

BioInfoDiag

Réseau Français de Bioinformatique pour le Diagnostic

Applications du RNA-Seq en diagnostic et bioinformatique: retour du GT BioinfoDiag

David Baux

01/04/2025

GTRNAseq: 3 Groupes de travail

▶ « Bio »

Définir les besoins du point de vue du biologiste pour les analyses RNASeq

▶ « Design »

Définir des recommandations liées à la conception des expérimentations de séquençage ARN

▶ « Méthodes »

Pratiques bioinformatiques liées à l'utilisation des techniques de séquençage ARN (RNA-seq) en diagnostic de génétique moléculaire.

Participants

GT Méthodes

David Baux – CHU de Montpellier
Aurélien Perrin – CHU de Montpellier
Florent Denoual – CHU de Rennes
Perrine Brunelle – CHU de Lille
Thomas Guignard – CHU de Montpellier
Wilfrid Carré – CHU de Rennes
Abdelhakim Bouazzaoui – CHU de Rennes
Simon Cabello-Aguilar – CHU de Montpellier
Amyra Aliouat – CHU de Rennes
Christophe Russo – CHU de Lille
Alexandra Lespagnol – CHU de Rennes
Laura Do Souto – CHU de Nantes
David Baux – CHU de Montpellier
Flora Ponelle-Chachuat – CLCC de Clermont-Ferrand
Christophe Habib – CHU de Toulouse
Jennifer Chiron – CLCC Bergonié
Anne-Sophie Dénommé-Pichon – CHU de Dijon
Anne-Sophie Jourdain – CHU de Lille
Camille Benoist – Institut Curie
Fabrice Bonte – CHU de Lille
Julie Bogoin – APHP
Elise Guéret – CHU de Montpellier
Julien Buratti – APHP
Claire Guissard – CHU de Nîmes
Jérôme Reboul – Inserm
Martin Broly – CHU de Montpellier
Charles Van Goethem – CHU de Montpellier
Luc Thomes – CHU de Lille
Eléonore Frouin – Institut Curie

Aurélien Bourdon – CLCC Bergonié
Seydi Thimbo – CHU de Lille
Laetitia Gaston – CHU de Bordeaux
Kahia Messaoudi – CHU d'Amiens
Laurence Lodé – CHU de Reims
Thérèse Commes – Inserm
Svetlana Gorokhova – Université d'Aix-Marseille
Corentin Marco – CHU de Nîmes
Rihab Azmani – CLCC Bergonié
Sylvain Mareschal – CHU de Lyon
Raphaël Leman – CLCC Baclesse
Valentin Vautrot – Université de Bourgogne
Christophe Russo – CHU de Lille
Jules Garreau – Université de Rennes
Ariane Mahieux – CHU de Brest
Justine Labory – INRAE
Frédéric Escudié – CHU de Toulouse
Franck Tirode – Inserm
L'équipe Bio2M notamment,
Thérèse Commes, Anthony Boureux, Benoit Guibert –
Inserm
Sylvie Tuffery – Inserm (partie épissage)
Benjamin Cogné – CHU de Nantes
Jean-Philippe Villemin – Inserm
Jean Muller – CHU de Strasbourg
Marie de Tayrac – CHU de Rennes

Mathieu Chopelet pour l'assistance technique d'une efficacité sans faille.

GT Design

Anne-Sophie Dénommé-Pichon
Anne-Sophie Jourdain
Aurélien Perrin
Christophe Habib
Bertrand Chesneau
Jennifer Chiron
Perrine Brunelle
Marie de Tayrac

GT Bio

Anne-Sophie Dénommé-Pichon
Claude Houdayer
François Lecoquierre
Jean Muller
Charles van Goethem
Céline Derambure
Myriam Vezain
Alexandra Lespagnol
Christophe Habib
John Rendu
Pierre-Antoine Rollat-Farnier
Svetlana Gorokhova
Wilfrid Carré
Anne-Sophie Jourdain
Marion Larrieux
Christel Vaché
Perrine Brunelle
Raphael Leman

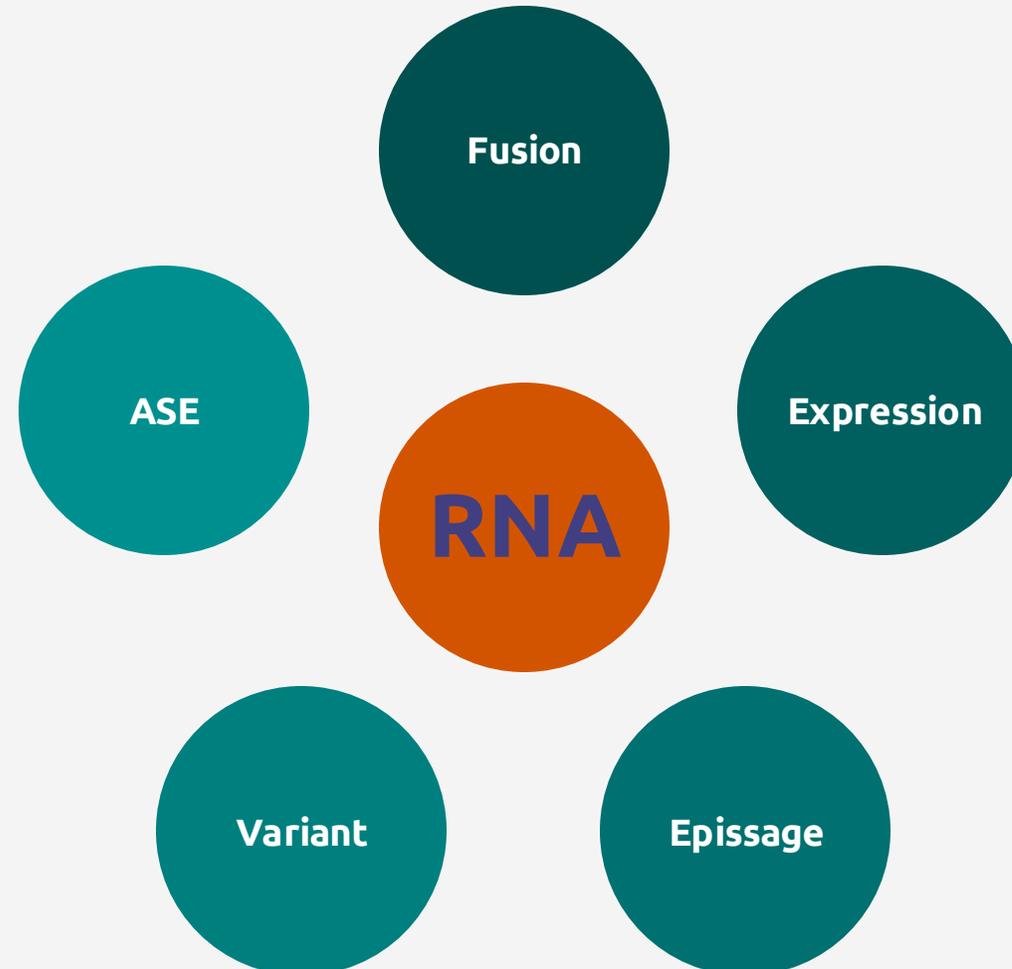
Les besoins du point de vue du biologiste pour les analyses RNASeq

Analytique

- Recherche de Fusions
- Expression génique
- Anomalies d'épissage
- Détection des SNV/INDEL
- Expression mono-allélique

Interprétation

- Visualisation
- Métriques Qualité run/patients

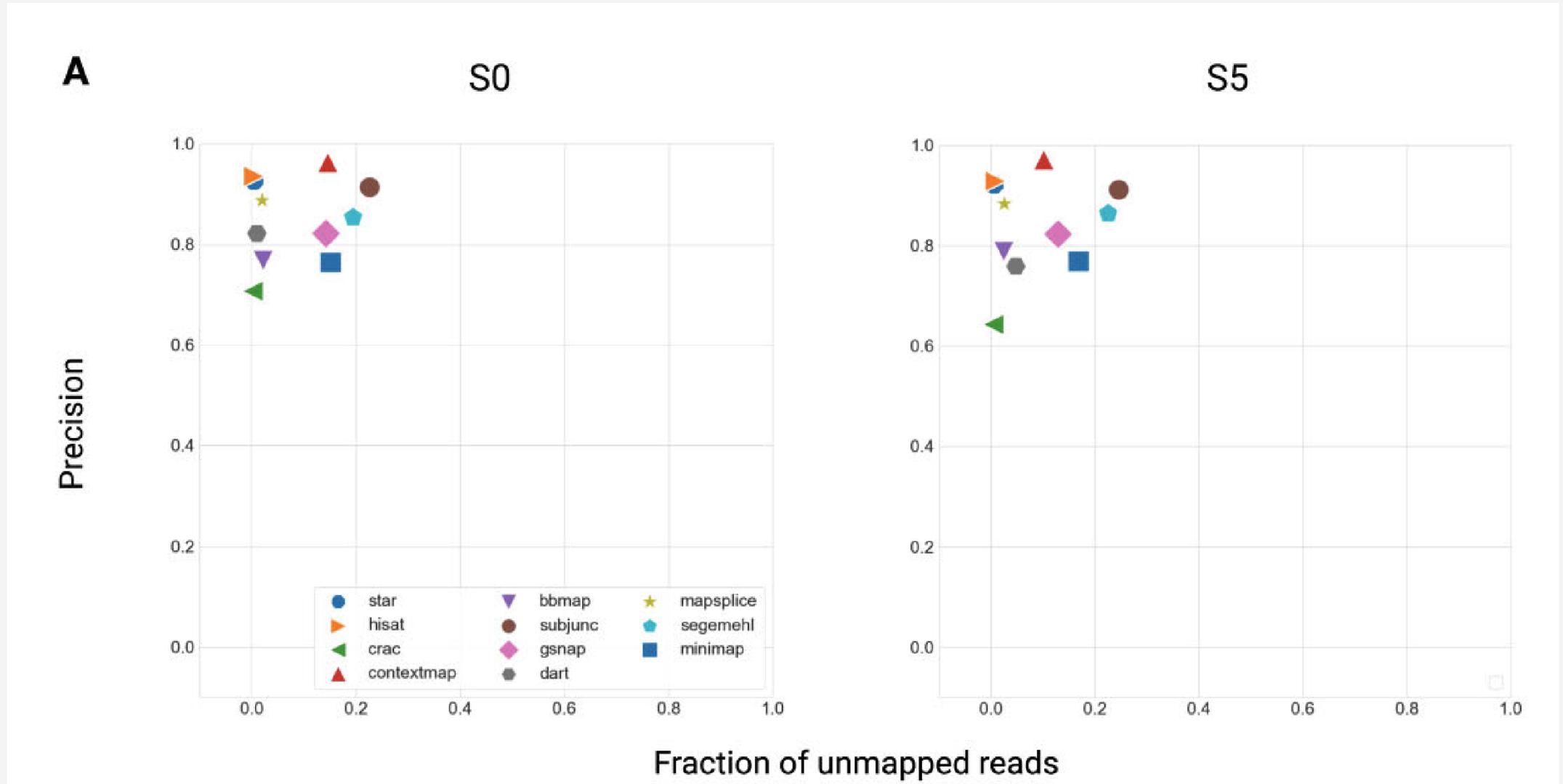


Spécificité du diagnostic moléculaire

- Patient VS cohorte
 - Cohorte patient? (epissage ~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

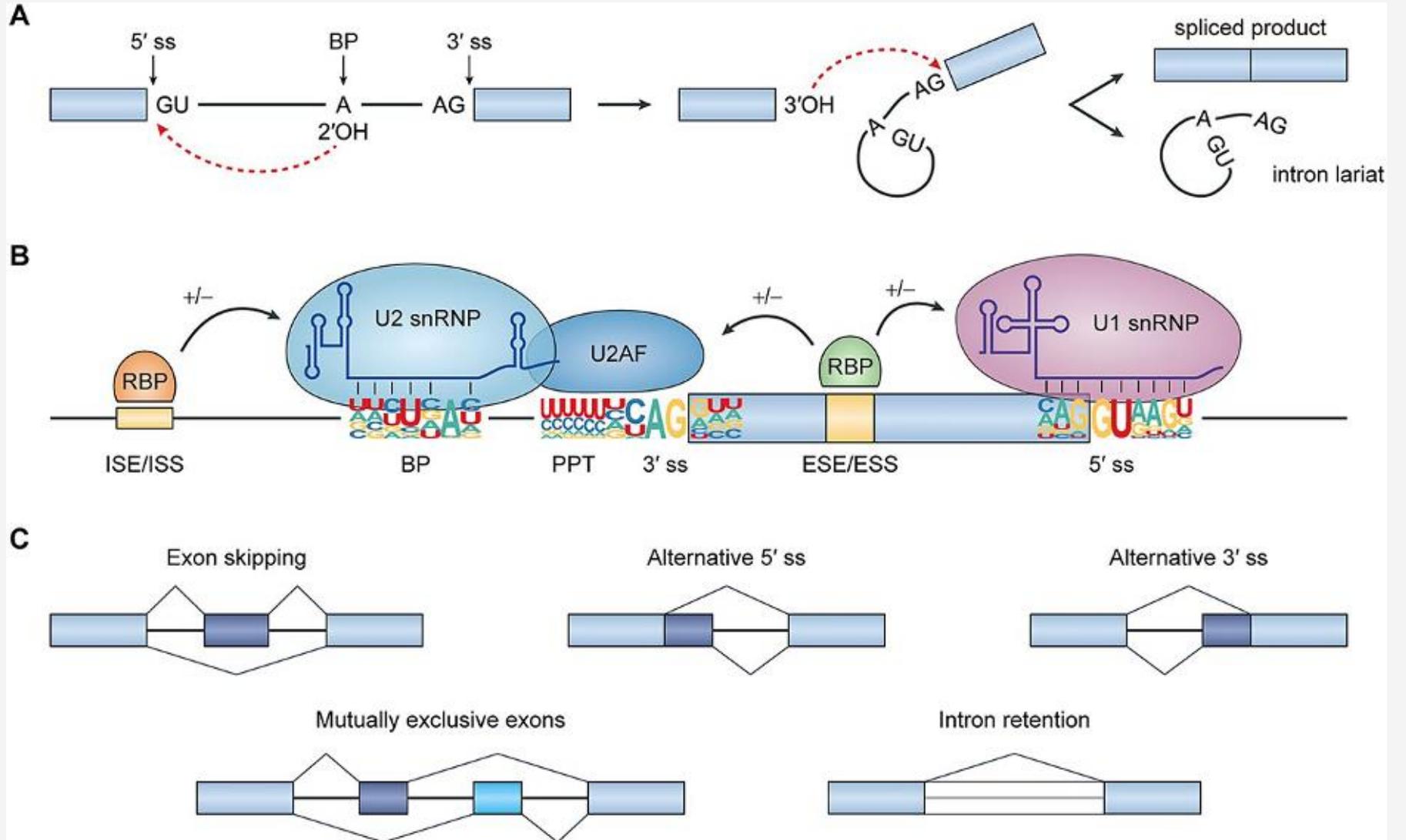


Fenn et al., 2023.

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

Anomalies de l'épissage

Wang et al., 2023



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

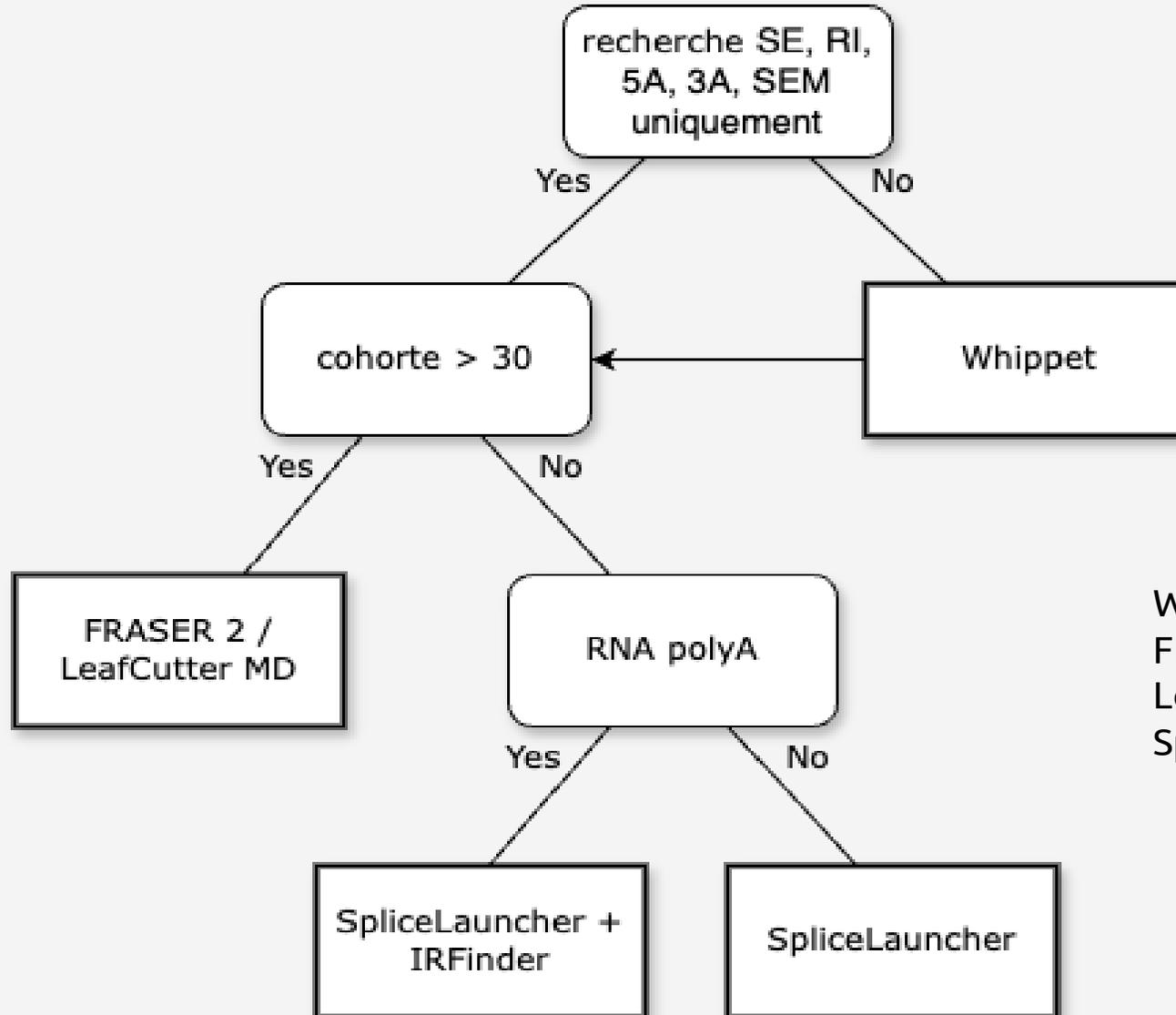
5A: site 5' d'épissage Alternatif

3A, site 3' d'épissage Alternatif

SEM: Saut d'Exon Multiple

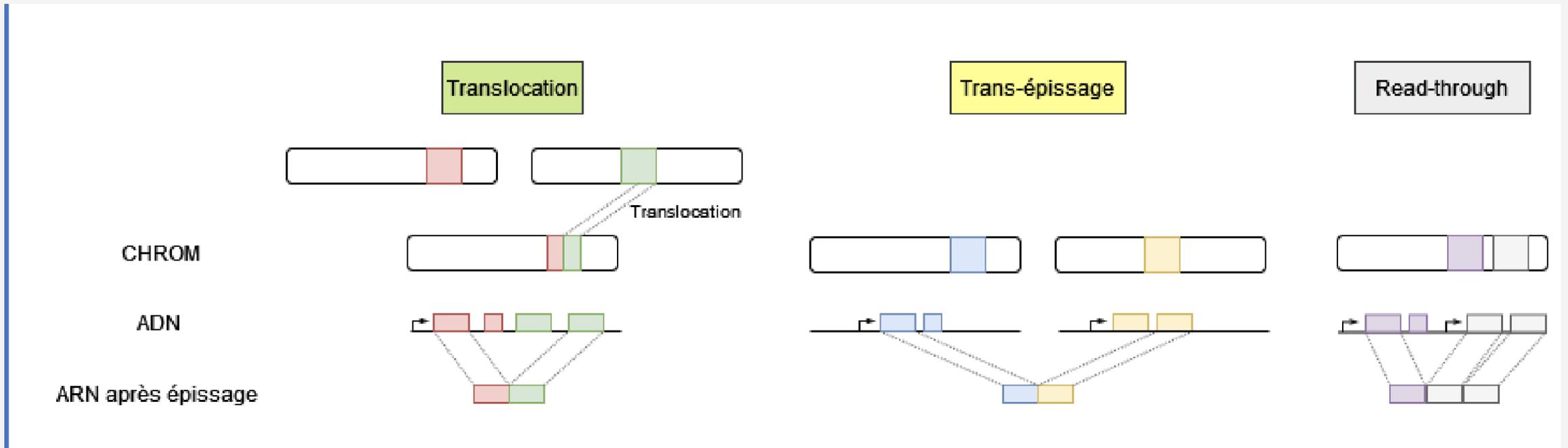
Adapté de Fenn et al., 2023

Anomalies de l'épissage: Ex d'arbre décisionnel



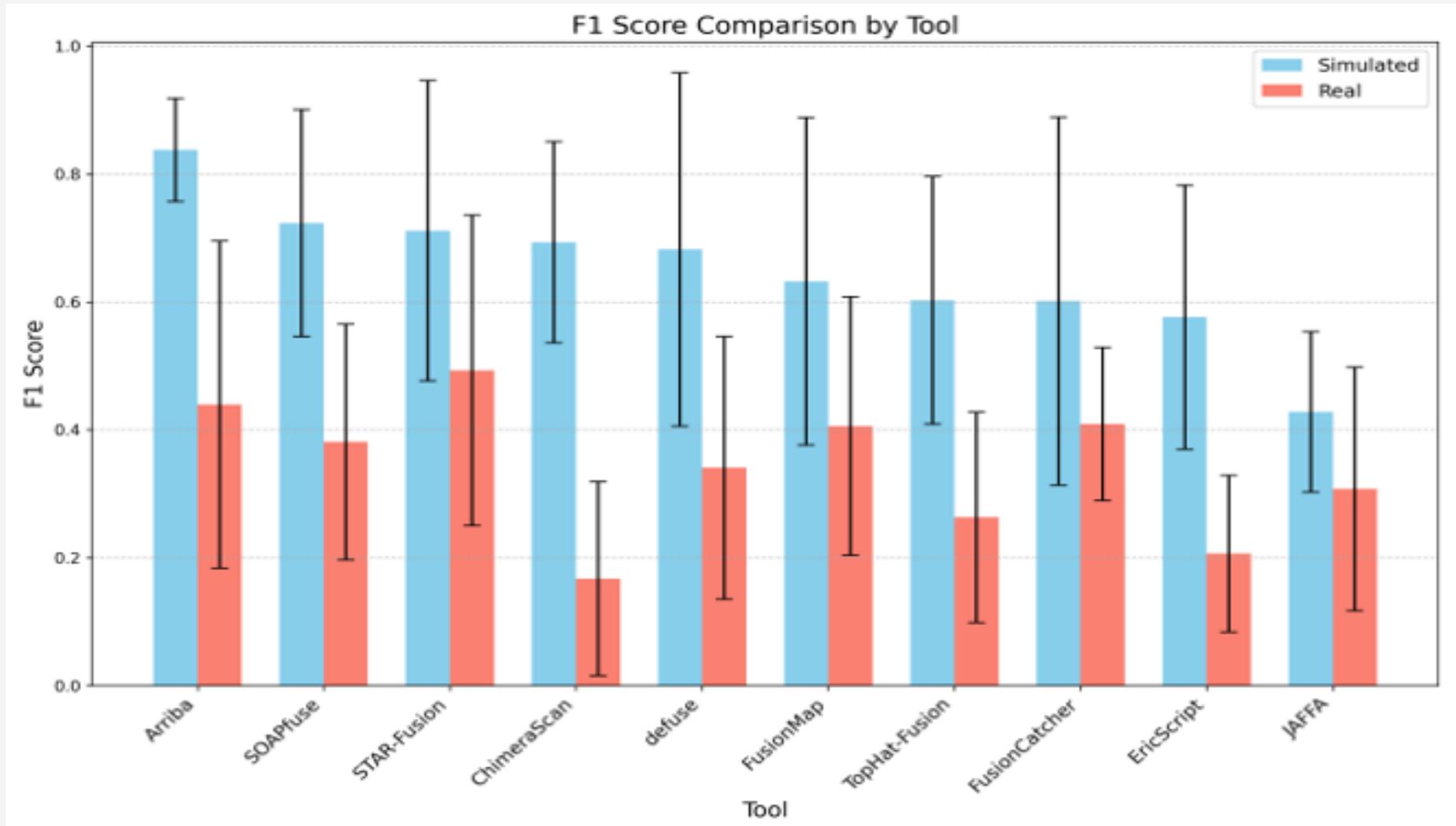
Whippet: Sterne-Weiler et al., 2018
FRASER2: Scheller et al., 2023
LeafCutterMD: jenkinson et al. 2020
SpliceLauncher: Leman et al., 2019

Transcrits de fusion



D'après Escudié F, communication personnelle

Transcrits de fusion



Expression différentielle

- 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

- Méthodes éprouvées en recherche mais problématique prégnante du "1 VS all"
 - Utilisation ACP pour détection valeurs aberrantes
 - OUT-RIDER (Brechtmann et al., 2018)
 - ABEILLE (Labory et al., 2022)
- En oncologie (ex: Centres Bérard et Bergonié):
 - classification de tumeurs par clustering (ex: sarcomes)

ARN en lectures longues

	GMAP	minimap2	SpAln	STAR (long)
Temps de traitement (CPU min)	631	15.9	2076	33.9
RAM max (GByte)	8.9	14.5	3.2	29.2
# introns alignés	692 275	693 553	692 945	78 603
# nouveaux introns	11 239	3113	8550	1214
% introns exacts	83.8	94.0	87.9	55.2

Bilan et perspectives

- Le document arrive!!!!
- Continuation du GT

1.3 Les différentes techniques de RNA-Seq et leurs applications

La grande majorité des ARNs présents dans un isolat cellulaire ou issu de tissu sont des ARNs ribosomiaux. Ces ARNs ne sont pas pertinents pour l'analyse de transcrits et doivent être éliminés. Pour cela, on peut soit effectuer une **sélection positive** (par sélection des ARN polyadénylés en fixant la queue polyA ou par capture des séquences exoniques) ou **négative** (élimination des ARNr par hybridation et séparation, ribo-déplétion).

Les méthodes positives sont recommandées pour l'analyse des gènes codants et des transcrits résultants alors que la ribo-déplétion permettra une meilleure étude des miARNs ou des ARNs non codants en général (incluant les lncARNs). La sélection des ARNs poly-adenylés peut être biaisée et montrer une couverture relativement plus importante pour les régions proches de la queue polyA (région terminale du transcrit) par rapport à une technique de capture qui sera plus homogène (Zhao et al., 2018). La capture polyA présente une meilleure élimination de l'ADN génomique résiduel par rapport à un traitement par DNase (Ura et al., 2022). Attention cependant à l'utilisation de sondes de capture, qui pourront avoir un périmètre de détection limité. Par exemple, si on utilise des sondes de capture exoniques uniquement, les rétentions introniques totales ne seront pas visibles directement mais identifiées par 2 rétentions partielles en 5' et 3' de l'intron. De plus, certains outils comme IRFinder seront inopérants ou avec des performances fortement réduites.

L'utilisation de technologies de capture offre une flexibilité importante au niveau des cibles :

- RNA-Seq ciblé sur un panel de gènes d'intérêts,
- analyse du transcriptome codant dans son ensemble (en utilisant un lot de sondes destinées à l'exome par exemple).

Un autre argument en faveur de cette technologie en diagnostic est la possibilité d'utiliser les mêmes sondes que pour le séquençage ADN, afin de réaliser le séquençage ARN, possibilité offerte par la plupart des fournisseurs actuels.

Enfin, valable pour une analyse d'expression génique uniquement, le 3' RNA-Seq se concentre sur l'analyse de l'extrémité 3' des transcrits. Cela présente plusieurs avantages, notamment celui de nécessiter beaucoup moins de quantité de séquences qu'une analyse classique pour quantifier les transcrits (environ 50 000 reads VS 30M), mais aussi de mieux caractériser les transcrits courts.

Participants

GT Méthodes

David Baux – CHU de Montpellier
Aurélien Perrin – CHU de Montpellier
Florent Denoual – CHU de Rennes
Perrine Brunelle – CHU de Lille
Thomas Guignard – CHU de Montpellier
Wilfrid Carré – CHU de Rennes
Abdelhakim Bouazzaoui – CHU de Rennes
Simon Cabello-Aguilar – CHU de Montpellier
Amyra Aliouat – CHU de Rennes
Christophe Russo – CHU de Lille
Alexandra Lespagnol – CHU de Rennes
Laura Do Souto – CHU de Nantes
David Baux – CHU de Montpellier
Flora Ponelle-Chachuat – CLCC de Clermont-Ferrand
Christophe Habib – CHU de Toulouse
Jennifer Chiron – CLCC Bergonié
Anne-Sophie Dénommé-Pichon – CHU de Dijon
Anne-Sophie Jourdain – CHU de Lille
Camille Benoist – Institut Curie
Fabrice Bonte – CHU de Lille
Julie Bogoin – APHP
Elise Guéret – CHU de Montpellier
Julien Buratti – APHP
Claire Guissard – CHU de Nîmes
Jérôme Reboul – Inserm
Martin Broly – CHU de Montpellier
Charles Van Goethem – CHU de Montpellier
Luc Thomes – CHU de Lille
Eléonore Frouin – Institut Curie

Aurélien Bourdon – CLCC Bergonié
Seydi Thimbo – CHU de Lille
Laetitia Gaston – CHU de Bordeaux
Kahia Messaoudi – CHU d'Amiens
Laurence Lodé – CHU de Reims
Thérèse Commes – Inserm
Svetlana Gorokhova – Université d'Aix-Marseille
Corentin Marco – CHU de Nîmes
Rihab Azmani – CLCC Bergonié
Sylvain Mareschal – CHU de Lyon
Raphaël Leman – CLCC Baclesse
Valentin Vautrot – Université de Bourgogne
Christophe Russo – CHU de Lille
Jules Garreau – Université de Rennes
Ariane Mahieux – CHU de Brest
Justine Labory – INRAE
Frédéric Escudié – CHU de Toulouse
Franck Tirode – Inserm
L'équipe Bio2M notamment,
Thérèse Commes, Anthony Boureux, Benoit Guibert –
Inserm
Sylvie Tuffery – Inserm (partie épissage)
Benjamin Cogné – CHU de Nantes
Jean-Philippe Villemin – Inserm
Jean Muller – CHU de Strasbourg
Marie de Tayrac – CHU de Rennes

Mathieu Chopelet pour l'assistance technique d'une efficacité sans faille.

GT Design

Anne-Sophie Dénommé-Pichon
Anne-Sophie Jourdain
Aurélien Perrin
Christophe Habib
Bertrand Chesneau
Jennifer Chiron
Perrine Brunelle
Marie de Tayrac

GT Bio

Anne-Sophie Dénommé-Pichon
Claude Houdayer
François Lecoquierre
Jean Muller
Charles van Goethem
Céline Derambure
Myriam Vezain
Alexandra Lespagnol
Christophe Habib
John Rendu
Pierre-Antoine Rollat-Farnier
Svetlana Gorokhova
Wilfrid Carré
Anne-Sophie Jourdain
Marion Larrieux
Christel Vaché
Perrine Brunelle
Raphael Leman