Statistics Applied to Bioinformatics

Study cases Gene expression data

Jacques van Helden

Jacques.van-Helden@univ-amu.fr
Aix-Marseille Université (AMU), France
Technological Advances for Genomics and Clinics
(TAGC, INSERM Unit U1090)
http://jacques.van-helden.perso.luminy.univmed.fr/

Two-channels microarray technology

Cell culture, tissue, ...

Sample 1

RNA

cDNA

Sample 2

RNA

cDNA

RNA extraction

Synthesis of fluorescent cDNA

Brightness ↔ Quantity Color ↔ Specificity

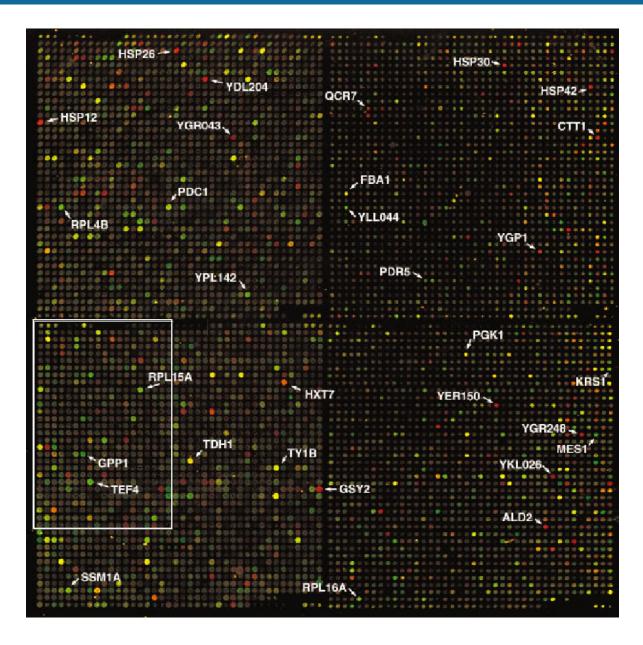
yellowish reddish greenish

not specific sample 1 - specific

sample 2 - specific

DNA chip

Complete microarray – two-channels, full cDNA, yeast



Two-channel microarrays - Raw intensity measurement

- Raw measurements
 - Red intensity
 - Red background
 - Green intensity
 - Green background
- *Intensity background* = level of expression
 - □ *Red* in experimental conditions
 - Green in control

Two-channel microarrays : log-ratio computation

The level of regulation is represented by the ratio

$$r = \frac{red - red.bg}{green - green.bg} \qquad r > 1 \quad \Rightarrow up\text{-}regulated$$

$$r < 1 \quad \Rightarrow down\text{-}regulated$$

- The log-ratio provides a more convenient statistic (we will see why during the course)
- log₂ is even more convenient because the scale is more intuitive than natural logarithms.

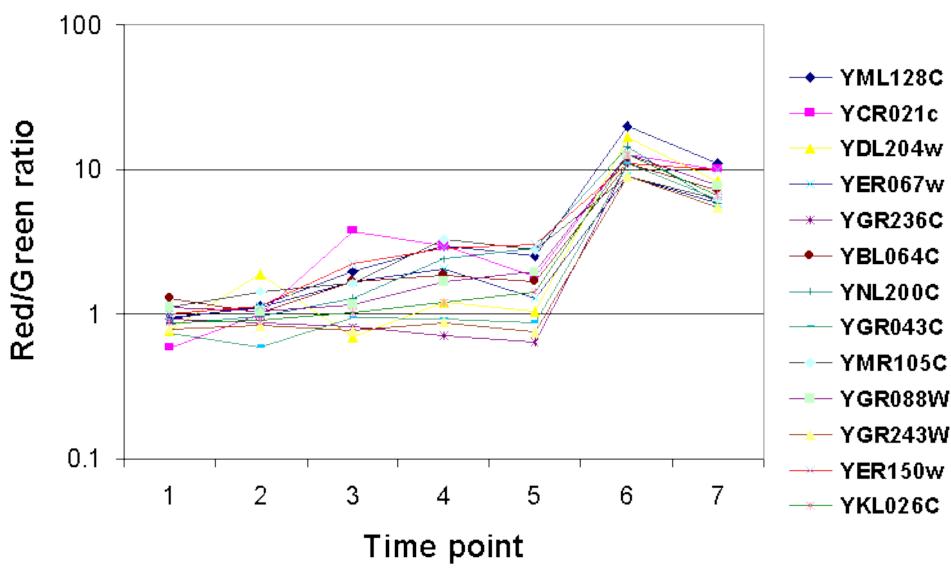
$$R = \log_2 \left(\frac{red - red.bg}{green - green.bg} \right) \qquad R < 0 \quad \Rightarrow down-regulated \\ R > 0 \quad \Rightarrow up\text{-}regulated$$

$$|R| > 1$$
 \rightarrow regulated by a factor of 2
 $|R| > 2$ \rightarrow regulated by a factor of 4
 $|R| > w$ \rightarrow regulated by a factor of 2^w

Examples of experimental conditions

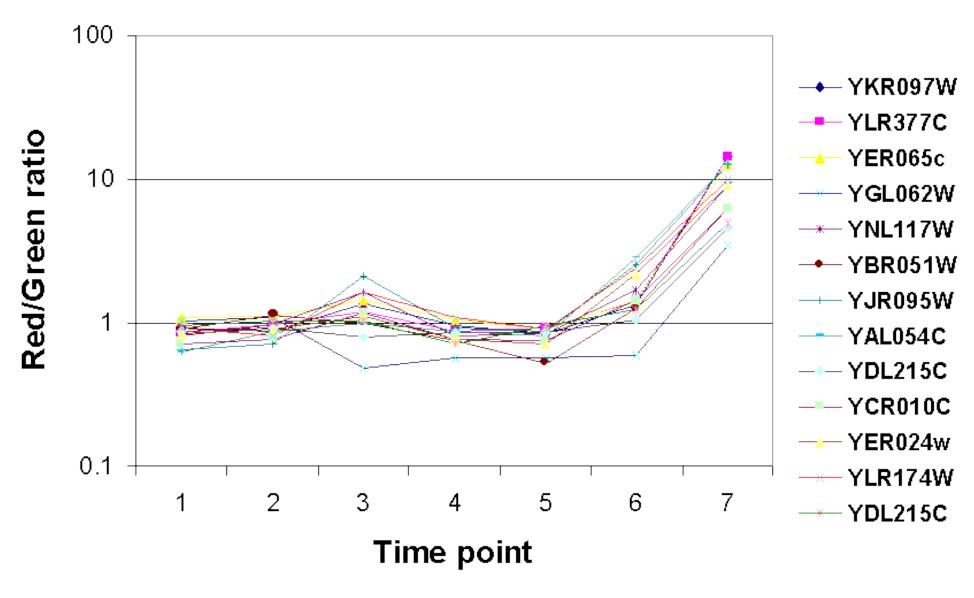
- Presence/absence of a metabolite
 - gal vs glucose
- Transcription factor mutants
 - Yap1p overexpression
 - TUP1 deletion
- Massive environmental changes
 - rich vs minimal medium
 - diauxic shift (7 time points during the shift)
- Cell differentiation
 - sporulation
 - mating type
- Cell cycle

Late activation during diauxic shift



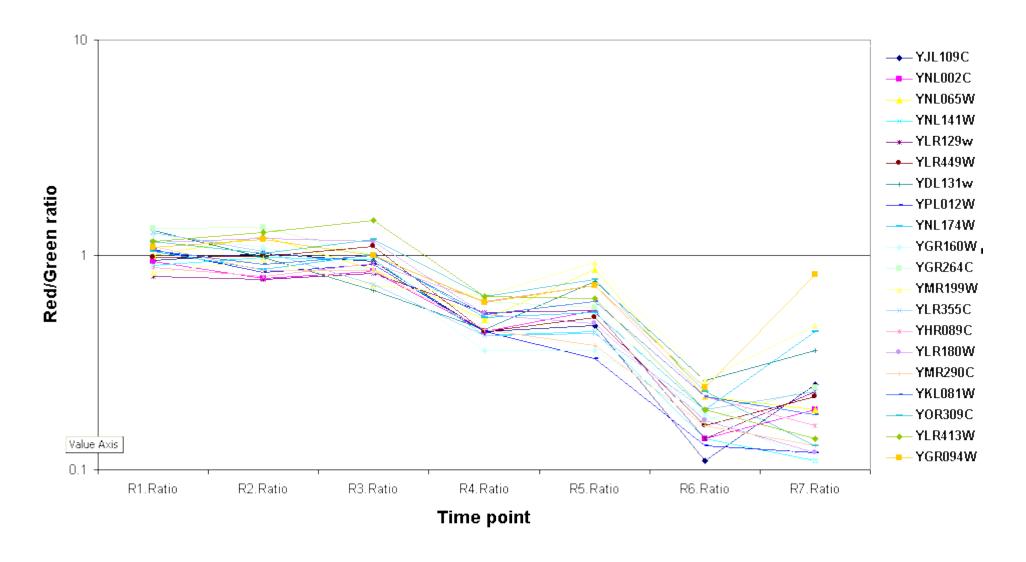
Data source: DeRisi JL, Iyer VR, Brown PO. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. Science 278(5338): 680-686.

Very late activation during diauxic shift



Data source: DeRisi JL, Iyer VR, Brown PO. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. Science 278(5338): 680-686.

Repression during diauxic shift



Data source: DeRisi JL, Iyer VR, Brown PO. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. Science 278(5338): 680-686.