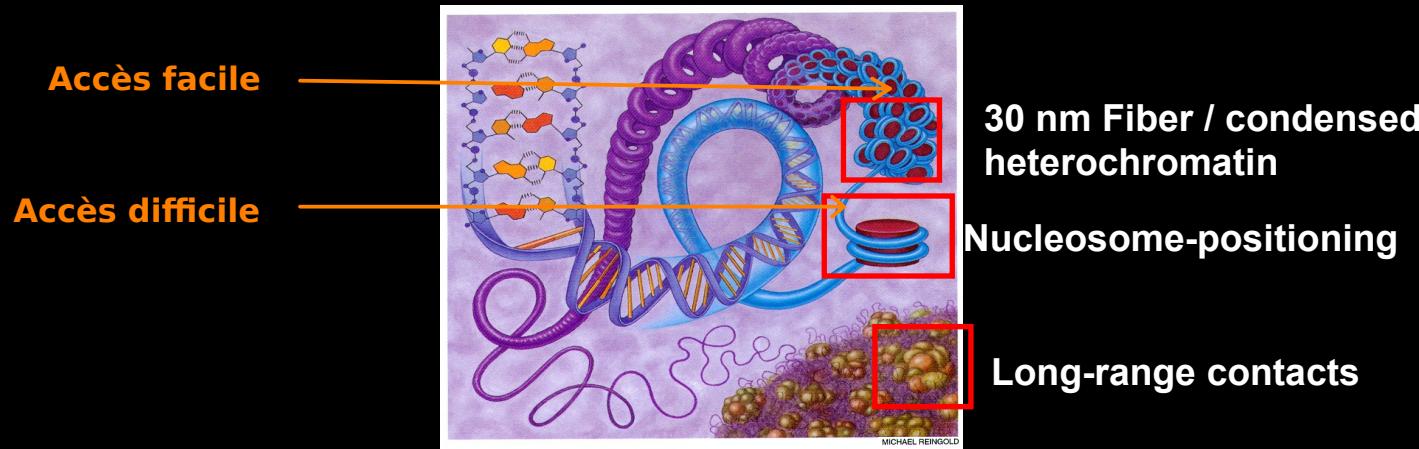


I. ChIP & Co.
The Technique - Biological Context

CNRS - Univ. of Toulouse
Chromatin Dynamics & Cell proliferation

Olivier Cuvier
Inserm

Chromatin Organization in eukaryotes: Nucleosome-positioning, condensation & Long-range interactions



Vogelmann, Valeri, Guillou, Cuvier* and Nollmann*, Nucleus, 2011

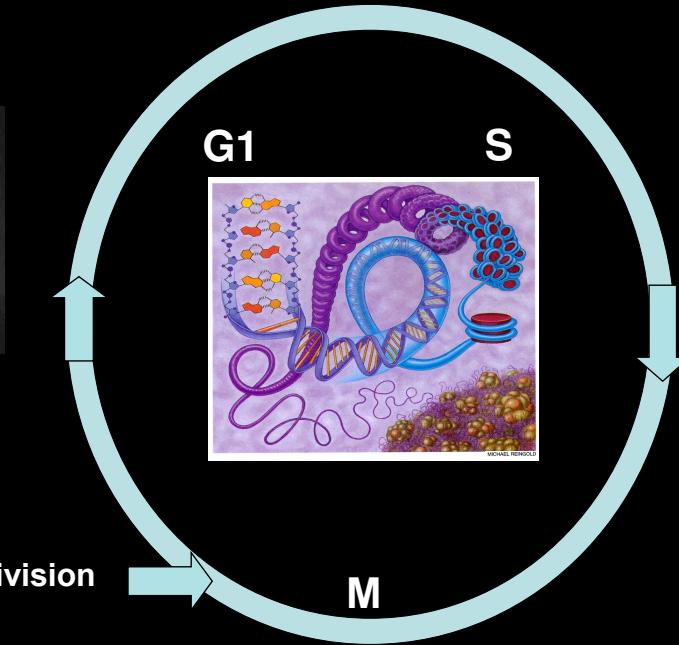
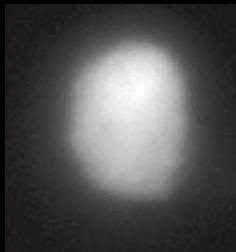
Nucléosome : - unité structurale de base
- en général rend inaccessible la zone où il est
- depuis 2006, se fixe plus sur GC que AT, voir article "A genomic code for nucleosome positioning"

30 nm Fieber : - accessibilité aux gènes difficile

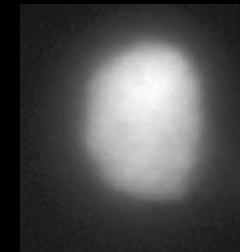
Long range contacts : - idée de territoire chrmosomique : zone condensée
- zone décondensé accessible

Major functions and chromosome dynamics through the Cell Cycle

Transcription

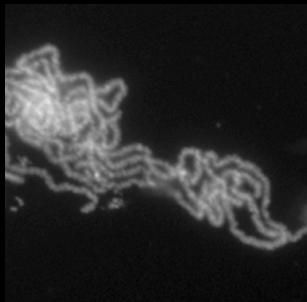


Replication



division

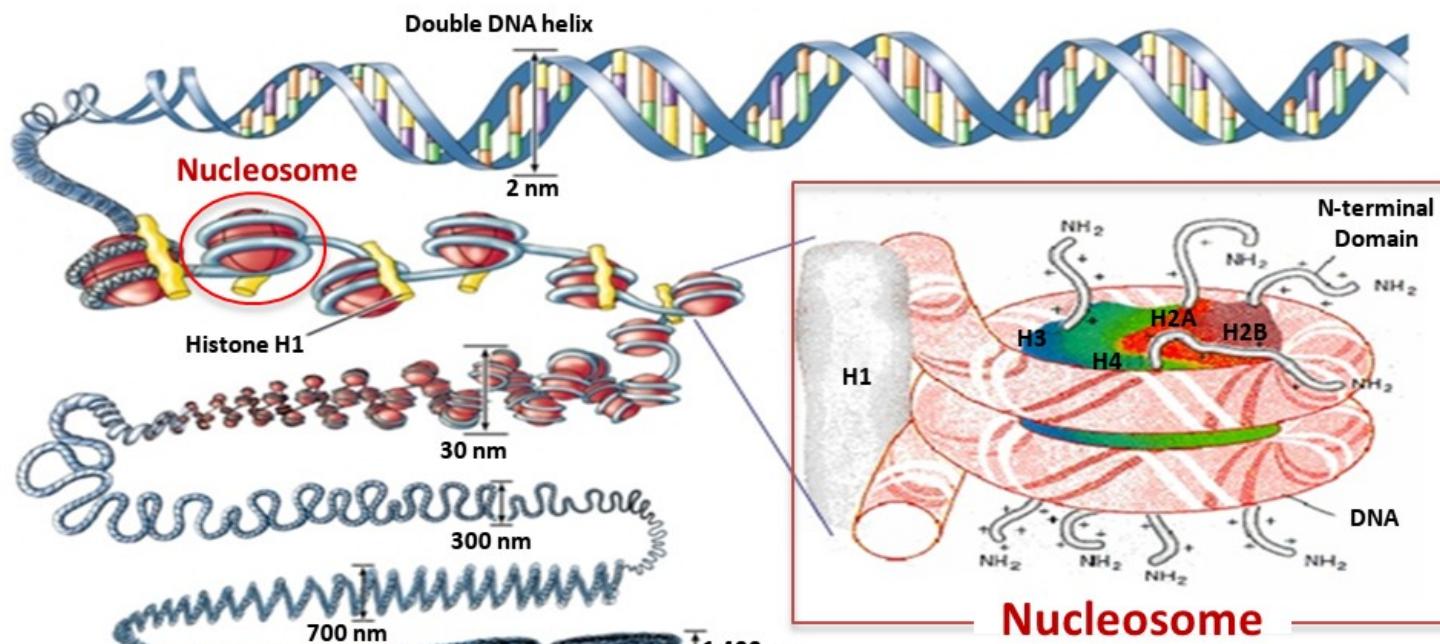
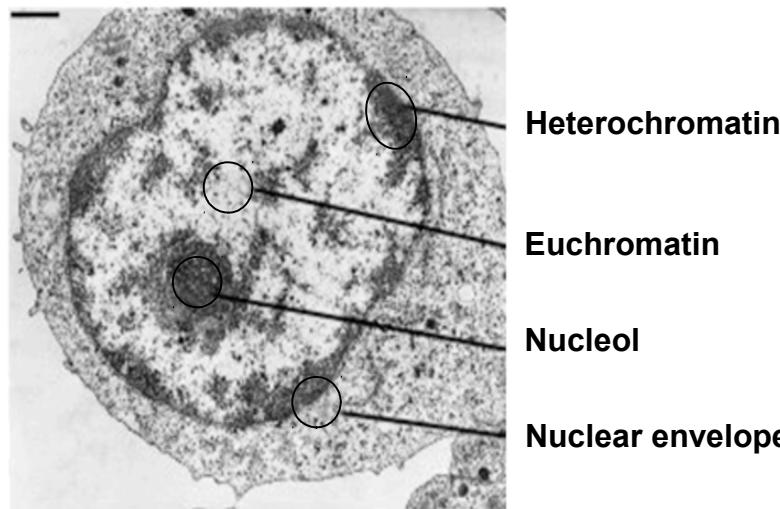
Condensation



**Compacte les chromosomes pour les répartir par zone
Permet de faire la ségrégation et d'éviter les cassures**

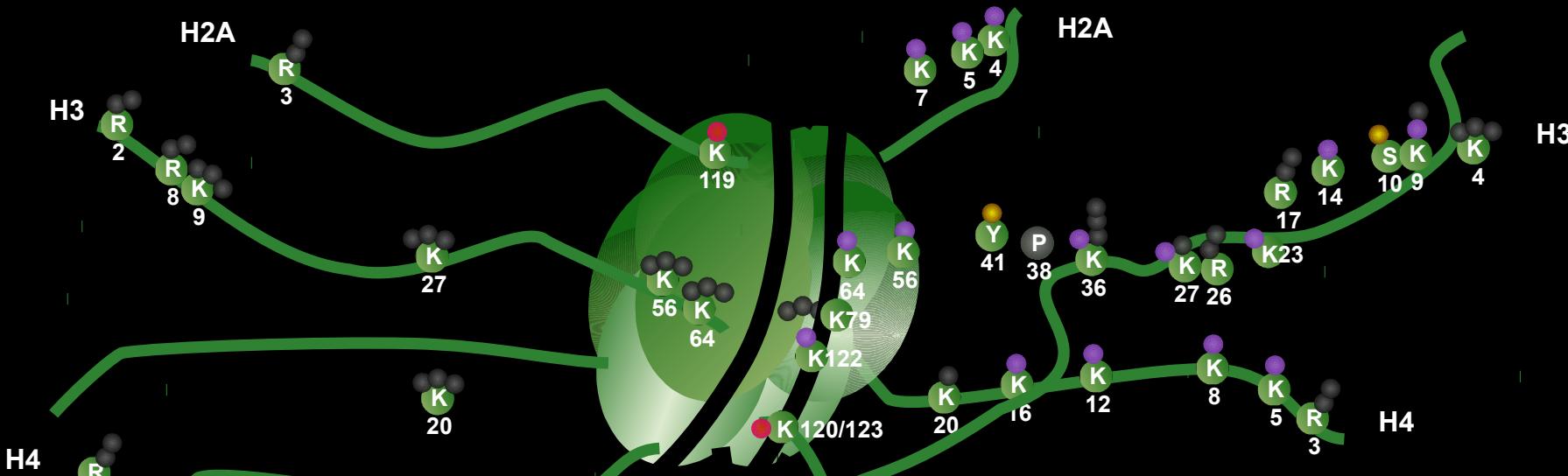
Condensation of Chromatin

=> heterochromatin or euchromatin



Epigenetic mechanisms & Post-translational modifications of Histones

Heterochromatin



H2B

Enzymes

HMT
HAT
Ubiquitin-ligase
Kinase

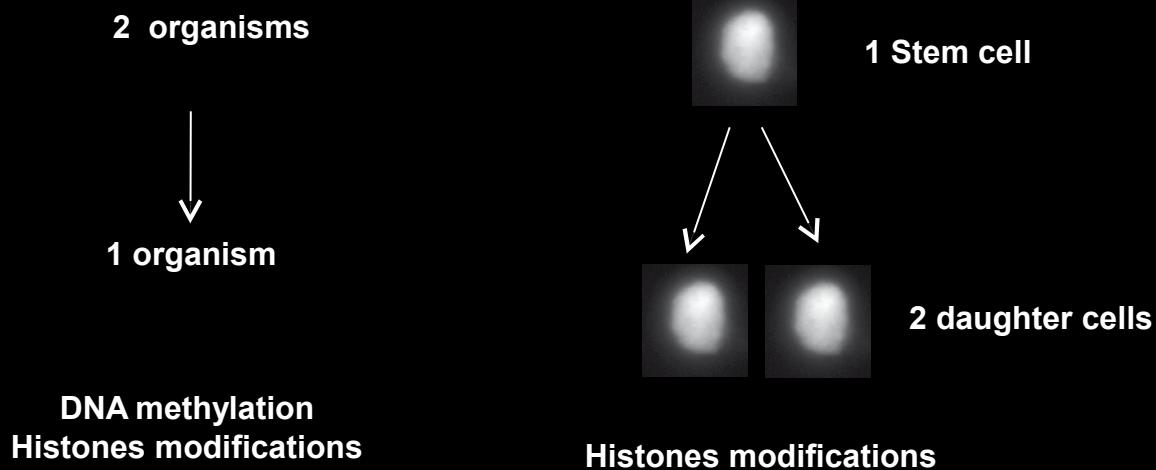
Enzymes

- Me HDM
- Ac HDAC
- Ub Deubiquitinase
- P Phosphatase
- Iso

ADM

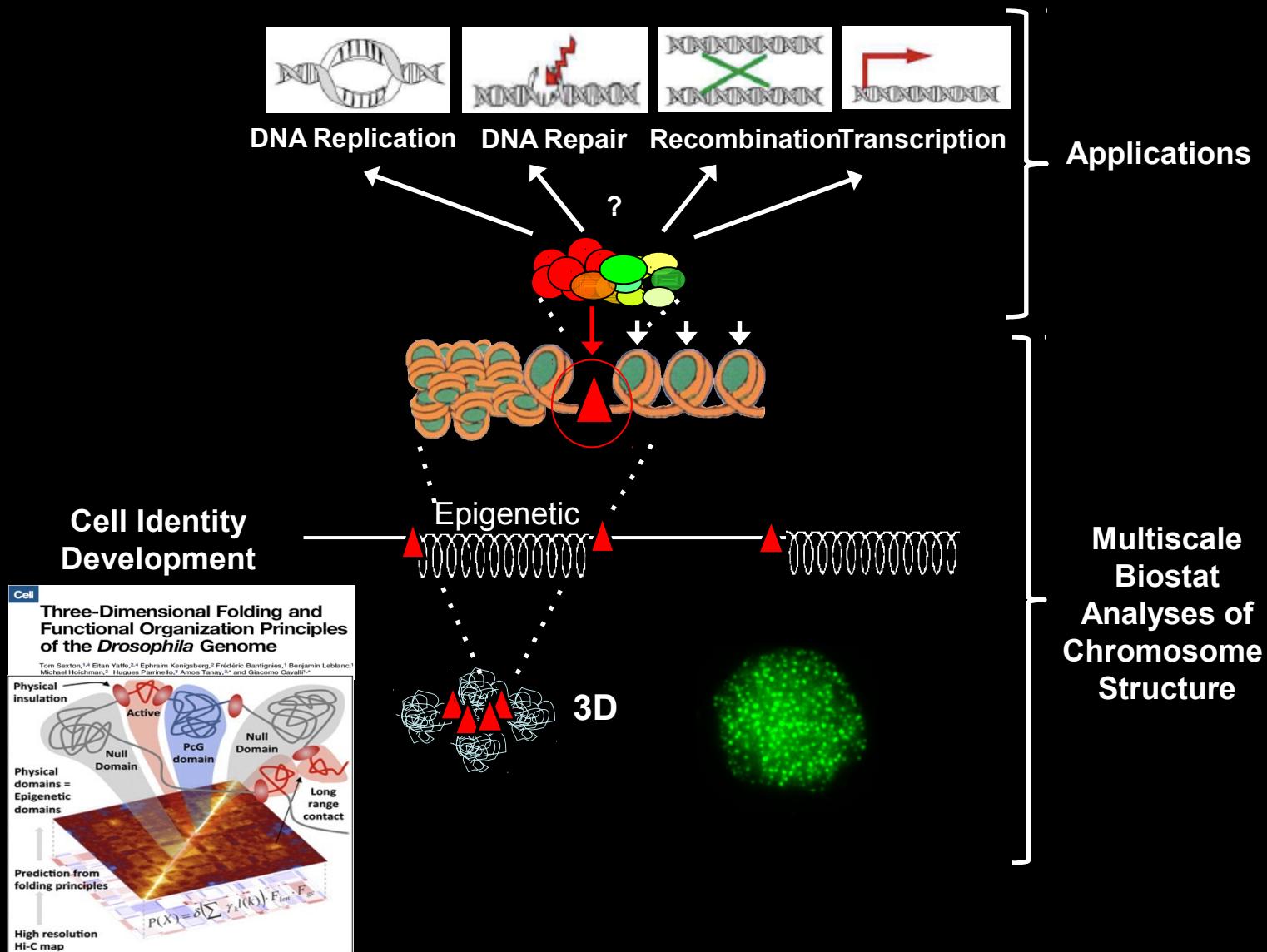
H2

Epigenetic mechanisms



Biological Context

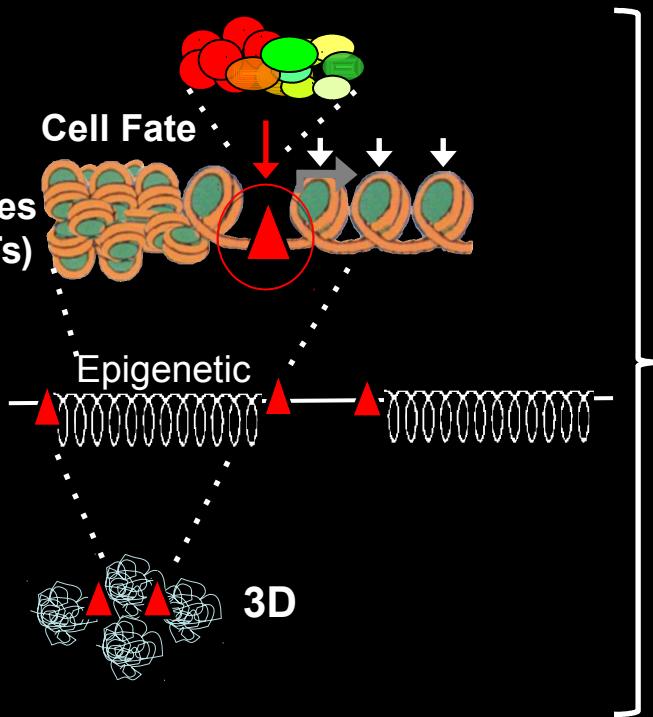
CBI - IBCG
LBME - LMGM- LBCMCP – CBD



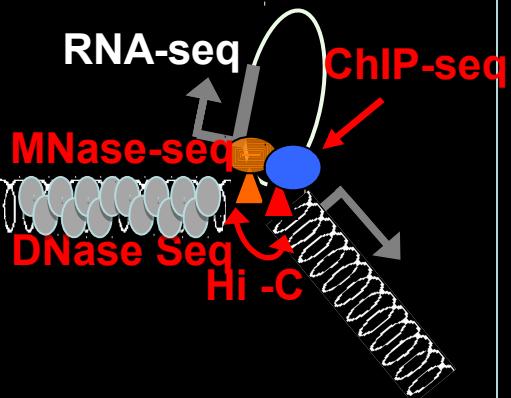
Floyd et al., Nature, 2013
 Dixon et al., Nature, 2012
 Sexton et al., Cell 2012
 Negre et al., Nature 2011

DNA recruitment modules Histone modifiers (HMTs)

Cancers
 Cell
 (re)programming
 Development



Multiscale Data mining



Lhoumaud, P., Hennion, M., Cuddapah, S., Morillon, P., Gamot, A., Queille, S., Lhoumaud, P., Severac, D., Dantec, C., Zhao, K., Embery, E., and O. Cuvier. Regulation of Nucleosome-dynamics by chromatin insulators and the HMT Mes4/NSD. **EMBO J.**, under revision

Jun Liang*, Laurent Lacroix*, Adrien Gamot*, Suresh Cuddapah*, Sophie Queille*, Priscillia Lhoumaud*, Pascal Martin*, Pierre Lepetit*, Saeed Saberi*, Franck Court*, Magali Hennion*, Serge Urbach*, Jutta Vogelmann*, Olivier Bouchez*, Marcelo Nöllmann*, Keji Zhao*, Eldon Embery* and O. Cuvier*[†] Chromatin Immunoprecipitation - Indirect peaks Highlight Functional Long-Range Interactions among Insulator Proteins and RNAPII Pausing Molecular Cell, under revision

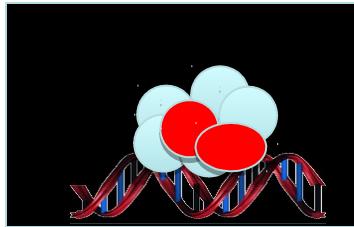
Chang, H., Cuvier, O. and J. Dekker (2009). Gene dates, parties, and galas: Chromatin dynamics and higher order organization. **EMBO rep.**, Vol 10, 689-693.

Jiang, W., Embery, E., O. Cuvier, Hart, CM (2009). Genome-wide mapping links BEAF binding sites to transcription. **MCB**, Vol 29, 3356-68.

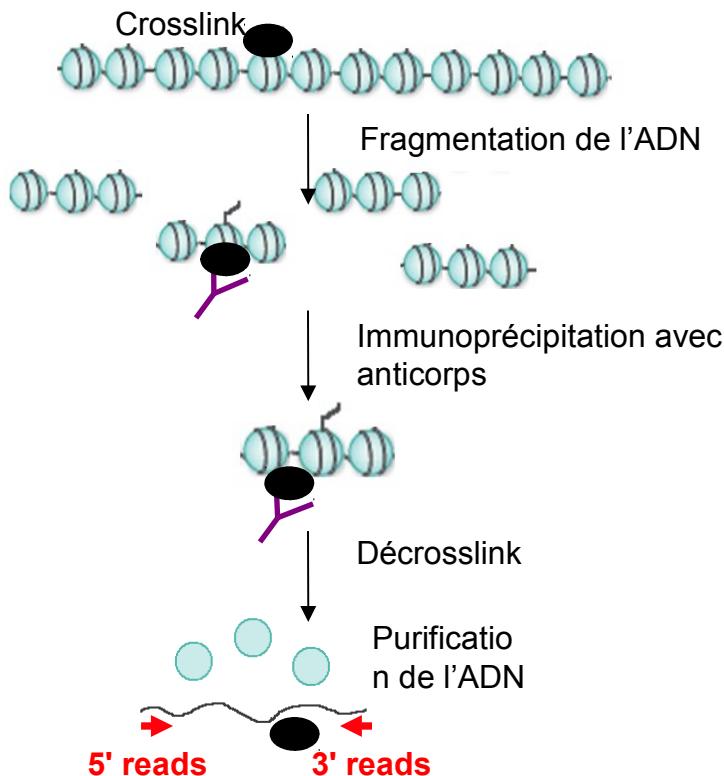
Eldon Embery,, Blattes, R., Schuttengruber, Kas, E., O. Cuvier (2009). BEAF regulates the Cell Cycle through controlling the deposition of H3K9me marks to nucleosome-associated dual-core elements. **PloS Biology**, Vol 6:2896-910

ChIP = ?

ChIP : requirements, conditions



ChIP (seq)



Specificity: Biology / Bioinfo

1) antibodies

- a) how to generate a good antibody ?
 - native protein conformation (3D), epitopes => rProt versus pept

b) specificity:

- affinity-purified / crude serum: check / WB

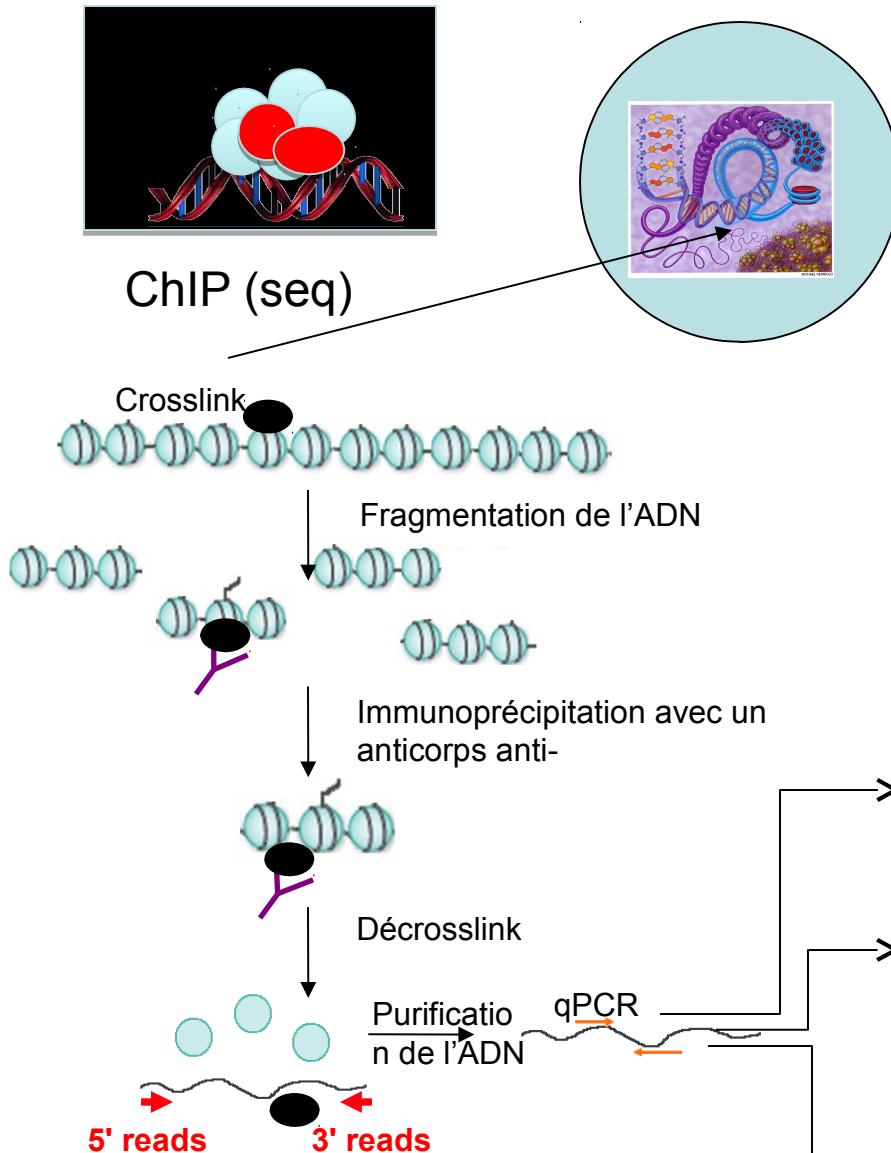
c) efficiency:

- test / immunoprecipitation exp

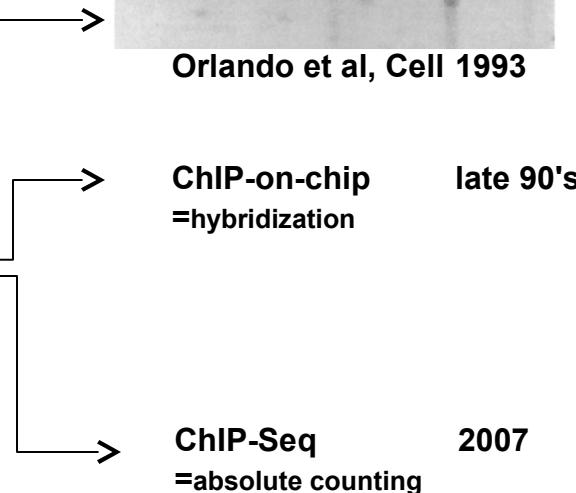
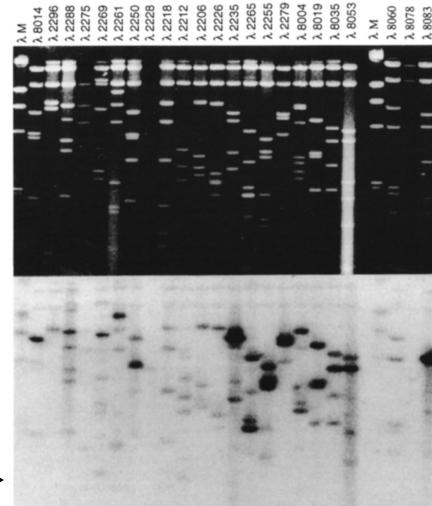
2) Variations: protocol, the experimentalist (sonication, washes etc) => input control, control antibody (IgG)

3)

ChIP : the origins of the signal

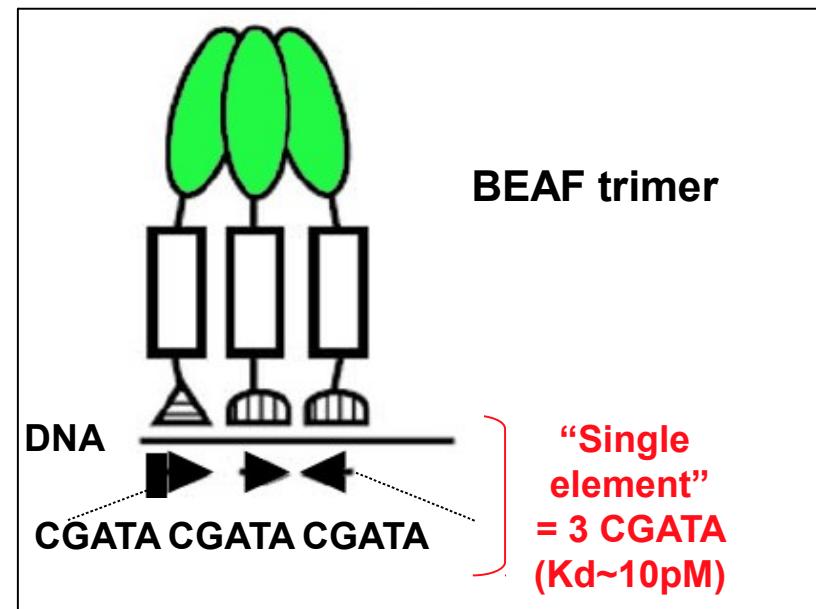
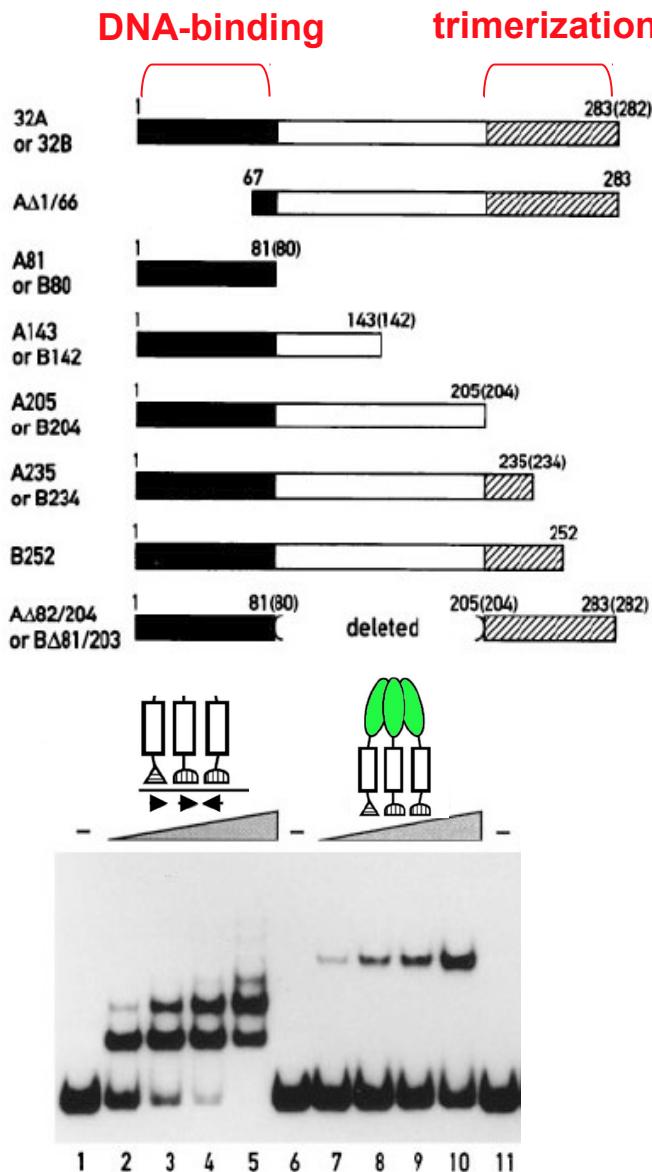


Living cell/organism $\times 10^3$ to 10^8



**Methodological context ?
Alternative techniques = ???**

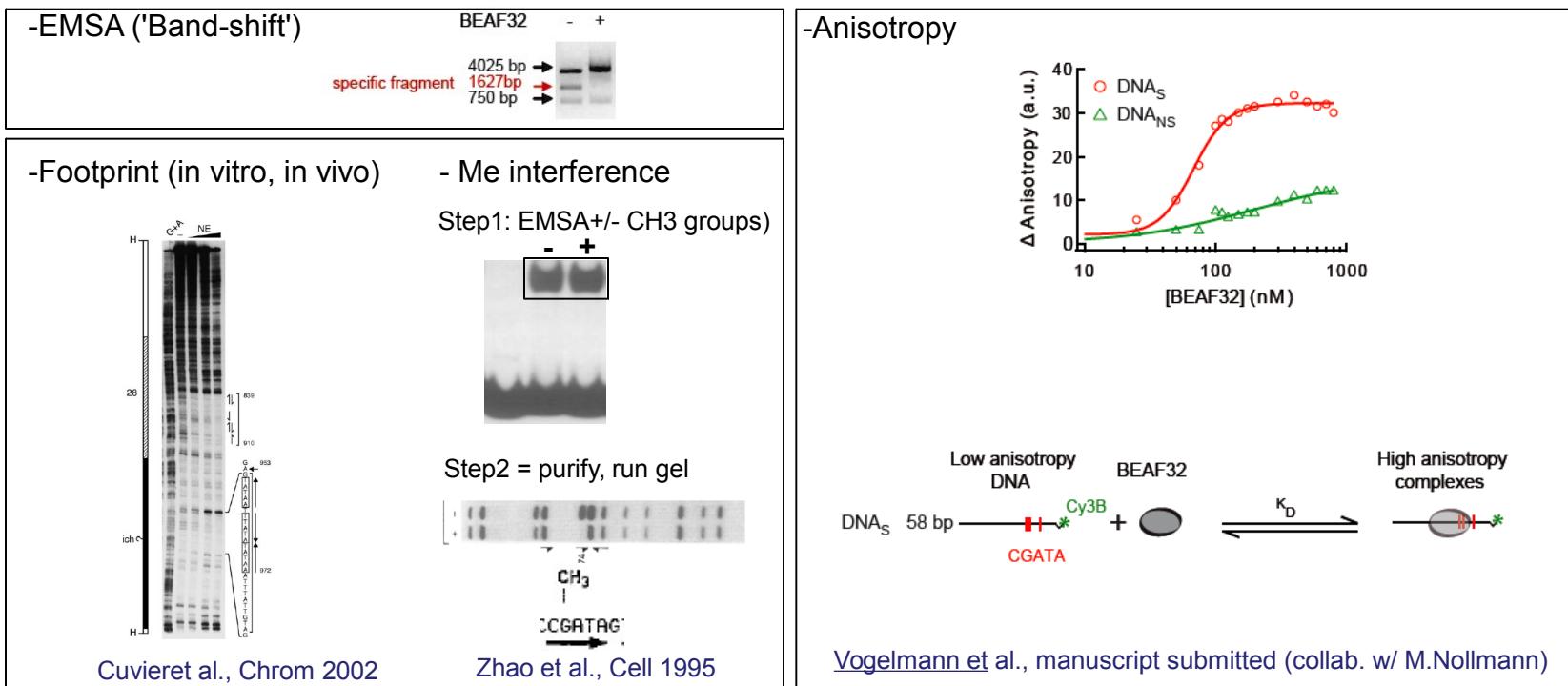
ChIP & other techniques / DNA-prot interaction studies



Cuvier et al., MCB, 1998 ; Hart et al., Chrom. 1999; Cuvier et al., Chrom., 2002

One BEAF subunit binds to one CGATA consensus motif

ChIP & other techniques / DNA-prot interaction studies



Advantages:

- *in vitro* allows, in general, to control conditions (salt, competitor, **factors** incubation time, buffers etc.)
- define a DNA consensus independently of any other parameter
- fast,

Limitations:

- binding does not mean it really binds
- in vivo*: DNA accessibility
- no binding does not mean it does not bind
- in vivo*...co-factors!

Why bother ? Advantages of ChIP ???

Advantages:

- in vivo*
- higher and higher resolution, precision etc.

Limitations:

- binding does not mean it really binds
- in vivo*: DNA accessibility
- no binding does not mean it does not bind
- in vivo*...co-factors!

*Those limitations may be exploited by direct comparisons
of predictions (*in vitro*) with measurements (*in vivo*)*

=

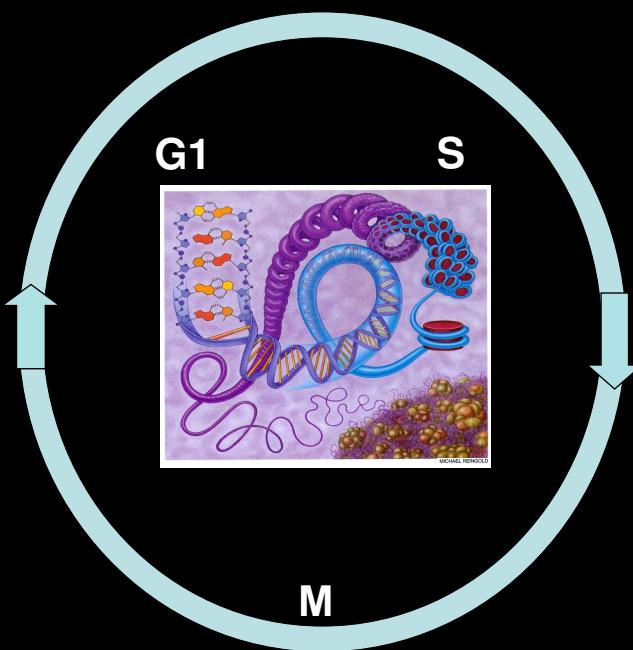
GOLDEN RULE / ChIP analyses

HETEROGENEITY

Spécificité !

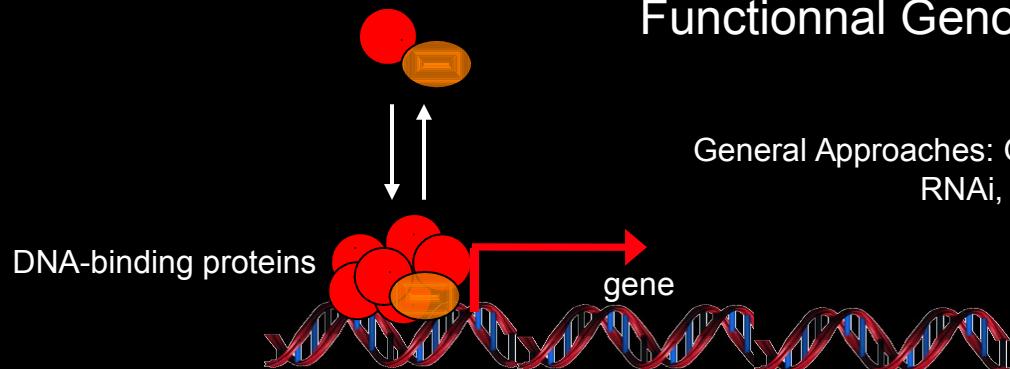
Transcription

Replication



Condensation

Technical context – Functionnal Genomics



General Approaches: Genetic Mutant, Depletions,
RNAi, shRNA etc.

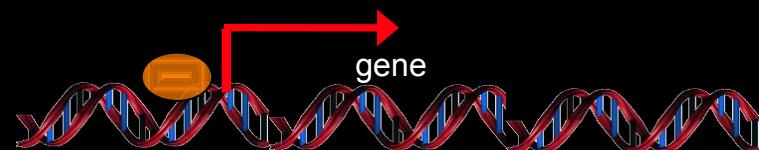
DNA-
protein
interaction

Single gene level:

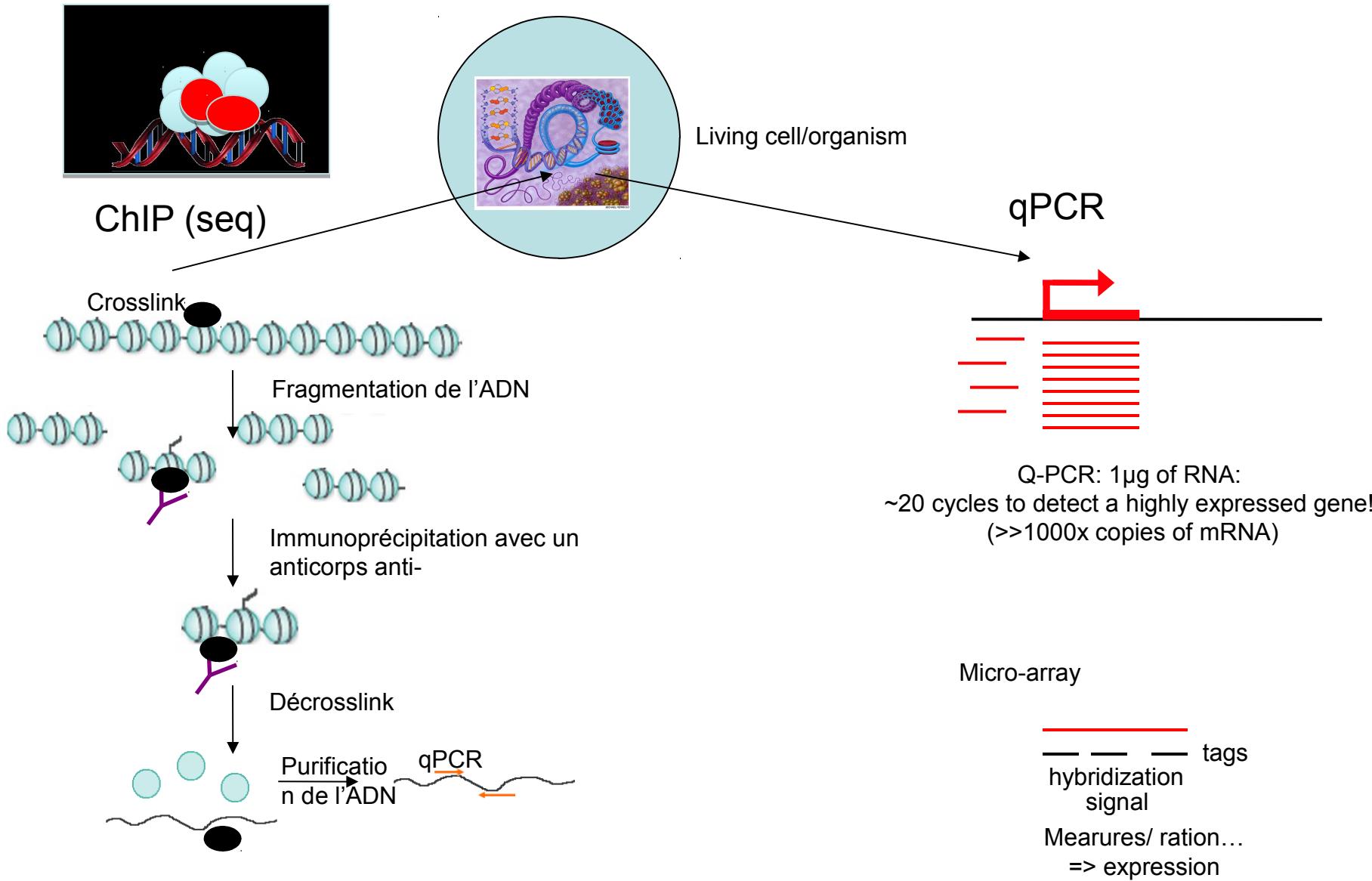
qPCR

Reporter Genes (transfections, transgenic animals etc.)
RTqPCR

Genomic context



The Basic of 'omics' : know what you're doing!



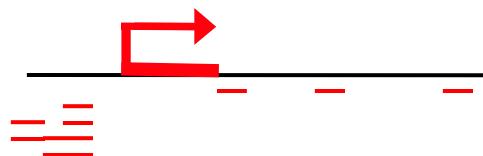
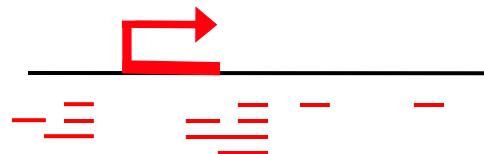
ChIP-seq

Quality controls

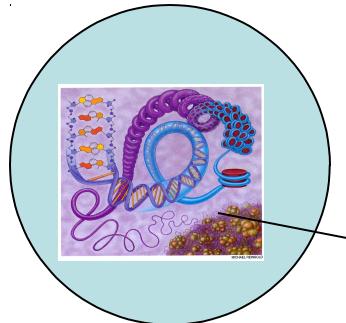


Antibodies: IgG !!!
-> pull down sequences => bias
(sequences easier to access, etc.)
⇒ **Ration of enrichment**

⇒ = **read density**

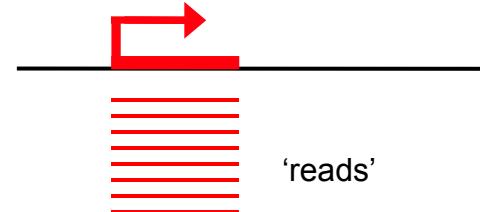


The Basic of 'omics' : know what you're doing!



Living cell/organism

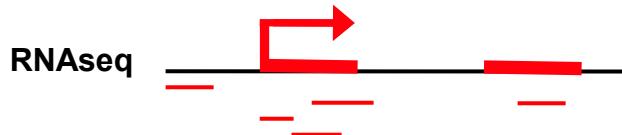
RNA (seq)



Bioinfo pipelines / tools:

- parsing (bwa, TopHat...)
- counting (averages, boxplots...)
- stats (fisher, wilcoxon, kolmogorov...)

Q-PCR: 1µg of RNA:
~20 cycles to detect a highly expressed gene!
(>>1000x copies of mRNA)



Resolution:

Mapping: exons, intergenic RNAs etc.

Micro-array

cDNA _____
— — — tags
hybridization signal
Measures/ ration...
=> expression

« Omics »

1. Généralités

Biological & Technical context

2. Techniques

~~quick and dirty~~

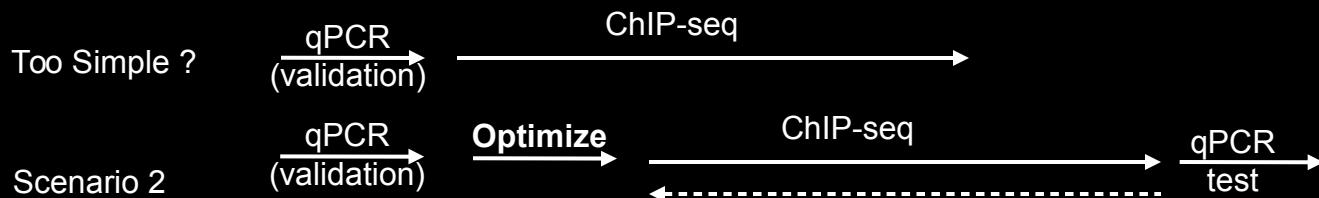
3. Applications

= lire, interpréter, imaginer, modéliser,
« moléculariser », réfléchir,
tester!!! = ur job

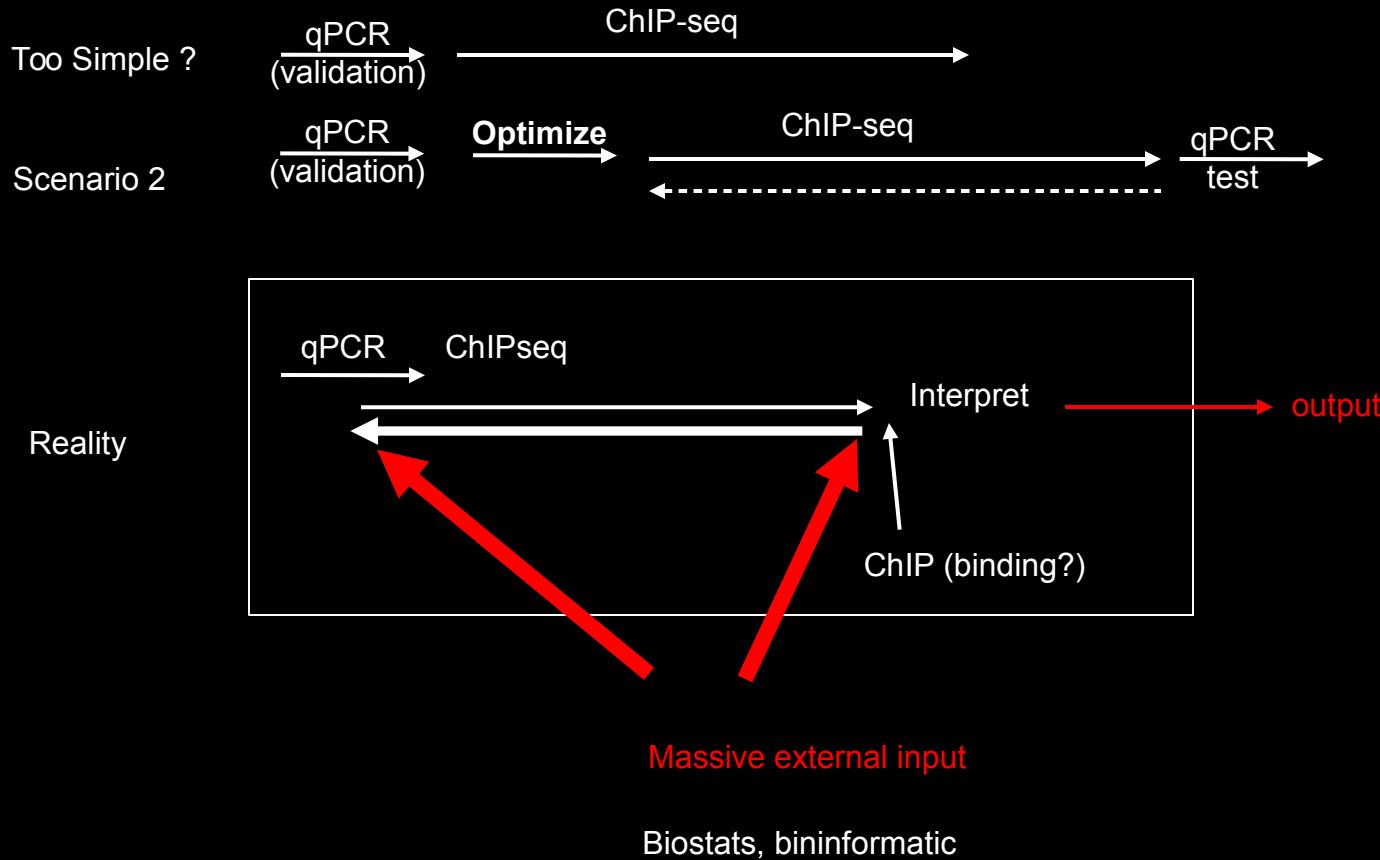
Submit your samples and get a good sleep ?



1. What are the limiting factors ???



1. What are the limiting factors ???



10 Gigabytes (10000 MegaB of text...)



'reads'

What are we going to do w/ all that data ????

A photograph of a man with a beard and short brown hair, wearing a light blue t-shirt and jeans, crouching in a grassy field. He is positioned next to a light-colored donkey. The background shows a green lawn and a fence in the distance.

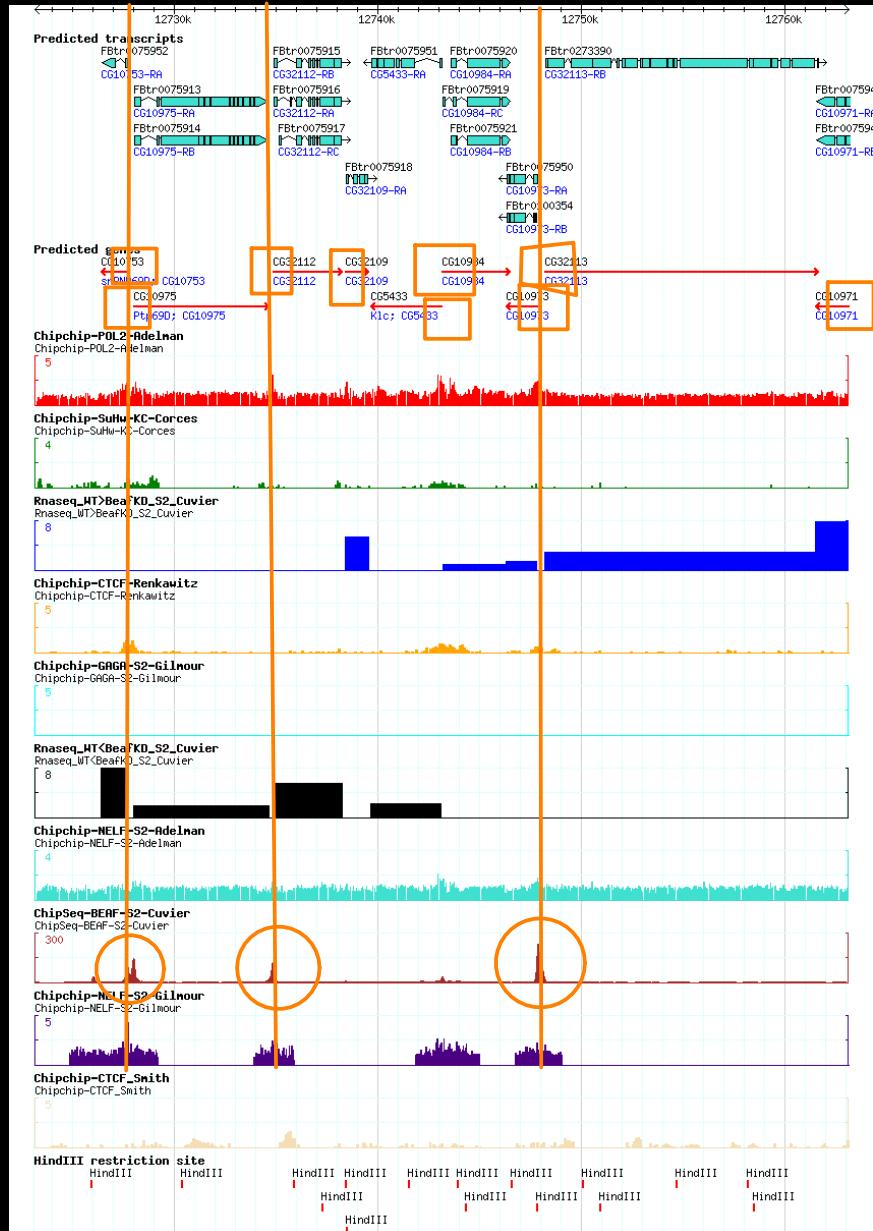
Genome browser

[http://insulators_chromosome-dynamics.biotoul.fr
/cgi-bin/gb2/gbrowse/fly5/](http://insulators_chromosome-dynamics.biotoul.fr/cgi-bin/gb2/gbrowse/fly5/)

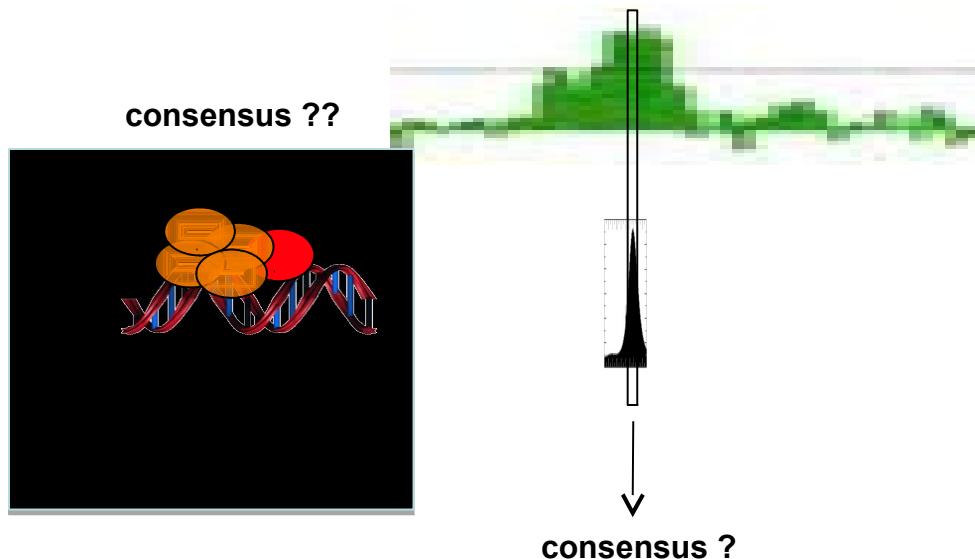
Examples:

- modEncode
-Flybase
- your personnal browser ??

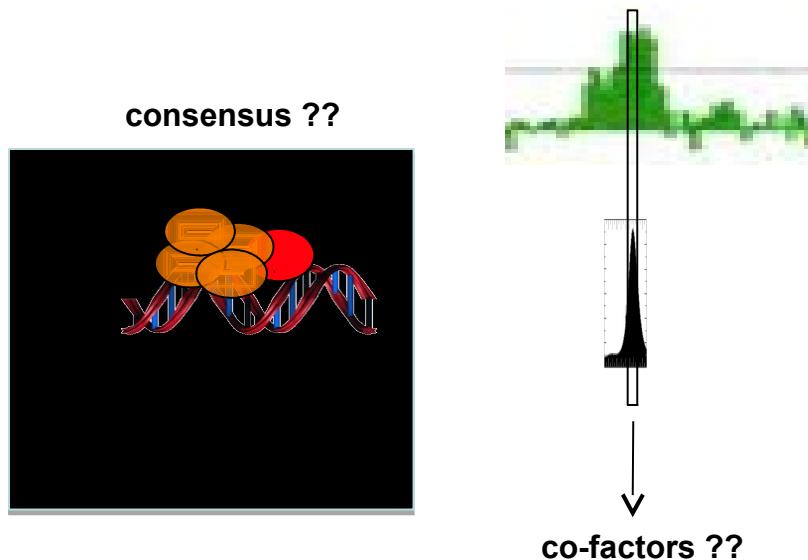
main idea = visualize !



MEME/bioinfo DNA-binding motif search



MEME/bioinfo DNA-binding motif search



Motif	Binding-Factor	E-value	
CGATA	BEAF site	1.2e-60	
tATCGATA	DRE	1.2e-60	
CAGTGTGACC	CP190 site	3.4e-59	
...			

« Omics »

1. Généralités

Biological & Technical context

2. Techniques

~~quick and dirty~~

3. Applications

= lire, interpréter, imaginer, modéliser,
« moléculariser », réfléchir,
tester!!! = ur job

4. Statistics...

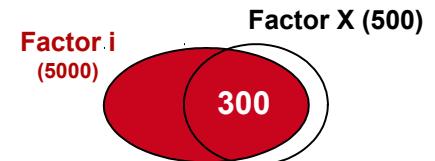
the basis for developing new tools

Binomial statistics:

In general, if the random variable K follows the binomial distribution with parameters n and p, we write $K \sim B(n, p)$. The probability of getting exactly k successes in n trials is given by the probability mass function:

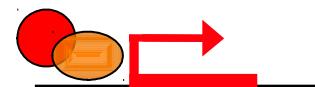
$$f(k; n, p) = \Pr(K = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$\binom{n}{k} = \frac{n!}{k!(n - k)!}$$



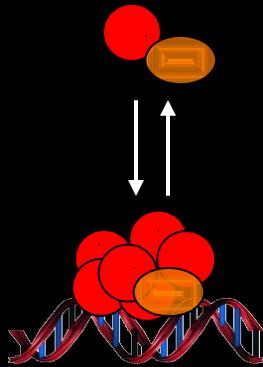
15,000 genes

$$\begin{aligned} p &= 5000/15000 = 1/3 \\ n &= 500 \\ k &= 300 \end{aligned}$$



Binomial distribution

Genomics – what (are the) limits ?



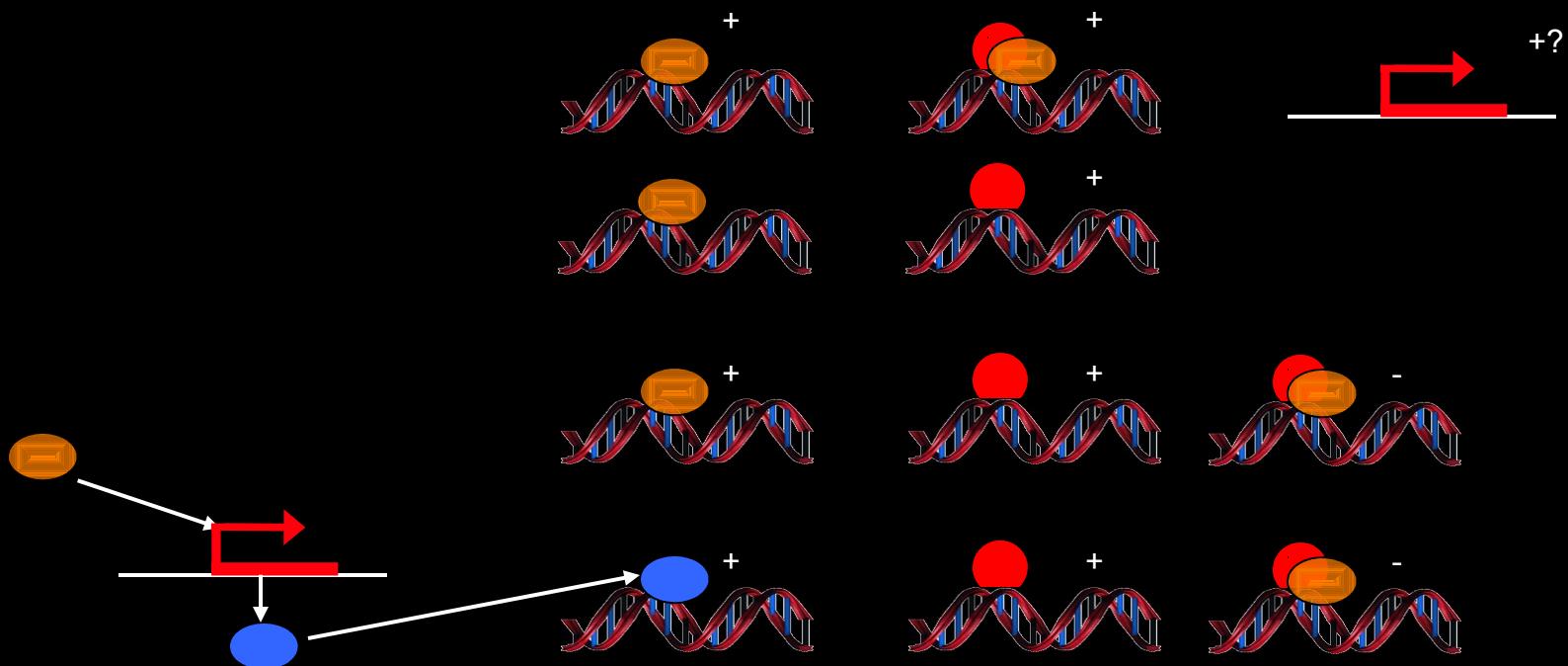
May be ok :

- Pricing + timing
- Statistics

May be NOT ok:

- Average / homogeneity
- time-resolution
- -> multiple tests PREDICT but no demonstration

-> Single molecule levels



« Omics »

1. Généralités

Biological & Technical context

2. Techniques

~~quick and dirty~~

3. Applications

= lire, interpréter, imaginer, modéliser,
« moléculariser », réfléchir,
tester!!! = ur job

4. Statistics...

NOT ur job (!) but you must
understand them

4. Development...

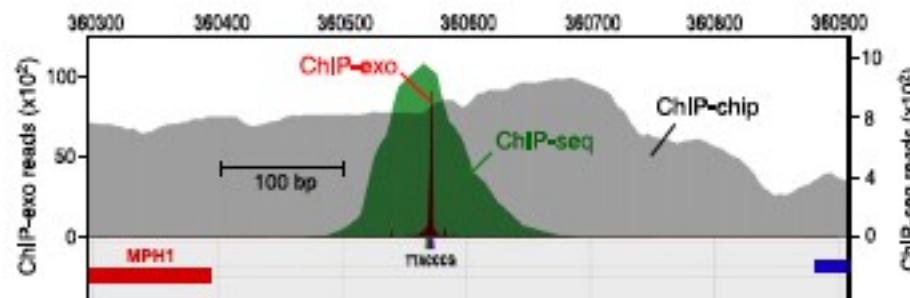
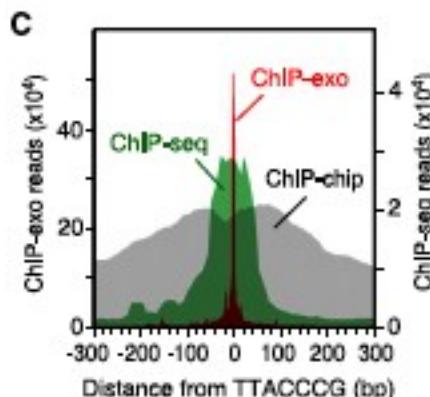
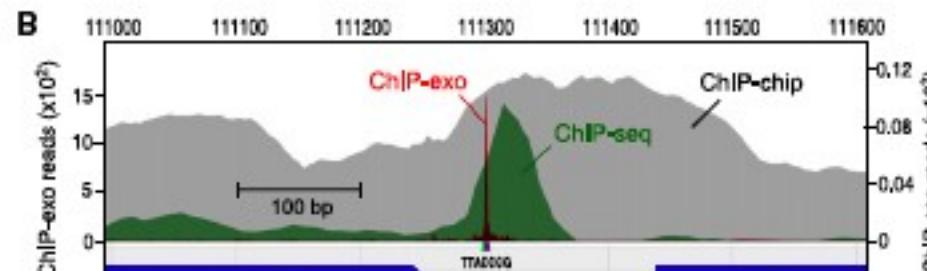
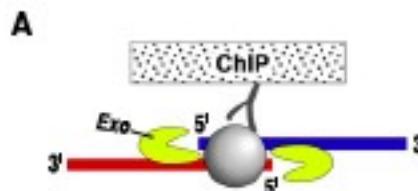
Comprehensive Genome-wide Protein-DNA Interactions Detected at Single-Nucleotide Resolution

Ho Sung Rhee¹ and B. Franklin Pugh^{1,*}

¹Center for Eukaryotic Gene Regulation, Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802, USA

*Correspondence: bfp2@psu.edu

DOI 10.1016/j.cell.2011.11.013

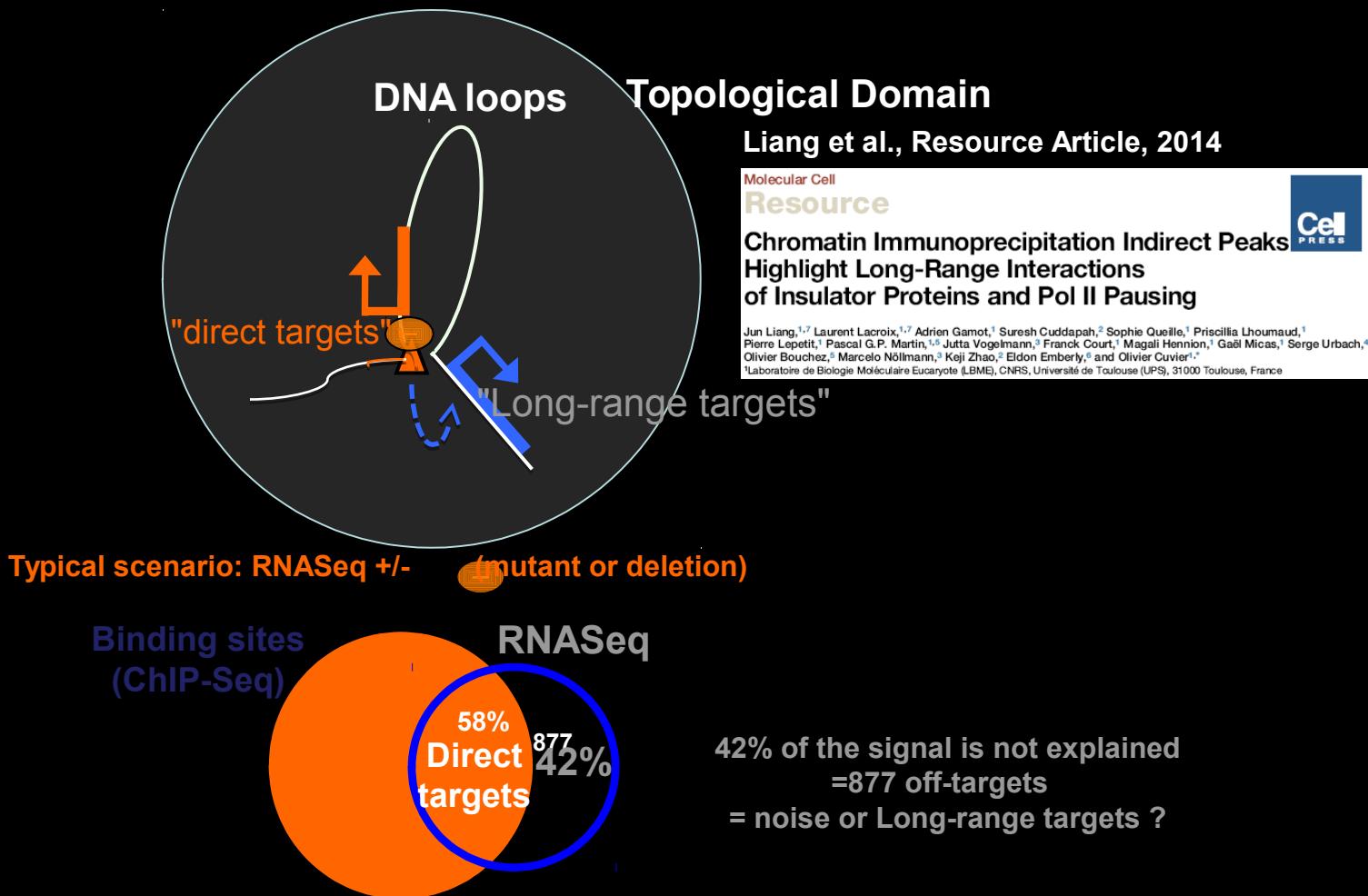


Chromatin Immunoprecipitation Indirect Peaks Highlight Long-Range Interactions of Insulator Proteins and Pol II Pausing

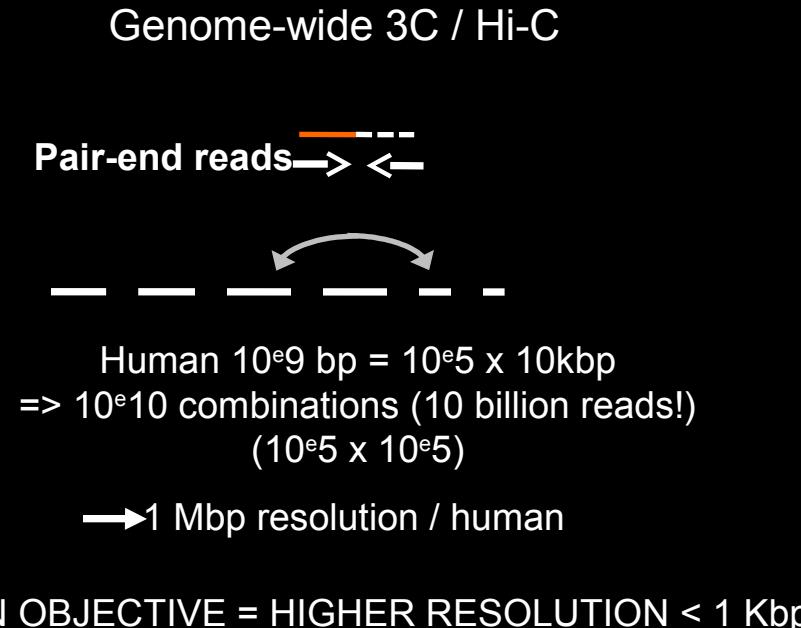
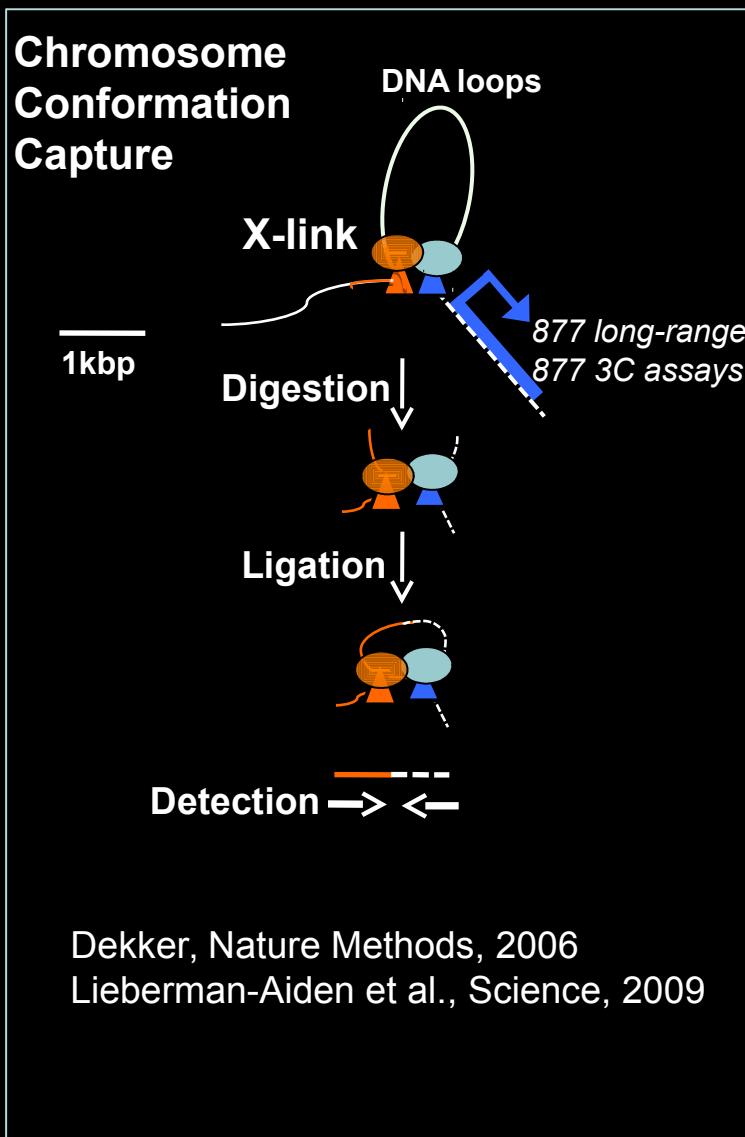
Jun Liang,^{1,7} Laurent Lacroix,^{1,7} Adrien Gamot,¹ Suresh Cuddapah,² Sophie Queille,¹ Priscillia Lhoumaud,¹ Pierre Lepetit,¹ Pascal G.P. Martin,^{1,5} Jutta Vogelmann,³ Franck Court,¹ Magali Hennion,¹ Gaël Micas,¹ Serge Urbach,⁴ Olivier Bouchez,⁵ Marcelo Nöllmann,³ Keji Zhao,² Eldon Emberly,⁶ and Olivier Cuvier^{1,*}

¹Laboratoire de Biologie Moléculaire Eucaryote (LBME), CNRS, Université de Toulouse (UPS), 31000 Toulouse, France

Transcriptional Regulation through Long-range contacts : Repression/activation: acting from a distance

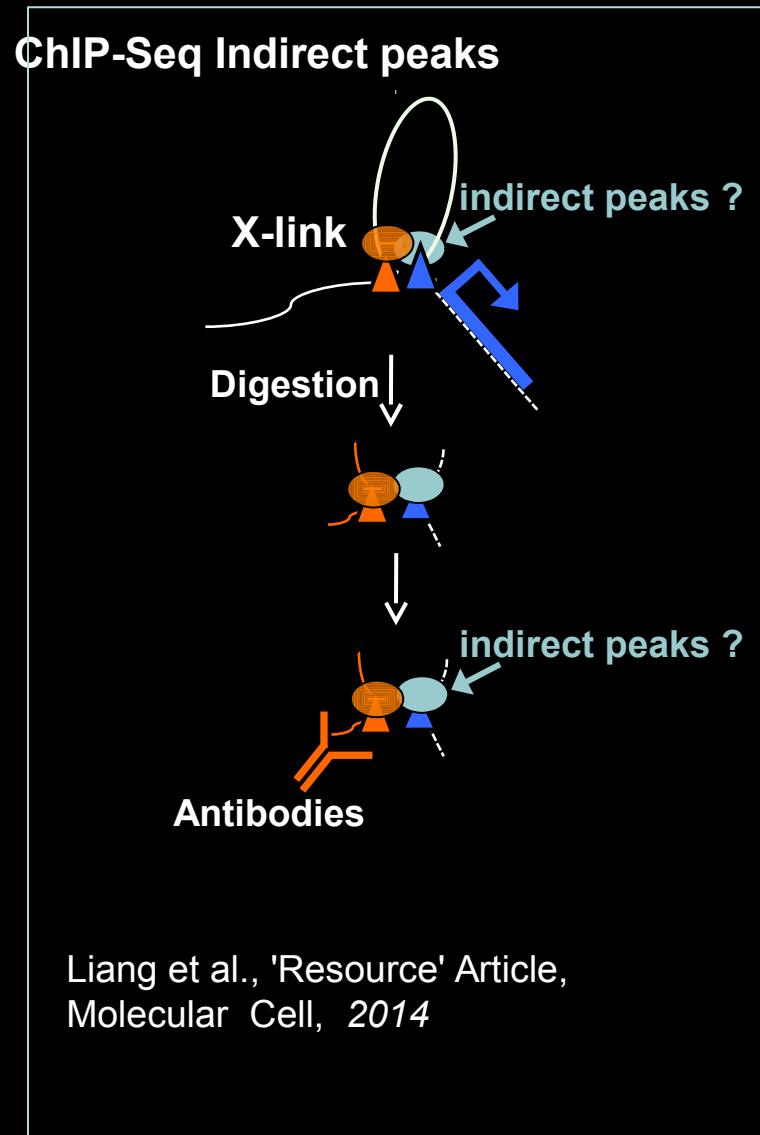
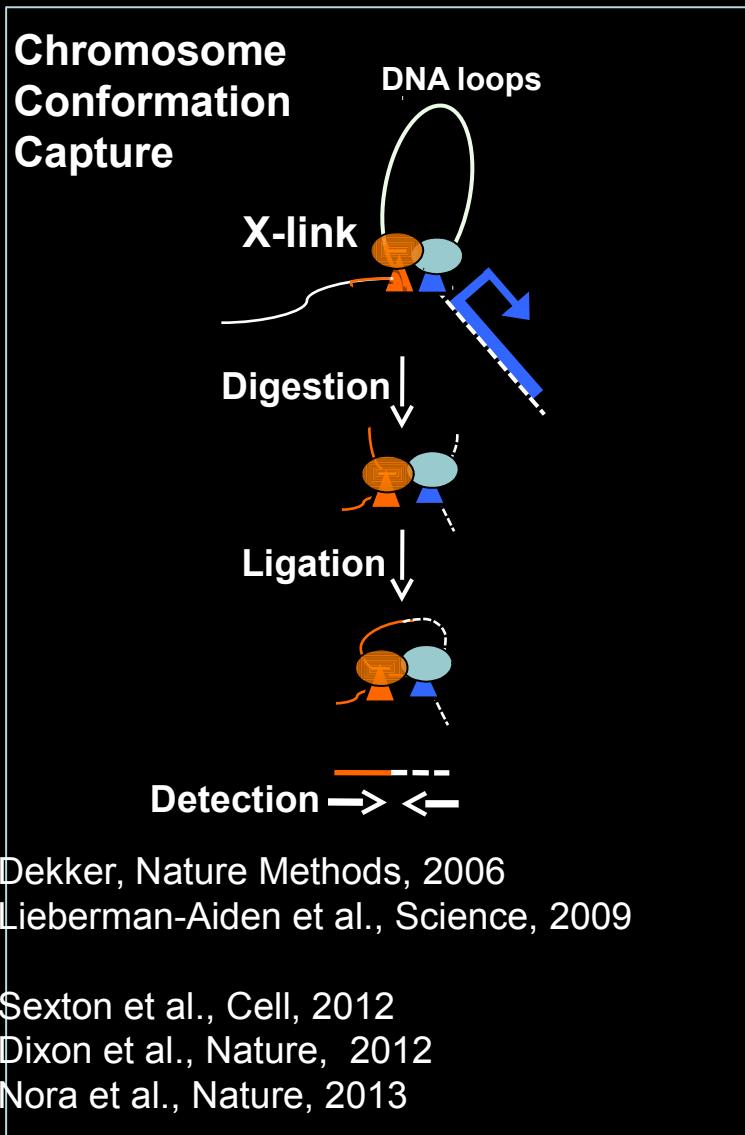


Detection of specific long-range interactions in chromatin 3C versus ChIP-Seq

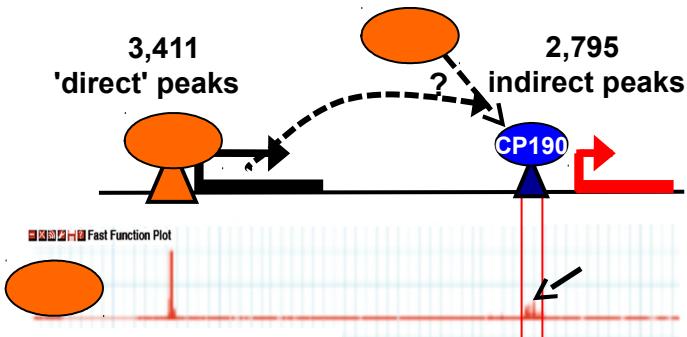


Dixon et al., Nature, 2012
Nora et al., Nature, 2012
Jin et al., Nature, 2013
Phillips-Cremins, Cell, 2013
Nora et al., Nature, 2013
Jin et al., Nature, 2013

Detection of specific long-range interactions in chromatin 3C versus ChIP-Seq

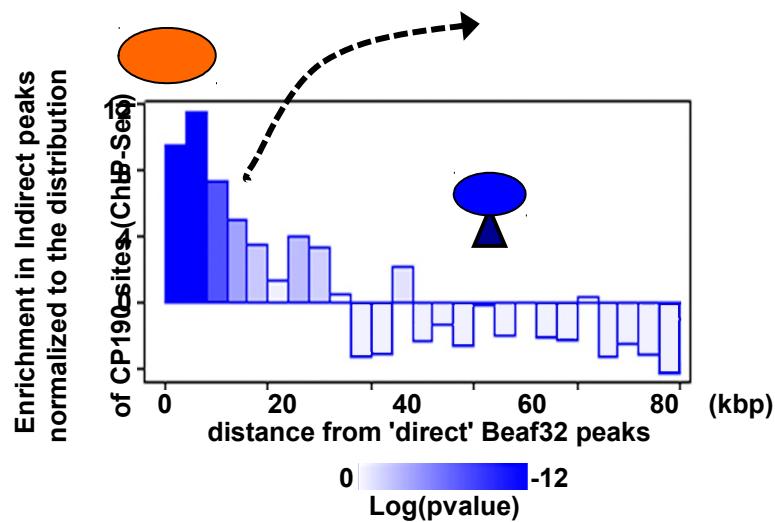


Biostatiscal analyses ChIP-Seq data

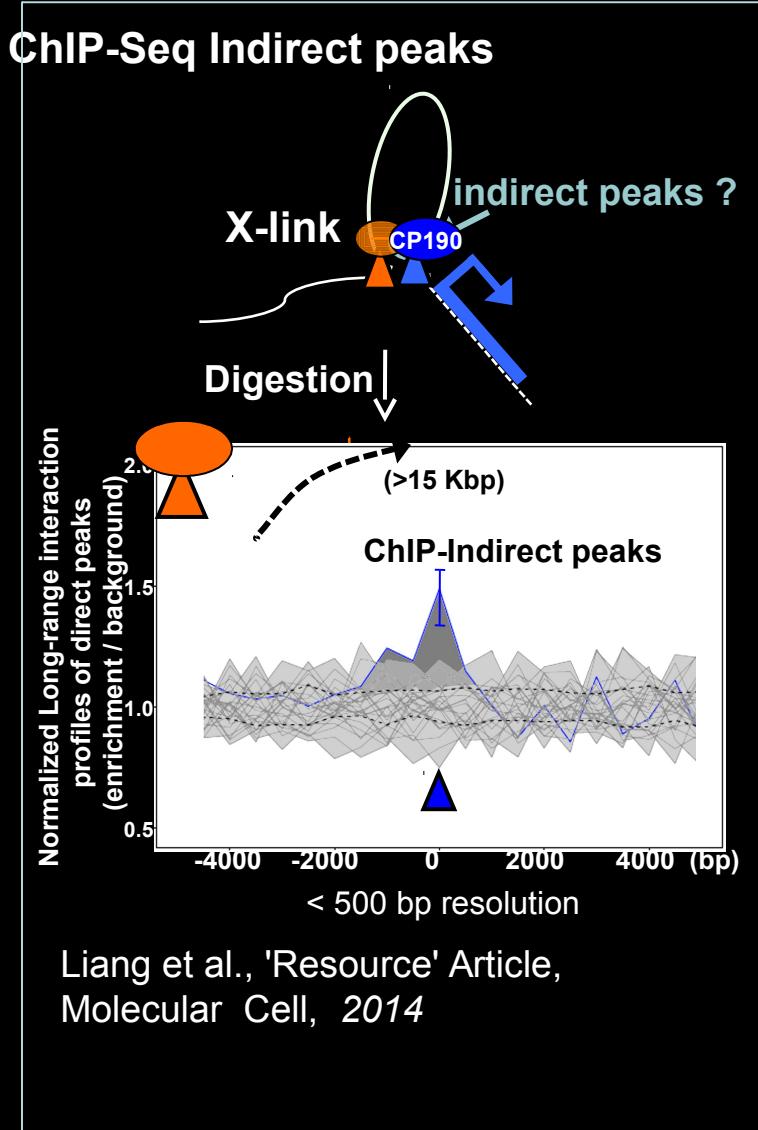
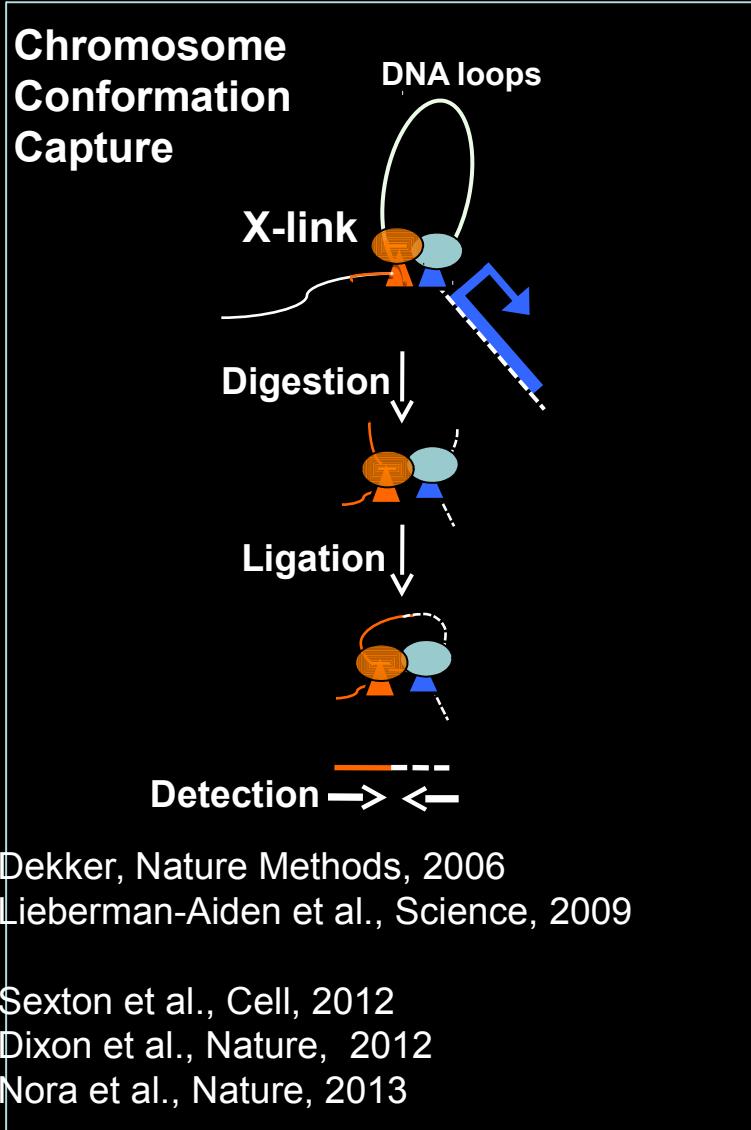


MoTif
Expected binding
to DNA

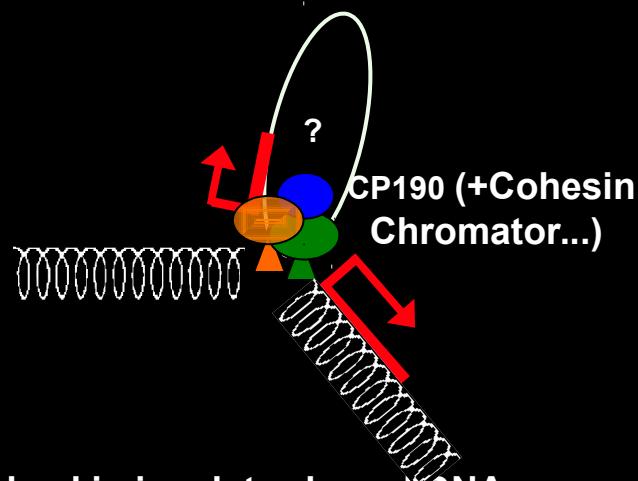
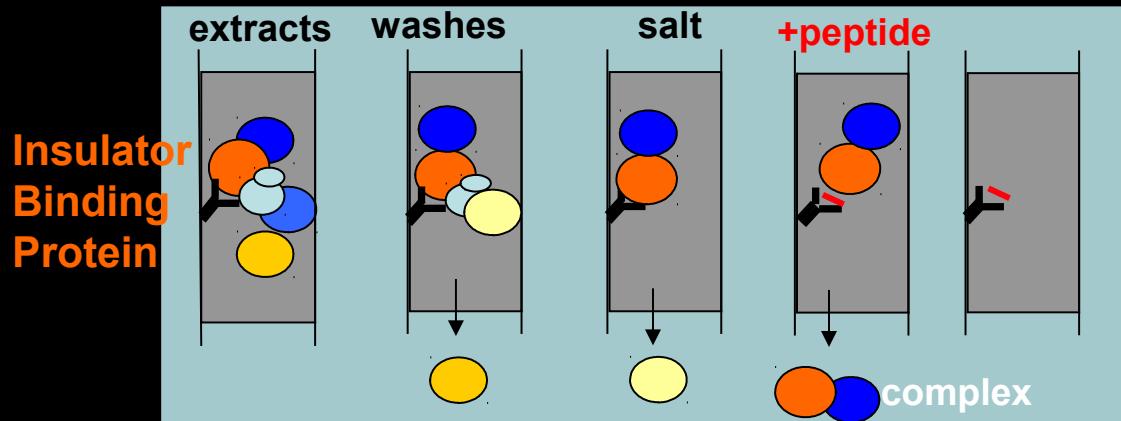
CGATA
no binding
expected



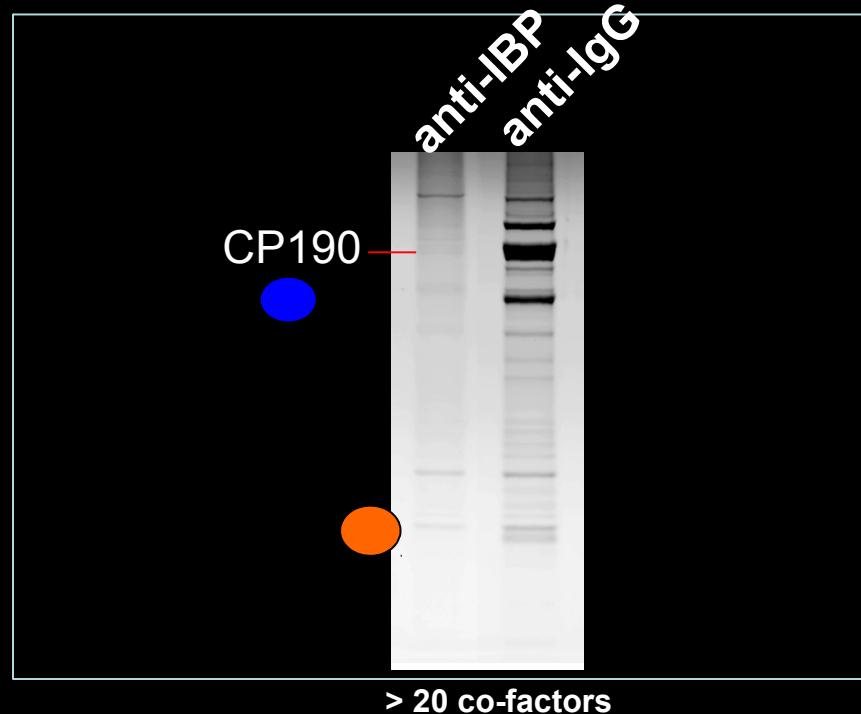
Detection of specific long-range interactions in chromatin 3C versus ChIP-Seq



Affinity Purification of Insulator Binding Protein (IBP) complexes



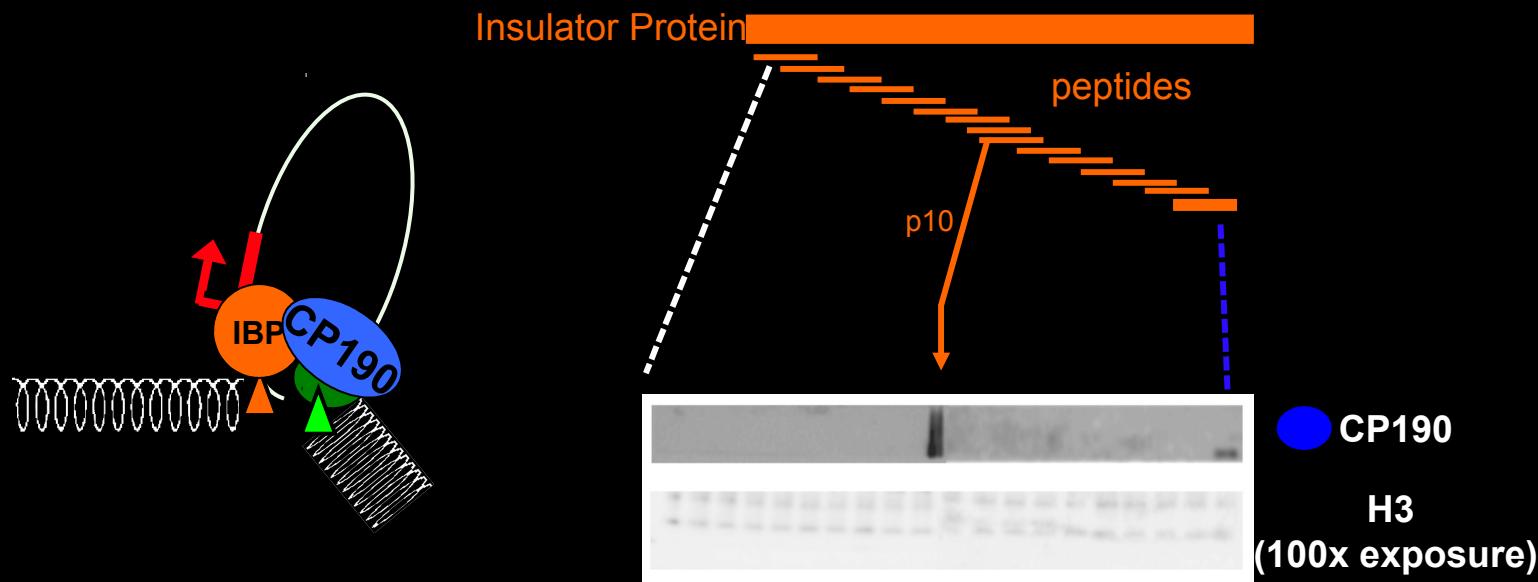
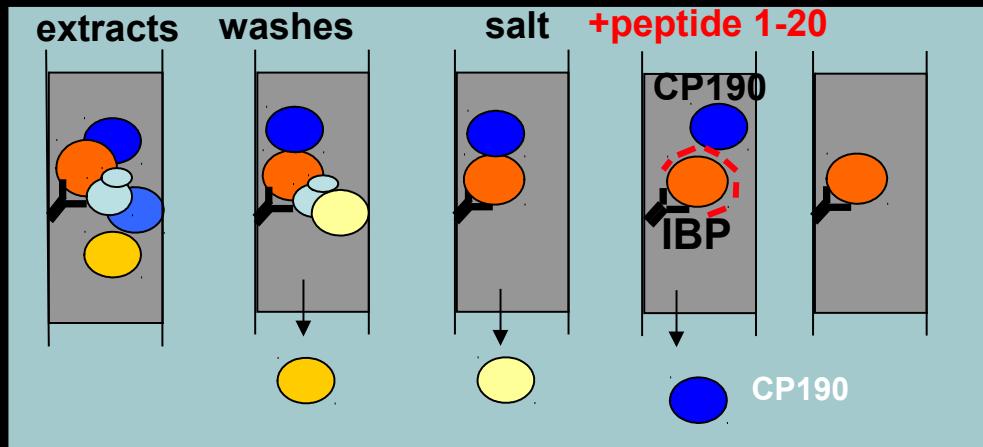
● = involved in insulator-based DNA looping ?



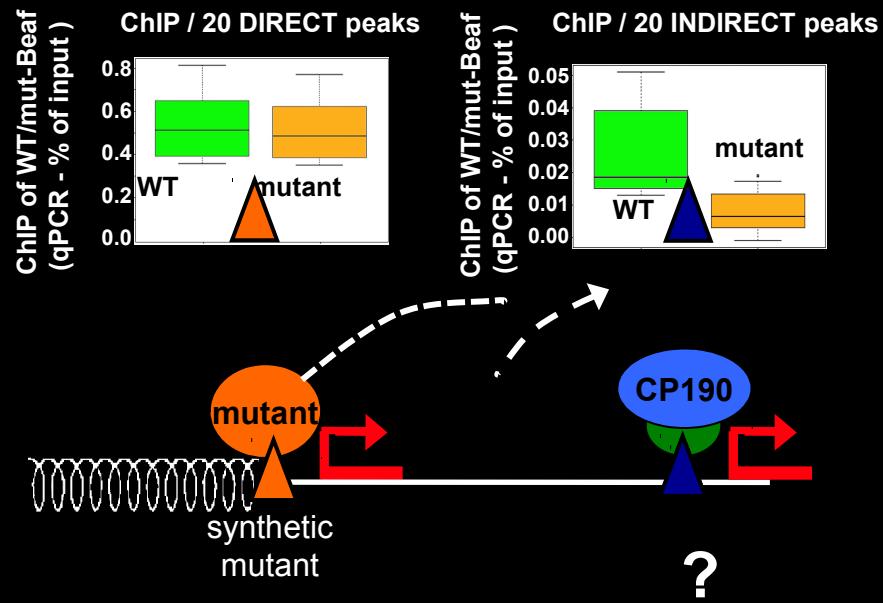
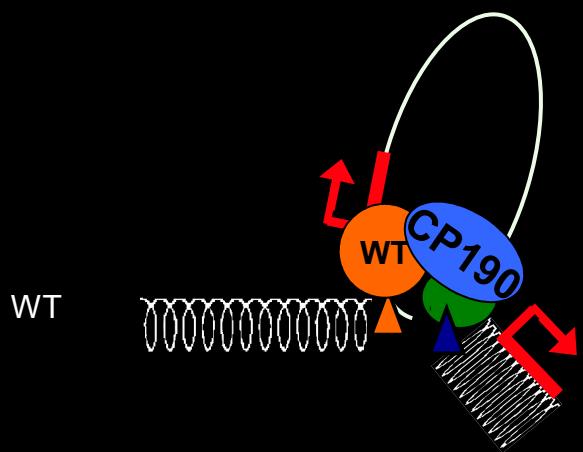
Hou et al., Mol Cell, 2012

Functionnal mutant of looping (a novel Biochemical technique)

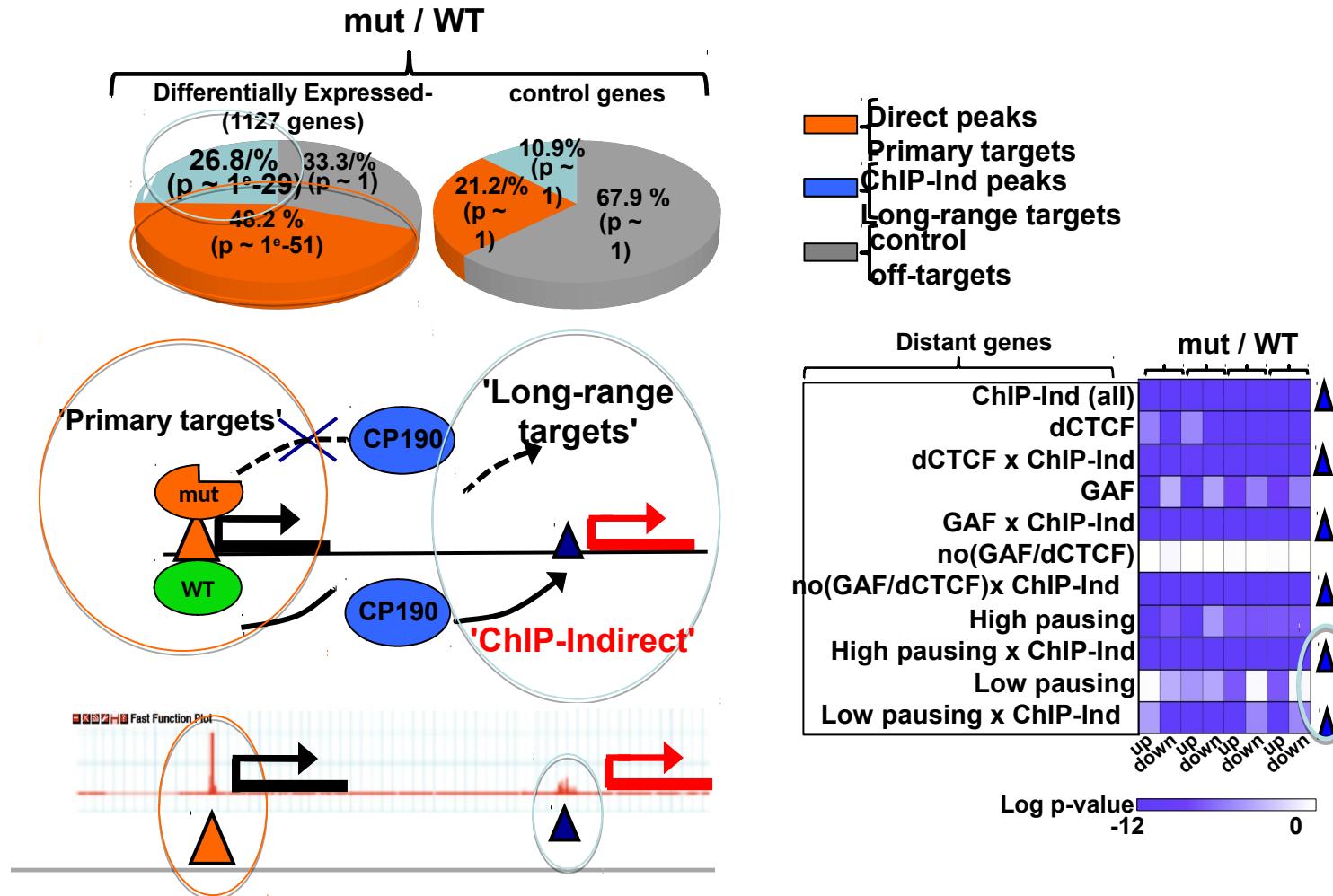
Affinity Chromatography



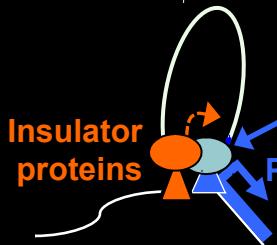
Functional KO of looping



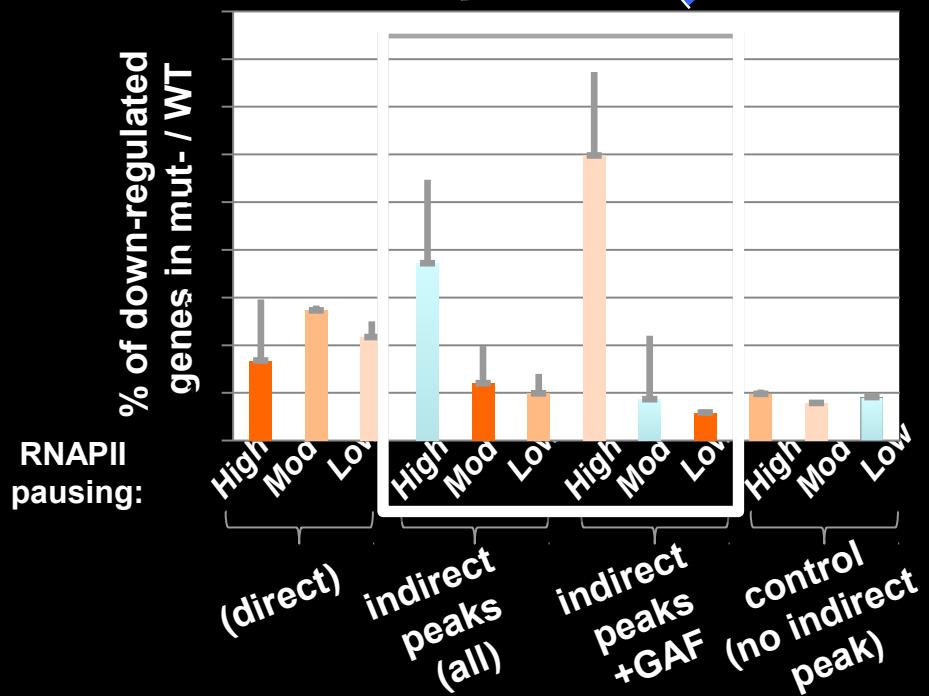
Functional KO of Looping



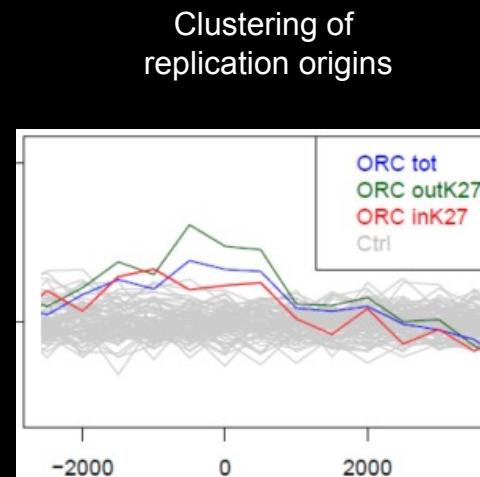
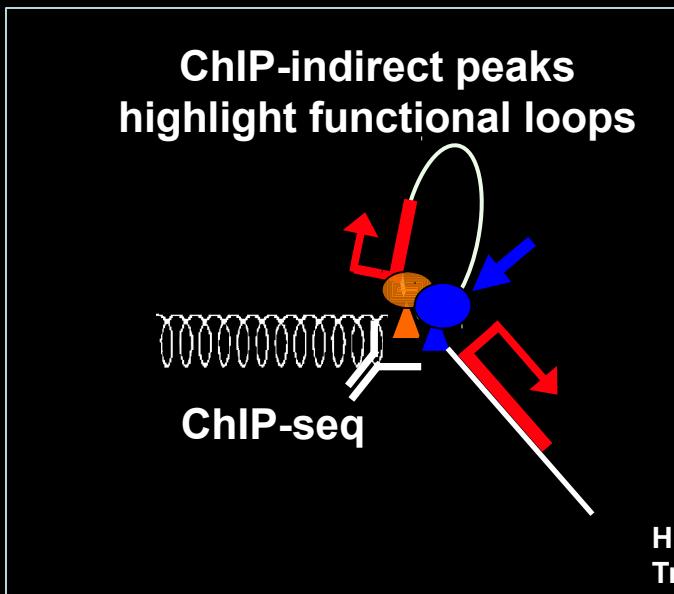
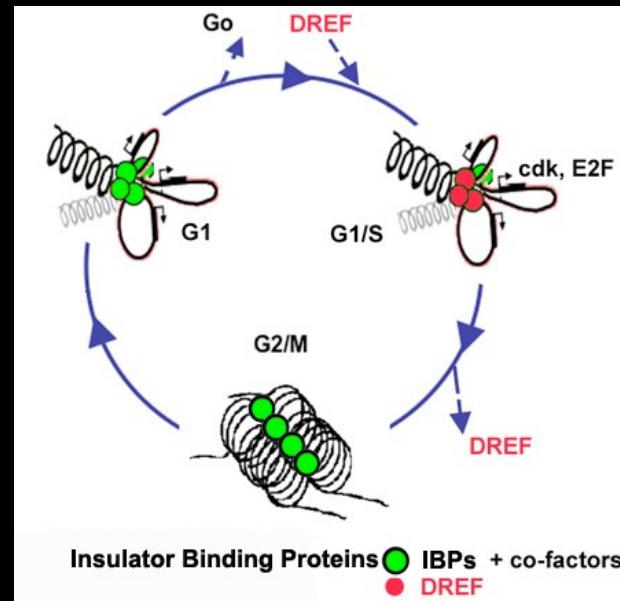
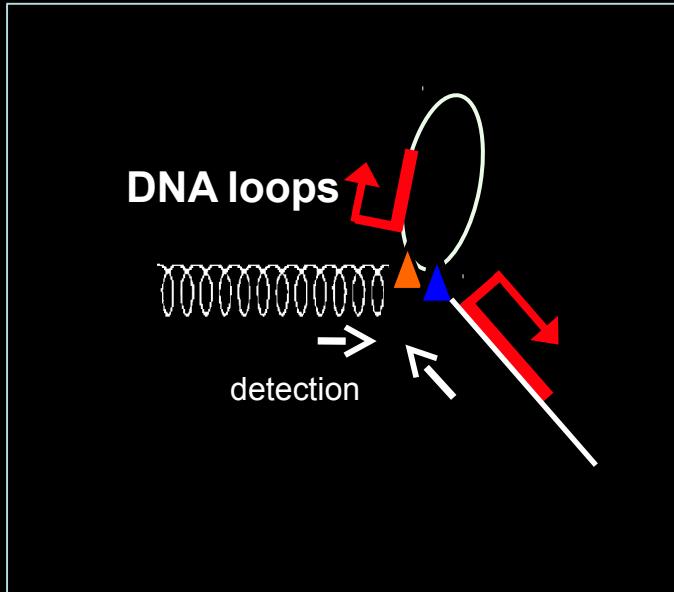
NELF-regulated RNA Polymerase II pausing prone genes to insulator-based Long-range contacts



Adelman and Lis, Nat. rev. Genet, 2012

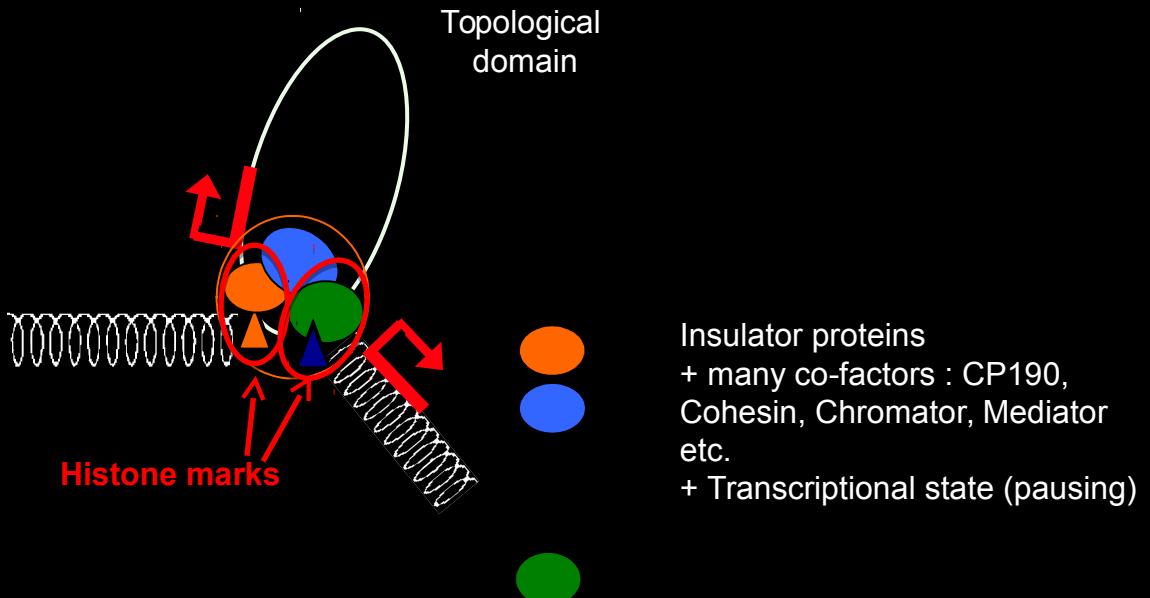


CONCLUSIONS & PERSPECTIVES / PART I



Humbert, Martin, Micas, Lhoumaud, Cucchi, Lacroix and Cuvier
Transcription-coupled clustering of Replication origins. *MS in prep.*

Model for insulator-based DNA looping

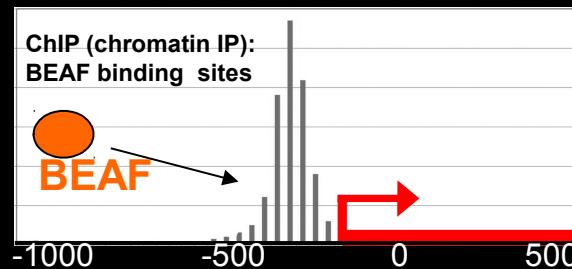
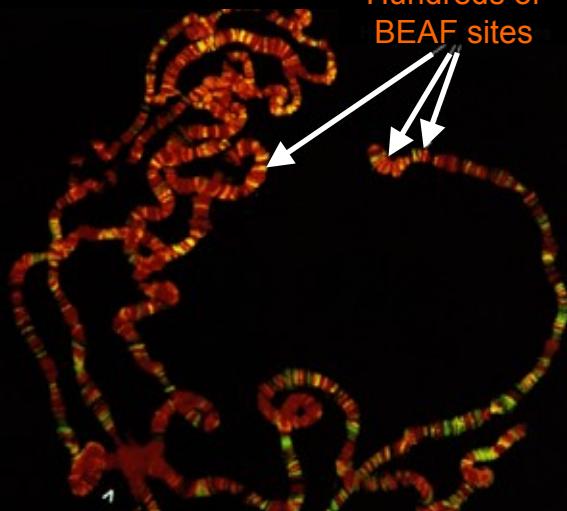
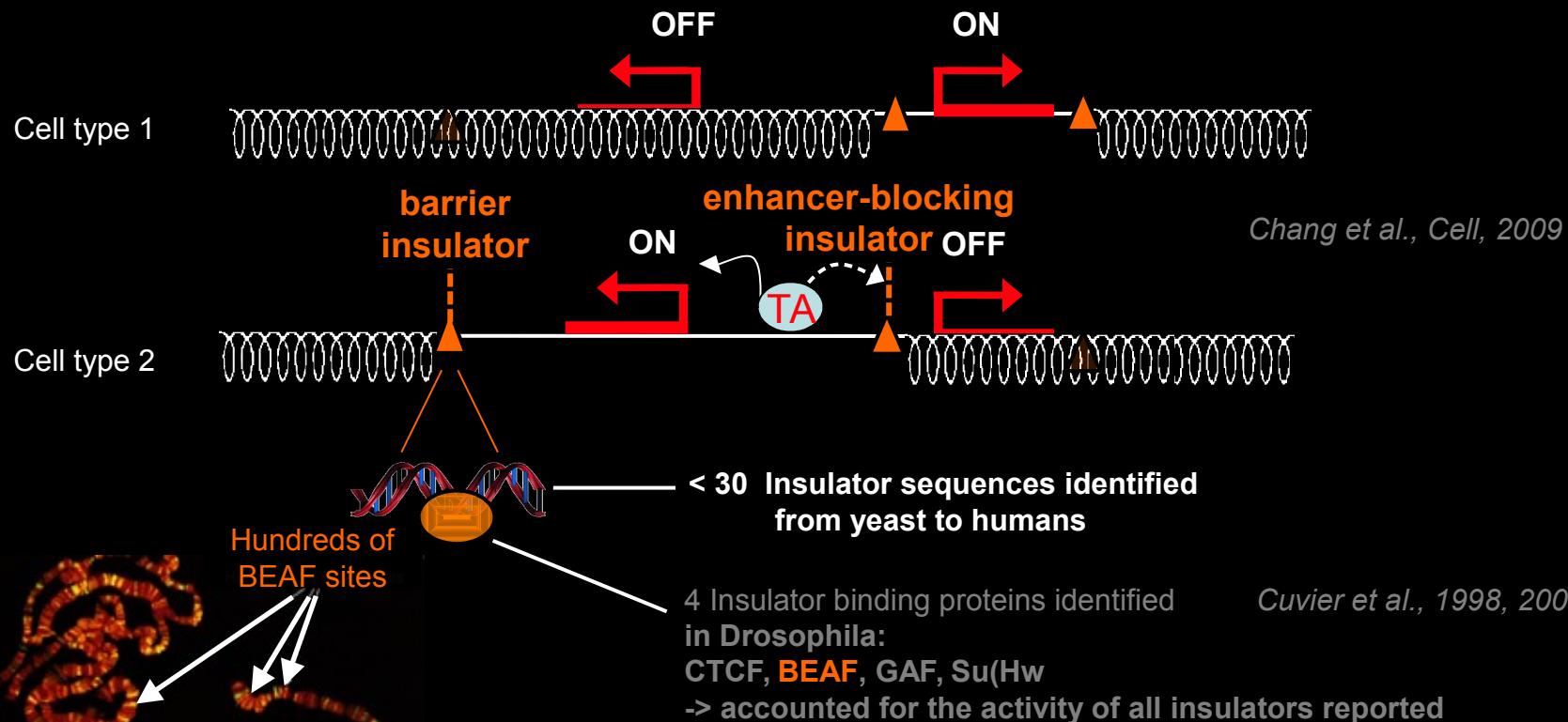


Liang et al., Mol Cell, 2014

Phillips-Cremins et al., Cell, 2013

Nora et al., Nature, 2013

Insulators specify Chromatin domains, Cell identity ?



>2000 genes
(Cell Prolif., Chrom. segregation...)

- Emberly, Blattes, Schuttengruber, Jiang, Hennion, Hart, Kas, Cuvier, PLoS Biol 2008*
Jiang, Emberly, Cuvier Hart MCB. 2009
Bushey et al. Corces : also true for CTCF Genes & Dev., 2009
Nègre et al. PloS Gen. 2009; Nature, 2011