


# Nucleic acid microarray data analyses in potato – virus interaction studies



Maruša Pompe-Novak, Špela Baebler, Hana Krečič-Stres,  
Kristina Gruden, Maja Kovač, Maarten Jongsma, Maja Ravnikar

National Institute of Biology, Večna pot 111, 1000 Ljubljana, Slovenia

Jožef Stefan Institute, Jamova 39, 1000 Ljubljana, Slovenia

Plant Research International, Postbus 16, 6700 AA Wageningen, The Netherlands

# Nucleic acid microarray data analyses in potato – virus interaction studies

---

- mainly microarray work in potato – virus interaction studies
- mention also other microarray work in our lab
- some optimizations done in our lab along the whole procedure of microarray work
- the main stress on the data analysis

# The main research areas of the Department of Plant Physiology and Biotechnology :

---

- plant – pathogen and plant – pest interactions
- development of high through-put molecular methods, including work on preparation of microarray for the detection of genetically modified organisms
  
- Microarrays in our lab
  - mainly used for identification of differentially expressed genes in plant – pathogen interactions
  - till now detection of plant viruses with microarrays side effect of our work
  - we joined in the ring test for validation of microarrays for detection of plant viruses

# Identification of differentially expressed genes in plant – pathogen interactions:

---

- potato – PVY<sup>NTN</sup> and PVY<sup>N</sup>
- grapevine – Bois noir and Flavescence doree
- colorado beetle – potato – PVY<sup>NTN</sup>
  - colorado beetle's adaptation to plant's defense mechanism is studied by characterization of digestive enzymes from guts of larvae

# Studies of potato – PVY interactions:

---

- discover mechanisms of disease development and resistance
- identify the genes, proteins and signaling molecules, which are involved in resistance mechanisms of potato and other plants to pathogens
- studying expression of potato genes involved in gene silencing
  - for searching for mechanism of RNAi in transgenic potato plants resistant to PVY<sup>NTN</sup>
  - in order to evaluate resistance in plants with inserted viral coat protein

The rest of the talk only about potato – PVY interaction studies in order to discover mechanisms of disease development and resistance

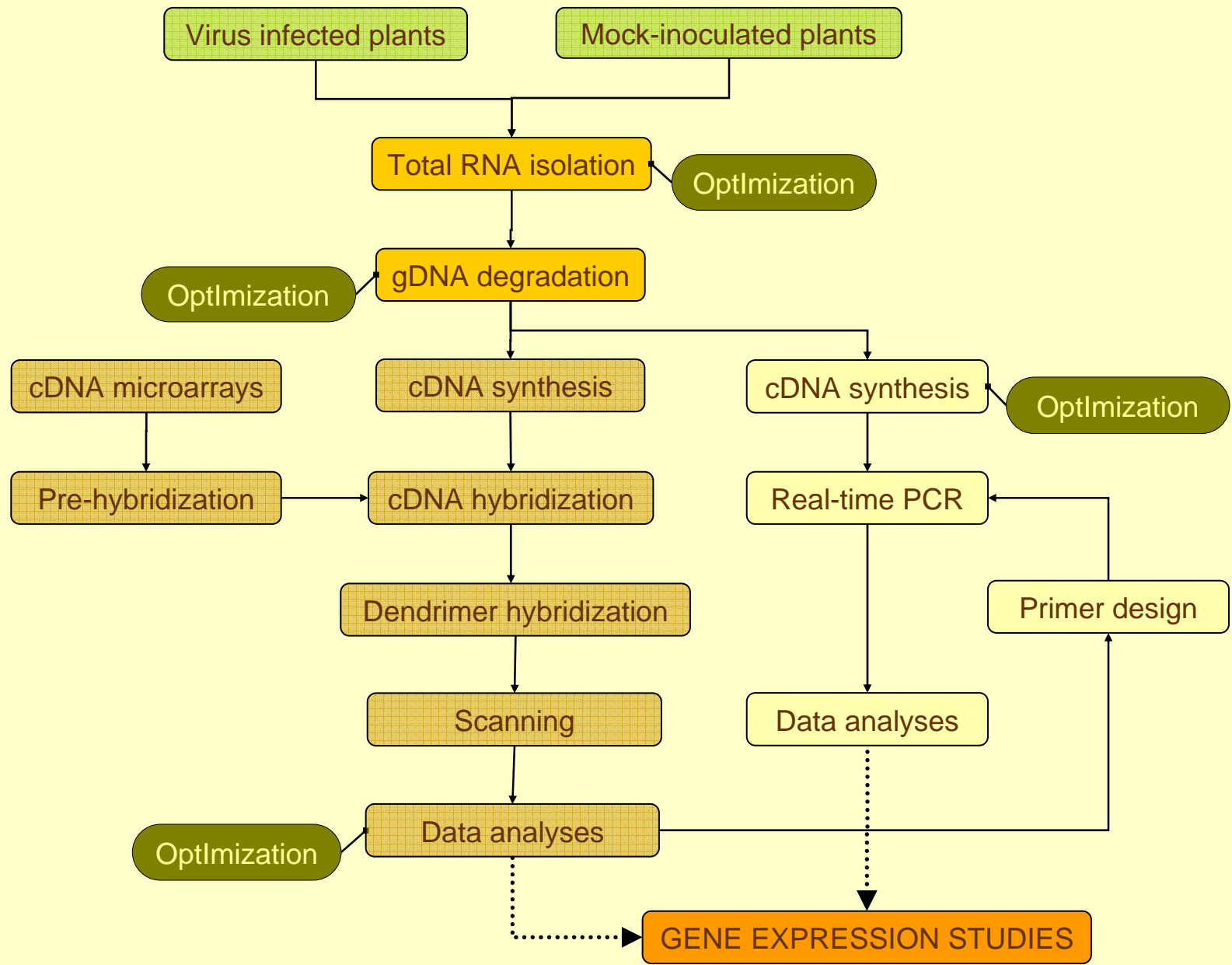
## Potato – PVY interactions:

highly sensitive cultivar Igor	PVY <sup>NTN</sup>	7 and 14 days and sec. inf. plants
resistant cultivar Sante	PVY <sup>NTN</sup>	30 minutes, 6, 12 and 24 hours
highly sensitive cultivar Igor resistant cultivar Sante	PVY <sup>NTN</sup>	30 minutes and 12 hours
highly sensitive cultivar Igor sensitive cultivar Desiree tolerant cultivar Pentland squire resistant cultivar Carlingford resistant cultivar Sante	PVY <sup>NTN</sup>	7 days
highly sensitive cultivar Igor sensitive cultivar Nadin	PVY <sup>N</sup> PVY <sup>NTN</sup>	30 minutes, 12 and 48 hours

## Sets of cDNA microarrays:

---

- ❑ self-made 400 clones potato cDNA microarrays spotted at Plant Research International in Wageningen in The Netherlands
- ❑ self-made 4000 clones potato cDNA microarrays spotted at Rikilt in Wageningen in The Netherlands
- ❑ self-made 400 clones colorado beetle cDNA microarrays spotted at Plant Research International in Wageningen in The Netherlands
- ❑ TIGR 10000 clones potato microarrays provided by The Institute for Genomic Research from Maryland, USA.
  
- ❑ differentially expressed gene libraries for producing self-made cDNA microarrays
  - ❑ from highly sensitive cultivar Igor to obtain genes involved in disease development
  - ❑ from resistant cultivar Sante to obtain genes responsible for resistance.



## Total RNA isolation:

---

- Isolation of total RNA from potato leaf tissue with RNeasy Plant Mini Kit (Qiagen) was improved to increase the capacity of the kit 3-fold.

Sample No.	totRNA from 100 mg		totRNA from 300 mg		Conc. ratio (300/100 mg)
	Conc. ( $\mu\text{g}/\mu\text{l}$ )	260/280 ratio	Conc. ( $\mu\text{g}/\mu\text{l}$ )	260/280 ratio	
1	0.17	2.16	0.72	2	4.1
2	0.2	1.97	0.71	2.01	3.6
3	0.19	2.04	0.59	1.99	3.2
4	0.26	2.09	0.78	2.01	3
average $\pm$ SD	$0.20 \pm 0.036$		$0.70 \pm 0.078$		

- 10 min of incubation was determined to be critical for efficient totRNA elution.

# Genomic DNA digestion:

---

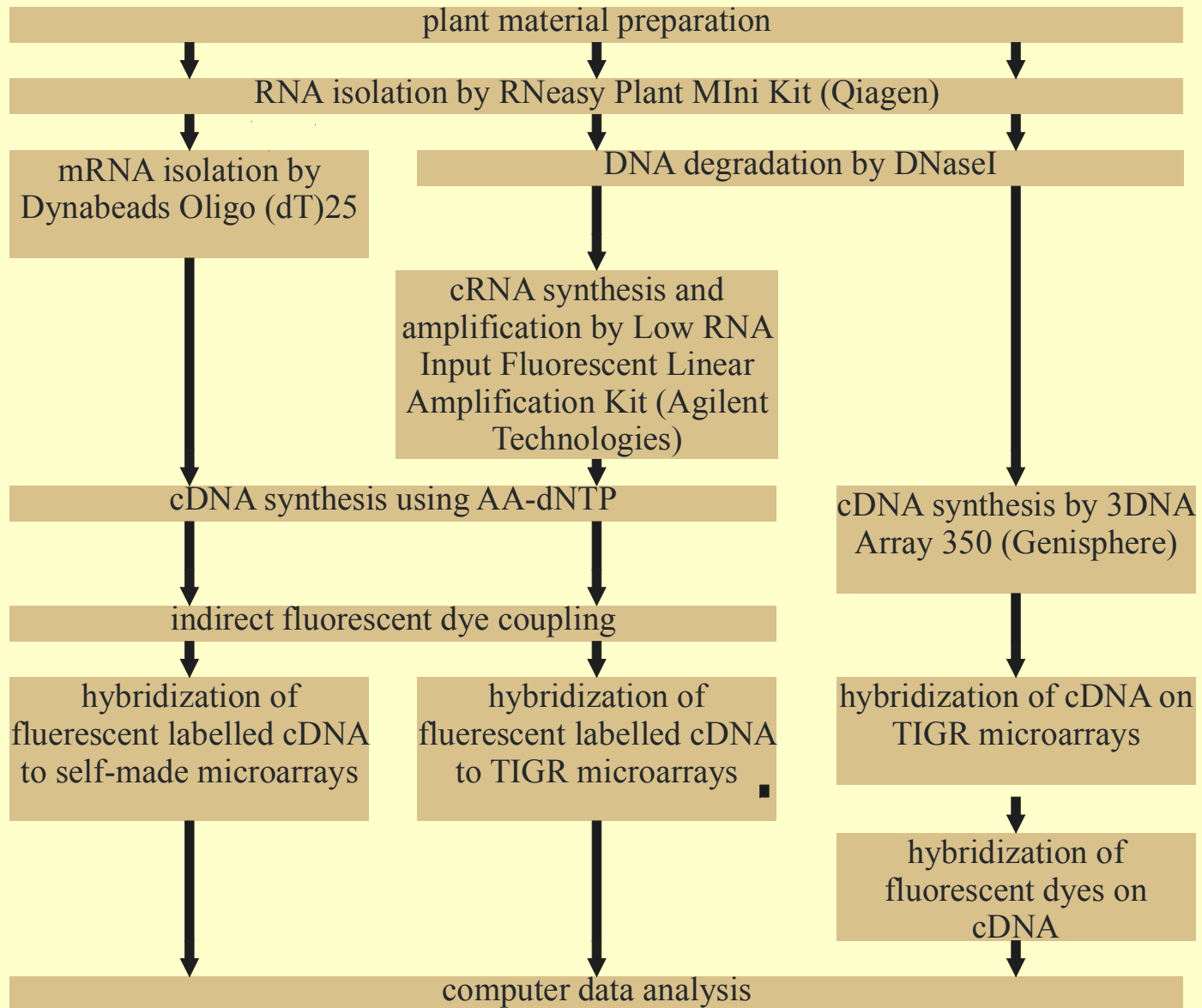
- Two approaches of genomic DNA digestion:
  - Invitrogen's Deoxyribonuclease I - Up to 25-times lower concentrations of DNase were sufficient to remove all residual genomic DNA in total RNA samples.
  - Qiagen's RNase free DNase set - Degradation of residual gDNA by on-column DNase I digestion was not complete.

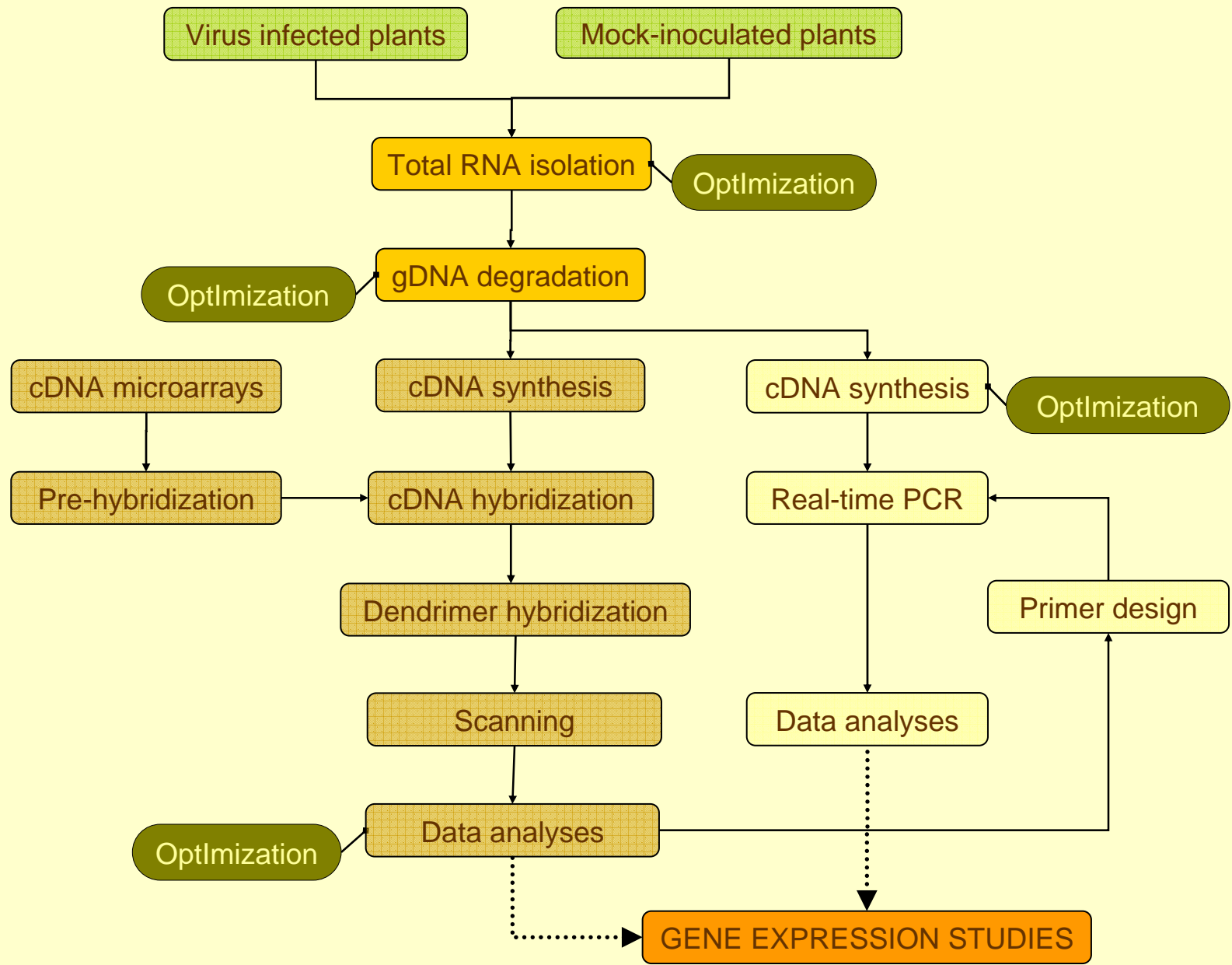
## cDNA synthesis:

---

- Two approaches of cDNA synthesis:
  - High-Capacity cDNA Archive Kit (Applied Biosystems) – optimal results achieved combined with oligo-d(T)<sub>16</sub> primers (Applied Biosystems)..
  - GeneAmp® RNA PCR system – not so efficient.

# Hybridization:





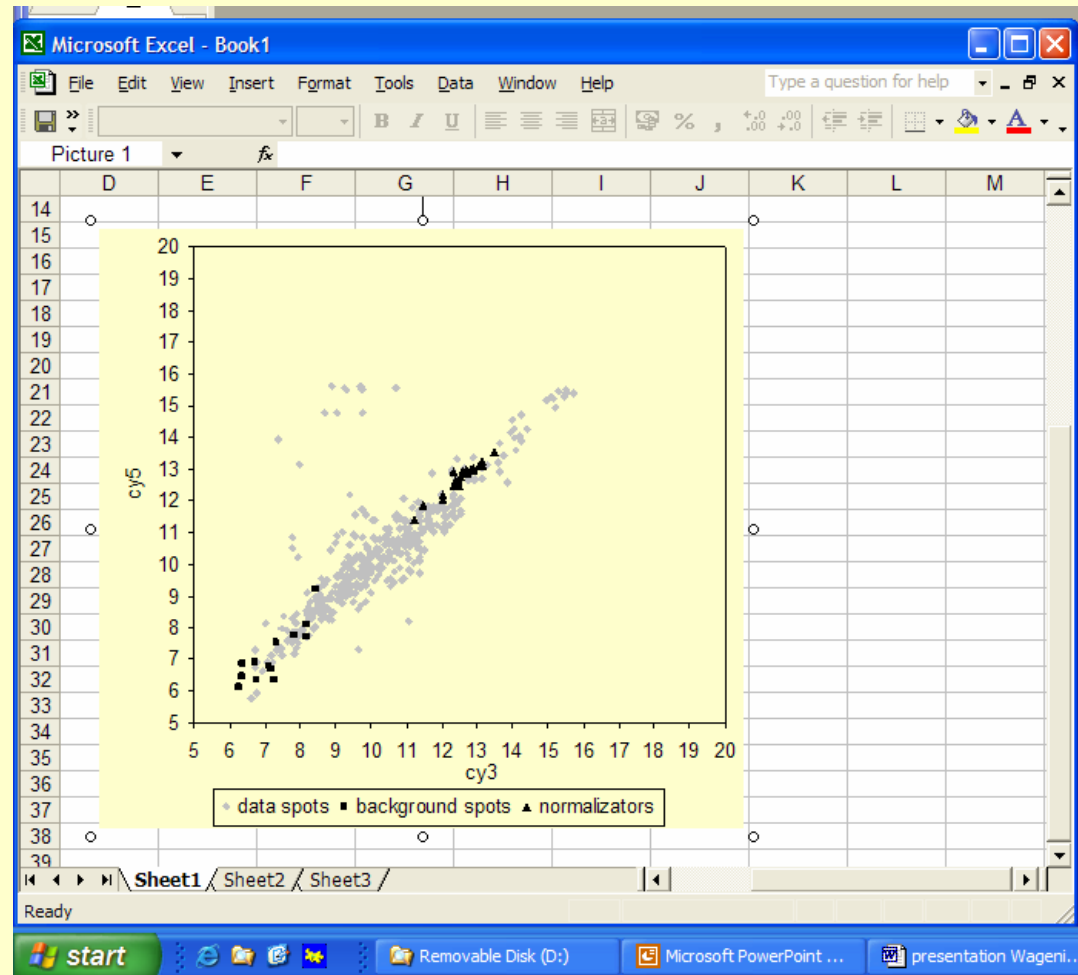
## Data analyses:

---

- in the opposite with diagnostic arrays, in the case of gene expression profiling experiments very complex
- a lot of effort in optimization of data analyses lately
  
- separate procedures for data analyses for
  - 10000 clones arrays
  - 400 clones arrays

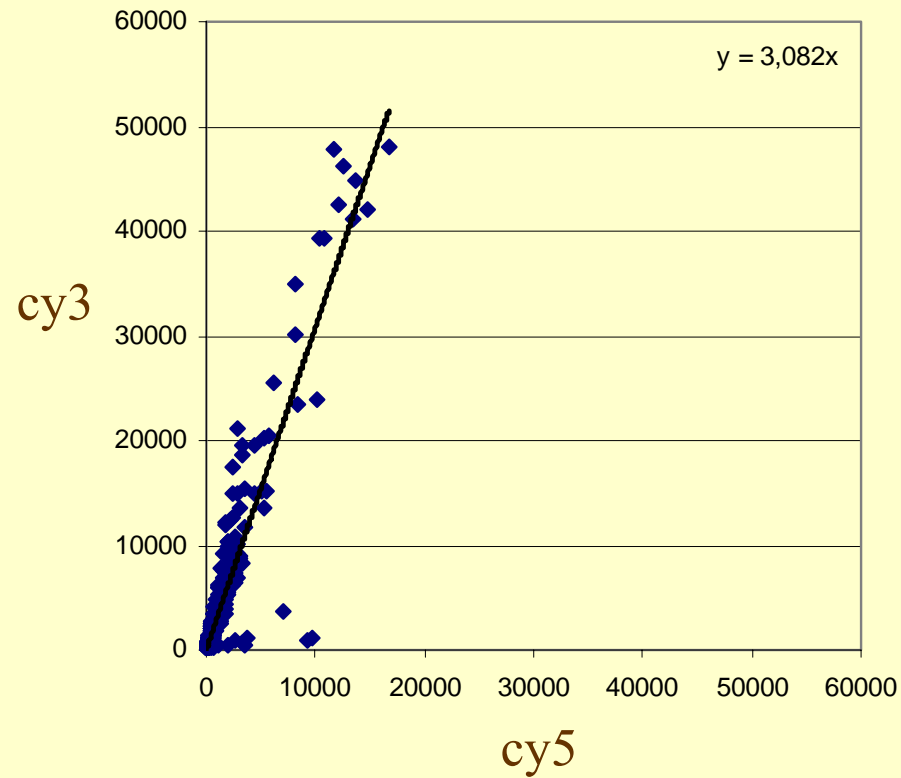
# Self-made 400 clones cDNA microarrays:

- Microsoft Excell



# Raw data:

---



unequal brightness of cy3 and cy5 dyes → normalization is necessary

# Normalization:

---

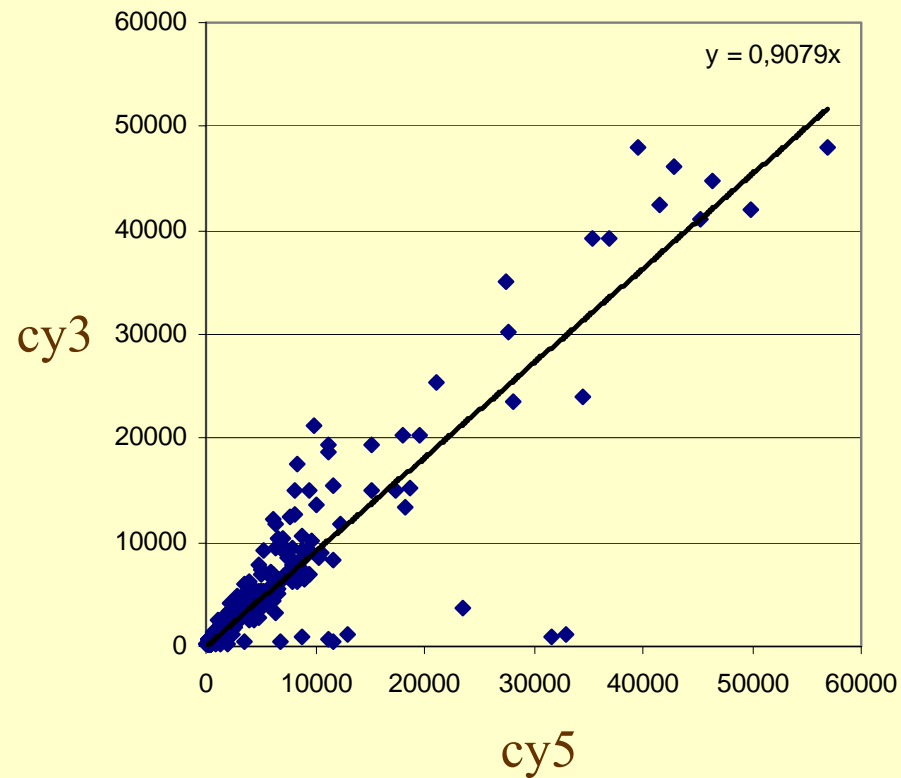
- group of normalization genes
  - **spike-in controls (luciferase...)**
  - internal controls (house keeping genes, 18S rRNA...)
- trendline
- by equalizing the distribution (Xpression, InforMax)
  - by setting average ratio to 1
  - by setting average to the reference

# Normalized data:

---

$$cy3_{\text{normalized}} = cy3_{\text{raw}}$$

$$cy5_{\text{normalized}} = cy5_{\text{raw}} * (\text{average luciferase intensity cy3} / \text{average luciferase intensity cy5})$$



# Background subtraction:

---

- ❑ **group of background spots (yeast DNA...)**
- ❑ background around spots
  - ❑ global background – average over the whole microarray area
  - ❑ local background – background around individual spots
- ❑ subtracted before normalization
- ❑ **subtracted after normalization**
- ❑ channel specific subtraction
- ❑ **subtraction of maximal background**

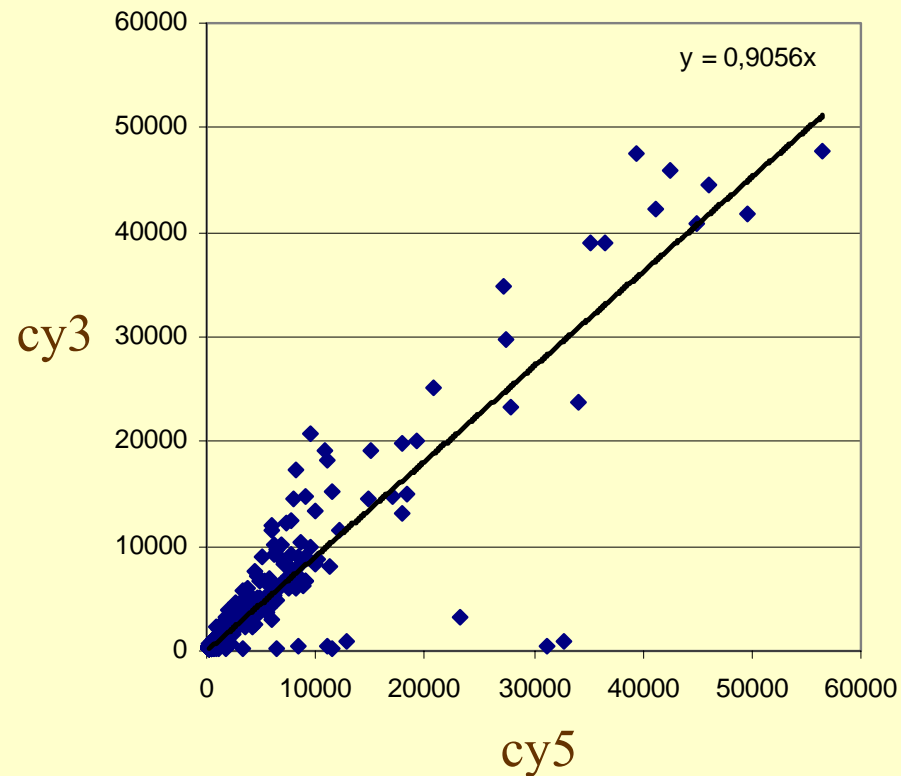
# Subtracted background:

---

background = MAX (average yeast gene intensity cy3; average yeast gene intensity cy5)

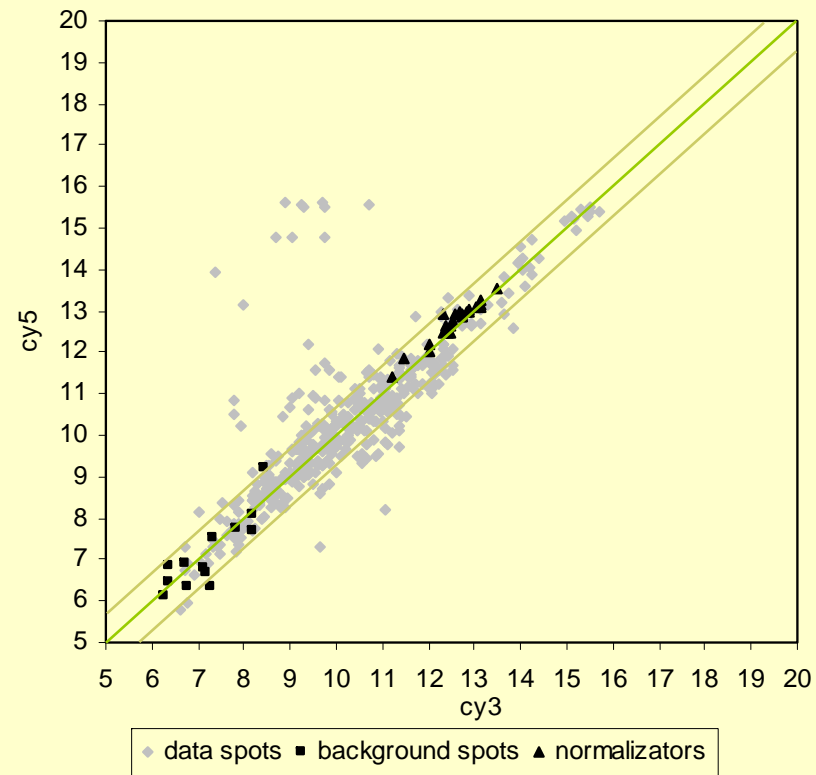
$cy3_{sub.back.} = IF (cy3_{norm.} - background > background; cy3_{norm.} - background; background)$

$cy5_{sub.back.} = IF (cy5_{norm.} - background > background; cy5_{norm.} - background; background)$



# Log<sub>2</sub> of processed data:

---



# Statistical significance:

---

- cv – Coefficient of Variability

cv = standard deviation / average

$$cv = \sqrt{cv_1^2 + cv_2^2 + \dots + cv_n^2 + cv_{\text{between}}^2}$$

- **Student's T-test** or Mann-Whitney U-test

p < 0.01 \*\*\*

p < 0.1 \*\*

p < 0.5 \*

# TIGR 10000 clones cDNA microarrays:

---

- ArrayPro
- R
- MEV (TM4)
- MapMan

BAEBLER, Špela, HREN, Matjaž, KOGOVŠEK, Polona, KREČIČ STRES, Hana, CURK, Tomaž, JUVAN, Peter, ZUPAN, Blaž, POMPE NOVAK, Maruša, GRUDEN, Kristina. 2005. *Laboratory and computer practice protocols*. Ljubljana: Department of Plant Physiology and Biotechnology, National Institute of Biology.

# ArrayPro:

---

Ignored spots:

- ❑ non-validated spots
- ❑ missing spots
- ❑ not uniform spots (smeared, doughnut shape...)
  - $\text{Raw\_Int\_TM\_cy3} / \text{Raw\_Int\_SD\_cy3} < 1$  OR
  - $\text{Raw\_Int\_TM\_cy5} / \text{Raw\_Int\_SD\_cy5} < 1$
- ❑ spots with not uniform background
  - $\text{Raw\_Int\_TM\_cy3} / \text{Back\_Int\_SD\_cy3} < 3$  OR
  - $\text{Raw\_Int\_TM\_cy5} / \text{Back\_Int\_SD\_cy5} < 3$
- ❑ spots with low signal intensity
  - $\text{Raw\_intensity\_TM\_cy3} - \text{Background\_TM\_cy3} < 1.5 * \text{Background\_TM\_cy3}$  AND
  - $\text{Raw\_intensity\_TM\_cy5} - \text{Background\_TM\_cy5} < 1.5 * \text{Background\_TM\_cy5}$

# R:

---

- Cross-channel normalization by Local regression (Loess)
- Background subtraction

Net\_Int

Raw\_intensity\_TM\_cy3 - Background\_TM\_cy3

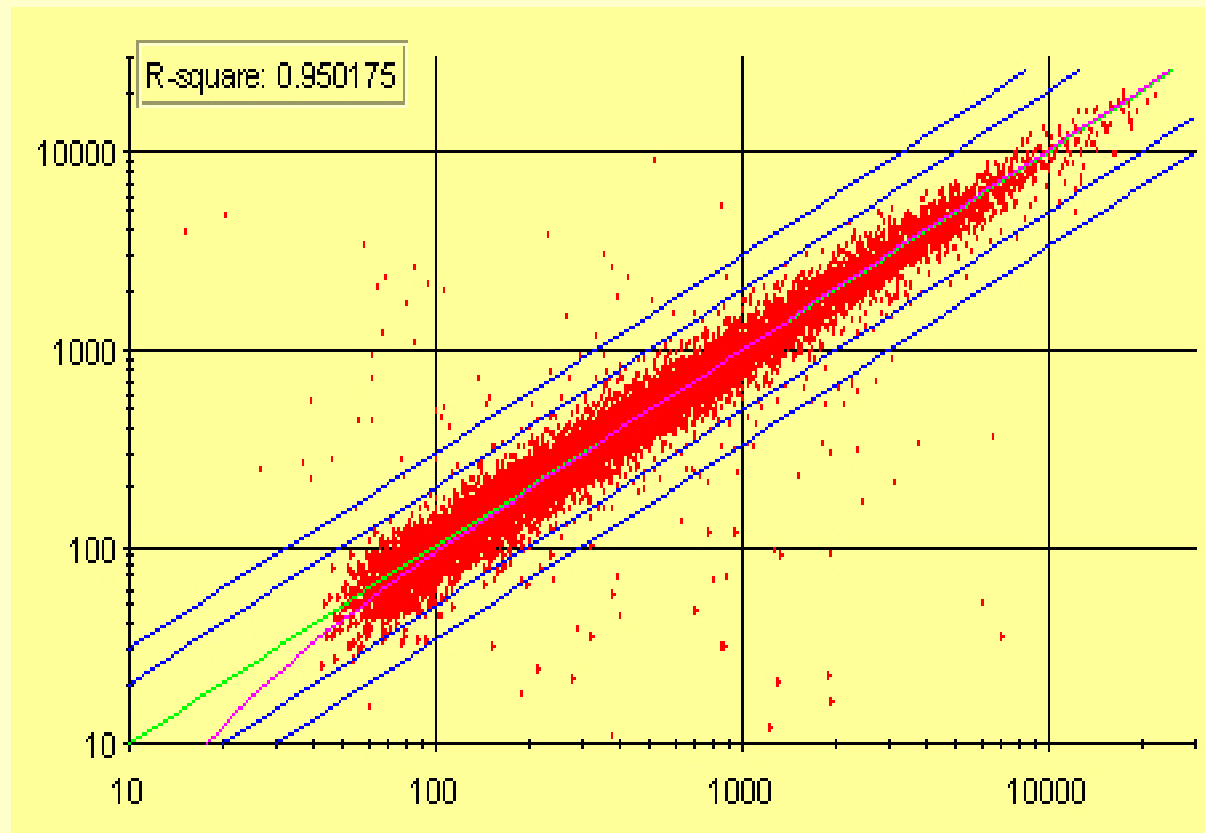
Raw\_intensity\_TM\_cy5 - Background\_TM\_cy5

- $M = \log_2(\text{Net\_Int\_cy3}/\text{Net\_Int\_cy5})$
- $A = \log_2(\text{Net\_Int\_cy3} * \text{Net\_Int\_cy5} / 2)$
- Replicated spots average
- Filtration

$-0.2 < M < 0.2$

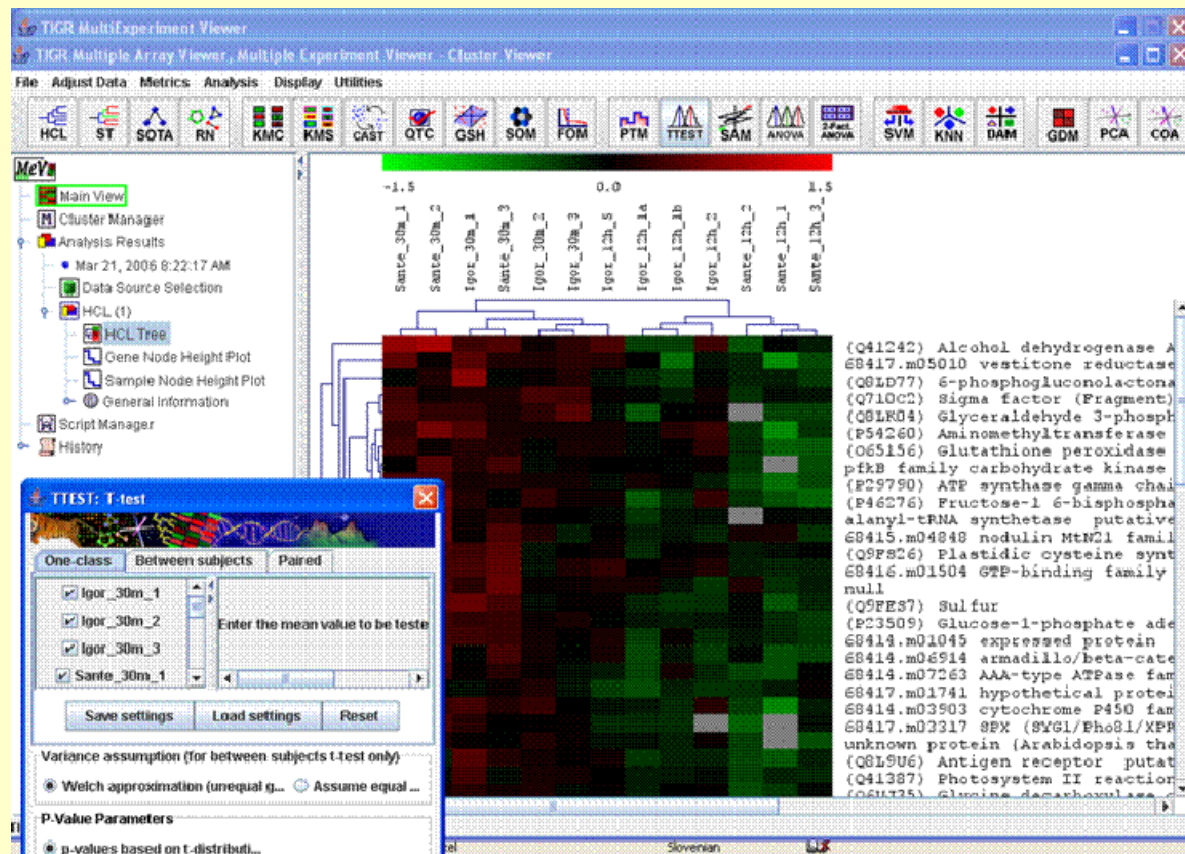
# Processed data:

---



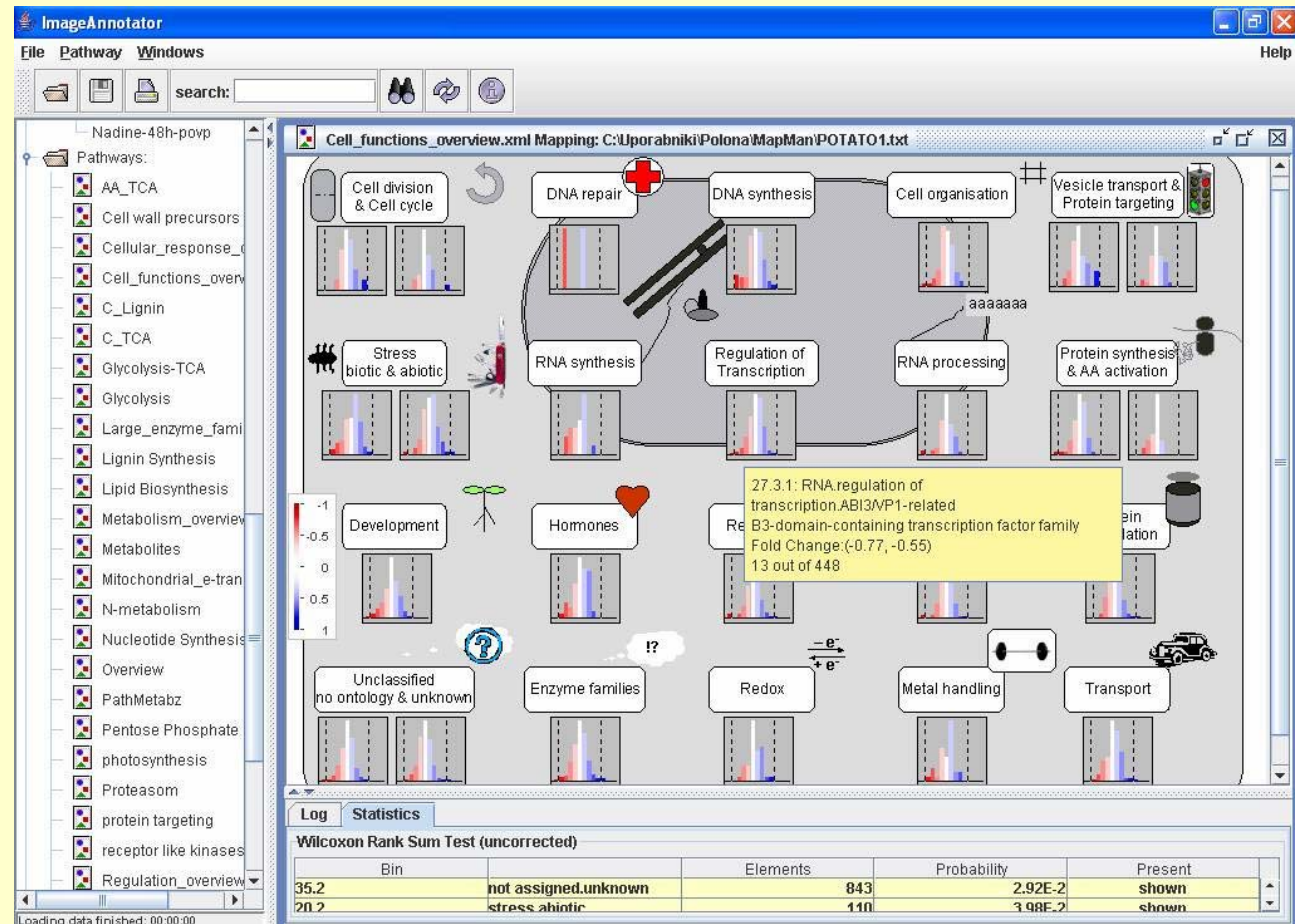
# Microarray Experiment Viewer (MEV):

- Programme package TM4
- T-test
- ANOVA
- data clustering according to results of statistical tests



# MapMan:

- gene grouping according to the metabolic pathways



## Some results:

---

- cultivars Igor and Nadine, infected with PVY<sup>N</sup> and PVY<sup>NTN</sup>
- highly sensitive cultivar Igor, infected with PVY<sup>NTN</sup>
  - after local lesions appearance 7 days after inoculation
  - in systemic infection 14 days after inoculation
  - in plants with established infection grown from infected tubers
- resistant cultivar Sante, infected with PVY<sup>NTN</sup>, 12 hours after infection

POMPE NOVAK, Maruša, GRUDEN, Kristina, BAEBLER, Špela, KREČIČ STRES, Hana, KOVAČ, Maja, JONGSMA, Maarten Anthonie, RAVNIKAR, Maja. 2006. Potato virus Y induced changes in the gene expression of potato (*Solanum tuberosum* L.). *Physiol. mol. plant pathol.* 67:237–247.

## Cultivar Igor:

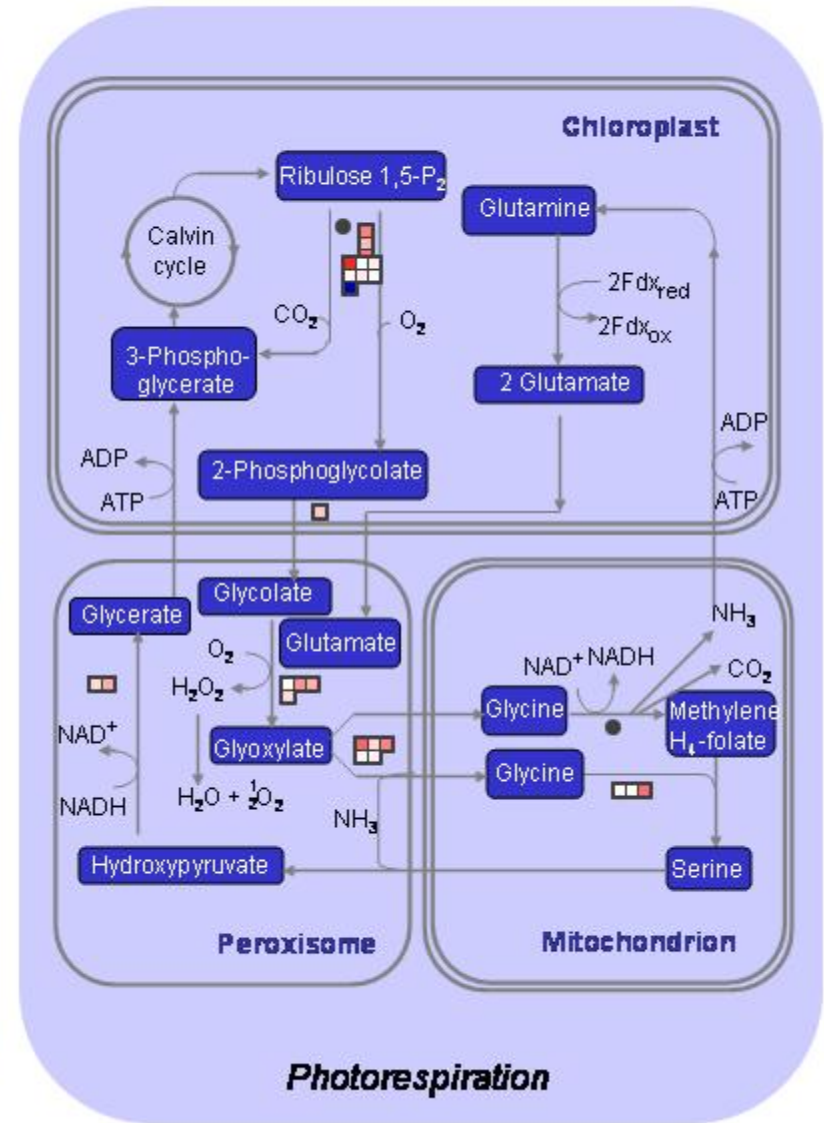
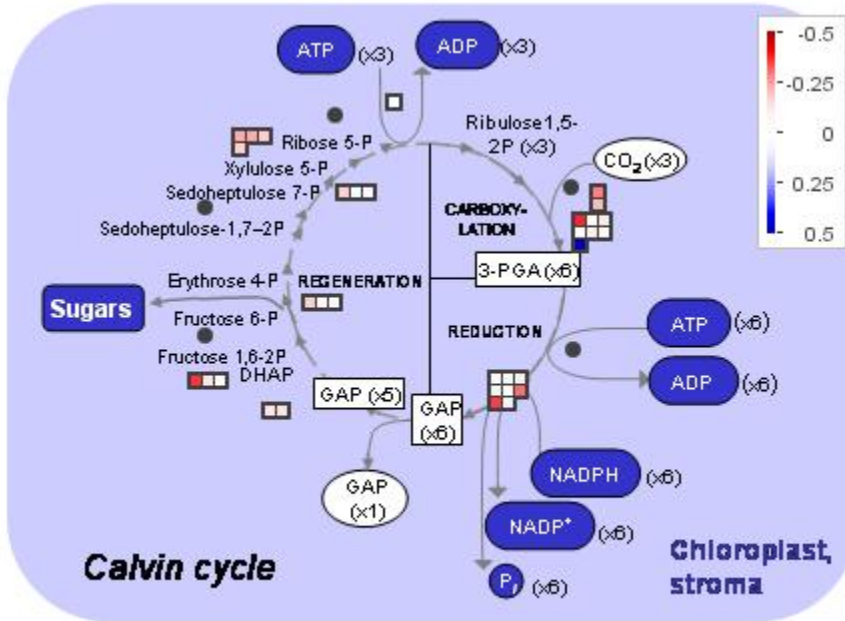
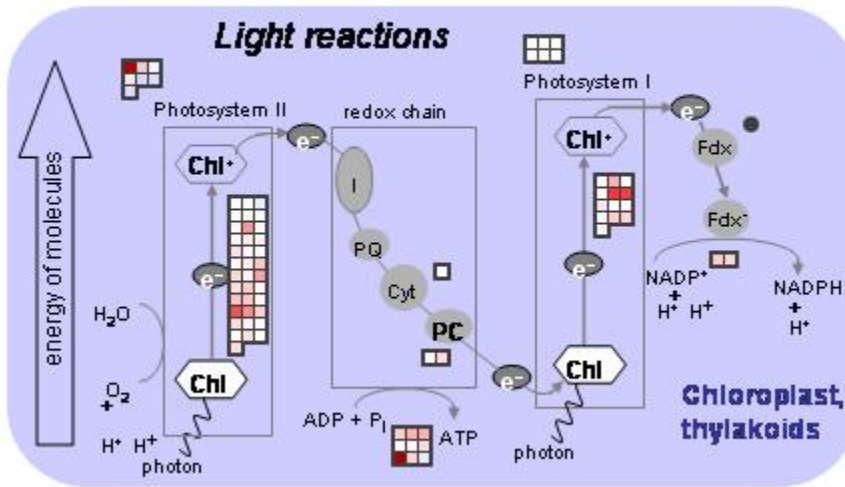
---

out of 184 genes	up-regulated	down-regulated	total regulated
non-inoculated leaves 7 dpi	2	20	22 (12%)
inoculated leaves 7 dpi	7	11	18 (10%)
non-inoculated leaves 14 dpi	3	32	35 (19%)
secondary infected plants	16	10	26 (14%)

# Cultivar Igor:

clone number	name of the most similar sequence form database	log2 of the ratio between gene expression in infected and healthy plants							
		upper non-inoculated leaves 7 days after inoculation		lower inoculated leaves 7 days after inoculation		lower non-inoculated leaves 14 days after inoculation		lower leaves of secondary infected plants	
		average	p	average	p	average	p	average	p
pot 190	heat shock protein 70 [2/2]	-0,2697	**	-0,7520	*	-1,2264	***	-0,8117	***
pot 015	glycine-rich RNA binding protein [2/2]	-0,5289	*	-0,7681	***	-0,5416		0,6403	*
pot 258	ABC transporter protein 1	0,3000		0,5449	**	1,3379	*	1,4043	**
pot 197	auxin repressed protein [1/1]	-0,6203	***	-1,1433	***			-0,1355	
pot 315	beta(1,3)glucanase	-0,2227	**	0,3671	*	0,9394		0,6021	*
pot 282	catalase 1 [1b/1]	0,1139		-0,1861		-1,7565	***	-0,2630	*
pot 031	unknown function [24/67]	-0,9003	**	-0,7164	*	-1,6731	***	0,2716	
pot 270	unknown function [43/67]	1,4043	***	1,2629	**			0,9445	

# Cultivar Sante:



---

National Institute of Biology, Ljubljana, Slovenia  
Plant Research International, Wageningen, The Netherlands  
Jožef Stefan Institute, Ljubljana, Slovenia  
Rikilt Institute of Food Safety, Wageningen, The Netherlands  
Central Science Laboratory, York, Great Britain  
University of Bristol, Bristol, Great Britain  
University of Poitiers, Poitier, France  
INRA, Montpellier, France  
University of Udine, Udine, Italy

Thank you!