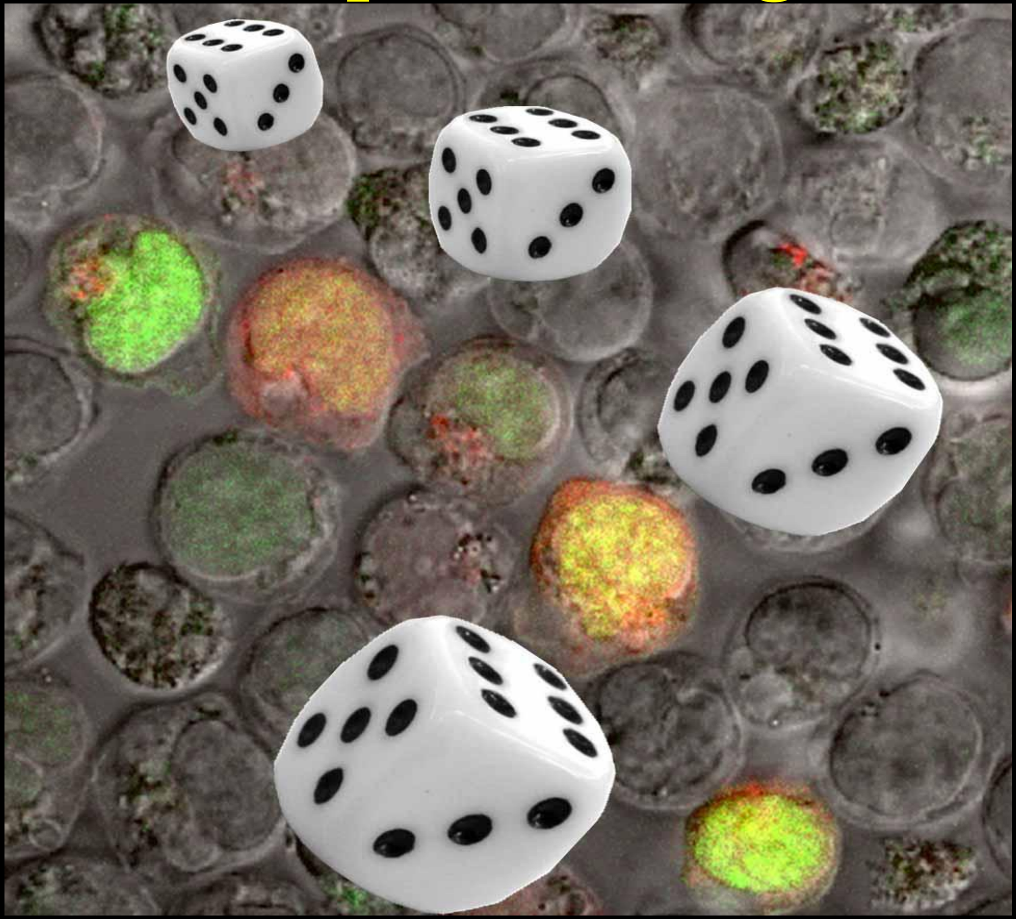
Stochastic Interactions Coordinate Transcriptional Programs



Dimitris Thanos
Biomedical Research Foundation
Academy of Athens
thanos@bioacademy.gr

Biology and Noise

Noise=Variability

TYPES OF NOISE

Noise can be classified as

internal : stochastic fluctuations in the number of protein molecules

external : fluctuations in environment and control parameters.

Noise also classified as

phenotypic noise: leading to qualitative differences in a cell phenotype (eg. lysis-lysogen pathway). Fluctuations cannot always be viewed as simply small perturbations as they can, in fact, induce different developmental pathways.

stablizable noise: leading to fluctuations in protein concentrations (robustness properties of biological systems).

Definition of Stochasticity

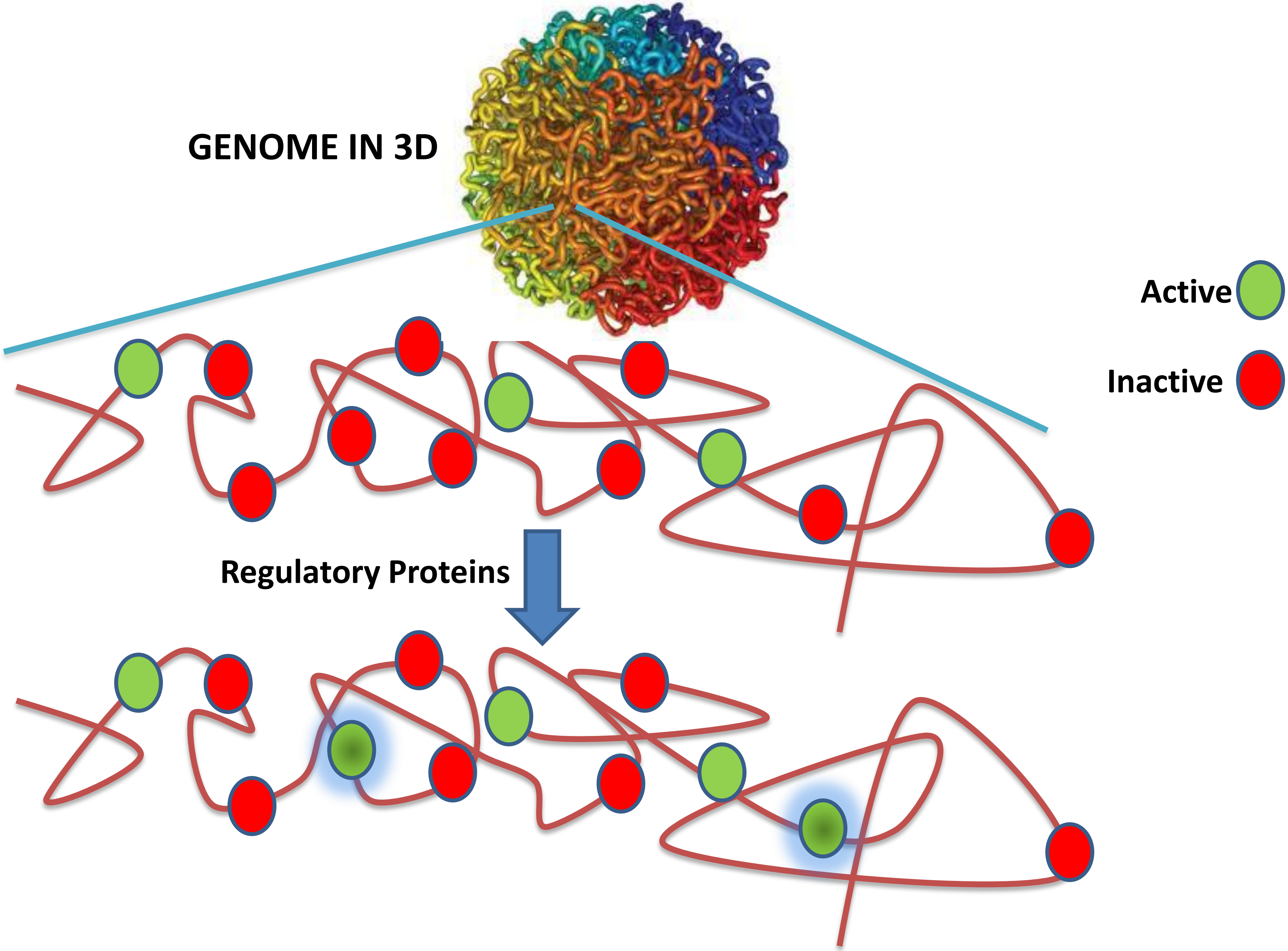
- I define stochasticity as randomness

- A stochastic process is one whose behavior is non-deterministic in that a state does not fully determine its next state

- Stochasticity arises because random intermolecular collisions make biochemical reactions stochastic

- Stochasticity is enhanced by low numbers of molecules (intrinsic) because low numbers changing by one or two can make individual events more significant

- Genetically identical cells can behave quite different from one another because of stochastic reactions



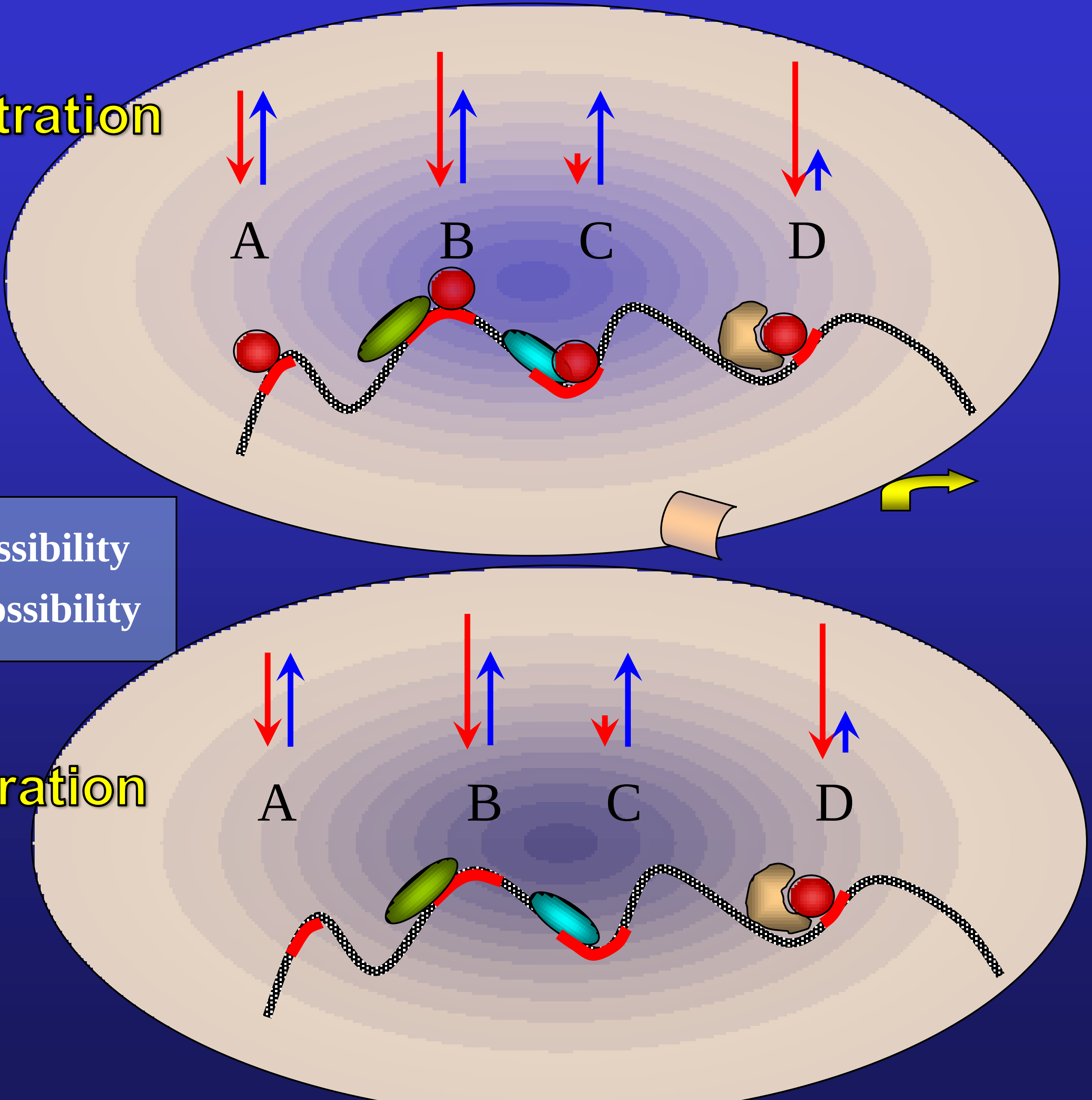
GENOME IN 3D

Active 🟢

Inactive 🔴

Regulatory Proteins

Dynamics of Transcription Factor DNA Binding

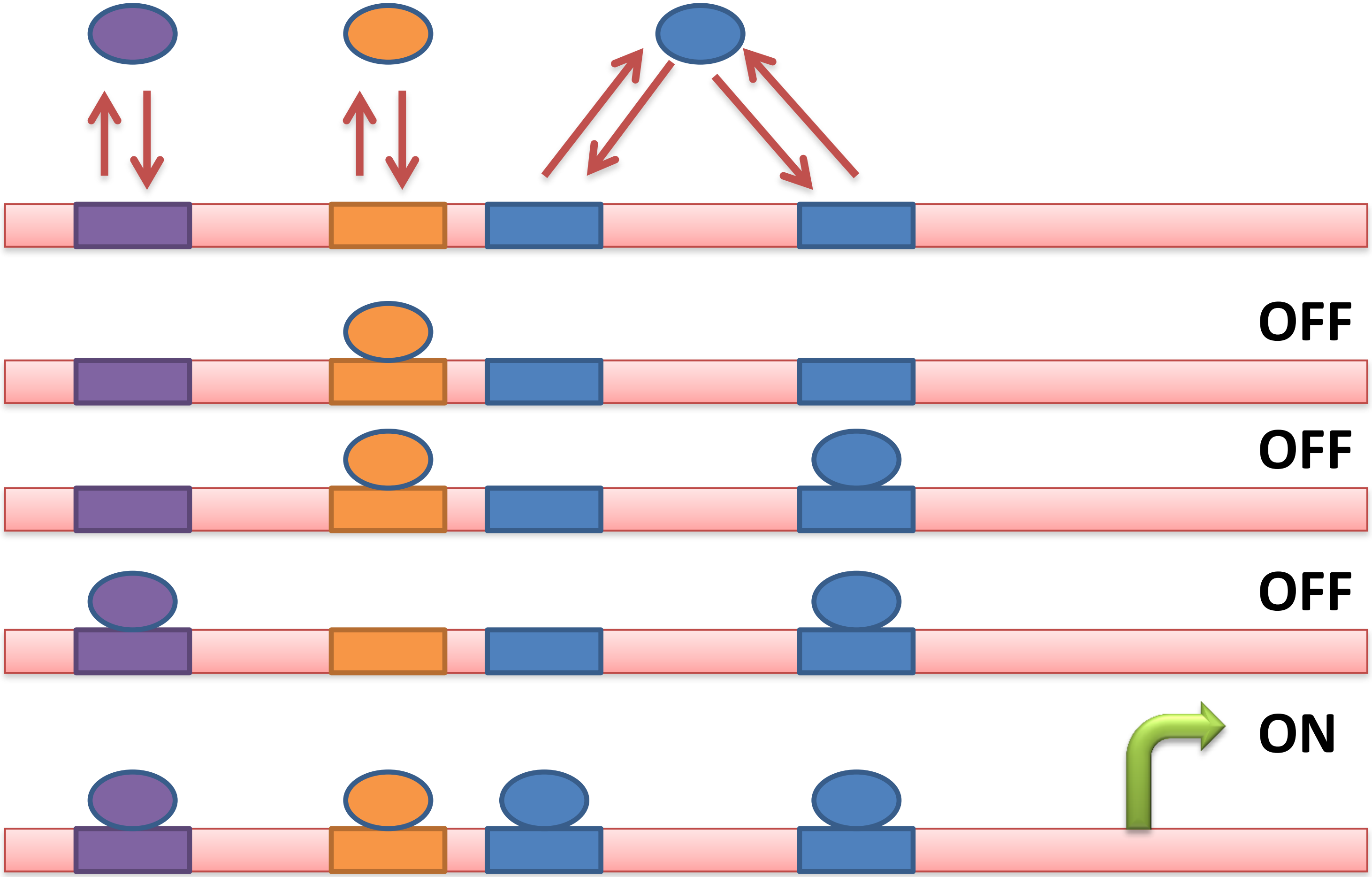


High concentration

Low concentration

- → Association possibility
- → Dissociation possibility

LOW CONCENTRATION OF TRANSCRIPTION FACTORS

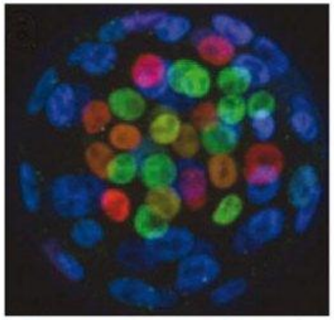


OFF

OFF

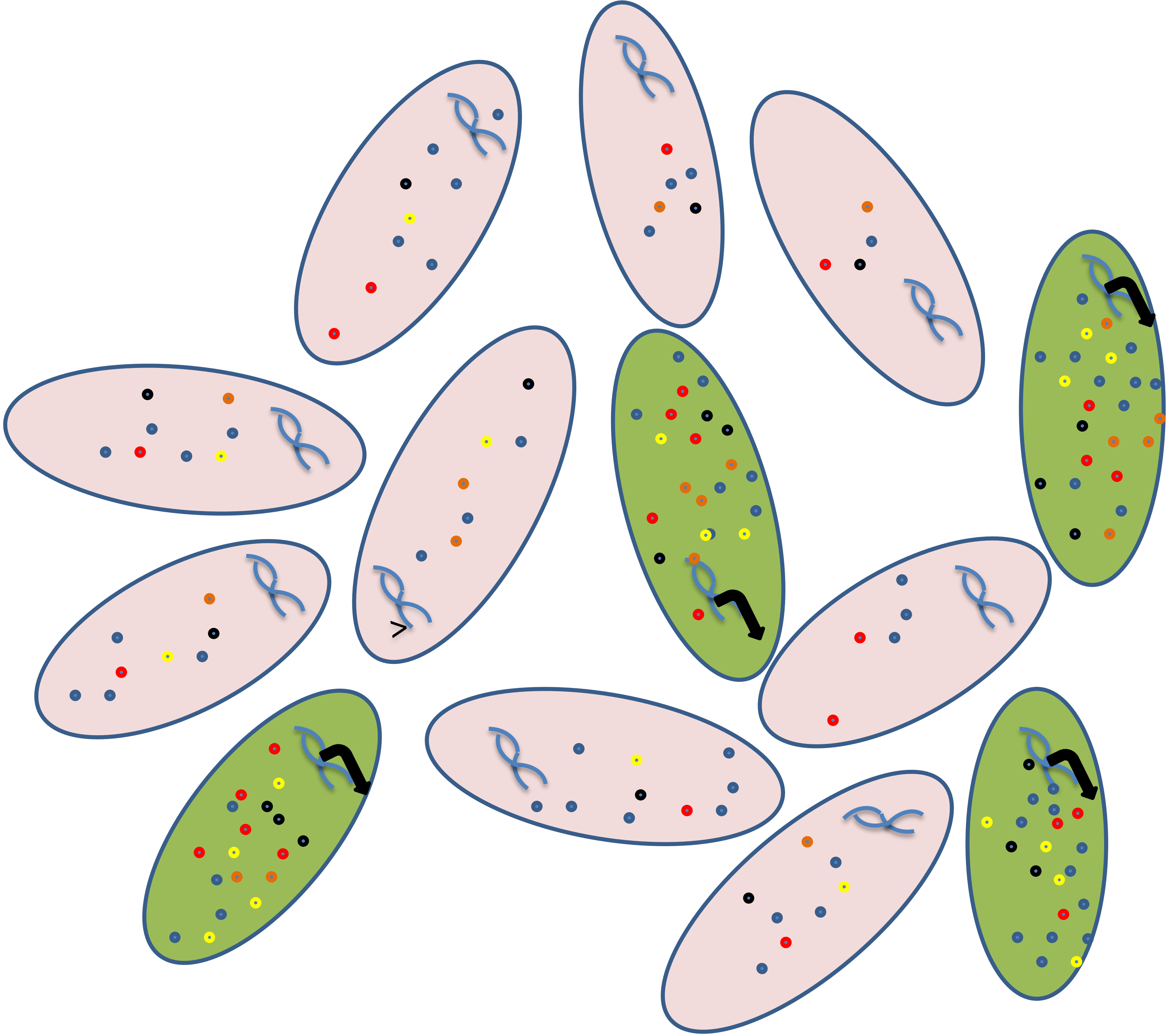
OFF

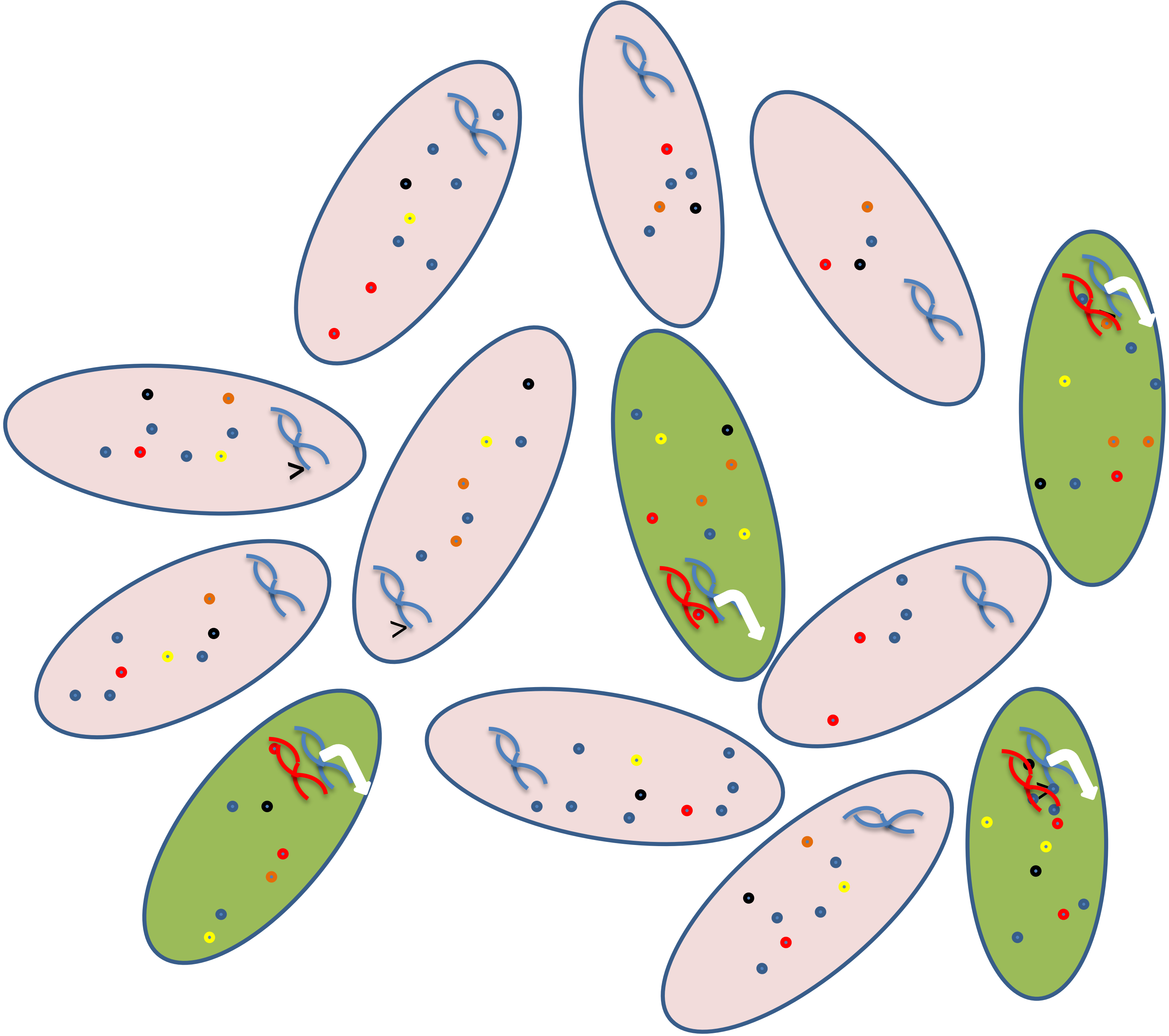
ON

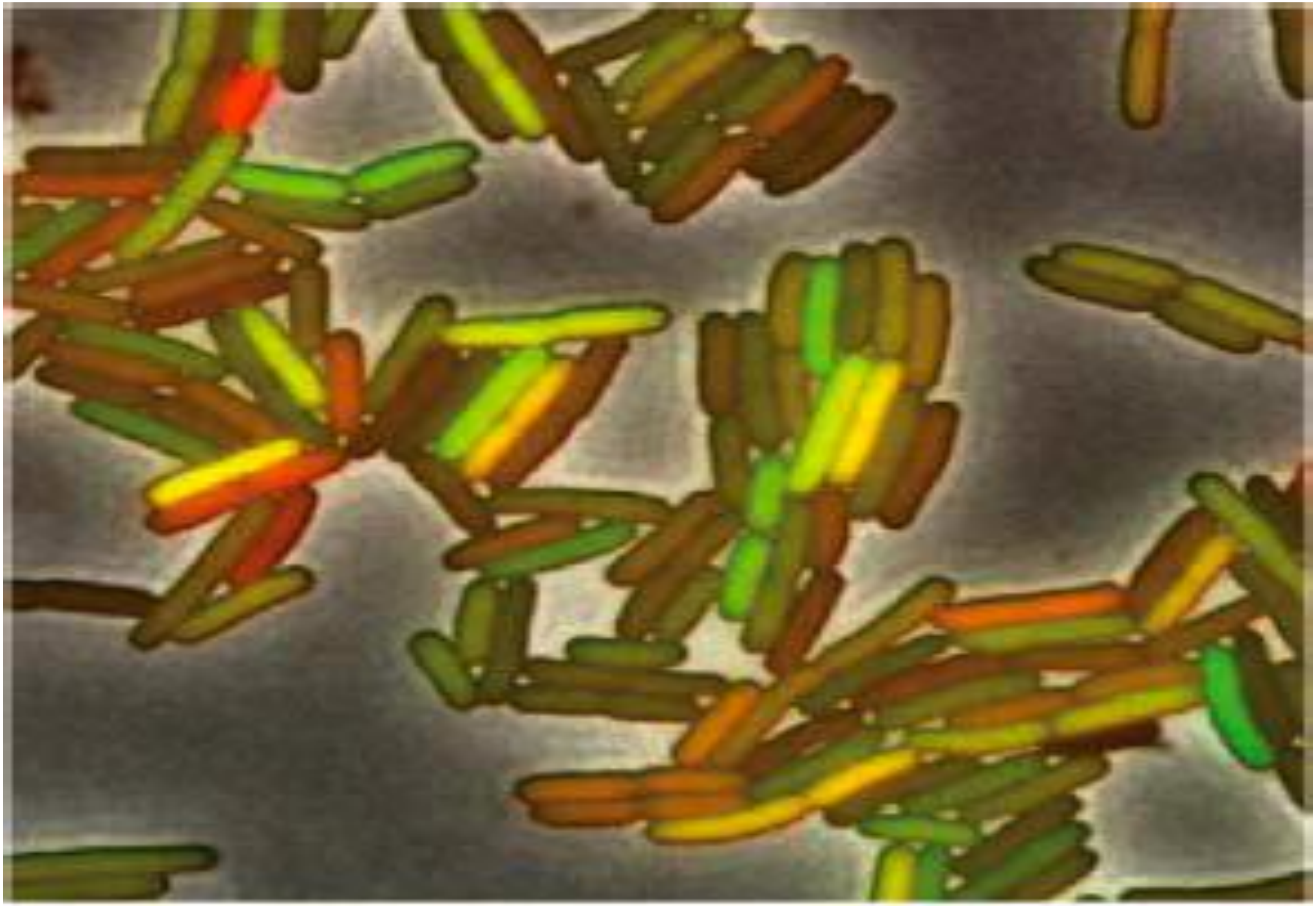


Nanog Gata6 DAPI

NOISY EXPRESSION
VARIABILITY FROM CELL TO CELL



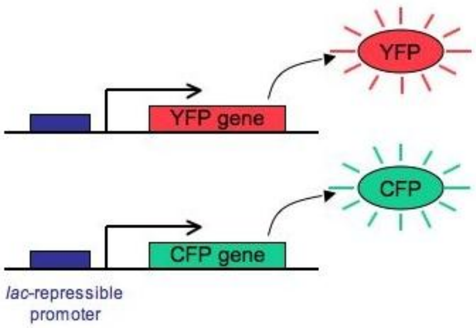




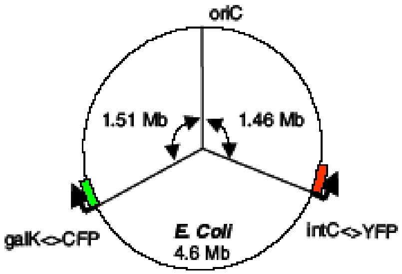
E. coli

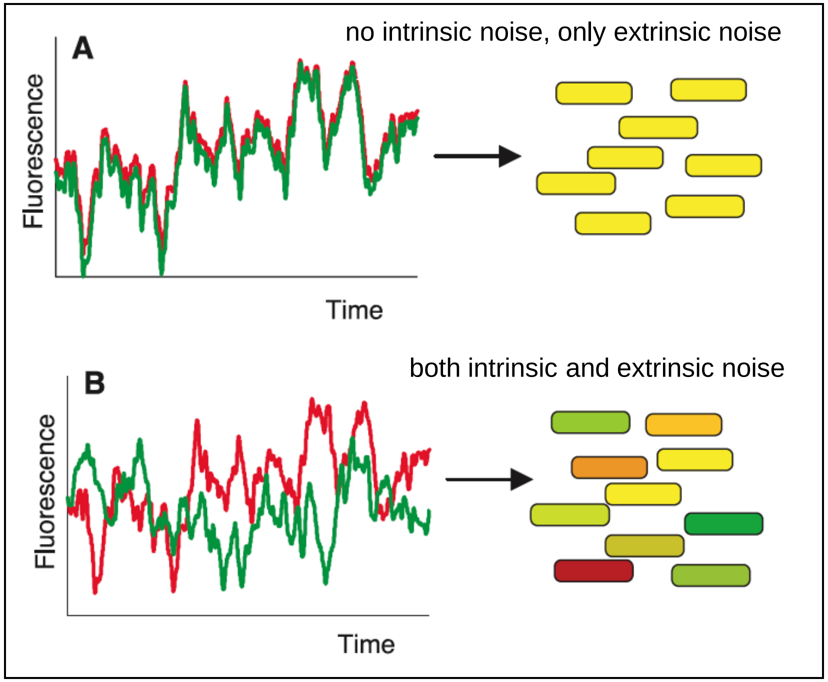
Stochastic gene expression in a single cell

Double reporter construction

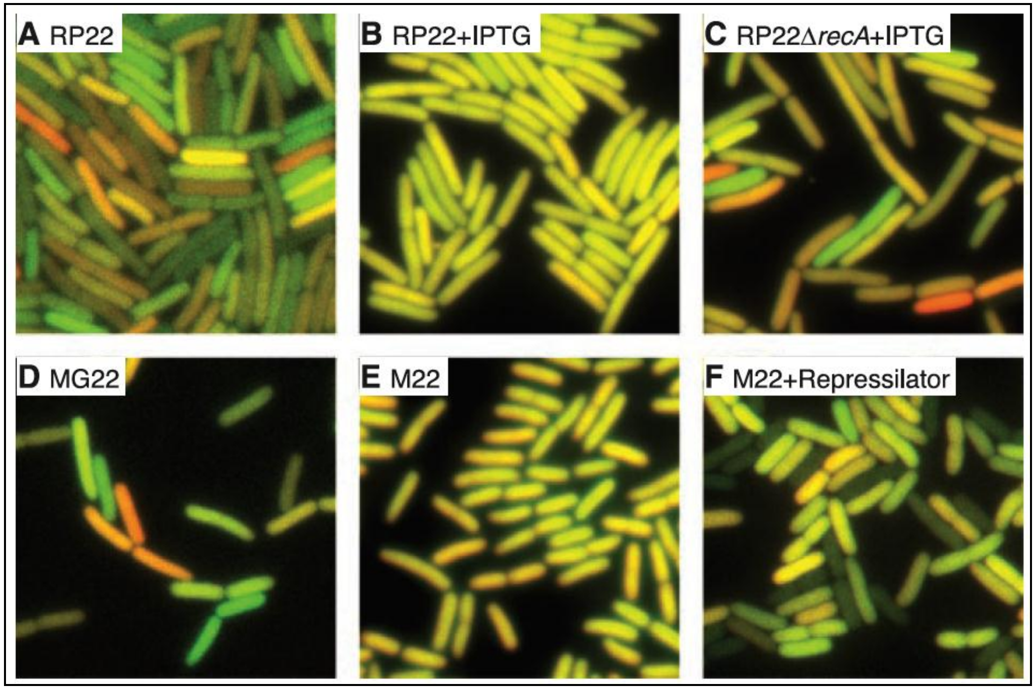


Experiment in *E. coli*

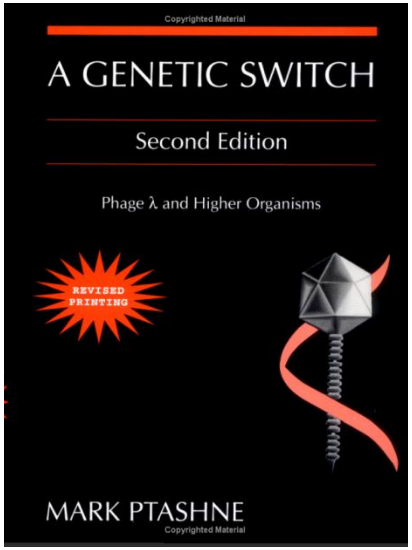



no intrinsic noise, only extrinsic noise

both intrinsic and extrinsic noise



Stochastic kinetic analysis of a developmental pathway bifurcation in phage-λ *Escherichia coli* cell



Copyrighted Material

A GENETIC SWITCH

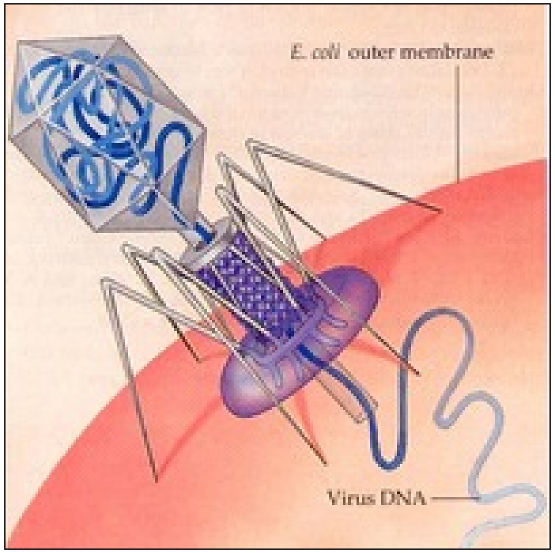
Second Edition

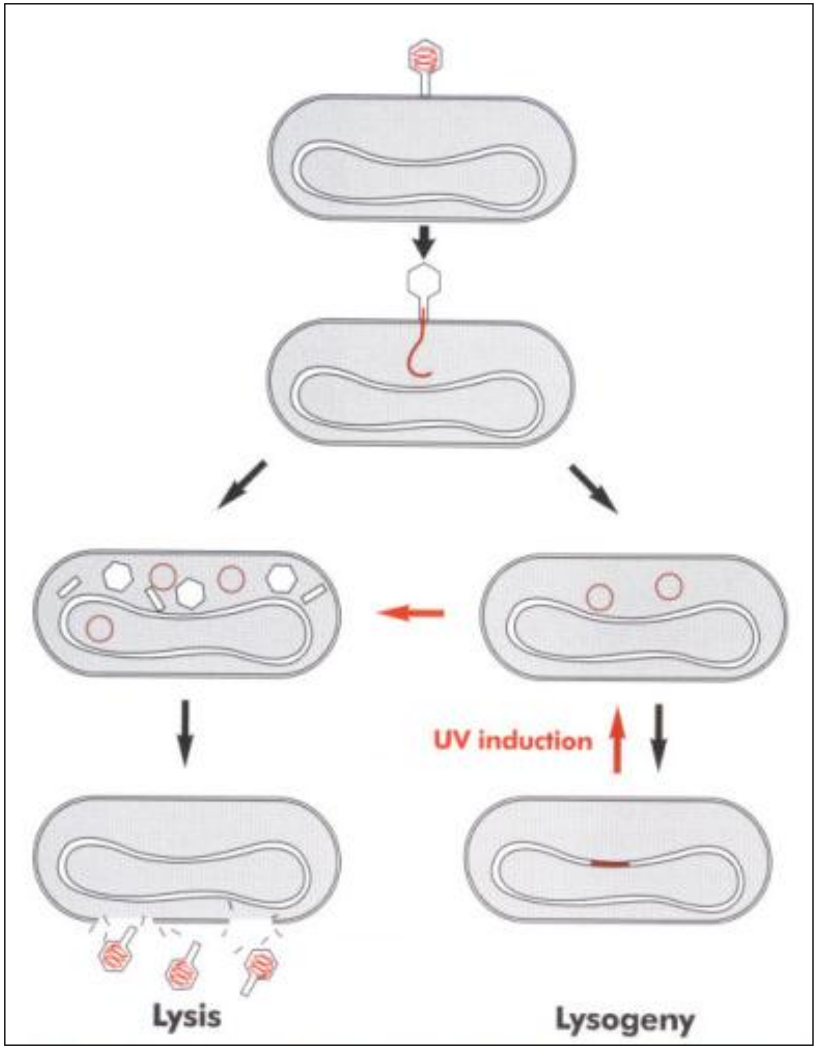
Phage λ and Higher Organisms

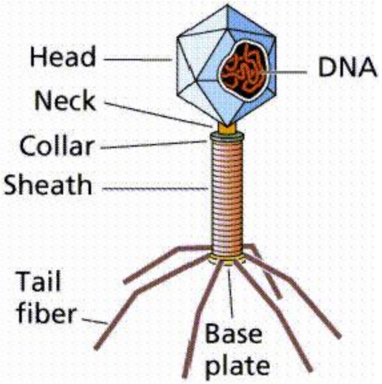
REVISED PRINTING

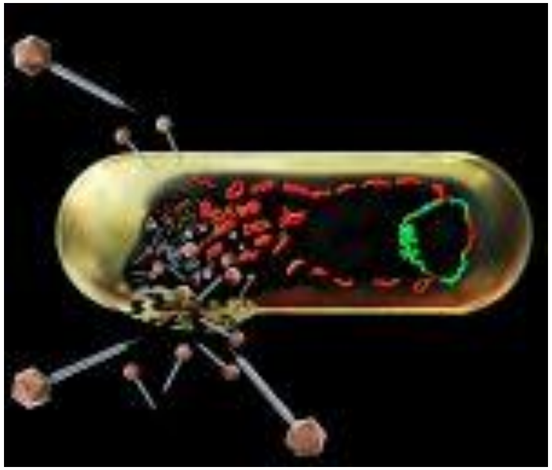
MARK PTASHNE

Copyrighted Material

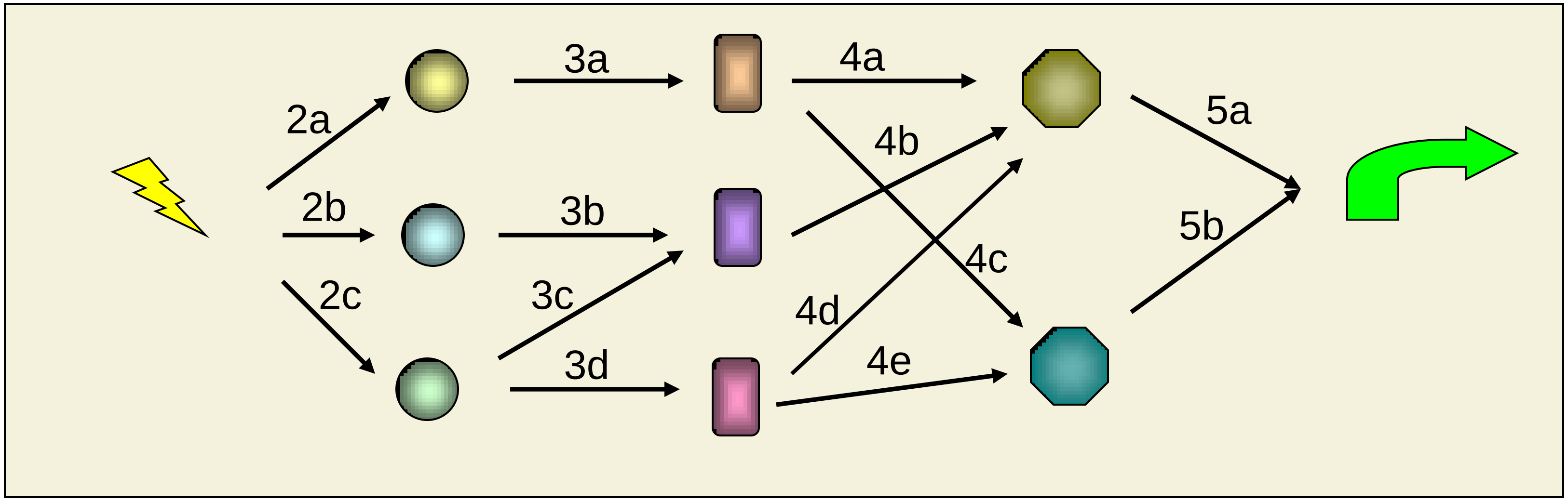








Stochasticity



1 → 2 → 3 → 4 → 5

stimulus

Signal transduction

TF activation

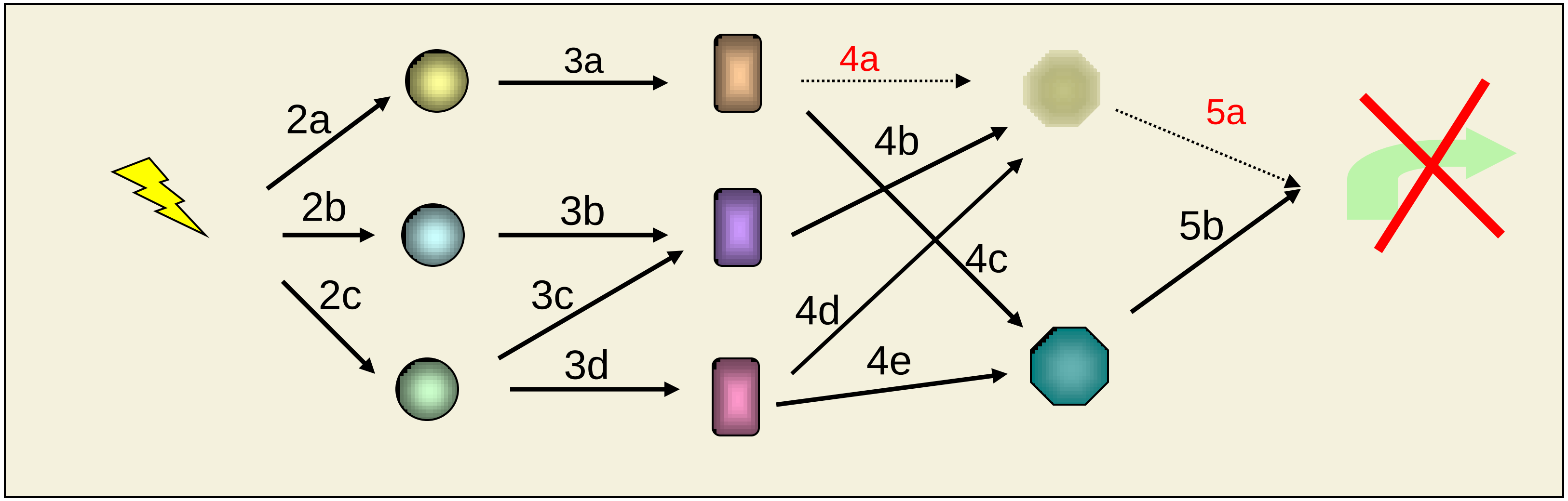
Cooperative binding on the enhancer of target gene

Gene activation

Every single step occurs in a probabilistic manner, it does not determine the next step but it is critical for the next step.

Minor changes in one step could affect dramatically the final result

Stochasticity



1 → 2 → 3 → 4 → 5

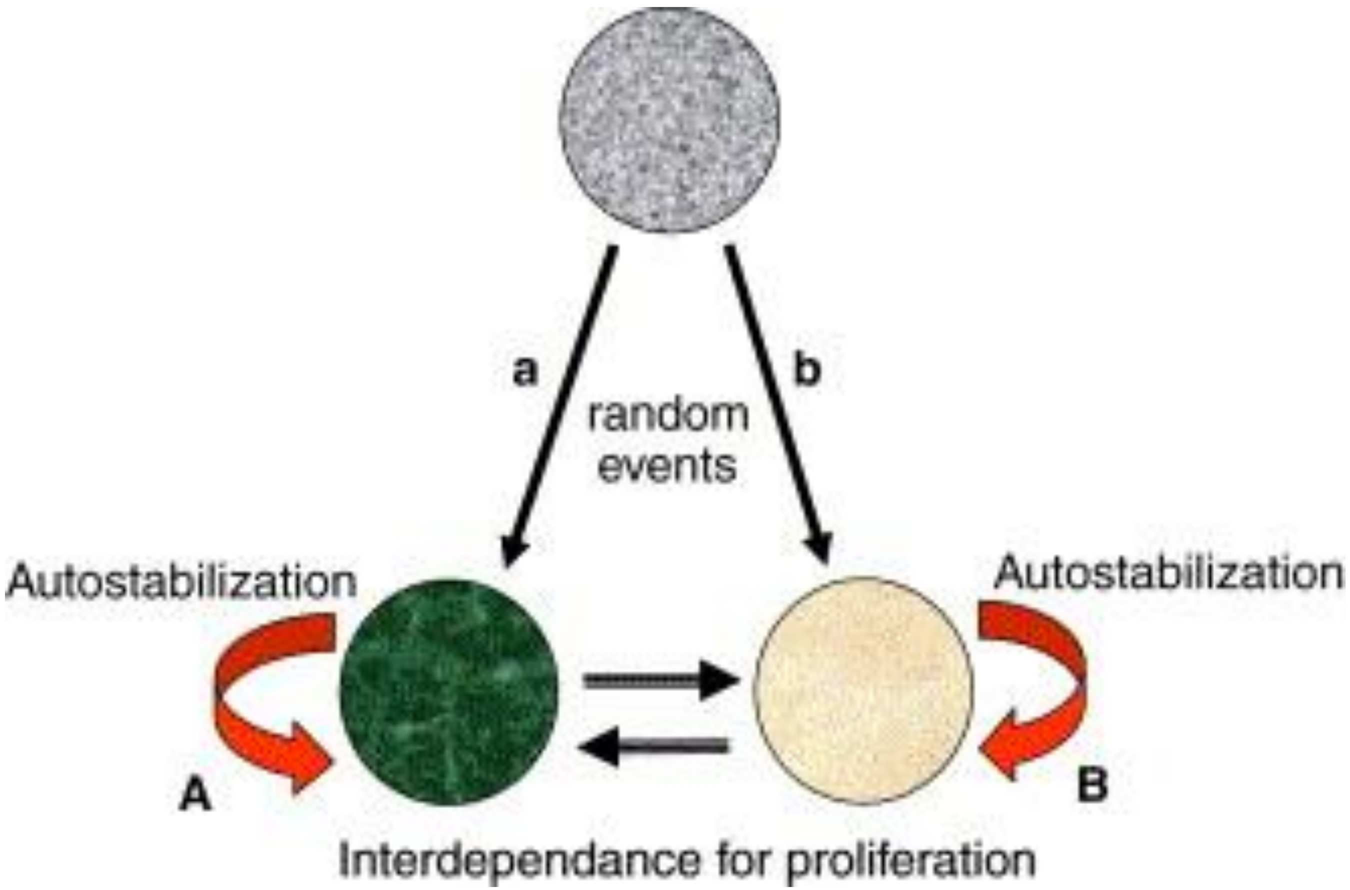
stimulus

Signal transduction

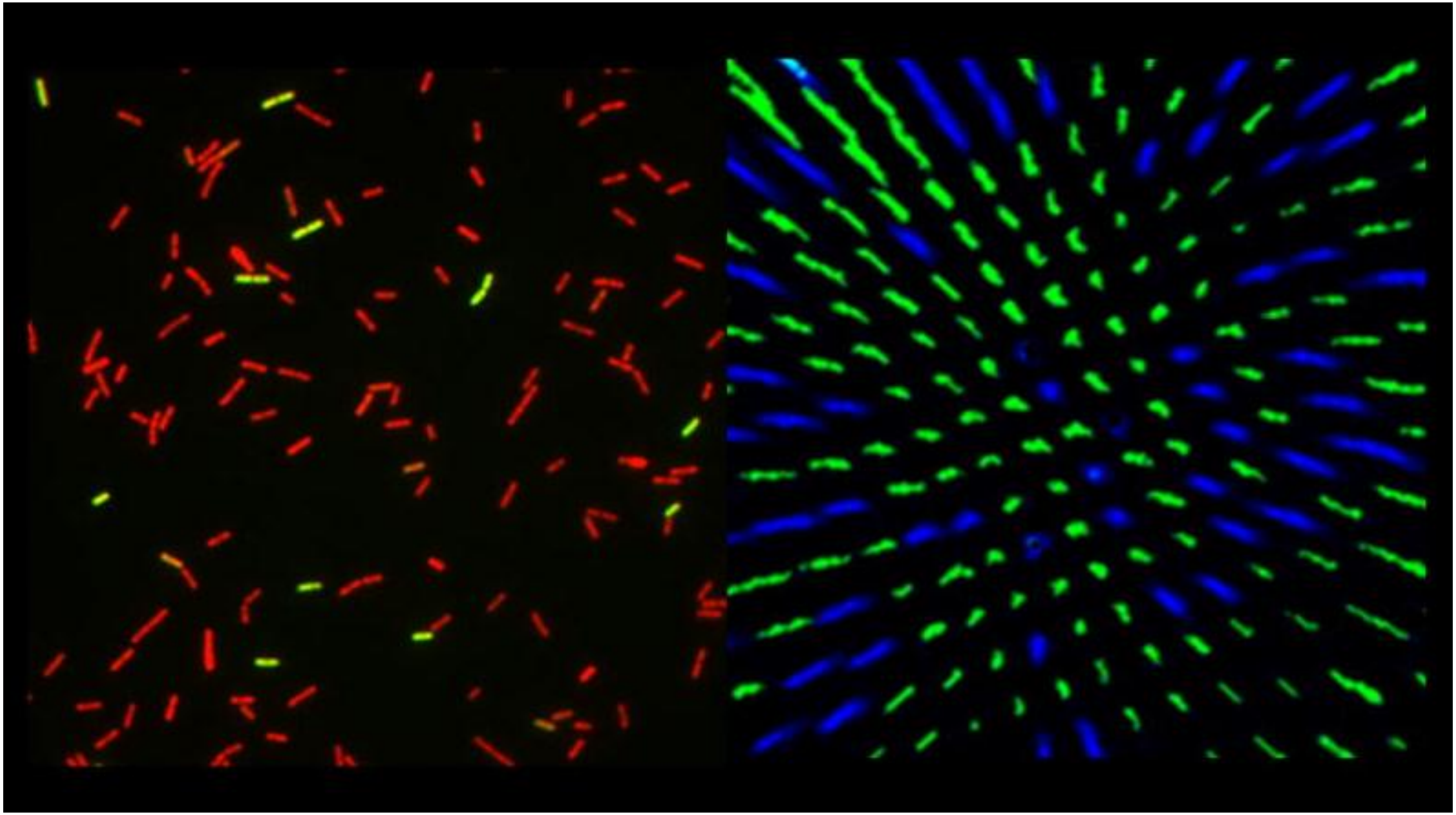
TF activation

Cooperative binding on the enhancer of target gene

Gene activation


a
b
random events
Autostabilization
Autostabilization
A
B
Interdependance for proliferation

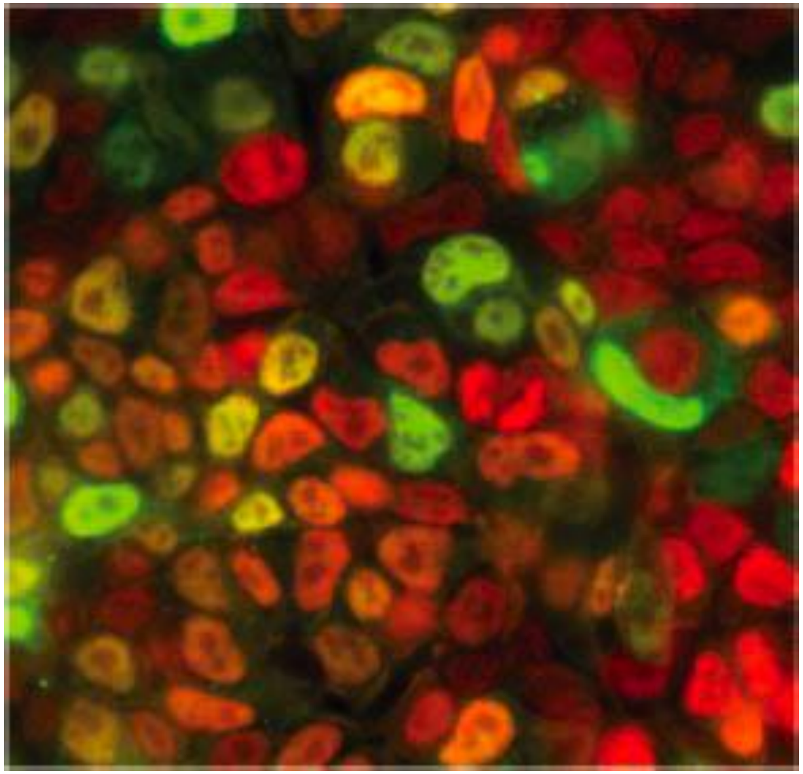
STOCHASTIC EXPRESSION-1



Bacteria

Drosophila Retina

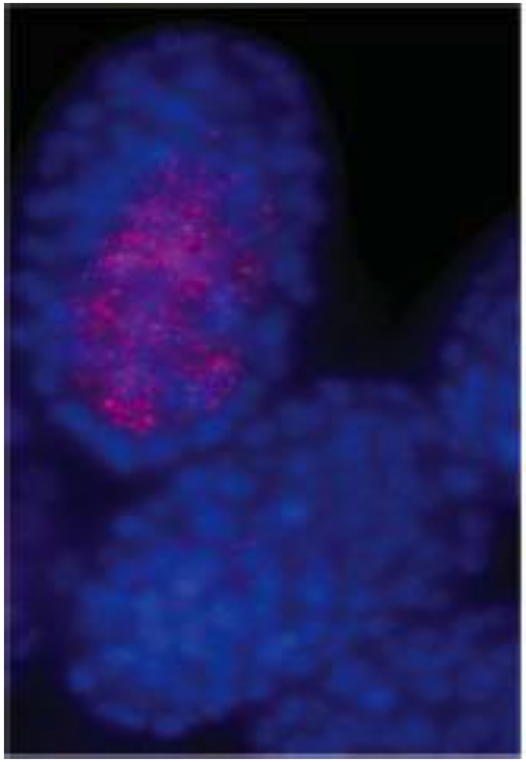
STOCHASTIC EXPRESSION-2



STEM CELLS

Oct4

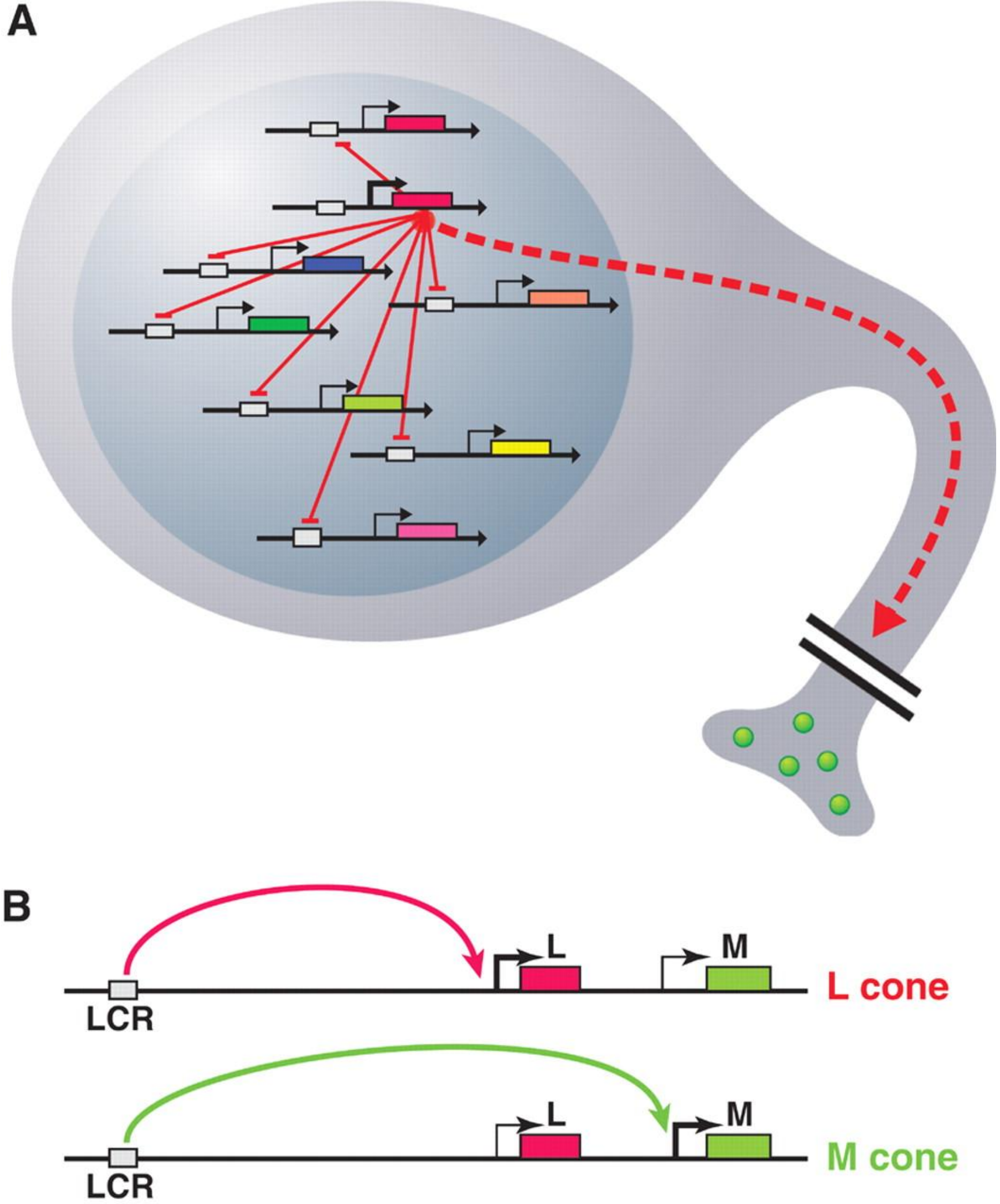
Nanog



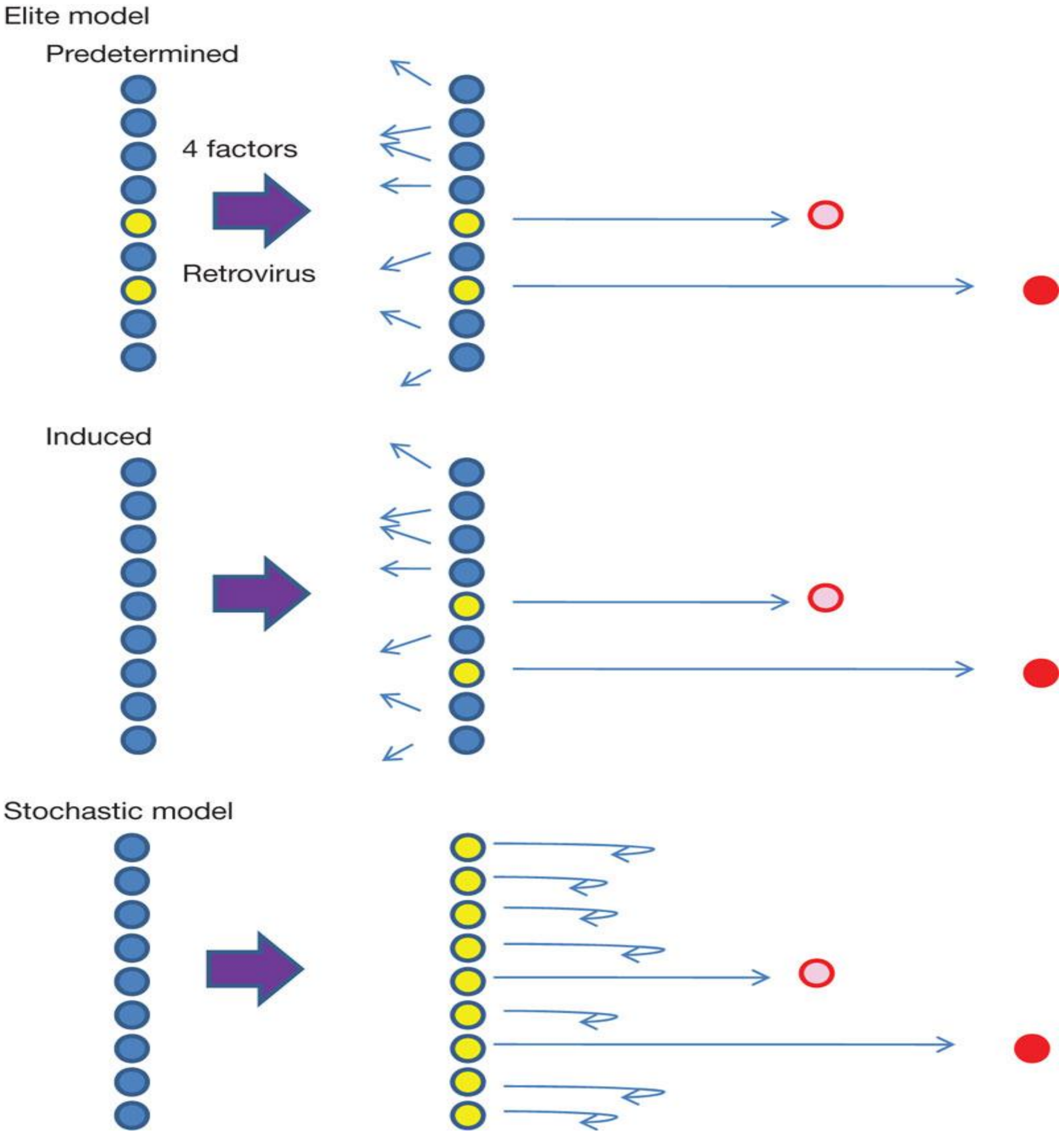
C. elegans skn-2

elt-1 → **gut**

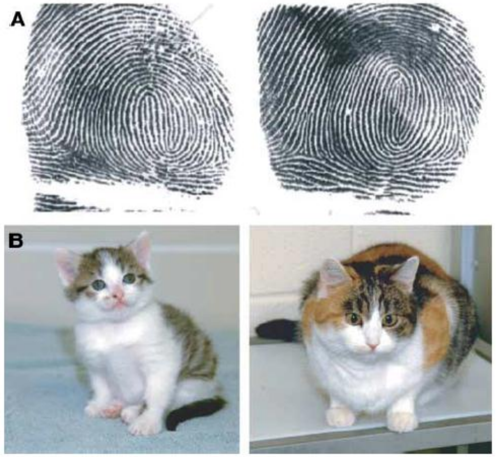
STOCHASTIC EXPRESSION-3



Two models explaining the low efficiency of iPS cell generation.



Stochastic influences on phenotype



Noise in Gene Expression: Origins, Consequences, and Control

Jonathan M. Raser[1,2] and Erin K. O'Shea[2]*†

Genetically identical cells and organisms exhibit remarkable diversity even when they have identical histories of environmental exposure. Noise, or variation, in the process of gene expression may contribute to this phenotypic variability. Recent studies suggest that this noise has multiple sources, including the stochastic or inherently random nature of the biochemical reactions of gene expression. In this review, we summarize noise terminology and comment on recent investigations into the sources, consequences, and control of noise in gene expression.

Any individual in a population of living organisms or cells is unique. Much of population variability is due to genetic differences, but environment and history also contribute to variability in cellular phenotype. Indeed, identical twin humans or cloned cats differ in appearance and behavior (Fig. 1). However, even cells or organisms with the same genes, in the same environment, with the same history, display variations in form and behavior that can be subtle or dramatic. Investigations have focused on the possibility that such variability is inevitable in biological systems because of the random nature of chemical reactions within a cell (*1*). When large numbers of molecules are present, chemical reactions may proceed in a predictable manner. However, when only a few molecules of a specific type exist in a cell, stochastic effects can become prominent.

Gene expression, as defined by the set of reactions that control the abundance of gene products, influences most aspects of cellular behavior, and its variation is often invoked to explain phenotypic differences in a population of cells. Because DNA, RNA, and proteins can be present and active at a few copies per cell, the abundance of gene products is theoretically sensitive to stochastic fluctuations. Four potential sources of variation in gene expression must be considered: (i) as described above, the inherent stochasticity of biochemical processes that are dependent on infrequent molecular events involving small numbers of molecules; (ii) variation in gene expression owing to differences in the internal states of a population of cells, either from predictable processes such as cell cycle progression or from a random process such as partitioning of mitochondria during cell division; (iii) subtle environmental differences, such as morphogen gradients in multicellular development; and (iv) ongoing genetic mutation, either random or directed. We use the term "noise" in gene expression to refer to the measured level of variation in gene expression among cells, regardless of source, within a supposedly identical population.

Fig. 1. Examples of possible stochastic influences on phenotype. (A) The fingerprints of identical twins are readily distinguished on close examination. Reprinted from (*37*) with permission from Elsevier. (B) Cc, the first cloned cat (left) and Rainbow, Cc's genetic mother (right), display different coat patterns and personalities (*38*). Photo credit, College of Veterinary Medicine and Biomedical Sciences, Texas A&M University.

Measurement Techniques and Definitions

Recent investigations have employed green fluorescent protein (GFP) variants, which allow the quantification of protein levels in living cells by flow cytometry or fluorescence microscopy. The coefficient of variation, or noise η, is defined as the ratio of the standard deviation to the mean of the population. Other metrics of variability can be useful as well (SOM Text).

Once genetic mutation and local microenvironments are eliminated as sources of noise, an elegant experimental method can assist in differentiating among the remaining sources (*2*). This method involves quantifying expression of two equivalent, independent gene reporters placed in the same cell, which then allows noise sources to be partitioned into two categories: intrinsic, meaning noise sources that create differences between the two reporters within the same cell (Fig. 2A), and extrinsic, referring to sources that affect the two reporters equally in any given cell but create differences between two cells (Fig. 2B). Stochastic events during the process of gene expression, from the level of promoter-binding to mRNA translation to protein degradation, will manifest as intrinsic noise. Differences between cells, either in local environment or in the concentration or activity of any factor that affects gene expression, will result in extrinsic noise. Extrinsic noise should be further subdivided into two categories (*3*, *4*): global noise, or fluctuations in the rates of the basic reactions that affect expression of all genes (Fig. 2C), and gene- or pathway-specific extrinsic noise (Fig. 2D), such as fluctuations in the abundance of a particular transcription factor or stochastic events in a specific signal transduction pathway. If a factor that causes extrinsic noise is experimentally manipulable, it is possible to eliminate such extrinsic noise by reduction of variability in that factor; for example, cell cycle synchronization will reduce extrinsic

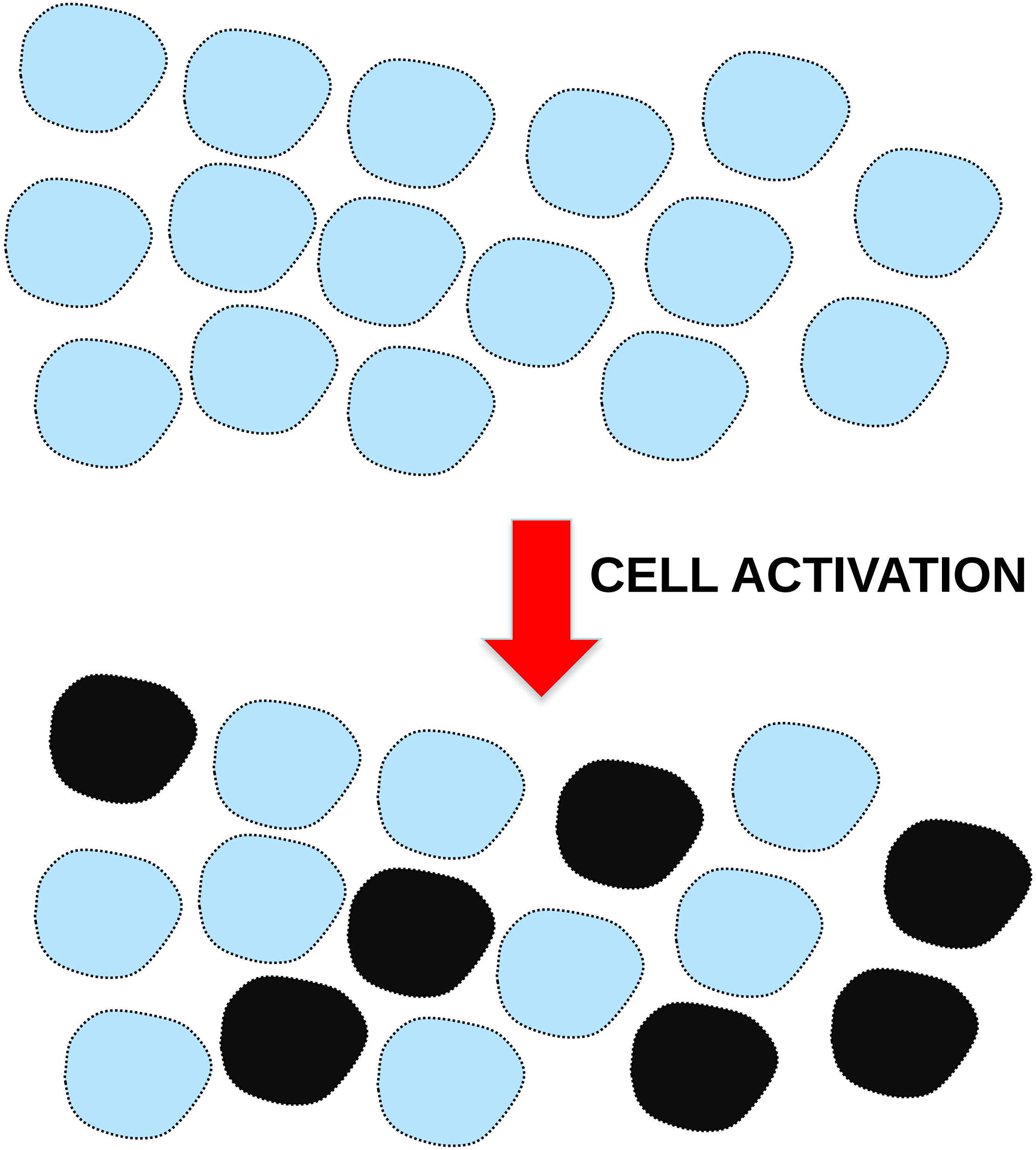
---

[1]Medical Scientist Training Program, [2]Howard Hughes Medical Institute, University of California–San Francisco, 600 16th Street, GH-S472D, San Francisco, CA 94143–2240, USA.

*Present address: Howard Hughes Medical Institute, Harvard University, 7 Divinity Avenue, Bauer 307, Cambridge, MA 02138, USA.

†To whom correspondence should be addressed. E-mail: erin_oshea@harvard.edu

STOCHASTIC EXPRESSION OF CYTOKINES



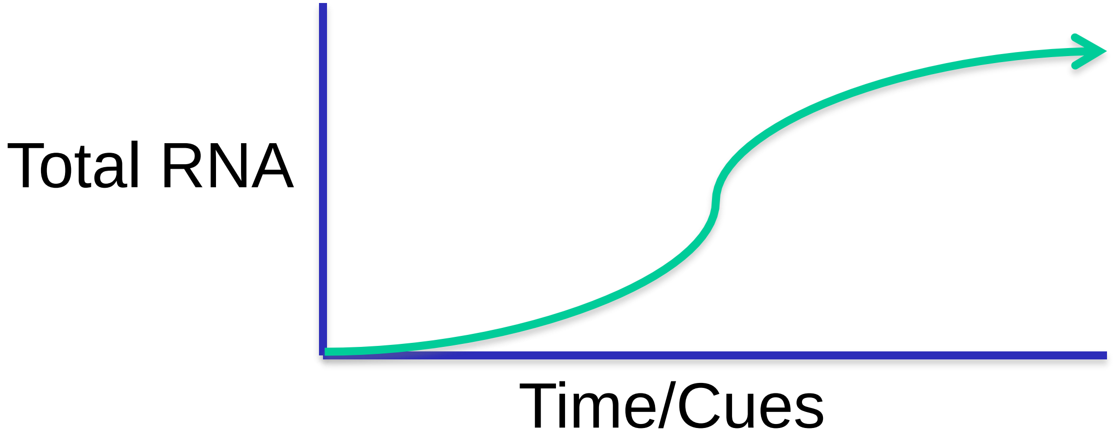
IFNβ
IFNγ
IFNα
IL-1
IL-2
IL-4
IL-6
IL-8
…….

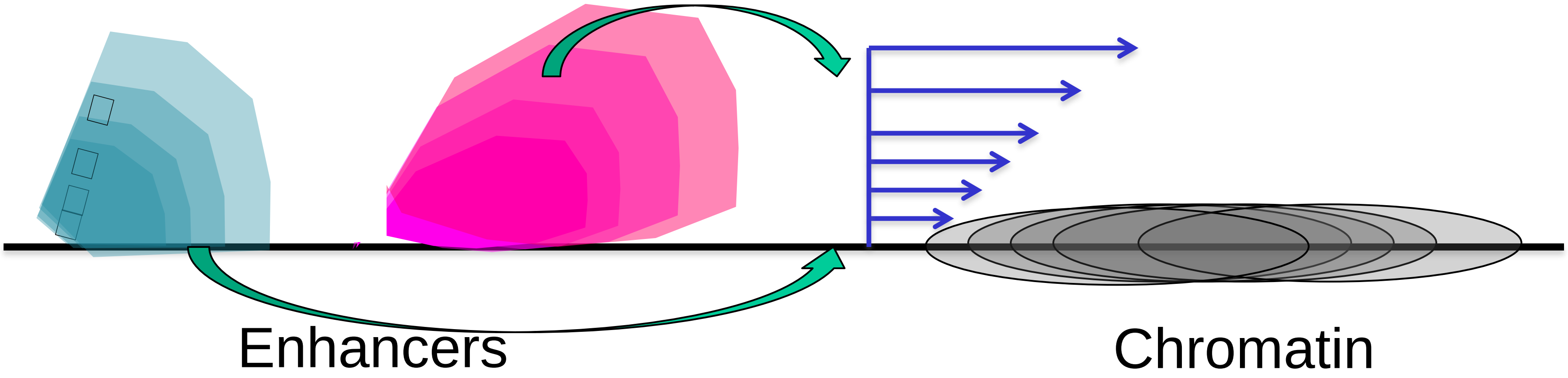
Stochasticity requires

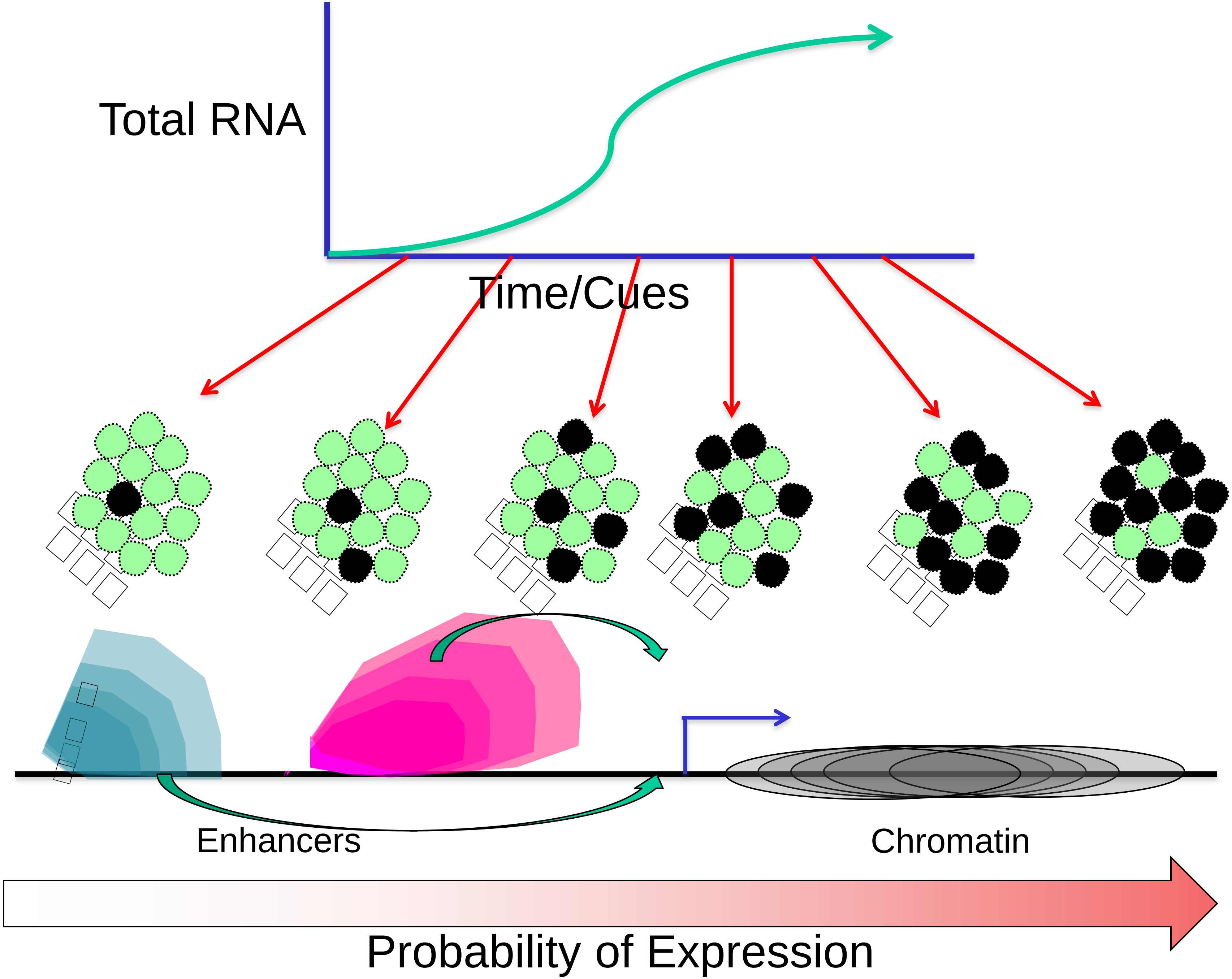
- A means to generate noise
(e.g biochemical reactions, protein-DNA interactions)

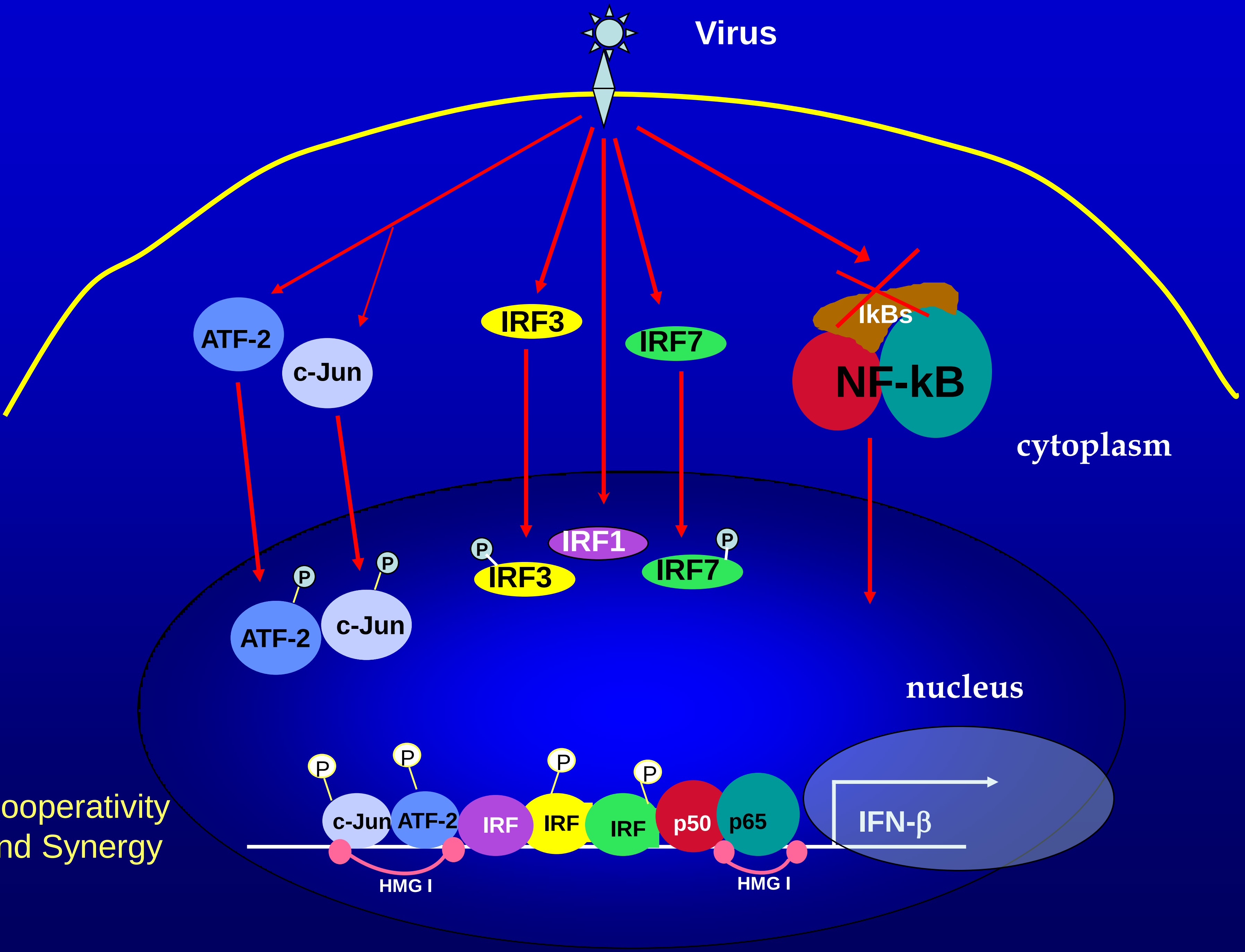
- Mechanisms to amplify the noise
(e.g hypersensitivity-cooperativity)

- Mechanisms to stabilize stochastic decisions
(e.g transcription factors, epigenetics)


Total RNA
Time/Cues

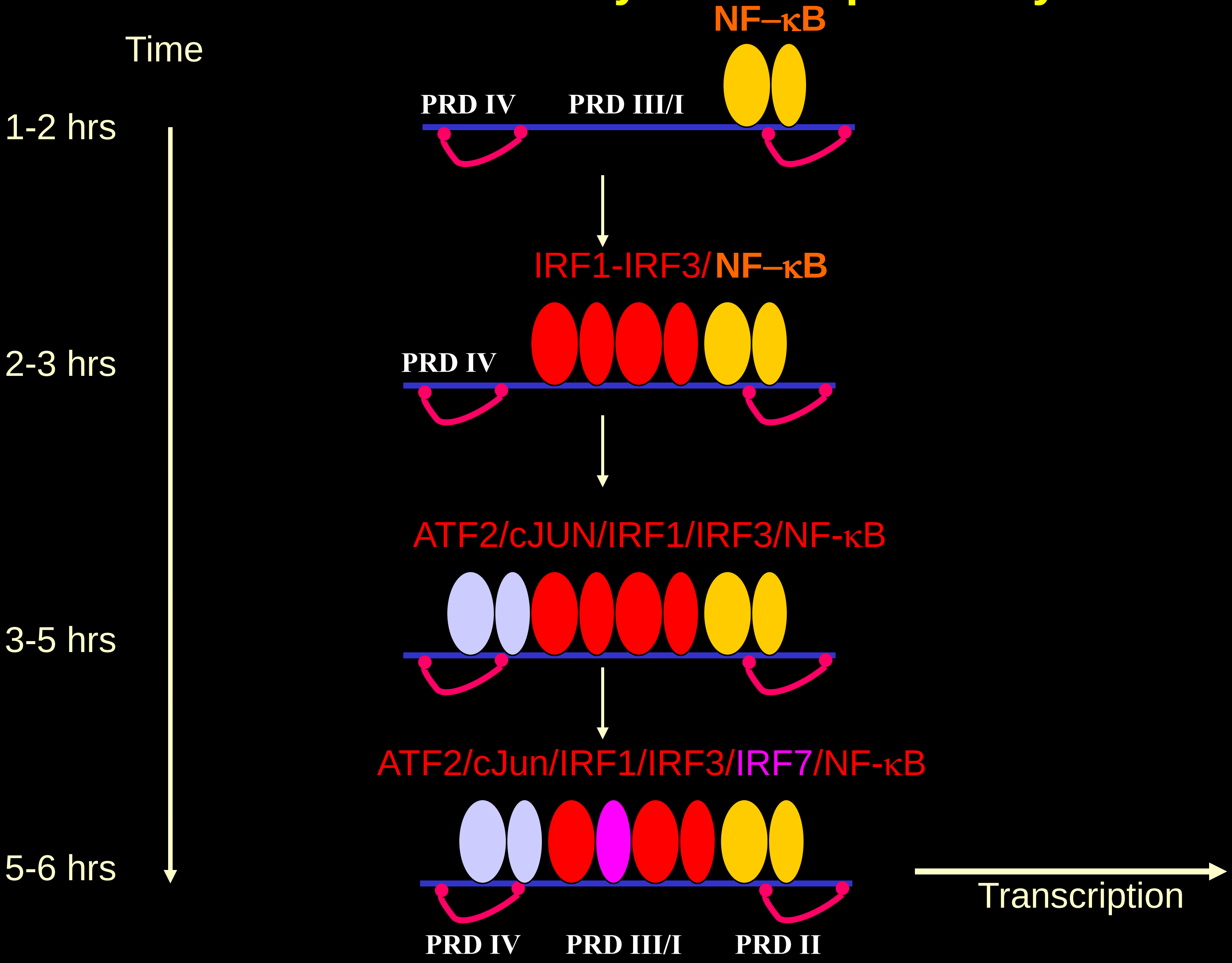

Enhancers
Chromatin


Total RNA
Time/Cues
Enhancers
Chromatin
Probability of Expression



Cooperativity
and Synergy

Allostery and Cooperativity



Time

1-2 hrs

NF–κB

PRD IV

PRD III/I

2-3 hrs

IRF1-IRF3/NF–κB

PRD IV

3-5 hrs

ATF2/cJUN/IRF1/IRF3/NF-κB

5-6 hrs

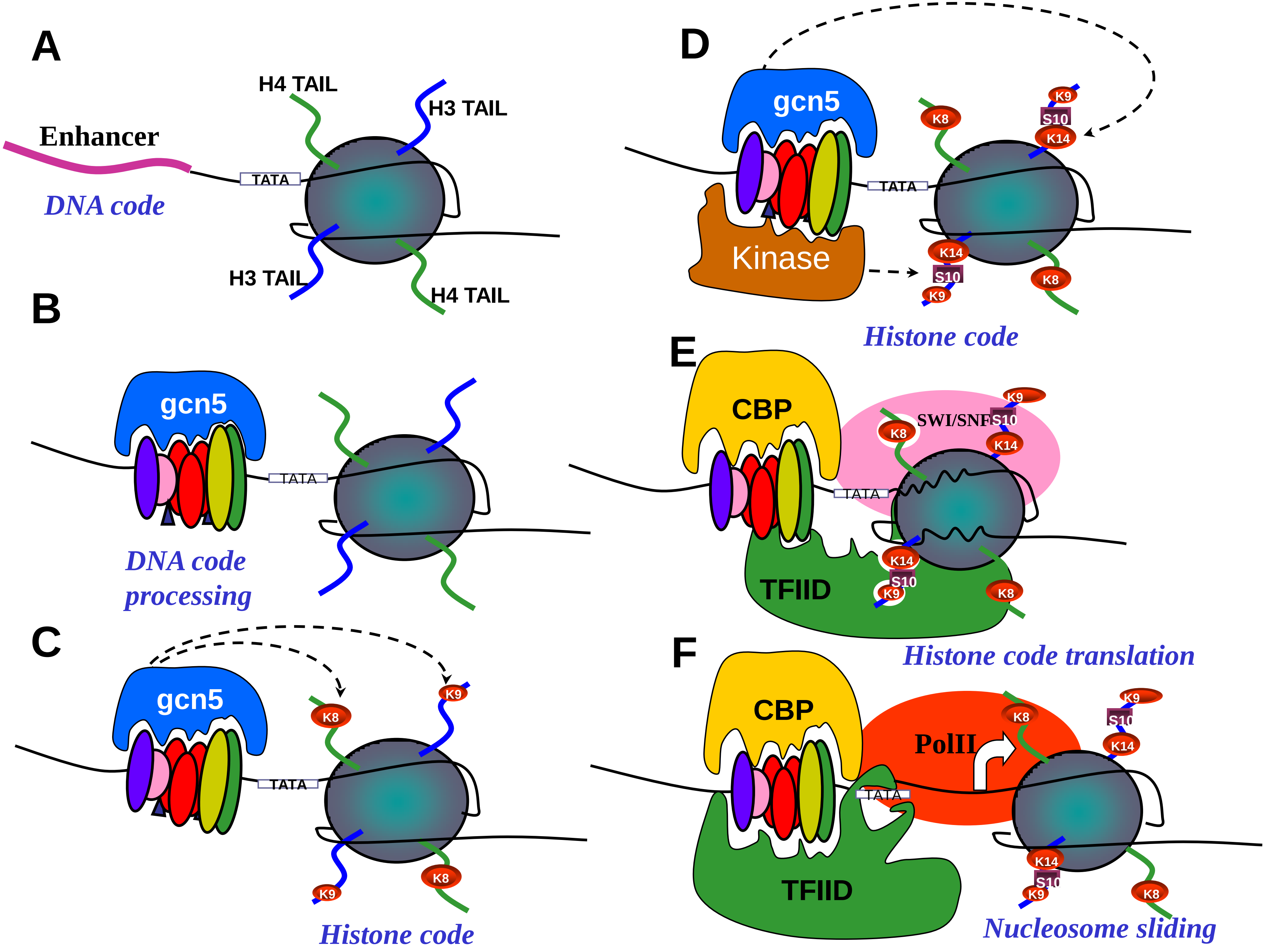
ATF2/cJun/IRF1/IRF3/IRF7/NF-κB

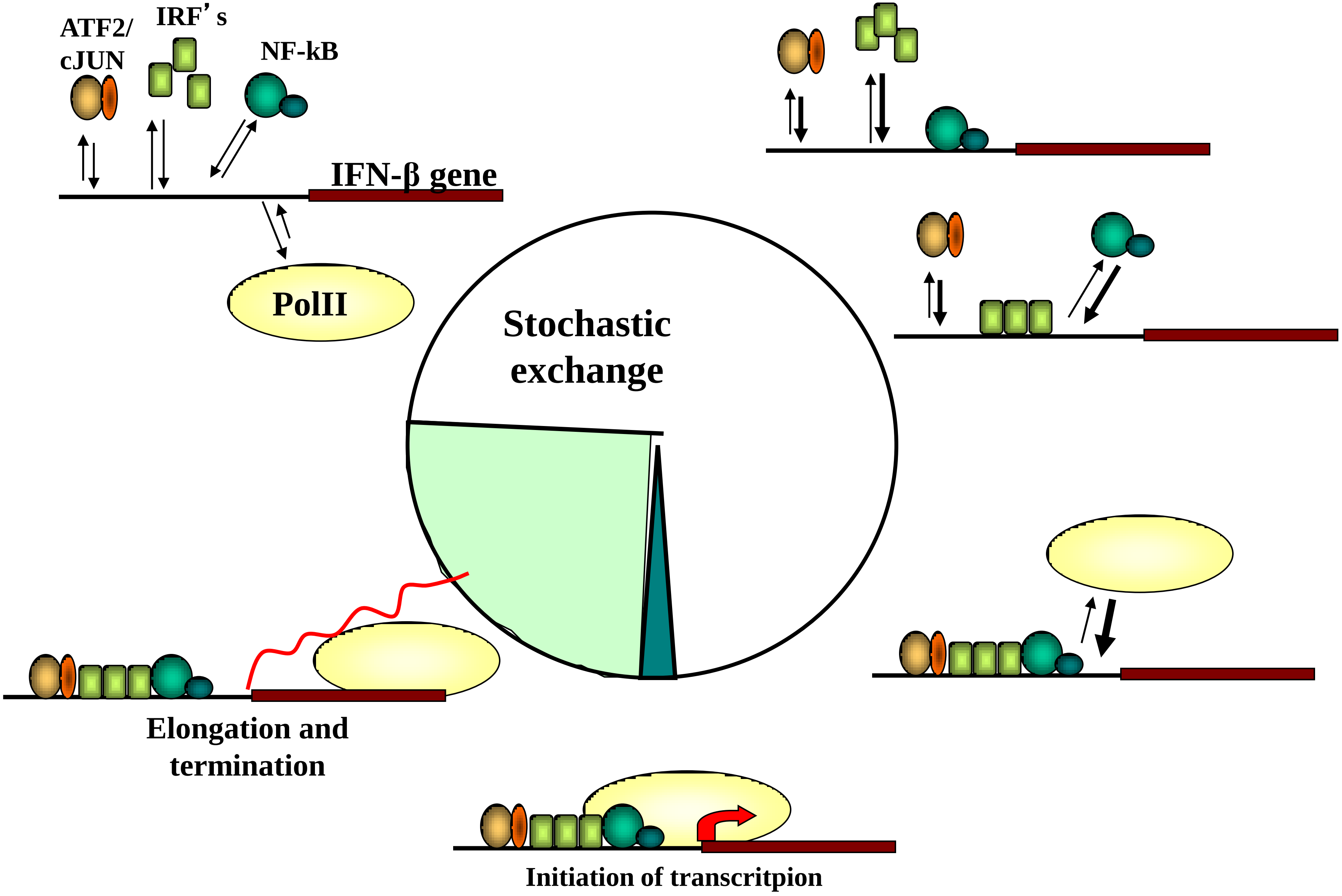
PRD IV

PRD III/I

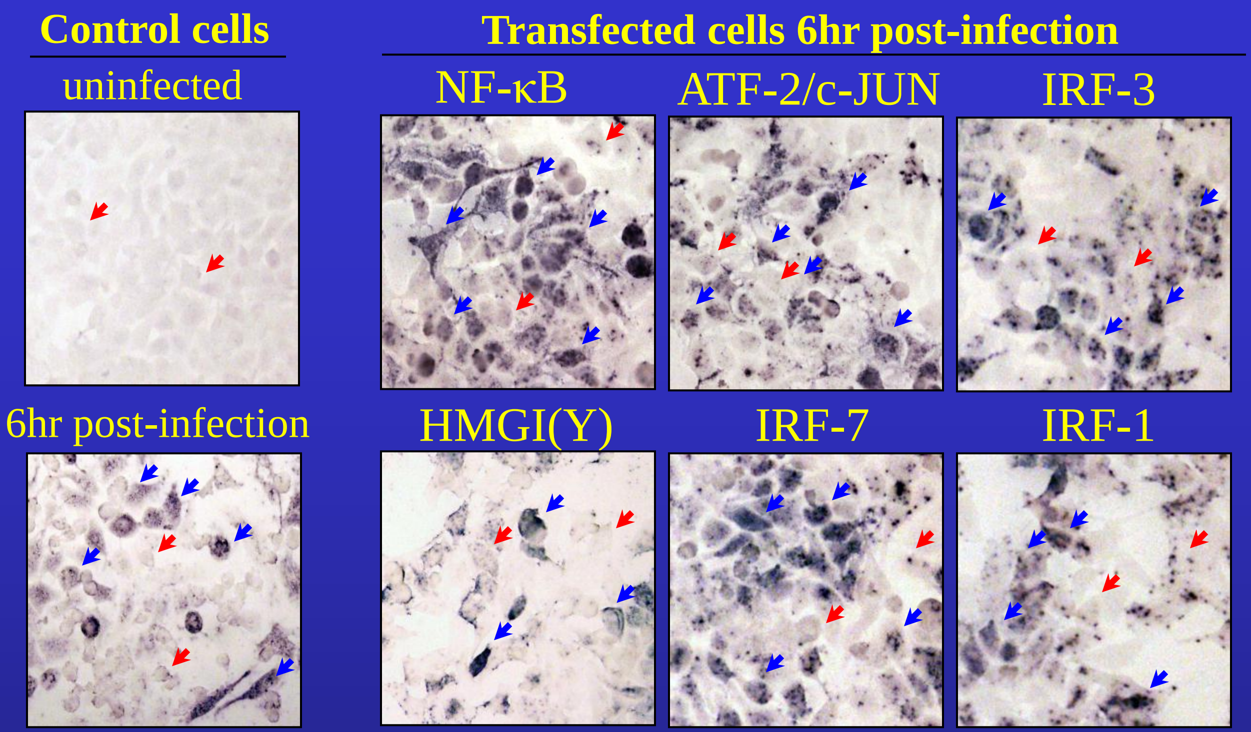
PRD II

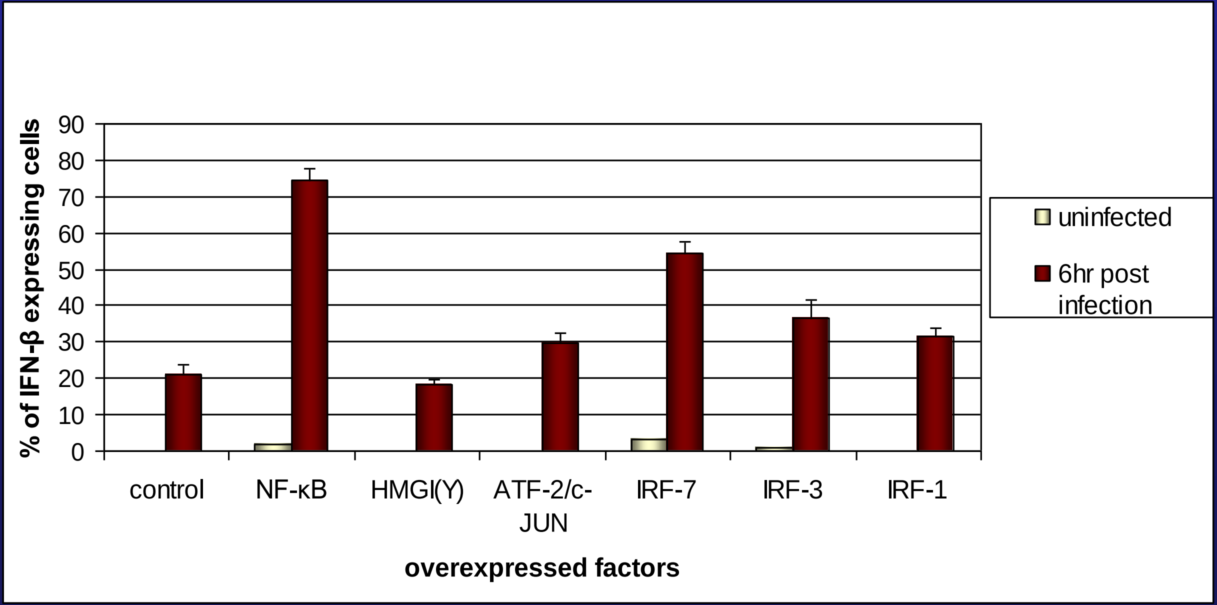
Transcription

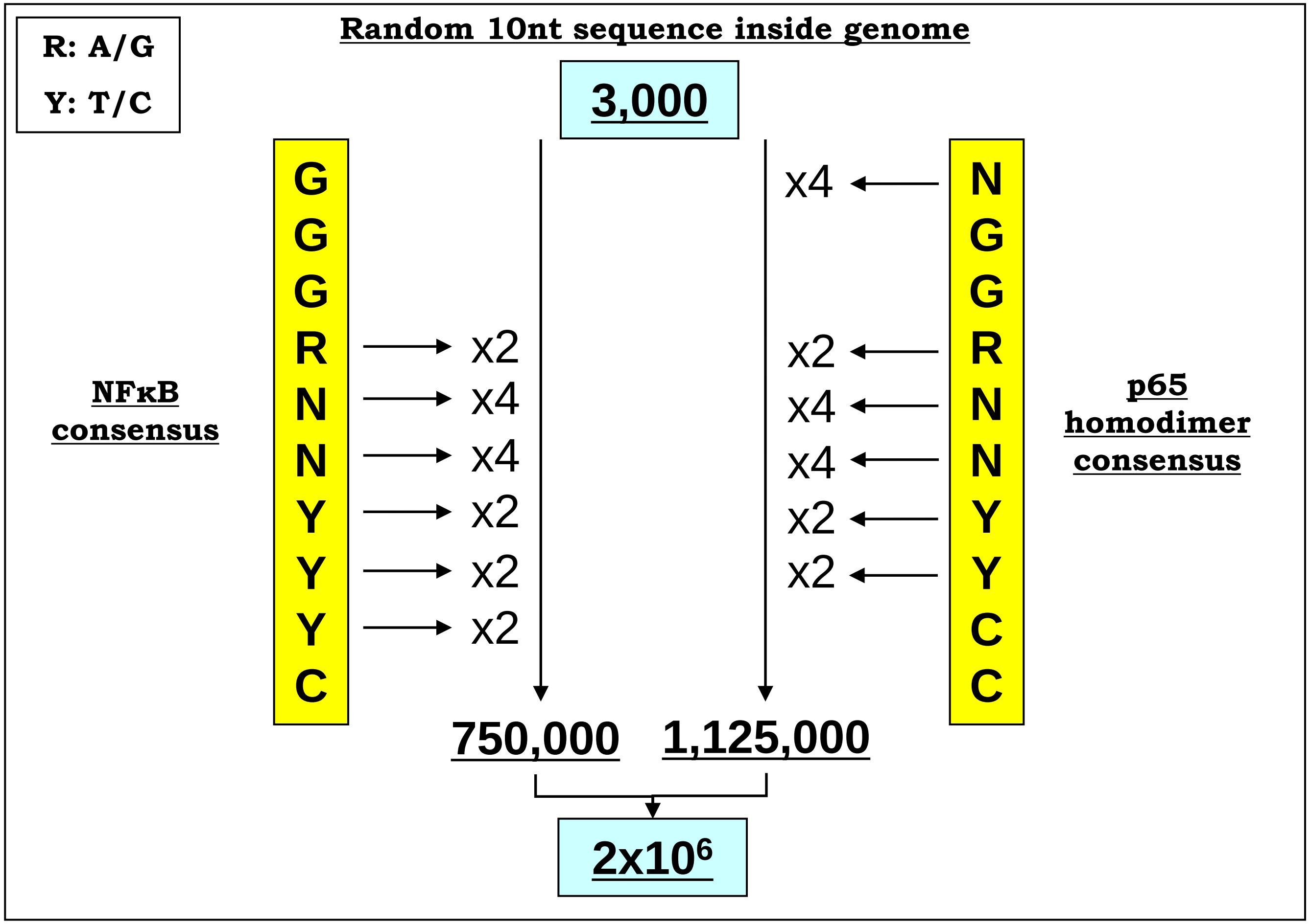

A
Enhancer
DNA code
TATA
H4 TAIL
H3 TAIL
H3 TAIL
H4 TAIL
B
gcn5
TATA
DNA code processing
C
gcn5
TATA
K8
K9
K9
K8
Histone code
D
gcn5
Kinase
TATA
K9
S10
K14
K8
K14
S10
K9
K8
Histone code
E
CBP
SWI/SNF
TATA
TFIID
K9
S10
K8
K14
K14
S10
K9
K8
Histone code translation
F
CBP
PolII
TATA
TFIID
K9
S10
K14
K8
K14
S10
K9
K8
Nucleosome sliding


ATF2/
cJUN
IRF’s
NF-kB
IFN-β gene
PolII
Stochastic
exchange
Elongation and
termination
Initiation of transcritpion

ISH experiments for IFN-β mRNA







R: A/G

Y: T/C

Random 10nt sequence inside genome

3,000

NFκB consensus

G G G R N N Y Y Y C

R → x2

N → x4

N → x4

Y → x2

Y → x2

Y → x2

p65 homodimer consensus

N G G R N N Y Y C C

N → x4

R → x2

N → x4

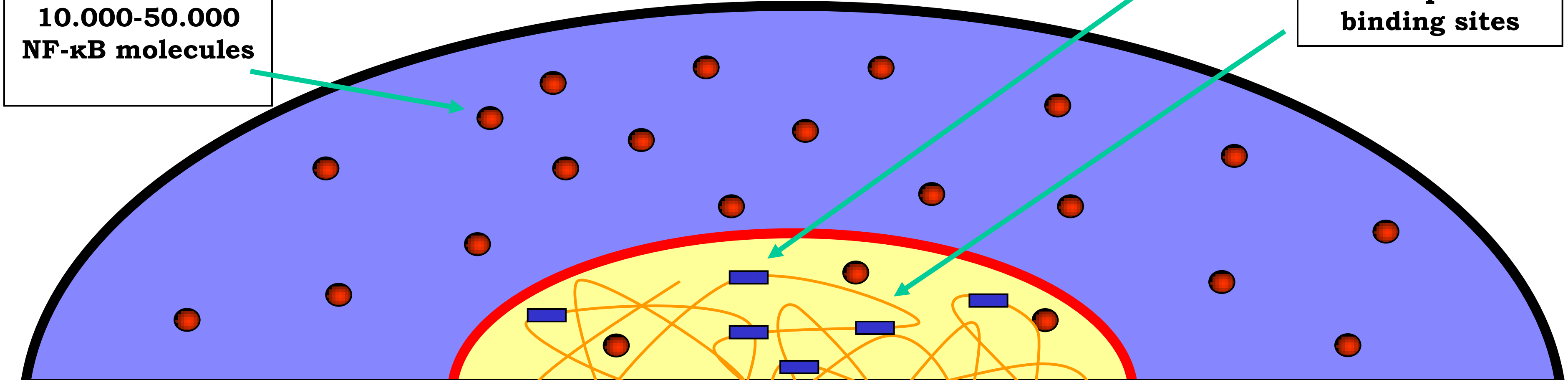
N → x4

Y → x2

Y → x2

750,000 **1,125,000**

$2x10^6$

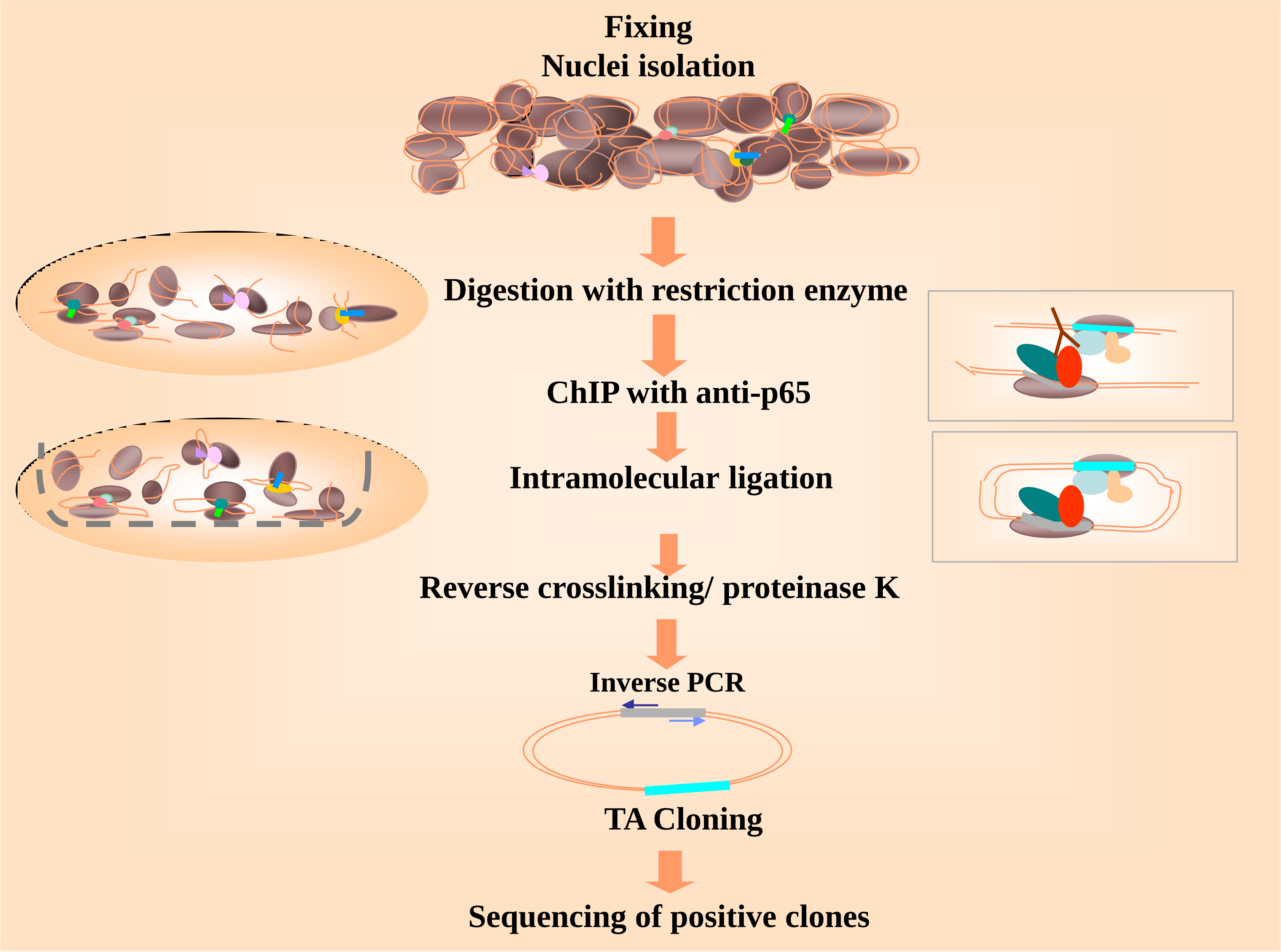


10.000-50.000 NF-κB molecules

~500 target genes

~$2x10^6$ putative binding sites

3C (Chromosome Conformation Capture)-ChIP Cloning



Fixing

Nuclei isolation

Digestion with restriction enzyme

ChIP with anti-p65

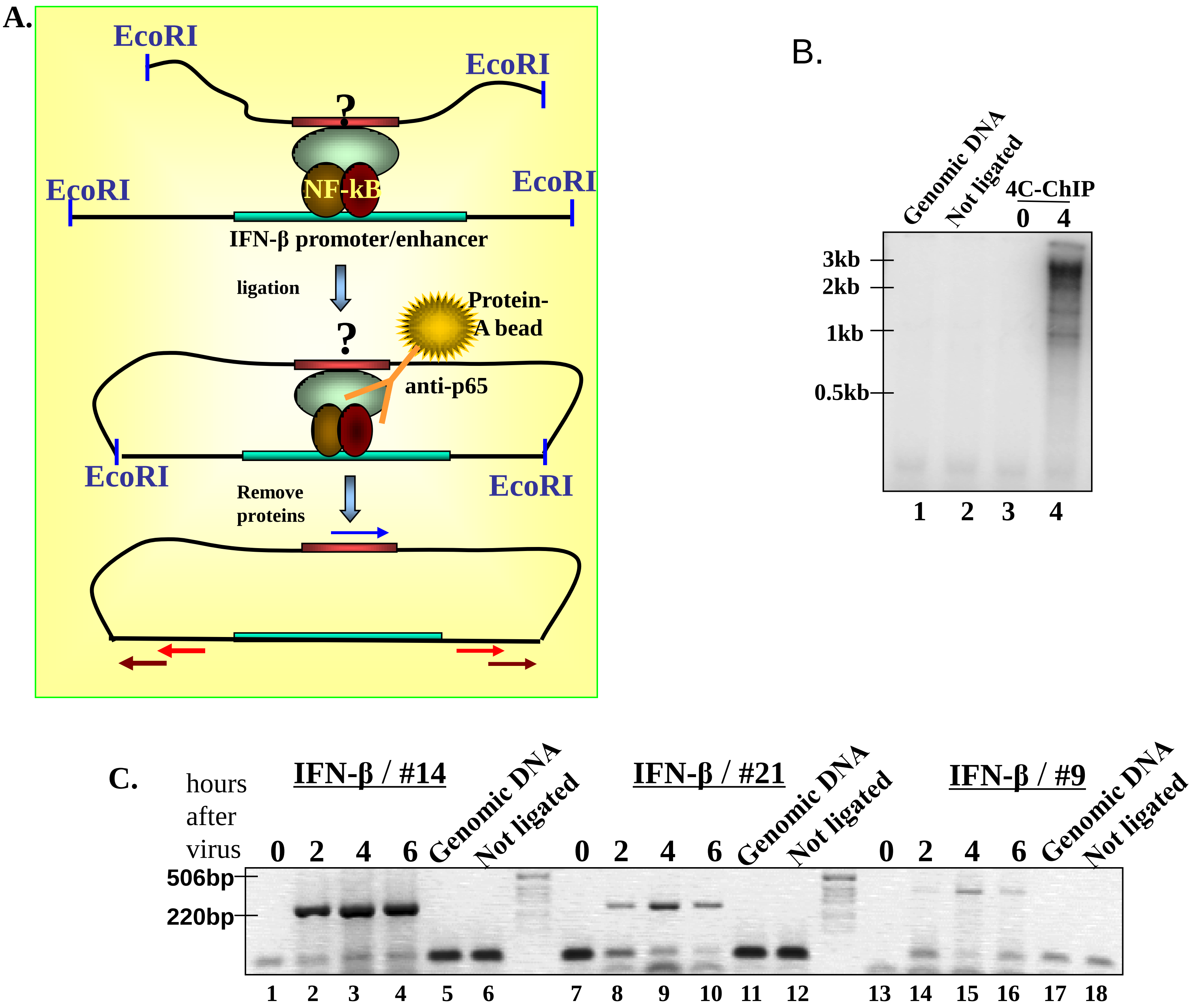
Intramolecular ligation

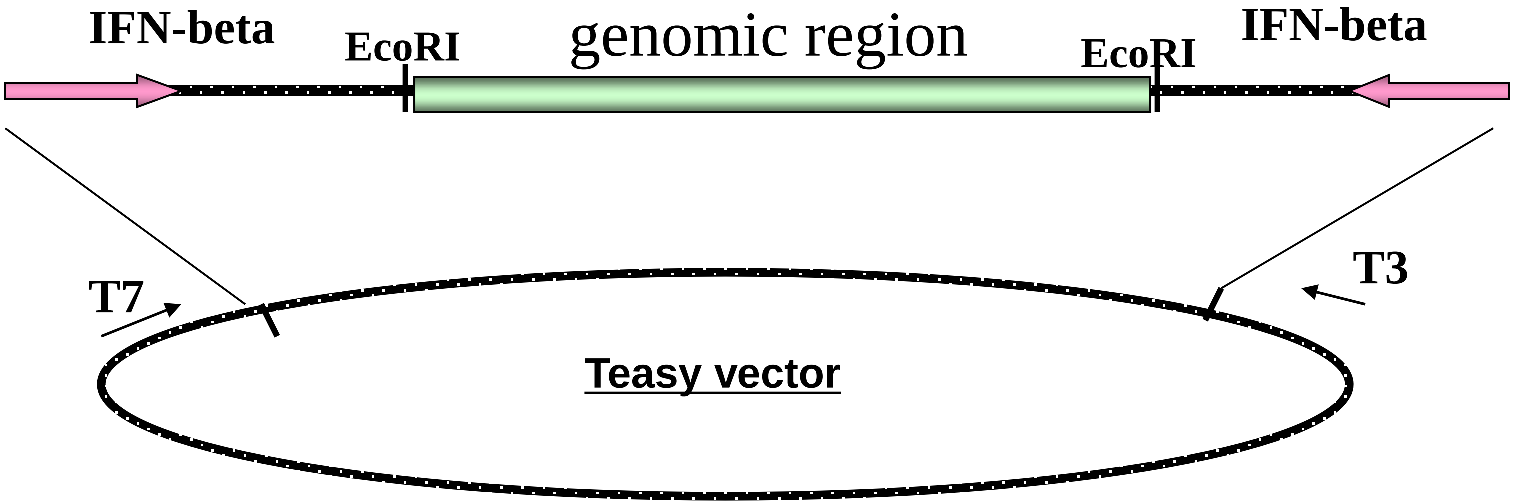
Reverse crosslinking/ proteinase K

Inverse PCR

TA Cloning

Sequencing of positive clones

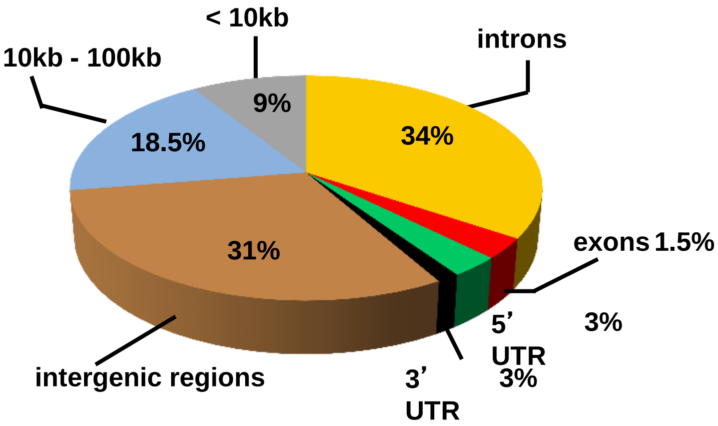

A.
EcoRI
EcoRI
?
NF-kB
EcoRI
EcoRI
IFN-β promoter/enhancer
ligation
?
Protein-A bead
anti-p65
EcoRI
EcoRI
Remove proteins
B.
Genomic DNA
Not ligated
4C-ChIP
0
4
3kb
2kb
1kb
0.5kb
1
2
3
4
C.
hours after virus
IFN-β / #14
0
2
4
6
Genomic DNA
Not ligated
IFN-β / #21
0
2
4
6
Genomic DNA
Not ligated
IFN-β / #9
0
2
4
6
Genomic DNA
Not ligated
506bp
220bp
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

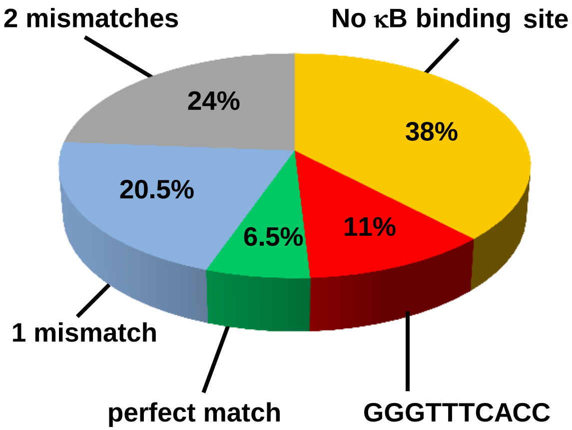


NF-κB Reception Centers-NRCs

4C-ChIP clones	*Precise position*	*Nearby genes*	*Alu element*
#21	4 p13	19kb upstream of **RHOH** 20kb downstream of **N4BP2**	Alu SX (**GGGTTTCACC**)
#N14	9q33	10kb upstream of **ABO**, 15kb downstream of **LCN1** (**lipocalin 1**)	Alu Y (**GGGTTTCACC**)
#K9	18q21	12kb upstream of **LOXHD1**	Alu Sx (**GGGTTTCACC**)

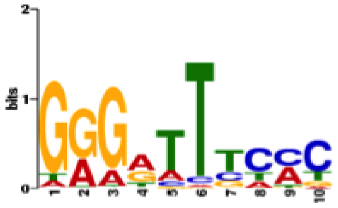
GENOME WIDE DISTRIBUTION OF NF-kB IN THE HUMAN GENOME



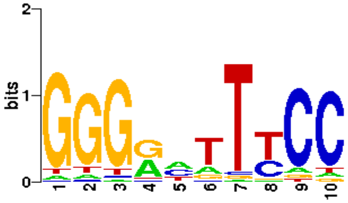


HUMAN	**GGGT*T*TCACC**	**GGGT*A*TCACC**
Chromosome 1	14726	203
Chromosome 2	12408	217
Chromosome 3	10007	187
Chromosome 4	8139	189
Chromosome 5	8545	169
Chromosome 6	8651	160
Chromosome 7	10037	163
Chromosome 8	7070	117
Chromosome 9	7497	100
Chromosome 10	8011	113
Chromosome 11	7106	126
Chromosome 12	8437	127
Chromosome 13	4409	69
Chromosome 14	5383	97
Chromosome 15	5865	62
Chromosome 16	6842	60
Chromosome 17	8449	58
Chromosome 18	3564	61
Chromosome 19	8220	47
Chromosome 20	4299	52
Chromosome 21	1879	18
Chromosome 22	3573	20
Chromosome X	6930	196
Chromosome Y	709	26

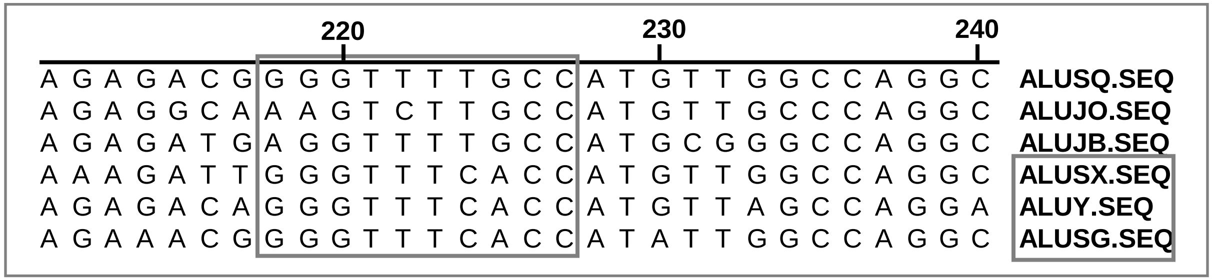
MOUSE	**GGGT*T*TCACC**	**GGGT*A*TCACC**
Chromosome 1	109	68
Chromosome 2	109	70
Chromosome 3	81	64
Chromosome 4	80	56
Chromosome 5	77	44
Chromosome 6	87	45
Chromosome 7	79	45
Chromosome 8	65	49
Chromosome 9	83	53
Chromosome 10	91	58
Chromosome 11	79	43
Chromosome 12	58	42
Chromosome 13	59	50
Chromosome 14	60	37
Chromosome 15	58	33
Chromosome 16	47	29
Chromosome 17	54	31
Chromosome 18	48	35
Chromosome 19	32	27
Chromosome X	69	85
Chromosome Y	0	0
Total	1425	964

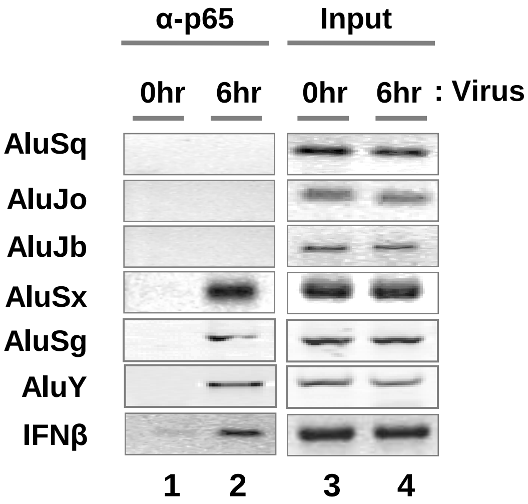


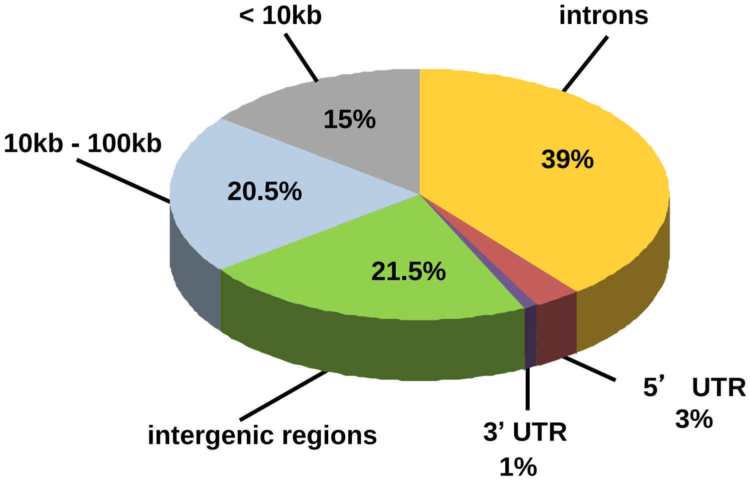
CHIP-cloning



CHIP-seq


220
230
240
A G A G A C G G G G T T T T G C C A T G T T G G C C A G G C ALUSQ.SEQ
A G A G G C A A A G T C T T G C C A T G T T G C C C A G G C ALUJO.SEQ
A G A G A T G A G G T T T T G C C A T G C G G G C C A G G C ALUJB.SEQ
A A A G A T T G G G T T T C A C C A T G T T G G C C A G G C ALUSX.SEQ
A G A G A C A G G G T T T C A C C A T G T T A G C C A G G A ALUY.SEQ
A G A A A C G G G G T T T C A C C A T A T T G G C C A G G C ALUSG.SEQ

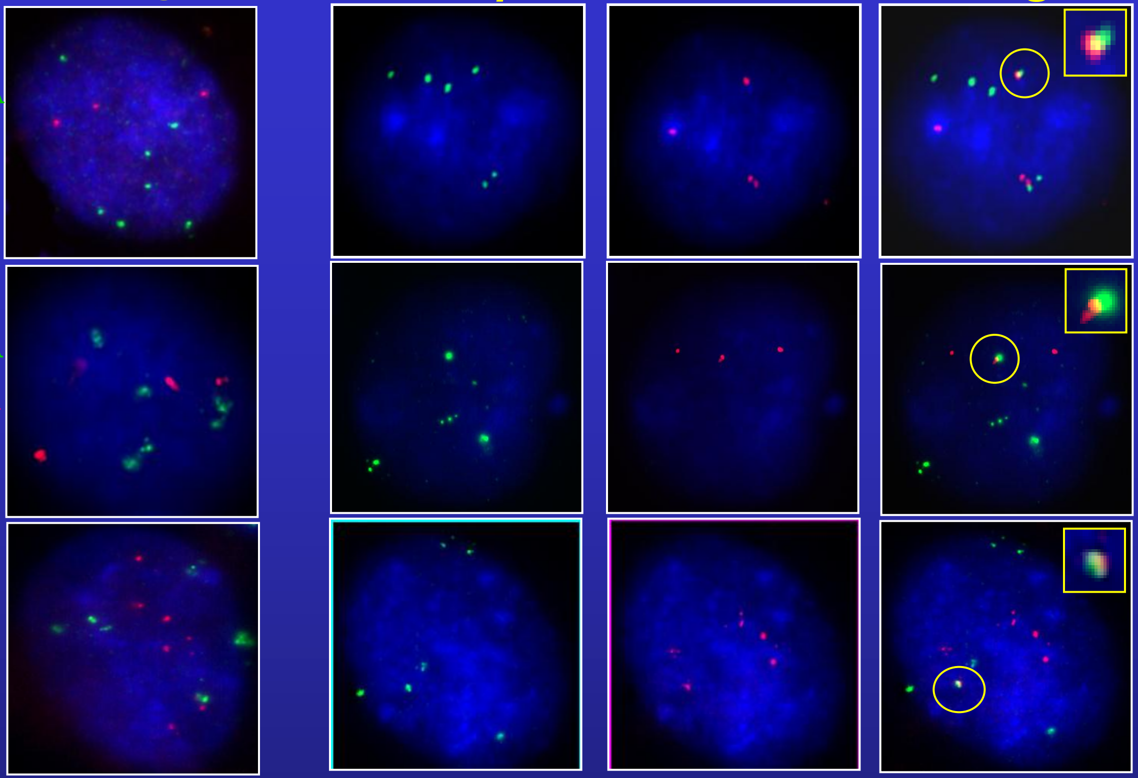

α-p65
Input
0hr
6hr
0hr
6hr
: Virus
AluSq
AluJo
AluJb
AluSx
AluSg
AluY
IFNβ
1
2
3
4


< 10kb
introns
15%
39%
10kb - 100kb
20.5%
21.5%
5' UTR
3%
3' UTR
1%
intergenic regions

uninfected

4h post-infection

merge **IFN-β** **NRC clone** **merge**

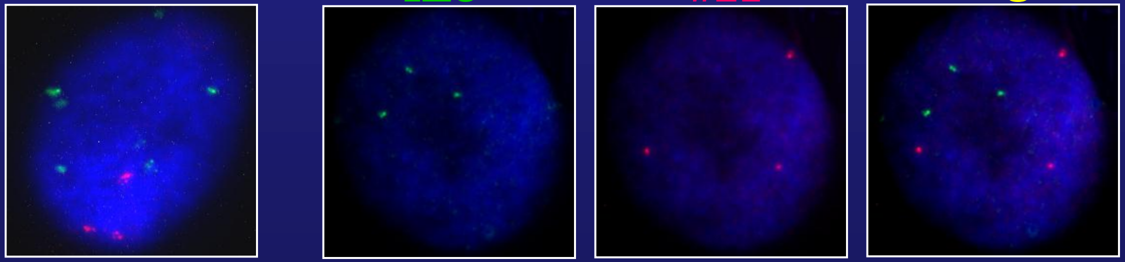


IFN-β DNA
NRC21 DNA

IFN-β DNA
NRC9 DNA

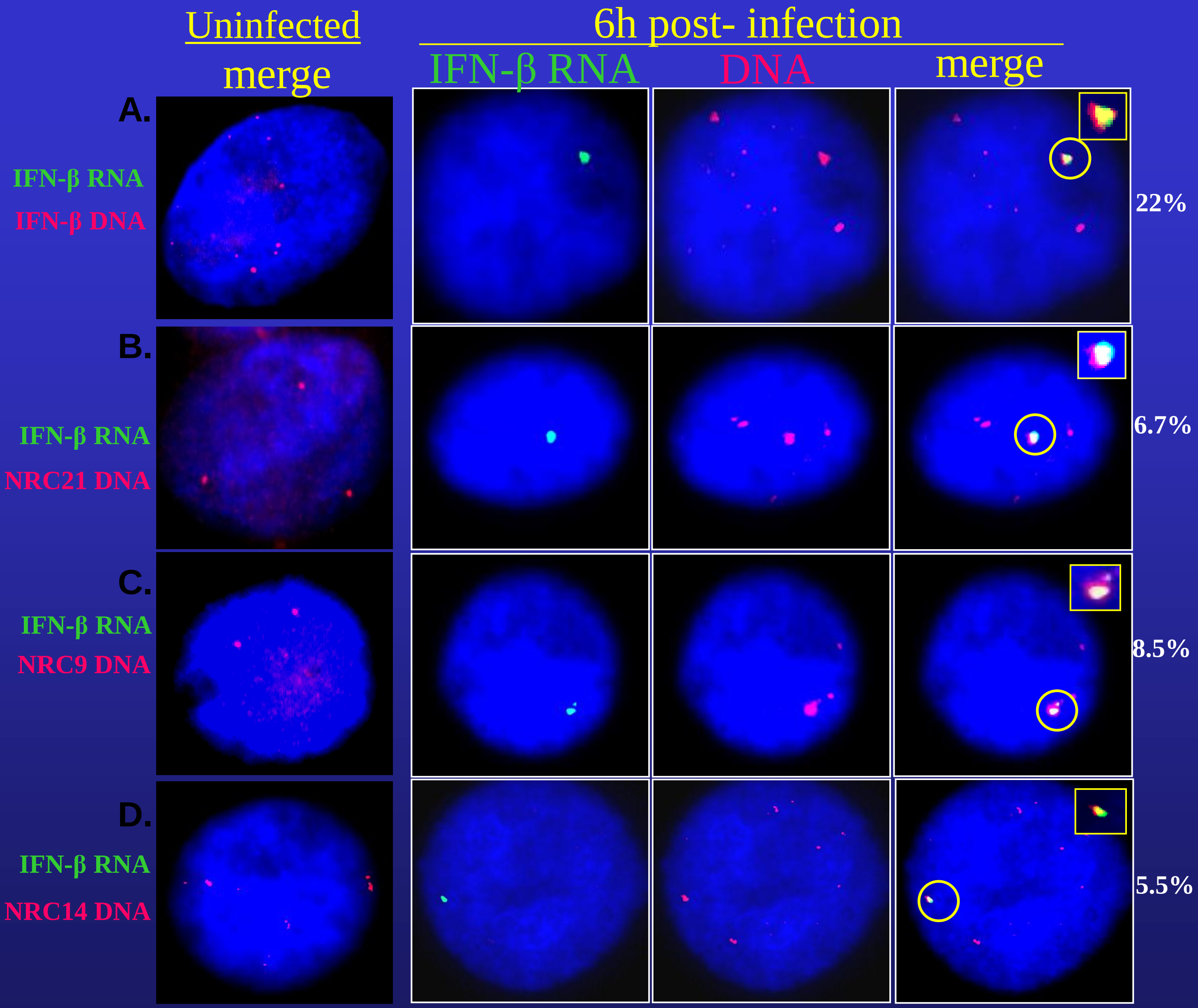
IFN-β DNA
NRC14 DNA

IL8 **#21** **Merged**



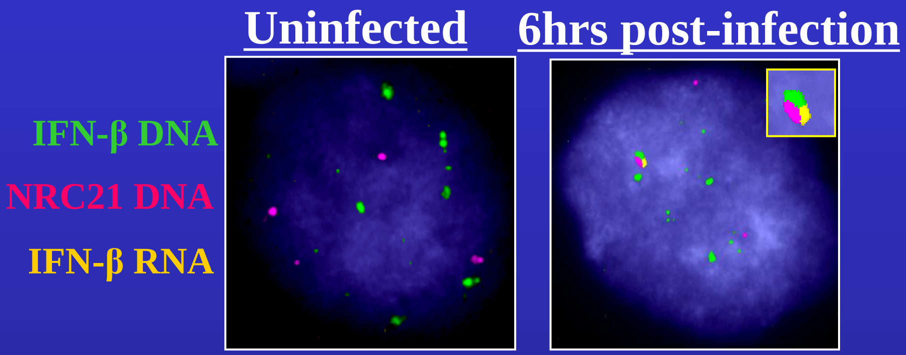
control

RNA/DNA FISH



NRCs colocalize with monoallelically expressed *IFN-β* RNA

Triple RNA/DNA FISH

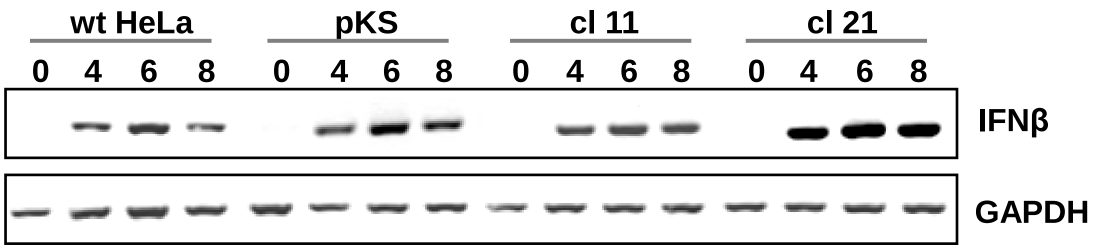


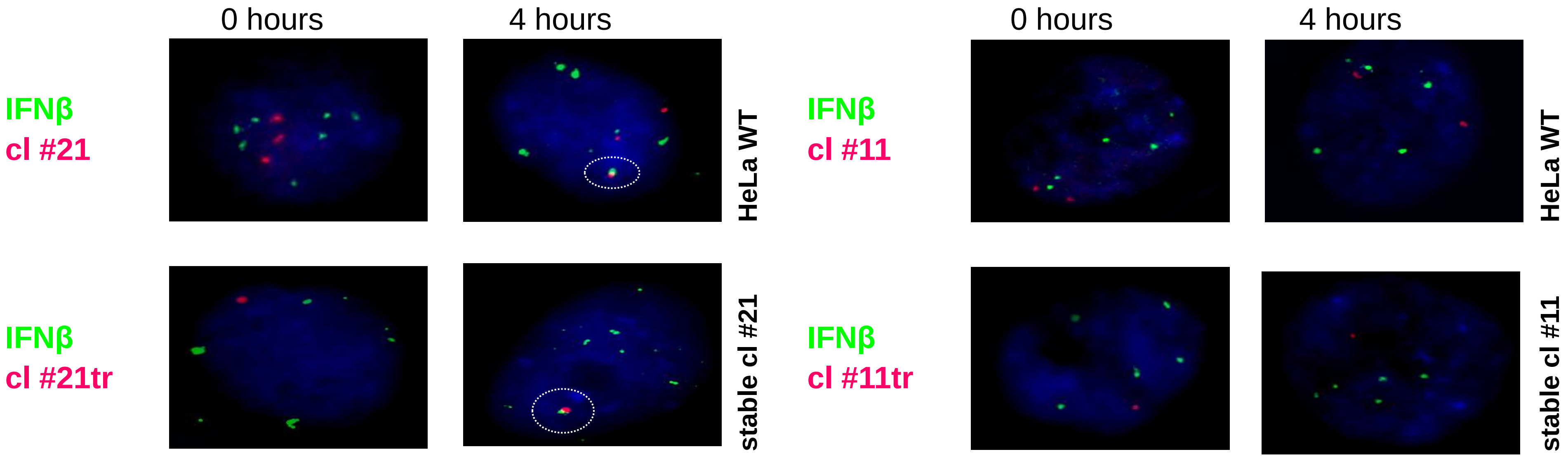
The single *IFN-β* allele interacting with NRCs is the one that expresses *IFN-β* mRNA

The stochastic choice of a single IFN-β allele for expression depends on stochastic inter-chromosomal associations.

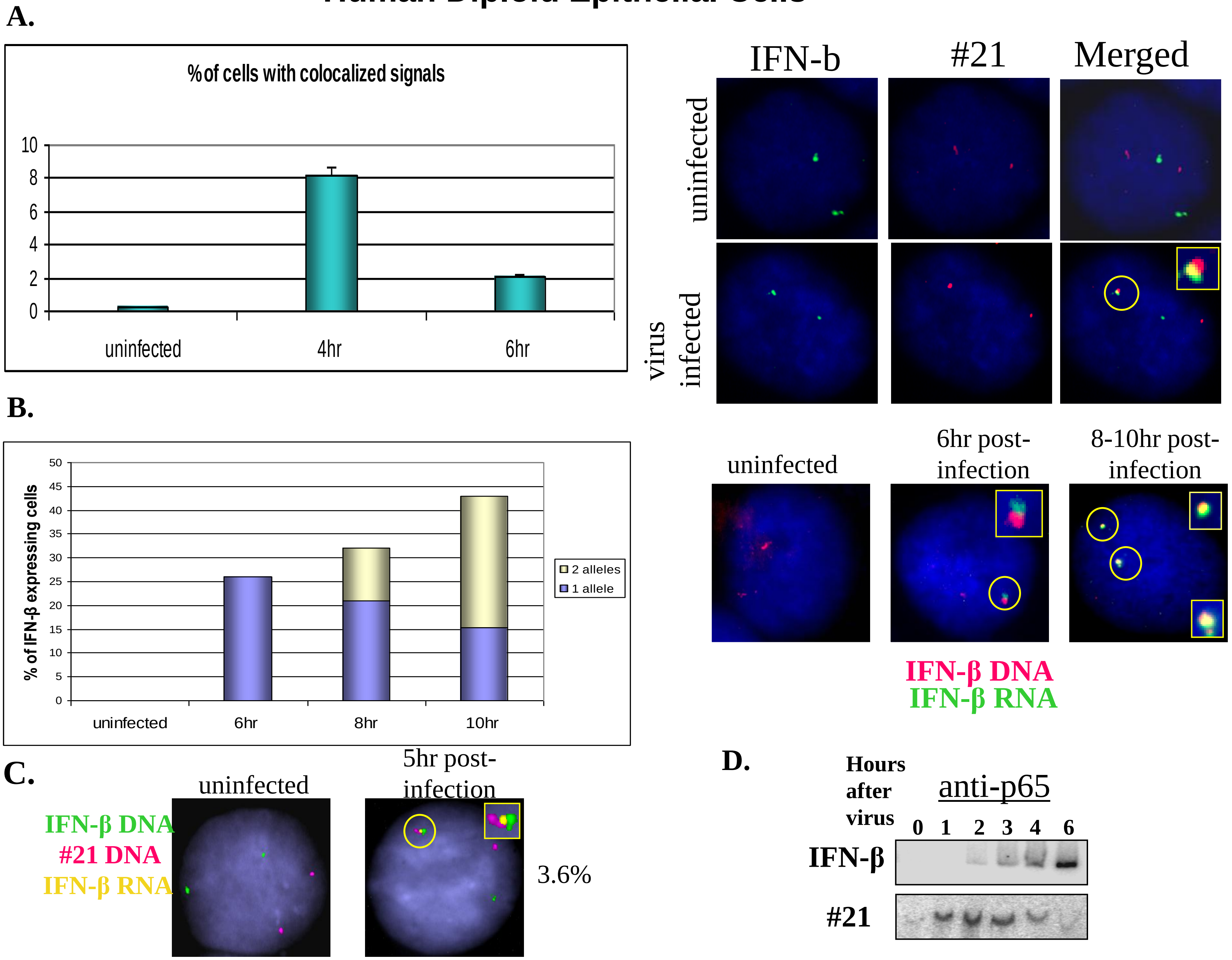
Hypothesis: Increased frequency of inter-chromosomal associations should lead to an increased probability of IFN-β expression

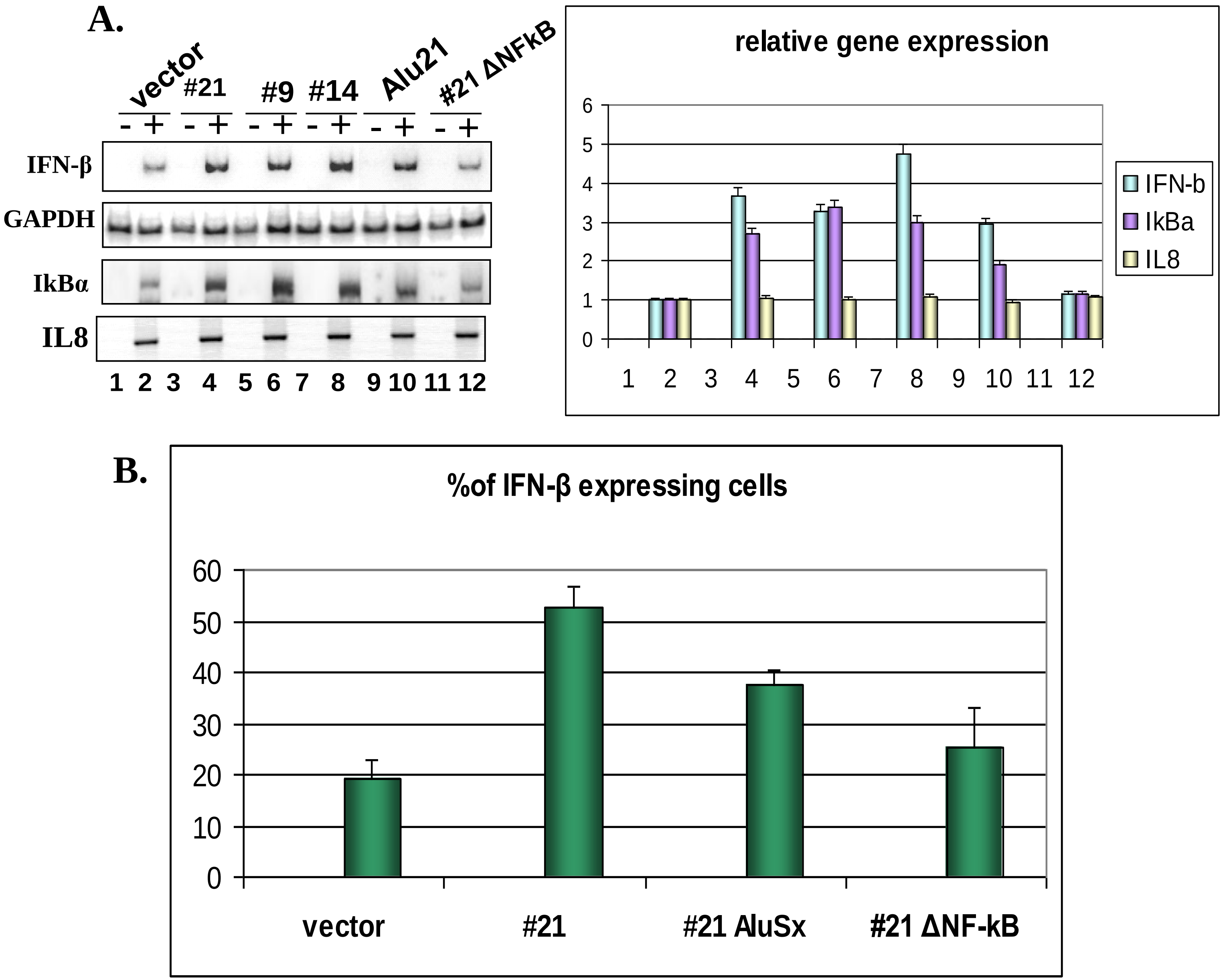

Alu-κB site
NRC # 21
NON NRC # 11
--GTGGATCACCTGAGGTCAGGAGTTCGAGATTAGCTTGGCCAACATGGTGAAACCC
GTGAGGATCA--TGAGGTCAGGAGATTGAGGCCATCCTGGCCAACATGGTGAAACCC
CGTTTCTACTAAAAAA-TACAAAAA--TTAGCTGCGCATGGTGGCAGGT
CATCTCTACTAAAAAAATACAAAAAAATTAGCTGGGTGTGGTGGCA---


wt HeLa
pKS
cl 11
cl 21
0
4
6
8
0
4
6
8
0
4
6
8
0
4
6
8
IFNβ
GAPDH


0 hours
4 hours
0 hours
4 hours
IFNβ
cl #21
IFNβ
cl #21tr
HeLa WT
stable cl #21
IFNβ
cl #11
IFNβ
cl #11tr
HeLa WT
stable cl #11

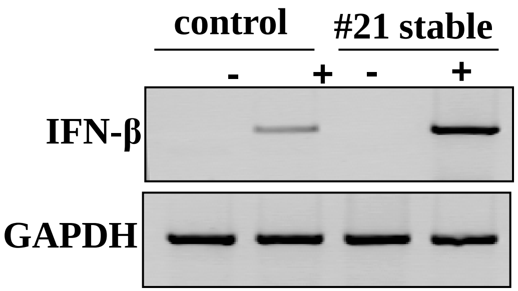
Human Diploid Epithelial Cells

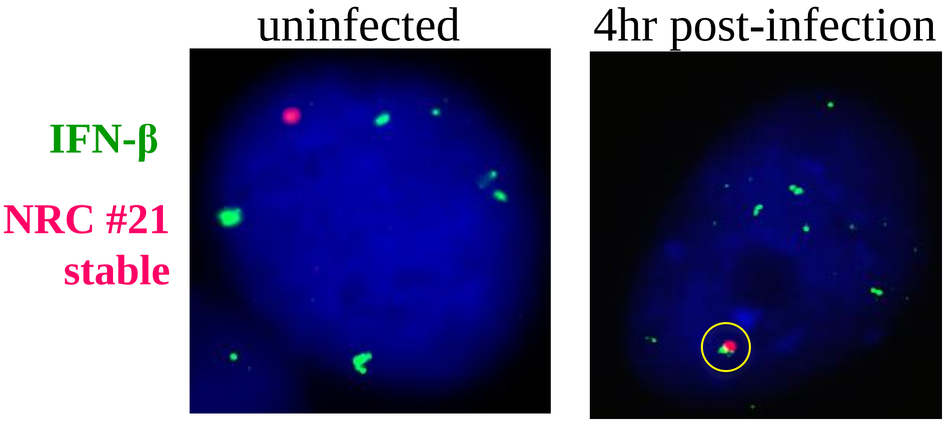




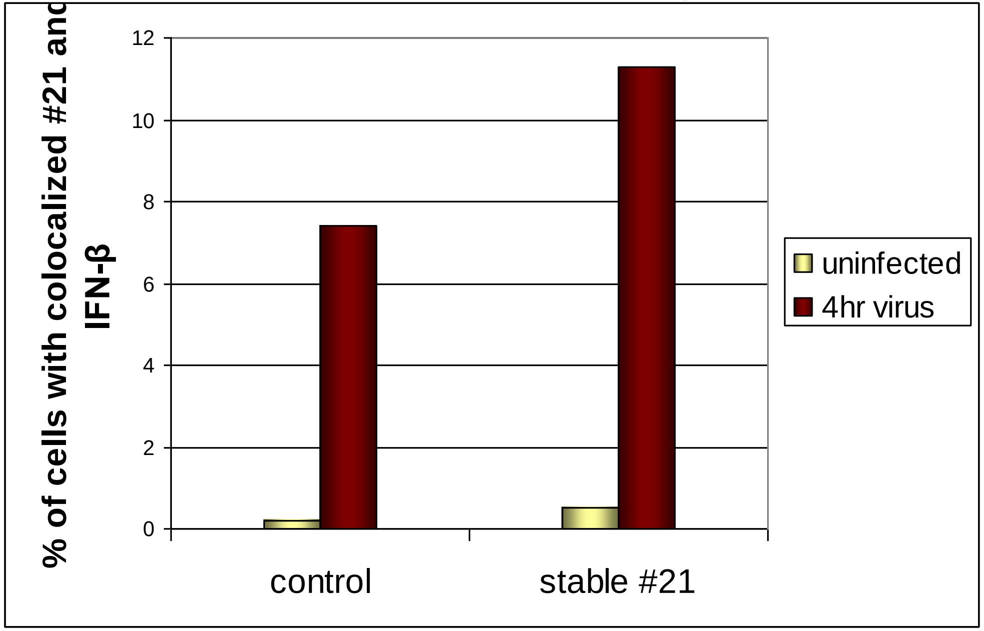
The NRC NF-kB site plays a critical role in determining the probability of IFN-β expression in each cell

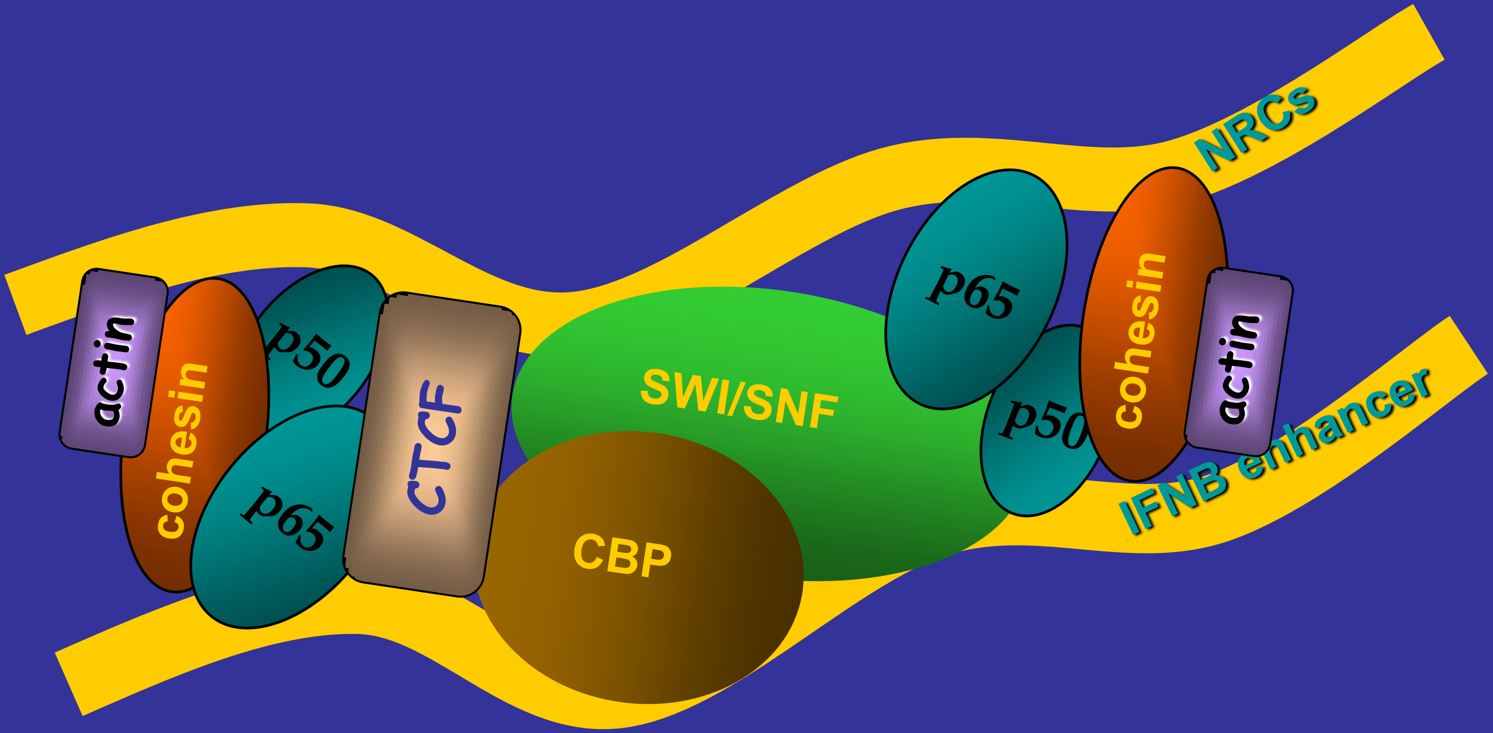
RT-PCR



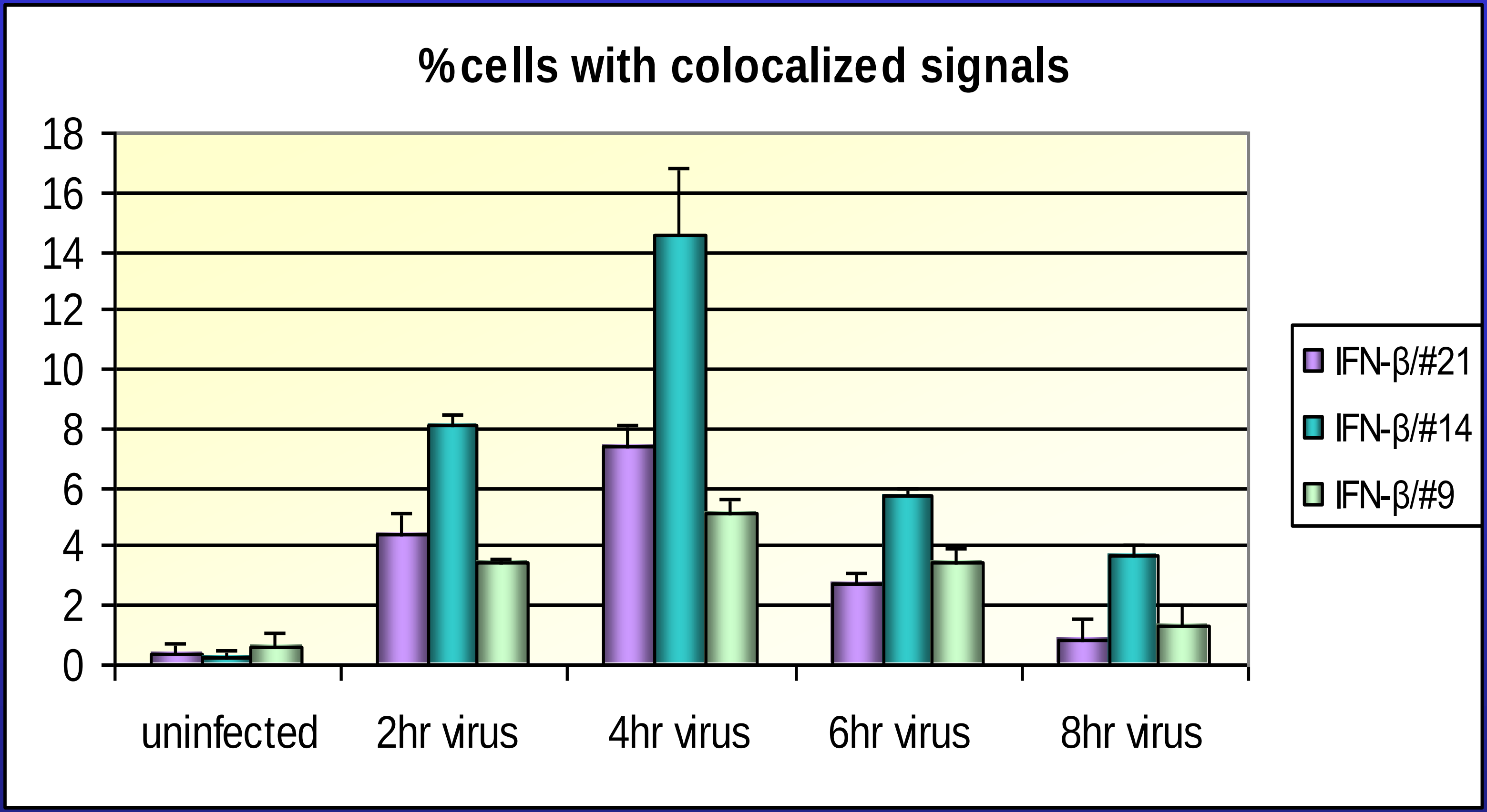


DNA FISH #21-IFN-β



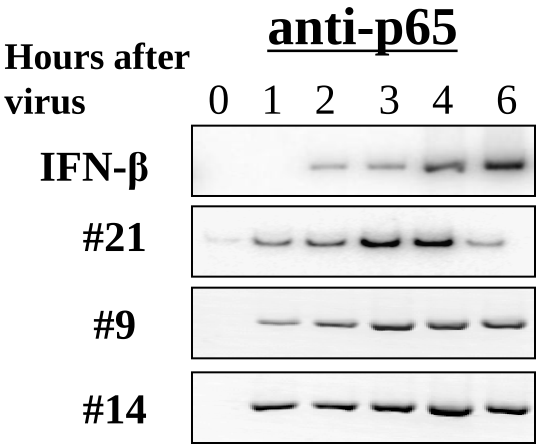

actin
cohesin
p50
p65
CTCF
SWI/SNF
CBP
p65
p50
cohesin
actin
NRCs
IFNB enhancer

HOW NRCs WORK?



The interchromosomal associations between the *IFN-β* locus and the NRCs occur at maximal frequencies before initiation of transcription and during enhanceosome assembly and they are significantly reduced at the time of initiation of transcription (~6 hrs)

ChIP



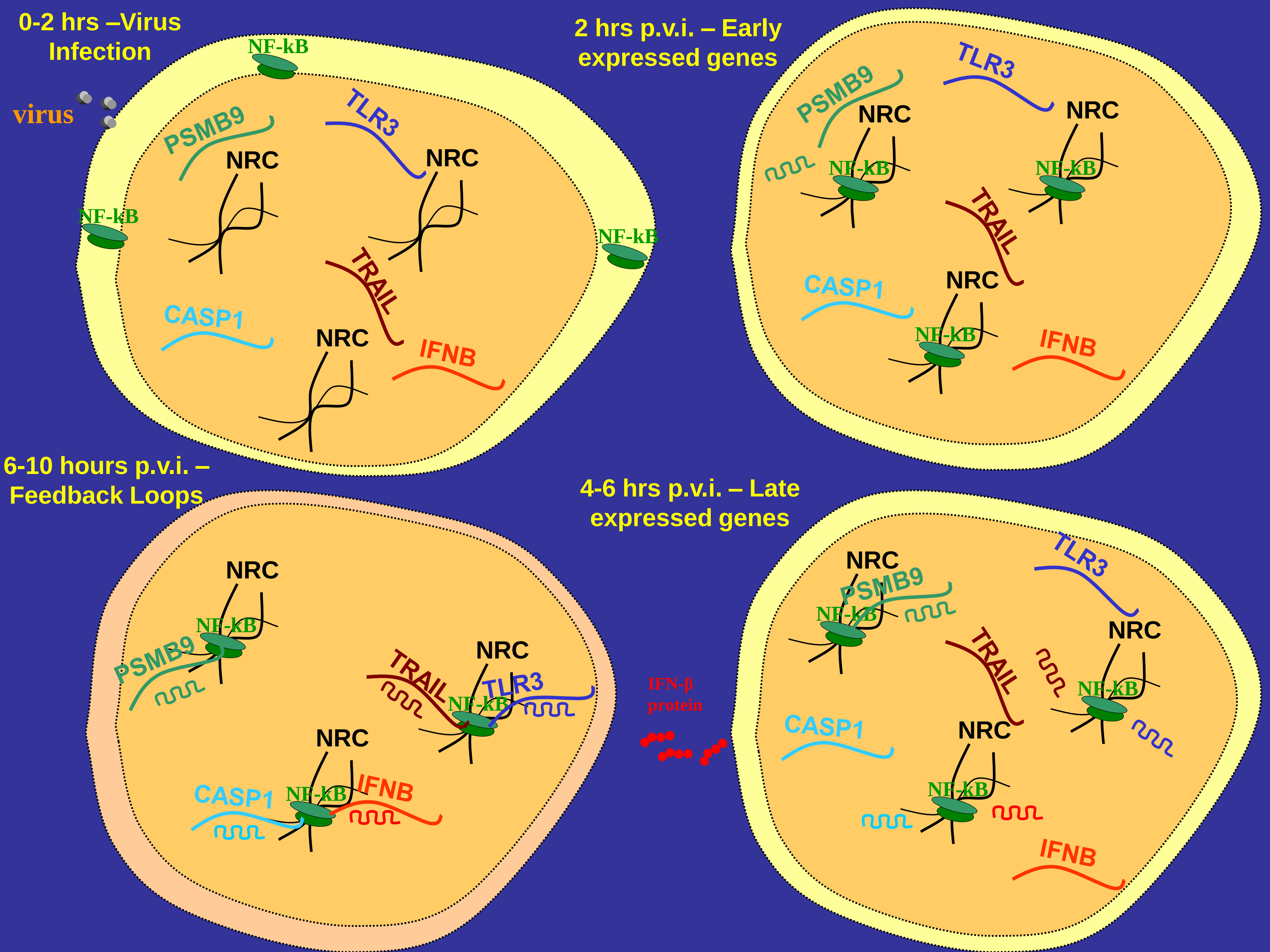
NRCs facilitate IFN-β expression by mediating the delivery of “limiting” NF-κB to the IFN-β enhancer

However,
despite this amazing complexity, the fraction of cells expressing these genes is always the same

However,
despite this amazing complexity, the fraction of cells expressing these genes is always the same

thus,

stochasticity in gene expression is genetically determined



0-2 hrs –Virus Infection

2 hrs p.v.i. – Early expressed genes

4-6 hrs p.v.i. – Late expressed genes

6-10 hours p.v.i. – Feedback Loops

Biomedical Research Foundation, Academy of Athens

Agelos Banos
Theodore Georgomanolis
Stefanos Tsiftsoglou
Maria Papathanasiou
Trantafyllos Paparountas

Effie Apostolou
Menie Merika
Maria Pliatsika
Depy Papadopoulou
Ethan Ford

Alexandros Polyzos
Chrysa Nikopoulou
Matthieu Lavigne
George Theodorou
George Sianidis