



COLLÈGE DE FRANCE  
—1530—

Collège de France; Chaire:  
Evolution du Développement et des Génomes  
[Denis.Duboule@college-de-france.fr](mailto:Denis.Duboule@college-de-france.fr)

## Cours 2025: Régulation des Gènes du Développement: Les Séquences 'Enhancers'

Leçon 4, vendredi 14 mars 2025



Anais Le Nabec, CIRB, Collège de France

[@denisduboule](https://twitter.com/denisduboule).blsk.social

1

Denis Duboule/2025
Régulation des Gènes du Développement

**Feature Review**  
**Enhancer Logic and Mechanics in Development and Disease**  
 Ryan Rickels<sup>1</sup> and Ali Shilatifard<sup>1\*</sup>  
 Trends in Cell Biology, August 2018, Vol. 28, No. 8 <https://doi.org/10.1016/j.tcb.2018.04.001>  
 © 2018 Published by Elsevier Ltd.

- \*Facteur pionnier, éviction de nucléosomes
- \*Promoteur (Pol II) en mode pause
- \*Looping et contact (cohesin, CTCF..)
- \*Recrutement modificateurs histones (p300)
- \*Recrutement du complexe Médiateur
- \*Démarrage de la Pol II, élongation

NELF: Negative elongation complex  
 SEC: Super elongation complex  
 TF: Transcription factors

H3K4me1/2 (deposited by MLL1/2)  
 H3K27ac (deposited by p300/CBP)  
 Serine 5 phosphorylation (associated with paused Pol II)  
 Serine 2 phosphorylation (associated with elongating Pol II)

2

### Composition (contenu) des enhanceurs

\*Niveau de l'*orthographe*  
(des lettres pour faire des mots, des nucléotides pour faire des motifs)

\*Niveau de la *syntaxe*  
(des mots pour faire des phrases, des motifs ADN pour donner un 'sens' spécifique)

\*Niveau de la *grammaire*  
(intégration dans une dimension fonctionnelle de la phrase/texte, de l'enhancer dans son contexte génomique et cellulaire)

### Topologie des paysages d'enhancers

3

### Composition (contenu) des enhanceurs

\*Niveau de l'*orthographe*  
(des lettres pour faire des mots, des nucléotides pour faire des motifs)

\*Niveau de la *syntaxe*  
(des mots pour faire des phrases, des motifs ADN pour donner un 'sens' spécifique)

\*Niveau de la *grammaire*  
(intégration dans une dimension fonctionnelle de la phrase/texte, de l'enhancer dans son contexte génomique et cellulaire)

4

### Composition (contenu) des enhancers

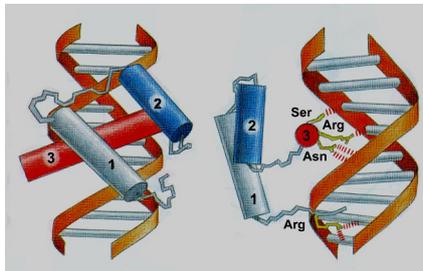
- \*De très nombreuses protéines peuvent lier l'ADN par reconnaissance de petites séquences
- \*Les différentes modalités de fixation à l'ADN par une protéine ne sont pas infinies et sont même relativement restreintes
- \*Certains motifs protéiques sont donc utilisés dans beaucoup de contextes différents
- \*Co-option, réutilisation d'exons...(bricolage..)

### 2 exemples

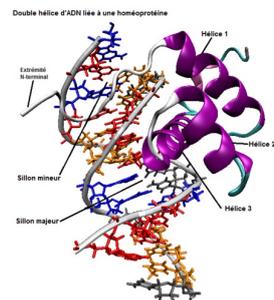
5

### 'Orthographe' des enhancers

Liaison à l'ADN par un motif 'homéodomaine'



- \*La protéine lie l'ADN par son homéodomaine (60 résidus)
- \*L'homéodomaine est replié en trois hélices alpha
- \*Les hélices 2 et 3 forment un motif 'hélice-tour-hélice'
- \*Ce motif reconnaît l'ADN dans le sillon majeur et établit les contacts



[https://fr.wikipedia.org/wiki/Boîte\\_homéotique](https://fr.wikipedia.org/wiki/Boîte_homéotique)



protéine à  
'homéodomaine' X



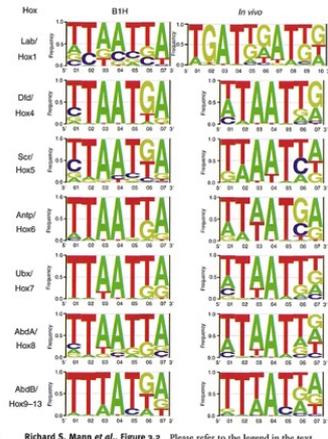
protéine à  
'homéodomaine' Y

6

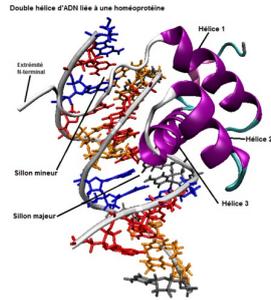


'Orthographe' des enhancers

Liaison à l'ADN par un motif 'homéodomaine' de la même catégorie



Richard S. Mann et al., Figure 3-2. Please refer to the legend in the text.



[https://fr.wikipedia.org/wiki/Boîte\\_homéotique](https://fr.wikipedia.org/wiki/Boîte_homéotique)



protéine à 'homéodomaine' X



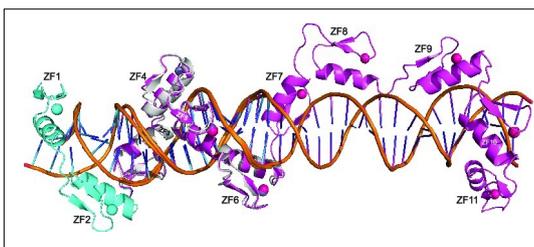
protéine à 'homéodomaine' Y



'Orthographe' des enhancers

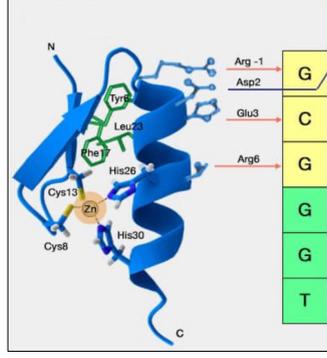
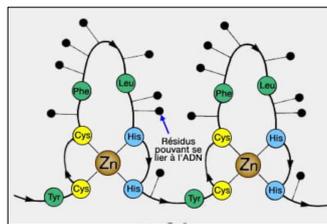
Liaison à l'ADN par des motifs 'doigts de zinc'

\*La protéine CTCF lie l'ADN grâce à 11 'doigts' de zinc. On estime entre 20K et 50K le nombre de sites de liaison à l'ADN dans le génome humain. Le site 'consensus' est: CCGCGNGGNGGCAG. Plusieurs doigts de zinc différents peuvent être engagés dans des contextes différents.



OXFORD UNIVERSITY PRESS Nucleic Acids Res, Volume 51, Issue 16, 8 September 2023, Pages 8447-8462

TFIIIA (9 'doigts', @Vetopsy)



## Composition (contenu) des enhanceurs

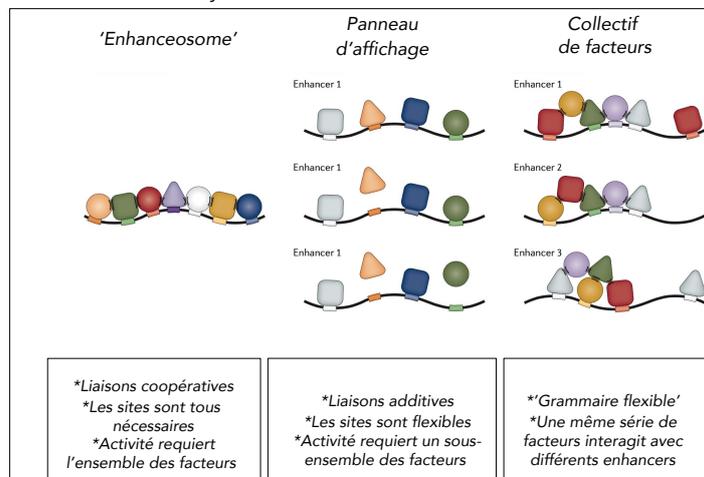
\*Niveau de l'*orthographe*'  
(des lettres pour faire des mots, des nucléotides pour faire des motifs)

\*Niveau de la *syntaxe*  
(des mots pour faire des phrases, des motifs ADN pour donner un 'sens' spécifique)

\*Niveau de la *grammaire*  
(intégration dans une dimension fonctionnelle de la phrase/texte, de l'enhancer dans son contexte génomique et cellulaire)

9

## Syntaxe des enhanceurs

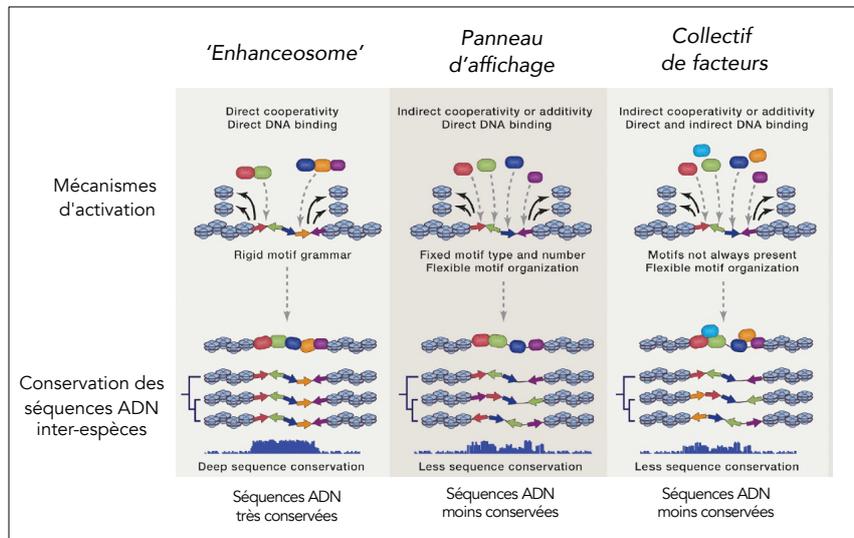


Les différents types de liaisons de facteurs correspondent à des niveaux différents de conservation des séquences cibles...

Modifié de: Spitz and Furlong, Nat. Rev. Genet. 2012

10

## Syntaxe des enhancers



Long, Prescott and Wysocka, Cell, 2016

11

## Composition (contenu) des enhancers

\*Niveau de l'*orthographe*  
(des lettres pour faire des mots, des nucléotides pour faire des motifs)

\*Niveau de la *syntaxe*  
(des mots pour faire des phrases, des motifs ADN pour donner un 'sens' spécifique)

\*Niveau de la *grammaire*  
(intégration dans une dimension fonctionnelle de la phrase/texte, de l'enhancer dans son contexte génomique et cellulaire)

12

## Composition (contenu) des enhancers

\*Y a-t-il un statut général d'enhancer, par exemple des propriétés que l'on retrouverait dans l'ensemble des séquences ADN ayant cette capacité d'activer la transcription?

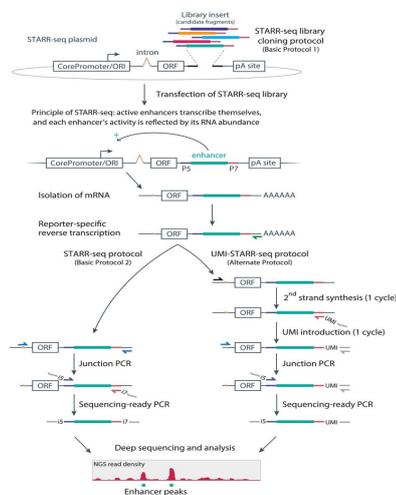
\*Trouve-t-on des enhancers pour différents types de promoteurs, et si oui, ont-ils des caractéristiques différentes? Par exemple entre des promoteurs associés à des 'gènes de ménage' (housekeeping) et des promoteurs associés à des 'gènes de développement' (spécifiques, pleiotropiques etc)?



Mise en place de l'approche STARR-seq en utilisant des promoteurs représentant ces deux catégories afin de voir si des enhancers différents sont sélectionnés

13

## Composition (contenu) des enhancers



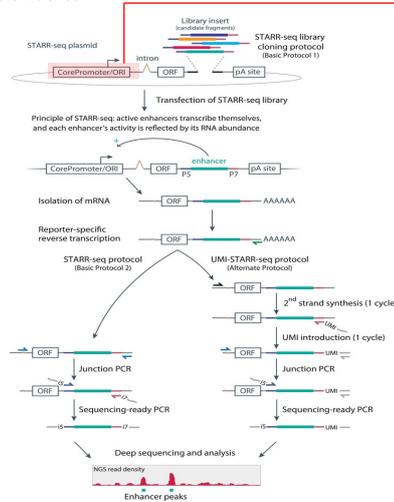
STARR-seq, UMI STARR-seq

UMI: *Identifiant Moléculaire Unique* (rajouté pour les bibliothèques moins complexes...) Marquage unique pour chaque cDNA préparé (par construction). Une sorte de code-barre mais moléculaire plutôt que cellulaire.

14

Syntaxe (règles grammaticales) des enhancers

STARR-seq and UMI-STARR-seq: Assessing Enhancer Activities for Genome-Wide-, High-, and Low-Complexity Candidate Libraries



Utiliser ce protocole avec deux 'core-promoteurs différents. L'un typique des gènes de 'ménage' (housekeeping, hk, ribosomal protein gene RpS12), l'autre typique des gènes du développement (eve).

L'idée étant de voir si l'on repêche des enhancers différents dans leurs séquences, positions etc.. donc de voir si la structure des promoteurs est en partie responsable des spécificités des contacts enhancers-promoteurs.

Dans des cellules de *Drosophila* (S2) en culture

CP Molecular Biology, Volume: 128, Issue: 1, First published: 09 September 2019, DOI: (10.1002/cpmb.105) © Wiley

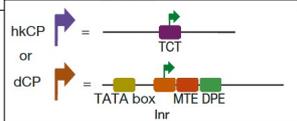
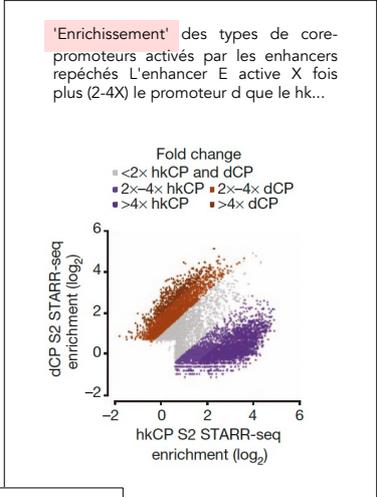
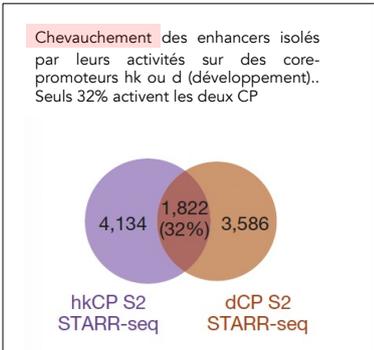
15

LETTER

Enhancer-core-promoter specificity separates developmental and housekeeping gene regulation

Muhammad A. Zaki<sup>1\*</sup>, Cosma D. Arnold<sup>1\*</sup>, Katharina Schenhubel<sup>1</sup>, Michaela Pappal<sup>1</sup>, Martina Roth<sup>1</sup>, Olga Pissal<sup>1</sup>, R. Alexander Stark<sup>1</sup>

354 | NATURE | VOL 518 | 16 FEBRUARY 2015



Ces deux types de promoteurs sont reconnaissables par la présence de petites séquences d'ADN différentes (depuis les années 2000...)

16

Denis Duboule/2023

COLLÈGE DE FRANCE  
—1530—

Régulation des Gènes du Développement

LETTER

doi:10.1038/s41588-023-0094-4

**Enhancer–core–promoter specificity separates developmental and housekeeping gene regulation**

Muhammad A. Zaidi<sup>1\*</sup>, Gomas D. Arnold<sup>1\*</sup>, Katharina Schenhubel<sup>1</sup>, Michael Pagan<sup>1</sup>, Martina Roth<sup>1</sup>, Olga Pissal<sup>1</sup> & Alexander Stark<sup>1</sup>

556 | NATURE | VOL 618 | 26 FEBRUARY 2023

Distribution dans le génome, proche des gènes, loin des gènes, et de quels gènes??

Quels sont les types de gènes impliqués (à proximité)?

Analyse par 'gene ontology' GO terms

GO term	hkCP	dCP
Nucleic acid metabolic process	Low	High
Spliceosomal complex	Low	High
Intracellular transport	Low	High
Cell cycle	Low	High
Cellular component organization or biogenesis	Low	High
Sequence-specific DNA binding RNA pol II TF activity	High	Low
Regulation of signalling	High	Low
Pattern specification process	High	Low
Morphogenesis of an epithelium	High	Low
Anatomical structure morphogenesis	High	Low

\*59% des 'enhancers hk' soit incluent, soit sont proches (moins de 200bp) des sites de démarrage de la transcription (TSS). Donc des régions promotrices peuvent agir comme enhancers.

\*A l'inverse, 56% des 'enhancers d' sont dans des introns de gènes et 27% dans des régions intergéniques.

Cette étude est ensuite étendue à d'autres core-promoteurs des deux catégories et à d'autres types cellulaires avec des résultats comparables. Peut-on prédire l'activité enhancer en ne considérant QUE la séquence d'ADN?

17

Denis Duboule/2023

COLLÈGE DE FRANCE  
—1530—

Régulation des Gènes du Développement

nature genetics ARTICLES

What I cannot create, I do not understand  
(R. Feynman)

**DeepSTARR predicts enhancer activity from DNA sequence and enables the de novo design of synthetic enhancers**

Bernardo P. de Almeida<sup>1,2</sup>, Franziska Reiter<sup>1,2</sup>, Michaela Pagan<sup>1</sup> and Alexander Stark<sup>1,2\*</sup>

NATURE GENETICS | VOL 54 | MAY 2022 | 613–624 | www.nature.com/naturegenetics

Enhancer sequences control gene expression and comprise binding sites (motifs) for different transcription factors (TFs). Despite extensive genetic and computational studies, the relationship between DNA sequence and regulatory activity is poorly understood, and de novo enhancer design has been challenging. Here, we built a deep-learning model, DeepSTARR, to quantitatively predict the activities of thousands of developmental and housekeeping enhancers directly from DNA sequence in *Drosophila melanogaster* S2 cells. The model learned relevant TF motifs and higher-order syntax rules, including functionally nonequivalent instances of the same TF motif that are determined by motif-flanking sequence and intermotif distances. We validated these rules experimentally and demonstrated that they can be generalized to humans by testing more than 40,000 wildtype and mutant *Drosophila* and human enhancers. Finally, we designed and functionally validated synthetic enhancers with desired activities de novo.

Dans cette étude, nous construisons un modèle d'apprentissage profond (DeepSTARR) afin de prédire de façon quantitative les activités enhancers de milliers de séquences à partir des séquences ADN seules (cellules S2 de Drosophile).

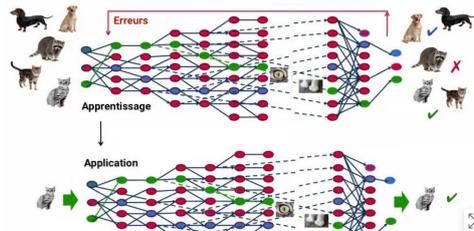
Basé sur ce modèle, nous produisons et validons de nouveaux enhancers synthétiques avec les spécificités souhaitées.

18

## L'apprentissage profond

Cours 2018-2019 du Professeur Stéphane Mallat (CdF; Sciences des données)

Le *deep Learning* s'appuie sur un réseau de **neurones artificiels** s'inspirant du **cerveau** humain. Ce réseau est composé de dizaines voire de centaines de « couches » de **neurones**, chacune recevant et interprétant les informations de la couche précédente. Le système apprendra par exemple à reconnaître les lettres avant de s'attaquer aux mots dans un texte, ou détermine s'il y a un visage sur une photo avant de découvrir de quelle personne il s'agit.



À TRAVERS UN PROCESSUS D'AUTOAPPRENTISSAGE, LE DEEP LEARNING EST CAPABLE D'IDENTIFIER UN CHAT SUR UNE PHOTO. À CHAQUE COUCHE DU RÉSEAU NEURONAL CORRESPOND UN ASPECT PARTICULIER DE L'IMAGE. © MAPR, C.D.

FUTURA

19

## L'apprentissage profond

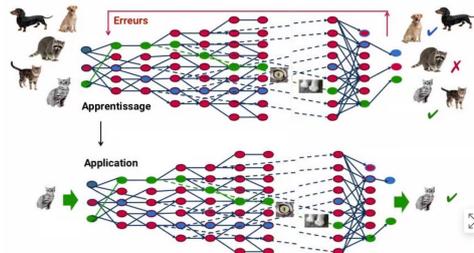
Un réseau neuronal convolusif se compose de deux types de neurones artificiels, agencés en « couches » traitant successivement l'information :

- les neurones de traitement, qui traitent une portion limitée de l'image (appelée « champ réceptif ») au travers d'une fonction de convolution ;
- les neurones de mise en commun des sorties dits de *pooling* (totale ou partielle).

Un traitement correctif non linéaire et ponctuel peut être appliqué entre chaque couche pour améliorer la pertinence du résultat<sup>5</sup>.

L'ensemble des sorties d'une couche de traitement permet de reconstituer une image intermédiaire, qui servira de base à la couche suivante.

wikipedia



À TRAVERS UN PROCESSUS D'AUTOAPPRENTISSAGE, LE DEEP LEARNING EST CAPABLE D'IDENTIFIER UN CHAT SUR UNE PHOTO. À CHAQUE COUCHE DU RÉSEAU NEURONAL CORRESPOND UN ASPECT PARTICULIER DE L'IMAGE. © MAPR, C.D.

FUTURA

20



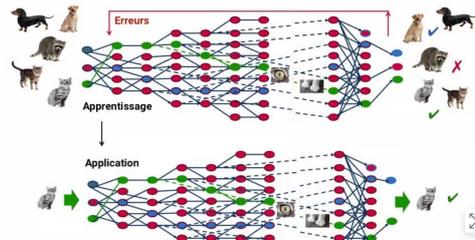
## L'apprentissage profond

### Traitement convolutif [ modifier | modifier le code ]

#### Traitement de profondeur 1 et sans chevauchement (facilitant la compréhension) [ modifier | modifier le code ]

Dans le cadre de la reconnaissance d'image, cette dernière est « pavée », c'est-à-dire découpée en petites zones (appelées tuiles). Chaque tuile sera traitée individuellement par un **neurone artificiel** (qui effectue une opération de filtrage classique en associant un poids à chaque pixel de la tuile). Tous les neurones ont les mêmes paramètres de réglage. Le fait d'avoir le même traitement (mêmes paramètres), sur une petite surface de l'image, légèrement décalé pour chaque champ récepteur, s'appelle une **convolution**<sup>7,8</sup>. Cette strate de neurones avec les mêmes paramètres est appelée « noyau de convolution ».

wikipedia



À TRAVERS UN PROCESSUS D'AUTOAPPRENTISSAGE, LE DEEP LEARNING EST CAPABLE D'IDENTIFIER UN CHAT SUR UNE PHOTO. À CHAQUE COUCHE DU RÉSEAU NEURONAL CORRESPOND UN ASPECT PARTICULIER DE L'IMAGE. © MAPR, C.D.

FUTURA

21



### DeepSTARR predicts enhancer activity from DNA sequence and enables the de novo design of synthetic enhancers

Bernardo P. de Almeida<sup>1,2</sup>, Franziska Reiter<sup>1,2</sup>, Michaela Pagani<sup>1</sup> and Alexander Stark<sup>1,2\*</sup>

NATURE GENETICS | VOL 54 | MAY 2022 | 613–624 | www.nature.com/naturegenetics

'What I cannot create, I do not understand'  
(R. Feynman)

**Enhancer sequences control gene expression and comprise binding sites (motifs) for different transcription factors (TFs). Despite extensive genetic and computational studies, the relationship between DNA sequence and regulatory activity is poorly understood, and de novo enhancer design has been challenging. Here, we built a deep-learning model, DeepSTARR, to quantitatively predict the activities of thousands of developmental and housekeeping enhancers directly from DNA sequence in *Drosophila melanogaster* S2 cells. The model learned relevant TF motifs and higher-order syntax rules, including functionally nonequivalent instances of the same TF motif that are determined by motif-flanking sequence and intermotif distances. We validated these rules experimentally and demonstrated that they can be generalized to humans by testing more than 40,000 wildtype and mutant *Drosophila* and human enhancers. Finally, we designed and functionally validated synthetic enhancers with desired activities de novo.**

Dans cette étude, nous construisons un modèle d'apprentissage profond (DeepSTARR) afin de prédire de façon quantitative les activités enhancers de milliers de séquences à partir des séquences ADN seules (cellules S2 de Drosophile).

Basé sur ce modèle, nous produisons et validons de nouveaux enhancers synthétiques avec les spécificités souhaitées.

22



COLLÈGE DE FRANCE  
—1530—

*Régulation des Gènes du Développement*

nature genetics ARTICLES  
https://doi.org/10.1038/s41588-022-01048-5

**DeepSTARR predicts enhancer activity from DNA sequence and enables the de novo design of synthetic enhancers**

Bernardo P. de Almeida<sup>1,2</sup>, Franziska Reiter<sup>1,2</sup>, Michaela Paganì<sup>1</sup> and Alexander Stark<sup>1,2\*</sup>

NATURE GENETICS | VOL 54 | MAY 2022 | 613–624 | www.nature.com/naturegenetics

### Points importants

<b>DeepSTARR reveals TF motifs required for enhancer activity.</b>	⇒ DeepSTARR révèle quels sont les motifs ADN nécessaires pour induire une activité enhancer.
<b>Nonequivalent instances of the same TF motif.</b>	⇒ Il existe des cas de non-équivalence du même site de liaison pour un facteur de transcription.
<b>Flanking sequence influences the importance of TF motifs.</b>	⇒ Les séquences ADN adjacentes à un site participent à l'importance de l'activité de ce site.
<b>In silico analysis reveals modes of motif cooperativity.</b>	⇒ Ces analyses in silico renseignent sur les façons dont des sites voisins peuvent coopérer.
<b>Motif syntax rules can be generalized to human enhancers.</b>	⇒ Les règles de syntaxe extraites de cette approche peuvent être généralisées (aussi chez les humains).
<b>Designing synthetic enhancers with desired activities.</b>	⇒ Cela ouvre la voie à la fabrication d'enhancers spécifiques entièrement synthétiques

This proof-of-concept experiment shows that the rules learned by DeepSTARR enable the a priori design of synthetic enhancers with desired activity levels.

25

COLLÈGE DE FRANCE  
—1530—

*Régulation des Gènes du Développement*

## Comment les enhancers fonctionnent à grandes distances?

### *Topologie des paysages d'enhancers*

**A. Les holo-enhancers**

**B. Les super-enhancers**

**C. Les archipels de régulation**

**D. Les paysages de régulations pleiotropiques**  
(non exclusifs...)

26

Très souvent, les paysages de régulation contiennent de nombreux enhanceurs qui contribuent à une (ou plusieurs) fonction(s) commune(s)

## An Integrated Holo-Enhancer Unit Defines Tissue and Gene Specificity of the *Fgf8* Regulatory Landscape

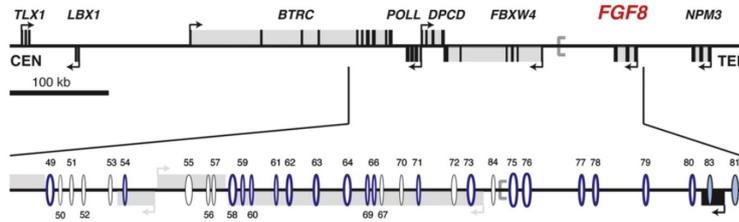
Mirna Marinić,<sup>1</sup> Tugce Aktas,<sup>1</sup> Sandra Ruf,<sup>1</sup> and François Spitz<sup>1,\*</sup>  
<sup>1</sup>Developmental Biology Unit, EMBL, Meyerhofstrasse 1, Heidelberg 69117, Germany

530 Developmental Cell 24, 530–542, March 11, 2013 ©2013 Elsevier Inc.

Developmental Cell  
Article

### A. Les holo-enhancers

Le gène *Fgf8* et sa régulation complexe..



27

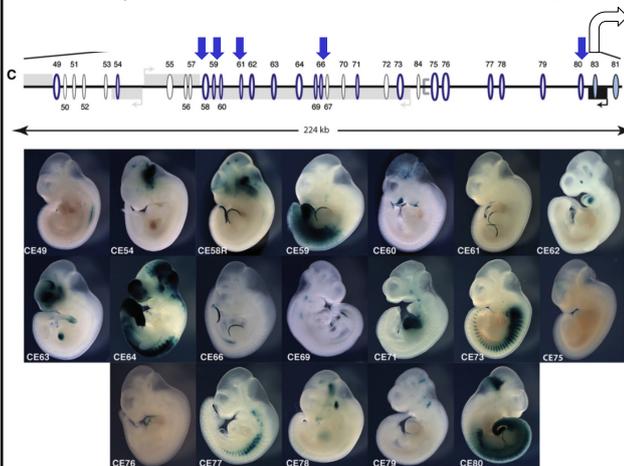
## Le gène *Fgf8* et sa régulation

**Figure 1. Structure and Expression of the *TLX1-NPM3* Interval**

(A) Expression patterns detected by in situ hybridization in E10.5 mouse embryos. Arrows indicate specific expression domains of *Lbx1* and *Fgf8*. AER, apical ectodermal ridge; BA, first branchial arch ectoderm; CP, forebrain commissural plate; MHB, midbrain-hindbrain boundary; MY, myoblasts; NE, nasal pit epithelium; SO, somites; TB, tail bud. See Figure S1 for qPCR data.

AER: Crête apicale du bourgeon de membre

*Fgf8* mRNAs



Comment intégrer ces différentes spécificités pour réguler la transcription aux promoteurs différents?

Scan de l'intervalle ADN par un 'enhancer 'sensor'

28

# Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor



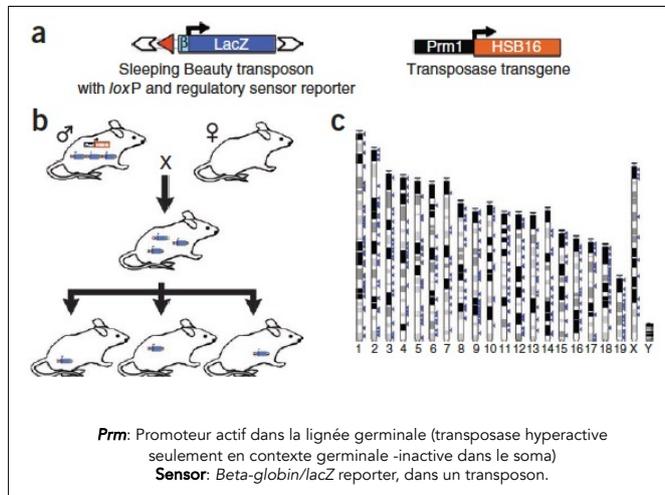
Sandra Ruf<sup>1</sup>, Orsolya Symmons<sup>1</sup>, Veli Vural Uslu<sup>1</sup>, Dirk Dolle<sup>2</sup>, Chloé Hot<sup>1</sup>, Laurence Ettwiller<sup>2</sup> & François Spitz<sup>1</sup>

NATURE GENETICS VOLUME 43 | NUMBER 4 | APRIL 2011

Mobilisation dans la lignée germinale et intégration de une ou X copie(s)

Ségrégation des copies

HOMING



29

# Large-scale analysis of the regulatory architecture of the mouse genome with a transposon-associated sensor



Sandra Ruf<sup>1</sup>, Orsolya Symmons<sup>1</sup>, Veli Vural Uslu<sup>1</sup>, Dirk Dolle<sup>2</sup>, Chloé Hot<sup>1</sup>, Laurence Ettwiller<sup>2</sup> & François Spitz<sup>1</sup>

NATURE GENETICS VOLUME 43 | NUMBER 4 | APRIL 2011

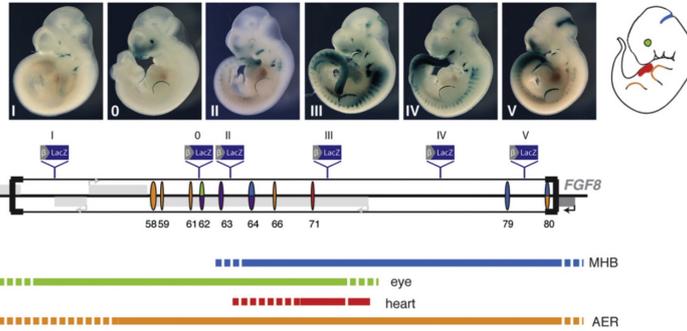
Fréquence de remobilisation des transposons à copie unique

Donor		Chr.	F1	SBlac	Remob <sup>b</sup>	Mapped	Intrachrom <sup>c</sup>	Local <sup>d</sup>
Insertion	Type							
177184c	Het	5	71	25	9	9	2	1
183038	Het	7	52	33	7	4	0	0
183036	Het	1	237	90	44	37	12	5
176599b	Het	2	104	46	11	10	4	3
176599bc <sup>a</sup>	Het	1, 2	53	41	19	17	7	2
176599b	Hom	2	61	57	12	11	2	1
176148b	Hom	16	24	22	7	3	0	0
178235	Hom	16	15	15	3	3	0	0
183041a	Het	3	261	106	41	24	5	4
178137a	Het	11	145	60	11	8	2	2
<b>Total</b>			<b>1,023</b>	<b>495</b>	<b>164 (33%)</b>	<b>126</b>	<b>34 (27%)</b>	<b>18 (14%)</b>

Chr., chromosome; het, heterozygous; hom, homozygous; remob, remobilized; intrachrom, intrachromosomal.  
<sup>a</sup>Male with two unlinked insertions (176599b and 176599c). The number of remobilizations is an underestimate as only F1 individuals, which were negative for both 176599b and 176599c, were considered as 'remobilized'.  
<sup>b</sup>Number of animals with a SBlac transposon in the genome but not at the starting position.  
<sup>c</sup>New insertions on the same chromosome than the donor site.  
<sup>d</sup>New insertions within 2 Mb of the position of the donor site.

↳ 'Homing'

30



\*Le transposon sensor intègre différentes influences de régulations selon son site d'insertion, sans présenter de logique particulière qui serait 'calculable' ou 'prévisible' (avec les moyens actuels...)

Marinic et al., *Developmental Cell*, 2013

31

### An Integrated Holo-Enhancer Unit Defines Tissue and Gene Specificity of the *Fgf8* Regulatory Landscape

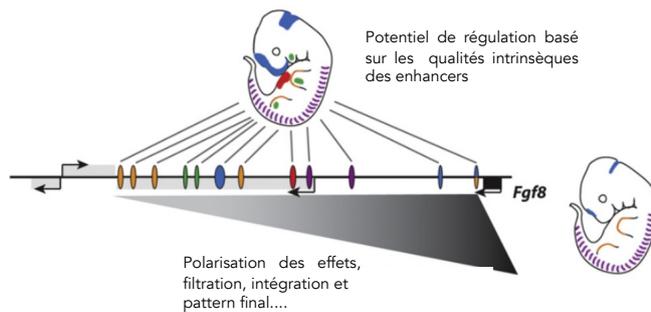
Mirna Marinić,<sup>1</sup> Tugce Aktas,<sup>1</sup> Sandra Ruf,<sup>1</sup> and François Spitz<sup>1,\*</sup>  
<sup>1</sup>Developmental Biology Unit, EMBL, Meyerhofstrasse 1, Heidelberg 69117, Germany

530 *Developmental Cell* 24, 530–542, March 11, 2013 ©2013 Elsevier Inc.

#### Figure 7. A Holo-Enhancer Region

The *Fgf8* gene-dense region comprises multiple enhancers (ovals) distributed within introns of the bystander genes (gray rectangles). Enhancers are color-coded to match the expression domains depicted on the mouse embryo outlines.

(A) Although individual enhancers possess large and diverse intrinsic autonomous activities (upper embryo), their collective output is not equal to their superposition: only a subset of these activities is displayed over distance to endogenous genes and polarized to the genomic position normally occupied by *Fgf8* (lower-right embryo).



32



Comment les enhancers fonctionnent à grandes distances?

\*Au niveau moléculaire?

\*Au niveau du 'choix' des gènes cibles et de l'organisation chromosomique?

- A. Les holo-enhancers
- B. Les super-enhancers**
- C. Les archipels de régulation
- D. Les paysages de régulations pleiotropiques (non exclusifs...)



Les 'super-enhancers'

Cellules ES

**B. Les super-enhancers I**

**Super-Enhancers in the Control of Cell Identity and Disease**

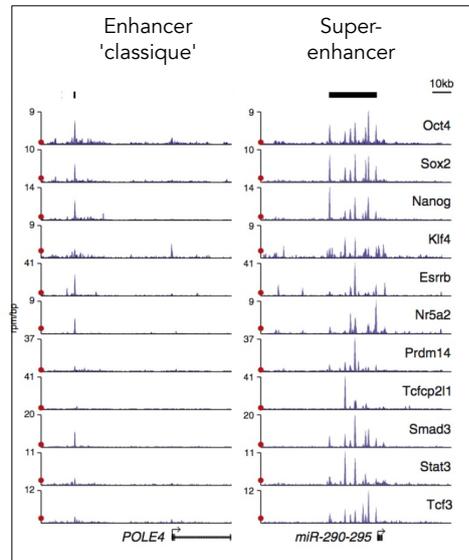
Denis Hnisz,<sup>1,2</sup> Brian J. Abraham,<sup>1,2</sup> Tong Ihn Lee,<sup>1,2</sup> Ashley Lau,<sup>1,2</sup> Violaine Saint-André,<sup>1</sup> Alla A. Sigova,<sup>1</sup> Heather A. Hoke,<sup>1,2</sup> and Richard A. Young<sup>1,2\*</sup>

<sup>1</sup>Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

<sup>2</sup>Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Cell 155, 934–947, November 7, 2013 ©2013 Elsevier Inc.

Les super-enhancers sont des clusters d'enhancers qui ensemble assemblent une haute densité de l'appareil transcriptionnel afin de produire une expression robuste de gènes cibles qui ont des rôles essentiels dans l'identité cellulaire.



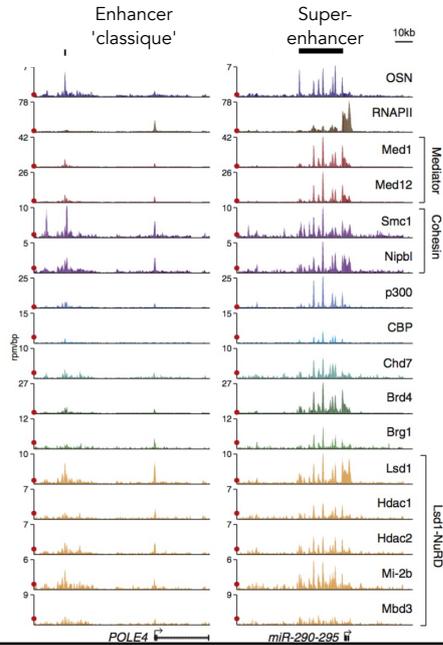
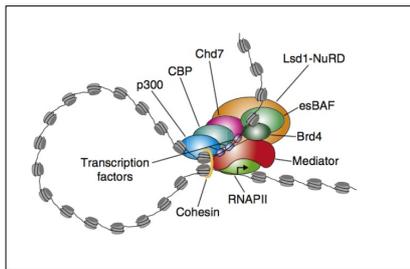


### Super-Enhancers in the Control of Cell Identity and Disease

Denes Hnisz,<sup>1,2</sup> Brian J. Abraham,<sup>1,2</sup> Tong Ihn Lee,<sup>1,2</sup> Ashley Lau,<sup>1,2</sup> Violaine Saint-André,<sup>1</sup> Alla A. Sigova,<sup>1</sup> Heather A. Hoke,<sup>1,2</sup> and Richard A. Young<sup>1,2\*</sup>

Cell 155, 934–947, November 7, 2013 ©2013 Elsevier Inc.

\*Les composants de l'appareil transcriptionnel sont présents sur les super-enhancers, ainsi que les complexes cohésines et les modificateurs de la chromatine



35



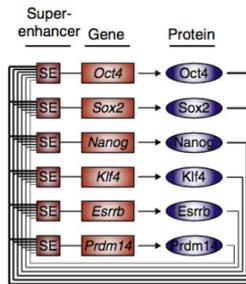
### Super-Enhancers in the Control of Cell Identity and Disease

Denes Hnisz,<sup>1,2</sup> Brian J. Abraham,<sup>1,2</sup> Tong Ihn Lee,<sup>1,2</sup> Ashley Lau,<sup>1,2</sup> Violaine Saint-André,<sup>1</sup> Alla A. Sigova,<sup>1</sup> Heather A. Hoke,<sup>1,2</sup> and Richard A. Young<sup>1,2\*</sup>

Cell 155, 934–947, November 7, 2013 ©2013 Elsevier Inc.

**SUMMARY**

Super-enhancers are large clusters of transcriptional enhancers that drive expression of genes that define cell identity. Improved understanding of the roles that super-enhancers play in biology would be afforded by knowing the constellation of factors that constitute these domains and by identifying super-enhancers across the spectrum of human cell types. We describe here the population of transcription factors, cofactors, chromatin regulators, and transcription apparatus occupying super-enhancers in embryonic stem cells and evidence that super-enhancers are highly transcribed. We produce a catalog of super-enhancers in a broad range of human cell types and find that super-enhancers associate with genes that control and define the biology of these cells. Interestingly, disease-associated variation is especially enriched in the super-enhancers of disease-relevant cell types. Furthermore, we find that cancer cells generate super-enhancers at oncogenes and other genes important in tumor pathogenesis. Thus, super-enhancers play key roles in human cell identity in health and in disease.



Cellules ES (feedback)

- Les super-enhancers stimulent un haut niveau de transcription
- Les super-enhancers sont associés à des gènes impliqués dans la différenciation cellulaire (linéage..)
- Des variations dans des super-enhancers peuvent être associées à des pathologies (cancer)

36

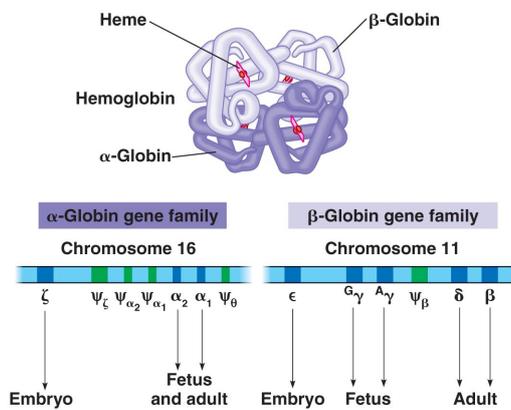
**B. Les super-enhancers II**

Un exemple (d'importance historique) d'une situation dans laquelle une protéine unique doit être produite en quantité suffisante pour assurer une fonction différenciée précise, en fin de parcours de différenciation (albumine, cristalline...). Pas de pléiotropie, mais une nécessité d'une activité soutenue et de longue durée.

37

**App 1Mio de globules rouges produites à chaque seconde chez l'adulte, contenant de l'hémoglobine**

The globin gene (blue) family (pseudo-genes are shown in green). The alpha and beta-globin gene clusters are on 2 different chromosomes (Hème: ion Fe complexé à une porphyrine- un par chaîne de globine-)



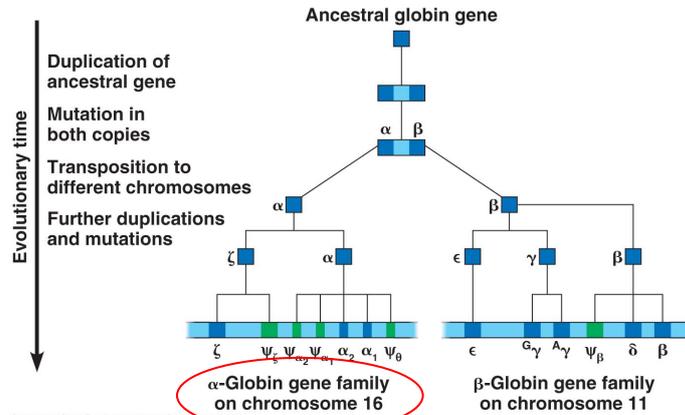
(b) The human  $\alpha$ -globin and  $\beta$ -globin gene families

Copyright © 2008 Pearson Education, Inc., publishing as Pearson Benjamin Cummings.

Campbell 8<sup>th</sup> ch 21

38

Human alpha and beta-globin gene clusters on 2 different chromosomes  
Evolution of these families, common ancestor. Pourquoi? Quel avantage sélectif?  
Evolve various affinities for O<sub>2</sub> (embryonic or fetal forms)



Copyright © 2008 Pearson Education, Inc., publishing as Pearson Benjamin Cummings

Campbell 8<sup>th</sup> ch 21

39

Received: 8 March 2023 | Revised: 3 May 2023 | Accepted: 5 May 2023  
DOI: 10.1002/Dev.2520047

**PROBLEMS & PARADIGMS**  
Prospects & Overviews

**Understanding fundamental principles of enhancer biology at a model locus**  
Analysing the structure and function of an enhancer cluster at the  $\alpha$ -globin locus

Mira Kassouf | Seren Ford | Joseph Blayney | Doug Higgs

BioEssays

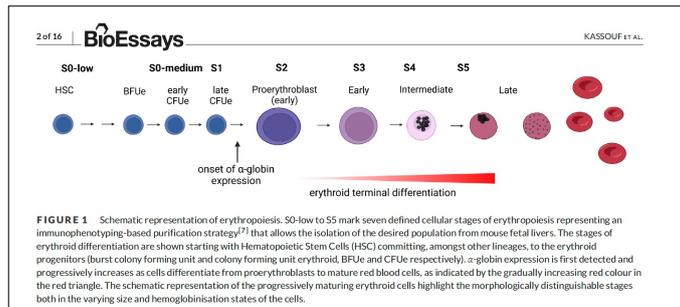
L'érythropoïèse dans les cellules de foie fœtales

Érythropoïèse

L'érythropoïèse est, chez les êtres humains, l'ensemble des processus de production des globules rouges dans la moelle osseuse. Les globules rouges assurent le transport de l'oxygène des poumons vers les organes et du dioxyde de carbone des organes vers les poumons.

Source: Wikipédia

- \*Matériel homogène et en quantité
- \*Bonne orthologie chez la souris
- \*Aspect médical (Thalassémies...-anémie-)
- \*Donc, position historique pour les enhancers...(beta-globine...)
- \*Régulation à distance, LCR..



40

COLLÈGE DE FRANCE  
—1530—

Régulation des Gènes du Développement

Received: 8 March 2023 | Revised: 3 May 2023 | Accepted: 5 May 2023  
DOI: 10.1002/bes.20230047

BioEssays

**PROBLEMS & PARADIGMS**  
Prospects & Overviews

**Understanding fundamental principles of enhancer biology at a model locus**  
Analysing the structure and function of an enhancer cluster at the  $\alpha$ -globin locus

Mira Kassouf | Seren Ford | Joseph Blayney | Doug Higgs

**\*Petit 'TAD' de 65kb, inclut dans un 'TAD' de 165kb (taille réduite due à la 'stratégie'?)**  
**\*Délimité par des sites CTCF**

**\*Un gène embryonnaire (epsilon, Hba-x)**  
**\*Une paire de gènes alpha similaires (Hba-1 et Hba-2)**

**\*Deux autres gènes globines de fonction inconnue**

**\*5 éléments de régulation (R1, R2, R3, Rm et R4) localisés entre -8 et -40kb, avec R1, R2, R3 et Rm localisés dans les introns du gène *Npr3* (housekeeping..).**

**\*Les enhancers contiennent des sites de liaison aux facteurs GATA1, TAL1, NFE2, KLF1, tous impliqués dans l'érythropoïèse...**

**\*Structure en 'Super-enhancer'**

**Structure du locus de l'alpha-globine**

41

COLLÈGE DE FRANCE  
—1530—

Régulation des Gènes du Développement

Received: 8 March 2023 | Revised: 3 May 2023 | Accepted: 5 May 2023  
DOI: 10.1002/bes.20230047

BioEssays

**PROBLEMS & PARADIGMS**  
Prospects & Overviews

**Understanding fundamental principles of enhancer biology at a model locus**  
Analysing the structure and function of an enhancer cluster at the  $\alpha$ -globin locus

Mira Kassouf | Seren Ford | Joseph Blayney | Doug Higgs

**Dynamique de la transcription: Le système *in vitro* permet de suivre les étapes de préparation et de démarrage de la transcription**

**PIC: Pre-initiation complex**

**GATA2 pionnier**

**Onset of  $\alpha$  globin expression**

**Recruitment of PIC, Pol II, and KLF1**

**Noisy transcriptional bursting** (S0, S1) → **Continuous transcriptional bursting** (S2, S3) → **Noisy transcriptional bursting** (S4)

**Legend:**  
 ▲ convergent CTCF  
 ■ enhancer  
 □ promoter  
 — transcription  
 — cohesin (mediating loop extrusion)  
 ● transcriptional pre-initiation complex

42



Received: 8 March 2023 | Revised: 3 May 2023 | Accepted: 5 May 2023  
 DOI: 10.1002/bes.202300047

**PROBLEMS & PARADIGMS**  
 Prospects & Overviews

**BioEssays**

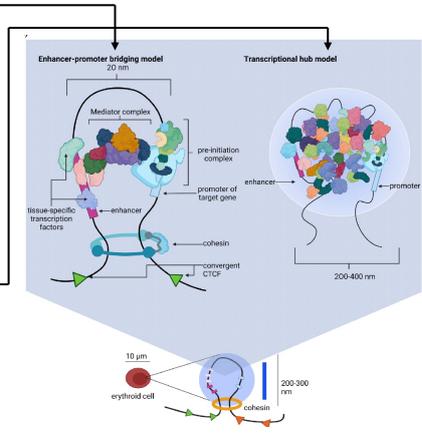
**Understanding fundamental principles of enhancer biology at a model locus**  
 Analysing the structure and function of an enhancer cluster at the  $\alpha$ -globin locus

Mira Kassouf | Seren Ford | Joseph Blayney | Doug Higgs

### Des modèles d'interactions différents

\*Modèle de 'bridging' (loop extrusion assure la proximité, créant un micro-domaine chromatinien, le médiateur fait le bridging).

\*Modèle du 'transcription hub', dans lequel la proximité est aussi assurée par loop extrusion mais une haute concentration de facteurs est observée, créant peut-être une séparation de phase et donc un environnement 'enrichi'. C'est donc un modèle moins déterministe, plus 'stochastique', qui peut expliquer certains résultats obtenus avec des délétions ponctuelles d'enhancers. Ce modèle est plus compatible avec des paysages d'holo-enhancers (par ex. *Fgf8* dans les membres) qu'avec des paysages additifs du genre *Shox2*.



43



Received: 8 March 2023 | Revised: 3 May 2023 | Accepted: 5 May 2023  
 DOI: 10.1002/bes.202300047

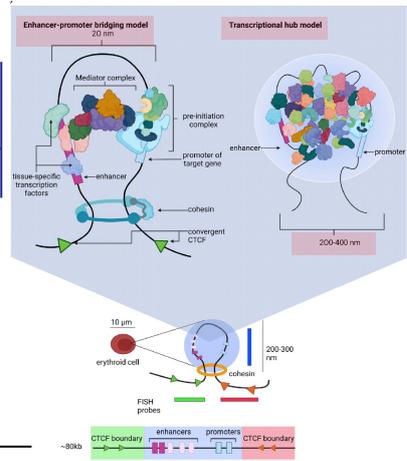
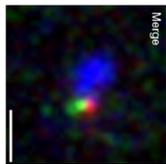
**PROBLEMS & PARADIGMS**  
 Prospects & Overviews

**BioEssays**

**Understanding fundamental principles of enhancer biology at a model locus**  
 Analysing the structure and function of an enhancer cluster at the  $\alpha$ -globin locus

Mira Kassouf | Seren Ford | Joseph Blayney | Doug Higgs

### Des modèles d'interactions différents, une structure générale visualisée par microscopie à super-résolution (nanometer scale..PN 2014).



44



Received: 8 March 2023 | Revised: 3 May 2023 | Accepted: 5 May 2023  
DOI: 10.1002/lev.202300047

PROBLEMS & PARADIGMS  
Prospects & Overviews

BioEssays

**Understanding fundamental principles of enhancer biology at a model locus**

Analysing the structure and function of an enhancer cluster at the  $\alpha$ -globin locus

Mira Kassouf | Seren Ford | Joseph Blayney | Doug Higgs

**Approches fonctionnelles du super-enhancer**

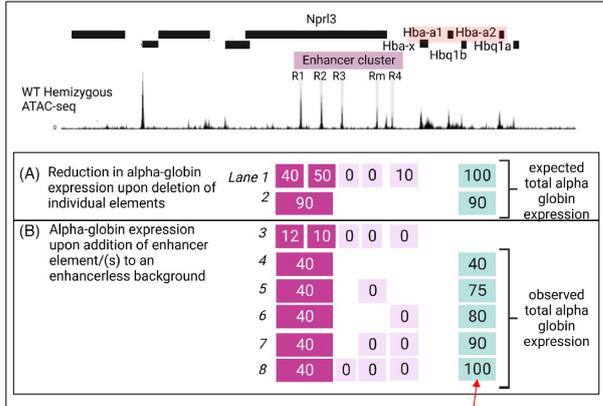
- \*Délétions variées
- \*Rajouts sur une délétion totale

\*Les enhancers n'ont pas la même importance.

\*Différence dans le résultat selon que les enhancers sont enlevés ou rajoutés... en particulier sur l'importance de R3, Rm et R4.

\*En effet, peu importants lorsqu'ils sont enlevés par unité, ils se révèlent critiques lorsque rajoutés à R1 et R2 (comparer les lignes 2 et 3 avec la ligne 8..).

\*Importance de la topologie des enhancers au sein du SE, pour induire une architecture de régulation



Fonction complète même si R3, Rm et R4 sont inactifs seuls....