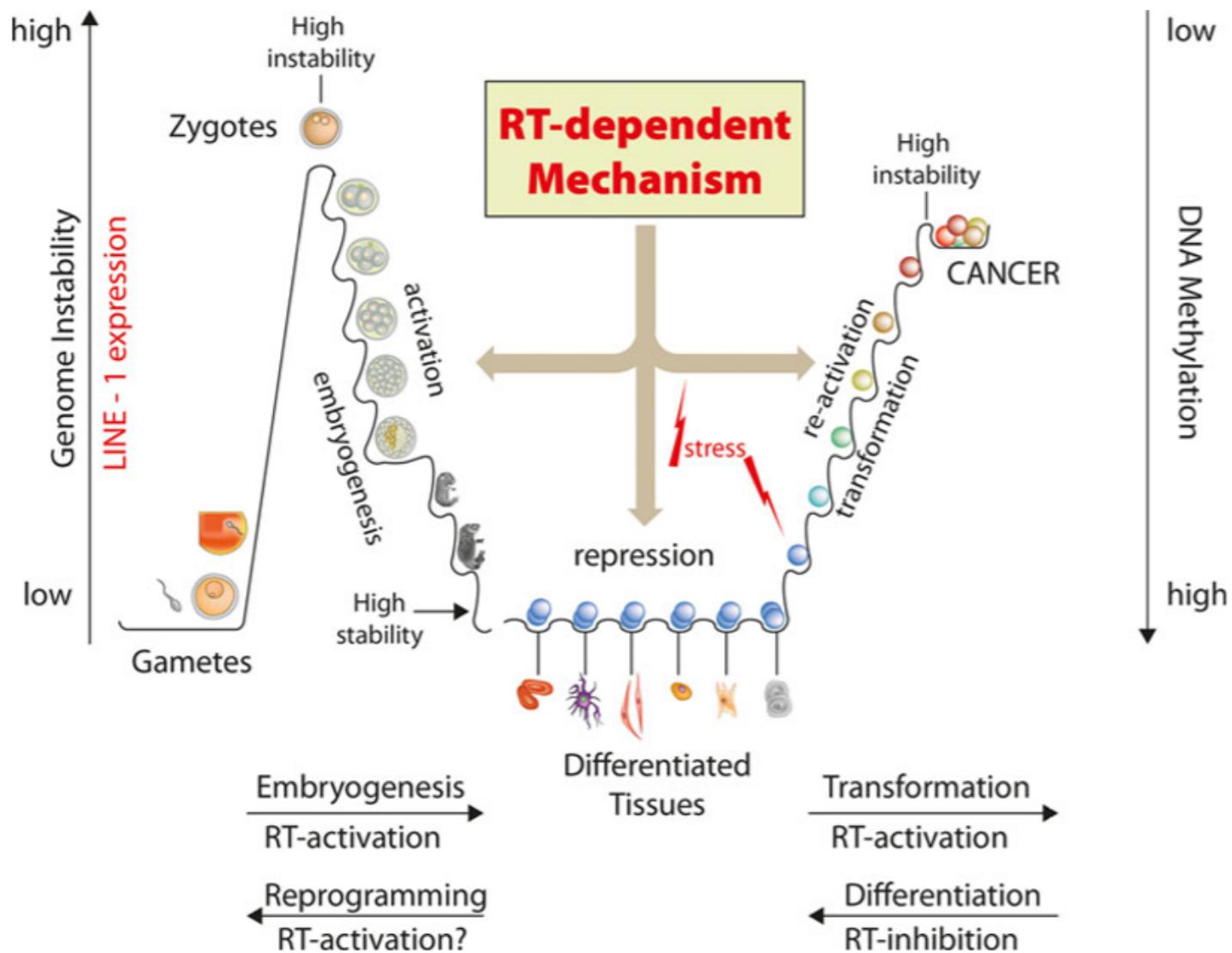


***PSEUDOGENE lncRNAs EXAMPLE 1***

***Pseudogene lncRNA that controls embryonic stem cell  
Self-renewal***

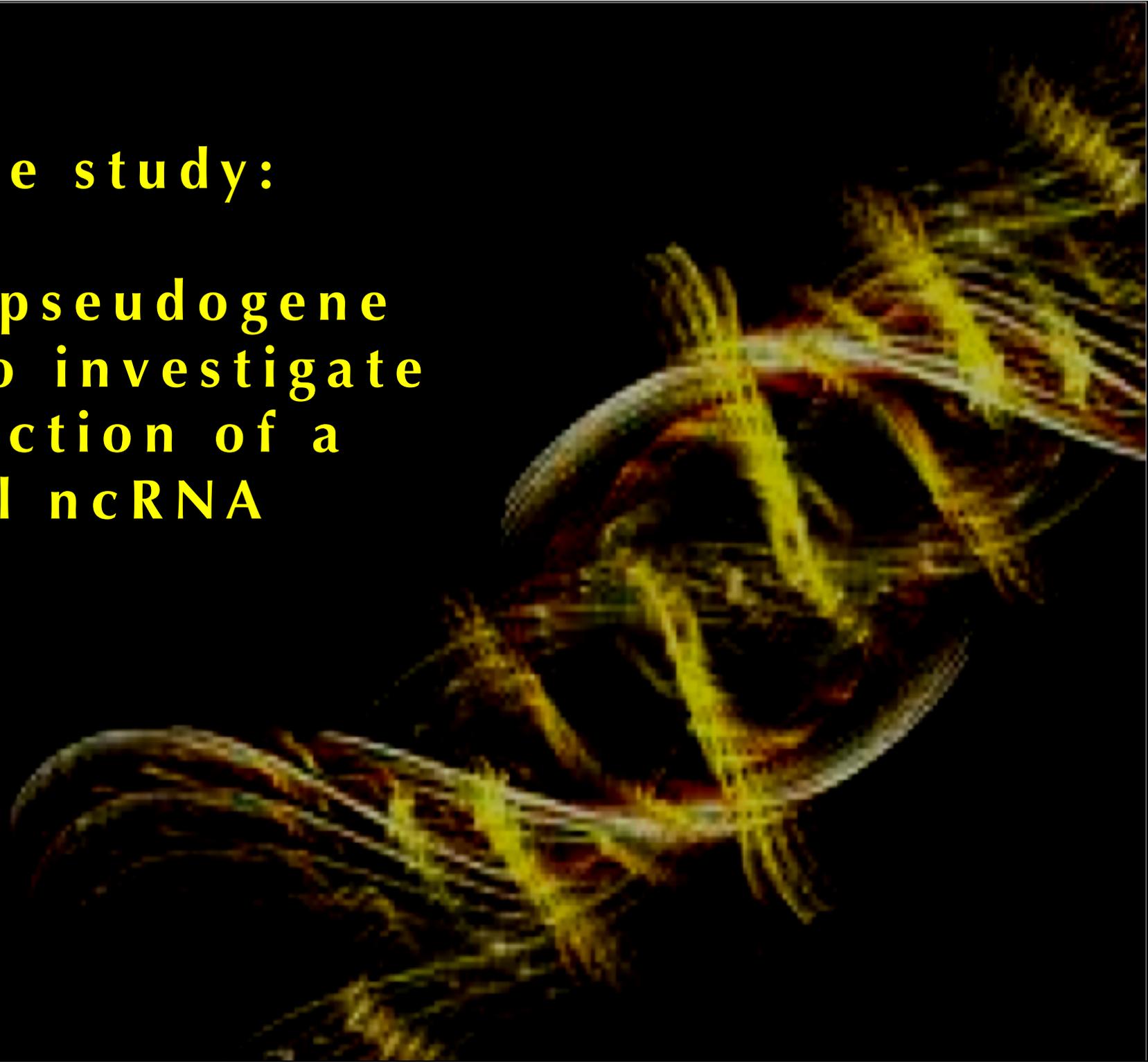
***A Oct4P4 pseudogene derived lncRNA silences the  
ancestral Oct4 gene in trans***

# Retrotransposon activity during development

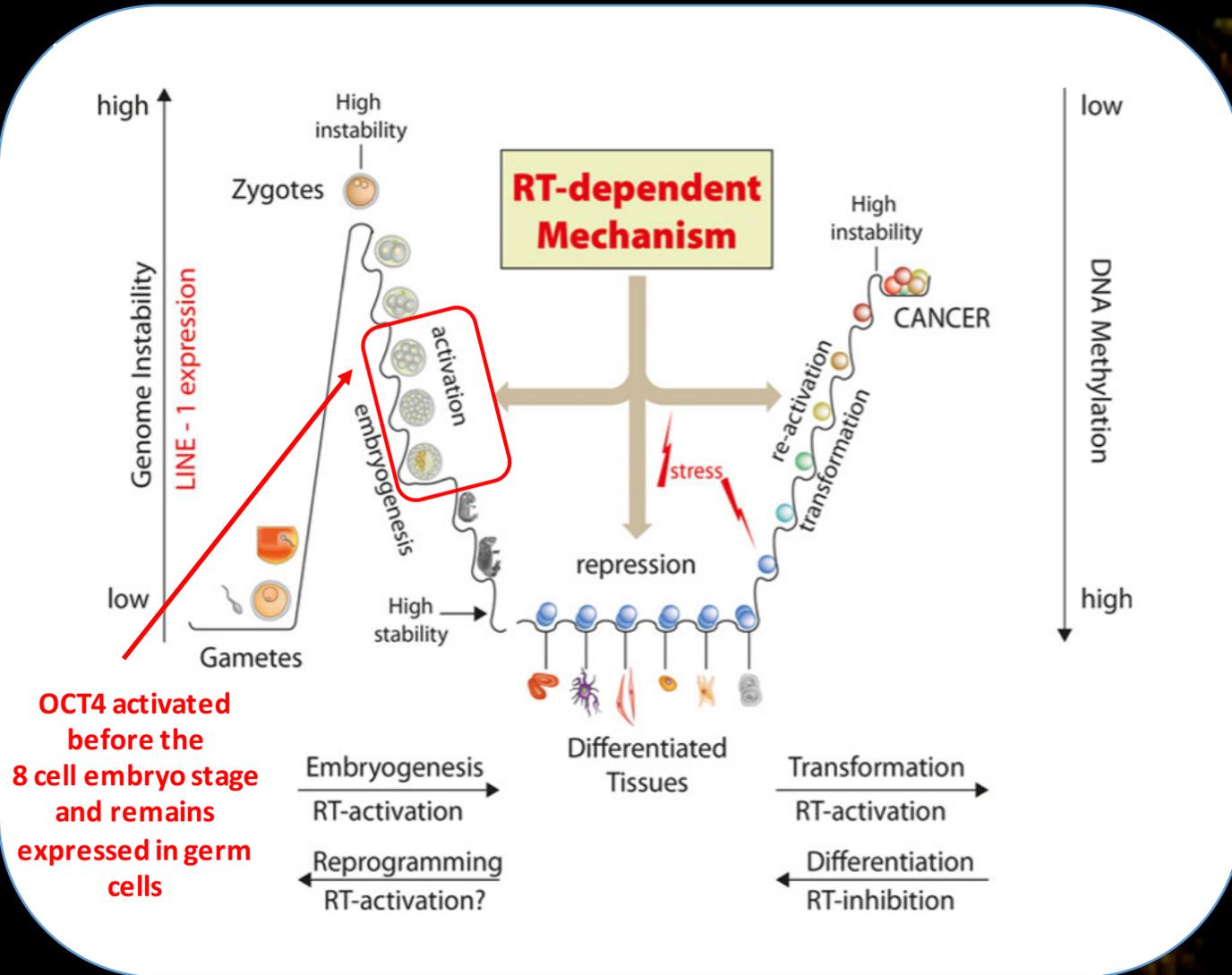


**A case study:**

**Using a pseudogene  
lncRNA to investigate  
the function of a  
novel ncRNA**



# Mouse and human contain several processed OCT4 pseudogenes

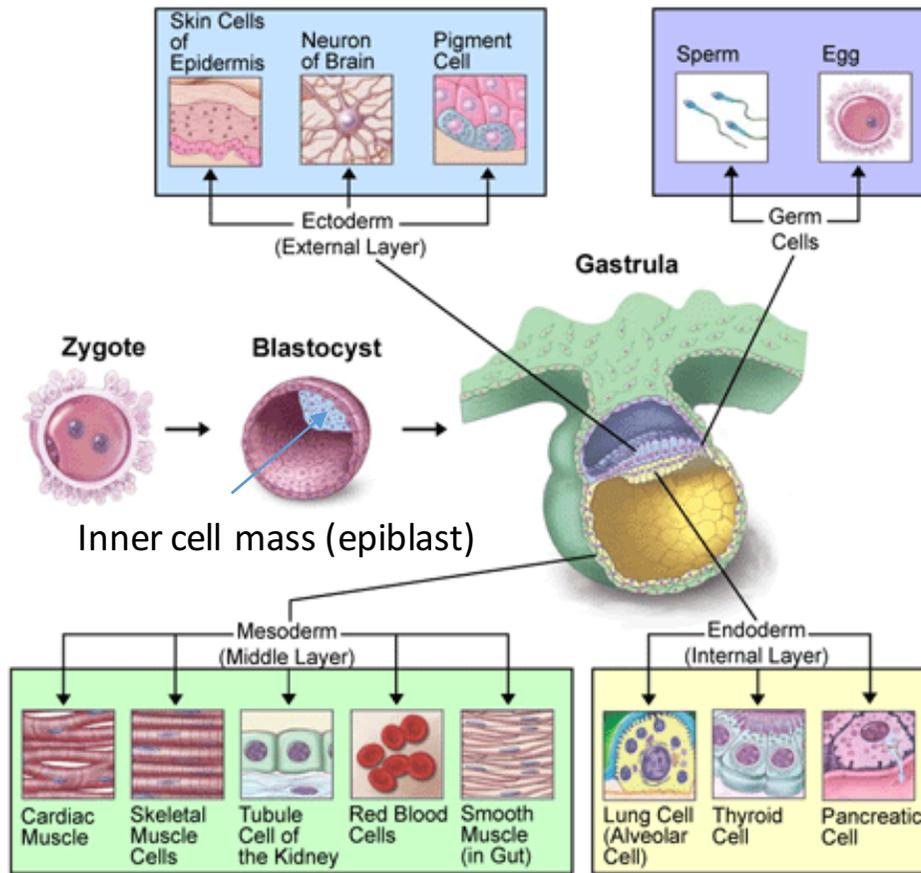


**Human:**  
 1 ancestral OCT4: Transcription factor  
 7 processed OCT4 pseudogenes

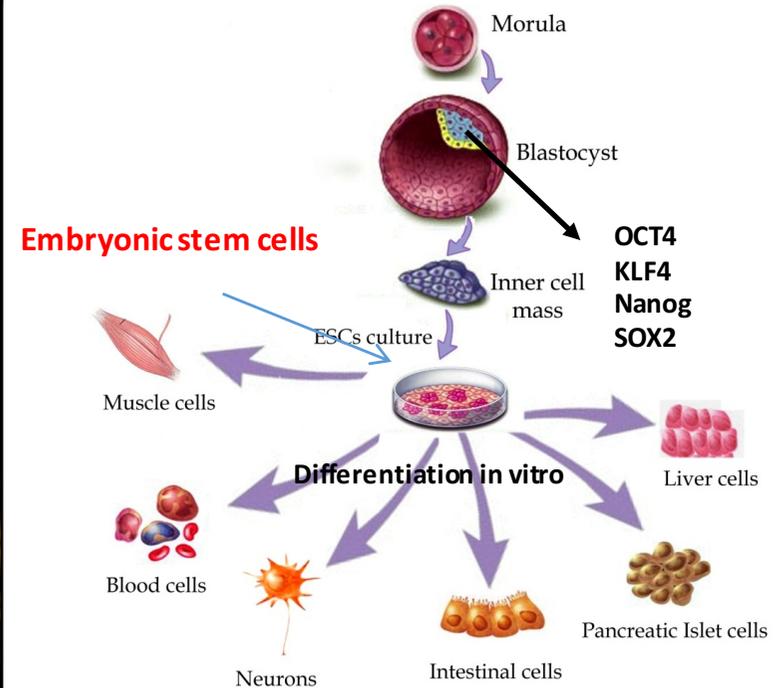
**Mouse:**  
 1 ancestral Oct4: Transcription factor  
 5 processed OCT4 pseudogenes

# The inner cell mass of the blastocyst are the source of pluripotent embryonic stem cells

Blastocyst cells give rise to all organs and cell types

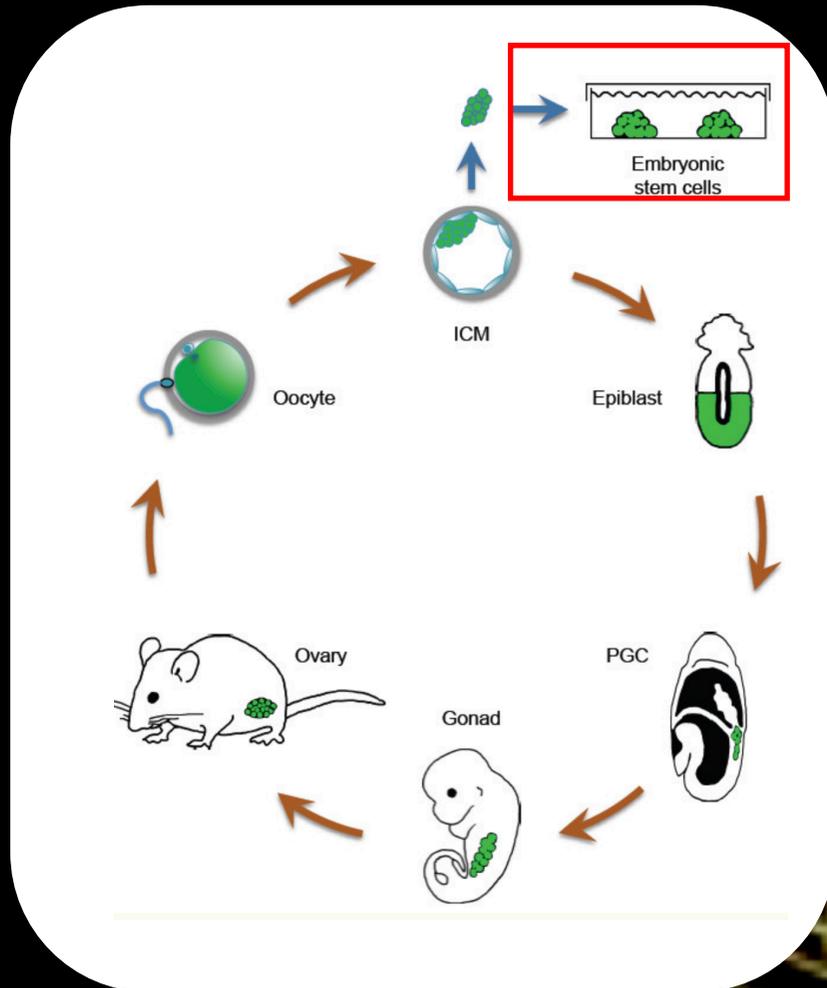


Blastocyst cells can be isolated and cultivated



# 1. OCT4 EXPRESSION

## Oct4 expression during the mouse life cycle

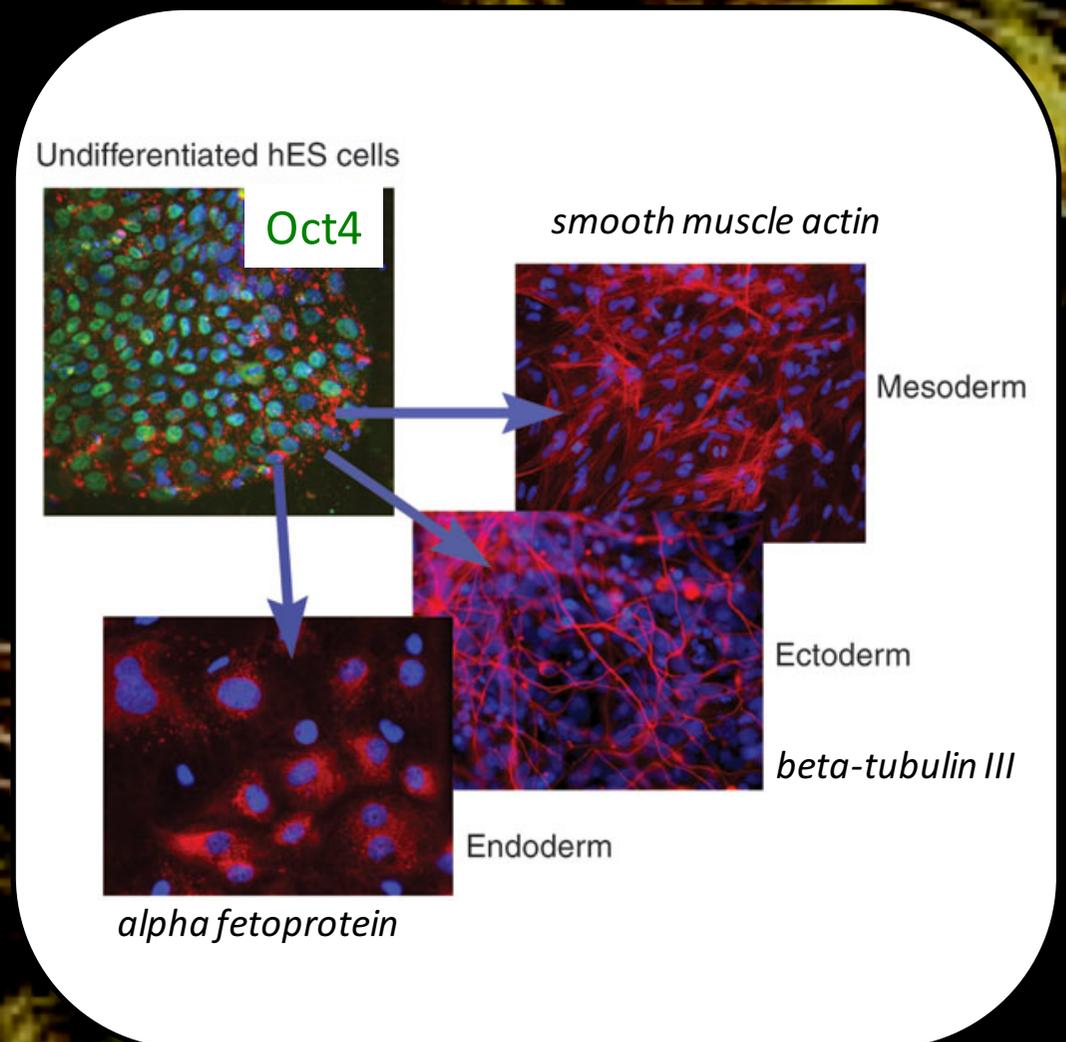
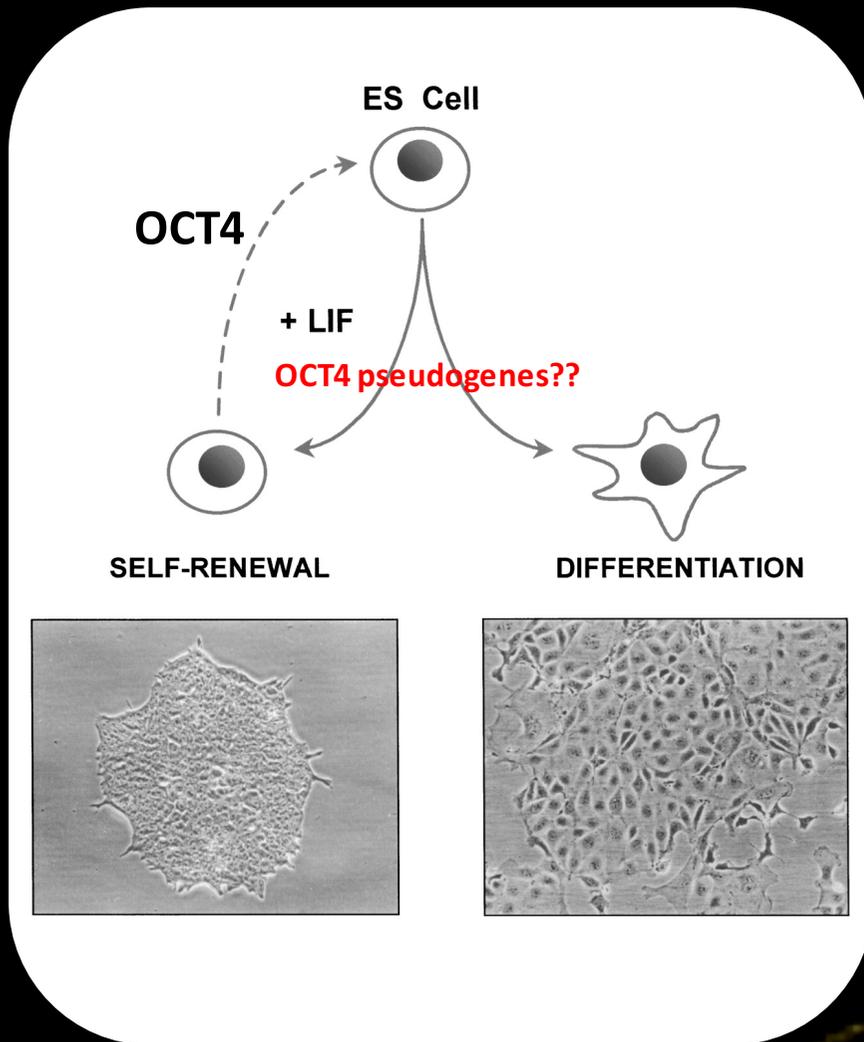


Cells and tissues expressing Oct4 are marked in green.

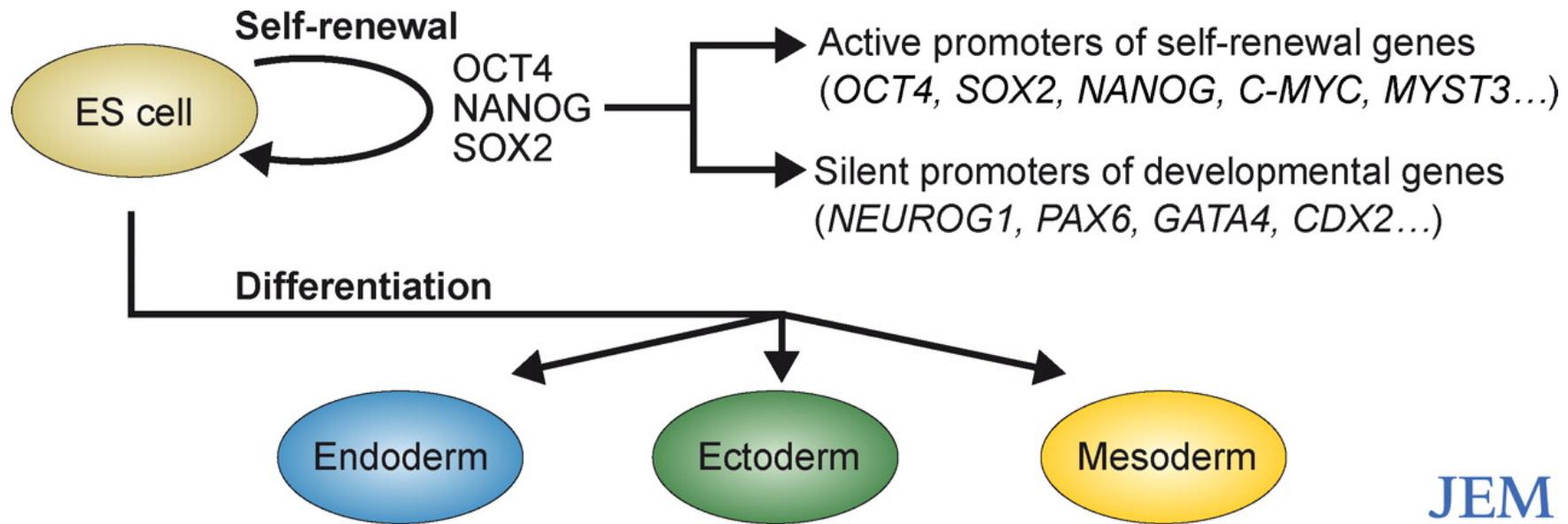
- Oct4 is expressed in mouse oocytes as a maternal transcript and protein.
- Zygotic Oct4 expression is activated prior to the 8-cell stage and is abundant and uniform in all cells of the embryo throughout the morula stage.
- However, as the outer cells of the embryo differentiate into the trophectoderm, Oct4 expression is restricted to cells of the ICM in the blastocyst.
- After implantation, Oct4 expression is maintained in the epiblast of the blastocyst
- Oct4 expression becomes restricted to primordial germ cells (PGCs), which are first specified in the extraembryonic mesoderm at the base of the allantoic bud during gastrulation. PGCs give rise to gametes, which following fertilization will develop into a new organism of next generation.

## 2. THE MODEL SYSTEM: ES CELL DIFFERENTIATION

**OCT4** expressing ES cells have self-renewing and differentiation potential *in vitro*



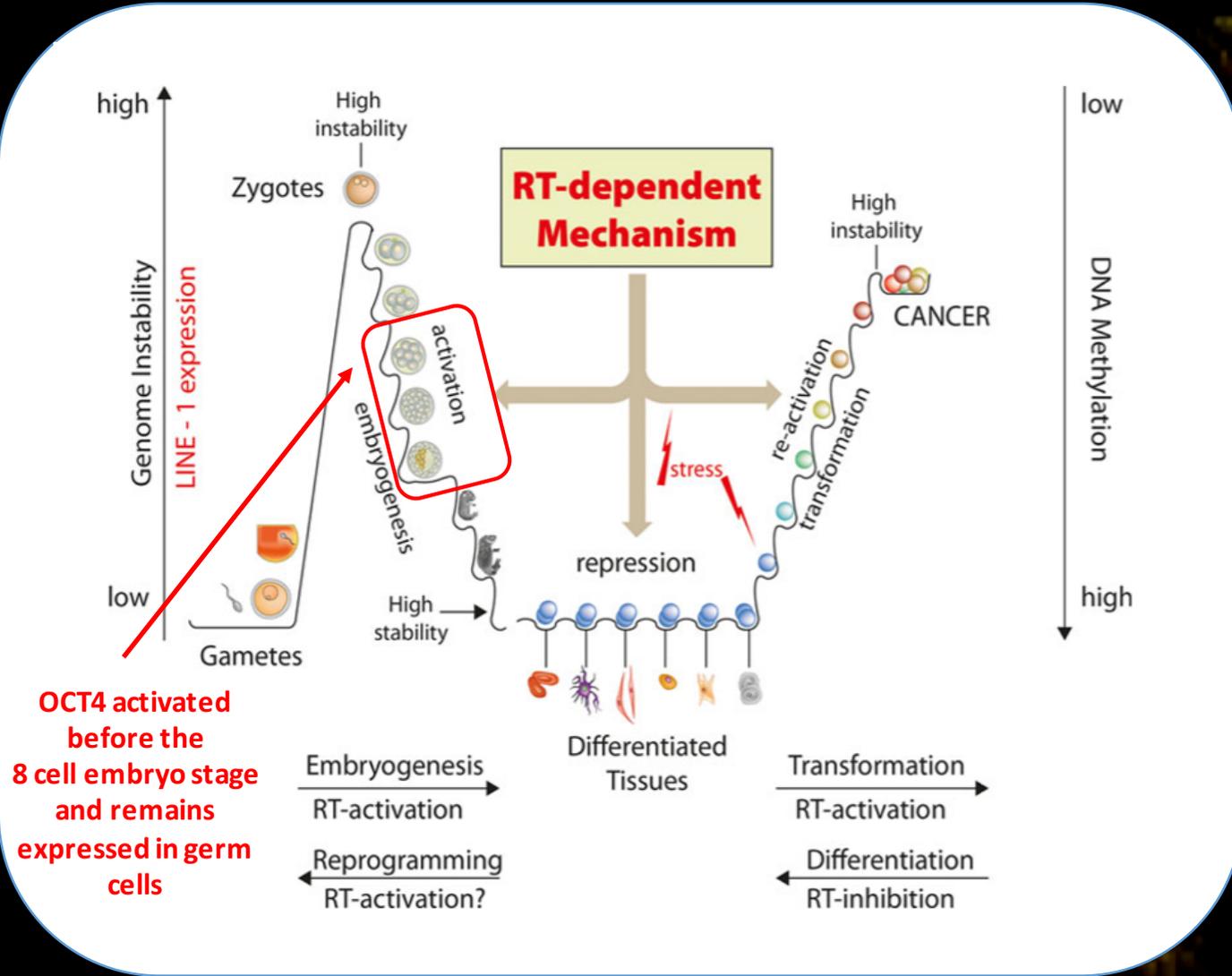
# The self-renewal transcription factor Oct4 is essential for embryonic stem cell self-renewal



JEM

Nicolaj Strøyer Christophersen, and Kristian Helin J Exp Med 2010;207:2287-2295

# Mouse and human contain several processed OCT4 pseudogenes



**Human:**  
 1 ancestral OCT4  
 7 processed OCT4  
 pseudogenes

**Mouse:**  
 1 ancestral Oct4  
 5 processed OCT4  
 pseudogenes

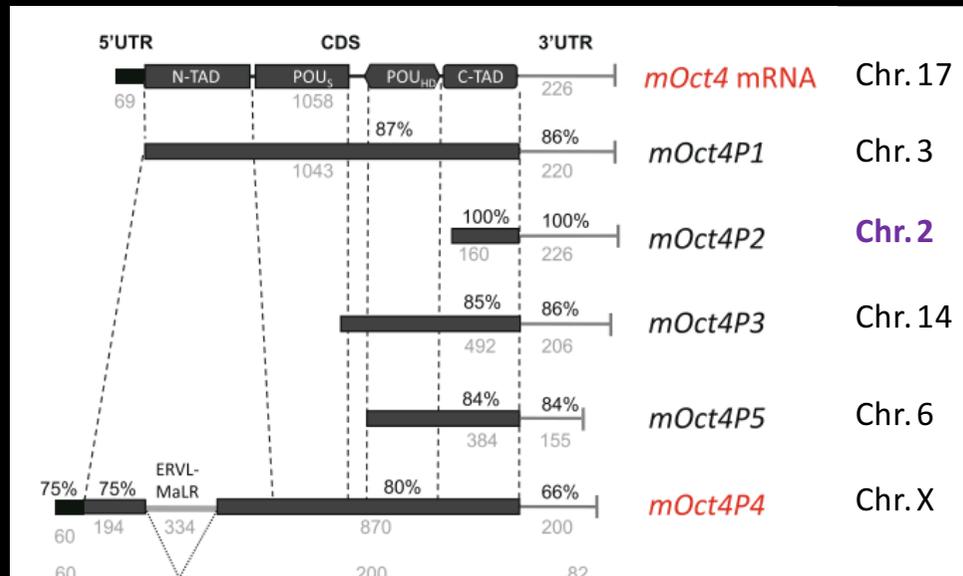
**OCT4 activated before the 8 cell embryo stage and remains expressed in germ cells**

# 3. Sequence analysis: Are there other sequences in the mouse genome that are similar to Oct4?

## Use Oct4 cDNA and make BlastN

Description	Graphic Summary	Alignments	Taxonomy					
Sequences producing significant alignments								
<input checked="" type="checkbox"/> select all 6 sequences selected								
Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
<input checked="" type="checkbox"/> Mus musculus strain C57BL/6J chromosome 3_GRCm39	Mus musculus	1177	2059	100%	0.0	87.15%	159745316	NC_000069.7
<input checked="" type="checkbox"/> Mus musculus strain C57BL/6J chromosome 17_GRCm39	Mus musculus	710	1977	100%	0.0	100.00%	95294699	NC_000083.7
<input checked="" type="checkbox"/> Mus musculus strain C57BL/6J chromosome X_GRCm39	Mus musculus	638	638	82%	2e-180	80.20%	169476592	NC_000086.8
<input checked="" type="checkbox"/> Mus musculus strain C57BL/6J chromosome 14_GRCm39	Mus musculus	520	520	46%	8e-145	85.71%	125139656	NC_000080.7
<input checked="" type="checkbox"/> Mus musculus strain C57BL/6J chromosome 6_GRCm39	Mus musculus	383	383	36%	1e-103	84.62%	149588044	NC_000072.7
<input checked="" type="checkbox"/> Mus musculus strain C57BL/6J chromosome 1_GRCm39	Mus musculus	303	303	15%	9e-80	100.00%	195154279	NC_000067.7

## 5 processed pseudogenes



## Unspliced sequence on Chr. 3 with high similarity to Oct4

**Mus musculus strain C57BL/6J chromosome 3, GRCm39**  
Sequence ID: [NC\\_000069.7](#) Length: 159745316 Number of Matches: 2

Range 1: 129227322 to 129228364 [GenBank](#) [Graphics](#) [Next Match](#) [Previous Match](#)

Score	Expect	Identities	Gaps	Strand
1177 bits(637)	0.0	936/1074(87%)	46/1074(4%)	Plus/Plus

Features: 101192 bp at 5' side: glutamyl aminopeptidase  
97965 bp at 3' side: elongation of very long chain fatty acids protein 6

```

Query 1 ATGGCTGGACACCTGGCTTCAGACTTCGCCTTCTCACCCACCAGGTGGGGGTGATGGG 60
Sbjct 129227322 ATGGCTGGACACCTGGCTTCAGACTTCGCCTTCTCA-CCCACCAGGTGGGGCGATGGG 129227380

Query 61 TCAGAGGGCTGGAGCCGGCTGGTGGATCCTCGAACCCTGGCTAAGCTTCACAGGGCCT 120
Sbjct 129227381 TCAGAGGGCTGGAGCCGGCTGGTGGACCTCGAACCCTGGCTAAGCTTCACAGGGCCT 129227440

Query 121 CCAGTGGCCCTGGAATCGGACCAGGCTCAGAGTATTGGGGATCTCCCATGTCCGCC 180
Sbjct 129227441 CTGGTGGCCCTGGAATAGGACCAGGCTTAGAGGTATT-GGGATCTCCCATGTCCACGC 129227499

Query 181 GCATACGAGTCTCGCGGAGGATGGCATACTGTGGACCTCAGGTTGGACTGGCCCTAGTC 240
Sbjct 129227500 GTATATGAGTCTCGCGGAGGATGGCATACTGTGGACCTCAG----- 129227541

Query 241 CCCCAGTTGGCCTGGAGACTTTGCAGCCTGAGGGCCAGGCAGGACACAGTGGAAAGC 300
Sbjct 129227542 -----GCTGGCCTGGAGACTTTGCAGCCTGAGAGCCAGGCAGGACAGTGTGGAGAGC 129227595

Query 301 AACTCAGAGGGAACCTCCTCGAGCCCTGTGCCAGCCGCCCAATGCCCTGAAGTTGGAG 360
Sbjct 129227596 AACTCAGAGGGAACCTCCTCGGCCTTGTGCCAGCCACCOCAGTGCCTGAAGTTGGAG 129227655

Query 361 AAGTGAACCAACTCCCGAGGAGTCCAGGACATGAAAGCCCTGCAGAGGAGCTAGAA 420
Sbjct 129227656 ATGAGGGAACCAAGTCCCGAGGAGTCCAGGACATGAAAGCTGCAGAGGAGCTAGAG 129227715

Query 421 CAGTTGCCAAGCTGCTGAAGCAGAAGAGATCACCTTGGGGTACACCCAGGCCAGCTG 480
Sbjct 129227716 AAGTTGCCAAGCTGCTGAAGCAGAAGAGATCACCTTGGGGTACACCCAGGCCAGCTG 129227775

Query 481 GGGCTCACCCCTGGGCGTTCTCTTTGGAAGGCTTTCAGCCAGACCAACATCTGTCCGCTC 540
Sbjct 129227776 GGGCTCACCCCTGGGCGTTCTCTCAGAAAGTG-TCAGCCAGACAACCATCTACTGCTTT 129227834

Query 541 GAGGCCCTGCAGCTCAGCCTTAAGAACATGTGTAGCTGCGGCCCTGCTGAGAAAGTGG 600
Sbjct 129227835 GAGGCCCGCAGCTCAGCCTTATGACCATGCATAAGCTGCGGCCCTGCTGAGAAAGTGG 129227894

Query 601 GTGAGGAAAGCCGACAACAATGAGAACCCTCAGGAGATATGCAAATCGGAGACCCTGGTG 660
Sbjct 129227895 GTAGAGGAAAGCCGAAAGACAATGAGAACCCTCAGGAGATATGCAAATCGGAGACCCTGGTG 129227954

Query 661 CAGGCCCGGAAGAAAAGCGAATAGCATTGAGAACCTGTGAGGTTGAGCTTGGAGACC 720
Sbjct 129227955 CAGGCCCGGAAGAAAAGCGGAGCATGAGAACCATGTGAGGAGGAGCCCTGGAGAA 129228014

Query 721 ATGTTTCTGAAGTGCCCGAAGCCCTCCCTACAGCAGATCACTACA-TCGC--CA--AT- 774
Sbjct 129228015 GAGTTTCTGAAGTGCCCGAGCCCTCCCGCAGCAGATCACTAGCATTTAGCATCACCATG 129228074

Query 775 CAGCTTGGGCTAGAGAAGGATGTGGTTCGAGTATGGTTCTGTAAACCGGCCAGAAAGGGC 834
Sbjct 129228075 CAGCTTGGGATGGAGAAGGACATGGTTCGAGGATGTGTTCTGTAAACCGGCCAGAAAGGGC 129228134

Query 835 AAAAGATCAACTATTGAGTATTCCTAACGAGAAGAGTATGAGGCTA-CAGG-ACACCTT 892
Sbjct 129228135 AAAAGATCAAGTATTGAGTATTCCTAACGAGAAAAGTATGAGGCTTCCCGGATGCTT 129228194

Query 893 TCC--CA-ggg-CTGTATCTTCTCTGCCCCAGGCTCCCACTTTGGCACCC 946
Sbjct 129228195 TCGGGCGTGGGCTGGGCTGTATCTTCTCTGCCCCAGG-CCCACTTTGGCACCC 129228253

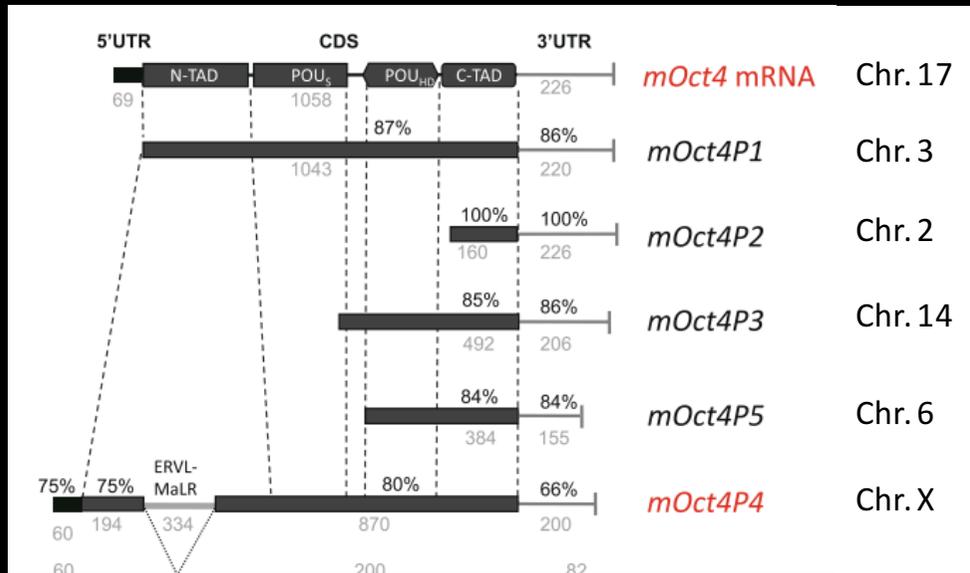
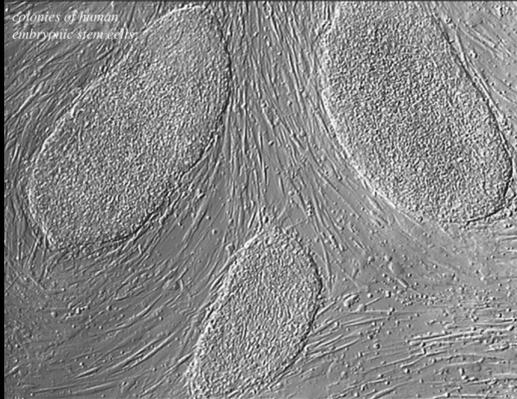
Query 947 CAGGCTATGGAAGCCCACTTCACCACACTTACTCAGTCCCTTTCTCGAGGGCGAGG 1006
Sbjct 129228254 CAAGTATGGGAGCCCACTTCACCACACT--ACTCGATACCTG-TCCTGAGGGCGAGG 129228310

Query 1007 CCTTTCCTCTGTTCC- GTCACCTGCTTGGGCTCTCCCATGCATTCAAACTGA 1059
Sbjct 129228311 CCTTTCCTCTGTTCCCTGCTGCTTGGGCTCTCCCATGCATTCAAACTGA 129228364
  
```

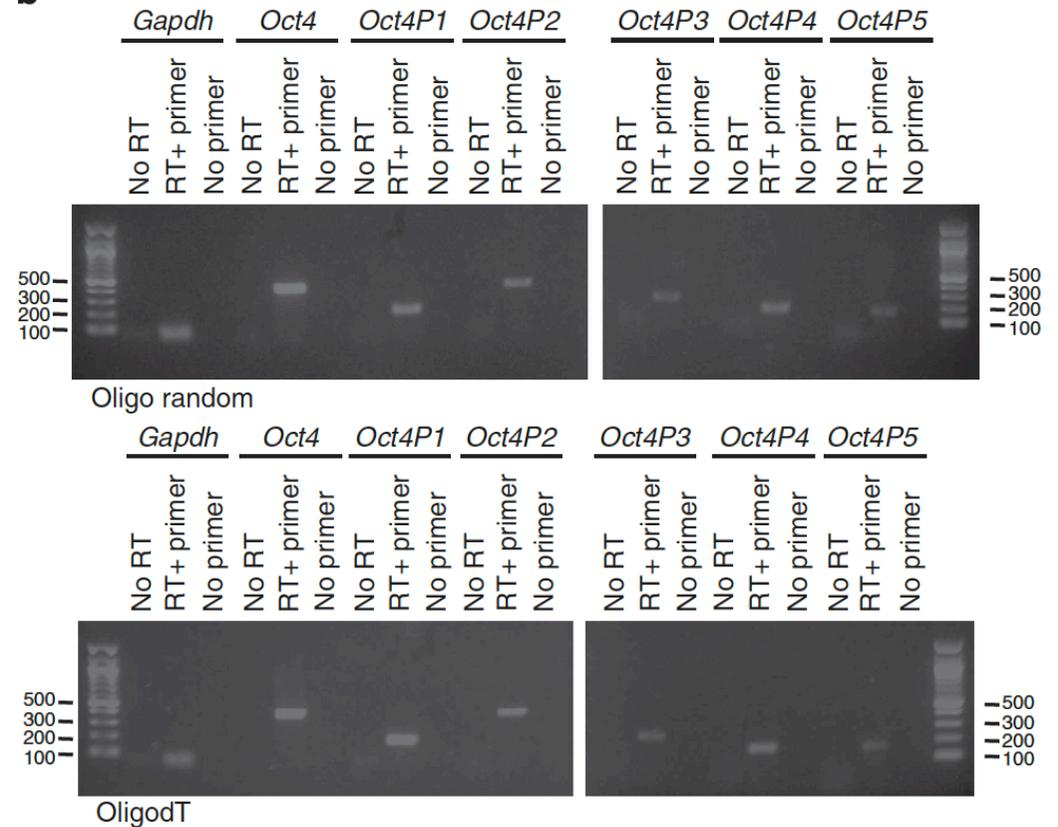
Oct4  
Chr3 sequence

# 4. EXPERIMENTAL EVALUATION OF lncRNA EXPRESSION IN MODEL CELLS

## Self-renewing mESCs



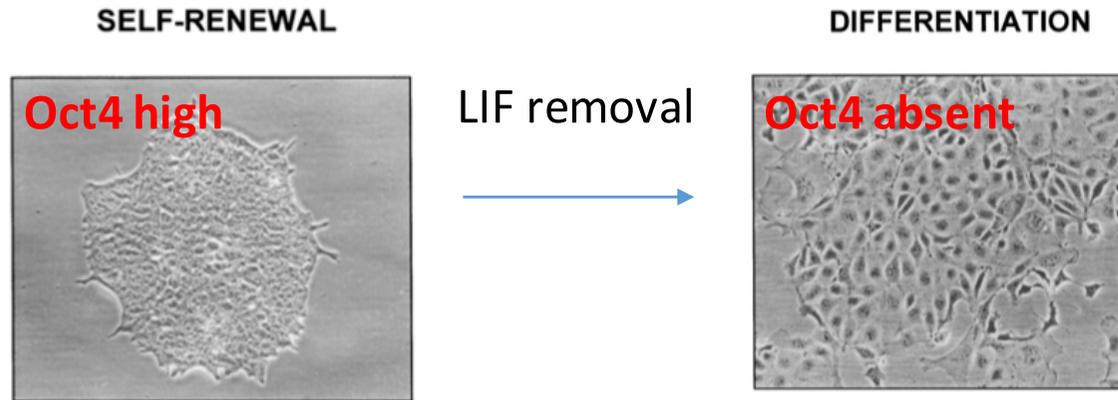
**b**



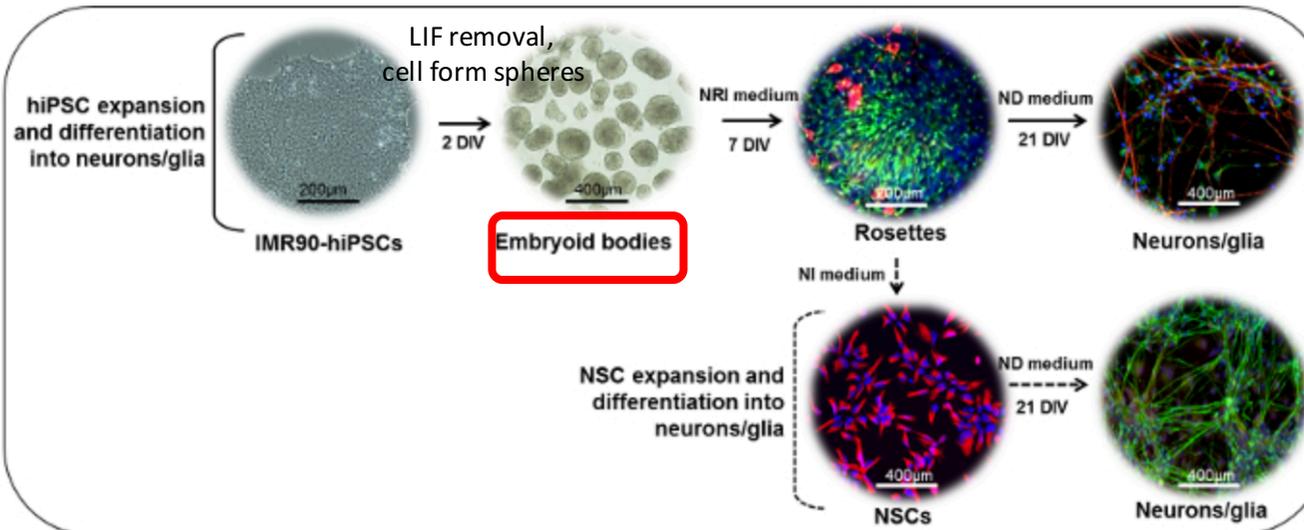
**Ancestral OCT4 gave rise to 5 processed pseudogenes that are expressed in mESCs**

# 4. EXPRESSION DURING IN VITRO DIFFERENTIATION

In vitro differentiation; 2D cell culture, attached on petri dishes

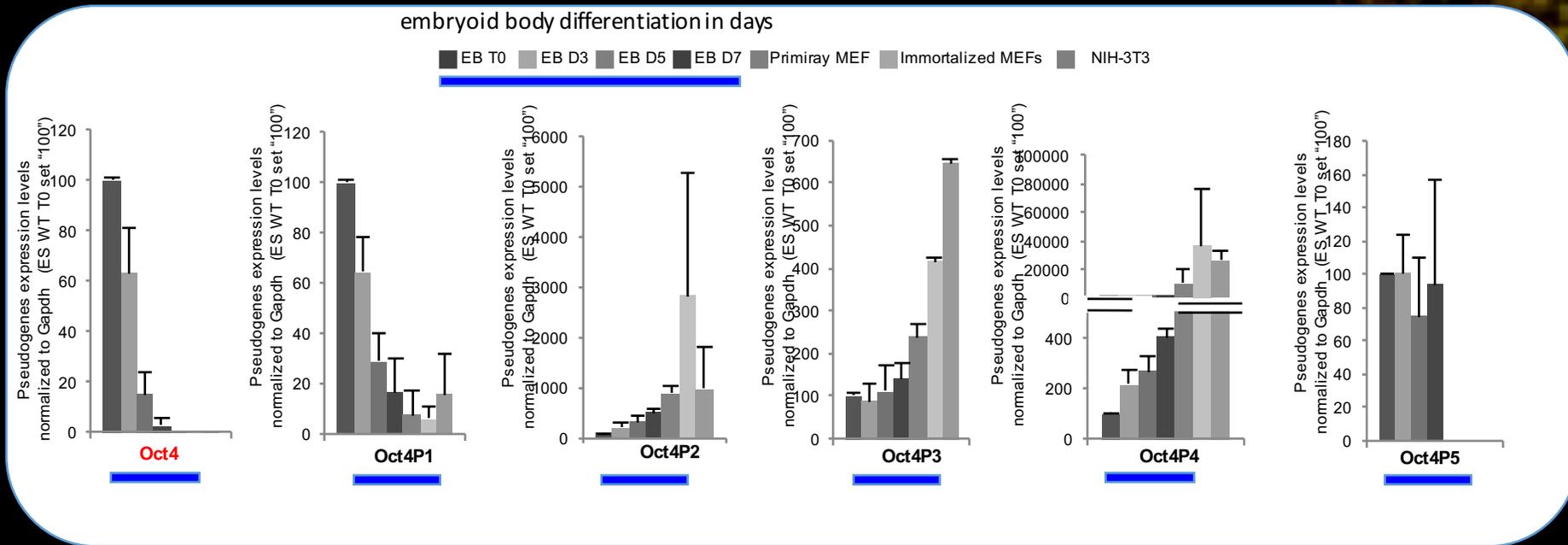


In vitro differentiation in 3D cell culture, embryoid body differentiation

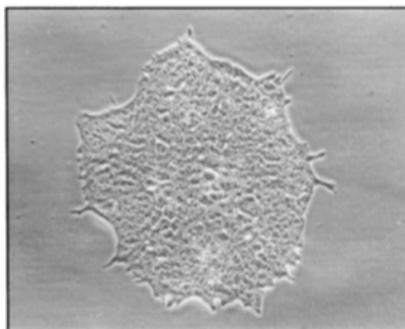


EB: embryoid body differentiation (a model for ES differentiation)

# Oct4 pseudogenes are tightly controlled during the differentiation of mESCs

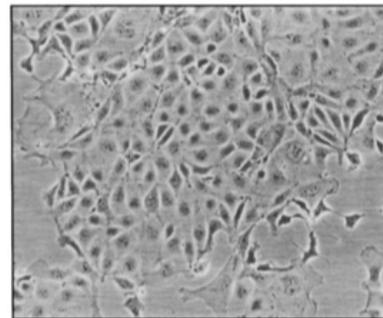


## SELF-RENEWAL



**Oct4**  
**Oct4P1 (-10X)**

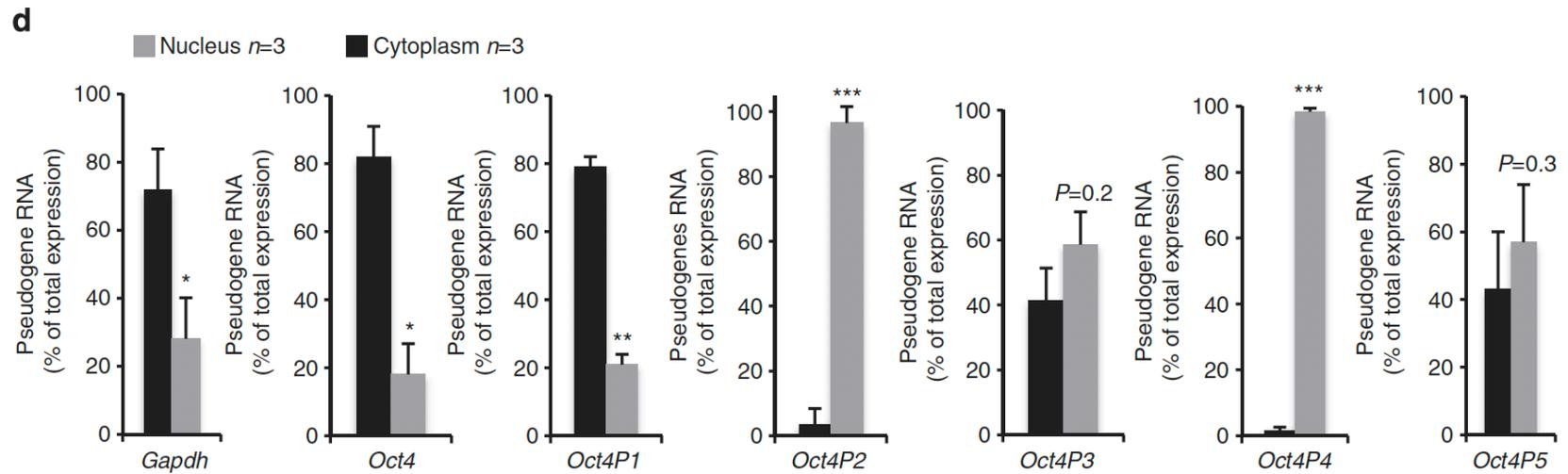
## DIFFERENTIATION



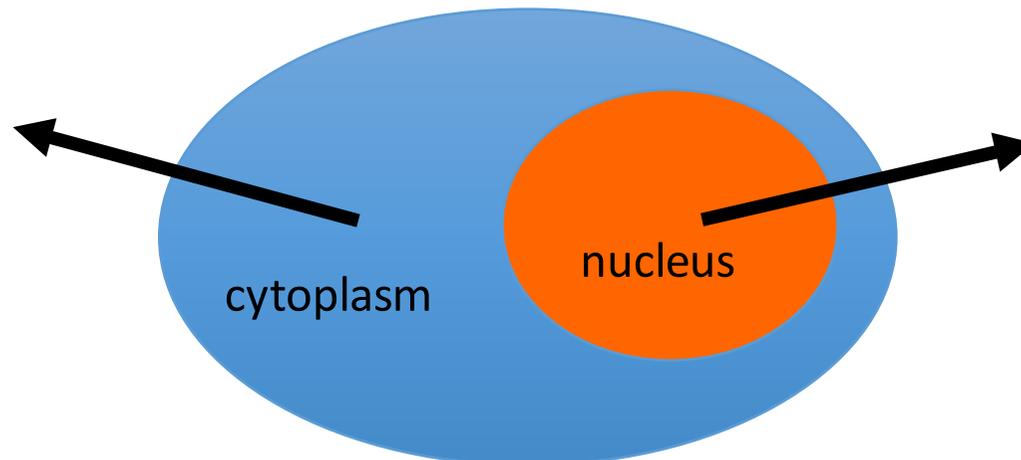
**Oct4P2 (+9x)**  
**Oct4P3 (+2x)**  
**Oct4P4 (+4x; Fibrobl. +200x)**

# 5. LOCALIZATION OF PSEUDOGENE lncRNA

## OCT4 pseudogenes are localized to nucleoplasm or cytoplasm

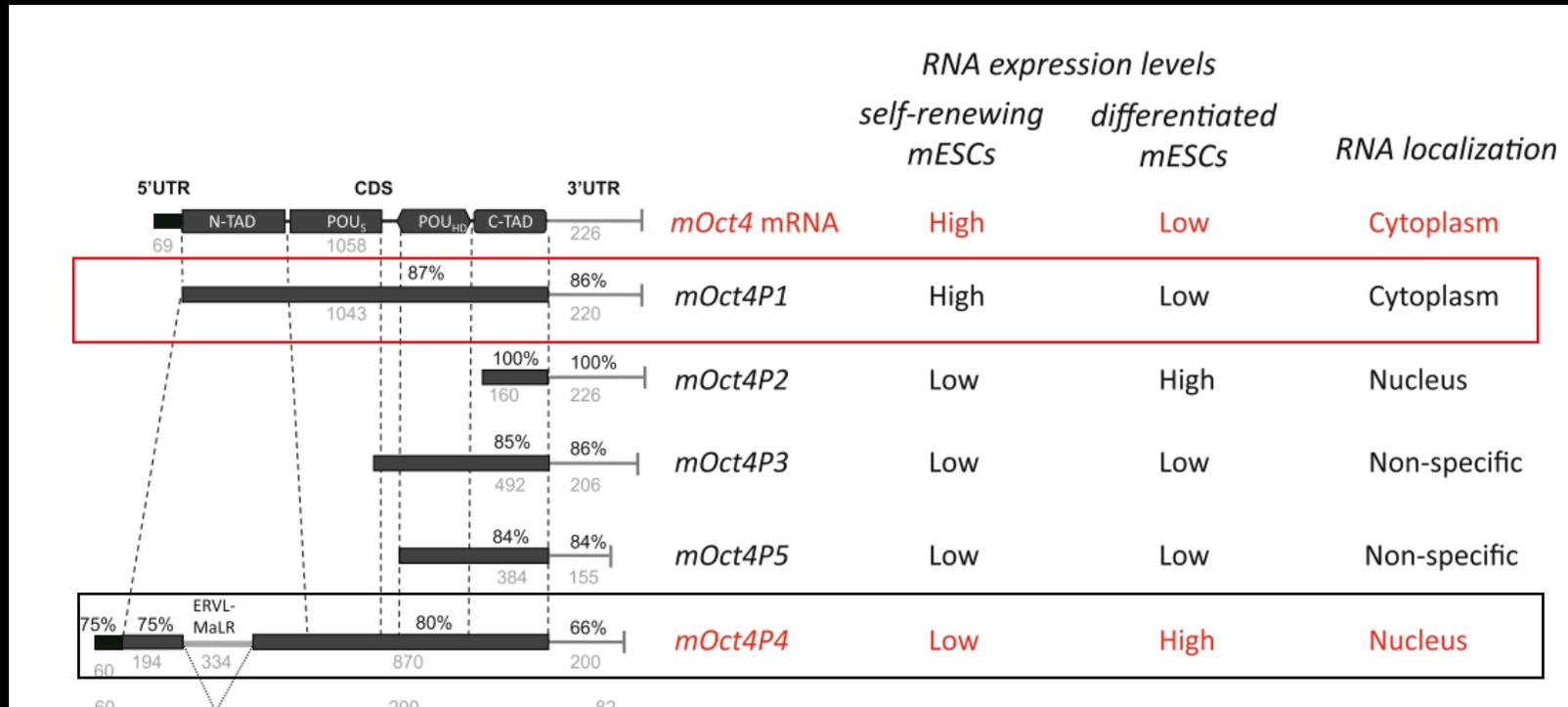


**Oct4**  
**Oct4P1**



Oct4P2  
**Oct4P4**

## 6. FUNCTIONAL ASSAY – lncRNA and self renewal



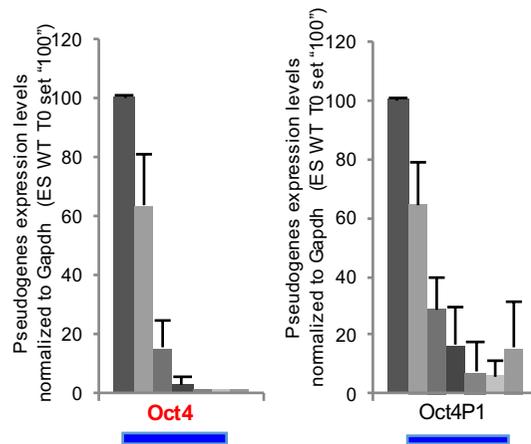
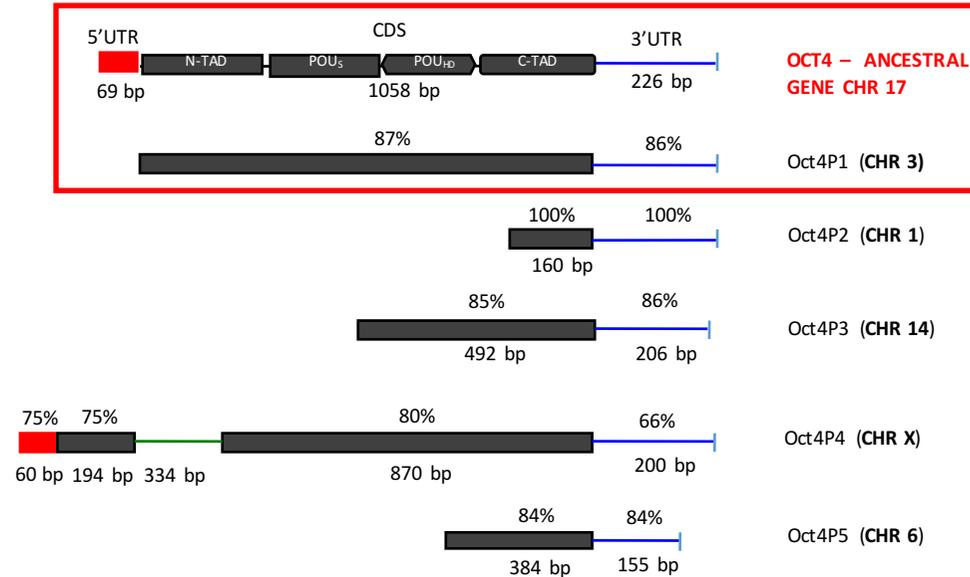
**Oct4 pseudogene derived RNAs have acquired distinct properties**

**How can we identify a biological function of pseudogene lncRNA  
A case study: Oct4P1 and Oct4P4**

# How to identify a biological function related to pseudogene derived lncRNA?



## Oct4P1: A cytoplasmatic lncRNA



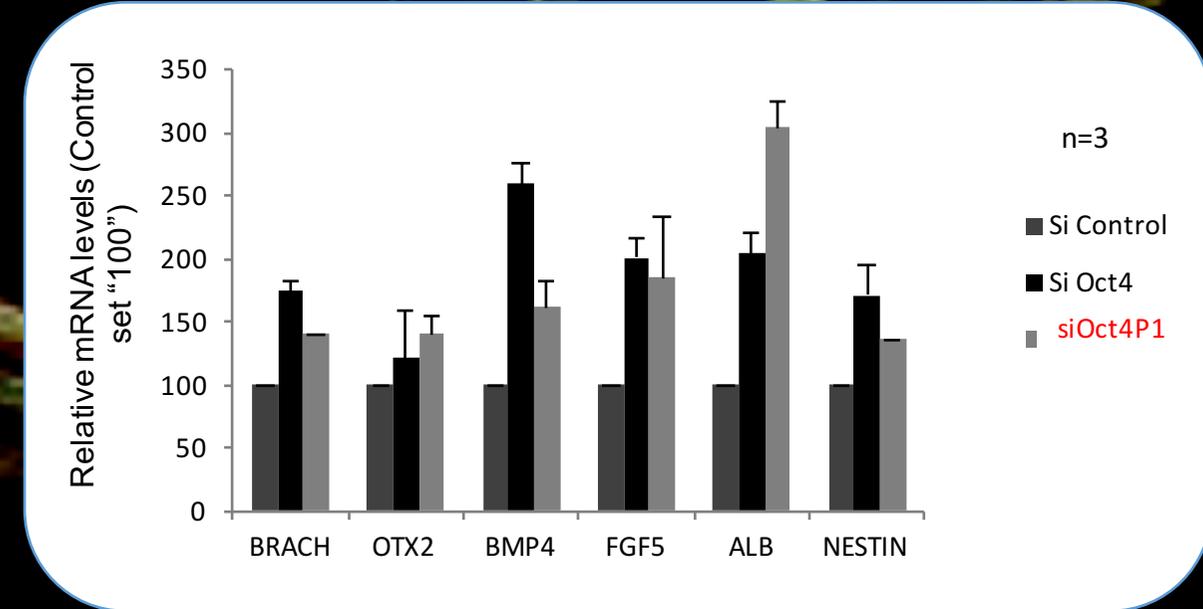
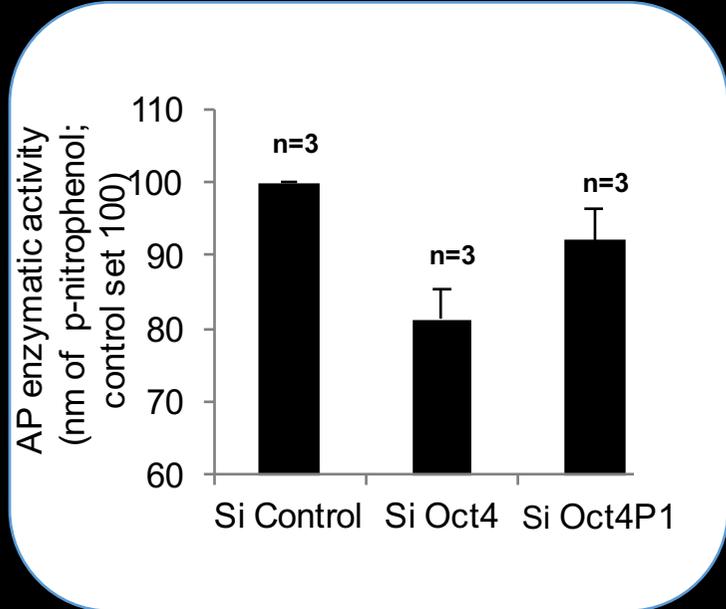
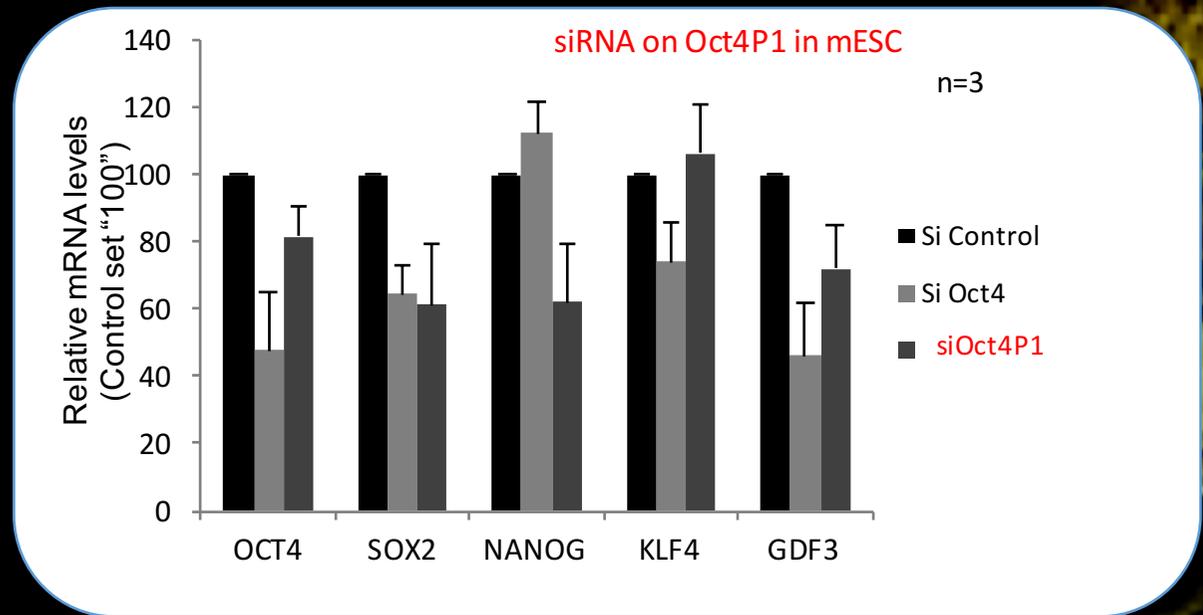
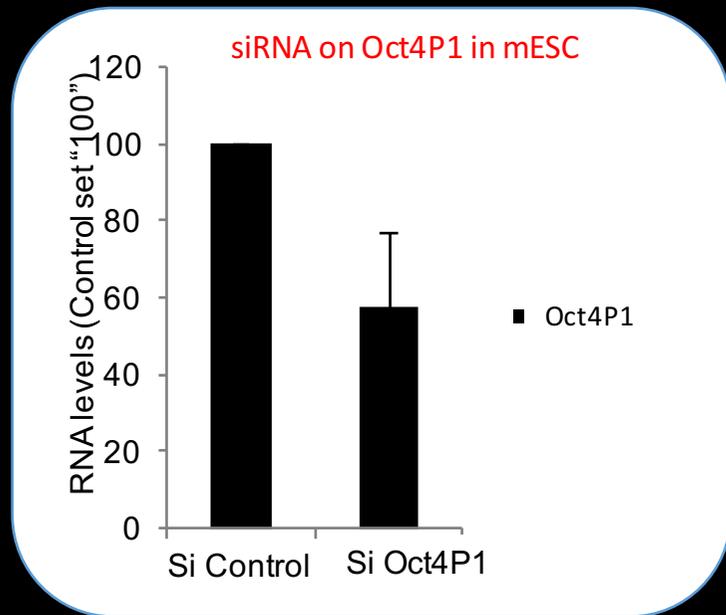
Oct4: high in ESC

Oct4P1: high in ESC, low in differentiated cells

siOct4: differentiation

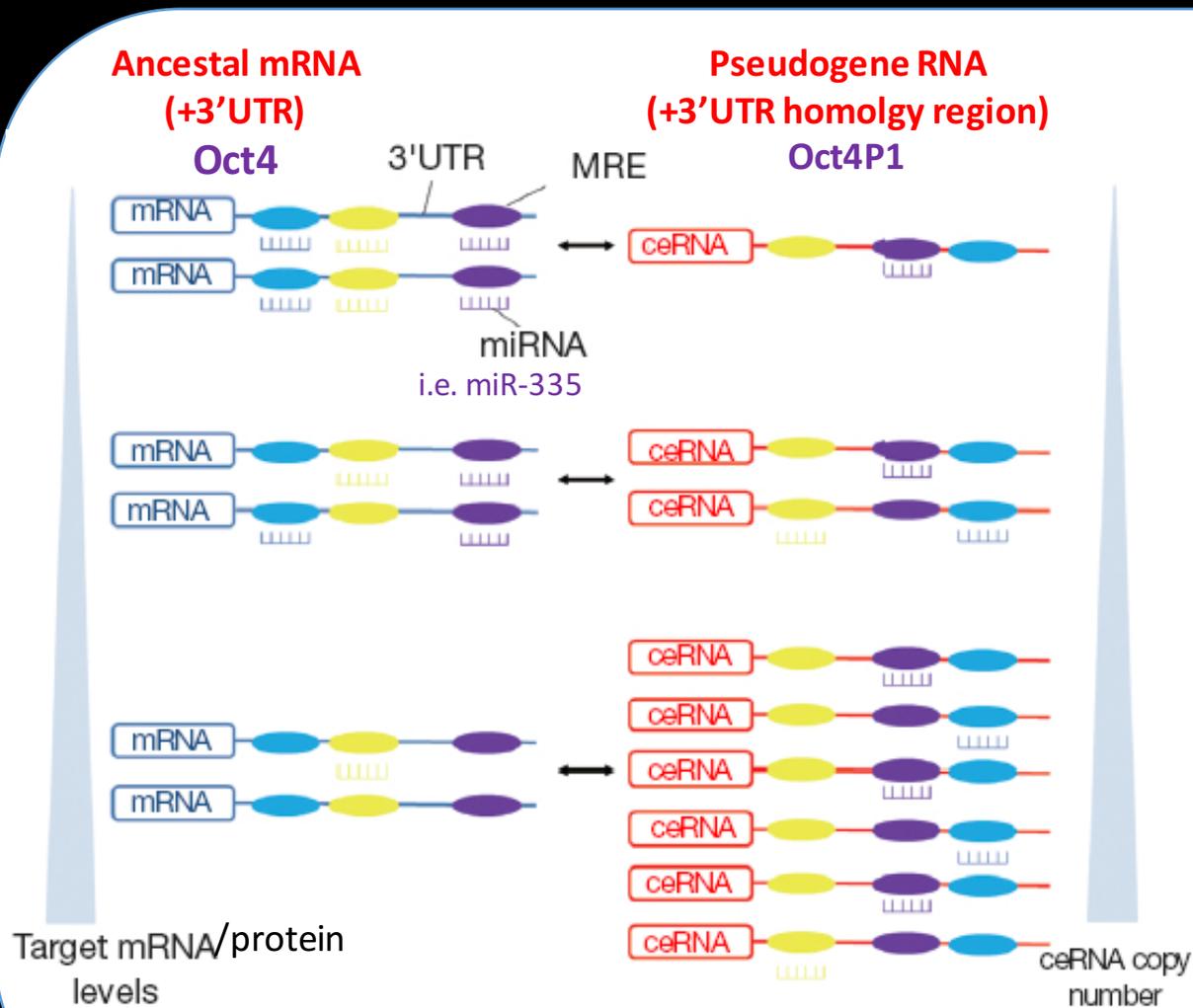
Hypothesis: Oct4P1 needed to keep high Oct4 expression in self renewing ESC??

# Oct4P1: A cytoplasmatic lncRNA



**Cytoplasmatic OCT4P1 promotes mESC self-renewal**

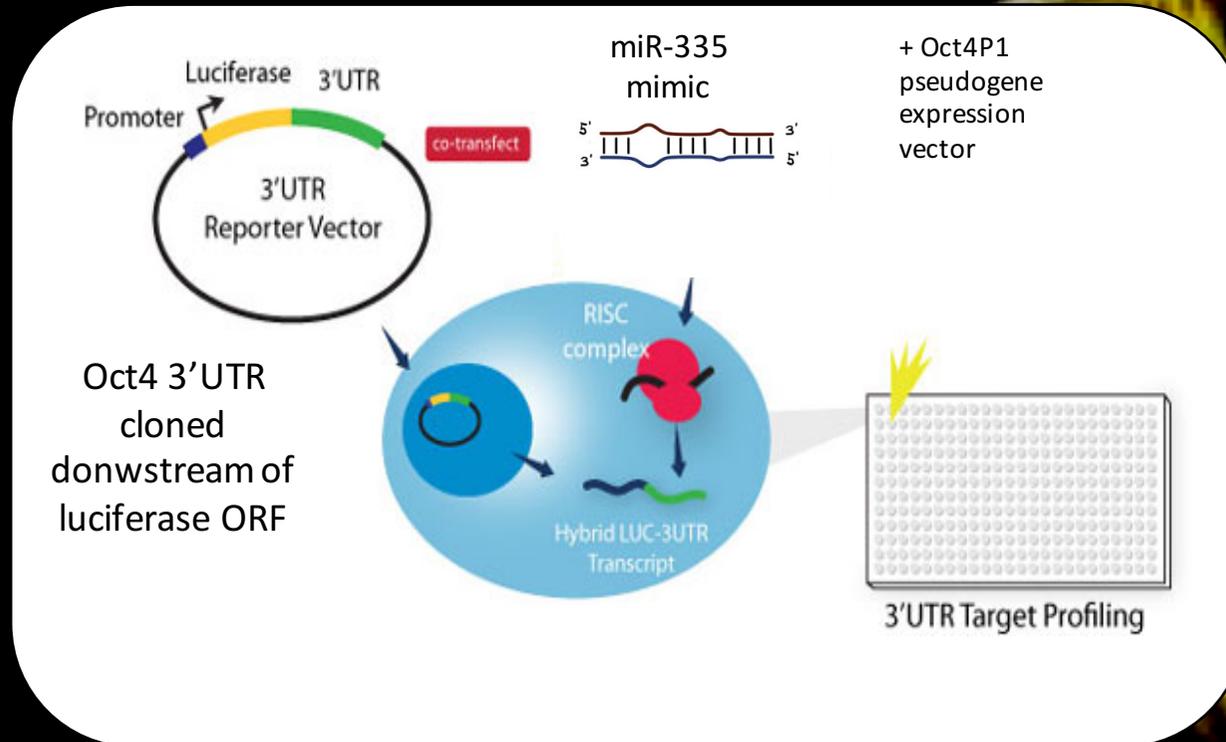
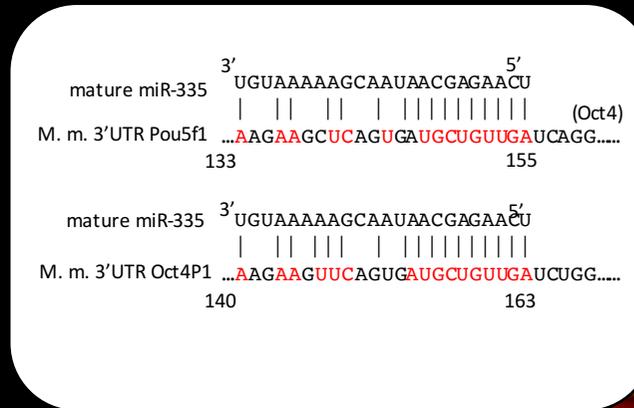
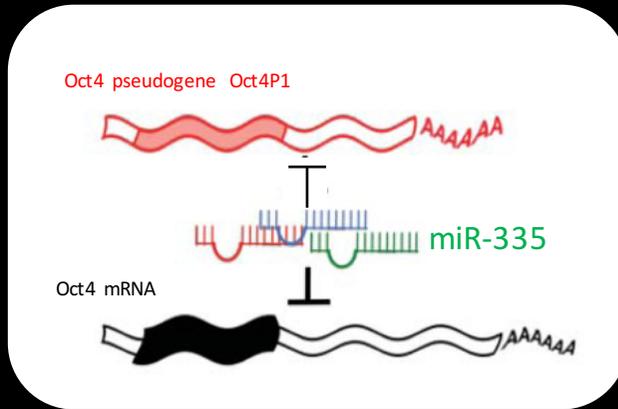
# Pseudogene sponge miRNAs that target the ancestral gene



miRNA that targets Oct4: i.e. miR-335  
MRE: miRNA response element

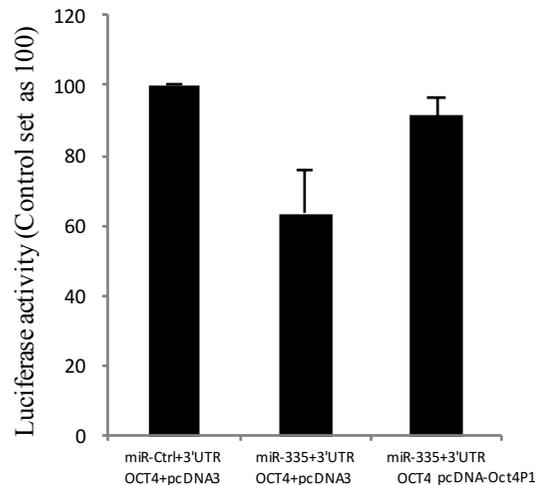
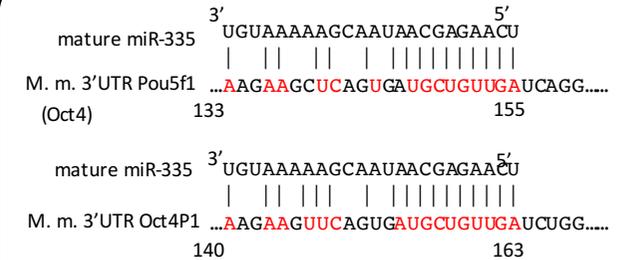
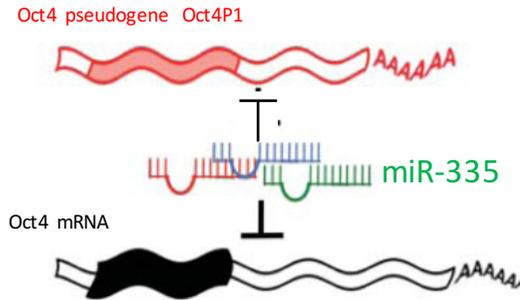
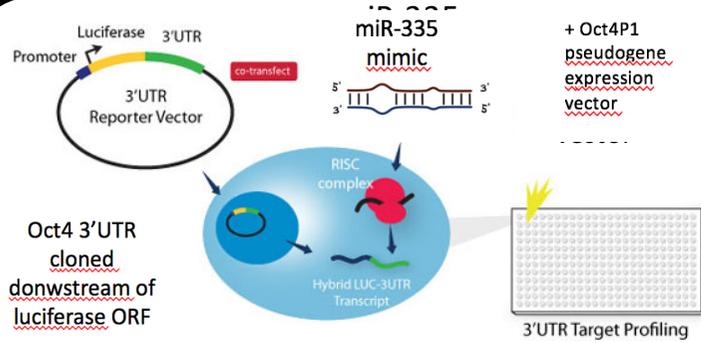
Evolution of  
pseudogenes  
to fine-tune the  
expression  
of ancestral genes

# Cytoplasmatic OCT4P1 acts as Oct4/Rb1 ceRNA

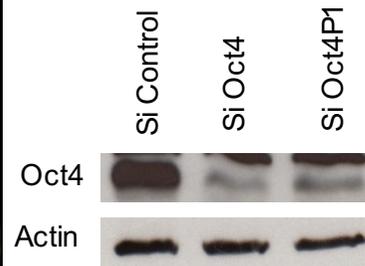


**OCT4P1 promotes self-renewal by sponging miRNAs that target Oct4**

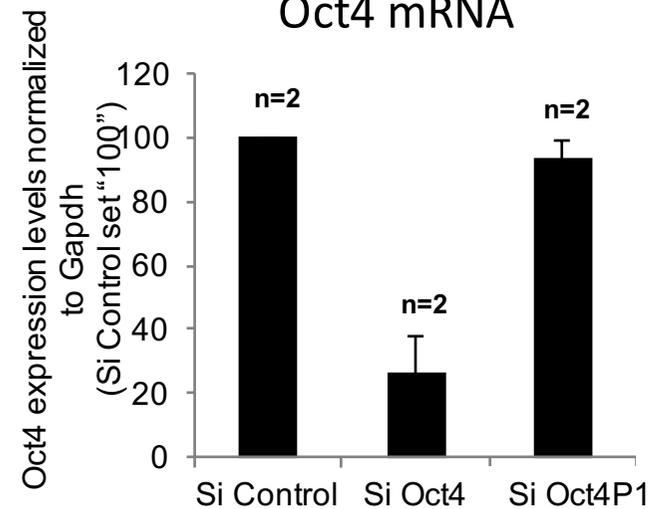
# Cytoplasmic OCT4P1 acts as Oct4 ceRNA



## Oct4 Protein



## Oct4 mRNA

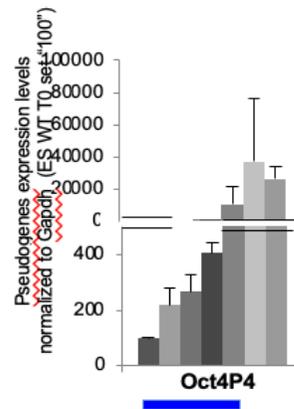
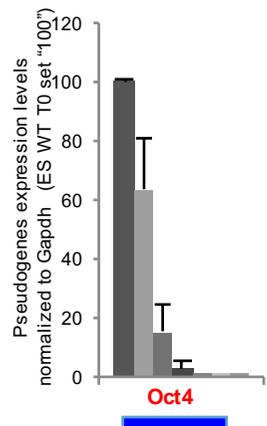
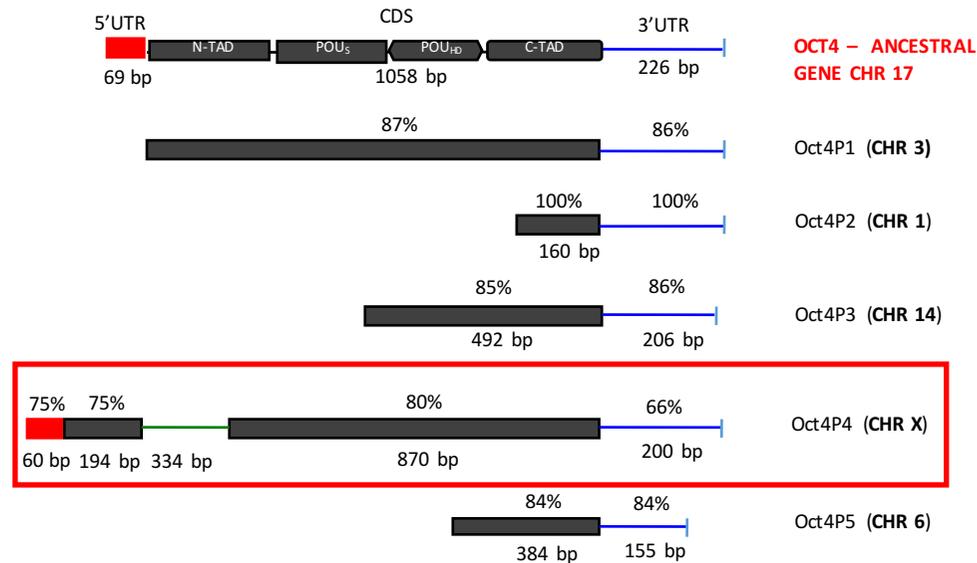


**OCT4P1 promotes self-renewal by sponging miRNAs that target Oct4**

# How to identify a biological function related to pseudogene derived lncRNA?



## Oct4P4: A cytoplasmatic lncRNA



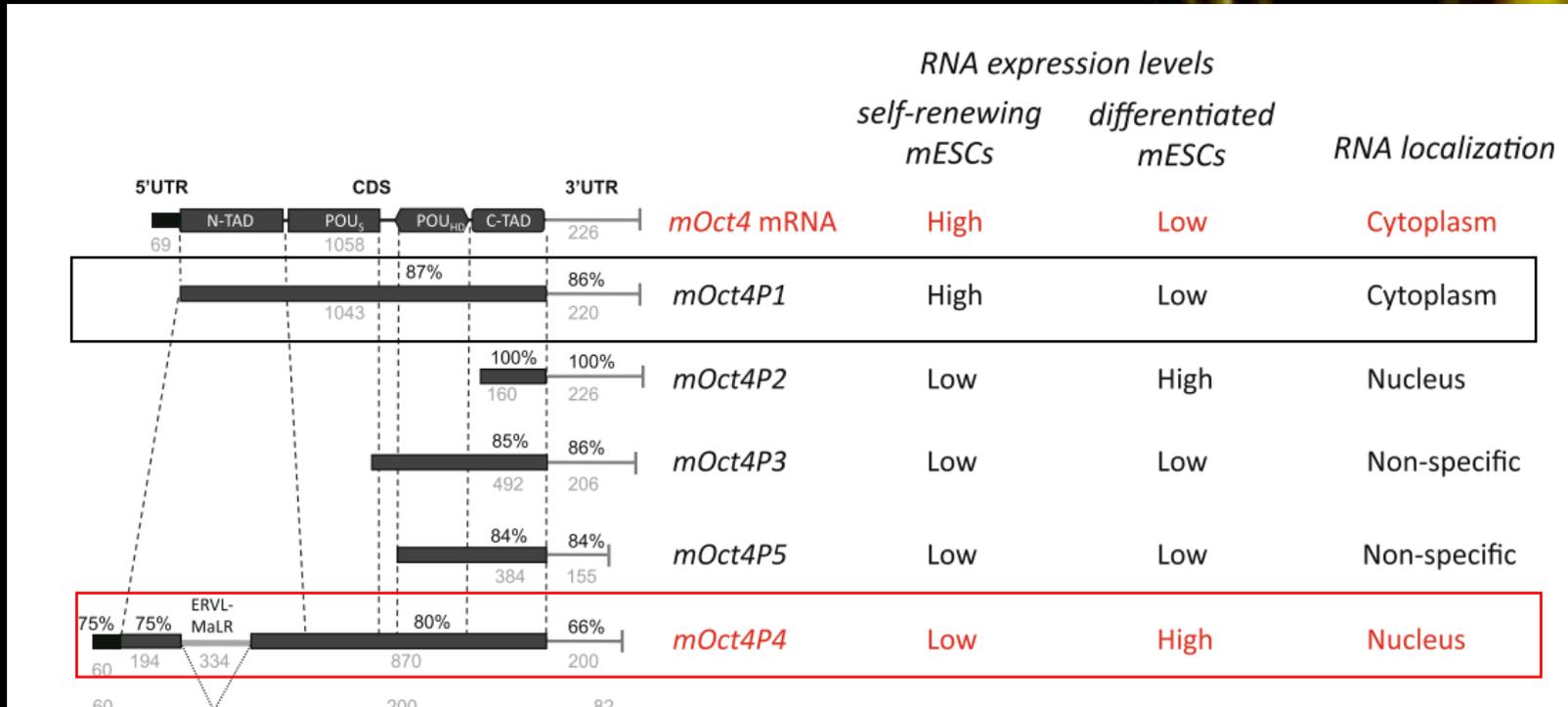
Oct4: high in ESC

Oct4P4: low in ESC, high in differentiated cells

Oct4: differentiation

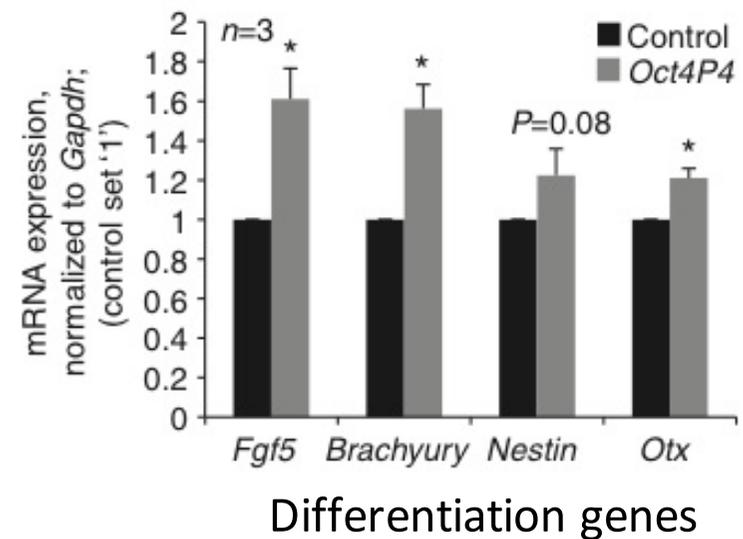
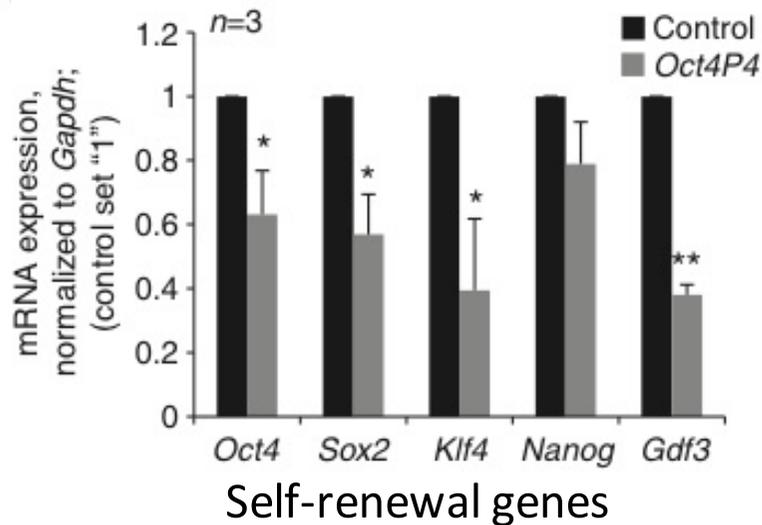
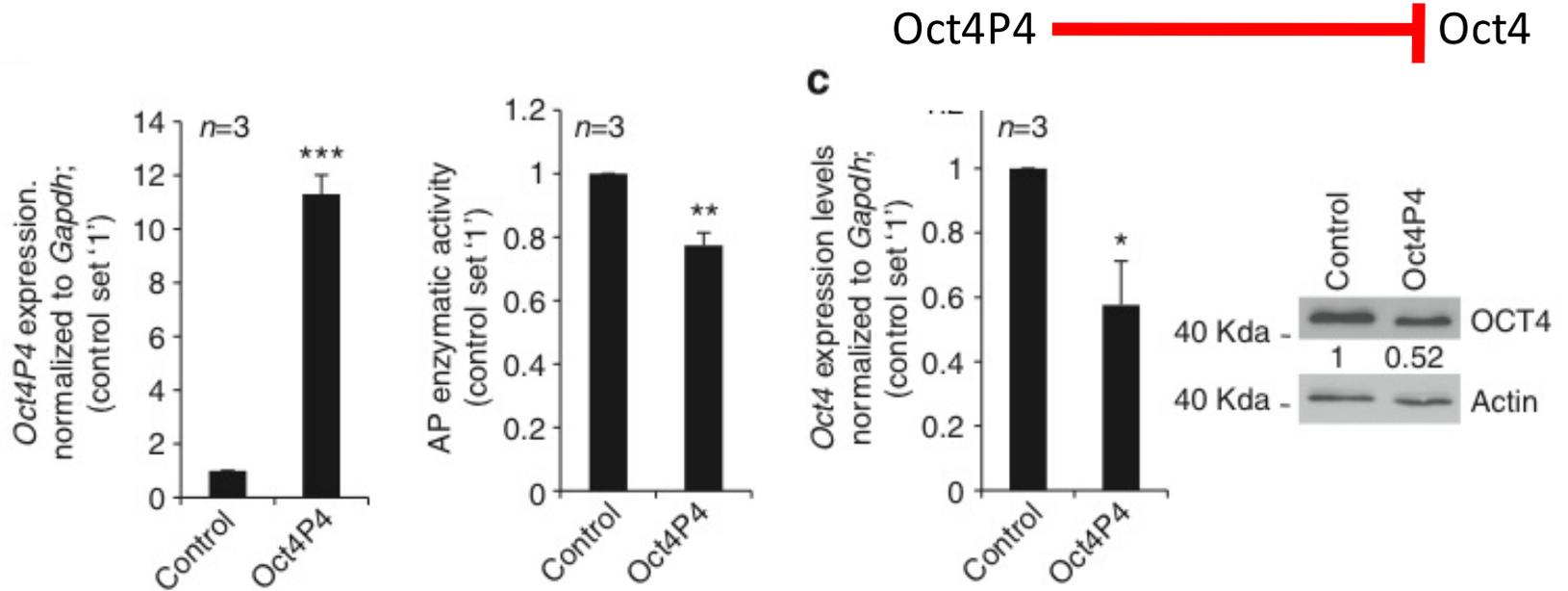
Hypothesis: Oct4P4 needed to silence Oct4 during differentiation?

# 6. FUNCTIONAL ASSAY – lncRNA and self renewal



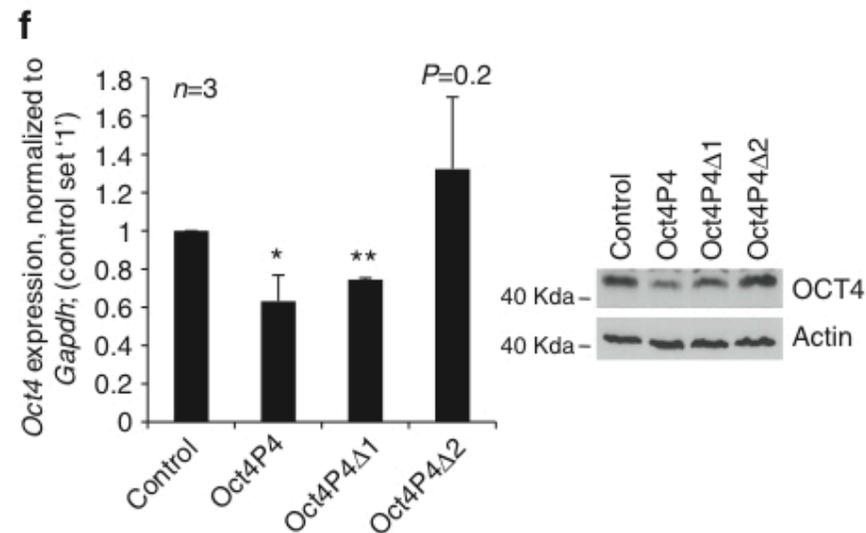
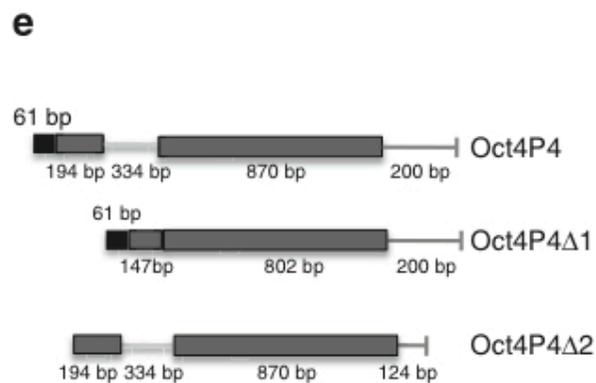
# Overexpression of OCT4P4 down regulates Oct4 to promote mESC differentiation

mESCs with stable *mOCT4P4* expression construct

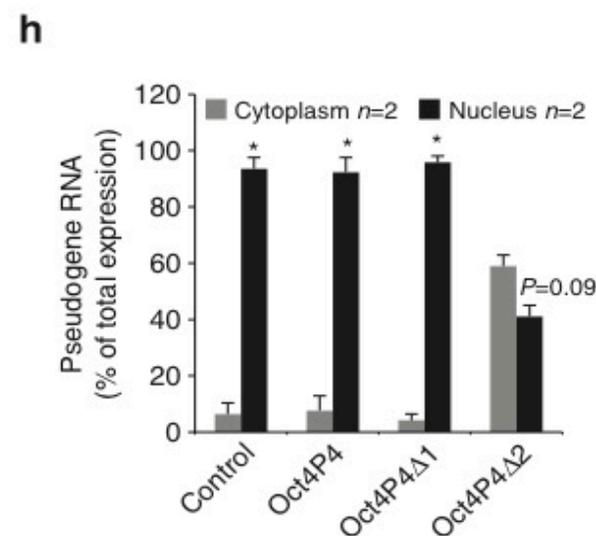
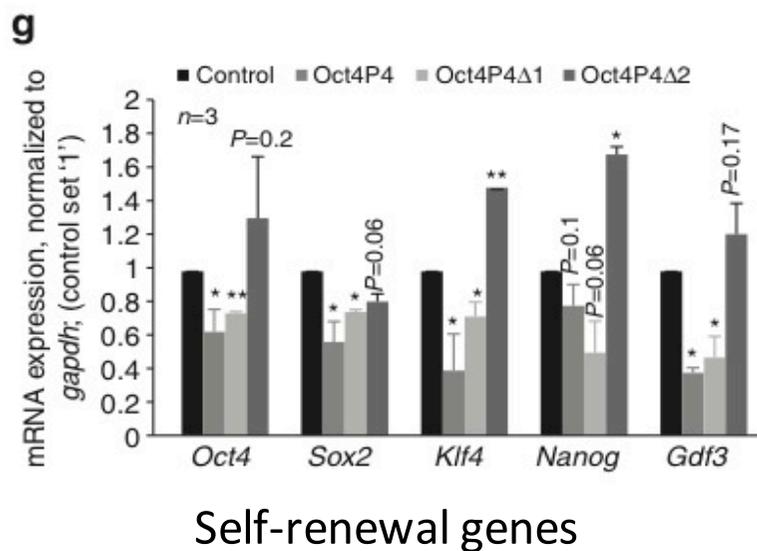


# 7. MOLECULAR ANALYSIS: WHAT ARE THE RELEVANT SEQUENCES??

Express Oct4P4 deletion constructs and follow parental gene expression



Express Oct4P4 deletion constructs and follow self-renewal programs



5' and 3' UTR homology regions are required to repress self-renewal marker genes

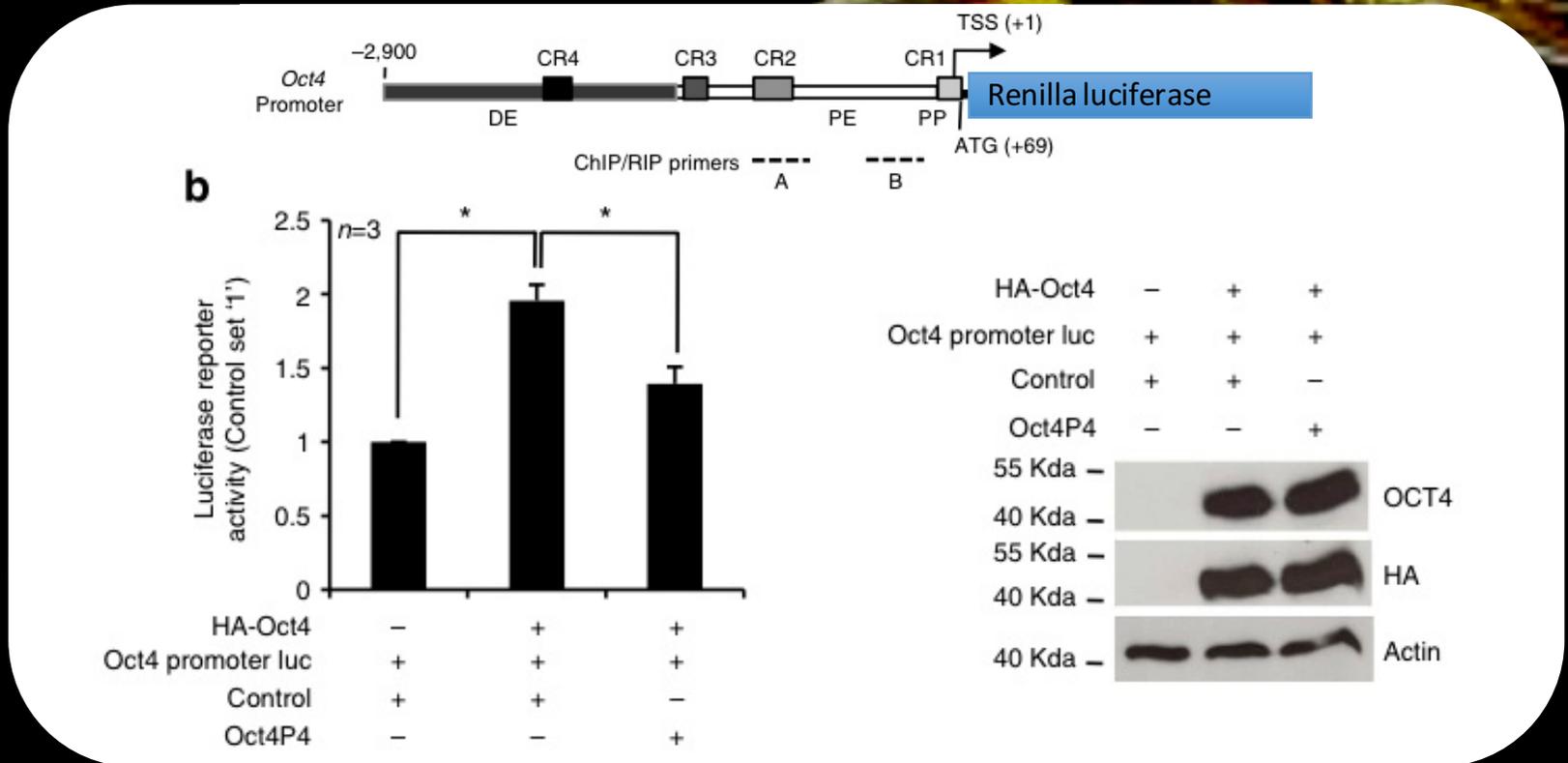
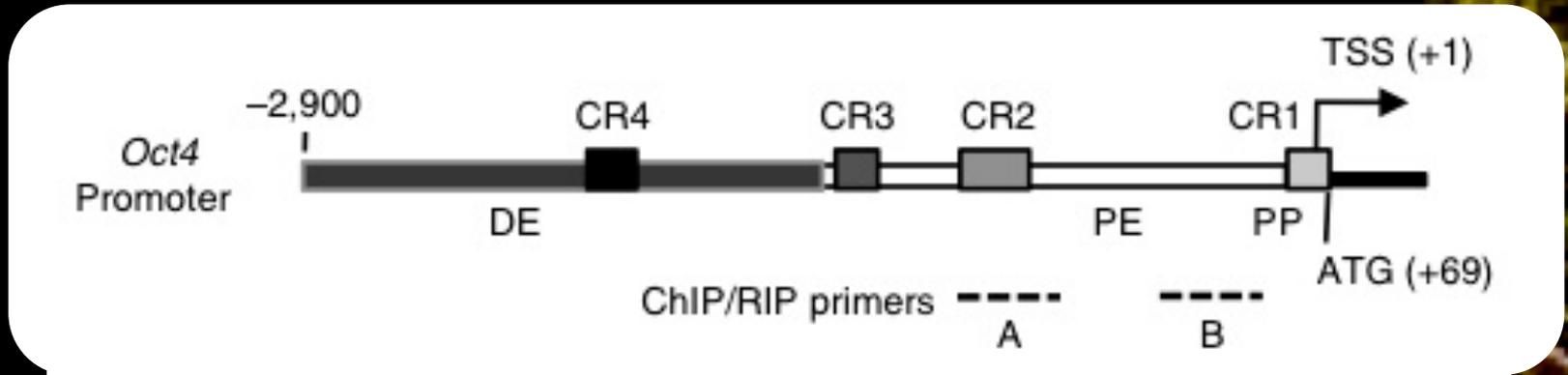
# 8. MOLECULAR ANALYSIS: USE SIMPLIFIED MODEL TO GET INDICATION ON MOLECULAR MECHANISM??

Oct4P4 is a nuclear RNA that suppresses Oct4 transcription

**Hypothesis:**  
transcriptional suppression of Oct4 promoter?

Use simplified models to support hypothesis:

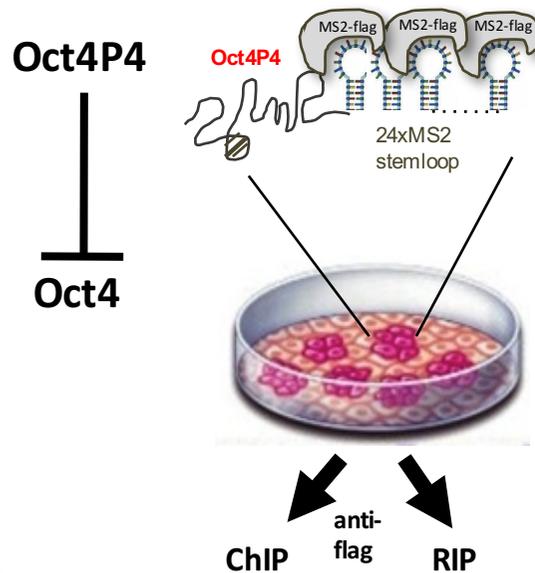
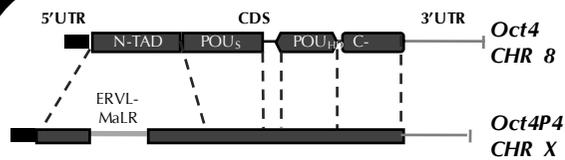
- Clone Oct4 promoter
- Clone luc reporter downstream
- Test whether Oct4P4 lncRNA suppresses reporter



**Oct4P4 may interfere with the ancestral Oct4 promoter**

# 9. MOLECULAR ANALYSIS: DEMONSTRATE MEACHINISM AT ENDOGENOUS Oct4 PROMOTER??

- Generate model system to:
1. study Oct4P4 lncRNA localization
  2. study chromatin at OCT4 promoter



Step 1: stably transfect mESCs with MS2-flag

Step 2: stably transfect mESCs with an expression vector encoding mOct4P4 lncRNA fused to 24 MS loop stem loops

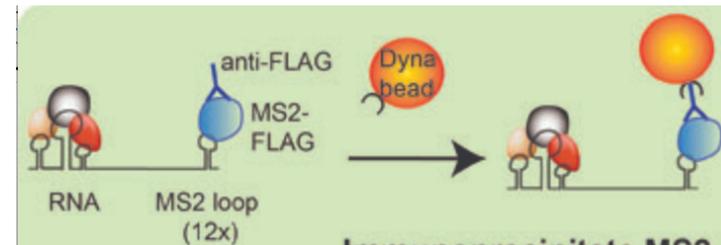
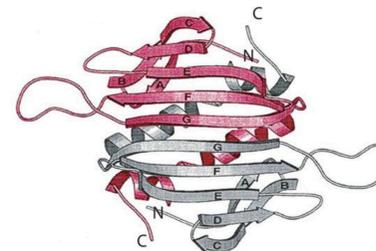
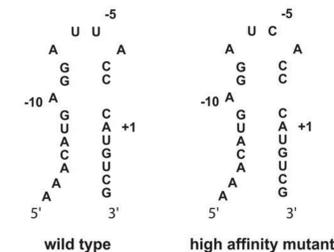


Diagram of MS2 coat protein dimer



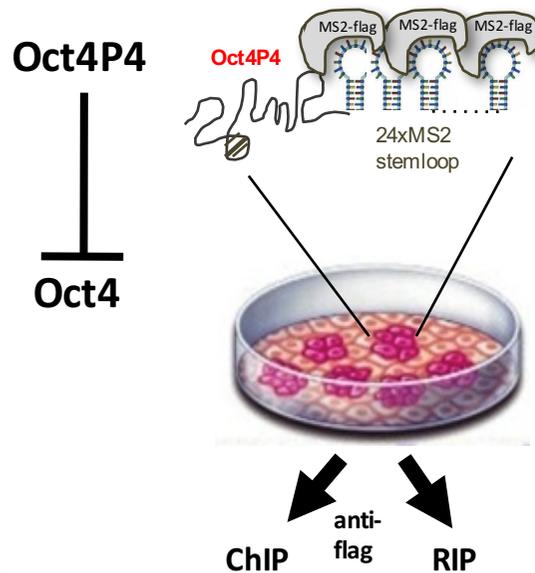
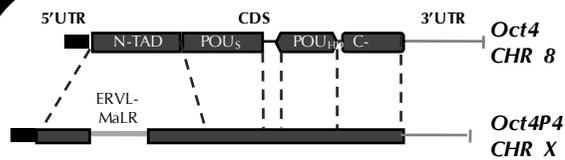
MS2 stem loop sequence



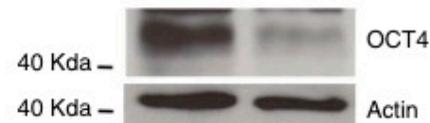
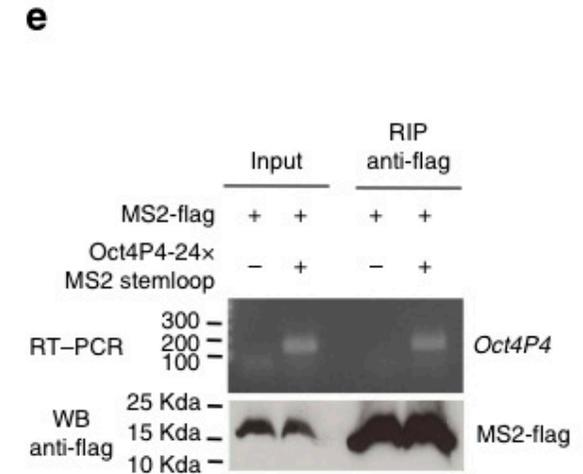
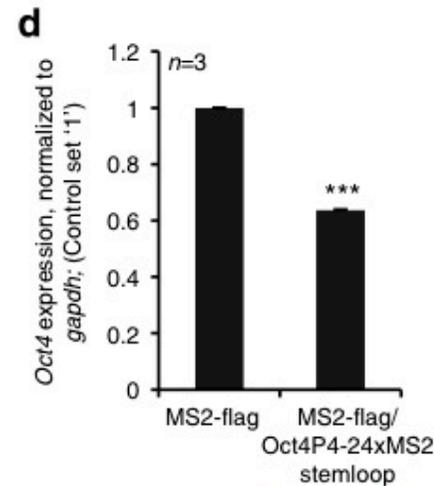
MS2 stem loop tagged Oct4P4 co-expressed with flag-MS2 in mESCs

# 9. MOLECULAR ANALYSIS: DEMONSTRATE MECHANISM AT ENDOGENOUS Oct4 PROMOTER??

- Generate model system to:
1. study Oct4P4 lncRNA localization
  2. study chromatin at OCT4 promoter

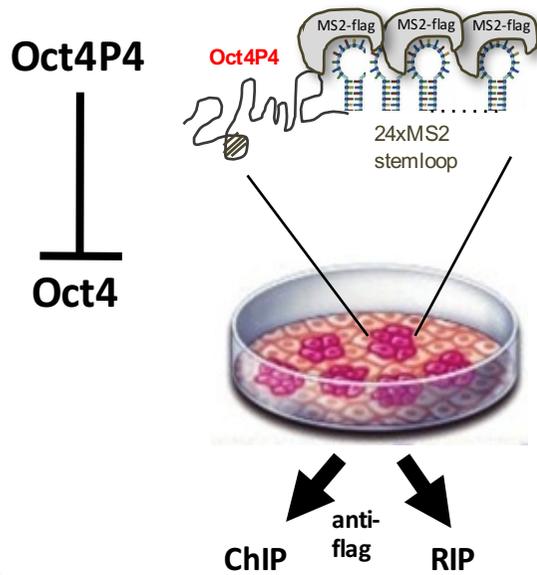
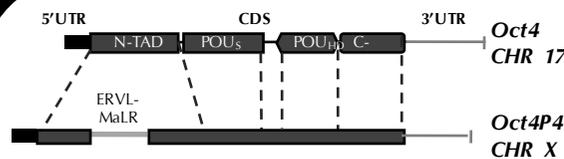
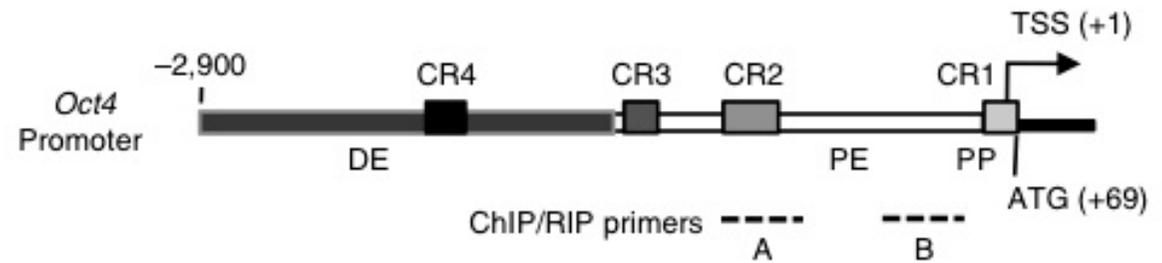


RNA immunoprecipitation anti-flag then RT-PCR for Oct4P4 lncRNA

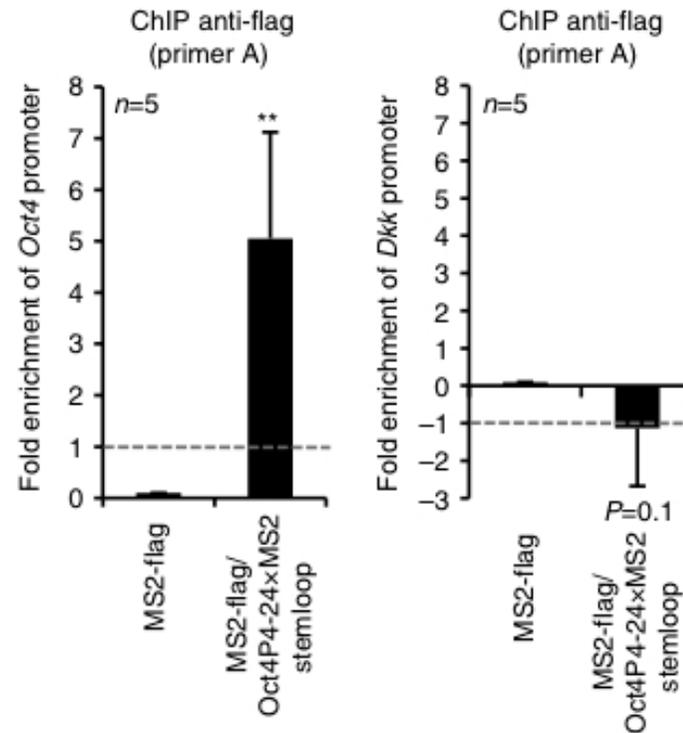


MS2 stem loop tagged Oct4P4 co-expressed with flag-MS2 in mESCs

# A model system to study Oct4P4 lncRNA localization



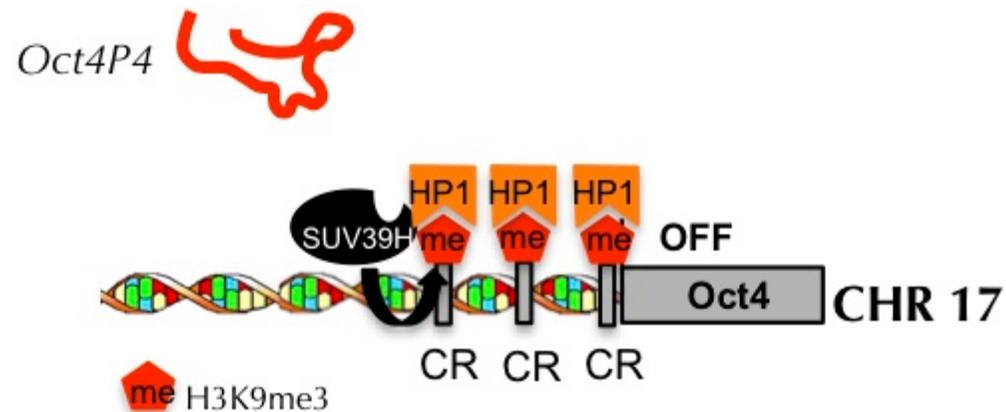
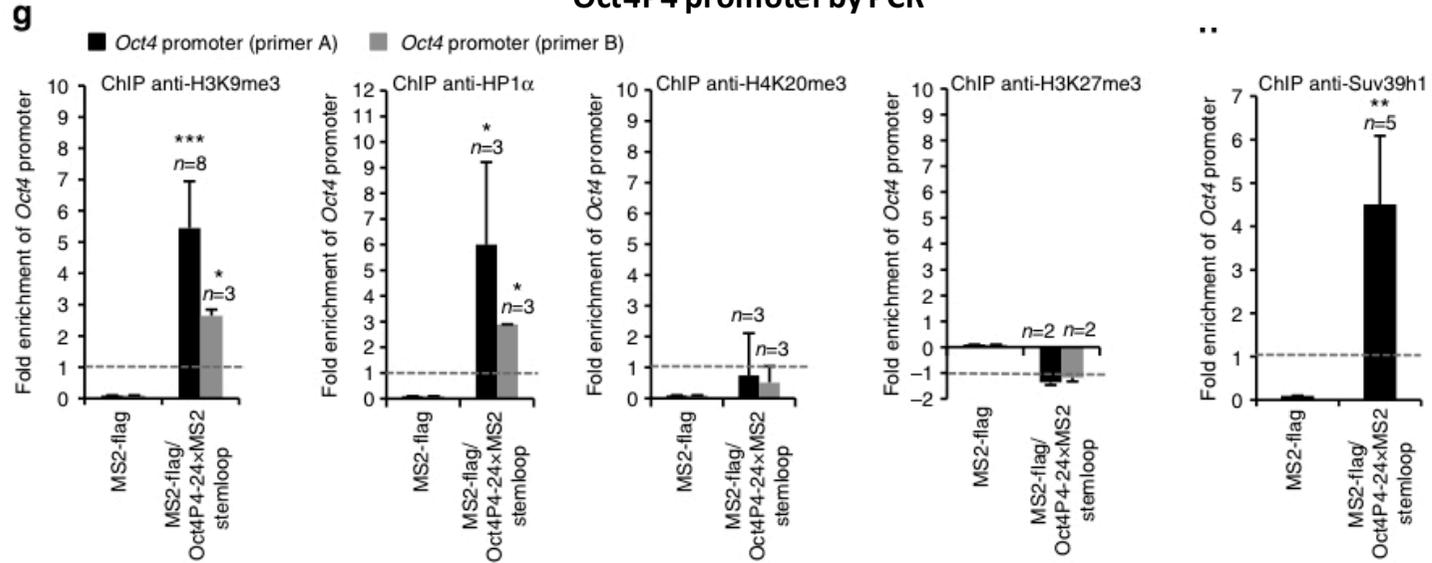
ChIP using anti-flag then use the immuno-precipitate to detect the Oct4P promoter by PCR



Oct4P4-MS2 lncRNA localizes to Oct4 promoter

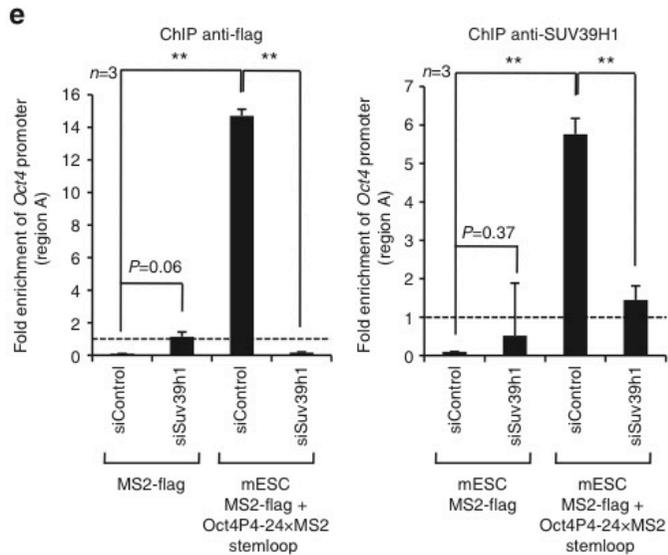
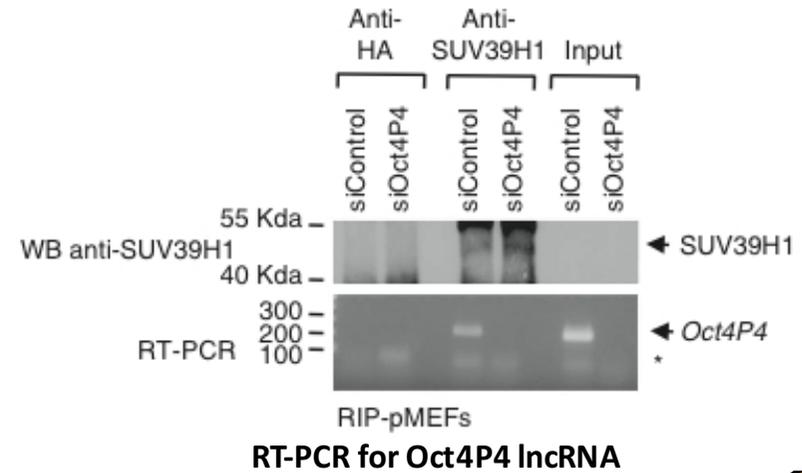
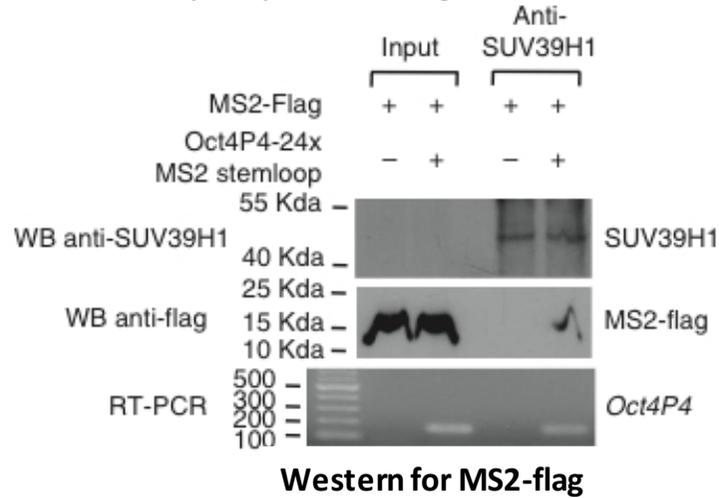
# Oct4P4-MS2 directs Suv39h1 to Oct4 promoter

ChIP using specific antibodies then use the immuno-precipitate to detect the Oct4P4 promoter by PCR

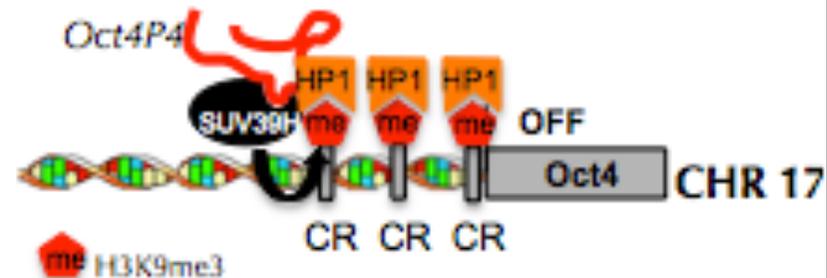


# Oct4P4-MS2 directly interacts with Suv39h1

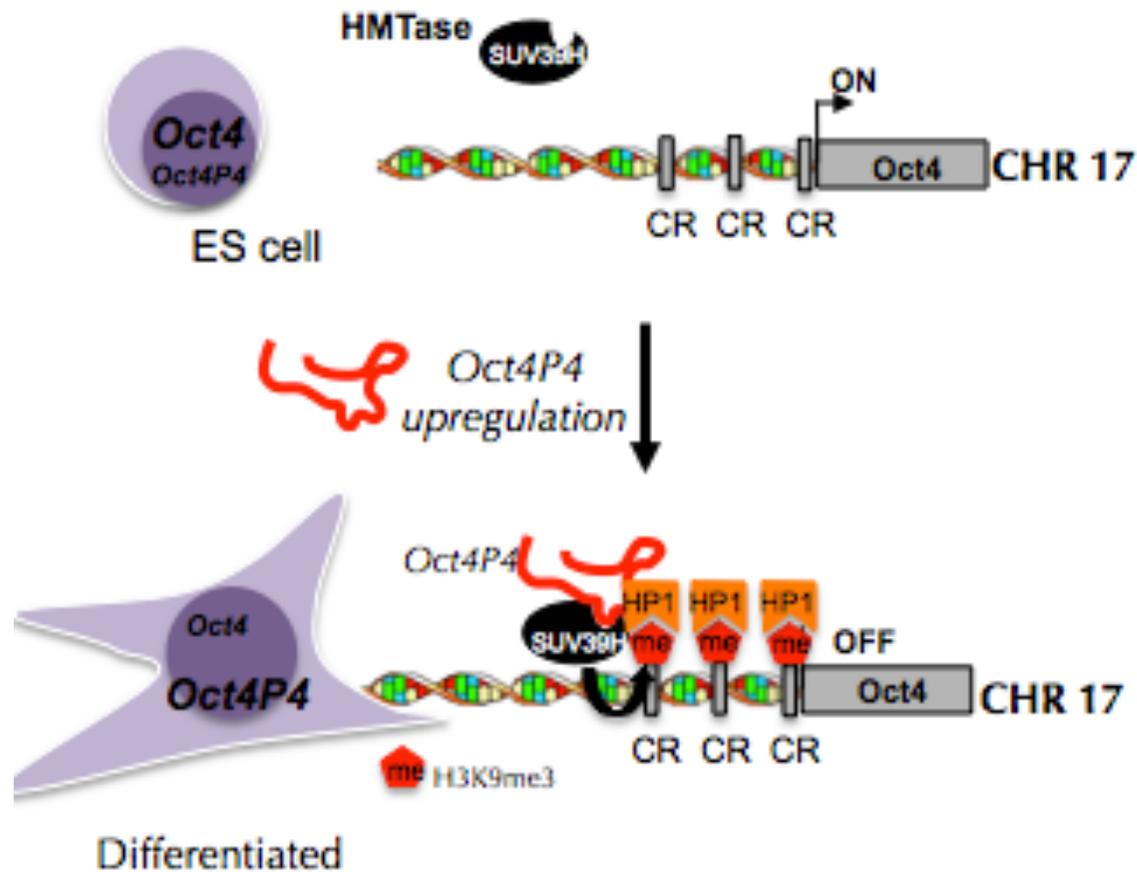
RNA immunoprecipitation using anti-SUV39h1;



ChIP using specific antibodies then use the immuno-precipitate to detect the Oct4P4 promoter by PCR



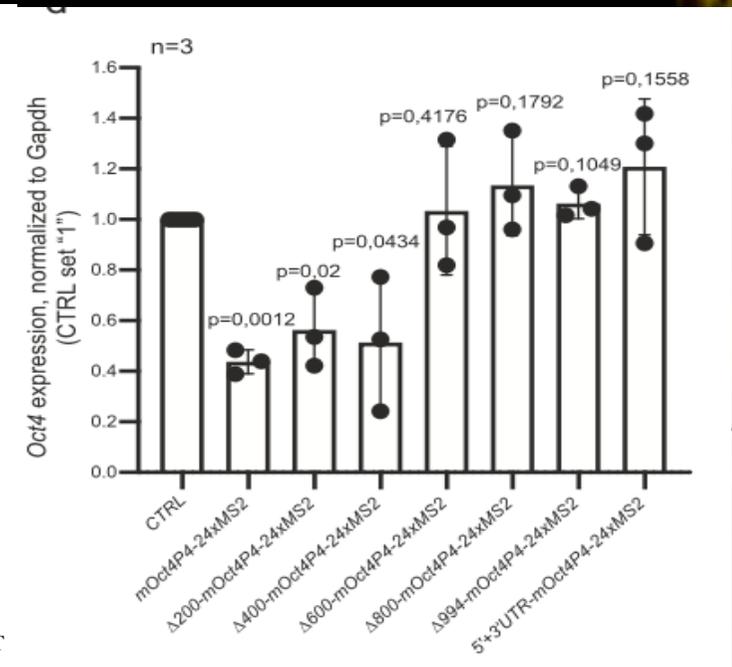
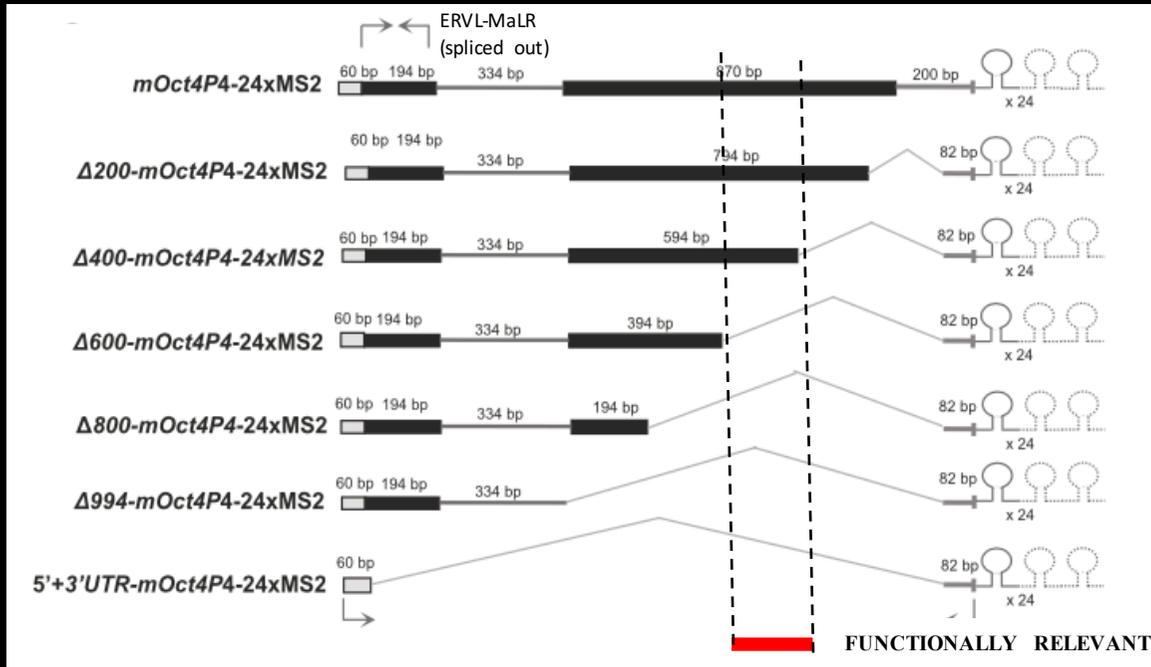
# Oct4P4-MS2 recruits Suv39h1 To direct silencing of the Oct4 promoter



**FUNCTIONAL  
RELEVANCE????**

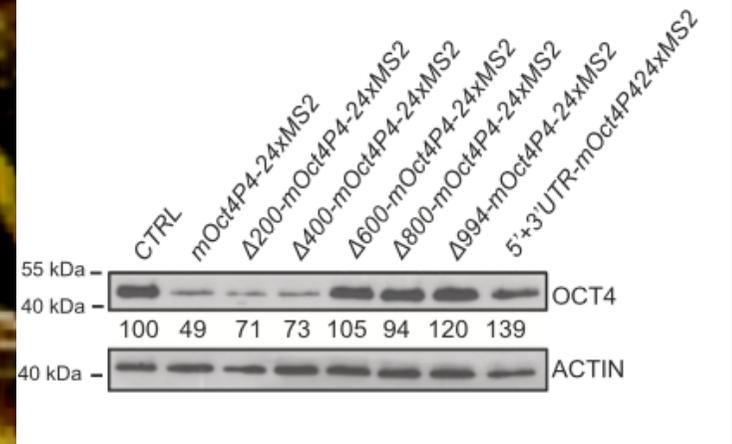
**REVERSIBILITY????**

# A detailed deletion analysis to identify functionally relevant RNA sequences



- Delete regions of Oct4P4
- Express in mESCs
- Monitor the expression of Oct4 on the protein and mRNA level

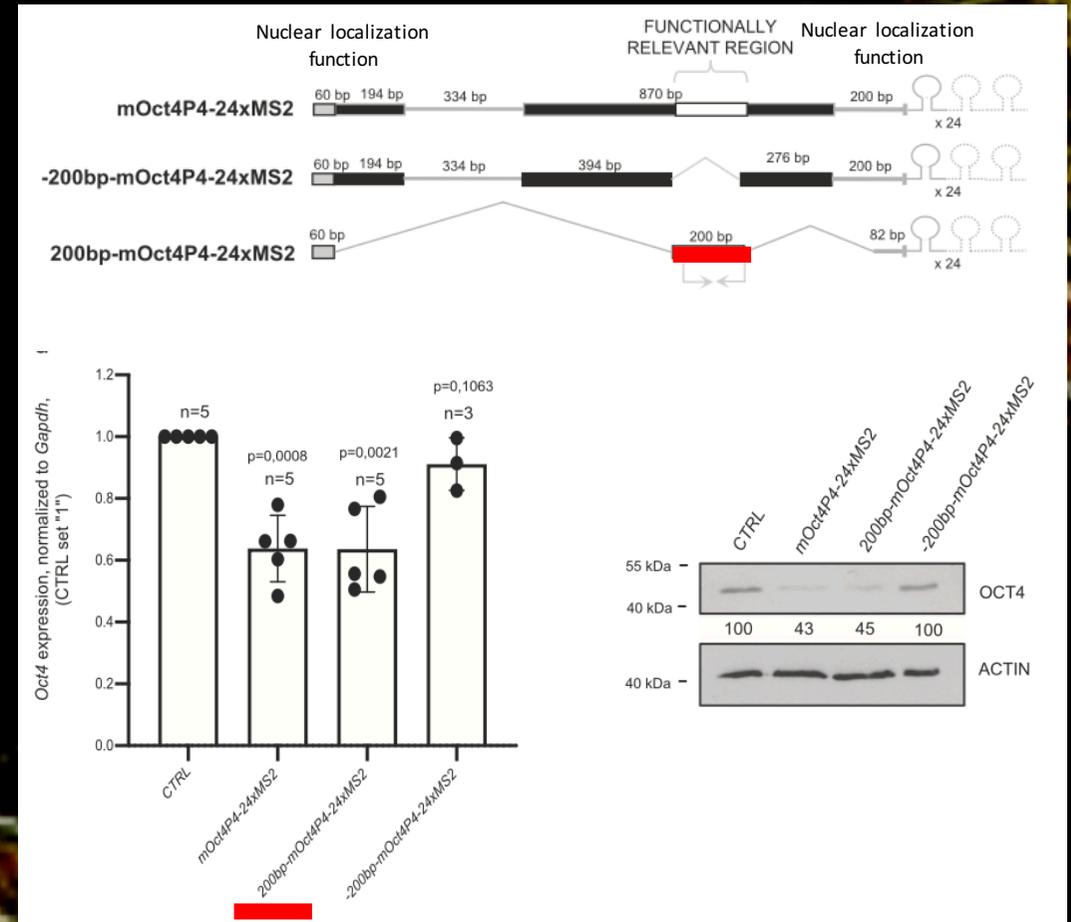
A 200 nucleotide element in *mOct4P4* is functionally relevant for *mOct4P4* function



# Validate that the 200 nt region is functionally relevant

- Generate a construct that lacks the 200 nt region
- Generate a construct only consists of the 200 nt region + sequences that determine nuclear localization
- Monitor expression of Oct4 on the protein and RNA level

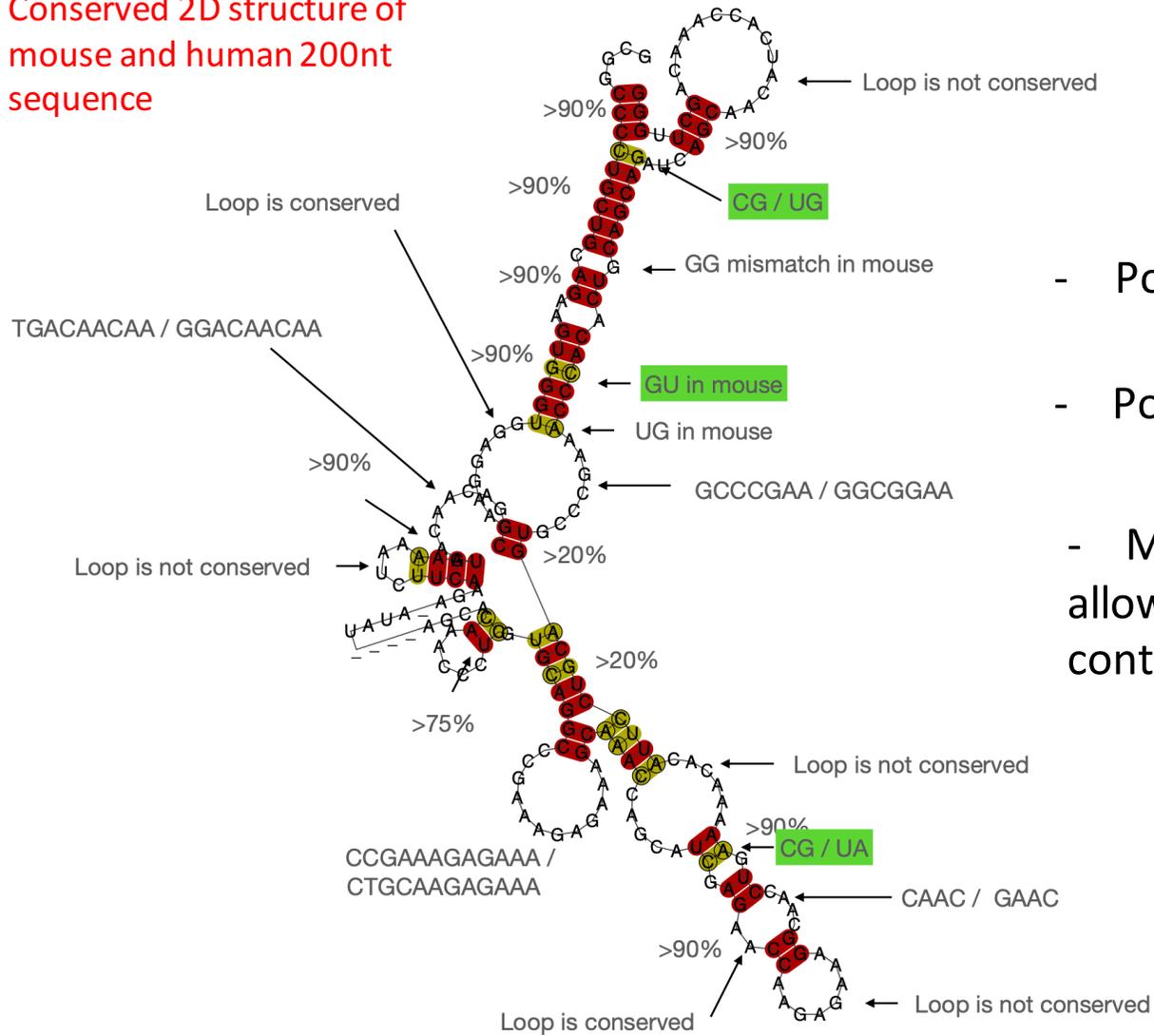
→ The 200nt motif is sufficient to reduce OCT4 expression



A 200 nt RNA sequence recapitulates the function of full length Oct4P4

## Next step: Structure function analysis of 200 nucleotide region of OCT4P1

Conserved 2D structure of mouse and human 200nt sequence



- Position of SUV39H1 interaction
- Position required for targeting of OCT4 promoter
- Modulation of ncRNA to target in order to allow targeting of other gene promoters in a controlled manner