



UMR7216 Epigénétique &  
Destin Cellulaire



## Introduction à la bioinformatique

atelier du 07/12/2022

Magali Hennion  
Olivier Kirsh



[bibs@parisepigenetics.com](mailto:bibs@parisepigenetics.com)

# Les sciences « omiques » étude globale

génomique

(épi)génomique

métagenomique

**ADN**

*Transcription*

Pré-ARNm

transcriptomique

**ARN messenger**

protéomique

*Traduction*

**Protéines**

séquençage

séquençage

spectrométrie  
de masse

RMN

métabolomique

**Métabolisme, constituants cellulaires...**

# Séquençage ADN

Qu'est-ce que le séquençage d'ADN?

→ Déterminer l'ordre des bases d'un fragment d'ADN

ATGCAGCGTTACCATG...

Comment ?

Depuis 2005, séquençage de 2<sup>de</sup> génération

-> très haut débit

## Après le séquençage

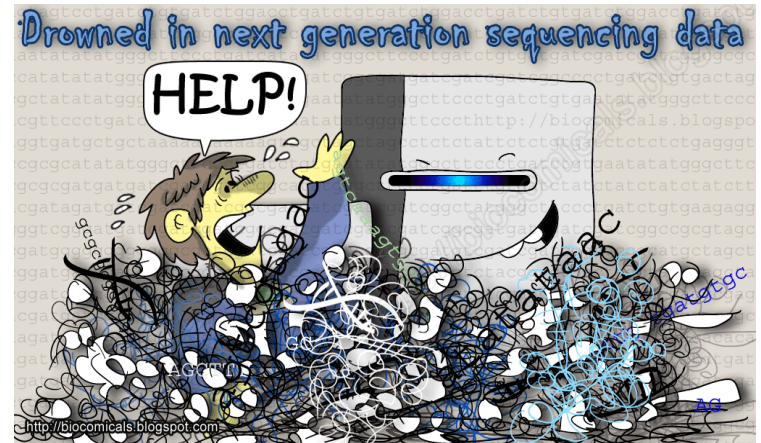
# Que faire avec les données ?



ExperimentX\_R1.fastq.gz

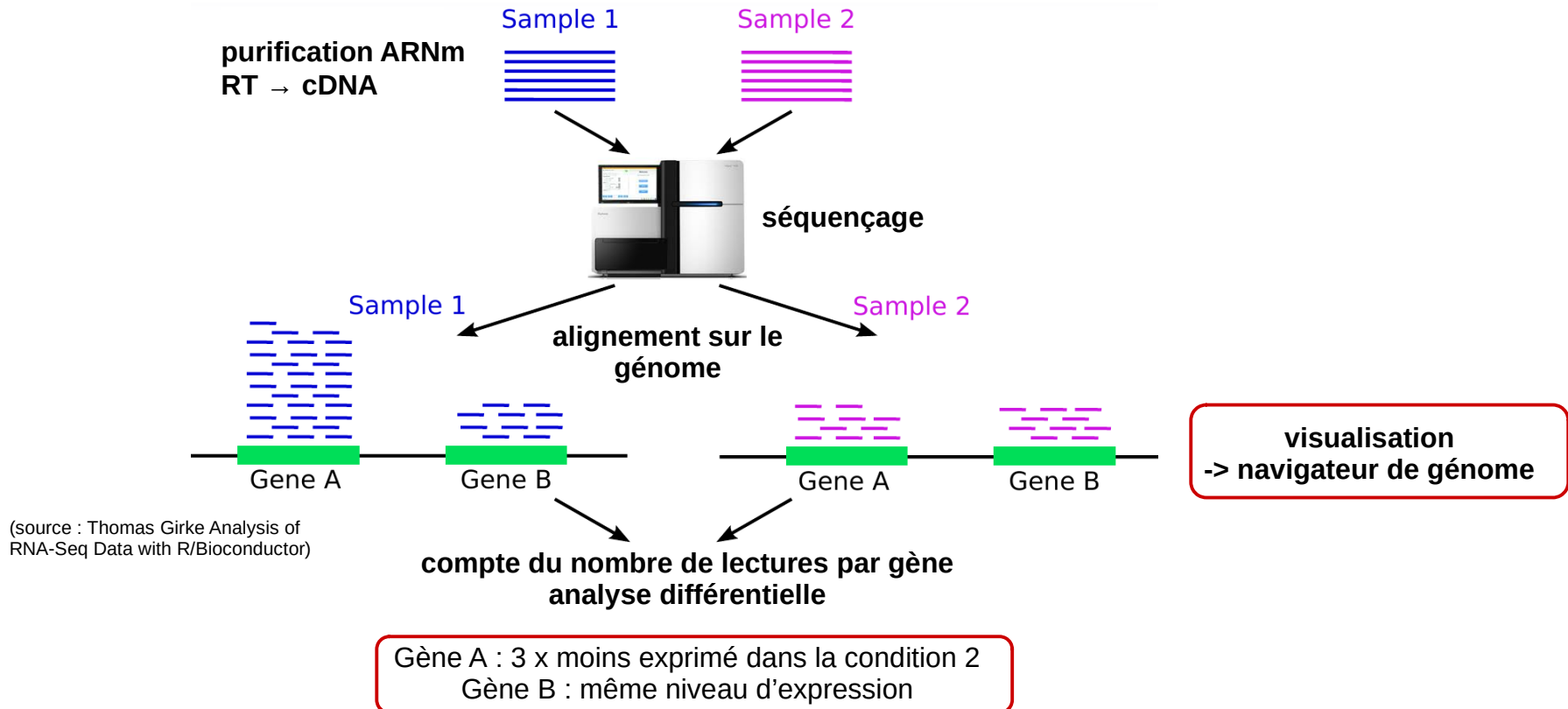
@HWI-ST980:107:D098EACXX:2:1101:1630:1986 1:N:0:TGACCA  
NTGGATCTGAGCAGTGATTTGAGGAAGTCTGCACAGCACAAACTGCTACAAATACAATAGTAATCCCCATTTGTGGTACAATGCGAAGTCACGTCAACT  
+  
#1=DDFFFHHHHHHJHGIIJJGIIJJIGIJJIIJJJJGIIJJIIJJIIHJJIIHHIIJJGIIJJGIIIDHHHIIJJJHFEFHFFFFFEDDDDDDDDDDDDDCDDD

Fichier avec des dizaines ou centaines de millions de séquences...



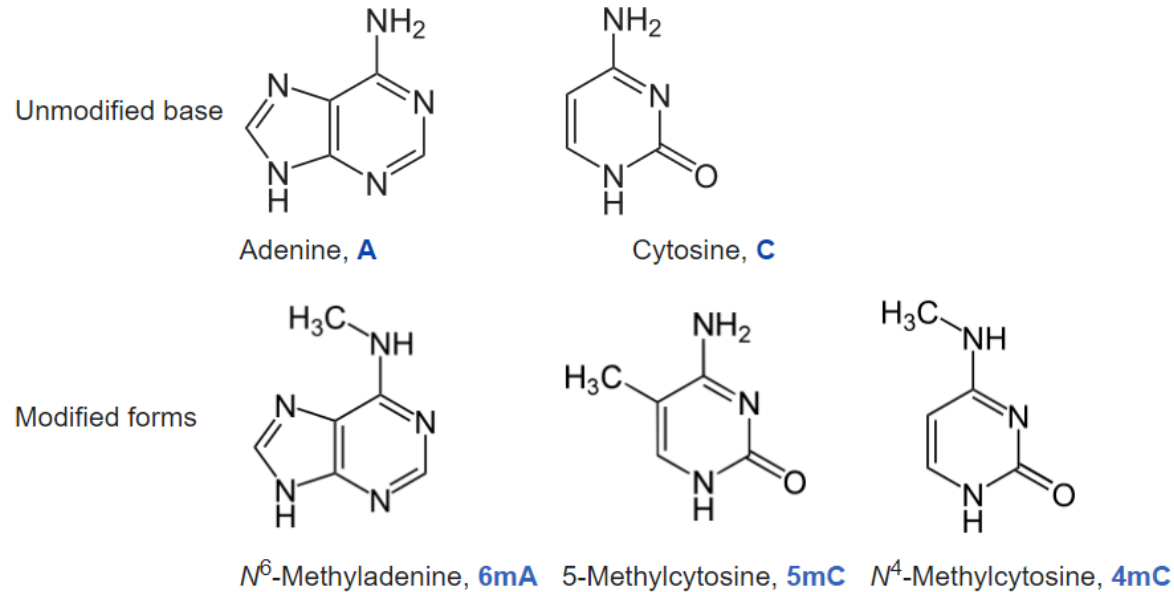
# Transcriptomique

Quantification de l'ensemble des transcrits dans différentes conditions ou différents types cellulaires



# Épigénétique

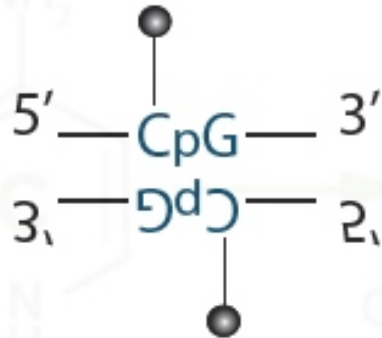
Étude à l'échelle du génome des marques épigénétiques  
-> exemple de la méthylation de l'ADN



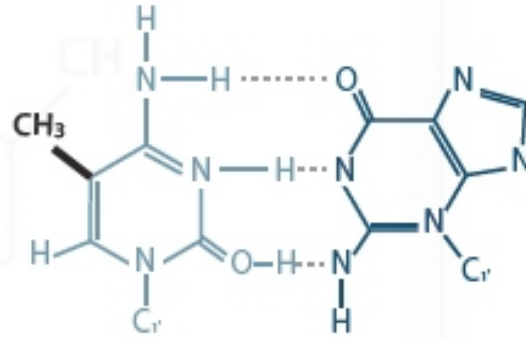
# Méthylation des cytosines

Mammifères : CpG

Autres organismes : CHG ou CHH (H = A, T ou C)



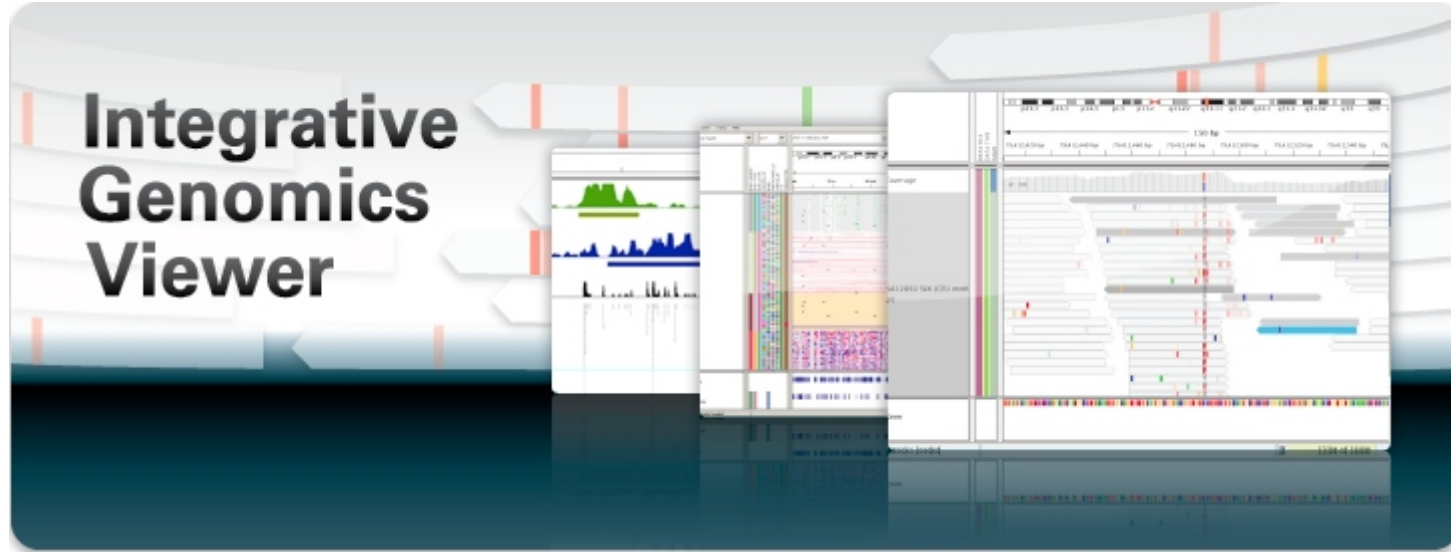
Cytosine  
méthylée



Guanine



# Visualisation des données

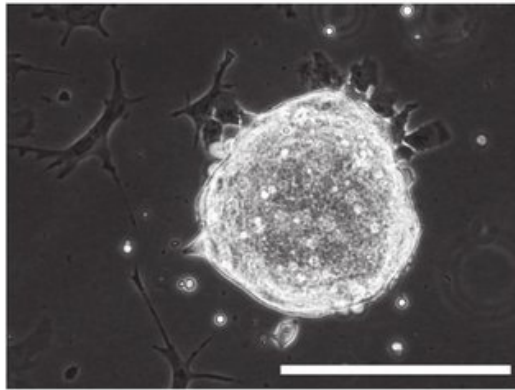


<https://software.broadinstitute.org/software/igv/>



# Comparaison de deux types cellulaires

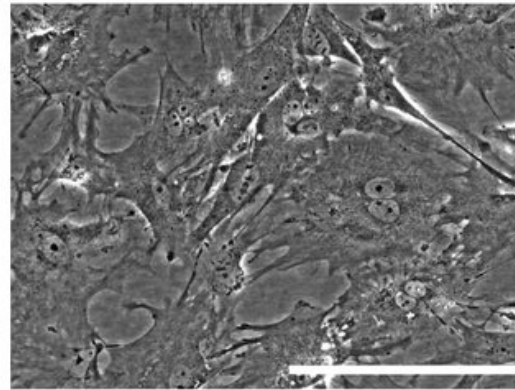
ESC



embryonic stem cells

cellules souches embryonnaires

MEF



mouse embryonic fibroblasts

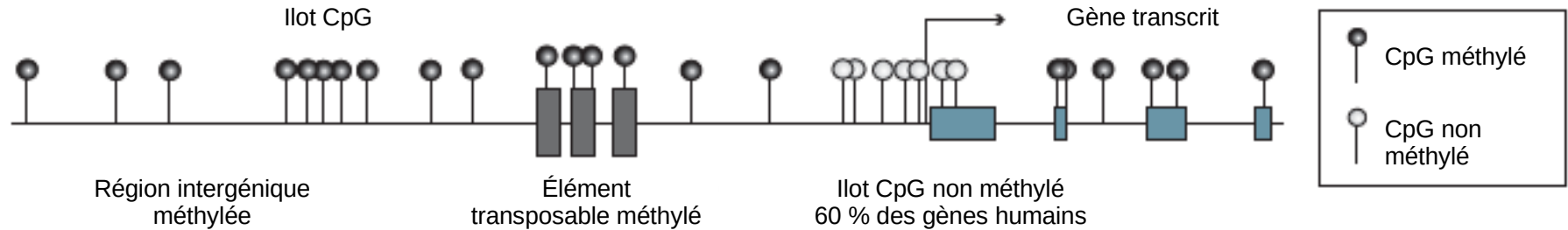
fibroblastes embryonnaires de souris

Boraas et al. 2016. PLoS ONE

Partie pratique !

# Méthylation des cytosines chez les mammifères

Où ?



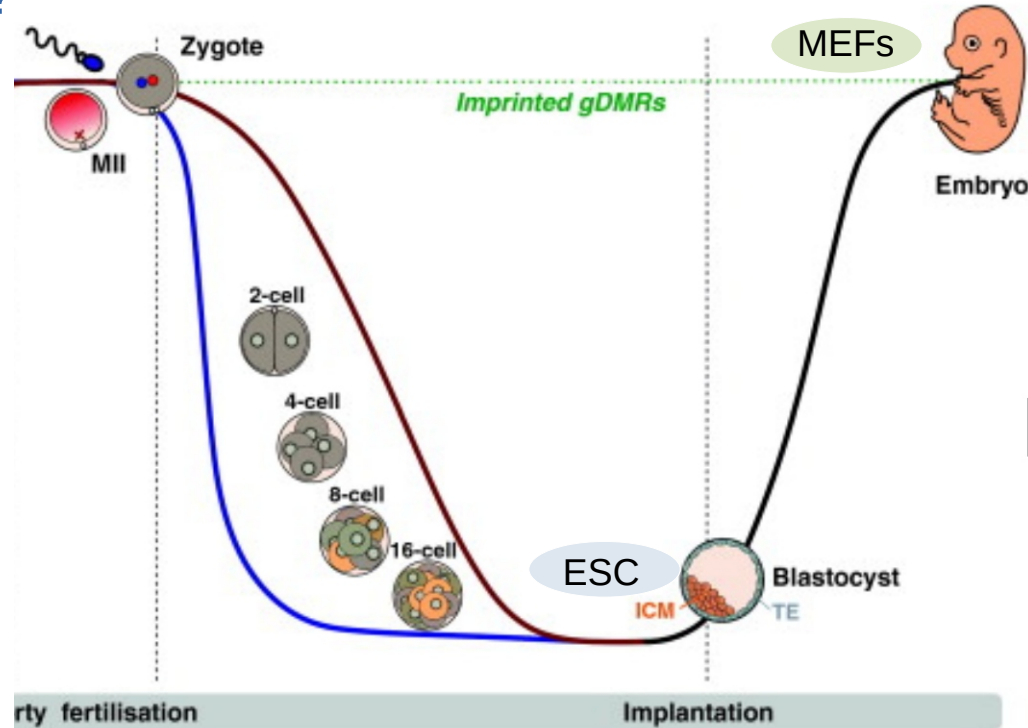
Comment ?

Enzymes

- DNMT (DNA methyl-transferase)
- TET (ten-eleven-translocation)

# Méthylation des cytosines chez les mammifères

Quand ?



# Annexes

A partir de 2005

# Séquençage de seconde génération

(ou NGS pour Next Generation Sequencing)

**Différence** majeure: plus besoin d'isoler un fragment d'ADN pour le séquencer

→ on séquence un **mélange**

1. **ADN fragmenté** → bibliothèque avec des adaptateurs synthétiques
2. **Immobilisation des fragments** (billes, surface de canaux microfluidiques)
3. **Génération de clusters clonaux (PCR)** : chaque fragment donne un cluster avec des milliers de molécules identiques
4. **Séquençage par cycles** en alternant réactions chimiques et imaging

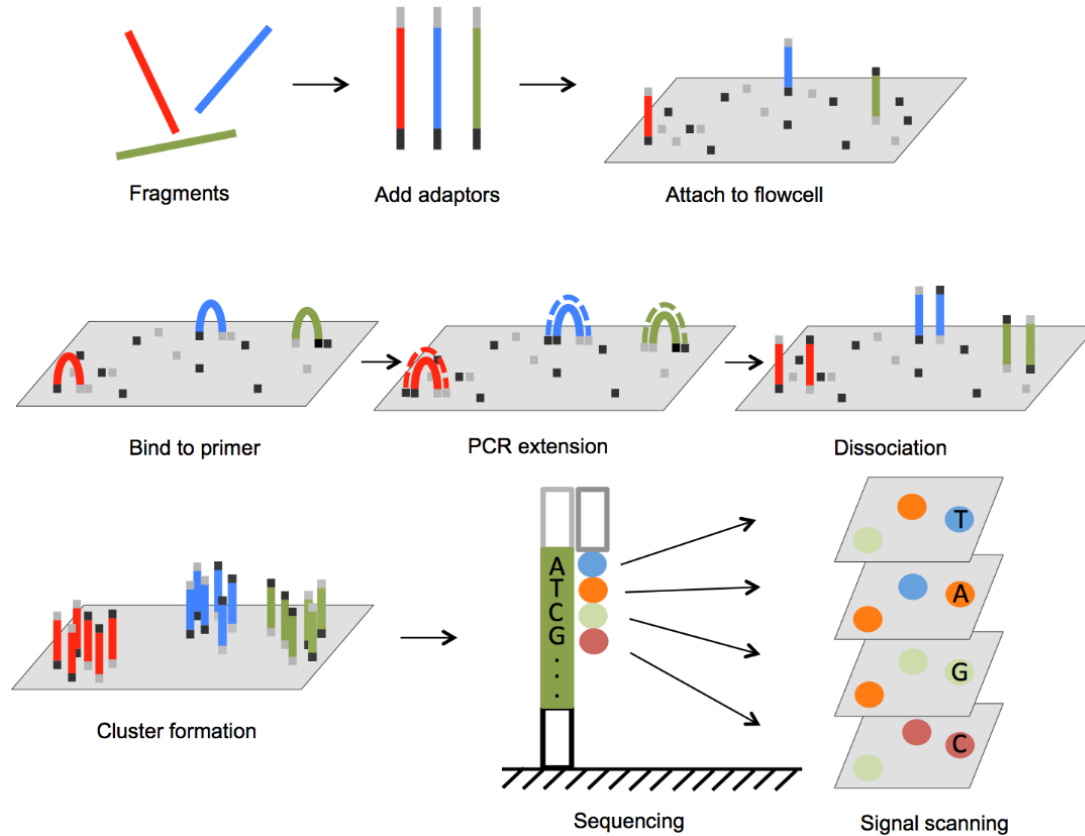
2005 : 454 (Life Sciences)

2006 : Solexa Genome Analyzer and SOLiD (Agencourt)

2010 : Ion Torrent

2006

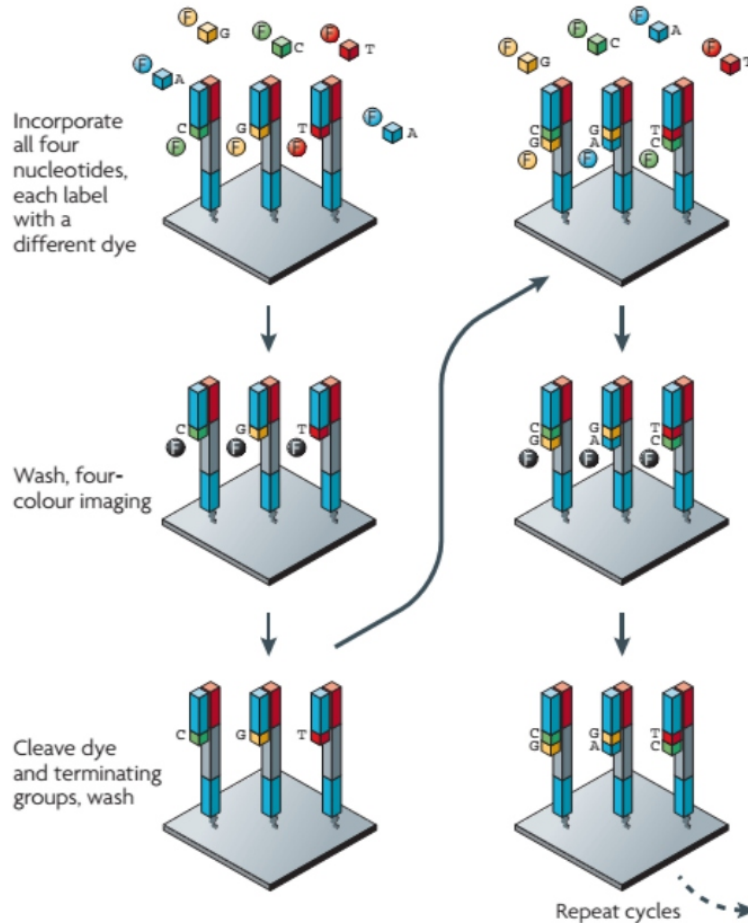
# Genome Analyzer Solexa (maintenant Illumina)



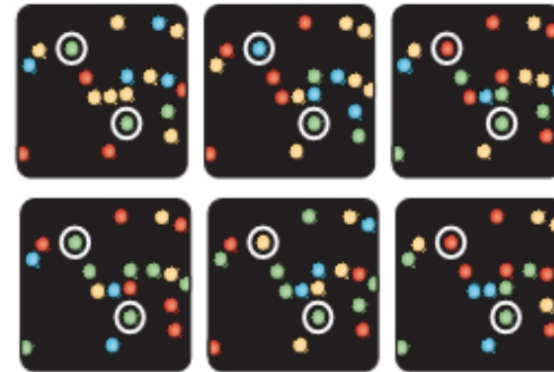
source : Illumina

2006

# Genome Analyzer Solexa (maintenant Illumina)



## 4. elongation & imaging

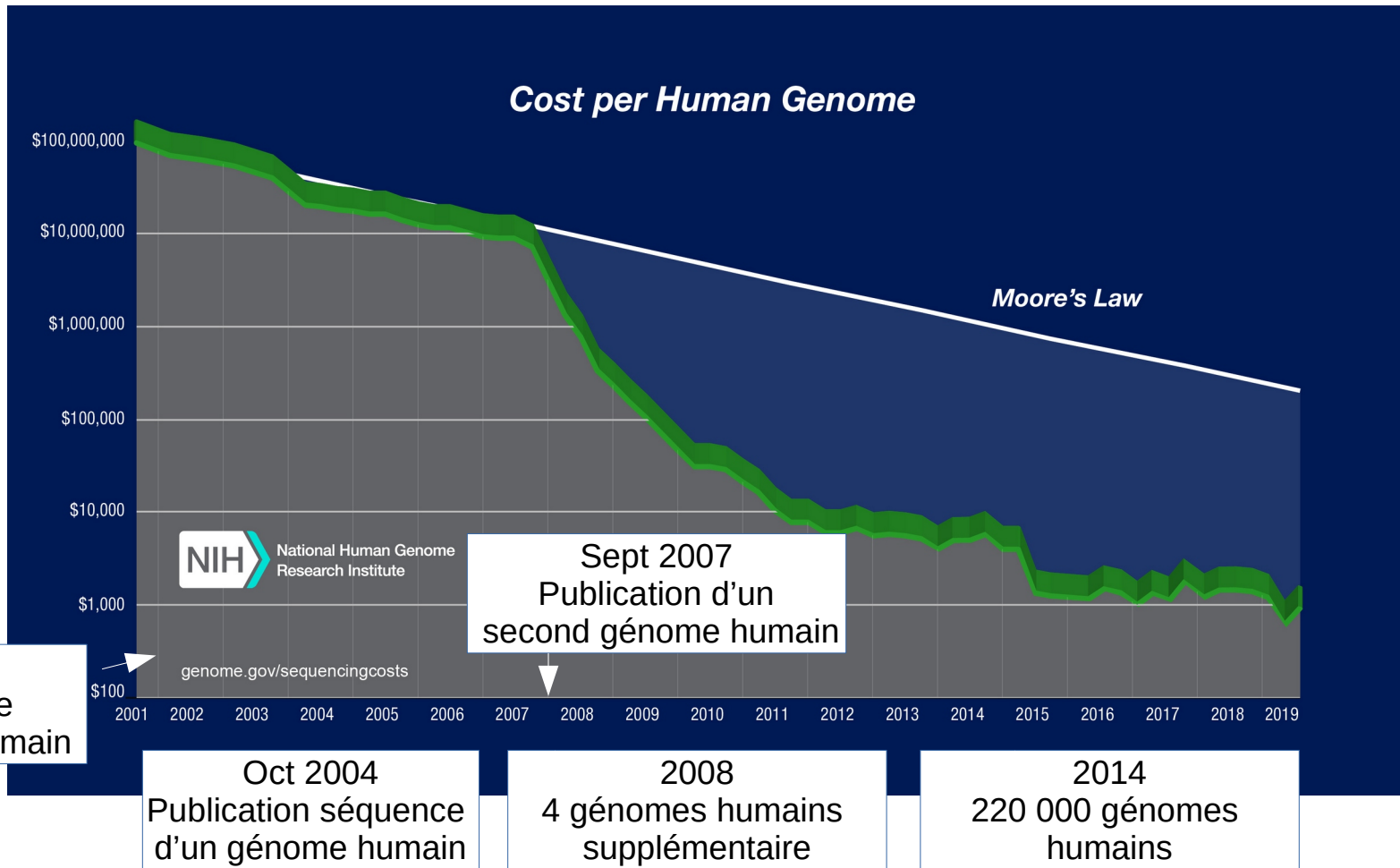


Top: CATCGT  
Bottom: CCCCCC

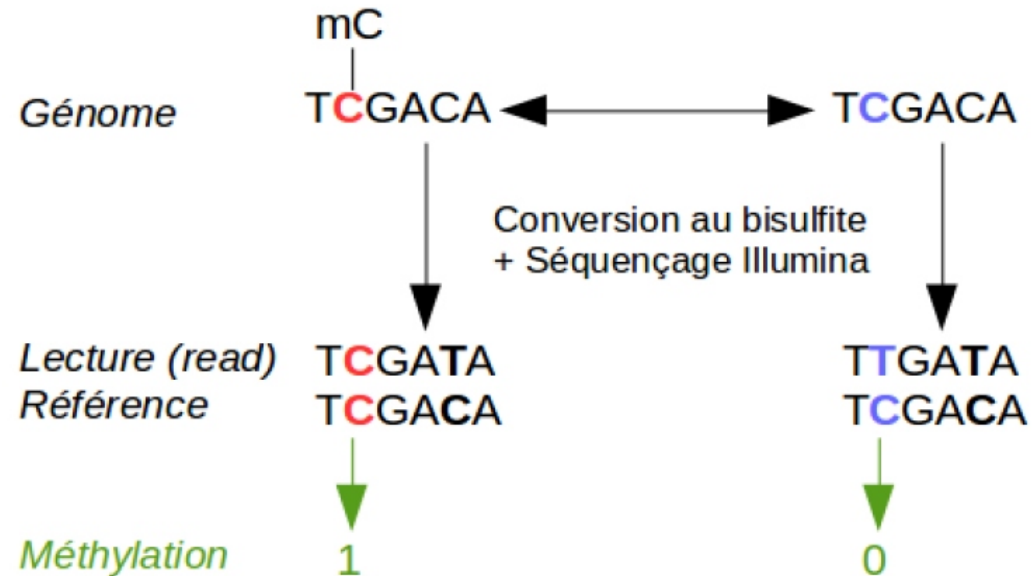
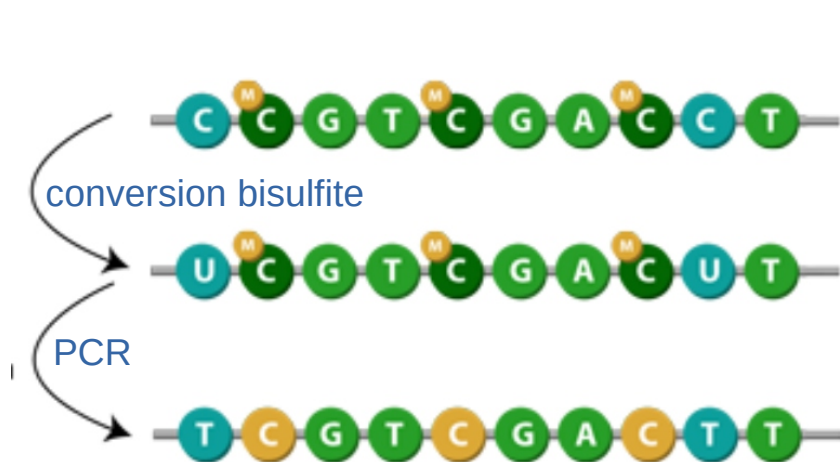
Metzker, 2010 (Nature Reviews Genetics)



# Coût du séquençage

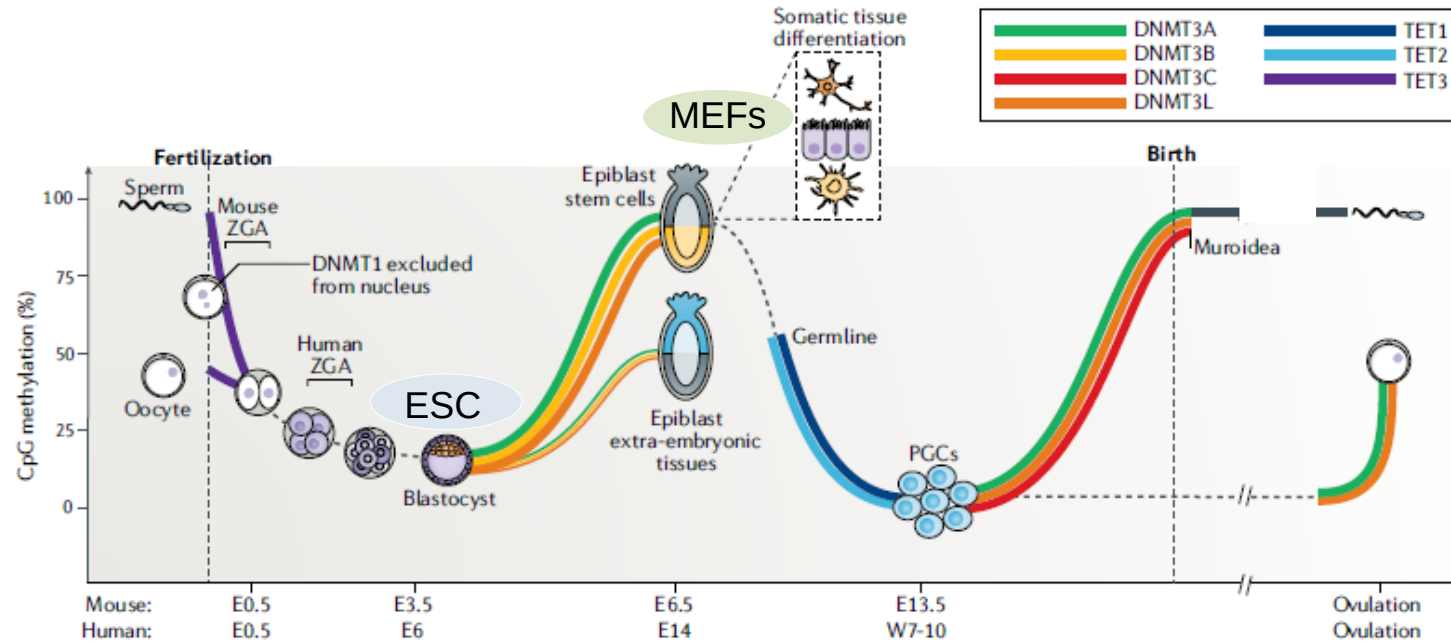


# Détermination du pourcentage de méthylation



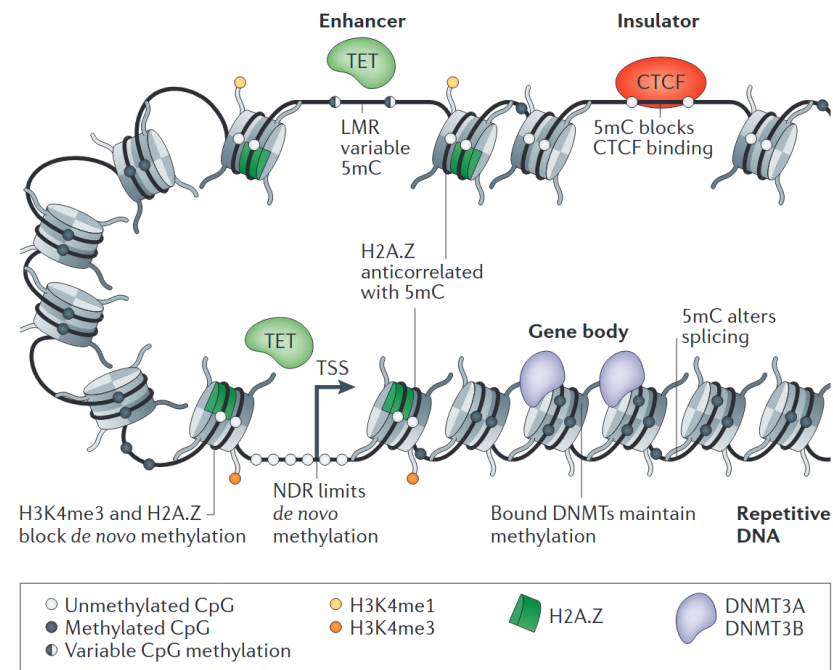
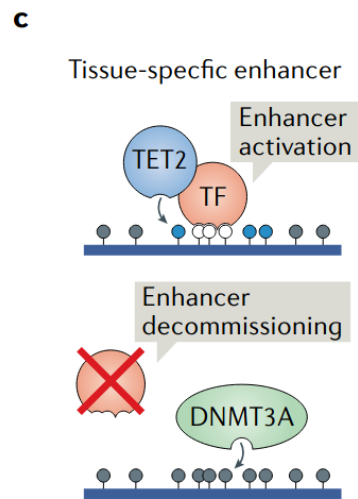
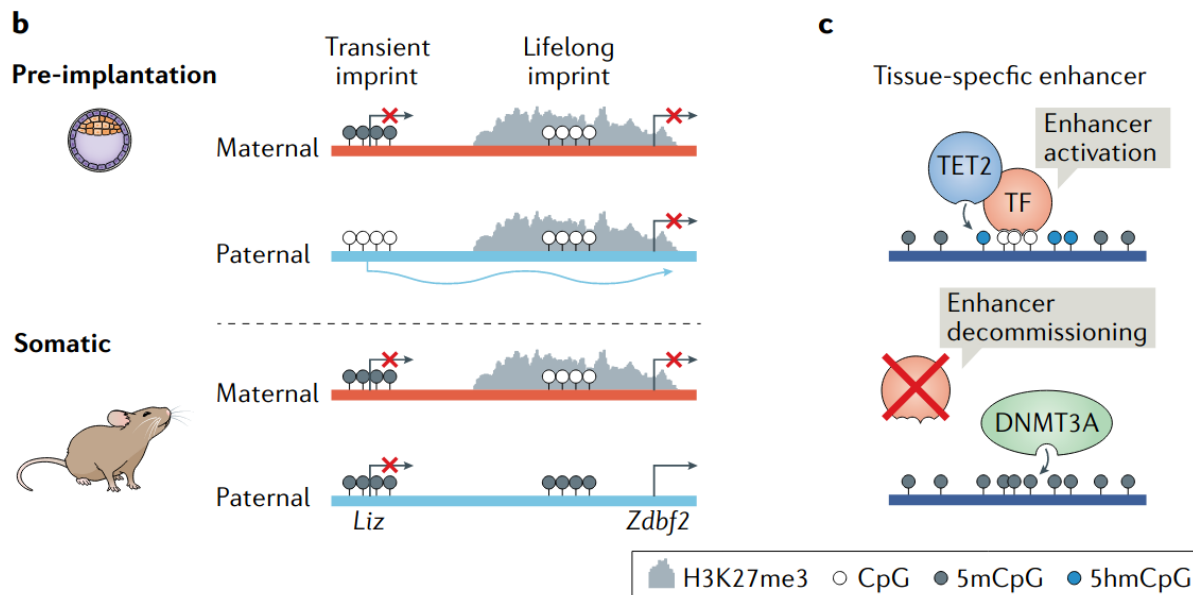
# Méthylation des cytosines chez les mammifères

Quand ?



# Méthylation des cytosines chez les mammifères

Pourquoi ?



# Méthylation des cytosines chez les mammifères

Pourquoi ?

... For fun and beauty ?



Morgan et al, 1999



Cubas et al, 1999