

Projektowanie eksperymentu mikromacierzowego

Zakładamy **POWTARZALNOŚĆ** badanego zjawiska

Nasz cel: detekcja **ISTOTNYCH** wzorców/zmian/wskaźników danego procesu

Niezbędny jest **KOMPROMIS** pomiędzy idealnym schematem eksperymentu, a rzeczywistymi możliwościami (dostępność mikromacierzy, dostęp do próbek, ilość materiału, koszt eksperymentów)

Określamy **KONKRETNE** pytania, na które będziemy starali się odpowiedzieć

Musimy uwzględnić istnienie **ŹRÓDEŁ ZMIENNOŚCI**: wpływ platformy, technologii, protokołu, czynników incydentalnych, czynników biologicznych

Pamiętamy, że nasze przewidywania opierają się na **STATYSTYCZNEJ ANALIZIE** wielokrotnych obserwacji (błędy fałszywie dodatnie, fałszywie ujemne, SIŁA przewidywania prawdziwie dodatniego wyniku)

1. Cel eksperymentu – ustalenie warunków eksperymentu

Porównanie genotypów? (odpowiednia kontrola)

Porównanie tkanek? (najlepiej z tego samego osobnika)

Stymulacja? (odpowiednia kontrola)

Badanie zmian w czasie (ile osobników, synchronizacja)

1-czynnikowy czy wieloczynnikowy?

2. Wybór platformy

Rodzaj sond, położenie sond, reprezentatywność dla danego organizmu i planowanego eksperymentu, producent, dostępność, koszt etc.

L1. Mikromacierz oligonukleotydowa do analizy transkryptomu *Nicotiana benthamiana*

praca bioinformatyczno-eksperymentalna (w stosunku 70%:30%). Jej celem jest otrzymanie i przetestowanie mikromacierzy oligonukleotydowej do badań ekspresji genów *Nicotiana benthamiana*.

Arabidopsis Gene Expression Microarray

Barley Gene Expression Microarray

Bovine Gene Expression Microarray

Brassica Gene Expression Microarray

C. elegans Gene Expression Microarray

Canine Gene Expression Microarray

Chicken Gene Expression Microarray

Cotton Gene Expression Microarray

Drosophila (V2) Gene Expression Microarray

E. coli Gene Expression Microarray

Horse Gene Expression Microarray

Magnaporthe (V2) Gene Expression Microarray

Mosquito Gene Expression Microarray

Porcine Gene Expression Microarray

Rabbit Gene Expression Microarray

Rhesus Macaque Gene Expression Microarray

Rice Gene Expression Microarray

Salmon Gene Expression Microarray

Sheep Gene Expression Microarray

Tobacco Gene Expression Microarray

Tomato Gene Expression Microarray

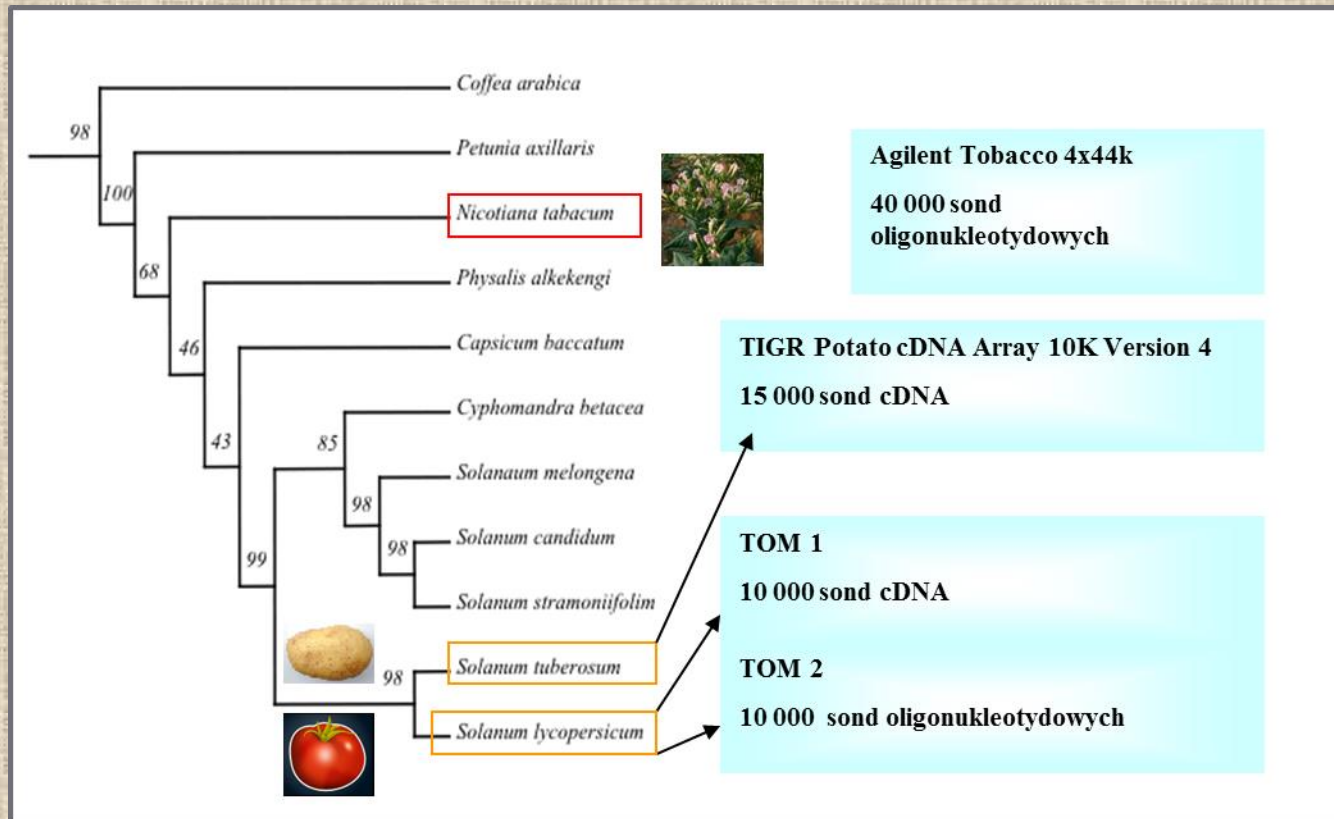
Wheat Gene Expression Microarray

Xenopus Gene Expression Microarray

Yeast Gene Expression Microarray

Zebrafish Gene Expression Microarray

o hybryzacji międzygatunkowej...



A.

I.



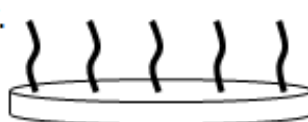
II.



III.



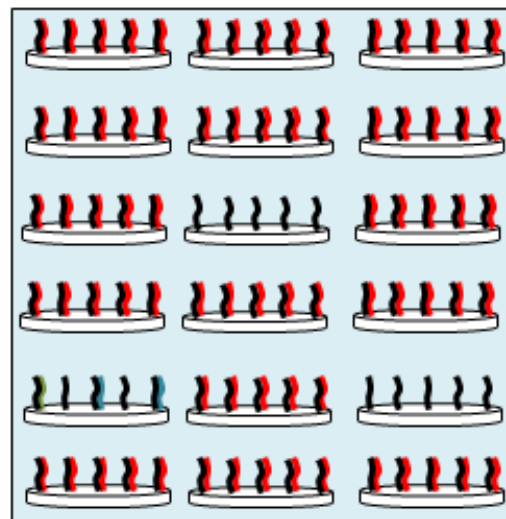
IV.



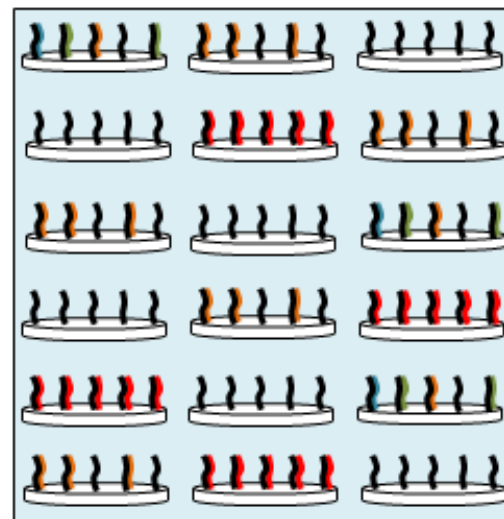
- > sonda
- > transkrypt o pełnym dopasowaniu
- > transkrypt o niepełnym dopasowaniu
- >> transkrypty inne niż docelowy

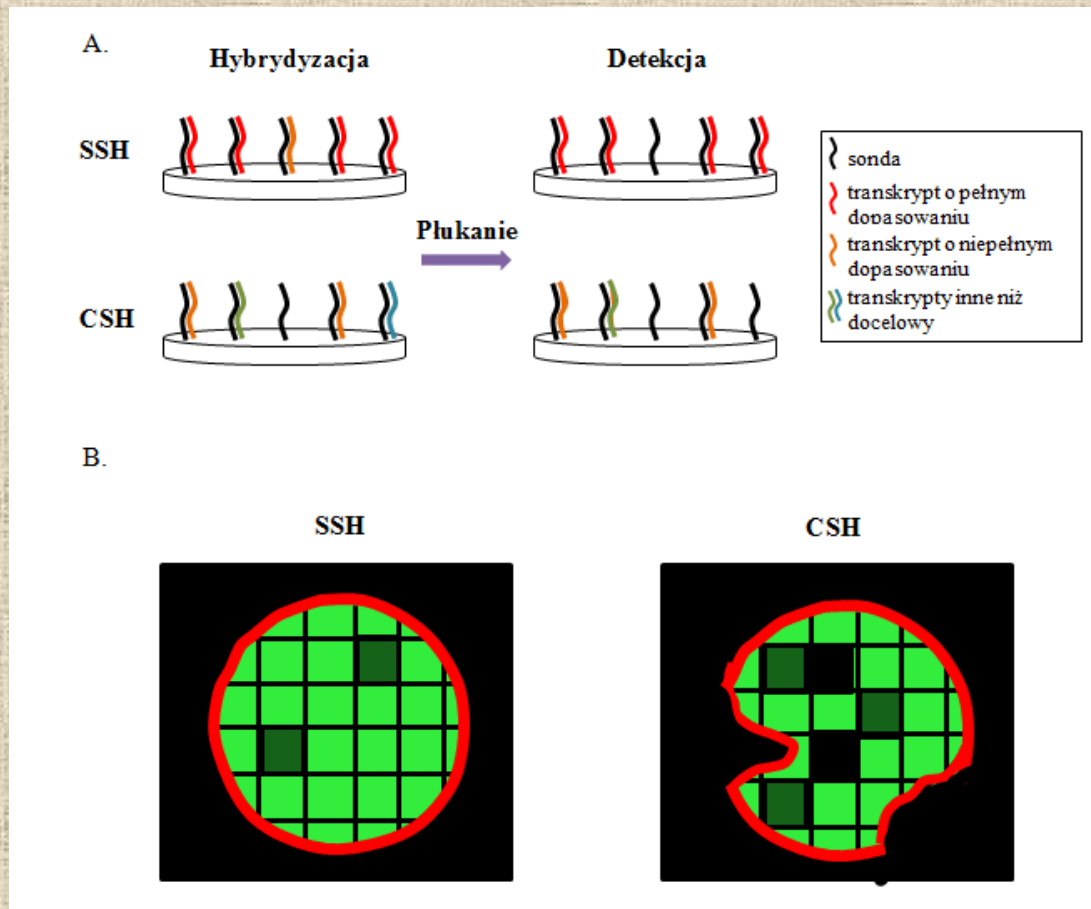
B.

SSH



CSH



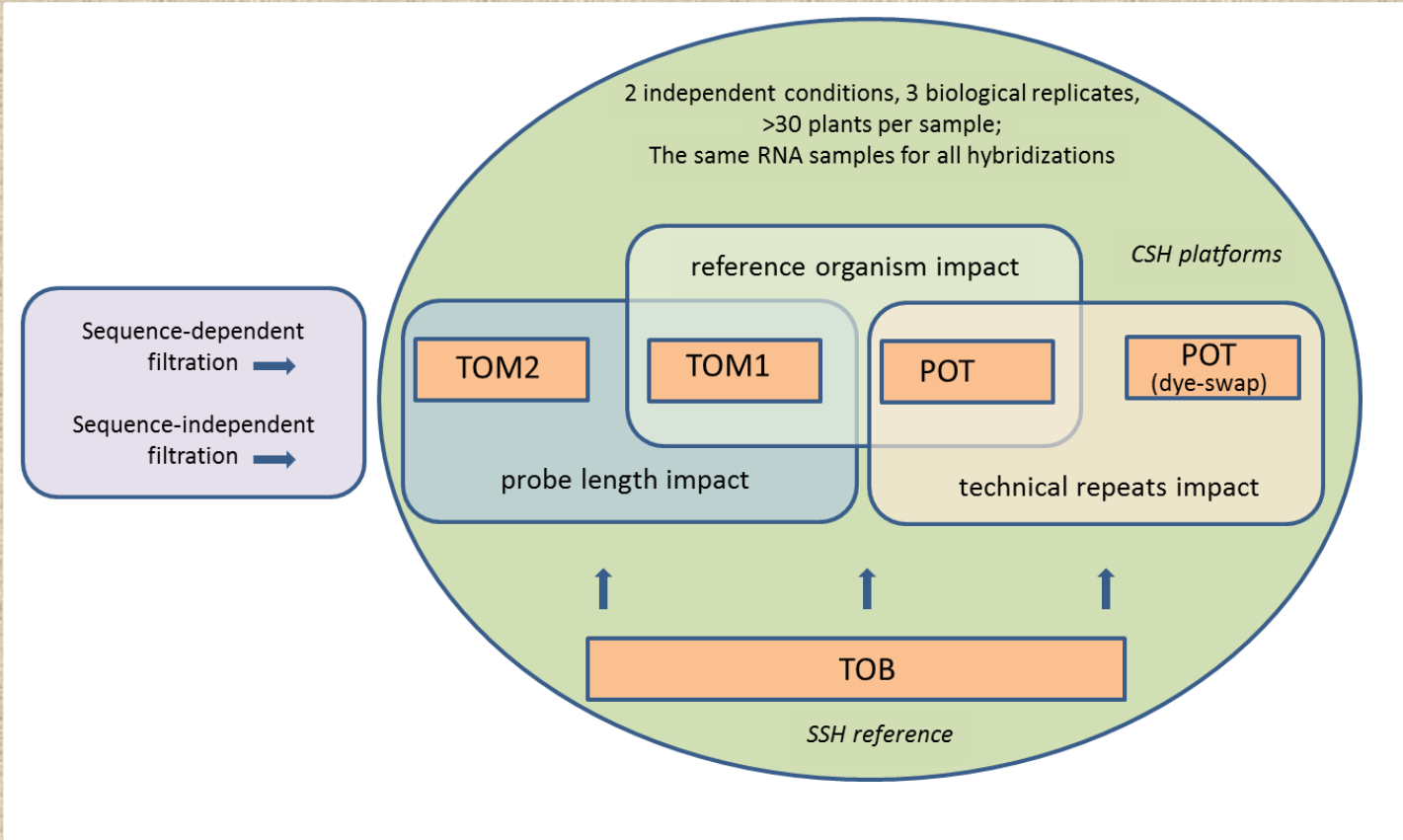


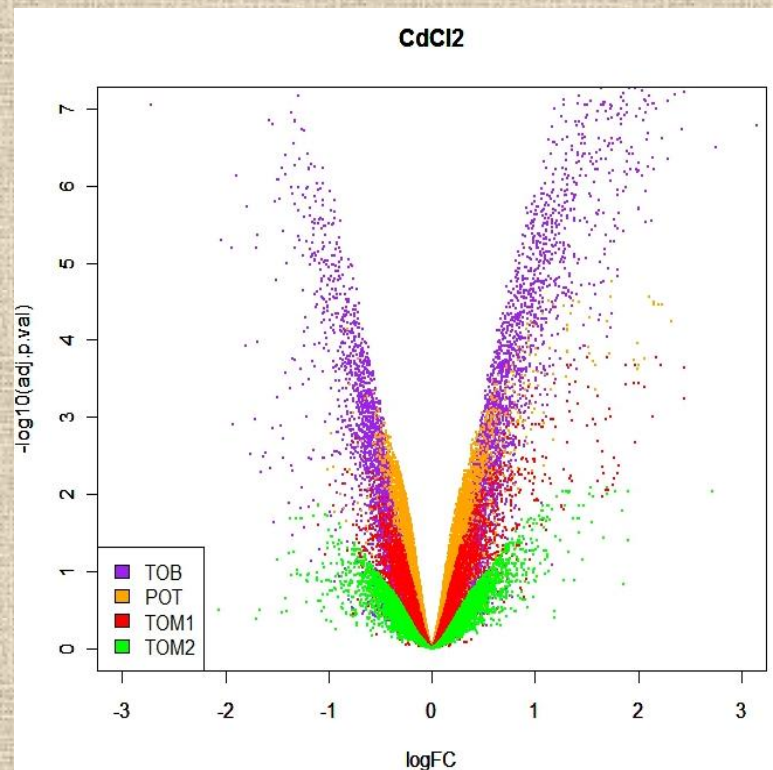
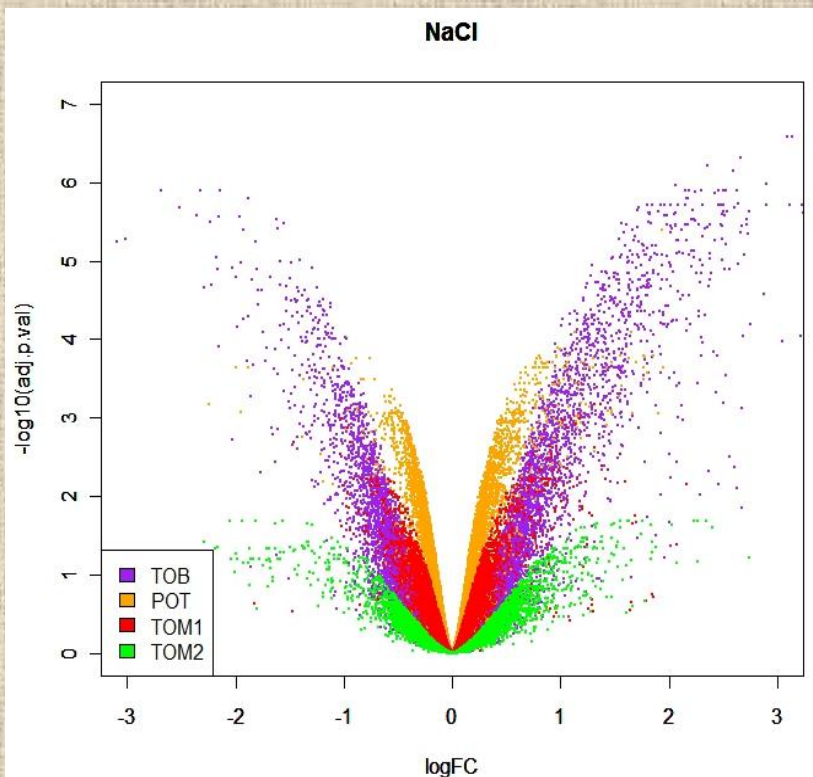
specyficzność



siła sygnału

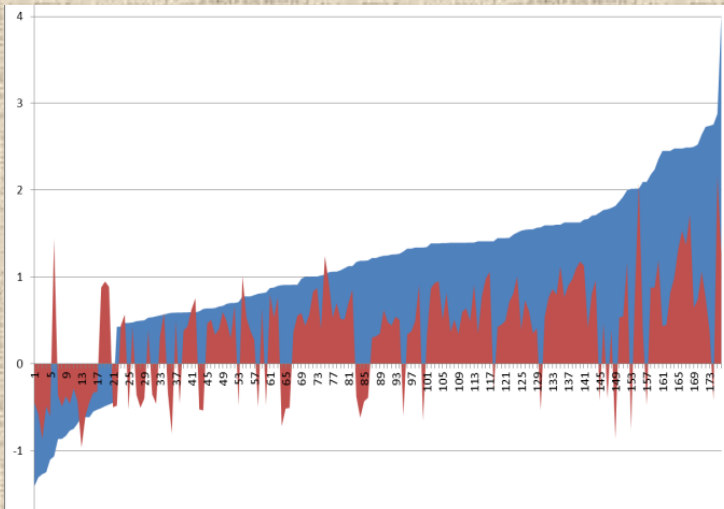
restrykcyjność odmywania



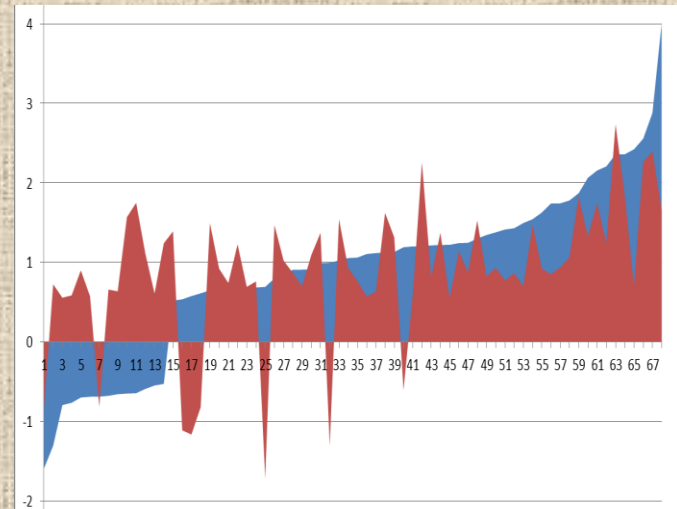


A)	NaCl	0	1	2	3	4
		ref	brak filtracji			
Platforma		TOB	TOM1	TOM2	POT-1powt	POT-2powt
Liczba przeanalizowanych sond		43730	8951	12020	10426	10426
Liczba sond o $p < 0,05$		2478 (9,4%)	530 (6,0%)	91 (0,7%)	3104 (29,8%)	991 (9,5%)
Sondy wskazujące na co najmniej dwukrotną zmianę ekspresji	Ogółem	1234	60	229	67	64
	W tym o ekspresji podwyższonej	75,0%	86,7%	55,9%	85,1%	78,1%

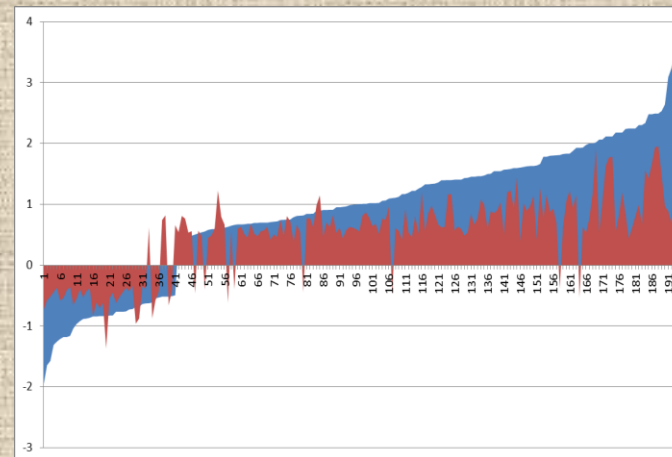
TOM1



TOM2



POT



PROBLEMY Z ADNOTACJĄ FUNKCJONALNĄ I BIOLOGICZNĄ INTERPRETACJĄ WYNIKÓW – MNÓSTWO WSTĘPNYCH ZAŁOŻEŃ

3. Uzyskanie materiału

Dobór próbek i sposobu uzyskania – minimalizacja naturalnej zmienności biologicznej:

-jeden czy wiele osobników

-(rośliny): sposób hodowli, temperatura, wilgotność, podlewanie, pora roku

-zbiór materiału: pora dnia, sposób zbioru, wielkość wycinka

4. Określenie ilości i rodzaju powtórzeń powtórzeń

organizm,
tkanka

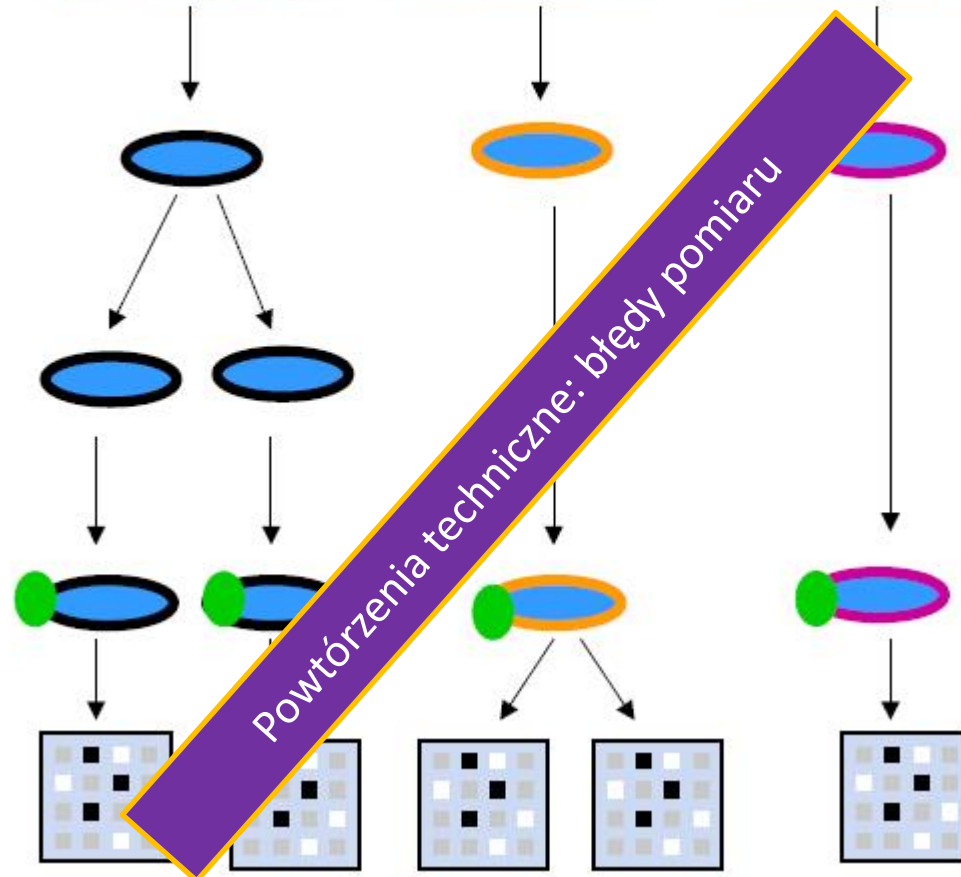
Powtórzenia biologiczne: wnioskowanie statystyczne

Izolacja RNA

amplifikacja

znakowanie

hybrydizacja



ile?
jakie?

5. Przygotowanie próbki

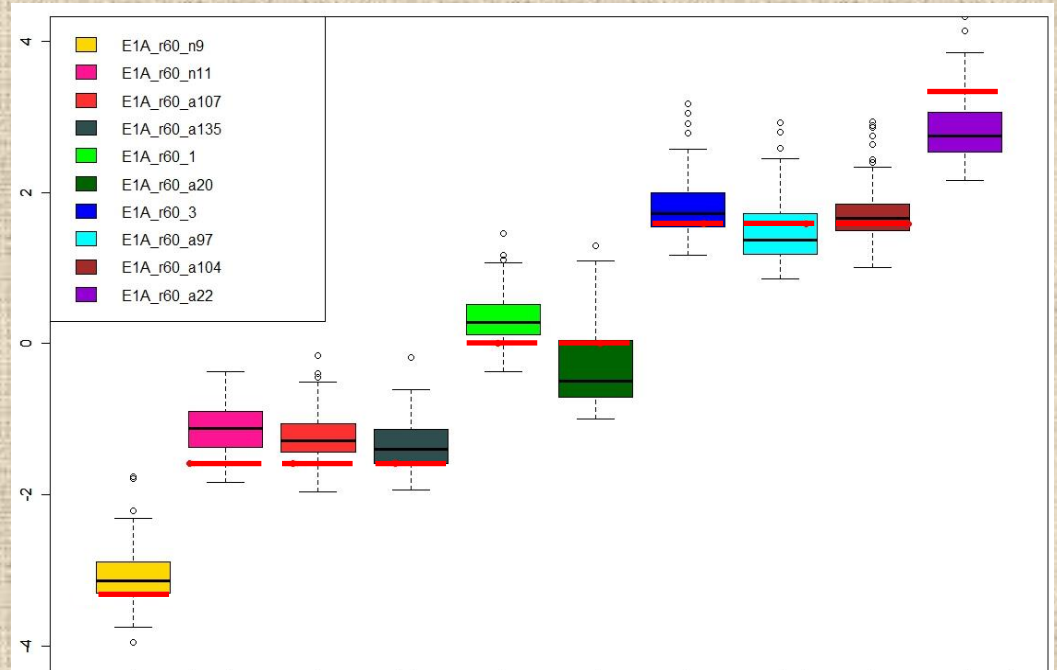
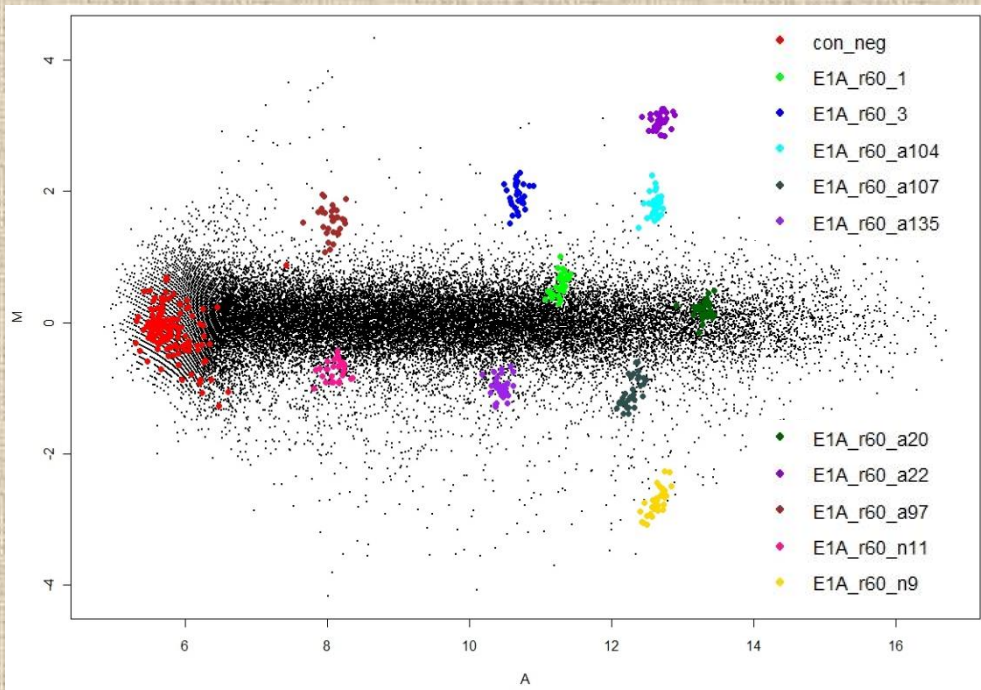
Standaryzacja protokołu

Amplifikacja

Kontrole zewnętrzne – „spike RNA”

Łączenie próbek / osobników w pulę

Kontrole zewnętrzne – „spike RNA”



Łączenie próbek / osobników w pulę

$T_0; T_1$

$T_0; T_1$

$T_0; T_1$

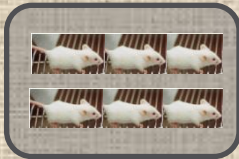
$T_0; T_1$

$T_0; T_1$

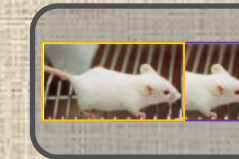
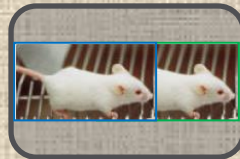
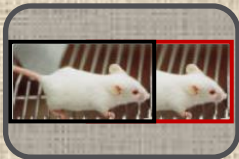
$T_0; T_1$



Tworzenie puli to „uśrednienie biologiczne” – redukuje zmienność biologiczną bez podnoszenia ilości mikromacierzy (=kosztów), przez co podnosi siłę statystyczną eksperymentu



ŁĄCZENIE W PULĘ NIE LIKWIDUJE KONIECZNOŚCI WYKONANIA POWTÓRZEŃ BIOLOGICZNYCH



$$\sigma_{Pool}^2 = \sigma^2 / \# \text{ in pool}$$

6. Opracowanie schematu hybrydyzacji

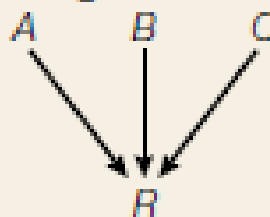
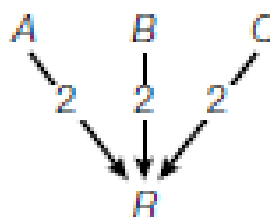
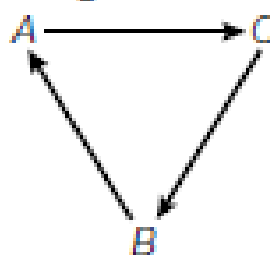
Hybrydyzacja 1-kanalowa



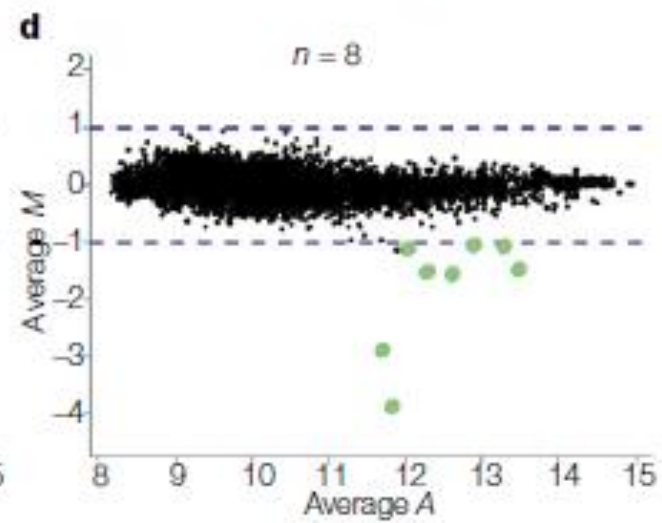
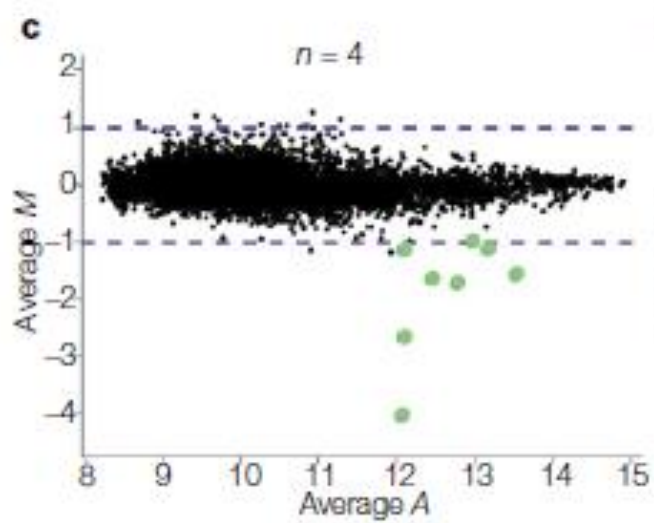
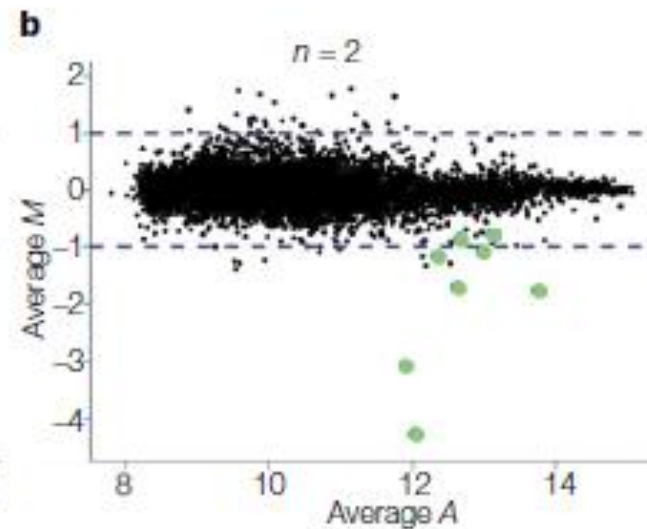
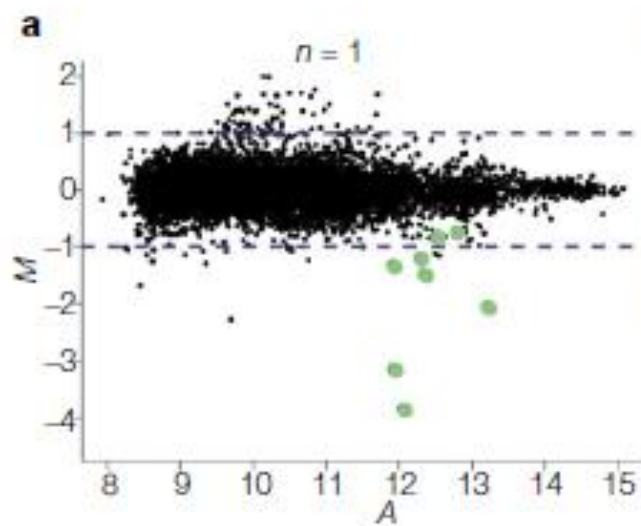
Hybrydyzacja 2-kanalowa



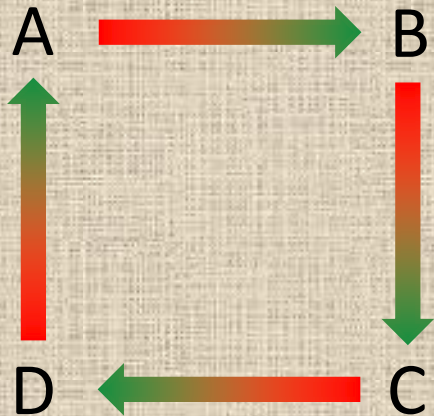
Table 1 | **Single-factor experiments**

Design choices	Number of slides	Units of material (number of samples)	Average variance
Indirect designs			
Design I 	3	$A = B = C = 1$	2.00
Design II 	6	$A = B = C = 2$	1.00
Direct design			
Design III 	3	$A = B = C = 2$	0.67

Variance of estimated effects for three different designs of single-factor experiments. σ^2 was set to 1 throughout.

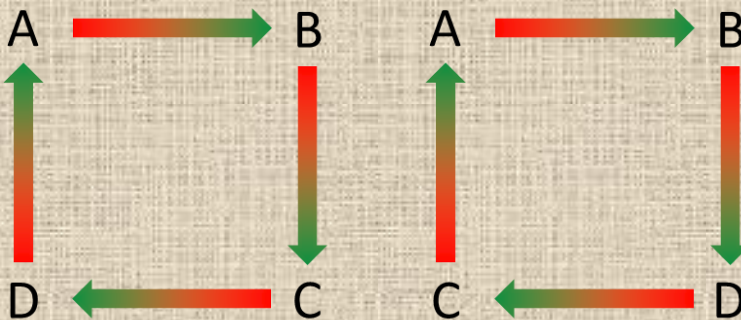


Pętle



wiele warunków / kombinacji do porównania,
Dye-swap pomiędzy różnymi próbkami,
analizy zmian w czasie
ale: nieudana hybrydyzacja wpływa na wszystkie porównania!

Pętla wielokrotna



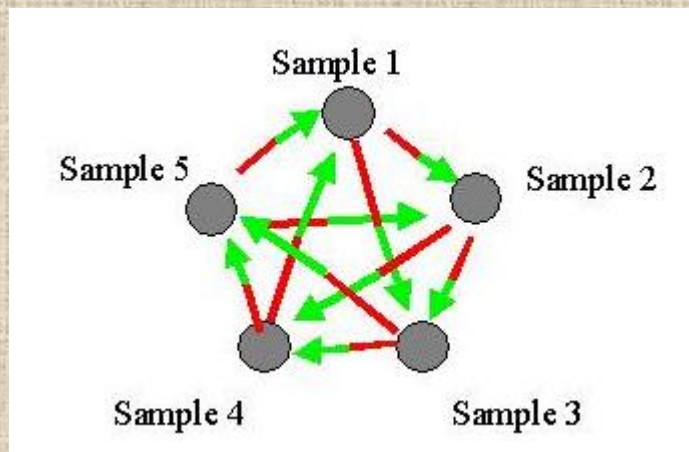
Zasada zbalansowanego eksperymentu:

Jednakowa ilość wszystkich próbek

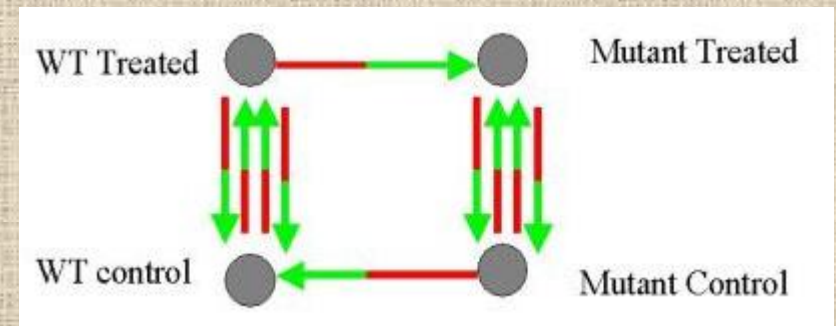
Jednakowa ilość znakowań Cy3 i Cy5

Porównania najbardziej interesujące w bezpośrednim porównaniu

Schemat wysyconych porównań



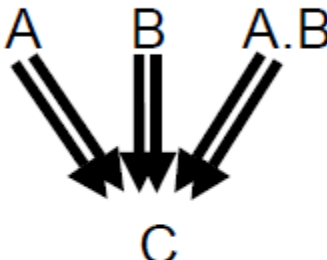
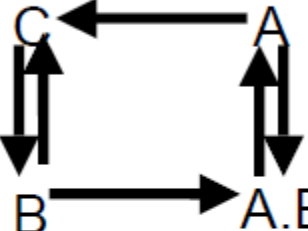
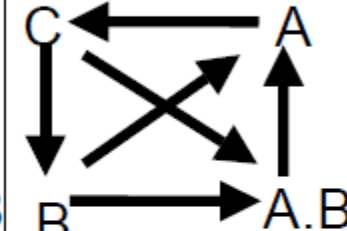
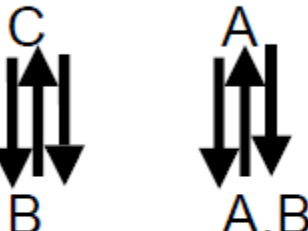
Schemat z wyborem najważniejszych porównań



Analizy zmian w czasie

Design choices in time series		t vs t+1			t vs t+2			Ave
		T1T2	T2T3	T3T4	T1T3	T2T4	T1T4	
N=3	A) T1 as common reference 	1	2	2	1	2	1	1.5
	B) Direct Hybridization 	1	1	1	2	2	3	1.67
N=4	C) Common reference 	2	2	2	2	2	2	2
	D) T1 as common ref + more 	.67	.67	1.67	.67	1.67	1	1.06
	E) Direct hybridization choice 1 	.75	.75	.75	1	1	.75	.83
	F) Direct Hybridization choice 2 	1	.75	1	.75	.75	.75	.83

2 x 2 factorial

	Indirect	A balance of direct and indirect		
	I) 	II) 	III) 	IV) 
# Slides	N = 6			
Main effect A	0.5	0.67	0.5	NA
Main effect B	0.5	0.43	0.5	0.3
Interaction A.B	1.5	0.67	1	0.67

Main effect A $\log(A / C) \text{ lub } \log(A.B / B)$

Interaction A.B $\log(A.B / B) - \log(A / C) = \log(A.B \times C / A \times B)$

High-Resolution Temporal Profiling of Transcripts during Arabidopsis Leaf Senescence Reveals a Distinct Chronology of Processes and Regulation

by Emily Breeze, Elizabeth Harrison, Stuart McHattie, Linda Hughes, Richard Hickman, Claire Hill, Steven Kiddle, Youn-sung Kim, Christopher A. Penfold, Dafyd Jenkins, Cunjin Zhang, Karl Morris, Carol Jenner, Stephen Jackson, Brian Thomas, Alexandra Tabrett, Roxane Legaie, Jonathan D. Moore, David L. Wild, Sascha Ott, David Rand, Jim Beynon, Katherine Denby, Andrew Mead, and Vicky Buchanan-Wollaston

Plantcell
Volume 23(3):873-894
April 25, 2011



Pytania:

jak wpływa proces starzenia na ekspresję genów?

czy cykl dobowy wpływa na ekspresję genów?

czy wpływ cyklu dobowego zmienia się w czasie starzenia?

jakie są powiązania pomiędzy profilem ekspresji a funkcją genów?

Próbka: 1 indywidualna roślina (L -7. liść rozety)



Punkty czasowe: 01 – 11 (19 DAS – 39 DAS)

Pory dnia: a, p

Powtórzenia biologiczne: a, b, c, d

Powtórzenia techniczne: czerwone, niebieskie

System: 2-kolorowy

$1 \times 11 \times 2 \times 4 \times 2 = 176$ próbek

Powtórzenia biologiczne (rano)

Powtórzenia biologiczne (popołudnie)

Dni



Supplemental Data. Breeze et al. (2011). Plant Cell 10.1105/pcp.111.083345.

Powtórzenia techniczne

Analiza (pakiet MAANOVA - MicroArray ANalysis Of Variance; Wu et al., 2003)

DANE

ANALIZA JAKOŚCI

wykresy, proste analizy statystyczne – korekta – powtórzenie hybrydyzacji

NORMALIZACJA WEWNĄTRZMACIERZOWA

transformacje LOWESS

NORMALIZACJA MIĘDZYMACIERZOWA

efekt mikromacierzy, barwnika oraz kombinacji tych dwóch czynników, uśredniony dla wszystkich genów

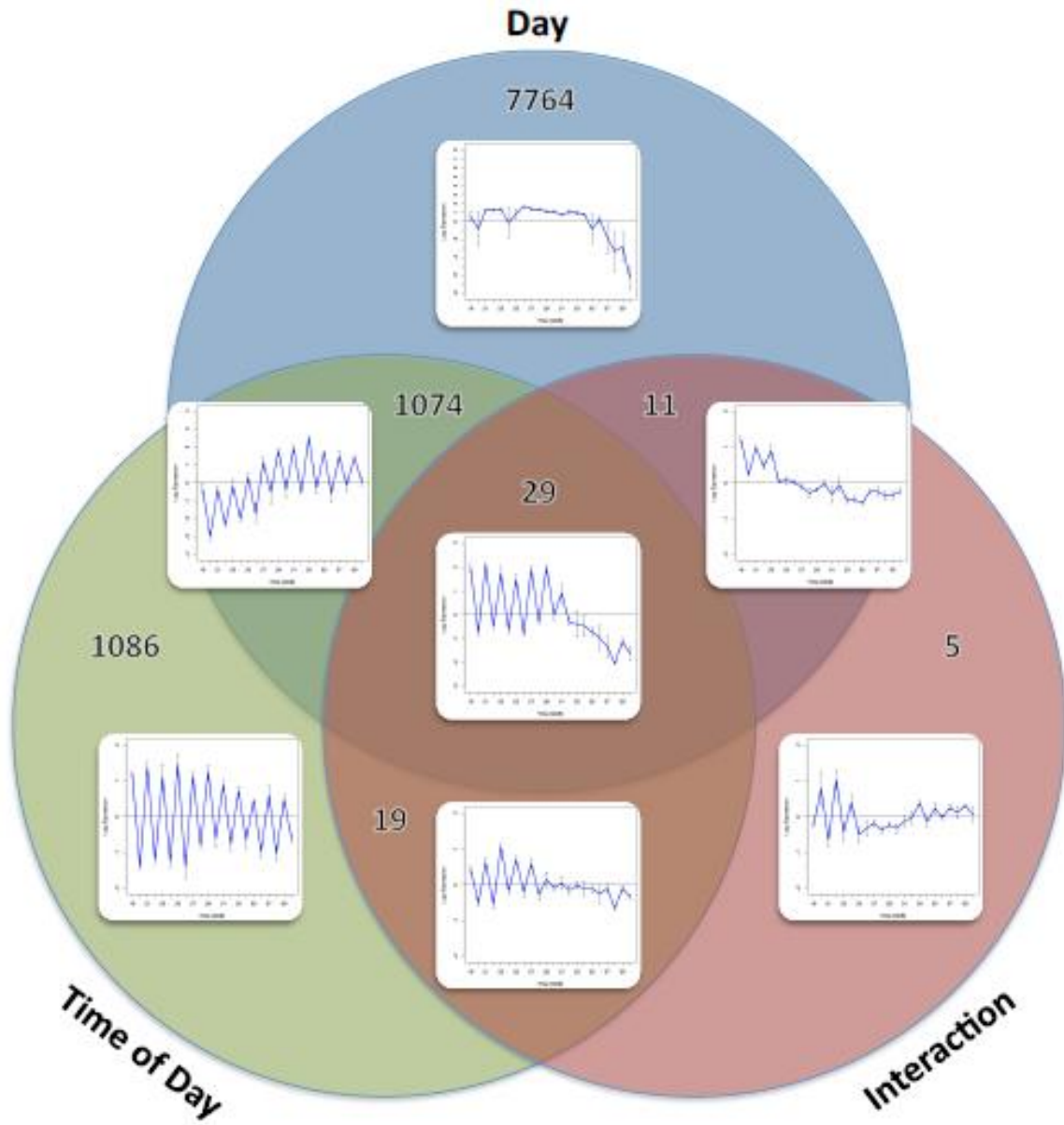
ANALIZA RÓŻNICOWA

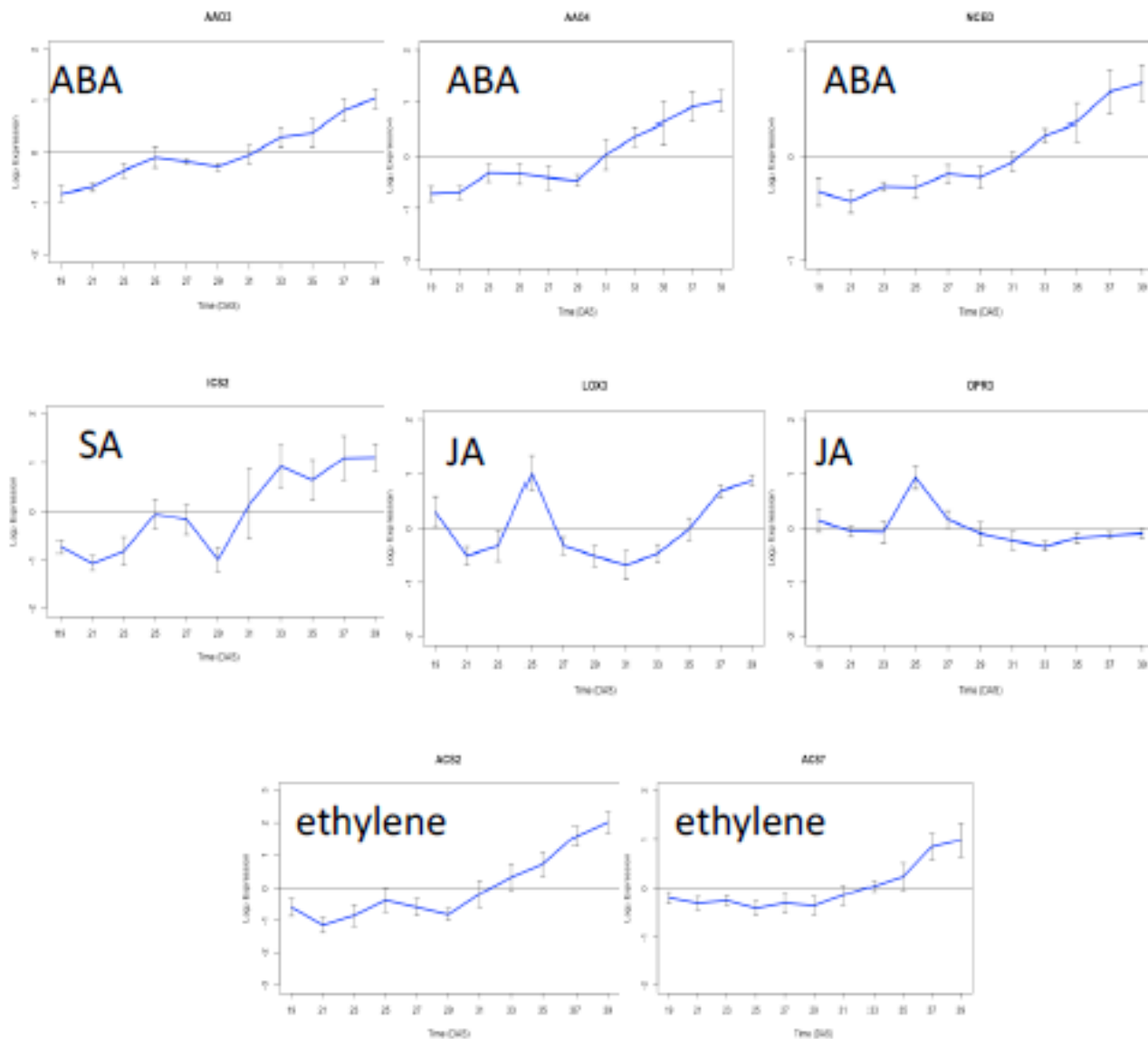
model mieszany $Y = \mu + \text{Dye} + \text{Array} + (\text{Day} * \text{TimeOfDay}) / \text{BioRep} + \epsilon$

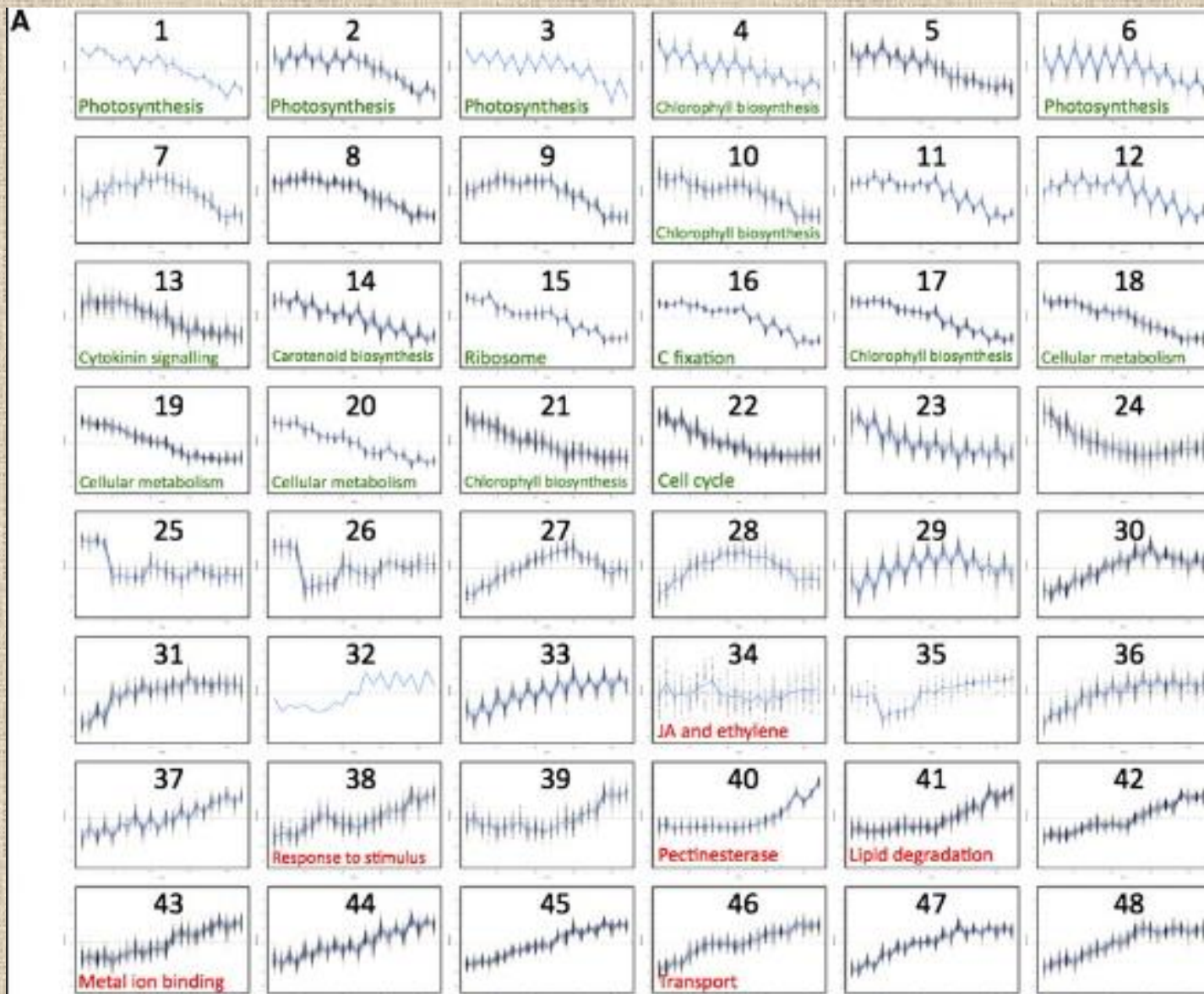
F-test + kontrola FDR

model uproszczony $Y = \mu + \text{Dye} + \text{Array} + \text{Day} / \text{BioRep} + \epsilon$

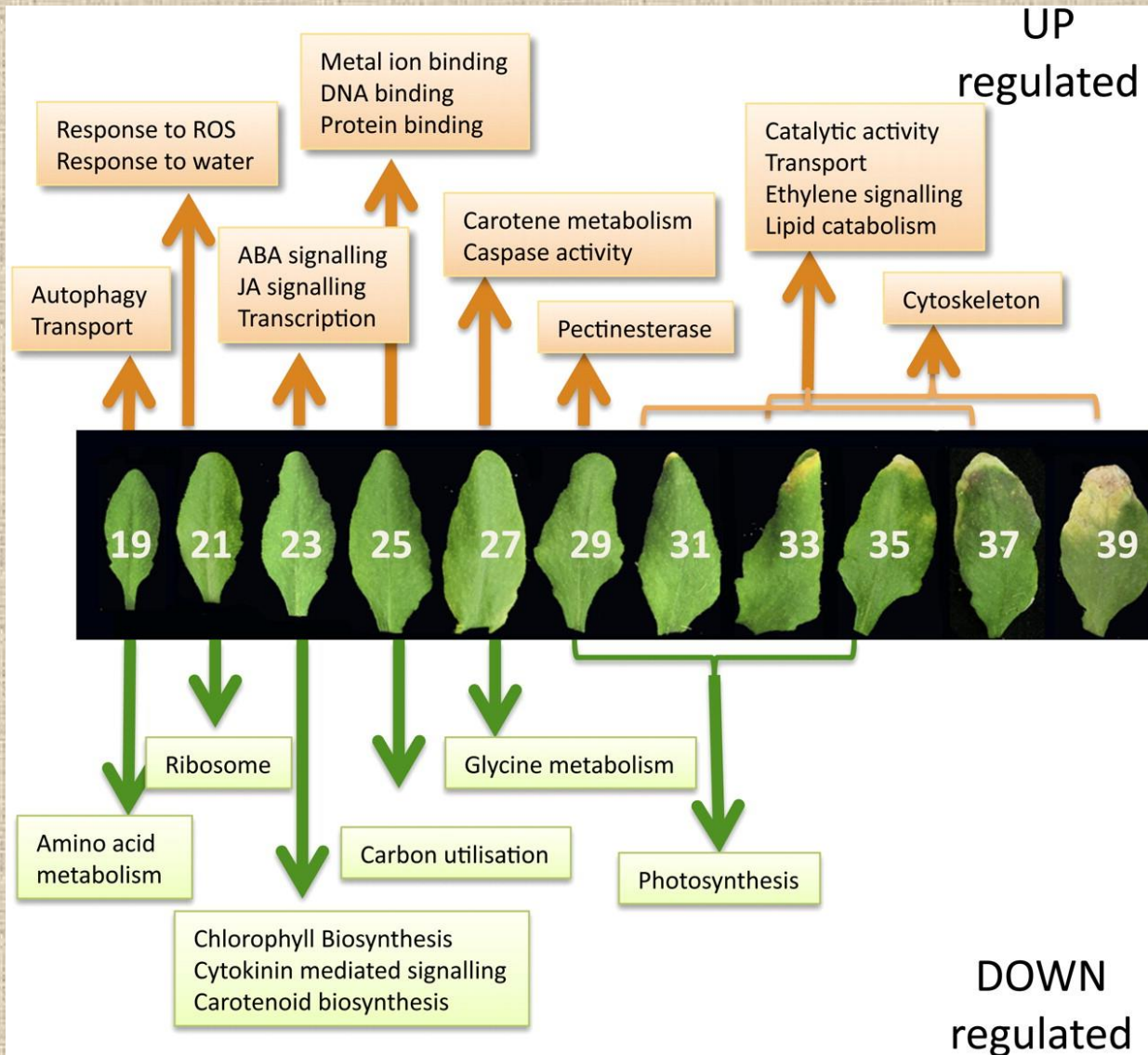
F-test + kontrola FDR







Geny o powiązanych funkcjach mają podobne profile ekspresji
nb. to leży u podstaw analiz funkcjonalnych typu „guilty by association”



GO: Biological Process

GO: Cellular Component

GO: Molecular Function

http://neurolex.org/wiki/Category:Resource:Gene_Ontology_Tools

Singular enrichment analysis

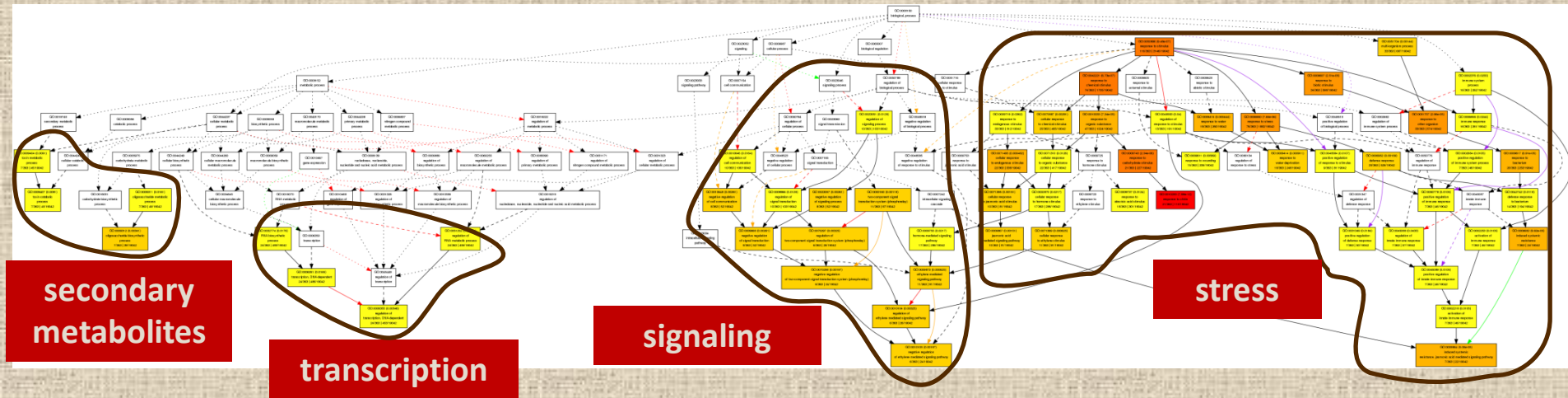
SEA is a traditional and widely used method. User only needs to prepare a list of gene/probe names, and enrichment GO terms will be found out after statistical test from pre-calculated background or customized one.

Maere S, Heymans K, Kuiper M (2005) BiNGO: a Cytoscape plugin to assess overrepresentation of Gene Ontology categories in biological networks. *Bioinformatics* 21, 3448-3449. ([PubMed](#))

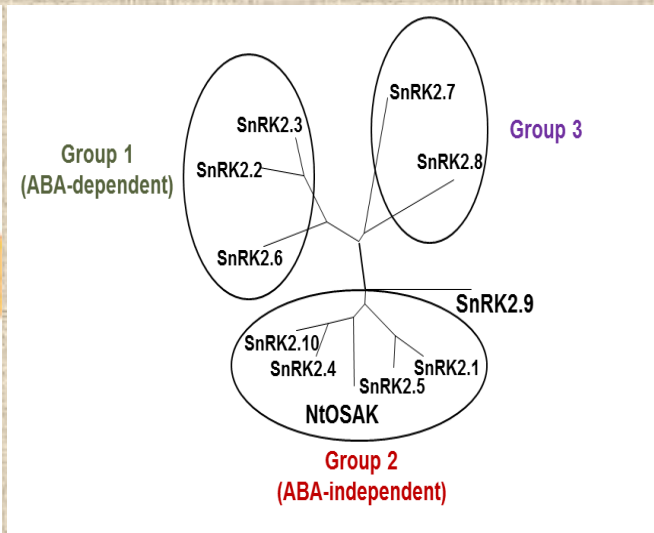
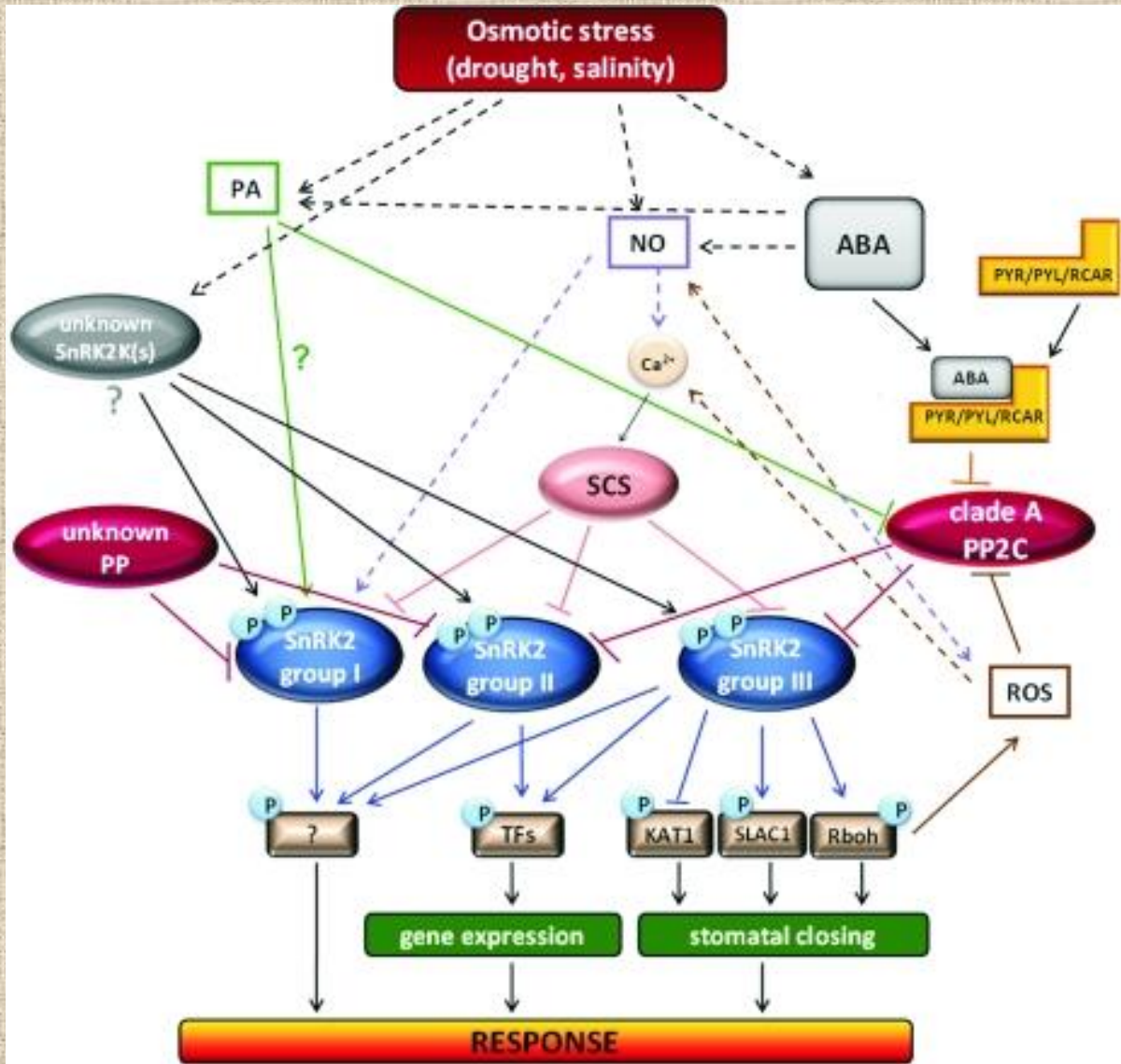
BiNGO : A Biological Network Gene Ontology tool.

BiNGO is a Java-based tool to determine which Gene Ontology (GO) categories are statistically overrepresented in a set of genes or a subgraph of a biological network. BiNGO is implemented as a plugin for Cytoscape, which is an open source bioinformatics software platform for visualizing and integrating molecular interaction networks. BiNGO maps the predominant functional themes of a given gene set on the GO hierarchy, and outputs this mapping as a Cytoscape graph. Gene sets can either be selected or computed from a Cytoscape network (as subgraphs) or compiled from sources other than Cytoscape (e.g. a list of genes that are significantly upregulated in a microarray experiment). The main advantage of BiNGO over other GO tools is the fact that it can be used directly and interactively on molecular interaction graphs. Another plus is that BiNGO takes full advantage of Cytoscape's versatile visualization environment. This allows you to produce customized high-quality figures.

Salt stress up-regulated genes in tobacco roots



Analiza GO może być problematyczna w przypadku organizmów niemodelowych!!!



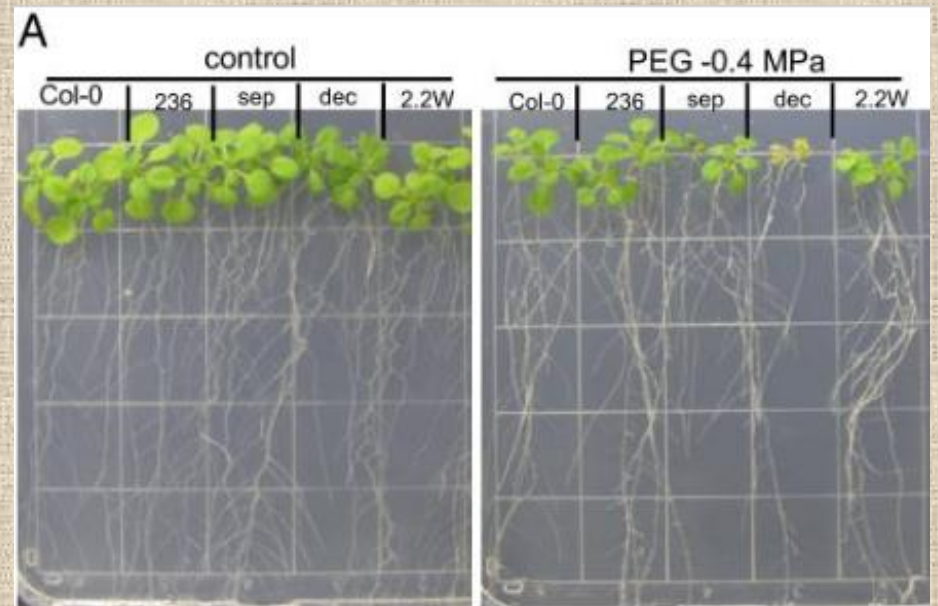
Rośliny:

<i>snrk2.4-1</i>	Salk_080588
<i>snrk2.4-2</i>	Salk_146522
<i>snrk2.10-1</i>	WiscDsLox233E9
<i>snrk2.10-2</i>	GABI_676G12
<i>wt</i>	

Traktowanie i kontrola

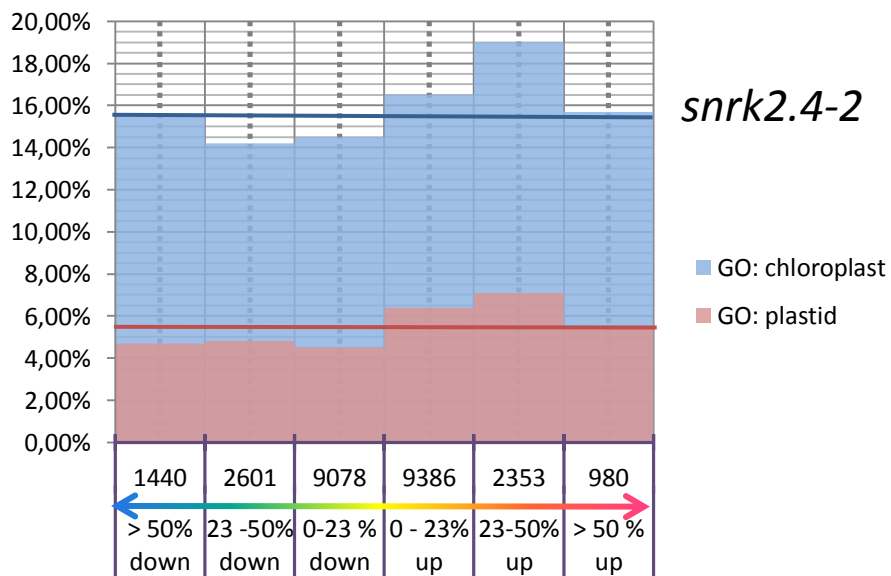
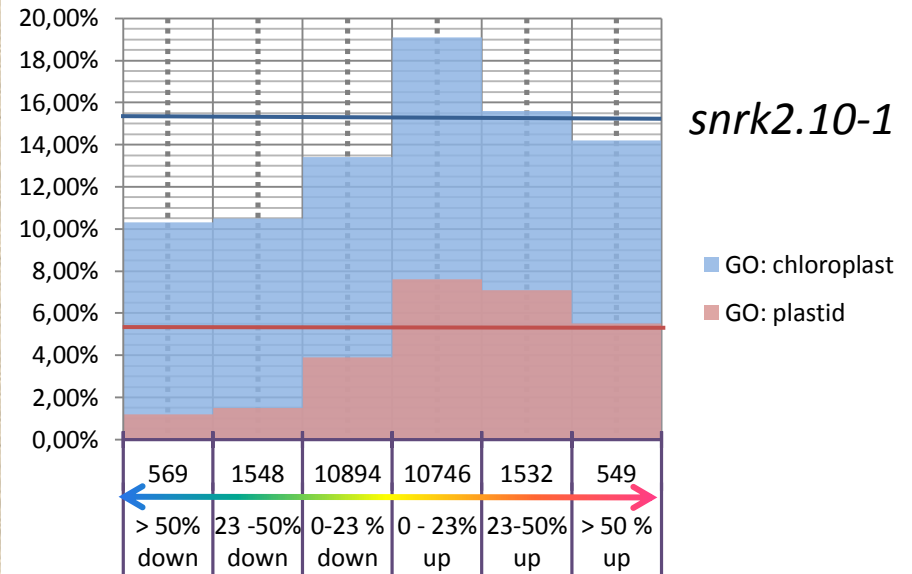
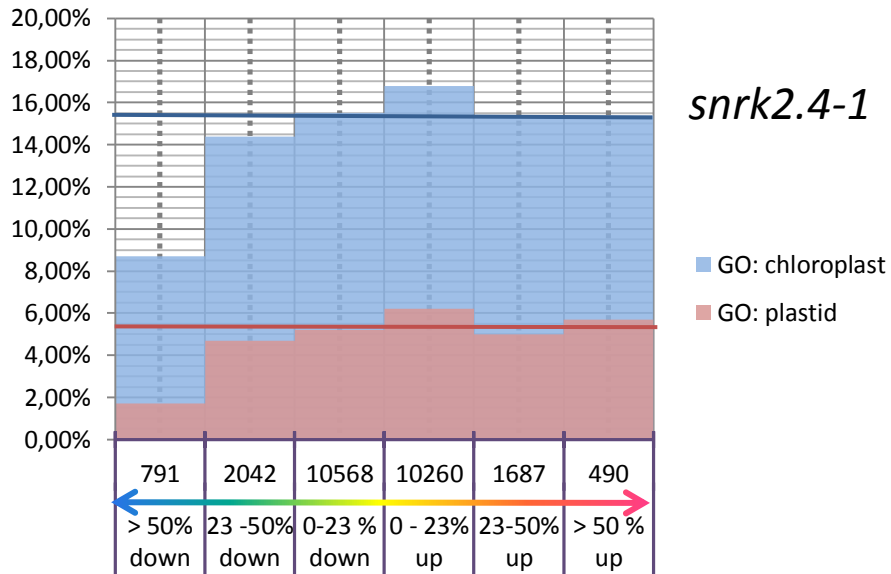
Siewki, liście i korzenie

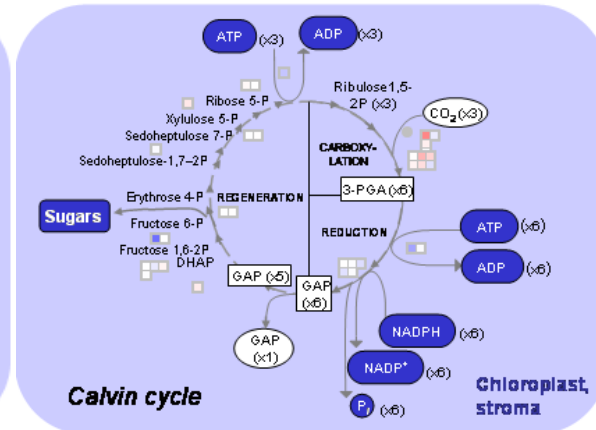
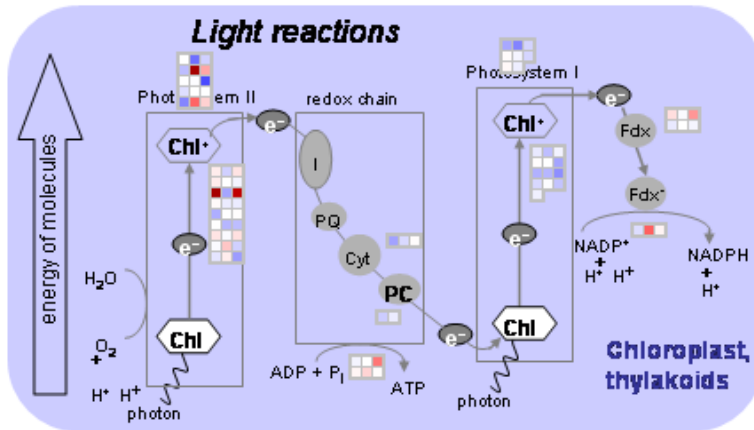
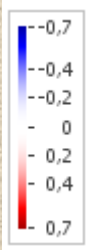
3-4 powtórzenia biologiczne



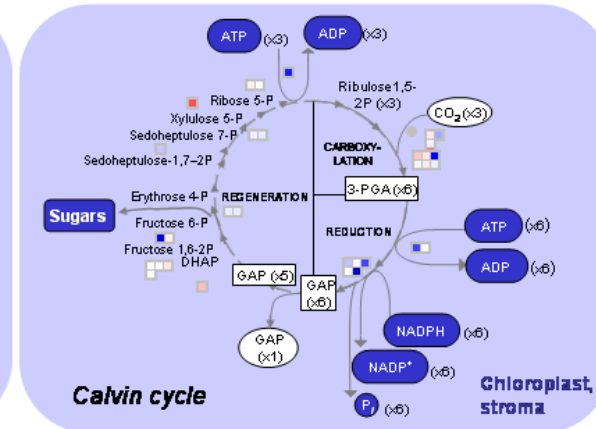
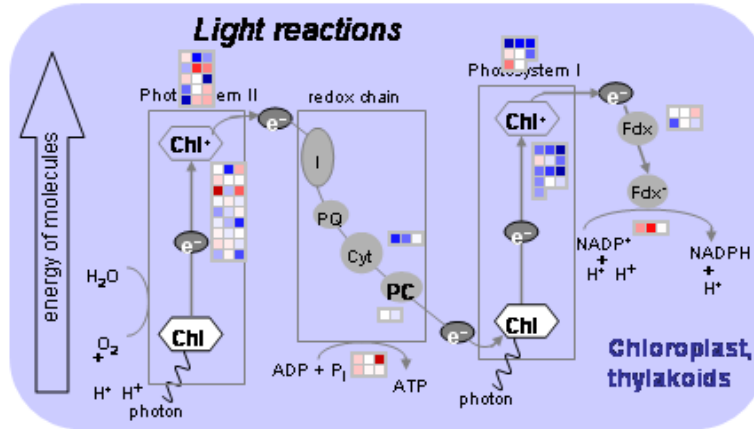
Fujii *et al.* Proc Natl Acad Sci U S A. 2011, 108(4): 1717–1722

Alterations in salt-stress response in multiple-snrk mutants relatively to wt plants	Decuple	Triple	Septuple
Disrupted SnRK2 groups	1,2,3	3 (ABA-dependent)	1 (ABA-independent)
Fresh weight	---	-	--
Root growth	---	-	-
ABA accumulation	---	+	-
ROS accumulation	0	0	0
Proline accumulation (ABA-dependent response)	--	--	+
IP ₃ accumulation (ABA-independent response)	---	0	--
Increase in osmotic stress-induced-genes expression	---	--	-
Global expression level of genes that are stressed-induced in wt	---	--	0

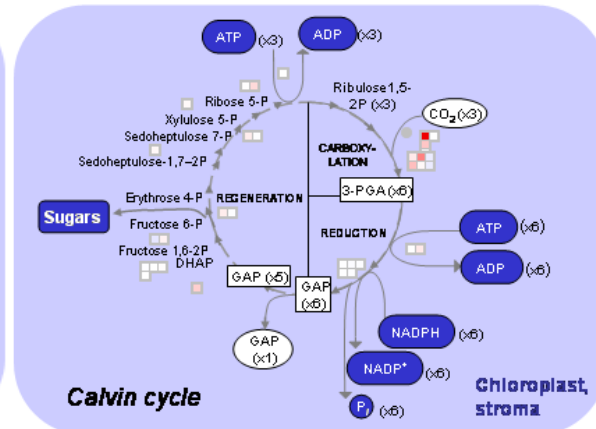
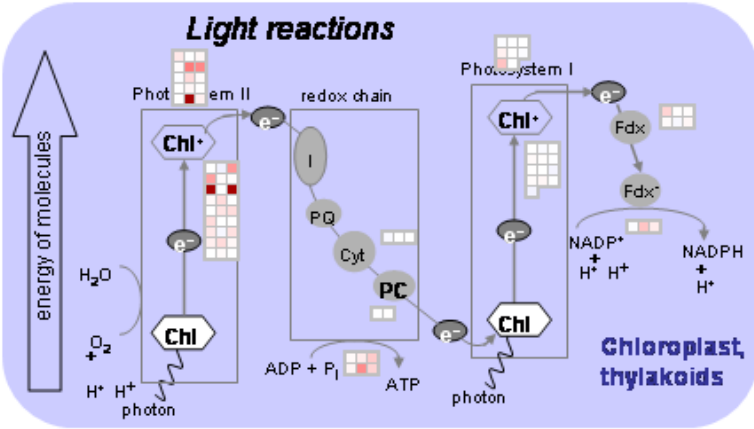




snrk2.4-1

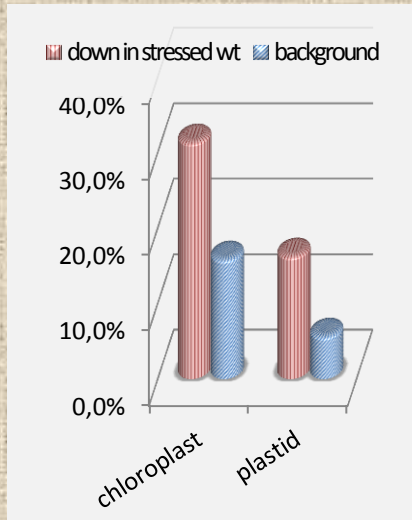


snrk2.4-2

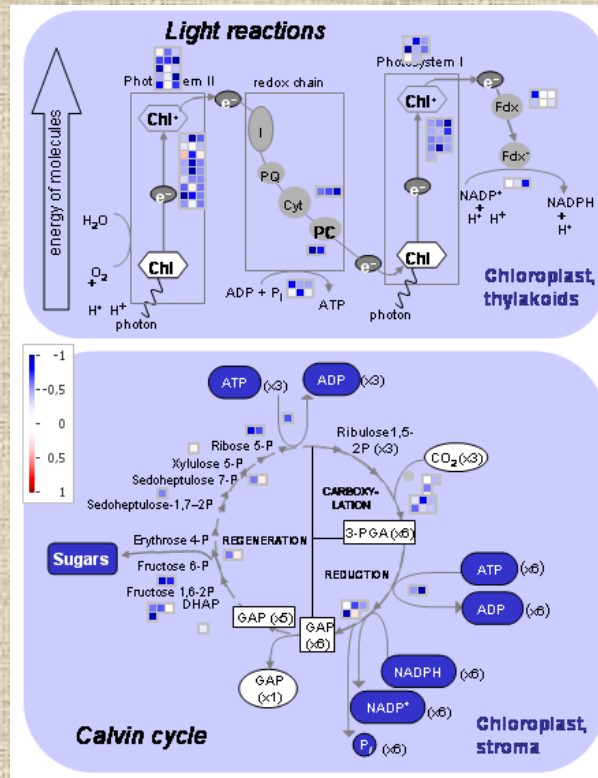


snrk2.10-1

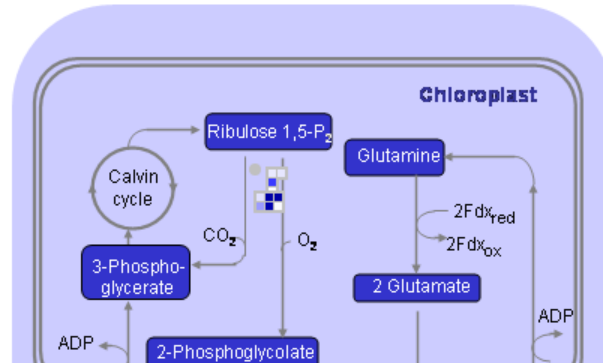
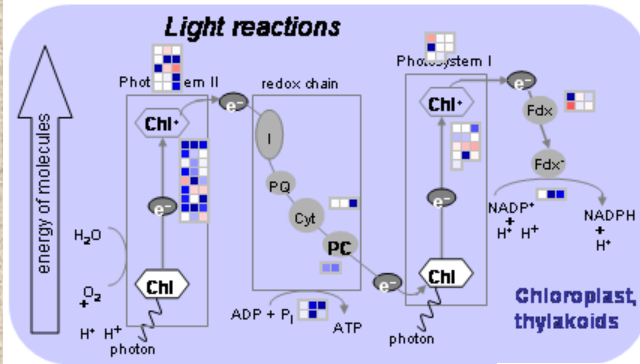
Stressed wt leaves



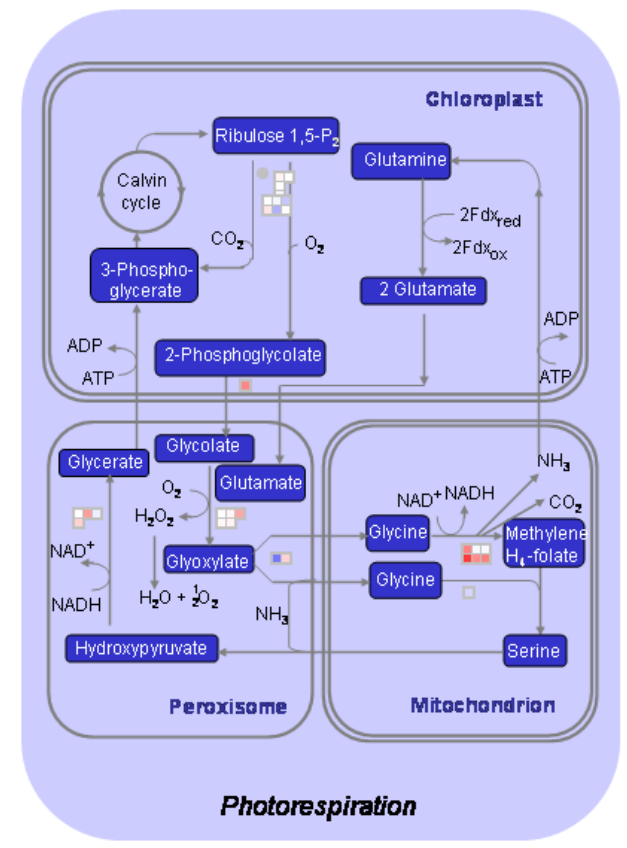
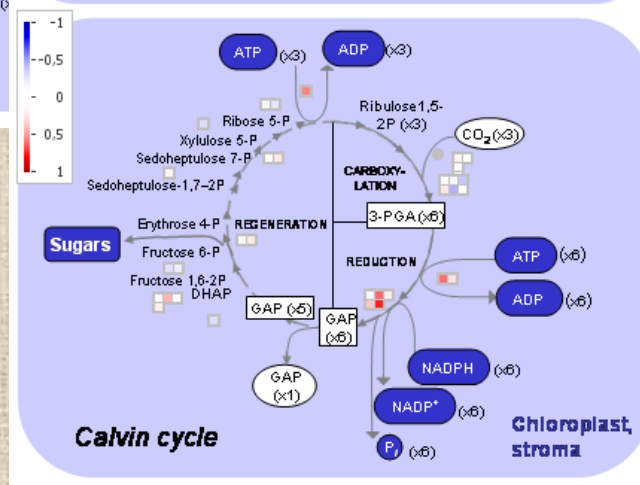
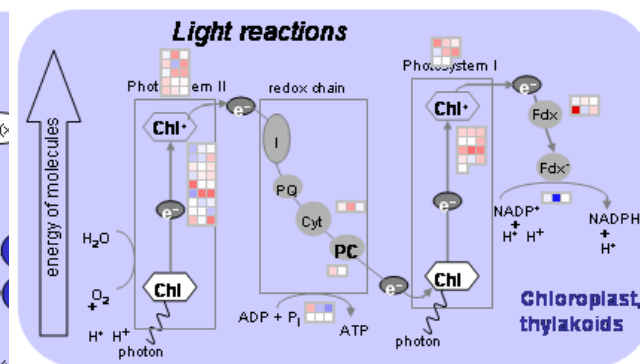
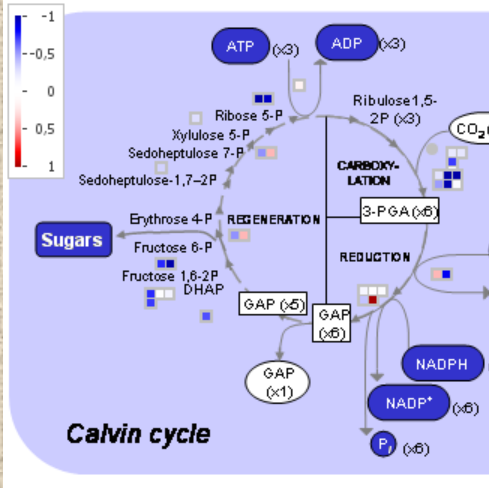
GO	Term	P-value	Genes	Enrichment
Biological Process				
GO:0009657	plastid organization	2.89e-08	36	3,8
GO:0009668	plastid membrane organization	2.02e-03	18	3,9
GO:0009765	photosynthesis, light harvesting	8.71e-03	8	8,4
GO:0009767	photosynthetic electron transport chain	7.05e-04	12	6,5
GO:0009773	photosynthetic electron transport in photosystem	5.24e-03	9	7,6
GO:0010027	thylakoid membrane organization	2.02e-03	18	3,9
GO:0010114	response to red light	5.40e-05	15	6,1
GO:0010207	photosystem II assembly	3.60e-07	22	5,4
GO:0010218	response to far red light	1.22e-04	14	6,2
GO:0015979	photosynthesis	1.26e-19	53	5,2
GO:0015994	chlorophyll metabolic process	4.84e-05	20	4,5
GO:0015995	chlorophyll biosynthetic process	8.93e-05	16	5,5
GO:0019684	photosynthesis, light reaction	2.96e-18	45	5,9



Stressed *snrk2.4-1* leaves



compared to *snrk2.4-1* under control conditions



compared to stressed wt

PAGE: Parametric Analysis of Gene Set Enrichment

Seon-Young Kim² and David J Volsky^{1,3}

Corresponding authors: Seon-Young Kim kimsy@kribb.re.kr - David J Volsky djv4@columbia.edu

▼ Author Affiliations

¹ Molecular Virology Division, St. Luke's-Roosevelt Hospital Center and Columbia University, New York, NY 10019, USA

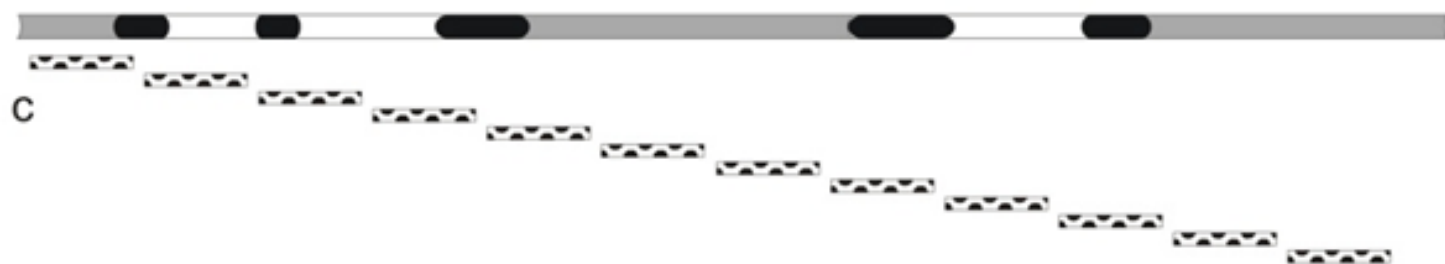
² Human Genomics Laboratory, Genome Research Center, Korea Research, Institute of Biosciences and Biotechnology 52 Eoeun-dong Yuseong-gu, Daejeon, 305-333, Korea

³ Molecular Virology Division, St. Luke's-Roosevelt Hospital Center 432 West 58th, Street Antenucci Building, Room 709 New York, NY 10019, USA

PAGE method is using Central Limit Theorem in statistics, this method is simple and efficient. Different to SEA, it takes expression level into account, and can deal with a long list of genes/probesets.

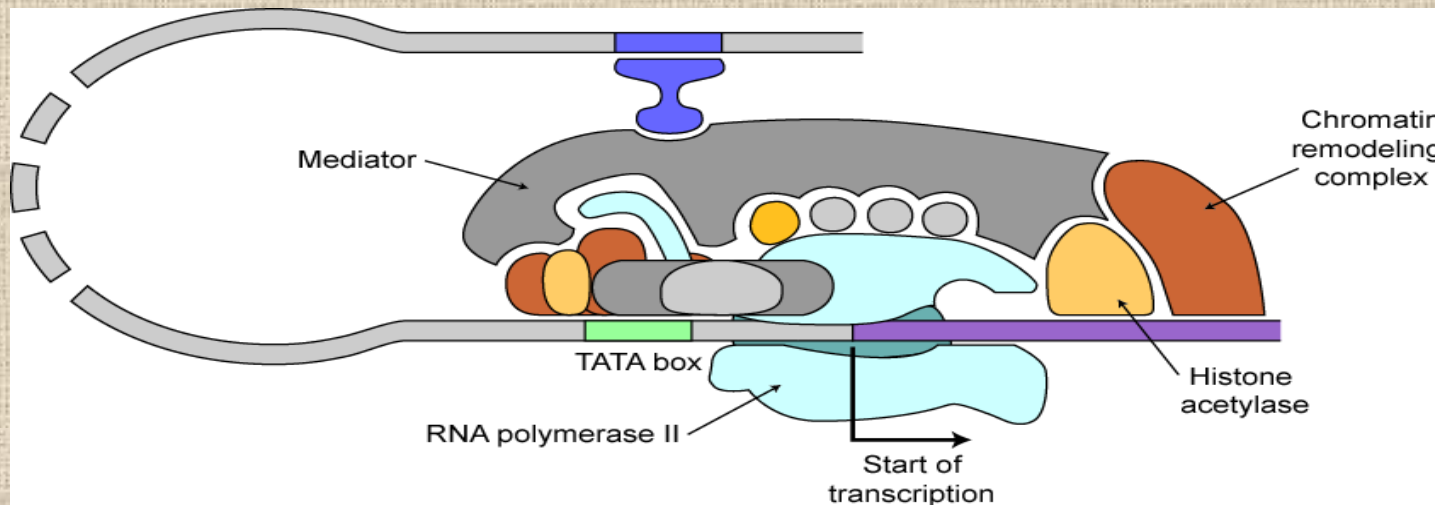
		osmotic stress response				expression in comparison with wt					
						basal (no stress)				during osmotic stress	
		wt S	wt L	2.4-1 L	2.10-1 L	2.4-1 S	2.10-1 S	2.4-1 L	2.10-1 L	2.4-1 L	2.10-1 L
C	cytosolic large ribosomal subunit	blue	light blue	grey	grey	yellow	red	orange	grey	red	yellow
C	cytosolic ribosome	blue	light blue	light blue	light blue	orange	red	red	grey	red	orange
C	cytosolic small ribosomal subunit	blue	light blue	grey	grey	red	orange	red	red	red	yellow
C	large ribosomal subunit	blue	light blue	light blue	light blue	yellow	red	orange	grey	red	yellow
C	organellar ribosome	grey	light blue	grey	grey	yellow	red	red	grey	red	grey
C	ribosomal subunit	blue	blue	light blue	light blue	yellow	red	red	grey	red	orange
C	ribosome	blue	blue	blue	light blue	orange	red	red	yellow	red	red
C	small ribosomal subunit	blue	light blue	light blue	grey	grey	red	orange	grey	red	yellow
C	structural constituent of ribosome	blue	blue	light blue	light blue	yellow	red	red	grey	red	orange
C	ribosome biogenesis	blue	light blue	light blue	grey	grey	red	orange	grey	red	orange
F	structural constituent of ribosome	blue	blue	light blue	light blue	orange	yellow	red	red	red	red
P	ribosome biogenesis	blue	light blue	light blue	grey	orange	yellow	red	orange	orange	red
P	translation	blue	blue	light blue	light blue	red	yellow	red	yellow	red	red
P	ribonucleoprotein complex biogenesis	blue	light blue	light blue	grey	orange	grey	orange	grey	orange	red
P	cellular protein metabolic process	light blue	grey	light blue	grey	yellow	grey	yellow	grey	yellow	red
P	protein targeting to chloroplast	grey	grey	light blue	grey	yellow	grey	yellow	grey	grey	yellow

		osmotic stress response				expression in comparison with wt					
						basal (no stress)				during osmotic stress	
		wt S	wt L	2.4-1 L	2.10-1 L	2.4-1 S	2.10-1 S	2.4-1 L	2.10-1 L	2.4-1 L	2.10-1 L
GO:0009507	chloroplast	light blue	blue	blue	light blue	red	red	red	red	red	red
GO:0009941	chloroplast envelope	light blue	blue	blue	light blue	yellow	red	yellow	red	yellow	red
GO:0009706	chloroplast inner membrane	grey	grey	grey	grey	yellow	red	red	red	red	yellow
GO:0031969	chloroplast membrane	grey	grey	grey	grey	grey	red	red	red	red	yellow
GO:0009707	chloroplast outer membrane	grey	grey	light blue	grey	grey	red	red	red	red	red
GO:0044434	chloroplast part	blue	blue	light blue	light blue	red	red	orange	red	orange	red
GO:0030095	chloroplast photosystem II	light blue	light blue	light blue	grey	grey	yellow	grey	yellow	yellow	orange
GO:0009570	chloroplast stroma	light blue	blue	blue	light blue	orange	red	orange	red	yellow	red
GO:0009534	chloroplast thylakoid	light blue	blue	blue	light blue	orange	red	grey	orange	orange	red
GO:0009543	chloroplast thylakoid lumen	light blue	blue	blue	light blue	yellow	orange	grey	yellow	yellow	red
GO:0009535	chloroplast thylakoid membrane	light blue	blue	light blue	light blue	orange	red	red	orange	red	red
GO:0030076	light-harvesting complex	blue	blue	light blue	grey	orange	red	grey	yellow	red	red
GO:0034357	photosynthetic membrane	blue	blue	light blue	light blue	orange	red	orange	orange	red	red
GO:0009521	photosystem	blue	light blue	light blue	light blue	orange	red	grey	orange	red	red
GO:0009522	photosystem I	blue	light blue	grey	grey	orange	red	orange	orange	red	orange
GO:0009523	photosystem II	blue	light blue	light blue	light blue	orange	red	orange	yellow	red	red
GO:0009536	plastid	light blue	blue	blue	light blue	red	red	orange	red	yellow	red
GO:0009508	plastid chromosome	grey	light blue	light blue	grey	grey	yellow	grey	grey	grey	yellow
GO:0009526	plastid envelope	light blue	blue	blue	light blue	yellow	red	yellow	orange	orange	red
GO:0009528	plastid inner membrane	grey	grey	grey	grey	grey	red	red	red	red	yellow
GO:0042170	plastid membrane	grey	grey	grey	grey	grey	red	red	red	red	yellow
GO:0042646	plastid nucleoid	grey	light blue	light blue	grey	grey	yellow	grey	grey	grey	yellow
GO:0044435	plastid part	light blue	blue	blue	light blue	red	red	orange	red	red	red
GO:0009532	plastid stroma	light blue	blue	blue	light blue	orange	red	red	red	yellow	red
GO:0031976	plastid thylakoid	blue	blue	light blue	light blue	orange	red	orange	orange	red	red
GO:0031978	plastid thylakoid lumen	light blue	blue	blue	light blue	yellow	orange	red	yellow	yellow	red
GO:0055035	plastid thylakoid membrane	blue	blue	light blue	light blue	orange	red	orange	orange	orange	red
GO:0010287	plastoglobule	blue	light blue	light blue	grey	orange	red	red	yellow	yellow	red
GO:0010319	stromule	light blue	light blue	grey	grey	yellow	orange	grey	orange	orange	red
GO:0009579	thylakoid	light blue	blue	blue	light blue	orange	red	red	orange	orange	red
GO:0031977	thylakoid lumen	light blue	blue	blue	light blue	grey	orange	grey	yellow	yellow	red
GO:0042651	thylakoid membrane	blue	blue	blue	light blue	red	red	yellow	orange	orange	red
GO:0044436	thylakoid part	blue	blue	blue	light blue	red	red	grey	orange	orange	red



- region międzygenowy
- ekson
- intron
- ☞ sonda

ChIP-on-chip czyli immunoprecytacja chromosomów połączona z analizą mikromacierzową



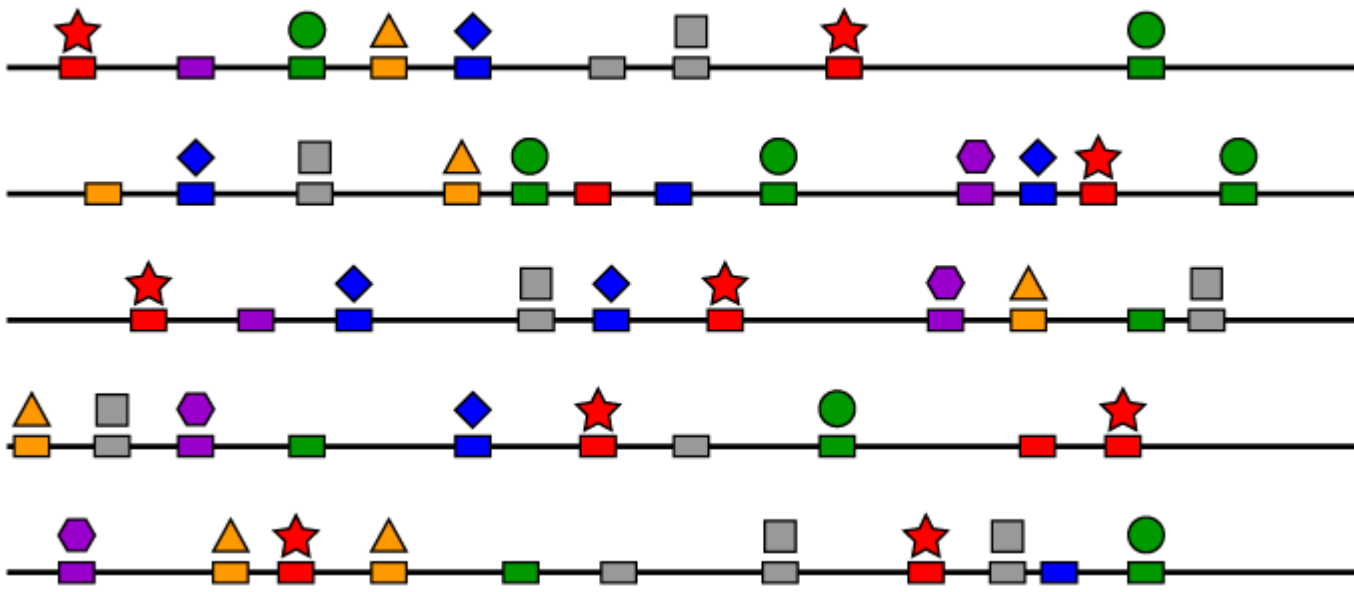
Cel: Badanie oddziaływań białek z DNA – rozpoznawanie rejonów chromatyny zawierających dane białko

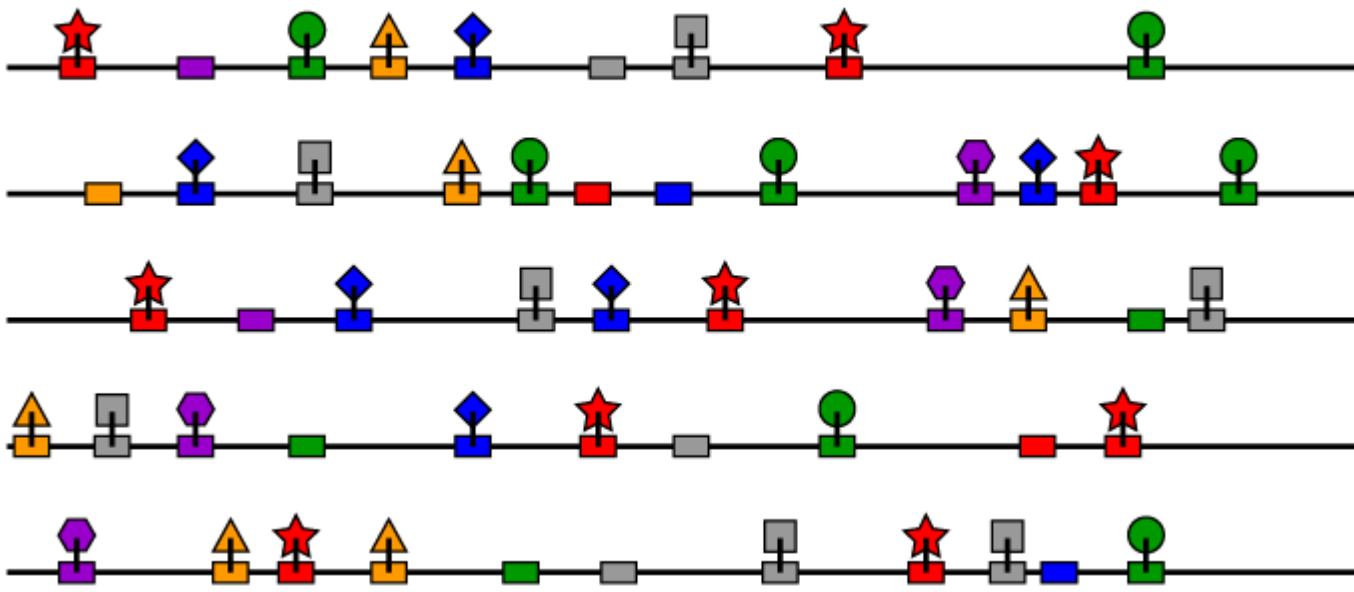
Białka:

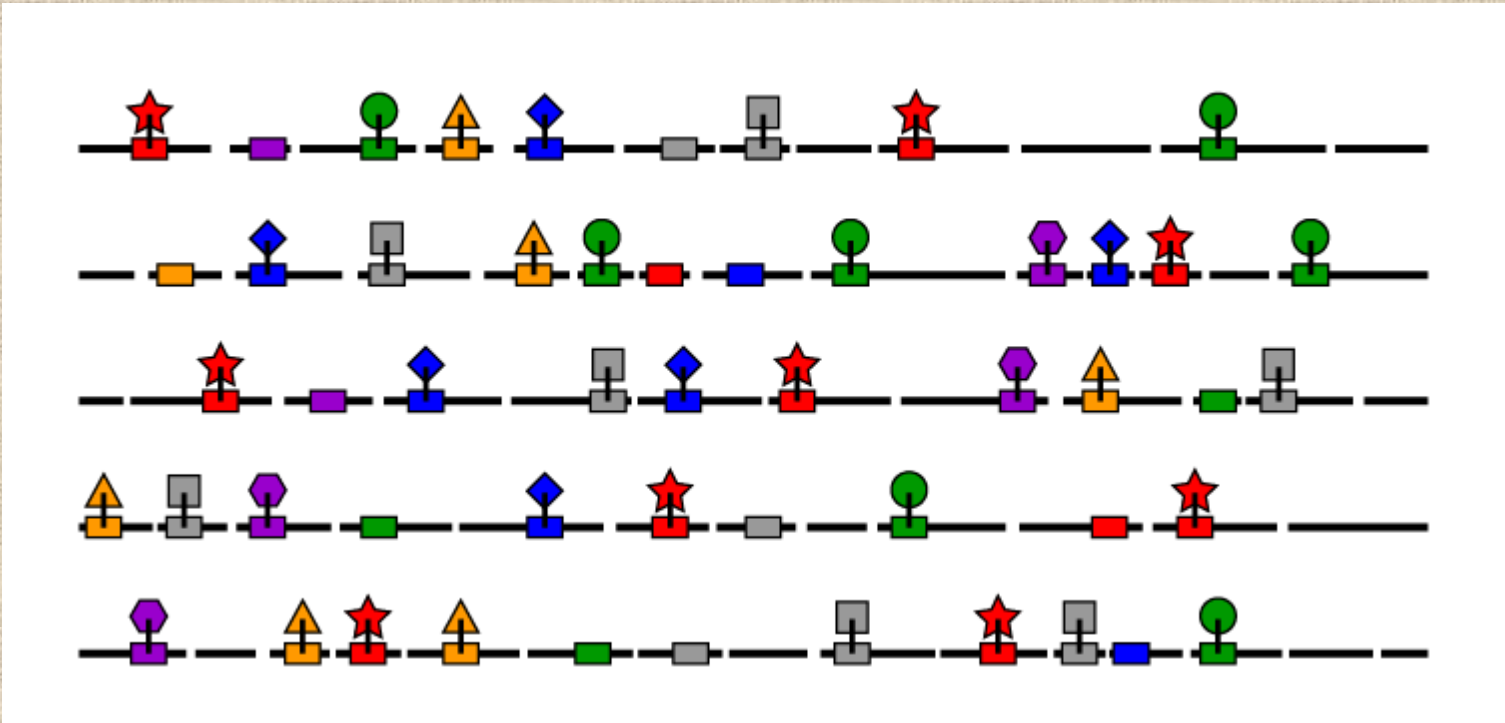
czynniki transkrypcyjne, histony, białka związane z replikacją, naprawą DNA etc.

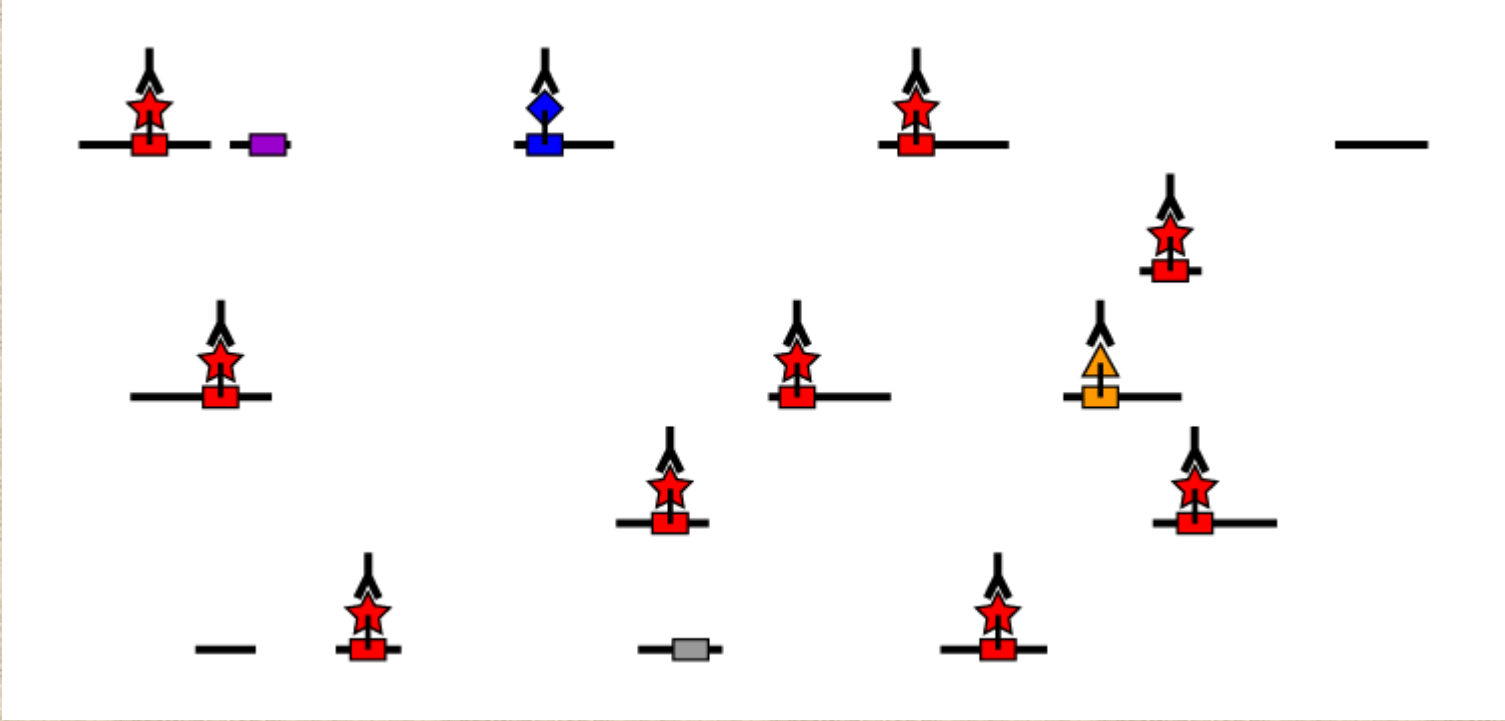
Regiony w DNA:

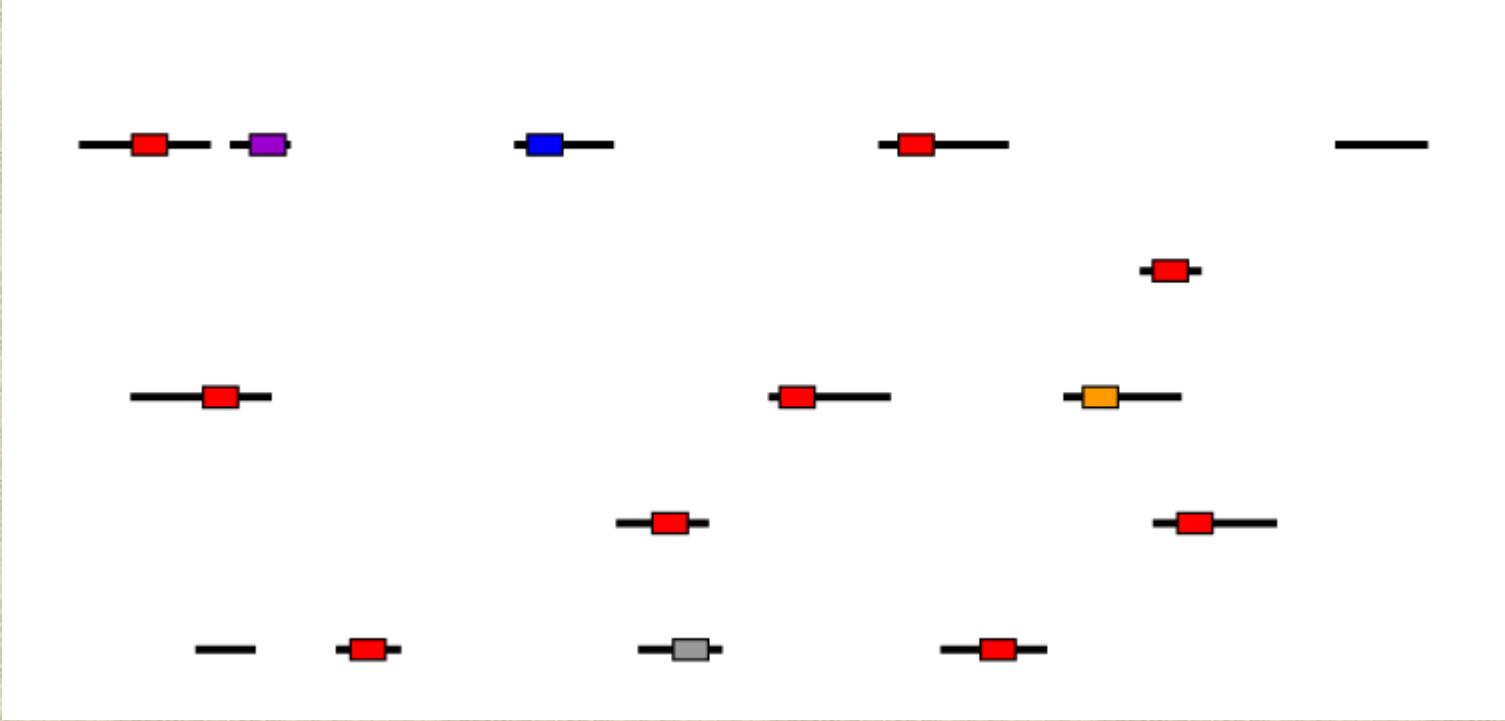
sekwencje wiążące TFs, miejsca startu replikacji, regiony startu transkrypcji, miejsc metylacji DNA etc.



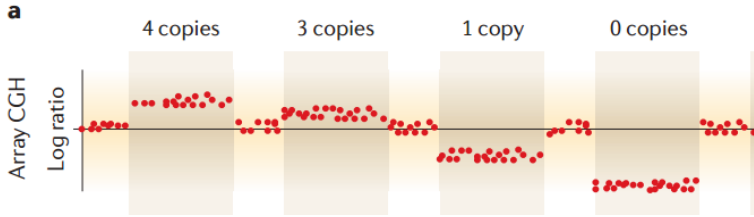
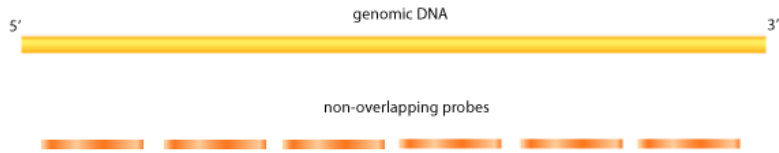
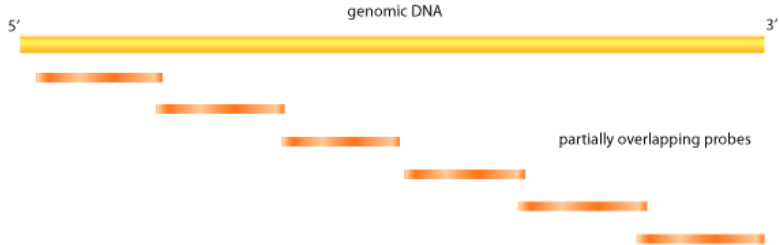
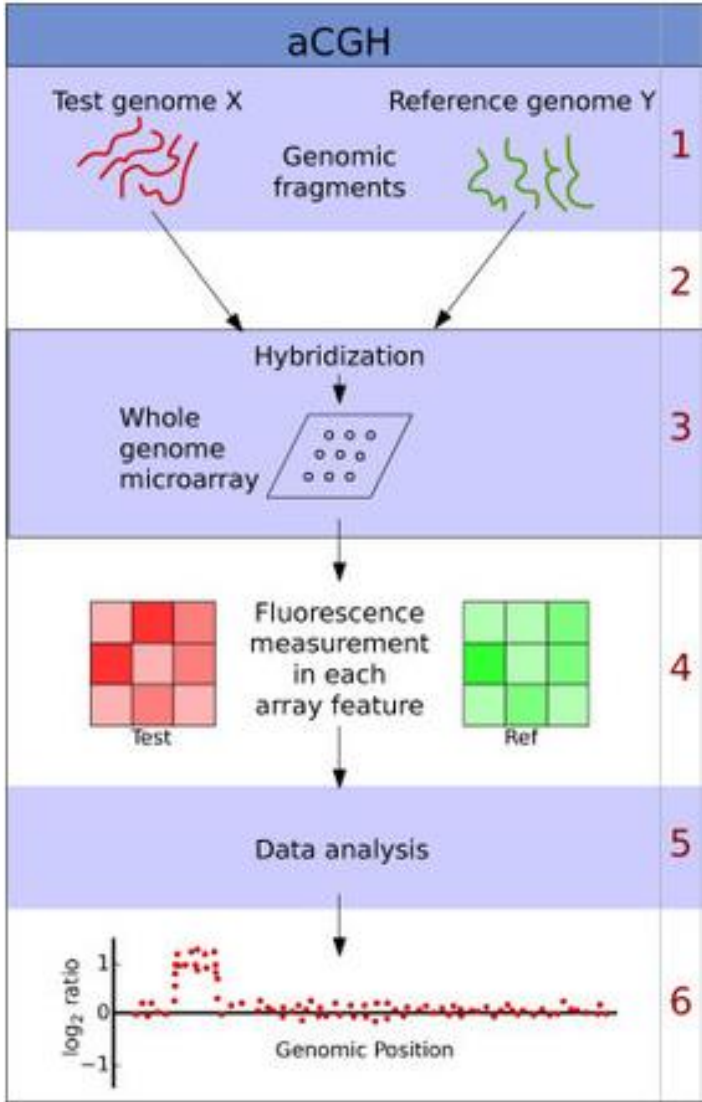






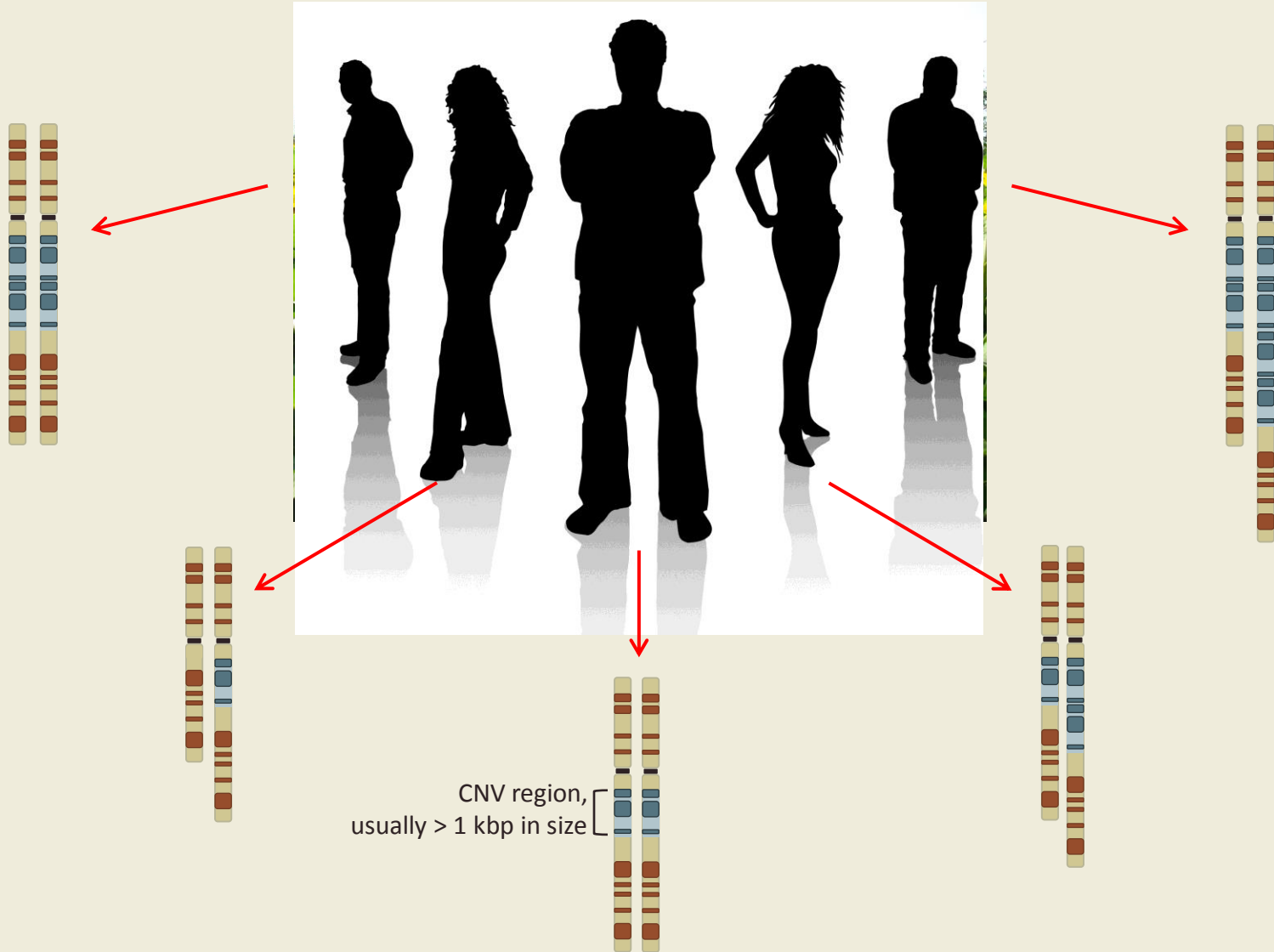


CGH = array-based comparative genomic hybridization

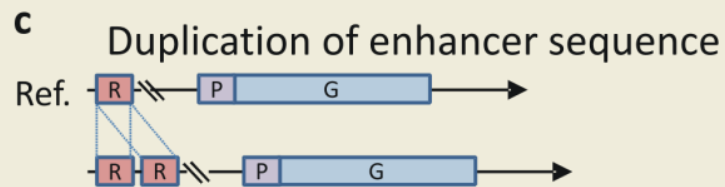
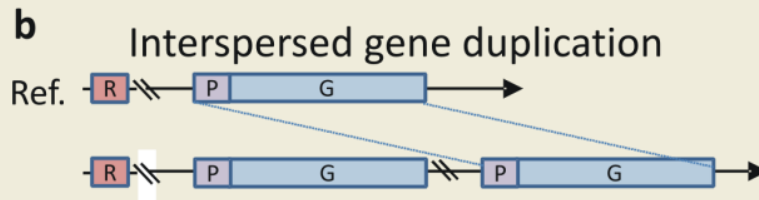
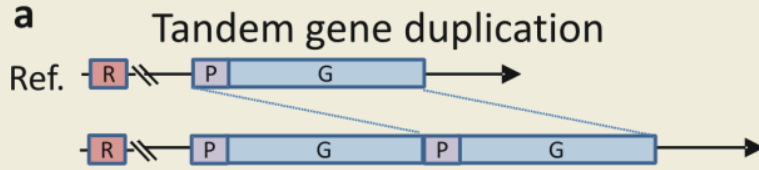


Alkan et al. 2011 Nat. Rev. Genet. 12: 363–376

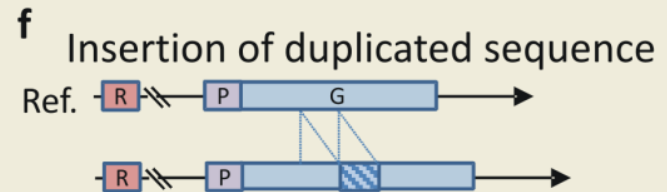
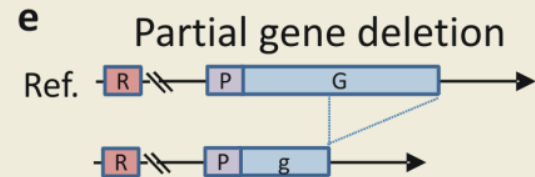
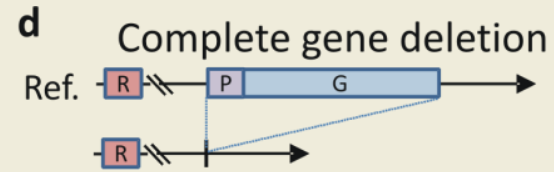
CNV = copy number variation



Significance:



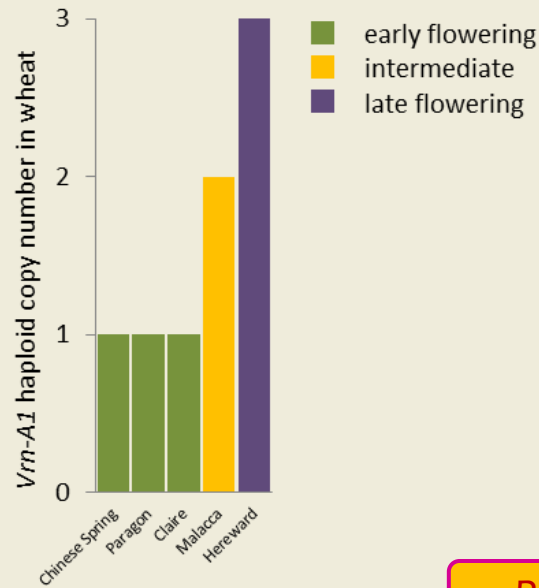
elevated level of full-length transcript



decreased level of full-length transcript

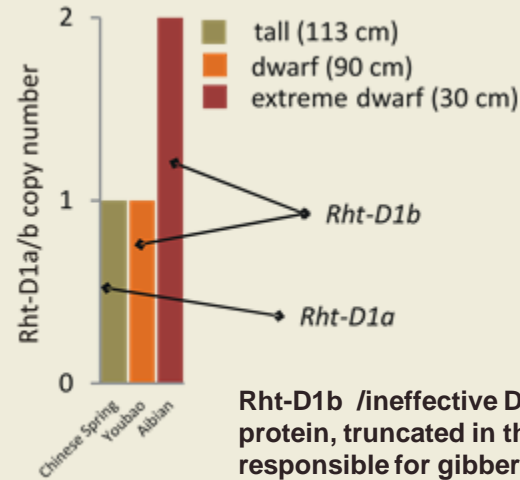
CNV controls agronomic traits in wheat

Flowering



Vrn-1 A / MADS-box transcription factor

Height



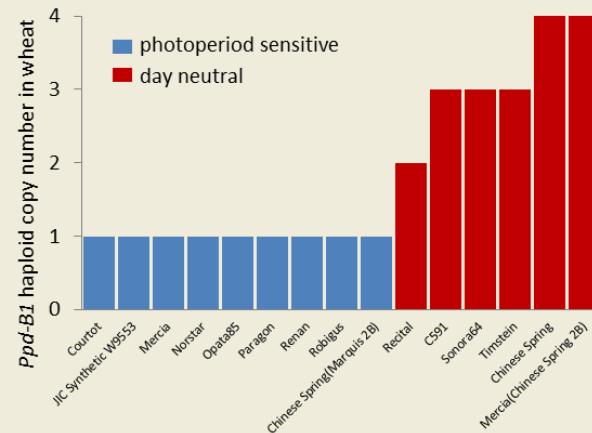
Rht-D1b / ineffective DELLA protein, truncated in the region responsible for gibberellic acid response



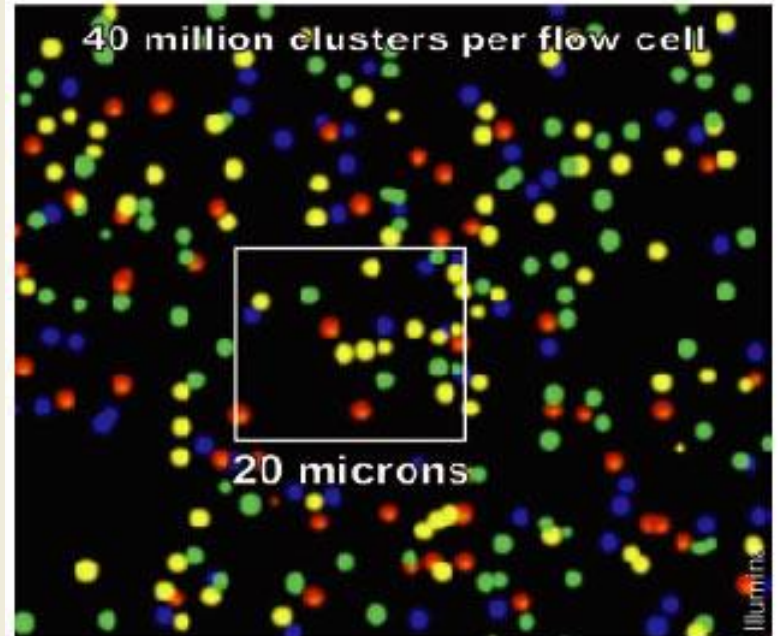
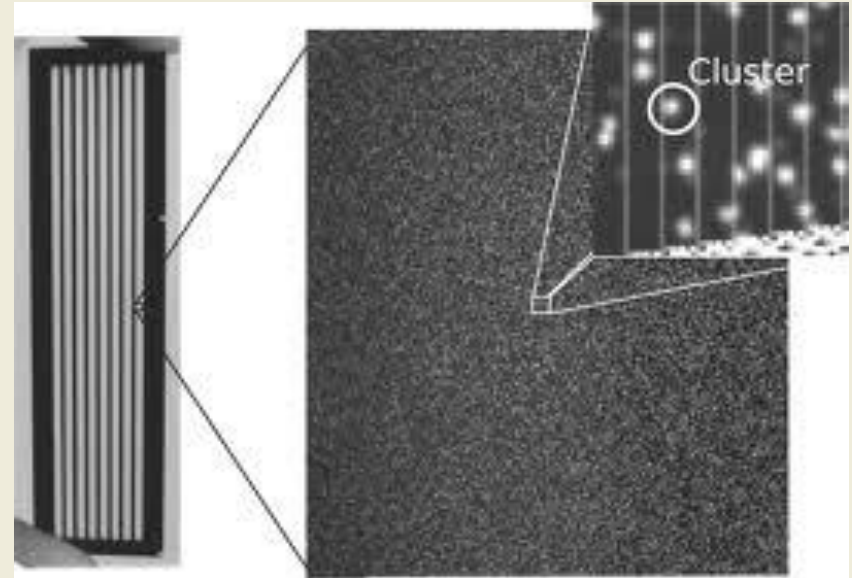
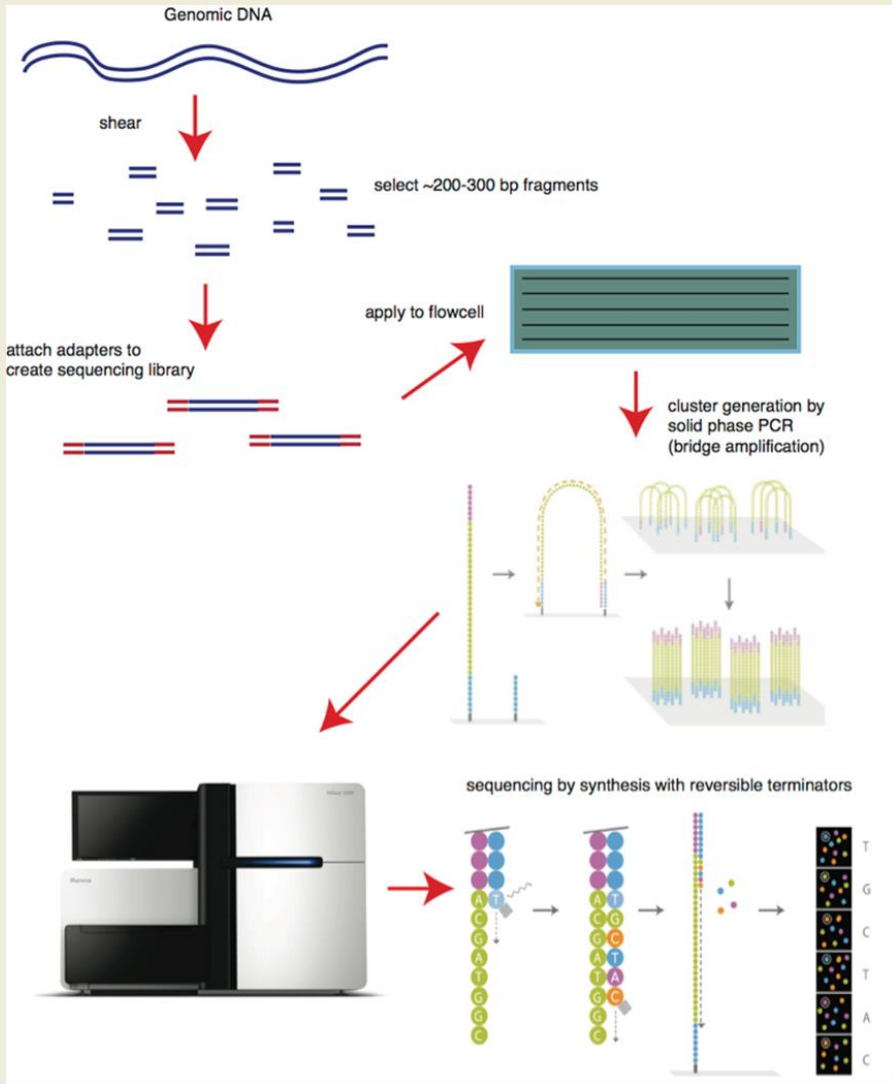
Li et al. 2012 New Phytol 196:282-291

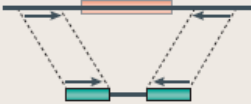


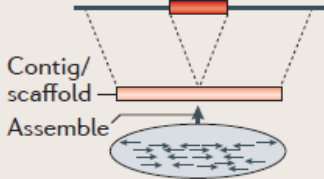

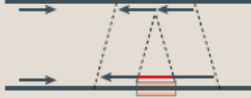
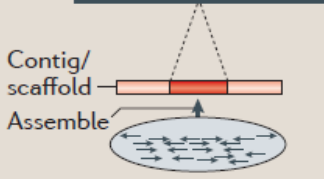
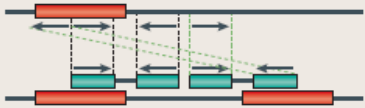
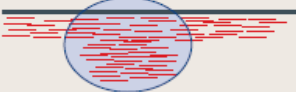
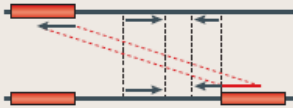
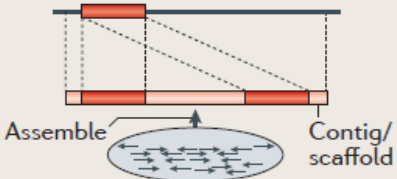

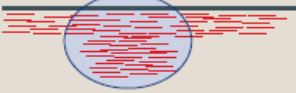

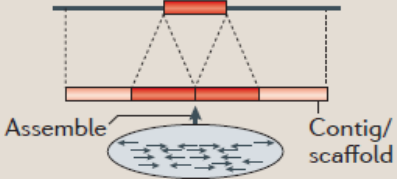
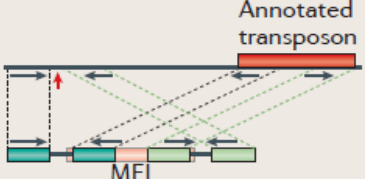
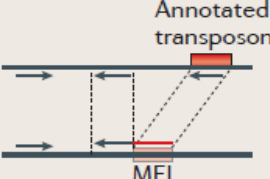
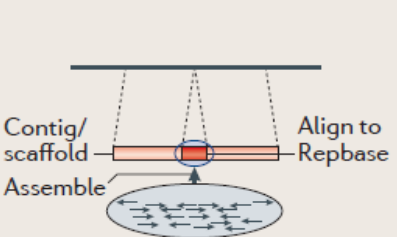
Photoperiod

Ppd-B1 / family member of pseudo response regulators (PRR)



NGS = whole genome-sequencing



SV classes	Read pair	Read depth	Split read	Assembly
Deletion				
Novel sequence insertion		Not applicable		
Interspersed duplication				
Tandem duplication				
Mobile-element insertion		Not applicable		

Methods:

Focused (specific CNV; a few or many samples)

quantitative PCR

fluorescence *in situ* hybridization (FISH)

the paralogue ratio test

multiplex amplifiable probe hybridization

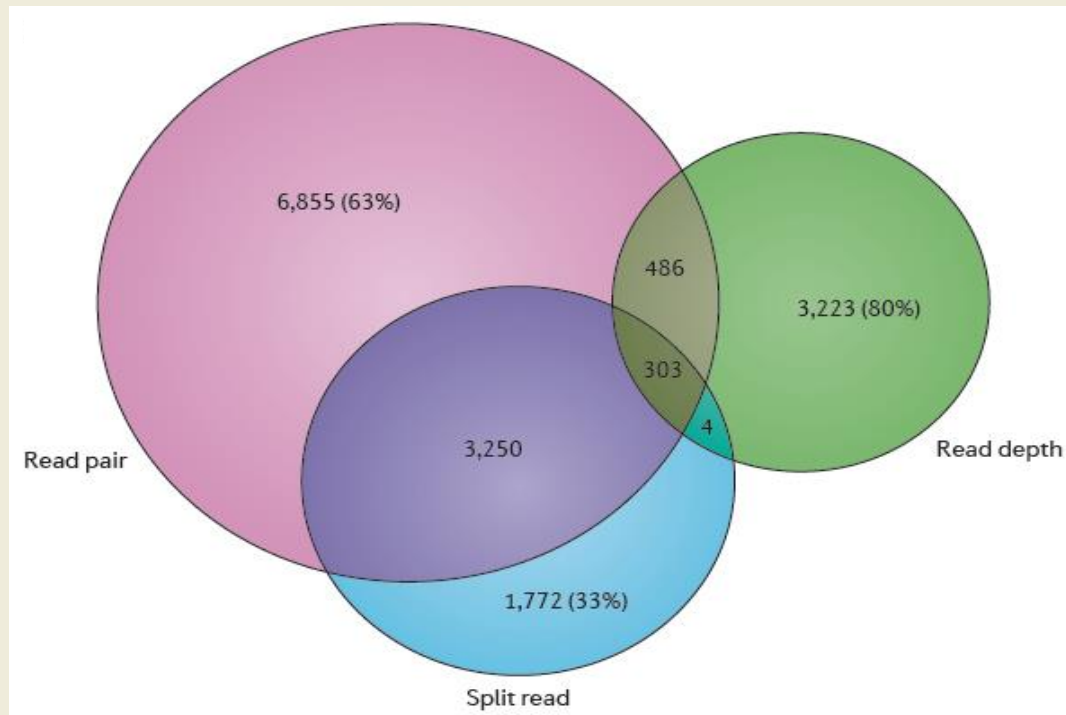
multiplex ligation-dependent probe amplification (MLPA)

Global (genome-scale CNV discovery; usually a limited number of samples)

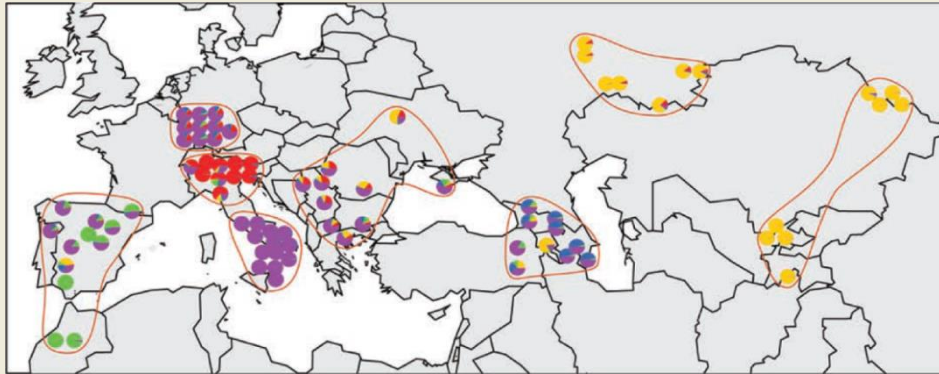
SNP arrays

array-based CGH

DNA sequencing (read pair, read depth, split read, *de novo* assembly)



CNV detection workflow



Cao et al. (2011) Nature Genetics 43: 956-963

MPICao2010 dataset (80 genomes)

Data filtering and quality check

SE mapping

PE mapping

Detection method evaluation
and CNV prediction

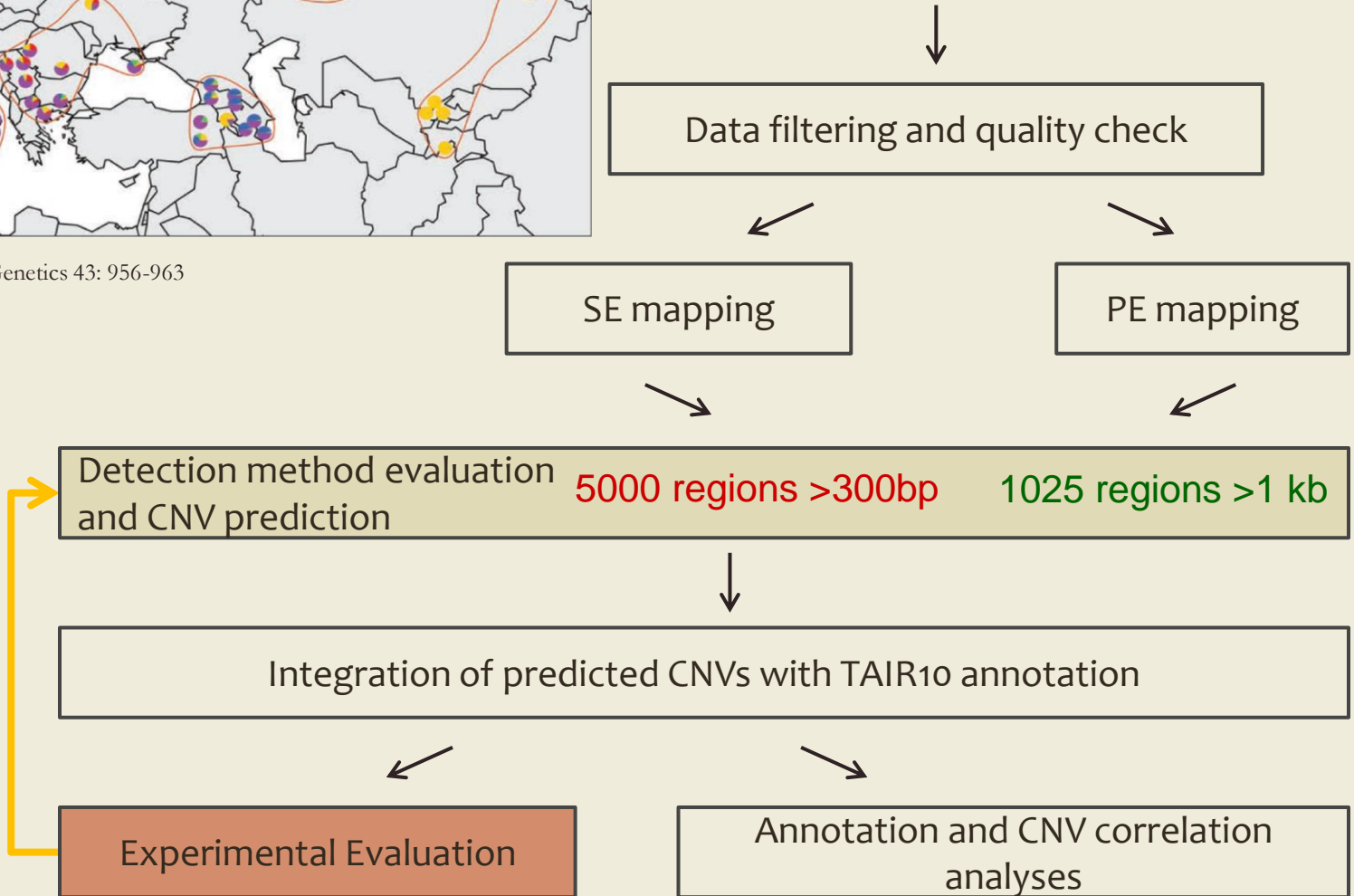
5000 regions >300bp

1025 regions >1 kb

Integration of predicted CNVs with TAIR10 annotation

Experimental Evaluation

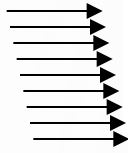
Annotation and CNV correlation
analyses



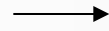
wt



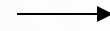
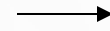
F₀



F₁

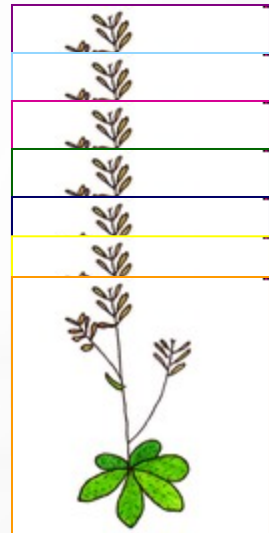
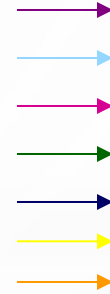
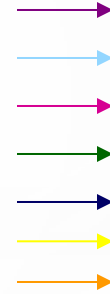
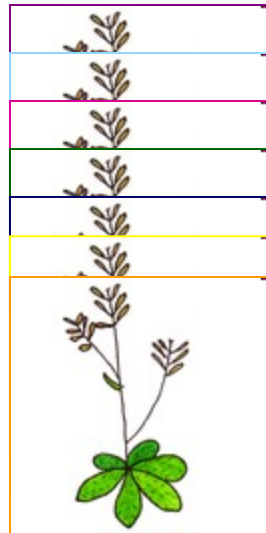
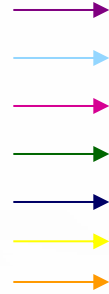
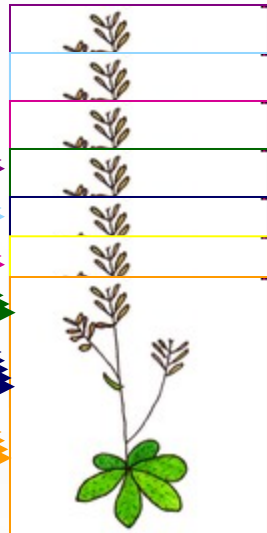
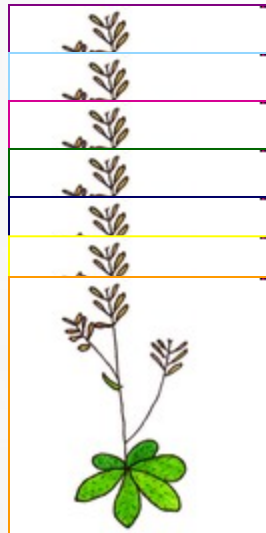


F₂



F_n

dcl



28 °C

22 °C

L2. Bioinformatyczne metody identyfikacji regionów o zmiennej liczbie kopii (CNV) w genomach roślinnych

Praca przeglądowa, w założeniu stanowi podstawę do projektowej pracy magisterskiej

L3. Globalna analiza bioinformatyczna genów podlegających naturalnej zmienności liczby kopii w ekotypach *Arabidopsis thaliana*

*Praca ukierunkowana na bioinformatyczną analizę struktury i funkcji wybranych genów zidentyfikowanych w regionach podlegających zmienności CNV u *Arabidopsis thaliana**