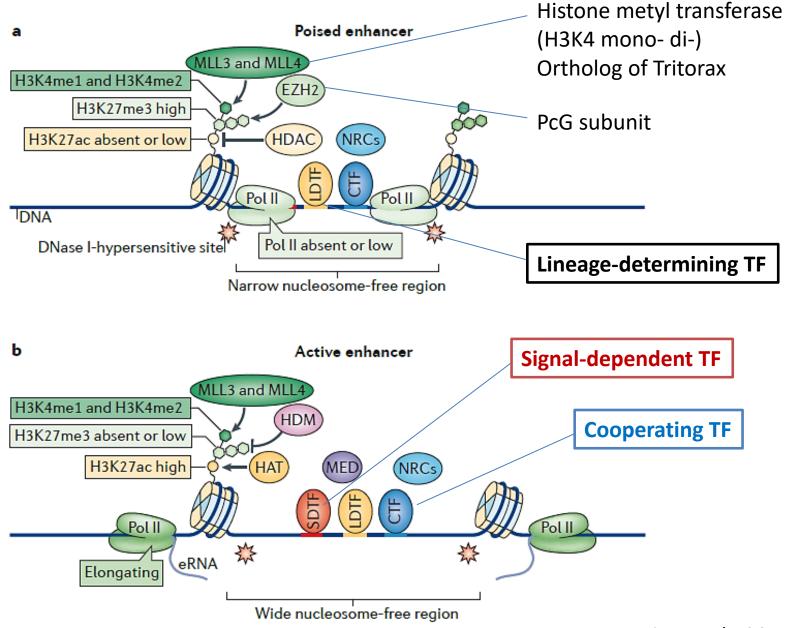
L4.6 – Transcriptional regulation (Enhancer Activation)

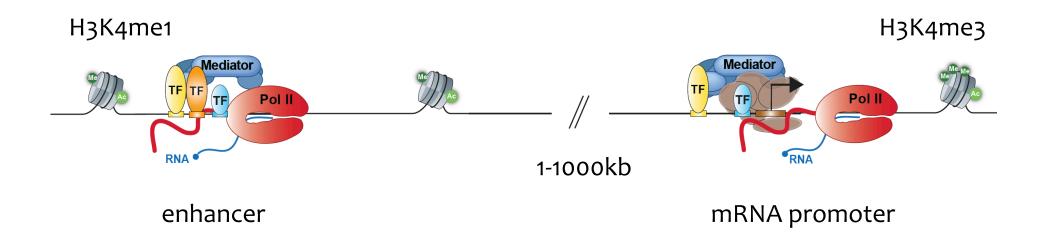
AGENDA

- 1. eRNA transcription
- 2. Enhancer-promoter looping
- 3. Topologically associated domains and CTCF



Heinz et al., 2015

Features of Enhancer-associated transcripts (eRNAs)



- eRNA transcription is associated with active enhancers
- eRNA transcription is responsive to stimulation
- eRNAs are unstable, long noncoding transcripts in both directions around enhancers
- eRNAs are **quantitatively correlated** with enhancer-regulated mRNA
- eRNAs are mainly poly(A-)

ARTICLES

Widespread transcription at neuronal activity-regulated enhancers

Tae-Kyung Kim¹*[†], Martin Hemberg²*, Jesse M. Gray¹*, Allen M. Costa¹, Daniel M. Bear¹, Jing Wu³, David A. Harmin^{1,4}, Mike Laptewicz¹, Kellie Barbara-Haley⁵, Scott Kuersten⁶, Eirene Markenscoff-Papadimitriou¹[†], Dietmar Kuhl⁷, Haruhiko Bito⁸, Paul F. Worley³, Gabriel Kreiman² & Michael E. Greenberg¹

We used genome-wide sequencing methods to study stimulus-dependent enhancer function in mouse cortical neurons. We identified ~12,000 neuronal activity-regulated enhancers that are bound by the general transcriptional co-activator CBP in an activity-dependent manner. A function of CBP at enhancers may be to recruit RNA polymerase II (RNAPII), as we also observed activity-regulated RNAPII binding to thousands of enhancers. Notably, RNAPII at enhancers transcribes bi-directionally a novel class of enhancer RNAs (eRNAs) within enhancer domains defined by the presence of histone H3 monomethylated at lysine 4. The level of eRNA expression at neuronal enhancers positively correlates with the level of messenger RNA synthesis at nearby genes, suggesting that eRNA synthesis occurs specifically at enhancers that are actively engaged in promoting mRNA synthesis. These findings reveal that a widespread mechanism of enhancer activation involves RNAPII binding and eRNA synthesis.

- Enhancers identified using ChIP-Seq with CBP antibodies (12,000, activated neurons)
- RNA Pol II colocalizes with several enhancers (ChIP-Seq)
- Level of eRNA expression correlates with level of expression of nearby genes

Research Paper – 3

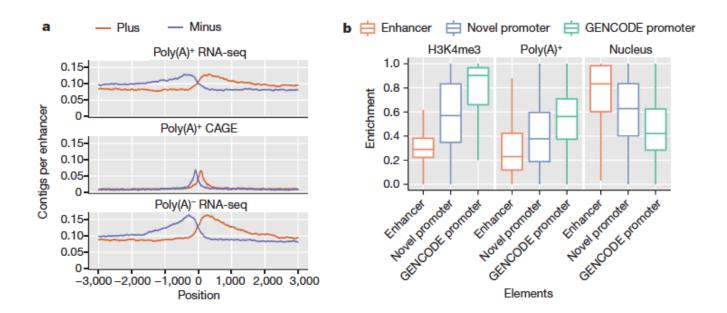
ARTICLE

Landscape of transcription in human cells Djebali et al., 2012 (ENCODE paper)

doi:10.1038/nature11233

Characterization of enhancer RNA

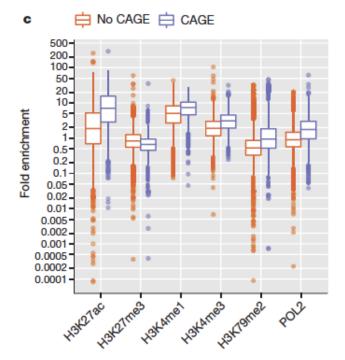
It has recently been reported that RNA polymerase II binds some distal enhancer regions and can produce enhancer-associated transcripts named eRNA^{33–35}. We used our RNA assays to detect and characterize transcriptional activity at enhancer loci predicted genome-wide from ENCODE chromatin immunoprecipitation and high-throughput sequencing (ChIP-seq) data^{20,36}.



Research Paper – 3

ARTICLE

Landscape of transcription in human cells Djebali et al., 2012 (ENCODE paper)



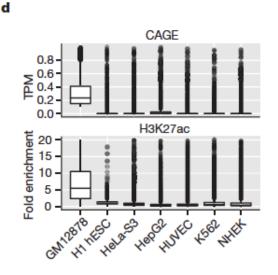


Figure 5 | **Transcription at enhancers.** a, The pattern of RNA elements around enhancer predictions^{20,36} containing DNase I hypersensitive sites. The lines represent the average frequency of RNA elements (top, polyadenylated long RNA contigs; middle, CAGE tag clusters; bottom, non-polyadenylated long RNA contigs) in a genomic window around the centre of the enhancer prediction as determined by DNase I hypersensitive sites. Elements on the plus strand are shown in red, and on the minus strand in blue. b, Enhancer transcripts differ from promoter transcripts. The box plots compare the features of transcripts at predicted enhancer loci compared to predicted novel intergenic promoters²⁰ and annotated promoters⁸. H3K4me3, poly(A)⁺ and nucleus denote the three following ratios: H3K4me3/(H3K4me3 + H3K4me1), polyadenylated /(polyadenylated + non-polyadenylated), nuclear/(nuclear +

H3K4me3 than novel or annotated promoters (left). Enhancer transcripts show higher levels of non-polyadenylated (middle) and nuclear (right) RNA relative to promoters. c, Chromatin state at transcribed enhancers. Enhancer predictions with evidence of transcription (in blue; Cage tags present at predicted locus) show a different pattern of histone modification and higher levels of RNA polymerase II binding than non-transcribed predictions (red). They are enriched for H3K27 acetylation, H3K4 methylation, H3K79 dimethylation and depleted for H3K27 trimethylation. d, Enhancer activity and transcription is cell-type specific. Loci predicted to be active transcribed enhancers in GM12878 cells show low signal for CAGE tags (top) and for H3K27 acetylation (bottom) in other cell lines. The whiskers are defined as Q1 $-1.5 \times IOR$ to Q3 $\pm 1.5 \times IOR$, where IOR is the interquartile range, and Q1

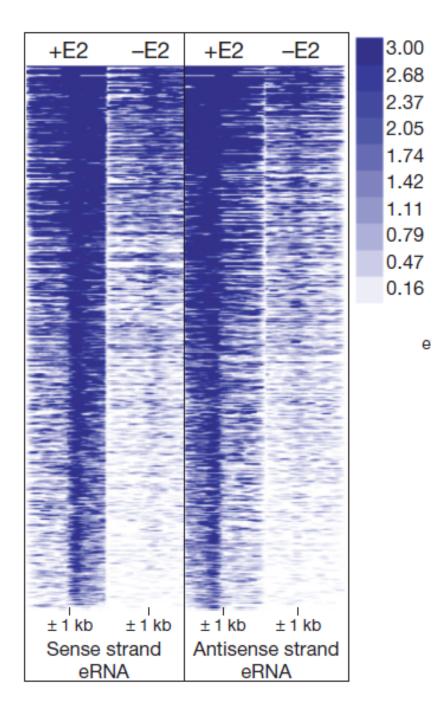
LETTER

Functional roles of enhancer RNAs for oestrogen-dependent transcriptional activation

Wenbo Li¹*, Dimple Notani¹*, Qi Ma^{1,2}, Bogdan Tanasa^{1,3}, Esperanza Nunez¹, Aaron Yun Chen¹, Daria Merkurjev^{1,2}, Jie Zhang¹, Kenneth Ohgi¹, Xiaoyuan Song¹, Soohwan Oh^{1,4}, Hong-Sook Kim¹, Christopher K. Glass⁵ & Michael G. Rosenfeld¹

The functional importance of gene enhancers in regulated gene expression is well established1-3. In addition to widespread transcription of long non-coding RNAs (lncRNAs) in mammalian cells4-6, bidirectional ncRNAs are transcribed on enhancers, and are thus referred to as enhancer RNAs (eRNAs)⁷⁻⁹. However, it has remained unclear whether these eRNAs are functional or merely a reflection of enhancer activation. Here we report that in human breast cancer cells 17β-oestradiol (E2)-bound oestrogen receptor a (ER-α) causes a global increase in eRNA transcription on enhancers adjacent to E2-upregulated coding genes. These induced eRNAs, as functional transcripts, seem to exert important roles for the observed ligand-dependent induction of target coding genes, increasing the strength of specific enhancer-promoter looping initiated by ER-a binding. Cohesin, present on many ER-aregulated enhancers even before ligand treatment, apparently contributes to E2-dependent gene activation, at least in part by stabilizing E2/ER-a/eRNA-induced enhancer-promoter looping. Our data indicate that eRNAs are likely to have important functions in many regulated programs of gene transcription.

- eRNA transcription from putative enhancers of E2-regulated genes is also induced by E2
- E2-induced eRNAs are required for the activation of the induction of corresponding coding genes.
- eRNAs might be required to promote/sustain promoter looping (NRIP and GREB loci)

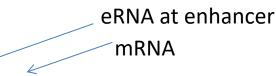


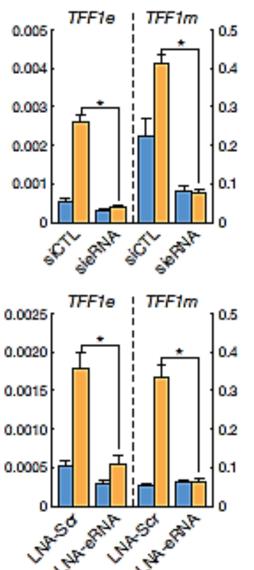
GRO-Seq (Nascent RNA) Strand-specific library prep

In this representation, each line is an enhancer, and the density of reads is depicted in the -1Kb to +1Kb interval, using a scale of blue color.

The lines are then ordered from the most transcribed to the less transcribed

Li et al., Nature 2013





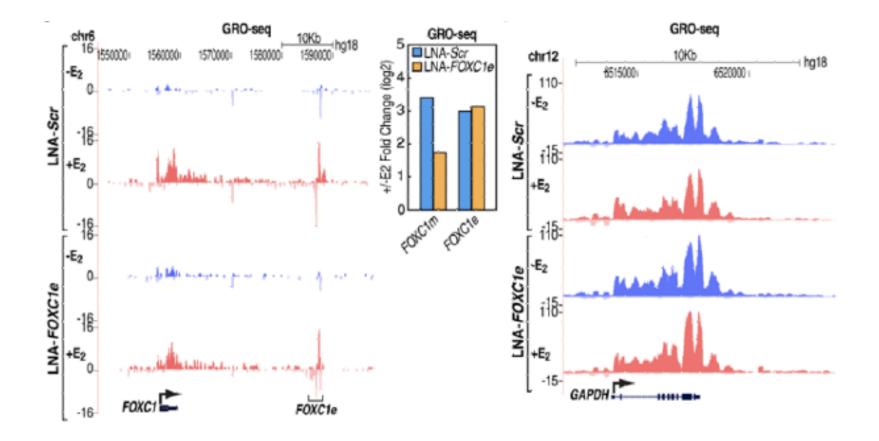
Based on GRO-seq analyses, we selected ten highly upregulated transcription units for further experimentation, each associated with adjacent UP-enhancers exhibiting ~2.5-5-fold E_2 -induction of eRNAs

Figure 2 | Importance of eRNA for target gene activation. a, b, siRNA/LNA knockdown of eRNAs. Efficacy and effects on coding gene transcription were assessed by qPCR for the TFF1, FOXC1 and CA12 eRNAs and corresponding coding transcription units.

Lower case 'e' and 'm' after gene names denote eRNA and gene mRNA, respectively. CTL, control; Scr, scramble.

LNA= Locked Nucleic Acid

Li et al., Nature 2013

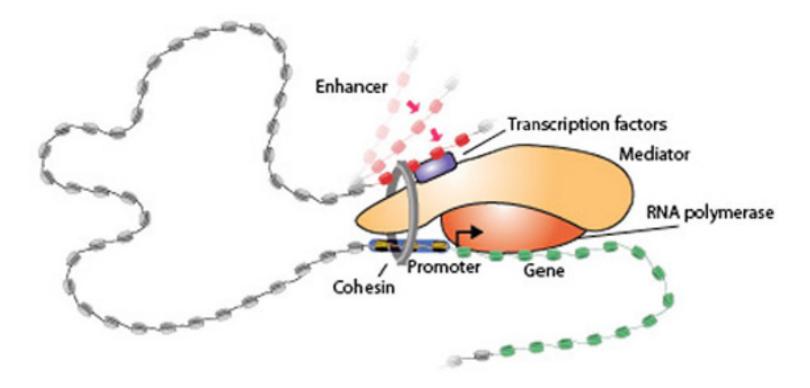


(e) GRO-seq data from FOXC1e LNA treated cells showing its inhibitory effect on the transcription of FOXC1 coding locus, but not on the targeted enhancer region itself. FOXC1e LNA does not affect GAPDH transcription.

AGENDA

- 1. eRNA transcription
- 2. Enhancer-promoter looping
- 3. Topologically associated domains and CTCF

Enhancer and promoter are thought to come together in the 3D space through formation of a loop

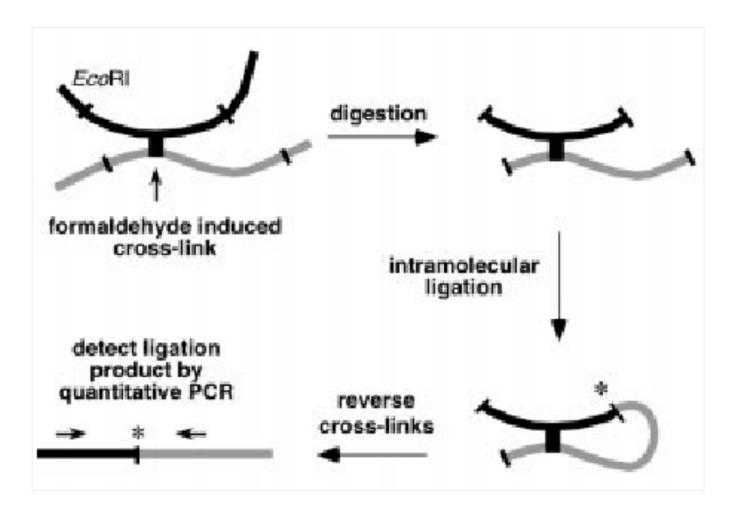


How do we study this process? Can we measure long-range interactions?

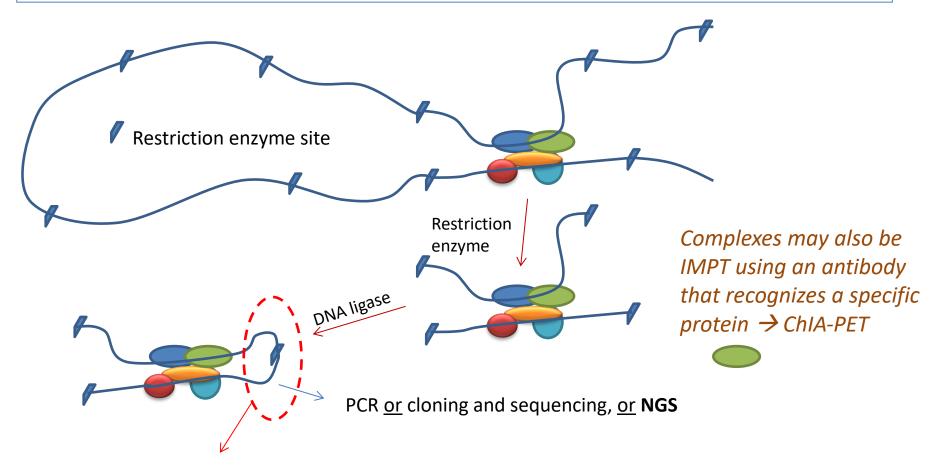
What are the mediators of interaction ?

Is the looping functional? Is it required?

The basic methodology to study Enhancer-Promoter interaction is 3C assay = chromosome conformation capture



Long-range interactions are studied with 3C (Chromatin Conformation Capture) or different genome-wide scale variants (4C, 5C, Hi-C, ChIA-PET).



PCR for single interaction. Generate libraries to NGS for genome-wide studies

Note: from this scheme nucleosomes are omitted

Looping and Interaction between Hypersensitive Sites in the Active β-globin Locus

Bas Tolhuis,² Robert-Jan Palstra,² Erik Splinter, Frank Grosveld, and Wouter de Laat¹ Department of Cell Biology and Genetics Faculty of Medicine Erasmus University, Rotterdam P.O. Box 1738 3000DR Rotterdam The Netherlands

Summary

Eukaryotic transcription can be regulated over tens or even hundreds of kilobases. We show that such long-range gene regulation in vivo involves spatial interactions between transcriptional elements, with intervening chromatin looping out. The spatial organization of a 200 kb region spanning the murine B-globin locus was analyzed in expressing erythroid and nonexpressing brain tissue. In brain, the globin cluster adopts a seemingly linear conformation. In erythroid cells the hypersensitive sites of the locus control region (LCR), located 40-60 kb away from the active genes, come in close spatial proximity with these genes. The intervening chromatin with inactive globin genes loops out. Moreover, two distant hypersensitive regions participate in these interactions. We propose that clustering of regulatory elements is key to creating and maintaining active chromatin domains and regulating transcription.

receives support from studies on transcriptional retion of many different prokaryotic genes. In fact model was originally based on work on bacterial phage repressor proteins, like the Gal, AraC, and pressor proteins, which were found to function when homomultimerized and bound to two sep: operator sites. Electron microscopy visually der strated the DNA in between to loop out (review) Ptashne, 1986). Thus, both types of mechanisms ap to function in bacteria. Eukaryotes have more congene clusters with regulatory elements functioning much greater distances. To date, there are no data unambiguously demonstrate one (or more or comitions) of the models to be correct for the regulation a given eukaryotic locus. Support for models has c from indirect and/or in vitro observations, and ofter distinction between the activation and actual trans tion of a locus is not made. However with respe transcription, a number of observations can only e be explained by the looping model. The first typ experiments involves studies on trans-activation, the ability of an enhancer to activate a promoter preon a physically separate DNA molecule. Most impo in this respect is the naturally occurring phenomof transvection in Drosophila (Bickel and Pirrotta, 1 In addition, Schaffner and coworkers demonstrate vitro that enhancers can stimulate transcription in t by coupling an enhancer- to a promoter-containing i mid via a biotin-streptavidin bridge (Mueller et al., 1 eenaa e'e ea "èe

In humans, there are two gene clusters direct the synthesis of hemoglobins: the α locus, which contains the embryonic ζ gene and the two adult α genes; and the β locus, which consists of the ε , ${}^{G}\gamma$, ${}^{A}\gamma$, δ , and β genes. Two globin gene switches occur during development: the embryonic to fetal globin switch, which coincides with the transition from embryonic (yolk sac) to definitive (fetal liver) hematopoiesis; and the fetal to adult switch, which occurs at the perinatal period. The switches from ε to γ and from γ to β globin gene expression are controlled exclusively at the transcriptional level. The LCR confers lineage-specific expression on the globin genes; it acts as the major enhancer of the β locus; it insulates the locus from surrounding inactive chromatin.

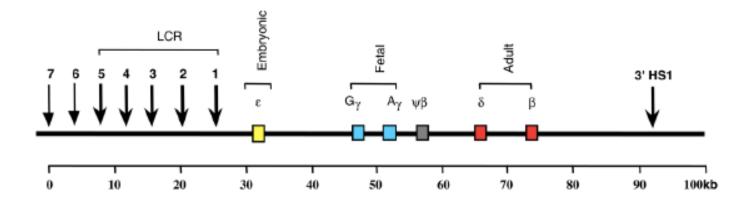


Figure 2. Diagram of the human β globin locus.

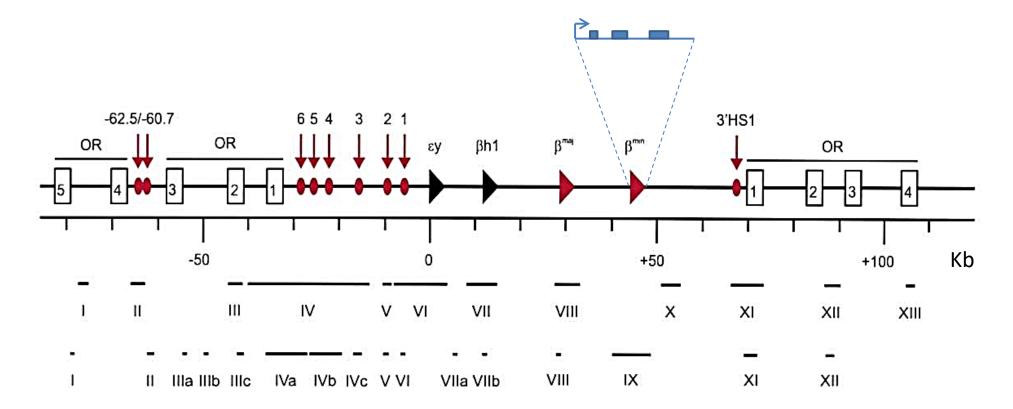
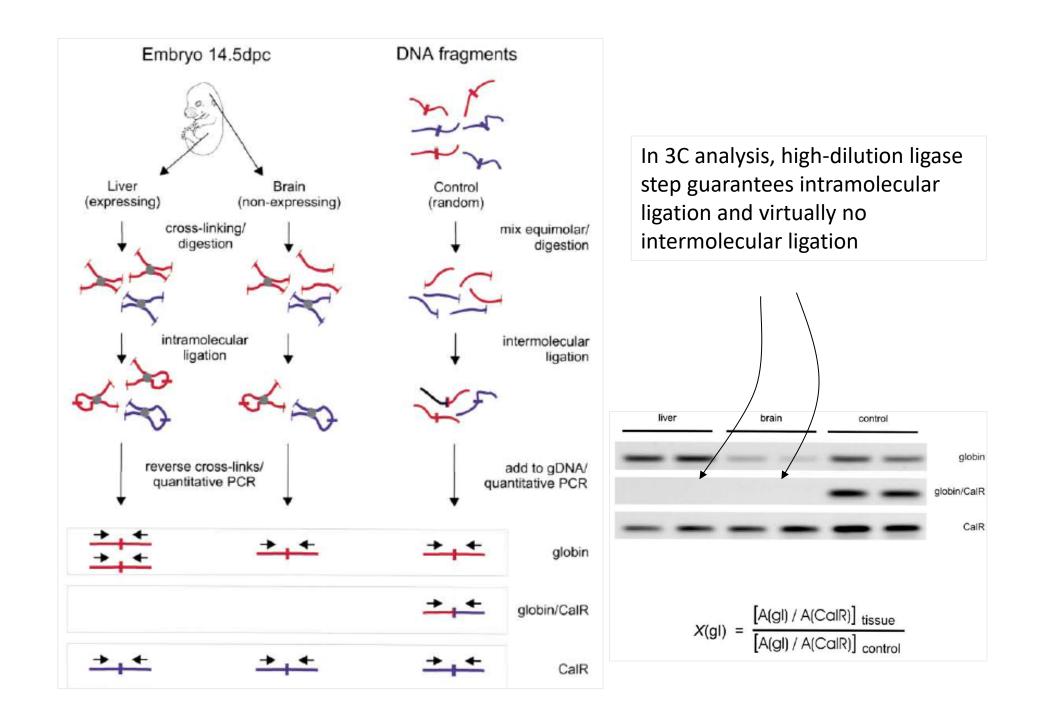


Figure 1. 3C Technology in the Murine beta-globin Locus

(A) Schematic presentation of the murine **beta-globin** locus. Red arrows and ellipses depict the individual HS. The globin genes are indicated by triangles, with **active genes (maj and min) in red** and **inactive genes (y and h1) in black**. The white boxes indicate the olfactory receptor (OR) genes (5OR1-5 and 3OR1-4). The two sets of restriction fragments (BgIII and HindIII) that were used for 3C analysis are shown below the locus. The individual fragments are indicated by Roman numerals. Identical numbering between BgIII and HindIII indicates that two fragments colocalize. Distances are in kb counting from the site of initiation of the y gene.



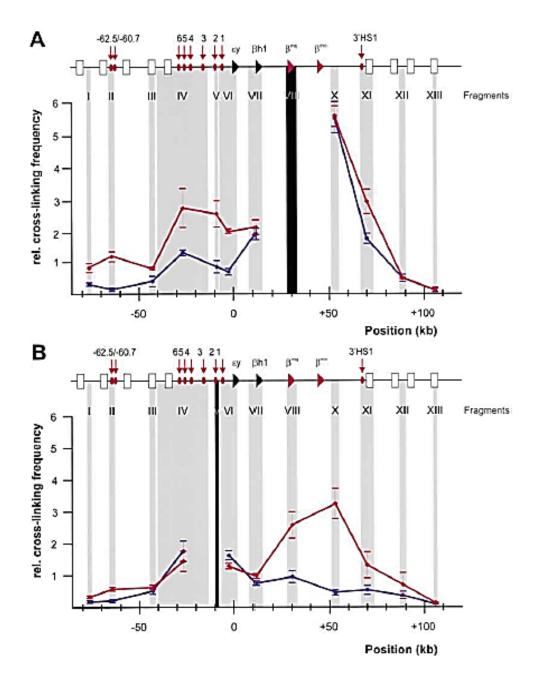


Figure 3. Erythroid-Specific Interaction and Looping between the LCR and an Active beta--globin Gene. Relative crosslinking frequencies observed in fetal liver are shown in red. For comparison, data obtained in brain are depicted in blue. Standard error of the mean is indicated. Crosslinking frequency with a value of 1 arbitrarily corresponds to the crosslinking frequency between two neighboring CalR control fragments (with restriction sites analyzed being 1.5 kb apart). Scaling on the *y* axis (from 0 to 6) allows direct comparison with Figures 2 and 4–6.

(A) Fixed BgIII fragment VIII (maj) versus the rest of the locus. (B) Fixed BgIII fragment V (5HS2) versus the rest of the locus. (C) Fixed BgIII fragment VII (h1) versus the rest of the locus.

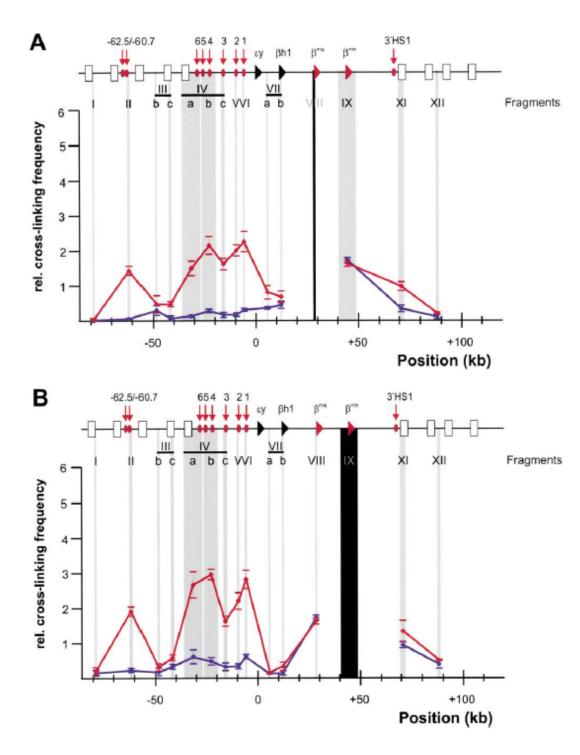


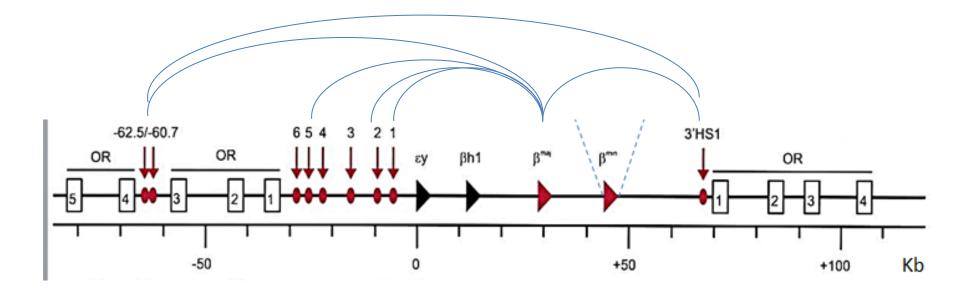
Figure 4. Erythroid-Specific Interactions between the Active beta-globin Genes and Individual Hypersensitive Sites in the LCR. Relative crosslinking frequencies observed in fetal liver (red) and brain (blue) are shown. Standard error of the mean is indicated. Crosslinking frequency with a value of 1 arbitrarily corresponds to the crosslinking frequency between two neighboring CalR control fragments (with restriction sites analyzed being 1.5 kb apart). Scaling on the y axis (from 0 to 6) allows direct comparison with other figures.

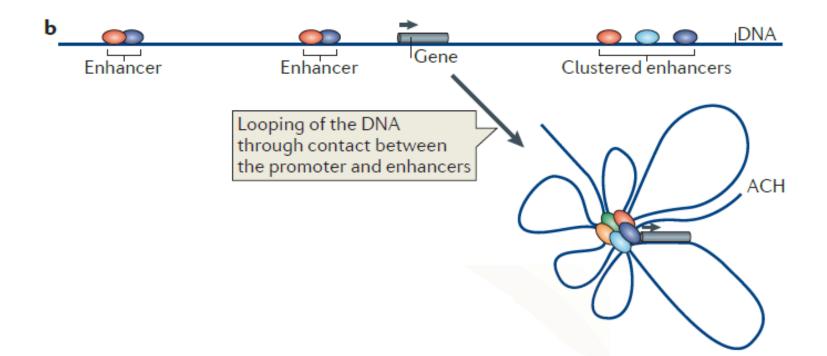
(A) Fixed HindIII fragment VIII Bmaj versus the rest of the locus.

(B) Fixed HindIII fragment IX (Bmin) versus the rest of the locus.

This data demonstrated multiple «looping» involving components of the super-enhancer LCR, other enhancers and promoters of the active B genes

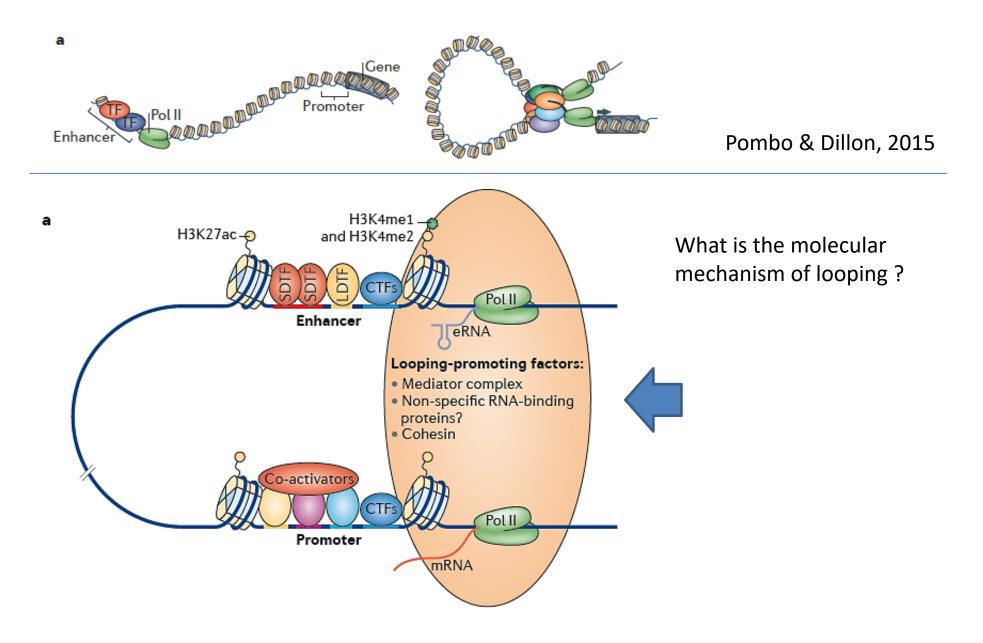
(the example below shows only some of the contacts verified)





Pombo & Dillon, 2015

BREAK \rightarrow Mechanisms to promote/support looping

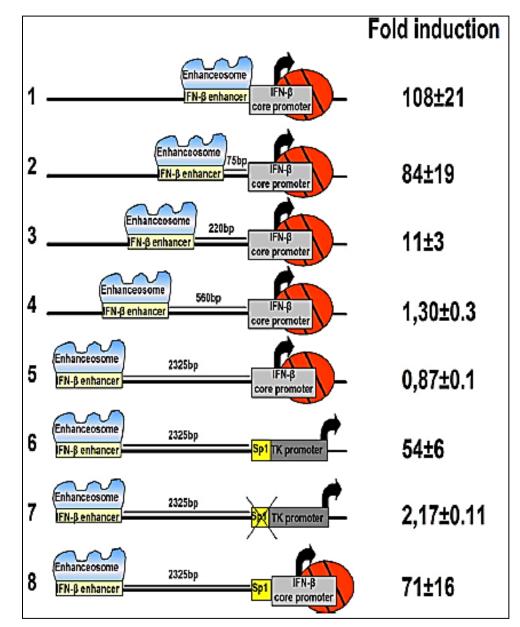


Nolis et al. **Transcription factors mediate long** *range enhancer–promoter interactions* PNAS 106:20222-227, 2009

Enhancer action from a distance requires upstream promoter elements \rightarrow TFs binding proximaly is required for enhancer function.

HeLa cells were transfected with the indicated chloramphenicol acetyl transferase (CAT) reporter plasmids. The cells were mock or virus infected for 24 h before being harvested. Then CAT activity was determined.

Line 1 is the natural arrangement.



LETTER

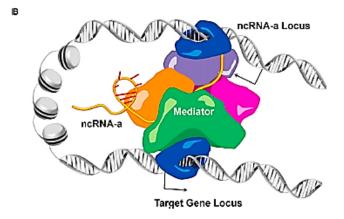
doi:10.1038/nature11884

Activating RNAs associate with Mediator to enhance chromatin architecture and transcription

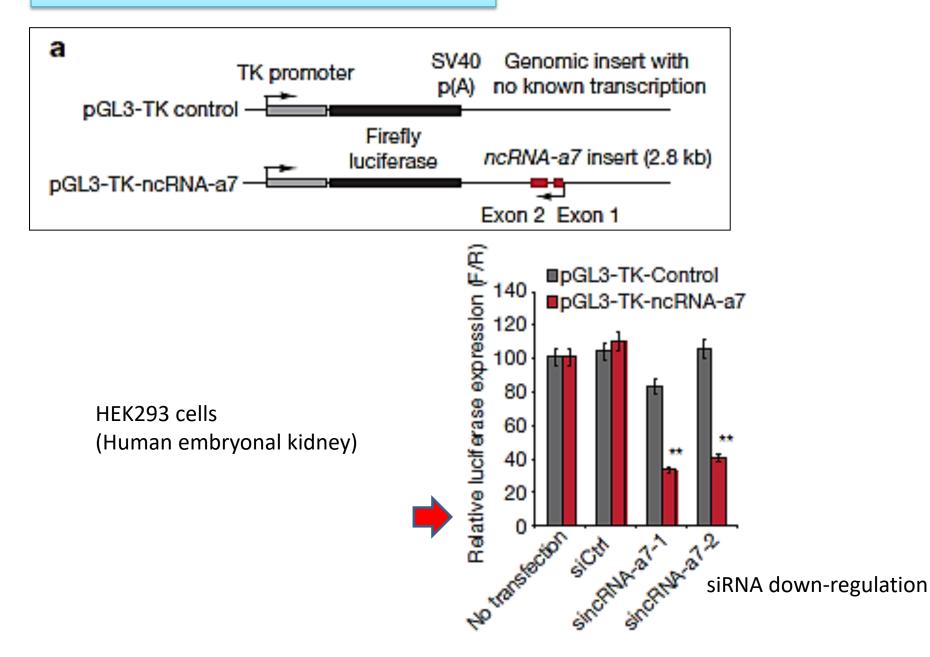
Fan Lai¹, Ulf A. Orom², Matteo Cesaroni¹, Malte Beringer³, Dylan J. Taatjes⁴, Gerd A. Blobel⁵ & Ramin Shiekhattar¹

In previous work they found IncRNAs with enhancerlike properties (Orom et al., Cell 2010): they identified a small subsets of IncRNAs, termed **ncRNA-activating** (ncRNA-a), that function to activate their neighbouring genes.

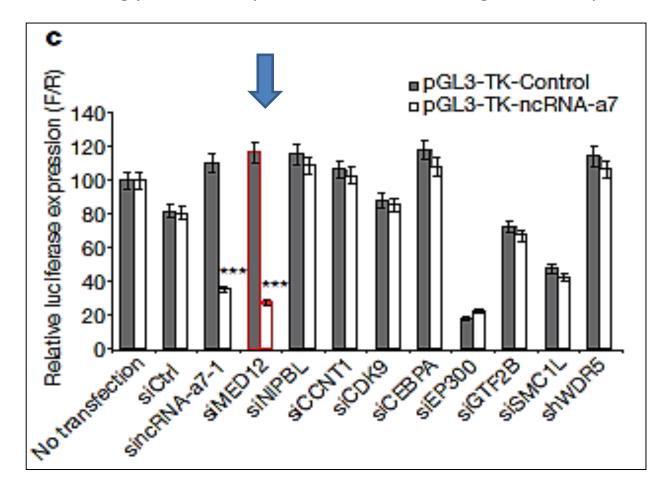
For 7 of them, they systhematically siRNA noncoding RNAs and identified neighbouring down-regulated genes



1st question: is activation ncRNA dependent?



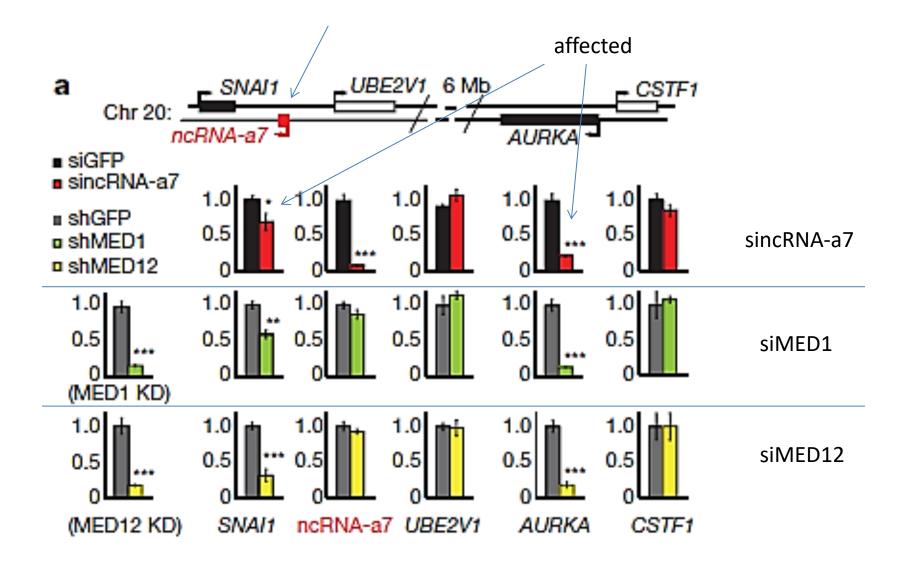
2nd question: which component of the transcriptional machinery is involved ?



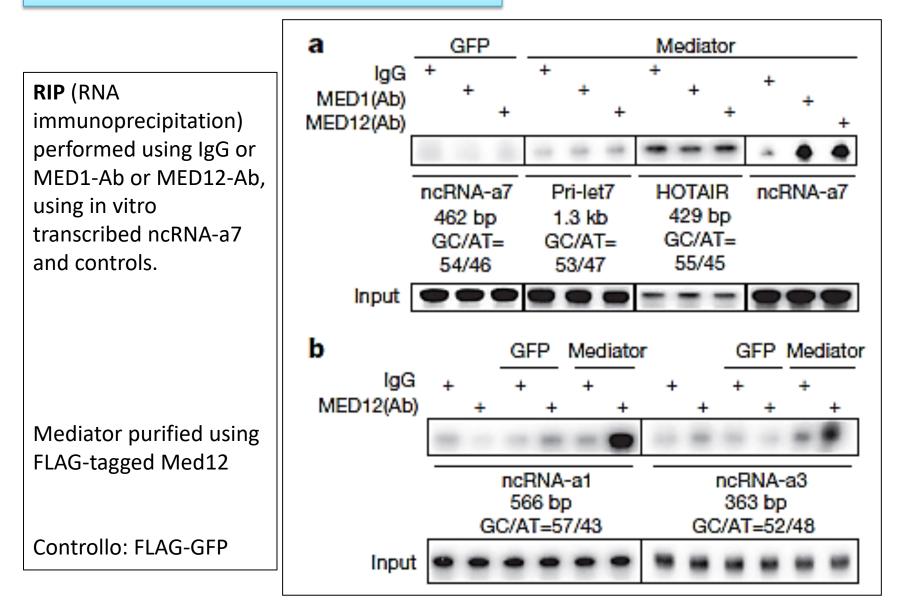
Screening protein components for function in gene activity

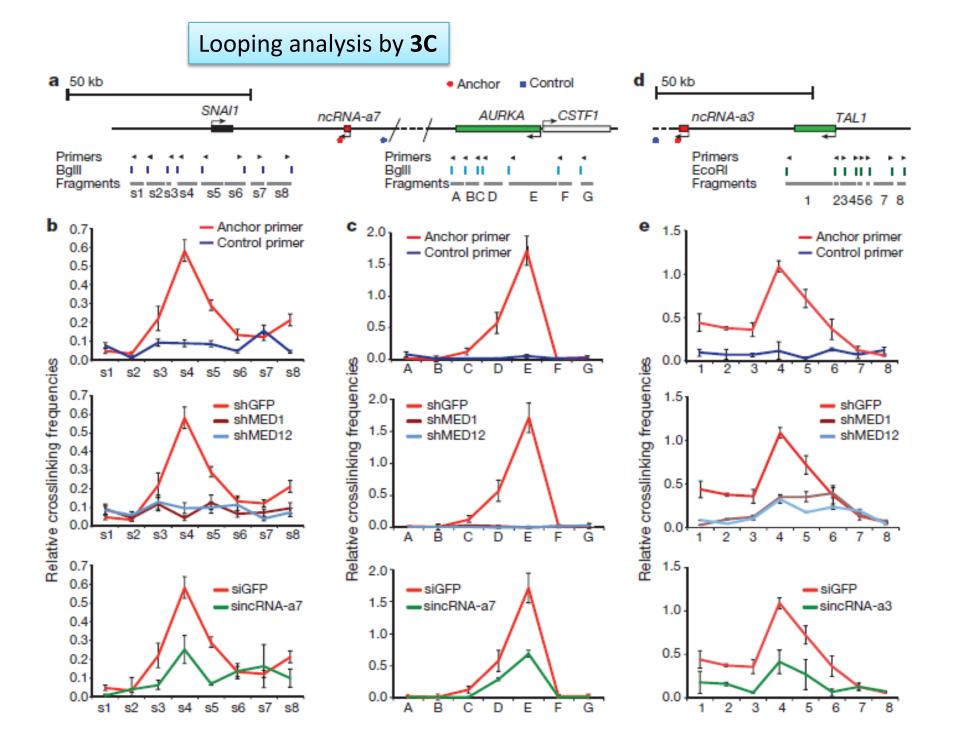
MED12 is the only protein, among those tested, that affects RNA-a function (not transcription of the reporter per se, relative transcription levels!)

3° question: is this effect reproducible on the endogenous loci?



Authors demonstrated ncRNA-a/MED binding

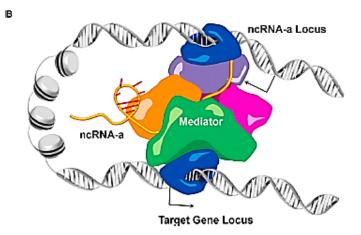




Conclusions

Activating nc-RNA binds to Mediator

They are required for supporting the activation of neighbouring genes through regulation of promoter-enhancer looping.



Q: Are these eRNAs?

From the paper Discussion:

A recent study uncovers a set of bi-directional transcripts (termed eRNA) that are derived from sites in the human genome that show occupancy by CBP, RNA polymerase II and are decorated by monomethyl Histone H3 lysine 4 (H3K4) (Kim et al., 2010). Moreover, they show that the expression of such transcripts is correlated with their nearest protein-coding genes. There are fundamental differences between their collection of ~2,000 transcripts and our GENCODE set of transcripts. First, while all their eRNAs are bidirectional, only about one percent of our ncRNAs show evidence of bidirectionality (see the example shown in the TAL1 locus). Second, our analysis of the histone modifications of a subset of ncRNAs that are expressed in lymph (Barski et al., 2007) indicates the presence of H3K4 trimethylation at the transcriptional start sites and H3K36 trimethylation at the body of the gene (Figure S1B and C). This is in stark contrast to eRNA loci where there is an absence of H3K4 trimethyl marks and the predominant chromatin signature is the monomethyl H3K4. Third, eRNAs are reported to be predominantly not polyadenylated. The majority of our collection of ncRNAs show evidence of polyadenylation as they were amplified using oligo-dT-primed reactions and furthermore 41 percent display the presence of a canonical polyadenylation site. Analysis of the protein-coding transcripts revealed that a similar proportion (52 percent) to that of our ncRNAs contain the canonical polyadenylation sites. Finally, while we show that a set of our ncRNAs function to enhance gene expression, there is no evidence provided for eRNAs exerting a biological function. While we believe that eRNAs designate a different class of ncRNAs than ncRNA-a described in our study, it is temping to speculate that many of the ncRNA-a and their promoters may correspond to mammalian enhancers or polycomb/trithorax response elements (PRE/TREs). In such a scenario, binding of polycomb or trithorax proteins to proximal promoters of ncRNA-a will regulate the expression of ncRNA-a which in turn impact the expression of the protein-coding gene at the distance.

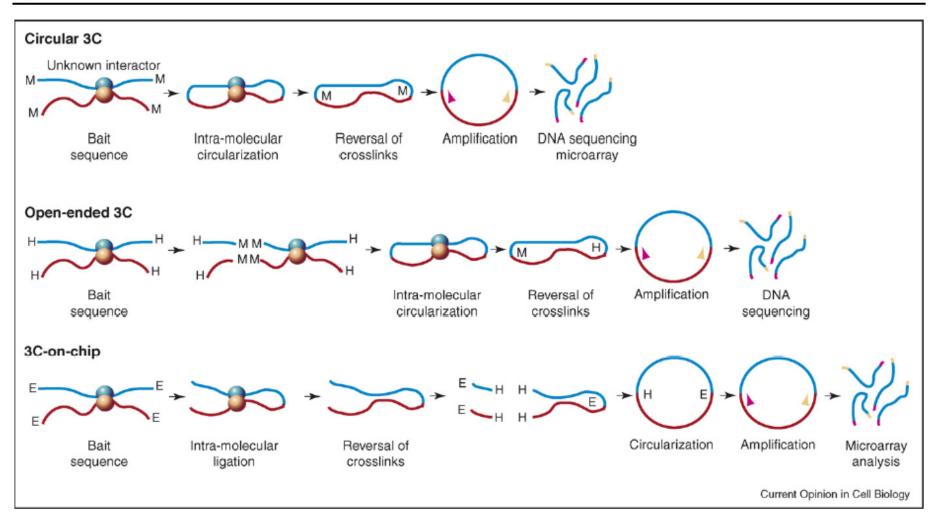
The next problem that we should consider is that the model of enhancer-promoter looping discussed is an over-simplification.

Indeed, in a real mammalian cell, each enhancer or promoters show many contacts and multiple long-range interactions.

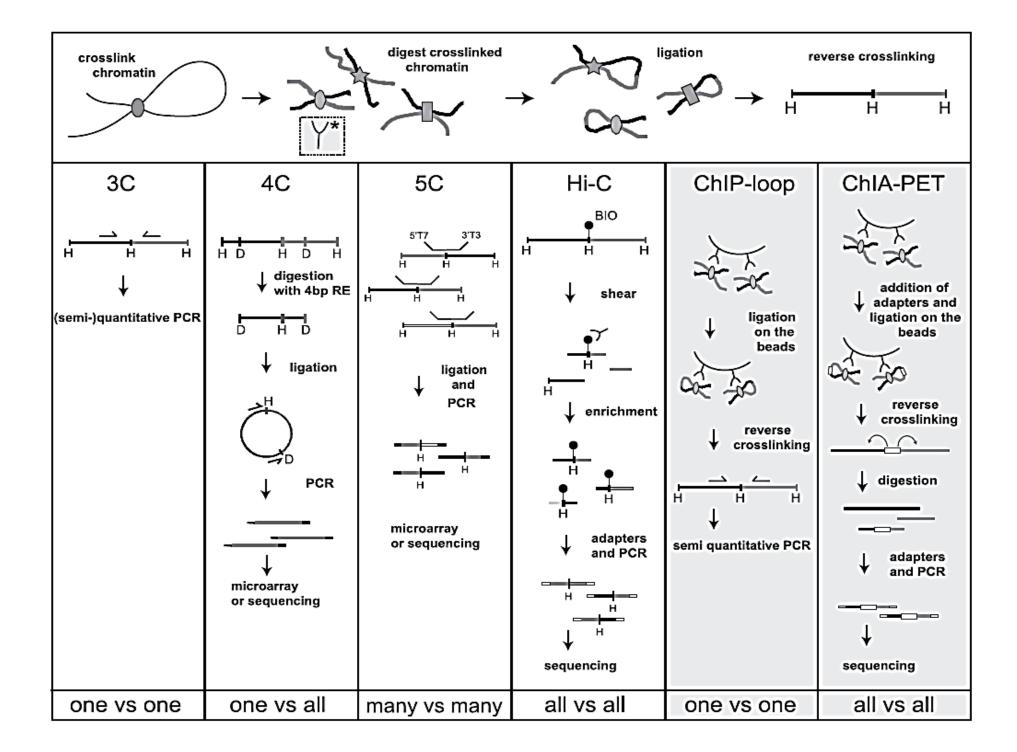
First discussed for the beta-globin locus (see slides before)

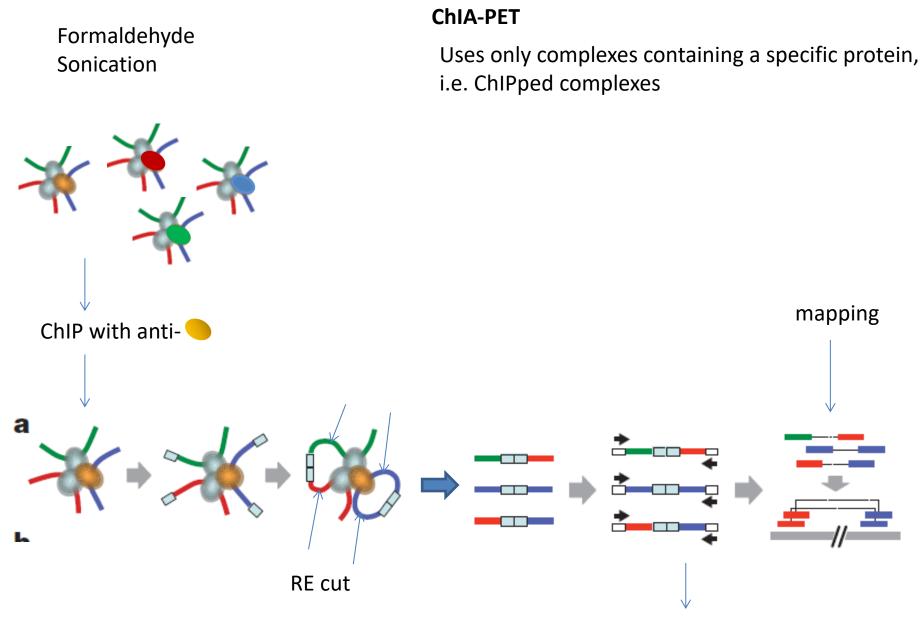
Now this has been clearly shown by many studies using Hi-C and other technologies.





A schematic comparison of the 4C approaches. All three methods shown use a circular DNA as a vehicle to identify unbiased interactors by means of inverse PCR primers. The 3C-on-chip generates the circular DNA structure after the reversal of crosslinking step, whereas the other methods generate circular DNA while the interacting sequences remain in a crosslinked DNA/protein complex. The methods also differ with respect to the use of restriction enzymes (M, *Msp* I; E, *Eco* RI; H, *Hind* III). A fourth 4C method, 'olfactory receptor' 3C [12*], did not specify the restriction enzymes used, but appears to be most closely related to the circular 3C principle.





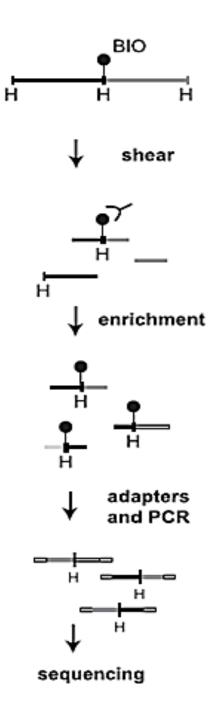
PCR amplify and mass-sequencing

HiC

Is a technique «all versus all»

After digestion and before ligation, sticky ends are filled using biotinylated nucleotides, so that ligation junctions remain marked with Biotin and can be enriched using streptavidin beads.

After this step, fragments are processed, amplified and NGSequenced as in other methods.



LETTER

Few conclusion from ENCODE studies (5C)

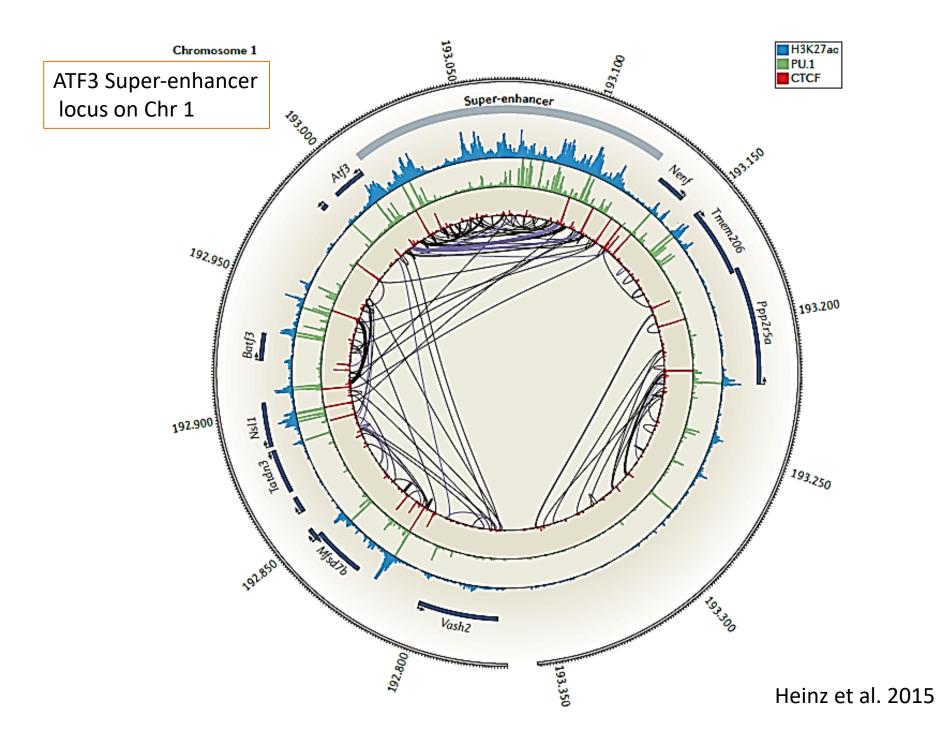
The long-range interaction landscape of gene promoters

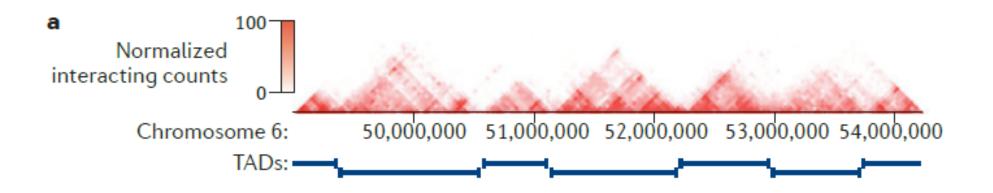
Amartya Sanyal¹*, Bryan R. Lajoie¹*, Gaurav Jain¹ & Job Dekker¹

In each cell line large numbers of statistically significant TSS–distal fragment interactions were identified, of which ca. 60% were observed <u>in</u> <u>only one</u> of the three cell lines (i.e. active enhancers show cell specificity)

Interactions are not limited to most proximal E/P couple: only 7% of interaction link an enhancer to the most proximal TSS.

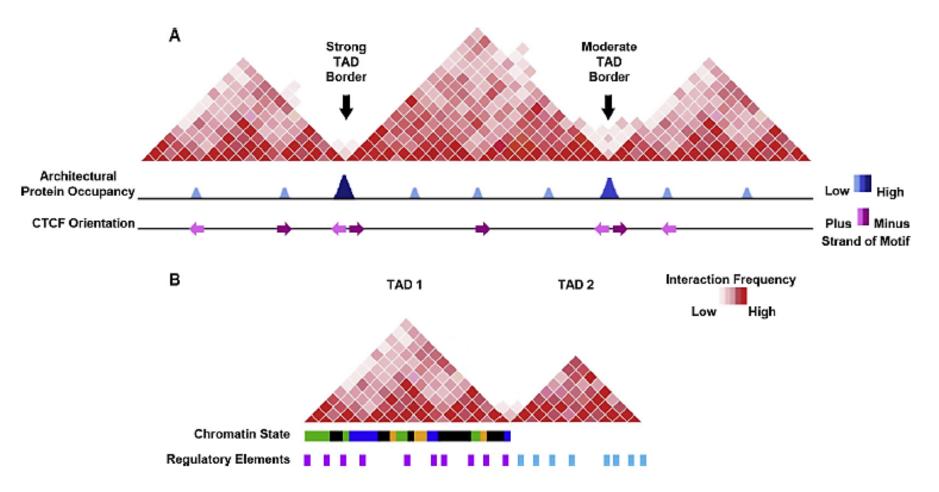
Very high number of interactions: some TSS show up to 20 contemporaneous interactions.





a | Hi-C profiles reveal that the mammalian genome is organized into **topologically associating domains (TADs)**: regions that show high levels of interaction within the region and little or no interaction with neighbouring regions. The heat map represents normalized Hi-C interaction frequencies.

Borders can be stronger or wealer, i.e. extra-TAD interactions are sometimes permitted



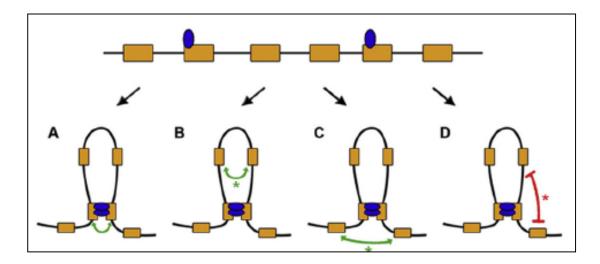
C. Cubeñas-Potts, V.G. Corces/FEBS Letters 589 (2015) 2923-2930

TADs are defined by interactions and are bordered by highly transcribed regions (housekeepers).

The large loops of chromain that define TADs are due to specific proteins that bind DNA at specific sequences and interact reciprocally.

They are defined «**Architectural Proteins**» and the sequences of DNA that are recognized by APS are called APBS.

By far, the most studied AP in Mammals is CTCF, but others exist. CTCF also interacts with cohesins, which are supposed to stabilize loops and that some studies have shown to be essential for enhancer activity.



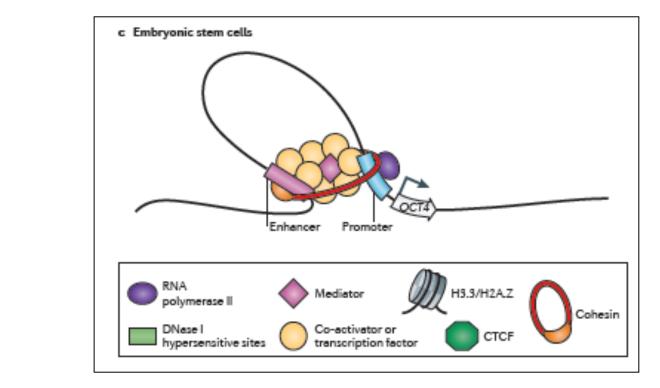
From Cabenas-Potts & Corces 2015

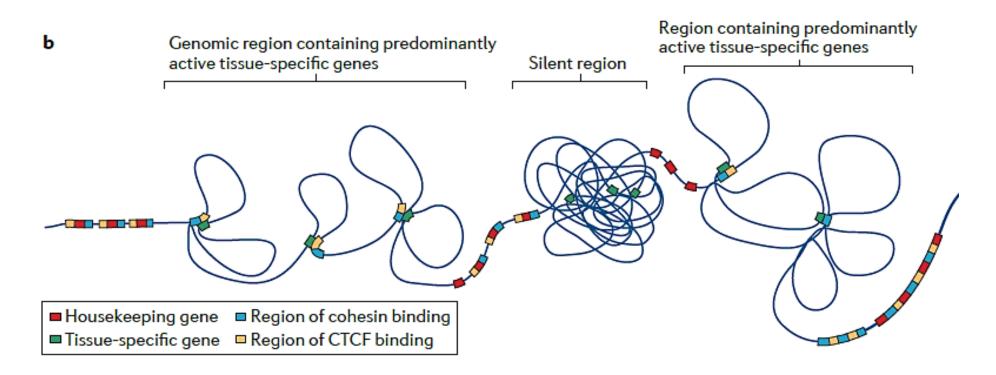
Cohesin stabilizes long-range interactions.

cohesins

Cohesins interact with the CTCF protein, a CCCTC-binding protein, methylation-sensitive, which binds to «domain boundaries» and is involved in intra- and inter-chromosomal interactions.

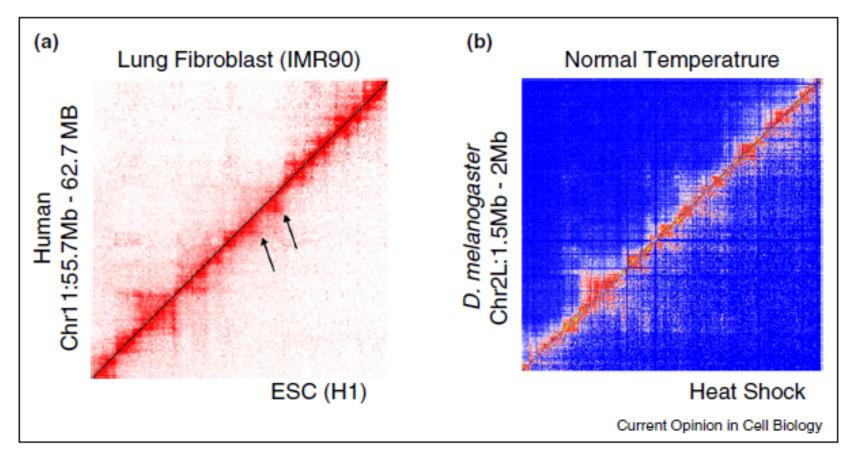
Cohesins mediate looping also in a CTCF-independent fashion, in ESC. In this case, the cohesin loading factor NIPBL and Mediators are found in complex with cohesins at enhancers.





Schematic of putative TAD structures. The central regions of TADs show high levels of chromatin interaction and coincide with the presence of tissue-specific genes and their associated enhancers, the interactions of which with their cognate promoters are facilitated by the presence of **cohesin** and CCCTC-binding factor (**CTCF**). The border regions between TADs are enriched for housekeeping genes, which are often clustered together and generally lack the widely dispersed distal enhancers that are found around tissue-specific genes. The border regions show high levels of CTCF and cohesin binding, although only CTCF seems to prevent interactions between TADs. Pombo & Dillon, 2015

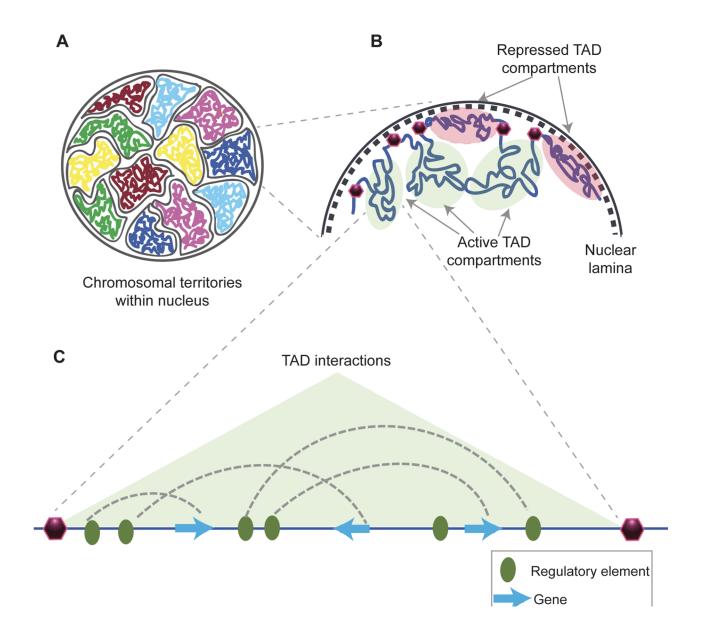
Developmental changes in TADs organization



Changes to domain organization. (a) Hi–C of human embryonic stem cells (H1 ESC – bottom right) compared to lung fibroblast cells (IMR90 – top left) [47]. Arrows indicate TAD structure changes. (b) Hi–C of D. melanogaster under heat shock (bottom right) compared to normal temperature (top left)

Rowley & Corces, 2016

TADs, intra-TADs, Promoter-Enhancer Looping



TADs, intra-TADs, Promoter-Enhancer Looping

