



GT RNA-seq méthodes

David Baux - séminaire BioinfoDiag 14/05/2024



Le GT

- ▶ 33 participants
- ▶ Objectif: faire l'état des lieux et établir des recommandations concernant la pratique du RNA-Seq en génétique médicale
- ▶ 1ère réunion le 12/10/2023
- ▶ 3 sous-groupes:
 - Anomalies de l'épissage (variation de structure)
 - Transcrits de fusion
 - Expression différentielle



Déroulement



- ▶ Pour chaque sous-groupe:
 - Veille bibliographique
 - Retours d'expérience
- ▶ 8 réunions en distanciel avec entre 10 et 25 participants
- ▶ Cloud bioinfodiag: CRs, diaporamas, captations vidéos des réunions...
- ▶ Un GT vivant!!!
- ▶ Productivité?

Spécificité du diagnostic moléculaire

- ▶ Patient VS cohorte
 - Cohorte patient? (épissage ~ok, différentiel d'expression oui/non)
 - Cohorte contrôle
- ▶ Construction de la librairie (gènes codants ++)

technique de sélection des ARNs	Épissage	Fusion	Expression	SNV/CNV calling	# d'échantillons par run
polyA	++	+	+	+	++
ribo déplétion	+/-	+/-	+/-	+/-	+
capture	+	+	+	+	++
3' RNA-Seq	-	-	++	-	++++

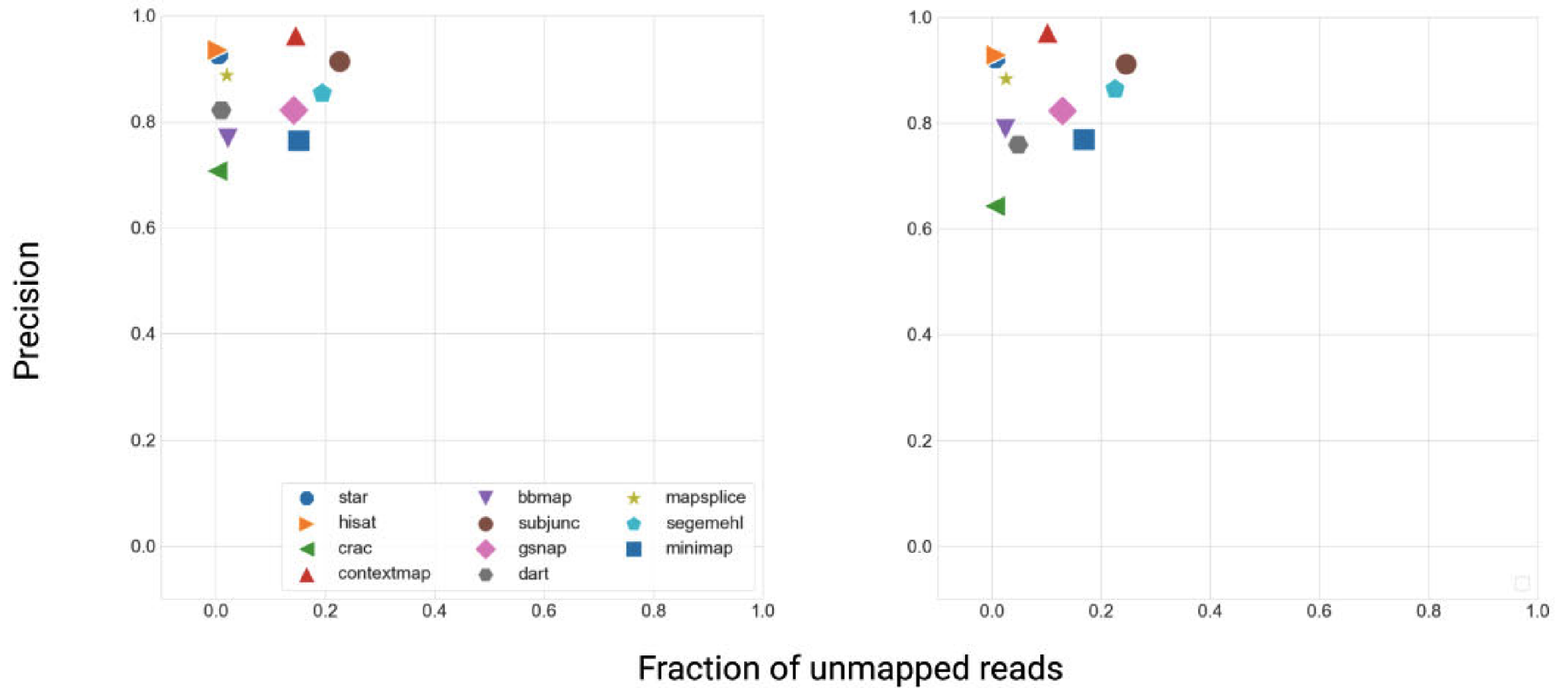
Alignement

A

S0

Fenn et al., 2023

S5

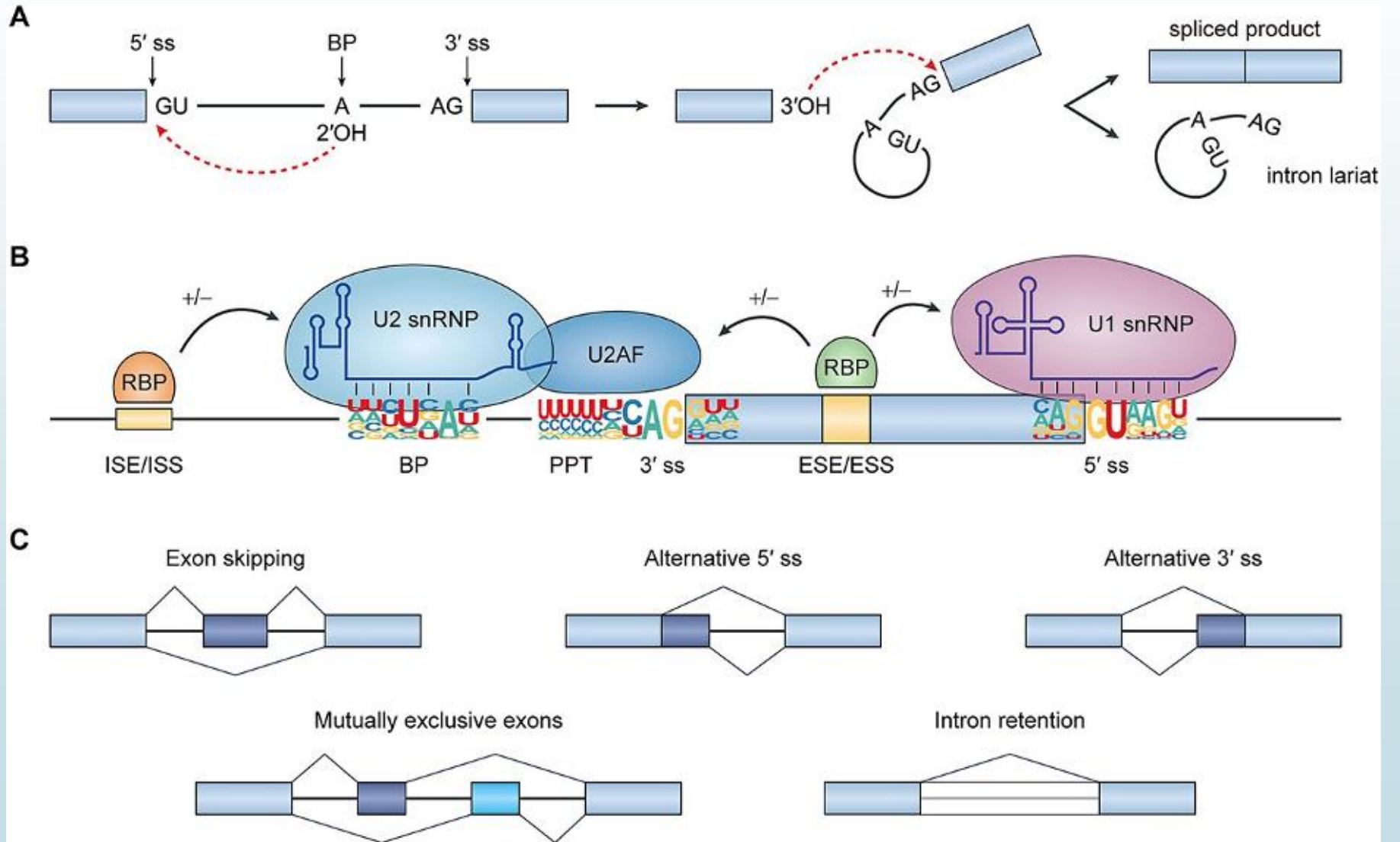


Precision: proportion of correctly mapped reads and junctions relative to the number of all mapped reads

Anomalies de l'épissage

F.Ponelle, C.Habib, J.Chiron, A.Perrin, L.Do Souto, R.Leman

Wang et al., 2023



Anomalies de l'epissage

- Outil intéressant pour le design expérimental: MRSD (Minimum Required Sequencing Depth) - <https://mcgm-mrtd.github.io/>

Gene list [About](#) [Code](#)

Splice junction read coverage

Proportion of splice junctions covered %

Confidence level %

PanelApp panel

Gene list

Anomalies de l'epissage

- Outil intéressant pour le design expérimental: MRSD

Summary

Sequencing reads	Genes covered (of 109)			
	Blood	Fibroblasts	LCL	Muscle
< 50M	8 (7%)	27 (25%)	20 (18%)	18 (17%)
50M to 100M	2 (2%)	1 (1%)	4 (4%)	5 (5%)
> 100M	5 (5%)	6 (6%)	6 (6%)	14 (13%)
Not covered	94 (86%)	75 (69%)	79 (72%)	72 (66%)

LCL: lymphoblastoid cell lines

Rowlands et al., 2022

Gene Symbol	Ensembl ID	Transcript Type	Panels	Minimum Read Sequencing Depth			
				Blood	Fibroblasts	LCL	Muscle
ABHD12	ENSG00000100997	MANE	Ataxia and cerebellar anomalies - narrow panel;Hereditary ataxia;Hereditary ataxia - adult onset;Hereditary ataxia and cerebellar anomalies - childhood onset;Hypotonic infant;Paediatric disorders;Severe Paediatric Disorders;White matter disorders - childhood onset;Hereditary neuropathy;Inborn errors of metabolism;Undiagnosed metabolic disorders;Retinal disorders;Hearing loss	56.8	7.2	6.88	36.16

Anomalies de l'épissage

Outil	Type d'évènement détecté
IRFinder / IRFinder-S	RI (totales)
FRASER / FRASER2	SE, RI, 5A, 3A, SEM
SpliceLauncher	SE, 5A, 3A
LeafCutterMD	SE, 5A, 3A, RI, SEM

types d'évènements détectés par outil.

SE: Saut d'Exon

RI: Rétention Intronique

5A: site 5' d'épissage Alternatif

3A, site 3' d'épissage Alternatif

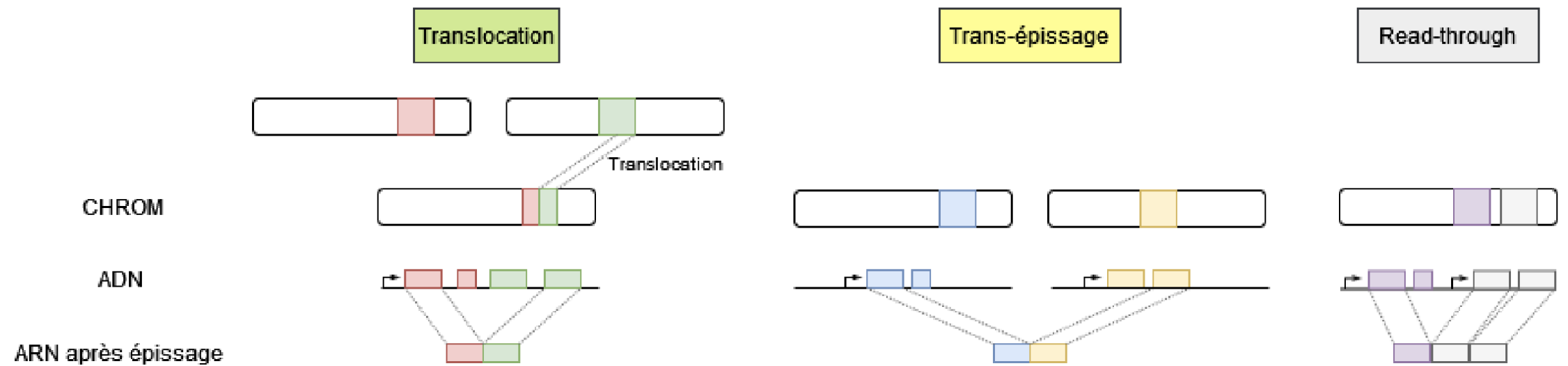
EME: Exon Mutuellement Exclusif

SEM: Saut d'Exon Multiple

Adapté de Fenn et al., 2022

Transcrits de fusion

C.Russo, A.Aliouat, J.Chiron, T.Guignard, F.Escudié, F.Denoual, R.Azmani, S. Cabello



Transcrits de fusion

C.Russo, A.Aliouat, J.Chiron, T.Guignard, F.Escudié, F.Denoual, R.Azmani, S. Cabello

Outils	Type de données d'entrées	F1 score Vicente-Garcès et al., 2023*	F1 score Creason et al., 2021**
Arriba	Alignement issu de STAR	0,89	0.73
Star-Fusion	Fastq ou alignement de STAR	0,81	0.70
FusionCatcher	Fastq	0,88	0.58
Cicero	Alignement issu de STAR	0,85	/
deFuse	Fastq	0,76	/

*Acute pediatric leukemia, 14 cell lines, 15 patients

**77 fusion detection entries and 65 isoform quantification entries on 51 synthetic tumors and 32 cell lines with spiked-in fusion constructs



Expression différentielle

S.Cabello, L. Do Souto, P.Brunelle, A.Mahieux, J.Garreau, R.Azman

- ▶ L'analyse typiquement utilisée en recherche et encore peu en diagnostic
- ▶ Applications possibles:
 - Contrôle qualité (ex: présence d'un effet batch dans un run)
 - Expression aberrante: signature d'un évènement
 - Signatures transcriptomiques / degré de sévérité de pathologies
 - Classification de variants?



Expression différentielle

S.Cabello, L. Do Souto, P.Brunelle, A.Mahieux, J.Garreau, R.Azmani

- ▶ Méthodes éprouvées en recherche mais problématique prégnante du "1 VS all"
 - Utilisation ACP pour détection valeurs aberrantes
 - OUT-RIDER
- ▶ En oncologie (ex: Centres Bérard et Bergonié):
 - classification de tumeurs par clustering (ex: sarcomes à Bordeaux)



Bilan

- ▶ Phase d'écriture des "recommandations" en cours pour les 3 thèmes
- ▶ Discussions pour partage de données (ex: matrices de comptage):
 - par quels biais (contacts directs, nextcloud bioinfodiag, autres?)
- ▶ Rejoignez-nous!!



Merci aux membres du GT!!!

- Adnani Yahia
- Anne-Sophie Denommé-Pichon
- Anne-Sophie Jourdain
- Hakim Bouazzaoui
- Camille Benoist
- Cyril Fournier
- Fabrice Bonte
- Flora Ponelle-Chachuat
- Florent Denoual
- Christophe Habib
- Jennifer Chiron
- Julie Bogoin
- Julien Buratti
- Laurent Castera
- Martin Broly
- Alexandra Lespagnol
- Perrine Brunelle
- Rihab Azmani
- Raphaël Leman
- Thomas Guignard
- Valentin Vautrot
- Wilfrid Carré
- Simon Cabello-Aguilar
- Amyra Aliouat
- Nada Maaziz
- Laura DoSouto-Ferreira
- Charles Van Goethem
- Christophe Russo
- Luc Thomes
- Seydi Thimbo
- Jules Garreau
- Frédéric Escudié