Transcription control in eucaryotes is complex:

- **Eukaryotic RNA-polymerase needs „general transcription factors“**
- **Eukaryotic includes promotor plus regulative DNA sequences**
- **Enhancer elements regulate genes in distance**

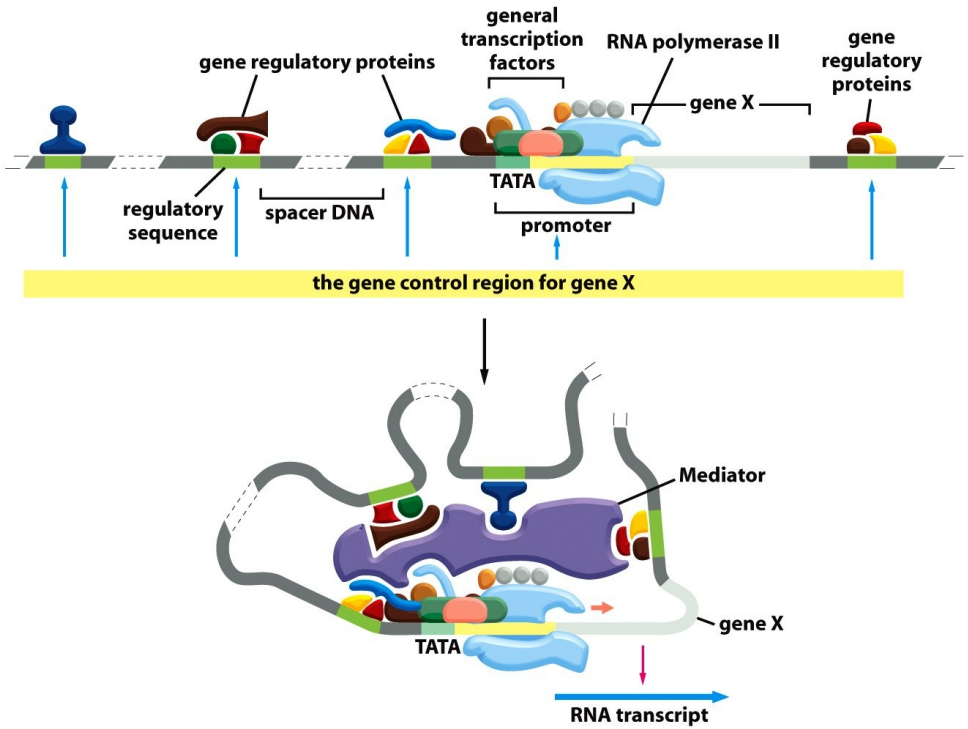
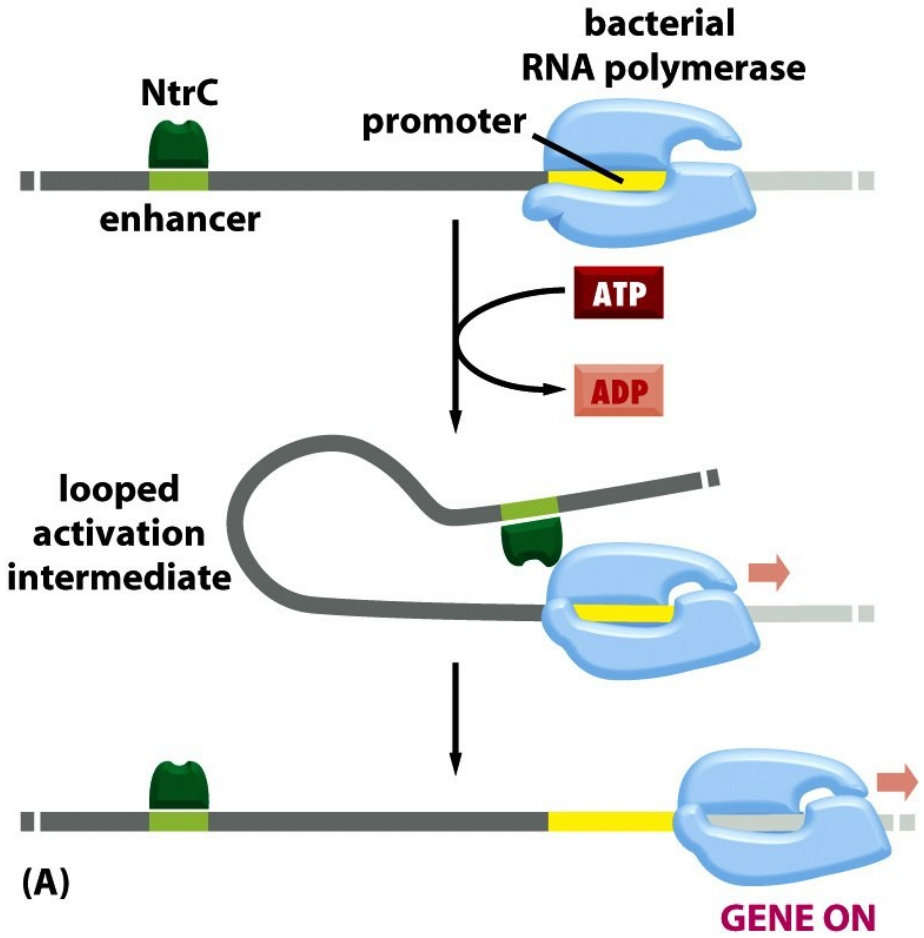


Figure 7-44 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Bacterial transcription is comparably simpler
However: Enhancer work on distance

W. Su et al PNAS (1990)



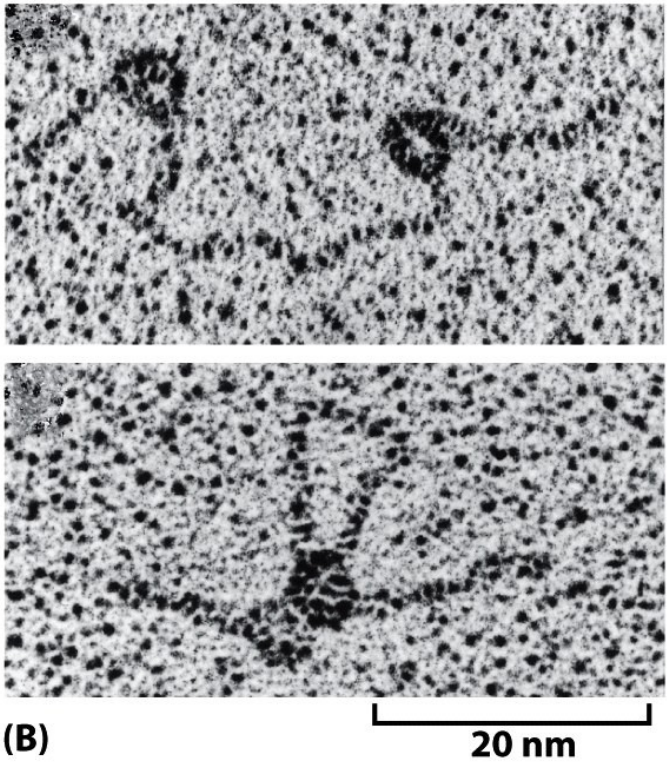
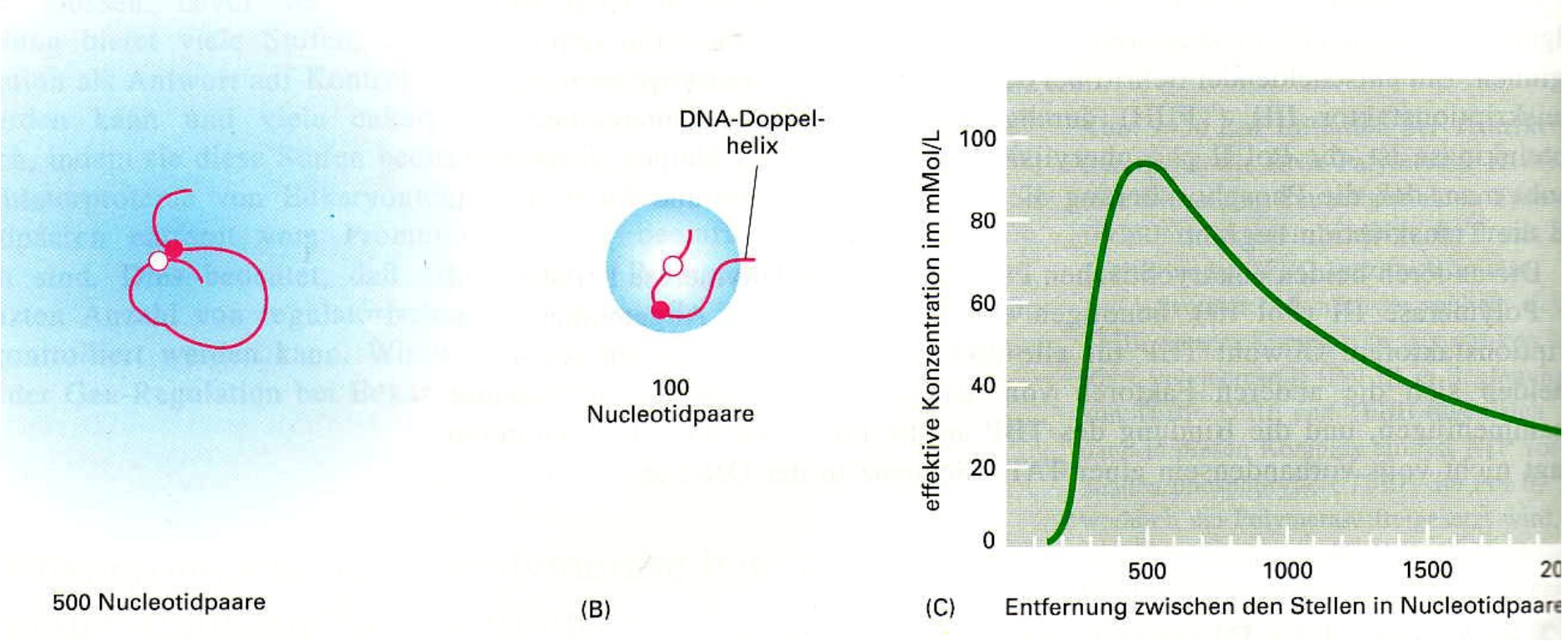


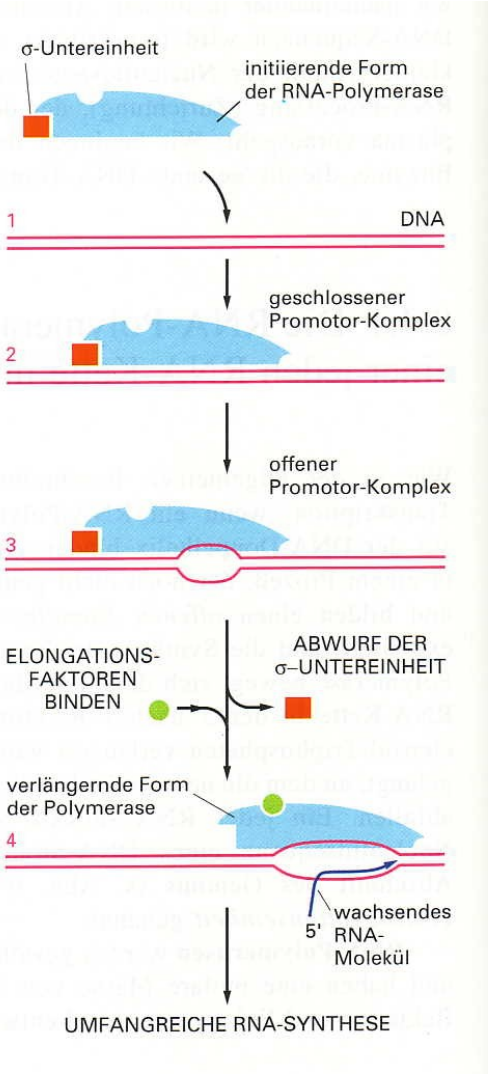
Figure 7-42 Molecular Biology of the Cell 5/e (© Garland Science 2008)

Loop formation increases interactions

Van Hippel



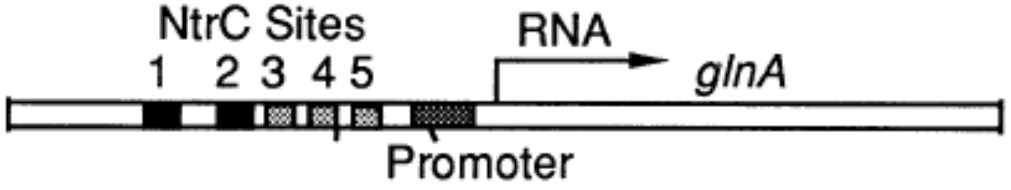
Example: NtrC (nitrogen regulatory Protein C)

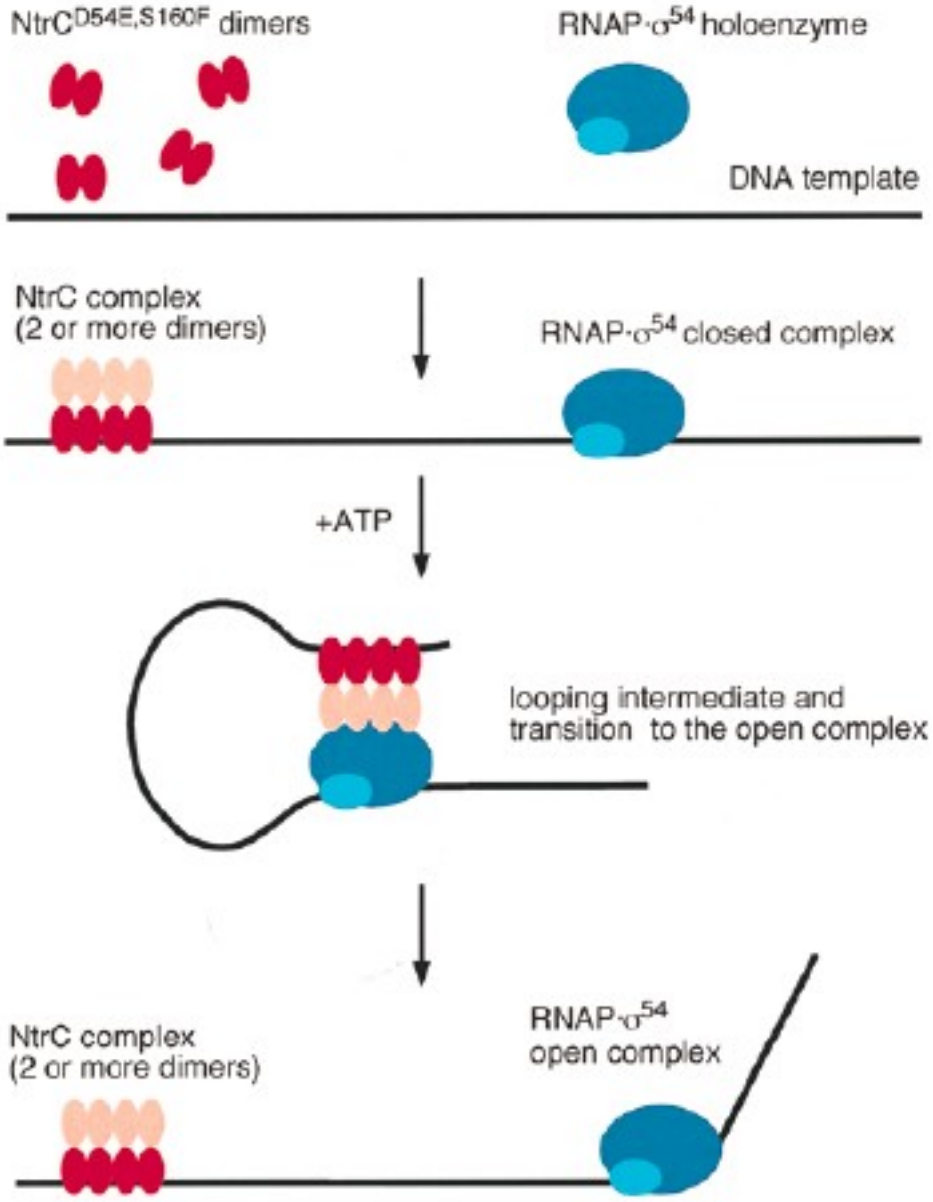


from enteric bacteria :

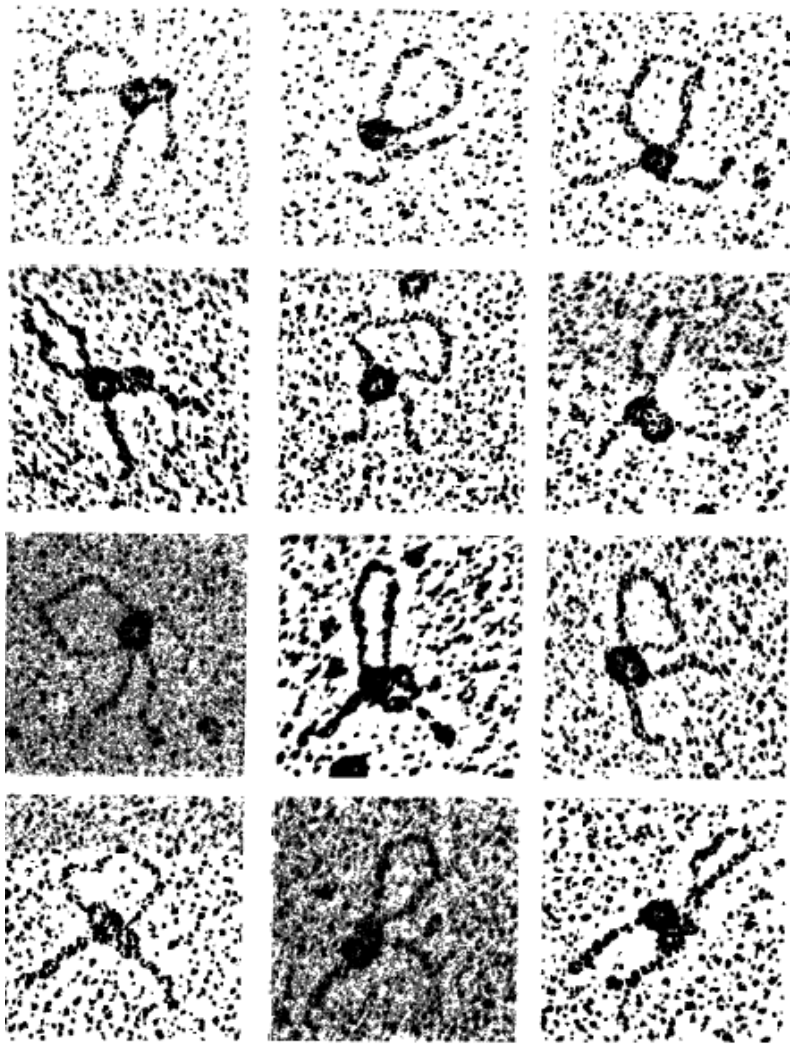
a transcription factor that activates a variety of genes that are involved in nitrogen utilization by contacting simultaneously a binding site on the DNA and RNA polymerase complexed with the σ54 sigma factor at the promoter.

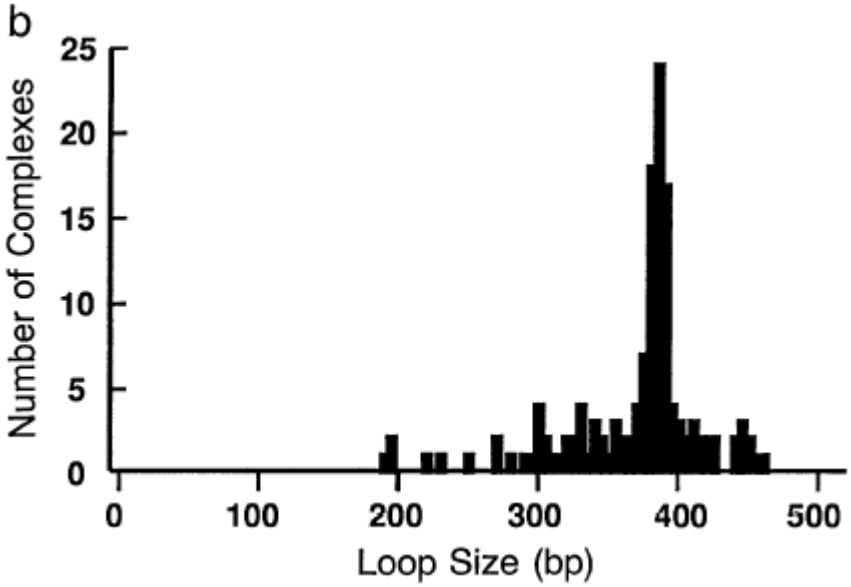
wild type *glnA* fragment:




NtrC^{D54E,S160F} dimers
RNAP·σ^{54} holoenzyme
DNA template
NtrC complex
(2 or more dimers)
RNAP·σ^{54} closed complex
+ATP
looping intermediate and
transition to the open complex
NtrC complex
(2 or more dimers)
RNAP·σ^{54}
open complex

Distribution of DNA loops formed of NtrC and Pol



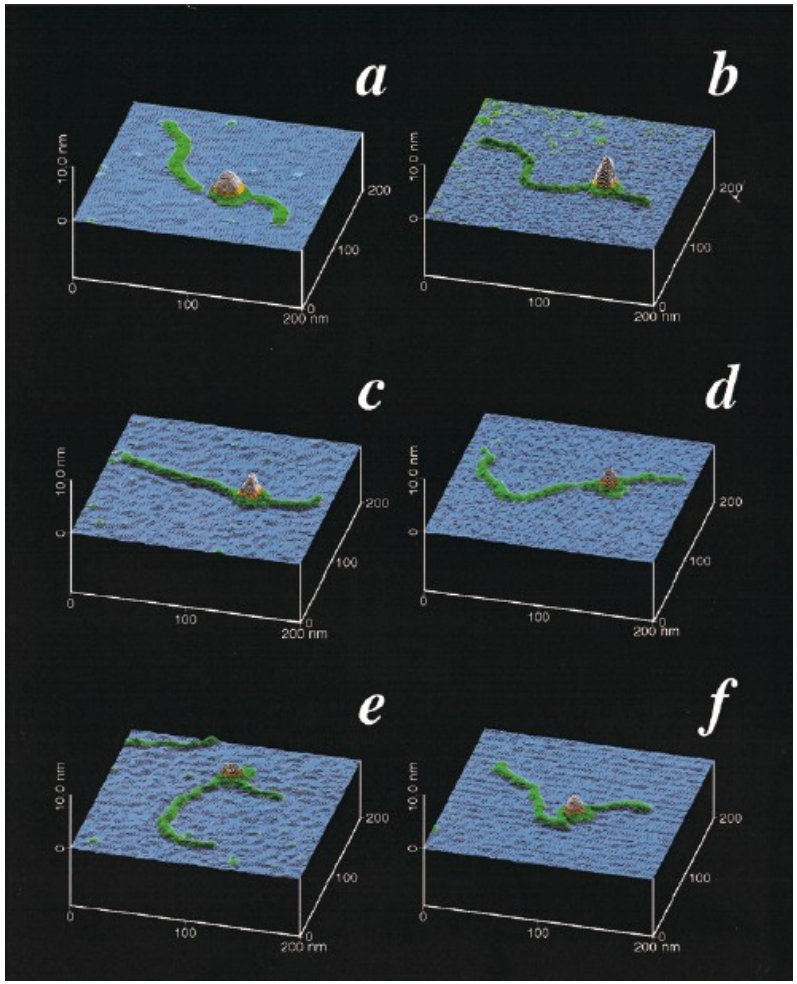


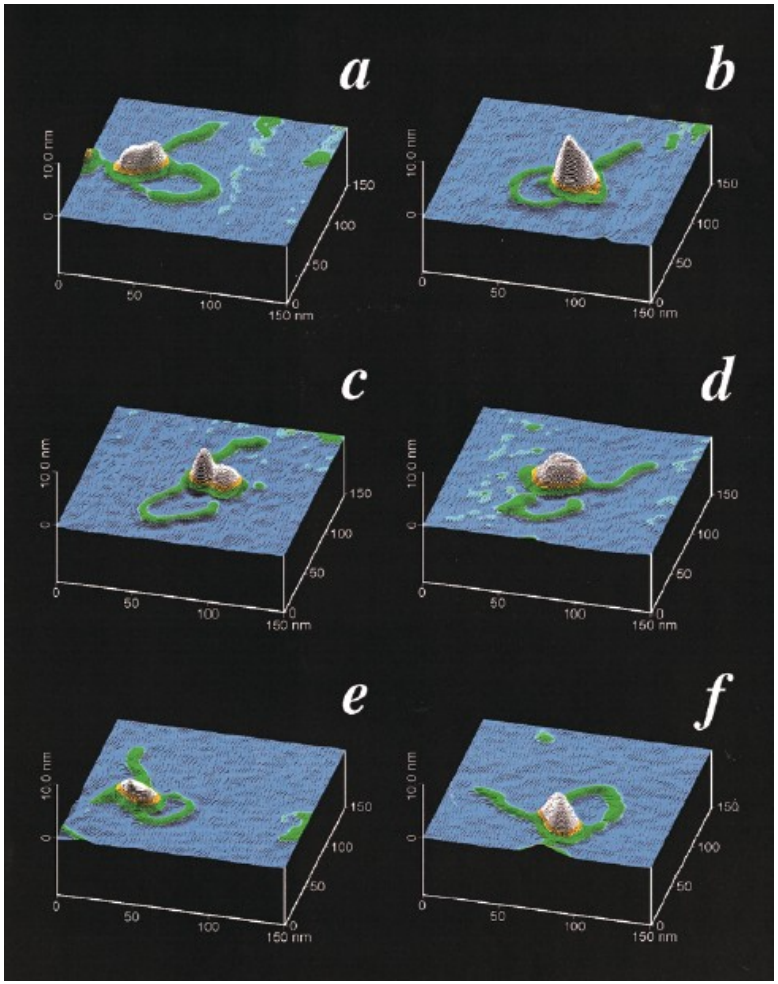
W. Su et al PNAS (1990)

Transcriptional Activation *via* DNA-looping: Visualization of Intermediates in the Activation Pathway of *E. coli* RNA Polymerase·σ^{54} Holoenzyme by Scanning Force Microscopy

Karsten Rippe[1*], Martin Guthold[2], Peter H. von Hippel[3] and Carlos Bustamante[4*]

J. Mol. Biol. (1997) 270, 125-138





Analysis of high throughput gene expression

Automated Discovery System

The Genome Project was the first

inherently digital, 1-dimensional, static
small (fits on one CD-ROM)

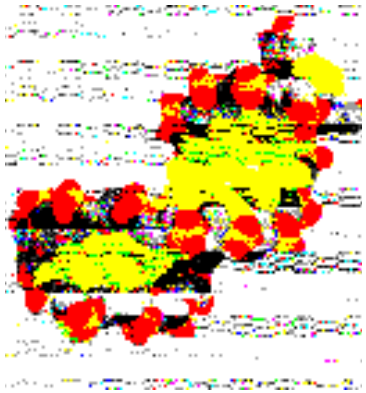
The "gene expression project"

clustering analysis yields "correlations" among genes
limited scope to infer causality from mRNA analysis

The genome and the proteome : a comparison

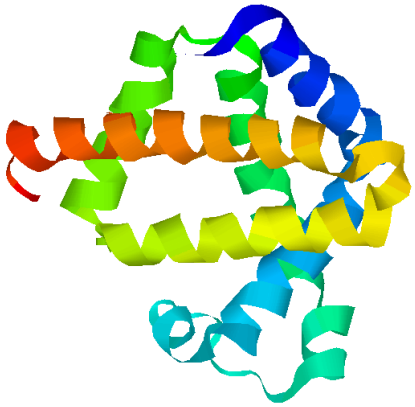
Genome

- static
- amplification possible (PCR)
- homogeneous
- no variability in amount



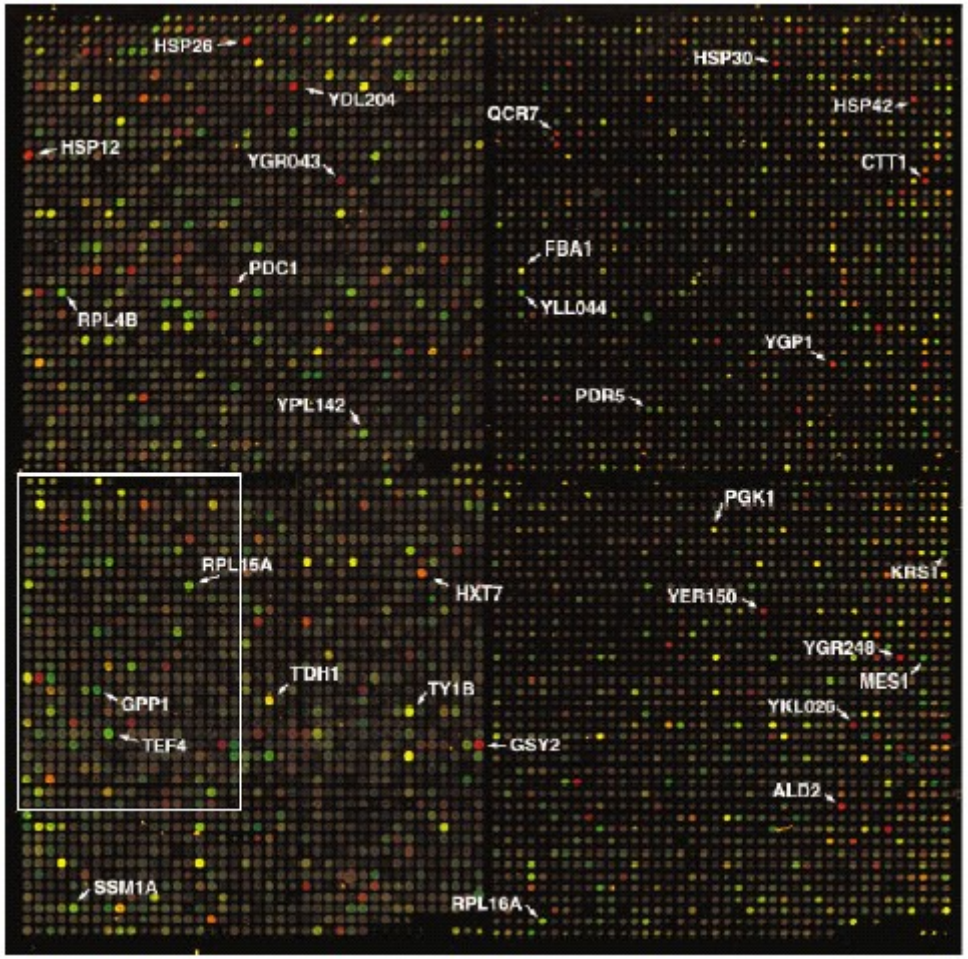
Proteome

- dynamic - condition dependent
- no amplification
- non-homogeneous
- high variablity in amount ($>10^6$)



The full yeast genome on a chip

Science *DeRisi et al. 278 (5338): 680*



Exploring the Metabolic and Genetic Control of Gene Expression on a Genomic Scale

Yeast genome microarray. The actual size of the microarray is
18 mm by 18 mm.

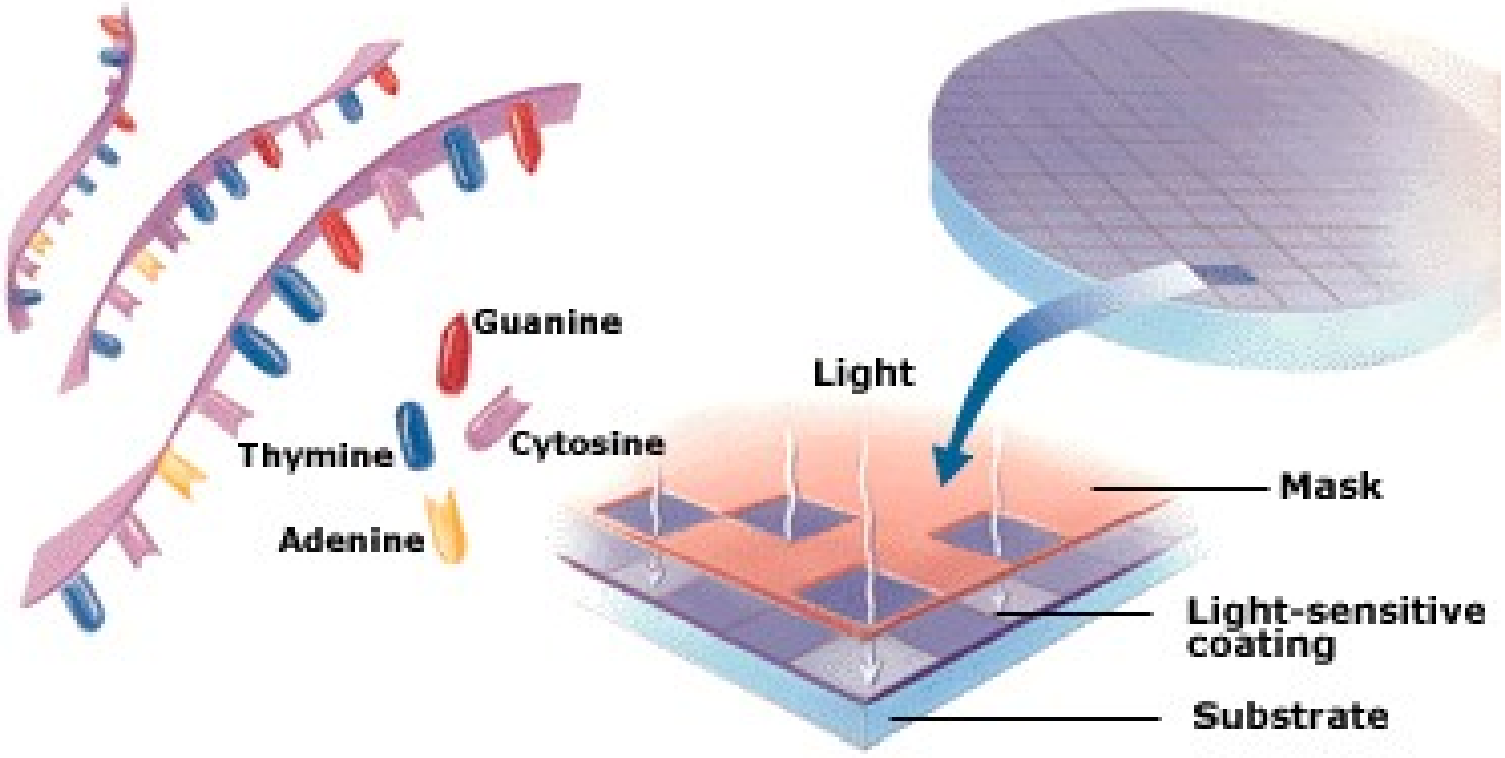
high-density arrays of oligonucleotides

Macroarrays : Pin spotted cDNAs or PCR products
on membranes, readout by radiation

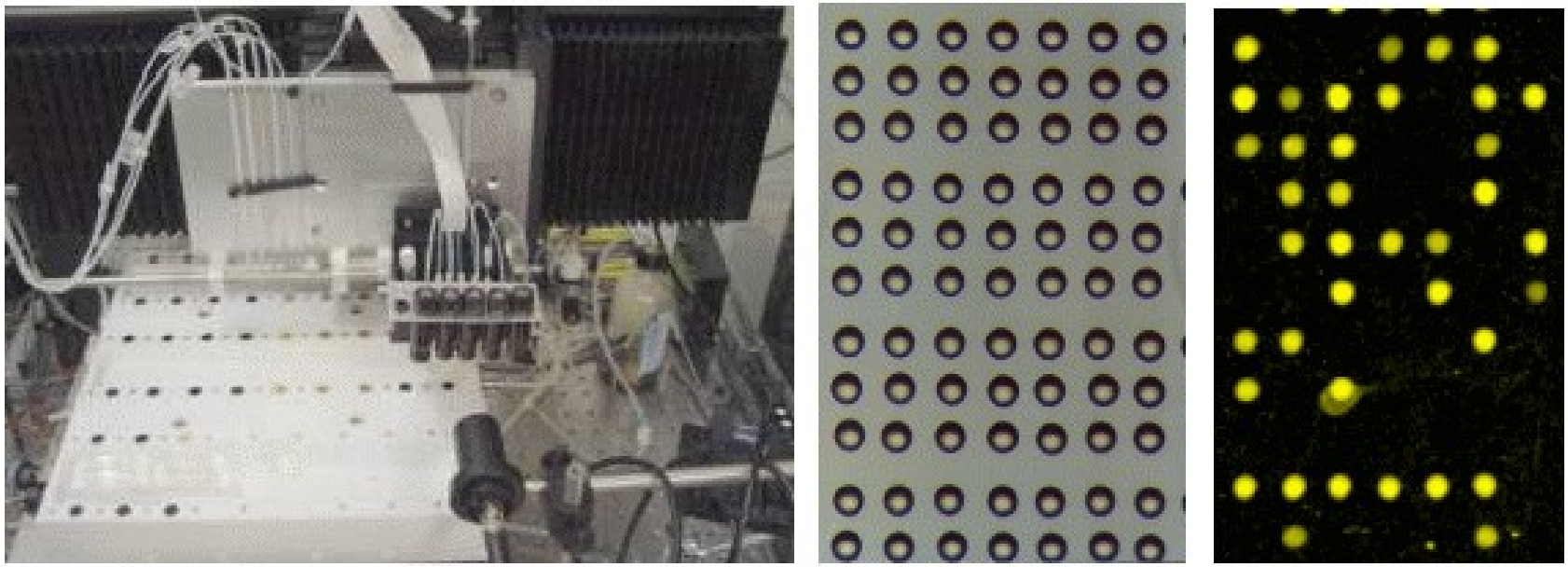
Microarrays : Pin spotted cDNAs or PCR products
on high density non-porous substrates
readout by high resolution fluorescence

microarrays allow study of gene expression
in a massively parallel way

How DNA Chips Are Made



ink-jet arrayer



The ISB "inkjet arrayer" (left) uses 192 piezoelectric nozzles to deposit picoliters of DNA monomer solution on to reactive glass slides. Each droplet, only 150 microns in diameter, is an individual reaction chamber for the synthesis of unique oligonucleotide (middle). When synthesis is completed, a DNA microarray can be hybridized with a fluorescence-labeled biological sample (right).

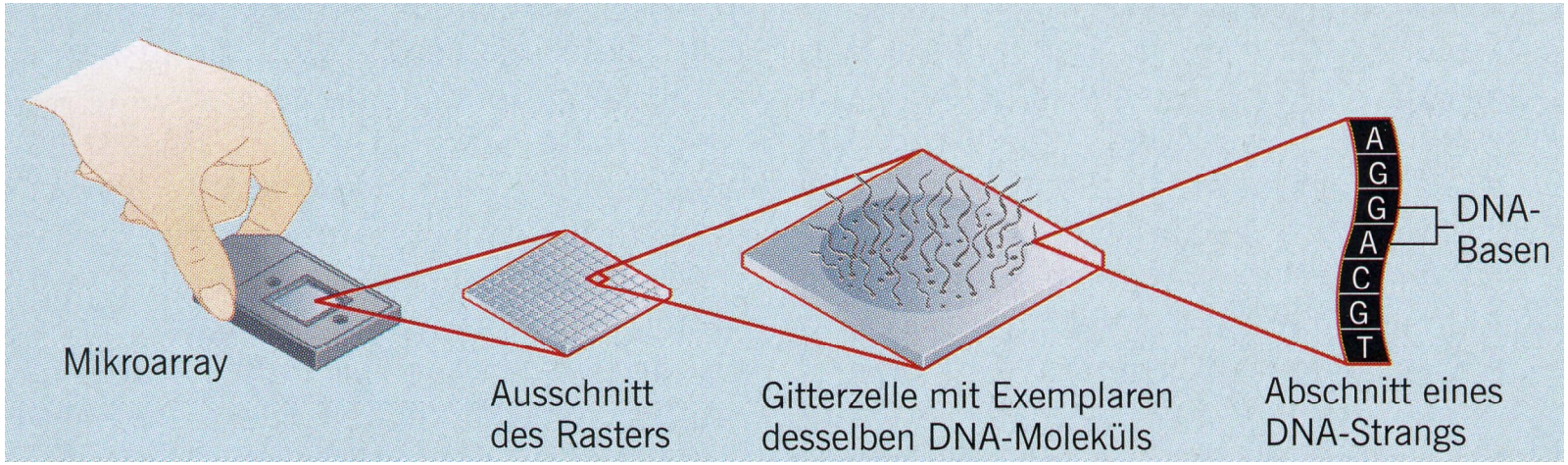
Reactive agent tests DNA-Chips

(Expression profiling)

Question: Does a reactive agent harm the liver?

Howto: Compare genes, that are activated by the new agent with genes activated by substances that are known to harm the liver

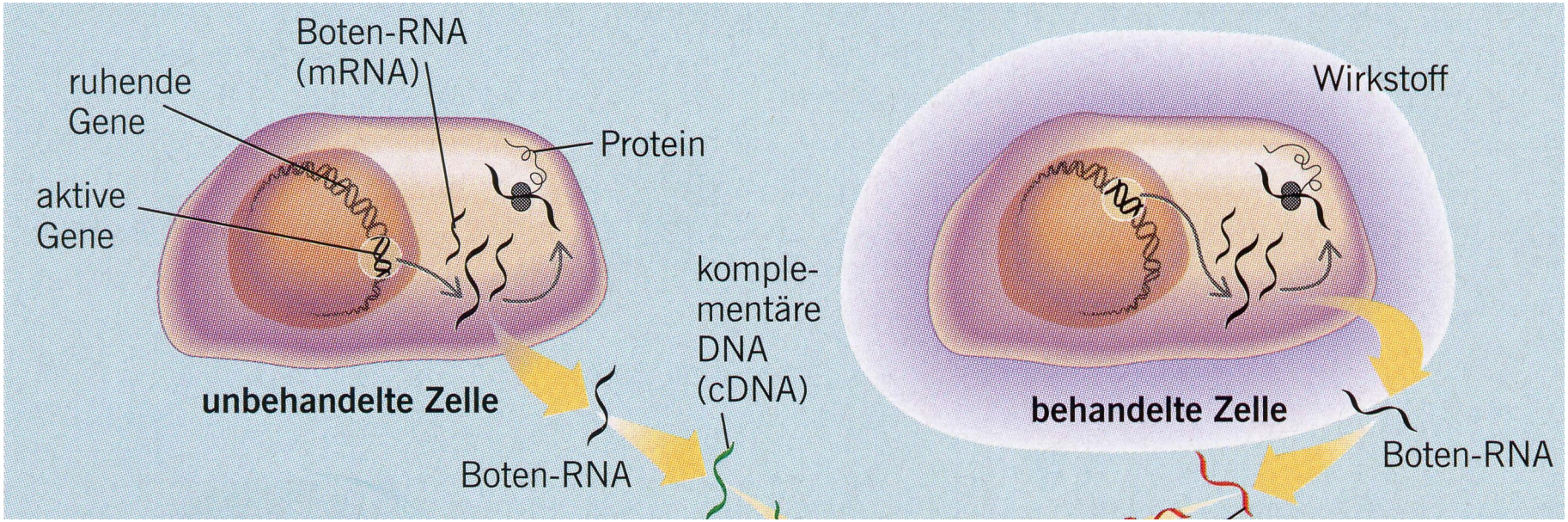
Technique: Chip, that is covered with different single strand DNA molecules in a chessboard manner (Mikroarray)



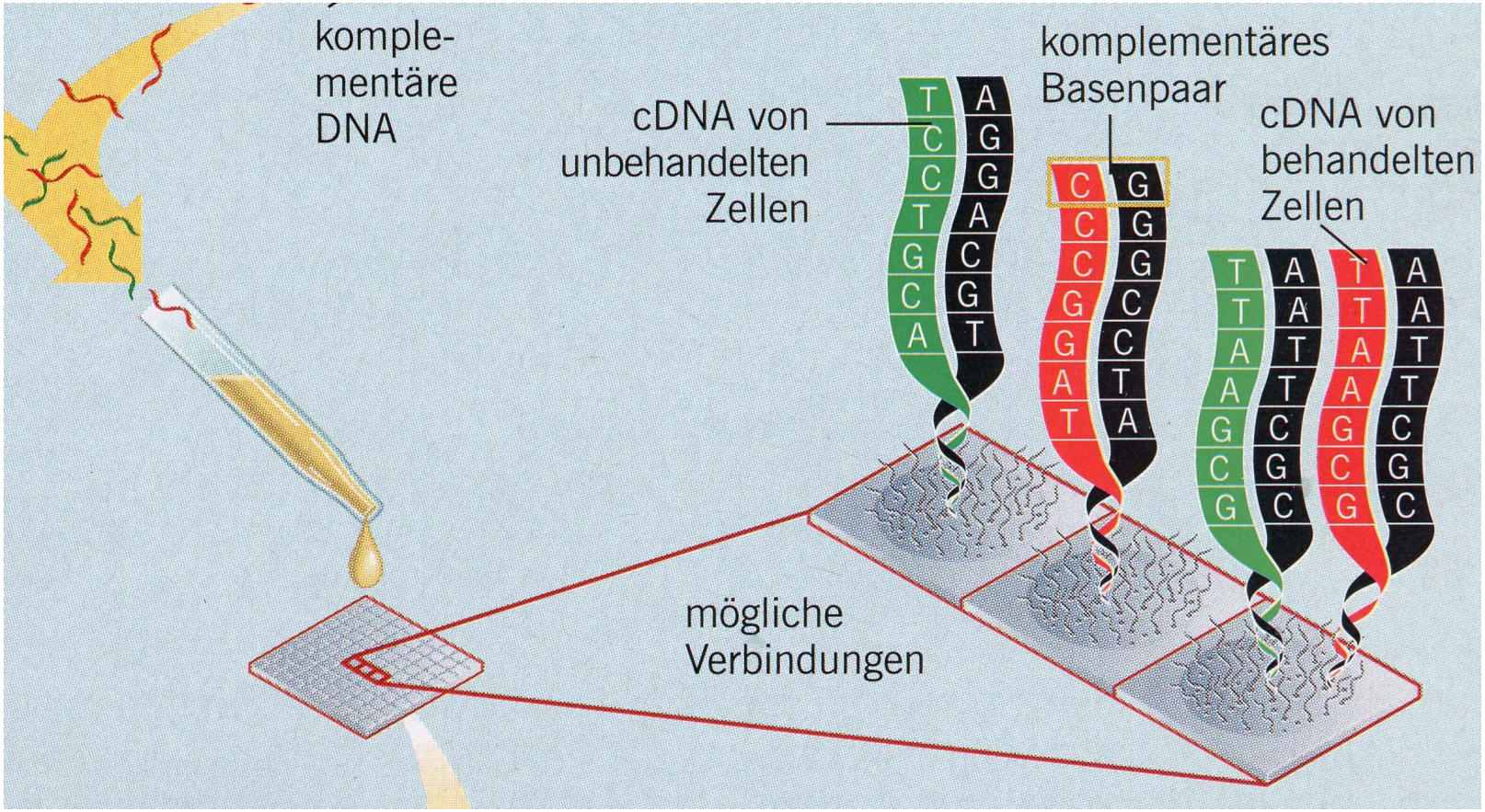
Protocol:

1.) Treat liver cells with the new agent, collect mRNA of this cells and mRNA of untreated cells Hint: Cells will mostly produce mRNA necessary to react on the new agent!

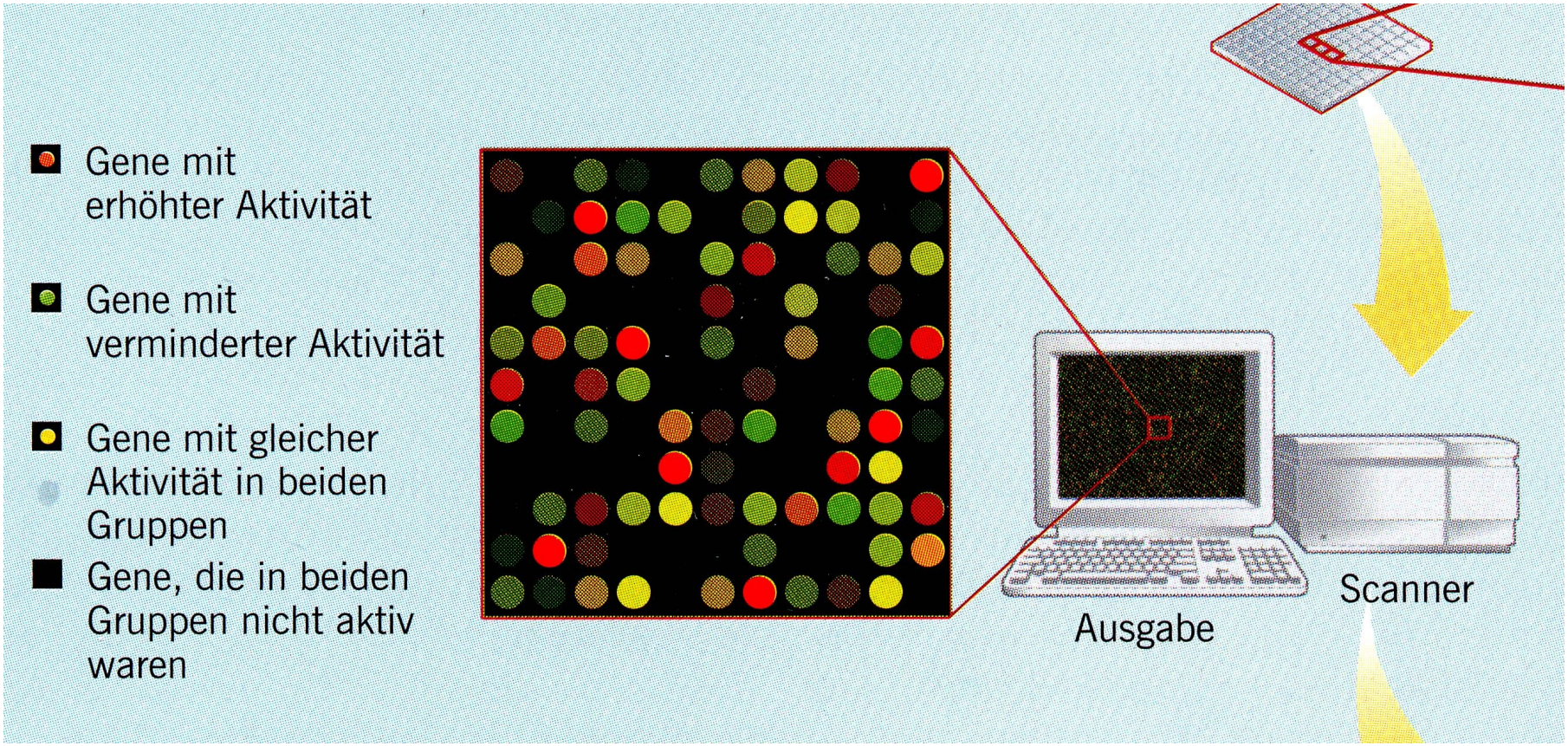
2.) Make new single stranded c-DNA complementary to both types of mRNA and dyed with different Fluorophores



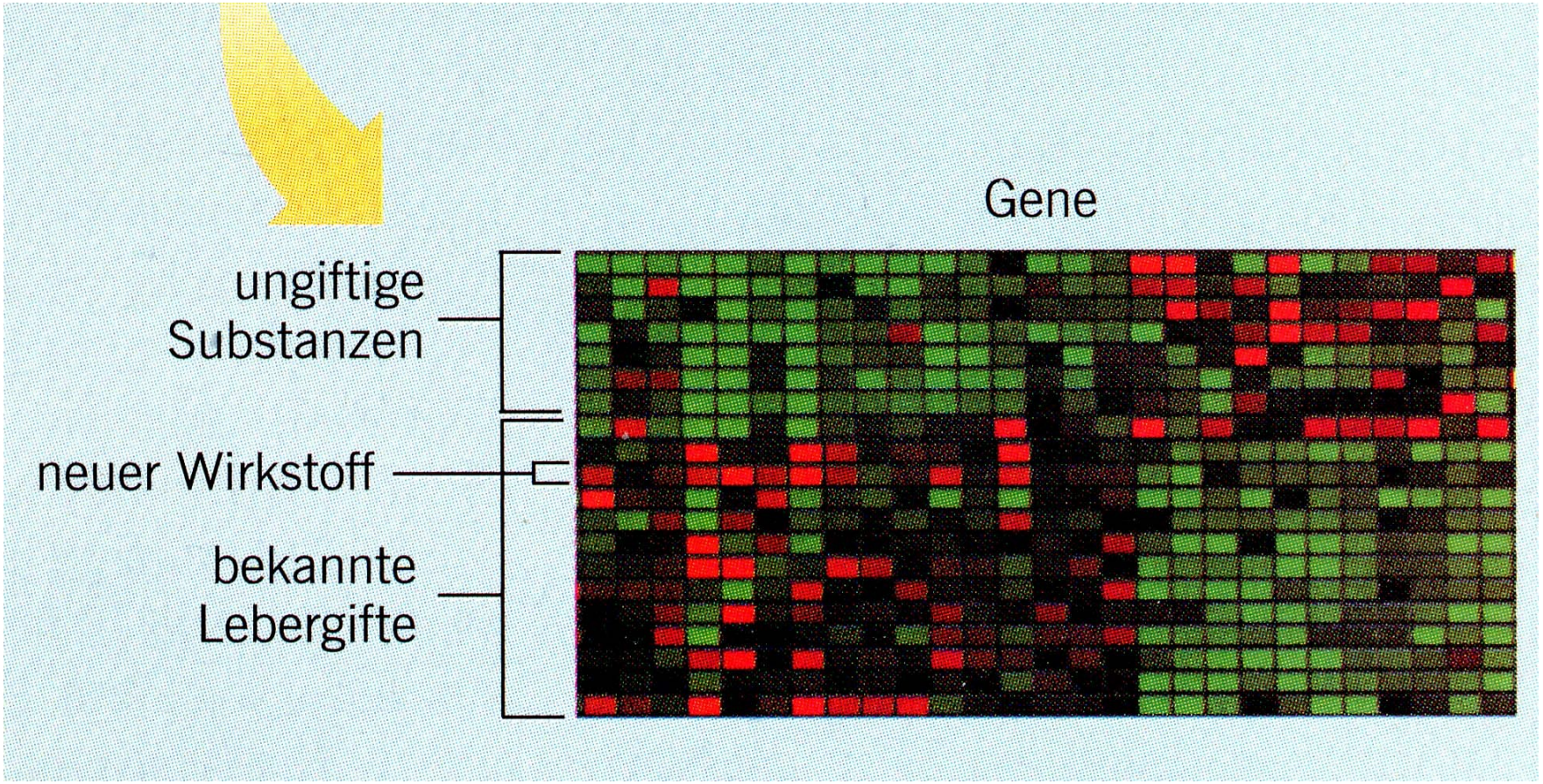
3.) c-DNA is brought to the chip and hybridizes to the complementary strands on the chip



4.) A scanner reads the fluorescence of the points (binding pattern) Now you have a „fingerprint“ of the new agent.



5.) The new binding pattern is compared to the binding pattern of all known agents:



The significance of expression data

„Fold-change“ Analyse:

x_i: Probe, y_i: Reference

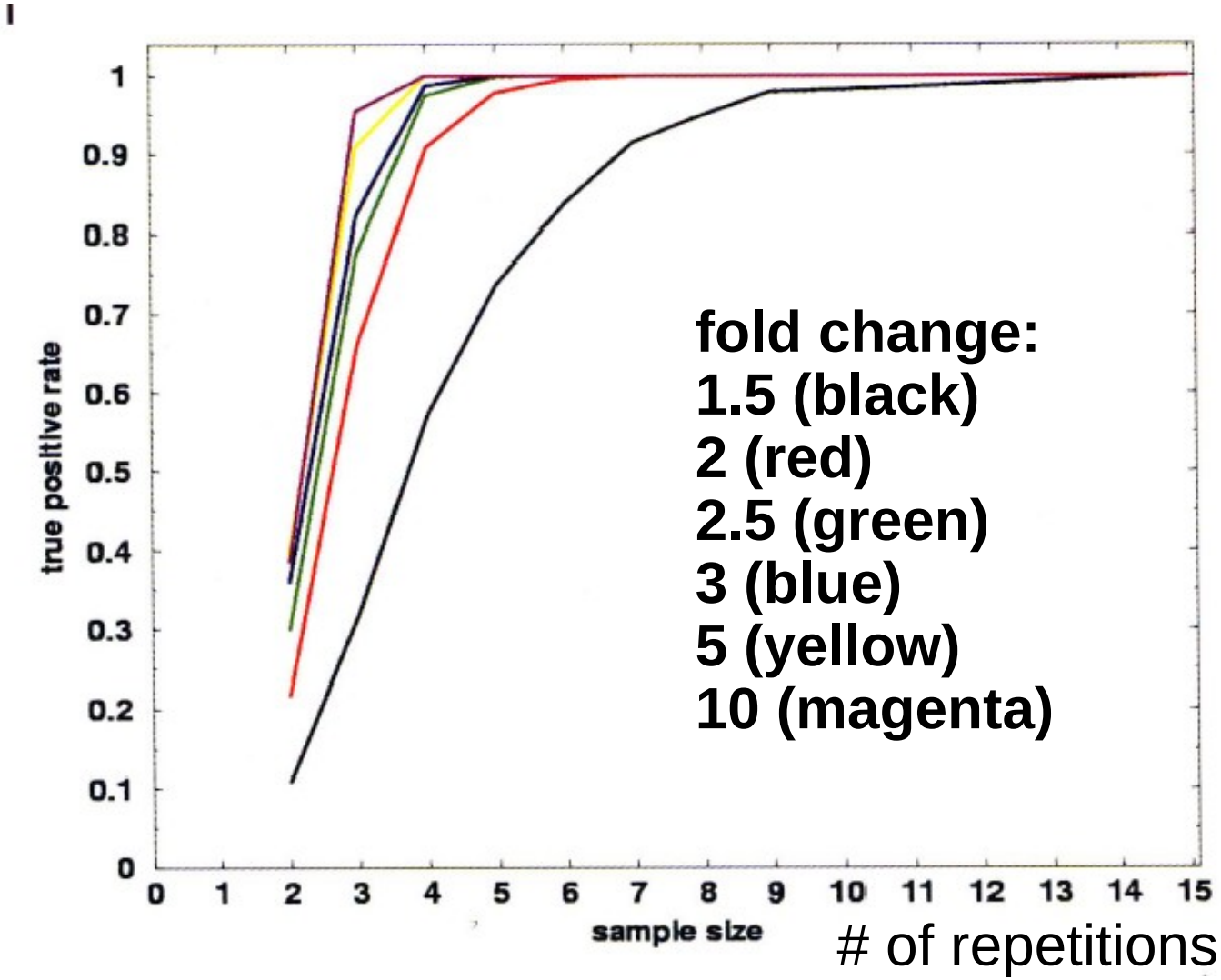
Standard deviation:

$$S_x = \sqrt{\frac{1}{(n-1)n}\sum_{i=1}^{n}\left(x_i - \bar{x}\right)^2} \qquad S_y = \sqrt{\frac{1}{(m-1)m}\sum_{i=1}^{m}\left(y_i - \bar{y}\right)^2}$$

Standard deviation of ratio

$$\frac{\bar{x}}{\bar{y}} \pm \frac{1}{\bar{y}^2}\sqrt{\bar{x}^2 S_y + \bar{y}^2 S_x}$$

Simulation of property for a correct detection



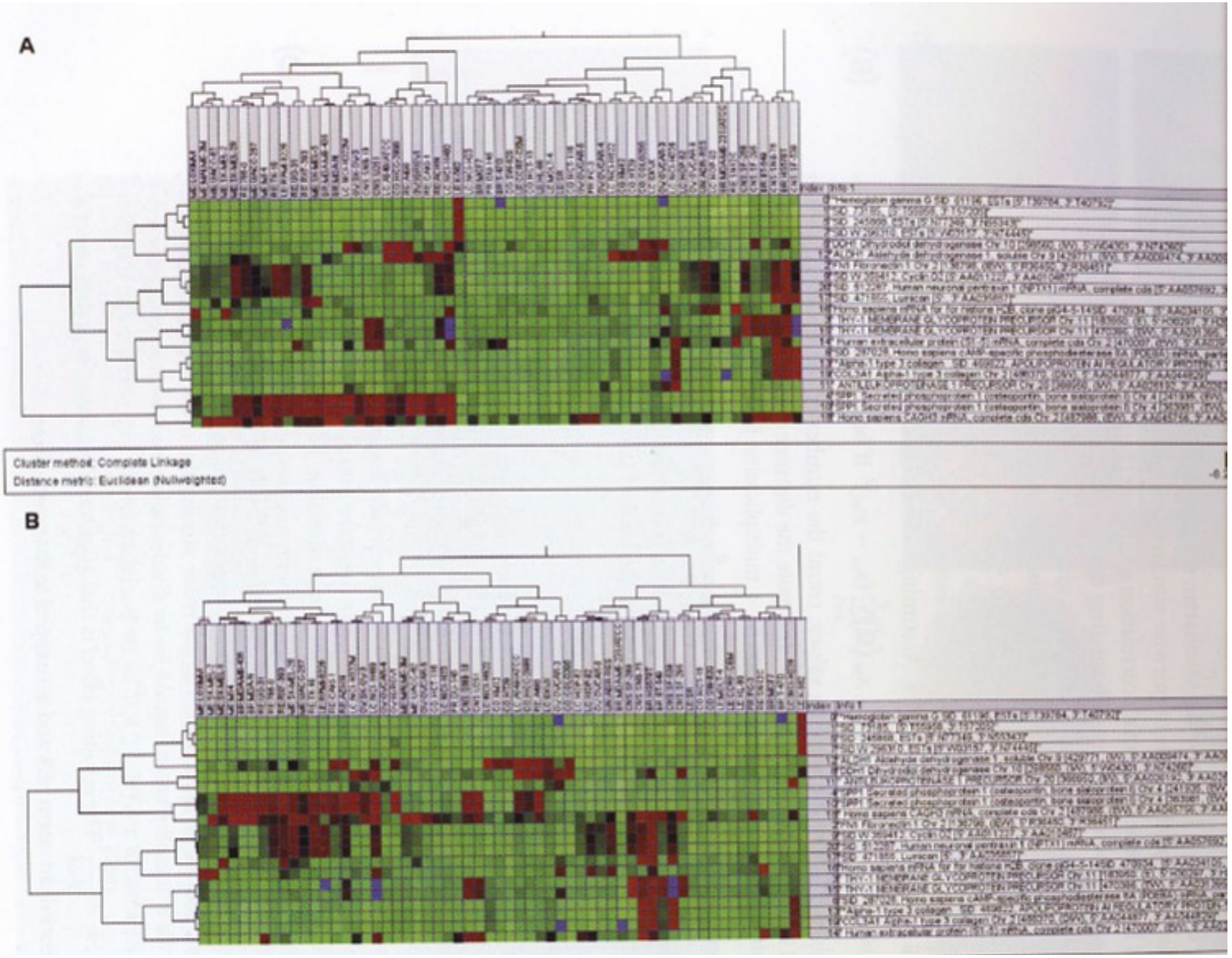
# of repetitions

Clusteranalysis

Similaryties of expressions are defined as „distances“ in expression space

$$d_q(x_n, x_m) = \left(\sum_{i=1}^{p} |x_{ni} - x_{mi}|^q \right)^{1/q}$$

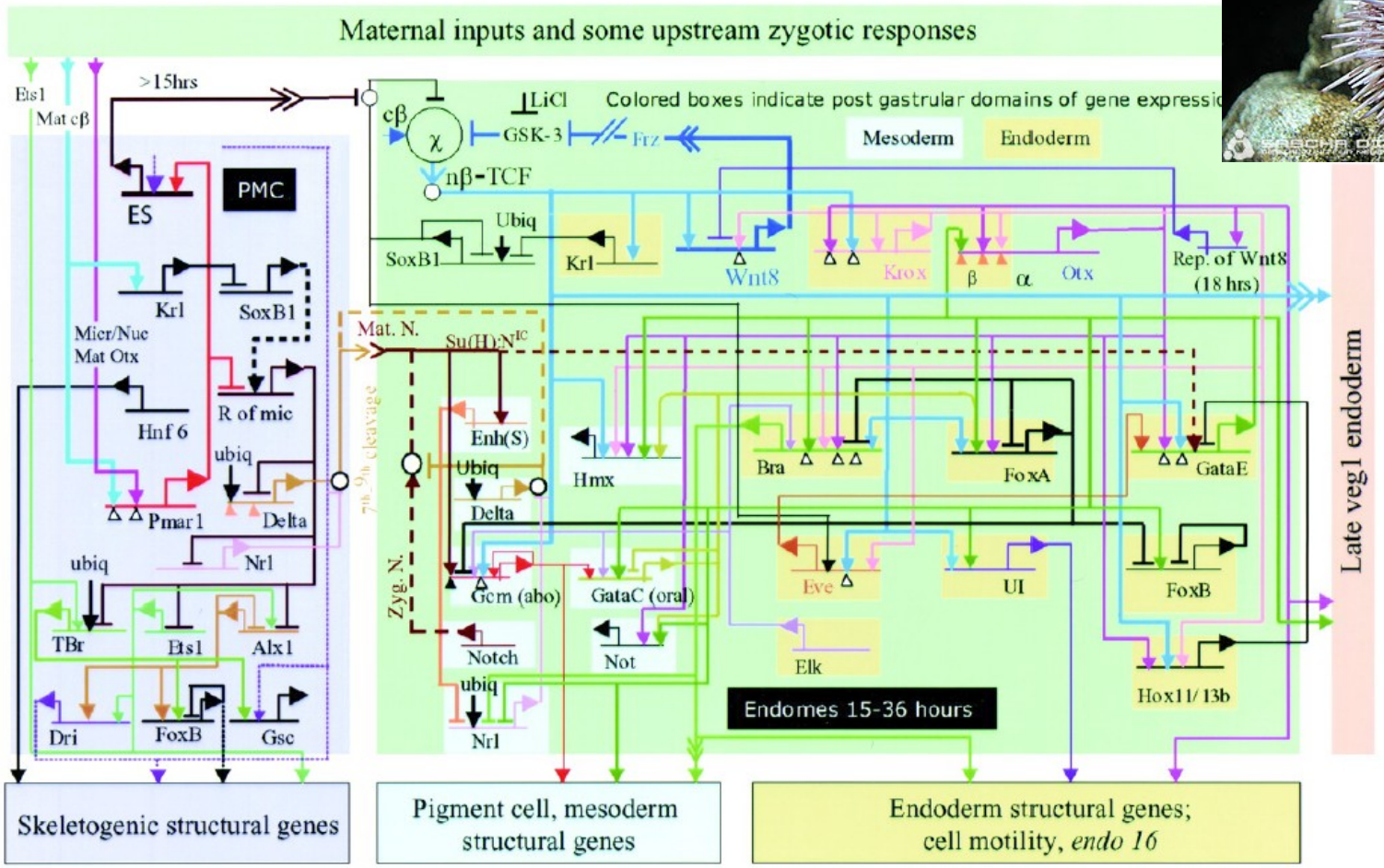
q=1 (manhattan), q=2 (euklidic)


A
Cluster method: Complete Linkage
Distance metric: Euclidean (Nulweighted)
B

Reverse Engineering Genetic Networks

Reverse engineering of Boolean networks aims to derive the Boolean interaction rules from time-dependent gene expression data (or from knockout experiments).

The genetic Network of embryonal development of sea uricin



Molecules to (functional) modules

From molecular to modular cell biology

Leland H. Hartwell, John J. Hopfield, Stanislas Leibler and Andrew W. Murray

(***Nature***, Dec 99)

To describe biological functions, we need a vocabulary that contains concepts such as amplification, adaptation, robustness, insulation, error correction and coincidence detection. For example, to decipher how the binding of a few molecules of an attractant to receptors on the surface of a bacterium can make the bacterium move towards the attractant (chemotaxis) will require understanding how cells robustly detect and amplify signals in a noisy environment.

notion of function. Therefore, in our opinion, the most effective language to describe functional modules and their interactions will be derived from the synthetic sciences, such as computer science or engineering, in which function appears naturally.

The essence of computational science is the capacity to engineer circuits that transform information from one form into

Network Motifs

Monod-Jacob (1961):

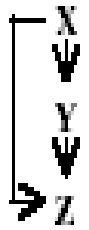
„It is obvious from the analysis of these [bacterial genetic regulatory] mechanisms that their known elements could be connected into a variety of „circuits“ endowed with any desired degree of stability.

(Wolf-Arkin, June 03)

Network motifs

- **are small subnetworks (max 5 nodes?)**
- **perform specific information processing tasks (= „natural circuits“)**
- **repeat (in a statistically significant way)**
- **are (probably) evolutionarily conserved**
- **are analogous to protein motifs**

GRN Motif example

Network	Nodes	Edges	N_{real}	$N_{rand} \pm SD$	Z score
Gene regulation (transcription)					Feed-forward loop
E. coli	424	519	40	7 ± 3	10
*S. cerevisiae**	685	1,052	70	11 ± 4	14

(Milo et al, Science 02)

Feedforward Loop

- **A regulator that controls a second Regulator**

and together they bind a common target gene

Function

- **A switch for rejecting transient input**

Motif classes (1)

Table 1

Proposed regulatory motifs classified on the basis of dynamic function.

Motif	Function	Mechanisms	Examples
Switches	Digital control Computation Signal integration, amplification and noise rejection	Transcriptional control, cooperativity [23,95], Zero-order [26] cascades [24,25] Multi-input [26] Cross-repressive feedback [30,31] Positive feedback [32,33**,35] Invertible DNA and ratio-based control [29*]	*fim* in *E. coli* Phage lambda Quorum sensing MAPK and c-Jun amino terminal kinase (JNK) pathways in *Xenopus* Synthetic switches [31,33**]
Oscillators	Temporal/sequence loop Synchronize to environment Reject noise Carry signal	Relaxation, harmonic, ring oscillators Negative feedback with high gain or a delay Positive feedback Combinations of positive and negative feedback [42,45,54].	Cell cycle cAMP Circadian rhythms Glycolysis [43] Cytosolic Ca^{2+} Synthetic oscillators [46,96–98]
Biphasic amplitude filters	Tune phenotype to environmental niche Auto-regulation Computation Amplitude multiplexing	Differentially activating binding affinity clusters [29*] Scaffolds [41] Concentration-dependent pathway activation/repression [39]	*fim* temperature tuning [29*,99,100] gltBDF [37] TBP [38]

[D.Wolf, A. Arkin]

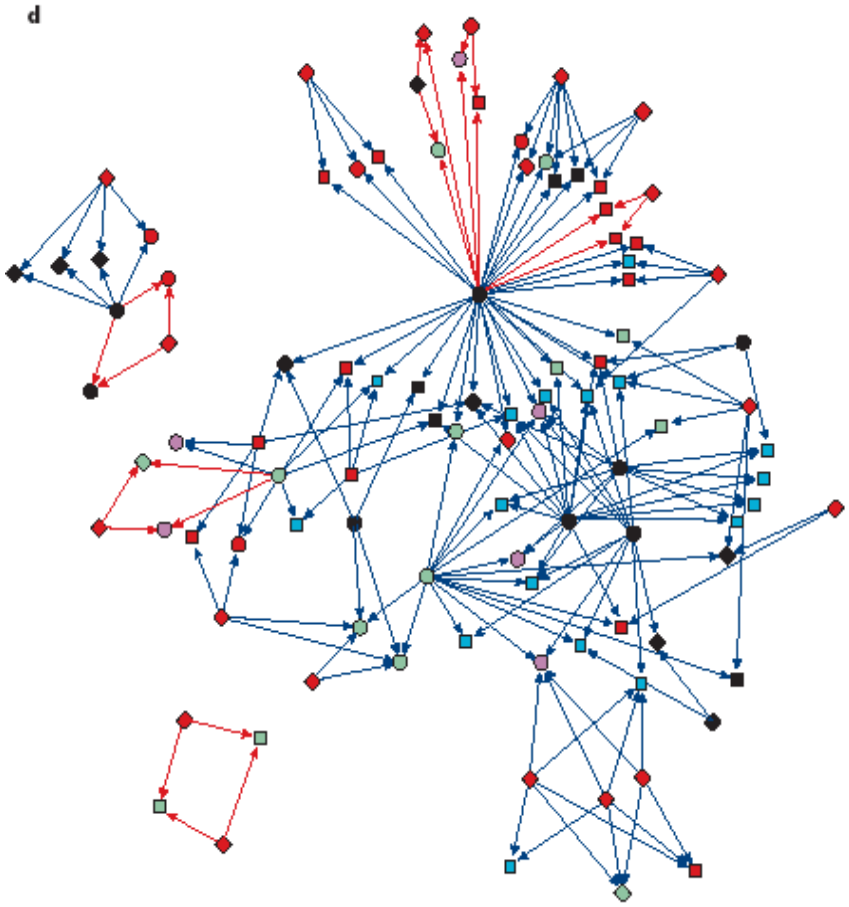
Motif classes (2)

Bandpass frequency filters	Interpret dynamic signals Filter noise Demodulate Demultiplex	Third-order chemical reactions Excitable media bandpass filter [53] Integral feedback [55] Saturated kinase and phosphatase activity Receptor desensitization [50,54,101]	Interleukin-2 activation by Ca^{2+} [52] Neural growth cones cAMP frequency decoding
Memory	Event tracking Sequencing Process control Temporal integration of signals	Multi-stability DNA inversion Receptor methylation DNA methylation [102] Histone acetylation Phosphorylation timers [103] Hysteresis and delays [29•,63•]	Developmental switches Cell cycle Sic1 [103] Shufflons Type 1 piliation, Chemotaxis
Noise filters	Precise regulation from noisy components.	Negative feedback Redundancy Cascades Checkpoints Delay lines [36,58,104] Frequency filters [53]	MAPK cascades [105] Cell cycle and flagellar synthesis checkpoints; Negative feedback [33••]
Noise amplifiers	Population heterogeneity, antigenic variation.	Noise controlled bistability [30] DNA rearrangement Slipped-strand mispairing [34]	Lambda phage [30] pap *fim* *his* Shufflons [34]

[D.Wolf, A. Arkin]

Motif clusters

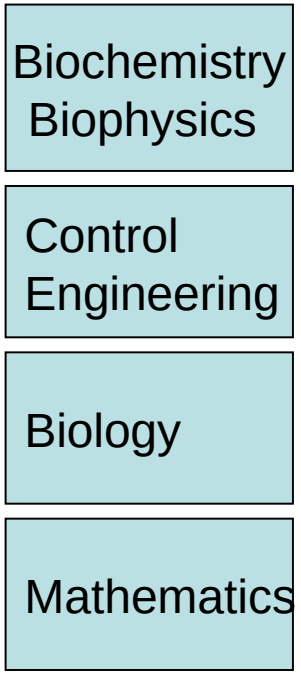
- Recent observation [Dobrin et al]: Specific motif types aggregate to form large motif clusters
- Example: in E.coli GRN, most motifs overlap, generating homologous motif clusers (→ specific motifs are no longer clearly separable)
- More research on motif interaction needed!



(Barabasi-Oltvai Feb 04)

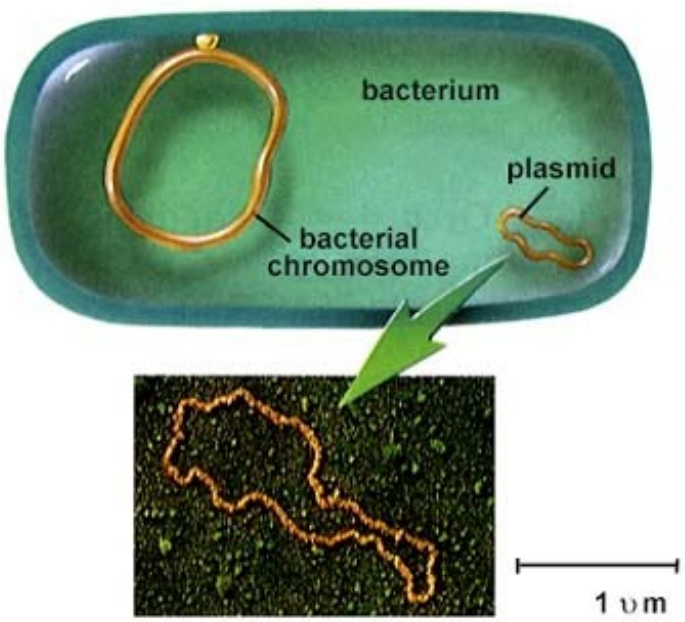
What are (functional) modules?

- **Diverse characteristics proposed:**
  - chemically isolated
  - operating on different time or spatial scales
  - robust
  - independently controlled
  - significant biological function
  - evolutionarily conserved
  - clustered in the graph theory sense
  - ...
  - any combination of the above



“Programming” Cells

plasmid = “user program”



Vision

- A new substrate for engineering: living cells
  - interface to the chemical world
  - cell as a factory / robot
- Logic circuit = process description
  - extend/modify behavior of cells
- Challenge:
  engineer complex, predictable behavior

Ron Weiss (Princeton)