

Shallow-Genome Sequencing :

La nouvelle ACPA ?

P. Chambon – Service de Génétique, CHU Rouen

N. Chatron – Service de Génétique, Hospices Civils de Lyon

26/11/2025 – Journée AchroPuce



Flux CGH-array



- De l'ADN à l'analyse : 3-4 jours de temps technique
- Jusqu'à 24 échantillons en parallèle (multiple de 8)
- Matériel spécifique :
 - Four à hybridation : suivi métrologique compliqué
 - Caisson ozone : plus de maintenance et pas de remplacement
 - Scanner CGH-array : seul usage
- Logiciels fournisseurs :
 - Analyse secondaire sur place : 1 an pour réaliser une montée de version par la DSN lyonnaise
 - Analyse tertiaire / bases de données cloud : interruption de service en septembre 2025

Points forts

Robuste

Logiciels

Visualisation

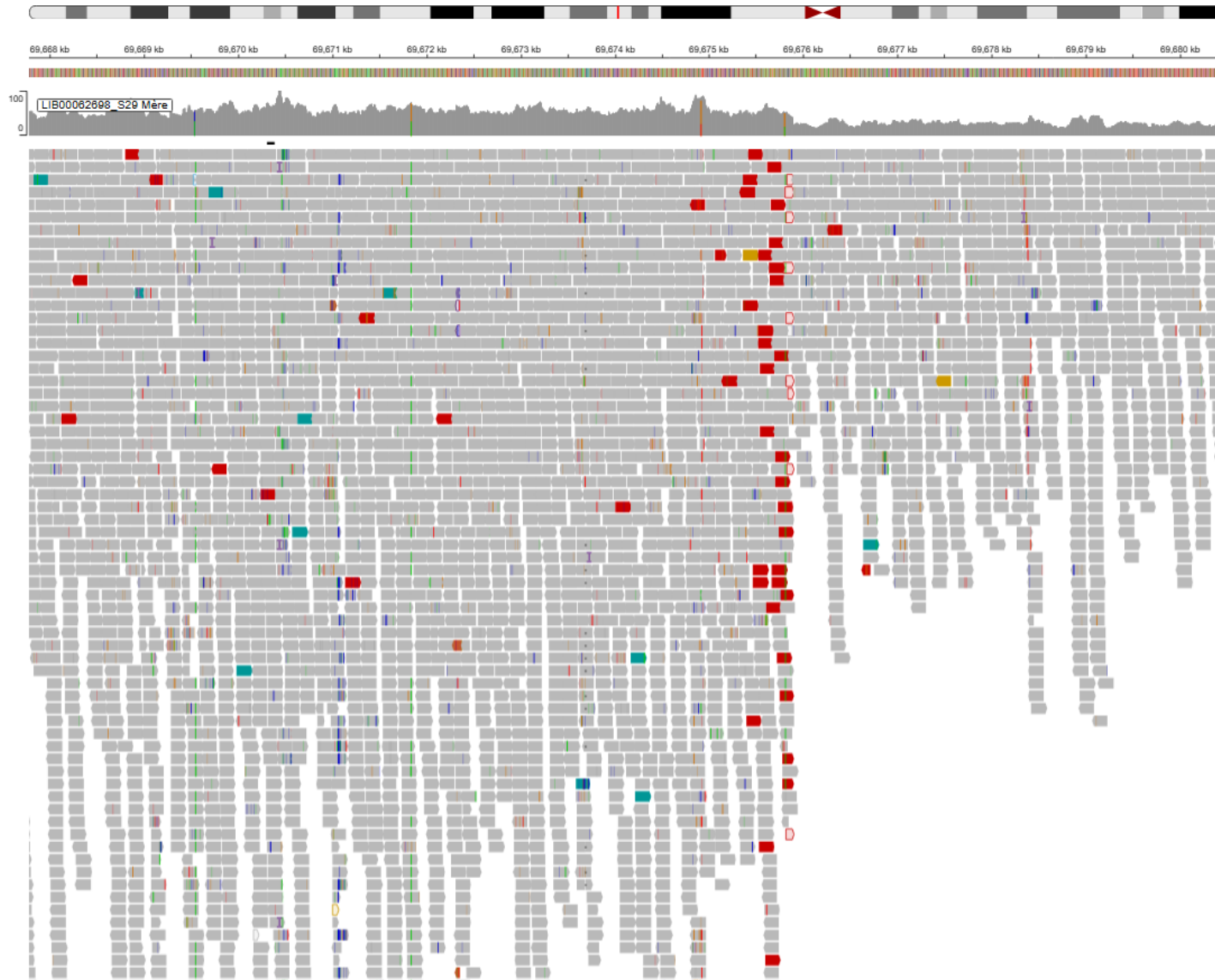
20 ans de recul

Changement technique ?

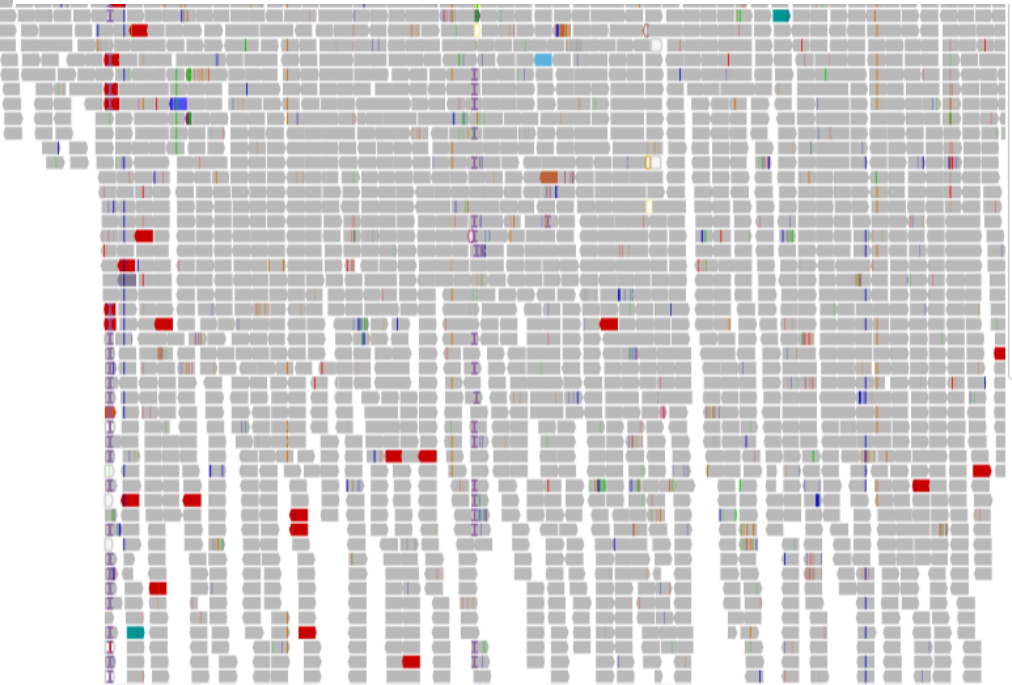
- Maintenir résolution
- Détection mosaïque
- Equipements déjà disponibles
- Internalisation bioinformatique
- Maîtrise temps technique
- Coût
- DPN : 400ng ADN max
- DPN : délai technique court



GENOME SEQUENCING « 30X »



- Maintenir résolution
- Détection mosaïque
- Equipements déjà disponibles
- Internalisation bioinformatique
- Quantité d'ADN max : 400ng
- Maîtrise temps technique
- Coût




Littérature

Received: 5 June 2023 | Revised: 20 September 2023 | Accepted: 22 September 2023
DOI: 10.1111/ahg.12532

ORIGINAL ARTICLE 

Low-pass whole genome sequencing is a reliable and cost-effective approach for copy number variant analysis in the clinical setting

Patricia C. Mazzonetto^{1,2} | Darine Villela²  | Silvia Souza da Costa¹ |

Genet Med. 2020 March ; 22(3): 500–510. doi:10.1038/s41436-019-0634-7.

Low-pass genome sequencing versus chromosomal microarray analysis: implementation in prenatal diagnosis

Huilin Wang, Ph.D.^{1,2,†}, Zirui Dong, Ph.D.^{2,3,†}, Rui Zhang, Ph.D.^{1,†}, Matthew Hoi Kin Chau, M.Sc.^{2,3,†}, Zhenjun Yang, M.Eng.², Kathy Yin Ching Tsang, M.Phil.², Hoi Kin Wong, M.Phil.²,

Copy-number variation

Original research

Validation and depth evaluation of low-pass genome sequencing in prenatal diagnosis using 387 amniotic fluid samples


Yeqing Qian,^{1,2} Yan Sun,³ Xueqin Guo,⁴ Lijie Song,^{5,6,7} Yixi Sun,^{1,2} Xiaoyang Gao,^{1,2}

Low-Pass Genome Sequencing 


Validation and Diagnostic Utility from 409 Clinical Cases of Low-Pass Genome Sequencing for the Detection of Copy Number Variants to Replace Constitutional Microarray

Alka Chaubey, Suresh Shenoy, Abhinav Mathur, Zeqiang Ma, C. Alexander Valencia, Babi R. Reddy Nallamilli, Edward Szekeres, Jr, Leah Stansberry, Ruby Liu, and Madhuri R. Hegde

Received: 8 April 2024 | Revised: 10 June 2024 | Accepted: 14 June 2024
DOI: 10.1002/ajmg.a.63802

ORIGINAL ARTICLE 

Low-pass whole genome sequencing as a cost-effective alternative to chromosomal microarray analysis for low- and middle-income countries

Patricia C. Mazzonetto^{1,2} | Darine Villela²  | Ana C. V. Krepischi¹ |

Human Genetics (2020) 139:1403–1415
<https://doi.org/10.1007/s00439-020-02185-9>

ORIGINAL INVESTIGATION 

Low-pass genome sequencing: a validated method in clinical cytogenetics

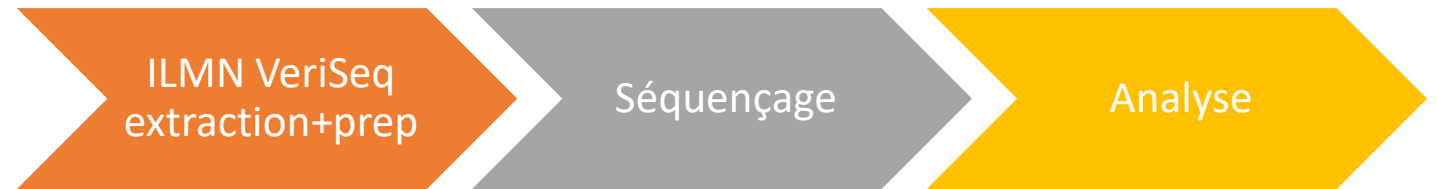
Matthew Hoi Kin Chau^{1,2,3}  · Huilin Wang⁴ · Yunli Lai^{5,6} · Yanyan Zhang^{1,2} · Fuben Xu^{5,6} · Yanqing Tang^{5,6} ·

Shallow genome sequencing ?

KEZAKO

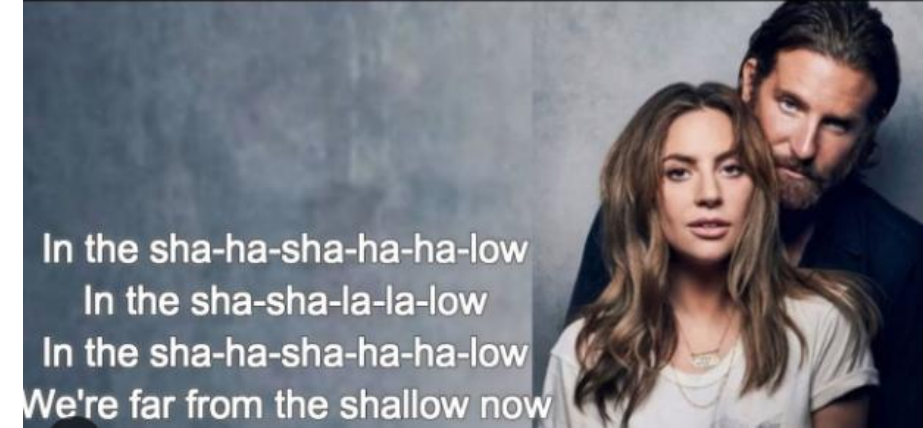


- = DPNI VeriSeq® Illumina
- Performances attendues « bonnes » car ADNlc = mosaïque
- Rendu en 1 jour
- « Routine » cytogénétique







Shallow genome sequencing

= low-pass genome sequencing



- Séquencer moins profond
- Pipelines dédiés pour rechercher des variations entre différentes fenêtres du génome

		Library	Read length	Nb of reads	Binsize	
Ghent		ILMN PCR-free	2*100	10-50M	15-50kb	
Gosselies		NextFlex Bioo	2*100	20M	15 kb	Mix DPNI+shallow
Anvers				6-18M		
Montreal		ILMN PCR-free		Génome 8x		

DNA PCR Free kit, Illumina : préparation des librairies pour shallow sequencing

0 – Pré requis

300 ng mini ADN dans 25µL RSB

Dosage Qubit BR

1 - Fragmentation ADN par transposase = tagmentation

Fragmentation par transposase couplée aux adaptateurs / saturation des billes

<https://emea.illumina.com/techniques/sequencing/ngs-library-prep/tagmentation.html>

2 – Purification

On ne garde que les fragments fixés aux billes

3 – Ligation des index

Index pour identifier le patient + adaptateurs p5 et p7 attachement à la flow cell

4 – Double purification

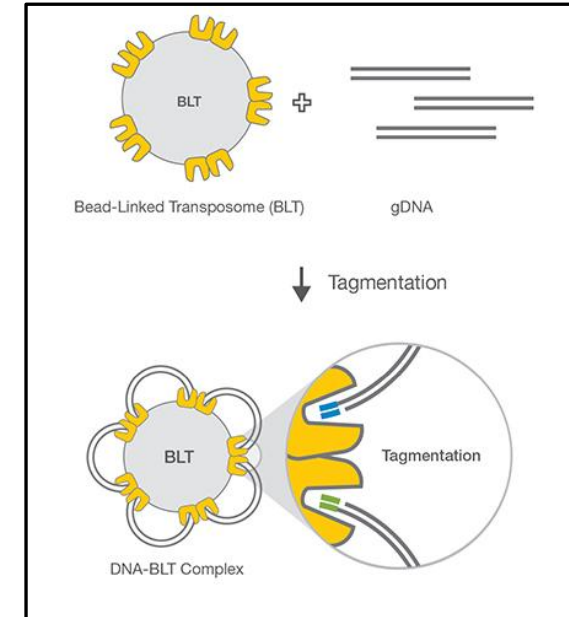
On ne garde que les fragments utiles (taille correcte et adaptateurs de chaque côté)

5 – Pool

Melange vol/vol car autonormalisation

Dosage Qubit ssDNA

6 – Séquençage



Expérience lyonnaise

- Objectif principal : **DPN**
- 16-24 échantillons par semaine.
- ADN extraits en semaine 0 – résultats rendus le jeudi de la semaine 1

- Contraintes locales :
 - Plateforme NGS partagé = réduire le temps d'occupation séquenceur
 - Réduire les coûts de fonctionnement pour permettre investissement


Lyon : choix techniques

- Préparation librairie : ILMN PCR-free
 - Utilisé par les centres belges
 - Utilisé par Auragen [proximité – échanges techniques]
 - Automatisable – 2h30 /24 échantillons
 - Tagmentation sans dosage pour normaliser à partir de 300ng
- Séquençage : 2x150 – 16 échantillons sur NS500 HighOutput
 - Maximaliste
- Bioinformatique : WISECONDORX
 - Expérience lointaine DPNI

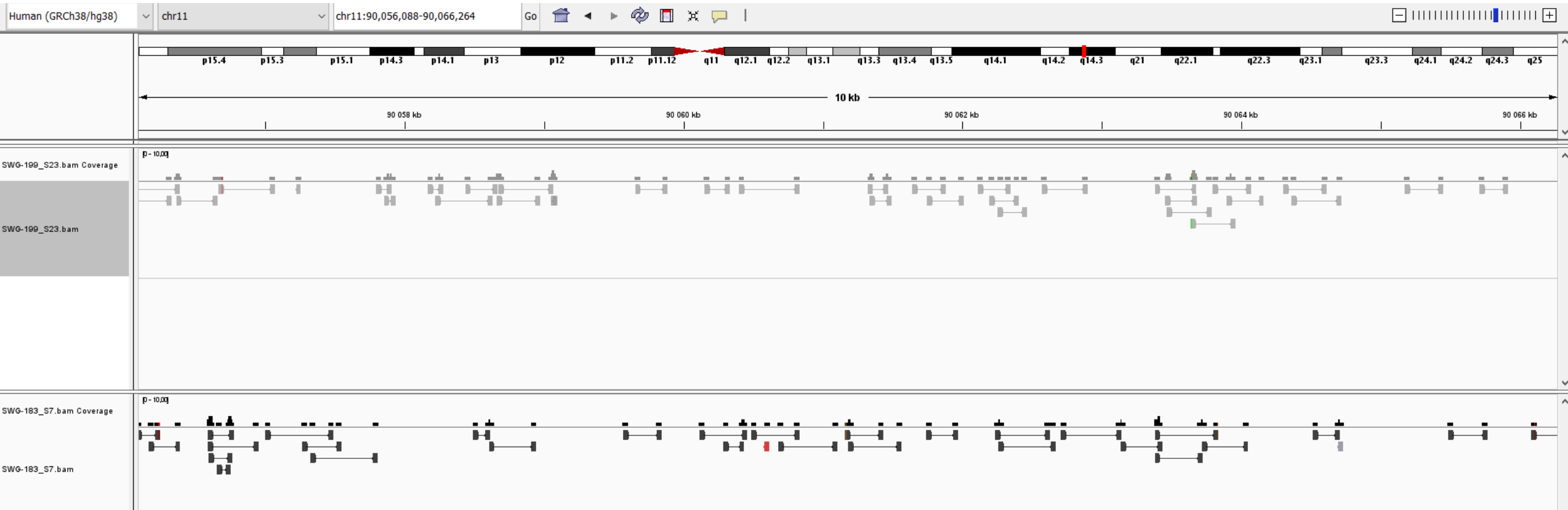
Lyon : CNV testés

CNV < 200 kb	200kb < CNV < 1Mb	CNV > 1Mb
12	26	32

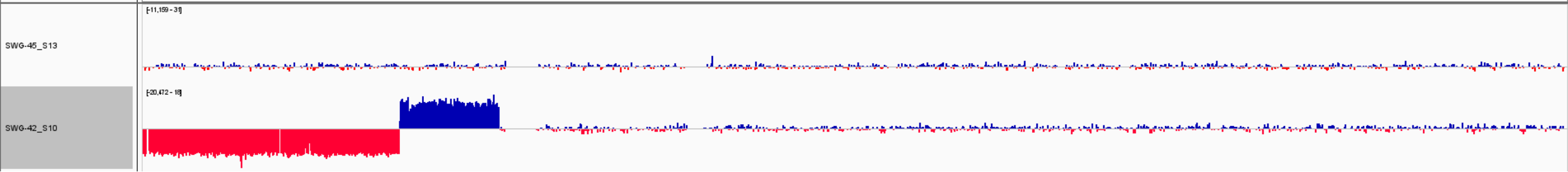
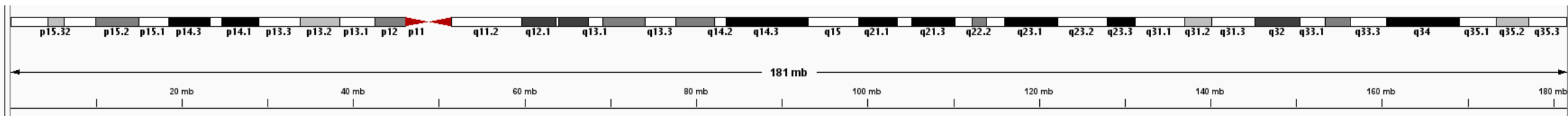
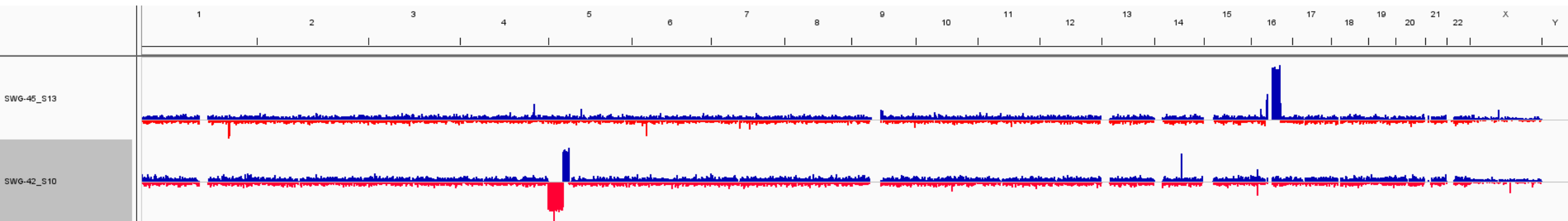
- CNV récurrents : 15q11.2, 15q13.3, 16p13.1, 16p11.2, 22q11.2
- Anneau de l’X en mosaïque à 40%
- Trisomie 8 en mosaïque à 15% ; à 27%
- Trisomie 15 en mosaïque à 9%
- Trisomie 14 en mosaïque à 10% + XXY en mosaïque à 8%
- Délétion 13qter en mosaïque à 20% (81Mb)

 Manque
délétion(s)
homozygote(s)

Lyon : signal brut = .bam



- Profondeur moyenne $\approx 0,2x \rightarrow 1\text{Go}$ [vs $0,65\text{Go}$.tiff CGH-array]
- Environ 30 paires de *reads* par bin de 10kb



Lyon : premiers résultats

OBSERVATIONS

Référence :

- Temps de calcul à **150 bp** :
> **3 jours** vs **50 bp** : < **6h**
 - Beaucoup de ressources nécessaires
 - **Même résultats à 35 bp**
- ➔ **35 bp validé (temps + coût)**

NOUVEAUX TESTS

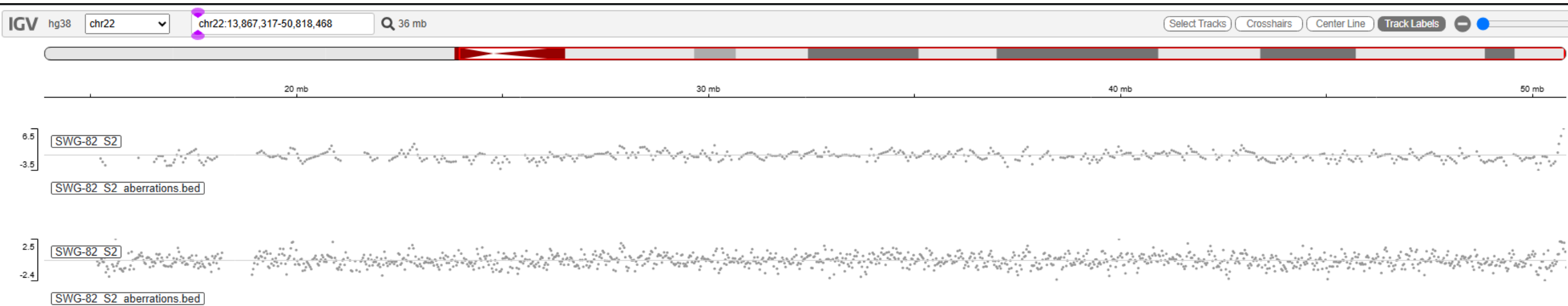
À 35 bp

756 combinaisons de paramètres testées :

- Nombre de reads
 - Sans downsampling
 - 1M 2M 5M 10M 20M
- Taille de bins
 - 5kb 10kb 15kb 25kb 50kb 100kb
- Zscore
 - 0,5 à 20

Lyon : premiers résultats

Faux négatifs : 1/70 !



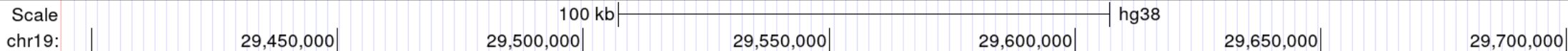
- Del 22qter : 127kb en ACPA

Lyon : nouveaux tests

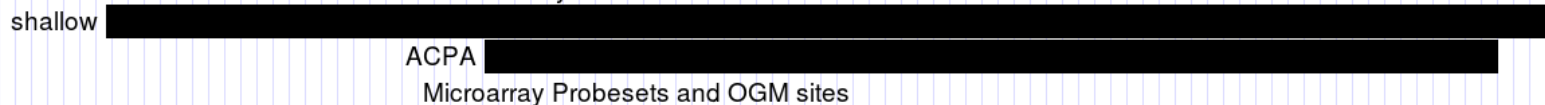
- 96 échantillons DPN consécutifs de 2024
= test en situation « réelle »
- 5/96 < 10M de reads → à invalider
→ Adaptation protocole : 350ng vs 300ng (seuil minimum ILMN pour éviter les normalisations)
- 1,6 CNV en moyenne [0-8] par analyse dont 0,3 jugé artefactuel (centromères, dispersion télomérique)
→ Un CNV classe 2 de 110kb détecté mais filtré.

Bornage CNVs

chr19 (q12) 19p13.3 19p13.2 13.12 19p13.11 19p12 19q12 q13.11 q13.12 19q13.2 q13.32 q13.33 13.41 q13.42 q13.43

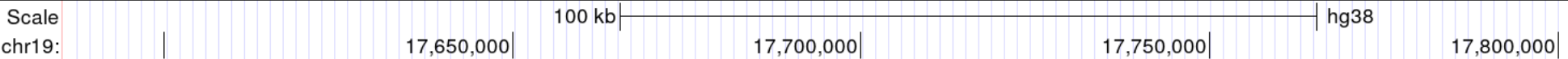


210 kb vs 147kb

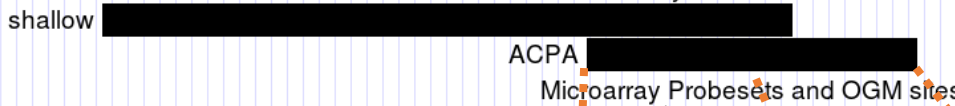


atiSure Cyto CGH 8x60

chr19 (p13.11) 19p13.3 19p13.2 13.12 19p13.11 19p12 19q12 q13.11 q13.12 19q13.2 q13.32 q13.33 13.41 q13.42 q13.43



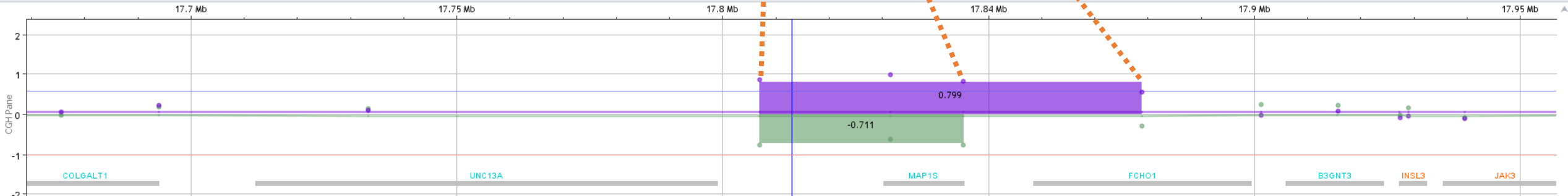
80 kb vs 38kb



atiSure Cyto CGH 8x60

p13.3 p13.2 p13.12 p13.11 p13.10 p13.9 p13.8 p13.7 p13.6 p13.5 p13.4 p13.3 p13.2 p13.1 q13.11 q13.12 q13.2 q13.31 q13.32 q13.33 q13.41 q13.42 q13.43

Gene View (chr19:17669296-17956761, Size=287 Kb)



Lyon : ce qu'il reste à faire

- Mosaïque *in silico* : limite de détection
- Validation de méthode standard : reproductibilité, inter-opérateurs,...
- Extension de la base de référence n=? 200 ?
- Ajustement visualisation
- Annotation et **base de données CNV prospective**
- Tests DRAGEN v4.x
- Appel d'offre réactifs en cours
- Formation techniciens

Expérience rouennaise

- Objectif principal : DPN et postnatal => 1 run hebdomadaire impératif
- 16-20 échantillons par semaine
- Plateforme de CGH vieillissante
- Opportunité locale :
 - Bonne expérience du DPNI whole genome
 - Bascule Nextseq > Novaseq => forte diminution des coûts de séquençage
 - Créneaux novaseq disponibles, runs partagés
 - Augmentation de l'activité urgente (exome anténatal, somatique, onco thérapeutique)

Rouen : choix techniques

- Préparation de librairie : ILMN PCR-free
 - Conseillé par Illumina
 - Similaire au DPNI
 - Input : \approx 300 ng pour autonormalisation
- Séquençage : Novaseq 6000
 - Flowcell SP : utilisation d'1 ou 2 lanes (16 ou 32 échantillons)
 - Objectif : utilisation de la 2^e lane pour exome ou somatique
⇒ PE 2x150pb (NB : DPNI = PE 2x36pb)
- Bioinformatique :
 - Serveur DRAGEN v4.3
 - WISECONDORX

Bioinformatique :



KESAKO DRAGEN ?

- DRAGEN Demultiplexing
- DRAGEN ORA Compression
- DRAGEN Map + Align
- **DRAGEN Germline**
- DRAGEN Somatic

DNA Mapping

Read Trimming

DRAGEN FASTQC

Sorting and Duplicate Marking

Small Variant Calling

Copy Number Variant Calling

Repeat Expansion Detection

Targeted Caller

Structural Variant Calling

VNTR Calling

Population Genotyping

Filter Duplicate Variants

Ploidy Calling

Multi Caller

QC Metrics Reporting

JSON Metrics Reporting

HLA Typing

Biomarkers

Downsampling

Unique Molecular Identifiers

Indel Re-aligner (Beta)

Star Allele Caller

High Coverage Analysis

CheckFingerprint

Population Haplotyping (Beta)

DUX4 Rearrangement Caller

Intérêts de DRAGEN

Architecture

FPGA dédiée au traitement génomique

Performance

Réduction des temps d'analyse

Facilité d'usage

Paramétrage minimal

Modalités d'implémentation DRAGEN

🖥️ Serveur local

Déploiement on-premises, contrôle total, données en interne

☁️ Cloud

Élasticité des ressources, collaboration facilitée, maintenance réduite

🧬 Embarqué sur séquenceur

Intégration directe, flux automatisé des données NGS vers analyses

Bioinformatique :

DRAGEN v4.3 - Analyse des CNV :



Principe

- **Définition des zones mappables**
- **Alignement** des reads
- **Binning** : découpage du génome en petits intervalles
- **Compte** du nb de reads dans chaque intervalle
- **Correction biais de GC**
- **Normalisation** :
 - Auto-normalization (intra run)
 - ou Panel de référence (PoN)

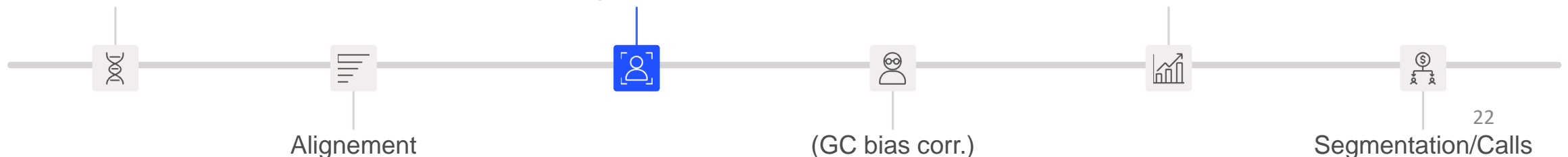
Paramètres clés

i WGS : Profondeur → Résolution	Valeurs conseillées
• --cnv-interval-width	≥30X → 1 kb
	10× → 5 kb
	5X → 10 kb
	1X → ?
• --cnv-filter-length :	200 Kb
• --cnv-segment-max-gap :	200 Kb
• --cnv-merge-distance	1 500 Kb

FASTQ/BAM/CRAM

Target Counts

Normalisation (Self ou PoN)



Rouen : premiers tests

Données

- **76 échantillons**
 - 16 anomalies
 - 60 normaux => Référence
- **20M paires de reads**

Objectifs

- **Paramétrage de Dragen**
- **Estimation de la résolution**
- **Outils de visualisation**
- **Durée du traitement des données**

- **3 Aneuploïdies et grands CNVs**

- **5 CNV de taille intermédiaire, PIEV : 217Kb à 2,57Mb**

- **8 petites délétions et duplications :**

Taille décroissante de 500 à 50Kb (500, 200, 100 et 50 Kb)

1 délétion, 1 duplication par taille

Rouen : premiers résultats

- **DRAGEN v4.3**

- **Utilisation d'une référence indispensable : \approx 70 CNV vs 16 CNV**
- **Optimisation des paramètres DRAGEN**

	initiaux	retenus
-cnv-interval-width :	50 000	10 000
-cnv-filter-length :	200 000	30 000
-cnv-segment-max-gap :	200 000	
-cnv-merge-distance :	1 500 000	

=> Détection des 16 CNV d'intérêt

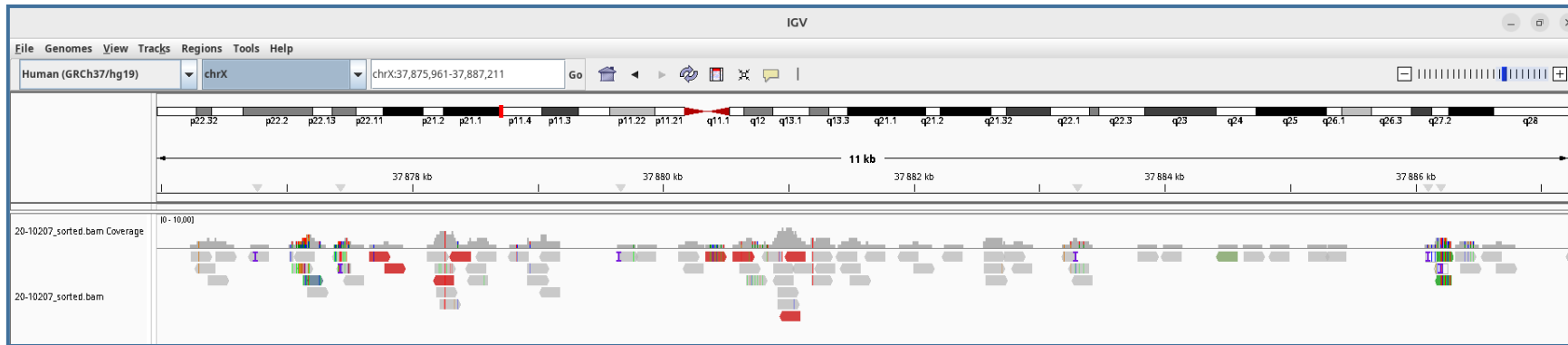
- **Renseignement du sexe patient nécessaire pour l'analyse des gonosomes**
- **Génération de fichiers gff ou xml pour visualisation des CNV dans IGV**
- **Temps de traitement DRAGEN : \approx 1 min par patient**

Rouen : premiers résultats

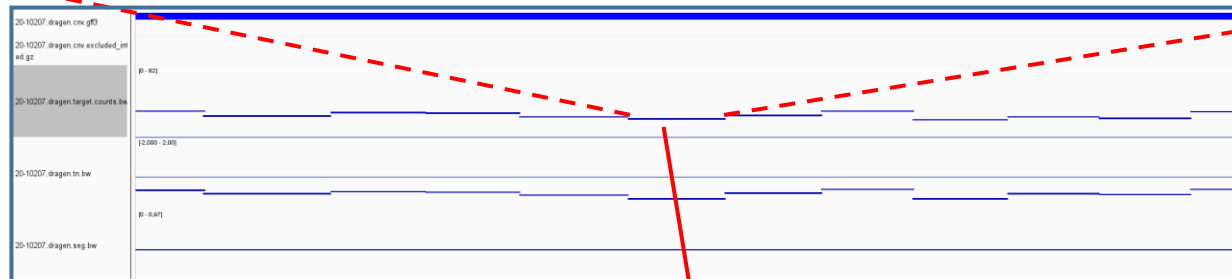
		Concordance	
Grands CNV	47,XXY	✓	
	46,X,rec(X)inv(X)(p11.23q21.1) => del (X)(pterp11.23) 47Mb + dup(X)(q21.1qter) 78Mb		
	46,XX,der(17)t(14;17)(q32.1;p13.3)dp at => dup(14)(q32.1qter) 14 Mb + del(17)(p13.3) 1,2 Mb		
PIEV, CNV intermédiaires	Dup 22q11.2 de 2,5Mb	✓	
	Del15q13.3 de 1,5 Mb		
	Del 1q21 de 1,3 Mb		
	Del 16p11.2 proximale de 547 Kb		
	Del 16p11.2 distale de 186 Kb		
Petits CNV	Del 15q15.3 de 500 Kb	Dup 8q24.3 de 502 Kb	✓
	Del 16q22.2 de 200 Kb	Dup 13q12.13 de 200 Kb	
	Del 22q12.1 de 101 Kb	Dup 17p11.2 de 101 Kb	
	Del 9q21.13 de 50 Kb	Dup 8q11.21 de 51 Kb	

Rouen : premiers résultats

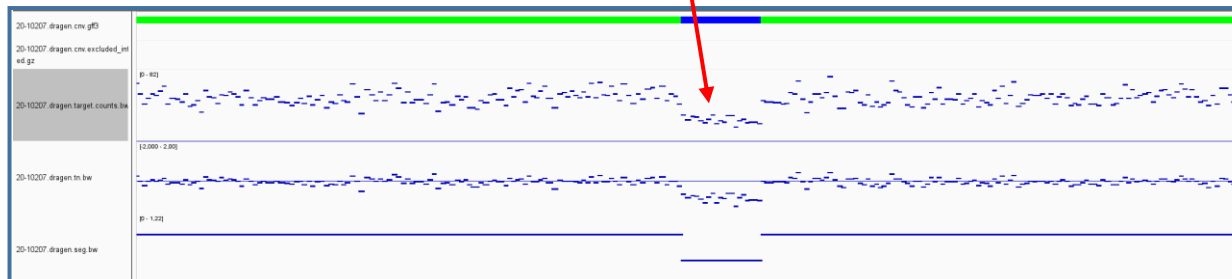
Visualisation IGV



80 reads / 10 Kb



17 M paires reads / genome
33M reads uniques



del 16q22.2
200 Kb

Rouen : premiers résultats

Visualisation IGV

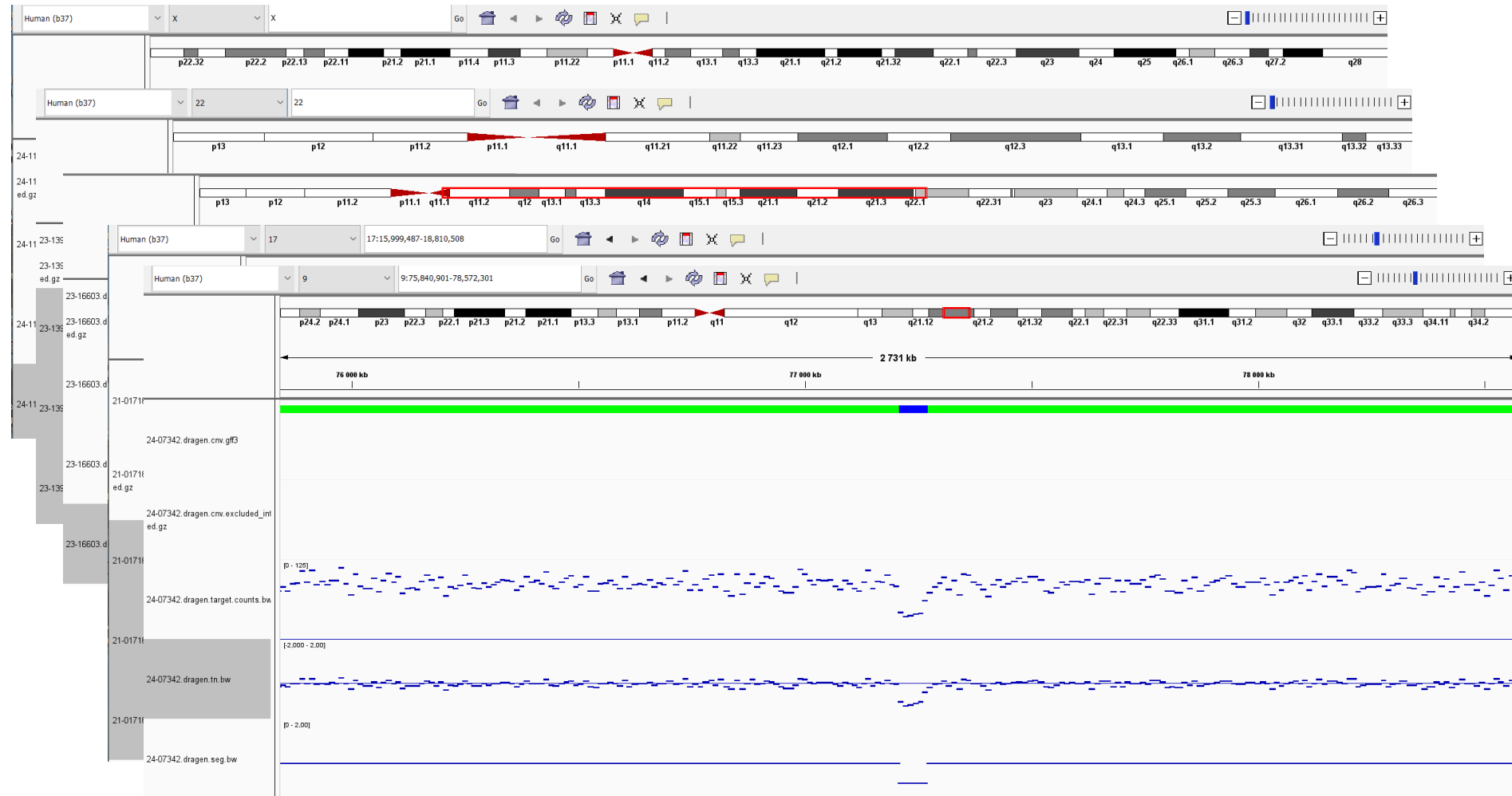
rec(X)inv(X)(p11.23q21.1)

dup 22q11.21 => 2,5 Mb

del 15q15.3 => 500 Kb

dup 17p11.2 => 100 Kb

del 9q21.1 => 50 Kb



Rouen : suite des tests

- **Intégration AnnotSV**
- **Test in silico**
 - **Performances pour les CNV en mosaïque**
 - Downsampling des données patients et référence pour simuler une mosaïque : 50% => 10%
 - DRAGEN : **pas de détection des CNV en mosaïque**
 - Test de DRAGEN 4.4
 - => ajout d'un pipeline basé sur la BAF : pas d'intérêt en shallow genome
 - => **version 4.5 optimisée Shallow Genome à venir**
 - Ajout de WisecondorX (ciblage des CNV >10 Mn en mosaïque)

Rouen : suite des tests

Caryotype / CGH / DPNI		Mosaïcisme				
CNV	Size (kb)	50,00 %	40,00 %	30,00 %	20,00 %	10,00 %
GRANDS CNV						
47,XXY		Trouvé	Trouvé	Trouvé	Trouvé	Trouvé
46,X,rec(X)inv(X)(p11.23q21.1)						
Dup Xq21.1	78 268	Trouvé	Trouvé	Trouvé	Trouvé	Trouvé
Del Xp11.23	47 679	Trouvé	Trouvé	Trouvé	Trouvé	Trouvé
46,XX,der(17)t(14;17)(q32.1;p13.3)dp17						
Del 17p13.3	1 240	Trouvé	X	X	X	X
Dup 14q32.1	17 549	Trouvé	Trouvé	X	X	X
PIEV						
Dup 22q11.21	2 569	X	X	X	X	X
Del 15q15.3	1 566	Trouvé	Trouvé	Trouvé	Trouvé	X
Del 1q21.1	1 317	Trouvé	Trouvé	Trouvé	X	X
Del 16p11.2 proximale	545	Trouvé	X	X	X	X
Del 16p11.2 distale	217	X	X	X	X	X
PETITS CNV						
Del 15q15.3	500	Trouvé	Trouvé	Trouvé	Trouvé	X
Dup 8q24.3	502	X	X	X	X	X
Del 16q22.2	200	Trouvé	Trouvé	X	X	X
Dup 13q12.13	200	X	X	X	X	X
Del 22q11.23	101	X	X	X	X	X
Dup 17p11.2	101	X	X	X	X	X
Del 9q21.13	50	X	X	X	X	X
Dup 8q11.21	51	X	X	X	X	X

Analyse WisecondorX

Excellente sensibilité pour les grands CNV

Meilleure sensibilité pour les délétions

Rouen : et après

- **Augmentation de la cohorte de CNV => 100 CNV**
- **Test de DRAGEN 4.5**
- **Intégration LIMS**
- **Optimisation des fichiers d'annotation**

Et pour rendre ?

- Nomenclature ISCN 2024

8 Microarray

8.1 Introduction

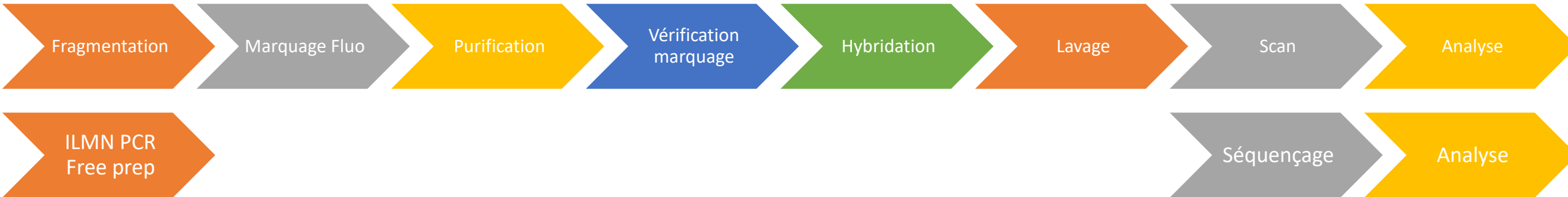
Chromosomal microarray is a genome-wide investigation for the detection of copy number change. In platforms that are supplemented with single nucleotide polymorphism (SNP) probes, zygosity and some forms of ploidy can be determined. Most microarray platforms cover only nonrepetitive regions. As such, the acrocentric short arms, centromeric and telomeric regions, and regions of heterochromatin are not detected by microarray.

Positional structural chromosome information is not provided by this technology and elucidation of the nature of the underlying rearrangement by alternative methods such as conventional karyotype or FISH may be required.

- l. The descriptive narrative, or interpretive comment in the report must indicate the platform used and the resolution. For targeted microarray analysis, see > [Chapter 10](#).
- m. If **shallow genome sequencing** is undertaken, the term **sseq** is used instead of **arr**.
- n. Some of the ISCN examples in this chapter do not represent observed data but are provided to demonstrate the nomenclature principles.

- Facturation : ACPA

Conclusions



- Technique « simple »
- Performances au moins équivalentes sur les CNVs homogènes
- Flux en partie commun avec les autres tests NGS
- Préparer les laboratoires au génome DPN

Remerciements

illumina®

N. Roserot

M. Martinez-Fresno

O. Alibeu

Questions ?



Pr C. Houdayer

Pr G. Nicolas

Dr K. Cassinari

Dr G. Joly

Dr B. Sudrie Arnaud

B. Berg Pinto

Dr A.M. Guerrot

Dr A. Goldenberg

Dr J. Coursimault

Dr G. Vera

Shallow-Genome Sequencing

Pr D. Sanlaville

Dr M. Pujalte

Dr M. Till

A. Labalme

C. Goujon

Dr C. Bardel

Dr S. Mareschal

G. Ben Hassen

Pr C. Vinciguerra

A. Thomas

