Modeling and simulation of gene regulatory networks 1

Hidde de Jong
IBIS
INRIA Grenoble – Rhône-Alpes
Hidde.de-Jong@inria.fr

December 4, 2013

INRIA Grenoble - Rhône-Alpes and IBIS

- IBIS: systems biology group at INRIA/Université Joseph Fourier/CNRS
  - Analysis of bacterial regulatory networks by means of models and experiments
  - Biologists, computer scientists, mathematicians, physicists, …

http://ibis.inrialpes.fr

Inria

UNIVERSITE JOSEPH FOURIER
SCIENCES TECHNOLOGIE MEDECINE

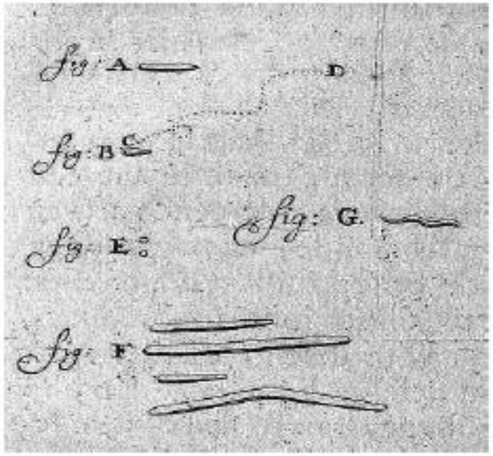
Overview

1. **Gene regulatory networks in bacteria**
2. Deterministic modeling of gene regulatory networks
3. Qualitative modeling of gene regulatory networks
4. Stochastic modeling of gene regulatory networks
5. Some current issues and perspectives

Bacteria

- Bacteria were first observed by Antonie van Leeuwenhoek, using a single-lens microscope of his own design

http://commons.wikimedia.org/



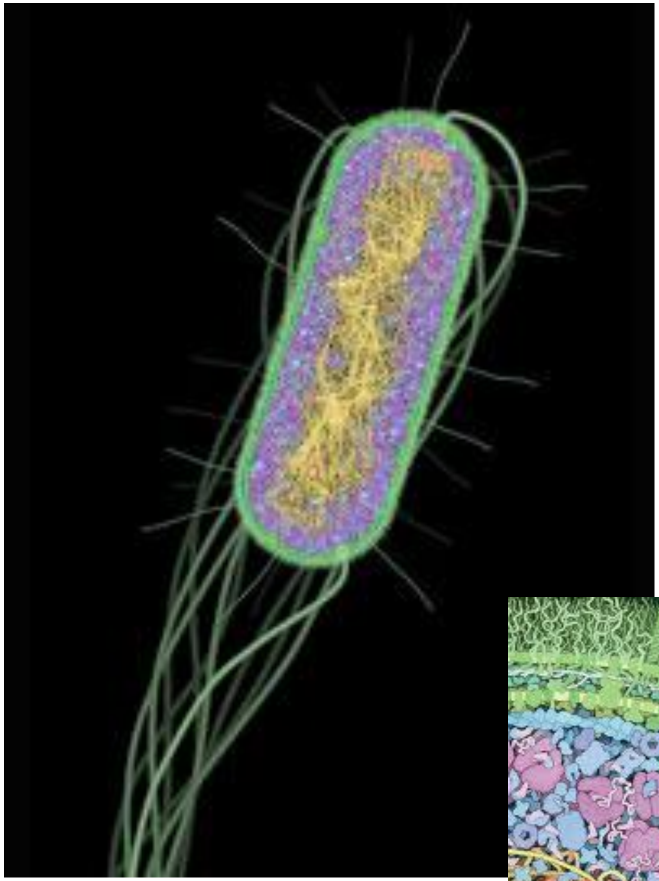
www.euronet.nl/users/warnar/leeuwenhoek.html.

van Leeuwenhoek A (1684), *Philosophical Transactions (1683–1775)* 14: 568–574

*"In the morning I used to rub my teeth with salt and rinse my mouth with water and after eating to clean my molars with a toothpick.... I then most always saw, with great wonder, that in the said matter there were many very **little living animalcules**, very prettily a-moving. The biggest sort had a very strong and swift motion, and shot through the water like a pike does through the water; mostly these were of small numbers."*

Bacteria are complex living systems

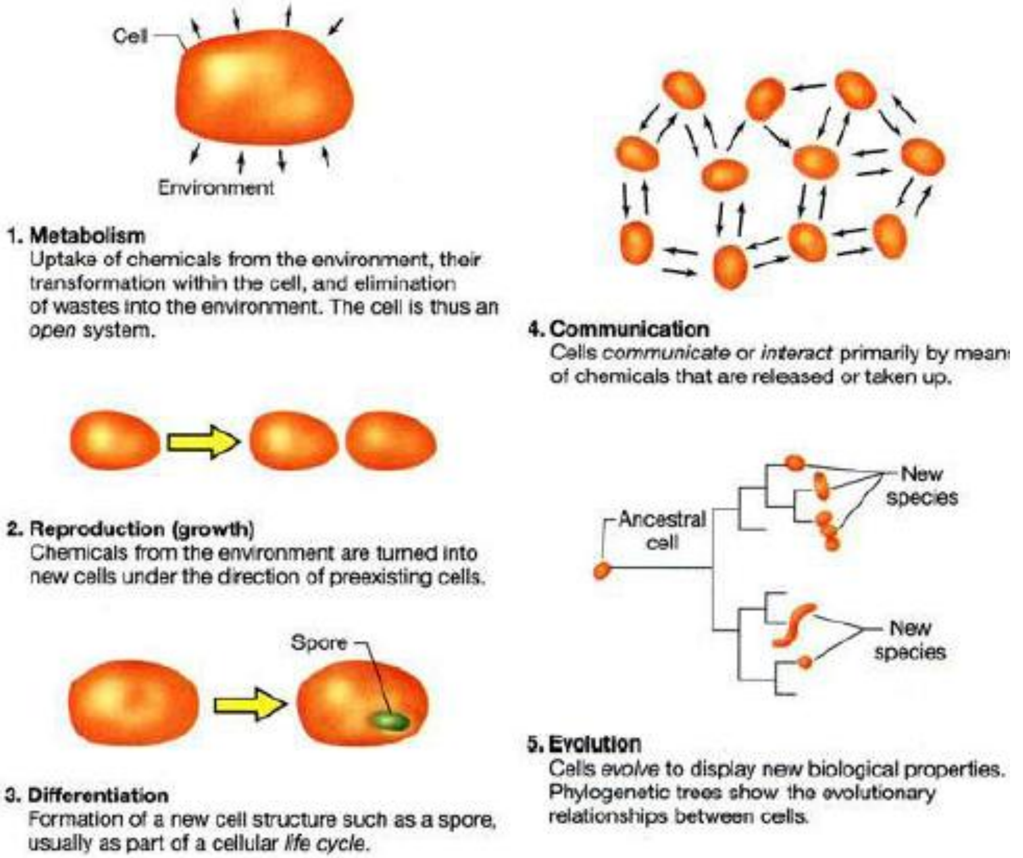
- Bacterial cells are complex biochemical and biophysical machines
  - Wide range of shapes, typically 0.5-5 µm in length
  - 10^6 bacterial cells in 1 ml of fresh water
  - 10 times as much bacterial cells as human cells in human body



Goodsell (2010), *The Machinery of Life, Springer,* 2nd ed.

Bacteria are complex living systems

- Bacterial cells are complex biochemical and biophysical machines
- Bacteria possess characteristics shared by most living systems:
  - Metabolism
  - Growth and reproduction
  - Differentiation
  - Communication
  - Evolution

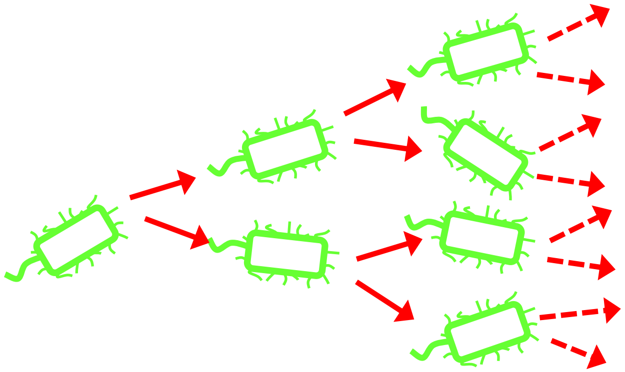


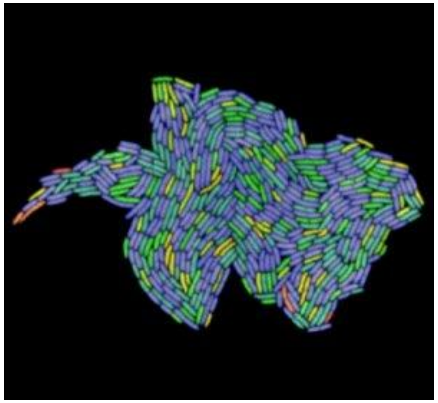
Madigan *et al.* (2003), *Brock Biology of Microorganisms*, Prentice Hall, 10th ed.

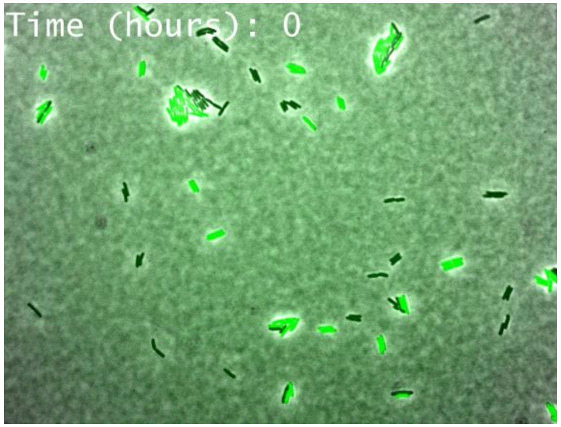
Bacterial growth and metabolism

- Bacteria are geared towards **growth** and **division**

  Escherichia coli cells have doubling times up to 20 min







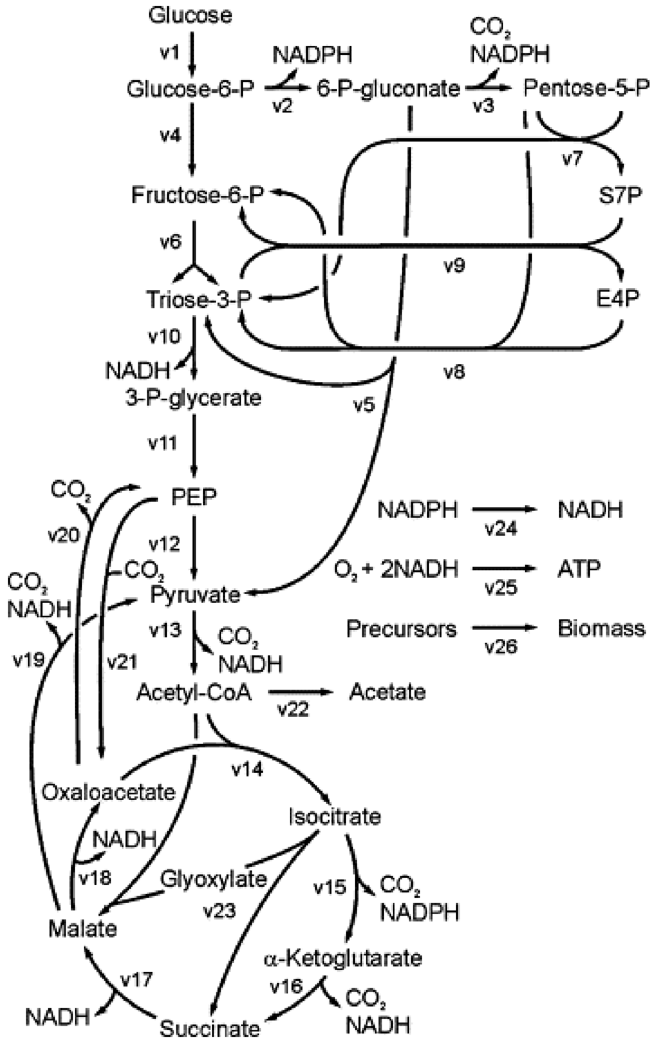
G. Baptist

- **Metabolism** fuels growth by production of energy and building blocks for macromolecules, using nutriments from environment

  ATP, amino acids, nucleotides, …

Bacterial growth and metabolism

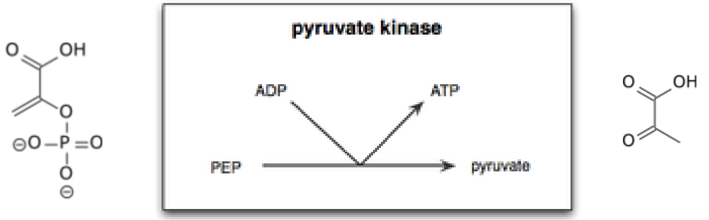
- Central **carbon metabolism** breaks down carbon sources for energy production and macromolecular synthesis
  - Glucose, acetate, lactose, …

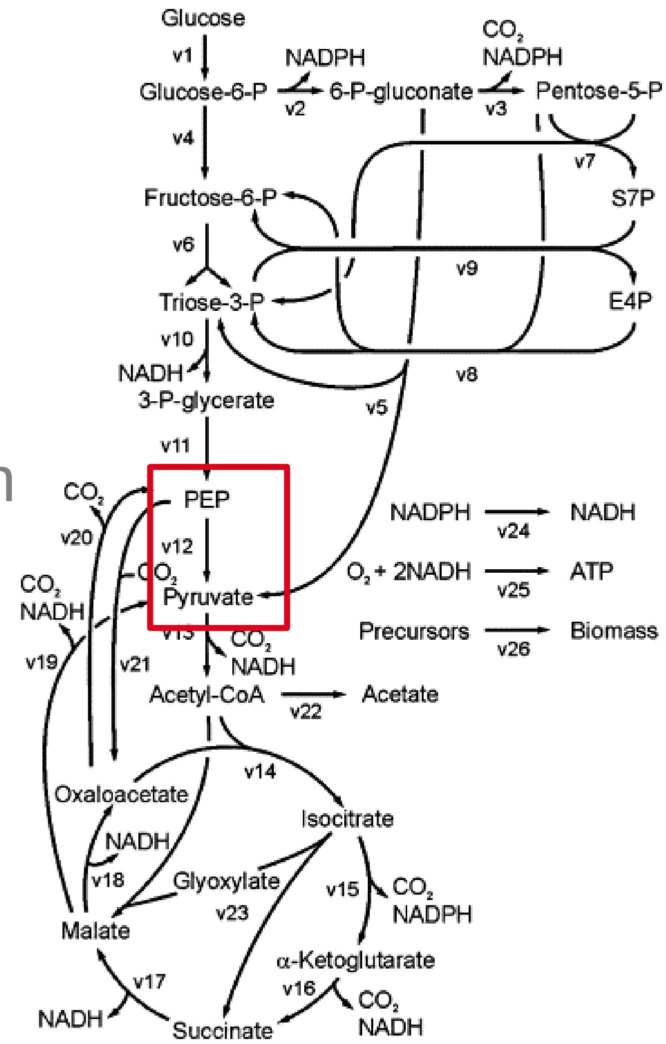


Fischer *et al.* (2004), *Anal. Biochem.*, 325(2):308–16

Bacterial growth and metabolism

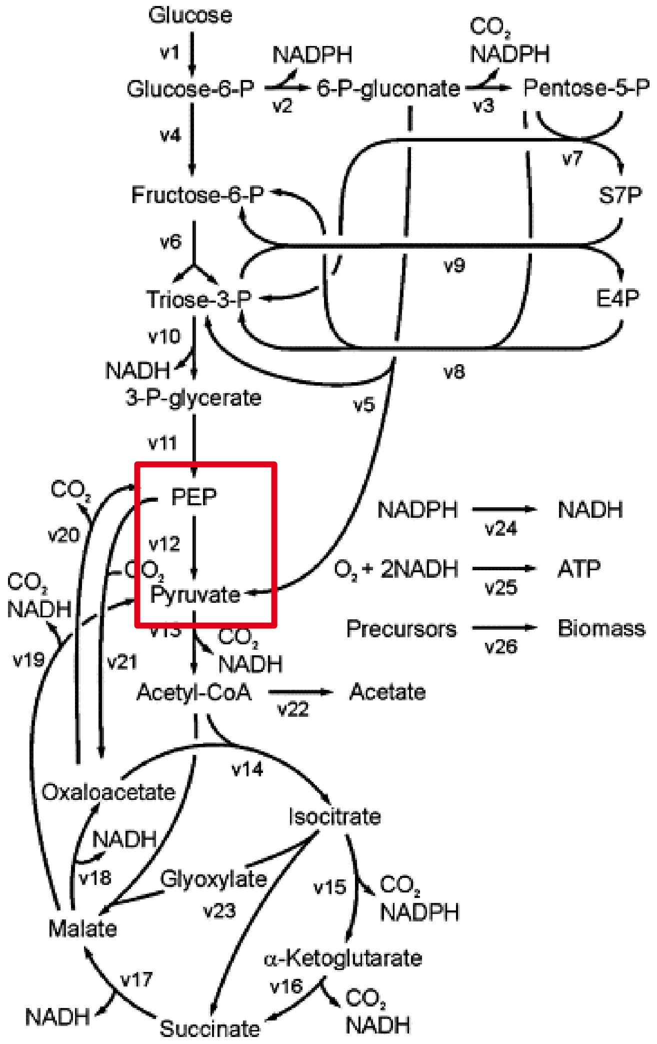
- Central **carbon metabolism** breaks down carbon sources for energy production and macromolecular synthesis
  - Glucose, acetate, lactose, …
- **Enzymes** catalyse individual steps in metabolic network
  - Pyruvate kinase transforms phosphoenolpyruvate (PEP) into pyruvate





Bacterial growth and metabolism

- Central **carbon metabolism** breaks down carbon sources for energy production and macromolecular synthesis
  - Glucose, acetate, lactose, …
- Enzymes produced from information encoded in **genes**
  - *pykF* is gene encoding pyruvate kinase

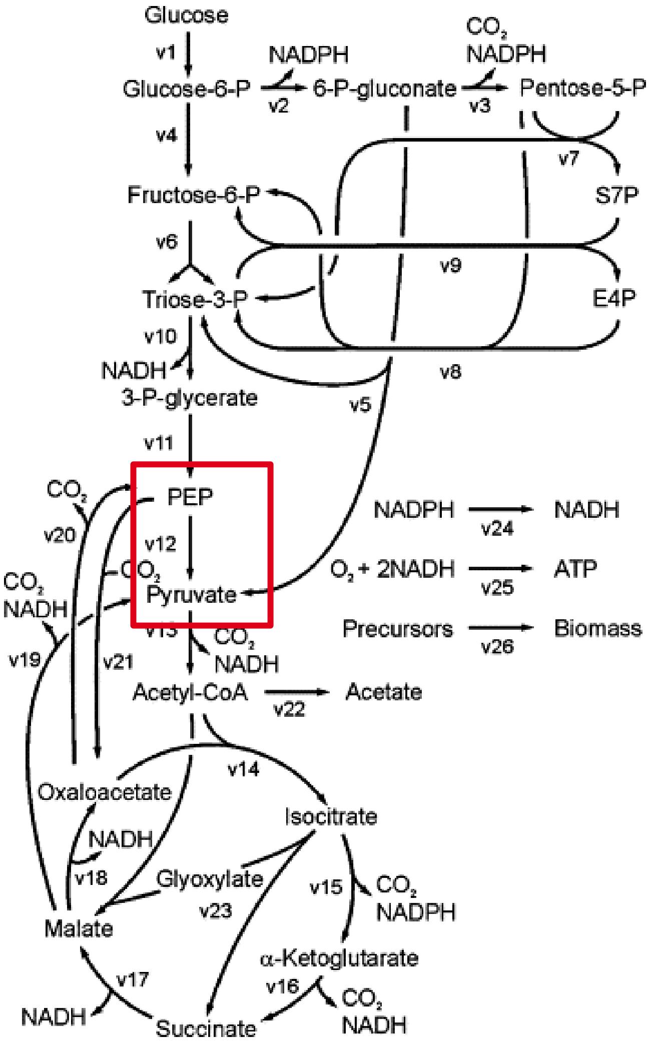


Bacterial growth and metabolism

- Central **carbon metabolism** breaks down carbon sources for energy production and macromolecular synthesis

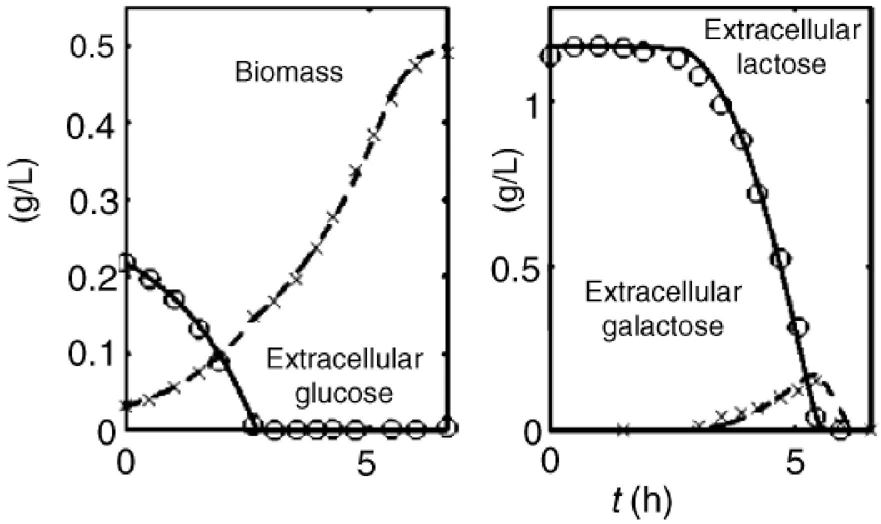
  Glucose, acetate, lactose, …

- Enzymes produced from information encoded in **genes**
  - *pykF* is gene encoding pyruvate kinase
  - Expression of *pykF* regulated by transcription factor Cra



Bacterial growth and metabolism

- Bacterial metabolism is **flexible**, allowing cells to grow on different carbon sources
    - Preferential utilisation: **diauxic growth** on glucose and lactose



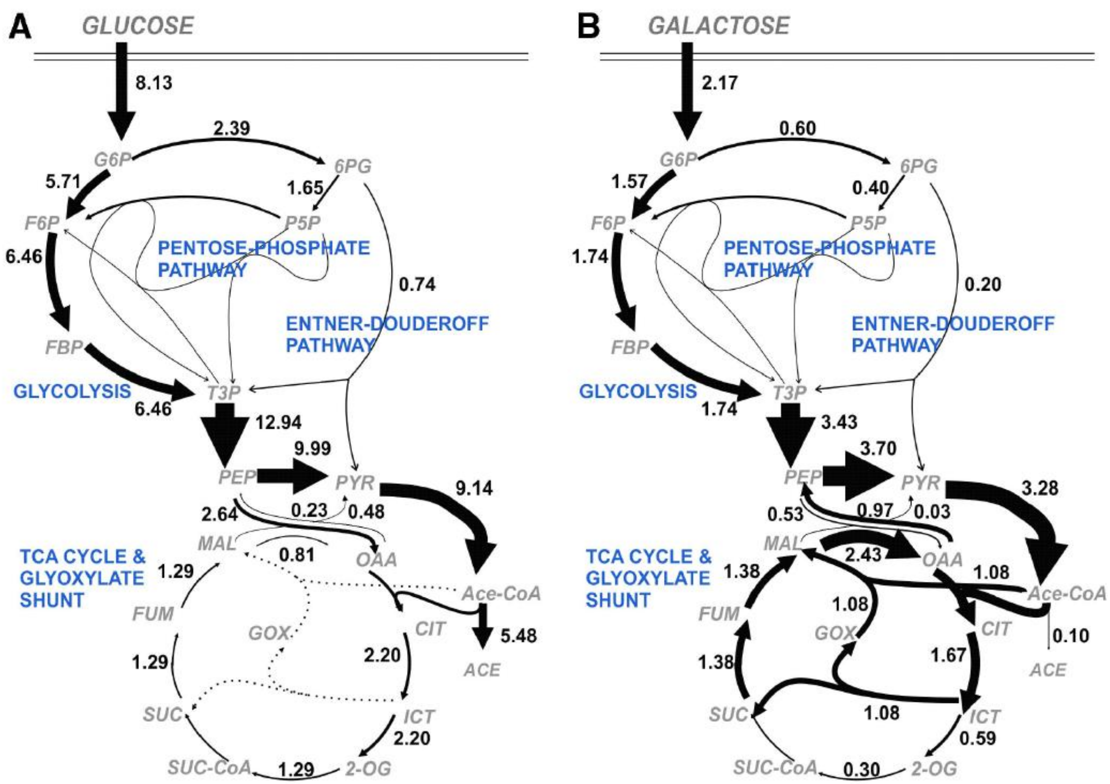
Bettenbrock *et al.* (2006), *J. Biol. Chem.*, 281(5):2578-84

- Adaptation of bacterial physiology to different carbon sources

Growth transition and metabolism

- Adaptation to different carbon source involves changes in **metabolic fluxes**

  Different flux distribution in central metabolism of *E. coli* during growth on glucose and galactose



Haverkorn van Rijsewijk *et al.* (2011), *Mol. Syst. Biol.*, 7:477

Growth transition and metabolism

- Adaptation to different carbon source involves adjustment of **metabolite concentrations**
  - Different metabolite concentrations in *E. coli* cells growing on glucose and acetate

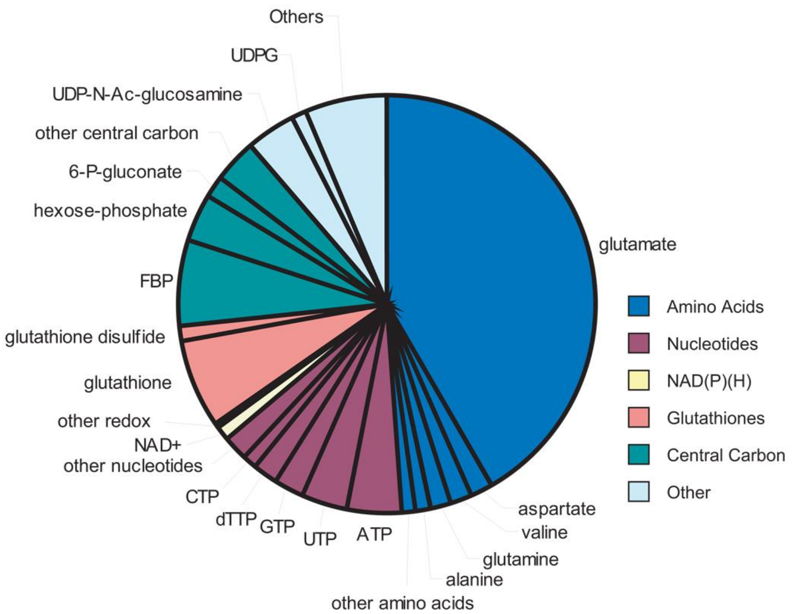


Table 1 Intracellular metabolite concentrations in glucose-fed, exponentially growing *E. coli*

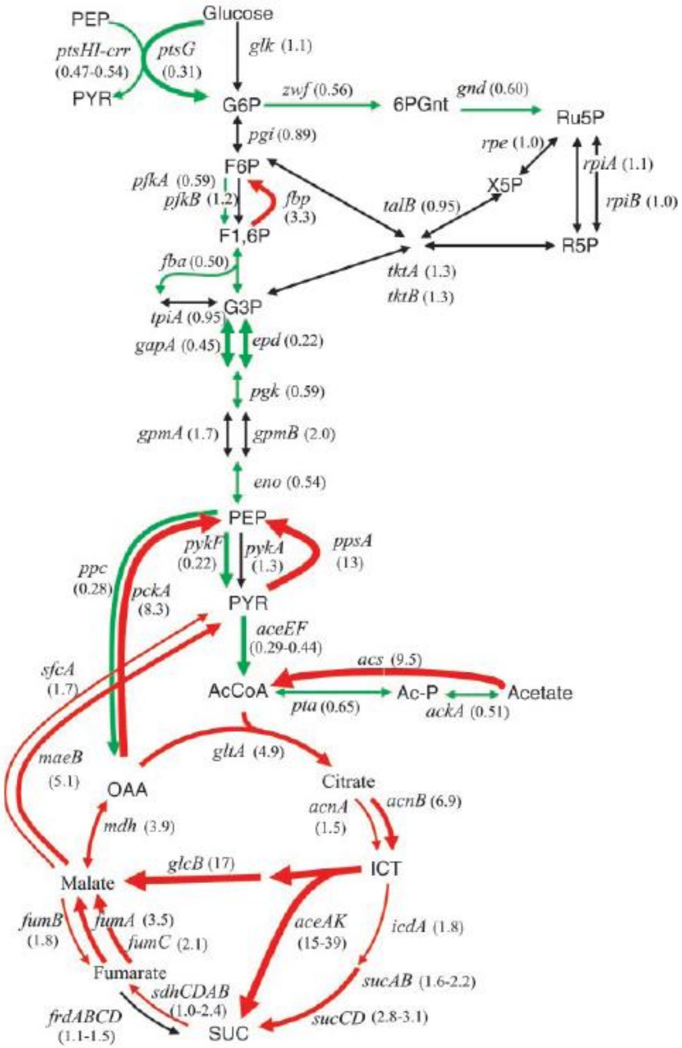
Metabolite	mol l^{-1}	Metabolite	mol l^{-1}
Glutamate	9.6×10^{-2}	UDP-glucuronate (51)	5.7×10^{-4}
Glutathione	1.7×10^{-2}	ADP	5.6×10^{-4}
Fructose-1,6-bisphosphate	1.5×10^{-2}	Asparagine (52)	5.1×10^{-4}
ATP	9.6×10^{-3}	α-Ketoglutarate	4.4×10^{-4}
UDP-*N*-acetylglucosamine (29)	9.2×10^{-3}	Lysine (53)	4.1×10^{-4}
Hexose-P[a]	8.8×10^{-3}	Proline (54)	3.9×10^{-4}
UTP (30)	8.3×10^{-3}	dTDP (55)	3.8×10^{-4}
GTP (31)	4.9×10^{-3}	Dihydroxyacetone phosphate	3.7×10^{-4}
dTTP	4.6×10^{-3}	Homocysteine (56)	3.7×10^{-4}
Aspartate	4.2×10^{-3}	CMP (57)	3.6×10^{-4}
Valine (32)	4.0×10^{-3}	Deoxyribose-5-P (58)	3.0×10^{-4}
Glutamine	3.8×10^{-3}	Isoleucine (59) + leucine (60)	3.0×10^{-4}
6-Phosphogluconate	3.8×10^{-3}	AMP	2.8×10^{-4}

Bennett *et al.* (2009), *Nat. Chem. Biol.*, 5(8):593-9

UNIVERSITE JOSEPH FOURIER SCIENCES TECHNOLOGIE MEDECINE

Growth transition and gene expression

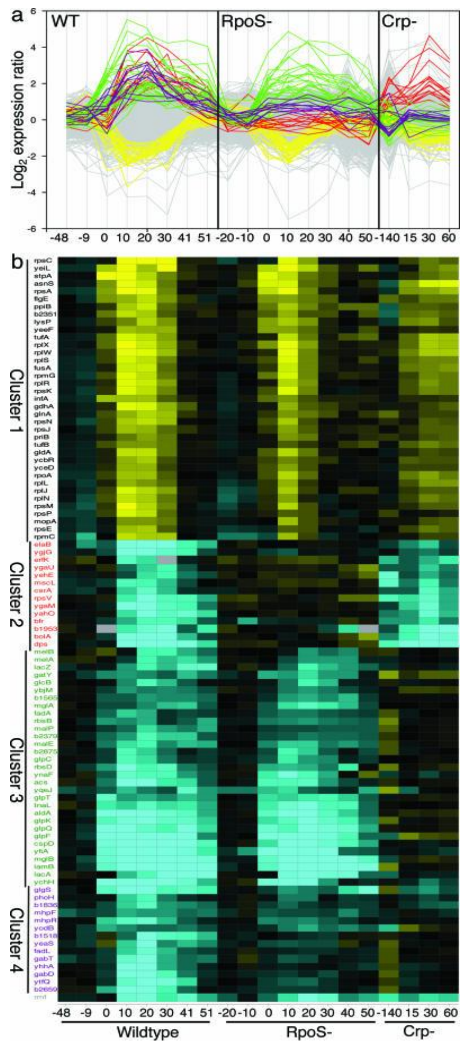
- Adaptation to different carbon source involves adjustment of **expression of enzymatic genes**
    - Difference in expression levels of genes encoding enzymes in central metabolism of *E. coli* during growth on glucose and acetate



Oh *et al.* (2002), *J. Biol. Chem.*, 277(15):13175–83

Growth transition and gene expression

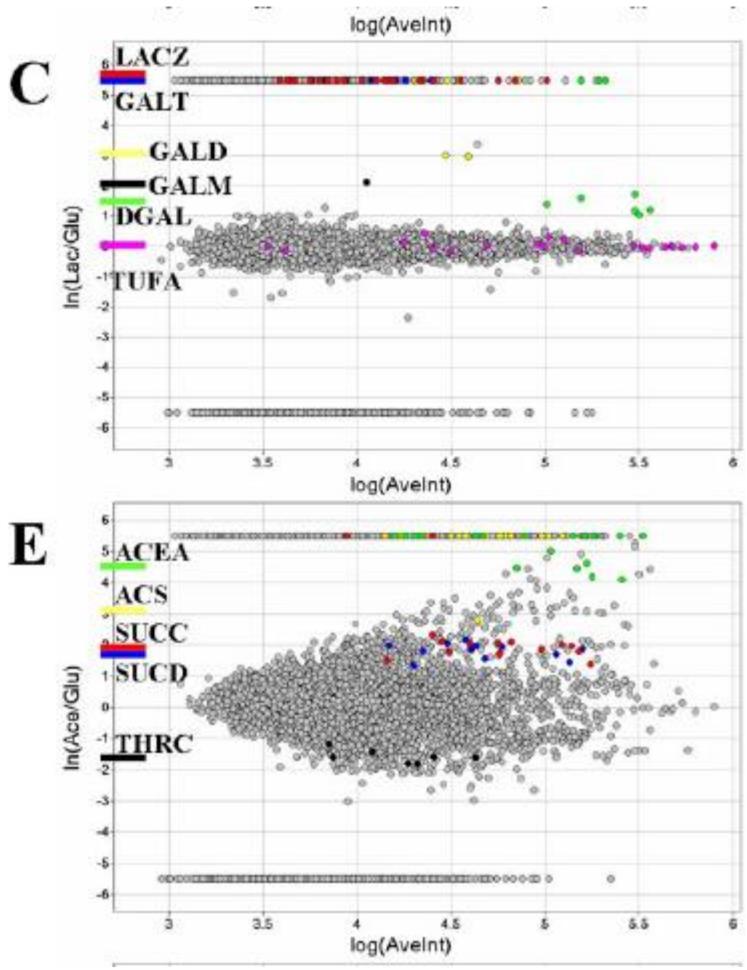
- Adaptation to different carbon source involves involves genome-wide **reorganization of gene expression**
    - Gene expression during glucose-acetate shift in *E. coli*



Traxler *et al.* (2006), *Proc. Natl. Acad. Sci. USA*, 103(7):2374–9

Growth transition and gene expression

- Adaptation to different carbon source involves involves adjustment of **protein levels**
  - Changes in relative protein abundance during growth on lactose vs glucose and acetate vs glucose

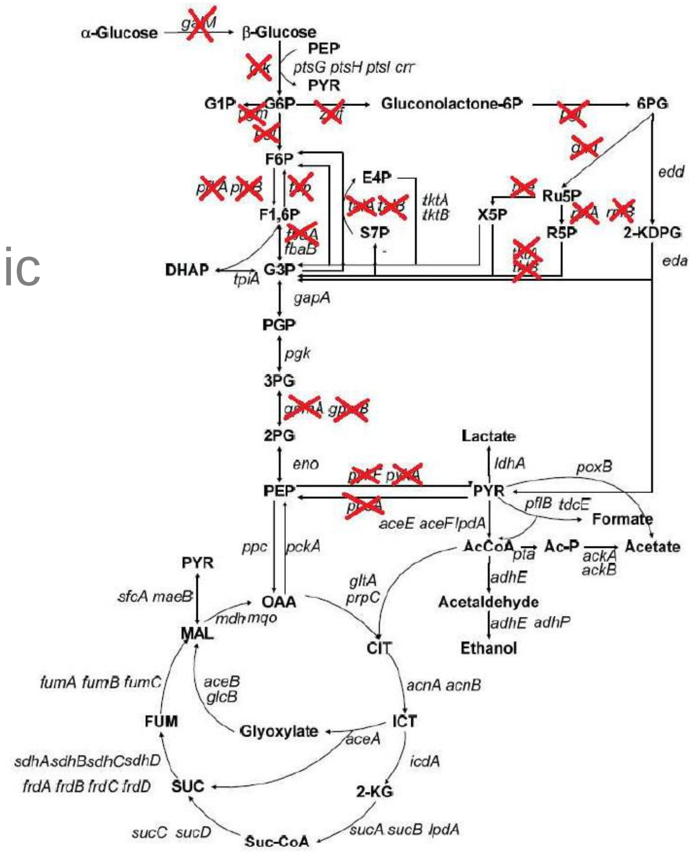


Silva *et al.* (2006), *Mol. Cell. Proteom.*, 5(4):589–607

Adaptation on multiple levels

- Adaptation to different carbon source involves **adjustments on multiple levels** at the same time!
  - Parallel measurement of enzyme and metabolite concentrations, and metabolic fluxes in a variety of experimental conditions

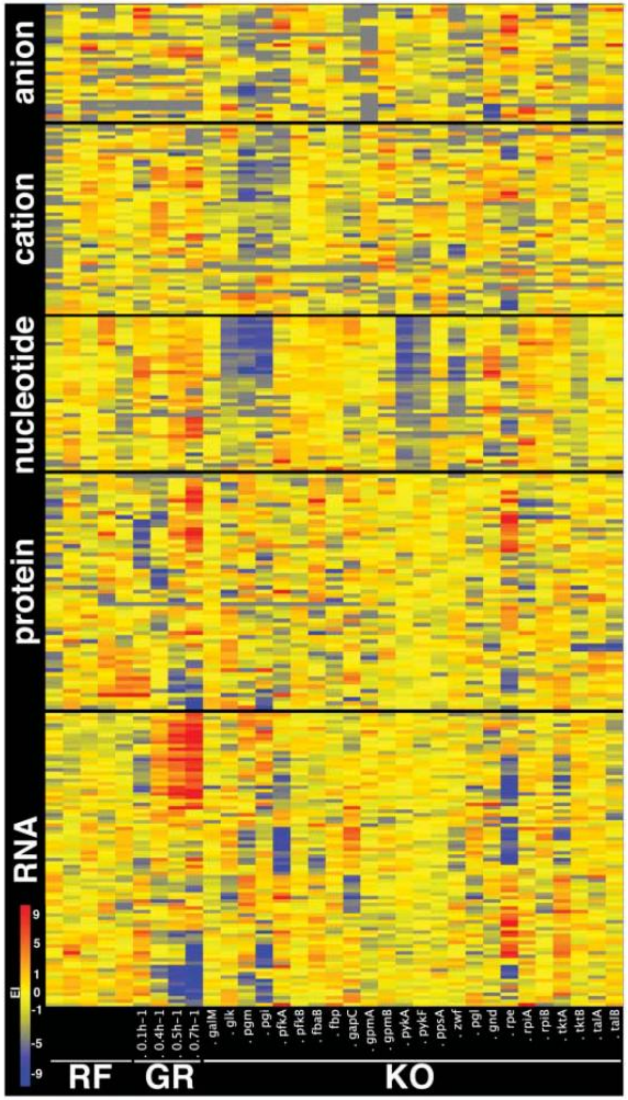
Ishii *et al.* (2007), *Science*, 316(5284):593-7



Adaptation on multiple levels

- Adaptation to different carbon source involves **adjustments on multiple levels** at the same time!
  - Parallel measurement of enzyme and metabolite concentrations, and metabolic fluxes in a variety of experimental conditions

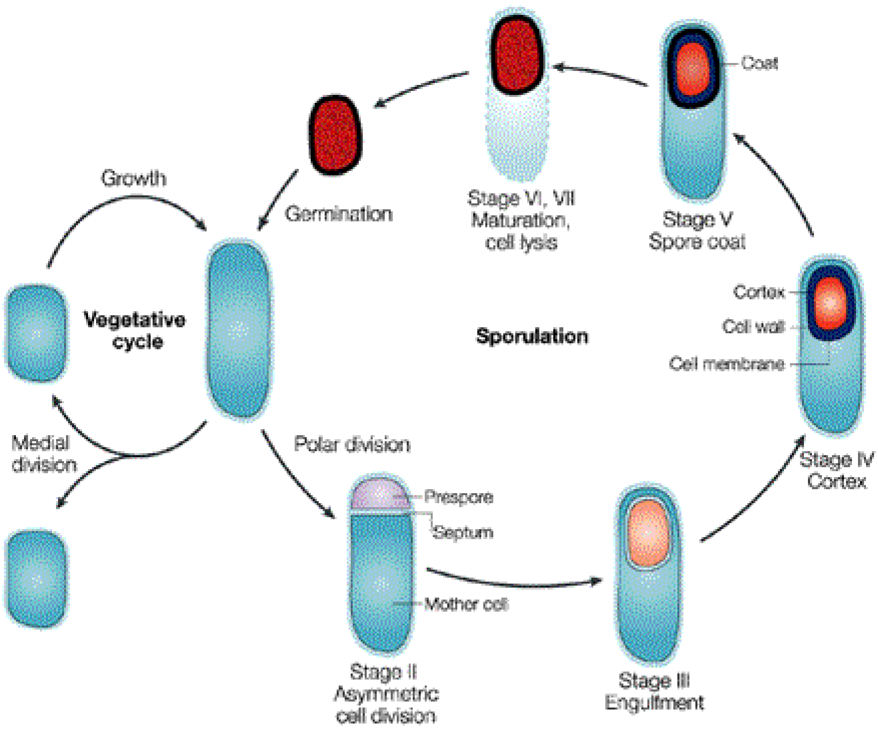
Ishii *et al.* (2007), *Science*, 316(5284):593-7



Sporulation and cell differentiation

- *Bacillus subtilis* cells can form resistant **spores** when environmental conditions become unfavorable (starvation)

  Ultimate response in repertoire of stress responses (motility, toxin release, competence, …)



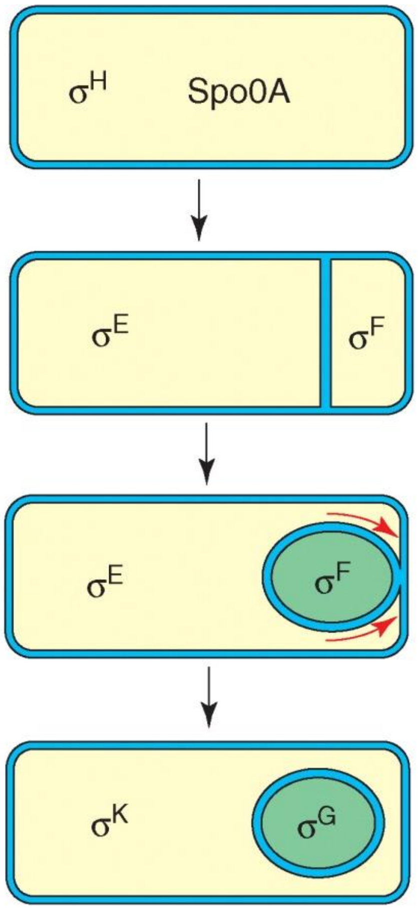
- Asymmetric cell division produces smaller forespore cell and larger mother cell

  Prototype of cell differentiation and intercellular signalling

Errington (2005), *Nat. Rev. Microbiol.*, 1(2):117-126

Sporulation and cell differentiation

- Precise temporal ordering of events on molecular level
- Specific proteins control differentiation processes in different stages of sporulation
  - Transcription factors, **sigma factors**



Piggot and Hilbert (2004), *Curr. Opin. Microbiol.*, 7(6):579-86

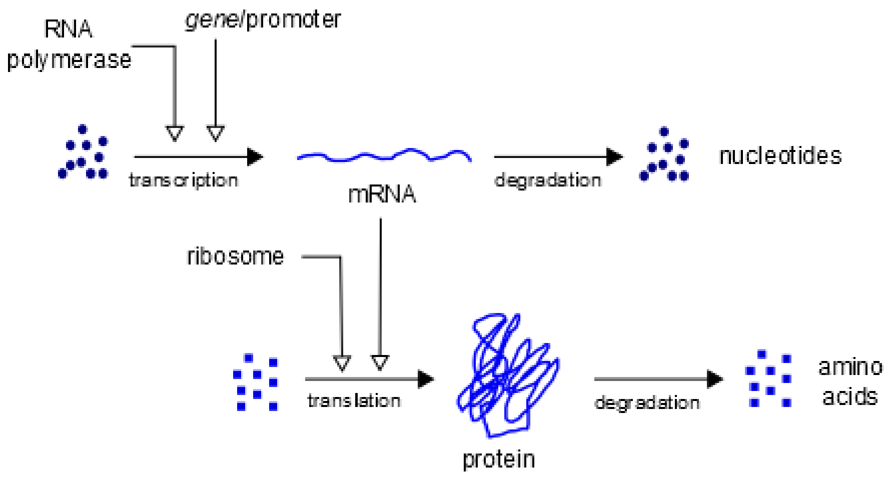
General questions on cellular adaptation

- Cells are capable of responding to a variety of changes in their environment by adapting their physiology
    - Change in carbon source, starvation, population density, …
- On the molecular level, these responses involve adjustment of protein concentrations in the cell
    - Enzymes, sigma factors, transcription factors, …
- **Question**: how can protein concentrations change in response to specific environmental changes?
- **Question**: how does cell coordinate changes in concentration of a variety of proteins?
- Changes in protein concentrations involve changes in gene expression

Gene expression

- Typically, and simplifying quite a bit, **gene expression** in bacteria involves:
  - Transcription by RNAP (mRNA)
  - Translation by ribosomes (proteins)
  - Degradation of mRNA and protein

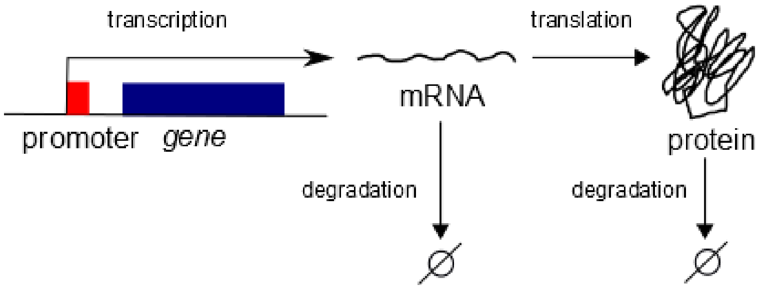
Biochemical view:



Gene expression

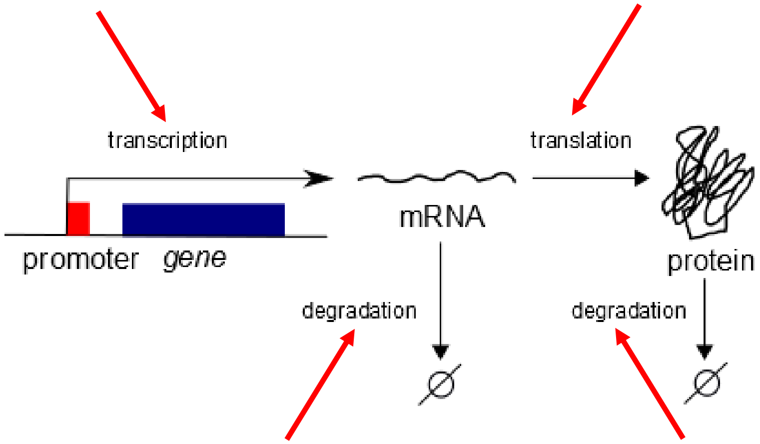
- Typically, and simplifying quite a bit, **gene expression** in bacteria involves:
  - Transcription by RNAP (mRNA)
  - Translation by ribosomes (proteins)
  - Degradation of mRNA and protein

Simplified view:



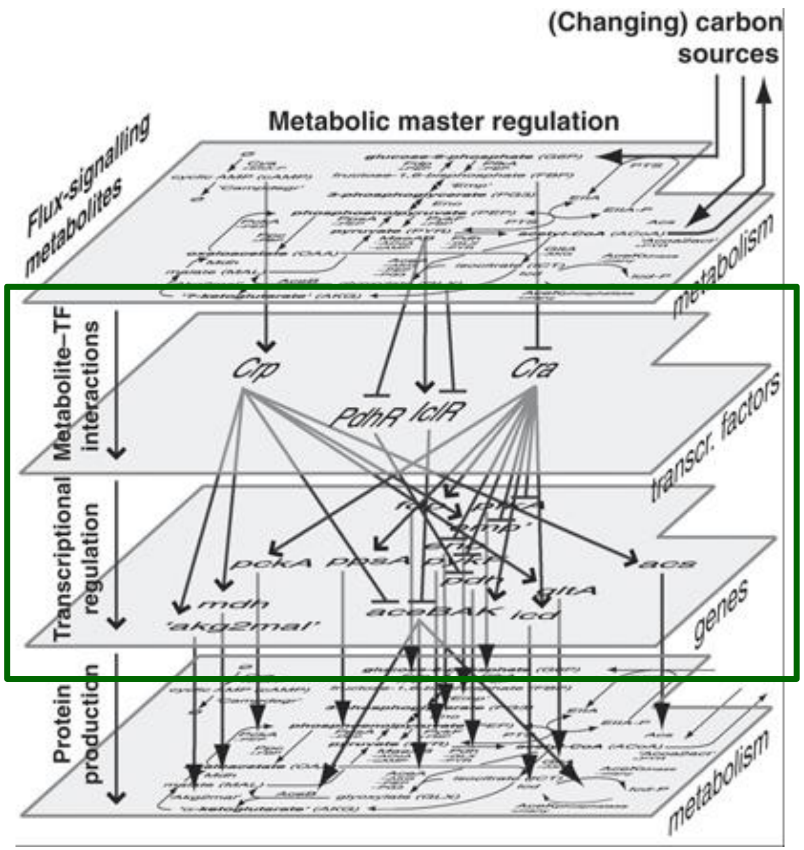
Regulation of gene expression

- Typically, and simplifying quite a bit, **regulation of gene expression** in bacteria involves:
  - Transcription regulation by transcription factors
  - Translation regulation by small RNAs
  - Regulation of degradation by proteases



Gene regulatory networks

- **Gene regulatory networks** control changes in gene expression levels in response to environmental perturbations



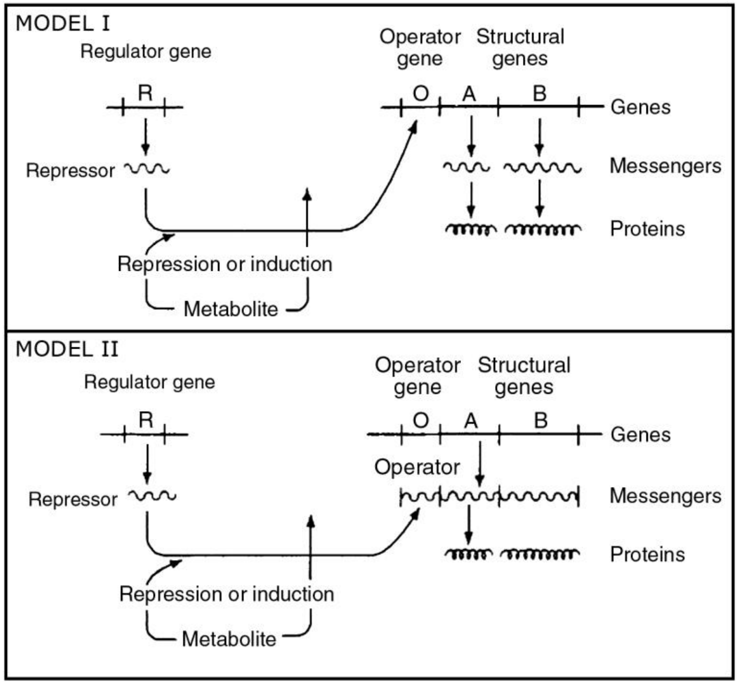
- Gene regulatory networks consist of genes, gene products, signalling metabolites, and their mutual regulatory interactions

  Global regulators of transcription involved in glucose-acetate diauxie in *E. coli*

Kotte *et al.* (2010), *Mol. Syst. Biol.*, 6:355

Gene regulatory networks

- **Gene regulatory networks** control changes in expression levels in response to environmental perturbations



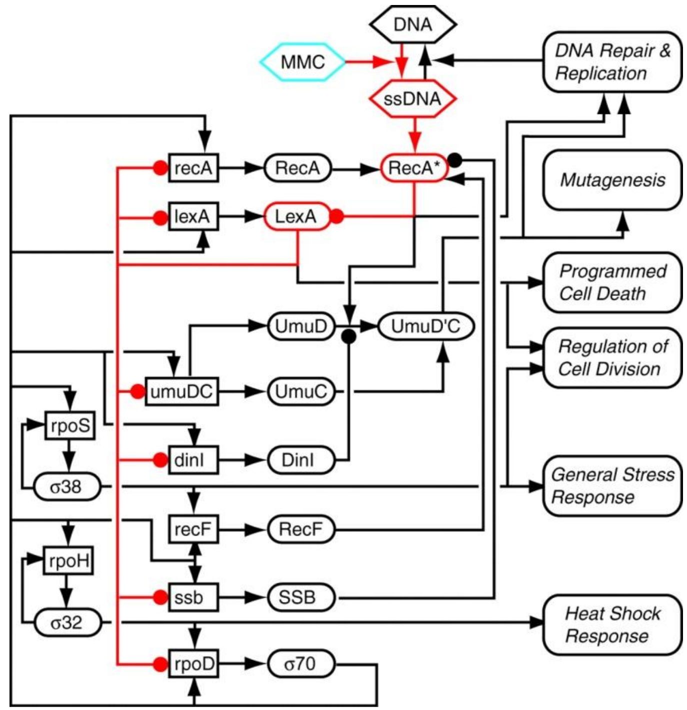
- Gene regulatory networks consist of genes, gene products, signalling metabolites, and their mutual regulatory interactions

Original *lac* operon model

Jacob and Monod (1961), *J. Mol. Biol.*, 3(3):318-56

Gene regulatory networks

- **Gene regulatory networks** control changes in expression levels in response to environmental perturbations



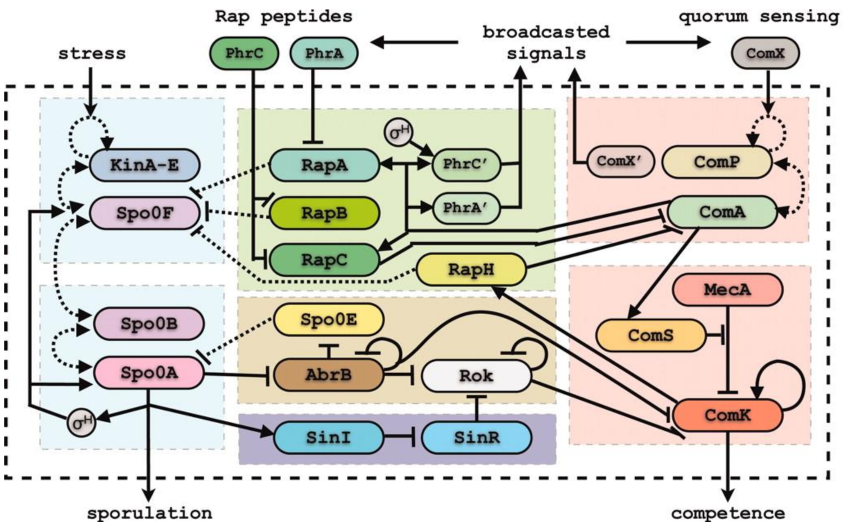
- Gene regulatory networks consist of genes, gene products, signalling metabolites, and their mutual regulatory interactions

SOS response network in *E. coli*

Gardner *et al.* (2011), *Science*, 301(5629):102-5

Gene regulatory networks

- **Gene regulatory networks** control changes in expression levels in response to environmental perturbations



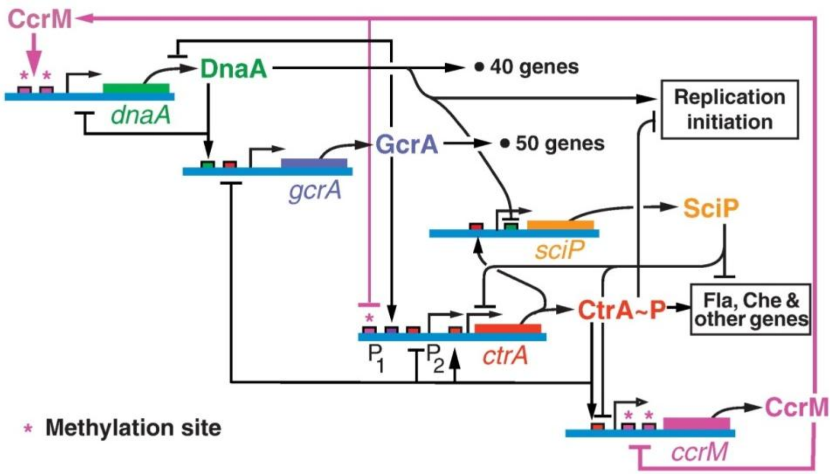
- Gene regulatory networks consist of genes, gene products, signalling metabolites, and their mutual regulatory interactions

Sporulation and competence network in *B. subtilis*

Schultz *et al.* (1961), *Proc. Natl. Acad. Sci. USA*, 106(50):21027-34

Gene regulatory networks

- **Gene regulatory networks** control changes in expression levels in response to environmental perturbations



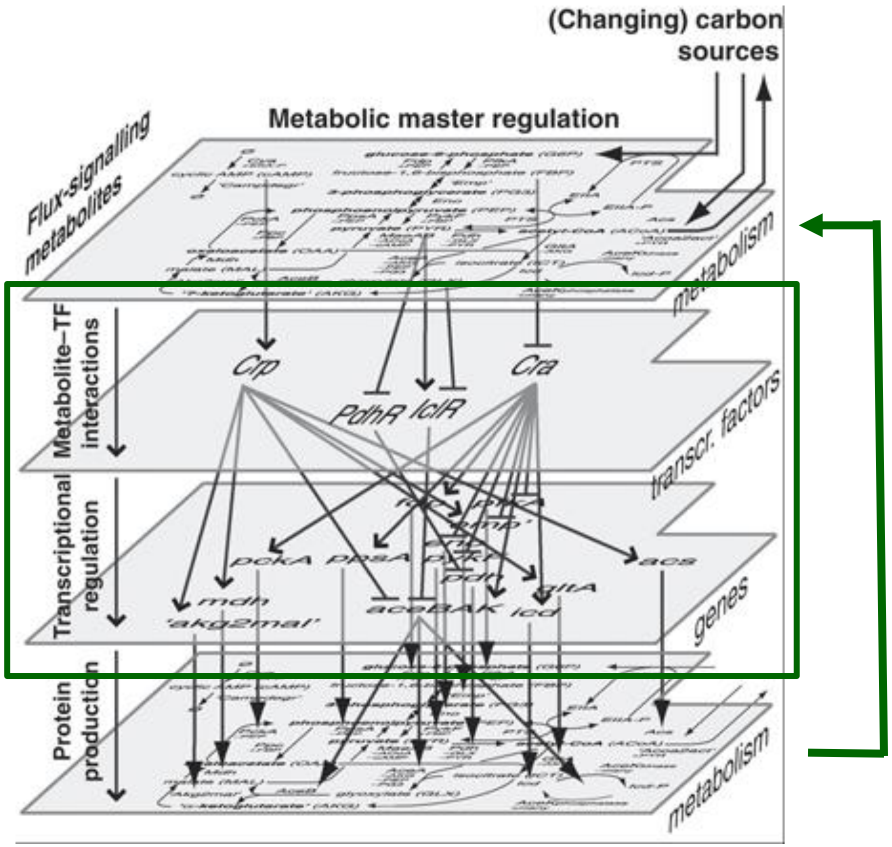
- Gene regulatory networks consist of genes, gene products, signalling metabolites, and their mutual regulatory interactions

Cauleobacter cell cycle network

McAdams and Shapiro (2011), *J. Mol. Biol.*, 409(1):28-35

Broader view on gene regulatory networks

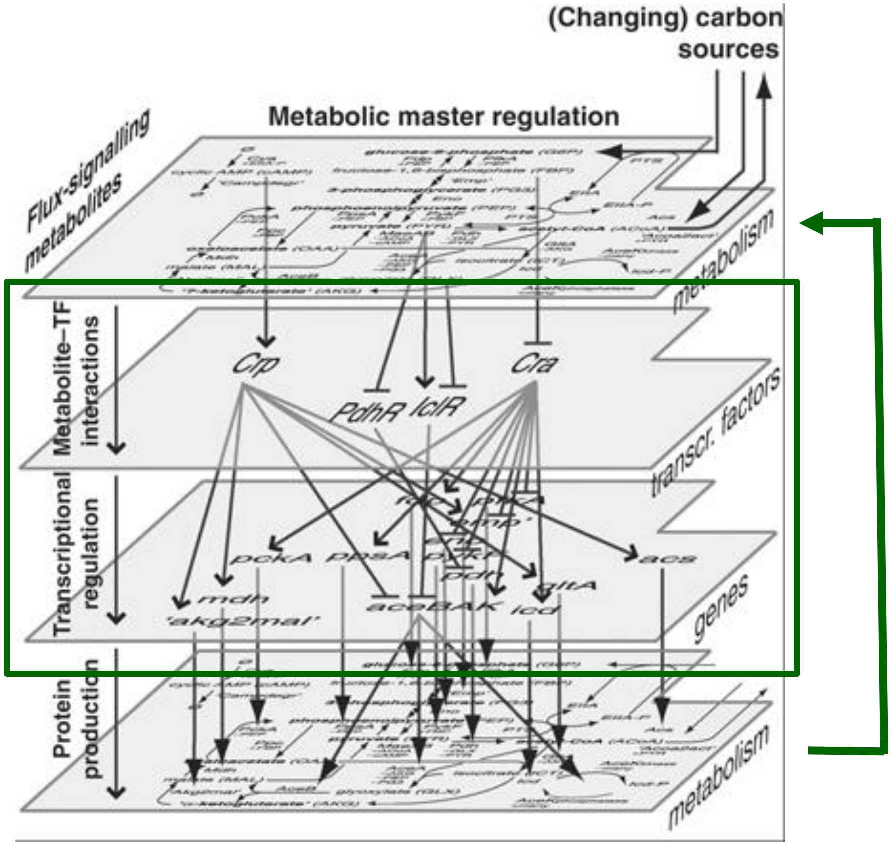
- Gene regulatory networks control changes in expression levels in response to environmental perturbations



- **But**: adaptation of gene expression leads to changes in metabolism which feed back into regulatory network
- Gene regulatory networks are intertwined with metabolic and signaling networks
    - Complex, heterogeneous systems evolving on different time-scales

Broader view on gene regulatory networks

- Gene regulatory networks control changes in expression levels in response to environmental perturbations

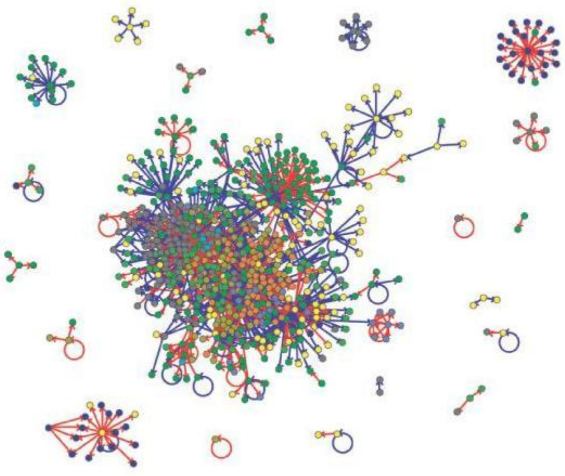


- Feedback through metabolism leads to indirect regulatory interactions: **metabolic coupling**
  - Regulatory effects of enzymes on gene expression

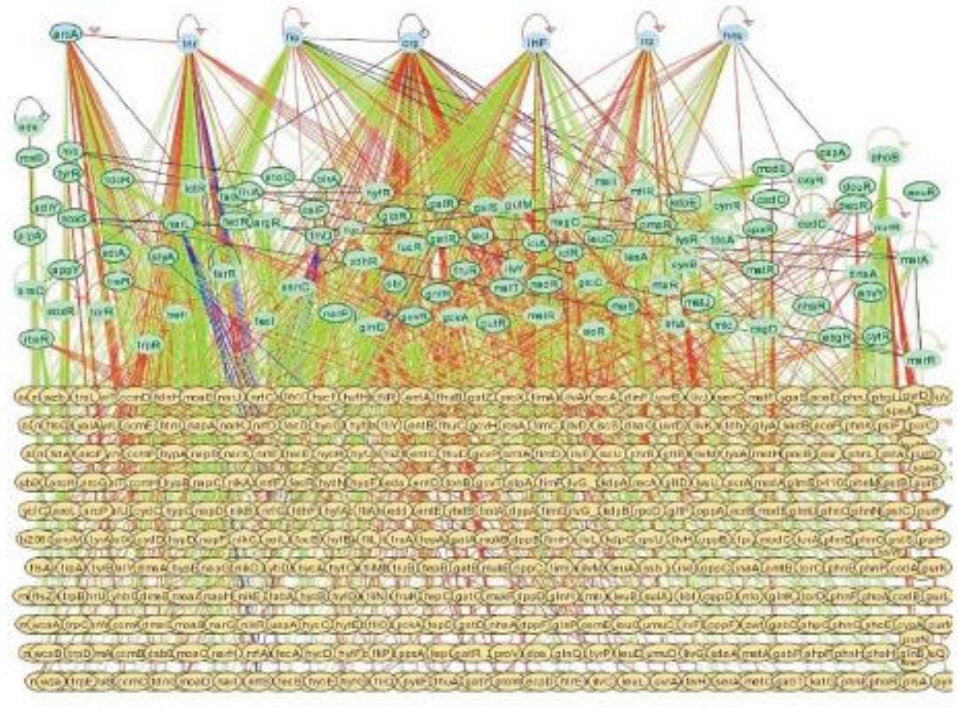
Baldazzi *et al.* (2010), *PLoS Comput. Biol.*, 6(6):e1000812

Complexity of gene regulatory networks

- Most gene regulatory networks of biological interest are large and complex
  - *E. coli* has 4200 genes coding for several hundreds of transcription factors



Cases and de Lorenzo (2005), *Nat. Rev. Microbiol.*, 3(2):105-18



Martinez-Antonio *et al.* (2003), *Curr. Opin. Microbiol.*, 6(5):482-9

Systems biology

- Most gene regulatory networks of biological interest are large and complex
- **No global view of functioning** of network available, despite abundant knowledge on network components
  - Understanding of dynamics requires **experimental tools** for monitoring gene expression over time
  - Understanding of dynamics requires **mathematical modeling** and **computer analysis and simulation**
  - Discipline now often referred to as **systems biology**

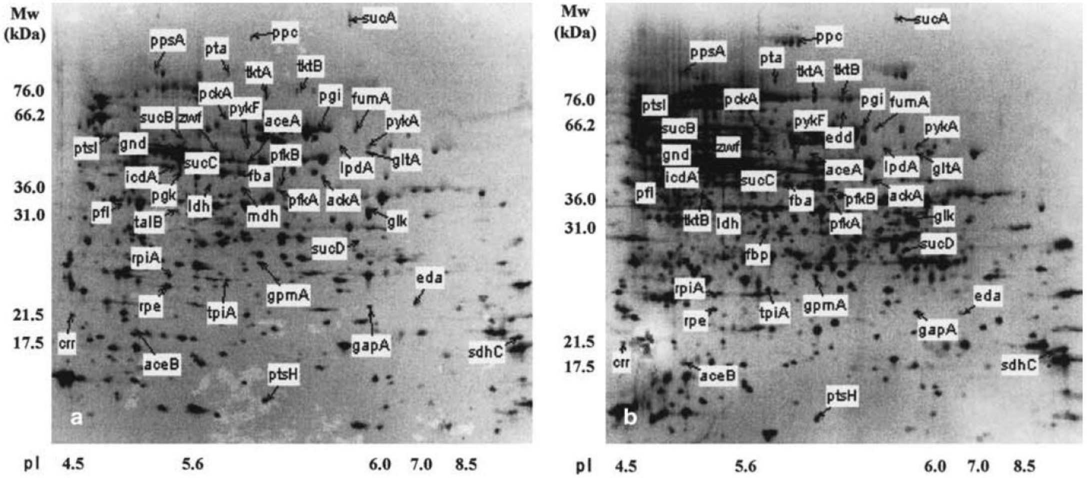
Alon (2007), *An Introduction to Systems Biology*, Chapman & Hall/CRC Press

Experimental tools

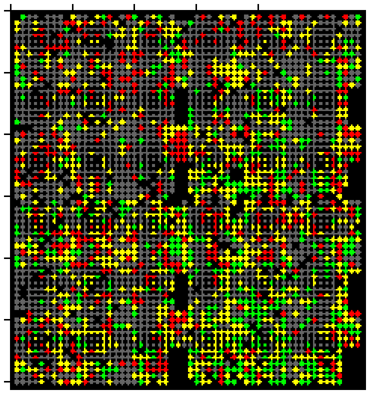
- A variety of experimental tools allow gene expression to be measured, by quantifying mRNA and protein abundancies

Peng and Shimizu (2003),
App. Microbiol. Biotechnol., 61:163-78

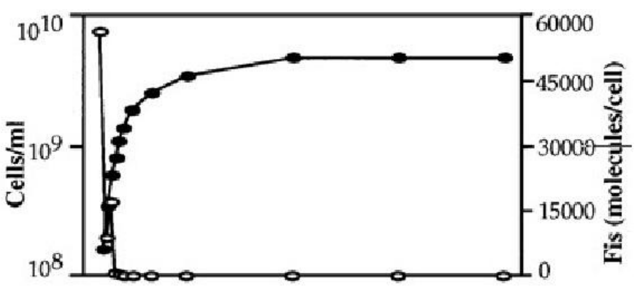
2D gels



DNA
microarrays



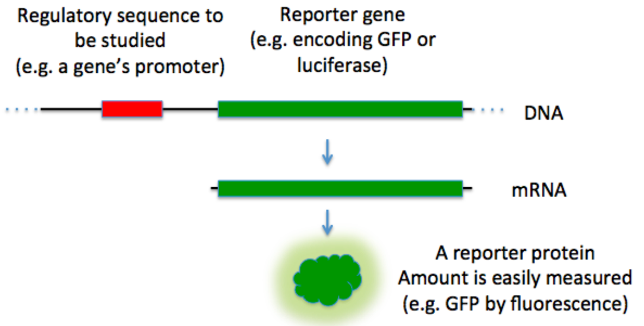
Western blots



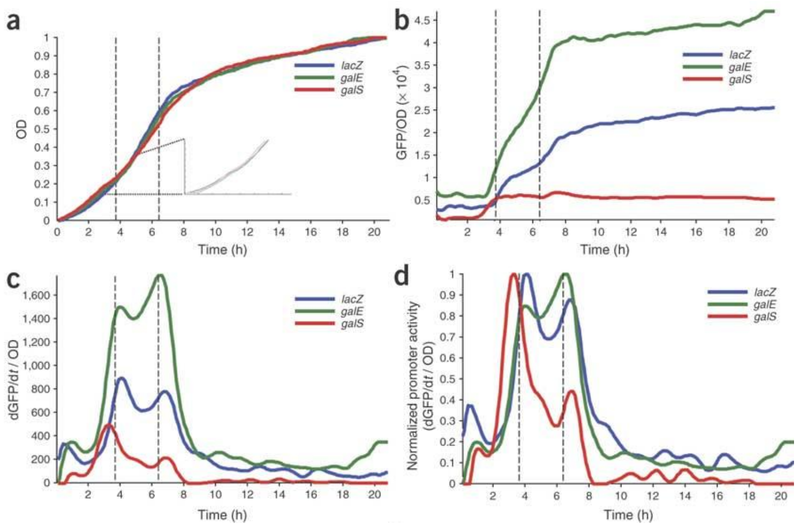
Ali Azam *et al.* (1999), *J. Bacteriol.*, 181(20):6361-70

Fluorescent reporter genes

- Use of **fluorescent reporter genes** allows expression from host promoter to be monitored *in vivo* and in real time
  - Different colors (emission peaks): GFP, YFP, RFP, …
  - Reporter genes on plasmids and on chromosome
  - Transcriptional and translational reporters



- Library of fluorescent transcriptional reporter genes in *E. coli*

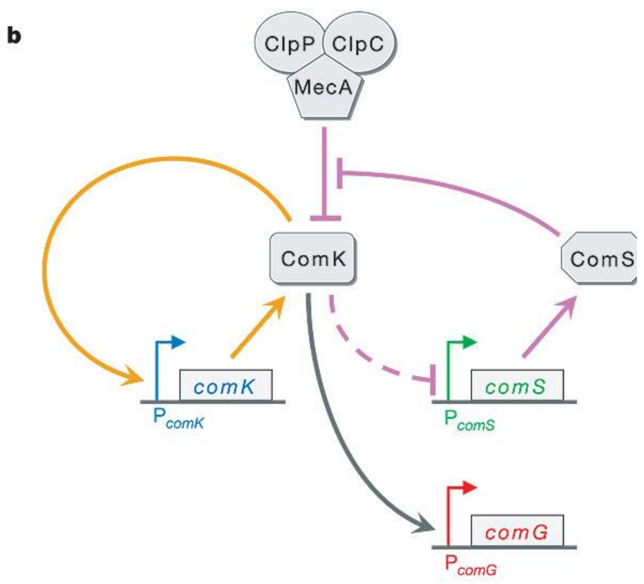


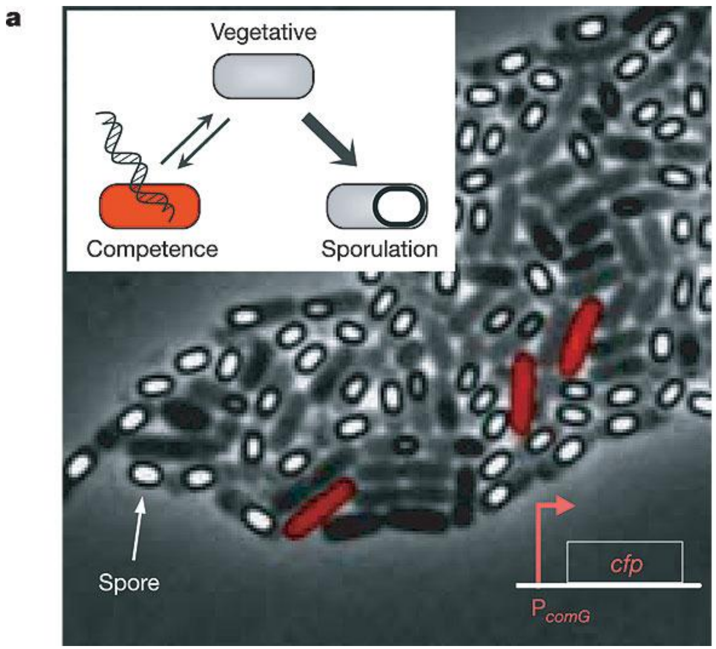
Zaslaver *et al.* (2006), *Nat. Methods,* 3(8):623-8

Single-cell microscopy

- Monitoring of gene expression in single cells using fluorescent reporters, **automated time-lapse microscopy**, and **image analysis**
- Monitoring onset of competence in *B. subtilis*

Süel *et al.* (2006), *Nature*, 440:545-50

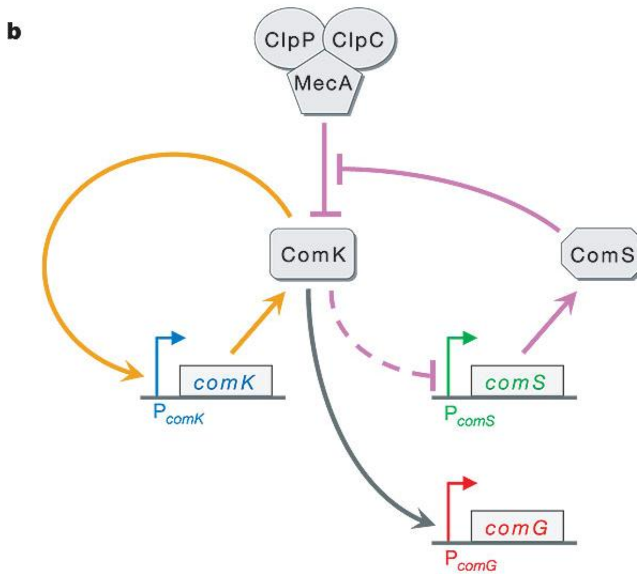


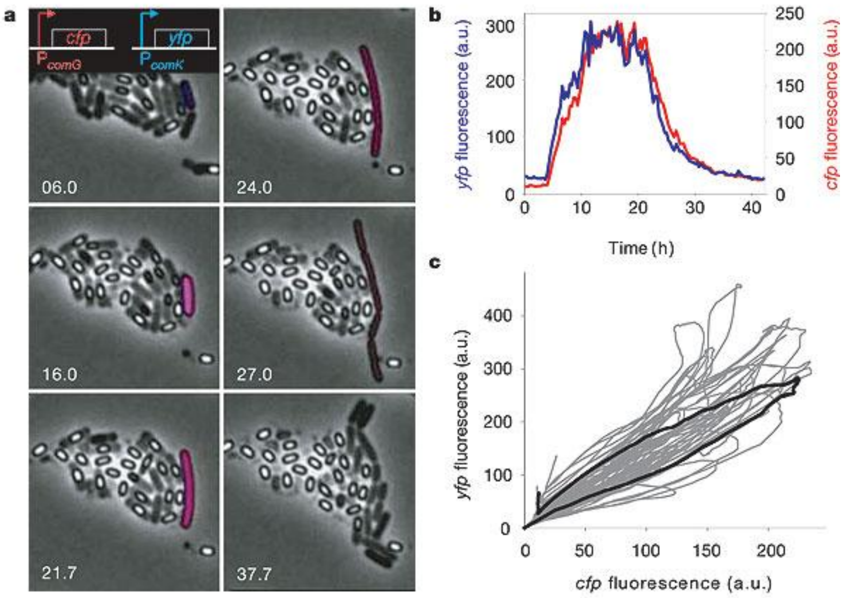


Single-cell microscopy

- Monitoring of gene expression in single cells using fluorescent reporters, **automated time-lapse microscopy**, and **image analysis**
- Monitoring onset of competence in *B. subtilis*

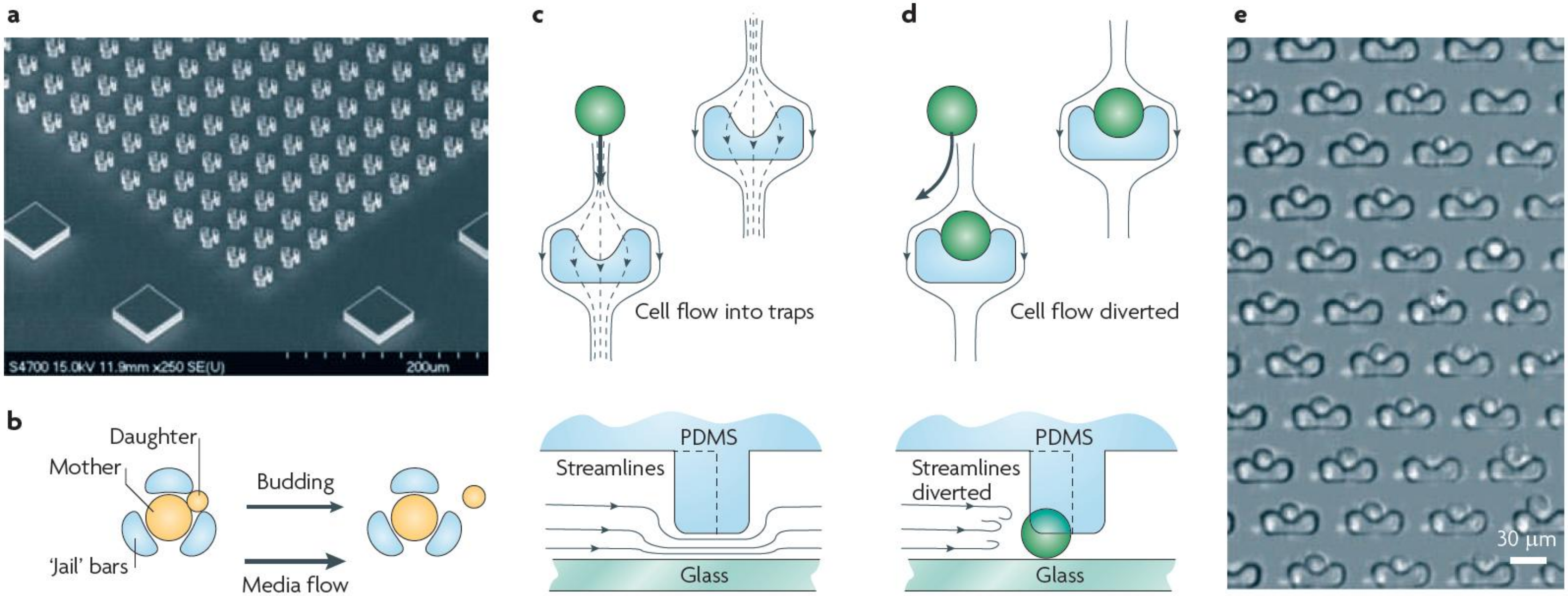
Süel *et al.* (2006), *Nature*, 440:545-50





Single-cell microscopy and microfluidics

- **Microfluidic** trapping devices for long-term acquisition of single-cell data

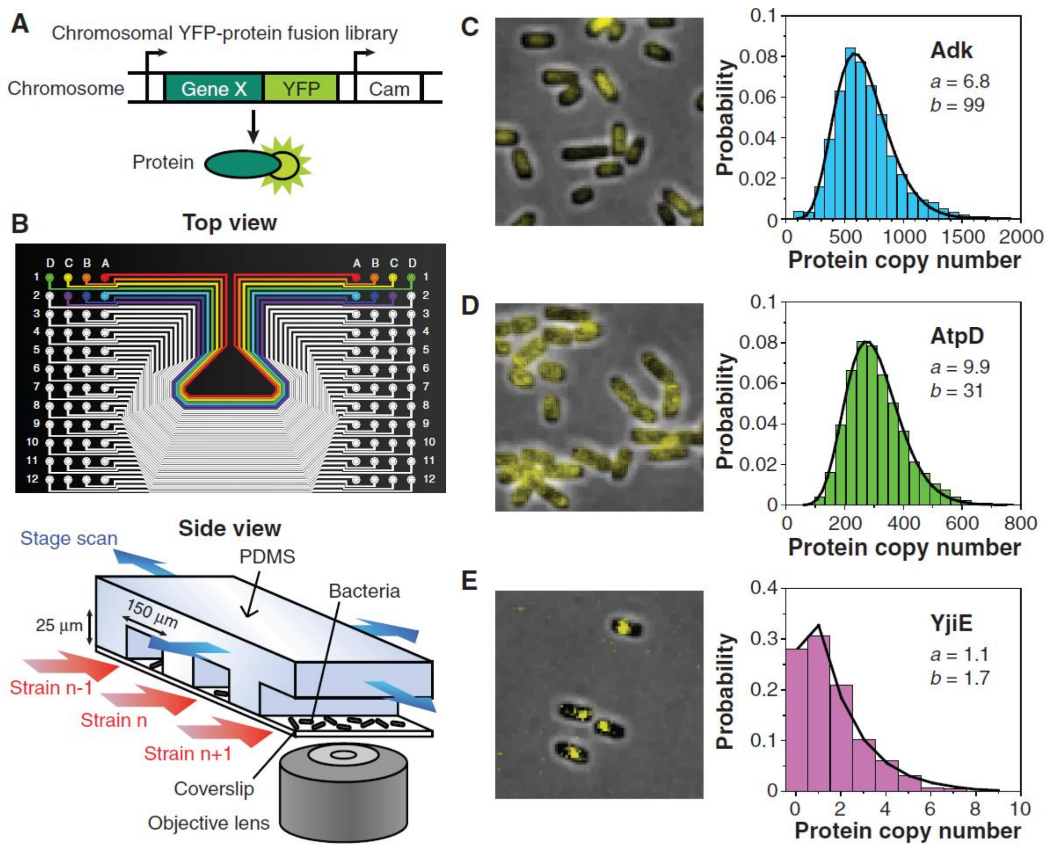


- Microfluidic devices allow tight control of environmental perturbations

Bennett and Hasty (2009), *Nat. Rev. Genet.*, 10(9):628-38

Single-molecule quantification

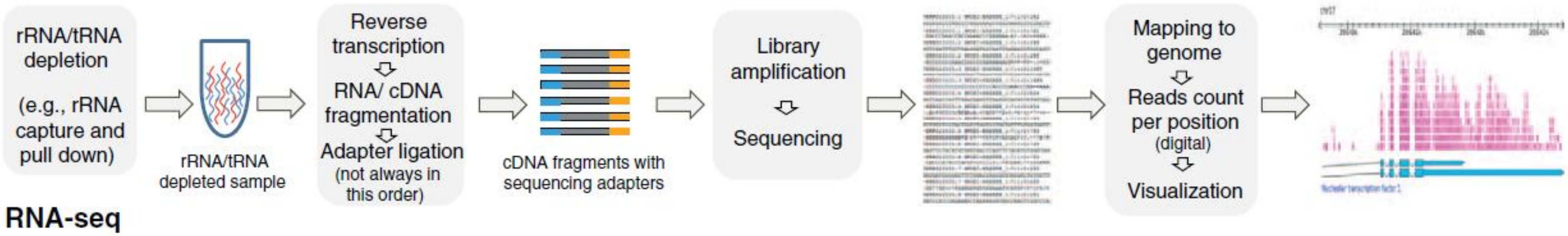
- Measurement of gene expression at **single-molecule level** using fluorescence reporter genes, microfluidic device, fluorescence microscopy, and calibration
- Measurement of expression of thousand *E. coli* genes using YFP-tagged chromosomal reporters



Taniguchi *et al.* (2010), *Science*, 329(5991):533-9

RNA sequencing

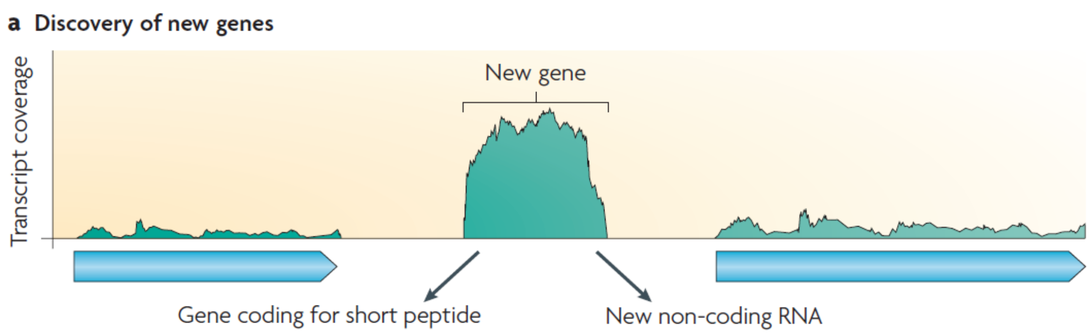
- **RNA sequencing (RNA-seq)** exploits new generation of sequencing technologies for quantifying RNA levels



Mäder *et al.* (2010), *Curr. Opin. Biotechnol.*, 22(1):32-41

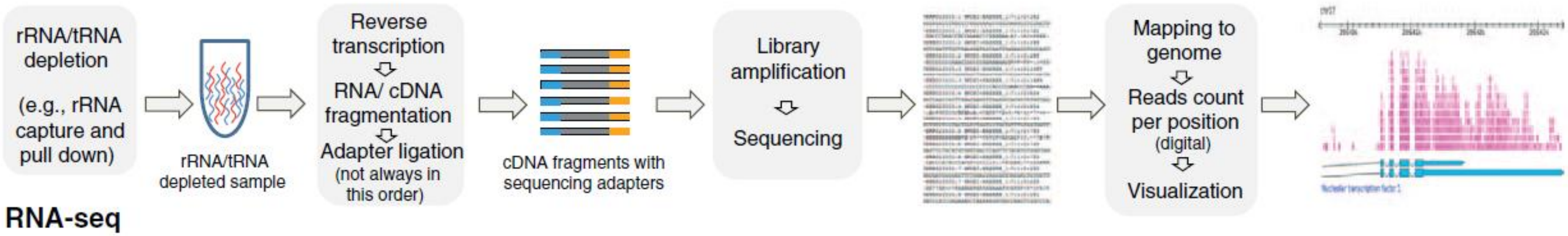
- Use of RNA-seq data to discover new genes and detect operon structure

Sorek and Cossart (2010), *Nat. Rev. Genet.*, 11(1):9-16



RNA sequencing

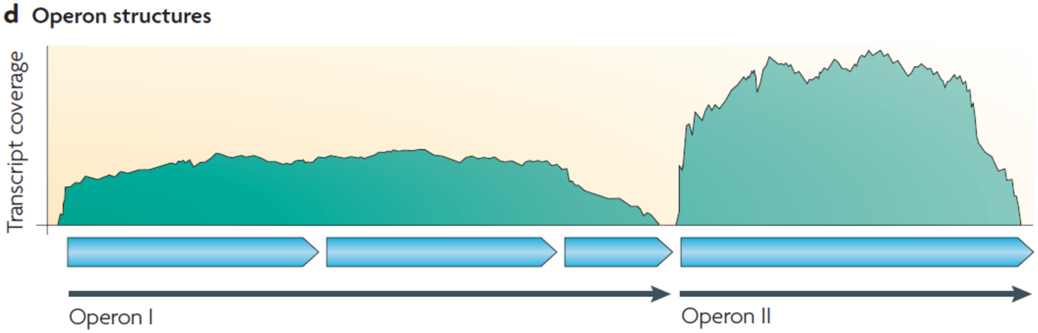
- **RNA sequencing (RNA-seq)** exploits new generation of sequencing technologies for quantifying RNA levels



Mäder *et al.* (2010), *Curr. Opin. Biotechnol.*, 22(1):32-41

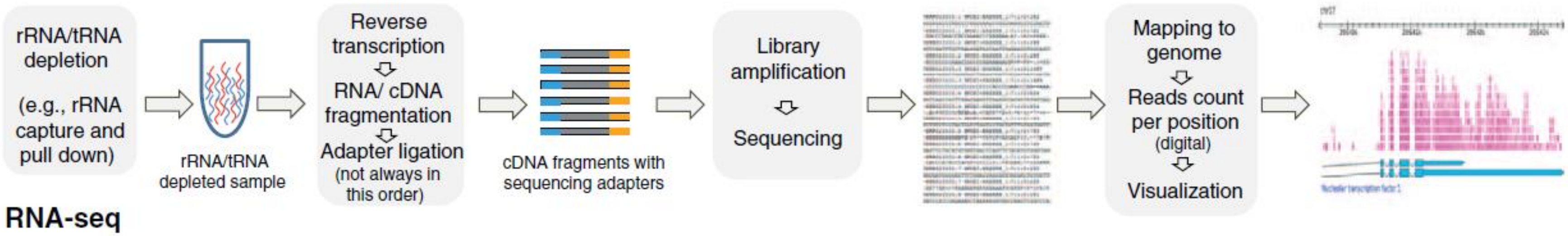
- Use of RNA-seq data to discover new genes and detect operon structure

Sorek and Cossart (2010), *Nat. Rev. Genet.*, 11(1):9-16



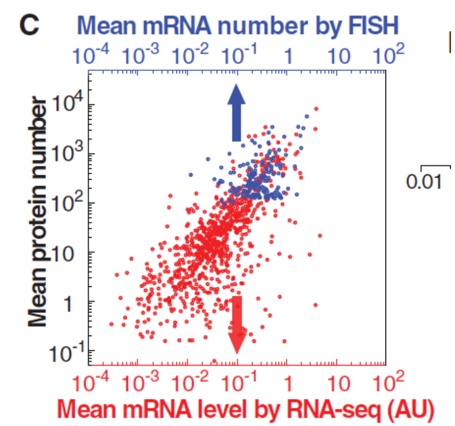
RNA sequencing

- **RNA sequencing (RNA-seq)** exploits new generation of sequencing technologies for quantifying RNA levels



Mäder *et al.* (2010), *Curr. Opin. Biotechnol.*, 22(1):32-41

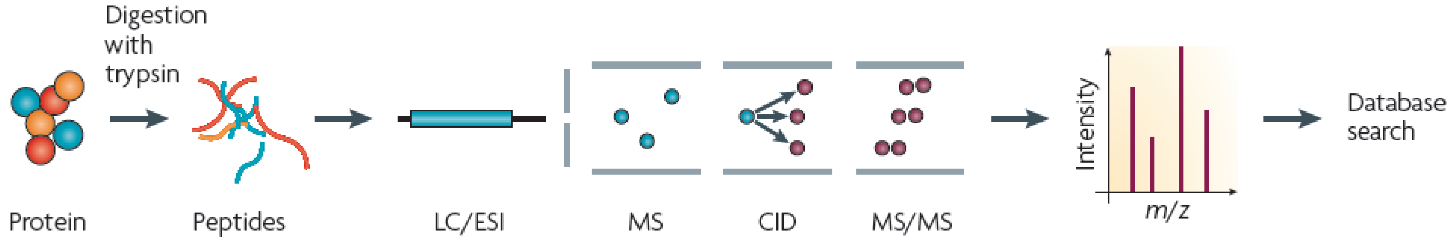
- Use of RNA-seq to determine correlation between mean RNA and mean protein levels



Taniguchi *et al.* (2010), *Science*, 329(5991):533-9

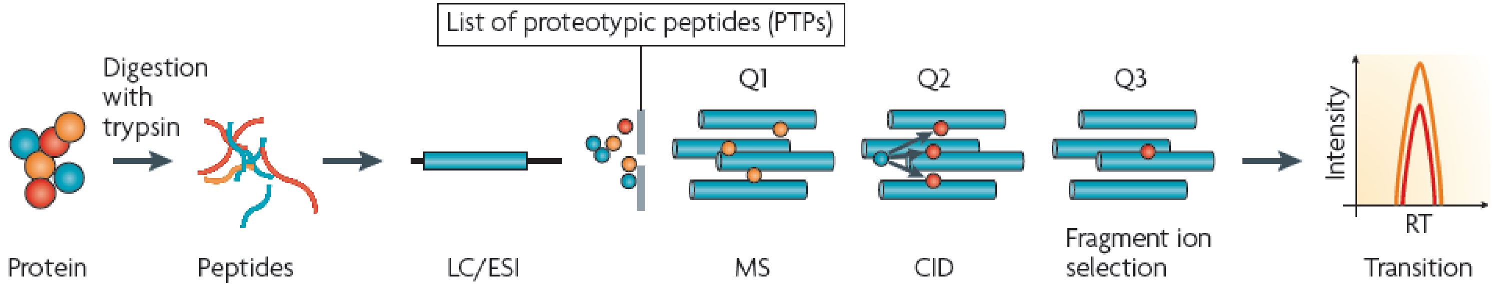
Quantitative proteomics

- Measurement of protein abundance using mass-spectrometry-based techniques (**quantitative proteomics**)
  - Use of calibration standards to achieve absolute quantification



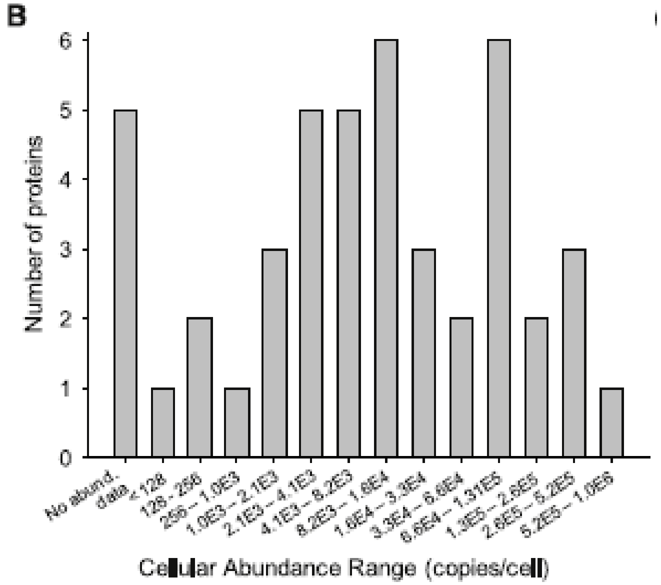
Gstaiger and Aebersold (2009), *Nat. Rev. Genet.*, 10:617-27

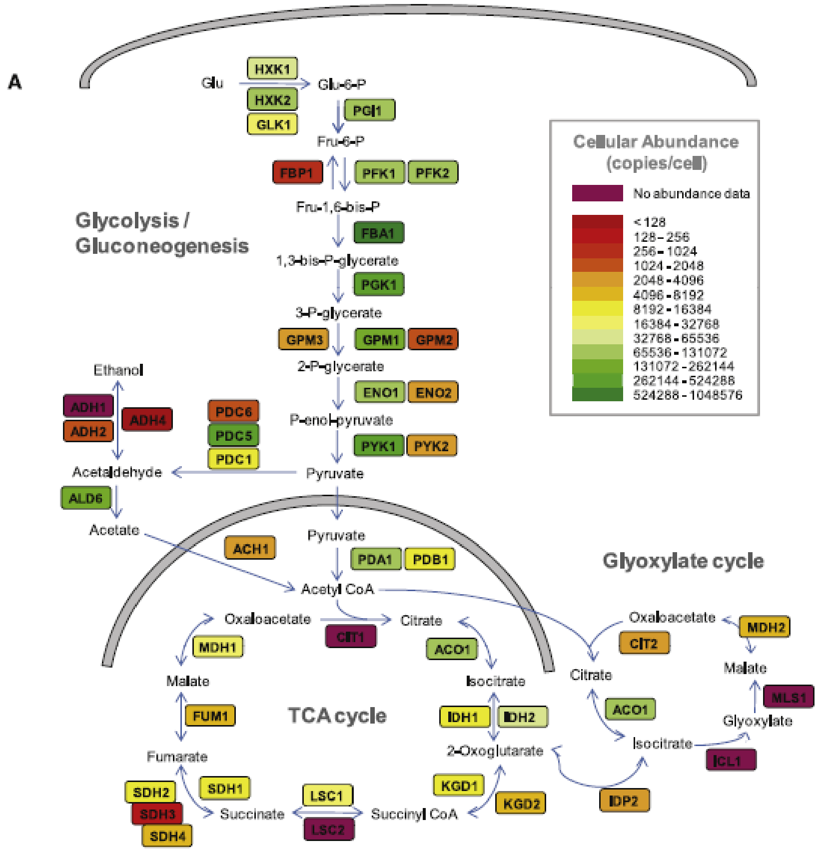
- Several targeted proteomics techniques developed to improve quantification of low-abundance proteins



Quantitative proteomics

- Absolute quantification of proteins in yeast carbon metabolism by means of **selected reaction monitoring (SRM)**





Picotti *et al.* (2009), *Cell*, 138:795-806

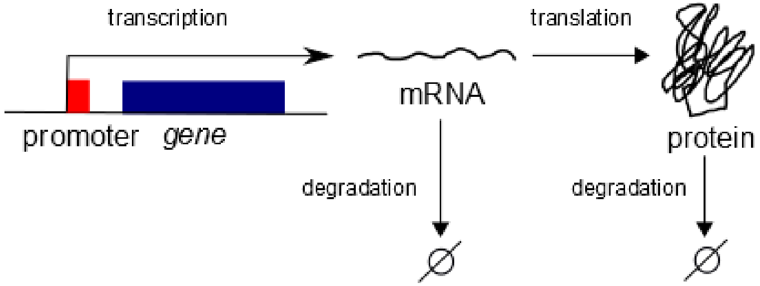
Systems biology

- Most gene regulatory networks of biological interest are large and complex
- **No global view of functioning** of network available, despite abundant knowledge on network components
  - Understanding of dynamics requires **experimental tools** for monitoring gene expression over time
  - Understanding of dynamics requires **mathematical modeling** and **computer analysis and simulation**
  - Discipline now often referred to as **systems biology**

Alon (2007), *An Introduction to Systems Biology,* Chapman & Hall/CRC Press

Modeling of gene regulatory networks

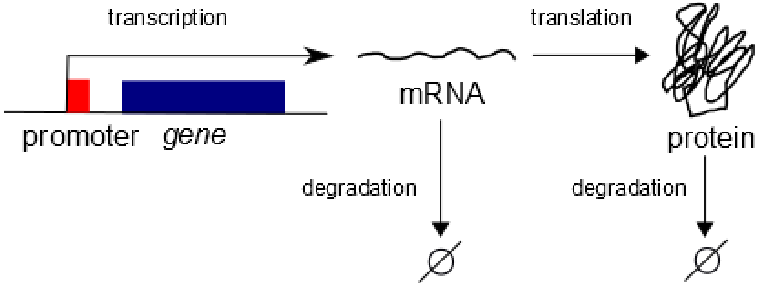
- Modeling of gene regulatory networks amount to modeling of gene expression and regulation of gene expression



- Possible **aims** of modeling of gene regulatory networks:
  - Understanding role of individual components and interactions
  - Suggesting missing components and interactions
- Advantages of **mathematical and computer tools**:
  - Precise and unambiguous description of network
  - Systematic derivation of predictions of network behavior

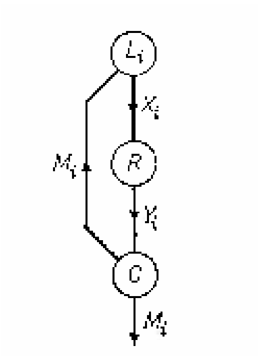
Modeling of gene regulatory networks

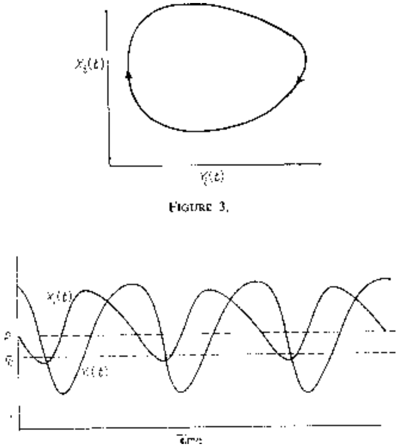
- Modeling of gene regulatory networks amount to modeling of gene expression and regulation of gene expression



- First models of gene regulatory networks date back to early days of molecular biology

    Feedback circuits and oscillators





Goodwin (1963), *Temporal Organization in Cells*

Modeling of gene regulatory networks

- Different modeling formalisms exist, describing gene expression on different levels of detail

detailed ⟷ coarse-grained

Stochastic master equations

Ordinary differential equations (ODEs)

Boolean networks

Smolen *et al.* (2000), *Bull. Math. Biol*., 62(2):247-292

Hasty *et al.* (2001), *Nat. Rev. Genet.*, 2(4):268-279

de Jong (2002), *J. Comput. Biol.*, 9(1): 69-105

Szallassi *et al.* (2006), *System Modeling in Cellular Biology*, MIT Press

Bolouri (2008), *Computational Modeling of Gene Regulatory Networks*, Imperial College Press

Karleback and Shamir (2008), *Nat. Rev. Mol. Cell Biol.*, 9(10):770-80

Conclusions

- Gene regulatory networks control adaptive response of bacteria to changes in environment
- Gene regulatory networks are intertwined with metabolic and signaling networks
- Technology for measuring gene expression over time, and thus functioning of gene regulatory networks, are rapidly developing
- Modeling necessary for understanding dynamics of complex networks: systems biology
- A variety of formalisms for modeling gene regulatory networks, in a detailed or coarse-grained way, have been developed

Merci

www.inrialpes.fr/ibis