

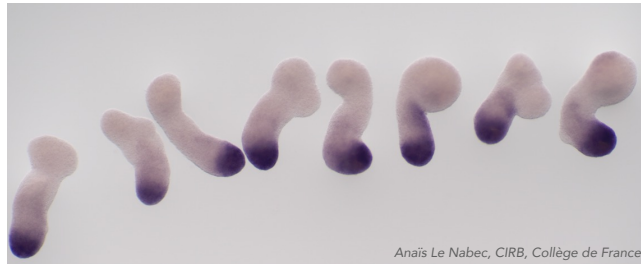


COLLÈGE DE FRANCE  
—1530—

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 2, vendredi 28 février 2025



Anais Le Nabec, CIRB, Collège de France

@denisduboule.blsk.social

1

Denis Duboule/2025



COLLÈGE DE FRANCE  
—1530—

Régulation des Gènes du Développement

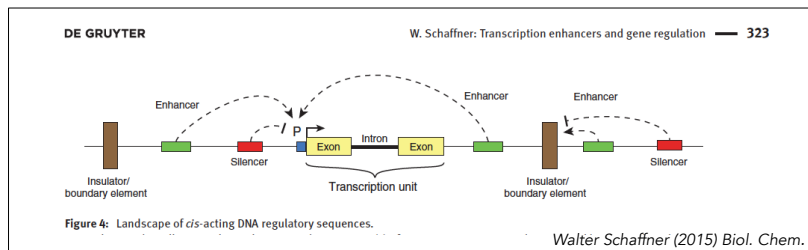
### Un bref historique (1981)

#### Expression of a $\beta$ -Globin Gene Is Enhanced by Remote SV40 DNA Sequences

Julian Banerji, Sandro Rusconi and  
Walter Schaffner  
Institut für Molekularbiologie II  
Universität Zürich  
Hönggerberg, CH-8093 Zürich, Switzerland

\*Les bases générales du concept d'enhancer' sont énoncées au début des années 1980

\*Une séquence d'ADN qui peut agir à distance pour augmenter la transcription à partir d'un promoteur cible, et cela indépendamment de son orientation



Fonction des enhancers: Ce sont des **interrupteurs** qui décident du temps et de l'endroit où un gène particulier est activé

2



## Approches pour identifier des enhanceurs

\*Comment trouver et valider fonctionnellement des enhanceurs?

Historiquement:

- 1) Comparaisons de séquences
- 2) Pièges à enhanceurs
- 3) Marques épigénétiques
- 4) Accessibilité de la chromatine
- 5) Contacts inter-chromatine
- 6) .....

3



## \*Piège à enhanceur, trappe à gène (*Drosophila*)

\*Piéger des enhanceurs en utilisant leur fonction comme signal de leur présence..

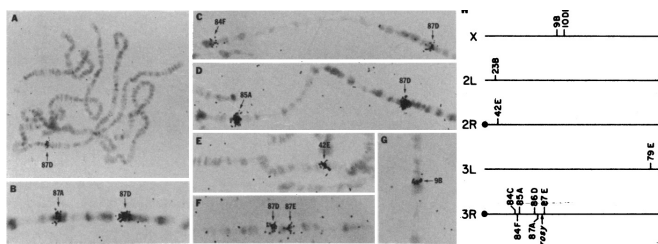
\*Un système basé sur le développement de la transgénèse chez les mouches

### Genetic Transformation of *Drosophila* with Transposable Element Vectors

Gerald M. Rubin and Allan C. Spradling

SCIENCE, VOL. 218, 22 OCTOBER 1982

**Summary.** Exogenous DNA sequences were introduced into the *Drosophila* germ line. A *rosy* transposon (*ry1*), constructed by inserting a chromosomal DNA fragment containing the wild-type *rosy* gene into a P transposable element, transformed germ line cells in 20 to 50 percent of the injected *rosy* mutant embryos. Transformants contained one or two copies of chromosomally integrated, intact *ry1* that were stably inherited in subsequent generations. These transformed flies had wild-type eye color indicating that the visible genetic defect in the host strain could be fully and permanently corrected by the transferred gene. To demonstrate the generality of this approach, a DNA segment that does not confer a recognizable phenotype on recipients was also transferred into germ line chromosomes.



Introduction du gène *rosy* qui est normalement localisé en 87D sur le chromosome 3 de la mouche. Visualisation directe des copies sommatrices grâce aux chromosomes polytènes des glandes salivaires.

4

\*Piège à enhancer, piège à gène (*Drosophila*)

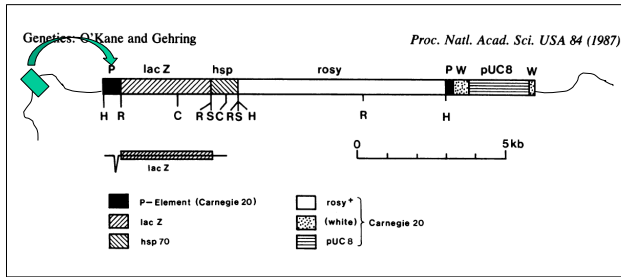
**Detection *in situ* of genomic regulatory elements in *Drosophila***

(position effects/P transposon/enhancers/ $\beta$ -galactosidase/cell-type markers)

CAHIR J. O'KANE AND WALTER J. GEHRING

Department of Cell Biology, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland

*Proc. Natl. Acad. Sci. USA*  
Vol. 84, pp. 9123-9127, December 1987  
Genetics

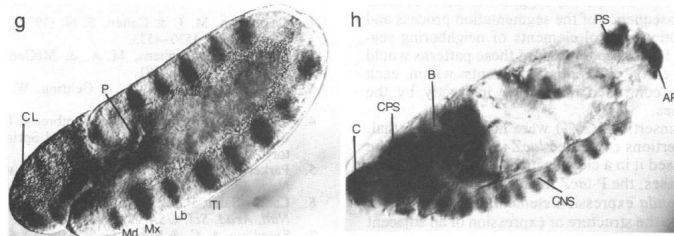


**ABSTRACT** We have developed an approach for the *in situ* detection of genomic elements that regulate transcription in *Drosophila melanogaster*. The approach is analogous to a powerful method of bacterial genetics, the random generation of operon fusions, that enables the isolation and characterization of genes simply by knowing or postulating their pattern of expression; it is not necessary initially to screen for mutant phenotypes. To apply this approach to *Drosophila*, we have used the expression of the *lacZ* gene of *Escherichia coli* from the P-element promoter in germ-line transformant flies to screen for chromosomal elements that can act at a distance to stimulate expression from this apparently weak promoter. Of 49 transformed fly lines obtained, ~70% show some type of spatially regulated expression of the *lacZ* gene in embryos; many of these express *lacZ* specifically in the nervous system. The P-*lacZ* fusion gene is, therefore, an efficient tool for the recovery of elements that may regulate gene expression in *Drosophila* and for the generation of a wide variety of cell-type-specific markers.

■ Enhancer (piégé) ayant une spécificité de 'demi-segment postérieur'

5

\*Piège à enhancer, piège à gène (*Drosophila*)

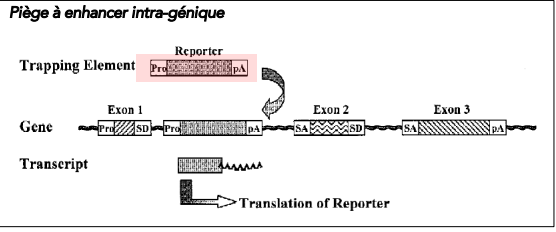


in the region of the cephalic sensory organs and in a number of cells posteriorly that we have not been able to identify. (g and h) Embryos carrying insertion 5 (cytological map position 56F) at stages 11 and 17, respectively.  $\beta$ -Galactosidase expression begins at stage 10 in segmentally repeated ventral ectodermal stripes. The stripes show a two-segment periodicity in the strength of expression and are in the posterior part of each segment. Other sites of expression at this stage are in the proctodeum, clypeolabrum, and the procephalic neurogenic region (not in focus). In later embryos

\*Détection d'un enhancer actif dans les parties postérieures de chaque segment

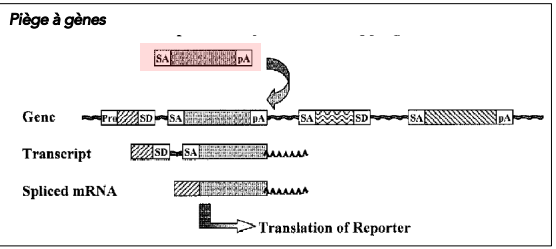
6

**\*Piège à enhancer, piège à gène (Drosophila)**



Durik et al., Genome Research, 1999

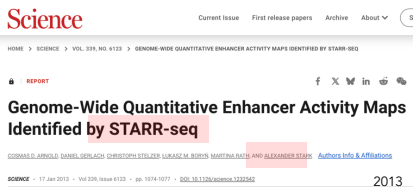
Un vecteur **enhancer-trap** peut également s'intégrer à l'intérieur d'introns (dans un gène). Dans ce cas, l'expression du reporter sera celle du gène et le gène sera interrompu (mutant hypomorphe?)



Un vecteur **enhancer-trap** modifié peut devenir une trappe à gène car un site accepteur d'épissage permet de remonter en séquence vers l'identité du gène en question grâce à un ARN chimérique.

Pro: Promoteur; pA: site de polyadénylation; SD: site donneur d'épissage; SA: site accepteur d'épissage

**\*Approches fonctionnelles à haut débit (Drosophila)**  
**Exemple: STARR-seq**



**WIKIPEDIA**  
The Free Encyclopedia

**STARR-seq**

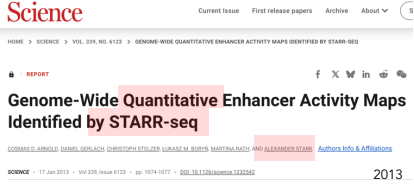
**STARR-seq** (short for **self-transcribing active regulatory region sequencing**) is a method to assay enhancer activity for millions of candidates from arbitrary sources of DNA. It is used to identify the sequences that act as transcriptional enhancers in a direct, quantitative, and genome-wide manner.<sup>[1]</sup>



Une méthode pour tester l'activité enhancer potentielle de millions de séquences de façon aléatoire et à partir de n'importe quelle source d'ADN (fragment, BAC, génome entier...)



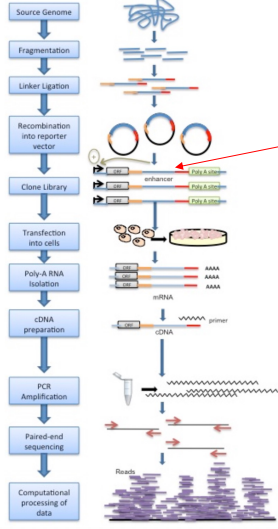
\*Approches fonctionnelles à haut débit (Drosophila) Exemple: STARR-seq (quantitatif)



**WIKIPEDIA**  
The Free Encyclopedia

**STARR-seq**

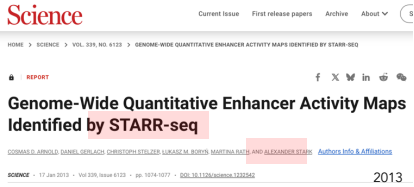
**STARR-seq** (short for **self-transcribing active regulatory region sequencing**) is a method to assay enhancer activity for millions of candidates from arbitrary sources of DNA. It is used to identify the sequences that act as transcriptional enhancers in a direct, quantitative, and genome-wide manner.<sup>[1]</sup>



- \*Source d'ADN (génomme)
- \*Fragmentation
- \*Addition de linkers
- \*Recombinaison dans un plasmide reporter (3' UTR) et préparation d'une librairie
- \*Transfection dans des cellules
- \*Isoler les mRNAs polyA+
- \*Production des cDNAs
- \*Amplification PCR
- \*Séquençage ADN
- \*Alignement sur le génome de départ



\*Approches fonctionnelles à haut débit (Drosophila) Exemple: STARR-seq



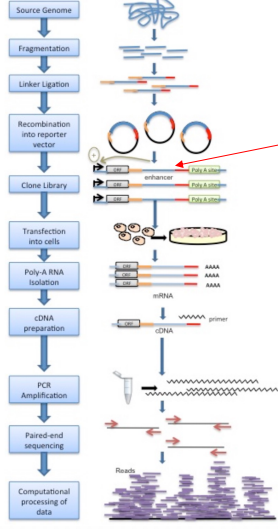
\*L'enhancer est 'auto-transcrit' car intégré dans le 3' UTR d'un gène rapporteur.

\*La quantité de cDNA pour un enhancer (nombre de reads) reflétera directement son activité (quantitatif)

\*Avantages d'une approche à haut débit, exhaustive et sans biais initiaux.

\*Inconvénients du système de détection car le choix des cellules va déterminer quelles séquences seront actives...

Les approches épigénétiques sont donc plus complètes mais elles n'ont pas de read out vraiment fonctionnel et/ou quantitatif..

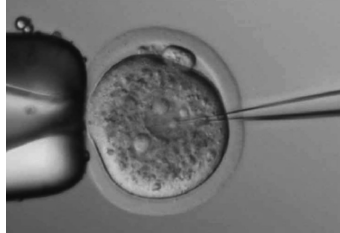


- \*Source d'ADN (génomme)
- \*Fragmentation
- \*Addition de linkers
- \*Recombinaison dans un plasmide reporter et préparation d'une librairie
- \*Transfection dans des cellules
- \*Isoler les mRNAs polyA+
- \*Production des cDNAs
- \*Amplification PCR
- \*Séquençage ADN
- \*Alignement sur le génome de départ



### \*Détection d'activité enhancer chez les mammifères (souris)

R<sup>6</sup> ResearchGate



\*Chez la souris, approche par injection de transgènes dans le zygote....

\*Procédure coûteuse et technique, mais relativement rapide (si pas de lignée) et efficace

\*Extraction de zygotes, injection, réimplantation dans des mères porteuses, établissement de lignées ou analyses des fœtus...

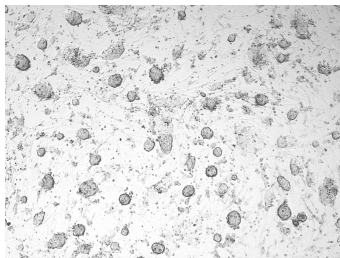
Mais! Une limitation importante de cette approche se trouve dans le côté aléatoire du site d'insertion et du grand nombre de copies généralement introduites...ce qui peut fausser le résultat de façon quantitative mais aussi (parfois) de façon qualitative. Les outils CRISPR/Cas pallient en partie ces problèmes par une grande précision d'insertion et de localisation. Mais cela reste une approche compliquée.

Comment maîtriser ces deux paramètres?

11



### \*Détection d'activité enhancer chez les mammifères (souris)



\*Utilisation des cellules ES ou iPS

\*Introduction de transgènes en utilisant des systèmes d'aéroports déterminés (landing sites) permettant de contrôler à la fois le site d'insertion et le nombre de copies insérées (une, si possible).

\*Analyse précise et comparative (le contexte est toujours le même).

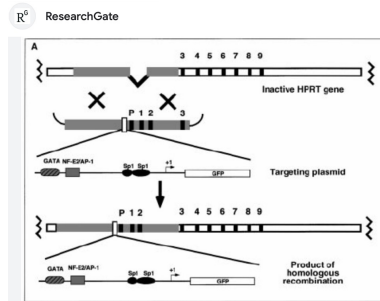
\*Mais! Nécessité de remettre ces cellules ES dans des conditions physiologiques permettant une approche fonctionnelle...(embryon, tissus, type cellulaire..)

\*Cela permet d'abord un travail en amont sur ces cellules et, une fois les conditions d'insertion maîtrisées et sélectionnées, le retour de ces cellules dans un contexte 'développemental', qui peut aller jusqu'à l'embryon lui-même.

12

\*Deux exemples de 'landing sites' pour transgènes en cellules ES

**Le locus *Hprt* muté et le locus *H11***  
(hypoxanthine-guanine phosphoribosyltransférase)



\*Le locus muté *Hprt* et le milieu de sélection HAT

\*L'Aminoptérine inhibe la synthèse d'ADN et donc les cellules meurent. En présence de Thymidine et d'Hypoxanthine, l'enzyme *Hprt* peut toutefois produire les éléments nécessaires (H et T sont des intermédiaires dans la synthèse des bases puriques A et G) et donc activer une 'salvage pathway'.

\*La bonne recombinaison (au locus) va corriger le gène *Hprt* muté et donc permettre le sauvetage de la cellule.

- \*Procédure très efficace et sélective (permet de sélectionner un évènement à basse fréquence)
- \*Procédure agressive pour les cellules..(drogue).

13

**Function-based identification of mammalian enhancers using site-specific integration**

Diane E Dickel<sup>1</sup>, Yiwen Zhu<sup>1</sup>, Alex S Nord<sup>1</sup>, John N Wylie<sup>2,3</sup>, Jennifer A Akiyama<sup>1</sup>, Veena Afzal<sup>1</sup>, Ingrid Plajzer-Frick<sup>1</sup>, Aileen Kirkpatrick<sup>4,5</sup>, Berthold Göttgens<sup>4,5</sup>, Benoit G Bruneau<sup>2,3,6,7</sup>, Axel Visel<sup>1,8,9</sup> & Len A Pennacchio<sup>1,8</sup>

**NATURE METHODS** | ADVANCE ONLINE PUBLICATION | 1 (2014)

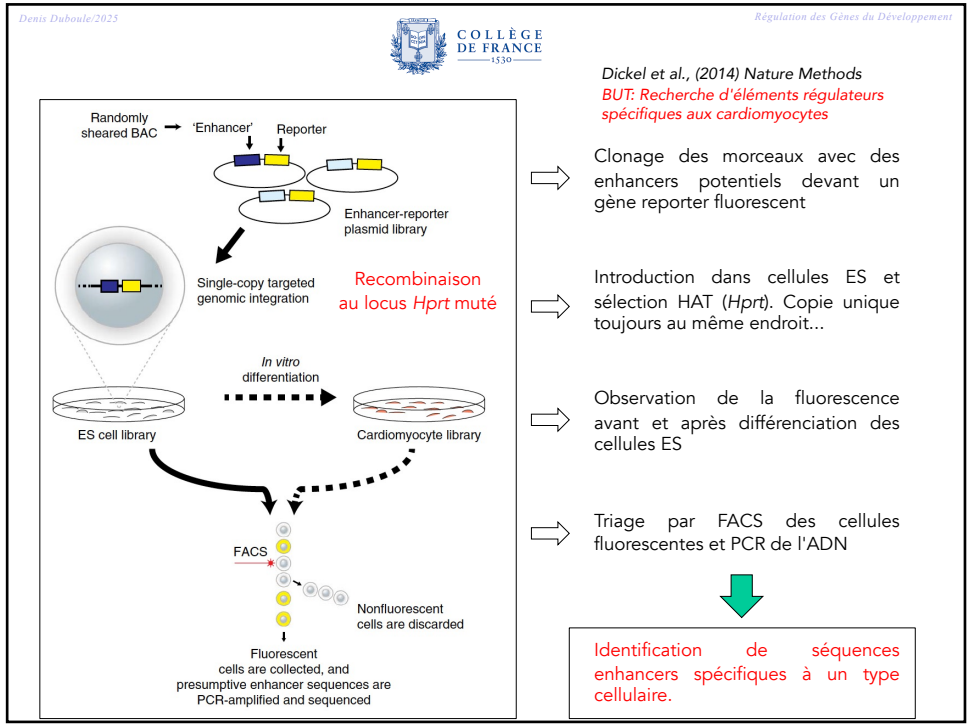
'piège à enhancers' ciblé, fonctionnel et à haut-débit

*Principe général:*

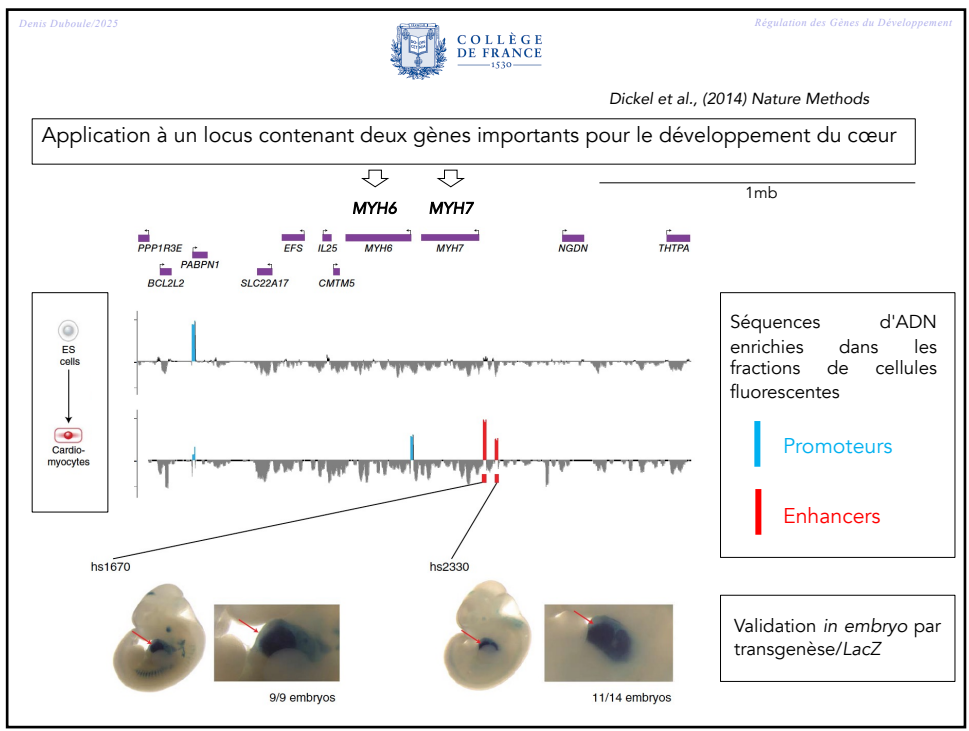
- \*Identifier un segment d'ADN d'intérêt (par exemple un BAC –*Bacterial Artificial Chromosome*- contenant un gène particulier).
- \*Fragmenter cet ADN et le cloner en amont d'un gène reporter (protéine fluorescente), dans un vecteur de recombinaison spécifique au locus *Hprt*.
- \*Trier les cellules fluorescentes, dans des conditions de différenciations variables
- \*Isoler les cellules intéressantes et séquencer l'ADN cloné (enhancer).

Mélange entre le côté aléatoire du piège à enhancer et une approche plus ciblée dans le choix de l'ADN et plus comparative dans la détection.

14



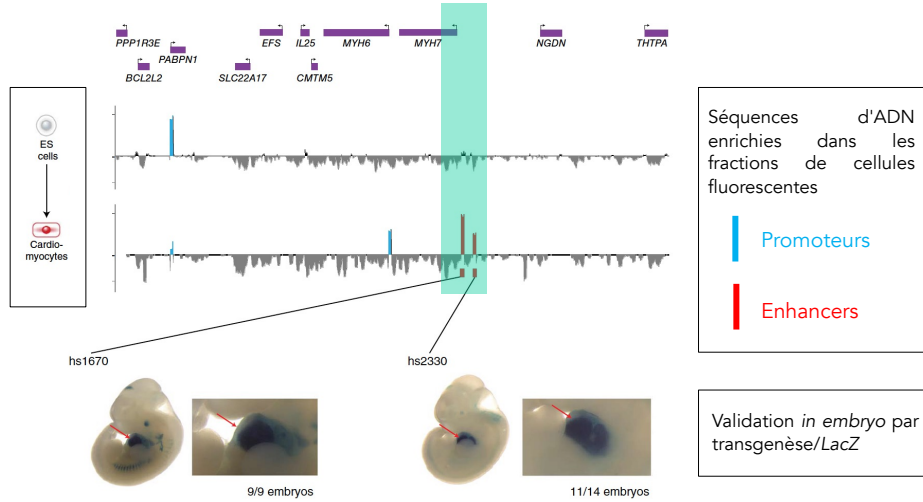
15



16



Dans ce cas précis, les deux enhanceurs se trouvent à proximité des gènes cibles, mais souvent la distance est bien plus grande...



17

\*Deux exemples de 'landing sites' pour transgènes en cellules ES

Le locus *Hprt* mutant et le locus *H11*

Published online 4 December 2013

Nucleic Acids Research, 2014, Vol. 42, No. 5, e34  
doi:10.1093/nar/gkt1290

**DICE, an efficient system for iterative genomic editing in human pluripotent stem cells**

Fangfang Zhu<sup>1,2</sup>, Matthew Gamba<sup>1</sup>, Alfonso P. Farruggio<sup>1</sup>, Simon Hippenmeyer<sup>3</sup>, Bosiljka Tasic<sup>3</sup>, Birgitt Schüle<sup>4</sup>, Yanru Chen-Tsai<sup>5,\*</sup> and Michele P. Calos<sup>1,\*</sup>

\*Dual Integrase Cassette Exchange (DICE)

\*Construction d'une plateforme d'échange à très haute spécificité, permettant l'étude rapide de constructions transgéniques dans des conditions comparables.



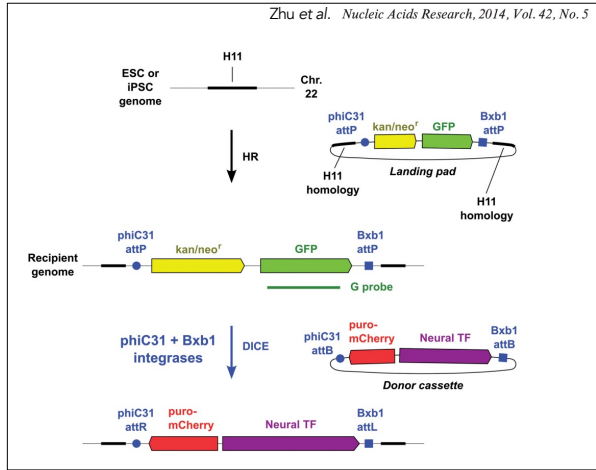
**ABSTRACT**

To reveal the full potential of human pluripotent stem cells, new methods for rapid, site-specific genomic engineering are needed. Here, we describe a system for precise genetic modification of human embryonic stem cells (ESCs) and induced pluripotent stem cells (iPSCs). We identified a novel human locus, *H11*, located in a safe, intergenic, transcriptionally active region of chromosome 22, as the recipient site, to provide robust, ubiquitous expression of inserted genes. Recipient cell lines were established by site-specific placement of a 'landing pad' cassette carrying *attP* sites for  $\phi$ IC31 and Bxb1 integrases at the *H11* locus by spontaneous or TALEN-assisted homologous recombination. Dual integrase cassette exchange (DICE) mediated by  $\phi$ IC31 and Bxb1 integrases was used to insert genes of interest flanked by  $\phi$ IC31 and Bxb1 *attB* sites at the *H11* locus, replacing the landing pad. This system provided complete control over content, direction and copy number of inserted genes, with a specificity of 100%. A series of genes, including mCherry and various combinations of the neural transcription factors LMX1a, FOXA2 and OTX2, were inserted in recipient cell lines derived from H9 ESC, as well as iPSC lines derived from a Parkinson's disease patient and a normal sibling control. The DICE system offers rapid, efficient and precise gene insertion in ESC and iPSC and is particularly well suited for repeated modifications of the same locus.

18

**\*Deux exemples de 'landing sites' pour transgènes en cellules ES**

**Le locus *Hprt* mutant et le locus *H11***



\*Procédure très efficace et très sélective (permet d'obtenir un évènement avec une fréquence proche de 100%).

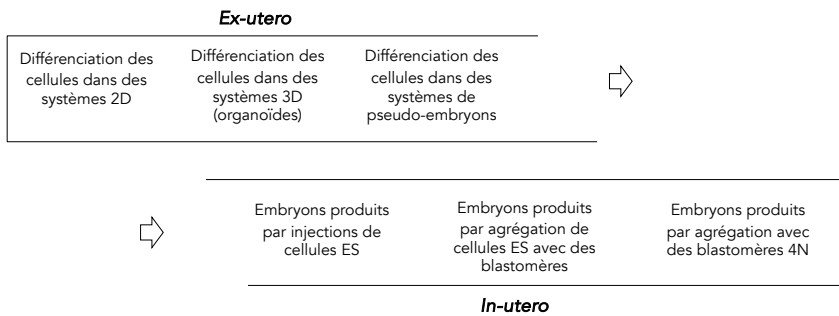
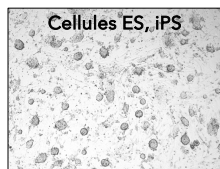
\*Procédure peu agressive pour les cellules (double sélection..).

\*Préparation d'une cassette d'échange nécessaire.

*\*...garder le locus comme point d'atterrissage mais utiliser simplement la recombinaison homologue (CRISPR..) et non pas une cassette d'échange. Avec ou sans sélection en parallèle...5-20% succès...*

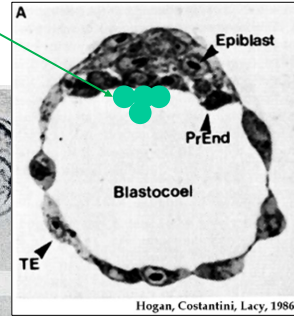
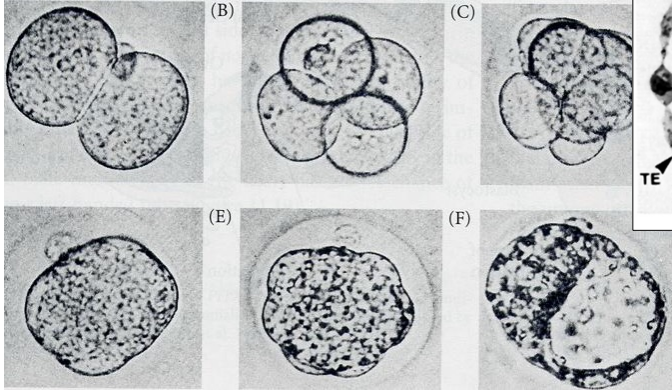
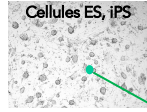
19

**\*Comment remettre les cellules ES (iPS) dans un contexte physiologiquement utile, informatif ?**



20

Faire des embryons par la route des cellules souches (injection)



'Developmental Biology', Scott Gilbert.

Les cellules ES ne font pas de dérivés extra-embryonnaires ⚠

21

Cellules souches embryonnaires et souris chimères

\*Production d'une souris à partir de cellules ES modifiées/ La route des cellules souches (d'autres 'routes' par les cellules souches sont possibles aujourd'hui; embryons 4n..)

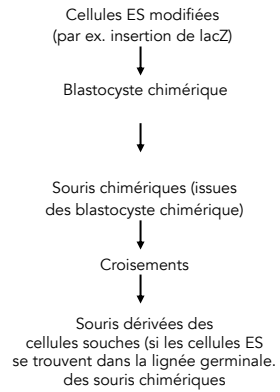
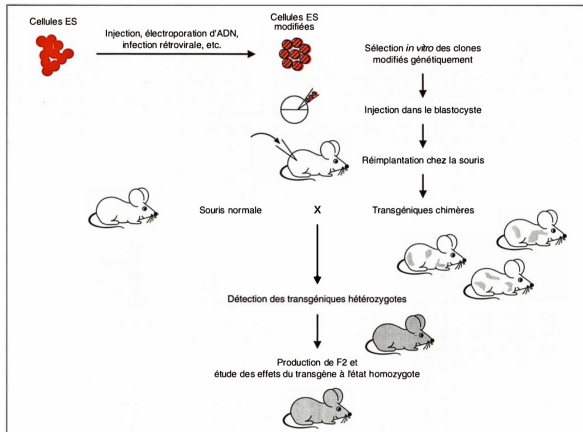


Figure 1. **Transgénèse utilisant les cellules ES.** Les cellules ES cultivées in vitro peuvent être modifiées génétiquement par divers moyens (mutagenèse, transfection d'ADN, infection rétrovirale, etc.). Les clones modifiés sont sélectionnés.

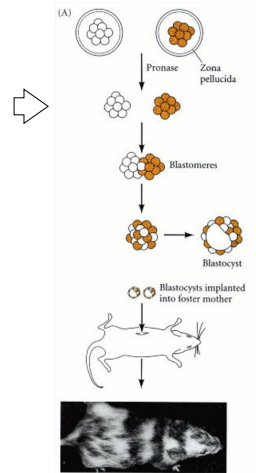
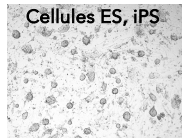
Charles Babinet (1992) Médecine/Sciences

DONC: production d'une souris modifiée à partir de cellules souches ES modifiées

22

Faire des embryons par la route des cellules souches (agrégation)

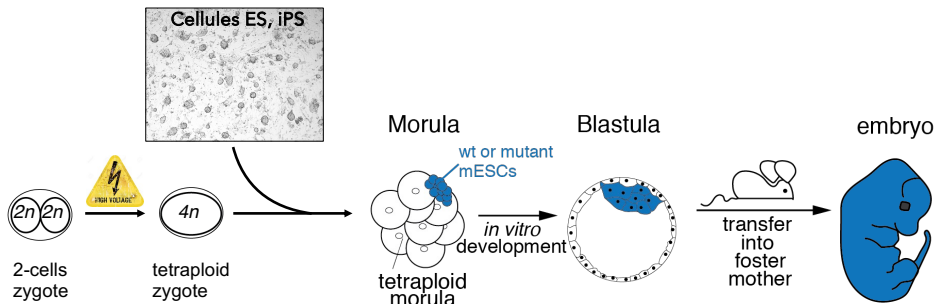
Des souris 'allophènes'  
(par agrégation de morula ou avec des cellules ES...)



23

Faire des embryons par la route des cellules souches (agrégation avec 4n)

Des souris 'allophènes' avec receveur tétraploïde (4n)  
(Par agrégation avec des cellules ES)



Résultats possibles en F0 déjà (mais alors pas de lignée établie...)  
(analyses quantitatives, temporelles..)

24



### \*Dynamique des séquences enhancers ?

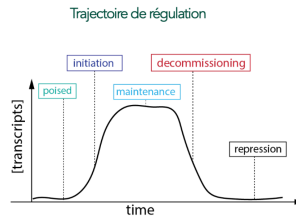
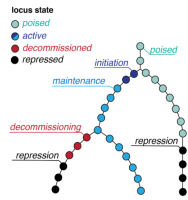
#### Temporal constraints on enhancer usage shape the regulation of limb gene transcription

Raquel Rouco, Antonella Rauseo, Guillaume Sapin, Olimpia Bompadre, Fabrice Darbellay, Guillaume Andrey

doi: <https://doi.org/10.1101/2024.03.22.585864>



bioRxiv  
THE PREPRINT SERVER FOR BIOLOGY

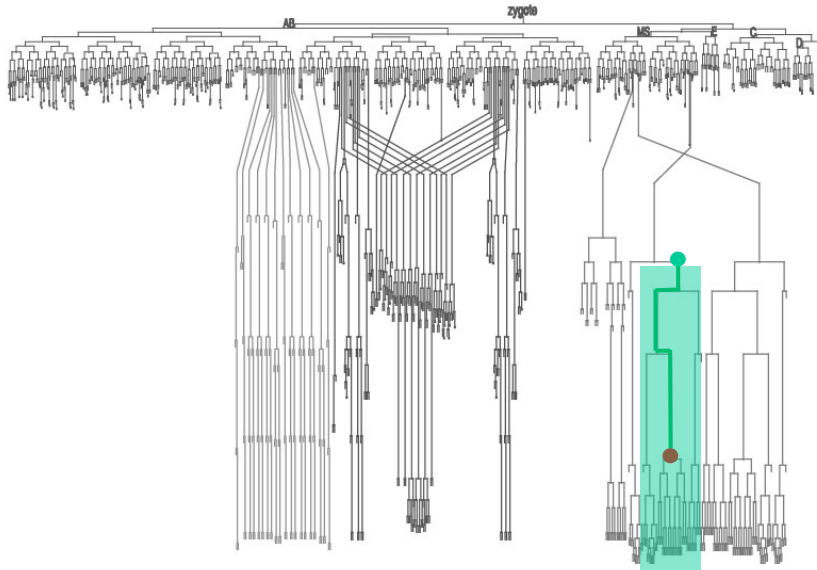


Raquel Rouco et Guillaume Andrey

25

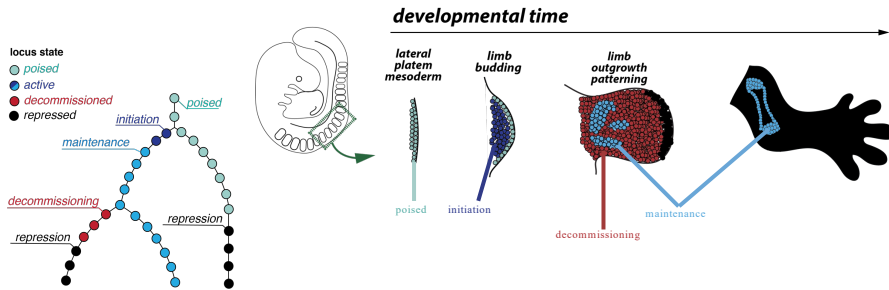


### Linéage cellulaire complet de *C. elegans* (Sulston et al.)



26

\*Dynamique des séquences enhancers  
Dynamique de la morphogenèse

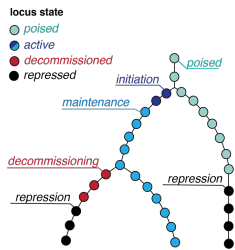


Peut-on suivre ces étapes de façon dynamique?  
Et isoler les cellules se trouvant à des stades différents, à des fins comparatives?

Raquel Rouco et Guillaume Andrey

27

\*Dynamique des séquences enhancers  
Dynamique de la morphogenèse



Le gène humain *SHOX* est localisé dans la région pseudo-autosomale 1 (sur le X et le Y), qui donc échappe à l'inactivation du X. Ce gène n'existe pas chez la souris, qui en revanche a *Shox1* et *Shox2*...

Les régions pseudo-autosomales (ou pseudo-autosomales, ou PAR pour Pseudo-Autosomal Region en anglais) sont des portions des chromosomes sexuels (ou gonosomes) homologues entre les deux chromosomes sexuels. Ainsi, elles sont héritées d'un point de vue statistique comme s'ils étaient portés par les autosomes.

Wikipedia  
[https://fr.wikipedia.org/wiki/R%C3%A9gion\\_pseudo-autoso...](https://fr.wikipedia.org/wiki/R%C3%A9gion_pseudo-autoso...)

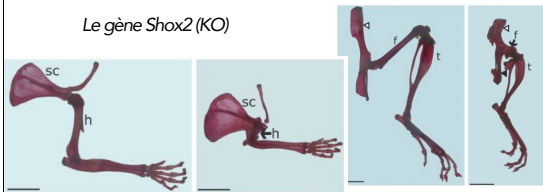
**A mouse model for human short-stature syndromes identifies *Shox2* as an upstream regulator of *Runx2* during long-bone development**

John Cobb<sup>1</sup>, Andriée Dierich<sup>1</sup>, Yolande Huss-García<sup>1</sup>, and Denis Duboule<sup>1\*</sup>

<sup>1</sup>Department of Zoology and Animal Biology and National Research Center "Frontiers in Genetics," Sciences III, University of Geneva, Quai Ernest Ansermet 30, 1211 Geneva 4, Switzerland and <sup>2</sup>Unité de Génétique et de Biologie Moléculaire et Cellulaire Institut Clinique de la Santé, Centre National de la Recherche Scientifique, Institut National de la Santé et de la Recherche Médicale/Université Louis Pasteur/Collège de France, BP 10142, 67080 Strasbourg, France

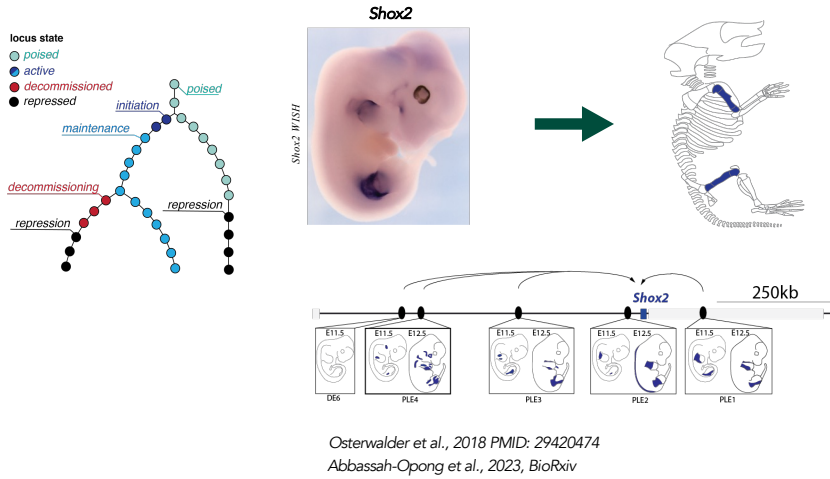
Deficiencies or mutations in the human pseudoautosomal *SHOX* gene are associated with a series of short-stature conditions, including Turner syndrome, Leri-Weill dyschondrosteosis, and Langer mesomelic dysplasia. Although this gene is absent from the mouse genome, the closely related paralogous gene *Shox2* displays a similar expression pattern in developing limbs. Here, we report that the conditional inactivation of *Shox2* in developing appendages leads to a strong phenotype, similar to the human conditions, although it affects a different proximodistal limb segment.

Le gène *Shox2* (KO)



28

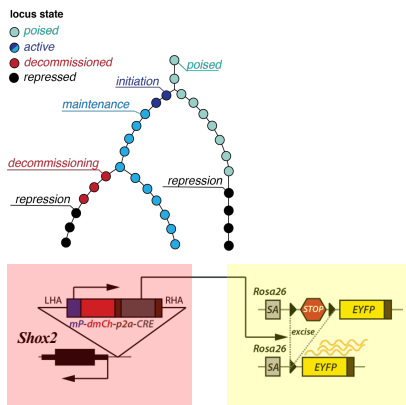
\*Comment étudier la dynamique des séquences enhancers ?



Osterwalder et al., 2018 PMID: 29420474  
Abbassah-Opong et al., 2023, BioRxiv

Raquel Rouco et Guillaume Andrey

\*Un système à deux fluorescences pour suivre une trajectoire de régulation



- \*Rosa26 est un gène exprimé dans toutes les cellules
- \*Une EYFP est placée à sa suite avec un site accepteur d'épissage (SA)
- \*Entre les deux une séquence de terminaison est introduite (STOP)...
- \*...flanquée de deux sites de recombinaison reconnus par la recombinaison CRE
- \*La recombinaison CRE est produite en même temps qu'une fluorescence rouge par un transgène inséré juste devant le gène Shox2 qui va donc repêcher les influences des enhancers
- \*La CRE va enlever le STOP et marquer de façon définitive toutes les cellules dérivant de cellules ayant exprimé Shox2 une fois ou l'autre...

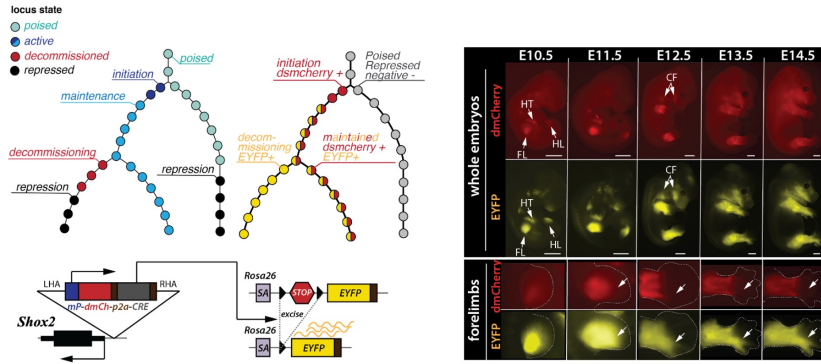
C. A double fluorophore approach to monitor Shox2 locus activity over time; mP: minimal  $\beta$ -globin promoter, dmCh: mCherry gene with a destabilized PEST sequence; P2A: self-cleavage sequence; CRE: CRE recombinase gene; SA: splice acceptor; STOP: floxed 3xSV40pA STOP signal; EYFP: EYFP gene.

Rouco et al., bioRxiv 2024

Raquel Rouco et Guillaume Andrey



\*Un système à deux fluorescences pour suivre une trajectoire de régulation



Rouco et al., bioRxiv 2024

Raquel Rouco et Guillaume Andrey