, 582-6, 588.
87. Xu W, Bak S, Decker A, Paquette SM, Feyereisen R, Galbraith DW. Microarray-based analysis of gene expression in very large gene families: the cytochrome P450 gene superfamily of *Arabidopsis thaliana*. *Gene*. 2001 Jul 11;272[1-2]:61-74.
88. Xing Y, Kapur K, Wong WH. Probe selection and expression index computation of Affymetrix Exon Arrays. *PLoS One*. 2006 Dec 20;1:e88
89. Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature*. 2005 Sep 1;437(7055):69-87.
90. Becher M., Talke IN., Krall L., Krämer U. Cross-species microarray transcript profiling reveals high constitutive expression of metal homeostasis genes in shoots of the zinc hyperaccumulator *Arabidopsis halleri*. *Plant J*.37(2). 2004:251-68
91. Zmieńko A, Guzowska-Nowowiejska M, Urbaniak R, Pląder W, Formanowicz P, Figlerowicz M. A tiling microarray for global analysis of chloroplast genome expression in cucumber and other plants. *Plant Methods*. 2011 Sep 28;7:29.
92. Rensink WA, Lee Y, Liu J, Iobst S, Ouyang S, Buell CR. Comparative analyses of six solanaceous transcriptomes reveal a high degree of sequence conservation and species-specific transcripts. *BMC Genomics*. 2005 Sep 14;6:124.
93. Smyth, G. K. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology*, 2004, Vol. 3, No. 1, Article 3
94. Ciarmiello L, Woodrow P., Fuggi A., Pontecorvo G., Carillo P. *Plant Genes for Abiotic Stress, Abiotic Stress in Plants - Mechanisms and Adaptations*, Prof. Arun Shanker (Ed.), ISBN: 978-953-307-394-1, InTech
95. Jiang Y, Deyholos MK. Comprehensive transcriptional profiling of NaCl-stressed *Arabidopsis* roots reveals novel classes of responsive genes. *BMC Plant Biol*. 2006 Oct 12;6:25.
96. Kreps JA, Wu Y, Chang HS, Zhu T, Wang X, Harper JF. Transcriptome changes for *Arabidopsis* in response to salt, osmotic, and cold stress. *Plant Physiol*. 2002 Dec;130(4):2129-41
97. Zhang M, Liu X, Yuan L, Wu K, Duan J, Wang X, Yang L. Transcriptional profiling in cadmium-treated rice seedling roots using suppressive subtractive hybridization. *Plant Physiol Biochem*. 2012 Jan;50(1):79-86.
98. Weber M, Trampczynska A, Clemens S. Comparative transcriptome analysis of toxic metal responses in *Arabidopsis thaliana* and the Cd(2+)-hypertolerant facultative metallophyte *Arabidopsis halleri*. *Plant Cell Environ*. 2006 May;29(5):950-63.
99. Yao D, Zhang X, Zhao X, Liu C, Wang C, Zhang Z, Zhang C, Wei Q, Wang Q, Yan H, Li F, Su Z. Transcriptome analysis reveals salt-stress-regulated biological processes and key pathways in roots of cotton (*Gossypium hirsutum* L.). *Genomics*. 2011 Jul;98(1):47-55.
100. Zhao CR, Ikka T, Sawaki Y, Kobayashi Y, Suzuki Y, Hibino T, Sato S, Sakurai N, Shibata D, Koyama H. Comparative transcriptomic characterization of aluminum, sodium chloride, cadmium and copper rhizotoxicities in *Arabidopsis thaliana*. *BMC Plant Biol*. 2009 Mar 23;9:32
101. Raman T, O'Connor TP, Hackett NR, Wang W, Harvey BG, Attiyeh MA, Dang DT, Teater M, Crystal RG. Quality control in microarray assessment of gene expression in human airway epithelium. *BMC Genomics*. 2009 Oct 24;10:493
102. Schroeder A, Mueller O, Stocker S, Salowsky R, Leiber M, Gassmann M, Lightfoot S, Menzel W, Granzow M, Ragg T. The RIN: an RNA integrity number for assigning integrity values to RNA measurements

103. PAGE: Parametric Analysis of Gene Set Enrichment Seon-Young Kim, David J Volsky
BMC Bioinformatics. 2005; 6: 144
104. Du Z, Zhou X, Ling Y, Zhang Z, Su Z. agriGO: a GO analysis toolkit for the agricultural community. Nucleic Acids Res. 2010 Jul;38(Web Server issue):W64-70.
105. Thimm O, Blaesing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY and M Stitt (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J. 37(6):914-39.
106. Li J, Besseau S, Törönen P, Sipari N, Kollist H, Holm L, Palva ET. Defense-related transcription factors WRKY70 and WRKY54 modulate osmotic stress tolerance by regulating stomatal aperture in Arabidopsis. New Phytol. 2013 Oct;200(2):457-72.
107. Yang Z, Tian L, Latoszek-Green M, Brown D, Wu K. Arabidopsis ERF4 is a transcriptional repressor capable of modulating ethylene and abscisic acid responses. Plant Mol Biol. 2005 Jul;58(4):585-96.
108. Fujimoto SY, Ohta M, Usui A, Shinshi H, Ohme-Takagi M. Arabidopsis ethylene-responsive element binding factors act as transcriptional activators or repressors of GCC box-mediated gene expression. Plant Cell. 2000 Mar;12(3):393-404
109. Mishina TE, Zeier J. The Arabidopsis flavin-dependent monooxygenase FMO1 is an essential component of biologically induced systemic acquired resistance. Plant Physiol. 2006 Aug;141(4):1666-75
110. Cheng H, Wang J, Chu S, Yan HL, Yu D. Diversifying selection on flavanone 3-hydroxylase and isoflavone synthase genes in cultivated soybean and its wild progenitors. PLoS One. 2013;8(1):e54154.
111. Chengjian X, De W, Fei K, Dan K, Xingyong Y, The Pathogen and Wound Induces Expression of Genes Related to Proanthocyanidins (PAs) Synthesis in Cotton Leaves; American Journal of Plant Sciences; Feb 2012, Vol. 3 Issue 2, p228
112. Cluis CP, Mouchel CF, Hardtke CS. The Arabidopsis transcription factor HY5 integrates light and hormone signaling pathways. Plant J. 2004 Apr;38(2):332-47.
113. Chen H, Xiong L. Role of HY5 in abscisic acid response in seeds and seedlings. Plant Signal Behav. 2008 Nov;3(11):986-8.
114. Iida A, Kazuoka T, Torikai S, Kikuchi H, Oeda K. A zinc finger protein RHL41 mediates the light acclimatization response in Arabidopsis. Plant J. 2000 Oct;24(2):191-203
115. Tognetti VB, Van Aken O, Morreel K, Vandenbroucke K, van de Cotte B, De Clercq I, Chiwocha S, Fenske R, Prinsen E, Boerjan W, Genty B, Stubbs KA, Inzé D, Van Breusegem F. Perturbation of indole-3-butyric acid homeostasis by the UDP-glucosyltransferase UGT74E2 modulates Arabidopsis architecture and water stress tolerance. Plant Cell. 2010 Aug;22(8):2660-79
116. Kimura M, Manabe K, Abe T, Yoshida S, Matsui M, Yamamoto YY. Analysis of hydrogen peroxide-independent expression of the high-light-inducible ELIP2 gene with the aid of the ELIP2 promoter-luciferase fusions. Photochem Photobiol. 2003 Jun;77(6):668-74.
117. Xu L, Ménard R, Berr A, Fuchs J, Cognat V, Meyer D, Shen WH. The E2 ubiquitin-conjugating enzymes, AtUBC1 and AtUBC2, play redundant roles and are involved in activation of FLC expression and repression of flowering in Arabidopsis thaliana. Plant J. 2009 Jan;57(2):279-88.
118. Vainonen JP, Jaspers P, Wrzaczek M, Lamminmäki A, Reddy RA, Vaahtera L, Brosché M, Kangasjärvi J. RCD1-DREB2A interaction in leaf senescence and stress responses in Arabidopsis thaliana. Biochem J. 2012 Mar 15;442(3):573-81.
119. Yamaguchi-Shinozaki K, Shinozaki K. Improving plant drought, salt and freezing tolerance by gene transfer of a single stress-inducible transcription factor. Novartis Found Symp. 2001;236:176-86; discussion 186-9.
120. Achard P, Gong F, Cheminant S, Alioua M, Hedden P, Genschik P. The cold-inducible CBF1 factor-dependent signaling pathway modulates the accumulation of the growth-repressing DELLA proteins via its effect on gibberellin metabolism. Plant Cell. 2008 Aug;20(8):2117-29.
121. Kim S, Kang JY, Cho DI, Park JH, Kim SY. ABF2, an ABRE-binding bZIP factor, is an essential component of glucose signaling and its overexpression affects multiple stress tolerance. Plant J. 2004 Oct;40(1):75-87. Erratum in: Plant J. 2005 Aug;43(3):467.
122. Roxas VP, Lodhi SA, Garrett DK, Mahan JR, Allen RD. Stress tolerance in transgenic tobacco seedlings that overexpress glutathione S-transferase/glutathione peroxidase. Plant Cell Physiol. 2000 Nov;41(11):1229-34.
123. Edwards R, Dixon DP, Walbot V. Plant glutathione S-transferases: enzymes with multiple functions in sickness and in health. Trends Plant Sci. 2000 May;5(5):193-8.
124. Li S, Fu Q, Chen L, Huang W, Yu D. Arabidopsis thaliana WRKY25, WRKY26, and WRKY33 coordinate induction of plant thermotolerance. Planta. 2011 Jun;233(6):1237-52.
125. Heavy Metal Stress. Activation of Distinct Mitogen-Activated Protein Kinase Pathways by Copper and Cadmium Claudia Jonak, Hirofumi Nakagami, Heribert Hirt Plant Physiol. 2004 October; 136(2): 3276–3283.

- 126 Bray EA. Genes commonly regulated by water-deficit stress in *Arabidopsis thaliana*. *J Exp Bot.* 2004 Nov;55(407):2331-41.
127. Kariola T, Brader G, Helenius E, Li J, Heino P, Palva ET. early responsive to dehydration 15, a negative regulator of abscisic acid responses in *Arabidopsis*. *Plant Physiol.* 2006 Dec;142(4):1559-73.
- 128 Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K. ERD15, a cDNA for a dehydration-induced gene from *Arabidopsis thaliana*. *Plant Physiol.* 1994 Dec;106(4):1707.
129. Moons A (2003a) OsPDR9 which encodes a PDR-type ABC transporter is induced by heavy metals, hypoxic stress and redox perturbations in rice roots. *FEBS Lett* 553:370–376
- 130 McClung CR. Regulation of catalases in *Arabidopsis*. *Free Radic Biol Med.*1997;23(3):489-96.
- 131 Wang W, Vinocur B, Shoseyov O, Altman A. Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci.* 2004 May;9(5):244-52.
- 132 Kovtun Y, Chiu WL, Tena G, Sheen J. Functional analysis of oxidative stress-activated mitogen-activated protein kinase cascade in plants. *Proc Natl Acad Sci U S A.* 2000 Mar 14;97(6):2940-5.
133. Jiang Y, Yang B, Harris NS, Deyholos MK. Comparative proteomic analysis of NaCl stress-responsive proteins in *Arabidopsis* roots. *J Exp Bot.* 2007;58(13):3591-607.
- 135 Sarry JE, Kuhn L, Ducruix C, Lafaye A, Junot C, Hugouvieux V, Jourdain A, Bastien O, Fievet JB, Vailhen D, Amekraz B, Moulin C, Ezan E, Garin J, Bourguignon J. The early responses of *Arabidopsis thaliana* cells to cadmium exposure explored by protein and metabolite profiling analyses. *Proteomics.* 2006 Apr;6(7):2180-98.
- 136 Nishizawa A, Yabuta Y, Yoshida E, Maruta T, Yoshimura K, Shigeoka S. *Arabidopsis* heat shock transcription factor A2 as a key regulator in response to several types of environmental stress. *Plant J.* 2006 Nov;48(4):535-47.
- 137 Delk NA, Johnson KA, Chowdhury NI, Braam J. CML24, regulated in expression by diverse stimuli, encodes a potential Ca²⁺ sensor that functions in responses to abscisic acid, daylength, and ion stress. *Plant Physiol.* 2005 Sep;139(1):240-53.
- 138 Cantero A, Barthakur S, Bushart TJ, Chou S, Morgan RO, Fernandez MP, Clark GB, Roux SJ. Expression profiling of the *Arabidopsis* annexin gene family during germination, de-etiolation and abiotic stress. *Plant Physiol Biochem.* 2006 Jan;44(1):13-24.
- 139 Reddy AS, Ali GS, Celesnik H, Day IS. Coping with stresses: roles of calcium- and calcium/calmodulin-regulated gene expression. *Plant Cell.* 2011 Jun;23(6):2010-32. doi: 10.1105/tpc.111.084988. Epub 2011 Jun 3. Review.
- 140 Porta H, Rocha-Sosa M. Plant lipoxygenases. Physiological and molecular features. *Plant Physiol.* 2002 Sep;130(1):15-21. Review.
- 141 Kravchik M, Bernstein N. Effects of salinity on the transcriptome of growing maize leaf cells point at cell-age specificity in the involvement of the antioxidative response in cell growth restriction. *BMC Genomics.* 2013 Jan 16;14:24.
- 142 Ouyang B, Yang T, Li H, Zhang L, Zhang Y, Zhang J, Fei Z, Ye Z. Identification of early salt stress response genes in tomato root by suppression subtractive hybridization and microarray analysis. *J Exp Bot.* 2007;58(3):507-20
- 143 Martínez-Ballesta MC, Aparicio F, Pallás V, Martínez V, Carvajal M. Influence of saline stress on root hydraulic conductance and PIP expression in *Arabidopsis*.
- 144 Liu Y, Xiong Y, Bassham DC. Autophagy is required for tolerance of drought and salt stress in plants. *Autophagy.* 2009 Oct;5(7):954-63.
- 145 Gong Z, Koiwa H, Cushman MA, Ray A, Bufford D, Kore-eda S, Matsumoto TK, Zhu J, Cushman JC, Bressan RA, Hasegawa PM. Genes that are uniquely stress regulated in salt overly sensitive (sos) mutants. *Plant Physiol.* 2001 May;126(1):363-75.
- 146 Llorente F, López-Cobollo RM, Catalá R, Martínez-Zapater JM, Salinas J. A novel cold-inducible gene from *Arabidopsis*, RCI3, encodes a peroxidase that constitutes a component for stress tolerance. *Plant J.* 2002 Oct;32(1):13-24.
147. Attia H, Arnaud N, Karray N, Lachaâl M. Long-term effects of mild salt stress on growth, ion accumulation and superoxide dismutase expression of *Arabidopsis* rosette leaves. *Physiol Plant.* 2008 Mar;132(3):293-305.
148. López-Millán AF, Sagardoy R, Solanas M, Abadía A, Abadía J. Cadmium toxicity in tomato (*Lycopersicon esculentum*) plants grown in hydroponics. *Environ Exp Bot* 2009; 65: 376-85.
149. I. Ogawa, H. Nakanishi, S. Mori, and N. K. Nishizawa, "Time course analysis of gene regulation under cadmium stress in rice, *Plant and Soil*, vol. 325, no. 1-2, pp. 97-108, Dec. 2009.
150. Expression and RNA splicing of the maize glutathione S-transferase Bronze2 gene is regulated by cadmium and other stresses. *K A Marrs, V Walbot Plant Physiol.* 1997 January; 113(1): 93–102
151. Ma S, Gong Q, Bohnert HJ. Dissecting salt stress pathways. *J Exp Bot.* 2006;57(5):1097-107.
152. de las Mercedes Dana M, Pintor-Toro JA, Cubero B. Transgenic tobacco plants overexpressing chitinases of fungal origin show enhanced resistance to biotic and abiotic stress agents. *Plant Physiol.*2006 Oct;142(2):722-30.

153. Kim YH, Bae JM, Huh GH. Transcriptional regulation of the cinnamyl alcohol dehydrogenase gene from sweet potato in response to plant developmental stage and environmental stress. *Plant Cell Rep.* 2010 Jul;29(7):779-91.
154. Knepper C, Savory EA, Day B. Arabidopsis NDR1 is an integrin-like protein with a role in fluid loss and plasma membrane-cell wall adhesion. *Plant Physiol.* 2011 May;156(1):286-300
155. Tuteja N, Mahajan S. Calcium signaling network in plants: an overview. *Plant Signal Behav.* 2007 Mar;2(2):79-85
156. Falara V, Akhtar TA, Nguyen TT, Spyropoulou EA, Bleeker PM, Schauvinhold I, Matsuba Y, Bonini ME, Schillmiller AL, Last RL, Schuurink RC, Pichersky E. The tomato terpene synthase gene family. *Plant Physiol.* 2011 Oct;157(2):770-89
157. Nieuwland J, Feron R, Huisman BA, Fasolino A, Hilbers CW, Derksen J, Mariani C. Lipid transfer proteins enhance cell wall extension in tobacco. *Plant Cell.* 2005 Jul;17(7):2009-19.
158. Geilfus CM, Zörb C, Mühlhng KH. Salt stress differentially affects growth-mediating β -expansins in resistant and sensitive maize (*Zea mays* L.). *Plant Physiol Biochem.* 2010 Dec;48(12):993-8
159. Lindemose S, O'Shea C, Jensen MK, Skriver K. Structure, function and networks of transcription factors involved in abiotic stress responses. *Int J Mol Sci.* 2013 Mar 13;14(3):5842-78
160. Nakashima K, Ito Y, Yamaguchi-Shinozaki K. Transcriptional regulatory networks in response to abiotic stresses in Arabidopsis and grasses. *Plant Physiol.* 2009 Jan;149(1):88-95
161. Jensen MK, Lindemose S, de Masi F, Reimer JJ, Nielsen M, Perera V, Workman CT, Turck F, Grant MR, Mundy J, Petersen M, Skriver K. ATAF1 transcription factor directly regulates abscisic acid biosynthetic gene NCED3 in Arabidopsis thaliana. *FEBS Open Bio.* 2013 Jul 29;3:321-7
162. Ludwig AA, Romeis T, Jones JD. CDPK-mediated signalling pathways: specificity and cross-talk. *J Exp Bot.* 2004 Jan;55(395):181-8
163. Ghanem ME, Albacete A, Martínez-Andújar C, Acosta M, Romero-Aranda R, Dodd IC, Lutts S, Pérez-Alfocea F. Hormonal changes during salinity-induced leaf senescence in tomato (*Solanum lycopersicum* L.). *J Exp Bot.* 2008;59(11):3039-50.
164. Salinity Effects on Root Growth and Senescence in Tomato and the Consequences for Severity of Phytophthora Root Rot Infection S.S. Snapp1 and C. Shennan JASHS May 1994 vol. 119 no. 3 458-463
165. Geilfus CM, Zörb C, Mühlhng KH. Salt stress differentially affects growth-mediating β -expansins in resistant and sensitive maize (*Zea mays* L.). *Plant Physiol Biochem.* 2010 Dec;48(12):993-8. doi: 10.1016/j.plaphy.2010.09.011
166. Guo W, Zhao J, Li X, Qin L, Yan X, Liao H. A soybean β -expansin gene GmEXPB2 intrinsically involved in root system architecture responses to abiotic stresses. *Plant J.* 2011 May;66(3):541-52
167. Zagorchev L, Kamenova P, Odjakova M. The Role of Plant Cell Wall Proteins in Response to Salt Stress. *ScientificWorldJournal.* 2014 Jan 19;2014:764089..
168. Xiong L, Lee H, Ishitani M, Zhu JK. Regulation of osmotic stress-responsive gene expression by the LOS6/ABA1 locus in Arabidopsis. *J Biol Chem.* 2002 Mar 8;277(10):8588-96
169. Zmienko A., Jackowiak P., Figlerowicz M.. Transcriptome sequencing : next generation approach to RNA functional analysis, *Biotechnologia,* 2011, 92(4), 311–319.
170. Hua J, Sakai H, Nourizadeh S, Chen QG, Bleeker AB, Ecker JR, Meyerowitz EM. EIN4 and ERS2 are members of the putative ethylene receptor gene family in Arabidopsis. *Plant Cell.* 1998 Aug;10(8):1321-32.
171. Kagaya Y, Hattori T. Arabidopsis transcription factors, RAV1 and RAV2, are regulated by touch-related stimuli in a dose-dependent and biphasic manner. *Genes Genet Syst.* 2009 Feb;84(1):95-9.
172. Cavalcanti FR, Lima JP, Ferreira-Silva SL, Viégas RA, Silveira JA. Roots and leaves display contrasting oxidative response during salt stress and recovery in cowpea. *J Plant Physiol.* 2007 May;164(5):591-600
173. Lin YF, Aarts MG. The molecular mechanism of zinc and cadmium stress response in plants. *Cell Mol Life Sci.* 2012 Oct;69(19):3187-206.
174. Skórzyńska-Polit E., Pawlikowska-Pawłęga B., Szczuka E., Drażkiewicz M., Krupa, Z. The Activity and Localization of Lipoxygenases in Arabidopsis thaliana under Cadmium and Copper Stresses. *Plant Growth Regulation,* 2006, 48(1), 29–39.
175. Kumar, S. P., Varman, P. A. M., & Kumari, B. D. R. Identification of differentially expressed proteins in response to Pb stress in *Catharanthus roseus*, 2011 5(September), 689–699.
176. DalCorso G, Farinati S, Furini A. Regulatory networks of cadmium stress in plants. *Plant Signal Behav.* 2010 Jun;5(6):663-7.
177. Iriti M, Faoro F. Chemical diversity and defence metabolism: how plants cope with pathogens and ozone pollution. *Int J Mol Sci.* 2009 Jul 30;10(8):3371-99.
178. Koprivova A, Kopriva S. Lessons from investigation of regulation of APS reductase by salt stress. *Plant Signal Behav.* 2008 Aug;3(8):567-9.
179. Balusamy SR, Kim YJ, Rahimi S, Senthil KS, Lee OR, Lee S, Yang DC. Transcript pattern of cytochrome P450, antioxidant and ginsenoside biosynthetic pathway genes under heavy metal stress in *Panax ginseng* Meyer. *Bull Environ Contam Toxicol.* 2013 Feb;90(2):194-202

180. Wan L, Zhang H. Cadmium toxicity: effects on cytoskeleton, vesicular trafficking and cell wall construction. *Plant Signal Behav.* 2012 Mar;7(3):345-8.
181. Martinoia E, Maeshima M, Neuhaus HE. Vacuolar transporters and their essential role in plant metabolism. *J Exp Bot.* 2007;58(1):83-102.
182. Kabała K. Struktura i regulacja aktywności roślinnych wakuolarnych H⁺-ATPaz, *Postępy Biologii Komórki*, tom 38, 2011, nr 3 (517-532)
183. Kim DY, Bovet L, Kushnir S, Noh EW, Martinoia E, Lee Y. AtATM3 is involved in heavy metal resistance in Arabidopsis. *Plant Physiol.* 2006 Mar;140(3):922-32.
184. Groppa MD, Tomaro ML, Benavides MP. Polyamines and heavy metal stress: the antioxidant behavior of spermine in cadmium- and copper-treated wheat leaves. *Biometals.* 2007 Apr;20(2):185-95
185. Hansey CN, Vaillancourt B, Sekhon RS, de Leon N, Kaepler SM, Buell CR. Maize (*Zea mays* L.) genome diversity as revealed by RNA-sequencing. *PLoS One.* 2012;7(3)
186. Garg AK, Kim JK, Owens TG, Ranwala AP, Choi YD, Kochian LV, Wu RJ. Trehalose accumulation in rice plants confers high tolerance levels to different abiotic stresses. *Proc Natl Acad Sci U S A.* 2002 Dec 10;99(25):15898-903
187. Denness L, McKenna JF, Segonzac C, Wormit A, Madhou P, Bennett M, Mansfield J, Zipfel C, Hamann T. Cell wall damage-induced lignin biosynthesis is regulated by a reactive oxygen species- and jasmonic acid-dependent process in Arabidopsis. *Plant Physiol.* 2011 Jul;156(3):1364-74
188. Knight H, Knight MR. Abiotic stress signalling pathways: specificity and cross-talk. *Trends Plant Sci.* 2001 Jun;6(6):262-7.
189. Lin CY, Trinh NN, Fu SF, Hsiung YC, Chia LC, Lin CW, Huang HJ. Comparison of early transcriptome responses to copper and cadmium in rice roots. *Plant Mol Biol.* 2013 Mar;81(4-5):507-22.
190. Yamaguchi H, Fukuoka H, Arao T, Ohyama A, Nunome T, Miyatake K, Negoro S. Gene expression analysis in cadmium-stressed roots of a low cadmium-accumulating solanaceous plant, *Solanum torvum*. *J Exp Bot.* 2010;61(2):423-37.
191. Moons A, Prinsen E, Bauw G, Van Montagu M. Antagonistic effects of abscisic acid and jasmonates on salt stress-inducible transcripts in rice roots. *Plant Cell.* 1997 Dec;9(12):2243-59
192. Barrero JM, Rodríguez PL, Quesada V, Piqueras P, Ponce MR, Micol JL. Both abscisic acid (ABA)-dependent and ABA-independent pathways govern the induction of NCED3, AAO3 and ABA1 in response to salt stress. *Plant Cell Environ.* 2006 Oct;29(10):2000-8.
193. Kiyosue T, Yamaguchi-Shinozaki K, Shinozaki K. ERD15, a cDNA for a dehydration-induced gene from Arabidopsis thaliana. *Plant Physiol.* 1994 Dec;106(4):1707.
194. Balazadeh S, Wu A, Mueller-Roeber B. Salt-triggered expression of the ANAC092-dependent senescence regulon in Arabidopsis thaliana. *Plant Signal Behav.* 2010 Jun;5(6):733-5
195. Ji H, Pardo JM, Batelli G, Van Oosten MJ, Bressan RA, Li X. The Salt Overly Sensitive (SOS) pathway: established and emerging roles. *Mol Plant.* 2013 Mar;6(2):275-86.
196. Oomen RJ, Wu J, Lelièvre F, Blanchet S, Richaud P, Barbier-Brygoo H, Aarts MG, Thomine S. Functional characterization of NRAMP3 and NRAMP4 from the metal hyperaccumulator *Thlaspi caerulescens*. *New Phytol.* 2009;181(3):637-50.
197. Kane MD, Jatkoa TA, Stumpf CR, Lu J, Thomas JD, Madore SJ. Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic Acids Res.* 2000 Nov 15;28(22):4552-7.
198. Kawasaki S, Borchert C, Deyholos M, Wang H, Brazille S, Kawai K, Galbraith D, Bohnert HJ. Gene expression profiles during the initial phase of salt stress in rice. *Plant Cell.* 2001 Apr;13(4):889-905.
199. Lu, S., Li, T., & Jiang, J. (2010). Effects of salinity on sucrose metabolism during tomato fruit development, 9(6), 842–849.
200. Rosenzweig, B. a., Pine, P. S., Domon, O. E., Morris, S. M., Chen, J. J., & Sistare, F. D. (2004). Dye Bias Correction in Dual-labeled cDNA Microarray Gene Expression Measurements. *Environmental Health Perspectives*, (4), 480–487.
201. Altman N. Replication, variation and normalisation in microarray experiments. *Appl Bioinformatics.* 2005;4(1):33-44.
202. Stuart JM, Segal E, Koller D, Kim SK. A gene-coexpression network for global discovery of conserved genetic modules. *Science.* 2003 Oct 10;302(5643):249-55.
203. Lu Y, Huggins P, Bar-Joseph Z. Cross species analysis of microarray expression data. *Bioinformatics.* 2009 Jun 15;25(12):1476-83.
204. Venet, D., Detours, V., & Bersini, H. (2012). A measure of the signal-to-noise ratio of microarray samples and studies using gene correlations. *PLoS one*, 7(12), e51013.

STRESZCZENIE

„Zastosowanie mikromacierzowej hybrydyzacji międzygatunkowej w badaniach transkryptomu *Nicotiana tabacum*”

Mikromacierzowa hybrydyzacja międzygatunkowa (ang. cross-species hybridization- CSH) jest metodą badawczą umożliwiającą ocenę ekspresji genów w przypadku braku dostępności mikromacierzy homologicznej dla gatunku badanego. Metoda jest powszechnie stosowana, jednak uważana za niestandardowe podejście ze względu na brak precyzyjnych wytycznych dotyczących projektowania i wykonania eksperymentu, a także analizy danych. Dotychczasowe, bardzo nieliczne badania na temat CSH nie dawały przekonującej odpowiedzi na pytania o zasadność stosowania tej metody.

Podstawowym celem niniejszej pracy była zatem kompleksowa ocena możliwości wykorzystania technik CSH oraz wyselekcjonowanie czynników wykazujących najistotniejszy wpływ na wynik tego typu analiz.

Zgodnie z przyjętymi założeniami osiągnięcie tak zarysowanego celu wymagało realizacji następujących zadań:

- ustalenia, w jaki sposób homologia pomiędzy genami gatunku badanego i reprezentowanego na mikromacierzy wpływa na wyniki CSH;
- określenia wpływu długości i rodzaju stosowanych sond na rezultaty badań
- opracowania metodyki analizy danych, ze szczególnym uwzględnieniem wyboru kryteriów wstępnej filtracji danych.

Równocześnie przyjęto, iż punktem odniesienia, pozwalającym na ocenę znaczenia poszczególnych czynników w hybrydyzacji CSH, powinny być wyniki uzyskane przy pomocy hybrydyzacji gatunkowo-specyficznej.

Jako model badawczy wybrano tytoń, reprezentujący rodzinę *Solanaceae*, z uwagi na wysoki stopień podobieństwa pomiędzy genami roślin należących do tej rodziny, dostępność różnych typów mikromacierzy dla kilku gatunków roślin, znaczenie gospodarcze wielu roślin psiankowatych oraz znaczną liczbę publikacji dotyczących zastosowania CSH w tej grupie roślin.

Dzięki użyciu różnych typów mikromacierzy (oligonukleotydowe, cDNA) oraz układów hybrydyzacji międzygatunkowej (tytoń: ziemniak; tytoń: pomidor), możliwe było systematyczne zbadanie wpływu poszczególnych czynników związanych z hybrydyzacją CSH na wynik eksperymentu.

Na podstawie wykonanych doświadczeń ustalono, że CSH jest metodą umożliwiającą uzyskanie wiarygodnych danych transkryptomicznych. Opracowano optymalną ścieżkę postępowania podczas wykonywania eksperymentów CSH i analizy

danych, uwzględniającą m.in. ocenę dostępnych zasobów laboratoryjnych i analitycznych, stosowanie sond cDNA, maksymalizację liczby powtórzeń, stosowanie filtracji danych opartej o parametry statystyczne. Istotnymi elementami, zwiększającymi szanse powodzenia eksperymentu są dostępność danych dotyczących stopnia podobieństwa sond do transkryptów oraz adnotacji funkcjonalnych. Ponadto, uzyskane wyniki eksperymentów na mikromacierzy homologicznej pozwoliły na opisanie profilów ekspresji genów w korzeniach *Nicotiana tabacum* w stresie abiotycznym.

Wydaje się, że w najbliższych latach metoda CSH powinna być stosowana jako badanie wstępne lub przesiewowe, pozwalające nakreślić ogólny profil ekspresji i wskazać potencjalnie zmienione miejsca w transkryptomie w porównaniu z istniejącymi gatunkami modelowymi.

ABSTRACT

„Utilizing of microarray cross-species hybridization in *Nicotiana tabacum* transcriptome analysis”

Cross-species hybridization (CSH) is the research method enabling assessment of gene expression in case the homologous microarray platform is not available. The method has been commonly used, but still considered as non-standard due to a lack of precise guidelines concerning experiment design and performance as well as data analysis. Few studies concerning CSH did not answer the questions concerning validity of the method.

The main purpose of this study was to perform comprehensive assessment of usefulness of CSH techniques and selection of most potent factors influencing results of such analyses.

We defined the specific objectives:

- establishing the extent to which sequence homology between genes of species of interest and species represented on the microarray influence on CSH;
- assessment of optimal length and type of probes for CSH research.
- development of data analysis method, taking into account data filtration.

Furthermore we assumed that the results of species-specific hybridization experiments should be performed as a reference, enabling assessment of CSH influencing factors.

We decided to perform experiments on tobacco (*Nicotiana tabacum*) of *Solanaceae* family, due to high sequence homology between its plant species, availability of several types of microarrays dedicated to different plant species, significance of these plants in agriculture and many publications concerning utilizing of CSH in this family.

Use of different types of microarrays (oligonucleotide, cDNA) and experimental designs (tobacco:potato; tobacco:tomato) enabled to study the influence of above mentioned factors on the CSH results.

We concluded that CSH is the method enabling to obtain reliable transcriptomic results. Optimal pipeline for performing CSH experiments and data analysis was established. The assessment of laboratory and analysis resources, use of cDNA probes, maximizing of number of replicates, use of data filtration based on statistical parameters are the most important aspects. The other important factors improving CSH results are the availability of data concerning probe-transcript homology and functional annotations. Furthermore, the results of species-specific experiments was used for description of gene expression profile in tobacco roots treated with abiotic stresses.

It seems that in next few years, CSH method should be utilized as preliminary or screening study for creating general expression profiles and indicating potentially altered fragments of transcriptome, comparing with model species.