

Mutations somatiques et trajectoires cancéreuses dans la muqueuse orale saine

Zinara LIDAMAHASOLO - M2 MISO

Encadrant : Pierre MARTINEZ

Tuteur : Mikaël SALSON

Centre de Recherche en Cancérologie de Lyon (CRCL)

Equipe « Analyse intégrée de la dynamique du cancer » dirigée par Dr Pierre SAINTIGNY

I. INTRODUCTION

A. Contexte

- **Mutations somatiques :**
 - Au cours du développement & maintenance.

I. INTRODUCTION

A. Contexte

- **Mutations somatiques :**
 - Au cours du développement & maintenance.
 - Potentiel tumorigène: **drivers du cancer**.

I. INTRODUCTION

A. Contexte

- **Mutations somatiques :**
 - Au cours du développement & maintenance.
 - Potentiel tumorigène: **drivers du cancer**.
- **Drivers et tissu sain :**
 - Mutations drivers omniprésentes.

I. INTRODUCTION

A. Contexte

- **Mutations somatiques :**
 - Au cours du développement & maintenance.
 - Potentiel tumorigène: **drivers du cancer**.
- **Drivers et tissu sain :**
 - Mutations drivers omniprésentes.
 - Sélection positive.
 - $dN/dS > 1$.

I. INTRODUCTION

A. Contexte

- **Mutations somatiques :**
 - Au cours du développement & maintenance.
 - Potentiel tumorigène: **drivers du cancer**.
- **Drivers et tissu sain :**
 - Mutations drivers omniprésentes.
 - Sélection positive.
 - $dN/dS > 1$.
 - Les plus incriminés: NOTCH1, TP53, FAT1, FAT3.

I. INTRODUCTION

A. Contexte

- **Mutations somatiques :**
 - Au cours du développement & maintenance.
 - Potentiel tumorigène: **drivers du cancer**.
- **Drivers et tissu sain :**
 - Mutations drivers omniprésentes.
 - Sélection positive.
 - $dN/dS > 1$.
 - Les plus incriminés: NOTCH1, TP53, FAT1, FAT3.
- **Muqueuse orale saine :**
 - Site le plus commun du principal cancer de la sphère tête-et-cou (carcinome épidermoïde).

I. INTRODUCTION

A. Contexte

- **Mutations somatiques :**
 - Au cours du développement & maintenance.
 - Potentiel tumorigène: **drivers du cancer**.
- **Drivers et tissu sain :**
 - Mutations drivers omniprésentes.
 - Sélection positive.
 - $dN/dS > 1$.
 - Les plus incriminés: NOTCH1, TP53, FAT1, FAT3.
- **Muqueuse orale saine :**
 - Site le plus commun du principal cancer de la sphère tête-et-cou (carcinome épidermoïde).
 - Évolution somatique mal caractérisée.

I. INTRODUCTION

B. Objectifs

Projet global

- **Caractériser** la dynamique évolutive somatique de la muqueuse orale saine.

I. INTRODUCTION

B. Objectifs

Projet global

- **Caractériser** la dynamique évolutive somatique de la muqueuse orale saine.
+++ Cellules épithéliales +++

I. INTRODUCTION

B. Objectifs

Projet global

- **Caractériser** la dynamique évolutive somatique de la muqueuse orale saine.
+++ Cellules épithéliales +++

Stage

- **Mettre en place** les solutions pour analyser les échantillons.

I. INTRODUCTION

B. Objectifs

Projet global

- **Caractériser** la dynamique évolutive somatique de la muqueuse orale saine.
+++ Cellules épithéliales +++

Stage

- **Mettre en place** les solutions pour analyser les échantillons.
- **Estimer la pertinence** des méthodes utilisées (prélèvement, analyse des données).

II. MATERIELS ET METHODES

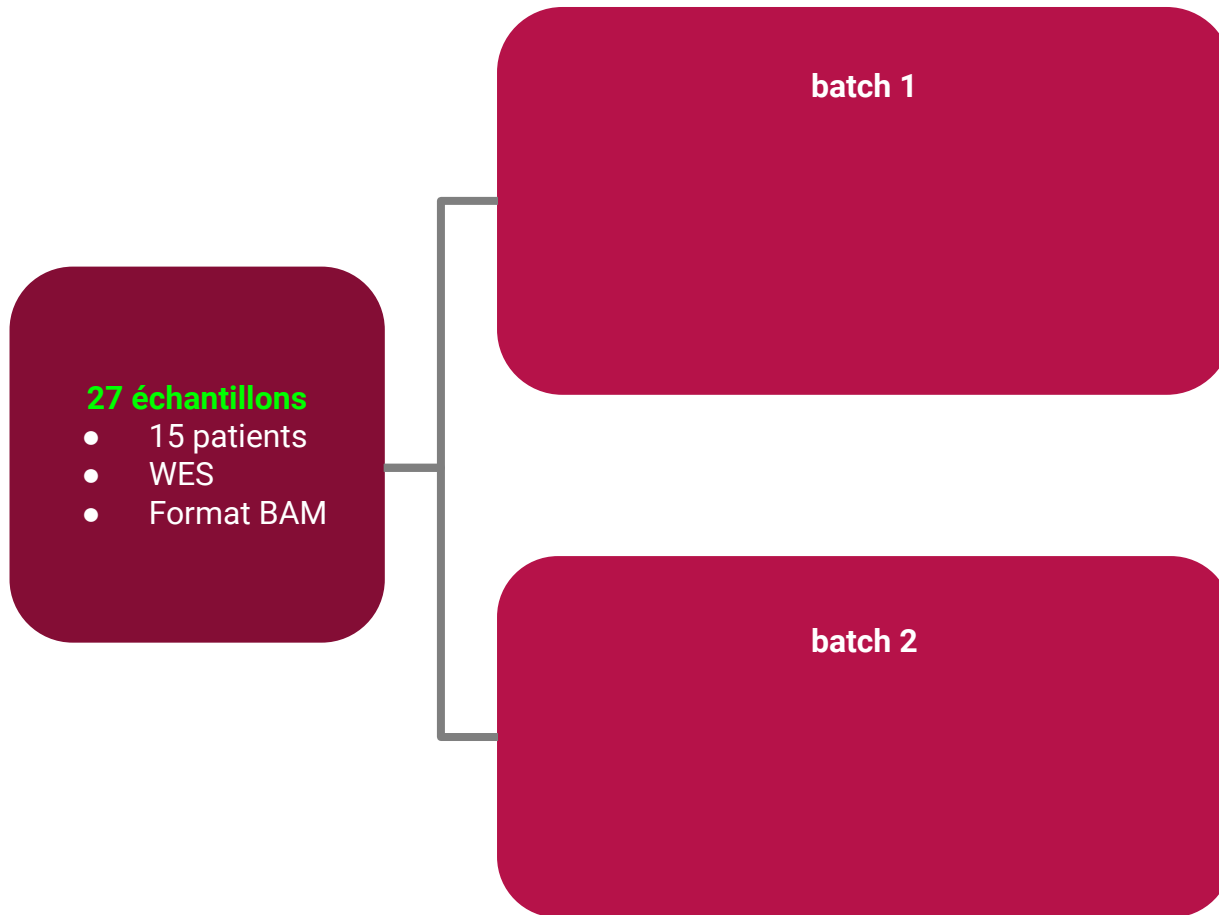
A. Nos jeux de données

27 échantillons

- 15 patients
- WES
- Format BAM

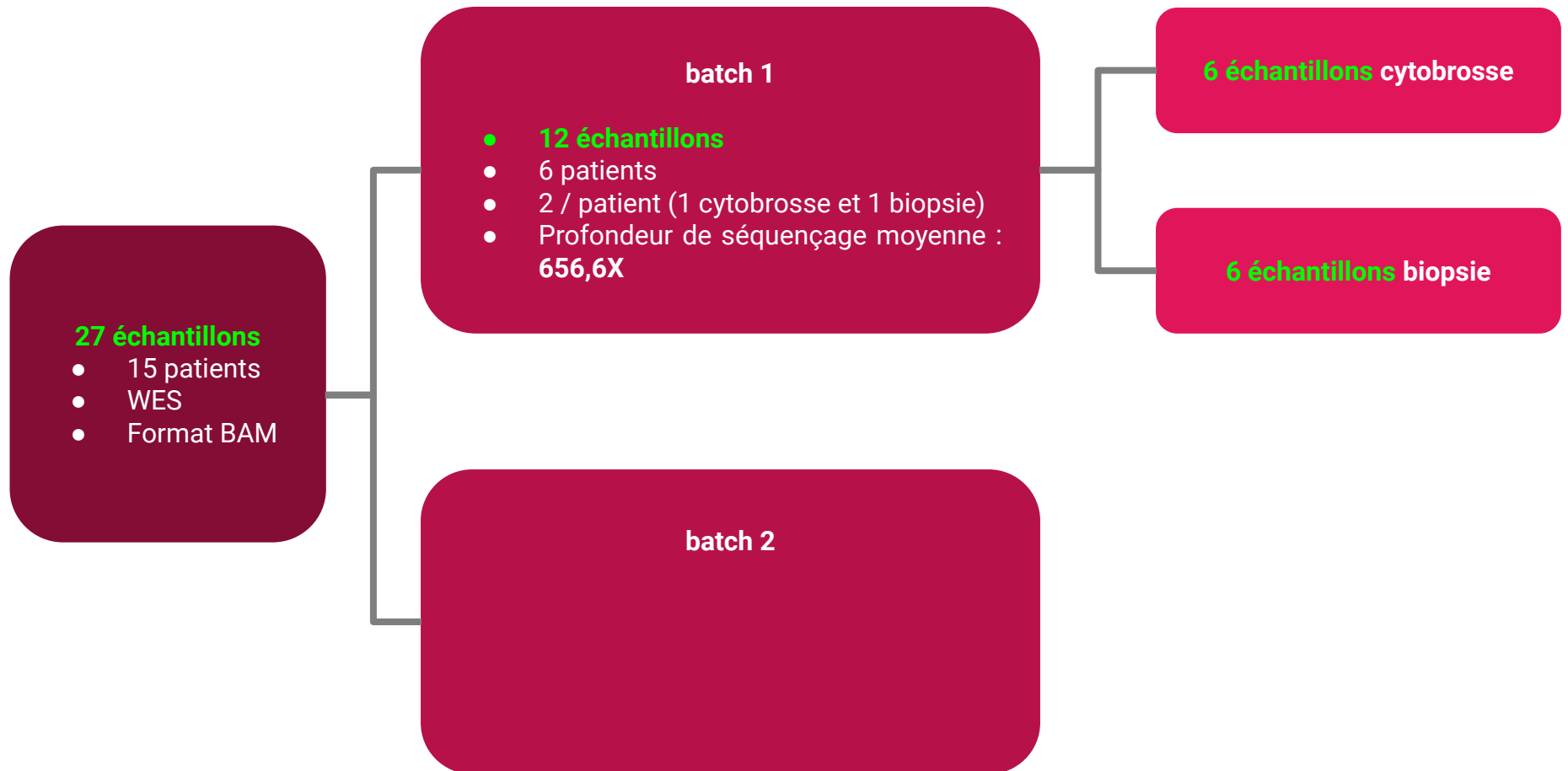
II. MATERIELS ET METHODES

A. Nos jeux de données



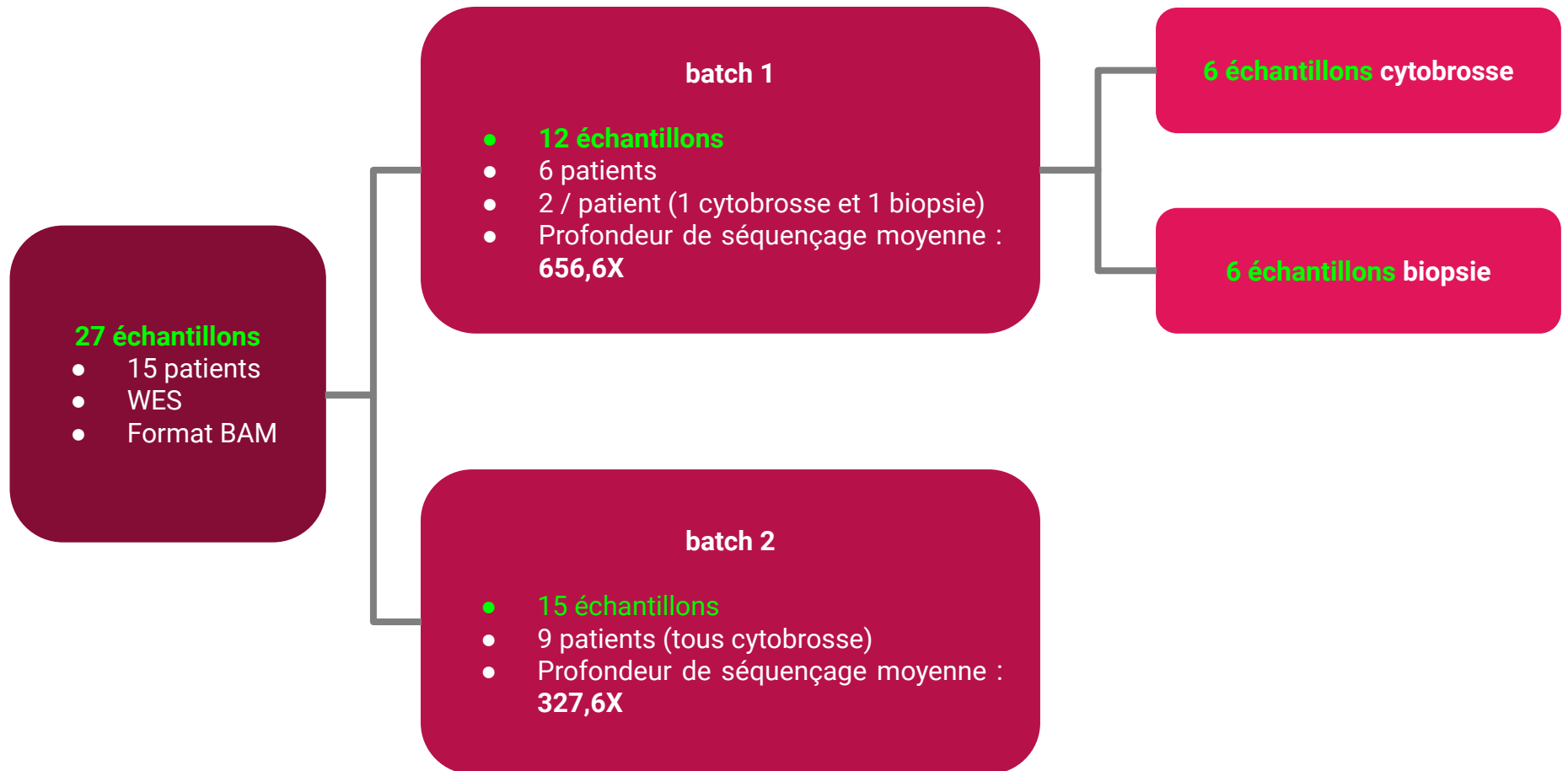
II. MATERIELS ET METHODES

A. Nos jeux de données



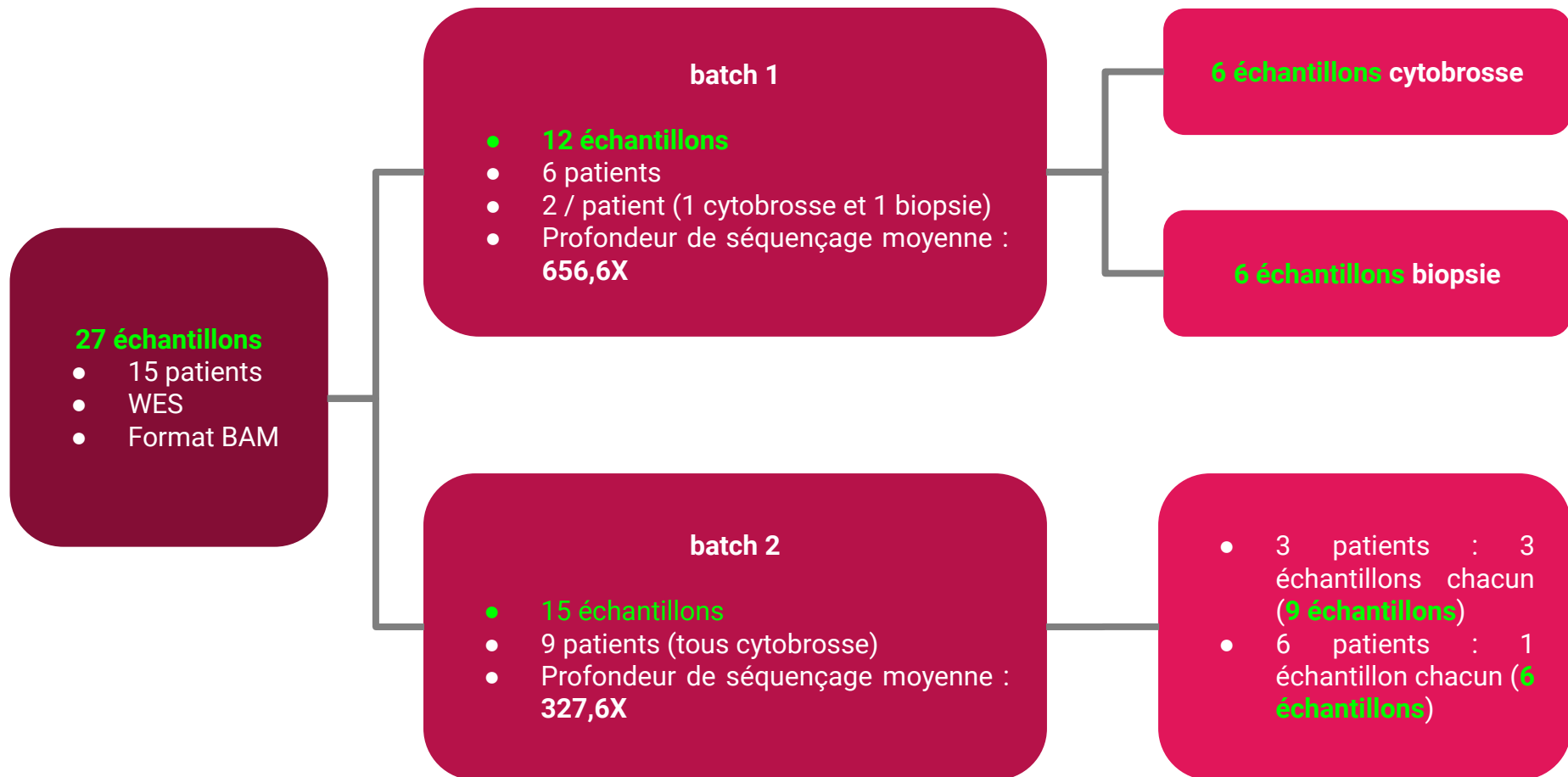
II. MATERIELS ET METHODES

A. Nos jeux de données



II. MATERIELS ET METHODES

A. Nos jeux de données



II. MATERIELS ET METHODES

B. Etapes

1

Appel de
variant

2

Filtre

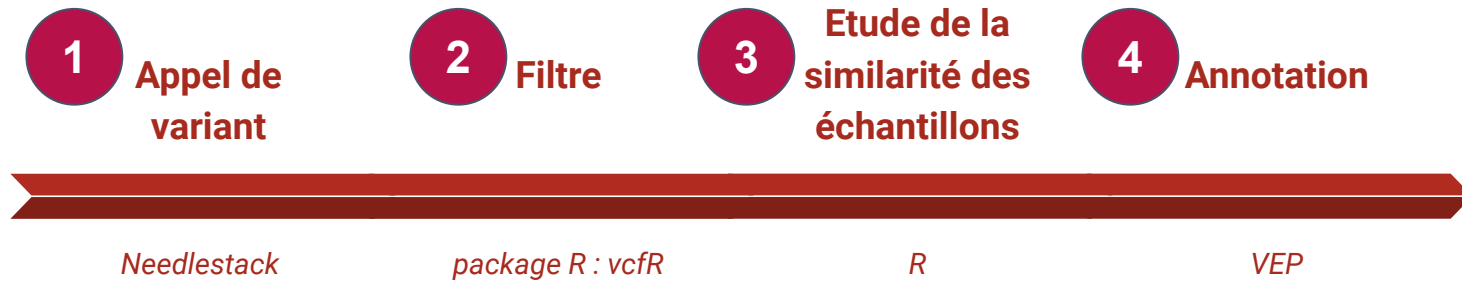


Needlestack

package R : vcfR

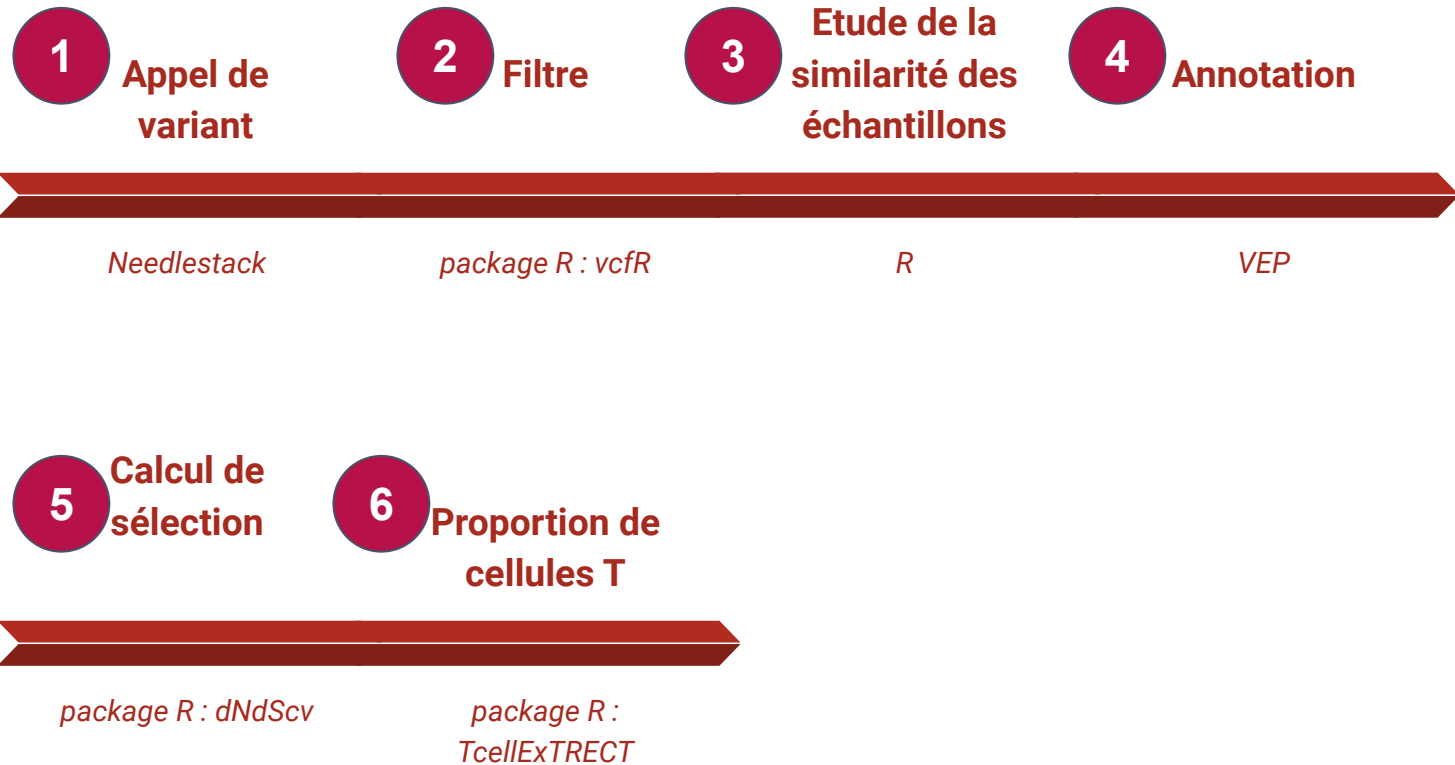
II. MATERIELS ET METHODES

B. Etapes



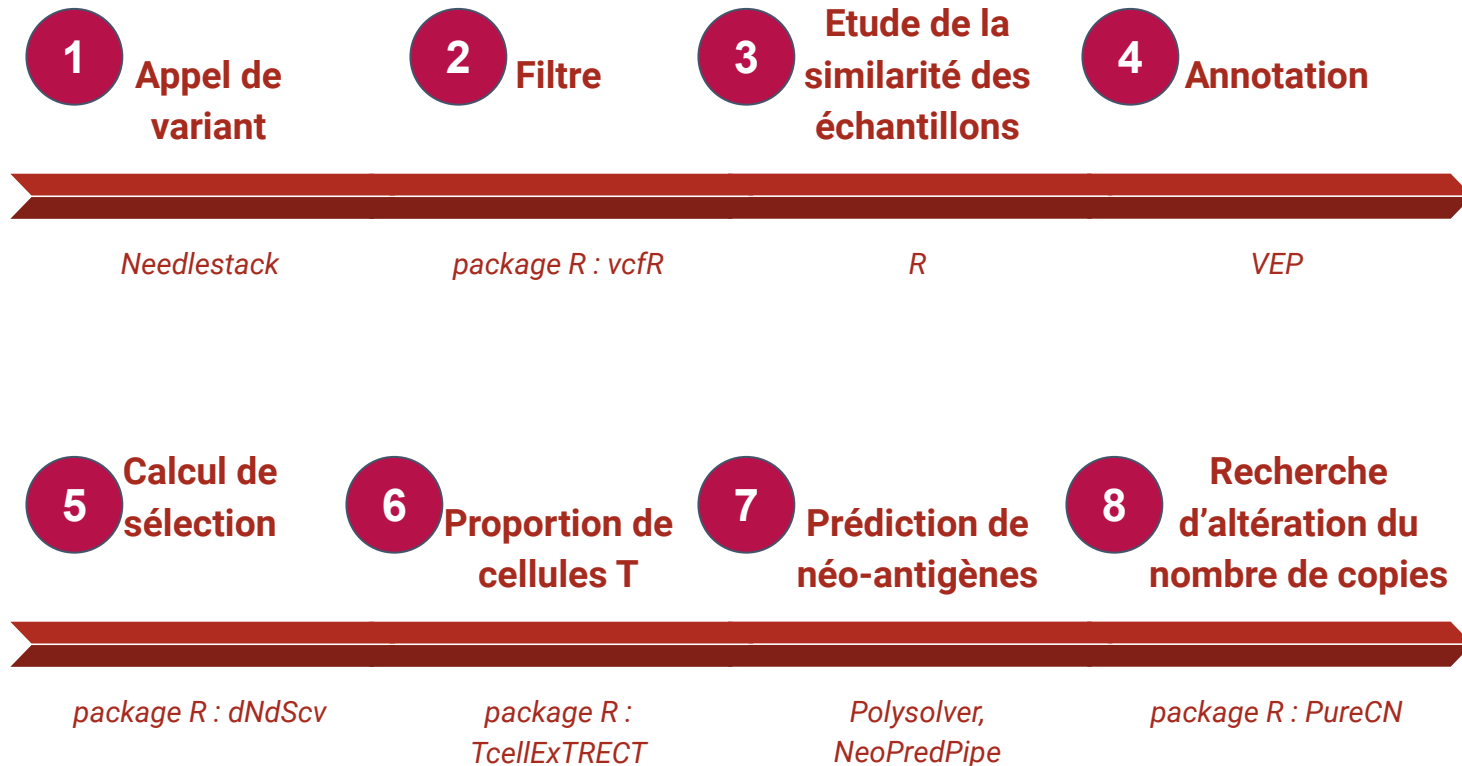
II. MATERIELS ET METHODES

B. Etapes



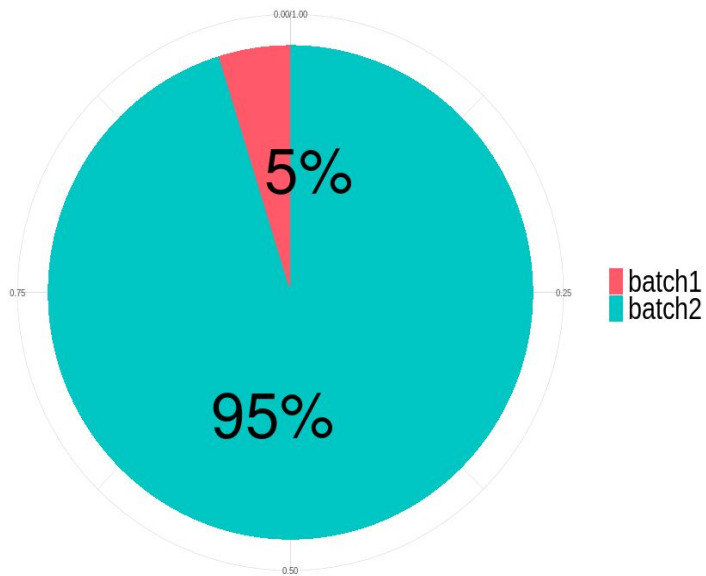
II. MATERIELS ET METHODES

B. Etapes



III. RESULTATS

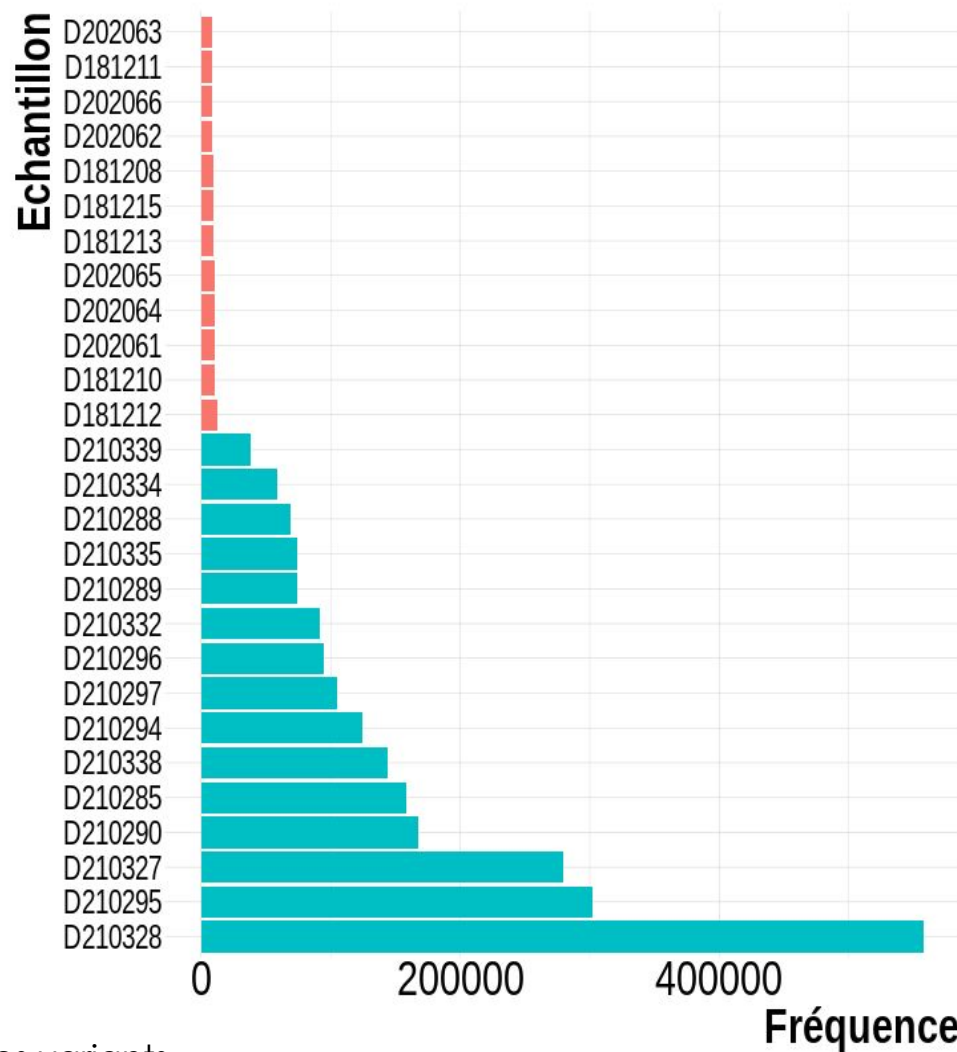
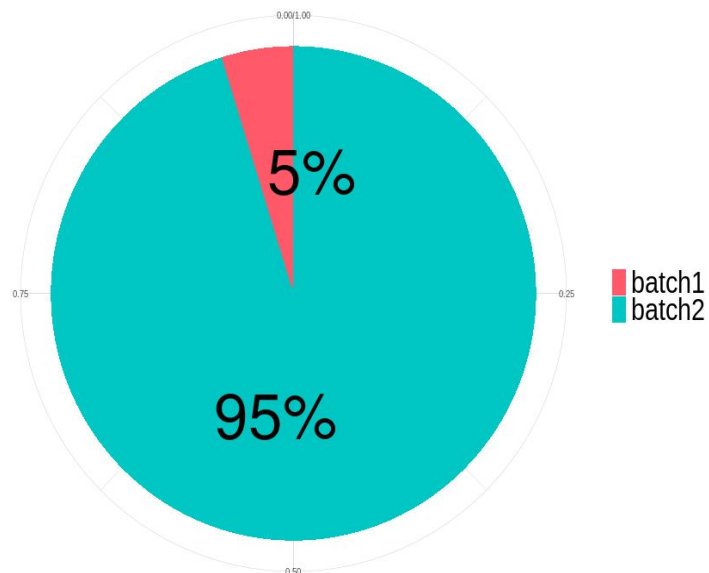
A. Variants avant filtre



Répartition des variants

III. RESULTATS

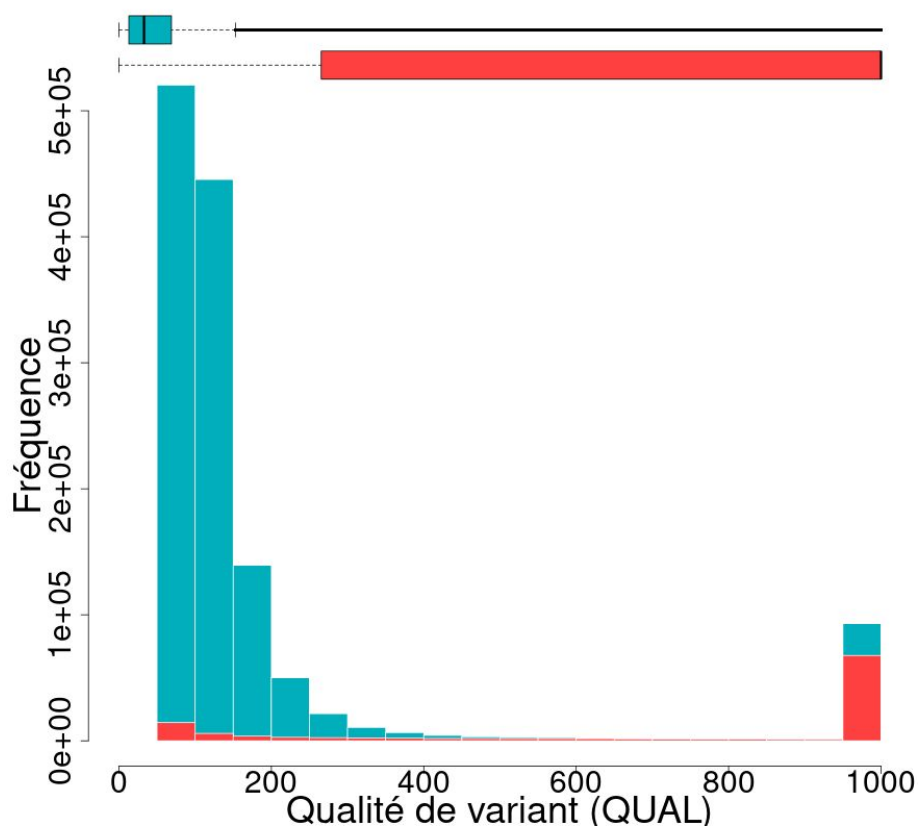
A. Variants avant filtre



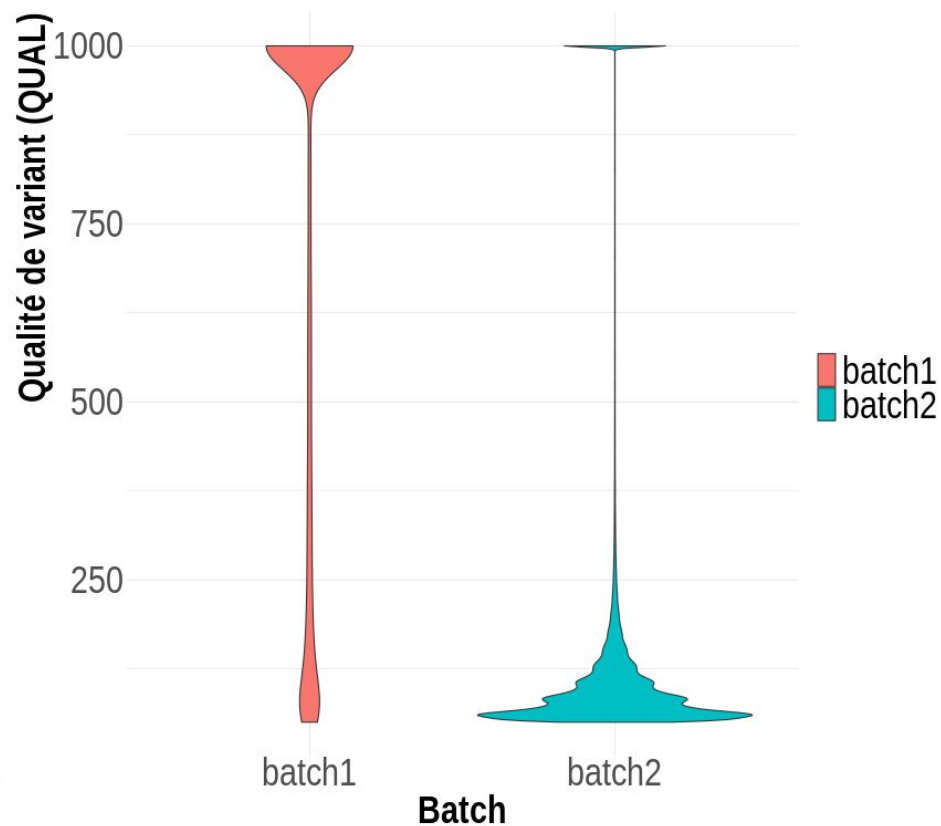
Répartition des variants

III. RESULTATS

A. Variants avant filtre



Qualité avant filtre



III. RESULTATS

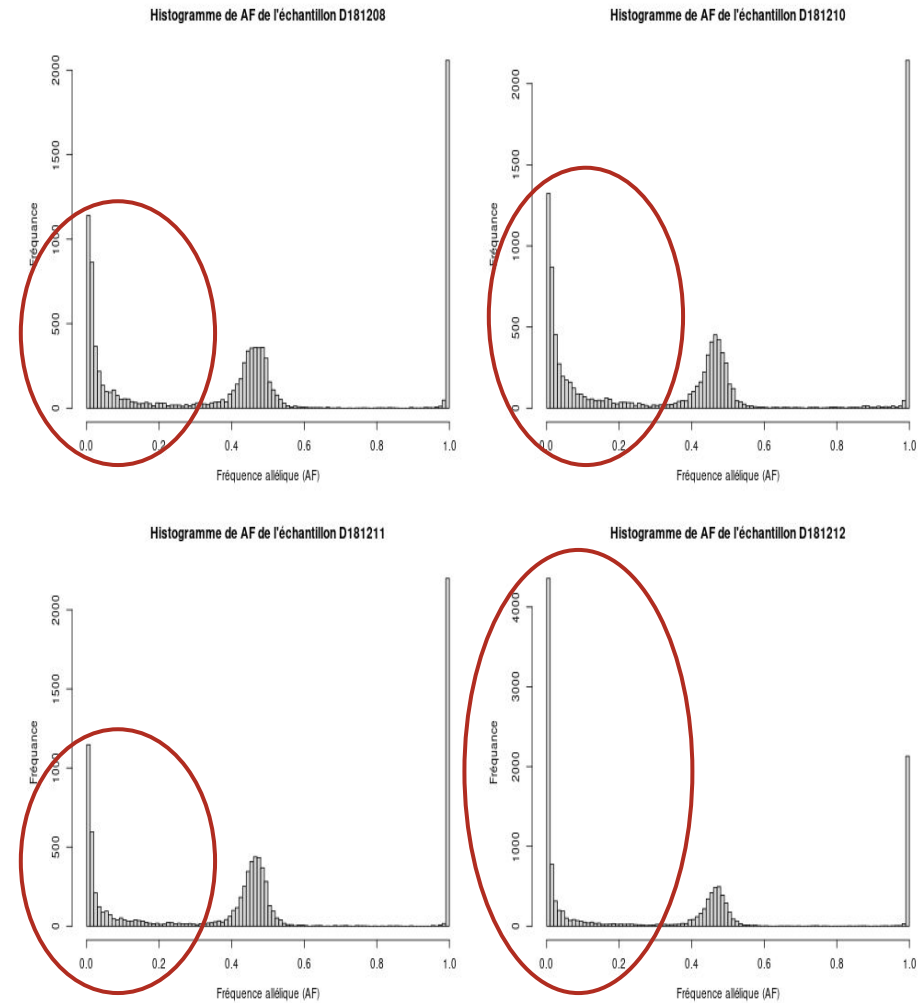
B. Filtre

- Fréquence allélique (AF) < **0.3**

III. RESULTATS

B. Filtre

- Fréquence allélique (AF) < 0.3

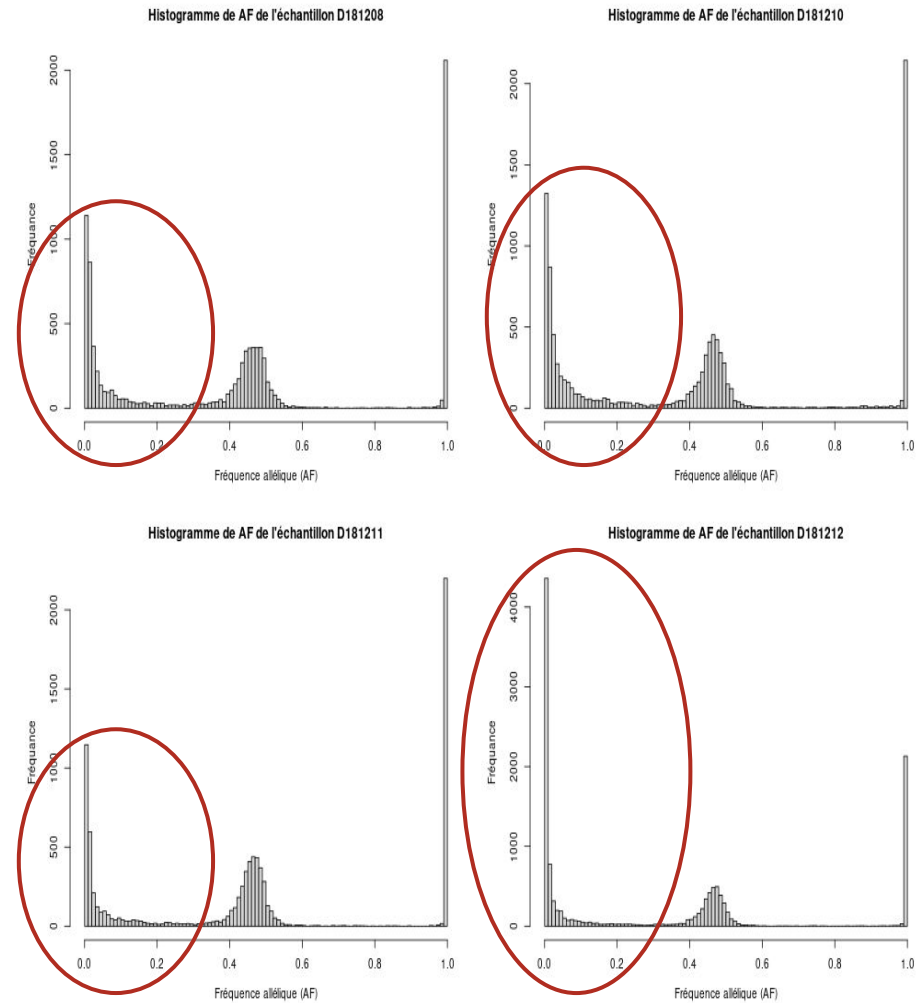


La fréquence allélique.

III. RESULTATS

B. Filtre

- Fréquence allélique (AF) < **0.3**
- Profondeur de lecture (DP) > **50**
- Nombre de reads portant le variant (AO) > **3**

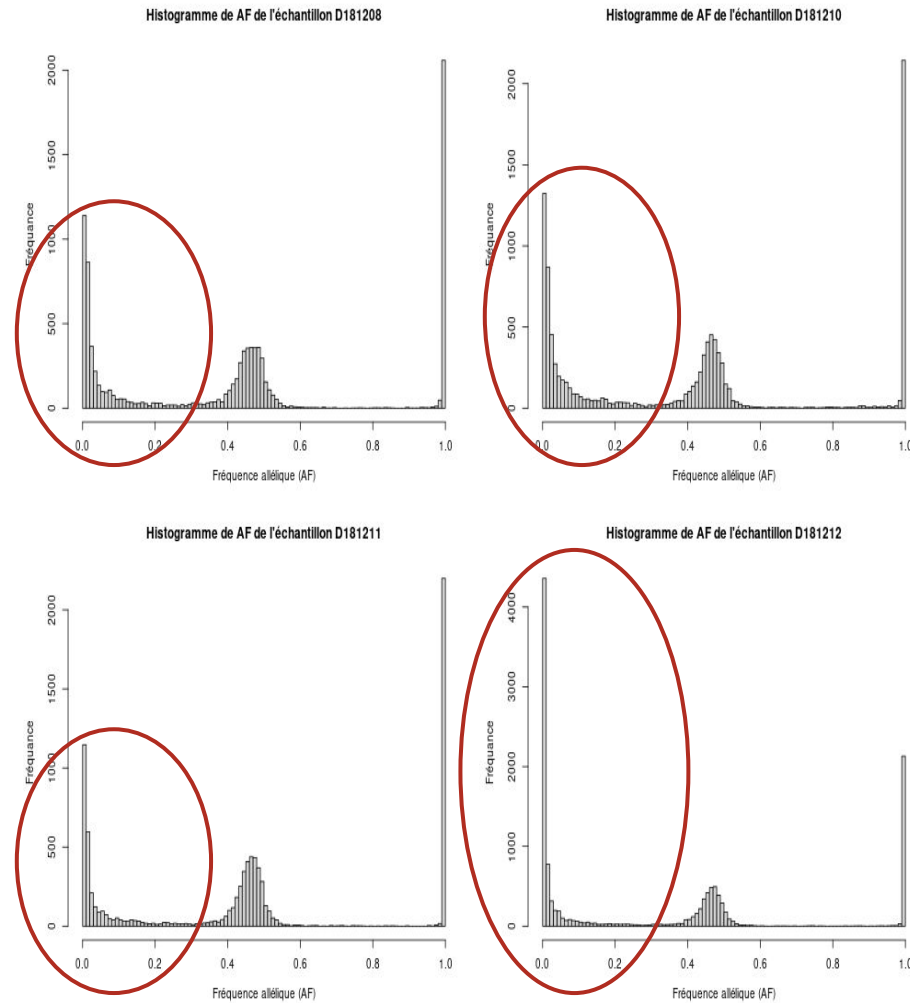


La fréquence allélique.

III. RESULTATS

B. Filtre

- Fréquence allélique (AF) < **0.3**
- Profondeur de lecture (DP) > **50**
- Nombre de reads portant le variant (AO) > **3**
- Qualité du variant (QUAL) :
 - batch1 : > **100**
 - batch2 : > **400**

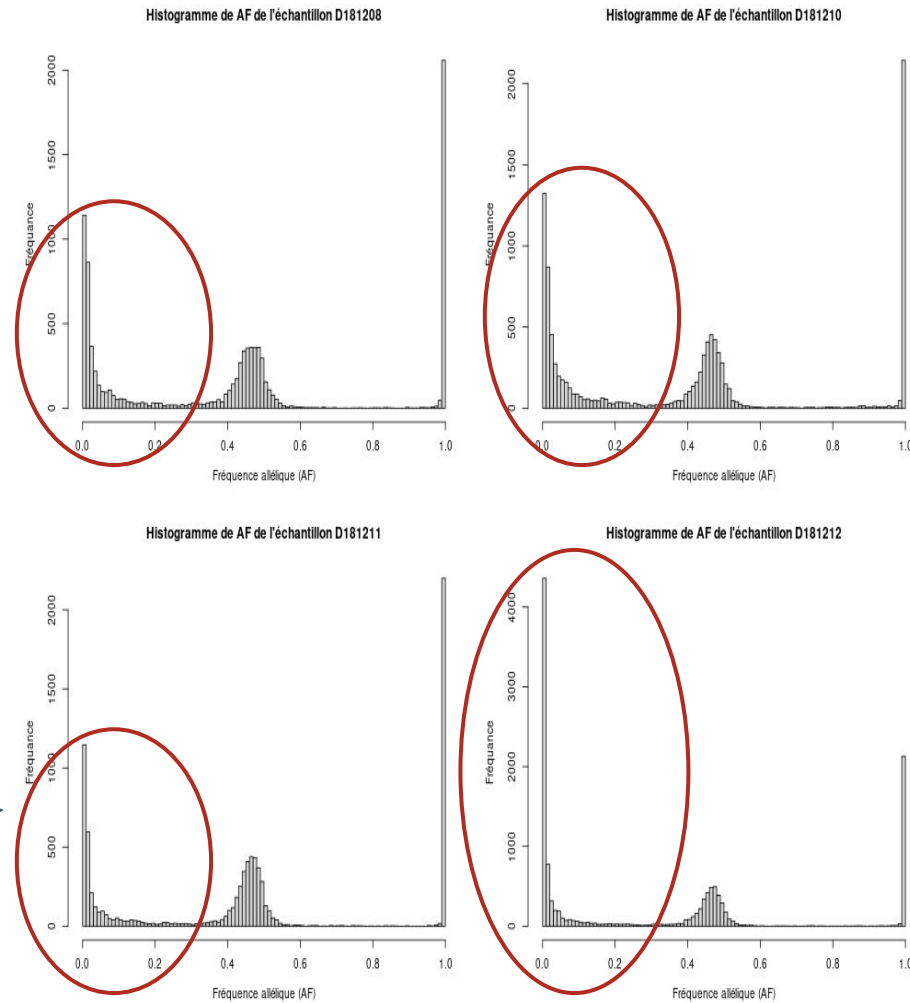
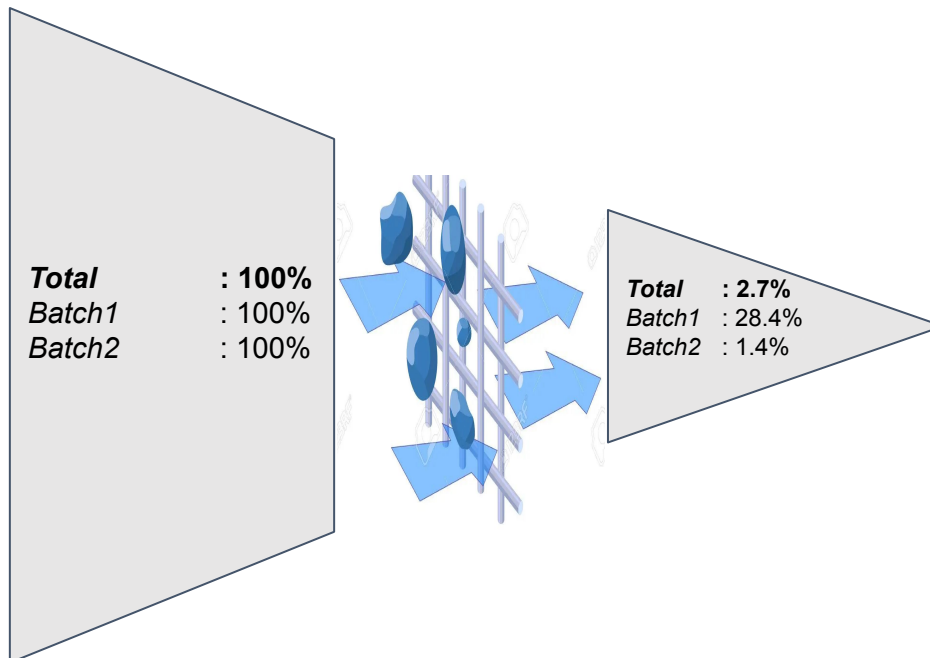


La fréquence allélique.

III. RESULTATS

B. Filtre

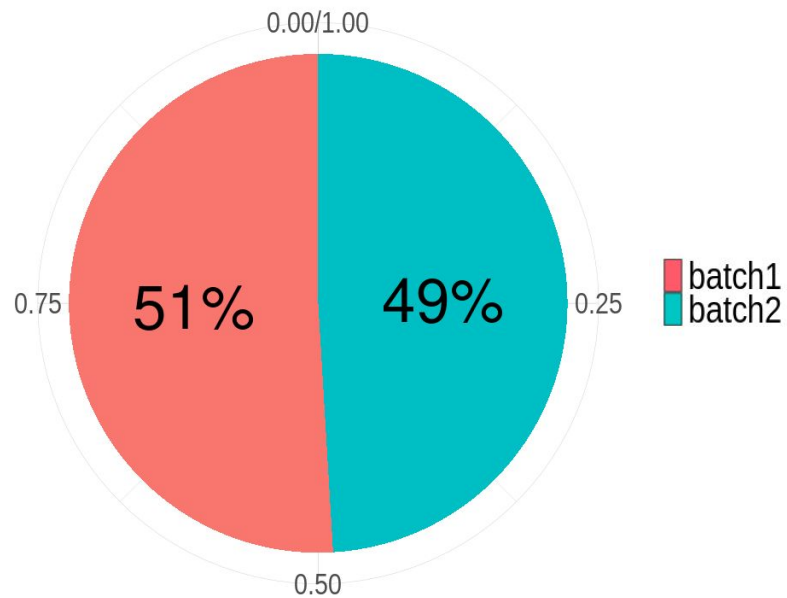
- Fréquence allélique (AF) < **0.3**
- Profondeur de lecture (DP) > **50**
- Nombre de reads portant le variant (AO) > **3**
- Qualité du variant (QUAL) :
 - batch1 : > **100**
 - batch2 : > **400**



La fréquence allélique.

III. RESULTATS

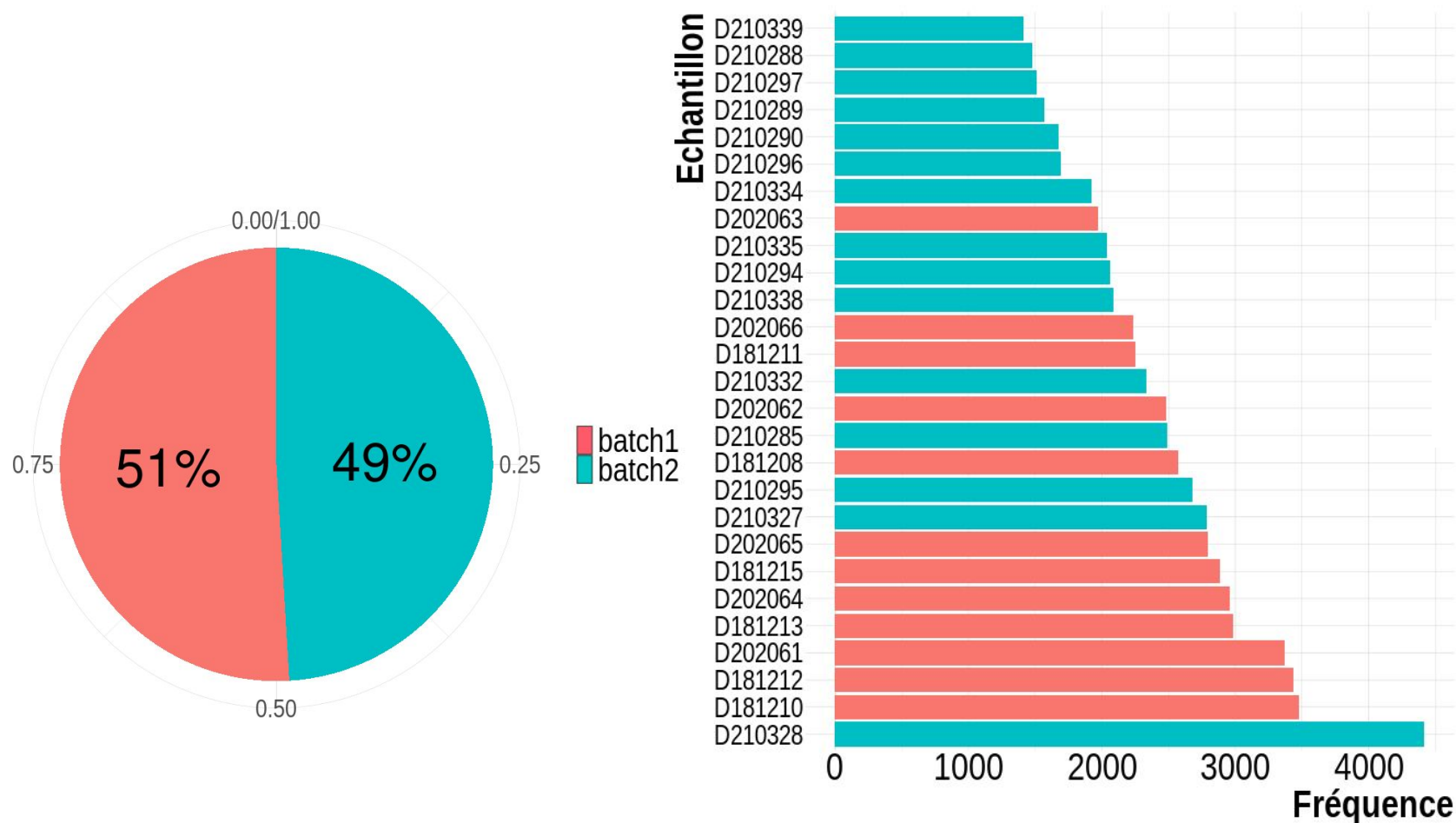
C. Variants après filtre



Répartition des variants après filtre

III. RESULTATS

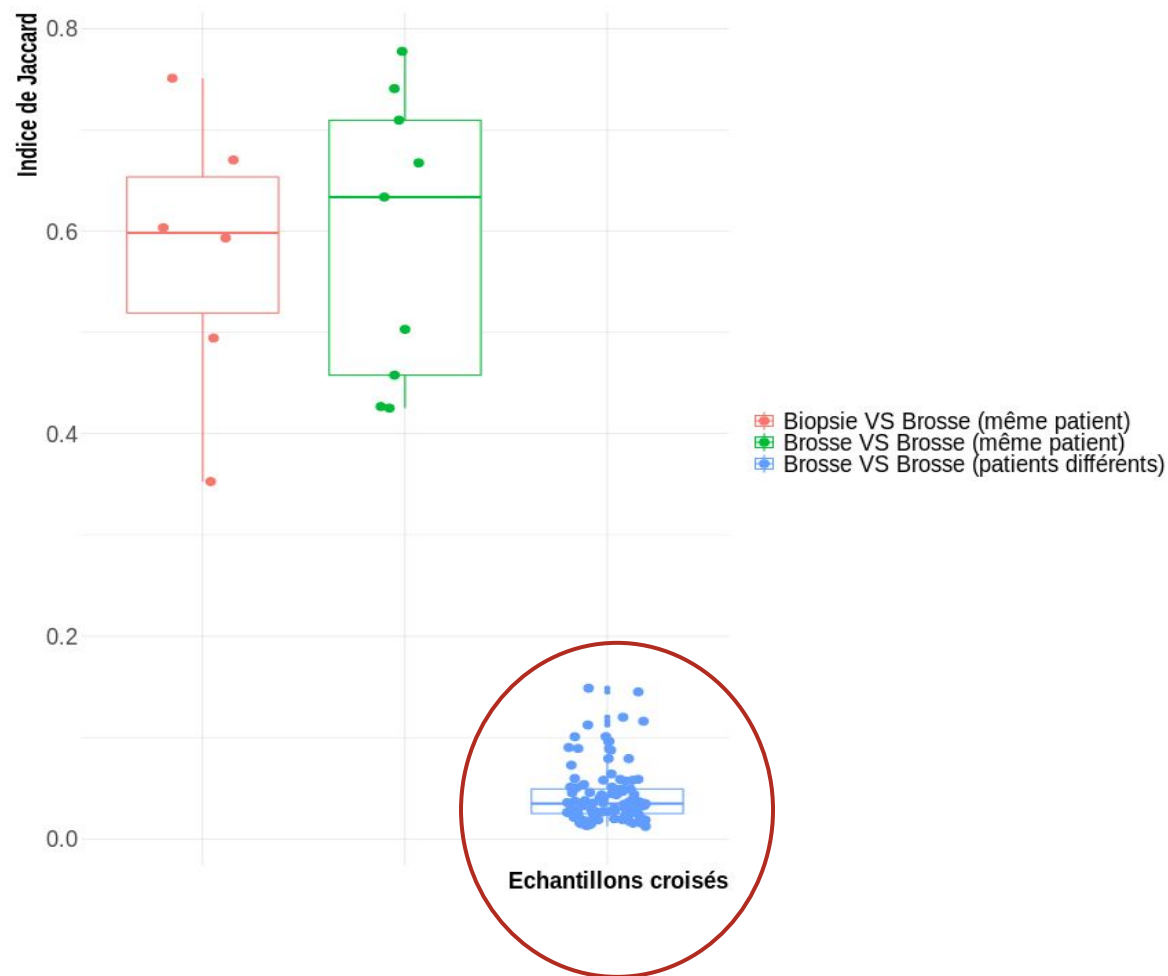
C. Variants après filtre



Répartition des variants après filtre

III. RESULTATS

D. Similarité des échantillons



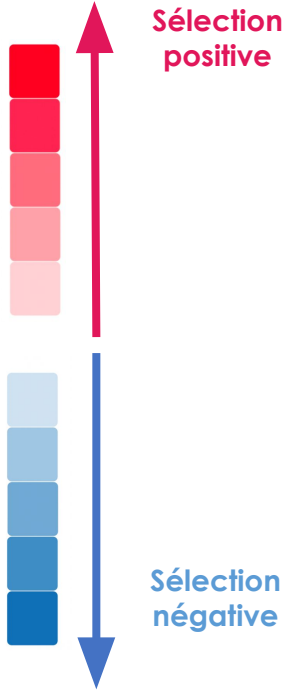
Répartition des indices de Similarité des échantillons

III. RESULTATS

E. Avantage sélectif des 62 gènes drivers du Carcinome Épidermoïde de la Tête et du Cou

- Significatifs :

Gène	nombre_de_mutations	dNdS_faux_sens	dNdS_non_sens	p_globale_ajustée
CREBBP	10	0.0000000	0.0000000	0.0006718274
CASP8	9	1.2133029	40.8414933	0.0006718274
FAT1	5	0.1296040	2.4311507	0.0007435047
NOTCH1	19	0.8875176	8.4462262	0.0032685884
KDM5C	8	0.0000000	0.0000000	0.0064674674
FAT3	11	0.2627164	0.0000000	0.0087696291
HLA-B	25	11.1075040	0.0000000	0.0217561387
HLA-A	45	1.2874269	0.0000000	0.0427825393
MYH9	15	0.1241240	0.4862293	0.0439383395



III. RESULTATS

E. Avantage sélectif des 62 gènes drivers du Carcinome Épidermoïde de la Tête et du Cou

- Significatifs :

Gène	nombre_de_mutations	dNdS_faux_sens	dNdS_non_sens	p_globale_ajustée
CREBBP	10	0.0000000	0.0000000	0.0006718274
CASP8	9	1.2133029	40.8414933	0.0006718274
FAT1	5	0.1296040	2.4311507	0.0007435047
NOTCH1	19	0.8875176	8.4462262	0.0032685884
KDM5C	8	0.0000000	0.0000000	0.0064674674
FAT3	11	0.2627164	0.0000000	0.0087696291
HLA-B	25	11.1075040	0.0000000	0.0217561387
HLA-A	45	1.2874269	0.0000000	0.0427825393
MYH9	15	0.1241240	0.4862293	0.0439383395



Sélection positive

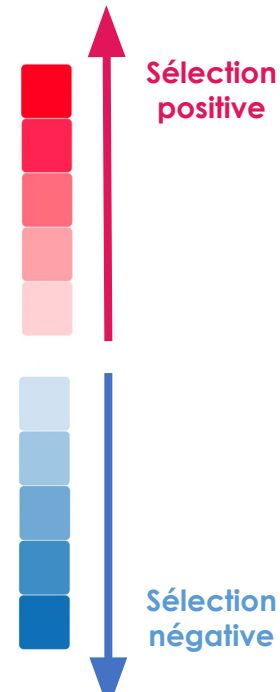
Sélection négative

III. RESULTATS

E. Avantage sélectif des 62 gènes drivers du Carcinome Épidermoïde de la Tête et du Cou

- Non-significatifs :

Gène	nombre_de_mutation	dNdS_faux_sens	dNdS_non_sens	p_valeur_ajustée	
KMT2C	12	0.6372499	0.00000	0.1466849	
TP53	6	2.7857984	18.93655	0.2102199	
BRCA2	5	0.4000035	0.00000	0.3675089	
PIK3CA	6	1.8106351	0.00000	0.8660165	
DDX3X	5	2.0057472	0.00000	0.8660165	
NOTCH2	14	0.7701560	0.00000	0.9401093	

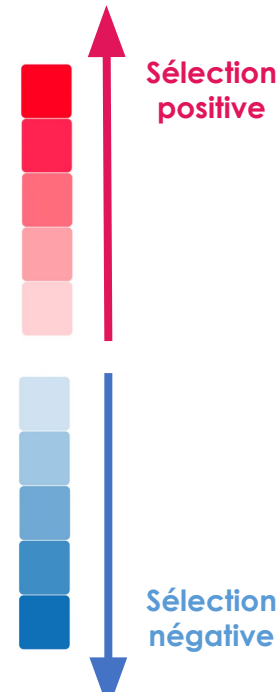


III. RESULTATS

E. Avantage sélectif des 62 gènes drivers du Carcinome Épidermoïde de la Tête et du Cou

- Non-significatifs :

Gène	nombre_de_mutation	dNdS_faux_sens	dNdS_non_sens	p_valeur_ajustée	
KMT2C	12	0.6372499	0.00000	0.1466849	
TP53	6	2.7857984	18.93655	0.2102199	
BRCA2	5	0.4000035	0.00000	0.3675089	
PIK3CA	6	1.8106351	0.00000	0.8660165	
DDX3X	5	2.0057472	0.00000	0.8660165	
NOTCH2	14	0.7701560	0.00000	0.9401093	



Sélection positive

Sélection négative

III. RESULTATS

F. Avantage sélectif des autres gènes non drivers

- Gènes (non-)exprimés dans les cellules épithéliales orales :

III. RESULTATS

F. Avantage sélectif des autres gènes non drivers

- Gènes (non-)exprimés dans les cellules épithéliales orales :
 - Human Oral Mucosa Cell Atlas (*Williams et al., Cell 2021*).

III. RESULTATS

F. Avantage sélectif des autres gènes non drivers

- Gènes (non-)exprimés dans les cellules épithéliales orales :
 - Human Oral Mucosa Cell Atlas (*Williams et al., Cell 2021*).
- 500 les plus mutés parmi les exprimés :
 - 1,4% sous sélection positive.

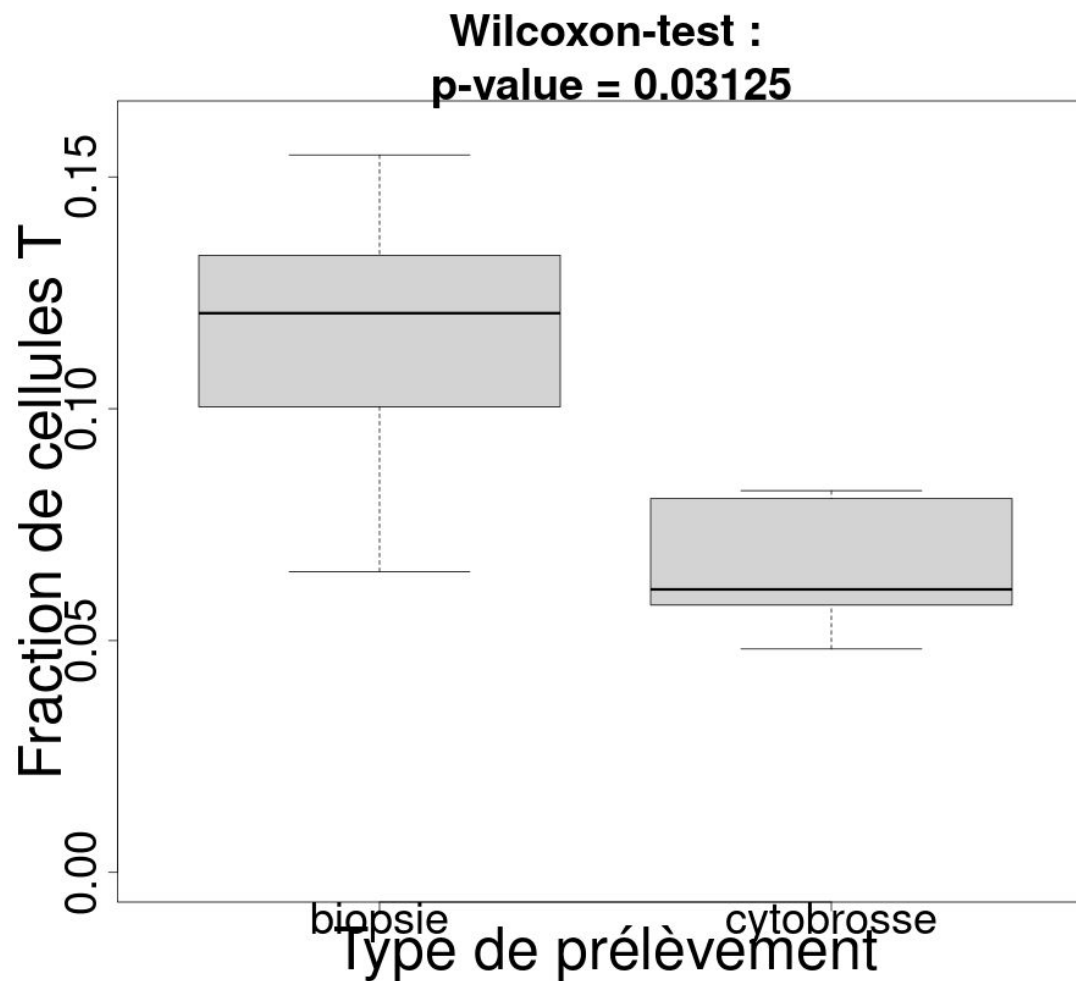
III. RESULTATS

F. Avantage sélectif des autres gènes non drivers

- Gènes (non-)exprimés dans les cellules épithéliales orales :
 - Human Oral Mucosa Cell Atlas (*Williams et al., Cell 2021*).
- 500 les plus mutés parmi les exprimés :
 - 1,4% sous sélection positive.
- 1000 les plus mutés parmi les non-exprimés :
 - 10% sous sélection positive.

III. RESULTATS

G. Fraction de cellules T



Distribution des fractions de cellules T entre biopsie et
cytobrosse de batch1

III. RESULTATS

H. Néo-antigènes & Altération du nombre de copies

- Par échantillon :

30,93% des **mutations** (10% dans gènes exprimés).

III. RESULTATS

H. Néo-antigènes & Altération du nombre de copies

- Par échantillon :

30,93% des **mutations** (10% dans gènes exprimés).



Néo-antigènes

III. RESULTATS

H. Néo-antigènes & Altération du nombre de copies

- Par échantillon :

30,93% des **mutations** (10% dans gènes exprimés).



Néo-antigènes

(2,31% Nouveaux néo-antigènes « **ligand fort** » par gènes exprimés).

III. RESULTATS

H. Néo-antigènes & Altération du nombre de copies

- Par échantillon :

30,93% des **mutations** (10% dans gènes exprimés).



Néo-antigènes

(2,31% Nouveaux néo-antigènes « **ligand fort** » par gènes exprimés).

- Absence d'altération du nombre de copies.

IV. DISCUSSION

- Différences entre batches :
 - **Nombre de variants** (+++ dans batch2)
 - **Qualité de variants** (+++ dans batch1)



- **Profondeur de séquençage** moyenne :
 - batch1 : 656,6X
 - batch2 : 327,6X

IV. DISCUSSION

- Différences entre batches :
 - **Nombre de variants** (+++ dans batch2)
 - **Qualité de variants** (+++ dans batch1)



- **Profondeur de séquençage** moyenne :
 - batch1 : 656,6X
 - batch2 : 327,6X



	Précision de la prédiction	Nombre	Vrais-positifs	Qualité
Batch1	+	-	+	+

IV. DISCUSSION

- Différences entre batches :
 - **Nombre de variants** (+++ dans batch2)
 - **Qualité de variants** (+++ dans batch1)



- **Profondeur de séquençage** moyenne :
 - batch1 : 656,6X
 - batch2 : 327,6X



	Précision de la prédiction	Nombre	Vrais-positifs	Qualité
Batch1	+	-	+	+
Batch2	-	+	-	-

IV. DISCUSSION

- Faible taille des échantillons :
 - Filtre :
 - Qualité non vérifiée.
 - Paramètres non calibrés.

IV. DISCUSSION

- Faible taille des échantillons :
 - Filtre :
 - Qualité non vérifiée.
 - Paramètres non calibrés.
 - Peu de gènes *drivers* sous sélection positive.

V. CONCLUSION

- **Mise en place des outils** pour la suite du projet.
 - Variants
 - Sélection
 - Néo-antigènes

V. CONCLUSION

- **Mise en place des outils** pour la suite du projet.
 - Variants
 - Sélection
 - Néo-antigènes
- **Méthodologie adéquate.**
 - Cytobrosses plus adaptées que biopsies.
 - Peu d'erreurs de séquençage récurrentes.
 - Gènes attendus retrouvés sous sélection positive.
 - + Nouveaux

V. CONCLUSION

- **Mise en place des outils** pour la suite du projet.
 - Variants
 - Sélection
 - Néo-antigènes
- **Méthodologie adéquate.**
 - Cytobrosses plus adaptées que biopsies.
 - Peu d'erreurs de séquençage récurrentes.
 - Gènes attendus retrouvés sous sélection positive.
 - + Nouveaux
- **Perspectives.**
 - + Grande cohorte avec annotations cliniques.
 - Fumeurs / Non-fumeurs / Patients avec cancer

**Merci pour votre
attention !**