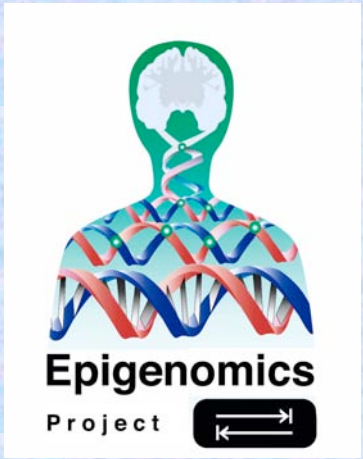
ÉPIGÉNOMIQUE DES RÉSEAUX DE RÉGULATION BIOLOGIQUE :
Le cas du modèle solénoïde des chromosomes

François Képès

Epigenomics Project / Genopole®
Evry, France



Centre d'Alembert, CIEFIST, mai 2006

OUTLINE

- **Start from a biological object ->**
- **Analyze unexpected features ->**
- **Inspire multidisciplinary research ->**
- **Uncover specific CS research tracks ->**
- **Raise increasingly theoretical issues …**

Start from a biological object

Biological issues

Although in broad outline, the fundamentals of molecular biology have been known for a long time, there remains a deep mystery.

How is the overall process of gene expression controlled?

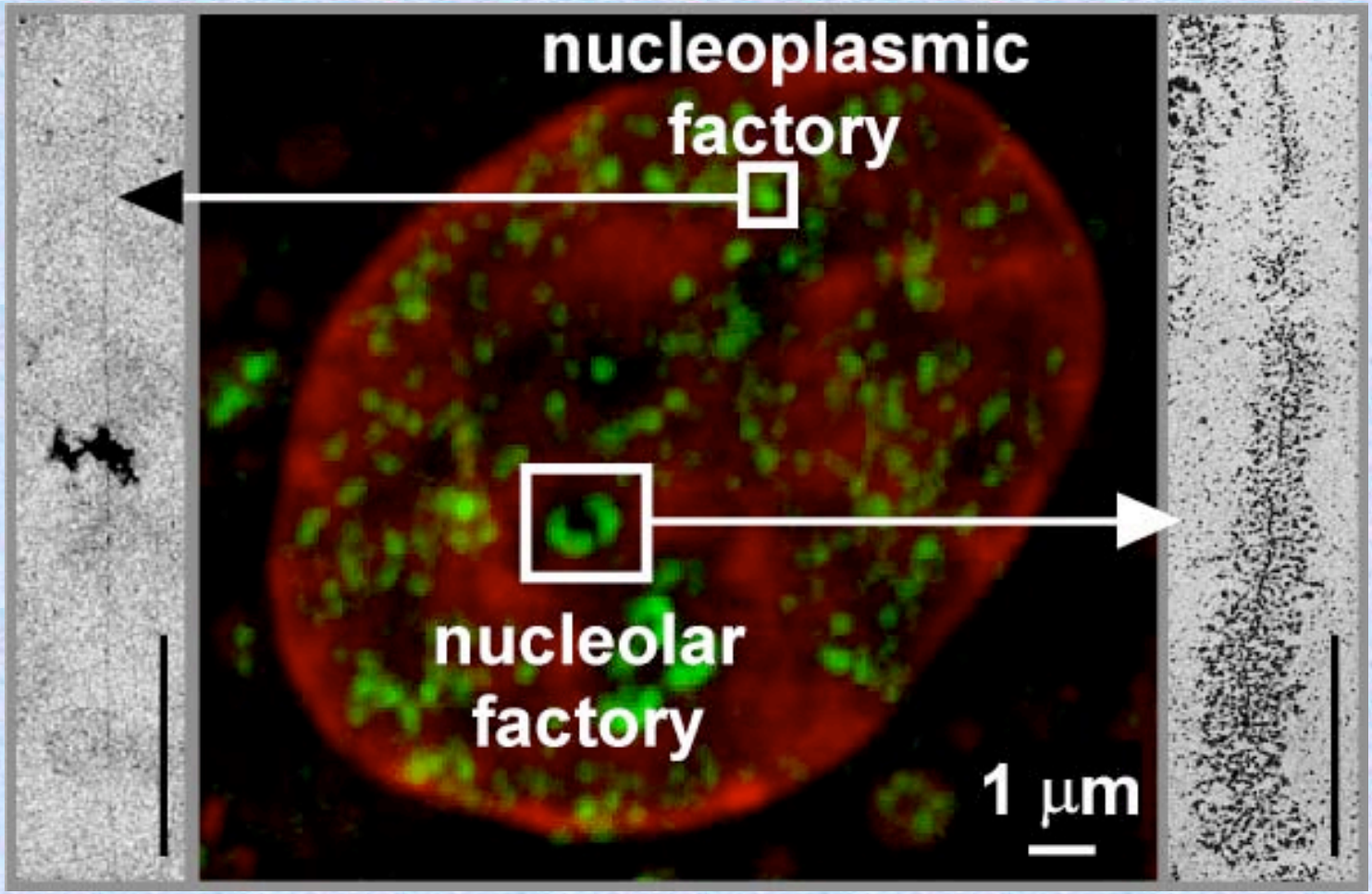
How are genes turned on and off with such delicate and adaptive precision?

How can many of them switch in such a concerted manner and in so little time?

Recent findings may shed some light on these old questions.
They fall in two categories: morphological and transcriptomic evidences.

Morphological evidence

Heterogeneous, multi-focal distribution of Transcription Factories



From Cook, P.R. (1999). Science *284*, 1790-1795

Transcriptional evidence

Is there an overall transcriptional scheme?

How does it achieve the goals set forth in previous slides?

Analyze unexpected features

Take-home message

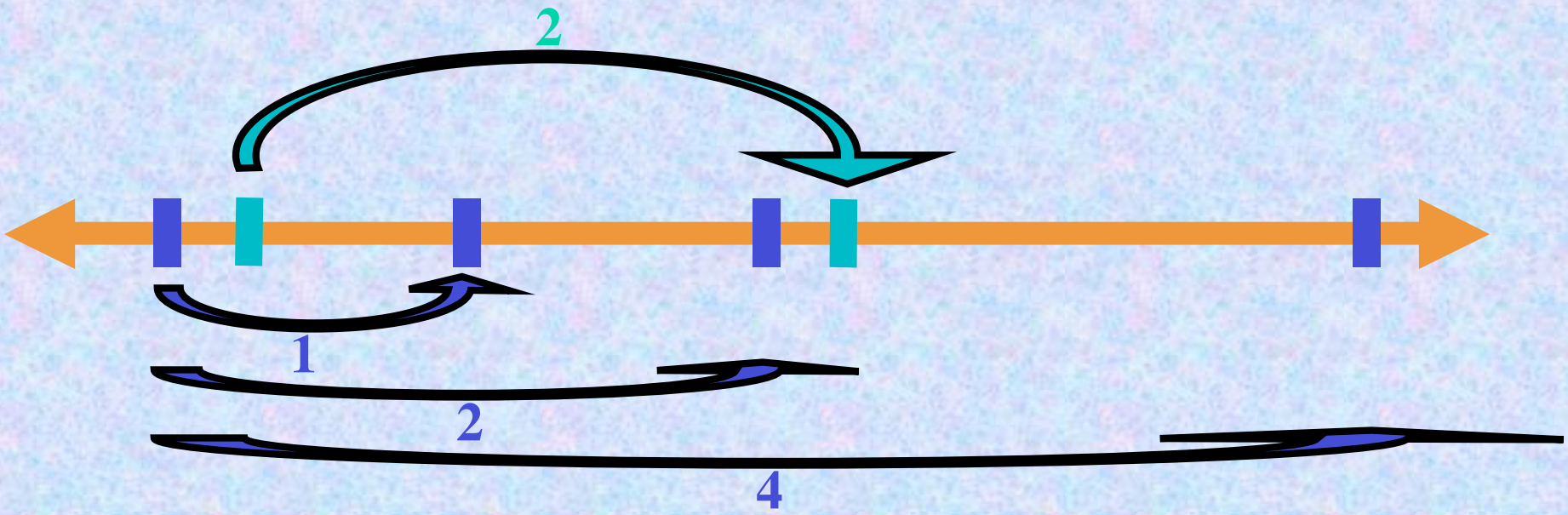
Co-regulated genes tend to spatially co-localize

Data

List of targets of the dedicated Transcription Factor Rap1p
(based on bench experiments,
either classical or CHromatin Immuno-Precipitation) :

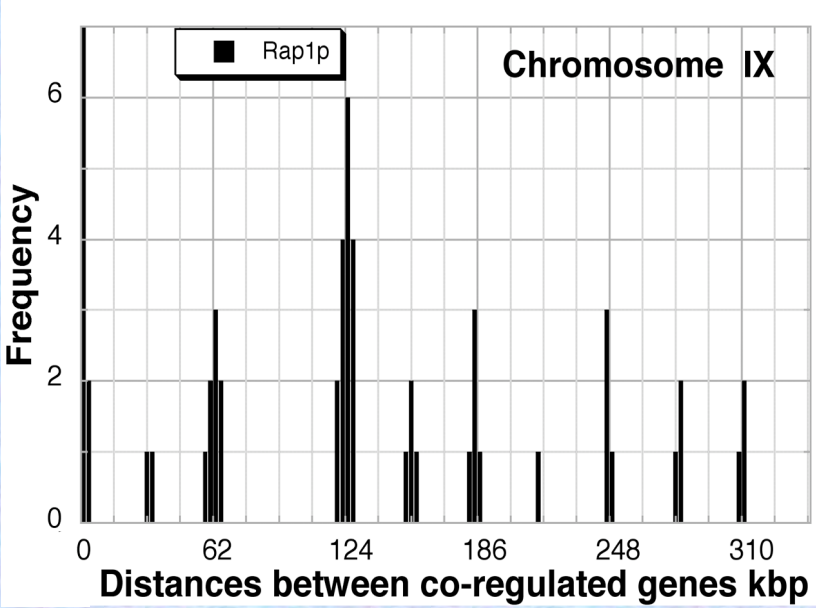
Gene	*Position on chromosome IX, left arm*
YIL012W	**409 325 bp**
YIL044C	**150 230 bp**
...	
YIL087C	**025 839 bp**

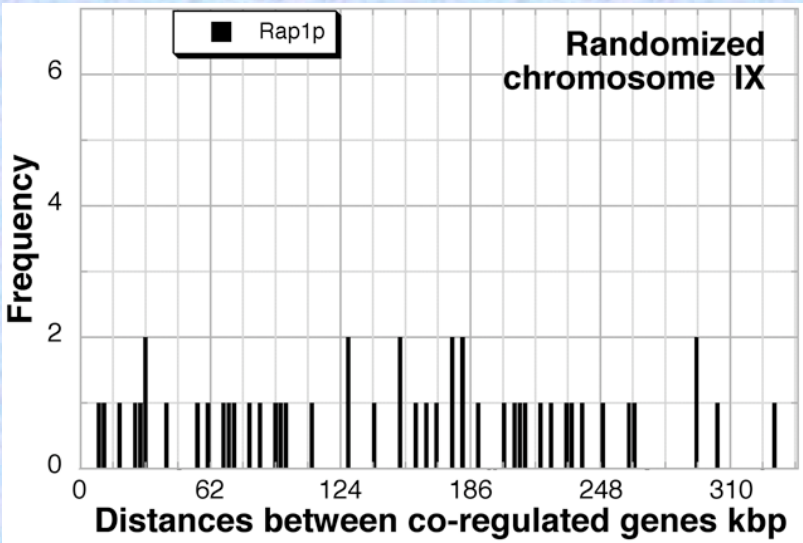
Method



Measuring distances between co-regulated genes
1 period <=> *n* base pairs

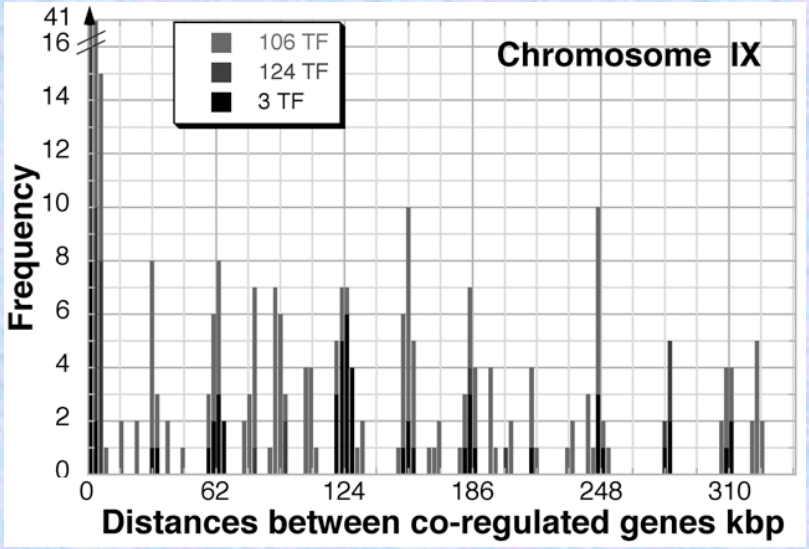
Yeast transcriptional scheme: one TF

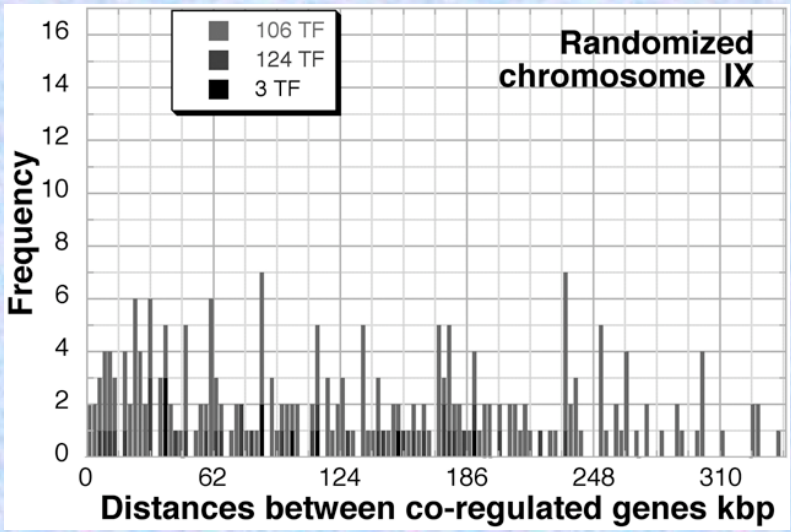




Yeast transcriptional scheme: all TFs

Guelzim, Bottani, Bourgine & Képès (2002) Nature Genetics 31: 60-63
Lee et al. (2002) Science 298, 799-804
Lieb, Liu, Botstein & Brown (2001) Nature Genetics 2 : 327-334





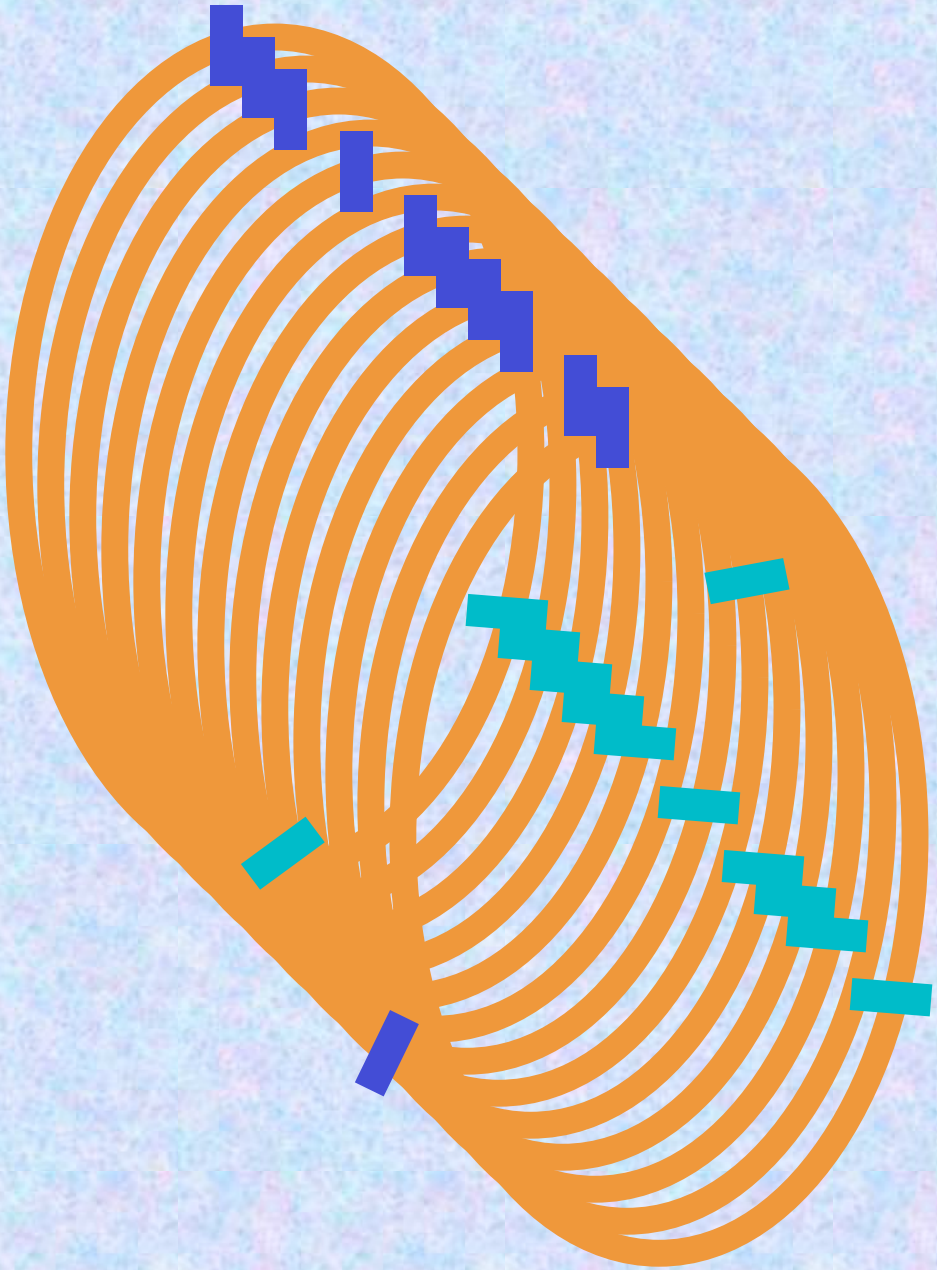
Yeast transcriptional scheme

Conclusion 1 :

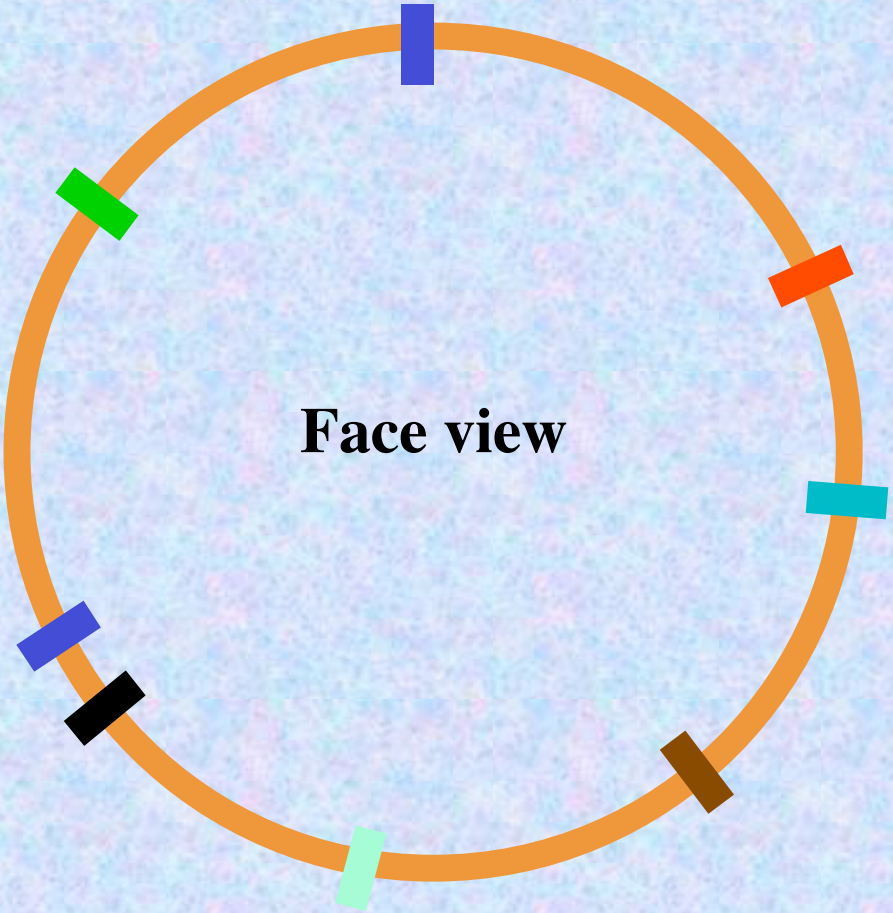
•Same period for all Transcription Factors

==> Solenoidal configuration of DNA

Solenoidal DNA



Solenoidal DNA



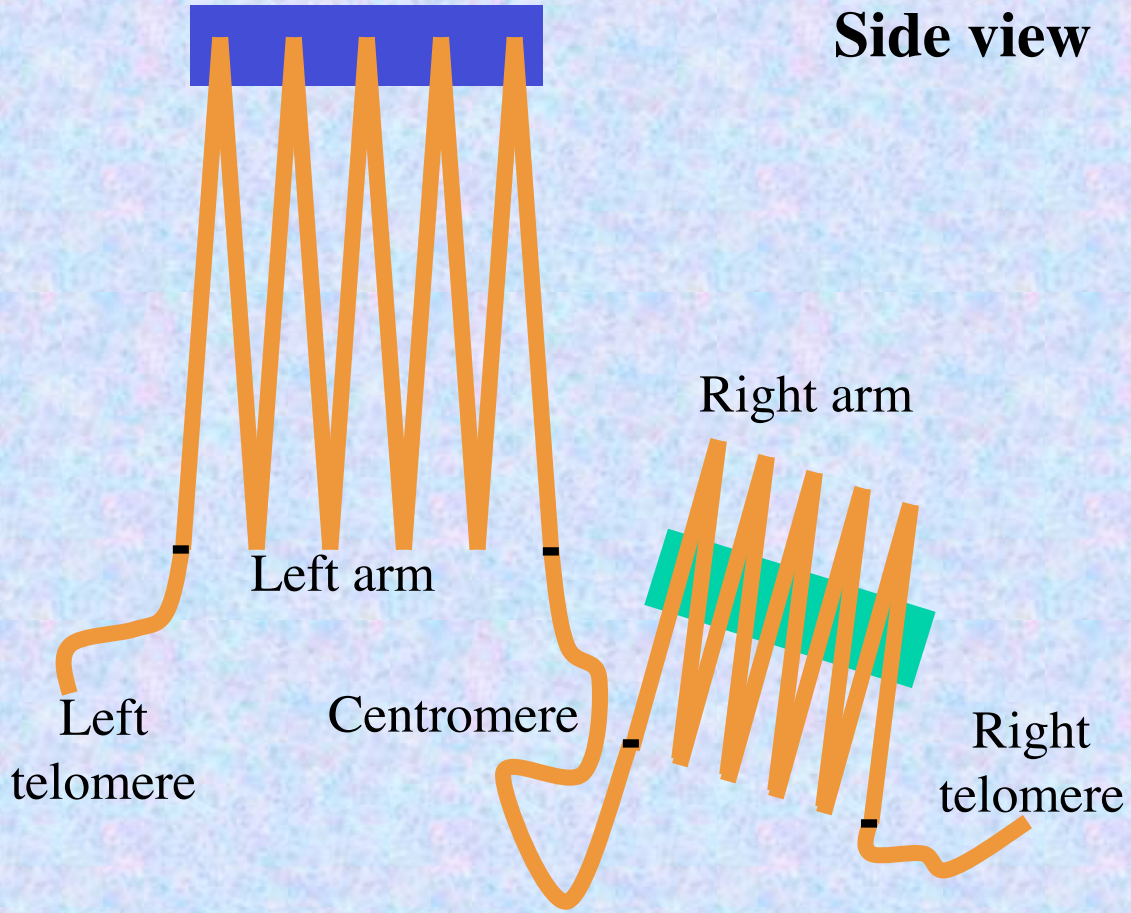
Yeast transcriptional scheme

Conclusion 2 :

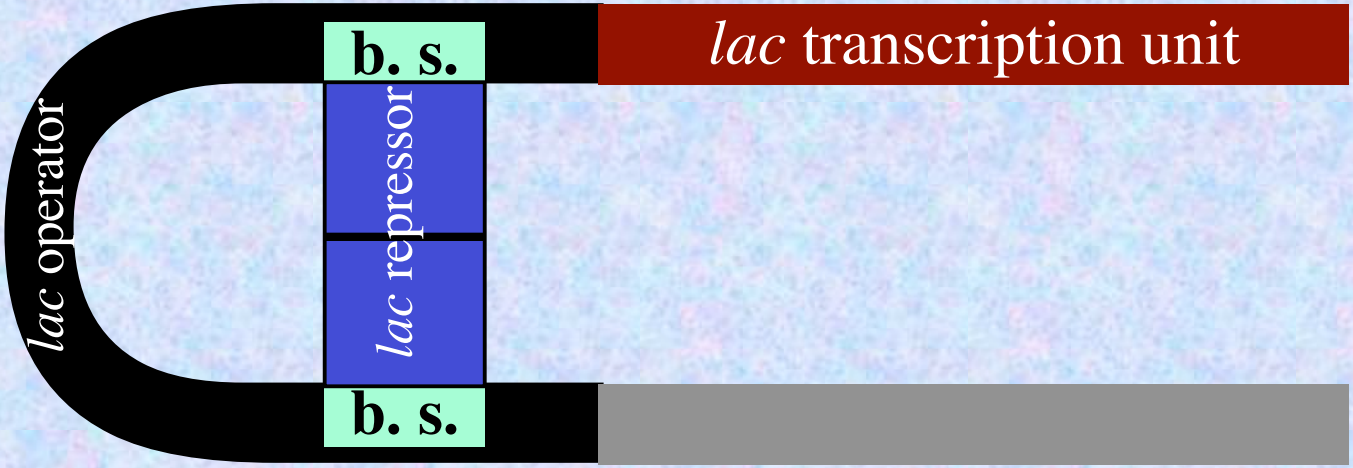
•Different periods for different chromosome arms

==> A consequence of transcription dynamics

Solenoidal DNA



Underlying mechanism : the case of the lactose repressor



A DNA loop induced by

- **bivalency of the repressor and**
- **the presence of 2 binding sites b.s.**

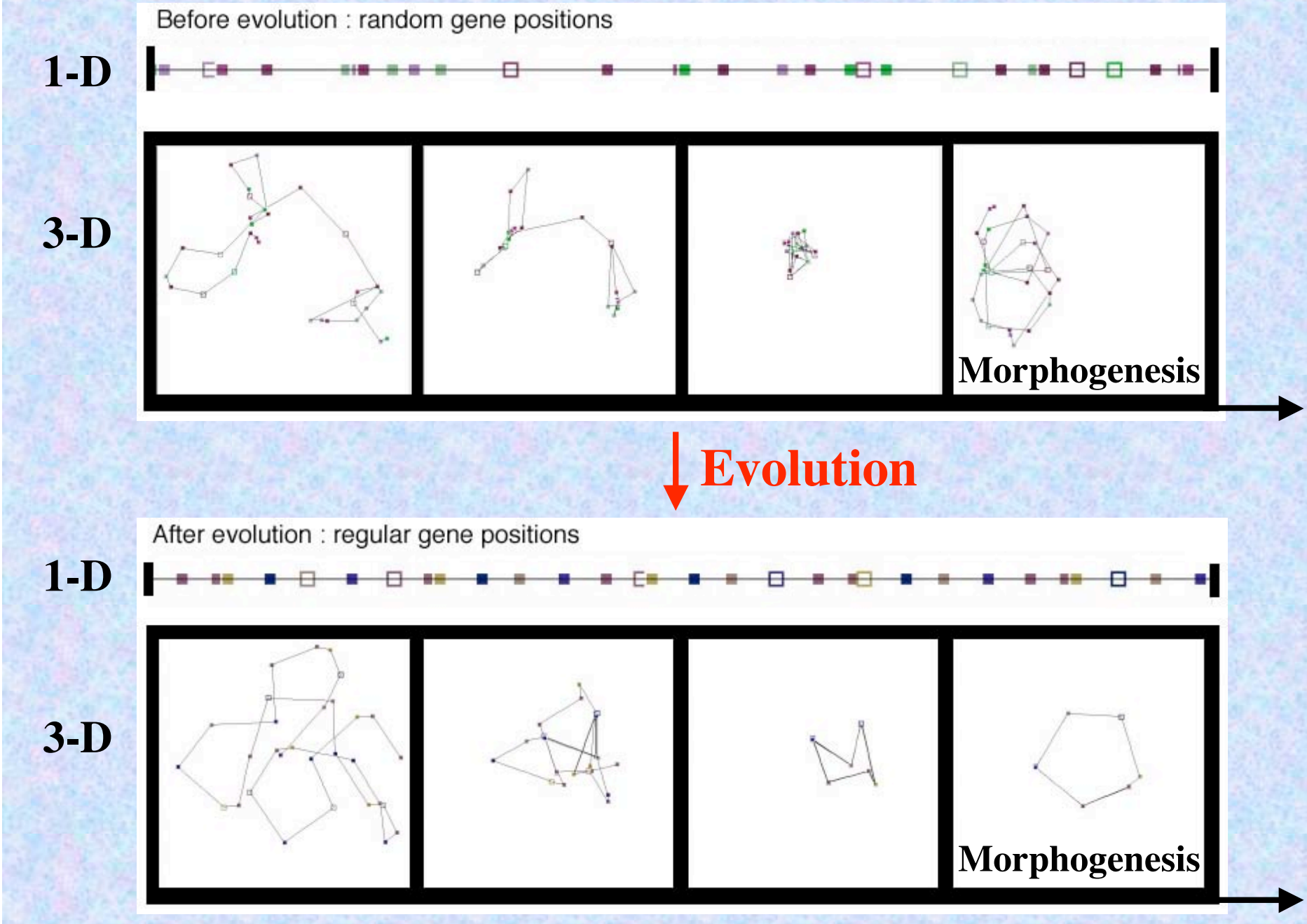
B. Müller-Hill

Inspire multidisciplinary research

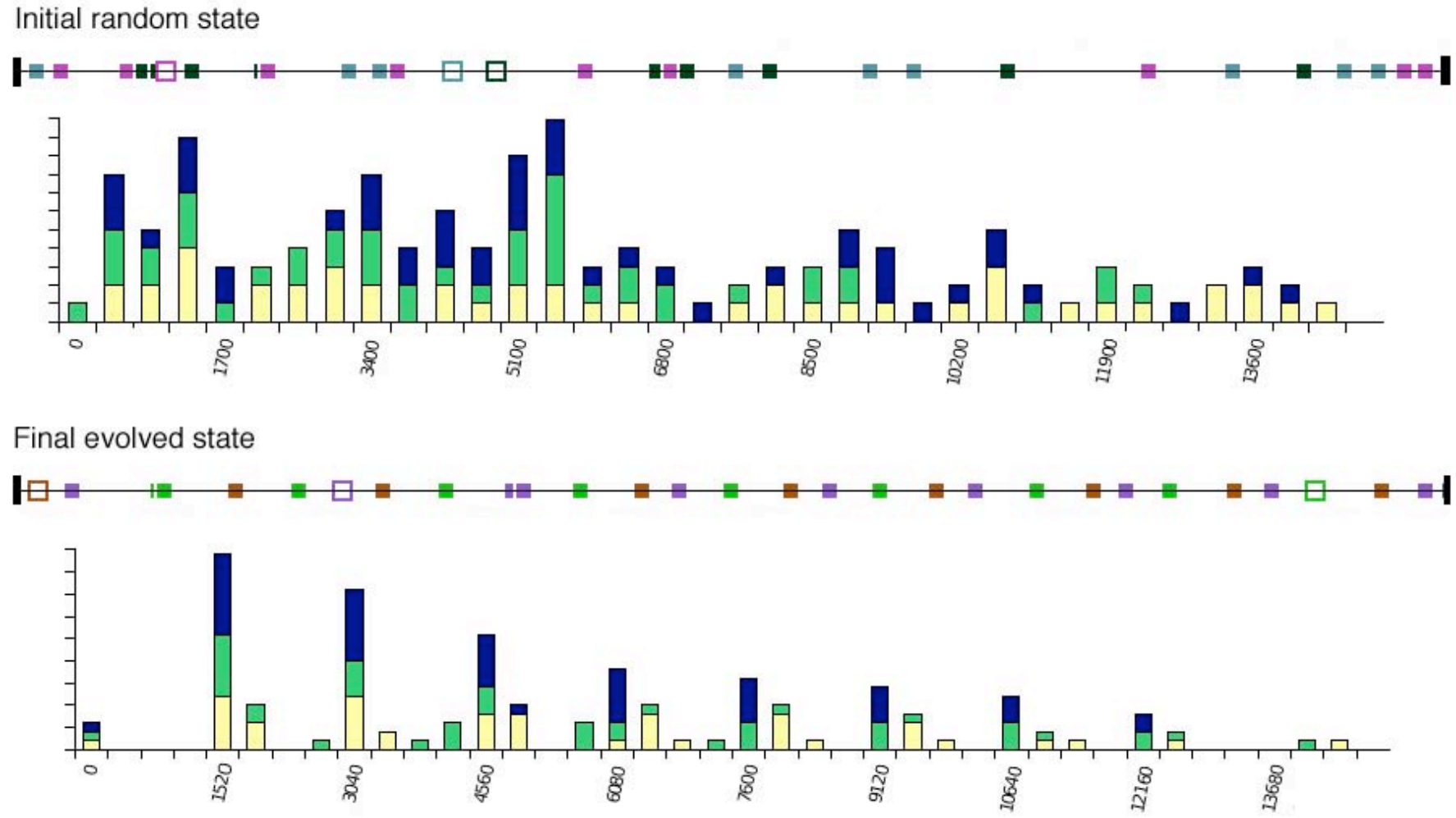
A biologist's question

How come such strong regularities are observed,
despite rapid gene shuffling over evolutionary times?

Genetic and Evolutionary Computation



Genetic and Evolutionary Computation

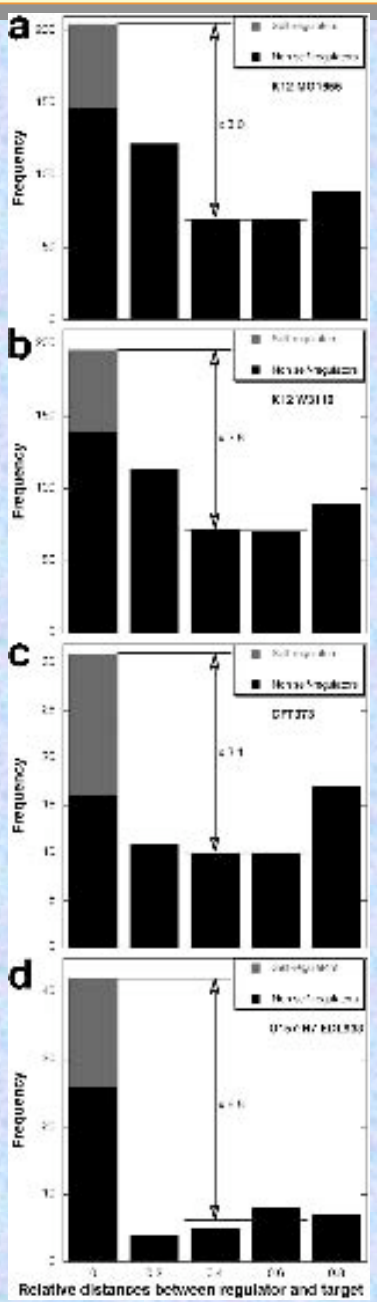


3 sets of 10 genes each

(W. Banzhaf , S. Leclercq & F. Képès)

Uncover specific CS research tracks

E. coli transcriptional scheme

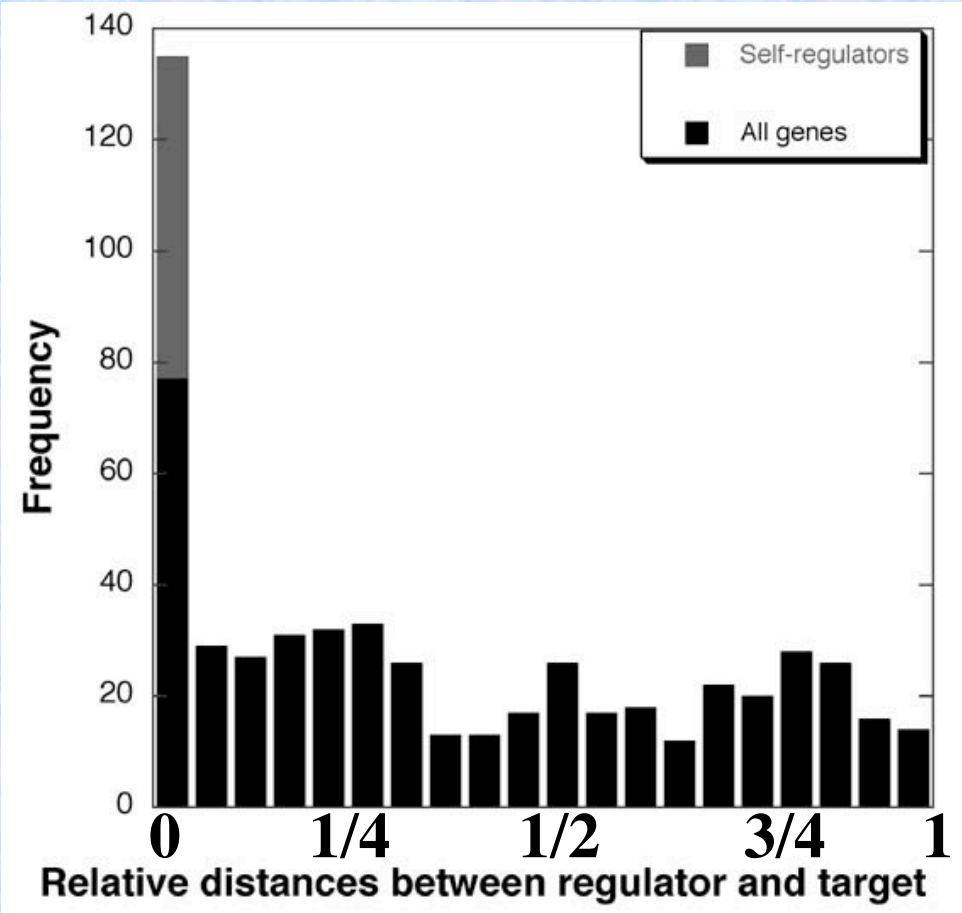


Conclusion 3 :

- **Several strains of *E. coli***

==> Several solutions to the same optimization problem

E. coli transcriptional scheme



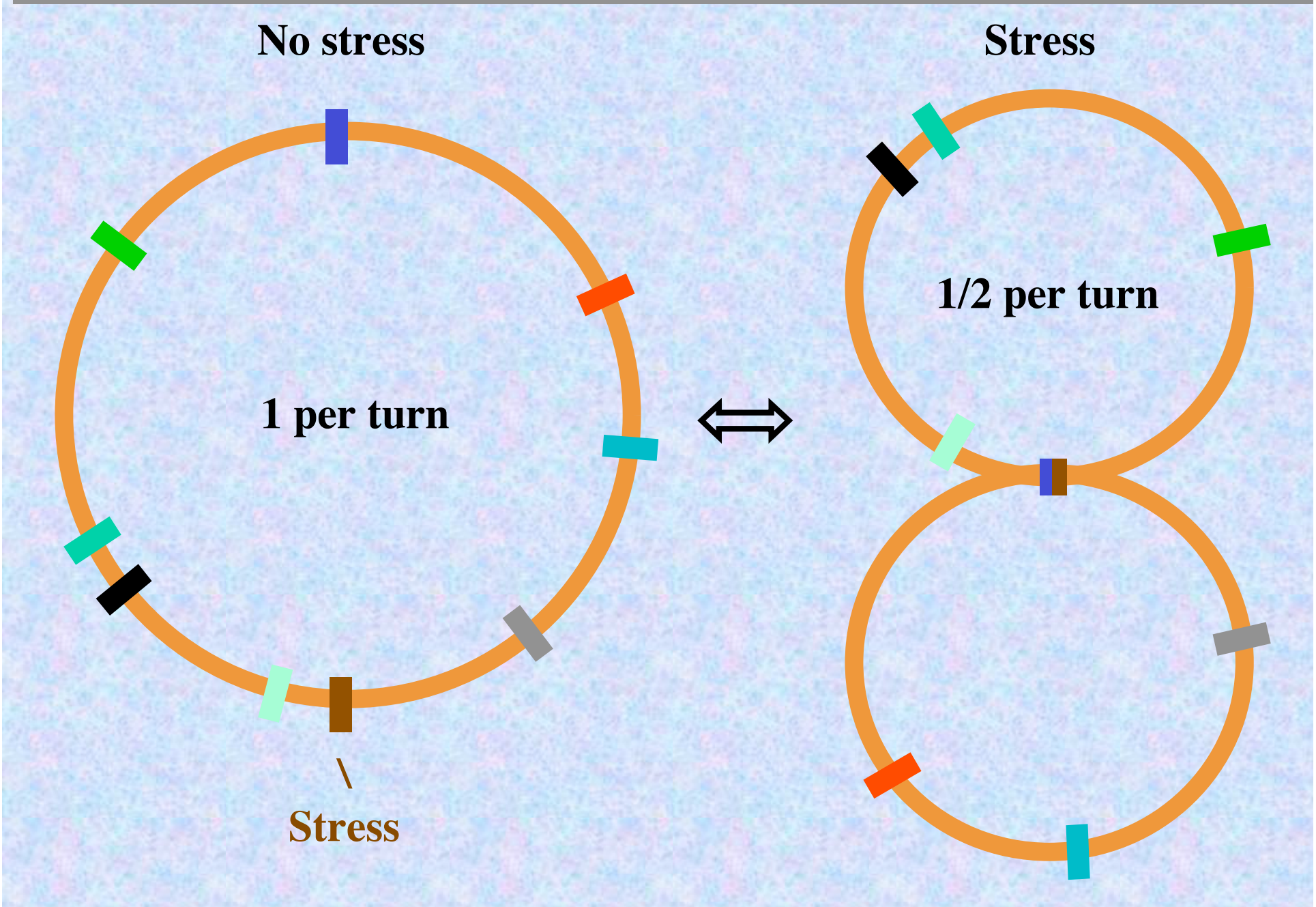
Conclusion 4 :

- **Several cell states**

==> One solution, several structures

(One 1-D DNA sequence, several 3-D structures)

Multi-objective optimization



Genetic and Evolutionary Computation (cont'd)

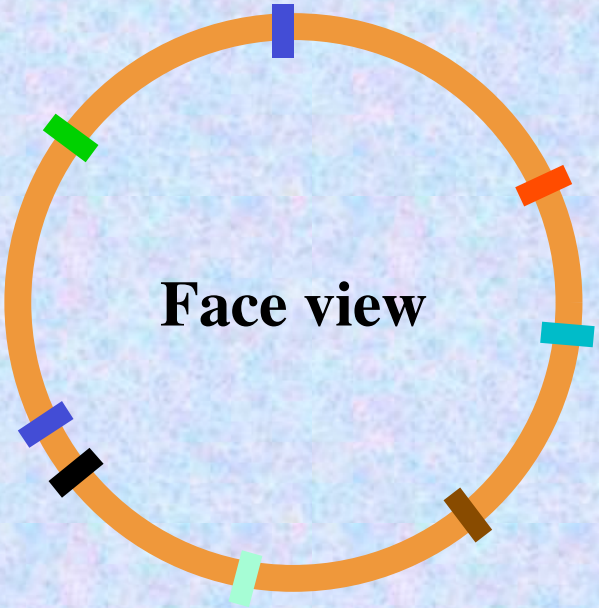
Genetic Programming:
a stochastic computational technique for the design of programs based on the paradigm of biological evolution.

One new research avenue, pioneered by W. Banzhaf (Memorial Univ, Canada) is to connect models of *artificial transcriptional networks* with techniques from Genetic Programming.

Genetic and Evolutionary Computation (cont'd)

Problem: The 'genes' in this model are not arranged along a 'chromosome'. This model has not yet been connected to a semantics of structures, an essential feature for GP.

Idea: (with W. Banzhaf & M. Schoenauer) Combine both models -> provide a structure for Banzhaf's artificial transcriptional network.



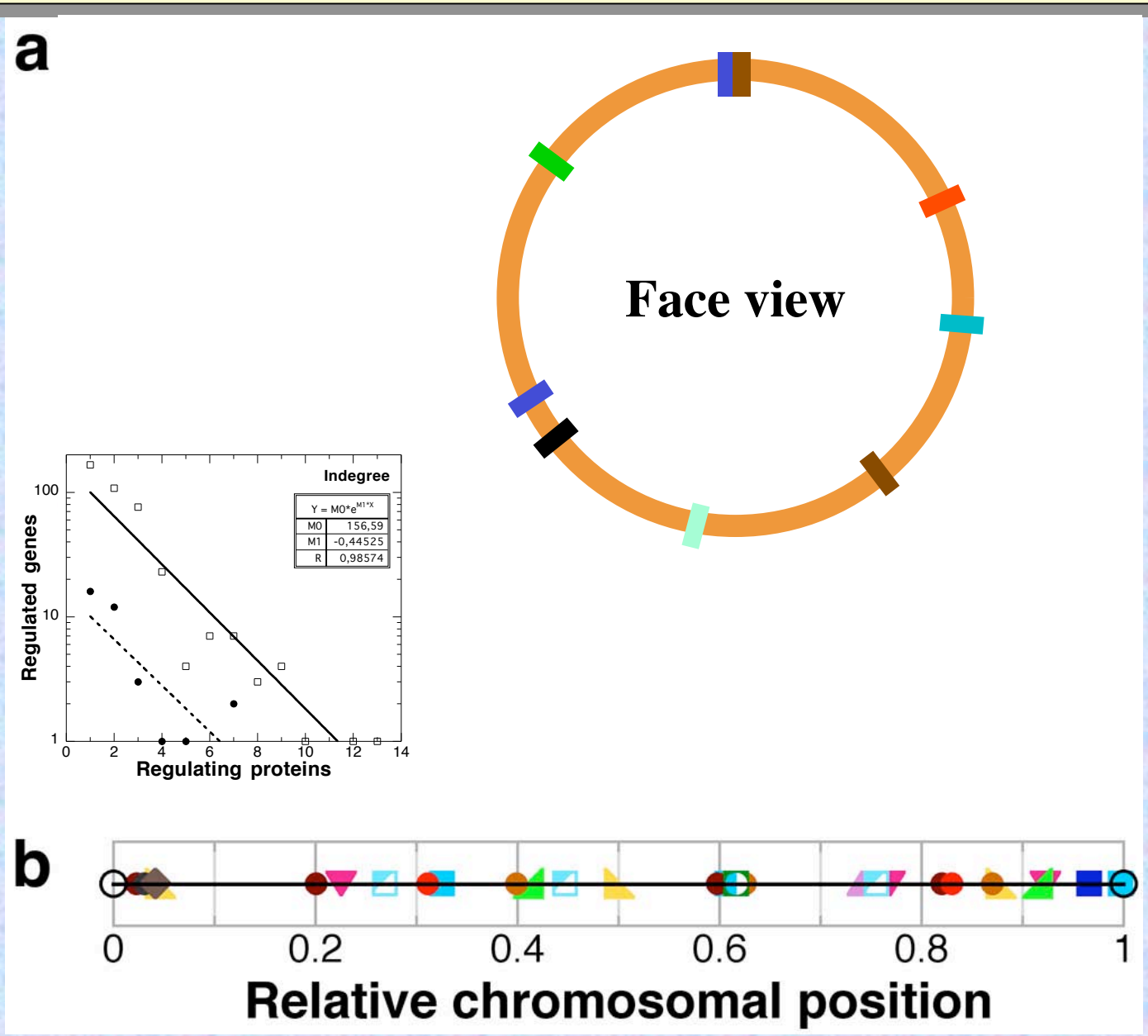
Multi-objective optimization

Gains:

1. Beyond the one solution <-> one structure case, GP now searches for one solution (1-D) that provides several alternative structures (3-D), thus allowing a whole new level of adaptiveness.
e.g. the whole program would now automatically switch in an adaptive way by changing the solenoidal period.

2. GP modeling in turn would allow to understand the minimal requirements for biological adaptiveness, in particular with respect to concerted transcriptional changes.

Combinatorial optimization



Raise increasingly theoretical issues