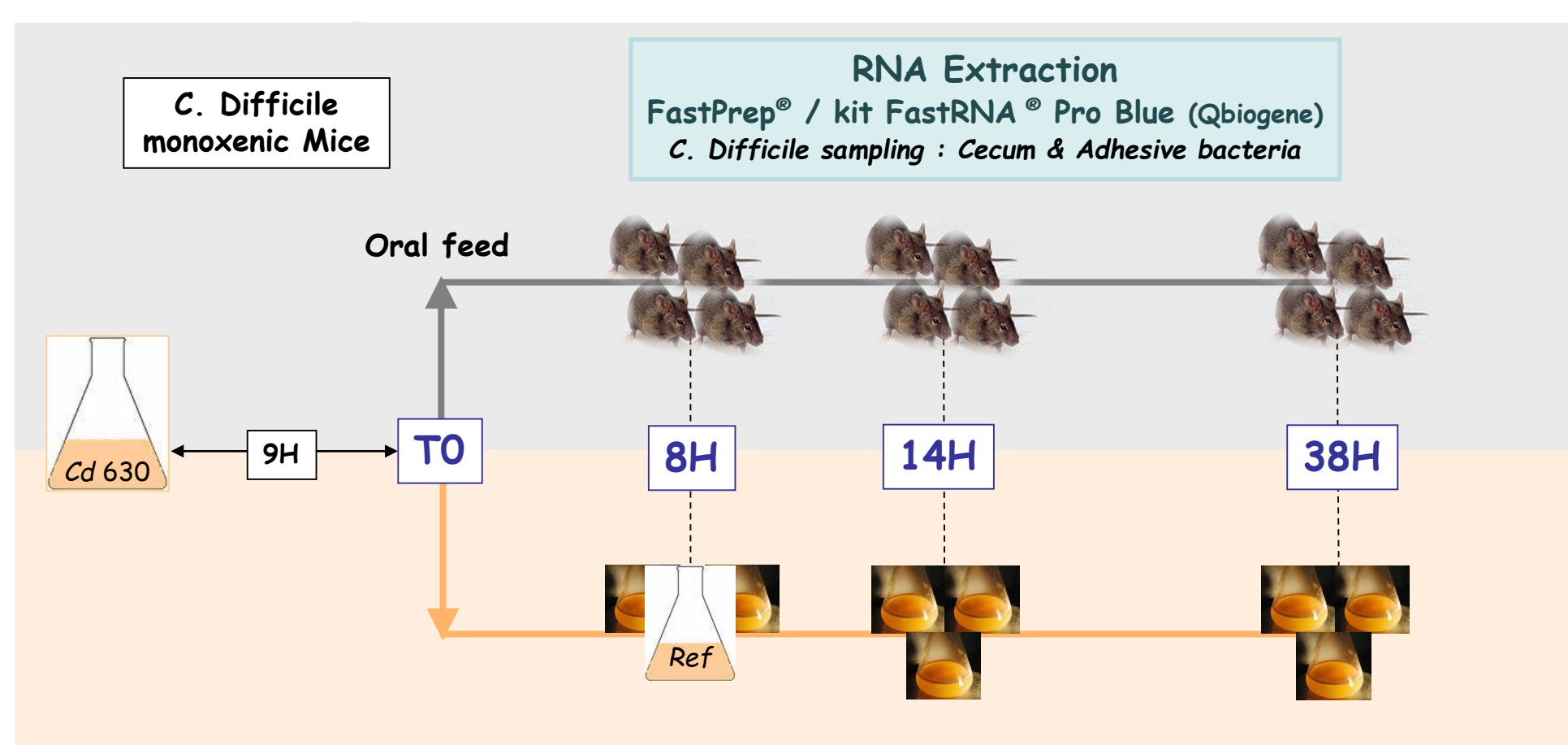
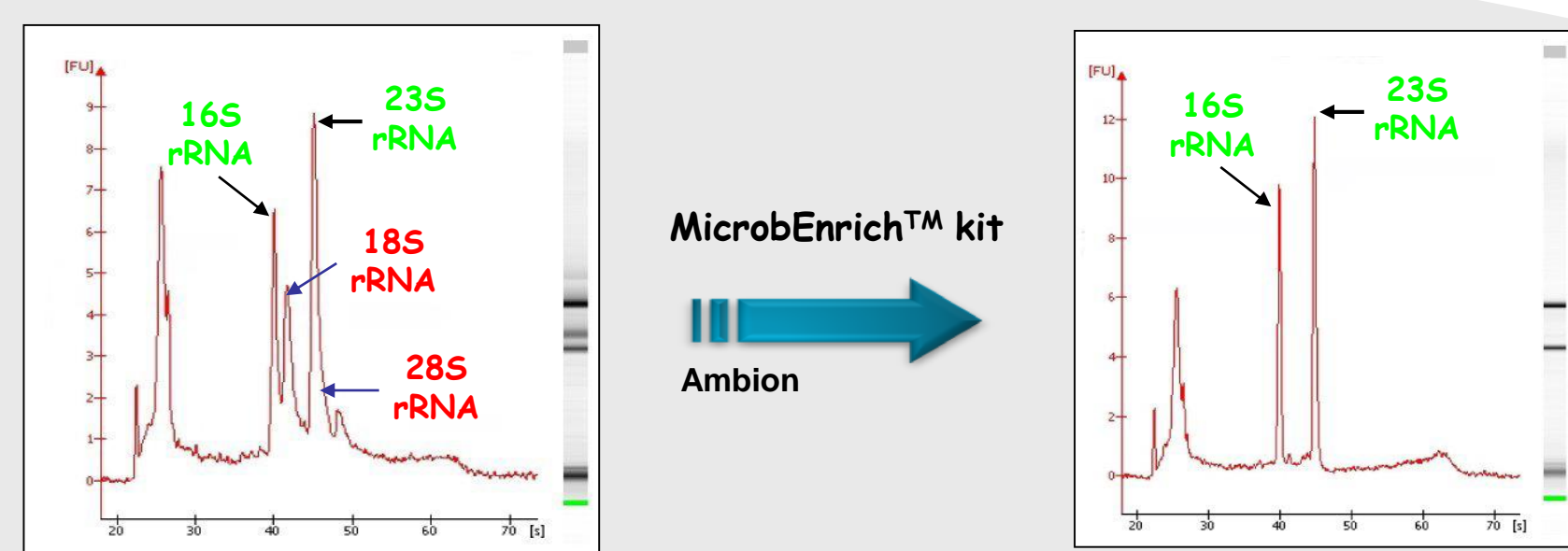


Experimental Design

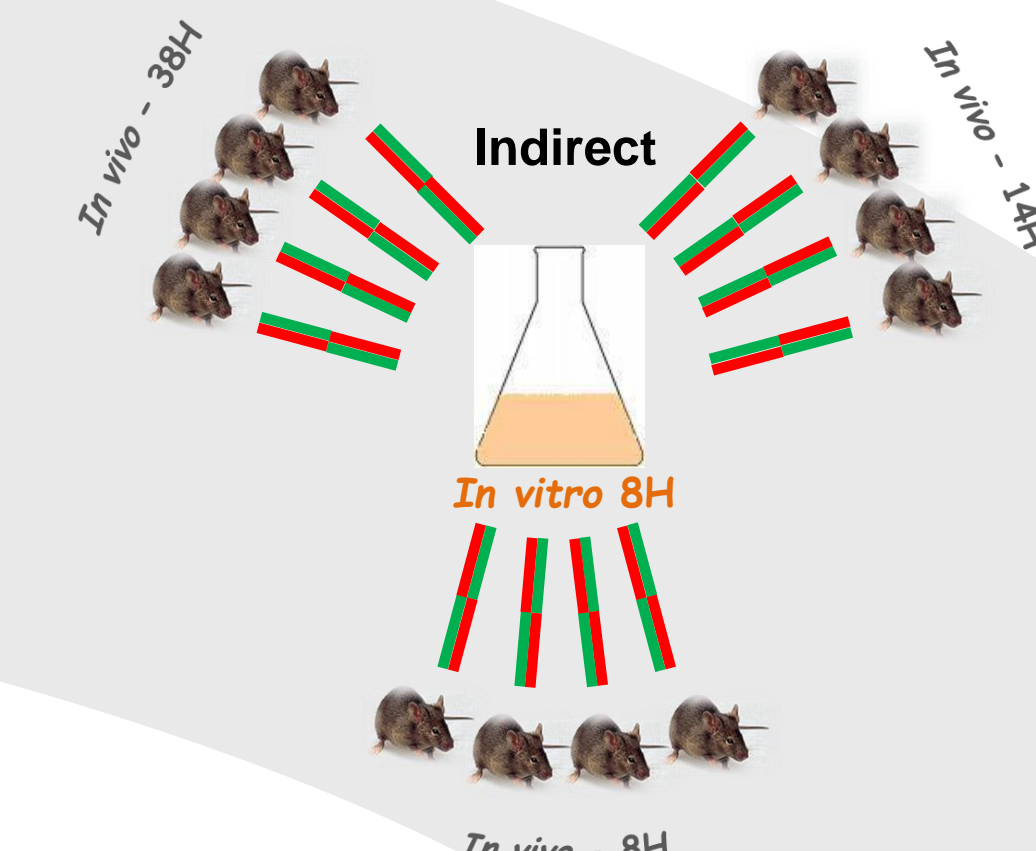


RNA preparation

- VIVO : Purification of Bacterial RNA
- VITRO : Quality Test BioAnalyser



Manipulation Design



Institut Pasteur

Unité des Toxines et Pathogénie Bactérienne

Global Gene Expression Analysis during *C. difficile* Infection

M. Monot¹, C. Denève², R. Tomé¹, C. Bourseaux-Eude³, L. Caleechurn¹, S. Rousseau³, I. Mozer³, C. Medigue⁴, D. Vallenet⁴, A. Collignon², C. Janoir² and B. Dupuy¹.

To understand the initial steps of the *C. difficile* colonization process and the relation with toxins production, we compared transcriptional profiles *in vivo* and *in vitro*. Axenic mice were challenged with *C. difficile* strain 630 and were sacrificed at 8, 14 and 38 hours post-infection. The same kinetics was done *in vitro*, followed by microarray manipulation.

Statistical analysis of generated data were performed. In parallel, to improve the relevance of our result we are performing a whole genome re-annotation using the MaGe platform. A graphical analysis interface usable within collaborative work, MA2HTML, has been created to decipher all the data obtained from the global gene expression study.

In summarize, 550 and 874 genes were found to be differentially expressed *in vivo* and *in vitro* respectively. More than 200 genes are specifically regulated *in vivo* and manually distributed in 18 classes and 250 functional groups. Among them three classes temporally expressed *in vivo* are particularly represented, i.e. sporulation (51), metabolism (99) and membrane transport (75).

Some gene regulated *in vivo* are currently inactivated and evaluated for their role of pathogenicity process of *C. difficile*.

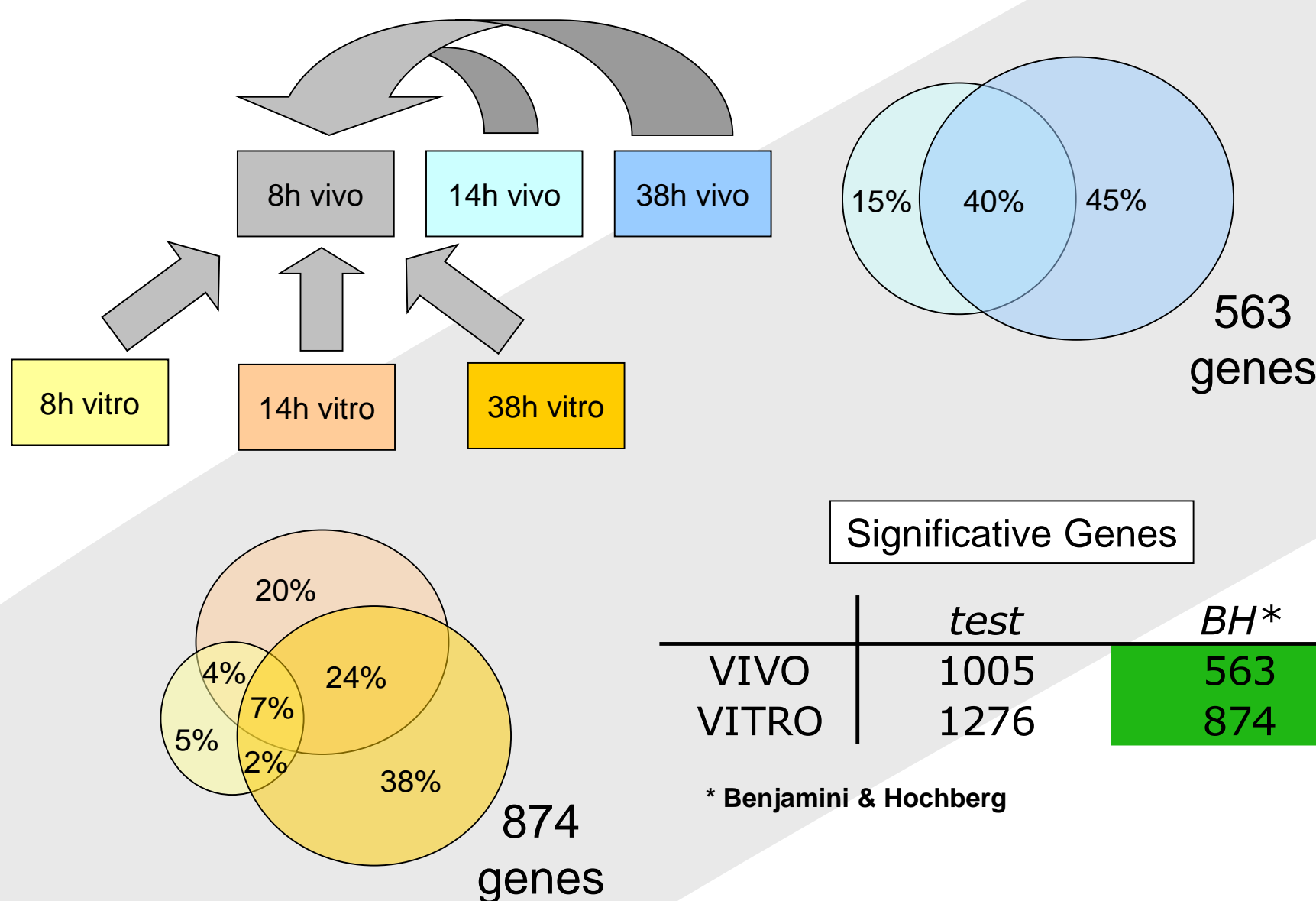


Smyth, G. K. et al. 2005

Statistical Analysis

Empirical Bayes Statistics
Benjamini & Hochberg Correction

Smyth, G. K. et al. 2004



Significant Genes

	test	BH*
VIVO	1005	563
VITRO	1276	874

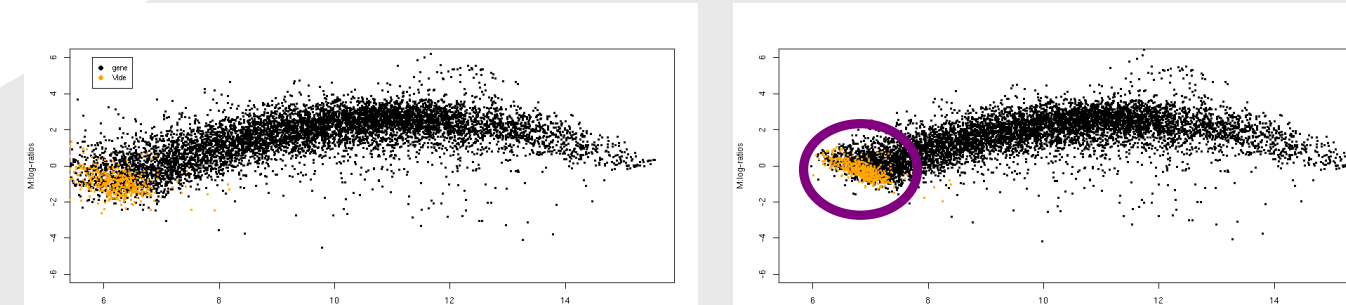
* Benjamini & Hochberg

Data normalization

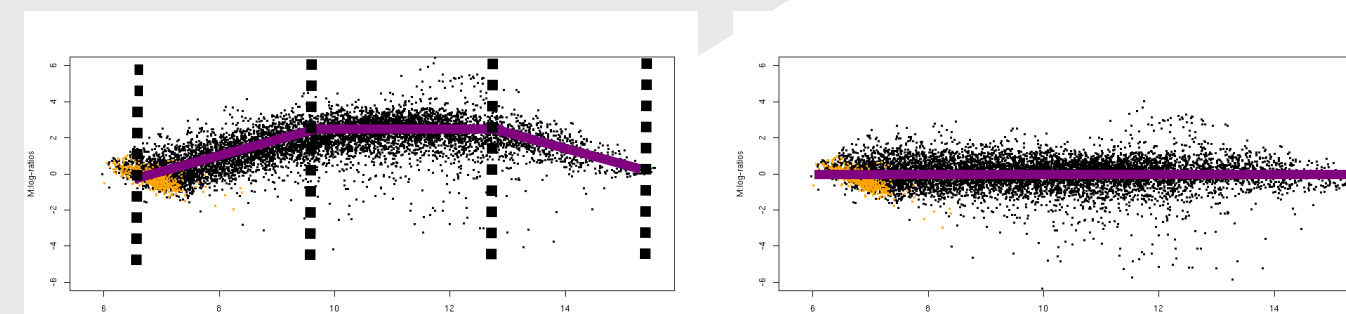
Background normalization¹
« loess » normalization²

Ritchie, M. E. et al. 2004

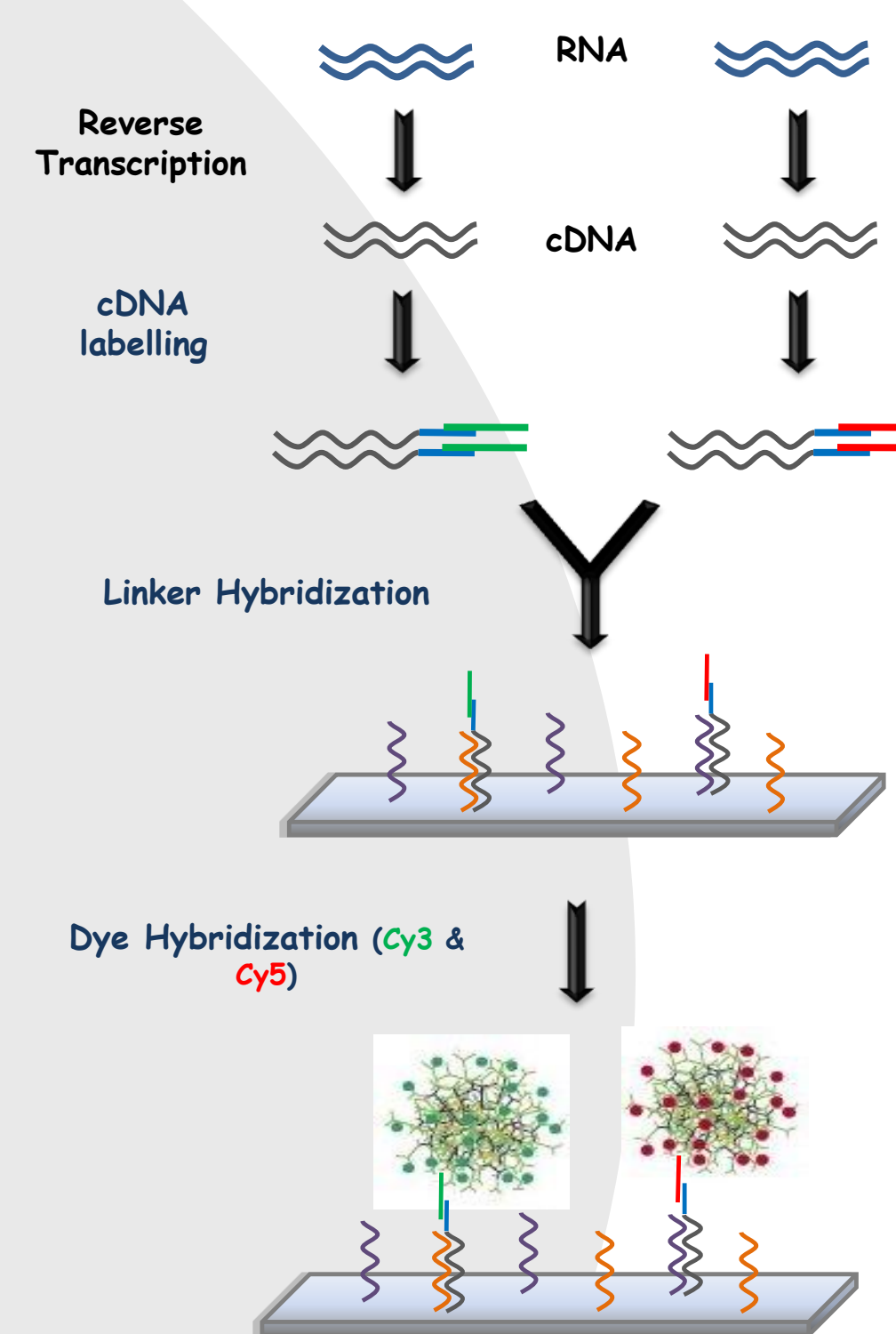
(1) Correct background to obtain positive value only



(2) Local linear regression to adjust scatter plot



MicroArray Protocols



PCR MicroArray - 3688 CDS

Stabler et al. 2006



Vallenet, D. 2006

Genome Reannotation

Improvement of MicroArray Analysis
Reconstruction of metabolic pathways

Summary	Known	Putative	Unknown
<i>C. difficile</i> annotation 2006*	34%	49%	17%
Reannotation 2009 (>50%)	42%	39%	19%

* Sebahia, M et al. 2006

Reannotation 2009	Known	Putative	Unknown
Known (775)	73%	24%	3%
Putative (729)	4%	78%	18%
Unknown (360)	0%	30%	70%
Total (1864 >50%)			

Analysis Result of *in vivo* kinetic

Up-regulated
Down-regulated

550 genes differentially expressed
63% up-regulated, 37% down-regulated
A majority are Reannotated

	All Fold change (FC > 2)	Manually Reannotated (%)	Manually Analysed (%)
Differentially expressed genes	550 (280)	431 (78%)	550 (100%)
Over expressed genes	349 (220)	300 (85%)	349 (100%)
Under expressed genes	201 (60)	131 (65%)	201 (100%)



Analysis web interface

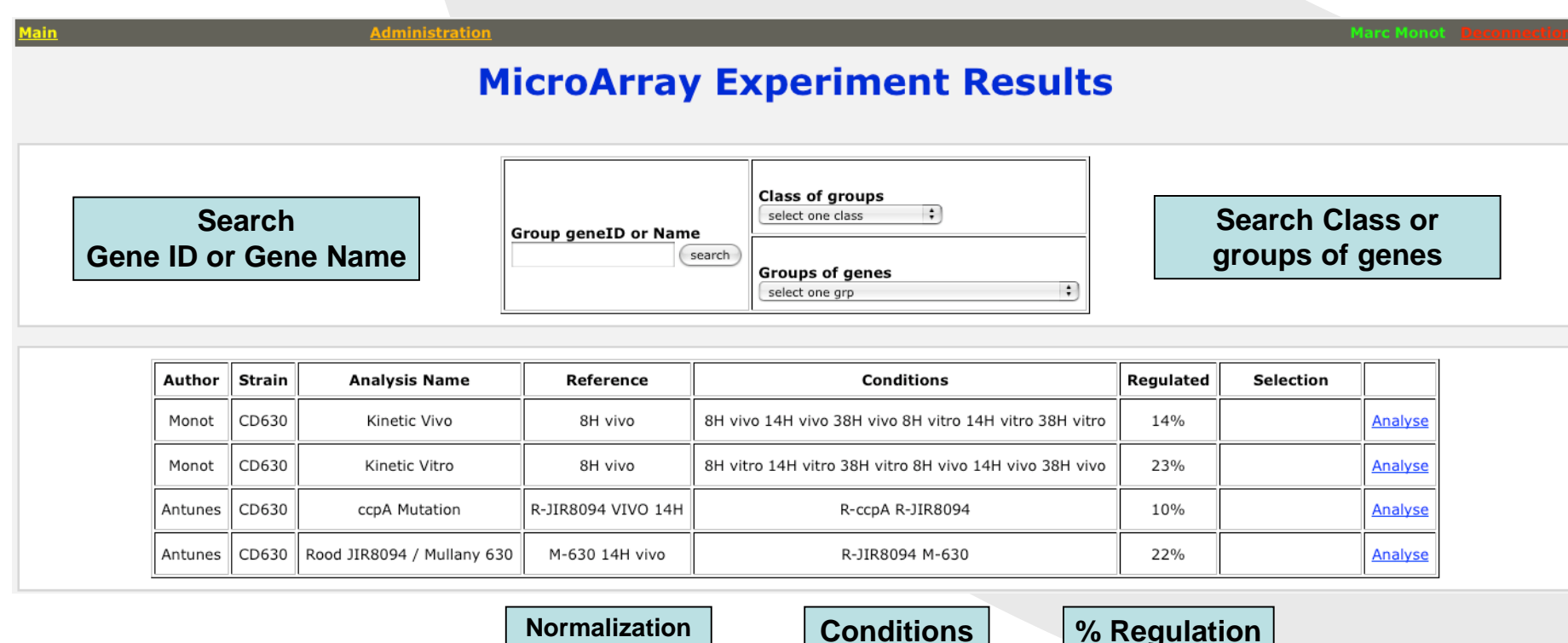
Reduce the amount of Data

- View only Significant results
- Groups genes (simple functional classification, Class & Grp)
 - Operon cluster*
 - Expert cluster

Shared Database

- Collaborative Work
- Link with reannotation project

* Dam, P. 2006



Mar's selected list of genes

- Cell Factor (6)
- Cell Wall (32)
- Fermentation (7)
- Membrane Transport (28)
- Metabolism Amino Acid (20)
- Metabolism Carbon (11)
- Metabolism Cofactor (5)
- Metabolism Fatty Acid (3)
- Metabolism Nucleic Acid (7)
- Mobile Element (5)
- Regulators (10)
- Respiration Anaerobic (8)
- Secretion (3)
- Sporulation (29)
- Stress (5)
- Translation (11)
- UNKNOWN (58)
- Virulence Factors (1)

437 genes selectionned (80%)

- 18 class of grp
- 250 groups of genes
- 113 genes non important (20%)
 - Considered by user

Kinetics Genome

Gene	8H	14H	38H	8H vitro	14H vitro	38H vitro
CD1906	5	6	6	6	6	6
CD1907	3	3	3	3	3	3
CD1908	6	12	12	6	12	12
CD1909	11	21	21	11	21	21
CD1910	7	12	12	7	12	12

Sporulation Class

54% Sporulation genes regulated
90% up-regulated, 10% down-regulated
Spore count in cecum
8H: 2.10²
14H & 38H: >4.10⁴
Comments:
Internal verification of manipulation
What about the sporulation kinetics ?

Gene ID	Gene Name	Description	8H vivo	14H vivo	38H vivo	8H vitro	14H vitro	38H vitro
CD1906	ethA	Ethanolamine utilization protein	1.00	2.12	0.56	0.35	0.43	0.43
CD1907	ethB	Ethanolamine utilization protein	1.00	2.07	0.28	0.26	0.34	0.34
CD1908	ethC	Ethanolamine utilization protein	1.00	2.49	3.96	0.27	0.37	0.45
CD1909	ethD	Ethanolamine utilization protein	1.00	1.90	0.84	0.21	0.45	0.45
CD1910	ethE	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1911	ethF	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1912	ethG	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1913	ethH	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1914	ethI	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1915	ethJ	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1916	ethK	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1917	ethL	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1918	ethM	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1919	ethN	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1920	ethO	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1921	ethP	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1922	ethQ	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1923	ethR	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1924	ethS	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1925	ethT	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1926	ethU	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1927	ethV	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1928	ethW	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1929	ethX	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1930	ethY	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43
CD1931	ethZ	Ethanolamine utilization protein	1.00	2.00	0.51	0.20	0.43	0.43

Ethanolamine and Sorbitol Groups

Ethanolamine utilization
• up-regulated *in vivo*
• down-regulated at all time in kinetics vitro

Sorbitol utilization
• up-regulated only at 14h in vivo
• down-regulated at 14h & 38h in vitro

- Sorbitol / Ethanolamine not present in *in vitro* media
- Sorbitol / Ethanolamine present in mice food uptake

Comments:
Why ethanolamine response increase over time in vivo ?
Why sorbitol response is only at 14h in vivo ?