

« *Approche haut débit, transcriptome, et cancer* »

Benjamin Bonneau

Benjamin Bonneau

benjamin.bonneau@universite-paris-saclay.fr

Maître de conférences

Université Paris Saclay

université
PARIS-SACLAY



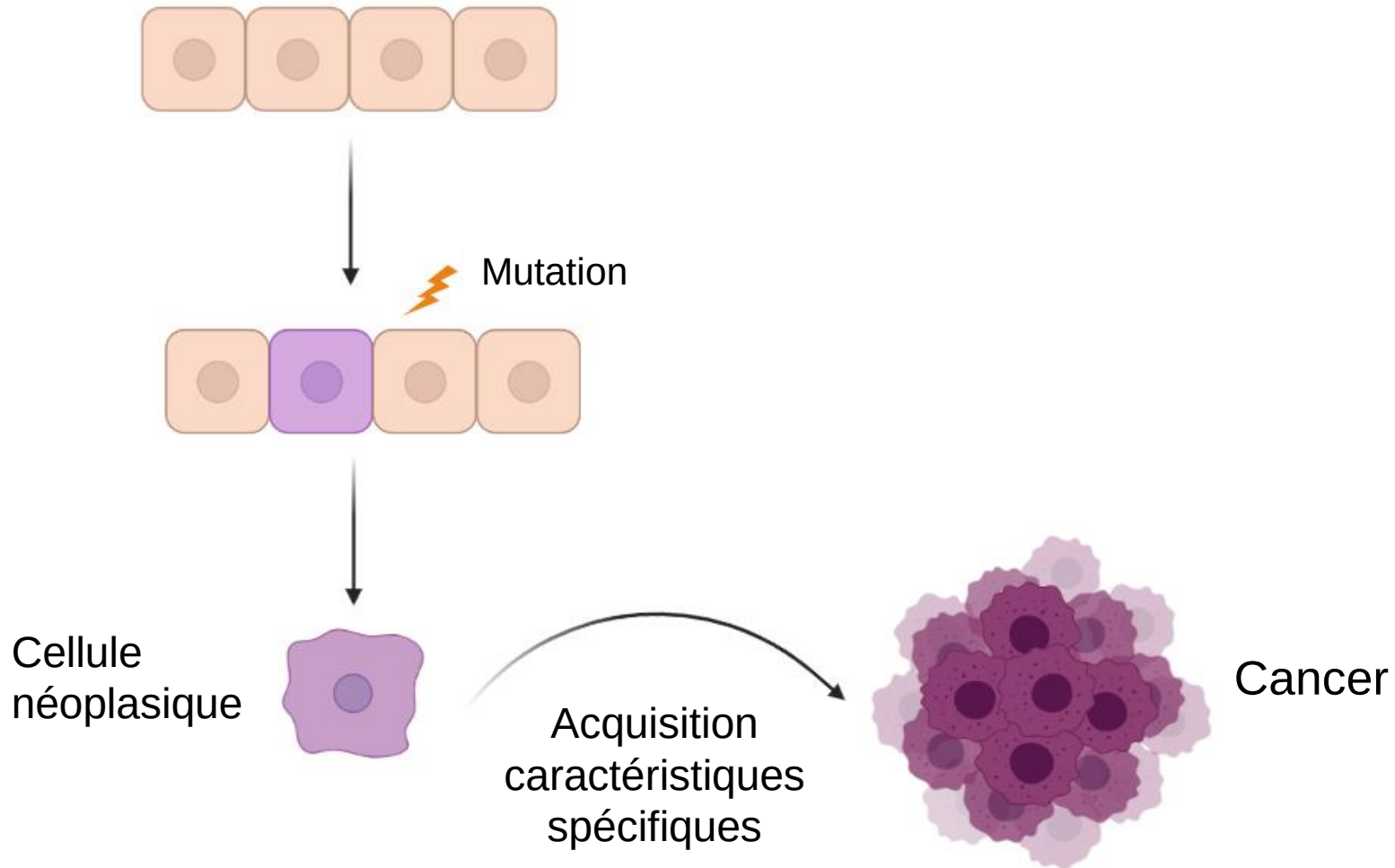
**Etude d'un cancer cérébral
pédiatrique, le médulloblastome**


institut
Curie

UMR3347 / U1021

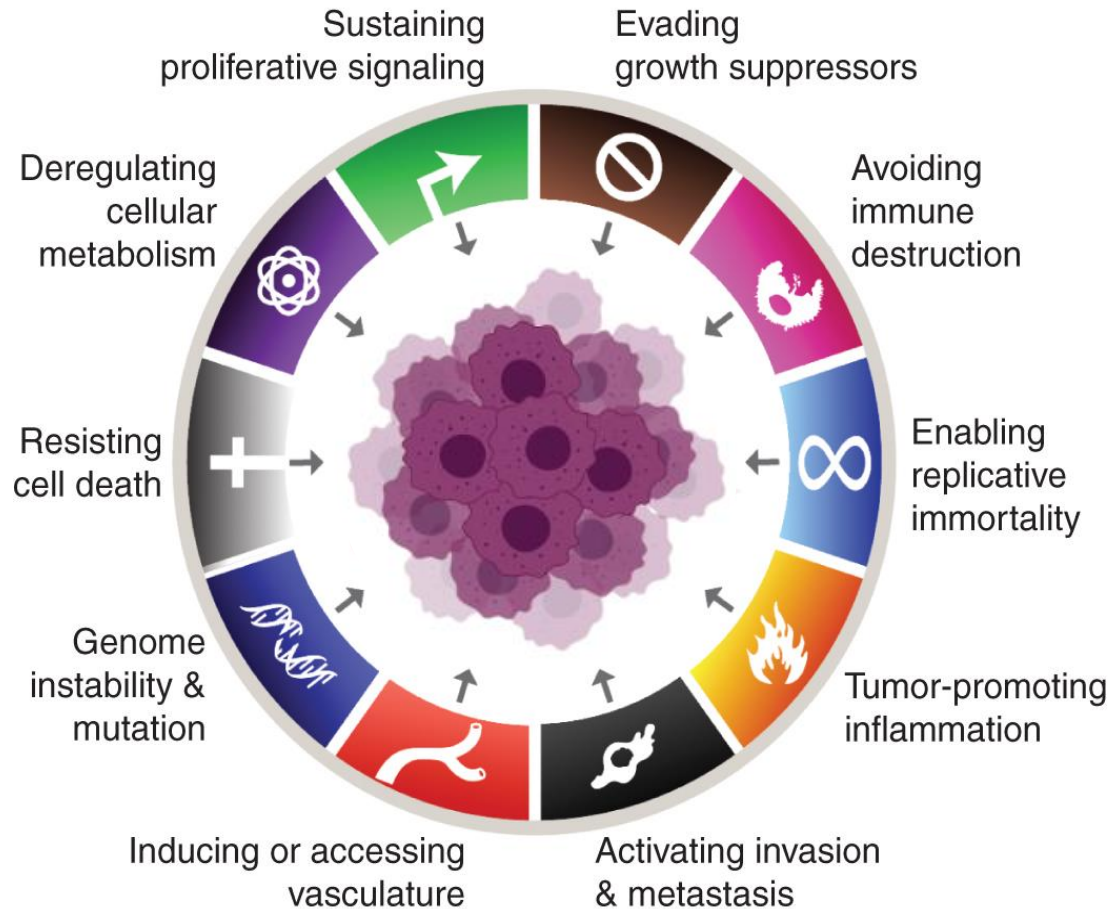
**Equipe Signalisation,
développement et tumeurs
cérébrales**

Transformation cellulaire



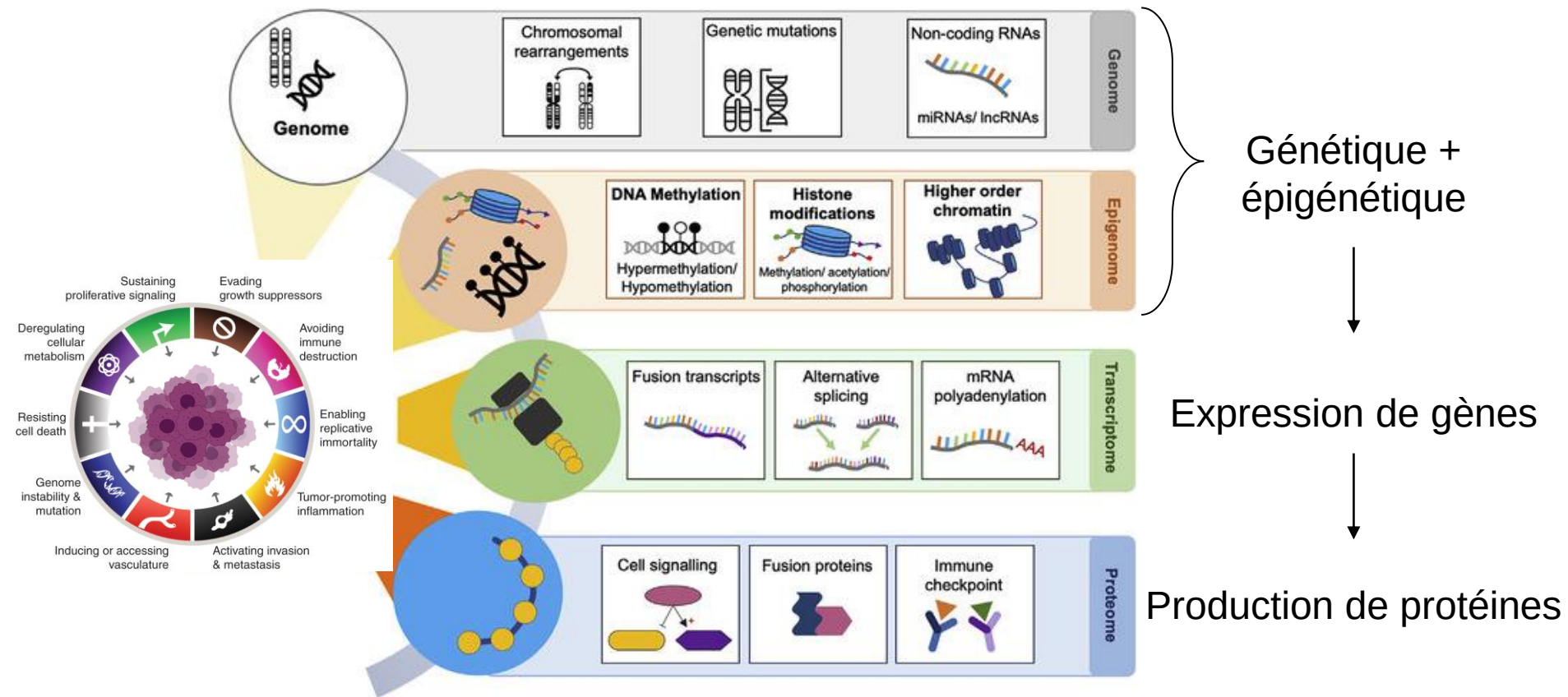
Caractéristiques spécifiques des Cancers

Hallmarks of cancer



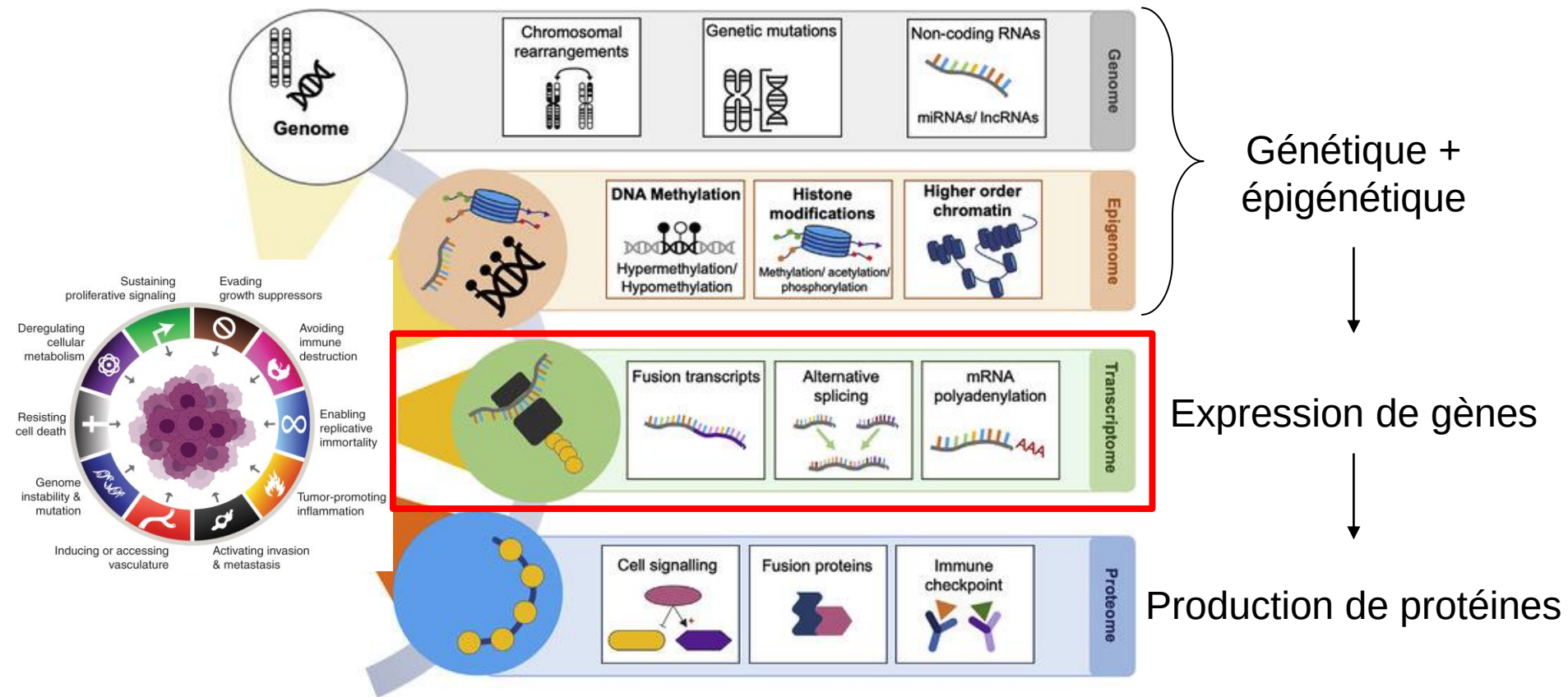
L'acquisition de nouvelles caractéristiques est rendue possible par diverses modifications au sein de la cellule

Etude des causes du cancer

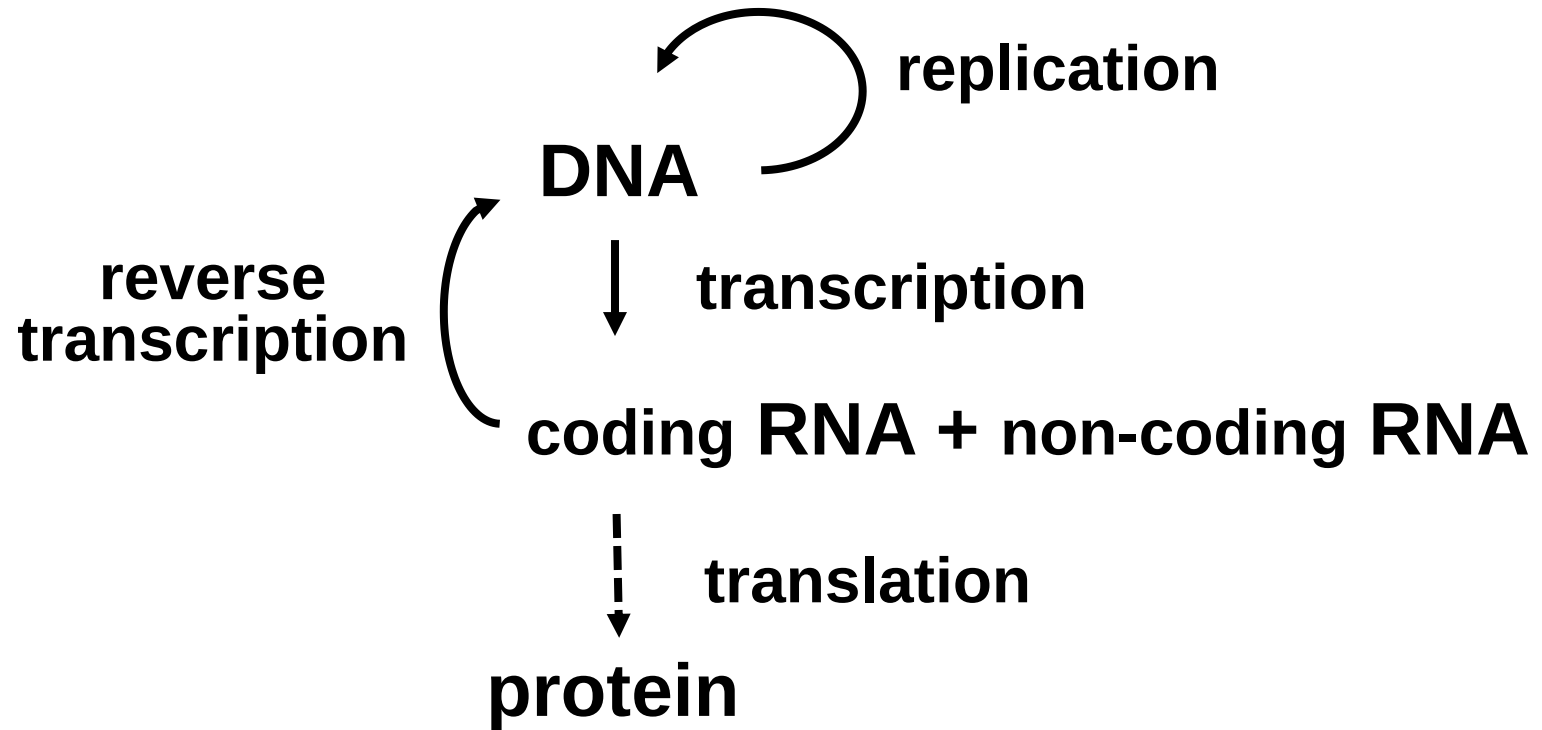


L'acquisition de nouvelles caractéristiques est rendue possible par diverses modifications au sein de la cellule

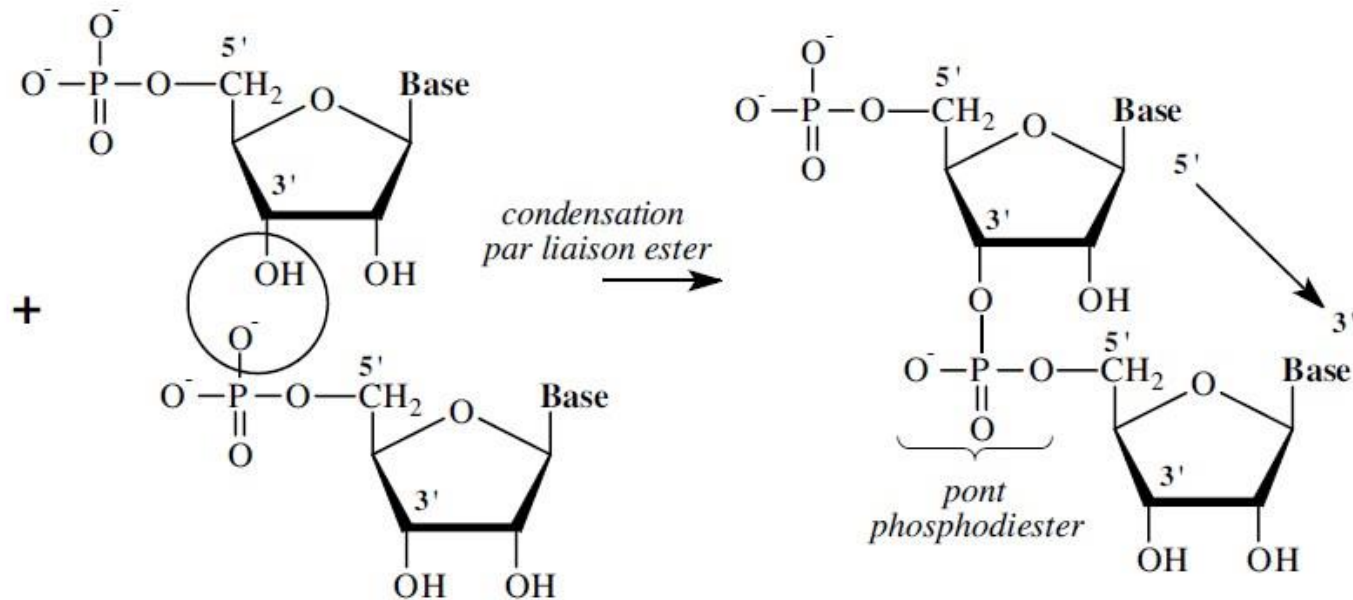
Etude des causes du cancer



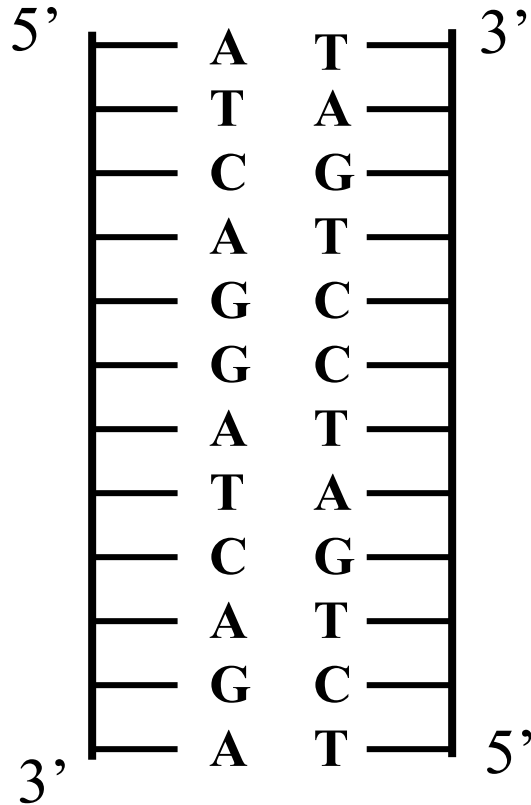
L'étude du transcriptome reflète les modifications génétiques et épigénétiques et renseigne sur l'état du protéome



ARN et ADN sont faits de nucléotides



base = A, G, C, T or U (T in DNA, U in RNA)



ADN est double brin

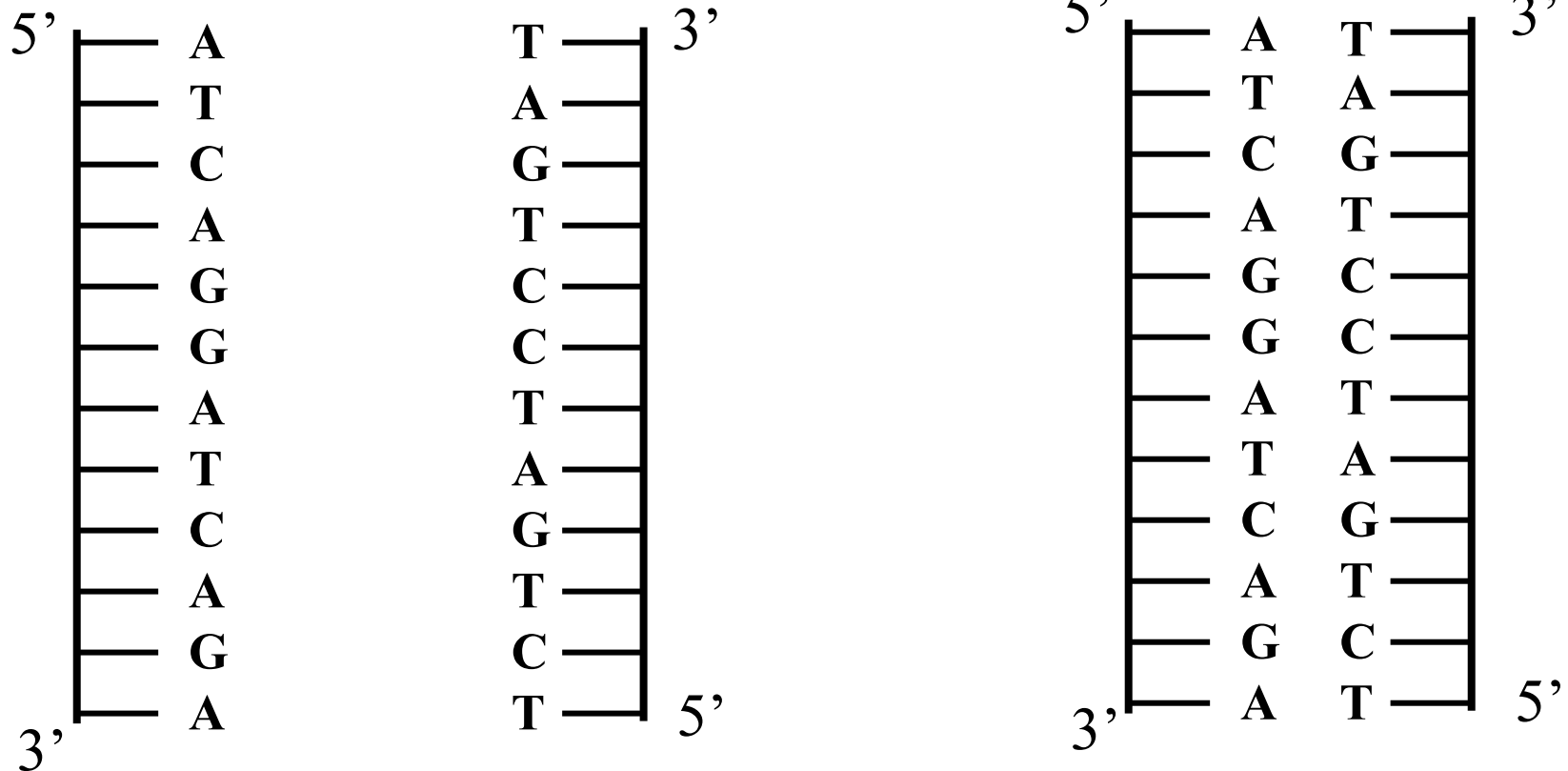


ARN est simple brin

**Les 2 brins ont des
séquences complémentaires**

Hybridation entre brins complémentaires:

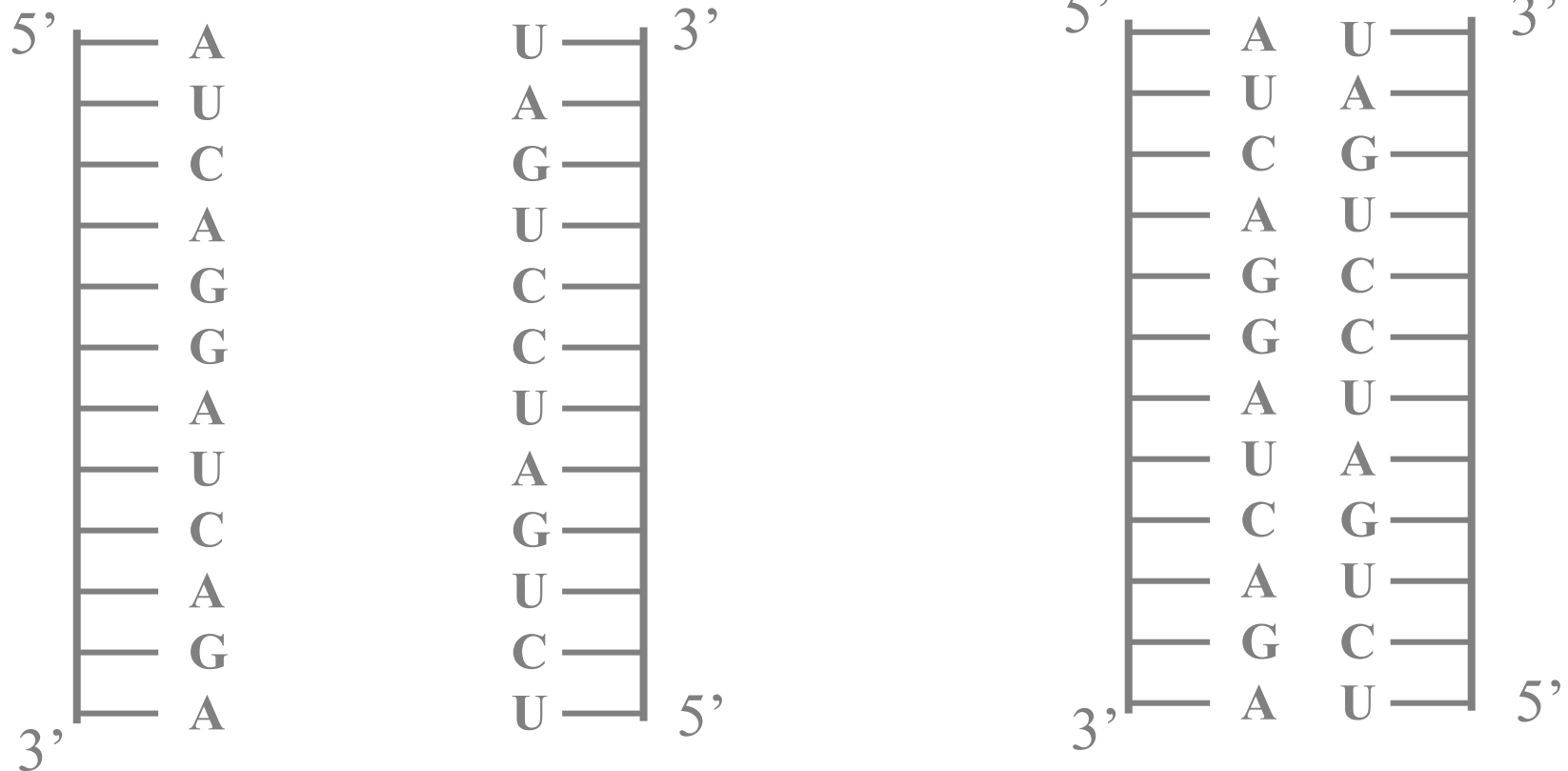
À la base des techniques de biologie moléculaire 4/4



Deux séquences complémentaires vont s'hybrider
-> 2 brins d'ADN peuvent s'hybrider

Hybridation entre brins complémentaires:

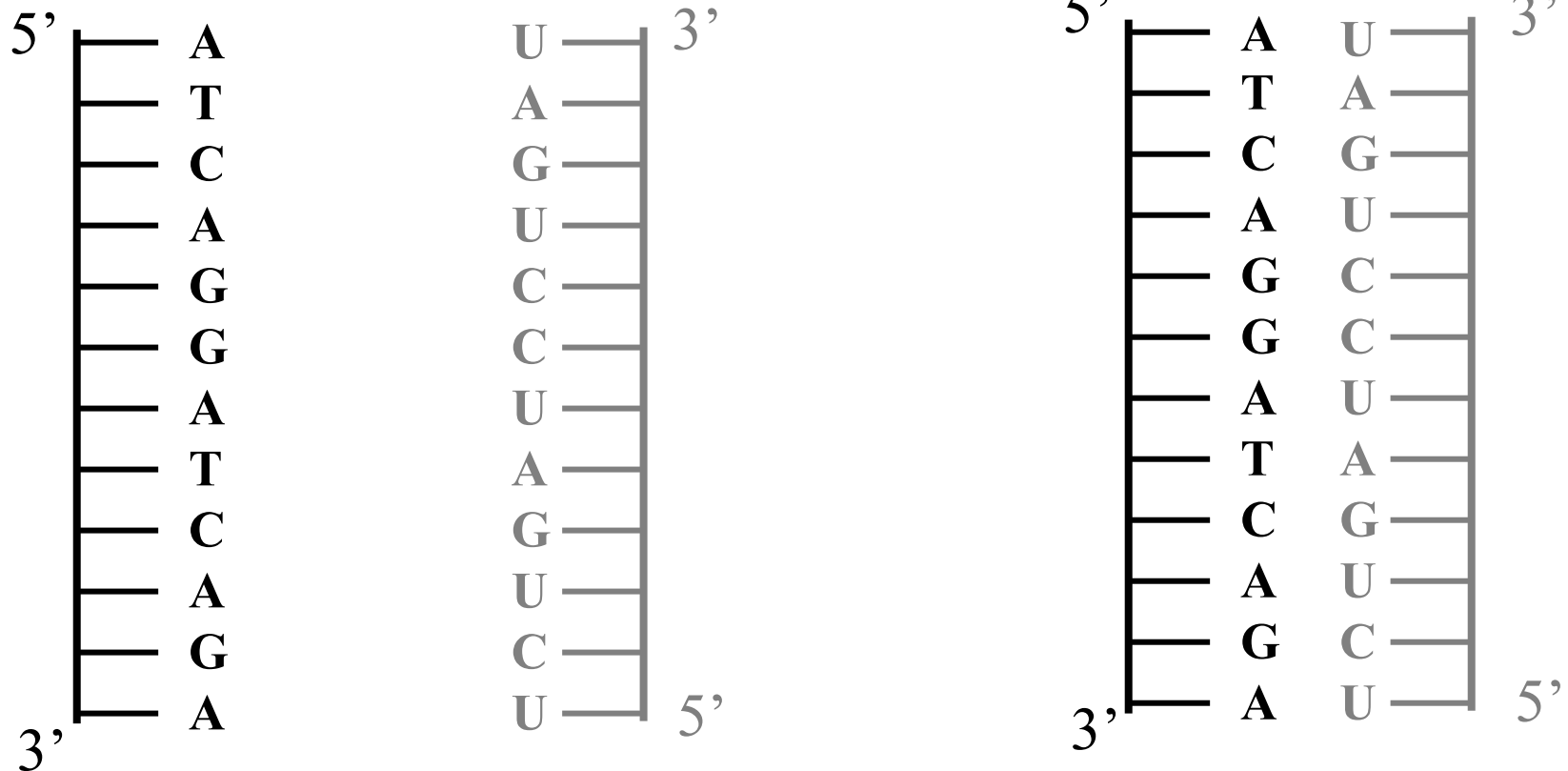
À la base des techniques de biologie moléculaire 4/4



Deux séquences complémentaires vont s'hybrider
-> 2 brins d'ARN peuvent s'hybrider

Hybridation entre brins complémentaires:

À la base des techniques de biologie moléculaire 4/4



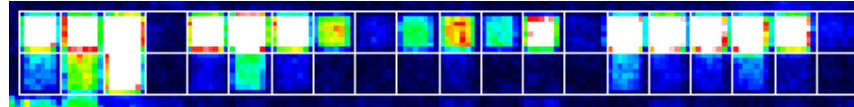
Deux séquences complémentaires vont s'hybrider
-> 1 brin d'ADN et 1 brin d'ARN peuvent s'hybrider

Outils pour l'analyse transcriptomique

2 technologies principales:

- **Micro-array (DNA chips ou puce à ADN)**
 - **RNA-seq par next-generation sequencing**
-

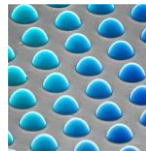
Différents types de puces pour étudier le transcriptome



Affymetrix

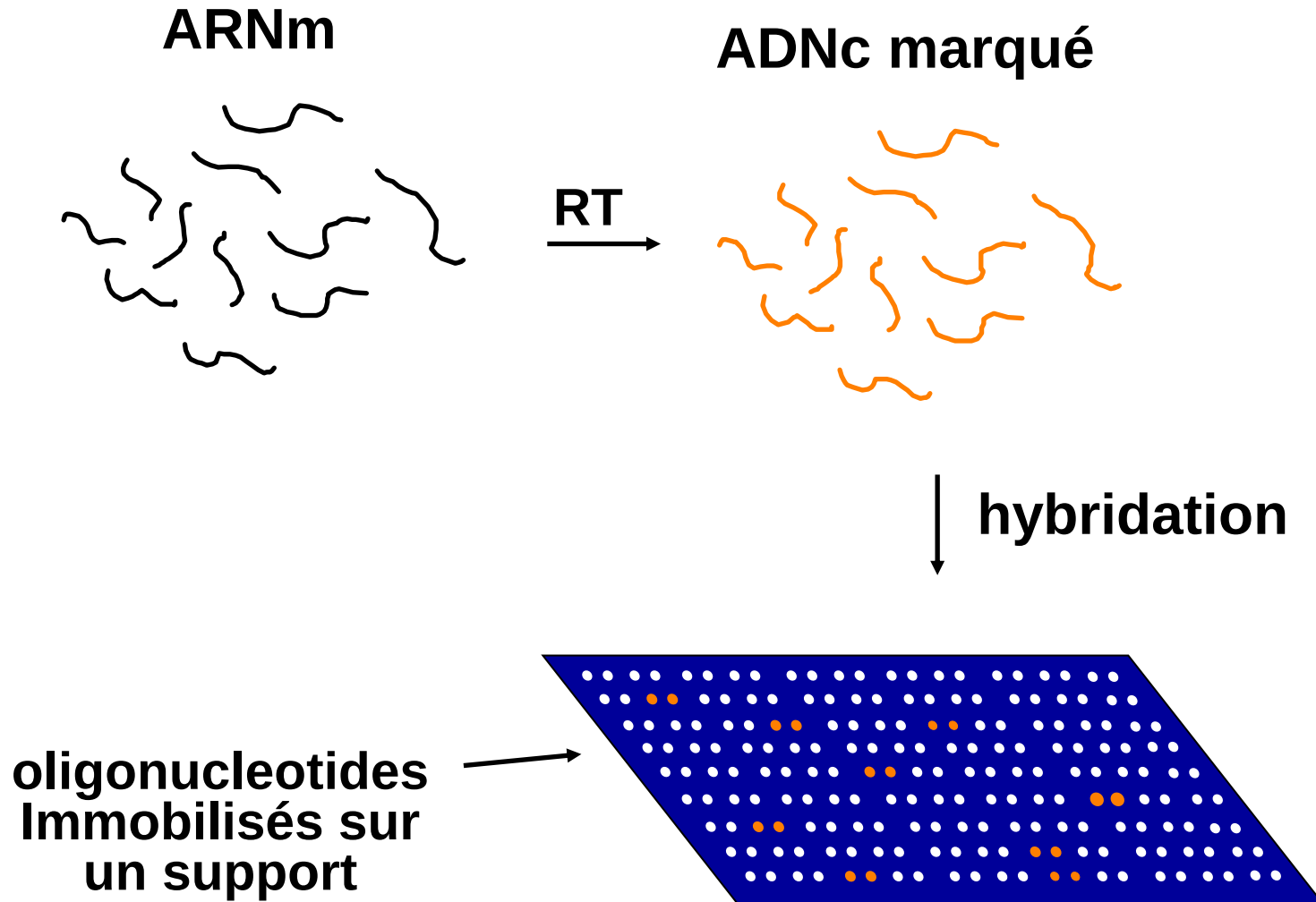


Agilent



Illumina

Puce à ADN: Principe

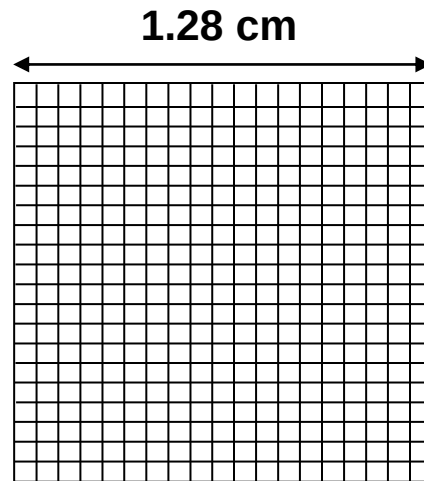


Affymetrix chips



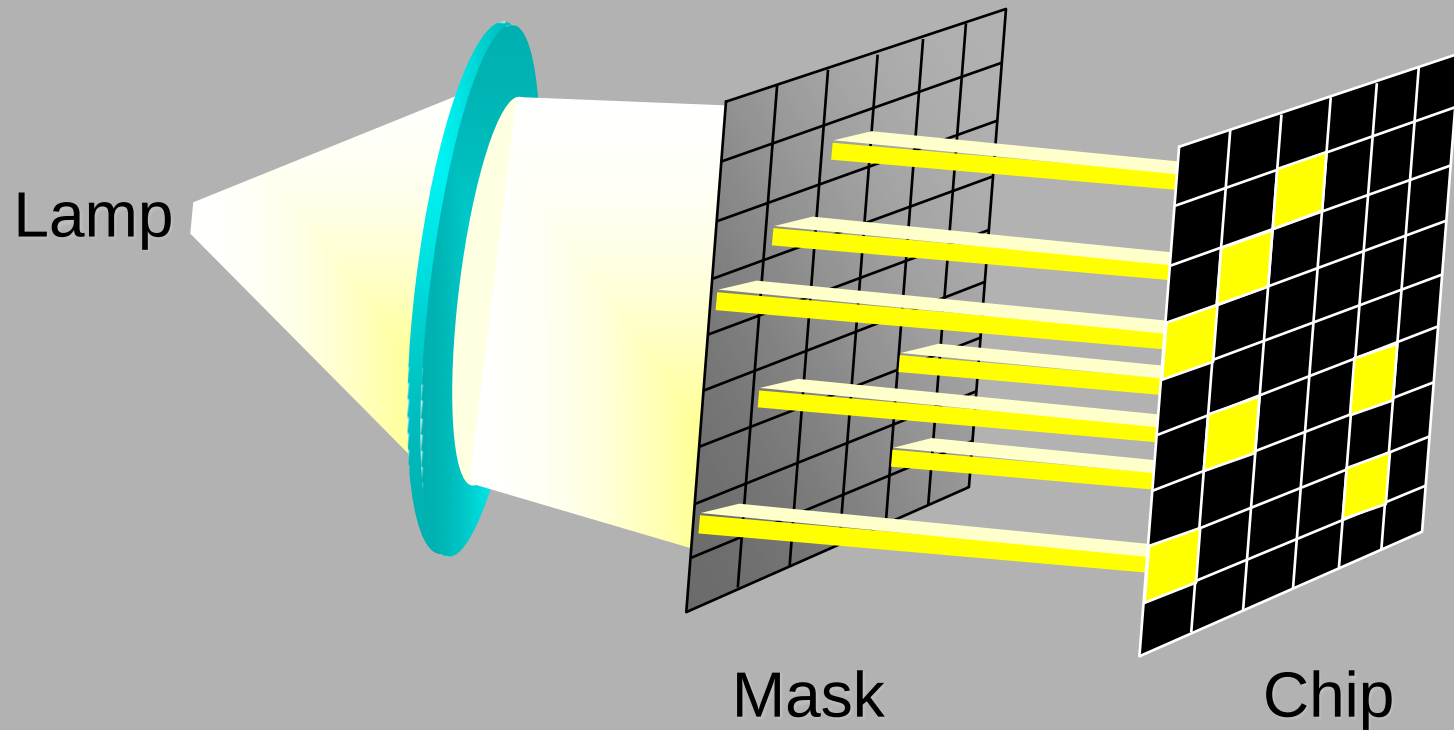
Puces Affymetrix

Puces avec de courts oligonucleotides (25 mers)



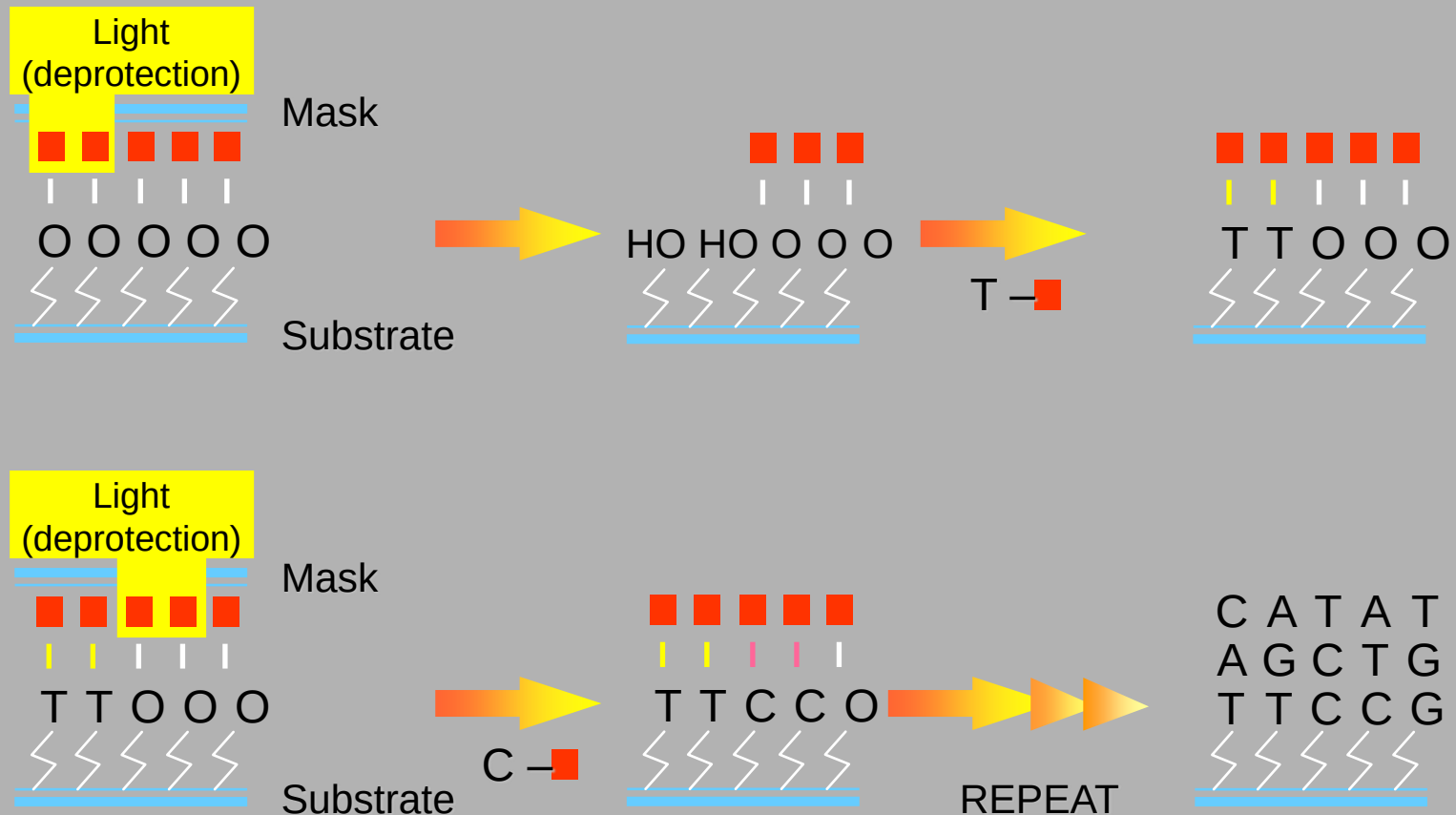
- 1 puce = plusieurs milliers à centaines de milliers de carrés de $50\mu\text{m} \times 50\mu\text{m}$, $24\mu\text{m} \times 24\mu\text{m}$, $20\mu\text{m} \times 20\mu\text{m}$, $18\mu\text{m} \times 18\mu\text{m}$
 - Chaque carré contient plusieurs millions de copies d'un oligonucléotide donné
-

Synthèse lithographique



Courtesy of Affymetrix

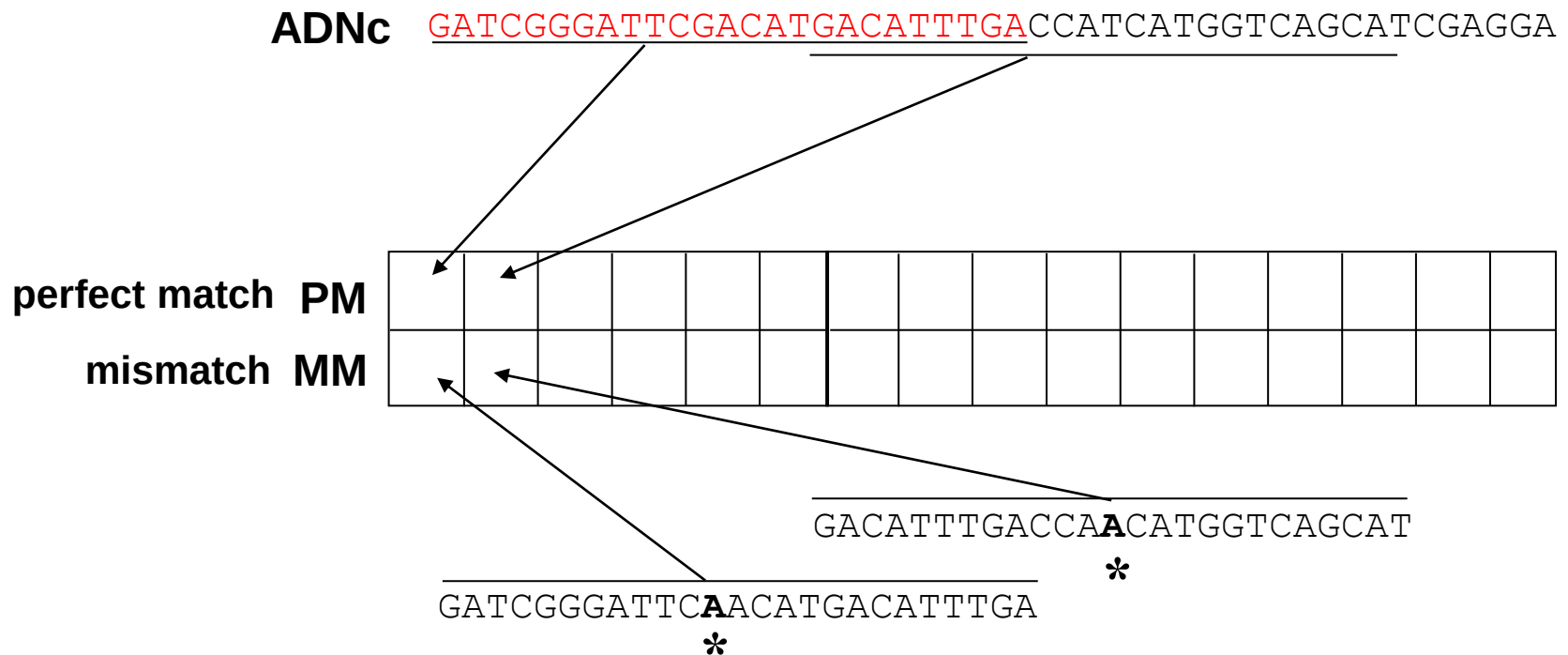
Synthèse d'oligonucléotides ordonnés



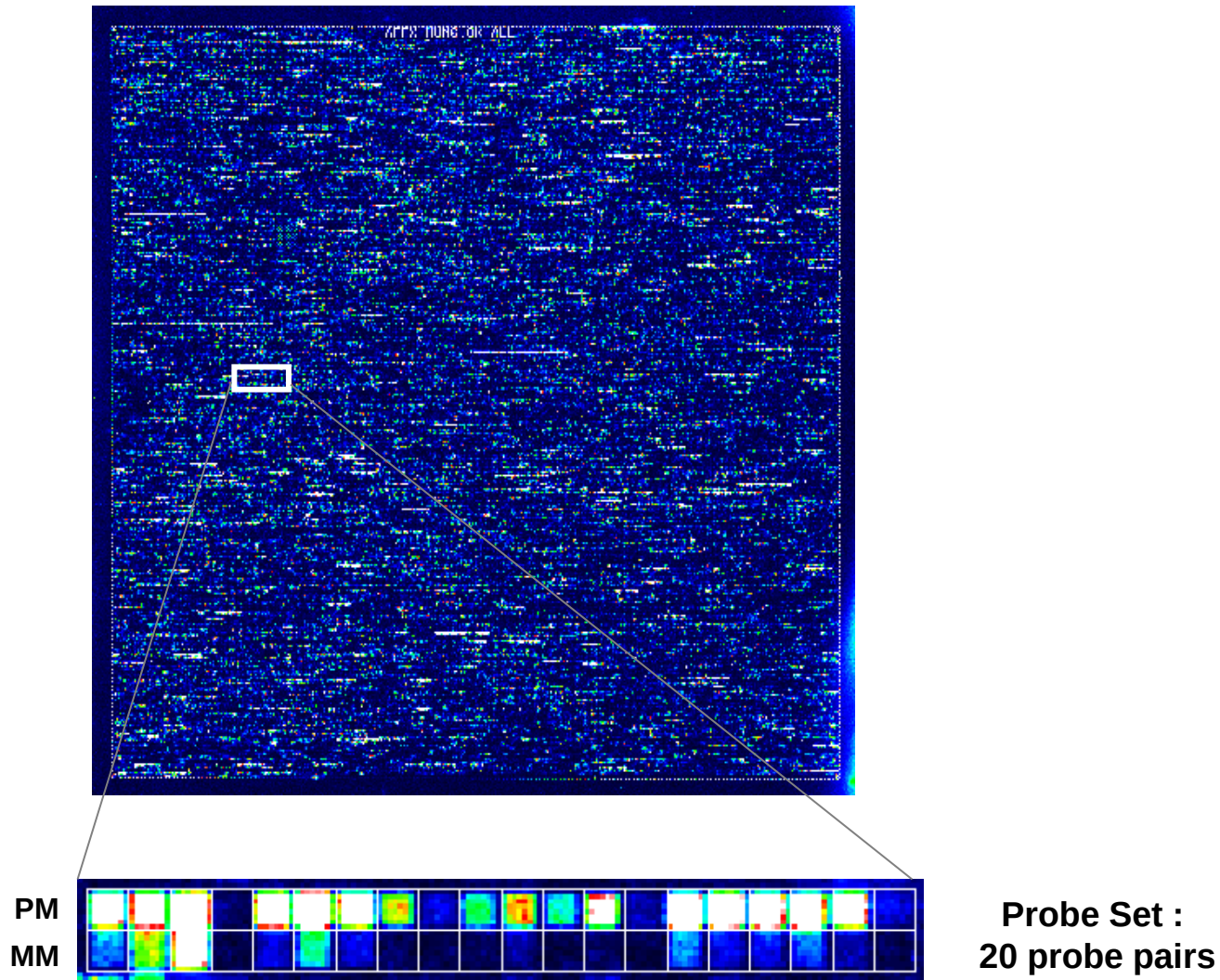
Courtesy of Affymetrix

Mesure d'expression, principe des puces Affymetrix

L'expression d'un gène est mesurée à l'aide de 11 à 20 paires de carré



Example of the Human HuGeneFl



Example of the Human HuGeneFl



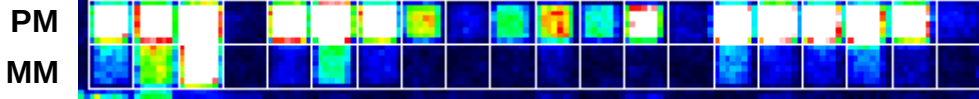
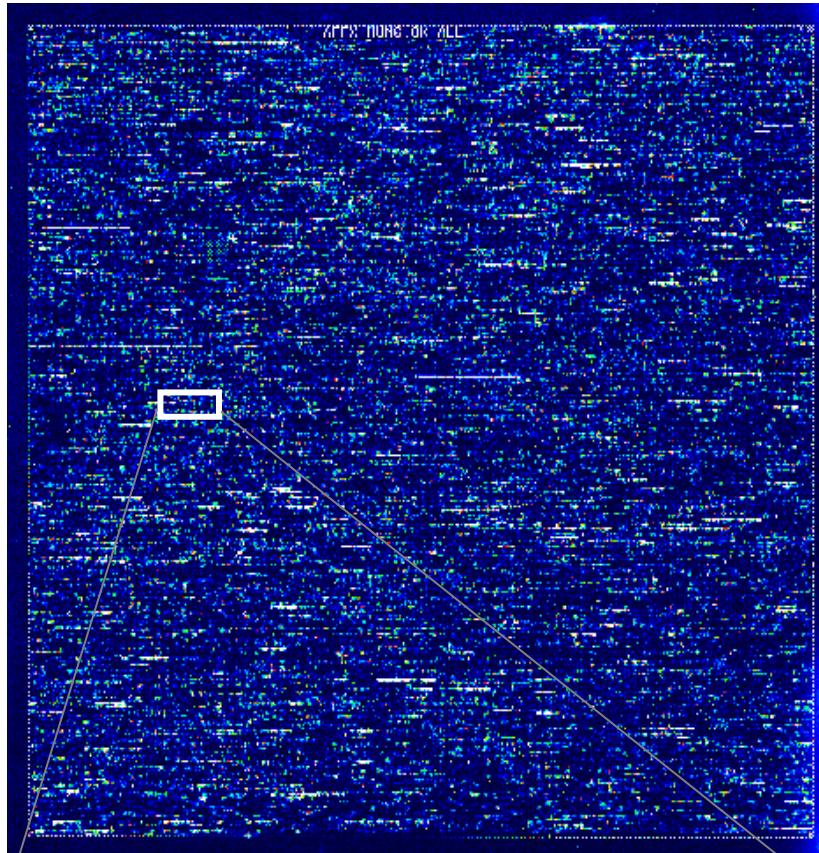
GeneChip® System
and Operating Software



Calcul un « score » pour chaque gène en fonction de l'intensité du signal dans les PM et les MM



« Score » reflète la quantité d'ADN hybridé et donc la quantité d'ARNm au départ



The human U133 Plus 2.0 chips (sept. 2003!)



- Carrés de 11 μM x 11 μM
- 11 paires PM - MM par gène ou EST
- ~1.300.000 carrés différents par puce
- 54.000 gènes + EST évalués

EST = expressed sequence Tag

Puces disponible pour l'étude du transcriptome

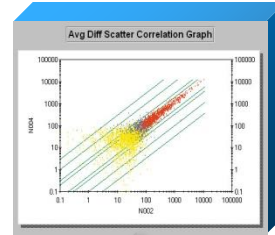
- *Arabidopsis thaliana*
 - *Caenorhabditis elegans*
 - *Canis familiaris*
 - *Danio rerio*
 - *Drosophila melanogaster*
 - *Xenopus laevis*
 - *E. coli*
 - *Saccharomyces cerevisiae*
 - mouse
 - rat
 - human
-

Affymetrix, avantages

- « ready to use »



GeneChip®
Probe Arrays



NetAffx™ Website

MAS 5.0
LIMS 3.0
MicroDB 3.0
DMT 3.0

- grand nombre de gènes
- robuste
- Puces pour épissage

Gene of interest	Affymetrix ratio TM/CM	TaqMan ratio TM/CM
Glycoprotein 39	13.7	13.6
Caspase 14	2.4	2.9
MRP8	108.2	81.4
Interleukin 1b	51.7	28.7
Leukoprotease inhibitor	11.7	19.1
CD44	3.6	1.1
Liver arginase	5.9	4.4
MMP3	7.4	3.0
Phosphoglycerate mutase	0.4	0.3
Serum Amyloid A3	250.8	161.3

Affymetrix, désavantages

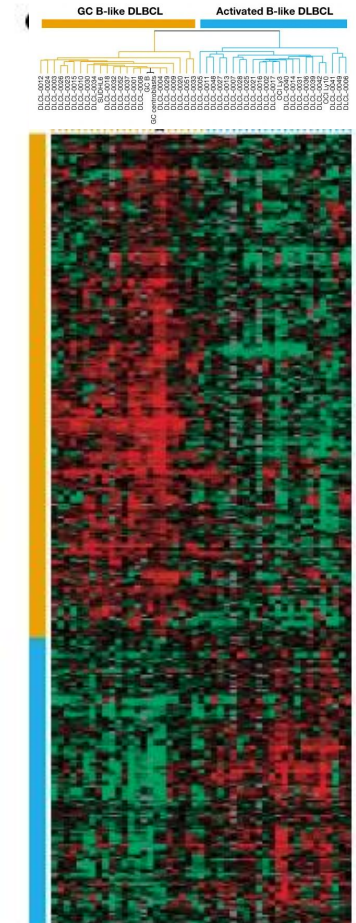
- **pas flexible**
 - **assez cher**
 - **ne peut pas être utilisé 2 fois**
 - **nécessite une grande quantité d'ADN**
-

Puces ADN: apport en cancérologie

Classification de tumeurs en sous-groupes moléculaires (cancer du sein, lymphome, medulloblastome, ...)

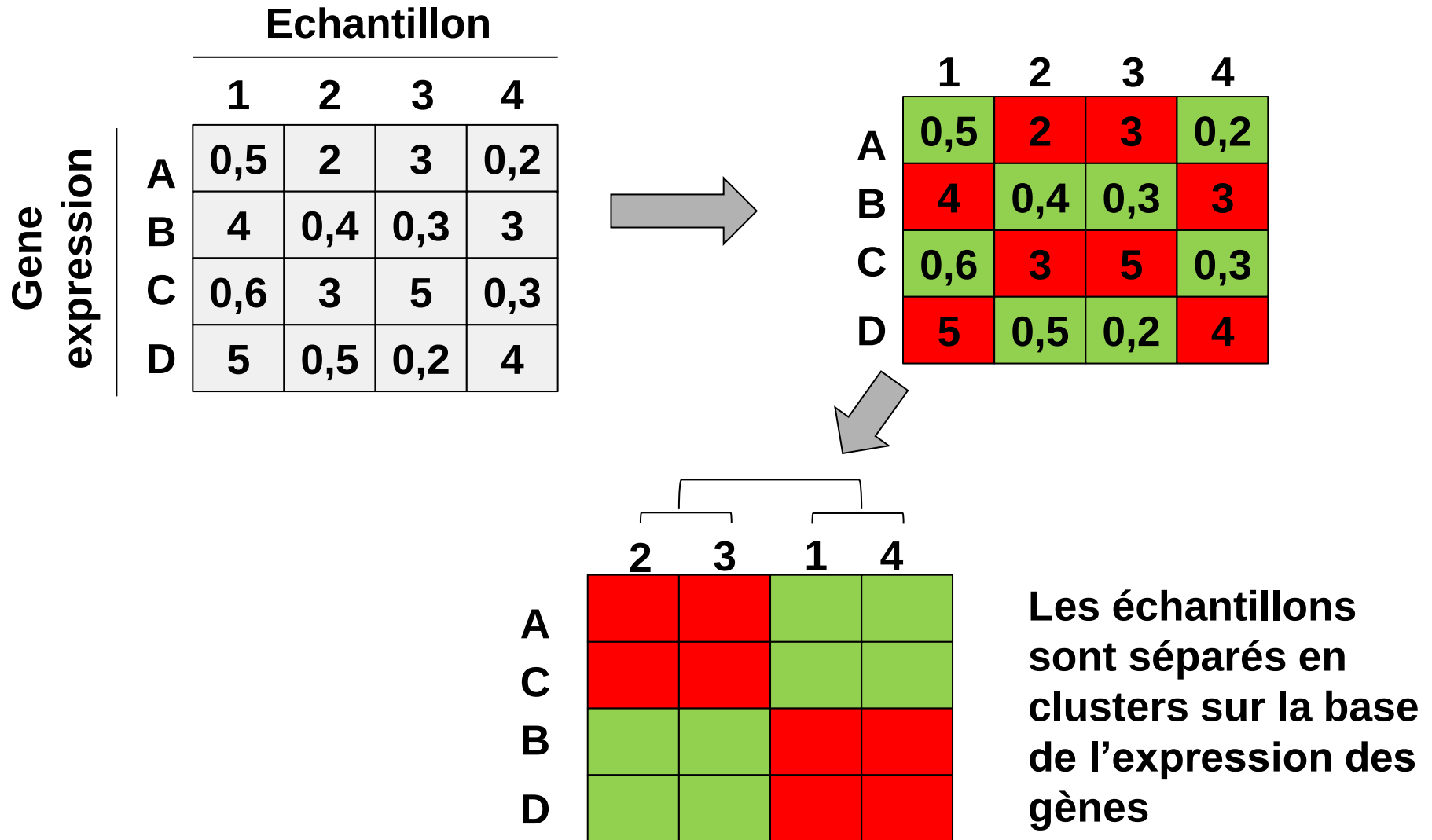
- **Importance pronostic:**
Exemple du Diffuse large B-cell lymphoma (DLBCL)
(Alizadeh et al., 2000, Nature)

Découverte de 2 sous-groupes dans les DLBCL (pas de classification connues précédemment)



Unsupervised hierachical clustering

Hierarchical clustering

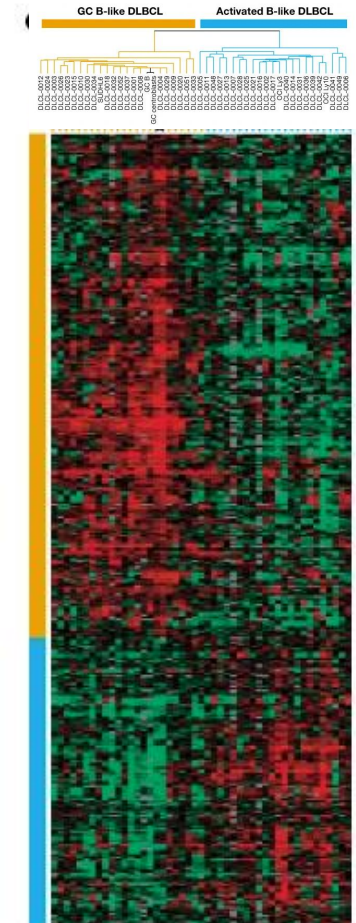
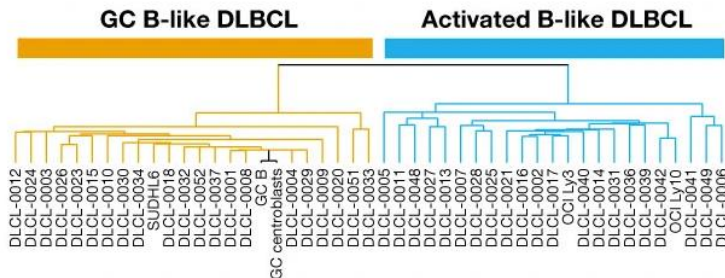


Puces ADN: apport en cancérologie

Classification de tumeurs en sous-groupes moléculaires (cancer du sein, lymphome, medulloblastome, ...)

- **Importance pronostic:**
Exemple du Diffuse large B-cell lymphoma (DLBCL)
(Alizadeh et al., 2000, Nature)

Découverte de 2 sous-groupes dans les DLBCL (pas de classification connues précédemment)



Unsupervised hierarchical clustering

Puces ADN : apport en cancérologie

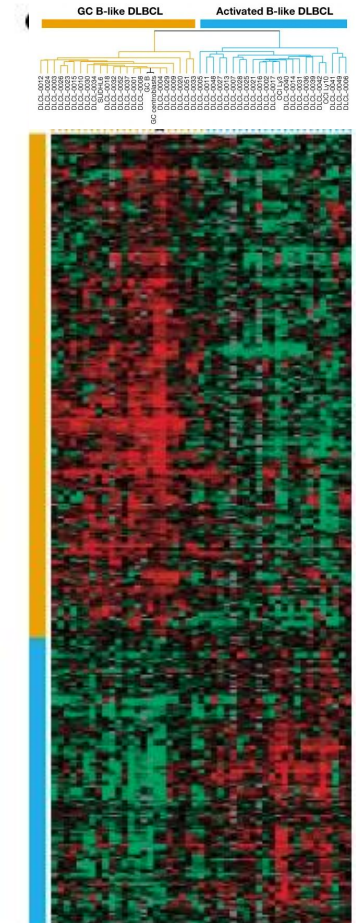
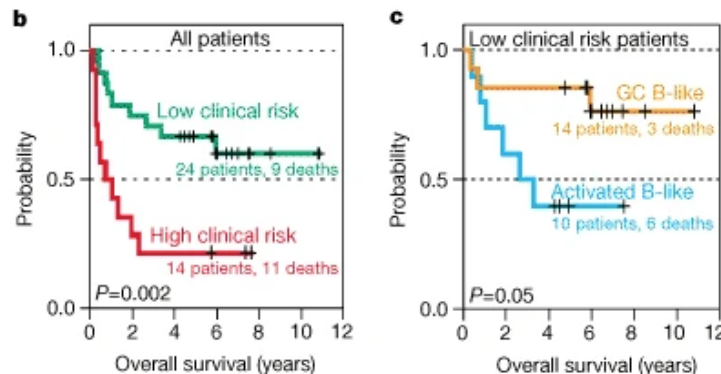
Classification de tumeurs en sous-groupes moléculaires (cancer du sein, lymphome, medulloblastome, ...)

- Importance pronostic:

Exemple du Diffuse large B-cell lymphoma (DLBCL)
(Alizadeh et al., 2000, Nature)

Avec le même traitement, 76% survie à 5 ans pour GC B-like
contre 16% pour Activated B-like

Meilleure classification du risque



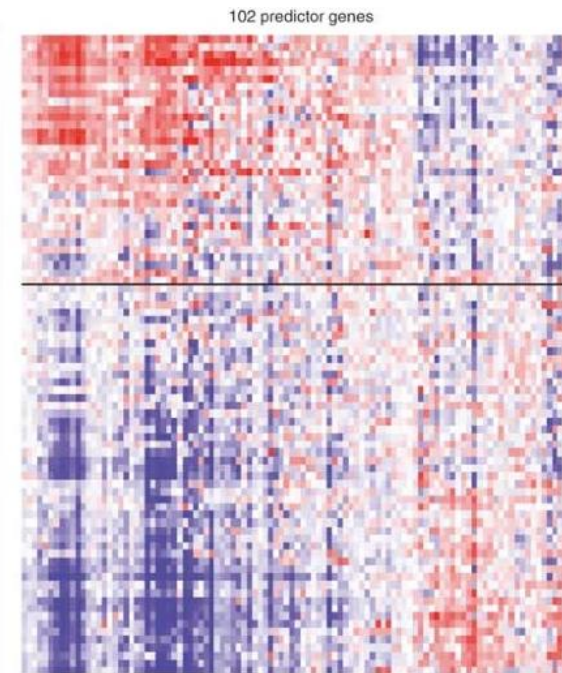
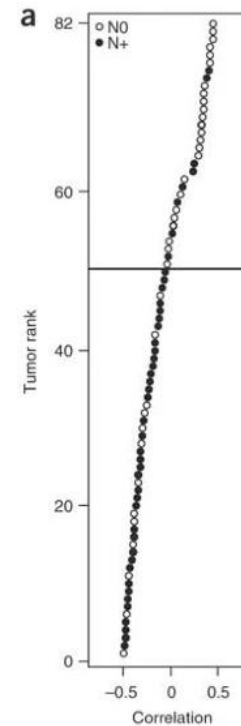
Unsupervised hierarchical clustering

Puces ADN : apport en cancérologie

Classification de tumeurs en sous-groupes moléculaires (cancer du sein, lymphome, medulloblastome, ...)

- **Ajustment traitement:**
Head and neck squamous cell carcinomas (HNSCCs)
(Roepman et al., 2005, Nature Genetics)

Mise en évidence de 102 gènes
différentiellement exprimés entre des tumeurs
ayant formées des métastases (N+)
ou pas (N0)



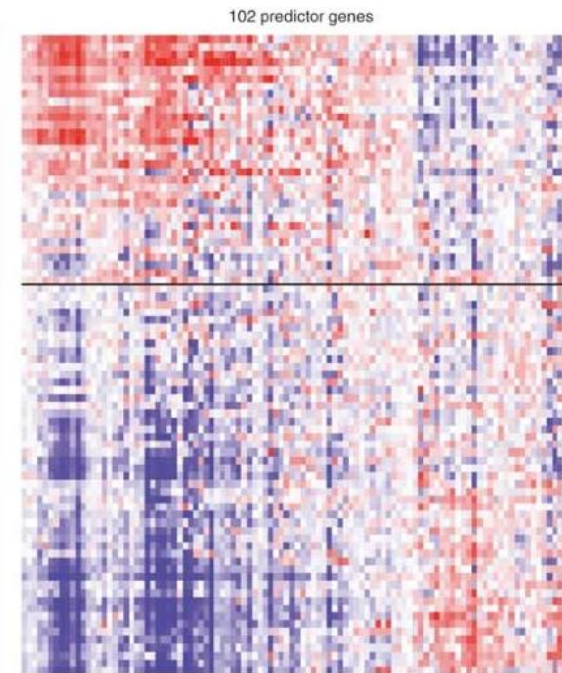
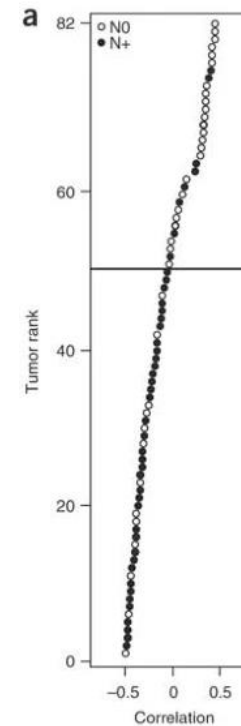
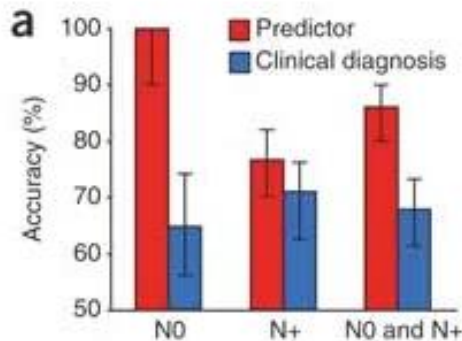
Supervised hierarchical clustering

Puces ADN : apport en cancérologie

Classification de tumeurs en sous-groupes moléculaires (cancer du sein, lymphome, medulloblastome, ...)

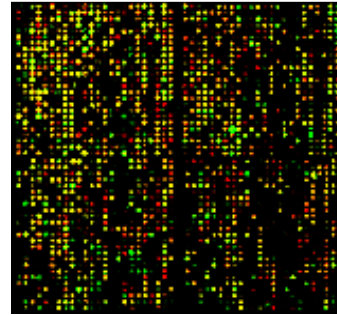
- **Ajustement traitement:**
Head and neck squamous cell carcinomas (HNSCCs)
(Roepman et al., 2005, Nature Genetics)

Prédiction du risque de métastases plus efficace que le diagnostic clinique

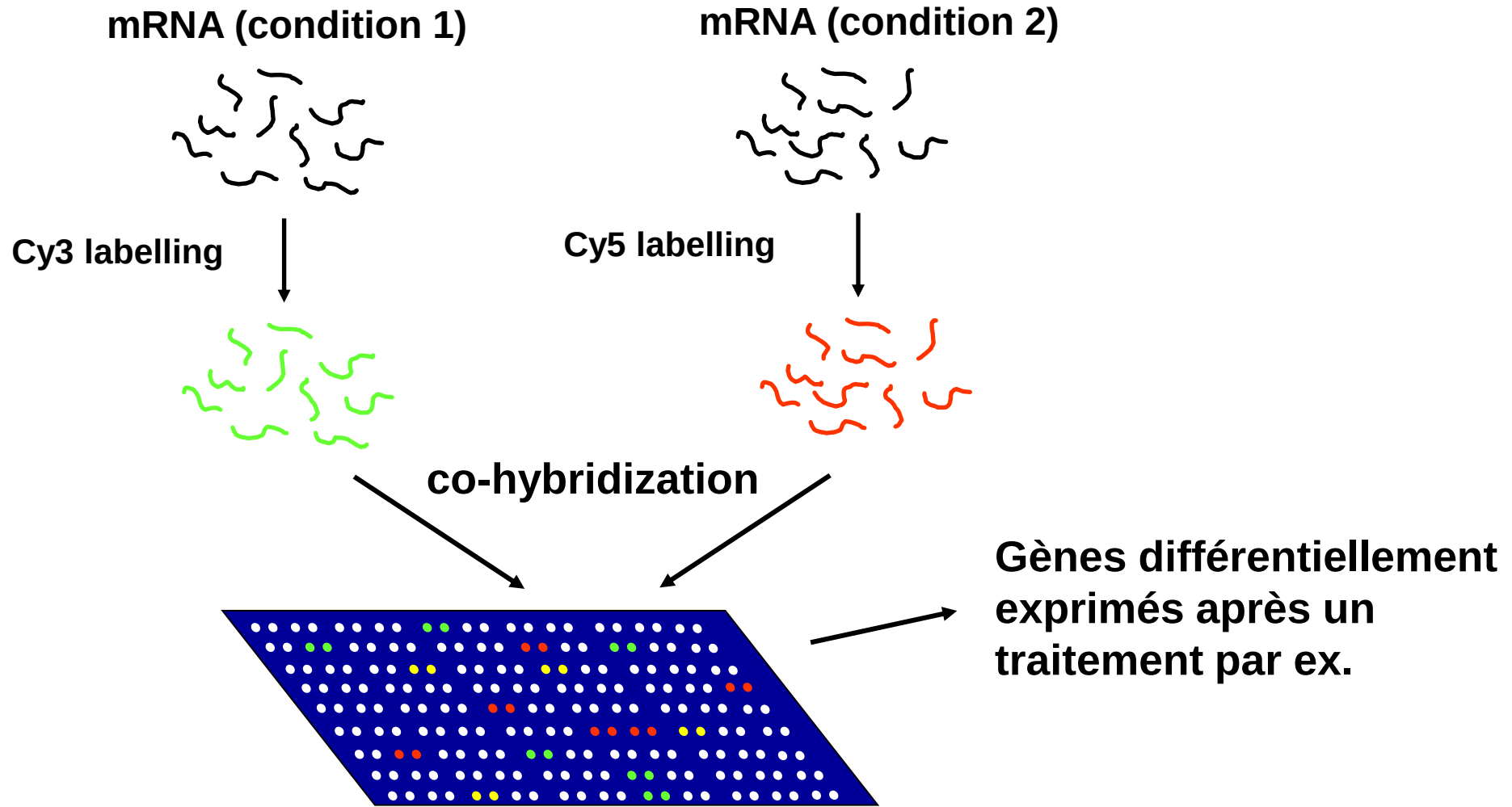


Supervised hierarchical clustering

Autres types de puces ADN



Double marquage



Puces à ADN pour l'épissage

23.000 genes

DNA

↓ transcription

30.000 – 50.000 pre mRNA

pre mRNA

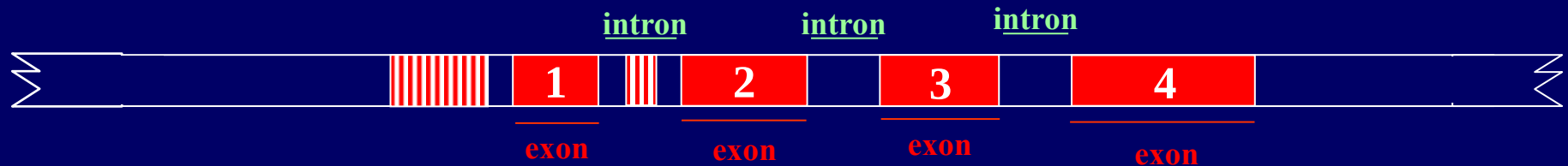
↓ splicing

150.000 – 250.000 mRNA

mature mRNA

Alternative splicing: one gene, several mRNAs

23.000 genes



Pre mRNA



alternative splicing



mRNAs

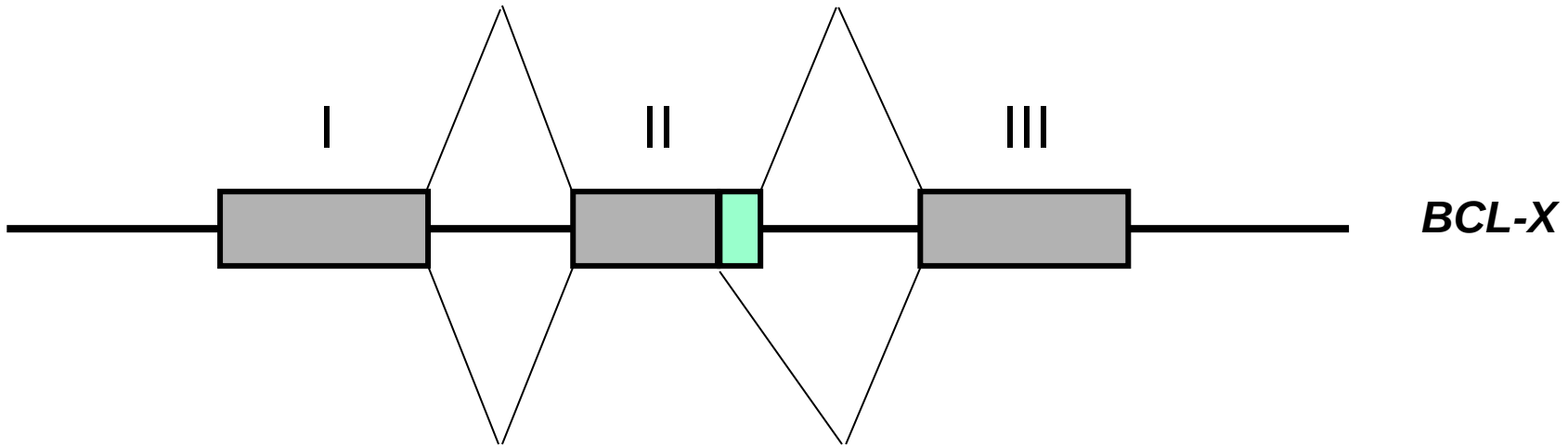


150.000 – 250.000 RNAs

Importance du splicing



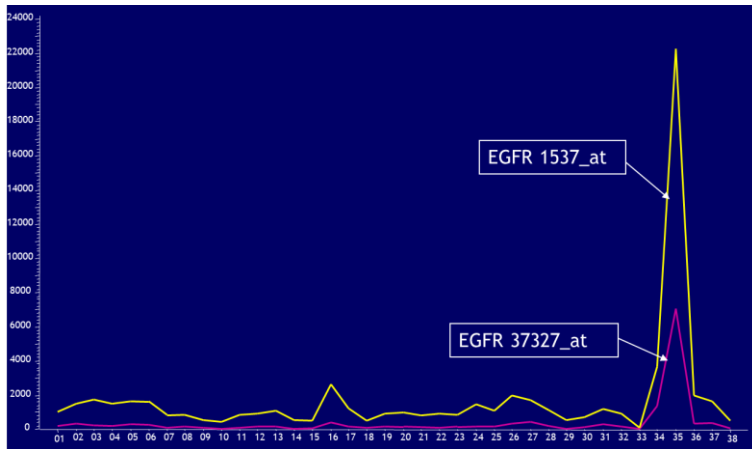
Bcl-x_L
antiapoptotic



Bcl-x_S
proapoptotic

Remarques sur les puces ADN

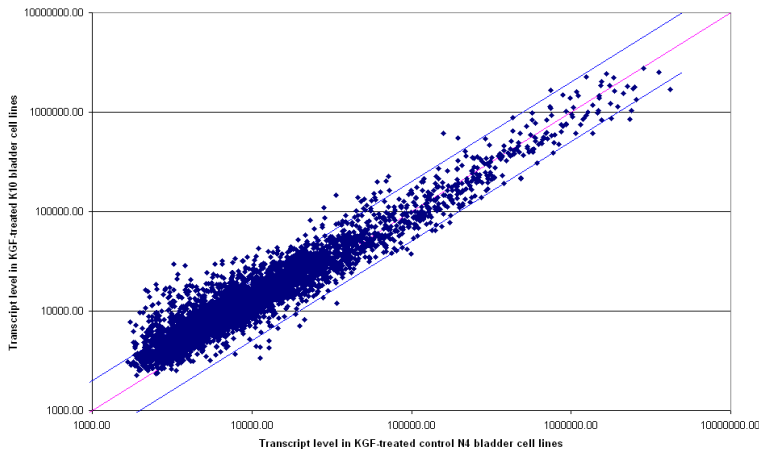
- Après un résultat obtenu par puce à ADN -> Avant de se focaliser sur un gène en particulier, il est nécessaire de confirmer le résultat par une autre approche (RT-qPCR par ex)
- Les sondes utilisées ont une influence sur la sensibilité des mesures



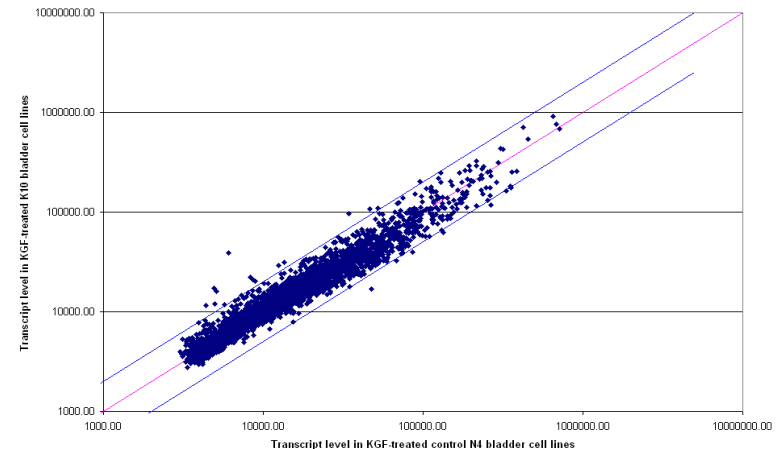
Remarques sur les puces ADN

- Après un résultat obtenu par puce à ADN -> Avant de se focaliser sur un gène en particulier, il est nécessaire de confirmer le résultat par une autre approche (RT-qPCR par ex)
- Les sondes utilisées ont une influence sur la sensibilité des mesures
- Nécessaire de répéter les mesures

**Scatter plot comparing the $\log_{10}$ (expression level)
from KGF-treated N4 and KGF-treated K10 bladder cell lines**



1 replicate / sample

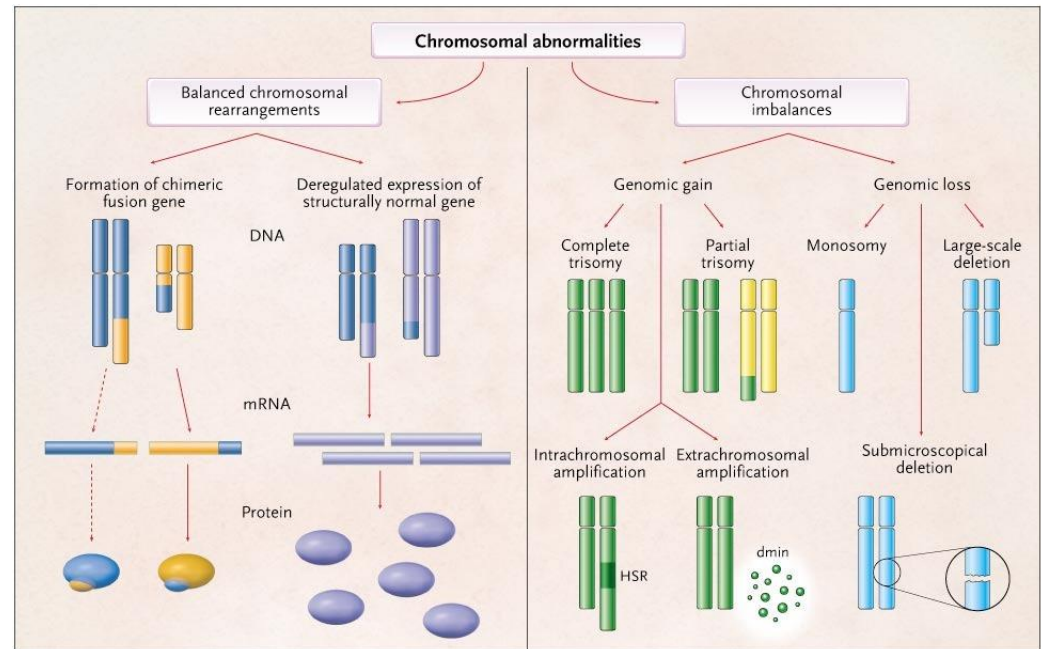
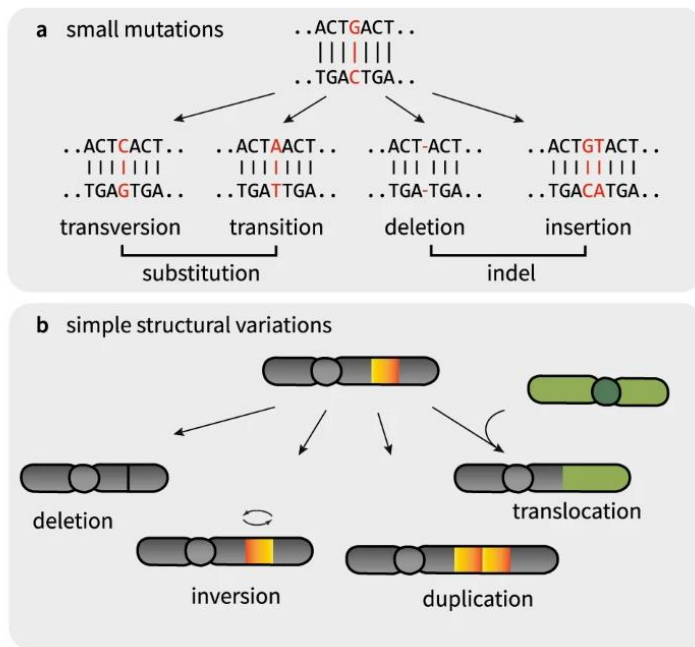


4 replicates / sample

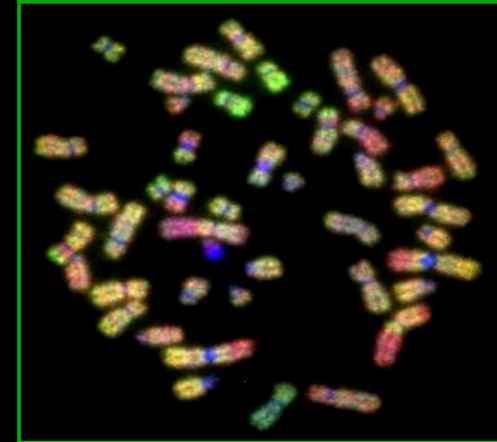
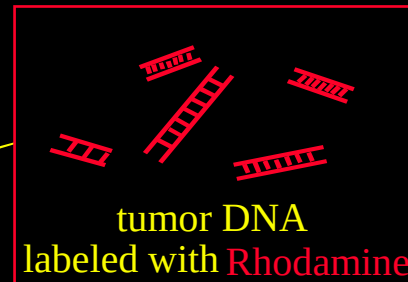
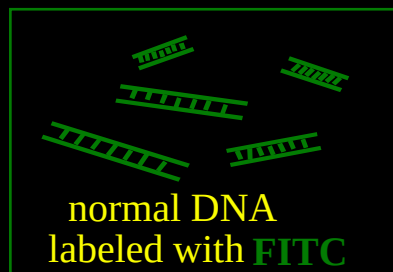
Puces à ADN pour étudier les variations génomiques

Diverses variations génomiques observées dans les cancers

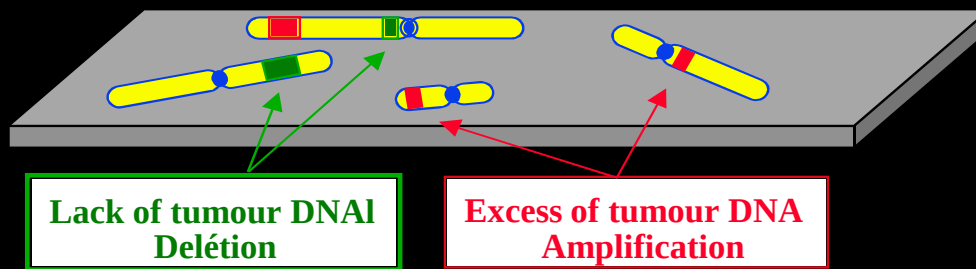
Fig. 2: Types of basic genomic variations.



Identification of DNA gains and losses using CGH (comparative genomic hybridization)

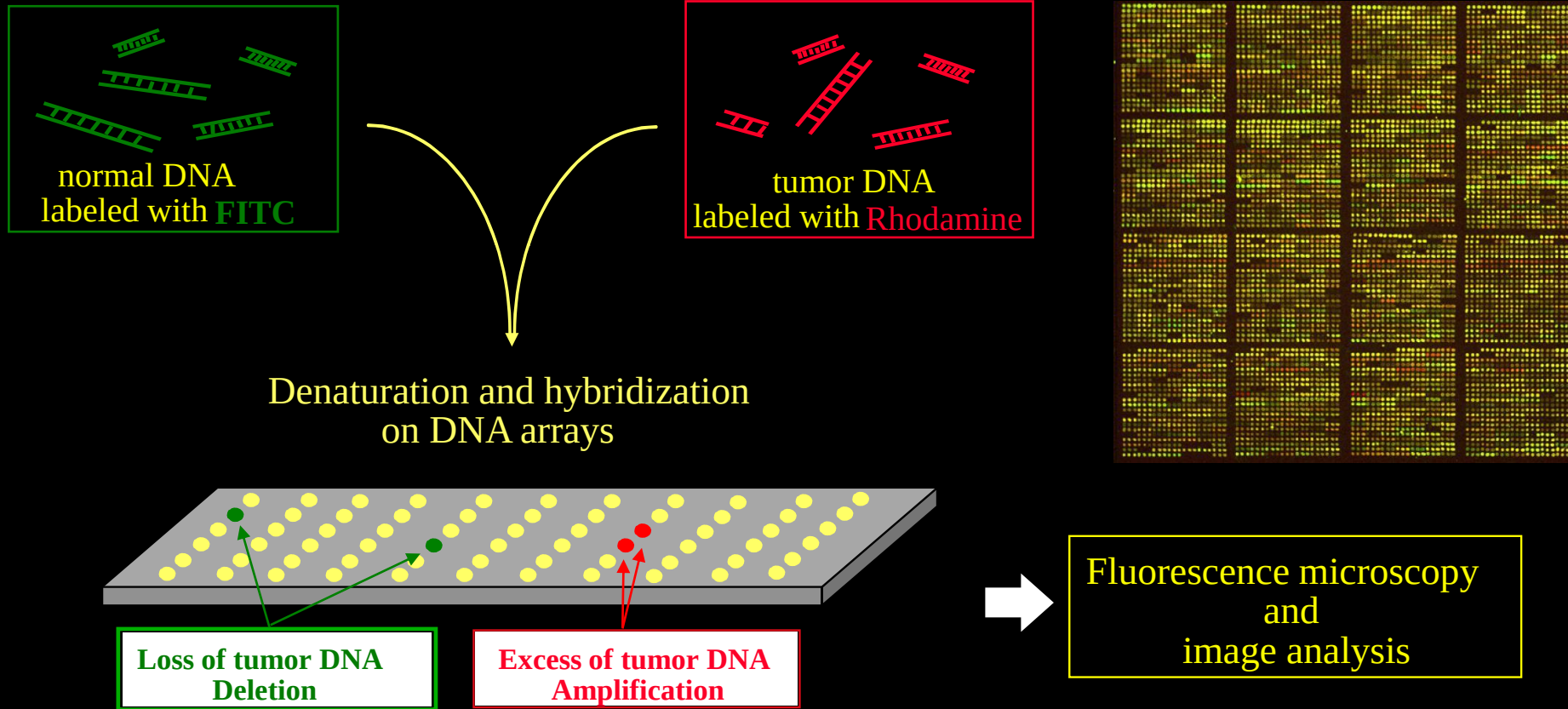


Denaturation and hybridization
on normal chromosomes



Fluorescence microscopy
and
image analysis

CGH (comparative genomic hybridization) on DNA arrays



Adapted from A. Aurias, INSERM U830, Institut Curie

Puces à ADN (CGH-array) : application pour le diagnostic de déséquilibres cytogénétiques cryptiques

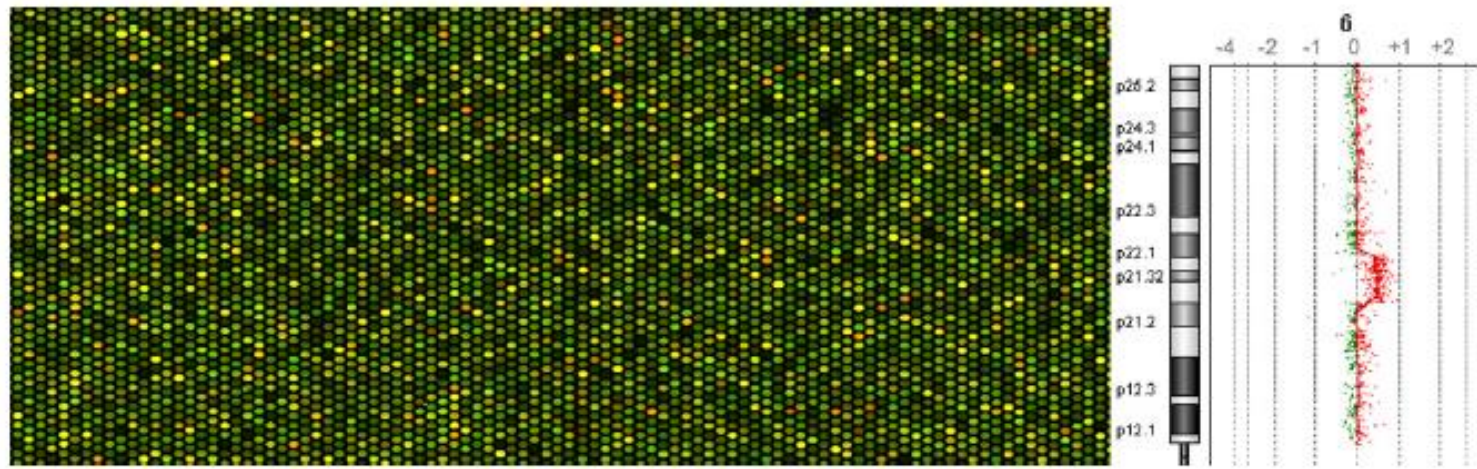


Fig. 1. Détail d'une puce ADN AgilentTM après hybridation. Ratios de fluorescence mesurés par le logiciel CGH-analyticsTM et présentés sous forme graphique (exemple d'une duplication 6p21.2p22.1).

Sur la puce: oligonucéotides de 100-200 bp dont la localisation sur le chromosome est connue

Analyse large échelle grâce aux puces

Transcriptome	cDNA array (≥ 1996) miRNA array ($\geq \sim 2004$)
Genome	CGH array, copy number (≥ 1998) CGH array, snp (≥ 2004)
DNA methylation	methylation arrays (≥ 2004)
High throughput sequencing ($\geq \sim 2008$)	
~ Proteome	protein array (≥ 2000)
~ Functional studies	microarrays of cells (≥ 2001)

Séquençage haut-débit

Ou comment faire (presque) tout

Identification des variants

Identification des mutations

Identification des translocations

Identification des gain et perte chromosomal

Quantification de l'expression des ARN

Identification des ilots CpG méthylés

Séquençage classique versus Séquençage à haut-débit

Sanger
1st generation sequencing

DNA → DNA cloning

Gel analysis (96 capillaries)

600 – 800 bases

75 kilo bases / experiment

10^3

highthroughput sequencing
2nd generation sequencing
(HTS, next-gen sequencing)

« no cloning step »

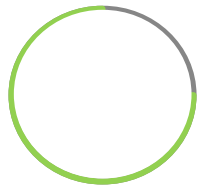
Millions of sequences analysed
simultaneously

100 – 200 bases

200 Giga bases / experiment

10^9

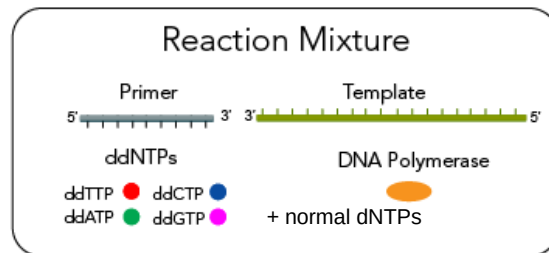
Sanger sequencing



ADN à séquencer cloné avec une séquence permettant la fixation d'amorce

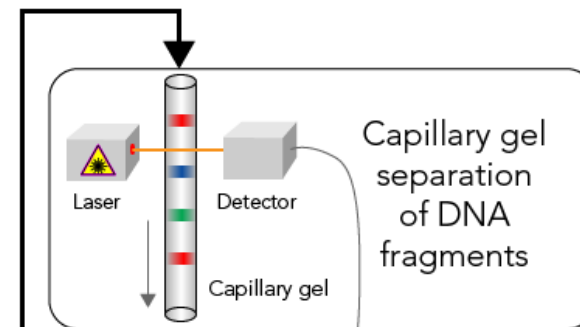
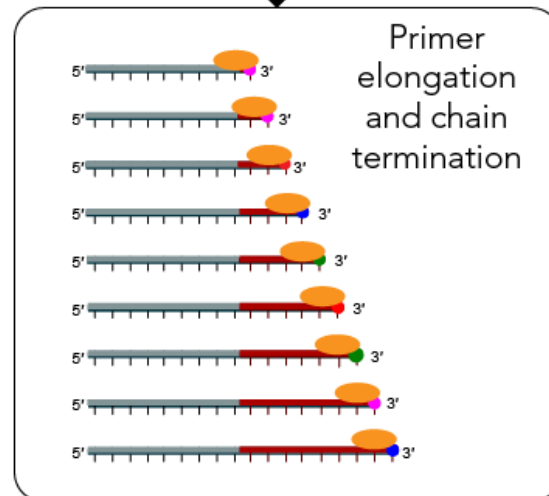
ddNTPs
ddTTP ● ddCTP ●
ddATP ● ddGTP ●

Utilisation de ddNTPs marqués avec des fluorochromes différents

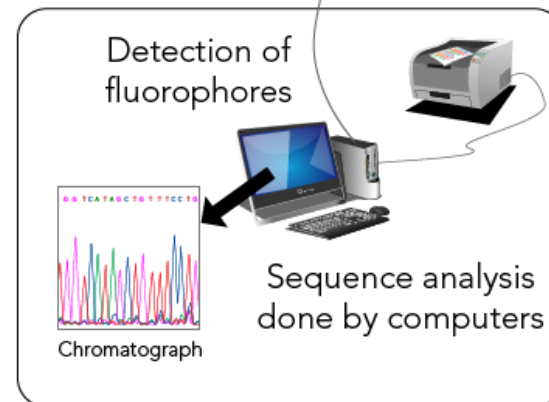


ADN amplifié
↓
Elongation arrêtée aléatoirement lors de l'ajout d'un ddNTP
↓

Génération de fragments de différentes tailles



Petits fragments passent en 1^{er}



The Human Genome Project

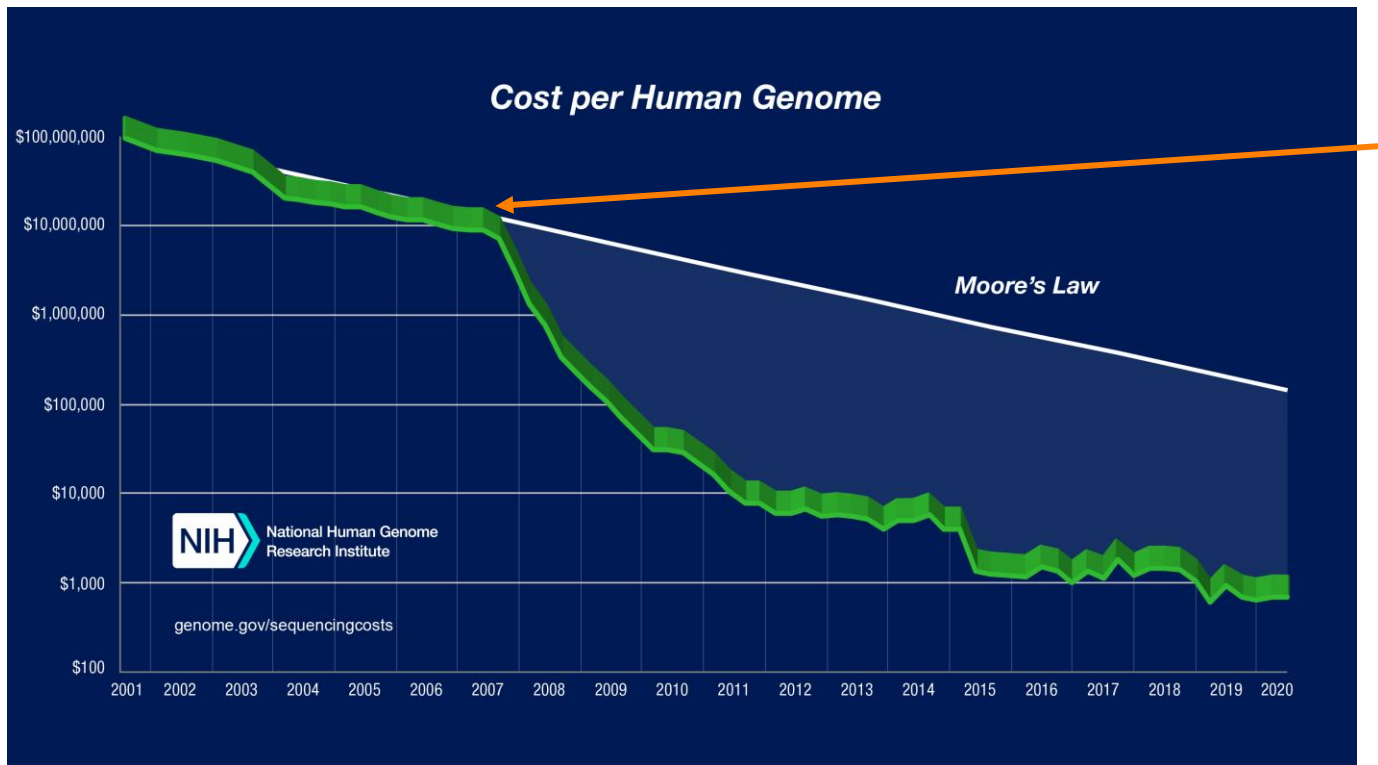
- Séquençage du génome humain = 3Gbp
- De 1990 à 2003
- ~3 Milliards \$
- 20 instituts, compagnies et laboratoires dans 6 pays
- Séquençage par méthode Sanger



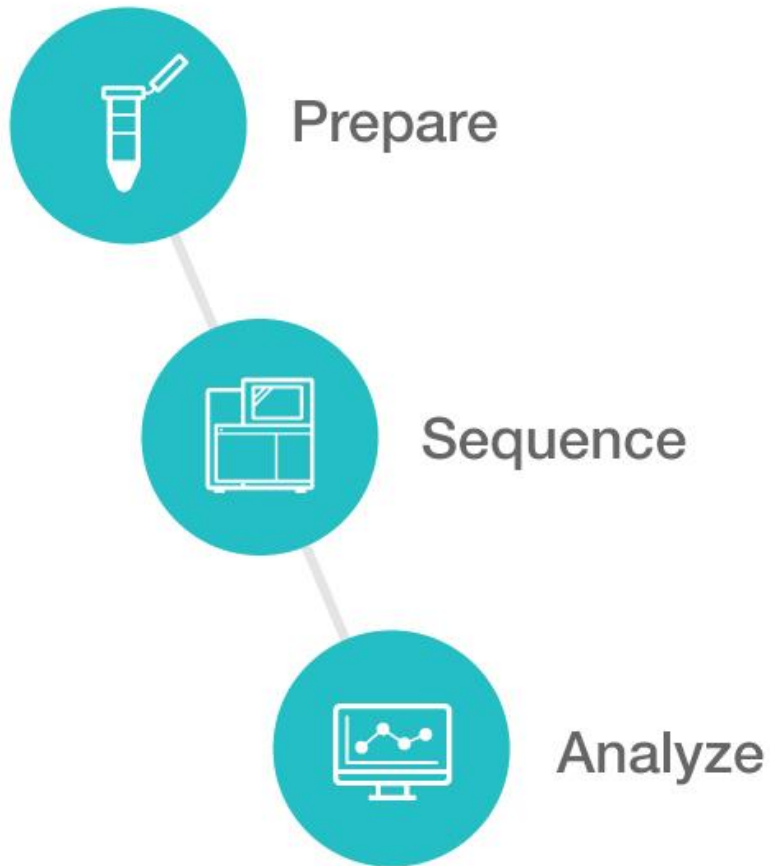
Early days: a DNA-sequencing lab in 1994.

The Human Genome Project

- 1000 Genomes Project : 2008-2012
- 100,000 Genomes Project : 2013-2018



Séquençage haut-débit (Next-gen sequencing or NGS)



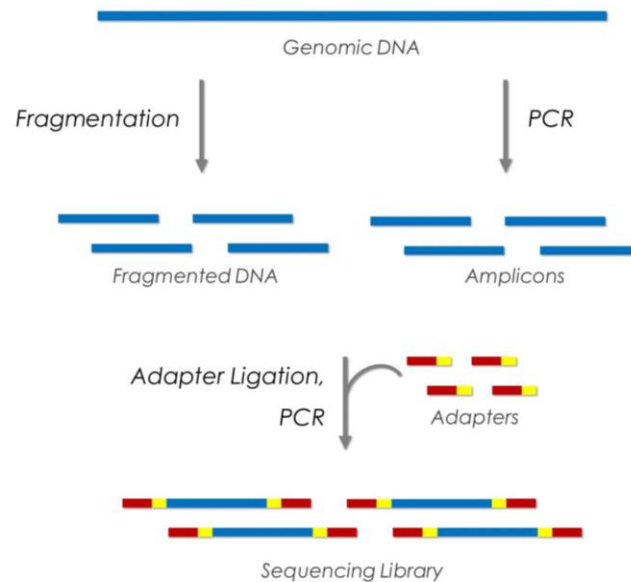
Séquençage haut-débit (Next-gen sequencing or NGS)



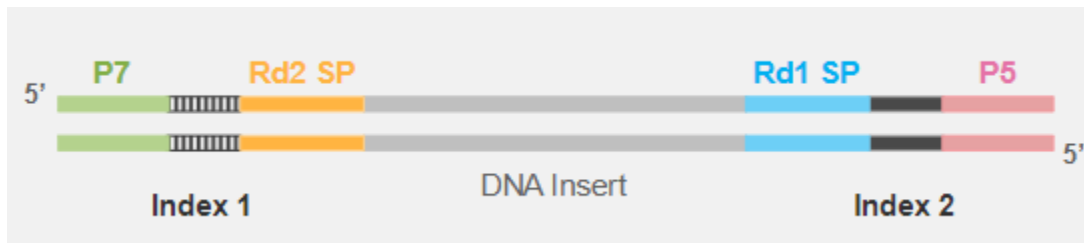
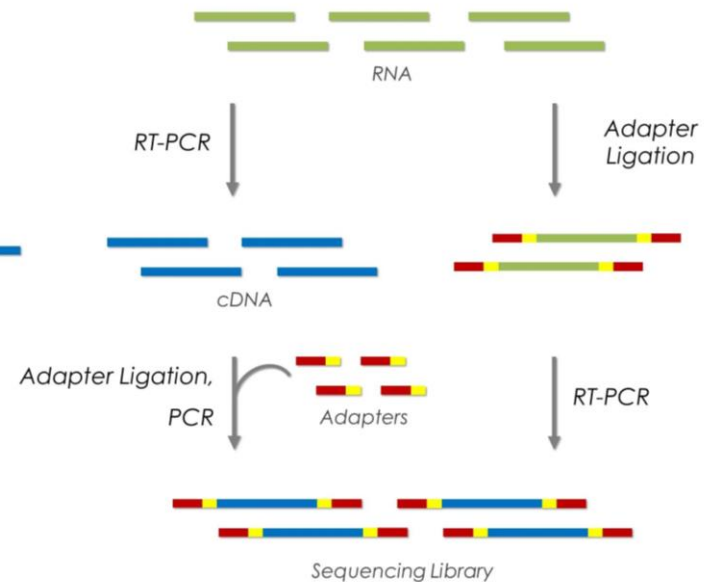
Prepare

50-200 bp

DNA Library Construction



RNA Library Construction



P5 et P7 : attachement au support
SP: sequencing primer

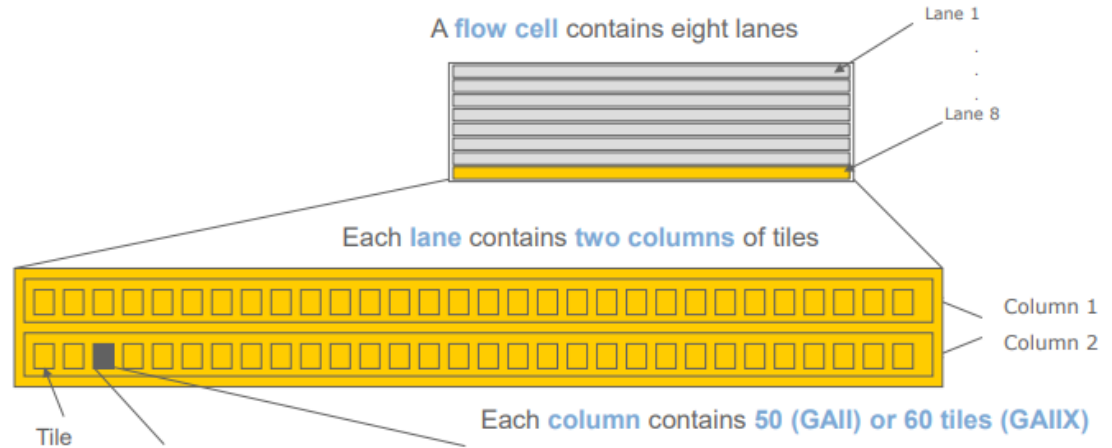
Séquençage haut-débit (Next-gen sequencing or NGS)



Prepare

ADN ajouté sur la « flow cell »

Flow cell : taille d'une lame de microscopie

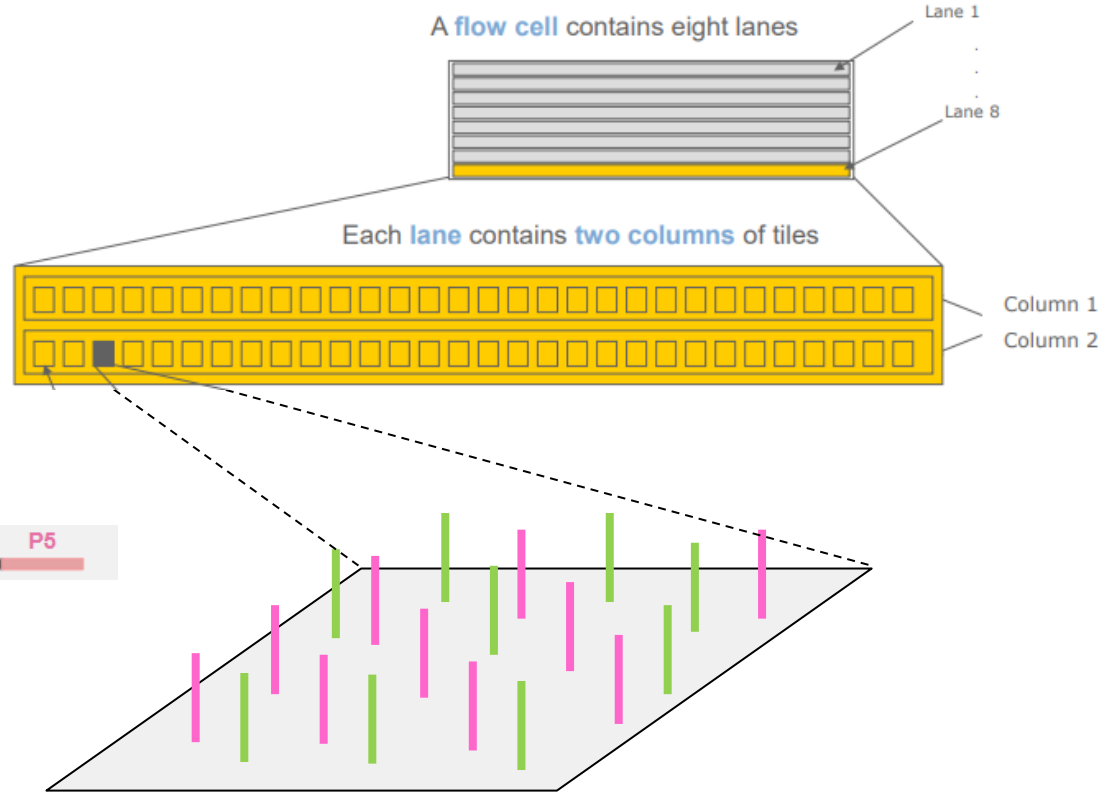


Séquençage haut-débit (Next-gen sequencing or NGS)



Prepare

ADN ajouté sur la « flow cell »



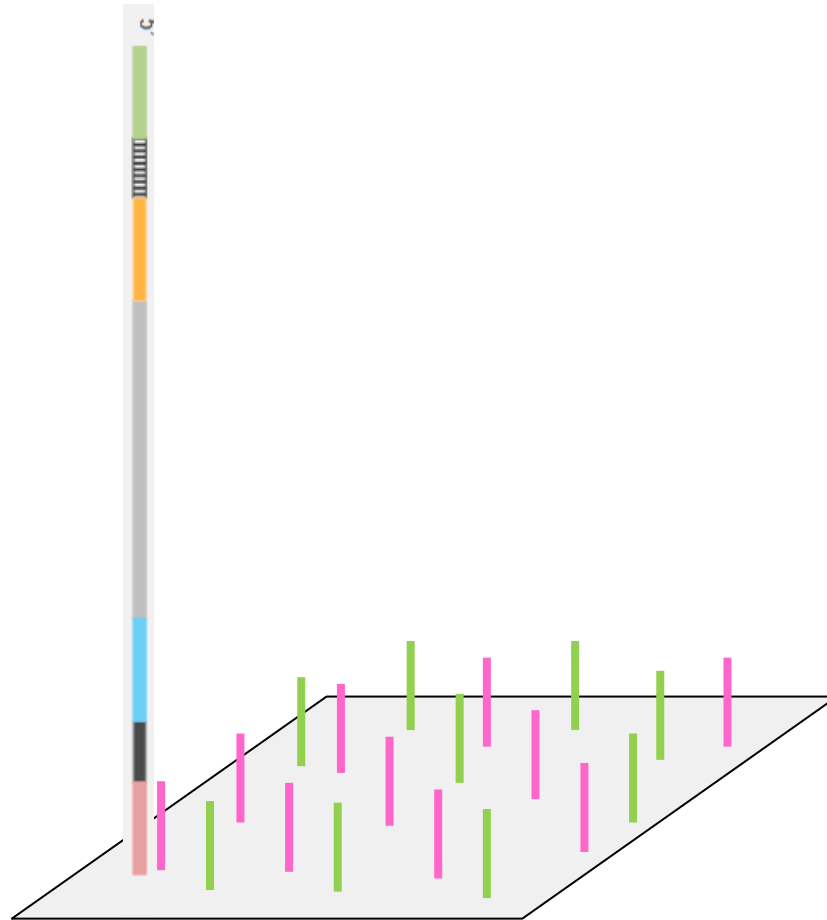
Plaques recouvertes d'oligonucléotides complémentaires aux séquences ajoutés dans les adapters

Séquençage haut-débit (Next-gen sequencing or NGS)



Prepare

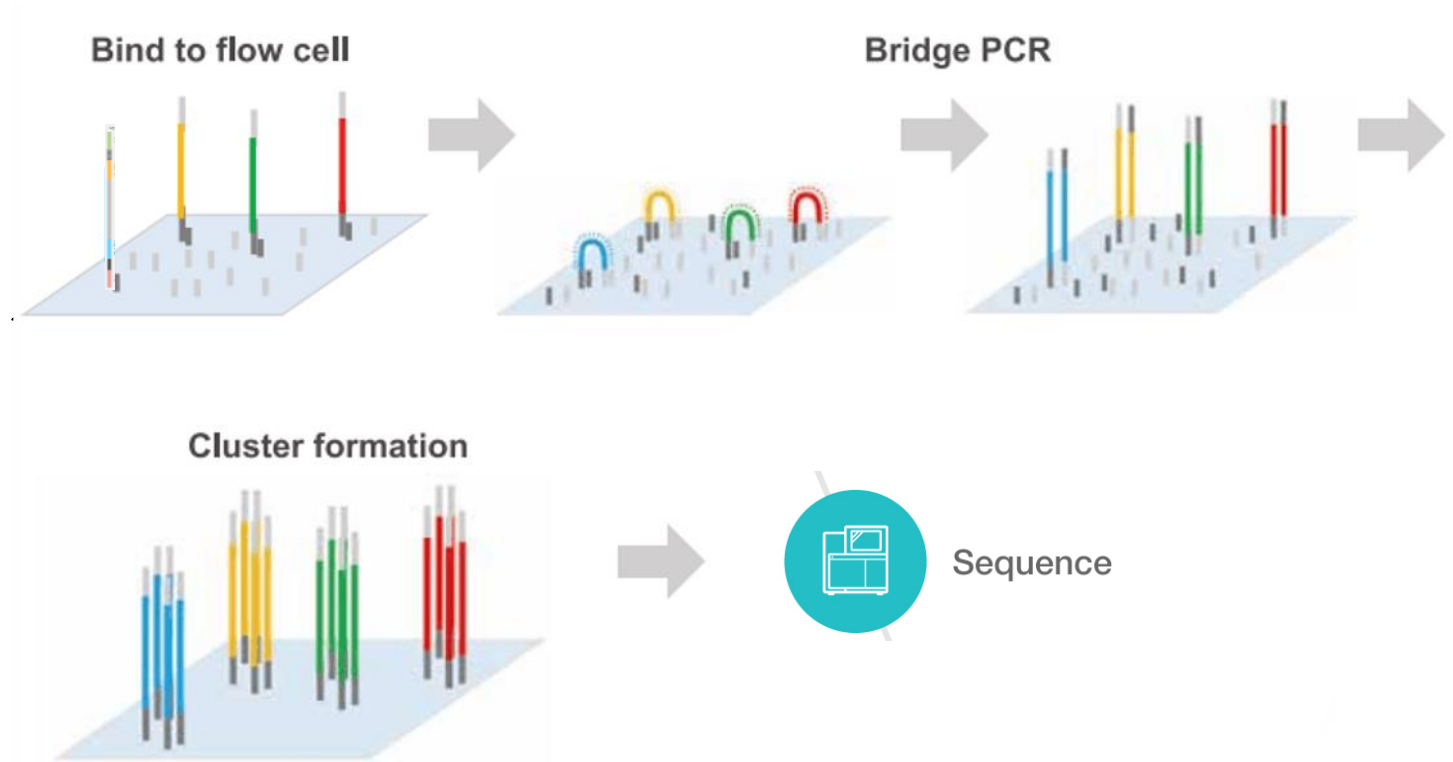
ADN ajouté sur la « flow cell »



Séquençage haut-débit (Next-gen sequencing or NGS)



Prepare

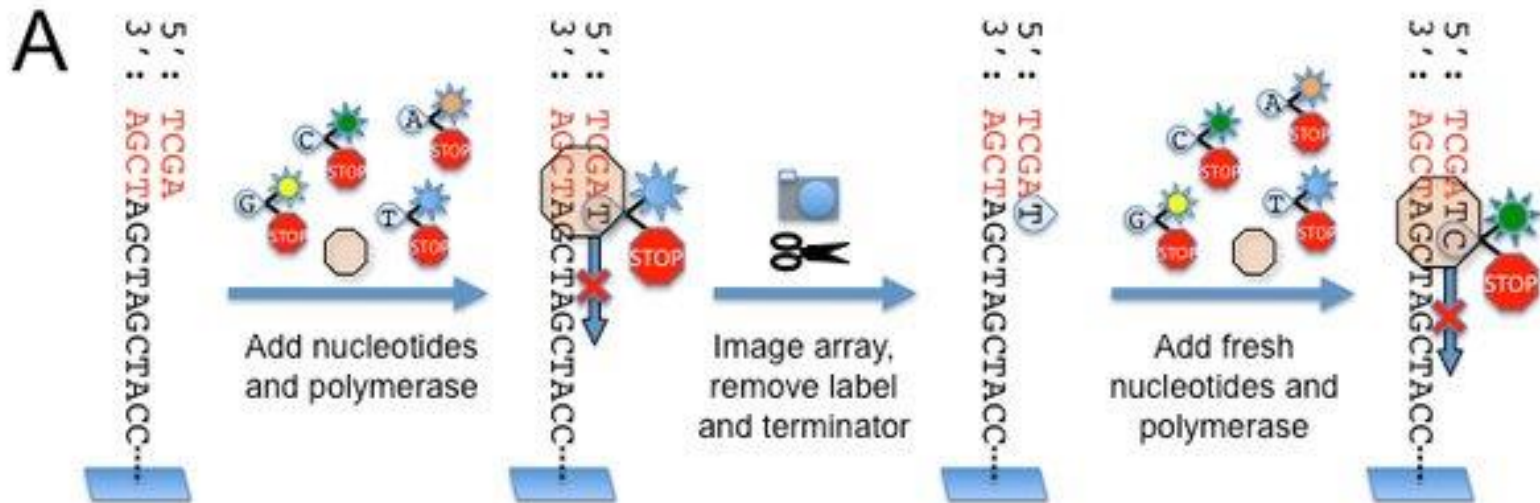


Chaque cluster correspond à un fragment généré au début de la préparation

Séquençage haut-débit (Next-gen sequencing or NGS)



Sequence

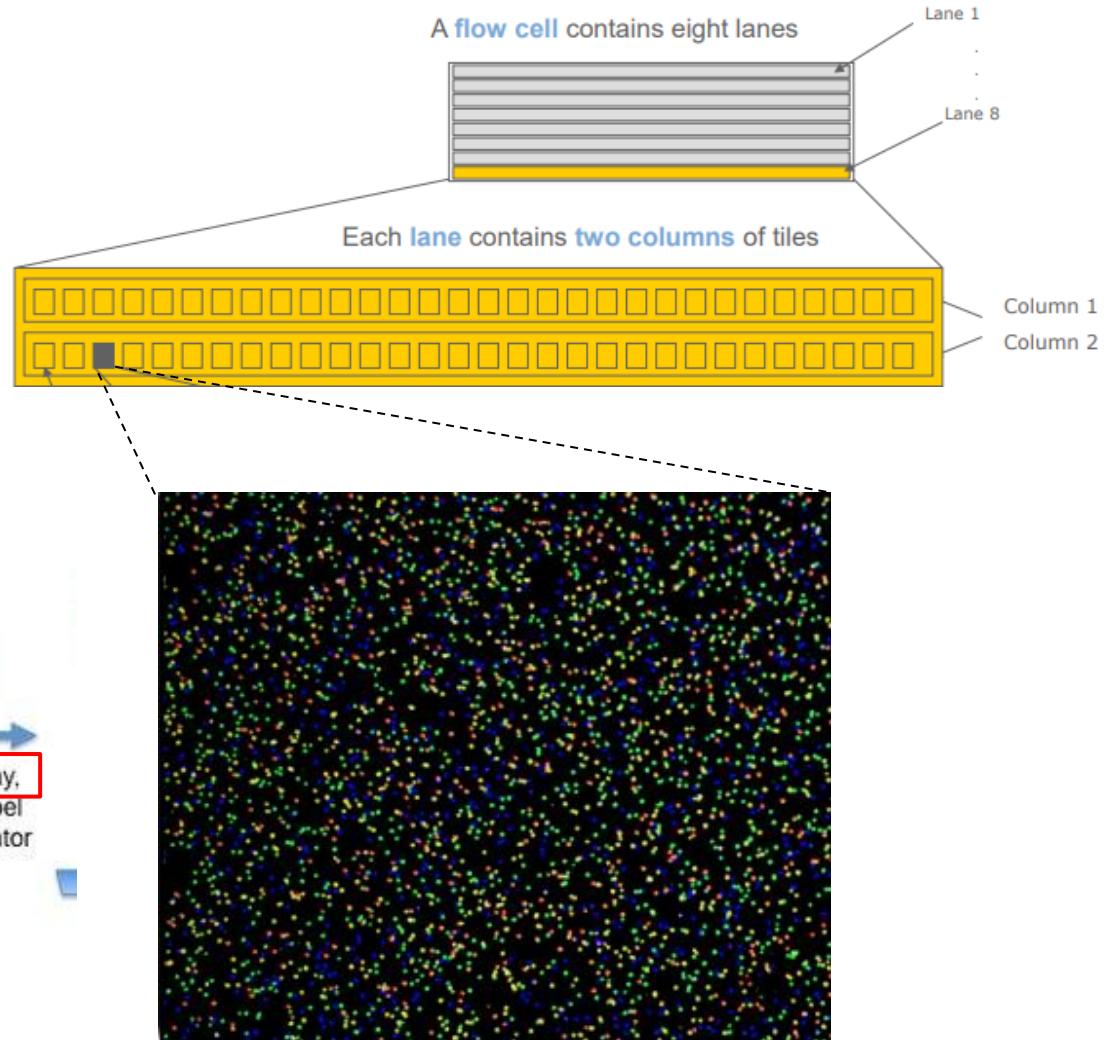


Nucléotides stoppant
l'élongation de
manière réversible

Séquençage haut-débit (Next-gen sequencing or NGS)



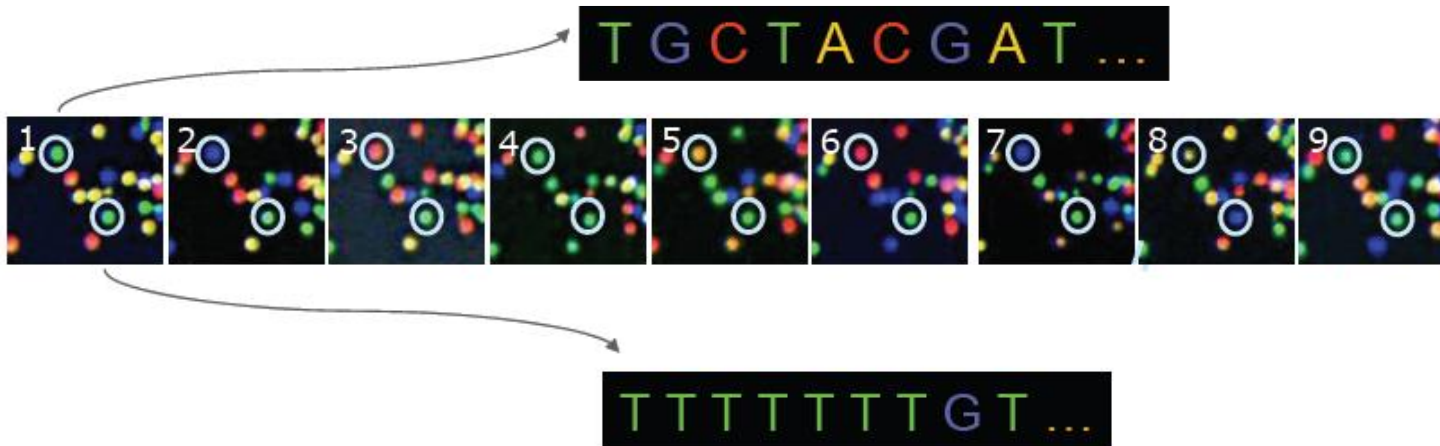
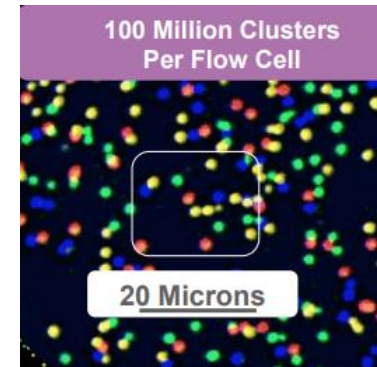
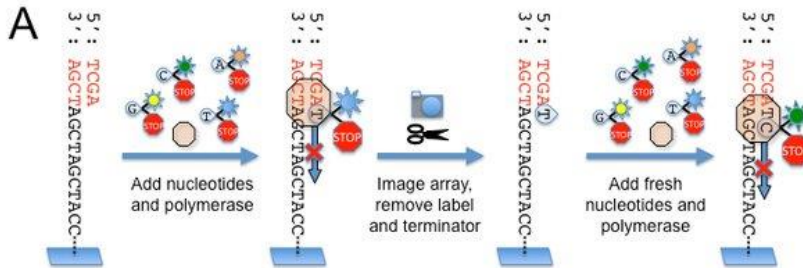
Sequence



Séquençage haut-débit (Next-gen sequencing or NGS)



Sequence



1 séquence
obtenue par
cluster
= 1 read

Séquençage haut-débit (Next-gen sequencing or NGS)



Analyze

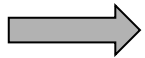
Dépend de l'objectif et du matériel de départ

RNA-seq : Séquençage cDNA issu de la RT des ARNm

- Expression différentielle
- Epissage
- Nouveau transcrit
- (Petite variation génomique)

Whole-genome or whole-exome sequencing (WGS or WES) : Séquençage de l'ADN génomique

- Variation génomique



Dans tous les cas, à la fin du séquençage : obtention de millions de petites séquences = « reads »

Séquençage haut-débit (Next-gen sequencing or NGS)

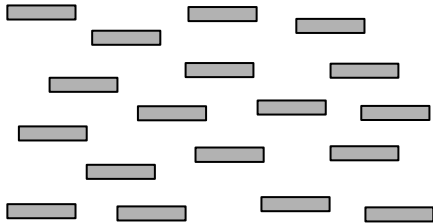


Analyse

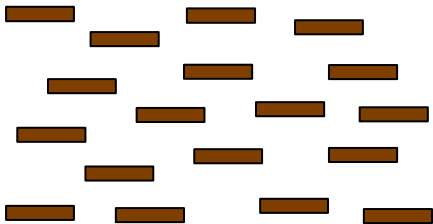
RNA-seq : Séquençage cDNA issu de la RT des ARNm

- Expression différentielle

« Reads » obtenus lors
du séquençage



Echantillon 1



Echantillon 2



Alignement
contre génome
de référence



Séquençage haut-débit (Next-gen sequencing or NGS)

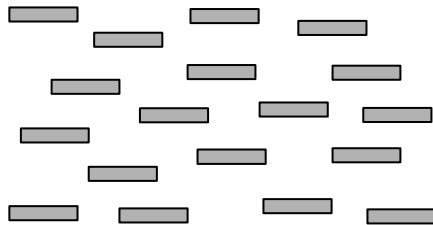


Analyse

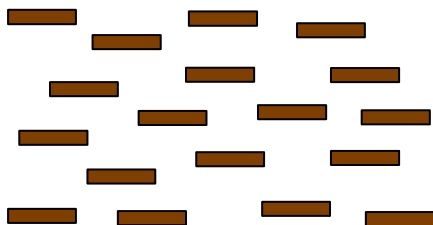
RNA-seq : Séquençage cDNA issu de la RT des ARNm

- Expression différentielle

« Reads » obtenus lors
du séquençage



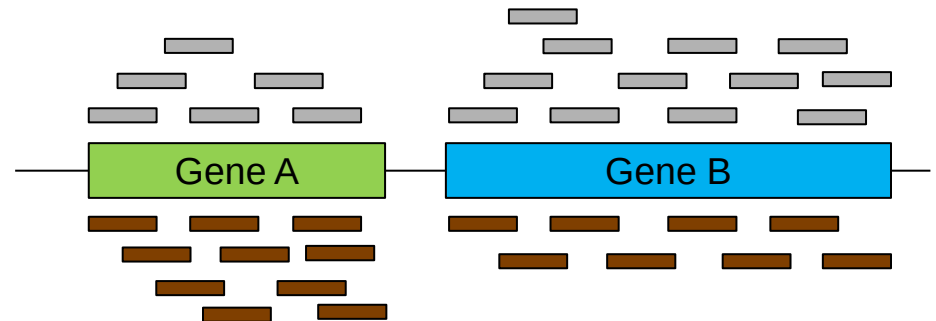
Echantillon 1



Echantillon 2



Alignement
contre génome
de référence



Expression gène A
augmentée dans
échantillon 2

Expression gène B
diminuée dans
échantillon 2

Séquençage haut-débit (Next-gen sequencing or NGS)

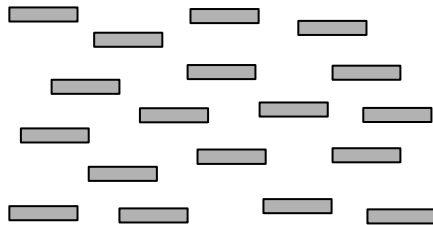


Analyse

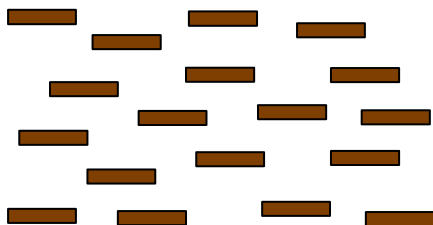
RNA-seq : Séquençage cDNA issu de la RT des ARNm

- Epissage

« Reads » obtenus lors
du séquençage



Echantillon 1

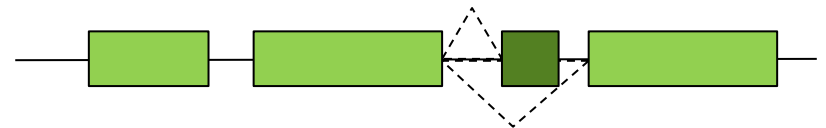


Echantillon 2

Alignement
contre génome
de référence

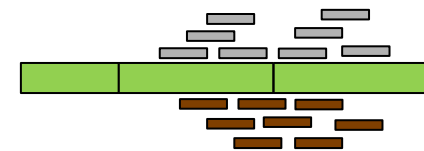
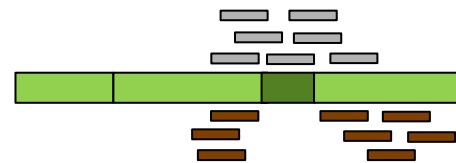


Gene A -> 2 isoformes



Isoforme A_L

Isoforme A_S



Isoforme A_L
exprimée dans
échantillon 1

Isoforme A_S
exprimée dans
échantillon 1

Séquençage haut-débit (Next-gen sequencing or NGS)

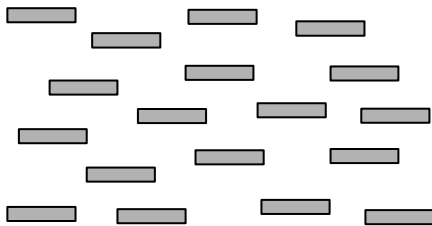


Analyse

RNA-seq : Séquençage cDNA issu de la RT des ARNm

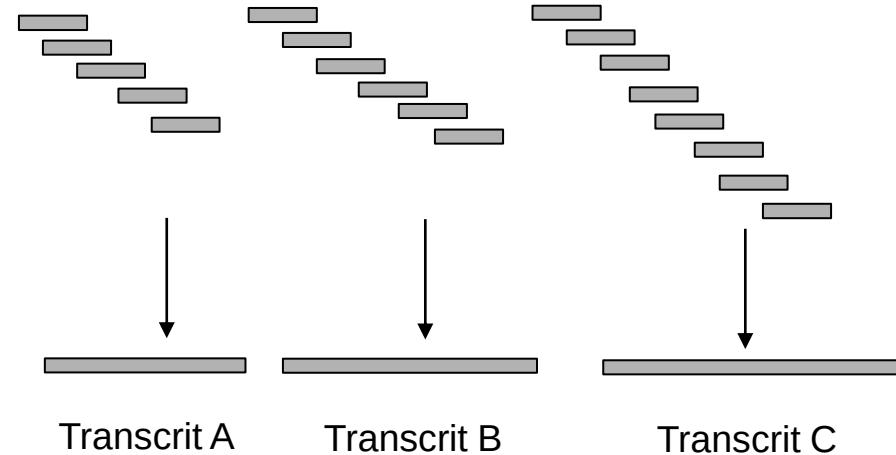
- Nouveau transcrit

« Reads » obtenus lors
du séquençage



Echantillon 1

Assemblage de
novo



**Possibilité de mettre en évidence des
transcrits non répertoriés**

Séquençage haut-débit (Next-gen sequencing or NGS)



Analyze

RNA-seq : Séquençage cDNA issu de la RT des ARNm

- Nouveau transcrit

Mittal and McDonald *BMC Medical Genomics* (2017) 10:53
DOI 10.1186/s12920-017-0289-7

BMC Medical Genomics

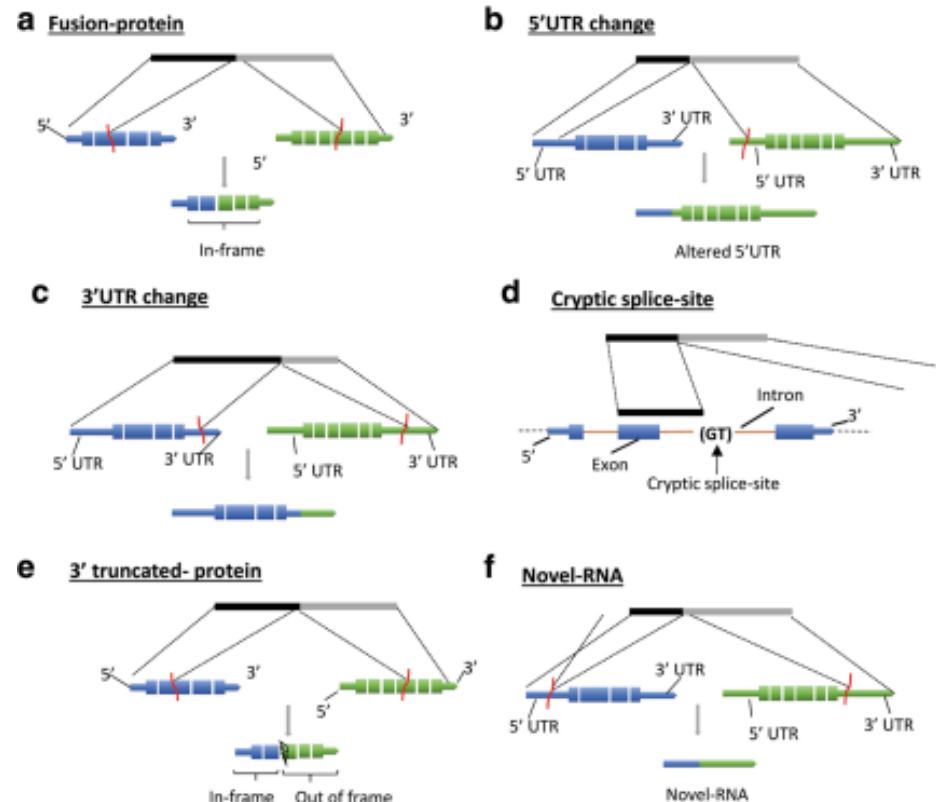
RESEARCH ARTICLE

Open Access



De novo assembly and characterization of breast cancer transcriptomes identifies large numbers of novel fusion-gene transcripts of potential functional significance

Vinay K. Mittal and John F. McDonald*



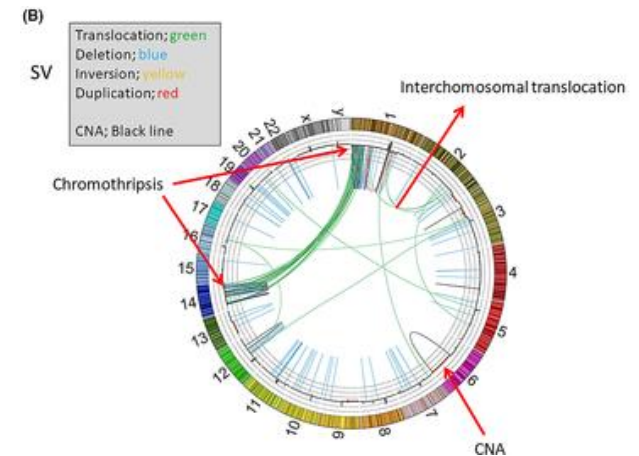
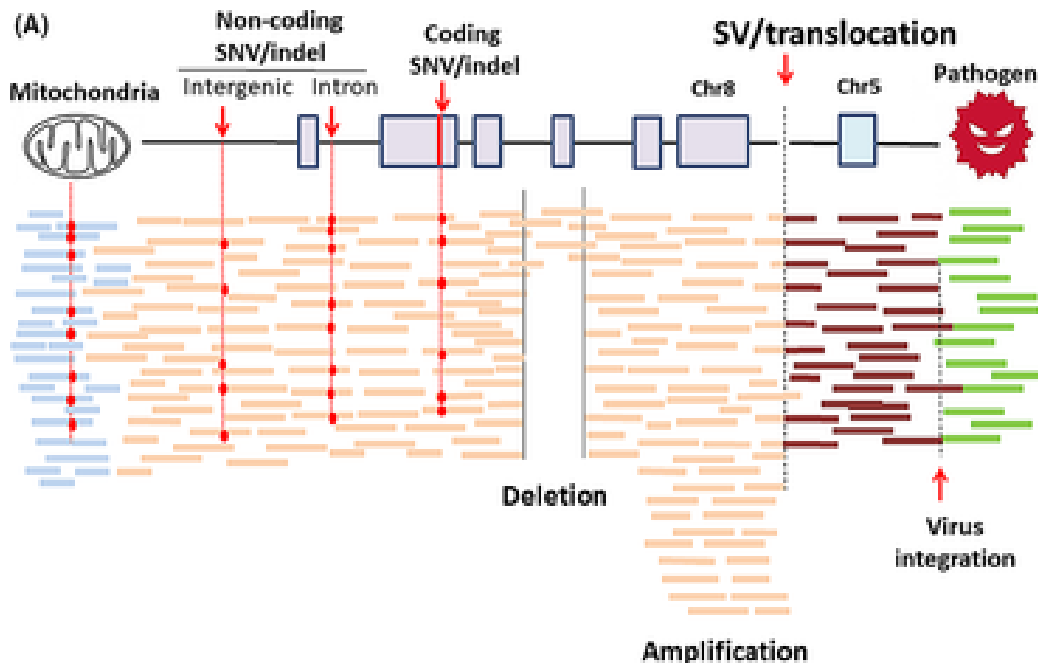
Séquençage haut-débit (Next-gen sequencing or NGS)



Analyze

Whole-genome or whole-exome sequencing (WGS or WES) : Séquençage de l'ADN génomique

- Variation génomique



A representative Circos plot of cancer genome structure from WGS analysis, which indicates SV and CNA in all human chromosomes (1-22+XY). Chromothripsis was observed in chromosomes 1 and 14. SNV, single nucleotide variants

Application NGS en cancérologie

Effort collectif pour le séquençage d'un très grand nombre de tumeurs

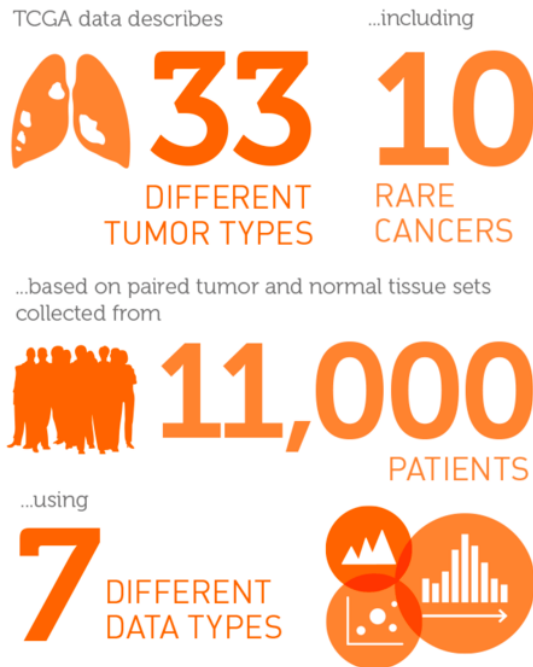
- **The cancer genome project (UK):** uses high-throughput genome sequencing to identify these somatically acquired mutations with the aim of characterising cancer genes, mutational processes and patterns of clonal evolution in human tumours
 - **The Cancer Genome Atlas (TCGA, USA) :** A comprehensive and coordinated effort to accelerate our understanding of the molecular basis of cancer through the application of genome analysis technologies, including large-scale genome sequencing.
 - **International Cancer Genome consortium (ICGC):** To obtain a comprehensive description of genomic, transcriptomic and epigenomic changes in 50 different tumor types and/or subtypes which are of clinical and societal importance across the globe. 17 countries currently participating and 42 cancer genome projects
-

Application NGS en cancérologie

En plus des données génomiques et transcriptomiques

-> Intégration de données de méthylome et protéome

-> Meilleure compréhension et description des cancers (mutations, sous-groupes)



- Le sous type basal-like du cancer du sein est très proche moléculairement des cancers séreux de l'ovaire -> Traitement similaire?
- Nouvelle classification du cancer de l'estomac
- Immense source de données pour la communauté des chercheurs

Application NGS en cancérologie

- Vers un traitement de précision?

La description des altérations génomiques retrouvés dans le « non-small cell lung » cancer par le TCGA a donné lieu à un essai clinique de médecine de précision Lung-MAP

-> Patients testés pour les diverses mutations possibles

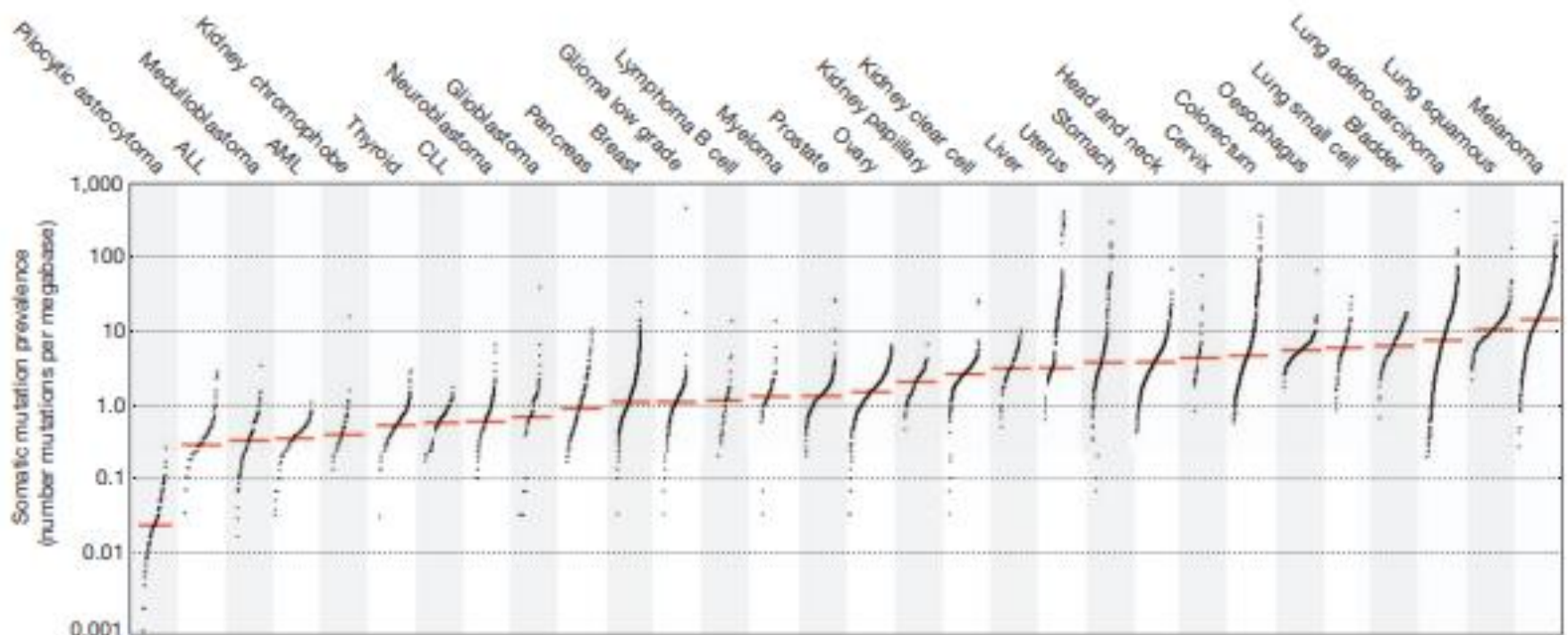
-> Traitement administré en conséquence

Application NGS en cancérologie

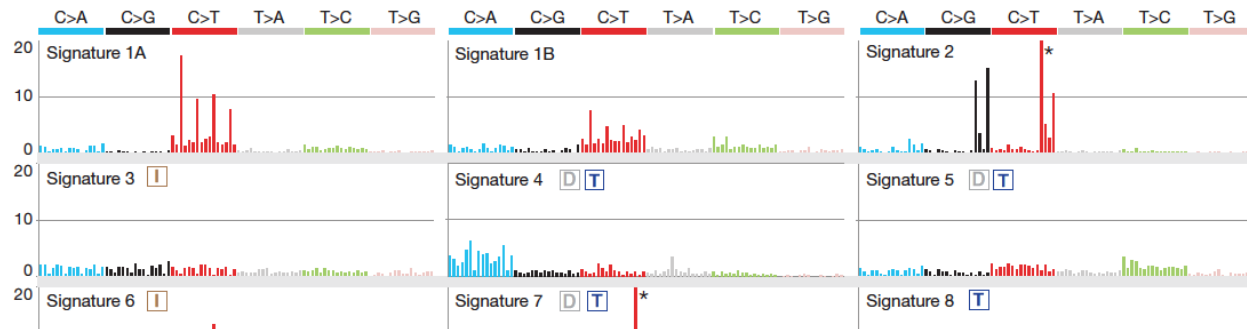
Signatures of mutational processes in human cancer

22 AUGUST 2013 | VOL 500 | NATURE | 415

Analyse de 4,938,362 mutations somatiques à partir 7,042 cancers de 30 types

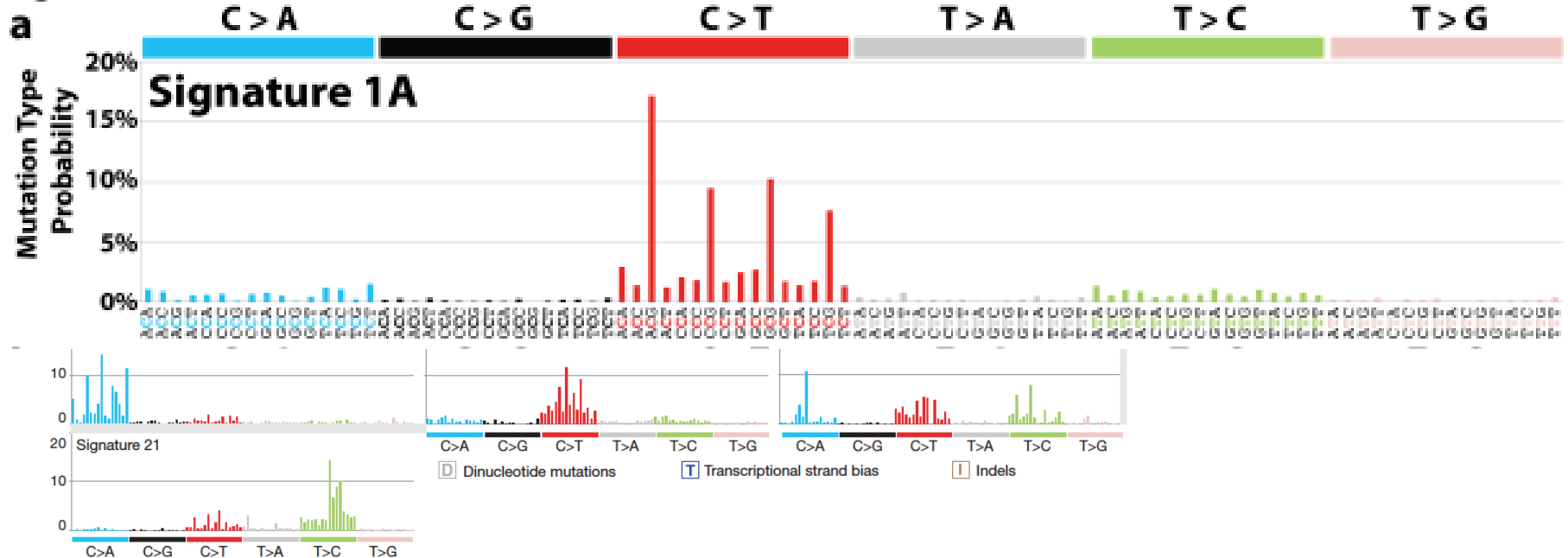


Application NGS en cancérologie

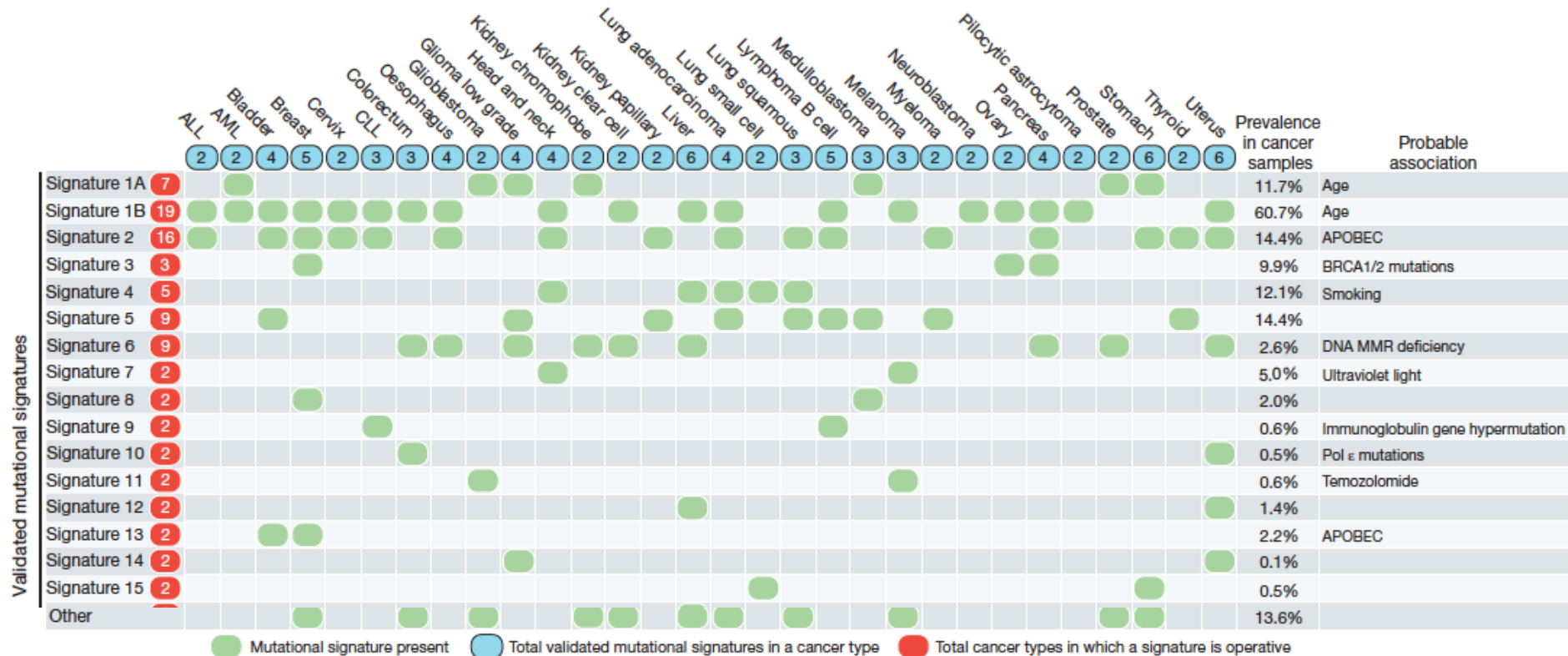


21 types de signatures peuvent caractériser les mutations

Figure S2



Application NGS en cancérologie



Types de signatures différents selon les cancers.

-> Possible lien avec l'âge, des expositions à des mutagènes ou des défauts de réparatations de l'ADN

Application NGS en cancérologie

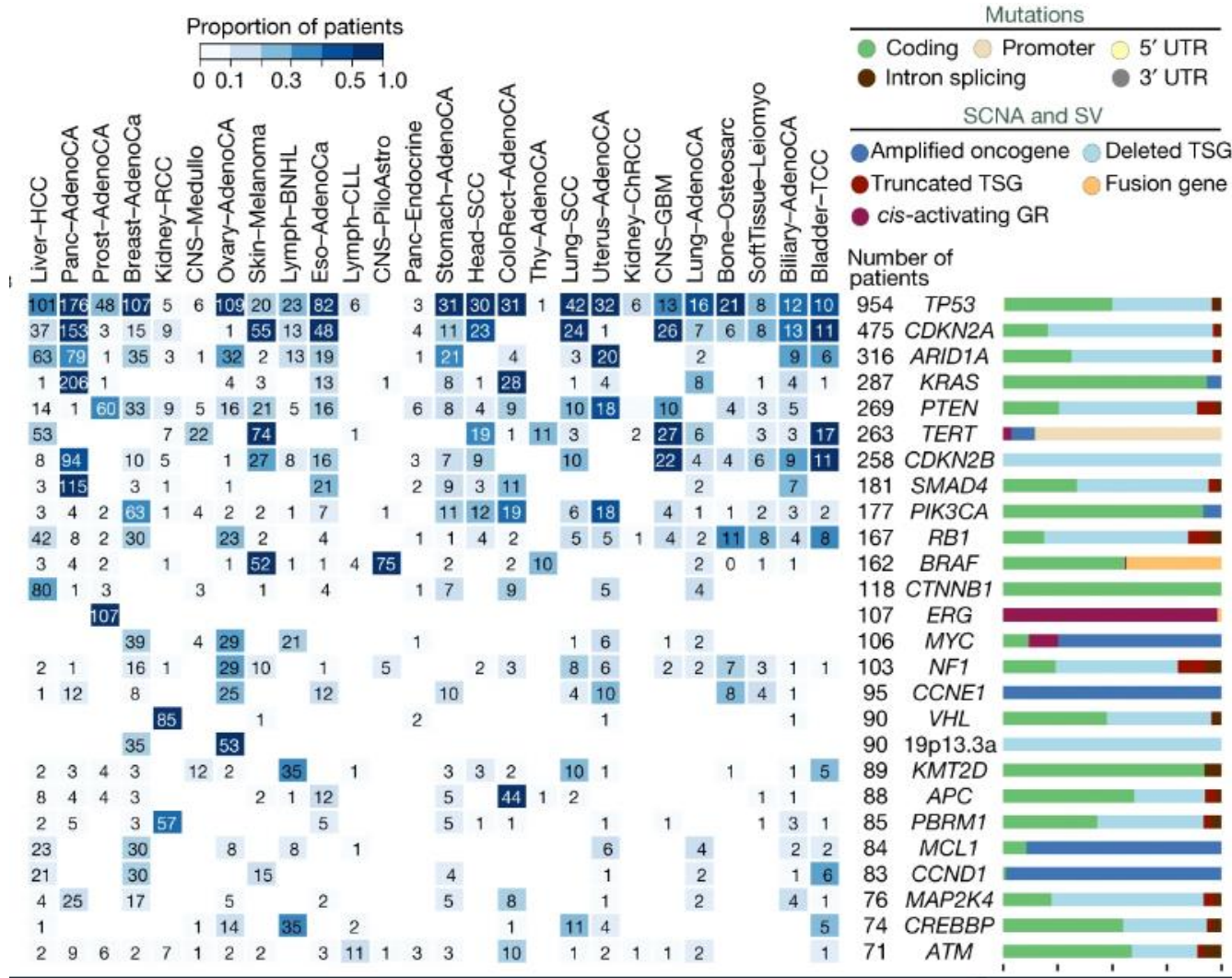
Pan-cancer analysis of whole genomes

[The ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium](#)

[Nature](#) **578**, 82–93 (2020) | [Cite this article](#)

From 2583 cancers: analysis of 43,778,859 somatic SNVs, 410,123 somatic multinucleotide variants, 2,418,247 somatic indels, 288,416 somatic SVs, 19,166 somatic retrotransposition events and 8,185 de novo mitochondrial DNA mutations

Application NGS en cancérologie



Mutations drivers probables

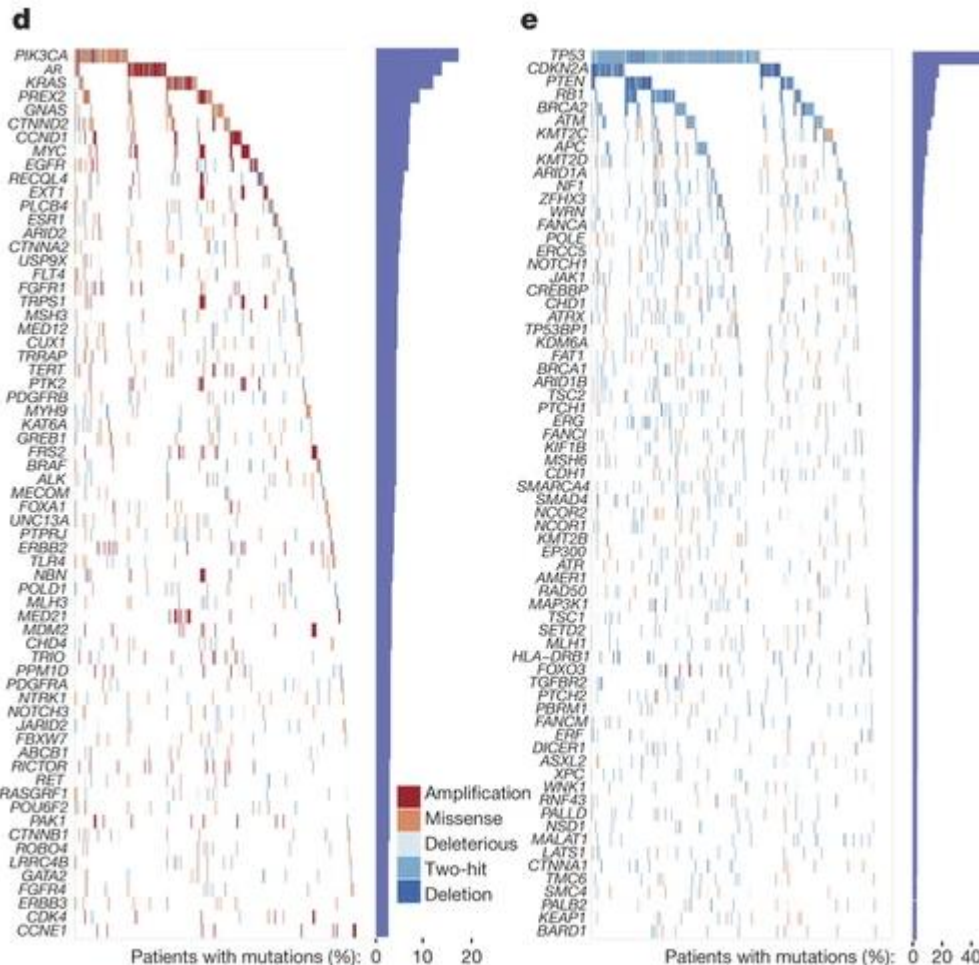


Au moins 1 mutation driver identifiée dans 91% des cancers étudiés

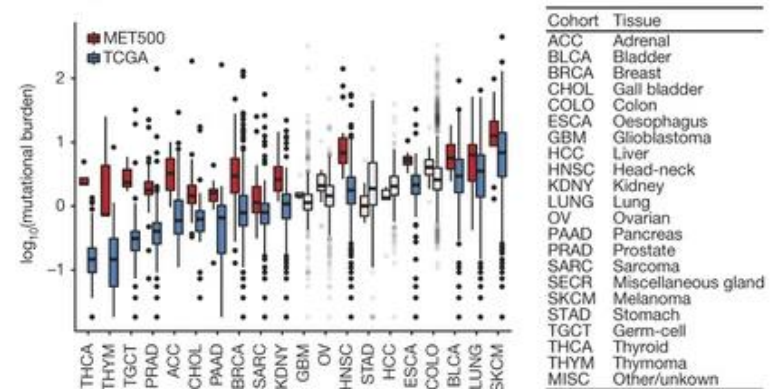
Application NGS en cancérologie

Integrative clinical genomics of metastatic cancer

17 AUGUST 2017 | VOL 548 | NATURE | 297



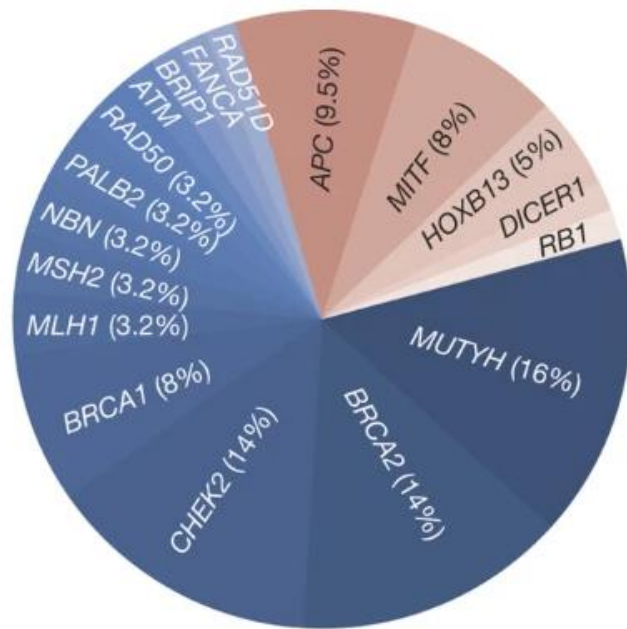
- Mutations drivers similaires observés lors de l'analyse de 500 métastases (20 types de cancers)
- Taux de mutations plus élevé dans métastases vs tumeurs primaires



Application NGS en cancérologie

Integrative clinical genomics of metastatic cancer

17 AUGUST 2017 | VOL 548 | NATURE | 297



En plus des mutations somatiques, mutations germinales dans ~15% des patients.

Principalement dans gènes impliqués dans la réparation de l'ADN

Application NGS en cancérologie

- **Beaucoup de mutations passagères**
 - **Beaucoup de mutations drivers à faible fréquence**
 - **Peu de mutations drivers à haute fréquence**
-

Cancer Genome Landscapes

Bert Vogelstein, Nickolas Papadopoulos, Victor E. Velculescu, Shibin Zhou,
Luis A. Diaz Jr., Kenneth W. Kinzler*

1. Most human cancers are caused by two to eight sequential alterations that develop over the course of 20 to 30 years.
2. Each of these alterations directly or indirectly increases the ratio of cell birth to cell death; that is, each alteration causes a selective growth advantage to the cell in which it resides.
3. The evidence to date suggests that there are ~140 genes whose intragenic mutations contribute to cancer (so-called Mut-driver genes). There are probably other genes (Epi-driver genes) that are altered by epigenetic mechanisms and cause a selective growth advantage, but the definitive identification of these genes has been challenging.
4. The known driver genes function through a dozen signaling pathways that regulate three core cellular processes: cell fate determination, cell survival, and genome maintenance.
5. Every individual tumor, even of the same histopathologic subtype as another tumor, is distinct with respect to its genetic alterations, but the pathways affected in different tumors are similar.

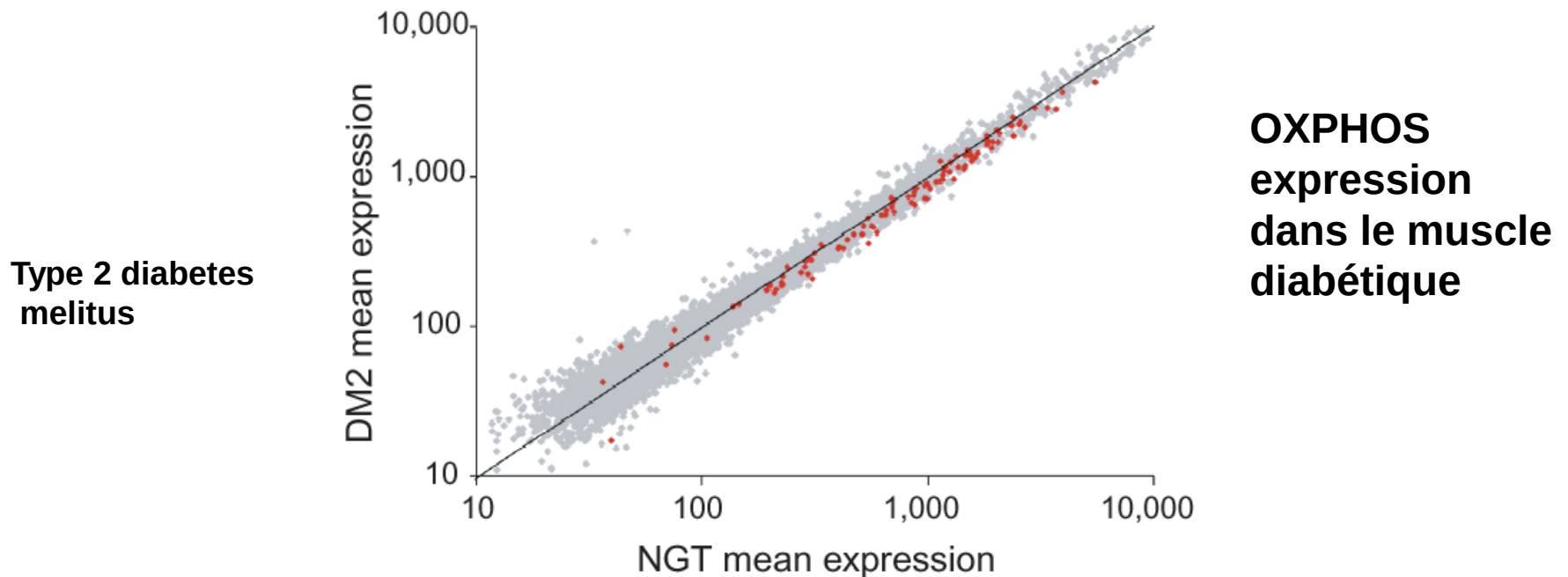
Science. 2013 March 29; 339(6127): 1546–1558.

Reasoning in terms of pathways and not in terms of genes

Des mutations peuvent affecter des composants de la même voie de signalisation

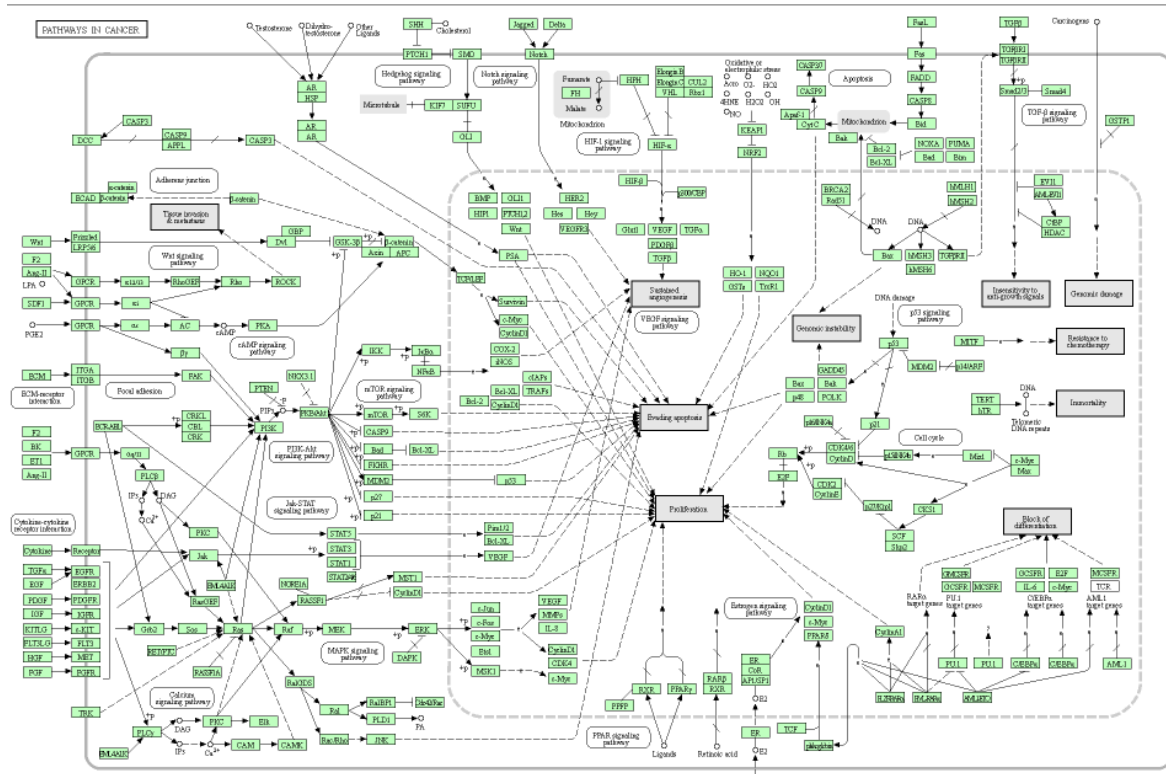
PGC-1 α -responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes

NATURE GENETICS VOLUME 34 | NUMBER 3 | JULY 2003



Application NGS en cancérologie

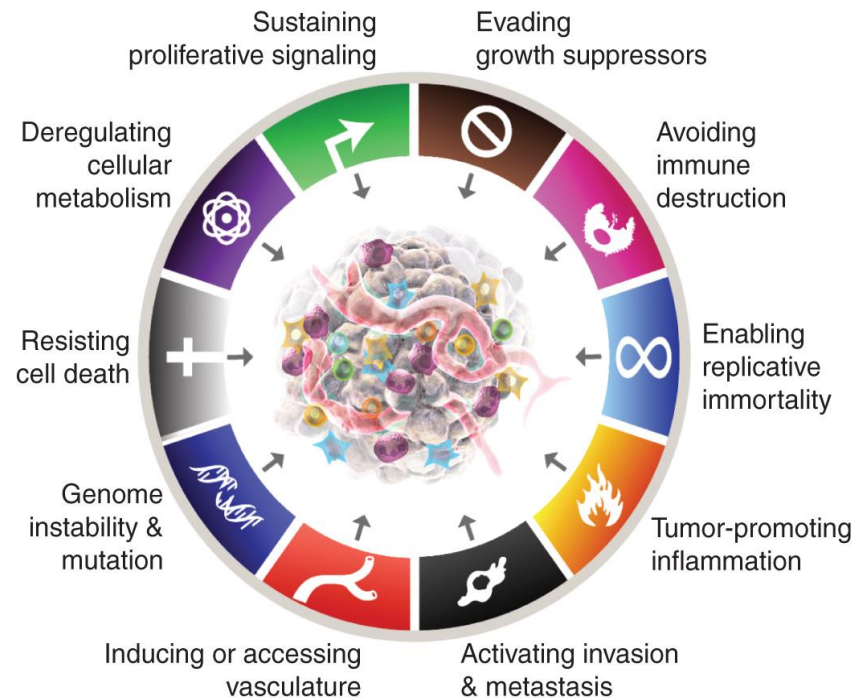
Les voies de signalisations sont liées les unes aux autres



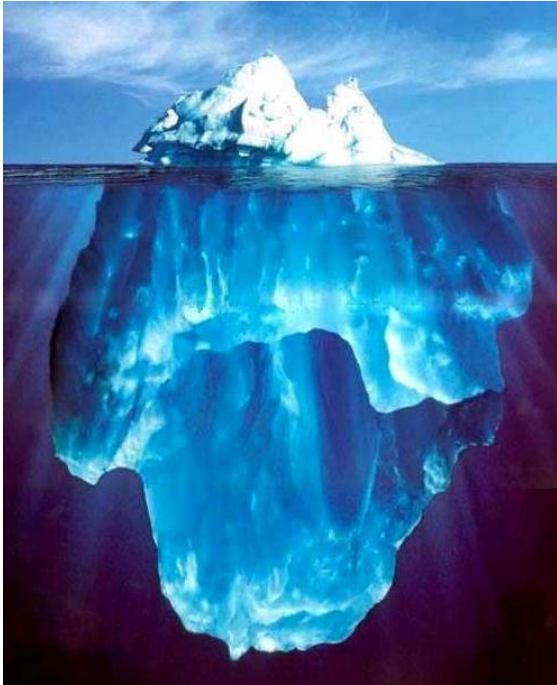
MAIS :

- Connaissance incomplète des voies de signalisation
- Régulation différente selon les tissus

Voie de signalisation et cancer



Application NGS en cancérologie



- **Identification mutations/altérations = 1^{ère} étape**
- **Etudes fonctionnelles sont essentielles pour repositionner ces modifications dans un contexte cellulaire plus large**

Cancer Genome Landscapes

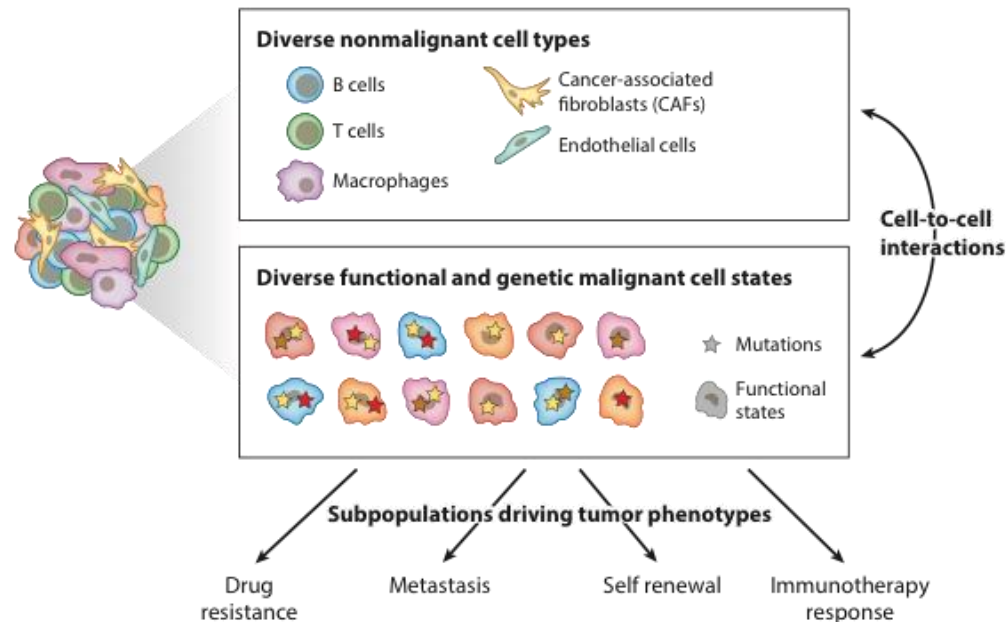
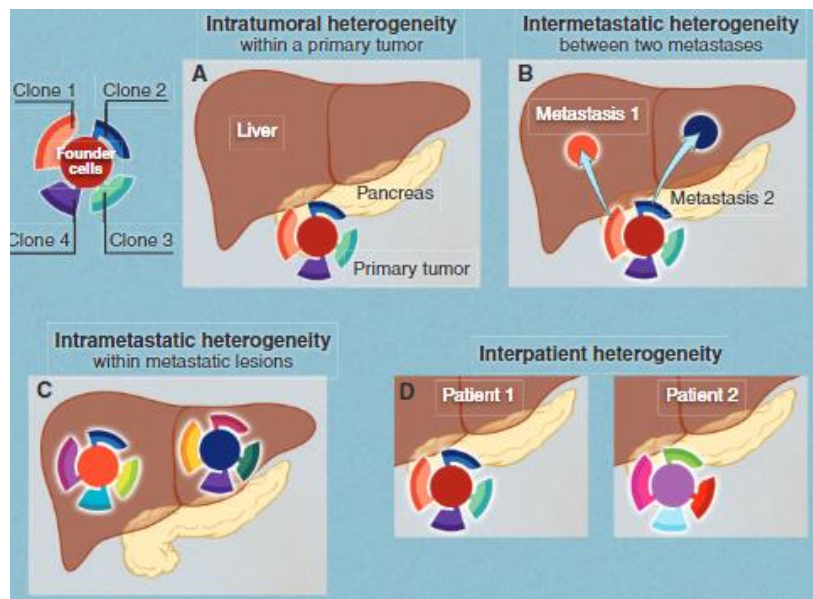
Bert Vogelstein, Nickolas Papadopoulos, Victor E. Velculescu, Shibin Zhou, Luis A. Diaz Jr., Kenneth W. Kinzler*

Science. 2013 March 29; 339(6127): 1546–1558.

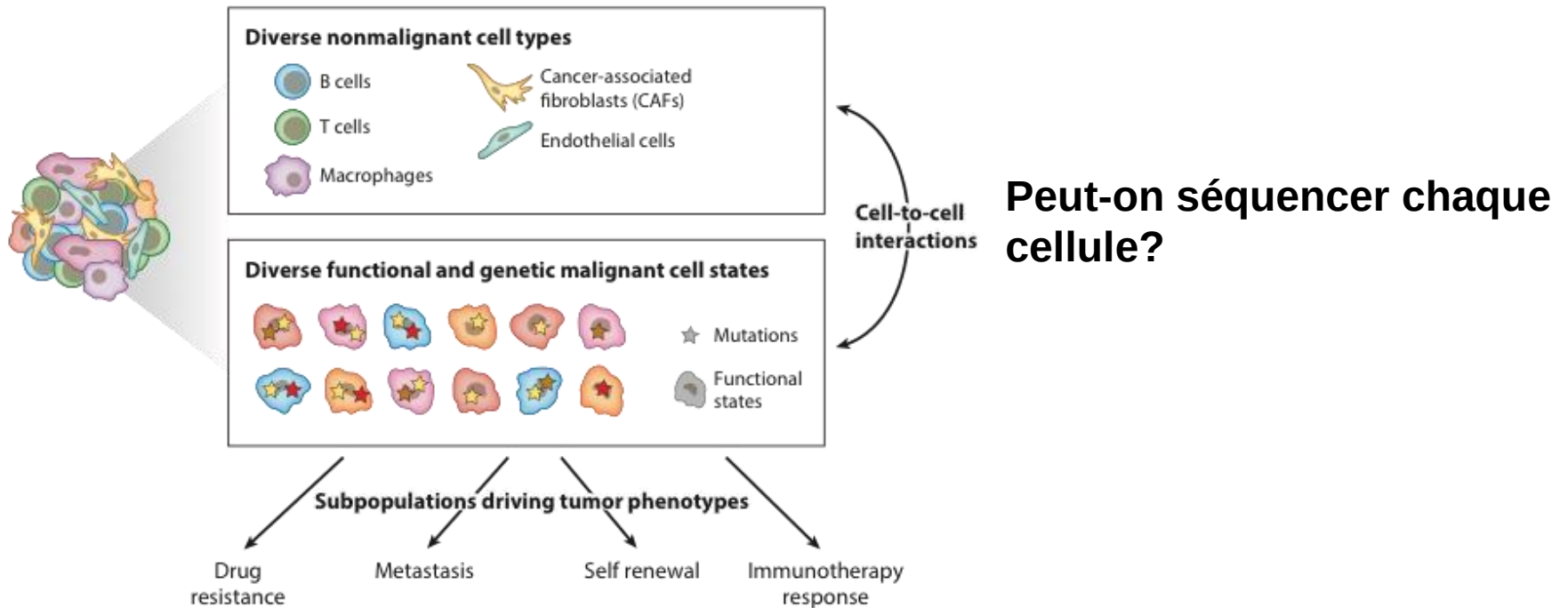
6. Genetic heterogeneity among the cells of an individual tumor always exists and can impact the response to therapeutics.

7. In the future, the most appropriate management plan for a patient with cancer will be informed by an assessment of the components of the patient's germline genome and the genome of his or her tumor.

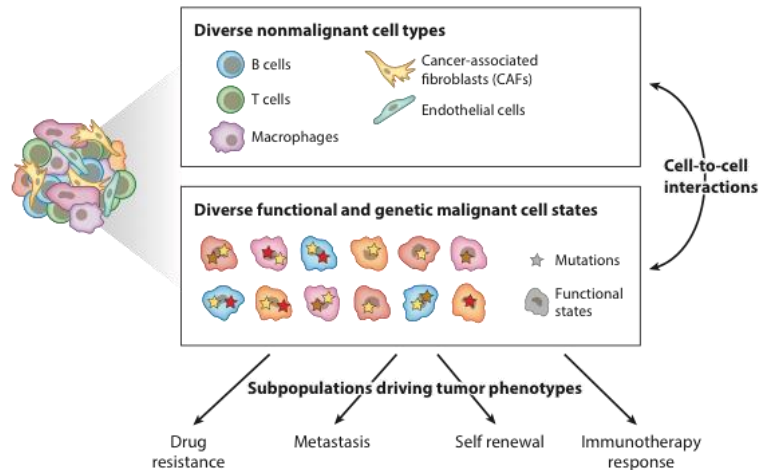
8. The information from cancer genome studies can also be exploited to improve methods for prevention and early detection of cancer, which will be essential to reduce cancer morbidity and mortality.



Problème de l'hétérogénéité (stroma and clones tumoraux)

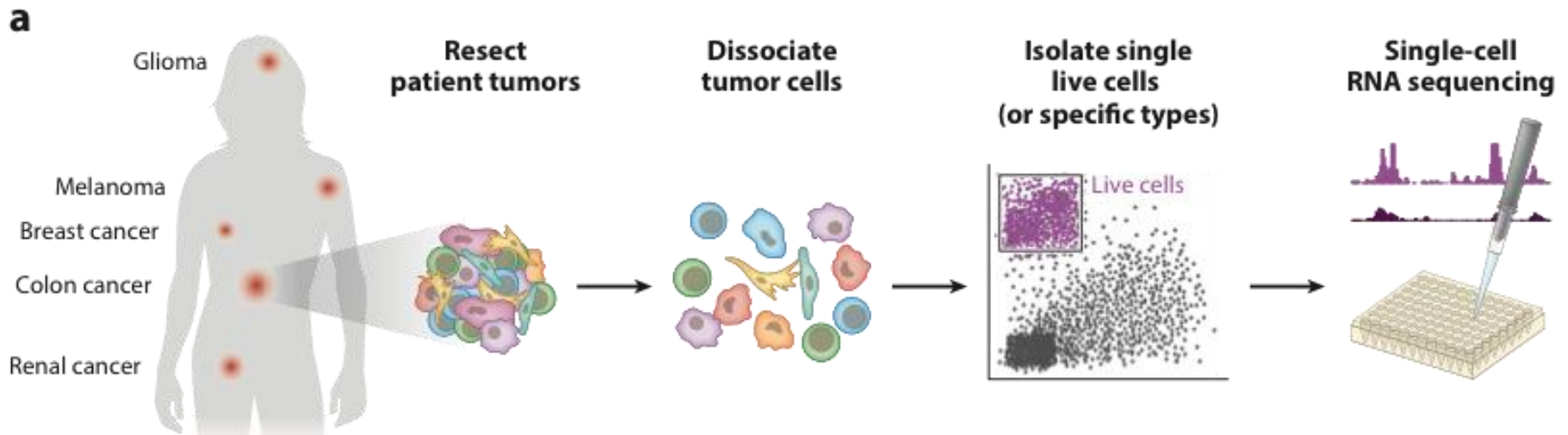


Problème de l'hétérogénéité (stroma and clones tumoraux)

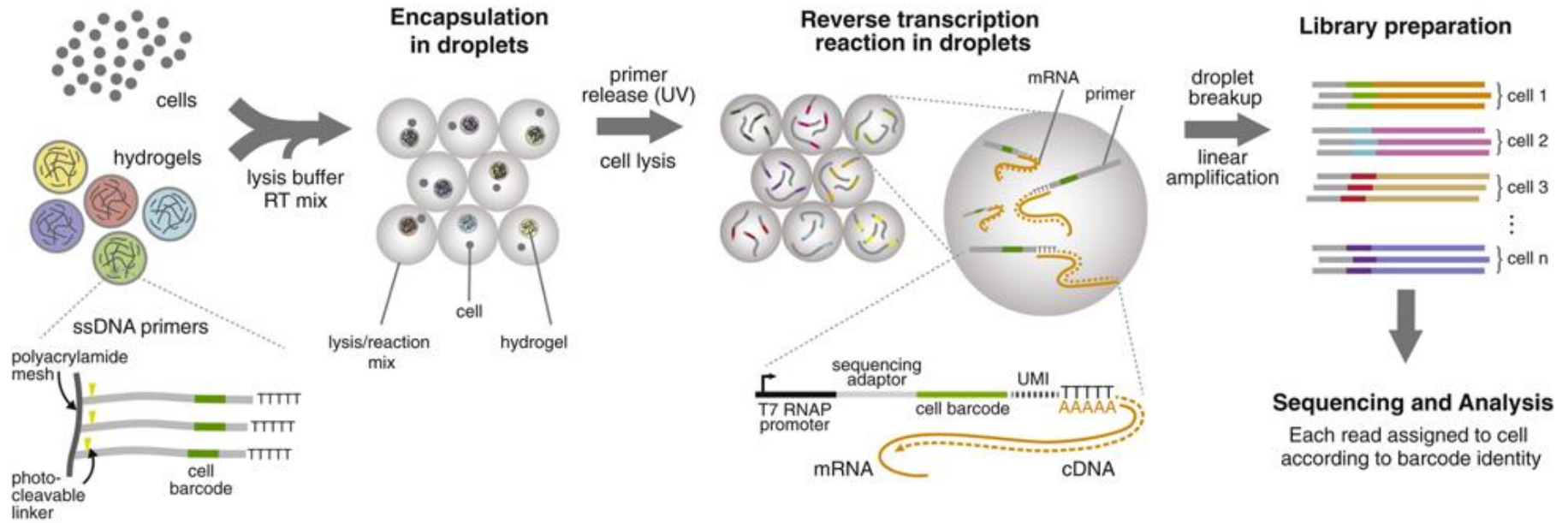


Peut-on séquencer chaque cellule?

OUI



Principe du single-cell RNA sequencing



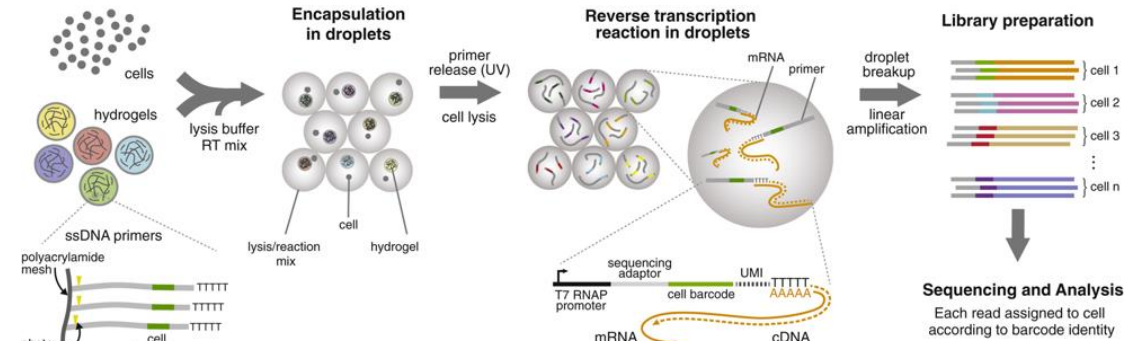
(Klein et al, 2015)

Barcode unique à chaque billes pour identifier chaque cellules

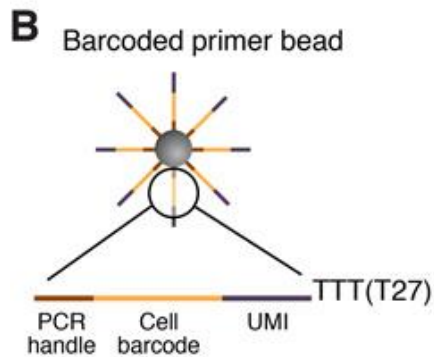
UMI (unique molecular identifier) permet l'identification de chaque ARNm initial

UMI : identifie les possibles erreurs dans la RT et l'amplification du cDNA

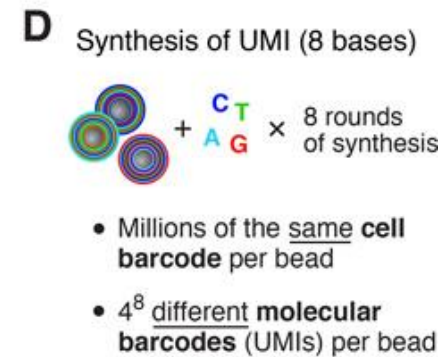
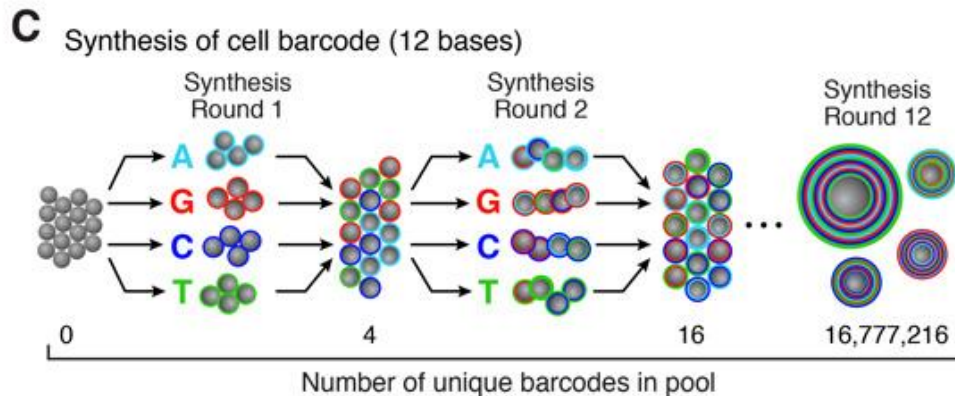
Principe du single-cell RNA sequencing



(Klein et al, 2015)



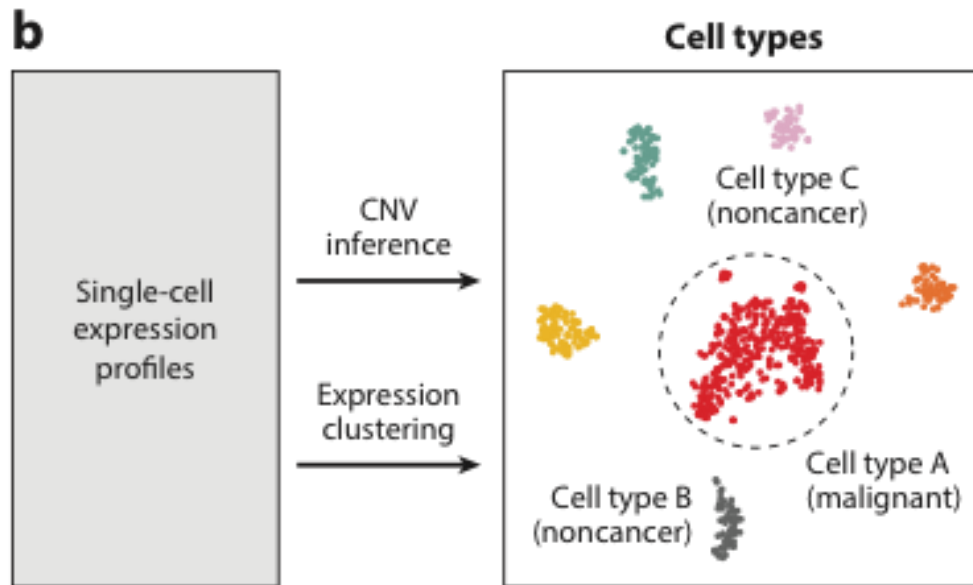
(Macosko et al, 2015)



Single-cell RNA sequencing en biologie du cancer

Au final, on obtient pour chaque cellule, le niveau d'expression de plusieurs milliers de gènes. Cela pour plusieurs milliers de cellules

-> comment représenter cela?

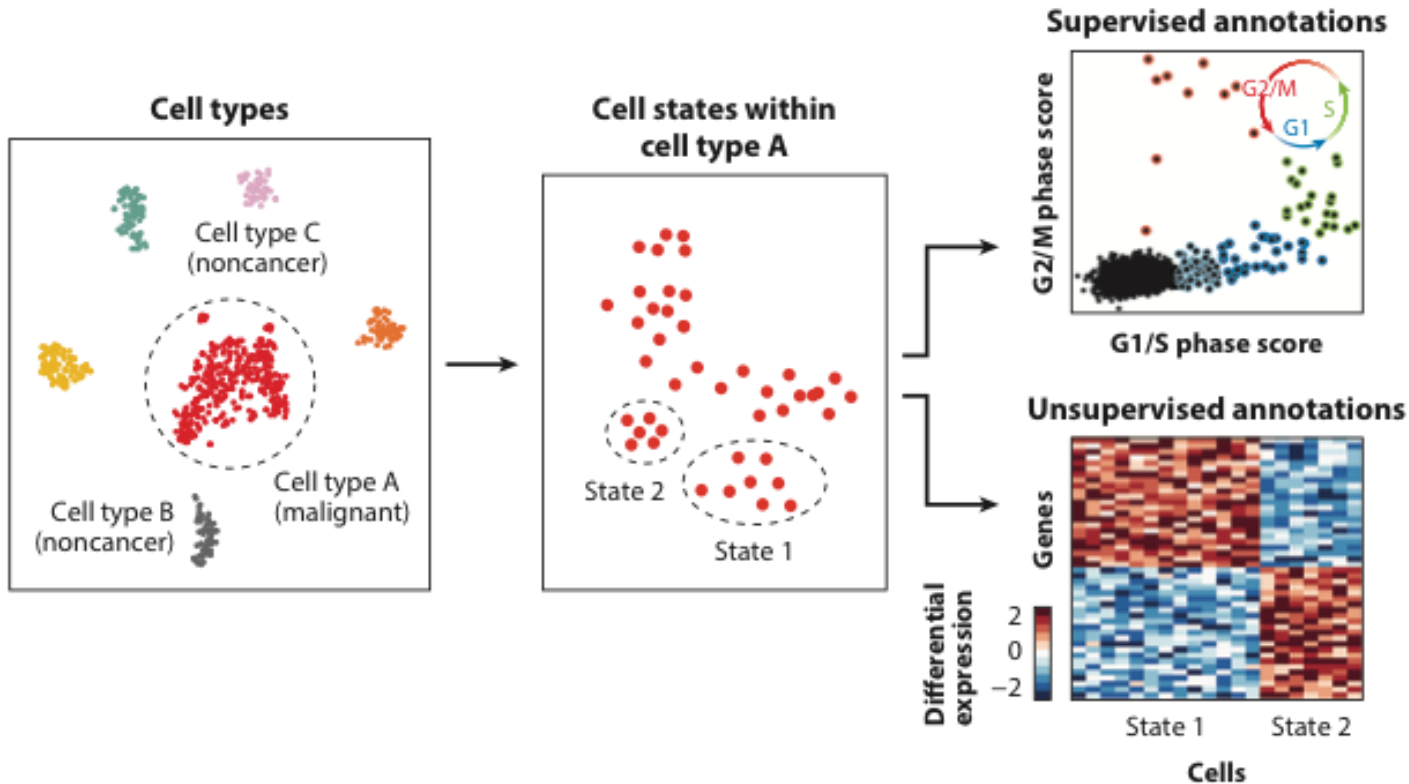


tSNE ou UMAP plot : réduction de dimensions

-> Chaque point représente une cellule

-> Groupe les cellules en fonction de similarité au niveau des ARNm détectés et de leur niveau d'expression

Single-cell RNA sequencing en biologie du cancer



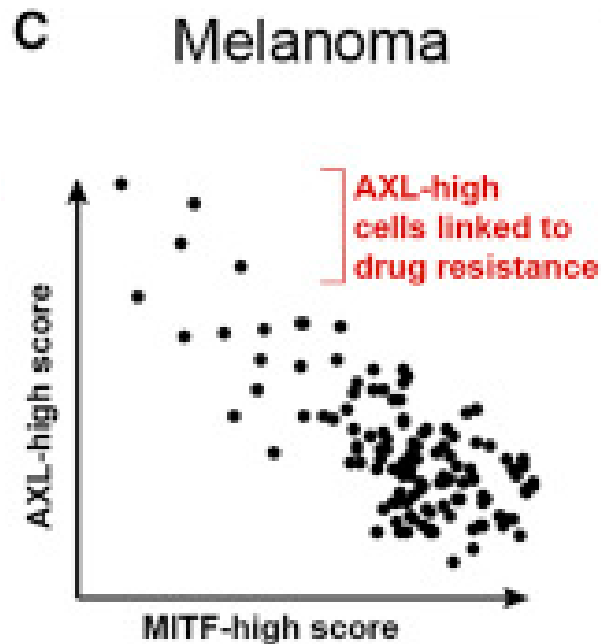
Identification de différents types cellulaires -> Contribution des cellules non cancéreuses (micro-environnement: Système immunitaire, Fibroblastes associés au cancer, ...)

Identification de différents états cellulaires dans la tumeur en fonction des gènes exprimés

Single-cell RNA sequencing en biologie du cancer

Identification de différents états cellulaires dans la tumeur en fonction des gènes exprimés

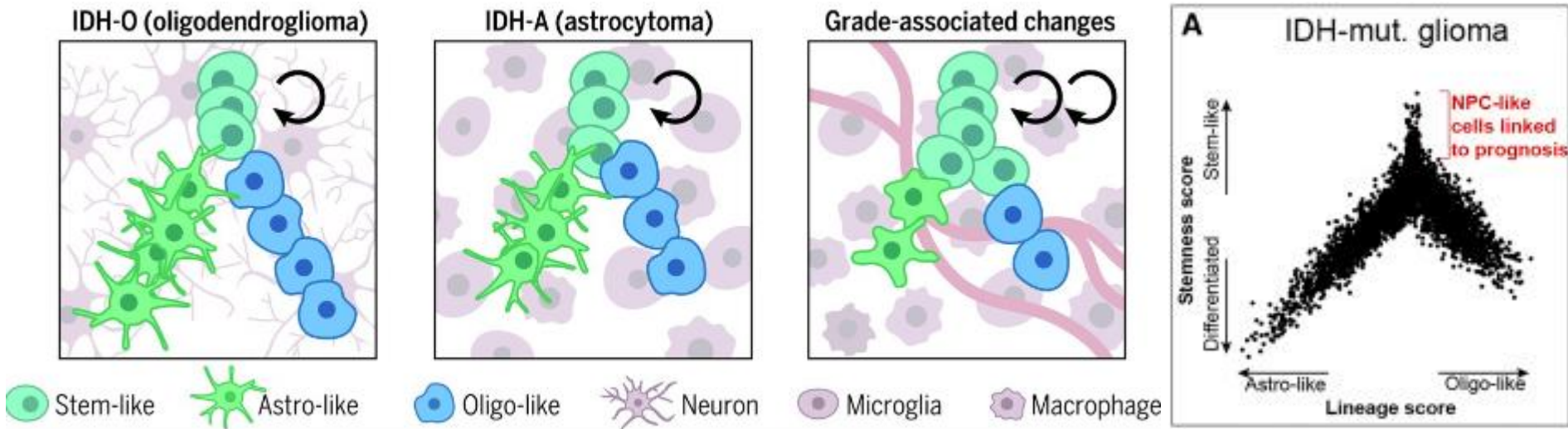
-> Présence d'un pool de cellules résistantes dans le mélanome



Single-cell RNA sequencing en biologie du cancer

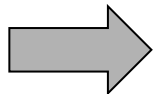
Identification de différents états cellulaires dans la tumeur en fonction des gènes exprimés

- Différents types cellulaires dans la tumeur et le micro-environnement des gliomes IDH-mutés
- Présence de cellules souches qui évolue avec la progression de la maladie



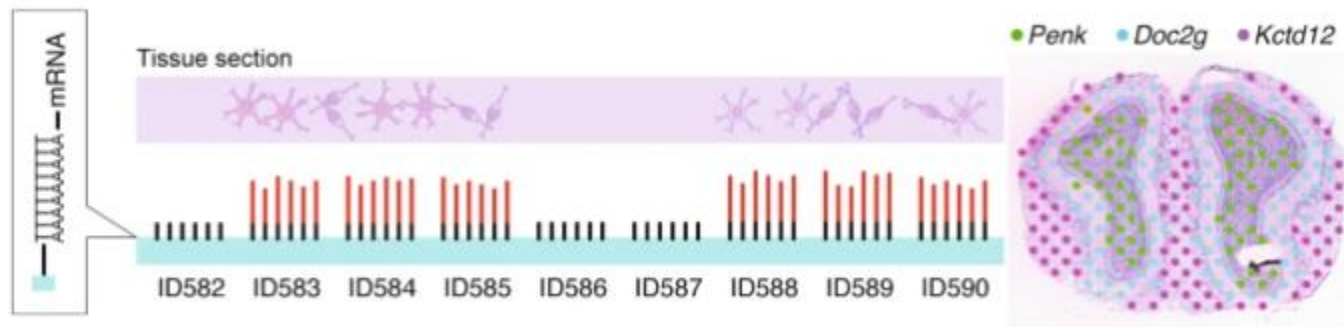
Limites du single-cell RNA sequencing

- Seul une part du transcriptome est séquencé (2000—8000 genes)
- Seule une portion de chaque mRNA est habituellement séquencé
- Presque impossible de détecté les mutations
- Séquençage des mRNA seulement
- Très dépendant de la qualité des échantillons
- Toujours très cher



**Plus utile en combinaison avec des analyse
génomiques et transcriptomique en « bulk »**

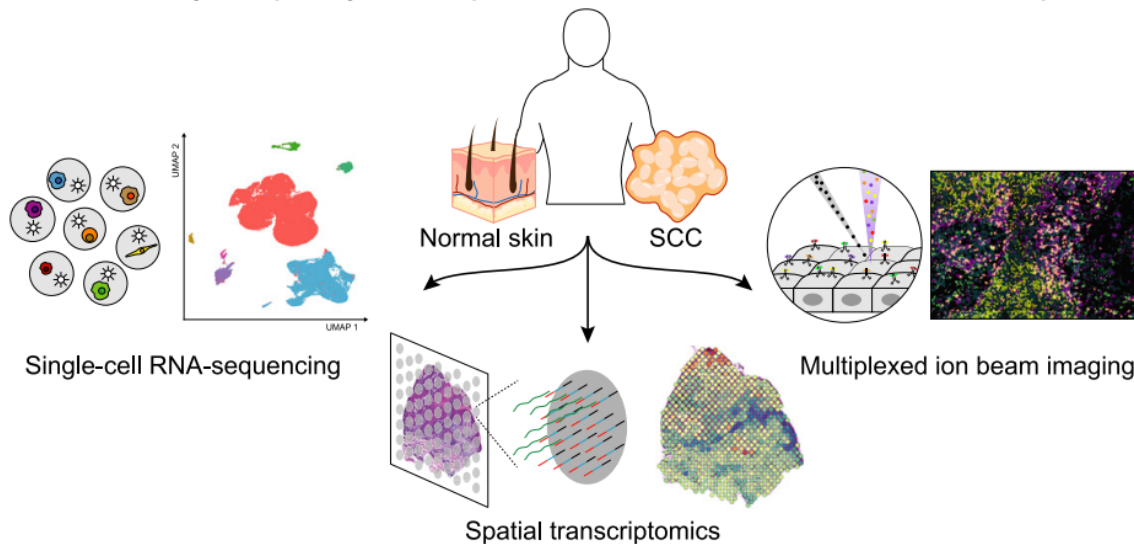
Vers une nouvelle étape: Spatial transcriptomic



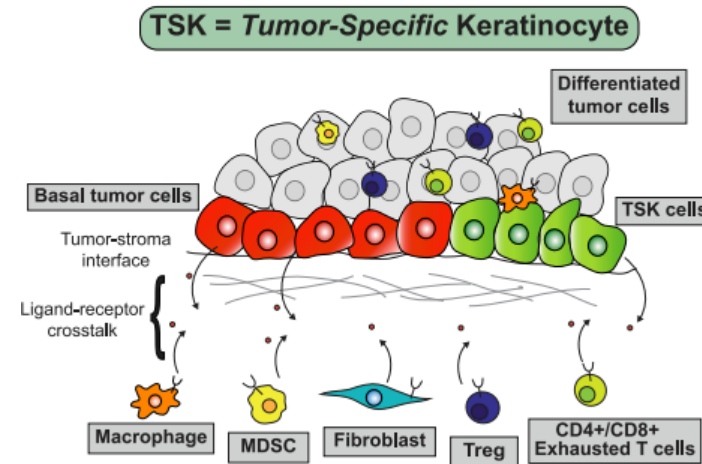
- Coupe de tissu imagée
 - ARN sont extrait de la coupe et déposé sur une grille possédant des oligos avec des identifiants de la localisation
 - ARN amplifié puis séquencé -> Grâce à l'identifiant présent dans la séquence on peut les repositionner sur la coupe
-

Vers une nouvelle étape: Spatial transcriptomic

Multimodal analysis of primary human squamous cell carcinoma and site-matched normal epidermis



Identification of cell types, localization, and signaling in tumor microenvironment



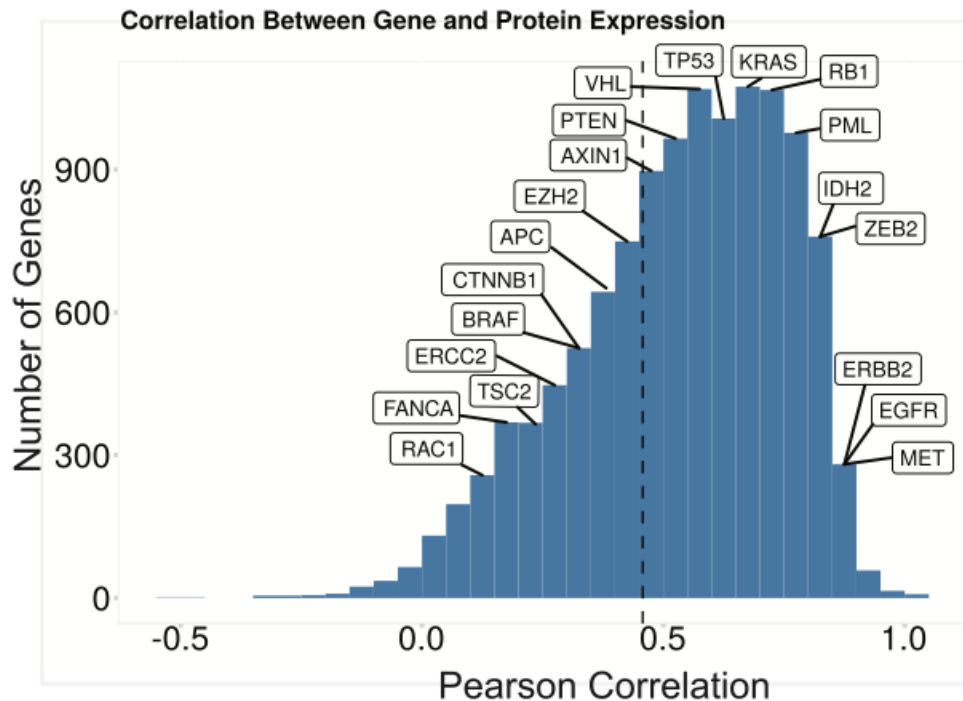
Mise en évidence d'une population de kératinocyte tumeur-spécifique dans des carcinomes squameux. Cette population située à la frontière de la tumeur exprime des gènes associés à la résistance à l'immunothérapie et des ligands

Au-delà de la transcriptomique : Protéomique

Cell

Quantitative Proteomics of the Cancer Cell Line Encyclopedia

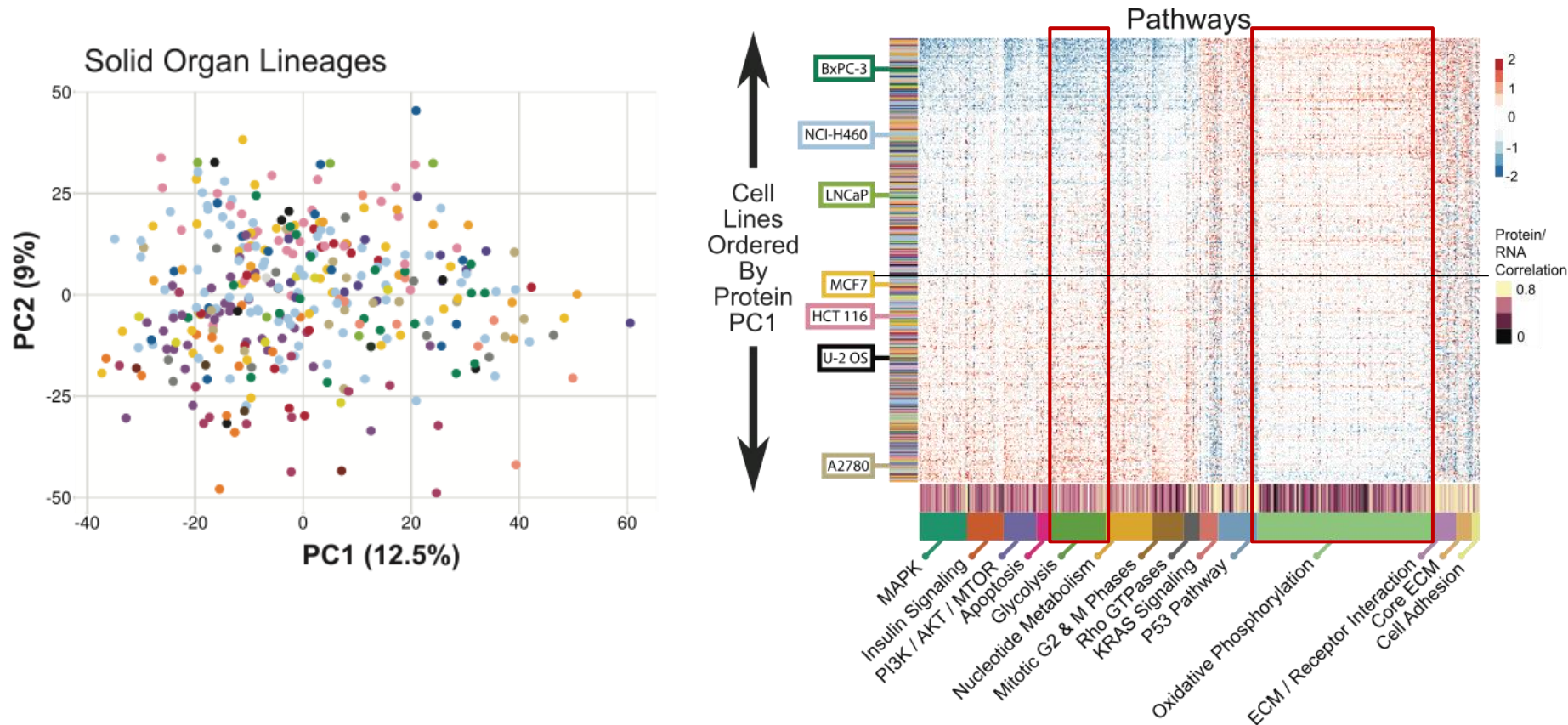
Pour 375 lignées cellulaires issues de cancer : Expression des mRNA évaluée par RNAseq et niveau des expressions des protéines mesuré par spectrométrie de masse



**Niveau ARNm et protéine
pas forcément bien corrélé**

**-> Mesure de la quantité
d'ARNm n'est pas
forcément indicatif de la
présence de la protéine**

Au-delà de la transcriptomique : Protéomique



Signature protéique ne correspond pas au type de cancer

Signature protéique permet en revanche de mettre en évidence des voies de signalisation up ou down-regulées (pas forcément de corrélation avec mRNA)

Toujours plus de -omics

- **Methylome**
 - **ATAC-seq : étude région accessible de l'ADN**
 - **ChIP-seq : modification des histones et liaisons de FT**
 - **Ribo-seq : étude des ARNm liés aux ribosomes**
 - **Lipidome**
 - **Métabolome**
-

Conclusion

- **Les données de omics permettent d'obtenir une vision globale de la tumeur**
- **Il faut également replacer la tumeur dans son environnement**
- **La technologie est maintenant accessible à toute équipe de recherche**
- **Intégration de données multi-omics pour une meilleure vision**
- **Quantité de données générées est gigantesque**
- **Nécessité de validation fonctionnelle**

