

BINF-F401: Projet

Étude de la glande pituitaire

Dejaeghere Justine - El Attar Sohail - Penelle Simon

1 Introduction

Dans le cadre du cours de "Computational Methods for Functional Genomics" nous avons étudié l'impact des données cliniques et des données génomiques sur l'aspect morphologique d'un organe. L'organe qui a été attribué à notre groupe est la glande pituitaire. La glande pituitaire ou hypophyse est une glande endocrine située à la base du cerveau qui joue un rôle de régulateur dans l'activité d'autres glandes et contrôle la sécrétion d'hormones impliquées dans la croissance, le métabolisme, la reproduction et la réponse au stress [5].

L'hypophyse est divisé en 2 lobes principaux : l'adénohypophyse et la neurohypophyse. Ces deux régions sécrètent des hormones différentes :

L'adénohypophyse (lobe antérieur) sécrète:

- **Hormone de croissance (GH)** : Stimule la croissance des os et des tissus ainsi que le métabolisme des lipides et des glucides.
- **Prolactine** : Favorise la production de lait dans les glandes mammaires après l'accouchement.
- **Hormone stimulante de la thyroïde (TSH)** : Régule la production d'hormones thyroïdiennes, influençant le métabolisme global.
- **Hormone adrénocorticotrope (ACTH)** : Stimule les glandes surrénales pour produire le cortisol.
- **Hormone folliculo-stimulante (FSH)** : Rôle dans le développement des follicules ovariens chez les femmes et la spermatogenèse chez les hommes.
- **Hormone lutéinisante (LH)** : Induit l'ovulation chez les femmes et stimule la production de testostérone chez les hommes.

La neurohypophyse (lobe postérieur) sécrète:

- **Ocytocine** : Rôle crucial dans l'accouchement et l'allaitement.
- **Hormone antidiurétique (ADH ou vasopressine)** : Régule l'équilibre hydrique du corps en contrôlant la concentration d'urine produite par les reins.

Un dysfonctionnement au niveau de l'hypophyse peut induire différentes maladies telles que l'adénome hypophysaire, l'hypopituitarisme, le syndrome de Cushing ou l'acromégalie [2].

2 Exploration des variables cliniques

Les données fournies proviennent de la base de données GTEx et sont composées d'un total de 245 échantillons provenant de différentes personnes et accumulant 14 observations chacunes. Dans les données cliniques, on peut retrouver comme informations le sexe, l'âge, la taille, le poids, le BMI, si la personne était donneuse ou si les organes ont été prélevés postmortem, le temps ischémique, le type de mort, la présence ou non de notes d'anatomo-pathologie et si la personne a été sous respirateur avant de mourir. Afin de visualiser la distribution des différentes variables et de détecter les possibles anomalies nous avons effectué un ensemble de graphes.

1. Distribution de l'âge

La majorité des données proviennent d'individus âgés de 35 à 65 ans, avec 157 personnes se trouvant dans cette tranche d'âge. Parmi les individus restants, on remarque que la plus grande partie est âgée de 65 à 80 ans, avec 78 personnes dans cette tranche d'âge et seulement 10 dans la catégorie restante, à savoir de 18 à 35 ans (Figure 13).

2. Distribution du sexe

On observe que les individus sont pour la plupart des hommes, 180 se trouvant dans cette catégorie (Figure 14).

3. Distribution de taille

Pour une classification optimale des individus, nous avons établi des catégories de taille sur la base de divisions arbitraires. Ces catégories sont définies comme suit : moins de 65 pouces, entre 65 et 70 pouces, et plus de 70 pouces. La majorité des individus se situent dans la catégorie supérieure, c'est-à-dire mesurant plus de 70 pouces. Parmi les autres, la plupart (96 personnes précisément) ont une taille se situant entre 65 et 70 pouces (Figure 15).

4. Distribution du poids

Nous avons divisé les poids entre 4 catégories pour pouvoir délimiter différents groupes (Figure 16) et nous avons pu mettre en évidence que les groupes majoritaires présentaient une masse de 150 à 200 livres ou de 200 à 250 livres (comprenant respectivement 108 et 89 individus).

5. Distribution du BMI

En divisant les résultats selon les intervalles de BMI utilisés par le système de classification de l'OMS (Figure 17), on observe que la majorité des individus se trouvent en surpoids, avec 105 individus ayant un BMI de cet ordre. Pour les individus restants, la répartition est similaire entre la catégorie de poids normal et d'obésité de classe 1 (comprenant respectivement 65 et 75 individus).

6. Distribution du type de prélevement

On peut observer que tous les échantillons ont été prélevés en postmortem (Figure 18).

7. Distribution du temps ischémique

Concernant les temps ischémiques, nous avons décidé de les séparer en différentes catégories afin de rendre le graphique plus lisible et cela de façon arbitraire. On peut voir que 12 échantillons présentent un temps

ischémique compris entre 0 et 500 min, 77 échantillons sont compris entre 501 et 750 min, 84 échantillons sont compris entre 751 et 1000 minutes, 51 échantillons sont compris entre 1001-1250 min et 21 échantillons sont présents dans la tranche la plus élevée (Figure 19).

8. Distribution du passage ou non par un respirateur avant le décès

On peut voir pour cette catégorie que l'on a 4 groupes, et cela même si la variable n'est censée fluctuer qu'entre 2 états. On retrouve un échantillon au statut inconnu et un autre où la valeur est manquante. On peut voir que l'on a également 221 échantillons qui ne sont pas passés par un respirateur avant leur mort et 22 pour qui c'est le cas (Figure 20).

9. Distribution du type de mort

En observant le nombre de morts par niveau de l'échelle de HARDY on observe la répartition suivante: 11 échantillons pour les cas de ventilateur, 15 pour les cas violents et rapides, 146 pour les cas rapides mais dont la mort est de cause naturelle, 22 situés sur le niveau intermédiaire de l'échelle et 51 échantillons dont la mort a été lente (Figure 21).

10. Distribution la présence ou non de notes d'anatomo-pathologie

On peut observer dans le graphique représentant le nombre de rapports d'anopath que dans notre échantillon il n'y en a aucun (Figure 22).

Afin de visualiser la distribution de nos données, nous avons décidé d'effectuer un plot de densité qui nous a permis de visualiser plus facilement la distribution et la dispersion de nos données mais également de juger l'impact de notre méthode de normalisation. Il a fallu d'abord éliminer les valeurs aberrantes trouvées au niveau de la colonne qui reporte le passage par un respirateur avant la mort. Nous avons choisi de supprimer la ligne (l'échantillon dont la valeur est manquante) pour l'échantillon contenant un statut inconnu et pour celui dont la valeur était manquante(NA). Si notre jeu de données était plus faible, nous aurions pu remplacer les valeurs manquantes par la médiane. La suppression de la colonne entière était aussi une possibilité envisageable, mais celle-ci n'est appliquée qu'en cas de manque de données massif.

Par la suite, il a été nécessaire de normaliser les variables à des intervalles comparables pour éviter que les variables possédant des intervalles plus grands dominent complètement les autres. Nous avons choisi d'utiliser la normalisation par z-score qui traite bien les valeurs aberrantes, en leur attribuant un score basé sur leur écart par rapport à la moyenne. Elle ne produit pas nécessairement des données normalisées ayant exactement la même échelle, mais elle garantit que les valeurs aberrantes n'auront pas un impact disproportionné.

On peut observer la distribution de nos variables normalisées sur le graphique de densité pour nos différentes variables (Figure 23).

2.1 Étude de la corrélation entre les variables numériques

Afin d'étudier la corrélation entre nos variables techniques et des variables démographiques, nous avons utilisé la corrélation de kendall qui est plus performante lorsque la taille de l'échantillon est petite et que'il y a beaucoup de rangs liés[1]. Pour cela, nous avons utilisé la fonction cor inclue dans R en spécifiant l'utilisation de la méthode 'kendall'. Nous avons résumé l'ensemble de nos valeurs de corrélations dans une heatmap afin que la visualisation des valeurs hautement corrélées soit visibles (Figure 1). La diagonale rouge représente les auto-corrélations et est donc égale à 1.

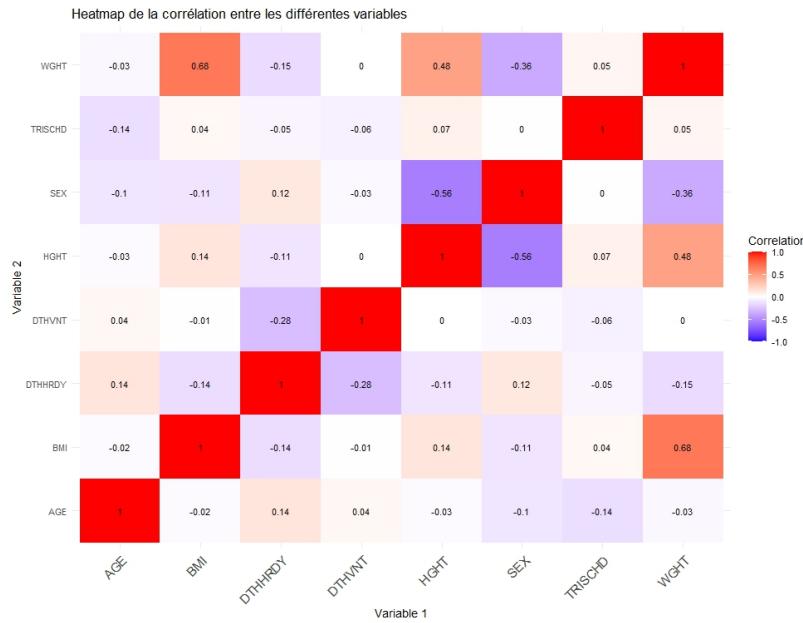


Figure 1: Heatmap des valeurs de corrélation entre les différentes variables

En ce qui concerne les variables démographiques et liées à la santé, on peut observer que le poids est corrélé à hauteur de 0,68 avec le BMI et de 0,48 avec la taille. Ces deux valeurs semblent logiques étant donné que le BMI est calculé en fonction du poids et que les individus de plus grande taille ont généralement un poids plus élevé. Le sexe quant à lui présente de multiples corrélations, aussi bien avec le BMI (-0,11) qu'avec le type de mort (0,12), la taille (-0,56) et le poids (-0,36).

Pour la taille, on peut voir que celle-ci est corrélée positivement à hauteur de 0,14 de avec le BMI et a une valeur de 0,48 avec le poids. Enfin, l'âge est légèrement corrélé négativement (-0,14) avec le temps ischémique et positivement (0,14) avec le type de mort.

Les variables techniques présentent également des corrélations. Concernant le temps ischémique, on peut voir qu'il corrélé négativement à hauteur de -0,14 avec l'âge. Le temps passée sous respirateur est corrélé négativement à hauteur de -0,28 avec le type de mort.

Le type de mort est quant a lui légèrement corrélé positivement(0,14) avec l'âge, et négativement (-0,14) avec le BMI, le temps passé sous respirateur avant le décès (-0,28) et le poids à hauteur de -0,15.

Afin de tester l'existence ou non de notre corrélation, il faut se référer à la p-valeur du test qui nous permet de savoir si la corrélation obtenue est due au hasard ou non. Nous avons donc décidé de faire une heatmap des différentes

valeurs de p-values en fonction des association de variables (Figure 2). Nous avons choisi dans ce cas de figure un seuil de p-value $\leq 0,05$ en dessous duquel nous avons décidé du rejet de H_0 et que donc le rejet de l'hypothèse nulle n'est pas dû au hasard (Dans le cas de la corrélation de Kendall, l'hypothèse nulle est que les deux variables étudiées sont indépendantes l'une de l'autre)[6]. Ainsi, nous avons pu vérifier si nos corrélations rejettent bel et bien H_0 car dans certains cas de figures même si la corrélation est élevée, la p-valeur nous informe que la corrélation est due au hasard.

Si l'on regarde pour les variables ayant des corrélation au dessus de 0,1 en valeur absolue regroupées dans le tableau 1 (Table 1), on peut voir que la totalité des corrélations possèdent une p-valeur associée inférieure à 0,05 (les valeurs ont été arrondies ce qui explique la présence de p-values = 0); on peut donc en conclure que le rejet de l'hypothèse nulle n'est pas du au hasard et que donc les corrélations existent bel et bien et ne sont pas le fruit du hasard.

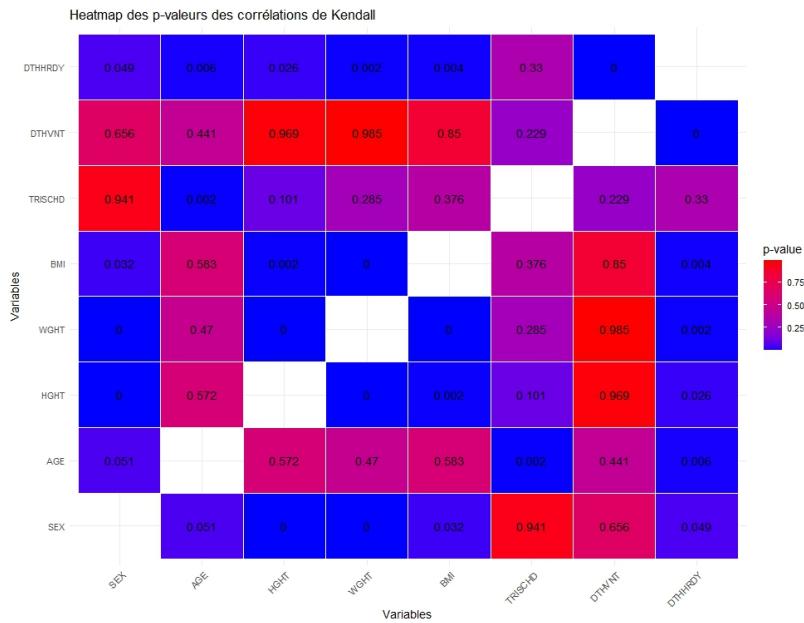


Figure 2: Heatmap des p-values pour les valeurs de corrélation entre les différentes variables

Table 1: Tableau des corrélation supérieures à 0,1 et leurs p-valeur associées

	Variable_1	Variable_2	Corr	P-val
2	AGE	DTHHRDY	0.14	0.01
5	AGE	SEX	-0.10	0.05
6	AGE	TRISCHD	-0.14	0.00
9	BMI	DTHHRDY	-0.14	0.00
11	BMI	HGHT	0.14	0.00
12	BMI	SEX	-0.11	0.03
14	BMI	WGHT	0.68	0.00
15	DTHHRDY	AGE	0.14	0.01
16	DTHHRDY	BMI	-0.14	0.00
17	DTHHRDY	DTHVNT	-0.28	0.00
18	DTHHRDY	HGHT	-0.11	0.03
19	DTHHRDY	SEX	0.12	0.05
21	DTHHRDY	WGHT	-0.15	0.00
24	DTHVNT	DTHHRDY	-0.28	0.00
30	HGHT	BMI	0.14	0.00
31	HGHT	DTHHRDY	-0.11	0.03
33	HGHT	SEX	-0.56	0.00
35	HGHT	WGHT	0.48	0.00
36	SEX	AGE	-0.10	0.05
37	SEX	BMI	-0.11	0.03
38	SEX	DTHHRDY	0.12	0.05
40	SEX	HGHT	-0.56	0.00
42	SEX	WGHT	-0.36	0.00
43	TRISCHD	AGE	-0.14	0.00
51	WGHT	BMI	0.68	0.00
52	WGHT	DTHHRDY	-0.15	0.00
54	WGHT	HGHT	0.48	0.00
55	WGHT	SEX	-0.36	0.00

2.2 Recherche de variables techniques confondantes

La présence d'une variable confondante peut être à l'origine d'une partie voire de l'entièreté d'une association entre variables. Elle peut donc entraîner une surévaluation de l'association (confusion positive), l'association est alors perçue comme présentant une force supérieure à la réalité. Cependant, il est aussi possible qu'elle soit à l'origine d'une sous-estimation de l'association (confusion négative) : l'association semble d'une force inférieure par rapport à ce qu'elle est réellement.

Trois conditions doivent être réunies pour qu'il y ait présence d'une variable confondante (Article [3]) :

1. Obligation d'association entre la variable confondante et à la fois le facteur de risque étudié et le résultat.
2. Présence d'une répartition inégale de la variable confondante entre les variables comparées.
3. Ne peut être une étape intermédiaire entre l'exposition et le résultat.

Nous avons étudié la présence ou non de variable confondante pour les variables techniques, à savoir le type de mort (échelle de Hardy), la présence ou non de ventilateur avant le décès et le temps ischémique. Pour cela, nous avons supprimé les variables techniques des données et avons sélectionné les paires de manière à ce qu'il n'y ait pas de doublons, en ne prenant que les éléments de la matrices de corrélation où l'indice de ligne est inférieur à l'indice de colonne (éléments situés en dessous de la diagonale de corrélations unitaires entre les valeurs et elles-mêmes).

Nous avons alors effectué, pour chaque paire de variables dépendante/indépendante, différentes régressions linéaires. Nous avons commencé par déterminer le coefficient lié à la régression linéaire simple de la variable dépendante en fonction de la variable indépendante. Ensuite, nous en avons fait de même pour quatre scénarios différents. Le premier consistait à inclure les trois possibles variables confondantes, et les trois derniers sont les coefficients résultants du retrait de respectivement la variable DTHVNT, TRISCHD et DTHHRDY. Cette étape nous a permis de mettre en évidence la variable technique qui présentait le plus d'impact sur la relation entre la variable dépendante et la variable indépendante, donc celle ayant induit la modification la plus élevée en terme de coefficient lors de son retrait.

On peut observer dans le tableau ci-dessus que dans 9 cas sur 10, l'impact maximal était dû au retrait de la variable de Hardy (Table 2).

	variable 1	variable 2	Max impact	m.c.DTHVNT (%)	m.c.TRISCHD (%)	m.c.DTHHRDY (%)	DTHVNT	TRISCHD	DTHHRDY
1	SEX	AGE	DTHHRDY	1.07	7.16	29.48	/	/	OUI
2	SEX	HGHT	TRISCHD	0.01	0.57	0.53	/	/	/
3	AGE	HGHT	DTHHRDY	3.95	32.45	56.57	/	OUI	OUI
4	SEX	WGHT	DTHHRDY	0.15	0.68	2.00	/	/	/
5	AGE	WGHT	DTHHRDY	0.87	37.76	81.66	/	OUI	OUI
6	HGHT	WGHT	DTHHRDY	0.15	0.27	0.83	/	/	/
7	SEX	BMI	DTHHRDY	1.89	0.47	16.65	/	/	OUI
8	AGE	BMI	DTHHRDY	7.58	69.08	168.73	/	OUI	OUI
9	HGHT	BMI	DTHHRDY	1.94	2.01	13.36	/	/	OUI
10	WGHT	BMI	DTHHRDY	0.19	0.19	1.29	/	/	/

Table 2: Étude de la possibilité de variable technique confondante, m.c. signifiant magnitude de la confondance

Par la suite, nous avons étudié la magnitude de la confondance pour chacune des variables techniques en appliquant la formule suivante :

$$\text{Magnitude de la confondance} = \left| \frac{C_{\text{brut}} - C_{\text{ajusté}}}{C_{\text{brut}}} \right|$$

Dans cette formule, C_{brut} représente le coefficient non ajusté à savoir le coefficient provenant du modèle de régression ne prenant en compte que les variables dépendante et indépendante et $C_{\text{ajusté}}$ représente le coefficient obtenu en ajoutant la variable possiblement confondante. Dans le cas où cette magnitude dépasse le seuil de 10%, nous avons conclu que la variable était confondante pour cette association. Ce cas de figure est représenté par les "OUI" (Table 2), et on peut donc observer que dans les cas 1, 7 et 9, la variable Hardy est une variable confondante pour l'association. Plus encore, dans les cas 3, 5, et 8 les variables du temps ischémiques et de Hardy sont concernées.

Pour avoir un meilleur aperçu de l'impact de ces variables confondantes, nous avons construit des graphiques de variation de la variable dépendante en fonction de la variable indépendante, en tenant compte de ces dernières.

Nous avons étudié en détails le cas numéro 8, à savoir la variation du BMI en fonction de l'âge en tenant compte des variables Trischd et Hardy car la magnitude de la confondance était la plus élevée avec une valeur de 69.08 et 168.73% pour TRISCHD et DTHHRDY respectivement (Figure 3). Dans ce graphique de dispersion, on peut observer les différentes catégories de Trischd et de Hardy, ainsi que les histogrammes marginaux représentants la distribution des variables du BMI et de l'âge. Ces deux variables présentent une droite de régression à la pente négative, impliquant que le BMI aurait tendance à diminuer avec l'âge. Cependant, les variables techniques correspondants au temps ischémique et au type de mort sont à prendre en compte dans cette analyse sachant qu'ils ont été mis en évidence dans l'étape précédente comme possibles variables confondantes.

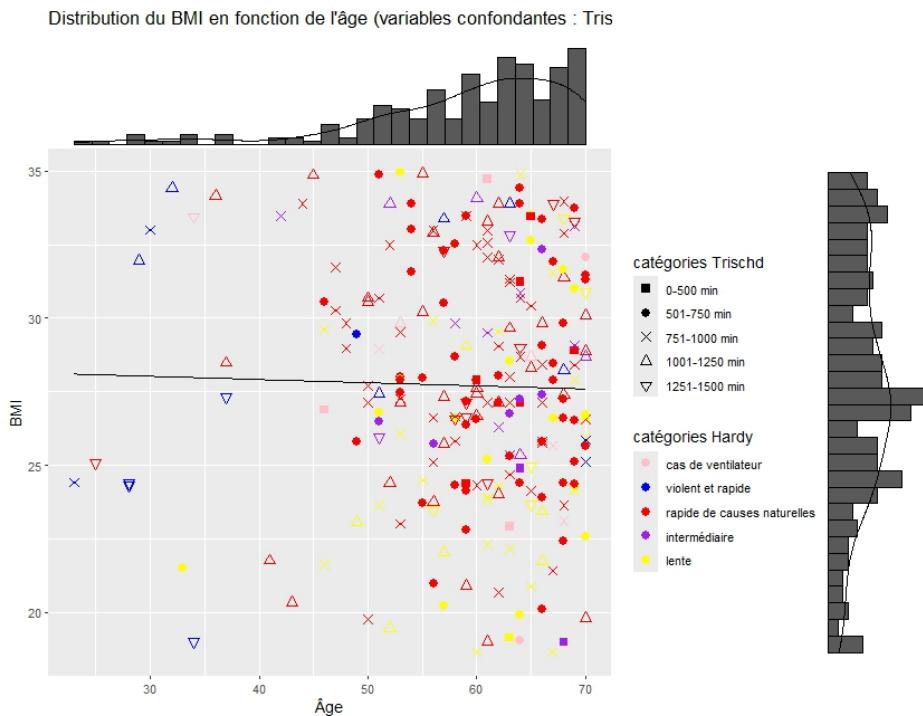


Figure 3: Mise en évidence de la confondance dans la relation BMI/âge

Pour étudier leur impact sur la relation entre le BMI et l'âge, nous avons ajusté la relation entre le BMI et l'âge pour TRISCHD en faisant une régression pour chacune des catégories de temps ischémique. On peut observer que la régression ne donne pas le même résultat que dans notre régression brute donc sans tenir compte de la variable confondante. En effet, les droites de régression ont des pentes assez variables selon le groupe considéré, allant de plus négatif pour le groupe des individus au temps de ventilation compris entre 0 et 500 minutes jusqu'à positif avec le groupe des temps de ventilation les plus élevés compris entre 1251 et 1500 minutes (Figure 4).

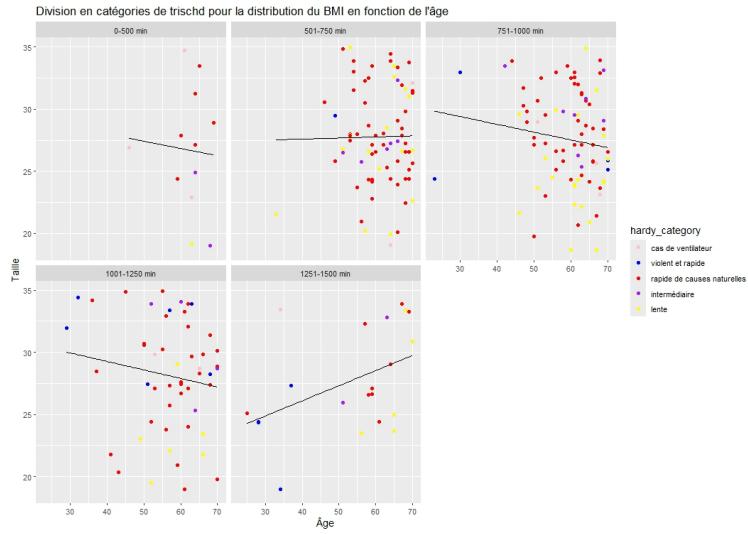


Figure 4: Division selon la variable TRISCHD pour la mise en évidence de confondance dans la relation BMI/âge

C'est également le cas pour la variable technique des types de mort HARDY. En effet, on peut observer des résultats similaires à ceux des régressions ajustées pour TRISCHD, avec une pente davantage négative pour les cas de ventilateurs et positive pour les morts dites lentes (Figure 5).

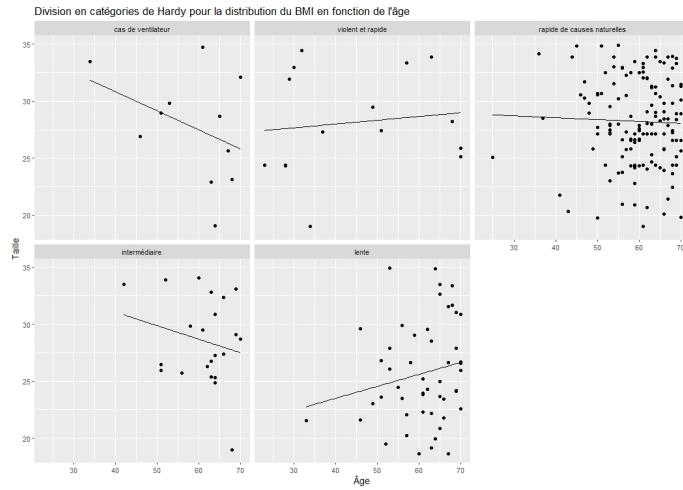


Figure 5: Division selon la variable DTHHRDY pour la mise en évidence de confondance dans la relation BMI/âge

En étudiant les autres associations possédant des variables confondantes, on peut faire les mêmes observations. Pour l'association entre l'âge et le sexe, on remarque qu'ajuster pour la variable du type de mort induit une variation au niveau de la pente de la droite de régression, qui devient même légèrement positive dans le cas de mort rapide de causes naturelles (Figure 24). De même, pour l'association entre la taille et l'âge (Figure 25), ainsi que pour l'association entre la masse et l'âge (Figure 26), nous observons des variations dans les valeurs des pentes. Ces variations sont telles qu'elles peuvent même devenir positives, à partir d'une pente brute initialement négative.

Il est toutefois à noter que l'effet de confondance est atténué lorsque la magnitude de confondance est plus faible. Par exemple, dans l'association entre le sexe et le BMI, où la magnitude de confondance est de 16.65 (Figure 27), ou dans l'association entre le BMI et la taille, avec une magnitude de confondance de 13.36 (Figure 28), les variations observées sont moins prononcées.

Ces observations soulignent l'importance de prendre en compte ces variables confondantes dans les analyses statistiques ultérieures. En effet, leur prise en compte permet d'obtenir des résultats plus précis et fiables, en évitant les biais potentiels introduits par ces variables confondantes.

3 Association entre variables cliniques et données morphologiques

En ce qui concerne la recherche d'association entre les variables cliniques et les données morphologiques, nous avons décidé de fusionner nos données morphologiques avec nos données cliniques (SEX, AGE, HIGHT, WGHT, BMI, TRISCHD, DTHVNT et DTHHRDY). Pour cela, il a fallu supprimer les lignes correspondant aux échantillons GTEX.16Z82.3126 et GTEX.1B8L1.3126 qui sont les échantillons ayant des valeurs inconnues pour DTHVNT.

Afin de procéder à la recherche en tant que telle des liens entre nos différentes variables, nous avons opté pour une approche consistant en une étude de corrélation en premier lieu suivie d'une recherche plus approfondie avec une analyse différentielle. Cette analyse différentielle est effectuée via DESeq2, qui permettra in fine de mettre en évidence les clusters morphologiques dont la présence est la plus influencée par nos variables cliniques.

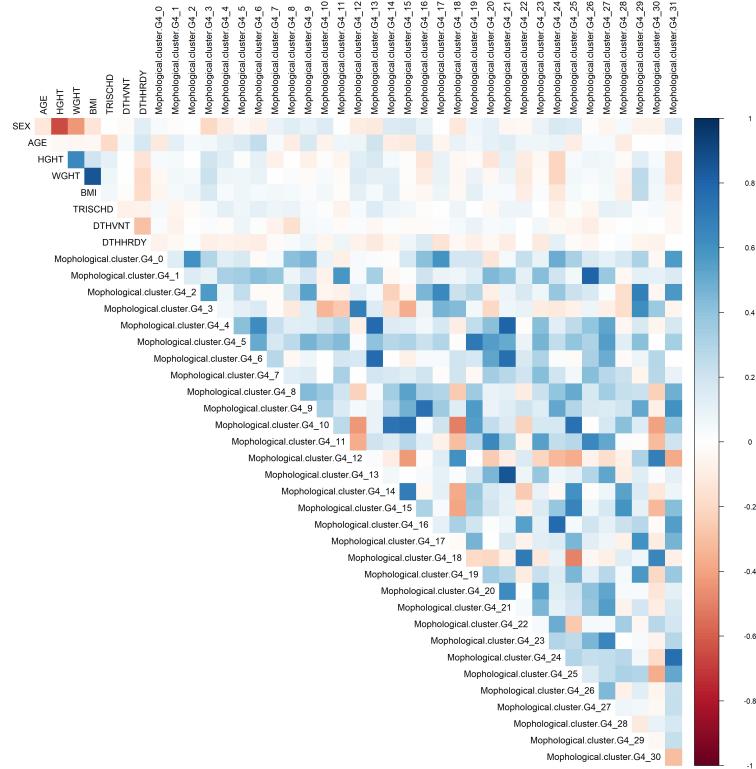


Figure 6: Matrice de corrélation des variables cliniques et des variables morphologiques

La matrice de corrélation ci-dessus (Figure 6) a été construite en utilisant comme méthode de calcul la corrélation de Spearman et non pas la corrélation de Kendall qui serait moins performante au vu de la quantité de variables à traiter. On peut voir que cette matrice reprend également les corrélations entre les variables cliniques, cependant cela ne sera pas pris en compte ici car nous l'avons déjà traité dans la section 2 avec une méthode plus appropriée. Nous ne parlerons ici que des corrélations entre les différentes variables cliniques et les différentes variables morphologiques, les valeurs de corrélations entre les différents groupes morphologiques ne seront pas traitées.

Si on examine les valeurs de corrélation pour les différentes variables démographiques et liées à la santé, on peut mettre en évidence différents clusters dont les variations sont liées à ces variables. Pour cela nous avons sélectionné les clusters dont les valeurs de corrélation sont supérieures ou égales à 0.1 en valeur absolue, ce qui permet de faire ressortir différents clusters morphologiques.

En ce qui concerne la variable liée au sexe, on peut observer 13 clusters présentant une corrélation significative (Table . Les variables de l'âge (Table 9) et de la taille (Table 10) présentent également une grande quantité de clusters

à la corrélation significative toutes en comptant 12.

Cela diffère pour les variables de la masse (Table 11) et du BMI (Table 12), qui présentent respectivement 9 et 6 clusters corrélés significativement. On peut noter que dans les deux cas, le cluster présentant la plus forte corrélation est le Morphological.cluster.G4.29, ce qui est prévisible au vu de la nécessité de connaître la masse pour déterminer le BMI (les deux variables sont corrélées positivement).

Par la suite, nous avons effectué une analyse différentielle avec DESeq2 pour mettre en évidence les liens entre les clusters et les variables techniques ou cliniques.

Pour effectuer cette analyse différentielle sur les variables techniques (temps ischémique, type de mort et présence ou non de ventilateur avant la mort), nous avons appliqué le protocole suivant à chacune des variables :

Nous avons d'abord utilisé la fonction DESeq pour effectuer différentes étapes, telles que l'estimation des dispersions et l'ajustement du modèle. Ensuite, la fonction results nous a permis d'extraire les résultats de l'objet DESeq ajusté. Ces résultats comprennent différentes valeurs, dont le log2FoldChange, qui est une mesure de la variation en terme de cluster morphologique entre la condition initiale et lorsque l'on tient compte de la variable technique.

Par la suite, nous avons généré des MA-Plot qui permettent d'identifier les variations en terme de clusters morphologiques entre les deux conditions. Dans ce type de graphe, chaque point représente un cluster, l'axe des x représente la moyenne des comptages normalisés pour ce cluster, et l'axe des y représente le log2 du foldchange pour ce cluster.

Les résultats ont enfin été filtrés pour ne conserver que ceux dont la valeur p ajustée était inférieure à 0.05, ce qui indique une variation significative. Les 10 premiers résultats ont ensuite été affichés. Ces résultats sont ceux qui présentent la plus grande variation en termes de fold change, qui est également statistiquement significative.

On peut observer dans le tableau ci-dessous (table 3) que la variable TRISCHD n'est liée à aucun cluster, que DTHVNT est liée avec 7 clusters différents et que DTHHRDY est liée avec 3 clusters. Nous avons utilisé le seuil arbitraire de p-valeur ajustée inférieure à 0,05, ce qui nous permet de dire que la différence exprimée est potentiellement réelle et non liée au hasard.

Les MA-plots permettent d'avoir une meilleure visualisation de ces résultats (Figure 7). On observe effectivement en bleu les clusters qui présentent une variation entre les deux conditions, donc entre la condition brute et celle tenant compte de la variation des variables techniques. Les résultats corroborent avec le tableau ci-dessus, on observe effectivement une absence de cluster régulé pour le temps ischémique, 7 clusters pour la présence ou non de ventilateur avant la mort et enfin 3 pour la variable technique représentant les types de mort.

Étant donné que l'analyse différentielle et la recherche des clusters les plus régulés pour les variables techniques a mis en évidence que le temps ischémique n'était associé à aucune population, nous avons décidé d'étudier l'impact

Table 3: Résultats de l'association entre données cliniques et morphologiques pour les variables techniques

	Variable	Regulated_Clusters	Top_10_Regulated_Clusters	p_value_ajusted
1	TRISCHD	0		
2	DTHVNT	7	Morphological.cluster.G4_1 Morphological.cluster.G4_4 Morphological.cluster.G4_7 Morphological.cluster.G4_8 Morphological.cluster.G4_16 Morphological.cluster.G4_24 Morphological.cluster.G4_26	0.0228922070008457 0.0132968435455 0.0475174621929639 0.00772433862386856 0.0132968435455 0.0132968435455 0.0231206376750739
3	DTHHRDY	3	Morphological.cluster.G4_10 Morphological.cluster.G4_16 Morphological.cluster.G4_24	0.0402665367258686 0.000113920285603039 0.00733806356088321

de ne pas le prendre en compte lors de l'ajustement des autres variables cliniques qui suit.

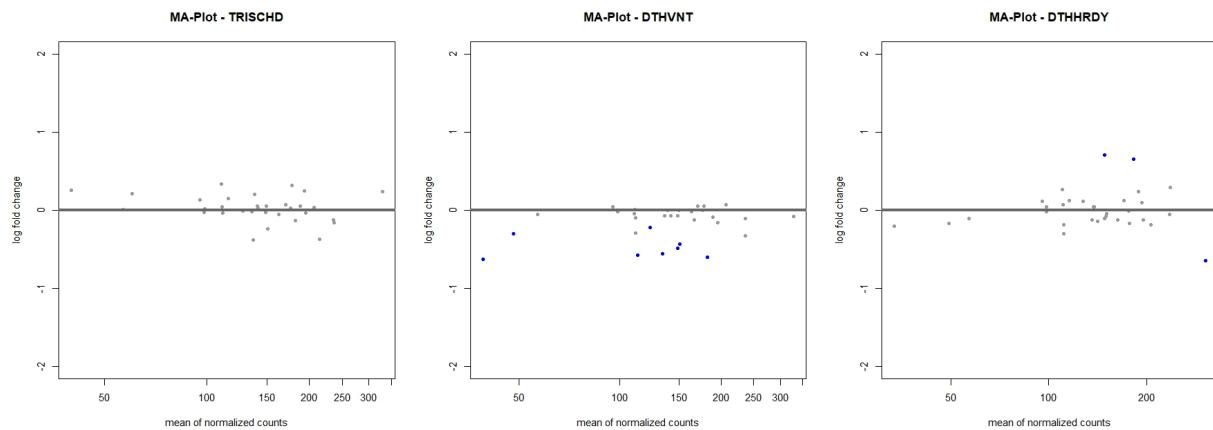


Figure 7: MA-plot des variables techniques

Pour cela, nous avons effectué le même protocole DESeq2 que précédemment pour les variables restantes, à savoir l'âge, le sexe, la masse, la taille et le BMI. Nous avons commencé par ajuster les données à la fois pour la variable de l'échelle de Hardy et pour le temps ischémique pour les variables pour lesquelles on observait une confondance dans nos analyses précédentes (Table 4).

Ensuite, nous avons effectué le même procédé mais cette fois-ci en ne prenant en compte que la variable confondante de l'échelle de Hardy (Table 5). En comparant les deux tableaux, on peut observer que les clusters régulés sont exactement identiques dans les deux cas, ce qui a confirmé notre intuition. L'ajustement par le temps ischémique n'est pas nécessaire et cela malgré son pouvoir de confondance révélé lors de l'analyse précédente, lui même ne semblant pas "réguler" de cluster morphologique.

Table 4: Résultats de l'association entre données cliniques et morphologiques en ajustant pour les variables confondantes

	Variables	Regulated_Clusters	Top_10_Regulated_Clusters	p_value_ajusted
	AGE			
1	+ TRISCHD	2	Morphological.cluster.G4_16	0.000821662975672862
	+ DTHHRDY		Morphological.cluster.G4_24	0.01301975180446
2	SEX	2	Morphological.cluster.G4_16	0.00031324414561557
	+ DTHHRDY		Morphological.cluster.G4_24	0.0155379669563199
	WGHT			
3	+ TRISCHD	2	Morphological.cluster.G4_16	9.7487401946439e-05
	+ DTHHRDY		Morphological.cluster.G4_24	0.0345182761855569
	HGHT			
4	+ TRISCHD	2	Morphological.cluster.G4_16	0.000678736209256733
	+ DTHHRDY		Morphological.cluster.G4_24	0.0259983446567536
	BMI			
5	+ TRISCHD	4	Morphological.cluster.G4_13	0.0413641579918321
	+ DTHHRDY		Morphological.cluster.G4_16	0.000100234558062738
			Morphological.cluster.G4_20	0.0336601829498327
			Morphological.cluster.G4_24	0.0181199855167048

En ce qui concerne les clusters mis en évidence en tant que tels, on peut voir que pour les différentes variables on retrouve toujours le cluster 16 et 24 sauf pour le BMI où l'on retrouve également le cluster 13 et 20. Pour chacun, les p-valeur ajustées inférieures à 0,05 laissent également penser que ces clusters morphologiques découverts lors de l'analyse ne seraient pas liée au hasard, comme dans le cas des variables techniques.

Table 5: Résultats de l'association entre données cliniques et morphologiques en ajustant pour les variables confondantes sans prendre en compte TRISCHD

	Variables	Regulated_Clusters	Top_10_Regulated_Clusters	p_value_ajusted
1	AGE + DTHHRDY	2	Morphological.cluster.G4_16 Morphological.cluster.G4_24	0.000821662975672862 0.01301975180446
2	SEX + DTHHRDY	2	Morphological.cluster.G4_16 Morphological.cluster.G4_24	0.00031324414561557 0.0155379669563199
3	WGHT + DTHHRDY	2	Morphological.cluster.G4_16 Morphological.cluster.G4_24	9.7487401946439e-05 0.0345182761855569
4	HGHT + DTHHRDY	2	Morphological.cluster.G4_16 Morphological.cluster.G4_24	0.000678736209256733 0.0259983446567536
5	BMI + DTHHRDY	4	Morphological.cluster.G4_13 Morphological.cluster.G4_16 Morphological.cluster.G4_20 Morphological.cluster.G4_24	0.0413641579918321 0.000100234558062738 0.0336601829498327 0.0181199855167048

Comme pour les variables techniques, les MA-Plots permettent d'avoir une meilleure visualisation des résultats et correspondent bien aux résultats attendus. En bleu, on peut effectivement observer les clusters qui présentent une variation entre les deux conditions, donc deux clusters régulés pour toutes les variables cliniques étudiées à l'exception du BMI qui en présente quatre (Figure 8).

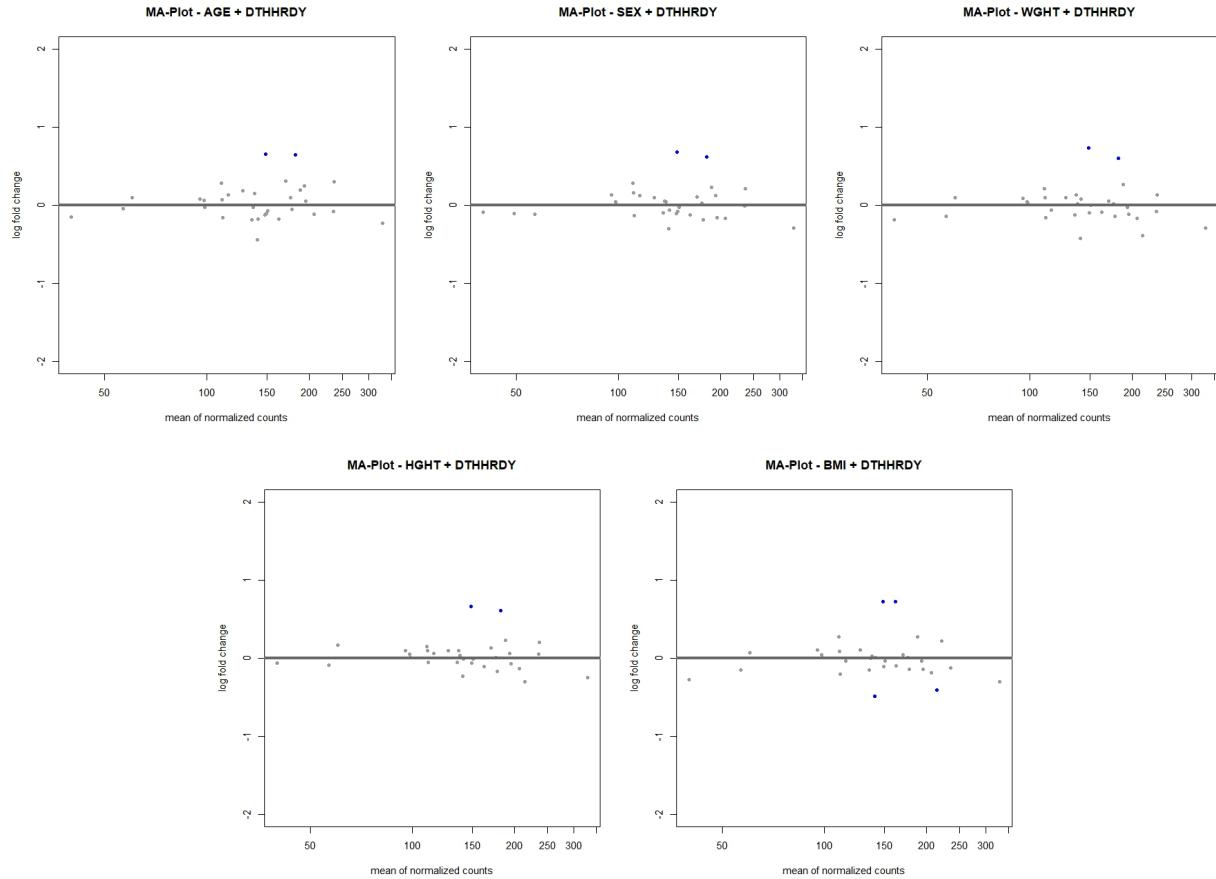


Figure 8: MA-plot des variables démographiques et liées à la santé

4 Comparaison entre expression des gènes, données de morphologie et voies biologiques

4.1 Association entre données morphologiques et expression des gènes

L'objectif dans cette section était de systématiquement comparer les résultats de nos clusters morphologiques à des données de RNA-seq qualifiant la régulation (up et down) de certains gènes et l'association de ces gènes à chaque cluster de données morphologiques. Nous avons donc programmé un script en R dans le but de trouver des signatures spécifiques de gènes pour chaque cluster. Notre objectif était de capturer, pour chacun des 32 clusters, le nombre de gènes ayant été régulé significativement au niveau de leur expression (vers le haut ou vers le bas), ainsi que les 10 gènes ayant le plus été différemment exprimés pour ce cluster.

Pour réaliser cet objectif nous avons procédé à une analyse systématique avec DESeq2, un outil très populaire pour les données de ('count') transcriptions, pour chaque cluster, identifiant les gènes régulés correspondants. Nous avons commencé par filtrer les exprimés trop faiblement simplement en utilisant un threshold (instancié à 10) de count minimum pour qu'un gene soit pris en compte dans l'analyse. Une fois le filtrage effectué, nous avons, pour des raisons de

lisibilité et de coherance, traduit chaque '*gene_id*' par le symbole HGNC associé (par exemple: 'ENG00000223972.5' est transformé en 'DDX11L1'). Cette transformation à été réalisé à partir de la colonne description de notre fichier *RNA_read_counts.tsv* qui contient pour chaque gene le nom correspondant.

Les résultats de cette analyse sont présentés dans la Table 6

Table 6: Résultats de l'association entre données de transcription et morphologiques

No.	Cluster	Number Regulated Genes	Top 10 Genes	P-value Adjusted
1	Morphological.cluster.G0	2158	KRTAP10-7, RNF17, LINC01456, CFC1B, IGFBP1, IGHV3-11, RP11-350G24.1, PRG4, OR9K1P, MEPE	0.021193012103821, 0.0371794746410133, 0.0421682732474655, 0.0002587919073789, 0.0210221495022811, 0.014091579580737, 0.0402811895915435, 0.0214351992922607, 0.0483175948325226, 0.00526260563267082
2	Morphological.cluster.G1	176	IGKV1-16, RP11-76C10.3, RP11-76C10.4, MT-TL1, SLC22A6, SLC5A5, LINC00890, PLAC9P1, SLC22A2, RP11-479J7.2	2.23040154225916e-33, 1.4499612908303e-21, 7.53071595136142e-15, 3.12585236765156e-19, 0.00191455593253214, 9.92682135835791e-05, 0.000117007776232616, 0.0146702606250464, 0.00494822524538865, 0.00839900112683786
3	Morphological.cluster.G2	2223	RP11-76C10.3, RP11-76C10.4, BRS3, IGFBP1, IGLV4-60, PSG4, APOBEC3H, CALCA, ANKRD33, RP11-244E17.1	1.0129525244131e-05, 0.000185005375763144, 6.45276944866429e-10, 0.000309224133353486, 0.0472320442832511, 0.00492703023986701, 1.09033647485925e-05, 0.000235698700974788, 0.00137029340059827, 0.0158587337132761
4	Morphological.cluster.G3	64	ADGRG7, CST2, KRT85, DCD, BRS3, KRT31, MUCL1, MFRP, PAX2, TTR	0, 0, 2.12958460663967e-06, 7.0048240367365e-07, 6.92239784178221e-22, 0.0307864889679675, 0.0085573179003579, 0.000228802384108589, 0.0232233911372789, 0.000300111751030236
5	Morphological.cluster.G4	299	ADGRG7, IGKV1-16, RP11-76C10.3, RP11-76C10.4, IGHV4-31, IGHGP, KRT16P2, KRT16, IGLV4-69, IGLV3-27	4.11633379629164e-08, 1.92541629461653e-40, 1.81009187710465e-14, 9.33184900043225e-08, 5.88613795635747e-08, 4.2994260753915e-16, 0.0016577765670507, 1.18206544616034e-11, 1.05504769291035e-08, 0.00572160964214409
6	Morphological.cluster.G5	349	IGKV1-16, RP11-76C10.3, RP11-76C10.4, SMR3B, IGHV3-11, IGHV4-31, TTR, IGLV3-19, IGLV4-69, MFRP	2.21389431947732e-23, 1.00985345087733e-08, 6.52682883309119e-05, 9.61009621325569e-06, 3.01973738236566e-05, 0.00198306876398684, 5.15091700395945e-08, 0.000361678064535235, 0.000860012730556655, 8.28668587658204e-05

Continued on next page

Table 6 continued from previous page

No.	Cluster	Number Regulated Genes	Top 10 Genes	P-value Adjusted
7	Morphological.cluster.G6	95	HTR2C, IGHV5-51, MT-TL1, IGKV1-16, IGLV4-69, IGKV3-20, IGKV3D-20, IGHV4-31, IGKJ1, COMP	7.67770133338512e-79, 9.84414189096538e-18, 6.70981469756844e-22, 3.29963471005101e-16, 2.47307374470577e-08, 3.73798022378449e-11, 0.000588414760990318, 5.54034418035342e-05, 3.53683924598218e-10, 2.95031258905967e-06
8	Morphological.cluster.G7	213	SPINK7, ADH7, TGM3, KRT78, MUC21, CRNN, IGKV1-16, KRT4, PADI1, CLCA4	1.19615609497281e-10, 2.58713051769885e-07, 2.20945255674589e-17, 9.4975205498307e-14, 3.04027380662024e-13, 2.23489551760007e-14, 1.69386176459638e-19, 4.1255148461282e-15, 3.38548716862811e-13, 1.04955718233149e-15
9	Morphological.cluster.G8	154	IGKV1-16, DMBT1, KRT78, TGM3, SPINK7, MUC21, ADH7, CRNN, KRT4, RP11-76C10.3	2.39062134254775e-45, 1.48012896997969e-77, 1.12808218715128e-62, 1.3339434580742e-76, 8.96818525659624e-42, 3.48351169458172e-57, 1.00342610672772e-25, 3.84017499526605e-61, 1.51313045648074e-64, 3.33878522008301e-205
10	Morphological.cluster.G9	194	RP11-76C10.3, IGHV3-11, RP11-76C10.4, C1QL2, MMP10, PROKR2, SLC5A5, IGLV8-61, TGM7, DNASE2B	0.000762314178367132, 5.26594097283441e-05, 0.00639589952015844, 9.86123301130364e-09, 0.0250184922370384, 0.00259545131830815, 0.026446077030289, 0.0339635899711583, 0.0227692974654373, 0.00163534635768335
11	Morphological.cluster.G10	417	RP11-76C10.3, RP11-76C10.4, CYP11B1, REG3A, SMR3B, PGA3, TFF2, GIF, C1QL2, PGC	1.82264978227675e-10, 1.77572068977218e-05, 4.04108923565823e-17, 0.00113327522508295, 7.71943816016276e-07, 1.26060390683791e-05, 0.00111315472056712, 2.09400409690583e-06, 8.04768566685584e-10, 8.64068837022992e-08
12	Morphological.cluster.G11	369	ADH7, KRT78, MUC21, DMBT1, KRT4, CRNN, TGM3, MMP3, SPRR3, SPINK7	2.64283249541853e-11, 5.01202473951909e-19, 1.58800403164835e-18, 9.52481434062943e-32, 1.561976974
13	Morphological.cluster.G12	45	PLAC9P1, IGKV1-16, RP11-567M16.1, CXCL6, RP11-244E17.1, NPY, SIAH3, CTD-255A7.2, IGHGP, TRH	4.02016723703676e-05, 2.63846844618986e-06, 1.15776273586451e-08, 0.0391175314582756, 0.00309822145965974, 2.32609865386122e-05, 5.19567208696312e-10, 2.63846844618986e-06, 0.0044577877496853, 0.00351178766623065
14	Morphological.cluster.G13	147	KRTAP10-7, IGKV1-16, IGHV5-51, AC053503.6, IGLV4-69, MT-TL1, IGKV3D-20, IGHV4-31, IGKV3-20, COMP	8.42806408677952e-07, 6.45793390711009e-27, 5.86567902823196e-17, 0.000915912626090031, 1.64803700301344e-10, 2.79894845037441e-12, 2.34922436343377e-05, 4.83604746050419e-06, 3.28521467728275e-11, 2.1158808407414e-085

Continued on next page

Table 6 continued from previous page

No.	Cluster	Number Regulated Genes	Top 10 Genes	P-value Adjusted
15	Morphological.cluster.G14	69	IGKV1-16, PLAC9P1, OLIG2, C5orf17, OLIG1, MIR219A2, RP11-17A4.2, MOG, IGLV3-19, CTD-2555A7.2	3.30117014331019e-30, 7.44492004065985e-12, 0.000122260123028395, 5.11313924302811e-05, 2.44259965411333e-05, 0.000538146450097263, 0.000117156957807013, 0.0186172710187711, 0.00610009322834877, 2.44259965411333e-05
16	Morphological.cluster.G15	55	MMP3, LHX2, SLTRK2, RP11-76C10.4, PNLLPRP1, CPA2, REG1A, CELA3B, REG1B, PNLLPRP2	0, 0, 1.70052100741554e-203, 2.00986024057433e-07, 8.83906761025864e-21, 4.29715207023192e-23, 4.36584795702809e-14, 7.36243449990389e-11, 1.31101087772668e-09, 1.51686181550193e-10
17	Morphological.cluster.G16	753	SLC22A8, SLC22A6, BRS3, SLC5A5, SLC13A4, CD177, KCNJ13, PROKR2, SLC22A2, IGLV4-69	9.13124215565146e-209, 7.81163957501151e-10, 8.4197778641755e-36, 3.3288043646552e-13, 4.3058669137408e-14, 3.98783948095696e-08, 2.30685597955621e-11, 1.15807565620328e-05, 0.00497567678889661, 0.000472335520793408
18	Morphological.cluster.G17	610	DMBT1, IGKV1-16, CYP11B1, KRTAP10-7, REG3A, RP11-76C10.3, RP11-76C10.4, PGA3, GIF, TFF2	1.6947293430545e-24, 1.42315671863487e-23, 1.43569707413907e-19, 8.74698599892336e-05, 8.14075703042383e-09, 2.21900859774908e-07, 0.000470134329541894, 1.37384190960478e-06, 3.06474023128566e-07, 0.000384861040445881
19	Morphological.cluster.G18	208	IGLV7-43, KCNJ13, AJ006998.2, CD177, MIR3142HG, IGFN1, CSRP3, XIRP2, RP11-96L7.2, GADL1	0.00874329309204064, 4.92393385144515e-08, 1.45668201295596e-06, 0.00064091886947543, 3.56837808587234e-08, 6.08134653291154e-08, 0.000161104917150318, 0.0091109110947328, 0.031628769765323, 0.0321436250182119
20	Morphological.cluster.G19	1004	IGKV1-16, DMBT1, SMR3B, BPIFB1, PGA3, PGC, TFF2, DNASE2B, C1QL2, GIF	6.39799569087428e-16, 7.41635847165989e-19, 1.28477988658115e-08, 0.000253158956758204, 0.00241766676877314, 0.000327777240810467, 0.0269903662670282, 0.000178798884601417, 5.52667393831311e-06, 0.00504647914829277
21	Morphological.cluster.G20	638	KRTAP10-7, CYP1A1, SFTPB, TTR, IGLV4-69, MFRP, IGHV3-11, PRG4, IGLV3-19, CHRNA6	4.22188090652723e-05, 3.36787703245996e-16, 1.28804779509494e-10, 7.87919290841631e-15, 2.04671570765809e-05, 5.05746369149396e-07, 0.000423746979959089, 2.89985281042188e-07, 0.00035914866234859, 0.000175669628485986

Continued on next page

Table 6 continued from previous page

No.	Cluster	Number Regulated Genes	Top 10 Genes	P-value Adjusted
22	Morphological.cluster.G21	303	RP11-76C10.3, RP11-76C10.4, KRTAP10-7, IGHV5-51, IGLV4-69, AC053503.6, RP5-978I12.1, IGHV4-31, IGHG4, IGHGP	9.6749476414314e-19, 7.22277185919017e-10, 0.000552140851800452, 1.10661684495602e-10, 4.56868145826252e-10, 0.00377123810801362, 2.02009734312632e-06, 9.86895380739565e-07, 1.37810615870201e-11, 2.41921070698356e-12
23	Morphological.cluster.G22	244	SLC22A6, IGLV7-43, SLC5A5, SLC13A4, SLC22A2, KCNJ13, IGLV3-27, SFTPB, IGKV3D-20, RP11-71N10.1	0, 3.44933694617138e-12, 3.41658043155526e-12, 5.23803314784136e-18, 2.16520649650953e-05, 5.91614998126327e-16, 0.0372210211084008, 0.000515949821074296, 0.000872121186767184, 0.000292839701273585
24	Morphological.cluster.G23	342	NKX2-1, IGKV1-16, LYPD2, IGHV5-51, MMP3, SLURP1, IGLV4-69, DSG3, IVL, TMPRSS11B	3.2335747237144e-111, 2.39326742815555e-27, 3.35831700400342e-06, 7.45881529241203e-09, 0.000319485022549318, 5.269748474214e-06, 3.01659559187416e-05, 0.000218048013960342, 2.54749949545628e-05, 7.00785483606495e-06
25	Morphological.cluster.G24	1280	SPINK7, TGM3, KRT78, MUC21, ADH7, CRNN, KRT4, CLCA4, PADI1, SPRR3	3.2335747237144e-111, 2.39326742815555e-27, 3.35831700400342e-06, 7.45881529241203e-09, 0.000319485022549318, 5.269748474214e-06, 3.01659559187416e-05, 0.000218048013960342, 2.54749949545628e-05, 7.00785483606495e-06
26	Morphological.cluster.G25	148	IGKV1-16, REG3A, RP11-76C10.3, DMBT1, CYP11B1, GP2, RP11-76C10.4, GIF, PGA3, TFF2	1.41827107298095e-18, 6.45330750046715e-24, 5.56654130319563e-19, 6.00293287054109e-19, 4.17583673376366e-09, 2.2318412445259e-19, 1.46802930340726e-19, 2.2540296896357e-21, 5.56274768365823e-17, 1.26907907900601e-18
27	Morphological.cluster.G26	106	IGKV1-16, KRT16P2, RP11-76C10.3, RP11-76C10.4, MT-TL1, RP11-527N22.1, IGLV4-69, C6orf58, MUC7, IGHV4-31	3.62169079648216e-38, 5.74071817952398e-21, 3.61359756333077e-13, 8.63113162448457e-31, 1.71502792713986e-23, 1.39401573629472e-19, 8.97221921144007e-07, 7.29511958052901e-14, 2.57307335374826e-12, 4.00057089381154e-08
28	Morphological.cluster.G27	429	IGKV1-16, DMBT1, ADH7, KRT78, CYP11B1, MUC21, GIF, CRNN, TFF2, TGM3	0.0480710429914863, 1.75756262821229e-47, 7.19563395434006e-13, 2.56898959726473e-22, 1.71270933164594e-15, 1.4332243914311e-22, 2.1953543389439e-07, 3.35334374460368e-06, 7.06852046979117e-08, 0.000389135810013407, 0.000112792085567494
				1.68392170951953e-47, 6.29907557645114e-30, 8.42428556945193e-17, 7.78004116718374e-34, 1.42268938255003e-23, 3.43677403639427e-31, 2.97814507438324e-22, 3.07580730143845e-29, 1.02431769645487e-11, 7.78004116718374e-34

Continued on next page

Table 6 continued from previous page

No.	Cluster	Number Regulated Genes	Top 10 Genes	P-value Adjusted
29	Morphological.cluster.G28	16	DMBT1, CYP11B1, CDH16, SFTPB, AOX2P, C5orf17, ADAMTS8, RP11-848P1.7, RP11-266K4.9, CD209	2.98842726746584e-14, 0.000227852310369974, 0.000227852310369974, 0.00289311050822501, 0.000227852310369974, 0.037761593079279, 0.00346943692242259, 0.0159201627164114, 0.0256672002424397, 0.00166952801593908
30	Morphological.cluster.G29	296	IGKV1-16, KRTAP10-7, KRT85, TTR, RP11-76C10.4, MFRP, DCD, IGHV4-31, IGHGP, APOBEC3H	2.08615379062149e-24, 0.00012168507219495, 3.28842240476628e-05, 3.21559633060964e-12, 0.00144643380312747, 1.99280558661472e-09, 0.0101136130667142, 0.00203315797610793, 3.87382324248738e-05, 4.51927549295466e-10
31	Morphological.cluster.G30	111	MT-TL1, KRT75, BRS3, SCGB1A1, XIRP2, ODAM, CSR3, RP11-567M16.1, IGHV4-31, SLC34A2	1.707084044128e-13, 0.000762251187954354, 3.64976070911259e-11, 0.00438335863644234, 4.3523061265865e-05, 0.0317618504529048, 1.41894047627548e-07, 4.3523061265865e-05, 0.0119831524160675, 0.00243399907591829
32	Morphological.cluster.G31	3724	STATH, KRT78, SPINK7, MUC21, TGM3, CRNN, KRT4, DMBT1, PADI1, CYP11B1	0, 1.09492470112271e-20, 5.35163762206483e-14, 7.42039893622442e-19, 1.65868983185837e-22, 1.99927865595812e-19, 9.67739526676716e-21, 8.66475577813409e-26, 1.28644428736206e-14, 7.74871671175166e-23

4.2 Analyse des Voies Biologiques à Partir des Données de Transcription

Dans cette section, nous avons tenté d'associer les gènes les plus régulés de chaque cluster avec des voies biologiques communes en utilisant l'ensemble de gènes REACTOME fourni. Cette association a été réalisée en comparant nos gènes à un fichier .gmt, en utilisant la bibliothèque fgsea du logiciel R, qui est spécialisée dans l'analyse d'enrichissement de voies géniques.

Le processus d'analyse commence par le prétraitement des données de transcription pour identifier les gènes significativement régulés dans chaque cluster morphologique. Nous avons utilisé la transformation log2 du Fold Change combinée à la valeur p-value pour classer les gènes selon leur importance dans les différences d'expression.

Ensuite, nous avons appliqué fgsea, qui permet d'évaluer l'enrichissement statistique de voies biologiques spécifiques dans une liste de gènes classée. Ce processus utilise les ensembles de gènes définis dans REACTOME pour identifier les voies significativement enrichies en gènes hautement régulés.

Les résultats de cette analyse sont présentés dans le tableau ci-dessous (Table 7), où nous rapportons les clusters morphologiques associés, le nombre de voies biologiques significativement enrichies, et les dix principales voies pour chaque cluster. Ces résultats illustrent comment les gènes régulés dans différents clusters sont impliqués dans des processus biologiques clés, offrant des indices sur les mécanismes moléculaires qui pourraient être perturbés dans les différentes conditions étudiées.

Table 7: Résultats de l'association entre données de transcription et les 'pathways' biologiques identifiés dans le fichier REACTOME.

No.	Cluster	Nbr significant Pathways	Top 10 Pathways
1	Morphological.cluster.G0	7	REACTOME_SR_P_DEPENDENT_COTRANSLATIONAL_PROTEIN_TARGETING_TO_MEMBRANE, REACTOME_NUCLEOTIDE_EXCISION_REPAIR, REACTOME_GPCR_LIGAND_BINDING, REACTOME_CLASS_A_1_RHODOPSIN_LIKE_RECEPTEORS, REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION, REACTOME_G_ALPHA_Q_SIGNALLING_EVENTS, REACTOME_PEPTIDE_LIGAND_BINDING_RECEPTEORS,
2	Morphological.cluster.G1	0	
3	Morphological.cluster.G2	1	REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION
4	Morphological.cluster.G3	0	
5	Morphological.cluster.G4	22	REACTOME_FC_EPSILON_RECECTOR_FCERI_SIGNALING, REACTOME_FCGAMMA_RECECTOR_FCGR_DEPENDENT_PHAGOCYTOSIS, REACTOME_FCERI_MEDIATED_NF_KB_ACTIVATION, REACTOME_SIGNALING_BY_THE_B_CELL_RECECTOR_BCR, REACTOME_LEISHMANIA_INFECTION, REACTOME_ANTI_INFLAMMATORY_RESPONSE_FAVOURING_LEISHMANIA_PARASITE_INFECTION, REACTOME_PARASITE_INFECTION, REACTOME_CELL_SURFACE_INTERACTIONS_AT_THE_VASCULAR_WALL, REACTOME_IMMUNOREGULATORY_INTERACTIONS_BETWEEN_A_LYMPHOID_AND_A_NON_LYMPHOID_CELL, REACTOME_COMPLEMENT CASCADE

Continued on next page

Table 7 continued from previous page

No.	Cluster	Nbr significant Pathways	Top 10 Pathways
6	Morphological.cluster.G5	26	REACTOME_FC_EPSILON_RECECTOR_FCERI_SIGNALING, REACTOME_FCGAMMA_RECECTOR_FCGR_DEPENDENT_PHAGOCYTOSIS, REACTOME_SIGNALING_BY_THE_B_CELL_RECECTOR_BCR, REACTOME_FCERI_MEDIATED_NF_KB_ACTIVATION, REACTOME_PARASITE_INFECTION, REACTOME_CELL_SURFACE_INTERACTIONS_AT_THE_VASCULAR_WALL, REACTOME_BINDING_AND_UPTAKE_OF_LIGANDS_BY_SCAVENGER_RECECTORS, REACTOME_FCGR3A_MEDIATED_IL10_SYNTHESIS, REACTOME_COMPLEMENT CASCADE, REACTOME_FCERI_MEDIATED_MAPK_ACTIVATION
7	Morphological.cluster.G6	15	REACTOME_FCGAMMA_RECECTOR_FCGR_DEPENDENT_PHAGOCYTOSIS, REACTOME_FCGR_ACTIVATION, REACTOME_ROLE_OF_LAT2_NTAL_LAB_ON_CALCIUM_MOBILIZATION, REACTOME_CREATION_OF_C4_AND_C2_ACTIVATORS, REACTOME_SCAVENGING_OF_HEME_FROM_PLASMA, REACTOME_FCERI_MEDIATED_MAPK_ACTIVATION, REACTOME_ANTIGEN_ACTIVATES_B_CELL_RECECTOR_BCR_LEADING_TO_GENERATION_OF_SECOND_MESSAGES, REACTOME_AMINE_LIGAND_BINDING_RECECTORS, REACTOME_CELLULAR_RESPONSE_TO_HEAT_STRESS, REACTOME_PROGRAMMED_CELL_DEATH
8	Morphological.cluster.G7	24	REACTOME_FORMATION_OF_THE_CORNIFIED_ENVELOPE, REACTOME_FCERI_MEDIATED_MAPK_ACTIVATION, REACTOME_PARASITE_INFECTION, REACTOME_KERATINIZATION, REACTOME_FC_EPSILON_RECECTOR_FCERI_SIGNALING, REACTOME_ROLE_OF_PHOSPHOLIPIDS_IN_PHAGOCYTOSIS, REACTOME_ANTIGEN_ACTIVATES_B_CELL_RECECTOR_BCR_LEADING_TO_GENERATION_OF_SECOND_MESSAGES, REACTOME_BINDING_AND_UPTAKE_OF_LIGANDS_BY_SCAVENGER_RECECTORS, REACTOME_ROLE_OF_LAT2_NTAL_LAB_ON_CALCIUM_MOBILIZATION, REACTOME_FCERI_MEDIATED_CA_2_MOBILIZATION
9	Morphological.cluster.G8	0	
10	Morphological.cluster.G9	0	
11	Morphological.cluster.G10	0	
12	Morphological.cluster.G11	3	REACTOME_SIGNAL_REGULATORY_PROTEIN_FAMILY_INTERACTIONS, REACTOME_TOLL_LIKE_RECECTOR_TLR1_TLR2 CASCADE, REACTOME_SIGNALING_BY_NODAL
13	Morphological.cluster.G12	22	REACTOME_PARASITE_INFECTION, REACTOME_FCGR_ACTIVATION, REACTOME_ROLE_OF_PHOSPHOLIPIDS_IN_PHAGOCYTOSIS, REACTOME_FCGAMMA_RECECTOR_FCGR_DEPENDENT_PHAGOCYTOSIS, REACTOME_INITIAL_TRIGGERING_OF_COMPLEMENT, REACTOME_FCGR3A_MEDIATED_IL10_SYNTHESIS, REACTOME_ROLE_OF_LAT2_NTAL_LAB_ON_CALCIUM_MOBILIZATION, REACTOME_FCERI_MEDIATED_MAPK_ACTIVATION, REACTOME_CREATION_OF_C4_AND_C2_ACTIVATORS, REACTOME_ANTIGEN_ACTIVATES_B_CELL_RECECTOR_BCR_LEADING_TO_GENERATION_OF_SECOND_MESSAGES
14	Morphological.cluster.G13	24	REACTOME_FC_EPSILON_RECECTOR_FCERI_SIGNALING, REACTOME_SIGNALING_BY_THE_B_CELL_RECECTOR_BCR, REACTOME_LEISHMANIA_INFECTION, REACTOME_FCERI_MEDIATED_NF_KB_ACTIVATION, REACTOME_FCGAMMA_RECECTOR_FCGR_DEPENDENT_PHAGOCYTOSIS, REACTOME_PARASITE_INFECTION, REACTOME_IMMUNOREGULATORY_INTERACTIONS_BETWEEN_A_LYMPHOID_AND_A_NON_LYMPHOID_CELL, REACTOME_ANTI_INFLAMMATORY_RESPONSE_FAVOURING_LEISHMANIA_PARASITE_INFECTION, REACTOME_COMPLEMENT CASCADE, REACTOME_CELL_SURFACE_INTERACTIONS_AT_THE_VASCULAR_WALL

Continued on next page

Table 7 continued from previous page

No.	Cluster	Nbr significant Pathways	Top 10 Pathways
15	Morphological.cluster.G14	4	REACTOME_CHROMATIN MODIFYING ENZYMES, REACTOME_NEUTROPHIL_DEGRANULATION, REACTOME_KERATINIZATION, REACTOME_FORMATION_OF_THE_CORNIFIED_ENVELOPE
16	Morphological.cluster.G15	41	REACTOME_GASTRIN_CREB_SIGNALLING_PATHWAY_VIA_PKC_AND_MAPK, REACTOME_ESR_MEDIATED_SIGNALING, REACTOME_GPCR_LIGAND_BINDING, REACTOME_RHO_GTPASE_CYCLE, REACTOME_CLASS_A_1_RHODOPSIN_LIKE_RECEPTEORS, REACTOME_RAC1_GTPASE_CYCLE, REACTOME_DISEASES_OF_GLYCOSYLATION, REACTOME_TCR_SIGNALING, REACTOME_INTERFERON_SIGNALING, REACTOME_RHOA_GTPASE_CYCLE
17	Morphological.cluster.G16	3	REACTOME_NGF_STIMULATED_TRANSCRIPTION, REACTOME_SIGNALING_BY_NTRKS, REACTOME_NUCLEAR_EVENTS_KINASE_AND_TRANSCRIPTION_FACTOR_ACTIVATION
18	Morphological.cluster.G17	26	REACTOME_CD22_MEDIATED_BCR_REGULATION, REACTOME_SCAVENGING_OFHEME_PLASMA, REACTOME_FCGR_ACTIVATION, REACTOME_ROLE_OF_LAT2_NTAL_LAB_ON_CALCIUM_MOBILIZATION, REACTOME_FCGAMMA_RECECTOR_FCGR_DEPENDENT_PHAGOCYTOSIS, REACTOME_FCER1_MEDIATED_NF_KB_ACTIVATION, REACTOME_METABOLIC_DISORDERS_OF_BIOLOGICAL_OXIDATION_ENZYMES, REACTOME_PARASITE_INFECTION, REACTOME_FCER1_MEDIATED_MAPK_ACTIVATION, REACTOME_FCGR3A_MEDIATED_IL10_SYNTHESIS
19	Morphological.cluster.G18	2	REACTOME_KERATINIZATION, REACTOME_FORMATION_OF_THE_CORNIFIED_ENVELOP
20	Morphological.cluster.G19	2	REACTOME_FCGAMMA_RECECTOR_FCGR_DEPENDENT_PHAGOCYTOSIS, REACTOME_CLASS_A_1_RHODOPSIN_LIKE_RECEPTEORS
21	Morphological.cluster.G20	21	REACTOME_NON_INTEGRIN_MEMBRANE_ECM_INTERACTIONS, REACTOME_THE_CANONICAL_RETINOID_CYCLE_IN_RODS_TWILIGHT_VISION, REACTOME_PHASE_I_FUNCTIONALIZATION_OF_COMPOUNDS, REACTOME_ARACHIDONIC_ACID_METABOLISM, REACTOME BIOSYNTHESIS_OF_SPECIALIZED_PROSOLVING_MEDIATORS_SPMS, REACTOME_VISUAL_PHOTOTRANSDUCTION, REACTOME_XENOBIOTICS, REACTOME_REGULATION_OF_LIPID_METABOLISM_BY_PPARALPHA, REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION, REACTOME_PROTEIN_LOCALIZATION
22	Morphological.cluster.G21	27	REACTOME_FCGAMMA_RECECTOR_FCGR_DEPENDENT_PHAGOCYTOSIS, REACTOME_LEISHMANIA_INFECTION, REACTOME_SIGNALING_BY_THE_B_CELL_RECEPTOR_BCR, REACTOME_FCER1_MEDIATED_NF_KB_ACTIVATION, REACTOME_FC_EPSILON_RECEPTOR_FCER1_SIGNALING, REACTOME_PARASITE_INFECTION, REACTOME_ANTI_INFLAMMATORY_RESPONSE_FAVOURING_LEISHMANIA _PARASITE_INFECTION, REACTOME_CELL_SURFACE_INTERACTIONS_AT_THE_VASCULAR_WALL, REACTOME_IMMUNOREGULATORY_INTERACTIONS_BETWEEN_A_LYMPHOID_AND_A_NON _LYMPHOID_CELL, REACTOME_FCGR3A_MEDIATED_IL10_SYNTHESIS
23	Morphological.cluster.G22	3	REACTOME_NOREpinephrine_NEUROTRANSMITTER_RELEASE_CYCLE, REACTOME_EXTRACELLULAR_MATRIX_ORGANIZATION, REACTOME_ECM_PROTEOGLYCANs
24	Morphological.cluster.G23	1	REACTOME_CD22_MEDIATED_BCR_REGULATION
25	Morphological.cluster.G24	3	REACTOME_FORMATION_OF_THE_CORNIFIED_ENVELOPE, REACTOME_KERATINIZATION, REACTOME_GPCR_LIGAND_BINDING
26	Morphological.cluster.G25	0	

Continued on next page

Table 7 continued from previous page

No.	Cluster	Nbr significant Pathways	Top 10 Pathways
27	Morphological.cluster.G26	0	
28	Morphological.cluster.G27	22	REACTOME_PARASITE_INFECTION, REACTOME_FC_EPSILON_RECECTOR_FCERI_SIGNALING, REACTOME_FCERI_MEDIATED_NF_KB_ACTIVATION, REACTOME_FCGAMMA_RECECTOR_FCGR_DEPENDENT_PHAGOCYTOSIS, REACTOME_SIGNALING_BY_THE_B_CELL_RECECTOR_BCR, REACTOME_FCGR3A_MEDIATED_IL10_SYNTHESIS, REACTOME_FCERI_MEDIATED_MAPK_ACTIVATION, REACTOME_FCGR_ACTIVATION, REACTOME_FCERI_MEDIATED_CA_2_MOBILIZATION, REACTOME_CD22_MEDIATED_BCR_REGULATION
29	Morphological.cluster.G28	14	REACTOME_SURFACTANT_METABOLISM, REACTOME_ENDOGENOUS_STEROLS, REACTOME_METABOLIC_DISORDERS_OF_BIOLOGICAL_OXIDATION_ENZYMES, REACTOME_DECTIN_2_FAMILY, REACTOME_ATF4_ACTIVATES_GENES_IN_RESPONSE_TO_ENDOPLASMIC_RETICULUM_STRESS, REACTOME_PERK_REGULATES_GENE_EXPRESSION, REACTOME_ACTIVATION_OF_MATRIX_METALLOPROTEINASES, REACTOME_INTERLEUKIN_10_SIGNALING, REACTOME_ANTIMICROBIAL_PEPTIDES, REACTOME_SIGNALING_BY_INTERLEUKINS
30	Morphological.cluster.G29	5	REACTOME_FC_EPSILON_RECECTOR_FCERI_SIGNALING, REACTOME_FCGAMMA_RECECTOR_FCGR_DEPENDENT_PHAGOCYTOSIS, REACTOME_FCERI_MEDIATED_NF_KB_ACTIVATION, REACTOME_PARASITE_INFECTION, REACTOME_BINDING_AND_UPTAKE_OF_LIGANDS_BY_SCAVENGER_RECEPTORS
31	Morphological.cluster.G30	10	REACTOME_FCGR_ACTIVATION, REACTOME_ROLE_OF_LAT2_NTAL_LAB_ON_CALCIUM_MOBILIZATION, REACTOME_FCERI_MEDIATED_MAPK_ACTIVATION, REACTOME_ROLE_OF_PHOSPHOLIPIDS_IN_PHAGOCYTOSIS, REACTOME_CREATION_OF_C4_AND_C2_ACTIVATORS, REACTOME_PARASITE_INFECTION, REACTOME_FCGR3A_MEDIATED_IL10_SYNTHESIS, REACTOME_SCAVENGING_OFHEME_FROM_PLASMA, REACTOME_FC_EPSILON_RECECTOR_FCERI_SIGNALING, REACTOME_NEUTROPHIL_DEGRANULATION
32	Morphological.cluster.G31	5	REACTOME_ENDOGENOUS_STEROLS, REACTOME_FORMATION_OF_THE_CORNIFIED_ENVELOPE, REACTOME_GPCR_LIGAND_BINDING, REACTOME_CLASS_A_1_RHODOPSIN_LIKE_RECEPTEORS, REACTOME_PEPTIDE_LIGAND_BINDING_RECEPTEORS

4.3 Discussion des résultats

Sur l'aspect technique de notre analyse, nous avons commencé par filtrer une grande partie des gènes exprimés en éliminant ceux qui étaient peu ou pas différentiellement exprimés. Au total, près de la moitié des gènes ont été écartés à cette étape. Ensuite, nous avons utilisé une combinaison de log₂ fold change et de p-value ajouté pour éviter les problèmes lors des tests multiples et 'ranker' les gènes les plus significativement exprimés dans chaque cluster morphologique. Nous avons implémenté une fonction R permettant de gérer l'analyse automatiquement pour chaque cluster. Les résultats de cette analyse ont ensuite été exportés en fichiers CSV pour garantir leur pérennité.

Enfin, après avoir harmonisé les noms des gènes pour obtenir un format standard entre les deux fichiers, nous avons utilisé cette liste de gènes pour effectuer une analyse FGSEA (Fast Gene Set Enrichment Analysis). Cette méthode nous permet de mettre en évidence les voies biologiques importantes ayant des liens significatifs avec chaque cluster morphologique. Les paramètres *minSize* et *maxSize* de la fonction *fgsea()* furent respectivement mis à 15 et 500.

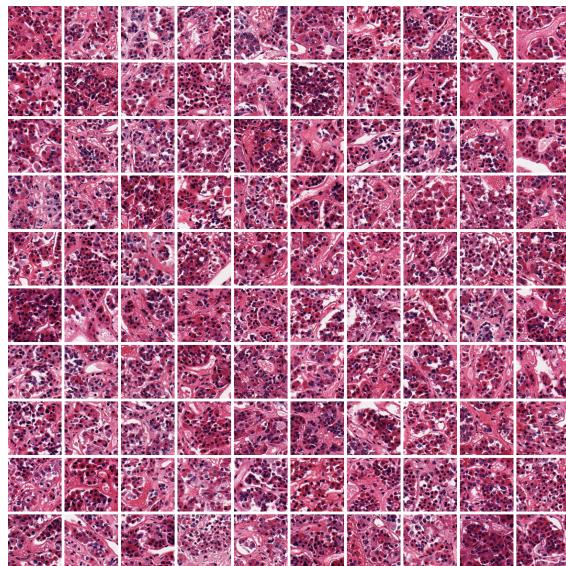
Notre analyse a permis de mettre en évidence une série de gènes dont la transcription est fortement corrélée avec l'analyse morphologique des tissus provenant de la glande pituitaire. Cette analyse morphologique provient d'un algorithme d'apprentissage automatique basé sur une architecture de réseaux de neurones. À partir de ce modèle, nos prédecesseurs ont appliqué un algorithme de "clusterisation" pour grouper chaque morceau (tile) d'une image en fonction de ses caractéristiques morphologiques. Nous nous attendions donc, dans cette analyse, à trouver des clusters dont la signature transcriptionnelle serait caractéristique de certains traits morphologiques propres à ce tissu.

La première constatation est que nous trouvons une très grande variance entre le nombre de gènes régulés par cluster, certains d'entre eux ayant plus de 2000 gènes régulés. Ces clusters sont numérotés 0, 2 et 31. Les images morphologiques associées à ces clusters sont représentées en (Figure: 9). D'après l'analyse FGSEA, ces clusters sont associés, en termes de voies biologiques, à l'organisation de la matrice extracellulaire et aux récepteurs cellulaires sur la membrane externe. À l'opposé, certains clusters n'ont que très peu de gènes régulés, comme les clusters 3, 12 ou 28 (Figure 10) qui sont plus liés au métabolisme ou à la réponse aux stress de la cellule.

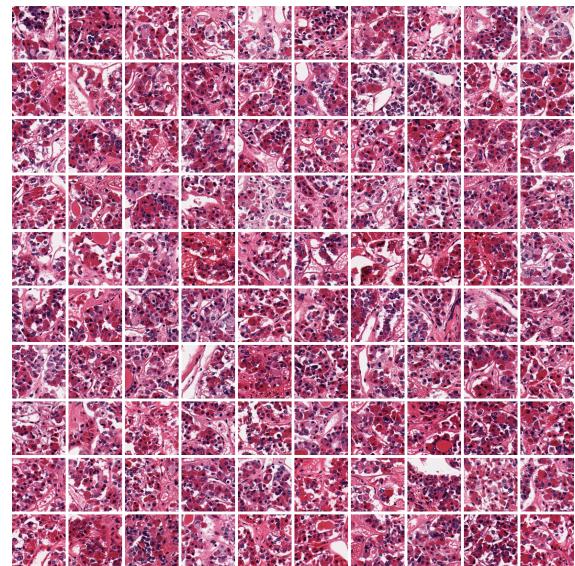
Nous avons également identifié deux clusters ayant une majorité d'images avec de nombreux artefacts ou extrémités de l'organe. Ces clusters, numérotés 21 et 23, régulent respectivement 303 et 342 gènes selon notre analyse. Ces images particulières peuvent être utiles comme contrôles négatifs ou positifs dans la suite de l'analyse (Figure 11).

A partir de ces images il est aussi possible de donner une estimation de la prévalence de l'inflammation dans le morceau de tissus. Les tissus inflammés peuvent être reconnaissable à partir de l'analyse des cellules (neutrophiles, nécroses, fibrose) mais aussi par la densité cellulaire ou la présence d'oedèmes (accumulation de fluides entre les cellules)[4]. Ceci est par exemple très marqué pour le cluster 21 (Figure 12).

Figure 9: Images morphologiques associé au Clusters 0 et 31 avec plus de 2000 gènes régulés (2158 et 3724)

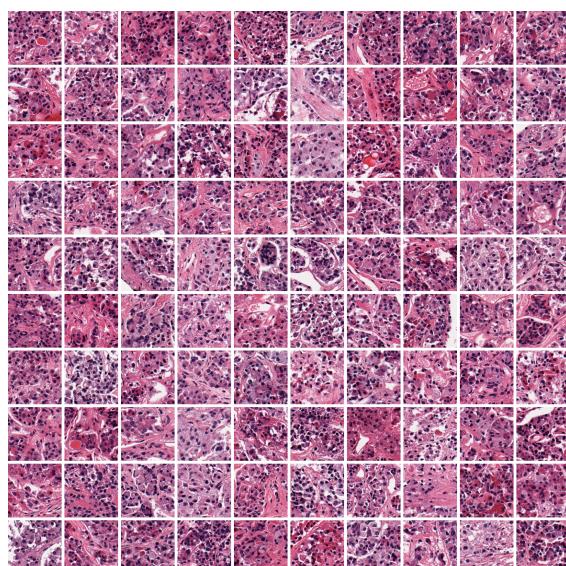


(a) Images morphologiques provenant du Cluster0

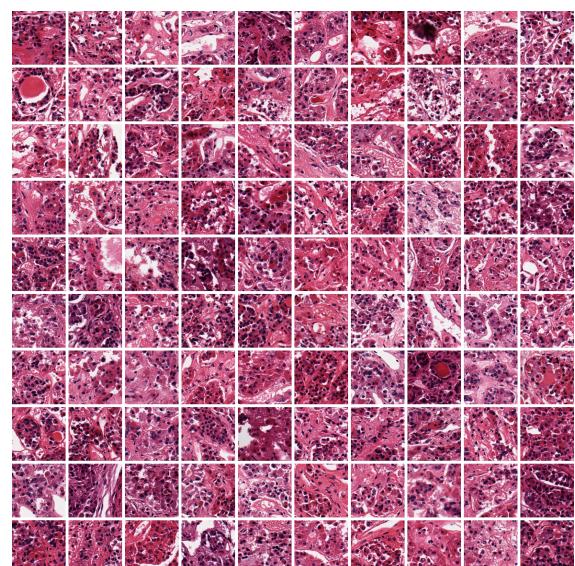


(b) Images morphologiques associé au Cluster31

Figure 10: Images morphologiques provenant des Clusters 3 et 28 avec moins de 100 gènes régulés (64 et 16)



(a) Images morphologiques associé au Cluster3

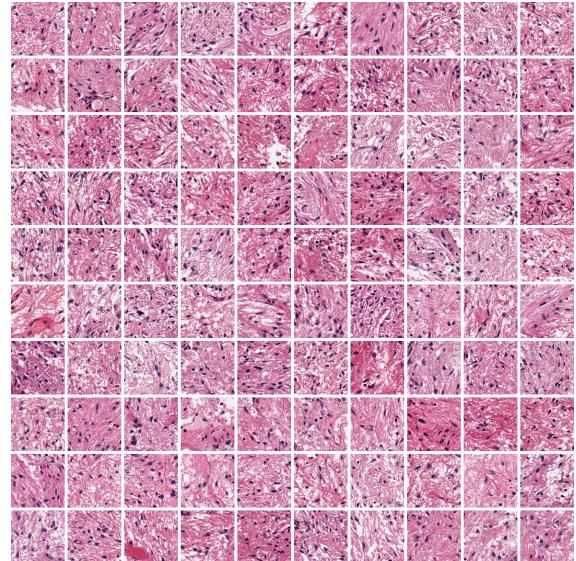


(b) Images morphologiques provenant du Cluster28

Figure 11: Images morphologiques associé aux Clusters 23 et 26, ces images très reconnaissable peuvent être utilisés comme contrôle positif ou négatif

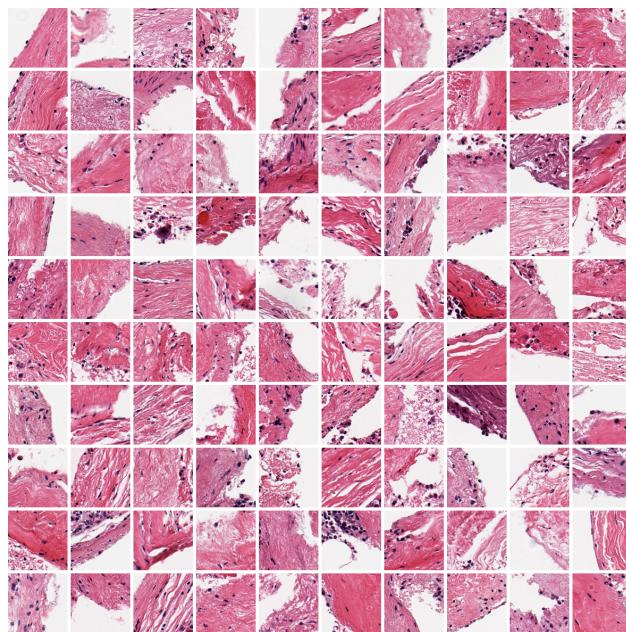


(a) Images morphologiques associé au Cluster23



(b) Images morphologiques associé au Cluster26

Figure 12: Cluster morphologique 21 montrant des signes inflammatoires



5 Figures

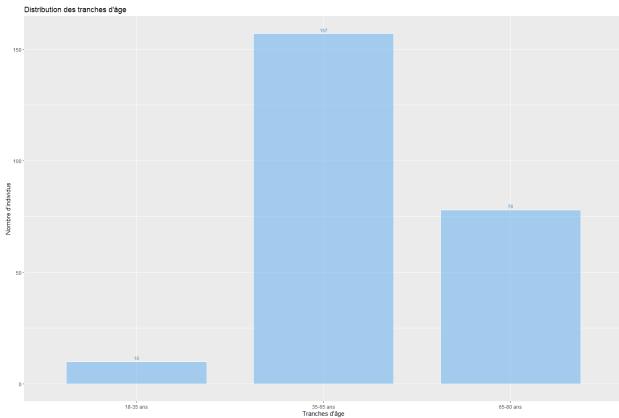


Figure 13: Barplot de la distribution de l'âge

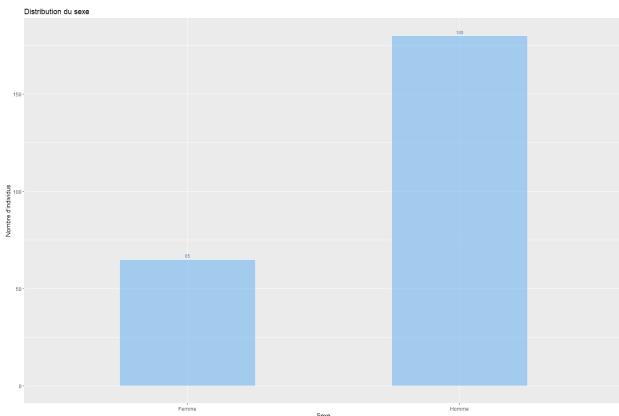


Figure 14: Barplot de la distribution des sexes

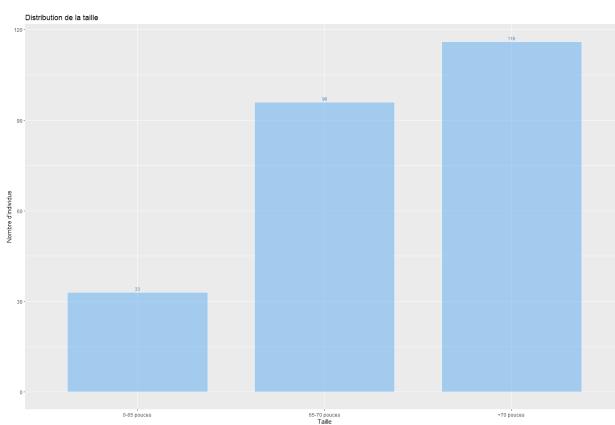


Figure 15: Barplot de la distribution de la taille

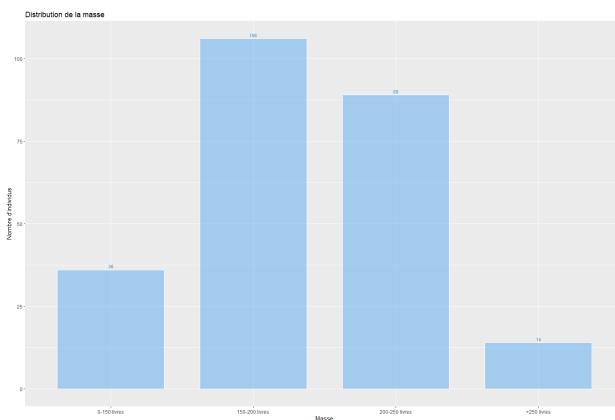


Figure 16: Barplot de la distribution du poids

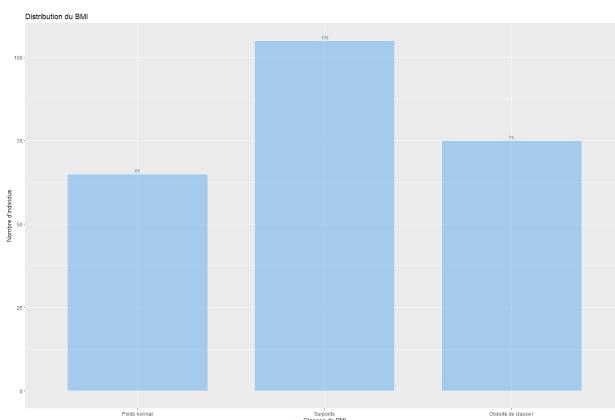


Figure 17: Barplot de la distribution des BMI

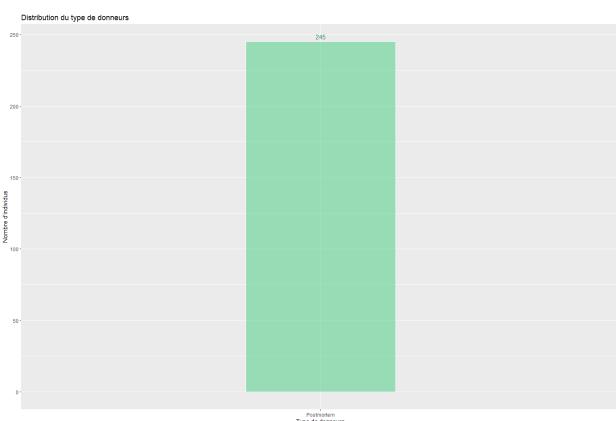


Figure 18: Barplot de la distribution du type de prélèvement

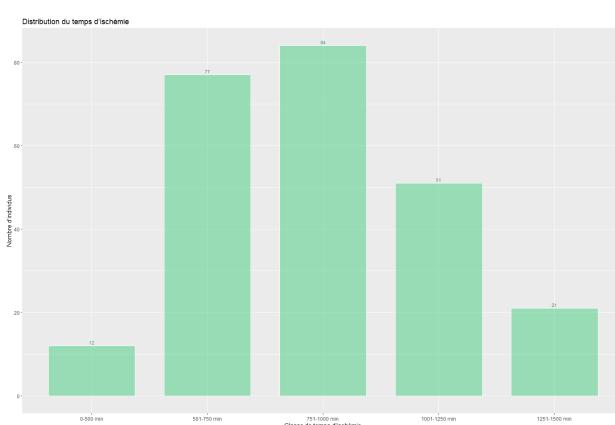


Figure 19: Barplot de la distribution du temps ischémique

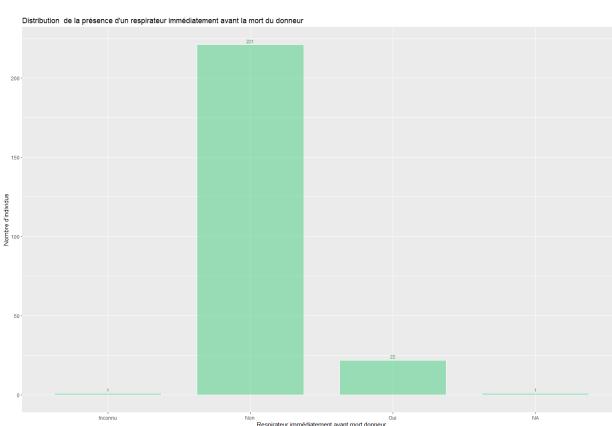


Figure 20: Barplot de la distribution d'individus sous respirateur avant leur décès

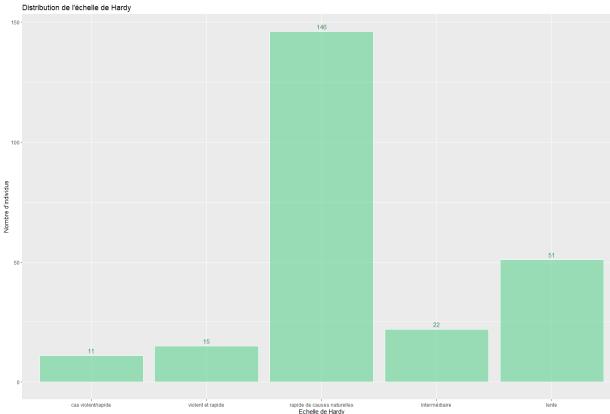


Figure 21: Barplot de la distribution des types de morts

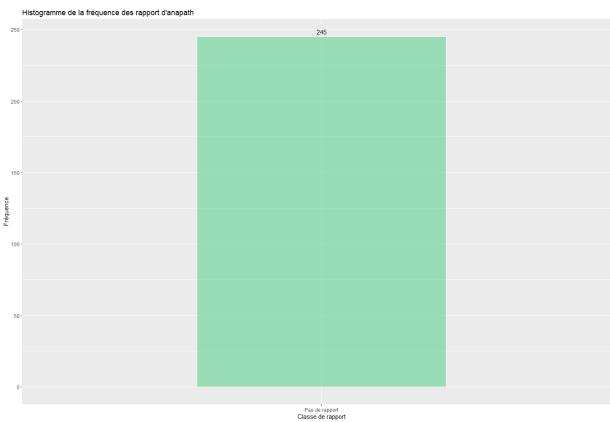


Figure 22: Barplot de la distribution de la présence de rapports d'anapath

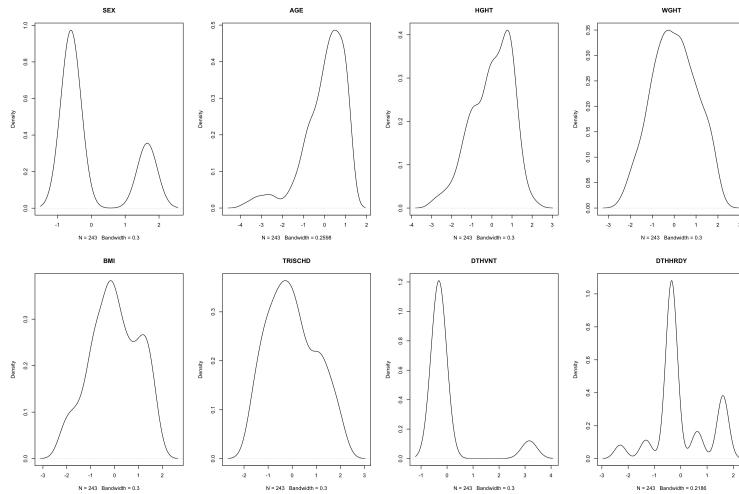
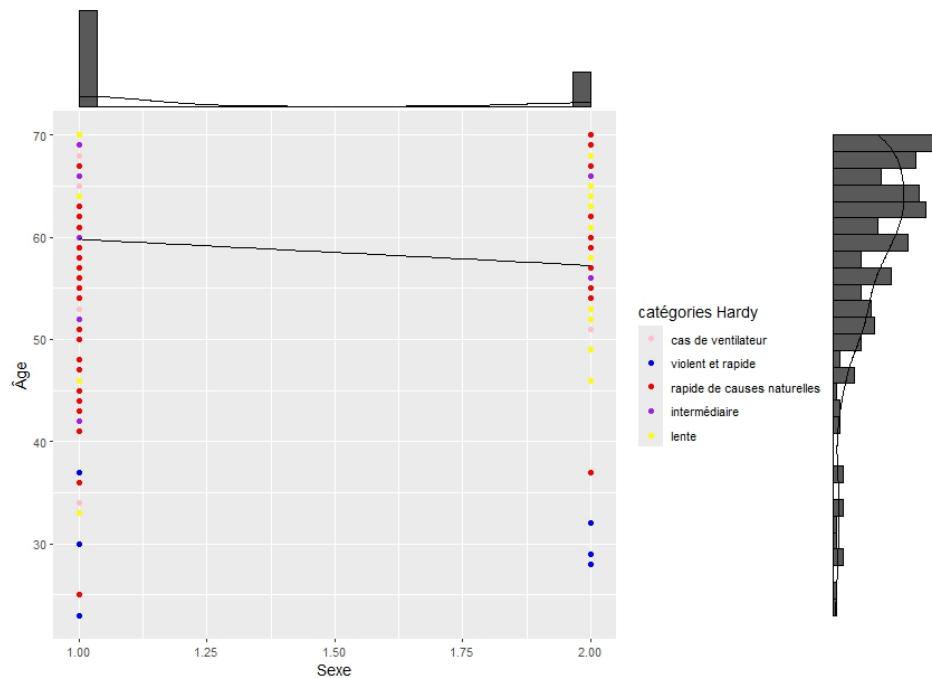


Figure 23: Graphiques de densité des variables en fonction des échantillons normalisé

Distribution de l'âge en fonction du sexe (variable confondante : Hardy)



Division en catégories de Hardy pour la distribution de l'âge en fonction du sexe

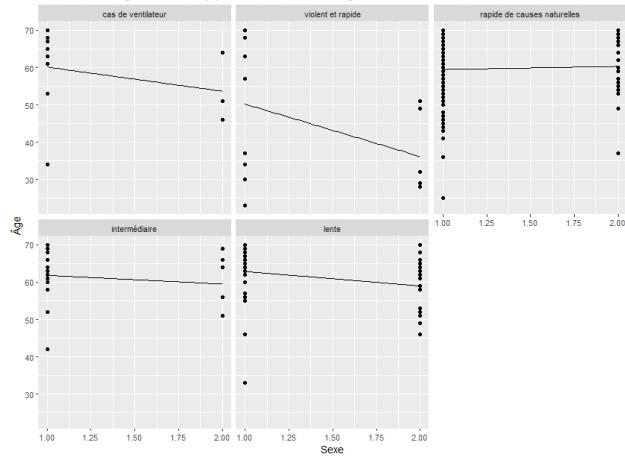


Figure 24: Analyse des variables confondantes pour la relation âge/sexe

Distribution de la taille en fonction de l'âge (variables confondantes : T)

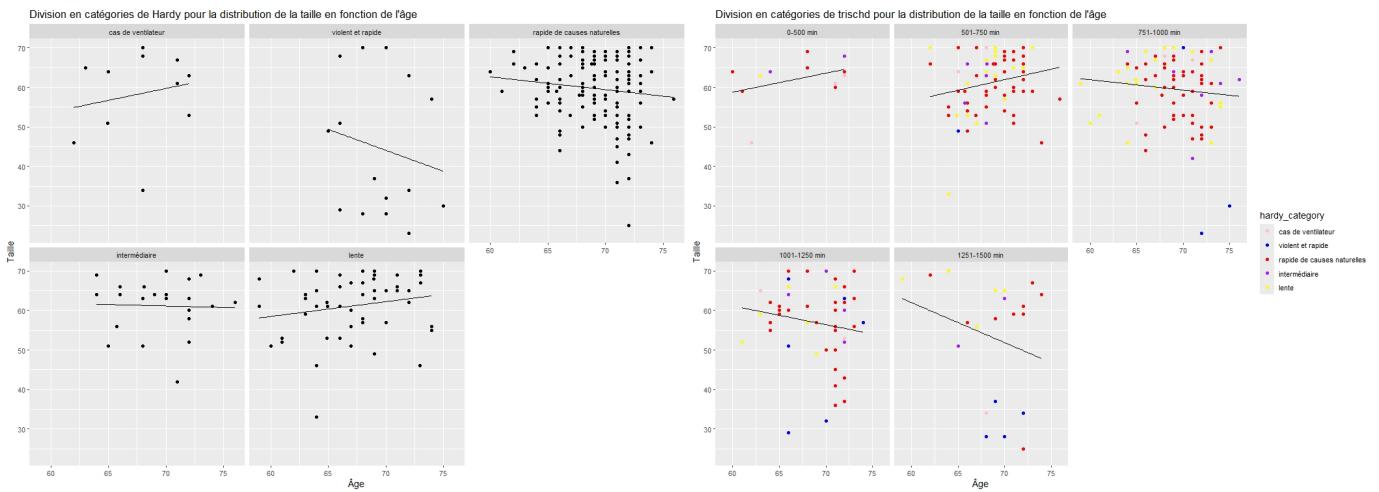
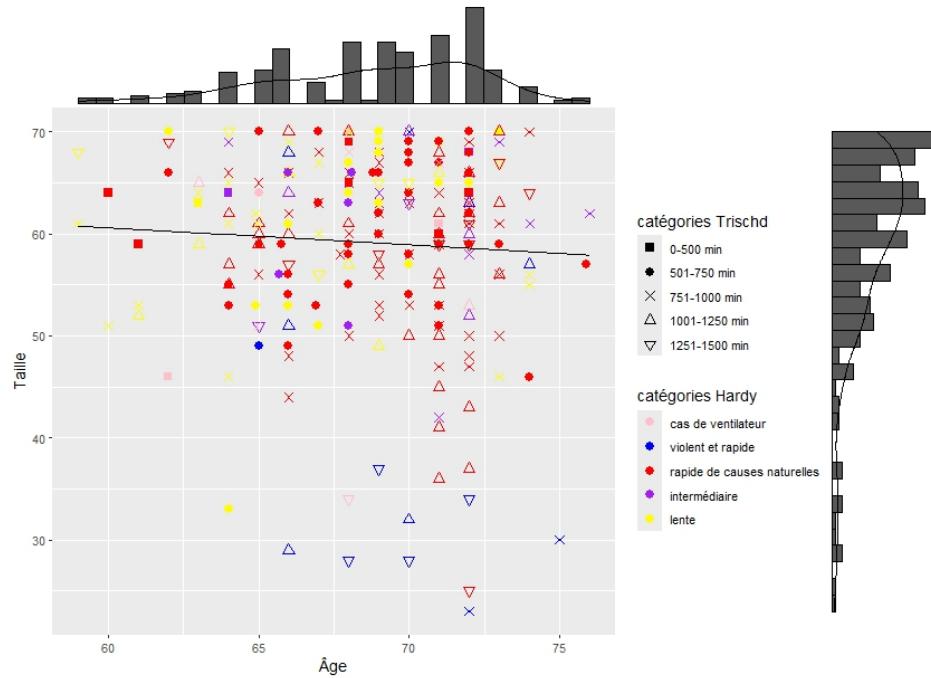


Figure 25: Analyse des variables confondantes pour la relation taille/âge

Distribution de la masse en fonction de l'âge (variables confondantes)

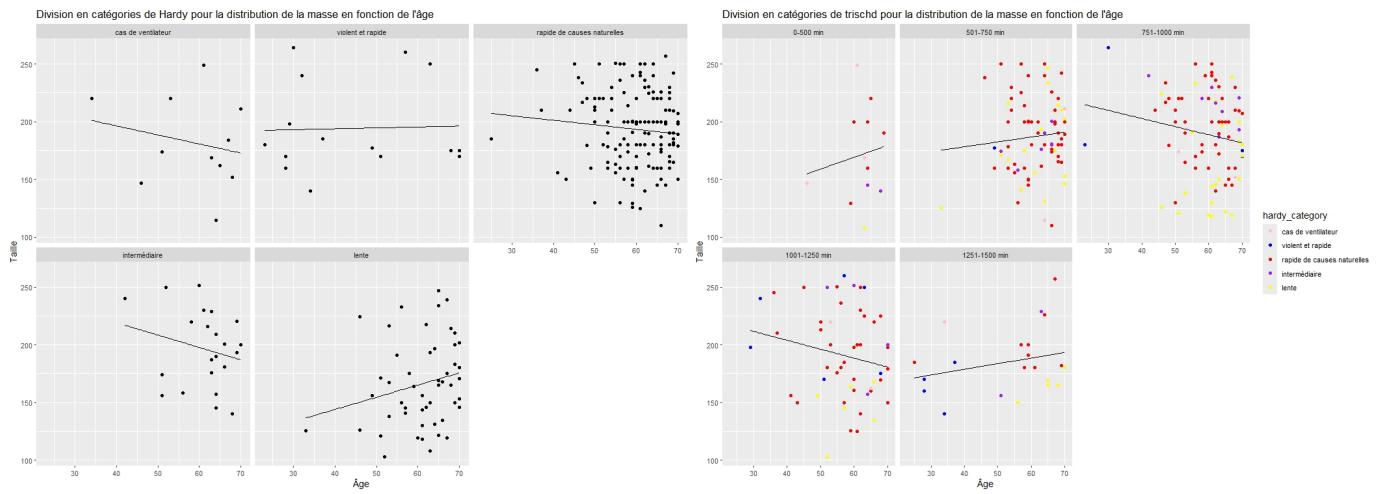
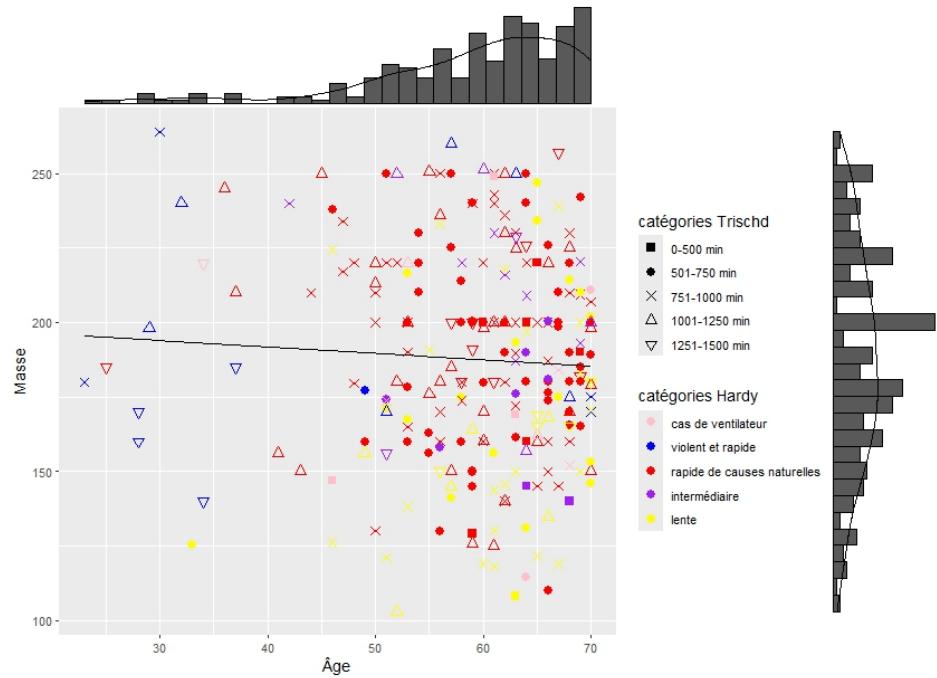


Figure 26: Analyse des variables confondantes pour la relation masse/âge

Distribution du BMI en fonction du sexe (variable confondante : Hardy)

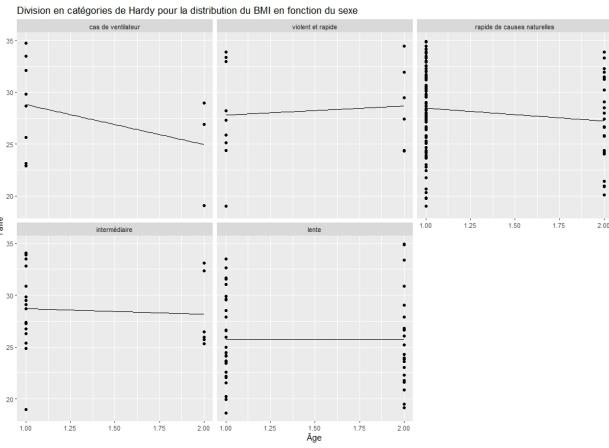
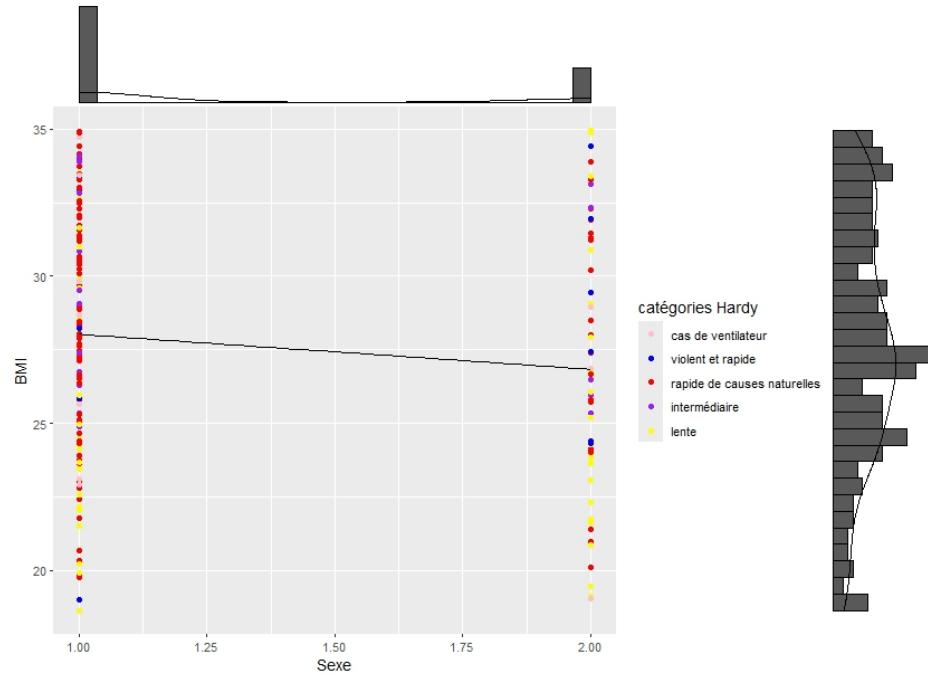


Figure 27: Analyse des variables confondantes pour la relation BMI/sexe

Distribution du BMI en fonction de la taille (variable confondante : Har)

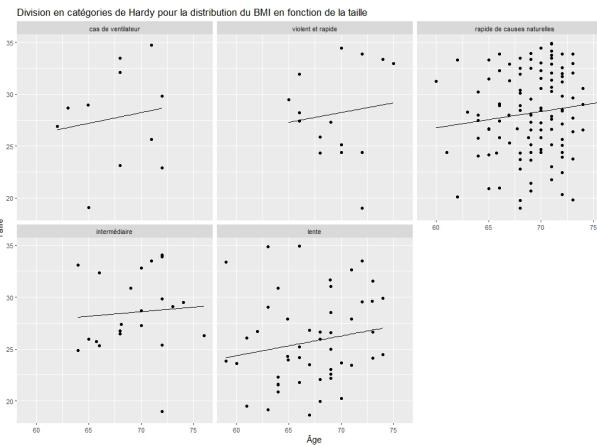
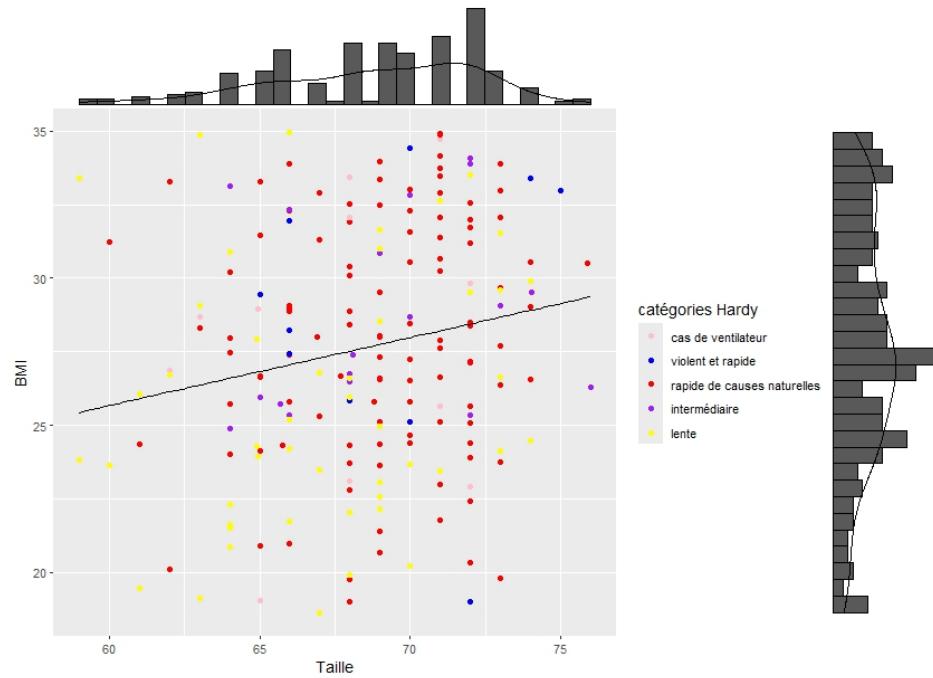


Figure 28: Analyse des variables confondantes pour la relation BMI/taille

Table 8: Valeurs des corrélations ($\geq 0,1$) entre le sexe et les différents clusters morphologiques

cluster	Corrélation
Morphological.cluster.G4_15	0.17
Morphological.cluster.G4_24	0.16
Morphological.cluster.G4_31	0.16
Morphological.cluster.G4_14	0.15
Morphological.cluster.G4_10	0.15
Morphological.cluster.G4_25	0.14
Morphological.cluster.G4_8	0.13
Morphological.cluster.G4_20	0.11
Morphological.cluster.G4_4	-0.10
Morphological.cluster.G4_12	-0.10
Morphological.cluster.G4_13	-0.11
Morphological.cluster.G4_18	-0.13
Morphological.cluster.G4_3	-0.21

Table 9: Valeurs des corrélations ($\geq 0,1$) entre l'âge et les différents clusters morphologiques

cluster	Corrélation
Morphological.cluster.G4_6	0.24
Morphological.cluster.G4_13	0.20
Morphological.cluster.G4_9	0.16
Morphological.cluster.G4_5	0.16
Morphological.cluster.G4_16	0.15
Morphological.cluster.G4_21	0.13 4
Morphological.cluster.G4_1	0.12
Morphological.cluster.G4_4	0.11
Morphological.cluster.G4_20	0.11
Morphological.cluster.G4_28	-0.10
Morphological.cluster.G4_15	-0.11
Morphological.cluster.G4_0	-0.12

Table 10: Valeurs des corrélations ($\geq 0,1$) entre la taille et les différents clusters morphologiques

cluster	Corrélation
Morphological.cluster.G4_3	0.19
Morphological.cluster.G4_13	0.16
Morphological.cluster.G4_12	0.13
Morphological.cluster.G4_29	0.13
Morphological.cluster.G4_4	0.12
Morphological.cluster.G4_21	0.11
Morphological.cluster.G4_6	0.10
Morphological.cluster.G4_22	-0.10
Morphological.cluster.G4_16	-0.13
Morphological.cluster.G4_20	-0.14
Morphological.cluster.G4_31	-0.15
Morphological.cluster.G4_24	-0.22

Table 11: Valeurs des corrélations ($\geq 0,1$) entre la masse et les différents clusters morphologiques

cluster	Corrélation
Morphological.cluster.G4_29	0.27
Morphological.cluster.G4_3	0.21
Morphological.cluster.G4_12	0.15
Morphological.cluster.G4_14	0.12
Morphological.cluster.G4_13	0.11
Morphological.cluster.G4_16	-0.11
Morphological.cluster.G4_22	-0.11
Morphological.cluster.G4_31	-0.16
Morphological.cluster.G4_24	-0.21

Table 12: Valeurs des corrélations ($\geq 0,1$) entre le BMI et les différents clusters morphologiques

cluster	Corrélation
Morphological.cluster.G4_29	0.25
Morphological.cluster.G4_3	0.16
Morphological.cluster.G4_14	0.12
Morphological.cluster.G4_12	0.12
Morphological.cluster.G4_19	0.11
Morphological.cluster.G4_24	-0.13

6 Github

Tout le code qui a été produit pour ce projet est disponible sur https://github.com/DJsimon93/Projet_genomique_401

7 Bibliographie

- [1] Zach Bobbitt. *How to Use cor() to Calculate Correlation Coefficients in R - Statology – statology.org*. [Accessed 24-05-2024].
- [2] Canadian Cancer Society / Société canadienne du cancer. *L'hypophyse – cancer.ca*. Consulté le 23-05-2024. 2024. URL: <https://cancer.ca/fr/cancer-information/cancer-types/pituitary-gland-tumours/what-are-pituitary-gland-tumours/the-pituitary-gland#:~:text=L%27hypophyse%20est%20divis%C3%A9e%20en,l%27arri%C3%A8re%20de%20la%20t%C3%A4te.>
- [3] *Confounding and Effect Measure Modification*. URL: https://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704-ep713_conounding-em/bs704-ep713_conounding-em_print.html (visited on 05/24/2024).
- [4] Dr. Ayser Ahmeed. *Morphology patterns of acute inflammation*. [Accessed 28-05-2024].
- [5] AP-HP. *Les pathologies hypophysaires - Endocrinologie, Diabétologie et Médecine de la reproduction – endocrino-sat.aphp.fr*. Consulté le 20-05-2024. 2024. URL: <https://endocrino-sat.aphp.fr/les-pathologies-en-endocrinologie/les-pathologies-hypophysaires/>.
- [6] Le Makistatheux. *Le coefficient de corrélation et le test de Kendall*. Consulté le 24 mai 2024. 2013.