



# L'insertion d'éléments mobiles, un mécanisme mutationnel ultra rare: exemple du syndrome de Bardet-Biedl

Jean Muller

Laboratoire de Génétique Médicale (UMR\_S 1112)

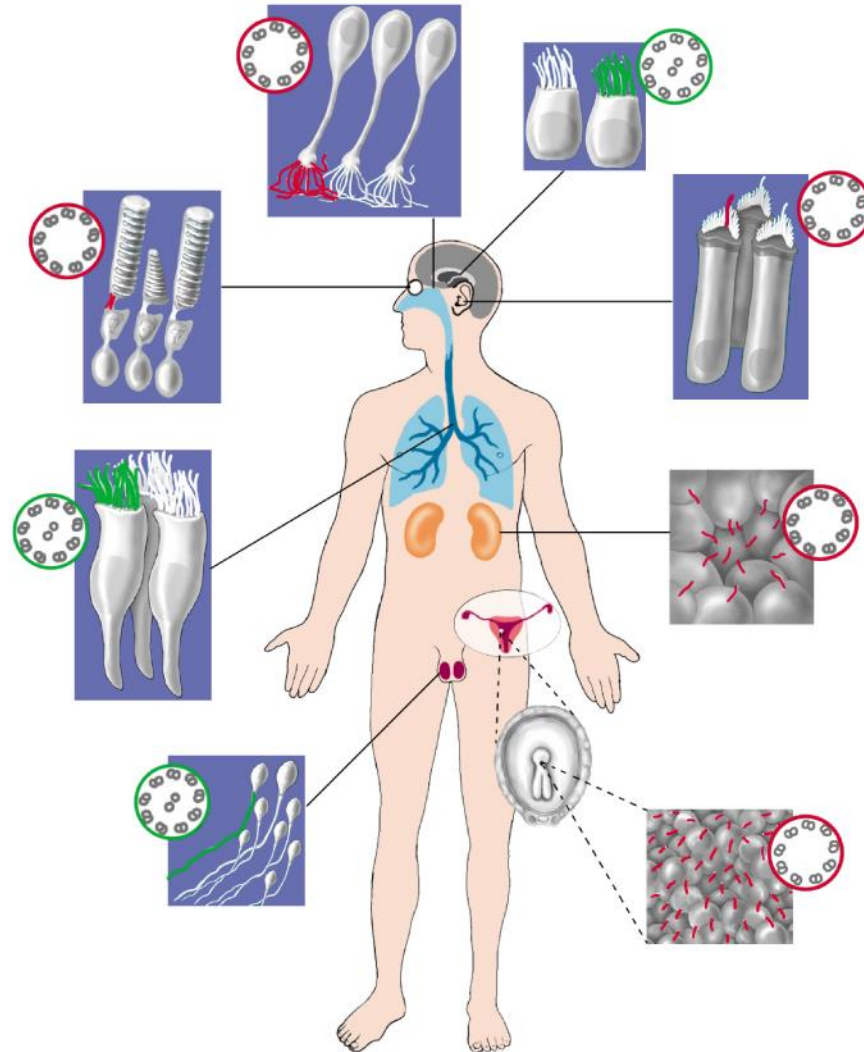
Laboratoires de Diagnostic Génétique

jeanmuller@unistra.fr



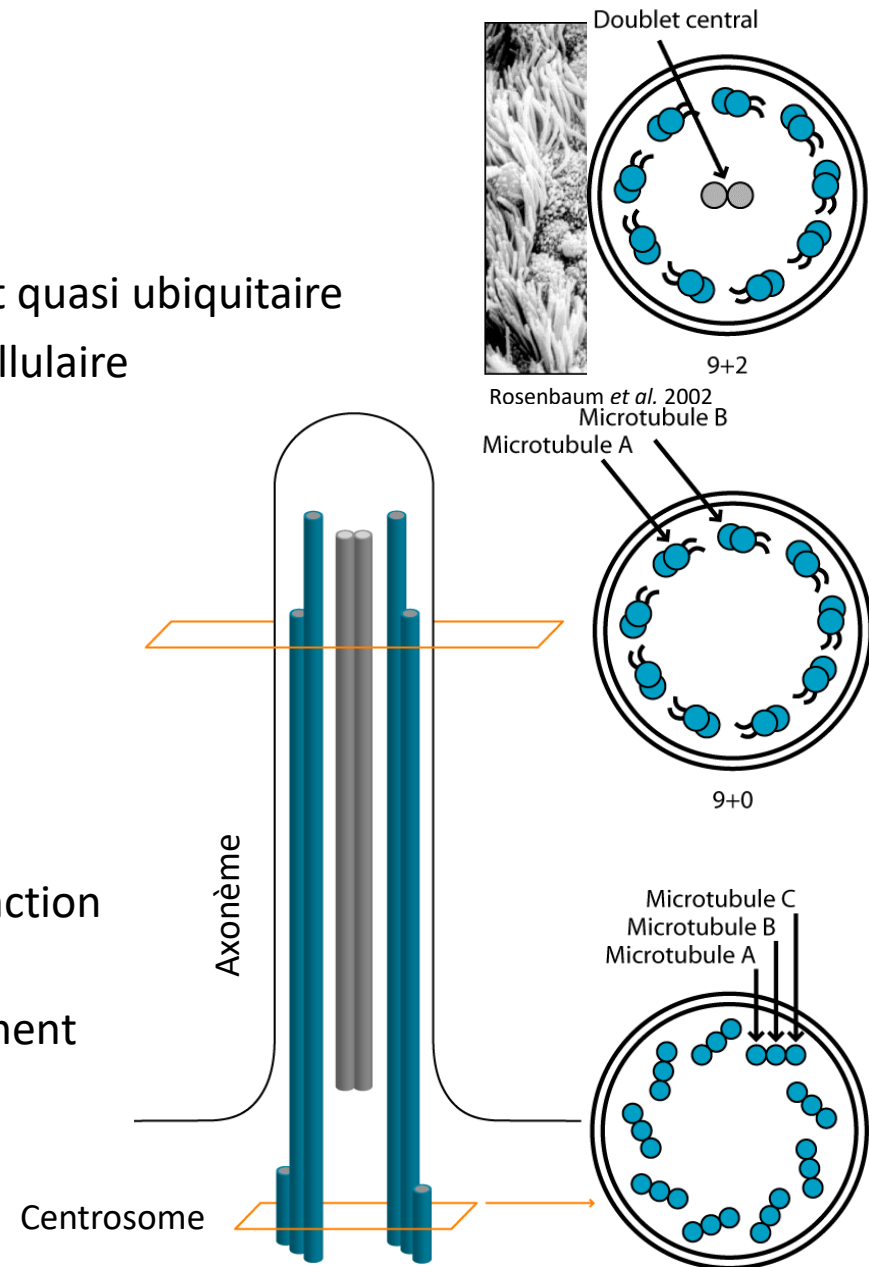
# Introduction

# Ciliopathies



# Le cil

- Organelle conservé au cours de l'évolution et quasi ubiquitaire
- Formation et résorption au cours du cycle cellulaire
  - Cil primaire « 9+0 »
  - Cil motile « 9+2 »
- Antenne cellulaire
  - Mécanorécepteur, chémorécepteur, photorécepteur
- Transmission des signaux à la cellule
- Activation de cascades de signalisation
- Rôle majeur dans le développement et la fonction des organes
- > 1000 protéines nécessaires au fonctionnement du cil

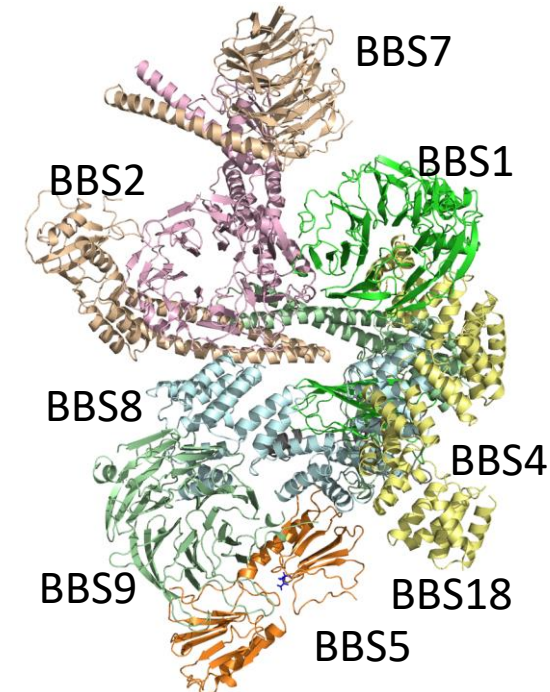


# Syndrome de Bardet-Biedl

- Rare  $\sim 1/150\,000$  (Europe du Nord)
- Transmission **autosomique récessive** (digénisme, oligogénisme, triallélisme)
- Hétérogénéité clinique:
  - variabilité intra et interfamiliale
- Hétérogénéité génétique:
  - **24 gènes BBS (417 exons,  $\sim 53$  kb CDS, 16 879 aa)**

## Modules fonctionnels:

- **BBSome**: BBS1, BBS4, BBS5...
- **Chaperonines-like**: BBS6, BBS10, BBS12
- **Transport intraflagellaire (IFT)**: IFT27, IFT172, IFT74



Chou *et al* 2019

# BBS – critères cliniques

Critères majeurs	Critères mineurs
Rétinite pigmentaire	Myopie sévère, strabisme, cataracte, astigmatisme
Polydactylie postaxiale	Syndactylie/brachydactylie
Obésité	Ataxie/troubles de la coordination
Lenteur d'idéation	Spasticité modérée des membres inférieurs
Reins hyperéchogènes ou kystiques	Retard psychomoteur
Hypogénitalisme masculine	Troubles du langage
Hypertension d'origine inconnue	Polyurie-polydipsie (diabète insipide néphrogénique)
	Anomale de l'émail dentaire/de l'éruption dentaire
	Intolérance glucidique
	Hypertrophie du ventricule gauche/cardiopathie
	Fibrose hépatique
4 critères majeurs	3 critères majeurs + 2 mineurs
<b>4M</b>	<b>3M + 2m</b>

Beales *et al* 1999

# BBS – distribution des gènes en cause



Postdoc

PhD

PhD

IR

**2 gènes majeurs (*BBS1/BBS10*)**  
15% patients non informatifs (négatifs)

- Identification

***BBS10***

***BBS12***

***LZTFL1/BBS17***

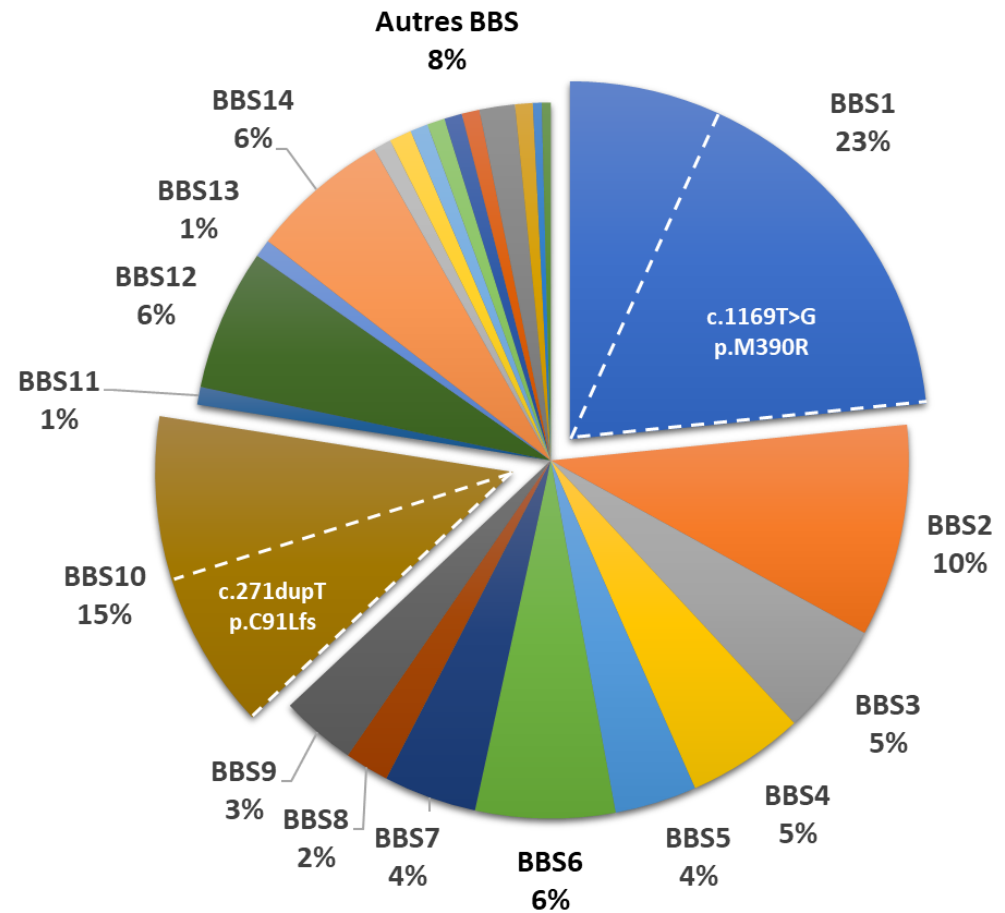
***BBIP1/BBS18***

- Confirmation

***IFT27/BBS19***

***IFT172/BBS20***

***IQCE***



(adapté de GeneReviews)

Stoetzel *et al*, Nature Genetics 2006

Stoetzel *et al*, AJHG 2007

Schaefer *et al*, CG 2013

Scheidecker *et al*, JMG 2013

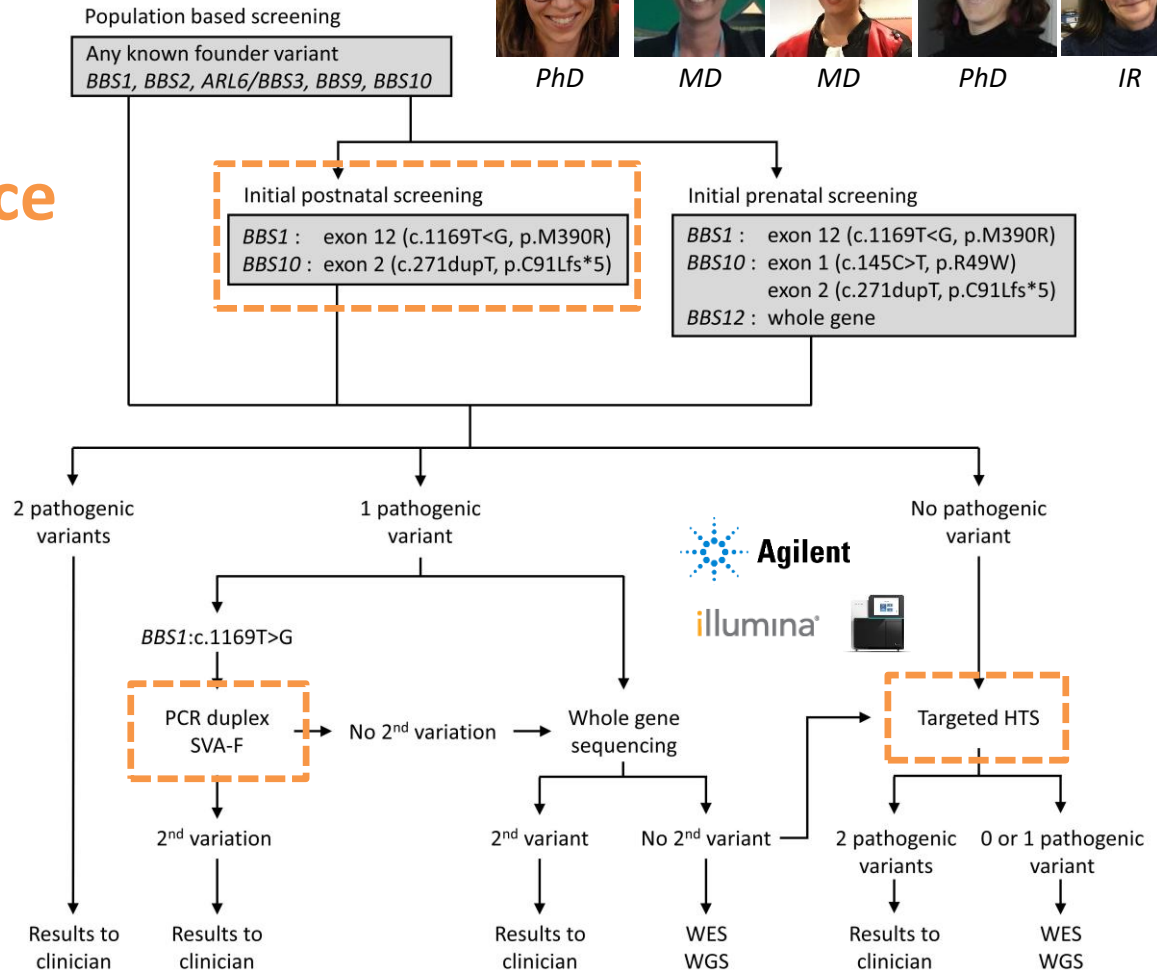
Schaefer *et al*, JHG 2016

Schaefer *et al*, FIG 2019 Estrada *et al*, HM 2019



# BBS – Arbre décisionnel diagnostique

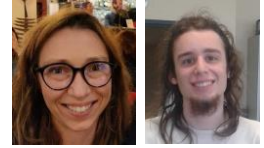
La recherche au service du patient



Muller *et al*, Human Genetics 2010  
 Redin *et al*, Journal Medical Genetics 2012  
 Mary *et al*, Clinical Genetics 2019  
 Gouronc *et al*, Clinical Genetics 2020  
 Delvallée *et al*, Clinical Genetics 2020



# BBS – A propos des insertions éléments mobiles



PhD

M2

## Importance de la détection des insertions éléments mobiles

- SVA F (*BBS1*), 8 cas
- AluYa5 (*BBS12*), 1 cas
- AluSx (*BBS10*), 1 cas\*

Redin *et al*, Journal Medical Genetics 2012

Delvallée *et al*, Clinical Genetics 2020

\* En cours de validation

# Les éléments mobiles du génome humain

# Mise en évidence

## Barbara McClintock (1902-1992)

Prix Nobel Physiologie ou Médecine en 1983

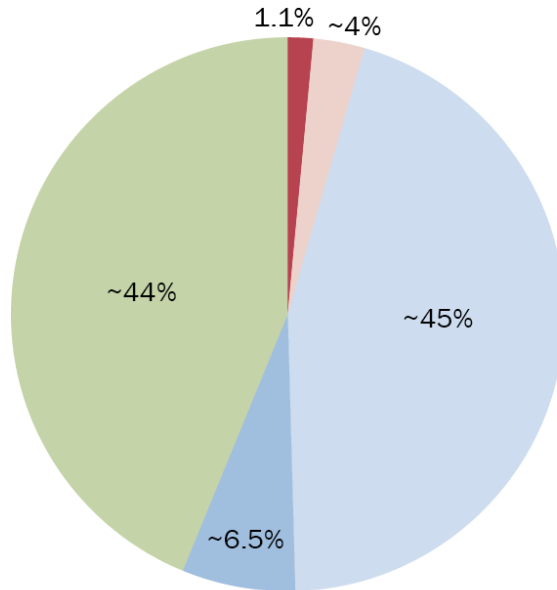
Raison: “pour la découverte des éléments génétiques mobiles”

Entre 1940 et 1950, elle a démontré que des éléments génétiques peuvent parfois **changer de position** sur un chromosome et que cela rend les gènes voisins **actifs ou inactifs**.



*Zea mays*

# Composition générale du génome

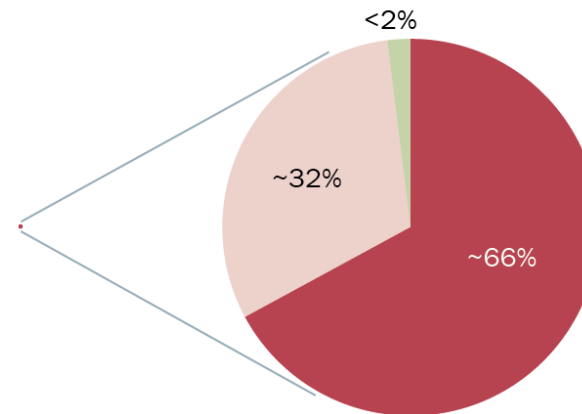


Génome nucléaire

Séquences très conservées

- Séquences codantes
- Séquences non codantes

>50% éléments répétés, faible complexité,



Génome mitochondrial

Séquences peu conservées

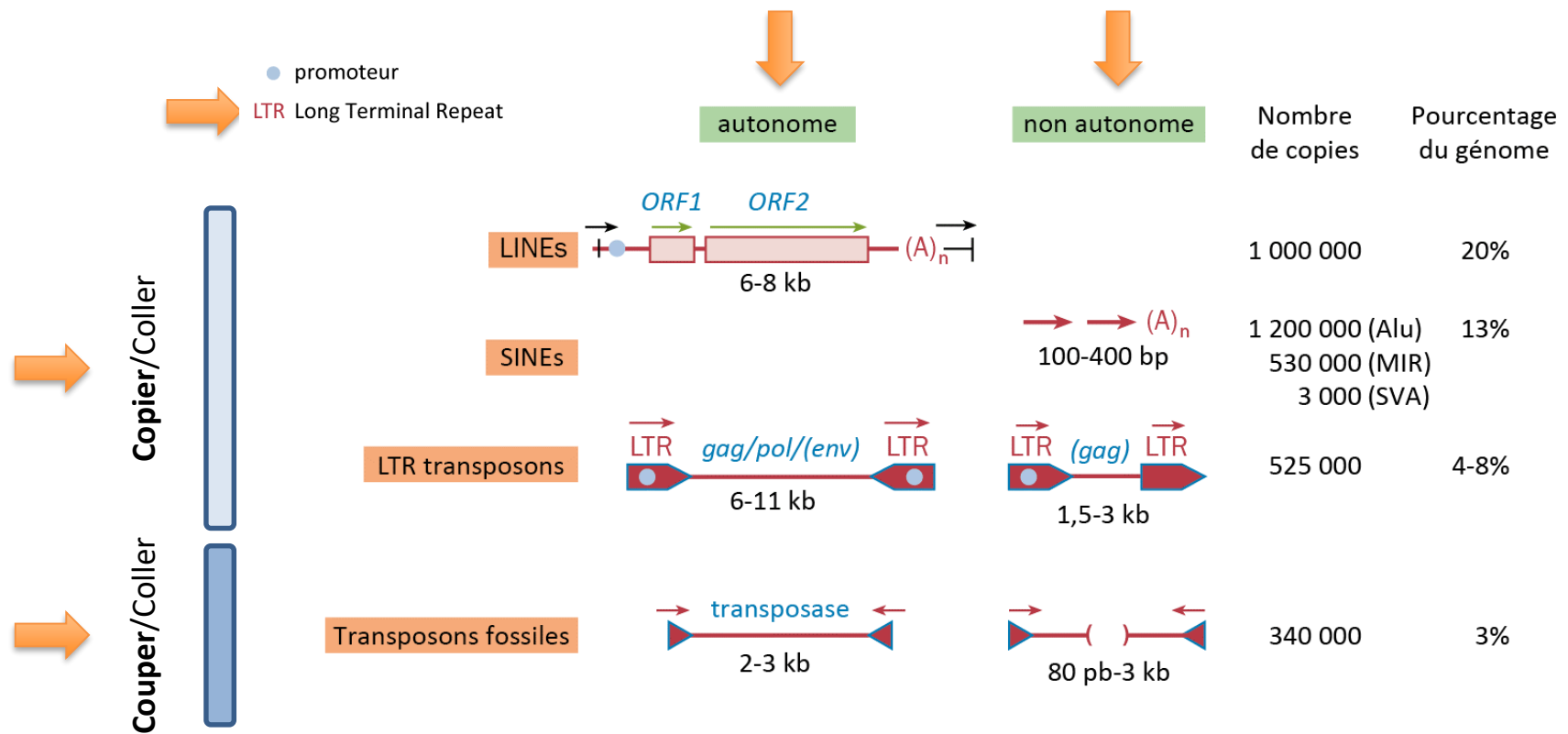
- Eléments répétés (transposons)
- Hétérochromatine constitutive (centromère, télomère)
- Autres séquences

# Élément répétés et mobiles pour certains

- **Éléments répétés en tandem**
  - 6,5% du génome
    - ADN satellites, minisatellites, microsatellites
- **Éléments répétés dispersés (« interspersed »)**
  - 40-45% du génome
    - Rétrotransposons
      - LINE, SINE (Long/Short Interspersed Elements), rétrovirus
    - Transposons
- **Pseudogènes**
  - 14,000 copies

# Éléments mobiles

- La plupart de ces éléments sont inactifs
- Un des mécanismes de l'évolution des génomes



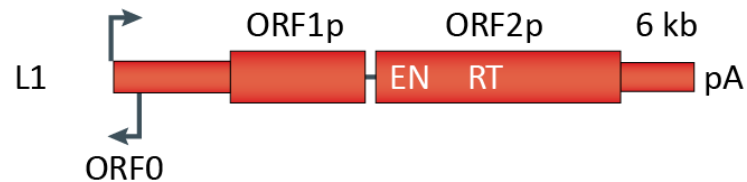




# Élément mobiles structure et mécanismes

LINE: « Long Interspersed Nuclear Elements »

- 3 sous-familles (Line-1, Line-2, Line-3), seule la première est encore active
- 1 000 000 de copies dont **<1% sont complètes**
- **~100** éléments encore **actifs** (Line-1)

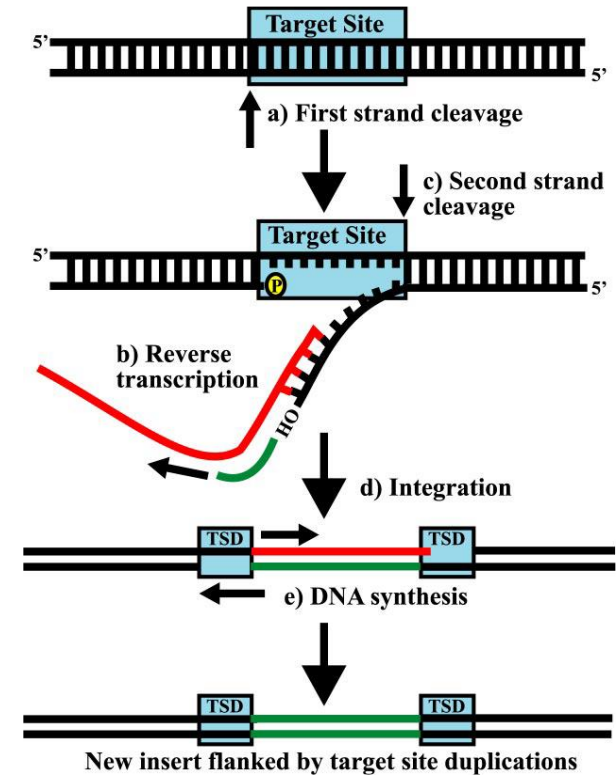
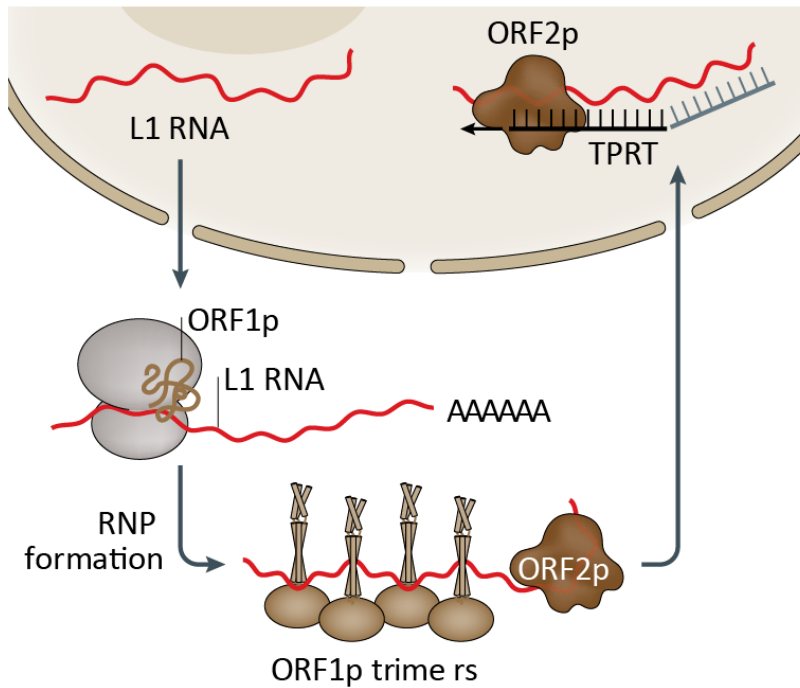


ORF1: RNA Binding protein

ORF2: Protéine activité endonucléase + reverse transcriptase

ORF: Open Reading Frame

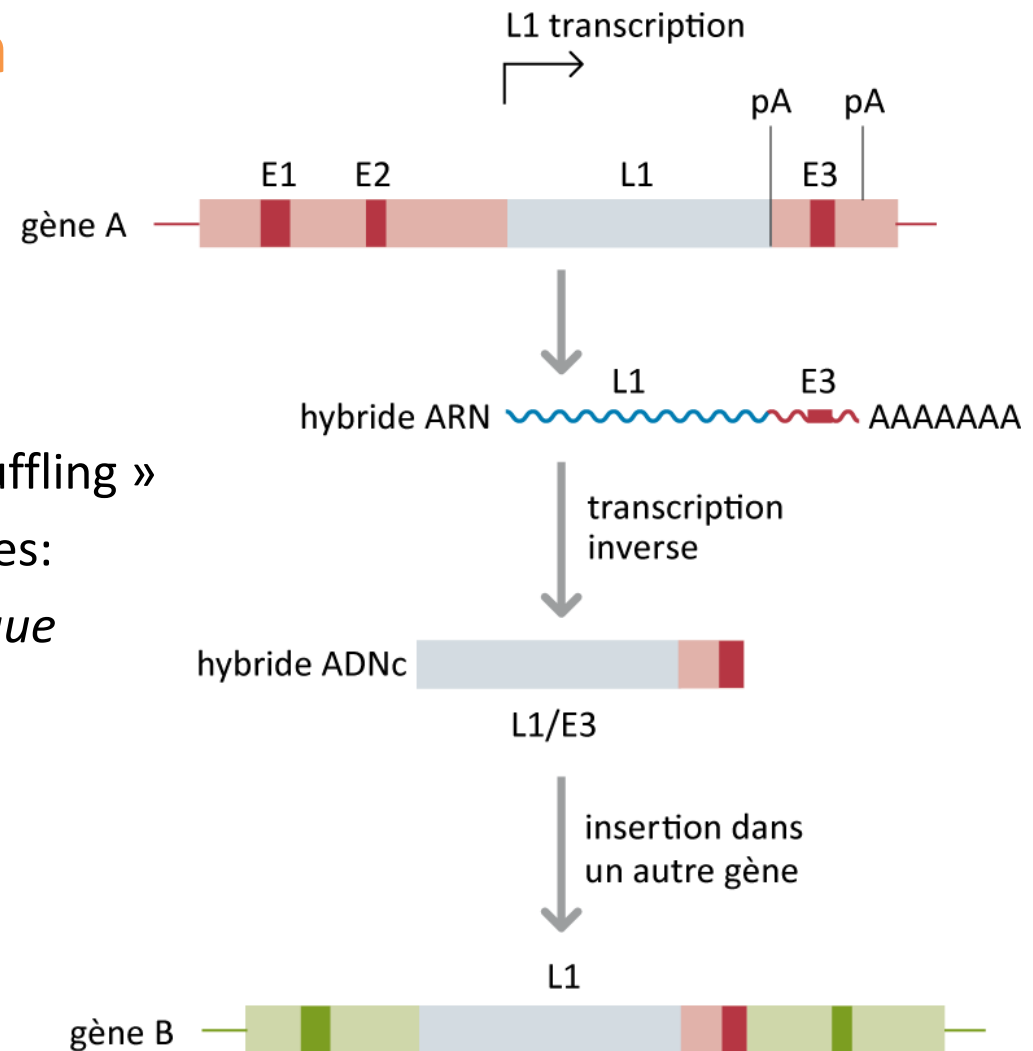
# Élément mobiles structure et mécanismes



TSD : Target Site Duplication

# Moteur de l'évolution

- Brassage d'exon ou exon « shuffling »
- Plusieurs mécanismes possibles:
  - *Recombinaison non-allélique*
  - Rétransposons (LINEs)



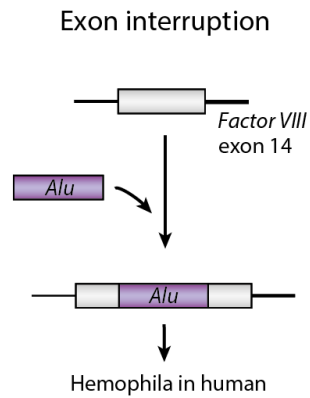
# Occurrence dans la population générale

- Fréquence d'apparition\* :
  - L1 1/63 à 1/231 naissances
  - Alu 1/40 à 1/42 naissances
  - SVA 1/63 à 1/309 naissances
- Régulation négative forte pour éviter un excès de MEI
  - Éléments incomplets (variations, réarrangements)
  - Méthylation (contrôle d'expression)
  - Modification des histones pendant embryogénèse

\*33 familles (599 WGS) et 2396 familles (9599 WGS)

# Mécanismes pathogénicité

# Quels mécanismes de pathogénicité?



# Occurrence en pathologie

- Fréquence d'apparition en pathologie:
    - L1 30 cas
    - Alu 76 cas
    - SVA <30 cas
- <200 cas décrits**
- Taux diagnostic très faible:
    - DDD 9 cas *de novo* sur 9 738 trio WES -> 0,04%
    - GeneDx 13 cas sur 38 871 trio WES -> 0,15%
    - Dijon 2 cas sur 2 410 WES -> 0,0008%

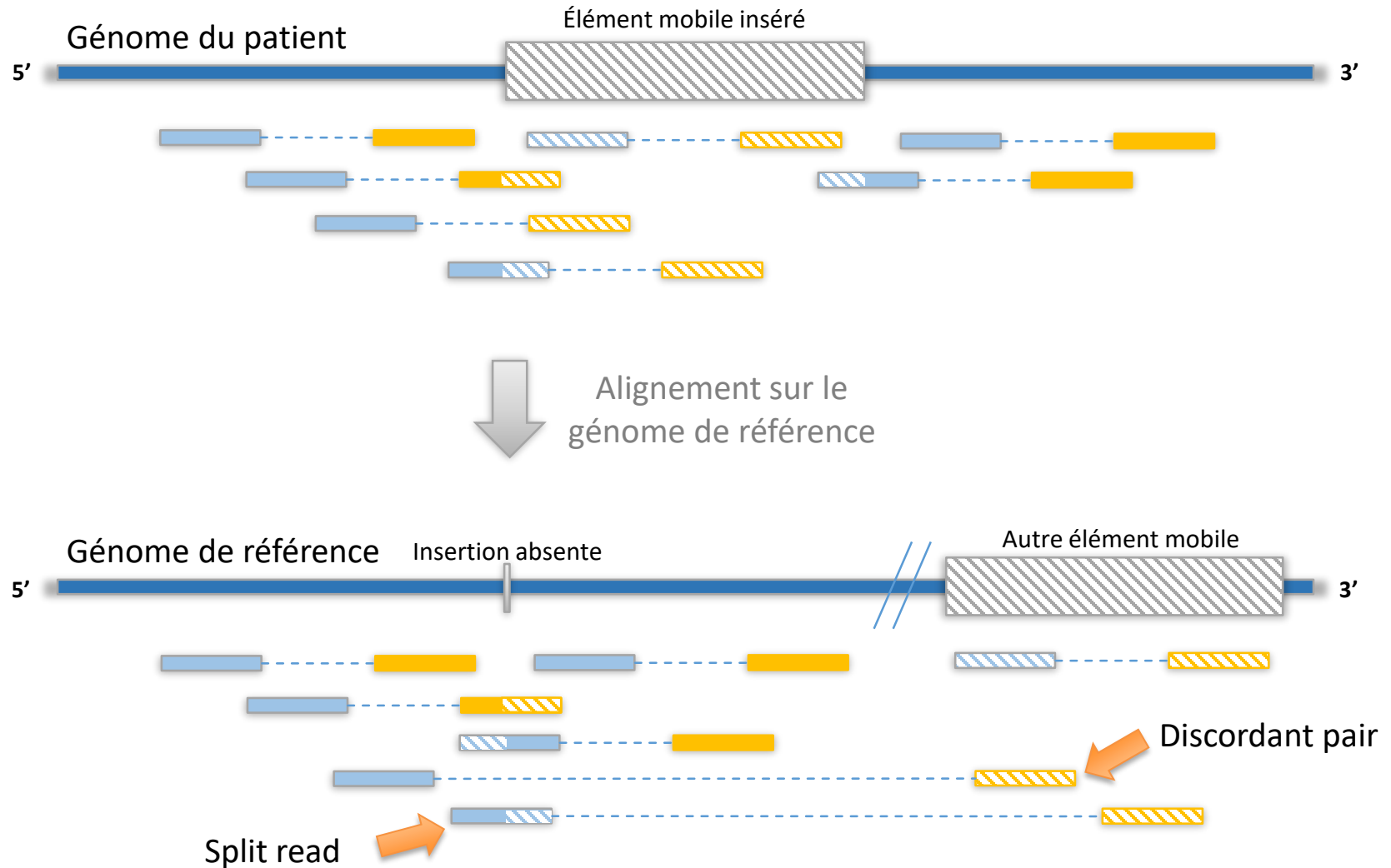
DDD Deciphering Developmental Disorders



# Insertion d'éléments mobile (MEI)

## Détection

# Détection des MEI



# Outils de détection

Thung et al. *Genome Biology* 2014, 15:488  
<http://genomebiology.com/2014/15/10/488>



METHOD

Open Access

## Mobster: accurate detection of mobile element insertions in next generation sequencing data

Djie Tjwan Thung<sup>1</sup>, Joep de Lig<sup>1,4</sup>, Lisenka EM Vissers<sup>1</sup>, Marloes Steehouwer<sup>1</sup>, Mark Kroon<sup>2</sup>, Petra de Vries<sup>1</sup>, Eline P Slagboom<sup>2</sup>, Kai Ye<sup>3</sup>, Joris A Veltman<sup>1,3</sup> and Jayne Y Hehir-Kwa<sup>1\*</sup>

### Mobster

2014 -> 56 citations

GRCh37 mais réactivation en 2022

Genetics  
inMedicine

BRIEF COMMUNICATION

Open 

### Mobile element insertion detection in 89,874 clinical exomes

Rebecca I. Torene, PhD, MMSc<sup>1</sup>, Kevin Galens, MS<sup>1</sup>, Shuxi Liu, PhD<sup>1</sup>, Kevin Arvai, MS<sup>1</sup>, Carlos Borroto, MS<sup>1</sup>, Julie Scuffins, MS<sup>1</sup>, Zhancheng Zhang, PhD<sup>1</sup>, Bethany Friedman, MS<sup>1</sup>, Hana Sroka, MS<sup>1</sup>, Jennifer Heeley, MD<sup>2</sup>, Erin Beaver, MS<sup>2</sup>, Lorne Clarke, MD<sup>3</sup>, Sarah Neil, MSc<sup>3</sup>, Jagdeep Walia, MBBS, FRCPC<sup>4</sup>, Danna Hull, MS<sup>4</sup>, Jane Juusola, PhD<sup>1</sup> and Kyle Retterer, MS<sup>1</sup>

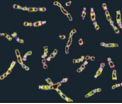
## The Mobile Element Locator Tool (MELT): population-scale mobile element discovery and biology



Eugene J. Gardner,<sup>1,2</sup> Vincent K. Lam,<sup>2,3</sup> Daniel N. Harris,<sup>1,2</sup> Nelson T. Chuang,<sup>1,2,4,5</sup> Emma C. Scott,<sup>1,2</sup> W. Stephen Pittard,<sup>6</sup> Ryan E. Mills,<sup>7,8</sup> The 1000 Genomes Project Consortium, and Scott E. Devine<sup>1,2,3,4</sup>

1000 Genomes

A Deep Catalog of Human Genetic Variation



### Melt

2017 -> 149 citations

Fin du support 05/2022

GeneDx

### Scramble

2020 -> 21 citations

Dernière version 04/2021

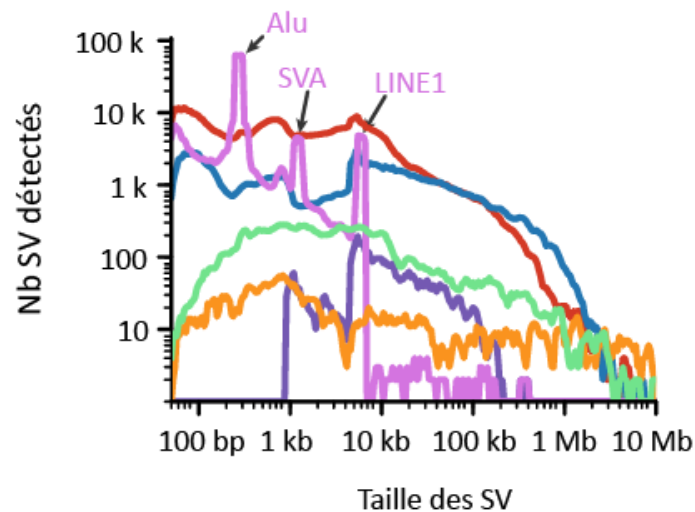
# Insertion d'éléments mobile (MEI)

## Base de données

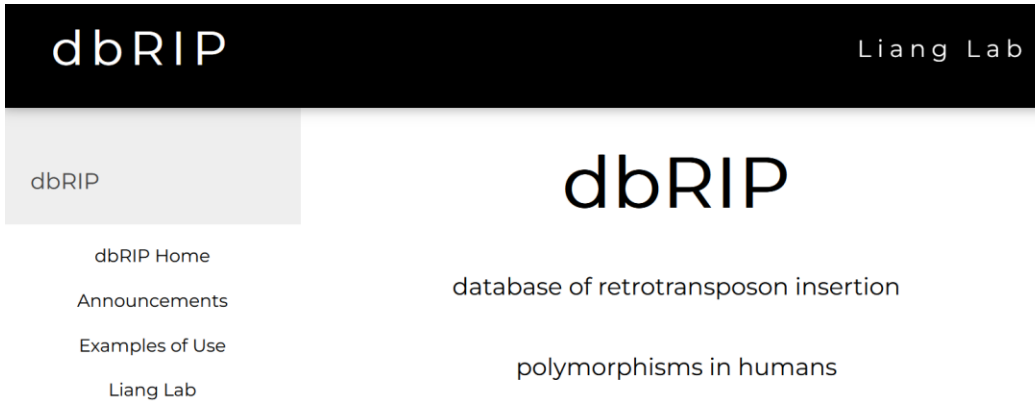
# Annotations – base de données

## gnomAD

- 14 891 WGS
- 498 257 SVs distincts (present dans 1 génome au moins)



# Annotations – base de données



RIP Class	# of loci (unique/total)
<b>Alu</b>	2086/2708
<b>L1</b>	598/800
<b>SVA</b>	77/87
<b>HERV</b>	10/10
<b>Total</b>	<b>2771/3605</b>

- Base de données d'insertion d'éléments mobiles polymorphiques humain (2006, nouvelle publi en cours)
- Anciennes données mais disponibles via genome browser de l'UCSC

# Annotations – base de données

Published online 25 February 2022

*Nucleic Acids Research*, 2022, Vol. 50, No. 5 2493–2508  
<https://doi.org/10.1093/nar/gkac128>

## Characterizing mobile element insertions in 5675 genomes

Yiwei Niu<sup>1,2,†</sup>, Xueyi Teng<sup>1,3,†</sup>, Honghong Zhou<sup>1,†</sup>, Yirong Shi<sup>1,3</sup>, Yanyan Li<sup>1,2</sup>,  
 Yiheng Tang<sup>1,3</sup>, Peng Zhang<sup>1</sup>, Huaxia Luo<sup>1</sup>, Quan Kang<sup>1</sup>, Tao Xu<sup>2,4,\*</sup> and  
 Shunmin He<sup>1,2,\*</sup>

<sup>1</sup>Key Laboratory of RNA Biology, Center for Big Data Research in Health, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China, <sup>2</sup>College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China, <sup>3</sup>University of Chinese Academy of Sciences, Beijing 100049, China and <sup>4</sup>National Laboratory of Biomacromolecules, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

Class	# of loci
Alu	26 553
LINE-1	7353
SVA	2667
HERV-K	126
<b>Total</b>	<b>36 699</b>

Database named HMEID:

<http://bigdata.ibp.ac.cn/HMEID/>



# Insertion d'éléments mobile (MEI)

## Syndrome de Bardet-Biedl

# Exemple 1 – Insertion d'un SVA dans *BBS1*

Importance de la détection des insertions éléments mobiles

- **SVA F (*BBS1*), 8 cas**
- **AluYa5 (*BBS12*), 1 cas**
- **AluX (*BBS10*), 1 cas\***

*BBS1*: c.1169T>G, p.M390R variation BBS la + fréquente

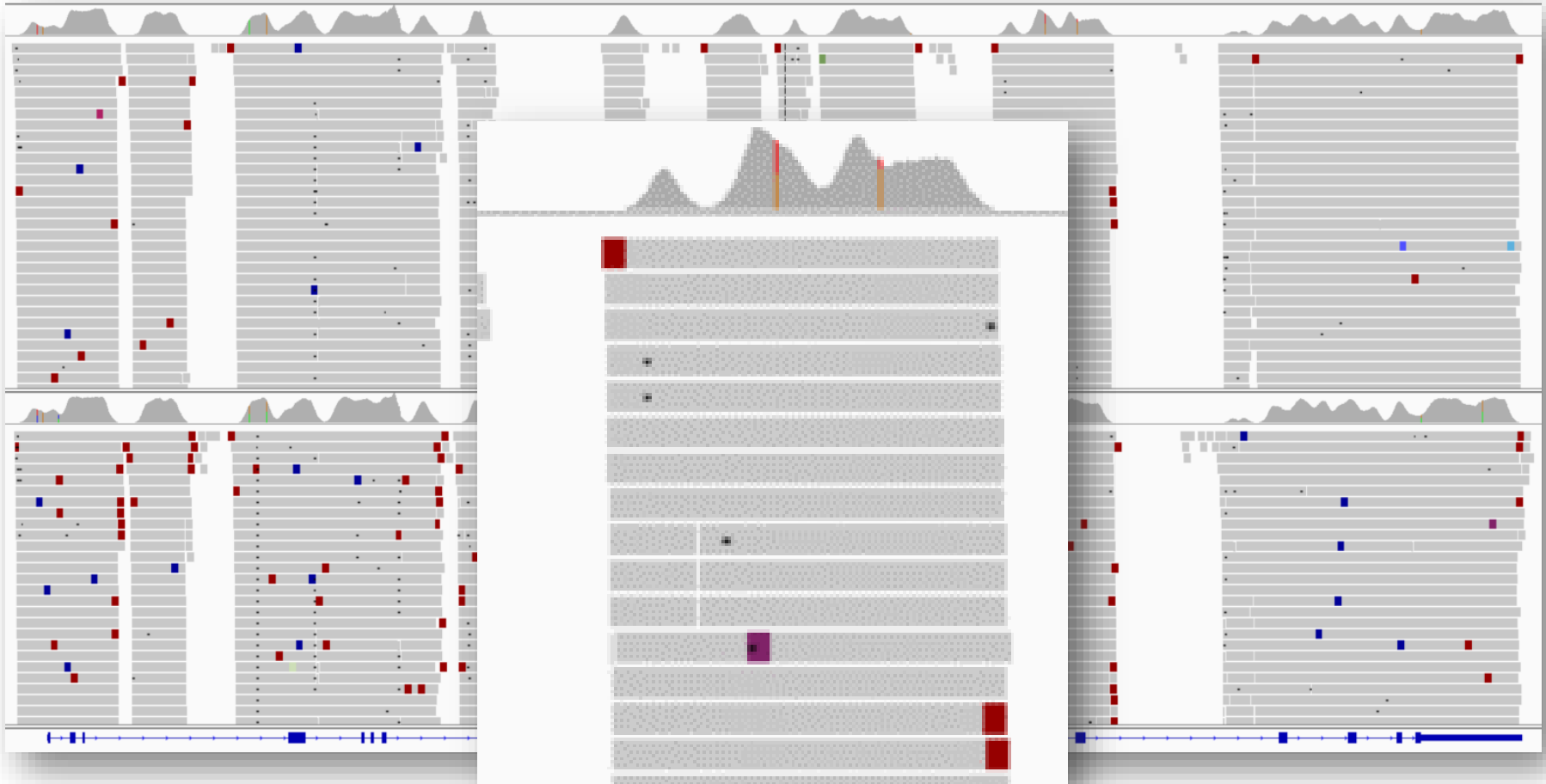
- 75% des patients *BBS1* = ~19% des BBS
- 0,27% dans la pop. générale

Excès de patients htz pour cette variation (sans second variant):

- *BBS1* est connu pour porter des grandes délétions (QMPSF, QPCR)
- Promoteur? Deep intronic splice?
- Hasard? Autre gène impliqué?
- Hypothèse de 2 M390R?

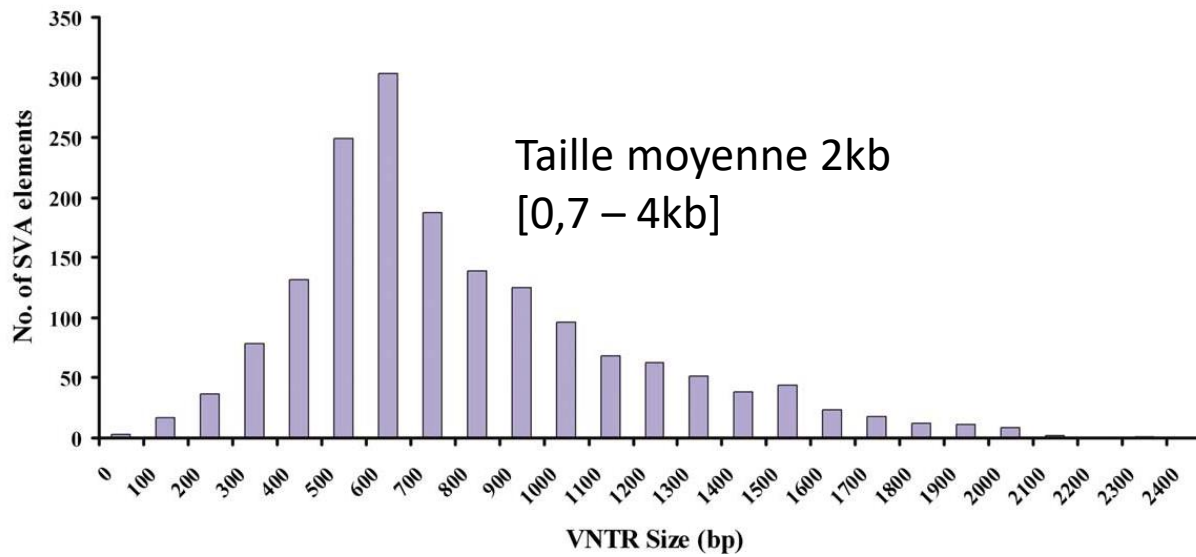
# Exemple 1 – Insertion d'un SVA

Pas de CNV, pas de second variant



# SVA - SINE-VNTR-Alus

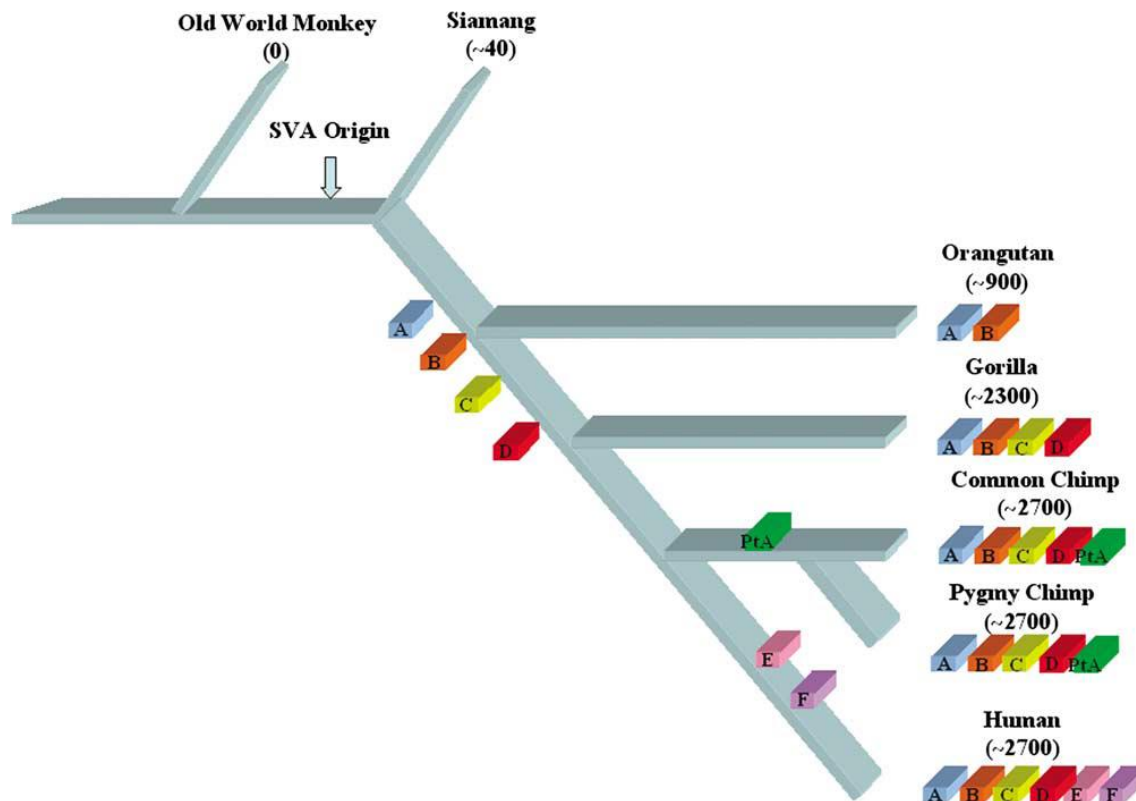
Structure complexe et taille variable



# SVA - SINE-VNTR-Alus

Plus jeune famille de rétrotransposons dans le génome humain

SVA\_F spécifique de l'Homme



# Exemple 1 – Insertion d'un SVA

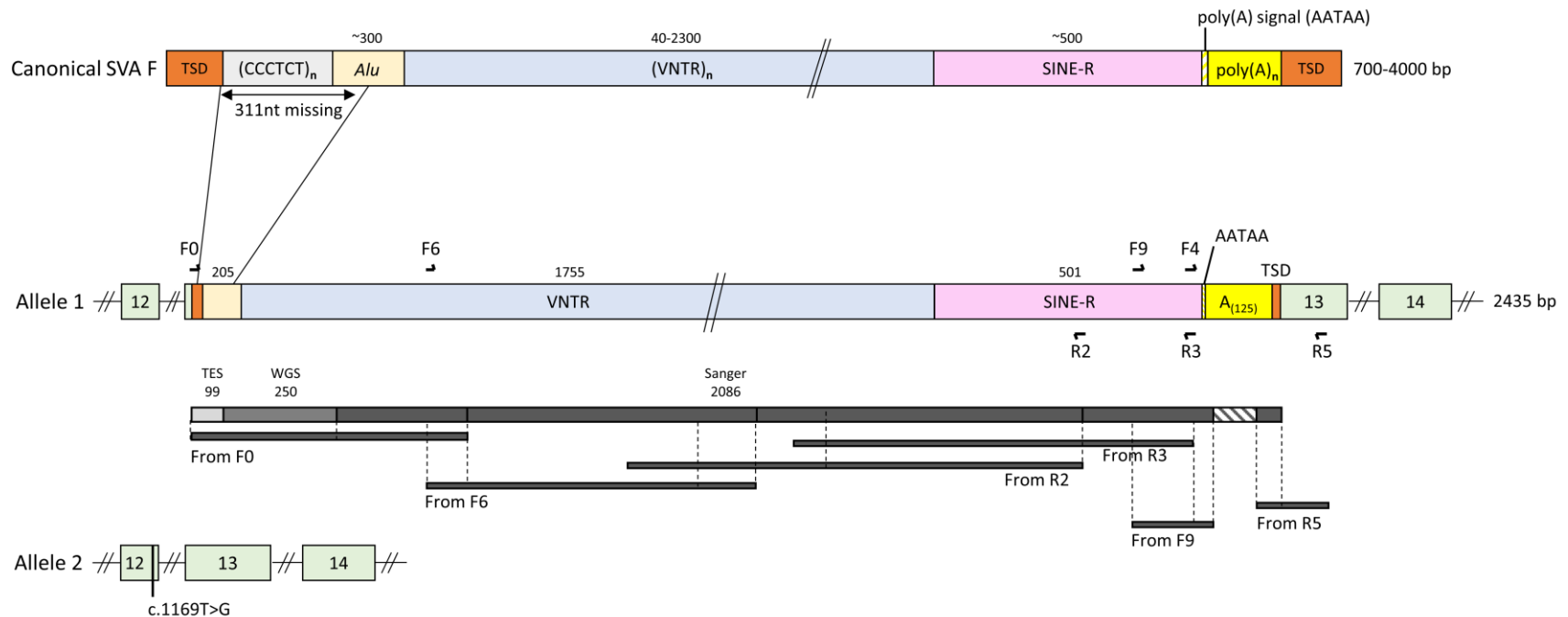


PhD

M2

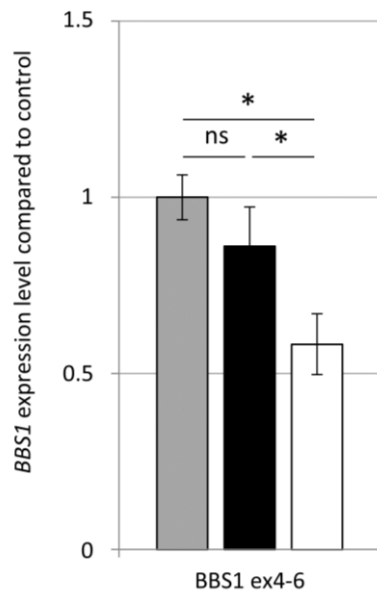
**Nomenclature HGVS: c.1214\_1215ins[MT113356], p.(Ala406Glnfs\*47)**  
**SVA type F de 2435 bp**

Panel, WGS, 10x genomics, reconstructions bioinfo (*de novo* assembly)  
 Sanger (marche sur le chromosome)

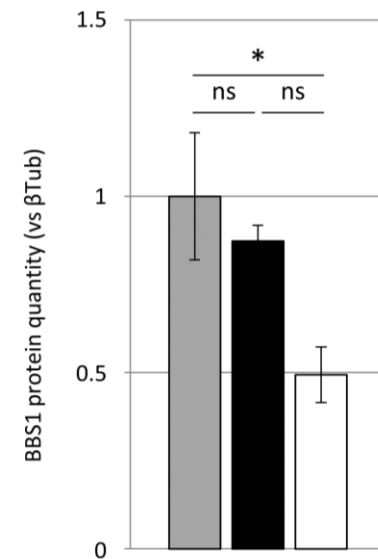
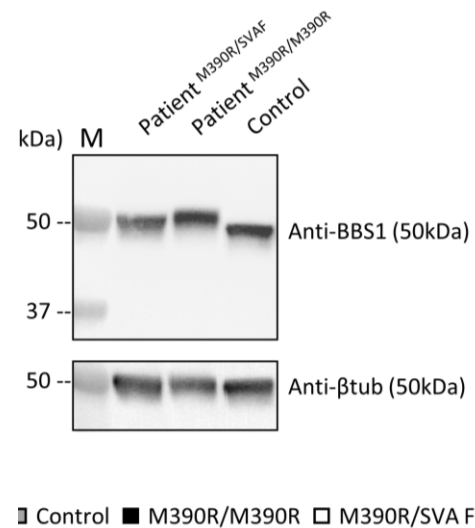


# Exemple 1 – Insertion d'un SVA

Impact de la variation sur le gène et la protéine BBS1?



Réduction significative de l'expression de *BBS1*



Absence de protéine tronquée



# Exemple 1 – Insertion d'un SVA

Screening de notre cohorte + collaborateur (Erica Davis, Chicago)

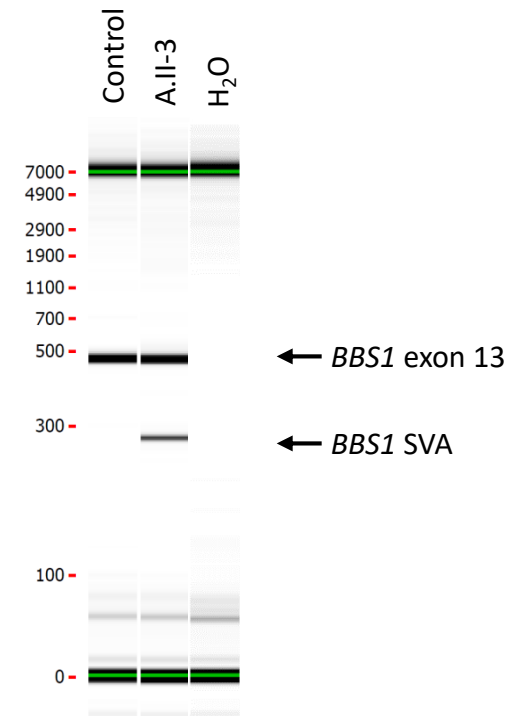
- Bioinformatique: 217 panels
- Bio. Mol.: 347 **PCR Duplex**

## 7 cas supplémentaires:

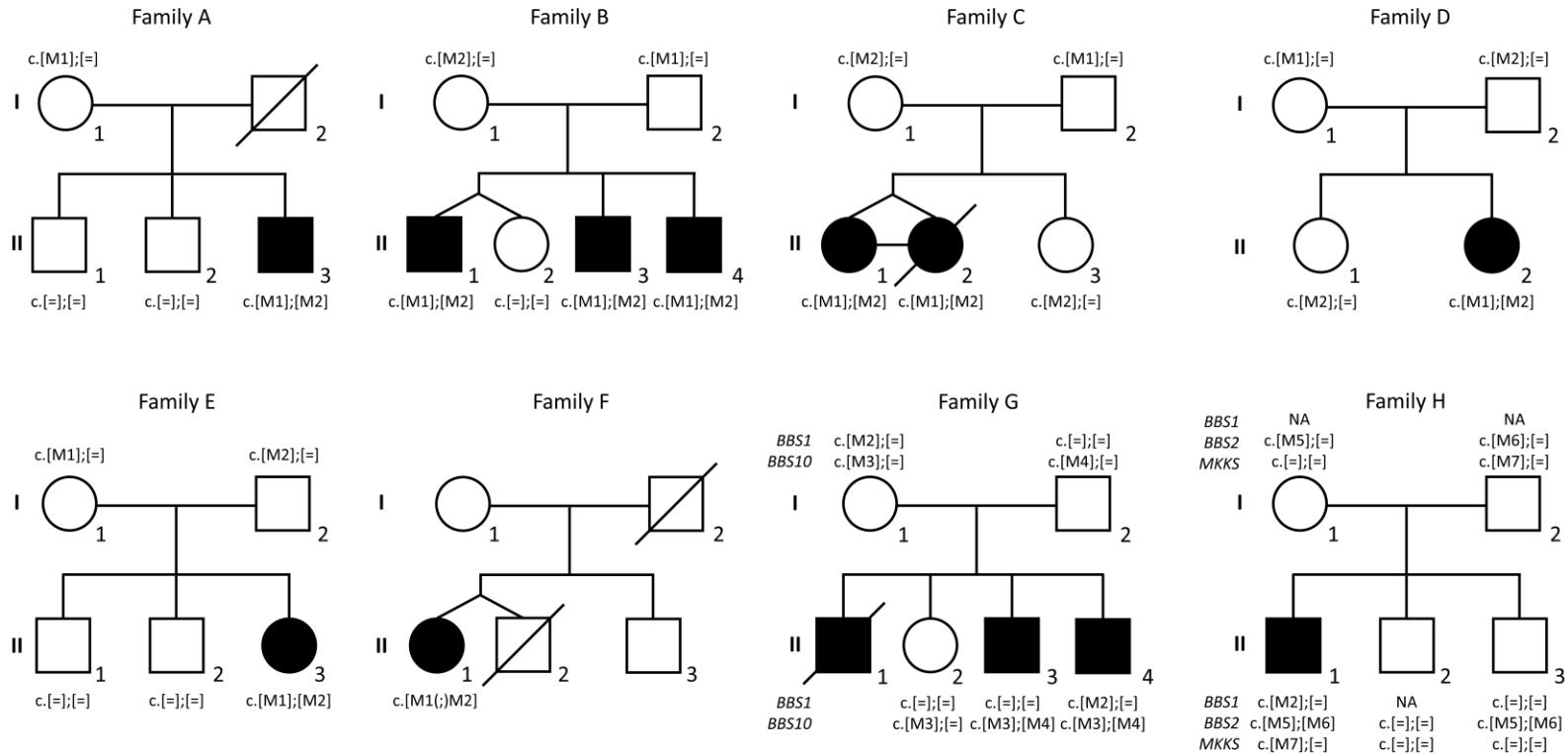
- Sanger a révélé le même point de cassure
- Effet fondateur ou bien mécanisme spécifique?

## Absent des populations contrôles:

- FranceGenRef dataset (862 WGS) - IGV
- 1k Genomes (2 504 WGS) - AnnotSV
- gnomAD SV 2.1 (14,891 WGS) - AnnotSV



# Exemple 1 – Insertion d'un SVA



M1: *BBS1* c.1169T>G, p.Met390Arg  
 M2: *BBS1* c.1214\_1215ins[MT113356], p.Ala406Glnfs\*47  
 M3: *BBS10* c.271dupT, p.Cys91Leufs\*5  
 M4: *BBS10* c.145C>T, p.Arg49Trp

M5: *BBS2* c.72C>T, p.Tyr24\*  
 M6: *BBS2* c.175C>T, p.Gln59\*  
 M7: *MKKS/BBS6* c.442C>T, p.Gln148\*

# Exemple 1 – Effet fondateur



E. Génin

## Effet fondateur en Europe

*BBS1* p.Met390Arg

Haplotype: ~1.614 Mb

Age: 93 générations 95% CI [49;183]

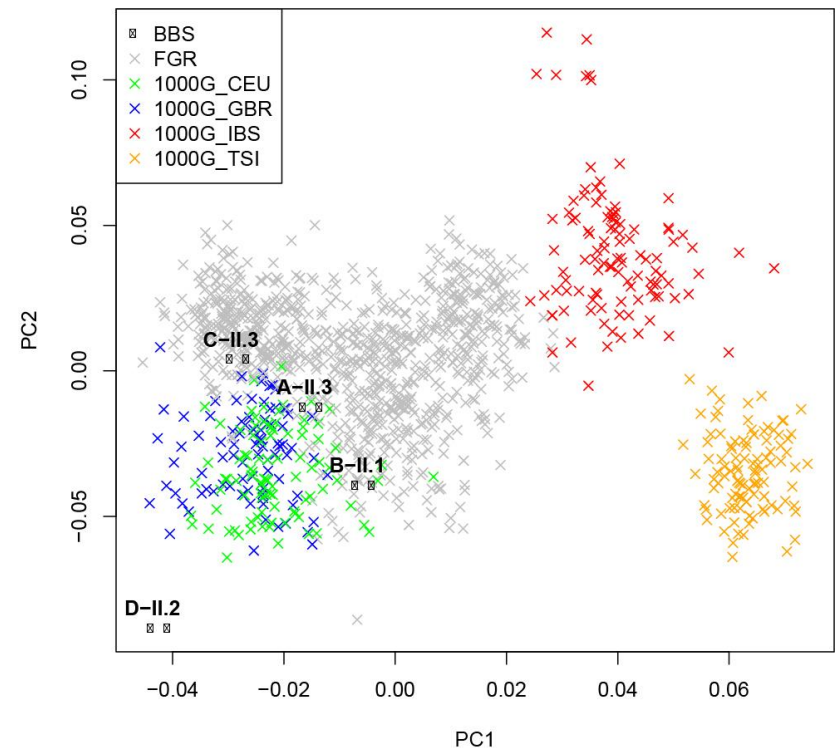
2325 ans - 95% CI [1225,4575 ans]

*BBS1* SVA F

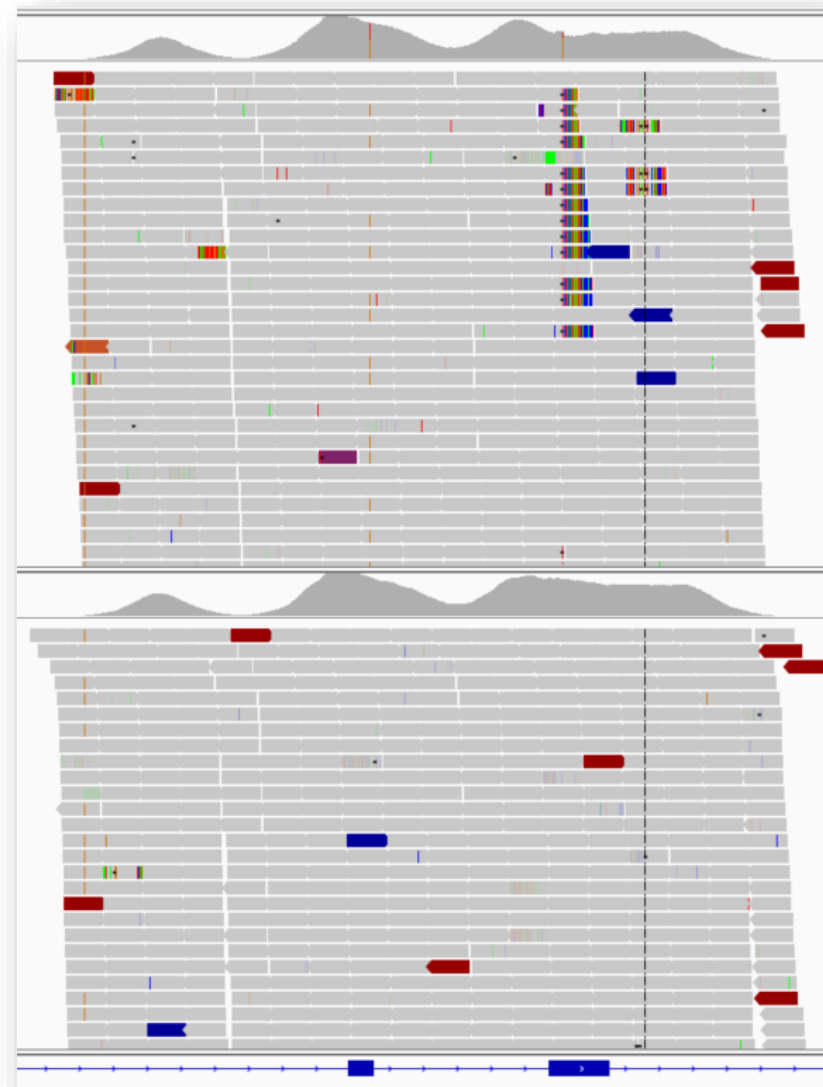
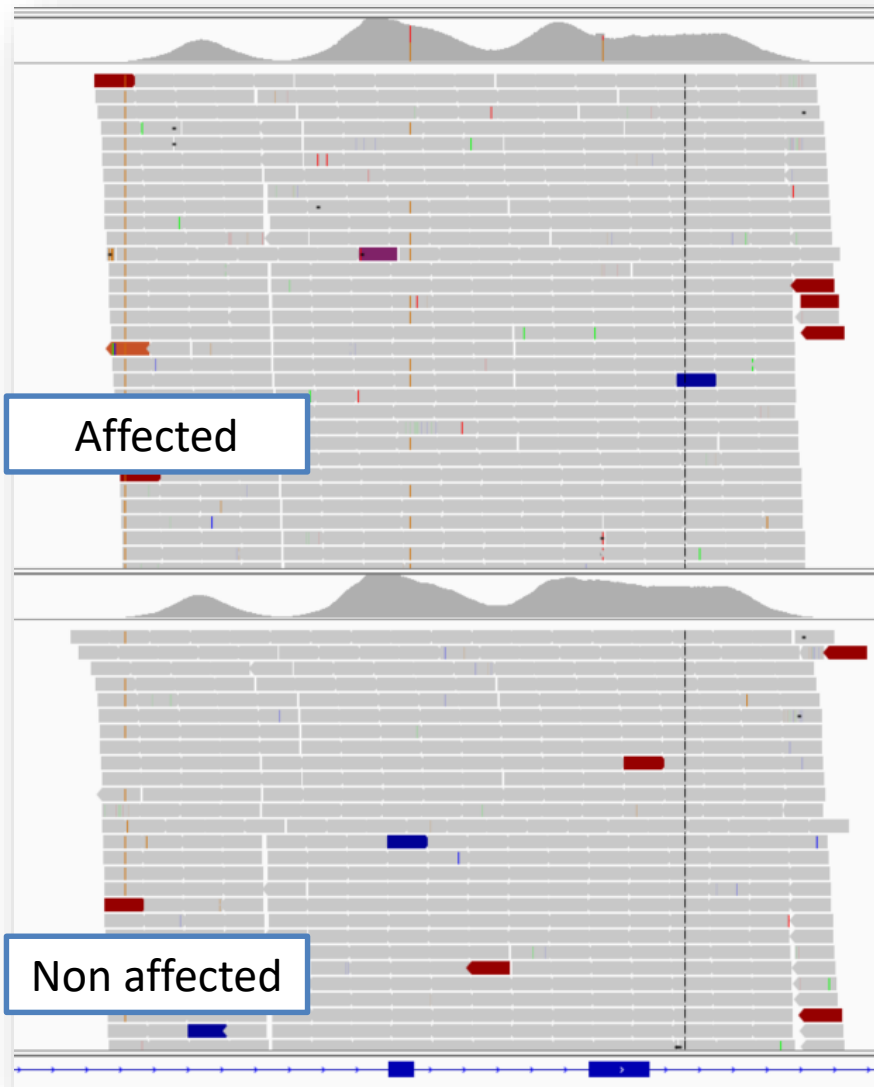
Haplotype: ~56.6 Kb

Age: 74 générations - 95% CI [33;190]

1852 ans - 95% CI [825,4750 ans]



# Exemple 1 – Insertion d'un SVA



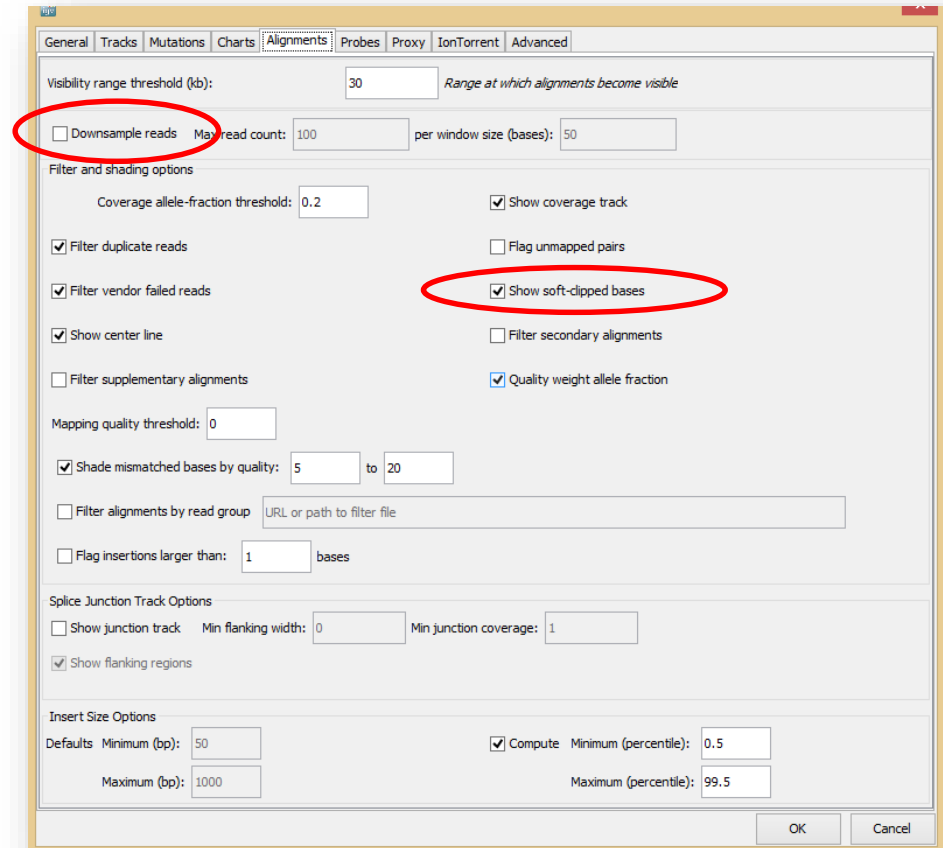


# Exemple 1 – Insertion d'un SVA

- Patient 1 :  $11/121 = 9\%$  reads
- Patient 2 :  $81/639 = 13\%$  reads
- Patient 3 :  $216/1108 = 19\%$  reads



<15% reads  
<10 reads



General Tracks Mutations Charts **Alignments** Probes Proxy IonTorrent Advanced

Visibility range threshold (kb): 30 *Range at which alignments become visible*

Downsample reads Max read count: 100 per window size (bases): 50

Filter and shading options

Coverage allele-fraction threshold: 0.2  Show coverage track

Filter duplicate reads  Flag unmapped pairs

Filter vendor failed reads  Show soft-clipped bases

Show center line  Filter secondary alignments

Filter supplementary alignments  Quality weight allele fraction

Mapping quality threshold: 0

Shade mismatched bases by quality: 5 to 20

Filter alignments by read group URL or path to filter file

Flag insertions larger than: 1 bases

Splice Junction Track Options

Show junction track Min flanking width: 0 Min junction coverage: 1

Show flanking regions

Insert Size Options

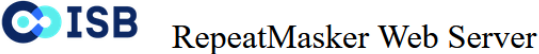
Defaults Minimum (bp): 50  Compute Minimum (percentile): 0.5

Maximum (bp): 1000 Maximum (percentile): 99.5

OK Cancel

# Exemple 1 – Insertion d'un SVA

<http://www.repeatmasker.org>



RepeatMasker screens DNA sequences in FASTA format against the Repbase-derived RepeatMasker library of repetitive elements or against the Dfam database and returns a masked query sequence ready for database searches. RepeatMasker also generates a table annotating the masked regions.

Reference: A.F.A. Smit, R. Hubley & P. Green, unpublished data. Current Version: open-4.0.9 (Dfam: 3.0 only\*)

**PLEASE READ FIRST:** As at this time we can only offer models for many transposable RepeatMasker locally. We are

[Check Current Queue Status](#)

Basic Options

Parcourir...

or

Sequence:

Matrix = 20p43g.matrix  
 Transitions / transversions = 0.0 (0 / 0)  
 Gap\_init rate = 0.00 (0 / 79), avg. gap s

Search Engine:

rmbblast  hmmer  cross\_match  
 abblast

Speed/Sensitivity:

rush  quick  default  slow

DNA source:

Human

Return Format:

html  tar file

Return Method:

html  email

Effacer Submit Sequence

score	% div.	% del.	% ins.	query sequence	begin	end	position in query- (left)	C matching + repeat	repeat class/family	-position in repeat- (left)	end	begin (left)	linkage id/graphic
721	0.0	0.0	0.0	UnnamedSequence	21	100	(0)	+ SVA_F	Retroposon/SVA	312	391	(984)	1
ANNOTATION EVIDENCE:													
721	0.00	0.00	0.00	UnnamedSequence	21	100	0	+ SVA_F	Retroposon/SVA	312	391	984	
721	0.00	0.00	0.00	UnnamedSequence	21	100	(0)	SVA_F#Retroposon/SVA	312 391 (984)	m_bis502i0			

UnnamedSequen 21 ACTCAGTGCTCAATGGTGCCAGGCTGGAGTGCAGTGGCGTGAICTCGGC 70

SVA\_F#Retropo 312 ACTCAGTGCTCAATGGTGCCAGGCTGGAGTGCAGTGGCGTGAICTCGGC 361

UnnamedSequen 71 TCACTACAACCTACACCTCCCAGCGCCTG 100

SVA\_F#Retropo 362 TCACTACAACCTACACCTCCCAGCGCCTG 100

**Class:** Retroposon

**Family:** SVA

**Length:** 1375 bp

**Source Species:** Homo\_sapiens

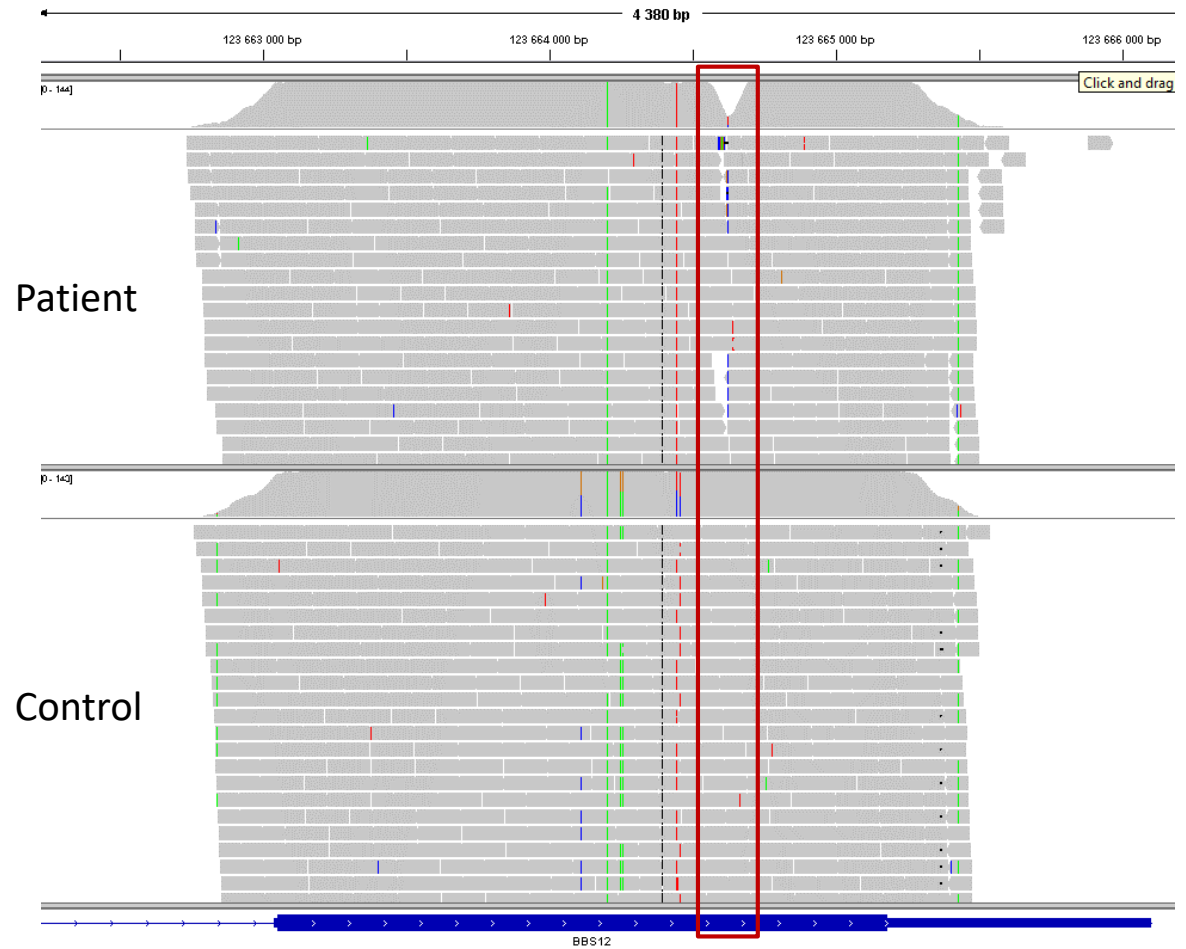
**Clades:** Homo\_sapiens

**Notes:** SVA subfamily consensus sequences from Wang et al. J. Mol. Biol. (2005) 354, 994-1007.

**Rebase Equivalent:** Not Defined

## Exemple 2 – Insertion homozygote d'un *Alu*

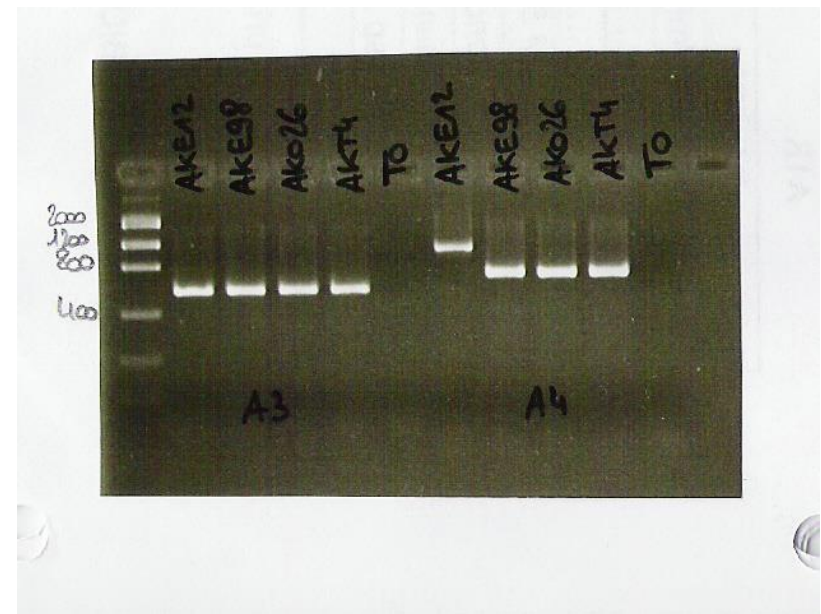
- 2010
- NGS ciblé
- Single read (72 pb)
- *BBS12*





## Exemple 2 – Insertion homozygote d'un *Alu*

AluYa5 (repeatmasker)

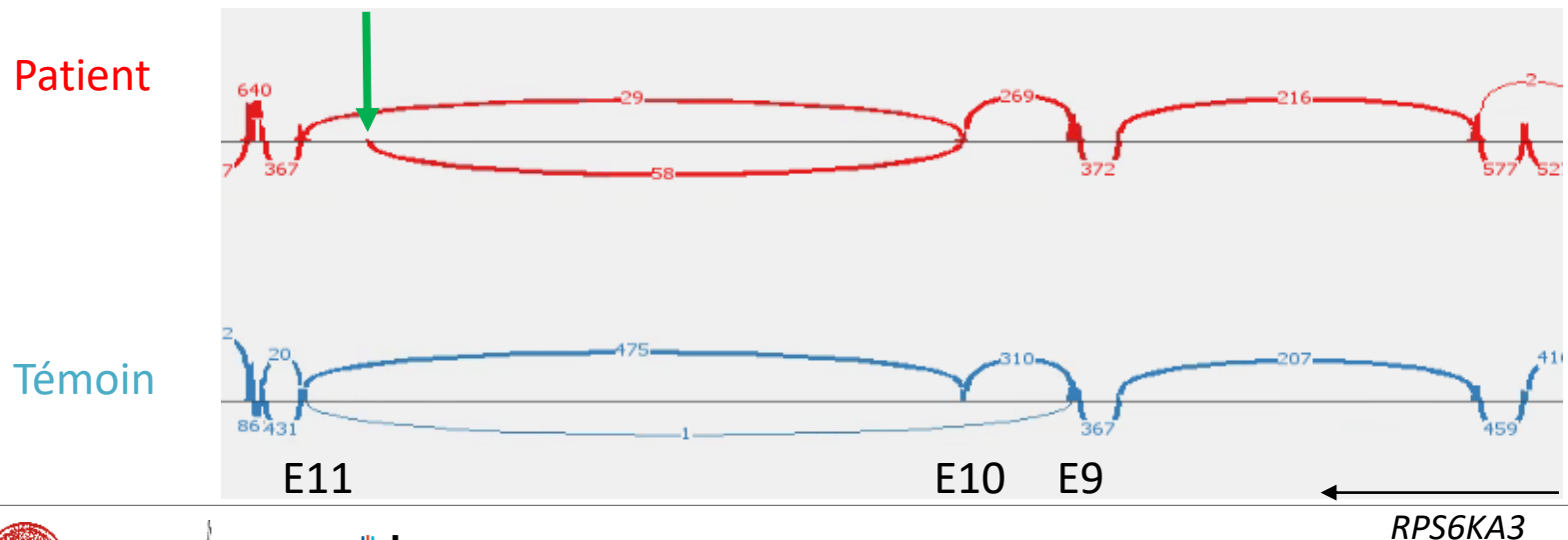


## Exemple 3 – Insertion hémizygote Line 1 *RPS6KA3*

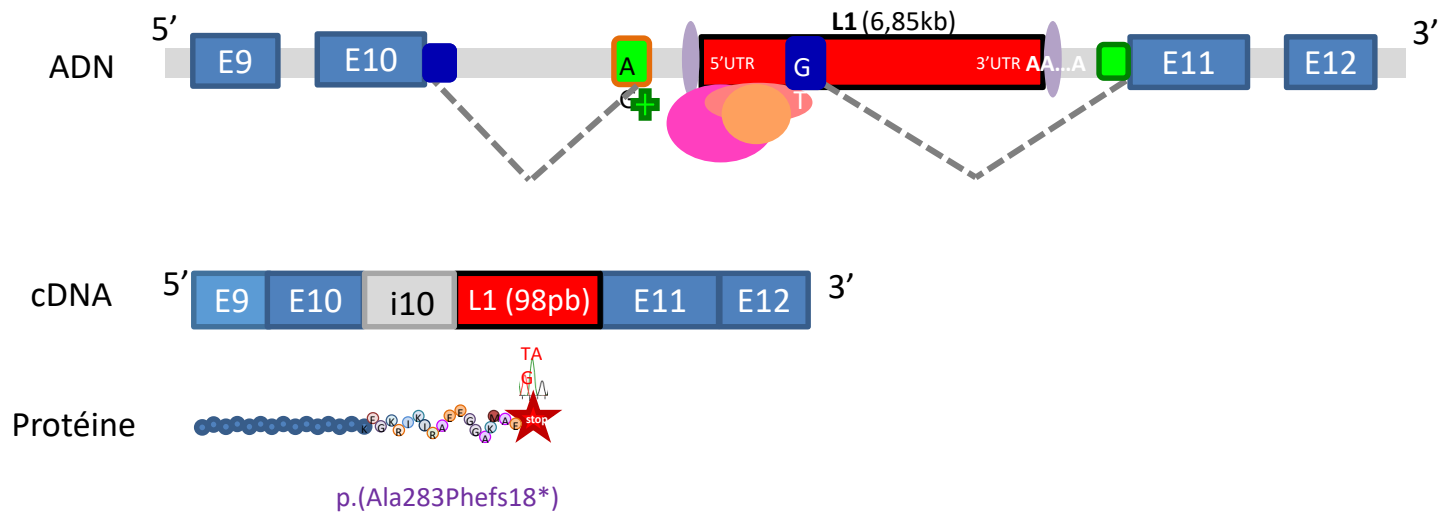
Margaux Biehler  
Bénédicte Gérard

- Diagnostic moléculaire d'un syndrome de Coffin-Lowry :
  - Panel, WGS et RNASeq
- Exonisation partielle d'un L1 dans les ARN de *RPS6KA3*

Epissage anormal dans l'intron 10 de *RPS6KA3* (-877)



## Exemple 3 – Insertion hémizygote Line 1 *RPS6KA3*



- Prédiction : apparition d'un codon stop prématuré et protéine RSK2 tronquée

# Limites

# Exemple 4: Faux positifs de MEI



# Exemple 4: Faux positifs de MEI - délétion homozygote



# Exemple 5: Les rétrocopies

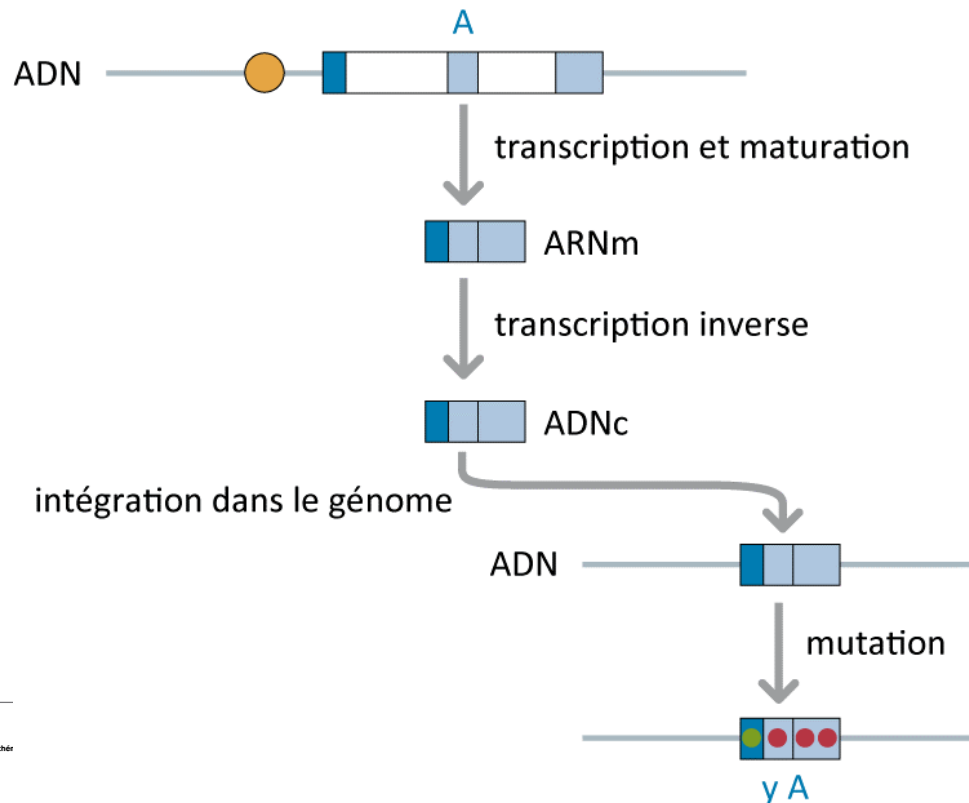
# Pseudogène

- **Définition:**
  - Copie généralement défectueuse d'un gène fonctionnel avec lequel le pseudogène partage une homologie de séquence significative.
- Concerne 10% des gènes codants, avec un nombre plus important pour des gènes activement transcrits
- >14000 décrits pour 3 catégories principales:
  - Pseudogène **processé** (~72%, ~10 000 copies)
  - Pseudogène **non processé** (~23%, ~3 500 copies)
  - Pseudogène **unitaire** (~1,4%, ~200 copies)



# Origine des pseudogènes

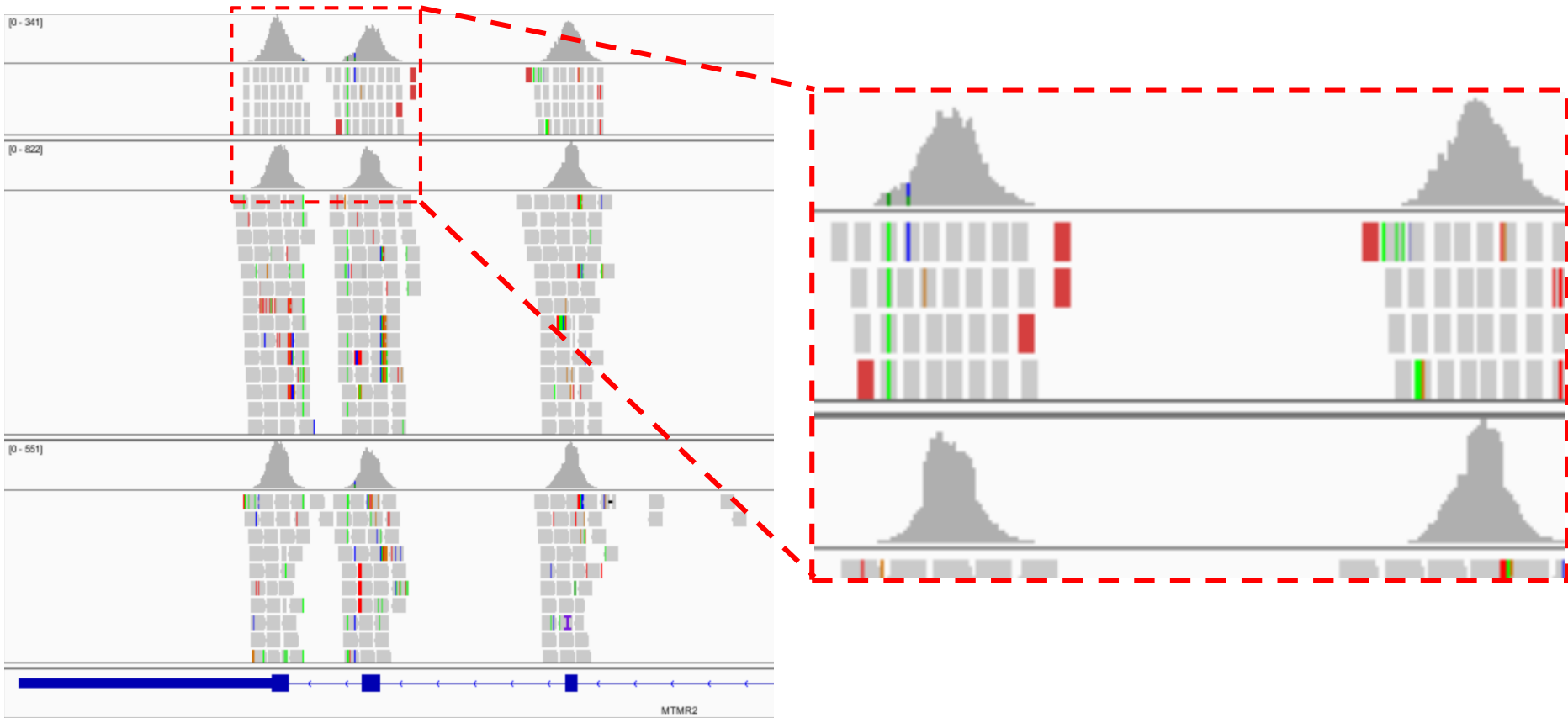
- Pseudogène dit **processé**
  - Mécanisme lié à la rétrotransposition d'un ARNm
  - Perte de la structure intron/exon (le pseudogène ne contient que de l'exon)
  - Perte de la région de régulation



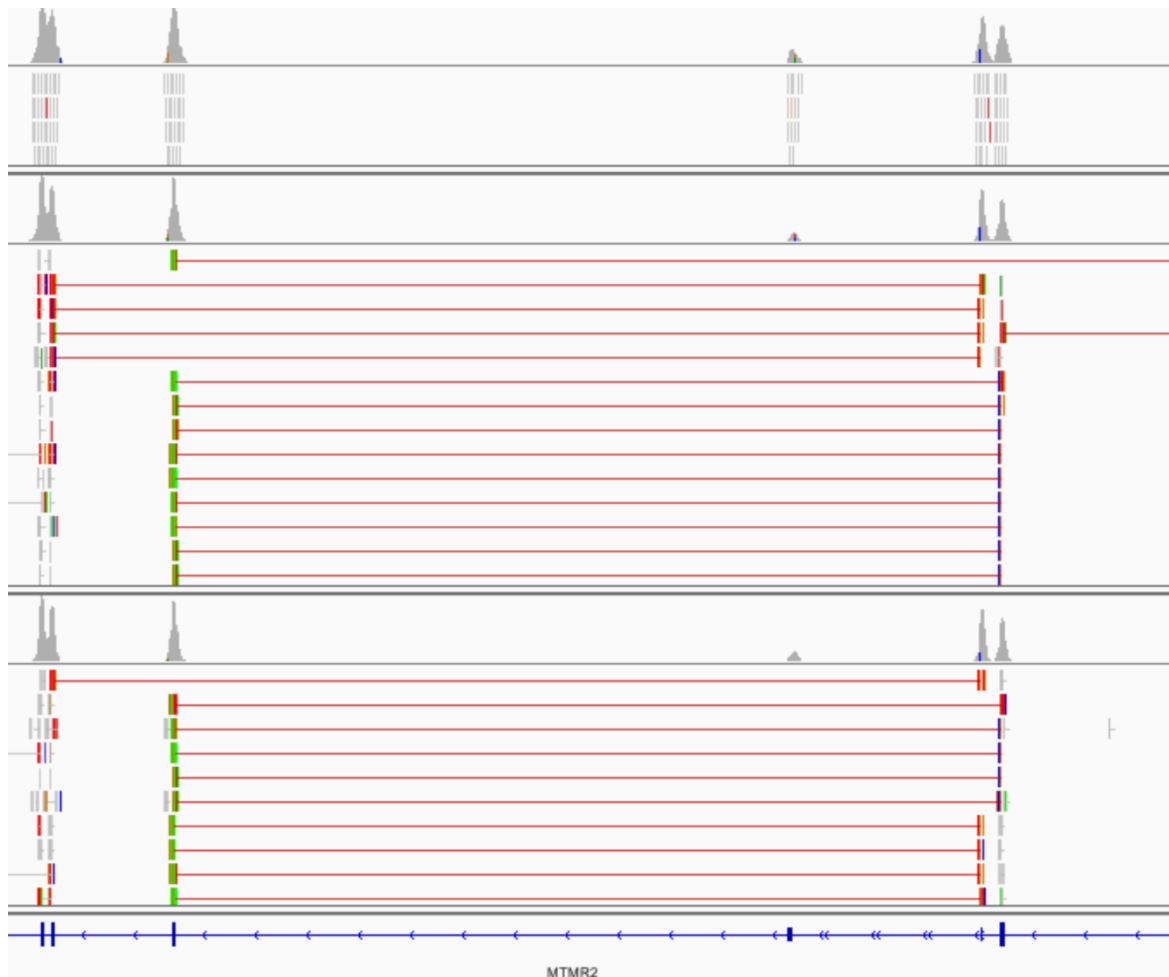
- promoteur
- variation bénigne
- variation inactivante

# Exemple 5: Les rétrocopies

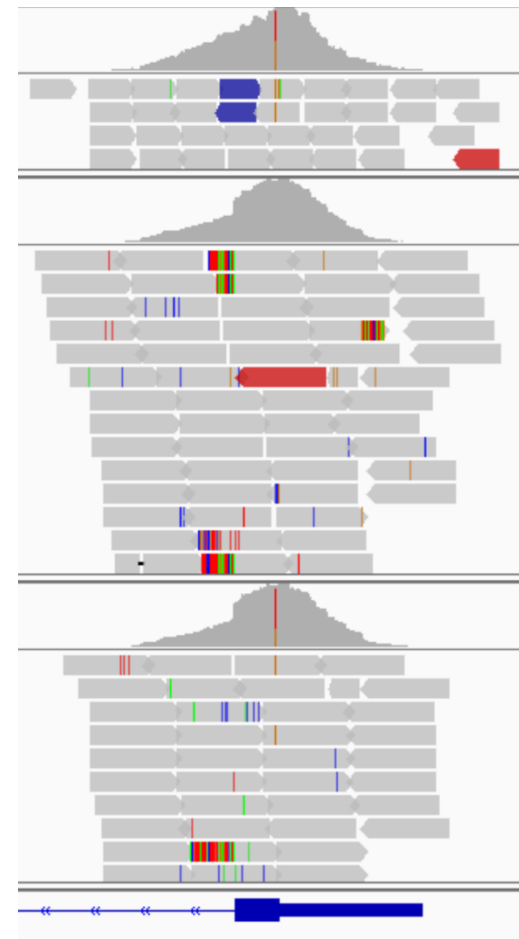
MTMR2



## Exemple 5: Les rétrocopies



Paired end  
Tri et coloration par « Insert size »

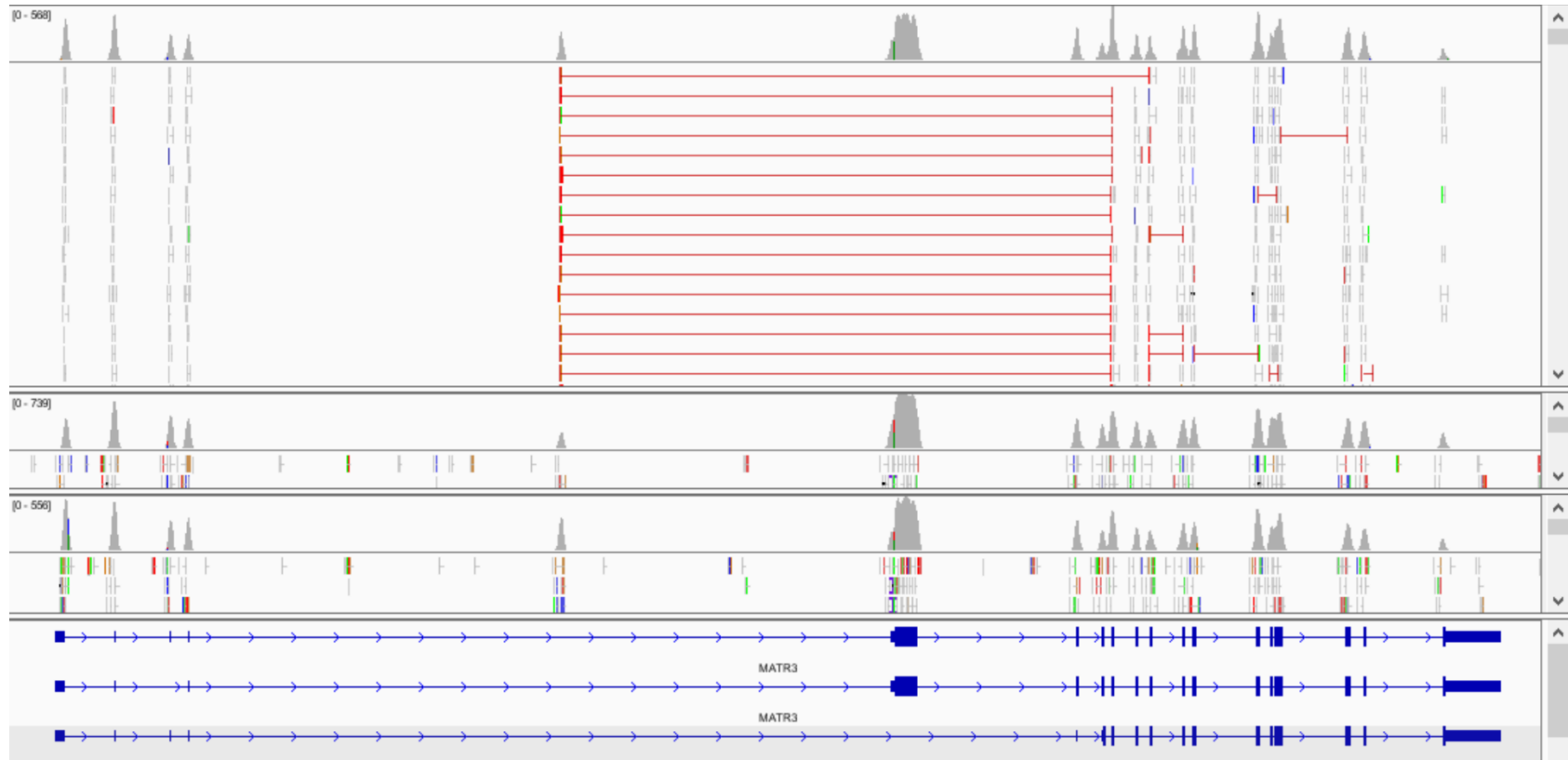


Ratio profondeur de lecture des variants biaisé

# Exemple 5: Les rétrocopies

*MATR3*

Isoforme spécifique



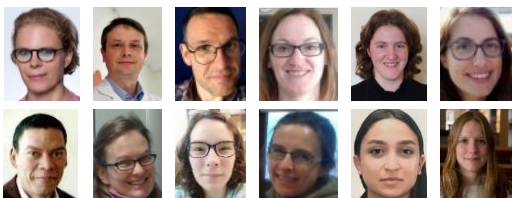
## Conclusion - perspectives

- Eléments mobiles du génome est un des mécanismes rares impliqué dans le BBS
- Seconde variation la plus fréquente du gène *BBS1*
- Contribution mineure dans le diagnostique
- Intégration dans le pipeline bioinformatique (CHU/Inserm)
  - Scramble et Mobster
- Difficulté à détecter de manière robuste
- Faux positifs et négatifs

## Laboratoire de Génétique Médicale (UMR\_S 1112)



**Hélène Dollfus** Adella Karam Salima El Chehadeh  
 Nicolas Le May Gaëlle Hayot **Corinne Stoetzel**  
**Véronique Geoffroy** Laurie Ruch **Clarisse Delvallée**  
 Sophie Scheidecker Amandine Girard **Samuel Nicaise**  
 Louise Porter Vincent Marion



## Chicago Lurie Children Hospital



Erica E. Davis, Nicholas Katsanis,  
Carmen C. Leitch, Georgios Kellaris

## Uppsala University



Joakim Klar, Niklas Dahl



## Collaborateurs cliniciens et biologistes



Florence Demurger, Koenraad Devriendt,  
Michèle Mathieu-Dramard, Christine Poitou-  
Bernert, Boris Keren, Christel Depienne,  
Sylvie Odent

## IGBMC - Génétique



**Jean-Louis Mandel**  
Amélie Piton

## CSTB - Systèmes Complexes et Bioinformatique Translationnelle



## Eye Diseases (ERN-EYE)



**Association Bardet-Biedl France**  
**Tous les patients qui nous font confiance**

## Hôpitaux Universitaires de Strasbourg



Laboratoires Diagnostic Génétique  
**Manuela Antin, Anne-Sophie Leuvrey, Elsa Nourisson**  
CARGO (Valérie Pelletier, Nathalie Goetz)



Service Génétique (Elise Schaefer, Salima El Chehadeh)  
Pédiatrie (Vincent Laugel)  
Bioinformatique (Antony & Co)



## CNRGH

Anne Boland,  
Steven McGinn  
Jean-François Deleuze & Co

## France GENREF

Emmanuelle Génin  
Jean-François Deleuze  
Richard Redon

