

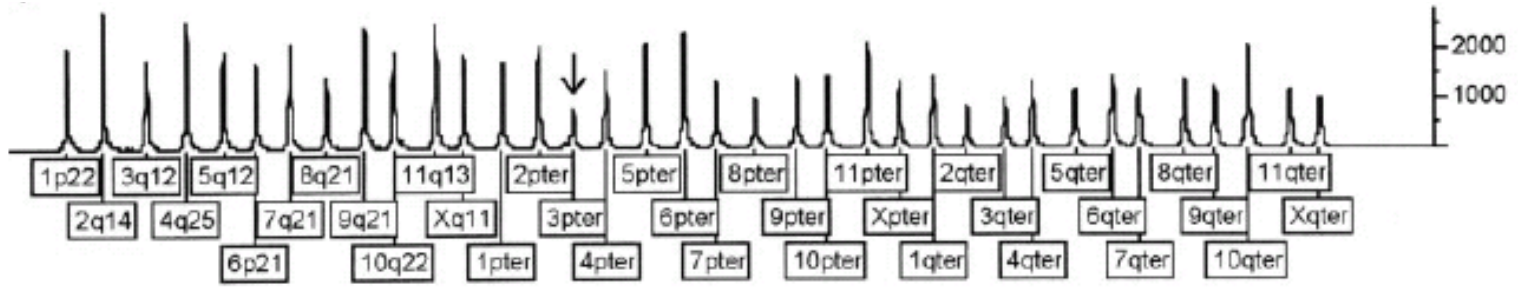
# Whole Genome Sequencing: Application à l'étude des chromosomes

Dr Caroline SCHLUTH-BOLARD

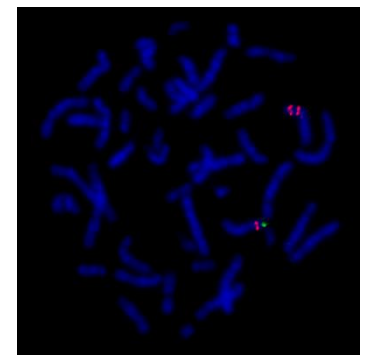
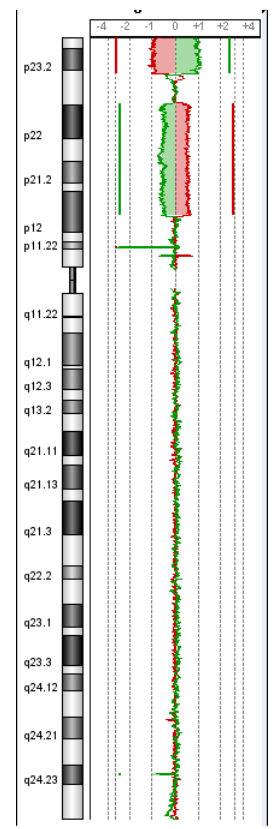
Laboratoire de Cytogénétique Constitutionnelle, Hospices Civils de Lyon  
Centre de Recherche en Neurosciences de Lyon

[caroline.schluth-bolard@chu-lyon.fr](mailto:caroline.schluth-bolard@chu-lyon.fr)





# Whole Genome Sequencing ?

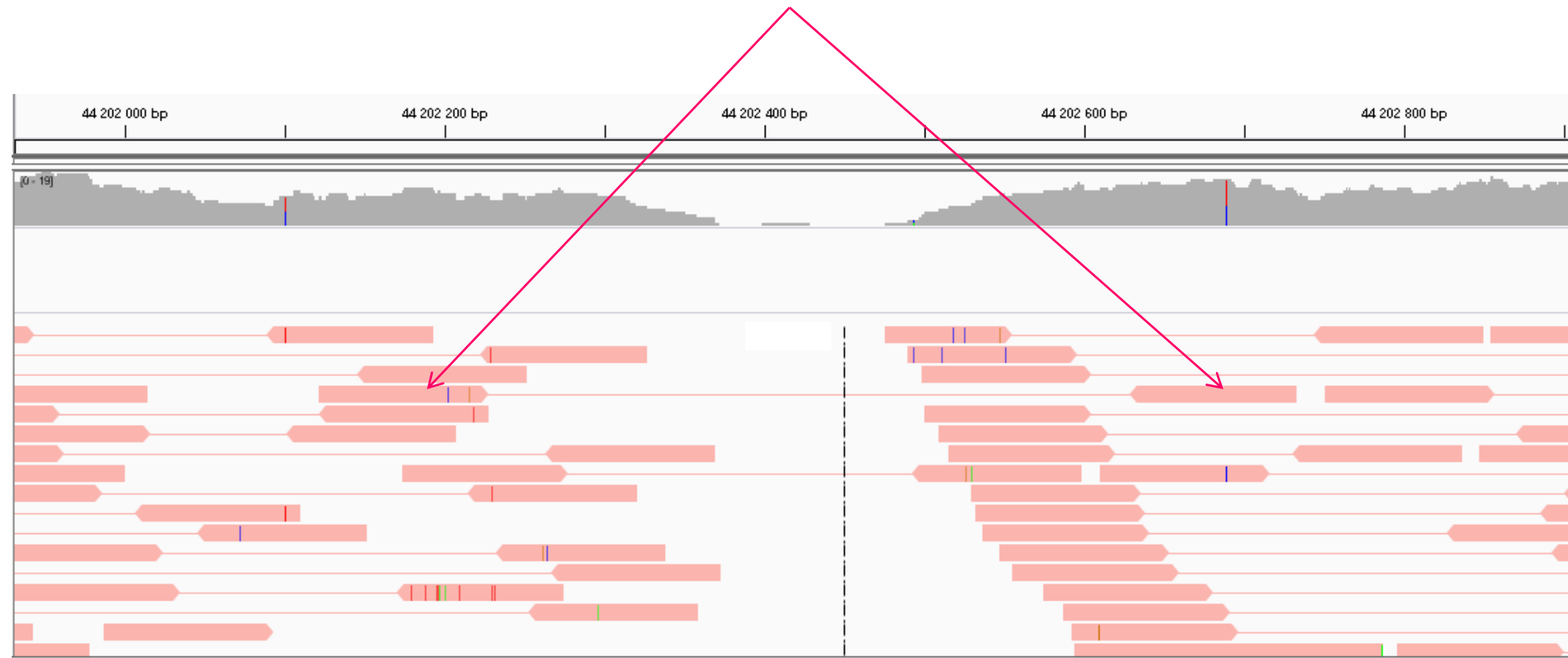


# Principe du WGS

## *Short reads*

- Préparation d'une librairie à partir de l'ADN génomique
- Séquençage massif en parallèle (reads) : 50-400 pb
- Alignement sur le génome de référence

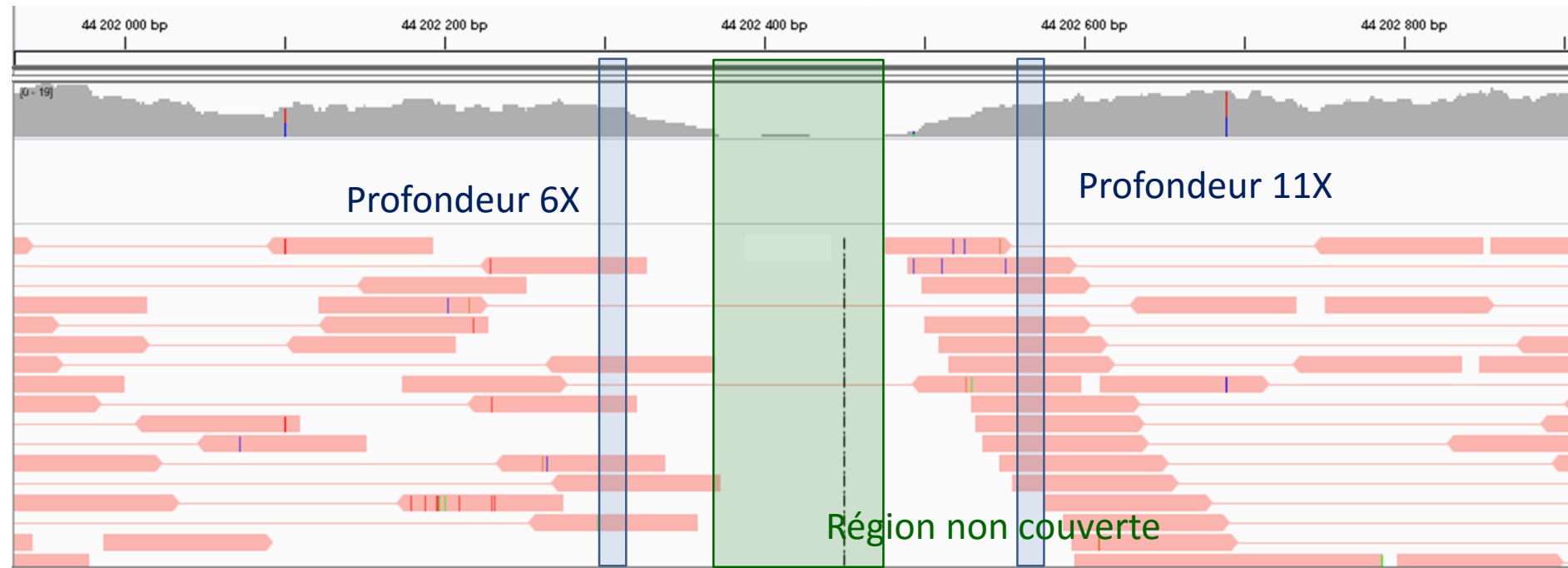
### Paire de reads



# Principe du WGS

## *Short reads*

- Préparation d'une librairie à partir de l'ADN génomique
- Séquençage massif en parallèle (reads) : 50-400 pb
- Alignement sur le génome de référence
- Profondeur : nb moyen de reads couvrant une région
- Couverture : taux de bases couvertes par des reads



# Principe du WGS

- Séquençage type paired-end

## Préparation de la librairie



ADN génomique



Fragments ≈ 400 pb  
(sonication)



Ligation des  
adaptateurs

## Séquençage



Amplification

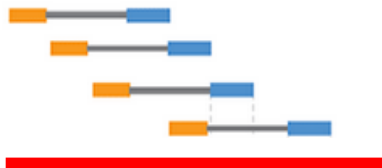


Séquençage read1



Séquençage read2

## Bioinformatique



hg19

- Alignement sur génome de référence
- Recherche de variants
- Interprétation

# Principe du WGS

- Séquençage type mate-pair

## Préparation de la librairie



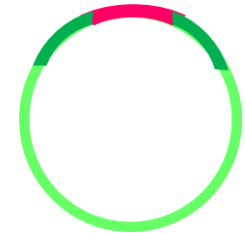
ADN génomique



Fragments  $\approx$  3000-10000 pb  
(sonication)

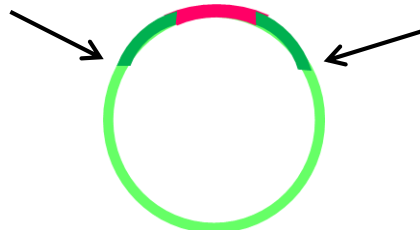


Ajout d'un adaptateur



Ligation et  
circularisation

## Préparation de la librairie



clivage



fragment linéaire



Ligation des  
adaptateurs

## Séquençage et bioinformatique

# Principe du WGS

## Paired-end

Préparation librairie simple

Bio-info plus simple

Résultats plus précis

Coût plus élevé

Séquences répétées et  
duplicons



## Mate-pair

Coût moins élevé

Meilleure couverture

Séquences répétées et gap

Préparation librairie délicate

Résultats moins précis

Séquences répétées et  
duplicons



# Principe du WGS

## Séquençage 3<sup>ème</sup> génération

### single molecule real-time long reads

- Reads de 10 à 200 kb  
Ex: PacBio, Nanopore (Minlon)
- Avantages :
  - Assemblage *de novo*
  - Structures complexes
- Inconvénients
  - Taux d'erreurs élevé
  - Coût
  - Débit

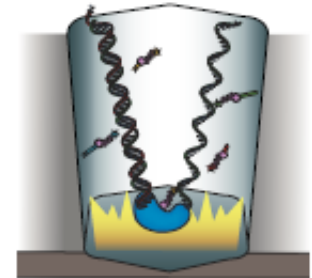
SMRTbell template  
Two hairpin adapters  
allow continuous  
circular sequencing



ZMW wells  
Sites where  
sequencing  
takes place

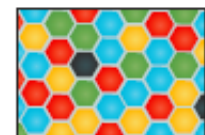


Labelled nucleotides  
All four dNTPs are  
labelled and available  
for incorporation



Modified polymerase  
As a nucleotide is  
incorporated by the  
polymerase, a camera  
records the emitted light

PacBio output  
A camera records the changing  
colours from all ZMWs; each  
colour change corresponds to  
one base



Ex : technologie PacBio



# Détection des variants structuraux

## *Détections des remaniements déséquilibrés (CNVs)*

- Analyse des paires de reads :

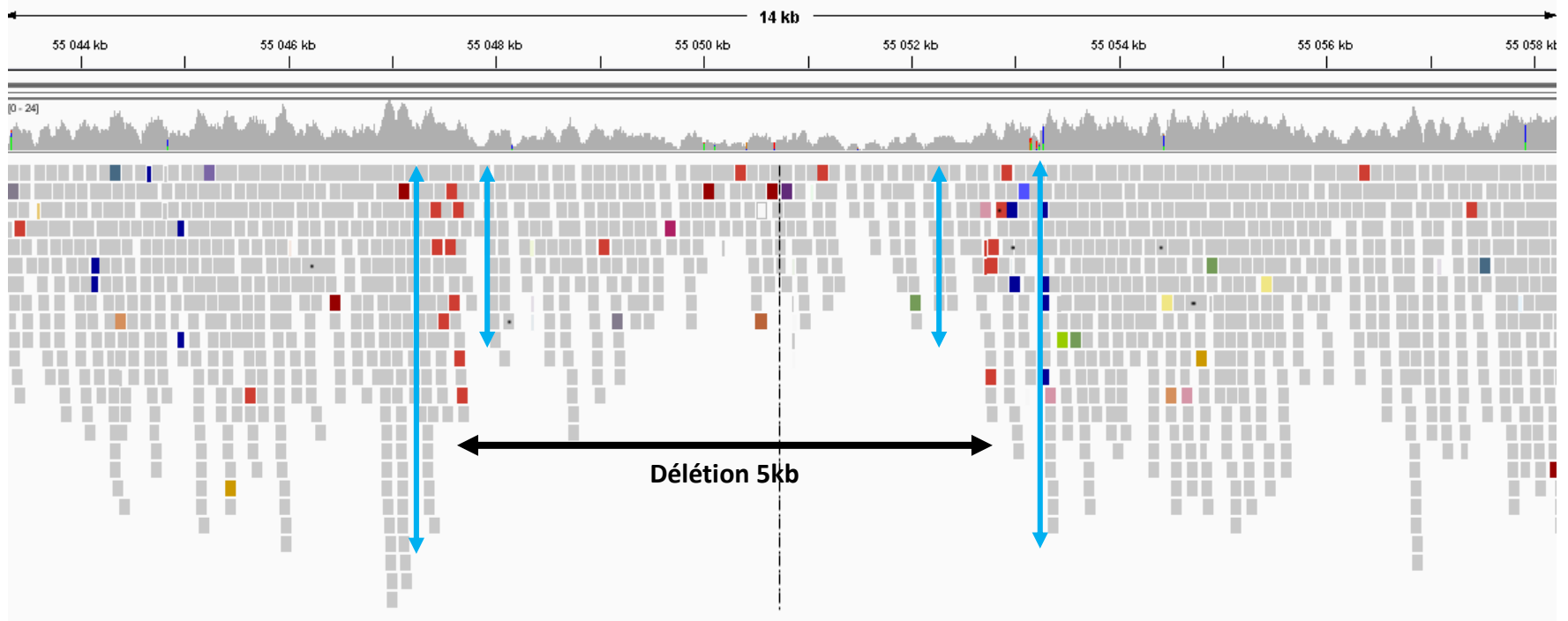
Distance entre 2 reads par rapport à la référence



# Détection des variants structuraux

## *Détection des remaniements déséquilibrés (CNVs)*

- Analyse de la profondeur (*read depth*) :  
densité de reads proportionnelle au nombre de copies

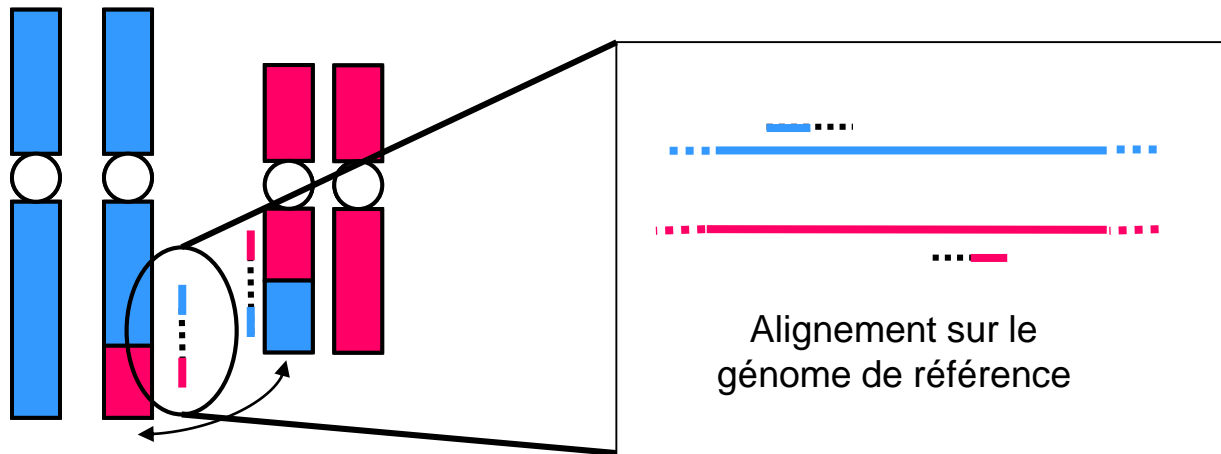


# Détection des variants structuraux

## *Détection des remaniements équilibrés*

- Analyse des paires de reads

Position, distance, orientation des reads : alignement aberrant

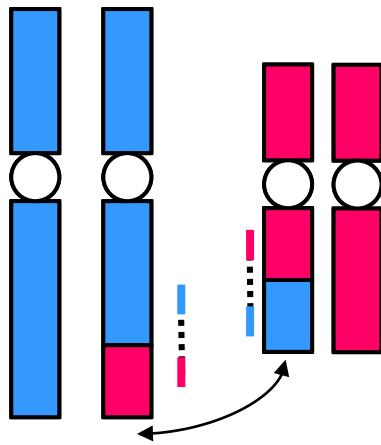


**Ex : Translocation**

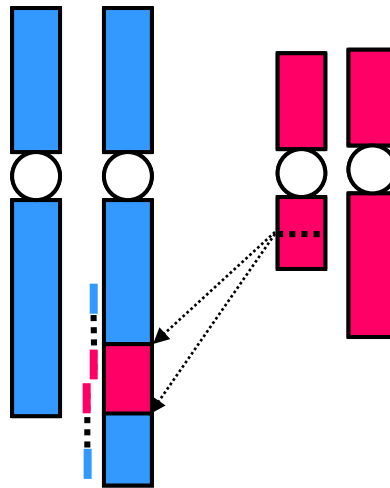
# Détection des variants structuraux

## *Détections des remaniements équilibrés*

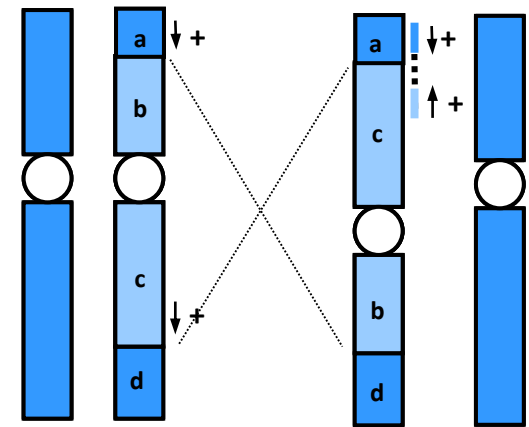
- Analyse des paires de reads



1) Translocation



2) Insertion

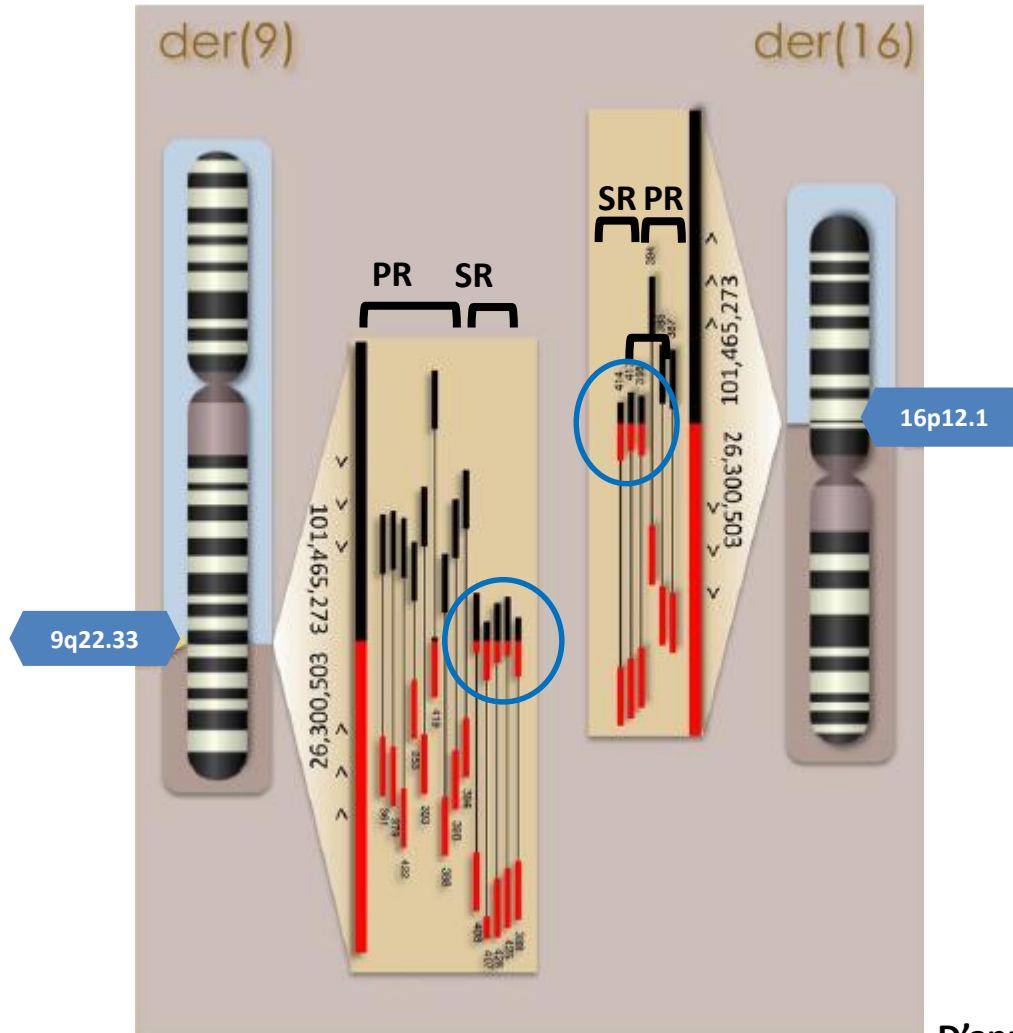


3) Inversion

# Détection des variants structuraux

## *Détections des remaniements équilibrés*

- Analyse des reads interrompus (split-reads)



# Les difficultés

*Une aiguille dans une botte de foin...*

- Pour une profondeur de 10-15X en paired-end :  
environ 4000 à 5000 remaniements détectés
- Nécessité de filtrer pour trouver les remaniements pertinents
- Base de données :
  - éliminer les évènements récurrents et les artefacts100 à 400 remaniements
- Autres filtres : nb de reads

# Les difficultés

## *Séquences répétées*

- Duplicons, télomères, centromères
- Impossibilité d'attribuer un read à une position unique
- Inconvénient majeur des techniques en reads courts

# Les difficultés

## *Séquences répétées*

Inversion inv(7)(q21.3q35) : duplicon en 7q35





# Les difficultés

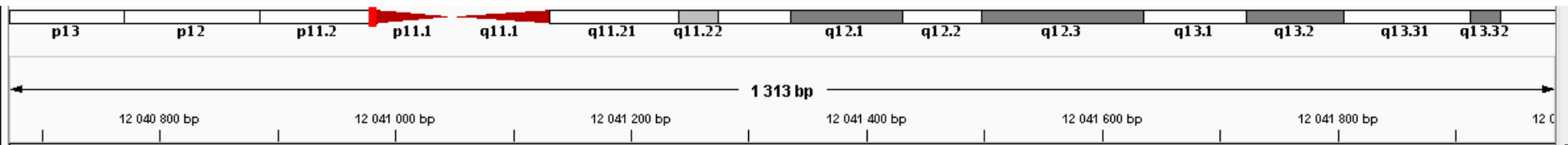
*Choix du génome de référence : hg19 ou hg38*

- La version hg38 intègre de nouvelles séquences, notamment au niveau des lacunes

ex t(3;22)(q13.3;p11.2)

chr22

Génome hg19

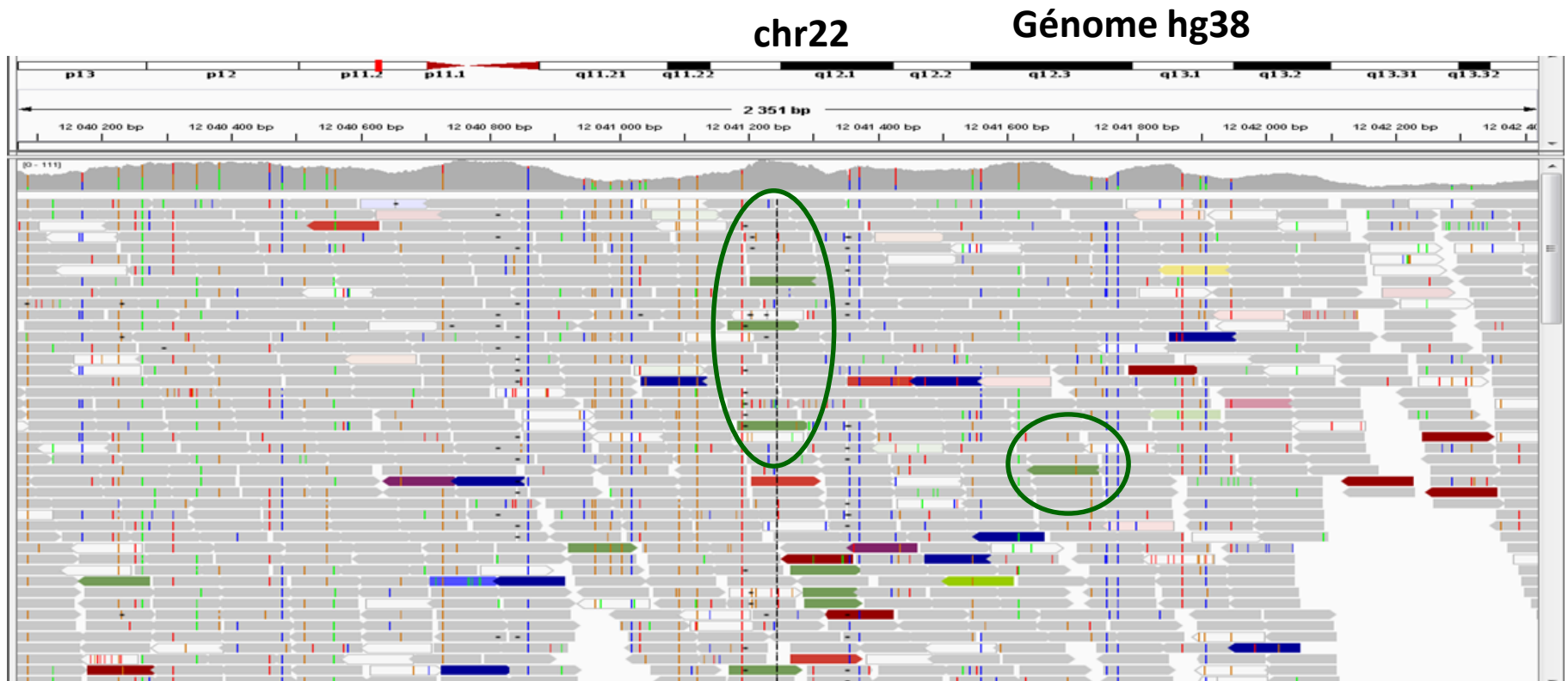


[0 - 10.00]

# Les difficultés

*Choix du génome de référence : hg19 ou hg38*

- $\text{ex t}(3;22)(\text{q13.3};\text{p11.2})$

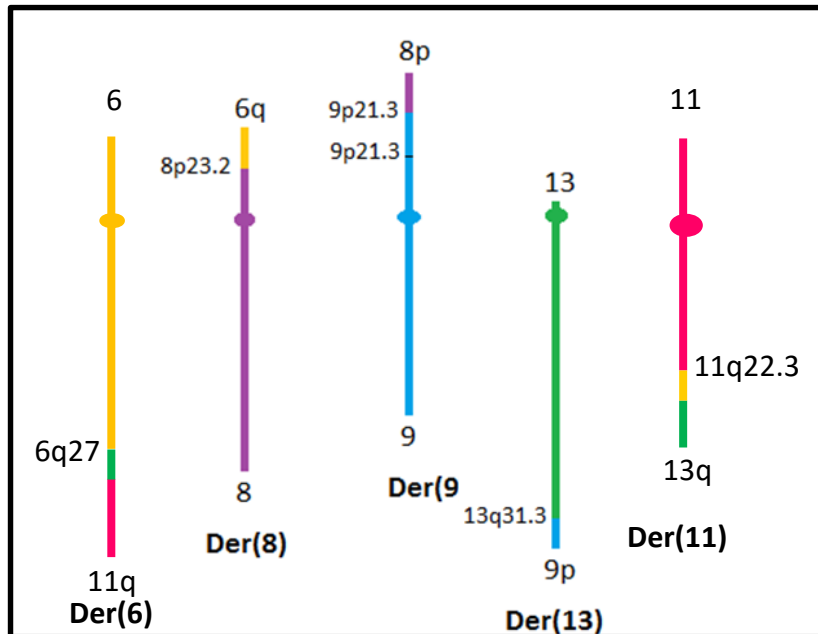


# Les difficultés

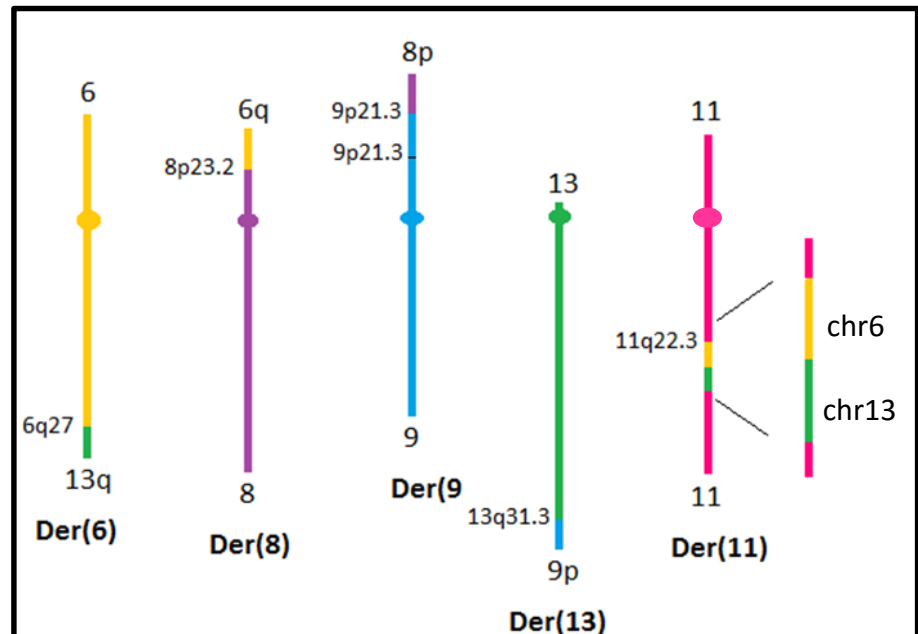
## *Interprétation*

- Breakdancer : remaniement complexe à 8 points de cassure impliquant les chr 6,8,9,13 et 11.

### Interprétation 1

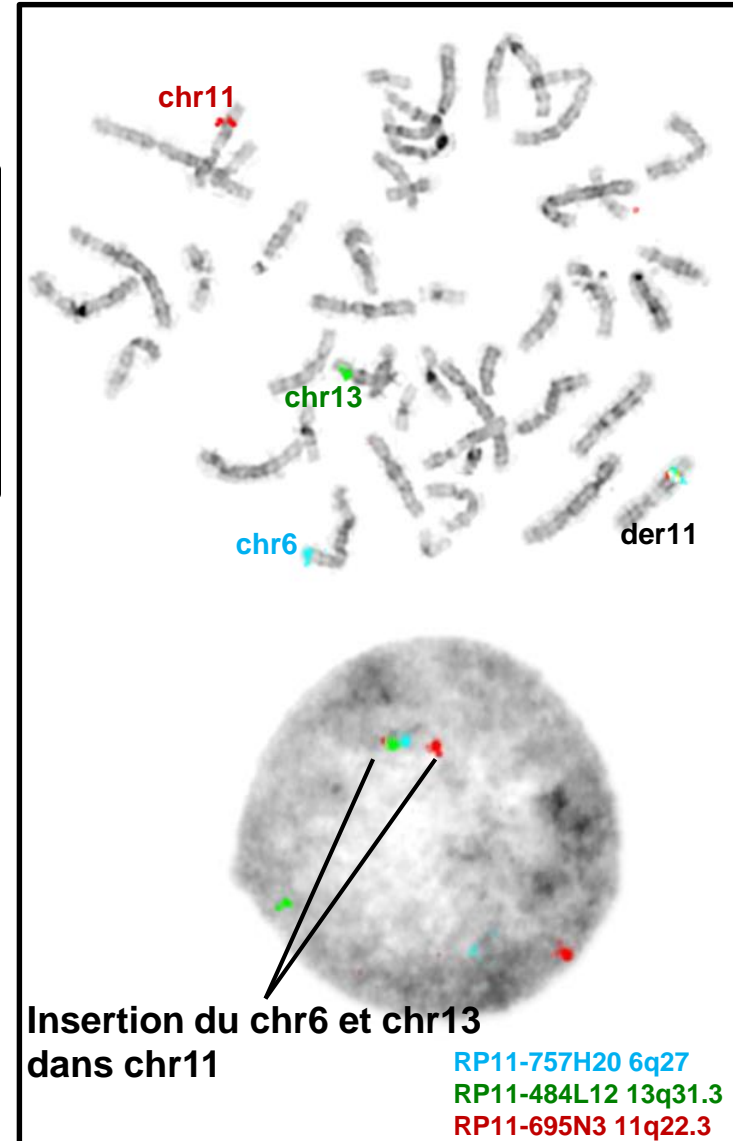
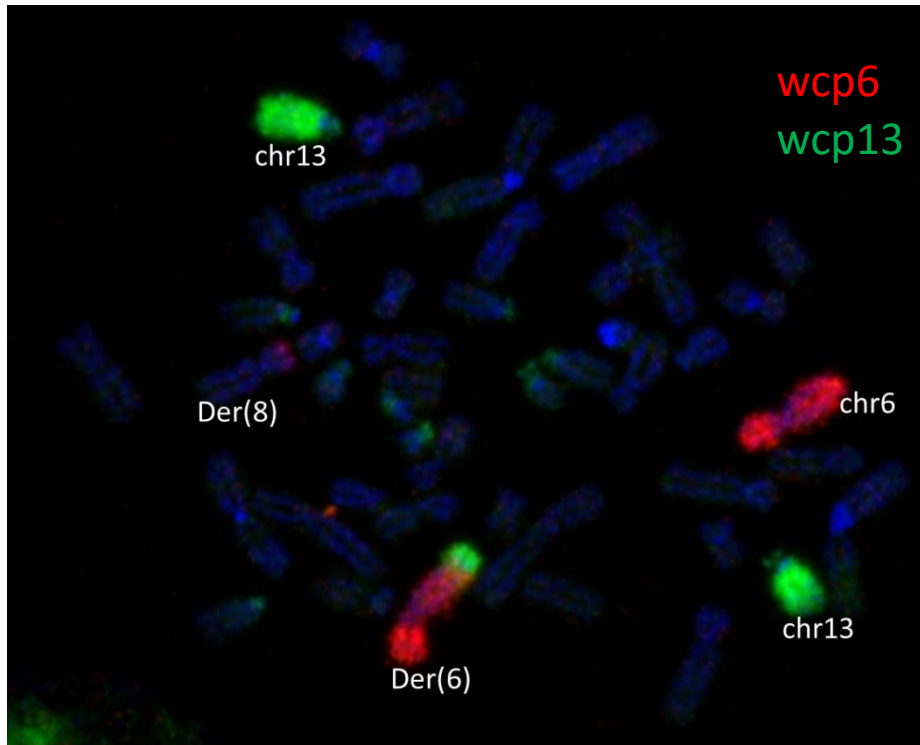
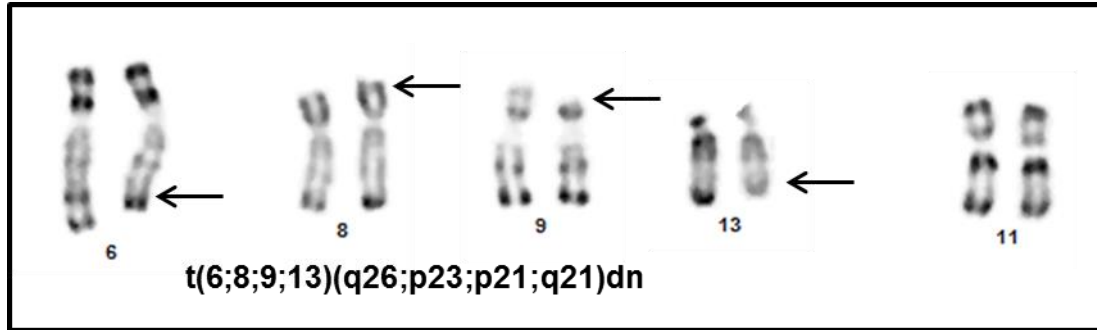


### Interprétation 2



# Les difficultés

## Interprétation



## Interprétation 2

# Les résultats

## *Cartographie des variants structuraux constitutionnels*

- Korb et al., Science, 2007

	NA15510	NA18505	
SVs	472	825	
indels	422	753	<b>90%</b>
inv	50	72	<b>10%</b>



- SV de petite taille  
65% < 10kb  
35% < 5kb
- Points chauds

# Les résultats

## *Cartographie des variants structuraux constitutionnels*

- 1000 genome project

Caractérisation des différents types de variants génotypiques par NGS, dans différentes populations issues de la collection HapMap

185 génomes étudiés pour SVs

	Nombre	% nouveaux
Délétion	13 826	58%
Duplication	501	89%
Insertion élément mobile	5370	87%
Insertion nouvelle séquence	174	93%

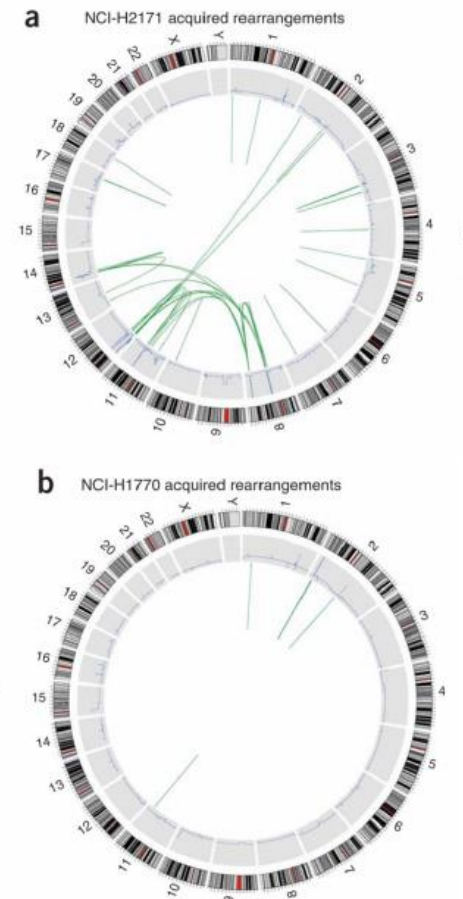
# Les résultats

## *Cartographie des variants structuraux somatiques*

- Campbell et al., Nat Genet, 2008

Caractérisation de deux lignées de cancer du poumon :  
variants constitutionnels et somatiques

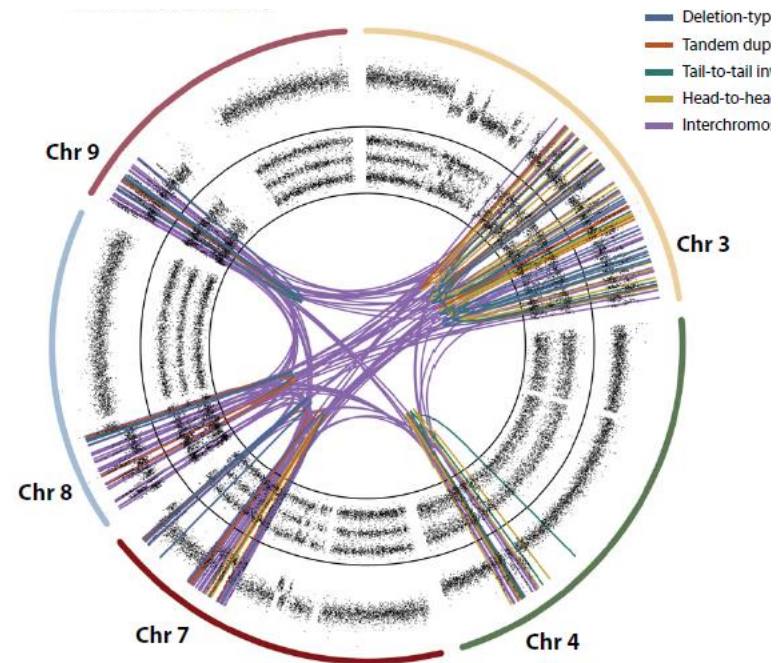
	NCI-H2171	NCI-H1770
SVs (acquis)	81	22
Del	2	0
Dup	11	2
Inv	2	1
Amplicon	44	19
Interchr	22	0



# Les résultats

## *Découverte de nouveaux remaniements et de leurs mécanismes: chromothripsis et chromoanasythesis*

- Remaniements très complexes
- Evènement catastrophique unique
- Pulvérisation d'un chromosome suivi de sa reconstitution
- Cancers : tumeurs cérébrales et osseuses
- DI/Malformations
- Individus normaux

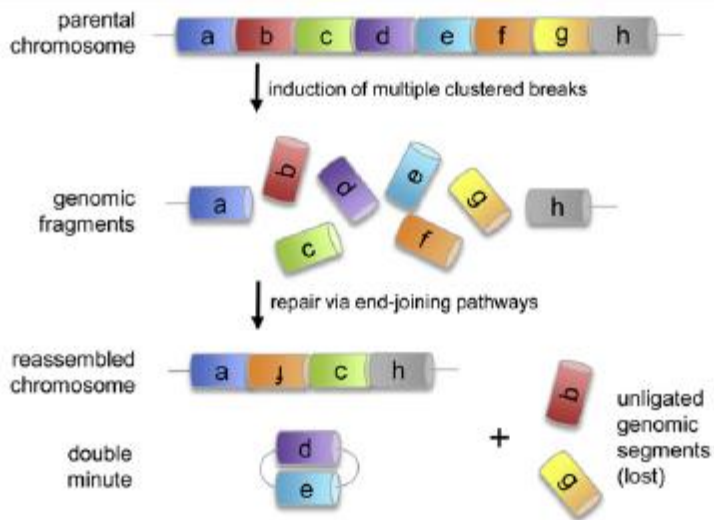




# Les résultats

## *Chromothripsis et chromoanasythesis*

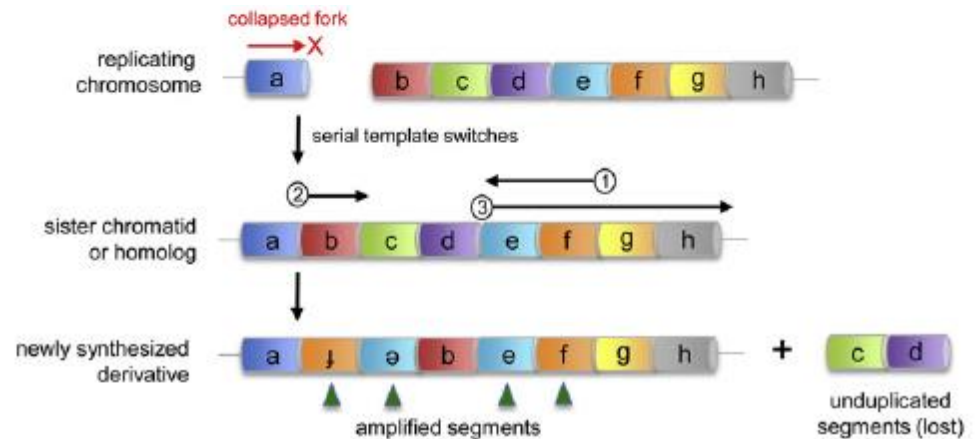
- Cassure double brin et NHEJ



↓

Chromothripsis

- Mécanisme réplcatif impliquant des régions de micro-homologies (MMBIR)



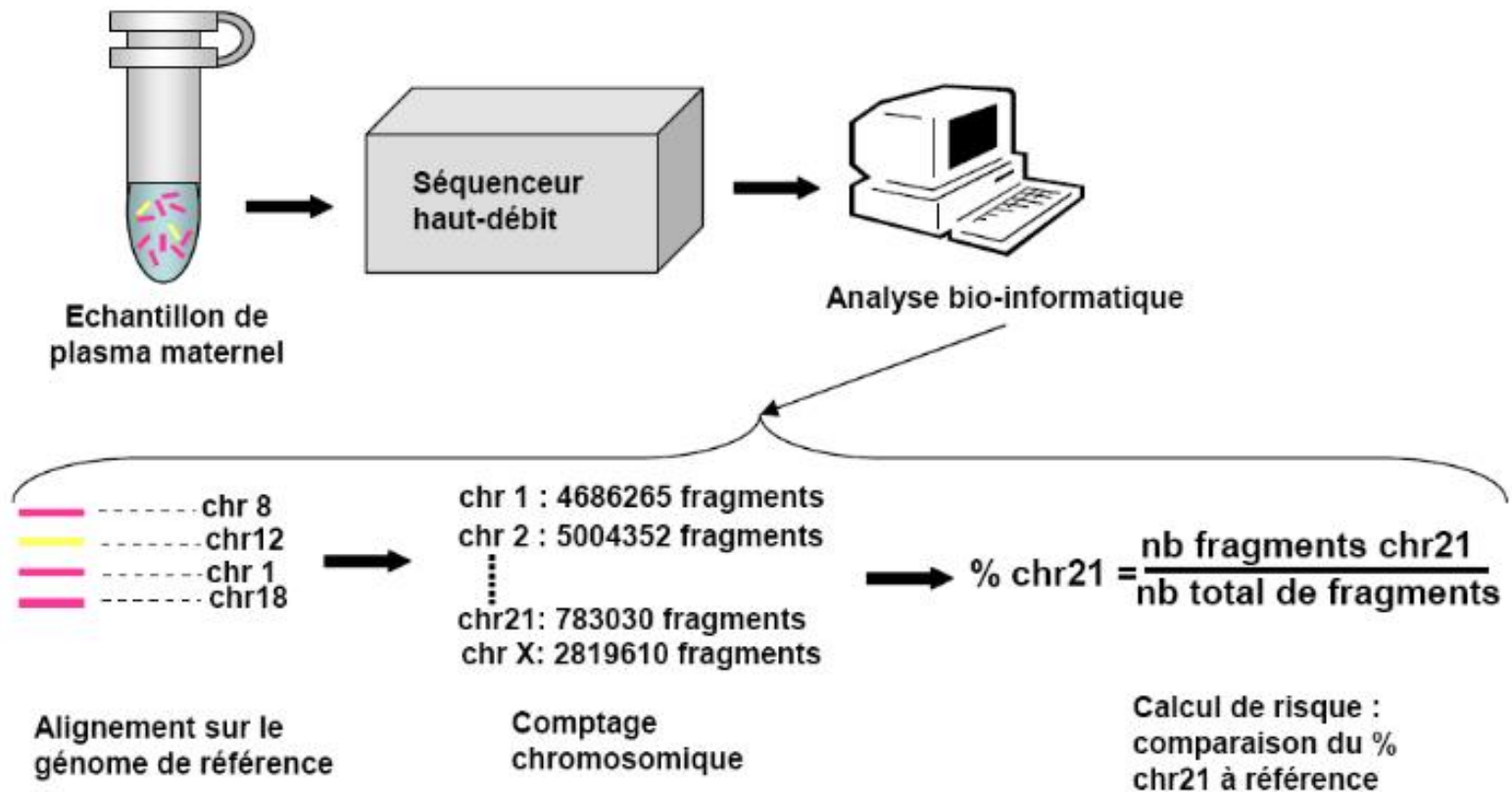
↓

Chromoanasythesis

# Applications diagnostiques

## *Dépistage prénatal non-invasif des aneuploïdies fœtales sur sang maternel*

(Chiu et al., PNAS, 2008; Fan et al., PNAS, 2008)



ADN plasmatique d'origine maternelle  
ADN plasmatique d'origine fœtale

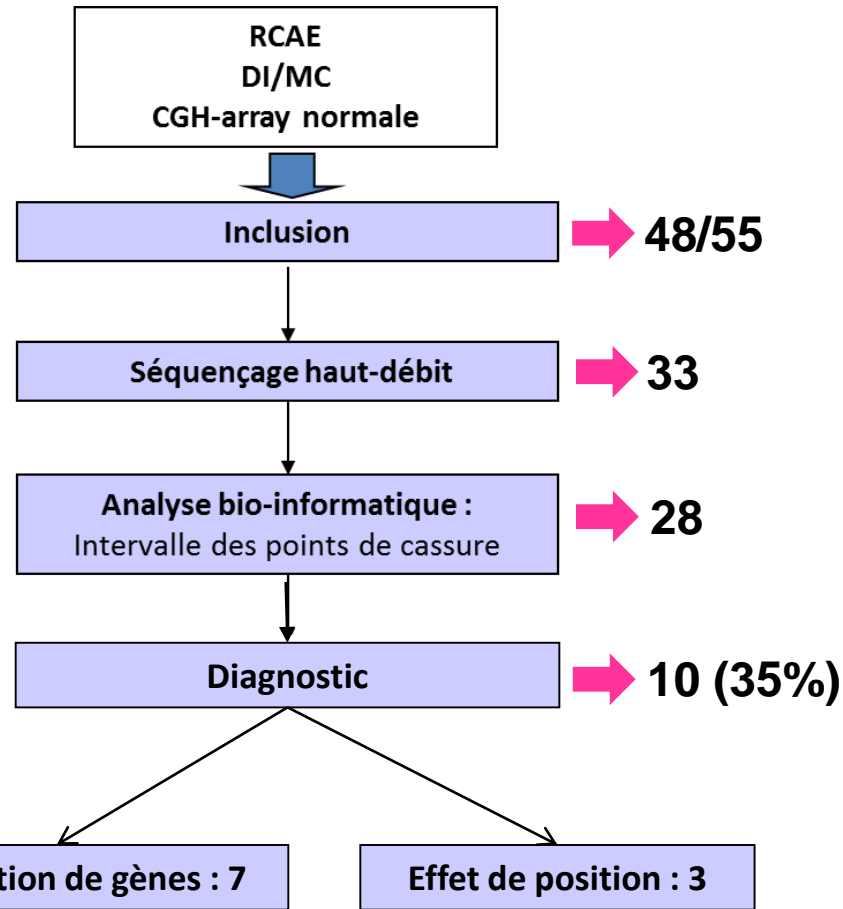
Détection trisomies 21, 13, 18

# Applications diagnostiques

## *Remaniements chromosomiques apparemment équilibrés à phénotype anormal*

- 6 à 9% des RCAE
  - Déséquilibre cryptique (ACPA)
  - Interruption de gène
  - Effet de position
  - Association fortuite
- Pathologie du point de cassure
- Whole genome sequencing : 1ères séries entre 2008-2012
  - Complexité supérieure /caryotype : 19% des patients
  - Interruption de gène : 80% des patients

# Remaniements chromosomiques apparemment équilibrés à phénotype anormal : projet ANI



- **136 points** de cassure en NGS (de 2 à 23 points de cassure) contre 70 points de cassure au caryotype
- **4 chromothripsis**

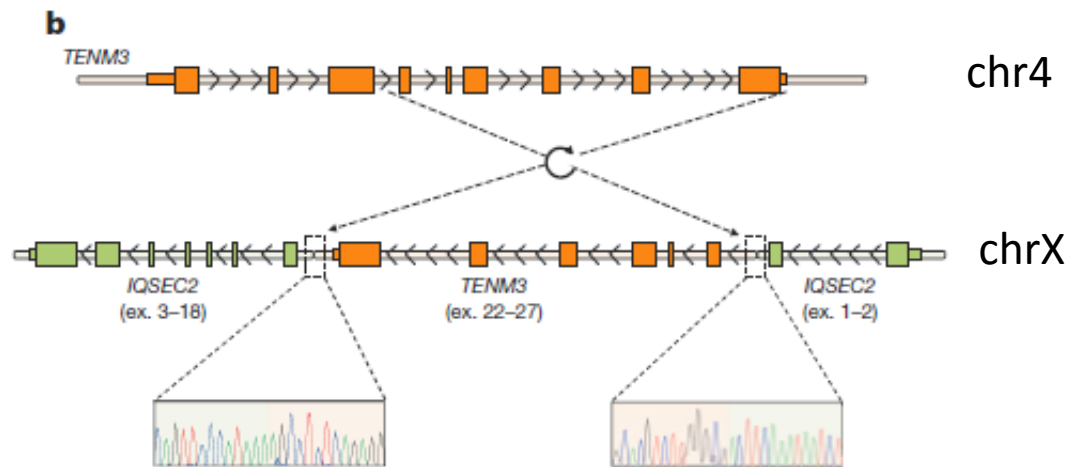
*KANSL1, SPRED1,  
MBD5 (x2), FOXP1,  
MEIS2, et DMD*

*MEF2C (x2)  
DLX5*

# Applications diagnostiques

## *Whole genome sequencing : analyse unique?*

- WGS de 50 trios atteints de DI sévère  
Recherche de SNVs et SVs  
84 mutations *de novo* dans les séquences codantes  
8 CNVs *de novo*, non vus en ACPA



Ex : Interruption du gène *IQSEC2* par une insertion de 62kb du gène *TENM3*

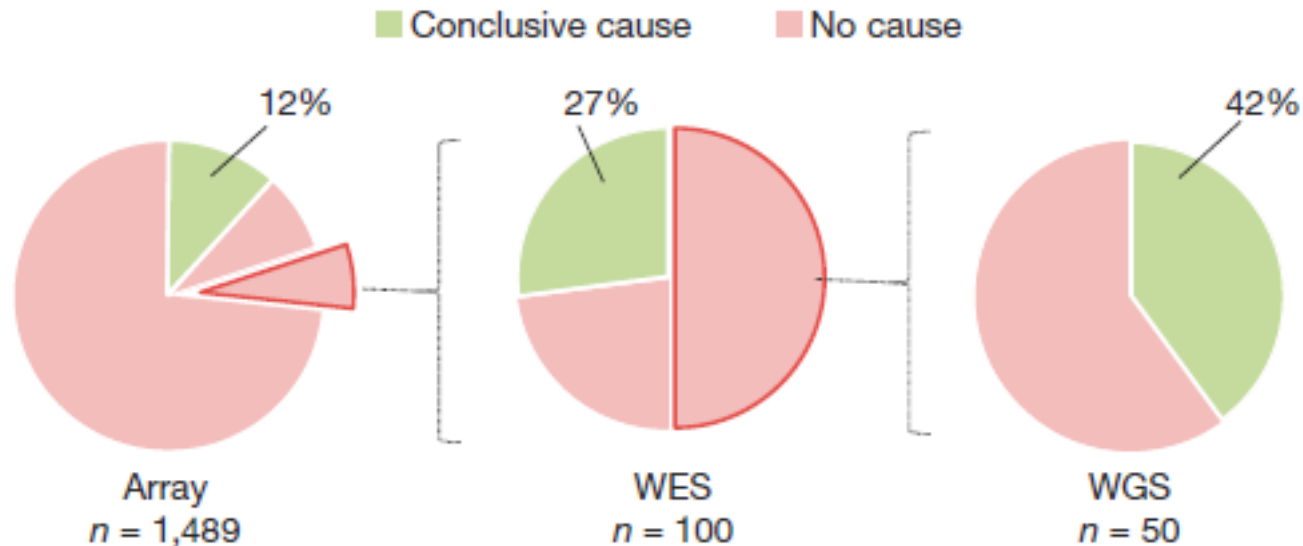
# Applications diagnostiques

## *Whole genome sequencing : analyse unique?*

- WGS de 50 trios atteint de DI sévère

21 diagnostics posés : 13 SNVs(AD), 1 SNV (AR), 7 CNVs

Identification de 8 gènes candidats



# Whole genome sequencing : analyse unique?

- Etude de faisabilité du transfert en diagnostic de la technique de WGS

6 patients avec diagnostic connu

6 patients sans diagnostic

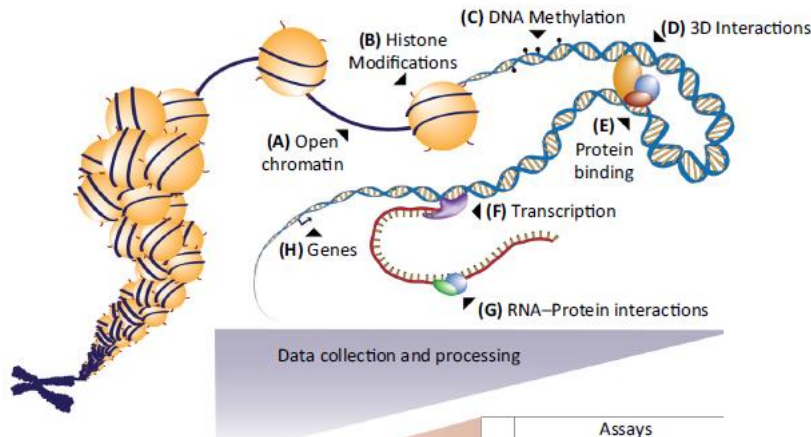
Statut	Diagnostics réalisés
P1 Mutation classique	5/8
P2 Mutation dans un exon difficile à séquencer en Sanger	4/8
P3 Mutation intronique profonde	0/7
P4 CNV à type de gain de grande taille	2/7
P5 CNV à type de perte de petite taille	6/8
P6 Remaniement chromosomique équilibré	3/7

- 2 patients : diagnostic réalisé
- 2 patients : gène candidat

# Perspectives

## *Vers une interprétation globale du génome?*

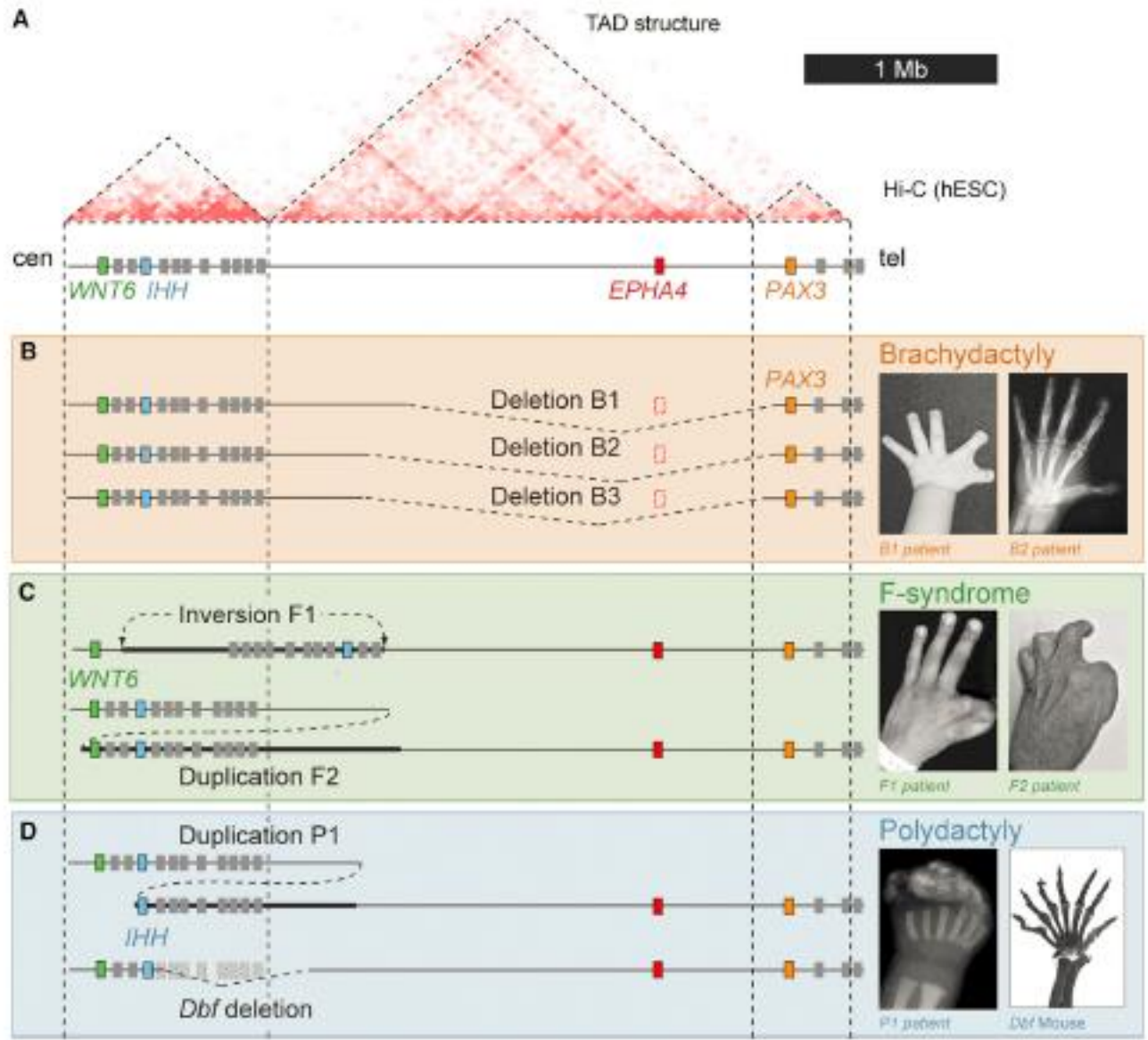
- Application du haut-débit à
- La méthylation de l'ADN : MedIP seq
- Les interactions ADN-protéines : Chip-Seq
- Les modifications des histones : ChIP-seq
- Les interactions intra-chromatine : capture de conformation (4C, HiC et dérivés)





# Vers une interprétation globale du génome?

- Exemple : rôle des TADs en pathologie



# Conclusion

Le whole genome sequencing

- Détection des remaniements de structure équilibrés et déséquilibrés
- Variations très nombreuses et souvent de petite taille
- Complémentarité des approches
- Applications diagnostiques : France Médecine Génomique 2025
- Intégration de l'ensemble des données haut-débit dans l'interprétation du génome

# Remerciements

- Service de Génétique, Laboratoire de Cytogénétique Constitutionnelle, HCL
- GENDEV, CRNL, INSERM U1028; CNRS UMR5292; UCBL1  
F Diguët, P-A Rollat-Farnier, L Pons, N Chatron, J Masson, D Sanlaville

- **Projet ANI**

Chambéry, Clermont-Ferrand, Dijon, Grenoble, Lille, Lyon, Marseille, Montpellier, Mulhouse, Nantes, Nice, Paris-Hôp. Trousseau, Paris-Hôp. Cochin, Bondy-Hôp. Jean Verdier, Paris-Hôp. Necker, Paris-Hôp. Pitié-Salpêtrière, Paris-Hôp. Robert Debré, Poitiers, Rouen, St-Etienne, Strasbourg, Toulon, Valence, Amiens, Bordeaux, Tours

- **Filière AnDDI-Rares, Filière DéfiScience**

C Vernin

Angers, Brest, Clermont-Ferrand, Dijon, Lyon, Marseille, Montpellier, Paris (Pitié), Poitiers, Rennes, Strasbourg