

# Transcription

# Central dogma

Describes the direction of information transfer in template processes

Template process - process of synthesizing one biomolecule based on a template of another one

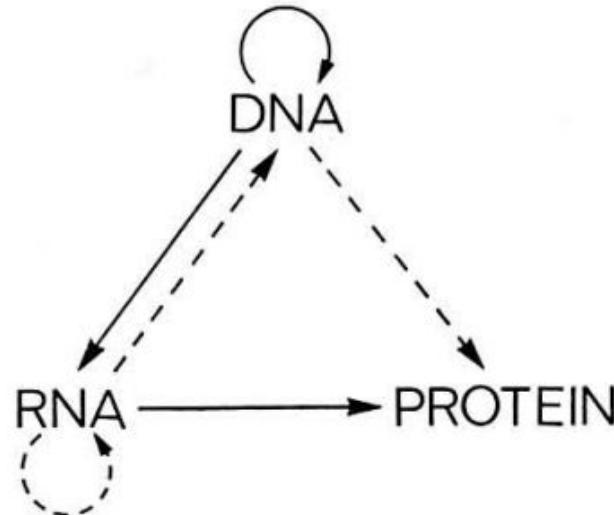


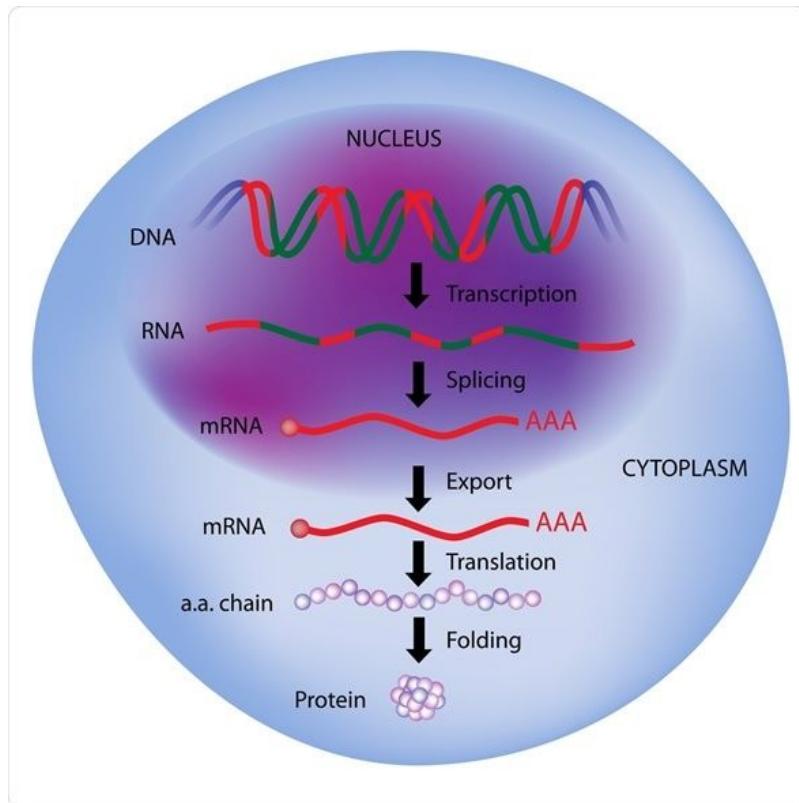
Fig. 3. A tentative classification for the present day. Solid arrows show general transfers; dotted arrows show special transfers. Again, the absent arrows are the undetected transfers specified by the central dogma.

# Gene expression

Process by which information from a gene is used in the synthesis of a functional gene product (usually a protein)

Several stages

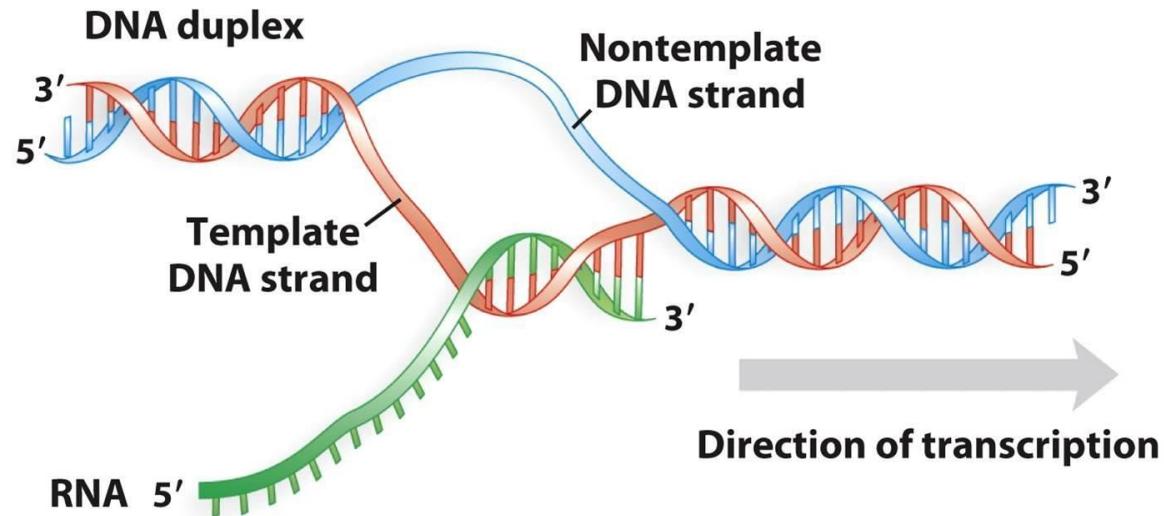
Each stage can be regulated



# Transcription

Template synthesis of RNA from DNA

Performed by DNA-dependent RNA-polymerase

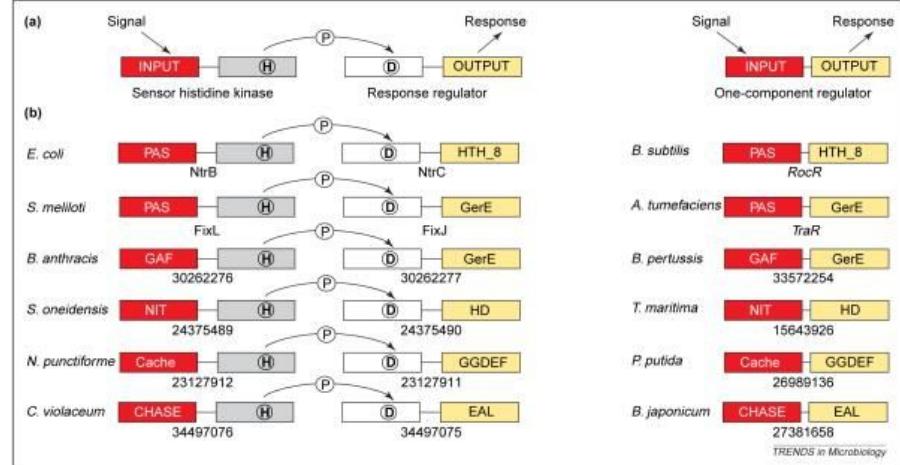


**Figure 15-1**  
*Molecular Biology: Principles and Practice*  
© 2012 W. H. Freeman and Company

# Why studying expression on RNA level makes sense?

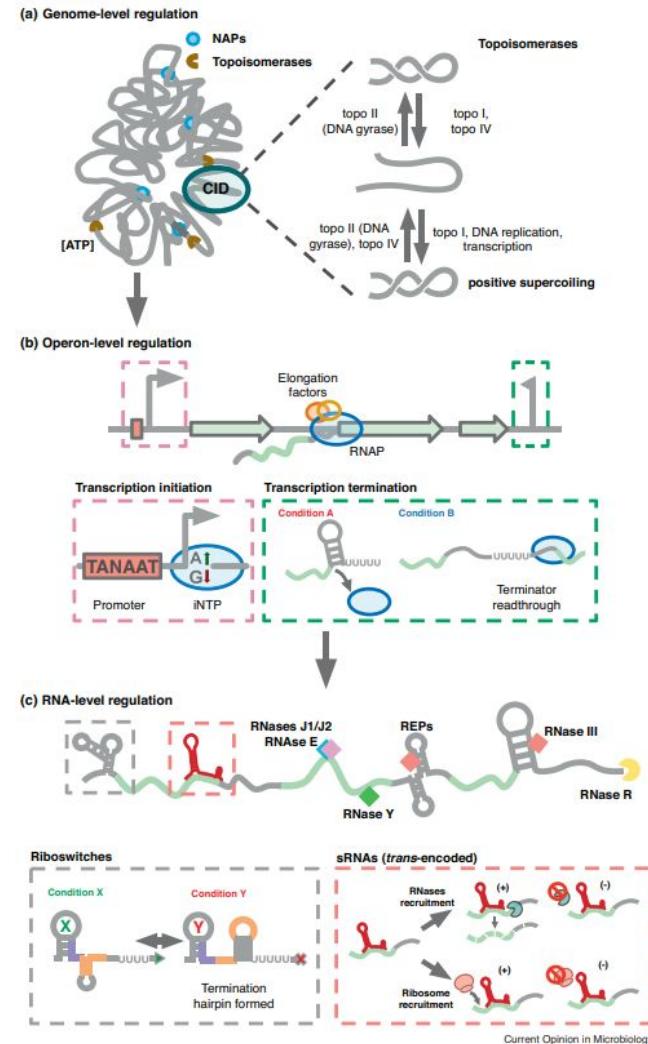
# Signal transduction

Life relies on signal uptake and interpretation



# Transcription regulation

Multiple levels of regulation starting with global regulation of genome structure and down to individual mRNAs



# DNA supercoiling

Transcription bubble changes DNA topology

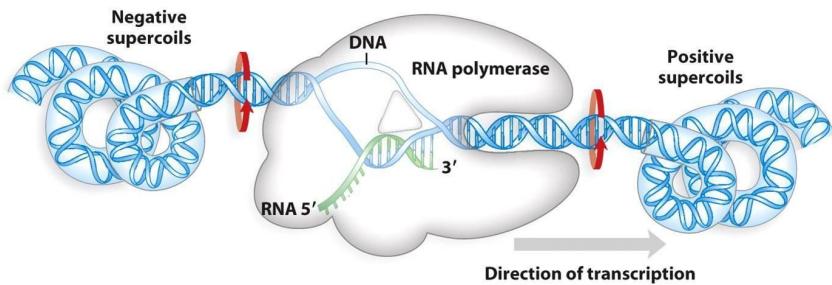
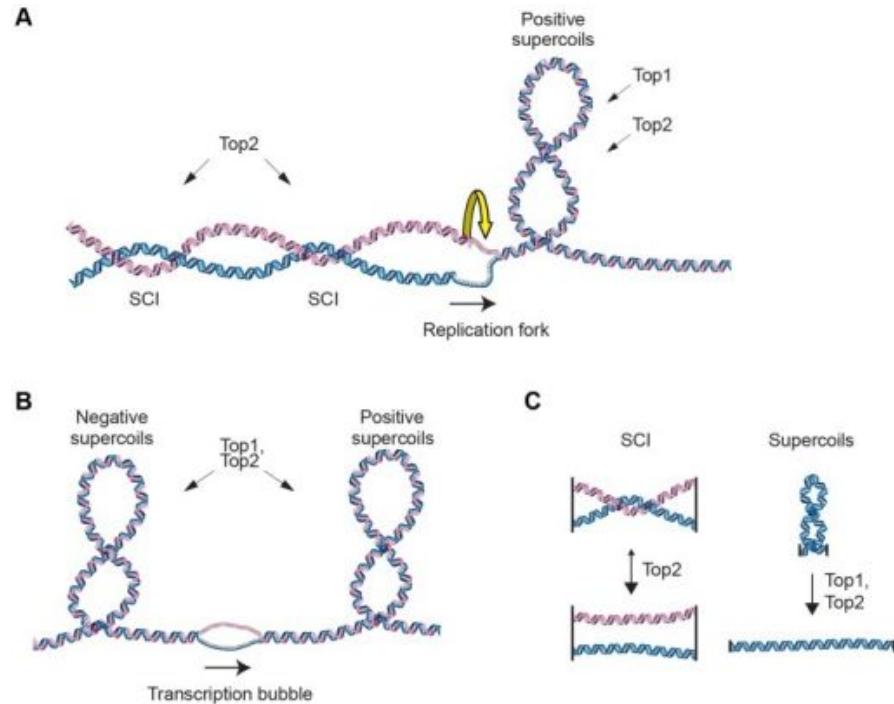


Figure 15-6  
Molecular Biology: Principles and Practice  
© 2012 W.H. Freeman and Company

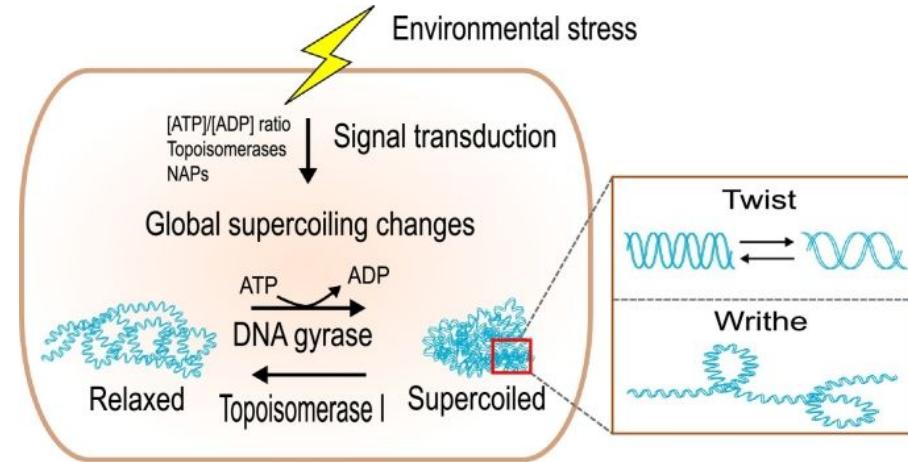


# Topoisomerases

Supercoiling is controlled by topoisomerases

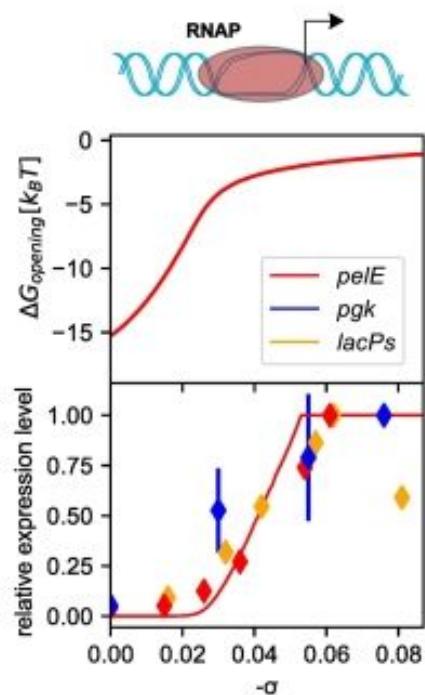
Gyrase induces negative supercoils

Topoisomerase I acts in the opposite direction

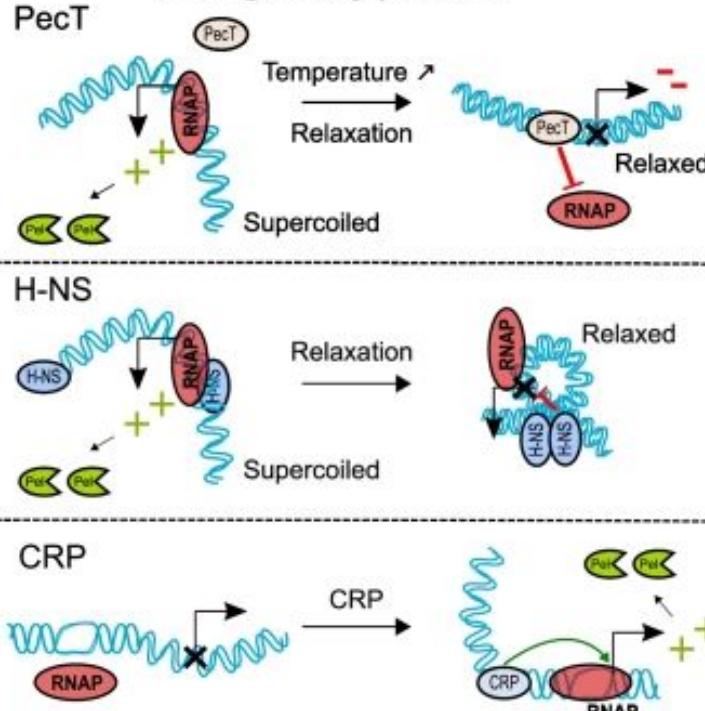


# DNA supercoiling as a regulator

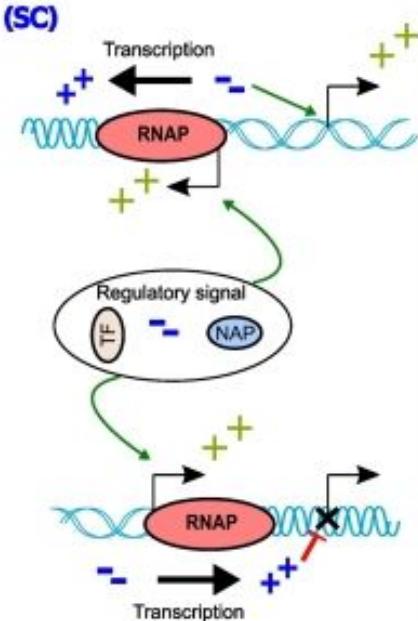
## A. Promoter opening



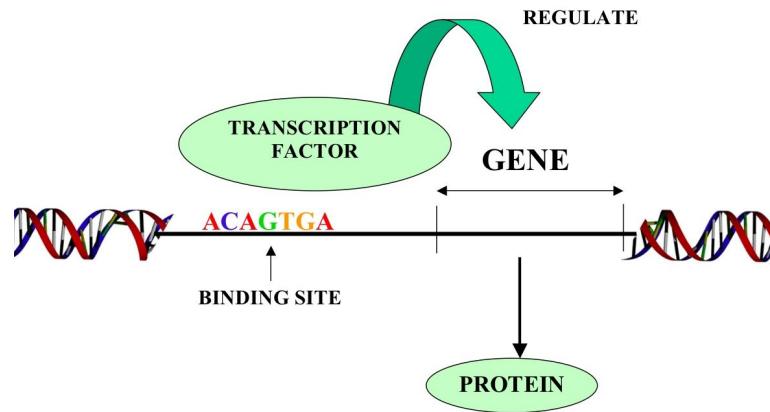
## B. Regulatory proteins



## C. Transcription-supercoiling coupling

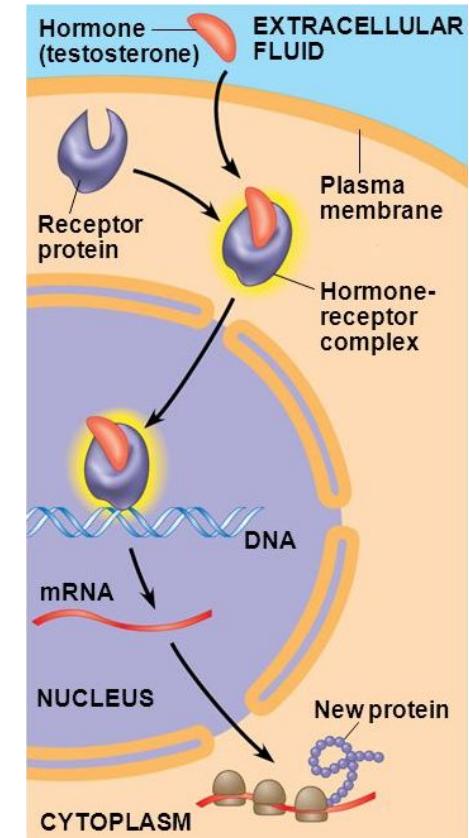


# Transcription regulation



Usually transcription is regulated on the initiation stage by transcription factors

Transcription factors - DNA binding proteins that regulate transcription



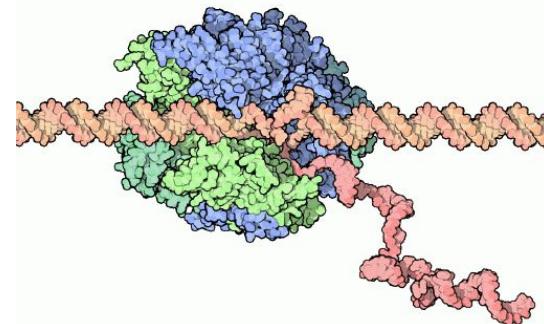
# RNA polymerase

Multisubunit enzyme

Subunits provide promoter binding,  
RNA synthesis, interaction with  
regulatory elements and DNA binding

Prokaryotic RNA Polymerase:  
Holoenzyme Enzyme

Subunit	Size	#/Molecule	Function
$\alpha$	36.5 kD	2	chain initiation and interaction with regulatory proteins
$\beta$	151 kD	1	chain initiation and elongation
$\beta'$	155 kD	1	DNA binding
$\sigma$	70 kD	1	promoter recognition

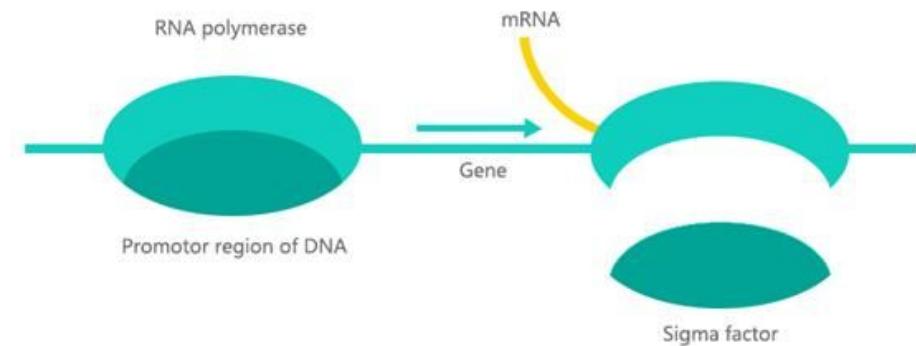


# Sigma factor

Transcription initiation factor

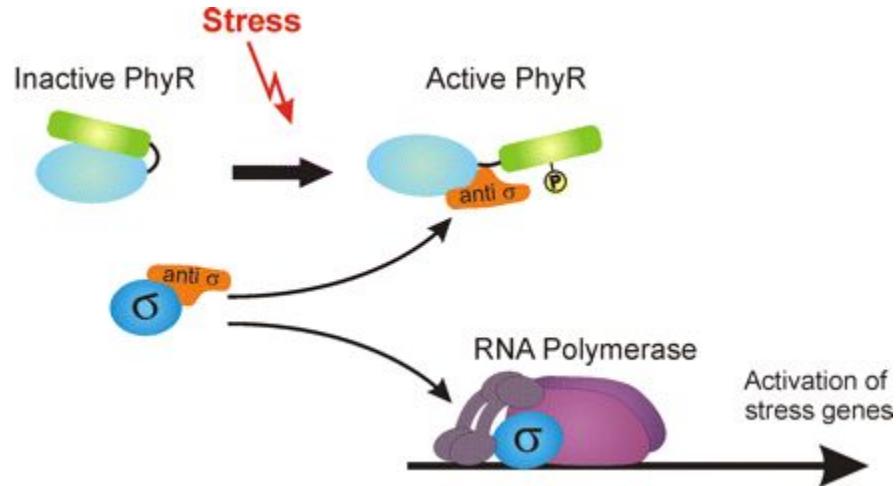
Bacteria may have multiple sigma factors with different controlled genes

*E. coli* has 7 different sigma factors.  $\sigma^{70}$  is the primary sigma factor. Serves for transcription of housekeeping genes



# Anti-sigma factors

Bind sigma factors and allow to start other transcription programs in response to change in environment



# Promoter

A place of transcription initiation and RNA polymerase binding

Prokaryotic promoter site:

DNA template	TTGACA	TATAAT	+1
	-35 Region	Pribnow box	Start of transcription

Eukaryotic promoter site:

DNA template	GGNCAATCT	TATAAA	+1
	CAAT box	TATA (Hogness) box	Start of transcription

# Activation strategies

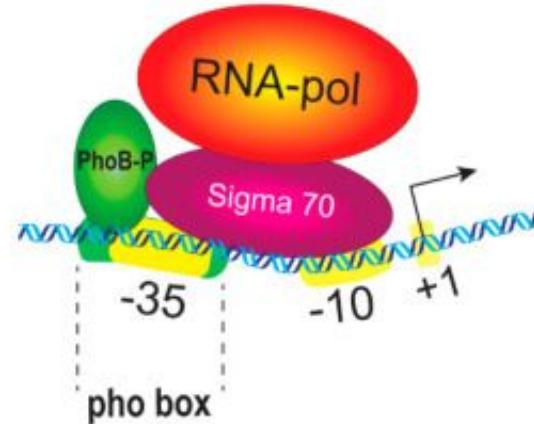
Several different ways of how to activate transcription

The simplest - to bind both promoter and RNAP to promote their interaction

PhoB reduces abortive transcription

A

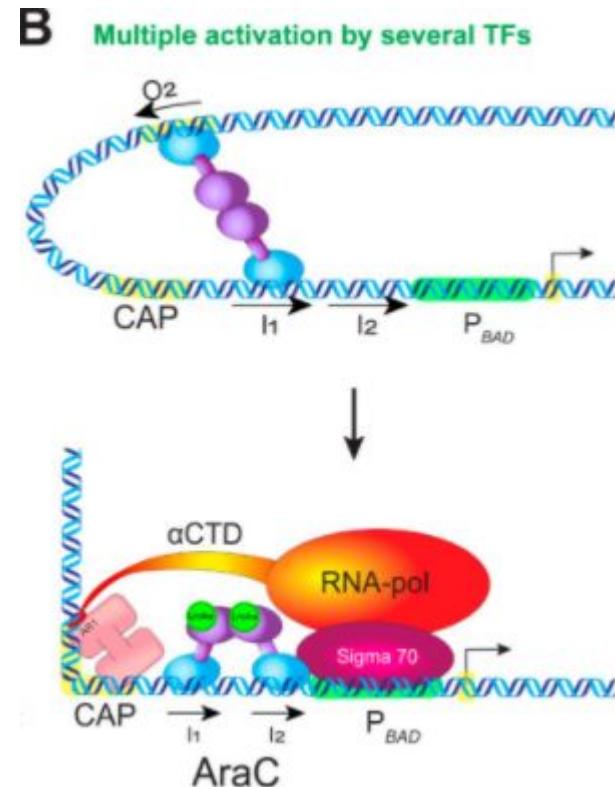
Direct activation by activator



# Multiple regulators

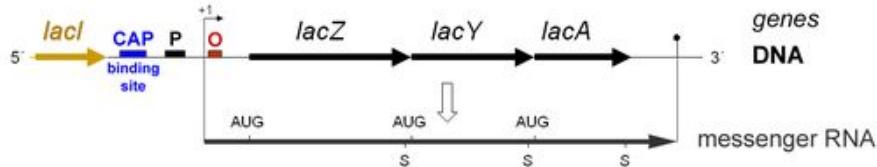
Two regulators

Without arabinose repression loop  
doesn't allow active transcription

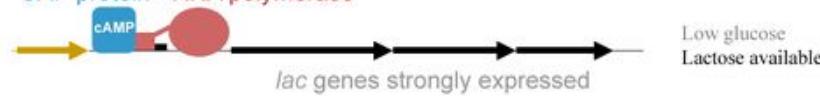


# Lac operon

## The lac Operon and its Control Elements



CAP protein RNA polymerase



Repressor protein



cAMP



CAP

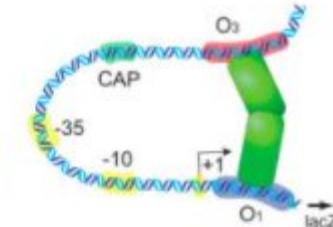
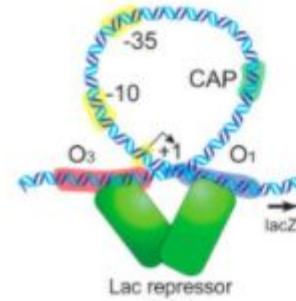
P

O

very low (basal) level of gene expression

C

## Repression loop

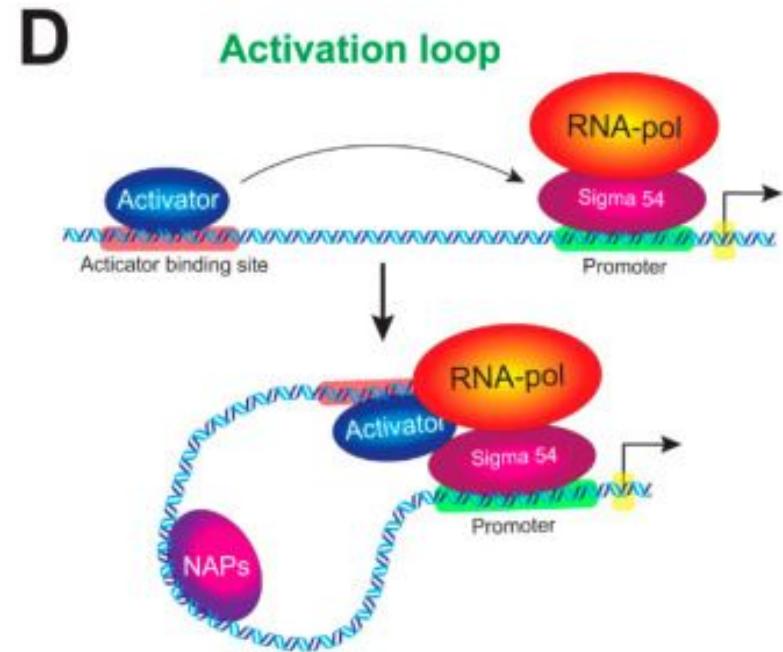


# Activation loop

Without nitrogen NtrC is phosphorylated

Loop interaction allows RNAP remodelling into active form

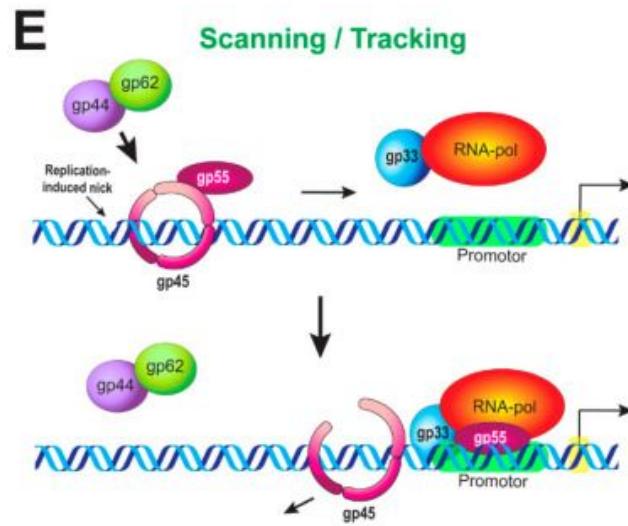
Nucleoid-associated proteins bend DNA to facilitate the interaction



# Sliding clamp activation

The ring is loaded onto DNA in the upstream region

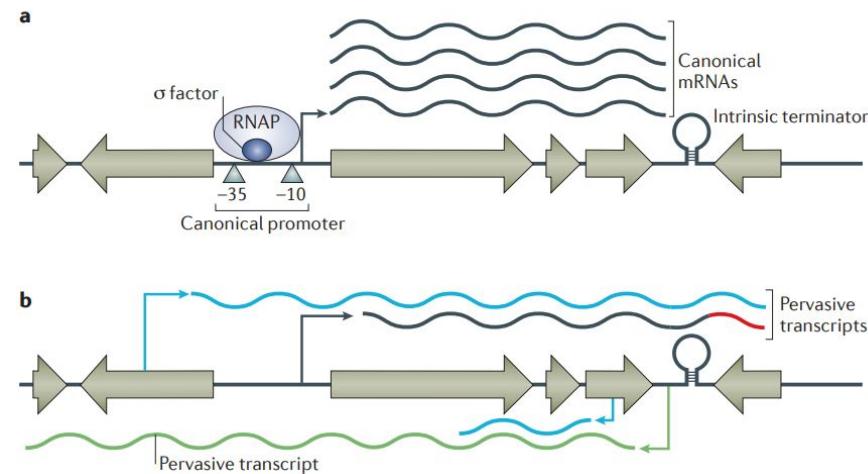
Scans DNA until promoter is found



# Pervasive transcription

Aside from canonical RNAs there is a multitude of non-canonical RNAs that are a) not constrained by gene boundaries and b) often non-coding and antisense

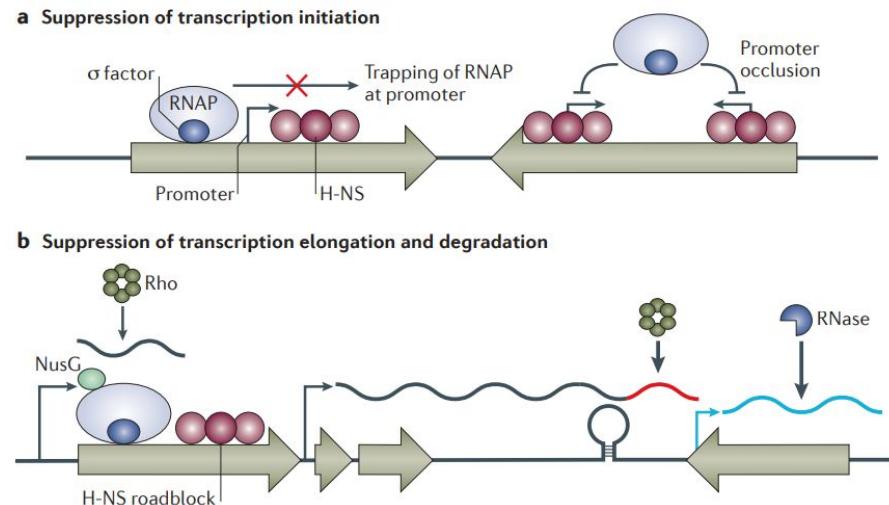
Sources of pervasive transcription - spurious promoters in coding sequences and transcriptional read-through



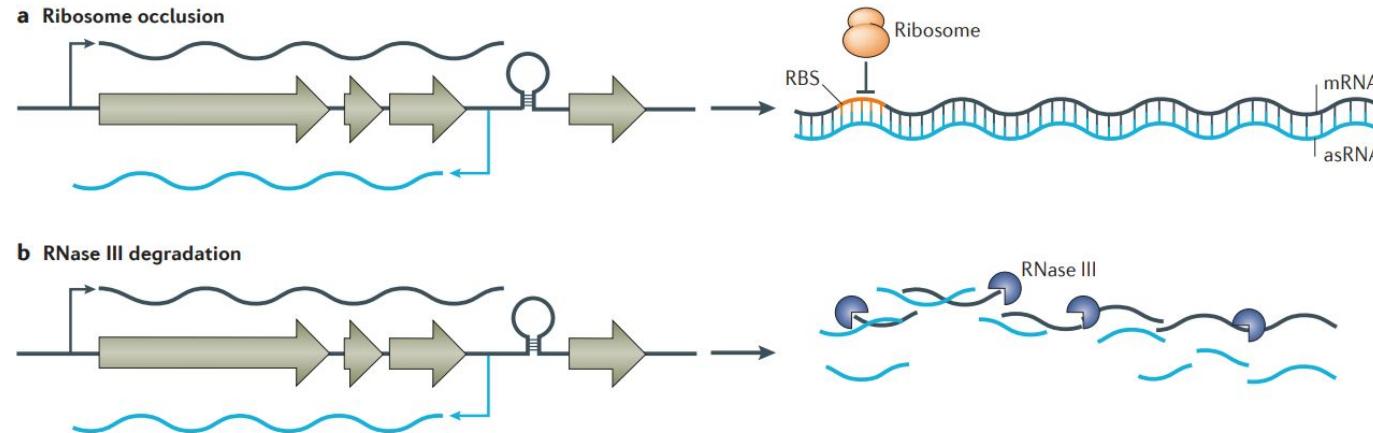
# Suppression of pervasive transcription

Histone-like nucleoid-structuring protein (H-NS) binds to AT-rich regions of the chromosome and suppresses initiation

H-NS can also suppress transcription elongation and make it more vulnerable to Rho-dependent termination



# Is pervasive transcription useful?



Can be a gene regulation mechanism, facilitates RNA degradation and translation suppression

Serves as a source of nanoRNAs that can initiate transcription

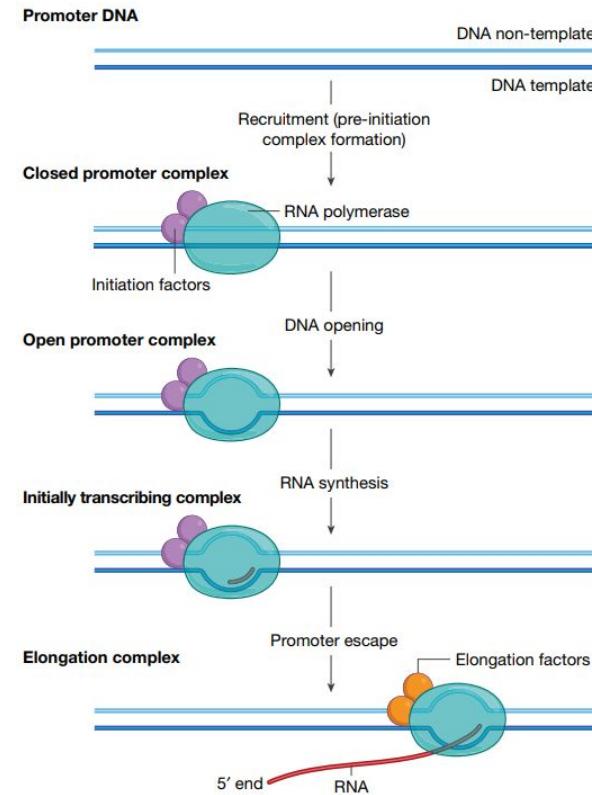
# Eukaryotic transcription

RNAP forms PIC on the promoter

Promoter complex opens

DNA-dependent RNA synthesis then generates an initially transcribing complex

RNAP escapes promoter



# Eukaryotic RNA-polymerases

Three nuclear RNAP:

I - ribosomal RNA

II - messenger RNA

III - small RNAs, transfer RNAs  
and 5S rRNA

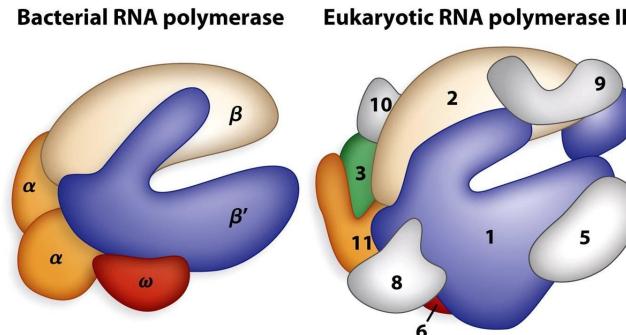


Figure 15-21  
Molecular Biology: Principles and Practice  
© 2012 W. H. Freeman and Company

RNA polymerase	I	II	III
Subunits	<u>14</u>	<u>12</u>	<u>16</u>
unique $\alpha\beta'\omega$ -like	5	5*	5
common	4	4	4
unique	5	3	7
Inhibition [ $\alpha$ -amanitin]	(resistant)	low	high
Products	pre-rRNA (28S, 5.8S, 18S)	mRNAs, 5 snRNAs	tRNAs, U6 snRNA, 5S rRNA, 7S RNA

# Eukaryotic promoters

Different RNA polymerases use different promoters

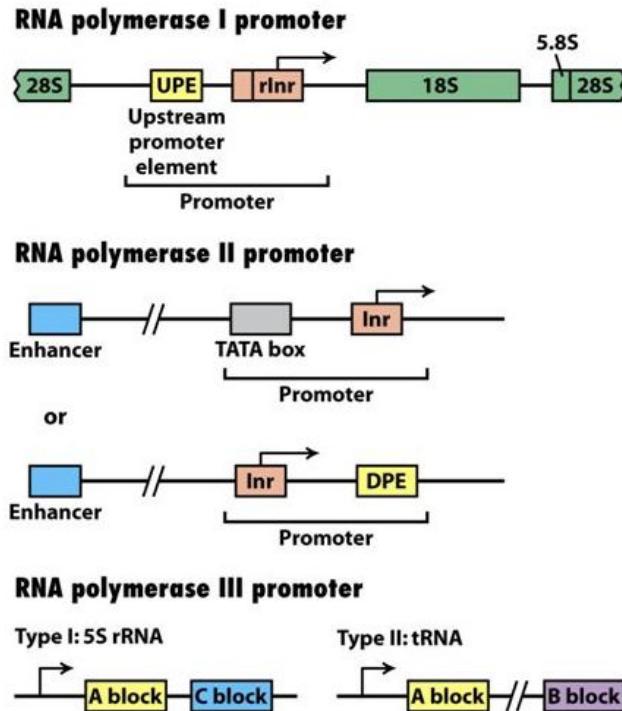
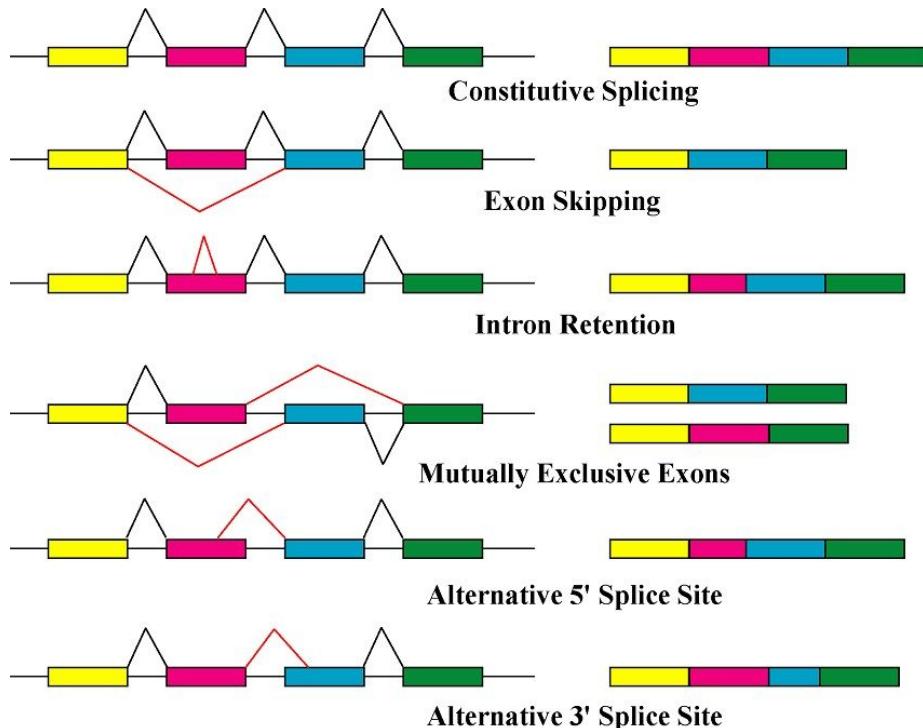


Figure 29-17  
Biochemistry, Sixth Edition  
© 2007 W.H. Freeman and Company

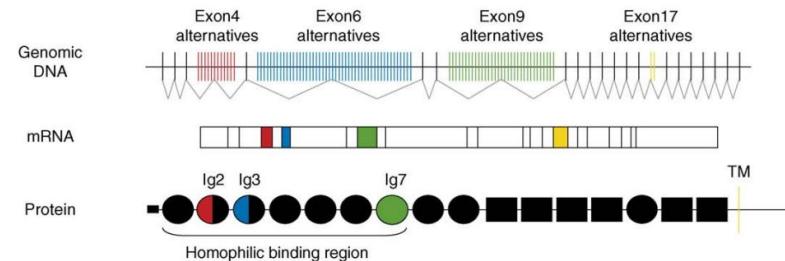
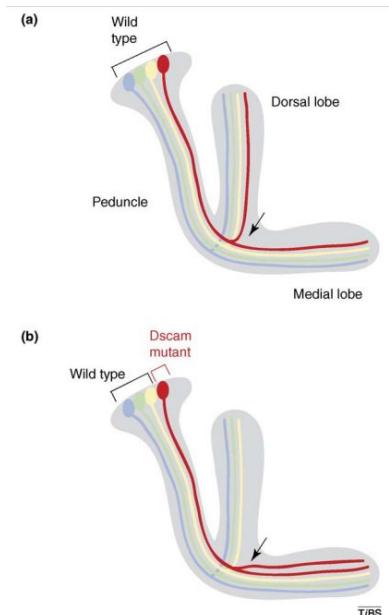
# Alternative splicing

Introns-exon structure allows multiple combinations of exons in the resultant mRNAs that correspond to different protein structures

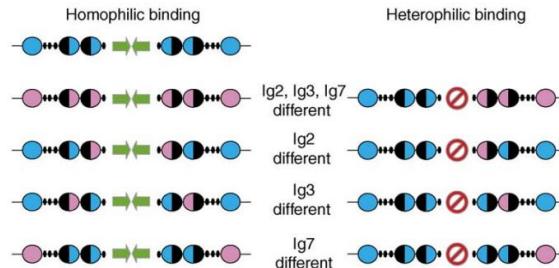


# Dscam alternative splicing in neural wiring

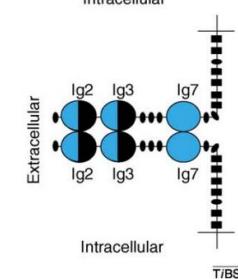
38016 different isoforms with 19008 different external domains



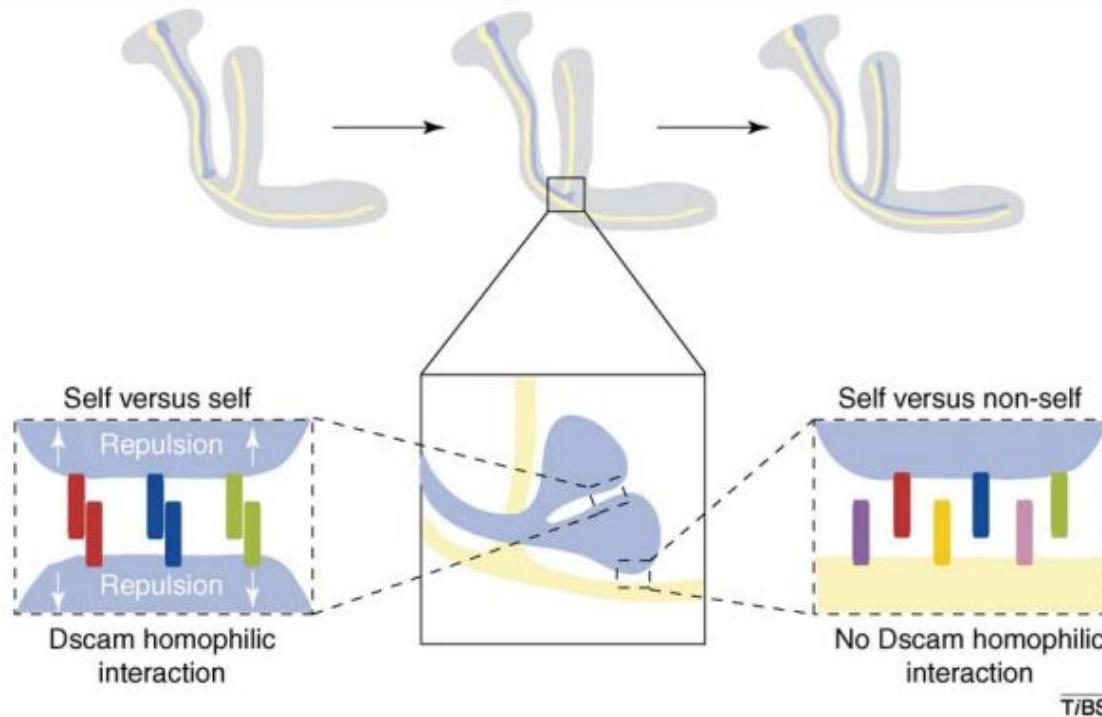
(b) Dscam binding is isoform specific



(c) Model of Dscam binding



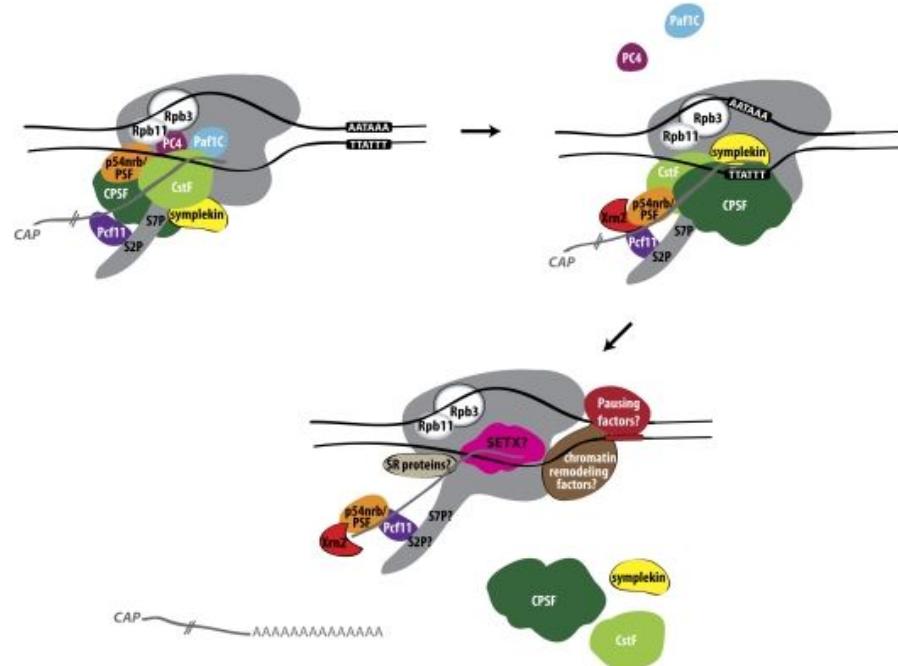
# Dscam alternative splicing in neural wiring



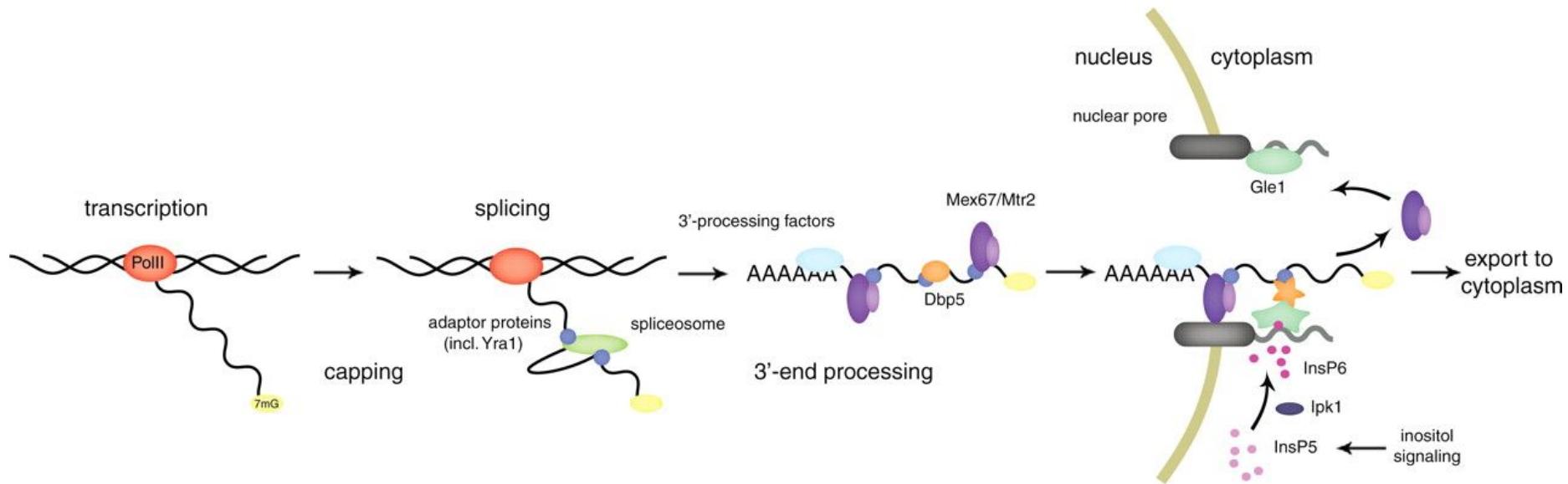
# Transcription termination with RNAPII

Termination happens after polyadenylation site is recognized by CPSF (cleavage and polyadenylation specificity factor) and CstF (cleavage stimulation factor)

Nascent mRNA is cleaved and polyA tail is added



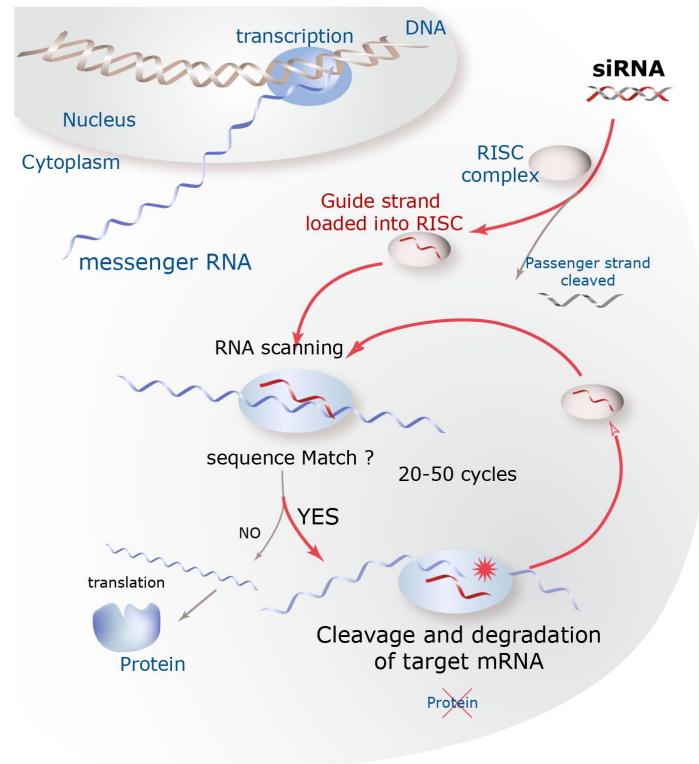
# mRNA nuclear export



# RNA interference

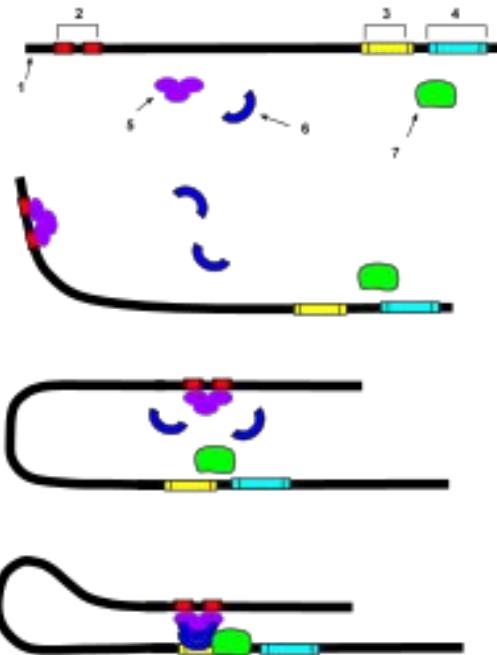
Mechanism of post-transcriptional regulation

Small RNAs are able to suppress translation or degrade target mRNAs

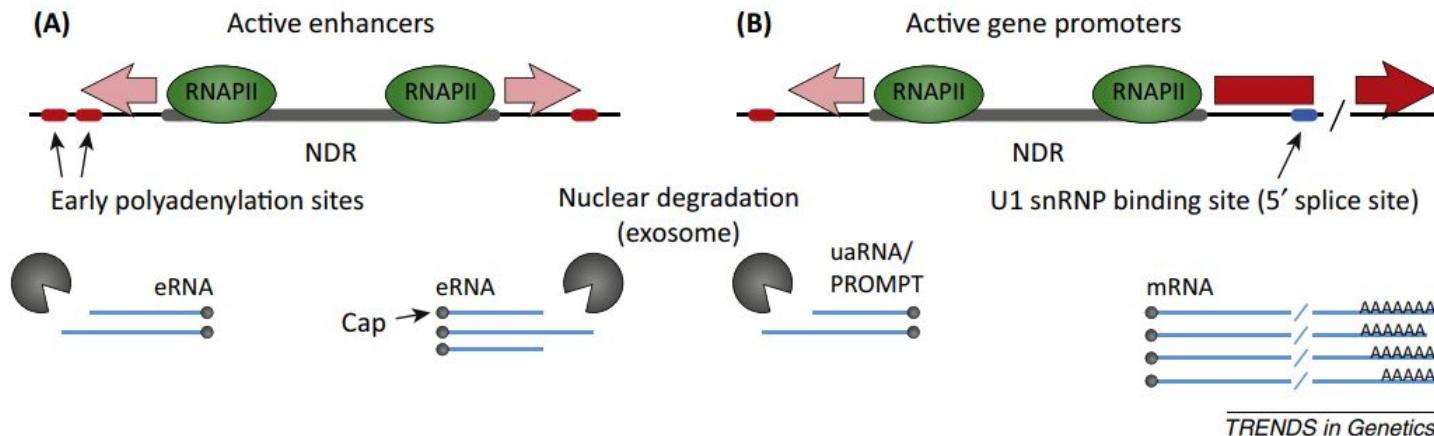


# Enhancer

DNA region that binds transcription factors (TFs) to facilitate transcription



# Enhancers and promoters



Promoters and enhancers share functional characteristics

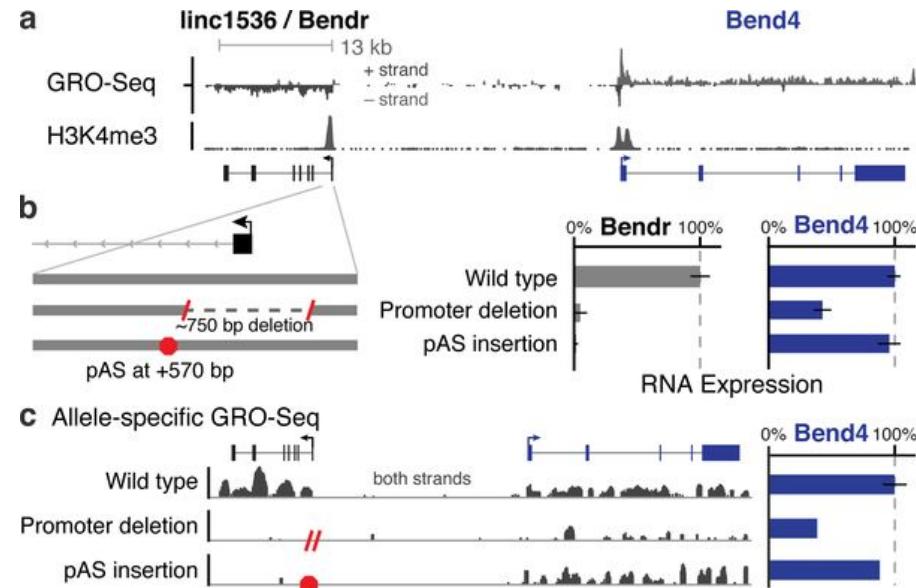
Enhancers probably evolved from promoters

# Promoters can be enhancers

lncRNA promoter regulates neighbouring gene

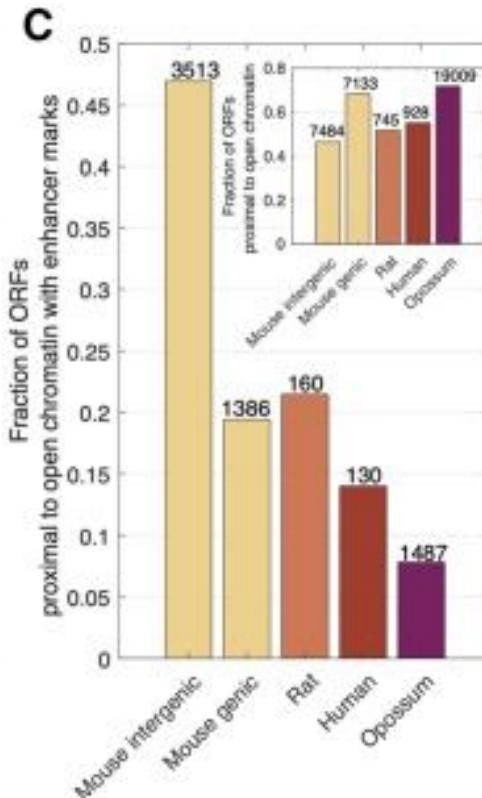
Early polyadenylation signal doesn't change it

Transcripts from this promoter are not even needed for this regulation to happen



# Enhancers drive gene evolution

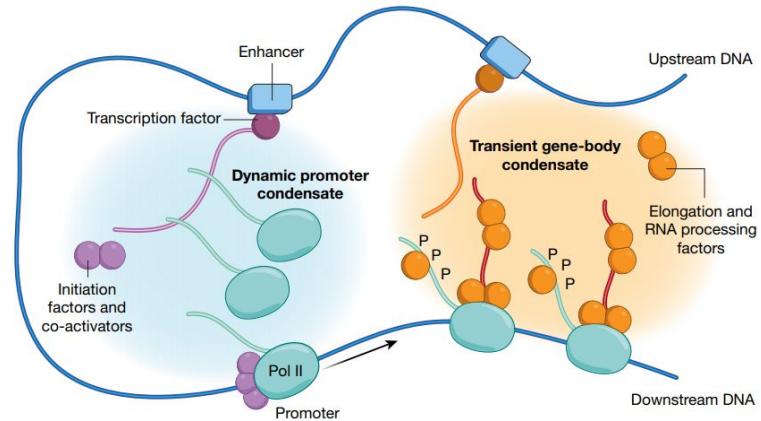
Novel intergenic ORFs are more likely to be occur near enhancers



# Condensate model

TFs form condensates to recruit RNAPII to promoters

RNAPII shuttles between promoter and gene-body condensates



# Transcription factors and genome architecture

Linear view of genome is outdated - 3D  
structure matters

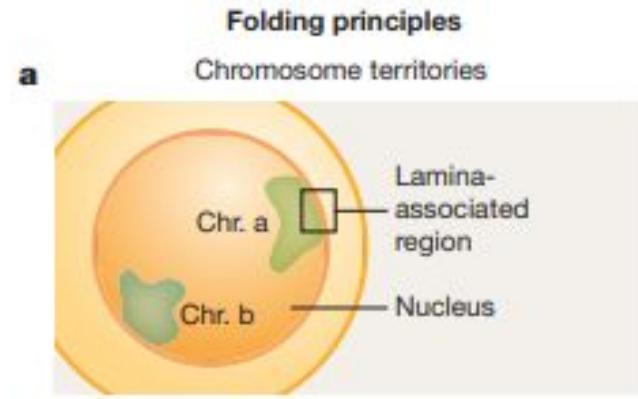
Transcription factors help shape 3D  
structure for altering functionality of  
genome

# Chromosome territories

Chromosomes are organized into specific compartments in the nucleus

Interactions with nuclear lamina important for their maintenance

Territories rarely intersect



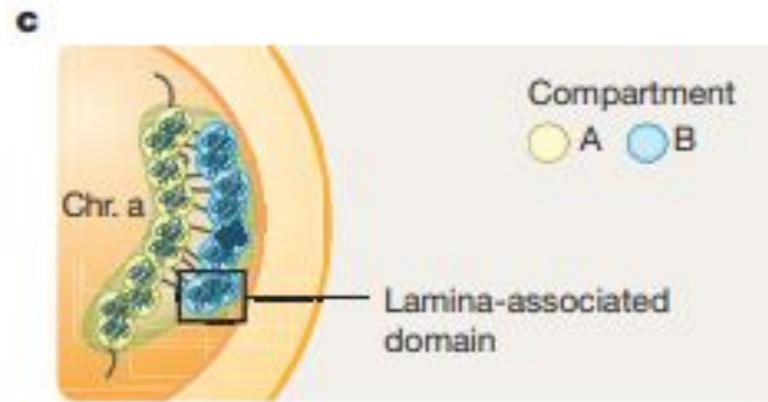
# Compartments in territories

Each territory can be divided into A and B compartments

A is mostly euchromatic

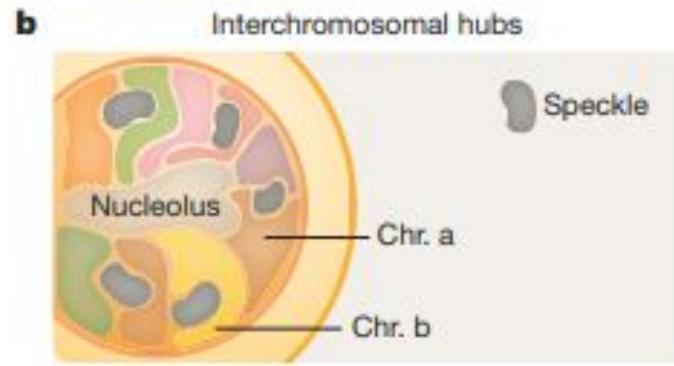
B is lamina-associated and mostly heterochromatic

Transcription factors (TFs) can reposition genes between compartments



# Nuclear hubs

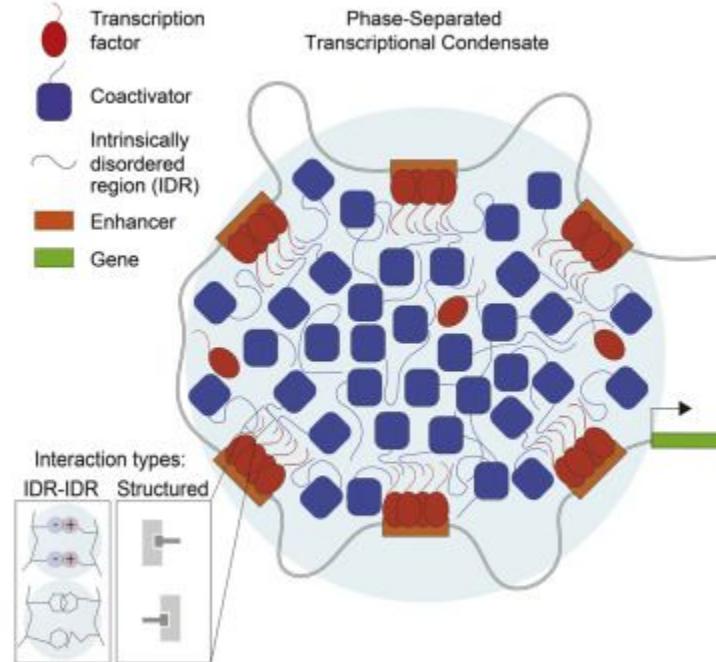
Higher order structures that form between different chromosomes



# TF-induced condensates

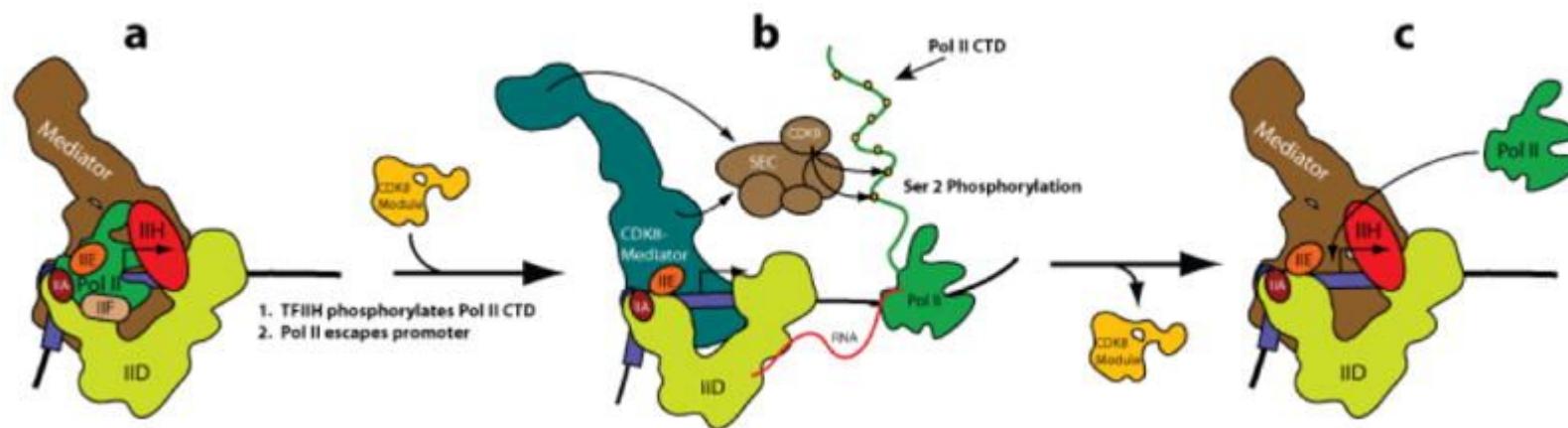
Transcription factors form phase-separated droplets via their activating domains

These droplets also contain Mediator



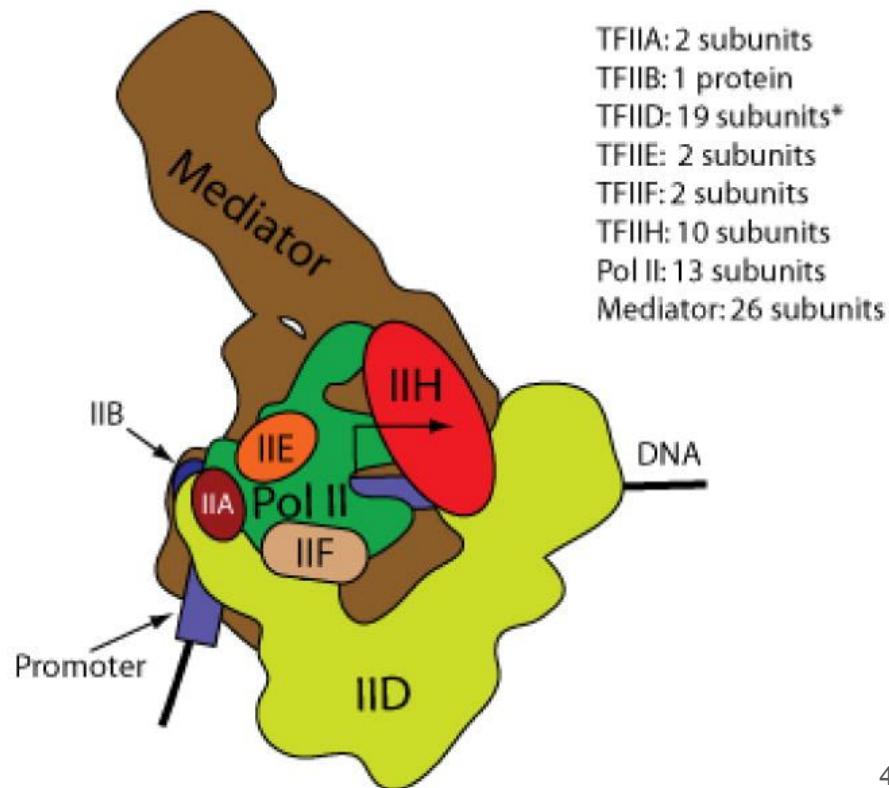
# Mediator

Communicates regulatory signals from DNA-bound TFs directly to the RNA polymerase II (pol II) enzyme



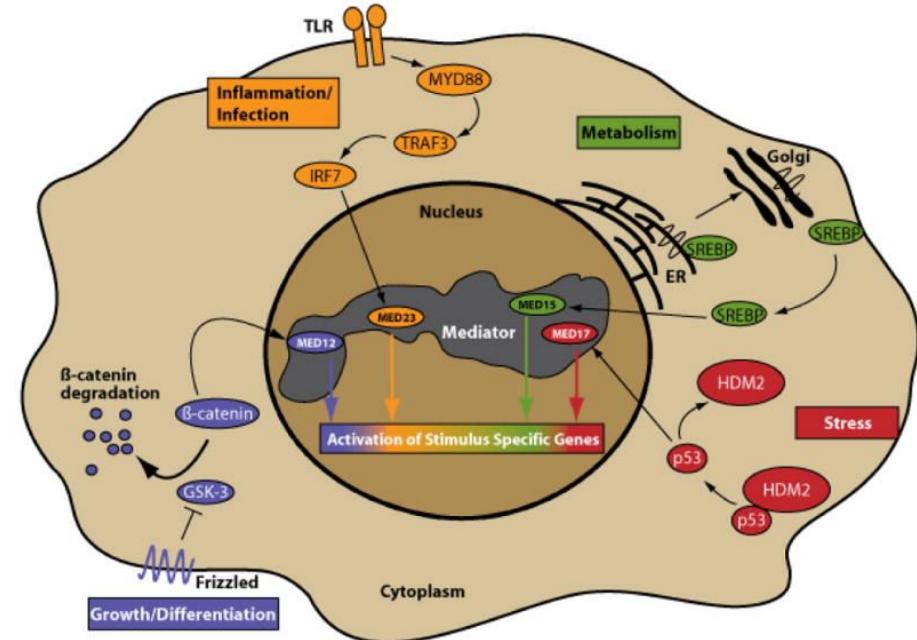
# Mediator

Huge multisubunit complex



# Mediator

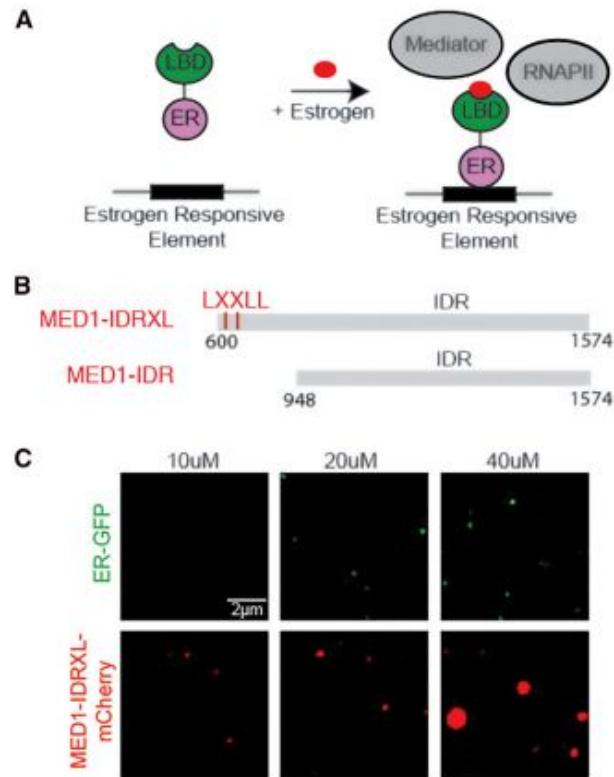
Different components of mediator are regulated via different signal pathways



# TF-induced condensates

Multiple different TFs form droplets with Mediator

Estrogen receptor forms droplets with Mediator in estrogen-dependent manner

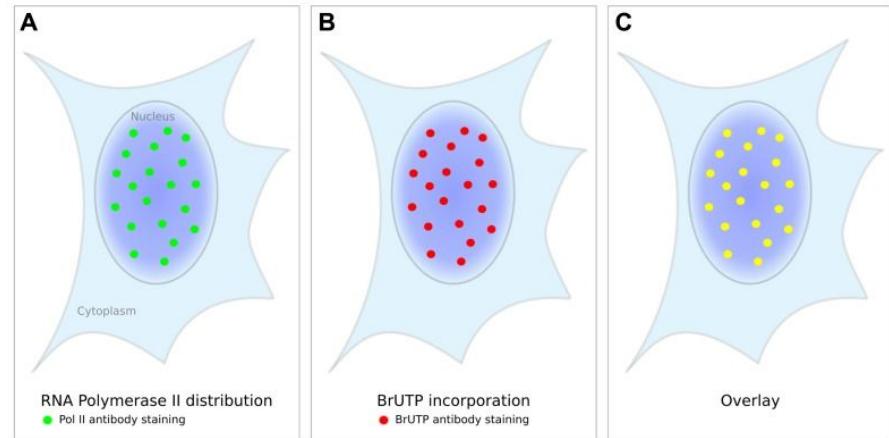


# Transcription factories

Colocalization of RNAPII and UTP incorporation

100-8000 depending on cell state

Transcription occurs at discrete sites called factories



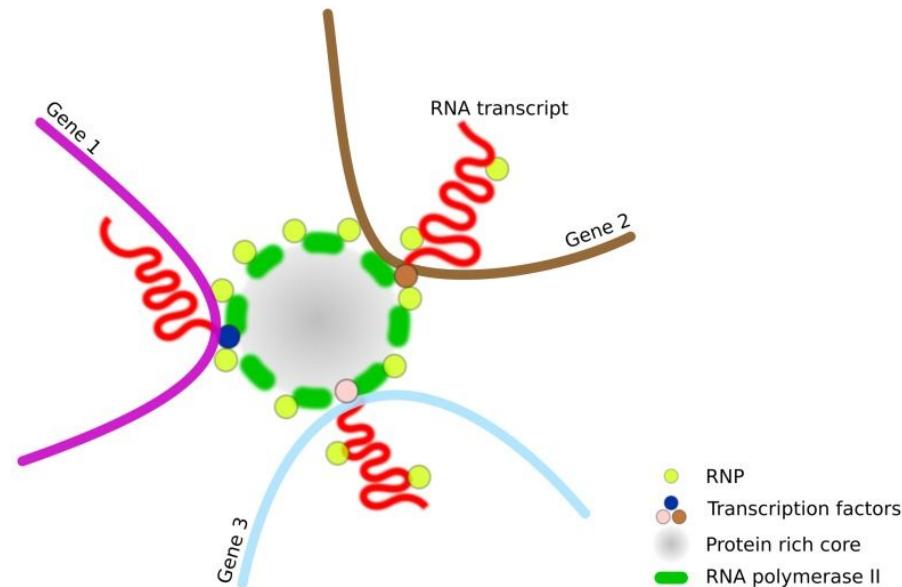
# Transcription factories

Contain 4–30 stationary RNA polymerase II molecules which are located on the surface of a protein-rich core

Contain co-activators, chromatin remodelers, transcription factors, histone modification enzymes, RNPs, RNA helicases, and splicing and processing factors

40 to 198 nm in size

Structure of a transcription factory

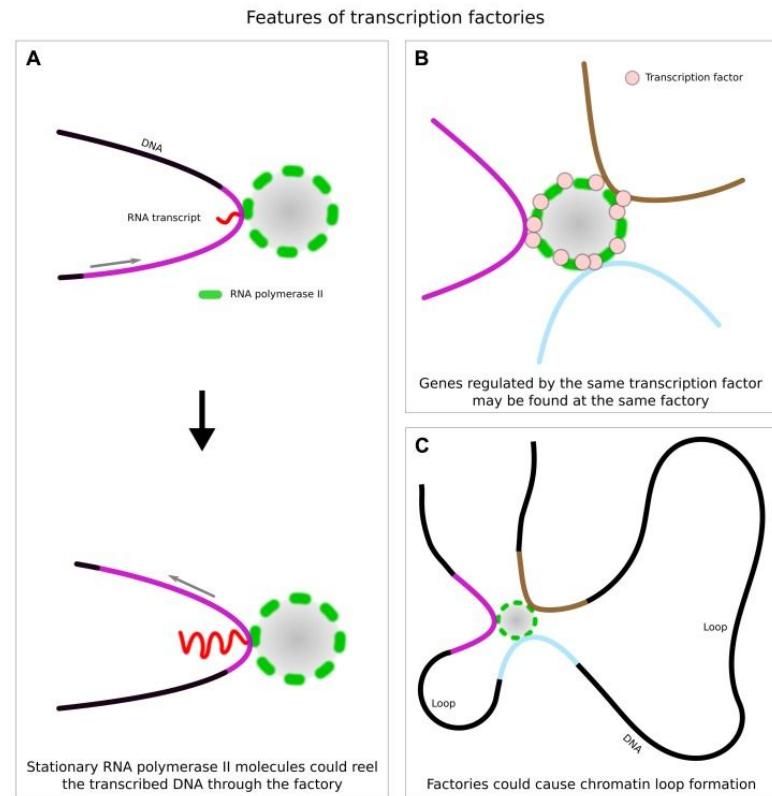


# Transcription factories

Transcription factories are probably stationary

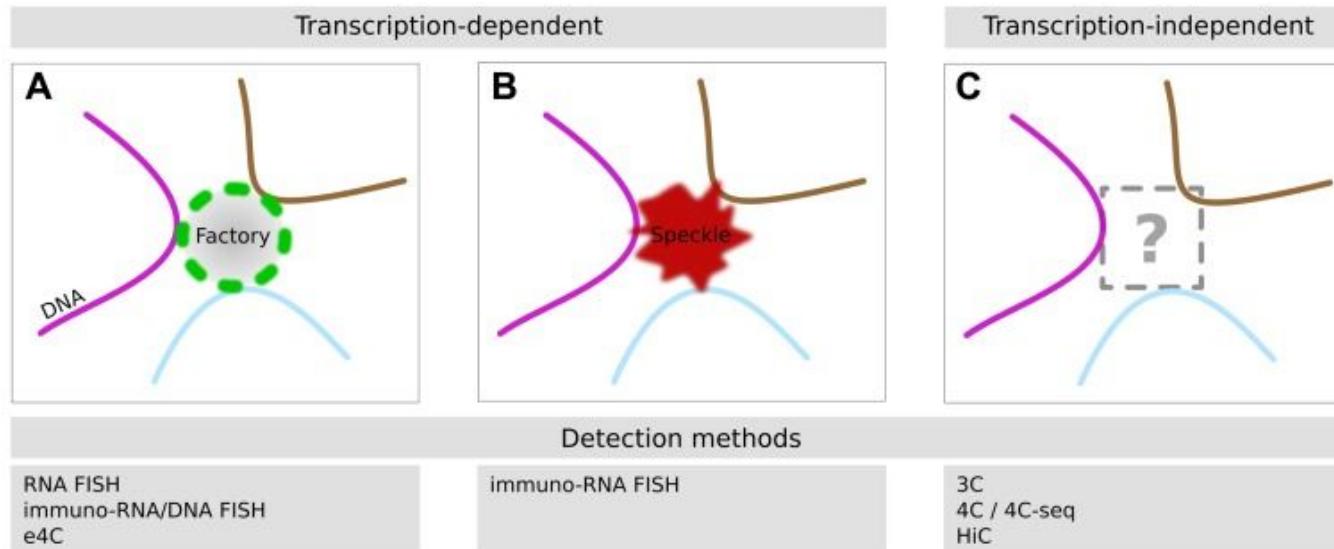
They may have TF preference

Can cause chromatin looping



# Several models of gene clustering

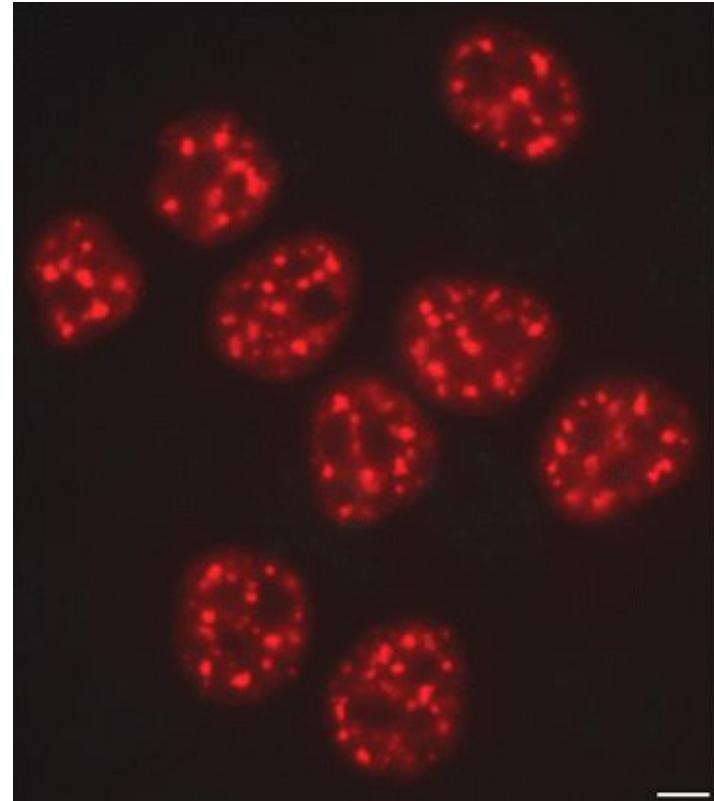
Models for gene clustering



# Nuclear speckles

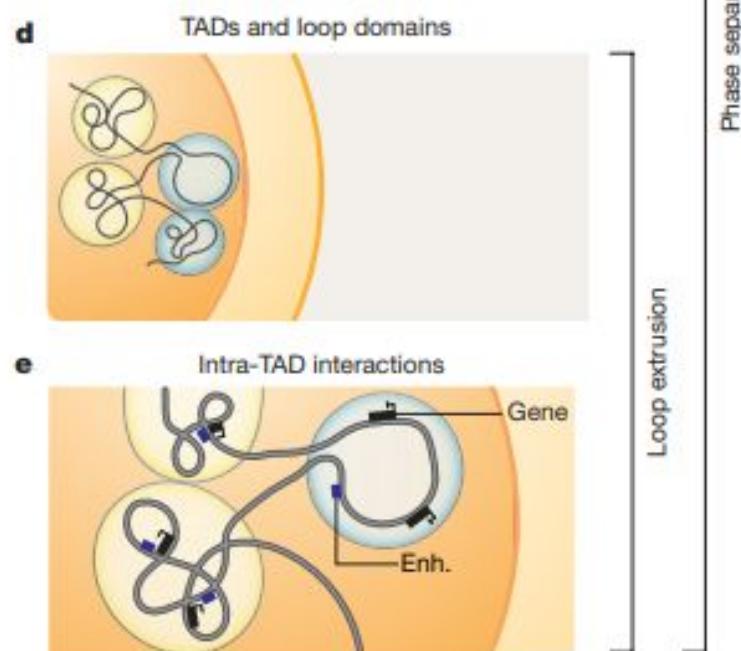
Contain splicing machinery (snRNP and SR proteins)

Some other proteins - translation factors, transcription factors



# Interchromosomal domains

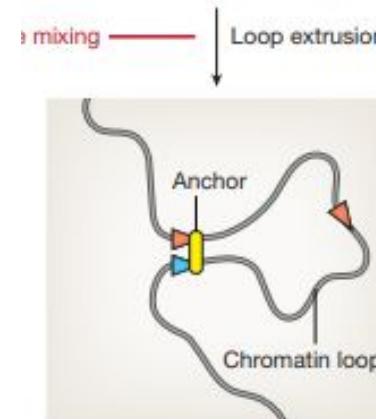
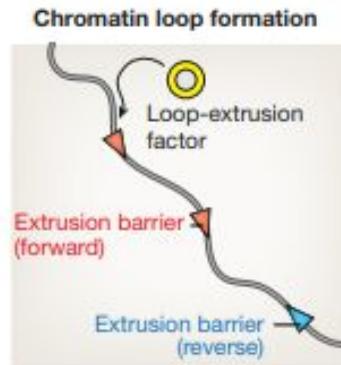
Topologically associating domains  
(TADs)



# TAD formation

Cohesin serves as a loop for DNA extrusion

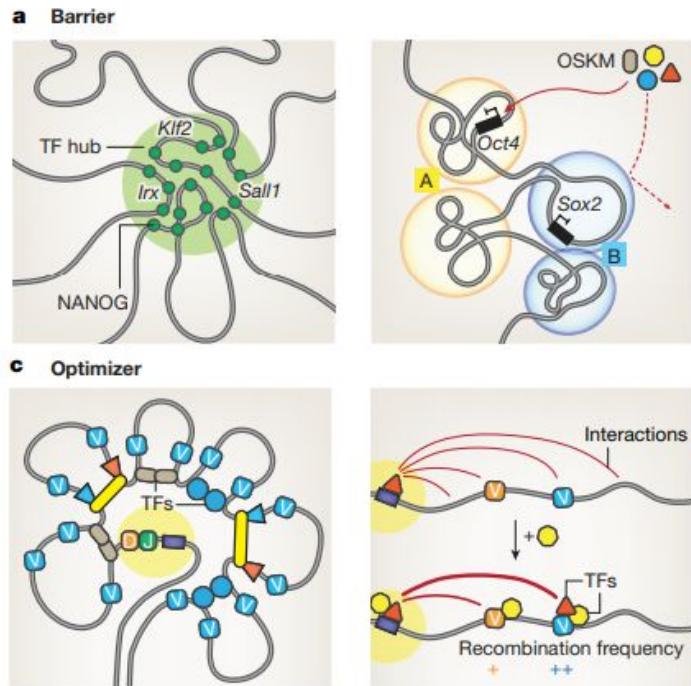
CTCF is a stopping signal. Separates heterochromatin



# Topology and gene functioning

A - maintaining expression levels to ensure stability of cell fate

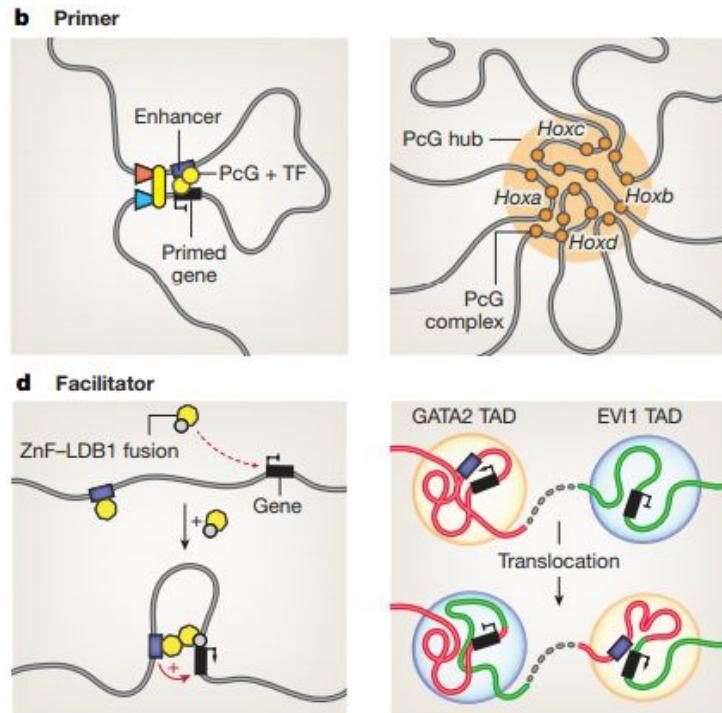
C - TFs initiate locus contraction to ensure that all VDJ segments are equally available



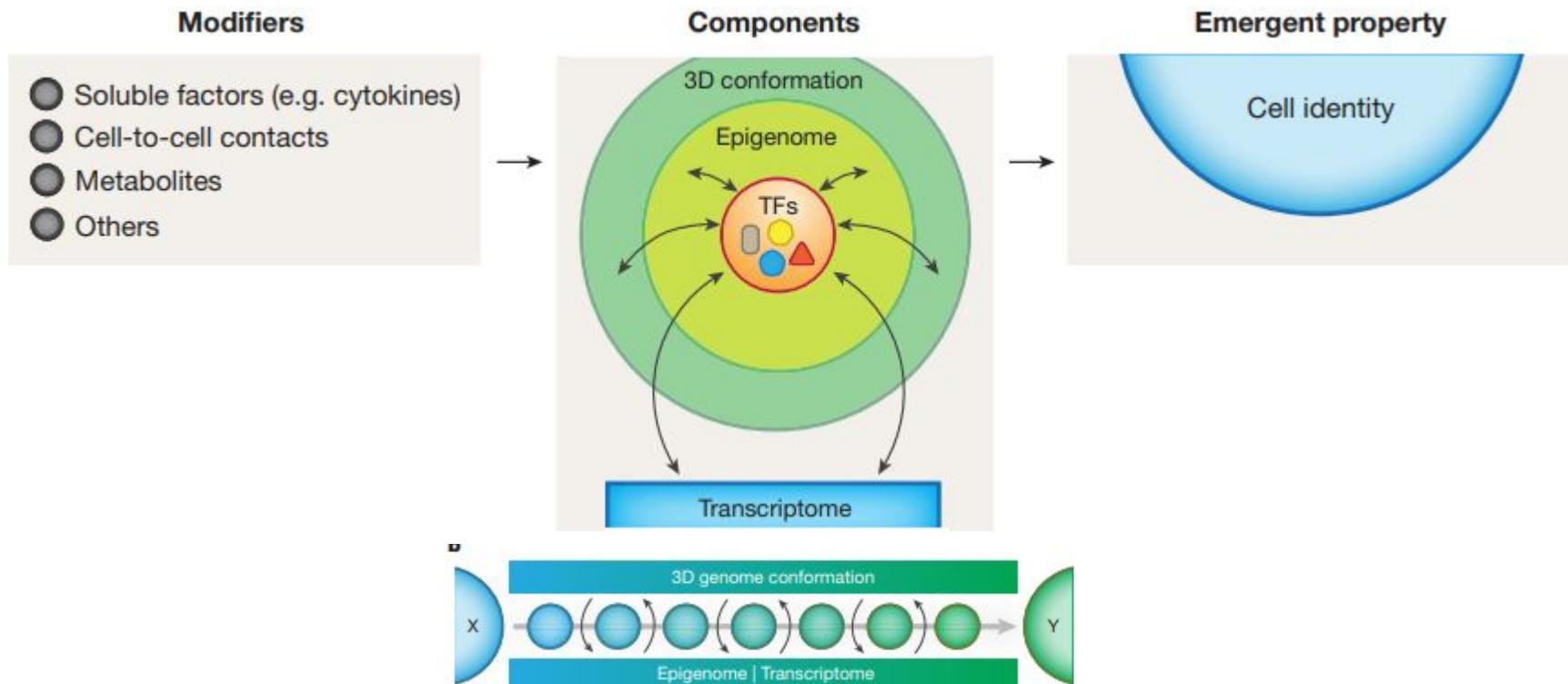
# Topology and gene functioning

B - TFs may ensure that genes are primed for rapid activation

D - Genome architecture alteration alone can activate genes without signal transmission

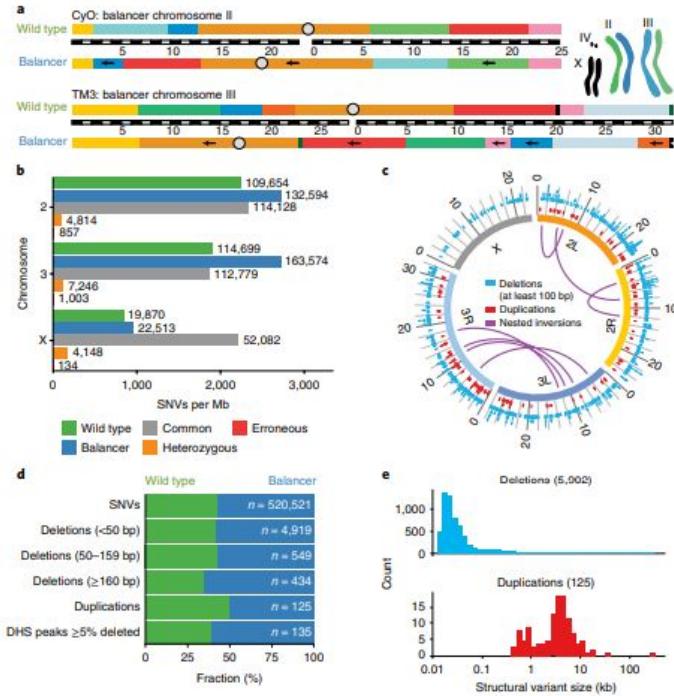


# Overall regulatory scheme



# 3D structure is not everything

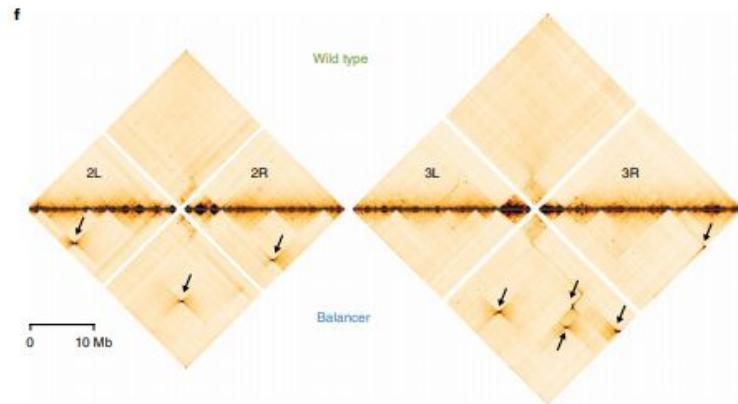
*Drosophila* lines with highly rearranged genome



14

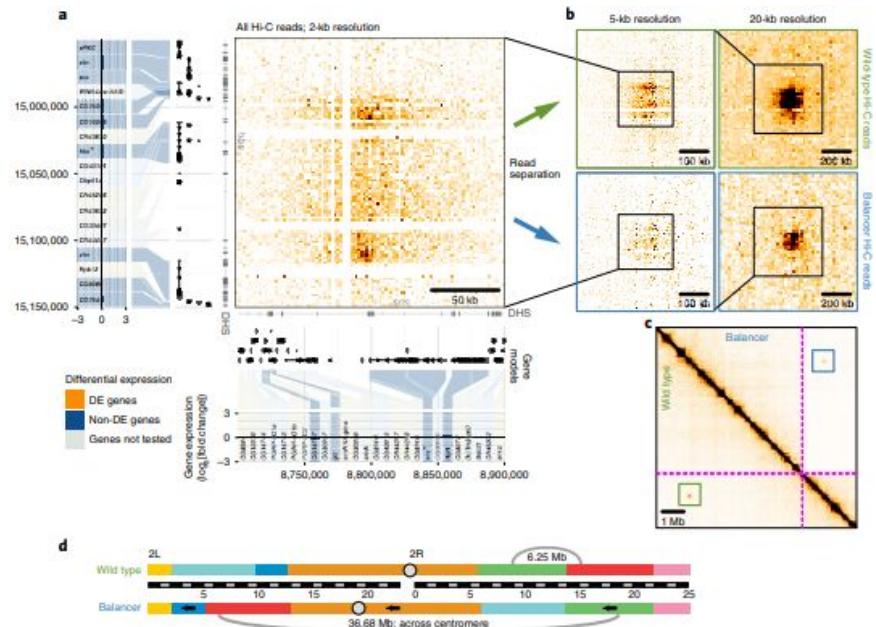
# TAD disruption

Disrupted both long and short-range interactions



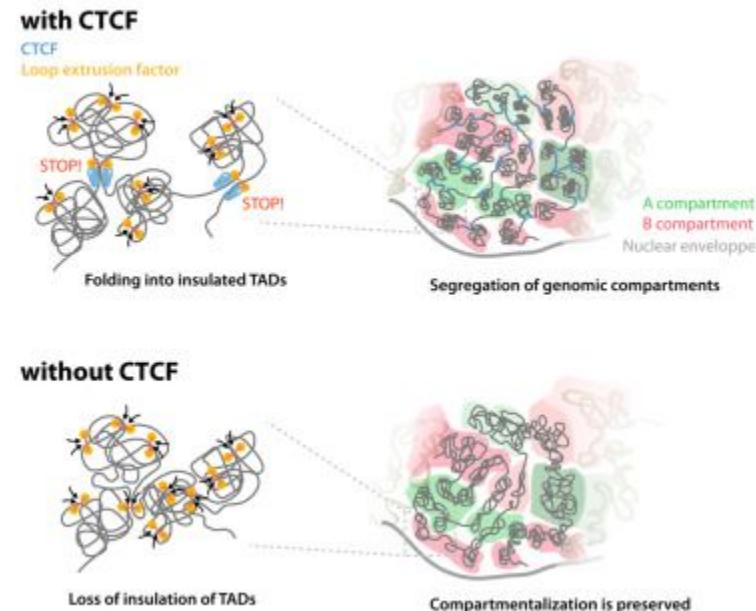
# Gene expression changes

Gene expression near rearrangements  
was mostly unaltered



# CTCF loss

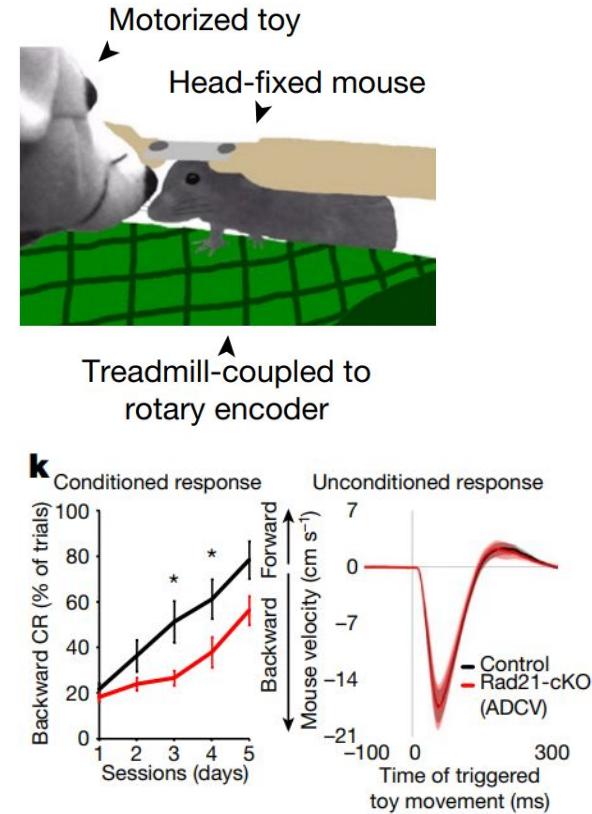
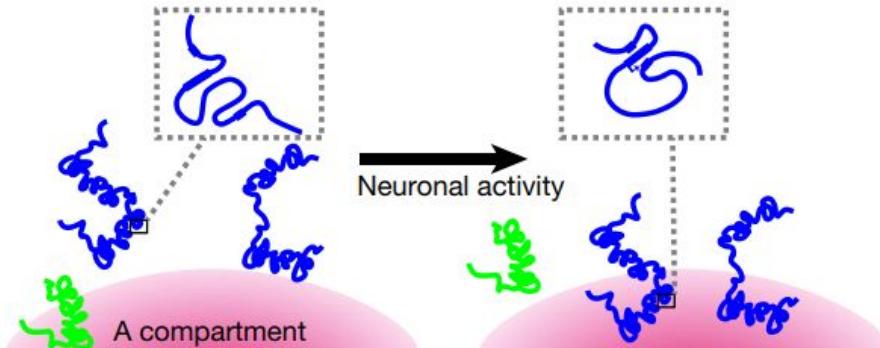
CTCF degradation leaves genomic compartmentalization mostly unaffected



# Manipulating 3D architecture to learn

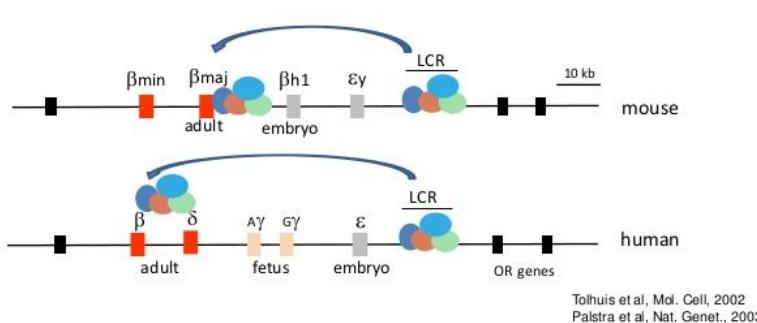
Motor learning causes chromatin rearrangement

Depletion of cohesin impairs learning



# Manipulating 3D architecture for therapy

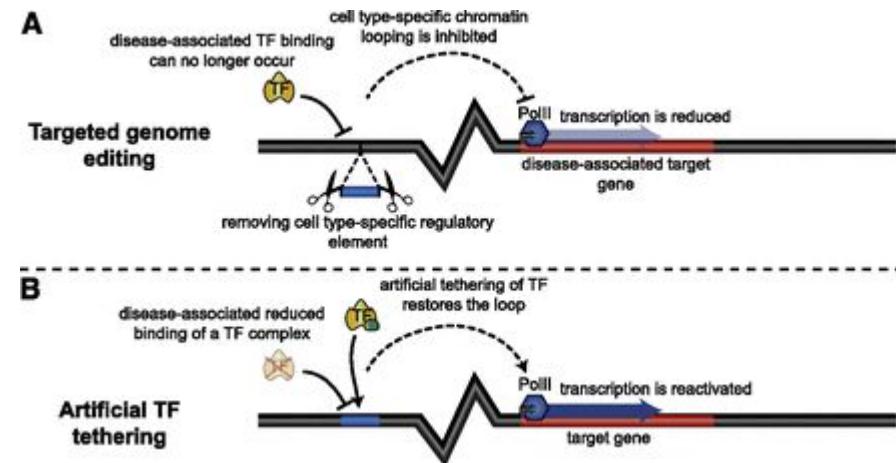
## The mammalian $\beta$ -globin loci



Locus control region (LCR) enhancer loops to genes determining which one is active as development proceeds



Drissen et al., *Genes Dev.*, 2004  
Vakoc et al., *Cell*, 2005  
Song et al., *Mol. Cell*, 2007  
Yun et al., *NAR*, 2014

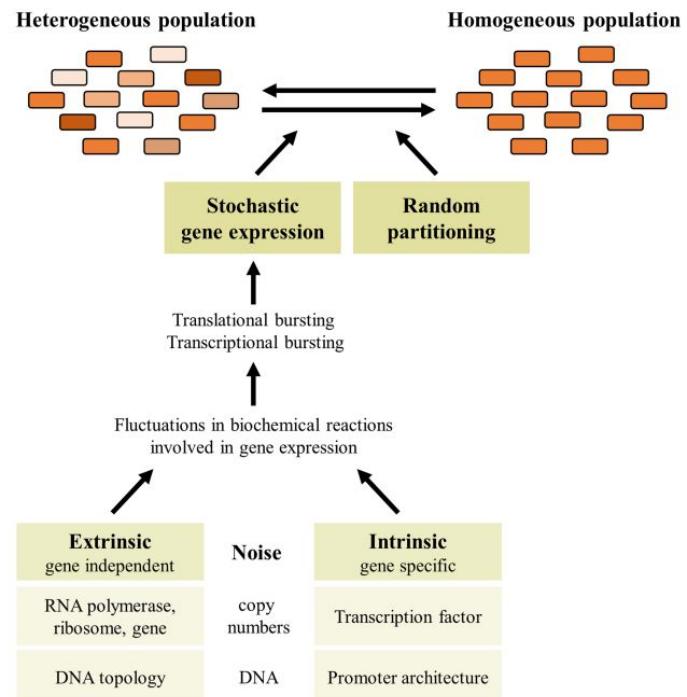


# Why do we need biological replicates?

# Cell-to-cell variation

Living cells are not homogenous even in the same environment

There are different reasons for that

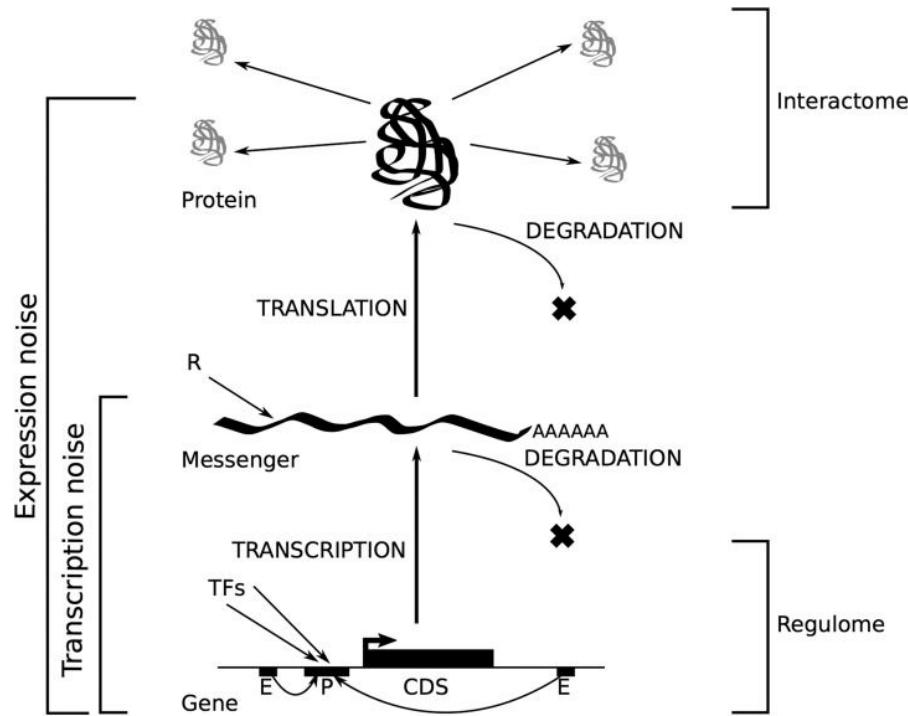


# Expression noise

Biochemical reactions rely on interacting molecules

Diffusion and binding are stochastic

Isogenic cells in the same environment can be different



# Stochasticity in cells

Intrinsic stochasticity is “generated by the dynamics of the system from the random timing of individual reactions” and extrinsic stochasticity is “generated by the system interacting with other stochastic systems in the cell or its environment.” (Shahrezaei and Swain, 2008)

# Transcription noise

Intrinsic - gene-specific (promoter architecture, transcription factors)

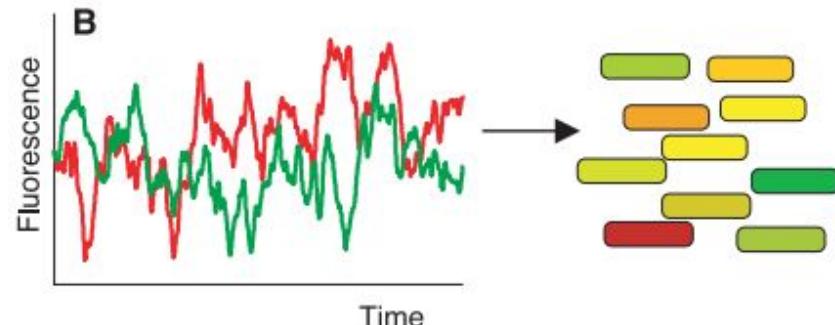
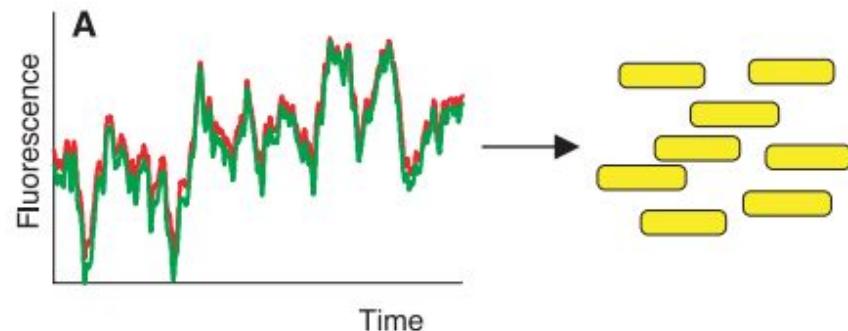
Extrinsic - cell-specific (availability of RNAP, DNA topology, ribosomes, copy number variation)

Extrinsic gene independent	Noise	Intrinsic gene specific
RNA polymerase, ribosome, gene	copy numbers	Transcription factor
DNA topology	DNA	Promoter architecture

# Measuring transcription noise

A - extrinsic noise. Cells may differ but all genes are equally affected

B - intrinsic noise. Genes differ



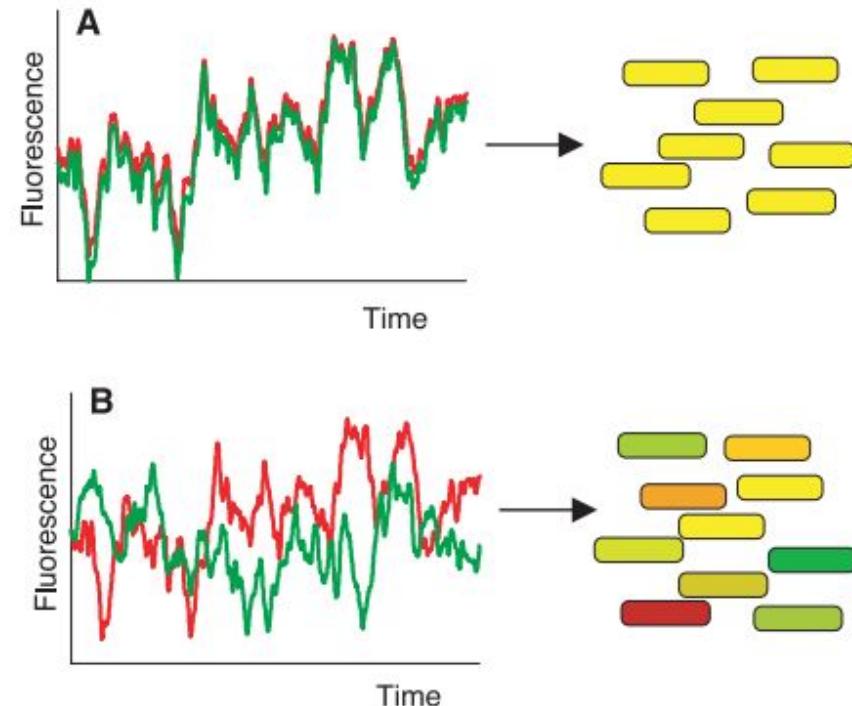
# Measure different noises

YFP and CFP controlled by lac promoter

Genes are equidistant from oriC

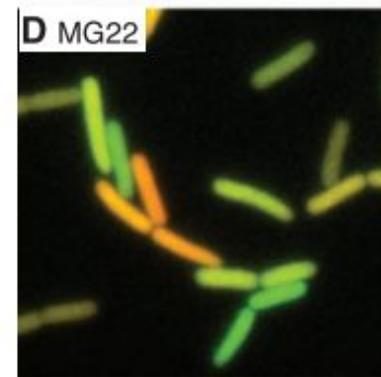
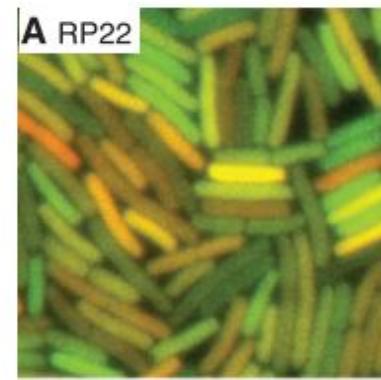
Fluorescent intensity difference tells about intrinsic noise

Total noise  $\wedge 2 =$  Intrinsic noise  $\wedge 2 +$  Extrinsic noise  $\wedge 2$



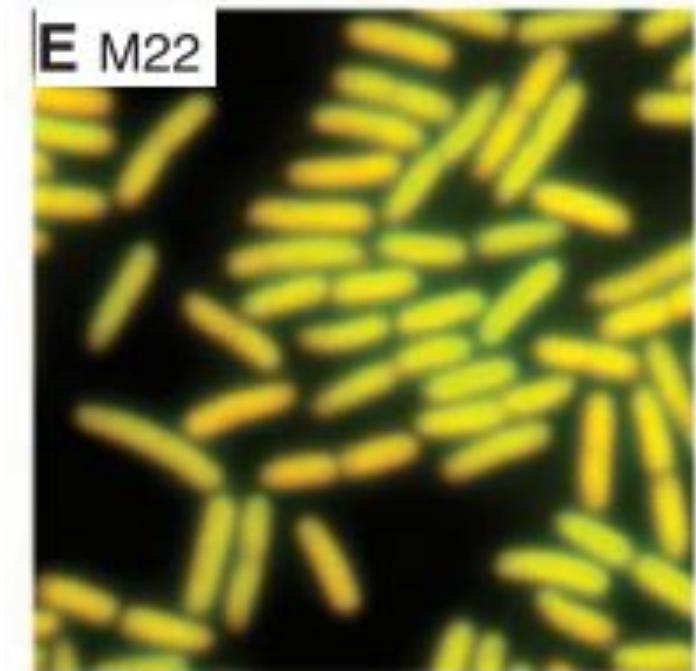
# Distinguishing between extrinsic and intrinsic noise

Wild-type (i.e. lacI<sup>+</sup>) strains are noisy



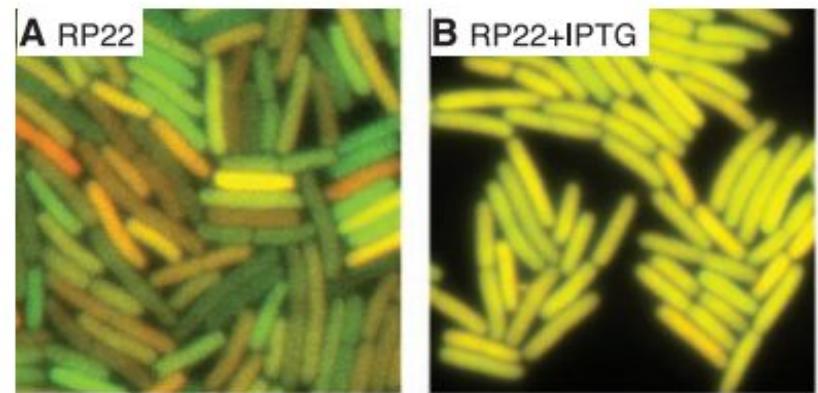
# Distinguishing between extrinsic and intrinsic noise

Strain without lacI is not noisy



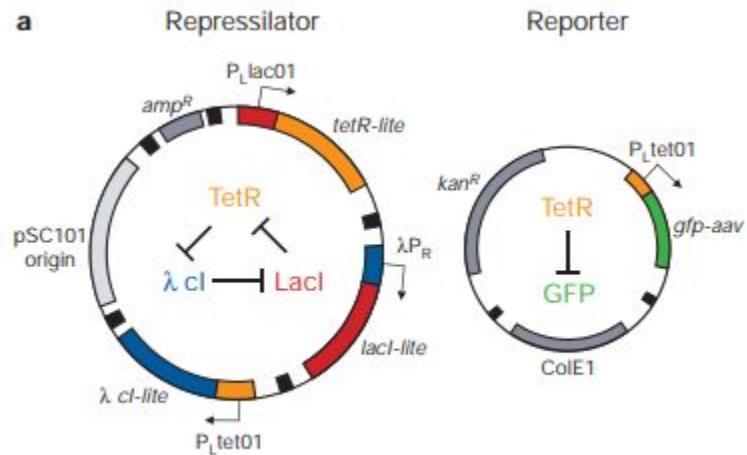
# Distinguishing between extrinsic and intrinsic noise

Blocking lacI with IPTG reduces noise



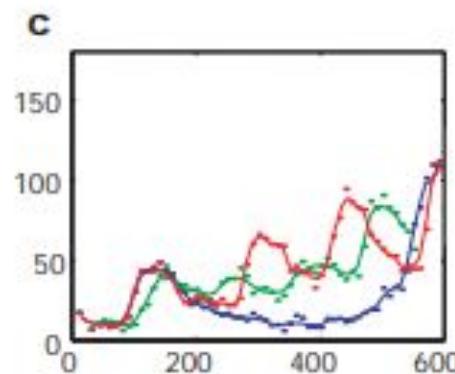
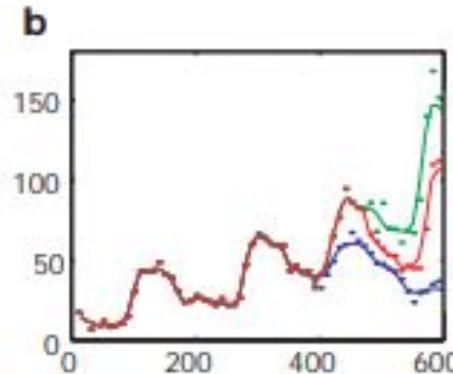
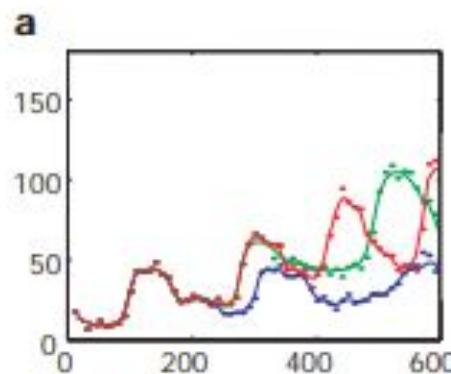
# Repressilator

Oscillatory behaviour because of mutually repressed genes



# Repressilator

Oscillatory behaviour rapidly falls apart :(

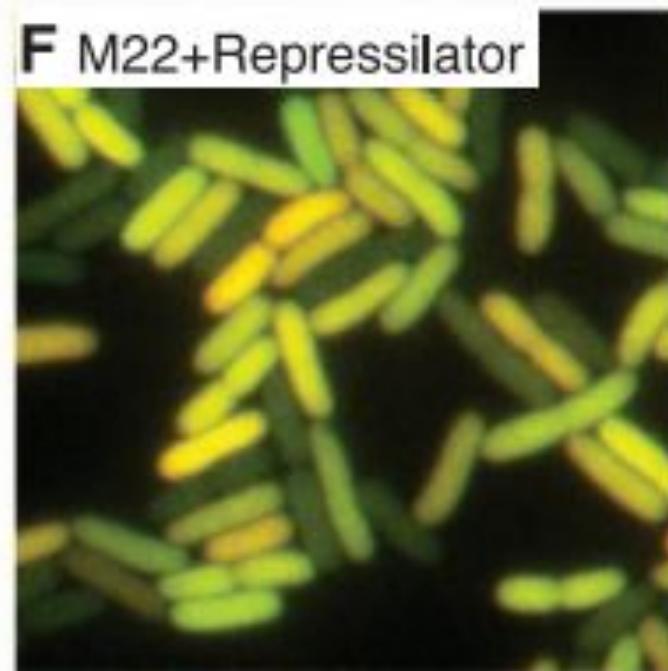


# Distinguishing between extrinsic and intrinsic noise

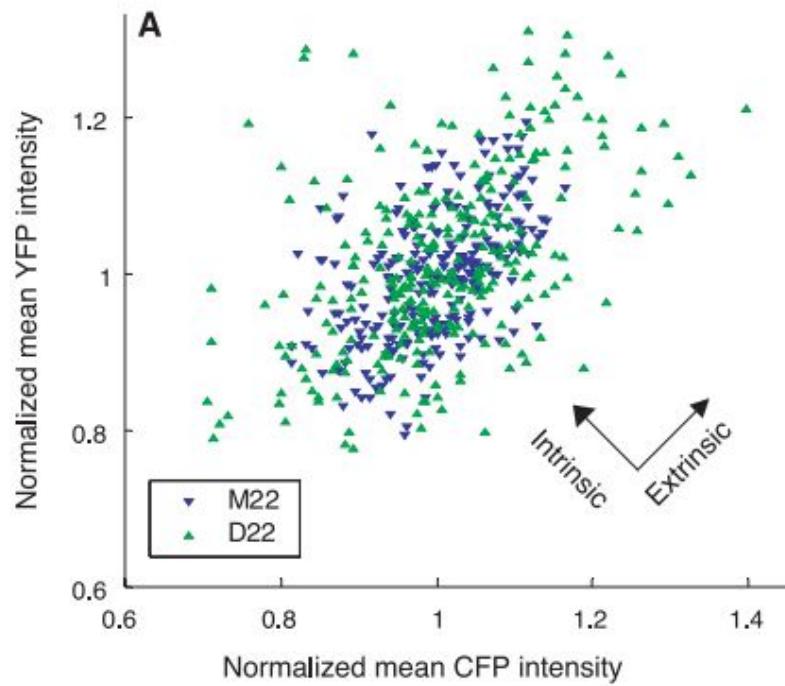
Repressilator induces noise

Also note increased total noise

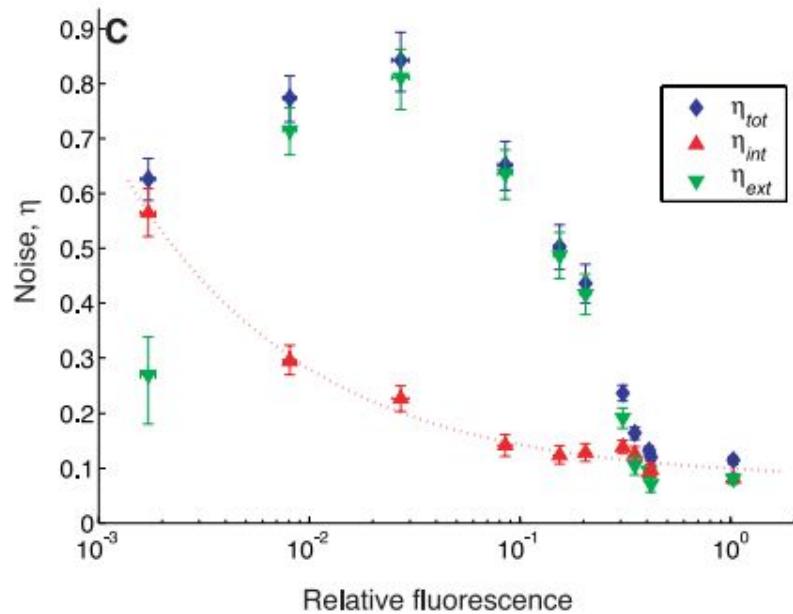
Intrinsic noise is increased when not in a steady state



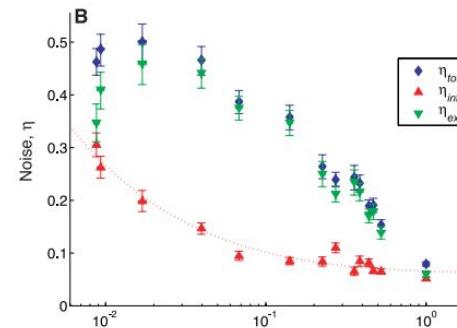
# Noisy vs non-noisy strain



# RecA decreases intrinsic noise



D22 (recA-, lacI- )

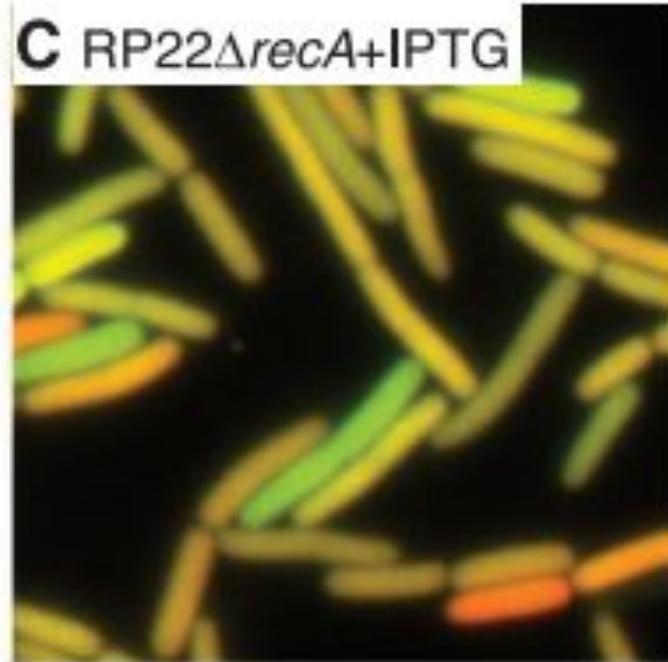


M22 (recA+, lacI- )

# Distinguishing between extrinsic and intrinsic noise

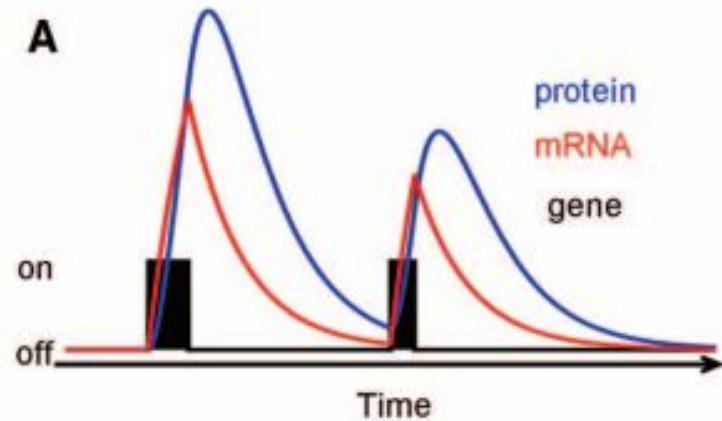
*recA* deletion increases intrinsic noise

C RP22 $\Delta$ *recA*+IPTG



# Transcription bursts

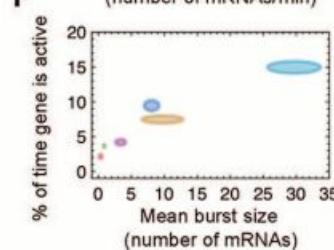
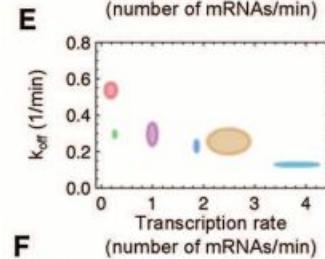
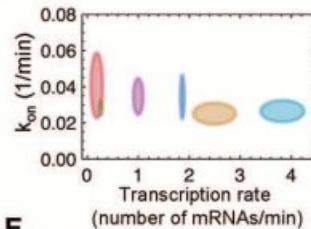
In prokaryotes and eukaryotes, most genes appear to be transcribed during short periods called transcriptional bursts, interspersed by silent intervals



# Transcription bursts

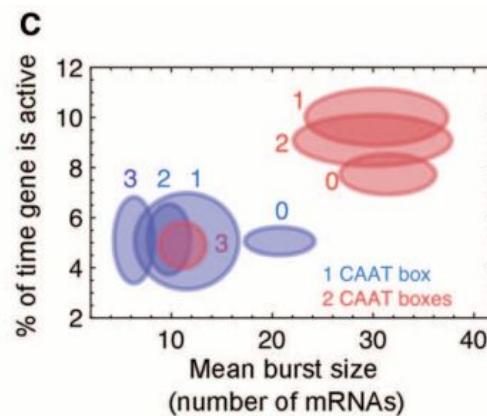
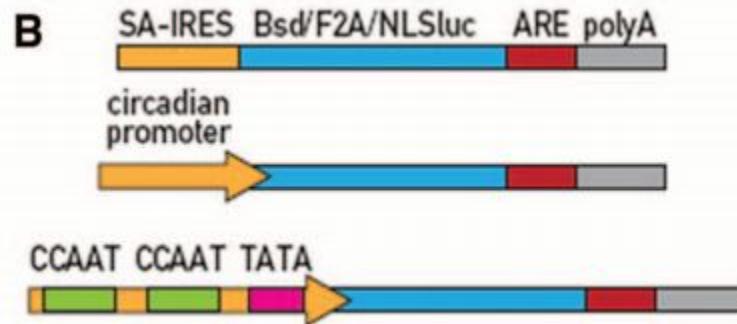
Burst kinetics is gene-specific

Bmal1a, GT:Glutaminase, GT:Prl2C2  
D GT:Serpine1, Sh3kbp1, GT:Plectin1



# Transcription bursts

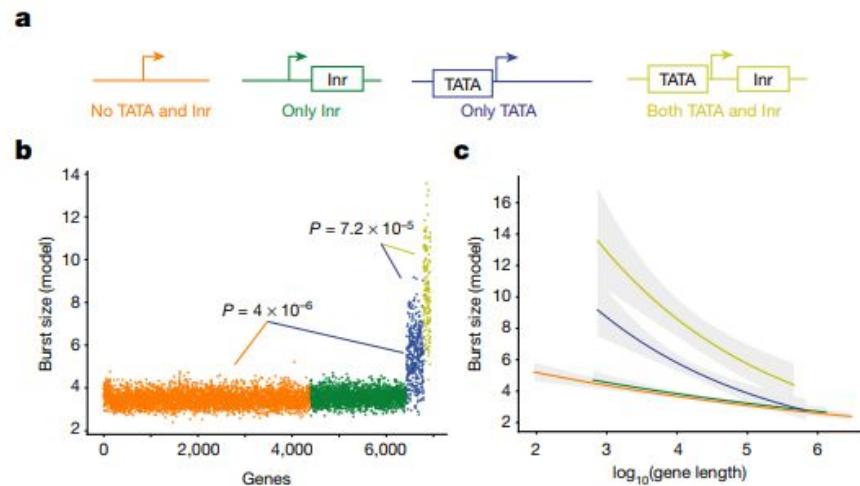
Promoter structure influences burst kinetics



# Core promoter regulates bursts

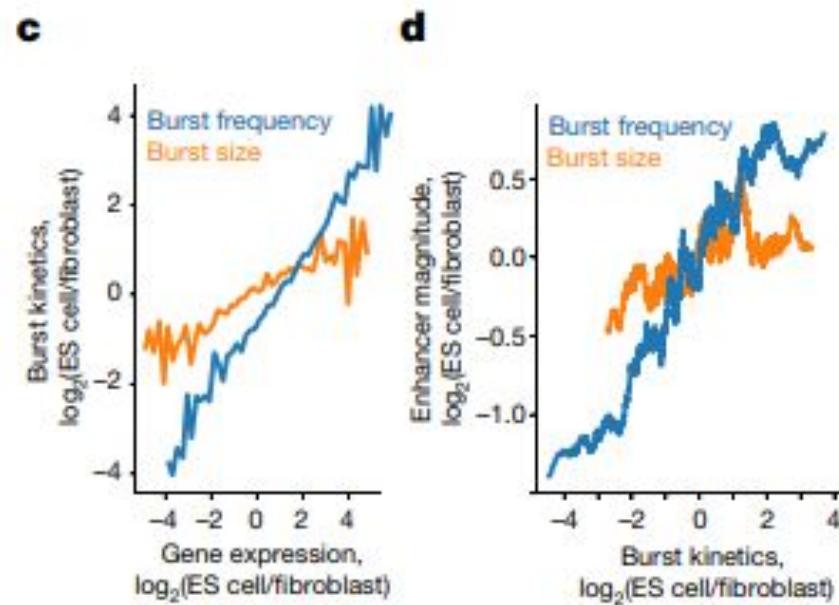
TATA increases bursts

Initiator increases bursts if TATA is present



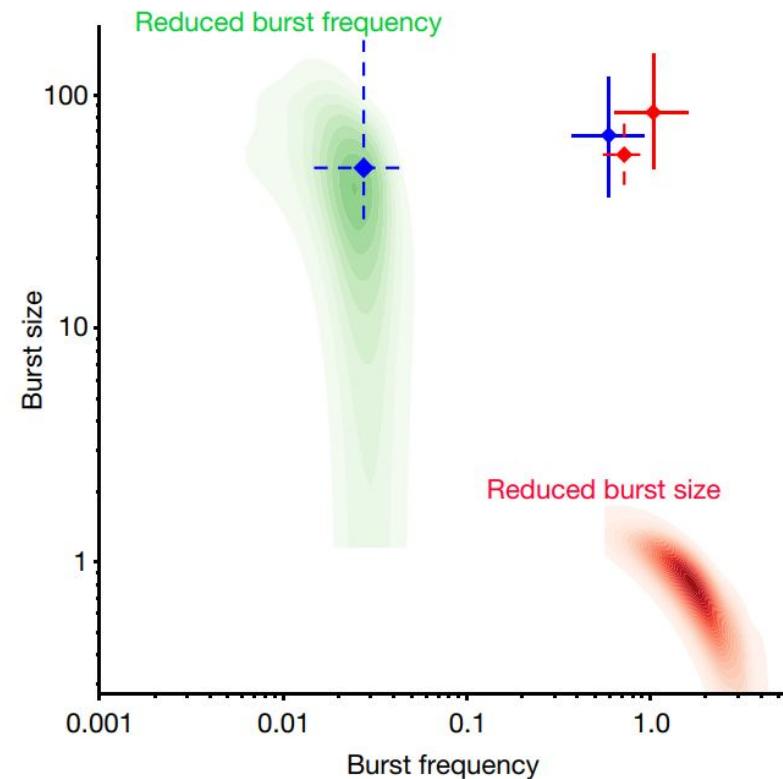
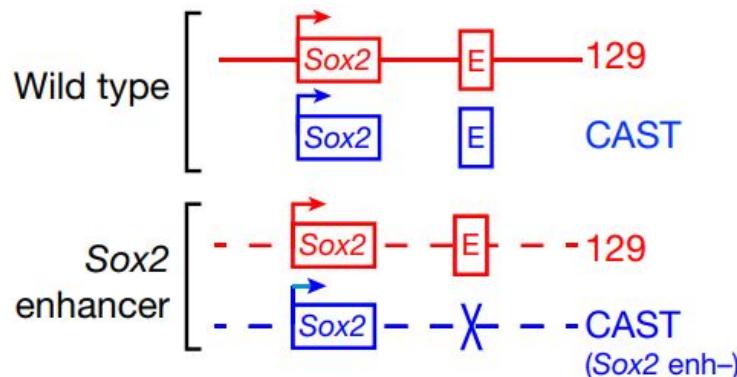
# Enhancers also regulate bursts

Transcription bursts are cell type-specific



# Enhancers also regulate bursts

Mutating enhancer can alter burst kinetics

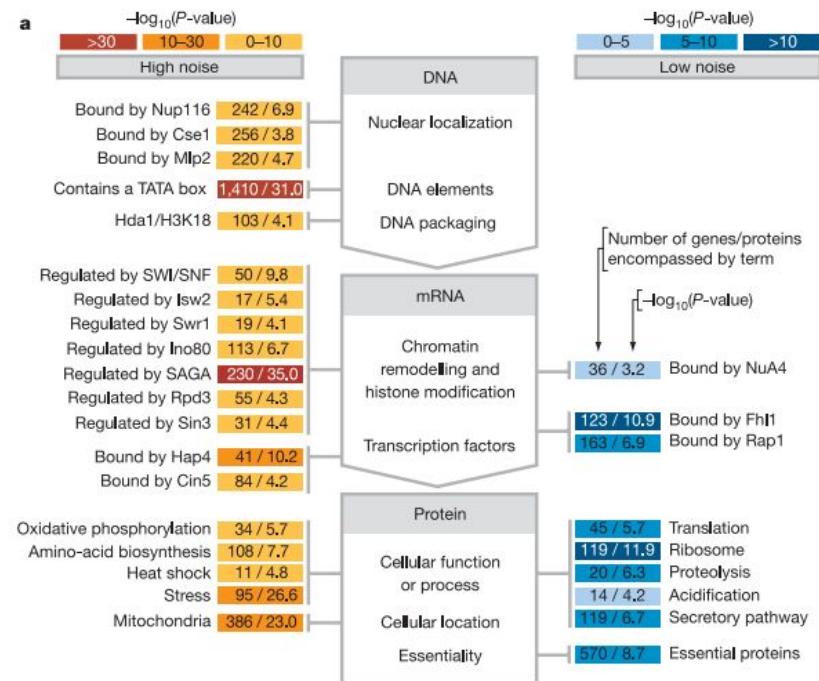


# Noise is random but noisiness is not

Biological noise differs highly across biological groups

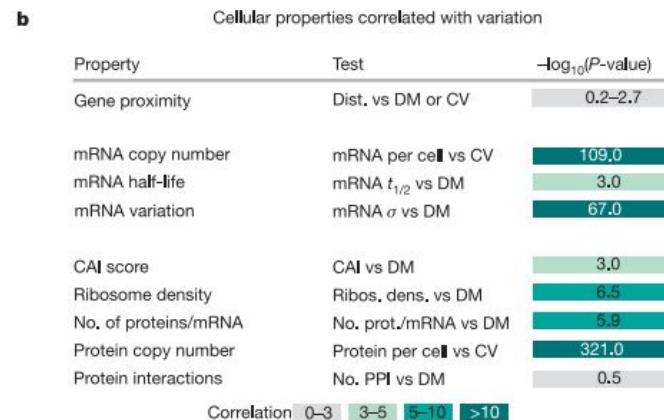
Housekeeping genes tend to be less noisy

Chromatin remodeling tends to induce noise in target genes



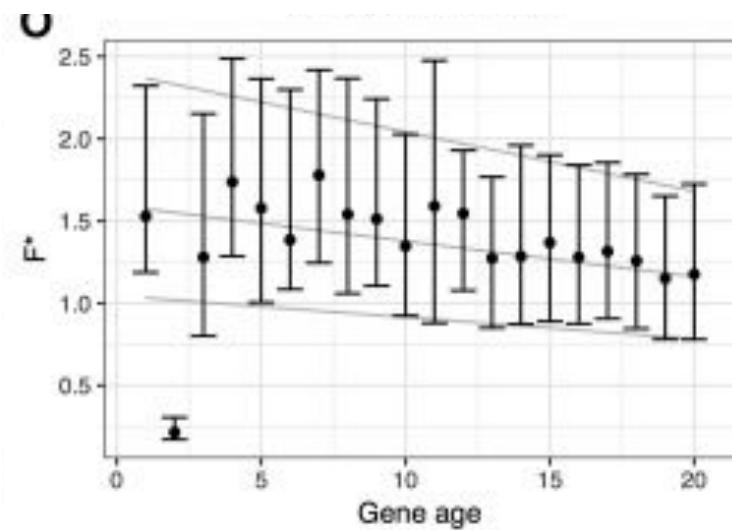
# Noise is random but noisiness is not

mRNA or protein copy number and the variation in mRNA expression are very strongly correlated with CV and DM values, respectively



# Noisiness is important in evolution

Older genes are less noisy

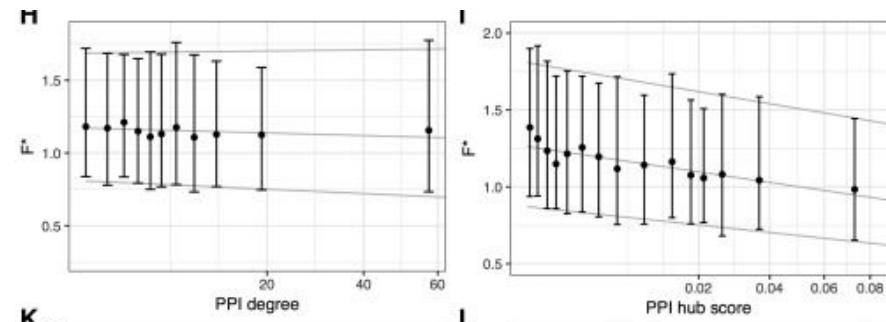


# Noisiness is important in evolution

Genes that encode highly connected proteins are less noisy

Protein position in PPI network defines transcriptional noise

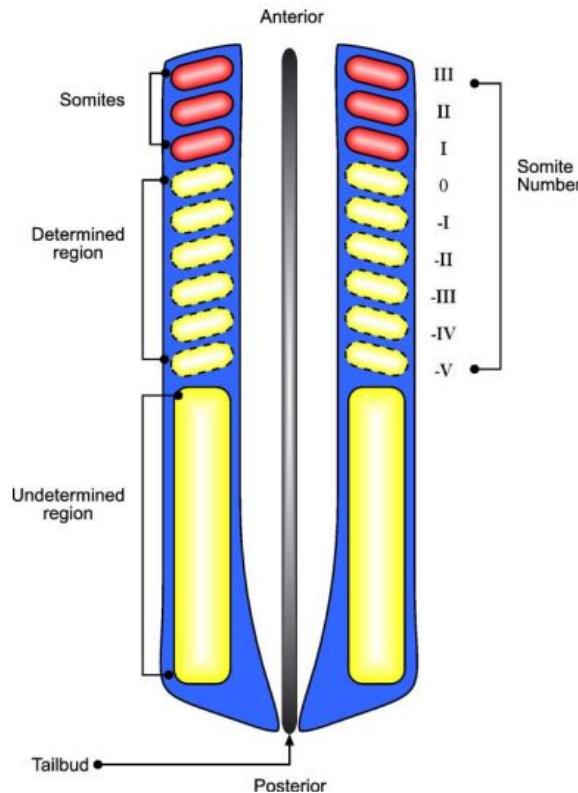
Noisiness and mean expression are separate parameters in evolution



# Cell fate vs noise

# Somite differentiation

Highly uniform process despite potential noise sources

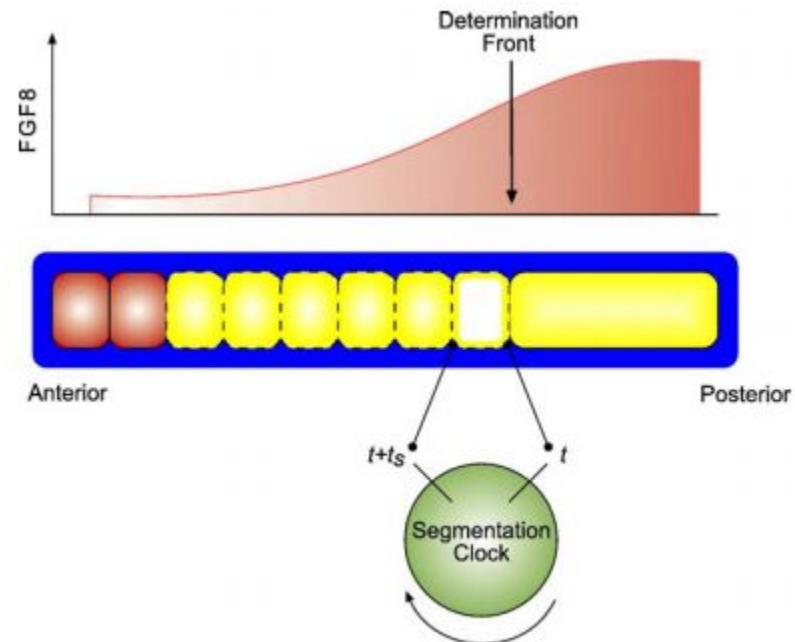


# Clock and wavefront model

Longitudinal positional information  
gradient down the AP

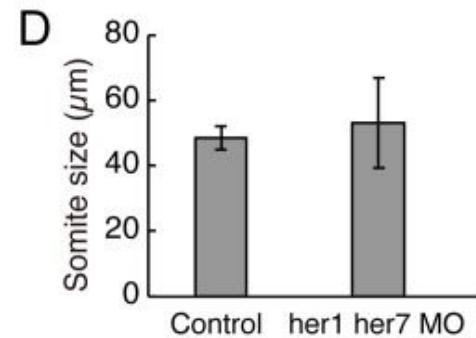
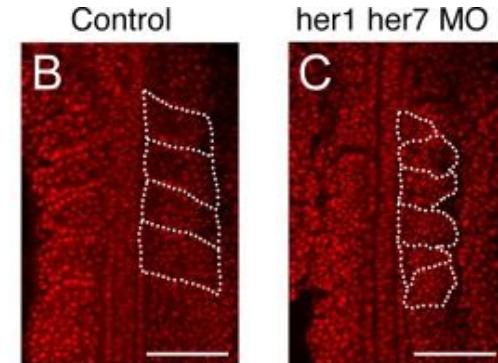
Interaction with cellular oscillator

Clock determines the time of  
catastrophe - rapid change of cell state



# Breaking clock disrupts somites

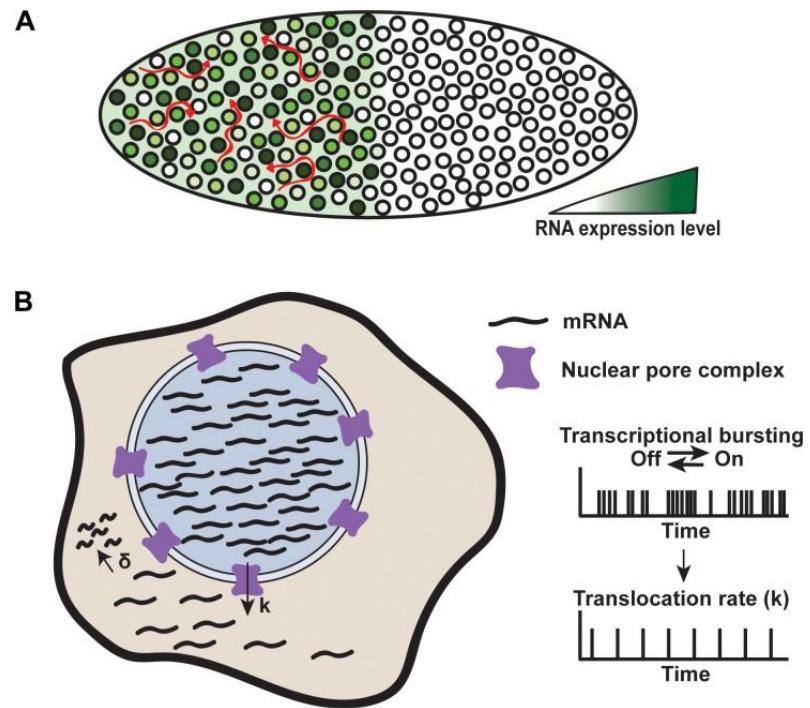
Zebrafish mutants with disrupted clock have larger variation in somite size



# Other ways to buffer noise

A - RNA levels are buffered across syncytium

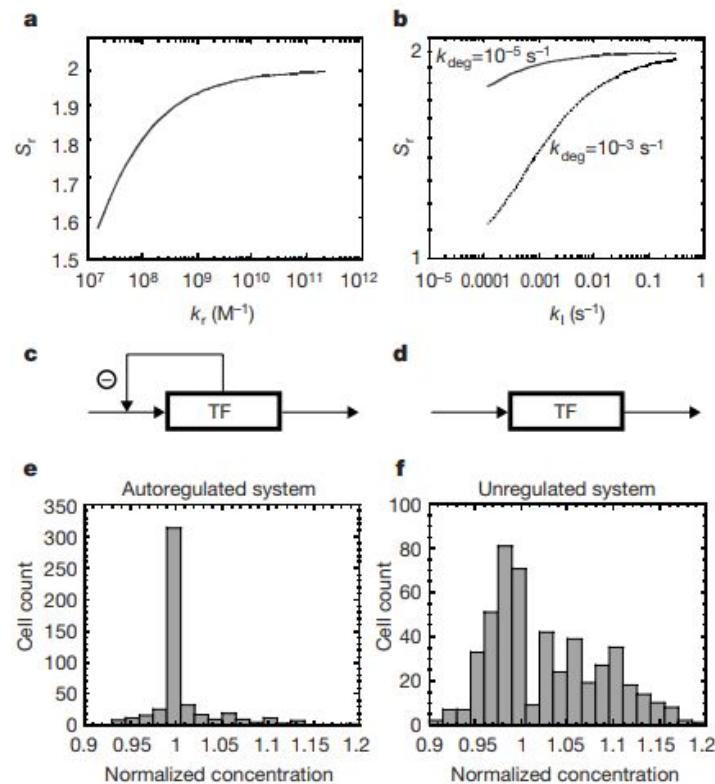
B - transcription is noisy but translocation is not



# Feedback regulation

Feedback regulation decreases noise

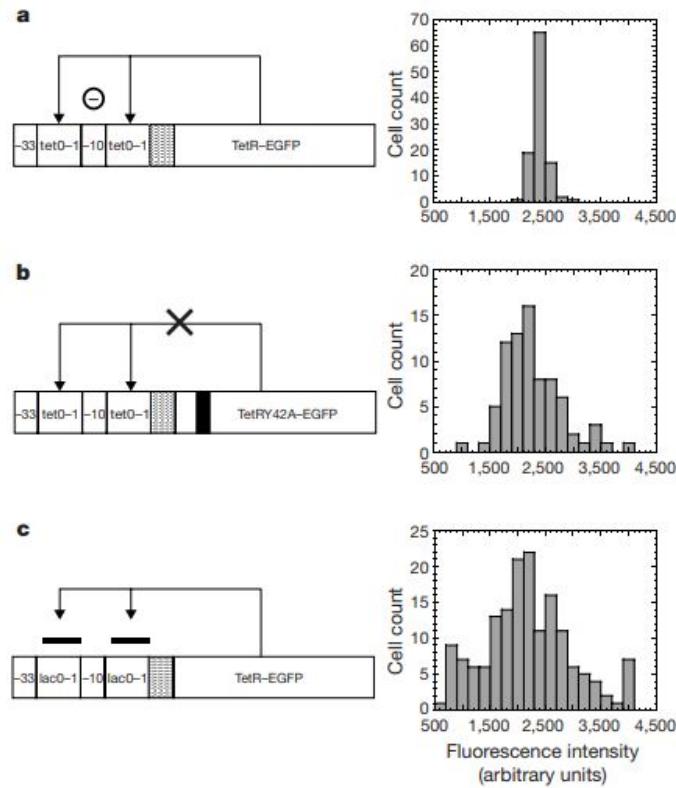
May be the reason why it is so popular  
in gene regulatory networks



# Feedback regulation

Repressing feedback regulation increases noise

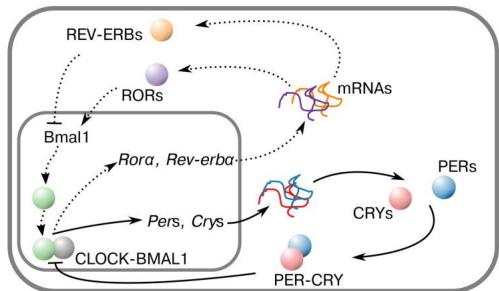
Can be done via mutating repressor or deleting operators



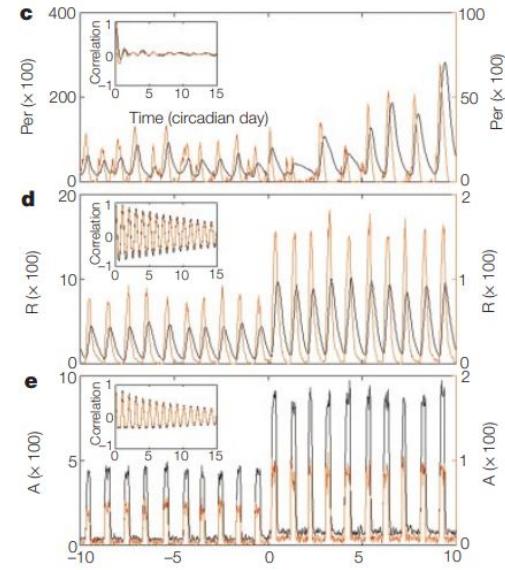
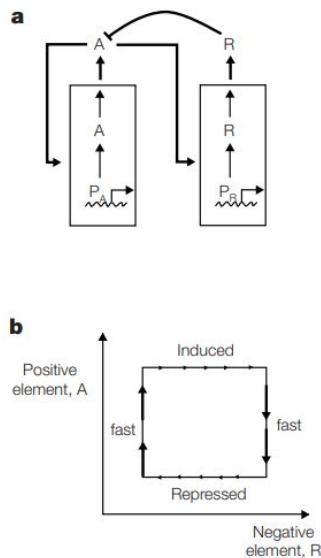
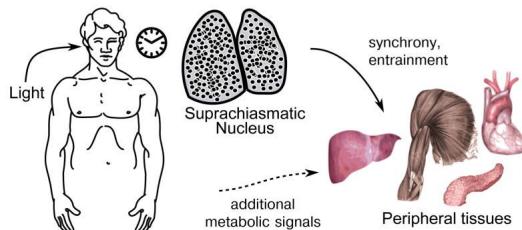
# Fighting noise in circadian clocks

Circadian systems should be noise resistant

A Mammalian Circadian Feedback Loops



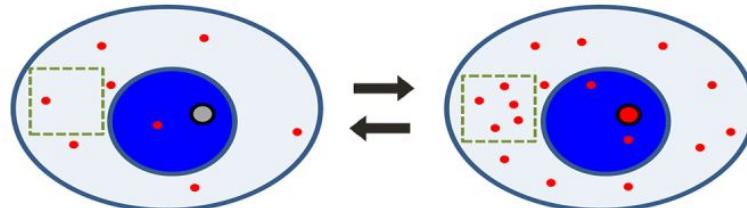
B Mammalian Circadian Hierarchy



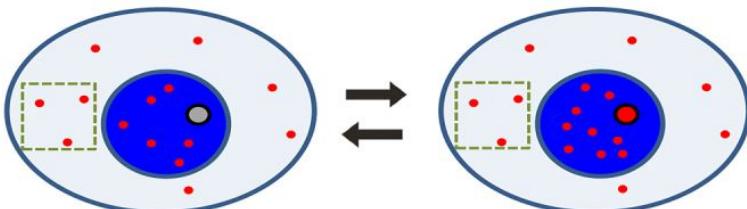
# Nuclear mRNA retention

Processed mRNA levels may be even higher in nucleus than in cytoplasm

Bursts of transcription generate cytoplasmic gene expression noise



Nuclear retention of mRNA reduces cytoplasmic gene expression noise



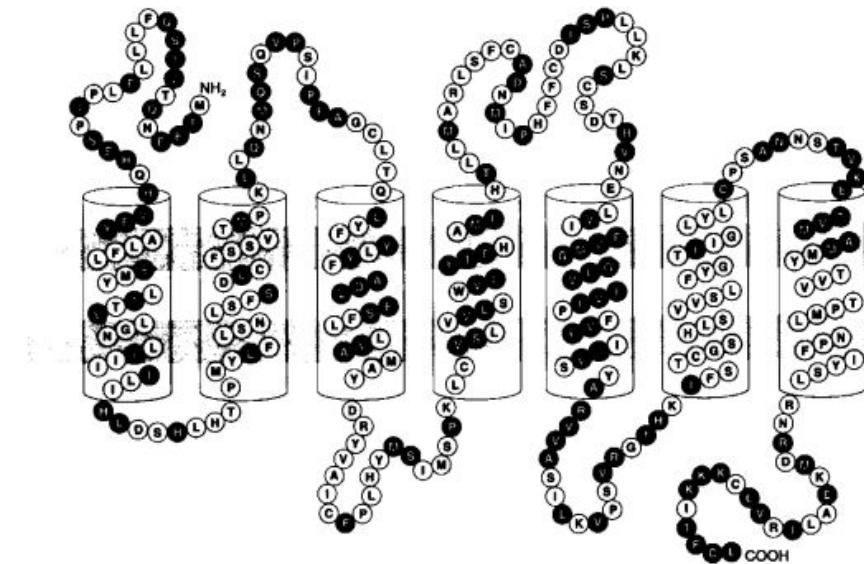
- Active transcription site
- Inactive transcription site
- mRNA

# Olfactory receptors (ORs)

Encoded by a huge multigene family

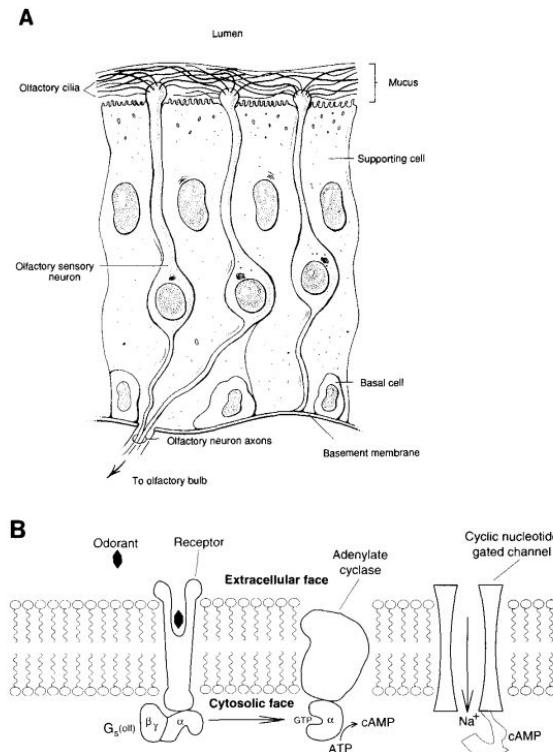
Mice have ~1000 functional genes (3% of genome)

Humans have around 400



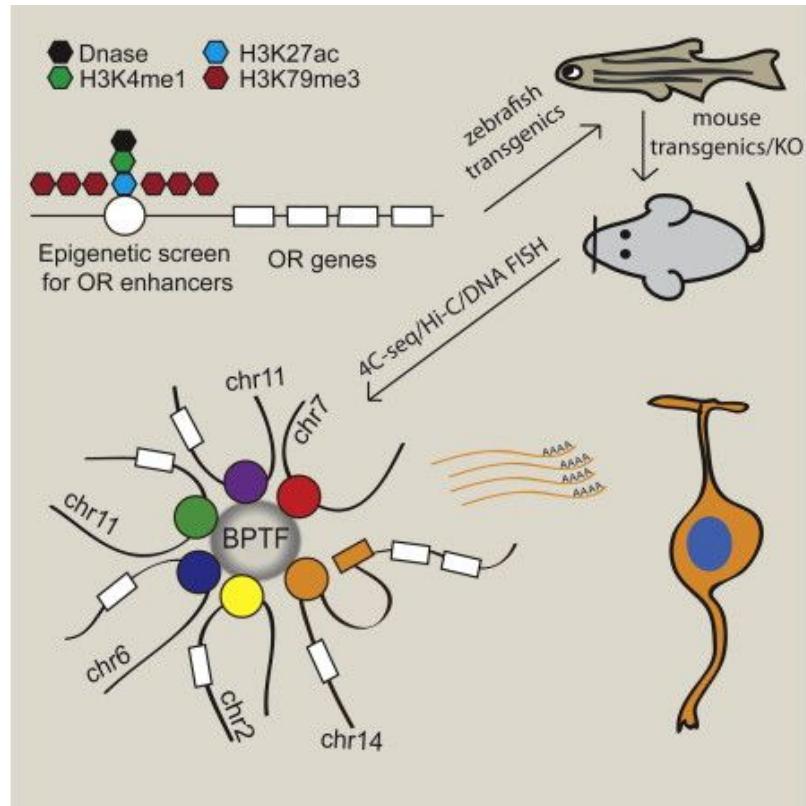
# Olfactory sensory neurons

Each mature olfactory sensory neuron expresses one olfactory receptor gene in monoallelic and stochastic fashion



# Greek islands

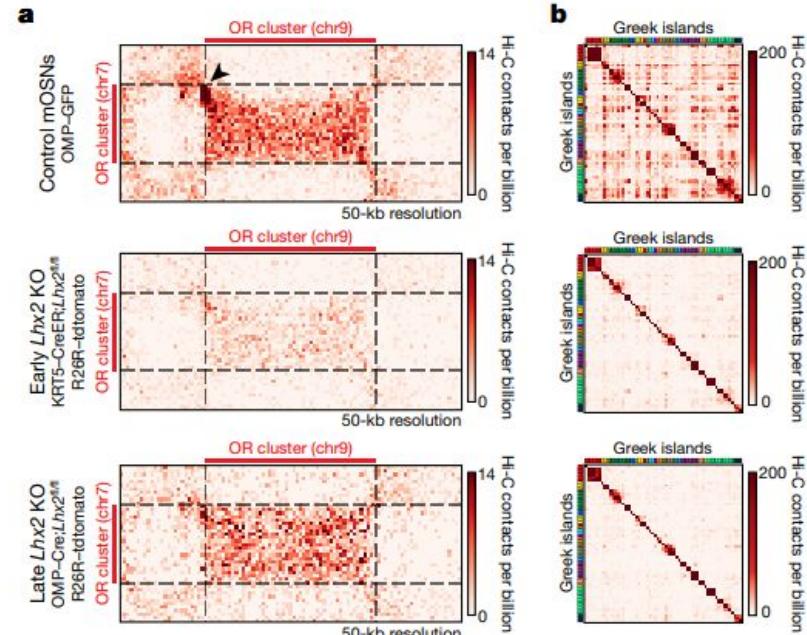
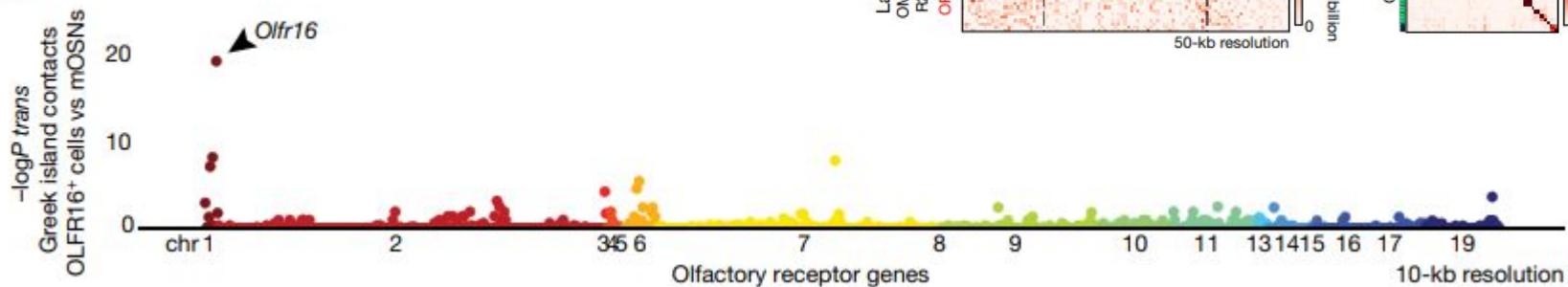
OR gene choice is controlled by multiple enhancers scattered across genome



# Greek islands and ORs

Greek islands are bound by LHX2 and LDB1

Greek islands only interact with active OR genes



# Questions?