

➤ Séquençage rationalisé 3' mRNA-seq

Plateforme de transcriptomique GeT-TRiX

Gaëlle Payros et Yannick Lippi

Biopuces

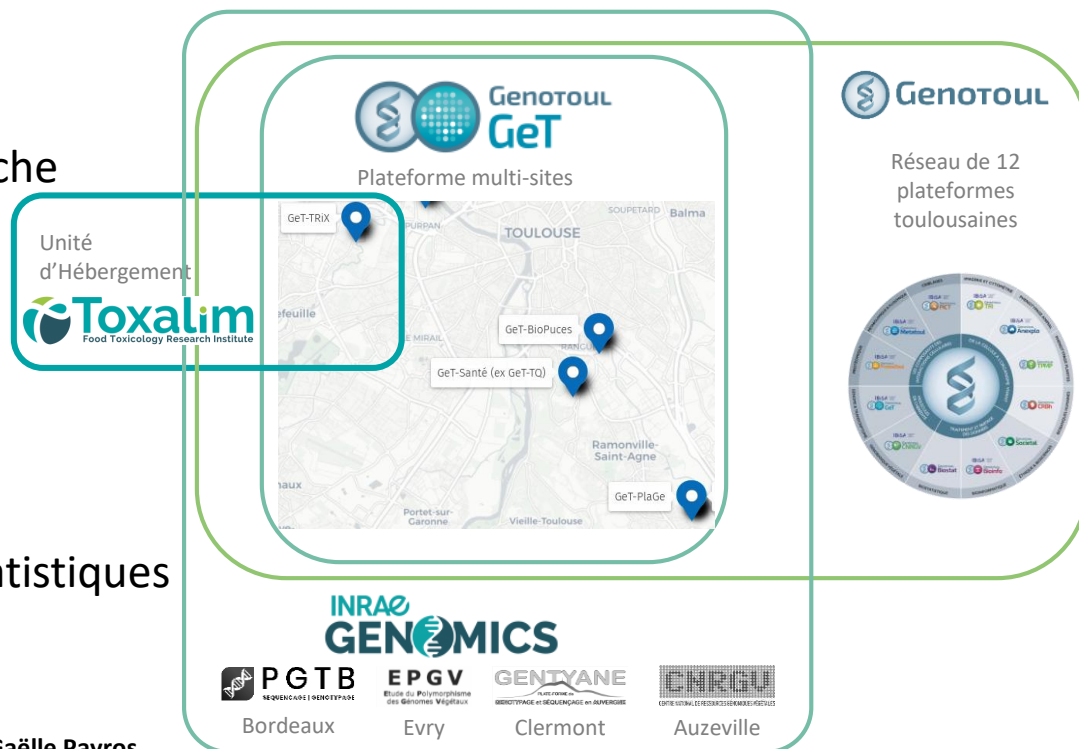
9 Octobre 2025

➤ La plateforme GeT-TRiX

Services intégrés d'analyses transcriptomiques

- Une plateforme dédiée aux études transcriptomiques
 - Mise à disposition d'équipements
 - Prestations « clés en main »
 - Collaboration R&D
 - Partenariat sur contrat de recherche
- Une expertise technologique
 - Analyses transcriptomiques
 - **RNA-seq**
 - **Microarray**
 - Analyses Bioinformatiques/Biostatistiques

- Environnement



Yannick Lippi

IE Bioinfo/Biostat
Pilotage



Claire Naylies

TR Biologie moléculaire
Utilisations



Gaëlle Payros

IE Biologie moléculaire
R&D NGS

➤ QuantSeq 3' mRNA-seq

- Contexte
- Principe
- Différence avec RNA-seq « classique » ?

> Contexte



Proposer un **nouveau service** d'étude du transcriptome, **adapté** pour l'expression génique différentielle sans *a priori* par séquençage des **ARN messagers** :

- Séquençage rationalisé des queues 3' des transcrits et non des transcrits pleine longueur



- Service **abordable** en termes de **coûts/masse de données** (séquençage ↘)



- Applicable sur un **grand nombre** d'individus : **multiplexage**



- **Performant sur sang total** via l'utilisation de « bloqueurs de globine » (disponible pour humain et cochon)



- Applicable à partir de matériel **faible qualité**

> Contexte



Proposer un **nouveau service** d'étude du transcriptome, **adapté** pour **l'expression génique différentielle** sans *a priori* par séquençage des **ARN messagers** :

- Séquençage rationalisé des queues 3' des transcrits et non des transcrits pleine longueur
- **Service abordable** en termes de **coûts/masse de données** (séquençage \searrow)
- Applicable sur un **grand nombre** d'individus : **multiplexage**
- **Performant sur sang total** via l'utilisation de « bloqueurs de globine » (disponible pour humain et cochon)
- Applicable à partir de matériel **faible qualité**



Financement Carnot F2E en collaboration avec GenPhySE. **Projet ATraGESeqR : Analyse Transcriptomique à Grande Echelle par Séquençage Rationalisé**



Principe

QuantSeq 3' mRNA-Seq V2 Library Prep Kit with UDI_Lexogen

LIBRARY GENERATION



3.5 hrs



1 hr

Reverse Transcription
(oligo(dT) priming)



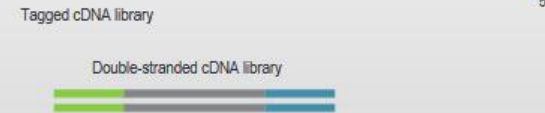
RNA Removal



Second Strand
Synthesis



Purification



1) Reverse Transcription : Synthèse du premier brin cDNA : oligo (dT) sur les queues Poly(A) des ARN.

2) Elimination de l'ARN

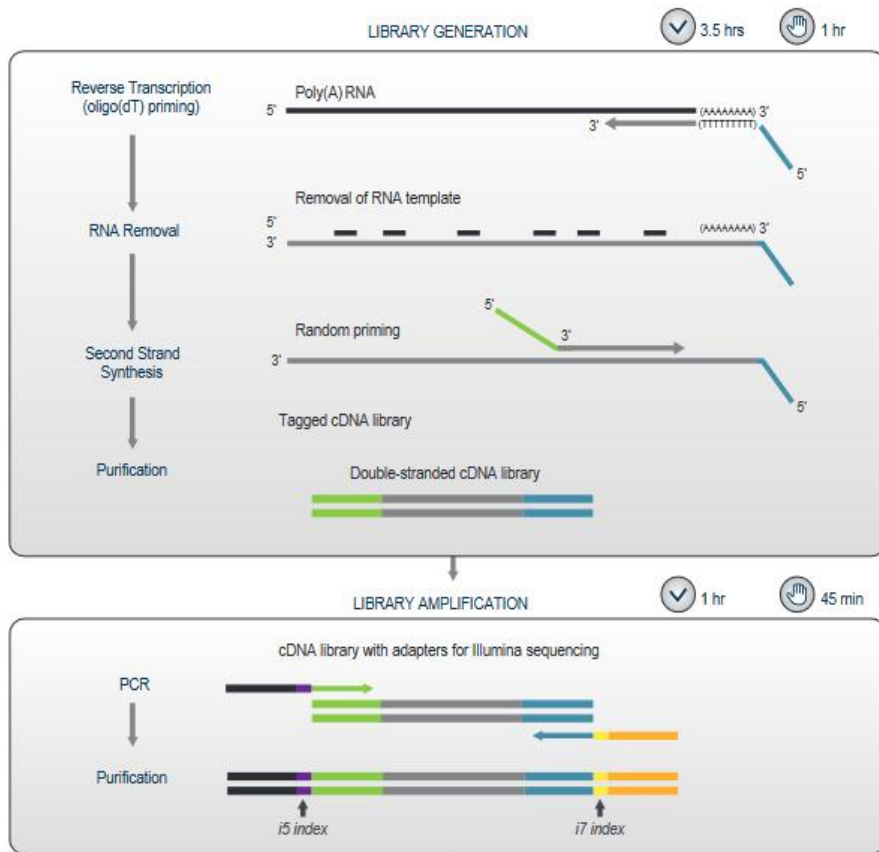
3) Synthèse du deuxième brin : Random primer : un fragment par transcrit

4) Purification sur billes



Principe

QuantSeq 3' mRNA-Seq V2 Library Prep Kit with UDI_Lexogen



1) Reverse Transcription : Synthèse du premier brin cDNA : oligo (dT) sur les queues Poly(A) des ARN.

2) Elimination de l'ARN

3) Synthèse du deuxième brin : Random primer : un fragment par transcrit

4) Purification sur billes

5) Amplification par PCR (+ ajout des index)

6) Purification finale sur billes

7) Contrôle qualité des librairies + Pool



Principe

QuantSeq 3' mRNA-Seq V2 Library Prep Kit with UDI_Lexogen

LIBRARY GENERATION



Reverse Transcription
(oligo(dT) priming)



RNA Removal



Second Strand
Synthesis



Tagged cDNA library

Double-stranded cDNA library

LIBRARY AMPLIFICATION



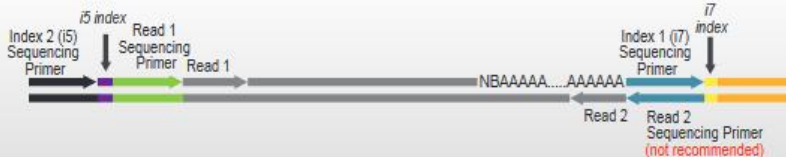
PCR



Purification



SEQUENCING - Read orientation for QuantSeq FWD



1) Reverse Transcription : Synthèse du premier brin cDNA : oligo (dT) sur les queues Poly(A) des ARN.

2) Elimination de l'ARN

3) Synthèse du deuxième brin : Random primer : un fragment par transcrit

4) Purification sur billes

5) Amplification par PCR (+ ajout des index)

6) Purification finale sur billes

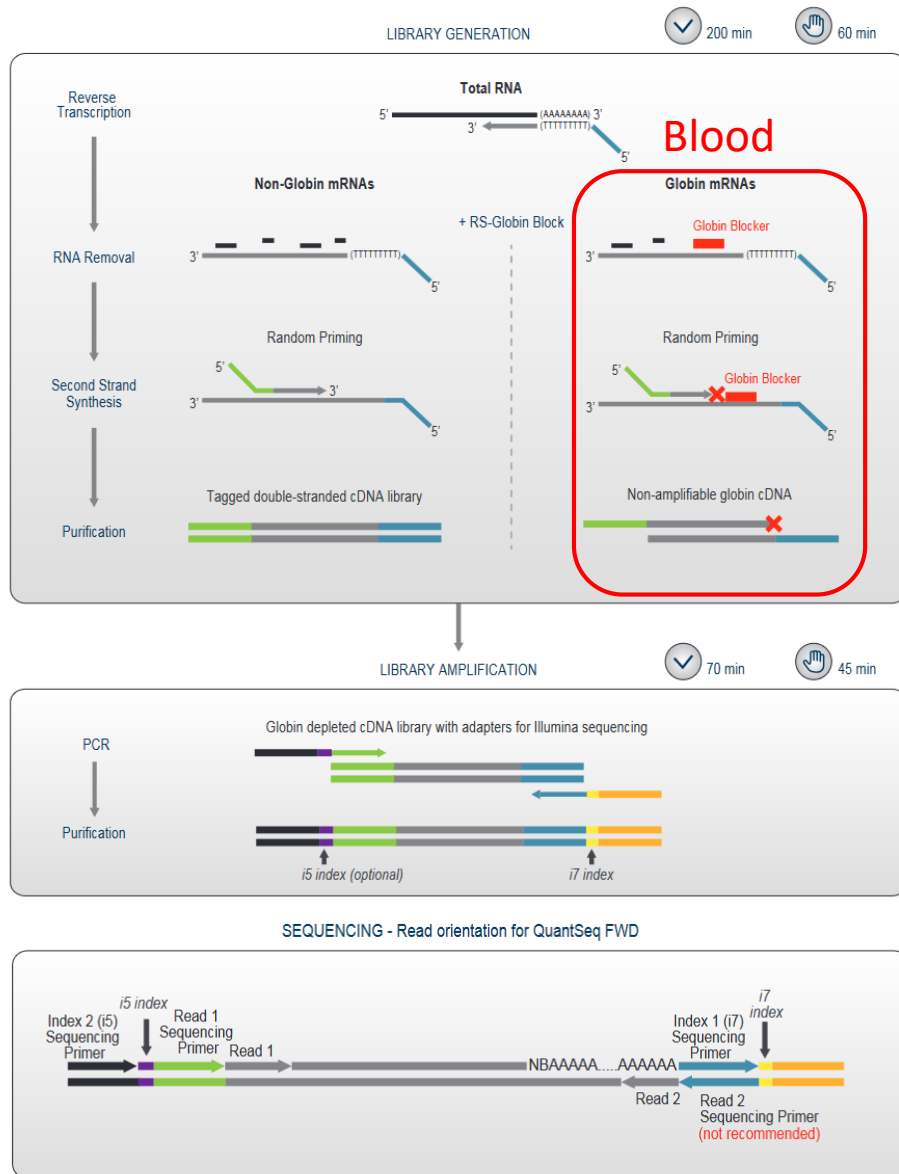
7) Contrôle qualité des librairies + Pool

8) Séquençage : Taille optimisée pour des longueurs de lectures courtes : SR50 – 100 pb.
3-5 M /lectures par échantillon



Principe

QuantSeq 3' mRNA-Seq V2 Library Prep Kit with UDI_Lexogen



1) Reverse Transcription : Synthèse du premier brin cDNA : oligo (dT) sur les queues Poly(A) des ARN.

2) Elimination de l'ARN + **RS-Globin Block** : empêche la génération de fragments à partir d'ARNm de globine, en se liant à l'ADNc et en empêchant la formation d'ADNc double brin

3) Synthèse du deuxième brin : Random primer : un fragment par transcrit

4) Purification sur billes

5) Amplification par PCR (+ ajout des index)

6) Purification finale sur billes

7) Contrôle qualité des librairies + Pool

8) Séquençage : Taille optimisée pour des longueurs de lectures courtes : SR50 – 100 pb. 3-5 M /lectures par échantillon

➤ Différence avec RNA-seq « classique »

3' mRNA-seq

RNA-seq « classique »

Pas de fragmentation des ARN messagers

Fragmentation des ARN messagers

mRNA-Seq
⌚ ~ 7 hrs

QuantSeq
⌚ ~ 4.5 hrs

Pre-processing

mRNA Purification → RNA Fragmentation

~ 1.5 hrs

Library Generation

First Strand Synthesis

Second Strand Synthesis

Clean-up

End Repair

Adapter Ligation

Clean-up

2.5 - 3 hrs

+

1 - 1.5 hrs

First Strand Synthesis

Second Strand Synthesis

Clean-up

Library Amplification

PCR

Clean-up

~ 1.5 hrs

PCR

Clean-up

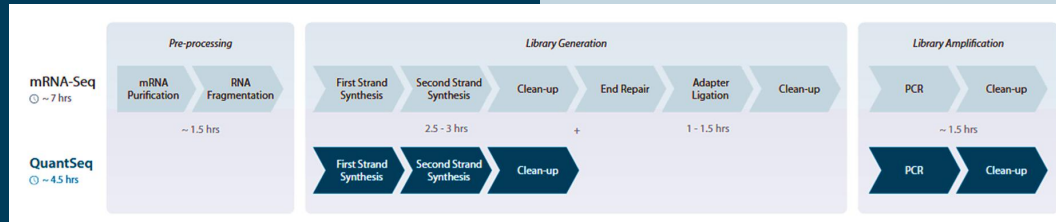
➤ Différence avec RNA-seq « classique »

3' mRNA-seq

RNA-seq « classique »

Pas de fragmentation des ARN messagers

Fragmentation des ARN messagers



Une seule lecture par transcrit



Plusieurs lectures par transcrit



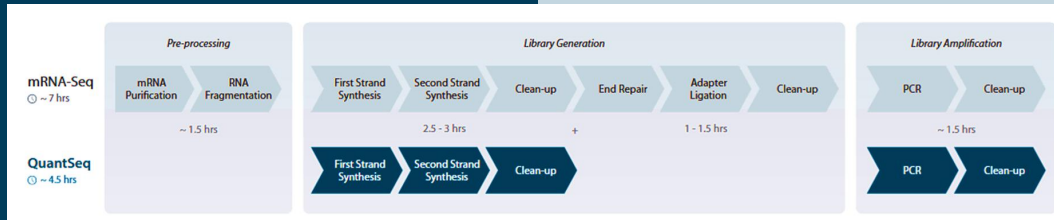
➤ Différence avec RNA-seq « classique »

3' mRNA-seq

RNA-seq « classique »

Pas de fragmentation des ARN messagers

Fragmentation des ARN messagers



Une seule lecture par transcrit



Plusieurs lectures par transcrit

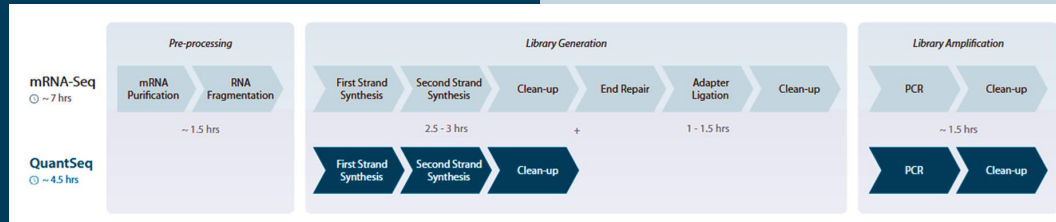


Quantifier l'expression des gènes codants
Analyser l'expression différentielle des gènes (DGE)

➤ Différence avec RNA-seq « classique »

3' mRNA-seq

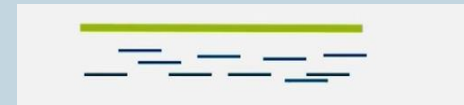
Pas de fragmentation des ARN messagers



Une seule lecture par transcrit



Plusieurs lectures par transcrit



Quantifier l'expression des gènes codants
Analyser l'expression différentielle des gènes (DGE)

- Meilleure détection des petits transcrits

Limitations :

- Découverte de nouveaux transcrits ou variants d'épissage
- Détecter les isoformes proches de l'extrémité 5'
- Analyser des transcrits complets, des gènes de fusion

- Transcriptome entier
- Découverte de nouveaux transcrits ou variants d'épissage
- Identifier les isoformes connues et nouvelles
- Détecter les fusions de gènes

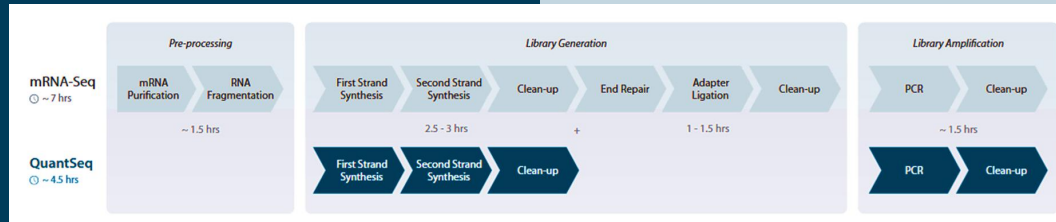
➤ Différence avec RNA-seq « classique »

3' mRNA-seq

RNA-seq « classique »

Pas de fragmentation des ARN messagers

Fragmentation des ARN messagers



Une seule lecture par transcrit



Plusieurs lectures par transcrit



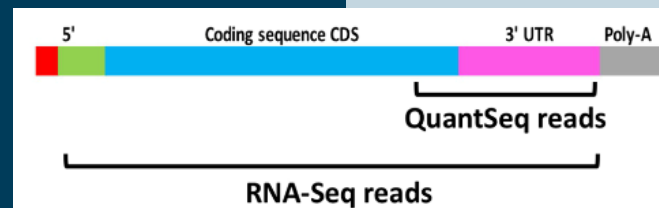
Quantifier l'expression des gènes codants
Analyser l'expression différentielle des gènes (DGE)

- Meilleure détection des petits transcrits
- Limitations :
- Découverte de nouveaux transcrits ou variants d'épissage
 - Détecter les isoformes proches de l'extrémité 5'
 - Analyser des transcrits complets, des gènes de fusion

- Transcriptome entier
- Découverte de nouveaux transcrits ou variants d'épissage
- Identifier les isoformes connues et nouvelles
- Détecter les fusions de gènes

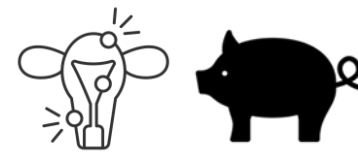
Profondeur de séquençage : 5M reads/éch
Séquençage rationalisé de l'extrémité
3' des ARN polyadénylés

Profondeur de séquençage : 30M à 100 M reads/éch
Séquençage sur toute la longueur du
transcrit



➤ QuantSeq 3' mRNA-seq sur tissus

- Conditions expérimentales validées
- Analyse comparative avec des données RNA-seq « classique »
- Identification de la profondeur de séquençage optimale



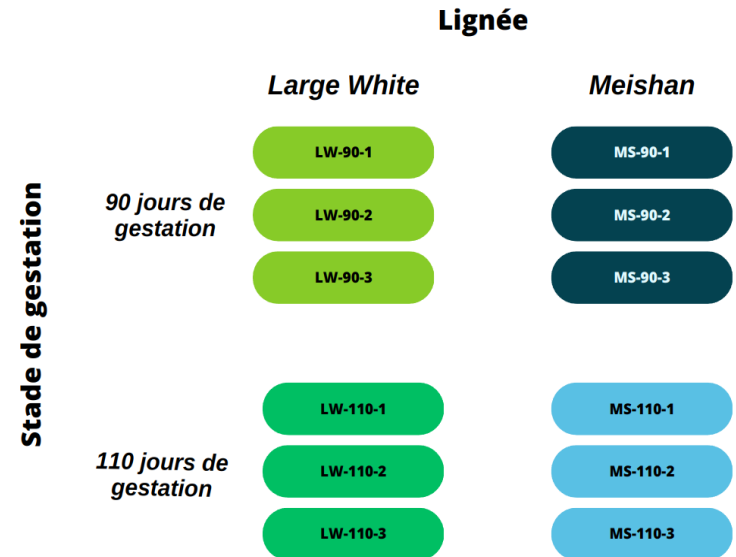
➤ Mise au point sur tissus

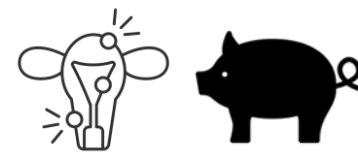
Conditions expérimentales

- 12 échantillons tissulaires **endomètre de truie** (cohorte PORCINET, GenpHySE) :
 - 2 génotypes :
 - Large White LW
 - Meishan MS
 - 2 stades de gestations (90-110j)
- Input : **500 ng ARN total**
- Séquençage **1 X 75 cycles**



- Environ **30M reads** par échantillon (pour voir l'impact de la profondeur)

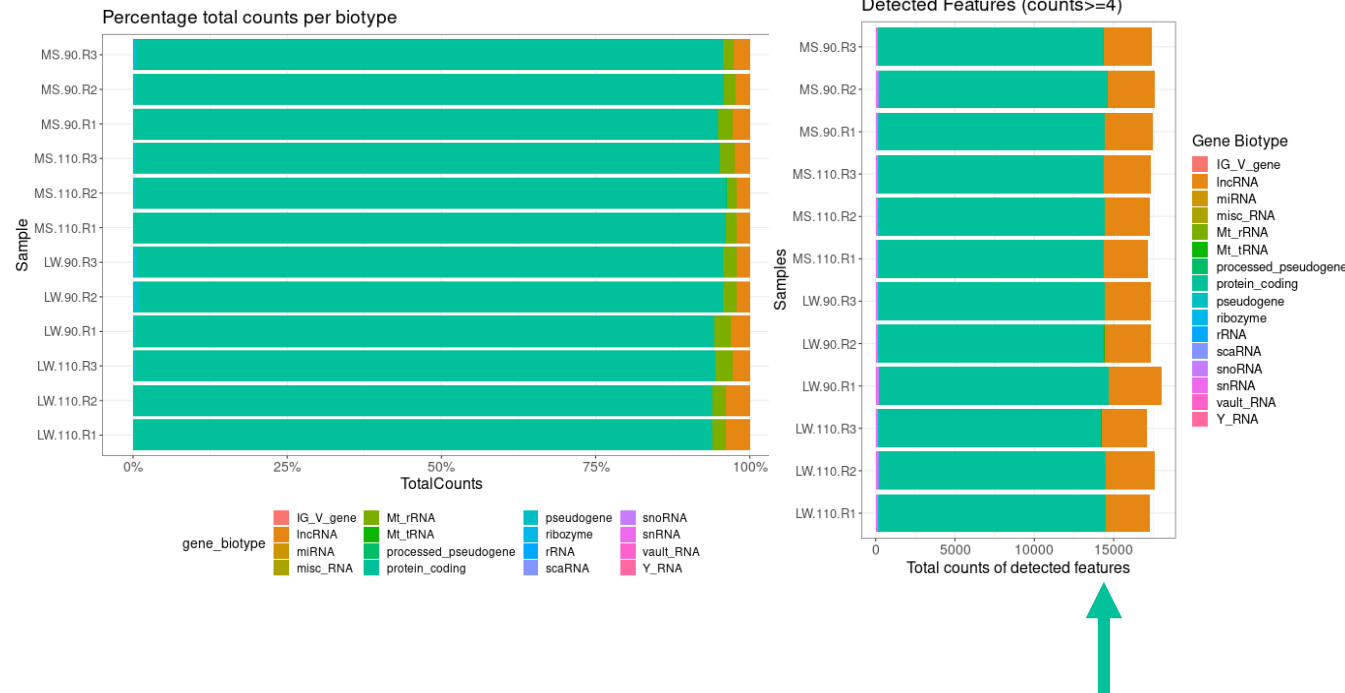




➤ Mise au point sur tissus

Résultats QuantSeq

- La majorité des gènes exprimés codent pour **des protéines** (environ 14 000).
- Une faible quantité, codent pour **des longs ARN non codants** et **des ARN ribosomiques mitochondriaux** (<5%).

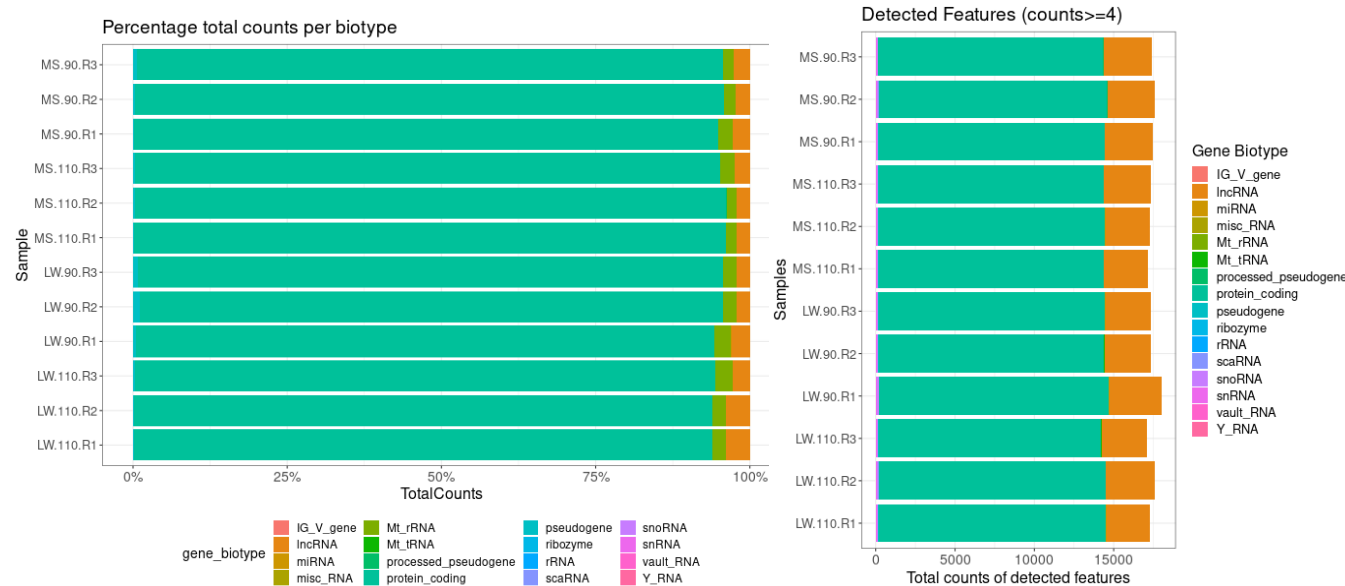




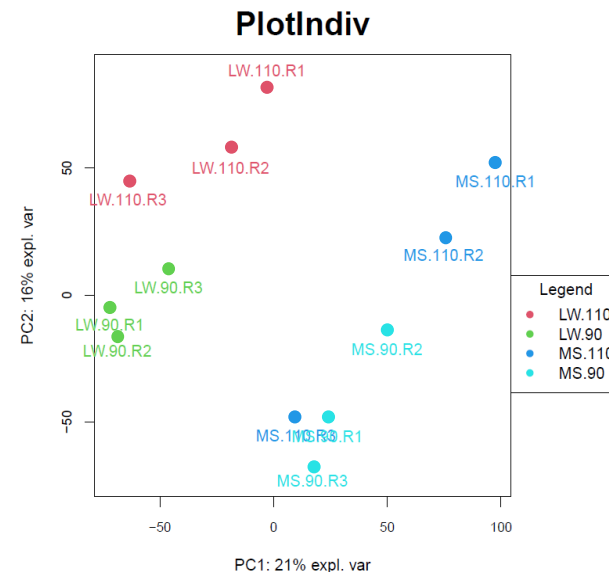
➤ Mise au point sur tissus

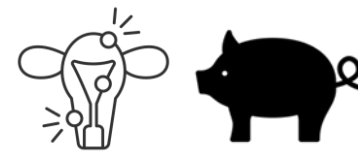
Résultats QuantSeq

- La majorité des gènes exprimés codent pour **des protéines** (environ 14 000).
- Une faible quantité, codent pour **des longs ARN non codants** et **des ARN ribosomiques mitochondriaux** (<5%).



- La variance entre les échantillons s'expliquent principalement :
 - par le génotype de la mère LW vs MS
 - puis le temps de gestation 90 vs 110 j

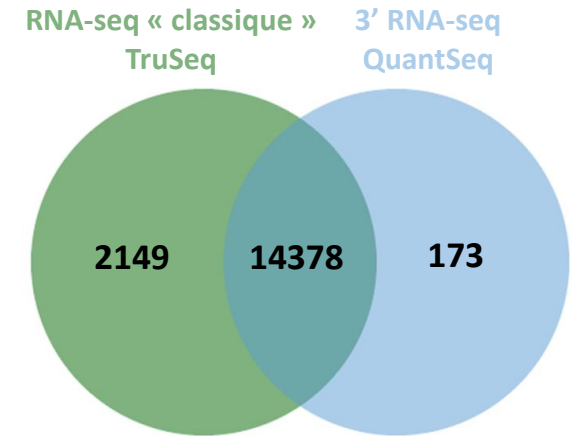




➤ Mise au point sur tissus

Analyse comparative avec les données RNA-seq « classiques » (TruSeq)

- **87%** des gènes codants pour des protéines identifiés en RNA-seq « classique » sont identifiés avec le 3' RNA-seq.
→ **14 378 gènes**





➤ Mise au point sur tissus

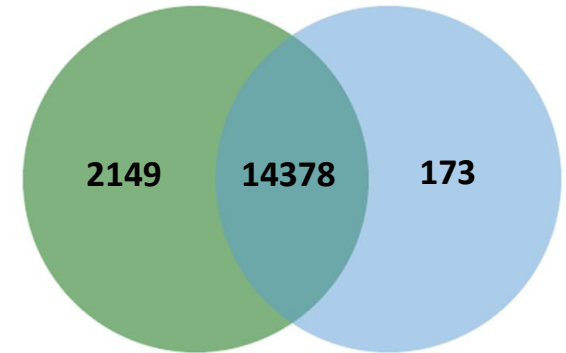
Analyse comparative avec les données RNA-seq « classiques » (TruSeq)

- **87%** des gènes codants pour des protéines identifiés en RNA-seq « classique » sont identifiés avec le 3' RNA-seq.

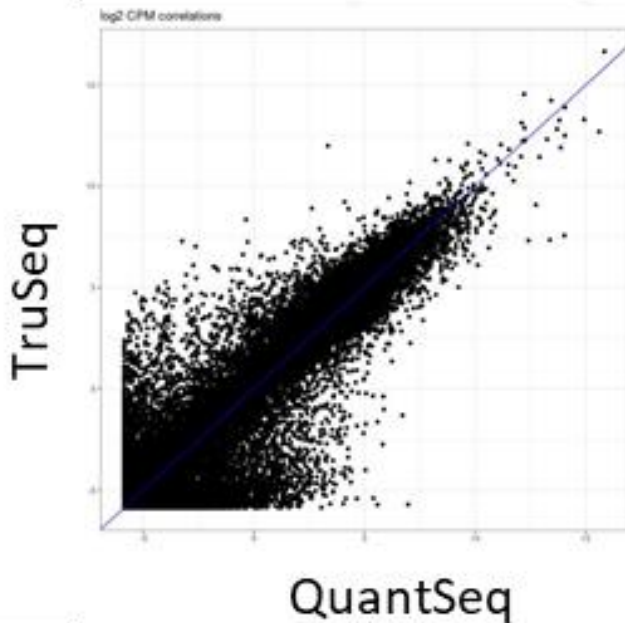
➔ **14 378 gènes**

RNA-seq « classique »
TruSeq

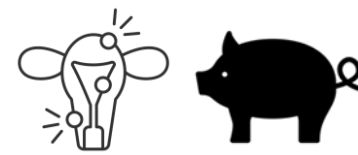
3' RNA-seq
QuantSeq



Pearson cor > 0,86



➔ Bonne corrélation des données

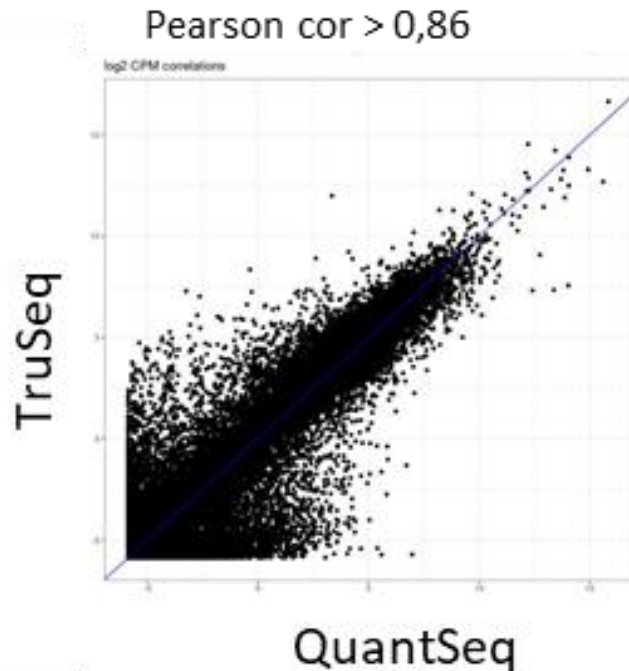
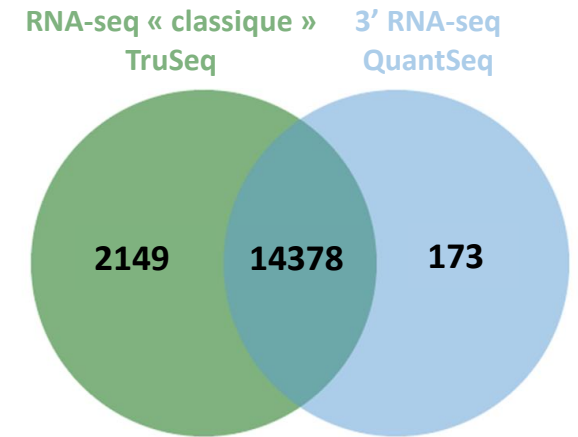


➤ Mise au point sur tissus

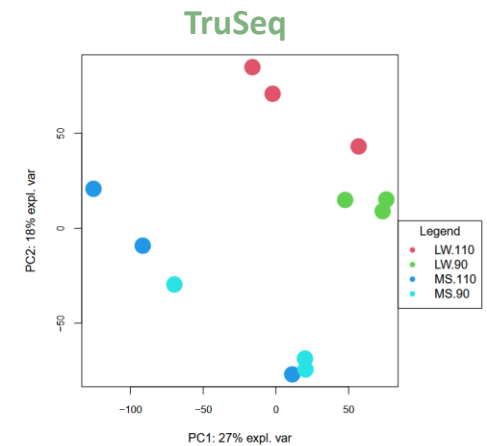
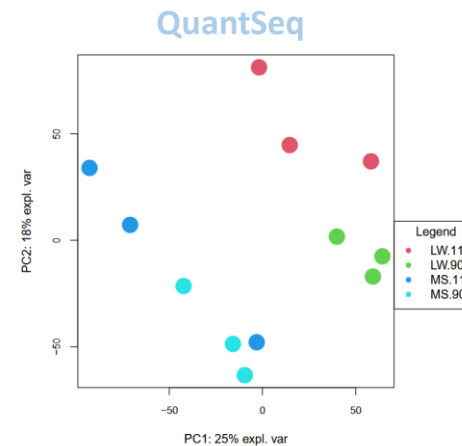
Analyse comparative avec les données RNA-seq « classiques » (TruSeq)

- **87%** des gènes codants pour des protéines identifiés en RNA-seq « classique » sont identifiés avec le 3' RNA-seq.

➔ **14 378 gènes**



➔ Bonne corrélation des données



➔ Structuration similaire des données

➔ **Le QuantSeq donne des données comparables au TruSeq**

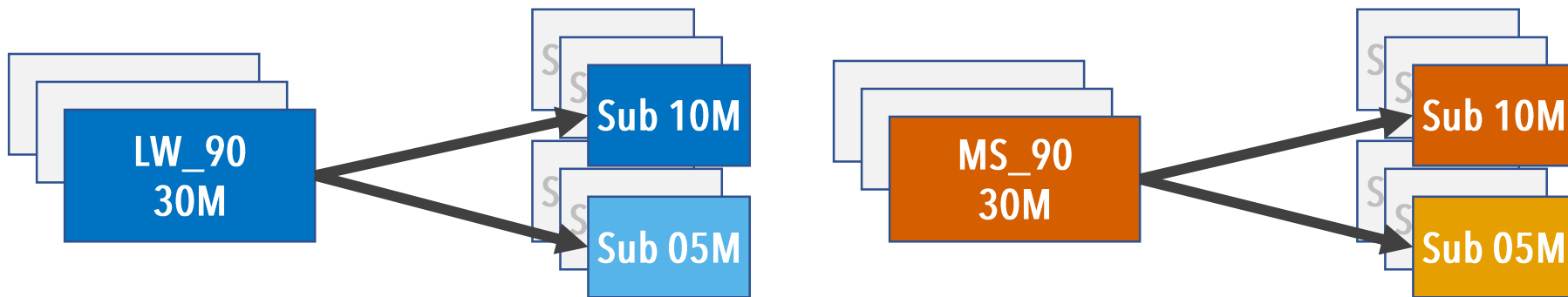


➤ Mise au point sur tissus

Identification de la profondeur de séquençage optimale

Comparer les résultats obtenus avec 05M et 10M de lectures

- **Etape 1 :** Définir le nombre de répétitions minimal et suffisant pour représenter la diversité des données initiales → **12 répétitions**
- **Etape 2 : Sous-échantillonnage :** 12 tirages aléatoires des lectures obtenues pour chaque individu à partir des 30M de lectures
 - pour la profondeur à **05M**
 - pour la profondeur à **10M**



- **Etape 3 :** Exploration et comparaison des données à **05M** et **10M** de lectures

➤ Mise au point sur tissus

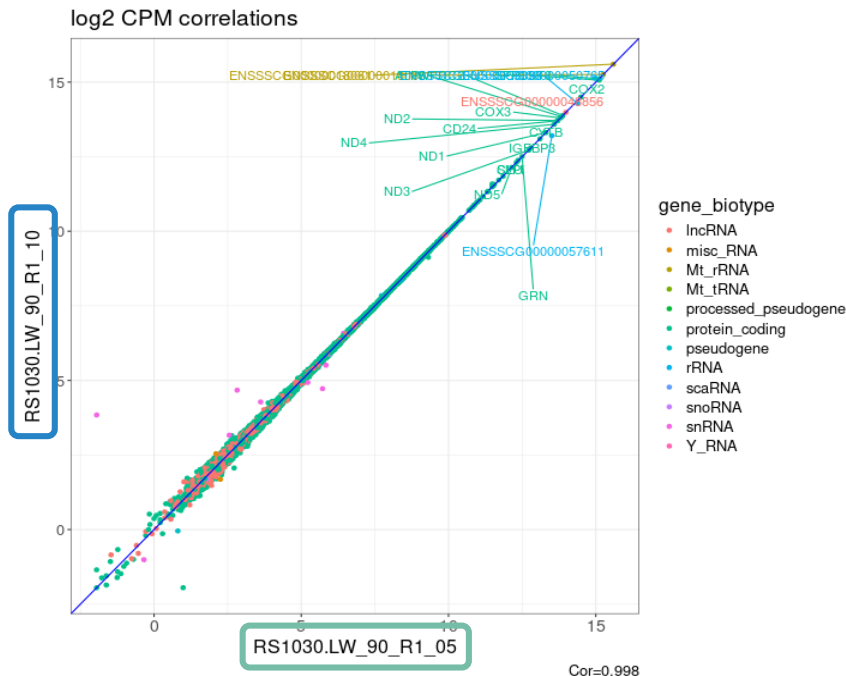
Données filtrées sur les
Moyennes des 12 itérations
Sous-échantillonnage
05 millions de lectures
10 millions de lectures



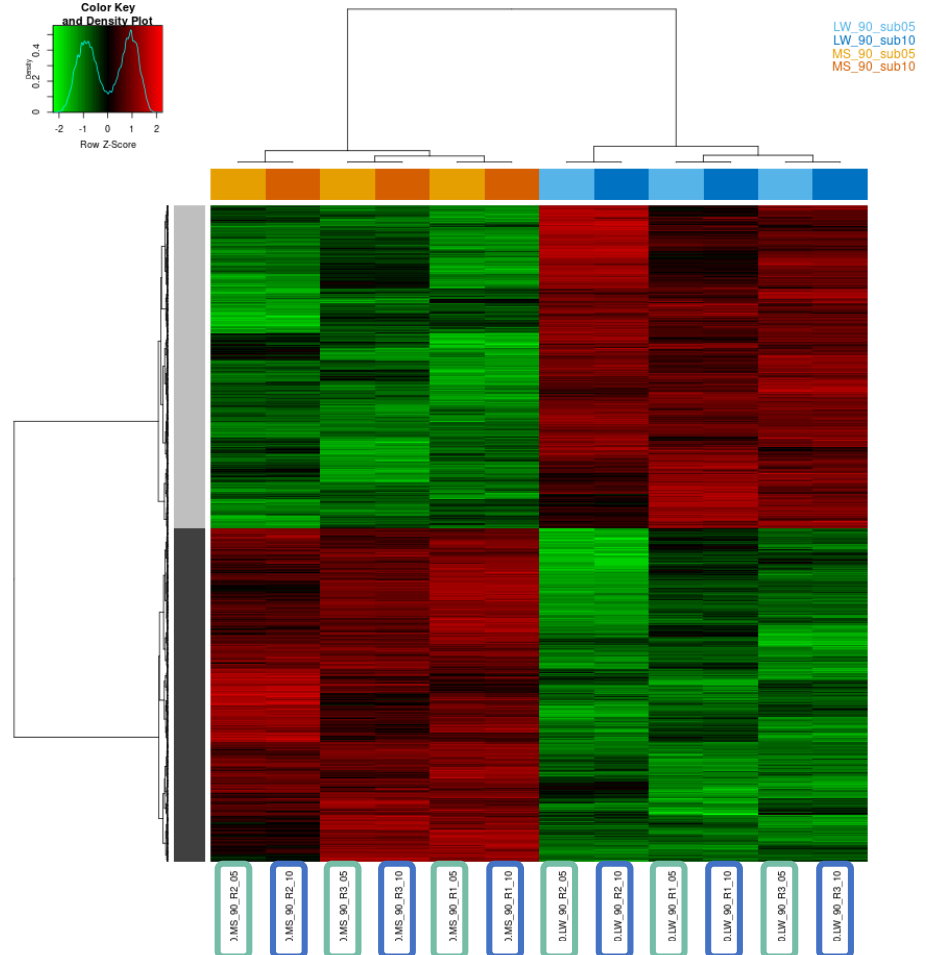
Identification de la profondeur de séquençage optimale

- Très bonnes corrélations 05M vs 10M, même pour les très faibles comptages !
- Pas de gènes spécifiquement détectés selon 05M ou 10M

Exemple pour un individu LW_90_R1



- Très bonnes corrélations 05M vs 10M, même pour les très faibles comptages !
- Pas de gènes spécifiquement détectés selon 05M ou 10M
- Réponses similaires à 05M et 10M



24

➤ QuantSeq 3' mRNA-seq sur sang total

- Conditions expérimentales validées
- Analyse comparative : efficacité du blocage des globines
- Analyse comparative avec des données MACE 3' RNA-seq et Microarray



➤ Mise au point sur sang total

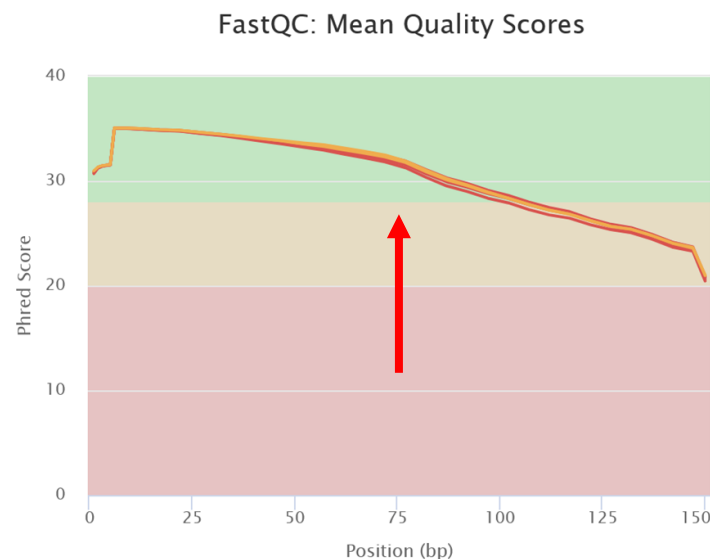
Conditions expérimentales

- 6 échantillons **sang total de cochon** (cohorte PigChange, GenPhySE) :
 - 3 génotypes (Large White, Creole, Croisés)
 - 2 conditions (Température 24C et 30C)
- Input : **200 ng ARN total**
- Utilisation **RS-Globin Block, *Sus scrofa*** : empêche la génération de fragments de librairies à partir d'ARNm de globine
- Séquençage **1 X 150 cycles**



→ **75 cycles suffit** perte de qualité après

		Genotype		
		LW	CR	Croisés
Condition	24 C	ST-2 ST-3	ST-1	ST-5
	30 C		ST-4	ST-6



➤ Mise au point sur sang total



Analyse comparative : Efficacité du blocage des globulines

Quant Seq 3' RNA-Seq

+ Bloqueur globine

15 156 gènes détectés

MACE 3' RNA-Seq

8 812 gènes détectés

Micro Array

5 771 gènes détectés

➤ Mise au point sur sang total



Analyse comparative : Efficacité du blocage des globulines

Quant Seq 3' RNA-Seq

+ Bloqueur globine

15 156 gènes détectés

13-30 % HB
≈ 12 000 gènes hors HB

MACE 3' RNA-Seq

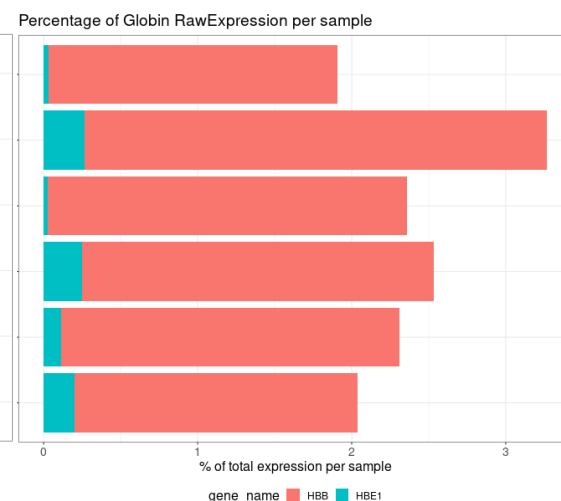
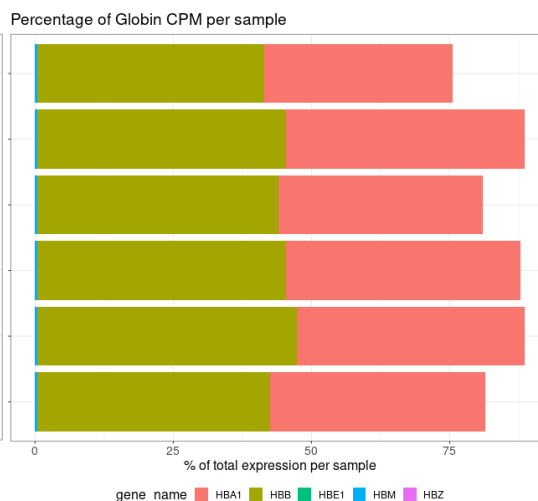
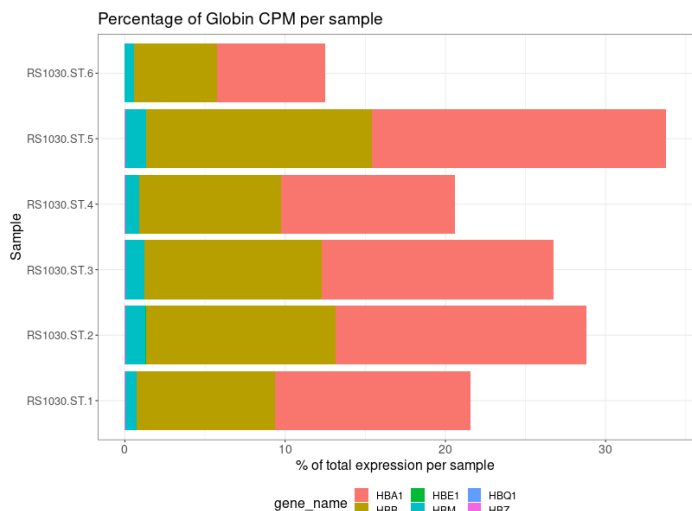
8 812 gènes détectés

>75% HB
≈ 2 000 gènes hors HB

Micro Array

5 771 gènes détectés

2-3% HB
≈ 5 500 gènes hors HB



➤ Mise au point sur sang total



Analyse comparative : Efficacité du blocage des globulines

Quant Seq 3' RNA-Seq

+ Bloqueur globine

15 156 gènes détectés

13-30 % HB

≈ 12 000 gènes hors HB

MACE 3' RNA-Seq

8 812 gènes détectés

>75% HB

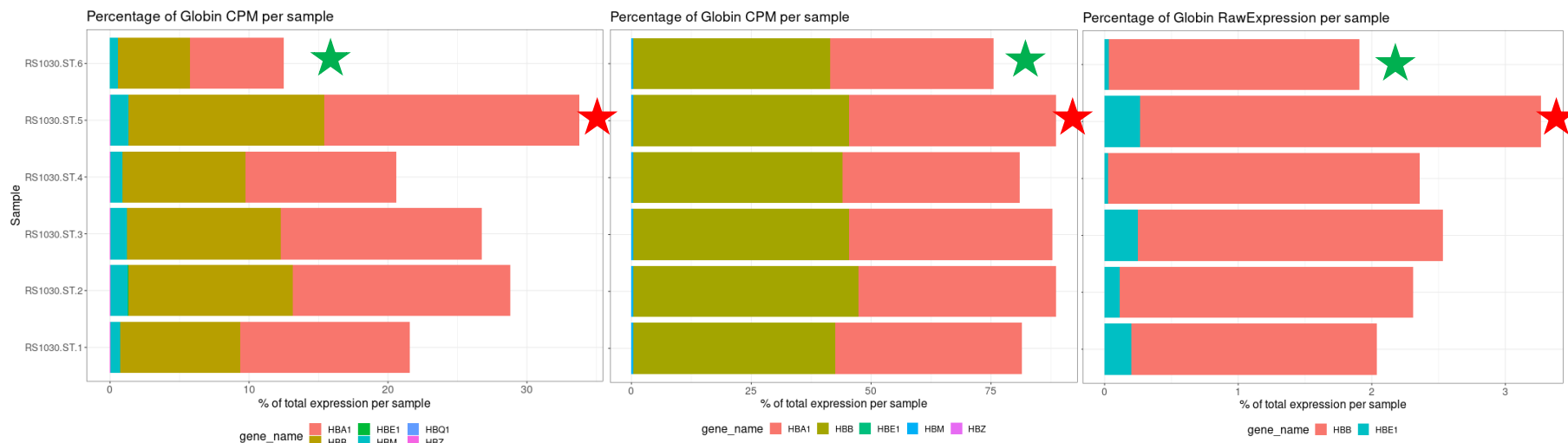
≈ 2 000 gènes hors HB

Micro Array

5 771 gènes détectés

2-3% HB

≈ 5 500 gènes hors HB



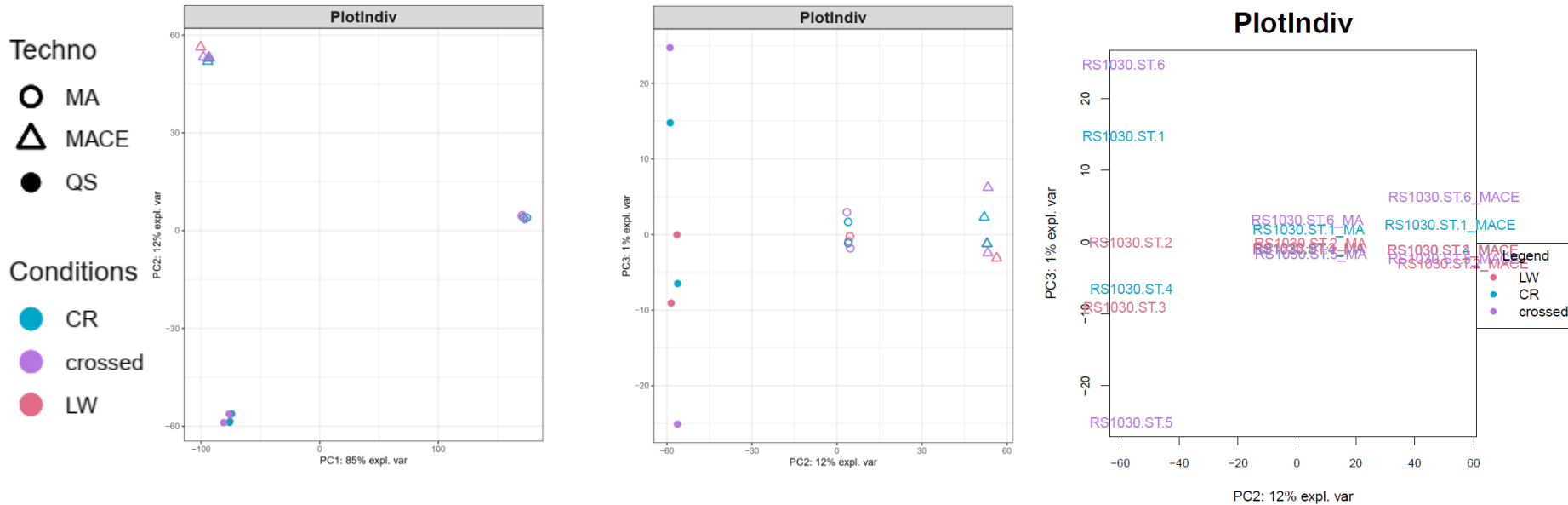
- Différences entre échantillons conservées entre les 3 technologies :
ST5 ★ le plus « contaminé » et ST6 ★ le moins en hémoglobine

➔ **Validation de l'efficacité du bloqueur de globine, plus de gènes exploitables**



➤ Mise au point sur sang total

Analyse comparative avec des données MACE 3' RNA-Seq et Micro array



PC1 (axe horizontal) = Microarray ○ vs QuantSeq ● et MACE △

PC2 (axe vertical) = QuantSeq vs MACE

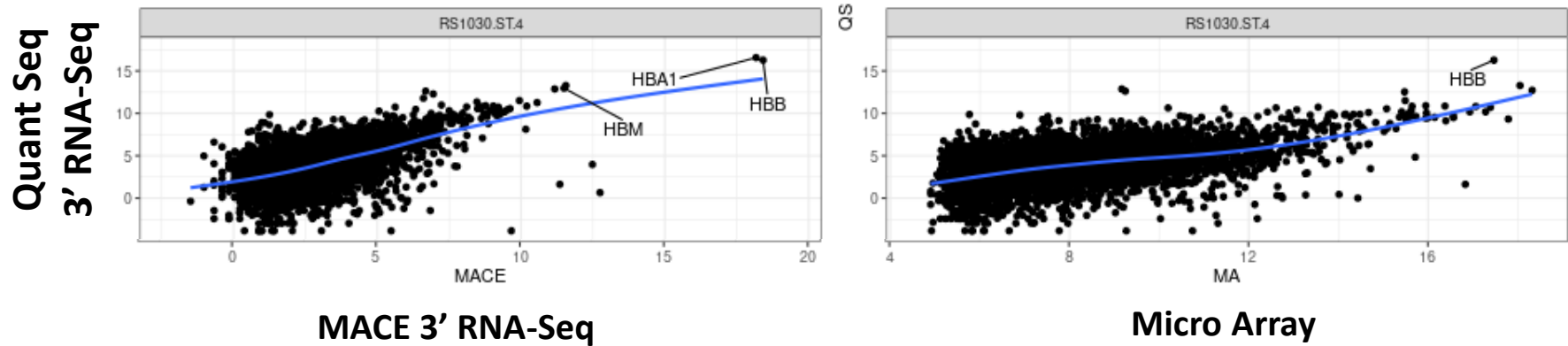
PC3 (Axe vertical) = individus intra techno

- **La technologie** structure les données (PC1 et PC2) mais des **effets échantillons** sont communs aux 3 technologies (PC3)

➤ Mise au point sur sang total



Analyse comparative avec des données MACE 3' RNA-Seq et Micro array

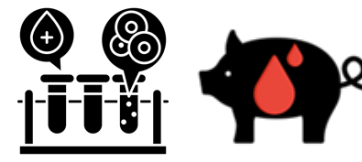


- Corrélations significatives entre les technologies ($R \sim 0,5$)
- Pas de gros biais de distribution
- HBB/HBA les plus exprimés quelle que soit la techno

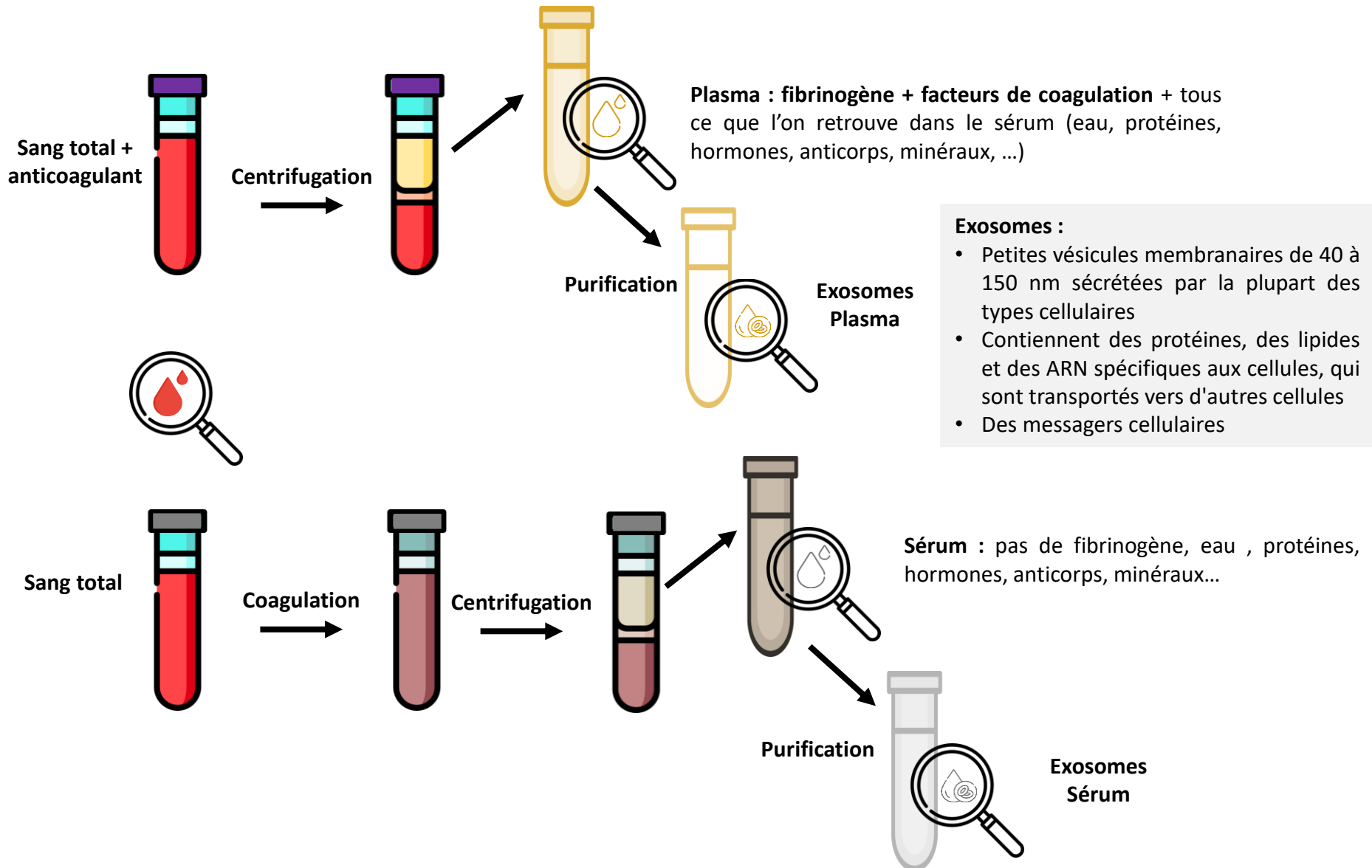
➔ Quant Seq sur sang donne des résultats exploitables et corrélés aux autres technologies MACE 3' RNA-Seq et Micro array

➤ QuantSeq 3' mRNA-seq sur différents compartiments sanguins : sang total, sérum, plasma, exosomes

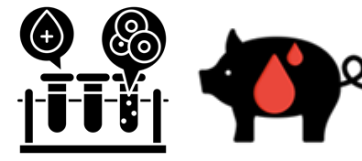
- Conditions expérimentales validées
- Résultats : exploration biostatistique



➤ Différents compartiments sanguins



➤ Différents compartiments sanguins



Plasma



Exosomes
Plasma



Sang total



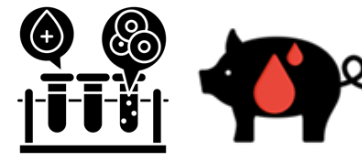
Sérum



Exosomes
Sérum

Intérêt :

- **Bio fluide facile** à obtenir
- **Prélèvement non létal**
- **Suivi longitudinal : Possibilité** mesures répétées
- Mise en évidence de **biomarqueurs**.....



➤ Différents compartiments sanguins

Conditions expérimentales

- 4 jeunes truies de 5 semaines (Duroc DanBred X Hybride DanBred) (BioToMyc, P.Pinton)
- 20 échantillons
- Différentes extractions ARN (GenPhySE) et préparations des librairies



Sang total



Plasma



Sérum



**Exosomes
Plasma**



**Exosomes
Sérum**

Extraction ARN PAXGENE

- Input : 200ng
- Nb de cycle PCR : **17**
- **Protocole Standard**
- **RS-Globin Block**

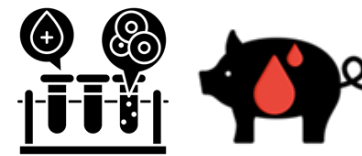
Extraction ARN STD Norgen

- Input : 5µl variable
- Nb de cycle PCR : **21**
- **Protocole low quality / input**

Purification des exosomes Extraction ARN Exosomes Norgen

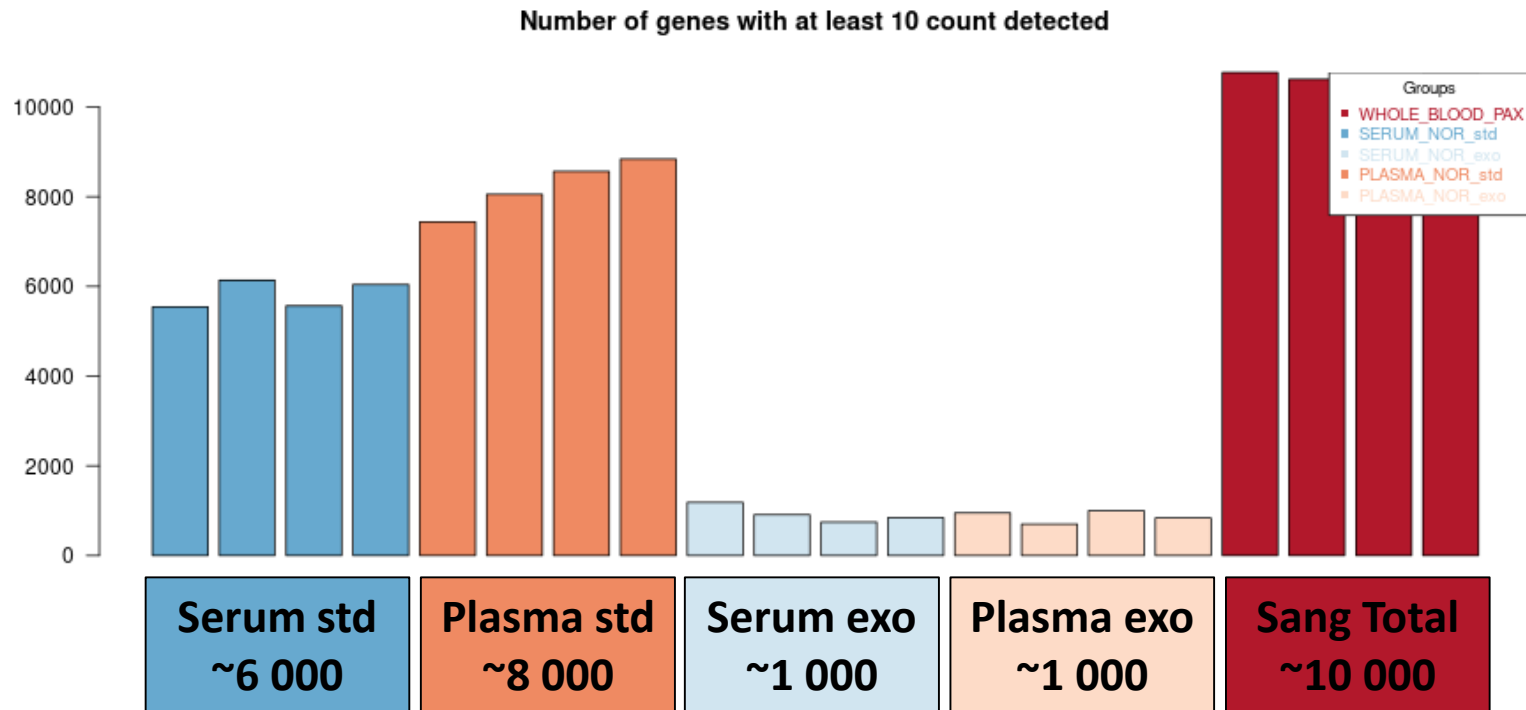
- Input : 5µl variable
- Nb de cycle PCR : **26**
- **Protocole low quality / input**

- Séquençage **1 X 76 cycles**

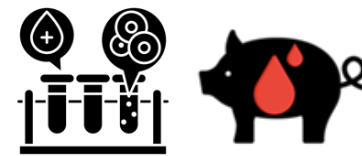


➤ Différents compartiments sanguins

Résultats : Exploration biostatistique : Nombre de gènes



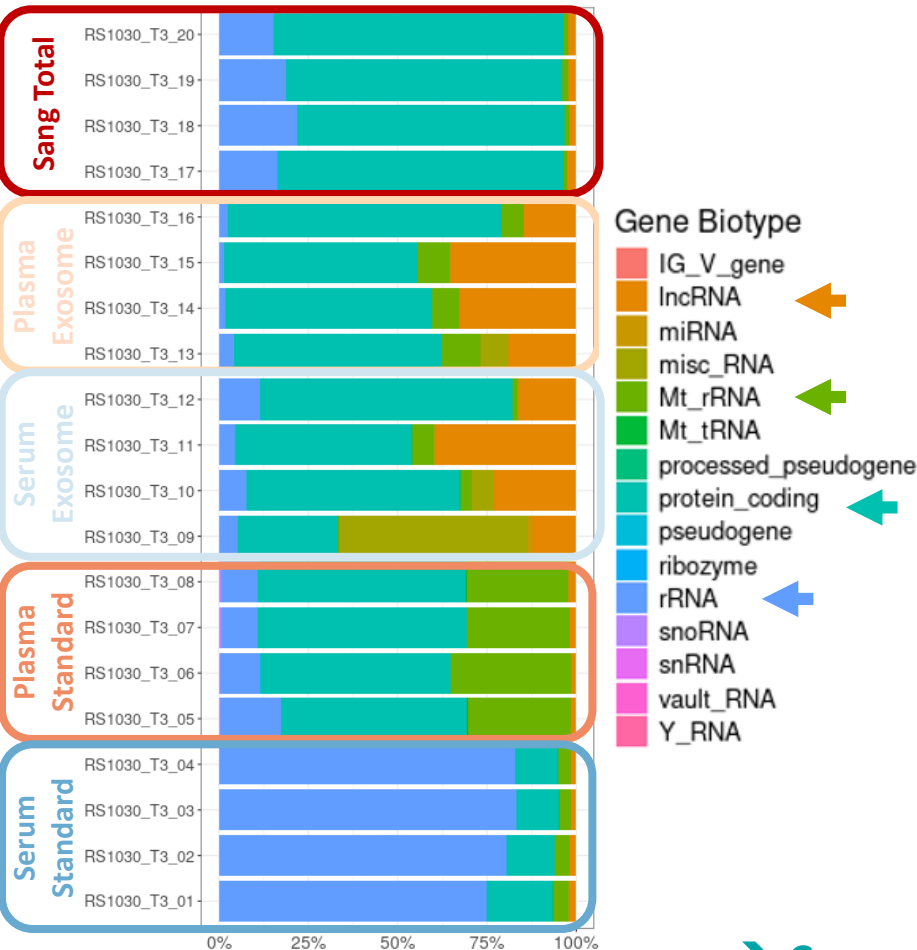
➔ En fonction du compartiment, le nombre de gènes détectés **varie** mais reste **reproductible pour les quatre individus**



➤ Différents compartiments sanguins

Résultats : Exploration biostatistique : Composition

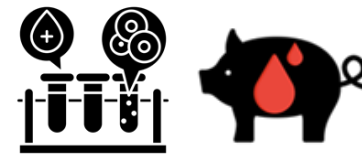
Percentage Total counts per biotype



	biotypes	# genes
Sang total	Protein > rRNA	~10 000
Plasma_Exo	Protein > lncRNA	~1 000
Serum_Exo	Protein > lncRNA > rRNA	~1 000
Plasma_std	Protein > Mt_rRNA	~8 000
Serum_std	rRNA > protein	~6 000

➔ **Compositions différentes :**

Protein coding majoritaires dans le Sang total et le Plasma mais pas dans le Serum

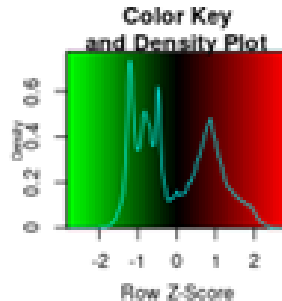


➤ Différents compartiments sanguins

Résultats : Exploration biostatistique : Analyse différentielle

WORK IN
PROGRESS

-
Gènes
Faiblement
exprimés

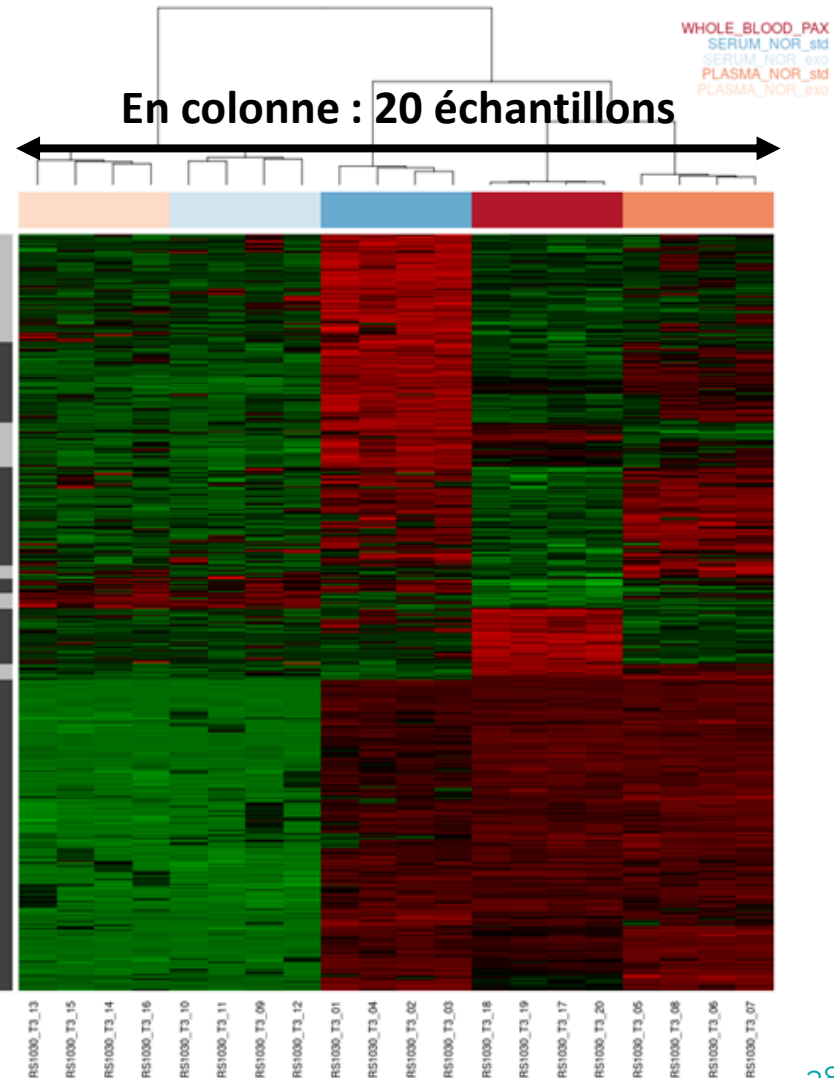


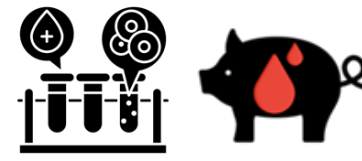
+
Gènes
Fortement
exprimés

En ligne :
Top des 500 gènes
10 clusters

Heatmap Top-500 FDR<5%

En colonne : 20 échantillons



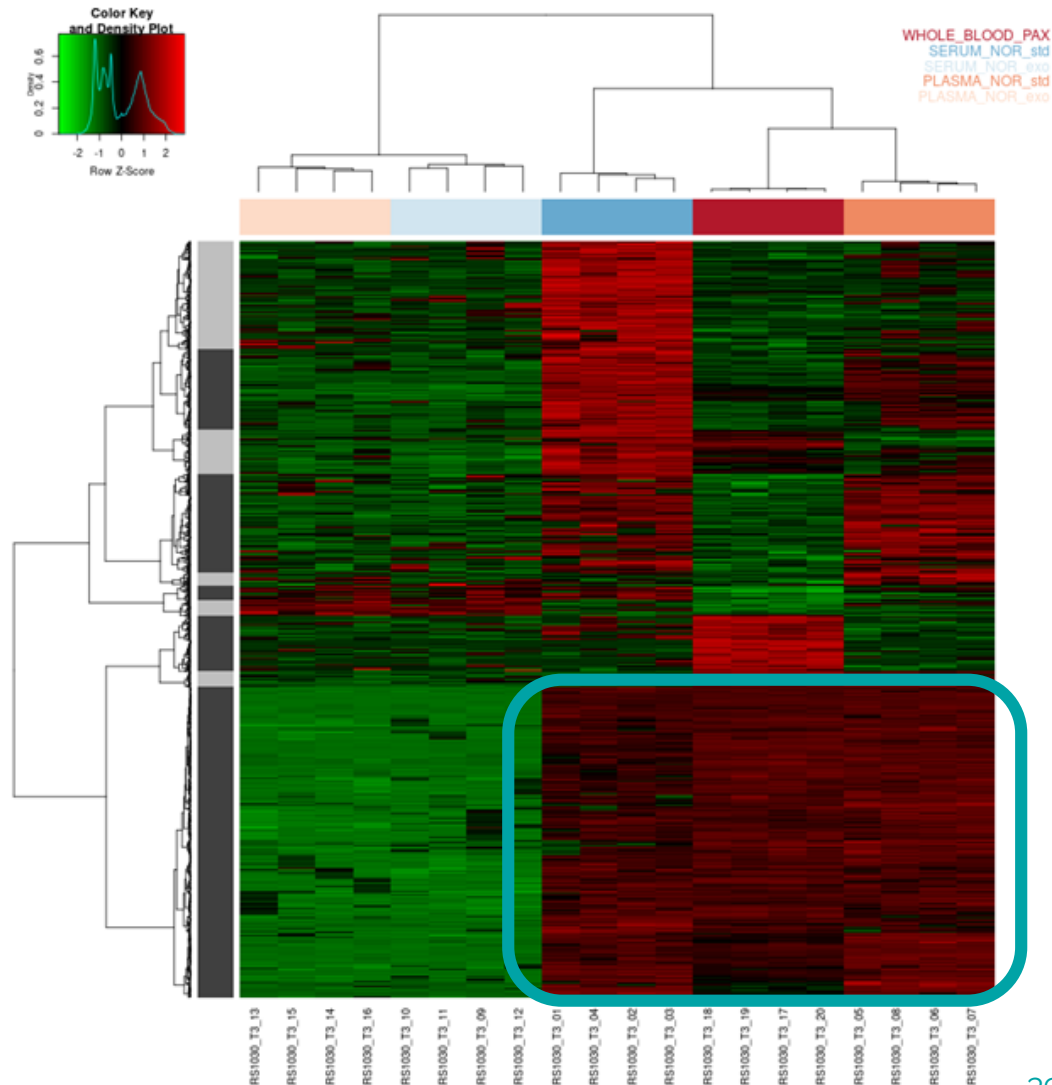


➤ Différents compartiments sanguins

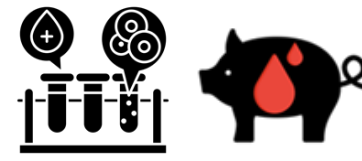
Résultats : Exploration biostatistique : Analyse différentielle

WORK IN
PROGRESS

Heatmap Top-500 FDR<5%



- De nombreux transcrits codant des protéines sont exprimés plus fortement dans le **Sang Total**, le **Plasma** et le **Serum**

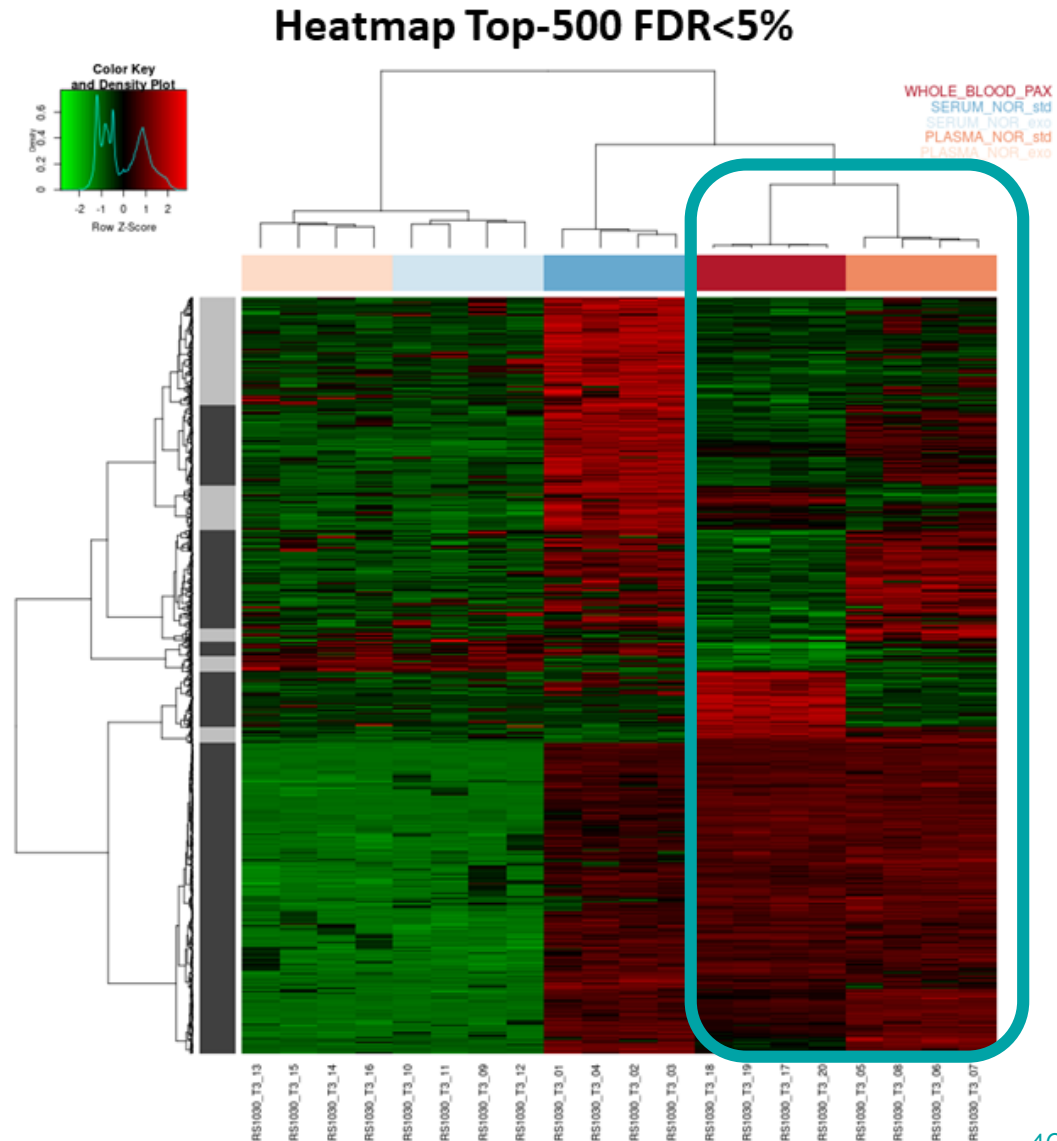


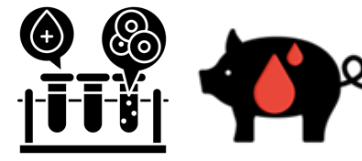
➤ Différents compartiments sanguins

Résultats : Exploration biostatistique : Analyse différentielle



- Le **Plasma** donne des profils d'expressions les plus proches du **Sang Total** comme on le voit avec la distance



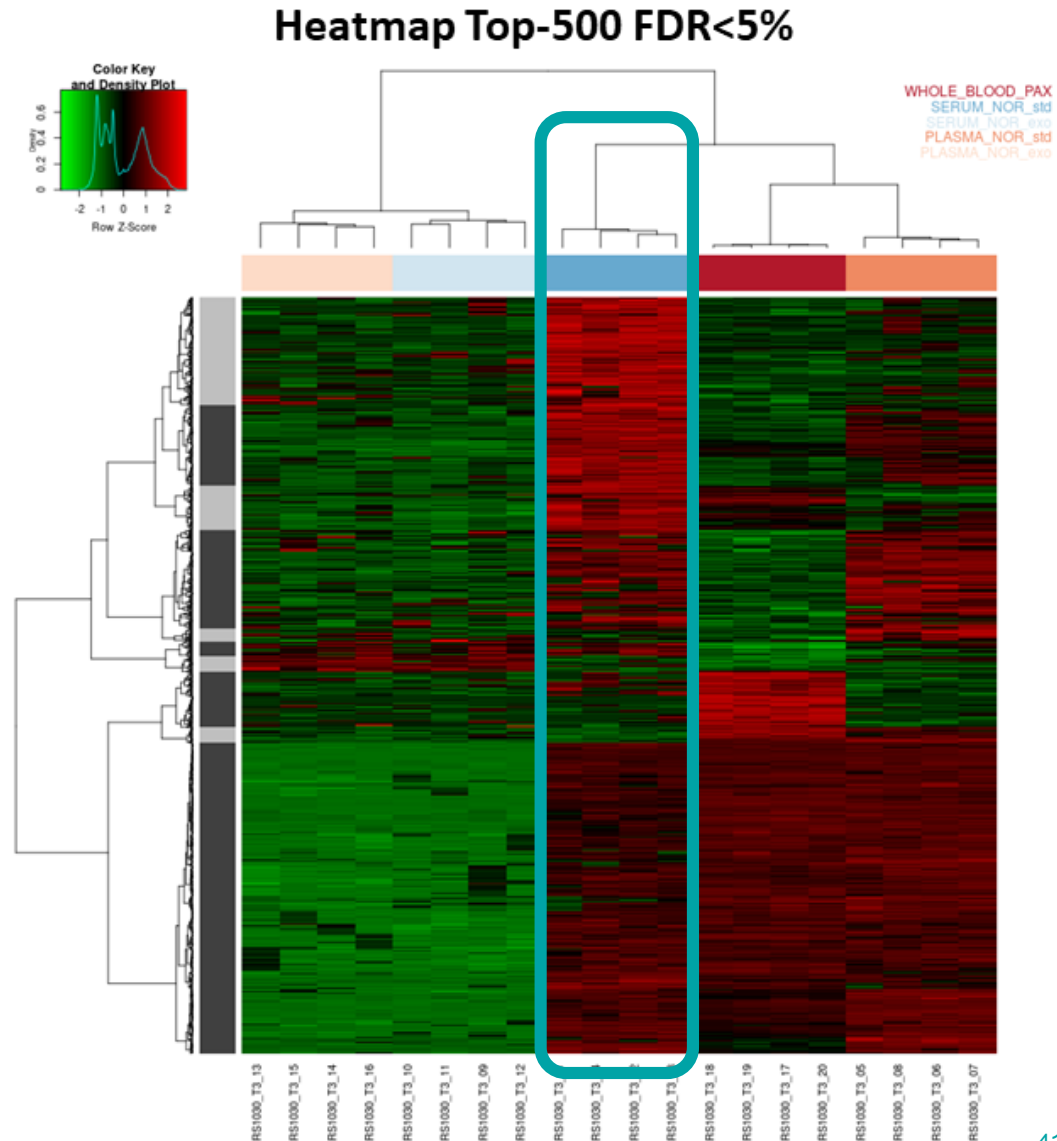


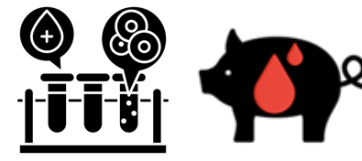
➤ Différents compartiments sanguins

Résultats : Exploration biostatistique : Analyse différentielle



- Le **Plasma** donne des profils d'expressions les plus proches du **Sang Total** comme on le voit avec la distance
- Le **Serum** présente des profils plus spécifiques (notamment lncRNA) par rapport au **Sang Total** ou au **Plasma**



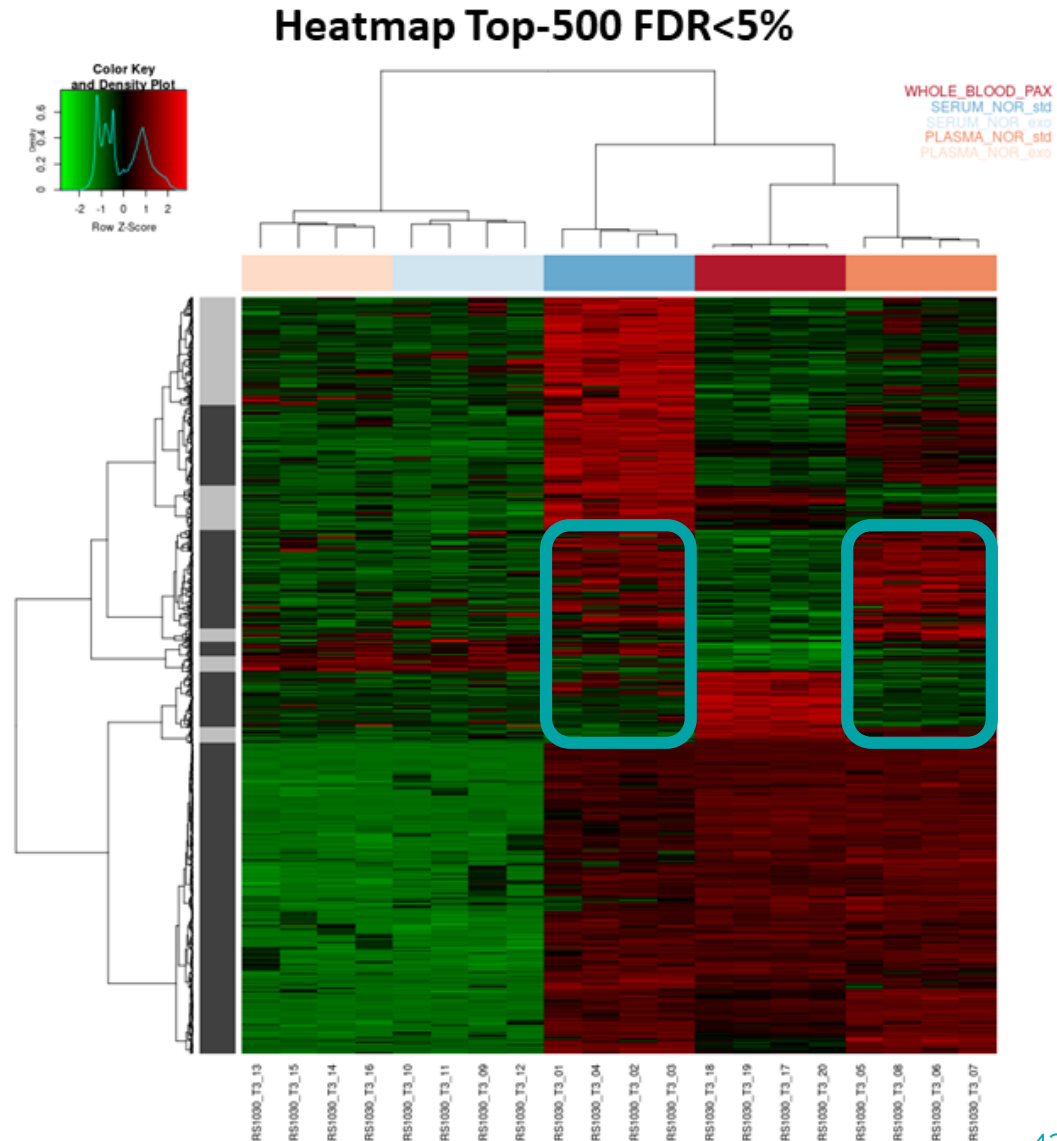


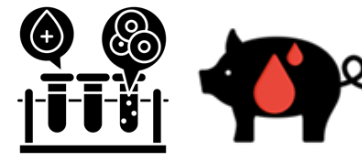
➤ Différents compartiments sanguins

Résultats : Exploration biostatistique : Analyse différentielle



- Des profils d'expression sont spécifiques et communs pour le **Plasma** et le **Serum** (protéines + lncRNA + snRNA)



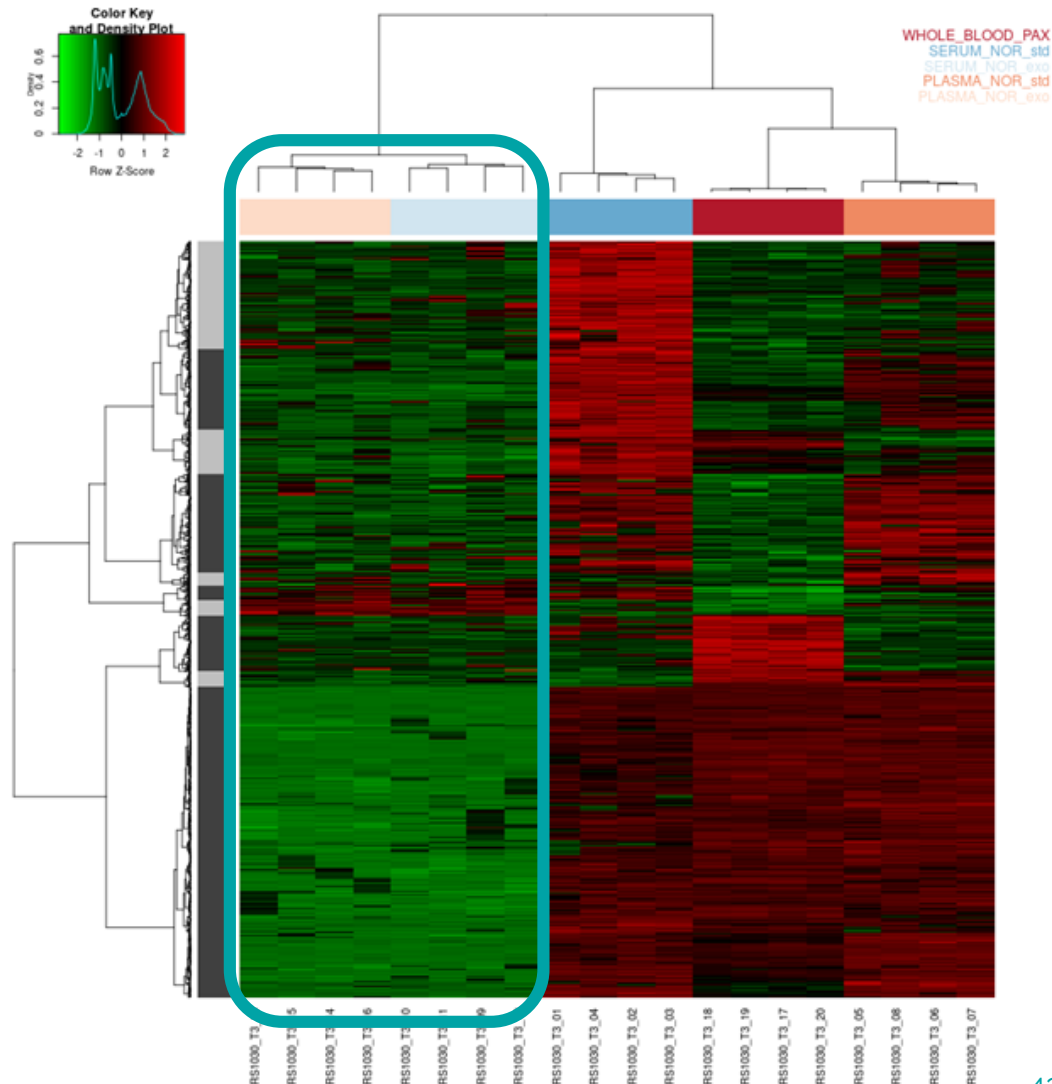


➤ Différents compartiments sanguins

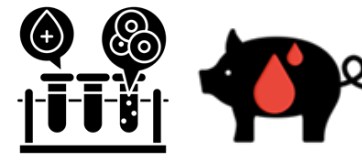
Résultats : Exploration biostatistique : Analyse différentielle



Heatmap Top-500 FDR<5%



- **Les Exosomes** présentent des profils plus spécifiques (lncRNA)



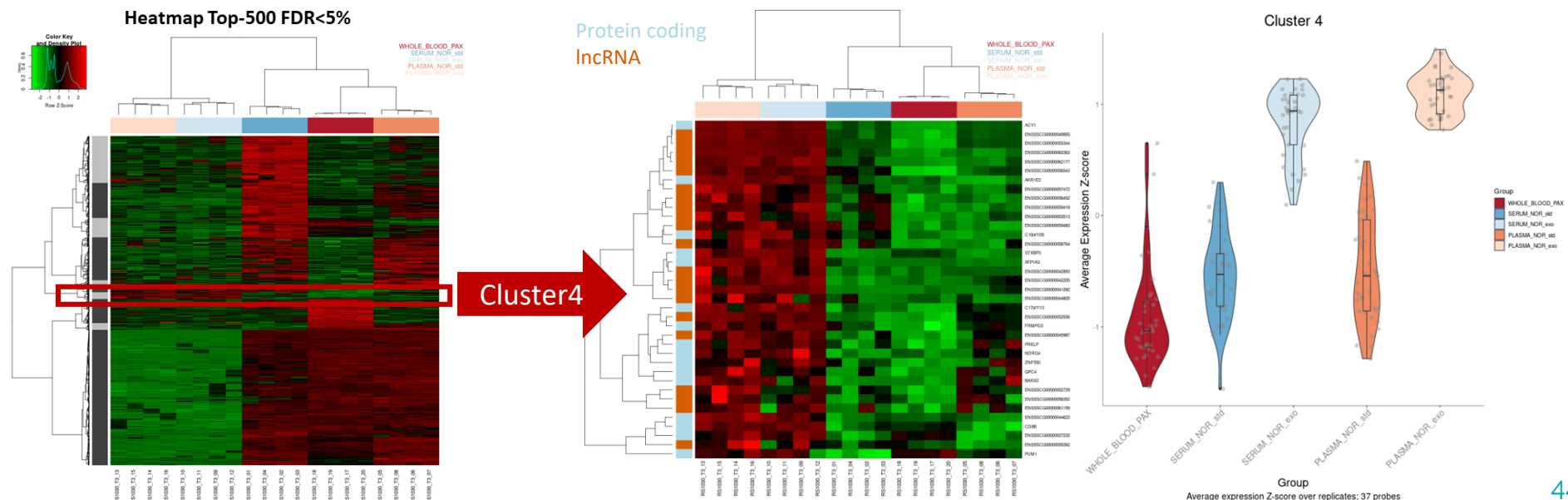
➤ Différents compartiments sanguins

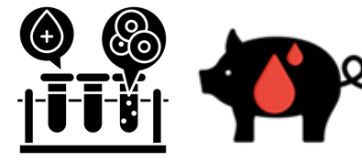
Résultats : Exploration biostatistique : Analyse différentielle



- Décryptage de chaque cluster et des gènes différentiellement exprimés associés :
 - Gènes **uniques** à un compartiment
 - Gènes **communs** à plusieurs compartiments
 - Intérêt particulier pour les longs ARN non codants et les gènes codants proches ?

Exemple du cluster 4 : spécifique des Exosomes avec 37 gènes

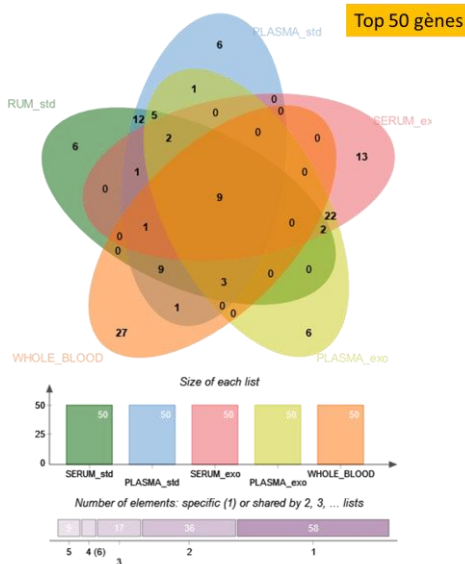




➤ Différents compartiments sanguins

Résultats : Exploration biostatistique : Analyse différentielle

- Exploration **des principales fonctions biologiques** basée sur l'enrichissement des annotations dans les **catégories GO** (Biological Process, Cellular Component, Molecular Function), ainsi que dans les bases de données **KEGG Pathways** et **Reactome**.
 - 2 approches en parallèle : analyse faite sur l'organisme *Sus scrofa* (le porc) et *Human*.



GO Molecular Function (Top 50 gènes)	WHOLE_BLOOD_PAX	SERUM_NOR_std	PLASMA_NOR_std	SERUM_NOR_exo	PLASMA_NOR_exo
Organic acid binding	x				
Haptoglobin binding	x				
Oxygen carrier activity	x				
Oxygen binding	x				
Translation regulator activity	x				
Angiotensin binding	x				
5-aminolevulinic synthase activity	x				
Structural constituent of ribosome	x	x	x		
Ubiquinol-cytochrome-c reductase activity	x	x	x		
Cytochrome-c oxidase activity	x	x	x		
2'-5'-oligoadenylate synthase activity (=apoptose, croissance)		x			
Platelet-derived growth factor binding		x			
Extracellular matrix structural constituent		x	x		
Hemoglobin alpha binding		x	x		
Large ribosomal subunit rRNA binding		x	x		
NADH dehydrogenase activity		x	x		
rRNA binding			x		
RNA binding			x		
Pyruvate synthase activity				x	
Myo-inositol symporter activity				x	
Interleukin-18 receptor activity				x	
Type I interferon receptor activity, binding				x	x
Aminocyclase activity				x	x
JAK pathway signal transduction adaptor activity				x	x
Prostaglandin-synthase activity				x	x
Heme binding					x

Réponse stress?

Marqueurs métaboliques?
Inflammatoires?

Fonction?

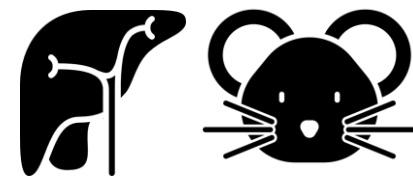
➔ **Objectif final : Data Paper pour description des données et cartographier les gènes retrouvés dans chaque compartiment et article scientifique en cours de rédaction**

➤ Projets 2025 Toxalim



- Foie Souris
- Muqueuse colon Souris
- Muqueuse jéjunum et colon Rat

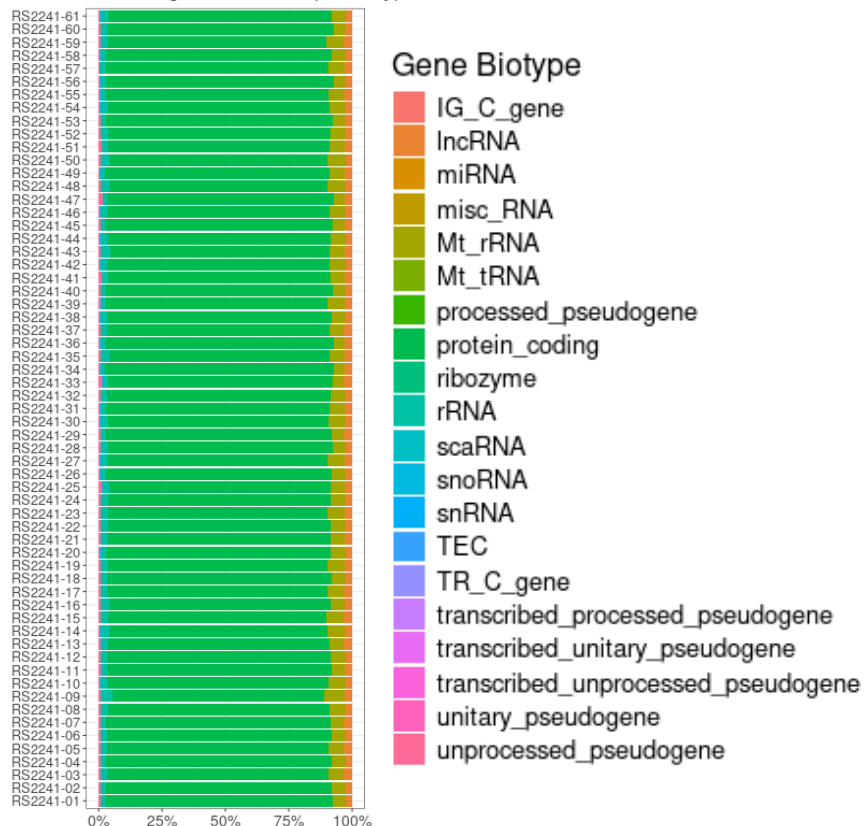
➤ Liver_Mus_musculus



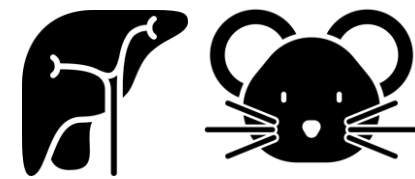
- 61 échantillons : 8 conditions (n=6 à 8), 20 M de reads/éch à la demande de l'équipe.
- **(A)** 90% des comptages correspondent à des gènes codant pour **des protéines** (puis ARNr mitochondriaux + long ARN non codants). Détection de **16 000 gènes**.

A

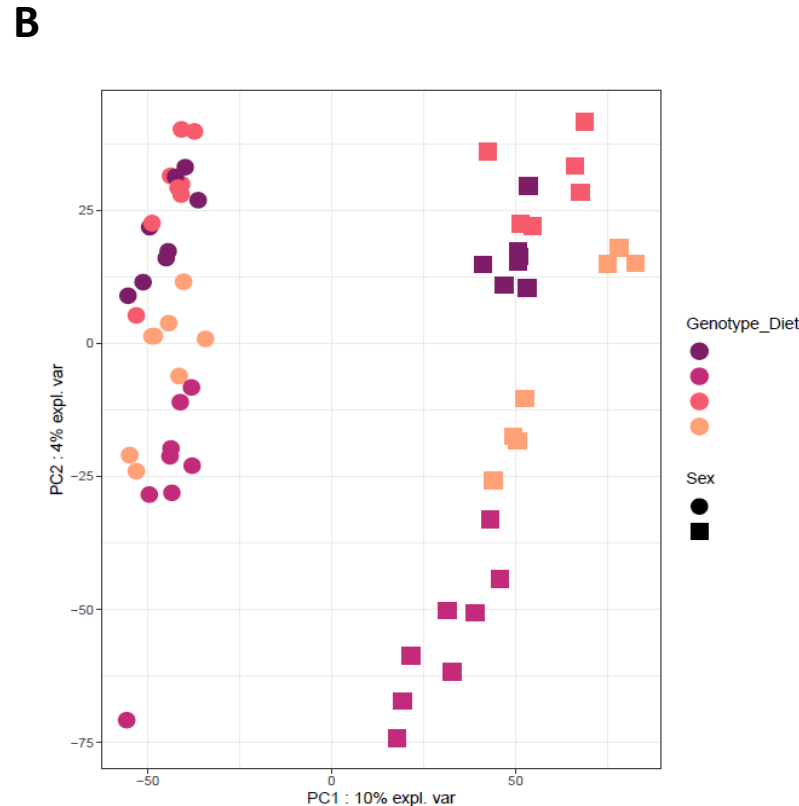
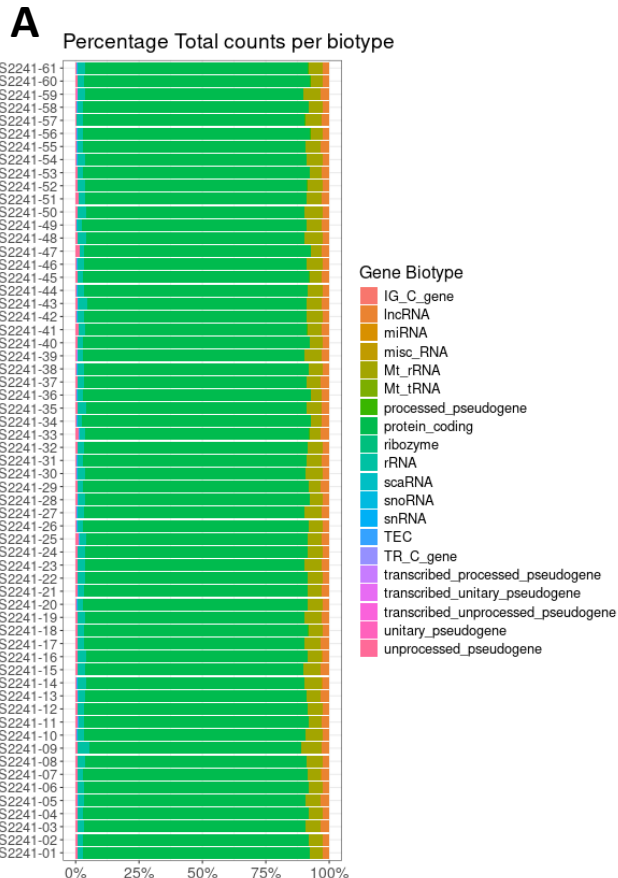
Percentage Total counts per biotype



➤ Liver_Mus_musculus



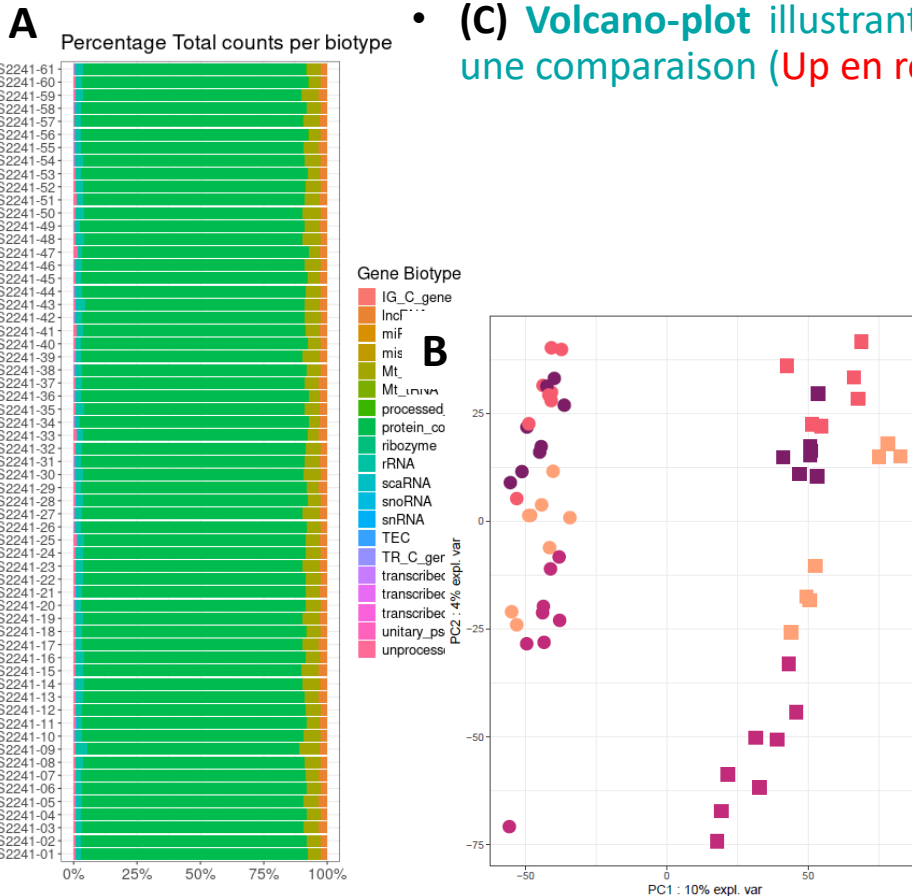
- 61 échantillons : 8 conditions (n=6 à 8), 20 M de reads/éch à la demande de l'équipe.
- **(A)** 90% des comptages correspondent à des gènes codant pour **des protéines** (puis ARNr mitochondriaux + long ARN non codants). Détection de **16 000 gènes**.
- **(B)** L'analyse en composante principale (ACP) montre une opposition des sexes (● ■) sur la composante 1. La composante 2 oppose les régimes et génotypes (● ● ● ●).



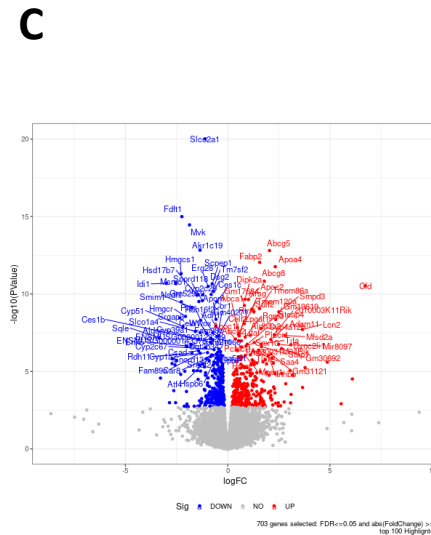
➤ Liver_Mus_musculus



- 61 échantillons : 8 conditions (n=6 à 8), 20 M de reads/éch à la demande de l'équipe.
- **(A)** 90% des comptages correspondent à des gènes codant pour **des protéines** (puis ARNr mitochondriaux + long ARN non codants). Détection de **16 000 gènes**.
- **(B)** L'analyse en composante principale (ACP) montre une opposition des sexes (● ■) sur la composante 1. La composante 2 oppose les régimes et génotypes (● ● ● ● ●).



- **(C)** Volcano-plot illustrant les gènes les plus différentiellement exprimés pour une comparaison (Up en rouge et Down en bleu).

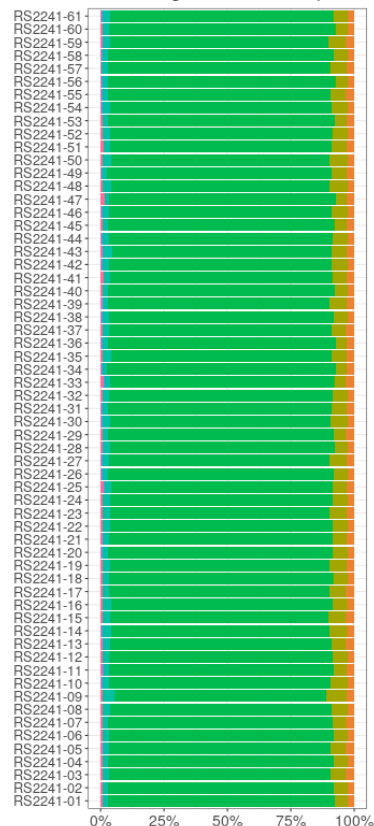


➤ Liver_Mus_musculus



- 61 échantillons : 8 conditions (n=6 à 8), 20 M de reads/éch à la demande de l'équipe.
- (A) 90% des comptages correspondent à des gènes codant pour **des protéines** (puis ARNr mitochondriaux + long ARN non codants). Détection de **16 000 gènes**.
- (B) **L'analyse en composante principale (ACP)** montre une opposition des sexes (● ■) sur la composante 1. La composante 2 oppose les régimes et génotypes (● ● ● ● ●).

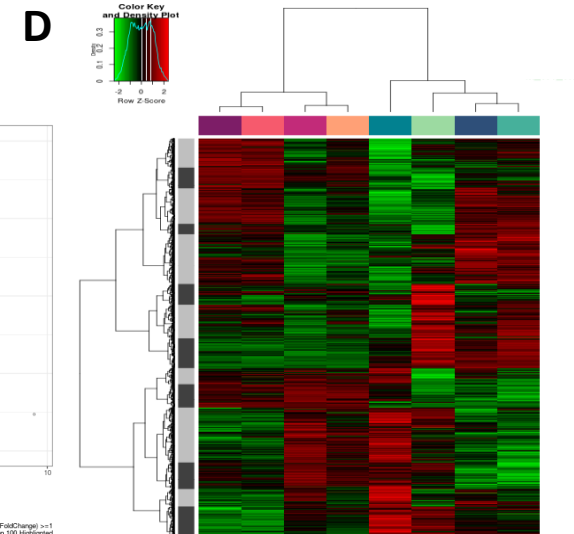
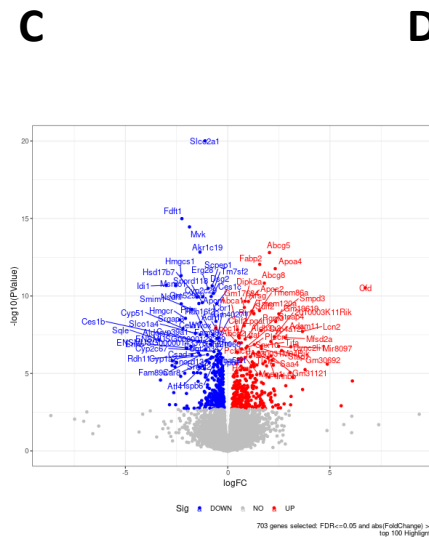
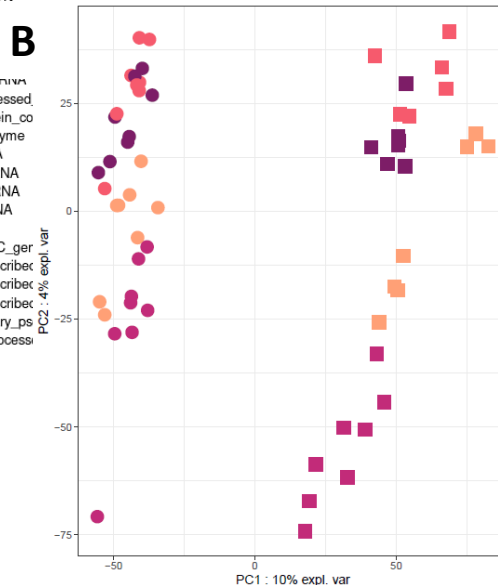
A Percentage Total counts per biotype



Gene Biotype

- IG_C_gene
- lncF...
- miF
- mis
- Mt
- Mt_rRNA
- processed
- protein_co
- ribozyme
- rRNA
- scaRNA
- snoRNA
- snRNA
- TEC
- TR_C_gene
- transcribed
- transcribed
- transcribed
- unitary_ps
- unprocess

- (C) **Volcano-plot** illustrant les gènes les plus différentiellement exprimés pour une comparaison (Up en rouge et Down en bleu).
- (D) **Heatmap** : Sélection des 1000 gènes les plus significativement régulés pour chaque comparaison. Choix de découper en 14 sous-groupes de gènes (clusters).



> Infos pratiques

3' mRNA-seq

RNA-seq « classique »

Quantifier l'expression des gènes codants
Analyser l'expression différentielle des gènes (DGE)
Rendu identique des résultats

- **ARN**

ARN : 2 tubes 5 µl à 135 ng/µL

Qualité des ARN : bonne à faible (homogène)

(adaptation du protocole)

Possibilité ARN de sang total *(utilisation bloqueur globine)*

- **ARN**

ARN : 2 tubes (5 µl et 30 µl) à 100 ng/µl

Qualité des ARN : bonne (homogène)

> Infos pratiques

3' mRNA-seq

RNA-seq « classique »

Quantifier l'expression des gènes codants
Analyser l'expression différentielle des gènes (DGE)
Rendu identique des résultats

• ARN

ARN : 2 tubes 5 µl à 135 ng/µL
Qualité des ARN : bonne à faible (homogène)
(adaptation du protocole)
Possibilité ARN de sang total (utilisation bloqueur globine)



Coût dépend du nombre d'échantillons et de la profondeur de séquençage (taille des flowcells du séquenceur)



• ARN

ARN : 2 tubes (5 µl et 30 µl) à 100 ng/µl
Qualité des ARN : bonne (homogène)

Estimation QC ARN + librairies + analyses

Pour 96 échantillons : 75 euros/éch
Pour 96 échantillons (sang) : 90 euros/éch

Pour 96 échantillons : 90 euros/éch

Estimation séquençage

Pour 48 échantillons : ~ 8 M lectures 35 euros/éch
Pour 96 échantillons : ~ 12 M lectures 30 euros/éch

Pour 48 échantillons : ~ 33 M lectures 90 euros/éch
Pour 96 échantillons : ~ 30 M lectures 70 euros/éch

Stockage des données

Pour 1 échantillon : ~ 0,5 giga octets

Pour 1 échantillon : ~ 2/3 giga octets

Merci pour votre attention

Personnes impliquées dans les travaux présentés



Equipe REGLISS
Laurence Liaubet
Laure Gress
Maëlle Gusmini



**Plateau Génome
Transcriptome de
Bordeaux**
Erwan Guichoux
Zoé Compagnie



**Plateau Génomique et
Transcriptomique,
UMR Inserm**
Carine Valle
Emeline Sarot



BioToMyc, Pinton Philippe



Plateforme GeT-TRiX
Claire Naylies
Kory Jason
Yannick Lippi



Financement