

Applications du RNA-Seq par capture dans le diagnostic des maladies rares

Laura Do Souto Ferreira, bioinformaticienne

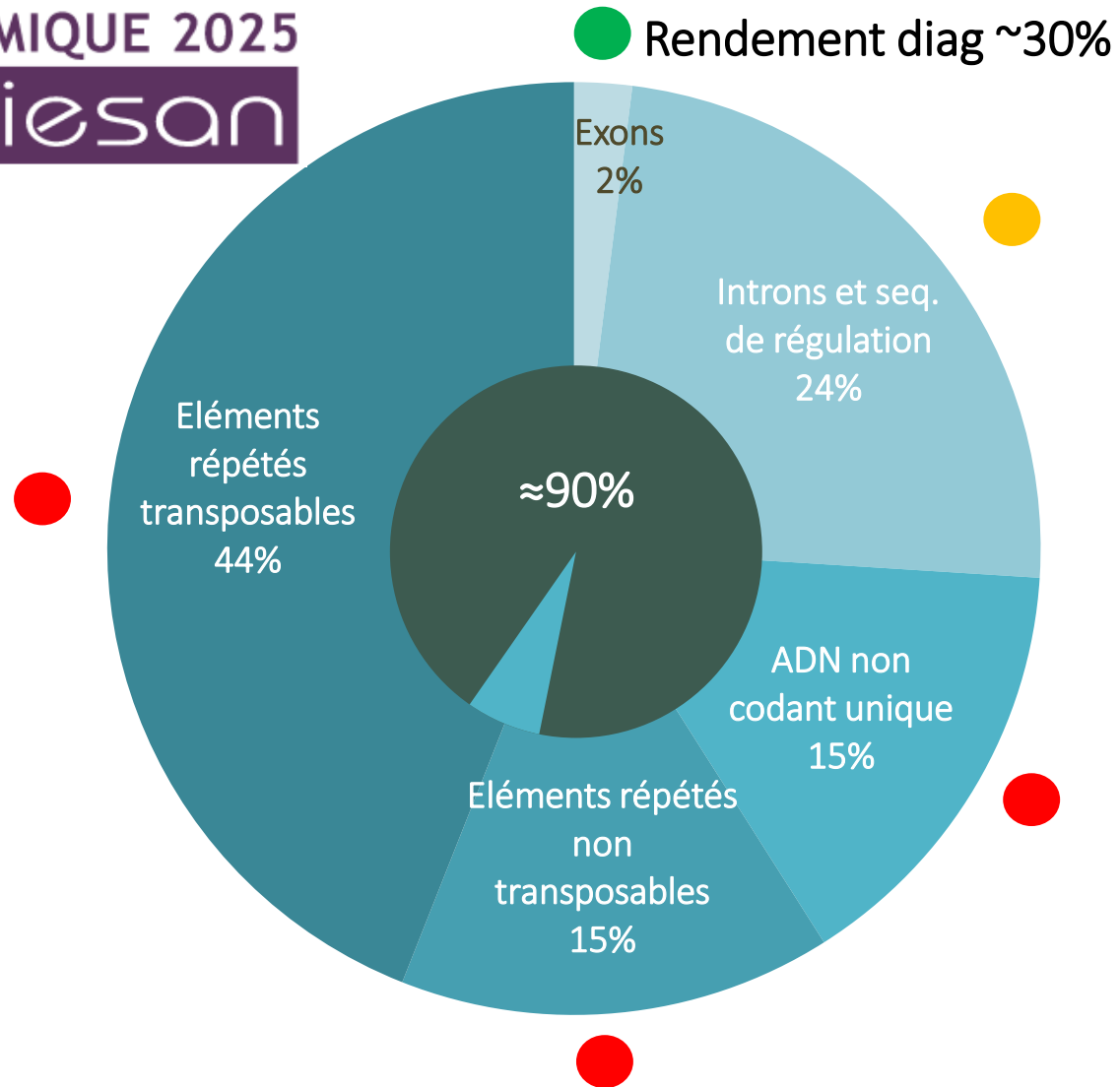
CHU de Nantes - Service de génétique médicale

Anomalies du développement

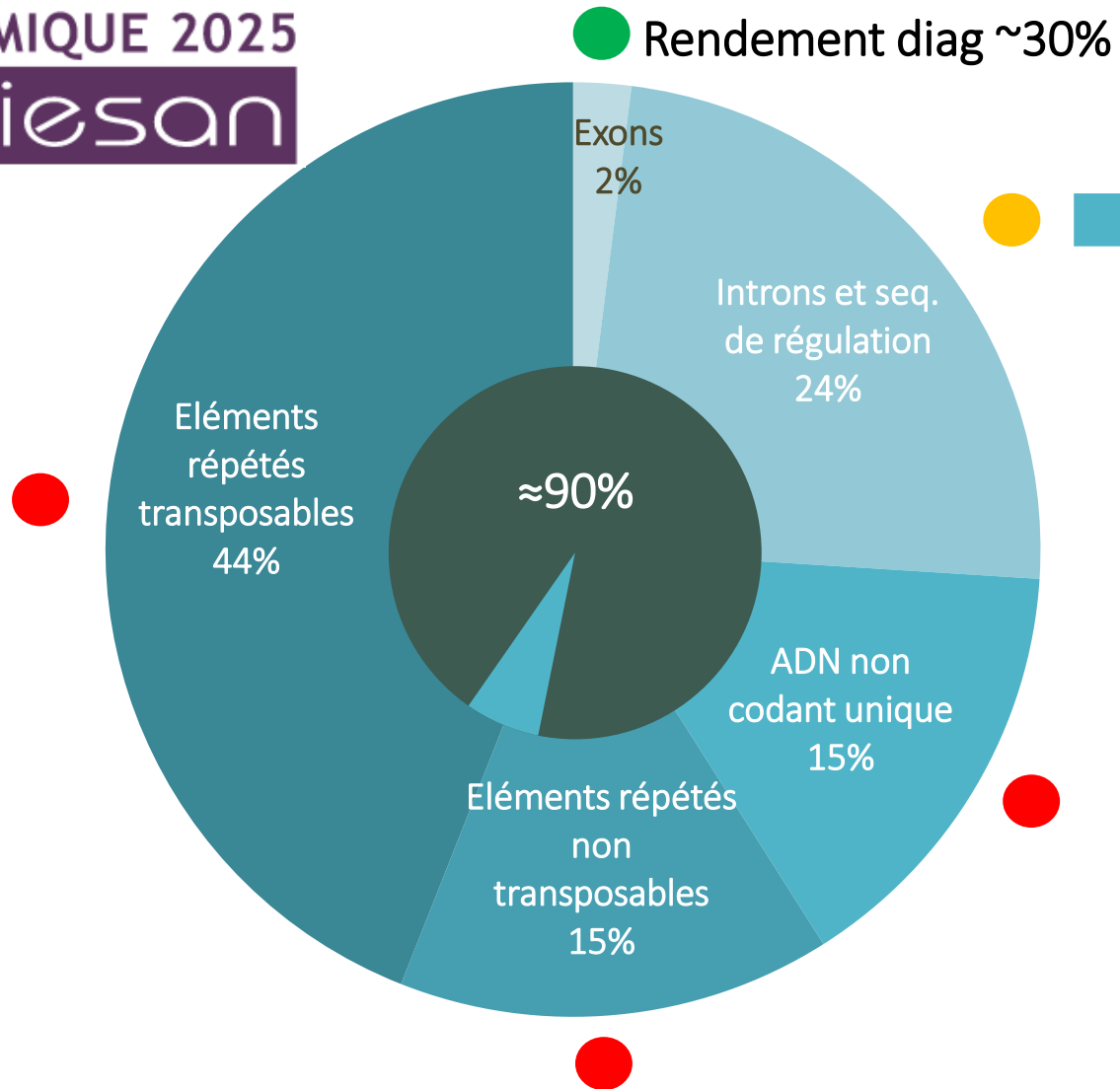
Anomalies du développement : vers le génome en 1^{ère} intention

FRANCE MÉDECINE
GÉNOMIQUE 2025

aviesan



Anomalies du développement : vers le génome en 1^{ère} intention



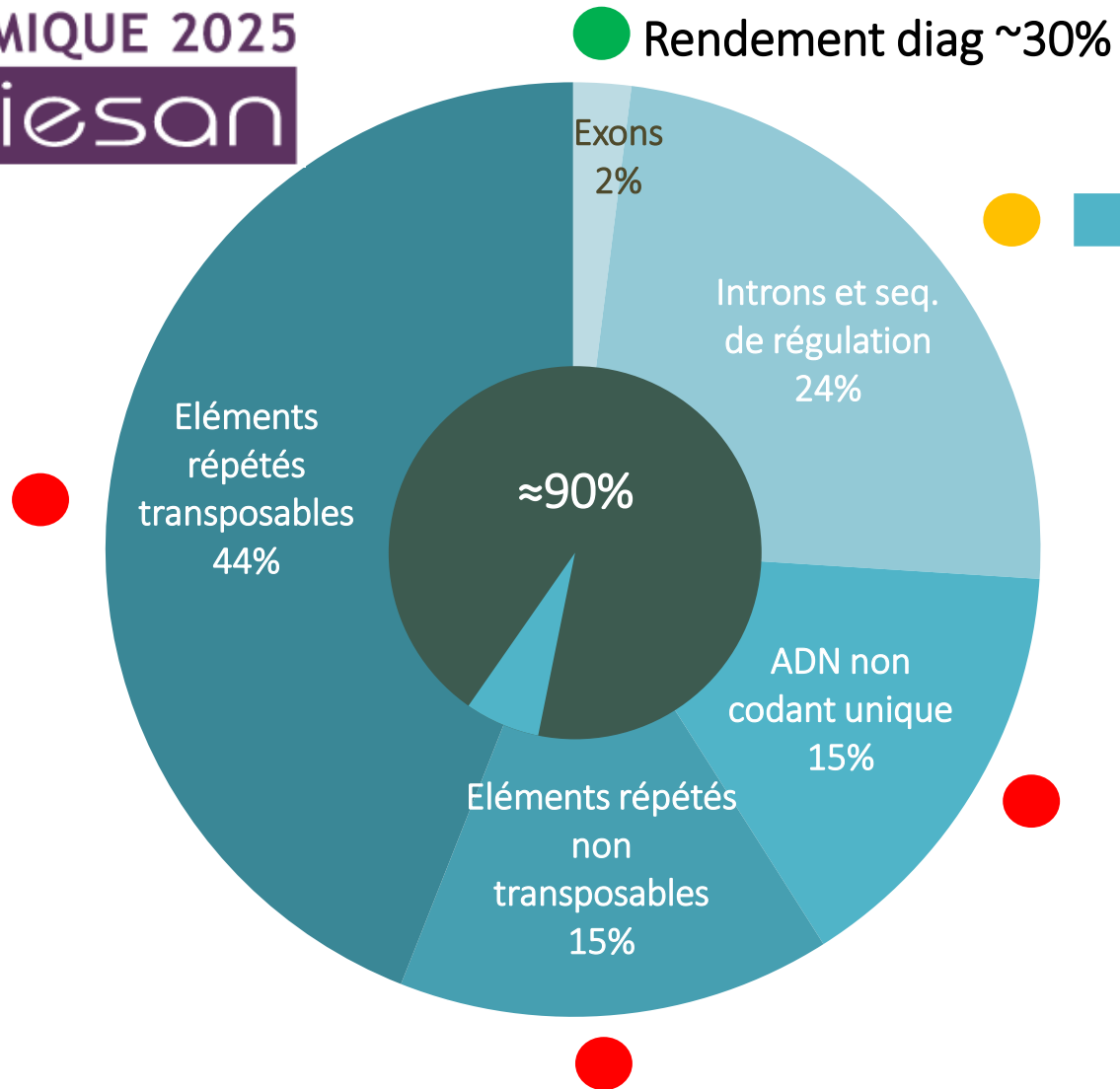
→ Difficulté de prédire l'effet



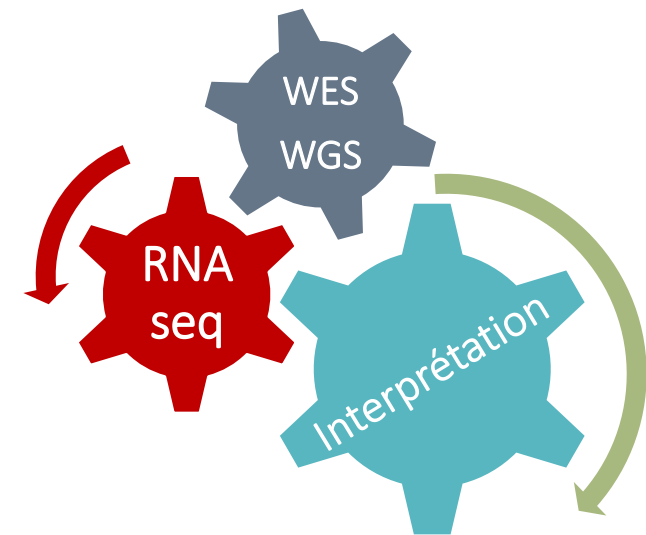
Anomalies du développement : vers le génome en 1^{ère} intention

FRANCE MÉDECINE
GÉNOMIQUE 2025

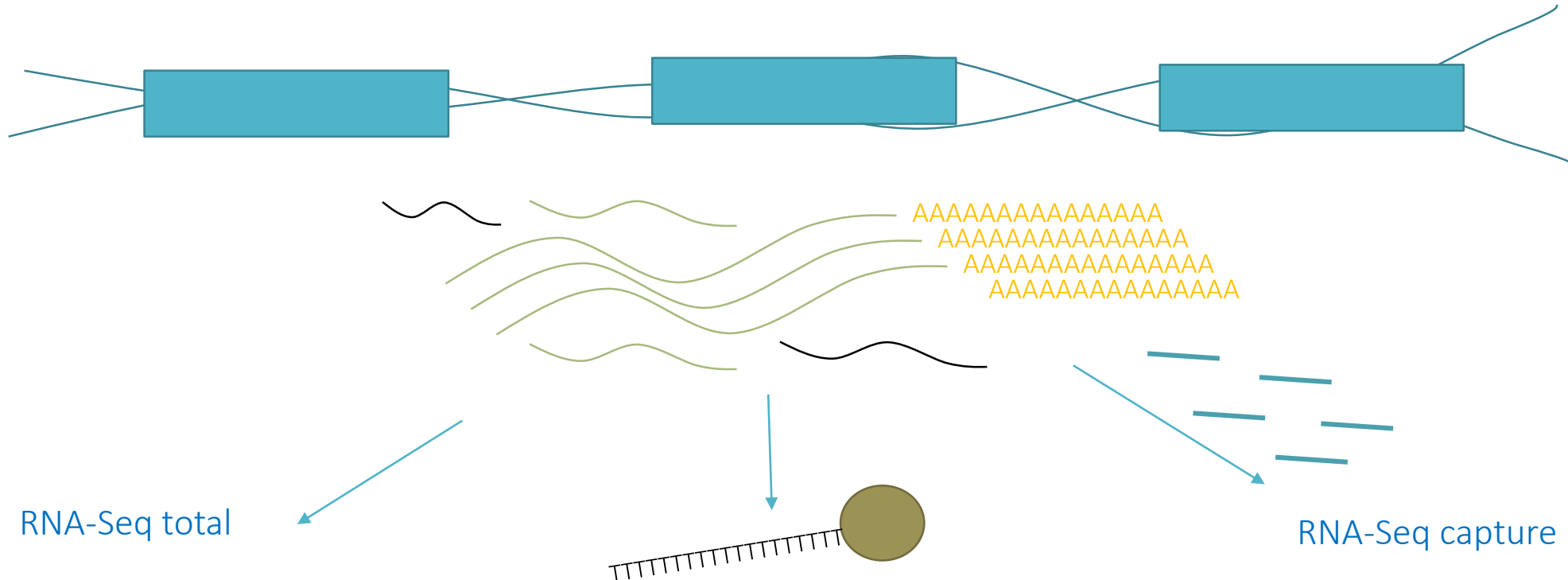
aviesan



● → Difficulté de prédire l'effet



Comment enrichir en ARNm



RNA-Seq total

- Besoin étape ribodéplétion
- Vision exhaustive des ARN (hors petits ARN comme miR)
- Besoin nombre de reads +++

RNA-Seq poly-A

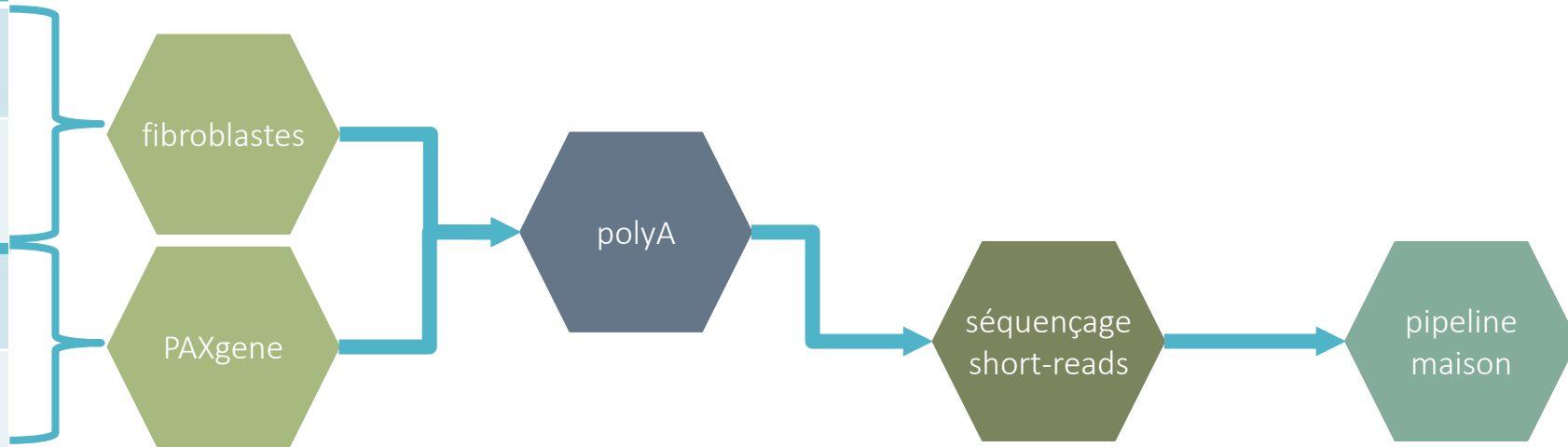
- Enrichissement poly-A: élimination des ARN ribosomiques, moins de lncRNA
- Couverture exhaustive des ARNm
- Saturation possible par ARNm très présent (ex: globine)
- Besoin reads +

RNA-Seq capture

- Capture des exons
- Meilleure couverture exonique: permet de diminuer le nombre de reads nécessaires
- Pas de couverture des UTR et introns
- Moins de saturation par des ARNm très exprimés (ex: globine dans PAXgene)

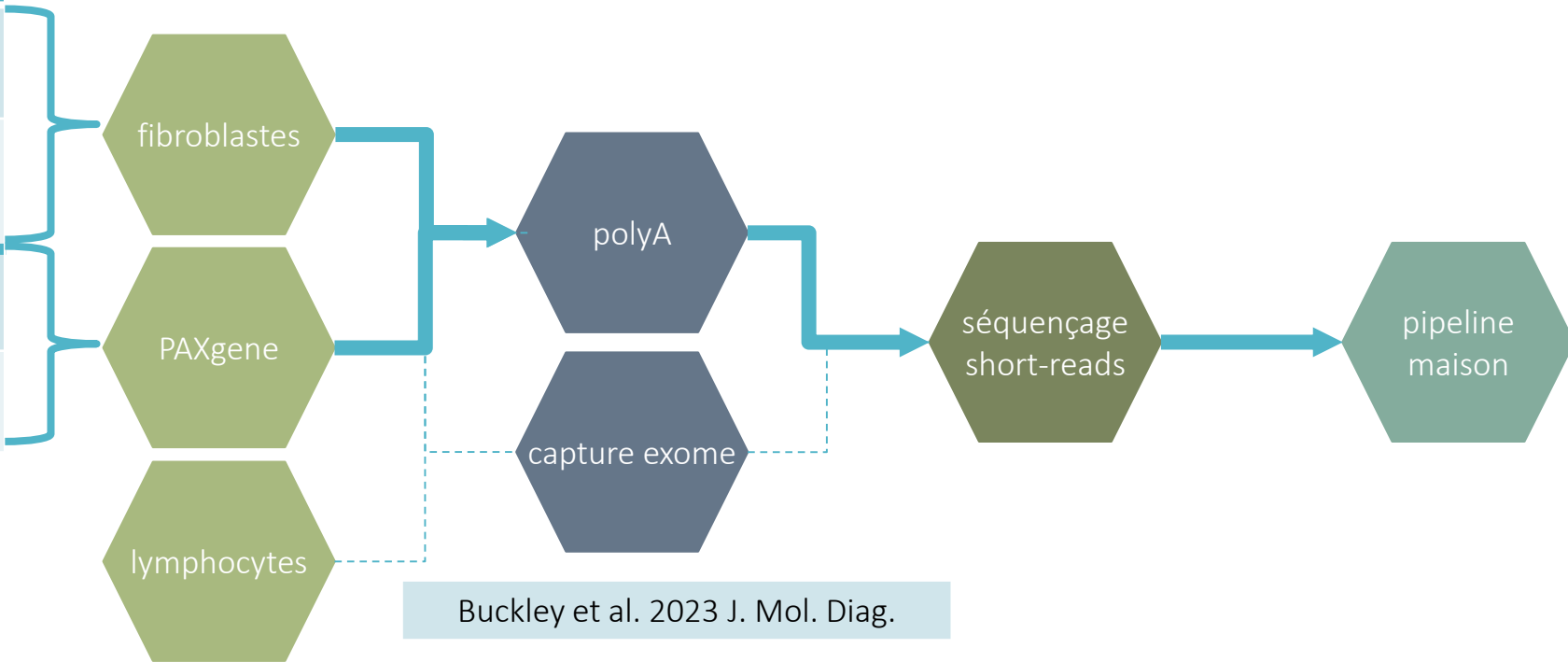
RNA-Seq appliqué aux maladies rares

Etude	Pathologie	Rendement
Yepes et al. 2022 Geno. medicine	Maladies mito. (WES neg)	32/205 (16%)
Dekker et al. 2023 AJHG	Troubles du ND (WES neg)	8/67 (12%)
Frésard et al 2019 Nat Med	Hétérogène (WES neg)	6/80 (7%)
Deshwar et al. 2023 AJHG	Hétérogène (WGS neg)	0/39



RNA-Seq appliqué aux maladies rares

Etude	Pathologie	Rendement
Yepes et al. 2022 Geno. medicine	Maladies mito. (WES neg)	32/205 (16%)
Dekker et al. 2023 AJHG	Troubles du ND (WES neg)	8/67 (12%)
Frésard et al 2019 Nat Med	Hétérogène (WES neg)	6/80 (7%)
Deshwar et al. 2023 AJHG	Hétérogène (WGS neg)	0/39

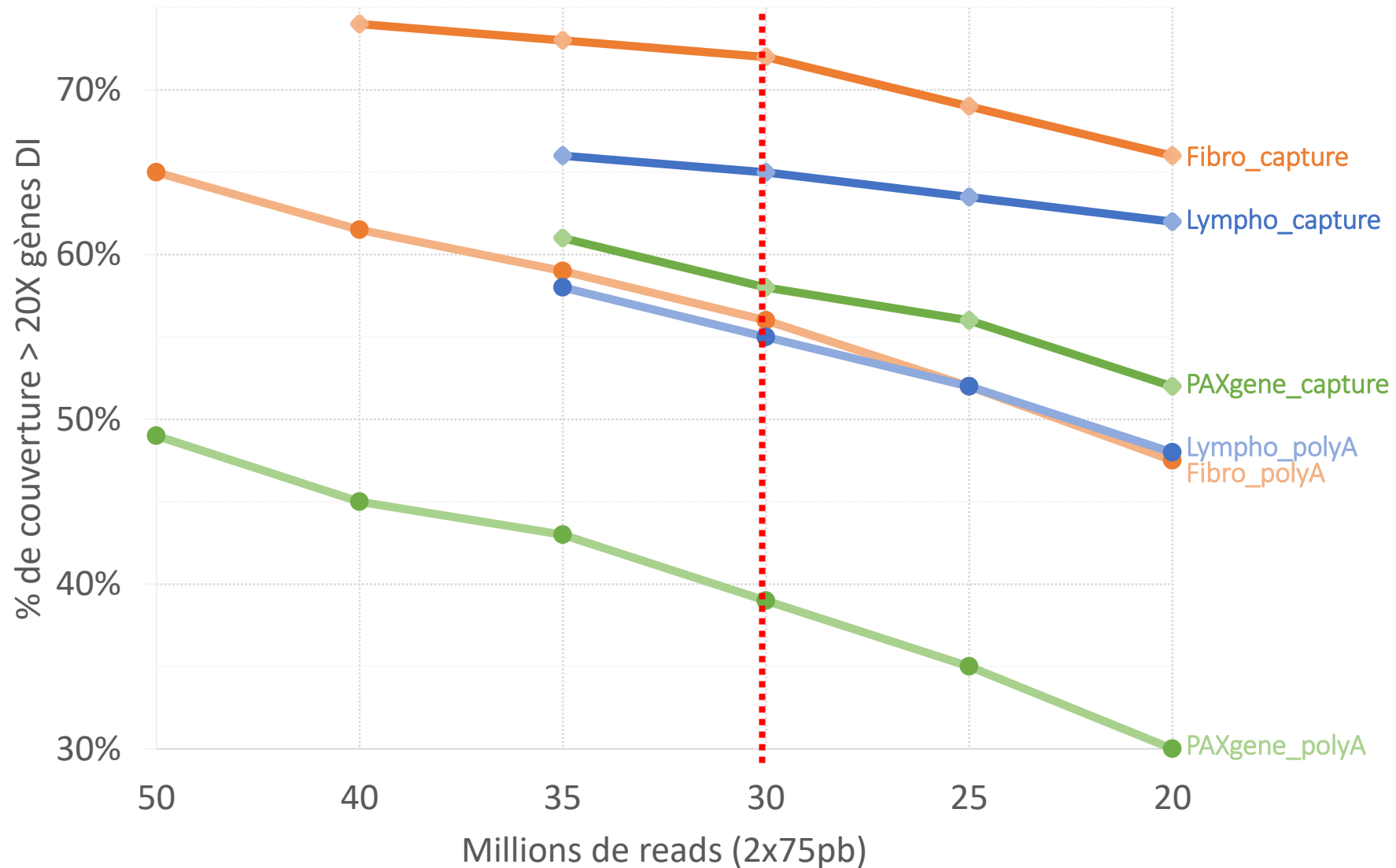


Adaptée laboratoire de diagnostic

Robuste / Reproductible

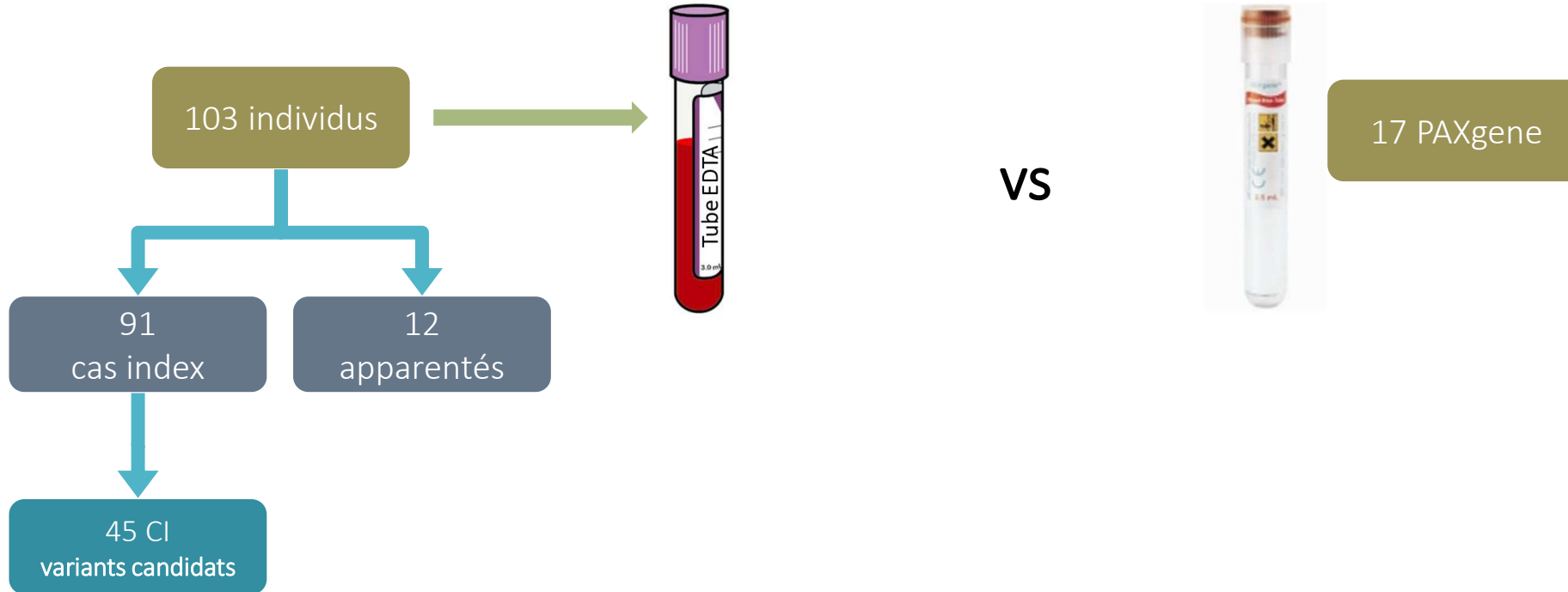
Expression des gènes du ND / DI

Comparaison Lympho vs Fibro vs PAXgene – polyA vs capture Gènes de DI (panelApp DI ~ 1300 gènes)

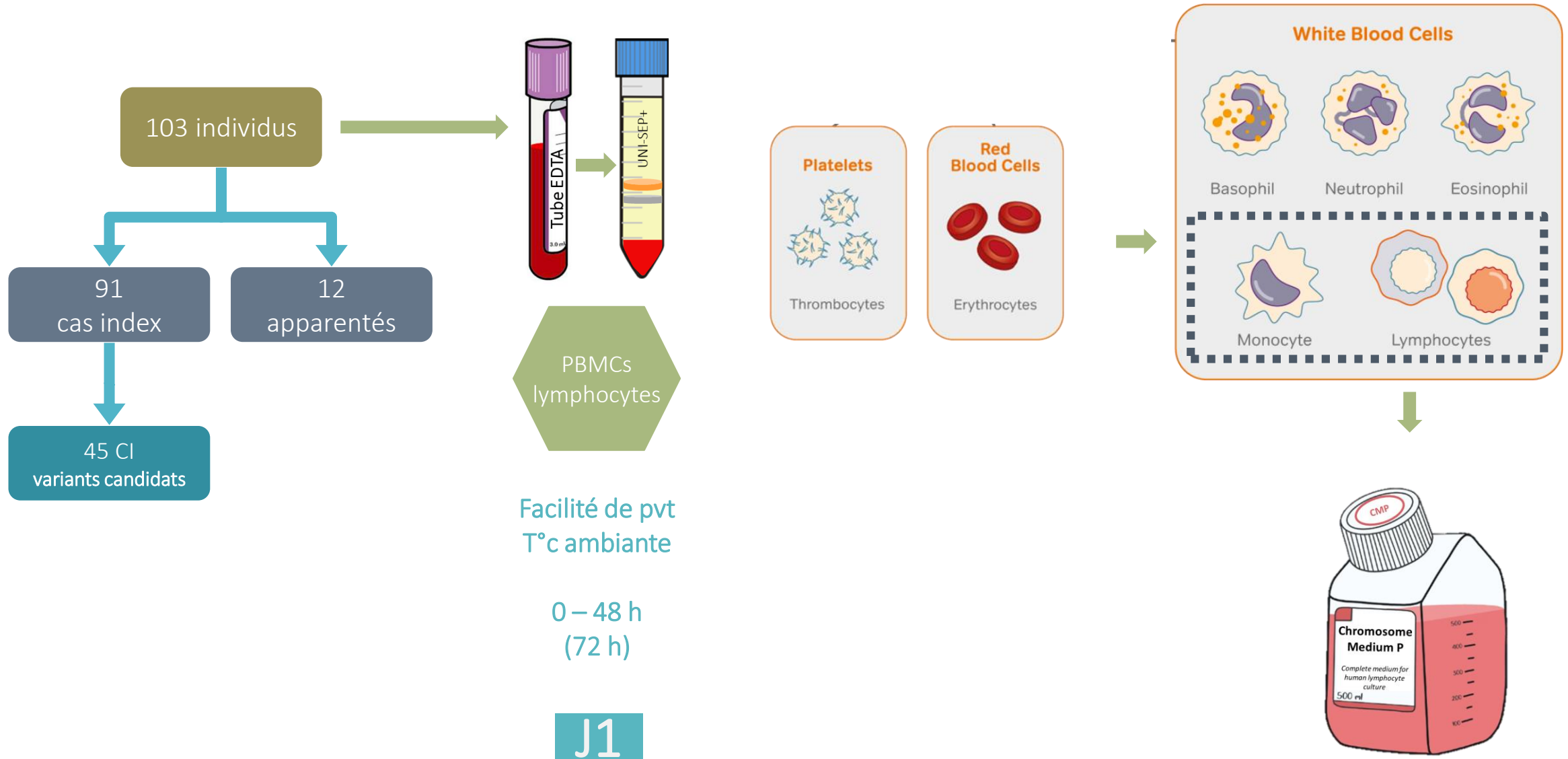


Choix pour une activité diagnostique et bonne couverture gènes DI -> Capture sur culture lymphocytaire

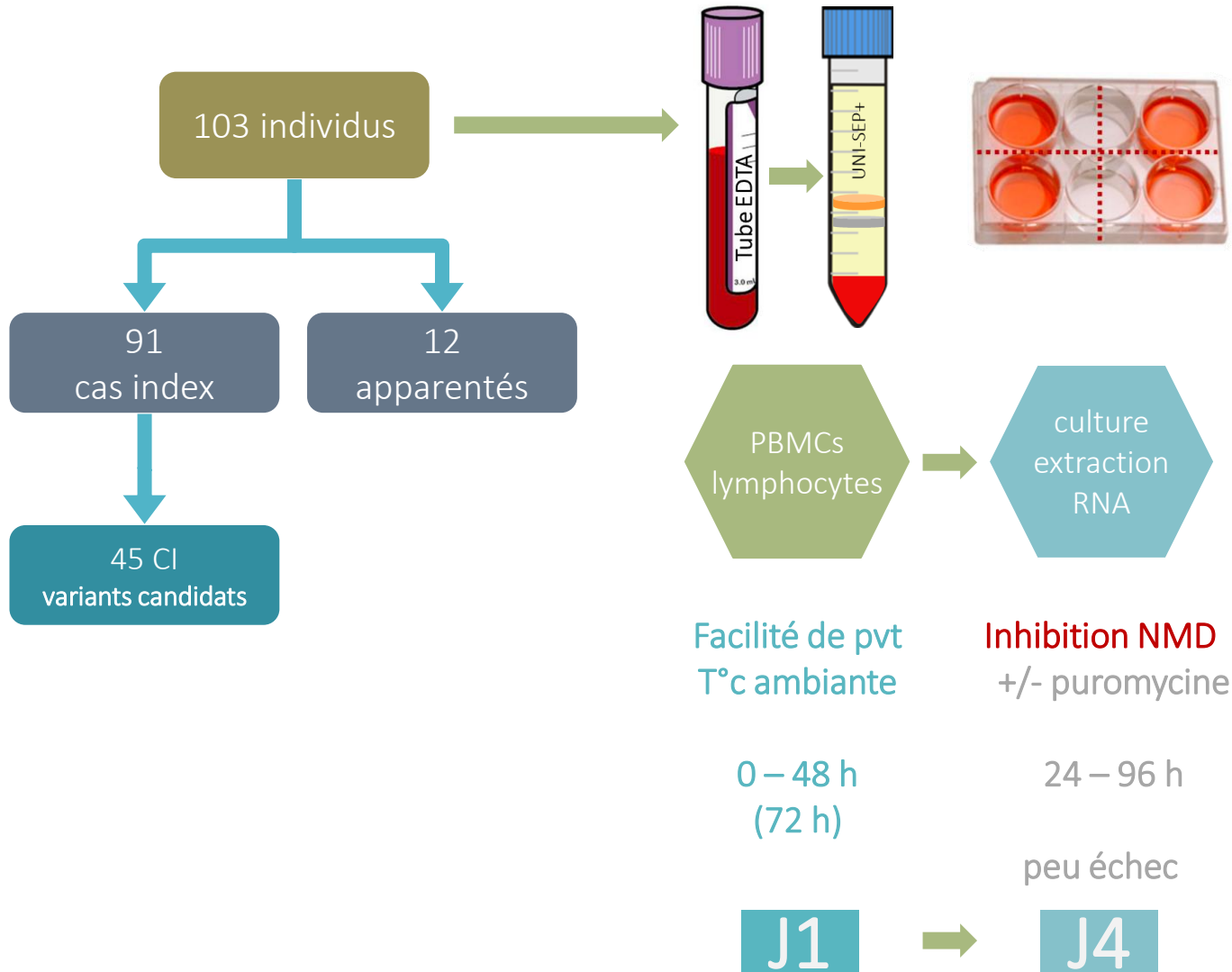
Méthode : Capture d'exome sur culture lymphocytaire



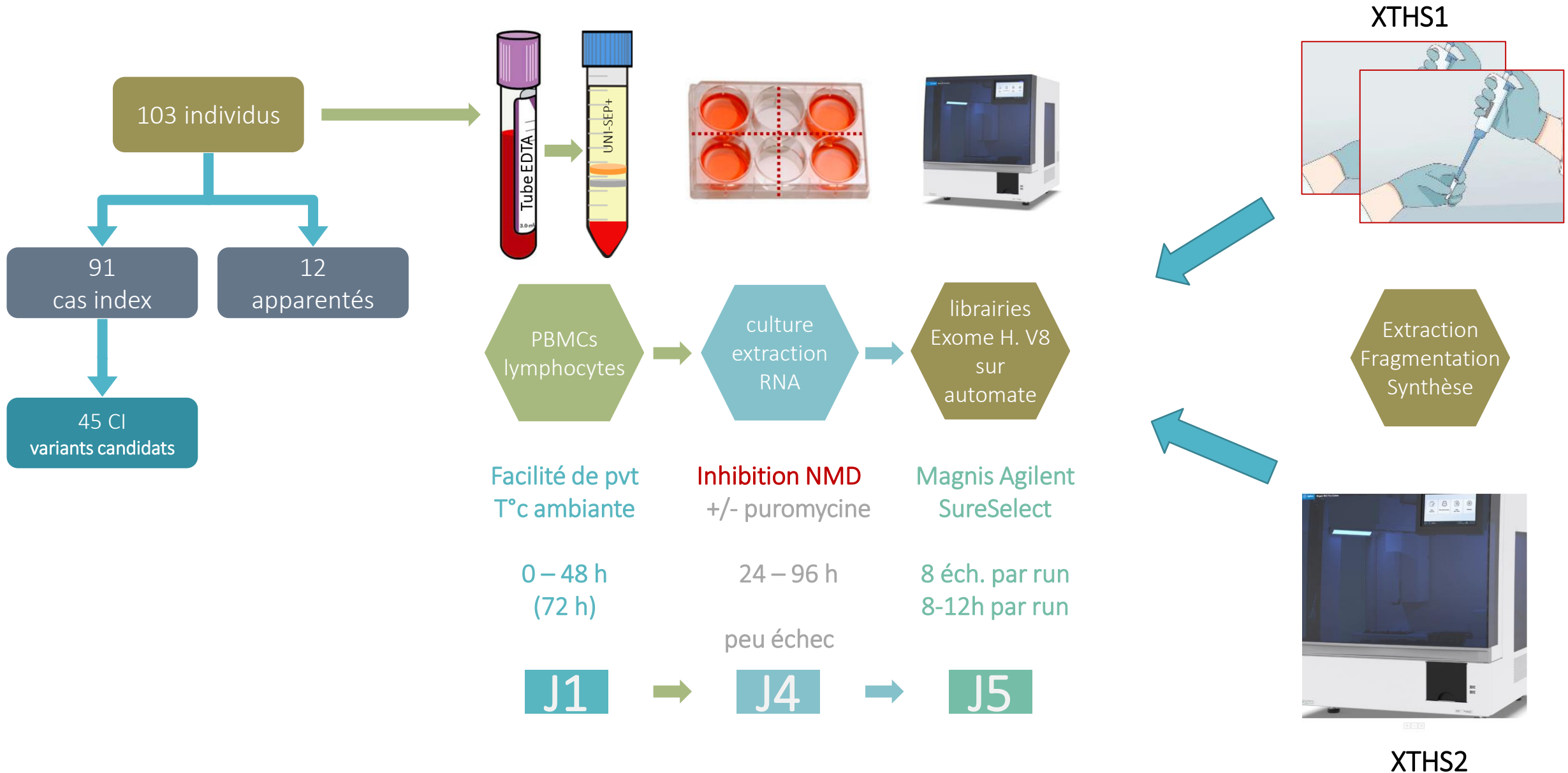
Méthode : Capture d'exome sur culture lymphocytaire



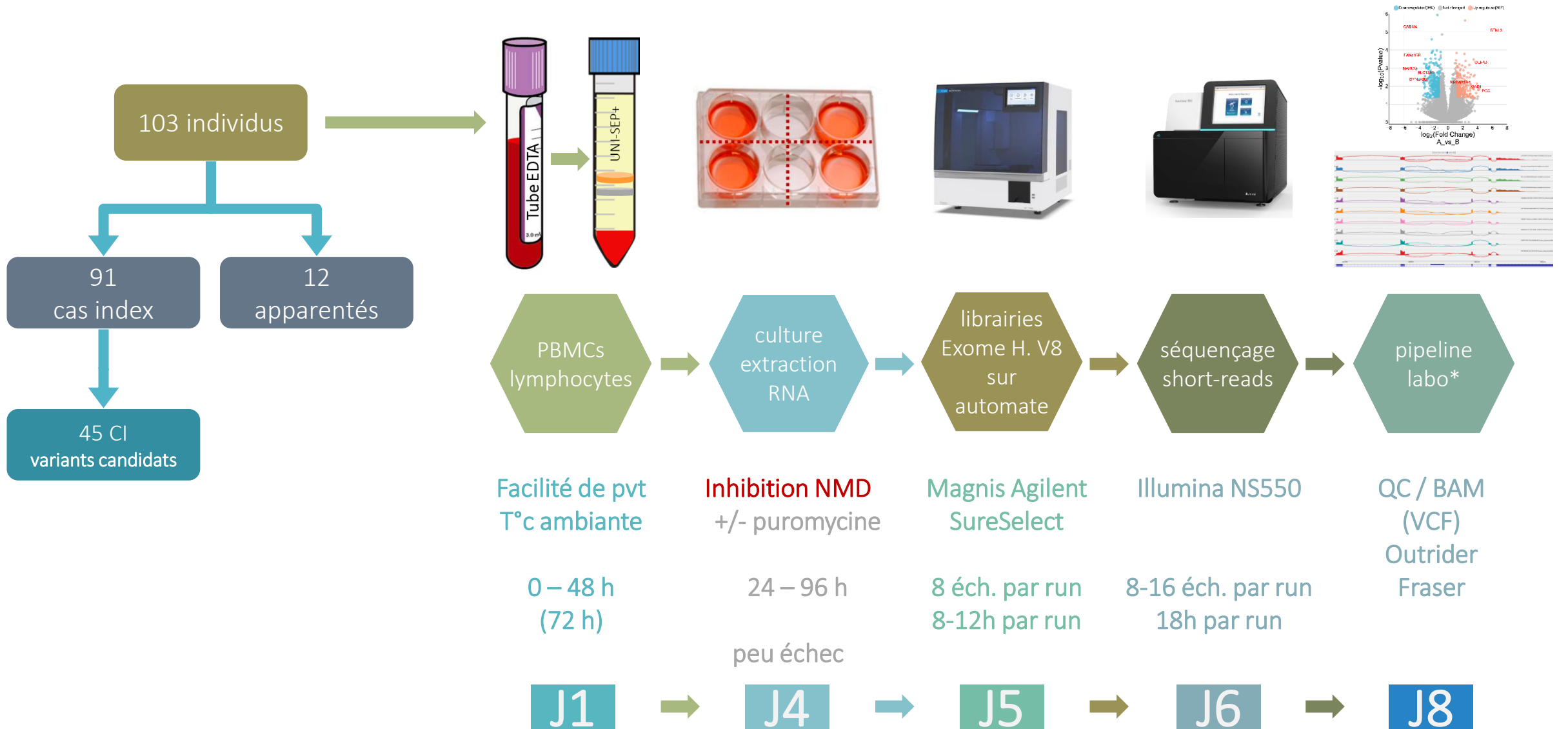
Méthode : Capture d'exome sur culture lymphocytaire



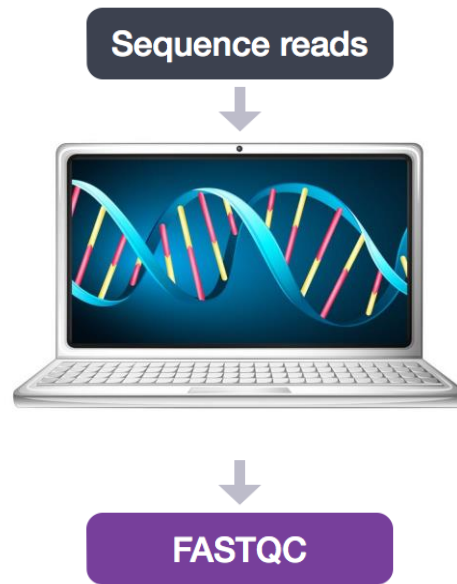
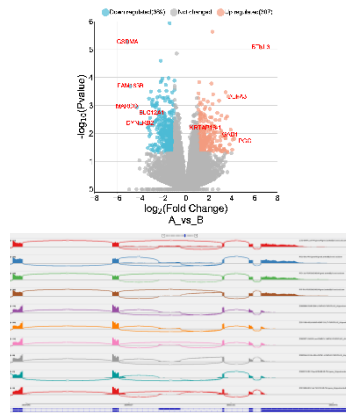
Méthode : Capture d'exome sur culture lymphocytaire



Méthode : Capture d'exome sur culture lymphocytaire



Méthode : Capture d'exome sur culture lymphocytaire



pipeline
labo*

QC / BAM
(VCF)
Outrider
Fraser

MultiQC

A modular tool to aggregate results from bioinformatics analyses across many samples into a single report.

Report generated on 2015-09-24, 09:09 based on data in `/Users/philewels/Desktop/MultiQC_testing/star/data`

Report location: `/Users/philewels/Desktop/MultiQC_testing/star/multiqc_report/multiqc_report.html`

General Statistics

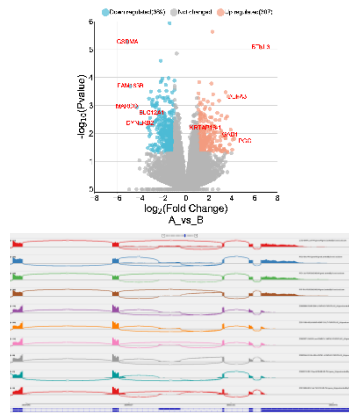
Show Key

Sample Name	% Assigned	M Assigned	% Mapped	M Mapped	Trimmed	% Dups	% GC	Length	M Seqs
SRR1067503_1	2.4%	0.9	63.2%	19.3	2.1%	12.9%	44%	35	30.5
SRR1067505_1	7.4%	1.5	79.1%	14.2	3.5%	7.8%	47%	35	18.0
SRR1067510_1	1.1%	0.6	50.6%	17.4	2.0%	11.4%	40%	35	34.3
SRR1067514_1	5.7%	1.9	70.2%	23.6	3.1%	6.6%	44%	35	33.6
SRR1067519_1	3.2%	0.9	81.1%	19.9	2.3%	5.8%	42%	35	24.6
SRR1067522_1	1.4%	0.7	61.8%	22.0	1.5%	13.3%	40%	36	35.7

Toolbox



Méthode : Capture d'exome sur culture lymphocytaire



Sequence reads



FASTQC

Adapter Trimming (Optional)

pipeline labo*

QC / BAM
(VCF)
Outrider
Fraser

- Analyse qualité des reads
- Qualité des bases
 - Qualité du reads
 - Taille reads
 - Adapteurs

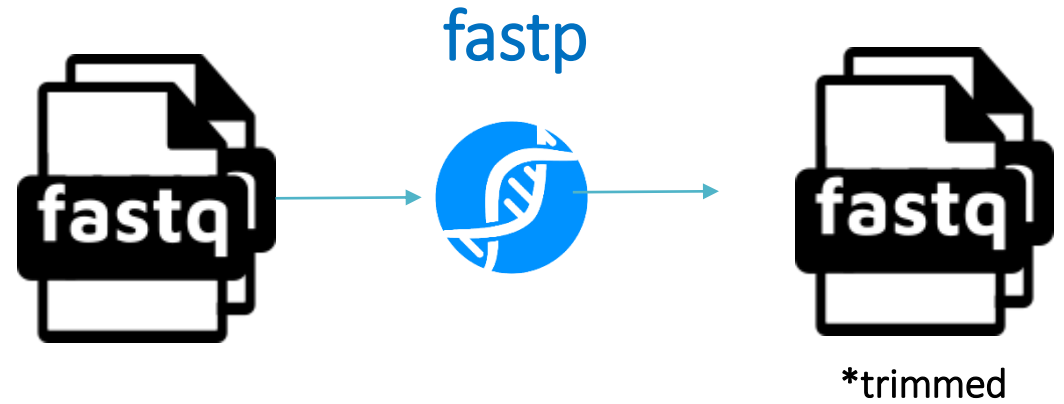
MultiQC

A modular tool to aggregate results from bioinformatics analyses across many samples into a single report.

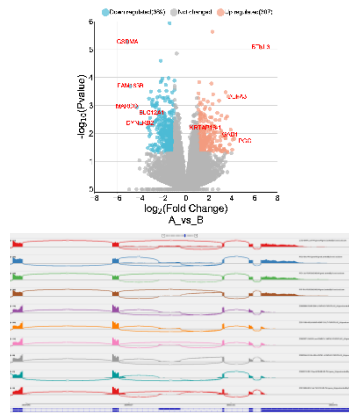
Report generated on 2015-09-24, 09:09 based on data in /Users/philewels/Desktop/MultiQC_testing/star/data
Report location: /Users/philewels/Desktop/MultiQC_testing/star/multiqc_report/multiqc_report.html

General Statistics

Sample Name	% Assigned	M Assigned	% Mapped	M Mapped	Trimmed	% Dups	% GC	Length	M Seqs
SRR1067503_1	2.4%	0.9	63.2%	19.3	2.1%	12.9%	44%	35	30.5
SRR1067505_1	7.4%	1.5	79.1%	14.2	3.5%	7.8%	47%	35	18.0
SRR1067510_1	1.1%	0.6	50.6%	17.4	2.0%	11.4%	40%	35	34.3
SRR1067514_1	5.7%	1.9	70.2%	23.6	3.1%	6.6%	44%	35	33.6
SRR1067519_1	3.2%	0.9	81.1%	19.9	2.3%	5.8%	42%	35	24.6
SRR1067522_1	1.4%	0.7	61.8%	22.0	1.5%	13.3%	40%	36	35.7

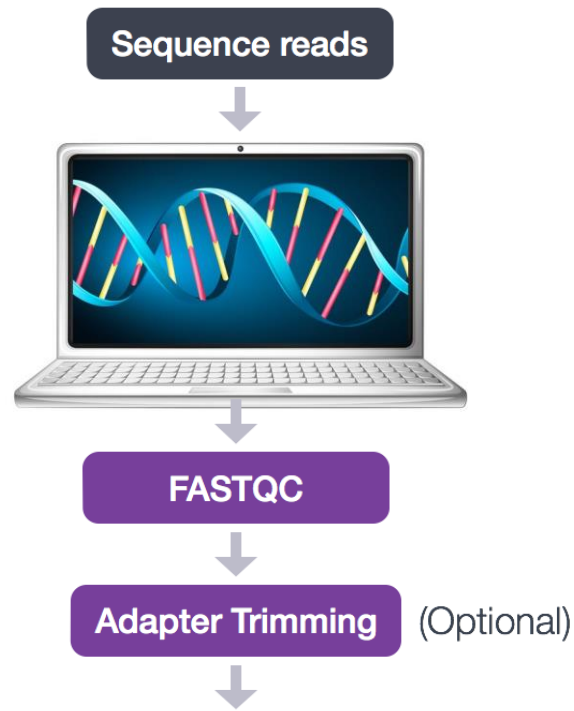


Méthode : Capture d'exome sur culture lymphocytaire



pipeline labo*

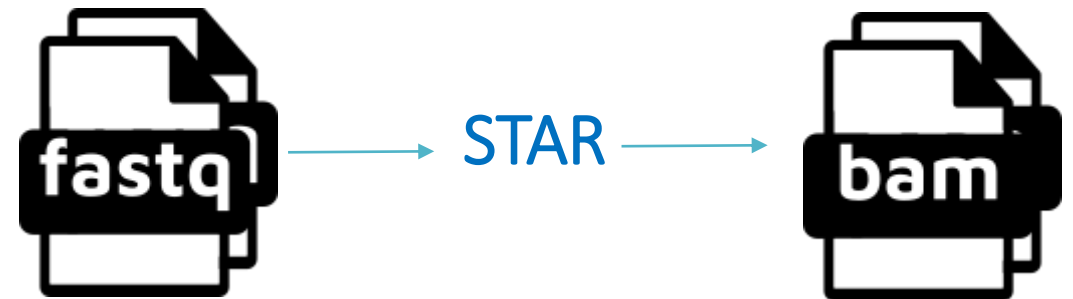
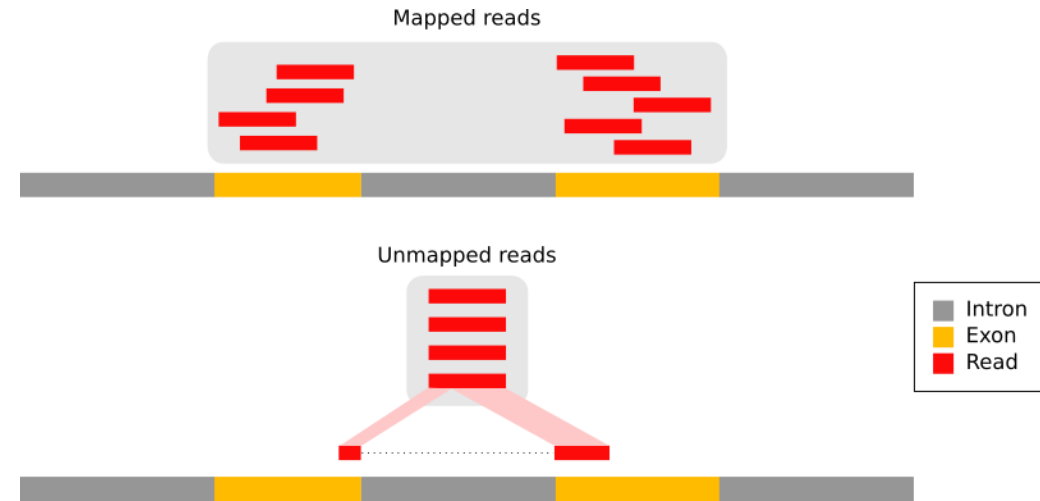
QC / BAM
(VCF)
Outrider
Fraser



Splice-aware mapping to genome

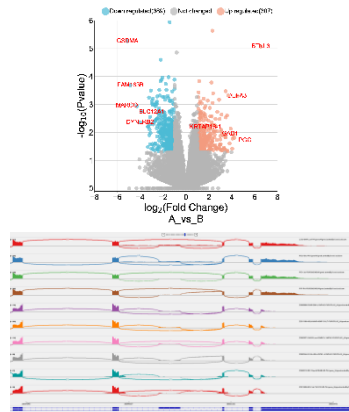
Alignement splice aware

- Référence : GRCh38
- 2Pass



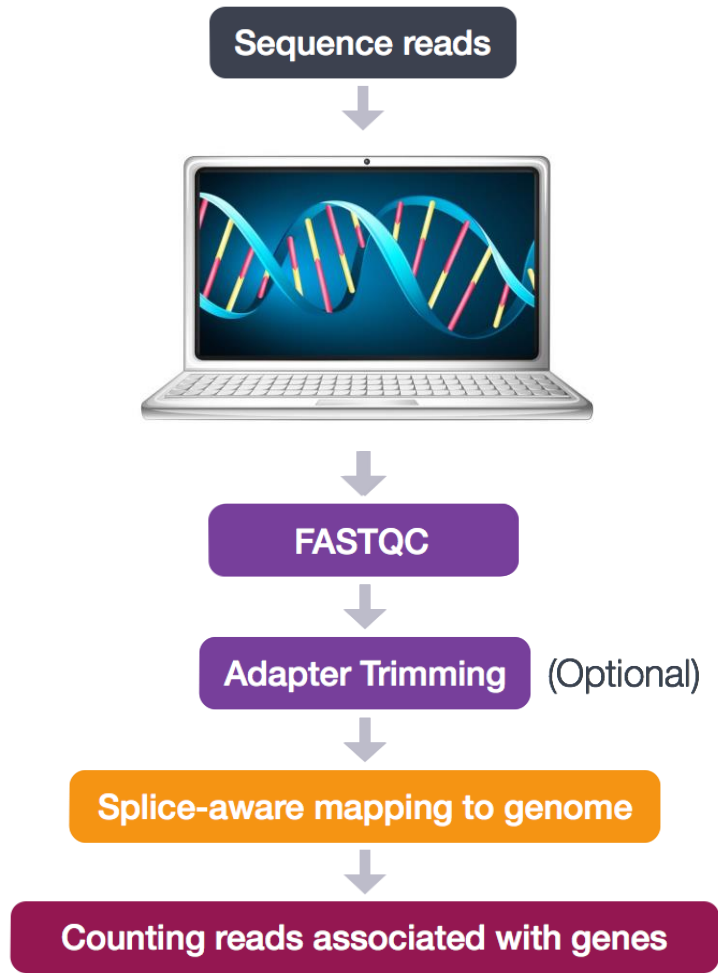
*trimmed

Méthode : Capture d'exome sur culture lymphocytaire



pipeline labo*

QC / BAM
(VCF)
Outrider
Fraser



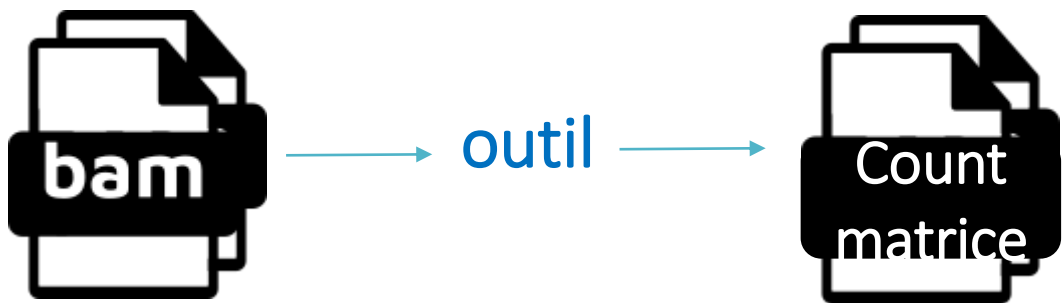
- Comptage**
- HTSeq -count
 - Kallisto

! Protocol stranded

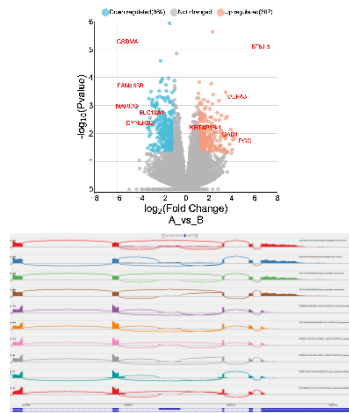
Each column is a sample

GENE ID	KD.2	KD.3	OE.1	OE.2	OE.3	IR.1	IR.2	IR.3
1/2-SBSRNA4	57	41	64	55	38	45	31	39
A1BG	71	40	100	81	41	77	58	40
A1BG-AS1	256	177	220	189	107	213	172	126
A1CF	0	1	1	0	0	0	0	0
A2LD1	146	81	138	125	52	91	80	50
A2M	10	9	2	5	2	9	8	4
A2ML1	3	2	6	5	2	2	1	0
A2MP1	0	0	2	1	3	0	2	1
A4GALT	56	37	107	118	65	49	52	37
A4GNT	0	0	0	0	1	0	0	0
AA06	0	0	0	0	0	0	0	0
AAA1	0	0	1	0	0	0	0	0
AAAS	2288	1363	1753	1727	835	1672	1389	1121
AACS	1586	923	951	967	484	938	771	635
AACSP1	1	1	3	0	1	1	1	3
AADAC	0	0	0	0	0	0	0	0
AADACL2	0	0	0	0	0	0	0	0
AADACL3	0	0	0	0	0	0	0	0
AADACL4	0	0	1	1	0	0	0	0
AADAT	856	539	593	576	359	567	521	416
AAGAB	4648	2550	2648	2356	1481	3265	2790	2118
AAK1	2310	1384	1869	1602	980	1675	1614	1108
AAMP	5198	3081	3179	3137	1721	4061	3304	2623
AANAT	7	7	12	12	4	6	2	7
AARS	5570	3323	4782	4580	2473	3953	3339	2666

Each row is a gene



Méthode : Capture d'exome sur culture lymphocytaire



Sequence reads



FASTQC

Adapter Trimming (Optional)

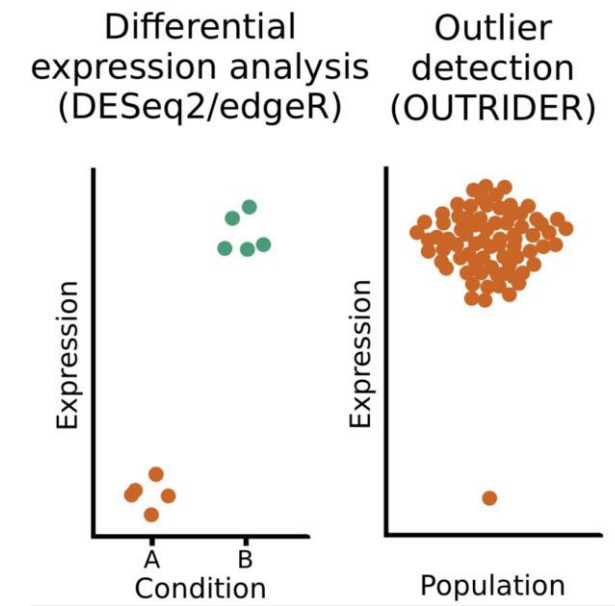
Splice-aware mapping to genome

Counting reads associated with genes

Statistical analysis to identify differentially expressed genes

pipeline labo*

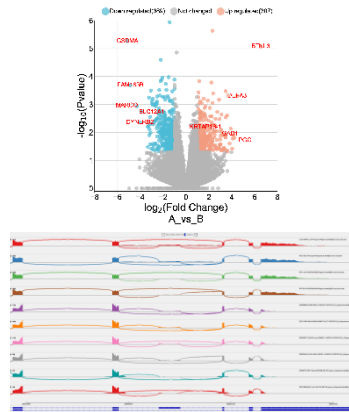
QC / BAM
(VCF)
Outrider
Fraser



→ outrider →

Gènes à expression aberrante

Méthode : Capture d'exome sur culture lymphocytaire



Sequence reads



FASTQC

Adapter Trimming (Optional)

Splice-aware mapping to genome

Counting reads associated with genes

Statistical analysis to identify differentially expressed genes

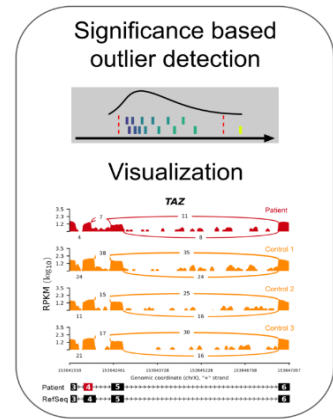
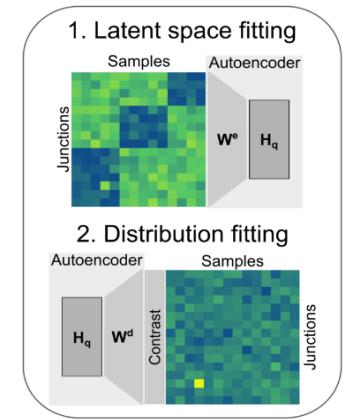
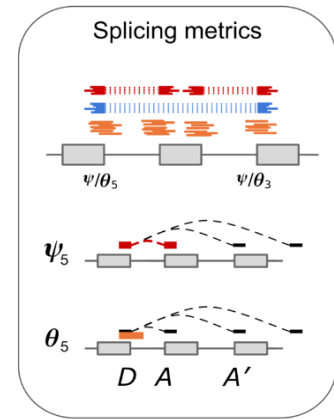
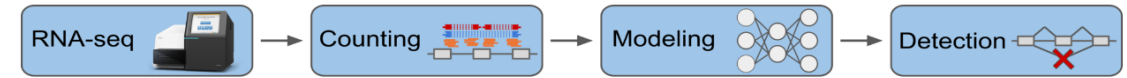
pipeline labo*

QC / BAM
(VCF)
Outrider
Fraser



Detection of aberrant splicing events in RNA-seq data using FRASER

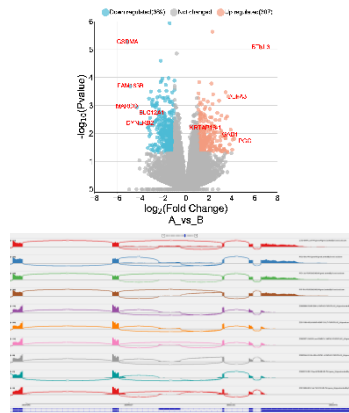
Christian Mertes^{1,6}, Ines F. Scheller^{1,2,6}, Vicente A. Yépez^{1,3}, Muhammed H. Çelik¹, Yingjiqiong Liang¹, Laura S. Kremer^{4,5}, Mirjana Gusic^{4,5}, Holger Prokisch^{4,5} & Julien Gagneur^{1,2,5}



FRASER

Epissages aberrants

Méthode : Capture d'exome sur culture lymphocytaire



Sequence reads



FASTQC

Adapter Trimming (Optional)

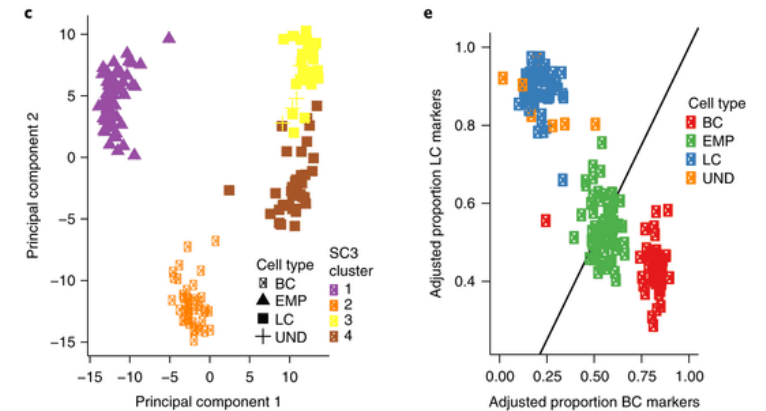
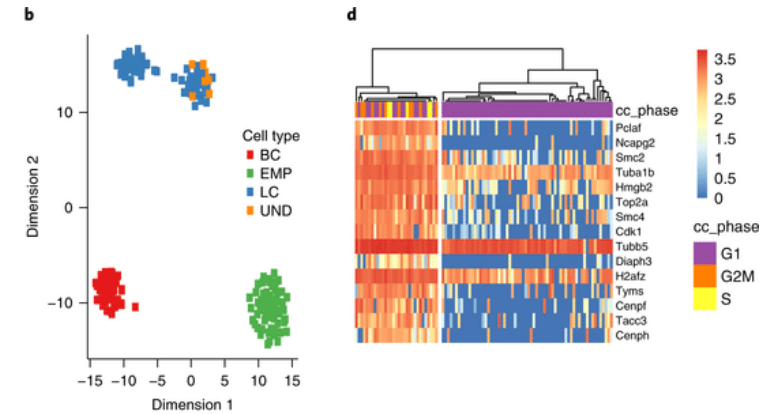
Splice-aware mapping to genome

Counting reads associated with genes

Statistical analysis to identify differentially expressed genes

pipeline labo*

QC / BAM
(VCF)
Outrider
Fraser



DESeq2

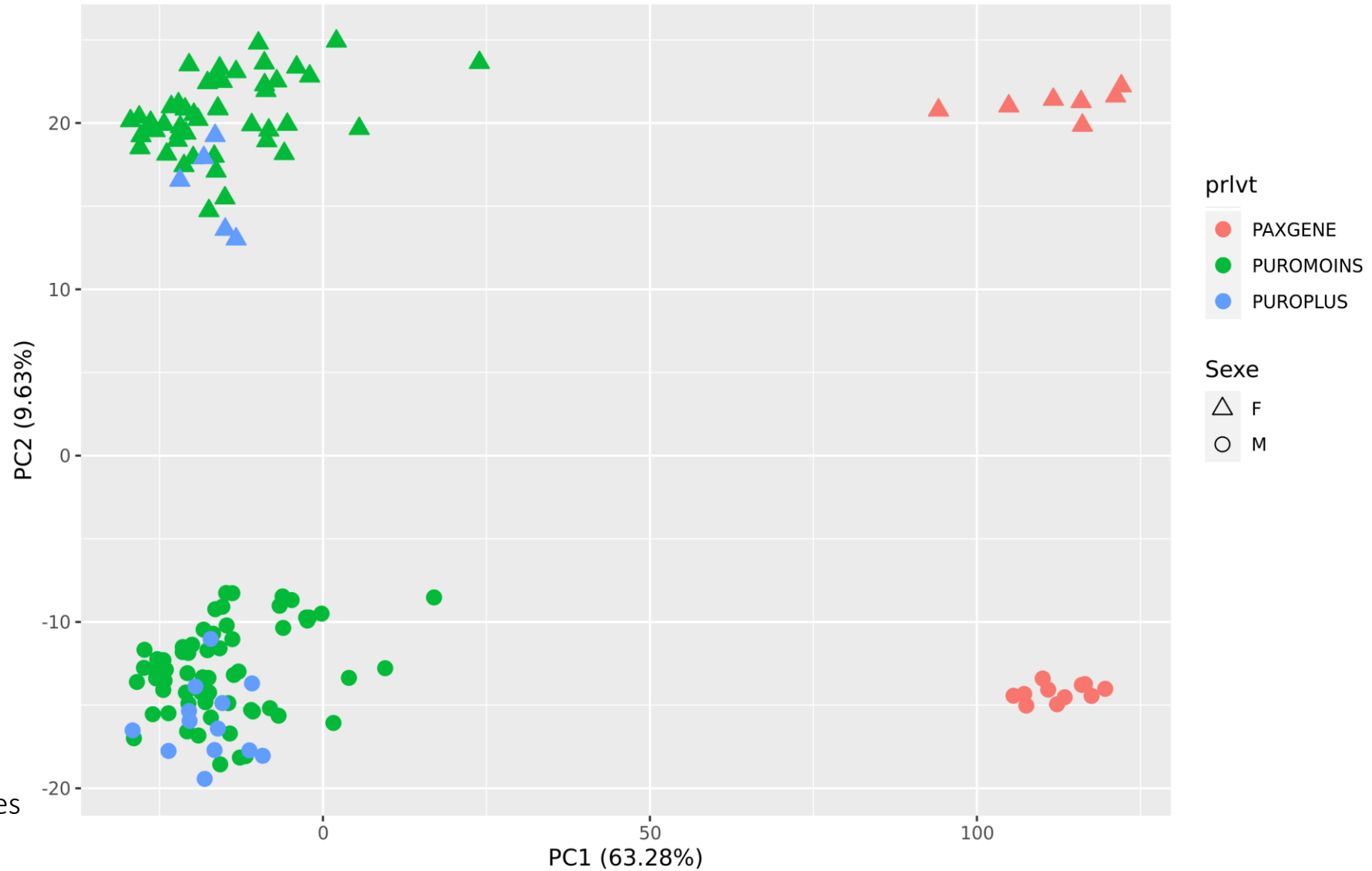
Signatures

Méthode : Capture d'exome sur culture lymphocytaire

- Culture lymphocytaire
 - Inhibition du NMD
- Capture Exome V8 Agilent (XTHS1 et XTHS2)
 - Séquençage Illumina 2x75

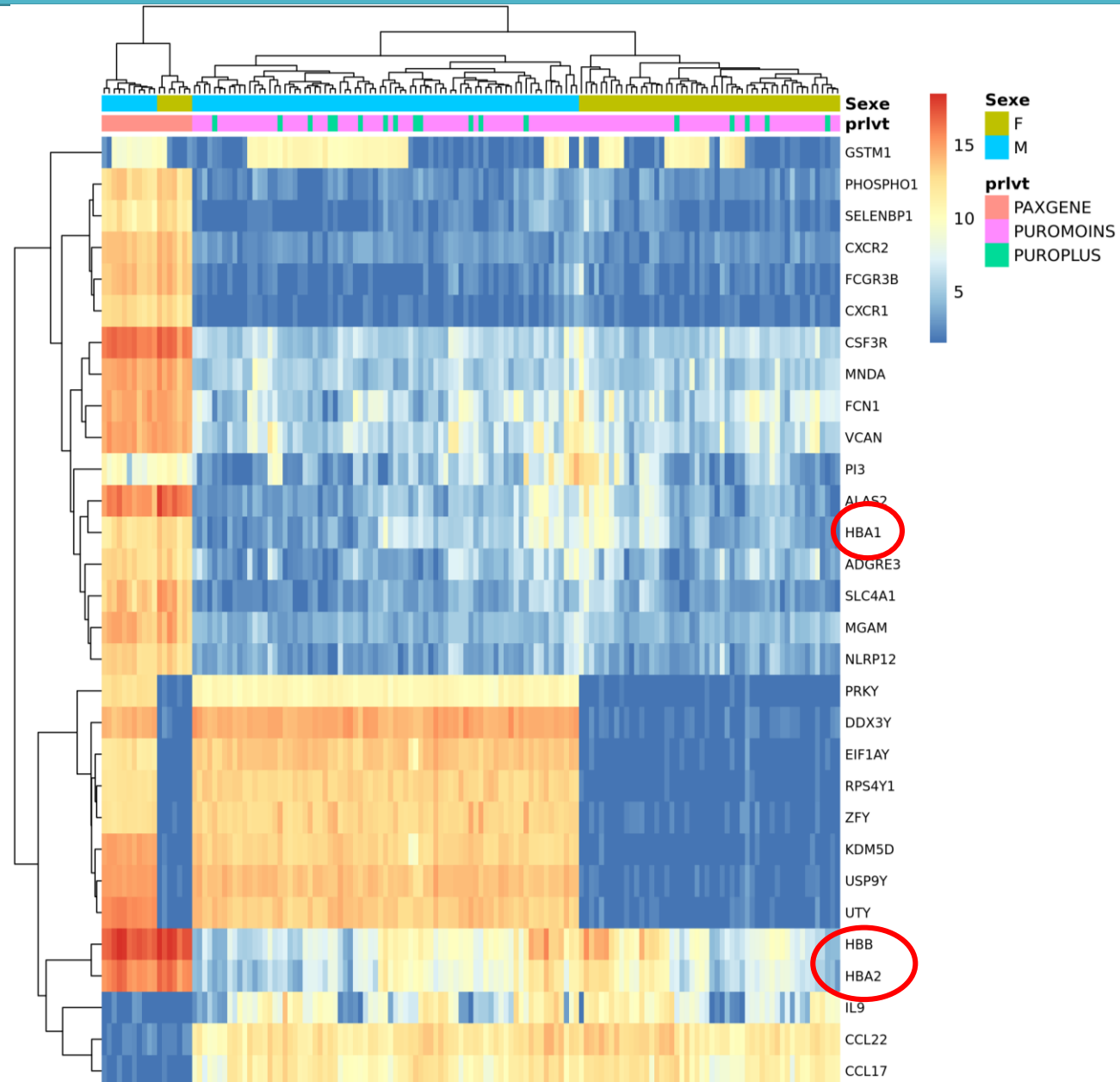
Analyse bioinformatique
QC
Alignement
FRASER / OUTRIDER / DESEQ2

ACP : PAXgene vs cultures lymphocytaires

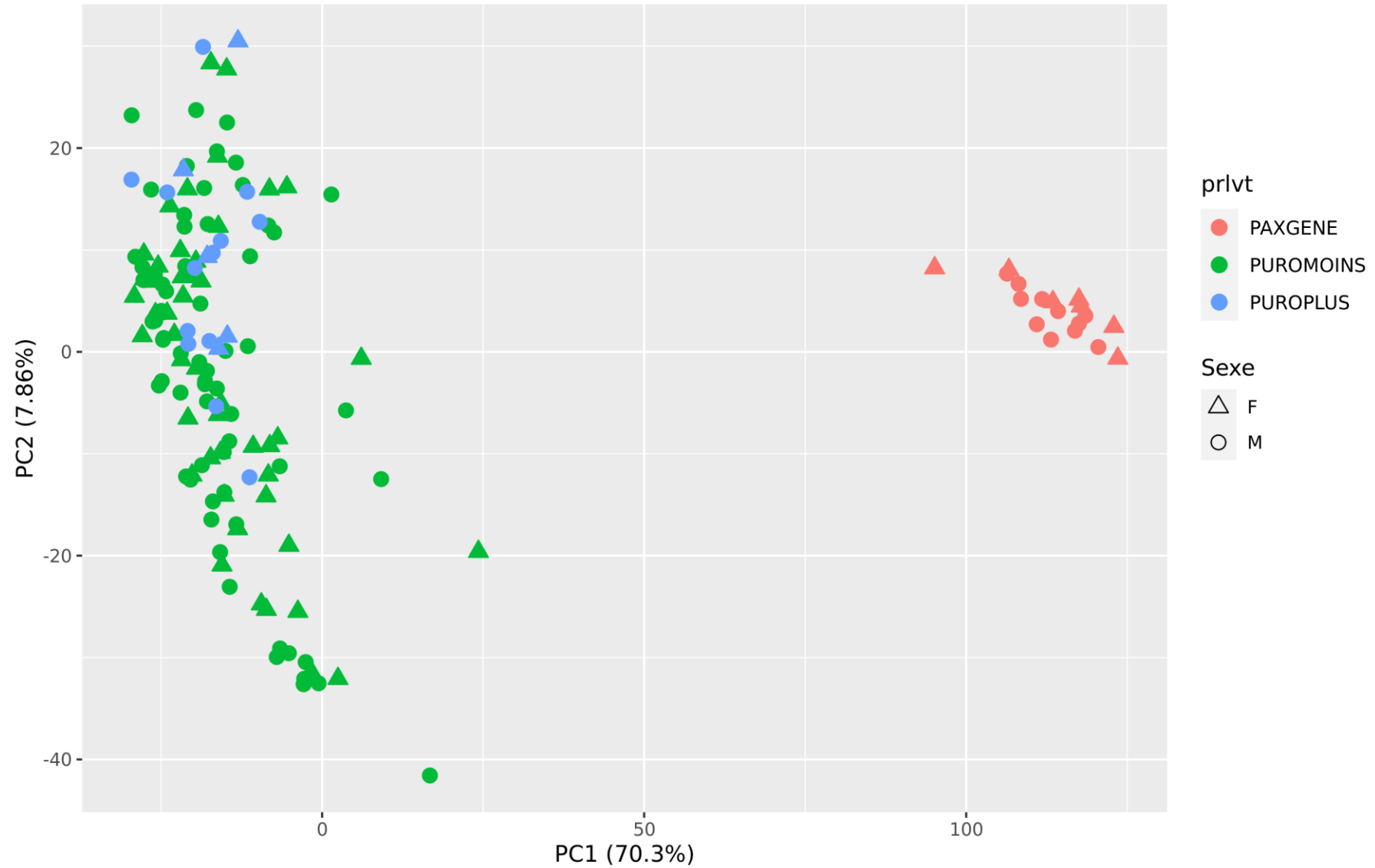


ACP à partir des comptages
Htseq-count avec DESeq2

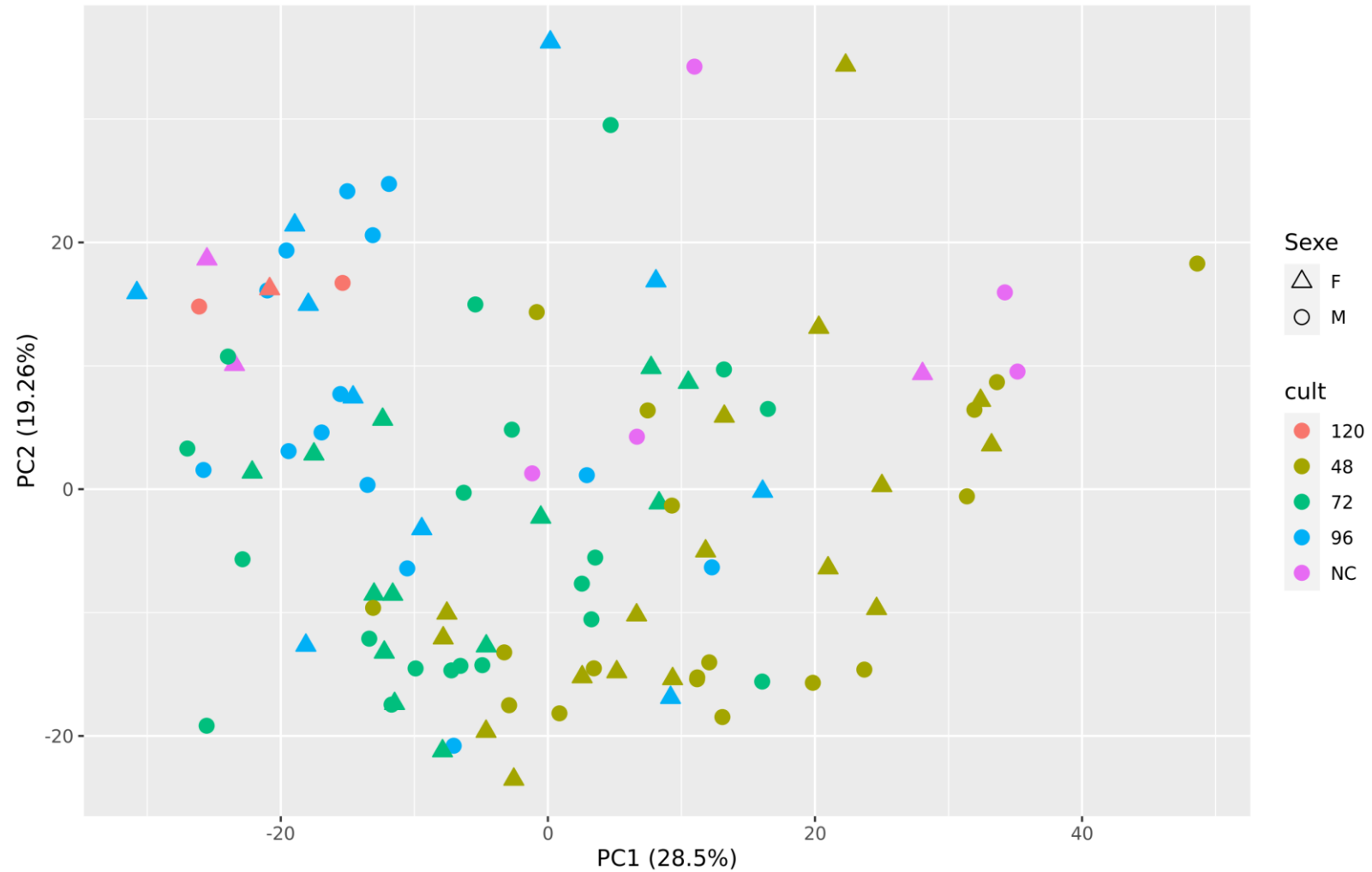
ACP : PAXgene vs cultures lymphocytaires



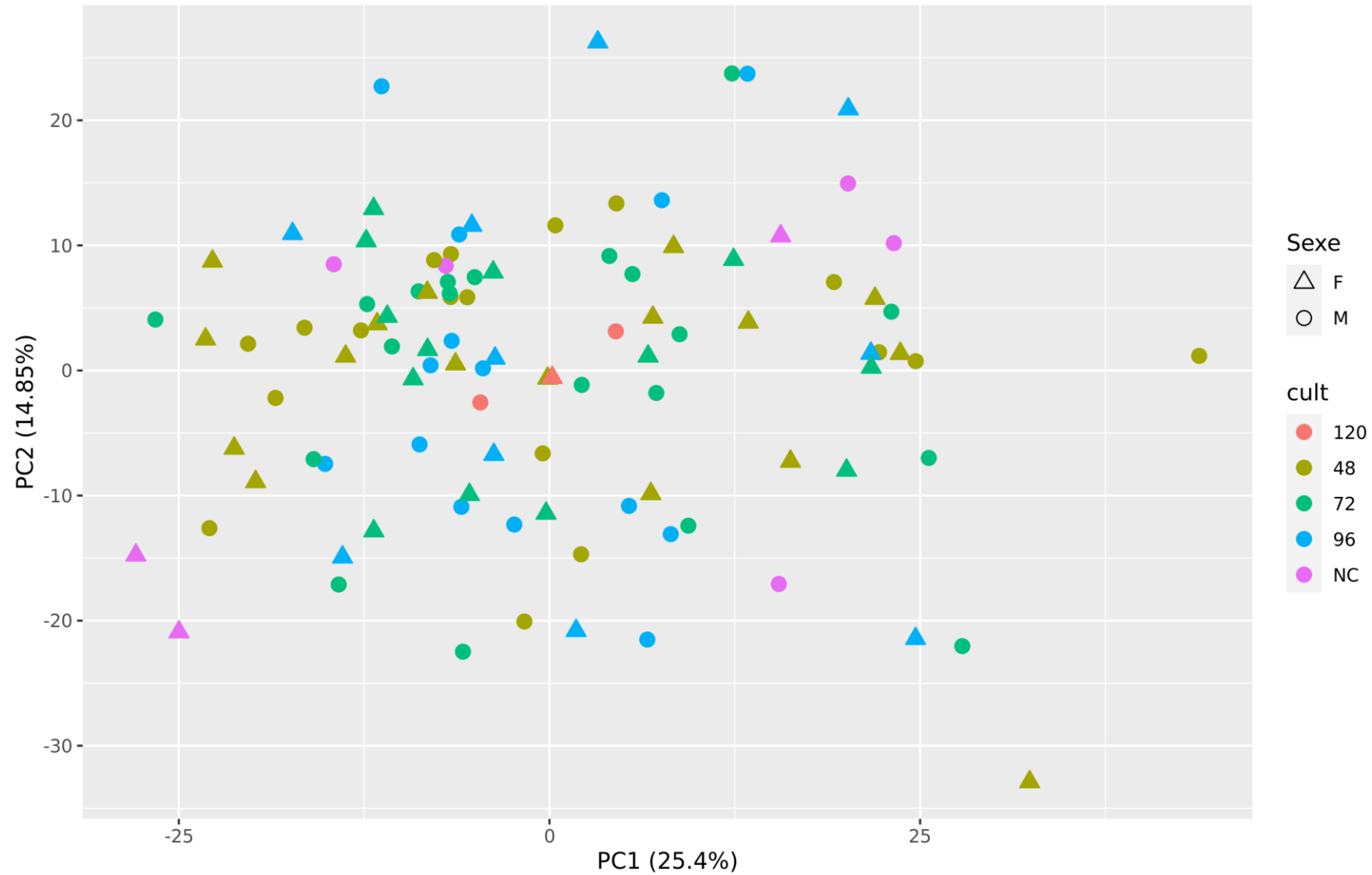
ACP : PAXgene vs cultures lymphocytaires (correction de l'effet sexe)



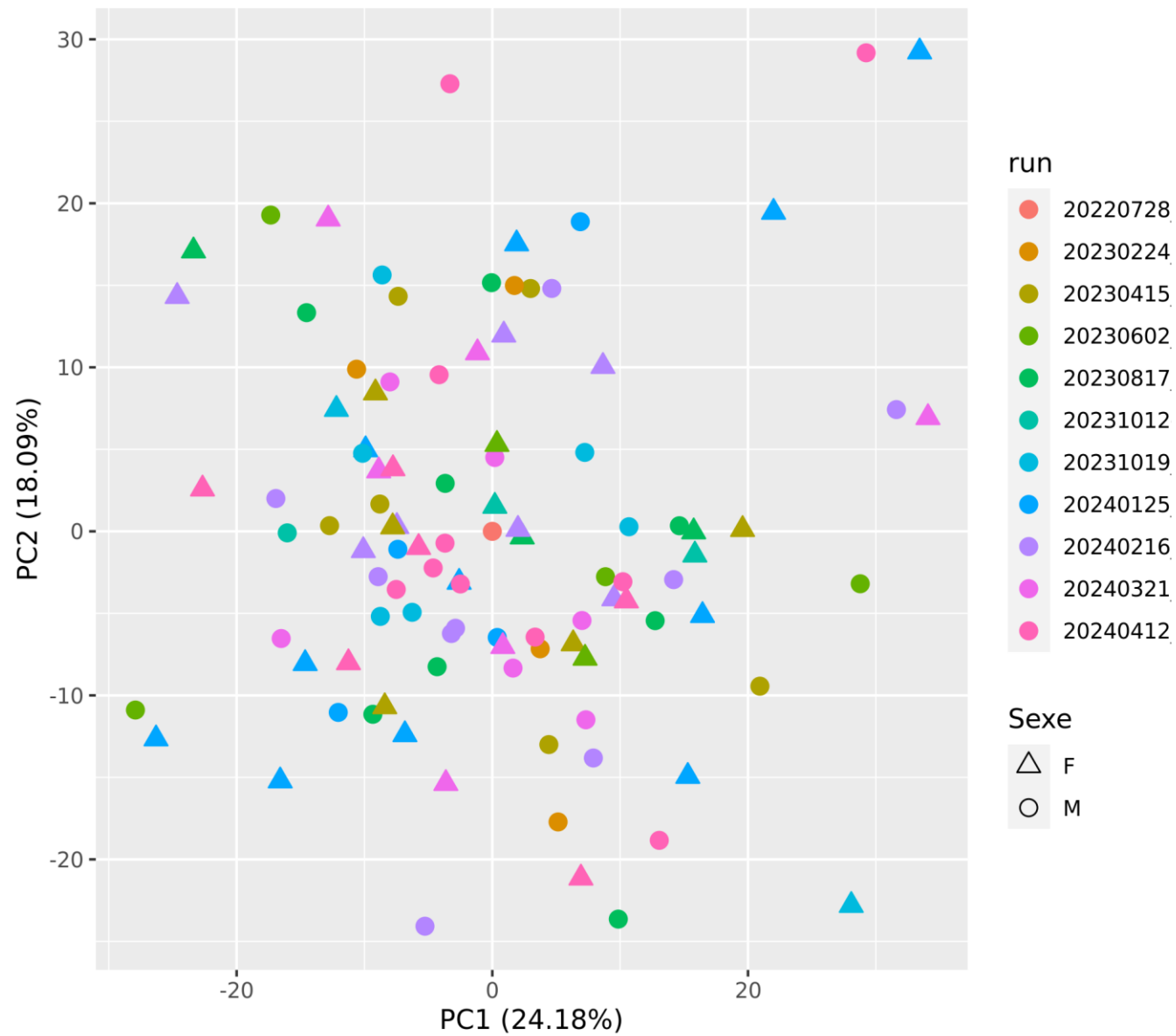
ACP : PAXgene vs cultures lymphocytaires (effet du temps de culture)



ACP : PAXgene vs cultures lymphocytaires (correction de l'effet batch)



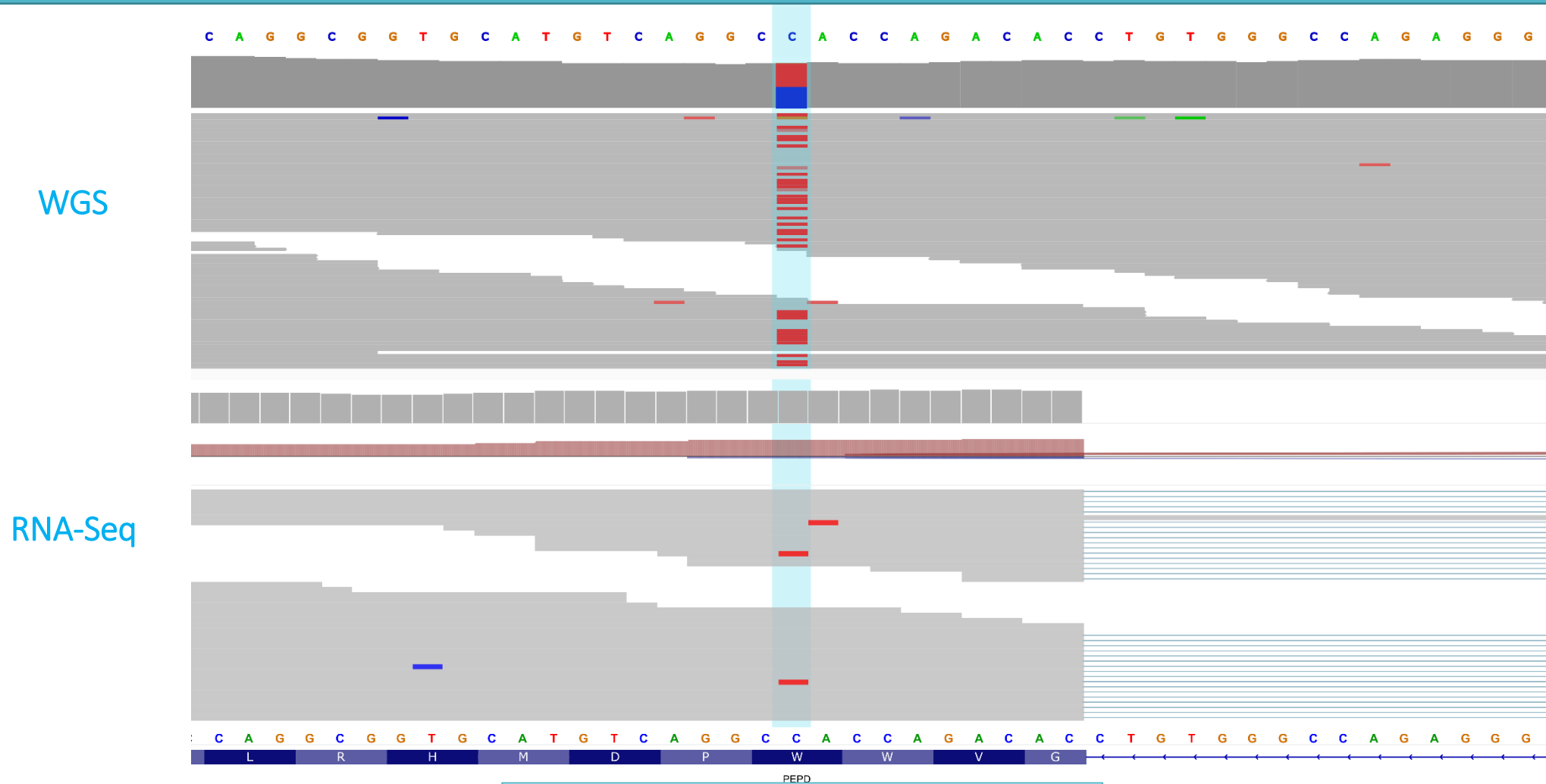
ACP : PAXgene vs cultures lymphocytaires (effet runs)



Corrections effets batch

- Effet sexe
- Effet du temps de culture
 - Pas d'effet de runs

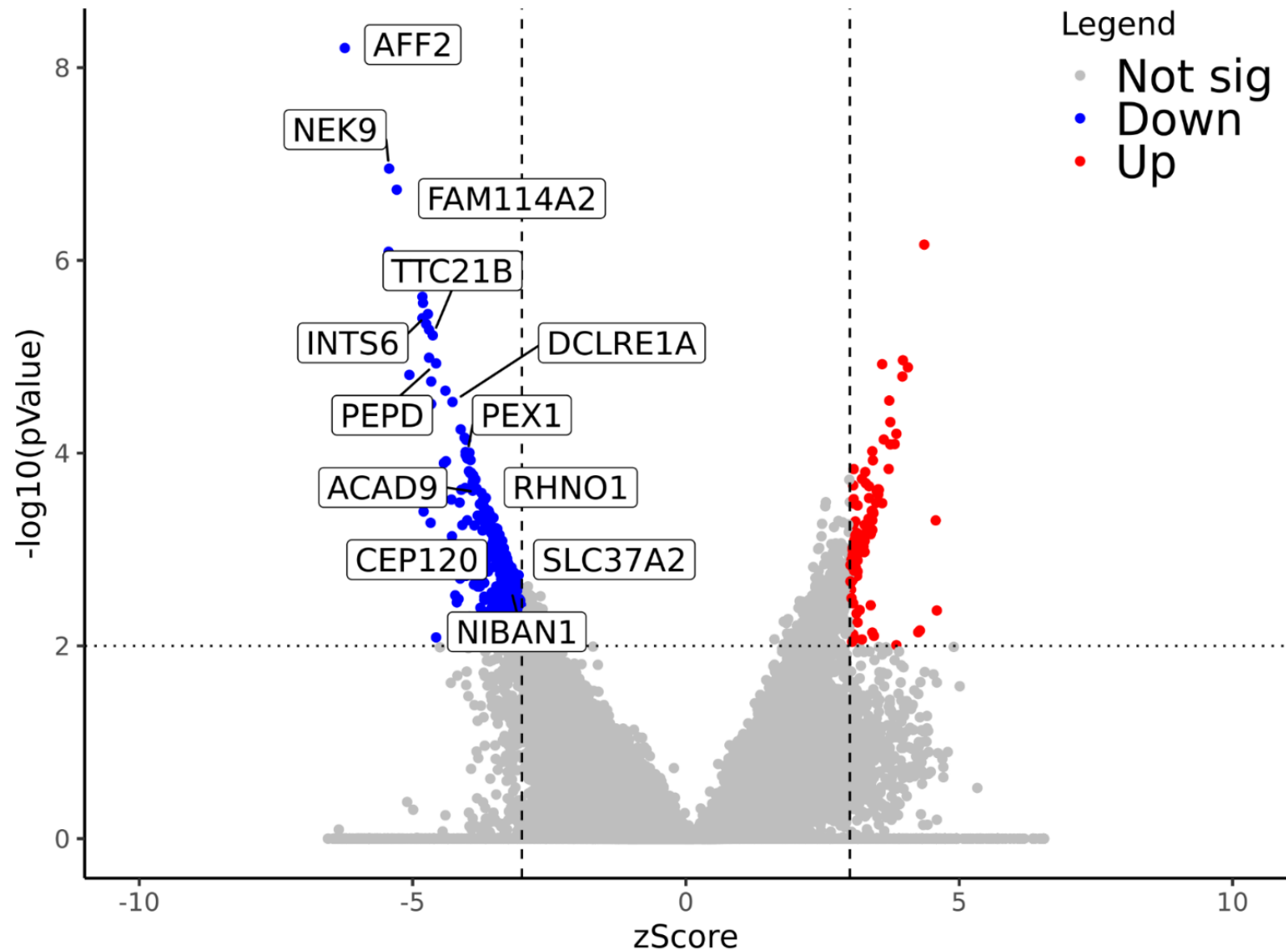
Recherche de témoins positifs d'expression aberrante



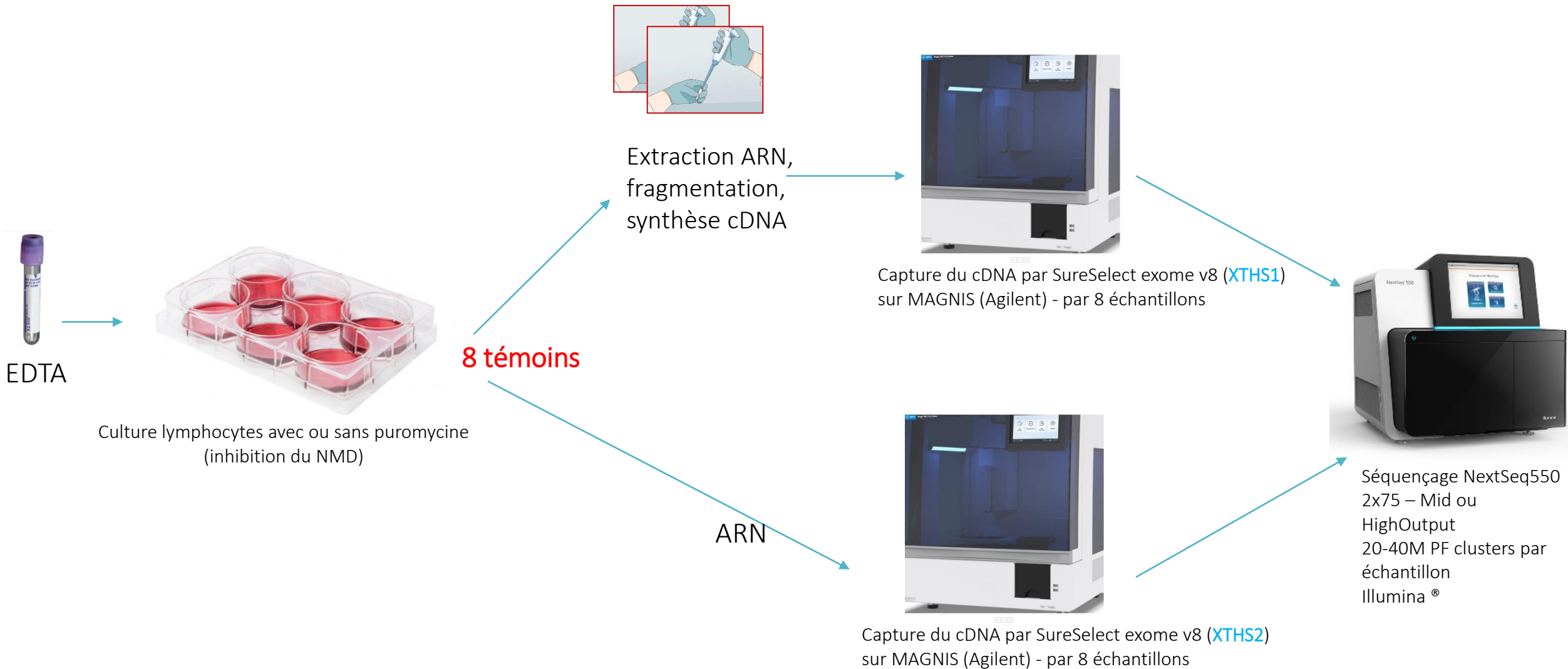
Perte de fonction hétérozygote (WGS)
Expression monoallélique (RNASeq)

Sélection de 8 témoins (XTHS1)

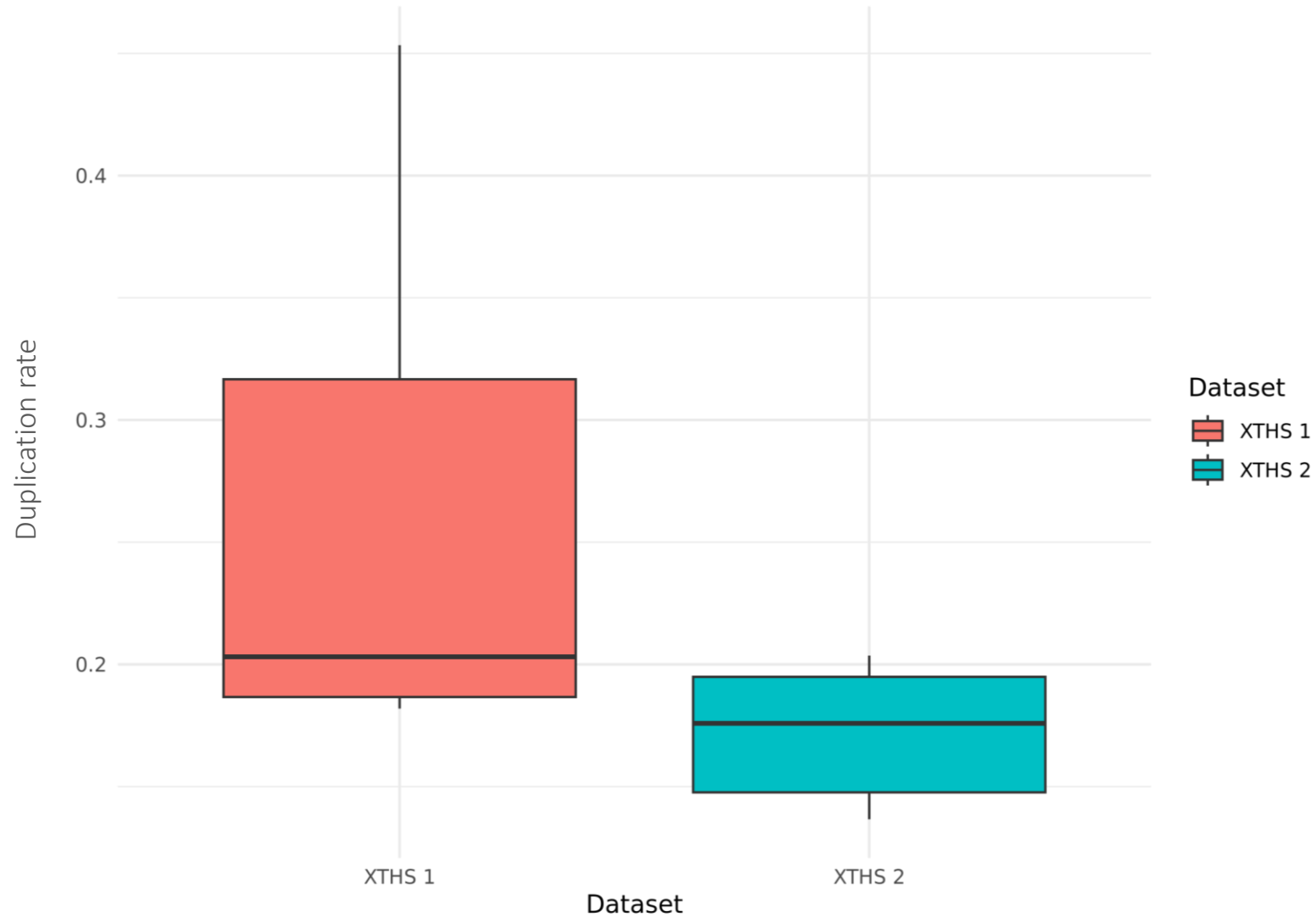
OUTRIDER: validation des 13 gènes témoins



Du sang au séquençage : étude du protocole XTHS2

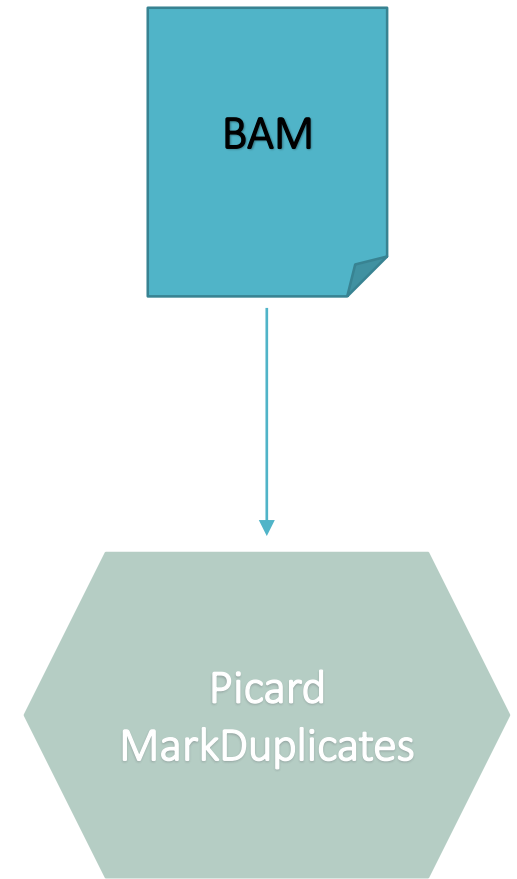


Taux de duplication plus bas en XTHS2

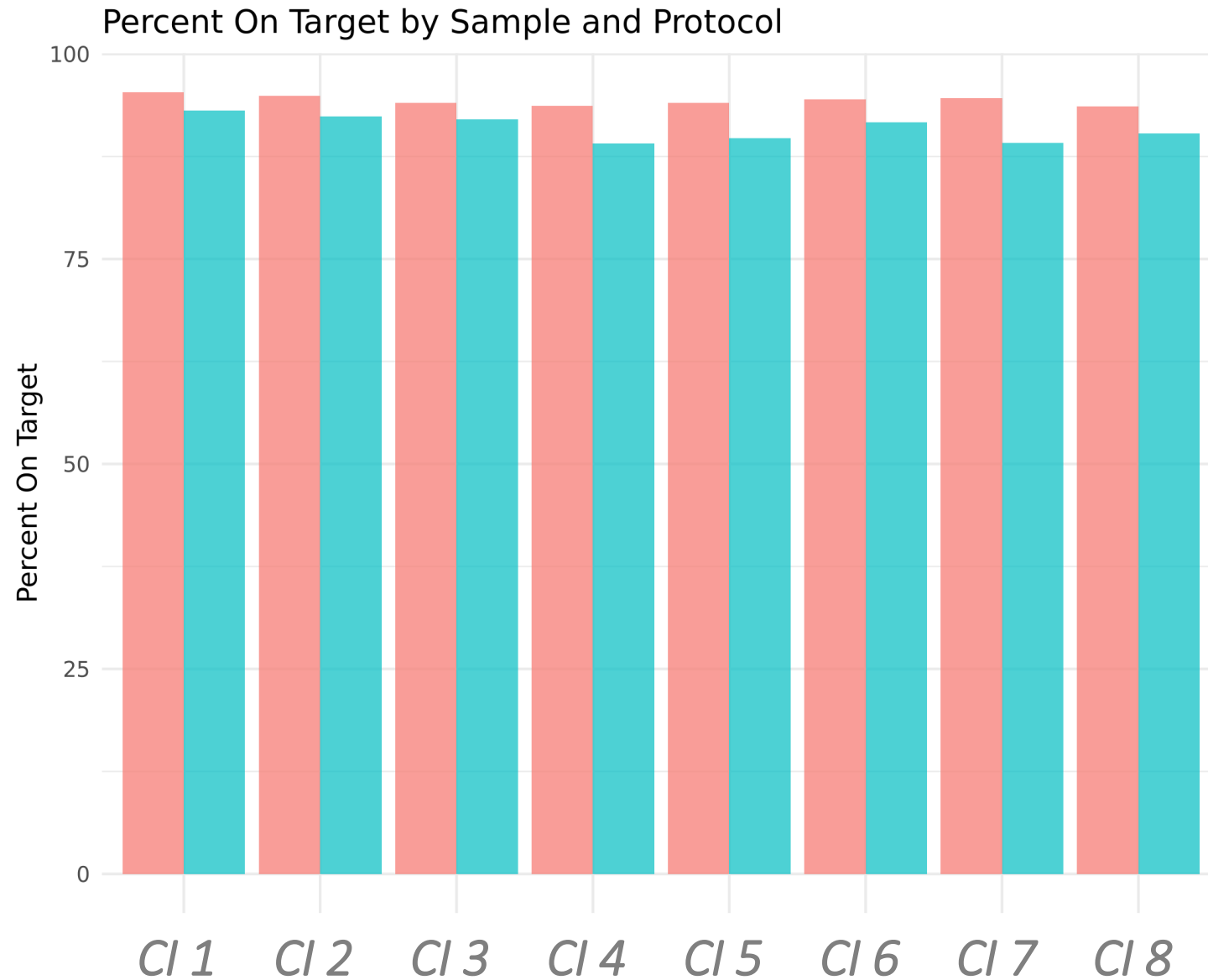


10 cycles PCR pré-capture
9 cycles PCR post-capture

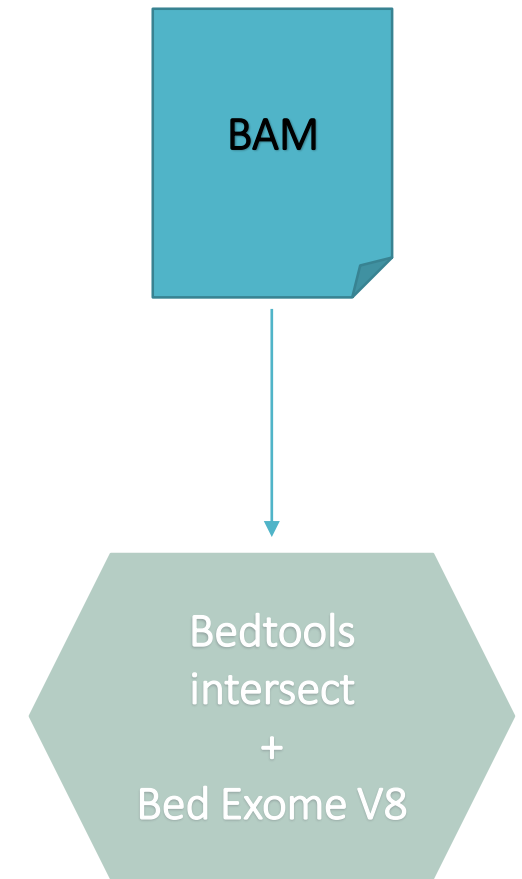
12 cycles PCR pré-capture
10 cycles PCR post-capture



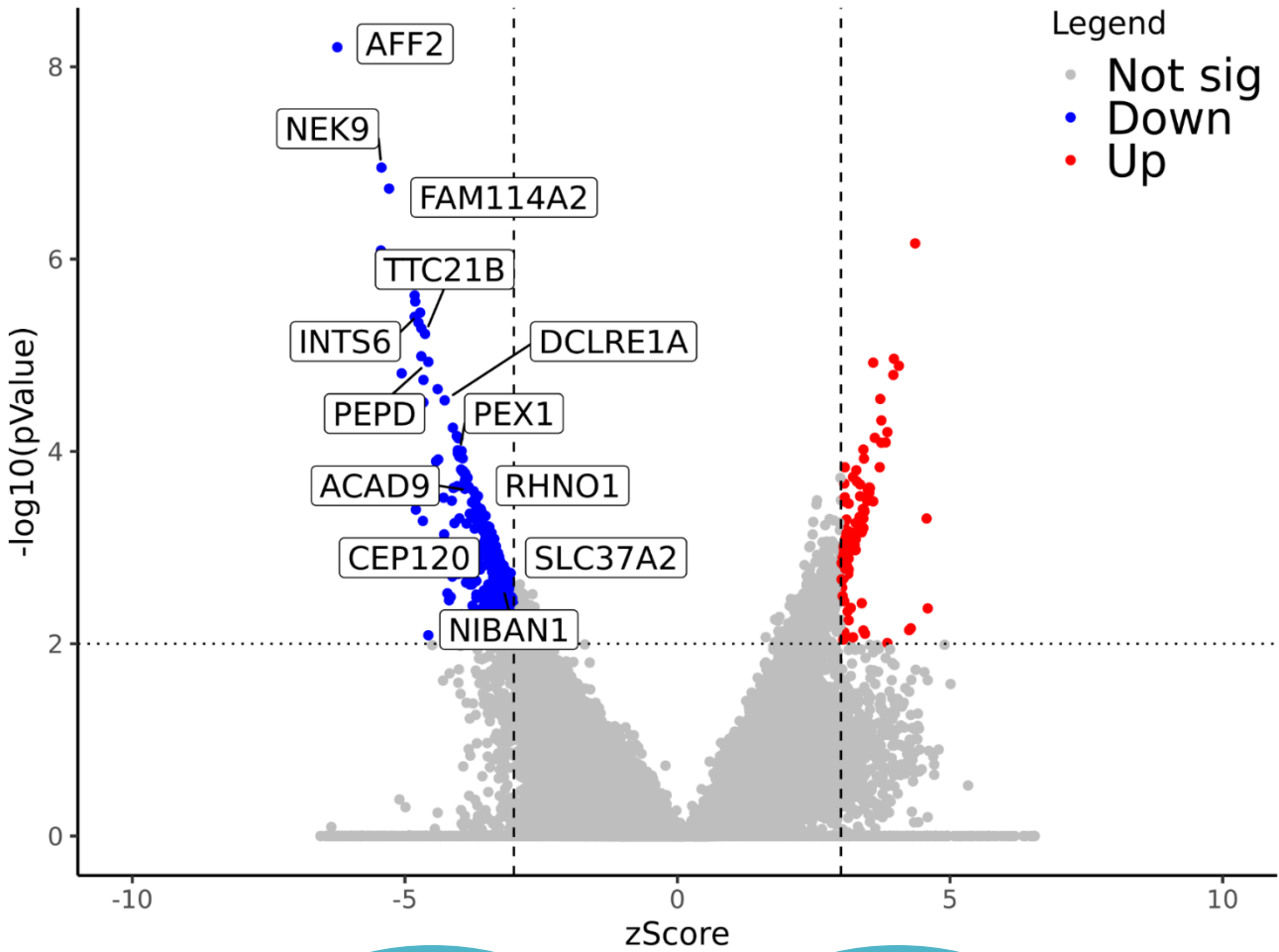
Pourcentage de couverture on target pour les 8 témoins



Protocol
XTHS1
XTHS2



Validation de témoins positifs en XTHS2

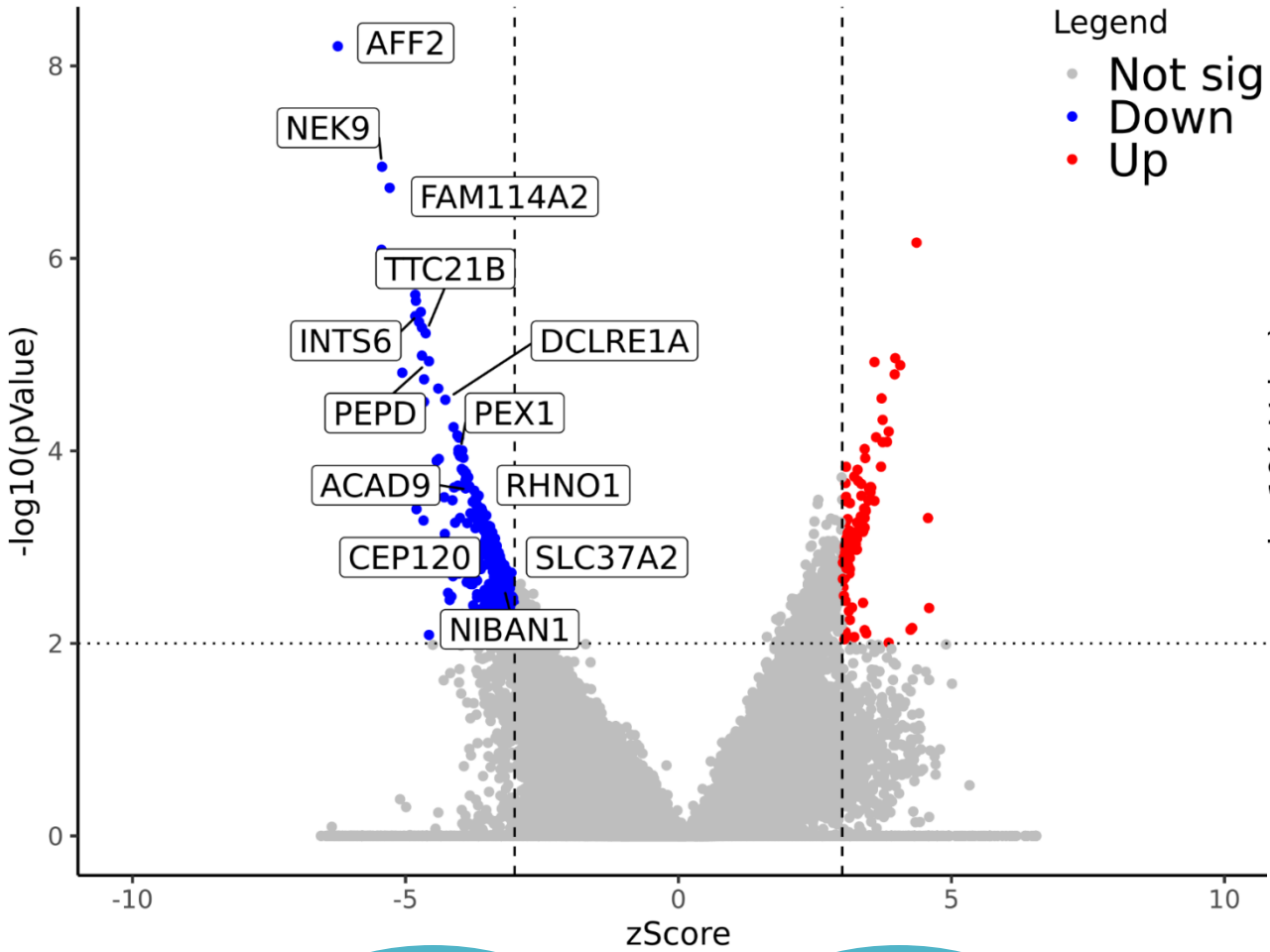


contrôles
XTHS1

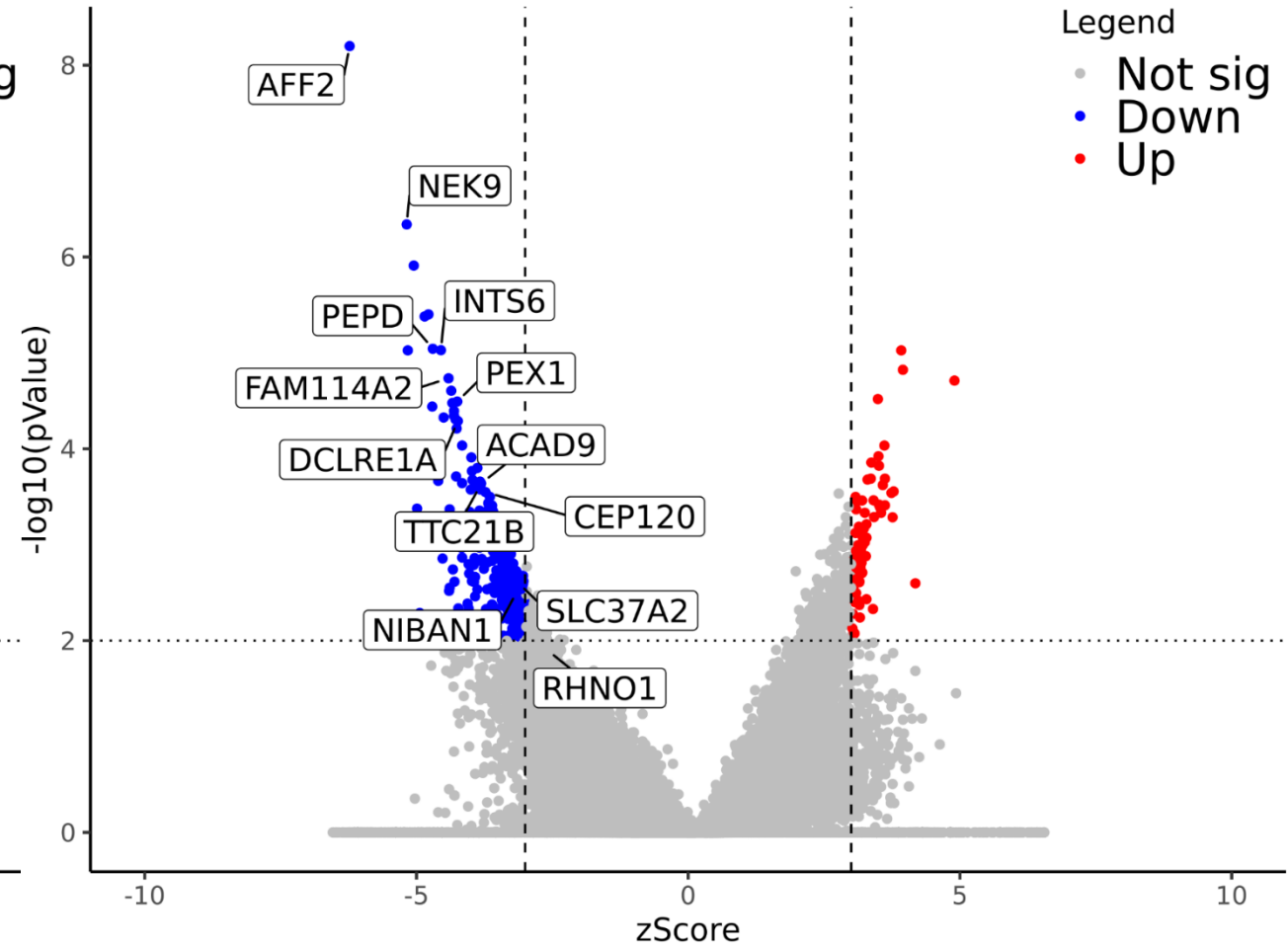
+

8 T+
XTHS1

Validation de témoins positifs en XTHS2



contrôles XTHS1 + 8 T+ XTHS1



contrôles XTHS1 + 8 T+ XTHS2

Conclusions : OUTRIDER

Stratégie RNA-seq capture exome sur culture lymphocytaire

→ permet d'être quantitatif

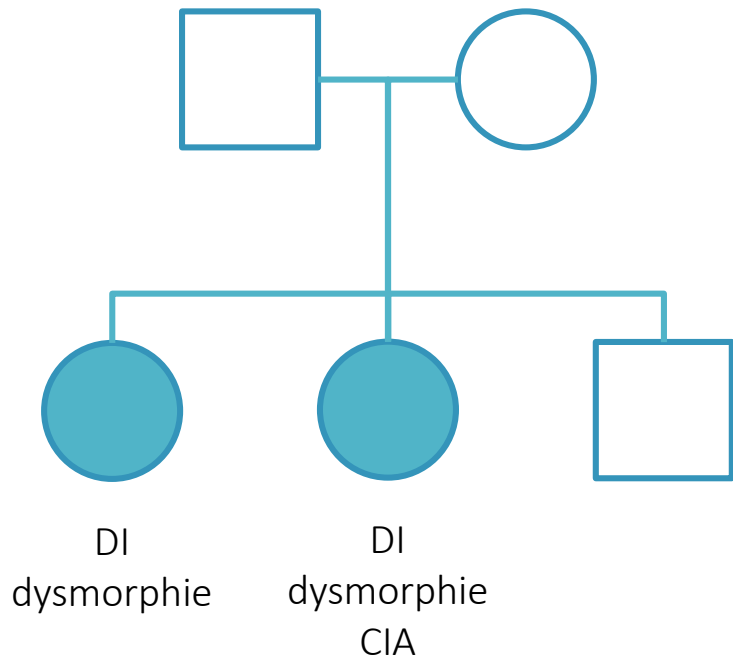
Utilisation d'une cohorte XTHS1 et XTHS2

→ XTHS2 : moins de duplicats, moins d'étapes manuelles, possibilité d'utiliser avec une cohorte XTHS1

→ Tests UMI en cours

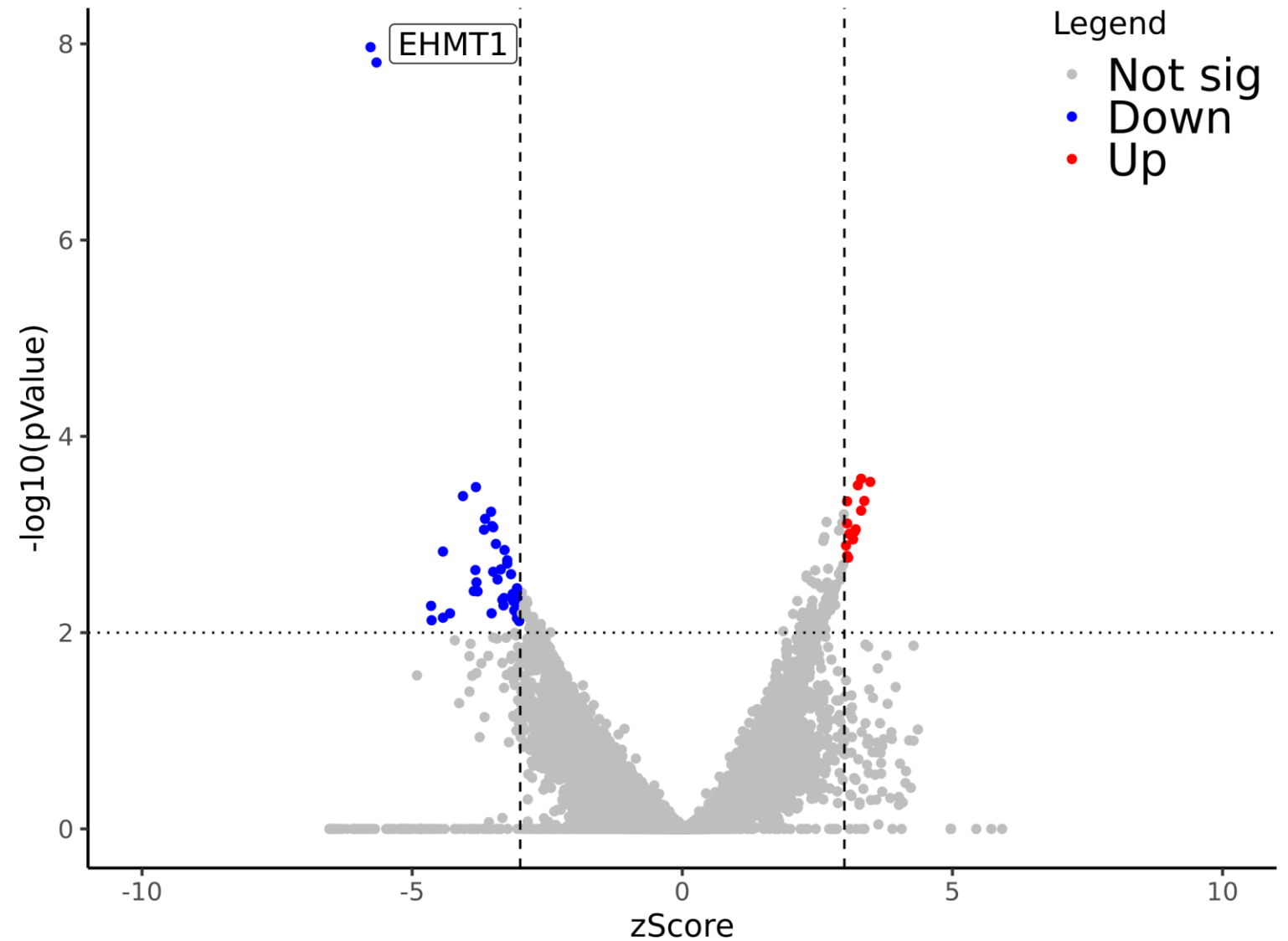
RNA-seq : Résolution d'un cas clinique

Famille suivie par B.Isidor (Nantes)

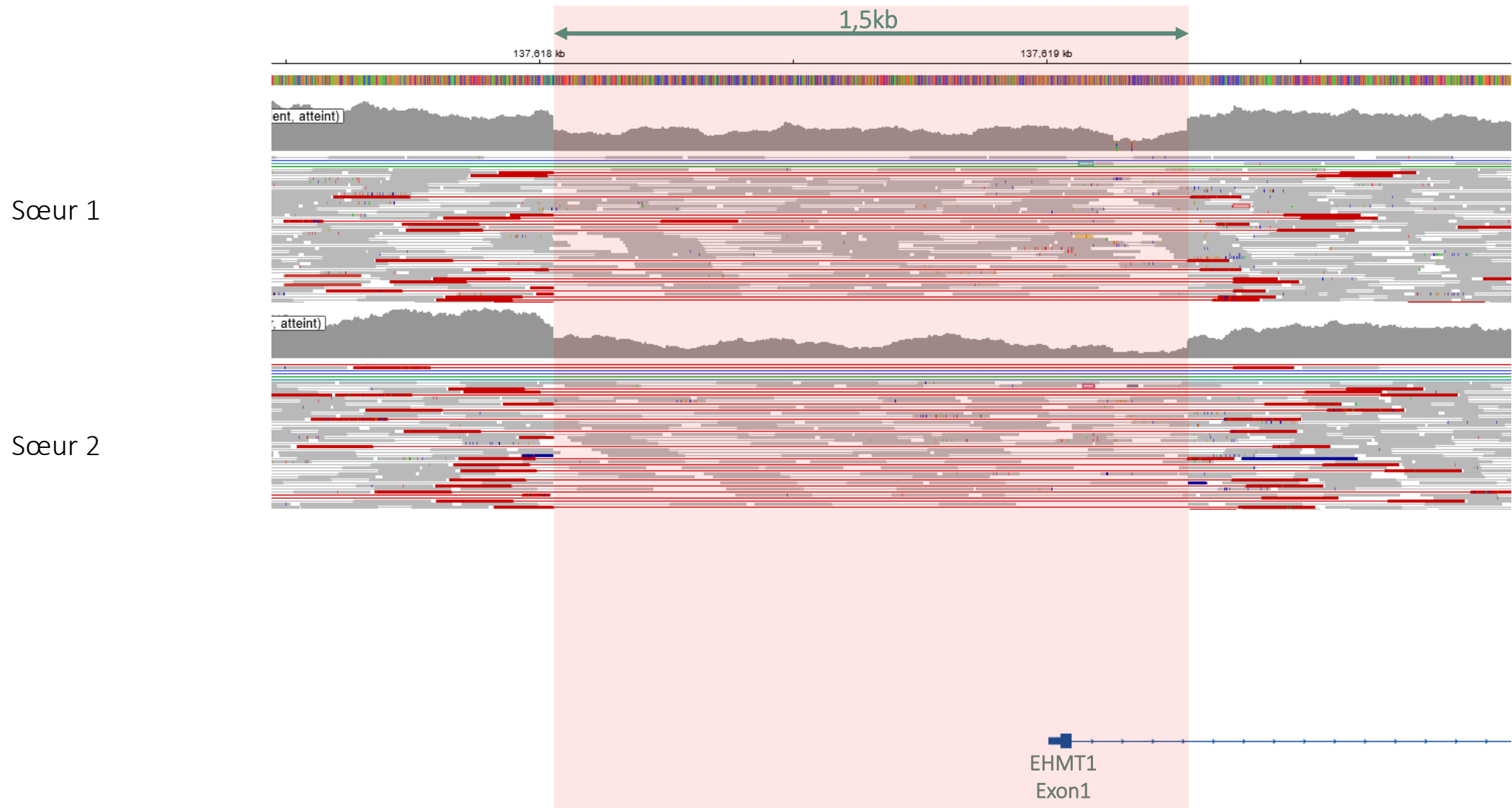


Exome négatif 2020

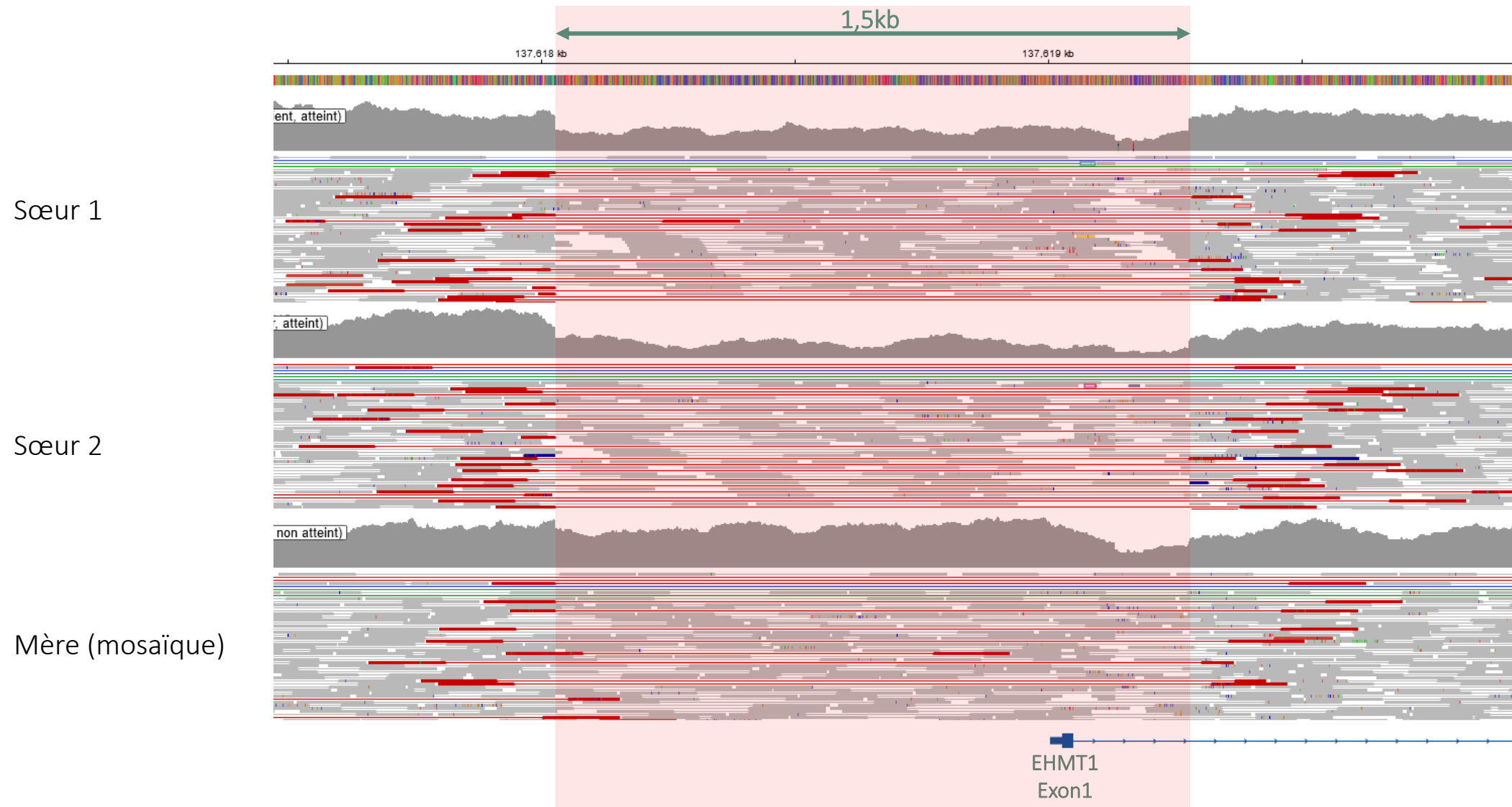
Génome quatuor
SeqOIA négatif 2022



Délétion du 1^{er} exon EHMT1



Délétion du 1^{er} exon EHMT1

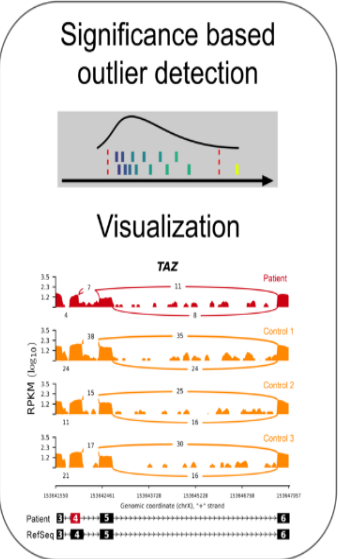
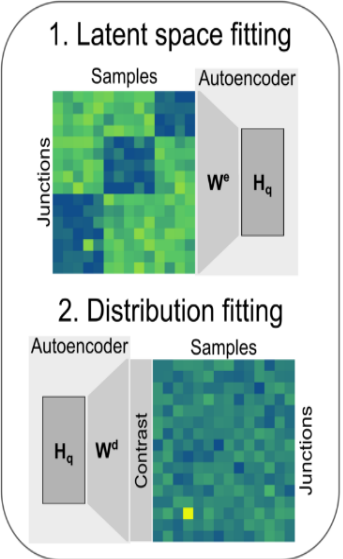
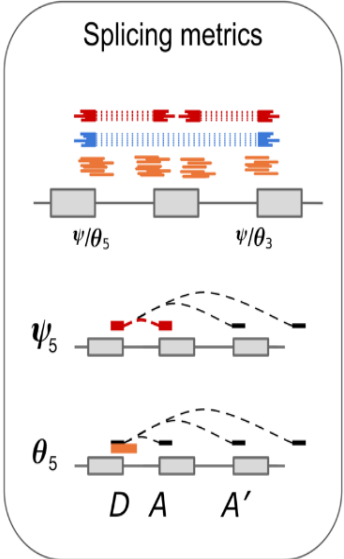
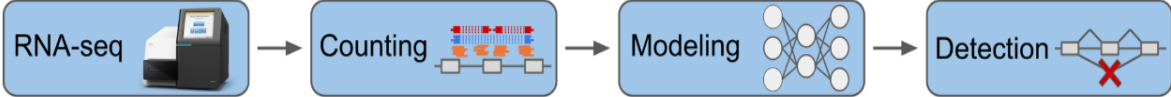


FRASER



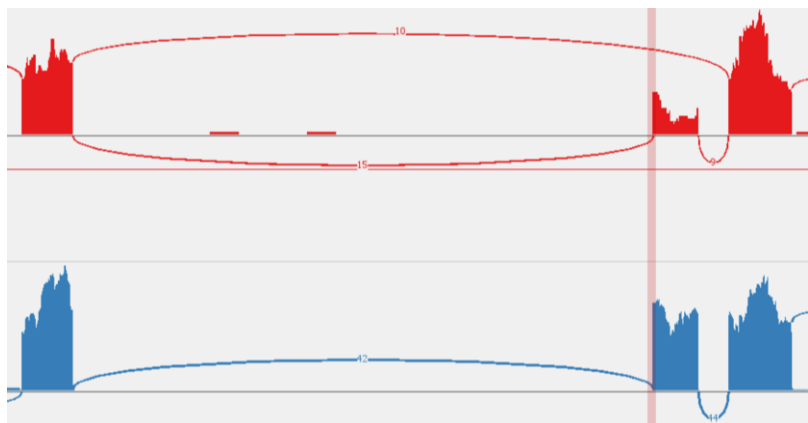
Detection of aberrant splicing events in RNA-seq data using FRASER

Christian Mertes ^{1,6}, Ines F. Scheller ^{1,2,6}, Vicente A. Yépez ^{1,3}, Muhammed H. Çelik ¹, Yingjiqiong Liang ¹, Laura S. Kremer ^{4,5}, Mirjana Gusic ^{4,5}, Holger Prokisch ^{4,5} & Julien Gagneur ^{1,2,5}

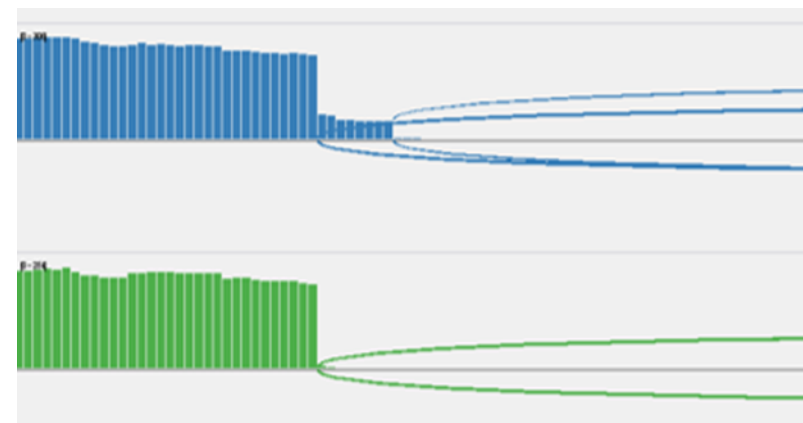


Validation des différents effets de variants sur l'épissage

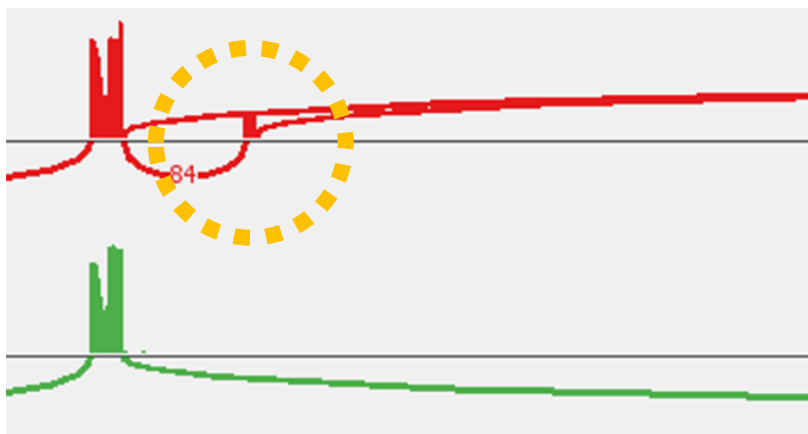
Saut d'exon



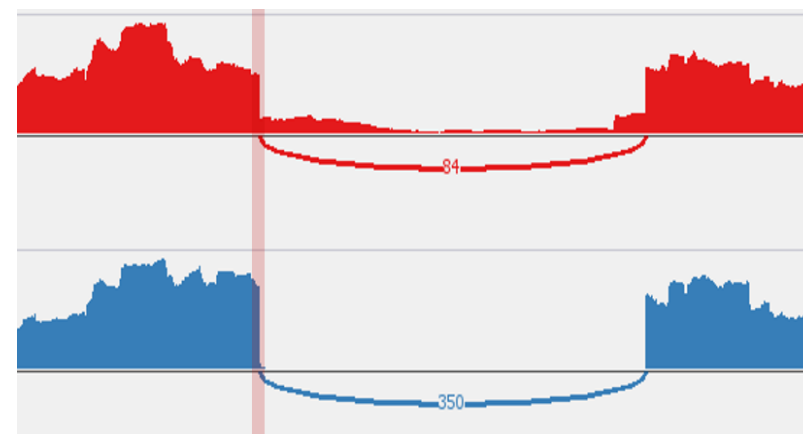
Site cryptique



Pseudo exon



Rétention intronique

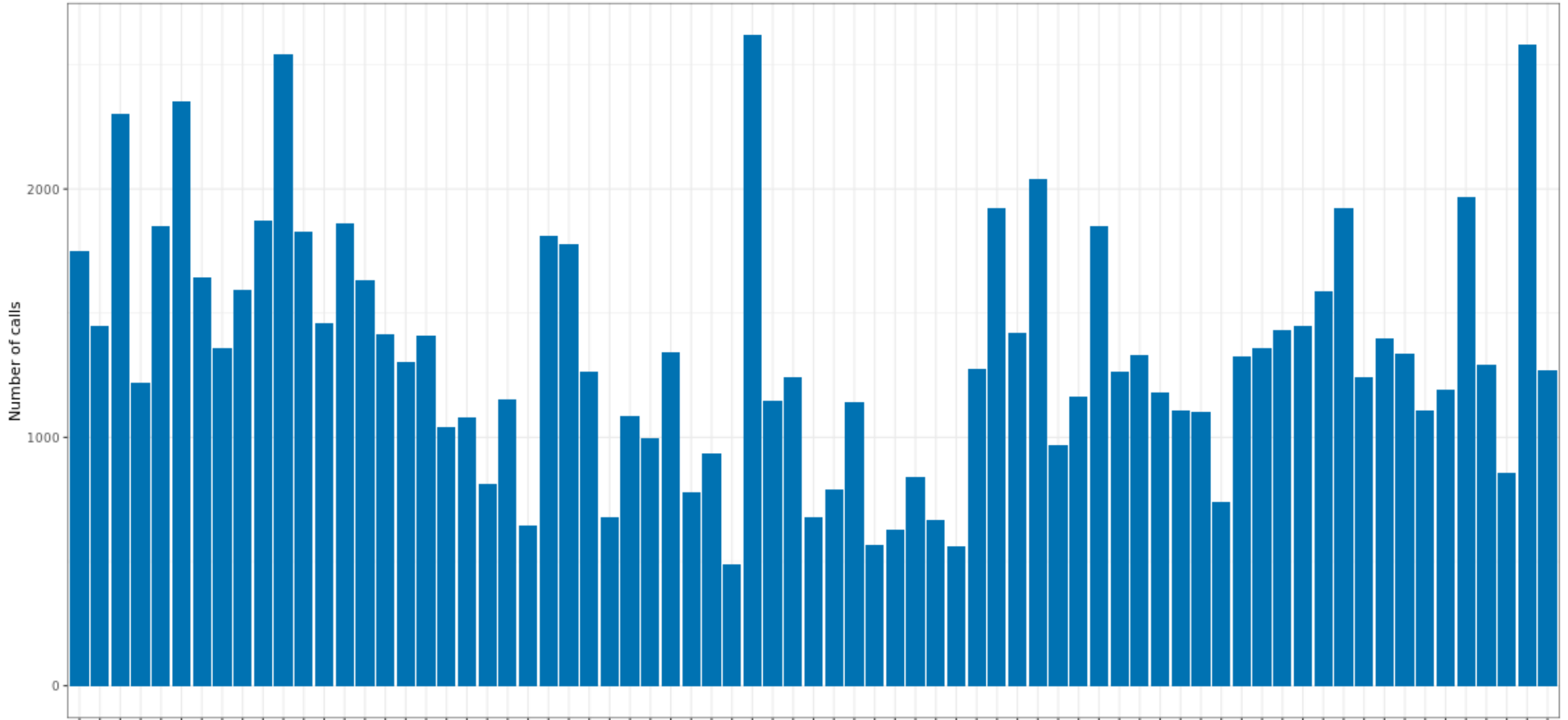


Confirmation de l'effet pour 8 variants

Les algorithmes permettent-ils d'identifier et prioriser ces évènements ?

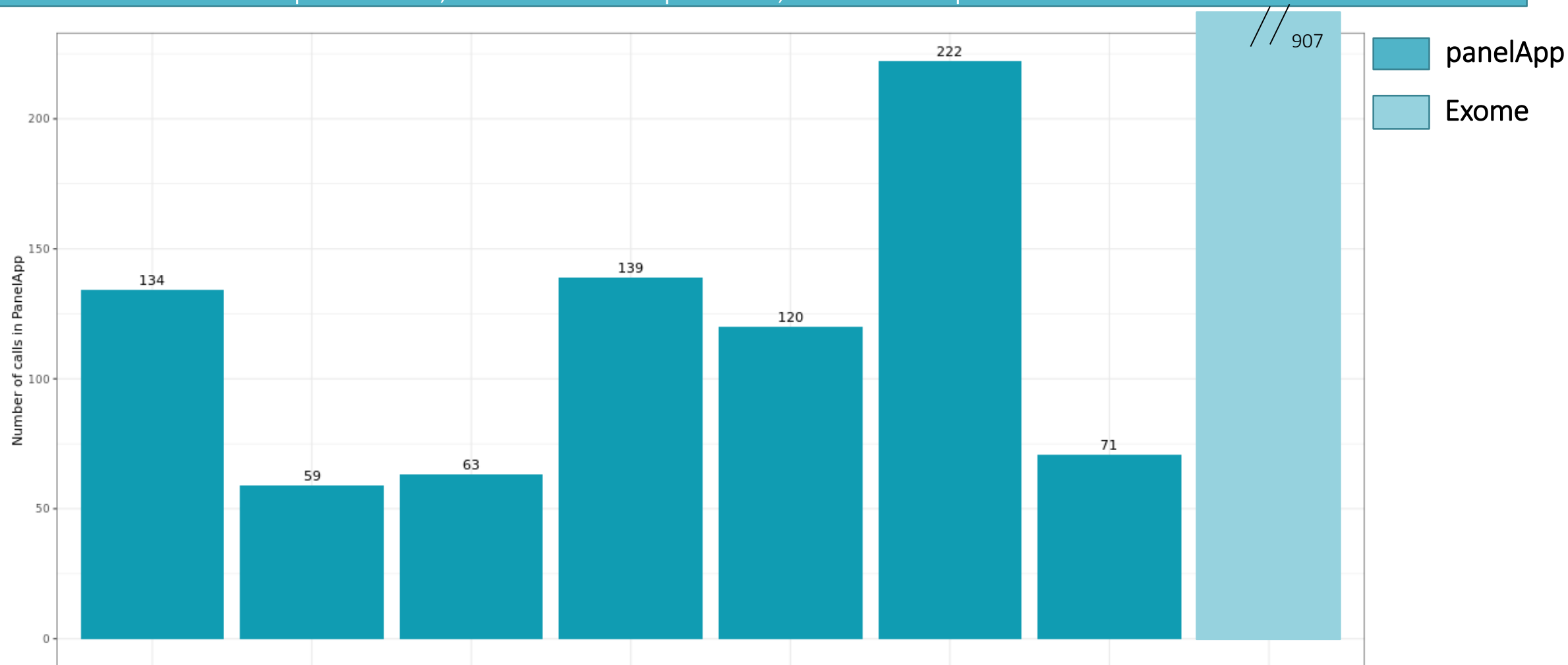
Priorisation FRASER : Nombre d'évènements aberrants d'épissage

pValue < 0,001 cohorte 73 patients



Priorisation FRASER : Nombre d'évènements aberrants d'épissage

pValue < 0,001 cohorte 73 patients, focus sur 8 patients témoins



Variant pathogène :

CI
ADNP

CI
DYNC1H1

CI
FOXP1

CI
KDM5C

CI
PSMC2

CI
USP9X

CI
TRIO

CI
INTS6

Rang :

126

7

57

1

1

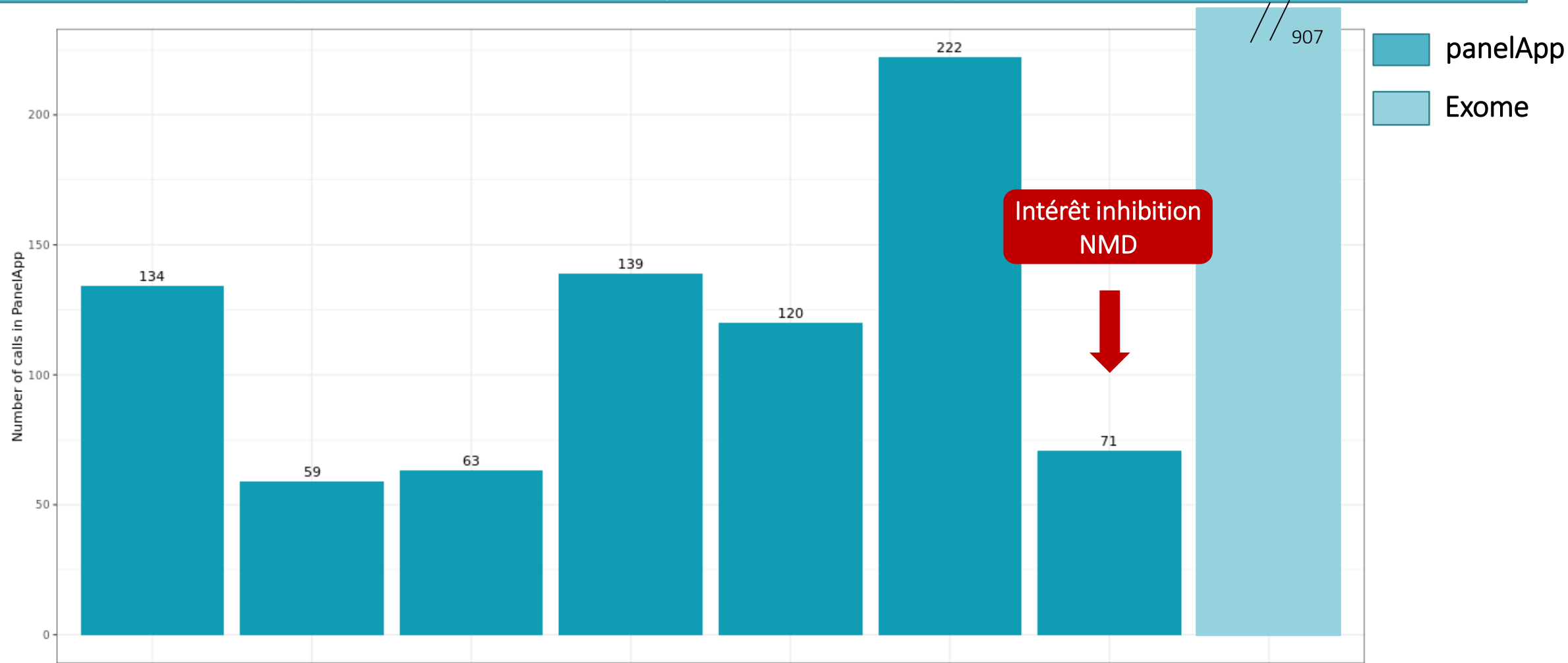
86

0

654

Priorisation FRASER : Nombre d'évènements aberrants d'épissage

pValue < 0,001 cohorte 73 patients, focus sur 8 patients témoins



Variant pathogène :

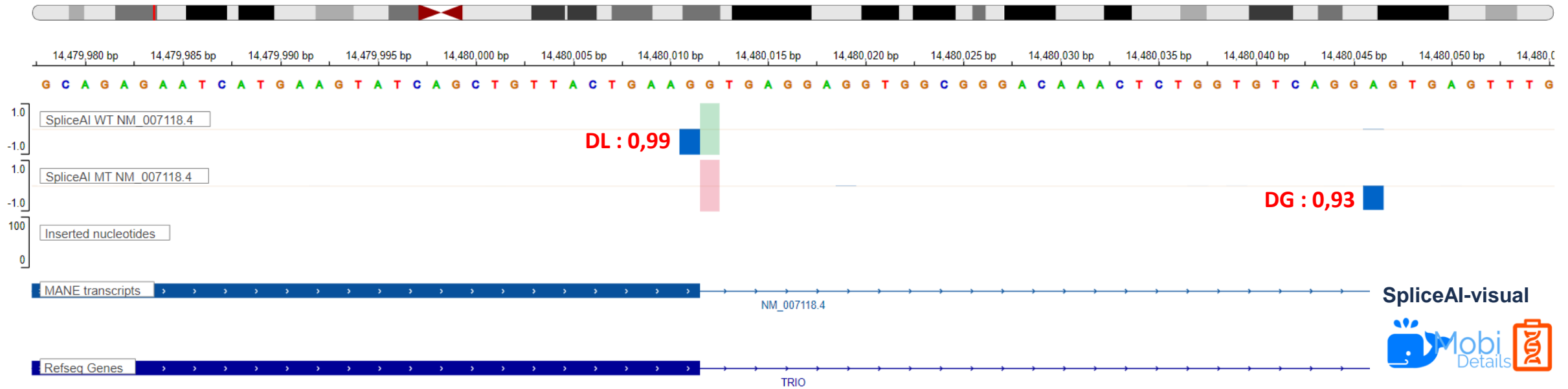
CI
ADNP *CI*
DYNC1H1 *CI*
FOXP1 *CI*
KDM5C *CI*
PSMC2 *CI*
USP9X *CI*
TRIO *CI*
INTS6

Rang :

126 7 57 1 1 86 0 654

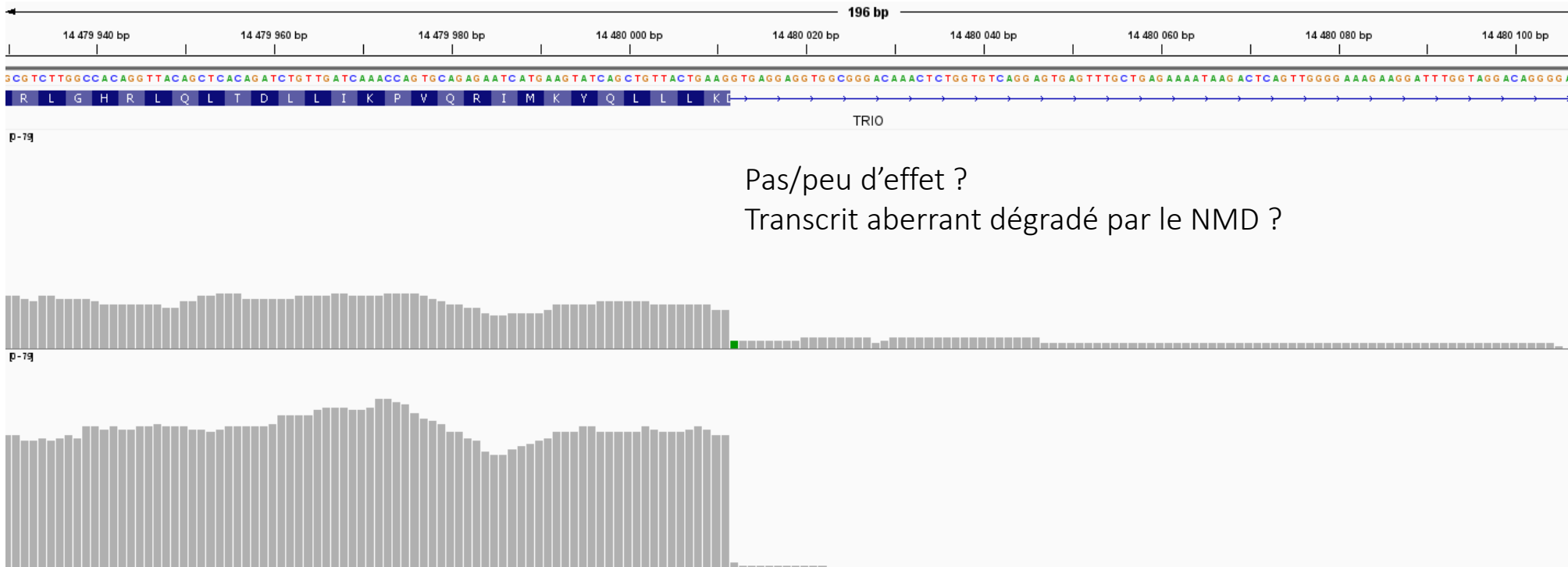
Intérêt de l'inhibition du NMD (traitement puromycine)

TRIO - NM_007118.4 - c.6336+1G>A



Intérêt de l'inhibition du NMD (traitement puromycine)

TRIO - NM_007118.4 - c.6336+1G>A

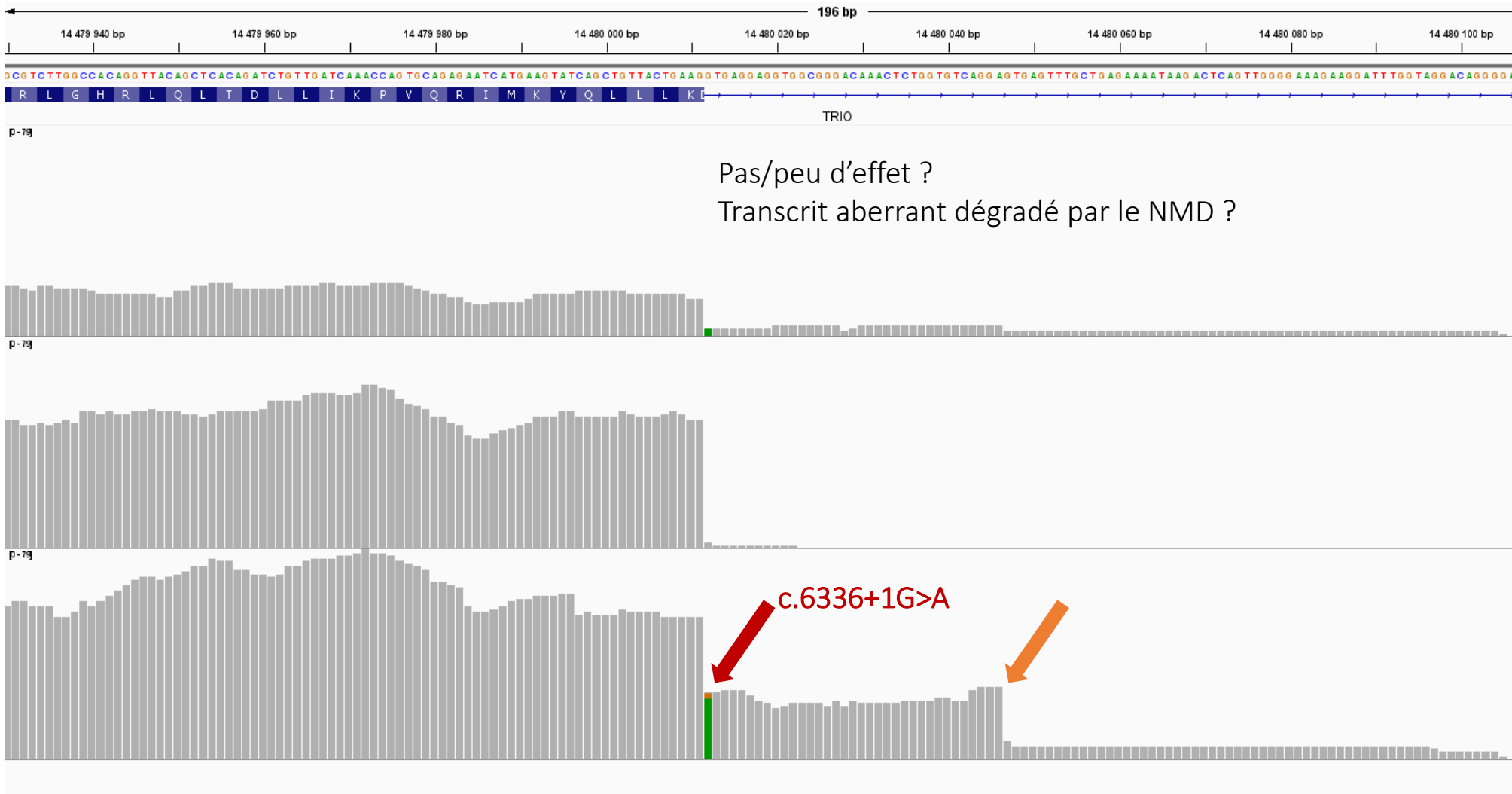


Cas index
sans puromycine

Témoin

Intérêt de l'inhibition du NMD (traitement puromycine)

TRIO - NM_007118.4 - c.6336+1G>A



Pas/peu d'effet ?
Transcrit aberrant dégradé par le NMD ?

Cas index
sans puromycine

Témoïn

Cas index
traité puromycine

Conclusions : FRASER

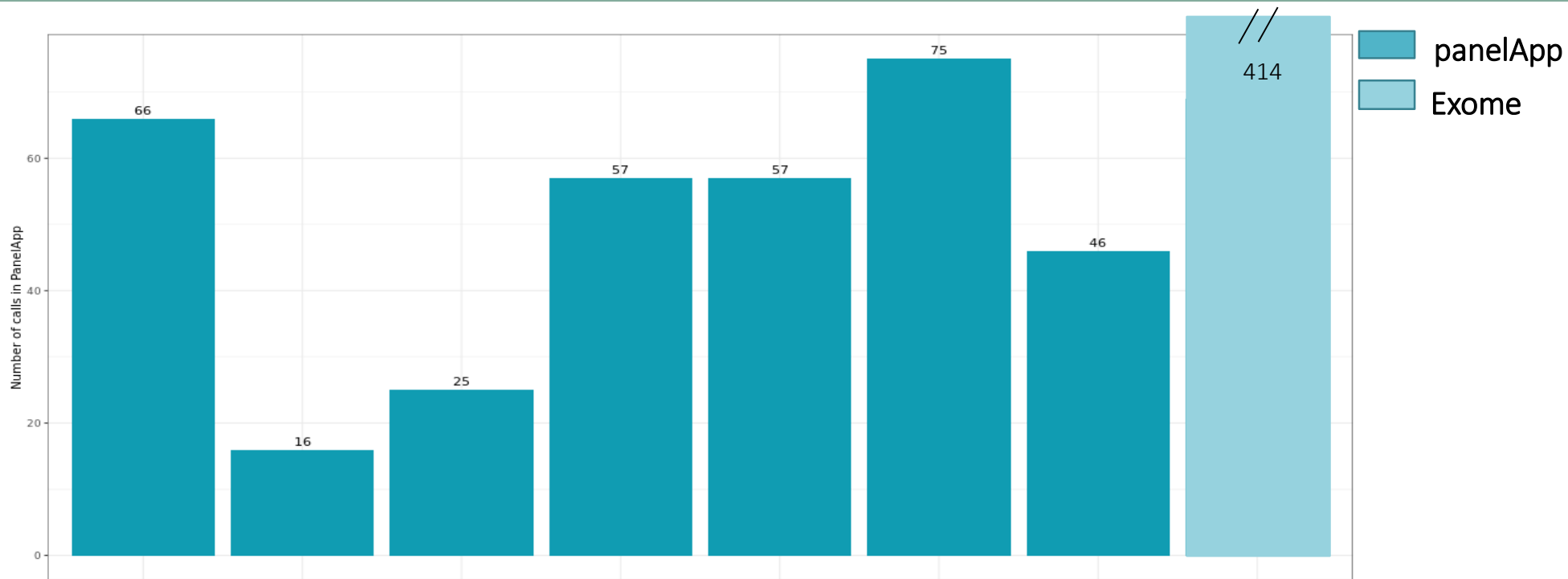
Validation de 7 sur 8 témoins d'épissage

Remarque : plus de bruits quand la cohorte augmente

Conclusions : FRASER

Validation de 7 sur 8 témoins d'épissage

Remarque : plus de bruits quand la cohorte augmente



Variant

pathogène :

Rang :

CI

ADNP

4 (126)

CI

DYNC1H1

7 (7)

CI

FOXP1

1 (57)

CI

KDM5C

7 (1)

CI

PSMC2

3 (1)

CI

USP9X

1 (86)

CI

TRIO

0 (0)

CI

INTS6

315 (654)

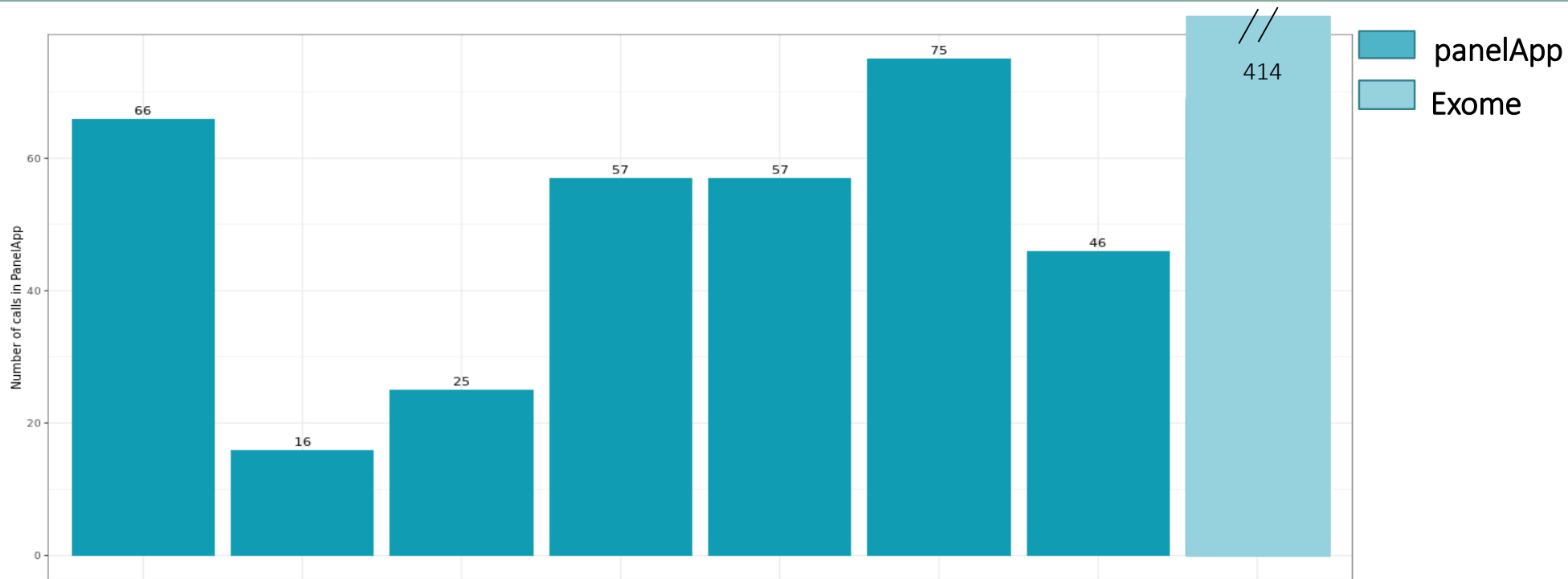
Rouge : cohorte 45 échantillons

Gris : cohorte 73 échantillons

Conclusions : FRASER

Validation de 7 sur 8 témoins d'épissage

Remarque : plus de bruits quand la cohorte augmente



Variant pathogène :

Rang :

CI
ADNP

4 (126)

CI
DYNC1H1

7 (7)

CI
FOXP1

1 (57)

CI
KDM5C

7 (1)

CI
PSMC2

3 (1)

CI
USP9X

1 (86)

CI
TRIO

0 (0)

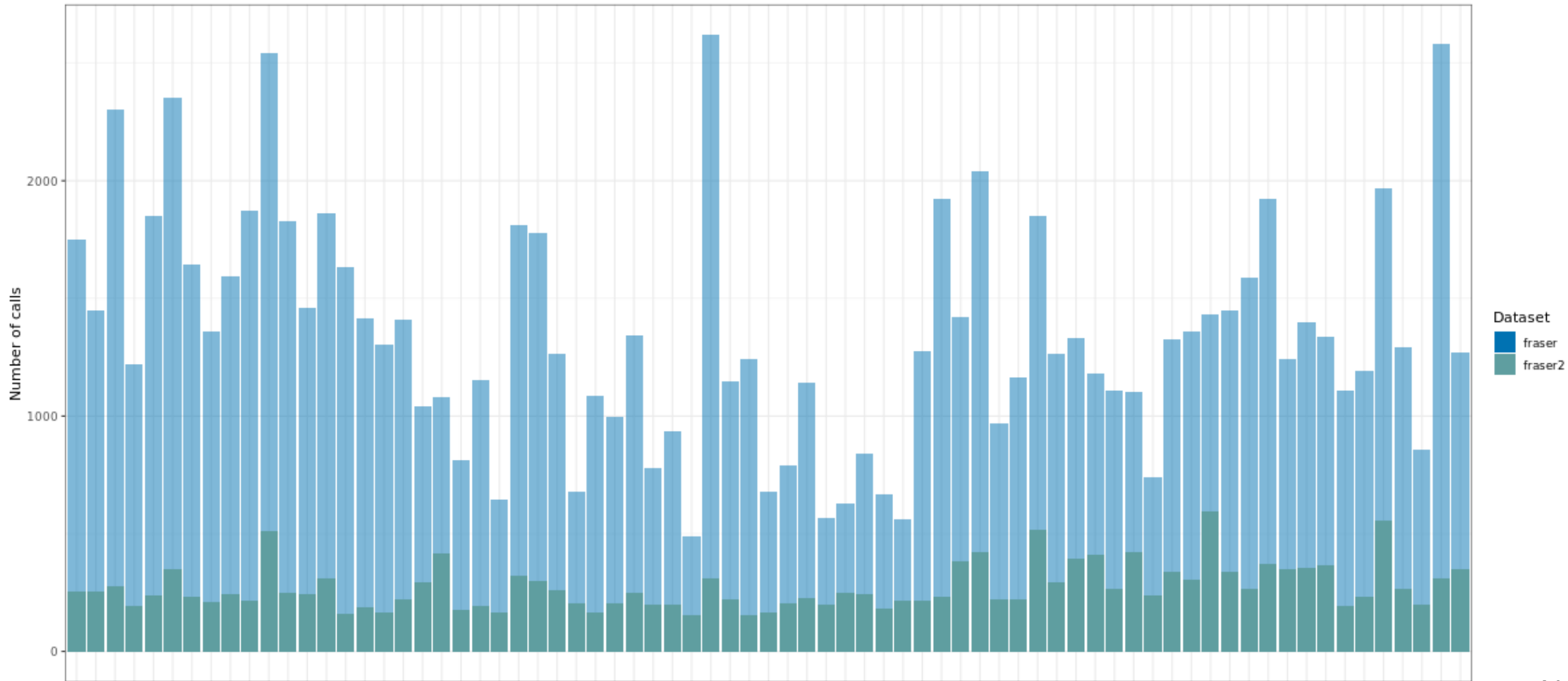
CI
INTS6

315 (654)

ET FRASER2 ?

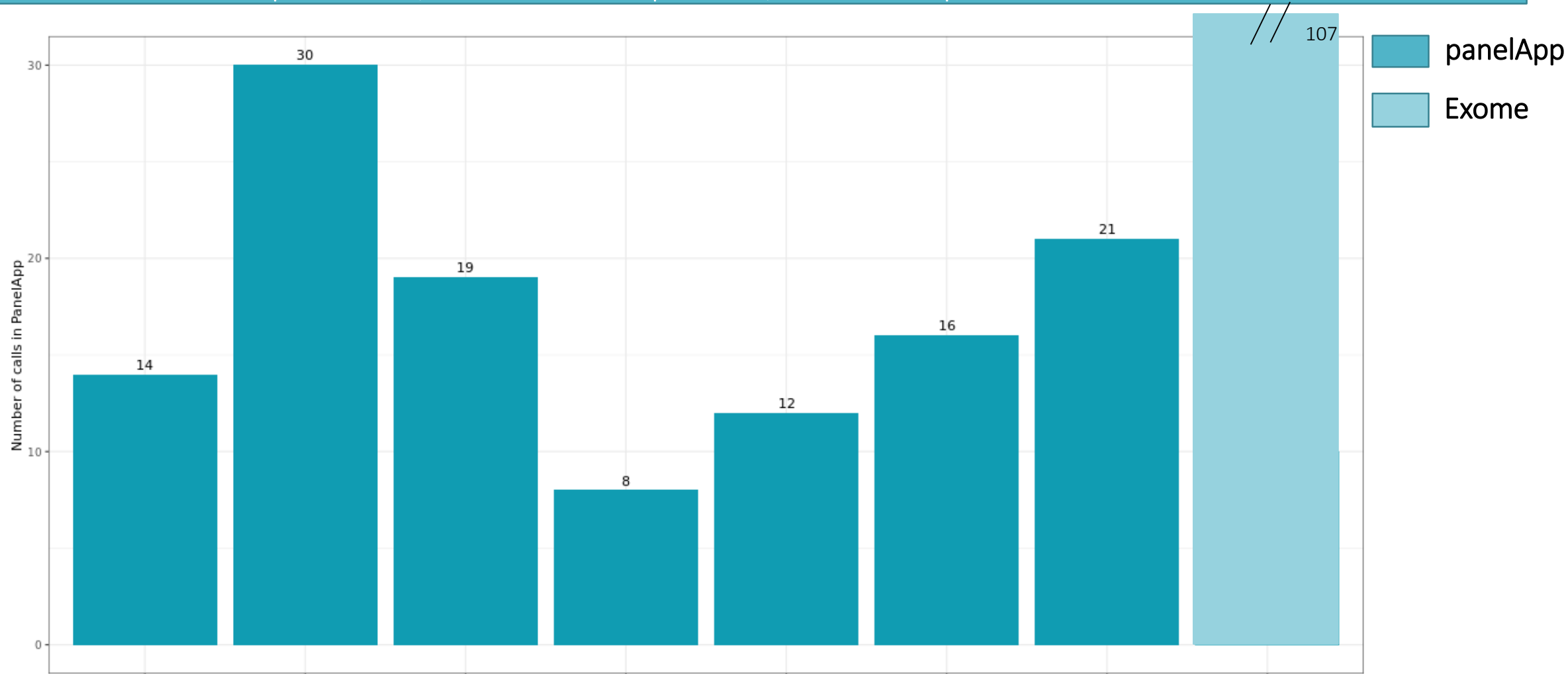
Nombre d'évènements aberrants d'épissage FRASER vs FRASER2

pValue < 0,001 cohorte 73 patients



Priorisation FRASER2 : Nombre d'évènements aberrants d'épissage

pValue < 0,001 cohorte 73 patients, focus sur 8 patients témoins



Variant pathogène :

Rang :

CI
ADNP

1 (126)

CI
DYNC1H1

2 (7)

CI
FOXP1

1 (57)

CI
KDM5C

0 (1)

CI
PSMC2

1 (1)

CI
USP9X

1 (86)

CI
TRIO

10 (0)

CI
INTS6

39 (654)

Conclusions

Stratégie RNA-seq capture exome sur culture lymphocytaire
→ adaptée au laboratoire de diagnostic / inhibition NMD

Détection et priorisation :

- FRASER → effets d'épissage aberrants
- OUTRIDER → quantités de transcrits aberrantes

Conclusions

Stratégie RNA-seq capture exome sur culture lymphocytaire
→ adaptée au laboratoire de diagnostic / inhibition NMD

Détection et priorisation :

- FRASER → effets d'épissage aberrants
- OUTRIDER → quantités de transcrits aberrantes

Analyse de 2^{ème} intention
après un génome négatif

Perspectives

Evaluer la sensibilité des outils :

Impact transcription

Création dataset perte de fonction sur gènes +/-
exprimés

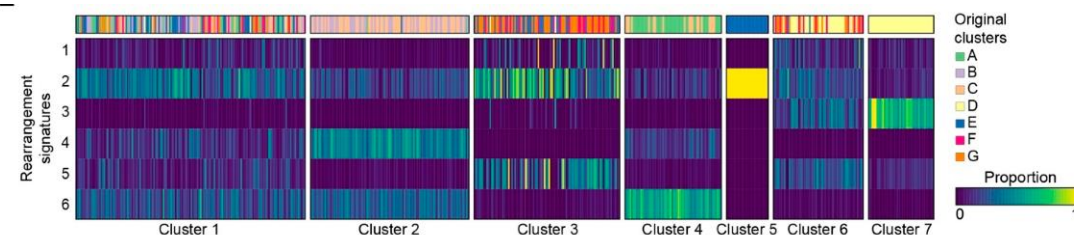
Analyse signatures :

Classifieur

Utilisation SVM / random forest

Datalake Seqoia

Projet national : RNU4-2



Remerciements



Thomas Besnard
Patricia Talarmain
Delphine Quinquis
Jérôme Buscail
Gaëlle Landeau
Eva Trochu
Maël Reynaud
Wallid Deb
Stéphane Béziau
Benjamin Cogné

Bertrand Isidor
Sandra Mercier
Mathilde Nizon
Marie Vincent
Solène Conrad

Virginie Vignard
Frédéric Ebstein
Sébastien Kury

Amandine Charreton
Emmanuelle Lecommandeur
Sylvie Odent
Et tous les cliniciens et
biologistes collaborateurs

Pierre Blanc
Alban Lermine
Pierre Marijon
Et leurs équipes



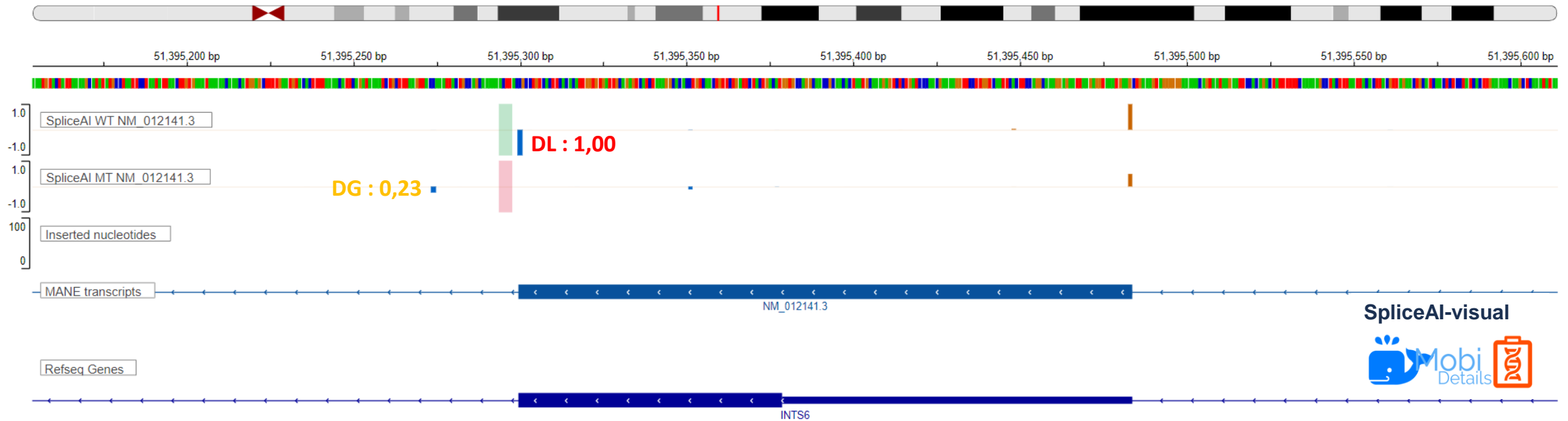
Remerciements



BiInfoDiag

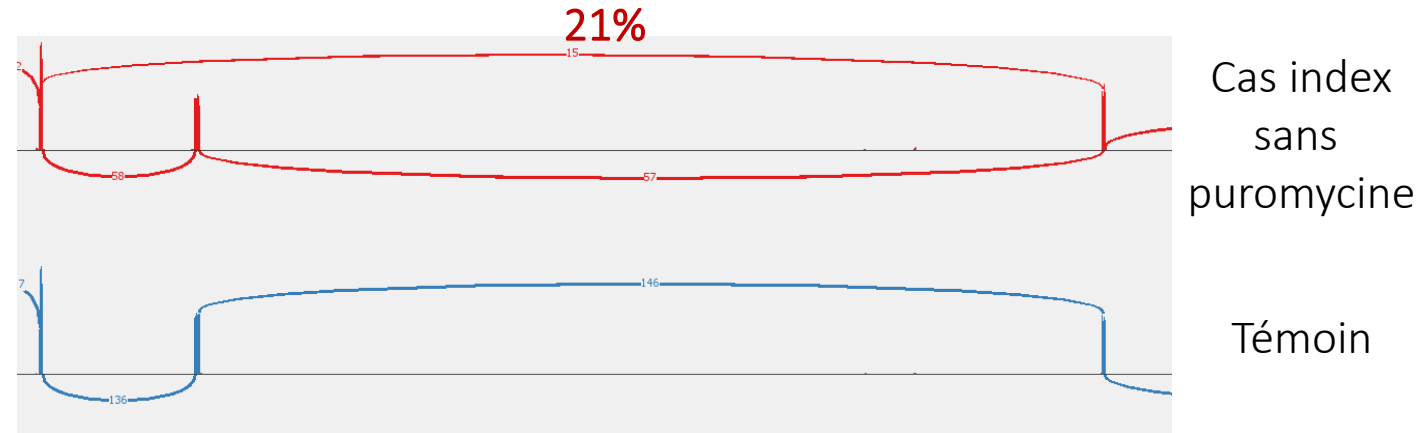
Détecter les aberrations de quantité de transcrits : OUTRIDER (Brechtmann et al, *AJHG*, 2018)

INTS6 - NM_012141.3 - c.613+3_613+6del



Détecter les aberrations de quantité de transcrits : OTRIDER (Brechtmann et al, *AJHG*, 2018)

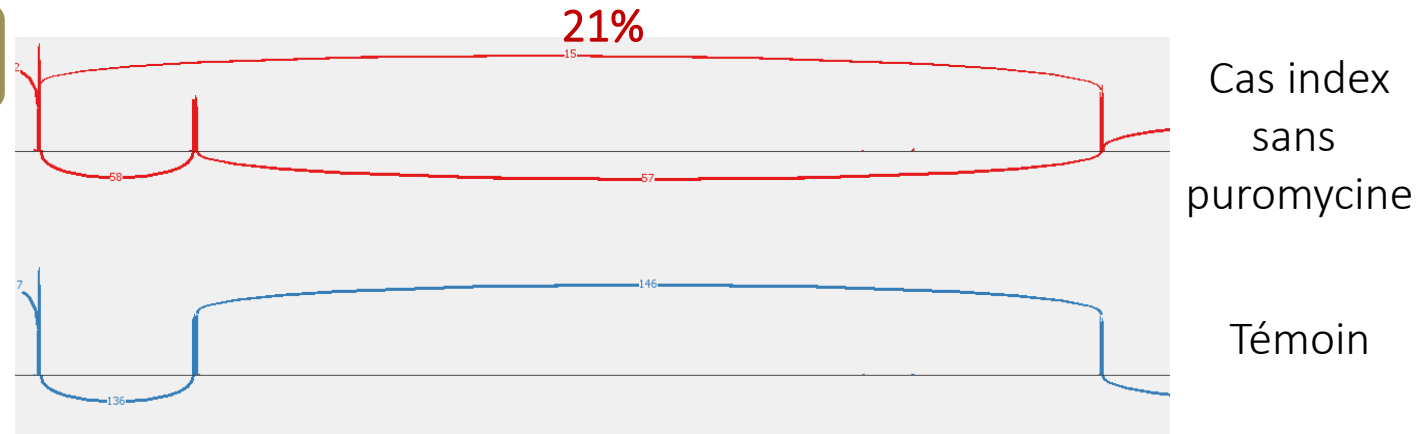
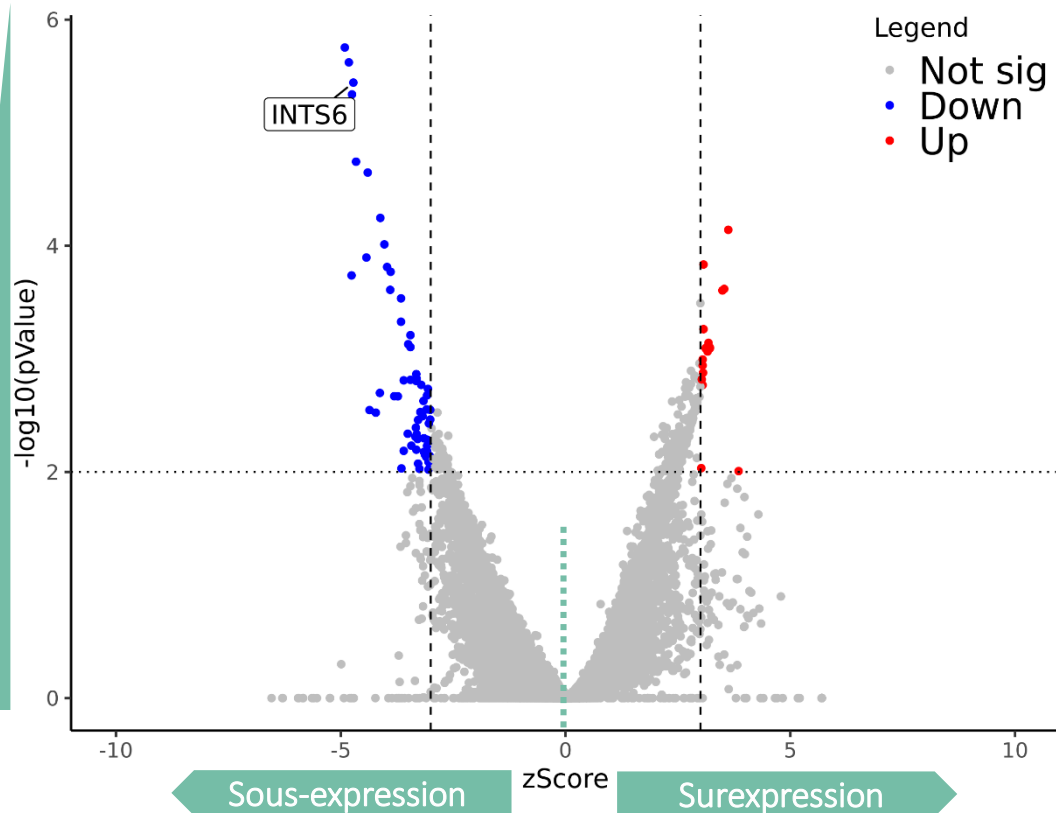
INTS6 - NM_012141.3 - c.613+3_613+6del



Détecter les aberrations de quantité de transcrits : OUTRIDER (Brechtmann et al, *AJHG*, 2018)

INTS6 - NM_012141.3 - c.613+3_613+6del

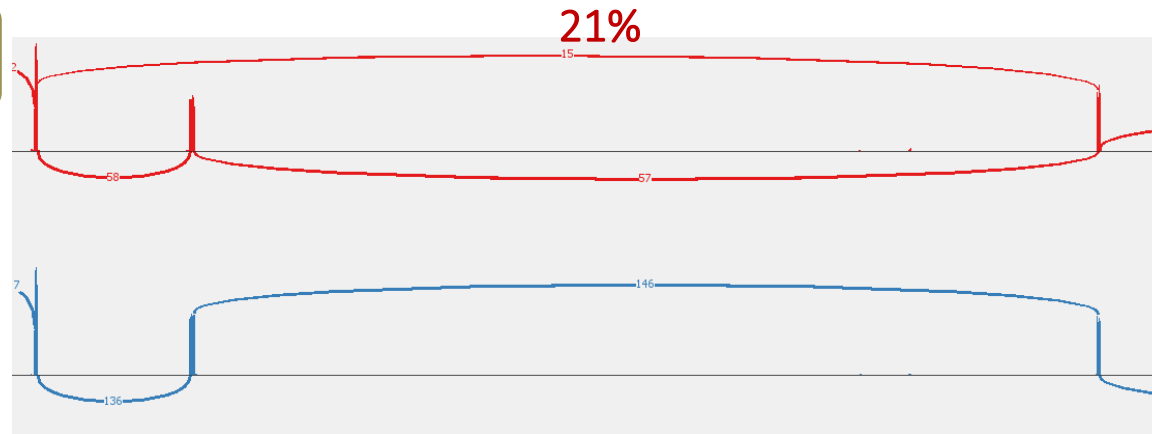
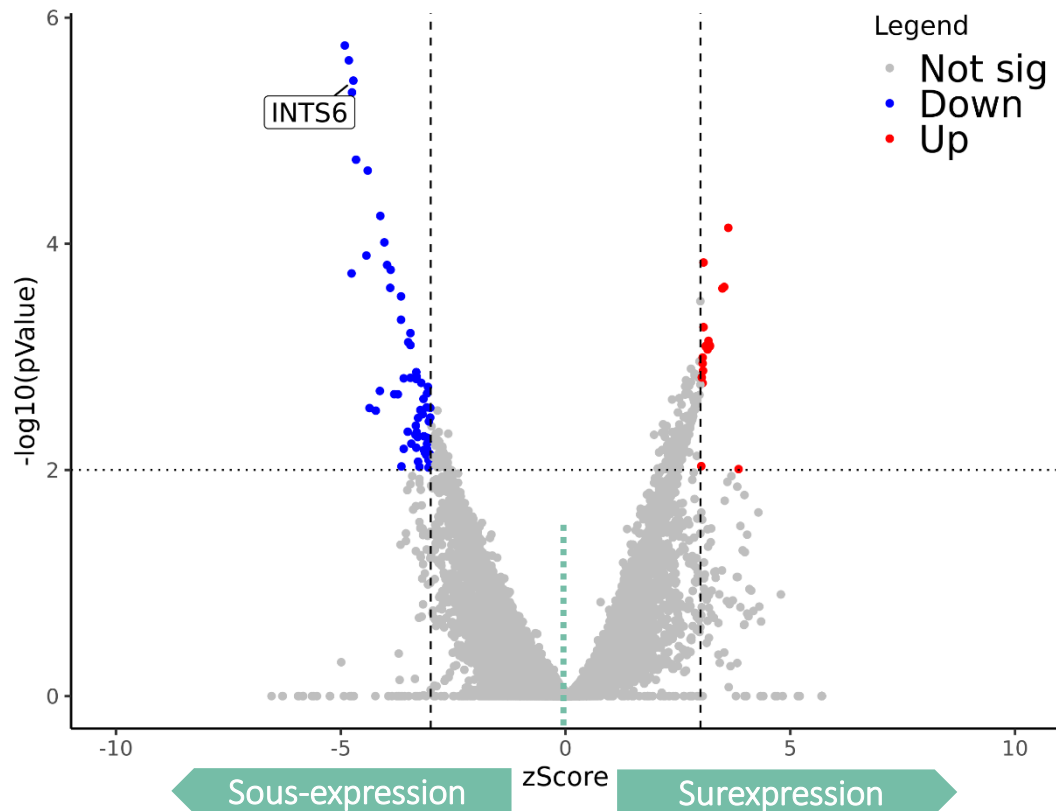
OUTRIDER Volcano plot : CIX *INTS6*



Détecter les aberrations de quantité de transcrits : OUTRIDER (Brechtmann et al, *AJHG*, 2018)

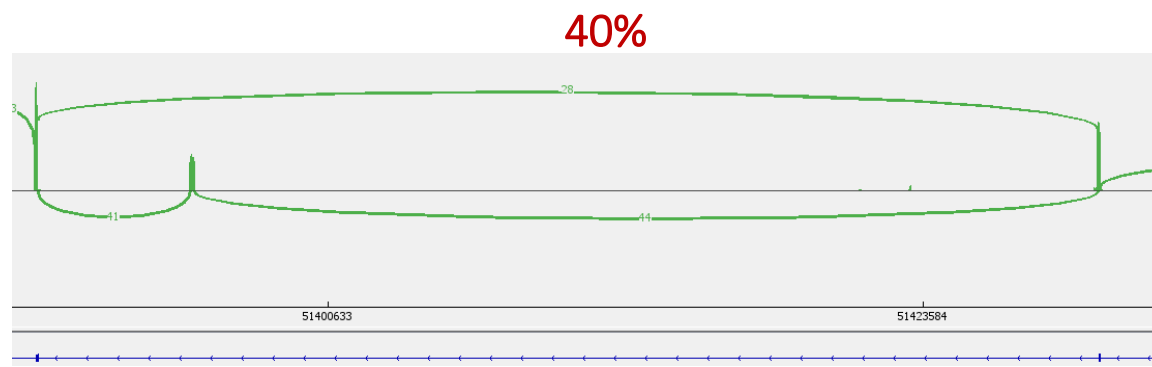
INTS6 - NM_012141.3 - c.613+3_613+6del

OUTRIDER Volcano plot : CIX *INTS6*



Cas index
sans
puromycine

Témoin



Cas index
traité
puromycine