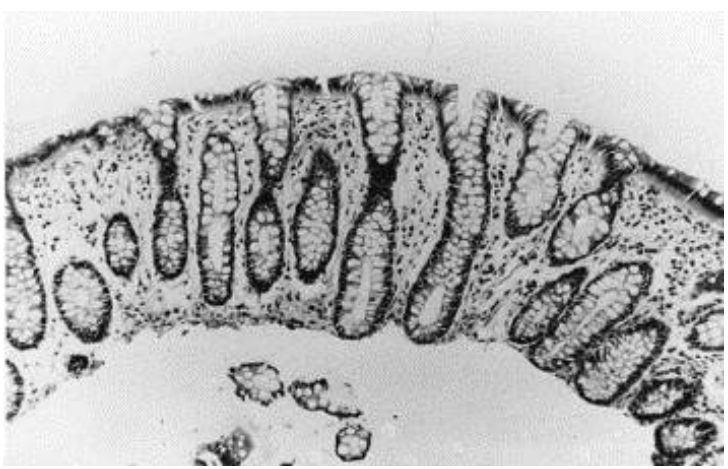
Radiotherapy induced late toxicity predicted by gene expression profiling

J. Peter Svensson
H. Vrieling, M. Giphart-Gassler

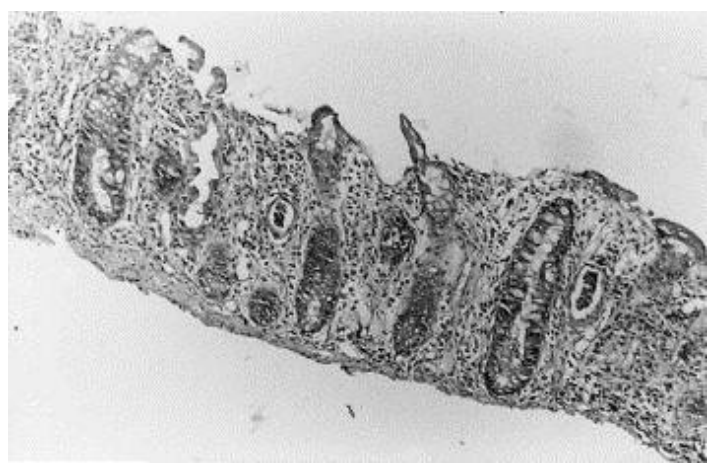
Columbia University, March 17, 2008

Radiation therapy

- 60% of cancer patients receive radiation therapy
- In addition to killing tumour cells, normal cells are damaged.



Normal



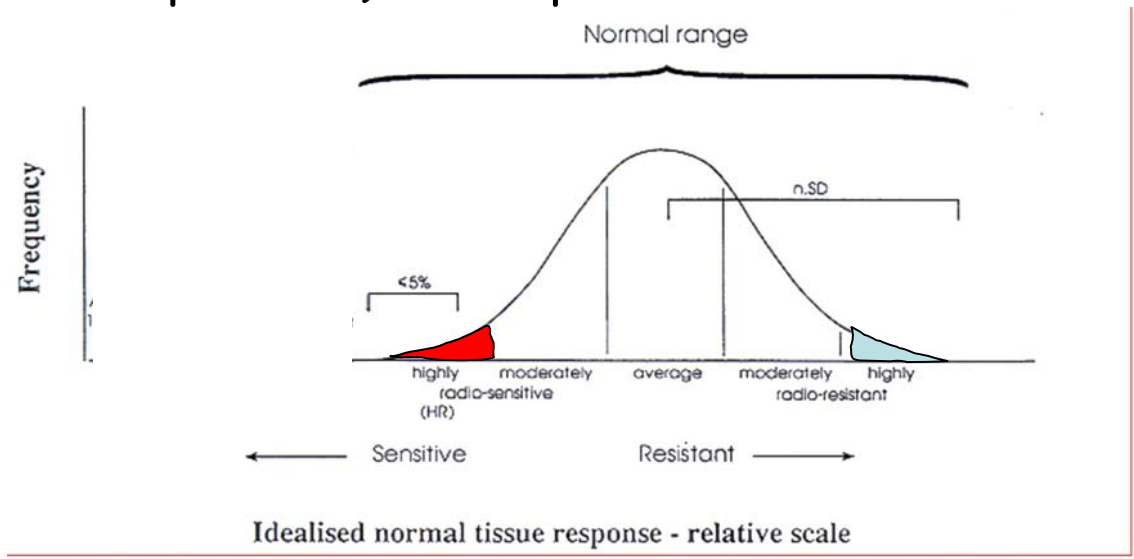
Radiation proctitis

From Vernia et al, Lancet 2001

Radiation toxicity

- Adverse radiation effects
  - acute - during treatment
  - late - after 3 months, chronic
- Although largely independent, a component of the late effects predicted by acute effects (consequential late effects, Heemsbergen et al, 2006)

- 5-10 % of treated patients (over-responders) develop severe late normal tissue response
  - Telangiectasia
  - Fibrosis
  - Proctitis



Burnet et al, Int. J. Cancer, 1998, 79, 606

Radiation toxicity

Hypothesis:

- Genetic susceptibility (Andreassen et al, 2005, Hoeller et al, 2003, Lee et al 2003)
- Subgroup of individuals more sensitive to radiation.
- Altered radiation response in normal tissue

Test:

- Determine radiation response by gene expression profiling
- *Ex vivo* irradiated and non-irradiated lymphocytes

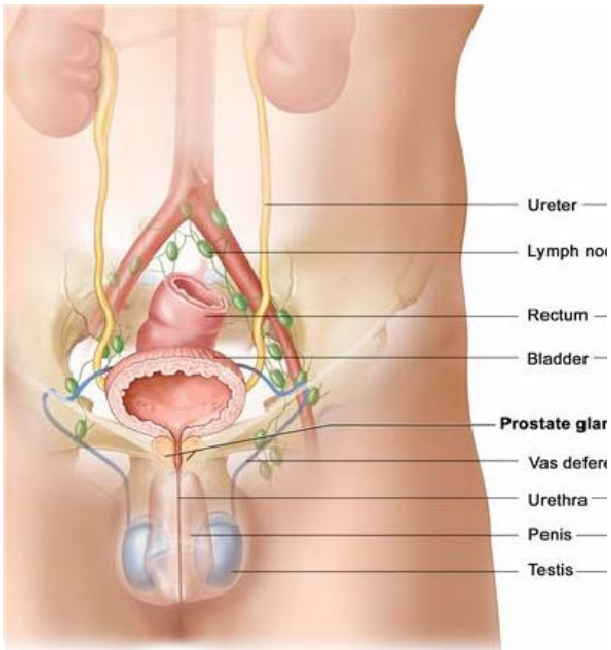
Goal:

- Identification of "high risk" individuals for radiation exposure
- Individual prognosis, tailor-made radiotherapy

Model: Prostate cancer patient group

- Large number of patients
  - Common tumour
  - Favorable prognosis
- Uniform treatment: External beam radiotherapy
- Homogenous group

- Large database (800 patients) for a retrospective study



From: www.nci.nih.gov

Patient group characteristics

- Patients with the most extreme cases of toxicity:
  - 27 over-responders (grade III, OR)
  - 23 non-responders (grade 0-I, NR)
- Groups matched for dose, volume, age
  - Dose-volume is a very predictive event for LRT (Peeters et al, 2005).

Table S1. Patient Characteristics of NRs and ORs

Variable	Subcategory	NRs (n = 17)				ORs (n = 21)				p-Value
		n	Range	Mean Value	95% C.I.	n	Range	Mean Value	95% C.I.	
Age		17	57–86	75.7	71.7–79.8	21	51–85	73.8	70.2–77.3	0.50 tt
PSA[a] before radiotherapy		17	5.3–70.0	21.0	13.0–28.9	21	2.0–43.8	14.4	9.2–19.6	0.12 tt
T-classification[b]	T1	2				5				0.35 χ^2
	T2	2				3				
	T3	11				13				
	T4	2				0				
Radiotherapy	Local only (70 Gy)	6				10				0.49 χ^2
	Locoregional (40 Gy + 30 Gy or 50 Gy + 20 Gy)	11				11				
Irradiated volume (local fields in cc)		17	414–1,231	749	622–878	20	489–1234	758	669–847	0.46 tt
Irradiated volume (pelvic fields in cc)		10	994–3,054	2,118	1,795–2,612	11	1,694–2,860	2,283	2,003–2,564	0.91 tt
Hormonal therapy	No	4				7				0.44 χ^2
	Yes	13				14				
Duration of follow-up (years)[c]		17	2.1–10.1	5.5	4.3–6.7	21	2.0–12.7	4.6	3.5–5.8	0.29 tt

[a]Prostate-specific antigen, an important biochemical marker for prostate cancer; normal value < 4 ng/ml.

[b]T-classification: clinical classification for primary tumor extension; T1, non-palpable; T2, palpable, within prostate; T3, extension beyond prostate; T4, invasion of adjacent organs.

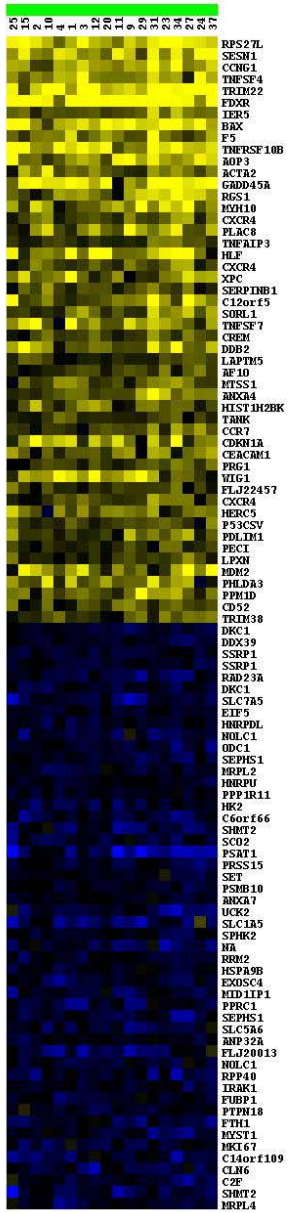
Experimental design

- Blood drawn from previous patients with no evidence of disease
- Stimulated peripheral lymphocytes irradiated (0 or 2 Gy X-rays)
- After 24 hours RNA isolated
- Gene expression profile using HG-U133A Affymetrix GeneChips (>22K transcripts)

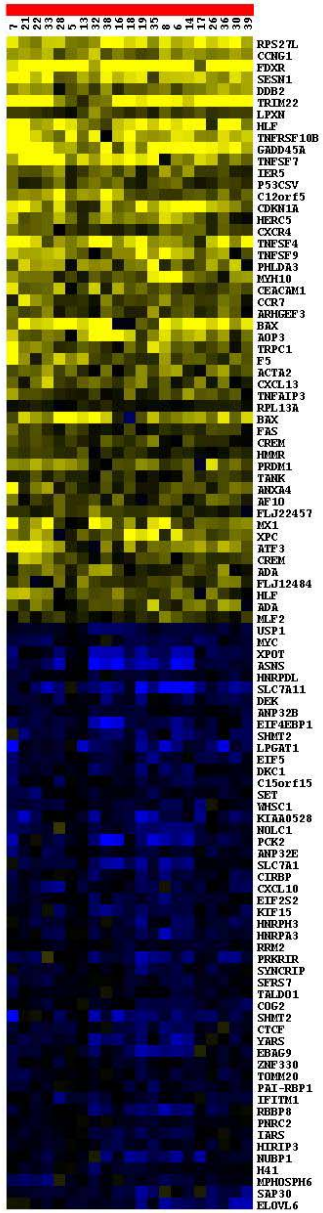
- Initial data preprocessing (Svensson et al, BMC Bioinformatics, 2005)
  - Data available at ArrayExpress (ETABM-90)
  - In R: all scripts validated and available at www.medgencentre.nl/pla/index.html

General radiation response

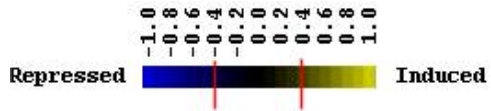
A NR



B OR

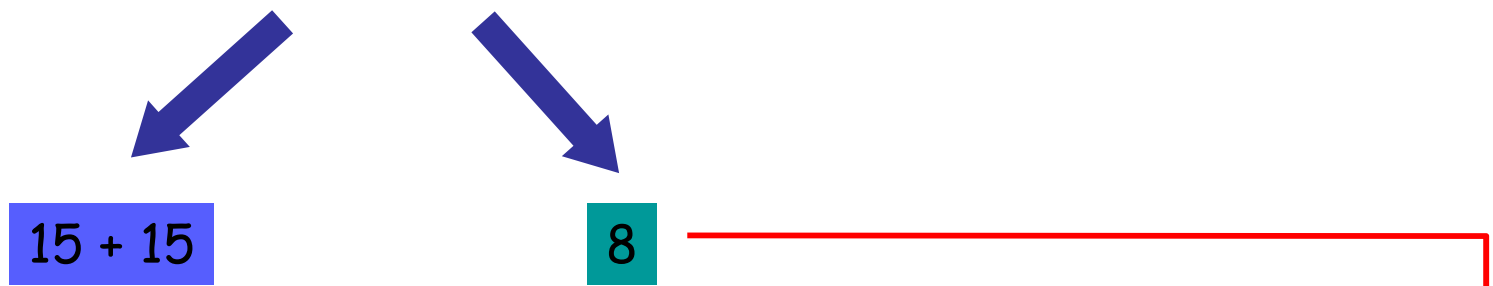


- Top 100 most significantly up- and down-regulated genes in OR and NR
- Contains characterized IR-responsive genes such as p21, DDB2, XP-C, Gadd45a…
- Heterogeneity not correlated to responder status



Classification by random cross-validation

38 samples (21 OR, 17 NR)

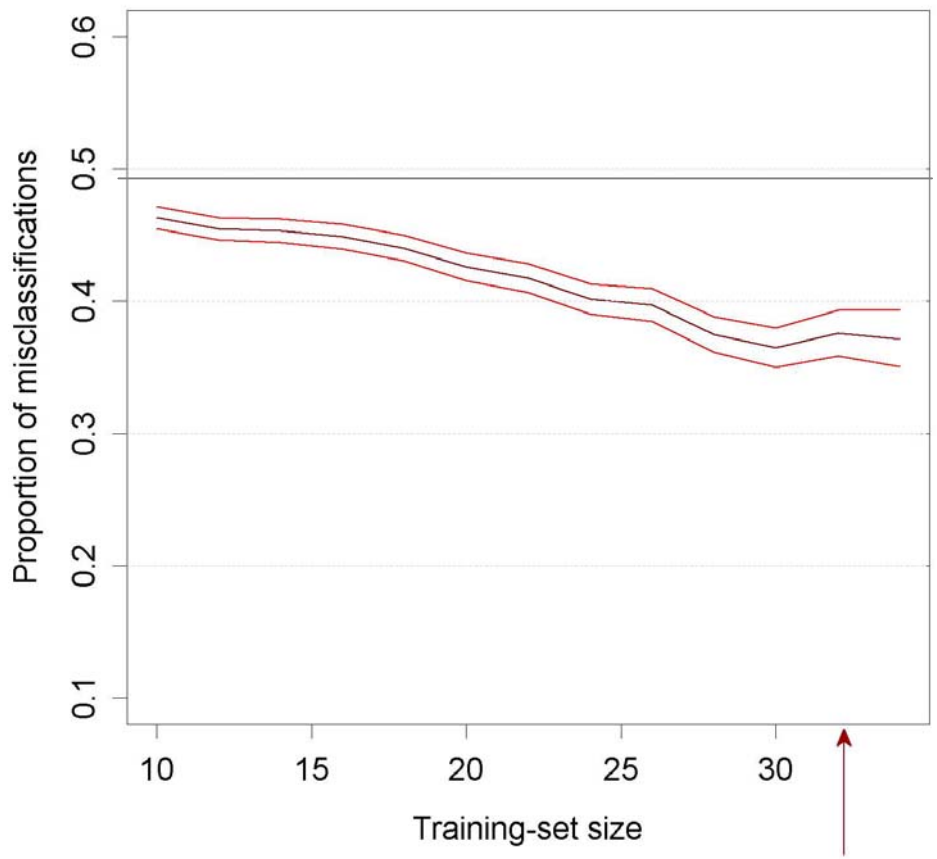


- ✓ Split patient group into a balanced training set and a remaining test set
- ✓ Define classifying signature as the 50 genes with the highest correlation between gene expression and responder status
- ✓ Test withheld samples for misclassification
- ✓ Repeat 500 times for each training set size
- ✓ Vary training set size from 10 – 34 patients

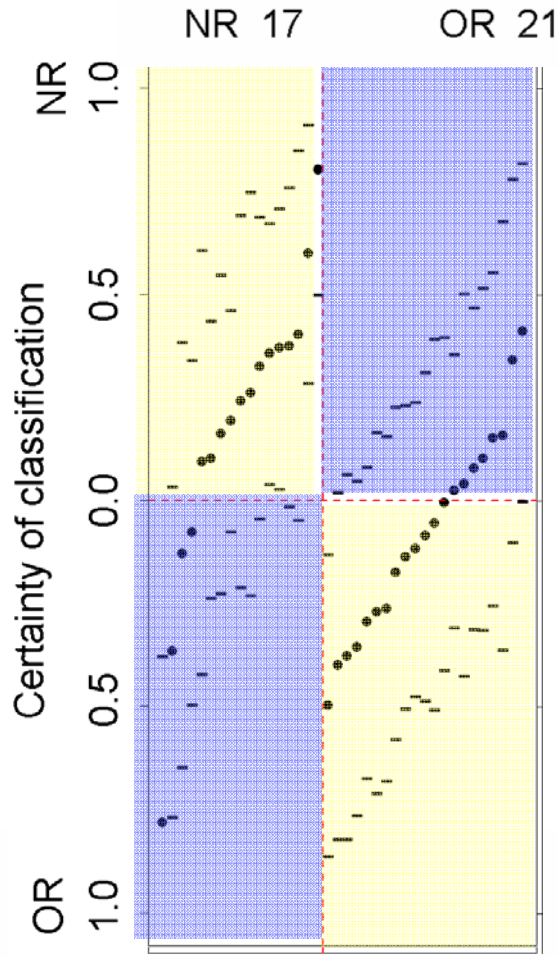
Michiels et al, Lancet, 2005

Classification of radiation response

- ~ 63% of the patients are correctly classified.
- Confidence interval wider for an independent study
- Population-based estimate



Certainty of classification for individual patients

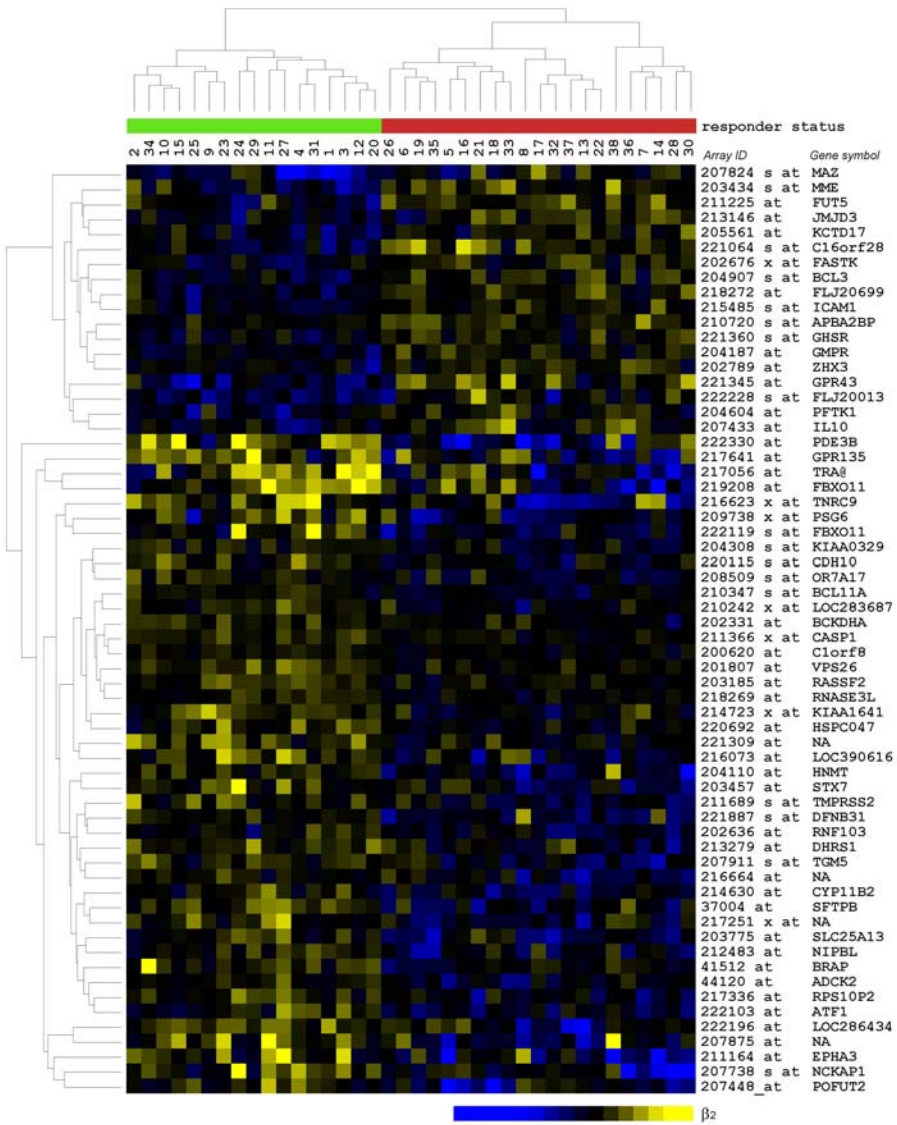


microarray			
NR	21	13 (4)	8 (0)
OR	17	4 (1)	13 (1)

- Based on difference in correlation between each patient and the NR and OR centroid at a training set size of 32 patients (16NR, 16 OR)
- Tolerance interval (95% confidence, 90% coverage) for each patient

- Classified with certainty if tolerance interval did not include zero

Classifying genes

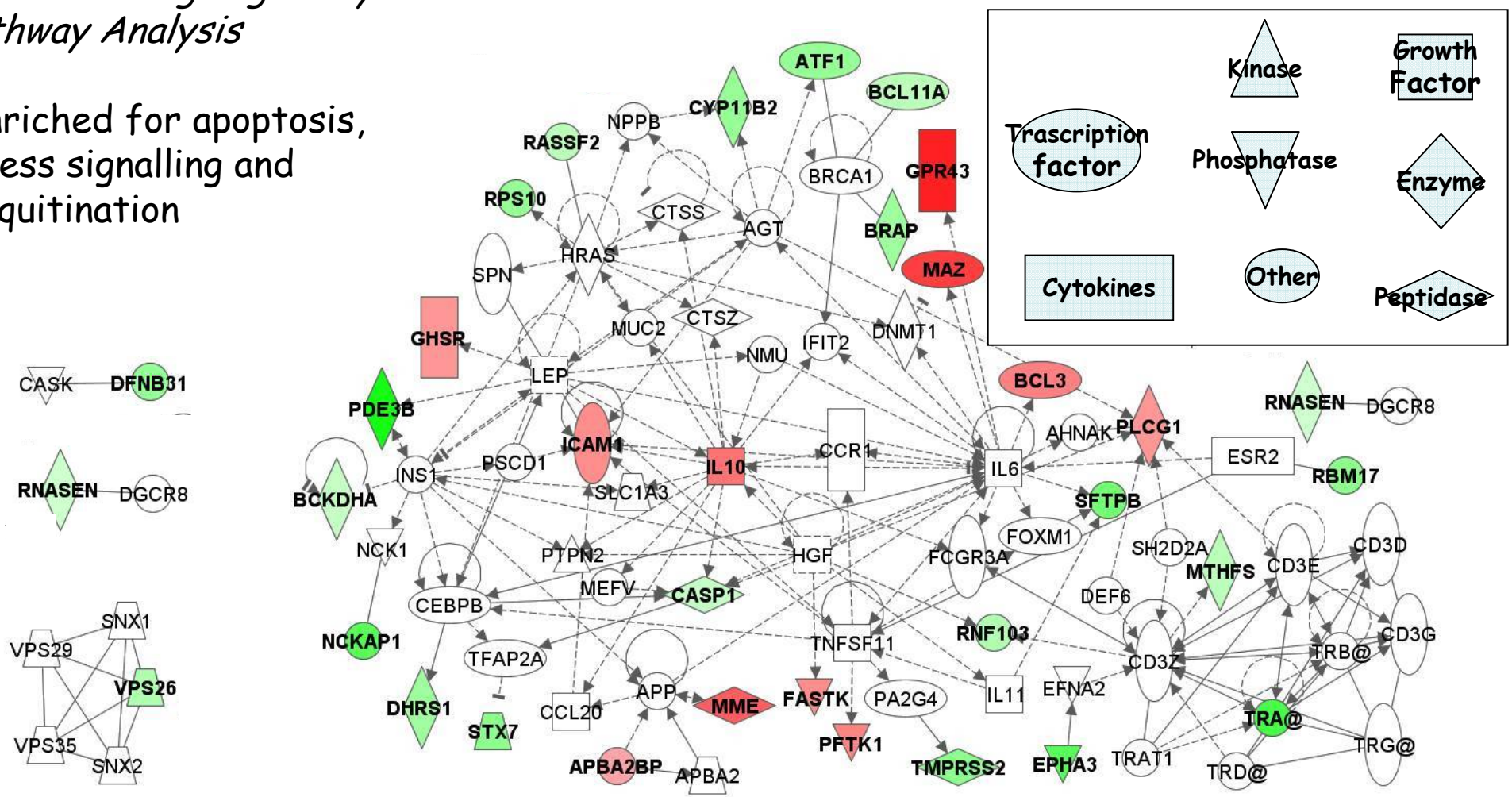


- 2D hierarchical clustering of the 62 genes present in >20% of the repeated assessments of the classifier
- No single gene separates the groups.
- List contains caspase, interleukins…

Classifying genes connected in literature based network

- *Created using Ingenuity Pathway Analysis*

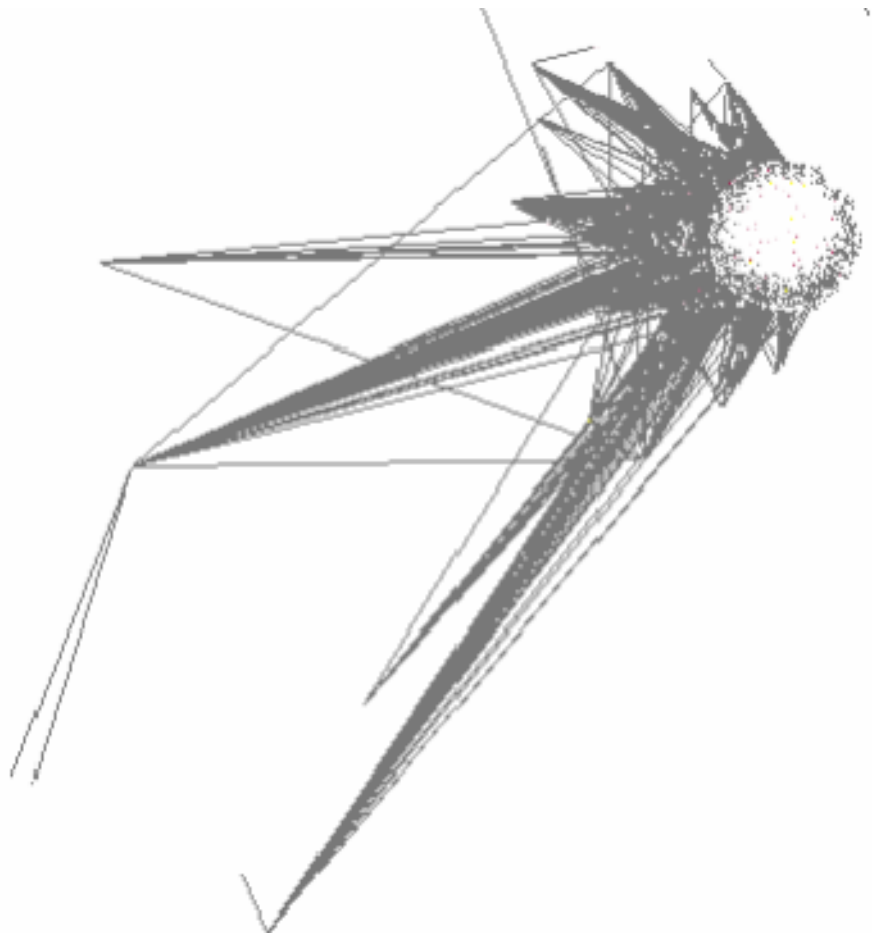
- Enriched for apoptosis, stress signalling and ubiquitination



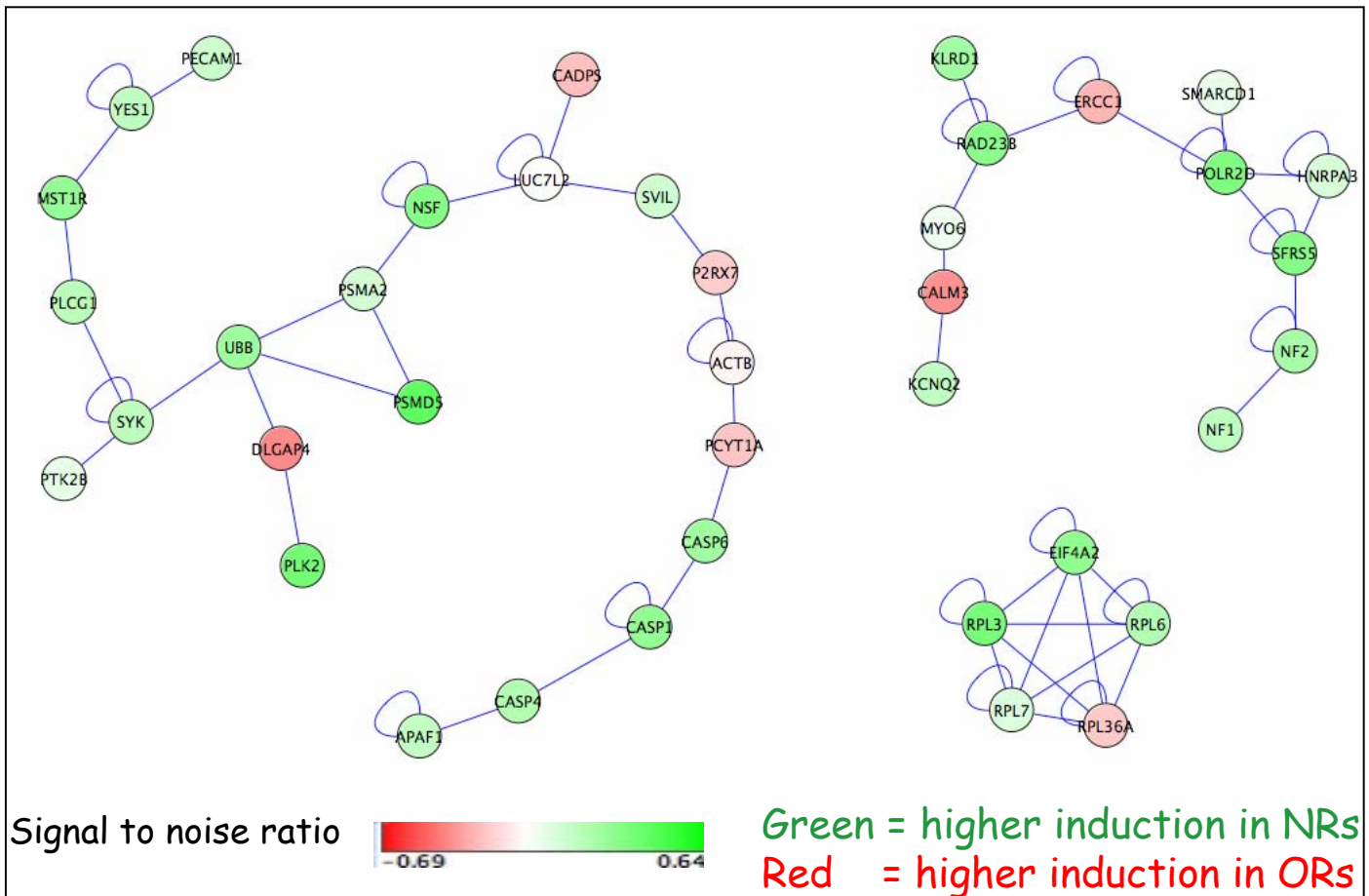
Green = higher induction in NRs
Red = higher induction in ORs

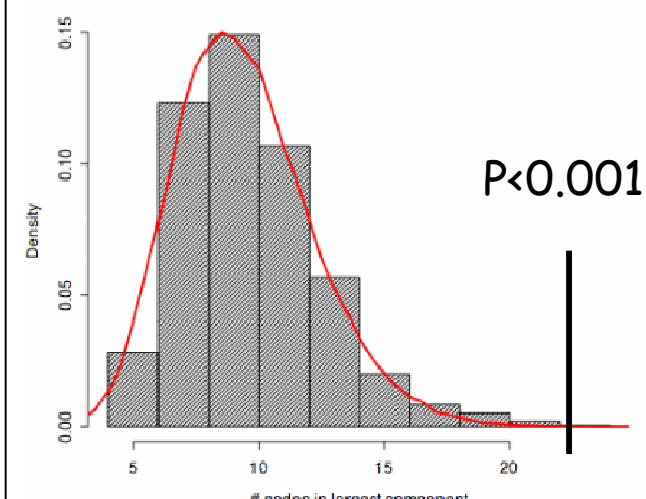
Experimental protein-protein interaction network

- High quality human protein-protein interaction network
  - experimental protein-protein interactions
  - 10,346 nodes, 53,538 edges
  - Sources: Stelzl et al (2005) and Rual et al (2005), Chaurasia et al, 2007
- Queried by Cytoscape (T Ideker, www.cytoscape.org)



Classifying genes connected protein-protein interaction network

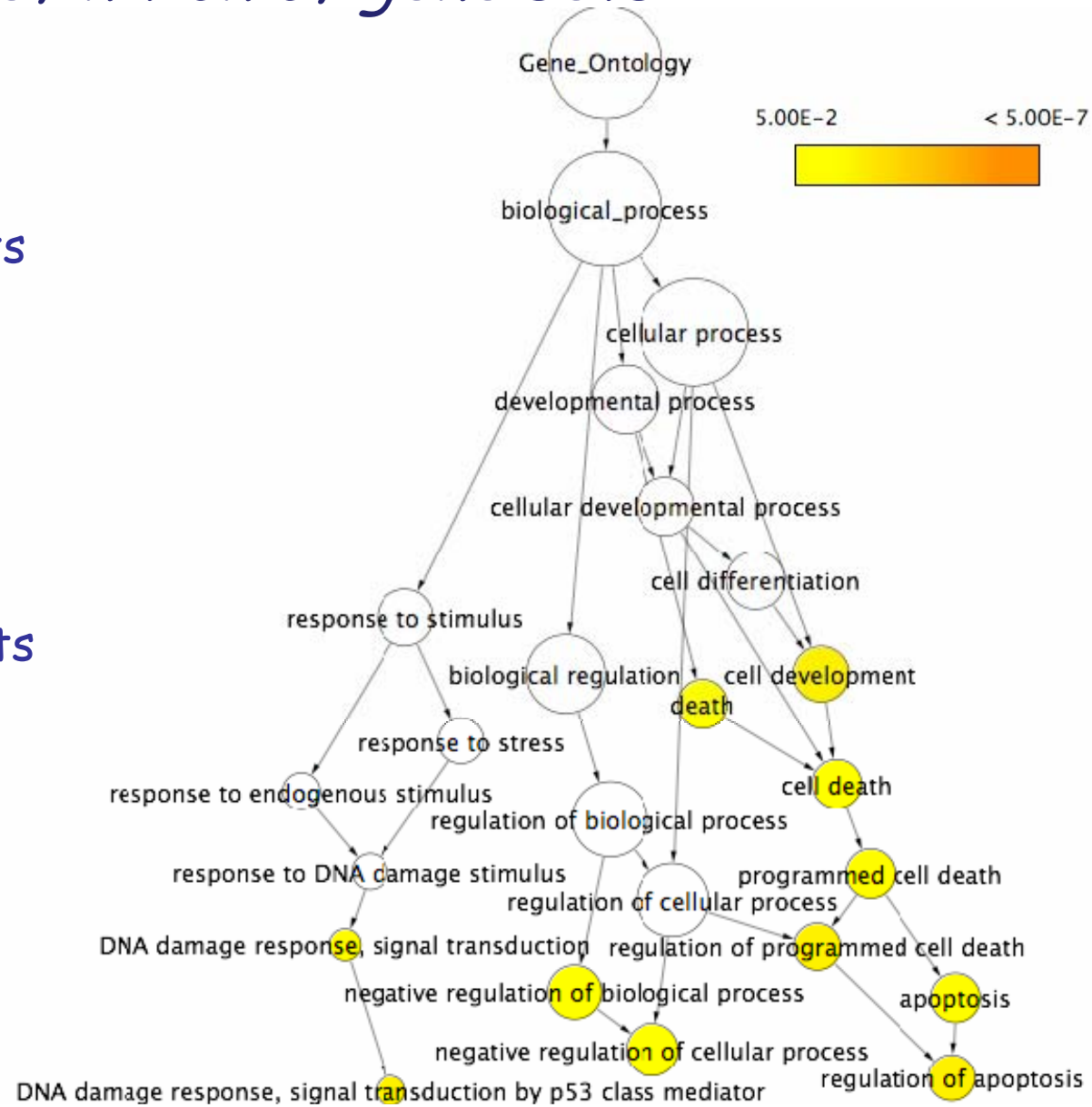




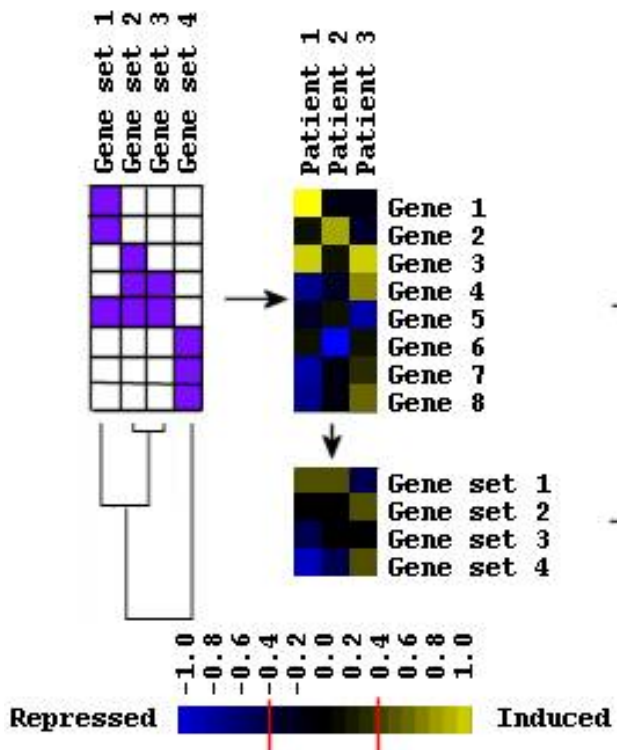
Definition of gene sets

- Genes grouped in gene sets based on functional and spatial relation (*Gene ontology consortium)*

- There are ~1.200 gene sets present on the array
  - >5 members
  - unique terms



The principle of gene set classification



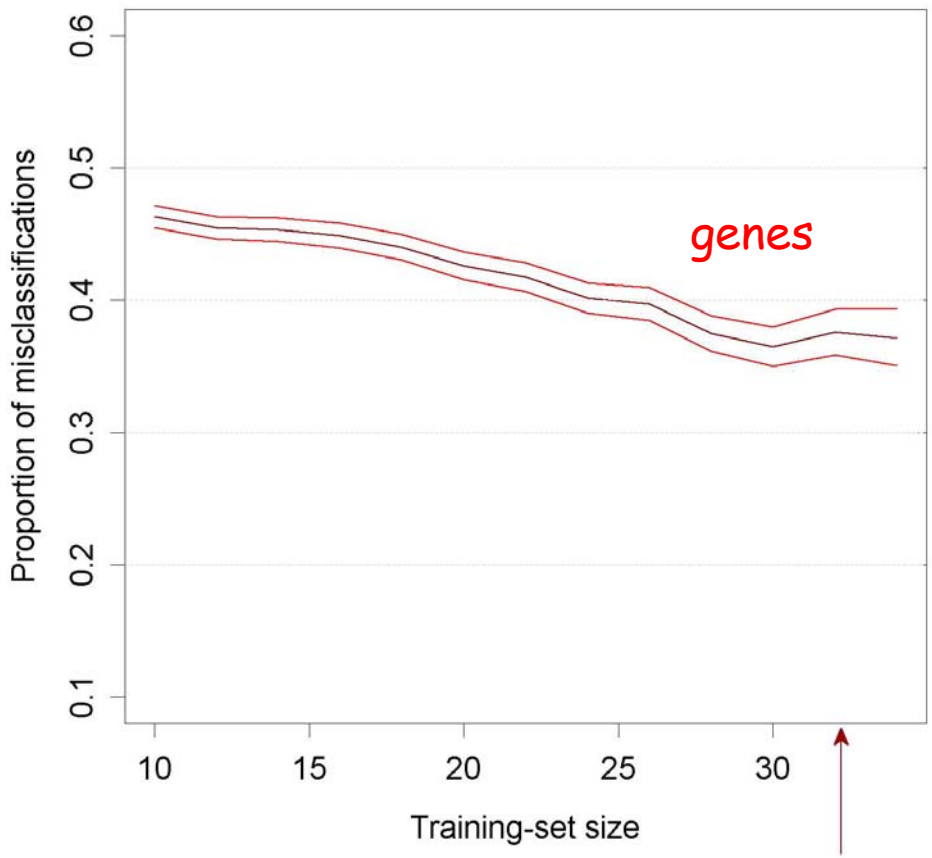
	Patient 1	Patient 2	Patient 3
Patient 1	1	0.07	-0.05
Patient 2		1	-0.21
Patient 3			1

Patient 1	1	0.95	-0.67
Patient 2		1	-0.85
Patient 3			1

$$r = \frac{\#up - \#down}{\#total}$$

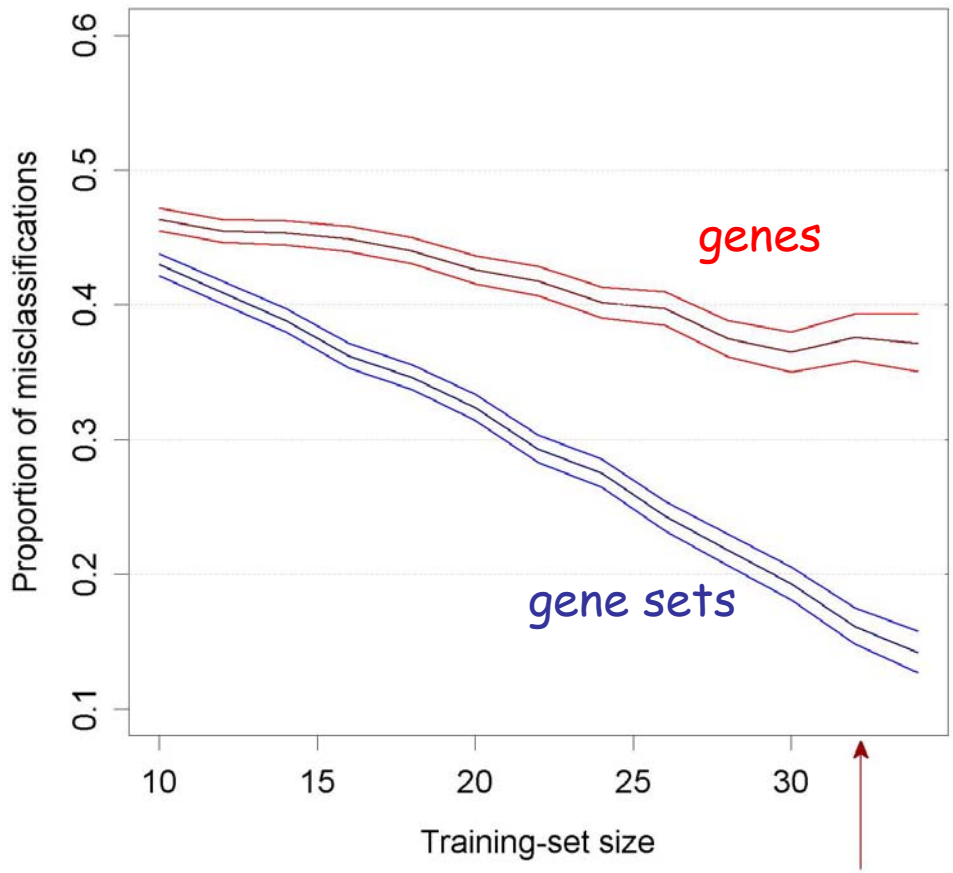
Functional classification results

- Random cross-validation



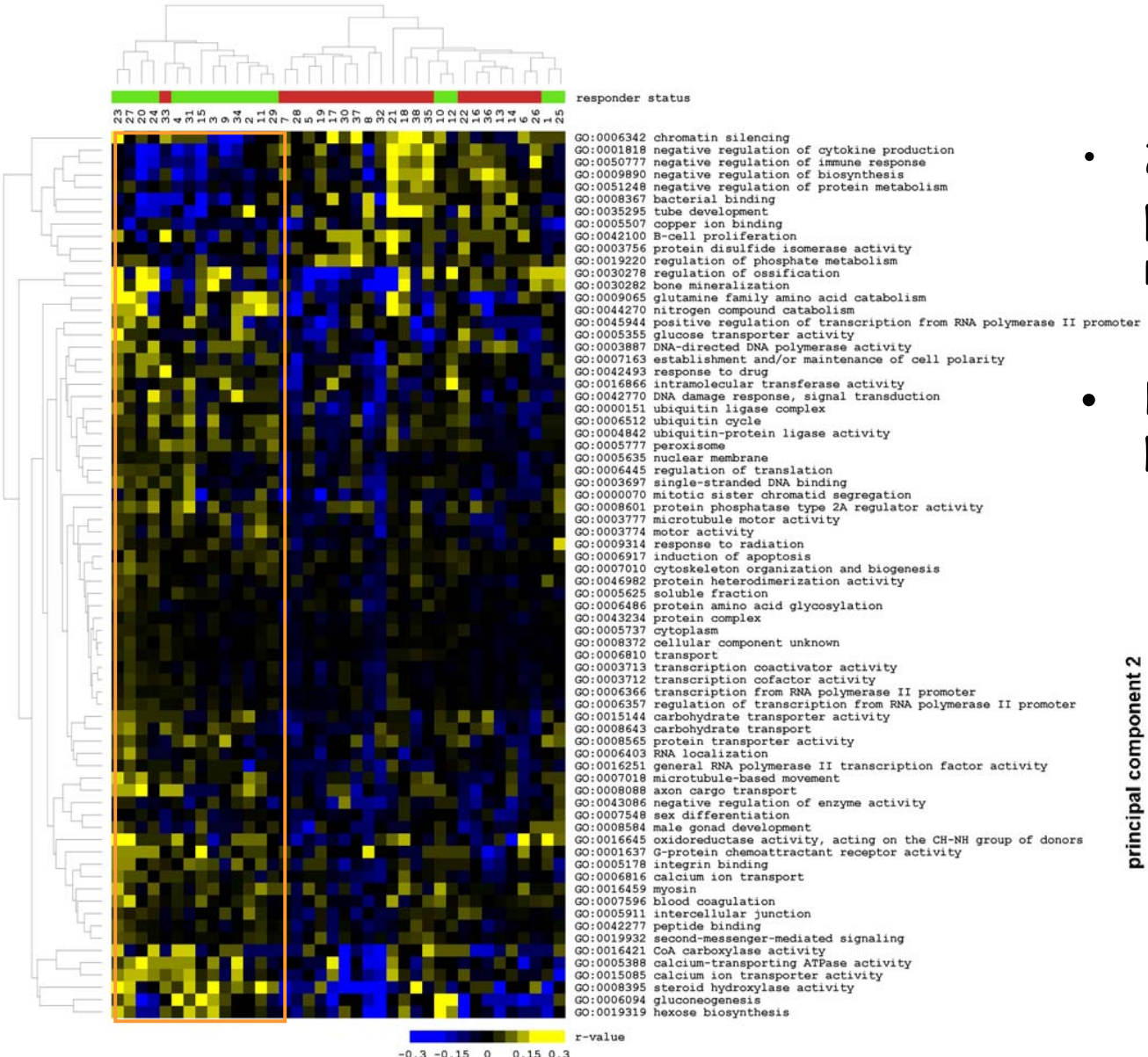
Functional classification results

- Random cross-validation
- Misclassification rate drops to 14% using Gene ontology gene sets

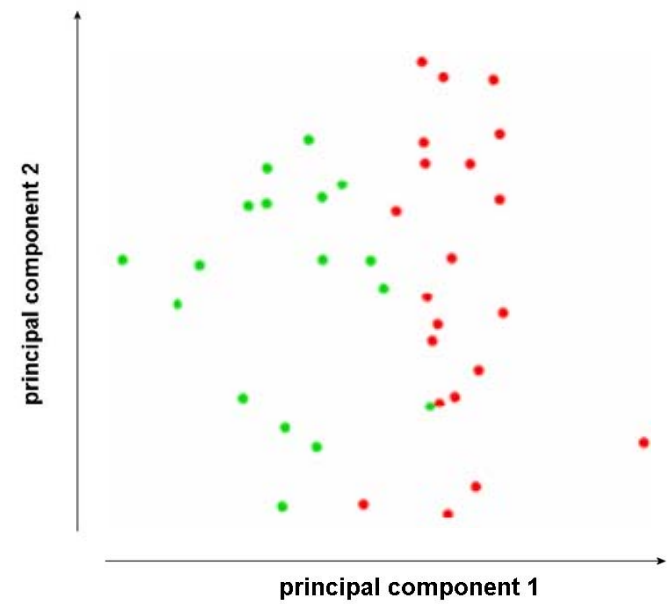


Svensson et al, 2006, PLoS Med

Classifying gene sets

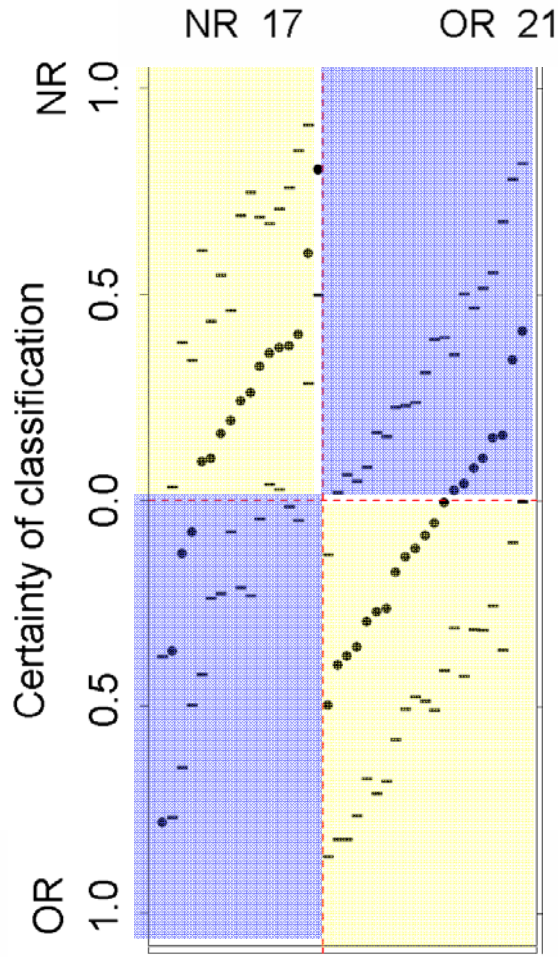


- 2D clustering based on the proportions of induced and repressed genes in gene sets
- Death more pronounced in NR (Barber et al, 2000)



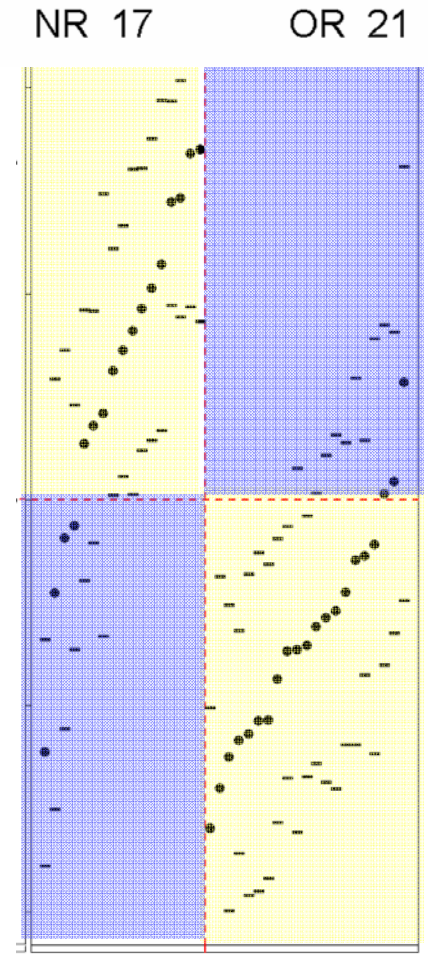
Certainty of classification for individual patients

genes



microarray	NR 17	OR 21
NR 21	13 (4)	8 (0)
OR 17	4 (1)	13 (1)

gene sets



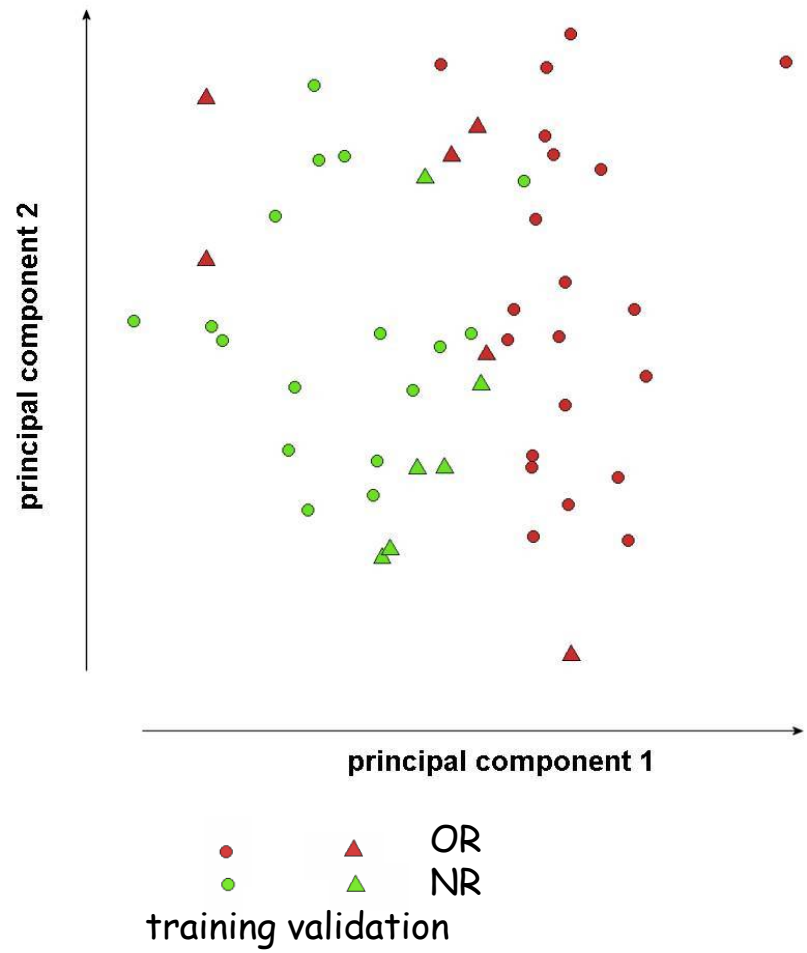
	NR 17	OR 21
NR 16	13 (10)	3 (0)
OR 22	4 (1)	18 (10)

Gene set classification of 12 new patients

A

microarray	physician: NR 6	physician: OR 6
NR 8	5 (3)	3 (2)
OR 4	1 (0)	3 (2)

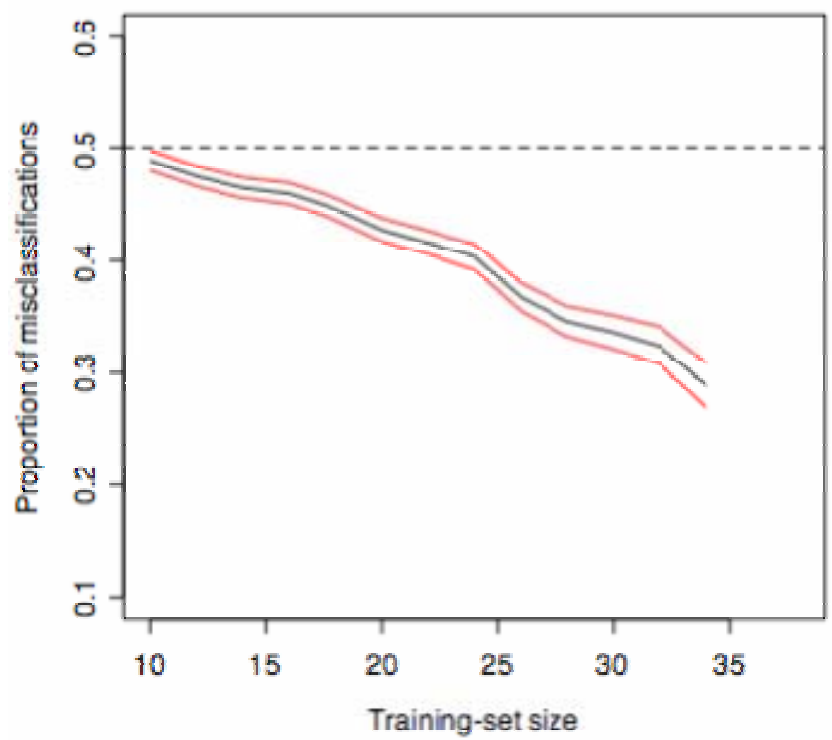
B



As time goes by…

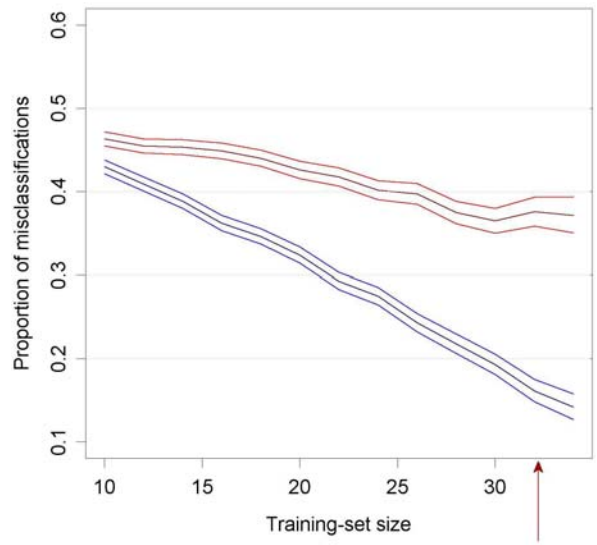
- As databases evolve, data are changed.
  - The last two years, GO gene sets changed on average 30%
  - Parameters needs adjusting.

- One way to get more objective gene sets is to look at protein interaction networks.
  - Results pending…

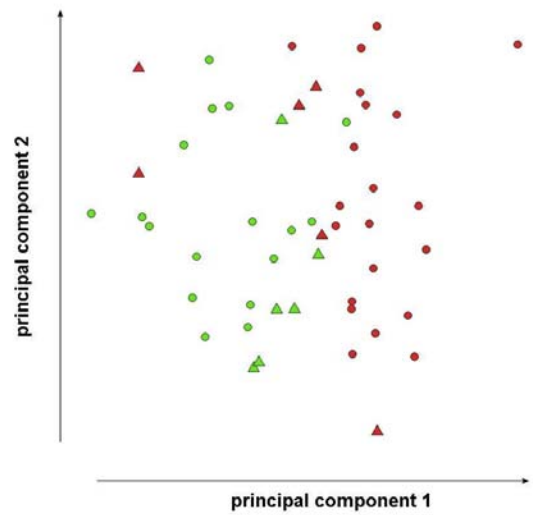


Conclusions and remarks

- Genes associated with late radiation toxicity.
- Genes that classify NR and OR are involved in e.g. apoptosis and ubiquitination, stress signalling
- Grouping of genes into functional modules drastically improved prediction results
- A subset of patients could be identified which could be predicted with certainty.



- Genetic effects are only part of the reason to develop late toxicity.
- Prospective study under way.



Acknowledgements

Leiden, The Netherlands
Toxicogenetics

Rebecca Esveldt-van Lange
Binie Klein
Dr. Harry Vrieling
Dr. Micheline Giphart-Gassler

Amsterdam, The Netherlands
Radiotherapy

Dr. Lukas Stalpers
Dr. Klaas Franken
Dr. Jaap Haveman

Bioinformatics
Dr. Perry Moerland

Uppsala, Sweden
Oncology

Prof. Ingela Turesson

MIT, US
Center for Environmental Health Studies

Prof. Leona Samson
Dr. Rebecca Fry

Funding by EU Risc-Rad, Swedish Cancer Society, CMSB Leiden, Swedish Research Council