



Université de Lille
Ecole Doctorale Biologie-Santé
Laboratoire PRISM, U1192 INSERM

THESE DE DOCTORAT

De l'hétérogénéité intra-tumorale à la recherche de vésicules
extracellulaires au sein de biopsie liquide en vue d'une
médecine personnalisée

Présentée et soutenue par

Lauranne Drelich

Le 12 octobre 2020

En vue de l'obtention du grade de Docteur en Sciences

De l'Université de Lille

Devant le jury composé de :

Rapportrice	Christine Carapito	Chargée de Recherche, Université de Strasbourg
Rapportrice	Hélène Castel	Directrice de Recherche (DR2), Université de Rouen
Examineur	Claude-Alain Maurage	Professeur, Université de Lille
Directrice de Thèse	Emilie Le Rhun	Chercheuse, Université de Zurich
Co-encadrant	Maxence Wisztorski	Maître de Conférences, Université de Lille

“Mary stood beside Wilbur, waiting as he sewed Henrietta’s abdomen closed. She wanted to run out of the morgue and back to the lab, but instead, she stared at Henrietta’s arms and legs—anything to avoid looking into her lifeless eyes. Then Mary’s gaze fell on Henrietta’s feet, and she gasped: Henrietta’s toenails were covered in chipped bright red polish. “When I saw those toenails,” Mary told me years later, “I nearly fainted. I thought, Oh jeez, she’s a real person. I started imagining her sitting in her bathroom painting those toenails, and it hit me for the first time that those cells we’d been working with all this time and sending all over the world, they came from a live woman. I’d never thought of it that way.”

Rebecca Skloot, The Immortal Life of Henrietta Lacks

Remerciements

Au terme de ce travail il est essentiel pour moi de remercier toutes les personnes ayant contribuées de près ou de loin à sa réalisation. Merci de m'avoir si bien entouré.

Au professeur Michel Salzet, directeur du laboratoire PRISM, et à la co-directrice, le professeur Isabelle Fournier. Merci infiniment d'avoir cru en moi et de m'avoir offert l'opportunité d'entreprendre ce travail. Merci pour votre détermination.

Au Docteur Emilie Le Rhun, directrice de ma thèse, pour ton aide et ton implication, merci de m'avoir permise de mener à bien ce projet. Au Docteur Maxence Wisztorski, co-encadrant de ma thèse, pour votre aide précieuse, votre disponibilité et vos encouragements. Merci d'avoir su me comprendre et me permettre d'avancer.

À Madame Christine Carapito, chargée de recherche à l'Université de Strasbourg et à Madame Hélène Castel, Directrice de Recherche à l'Université de Rouen, pour avoir accepté d'examiner mon travail.

Au Professeur Claude-Alain Maurage, pour avoir participé à mes travaux et pour votre bienveillance durant ces trois ans.

A l'ensemble des personnels du laboratoire, pour être toujours disponibles et à l'écoute. Notamment à Annie, Jean-Pascal, Christelle et Lucie, pour être toujours prêts à aider.

En particulier, Soulaïmane, je ne saurais compter le nombre de fois où tu as été là pour m'aider. Tu as été indispensable à ce travail, en plus d'avoir été un véritable soutien. Marie, pour ton implication dans le projet depuis mon stage de M2, tu es assurément la personne la plus intelligente que je connaisse. Christophe, pour avoir été mon directeur au commencement. À Irène, Nina, Philippe, Antonella, Tristan, Etienne et Estelle, pour votre amitié bienveillante et pour les fous rires.

À Mélanie et Flore, mes deux collègues et finalement mes grandes amies. Je ne sais pas si on est une entité, une meute, capitaines d'un bateau pirate ou trois hyènes ... ce qui est sûr c'est que parce que vous étiez là je ne me suis jamais sentie seule. Chaque semaine j'ai compris un peu plus que sans vous ça n'aurait pas été possible. Je ne pensais pas rencontrer deux personnes aussi fortes, intelligentes et drôles que vous.

À mes meilleures amies. Marion pour être là depuis toutes ces années, tu es mon pilier chaque jour. Inès, pour tes mots d'encouragement, pour toujours avoir cru en moi.

À Hugo, pour me soutenir au quotidien et dans toutes les nouvelles causes que j'ai envie de défendre. Merci de m'avoir apporté tout l'amour dont j'avais besoin. On s'est rencontré presque tout pile au début de ma thèse, cette expérience aurait été plus difficile sans toi. Un grand merci également à ta famille, pour tous ces bons moments.

A l'ensemble de ma famille, mon plus grand repère et mon soutien le plus total. Merci d'être mon refuge en toute circonstance.

Particulièrement à mes parents. J'ai longtemps répété que vous ne compreniez rien à ce que je faisais depuis bien longtemps. La vérité c'est que je n'ai jamais vraiment su ce que je faisais. Merci infiniment de m'avoir toujours laissé faire.

A mes grands-parents, qui étaient/sont tous les quatre assurément plus intelligents et cultivés que je ne le serai jamais. Je sais la chance que j'ai eu de pouvoir faire les études que vous auriez dû pouvoir faire.

Productions Scientifiques

Publications en cours :

Lauranne Drelich, Marie Duhamel, Maxence Wisztorski, Soulaymane Aboulouard, Jean-Pascal Gimeno, Pierre-Damien Caux, Nina Ogrinc, Patrick Devos, Tristan Cardon, Michael Weller, Fabienne Escande, Fahed Zairi, Claude-Alain Maurage, Isabelle Fournier, Emilie Le Rhun and Michel Salzet. New glioma classification for precise therapeutic decision from spatially-resolved proteomics guided by MALDI-MSI and clinical data integration. (soumis juillet 2020)

Lauranne Drelich, Soulaymane Aboulouard, Julien Franck, Michel Salzet, Isabelle Fournier, Maxence Wisztorski. Expanded Spatially Resolved Proteomics using tissue expansion. (soumis aout 2020)

Communications orales :

Lauranne Drelich, Marie Duhamel, Emilie Le Rhun, Maxence Wisztorski, Christophe Lefebvre, Claude-Alain Maurage, Michel Salzet, Isabelle Fournier. Evaluation of non-supervised MALDI MSI combined with microproteomics for determination of glioblastoma heterogeneity. *European Graduate School of Neuroscience, EURON 2019 23-24 Septembre 2019 – Université du Luxembourg, Luxembourg*

Lauranne Drelich, Marie Duhamel, Emilie Le Rhun, Maxence Wisztorski, Christophe Lefebvre, Fahed Zairi, Nicolas Reyns, Claude-Alain Maurage, Fabienne Escande, Michel Salzet, Isabelle Fournier. Evaluation of non-supervised MALDI MSI combined with microproteomics for determination of glioblastoma heterogeneity - *European Mass Spectrometry Conference, EMSC 2018, 10-14 Mai 2017 - Saarbrücken, Allemagne*

Lauranne Drelich, Marie Duhamel, Emilie Le Rhun, Maxence Wisztorski, Christophe Lefebvre, Fahed Zairi, Nicolas Reyns, Claude-Alain Maurage, Fabienne Escande, Michel Salzet, Isabelle Fournier. Evaluation of non-supervised MALDI MSI combined with microproteomics

for determination of glioblastoma heterogeneity. Club Jeune de la Société Française de Spectrométrie de Masse (SFSM) 24-30 Mars 2017 - Club Jeune SFSM – Montpellier, France

Posters :

Lauranne Drelich, Nicolas-Xavier Bonne, Marie Duhamel, Dominique Lallemand, Michel Salzet, Isabelle Fournier. Characterization of vestibular schwannomas in NF2 using MALDI mass spectrometry imaging combined with spatially resolved proteomic. Global Neurofibromatosis, 2-6 Novembre 2018 – Paris, France

Participation à l'encadrement de stagiaires :

- Pierre-Gilles Ferron, 2019 – Stage de 2ème année de Master, 6 mois

« Développement d'une méthode de microprotéomique basée sur l'expansion en gel de coupes de tissus FFPE et frais congelés »

- Sylvain Osien, 2017 – Stage de 1ère année de Master, 6 semaines

« Enrichissement en vésicules extracellulaires à partir d'un cerveau de rat en vue de nouvelles perspectives diagnostiques pour les patients atteints de glioblastomes »

Résumé

Les gliomes représentent 80% des tumeurs cérébrales malignes primitives et sont classés selon différents grades de malignité. Les glioblastomes, groupe le plus agressif, représentent plus de la moitié de tous les gliomes. Il s'agit de tumeurs dont la composition est hétérogène, présentant notamment de zones de nécrose, de prolifération vasculaire, de prolifération cellulaire. La survie des patients peut aller de quelques mois à quelques années après la chirurgie et la chimiothérapie. L'imagerie par spectrométrie de masse MALDI est une technique intéressante pour l'étude de ces tumeurs, car elle permet de prendre en considération l'hétérogénéité intra-tumorale. Dans cette étude, l'imagerie MALDI MS est couplée à la microprotéomique localisée sur tissu dans l'objectif d'identifier des sous-groupes de glioblastomes afin d'aider au diagnostic et au pronostic.

Les images moléculaires ont été générées à partir de coupes fines de tissus dans le but de déterminer la localisation de peptides digérés. Ensuite, des analyses statistiques non supervisées nous ont permis de générer un regroupement hiérarchique de régions moléculaires homogènes au sein du tissu. Une analyse restreinte à ces différentes régions à nous a donné accès à leur contenu protéique. Le regroupement hiérarchique a révélé trois régions moléculaires dans l'ensemble des échantillons. À partir de ces régions, 3 groupes de patients ont été identifiés : le groupe A (13/50 patients), le groupe B (9/50 patients) et le groupe C (23/50 patients) . Cinq patients sont restés non classables. Chaque groupe présentait des signatures moléculaires spécifiques. Les analyses microprotéomiques ont montré un panel de protéines spécifiquement surexprimées dans chaque groupe : les protéines surexprimées dans le groupe A sont associées à la neurogenèse ; celles du groupe B sont associées à une activation du système immunitaire et celles du groupe C sont impliquées dans des processus viraux. Enfin, nous avons identifié 6 nouveaux marqueurs pronostiques des glioblastomes qui pourraient aider à stratifier les patients et aider à la décision thérapeutique.

En parallèle, nous nous sommes intéressés au développement d'une nouvelle technique de microprotéomique localisée basée sur l'expansion de tissu. Agrandir le tissu permet, tout en conservant les méthodes de protéomique conventionnelle, d'analyser des régions plus petites que celles atteignables actuellement. Jusqu'à 655 protéines ont pu être identifiées dans une région de 460 μm de diamètre, ce qui correspond à une moyenne de 940 cellules. De plus, cette stratégie est intéressante pour l'imagerie MS, car elle donne la possibilité de cartographier facilement un grand nombre de protéines sur la base de leur quantification avec une résolution spatiale inférieure à 350 μm . En ce sens, l'expansion de tissu est d'un grand intérêt pour augmenter la résolution spatiale de la protéomique localisée.

Finalement, étudier les vésicules extracellulaires circulant dans le sang pourrait à terme permettre d'aider au dépistage, au diagnostic ainsi qu'au suivi des glioblastomes. L'analyse protéomique des vésicules du sang des patients pourrait permettre d'identifier des protéines ou des réseaux déjà identifiés dans les échantillons tumoraux. Les résultats préliminaires à partir de plasma témoin montrent qu'il est possible d'isoler des vésicules et de réaliser des analyses protéomiques. Ainsi, il serait intéressant de pouvoir associer un profil de vésicule extracellulaire pour chaque groupe de tumeurs. De plus, nous avons été capables d'isoler spécifiquement les vésicules contenues dans une coupe fine de cerveau de rat. Ces résultats sont encourageants et à terme il sera possible d'identifier les vésicules provenant de la tumeur.

Summary

Gliomas account for 80% of all malignant brain tumors and are classified within different malignity grades. Glioblastomas, the most aggressive group, represent more than half of all gliomas but remain a heterogeneous group. Indeed, patient survival ranges from several months to a few years after surgery and chemotherapy. To study gliomas, MALDI mass spectrometry imaging is a technique of choice, allowing for tumor heterogeneity analysis. In this study, MALDI-MSI is coupled with spatially-resolved microproteomic aiming at identifying subgroups of glioblastomas patients in order to help for diagnosis and prognosis. Molecular images are generated from thin tissue sections to determine digested peptides spatial localizations. Based on unsupervised statistical analysis, we generated hierarchical clustering of homogeneous molecular regions. According to these regions, spatially resolved proteomic provided a broad range of protein identification and their relative quantifications.

The hierarchical clustering reveals three molecular regions within all the tumor samples. Based on these regions, three groups of patients can be determined: group A (13/50 patients), group B (9/50 patients) and group C (23/50 patients) and 5 non-classifiable patients, each group with specific molecular signatures. Microproteomic analyzes show a panel of proteins specifically overexpressed in each group: proteins overexpressed in group A are associated with neurogenesis; those of group B are linked to immune system activation; and those of group C are involved in viral processes. Finally, we have identified 6 new prognosis markers for glioblastomas that could help stratifying patients and orient clinician in choosing therapeutics.

In parallel, we were interested in developing a novel spatially resolved microproteomic technique based on tissue expansion. Tissue enlargement allows, while maintaining conventional proteomics methods, to analyze regions smaller than those currently achievable. Up to 655 proteins were identified within a region of 460 μm diameter, which corresponds to an average of 940 cells. In addition, this strategy is relevant for MS imaging as it gives the possibility to easily map a large number of proteins based on their quantification within a spatial resolution of less than 350 μm . Thus, tissue expansion is of great interest in increasing the spatial resolution of localized proteomics.

Finally, studying circulating extracellular vesicles in blood may ultimately lead to an earlier diagnosis of glioblastomas. Proteomic analysis of patients' blood vesicles can identify proteins or networks already identified in tumor mass. Preliminary results from control plasma show that it is possible to isolate and perform proteomic analyzes on circulating vesicles. Thus, it would be interesting to be able to associate an extracellular vesicle profile for each group of tumors. In addition, we have been able to specifically isolate the vesicles embedded in a thin section of rat brain. These results are encouraging and, in the long term, will make it possible to identify vesicles originating from the tumor.

Table des matières

Remerciements	I
Productions Scientifiques	III
Résumé	V
Summary	VI
Table des matières	VII
Liste des figures	XI
Liste des tableaux	XIV
Liste des abréviations	XV
Avant-propos	XVI
ETAT DE L'ART	1
I. Les gliomes	1
A. Généralités	1
1. Épidémiologie	1
2. Diagnostic	3
3. Pronostic	4
4. Traitements	5
B. Classification des gliomes	8
1. Classification des tumeurs cérébrales	8
2. Classification histologique 2007	9
3. Classification moléculaire 2016	10
4. Classification des glioblastomes	14
C. Protéomique des gliomes	17
1. Marqueurs protéiques identifiés dans l'étude des échantillons tumoraux 17	
2. Études sur les fluides biologiques	18
II. Implication de la spectrométrie de masse pour l'analyse des cancers	21
A. La spectrométrie de masse	21
B. Imagerie par spectrométrie de masse MALDI	22
1. Principe	22
2. Traitements des données	26
3. Intérêt pour l'analyse de tissus cancéreux	29
C. Protéomique localisée	33
1. Capture par microdissection laser	33

2.	Digestion <i>in situ</i> et extraction par microjonction liquide (<i>Liquid Extraction Surface Analysis</i> LESA)	34
3.	Microdissection assistée par parafilm	35
4.	Digestion des protéines par migration passive dans un hydrogel	36
III.	Les vésicules extracellulaires (VEs)	38
A.	Généralités	38
1.	Les différents types de vésicules extracellulaires.....	38
2.	Rôles dans les cancers	41
B.	Intérêts des vésicules extracellulaires dans la prise en charge des cancers 42	
1.	Diagnostic	42
2.	Évaluation de la réponse aux traitements.....	43
3.	Stratégies thérapeutiques	43
	Objectifs des travaux de Thèse	45
	CHAPITRE 1 Utilisation de la protéomique localisée guidée par imagerie MALDI pour la stratification des patients atteints de glioblastome	47
I.	Introduction.....	47
II.	Matériel et Méthodes.....	52
A.	Recrutement des patients.....	52
B.	Immunohistochimie (IHC)	52
C.	Extraction d'ADN et quantification	53
D.	Imagerie par spectrométrie de masse MALDI	54
1.	Préparation du tissu et acquisition de l'image	54
2.	Traitement et analyses des images	55
E.	Analyses SpiderMass.....	56
1.	Acquisition des spectres	56
2.	Construction du modèle de classification	58
F.	Protéomique localisée sur tissus	58
1.	Digestion <i>in situ</i>	58
2.	Extraction par microjonction liquide (LESA)	59
3.	Analyses NanoLC-MS/MS.....	59
4.	Analyse des données	60
G.	Analyses biologiques.....	61

1.	Analyses des voies de signalisation	61
2.	Comparaison avec <i>Human Pathology Atlas</i>	61
3.	Identification de protéines issues de cadres de lectures alternatifs.....	61
H.	Analyses statistiques	62
I.	Confirmation par immunofluorescence (IF)	62
III.	Résultats et Discussion	64
A.	Caractéristiques cliniques	64
B.	L'imagerie par spectrométrie de masse MALDI permet de regrouper les patients en fonction de leurs caractéristiques moléculaires	69
C.	L'analyse SpiderMass confirme le groupement obtenu par imagerie	73
D.	Caractéristiques cliniques des groupes identifiés.....	75
E.	Identification de voies de signalisation spécifiques de chaque groupe	78
F.	Identification de protéines alternatives	89
G.	Corrélation entre TCGA et données protéomiques.....	92
H.	Corrélation avec les données cliniques.....	94
IV.	Conclusion et Perspectives	101
 CHAPITRE 2 Développement d'une nouvelle approche de microprotéomique localisée basée sur l'expansion de tissu		
		105
I.	Introduction.....	105
II.	Matériel et Méthodes.....	112
A.	Section des tissus	112
B.	Expansion de tissu.....	112
1.	Gélification des tissus	112
2.	Homogénéisation mécanique.....	112
3.	Expansion et coupe.....	113
C.	Analyse protéomique.....	113
D.	Digestion <i>in situ</i> et extraction par microjonction liquide	114
E.	Analyse NanoLC-MS et MS / MS	114
F.	Analyse des données.....	115
III.	Résultats et Discussion	117
A.	L'homogénéisation à la protéinase K permet d'identifier peu de protéines et implique la migration des protéines dans le gel	117
B.	Analyse comparative de différentes alternatives pour la protéinase K ...	118

C.	Optimisation et reproductibilité du protocole d'homogénéisation au SDS pour l'analyse microprotéomique.....	121
D.	Utilisation d'un emporte-pièce pour biopsie pour l'échantillonnage de petites régions.....	126
E.	Comparaison avec la digestion <i>in situ</i> et microjonction liquide.....	128
F.	« <i>Imaging-like</i> » : Profil d'expression protéique basé sur la quantification après expansion.....	131
IV.	Conclusion et Perspectives	134
CHAPITRE 3 Vers un suivi des patients atteints de glioblastome par biopsies liquides grâce à la caractérisation des vésicules extracellulaires du sang.....		
I.	Introduction.....	142
II.	Matériel et Méthodes.....	146
A.	Enrichissement en vésicules extracellulaires à partir de sang humain	146
B.	Enrichissement en vésicules extracellulaires à partir d'un cerveau de rat	146
C.	Analyse NTA (<i>Nanoparticle Tracking Analysis</i>)	147
D.	Analyse protéomique.....	148
E.	Analyses NanoLC-MS/MS.....	148
F.	Analyses des données	148
III.	Résultats et Discussion	150
A.	Enrichissement à partir d'échantillons sanguins	150
B.	Analyse protéomique des vésicules du sang	151
C.	Comparaison aux protéines identifiées dans la littérature	154
D.	Enrichissement à partir de tissus	155
IV.	Conclusion et Perspectives	157
CONCLUSION et PERSPECTIVES.....		161
BIBLIOGRAPHIE.....		166
ANNEXES.....		185

Liste des figures

Figure 0-1 : Distribution des gliomes du système nerveux central par sous-types histologiques.	1
Figure 0-2 Principe général de l'imagerie MALDI.....	23
Figure 0-3 Evolution de la résolution spatiale des images MALDI.....	26
Figure 0-4 Simplification de spectres de masses par représentation dans un espace multidimensionnel	27
Figure 0-5 Principe de la segmentation <i>bisecting k mean</i> et représentation sur un tissu	29
Figure 0-6 Segmentation spatiale des régions musculaires traumatisées (tm), adjacentes au traumatisme (tam) et saines (hm) pour chaque grat biologique.....	30
Figure 0-7 Distribution du paclitaxel par imagerie MALDI (panneau supérieur) et coloration H&E d'une coupe consécutive (panneau inférieur)	32
Figure 0-8 Workflow schématique de la microdissection laser suivi de la collecte et du transfert de cellules dans les nano puits.....	33
Figure 0-9 Principe de l'extraction par micro-jonction liquide	35
Figure 0-10 Microdissection assistée parafilm.....	36
Figure 0-11 Principe de la digestion des protéines par migration passive dans un hydrogel .	37
Figure 0-12 Représentation schématique de la biogenèse et de la composition des trois principales classes de vésicules extracellulaires	39
Figure 1-1 Classification des gliomes de grade III basée sur l'imagerie MALDI et la microprotéomique	49
Figure 1-2 Voies de signalisations associées aux protéines surexprimées dans les 3 groupes identifiées par imagerie MALDI.....	50
Figure 1-3 Schéma de la configuration du SpiderMass et déroulement d'une expérience standard.	57
Figure 1-4 Annotations histologiques des échantillons de tumeurs de l'étude Gliomic.	64
Figure 1-5 Courbes de survie en fonction des différents facteurs pronostics.....	68
Figure 1-6 Les annotations histologiques ne révèlent pas la complexité moléculaire des échantillons.	69
Figure 1-7 Segmentation globale et regroupement des patients.....	71
Figure 1-8 Ions caractéristiques de chaque groupe	72

Figure 1-9 Modèle de classification PCA-LDA basé sur les 3 groupes de gliomes	74
Figure 1-10 Annotations histologiques des échantillons tumoraux par groupe	75
Figure 1-11 Correspondance entre la classification obtenue suite aux analyses d'imagerie MALDI-MSI et les analyses de microprotéomique.....	79
Figure 1-12 Voies de signalisations surexprimées dans le groupe A (cluster 2)	82
Figure 1-13 Représentation de l'infiltration microgliale au sein du groupe B.....	84
Figure 1-14 Voies de signalisations identifiées dans le groupe B	86
Figure 1-15 Voies de signalisation identifiées dans le groupe C.....	88
Figure 1-16 L'analyse des protéines alternatives met en lumière des clusters de protéines surexprimées dans les trois groupes déterminés par imagerie.....	90
Figure 1-17 Courbes de survie en fonction des groupes A, B et C.....	94
Figure 1-18 Regroupement basé sur les 6 protéines les plus significatives de la survie	95
Figure 1-19 Courbe de survie des 3 groupes identifiées sur la base des 6 protéines.....	96
Figure 1-20 Expression des protéines significatives de la survie dans les différents groupes de patients.....	97
Figure 1-21 Immunohistofluorescence des 6 marqueurs de survie identifiés	98
Figure 1-22 Heatmap des protéines alternatives montrant différents profils de régulation dans les trois groupes de survie	99
Figure 1-23 Modèle de classification PCA-LDA basé sur 3 groupes de survie	100
Figure 1-24 Comparaison des segmentations spatiales avec et sans les patients présentant une mutation IDH.....	102
Figure 2-1 Concept de la microscopie par expansion	107
Figure 2-2: Les différentes étapes de la microscopie par expansion.....	108
Figure 2-3 Microscopie par expansion adapté à l'analyse microprotéomique	109
Figure 2-4 Diagramme de Venn des résultats d'identifications sur 3 régions grâce au kit commercial.....	110
Figure 2-5 : Diagrammes de Venn sur les données d'identification en fonction de la concentration en Protéinase K.....	111
Figure 2-6 Nombre d'identification après homogénéisation à la protéinase K.....	117
Figure 2-7 Analyse comparative de différentes alternatives à la protéinase K.....	119
Figure 2-8 Analyse de la reproductibilité protéomique avec le protocole d'homogénéisation utilisant le SDS.....	123

Figure 2-9 Comparaison des réplicats deux à deux	124
Figure 2-10 Variation des tailles de découpe de gel avec un emporte-pièce 1,5 mm	126
Figure 2-11 Analyse de la reproductibilité de l'échantillonnage par emporte-pièce	127
Figure 2-12 Utilisation d'un emporte-pièce : Comparaison des réplicats deux à deux.....	127
Figure 2-13 Comparaison des protéines identifiées après expansion et avec digestion in situ et extraction par micro jonction liquide.....	129
Figure 2-14 Comparaison qualitative des protéines identifiées après expansion et avec digestion in situ et extraction LESA.....	130
Figure 2-15 Reconstitution d'images basées sur les profils d'expression de différentes protéines après expansion de tissu.....	132
Figure 2-16 Coupe de cerveaux frais après expansion et marquage au Hoescht observée sous lampe ultraviolet	136
Figure 2-17 Principe de la microscopie par Expansion itérative (iExM).....	138
Figure 2-18 Cerveau de rat avant et après Expansion x10.....	139
Figure 3-1 Analyse des vésicules extracellulaires par NTA (<i>Nanoparticle Tracking Analysis</i>)	147
Figure 3-2 Analyses NTA de vésicules extracellulaires issus de plasma.....	151
Figure 3-3 Nombre de protéines identifiées selon les deux méthodes d'enrichissement	152
Figure 3-4 Comparaison des protéines identifiées dans l'enrichissement en vésicules et dans le contrôle pour chaque méthode	153
Figure 3-5 Comparaison des protéines surexprimées dans les échantillons de vésicules avec les protéines communément identifiées dans la littérature	154
Figure 3-6 Caractérisation NTA des nanoparticules enrichies à partir de cerveau de rat.....	155
Figure 3-7 Analyse NTA des nanoparticules enrichies à partir de coupes fines cerveau de rat	156
Figure 3-8 Enrichissement et collection de VEs par chromatographie d'exclusion stérique.	158

Liste des tableaux

Tableau 0-1 Taux de survie à 1 an, 2 ans, 5 ans et 10 ans de différents types de gliomes	4
Tableau 0-2 Classification OMS 2007 des gliomes.....	9
Tableau 0-3 Classification OMS 2016 des gliomes.....	11
Tableau 0-4 Recommandations du groupe cIMPACT NOW.....	14
Tableau 1-1 Caractéristiques cliniques	65
Tableau 1-2 Caractéristiques cliniques des patients en fonction des groupes identifiés par imagerie MALDI.....	76
Tableau 1-3 Nombres d'extractions réalisées dans chacun des groupes	78
Tableau 1-4 Protéines pronostics identifiées dans les groupes A, B et C	93
Tableau 1-5 Protéines significatives de la survie des patients utilisées pour le regroupement des patients	95
Tableau 2-1 Facteur d'expansion et nombre de protéines identifiées selon l'agent d'homogénéisation	120
Tableau 2-2 Facteur d'expansion et nombre de protéines identifiées après homogénéisation au SDS dans chaque triplicata	123
Tableau 2-3 Facteur d'expansion et nombre de protéines identifiées à l'aide d'un emporte-pièce pour biopsie dans chaque triplicata	126
Tableau 2-4 Facteur d'expansion et nombre de protéines identifiées dans chaque triplicata à partir de coupes congelées	136
Tableau 3-1 Concentrations et tailles des particules analysées à partir d'échantillons de sang	150
Tableau 3-2 Concentrations et tailles des particules analysées à partir de tissus.....	155

Liste des abréviations

ADN	Acide désoxyribonucléique
ANI	Aniline
ARN	Acide Ribonucléique
ATRX	Alpha Thalassemia/Mental Retardation Syndrome X-linked
CBTRUS	Central Brain Tumor Registry of the United States
C-IMPACT NOW	Consortium to Inform Molecular and Practical Approaches to CNS Tumor Taxonomy - Not Official WHO
DDA	Data Dependant Acquisition
DHB	2,5-Dihydroxybenzoic acid
DMSO	Diméthylsulfoxyde
EANO	European Association of Neuro-Oncology
EGFRvIII	Epidermal Growth Factor Receptor variant III
ESI	Electrospray
ExM	Expansion Microscopy
FFPE	Formalin Fixed Paraffin Embedded
FTICR	Fourier transform ion cyclotron resonance
GBM	Glioblastome
HCCA	α -Cyano-4-Hydroxycinnamic Acid
HES	Hématoxyline Eosine Safran
IDH	Isocitrate Dehydrogenase
IRM	Imagerie par Résonance Magnétique
ITO	Indium Tin Oxide
LC-MS	Liquid Chromatography-Mass Spectrometry
LCR	Liquide Céphalo-Rachidien
LESA	Liquid Extraction Surface Analysis
MALDI	Matrix Assisted Laser Desorption Ionization
MGMT	O6-methylguanine DNA methyltransferase
MS	Mass Spectrometry
MSI	Mass Spectrometry iMAGING
NF1	Neurofibromatose 1
NOS	Not Oserwise Specified
OMS	Organisation mondiale de la santé
PTEN	Phosphatase and Tensin homolog
ROIs	Regions Of Interest
SNC	Système Nerveux Central
TCGA	The Cancer Genome Atlas
TEP	Tomographie par émission de positons
TIC	Total ion count
TOF	Time of Flight
UC	Ultracentrifugation
VEs	Vésicules Extracellulaires

Avant-propos

Les tumeurs cérébrales sont classées selon les recommandations internationales de l'Organisation Mondiale de la Santé (OMS). Les gliomes forment un groupe très hétérogène qui peut présenter de grandes variations histologiques. Les glioblastomes représentent les tumeurs cérébrales malignes primitives les plus fréquentes. C'est à ce groupe de tumeurs que nous nous sommes intéressés dans ce travail.

La classification OMS permet de classer les gliomes selon le sous-type tumoral et selon le grade de malignité. Elle a pour rôle principal de guider le choix de la thérapie, et de donner des informations sur le pronostic des patients (Ostrom et al., 2014; Louis et al., 2016). En 2007, cette classification reposait principalement sur des aspects histologiques. Le type de gliome était alors défini par le type cellulaire majoritaire (astrocytome, oligodendrogliome, oligo-astrocytome, glioblastome) et le grade (I à IV) défini grâce aux données de différenciation et de densité cellulaire, la présence de nécrose, de prolifération vasculaire et l'activité mitotique (Louis et al., 2007).

Une mise à jour de la classification OMS a été publiée en 2016 (Louis et al., 2016). En effet, il est apparu que la classification OMS 2007 entraînait d'importants problèmes de reproductibilité inter- et intra-observateurs et un manque de précision pour le diagnostic, par ailleurs les progrès technique ont permis ont mis en évidence de nouveaux marqueurs, notamment en biologie moléculaire permettant de distinguer de nouveaux sous-groupes de gliomes. La classification de 2016 prend en compte les progrès techniques réalisés et intègre désormais des données génétiques et de biologie moléculaire. En effet, les profils moléculaires et génétiques influencent le comportement biologique des tumeurs et leurs susceptibilités aux différentes thérapies. Parmi les marqueurs essentiels utilisés en routine dans la classification OMS, on note le statut de la mutation du gène IDH1 et de la co-délétion des bras chromosomiques 1p et 19q sont notamment utilisés (Weller et al., 2017a). L'utilisation de ces marqueurs améliore et redéfinit donc les différents sous-types de tumeurs. Le groupe C-IMPACT NOW a été créé dans le but d'actualiser régulièrement, de manière consensuelle et entre les mises à jour de la classification OMS, les marqueurs qui améliorent

le classement et la prise en charge des tumeurs cérébrales (Brat et al., 2018, 2020; Ellison et al., 2020; Louis et al., 2016, 2017).

De nombreuses études protéomiques ont également fourni des informations importantes sur la physiopathologie des glioblastomes (Ghantasala et al., 2020; Jovčevska et al., 2017; Patil and Mahalingam, 2019; Stetson et al., 2016). Les protéines sont en effet affectées par les dysfonctionnements cellulaires, dans leur régulation et leur structure. La protéomique permet l'analyse de ces changements par l'observation directe des protéines et de leurs modifications post-traductionnelles (phosphorylation, glycosylation et acétylation). Néanmoins, la plupart des études menées actuellement analysent de façon globale la totalité des tumeurs alors qu'elles sont très hétérogènes. Ces études ne permettent donc pas de distinguer chaque sous-région au sein d'une tumeur d'un même patient. De plus, peu d'études prennent en compte les données cliniques des patients et la plupart n'ont pu évaluer qu'un nombre restreint de patients ce qui limite l'interprétation statistique des données.

En ce sens, les travaux réalisés par le laboratoire PRISM se sont distingués en prenant en compte l'hétérogénéité tumorale. Une première étude a démontré l'intérêt de la protéomique localisée couplée à l'imagerie moléculaire par spectrométrie de masse MALDI (Matrix-Assisted Laser Desorption Ionization) pour établir une classification des gliomes de grade III, basée sur les données protéomiques et les données cliniques (Le Rhun et al., 2016). L'imagerie MALDI MS donne en effet la possibilité de localiser différentes molécules au sein d'un tissu et permet donc de déterminer des signatures moléculaires en fonction de l'hétérogénéité de la tumeur (Bonnell et al., 2011; Franck et al., 2009). Le couplage de l'imagerie MALDI et de la protéomique localisée sur tissu permet ensuite d'obtenir des informations nouvelles sur les glioblastomes.

Ainsi, le premier chapitre de ce projet de thèse est de proposer une classification moléculaire des glioblastomes, tumeurs cérébrales malignes primitives les plus fréquentes, basée sur les signatures moléculaires obtenues par imagerie par spectrométrie de masse et microprotéomique.

La microdigestion *in situ* au niveau des régions d'intérêts permet d'identifier les protéines spécifiquement sur- et sous-exprimées dans les différentes régions tumorales. Le développement des techniques de microprotéomique spatialement définie paraît intéressant

pour étudier plus précisément un nombre limité de cellules et ainsi observer les différents mécanismes physiopathologiques. Dans ce contexte, nous avons développé une méthode de microéchantillonnage basée sur l'expansion de tissu (ExM) (Chen et al., 2015a). L'originalité de notre approche réside dans le fait d'étendre un morceau de tissu ou une coupe de tissu en 3 dimensions pour sélectionner facilement les régions d'intérêt à analyser. Ainsi, cette méthode a pour but d'agrandir les coupes de tissu et de réaliser une microdissection manuelle sur des régions d'intérêts pour ensuite réaliser l'analyse protéomique.

La deuxième partie de ce projet de thèse porte sur le développement de cette nouvelle technique de protéomique localisée.

La mise en corrélation des données moléculaires avec les données cliniques des patients est un enjeu crucial, l'objectif à terme étant d'adapter et de personnaliser les approches thérapeutiques. À l'heure actuelle, les principales options thérapeutiques pour les gliomes intègrent la résection chirurgicale maximale de la tumeur, en respectant la sécurité clinique, la radiothérapie et les pharmacothérapies (Weller et al., 2017a). Le traitement initial standard d'un glioblastome repose actuellement sur une chirurgie complétée par une période de radiothérapie et chimiothérapie par témozolomide (Témodal) concomitante puis par une période de chimiothérapie par témozolomide. Les réponses aux traitements sont cependant aux mieux modestes. Au moment de la récurrence, les réponses au traitement sont encore plus faibles avec des tumeurs particulièrement résistantes. Le pronostic des patients est de ce fait très faible.

Le suivi des patients atteints de glioblastome est réalisé par un suivi clinique, un suivi par imagerie par résonance magnétique (IRM) et par tomographie par émission de positons (TEP). Aucun biomarqueur n'est actuellement identifié pour surveiller la réponse au traitement et la progression tumorale. Ainsi, détecter ce que les cellules tumorales produisent en temps réel est un objectif majeur. Dans cette optique, l'étude des vésicules extracellulaires (VEs) est primordiale.

Les vésicules extracellulaires comme les exosomes ou les microvésicules sont des vésicules sécrétées à la fois par les cellules saines et cancéreuses (Abels and Breakefield, 2016). Par leur contenu en protéines, ADN et ARN, elles donnent des informations sur leur cellule d'origine. Elles sont aussi d'importantes médiatrices de la communication cellulaire

dans les cancers, notamment par le transport d'oncogène et de miARN. Elles sont également impliquées dans la progression tumorale, avec par exemple les miARN miR-1 ou miR-21 identifiés comme abondants dans ces vésicules et directement impliqués dans la progression et l'invasion tumorale (Bronisz et al., 2014; Gourlay et al., 2017).

Les VEs reflètent à la fois ce que les cellules cancéreuses produisent, mais aussi ce que le microenvironnement sécrète en réponse. Ces vésicules sont donc une source de biomarqueurs potentiels pour le diagnostic et le pronostic des patients, mais également potentiellement pour le suivi thérapeutiques (Minciacchi et al., 2015). Elles peuvent enfin représenter une cible thérapeutique potentielle. L'intérêt de l'étude de ces vésicules réside dans le fait qu'elles sont détectables dans les différents fluides biologiques comme le sang ou le liquide céphalo-rachidien (LCR) (Santiago-Dieppa et al., 2014). L'étude de ces vésicules donne donc accès à de nombreuses informations sans avoir besoin de réaliser de chirurgie avec les risques de complications potentielles inhérents.

Ainsi, la dernière partie de cette thèse s'intéresse aux protéines contenues dans les vésicules extracellulaires du sang de patients atteints de glioblastome dans le but de les relier à la tumeur et à la classification établie par imagerie MALDI. De plus, une approche sur tissus sera également envisagée dans l'optique de relier les vésicules à leur origine tumorale.

ETAT DE L'ART

I. Les gliomes

Les gliomes ou tumeurs gliales constituent un groupe de tumeurs cérébrales provenant de la glie, le tissu de soutien neuronal. Les glioblastomes sont les tumeurs cérébrales malignes primitives les plus fréquentes et les plus agressives.

A. Généralités

1. Épidémiologie

Les tumeurs cérébrales primitives, bénignes et malignes confondues, ont une incidence de 21,28 cas par 100 000 habitants par an selon les données du Central Brain Tumor Registry of the United States (CBTRUS) 2012-2016 (Ostrom et al., 2019). Les gliomes, qui émanent de la glie, représentent 27 % de ces tumeurs (Louis et al., 2016). Plusieurs sous-types histologiques sont identifiés au sein des gliomes, avec notamment les glioblastomes, les astrocytomes, les oligodendrogliomes et les tumeurs épendymaires.

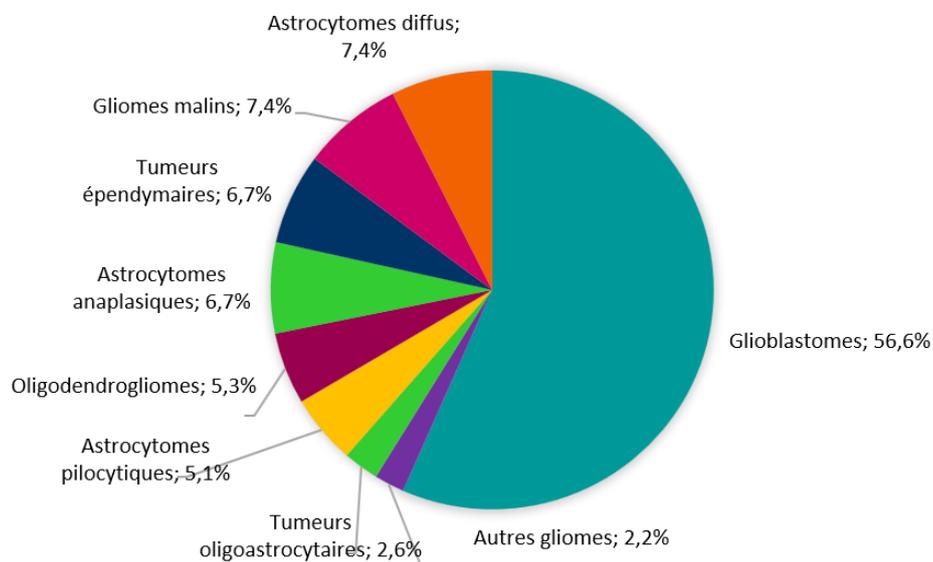


Figure 0-1 : Distribution des gliomes du système nerveux central par sous-types histologiques.

Les glioblastomes, tumeurs cérébrales primitives les plus agressives, représentent plus de la moitié des gliomes. Adapté de (Ostrom et al., 2019)

Les tumeurs de bas grades, comme les gangliogliomes, les épendymomes et les astrocytomes pilocytiques, sont généralement bénignes, limitées et nécessitent une résection complète de la tumeur par chirurgie. Dans le cas de tumeurs plus diffuses, il est souvent difficile d'extraire la totalité des tumeurs dont les cellules cancéreuses s'infiltrent dans le tissu cérébral sain.

Au sein des gliomes de haut-grades, les glioblastomes sont les tumeurs primitives du système nerveux central (SNC) les plus communes et les plus agressives (Figure 0-1). Ces tumeurs représentent 56,6 % des gliomes et 14,6 % de toutes les tumeurs du SNC confondues (Ostrom et al., 2019). En France, l'incidence globale de ces tumeurs est de 4 pour 100 000 habitants par an (Zouaoui et al., 2012), 3,22 pour 100 000 habitants selon les dernières données du CBTRUS (Ostrom et al., 2019). L'âge médian des patients au moment du diagnostic de glioblastomes est de 64 ans, il existe une légère prédominance masculine (Ostrom et al., 2019).

La majorité des gliomes apparaissent de façon sporadiques sans pathologie familiale sous-jacente mais il existe cependant une forte association entre ces tumeurs et certains syndromes héréditaires rares (Ostrom et al., 2015). En particulier, on observe 100 fois plus de gliomes de hauts grades chez les patients atteints de neurofibromatose de type 1 (NF1, maladie de Recklinghausen) (Albers and Gutman, 2009).

L'étiologie des gliomes est mal connue, seule l'exposition aux radiations ionisantes a été montrée comme pouvant augmenter le risque de tumeurs (Ellor et al., 2014; Ostrom et al., 2019). Les causes environnementales comme la pollution ou l'exposition aux pesticides sont souvent suggérées, cependant il n'existe aucune preuve formelle de leurs implications dans ces tumeurs. Bien qu'aucun lien de causalité n'ait été établi, il est intéressant de noter qu'une commune du Gard (France) accueillant un important bassin industriel, observe trois fois plus de cas de glioblastome par rapport à la moyenne départementale (Rivière Stéphanie, 2020). A noter également que les antécédents d'allergie semblent réduire le risque de tumeur cérébrale (Ostrom et al., 2019).

L'origine virale des gliomes est un sujet émergent. Les virus sont en effet connus pour être responsables de 15 % des cancers et le papillomavirus humain (HPV) est par exemple une

cause directe du cancer du col de l'utérus (Morales-Sánchez and Fuentes-Pananá, 2014). En effet, étant donné leur capacité à modifier les cellules normales pour produire plus de virus ils peuvent notamment induire des dommages sur l'ADN et ainsi augmenter l'instabilité génétique de ces cellules. De plus, ils sont susceptibles de causer une inflammation chronique des tissus, favorisant le développement de cancer. En particulier, le virus Epstein-Barr est reconnu comme étant impliqué dans différents cancers comme les lymphomes, les cancers du sein et de la prostate, mais son implication dans les gliomes est encore largement débattue (Akhtar et al., 2018; Morales-Sánchez and Fuentes-Pananá, 2014). Aucune donnée avec un niveau de preuve suffisant n'est disponible concernant les gliomes à l'heure actuelle.

2. Diagnostic

Les principaux symptômes incluent des céphalées, des déficits neurologiques et possiblement des troubles du comportement. Ces symptômes varient cependant selon la taille et la localisation de la tumeur (Porter, 2012; Wirsching et al., 2016). Les signes cliniques sont généralement associés aux zones fonctionnelles atteintes du cerveau. Par exemple, des troubles du comportement sont fréquemment retrouvés dans les localisations frontales, des troubles de la vision dans les localisations occipitales, des troubles du langage dans les lésions temporales gauches chez les droitiers. L'installation des symptômes dépend également de la rapidité de croissance du gliome. En cas de glioblastome, le diagnostic est généralement posé rapidement quelques semaines à quelques mois après les premiers signes. Dans le cas d'une croissance tumorale lente, les symptômes peuvent survenir plusieurs années après l'apparition de la tumeur (Alexander and Cloughesy, 2017).

L'imagerie par résonance magnétique (IRM) est le premier examen à être réalisé en cas de suspicion de tumeur cérébrale. En tant qu'examen de référence, il permet de localiser et d'évaluer la taille et la présentation de la tumeur, de réaliser le suivi et d'évaluer l'efficacité des traitements. À l'IRM, presque tous les glioblastomes observent une prise de contraste avec du gadolinium et montrent une masse de formes irrégulières avec un centre nécrosé (Lapointe et al., 2018). Le scanner cérébral et la tomographie par émission de positons (TEP) peuvent être réalisés à la demande du chirurgien. Ces examens ont pour but principal de préparer le geste opératoire pour, entre autres, délimiter les zones à biopsier. De plus, la TEP

peut également avoir un intérêt pour le suivi du patient, notamment en cas doute de radionécrose (effet secondaire de la radiothérapie) ou progression tumorale.

Le diagnostic est confirmé par les analyses anatomopathologiques et de biologie moléculaire effectuées sur un prélèvement chirurgical de tissu (exérèse ou biopsie). En effet, la nature exacte de la tumeur ne peut être connue sans ces analyses complémentaires. L'examen histologique d'un glioblastome est caractérisé par de la nécrose et de la prolifération endothélio-capillaire avec parfois présence d'hémorragie (Lapointe et al., 2018; Louis et al., 2016; Mandoj et al., 2019).

3. Pronostic

Les gliomes bien circonscrits, tels que les astrocytomes pilocytiques, ont un bien meilleur pronostic, la qualité de la résection chirurgicale étant un élément déterminant pour le pronostic. De manière générale, les gliomes 2 à 4 sont des pathologies incurables. Une chirurgie d'exérèse "totale" ou "subtotale" étant souvent difficile, les cellules tumorales infiltrantes restent inévitablement dans le cerveau environnant et entraînent inmanquablement la progression ou la récurrence de la tumeur plus tard (Wilson et al., 2014).

Tableau 0-1 Taux de survie à 1 an, 2 ans, 5 ans et 10 ans de différents types de gliomes

	Survie à 1 an	Survie à 2 ans	Survie à 5 ans	Survie à 10 ans
<i>Astrocytomes diffus</i>	74,7 %	64,1 %	51,6 %	41,3 %
<i>Astrocytomes anaplasiques</i>	64,3 %	46,0 %	30,2 %	22,3 %
<i>Oligodendrogliomes</i>	94,5 %	90,1 %	82,7 %	68,7 %
<i>Oligodendrogliomes anaplasiques</i>	85,8 %	74,3 %	60,2 %	47,3 %
<i>Glioblastomes</i>	40,8 %	18,5 %	6,8 %	4,7 %

d'après (Ostrom et al., 2019)

Les glioblastomes sont les tumeurs cérébrales primitives dont le pronostic est le plus sévère. Sans traitement la survie des patients n'excède pas quelques mois. Même avec

traitement, plus de la moitié des patients ne survivent pas après un an. Après 2 ans, on observe moins de 20 % de survie et seulement 4 % après 10 ans (Ostrom et al., 2019) (Tableau 0-1).

Les marqueurs pronostiques peuvent donner des indications sur l'évolution naturelle d'un cancer, indépendamment du traitement. Ils peuvent aider à définir quel patient doit être traité et à quel point le traitement doit être agressif par exemple. L'identification de biomarqueurs pronostiques et prédictifs est nécessaire pour classer les tumeurs selon des sous-groupes de pronostiques distincts, individualiser les soins et améliorer la survie et la qualité de vie des patients.

Des facteurs de mauvais pronostic sont connus, entre autres un âge supérieur à 40 ans, des signes cliniques sévères notamment avec une hypertension intracrânienne, une épilepsie difficilement contrôlable, un volume tumoral important et une évolutivité rapide de la tumeur. L'état général du patient (indice de Karnofsky), la localisation de la tumeur et les possibilités de chirurgie sont également des facteurs pronostiques. Les marqueurs moléculaires ont également permis de mieux classer les gliomes et d'améliorer les connaissances sur le pronostic.

4. Traitements

Le traitement est défini dans les directives EANO (European Association of Neuro Oncology) (Weller et al., 2017a). La dernière version date de 2017 et une nouvelle version est en cours de préparation. Les décisions thérapeutiques sont prises lors de réunions de concertation pluridisciplinaires. La chirurgie, première étape du traitement, permet également de poser le diagnostic. Une résection maximale en toute sécurité de la tumeur est effectuée. Si une exérèse ne peut pas être réalisée en raison de la localisation anatomique notamment, des biopsies sont généralement effectuées pour confirmer le diagnostic et décider de la prise en charge. À cause de l'aspect diffus des glioblastomes, l'exérèse complète est difficile et les cellules tumorales résiduelles sont souvent à l'origine de la récurrence.

Depuis 2005, les patients sont généralement traités après la chirurgie (résection ou biopsie si la résection est impossible) par radiothérapie en combinaison avec de la chimiothérapie par témozolomide pendant 6 semaines puis par chimiothérapie par

témozolomide pendant 6 mois (correspondant à 6 cycles de 5 jours effectués tous les 28 jours) (Stupp et al., 2005, 2009). La combinaison de ces traitements permet d'augmenter la moyenne de survie de 12,1 à 14,6 mois par rapport à la radiothérapie seule (Stupp et al., 2005, 2009). Le témozolomide est un agent alkylant de l'ADN qui pénètre facilement la barrière hémato-encéphalique. Il induit la méthylation de la guanine en O6 et N7 de l'ADN ce qui cause des dommages irréversibles et provoque la mort cellulaire par apoptose (Zhang, F.G. Stevens and D. Bradshaw, 2011). Le statut de méthylation du promoteur du gène de la méthylguanine méthyltransférase (MGMT) est un facteur prédictif de la réponse à ce traitement (Hegi et al., 2004). Les patients sont généralement suivis tous les 3 mois par une évaluation clinique et par IRM cérébrale.

Au moment de la progression, une nouvelle chirurgie peut être discutée, généralement si la tumeur récidive plusieurs mois après le diagnostic. Un essai est en cours pour évaluer cette indication (RESURGE ; NCT02394626). De nouvelles irradiations peuvent être discutées au cas par cas, mais il n'y a pas de preuve formelle de leur intérêt et il existe un risque de radionécrose. Une reprise du traitement au témozolomide peut être parfois proposée. La lomustine, nitrosourée, est considérée comme le traitement de référence lors de la première récidive (Weller and Le Rhun, 2020). Une inclusion dans des essais cliniques est proposée lorsque c'est possible. Les patients sont généralement suivis tous les 3 mois.

Malgré ces traitements, la médiane de survie des patients reste faible et les cellules tumorales infiltrées dans le cerveau induisent fatalement la récidive de la tumeur. Dans cette optique, de nouveaux traitements sont en développement constant.

Actuellement 2 études sponsorisées par l'EORTC (European Organisation for Research and Treatment of Cancer) sont ouvertes. L'étude de phase III EORTC 1709 (NCT03345095) évalue l'efficacité du marizomib, un inhibiteur du protéasome chez les patients présentant un nouveau diagnostic de glioblastome. L'étude de phase I/II EORTC 1608 (NCT03224104) évalue la faisabilité de la combinaison du TG02, un inhibiteur des CDK (notamment CDK9) à la radiothérapie ou au temozolomide chez les patients présentant un glioblastome nouvellement diagnostiqué et explore l'efficacité du TG02 chez les patients en première récidive.

Les immunothérapies prennent une place grandissante dans les traitements émergents. Les résultats de l'étude Checkmate 548 trial (NCT02667587), évaluant le nivolumab, anti-PD1, en première ligne chez les patients atteints de glioblastome avec méthylation du promoteur du gène MGMT sont en attente. Les immunothérapies n'agissent pas directement sur la tumeur mais stimulent le système immunitaire pour lutter efficacement contre les cellules tumorales (Tivnan et al., 2017). Les tumeurs gliales sont composées de 30 à 50 % de cellules immunitaires et ces cellules ont un rôle important dans la tumorigenèse (Gieryng et al., 2017). Les glioblastomes altèrent les cellules immunitaires du système nerveux et notamment les macrophages recrutés par la tumeur permettent d'augmenter la migration et l'invasion des cellules tumorales (Vollmann-Zwerenz et al., 2020). Ainsi, une stratégie développée au laboratoire se base sur la réactivation des macrophages pour empêcher la progression tumorale et éradiquer la tumeur (Duhamel et al., 2018). De plus, différents essais d'immunothérapies émergent, notamment l'utilisation de CAR-T, des cellules T modifiées pour reconnaître et induire la lyse de cellules tumorales spécifiques (Bagley et al., 2018).

Des stratégies vaccinales se développent également et ciblent entre autres des protéines anti-apoptotiques comme la survivine (SurVaxM – phase II) ou des facteurs de transcription comme la protéine tumorale de Wilm (DSP-7888 -phase II) (Fenstermaker et al., 2016; De Groot et al., 2018). Il existe également des vaccins peptidiques personnalisés qui prennent en compte la diversité génétique, les biomarqueurs tumoraux et les profils cliniques de chaque patient (Cuoco et al., 2018). Actuellement en phase I, le vaccin peptidique GAVPAC se base sur les mutations et l'analyse du transcriptome de chaque tumeur individuellement (Hilf et al., 2019). Cette approche permet de créer des thérapies ciblées individualisées.

Les thérapies personnalisées sont intéressantes pour le traitement des gliomes qui constituent un groupe de tumeurs très hétérogène. Les glioblastomes peuvent en effet présenter de grandes variations histologiques et moléculaires (Brennan et al., 2013). Ces variations se retrouvent au niveau clinique et il est important de noter que tous les patients ne réagissent pas de la même façon face au traitement, certains présentant des résistances (Eder and Kalman, 2014). En ce sens, le développement de traitement personnalisé apparaît comme une évidence.

B. Classification des gliomes

1. Classification des tumeurs cérébrales

La classification des gliomes est un élément clé pour orienter les thérapies. Les tumeurs cérébrales sont classées selon la classification internationale de l'Organisation Mondiale de la Santé (OMS). Celle-ci permet de classer les tumeurs selon leur type cellulaire et selon leur grade de malignité, elle a pour principaux rôles de prédire le comportement des tumeurs et de donner des informations sur le pronostic des patients (Louis et al., 2007).

Depuis sa création, la classification des tumeurs cérébrales a subi de nombreuses améliorations et plusieurs versions ont été publiées. En 1926, Bailey et Cushing, deux neurochirurgiens, ont soumis pour la première fois leurs observations dans leur manuscrit « A Classification of the Tumors of the Glioma Group on a Histogenetic Basis with a Correlated Study of Prognosis » (Ferguson and Lesniak, 2005; Percival Bailey and Harvey Cushing, 1927). Cette première classification est la base de la neuro-oncologie moderne. Auparavant, les gliomes étaient couramment regroupés ensemble et considérés comme « fatal ». Cependant, Bailey et Cushing ont noté de grandes variations entre les patients atteints de gliomes : après chirurgie, la survie des patients paraissait aléatoire et de nombreux patients survivaient beaucoup plus longtemps que prévu (Percival Bailey and Harvey Cushing, 1927). Cette observation les a amenés à effectuer une étude approfondie de la structure de 254 biopsies de gliome. Par observation microscopique, ils classent ces gliomes selon leurs compositions et les désignent en accord avec leurs caractéristiques cytologiques : les astrocytomes sont des tumeurs composées d'astrocytes, les oligodendrogliomes sont composées d'oligodendrocytes et les oligoastrocytomes sont des tumeurs mixtes (Percival Bailey and Harvey Cushing, 1927).

La première édition de la classification de l'OMS, « Histological Typing of Tumors of the Central Nervous System », paraît en 1979 et se base sur le consensus d'un groupe de travail international composé d'experts. Ce premier « blue book », couleur distinctive de la série, contient une nomenclature et des définitions précises (Kleihues et al., 1993; Scheithauer, 2009). En plus d'une classification et de définitions générales, la notion de « grades » de malignité est essentielle à des fins thérapeutiques et pronostiques. Ainsi, quatre grades (I, II,

III et IV) ont été définis. Les tumeurs de grade I et II étaient alors généralement considérées comme bénignes et les tumeurs de grade III à IV plus agressives (Scheithauer, 2009).

Depuis l'acceptation internationale de la première classification, des classifications actualisées sont publiées régulièrement.

2. Classification histologique 2007

En 2007 paraît la quatrième classification des tumeurs du système nerveux central (Louis et al., 2007). Elle est principalement basée sur des données histologiques, le type de tumeur étant défini par le type cellulaire majoritaire.

Les gliomes se développent à partir des cellules gliales, en particulier des astrocytes et des oligodendrocytes. La classification détermine également l'échelle de malignité par la différenciation et la densité cellulaire, la présence de nécrose, de prolifération vasculaire, d'atypies nucléaires et d'activité mitotique (Tableau 0-2). Ainsi, le grade I (astrocytome pilocytique) caractérise des tumeurs avec une faible prolifération cellulaire souvent bien limitée. À l'inverse, le grade IV (glioblastome), qui correspond au grade le plus sévère, s'applique à des tumeurs infiltrantes avec forte activité mitotique, sujettes à la nécrose et à la prolifération vasculaire et avec évolution rapide.

Tableau 0-2 Classification OMS 2007 des gliomes

	I	II	III	IV	Description
<i>Astrocytome Pilocytique</i>	•				Cellules avec de longues projections ressemblant à des cheveux
<i>Astrocytome diffus</i>		•			Infiltrant, faible activité mitotique
<i>Oligodendrogliome</i>		•			
<i>Astrocytome anaplasique</i>			•		Très infiltrant, présence d'atypie nucléaire et activité mitotique importante
<i>Oligodendrogliome anaplasique</i>			•		
<i>Glioblastome</i>				•	Activité mitotique intense, prolifération endothélio- capillaire, nécrose, généralement associé à une évolution rapide

Adapté de (Louis et al., 2007)

Les gliomes de grade I, comme les astrocytomes pilocytiques, sont généralement peu agressifs. Une exérèse complète peut permettre une guérison. Dans les gliomes diffus (grades II et III), il n'existe pas de frontières nettes entre la tumeur et les tissus environnant et la chirurgie est par conséquent difficile. Les gliomes diffus sont très infiltrants et les cellules tumorales résiduelles après la chirurgie conduisent souvent à la récurrence de la tumeur. Les tumeurs anaplasiques (III) ont une néovascularisation importante, une activité mitotique intense et une grande densité cellulaire. Les gliomes de grades IV, les glioblastomes, sont marqués par la présence de nécrose et de prolifération endothélio-capillaire intense.

Ce système de classification entraîne cependant un manque de précision pour le diagnostic. Les critères visuels histologiques sont également sujets à d'importants problèmes de reproductibilité inter et intra-observateurs (van den Bent, 2010). Par ailleurs, les progrès en génétique de ces dernières années ont permis d'identifier de nombreuses altérations génétiques impliquées dans la formation et la progression des gliomes (Parsons et al., 2008). Les profils moléculaires et génétiques influencent le comportement biologique des tumeurs et leurs susceptibilités aux différentes thérapies (Parsons et al., 2008; Yan et al., 2009). Les biomarqueurs deviennent donc essentiels pour le diagnostic des glioblastomes et ainsi une mise à jour de la classification a été publiée en 2016 (Louis et al., 2016).

3. Classification moléculaire 2016

a) Mise à jour de la classification

La dernière mise à jour de cette classification intègre désormais des données génétiques et de biologie moléculaire. L'utilisation de ces biomarqueurs ajoute un degré supplémentaire d'objectivité dans le diagnostic et permet de définir des groupes plus homogènes d'un point de vue biologique. Elle permet également d'orienter le choix des thérapies et d'identifier de nouveaux facteurs pronostics et prédictifs de la survie et de la réponse aux traitements (Louis et al., 2016).

Même si l'analyse histologique est toujours nécessaire notamment pour la détermination du grade de malignité, on observe une restructuration majeure en ce qui concerne les gliomes (Louis et al., 2016). Les trois marqueurs moléculaires utilisés pour la

classification sont : la mutation du gène Isocitrate Déshydrogénase (IDH-mutant, IDH non muté), la co-délétion 1p/19q et la mutation H3-K27M (Tableau 0-3).

Une tumeur sans mutation diagnostique reçoit la désignation *not otherwise specified* (NOS). Par exemple, les oligoastrocytomes, qui ont souffert d'une grande discordance inter-observateurs, sont dorénavant définies en tant que NOS et leurs diagnostic doit être limité aux rares cas où le profil moléculaire ne permet pas de conclure entre oligodendrogliomes et astrocytomes (van den Bent, 2010; Louis et al., 2016).

Tableau 0-3 Classification OMS 2016 des gliomes

	I	II	III	IV
<i>Astrocytome pilocytique</i>	•			
<i>Astrocytome diffus,</i> - <i>IDH mutant</i> - <i>IDH non muté</i> - <i>NOS</i>		•		
<i>Oligodendrogliomes,</i> - <i>IDH mutant avec co-délétion 1p/19q</i> - <i>NOS</i>		•		
<i>Astrocytome anaplasique,</i> - <i>IDH mutant</i> - <i>IDH non muté</i> - <i>NOS</i>			•	
<i>Oligodendrogliomes anaplasiques,</i> - <i>IDH mutant avec co-délétion 1p/19q</i> - <i>NOS</i>			•	
<i>Glioblastome,</i> - <i>IDH non muté</i> - <i>IDH mutant</i> - <i>NOS</i>				•
<i>Gliome diffus de la ligne médiane, H3-K27M-muté</i>				•

Adapté de (Louis et al., 2016)

La mutation IDH est la modification majeure intervenant dans la nouvelle classification des gliomes. Le statut IDH permet en effet de distinguer les différents sous-types et définit des tumeurs spécifiques d'un point de vue biologique et clinique. La mutation du gène IDH1 est une mutation qui survient tôt dans le développement des gliomes et est principalement retrouvée chez les patients atteints de gliome de plus bas grade (Weller et al., 2013; Yan et

al., 2009). IDH1 et IDH2 sont des gènes codant pour l'enzyme isocitrate déshydrogénase qui catalyse la conversion de l'isocitrate en α -cétoglutarate. La mutation induit la production de 2-hydroxyglutarate (2-HG) et la diminution de Nicotinamide adénine dinucléotide phosphate (NADPH) ce qui favorise la méthylation à l'échelle du génome dans la cellule. Les gliomes IDH mutés montrent une hyperméthylation des îlots CpG, constituant un sous-ensemble appelé « glioma CpG-island methylator phenotype (G-CIMP) » (Turcan et al., 2012). Cette mutation est marqueur de bon pronostic.

La co-délétion des bras chromosomiques 1p/19q est caractéristique des oligodendrogliomes. Les patients avec co-délétion 1p/19q ont une survie significativement plus longue que les patients sans cette altération et sont plus sensibles à la chimiothérapie et à la radiothérapie (Van Den Bent et al., 2013; Liu et al., 2016)

Finalement, un groupe restreint de tumeur est défini par la mutation K27M dans le gène de l'histone H3 H3F3A, ou moins fréquemment dans le gène HIST1H3B. Ce nouveau sous-groupe, les gliomes diffus de la ligne médiane H3-K27M-muté, apparaît principalement chez les enfants et sont de pronostic très défavorable (Louis et al., 2016).

b) Les marqueurs génétiques d'intérêt

D'autres marqueurs sont également fréquemment recherchés étant donné leurs intérêts pronostic et pour orienter le choix des thérapies.

Parmi ces marqueurs, la méthylation du promoteur de la méthylguanine méthyltransférase (MGMT) est un facteur pronostic important qui augmente l'effet de la chimiothérapie et la survie des patients (Weller et al., 2013). MGMT est une enzyme de réparation de l'ADN et la méthylation de son promoteur a pour conséquence d'inhiber la transcription du gène (Cankovic et al., 2013). La sensibilité aux agents alkylants est réduite en présence de MGMT, par conséquent la tumeur devient chimiosensible en son absence.

L'amplification EGFR est l'une des altérations observées le plus tôt dans les glioblastomes. L'activation d'EGFR déclenche des voies de signalisation provoquant la prolifération et la migration cellulaire.

Le gène codant pour la protéine tumorale 53 (tp53) est l'un des plus connus pour être impliqué dans la tumorigenèse, p53 régulant notamment le cycle cellulaire et l'apoptose. La mutation du gène implique la perte de la protéine et de ses fonctions anti proliférative et apoptotique.

La perte d'expression de la protéine alpha-thalassemia/mental retardation X-linked (ATRX), intervenant dans le remodelage de la chromatine, est plus souvent retrouvée dans les gliomes de bas grade (Kannan et al., 2012; Liu et al., 2019). Avec la mutation p53, ces deux altérations sont caractéristiques des astrocytomes diffus IDH mutant (Louis et al., 2016).

Le gain du chromosome 7 et la perte du chromosome 10 sont retrouvés dans 80 % des glioblastomes mais la perte semble limitée au bras long dans les GBM secondaires (Otani et al., 2017). La perte de ce chromosome implique directement la protéine suppresseur de tumeur PTEN (10q23.3). En cas de différenciation cellulaire excessive, PTEN régule le cycle cellulaire et provoque l'apoptose, empêchant ainsi la croissance tumorale.

Le gène TERT (telomerase reverse transcriptase) est essentiel dans le maintien des télomères. Dans les cancers, les mutations du promoteur de TERT sont connues pour augmenter l'activité de la télomérase, permettant une extension indéfinie des télomères et l'immortalisation des cellules (Arita et al., 2013). Au sein des gliomes, on observe cette mutation dans 70–75 % des glioblastomes (Ohgaki and Kleihues, 2013).

L'ensemble de ces biomarqueurs a un intérêt diagnostique important et il a été démontré que la nouvelle classification est plus représentative du pronostic des patients que la classification histologique (Weller et al., 2015). Cette nouvelle classification est susceptible d'être encore affinée par l'intégration de biomarqueurs supplémentaires. La classification OMS est régulièrement mise à jour selon l'avancée des connaissances et à l'heure actuelle, de nombreuses études s'intéressent notamment au profil de méthylation de l'ADN dans l'optique de stratifier les patients (Lee et al., 2015; Orozco et al., 2018).

Le consortium cIMPACT-NOW (*Consortium to Inform Molecular and Practical Approaches to CNS Tumor Taxonomy*) a été créé sous le parrainage de la Société internationale de neuropathologie à la suite de la mise à jour de 2016 de la classification OMS (Louis et al., 2017). Le but de ce consortium est de proposer des changements quant aux futures

classifications des tumeurs du SNC. Les mises à jour du cIMPACT-NOW ne visent pas à remplacer la classification OMS existante (cIMPACT-NOW est «*Not Official WHO*»), mais à fournir des directives possibles pour les praticiens et les futures mises à jour de la classification OMS (Brat et al., 2018, 2020; Ellison et al., 2019, 2020; Louis et al., 2018a, 2018b, 2020) (Tableau 0-4).

Tableau 0-4 Recommandations du groupe cIMPACT NOW

Mises à jour du groupe cIMPACT NOW	
<i>cIMPACT 1</i> (Louis et al., 2018a)	Clarification de l'utilisation du terme NOS (<i>Not Oserwise Specified</i>) et proposition de l'utilisation du terme NEC (<i>Not Elsewhere Classified</i>) quand le diagnostic est impossible ou pour une nouvelle entité émergente.
<i>cIMPACT 2</i> (Louis et al., 2018b)	Clarification du terme « gliome diffus de la ligne médiane, mutant H3 K27M » Diagnostic d'astrocytome IDH mutant possible en l'absence de test 1p / 19q si perte d'expression nucléaire ATRX et/ou immunopositivité p53 forte.
<i>cIMPACT 3</i> (Brat et al., 2018)	Proposition de la dénomination « gliomes astrocytaires diffus, IDH non mutant, avec les caractéristiques moléculaires du glioblastome, grade IV » pour les astrocytomes diffus de grade II et III IDH non muté avec amplification EGFR et/ou gain du chromosome 7 entier et perte du chromosome 10 entier (+ 7 / -10) et/ou mutation du promoteur TERT.
<i>cIMPACT 4</i> (Ellison et al., 2019)	Recommandation d'un système intégré de diagnostic, combinant les caractéristiques histologiques et génétiques.
<i>cIMPACT 5</i> (Brat et al., 2020)	Recommandation d'ajouter la délétion homozygote de CDKN2A/B comme critère de grade IV pour les astrocytomes mutants IDH. Et recommandation de la dénomination « astrocytome, mutant IDH, grade IV » à la place de « glioblastome, IDH mutant, grade IV ».
<i>cIMPACT 6</i> (Louis et al., 2020)	Recommandation de l'utilisation de nouveaux types et sous-types.
<i>cIMPACT 7</i> (Ellison et al., 2020)	Recommandation d'une nouvelle classification pour les épendymomes.

4. Classification des glioblastomes

Les glioblastomes sont les gliomes les plus agressifs, initialement qualifiés de multiformes, ils constituent un groupe très hétérogène de tumeurs. Selon la classification de

2016, les glioblastomes peuvent être classés selon leurs statuts IDH. Les glioblastomes IDH non mutés représentent 90 % des glioblastomes et incluent comme variantes morphologiques le glioblastome à cellules géantes, le gliosarcome et le glioblastome épithélioïde. Les IDH mutants ne représentent que 10 % des cas (Louis et al., 2016). Cette répartition est souvent associée à la répartition entre glioblastome primaire et secondaire, les glioblastomes IDH non muté correspondant aux glioblastomes primaires et les IDH mutants aux glioblastomes secondaires qui par définition se développeraient à partir de gliome de bas grade.

Leurs développements étant différents, les deux entités expriment des mutations différentes et ont des médianes de survie différentes (Louis et al., 2016). En particulier, les mutations PTEN et TERT et l'amplification d'EGFR sont plus fréquentes dans les glioblastomes IDH non mutés. Au contraire, on observe plus souvent les mutations TP53 et ATRX dans les glioblastomes secondaires IDH mutants. Il est également intéressant de noter que ces différences impactent le devenir des patients, les GBM IDH mutant ayant une médiane de survie beaucoup plus longue.

Cependant, même au sein d'une même catégorie les glioblastomes sont très différents du point de vue clinique et ne répondent pas de la même manière aux traitements. Les profils moléculaires peuvent expliquer ces disparités et mettre en lumière les différents mécanismes de progression tumorale. Ainsi, les glioblastomes peuvent être classés en sous-types sur la base de différents marqueurs (Lee et al., 2018).

Cette nomenclature sera finalement mise à jour lors de la prochaine révision de la classification à partir des recommandations cIMPACT-NOW. Les modifications principales incluent une définition moléculaire des glioblastomes basée sur le statut IDH ainsi que sur l'amplification EGFR, le gain du chromosome 7 et la perte du chromosome 10, et la mutation du promoteur TERT (Brat et al., 2018).

Actuellement, les efforts pour sous-classifier davantage les glioblastomes se concentrent principalement sur le génome, le méthylome et le transcriptome de ces tumeurs. Le Cancer Genome Atlas (TCGA) a permis de répertorier les anomalies génomiques récurrentes des glioblastomes (Verhaak et al., 2010). Ainsi, quatre sous-groupes de glioblastomes ont été mis en évidence : proneural, neural, classique et mésenchymateux (Verhaak et al., 2010). Le type proneural est peu fréquent et apparaît principalement chez les

patients jeunes atteints de glioblastome IDH mutant. Ce type a des caractéristiques communes avec les oligodendrogliomes, par exemple la surexpression des gènes PDGFRA, NKX2-2, et OLIG2. Le type neural présente l'expression de marqueurs neuronaux. Ces deux sous-types ont une amplification d'EGFR et sont les plus similaires au tissu cérébral sain. Les glioblastomes classiques expriment des marqueurs de cellules souches de neurones comme la nestine et des voies de signalisation NOTCH et Sonic Hedgehog. C'est dans ce sous-type qu'on observe le plus fréquemment l'amplification du chromosome 7 et la perte du chromosome 10. Finalement le type mésenchymateux est caractérisé par la mutation du gène NF1 et exprime des marqueurs qui rappellent la transition épithélio-mésenchymateuse. Ce sous-type est considéré comme le plus agressif (Wang et al., 2017). Plus récemment, le sous-type neural a été associée à la localisation à la marge de la tumeur où il y a une augmentation du tissu neural sain ce qui suggère que ce sous-type n'est pas tumeur spécifique (Gill et al., 2014; Wang et al., 2017).

L'analyse du profil de méthylation de l'ADN peut également être utilisée pour établir une sous-classification des glioblastomes. Les patients avec méthylation G-CIMP ont été associés à une survie plus longue (Noushmehr et al., 2010). Le profilage de la méthylation du génome chez les patients adultes atteints de glioblastome IDH non muté a permis de définir trois sous-types épigénétiques. Le premier sous-type est le sous-type récepteur tyrosine kinase (RTK) I avec amplification PDGFR, le deuxième sous-type est le sous-type classique RTK II souvent avec amplification EGFR, suppression CDKN2A / B, et mutation PTEN. Le dernier sous-type est le sous-type mésenchymateux (Sturm et al., 2012).

Aussi, la localisation de la tumeur constitue une information importante. Les gliomes diffus H3 K27M mutants de la ligne médiane ont été ajoutés à la liste des gliomes diffus dans la classification 2016 (Louis et al., 2016). Faisant anciennement partie de la catégorie des glioblastomes, ces tumeurs sont maintenant définies comme gliomes intermédiaires de haut grade avec une différenciation astrocytaire prédominante et une mutation K27M dans les gènes H3F3A ou HIST1H3B/C. Ces tumeurs surviennent principalement chez les enfants et sont le plus souvent localisées dans le tronc cérébral, le thalamus et/ou la moelle épinière. Tant chez l'enfant que chez l'adulte, leur pronostic est généralement mauvais (OMS grade IV) (Ostrom et al., 2019). De plus, les tumeurs proches de la zone sous-ventriculaire, qui est le site

majeur de la neurogénèse dans le cerveau, apparaissent plus agressives que celles éloignées de cette zone (Gollapalli et al., 2017).

À l'heure actuelle, ces classifications sont peu utilisées en clinique et n'ont pas conduit à l'amélioration de la survie des patients. Les glioblastomes étant très hétérogènes, la définition de sous-type est également contestée. Cette hétérogénéité fait qu'il n'existe aucun consensus sur les marqueurs moléculaires à prendre en compte dans l'examen de la progression tumorale. De nouvelles approches sont donc nécessaires afin d'une part de mieux appréhender la pathologie pour la recherche de traitement et d'autre part pour améliorer la qualité de vie des patients.

Finalement, les approches protéomiques restent rares. L'analyse du protéome permet pourtant d'identifier et de quantifier le produit final de la génomique et de la transcriptomique et peut caractériser l'activation de voies de signalisation spécifiques (Deighton et al., 2014; Kalinina et al., 2011). La détermination de signatures protéomiques spécifiques pourrait aider à améliorer la distinction entre les différents sous-types de glioblastome et à guider la prise en charge thérapeutique.

C. Protéomique des gliomes

1. Marqueurs protéiques identifiés dans l'étude des échantillons tumoraux

L'identification de nouveaux biomarqueurs pourrait aider au diagnostic et à mieux connaître le pronostic, guider la décision de traitement, et aider à l'évaluation de la réponse au traitement. L'analyse à grande échelle des protéines, la protéomique, permet d'étudier directement les processus cellulaires et pathologiques. De nombreuses études protéomiques ont ainsi fourni des informations importantes sur la physiopathologie des glioblastomes.

L'analyse de 75 échantillons de gliomes humains, en utilisant une stratégie en gel bidimensionnel, a permis d'identifier 37 protéines associées aux gliomes. Parmi ces protéines, 27 sont marqueurs de gliomes de "haut grade" et 10 marqueurs de gliomes de "bas grades" (Iwadate et al., 2004). Plus récemment un ensemble de 13 protéines ayant une valeur pronostic a été identifié à la suite d'une analyse sur plus de 200 patients (Stetson et al., 2016).

Parmi elles, quatre sont marqueurs de longue survie (GAB2, SMAD1, phosphorylated Src, et Src) tandis que neuf sont de mauvais pronostic (AMPK, ANXA1, phosphorylated c-Jun, phosphorylated c-Met, phosphorylated CHEK1, Cox-2, IGFBP2, phosphorylated NFkB1, et TGM2).

Différentes protéines ont également été associées aux sous-types mis en évidence par transcriptomique (Patil and Mahalingam, 2019; Verhaak et al., 2010). En effet, EGFR a été identifiée comme surexprimée dans le sous-type classique, SERPINE1, phospho-MYH9 et TAZ dans le type mésenchymateux et dans le type proneural une surexpression de phospho-GSK3 est observée (Patil and Mahalingam, 2019).

De plus, une analyse comparative entre des patients jeunes et âgés atteints de glioblastomes a identifié plusieurs protéines différentiellement exprimées et impliquées dans la régulation de la tumorigenèse. Par exemple, la protéine PEPB1 est surexprimée chez les patients jeunes et sous exprimée chez les patients âgés (Deighton et al., 2014). Des analyses complémentaires pourraient expliquer les différences de pronostiques entre les patients.

Il est intéressant de noter que selon les méthodes d'analyse, les protéines d'intérêt sont différentes. Aussi, on peut souligner que l'agressivité des gliomes ne peut être prédite par une seule protéine biomarqueur, mais plutôt par une combinaison de protéines. Cette disparité peut également provenir de l'hétérogénéité de ces tumeurs. En effet, ces études analysent de façon globale la totalité des protéines du tissu sans prendre en compte l'hétérogénéité intra-tumorale. Elles ne permettent donc pas de distinguer chaque sous-région au sein d'une tumeur d'un même patient et engendrent des phénomènes de dilution des protéines par les protéines majoritaires. Cette dilution empêche l'analyse de l'ensemble des voies de signalisation impliquées dans la physiopathologie des glioblastomes.

2. Études sur les fluides biologiques

L'analyse du tissu tumoral est nécessaire pour le diagnostic de la pathologie. Le suivi de l'évolution de la tumeur est réalisé par un suivi clinique et IRM. Des TEP sont également parfois proposées notamment pour le suivi après radiothérapie et témozolomide concomitant lorsqu'il est difficile de trancher entre effet secondaire (radionécrose) et progression précoce.

À l'heure actuelle, il n'y a pas de marqueurs biologiques pour le suivi des glioblastomes. Néanmoins, l'analyse des fluides biologiques apparaît comme une piste intéressante, pouvant potentiellement aider à la décision thérapeutique, au suivi de l'évolution de la maladie et à l'identification de résistances précoces aux traitements, voire au dépistage précoce dans les formes familiales ou chez les sujets à risque génétique ou après exposition à des radiations ionisantes.

Dans l'étude des tumeurs cérébrales, le liquide céphalo-rachidien (LCR) est le fluide biologique le plus proche de la tumeur et de ce qu'elle produit (sécrétome). Le LCR étant faiblement concentré en cellule, le sécrétome de la tumeur est facile d'accès et n'est pas masquée par des protéines abondantes. Ainsi, chez les patients atteints de glioblastomes, le liquide céphalo-rachidien montre des taux faibles de GAL, HSPA5 et WNT4 et un taux élevé de IL-6 par rapport aux gliomes de bas grades (Shen et al., 2014). Cependant, la réalisation d'une ponction lombaire est une procédure souvent contre-indiquée en raison d'un effet de masse chez les patients porteurs d'un gliome. De plus, il est impossible de collecter ce fluide de façon régulière et la très faible teneur en protéines du LCR en fait un échantillon difficile à étudier.

Le fluide biologique le plus accessible est le sang où circulent de nombreuses informations biologiques. Une analyse réalisée sur le sérum de 65 patients atteints de tumeur cérébrale a permis de mettre en évidence la surexpression d'EGFR chez les patients atteints de glioblastomes par rapport à celui des patients ayant un astrocytome anaplasique (Quaranta et al., 2007). Une autre étude a identifié 61 protéines impliquées dans les réactions inflammatoires et la progression tumorale en général dont l'expression est dérégulée dans le sang de patient atteint de glioblastome (Gautam et al., 2012).

Le sang est accessible très facilement et il est possible de réaliser plusieurs prises de sang au cours du temps ce qui en fait un outil indispensable pour le suivi du patient. Ainsi, le taux élevé de IGFBP-2 dans le plasma est un indicateur de la survie et de la rechute des patients (Lin et al., 2009) et plus récemment les taux plasmiques de CRP, C9 et LRG1 ont été démontrés comme positivement corrélés à la taille de la tumeur (Miyachi et al., 2018). La principale limite de l'analyse du sang réside dans la présence de protéines très abondantes qui peut empêcher la détection de protéines peu abondantes.

Finalement, il apparaît essentiel de découvrir de nouveaux biomarqueurs qui pourraient à la fois servir au diagnostic, guider le suivi mais aussi orienter la recherche vers de nouvelles cibles thérapeutiques. La recherche de biomarqueurs par analyse protéomique implique l'identification, la caractérisation et la quantification des protéines contenues dans un échantillon. La spectrométrie de masse s'est progressivement implantée comme la technique de choix pour ce type d'analyse.

II. Implication de la spectrométrie de masse pour l'analyse des cancers

Les techniques de spectrométrie de masse sont appliquées dans divers domaines du diagnostic médical comme l'identification microbiologique, la toxicologie ou l'immunologie. L'étude des protéines par spectrométrie de masse est également un domaine intéressant pour mieux comprendre les dysfonctionnements cellulaires au sein d'une pathologie.

A. La spectrométrie de masse

La spectrométrie de masse (MS) est la technique analytique centrale pour la recherche et l'étude des protéines et des biomolécules en général. La MS permet de détecter, d'identifier et de quantifier des molécules d'intérêt par mesure de la masse. Les ions formés sont séparés en fonction de leur rapport masse/charge (m/z). Un spectromètre de masse se compose de trois parties : une source d'ionisation, un analyseur de masse et un détecteur. La source d'ionisation permet la création d'ions en phase gazeuse, donc chargés. Les sources les plus utilisées sont les sources par électro-nébulisation (*Electrospray Ionization*, ESI) et la Désorption/Ionisation Laser Assistée par Matrice (*Matrix Assisted Desorption Ionization* MALDI). Les ions en phase gazeuse sont injectés dans l'analyseur de masse pour être séparés en fonction de leurs rapports masse/charge. Par exemple dans un analyseur temps de vol (*Time of Flight*, TOF), après avoir été accélérés dans un champ électrique, le temps de déplacement des ions jusqu'au détecteur est mesuré pour permettre la détermination de la masse.

Cette technique est particulièrement intéressante pour l'identification et la quantification de protéines et la détermination de leurs structures. Dans ce but, différentes stratégies existent dont les approches dites « bottom-up » qui sont les plus courantes. Ces approches impliquent le clivage enzymatique de la protéine entière pour générer un mélange de peptides. Généralement, les peptides sont séparés par chromatographie en phase liquide (LC) avant d'être analysés. Les peptides sont séparés suivant leur interaction avec une phase stationnaire selon leur degré d'hydrophobie avec un gradient croissant de solvant apolaire. Suivant leur temps de rétention, les peptides injectés dans le spectromètre de masse par un couplage avec une source ESI. Après une première mesure, les signaux les plus intenses (mode d'acquisition dépendant des données (*Data Dependant Acquisition*, DDA) sont fragmentés les

uns à la suite des autres, pour obtenir des informations sur leurs structures. Pour les peptides, cela va permettre de déterminer leur séquence en acides aminés et remonter ainsi à la protéine d'origine. Cette identification se fait par le biais de moteurs de recherche contre une banque de données regroupant les séquences des protéines connues. Des méthodes de quantification relatives vont permettre de comparer les abondances de chaque protéine entre différents échantillons. Ces techniques peuvent être réalisées avec ou sans marquage isotopique.

La spectrométrie de masse permet donc d'identifier et d'analyser un grand nombre de protéines contenues dans un échantillon biologique et par conséquent permet de comprendre de nombreux processus biologiques dont les mécanismes de tumorigenèse.

B. Imagerie par spectrométrie de masse MALDI

1. Principe

La source de production d'ions MALDI a émergé dans les années 1980. Elle permet une ionisation douce des molécules, c'est-à-dire peu de fragmentation, et donc une énergie transférée faible pour obtenir des ions en phase gazeuse à partir de l'état solide (Karas and Hillenkamp, 1988). L'interaction entre le faisceau laser et la matrice, contenant les composés à analyser, entraîne une éjection de matière permettant un passage en phase gazeuse. De plus, les composés vont acquérir une charge permettant leur séparation dans l'analyseur de masse associé. La matrice est un petit composé chimique absorbant à la longueur d'onde du laser utilisé. La majorité des matrices ont été déterminées de manière empirique, les premières utilisées sont les acides nicotiques, acides caféiniques et acide vanillique (Hillenkamp and Karas, 1990). La matrice, qui est en large excès, est co-cristallisée avec l'analyte avant d'être irradiée par le faisceau laser. Selon le type de matrice utilisée, différents types de composés pourront être détectés.

Une nouvelle application de la source MALDI a été développée. L'imagerie MALDI consiste à cartographier une surface comme une coupe de tissu en mode point par point. Chaque point détermine une coordonnée (x,y) et un spectre de masse est enregistré. L'échantillon est ensuite déplacé d'une distance définissant la résolution spatiale de l'image.

Un nouveau spectre est alors enregistré et l'opération est répétée pour couvrir la totalité de la surface à analyser. L'ensemble de ces spectres permet ensuite de reconstituer des « cartes de densité ionique » de la distribution des différents composés au sein de la coupe de tissu. Pour chaque composé de rapport m/z spécifique, les intensités de chaque spectre sont enregistrées avec leurs coordonnées sur le tissu. Les variations d'intensité sont reportées sur une échelle de couleur, de plus il est également possible de visualiser plusieurs composés sur une même image afin d'apprécier leurs localisations respectives (Figure 0-2). Ces dernières années, l'imagerie par spectrométrie de masse MALDI est devenue un outil important pour la recherche par sa capacité à localiser des centaines de molécules sans a priori et sans marquage préalable. Une grande diversité de biomolécules endogènes ou exogènes, telles que des métabolites, des lipides, des peptides et des protéines peuvent être détectés dans des échantillons biologiques divers (plantes, tissus, organoïdes, cheveux) (Buchberger et al., 2018).

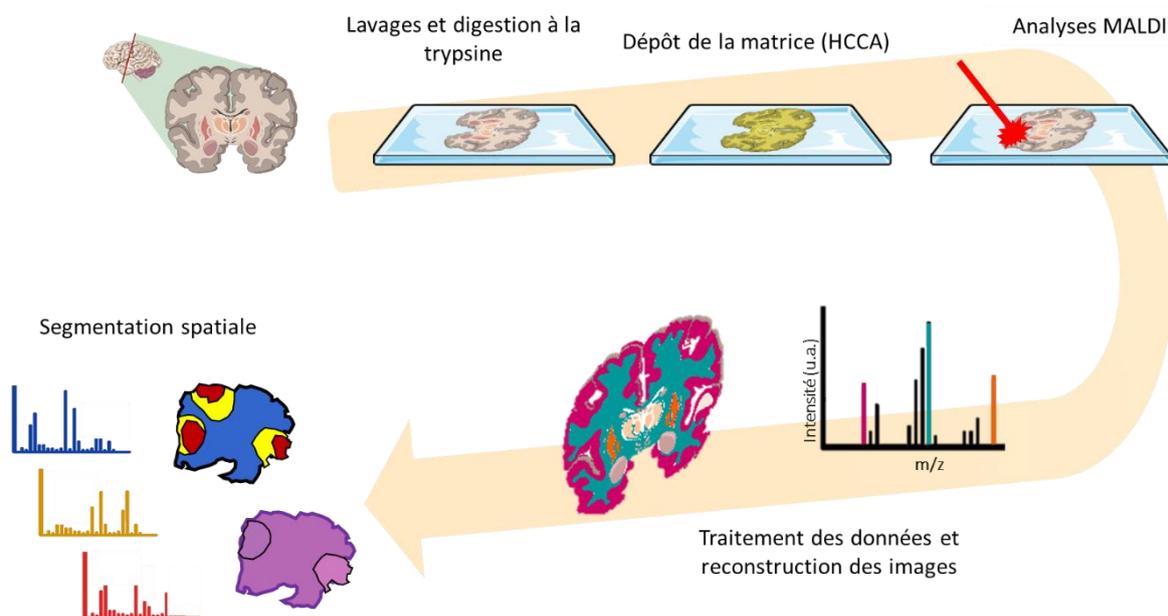


Figure 0-2 Principe général de l'imagerie MALDI

Après avoir réalisé des coupes de tissus congelés et les avoir transférés sur lames conductrices ITO, les sections sont lavées pour retirer les lipides et la matrice est déposée. Le laser du spectromètre de masse balaye la surface du tissu et des spectres de masse sont enregistrés pour chaque point. Les spectres sont ensuite enregistrés avec leurs coordonnées sur le tissu et les images moléculaires peuvent être générées.

Pour l'analyse MALDI, la préparation des échantillons est une étape cruciale. L'épaisseur des tissus est souvent comprise entre 7 et 12 μm . Il est nécessaire de déposer le tissu sur une lame conductrice pour éviter l'accumulation de charge pendant l'analyse. La propriété conductrice des lames est indispensable pour l'analyse avec un analyseur TOF, des différences de tension sur la surface de l'échantillon peuvent induire des variations dans le champ électrique et donc dans l'accélération des ions ce qui induit des modifications dans les mesures de temps de déplacement et fausser l'analyse. En ce sens, on peut recouvrir d'or la lame ou utiliser des lames ITO (indium tin oxide - ITO) qui sont des lames de verre rendues conductrice par dépôt d'une fine couche d'oxyde d'étain et d'indium.

Une fois des coupes d'une épaisseur comprise entre 7 et 12 μm obtenues, un étalement sur lame conductrices est effectué et différentes étapes de préparation sont réalisées. Pour les tissus FFPE, en plus des étapes de déparaffinage, une étape de démasquage des antigènes est indiquée pour améliorer l'accessibilité des protéines et faciliter leur digestion enzymatique. Pour les tissus frais, l'utilisation de différents lavages permet d'améliorer le signal avec en premier lieu des lavages avec des solutions de concentration différentes d'éthanol pour éliminer les sels présents, puis des bains d'acétone peut également être indiqué pour précipiter les protéines à la surface du tissu et également le chloroforme afin de retirer les lipides qui pourraient masquer le signal protéique (Lemaire et al., 2006a). Ces lavages n'engendrent pas de délocalisation des analytes.

Les tissus frais congelés sans fixation chimique permettent d'analyser de nombreuses molécules. Cependant il est difficile et parfois impossible d'obtenir des échantillons frais, la plupart des échantillons conservés étant des tissus fixés au paraformaldéhyde puis inclus dans la paraffine (FFPE). Sur les tissus FFPE, il est impossible de rompre les liaisons formées entre les protéines avec le faisceau laser. Cependant, le développement de nouvelles méthodes a permis de rendre accessible l'utilisation de ces tissus, notamment via l'analyse des peptides de digestion (Wisztorski et al., 2007). Pour ce type de préparation, le tissu est digéré grâce au dépôt de trypsine sur la surface de la coupe et après un temps d'incubation, la matrice est déposée. Cette technique est également applicable sur les tissus frais congelés et elle permet d'accéder à des protéines de plus hauts poids moléculaires.

Les matrices sont des molécules organiques acides de faible poids moléculaire et le choix de la matrice dépend de la nature des molécules à analyser. Par exemple, l'acide 2,5-dihydroxybenzoïque (DHB) est utilisé pour l'analyse des petites molécules comme les métabolites alors que l'acide α -cyano-4-hydroxycinnamique (HCCA) est plus destiné à analyser les peptides. L'acide sinapique (SA) est préféré pour la détection des protéines. La matrice HCCA couplée avec la base organique aniline (HCCA/ANI) a été développée au sein du laboratoire et permet d'obtenir des signaux intenses pour l'analyse des peptides (Lemaire et al., 2006a). Pour la réalisation d'image, la matrice est déposée à la surface du tissu grâce à un nébuliseur automatique. L'utilisation de robot permet une application homogène et reproductible pour recouvrir la totalité de la zone à analyser. La matrice est la plupart du temps déposée sous forme liquide en mélange avec des solvants organiques. Par conséquent, le dépôt ne doit pas être trop humide pendant le dépôt pour éviter la délocalisation des molécules d'intérêt sur le tissu. Il est également possible de déposer la matrice par sublimation pour une application homogène (Murphy et al., 2011).

Le dépôt solide obtenu est par la suite irradié par un laser ce qui permet la désorption/ionisation des molécules. En imagerie, les analyseurs TOF sont les plus couramment utilisés mais des couplages avec des analyseurs de type Orbitrap ou FTICR sont possibles et permettent d'obtenir un meilleur pouvoir résolutif et donc une meilleure séparation des ions proches en masse (Montero et al., 2020; Spraggins et al., 2016).

Les spectres sont ensuite enregistrés avec leurs coordonnées sur le tissu et les images de localisation peuvent être générées pour chaque m/z. Ces images représentent la localisation des molécules sur le tissu de façon non ciblée.

En ce qui concerne la résolution spatiale, la distance de déplacement entre deux points va définir la taille des pixels (ou voxels) de l'image. Cette résolution dépend de plusieurs paramètres dont la taille des cristaux de matrices et celle du faisceau laser. Actuellement des analyses peuvent couramment être réalisées autour de 50 μ m à 70 μ m. Néanmoins le développement de nouveaux instruments, notamment avec des avancées au niveau de la focalisation du faisceau laser, a permis d'augmenter la résolution spatiale jusqu'à 5 μ m (Feenstra et al., 2017). Ainsi, des images de plus en plus précises pourront à terme permettre l'analyse au niveau cellulaire (Ščupáková et al., 2020). (Figure 0-3)

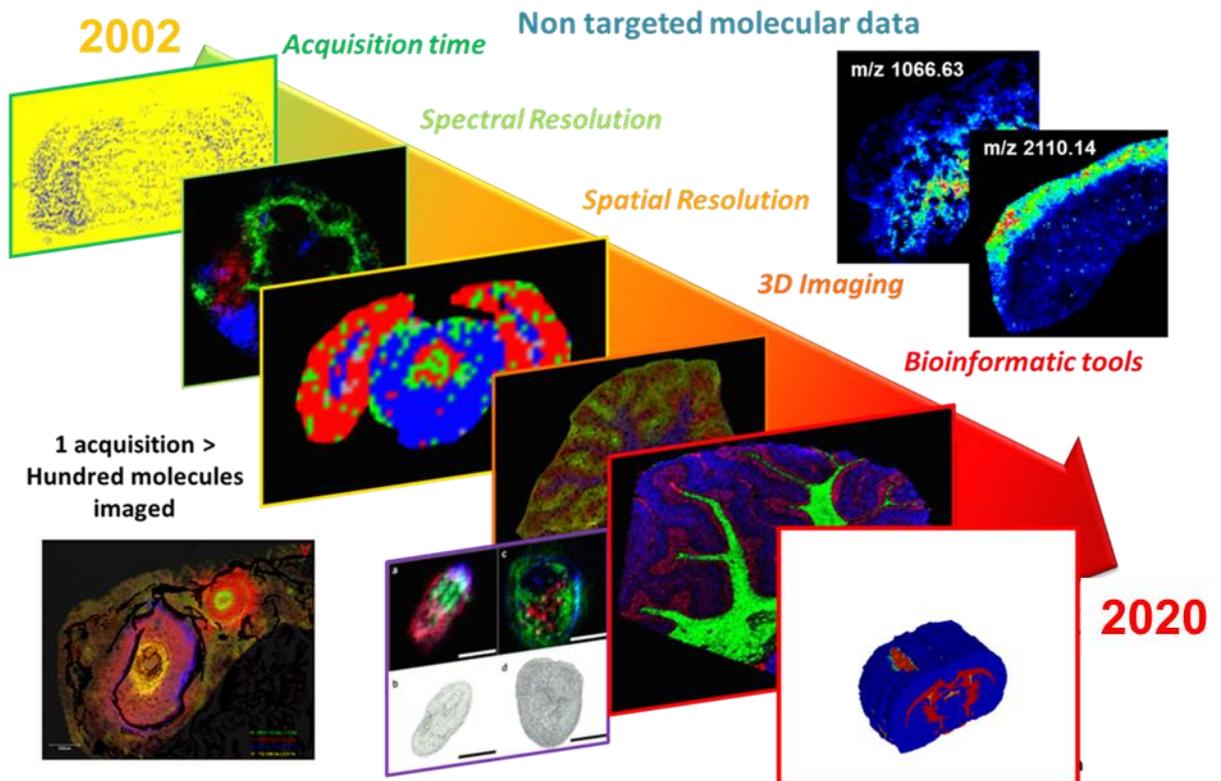


Figure 0-3 Evolution de la résolution spatiale des images MALDI
 Exemple de cerveau de rat de 2002 jusqu'à l'imagerie en 3 dimensions aujourd'hui

2. Traitements des données

Dans un premier temps le bruit de fond est soustrait et les spectres sont ensuite réalignés pour corriger les variations des pics qui peuvent être liés à la position du porte échantillon dans l'instrument et à la rugosité de la surface étudiée. Une étape de normalisation est réalisée pour corriger les variations d'intensité et diminuer les variations inter-spectres. La méthode de normalisation la plus utilisée est la normalisation du nombre total d'ion (TIC), qui normalise les spectres séparément en divisant l'intensité de chaque spectre par la somme de toutes ses intensités (Deininger et al., 2011). D'autres méthodes de normalisation existent, notamment le *Root Mean Square* (RMS), où chaque spectre est divisé par la racine carrée de la moyenne de toutes les valeurs d'intensité au carré. Il est également possible de normaliser en divisant chaque spectre par la médiane ou la valeur d'intensité maximale du spectre (Deininger et al., 2011).

Les données d'imagerie sont difficiles à exploiter notamment à cause du volume d'informations obtenu. En effet l'augmentation de la résolution spatiale a inmanquablement provoqué l'augmentation de la taille des fichiers, chaque pixel d'une image contenant un spectre de masse. Des stratégies de représentation et de réduction des données sont ainsi nécessaires.

Les analyses multivariées permettent la simplification des données d'imagerie. Dans un premier temps les spectres sont simplifiés par une représentation dans un espace multidimensionnel. Pour un spectre de masse contenant n pics, un espace à n dimensions est créé, chaque axe correspondant à un pic. La valeur d'intensité de chaque pic est ensuite reportée selon ces axes, ainsi l'ensemble des spectres contenus dans une image sont représentés par un nuage de point (Figure 0-4).

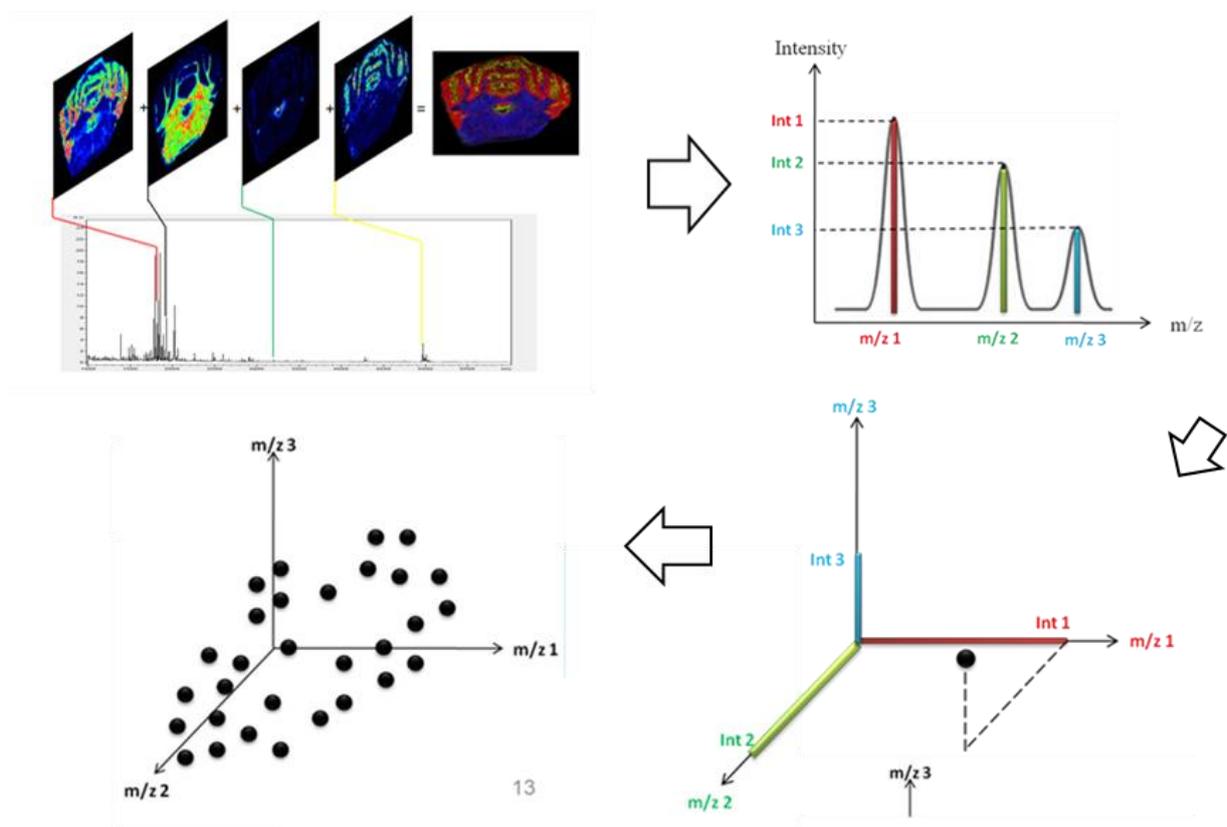


Figure 0-4 Simplification de spectres de masses par représentation dans un espace multidimensionnel

Un espace multidimensionnel peut être créé à partir des spectres générés dans une image MALDI. Cet espace permet la représentation des données sous forme de nuage de point.

Les spectres provenant de régions de compositions moléculaires particulières étant similaires, les points les représentant seront proches dans l'espace multidimensionnel. Par la suite, les axes de l'espace sont simplifiés par l'application de nouvelles coordonnées basées sur les composantes principales. La première composante principale (PC1) couvre la plus grande échelle de valeurs dans le nuage de points et ainsi de suite (PC2 ; PC3 ...). Les premières composantes permettent la meilleure séparation des spectres. Ensuite, à partir du *loading plot* qui représente la variance des m/z , il est possible d'identifier les m/z spécifiques des groupes de spectres d'intérêt.

Finalement des analyses statistiques par regroupement non supervisé permettent de définir des régions d'intérêt au sein du tissu. La segmentation spatiale regroupe les spectres similaires en régions d'intérêts, chaque groupe de spectre correspondant à une composition moléculaire particulière. La comparaison des spectres est faite à partir de l'espace à n dimensions créé. Les données peuvent être segmentées spatialement grâce à des algorithmes (Alexandrov et al., 2010). Par exemple, l'algorithme de segmentation « *bisecting k-means* » regroupe les spectres en k groupes, chaque groupe étant basé sur les distances spatiales entre les spectres. Cette méthode de groupement utilise une ligne bissectrice pour séparer le nuage de points initial en deux groupes distincts (Figure 0-5). Par la suite chaque groupe de points peut être séparé à nouveau selon le même principe.

Les résultats de segmentation sont exprimés sous la forme d'une arborescence (dendrogramme) où chaque branche correspond à un groupe de spectres regroupés par l'algorithme. Les dendrogrammes permettent de représenter graphiquement la séparation entre les groupes et une carte de segmentation du tissu est alors générée. La longueur des branches du dendrogrammes représente la distance entre les regroupements.

Sur la carte, chacun des groupes identifiés est représenté par une couleur suivant le dendrogramme. Ainsi, des régions présentant des signatures moléculaires différentes peuvent être mises en évidence. En ce sens, cette technique d'analyse est complémentaire à l'histologie. L'analyse de la composition moléculaire du tissu donne accès à un niveau d'information supplémentaire, les régions identifiées pouvant ne pas être différentiable d'un point de vue morphologique. On parle alors d'« histologie moléculaire ».

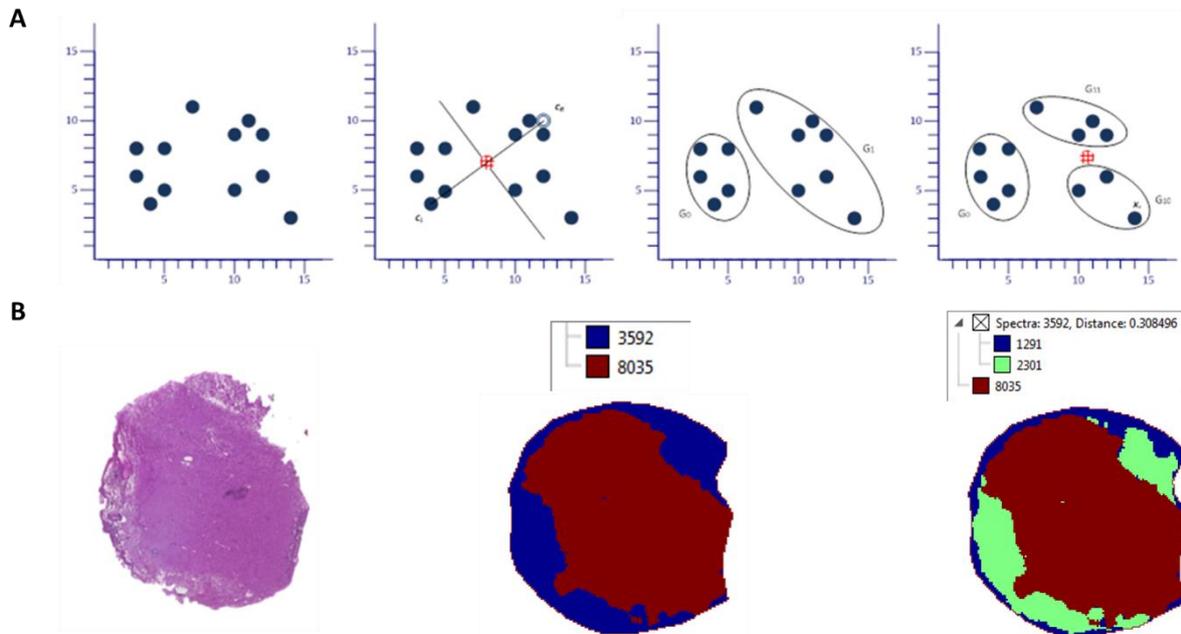


Figure 0-5 Principe de la segmentation *bisecting k mean* et représentation sur un tissu

A. Principe de la segmentation par regroupement des B. points Tissu cancéreux homogène d'un point de vue histologique et carte de segmentation de ce tissu après segmentation par l'algorithme. La distance entre chaque région est visible via l'arborescence, ici les régions bleu et verte sont similaires entre elles mais éloignées de la région rouge.

En clinique, cette technique permet de déterminer la complexité moléculaire qui peut exister au sein d'une tumeur par la distribution spatiale des molécules d'intérêt. Ainsi, l'imagerie par spectrométrie de masse MALDI a notamment un intérêt pour la recherche de biomarqueurs diagnostic et pronostic dans les cancers (Kriegsmann et al., 2018).

3. Intérêt pour l'analyse de tissus cancéreux

L'imagerie par spectrométrie de masse fait émerger de nombreuses perspectives en cancérologie clinique, notamment concernant le diagnostic, la détection précoce et le développement de thérapies personnalisées.

Les tumeurs sont des entités complexes qui dépendent des cellules environnantes, des vaisseaux sanguins et également des cellules du système immunitaire. L'homogénéisation de l'échantillon de tissu entraîne donc une perte d'informations importante. En effet, les tissus adjacents à la tumeur qui semblent histologiquement normaux peuvent déjà être compromis

sur le plan moléculaire. En ce sens, l'imagerie MALDI peut aider au diagnostic du pathologiste et donner des informations sur le pronostic du patient et le stade de développement de la tumeur. Elle permet de détecter les frontières de la tumeur et son infiltration et pourrait permettre d'établir un profil moléculaire spécifique suffisant pour diagnostiquer de façon précoce des lésions cancéreuses.

Pour les grades III et IV du cancer de l'ovaire, cette technique a permis de mettre en évidence un biomarqueur de m/z 9744 correspondant au fragment C-terminal de Reg- α (Lemaire et al., 2007). Ce marqueur est spécifique du cancer de l'ovaire et permet donc de détecter les marges de la tumeur. L'imagerie MALDI permet également de détecter les marges des régions traumatiques de blessures musculaires par rapport aux tissus adjacents (Figure 0-6) (Klein et al., 2014). De même, une équipe a montré la possibilité de discriminer un adénome de la glande saine adjacente en moins de 30 minutes, facilitant la préservation des tissus sains pendant la chirurgie (Calligaris et al., 2015).

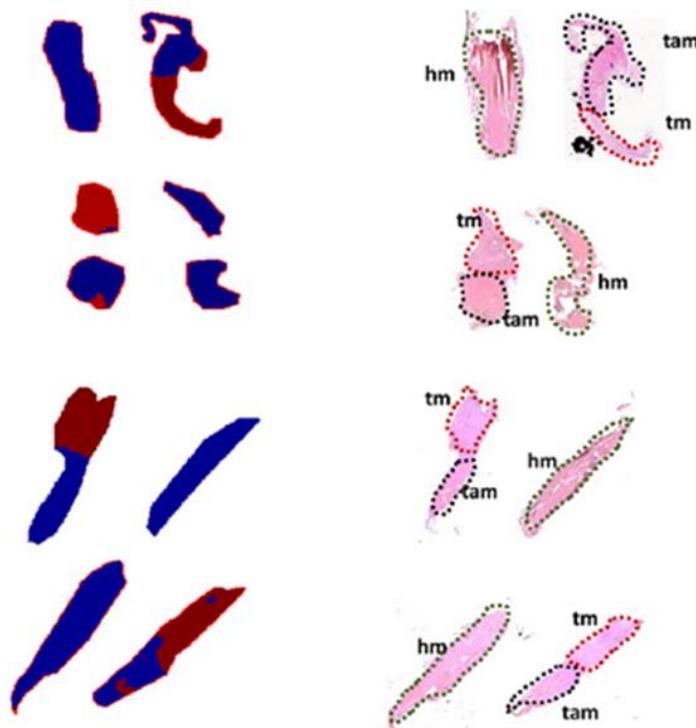


Figure 0-6 Segmentation spatiale des régions musculaires traumatisées (tm), adjacentes au traumatisme (tam) et saines (hm) pour chaque réplicat biologique

Coloration HE des tissus sains et traumatisés. La segmentation spatiale en deux clusters montre une discrimination claire entre les tissus sains (hm) et traumatiques adjacents (ta), qui sont représentés en bleu, de la région traumatisée (tm, rouge). Adapté de (Klein et al., 2014)

Il est également possible de déterminer le grade de malignité de la tumeur, par exemple les profils protéiques peuvent discriminer les gliomes de grades II, III et IV (Schwartz et al., 2004). Plus récemment, une équipe a identifié des signaux associés aux différents grades de carcinomes des cellules rénales (Stella et al., 2019). De plus la surexpression significative de 3 protéines dans les carcinomes hépatocellulaires par rapport aux tissus non cancéreux a permis de générer un modèle de classification prédictif des lésions bénignes et malignes du foie (Marquardt et al., 2015). De même un modèle prédictif de discrimination entre les adénocarcinomes canaux pancréatiques et les tumeurs neuroendocrines pancréatiques a récemment été publié (Casadonte et al., 2019). Ces classifications permettent de simplifier le diagnostic et d'initier rapidement et précisément les traitements des patients.

Les signatures moléculaires obtenues par imagerie peuvent également être utilisées pour classer des tumeurs différentes, d'un point de vue clinique cela permet d'identifier la source d'une tumeur métastatique quand aucune tumeur primaire apparente n'est présente. Ainsi, il a été montré que l'imagerie MALDI permet de discriminer différents types de tumeurs localisées sur des organes différents (Meding et al., 2012).

De plus, l'utilisation de l'imagerie MS peut également être intéressante pour étudier la pénétration d'un médicament et sa localisation dans la tumeur. Par exemple, une équipe a montré que les composés néphrotoxiques polymyxine B1 et colistine s'accumulaient préférentiellement dans la région corticale du rein tandis que le nonapeptide polymyxine B qui est moins toxique était réparti plus uniformément dans le rein (Nilsson et al., 2015). Plus récemment, la distribution et la concentration de paclitaxel ont été mesurées dans un modèle 3D de mésothéliome pleural malin, qui est connu pour être très hétérogène (Giordano et al., 2016). Dans cette étude le paclitaxel était retrouvé plus largement dans les zones de prolifération tandis que le signal était plus faible dans les régions nécrotiques (Figure 0-7). Ces résultats confirment que l'hétérogénéité d'un tissu impacte la distribution de médicaments et peut par conséquent affecter l'effet du traitement ce qui peut expliquer pourquoi certains patients répondent mieux à la chimiothérapie que d'autres.

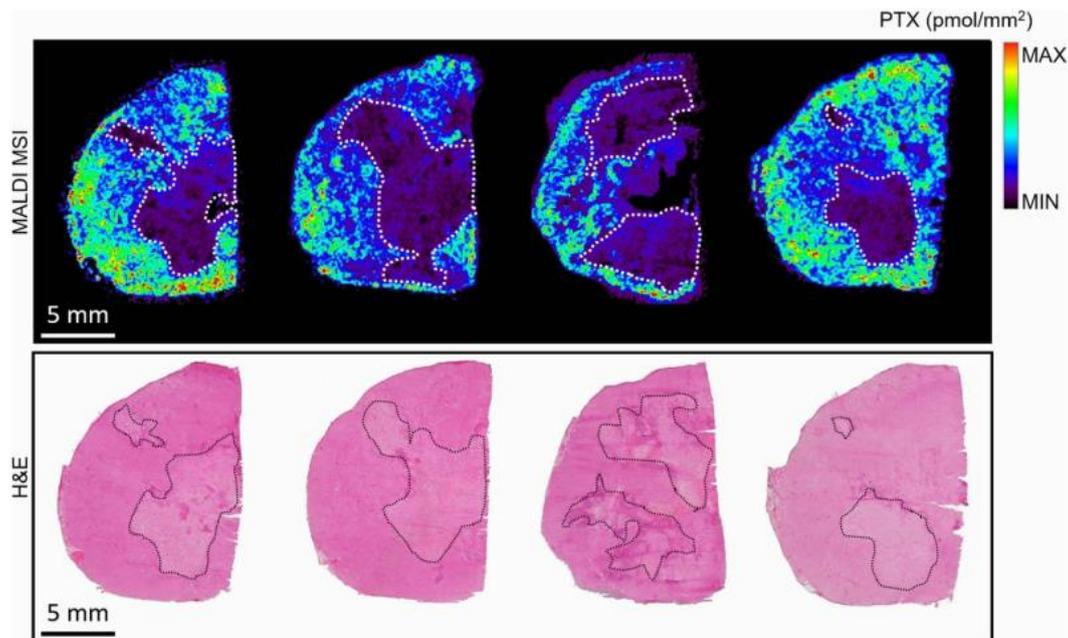


Figure 0-7 Distribution du paclitaxel par imagerie MALDI (panneau supérieur) et coloration H&E d'une coupe consécutive (panneau inférieur)

Les zones nécrotiques, en pointillées, sont celles où le signal médicamenteux est le plus faible. (Giordano et al., 2016)

Finalement l'imagerie de type MALDI donne la possibilité de localiser différentes molécules au sein d'un tissu et ainsi déterminer des signatures moléculaires selon l'hétérogénéité de la tumeur. Cette technique d'imagerie permet de mettre en évidence des marqueurs de phénotypes cellulaires particuliers associés à des mécanismes physiopathologiques mais ne permet pas d'identifier ces marqueurs. L'identification des protéines est possible via des analyses standardisées en MS/MS au moment de l'acquisition des images mais reste complexe et le nombre de protéines identifiées reste limité aux protéines les plus abondantes (Minerva et al., 2008; Stauber et al., 2010). En effet, l'identification des protéines présentes en faible quantité est difficile à cause de la difficulté à isoler des quantités suffisantes de protéine pour la fragmentation, en plus de l'effet de suppression des ions.

Au contraire, les techniques conventionnelles permettent d'identifier et de quantifier les protéines et par conséquent d'obtenir des informations indispensables à la compréhension des processus physiopathologiques. Pour accéder à l'identification des protéines situées dans

les régions d'intérêt mises en évidence par l'imagerie, de nouvelles stratégies se sont développées ces dernières années.

C. Protéomique localisée

1. Capture par microdissection laser

La microdissection laser a été l'une des premières technologies développées pour appréhender l'hétérogénéité des tissus, c'est une technique de préparation d'échantillon utilisant un microscope inversé et un laser infrarouge ou ultraviolet (Banks et al., 1999). Pour micro-disséquer un tissu, le laser suit une trajectoire prédéfinie par l'utilisateur. Ainsi, il est possible de cibler précisément des microstructures de tissus en fonction de la morphologie, de la coloration HES, de l'immunofluorescence ou de l'imagerie MS. Les cellules disséquées sont alors éjectées jusqu'à plusieurs millimètres et peuvent être récupérées dans le capuchon d'un microtube. La reproductibilité de cette technique n'est pas optimale, principalement en raison de l'expulsion de la zone sélectionnée qui engendre des pertes plus ou moins importantes (Longuespée et al., 2016). Récemment, cette technique a été combinée à la méthode NanoPOTS (Nanodroplet Processing in One pot for Trace Samples) qui consiste à récupérer les cellules éjectées dans une goutte de DMSO pour ensuite les transférer dans un nano puits et réduire les volumes de digestion au nanolitre, minimisant ainsi considérablement les pertes de protéines et de peptides (Figure 0-8) (Zhu et al., 2018a).

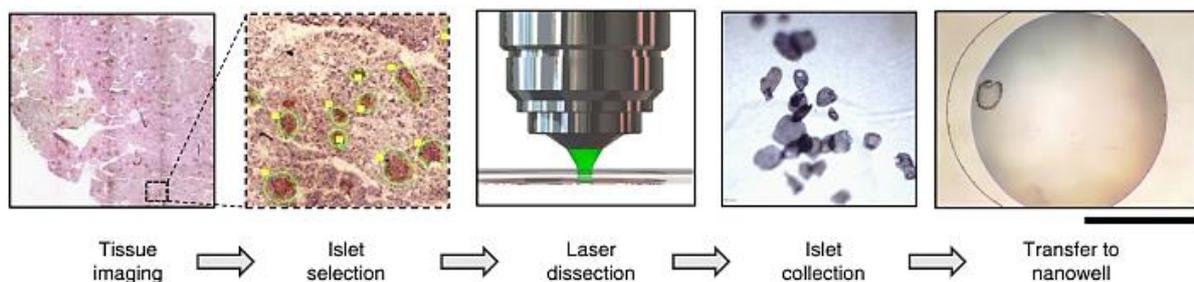


Figure 0-8 Workflow schématique de la microdissection laser suivie de la collecte et du transfert de cellules dans les nano puits.

Coupes de 10 µm d'épaisseur d'îlots individuels isolés par microdissection laser à partir de tissu pancréatique, toute la procédure de traitement nanoPOTS est effectuée dans une gouttelette de 200 nL ou moins. Adapté de (Zhu et al., 2018b)

Cette combinaison a permis l'identification de 180 protéines dans une région de 50µm, ce qui correspond à 2 – 6 cellules. Cette approche permet également de générer des images basées sur les quantifications relatives des protéines identifiées avec une résolution spatiale de 100 µm (Piehowski et al., 2020).

La capture par microdissection laser est une technique régulièrement utilisée dans l'étude des cancers. Cette combinaison est notamment intéressante dans l'analyse de l'hétérogénéité intra-tumorale (Longuespée et al., 2018). Ainsi, l'analyse de différentes zones du cancer de la prostate a révélé des mécanismes de communications cellulaires entre les cellules tumorales et leur microenvironnement (Pin et al., 2016). Plus récemment, la technique a permis d'isoler spécifiquement des fibroblastes associés au cancer (CAFs) et sains et par conséquent d'isoler des facteurs clés de l'implication des CAFs dans la progression tumorale (Chatterjee et al., 2019). La microdissection laser peut également être réalisée directement après une analyse MALDI, sur la même coupe de tissu (Dilillo et al., 2017).

2. Digestion in situ et extraction par microjonction liquide (*Liquid Extraction Surface Analysis* LESA)

De nombreuses techniques de microéchantillonnage se développent pour accéder à l'identification à grande échelle des protéines dans les régions d'intérêts mises en évidence par imagerie. En ce sens, la microdigestion *in situ* à l'aide d'une imprimante chimique au niveau des régions d'intérêts est intéressante. Cette approche permet de déposer de la trypsine directement sur le tissu pour le digérer de façon très localisée (<300µm) et reproductible. Une fois les zones d'intérêts définies, quelques picolitres d'une solution contenant une enzyme de digestion (le plus souvent de la trypsine) sont déposés à intervalle régulier sur la coupe pendant 2 heures afin de maintenir une goutte sur la zone à digérer. Par la suite, il est possible d'extraire les peptides de digestion par microjonction liquide (LESA) grâce à différents solvants (Figure 0-9) (Quanico et al., 2013). Ces peptides peuvent par la suite être analysés par spectrométrie de masse LC-MS/MS en vue de leur identification par interrogation des banques de données. Cette stratégie permet d'identifier plus de 1300 protéines dans une surface de digestion de 650µm correspondants environ à 1900 cellules de façon reproductible (Quanico et al., 2013).

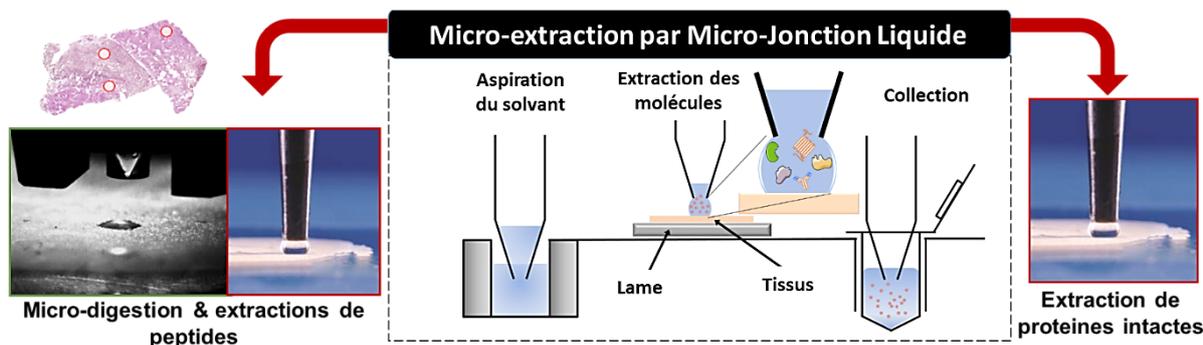


Figure 0-9 Principe de l'extraction par micro-jonction liquide

Les régions d'intérêt sont microdigérées par dépôt de trypsine et les peptides contenus dans ces régions sont micro-extraits puis analysés par NanoLC-MS/MS. Les protéines intactes peuvent être extraites directement à partir de tissus frais. Adaptée de (Wisztorski et al., 2017)

Il est également possible d'extraire directement les protéines intactes ce qui permet d'avoir accès à leurs modifications post-traductionnelles (Delcourt et al., 2018a; Wisztorski et al., 2016). Récemment, il a été démontré qu'il était possible d'extraire et d'analyser des protéines intactes de taille supérieure à 70 kDa (Griffiths et al., 2019).

Cette stratégie a notamment été utilisée pour la caractérisation de tumeurs de l'ovaire et a permis d'identifier différentes protéines spécifiquement surexprimées et sous-exprimées dans les régions tumorales (Quanico et al., 2017a; Wisztorski et al., 2013).

De plus, une étude réalisée au laboratoire sur 10 gliomes de grade III a démontré l'intérêt de la microjonction liquide combinée à l'imagerie par spectrométrie de masse pour établir une classification des gliomes (Le Rhun et al., 2016). Cette étude a pu fournir de nouvelles informations sur l'organisation des gliomes de grade III et pourrait permettre l'élaboration d'une classification plus précise de ces tumeurs. Le projet de thèse présenté dans ce manuscrit se place dans la continuité de cette étude pilote (voir CHAPITRE 1).

3. Microdissection assistée par parafilm

La microdissection assistée par parafilm est une stratégie développée au laboratoire qui permet de découper manuellement une région d'intérêt sur une coupe de tissu et de réaliser diverses analyses à partir de ce fragment (Franck et al., 2013; Quanico et al., 2015). Pour cela les coupes de tissus sont étalées sur une lame de verre recouverte de parafilm ce

qui permet la microdissection manuelle des ROIs au microscope à l'aide d'un scalpel. Les ROIs auront été préalablement identifiées en soumettant une coupe de tissu consécutive à une analyse MALDI MSI (Quanico et al., 2017b). Après excision, les protéines sont extraites et digérées pour procéder à leurs identifications par chromatographie liquide et interrogation des banques de données.

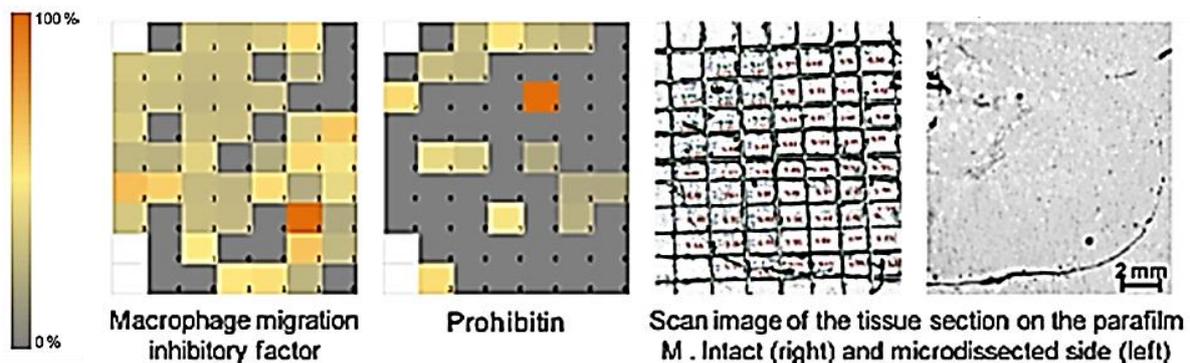


Figure 0-10 Microdissection assistée parafilm

Image de la section de cerveau de rat dont la moitié a été soumise à une dissection assistée par parafilm et reconstruction de la distribution de deux protéines représentatives basées sur le nombre de spectre identifiés par l'analyse nanoLC-MS / MS. Adapté de (Franck et al., 2013)

Cette approche peut être appliquée à des coupes de tissus entières dans le but de reconstruire des images basées sur les quantifications relatives obtenues à partir des données d'identification (Figure 0-10) (Franck et al., 2013). Cette approche est compatible avec l'analyse des lipides et des métabolites simultanément à celle des protéines (Quanico et al., 2017c). Il est également possible d'adapter cette méthode à la recherche de biomarqueurs, en ce sens des biomarqueurs ont déjà pu être identifiés dans le cancer de la prostate, en plus des marqueurs déjà connus (Quanico et al., 2015).

4. Digestion des protéines par migration passive dans un hydrogel

Une autre approche basée sur des hydrogels contenant de la trypsine a été mise au point (Harris et al., 2013). Ces hydrogels sont déposés directement sur la coupe de tissu et la digestion des protéines se fait après diffusion passive dans l'hydrogel (Figure 0-11). Les peptides de digestion sont ensuite extraits de l'hydrogel pour être analysés par LC-MS. Avec cette technique, plus de 600 protéines ont été identifiées dans un hydrogel de 260 µm de

diamètre (Rizzo et al., 2017). Il est cependant difficile d'obtenir et de manipuler des hydrogels de quelques millimètres de diamètre.

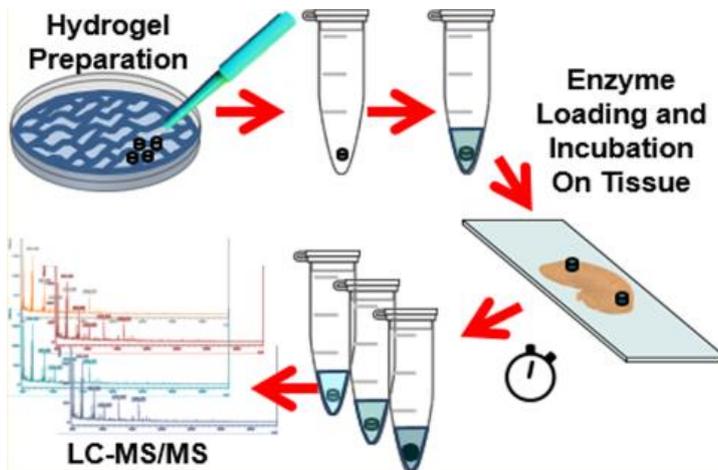


Figure 0-11 Principe de la digestion des protéines par migration passive dans un hydrogel

Les hydrogels sont découpés à l'aide d'emporte-pièce pour biopsies de différents diamètres puis incubés avec de la trypsine. Les hydrogels sont ensuite placés sur des coupes tissues pour digérer spécifiquement des régions d'intérêt puis analysés par LC-MS/MS. Adapté de (Rizzo et al., 2017)

Cette technique est assez récente mais a néanmoins déjà été utilisée sur des coupes de tissus frais congelés et des tissus FFPE de myxome cardiaque et a permis de différencier des régions du tissu grâce aux nombreux signaux protéiques identifiés (Taverna et al., 2019).

Finalement, l'ensemble de ces stratégies permet d'identifier de nombreuses protéines de façon très localisée et donc de mieux appréhender les différents processus physiopathologiques, notamment dans l'étude des tumeurs et du microenvironnement tumoral. L'identification de sous-type tumoral est également possible et ainsi différents biomarqueurs pourraient être identifiés et apporter des informations pour aider aux décisions thérapeutiques et améliorer les connaissances concernant le pronostic. Ces informations reposent cependant sur l'analyse du tissu tumoral et ne sont pas adaptées pour le suivi du patient. Aucun biomarqueur n'est actuellement identifié pour surveiller la progression tumorale et la réponse au traitement. Ainsi, détecter ce que les cellules tumorales produisent en temps réel est un objectif majeur. Dans cette optique, l'étude des vésicules extracellulaires (VEs) est primordiale.

III. Les vésicules extracellulaires (VEs)

Les vésicules extracellulaires transportent ce que les cellules tumorales sécrètent et ce que les cellules environnementales produisent en réponse. Retrouvées facilement dans les fluides biologiques, l'analyse des VEs permet d'obtenir des informations sur la tumeur elle-même. Le terme de vésicules extracellulaires caractérise un groupe hétérogène comprenant les corps apoptotiques, les exosomes et les microvésicules. Considérées comme un moyen de communication entre les cellules, les VEs sont impliquées dans de nombreux processus physiologiques et physiopathologiques (Becker et al., 2016; Bronisz et al., 2016; Gourlay et al., 2017; Zhao et al., 2018).

A. Généralités

1. Les différents types de vésicules extracellulaires

Les vésicules extracellulaires (VEs) regroupent plusieurs sous-populations actuellement classées en fonction de leurs biogenèses et de leurs tailles.

On distingue parmi les VEs les microvésicules (aussi appelées ectosomes ou microparticules) qui sont issues du bourgeonnement extérieur de la membrane plasmique de la cellule. La formation de ces bourgeons est due aux réarrangements de certains phospholipides, dont la translocation de la phosphatidylsérine à la membrane externe, et aux contractions des protéines du cytosquelette (Abels and Breakefield, 2016). Les microvésicules observent des tailles comprises entre 50 μ m et 1 μ m.

Les exosomes dérivent de l'endosome et impliquent la formation de corps multivésiculaires. Ils sont libérés par la fusion du corps multivésiculaire avec la membrane plasmique de la cellule (Abels and Breakefield, 2016). Ces vésicules peuvent être formées à la suite de la réorganisation de la membrane endosomale de telle sorte qu'elle s'enrichit fortement en tétraspanines (Pols and Klumperman, 2009). Le deuxième mécanisme conduisant à la formation de ces vésicules implique la machinerie des protéines ESCRT (*Endosomal Sorting Complex Required for Transport*). Les exosomes sont plus petits que les microvésicules, leur taille étant généralement comprise entre 30nm et 100nm.

Les vésicules extracellulaires incluent également les corps apoptotiques qui sont formés quand la cellule rentre en mort cellulaire programmée.

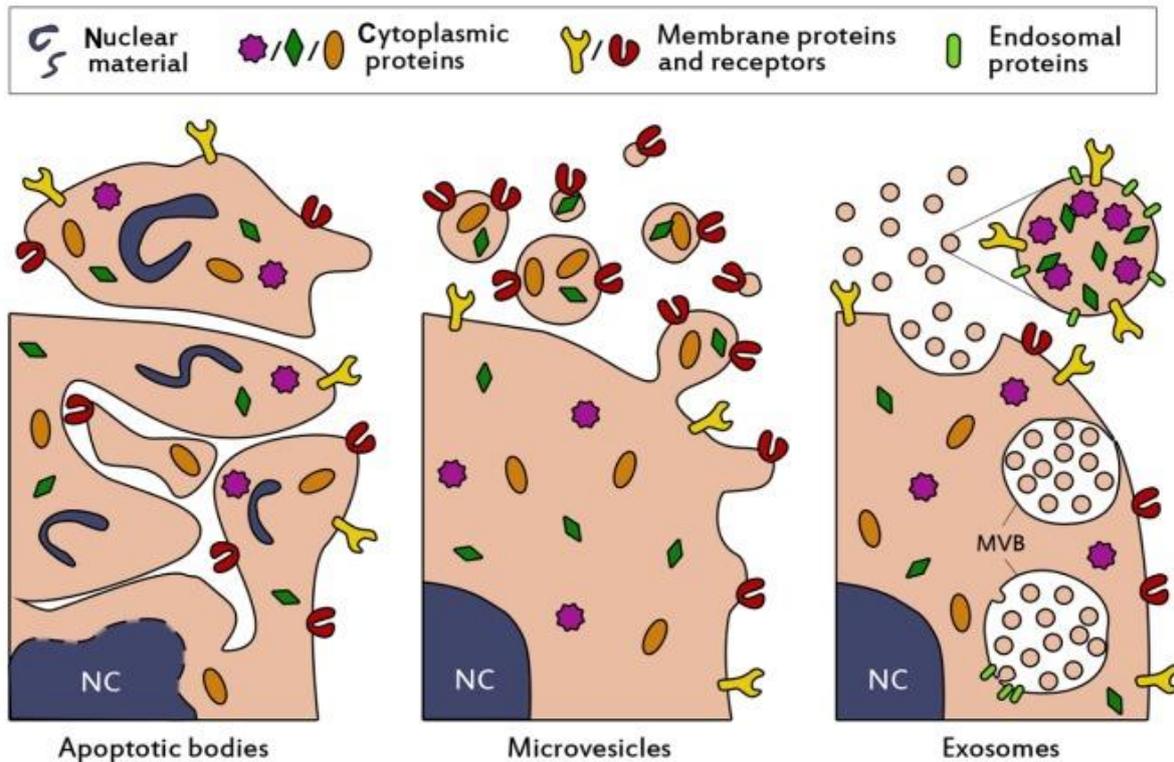


Figure 0-12 Représentation schématique de la biogenèse et de la composition des trois principales classes de vésicules extracellulaires

Représentation schématique de la biogenèse et de la composition des trois principales classes de vésicules extracellulaires. Les corps apoptotiques (gauche) se forment lorsque les cellules entrent en apoptose. Les microvésicules (milieu) sont formées par bourgeonnement et fission de la membrane plasmique. Les exosomes (droite) proviennent de structures endosomales appelées corps multivésiculaires (MVB). NC = Noyau (Kooijmans et al., 2012)

Les vésicules extracellulaires sont des actrices clés de la communication cellulaire. Leur contenu est fortement influencé par leur biogenèse ainsi que par leurs origines cellulaires et les conditions physiologiques ou pathologiques. D'une manière générale, les vésicules peuvent contenir des protéines, des lipides et des acides nucléiques (ADN et ARN). L'ensemble du contenu vésiculaire identifié a été répertorié dans plusieurs bases de données, dont Vesiclepedia (<http://microvesicles.org/>) et Exocarta (<http://www.exocarta.org/>). Ainsi, un total de 349 988 protéines, 27 646 mARN, 10 520 miARN et 639 lipides a été identifié dans les vésicules provenant de 41 espèces différentes sur un total de 1 254 études selon Vesiclepedia

(données de juin 2020). Ces données montrent le contenu riche et diversifié des vésicules extracellulaires.

Parmi les protéines contenues dans ces vésicules on retrouve notamment celles associées à leurs biogenèses comme les tétraspanines (CD63, CD81 et CD9) ou les protéines de la machinerie ESCRT comme Alix et TSG101 (Choi et al., 2015; Raposo and Stoorvogel, 2013). Concernant les lipides, notamment impliqués dans la régulation des membranes, on retrouve principalement les sphingolipides, la phosphatidylsérine, le cholestérol et la céramide (Durcin et al., 2017).

La découverte de matériel génétique dans les VEs a permis d'observer pour la première fois le transfert intercellulaire d'ARNm et de miARN (Valadi et al., 2007). En effet, ces acides nucléiques peuvent être transmis à une autre cellule et être fonctionnels dans ce nouvel emplacement, induisant ainsi une régulation de l'expression des gènes de la cellule receveuse. D'autres études ont également mis en évidence la présence d'ARN longs non-codants, d'ADN mitochondrial et d'ADN double brins (Kalluri and Lebleu, 2016).

Cependant, il reste difficile de déterminer précisément la composition spécifique des différents types de vésicules. En effet, il existe de grandes variations concernant les techniques d'isolement des vésicules et pas de méthodes standards, de plus elles proviennent souvent de sources différentes (cultures cellulaires, fluides biologiques, tissus ...). La principale technique utilisée actuellement est l'ultracentrifugation (UC) associée à un gradient de densité ou une chromatographie d'exclusion stérique. Il existe également plusieurs kits commerciaux disponibles, la plupart reposant sur la précipitation ou l'immunocapture des vésicules. Ces approches doivent être prises avec prudence, car elles ne parviennent souvent pas à différencier les VEs de tailles différentes et les agrégats protéiques. En plus des vésicules extracellulaires, les fluides biologiques contiennent également des concentrations élevées de protéines et de lipides. La présence de ces molécules complique la purification des vésicules.

Il est donc crucial de normaliser les méthodes d'isolement et d'analyse de l'ensemble de ces vésicules, le terme VEs devant être privilégié à moins d'identifier des marqueurs spécifiques permettant de classer les différents sous-types.

2. Rôles dans les cancers

Les vésicules extracellulaires sont sécrétées à la fois par les cellules saines et cancéreuses, elles sont par conséquent impliquées à plusieurs niveaux dans la tumorigenèse.

Les VEs sont d'importantes médiatrices de la communication cellulaire dans les cancers, notamment par le transport d'oncogène et de microARN entre les cellules tumorales (Bronisz et al., 2016; Minciacchi et al., 2015). Particulièrement, le microARN miR-21 a été identifié comme abondant dans les vésicules de gliomes et directement impliqué dans la progression et l'invasion tumorale (Santiago-Dieppa et al., 2014). De plus, il a été démontré que les cellules de gliomes exprimant la protéine oncogène EGFRvIII sécrètent des microvésicules contenant EGFRvIII pour la transférer à des cellules cancéreuses EGFRvIII négatives dans la même tumeur primaire (Al-Nedawi et al., 2008). Dans les cellules réceptrices, EGFRvIII va finalement induire l'accélération de la croissance tumorale.

De la même manière, les cellules cancéreuses peuvent communiquer avec des cellules saines. Les VEs dérivées de tumeurs peuvent modifier fonctionnellement les fibroblastes en les convertissant en myofibroblastes contribuant ainsi à la vascularisation et la croissance tumorale (Webber et al., 2010). Elles peuvent également avoir des propriétés immunorégulatrices, notamment en augmentant la prolifération et l'activité immunosuppressive des lymphocytes T régulateurs (Szajnik et al., 2010).

Aussi, la communication longue distance des VEs permet à la tumeur de s'éloigner de son foyer initial. Les VEs issues de cellules tumorales dressent un microenvironnement favorable pour générer des métastases de façon non aléatoire, via notamment l'expression d'intégrines spécifiques (Hoshino et al., 2015).

Le rôle des exosomes dans le transfert de résistance aux médicaments est également étudié. Les VEs issues de lignées cellulaires de cancer du sein peuvent transférer leur résistance au docétaxel à des cellules sensibles à ce médicament (Lv et al., 2014).

Les VEs reflètent à la fois ce que les cellules cancéreuses produisent, mais aussi ce que le microenvironnement sécrète en réponse. Les vésicules extracellulaires représentent donc

une source d'informations importante pour la prise en charge des patients et pouvant être obtenues à partir de différents fluides biologiques comme le sang ou le LCR.

B. Intérêts des vésicules extracellulaires dans la prise en charge des cancers

Les vésicules extracellulaires représentent une nouvelle opportunité pour la prise en charge des patients atteints de cancer.

1. Diagnostic

En tant que marqueurs diagnostic, les VEs sont disponibles facilement et fournissent des informations importantes concernant la progression de la pathologie. Les vésicules sont retrouvées dans tous les fluides biologiques comme le sang, le liquide céphalo-rachidien, le lait maternel et l'urine (Raposo and Stoorvogel, 2013). Elles sont donc accessibles de façon non invasive pour le patient. De plus, leur contenu en protéine et acide nucléique est stable, car protégé des protéases et autres enzymes par la membrane bi lipidique.

De nombreuses études ont mis en évidence l'intérêt du contenu vésiculaire pour la détection précoce du cancer. Ainsi, le microARN mir-21 contenu dans les VEs du LCR pourrait être utilisé pour identifier les patients atteints d'un glioblastome des patients sains (Akers et al., 2013). De la même manière la protéine EGFRvIII issue des VEs du sang pourrait être un marqueur diagnostic pour les patients atteints de glioblastome (Al-Nedawi et al., 2008; Skog et al., 2008).

De plus, les VEs représentent un outil potentiellement important quand la tumeur est difficilement ou pas accessible. Elles pourraient également aider à déterminer le sous-type tumoral. Pouvoir accéder à ce que produit la tumeur par les VEs, sans réaliser de biopsies, serait un progrès majeur dans la prise en charge des patients. Récemment, l'analyse protéomique des VEs issues de 10 lignées cellulaires de cancer du sein a montré une concordance avec leurs sous-types moléculaires respectifs (Rontogianni et al., 2019).

Enfin, la facilité d'accès aux VEs dans les fluides biologiques, notamment dans le sang, aurait un réel intérêt pour le dépistage de certains cancers, a fortiori dans les populations

asymptomatiques et les population à risque à cause de formes familiales ou d'exposition à des radiations (Sandfeld-Paulsen et al., 2016; Taylor and Gerchel-Taylor, 2008).

2. Évaluation de la réponse aux traitements

En plus du diagnostic, il est nécessaire de pouvoir contrôler la réponse du patient au traitement, surveiller l'apparition éventuelle d'un mécanisme de résistance ou évaluer la maladie résiduelle. Ce suivi est possible grâce aux VEs et à l'accès facile aux fluides biologiques du patient.

Par exemple, dans le cancer du pancréas les variations du niveau d'EpCAM intravésiculaire pendant le traitement sont associées aux pronostics des patients (Giampieri et al., 2019). Les patients avec un taux augmenté de EpCAM pendant le traitement avaient une meilleure survie tandis que ceux dont le taux baisse ont une survie plus faible.

Dans le mélanome, le taux de PD-L1 exosomal avant et pendant le traitement au pembrolizumab pourrait refléter l'activité anti-tumorale et par conséquent servir à prédire les résultats cliniques d'un traitement anti-PD-1 (Chen et al., 2018). Connaître rapidement les effets d'un traitement avant les premiers signes cliniques pourrait permettre d'anticiper et d'améliorer les décisions thérapeutiques. Des études cliniques utilisent les exosomes dans ce sens. Par exemple, une étude a montré que les variations d'expression d'EGFR observées dans les cellules cancéreuses de poumon en réponse au traitement sont également observées dans les exosomes circulants dans le sang des patients (Yu et al., 2017). Dans le cancer de la prostate, le contenu moléculaire des vésicules sécrétées par les cellules tumorales peut également être utilisé comme biomarqueurs pharmacodynamiques de la réponse au traitement (Kharaziha et al., 2015).

Ces résultats prometteurs nécessitent d'être confirmés dans de larges essais cliniques.

3. Stratégies thérapeutiques

Si les VEs ont un intérêt diagnostique et pronostic, il est également possible de les utiliser en tant que thérapie. Il existe deux approches différentes : « *Use them or target them ?* ».

Dans un premier temps, il paraît intéressant de les cibler pour entraver la communication cellulaire qui permet à la tumeur de progresser. Il est possible de limiter leurs assimilations par la cellule réceptrice grâce à des inhibiteurs d'endocytose. Par exemple le traitement à la Cytochalasin D conduit à une réduction de l'absorption des vésicules de 36 % (Escrevente et al., 2011). De même, le Dynasore peut interférer avec l'internalisation des VEs dans les cellules réceptrices (Macia et al., 2006).

Avant même d'impacter la réception des vésicules, il est possible d'impacter leurs biogénèses. Le *knockdown* de Rab27a réduit de 50 % la production d'exosomes des cellules de mélanome conduisant à la réduction de la croissance tumorale (Peinado et al., 2012).

Dans un second temps, les VEs peuvent être utilisées comme vecteurs de traitements. La capacité de ces vésicules à traverser la barrière hématoencéphalique étant particulièrement intéressante pour la thérapie de tumeurs cérébrales.

Les VEs peuvent cibler spécifiquement les cellules tumorales et peuvent être chargées de drogues ou anticorps. Récemment, il a été montré que l'effet antitumoral du paclitaxel pouvait être amélioré quand il était transporté dans des exosomes de macrophages pro-inflammatoires (Wang et al., 2019). De même, l'effet de la doxorubicin est augmenté quand elle est véhiculée par des exosomes dérivées de cellules tumorales (Hadla et al., 2016).

Pour l'analyse des gliomes, l'étude de ces vésicules donne accès à de nombreuses informations sans avoir besoin de réaliser de chirurgie. Il paraît envisageable dans le futur d'intégrer l'analyse des VEs pour le diagnostic précis des tumeurs cérébrales. De plus, l'analyse des composants contenus dans les vésicules pourrait rendre possible le suivi de la progression de la tumeur et de la réponse aux traitements, ainsi que le dépistage des populations à risque.

Objectifs des travaux de Thèse

Objectif 1 : Le premier objectif de ce projet se concentre sur l'établissement d'une classification des glioblastomes basée l'imagerie par spectrométrie de masse et les données de protéomique. Ainsi, l'étude « Gliomic » (NCT02473484) intègre une cohorte prospective de 50 patients nouvellement diagnostiqués pour un glioblastome selon la classification de l'OMS 2016. L'ensemble des tumeurs a été analysé, d'une part par imagerie MALDI MSI afin d'obtenir la distribution des protéines au sein de chaque tissu et d'autre part via des analyses microprotéomiques localisées permettant d'accéder à l'identification à large échelle des protéines. L'addition des données génétiques et cliniques des patients a permis d'identifier des marqueurs moléculaires de pronostic des gliomes au sein des prélèvements tumoraux. Ainsi, cette étude a rendu possible l'établissement d'une classification des glioblastomes plus complémentaire à celle existante. D'autre part, de nouveaux marqueurs pronostiques pourront potentiellement être identifiés.

Objectif 2 : Le deuxième objectif de cette thèse est le développement d'une nouvelle technique de microprotéomique. Cette approche se base sur la microscopie par expansion (ExM) qui permet d'agrandir une coupe de tissu en 3 dimensions de façon isotropique. Le protocole d'expansion de base a été adapté afin de le rendre compatible avec une analyse protéomique. Notamment, l'étape du protocole dite « d'homogénéisation mécanique » permettant un agrandissement isotrope des coupes a été amélioré pour augmenter le rendement d'identification protéique.

Objectif 3 : La finalité de ce travail est d'associer le profil protéique des VEs du sang des patients de l'étude « Gliomic » à la classification des tumeurs réalisée par imagerie par spectrométrie de masse. Ainsi, la dernière partie du projet porte sur l'enrichissement et l'analyse protéomique de vésicules extracellulaires. Le développement du protocole d'enrichissement des vésicules constitue la base de cette partie. Les vésicules seront caractérisées par leurs contenus protéiques grâce à des analyses par spectrométrie de masse.

CHAPITRE 1

Utilisation de la protéomique localisée guidée par imagerie MALDI pour la stratification des patients atteints de glioblastome

I. Introduction

Les glioblastomes sont les tumeurs primitives malignes principales du système nerveux. Ces tumeurs ont un très mauvais pronostic : seulement 5 % des patients survivent plus de 5 ans (Ostrom et al., 2019). Dans la plupart des études cliniques, la médiane de survie est estimée à 16 mois (Chinot et al., 2014; Gilbert et al., 2014; Stupp et al., 2009, 2017; Weller et al., 2017b). Dans les études de population, la médiane de survie est estimée à 12 mois (Gramatzki et al., 2016). Les facteurs pronostiques favorables sont un âge inférieur à 40 ans et un indice de Karnofsky élevé au moment du diagnostic. De plus, on observe un meilleur pronostic quand le volume tumoral résiduel après chirurgie est faible. Le traitement standard du glioblastome intègre la résection maximale de la tumeur suivie d'une radiothérapie avec témozolomide concomitant et d'entretien (Stupp et al., 2005, 2009; Weller et al., 2017a).

Les critères morphologiques pour le diagnostic de glioblastome selon la classification actuelle des tumeurs du système nerveux central comprennent une activité mitotique intense, des atypies nucléaires, une prolifération microvasculaire importante et la présence de nécrose (Louis et al., 2016). Ainsi, trois groupes de glioblastomes sont définis au sein de la classification OMS 2016: les glioblastomes IDH non mutés qui représentent plus de 90 % des glioblastomes, les glioblastomes IDH mutés qui représentent moins de 10 % des glioblastomes et les glioblastomes NOS (*not otherwise specified*) lorsque le statut IDH est inconnu.

Cependant, ces tumeurs restent très hétérogènes et au sein d'un même groupe des évolutions cliniques différentes sont observées. Ainsi, il apparaît nécessaire de redéfinir des sous-groupes de glioblastomes.

Les différentes sous-classifications actuellement publiées se concentrent principalement sur le génome, le méthylome et le transcriptome. Ces classifications sont néanmoins peu utilisées en clinique et n'ont pas conduit à l'amélioration de la survie des

patients. Les approches protéomiques restent rares, pourtant de nombreuses études protéomiques ont déjà fourni des informations importantes sur la physiopathologie des glioblastomes (Deighton et al., 2010; Ghantasala et al., 2020; Jovčevska et al., 2017; Stetson et al., 2016). Les protéines sont en effet affectées par les dysfonctionnements cellulaires, dans leur régulation et leur structure. La protéomique permet l'analyse de ces changements par l'observation directe des protéines et de leurs modifications post-traductionnelles (phosphorylation, glycosylation, acétylation...). La détermination de signatures protéomiques spécifiques pourrait finalement aider à améliorer la distinction entre les différents sous-types de glioblastome et à guider la prise en charge thérapeutique.

La plupart des études menées actuellement analysent de façon globale la totalité des protéines du tissu. Compte tenu de l'hétérogénéité de ces tumeurs, ces études ne permettent pas de distinguer chaque sous-région tumorale. Les glioblastomes étant très hétérogènes, plusieurs échantillons provenant d'une même tumeur peuvent comporter différents des sous-régions avec des proportions variables de zones de prolifération tumorale, nécrose, prolifération capillaire. Aussi, peu d'études prennent en compte les données cliniques des patients, comme la survie après chirurgie, et il est nécessaire d'avoir une cohorte de patients suffisante pour établir des données statistiques fiables. De nouvelles approches sont donc nécessaires afin d'une part de mieux appréhender les différents sous-groupes de gliomes pour la recherche de traitement et d'autre part pour améliorer la qualité de vie des patients.

En ce sens, une première étude réalisée au laboratoire sur 10 gliomes de grade III a démontré l'intérêt de l'imagerie par spectrométrie de masse pour établir une classification de ces tumeurs (Le Rhun et al., 2016). Cette étude a fourni de nouvelles informations sur l'organisation des gliomes de grade III et pourrait permettre l'élaboration d'une classification plus précise de ces tumeurs. Pour chaque tumeur, les images ont été générées après digestion à la trypsine. Par la suite, la segmentation spatiale individuelle a fourni des informations concernant l'hétérogénéité moléculaire de chaque échantillon et a permis de déterminer des régions d'intérêt spécifiques intra-tumoral. Pour certains tissus on observe des différences entre la segmentation et les annotations histologiques. Notamment, l'imagerie révèle que les tumeurs annotées comme homogènes par le pathologiste sont en réalité complexes avec plusieurs zones moléculaires distinctes.

Ensuite, la segmentation globale de cinq tissus a permis de regrouper les patients ayant une signature moléculaire similaire. Ainsi, trois groupes de tumeurs ont été identifiés alors même que ces tumeurs sont toutes classées en grade III par la classification de l’OMS de 2007. Par la suite, les protéines de ces régions ont été identifiées par microdigestion et micro extraction in situ suivi d’une analyse LC-MS/MS. Finalement, les analyses statistiques des données de microprotéomique ont montré le même regroupement que celui déterminé par imagerie MALDI (Figure 1-1).

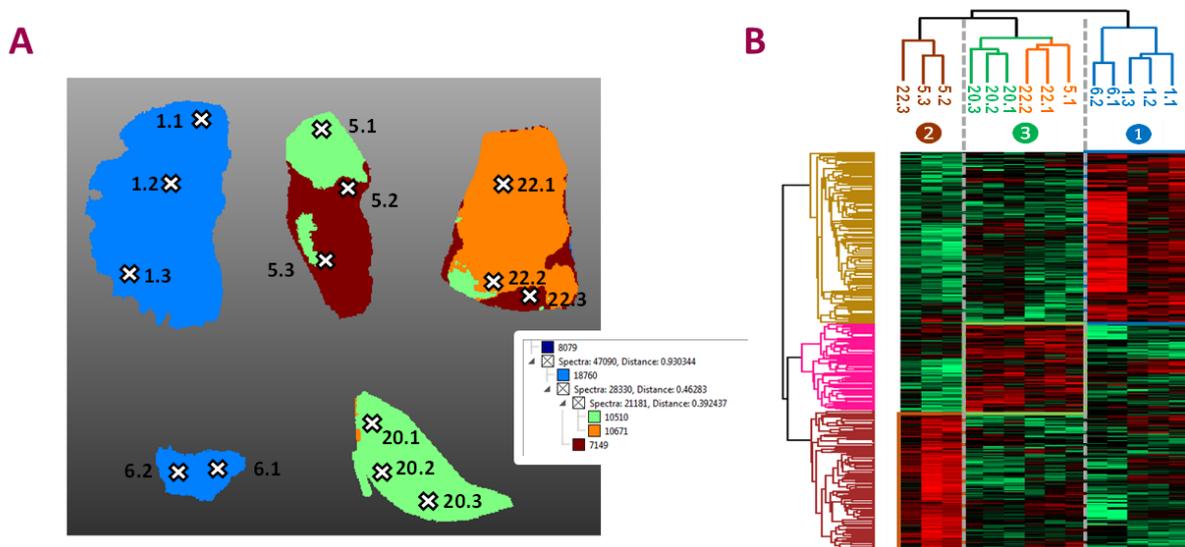


Figure 1-1 Classification des gliomes de grade III basée sur l’imagerie MALDI et la microprotéomique

A. Segmentation globale de l’ensemble des tissus. Chaque couleur correspond à une région moléculaire conformément au dendrogramme. Les échantillons ont été regroupés selon leur similarité de profil moléculaire. Trois groupes ont été déterminés : bleu, marron et orange/vert. Les croix blanches représentent la localisation des points de microextractions sur chacun des cinq tissus. C. Heatmap des protéines significatives entre les trois groupes, mettant en évidence la présence de 3 branches de patients en accord avec la segmentation globale des tissus réalisée par imagerie. Adapté de (Le Rhun et al., 2016)

Pour chaque groupe un set de protéines surexprimées a été identifié. L’analyse des fonctions biologiques associées à ces protéines a permis de révéler des réseaux protéiques différents associés à chaque groupe : la néoplasie pour le groupe 1, l’inflammation pour le deuxième groupe et à la neurogénèse pour le troisième groupe (Figure 1 2). Cette étude ayant été réalisée sur un nombre limité de patient, il est nécessaire d’augmenter la cohorte pour avoir une analyse statistique plus fiable.

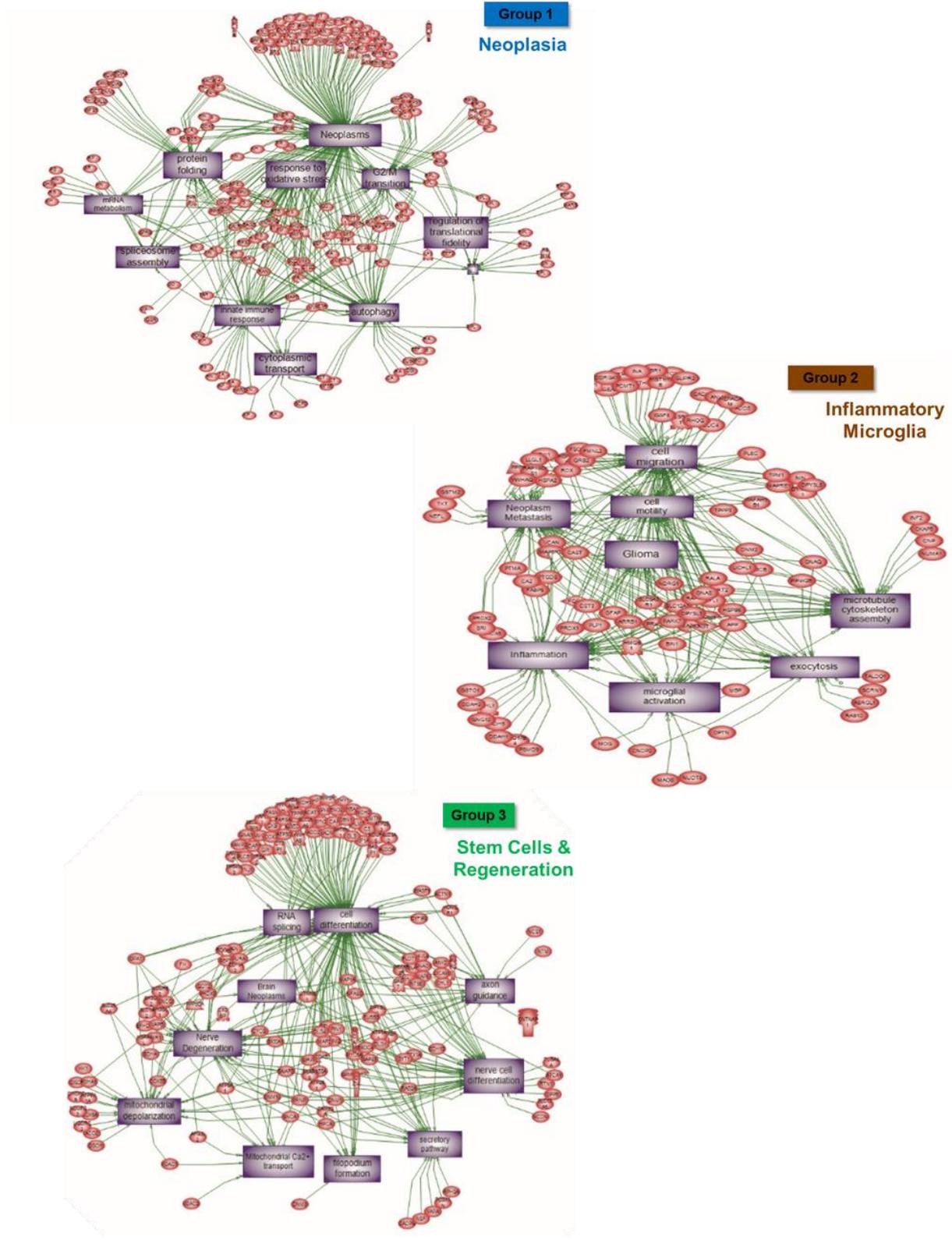


Figure 1-2 Voies de signalisations associées aux protéines surexprimées dans les 3 groupes identifiées par imagerie MALDI

Réseaux protéiques différents associés à chaque groupe : la néoplasie pour le premier groupe, l'inflammation pour le deuxième groupe et à la neurogénèse pour le troisième groupe. Adapté de (Le Rhun et al., 2016).

Le projet de thèse présenté dans ce manuscrit se place dans la continuité de cette étude pilote. La méthodologie d'analyse sera utilisée sur une cohorte de 50 patients atteint de glioblastome pour améliorer la classification actuelle de ces tumeurs.

Étude GLIOMIC

L'étude GLIOMIC (NCT02473484) est le sujet principal de ce manuscrit. Cette étude est une cohorte prospective de 50 patients nouvellement diagnostiqués pour un glioblastome. Les données cliniques, les échantillons tumoraux et sanguins des patients pour lesquels un diagnostic de glioblastome selon la classification OMS 2016 était confirmé par l'anatomopathologiste ont été collectés à partir de patients recrutés au Centre Hospitalier Universitaire de Lille (CHU) de Lille selon le protocole mis en place pour l'étude.

En plus d'améliorer la classification de ces tumeurs, cette stratégie expérimentale pourrait permettre d'améliorer la compréhension des processus physiopathologiques et ainsi de découvrir nouveaux marqueurs et potentielles cibles thérapeutiques. En effet, à l'heure actuelle, les réponses aux différents traitements sont peu efficaces.

II. Matériel et Méthodes

A. Recrutement des patients

Les patients nouvellement diagnostiqués pour un glioblastome ont été recrutés prospectivement entre septembre 2014 et novembre 2018 au Centre Hospitalier Universitaire (CHU) de Lille, France. La participation à l'étude GLIOMIC a été proposée aux patients avec une suspicion de gliome de haut grade lors des consultations réalisées dans le cadre du soin avant la chirurgie. Les principaux critères d'inclusion sont :

- Âge supérieur à 18 ans
- Absence d'antécédents d'autre cancer, sauf cancer basocellulaire cutané ou cancer in situ du col de l'utérus
- Absence d'antécédent de radiothérapie encéphalique
- Absence de traitement systémique antitumoral (chimiothérapie ou thérapie ciblée) antérieur à la chirurgie du gliome
- Absence de maladie génétique pouvant être à l'origine de tumeur cérébrale, tumeur potentiellement radio-induite
- Absence d'antécédents de maladie neurologique dégénérative antérieure au diagnostic

Enfin, les échantillons devaient être congelés à la tumorothèque de Lille dans les 2 heures après la chirurgie.

L'accord du Comité de Protection des Personnes a été obtenu le 08/04/2014 (ID-RCB 2014-A00185-42) avant l'initiation de l'étude. Cette étude adhère aux principes de la déclaration d'Helsinki et des règles de bonne pratique clinique, tous les patients ont signé un accord de consentement. L'étude est enregistrée au NCT 02473484.

B. Immunohistochimie (IHC)

Les mutations IDH1-R132H et l'expression nucléaire ATRX ont été déterminées par immunohistochimie sur des échantillons de tissus tumoraux fixés et inclus en paraffine (FFPE). Pour cela, des coupes de 4µm ont été marquées dans un automate Ultra (Ventana-Roche

Tissue Diagnostics, Tucson AZ) après démasquage (IDH1 et ATRX : citrate pH 6,0) selon les protocoles des fournisseurs. Un contrôle positif (tissu tumoral d'un astrocytome muté IDH1 R132H, non codé-lété) a été coloré sur la même lame pour le marquage IDH1. La mutation ATRX a été déterminée comme positive quand plus de 10 % de cellules tumorales positives ont été notés (expression ATRX). Pour la mutation IDH1 R132H le clone Diavona H09, monoclonal de souris a été utilisé à 1/40 de dilution, pour ATRX l'anticorps HPA001906 Lapin polyclonal de Sigma a été utilisé à dilution 1/200. Ces analyses ont été effectuées au laboratoire U1172 - LiNCog - Lille Neuroscience & Cognition.

C. Extraction d'ADN et quantification

Des analyses moléculaires ont été réalisées sur des tissus FFPE. Les tests suivants ont été effectués : tableau d'hybridation génomique comparative (CGH), méthylation du promoteur O6-méthylguanine-ADN méthyltransférase (MGMT) et séquençage IDH1/2. Tous les tissus utilisés pour l'extraction d'ADN ont été évalués histologiquement pour déterminer le contenu des cellules tumorales. L'analyse a été effectuée sur tous les échantillons de tissus. Les échantillons contenant moins de 40 % de cellules tumorales ont été considérés comme non interprétables lorsqu'aucune anomalie moléculaire n'a été trouvée. L'extraction d'ADN à partir des tissus FFPE a été réalisée à l'aide du kit QIAamp DNA FFPE Tissue (Qiagen). Les profils CGH ont été déterminés en utilisant le kit de microarray CGH humain SurePrint G3, 8x60K (Agilent) et le logiciel CytoGenomics v2.7. La limite de résolution était de 1 Mb. La présence d'une codé-létion 1p / 19q, le gain du chromosome 7, la perte du chromosome 10, l'amplification du gène EGFR et la suppression homozygote du gène Cyclin-Dependent Kinase Inhibitor 2A (CDKN2A) ont été systématiquement évalués. Les mutations affectant le codon 132 de IDH1 (ref seq NM_005896.2), le codon 172 d'IDH2 (ref seq NM_002168.2) ont été validées par la séquence PCR Sanger lorsque l'immunohistochimie était négative. Le statut de méthylation du promoteur MGMT (CpGs 74-78) a été déterminé après traitement au bisulfite par pyroséquençage sur un PyroMark Q96 avec le kit MGMT PyroMark (Qiagen). La présence d'une méthylation était positive lorsqu'un minimum de 8 % de méthylation a été observé. Ces analyses ont été effectuées au CHU de Lille, Service de biochimie et biologie moléculaire.

D. Imagerie par spectrométrie de masse MALDI

1. Préparation du tissu et acquisition de l'image

Des coupes consécutives (12µm d'épaisseur) de chaque échantillon ont été réalisées à l'aide d'un cryostat (Leica Microsystems, Nanterre, France) pour réaliser les images MALDI. Les coupes ont été déposées sur lame ITO (LaserBio Labs, Valbonne, France) et séchées 15 minutes au dessiccateur dès leur sortie du congélateur. Plusieurs étapes de lavages ont été réalisées afin de retirer les lipides qui pourraient masquer le signal protéique : (1) 1 minute dans l'éthanol 70 % (v/v), (2) 1 minute dans l'éthanol 100 % (3) 30 secondes dans l'acétone et (4) 30 secondes dans le chloroforme (Lemaire et al., 2006b). Ces deux dernières étapes ont été répétées deux fois. De la trypsine (60 µg/mL, diluée dans du tampon NH_4HCO_3 50 mM) a ensuite été déposée uniformément sur toute la surface du tissu pendant quinze minutes à l'aide d'un nébuliseur relié à une seringue à un débit de 180 nl/min. L'incubation a ensuite été effectuée dans l'ImagePrep (Bruker Daltonics, Breme, Allemagne) permettant la digestion des protéines. L'instrument nébulise de l'eau chauffée à 37°C sur le tissu pendant 2 heures. La procédure de digestion comprend 60 cycles correspondant chacun à 2 secondes de nébulisation, 180 secondes d'incubation et 60 secondes de séchage. Durant toute la durée de la digestion, un réservoir contenant de l'eau chauffée, cette fois-ci, à 95°C a été placé dans l'ImagePrep pour maintenir une atmosphère humide. Cette eau ainsi que l'eau à 37°C ont été remplacées toutes les 30 minutes. La matrice HCCA/ANI (Lemaire et al., 2006a) a été préparée juste avant utilisation en ajoutant 36 µL d'aniline à 5 mL d'une solution d'HCCA (10 mg/mL dilués dans ACN / TFA 0.1 % aq. (7:3, v/v)). Après 5 minutes de sonication, la matrice a été déposée sur le tissu en utilisant l'ImagePrep. Une méthode spécialement développée pour la matrice HCCA/ANI a été utilisée. Cette méthode comprend 4 étapes de cycles de pulvérisation. En bref, l'étape 1 comprend 5 cycles à 0,1 V avec 1 seconde de pulvérisation, 10 secondes d'incubation et 180 secondes de séchage. L'étape 2 comprend 8 cycles à 0,2 V avec 1,25 secondes de pulvérisation, 30 secondes d'incubation et 120 secondes de séchage. À l'étape 4, 37 cycles à 0,3 V ont été effectués avec une pulvérisation de 1,55 seconde, une incubation de 20 seconde et un temps de séchage adapté selon le capteur optique (25%). Enfin, 3 cycles ont été réalisés à l'étape 4 avec 2 secondes de pulvérisation, 20 secondes d'incubation et un temps

de séchage adapté selon le capteur optique (25%) et pour le dernier cycle jusqu'à séchage complet de la lame. Ces paramètres ont été optimisés afin d'éviter la délocalisation des peptides et permettre la cristallisation des peptides avec la matrice. Après le dépôt de la matrice, la lame a été directement transférée dans le spectromètre de masse. Les images ont été réalisées avec un spectromètre de masse Ultraflex II MALDI-TOF/TOF (Bruker Daltonics, Breme, Allemagne) équipé d'un laser SmartBeam II. Les spectres de masse des peptides digérés ont été obtenus en mode réflecteur positif sur une gamme de masse entre m/z 800 et 4000. Les spectres ont été générés à 400 tirs lasers par point. Les images ont été réalisées avec une résolution spatiale de 70 μm .

2. Traitement et analyses des images

Le logiciel SCiLS Lab (SCiLS Lab 2015b, SCiLS GmbH) a été utilisé pour la segmentation spatiale non supervisée des gliomes. Les données brutes ont d'abord été importées sur le logiciel. Un prétraitement des données a ensuite été réalisé afin de ramener tous les spectres au niveau de la ligne de base et ainsi permettre de les comparer les uns aux autres. Les spectres ont été normalisés pour corriger les variations d'intensité par rapport au courant ionique total (TIC) afin de diminuer les variations interspectres au cours d'une même acquisition (D et al., 2012; Klein et al., 2014). Un filtre a ensuite été appliqué afin de réduire le bruit de fond. Les données générées ont alors été spatialement segmentées grâce à un algorithme « bisecting k means » (cf. État de l'art) (Alexandrov et al., 2010).

En premier lieu, chaque tissu a été segmenté individuellement ; la segmentation individuelle fournit des informations concernant l'hétérogénéité du tissu et permet de déterminer des régions d'intérêt pour chaque échantillon. Par la suite, tous les tissus ont été segmentés ensemble (segmentation globale) afin de regrouper les patients ayant une signature moléculaire similaire. En bref, l'algorithme de segmentation a regroupé les spectres selon leurs similarités. Les résultats sont exprimés sous la forme d'un dendrogramme, où chaque branche correspond à un groupe de spectres pour lequel une couleur est attribuée. Les branches du dendrogramme ont été définies sur la base d'un calcul de distance Euclidien (simple distance entre deux points) entre chaque cluster. Une carte de segmentation a alors été générée où les régions de composition moléculaire différente sont représentées par une

couleur suivant le dendrogramme. À titre de comparaison, la segmentation globale a également été effectuée en utilisant la méthode de clustering de Ward avec IMAGEREVEAL MS Ver.1.1 (Shimadzu). Les segmentations spatiales ont permis de déterminer des régions d'intérêt. Ces régions seront ensuite soumises à une microdigestion suivie d'une micro-extraction afin d'avoir accès à l'identification des protéines qu'elles contiennent.

E. Analyses SpiderMass

1. Acquisition des spectres

La configuration globale de l'instrument a été décrite par (Ogrinc et al., 2019; Saudemont et al., 2018) (Figure 1-3). Brièvement, le système est composé de trois parties : une sonde de microéchantillonnage constituée d'un laser fibré et accordable en longueur d'onde, une ligne de transfert et un spectromètre de masse (Fatou et al., 2016). La microsonde laser est réglée sur la longueur d'onde de 2,94 μ m (Radiant version 1.0.1, OPOTEK Inc., Carlsbad, CA, USA) pompé par un laser Nd: YAG pulsé (durée d'impulsion: 5 ns , λ = 1064 nm, Quantel, Les Ulis, France). Une fibre laser biocompatible (450 μ m de diamètre intérieur ; longueur de 1 m ; Infrared Fiber Systems, Silver Spring, CO, USA) est connectée à la sortie du système laser et à la sonde de microéchantillonnage comprenant une lentille de focalisation de 4 cm est fixée à l'extrémité de la fibre laser. La sonde de microéchantillonnage permet à l'utilisateur de tenir le système et de balayer la surface des tissus à une résolution de 400 μ m. Dans ces expériences, le temps d'irradiation a été fixé à 1 seconde à une énergie laser de 5 mJ / impulsion correspondant à une fluence laser d'environ 3 J / cm². L'énergie laser a été mesurée au point focal de la lentille de focalisation à l'aide d'un puissance-mètre (ThorLabs, Maisons-Laffitte, France). La deuxième partie du système est une ligne de transfert de trois mètres de long réalisée avec un tube Tygon ND 100-65 (2,4 mm de diamètre intérieur, 4 mm de diamètre extérieur, Akron, USA). La ligne de transfert est fixée d'une part sur la sonde de microéchantillonnage à l'extrémité de la fibre laser et d'autre part directement au spectromètre de masse (Xevo, Waters, Manchester, Royaume-Uni) où la source électrospray conventionnelle a été retirée et remplacée par une interface de pression atmosphérique (Fatou et al., 2016). Chaque acquisition a été réalisée avec une infusion d'isopropanol à 150

$\mu\text{L}/\text{min}$ grâce à une jonction en T à l'entrée du spectromètre. Le matériel ablaté et l'isopropanol sont ainsi mélangés dans la jonction avant d'être analysés. L'acquisition des spectres a été effectuée en mode positif avec un temps de balayage de 1 seconde.

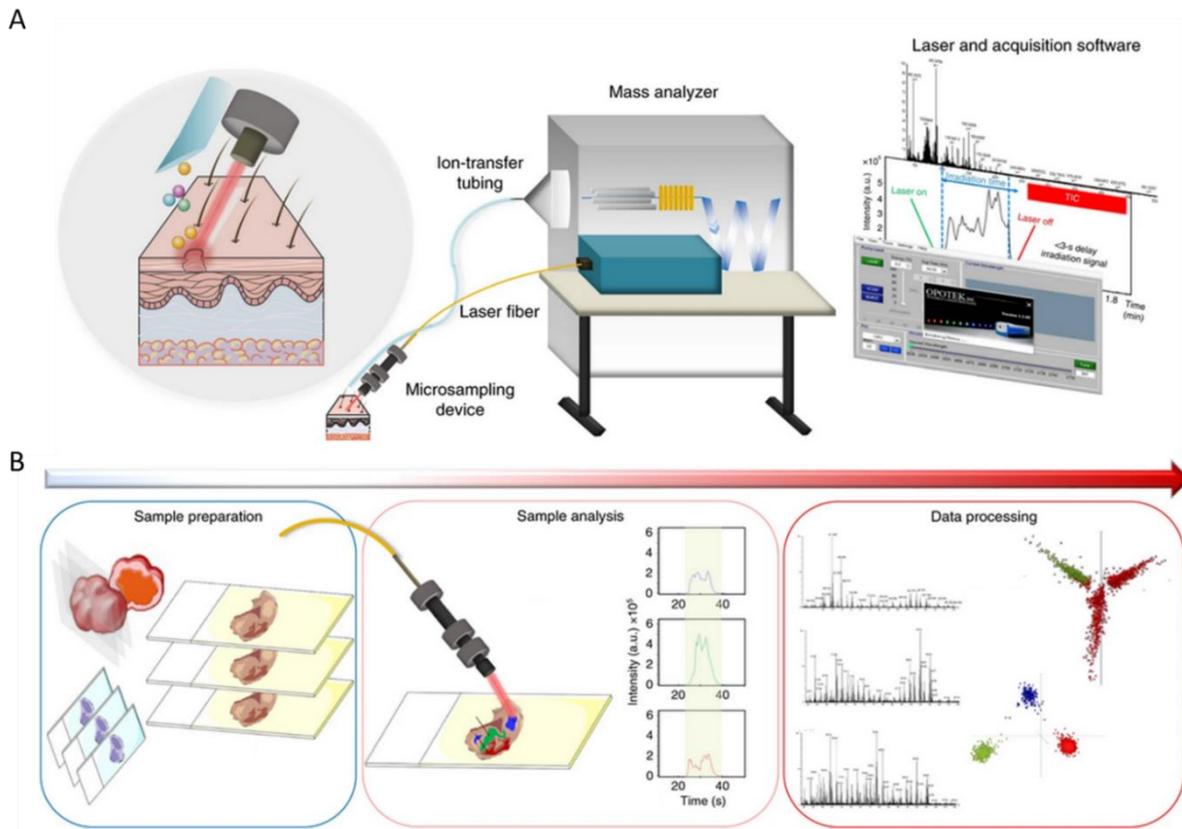


Figure 1-3 Schéma de la configuration du SpiderMass et déroulement d'une expérience standard.

A. Le SpiderMass est équipée d'un laser fibré, d'un tube de transfert d'ions et d'un spectromètre de masse. B. L'analyse SpiderMass est généralement effectuée sur une période d'irradiation de 10 secondes à 3 reprises pour chacune des zones d'intérêt. Pour reconstruire des modèles de classification, les spectres enregistrés sont affectés à une classe et sont ensuite soumis à une analyse multivariée non supervisée (par exemple, PCA) ou supervisée (par exemple, LDA) pour étudier la séparation des données et construire des modèles de classification. Adapté de Ogrinc et al., 2019.

Avant l'analyse SpiderMass, les échantillons ont été décongelés à température ambiante pendant 30 s. La séquence d'acquisition spectrale était composée de 2 ou 3 acquisitions avec des périodes d'irradiation d'une seconde. Les ROIs ont été sélectionnées à partir des données d'imagerie MALDI avant chaque acquisition de SpiderMass afin de s'assurer que chaque acquisition a été réalisée sur la même zone du tissu (Ogrinc et al., 2019).

2. Construction du modèle de classification

Pour l'analyse des données, tous les fichiers bruts produits avec le SpiderMass ont été importés dans le logiciel Abstract Model Builder (AMX v.0.9.2092.0). Après l'importation, les spectres ont d'abord été prétraités. Les étapes de prétraitement comprennent la soustraction de la ligne de base, la normalisation du nombre total d'ions (TIC), la calibration de la correction de masse et le binning de 0,1 ou 0,2 Da. L'analyse a été réalisée sur 30 échantillons, la totalité des échantillons (50) n'a pas pu être analysé, faute de quantité de coupes suffisante ou d'état général des coupes ne permettant pas d'obtenir de mesure satisfaisante. L'ensemble des spectres MS obtenus à partir des 30 échantillons tumoraux a ensuite été utilisé pour construire un modèle de classification par analyse en composantes principales et analyse discriminante linéaire (PCA-LDA) (Ogrinc et al., 2019). La première étape consiste à réaliser une PCA pour réduire la multidimensionnalité des données en générant des caractéristiques expliquant une grande partie de la variance observée (cf. ÉTAT DE L'ART). Ces caractéristiques ont ensuite été soumises à une analyse supervisée par LDA en définissant les classes sur lesquelles le modèle sera basé. La LDA permet de classer les spectres de chaque échantillon et d'évaluer le modèle par validation croisée. La validation croisée a été effectuée soit en utilisant les méthodes «20 % out» ou «leave one patient out». Pour la première méthode, 20 % des spectres MS sont retirés au hasard du spectre total et le modèle est reconstruit à partir des 80 % restants. Les 20 % de spectres restants sont utilisés pour interroger le modèle reconstruit. La permutation est automatiquement réitérée pendant 5 cycles avant de rendre compte des résultats de la validation croisée. Pour la seconde méthode, les spectres issus d'un même patient sont retirés de la banque de données. La nouvelle banque de données ainsi créée sera interrogée pour classer les spectres du patient. Cette opération a été répétée pour l'ensemble des 30 patients afin de valider le modèle.

F. Protéomique localisée sur tissus

1. Digestion *in situ*

Au total, 147 ROI ont été sélectionnés à partir des données d'imagerie MALDI. Les étapes de protéomique localisées ont été réalisées au niveau de ces régions prédéfinies selon

le protocole précédemment publié (Quanico et al., 2013). Les coupes de tissu (20 µm) ont été séchées 15 minutes au dessiccateur dès leur sortie du congélateur et plusieurs étapes de lavages ont été réalisées afin de retirer les lipides. Une solution de LysC-Trypsine (40 µg/mL, dans Tris-HCl 50 mM, pH 8.0) a ensuite été déposée grâce à une imprimante chimique (CHIP-1000, Shimadzu, CO, Kyoto, Japon). L'aire de digestion a été définie à 1 mm² (4x4 spots de 200µm). L'enzyme a été déposée pendant 2 heures de manière automatique puis une solution de TFA 0.1 % aq a été déposée selon les mêmes paramètres durant 25 cycles pour stopper la digestion.

2. Extraction par microjonction liquide (LESA)

Après la microdigestion, les peptides ont été extraits à l'aide de l'instrument TriVersa Nanomate (Advion Biosciences Inc., Ithaca, NY, États-Unis) avec l'option « Liquid Extraction Surface Analysis » (LESA). En résumé, 0,6 µL de solvant ont été aspirés et déposés sur la région digérée afin de former une microjonction liquide entre la surface du tissu et la pointe du cône. Le liquide a ensuite été aspiré à nouveau pour être distribué dans un tube collecteur. Trois étapes ont été effectuées pour chaque extraction : (1) TFA 0.1 % aq, (2) ACN / TFA 0.1 % aq. (8:2, v/v), (3) MeOH: TFA 0,1 % aq. (7:3, v/v). Chaque étape a été répétée deux fois pour augmenter la quantité de peptides recueillie.

3. Analyses NanoLC-MS/MS

Après la micro-extraction, les échantillons ont été séchés, reconstitués puis dessalés sur une colonne Ziptip C-18 (Millipore, Saint-Quentin-en-Yvelines, France) avant d'être séchés à nouveau. Les échantillons secs ont ensuite été repris avec une solution ACN / AF 0.1 % aq (2/98, v/v). L'analyse a été effectuée via une chromatographie NanoAcquity UPLC (Waters) couplée à un spectromètre de masse Q-exactive (Thermo Scientific) via une source électrospray. Une colonne de préconcentration (nanoAcquity Symmetry C18, 5 µm, 180 µm x 20 mm) et une colonne analytique (nanoAcquity BEH C18, 1,7 µm, 75 µm x 250 mm) ont été utilisées. Les peptides ont été séparés en appliquant un gradient linéaire d'acétonitrile (5 %-35 %) pendant 2 heures à un débit de 300 nl/min. La fragmentation MS/MS a été réalisée en mode dépendante des données (DDA) des 10 ions les plus intenses de l'analyse MS (Top 10).

L'analyse MS a été réalisée avec une gamme de masse m/z comprise entre 300 et 1600, une résolution de 70 000 FWHM, un AGC de 3×10^6 ions et un temps d'injection maximum de 120 ms. L'analyse MS/MS a été effectuée avec une gamme de masse située entre 200 et 2000, un AGC de 5×10^4 ions, un temps d'injection maximum de 60 ms et la résolution a été fixée à 17 500 FWHM. Pour éviter tout effet de lot lors de l'analyse, les extractions ont été choisies au hasard pour créer des séquences d'analyse.

4. Analyse des données

Les données de spectrométrie de masse ont été traitées avec le logiciel MaxQuant (Cox and Mann, 2008; Tyanova et al., 2015) (Version 1.5.3.30) utilisant le moteur de recherche Andromeda (Cox et al., 2011). L'identification des protéines a été réalisée par interrogation de la base de données UniProt Homo Sapiens (Juillet 2018, 20 412 entrées) avec comme critère d'identification 2 peptides minimum par protéine dont au moins 1 unique. La trypsine a été choisie pour le mode de digestion. L'acétylation en position N-terminale et l'oxydation des méthionines ont été choisies comme modifications variables. Les analyses ont été réalisées avec une précision de mesure de 6 ppm en MS et 20 ppm en MS/MS. Le taux de faux positifs (FDR) a été défini à 1 % en peptide et protéine. La quantification sans marquage au préalable des protéines (*label-free quantification*) a été réalisée grâce à l'algorithme MaxLFQ intégré dans le logiciel MaxQuant avec les paramètres par défaut (Cox et al., 2014).

L'analyse statistique des protéines identifiées a été réalisée grâce au logiciel Perseus (version 1.6.0.7). Tout d'abord, les résultats issus de la base de données inversée, les protéines contenant uniquement des peptides modifiés et les contaminants potentiels ont été supprimés. Des analyses statistiques ont été effectuées en utilisant un test ANOVA avec une valeur seuil "Benjamini Hocheberg FDR" de 5 %. Trois regroupements différents ont été utilisés pour l'ANOVA : (1) le groupe de couleurs basé sur les trois couleurs de la segmentation globale Scils des 50 échantillons (rouge; jaune et bleu), (2) le groupe de patients qui est déterminé par la couleur majoritaire présente dans chaque échantillon de tumeur (groupes A, B, C) et (3) le temps de survie des patients (patients avec une OS > au troisième quartile ; patients avec une OS entre le premier et le troisième quartile et patients avec une OS < au premier quartile). Des protéines significativement différentes ont été sélectionnées et

normalisées par un Z-score. Pour la représentation, un regroupement hiérarchique a été réalisé.

G. Analyses biologiques

1. Analyses des voies de signalisation

L'annotation des protéines identifiées a été réalisée à l'aide du système de classification PANTHER (version 14.1, <http://www.pantherdb.org>), de FunRich (version 3.1.3) (Pathan et al., 2017) et de la base de données STRING (version 11.0, www.string-db.org) (Szklarczyk et al., 2019). Le réseau d'interaction potentielle a ensuite été chargé dans Cytoscape 3.7.2. L'extension Reactome FI a également été utilisé. Les protéines différentiellement exprimées dans chaque condition ont également été décrites grâce à la base de données Ariadne ResNet (Yuryev et al., 2009) à l'aide de Pathway Studio (version 11.0, Elsevier). L'algorithme d'analyse d'enrichissement de sous-réseau (SNEA *subnetwork Enrichment Analysis*) a été utilisé pour détecter les voies de signalisations significatives dans lesquelles les protéines identifiées sont impliquées.

2. Comparaison avec *Human Pathology Atlas*

Les données sur les gliomes contenues dans l'atlas de pathologie humaine (Human pathology atlas) (Uhlen et al., 2017) ont été utilisées. La base de données transcriptomique TCGA provenant de 153 patients et qui a permis d'identifier 268 gènes potentiellement pronostiques (201 pronostics défavorables et 67 pronostics favorables) a été comparé aux protéines identifiées dans notre étude.

3. Identification de protéines issues de cadres de lectures alternatifs

Les données brutes obtenues par analyse nanoLC-MS / MS ont été analysées en utilisant la quantification LFQ (nœud de calcul Minora) de Proteome Discoverer V2.3 (Thermo Scientific) avec les paramètres suivants : trypsine comme enzyme de digestion, 2 clivages manqués, oxydation à la méthionine comme modification variable et carbamidométhylation

des cystéines comme modification statique, tolérance de masse du précurseur : 10 ppm et tolérance de masse du fragment : 0,6 Da. La validation a été effectuée à l'aide du Percolator avec un FDR réglé à 0,001 %. Un consensus a ensuite été appliqué pour l'arrangement statistique en utilisant un indice de confiance d'identification de protéines élevé. La base de données de protéines a été téléchargée depuis Openprot (<https://openprot.org/>) et comprend des protéines de références (RefProt), de nouvelles isoformes et des protéines alternatives (AltProts) prédites à partir des annotations Ensembl et RefSeq (GRCh38.83, GRCh38.p7) (Brunet et al., 2019; Delcourt et al., 2018b; Vanderperre et al., 2012, 2013) pour un total de 658 263 entrées. Les données de quantification des protéines ont été extraites et utilisées dans Perseus pour réaliser une analyse statistique et une représentation graphique.

H. Analyses statistiques

Tout d'abord, des analyses descriptives ont été effectuées sur les données cliniques. Les patients ont ensuite été divisés en 3 groupes selon les quartiles de la survie globale (<Q1, Q1-Q3,> Q3). Le modèle de Cox a été utilisé pour déterminer quelles protéines étaient les plus associées à la survie globale. Une analyse par étapes et des méthodes de bootstrap (500 échantillons) ont été utilisées pour garantir la robustesse des résultats. Les protéines sélectionnées après cette étape ont été utilisées pour effectuer une classification hiérarchique (distance euclidienne et méthode de Ward) sur les 50 patients afin de déterminer s'il y avait des sous-groupes (*clusters*). Enfin, les variables cliniques ont été analysées selon les différents clusters afin de fournir une description clinique des clusters obtenus. Ces analyses statistiques ont été effectuées à l'aide du logiciel SAS, V9.4.

I. Confirmation par immunofluorescence (IF)

Les groupes de survie ont été validés à l'aide d'anticorps dirigés contre MAOB, IGHM, HSPD1, LASP1, CFH et ANXA6. Les tissus ont été incubés avec l'anticorps primaire à 4 °C pendant une nuit, puis avec un anticorps secondaire (Alexa fluor conjugué anticorps, dilution 1/1 000) pendant 1 heure à température ambiante. Les anticorps primaires suivants ont été utilisés : MAOB (Abxexa; dilution 1/100), IGHM (Abcam, dilution 1/50), HSPD1 (Abcam, dilution 1/200), LASP1 (Santa Cruz, dilution 1/50), CFH (Abcam, dilution 1/200) et ANXA6

(Abcam, dilution 1/50). Toutes les lames ont été observées sur un microscope confocal Zeiss LSM700. Trois à quatre photos ont été prises pour chaque section tumorale. Le traitement des images et la quantification de l'intensité de fluorescence ont été effectués à l'aide du logiciel ImageJ.

L'ensemble des caractéristiques cliniques explorées ont été reportées dans le Tableau 1-1. Parmi les 50 patients de l'étude, 35 patients (70 %) étaient des hommes. Au moment du diagnostic, l'âge médian était de 59 ans. L'indice de Karnofsky médian, reflétant l'état général, des patients était de 90 (100 représente un état de santé sans pathologie ; 0 correspond au décès).

La totalité des patients a bénéficié d'une chirurgie, 29 (58 %) ont eu une résection complète, 20 (40 %) une résection partielle et un patient a été biopsié. Le traitement standard, radiothérapie et chimiothérapie par témozolomide, a été initié pour 45 patients (90 %). Des études cliniques ont été proposées pour 2 patients. Deux patients ont reçu un autre traitement : l'un n'a reçu qu'une radiothérapie accélérée en raison de son âge et de son état général et le second a reçu 6 cycles de témozolomide puis une radiothérapie stéréotaxique. Un patient n'a pas reçu de traitement.

Une progression tumorale a été observée pour 40 patients (80 %), la médiane de survie sans progression est de 10,6 mois. Une nouvelle ligne de traitement a été initiée pour 35 patients incluant une chirurgie dans 5 cas. Après un suivi médian de 21 mois, la majorité des patients (82 %) étaient décédés. La survie globale médiane est de 20,5 mois.

Tableau 1-1 Caractéristiques cliniques

	Total population (n=50)
<i>Sexe</i>	
<i>Femme, n (%)</i>	15 (30)
<i>Homme, n (%)</i>	35 (70)
<i>Âge au diagnostic (année)</i>	59 (50-65)
<i>Indice de Karnofsky au diagnostic médiane (IQ)</i>	90 (80-100)
<i>0-80, n (%)</i>	15 (30)
<i>90-100, n (%)</i>	35 (70)
<i>Localisation</i>	
<i>frontal, n (%)</i>	15 (30)
<i>occipital, n (%)</i>	3 (6)
<i>pariétal, n (%)</i>	12 (24)
<i>temporal, n (%)</i>	20 (40)
<i>Qualité de la chirurgie</i>	
<i>complète, n (%)</i>	29 (58)
<i>partielle, n (%)</i>	20 (40)
<i>biopsies, n (%)</i>	1 (2)
<i>Statut IDH1/2*</i>	

Non muté, n (%)	46 (92)
mutant, n (%)	4 (8)
ATRX	
marquage non conservé, n (%)	7 (14)
marquage conservé, n (%)	43 (86)
Statut de méthylation du promoteur MGMT	
non méthylé, n (%)	33 (66)
méthylé, n (%)	17 (34)
Amplification EGFR	
non, n (%)	26 (52)
oui, n (%)	24 (48)
Gain chromosome 7 et perte chromosome 10	
non, n (%)	15 (30)
oui, n (%)	35 (70)
Amplification EGFR ou gain 7 / perte 10	42 (84)
Amplification EGFR sans gain 7 / perte 10	7 (14)
Amplification EGFR et gain 7 / perte 10	17 (34)
Gain 7 / perte 10 sans amplification EGFR	18 (36)
Délétion homozygote CDKN2A	
non, n (%)	19 (39)
oui, n (%)	29 (59)
autre: amplification de CDKN2A, n (%)	1 (2)
Traitement initial	
RT/TMZ puis 6 cycles de TMZ, n (%)	20 (40)
RT/TMZ puis au moins 6 mois de TMZ, n (%)	4 (8)
RT/TMZ puis moins de 6 mois de TMZ, n (%)	21 (42)
Autre traitement**, n (%)	2 (4)
Etude clinique, n (%)	2 (4)
Pas de traitement, n (%)	1 (2)
Progression	
oui, n (%)	40 (80)
non, n (%)	5 (10)
non connue, n (%)	5 (10)
Médiane de survie sans progression (mois)	10.6 (6.9-16.8)
Traitement à la première progression	
oui, n (%)	35*** (87.5)
non, n (%)	5 (12.5)
Décès	
oui, n (%)	41 (82)
non, n (%)	9 (18)
Médiane de survie après chirurgie (mois)	20.5 (13.5-34.5)
Survie	
Quartile supérieur, n (%)	12 (24)
Interquartile, n (%)	26 (52)
Quartile inférieur, n (%)	12 (24)

* selon la classification OMS 2016; ** un patient: RT seulement, un patient: 6 cycles TMZ puis SRT ; *** incluant la chirurgie dans 5 cas. RT: radiothérapie, TMZ: temozolomide

Pour chaque patient, un échantillon de tissu tumoral a été récupéré et soumis à des analyses immunohistochimiques.

L'expression de IDH1 (R132H) a été mise en évidence dans 4 tumeurs (8 %), l'ensemble des autres tumeurs étant IDH1 (R132H) négatives. Aucune analyse de séquençage n'a mis en évidence une autre mutation IDH non analysable en immunohistochimie. Ainsi, 46 tumeurs (92 %) sont classées IDH négatives. La perte de l'expression d'ATRX a été observée dans 7 cas. La méthylation du promoteur MGMT a été observée dans 17 cas (34 %). Ainsi, 33 tumeurs (66 %) n'observent pas de méthylation du promoteur de la MGMT. L'amplification du gène EGFR a été mise en évidence dans 24 cas (48 %). Le gain du chromosome 7 associé à la perte du chromosome 10 a été retrouvé dans 35 cas (70 %). La délétion homozygote CDKN2A a été retrouvée dans 29 cas (59 %) et une amplification CDKN2A a été retrouvée dans seulement 1 cas. A noter que l'étude de la mutation du promoteur TERT n'est pas réalisée au laboratoire du CHU.

Selon les dernières données épidémiologiques, l'âge médian des patients atteints de glioblastome est de 65 ans (Ostrom et al., 2018). L'âge médian des patients de l'étude est plus faible (59 ans), principalement dû au biais de sélection des patients, en effet pour participer à l'étude les patients doivent être en assez bon état général pour supporter la neurochirurgie. La comparaison avec les données issues de la littérature montre le même ratio de patient IDH muté / non muté que dans notre étude, à savoir 10 % / 90 %.

Par ailleurs, les données de survie des patients de l'étude sont en accord avec les facteurs de bon pronostic avérés. Ainsi, l'indice de Karnofsky et le statut de méthylation du promoteur MGMT sont des indicateurs de la survie (Figure 1-5). La qualité de la résection ne semble cependant pas déterminante dans la survie globale des patients de l'étude (Figure 1-5).

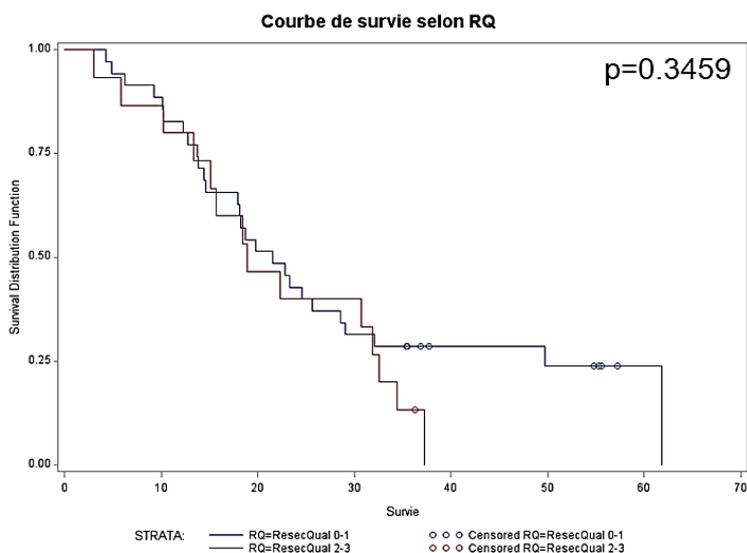
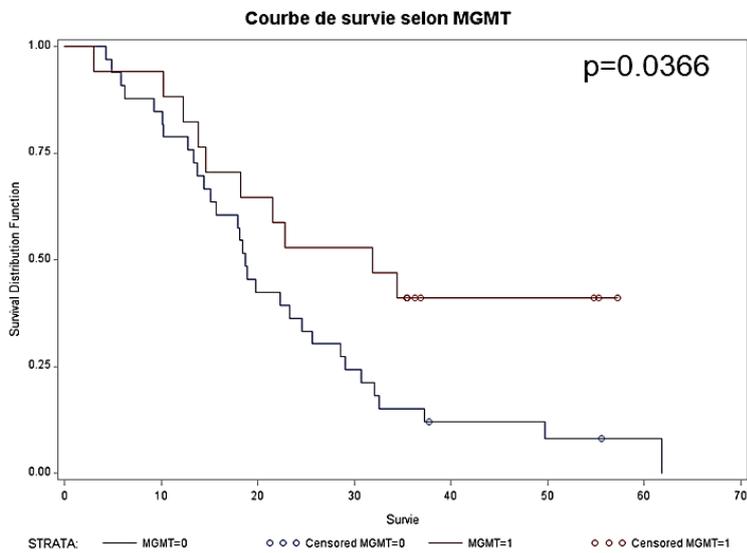
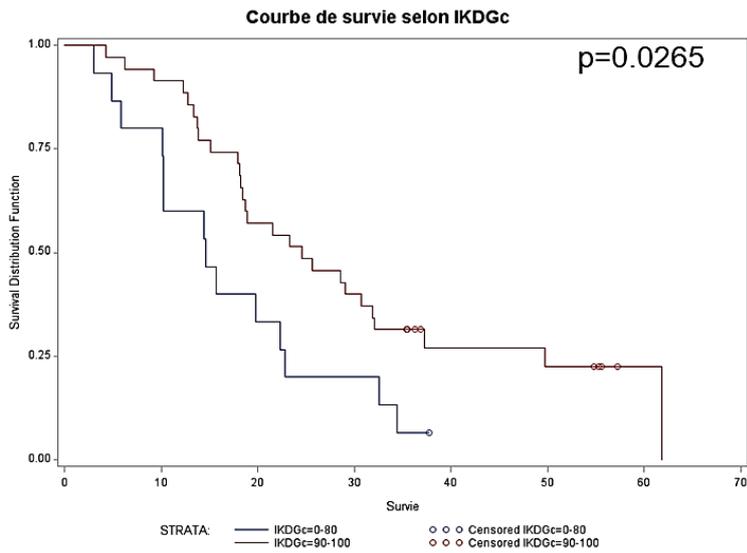


Figure 1-5 Courbes de survie en fonction des différents facteurs pronostics

Courbe de survie globale de l'ensemble des patients selon l'indice de Karnofsky, le statut MGMT et la qualité de la résection. 0-1 = résection totale, sous-totale 2-3 = biopsies de résection partielle

En dépit de ces indicateurs, la survie globale des patients reste très variable et est impossible à prévoir. De nouveaux biomarqueurs pronostics sont donc requis pour améliorer la prise en charge des patients.

L'ensemble des données cliniques et histologiques montre l'hétérogénéité des glioblastomes. Compte tenu de cette hétérogénéité, nous avons par la suite réalisé l'image MALDI de chacune des tumeurs afin de déterminer leurs compositions protéiques.

B. L'imagerie par spectrométrie de masse MALDI permet de regrouper les patients en fonction de leurs caractéristiques moléculaires

Tout d'abord, les annotations histologiques ont été comparées aux segmentations individuelles obtenues par imagerie MALDI réalisées sur l'ensemble des tumeurs. La segmentation individuelle correspond à l'analyse individuelle de chaque échantillon. Pour les segmentations individuelles, tous les spectres de masse générés au sein de chaque tissu sont comparés dans le but de distinguer les régions de compositions protéiques différentes. Dans ce cas, les échantillons ne peuvent pas être comparés entre eux.

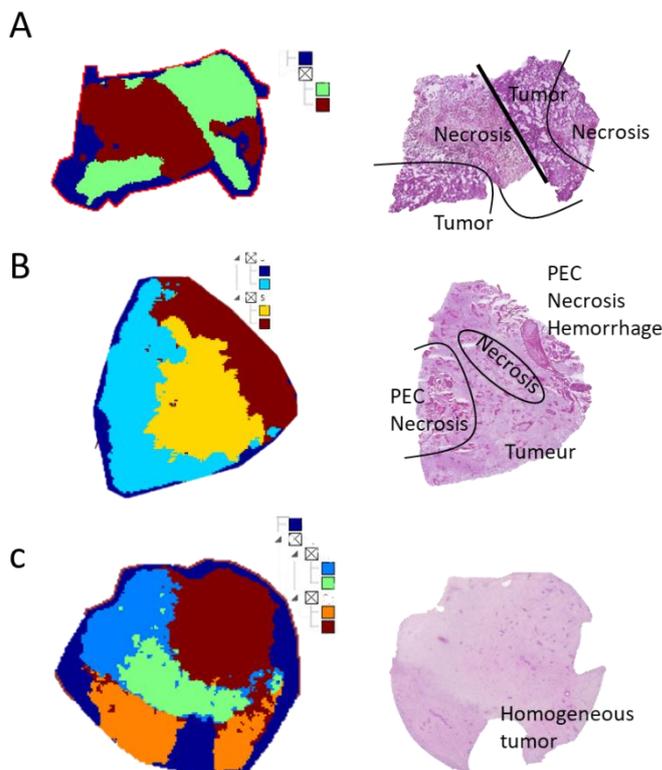


Figure 1-6 Les annotations histologiques ne révèlent pas la complexité moléculaire des échantillons.

Résultats représentatifs de la comparaison. A. Les annotations histologiques sont similaires à la segmentation spatiale obtenue après imagerie MALDI. B. L'imagerie révèle des zones moléculaires différentes des annotations. C. Les tumeurs dites homogènes sont en réalité complexes.

La comparaison entre les annotations histologiques et les images moléculaires a permis de mettre en évidence des différences pour de nombreux échantillons. En effet, de nouvelles régions d'intérêt ont pu être mises en évidence grâce à l'analyse MALDI par rapport à l'analyse histologique classique (Figure 1-6). Une faible concordance est observée entre ces deux techniques d'analyse. Les tumeurs paraissant plutôt homogènes d'un point de vue histologique se sont révélées complexes d'un point de vue moléculaire.

Suivant ces observations, une segmentation globale des 50 échantillons a été réalisée dans le but de comparer les cas entre eux et de déterminer si certaines caractéristiques étaient communes. Ainsi, trois grandes régions ont été identifiées : la région rouge, la région jaune et la bleue (Figure 1-7). Chacune de ces régions est constituée de signatures spectrales similaires et représente donc une zone moléculaire homogène. Les trois régions correspondent à trois zones moléculaires différentes. Nous avons ensuite affecté chaque échantillon de tumeur à un groupe, le regroupement des patients se faisant à partir des régions identifiées. Chaque zone moléculaire est représentative d'un groupe. Compte tenu de l'hétérogénéité des tissus et pour inclure le maximum de patient, un seuil de 50% a été déterminé, si une région constitue la majorité (+ 50 %) de la surface de l'échantillon tumoral, le patient est considéré comme faisant partie du groupe associé à cette région. Ainsi, les échantillons constitués en majorité de la région rouge forment le groupe A ; ceux constitués en majorité de la région jaune forment le groupe B et ceux constitués en majorité de bleu forment le groupe C. Les patients d'un groupe donné partagent donc des caractéristiques communes et sont très similaires entre eux.

Le groupe A est formé par les patients dont la tumeur apparaît majoritairement rouge. Pour les 13 patients faisant partie de ce groupe on observe 96.8 % de zone rouge, 0.35 % de jaune et 2.8 % de bleu. Le groupe B contient 9 patients et est à 75.9 % de couleur jaune (7.7 % rouge et 16.3 % bleu). Le groupe C contient 23 patients à dominante bleu (83.87 % bleu ; 8.4 % jaune et 7.7 % rouge).

Cette méthode ne permet pas de grouper tous les patients. Cinq patients présentent un mélange hétérogène des trois couleurs, sans aucune couleur réellement majoritaire. Ces patients ne sont donc pas classables avec cette technique. Pour la suite des analyses nous avons écarté ces 5 échantillons.

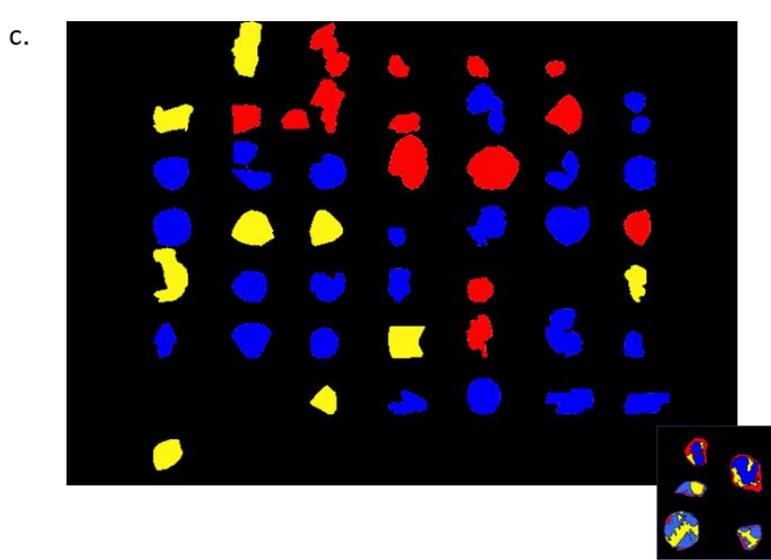
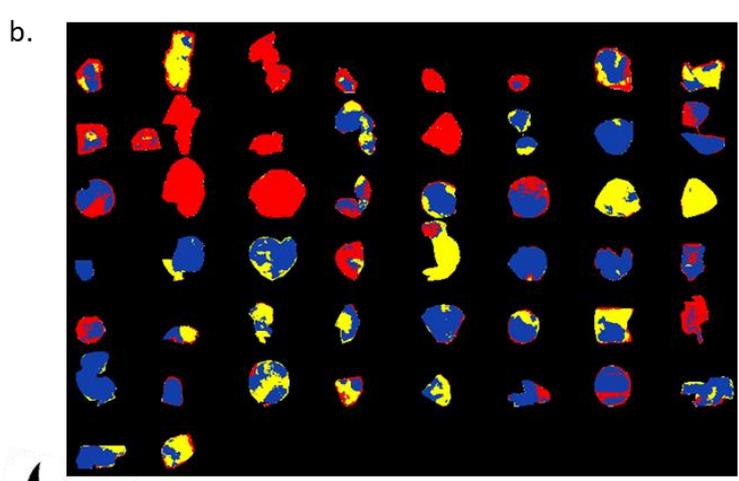


Figure 1-7 Segmentation globale et regroupement des patients

(a,b) La segmentation globale a permis de mettre en évidence 3 zones moléculaires distinctes au sein de l'ensemble des tumeurs. (c) Les patients ont ensuite été regroupés suivant la prédominance de chaque zone. 5 patients n'ont pas pu être regroupé selon cette méthode

Au sein d'un groupe, tous les patients partagent des caractéristiques moléculaires communes, ce qui signifie que les spectres dans chacune des régions colorées sont très similaires. Ainsi, des ions spécifiques peuvent être attribués à chaque région : les ions m/z 967.621 et 1492.916 sont spécifiquement présents dans le groupe A tandis que les m/z 1914.591, 2375.074 et 2376.274 sont spécifiques du groupe B et finalement m/z 1473.312, 2045.815, 2046.615 et 2237.849 sont spécifiques du groupe C (Figure 1-8).

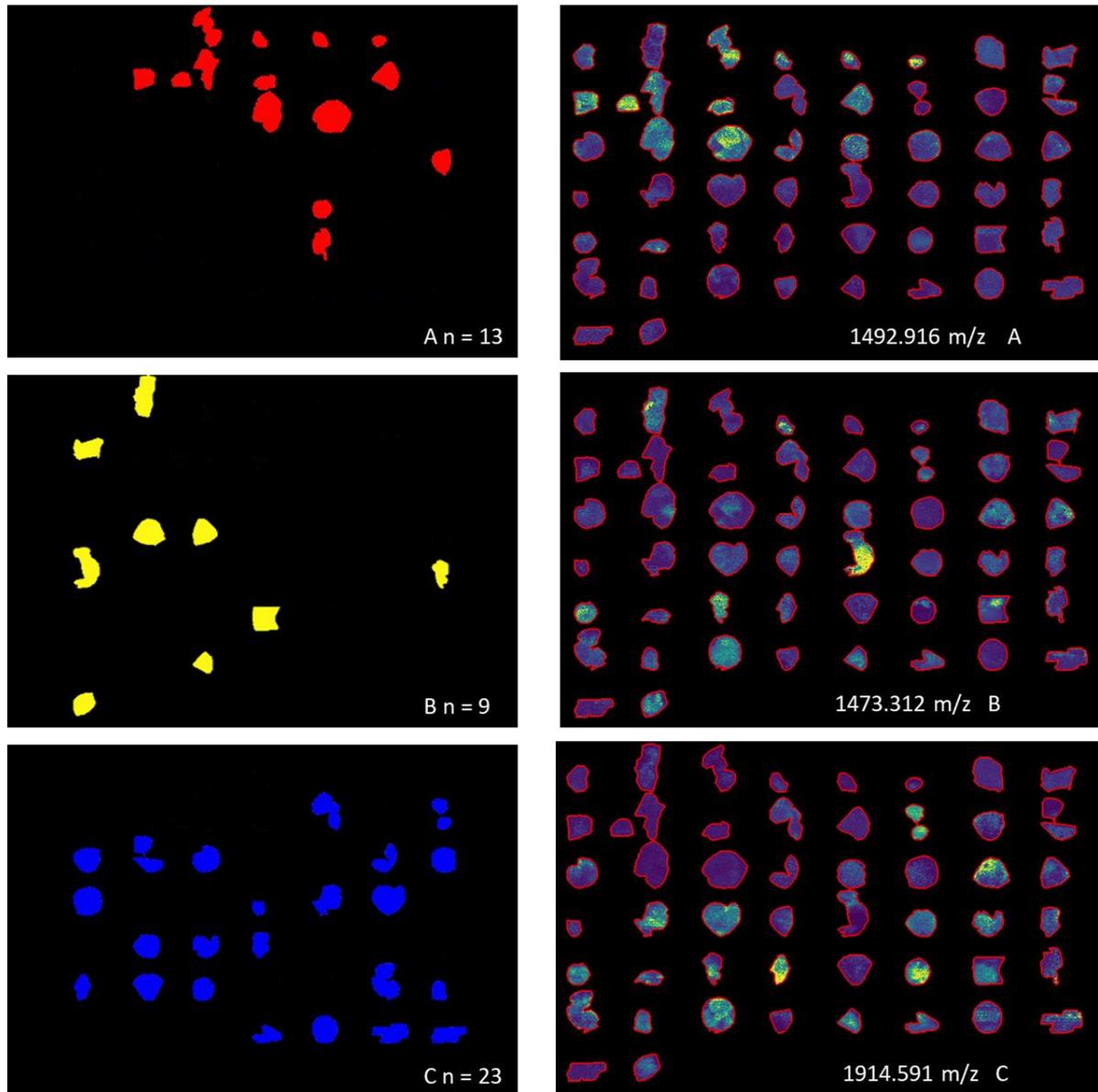


Figure 1-8 Ions caractéristiques de chaque groupe

A gauche, les échantillons de tumeurs dans chaque groupe avec à droite un m/z caractéristique du groupe.

Finalement, l'imagerie MALDI permet d'observer l'hétérogénéité des glioblastomes au niveau protéique et démontre l'existence de nouvelles régions d'intérêt. A partir de cette hétérogénéité il est facile de regrouper les patients selon leurs caractéristiques communes. Ainsi 3 groupes ont pu être définis sur la base de la segmentation.

Dans un deuxième temps, nous avons voulu consolider cette classification avec les données obtenues par SpiderMass. Le SpiderMass est une technologie utilisant la spectrométrie de masse et proche de la technologie MALDI. Ce système utilise un laser fibré qui est réglé pour exciter les molécules d'eau. L'eau agissant comme une matrice on parle de désorption / ionisation laser assistée par l'eau (WALDI : *Water Assisted Laser Desorption Ionization*) (Ogrinc et al., 2019; Saudemont et al., 2018).

C. L'analyse SpiderMass confirme le groupement obtenu par imagerie

Pour valider la classification obtenue, un total de 30 échantillons a été analysé via la technologie SpiderMass. L'analyse de l'ensemble des tumeurs n'était pas possible, certains échantillons ayant été utilisés pour d'autres expérimentations. Dans cette analyse, les molécules d'intérêt sont les lipides. Sur chaque tissu des points d'intérêts ont été définis selon les zones d'intérêt identifiées par la segmentation globale obtenue par imagerie MALDI. Chaque point a été irradié par le laser pour finalement générer les spectres lipidiques.

Une classification PCA-LDA (analyse composante principale-analyse discriminante linéaire) a été réalisée pour analyser les spectres obtenus par l'analyse des tumeurs au SpiderMass. La première étape de ce test, PCA, consiste à réduire la dimensionnalité des données en générant de nouvelles variables, appelés composante principale, afin d'expliquer la plupart des variabilités observées. Une analyse supervisée, LDA, est ensuite réalisée à partir des composantes principales créées en PCA. La classification PCA-LDA se base sur les 3 groupes prédéterminés par imagerie MALDI : A ; B et C.

A partir des spectres obtenus et des groupes définis, le modèle de classification PCA-LDA détermine les paramètres spécifiques à chaque groupe. Ainsi, le premier axe de l'analyse LDA (LDA) permet de discriminer le groupe A du groupe C et le deuxième axe (LDA) de séparer le groupe B des groupes A et C (Figure 1-9). Les ions les plus discriminants de chaque groupe

ont été définis. Par exemple les pics les plus discriminants pour le groupe A sont m/z 746,75, 810,65 et 916,65. Pour le groupe B, les m/z 718,55, 724,65, 751,55 et 778,55 sont les plus discriminants. Finalement, les ions les plus discriminants pour le groupe C, correspondent aux m/z 725,4, 754,6, 788,65 et 850,85 (Figure 1-9).

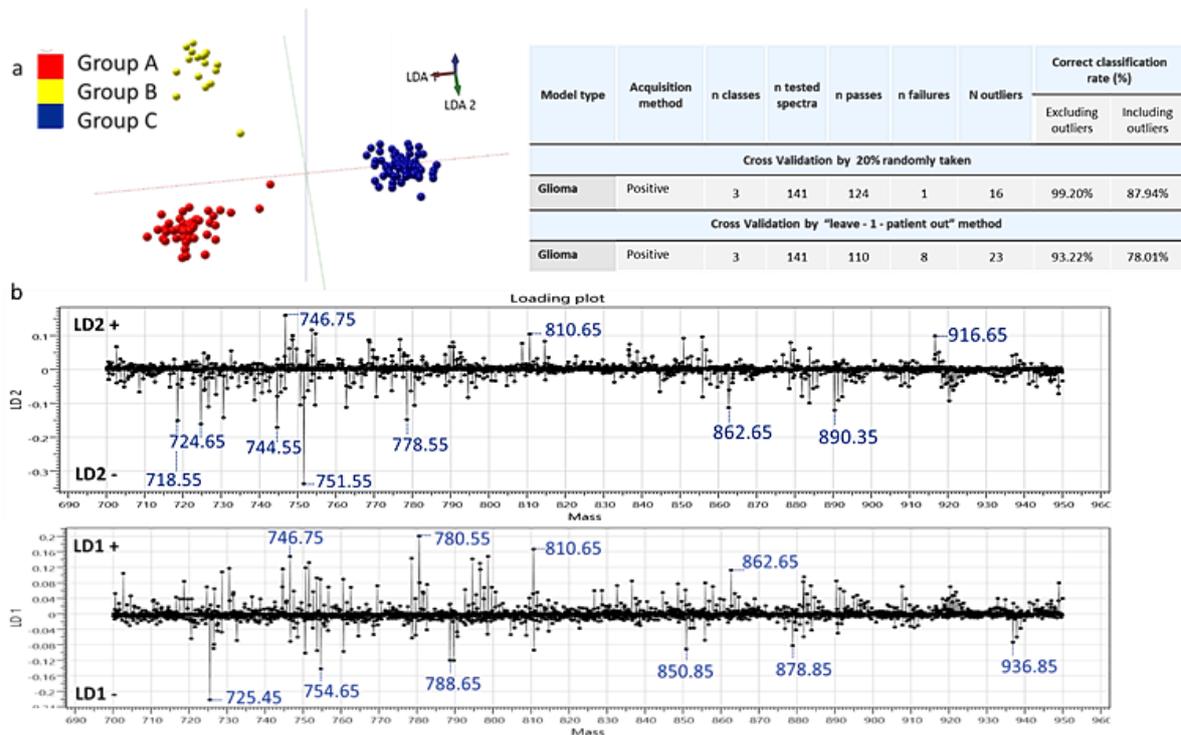


Figure 1-9 Modèle de classification PCA-LDA basé sur les 3 groupes de gliomes

a) Représentation LDA de l'analyse PCA-LDA à 3 classes. Le tableau (à droite) représente les résultats de la validation croisée «20 % out» et «Leave-one-patient-out» du modèle de classification. b) Les spectres de chargement LD2 (en haut) indiquent les pics discriminants entre le groupe A (rouge) et le groupe B (jaune). Les 10 pics lipidiques les plus discriminants sont indiqués par le trait bleu. Les spectres de chargement LD1 (en bas) indiquent les pics discriminants entre le groupe A (rouge) et le groupe C (bleu). Les 10 pics lipidiques les plus discriminants sont indiqués par la ligne bleue en pointillés

Pour consolider la classification, une validation a été effectuée en utilisant les méthodes « 20 % out » et « leave one patient out ». La première méthode du retrait aléatoire de 20 % des patients montre des taux de classification corrects de 99,20 % et 87,94 % respectivement avec et sans valeurs aberrantes. Les valeurs aberrantes représentent les valeurs qui n'ont pas pu être réattribuées lors de la classification. La deuxième méthode de classification, la méthode de sortie d'un patient, montre un score de classification correct de 93,22 % et 78,01 %, avec ou sans les valeurs aberrantes respectivement.

Ces analyses permettent de conclure que les profils moléculaires obtenus en MS sont différents selon le groupement. Ainsi, l'analyse LDA des données SpiderMass a permis de regrouper les échantillons de la même manière que la classification MALDI-MSI.

D. Caractéristiques cliniques des groupes identifiés

Dans un premier temps, les données d'imagerie ont été comparées à l'analyse histopathologiques des tumeurs (Figure 1-10).

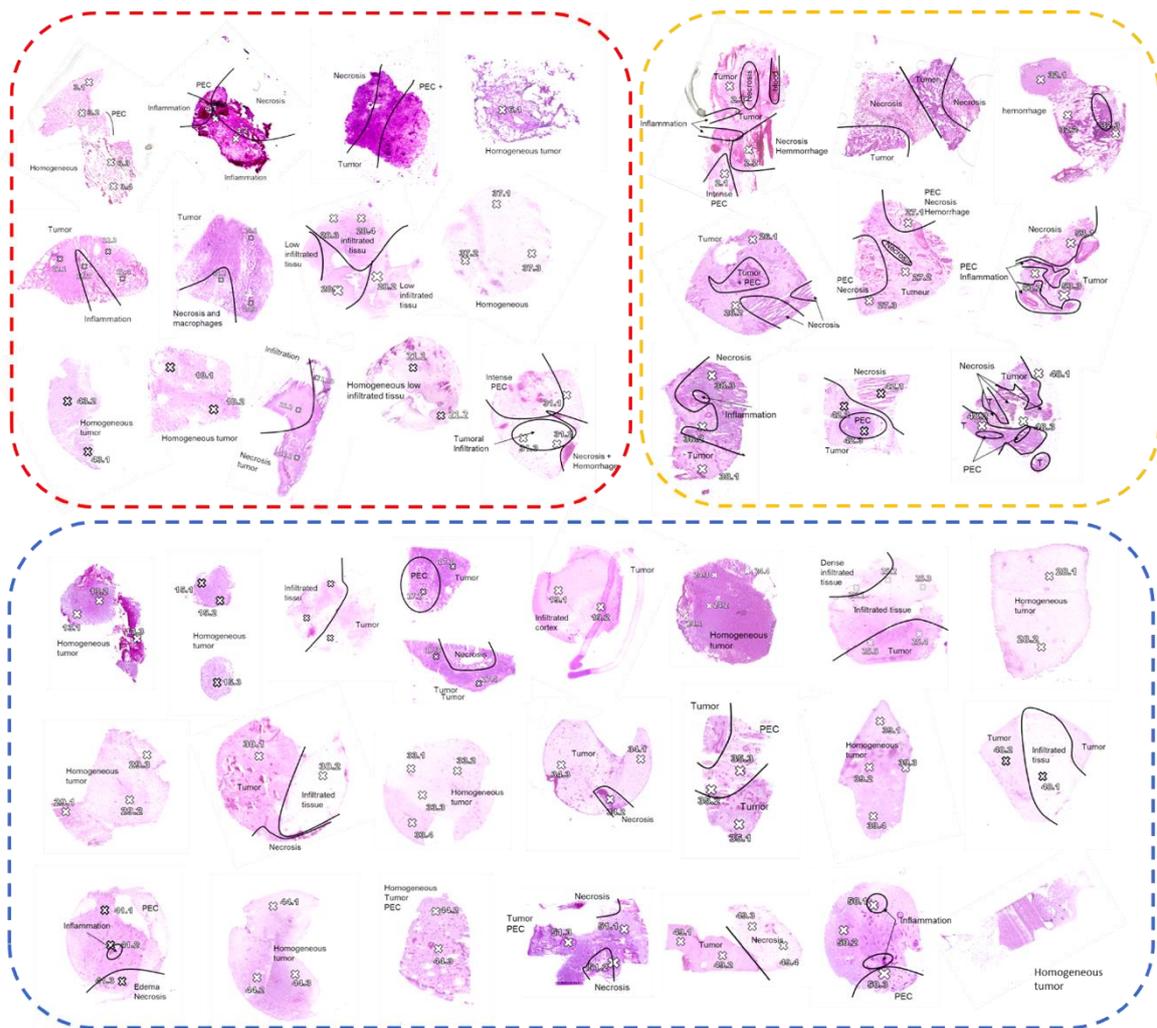


Figure 1-10 Annotations histologiques des échantillons tumoraux par groupe

Les échantillons sont répartis selon les groupes A, B et C identifiés par imagerie. Chaque groupe est encadré en pointillés : rouge pour le groupe A, jaune pour le groupe B et bleu pour le groupe C.

Ainsi, l'annotation anatomopathologique révèle que le groupe A est hétérogène et présente des zones de nécrose, d'inflammation et d'infiltration de cellules tumorales. Le

groupe B, également hétérogène, observe une grosse proportion de tumeurs avec des zones intenses de prolifération endothélio-capillaire accompagnées de zones d'inflammation avec parfois présence d'hémorragie. Au contraire, les tumeurs du groupe C sont le plus souvent homogène (Figure 1-10).

Les caractéristiques cliniques des patients ont par la suite été reportées selon les trois groupes (Tableau 1-2). Ainsi, les patients du groupe A sont plus jeunes que les patients des groupes B et C et présentent également moins de méthylation au niveau du promoteur MGMT. Plus de la moitié de ces patients (62 %) ne présente pas d'amplification EGFR. Le groupe B a le plus haut niveau de méthylation MGMT et d'amplification EGFR (89 %). Enfin, 57 % des patients du groupe C présentent une amplification EGFR.

Tableau 1-2 Caractéristiques cliniques des patients en fonction des groupes identifiés par imagerie MALDI

	A (n=13) (rouge)	B (n=9) (jaune)	C (n=23) (bleu)
Sexe			
<i>Femme, n (%)</i>	4 (31)	2 (22)	7 (30)
<i>Homme, n (%)</i>	9 (69)	7 (78)	16 (70)
Age au diagnostic (année) Médiane (IQ)	48 (43-55)	60 (59-64)	62 (51-68)
Indice de Karnofsky au diagnostic médiane (IQ)	90 (70-100)	90 (90-100)	90 (80-95)
<i>0-80, n (%)</i>	5 (38)	0 (0)	8 (35)
<i>90-100, n (%)</i>	8 (62)	9 (100)	15 (65)
Localisation			
<i>frontal, n (%)</i>	4 (31)	1 (11)	9 (39)
<i>occipital, n (%)</i>	1 (8)	0 (0)	2 (9)
<i>pariétal, n (%)</i>	4 (31)	3 (33)	4 (17)
<i>temporal, n (%)</i>	4 (31)	5 (55)	7 (30)
Qualité de la chirurgie			
<i>complète, n (%)</i>	6 (46)	8 (89)	10 (43)
<i>partielle, n (%)</i>	7 (54)	1 (11)	12 (52)
<i>biopsies, n (%)</i>	0	0	1 (4)
Statut IDH1/2*			
<i>Non muté, n (%)</i>	11 (85)	9 (100)	21 (91)
<i>mutant, n (%)</i>	2 (15)	0 (0)	2 (9)
ATRX			
<i>marquage non conservé, n (%)</i>	2 (15)	0 (0)	4 (17)
<i>marquage conservé, n (%)</i>	11 (85)	9 (100)	19 (83)
Statut de méthylation du promoteur MGMT			
<i>non méthylé, n (%)</i>	11 (85)	4 (44)	13 (57)
<i>méthylé, n (%)</i>	2 (15)	5 (56)	10 (43)

<i>Amplification EGFR</i>			
<i>non, n (%)</i>	8 (62)	5 (56)	10 (43)
<i>oui, n (%)</i>	5 (38)	4 (44)	13 (57)
<i>Gain chromosome 7 et perte chromosome 10</i>			
<i>non, n (%)</i>	5 (38)	1 (11)	8 (35)
<i>oui, n (%)</i>	8 (62)	8 (89)	15 (65)
<i>Amplification EGFR ou gain 7 / perte 10</i>	8 (62)	8 (89)	20 (87)
<i>Amplification EGFR sans gain 7 / perte 10</i>	0 (0)	0 (0)	6 (26)
<i>Amplification EGFR et gain 7 / perte 10</i>	5 (38)	4 (44)	7 (30)
<i>Gain 7 / perte 10 sans amplification EGFR</i>	3 (2)	4 (44)	7 (30)
<i>Délétion homozygote CDKN2A</i>			
<i>non, n (%)</i>	6 (46)	3 (33)	8 (35)
<i>oui, n (%)</i>	6 (46)	6 (67)	15 (65)
<i>autre: amplification de CDKN2A, n (%)</i>	1 (8)	0 (0)	0 (0)
<i>Traitement initial</i>			
<i>RT/TMZ puis 6 cycles de TMZ, n (%)</i>	5 (38)	6 (67)	8 (35)
<i>RT/TMZ puis au moins 6 mois de TMZ, n (%)</i>	1 (8)	0 (0)	3 (13)
<i>RT/TMZ puis moins de 6 mois de TMZ, n (%)</i>	4 (31)	2 (22)	11 (48)
<i>Autre traitement**, n (%)</i>	2 (15)	0 (0)	0 (0)
<i>Pas de traitement, n (%)</i>	0 (0)	0 (0)	1 (4)
<i>Etude clinique, n (%)</i>	1 (8)	1 (11)	0 (0)
<i>Progression</i>			
<i>oui, n (%)</i>	12 (92)	7 (78)	17 (74)
<i>non, n (%)</i>	1 (8)	1 (11)	3 (13)
<i>non connue, n (%)</i>	0 (0)	1 (11)	3 (13)
<i>Médiane de survie sans progression (mois)</i>	10.5 (6.5-32.5)	12.5 (10-19)	12 (6-14.5)
<i>Traitement à la première progression</i>			
<i>oui, n (%)</i>	11 (92)	7 (100)	14 (82)
<i>non, n (%)</i>	1 (8)	0 (0)	3 (18)
<i>Décès</i>			
<i>oui, n (%)</i>	11 (85)	7 (78)	18 (78)
<i>non, n (%)</i>	2 (15)	2 (22)	5 (22)
<i>Médiane de survie après chirurgie (mois)</i>	24.5 (13.5-49.5)	23 (15-29)	20 (17-33)
<i>Survie</i>			
<i>Quartile supérieur, n (%)</i>	4 (31)	2 (22)	4 (17)
<i>Interquartile, n (%)</i>	4 (31)	6 (67)	14 (61)
<i>Quartile inférieur, n (%)</i>	5 (38)	1 (11)	5 (22)
<i>Epilepsie</i>			
<i>non</i>	6 (46)	2 (22)	9 (39)
<i>Avant le diagnostic ou au diagnostic</i>	6 (46)	7 (77)	14 (60)
<i>Pendant le suivi</i>	1 (8)	0 (0)	0 (0)
<i>Evènement thrombotique</i>			
<i>non</i>	13 (100)	8 (89)	21 (91)
<i>Avant le diagnostic ou au diagnostic</i>	0 (0)	1 (11)	1 (4)
<i>Pendant le suivi</i>	0 (0)	0 (0)	1 (4)

Afin de mieux comprendre les différences moléculaires observées entre les groupes, l'analyse de microprotéomique localisée sur tissu a été réalisée sur les 45 patients. L'imagerie de type MALDI donne la possibilité de localiser différentes molécules au sein d'un tissu et ainsi déterminer des signatures moléculaires selon l'hétérogénéité de la tumeur. Cependant elle ne donne pas accès directement à l'identification de ces marqueurs.

E. Identification de voies de signalisation spécifiques de chaque groupe

L'imagerie MALDI permet de guider les analyses par microprotéomique de façon localisée sur les coupes de tissu. Il est ainsi possible de garder l'information de localisation des protéines identifiées. Ces études de protéomique localisée consistent en des digestions enzymatiques localisées suivies de micro-extractions liquides sur des zones d'intérêt pour effectuer ensuite une analyse protéomique à large échelle basée sur la spectrométrie de masse (ETAT DE L'ART Figure 0-9). Cette analyse se fait à l'aide d'un système nanoLC couplé à un spectromètre de masse de haute résolution spectrale. Ainsi, sur chaque tissu, des points de microprotéomique ont été définis en fonction de l'hétérogénéité des tumeurs afin de mieux comprendre les mécanismes biologiques impliqués dans chaque zone.

L'analyse protéomique a été réalisée sur les 45 patients qui ont pu être groupés par imagerie MS. Ainsi, 147 régions ont été sélectionnées et microdigérées à partir des données d'imagerie MALDI. Chaque point d'extraction a donc été associé à chacune des trois zones et chacun des trois groupes. Sur chaque tissu, entre 2 et 5 points d'extractions ont été sélectionnés afin d'analyser l'hétérogénéité des tumeurs. Parmi l'ensemble des tissus, 32 extractions ont été faites dans la zone Rouge, 30 dans la zone Jaune et 85 dans la Bleue. Concernant les groupes de patient, on retrouve 35 extractions dans les tissus du groupe A (dont 26 dans la zone Rouge), 26 dans le groupe B (dont 15 dans la zone Jaune) et 70 dans le groupe C (dont 60 dans la zone Bleu) (Tableau 1-3).

Tableau 1-3 Nombres d'extractions réalisées dans chacun des groupes

	A	B	C	Inclassables	TOTAL
<i>Bleu</i>	7	10	60	7	84
<i>Jaune</i>	2	15	6	8	31
<i>Rouge</i>	26	1	4	1	32
TOTAL	35	26	70	16	147

Les peptides extraits dans chacune de ces régions ont été analysés par LC-MS/MS en vue de leurs identifications. Après interrogation des banques de données, un total de 4936 protéines a été identifié. Pour la suite des analyses nous nous sommes focalisés sur les extractions provenant des régions majoritaires correspondant aux groupes de patients. C'est-à-dire les extractions de la zone rouge des échantillons de patients du groupe A ; de la zone jaune du groupe B et de la zone bleue du groupe C. Ces extractions ont servi de point de départ à l'analyse statistique.

Un test ANOVA (FDR Benjamini Hochberg 0,05) a été réalisé en groupant les extractions selon les annotations Rouge_A ; Jaune_B et Bleu_C. Au total, 1231 protéines ont montré une différence d'expression significative entre les trois groupes, ces protéines ont été représentées sous la forme d'un HeatMap, où les protéines surexprimées sont en rouge et les sous-exprimées en vert (Figure 1-11).

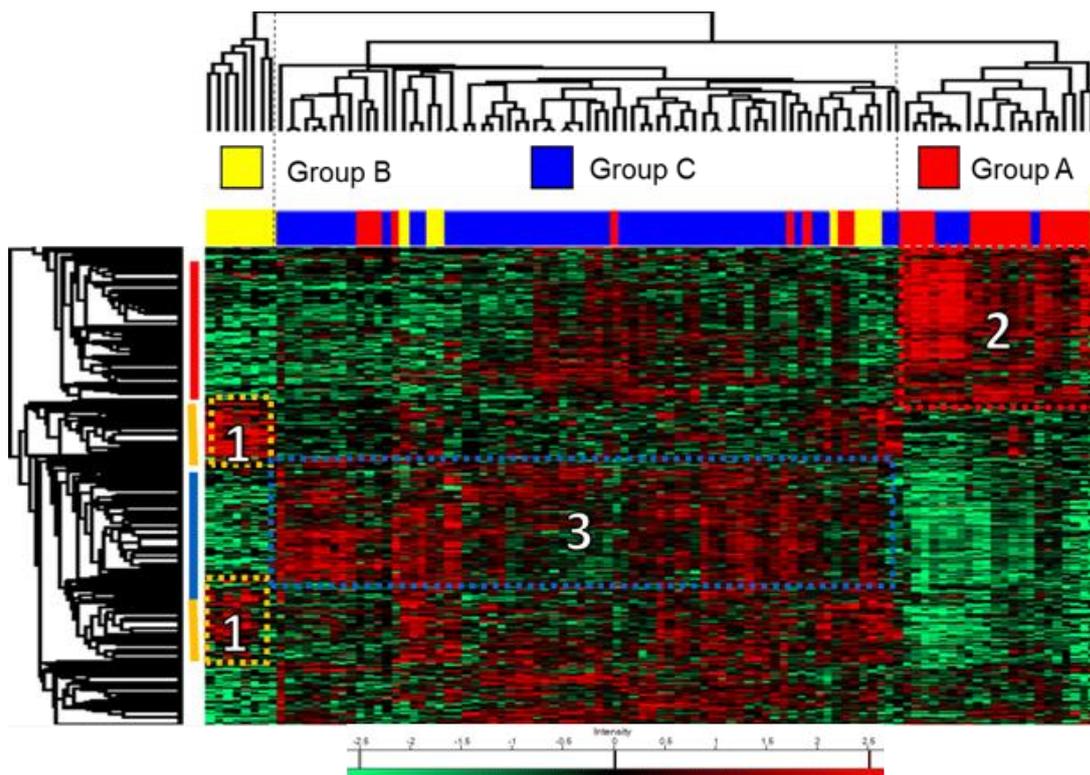


Figure 1-11 Correspondance entre la classification obtenue suite aux analyses d'imagerie MALDI-MSI et les analyses de microprotéomique

Analyse statistique (ANOVA, Benjamini-Hochberg $FDR < 0.05$) des protéines identifiées sur les points de micro-extractions correspondant à la couleur représentative du groupe de patient (Rouge pour le groupe A ; jaune pour le groupe B et bleu pour le groupe C). On observe pour chaque groupe un cluster de protéines surexprimées.

Dans un premier temps, on observe que les échantillons se regroupent selon 3 branches principales qui représentent les groupes de patients. La première branche du regroupement des échantillons (la plus à gauche) contient exclusivement des extractions issues de la zone jaune du groupe B. La deuxième branche contient 77 % d'extractions de la zone bleue du groupe C et 10 % de la zone jaune du groupe B et 13 % de la zone rouge du groupe A. Finalement la troisième branche contient 77 % d'extractions de la zone rouge du groupe A et 23 % de la zone bleu du groupe C. Ces 3 branches sont cohérentes avec les régions identifiées par la segmentation spatiale MALDI. Ainsi, les échantillons issus de la même région colorée ont effectivement le même profil protéomique.

Le Heatmap généré permet de mettre en lumière des clusters de protéines surexprimées et sous exprimées dans chacun des trois groupes. Ainsi, trois clusters de protéines se démarquent. Le cluster 1 correspond aux protéines spécifiquement surexprimées dans la branche regroupant majoritairement des échantillons du groupe B. Le cluster 2 aux protéines surexprimées dans la branche regroupant majoritairement des échantillons du groupe A. Le cluster 3 représente les protéines surexprimées dans la branche regroupant majoritairement des échantillons du groupe C. La liste des protéines de chaque cluster est reportée en ANNEXE. L'analyse de ces protéines permet finalement d'identifier les processus biologiques spécifiques de chaque groupe.

Les protéines surexprimées dans le groupe A (cluster 2) sont associées au système nerveux. En effet, la plupart des protéines de ce cluster sont impliquées dans la neurogénèse et la guidance axonale comme les neurofilaments (NEFH ; NEFL et NEFM), les dihydropyrimidinase-related protein 1 ; 4 et 5 (CRMP1, DPYSL4 et DPYSL5), la neuromodulin (GAP43), microtubule-associated protein tau (MAPT), kinesin-like protein KIF2A (KIF2A), unconventional myosin-XVIIIa (MYO18A) et alpha-internexin (INA). De plus certaines protéines de ce groupe ont déjà été identifiées surexprimées dans des glioblastomes comme les protéines CRMP1, DPYSL2 (i.e. CRMP2), (Jovčevska et al., 2017), DPYSL5 (i.e. CRMP5) (Moutal et al., 2015), GAP43 (Gutmann et al., 2002; Huang et al., 2003; Voigt et al., 2017).

Par la suite, l'analyse de ces protéines via String permet d'identifier les réseaux d'interaction protéiques spécifiques. String est une base de données recensant les interactions protéine-protéine connues et prédites, ces interactions comprennent les

associations directes (physiques) et indirectes (fonctionnelles). Ainsi, les voies de signalisation principales identifiées dans ce cluster sont le développement du système nerveux, la projection neuronale et le guidage axonale (Figure 1-12).

L'analyse des principales voies de signalisation a également permis de confirmer l'implication des protéines identifiées dans la croissance des neurites, la synaptogénèse et la plasticité synaptique (Figure 1-12). Ces analyses permettent de faciliter l'interprétation des données et d'avoir un aperçu des voies de signalisation impliquées dans chaque groupe.

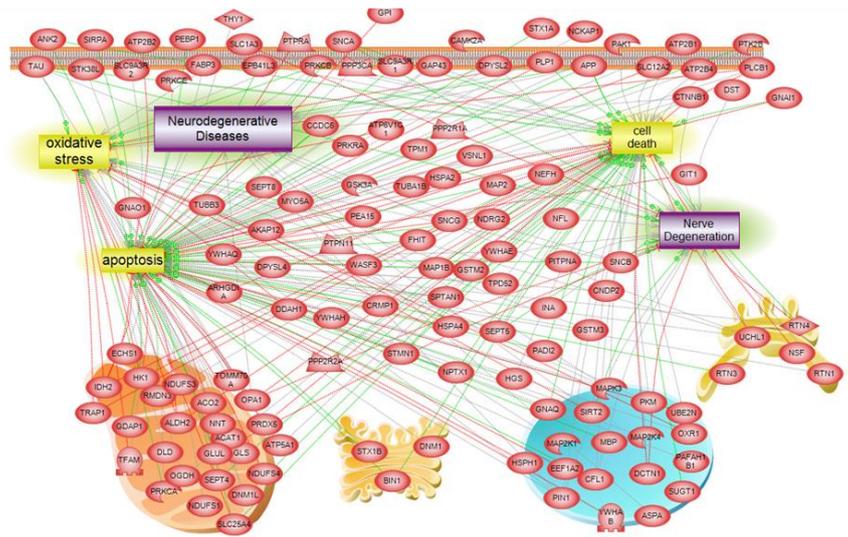
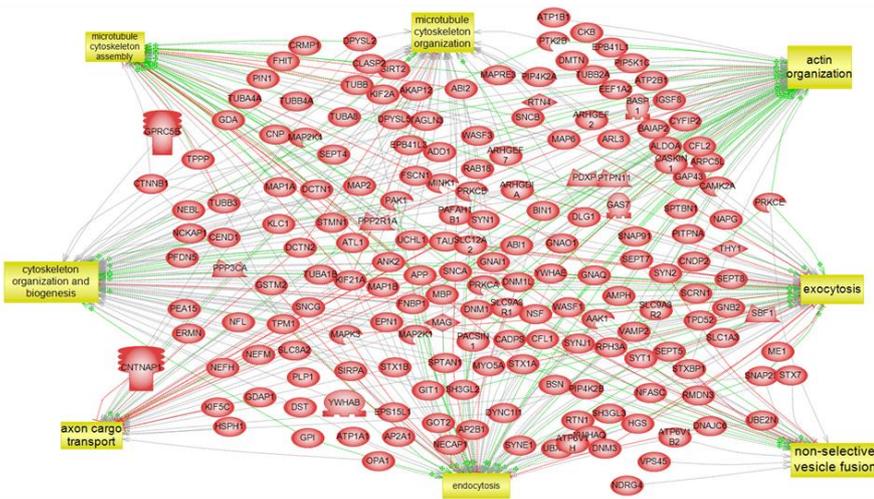
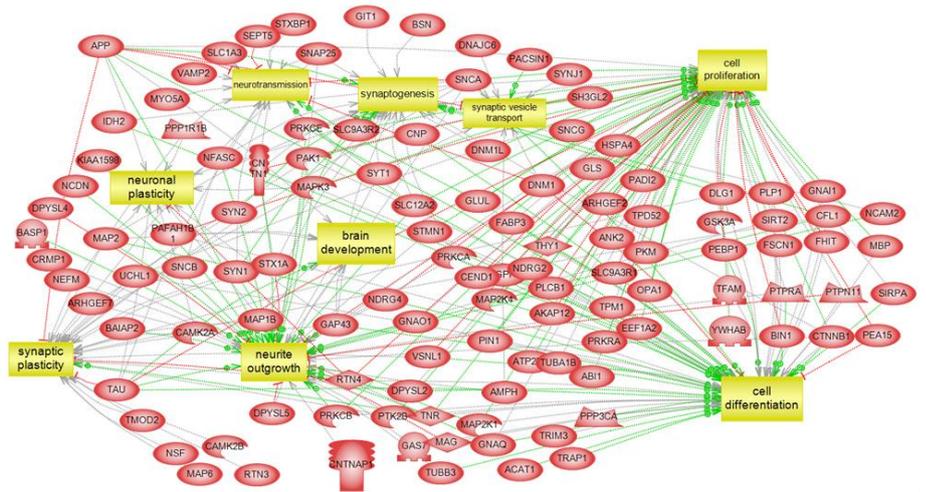


Figure 1-12 Voies de signalisations surexprimées dans le groupe A (cluster 2)

L'analyse des voies de signalisations du groupe A révèle que les protéines surexprimées dans ce groupe sont impliquées dans a) la différenciation cellulaire, la croissance des neurites, la plasticité neuronale et b) l'organisation de l'actine et du cytosquelette et du transport axonale. De plus, une partie est également associée aux maladies neurodégénératives et plus généralement à la dégénération des nerfs.

Finalement, ce groupe de patient a un profil neurologique, avec de nombreuses protéines impliquées dans les processus de développement du système nerveux. En effet, on retrouve des marqueurs du développement du système nerveux comme les dihydropyrimidinase-related protein (CRMP), neuromodulin (GAP43), microtubule-associated protein tau (MAPT) ainsi que des marqueurs de développement et de différenciation des oligodendrocytes comme abl interactor 1 (ABI1), aspartoacylase (ASPA), 2',3'-cyclic-nucleotide 3'-phosphodiesterase (CNP) et contactin-associated protein 1 (CNTNAP1).

Il est également intéressant de noter la présence de marqueurs de cellules souches et cellules progénitrices comme tripartite motif-containing protein 2 (TRIM2). Les cellules progénitrices neurales impliquent la présence de marqueurs de neurones immatures comme tubulin beta-3 chain (TUBB3), de marqueurs de neurones matures tel que NeuN (ARN binding protein fox-1 homolog 3, RBFOX3) et des marqueurs de synapses comme synaptophysin (SYP) et synaptic vesicle glycoprotein 2A (SV2A) (Beier et al., 2018). Parmi les protéines identifiées comme surexprimées dans ce groupe on retrouve notamment stathmin (STMN1) et les neurofilaments NEFH, NEFM et NEFL qui sont également des marqueurs des cellules progénitrices neurales et des cellules souches de gliomes (Neftel et al., 2019). Les cellules souches gliomales possèdent de nombreuses caractéristiques communes avec les cellules souches neurales normales (NSC) : la capacité de se renouveler, de proliférer indéfiniment et de se différencier en différents types de cellules. Ces cellules souches cancéreuses ont déjà été identifiées et isolées au sein de glioblastomes (Lathia et al., 2015).

Les protéines surexprimées dans le groupe B (principalement représenté dans le cluster 1) correspondent à une activation microgliale et de façon plus générale à l'activation du système immunitaire. On observe la présence d'immunoglobuline dans ce groupe (Ig mu chain C region ; Ig gamma-3 chain C region ; Ig gamma-2 chain C region ; Ig kappa chain C region ; Ig gamma-1 chain C region ; Ig lambda-6 chain C region ; Ig alpha-1 chain C region) ainsi que des marqueurs de cellules immunitaires comme les macrophages (Macrophage-capping protein). Parmi ces protéines, on retrouve également des protéines impliquées dans la réponse immunitaire comme le complément C1q subcomponent subunit C et B (C1QC et C1QB), haptoglobine (HP), kininogène 1 (KNG1), histidine-rich glycoprotein (HRG), transthyrétine

(TTR), grancalcin (GCA), protéines S100-A9 (S100A9) et S100-A12 (S100A12), erythrocyte band 7 integral membrane protein (STOM) et galectin-3-binding protein (LGALS3).

Cette implication du système immunitaire a été validée par une immunohistochimie contre la Galectine-3 qui est un marqueur de microglie et de macrophage. Le marquage montre une correspondance avec la zone moléculaire jaune qui est la région la plus représentée dans le groupe B (Figure 1-13).

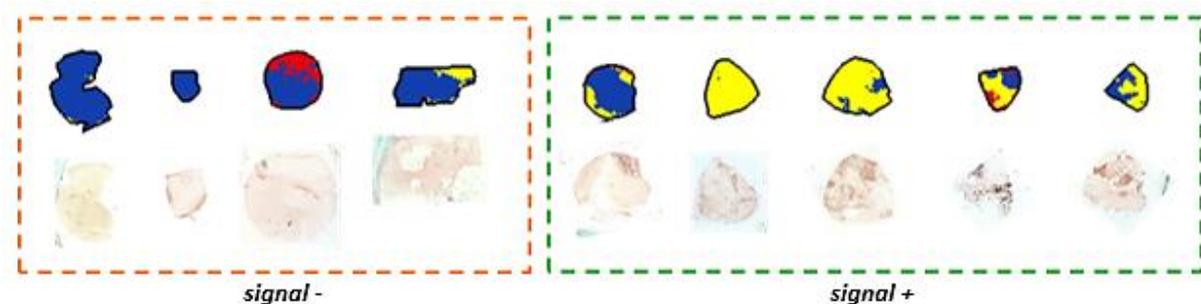


Figure 1-13 Représentation de l'infiltration microgliale au sein du groupe B
Immunohistochimie avec l'anticorps anti-Galectine 3. Le marquage anti-Galectine 3 correspond à la zone moléculaire jaune qui est associé à l'activation du système immunitaire.

En parallèle de l'activation du système immunitaire, de nombreuses protéines intervenant dans le processus de coagulation sont aussi surexprimées. Par exemple les chaînes de fibrinogène (FGA, FGB, FGG), le kininogène-1 (KNG1) et la prothrombine participent à l'initiation de la coagulation. On retrouve également le plasminogène (PLG), facteur clé de la fibrinolyse, c'est-à-dire de la dégradation du caillot et l'antithrombine 3 qui inhibe la coagulation. Quelques protéines sont aussi impliquées dans le transport du fer comme la ceruloplasmine (CP), la serotransferrin (TF) et hemopexin (HPX). Les mécanismes de coagulation sont connus pour être favorable aux cellules cancéreuses. La formation de fibrine facilite la migration de ces cellules et donc l'invasion. La coagulation est également un problème majeur chez les patients atteints de glioblastomes qui sont des patients à haut risque de thrombose (Taillibert et al., 2015).

L'annotation anatomopathologique révèle que la quasi-totalité (78 %) des coupes présente des zones de prolifération endothélio-capillaire intenses ou des zones d'inflammation avec parfois présence d'hémorragie. Ces annotations sont en accord avec les

résultats protéomiques. La prolifération endothélio-capillaire et les zones d'hémorragie indiquent que ces tumeurs sont très vascularisées et il est donc logique d'identifier des protéines de la coagulation dans ce groupe. De plus, 11 % des patients du groupe B présentent des événements thromboemboliques (Tableau 1-2). Aussi, les zones d'inflammation correspondent à la réaction du système immunitaire face à la tumeur.

L'analyse des voies de signalisation via Elsevier Pathways confirme que la plupart des protéines surexprimées dans ce groupe sont impliquées dans la cascade de coagulation, l'inflammation, l'ischémie, la vascularisation et le cancer. Certaines de ces protéines ont déjà été détectées dans la base de données des gliomes du *Cancer Genome Atlas* (TCGA) et sont pour la plupart des indicateurs de pronostic défavorables pour la survie des patients, par exemple grancalcin et CAPG.

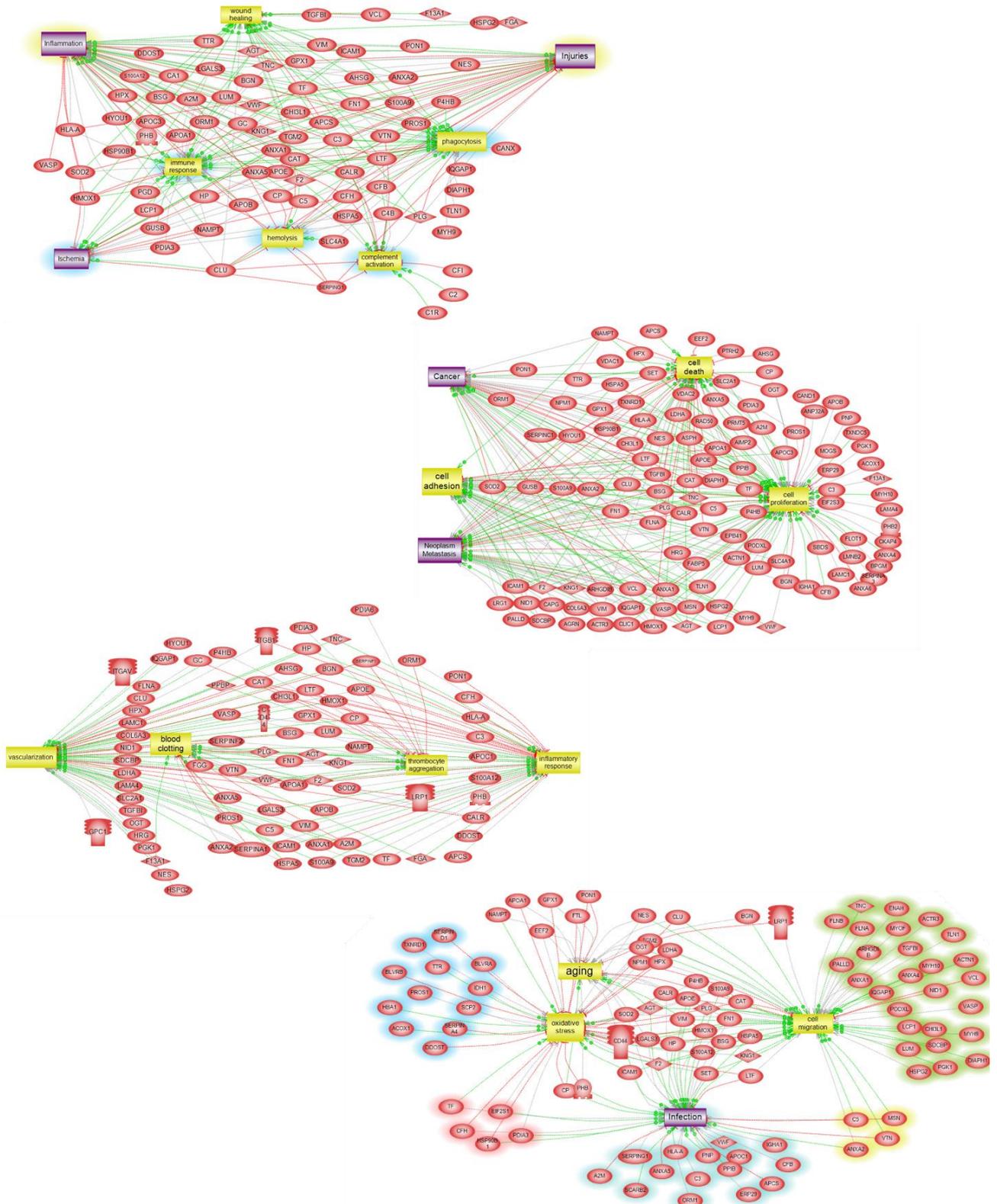


Figure 1-14 Voies de signalisations identifiées dans le groupe B

L'analyse des voies de signalisations indique que les protéines surexprimées dans ce groupe sont impliquées dans a) l'inflammation, l'ischémie b) le cancer, la prolifération cellulaire c) la réponse inflammatoire et la vascularisation. Une grande partie de ces protéines est également impliquée dans le stress oxydatif et l'infection.

Les patients du groupe B expriment des protéines liées au statut immunitaire avec une infiltration de macrophages indiqué par CAPG, la présence de chaînes lourdes et légères des immunoglobulines (IGHG2; IGKC; IGHM; IGHG1; IGLC6 et IGHA1) et des protéines de la cascade de coagulation comme HP, KNG1, HRG, TTR, GCA, S100A9 et STOM. Ces résultats sont en accord avec une étude de (Cheng et al., 2016), dans laquelle huit gènes liés au système immunitaire (FOXO3, IL6, IL10, ZBTB16, CCL18, AIMP1, FCGR2B et MMP9) ont été identifiés et utilisés comme marqueurs pronostiques défavorables dans les glioblastomes. Parmi les 8 gènes signatures, AIMP a également été identifié dans le groupe B. De plus les marqueurs CD44, nestin (NES) et vimentin (VIM), surexprimées dans le groupe B, sont des marqueurs de l'état de type mésenchymateux.

Le groupe C constitue la moitié de la cohorte et contient des protéines majoritairement impliquées dans la croissance tumorale. Par exemple, on retrouve les protéines hepatoma-derived growth factor (HDGF) et developmentally-regulated GTP-binding protein 2 (DRG2). Il est intéressant de noter la présence de protéines impliquées dans les processus viraux. En effet, Zinc finger CCCH-type antiviral protein 1 (ZC3HAV1) est une protéine antivirale qui inhibe la réplication de virus en dégradant les ARNm viraux et E3 ubiquitin/ISG15 ligase TRIM25 (TRIM25) est impliquée dans l'immunité contre les virus. La protéine EIF2AK2 joue également un rôle dans la réponse immunitaire contre les infections virales et est notamment impliquée dans la régulation de la prolifération et la différenciation cellulaire.

L'analyse des voies de signalisation indique que ce groupe est lié au cancer, à la croissance cellulaire et à la réparation de l'ADN. Ces analyses confirment également l'implication de ces protéines dans la traduction et l'activité ribosomale.

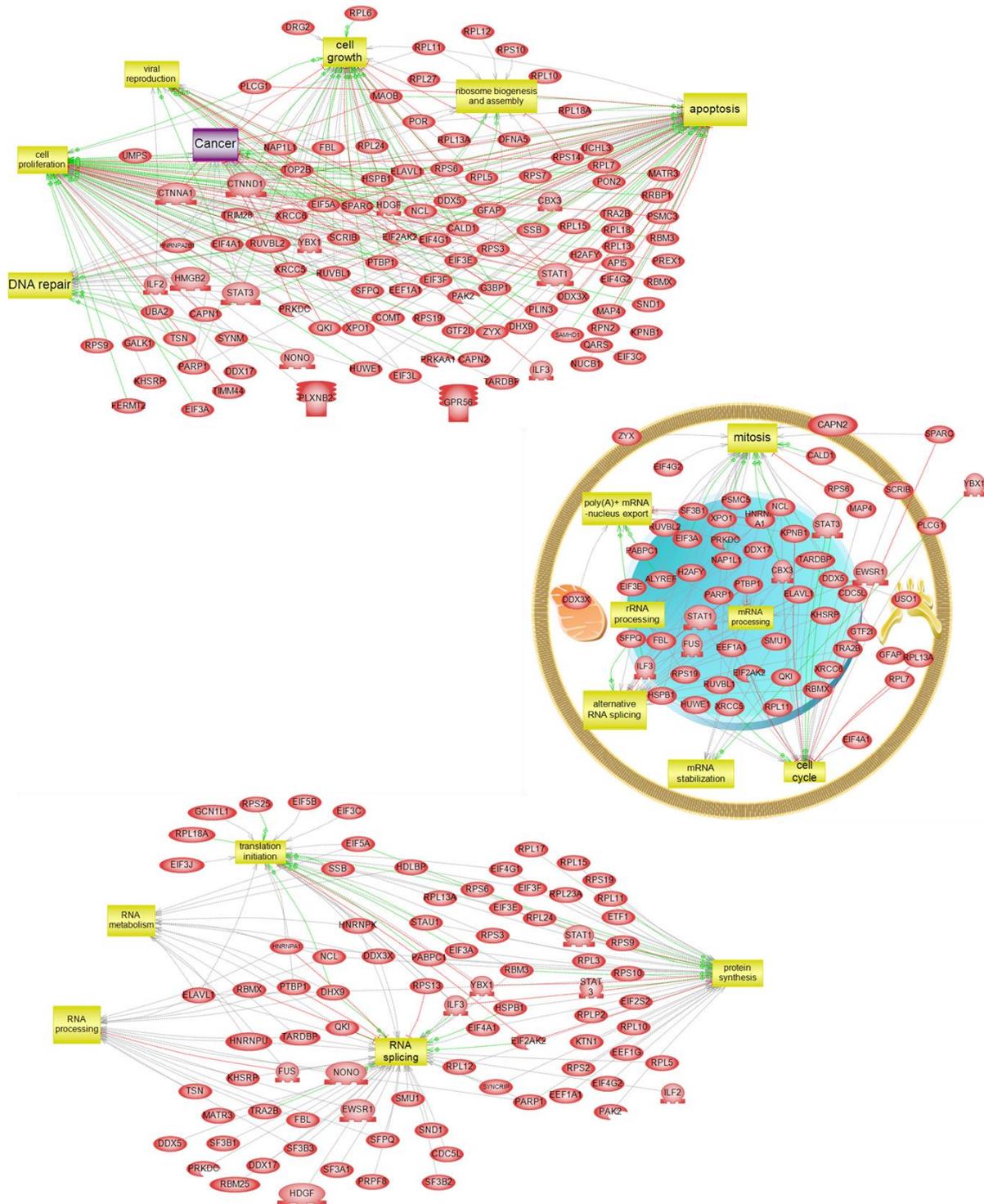


Figure 1-15 Voies de signalisation identifiées dans le groupe C

L'analyse des voies de signalisation montre l'implication des protéines surexprimées dans ce groupe dans a) le cancer, la prolifération cellulaire et la reproduction virale. Ces protéines sont également liées à traduction et la synthèse protéique.

Le groupe C est lié à une réponse immunitaire antivirale et à une infection virale. Des études récentes ont reporté un lien entre les glioblastomes et l'exposition virale périnatale (Akhtar et al., 2018; Dickinson et al., 2002; Limam et al., 2019; Strojnik et al., 2017). Il a récemment été évoqué que le virus Epstein-Barr pourrait être impliqué dans l'étiologie des glioblastomes (Zavala-Vega et al., 2019). De plus, certains travaux récents ont également signalé que le cytomégalovirus (CMV) favorise la croissance des glioblastomes chez la souris via le recrutement des péricytes et l'angiogenèse (Krenzlin et al., 2019). Chez l'homme, des acides nucléiques et des protéines CMV ont été observés dans le tissu tumoral GBM (Rahman et al., 2019).

Cette analyse protéomique permet de mettre en évidence les différents processus biologiques de chaque groupe. Chaque groupe de patients est très différent, le premier correspond à un profil neurologique, le deuxième un profil immunologique tandis que le dernier possède une implication virale.

Ces stratégies reposent sur l'interrogation de la base de données de protéines, par conséquent, seules les protéines référencées peuvent être identifiées. Les protéines alternatives sont des protéines issues de séquences codantes alternatives. Elles sont différentes des protéines de références et contribuent à la diversité protéique. Il est nécessaire d'étudier ces protéines et leurs implications dans les différents processus biologiques pour une meilleure compréhension des mécanismes pathologiques (Delcourt et al., 2018b).

F. Identification de protéines alternatives

Les protéines alternatives sont des protéines qui ne sont pas répertoriées dans les banques de données protéiques, en effet, ces protéines sont issues de cadres de lectures alternatifs (Vanderperre et al., 2013). Traditionnellement, on considère que les protéines proviennent des cadres de lecture les plus longs. Ainsi, les protéines alternatives sont issues de petits cadres de lecture, de décalage de cadre lecture, de chevauchement avec une région non codante, de cadre de lecture situé sur les régions 3' et 5' non codantes et également d'ARNs annotés comme non codants (ARNnc) (Vanderperre et al., 2013). Les protéines alternatives constituent un nouveau protéome à explorer.

En utilisant la base de données de protéines alternatives OpenProt (Brunet et al., 2018), nous avons pu identifier 257 protéines alternatives (AltProt) et 170 de ces protéines ont pu être quantifié.

L'ensemble des échantillons a été regroupé selon la couleur et le groupe de chaque tumeur, comme précédemment. Ainsi, l'analyse statistique (ANOVA p-value 0.05) a permis de mettre en évidence 41 AltProt différentiellement exprimées entre chaque groupe.

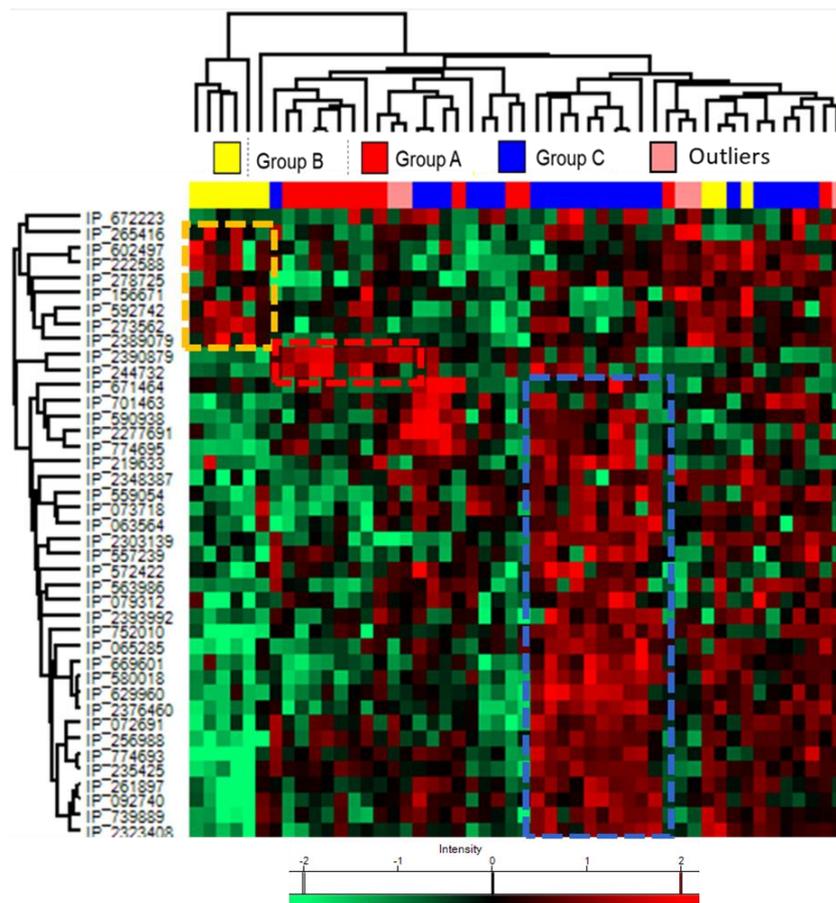


Figure 1-16 L'analyse des protéines alternatives met en lumière des clusters de protéines surexprimées dans les trois groupes déterminés par imagerie.

Analyse statistique (ANOVA, p value 0.05) des protéines alternatives identifiées sur les points de micro-extractions correspondant à la couleur représentative du groupe de patient (Rouge pour le groupe A ; jaune pour le groupe B et bleu pour le groupe C). Dans chaque groupe, un cluster de protéine est surexprimé.

On observe deux AltProt surexprimées dans le groupe A : IP_2390879 issu de l'ARN non codant (ncARN) LOC107985743 et IP_244732 provenant de la partie 5'UTR de l'ARNm

codant pour KIFC3 qui est impliqué dans l'adhésion cellulaire. Dans le groupe B, on retrouve neuf AltProts surexprimées : quatre sont transcrites à partir de d'ARN non codant, deux sont situés sur la région 5'UTR d'un ARNm, un au niveau 3'UTR et deux résultent du décalage de cadre de lecture dans la séquence codante. Dans le groupe C, 30 AltProts sont surexprimées : 16 issues d'ARN non codants, 2 provenant de la région 5'UTR, 8 de la région 3'UTR et 4 résultent du décalage de trame dans la séquence codante.

Certaines AltProts ont également déjà été identifiées dans une étude précédente comme la protéine alternative IP_2323408, provenant du gène LOC105376126 non caractérisé et actuellement décrite comme ARN non codant (Cardon et al., 2020a). De même, les AltProts IP_774693, IP_572422 et IP_671464 ont déjà été identifiées et correspondent à des pseudogènes : TUBAP2, TUBBP1 et TPI1P1 respectivement. Ces pseudogènes, pour lesquels aucune protéine n'a encore été observée, ont une expression de leurs transcrits dans des lignées cellulaires de gliome (Atlas d'expression) (Petryszak et al., 2016). D'autre part, il est intéressant de noter que IP_079312, provenant de l'ARNm codant EDARADD a été corrélée à un faible taux de survie chez les patientes atteintes d'un cancer de l'ovaire (Cardon et al., 2020b). De plus, une étude sur les partenaires d'interaction d'AltProts dans des cellules NCH82 a mis en évidence 5 protéines alternatives différentiellement exprimées après transition épithélio-mésenchymateuse (Cardon et al., 2020a). L'une de ces protéines a été identifiée comme surexprimée dans le groupe B : IP_156671 qui provient de la région 3'UTR du transcrit codant pour solute carrier family 13 member 1 (SLC13A1). Les 4 autres sont surexprimés dans le groupe C: IP_261897 provenant d'un ARN non codant, IP_063564 et IP_256988 issus respectivement des régions 3'UTR des gènes CLDN19 et TBX21 et IP_073718 provenant d'un décalage dans le cadre de lecture du gène CCDC181.

Plus de la moitié des AltProts identifiées ici proviennent de la traduction d'ARN non codants transcrites à partir de pseudogènes. Récemment, il a été démontré que les pseudogènes peuvent être utilisés comme signatures pour le pronostic des gliomes. Ainsi, 6 pseudogènes (SP3P, ANXA2P3, PTTG3P, LPAL2, CLCA3P et TDH) se sont révélés être associés à la survie des patients atteints de gliome (Gao et al., 2015). Un ensemble de 9 autres pseudogènes (TP73-AS1, AC078883.3, RP11-944L7.4, HAR1B, RP4-635E18.7, HOTAIR, SAPCD1-AS1, AC104653.1 et RP5-1172N10.2.) ont été identifiés comme étant des marqueurs

de pronostic de la survie des patients atteints de gliome (Lei et al., 2018). Ces résultats fournissent de nouvelles perspectives sur le rôle biologique des pseudogènes dans le cancer et en particulier dans le gliome. De plus, les protéines alternatives identifiées et traduites à partir de ces ARN non codants fournissent des informations supplémentaires quant à la physiopathologie des gliomes.

Il est intéressant de noter qu'il existe parmi les protéines alternatives, des protéines spécifiques à chacun des trois sous-groupes de glioblastome. Il sera nécessaire d'étudier plus profondément la fonction de ces protéines, en effet, elles peuvent être impliquées dans des voies de signalisations spécifiques et interagir avec d'autres protéines. Par la suite, il paraît envisageable de les utiliser pour améliorer le diagnostic et de les considérer comme de potentielles cibles thérapeutiques.

G. Corrélation entre TCGA et données protéomiques

Nous avons ensuite comparé la totalité des protéines identifiées (presque 5000) à la base de données TCGA. Le *Cancer Genome Atlas* (TCGA) est un projet de séquençage génomique du cancer qui a permis la caractérisation de 33 types de cancers différents. Dans les gliomes, 682 gènes ont été identifiés comme ayant une expression élevée. Au sein de nos échantillons, nous avons identifié les protéines correspondant à 282 de ces gènes.

Parmi les 682 gènes montrant une expression élevée dans les gliomes, 268 sont suggérés comme pronostique sur la base des données transcriptomiques de 153 patients. Ainsi, 201 gènes sont associés à un pronostic défavorable et 67 gènes à un pronostic favorable. À partir des données protéomiques de chaque groupe, nous avons identifié 10 protéines associées à un pronostic défavorable (Tableau 1-4). L'une de ces protéines est surexprimée dans le groupe C : NADPH-cytochrome P450 reductase (POR). Deux protéines sont surexprimées dans le groupe B : Adipocyte enhancer-binding protein 1 (AEBP1) et Protein disulfide-isomerase A4 (PDIA4) et sont impliquées dans l'organisation et le remodelage de la matrice extracellulaire. Finalement, 7 protéines sont surexprimées dans le groupe A : cell cycle exit and neuronal differentiation protein 1 (CEND1) ; dematin (DMTN) ; serine/threonine-protein kinase PAK 1 (PAK1) ; dual specificity mitogen-activated protein kinase kinase 1 (MAP2K1) ; Thy-1 membrane glycoprotein (THY1) ; Visinin-like protein 1 (VSNL1) et NADP-

dependent malic enzyme (ME1). Ces protéines sont majoritairement impliquées dans la différenciation et la projection neuronale.

Nous avons également trouvé 6 protéines associées à un pronostic favorable (Tableau 1-4). Quatre protéines sont notamment impliquées dans le développement du système nerveux et surexprimées dans le groupe A : glutamate dehydrogenase 1 (GLUD1) ; Rab GDP dissociation inhibitor alpha (GDI2) ; serine-tRNA ligase, cytoplasmic (SARS) et Rho guanine nucleotide exchange factor 7 (ARHGEF7). Une protéine est surexprimée dans le groupe B Importin subunit alpha-4 (KPNA3) et une protéine est surexprimée dans le groupe C Polyadenylate-binding protein 1 (PABPC1). Ces protéines sont notamment impliquées dans l'infection virale.

Tableau 1-4 Protéines pronostics identifiées dans les groupes A, B et C

	Protéine	Gene	Groupe
<i>Pronostic défavorable</i>	Cell cycle exit and neuronal differentiation 1	CEND1	A
	Dematin actin binding protein	DMTN	A
	P21 (RAC1) activated kinase 1	PAK1	A
	Mitogen-activated protein kinase kinase 1	MAP2K1	A
	Thy-1 cell surface antigen	THY1	A
	Visinin like 1	VSNL1	A
	Malic enzyme 1	ME1	A
	AE binding protein 1	AEBP1	B
	Protein disulfide isomerase family A member 4	PDIA4	B
	Cytochrome p450 oxidoreductase	POR	C
<i>Pronostic favorable</i>	Glutamate dehydrogenase 1	GLUD1	A
	Rho guanine nucleotide exchange factor 7	ARHGEF7	A
	GDP dissociation inhibitor 2	GDI2	A
	Seryl-tRNA synthetase	SARS	A
	Karyopherin subunit alpha 3	KPNA3	B
	Poly(A) binding protein cytoplasmic 1	PABPC1	C

Des protéines biomarqueurs pronostiques ont donc été retrouvées dans chacun des 3 groupes. La plupart des protéines ont été identifiées dans le groupe A et sont liées au développement du système nerveux et à la différenciation des neurones.

H. Corrélation avec les données cliniques

Dans l'ensemble de la cohorte, le statut MGMT et l'indice de Karnofsky sont corrélés avec la survie globale, mais pas la qualité de la résection (Figure 1-5). Finalement, nous avons comparé le groupement aux données de survie des patients (Figure 1-17). Il n'existe pas de différence significative en survie entre les trois groupes malgré les différents processus physiopathologiques identifiés.

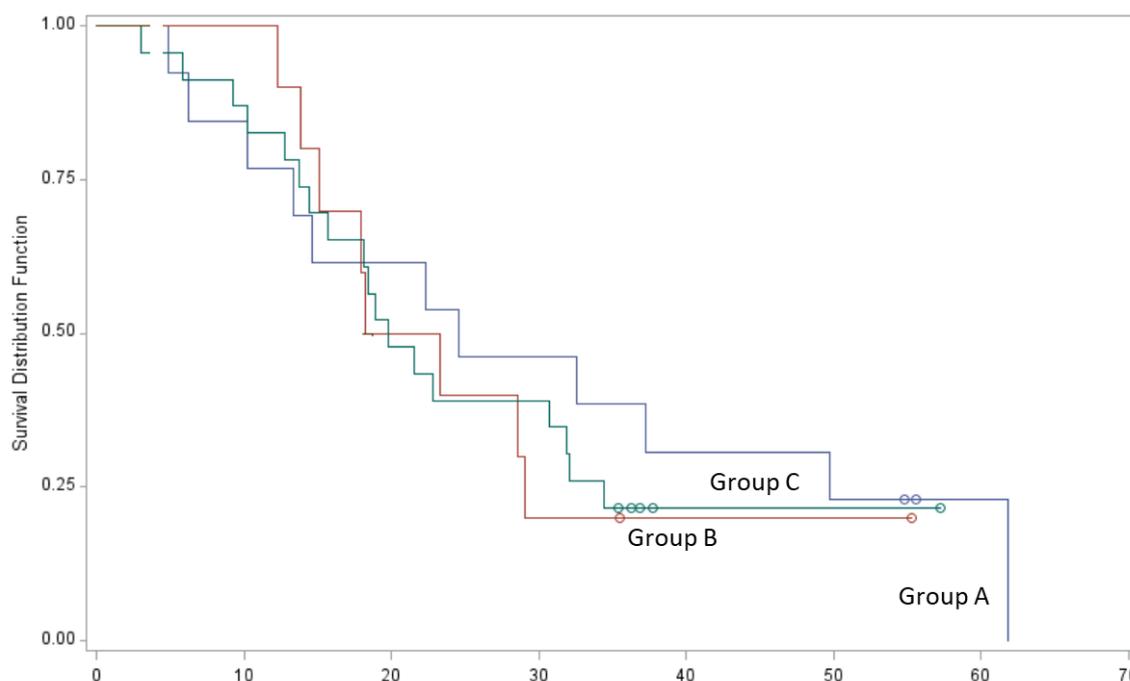


Figure 1-17 Courbes de survie en fonction des groupes A, B et C

L'analyse de la survie globale des patients ne révèle aucune différence significative entre les 3 groupes

Dans l'optique de trouver de nouvelles protéines pronostiques, nous avons effectué un test ANOVA sur l'ensemble des données protéomiques de la totalité des patients en fonction de leurs données de survie. Ainsi, la cohorte de 50 patients a été divisée en 3 groupes : 12 patients (24 %) ayant une survie inférieure au troisième quartile, 26 patients (52 %) avec une survie entre le premier et le troisième quartile et 12 patients (24 %) avec une survie supérieure au premier quartile. A partir de ces données, l'analyse statistique (ANOVA p-value 0.05) a permis d'identifier 126 protéines significatives entre ces 3 groupes de patients définis par leur survie. Ensuite, en utilisant le modèle de Cox, 30 protéines ont été identifiées comme significatives à $p = 0,05$ et 9 à $p = 0,01$. Après une analyse étape par étape et une procédure

de bootstrap, 6 protéines sont restées significativement corrélées à la survie : MAOB, IGHM, HSPD1, LASP1, CFH et ANXA6 (Tableau 1-5).

Tableau 1-5 Protéines significatives de la survie des patients utilisées pour le regroupement des patients

Parameter	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95 % Hazard Ratio Confidence Limits
MAOB	0,626	0,181	12,002	0,001	1,870	1,312 2,665
IGHM	0,321	0,136	5,533	0,019	1,378	1,055 1,800
HSPD1	-1,204	0,568	4,493	0,034	0,300	0,099 0,913
LASP1	-0,582	0,282	4,257	0,039	0,559	0,321 0,971
CFH	-0,257	0,123	4,367	0,037	0,774	0,608 0,984
ANXA6	0,458	0,191	5,735	0,017	1,580	1,087 2,298

Les trois protéines IGHM, ANXA6 et MAOB sont de mauvais pronostic. En effet, le « *hazard ratio* » de ces 3 protéines est supérieur à 1 et plus la valeur est élevée, plus la survie est faible (Tableau 1-5). Les trois autres protéines, LASP1, CFH et HSPD1 sont de bon pronostic, leurs « *hazard ratio* » étant inférieur à 1 (Tableau 1-5). Sur la base de l'expression de ces 6 protéines, 3 groupes de patients ont été identifiés (respectivement les groupes 1, 2 et 3) (Figure 1-18).

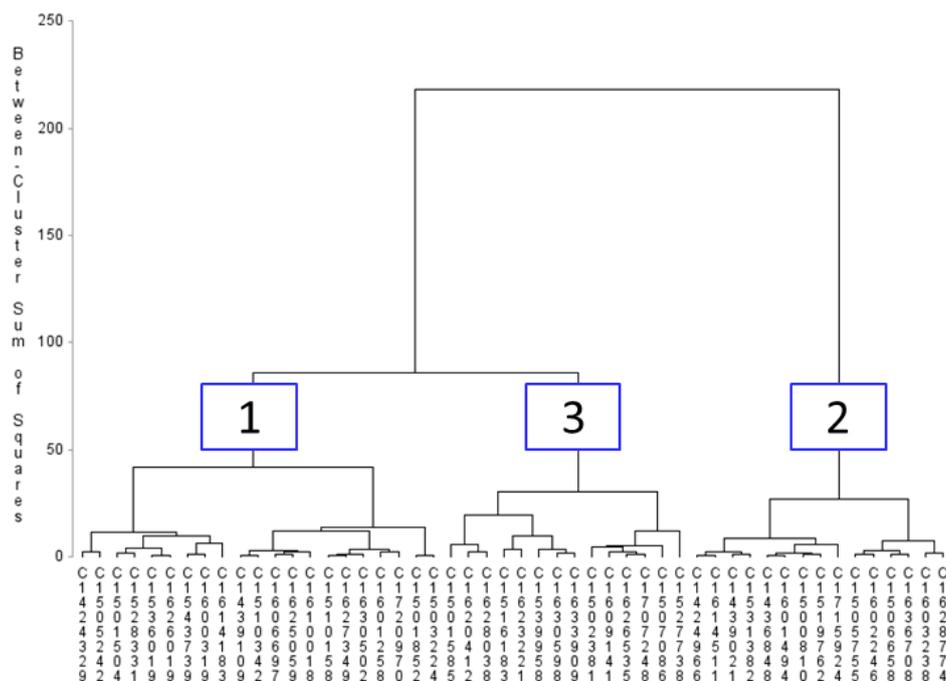


Figure 1-18 Regroupement basé sur les 6 protéines les plus significatives de la survie
Regroupement basé sur les valeurs de quantification des protéines MAOB, IGHM, HSPD1, LASP1, CFH et ANXA6.

Les courbe de survie des patients de chacun des 3 groupes ont été déterminées (Figure 1-19). Les patients du groupe 2 (n = 15) ont la survie la plus faible, tandis que les patients du groupe 3 (n = 14) ont la plus élevée. Le groupe 1 (n = 21) est intermédiaire.

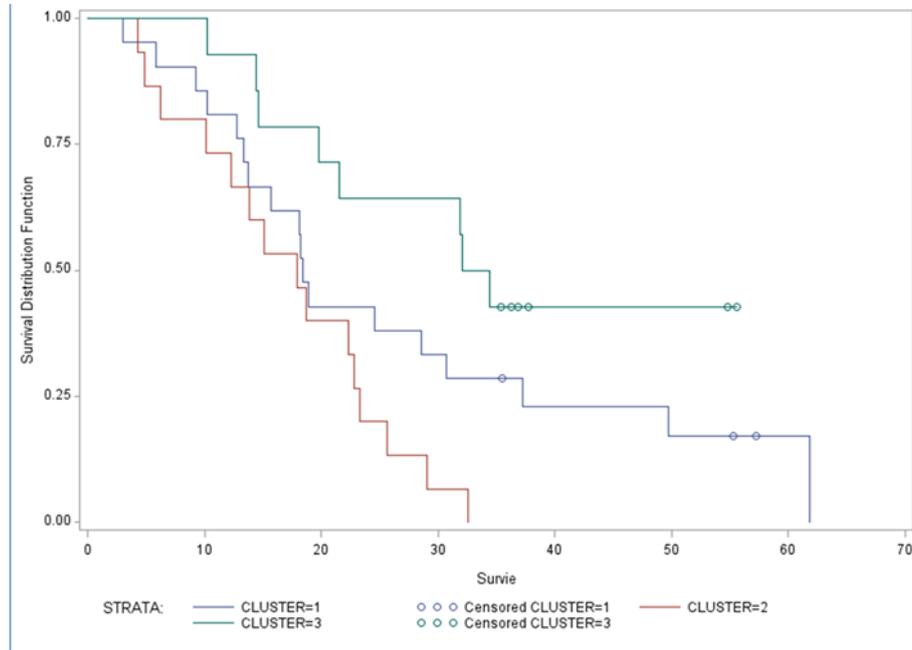


Figure 1-19 Courbe de survie des 3 groupes identifiées sur la base des 6 protéines

La survie globale de chaque groupe identifié révèle que le groupe 2 observe le meilleur pronostic tandis que le groupe 3 a la survie la plus faible

Par la suite, nous avons voulu comparer ces 3 groupes aux groupes définis par imagerie. En ce sens, nous avons comparé l'expression des 30 marqueurs pronostiques identifiés entre les 3 groupes de survie (1 ; 2 et 3) et les 3 groupes définis précédemment (A ; B et C). Ainsi, 17 de ces protéines sont surexprimées à la fois dans le groupe B (activation du système immunitaire) et dans le groupe 2 (survie la plus faible). Parmi ces protéines on retrouve 3 des protéines utilisées pour le regroupement des patients : IGHM, ANXA6 et MAOB. Ces 3 protéines sont des indicateurs de mauvais pronostic (*Hazard Ratio* > 1). Ces protéines sont principalement liées aux macrophages et à l'activation des microglies.

Un cluster de protéines est surexprimé dans le groupe 3 (survie la plus favorable) et très représenté dans le groupe A. Parmi ces protéines, on retrouve LASP1 et HSPD1 qui sont des indicateurs de bon pronostic (*Hazard Ratio* < 1).

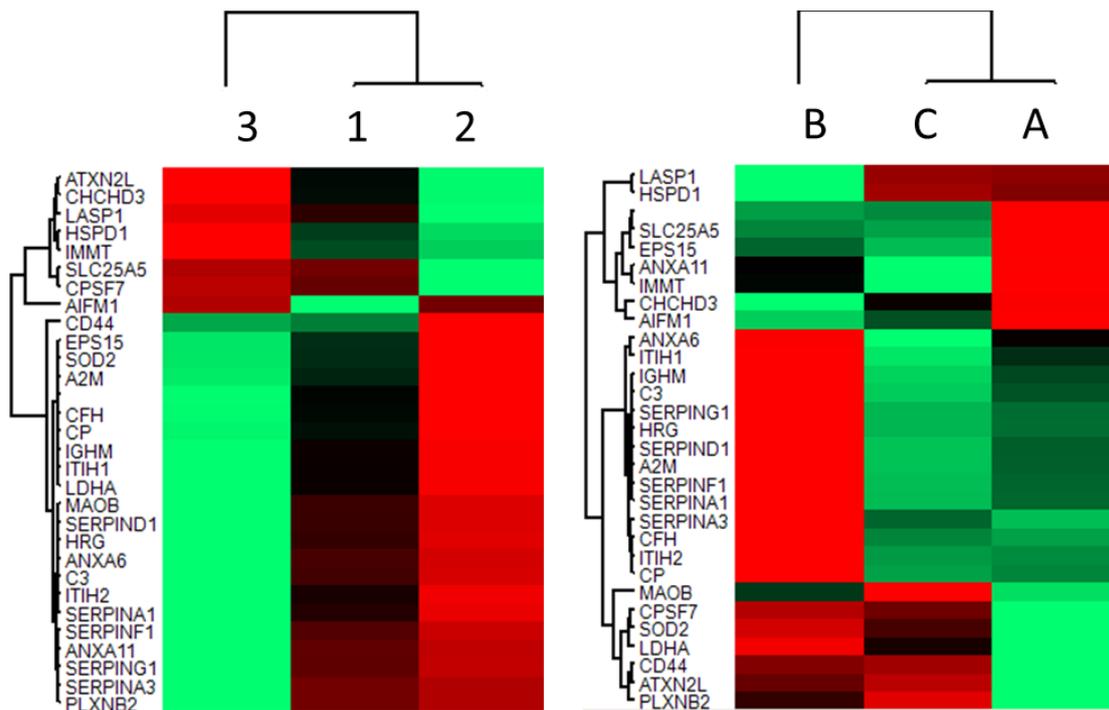


Figure 1-20 Expression des protéines significatives de la survie dans les différents groupes de patients

Heatmap des 30 protéines significatives après le modèle de Cox ($p = 0,05$) entre les 3 groupes de patients définis par leur survie 1, 2 et 3 (gauche) et expression de ces 30 protéines dans chaque groupe d'imagerie groupe A B et C (droite)

La protéine CFH est plus exprimée dans le groupe 2, bien qu'elle soit un marqueur de bon pronostic. Cette protéine a une valeur de risque de 0,77 ; c'est le marqueur le plus faible par rapport aux deux autres protéines de bon pronostic (HSPD1 et LASP1).

Ces 6 marqueurs spécifiques (MAOB, IGHM, HSPD1, LASP1, CFH et ANXA6) ont été validés par immunohistofluorescence (Figure 1-21). Nous avons confirmé chez les patients du groupe 2 (survie la plus faible), une expression plus élevée d'ANXA6, IGHM et MAOB, qui sont en bonne corrélation avec les données protéomiques. Au contraire, HSPD1 et LASP1 apparaissent plus exprimés dans les tumeurs du cluster 3 (survie la plus longue), validant à nouveau nos résultats précédents. La protéine CFH est légèrement plus exprimée dans les tumeurs du groupe 3, ce qui est attendu car c'est un marqueur de bon pronostic. Enfin, les patients appartenant au cluster OS intermédiaire (groupe 1) présentent une signature mixte avec une expression élevée à la fois des marqueurs de bon pronostic (HSPD1 et LASP1) et de mauvais pronostic (ANXA6 et IGHM).

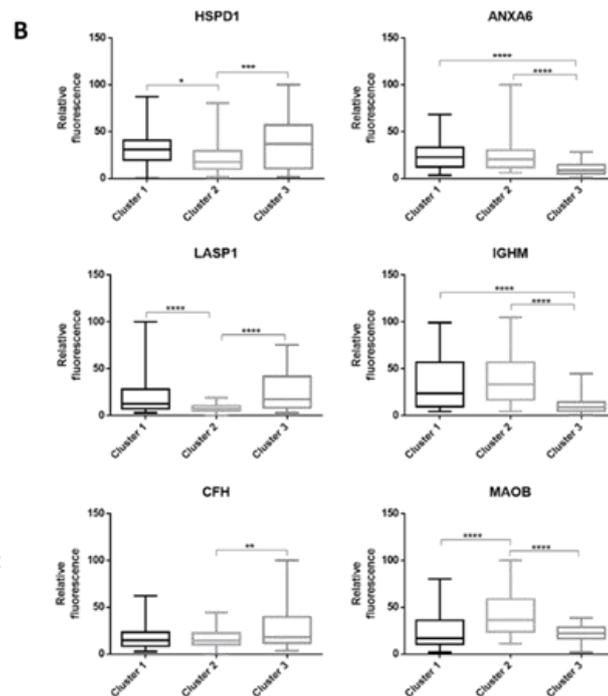
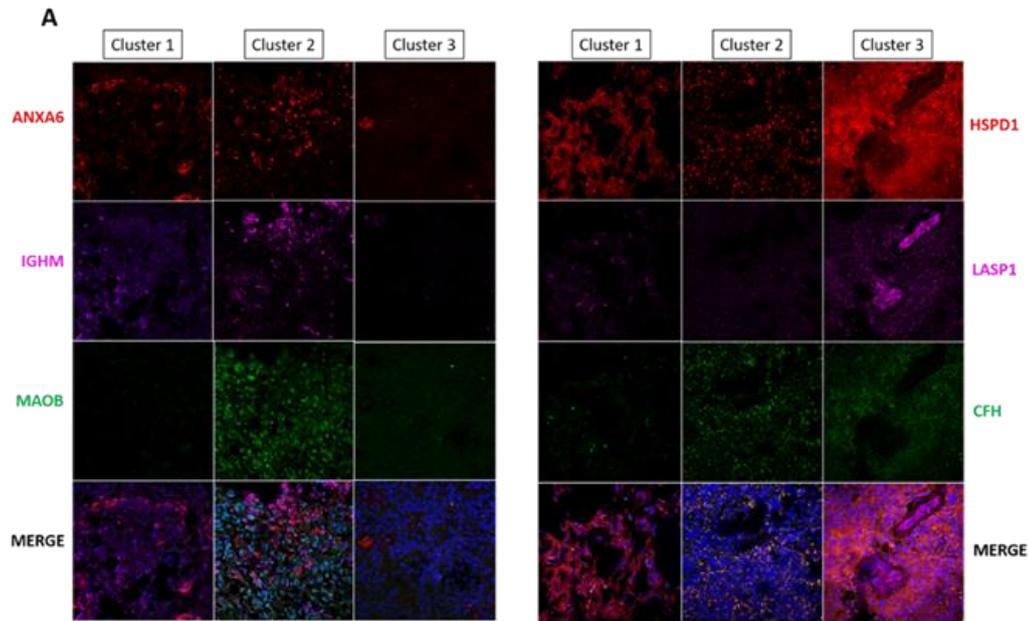


Figure 1-21 Immunohistofluorescence des 6 marqueurs de survie identifiés

A) Images de fluorescences représentatives des 6 protéines dans les trois groupes de survie. MAOB, IGHM et ANXA6 sont associés à un mauvais pronostic tandis que CFH, LASP1 et HSPD1 sont liés à un bon pronostic. Les images ont été acquises avec un microscope confocal à un grossissement de 40x. B) Quantification de l'intensité de fluorescence des 6 protéines dans les 3 groupes de survie. Les images prises à partir de 15 tumeurs du groupe 1, 14 tumeurs du groupe 2 et 10 tumeurs du groupe 3 ont été quantifiées. Pour chaque tumeur, 3 à 4 images ont été acquises et quantifiées. Des différences significatives ont été identifiées en utilisant le test t non apparié avec **** $p < 0,0001$; *** $p < 0,001$; ** $p < 0,01$ et * $p < 0,05$.

L'analyse des protéines alternatives de chaque groupe de survie, avec la base de données alternative OpenProt, a par la suite permis la détection de 29 AltProts significatives (Figure 1-22). Parmi ces 29 AltProts, 22 sont surexprimées dans le groupe de patients ayant la survie la plus faible (groupe 2). Les 7 autres, malgré une variation importante, ne permettent pas de différencier les groupes de survie 1 et 3. Sur les 29 AltProts, 15 sont issues d'ARN non codant (51 %), soulignant une fois de plus l'importance des ARN non codants dans le développement du cancer.

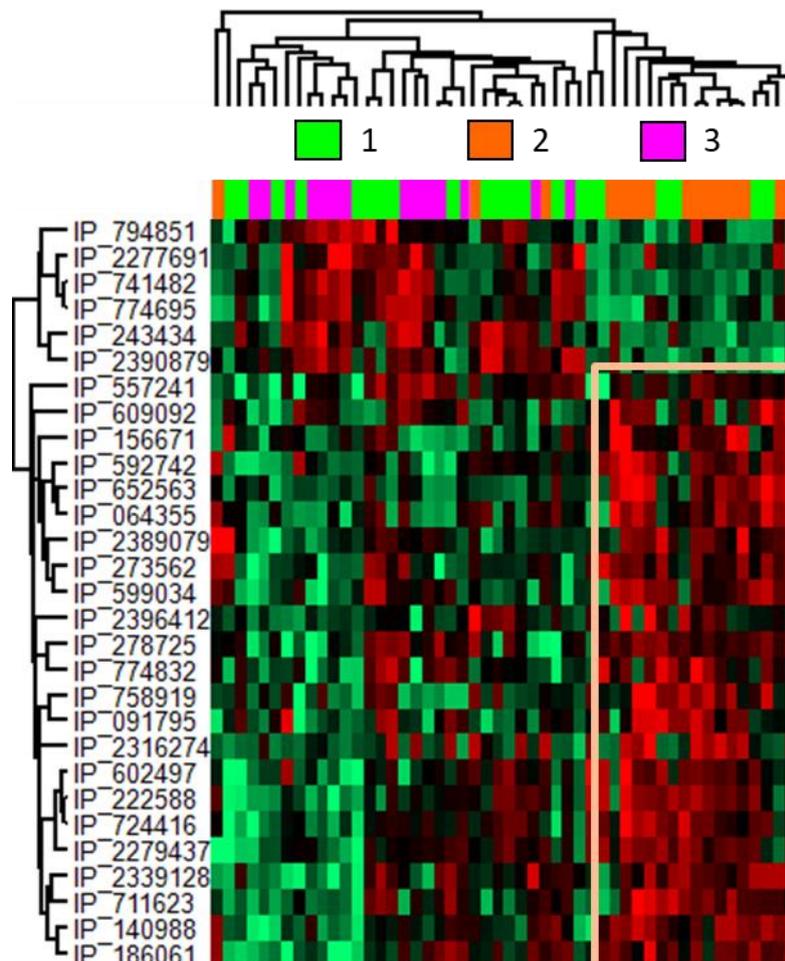
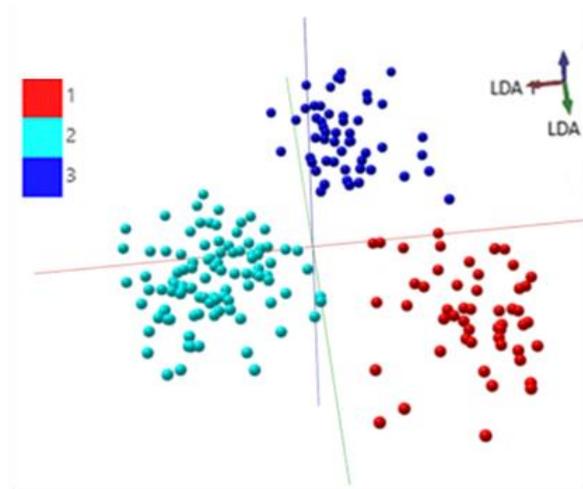


Figure 1-22 Heatmap des protéines alternatives montrant différents profils de régulation dans les trois groupes de survie

L'encadré beige montre un cluster de protéines surexprimées dans le groupe 2, associé à une mauvaise survie.

Finalement une seconde classification a été réalisée grâce au SpiderMass en prenant en compte les 3 groupes de survie et montre un excellent taux de séparation (Figure 1-23).

L'analyse LDA montre une bonne séparation des 3 groupes. La validation par retrait aléatoire de 20 % des patients a donné 92 % de bonne classification en excluant les valeurs aberrantes et 90,64 % en les incluant, confirmant la spécificité de la classification basée sur la survie (Figure 1-23).



Group	Number of spectra	Number of passes	Number of failures	Number of outliers	Correct Classification Rate	
					Excluding outliers	Including outliers
Total	203	184	16	3	92.00%	90.64%

Figure 1-23 Modèle de classification PCA-LDA basé sur 3 groupes de survie

Le tableau représente les taux de validation croisée «20 % out» du modèle construit. Représentation LDA des groupes de survie ; Groupe 1 (rouge), Groupe 2 (cyan) et Groupe 3 (bleu). La LDA 1 sépare le cluster 1 et le cluster 2, tandis que la LDA sépare le cluster 1 et le cluster 3.

Considérant l'ensemble de nos résultats, nous avons établi une classification des glioblastomes basée sur les caractéristiques moléculaires et qui peut être très utile pour le développement de nouvelles thérapies et de la médecine personnalisée pour les patients.

IV. Conclusion et Perspectives

Les glioblastomes représentent une pathologie complexe et de nouveaux critères moléculaires sont nécessaires pour améliorer leur classification et leur prise en charge. Ainsi, ce projet porte sur l'élaboration d'une nouvelle classification des patients atteints de glioblastomes basée sur des techniques de protéomique. L'utilisation de l'imagerie par spectrométrie de masse MALDI est déterminante dans l'étude de ces tumeurs puisqu'elle permet de rendre compte de l'hétérogénéité tumorale. Cette technique d'analyse non supervisée a ici été couplée à la protéomique de tissu dans le but de concevoir une classification de ces tumeurs. L'analyse par imagerie MALDI et segmentation spatiale fait émerger une nouvelle sous classification de cette pathologie. Nous avons validé ces observations grâce au SpiderMass avec une bonne classification de 93 %.

Ainsi, 3 sous-groupes ont été identifiés (A-rouge, B-jaune et C-bleu) et, dans l'optique d'investiguer les voies biologiques impliquées dans ces trois groupes, la protéomique localisée sur tissus a confirmé le groupement. Les signatures moléculaires des différents sous-groupes de tumeurs ont été identifiées. Ces données confirment qu'il existe une hétérogénéité moléculaire inter-patient qui peut être liée au phénotype tumoral et à la plasticité cellulaire (Neftel et al., 2019).

Cette étude a permis de mettre en évidence la présence de 6 nouvelles protéines pronostiques pour les glioblastomes : LASP1, HSPD1 et CFH sont des marqueurs pronostiques favorables tandis que IGHM, ANXA6 et MAOB sont des marqueurs pronostiques défavorables. L'expression de ces marqueurs protéiques défavorables est plus élevée dans les tumeurs appartenant au groupe 2 (survie la plus faible) et au groupe B. De plus, 2 protéines pronostiques défavorables connues, sont également surexprimées dans le groupe B (AEBP1 et PDIA4). Les 2 autres groupes sont associés à une meilleure survie globale, les marqueurs associés étant de bon pronostic.

La signature moléculaire des patients du groupe B montre que les protéines de ce groupe sont principalement impliquées dans l'inflammation, la phagocytose, le stress oxydatif et l'ischémie et sont liées à l'infiltration des macrophages. Les cellules tumorales pourraient influencer l'état de polarisation des macrophages associés aux tumeurs (Guadagno et al.,

2018). Les cellules immunitaires ont un rôle décisif dans la communication avec les cellules cancéreuses au sein du microenvironnement tumoral. Compte tenu cet aspect immunitaire, il paraît adapté de proposer les stratégies d'immunothérapie spécifiquement aux patients du groupe B. Dans ce contexte, notre étude pourrait être un point de départ intéressant pour guider le développement de nouvelles stratégies thérapeutiques personnalisées et de meilleures décisions de traitement. En ce sens, une stratégie développée au laboratoire se base sur la réactivation des macrophages pour empêcher la progression tumorale et éradiquer la tumeur grâce à des inhibiteurs de proprotéines convertases (Duhamel et al., 2018; Rose et al., 2020).

Notre étude ayant démarré sur la base de la classification OMS de 2016, comprend à la fois des glioblastomes IDH non mutés et IDH mutés. Cependant, des études très récentes suggèrent que ces deux entités représentent des cas cliniques et génétiques distincts (Brat et al., 2020). Bien que la cohorte ne comprenne que 4 tumeurs IDH mutantes, leurs protéomes n'ont pas impacté la classification, le même regroupement est observé en l'absence de ces tumeurs (Figure 1-24). De plus, dans la cohorte, la survie des patients IDH mutés n'est pas supérieure à celle des patients IDH non mutés. Ainsi, au sein de notre cohorte aucune différence majeure n'est observée entre ces deux entités.

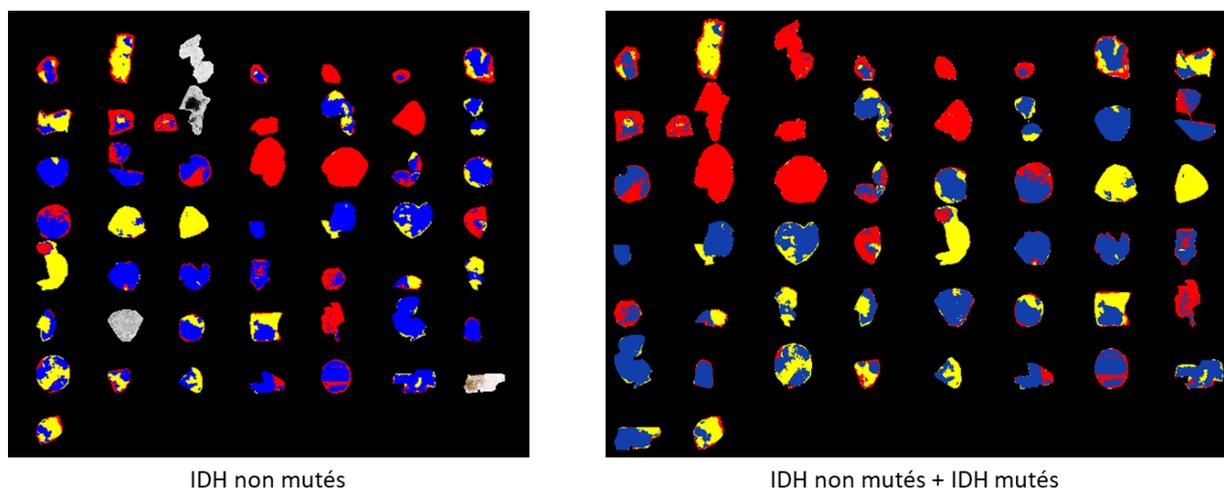


Figure 1-24 Comparaison des segmentations spatiales avec et sans les patients présentant une mutation IDH

La segmentation spatiale des patients de la cohorte et par conséquent, le regroupement des patients n'est pas affecté par la présence des tumeurs IDH mutés.

L'ensemble de ces résultats constitue de premières observations et des études futures, incluant notamment une plus grande cohorte de patient, seront nécessaires pour confirmer nos observations.

Pour conclure, nous avons établi que les glioblastomes peuvent être divisés en 3 groupes. Chaque groupe présente des signatures moléculaires spécifiques, reflétant un phénotype moléculaire spécifique des tumeurs. Le groupe A correspond à un profil plutôt neurologique avec notamment des protéines impliquées dans la neurogénèse. Le groupe B correspond à une activation du système immunitaire et une dérégulation de la cascade de la coagulation. Le dernier groupe, le groupe C, présente une implication virale forte. Chaque groupe met donc en évidence des voies de signalisation très spécifiques. Il est nécessaire de prendre en compte ces différences pour cibler les traitements. L'approche protéomique pourrait à terme permettre le développement de stratégies thérapeutiques personnalisées.

L'étude de l'hétérogénéité des glioblastomes est un enjeu majeur pour la compréhension des mécanismes physiopathologiques et pour la recherche thérapeutique. De ce fait, il apparaît primordial d'étudier les multiples types cellulaires composant la tumeur et leurs interactions avec le microenvironnement tumoral.

La protéomique localisée sur tissu offre l'occasion de comprendre les mécanismes intervenant dans la croissance tumorale tout en conservant leurs localisations sur les tissus. Un des objectifs de la protéomique localisée est l'analyse de régions de plus en plus petites tout en gardant l'information de localisation. Le développement de ces techniques ouvre finalement de nouvelles possibilités pour identifier des populations cellulaires distinctes.

Ainsi, la deuxième partie de ce projet de thèse se concentre sur le développement d'une nouvelle technique de protéomique localisée, qui à terme pourrait être utile pour l'étude des gliomes.

CHAPITRE 2

Développement d'une nouvelle approche de microprotéomique localisée basée sur l'expansion de tissu

I. Introduction

Les protéines sont affectées dans leurs structures et leurs régulations par les changements et les dysfonctionnements cellulaires. La protéomique permet l'analyse de ces fluctuations par observation directe des protéines. Dans un contexte pathologique, l'étude des protéines permet une meilleure compréhension des mécanismes intervenant dans la maladie. Cependant, l'analyse du tissu dans sa globalité empêche de relier la protéine à leurs emplacements. Ces types d'analyses ne tiennent pas compte de l'hétérogénéité tumorale et de l'infiltration du microenvironnement à l'intérieur de la tumeur.

Il est crucial de pouvoir distinguer chaque sous-région tumorale et les expressions différentielles des protéines entre ces régions pour identifier les mécanismes spécifiques se produisant dans chacune.

Ainsi, de nouvelles techniques ont été développées et ont permis le développement d'une nouvelle branche à la protéomique : la microprotéomique. Particulièrement, la microprotéomique localisée sur tissu permet d'identifier les protéines et donc les voies de signalisation au sein de petites régions de tissu en conservant l'information de localisation. Pour isoler exclusivement la population cellulaire, plusieurs stratégies peuvent être envisagées.

- Capture par microdissection laser
- Digestion in situ et extraction par microjonction liquide (LESA)
- Microdissection assistée par parafilm
- Digestion des protéines par migration passive dans un hydrogel

C'est dans ce contexte que se place notre projet. Comparée aux techniques précédemment décrites, notre approche est originale puisqu'elle implique l'agrandissement

du tissu en 3 dimensions pour sélectionner facilement les régions d'intérêt à analyser. Ainsi, cette méthode a pour but d'agrandir les coupes de tissu et de réaliser une microdissection manuelle sur des régions d'intérêts pour ensuite réaliser l'analyse protéomique

La microscopie par expansion (ExM)

Notre approche se base sur la microscopie par expansion (ExM) dans laquelle une coupe de tissu est agrandie de 3 à 100 fois (Chen et al., 2015a). Cette technique fut mise au point au Massachusetts Institute of Technology par Edward Boyden et son équipe en 2015 (Chen et al., 2015a). La microscopie par expansion permet, par des techniques de microscopie conventionnelles, l'observation de tissus biologiques, de cellules à l'échelle du nanomètre. Cette méthode contourne la limite de diffraction des microscopes optiques conventionnels qui offrent une résolution de l'ordre de la centaine de nanomètres en agrandissant physiquement l'échantillon pour descendre à une résolution de 70nm (Chen et al., 2015b). Relativement simple, elle peut être appliquée à différents types d'échantillons, des cultures cellulaires aux coupes de tissus en passant par différents types de préparation comme les tissus frais, congelés et fixés au préalable (Chen et al., 2015a). Elle ne nécessite aucun instrument sophistiqué ou coûteux.

La technique d'expansion repose sur la capacité des polyélectrolytes comme le sodium acrylate d'absorber une grande quantité d'eau entraînant un gonflement de cet hydrogel et donc son expansion (Chen et al., 2015a). Les polyélectrolytes permettent au gel de polyacrylamide de s'étendre et d'atteindre jusqu'à 4,5 fois sa taille originale (Figure 2-1 A). La méthode ExM permet d'observer spécifiquement une structure cible grâce à un système d'étiquette fluorescente. Cette étiquette permet également de lier les cibles marquées au réseau de polymère (Figure 2-1 B). Une fois l'échantillon intégré au gel, l'étape d'homogénéisation permet de supprimer les contraintes mécaniques entre les cellules pour conserver l'organisation et la structure du tissu lors de l'expansion. Cette étape est réalisée avec des protéases non spécifiques, le plus souvent la Protéinase K. Le gel est ensuite immergé dans de l'eau ultra pure. L'ajout d'eau en excès permet à l'hydrogel, et donc au tissu biologique, de s'étendre de façon isotropique, sans cassures ni déformation. Enfin, la

fluorescence du fluorochrome de l'étiquette ExM peut être visualisée au microscope, il est possible d'utiliser plusieurs fluorochromes différents pour des colorations multiples.

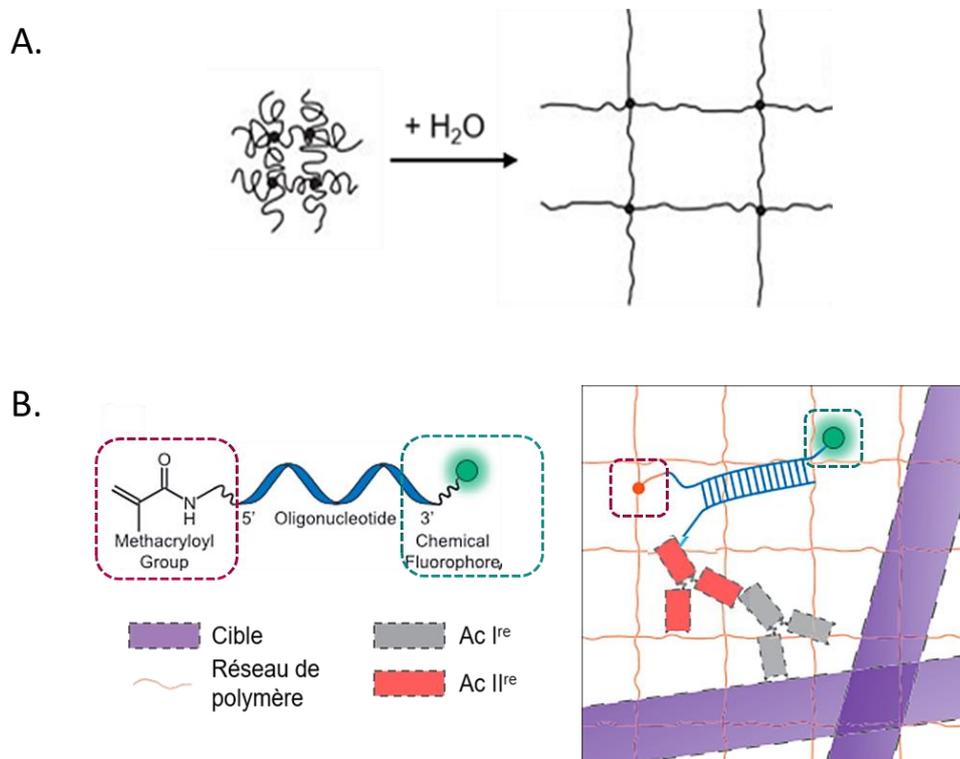


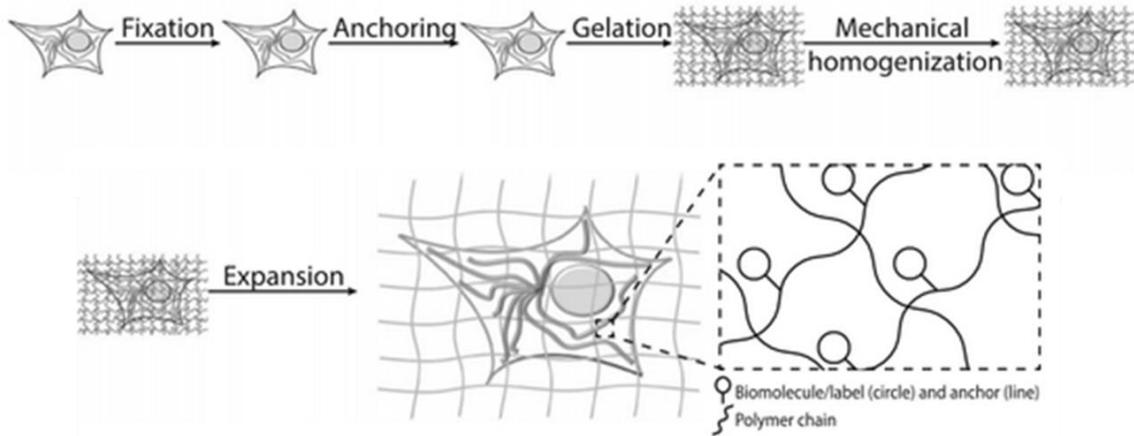
Figure 2-1 Concept de la microscopie par expansion

A. *Principe d'expansion des polyélectrolytes. Les gels de polyélectrolytes consistent en un réseau de molécules reliées entre elles. Ces gels ont la capacité d'absorber de grandes quantités d'eau en élargissant leur réseau.* B. *L'échantillon biologique est dans un premier temps incubé avec un anticorps primaire (gris) spécifique de la structure cible (violet). L'anticorps secondaire (rouge) est couplé avec un oligonucléotide qui lie une étiquette fluorescente (vert) et un groupe méthacryloyle (point orange). L'incorporation dans le réseau de polymère (lignes orange) est conservée après protéolyse. Adapté de (Chen et al., 2015a)*

Des variantes ont par la suite été développées. Notamment, la méthode ProExM (*Protein retention Expansion Microscopy*) permet d'ancrer les protéines de l'échantillon biologique sans avoir recours à une étiquette (Tillberg et al., 2016a). L'ancrage des protéines se fait grâce à l'acryloyl-X (AcX) qui forme un lien entre les protéines et le maillage d'acrylamide. L'AcX possède un groupement ester succinimide qui permet la liaison aux amines libres des protéines. L'autre extrémité de l'AcX va s'incorporer au gel d'acrylamide lors de la polymérisation grâce au groupement acryloyl. Ainsi, la localisation des protéines sera

conservée pendant l'expansion car elles vont s'étendre uniformément en même temps que le gel (Figure 2-2) (Gao et al., 2017).

A.



B.

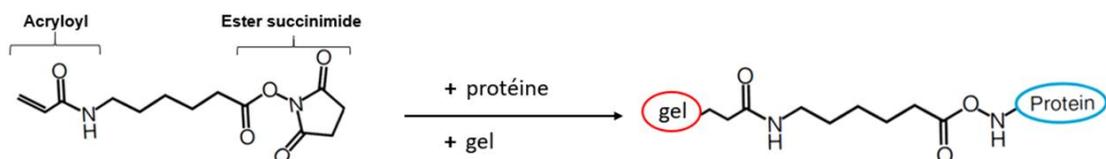


Figure 2-2: Les différentes étapes de la microscopie par expansion

A. Le tissu est dans un premier temps fixé avant de procéder. Une fois sur lame de verre, la coupe de tissu est incubée avec d'une molécule d'ancrage, l'Acryloyl-X, qui va ancrer les protéines du tissu au gel d'acrylamide. Après polymérisation du gel, une homogénéisation mécanique est réalisée avec la protéinase K ce qui permet de conserver les structures du tissu une fois étendu. La coupe incluse dans le gel est par la suite plongée dans plusieurs bains successifs d'eau jusqu'à expansion maximale. B. Structure et mode de liaison aux protéines de l'acryloyl-X. Adapté de (Gao et al., 2017)

L'ExFISH est une autre variante de la microscopie par expansion qui s'appuie la technique d'hybridation in situ permettant de cibler les ARNs (*Fluorescent In Situ Hybridization*, FISH) (Asano et al., 2018; Chen et al., 2016). Le protocole ExFISH utilise une molécule permettant d'ancrer l'ARN aux mailles du gel, le LabelX. Le LabelX permet la liaison au gel et est capable de réagir avec les guanines des ARNs. Ainsi, l'échantillon biologique est incubé avec le LabelX, et après gélification et expansion, l'échantillon est incubé avec les sondes complémentaires de l'ARN cible. Finalement, l'ExFISH permet l'observation d'ARNs par imagerie à l'échelle du nanomètre.

Application pour la microprotéomique localisée

Cette approche est pertinente en tant que nouvelle stratégie de microprotéomique. En effet, les régions d'intérêt peuvent être excisées après expansion et digérées de façon conventionnelle afin d'identifier les protéines spécifiques de ces régions (Figure 2-3 A). En divisant la taille de la zone découpée par le facteur d'expansion, il est possible de retrouver la taille d'origine de la zone digérée. Le facteur d'expansion est calculé en divisant la mesure après expansion divisé par la mesure avant expansion (Figure 2-3 B).

Les tailles des régions découpées, une fois rapportées à leur taille avant expansion, permettent d'analyser des régions de tissus parmi les plus petites décrites à ce jour en microprotéomique. En effet, sur un tissu étendu, une extraction de 1mm² correspond à une région de 200 à 300µm de côté avant expansion suivant le facteur d'expansion.

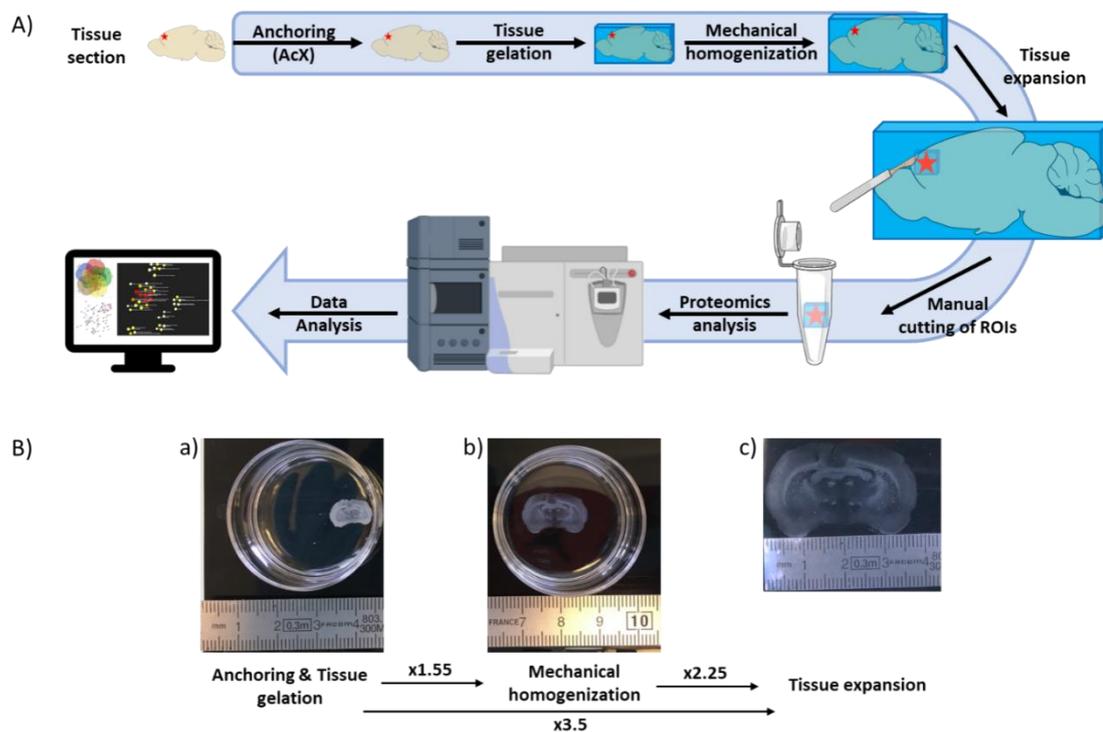


Figure 2-3 Microscopie par expansion adapté à l'analyse microprotéomique

A. Le tissu biologique est agrandi de façon isotropique par microscopie par expansion et les régions d'intérêts sont sélectionnées et excisées manuellement. Après découpe les protéines contenues dans les morceaux de gels sont digérées et les peptides extraits en vue d'une analyse MS. B. Une coupe de cerveau de rat dans sa taille d'origine de 1,2 cm, avant expansion et B. la même section de tissu après expansion avec une taille de 3,8 cm. Le facteur d'expansion est calculé en divisant la mesure après expansion (c) par la mesure avant expansion (a). Dans ce cas, nous pouvons estimer que le facteur d'expansion est de 3,5.

Dans ce chapitre, le protocole d'expansion (proExM) a été optimisé en vue de son utilisation en tant que technique robuste en microprotéomique. Des premiers tests ont été réalisés en utilisant le kit commercialisé par le MIT (Expansion Technology) et dédié à l'étude des protéines (ProExM). Ainsi, après l'expansion (x3.6) d'une coupe de cerveau de rat fixée et cryoprotégée d'une épaisseur de 60µm, 97 protéines ont été identifiées au sein de trois régions de 1mm² différentes (Figure 2-4). Même si peu de protéines sont identifiées, les protéines majoritaires du cerveau comme la myelin basic protein (MBP) ou encore l'actine ont pu être identifiées. Ces résultats représentent le point de départ pour le développement de la technique d'expansion de tissu pour les études protéomiques.

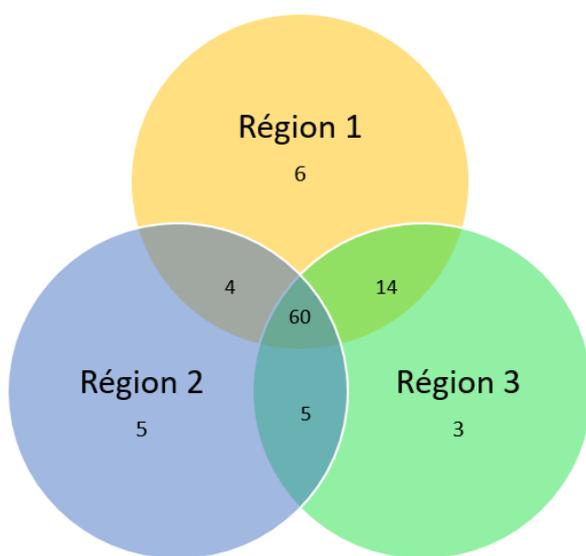


Figure 2-4 Diagramme de Venn des résultats d'identifications sur 3 régions grâce au kit commercial.

84 protéines sont identifiées dans la région 1, 74 pour la région 2 et 82 pour la région 3.

Le protocole a ultérieurement été adaptée à partir l'article de Paul W. Tillberg et de ses collaborateurs (Tillberg et al., 2016b). Notamment, du fait de sa grande toxicité, l'acrylamide en poudre utilisé dans le protocole original a été remplacé par de l'acrylamide sous forme liquide déjà mélangé au cross-linker. De plus, différents types de tissus ont été testés pour améliorer le nombre de protéines identifiées. Ainsi, des coupes de tissus FFPE de cerveau d'une épaisseur de 10µm ont permis l'identification de 61 protéines. Finalement, l'étape d'homogénéisation a été optimisée et la concentration de la protéinase K a été réduite de 8U/mL à 4U/mL. En effet, cette enzyme étant non spécifique, l'identification des protéines est rendue difficile par la génération de fragments aléatoires pouvant être perdu au cours des différentes étapes de préparation. Diminuer la concentration de protéinase K permet de

diminuer le nombre de clivages et par conséquent de multiplier par 2 le rendement d'identification (Figure 2-5). Ainsi, l'ensemble de ces améliorations ont permis d'identifier 135 protéines dans une région de 1mm² après expansion d'un facteur de 3.7.

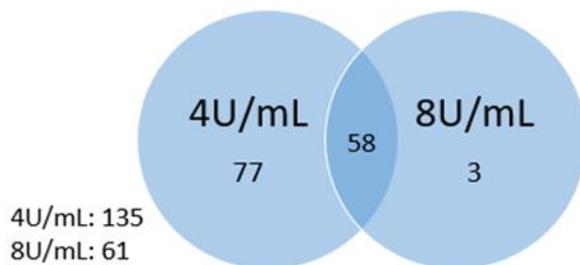


Figure 2-5 : Diagrammes de Venn sur les données d'identification en fonction de la concentration en Protéinase K.

Diminuer la concentration de protéinase K de 8U/mL à 4U/mL permet d'augmenter l'identification de 61 à 135 protéines identifiées.

Malgré ces optimisations, de nombreuses protéines sont encore retrouvées dans l'échantillon contrôle qui est une zone prélevée dans le gel hors de la coupe. La présence de ces protéines à l'extérieur du tissu est signe de leur migration au sein du gel. De ce fait les protéines ne peuvent pas être reliées à leurs localisations sur le tissu. Suite à ces résultats préliminaires, le développement s'est concentré sur l'amélioration du protocole d'expansion de tissu pour augmenter le nombre de protéines identifiées en conservant leur localisation ainsi que la reproductibilité de l'analyse.

II. Matériel et Méthodes

A. Section des tissus

L'ensemble des analyses ont été réalisées sur des cerveaux de rat, le premier fixé et inclus en paraffine (FFPE) et le second frais congelé. Des coupes de 12 μm de tissus frais et FFPE ont été utilisées.

B. Expansion de tissu

1. Gélification des tissus

L'expansion du tissu a été réalisée conformément au protocole publié par Tillberg et al., 2016. En bref, l'acryloyl-X (AcX) est déposé sur les tissus et laissé en incubation pendant une nuit dans une chambre humide à température ambiante. La solution gélifiante est préparée au dernier moment en ajoutant 8 μL de 4-Hydroxy TEMPO (0,625 %, inhibiteur de polymérisation des acrylamides), 8 μL d'APS et 8 μL de TEMED à 376 μL de solution monomère (acrylate de sodium 85,5mg / mL, 30 % acrylamide / bis -acrylamide 30 %; 37,5: 1, chlorure de sodium 1,02 g / ml et 10 % PBS 1X). Cette solution est finalement déposée sur le tissu puis piégé à l'intérieur d'une chambre de gélification fermée à l'aide d'une lamelle de verre et placée dans une chambre humide à 37 ° C pendant 2 heures.

2. Homogénéisation mécanique

Le protocole utilise la protéinase K (8U/mL) diluée dans un tampon d'homogénéisation (50 mM TrisHCl pH8, 1 mM EDTA, 0,5 % Triton X-100, 1 M NaCl). Le tissu imbriqué dans le gel est immergé dans 2 ml de protéinase K pendant 3 heures à 60 ° C. Plusieurs processus d'homogénéisation ont été testés pour remplacer la protéinase K. Tout d'abord, une solution de protéinase K diluée à 4 unités / ml est utilisée. Dans un second temps différentes enzymes sont utilisées et, pour minimiser la quantité totale d'enzyme, la solution d'homogénéisation est déposée avec un volume suffisant pour couvrir la surface de la coupe de tissu (généralement 500 μl) à la place d'une immersion totale. La première solution d'enzyme est une solution de LysC (20 μg / mL) diluée dans une solution d'eau déionisée avec de l'urée 6M

et 1 mM d'EDTA. Ainsi, 500 μ L sont déposés sur le tissu et placés à 37 ° C pendant une nuit. La deuxième enzyme est une solution de 1 ml de trypsine / EDTA (0.025% trypsine et 0.01% EDTA dans du *Phosphate Buffer Saline* (PBS)) utilisée couramment pour la dissociation des cellules en culture et provenant de Thermo Fisher Scientific (Courtabœuf, France). Les 1mL de cette solution sont déposées directement sur le tissu et placée à 37 ° C pendant 20 minutes. Finalement, un détergent, le SDS, a été utilisé comme dernier agent d'homogénéisation. Des optimisations ont été nécessaires pour l'utilisation de ce détergent : dans un premier temps 5 % de SDS dilué dans 1mL d'eau sont placés sur le tissu et incubés 30 minutes à 95 ° C ; dans un deuxième temps placé en chambre humide une nuit à 58 ° C.

3. Expansion et coupe

Ensuite, chaque coupe incluse dans un gel est séparée de la lame par immersion dans de l'eau HPLC pendant 1 h. L'eau est remplacée 3 fois toutes les 15 min. Le facteur d'expansion est calculé en divisant la taille finale du tissu après expansion par sa taille initiale. Une fois la taille d'expansion maximale atteinte, des régions spécifiques sont excisées précisément sur le tissu. Ici, des régions carrées de 1x1mm² et 5x5mm² sont découpées manuellement à l'aide d'un scalpel et transférés individuellement dans des tubes de micro centrifugation (LoBind 1.5mL microcentrifuge tubes, Eppendorf). Pour chaque coupe une région contrôle, à l'extérieur du tissu, est également découpée pour évaluer la diffusion éventuelle de protéines et de peptides dans le gel. Le facteur d'expansion est généralement compris entre 2,9 et 3,4. Les résultats en nombre d'identification seront exprimés en tenant compte de la taille de la région après expansion (5x5 mm² et 1x1 mm²) puis la taille réelle sera calculée pour chaque échantillon. Nous avons également utilisé un poinçon de biopsie d'un diamètre de 1,5 mm pour l'échantillonnage des régions après expansion. Par la suite, une analyse de type imagerie a été réalisée, ainsi 15 régions consécutives, de 1x1 mm² chacune, ont été excisées à travers le tissu afin d'assimiler chaque pièce de gel à un pixel d'image.

C. Analyse protéomique

Les morceaux de gel contenus dans les tubes de micro centrifugation sont recouverts de 50 μ L de NH₄HCO₃ (50 mM), le processus de réduction est réalisé par incubation des

morceaux dans 10 μ L de DTT (45 mM dans NH_4HCO_3 50 mM) 15 min à 50 ° C. Pour l'alkylation, on réalise une incubation avec 10 μ L d'IAA (100 mM dans NH_4HCO_3 50 mM) 15 min à température ambiante dans l'obscurité. Enfin, 20 μ L de trypsine (20 μ g / mL dans NH_4HCO_3 50 mM) sont ajoutés dans chaque échantillon et incubés pendant une nuit à 37 ° C. La digestion est arrêtée en ajoutant du TFA à 1 % du volume final.

D. Digestion *in situ* et extraction par microjonction liquide

Pour comparer à une autre méthode de protéomique localisée, une expérience de digestion *in situ* et extraction par microjonction liquide a été réalisée selon le protocole publié précédemment (Quanico et al., 2013). En bref, la région d'intérêt a été digérée avec de la trypsine à l'aide d'une imprimante chimique (CHIP-1000, Shimadzu, CO, Kyoto, Japon) et les peptides ont été extraits à l'aide de la plateforme TriVersa Nanomate (Advion Biosciences Inc., Ithaca, NY, USA) avec option analyse de surface et extraction liquide (LESA).

E. Analyse NanoLC-MS et MS / MS

Après digestion, les échantillons ont été dessalés sur colonnes C-18 Ziptip (Millipore, Saint-Quentin-en-Yvelines, France) élués par 80 % d'ACN et séchés sous vide. Les échantillons séchés ont été reconstitués dans 0,1 % d'AF / ACN (98 : 2, v / v). Deux systèmes ont été utilisés pour l'analyse. Les échantillons utilisant la protéinase K comme agent d'homogénéisation et ceux pour l'imagerie ont été analysés sur une nanoACQUITY UPLC (Waters). L'ensemble des autres échantillons a été analysé via EASY-nLC 1000 (Thermo Scientific). Chaque système de chromatographie nanoLC a été réglé pour obtenir les mêmes performances. Pour la nanoACQUITY UPLC, les échantillons ont été séparés par phase inverse sur une colonne de pré-concentration (nanoAcquity Symmetry C18, 5 μ m, 180 μ m x 20 mm) et une colonne analytique (nanoAcquity BEH C18, 1,7 μ m, 75 μ m x 250 mm). Les peptides ont été séparés en appliquant un gradient linéaire d'acétonitrile (5 % -35 %) dans 0,1 % d'acide formique pendant 2 heures, à un débit de 300 nl / min. Pour la EASY-nLC1000, les échantillons ont été séparés en phase inverse en utilisant une colonne Proxeon (75 μ m ID x 2 cm, 3 μ m, Thermo Scientific) et une colonne C18 (Acclaim PepMap, 75 μ m ID x 50 cm, 2 μ m, Thermo Scientific). Les peptides digérés ont été séparés en utilisant une quantité croissante d'acétonitrile (2 à 30 %) dans 0,1

% d'acide formique pendant 2 heures à un débit de 300 nL / min. Une tension de 2,4 kV a été appliquée pour électro-pulvériser l'éluant via la source nanospray. Les analyses MS ont été réalisées sur un spectromètre de masse haute résolution Q-Exactive Thermo Scientific selon une gamme de masse m/z comprise entre 300 et 1600, une résolution de 70 000 FWHM, un AGC de $3e6$ ions et un temps d'injection maximum de 120 ms. L'analyse MS/MS est réalisée en mode data dépendant, défini pour analyser les 10 ions les plus intenses de l'analyse MS. la plage de balayage était comprise entre m/z 200 et m/z 2000, un AGC a été fixé à $5e4$ ions et la résolution a été fixée à 17 500 FWHM. L'état de charge par défaut a été fixé à 2, les états de charge non attribués et +1 ont été rejetés, le HCD avec une énergie normalisée à 30 a été utilisé et l'exclusion dynamique a été activée pendant 25 s.

F. Analyse des données

Les données générées par l'analyse nanoLC-MS/MS ont été traitées avec le logiciel MaxQuant (Cox and Mann, 2008; Tyanova et al., 2015) (Version 1.6.1) à l'aide du moteur de recherche Andromeda (Cox et al., 2011). Les requêtes ont été faites en utilisant la banque de données *Rattus norvegicus* d'UniprotKB/Swiss-Prot (8,168 entrées, juillet 2019). Pour les paramètres de recherche, l'enzyme de digestion sélectionnée est la trypsine, un maximum de deux sites de clivage manqués autorisés. Les modifications variables choisies sont l'oxydation des méthionines et l'acétylation des protéines en N-terminal. La carbamidométhylation a été définie comme modification fixe pour l'analyse des échantillons après expansion. Une précision de masse initiale de 6 ppm a été sélectionnée pour les spectres MS et la tolérance MS / MS a été fixée à 20 ppm pour les données HCD. Pour la validation des peptides identifiés, un taux de faux positif (*false discovery rate* (FDR)) de 1 % a été utilisé. La quantification relative et sans marquage des protéines a été réalisée dans MaxQuant en utilisant l'algorithme MaxLFQ (Cox et al., 2014) avec les paramètres par défaut. La fonction de correspondance entre les analyses (*match between run* (MBR)), avec une fenêtre de correspondance de 0,7 min et une fenêtre d'alignement de 20 min, a été activée pour augmenter l'identification peptidique / protéique des échantillons. L'analyse des protéines identifiées a été réalisée à l'aide du logiciel Perseus (<http://www.perseus-framework.org/>) (version 1.6.0.7). Les protéines contenant uniquement des peptides modifiés et les contaminants potentiels ont été

supprimés. Les différentes méthodes ont été évaluées en termes de chevauchement des identifications des protéines (diagrammes de Venn) et également exprimées sous forme de coefficients de corrélation de Pearson (dot plot et valeur r). La visualisation par Diagramme de Venn a été réalisée avec BioVenn (<http://www.biovenn.nl/>) (Hulsen et al., 2008). Pour l'imagerie basée sur les données de quantification, les valeurs LFQ des protéines ont été utilisées pour reconstruire les images avec TIGR Multiexperiment (MEV v 4.9).

III. Résultats et Discussion

A. L'homogénéisation à la protéinase K permet d'identifier peu de protéines et implique la migration des protéines dans le gel

Dans un premier temps, nous avons basé notre protocole d'expansion de tissus sur les la publication de Tillberg concernant la microscopie d'expansion à rétention de protéines (proEXM) (Tillberg et al., 2016b). En bref, des coupes de tissus FFPE ont été traitées pour éliminer la paraffine et l'acryloyl-X (AcX) a été déposé directement sur la coupe de tissu. Comme expliqué précédemment, l'AcX lie les protéines au maillage d'acrylamide qui va ensuite copolymériser avec le polymère d'hydrogel expansible pendant le processus de gélification (Figure 2-2). Ensuite, l'étape d'homogénéisation mécanique consiste à supprimer les propriétés mécaniques de l'échantillon afin de conserver l'intégrité structurale et l'organisation structurale de l'échantillon lors de l'expansion. L'homogénéisation de l'échantillon incorporé dans le gel est réalisée grâce à une digestion enzymatique avec la protéinase K. Le gel est finalement immergé dans l'eau jusqu'à expansion maximale. Dans cette configuration, nous avons obtenu un facteur d'expansion d'environ 2,3 fois la taille d'origine.

À partir du tissu agrandi, des régions de tissu de 5x5 mm² ont été découpées manuellement avec une lame de scalpel (Figure 2-3). Une région a également été découpée à l'extérieur du tissu (hydrogel seul) comme contrôle pour évaluer la migration possible des molécules dans le gel. Les morceaux de gel ont ensuite été soumis à un protocole de digestion protéomique classique et les peptides extraits ont été analysés par spectrométrie de masse. Au sein du tissu, 14 protéines ont été identifiées et 9 protéines dans la région contrôle (en retirant les principaux contaminants) (Figure 2-6).

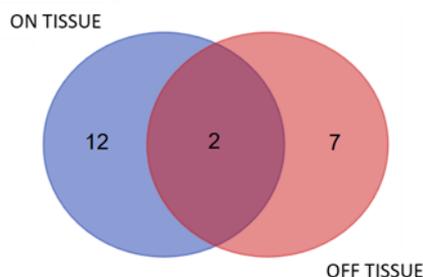


Figure 2-6 Nombre d'identification après homogénéisation à la protéinase K

Peu de protéines sont identifiées (14) et une migration importante est observée dans l'échantillon contrôle (9 protéines).

Le nombre de protéines identifiées est faible par rapport à ce qui est attendu pour cette surface de tissu. De plus, un nombre équivalent de protéines est identifié en dehors de la section de tissu. L'observation de peptides à l'extérieur du tissu pourrait indiquer une perte de certains peptides produits lors de l'étape d'homogénéisation à l'aide de la protéinase K. En effet, la protéinase K clive les protéines de façon non spécifique, elle coupe les liaisons peptidiques au niveau du groupe carboxyle des acides aminés et par conséquent conduit à la formation de petits peptides non ancrés au gel. Ces peptides peuvent alors diffuser dans et à l'extérieur de l'hydrogel lors de l'expansion. Cette diffusion entraîne une perte de localisation des protéines et réduit considérablement le nombre de protéines identifiées.

Néanmoins, certaines protéines ont pu être identifiées après l'expansion. Ces résultats sont encourageants et nous ont incités à remplacer la protéinase K lors de l'étape d'homogénéisation mécanique pour réduire les éventuelles pertes de protéines.

B. Analyse comparative de différentes alternatives pour la protéinase K

Différentes stratégies d'homogénéisation ont été explorées. Ainsi, nous avons testé deux enzymes avec des sites de clivage spécifiques, la trypsine et la LysC, et un détergent anionique, le SDS, qui perturbe les liaisons protéiques en linéarisant les protéines.

Pour la trypsine et la LysC, une homogénéisation incomplète est observée avec de nombreuses fissures apparues pendant l'expansion (Figure 2-7). En utilisant le SDS, l'expansion est plus homogène et l'expansion est assez similaire à celle obtenue avec la protéinase K (Figure 2-7). Le facteur d'expansion a été estimé à 2,1 en utilisant la trypsine et la LysC; à 2,3 avec le SDS et à 2,4 avec la protéinase K. La corrélation entre la taille des régions échantillonnées et leur taille réelle selon le facteur d'expansion est présentée dans le Tableau 2-1 avec une estimation du nombre correspondant de cellules.

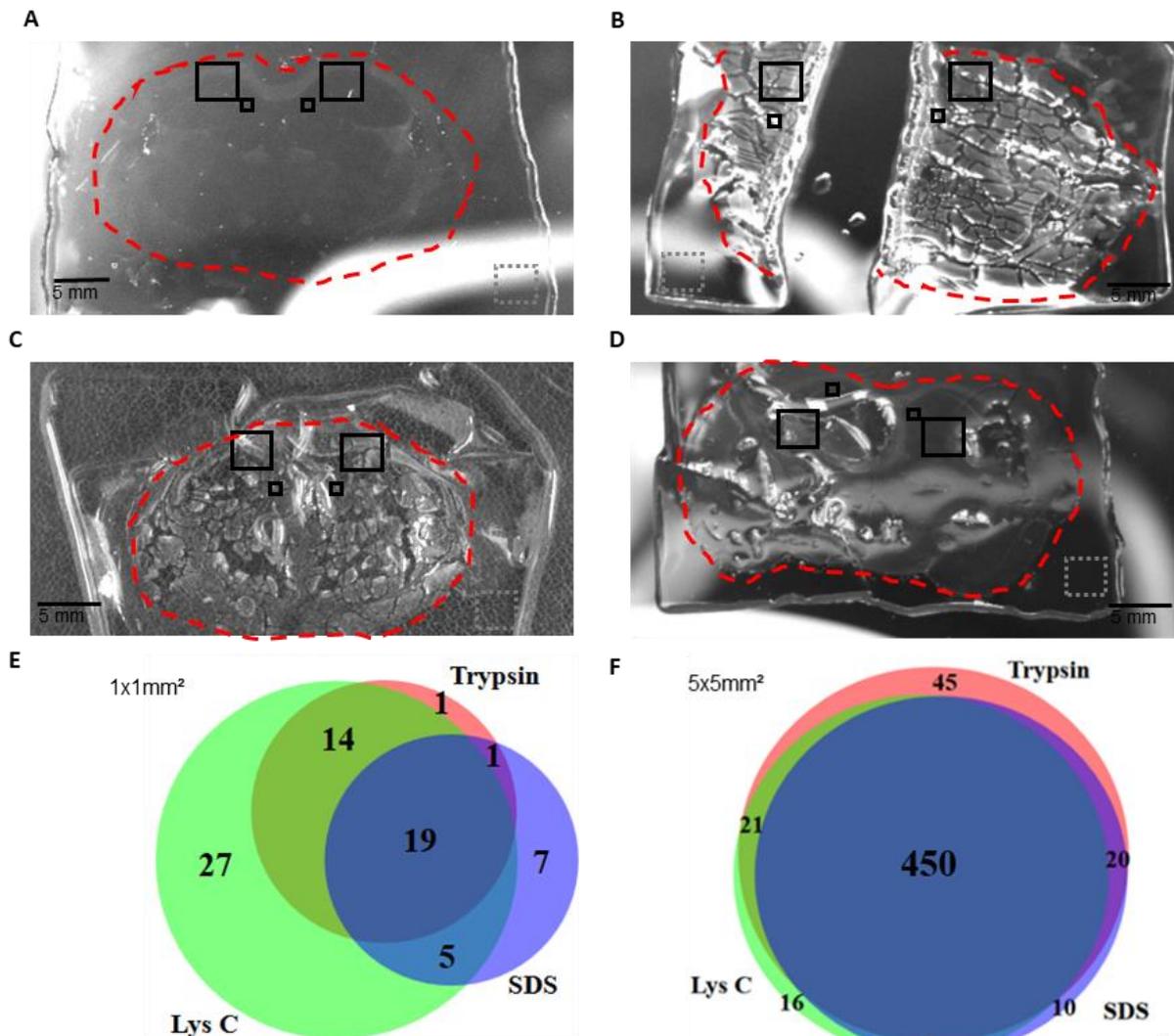


Figure 2-7 Analyse comparative de différentes alternatives à la protéinase K

A B C D. Cerveau de rat élargi après différents processus d'homogénéisation. A. Protéinase K. B. Trypsine. C. Lys C et D. SDS. Les carrés correspondent aux différentes zones coupées : 5x5 mm² et 1x1 mm² à l'intérieur du tissu et une zone de contrôle (5x5 mm²) dans le gel uniquement. Les protéines dans ces zones ont été digérées et identifiées par spectrométrie de masse. E / F. Le nombre de protéines identifiées dans chaque zone est représenté dans les diagrammes de Venn. E. Pour 5x5 mm, 536 protéines ont été quantifiées en utilisant de la trypsine comme agent d'homogénéisation (B), 495 pour Lys C (C) et 488 pour SDS (D). F. Pour les régions de 1 x 1 mm², 35 protéines ont été quantifiées pour la trypsine (B), 65 protéines avec Lys-C (C) et 32 protéines pour SDS (D).

Tableau 2-1 Facteur d'expansion et nombre de protéines identifiées selon l'agent d'homogénéisation

	Homogénéisation	Facteur d'expansion	Taille réelle (Nombre de cellules)	Nombre de protéines identifiées (contrôle)
5x5mm ²	Proteinase K	2.4	2.08x2.08mm ² (~24 500)	2 (2)
	Trypsine	2.1	2.38x2.38mm ² (~32 000)	536 (2)
	LysC	2.1	2.38x2.38mm ² (~32 000)	495 (0)
	SDS	2.3	2.17x2.17mm ² (~26 600)	488 (0)
1x1mm ²	Proteinase K	2.4	416x416µm ² (~980)	0 (0)
	Trypsine	2.1	476x476µm ² (~1 280)	35 (2)
	LysC	2.1	476x476µm ² (~1 280)	65 (0)
	SDS	2.3	435x435µm ² (~1 070)	32 (0)

Concernant le nombre de protéines identifiées, les mêmes résultats sont obtenus en utilisant la protéinase K avec un faible nombre de protéines. Pour les autres conditions, le nombre de protéines identifiées est significativement supérieur. Dans les régions de 5x5 mm², jusqu'à 536 protéines ont été quantifiées en utilisant la trypsine comme agent d'homogénéisation, 488 pour le SDS et 495 pour la Lys C (Figure 2-7). Au total, 450 protéines communes ont été identifiées dans toutes les conditions ce qui correspond à 79% de chevauchement. Pour les régions de 1x1 mm², 35 protéines ont été quantifiées pour la trypsine et 32 protéines pour le SDS. Le nombre d'identifications augmente jusqu'à 65 protéines avec l'homogénéisation avec la LysC. Pour ces régions on observe 19 protéines communes, ce qui correspond à seulement 26% de chevauchement (Figure 2-7).

Dans la région contrôle, seulement 2 protéines ont été quantifiées pour l'homogénéisation avec la LysC, 1 pour la trypsine et aucune lorsque le SDS est utilisé. Ces résultats révèlent une migration nulle ou très minime à l'intérieur du gel.

Cependant, comme l'homogénéisation avec la trypsine et la LysC n'est pas homogène, il est difficile d'estimer précisément la taille réelle de la région digérée (Figure 2-7). Sur la base de la mesure du facteur d'expansion, il existe une variation d'environ 20 % du nombre de

cellules analysées entre les différents tests (Tableau 2-1). Le nombre de cellules est également difficile à estimer lorsque l'expansion n'est pas isotrope comme pour la trypsine et la LysC. Dans ces 2 cas, on observe la présence d'un nombre élevé de fissures et de déformations, ce qui signifie que certains groupes de cellules ont conservé leur taille d'origine et par conséquent que la densité de cellules est plus élevée que prévu dans la région échantillonnée (Figure 2-7).

Compte tenu de ces résultats, nous avons choisi de concentrer notre développement sur le SDS pour remplacer la protéinase K dans le processus d'homogénéisation. Le SDS semble être une alternative intéressante car il ne clive pas les protéines en peptides et de ce fait les protéines complètes sont ancrées au gel (Ku et al., 2016; Tillberg et al., 2016b; Treweek et al., 2015). En tant qu'agent d'homogénéisation, le SDS permet l'élimination des lipides, conserve l'intégrité des protéines et n'affecte que leurs conformations et leurs charges. De plus, le SDS est bon marché et facile à utiliser contrairement aux enzymes qui ont besoin d'une température spécifique pour conserver leur activité.

C. Optimisation et reproductibilité du protocole d'homogénéisation au SDS pour l'analyse microprotéomique

Dans un premier temps, nous avons optimisé le protocole en augmentant le temps d'homogénéisation pendant une nuit dans une chambre humide et en diminuant la température à 58 ° C pour éviter l'évaporation trop rapide de la solution et l'altération de l'hydrogel. Ces améliorations se traduisent par un facteur d'expansion supérieur, jusqu'à 3,1 fois la taille initiale du tissu.

D'autres paramètres ont été modifiés pour augmenter le facteur d'expansion et le nombre de protéines identifiées mais n'ont pas conduit à une amélioration significative. Par exemple, le volume de SDS en excès (3ml) ne permet pas une meilleure homogénéisation qu'un volume à peine suffisant pour couvrir la coupe incluse dans le gel (500µl). Une étape de démasquage d'antigène (tampon TrisHCl 20 mM pH 9, 30 minutes 90°C) a également été ajoutée au protocole après le déparaffinage de la coupe FFPE afin d'améliorer l'efficacité de la digestion des protéines et donc augmenter le nombre d'identifications finales. En effet, dans les tissus FFPE de nombreux pontages peuvent s'établir entre le paraformaldéhyde et les

protéines, rendant certaines séquences difficilement accessibles et pouvant limiter l'action des enzymes de digestion par exemple. Le démasquage, par action d'une forte chaleur (90°C) et d'un tampon (Tris HCl), va augmenter la linéarisation des protéines libérant ainsi l'accès à des séquences d'intérêts. Cette étape n'a cependant entraîné aucune augmentation nette du nombre d'identifications et le tissu était plutôt plus fragile à manipuler. Un protocole de digestion en gel (Shevchenko et al., 2007) des morceaux de gel a été évalué pour améliorer l'extraction des peptides du réseau d'acrylamide, il est en effet possible qu'il y ait des peptides qui restent dans le gel et soient perdus lors d'une digestion en solution. Aucune différence significative n'a été observée. Le protocole de digestion en gel possède plus d'étapes et présente un risque plus important de perdre les morceaux d'hydrogel entre les différents lavages. Il est donc plus contraignant et il est préférable de conserver la digestion en solution pour les prochaines expériences.

La reproductibilité a été évaluée en comparant les régions échantillonnées en trois exemplaires sur des sections consécutives du mésencéphale (S1, S2 et S3) (Figure 2-8). Une seule protéine a été identifiée dans la zone de contrôle de la section S3 et aucune protéine n'a été identifiée dans les zones de contrôle S1 et S2. Ces résultats indiquent que les protéines conservent leur emplacement dans le tissu malgré le temps de préparation plus long. Ces résultats sont conformes à nos attentes, les protéines sont ancrées au réseau d'acrylamide et par conséquent ne migrent pas à l'extérieur du tissu.

La représentation en diagramme de Venn a été utilisée pour évaluer la reproductibilité qualitative et les coefficients de corrélation de Pearson ont été calculés (dot plot et valeur r) pour la reproductibilité quantitative (Figure 2-8). Le nombre de protéines identifiées dans chaque réplicat, quelle que soit leurs abondances, est présenté dans les diagrammes de Venn, tandis que les valeurs quantitatives (valeurs LFI) ont été utilisées pour calculer les diagrammes de corrélation. Pour les régions de 5x5 mm², un total de 1259 protéines différentes a été identifié avec 898, 945 et 994 protéines respectivement dans les réplicats 1; 2 et 3 (Figure 2-8 ; Tableau 2-2). L'optimisation a permis de doubler le nombre d'identification de protéines par rapport aux expériences précédentes. Sur la Figure 2-8, le diagramme de Venn montre que 675 protéines sont partagées par toutes les réplicats ce qui correspond à 53,6 % de protéines communes avec une petite variation individuelle (environ 5 %).

Tableau 2-2 Facteur d'expansion et nombre de protéines identifiées après homogénéisation au SDS dans chaque triplicata

	Section (1,2cm)	Taille après expansion	Facteur d'expansion	Taille réelle (Nombre de cellules)	Nombre de protéines identifiées (control)
5x5mm ²	S1	3,7 cm	3,08	1.62x1.62mm ² (~14 860)	898 (0)
	S2	3,65 cm	3,04	1.64x1.64mm ² (~15 229)	945 (0)
	S3	3,65 cm	3,04	1.64x1.64mm ² (~15 229)	994 (1)
1x1mm ²	S1	3,7 cm	3,08	325x325µm ² (~598)	335 (0)
	S2	3,65 cm	3,04	329x329µm ² (~613)	392 (0)
	S3	3,65 cm	3,04	329x329µm ² (~613)	454 (1)

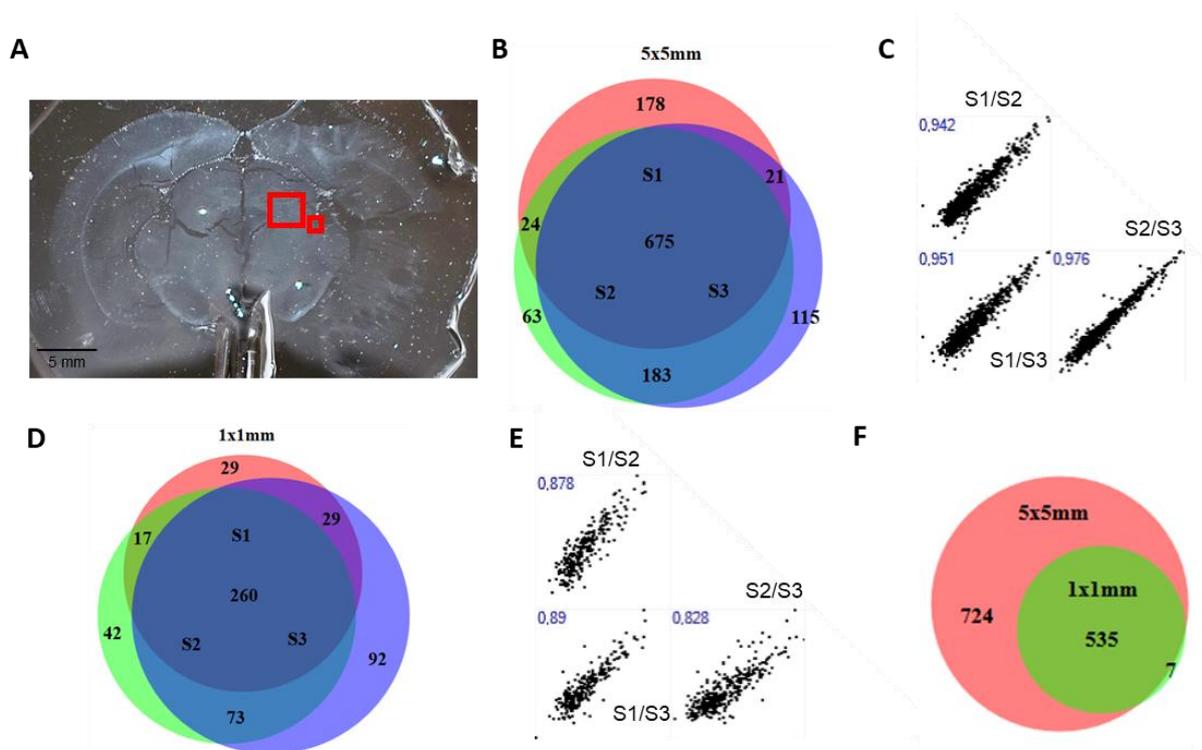


Figure 2-8 Analyse de la reproductibilité protéomique avec le protocole d'homogénéisation utilisant le SDS

A. Cerveau de rat agrandi après homogénéisation au SDS, les carrés rouges représentent les régions analysées. B / D. L'expérience a été réalisée en trois exemplaires sur 3 coupes de tissu consécutives, représentation en diagramme de Venn du nombre de protéines identifiées dans les extractions de (B) 5x5 mm et (D) 1x1 mm. C / E. Corrélation de Pearson de la quantification sans marquage des protéines identifiées dans les extractions de (C) 5x5 mm et (E) 1x1 mm. F. Représentation en diagramme de Venn des protéines identifiées dans les extractions 1x1 mm et 5x5mm.

En comparant les réplicats deux à deux il est intéressant de noter que les réplicats S2 et S3 apparaissent très proches avec une grosse proportion de protéines communes (environ 80 %) (Figure 2-9). Le réplicat S1 contient une plus faible proportion de protéine que les deux autres réplicats (environ 60 % des chevauchements). Ces résultats peuvent être expliqués par le facteur d'expansion plus élevé pour le réplicat 1 (3.08) que les réplicats 2 et 3 (3.04) (Tableau 2-2). Les régions découpées dans les réplicats 2 et 3 ont une taille réelle de 1.64x1.64mm², en admettant que la taille moyenne d'une cellule soit de 15 µm de diamètre, il est possible d'estimer à plus de 15 200 le nombre de cellules digérées (Tableau 2-2). Dans le réplicat 1, la région découpée a une taille réelle de 1.62x1.62mm² ce qui correspond à 14 860 cellules (Tableau 2-2). Il est possible que la variation observée provienne de cette différence dans le nombre de cellules analysées mais aussi de la variation des caractéristiques histologiques des coupes consécutives.

Malgré cette variation, les valeurs du coefficient de corrélation de Pearson montrent que les protéines identifiées ont la même valeur de quantification entre les réplicats (moyenne de 0,956) (Figure 2-8).

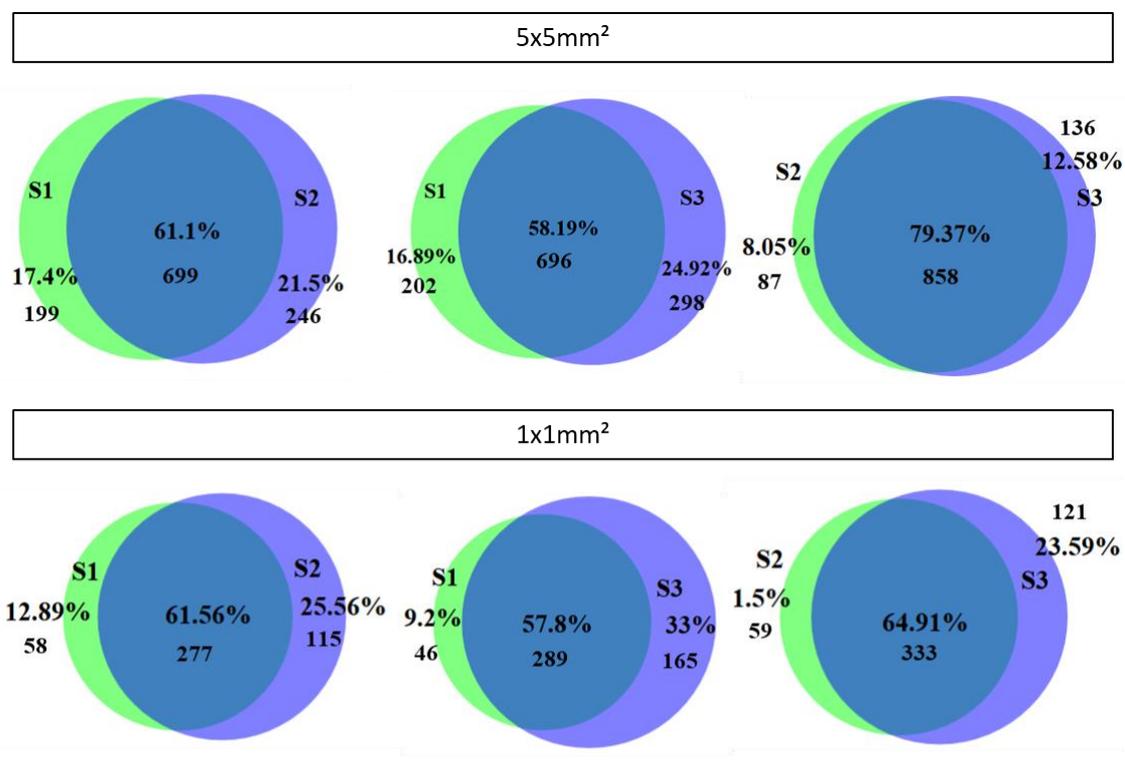


Figure 2-9 Comparaison des réplicats deux à deux

La comparaison directe montre une plus grande similarité entre les réplicats S2 et S3 qu'avec le réplicat S1.

Pour les échantillons provenant des régions taille 1x1 mm², 335 protéines ont été identifiées dans la section 1 ; 392 dans la section 2 et 454 dans la section 3 pour un total de 542 protéines différentes (Tableau 2-2). La variabilité est plus élevée que pour les régions de 5x5 mm², en effet 260 protéines sont partagées par toutes les répliquats ce qui représente 47,97 % de protéines communes (Figure 2-8) (environ 60 % pour la comparaison deux à deux (Figure 2-9)). Les mêmes variations sont observées dans l'analyse de la reproductibilité quantitative avec un coefficient de corrélation de Pearson variant de 0,83 à 0,89 (Figure 2-8).

Il est intéressant de noter que plus de 50% des protéines identifiées dans les régions de 1x1mm² sont identifiées grâce à la fonctionnalité *match between run*. Cette fonctionnalité permet d'identifier des peptides présents dans des échantillons sur la base de leur mesure en MS, mais non identifiés par MS/MS. Ainsi, les peptides sont identifiés par correspondance entre les analyses (temps de rétention et masses similaires). Plus de 98 % des protéines identifiées dans ces régions sont également identifiées dans les régions de 5x5 mm² (Figure 2-8) ce qui suggère que l'utilisation d'un échantillon de référence plus grand (5x5 mm²) permet d'identifier et de quantifier plus de protéines à partir de plus petite région.

L'utilisation du SDS comme agent d'homogénéisation engendre une expansion isotrope avec un facteur d'expansion d'environ 3,1 pour une section de tissu FFPE de 12 µm d'épaisseur. Il est alors facile de sélectionner spécifiquement une région d'intérêt à échantillonner à l'aide d'une simple lame de scalpel. L'analyse protéomique d'une région de 5x5 mm² permet d'identifier jusqu'à 994 protéines. Compte tenu du facteur d'expansion, la zone réelle digérée est d'environ 1,6x1,6 mm² (Tableau 2-2). La réduction de la zone d'échantillonnage à 1x1 mm² après expansion correspond à une véritable zone digérée de 330x330µm² (Tableau 2-2). Cette zone correspond à environ 600 cellules et jusqu'à 454 protéines ont été identifiées dans cette région. Finalement, pour cette position donnée, un pourcentage de variation individuelle de 10 % a été observé avec 47,97 % de protéine commune (Figure 2-8). Une bonne reproductibilité quantitative est observée avec une corrélation de Pearson moyenne de 0,87 (Figure 2-8).

La reproductibilité à la fois qualitative et quantitative observée est très satisfaisante même s'il existe des variations. Ces variations peuvent être expliquées par l'évolution des caractéristiques histologiques sur les coupes de tissus consécutives, mais aussi par la difficulté

d'échantillonnage précis de la même région en raison de la transparence du tissu après expansion. La reproductibilité des petites régions est plus faible que celle des plus grandes régions de tissu, probablement en raison de la découpe manuelle et de la variation du facteur d'expansion. En effet, l'échantillonnage manuel des régions avec une lame de scalpel peut entraîner des différences dans la taille des morceaux gels. En ce sens, un outil capable de découper une région spécifique de taille reproductible a ensuite été utilisé.

D. Utilisation d'un emporte-pièce pour biopsie pour l'échantillonnage de petites régions

Pour améliorer la reproductibilité, un emporte-pièce pour biopsie de 1,5mm de diamètre a été utilisé pour couper et extraire spécifiquement un disque d'hydrogel contenant du tissu. L'utilisation de cet outil permet de couper avec précision un disque de gel de 1400 μm de diamètre (Figure 2-10). Cette zone correspond à une région de tissu d'environ 460 μm de diamètre avant expansion (environ 940 cellules) (Tableau 2-3).

Tableau 2-3 Facteur d'expansion et nombre de protéines identifiées à l'aide d'un emporte-pièce pour biopsie dans chaque triplicata

	Section (1,2cm)	Taille après expansion	Facteur d'expansion	Taille réelle (Nombre de cellules)	Nombre d'identification (contrôle)
Punch Biopsy	S1	3,7 cm	3,08	\emptyset 454 μm^2 (~916)	582 (0)
	S2	3,65 cm	3,04	\emptyset 460 μm^2 (~941)	682 (0)
	S3	3,65 cm	3,04	\emptyset 460 μm^2 (~941)	700 (1)

Pour trois régions consécutives, nous avons obtenu un diamètre de $1400 \pm 37\mu\text{m}$ (Figure 2-10).

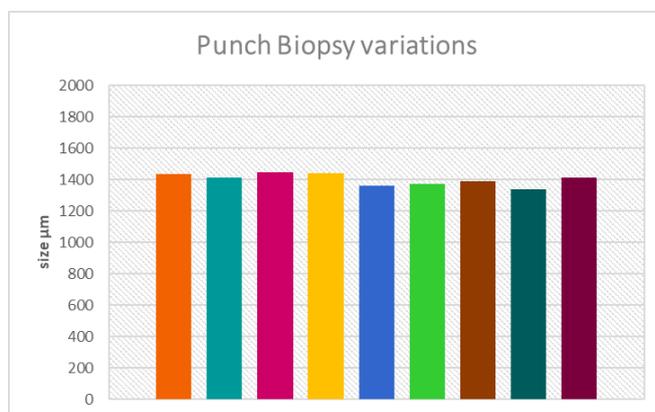


Figure 2-10 Variation des tailles de découpe de gel avec un emporte-pièce 1,5 mm

La valeur moyenne d'un gel découpé à l'aide de l'emporte-pièce de 1,5mm est de 1400 μm , les tailles se situent entre 1360 μm et 1445 μm .

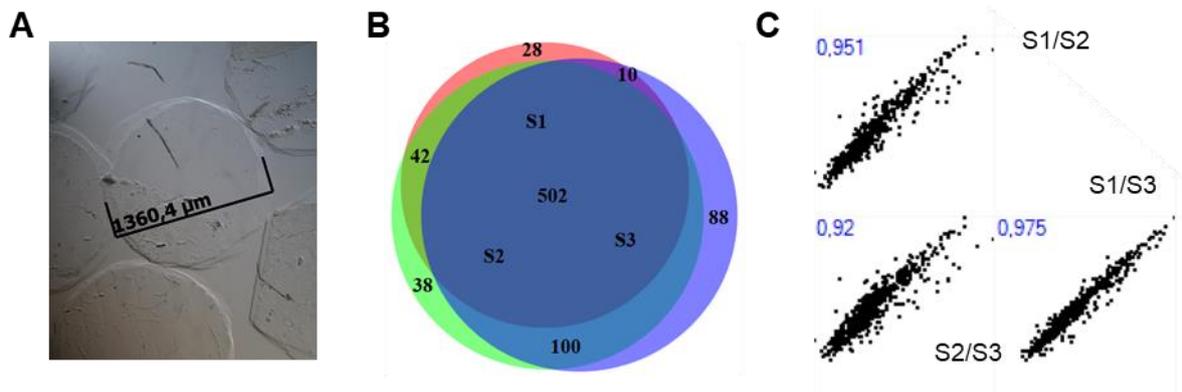


Figure 2-11 Analyse de la reproductibilité de l'échantillonnage par emporte-pièce

A. Disque de gel coupé à l'aide d'un emporte-pièce de 1350 μm de diamètre. B. Représentation sous forme de diagramme de Venn du nombre de protéines identifiées dans les trois réplicats. C. Corrélation de Pearson dans chaque réplicat.

Un total de 808 protéines a été obtenu dans l'ensemble des réplicats (582, 682 et 700 dans les réplicats S1, S2 et S3 respectivement (Figure 2-11 ; Tableau 2-3).

Le diagramme de Venn montre une bonne reproductibilité entre les trois réplicats avec 62,13 % de protéine commune. La reproductibilité quantitative est également élevée avec une corrélation de Pearson moyenne de 0,95 (Figure 2-11). Pour la variabilité individuelle entre les réplicats, les mêmes observations que ci-dessus sont faites, on observe en effet des différences plus importantes entre les réplicats S1 et S3 (Figure 2-12). Avec l'emporte-pièce, le coefficient de corrélation de Pearson augmente jusqu'à 0,97 (Figure 2-11) montrant que réaliser des découpes reproductibles permet de diminuer la variation.

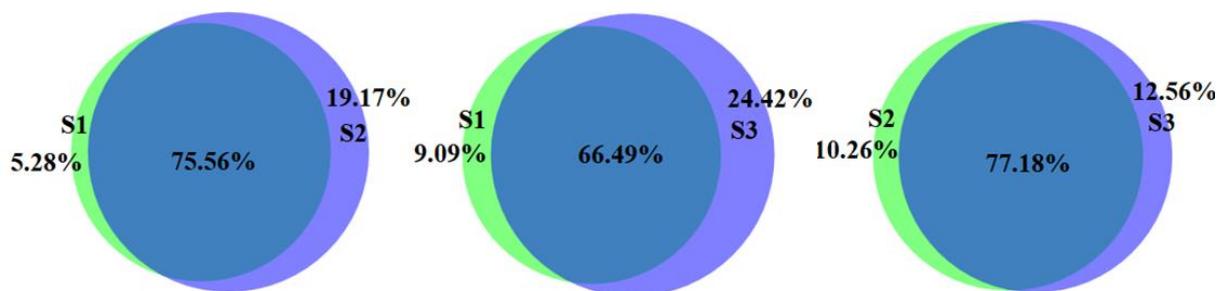


Figure 2-12 Utilisation d'un emporte-pièce : Comparaison des réplicats deux à deux

Comme précédemment, la comparaison directe montre une plus grande similarité entre les réplicats S2 et S3 qu'avec le réplicat S1.

Ces résultats mettent en évidence que la variation observée dans le contenu protéique est principalement due à la découpe manuelle de la région d'intérêt. En effet, il est difficile d'échantillonner manuellement des morceaux de gel de même taille. La reproductibilité est considérablement améliorée avec l'utilisation de l'emporte-pièce. Cette lame circulaire permet de découper des régions de même taille facilement et avec précision. Par conséquent, la variabilité individuelle diminue à 6,3 % et le nombre de protéines communes augmente de façon considérable à 62,13 % (Figure 2-11). De même, la reproductibilité quantitative augmente avec un facteur de corrélation Pearson moyen de 0,95 (Figure 2-11).

Pour un échantillon correspondant à une région de 460 μm de diamètre, nous avons pu identifier une moyenne de 655 protéines provenant d'environ 940 cellules (Tableau 2-3). Une faible variation individuelle est encore observée, certainement due à une variation mineure dans la position de l'échantillon en raison de la transparence de la coupe de tissu. L'évolution des caractéristiques histologiques entre les coupes de tissus et la variation du facteur d'expansion peuvent également être source de ces variations.

Ainsi, en tenant compte du facteur d'expansion, diminuer le diamètre de l'emporte-pièce jusqu'à 300 μm pourrait permettre de réduire considérablement la région d'intérêt à 100 μm . Récemment, des hydrogels délivrant des enzymes ont été utilisés pour digérer et extraire des peptides de façon localisée sur un échantillon de tissu (Harris et al., 2013; Rizzo et al., 2017). Dans ce cas, les hydrogels sont obtenus à l'aide du même type d'emporte-pièce que nous avons utilisé. Cette technique a permis d'identifier environ 700 protéines en utilisant un diamètre d'hydrogel de 357 μm (Rizzo et al., 2017). La principale limite de cette approche est la manipulation délicate de ces hydrogels de quelques millimètres de diamètre. En ce sens, inclure le tissu dans le gel permet d'atteindre et de diminuer facilement la taille des régions d'intérêt.

E. Comparaison avec la digestion *in situ* et microjonction liquide

Par la suite, une comparaison a été réalisée avec la méthode de digestion in-situ et extraction par microjonction liquide (LESA) (ETAT DE L'ART, Figure 0-9). Cette stratégie de micro extraction a été développée dans notre équipe et est utilisée notamment pour la caractérisation du microenvironnement de tissu (Quanico et al., 2013, 2017c; Wisztorski et al.,

2017). Les résultats obtenus après expansion ont été comparés à ceux obtenus en utilisant la meilleure configuration pour la microdigestion *in situ*, *i.e.* sur tissu frais congelé.

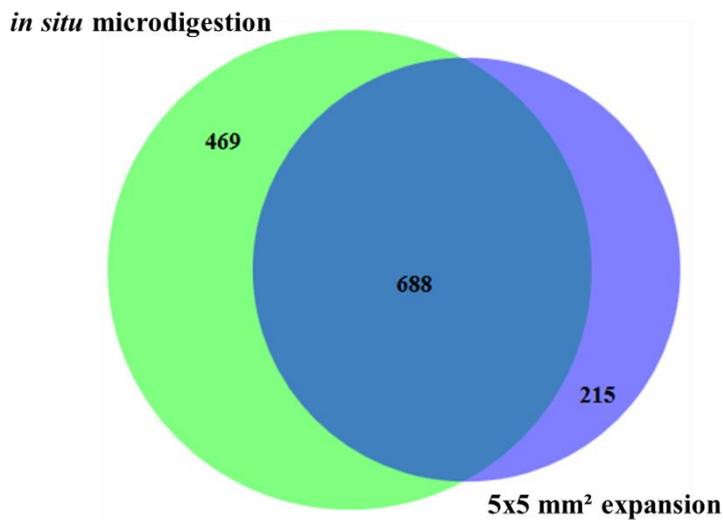


Figure 2-13 Comparaison des protéines identifiées après expansion et avec digestion *in situ* et extraction par micro jonction liquide

Un nombre comparable de protéine est identifié après expansion (903, au moins 2 valeurs dans les trois réplicats de 5x5mm²) et après microdigestion in situ et microextraction (1157 protéines).

La digestion *in situ* combinée à la microextraction par LESA effectuée indépendamment permet l'identification de 1157 protéines dans une région de 900 μm de diamètre. Dans la même région, 903 protéines ont été identifiées au moins deux fois dans les 3 réplicats après expansion (Figure 2-13). Un nombre équivalent de protéines est identifié dans les deux techniques, avec un total de 688 protéines communes (Figure 2-13). Il est intéressant de noter que 469 protéines sont uniques à l'extraction par LESA tandis que 215 sont uniques à la stratégie d'expansion de tissu.

Les protéines identifiées dans les deux cas sont de masses moléculaires variées et assez similaires entre les deux techniques et se situent entre 5 et 619 kDa (Figure 2-14). Des protéines de poids moléculaire élevé comme la serine / thréonine-protéine kinase mTOR (288 kDa) et ankyrin-3 (284 kDa) ont été identifiées dans les deux méthodes. Les protéines huntingtin (343 kDa) et fibronectine (272 kDa) ont été identifiées respectivement avec le LESA et l'expansion. Des protéines plus petites ont également été identifiées, la myotrophine (12 kDa) et la thiorédoxine (11 kDa) respectivement avec LESA et l'expansion de tissu. La complexin-2 (15 kDa) a été identifiée dans les deux méthodes.

La corrélation moyenne de Pearson observée entre les deux techniques est de 0,678 (Figure 2-14). Bien que la taille de la région et les méthodes de digestion soient légèrement

différentes, il semble que les protéines identifiées ont une valeur de quantification équivalente. La variation observée est sans aucun doute due à la fixation des tissus, en effet l'extraction LESA a été réalisée sur du tissu frais congelé, les protéines sont donc plus accessibles que dans un tissu fixé et inclus en paraffine (Wiśniewski, 2013).

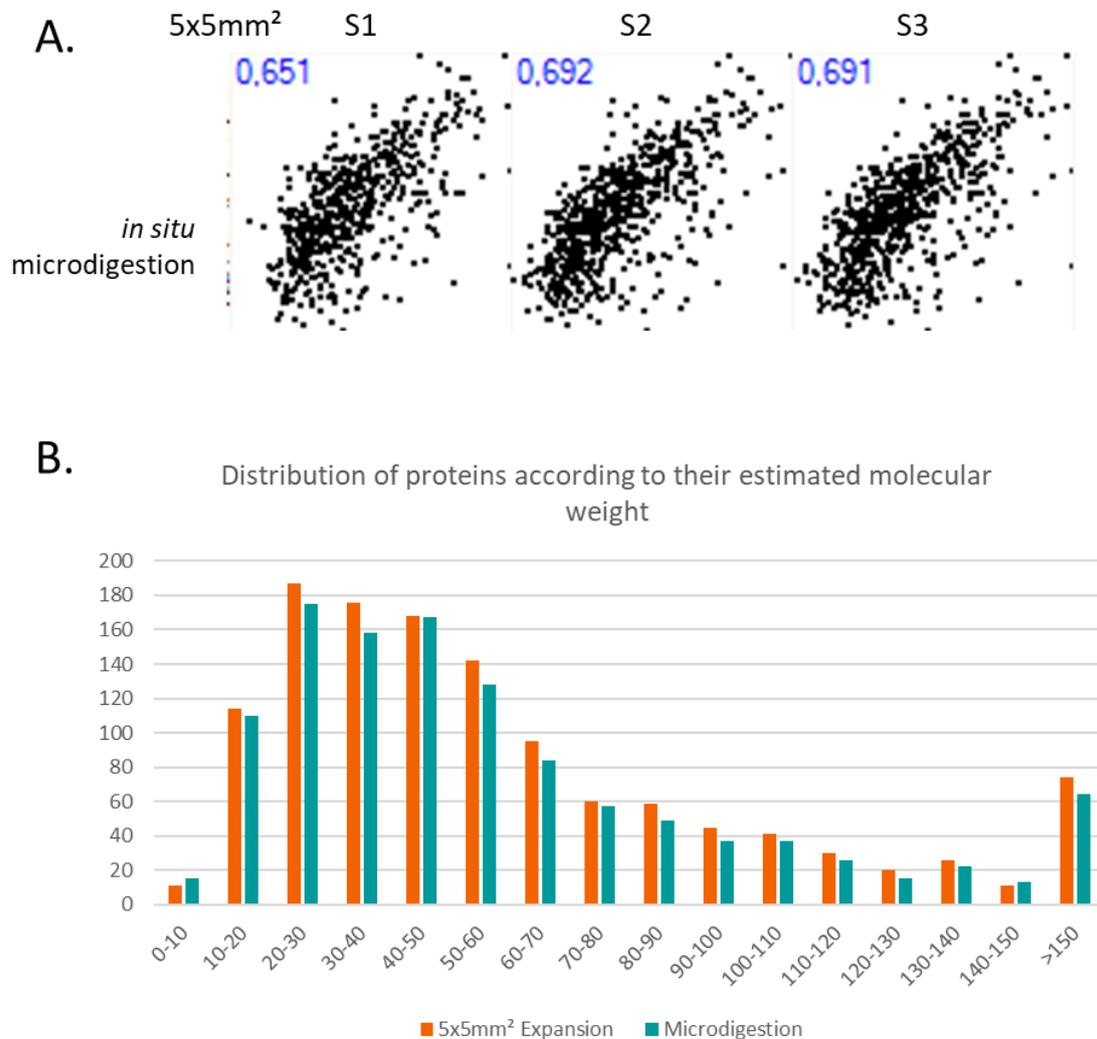


Figure 2-14 Comparaison qualitative des protéines identifiées après expansion et avec digestion *in situ* et extraction LESA

A. Corrélation de Pearson entre les valeurs LFQ et B. Poids moléculaire estimé des protéines détectées dans chaque condition.

La protéomique localisée sur tissu permet d'étudier le microenvironnement et l'hétérogénéité des tumeurs (Quanico et al., 2017a; Wisztorski et al., 2013). La technique de la microdigestion *in situ* et microextraction par microjonction liquide permet de comprendre des processus physiologiques spécifiques dans des régions discrètes, notamment pour l'étude

des gliomes (Le Rhun et al., 2016). Ainsi, la microprotéomique par expansion de tissu est une technique prometteuse pour atteindre le même objectif et potentiellement identifier des biomarqueurs spécifiques. Entre les deux stratégies, l'expansion semble plus accessible car elle ne nécessite aucun instrument et peut être effectuée facilement sans formation.

F. « *Imaging-like* » : Profil d'expression protéique basé sur la quantification après expansion

Par la suite, nous avons réalisé un profilage des protéines du tissu basé sur la quantification en analysant des points adjacents consécutifs. L'imagerie basée sur les données de quantification a déjà été réalisée avec d'autres techniques au sein du laboratoire (Franck et al., 2013; Quanico et al., 2015; Wisztorski et al., 2016). Pour ce faire, nous avons échantillonné 15 régions consécutives de 1x1 mm² chacune le long d'une ligne à travers le tissu afin d'assimiler chaque pièce de gel à un pixel d'image (Figure 2-15). Il est alors possible d'utiliser des données de quantification pour la reconstruction d'images afin de fournir une image protéique basée sur les données de quantification.

L'ensemble des extraits a été digéré et analysé par LC-MS/MS. Un total de 511 protéines a été identifié dans ces extractions. Les images ont été reconstruites et ainsi chaque pixel représente un morceau de gel avec sa position sur le tissu. La couleur de chaque pixel dépend des différences d'expressions des protéines : le rouge représente la surexpression des protéines tandis que le bleu représente la sous-expression (Figure 2-15).

Compte tenu du facteur d'expansion, la résolution spatiale de ces images se situe aux alentours de 300µm. Sur la (Figure 2-15) il est facile d'observer l'expression de chaque protéine à travers le cerveau de rat étudié.

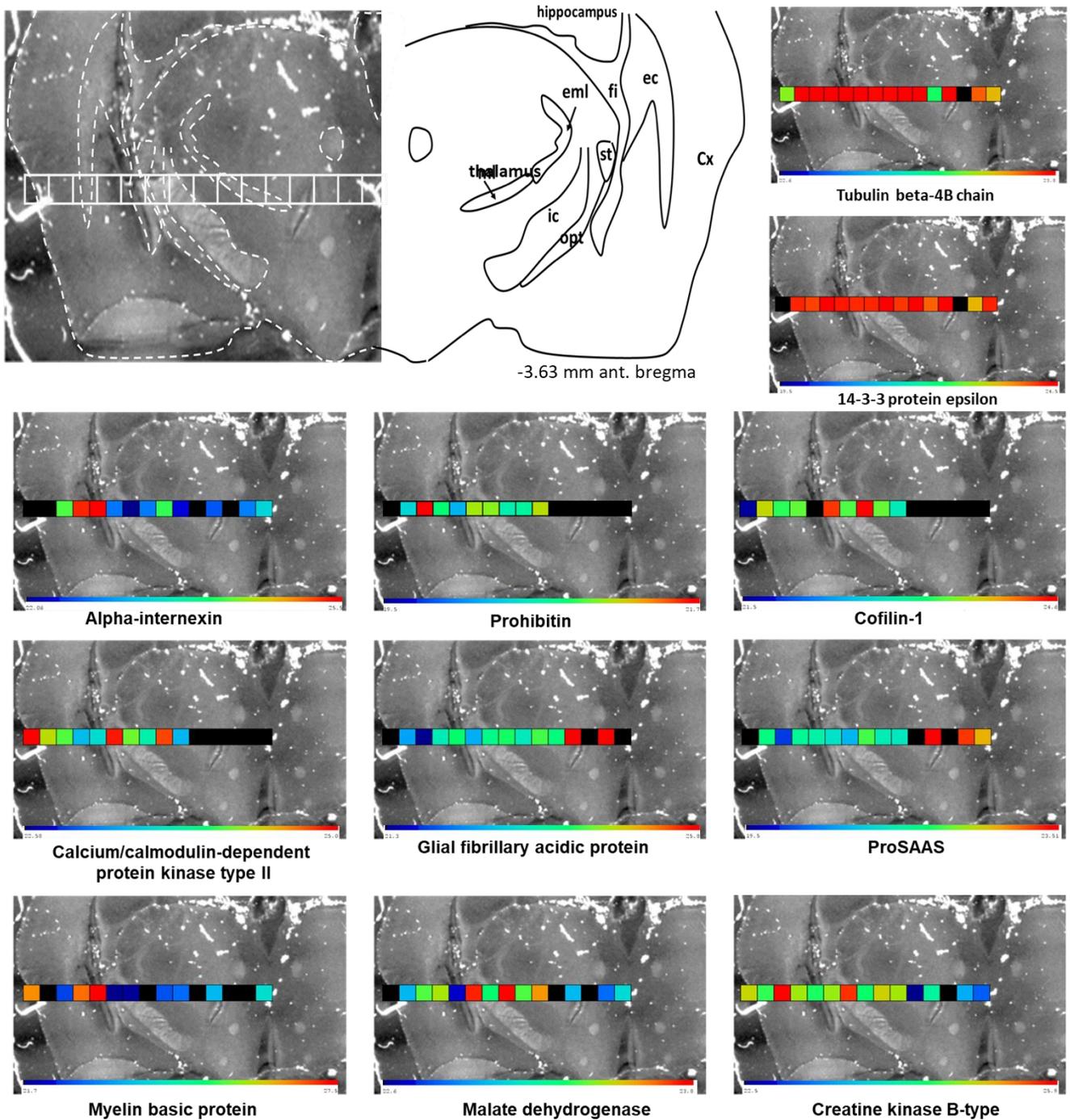


Figure 2-15 Reconstitution d'images basées sur les profils d'expression de différentes protéines après expansion de tissu

Image d'une section de cerveau de rat après expansion, les pixels représentent les extractions pour la microprotéomique. Les quantifications relatives sans marquage obtenues après analyse nanoLC-MS/MS permettent de reconstruire des images protéiques de la zone étudiée. ec: external capsule; fi: fimbria of the hippocampus; ic: internal capsule; Cx: cerebral cortex; st: stria terminalis; opt: optic tract; eml: external medullary lamina; ml: medial lemniscus

L'expression de protéines spécifiques dans chaque région du cerveau a ensuite été examinée. Ainsi, la cartographie spatiale obtenue est notamment confirmée par la surexpression de la créatine kinase (CKB) et de calcium/calmodulin-dependent protein kinase kinase 2 (Camkk2) dans le cortex et l'hippocampe qui sont connus pour contenir ces protéines (Erondu and Kennedy, 1985; Katagiri et al., 2010). Dans ces régions, nous observons également la surexpression d'alpha-internexin (Ina) par rapport au thalamus (Figure 2-15). L'alpha-internexin est un filament intermédiaire impliqué dans la morphogenèse des neurones et la croissance des neurites (Bott and Winckler, 2020). La cofilin-1 (Cfl1) apparaît préférentiellement exprimée dans la région de l'hippocampe. Cette protéine spécifique, en tant que régulateur de la dynamique de l'actine, peut contribuer aux processus dégénératifs (Bamburg et al., 2010). Une distribution préférentielle de GFAP et de Pro SAAS est observée dans le thalamus, ce qui est conforme aux publications précédentes (Franck et al., 2013; Katagiri et al., 2010). En particulier, la quantification de la MBP est conforme aux données d'immunohistologie (Wisztorski et al., 2016). Des protéines comme la tubulin beta-4 et la protéine 14-3-3 epsilon ont également été identifiées. Ces protéines sont connues pour être exprimées de manière homogène au sein du cerveau. Nous avons observé ici un faible coefficient de variation de 2 % et 1 % respectivement pour ces 2 protéines, ce qui est conforme à la littérature (Figure 2-15).

Finalement, l'analyse de régions spécifiques par microprotéomique par expansion permet de reconstruire des images moléculaires indirectes basées sur l'identification et la quantification de protéines. Dans cette configuration, la résolution spatiale est inférieure à 350 μm . L'expansion de tissu adapté à la protéomique représente une nouvelle opportunité intéressante pour l'imagerie par spectrométrie de masse, car elle permet de cartographier un grand nombre de protéines de forte et faible abondance. La technique a été développée sur une section de tissu cérébral de rat, mais pourrait être facilement appliquée à un échantillon tumoral pour étudier l'hétérogénéité des protéines dans la tumeur et identifier des biomarqueurs spécifiques.

IV. Conclusion et Perspectives

La microscopie par expansion a été initialement développée pour augmenter physiquement la taille des tissus de manière isotrope pour les observer à une échelle nanométrique avec un microscope conventionnel (Chen et al., 2015a, 2015b). Cette technique convient également pour l'observation de protéines (proExm) à l'aide d'anticorps fluorescents (Tillberg et al., 2016b). Cependant, cette technique n'a jamais été appliquée à l'analyse protéomique à grande échelle. L'expansion de tissu est facile à utiliser et adaptable pour la protéomique localisée, car elle permet de réaliser une extraction précise dans un échantillon de tissu.

Nous avons démontré que l'expansion de tissu est compatible avec la protéomique localisée sur coupes de tissus FFPE. Cette approche permet l'identification de protéines sur une zone précisément définie inférieure à 350 μ m. Cette stratégie, en combinaison avec une analyse protéomique conventionnelle, permet l'identification d'environ 655 protéines pour une surface équivalente à environ 940 cellules. L'identification de protéine à partir de tissu expansé montre une bonne reproductibilité à la fois qualitative et quantitative.

La protéinase K est la stratégie la plus courante pour réaliser l'étape d'homogénéisation mécanique (Tillberg et al., 2016b). Cependant, cette enzyme a une faible spécificité et génère de petits peptides qui diffusent à travers l'hydrogel. Au contraire, en tant que détergent, le SDS dénature les protéines natives en perturbant les liaisons non covalentes responsables de la structure tridimensionnelle de la protéine. Ainsi, les protéines sont linéarisées et il est possible d'étendre le tissu. De plus, bon marché et largement utilisé en laboratoires, le SDS permet d'éliminer les lipides qui pourraient avoir un impact/masquer sur l'identification des protéines. L'utilisation du SDS comme agent d'homogénéisation permet également de déstructurer les membranes pour faciliter l'expansion, on observe une expansion de facteur 3 pour une section de tissu FFPE de 12 μ m d'épaisseur.

L'expansion de tissu peut être combinée avec ExFISH qui implique l'utilisation d'un lien qui permet à l'ARN d'être attaché de manière covalente au gel (Asano et al., 2018; Chen et al., 2016). Ensuite, l'imagerie par hybridation fluorescente in situ (FISH) de l'ARNm peut être réalisée (Chen et al., 2016). La région contenant le signal de fluorescence peut être délimitée,

coupée et analysée pour observer la présence de la protéine traduite par spectrométrie de masse. La relation ARNm / protéine correspondante a déjà été étudiée et permet d'évaluer le niveau de corrélation de régulation (Budnik et al., 2018). Il est également possible d'identifier d'autres protéines dans la même région et par conséquent d'identifier des voies de signalisation potentielles. De plus, il est possible de se concentrer sur des ARN non codants particuliers qui peuvent en réalité coder pour des protéines alternatives (Cardon et al., 2019, 2020b).

Dans l'ensemble, ces résultats démontrent le potentiel de cette stratégie pour l'identification et la localisation des protéines. Toutefois, potentiellement plus de protéines pourraient être identifiées. La fixation au formaldéhyde est avantageuse pour la préservation et la conservation de l'architecture des coupes de tissu, mais entraîne des réticulations de formaldéhyde et de nombreuses protéines ne sont toujours pas détectables en utilisant ces approches (Wiśniewski, 2013). De ce fait et compte tenu des résultats très encourageants obtenus sur les tissus fixés, le protocole développé a été appliqué à une coupe de tissu frais congelé. Ainsi, le protocole d'expansion a été utilisé sur une coupe de cerveau de rat frais congelé de 20µm d'épaisseur avec deux optimisations majeures. La première optimisation concerne la durée d'homogénéisation qui a été ramenée à 1h en raison de la plus grande accessibilité des protéines dans un tissu frais congelé comparé à un tissu FFPE. La seconde optimisation consiste à utiliser des lames sans revêtement poly-Lysine. En effet, ce revêtement entraîne une forte adhérence de la coupe de tissu à la lame et il est impossible de la décoller pour procéder à l'expansion. Ainsi, pour la suite des expériences, les tissus ont préalablement été déposés sur lame sans revêtement poly-Lysine (lame ITO).

Les premiers résultats montrent qu'il est possible d'identifier des protéines à partir de tissus frais mais la reproductibilité est très faible (Tableau 2-4). En effet, le tissu n'étant pas fixé, les différentes étapes, notamment l'homogénéisation mécanique, conduisent à une migration importante des protéines au sein du gel entraînant la perte de la localisation réelle. Ceci est illustré par une grande variation entre les réplicats et la présence de nombreuses protéines dans l'échantillon témoin (19 ; 55 et 1) (Tableau 2-4).

Tableau 2-4 Facteur d'expansion et nombre de protéines identifiées dans chaque triplicata à partir de coupes congelées

Gel + tissu (1,5cm)	Taille après expansion	Facteur d'expansion	Taille réelle (Nombre de cellules)	Nombre d'identification (contrôle)
S1	4,9 cm	3,26	1.53x1.53mm ² (~13 248)	136 (10)
S3	5 cm	3,33	1.50x1.50mm ² (~12 733)	78 (26)
S2	4,8 cm	3,2	1.56x1.56mm ² (~13 772)	788 (0)

Les tissus frais ont l'avantage de n'avoir subi aucun traitement mais sont par conséquent plus fragiles et plus sujets à la diffusion des protéines labiles. La post fixation au paraformaldéhyde pourrait être intéressante pour stabiliser les protéines, également un temps plus court d'homogénéisation au SDS pourrait être appliqué afin de limiter leur délocalisation.

L'expansion de tissus frais congelés entraîne cependant un autre problème majeur : la clarification. En effet, la coupe de tissu devient invisible à l'œil nu. Il est de ce fait impossible de distinguer le tissu à l'intérieur du gel et par conséquent d'effectuer une découpe précise et reproductible. Néanmoins, il est possible d'utiliser un colorant comme le Fast green qui permet de voir la bordure des tissus ou un colorant fluorescent comme le Hoechst qui colore spécifiquement le noyau (Figure 2-16). Il reste cependant à évaluer la compatibilité du colorant fluorescent Hoechst avec l'analyse en spectrométrie de masse.

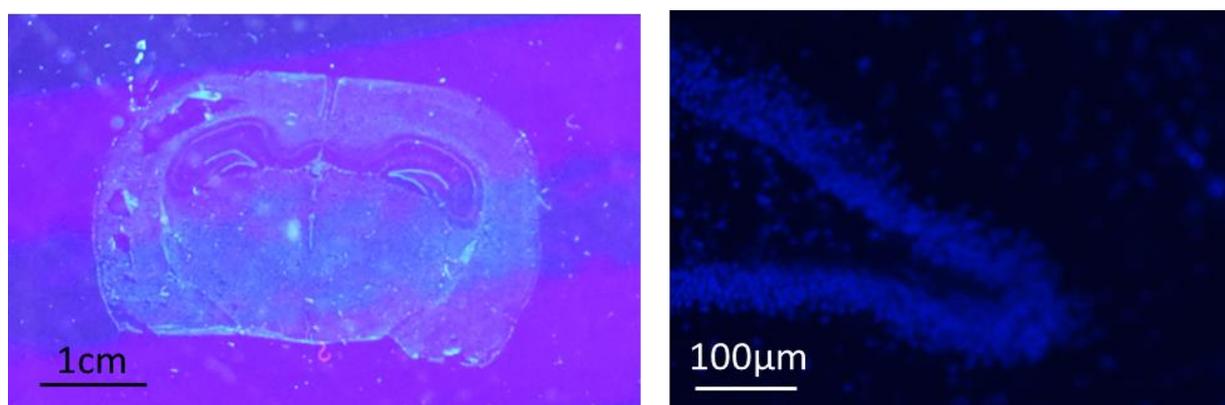


Figure 2-16 Coupe de cerveaux frais après expansion et marquage au Hoescht observée sous lampe ultraviolet

Une coupe de cerveau de rat frais congelée après expansion et colorée au Hoeschst

Concernant l'analyse sur tissu FFPE, d'autres paramètres peuvent être optimisés pour améliorer le rendement en protéine des régions analysées. En effet, la présence du gel est un facteur pouvant être limitant pour la digestion des protéines et également pour l'extraction des peptides de digestion. De plus, les approches en Top-Down ne sont pas compatibles avec ce type de préparation. Afin d'améliorer ces différents points, l'utilisation d'un gel réversible a été envisagée. Le N-N'Bis(acryloyl)Cystamine (BAC) est un lien qui permet les pontages entre les filaments d'acrylamide. Le BAC possède un pont disulfure qu'il est possible de cliver en conditions réductrices. Après validation de la technique sur un standard protéique (ubiquitine), un premier test sur une coupe de cerveau FFPE de 10 μ m a été réalisé. Les morceaux de gels après expansion ont finalement subi une réduction au DTT 15 minutes à 50°C. Ces analyses n'ont cependant permis d'identifier qu'une vingtaine de protéines dont la majorité sont des contaminants. Le faible rendement d'identification peut être dû à différents facteurs. Ces tests ont été réalisés avant les différentes optimisations exposées dans ce chapitre, donc en utilisant de la protéinase K comme agent d'homogénéisation qui entraîne un faible nombre d'identifications. Également, les nombreux polymères présents dans la solution pourraient masquer le signal en spectrométrie de masse. Un autre paramètre à prendre en compte pour ces expériences avec gels réversibles est la possible formation de complexes « protéine-AcryloylX-acrylamide » ce qui va entraîner un décalage de masse et donc une diminution de l'identification en spectrométrie de masse en tandem.

L'utilisation de l'acryloylX pour ancrer les protéines dans le gel est nécessaire pour conserver la localisation des protéines au sein du tissu. Ce lien réagit d'un côté avec les groupements amines des protéines et de l'autre avec les groupements acrylamide. Cependant, l'ancrage irréversible des protéines est problématique pour l'identification en spectrométrie de masse. Les peptides liés au réseau d'acrylamide restent ancrés dans le gel et ce même après digestion. En ce sens, l'utilisation d'un lien clivable en conditions réductrices est intéressante. En effet, rendre l'ancrage réversible permettrait d'augmenter le rendement d'identification protéique. Des tests préliminaires ont été entrepris pour réaliser un lien clivable. Ce lien a été synthétisé par l'ajout en concentrations égales (1:1) de 3-butène-1amine, contenant un groupement acryloyl, et de DTSP (3,3'-Dithiodipropionic acid di(N-hydroxysuccinimide ester) contenant deux esters succinimides. Le produit de cette réaction

contient un groupement acryloyl pour la liaison au gel et un ester succinimide pour la liaison aux protéines. Ce premier test a été réalisé en parallèle d'une homogénéisation à la protéinase K, par conséquent très peu de protéines ont pu être identifiées. Il serait intéressant de combiner l'utilisation de ce lien avec l'homogénéisation au SDS dans le but d'augmenter le niveau d'identification.

Dans un tissu agrandi 3 fois, il est possible de minimiser considérablement la taille de la zone analysée. Ainsi, une façon de diminuer encore la résolution spatiale consisterait à augmenter encore la taille du tissu. Pour atteindre cet objectif, différents protocoles ont déjà été développés pour augmenter le facteur d'expansion. Une nouvelle technique d'expansion nommée ZOOM (Zoom by hydrOGel cONversion Microscopy) a été en mesure d'améliorer considérablement le facteur d'expansion jusqu'à 8 (Park et al., 2019). Cette technique utilise la réaction de conversion d'hydrogel qui implique l'hydrolyse alcaline des amides primaires non ioniques de l'hydrogel pour former des groupes carboxylates ioniques. L'hydrolyse permet d'introduire des résidus ioniques dans le réseau d'hydrogel pour faciliter l'expansion (Park et al., 2019).

De plus, la microscopie par expansion itérative (iExM), composée de plusieurs étapes d'expansion successives, permet d'agrandir l'échantillon biologique jusqu'à 20 fois (Chang et al., 2017). Dans cette technique, un second hydrogel est synthétisé dans l'espace créé par la première expansion (Figure 2-17). Le premier gel est réalisé avec un gel réversible qui est clivé après la polymérisation du second gel, permettant ainsi une seconde expansion.

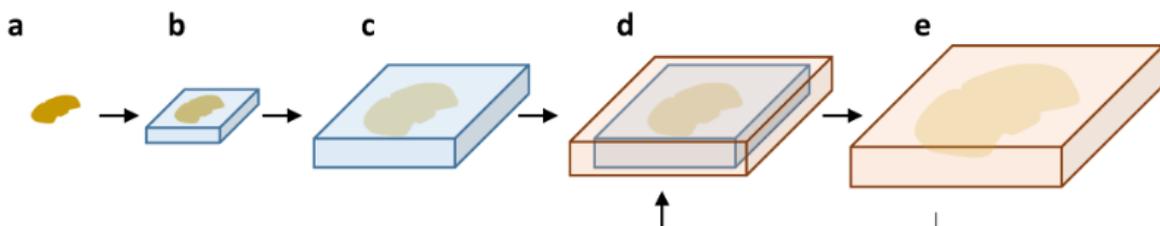


Figure 2-17 Principe de la microscopie par Expansion itérative (iExM)

Schéma de l'expansion itérative, montrant comment une coupe de cerveau peut être agrandie plusieurs fois. Tout d'abord, le réseau de gel de polyélectrolyte gonflable contenant un lien clivable est formé à travers la coupe de cerveau (b). La coupe est étendue après homogénéisation mécanique comme dans le protocole ExM (c). Après l'expansion, un second gel est formé dans le premier (d) puis étendu après dissolution du premier gel (e). Ce processus (d, e) peut être appliqué à plusieurs reprises pour augmenter encore le grossissement physique. (Chang et al., 2017)

L'optimisation de la composition du gel peut permettre d'agrandir l'échantillon de 10 fois sa taille initiale (Truckenbrodt et al., 2018). La principale modification dans la composition de ces gels est la présence de N,N-diméthylacrylamide (DMAA) à la place de l'AA/BAA. Ce composé permet la formation d'un gel dont les mailles sont très espacées. Ainsi une extraction de $1 \times 1 \text{ mm}^2$ sur une coupe de cerveau après expansion 10x correspond en réalité à une zone d'environ $100 \times 100 \mu\text{m}^2$. Ce protocole a été envisagé pour diminuer la taille d'analyse et les premiers résultats sont encourageants. En effet, le premier test réalisé et a permis l'identification d'environ 20 protéines (en retirant les principaux contaminants) (Figure 2-18). Cependant, le gel est très fragile et difficile à manipuler ce qui implique la formation de nombreuses cassures. De plus, compte tenu le facteur d'expansion, le tissu n'est pas visible à l'œil nu. Trouver un colorant compatible avec la spectrométrie de masse est primordial pour réaliser des extractions de manière contrôlée.

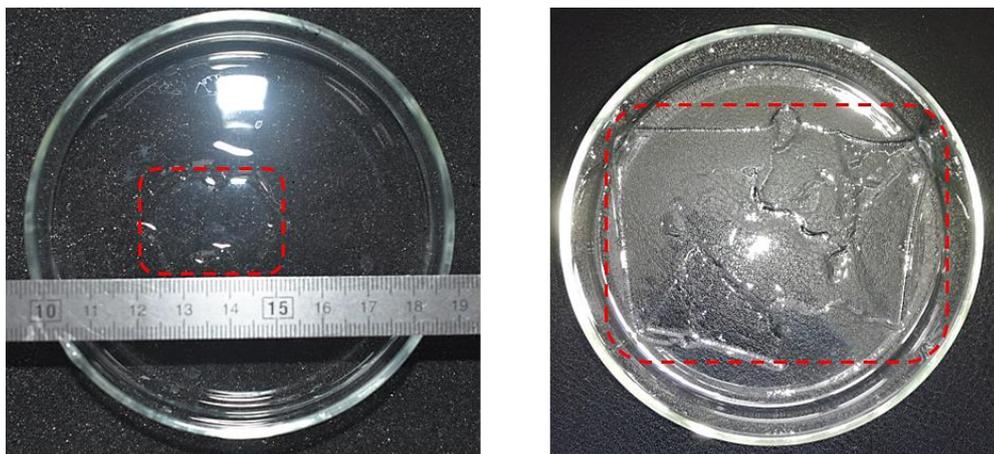


Figure 2-18 Cerveau de rat avant et après Expansion x10
A. Gel avant expansion B. Gel étendu 10 fois.

Ce premier test démontre la possibilité de réduire encore la taille d'analyse, bien que l'échantillon devienne par conséquent plus fragile à manipuler. L'utilisation de ce type de gel pourrait permettre d'aller plus loin dans les analyses de type cellules uniques. Dans un tissu agrandi 10 fois, l'utilisation d'un emporte-pièce de $200 \mu\text{m}$ de diamètre permet de couper spécifiquement une région de $20 \mu\text{m}$ ce qui correspond approximativement au diamètre d'une cellule. Ainsi, cette stratégie pourrait être utilisée comme un outil pour l'analyse protéomique cellule unique dans le but d'identifier des variations au niveau cellulaire et de comprendre des

mécanismes physiopathologiques très spécifiques. Des optimisations restent cependant nécessaires pour une utilisation robuste en microprotéomique.

Facile à réaliser et ne nécessitant pas d'instruments coûteux, l'expansion de tissu est une technique prometteuse pour la microprotéomique localisée. Bien que développée initialement sur une section de tissu cérébral de rat, cette stratégie peut être appliquée à d'autres tissus, y compris du tissu tumoral. Par exemple, dans l'étude de l'hétérogénéité des gliomes, agrandir le tissu tumoral permettrait, tout en conservant les méthodes de protéomique conventionnelle, d'analyser des régions plus petites que celle atteignable actuellement et permettre la découverte de nouveaux biomarqueurs.

La protéomique localisée est intéressante car elle permet d'obtenir des informations sur les mécanismes physiopathologiques d'une tumeur. Au sein des gliomes, la découverte de voies de signalisation à l'intérieur de régions très localisée pourrait amener des connaissances nouvelles pour le diagnostic des patients et le choix des thérapies.

Malgré les avancées dans la compréhension de la pathologie, ces analyses ne permettent pas de suivre l'évolution de la tumeur. En effet, après chirurgie, le suivi est réalisé par un examen clinique et des examens d'imagerie. Il n'existe pas de biomarqueurs pour contrôler la réponse au traitement. En ce sens l'analyse des vésicules extracellulaires est primordiale.

Ainsi, la dernière partie de ce projet de thèse porte sur la caractérisation protéique des VEs pour le suivi du traitement.

CHAPITRE 3

Vers un suivi des patients atteints de glioblastome par biopsies liquides grâce à la caractérisation des vésicules extracellulaires du sang

I. Introduction

La surveillance des patients atteints de glioblastomes est principalement réalisée par une évaluation clinique et un suivi en imagerie par résonance magnétique (IRM) et tomographie par émission de positons (TEP) (Johnson et al., 2019). Par IRM, il est possible d'observer par exemple la présence de nouvelles lésions, cependant ces observations doivent être interprétées avec précaution. En effet un phénomène de « pseudo progression » peut apparaître, c'est-à-dire une augmentation de la masse tumorale ou l'apparition de nouvelles lésions (Zikou et al., 2018). Notamment, les traitements par immunothérapies peuvent être responsables d'une réponse inflammatoire pouvant imiter les caractéristiques radiologiques de la progression tumorale (Okada et al., 2015).

Il n'existe pas à l'heure actuelle de biomarqueurs permettant de suivre et de caractériser la réponse aux traitements, la progression tumorale ou la rechute. En ce sens la recherche de vésicules extracellulaires au sein de biopsie liquide est une approche intéressante. Les vésicules extracellulaires représentent une population hétérogène de structures membranaires dérivées des cellules. Elles sont présentes dans les fluides biologiques comme le sang ou le liquide céphalorachidien (Raposo and Stoorvogel, 2013). Elles sont notamment impliquées dans la communication intercellulaire, ce qui permet aux cellules d'échanger des protéines, des lipides et du matériel génétique (Minciacchi et al., 2015).

Les VEs peuvent fournir des informations importantes sur les cellules tumorales dont elles proviennent, de plus, le taux et le contenu des vésicules est notamment impacté par la thérapie. En effet, une étude récente sur des patients atteints de carcinomes épidermoïdes de la tête et du cou a montré que les caractéristiques des exosomes pouvaient discriminer les patients qui répondent ou non à la thérapie (Theodoraki et al., 2019). De même, le taux

d'expression de PD-L1 dans les exosomes provenant du sang de patients traités avec un anti-PD-1 diminue significativement en cas de réponse au traitement et, au contraire, augmente en cas de progression de la maladie (Del Re et al., 2018). Aucun changement n'est observé quand la maladie est stable.

La réponse au traitement observé via les VEs peut également servir à prédire la survie des patients. Par exemple, dans le mélanome malin cutané, les patients avec des taux élevés de miR-497-5p dans les exosomes du plasma durant un traitement MAPKi (inhibiteur de la voie MAPK) ont une survie significativement plus longue (Svedman et al., 2018).

Les VEs ont également déjà été étudiées dans le suivi des glioblastomes. Ainsi, on observe une concentration de vésicules plasmatiques plus élevée chez les patients atteints de glioblastome par rapport aux témoins sains. Après chirurgie ce taux diminue et devient comparable au taux témoins, cependant, en cas de récurrence le taux augmente à nouveau (Osti et al., 2019). Ces résultats montrent qu'il est possible d'utiliser les vésicules extracellulaires comme biomarqueur capable de décrire l'état de la tumeur à travers un échantillon de sang.

Dans ce contexte, il paraît envisageable dans le futur d'intégrer l'analyse des composants contenus dans les vésicules pour le suivi de la progression de la tumeur et de la réponse aux traitements. Notamment, il paraît intéressant de caractériser ces vésicules, en se focalisant sur leurs contenus protéiques, via extractions protéiques et analyses par spectrométrie de masse. Cette analyse pourrait, à terme, permettre d'établir un profil de vésicule extracellulaire pour chaque groupe de tumeur précédemment identifié par imagerie MALDI (cf. CHAPITRE 1).

Dans ce chapitre, des études préliminaires d'enrichissement en vésicule ont été réalisés sur des échantillons de sang de patients sains.

Méthodes d'enrichissement des VEs à partir de fluides biologiques

Dans l'objectif de caractériser les vésicules extracellulaires, et notamment leurs compositions, il est nécessaire de séparer les VEs des agrégats protéiques et lipoprotéines

contenues dans les fluides biologiques. En ce sens, les méthodes d'enrichissements doivent être fiables et des contrôles de pureté doivent être réalisés.

La Société Internationale des Vésicules Extracellulaires (*International Society of Extracellular Vesicle*, ISEV) publie régulièrement des recommandations expérimentales concernant l'isolement et l'étude des vésicules extracellulaires (Lötvald et al., 2014; Théry et al., 2018). La dernière mise à jour de ces recommandations implique notamment que la méthode d'enrichissement doit être rapportée en détail pour permettre de la reproduire et que les composants associés aux VEs nouvellement identifiés doivent être validés en montrant leur association spécifique avec les VEs (Théry et al., 2018).

L'ultracentrifugation est la méthode la plus utilisée, les VEs étant concentrées avec une force centrifuge de 100 000g pendant 1 à 2 heures (Willms et al., 2018). Cette approche, en plus d'être assez longue et de nécessiter des instruments coûteux, ne permet pas de séparer les VEs des agrégats protéiques. Ainsi, il est possible de la coupler avec d'autres techniques comme le gradient de densité qui permet d'obtenir une population plus pure (Lemaire et al., 2019; Willms et al., 2018). Les VEs sont séparées dans un gradient d'une solution, iodixanol ou saccharose par exemple, dont la densité augmente du sommet vers le fond du tube. Si cette approche permet d'obtenir une population de VEs pure, elle reste longue et fastidieuse.

Deux méthodes d'enrichissement seront évaluées au sein de ce chapitre afin d'évaluer l'enrichissement en VEs à partir d'échantillons de plasma. Dans un premier temps, un kit commercial, *Exosomes Precipitation Solution (Serum/Plasma)* provenant de Macherey Nagel a été utilisé. Ce kit repose sur une co-précipitation avec un polymère, la solubilité des VEs est réduite par l'ajout d'un réactif par conséquent, elles peuvent être précipitées par centrifugation à faible vitesse. La deuxième méthode, la méthode PROSPR (PROtein Organic Solvent PREcipitation), se base sur la précipitation des protéines contaminantes par l'utilisation d'un solvant organique (Gallart-Palau et al., 2015). Après ajout d'un volume d'acétone en excès, les protéines sont centrifugées à faible vitesse. La population de VEs est alors contenue dans le surnageant. Après enrichissement via les deux méthodes, le contenu protéique des vésicules sera analysé et les résultats des deux stratégies seront évalués.

Enrichissement des VEs à partir de tissus

En parallèle de ces analyses et dans le but de démontrer que les VEs isolées à partir du plasma sont réellement issues de la tumeur, des enrichissements à partir de tissu seront réalisés (Gallart-Palau et al., 2016). Les vésicules contenues dans les interstices cellulaires d'un cerveau de rat seront enrichies grâce au protocole PROSPR afin d'évaluer la possibilité d'isoler des VEs à partir d'échantillons tumoraux.

Isoler des vésicules directement à partir de tumeur permet de se situer au plus près des cellules tumorales et ainsi d'étudier spécifiquement les VEs provenant de ces cellules. Les vésicules extracellulaires sont des médiateurs clés de la communication intercellulaire. En ce sens, étudier la signalisation basée sur ces vésicules permettrait de faire progresser considérablement notre compréhension de la tumeur. De plus, il apparaît intéressant de comparer ces vésicules aux vésicules circulants dans le plasma des patients. Le protocole d'enrichissement sera adapté afin d'évaluer la quantité minimale de tissu nécessaire pour l'enrichissement, ainsi, des coupes fines de tissu seront utilisées.

II. Matériel et Méthodes

A. Enrichissement en vésicules extracellulaires à partir de sang humain

Pour les deux méthodes, des échantillons (500 μ L) de sang humain qualifiés (négatifs pour : les anticorps anti-VIH 1+2, anti-VHC, anti-Hbc, antigène HBs, ARN du VIH1 et du VHC, l'ADN du VHB et anticorps anti-syphilis) provenant du don du sang (Établissement Français du Sang, EFS) ont été centrifugés 15 minutes à 3 000 g pour récupérer spécifiquement 300 μ L de plasma. Le premier enrichissement a été réalisé selon le protocole du kit Exosomes Précipitation Solution (Serum/Plasma) - Macherey Nagel. Ainsi, 120 μ L de solution de précipitation sont ajoutés au plasma et, après un temps d'incubation de 30 minutes à 4°C, la solution est centrifugée 5 minutes à 3000 g. Les vésicules sont contenues dans le culot. La deuxième technique d'isolement est adaptée du protocole PROSPR développé par Gallart-Palau et al., 2015. Ainsi, 1,2 mL d'acétone à -20°C sont ajoutés au plasma et la solution finale est centrifugée 5 minutes à 3 000 g. Le surnageant, contenant les vésicules est récupéré et séché à l'aide d'un concentrateur sous vide. Dans les deux cas les vésicules sont conservées à -20°C jusqu'à utilisation.

B. Enrichissement en vésicules extracellulaires à partir d'un cerveau de rat

Cette technique d'isolement est adaptée du protocole développé par Gallart-Palau et al., 2016. À partir d'un cerveau de rat frais congelé, des morceaux de différentes tailles ont été découpés au scalpel et pesés. Après un lavage au PBS, les échantillons sont placés dans des tubes contenant des billes en plastique et une solution d'acétate d'ammonium (AA, 100mM) avec inhibiteur de protéases (IP, 10x). Les tubes sont centrifugés à une vitesse de 15 000 x g à 4°C pendant 10 min. Les surnageants, contenant les vésicules, sont récupérés et les culots sont remis en suspension dans une solution d'AA+IP pour une seconde centrifugation à dans les mêmes conditions. Les surnageants sont regroupés et une solution d'acétone à -20°C est ajouté en excès (4 fois le volume). La solution finale est centrifugée à 5 000 x g à 4°C pendant 3 min. Enfin, les surnageants contenant les vésicules sont collectés, séchés par un concentrateur sous vide et conservés à -20°C jusqu'à utilisation.

C. Analyse NTA (Nanoparticle Tracking Analysis)

L'analyse NTA permet de visualiser des nanoparticules, de mesurer leurs tailles et leurs concentrations dans un échantillon donné. Les vésicules obtenues dans le sang et les tissus ont été analysées avec un analyseur de particules NS300 (Malvern Instruments Ltd). Les vésicules issues des échantillons sanguins et de tissus ont été reprises dans 150µL de PBS filtré sur filtre seringue 0,22µm. Les échantillons ont été dilués au 1/100^{ème} puis injectés dans la chambre d'analyse grâce à un pousse seringue réglé à un débit constant. L'échantillon circule dans une chambre traversée par un faisceau lumineux et une caméra enregistre le trajet et la taille de chaque particule grâce aux mouvements browniens (Figure 3-1). Pour chaque échantillon, cinq enregistrements d'une minute ont été réalisés, avec un niveau de caméra (luminosité) fixé à 7. Le seuil de détection a été fixé à 3 pour chaque enregistrement et le focus de la caméra a été réglé pour que les particules apparaissent distinctement sous forme de cercles clairs. Les autres paramètres ont été réglés par défaut. Les particules ont été analysées (tailles médianes et concentrations) par le logiciel NanoSight version 3.2.

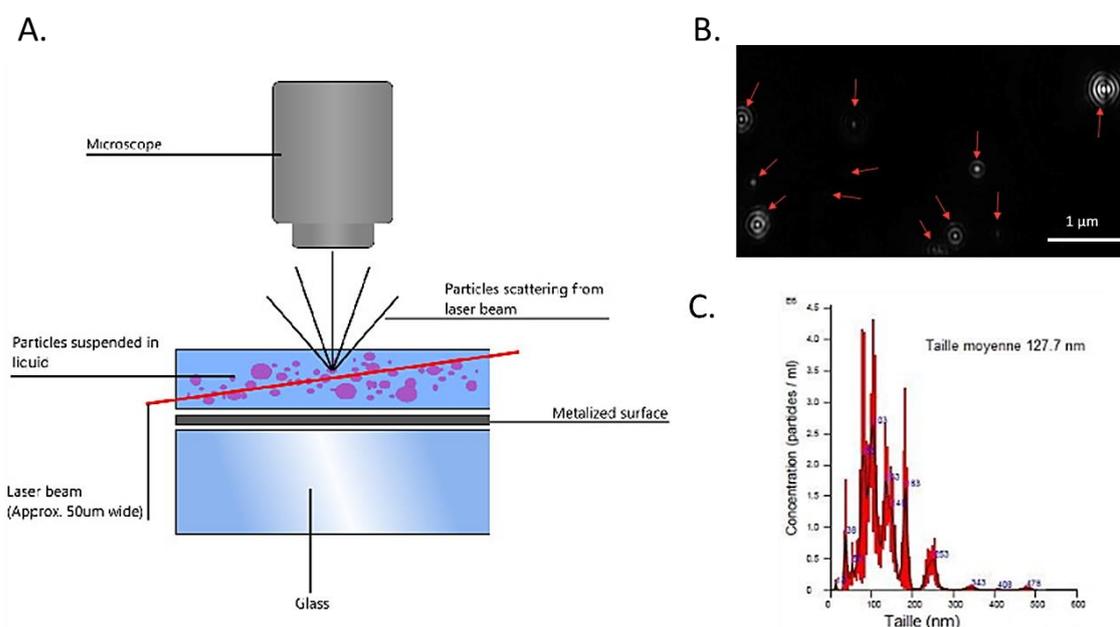


Figure 3-1 Analyse des vésicules extracellulaires par NTA (Nanoparticle Tracking Analysis)

A. Les particules en suspension dans la chambre diffusent la lumière du laser de telle sorte qu'elles peuvent être facilement visualisées avec un microscope équipé d'une caméra. Une vidéo des particules est enregistrée et, grâce aux mouvements browniens, le logiciel repère les particules individuellement et en calcule le diamètre. B. Particules de l'échantillon observables à la caméra. C. Les informations sont par la suite utilisées par le logiciel pour déterminer la concentration (particule/mL) et la taille (nm) des vésicules. A. www.malvernanalytical.com

D. Analyse protéomique

Les vésicules ont été re suspendues dans 100µL de PBS et ensuite soumises à sonication (3 x 30 secondes) pour briser les membranes lipidiques et ainsi extraire les protéines. Le processus de réduction est réalisé par ajout de 50 µL de DTT (10 mM dans NH₄HCO₃ 50 mM) 40 min à 56 ° C. Pour l'alkylation, on réalise une incubation avec 50 µL d'IAA (55 mM dans NH₄HCO₃ 50 mM) 40 min à température ambiante dans l'obscurité. Enfin, 25 µL de trypsine (40 µg / mL dans NH₄HCO₃ 50 mM) sont ajoutés dans chaque échantillon et incubés pendant une nuit à 37 ° C. La digestion est arrêtée en ajoutant du TFA à 1 % du volume final.

E. Analyses NanoLC-MS/MS

Les échantillons sont dessalés sur colonne Ziptip C-18 (Millipore, Saint-Quentin-en-Yvelines, France) avant d'être séchés à nouveau. Les échantillons secs ont ensuite été repris avec une solution ACN / AF 0.1 % aq (2/98, v/v). L'analyse a été effectuée via une chromatographie NanoAcquity UPLC (Waters) couplée à un spectromètre de masse Q-exactive (Thermo Scientific) via une source électrospray. Une colonne de préconcentration (nanoAcquity Symmetry C18, 5 µm, 180 µm x 20 mm) et une colonne analytique (nanoAcquity BEH C18, 1,7 µm, 75 µm x 250 mm) ont été utilisées. Les peptides ont été séparés en appliquant un gradient linéaire d'acétonitrile (5%-35%) pendant 2heures à un débit de 300 nl/min. La MS/MS a été réalisée en mode d'acquisition dépendante des données pour les 10 ions les plus intenses de l'analyse MS (DDA-Top 10). L'analyse MS a été réalisée avec une gamme de masse m/z comprise entre 300 et 1600, une résolution de 70 000 FWHM, un AGC de 3e6 ions et un temps d'injection maximum de 120 ms. L'analyse MS / MS a été effectuée avec une gamme de masse située entre 200 et 2000, un AGC de 5e4 ions, un temps d'injection maximum de 60 ms et la résolution a été fixée à 17 500 FWHM.

F. Analyses des données

Les données de spectrométrie de masse ont été traitées avec le logiciel MaxQuant (Cox and Mann, 2008; Tyanova et al., 2015) (Version 1.5.3.30) utilisant le moteur de recherche Andromeda (Cox et al., 2011). Les identifications de protéines ont été réalisées par

interrogation des bases de données UniProt Homo Sapiens et Rattus Norvegicus avec comme critère d'identification 2 peptides minimum par protéine dont au moins 1 unique. La trypsine a été choisie pour le mode de digestion. L'acétylation en position N-terminale et l'oxydation des méthionines ont été choisies comme modifications variables, la carbamidométhylation a été définie comme modification fixe. Les analyses ont été réalisées avec une précision de mesure de 6 ppm en MS et 20 ppm en MS/MS. Le taux de faux positifs (FDR) a été défini à 1 % en peptide et protéine. La quantification sans marquage au préalable des protéines (*label-free quantification*) a été réalisée grâce à l'algorithme MaxLFQ intégré dans le logiciel MaxQuant avec les paramètres par défaut (Cox et al., 2014). L'analyse statistique des protéines identifiées a été réalisée grâce au logiciel Perseus (version 1.6.0.7). Tout d'abord, les résultats issus de la base de données inversée, les protéines contenant uniquement des peptides modifiés et les contaminants potentiels ont été supprimés. Des analyses statistiques ont été effectuées en utilisant un t-test avec une p-value de 0.05. Pour la représentation, un regroupement hiérarchique a été réalisé. Les différentes méthodes ont été évaluées en termes de chevauchement des identifications des protéines (diagrammes de Venn).

III. Résultats et Discussion

A. Enrichissement à partir d'échantillons sanguins

Les échantillons de sang utilisés humains ont été centrifugés 15 minutes à 3 000 g pour récupérer spécifiquement 300 µL de plasma. Deux modes d'enrichissement ont été comparés : un kit commercial et la méthode PROSPR. Les vésicules isolées dans les deux conditions ont été séchées puis re-suspendues dans 150 µL de PBS préalablement filtré (0,22 µm).

Pour analyser la concentration et la répartition des tailles des vésicules enrichies dans chaque condition, des analyses NTA sont réalisées. L'analyse NTA consiste à suivre des nanoparticules en suspension et permet de déterminer leurs tailles et leurs nombres.

Les échantillons ont préalablement été dilués au 1/100^{ème} avant analyse NTA. Les résultats de l'analyse (Tableau 3-1, Figure 3-2), montrent que les deux méthodes sont comparables en termes de quantité de vésicules extracellulaires enrichies ($1,8 \times 10^{10}$ particules/mL). Les contrôles PBS indiquent une concentration négligeable de particules ($6,5 \times 10^6$ particules/mL).

Tableau 3-1 Concentrations et tailles des particules analysées à partir d'échantillons de sang

	Nombre de particules /mL	Taille la plus représentée (nm)
<i>blanc PBS</i>	6,48E+06	67,8
<i>KIT</i>	1,83E+10	101,7
<i>PROSPR</i>	1,78E+10	125,7

Les vésicules extracellulaires isolées grâce au kit apparaissent plus petites et de tailles plus homogènes (50-200 nm) que celles isolées via la méthode PROSPR (100-350 nm) (Figure 3-2). Une partie des particules isolées avec la méthode PROSPR est de taille supérieure à 300 nm. En effet, certains pics représentant des tailles allant de 250 à 500 nm sont visibles mais la concentration ne dépasse pas 2×10^6 particules/mL. La taille élevée de ces particules peut indiquer la présence de liposomes ou de corps apoptotiques.

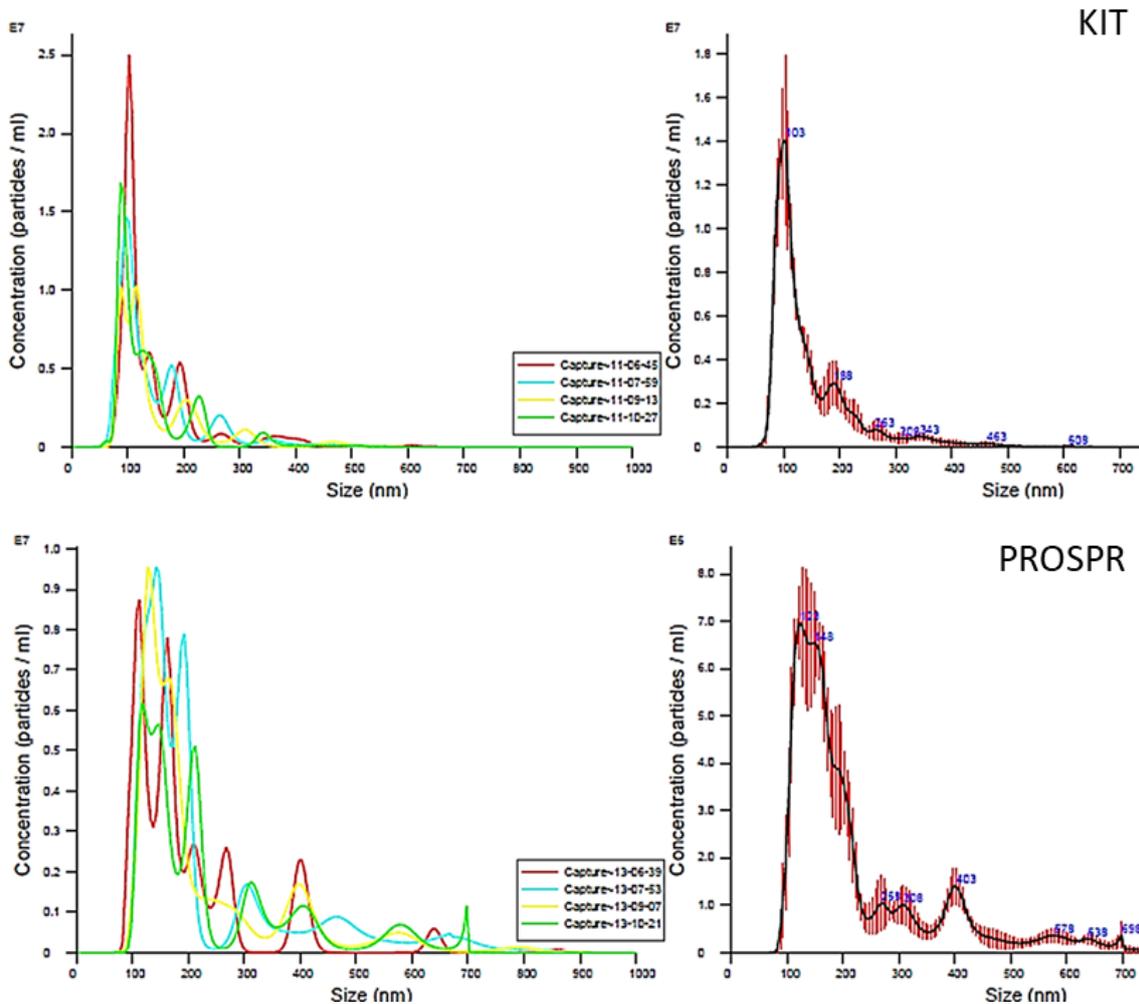


Figure 3-2 Analyses NTA de vésicules extracellulaires issus de plasma

Analyses NTA de plasmas enrichis en vésicules extracellulaires via les 2 méthodes. En haut, distribution des tailles et concentrations des VEs selon le kit et en bas, selon la méthode PROSPR. Pour chaque analyse, 5 captures consécutives d'une minute sont réalisées. A droite, chaque courbe correspond à un enregistrement et le graphique de gauche correspond à la moyenne de ces analyses.

B. Analyse protéomique des vésicules du sang

Par la suite, des analyses protéomiques ont été réalisées afin d'observer, en premier lieu, s'il était possible d'identifier des protéines à partir de ces enrichissements. Un total de 205 protéines a été identifié à partir des vésicules enrichies avec les deux méthodes (Figure 3-3). Presque trois fois plus de protéines sont identifiées avec le kit commercial (204) qu'avec la méthode PROSPR (75). On observe 74 (36 %) protéines communes entre les deux conditions, une seule protéine (Beta-Ala-His dipeptidase) est uniquement retrouvée avec la méthode PROSPR.

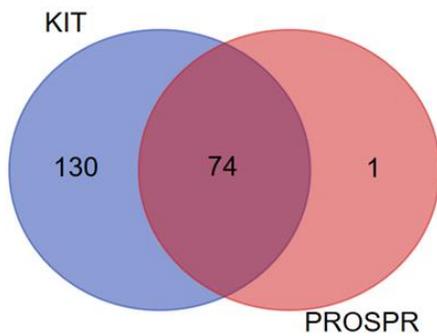


Figure 3-3 Nombre de protéines identifiées selon les deux méthodes d'enrichissement
Analyse réalisée sur quatre échantillons de plasma.

Dans un second temps, la pureté des vésicules a été investigué en comparant les protéines identifiées après enrichissement aux protéines issues des échantillons contrôles.

Pour le kit, les protéines contenues dans le surnageant, après précipitation des vésicules, ont été analysées. Ainsi, plus de la moitié (62 %) des protéines identifiées dans le contrôle sont également identifiées dans l'enrichissement en vésicules (Figure 3-4). Pour la méthode PROSPR, les protéines précipitées par l'acétone ont été comparées à l'échantillon de vésicules. Dans cette configuration uniquement 2 protéines sont retrouvées uniquement dans le surnageant contenant les vésicules (Figure 3-4). Dans les deux méthodes d'enrichissement, une faible proportion de protéines est spécifique des échantillons de vésicules, une grande proportion (47 %) est retrouvée dans les deux échantillons.

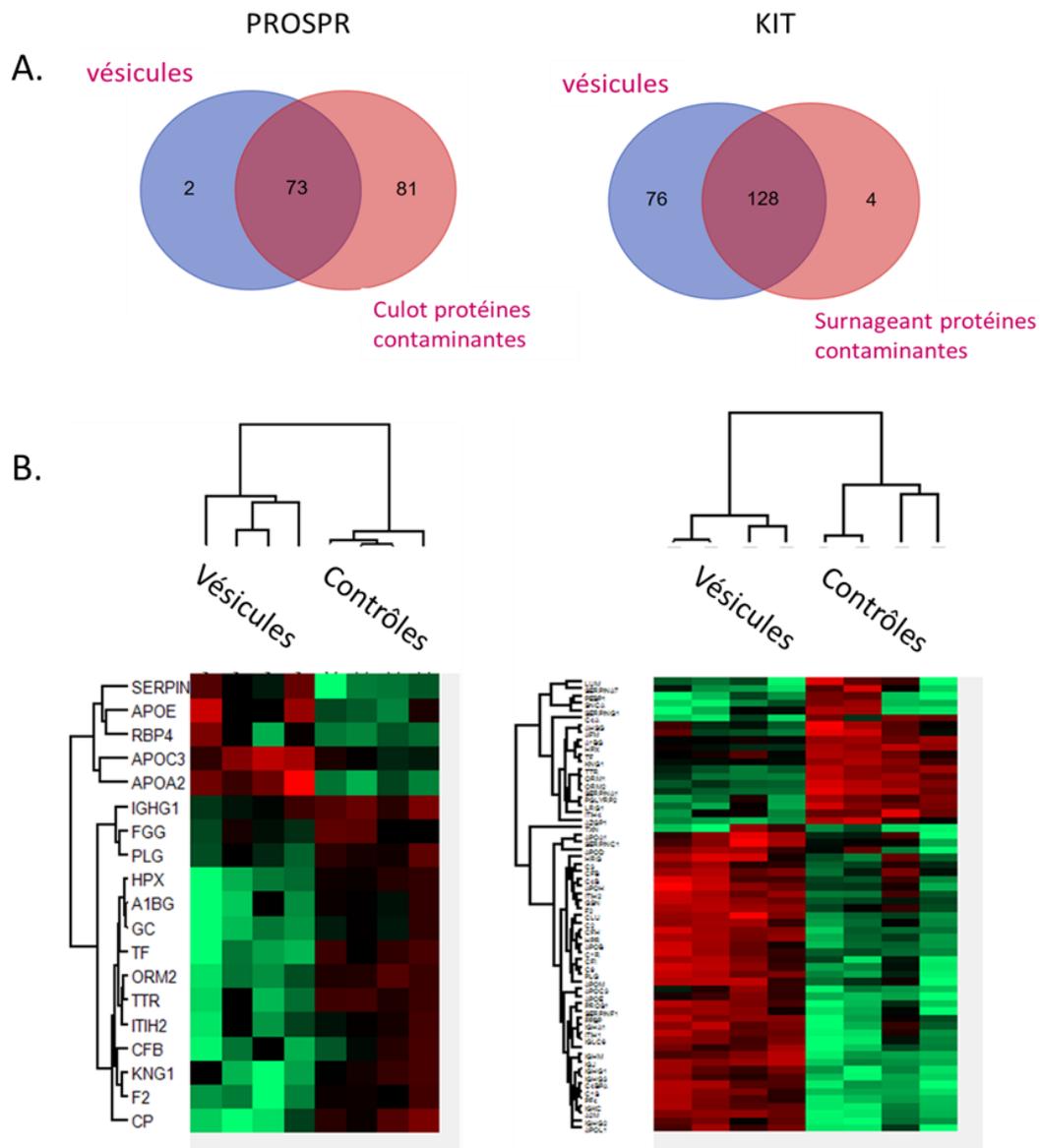


Figure 3-4 Comparaison des protéines identifiées dans l'enrichissement en vésicules et dans le contrôle pour chaque méthode

A. Nombre de protéines identifiées dans les échantillons vésicules et contrôle pour chaque méthode d'enrichissement. B. Analyse statistique (t-test p -value 0.05) révèle des clusters de protéines surexprimées dans les échantillons enrichis en vésicules et les contrôles.

On observe des protéines surexprimées spécifiquement dans les échantillons de vésicules (Figure 3-4). En effet, 5 protéines sont spécifiquement surexprimées dans l'échantillon enrichi en vésicules par la méthode PROSPR. En utilisant le kit commercial, 41 protéines sont surexprimées dans l'échantillon enrichi (Figure 3-4).

C. Comparaison aux protéines identifiées dans la littérature

Les protéines surexprimées dans chaque condition ont été comparées aux TOP 100 des protéines identifiées dans les bases de données Exocarta, EVpedia et Vesiclepedia qui recensent l'ensemble du protéome connu des vésicules extracellulaires (Figure 3-5). On ne retrouve que 3 protéines identifiées dans ces listes. La protéine alpha-2-macroglobulin (A2M) est identifiée à la fois dans l'échantillon vésicule enrichi à partir du kit et dans les bases de données EVpedia et Vesiclepedia et les protéines complément C3 et gelsolin sont identifiées dans les vésicules enrichies avec la méthode PROSPR et dans la base Vesiclepedia.

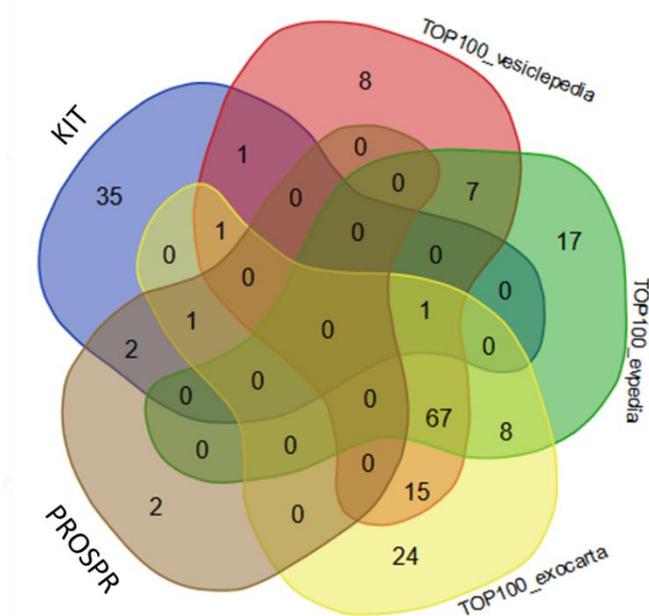


Figure 3-5 Comparaison des protéines surexprimées dans les échantillons de vésicules avec les protéines communément identifiées dans la littérature

Les TOP 100 de protéines des trois bases de données Vesiclepedia, EVpedia et Exocarta montre uniquement 3 protéines identifiées dans les échantillons de vésicules.

Le faible nombre de protéines identifiées couramment dans la littérature peut indiquer que dans les deux méthodes, l'enrichissement n'est pas concluant. D'autres méthodes de validation seront nécessaires pour confirmer la présence de vésicules. Notamment, des analyses de microscopie électronique sont envisagées pour s'assurer de la présence de vésicules via l'observation de la membrane lipidique.

En parallèle de ces analyses et dans le but de démontrer que les VEs isolées à partir du plasma sont réellement issues de la tumeur, des enrichissements à partir de tissu ont été

réalisés (Gallart-Palau et al., 2016). Les vésicules contenues dans les interstices cellulaires d'un cerveau de rat seront enrichies grâce au protocole PROSPR afin d'évaluer la possibilité d'isoler des VEs à partir d'échantillons tumoraux.

D. Enrichissement à partir de tissus

Le protocole PROSPR a été utilisé, ce protocole a été initialement développé pour enrichir des tissus de 40mg (Gallart-Palau et al., 2016). Les protéines contaminantes ont été précipitées par de l'acétone comme vu précédemment. Les vésicules contenues dans le surnageant sont séchées et re suspendue dans du PBS avant d'être analysées (Figure 3-6).

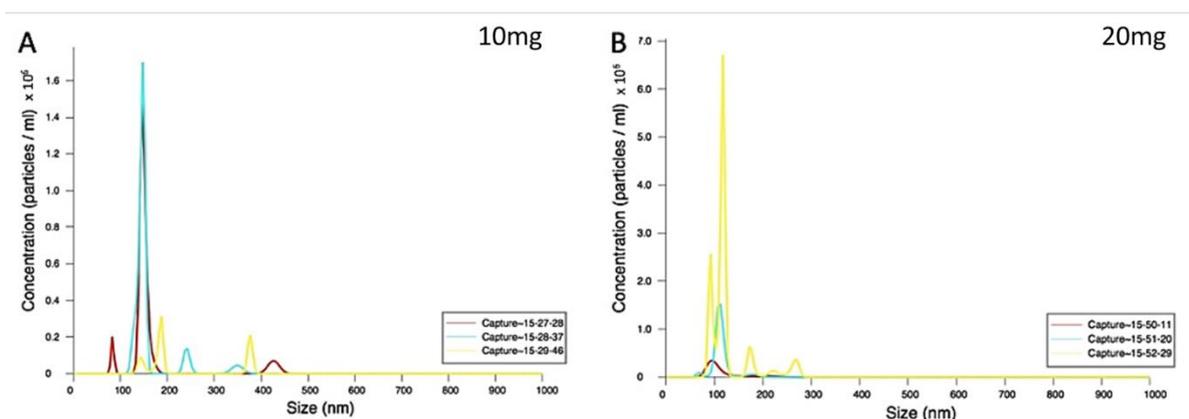


Figure 3-6 Caractérisation NTA des nanoparticules enrichies à partir de cerveau de rat
Distribution de la taille des nanoparticules selon la taille des échantillons de tissus

Dans notre cas, l'analyse a été réalisée sur des échantillons de tissu cérébral de rat de 10 et 20 mg. Aucune particule n'a été identifiée dans le contrôle PBS. Les tailles des particules les plus représentées sont centrées entre 90 et 150 nm (Figure 3-6). On observe une concentration plus élevée de vésicule dans l'échantillon de 20 mg comparé à l'échantillon de 10 (Tableau 3-2).

Tableau 3-2 Concentrations et tailles des particules analysées à partir de tissus

	Nombre de particules /mL	Taille la plus représentée (nm)
10mg	4,65E+07	147,5
20mg	1,15E+08	116,5

Différents enrichissements ont été réalisés afin d'évaluer la quantité minimale de tissu nécessaire. Des coupes fines de tissu de 20µm d'épaisseur (1,5 cm² ; 2,1 mg) ont été utilisées. Dans un premier temps, les analyses NTA ont montré que des particules allant de 90nm à 300nm ont pu être enrichies dans cette coupe (Figure 3-7). La concentration de ces nanoparticules est de 1,88 x 10⁸/mL et la taille la plus représentée est 158 nm (Figure 3-7).

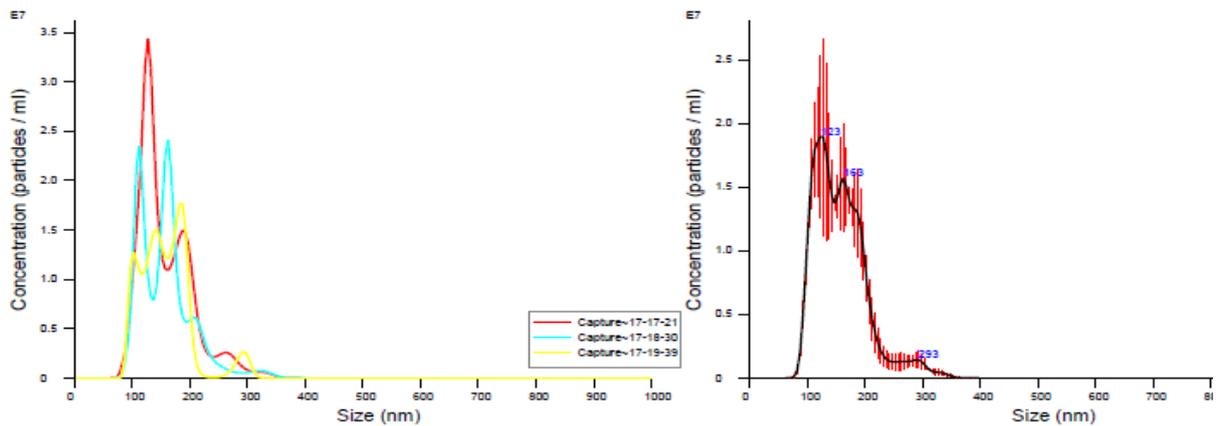


Figure 3-7 Analyse NTA des nanoparticules enrichies à partir de coupes fines de cerveau de rat
Distribution de la taille des nanoparticules selon la taille des échantillons de tissus

Cette méthode est encore en phase d'optimisation mais les premiers résultats sont encourageants. L'analyse NTA permet d'observer des nanoparticules et non spécifiquement des vésicules extracellulaires. Ainsi il est primordial de confirmer la présence de vésicules grâce à différentes méthodes (microscopie électronique, analyse protéomique).

IV. Conclusion et Perspectives

Ces dernières années, les vésicules extracellulaires ont pris une grande place dans la recherche en neurosciences au vu de leur rôle dans la physiopathologie de plusieurs maladies impliquant le système nerveux central. Au vu de leur richesse en information (protéines, lipides, acides nucléiques) il devient pertinent de les exploiter pour étudier la progression d'une tumeur. À ce jour, deux méthodes d'enrichissement en vésicules extracellulaires ont été testées sur des échantillons de sang de patients sains.

Les différents protocoles nous ont permis d'enrichir des particules à partir d'échantillon de plasma humain. Ces particules ont été observées et quantifiées par analyse NTA, une quantité similaire de particules a été identifiée via les deux méthodes, le kit permettant une isolation plus homogène (50-200 nm) que la méthode PROSPR (100-350 nm). L'utilisation de kit est assez coûteuse et n'est pas envisageable pour une utilisation à grande échelle. De plus, la présence de polymère peut être un facteur limitant pour l'analyse par spectrométrie de masse. La méthode PROSPR est simple d'utilisation cependant il est intéressant de noter que l'acétone peut interférer avec les propriétés fonctionnelles des membranes des vésicules et de ce fait induire la fusion des VEs en structures multivésiculaires (Gámez-Valero et al., 2016).

Finalement, si des nanoparticules ont pu être identifiées dans les deux cas, nous n'avons pas pu confirmer la présence de vésicules extracellulaires. En effet, très peu de protéines spécifiques des vésicules ont été identifiées. Le microscope électronique est le moyen le plus fiable pour l'identification de vésicules extracellulaires, en effet, de cette façon il est possible d'observer la double membrane caractéristique des vésicules et donc de les distinguer d'éventuels agrégats protéiques.

En parallèle, une méthode de séparation par exclusion stérique a été développée au sein du laboratoire (Lemaire et al., 2020) (Figure 3-8). La chromatographie d'exclusion stérique (*Size Exclusion Chromatography*, SEC) permet la séparation des vésicules en fonction de leurs tailles grâce à des billes poreuses dont les pores ont une taille spécifique (Figure 3-8).

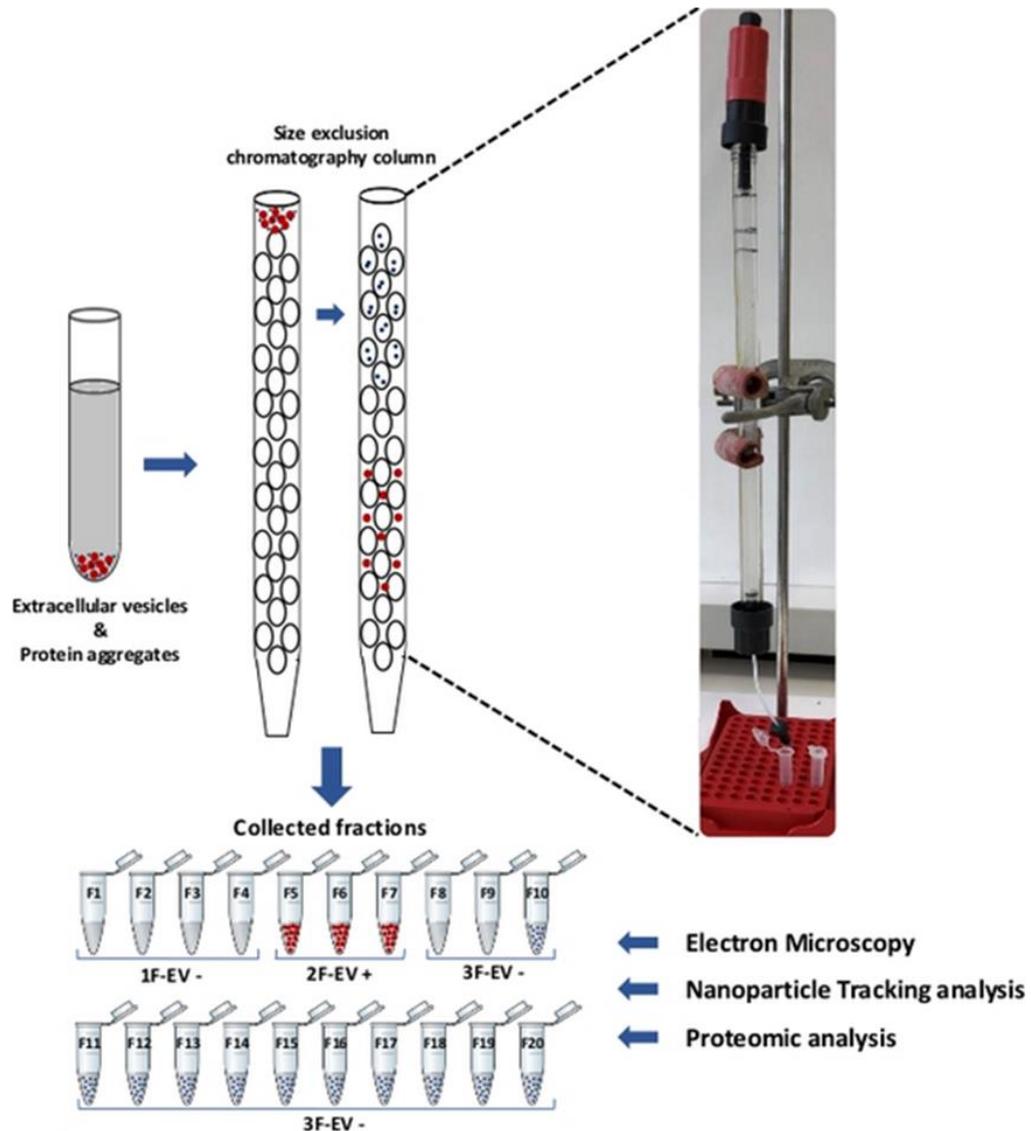


Figure 3-8 Enrichissement et collection de VEs par chromatographie d'exclusion stérique

Les VE sont dans un premier temps, enrichies par ultracentrifugation (UC). Le culot UC contenant les VEs est chargé sur une colonne et séparé par chromatographie d'exclusion stérique (SEC) dans 20 fractions éluées différentes. Comme révélé dans les étapes suivantes (microscopie électronique, analyse de suivi des nanoparticules et analyse protéomique), les fractions SEC ont été regroupées et organisées en 1F-EV- (ne contenant pas de vésicules), 2F-EV + (enrichis en vésicules) et 3F-EV- (ne contenant pas de vésicules mais des protéines solubles). Adapté de (Lemaire et al., 2020)

Appliquée à des VEs issues de cultures cellulaires, cette méthode a permis d'isoler $1,6 \times 10^9$ particules / mL dans les fractions enrichies en vésicules. De plus une analyse protéomique a été réalisée par la suite pour confirmer la présence de vésicules. Près de 700 protéines ont été identifiées, dont 536 dans les fractions contenant les vésicules (Lemaire et

al., 2020). Par ailleurs, après comparaison au TOP 100 de la base de données ExoCarta, 86 protéines ont été identifiées.

Cette approche apparaît donc assez robuste pour l'enrichissement et l'analyse protéomiques des VEs et sera évaluée sur des échantillons de plasma. Finalement, si les résultats sont concluants, la méthode pourra être appliquée à l'analyse du sang des patients de la cohorte Gliomic.

Pour chaque patient de l'étude Gliomic, du sang a été prélevé au moment de l'hospitalisation pour la chirurgie. L'analyse des composants contenus dans les vésicules de ces échantillons sanguins permettra d'établir un profil diagnostique/pronostique des patients et rendra possible le suivi de la progression de la tumeur et de la réponse aux traitements. En effet, il serait intéressant d'associer un profil de vésicule extracellulaire à chaque groupe de tumeur. Aussi, il est possible d'intégrer l'analyse des protéines alternatives à la caractérisation des VEs (Murgoci et al., 2020). La fonction de ces protéines reste à explorer cependant nous avons précédemment montré leurs implications dans le groupement et la survie des glioblastomes (cf. CHAPITRE 1).

Pour aller plus loin et afin d'identifier des biomarqueurs des glioblastomes, il est nécessaire de pouvoir réaliser un diagnostic plus précoce et sans devoir nécessairement accéder au tissu tumoral. De plus, les données issues de la caractérisation des vésicules pourraient être corrélées avec le pronostic des patients. En comparant les VEs des patients traités en première intention et ceux présentant une résistance, il sera par la suite possible d'évaluer la résistance ou non aux traitements à partir d'une prise de sang.

De plus, le protocole PROSPR nous a permis d'enrichir des particules à partir d'un cerveau de rat. Les premiers résultats sont encourageants mais il sera par la suite nécessaire, comme pour les enrichissements à partir de sang, de confirmer la présence de vésicules avec des marqueurs spécifiques de VEs et des analyses de microscopie. Aussi, il serait intéressant d'utiliser la chromatographie d'exclusion stérique pour améliorer la qualité d'enrichissement.

L'enrichissement à partir de tissu permettra, à termes, d'identifier des vésicules provenant de la tumeur plus facilement. En effet les vésicules contenues dans le sang peuvent avoir plusieurs origines et les vésicules tumorales pourraient être potentiellement diluées. Analyser les vésicules à partir de la tumeur permet de se placer au plus proche des cellules tumorales. De cette façon il sera possible d'être sûr que les vésicules retrouvées dans le sang proviennent directement de la tumeur et non d'un mécanisme de réponse ou d'autres cellules. Dans le futur, l'analyse des vésicules contenues dans la masse tumorale apportera de nouvelles connaissances concernant la communication cellulaire au sein de la tumeur. A partir de ces connaissances, il est envisageable de cibler les vésicules pour empêcher la communication entre les cellules tumorales et le microenvironnement. Il est par exemple possible de limiter leurs assimilations par les cellules réceptrices ou encore d'impacter leurs biogénèse (Escrevente et al., 2011; Macia et al., 2006; Peinado et al., 2012). Aussi, l'étude des vésicules extracellulaires permet de les utiliser comme vecteur de traitement (Hadla et al., 2016)

L'objectif *in fine* de ce projet est d'ouvrir la porte vers un suivi de la progression tumorale et des traitements par biopsie liquide en vue de la mise en place d'une médecine personnalisée.

CONCLUSION et PERSPECTIVES

Les gliomes représentent une pathologie complexe et la prise en compte de nouveaux critères moléculaires pourrait améliorer leur classification. Ainsi, ce projet porte sur l'élaboration d'une nouvelle classification des gliomes basée sur des techniques de protéomique. L'utilisation de l'imagerie par spectrométrie de masse MALDI est déterminante dans l'étude de ces tumeurs puisqu'elle a permis de rendre compte de l'hétérogénéité tumorale et d'établir également un regroupement entre les patients. Cette technique d'analyse non supervisée a ici été couplée à une analyse protéomique prenant en considération la localisation des régions sur les coupes dans le but de concevoir une classification de ces tumeurs.

L'étude Gliomic, sujet principal de ce manuscrit, se compose d'une cohorte de cinquante patients présentant un glioblastome. L'analyse par imagerie MALDI et la segmentation spatiale individuelle de ces cas ont dans un premier temps montré une réelle hétérogénéité inter- et intratumoral. En effet, une faible concordance a été observée entre l'imagerie et l'annotation anatomopathologique. La plupart des tumeurs considérées comme homogènes d'un point de vue histologique se sont révélées complexes d'un point de vue moléculaire. Par la suite, une segmentation globale de l'ensemble des tumeurs a permis de comparer les cas et de mettre en évidence 3 regroupements de patients présentant des signatures moléculaires spécifiques. Le regroupement des cas a été validé avec la technologie SpiderMass qui permet des analyses sans préparation de l'échantillon (Ogrinc et al., 2019). Un taux de 99% de validation de la classification a ainsi été observé. Au sein de chaque coupe de tissu, les analyses de microprotéomique localisées ont permis d'identifier au total près de 5 000 protéines. Parmi ces protéines, 1231 ont montré une expression différentielle dans les trois groupes identifiés par imagerie MS. L'analyse de la fonction biologique et notamment des voies de signalisation a permis de mettre en évidence la présence de processus biologiques spécifiques pour chacun des groupes.

Le groupe A (n=13) correspond à un profil neurologique avec notamment des protéines impliquées dans le développement du système nerveux, des marqueurs de différenciation et

des signatures de cellules souches et progénitrices. Le groupe B (n=9) correspond à une activation du système immunitaire et une dérégulation de la cascade de la coagulation. On retrouve notamment dans ce groupe des protéines liées au statut immunitaire avec une infiltration de macrophages. Le dernier groupe, le groupe C (n=23), présente une forte implication virale, avec des protéines impliquées dans la réponse immunitaire antivirale et l'infection virale.

De plus, l'analyse des protéines alternatives a également révélé des différences entre les trois groupes. Les protéines alternatives sont des protéines issues de cadres de lectures alternatifs (Vanderperre et al., 2013). La fonction de ces protéines reste à étudier. Elles peuvent interagir avec des protéines de références et être impliquées dans des voies de signalisations particulières. Il est intéressant de noter que plus de 50% des AltProt identifiées proviennent de la traduction d'ARN non codants, révélant de nouvelles perspectives sur le rôle biologique de ces ARNs.

La corrélation de ces observations avec les données cliniques des patients de l'étude a permis d'avoir une vue d'ensemble de la pathologie. L'analyse des données de survie a mis en évidence 6 protéines statistiquement associées au pronostic. Sur la base de l'expression relative d'un ensemble de 6 protéines, 3 groupes de patients ont été déterminés : le groupe 3 (n=14) observe une survie significativement plus longue, le groupe 2 (n=15) possède la survie la plus faible tandis que le groupe 1 (n=21) présente une durée de survie intermédiaire. Si aucune corrélation directe n'est observée entre le regroupement réalisé par imagerie MALDI et celui basé sur les données de survie, il est néanmoins intéressant de noter que les protéines surexprimées dans le groupe 2 sont également retrouvées dans le groupe B.

Chaque groupe met en évidence des voies de signalisation très spécifiques. Il est nécessaire de prendre en compte ces différences pour cibler les traitements. L'approche protéomique pourrait permettre le développement de stratégies thérapeutiques personnalisées. Par exemple, les protéines surexprimées dans le groupe B sont majoritairement impliquées dans la réponse immunitaire, ainsi, il paraît adapté de proposer des stratégies d'immunothérapies aux patients de ce groupe. En ce sens, des stratégies développées au sein du laboratoire se basent sur la réactivation des macrophages pour

empêcher la progression tumorale et éradiquer la tumeur grâce à des inhibiteurs de protéines convertases (Duhamel et al., 2018; Rose et al., 2020).

L'imagerie MALDI a permis de faire ressortir l'hétérogénéité de ces tumeurs. Si des points communs ont permis de constituer des groupes de patients, aucune tumeur n'est identique. Dans notre cohorte, 5 cas n'ont pas pu être classés, car présentant la présence équivalente des trois zones. L'analyse étant réalisée sur une seule partie de la masse tumorale, il paraît intéressant d'analyser d'autres parties des tumeurs déjà étudiées afin d'apporter des informations complémentaires et de confirmer la classification.

Aussi, augmenter la cohorte de patients pourra améliorer la robustesse de l'analyse statistique et confirmer nos observations. Pour de futures analyses, il sera également nécessaire de prendre en compte le statut IDH des patients. En effet, nos analyses ont été réalisées avant les dernières recommandations CIMPACT NOW suggérant de séparer ces deux entités (Brat et al., 2020). D'après ces recommandations, les glioblastomes IDH mutés et IDH non mutés représentent des cas cliniques et génétiques distincts, les glioblastomes mutés ayant une médiane de survie supérieure. Cependant, au sein de la cohorte de l'étude Gliomic, nos premières observations ne permettent pas de différencier les tumeurs selon leurs statuts IDH.

Dans un deuxième temps, d'un point de vue technique, l'expansion de tissu adaptée à la protéomique localisée permet d'analyser de façon encore plus localisée les protéines d'un tissu.

L'expansion de tissu permet d'agrandir une coupe de tissu jusqu'à 4 fois sa taille initiale. À partir du tissu agrandi, il est possible de prélever manuellement des zones précises au sein du tissu. La protéinase K est le plus souvent utilisée comme agent d'homogénéisation avant expansion (Tillberg et al., 2016b). Cette protéase non spécifique n'est cependant pas adaptée à l'analyse protéomique à grande échelle, en effet, elle génère de petits peptides qui vont migrer hors du gel. Le SDS a été choisi comme alternative à la protéinase K. En tant que détergent, le SDS dénature les protéines en perturbant leurs structures tridimensionnelles. De plus, il permet d'éliminer les lipides du tissu qui pourraient masquer le signal protéique.

L'utilisation du SDS comme agent d'homogénéisation a rendu possible l'analyse protéomique robuste des zones étudiées. Cette stratégie permet l'identification de près de 655 protéines dans une zone définie de 460µm ce qui correspond à environ 940 cellules. L'identification de protéines à partir de tissu étendu montre une bonne reproductibilité à la fois qualitative et quantitative.

La technique est en cours d'optimisation pour être utilisée sur des tissus frais. En effet, la fixation des tissus FFPE induit une moins bonne accessibilité aux protéines. Sur tissus frais, un total de 1 000 protéines a pu être identifié, cependant une importante variation du nombre de protéines ainsi qu'une migration importante est observable. La transparence du tissu après expansion rend difficile le prélèvement d'une région précise du tissu. L'utilisation de colorants compatibles avec la spectrométrie de masse sera indispensable à la réalisation de la méthode.

Finalement, augmenter le facteur d'expansion permettrait d'analyser des régions de quelques dizaines de micromètres. L'optimisation du protocole pour agrandir un tissu entre 10 à 100 fois pourrait permettre son utilisation pour l'analyse protéomique de cellule unique.

Les développements ont été réalisés sur des coupes de cerveau de rat. À terme, cette stratégie pourra être appliquée à d'autres tissus, y compris du tissu tumoral. Par exemple, dans l'étude de l'hétérogénéité des gliomes, agrandir le tissu tumoral permettrait, tout en conservant les méthodes de protéomique conventionnelle, d'analyser des régions plus petites que celles atteignables actuellement. Finalement, cette stratégie pourrait améliorer notre compréhension de la pathologie et permettre la découverte de nouveaux biomarqueurs.

Finalement, pour aller plus loin et afin d'identifier des biomarqueurs de glioblastomes, le développement des biopsies liquides et des biomarqueurs pourraient aider au diagnostic, au suivi et au dépistage des patients à risque .

En ce sens, la caractérisation de vésicules extracellulaires à partir de biopsie liquide est déterminante. Les mises au point développées pourront permettre d'analyser les vésicules extracellulaires issues du sang des patients de la cohorte Gliomic. L'analyse de ces VEs permettra de révéler si des biomarqueurs peuvent être identifiés.

Finalement, il est possible d'envisager de pouvoir associer un profil de vésicule à chacun des trois groupes de patients identifiés par imagerie MALDI. Aussi, la caractérisation de ces VEs pourra permettre de suivre la progression tumorale et la réponse au traitement. Par ailleurs, elles pourront définir la réponse aux traitements.

De plus, l'analyse des vésicules extracellulaires contenues dans la masse tumorale permettrait de se situer au plus près des cellules tumorales. Des optimisations concernant l'enrichissement sur tissus doivent encore être effectuées cependant les premiers résultats obtenus sur cerveau de rat sont encourageants.

BIBLIOGRAPHIE

Abels, E.R., and Breakefield, X.O. (2016). Introduction to Extracellular Vesicles: Biogenesis, RNA Cargo Selection, Content, Release, and Uptake. *Cell. Mol. Neurobiol.* *36*, 301–312.

Akers, J.C., Ramakrishnan, V., Kim, R., Skog, J., Nakano, I., Pingle, S., Kalinina, J., Hua, W., Kesari, S., Mao, Y., et al. (2013). miR-21 in the Extracellular Vesicles (EVs) of Cerebrospinal Fluid (CSF): A Platform for Glioblastoma Biomarker Development. *PLoS One* *8*.

Akhtar, S., Vranic, S., Cyprian, F.S., and Al Moustafa, A.-E. (2018). Epstein–Barr Virus in Gliomas: Cause, Association, or Artifact? *Front. Oncol.* *8*, 123.

Al-Nedawi, K., Meehan, B., Micallef, J., Lhotak, V., May, L., Guha, A., and Rak, J. (2008). Intercellular transfer of the oncogenic receptor EGFRvIII by microvesicles derived from tumour cells. *Nat. Cell Biol.* *10*, 619–624.

Albers, A.C., and Gutman, D.H. (2009). Gliomas in patients with neurofibromatosis type. *Expert Rev. Neurother.* *9*, 535–539.

Alexander, B.M., and Cloughesy, T.F. (2017). Adult glioblastoma. *J. Clin. Oncol.* *35*, 2402–2409.

Alexandrov, T., Becker, M., Deininger, S.-O., Ernst, G., Wehder, L., Grasmair, M., von Eggeling, F., Thiele, H., and Maass, P. (2010). Spatial Segmentation of Imaging Mass Spectrometry Data with Edge-Preserving Image Denoising and Clustering. *J. Proteome Res.* *9*, 6535–6546.

Arita, H., Narita, Y., Fukushima, S., Tateishi, K., Matsushita, Y., Yoshida, A., Miyakita, Y., Ohno, M., Collins, V.P., Kawahara, N., et al. (2013). Upregulating mutations in the TERT promoter commonly occur in adult malignant gliomas and are strongly associated with total 1p19q loss. *Acta Neuropathol.* *126*, 267–276.

Asano, S.M., Gao, R., Wassie, A.T., Tillberg, P.W., Chen, F., and Boyden, E.S. (2018). Expansion Microscopy: Protocols for Imaging Proteins and RNA in Cells and Tissues. *Curr. Protoc. Cell Biol.* *80*, e56.

Bagley, S.J., Desai, A.S., Linette, G.P., June, C.H., and O’Rourke, D.M. (2018). CAR T-cell therapy for glioblastoma: Recent clinical advances and future challenges. *Neuro. Oncol.*

Bamburg, J.R., Bernstein, B.W., Davis, R., Flynn, K.C., Goldsbury, C., Jensen, J.R., Maloney, M.T., Marsden, I.T., Minamide, L.S., Pak, C.W., et al. (2010). ADF/Cofilin-Actin Rods in Neurodegenerative Diseases. *Curr. Alzheimer Res.*

Banks, R.E., Dunn, M.J., Forbes, M.A., Stanley, A., Pappin, D., Naven, T., Gough, M., Harnden, P., and Selby, P.J. (1999). The potential use of laser capture microdissection to selectively obtain distinct populations of cells for proteomic analysis — Preliminary findings. *20*, 689–700.

Becker, A., Thakur, B.K., Weiss, J.M., Kim, H.S., Peinado, H., and Lyden, D. (2016). Extracellular Vesicles in Cancer: Cell-to-Cell Mediators of Metastasis. *Cancer Cell* 30, 836–848.

Beier, C.P., Rasmussen, T., Dahlrot, R.H., Tenstad, H.B., Aarø, J.S., Sørensen, M.F., Heimisdóttir, S.B., Sørensen, M.D., Svenningsen, P., Riemenschneider, M.J., et al. (2018). Aberrant neuronal differentiation is common in glioma but is associated neither with epileptic seizures nor with better survival. *Sci. Rep.*

van den Bent, M.J. (2010). Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective. *Acta Neuropathol.* 120, 297–304.

Van Den Bent, M.J., Brandes, A.A., Taphoorn, M.J.B., Kros, J.M., Kouwenhoven, M.C.M., Delattre, J.Y., Bernsen, H.J.J.A., Frenay, M., Tijssen, C.C., Grisold, W., et al. (2013). Adjuvant procarbazine, lomustine, and vincristine chemotherapy in newly diagnosed anaplastic oligodendroglioma: Long-term follow-up of EORTC brain tumor group study 26951. *J. Clin. Oncol.* 31, 344–350.

Bonnel, D., Longuespee, R., Franck, J., Roudbaraki, M., Gosset, P., Day, R., Salzet, M., and Fournier, I. (2011). Multivariate analyses for biomarkers hunting and validation through on-tissue bottom-up or in-source decay in MALDI-MSI: Application to prostate cancer. *Anal. Bioanal. Chem.* 401, 149–165.

Bott, C.J., and Winckler, B. (2020). Intermediate filaments in developing neurons: Beyond structure. *Cytoskeleton.*

Brat, D.J., Aldape, K., Colman, H., Holland, E.C., Louis, D.N., Jenkins, R.B., Kleinschmidt-DeMasters, B.K., Perry, A., Reifenberger, G., Stupp, R., et al. (2018). cIMPACT-NOW update 3: recommended diagnostic criteria for “Diffuse astrocytic glioma, IDH-wildtype, with molecular features of glioblastoma, WHO grade IV.” *Acta Neuropathol.* 136, 805–810.

Brat, D.J., Aldape, K., Colman, H., Figarella-Branger, D., Fuller, G.N., Giannini, C., Holland, E.C., Jenkins, R.B., Kleinschmidt-DeMasters, B., Komori, T., et al. (2020). cIMPACT-NOW update 5: recommended grading criteria and terminologies for IDH-mutant astrocytomas. *Acta Neuropathol.* 139, 603–608.

Brennan, C.W., Verhaak, R.G.W., McKenna, A., Campos, B., Nounshmehr, H., Salama, S.R., Zheng, S., Chakravarty, D., Sanborn, J.Z., Berman, S.H., et al. (2013). The Somatic Genomic Landscape of Glioblastoma. *Cell* 155, 462–477.

Bronisz, A., Wang, Y., Nowicki, M.O., Peruzzi, P., Ansari, K.I., Ogawa, D., Balaj, L., De Rienzo, G., Mineo, M., Nakano, I., et al. (2014). Extracellular Vesicles Modulate the Glioblastoma Microenvironment via a Tumor Suppression Signaling Network Directed by miR-1. *Cancer Res.* 74, 738–750.

Bronisz, A., Godlewski, J., and Chiocca, E.A. (2016). Extracellular Vesicles and MicroRNAs: Their Role in Tumorigenicity and Therapy for Brain Tumors. *Cell. Mol. Neurobiol.* 36, 361–376.

Brunet, M.A., Brunelle, M., Lucier, J.F., Delcourt, V., Levesque, M., Grenier, F., Samandi, S., Leblanc, S., Aguilar, J.D., Dufour, P., et al. (2019). OpenProt: A more comprehensive guide to explore eukaryotic coding potential and proteomes. *Nucleic Acids Res.*

Budnik, B., Levy, E., Harmange, G., and Slavov, N. (2018). SCoPE-MS: mass spectrometry of single mammalian cells quantifies proteome heterogeneity during cell differentiation. *Genome Biol.* *19*.

Calligaris, D., Feldman, D.R., Norton, I., Olubiyi, O., Changelian, A.N., Machaidze, R., Vestal, M.L., Laws, E.R., Dunn, I.F., Santagata, S., et al. (2015). MALDI mass spectrometry imaging analysis of pituitary adenomas for near-real-time tumor delineation. *Proc. Natl. Acad. Sci. U. S. A.* *112*, 9978–9983.

Cankovic, M., Nikiforova, M.N., Snuderl, M., Adesina, A.M., Lindeman, N., Wen, P.Y., and Lee, E.Q. (2013). The role of MGMT testing in clinical practice: A report of the association for molecular pathology. *J. Mol. Diagnostics* *15*, 539–555.

Cardon, T., Salzet, M., Franck, J., and Fournier, I. (2019). Nuclei of HeLa cells interactomes unravel a network of ghost proteins involved in proteins translation. *Biochim. Biophys. Acta - Gen. Subj.* *1863*, 1458–1470.

Cardon, T., Franck, J., Coyaud, E., Laurent, E.M.N., Damato, M., Maffia, M., Vergara, D., Fournier, I., and Salzet, M. (2020a). Alternative proteins are functional regulators in cell reprogramming by PKA activation. *Nucleic Acids Res.*

Cardon, T., Hervé, F., Delcourt, V., Roucou, X., Salzet, M., Franck, J., and Fournier, I. (2020b). Optimized Sample Preparation Workflow for Improved Identification of Ghost Proteins. *Anal. Chem.* *92*, 1122–1129.

Casadonte, R., Kriegsmann, M., Perren, A., Baretton, G., Deininger, S.O., Kriegsmann, K., Welsch, T., Pilarsky, C., and Kriegsmann, J. (2019). Development of a Class Prediction Model to Discriminate Pancreatic Ductal Adenocarcinoma from Pancreatic Neuroendocrine Tumor by MALDI Mass Spectrometry Imaging. *Proteomics - Clin. Appl.* *13*.

Chang, J.B., Chen, F., Yoon, Y.G., Jung, E.E., Babcock, H., Kang, J.S., Asano, S., Suk, H.J., Pak, N., Tillberg, P.W., et al. (2017). Iterative expansion microscopy. *Nat. Methods* *14*, 593–599.

Chatterjee, A., Jana, S., Chatterjee, S., Wastall, L.M., Mandal, G., Nargis, N., Roy, H., Hughes, T.A., and Bhattacharyya, A. (2019). MicroRNA-222 reprogrammed cancer-associated fibroblasts enhance growth and metastasis of breast cancer. *Br. J. Cancer* *121*, 679–689.

Chen, F., Tillberg, P.W., and Boyden, E.S. (2015a). Expansion microscopy. *Science* (80-). *347*, 543–548.

Chen, F., Tillberg, P.W., and Boyden, E.S. (2015b). Optical imaging. Expansion microscopy. *Science* *347*, 543–548.

Chen, F., Wassie, A.T., Cote, A.J., Sinha, A., Alon, S., Asano, S., Daugharthy, E.R., Chang, J.B., Marblestone, A., Church, G.M., et al. (2016). Nanoscale imaging of RNA with expansion microscopy. *Nat. Methods* *13*, 679–684.

Chen, G., Huang, A.C., Zhang, W., Zhang, G., Wu, M., Xu, W., Yu, Z., Yang, J., Wang, B., Sun, H., et al. (2018). Exosomal PD-L1 contributes to immunosuppression and is associated with anti-PD-1 response. *Nature* *560*, 382–386.

Cheng, W., Ren, X., Zhang, C., Cai, J., Liu, Y., Han, S., and Wu, A. (2016). Bioinformatic profiling identifies an immune-related risk signature for glioblastoma. *Neurology*.

Chinot, O.L., Wick, W., Mason, W., Henriksson, R., Saran, F., Nishikawa, R., Carpentier, A.F., Hoang-Xuan, K., Kavan, P., Cernea, D., et al. (2014). Bevacizumab plus radiotherapy-temozolomide for newly diagnosed glioblastoma. *N. Engl. J. Med.*

Choi, D.-S., Kim, D.-K., Kim, Y.-K., and Ghoo, Y.S. (2015). Proteomics of extracellular vesicles: Exosomes and ectosomes: PROTEOMICS OF EXTRACELLULAR VESICLES. *Mass Spectrom. Rev.* *34*, 474–490.

Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* *26*, 1367–1372.

Cox, J., Neuhauser, N., Michalski, A., Scheltema, R.A., Olsen, J. V, and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. *J. Proteome Res.* *10*, 1794–1805.

Cox, J., Hein, M.Y., Lubner, C.A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol. Cell. Proteomics* *13*, 2513–2526.

Cuoco, J.A., Benko, M.J., Busch, C.M., Rogers, C.M., Prickett, J.T., and Marvin, E.A. (2018). Vaccine-Based Immunotherapeutics for the Treatment of Glioblastoma: Advances, Challenges, and Future Perspectives. *World Neurosurg.* *120*, 302–315.

D, T., JH, K., J, O., H, T., P, M., and T, A. (2012). On the Importance of Mathematical Methods for Analysis of MALDI-imaging Mass Spectrometry Data. *J. Integr. Bioinform.* *9*.

Deighton, R.F., McGregor, R., Kemp, J., McCulloch, J., and Whittle, I.R. (2010). Glioma pathophysiology: Insights emerging from proteomics. *Brain Pathol.*

Deighton, R.F., Le Bihan, T., Martin, S.F., Barrios-Llerena, M.E., Gerth, A.M.J., Kerr, L.E., McCulloch, J., and Whittle, I.R. (2014). The proteomic response in glioblastoma in young patients. *J. Neurooncol.* *119*, 79–89.

Deininger, S.O., Cornett, D.S., Paape, R., Becker, M., Pineau, C., Rauser, S., Walch, A., and Wolski, E. (2011). Normalization in MALDI-TOF imaging datasets of proteins: Practical considerations. *Anal. Bioanal. Chem.* *401*, 167–181.

Delcourt, V., Franck, J., Quanico, J., Gimeno, J.-P., Wisztorski, M., Raffo-Romero, A., Kobeissy, F., Roucou, X., Salzet, M., and Fournier, I. (2018a). Spatially-Resolved Top-down Proteomics Bridged to MALDI MS Imaging Reveals the Molecular Physiome of Brain Regions. *Mol. Cell. Proteomics* *17*, 357–372.

Delcourt, V., Staskevicius, A., Salzet, M., Fournier, I., and Roucou, X. (2018b). Small Proteins Encoded by Unannotated ORFs are Rising Stars of the Proteome, Confirming Shortcomings in Genome Annotations and Current Vision of an mRNA. *Proteomics*.

Dickinson, H.O., Nyari, T.A., and Parker, L. (2002). Childhood solid tumours in relation to infections in the community in Cumbria during pregnancy and around the time of birth. *Br.*

J. Cancer.

Dilillo, M., Pellegrini, D., Ait-Belkacem, R., De Graaf, E.L., Caleo, M., and McDonnell, L.A. (2017). Mass Spectrometry Imaging, Laser Capture Microdissection, and LC-MS/MS of the Same Tissue Section. *J. Proteome Res.* *16*, 2993–3001.

Duhamel, M., Rose, M., Rodet, F., Murgoci, A.N., Zografidou, L., Régnier-Vigouroux, A., Abeele, F., Vanden, Kobeissy, F., Nataf, S., Pays, L., et al. (2018). Paclitaxel Treatment and Proprotein Convertase 1/3 (PC1/3) Knockdown in Macrophages is a Promising Antiglioma Strategy as Revealed by Proteomics and Cytotoxicity Studies. *Mol. Cell. Proteomics* *17*, 1126–1143.

Durcin, M., Fleury, A., Taillebois, E., Hilairet, G., Krupova, Z., Henry, C., Truchet, S., Trötz Müller, M., Köfeler, H., Mabileau, G., et al. (2017). Characterisation of adipocyte-derived extracellular vesicle subtypes identifies distinct protein and lipid signatures for large and small extracellular vesicles. *J. Extracell. Vesicles* *6*, 1305677.

Eder, K., and Kalman, B. (2014). Molecular Heterogeneity of Glioblastoma and its Clinical Relevance. *Pathol. Oncol. Res.* *20*, 777–787.

Ellison, D.W., Hawkins, C., Jones, D.T.W., Onar-Thomas, A., Pfister, S.M., Reifenberger, G., and Louis, D.N. (2019). cIMPACT-NOW update 4: diffuse gliomas characterized by MYB, MYBL1, or FGFR1 alterations or BRAF V600E mutation. *Acta Neuropathol.* *137*, 683–687.

Ellison, D.W., Aldape, K.D., Capper, D., Fouladi, M., Gilbert, M.R., Gilbertson, R.J., Hawkins, C., Merchant, T., Pajtler, K., Venneti, S., et al. (2020). cIMPACT-NOW Update 7: advancing the molecular classification of ependymal tumors. *Brain Pathol.*

Ellor, S. V., Pagano-Young, T.A., and Avgeropoulos, N.G. (2014). Glioblastoma: Background, standard treatment paradigms, and supportive care considerations. *J. Law, Med. Ethics* *42*, 171–182.

Erondu, N.E., and Kennedy, M.B. (1985). Regional distribution of type II Ca²⁺/calmodulin-dependent protein kinase in rat brain. *J. Neurosci.*

Escrevente, C., Keller, S., Altevogt, P., and Costa, J. (2011). Interaction and uptake of exosomes by ovarian cancer cells. *BMC Cancer* *11*, 1–10.

Fatou, B., Saudemont, P., Leblanc, E., Vinatier, D., Mesdag, V., Wisztorski, M., Focsa, C., Salzet, M., Ziskind, M., and Fournier, I. (2016). In vivo Real-Time Mass Spectrometry for Guided Surgery Application. *Sci. Rep.* *6*.

Feenstra, A.D., Dueñas, M.E., and Lee, Y.J. (2017). Five Micron High Resolution MALDI Mass Spectrometry Imaging with Simple, Interchangeable, Multi-Resolution Optical System. *J. Am. Soc. Mass Spectrom.* *28*, 434–442.

Fenstermaker, R.A., Ciesielski, M.J., Qiu, J., Yang, N., Frank, C.L., Lee, K.P., Mechtler, L.R., Belal, A., Ahluwalia, M.S., and Hutson, A.D. (2016). Clinical study of a survivin long peptide vaccine (SurVaxM) in patients with recurrent malignant glioma. *Cancer Immunol. Immunother.* *65*, 1339–1352.

Ferguson, S., and Lesniak, M.S. (2005). Percival Bailey and the classification of brain

tumors. *Neurosurg. Focus* 18, e7.

Franck, J., Arafah, K., Elayed, M., Bonnel, D., Vergara, D., Jacquet, A., Vinatier, D., Wisztorski, M., Day, R., Fournier, I., et al. (2009). MALDI imaging mass spectrometry state of the art technology in clinical proteomics. *Mol. Cell. Proteomics* 8, 2023–2033.

Franck, J., Quanico, J., Wisztorski, M., Day, R., Salzet, M., and Fournier, I. (2013). Quantification-based mass spectrometry imaging of proteins by parafilm assisted microdissection. *Anal. Chem.*

Gallart-Palau, X., Serra, A., Wong, A.S.W., Sandin, S., Lai, M.K.P., Chen, C.P., Kon, O.L., and Sze, S.K. (2015). Extracellular vesicles are rapidly purified from human plasma by PROSPR Organic Solvent PREcipitation (PROSPR). *Sci. Rep.* 5.

Gallart-Palau, X., Serra, A., and Sze, S.K. (2016). Enrichment of extracellular vesicles from tissues of the central nervous system by PROSPR. *Mol. Neurodegener.* 11, 41.

Gámez-Valero, A., Monguió-Tortajada, M., Carreras-Planella, L., Franquesa, M., Beyer, K., and Borràs, F.E. (2016). Size-Exclusion Chromatography-based isolation minimally alters Extracellular Vesicles' characteristics compared to precipitating agents. *Sci. Rep.* 6.

Gao, K.M., Chen, X.C., Zhang, J.X., Wang, Y., Yan, W., and You, Y.P. (2015). A pseudogene-signature in glioma predicts survival. *J. Exp. Clin. Cancer Res.*

Gao, R., Asano, S.M., and Boyden, E.S. (2017). Q & A: Expansion microscopy. *BMC Biol.*

Gautam, P., Nair, S.C., Gupta, M.K., Sharma, R., Polisetty, R.V., Uppin, M.S., Sundaram, C., Puligopu, A.K., Ankathi, P., Purohit, A.K., et al. (2012). Proteins with Altered Levels in Plasma from Glioblastoma Patients as Revealed by iTRAQ-Based Quantitative Proteomic Analysis. *PLoS One* 7.

Ghantasala, S., Gollapalli, K., Epari, S., Moiyadi, A., and Srivastava, S. (2020). Glioma tumor proteomics: clinically useful protein biomarkers and future perspectives. *Expert Rev. Proteomics.*

Giampieri, R., Piva, F., Occhipinti, G., Bittoni, A., Righetti, A., Pagliaretta, S., Murrone, A., Bianchi, F., Amantini, C., Giuliotti, M., et al. (2019). Clinical impact of different exosomes' protein expression in pancreatic ductal carcinoma patients treated with standard first line palliative chemotherapy. *PLoS One* 14, e0215990.

Giering, A., Pszczolkowska, D., Walentynowicz, K.A., Rajan, W.D., and Kaminska, B. (2017). Immune microenvironment of gliomas. *Lab. Investig.* 97, 498–518.

Gilbert, M.R., Dignam, J.J., Armstrong, T.S., Wefel, J.S., Blumenthal, D.T., Vogelbaum, M.A., Colman, H., Chakravarti, A., Pugh, S., Won, M., et al. (2014). A randomized trial of bevacizumab for newly diagnosed glioblastoma. *N. Engl. J. Med.*

Gill, B.J., Pisapia, D.J., Malone, H.R., Goldstein, H., Lei, L., Sonabend, A., Yun, J., Samanamud, J., Sims, J.S., Banu, M., et al. (2014). MRI-localized biopsies reveal subtype-specific differences in molecular and cellular composition at the margins of glioblastoma. *Proc. Natl. Acad. Sci. U. S. A.* 111, 12550–12555.

Giordano, S., Morosi, L., Veglianese, P., Licandro, S.A., Frapolli, R., Zucchetti, M., Cappelletti, G., Falciola, L., Pifferi, V., Visentin, S., et al. (2016). 3D Mass Spectrometry Imaging Reveals a Very Heterogeneous Drug Distribution in Tumors. *Sci. Rep.* 6.

Gollapalli, K., Ghantasala, S., Kumar, S., Srivastava, R., Rapole, S., Moiyadi, A., Epari, S., and Srivastava, S. (2017). Subventricular zone involvement in Glioblastoma-A proteomic evaluation and clinicoradiological correlation. *Sci. Rep.* 7, 1449.

Gourlay, J., Morokoff, A.P., Luwor, R.B., Zhu, H.-J., Kaye, A.H., and Stylli, S.S. (2017). The emergent role of exosomes in glioma. *J. Clin. Neurosci. Off. J. Neurosurg. Soc. Australas.* 35, 13–23.

Gramatzki, D., Dehler, S., Rushing, E.J., Zaugg, K., Hofer, S., Yonekawa, Y., Bertalanffy, H., Valavanis, A., Korol, D., Rohrmann, S., et al. (2016). Glioblastoma in the Canton of Zurich, Switzerland revisited: 2005 to 2009. *Cancer* 122, 2206–2215.

Griffiths, R.L., Konijnenberg, A., Viner, R., and Cooper, H.J. (2019). Direct Mass Spectrometry Analysis of Protein Complexes and Intact Proteins up to >70 kDa from Tissue. *Anal. Chem.* 91, 6962–6966.

De Groot, J.F., Cloughesy, T.F., Pitz, M.W., Narita, Y., and Nonomura, T. (2018). A randomized, multicenter phase 2 study of DSP-7888 dosing emulsion in combination with bevacizumab (Bev) versus Bev alone in patients with recurrent or progressive glioblastoma. *J. Clin. Oncol.* 36, TPS2071–TPS2071.

Guadagno, E., Presta, I., Maisano, D., Donato, A., Pirrone, C.K., Cardillo, G., Corrado, S.D., Mignogna, C., Mancuso, T., Donato, G., et al. (2018). Role of macrophages in brain tumor growth and progression. *Int. J. Mol. Sci.*

Gutmann, D.H., Huang, Z.Y., Hedrick, N.M., Ding, H., Guha, A., and Watson, M.A. (2002). Mouse glioma gene expression profiling identifies novel human glioma-associated genes. *Ann. Neurol.*

Hadla, M., Palazzolo, S., Corona, G., Caligiuri, I., Canzonieri, V., Toffoli, G., and Rizzolio, F. (2016). Exosomes increase the therapeutic index of doxorubicin in breast and ovarian cancer mouse models. *Nanomedicine* 11, 2431–2441.

Harris, G.A., Nicklay, J.J., and Caprioli, R.M. (2013). Localized in situ hydrogel-mediated protein digestion and extraction technique for on-tissue analysis. *Anal. Chem.*

Hilf, N., Kuttruff-Coqui, S., Frenzel, K., Bukur, V., Stevanović, S., Gouttefangeas, C., Platten, M., Tabatabai, G., Dutoit, V., van der Burg, S.H., et al. (2019). Actively personalized vaccination trial for newly diagnosed glioblastoma. *Nature* 565, 240–245.

Hoshino, A., Costa-Silva, B., Shen, T.L., Rodrigues, G., Hashimoto, A., Tesic Mark, M., Molina, H., Kohsaka, S., Di Giannatale, A., Ceder, S., et al. (2015). Tumour exosome integrins determine organotropic metastasis. *Nature* 527, 329–335.

Huang, Z. yong, Wu, Y.L., Burke, S.P., and Gutmann, D.H. (2003). The 43,000 growth-associated protein functions as a negative growth regulator in glioma. *Cancer Res.*

Hulsen, T., de Vlieg, J., and Alkema, W. (2008). BioVenn - A web application for the

comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* 9, 488.

Iwadate, Y., Sakaida, T., Hiwasa, T., Nagai, Y., Ishikura, H., Takiguchi, M., and Yamaura, A. (2004). Molecular Classification and Survival Prediction in Human Gliomas Based on Proteome Analysis. *CANCER Res.* 64, 2496–2501.

Johnson, D.R., Guerin, J.B., Ruff, M.W., Fang, S., Hunt, C.H., Morris, J.M., Morris, P.P., and Kaufmann, T.J. (2019). Glioma response assessment: Classic pitfalls, novel confounders, and emerging imaging tools. In *British Journal of Radiology*, (British Institute of Radiology), p.

Jovčevska, I., Zupanec, N., Urlep, Ž., Vranic, A., Matos, B., Stokin, C.L., Muyldermans, S., Myers, M.P., Buzdin, A.A., Petrov, I., et al. (2017). Differentially expressed proteins in glioblastoma multiforme identified with a nanobody-based anti-proteome approach and confirmed by OncoFinder as possible tumor-class predictive biomarker candidates. *Oncotarget* 8, 44141–44158.

Kalinina, J., Peng, J., Ritchie, J.C., and Van Meir, E.G. (2011). Proteomics of gliomas: Initial biomarker discovery and evolution of technology. *Neuro. Oncol.*

Kalluri, R., and Lebleu, V.S. (2016). Discovery of double-stranded genomic DNA in circulating exosomes. *Cold Spring Harb. Symp. Quant. Biol.* 81, 275–280.

Kannan, K., Inagaki, A., Silber, J., Gorovets, D., Zhang, J., Kasthuber, E.R., Heguy, A., Petrini, J.H., Chan, T.A., and Huse, J.T. (2012). Whole exome sequencing identifies ATRX mutation as a key molecular determinant in lower-grade glioma. *Oncotarget* 3, 1194–1204.

Katagiri, T., Hatano, N., Aihara, M., Kawano, H., Okamoto, M., Liu, Y., Izumi, T., Maekawa, T., Nakamura, S., Ishihara, T., et al. (2010). Proteomic analysis of proteins expressing in regions of rat brain by a combination of SDS-PAGE with nano-liquid chromatography-quadrupole-time of flight tandem mass spectrometry. *Proteome Sci.* 8, 41.

Kharaziha, P., Chioureas, D., Rutishauser, D., Baltatzis, G., Lennartsson, L., Fonseca, P., Azimi, A., Hultenby, K., Zubarev, R., Ullén, A., et al. (2015). Molecular profiling of prostate cancer derived exosomes may reveal a predictive signature for response to docetaxel. *Oncotarget* 6, 21740–21754.

Kleihues, P., Burger, P.C., and Scheithauer, B.W. (1993). *Histological Typing of Tumours of the Central Nervous System* (Springer Berlin Heidelberg).

Klein, O., Strohschein, K., Nebrich, G., Oetjen, J., Trede, D., Thiele, H., Alexandrov, T., Giavalisco, P., Duda, G.N., von Roth, P., et al. (2014). MALDI imaging mass spectrometry: Discrimination of pathophysiological regions in traumatized skeletal muscle by characteristic peptide signatures. *14*, 2249–2260.

Kooijmans, S.A.A., Vader, P., van Dommelen, S.M., van Solinge, W.W., and Schiffelers, R.M. (2012). Exosome mimetics: A novel class of drug delivery systems. *Int. J. Nanomedicine* 7, 1525–1541.

Krenzlin, H., Behera, P., Lorenz, V., Passaro, C., Zdioruk, M., Nowicki, M.O., Grauwet, K., Zhang, H., Skubal, M., Ito, H., et al. (2019). Cytomegalovirus promotes murine glioblastoma

growth via pericyte recruitment and angiogenesis. *J. Clin. Invest.*

Kriegsmann, J., Casadonte, R., Kriegsmann, K., Longuespée, R., and Kriegsmann, M. (2018). Mass spectrometry in pathology – Vision for a future workflow. *Pathol. Res. Pract.* *214*, 1057–1063.

Ku, T., Swaney, J., Park, J.Y., Albanese, A., Murray, E., Hun Cho, J., Park, Y.G., Mangena, V., Chen, J., and Chung, K. (2016). Multiplexed and scalable super-resolution imaging of three-dimensional protein localization in size-adjustable tissues. *Nat. Biotechnol.* *34*, 973–981.

Lapointe, S., Perry, A., and Butowski, N.A. (2018). Primary brain tumours in adults. *Lancet* *392*, 432–446.

Lathia, J.D., Mack, S.C., Mulkearns-Hubert, E.E., Valentim, C.L.L., and Rich, J.N. (2015). Cancer stem cells in glioblastoma. *Genes Dev.* *29*, 1203–1217.

Lee, E.-J., Rath, P., Liu, J., Ryu, D., Pei, L., Noonepalle, S.K., Shull, A.Y., Feng, Q., Litofsky, N.S., Miller, D.C., et al. (2015). Identification of Global DNA Methylation Signatures in Glioblastoma-Derived Cancer Stem Cells. *J. Genet. Genomics* *42*, 355–371.

Lee, E., Yong, R.L., Paddison, P., and Zhu, J. (2018). Comparison of glioblastoma (GBM) molecular classification methods. *Semin. Cancer Biol.* *53*, 201–211.

Lei, B., Yu, L., Jung, T.A., Deng, Y., Xiang, W., Liu, Y., and Qi, S. (2018). Prospective Series of Nine Long Noncoding RNAs Associated with Survival of Patients with Glioblastoma. *J. Neurol. Surgery, Part A Cent. Eur. Neurosurg.*

Lemaire, Q., Raffo-Romero, A., Arab, T., Van Camp, C., Drago, F., Forte, S., Gimeno, J.P., Begard, S., Colin, M., Vizioli, J., et al. (2019). Isolation of microglia-derived extracellular vesicles: Towards miRNA signatures and neuroprotection. *J. Nanobiotechnology* *17*.

Lemaire, Q., Duhamel, M., Raffo-Romero, A., Salzet, M., and Lefebvre, C. (2020). Characterization of immune cell-derived extracellular vesicles and studying functional impact on cell environment. *J. Vis. Exp.* *2020*, 1–11.

Lemaire, R., Tabet, J.C., Ducoroy, P., Hendra, J.B., Salzet, M., and Fournier, I. (2006a). Solid Ionic Matrixes for Direct Tissue Analysis and MALDI Imaging. *Anal. Chem.* *78*, 809–819.

Lemaire, R., Wisztorski, M., Desmons, A., Tabet, J.C., Day, R., Salzet, M., and Fournier, I. (2006b). MALDI-MS direct tissue analysis of proteins: Improving signal sensitivity using organic treatments. *Anal. Chem.* *78*, 7145–7153.

Lemaire, R., Menguellet, S.A., Stauber, J., Marchaudon, V., Lucot, J.P., Collinet, P., Farine, M.O., Vinatier, D., Day, R., Ducoroy, P., et al. (2007). Specific MALDI imaging and profiling for biomarker hunting and validation: Fragment of the 11S proteasome activator complex, reg alpha fragment, is a new potential ovary cancer biomarker. *J. Proteome Res.*

Limam, S., Missaoui, N., Mestiri, S., Yacoubi, M.T., Krifa, H., Selmi, B., and Mokni, M. (2019). Epstein-Barr virus infection in gliomas. *Curr. Res. Transl. Med.*

Lin, Y., Jiang, T., Zhou, K., Xu, L., Chen, B., Li, G., Qiu, X., Jiang, T., Zhang, W., and Song, S.W. (2009). Plasma IGFBP-2 levels predict clinical outcomes of patients with high-grade

gliomas. *Neuro. Oncol.* *11*, 468–476.

Liu, J., Zhang, X., Yan, X., Sun, M.E.I., Fan, Y., and Huang, Y. (2019). Significance of TERT and ATRX mutations in glioma. *Oncol. Lett.* *17*, 95–102.

Liu, Y., Hu, H., Zhang, C., Wang, Z., Li, M., and Jiang, T. (2016). Integrated analysis identified genes associated with a favorable prognosis in oligodendrogliomas. *Genes, Chromosom. Cancer* *55*, 169–176.

Longuespée, R., Alberts, D., Pottier, C., Smargiasso, N., Mazzucchelli, G., Baiwir, D., Kriegsmann, M., Herfs, M., Kriegsmann, J., Delvenne, P., et al. (2016). A laser microdissection-based workflow for FFPE tissue microproteomics: Important considerations for small sample processing. *Methods* *104*, 154–162.

Longuespée, R., Alberts, D., Baiwir, D., Mazzucchelli, G., Smargiasso, N., and De Pauw, E. (2018). MALDI Imaging Combined with Laser Microdissection-Based Microproteomics for Protein Identification: Application to Intratumor Heterogeneity Studies. In *Methods in Molecular Biology*, (Humana Press Inc.), pp. 297–312.

Lötvall, J., Hill, A.F., Hochberg, F., Buzás, E.I., Vizio, D. Di, Gardiner, C., Gho, Y.S., Kurochkin, I. V., Mathivanan, S., Quesenberry, P., et al. (2014). Minimal experimental requirements for definition of extracellular vesicles and their functions: A position statement from the International Society for Extracellular Vesicles. *J. Extracell. Vesicles* *3*.

Louis, D.N., Ohgaki, H., Wiestler, O.D., Cavenee, W.K., Burger, P.C., Jouvet, A., Scheithauer, B.W., and Kleihues, P. (2007). The 2007 WHO Classification of Tumours of the Central Nervous System. *Acta Neuropathol.* *114*, 97–109.

Louis, D.N., Perry, A., Reifenberger, G., von Deimling, A., Figarella-Branger, D., Cavenee, W.K., Ohgaki, H., Wiestler, O.D., Kleihues, P., and Ellison, D.W. (2016). The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathol.* *131*, 803–820.

Louis, D.N., Aldape, K., Brat, D.J., Capper, D., Ellison, D.W., Hawkins, C., Paulus, W., Perry, A., Reifenberger, G., Figarella-Branger, D., et al. (2017). cIMPACT-NOW (the consortium to inform molecular and practical approaches to CNS tumor taxonomy): a new initiative in advancing nervous system tumor classification. *Brain Pathol.* *27*, 851–852.

Louis, D.N., Wesseling, P., Paulus, W., Giannini, C., Batchelor, T.T., Cairncross, J.G., Capper, D., Figarella-Branger, D., Lopes, M.B., Wick, W., et al. (2018a). cIMPACT-NOW update 1: Not Otherwise Specified (NOS) and Not Elsewhere Classified (NEC). *Acta Neuropathol.* *135*, 481–484.

Louis, D.N., Giannini, C., Capper, D., Paulus, W., Figarella-Branger, D., Lopes, M.B., Batchelor, T.T., Cairncross, J.G., van den Bent, M., Wick, W., et al. (2018b). cIMPACT-NOW update 2: diagnostic clarifications for diffuse midline glioma, H3 K27M-mutant and diffuse astrocytoma/anaplastic astrocytoma, IDH-mutant. *Acta Neuropathol.* *135*, 639–642.

Louis, D.N., Wesseling, P., Aldape, K., Brat, D.J., Capper, D., Cree, I.A., Eberhart, C., Figarella-Branger, D., Fouladi, M., Fuller, G.N., et al. (2020). cIMPACT-NOW update 6: new entity and diagnostic principle recommendations of the cIMPACT-Utrecht meeting on future

CNS tumor classification and grading. *Brain Pathol.* *30*.

Lv, M. meng, Zhu, X. ya, Chen, W. xian, Zhong, S. liang, Hu, Q., Ma, T. fei, Zhang, J., Chen, L., Tang, J. hai, and Zhao, J. hua (2014). Exosomes mediate drug resistance transfer in MCF-7 breast cancer cells and a probable mechanism is delivery of P-glycoprotein. *Tumor Biol.* *35*, 10773–10779.

Macia, E., Ehrlich, M., Massol, R., Boucrot, E., Brunner, C., and Kirchhausen, T. (2006). Dynasore, a Cell-Permeable Inhibitor of Dynamin. *Dev. Cell* *10*, 839–850.

Mandoj, C., Tomao, L., and Conti, L. (2019). Coagulation in Brain Tumors: Biological Basis and Clinical Implications. *Front. Neurol.* *10*, 181.

Marquardt, C., Tolstik, T., Bielecki, C., Kaufmann, R., Crecelius, A.C., Schubert, U.S., Settmacher, U., Stallmach, A., and Dirsch, O. (2015). MALDI imaging-based classification of hepatocellular carcinoma and non-malignant lesions in fibrotic liver tissue. *Z. Gastroenterol.* *53*, 33–39.

Meding, S., Nitsche, U., Balluff, B., Elsner, M., Rauser, S., Schöne, C., Nipp, M., Maak, M., Feith, M., Ebert, M.P., et al. (2012). Tumor classification of six common cancer types based on proteomic profiling by MALDI imaging. In *Journal of Proteome Research*, (American Chemical Society), pp. 1996–2003.

Minciacchi, V.R., Freeman, M.R., and Di Vizio, D. (2015). Extracellular Vesicles in Cancer: Exosomes, Microvesicles and the Emerging Role of Large Oncosomes. *Semin. Cell Dev. Biol.* *40*, 41–51.

Minerva, L., Clerens, S., Baggerman, G., and Arckens, L. (2008). Direct profiling and identification of peptide expression differences in the pancreas of control and ob/ob mice by imaging mass spectrometry. *Proteomics* *8*, 3763–3774.

Miyauchi, E., Furuta, T., Ohtsuki, S., Tachikawa, M., Uchida, Y., Sabit, H., Obuchi, W., Baba, T., Watanabe, M., Terasaki, T., et al. (2018). Identification of blood biomarkers in glioblastoma by SWATH mass spectrometry and quantitative targeted absolute proteomics. *PLoS One* *13*.

Montero, R., Abad-García, B., Garate, J., Martín-Saiz, L., Barceló-Coblijn, G., and Fernandez, J.A. (2020). Improving spatial resolution of a LTQ Orbitrap MALDI source. *J. Am. Soc. Mass Spectrom.* *jasms.0c00133*.

Morales-Sánchez, A., and Fuentes-Pananá, E.M. (2014). Human viruses and cancer. *Viruses* *6*, 4047–4079.

Moutal, A., Honnorat, J., Massoma, P., Désormeaux, P., Bertrand, C., Malleval, C., Watrin, C., Chounlamountri, N., Mayeur, M.E., Besançon, R., et al. (2015). CRMP5 controls glioblastoma cell proliferation and survival through Notch-dependent signaling. *Cancer Res.*

Murgoci, A.N., Cardon, T., Aboulouard, S., Duhamel, M., Fournier, I., Cizkova, D., and Salzet, M. (2020). Reference and Ghost Proteins Identification in Rat C6 Glioma Extracellular Vesicles. *IScience* *23*.

Murphy, R.C., Hankin, J.A., Barkley, R.M., and Zemski Berry, K.A. (2011). MALDI imaging

of lipids after matrix sublimation/deposition. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* 1811, 970–975.

Neftel, C., Laffy, J., Filbin, M.G., Hara, T., Shore, M.E., Rahme, G.J., Richman, A.R., Silverbush, D., Shaw, M.L., Hebert, C.M., et al. (2019). An Integrative Model of Cellular States, Plasticity, and Genetics for Glioblastoma. *Cell*.

Nilsson, A., Goodwin, R.J.A., Swales, J.G., Gallagher, R., Shankaran, H., Sathe, A., Pradeepan, S., Xue, A., Keirstead, N., Sasaki, J.C., et al. (2015). Investigating Nephrotoxicity of Polymyxin Derivatives by Mapping Renal Distribution Using Mass Spectrometry Imaging. *Chem. Res. Toxicol.* 28, 1823–1830.

Noushmehr, H., Weisenberger, D.J., Diefes, K., Phillips, H.S., Pujara, K., Berman, B.P., Pan, F., Pelloski, C.E., Sulman, E.P., Bhat, K.P., et al. (2010). Identification of a CpG Island Methylator Phenotype that Defines a Distinct Subgroup of Glioma. *Cancer Cell* 17, 510–522.

Ogrinc, N., Saudemont, P., Balog, J., Robin, Y.M., Gimeno, J.P., Pascal, Q., Tierny, D., Takats, Z., Salzet, M., and Fournier, I. (2019). Water-assisted laser desorption/ionization mass spectrometry for minimally invasive in vivo and real-time surface analysis using SpiderMass. *Nat. Protoc.* 14, 3162–3182.

Ohgaki, H., and Kleihues, P. (2013). The Definition of Primary and Secondary Glioblastoma. *Clin. Cancer Res.* 19, 764–772.

Okada, H., Weller, M., Huang, R., Finocchiaro, G., Gilbert, M.R., Wick, W., Ellingson, B.M., Hashimoto, N., Pollack, I.F., Brandes, A.A., et al. (2015). Immunotherapy response assessment in neuro-oncology: A report of the RANO working group. *Lancet Oncol.* 16, e534–e542.

Orozco, J.I.J., Knijnenburg, T.A., Manughian-Peter, A.O., Salomon, M.P., Barkhoudarian, G., Jalas, J.R., Wilmott, J.S., Hothi, P., Wang, X., Takasumi, Y., et al. (2018). Epigenetic profiling for the molecular classification of metastatic brain tumors. *Nat. Commun.* 9.

Osti, D., Bene, M. Del, Rappa, G., Santos, M., Matafora, V., Richichi, C., Faletti, S., Beznoussenko, G. V., Mironov, A., Bachi, A., et al. (2019). Clinical significance of extracellular vesicles in plasma from glioblastoma patients. *Clin. Cancer Res.* 25, 266–276.

Ostrom, Q.T., Gittleman, H., Stetson, L., Virk, S.M., and Barnholtz-Sloan, J.S. (2015). Epidemiology of Gliomas. *Cancer Treat. Res.* 163, 1–14.

Ostrom, Q.T., Gittleman, H., Truitt, G., Boscia, A., Kruchko, C., and Barnholtz-Sloan, J.S. (2018). CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the United States in 2011-2015. *Neuro. Oncol.*

Ostrom, Q.T., Cioffi, G., Gittleman, H., Patil, N., Waite, K., Kruchko, C., and Barnholtz-Sloan, J.S. (2019). CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012-2016. *Neuro. Oncol.*

Otani, R., Uzuka, T., and Ueki, K. (2017). Classification of adult diffuse gliomas by molecular markers—a short review with historical footnote. *Jpn. J. Clin. Oncol.* 47, 2–6.

Park, H., Choi, D., Park, J.S., Sim, C., Park, S., Kang, S., Yim, H., Lee, M., Kim, J., Pac, J., et al. (2019). Scalable and Isotropic Expansion of Tissues with Simply Tunable Expansion Ratio. *Adv. Sci.* *6*, 1901673.

Parsons, D.W., Jones, S., Zhang, X., Lin, J.C.-H., Leary, R.J., Angenendt, P., Mankoo, P., Carter, H., Siu, I.-M., Gallia, G.L., et al. (2008). An Integrated Genomic Analysis of Human Glioblastoma Multiforme. *Science* (80-). *321*, 1807–1812.

Patil, V., and Mahalingam, K. (2019). Comprehensive analysis of Reverse Phase Protein Array data reveals characteristic unique proteomic signatures for glioblastoma subtypes. *Gene* *685*, 85–95.

Peinado, H., Alečković, M., Lavotshkin, S., Matei, I., Costa-Silva, B., Moreno-Bueno, G., Hergueta-Redondo, M., Williams, C., García-Santos, G., Ghajar, C.M., et al. (2012). Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nat. Med.* *18*, 883–891.

Percival Bailey and Harvey Cushing (1927). A classification of the tumours of the glioma group on a histogenetic basis, with a correlated study of prognosis. (Wiley).

Petryszak, R., Keays, M., Tang, Y.A., Fonseca, N.A., Barrera, E., Burdett, T., Füllgrabe, A., Fuentes, A.M.P., Jupp, S., Koskinen, S., et al. (2016). Expression Atlas update - An integrated database of gene and protein expression in humans, animals and plants. *Nucleic Acids Res.*

Piehowski, P.D., Zhu, Y., Bramer, L.M., Stratton, K.G., Zhao, R., Orton, D.J., Moore, R.J., Yuan, J., Mitchell, H.D., Gao, Y., et al. (2020). Automated mass spectrometry imaging of over 2000 proteins from tissue sections at 100- μ m spatial resolution. *Nat. Commun.* *11*.

Pin, E., Stratton, S., Belluco, C., Liotta, L., Nagle, R., Hodge, K.A., Deng, J., Dong, T., Baldelli, E., Petricoin, E., et al. (2016). A pilot study exploring the molecular architecture of the tumor microenvironment in human prostate cancer using laser capture microdissection and reverse phase protein microarray. *Mol. Oncol.* *10*, 1585.

Pols, M.S., and Klumperman, J. (2009). Trafficking and function of the tetraspanin CD63. *Exp. Cell Res.* *315*, 1584–1592.

Porter, A. (2012). A dead end: A review of glioblastoma multiforme. *Eukaryon* *8*, 64–68.

Quanico, J., Franck, J., Dauly, C., Strupat, K., Dupuy, J., Day, R., Salzet, M., Fournier, I., and Wisztorski, M. (2013). Development of liquid microjunction extraction strategy for improving protein identification from tissue sections. *J. Proteomics* *79*, 200–218.

Quanico, J., Franck, J., Gimeno, J.P., Sabbagh, R., Salzet, M., Day, R., and Fournier, I. (2015). Parafilm-assisted microdissection: A sampling method for mass spectrometry-based identification of differentially expressed prostate cancer protein biomarkers. *Chem. Commun.*

Quanico, J., Franck, J., Cardon, T., Leblanc, E., Wisztorski, M., Salzet, M., and Fournier, I. (2017a). NanoLC-MS coupling of liquid microjunction microextraction for on-tissue proteomic analysis. *Biochim. Biophys. Acta - Proteins Proteomics.*

Quanico, J., Franck, J., Wisztorski, M., Salzet, M., and Fournier, I. (2017b). Combined

MALDI mass spectrometry imaging and parafilm-assisted microdissection-based LC-MS/MS workflows in the study of the brain. In *Methods in Molecular Biology*, p.

Quanico, J., Franck, J., Wisztorski, M., Salzet, M., and Fournier, I. (2017c). Integrated mass spectrometry imaging and omics workflows on the same tissue section using grid-aided, parafilm-assisted microdissection. *Biochim. Biophys. Acta - Gen. Subj.* *1861*, 1702–1714.

Quaranta, M., Divella, R., Daniele, A., Di Tardo, S., Venneri, M.T., Lolli, I., and Troccoli, G. (2007). Epidermal Growth Factor Receptor Serum Levels and Prognostic Value in Malignant Gliomas. *Tumori J.* *93*, 275–280.

Rahman, M., Dastmalchi, F., Karachi, A., and Mitchell, D. (2019). The role of CMV in glioblastoma and implications for immunotherapeutic strategies. *Oncoimmunology*.

Raposo, G., and Stoorvogel, W. (2013). Extracellular vesicles: Exosomes, microvesicles, and friends. *J. Cell Biol.* *200*, 373–383.

Del Re, M., Marconcini, R., Pasquini, G., Rofi, E., Vivaldi, C., Bloise, F., Restante, G., Arrigoni, E., Caparello, C., Grazia Bianco, M., et al. (2018). PD-L1 mRNA expression in plasma-derived exosomes is associated with response to anti-PD-1 antibodies in melanoma and NSCLC. *Br. J. Cancer* *118*, 820–824.

Le Rhun, E., Duhamel, M., Wisztorski, M., Gimeno, J.-P., Zairi, F., Escande, F., Reyns, N., Kobeissy, F., Maurage, C.-A., Salzet, M., et al. (2016). Evaluation of non-supervised MALDI mass spectrometry imaging combined with microproteomics for glioma grade III classification. *Biochim. Biophys. Acta - Proteins Proteomics*.

Rivière Stéphanie, C.O.M.D. (2020). Suspicion d'excès de cas de glioblastomes dans les communes gardoises de Salindres et Rousson : mise à jour des données de surveillance et premières investigations de l'environnement. Rapport d'étape.

Rizzo, D.G., Prentice, B.M., Moore, J.L., Norris, J.L., and Caprioli, R.M. (2017). Enhanced Spatially Resolved Proteomics Using On-Tissue Hydrogel-Mediated Protein Digestion. *Anal. Chem.*

Rontogianni, S., Synadaki, E., Li, B., Liefwaard, M.C., Lips, E.H., Wesseling, J., Wu, W., and Altelaar, M. (2019). Proteomic profiling of extracellular vesicles allows for human breast cancer subtyping. *Commun. Biol.* *2*, 1–13.

Rose, M., Duhamel, M., Aboulouard, S., Kobeissy, F., Le Rhun, E., Desmons, A., Tierny, D., Fournier, I., Rodet, F., and Salzet, M. (2020). The Role of a Proprotein Convertase Inhibitor in Reactivation of Tumor-Associated Macrophages and Inhibition of Glioma Growth. *Mol. Ther. - Oncolytics* *17*, 31–46.

Sandfeld-Paulsen, B., Jakobsen, K.R., Bæk, R., Folkersen, B.H., Rasmussen, T.R., Meldgaard, P., Varming, K., Jørgensen, M.M., and Sorensen, B.S. (2016). Exosomal proteins as diagnostic biomarkers in lung cancer. *J. Thorac. Oncol.* *11*, 1701–1710.

Santiago-Dieppa, D.R., Steinberg, J., Gonda, D., Cheung, V.J., Carter, B.S., and Chen, C.C. (2014). Extracellular vesicles as a platform for 'liquid biopsy' in glioblastoma patients. *Expert Rev. Mol. Diagn.* *14*, 819–825.

Saudemont, P., Quanico, J., Robin, Y.M., Baud, A., Balog, J., Fatou, B., Tierny, D., Pascal, Q., Minier, K., Pottier, M., et al. (2018). Real-Time Molecular Diagnosis of Tumors Using Water-Assisted Laser Desorption/Ionization Mass Spectrometry Technology. *Cancer Cell* 34, 840-851.e4.

Scheithauer, B.W. (2009). Development of the WHO Classification of Tumors of the Central Nervous System: A Historical Perspective. *Brain Pathol.* 19, 551–564.

Schwartz, S.A., Weil, R.J., Johnson, M.D., Toms, S.A., and Caprioli, R.M. (2004). Protein Profiling in Brain Tumors Using Mass Spectrometry: Feasibility of a New Technique for the Analysis of Protein Expression. *Clin. Cancer Res.* 10, 981–987.

Ščupáková, K., Balluff, B., Tressler, C., Adelaja, T., Heeren, R.M.A., Glunde, K., and Ertaylan, G. (2020). Cellular resolution in clinical MALDI mass spectrometry imaging: The latest advancements and current challenges. *Clin. Chem. Lab. Med.* 58, 914–929.

Shen, F., Zhang, Y., Yao, Y., Hua, W., Zhang, H.S., Wu, J.S., Zhong, P., and Zhou, L.F. (2014). Proteomic analysis of cerebrospinal fluid: toward the identification of biomarkers for gliomas. *Neurosurg. Rev.* 37, 367–380.

Shevchenko, A., Tomas, H., Havliš, J., Olsen, J. V., and Mann, M. (2007). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* 1, 2856–2860.

Skog, J., Würdinger, T., van Rijn, S., Meijer, D.H., Gainche, L., Curry, W.T., Carter, B.S., Krichevsky, A.M., and Breakefield, X.O. (2008). Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. *Nat. Cell Biol.* 10, 1470–1476.

Spraggins, J.M., Rizzo, D.G., Moore, J.L., Noto, M.J., Skaar, E.P., and Caprioli, R.M. (2016). Next-generation technologies for spatial proteomics: Integrating ultra-high speed MALDI-TOF and high mass resolution MALDI FTICR imaging mass spectrometry for protein analysis. *Proteomics* 16, 1678–1689.

Stauber, J., MacAleese, L., Franck, J., Claude, E., Snel, M., Kaletas, B.K., Wiel, I.M.V.D., Wisztorski, M., Fournier, I., and Heeren, R.M.A. (2010). On-Tissue Protein Identification and Imaging by MALDI-Ion Mobility Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* 21, 338–347.

Stella, M., Chinello, C., Cazzaniga, A., Smith, A., Galli, M., Piga, I., Grasso, A., Grasso, M., Del Puppo, M., Varallo, M., et al. (2019). Histology-guided proteomic analysis to investigate the molecular profiles of clear cell Renal Cell Carcinoma grades. *J. Proteomics* 191, 38–47.

Stetson, L.C., Dazard, J.E., and Barnholtz-Sloan, J.S. (2016). Protein markers predict survival in glioma patients. *Mol. Cell. Proteomics* 15, 2356–2365.

Strojnik, T., Duh, D., and Lah, T.T. (2017). Prevalence of neurotropic viruses in malignant glioma and their onco-modulatory potential. *In Vivo (Brooklyn)*.

Stupp, R., Mason, W.P., Van Den Bent, M.J., Weller, M., Fisher, B., Taphoorn, M.J.B., Belanger, K., Brandes, A.A., Marosi, C., Bogdahn, U., et al. (2005). Radiotherapy plus concomitant and adjuvant temozolomide for glioblastoma. *N. Engl. J. Med.* 352, 987–996.

Stupp, R., Hegi, M.E., Mason, W.P., van den Bent, M.J., Taphoorn, M.J., Janzer, R.C., Ludwin, S.K., Allgeier, A., Fisher, B., Belanger, K., et al. (2009). Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *Lancet Oncol.* *10*, 459–466.

Stupp, R., Taillibert, S., Kanner, A., Read, W., Steinberg, D.M., Lhermitte, B., Toms, S., Idbaih, A., Ahluwalia, M.S., Fink, K., et al. (2017). Effect of tumor-treating fields plus maintenance temozolomide vs maintenance temozolomide alone on survival in patients with glioblastoma a randomized clinical trial. *JAMA - J. Am. Med. Assoc.*

Sturm, D., Witt, H., Hovestadt, V., Khuong-Quang, D.A., Jones, D.T.W., Konermann, C., Pfaff, E., Tönjes, M., Sill, M., Bender, S., et al. (2012). Hotspot Mutations in H3F3A and IDH1 Define Distinct Epigenetic and Biological Subgroups of Glioblastoma. *Cancer Cell.*

Svedman, F.C., Lohcharoenkal, W., Bottai, M., Brage, S.E., Sonkoly, E., Hansson, J., Pivarcsi, A., and Eriksson, H. (2018). Extracellular microvesicle microRNAs as predictive biomarkers for targeted therapy in metastatic cutaneous malignant melanoma. *PLoS One* *13*.

Szajnik, M., Czystowska, M., Szczepanski, M.J., Mandapathil, M., and Whiteside, T.L. (2010). Tumor-derived microvesicles induce, expand and up-regulate biological activities of human regulatory T cells (Treg). *PLoS One* *5*.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Bork, P., et al. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.*

Taillibert, S., Taillandier, L., and Le Rhun, E. (2015). Venous thrombosis in patients with high-grade glioma. *Curr. Opin. Oncol.* *27*, 516–521.

Taverna, D., Mignogna, C., Santise, G., Gaspari, M., and Cuda, G. (2019). On-Tissue Hydrogel-Mediated Nondestructive Proteomic Characterization: Application to fr/fr and FFPE Tissues and Insights for Quantitative Proteomics Using a Case of Cardiac Myxoma. *Proteomics - Clin. Appl.* *13*.

Taylor, D.D., and Gercel-Taylor, C. (2008). MicroRNA signatures of tumor-derived exosomes as diagnostic biomarkers of ovarian cancer. *Gynecol. Oncol.* *110*, 13–21.

Theodoraki, M.N., Yerneni, S., Gooding, W.E., Ohr, J., Clump, D.A., Bauman, J.E., Ferris, R.L., and Whiteside, T.L. (2019). Circulating exosomes measure responses to therapy in head and neck cancer patients treated with cetuximab, ipilimumab, and IMRT. *Oncoimmunology* *8*.

Théry, C., Witwer, K.W., Aikawa, E., Alcaraz, M.J., Anderson, J.D., Andriantsitohaina, R., Antoniou, A., Arab, T., Archer, F., Atkin-Smith, G.K., et al. (2018). Minimal information for studies of extracellular vesicles 2018 (MISEV2018): a position statement of the International Society for Extracellular Vesicles and update of the MISEV2014 guidelines. *J. Extracell. Vesicles* *7*.

Tillberg, P.W., Chen, F., Piatkevich, K.D., Zhao, Y., Yu, C.C., English, B.P., Gao, L., Martorell, A., Suk, H.J., Yoshida, F., et al. (2016a). Protein-retention expansion microscopy of

cells and tissues labeled using standard fluorescent proteins and antibodies. *Nat. Biotechnol.*

Tillberg, P.W., Chen, F., Piatkevich, K.D., Zhao, Y., Yu, C.-C., English, B.P., Gao, L., Martorell, A., Suk, H.-J., Yoshida, F., et al. (2016b). Protein-retention expansion microscopy of cells and tissues labeled using standard fluorescent proteins and antibodies. *Nat. Biotechnol.* *34*, 987–992.

Tivnan, A., Heilinger, T., Lavelle, E.C., and Prehn, J.H.M. (2017). Advances in immunotherapy for the treatment of glioblastoma. *J. Neurooncol.* *131*.

Treweek, J.B., Chan, K.Y., Flytzanis, N.C., Yang, B., Deverman, B.E., Greenbaum, A., Lignell, A., Xiao, C., Cai, L., Ladinsky, M.S., et al. (2015). Whole-body tissue stabilization and selective extractions via tissue-hydrogel hybrids for high-resolution intact circuit mapping and phenotyping. *Nat. Protoc.* *10*, 1860–1896.

Truckenbrodt, S., Maidorn, M., Crzan, D., Wildhagen, H., Kabatas, S., and Rizzoli, S.O. (2018). X10 expansion microscopy enables 25-nm resolution on conventional microscopes. *EMBO Rep.* *19*.

Turcan, S., Rohle, D., Goenka, A., Walsh, L.A., Fang, F., Yilmaz, E., Campos, C., Fabius, A.W.M., Lu, C., Ward, P.S., et al. (2012). IDH1 mutation is sufficient to establish the glioma hypermethylator phenotype. *Nature* *483*, 479–483.

Tyanova, S., Temu, T., Carlson, A., Sinitcyn, P., Mann, M., and Cox, J. (2015). Visualization of LC-MS/MS proteomics data in MaxQuant. *Proteomics* *15*, 1453–1456.

Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhori, G., Benfeitas, R., Arif, M., Liu, Z., Edfors, F., et al. (2017). A pathology atlas of the human cancer transcriptome. *Science* (80-.). *357*.

Valadi, H., Ekström, K., Bossios, A., Sjöstrand, M., Lee, J.J., and Lötvall, J.O. (2007). Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* *9*, 654–659.

Vanderperre, B., Lucier, J.F., and Roucou, X. (2012). HALtORF: A database of predicted out-of-frame alternative open reading frames in human. *Database*.

Vanderperre, B., Lucier, J.F., Bissonnette, C., Motard, J., Tremblay, G., Vanderperre, S., Wisztorski, M., Salzet, M., Boisvert, F.M., and Roucou, X. (2013). Direct Detection of Alternative Open Reading Frames Translation Products in Human Significantly Expands the Proteome. *PLoS One*.

Verhaak, R.G.W., Hoadley, K.A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M.D., Miller, C.R., Ding, L., Golub, T., Mesirov, J.P., et al. (2010). Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *Cancer Cell* *17*, 98–110.

Voigt, A., Nowick, K., and Almaas, E. (2017). A composite network of conserved and tissue specific gene interactions reveals possible genetic interactions in glioma. *PLoS Comput. Biol.*

Vollmann-Zwerenz, A., Leidgens, V., Feliciello, G., Klein, C.A., and Hau, P. (2020). Tumor

cell invasion in glioblastoma. *Int. J. Mol. Sci.* *21*, 1932.

Wang, P., Wang, H., Huang, Q., Peng, C., Yao, L., Chen, H., Qiu, Z., Wu, Y., Wang, L., and Chen, W. (2019). Exosomes from M1-polarized macrophages enhance paclitaxel antitumor activity by activating macrophages-mediated inflammation. *Theranostics* *9*, 1714–1727.

Wang, Q., Hu, B., Hu, X., Kim, H., Squatrito, M., Scarpace, L., deCarvalho, A.C., Lyu, S., Li, P., Li, Y., et al. (2017). Tumor Evolution of Glioma-Intrinsic Gene Expression Subtypes Associates with Immunological Changes in the Microenvironment. *Cancer Cell* *32*, 42–56.e6.

Webber, J., Steadman, R., Mason, M.D., Tabi, Z., and Clayton, A. (2010). Cancer exosomes trigger fibroblast to myofibroblast differentiation. *Cancer Res.* *70*, 9621–9630.

Weller, M., Pfister, S.M., Wick, W., Hegi, M.E., Reifenberger, G., and Stupp, R. (2013). Molecular neuro-oncology in clinical practice: a new horizon. *Lancet Oncol.* *14*, e370–e379.

Weller, M., Weber, R.G., Willscher, E., Riehm, V., Hentschel, B., Kreuz, M., Felsberg, J., Beyer, U., Löffler-Wirth, H., Kaulich, K., et al. (2015). Molecular classification of diffuse cerebral WHO grade II/III gliomas using genome- and transcriptome-wide profiling improves stratification of prognostically distinct patient groups. *Acta Neuropathol.* *129*, 679–693.

Weller, M., van den Bent, M., Tonn, J.C., Stupp, R., Preusser, M., Cohen-Jonathan-Moyal, E., Henriksson, R., Rhun, E. Le, Balana, C., Chinot, O., et al. (2017a). European Association for Neuro-Oncology (EANO) guideline on the diagnosis and treatment of adult astrocytic and oligodendroglial gliomas. *Lancet Oncol.* *18*, e315–e329.

Weller, M., Butowski, N., Tran, D.D., Recht, L.D., Lim, M., Hirte, H., Ashby, L., Mechtler, L., Goldlust, S.A., Iwamoto, F., et al. (2017b). Rindopepimut with temozolomide for patients with newly diagnosed, EGFRvIII-expressing glioblastoma (ACT IV): a randomised, double-blind, international phase 3 trial. *Lancet Oncol.*

Willms, E., Cabañas, C., Mäger, I., Wood, M.J.A., and Vader, P. (2018). Extracellular vesicle heterogeneity: Subpopulations, isolation techniques, and diverse functions in cancer progression. *Front. Immunol.* *9*.

Wirsching, H.G., Galanis, E., and Weller, M. (2016). Glioblastoma. In *Handbook of Clinical Neurology*, (Elsevier), pp. 381–397.

Wiśniewski, J.R. (2013). Proteomic Sample Preparation from Formalin Fixed and Paraffin Embedded Tissue. *J. Vis. Exp.*

Wisztorski, M., Lemaire, R., Stauber, J., Menguellet, S.A., Jardin-Mathé, O., Day, R., Salzet, M., and Fournier, I. (2007). Imagerie MALDI-Une nouvelle technologie pour découvrir et valider de nouveaux biomarqueurs. *Médecine/Sciences* *23*, 31–38.

Wisztorski, M., Fatou, B., Franck, J., Desmons, A., Farré, I., Leblanc, E., Fournier, I., and Salzet, M. (2013). Microproteomics by liquid extraction surface analysis: Application to FFPE tissue to study the fimbria region of tubo-ovarian cancer. *Proteomics - Clin. Appl.* *7*, 234–240.

Wisztorski, M., Desmons, A., Quanico, J., Fatou, B., Gimeno, J.-P., Franck, J., Salzet, M., and Fournier, I. (2016). Spatially-resolved protein surface microsampling from tissue sections using liquid extraction surface analysis. *Proteomics* *16*, 1622–1632.

Wisztorski, M., Quanico, J., Franck, J., Fatou, B., Salzet, M., and Fournier, I. (2017). Droplet-Based Liquid Extraction for Spatially-Resolved Microproteomics Analysis of Tissue Sections. In *Methods in Molecular Biology* (Clifton, N.J.), pp. 49–63.

Yan, H., Parsons, D.W., Jin, G., McLendon, R., Rasheed, B.A., Yuan, W., Kos, I., Batinic-Haberle, I., Jones, S., Riggins, G.J., et al. (2009). IDH1 and IDH2 mutations in gliomas. *N. Engl. J. Med.* *360*, 765–773.

Yu, H.A., Perez, L., Chang, Q., Gao, S.P., Kris, M.G., Riely, G.J., and Bromberg, J. (2017). A Phase 1/2 Trial of Ruxolitinib and Erlotinib in Patients with EGFR-Mutant Lung Adenocarcinomas with Acquired Resistance to Erlotinib. In *Journal of Thoracic Oncology*, (Elsevier Inc), pp. 102–109.

Yuryev, A., Kotelnikova, E., and Daraselia, N. (2009). Ariadne's ChemEffect and Pathway Studio knowledge base. *Expert Opin. Drug Discov.* *4*, 1307–1318.

Zavala-Vega, S., Palma-Lara, I., Ortega-Soto, E., Trejo-Solís, C., de Arellano, I.T.R., Ucharima-Corona, L.E., García-Chacón, G., Ochoa, S.A., Xicohtencatl-Cortes, J., Cruz-Córdova, A., et al. (2019). Role of epstein-barr virus in glioblastoma. *Crit. Rev. Oncog.*

Zhao, H., Achreja, A., Iessi, E., Logozzi, M., Mizzoni, D., Di Raimo, R., Nagrath, D., and Fais, S. (2018). The key role of extracellular vesicles in the metastatic process. *Biochim. Biophys. Acta - Rev. Cancer* *1869*, 64–77.

Zhu, Y., Dou, M., Piehowski, P.D., Liang, Y., Wang, F., Chu, R.K., Chrisler, W.B., Smith, J.N., Schwarz, K.C., Shen, Y., et al. (2018a). Spatially Resolved Proteome Mapping of Laser Capture Microdissected Tissue with Automated Sample Transfer to Nanodroplets. *Mol. Cell. Proteomics* *17*, 1864–1874.

Zhu, Y., Piehowski, P.D., Zhao, R., Chen, J., Shen, Y., Moore, R.J., Shukla, A.K., Petyuk, V.A., Campbell-Thompson, M., Mathews, C.E., et al. (2018b). Nanodroplet processing platform for deep and quantitative proteome profiling of 10-100 mammalian cells. *Nat. Commun.* *9*.

Zikou, A., Sioka, C., Alexiou, G.A., Fotopoulos, A., Voulgaris, S., and Argyropoulou, M.I. (2018). Radiation Necrosis, Pseudoprogression, Pseudoresponse, and Tumor Recurrence: Imaging Challenges for the Evaluation of Treated Gliomas. *Contrast Media Mol. Imaging* *2018*.

Zouaoui, S., Rigau, V., Mathieu-Daudé, H., Darlix, A., Bessaoud, F., Fabbro-Peray, P., Bauchet, F., Kerr, C., Fabbro, M., Figarella-Branger, D., et al. (2012). Recensement national histologique des tumeurs primitives du système nerveux central : Résultats généraux sur 40 000 cas, principales applications actuelles et perspectives. *Neurochirurgie* *58*, 4–13.

ANNEXES

CHAPITRE 1 : Protéines surexprimées dans les trois groupes

A

ID	Protein names	Gene names
Q96TC7	Regulator of microtubule dynamics protein RMDN3	RMDN3
O75569	Interferon-inducible double-stranded RNA-dPRKRA	dPRKRA
Q9NRW7	Vacuolar protein sorting-associated protein VPS45	VPS45
P05165-3	Propionyl-CoA carboxylase alpha chain, mit FCCA	FCCA
Q8NF91	Nesprin-1	SYNE1
P62140	Serine/threonine-protein phosphatase PP1-PPP1CB	PPP1CB
O76041-2	Nebulette	NEBL
P21266	Glutathione S-transferase Mu 3	GSTM3
P43007	Neutral amino acid transporter A	SLC1A4
P18433-6	Receptor-type tyrosine-protein phosphatase PTPRA	PTPRA
P43003-2	Excitatory amino acid transporter 1	SLC1A3
P21579	Synaptotagmin-1	SYT1
Q9NYB9-2	Abl interactor 2	ABI2
Q81W45	ATP-dependent (S)-NAD(P)H-hydrate dehydCARKD	CARKD
P43155-2	Carnitine O-acetyltransferase	CRAT
Q9Y6K8-3	Adenylate kinase isoenzyme 5	AK5
O14576-2	Cytoplasmic dynein 1 intermediate chain 1	DYNC1I1
P05067-10	Amyloid beta A4 protein	APP
P04216	Thy-1 membrane glycoprotein	THY1
O15394	Neural cell adhesion molecule 2	NCAM2
Q92752	Tenascin-R	TNR
Q9Y273	Guanine deaminase	GDA
P48163	NADP-dependent malic enzyme	ME1
Q9BWD1	Acetyl-CoA acetyltransferase, cytosolic	ACAT2
Q9H0B6-2	Kinesin light chain 2	KLC2
P16949	Stathmin	STMN1
P23528	Cofilin-1	CFL1
Q9ULU8-2	Calcium-dependent secretion activator 1	CADPS
O43815	Striatin	STRN
Q15818	Neuronal pentraxin-1	NPTX1
Q07866-7	Kinesin light chain 1	KLC1
Q3KQC3-2	MAP7 domain-containing protein 1	MAP7D1
Q9NQC3	Reticulon-4	RTN4
Q96RU3-4	Formin-binding protein 1	FNBP1
Q5JSH3-2	WD repeat-containing protein 44	WDR44
P46821	Microtubule-associated protein 1B	MAP1B
Q02952	A-kinase anchor protein 12	AKAP12
O60282	Kinesin heavy chain isoform 5C	KIF5C
P51970	NADH dehydrogenase [ubiquinone] 1 alpha NDUFAS8	NDUFAS8
Q99471	Prefoldin subunit 5	PFDN5
Q12959-4	Disks large homolog 1	DLG1

B

ID	Protein names	Gene names
O75955	Flotillin-1	FLOT1
Q14254	Flotillin-2	FLOT2
Q15223	Nectin-1	PVRL1
P80511	Protein S100-A12	S100A12
Q70U00	Inhibitor of nuclear factor kappa-B kinase-1 IKBIP	IKBIP
O15294-3	UDP-N-acetylglucosamine-peptide N-acetylOGT	OGT
P02788-2	Lactotransferrin	LTF
Q02809	Procollagen-lysine,2-oxoglutarate 5-dioxygePLOD1	PLOD1
P05556	Integrin beta-1	ITGB1
O00469	Procollagen-lysine,2-oxoglutarate 5-dioxygePLOD2	PLOD2
O14810	Complexin-1	CPLX1
P08236	Beta-glucuronidase	GUSB
P08133	Annexin A6	ANXA6
P35052	Glypican-1	GPC1
Q93077	Histone H2A type 1-C	HIST1H2AC
Q13557-8	Calcium/calmodulin-dependent protein kinase CAMK2D	CAMK2D
P06748-2	Nucleophosmin	NPM1
P35613-2	Basigin	BSG
Q9UHG3	Prenylcysteine oxidase 1	PCYOX1
Q9Y277	Voltage-dependent anion-selective channel VDAC3	VDAC3
P45880-2	Voltage-dependent anion-selective channel VDAC2	VDAC2
P21796	Voltage-dependent anion-selective channel VDAC1	VDAC1
P35232	Inhibitin	PHB
P02775	Platelet basic protein	PPBP
P35908	Keratin, type II cytoskeletal 2 epidermal	KRT2
O75348	V-type proton ATPase subunit G 1	ATP6V1G1
P27105	Erythrocyte band 7 integral membrane protein STOM	STOM
Q6PCB0	von Willebrand factor A domain-containingVWA1	VWA1
Q99715-4	Collagen alpha-1(XII) chain	COL12A1
Q9NZM1-6	Myoferlin	MYOF
Q81UX7	Adipocyte enhancer-binding protein 1	AEBP1
Q15582	Transforming growth factor-beta-induced TGFBI	TGFBI
P51884	Lumican	LUM
P12110	Collagen alpha-2(VI) chain	COL6A2
P21810	Biglycan	BGN
P51888	Prolargin	PRELP
P21980	Protein-glutamine gamma-glutamyltransferase TGM2	TGM2
P12109	Collagen alpha-1(VI) chain	COL6A1
P12111	Collagen alpha-3(VI) chain	COL6A3
Q16363-2	Laminin subunit alpha-4	LAMA4
Q14112-2	Nidogen-2	NID2

C

ID	Protein names	Gene names
Q02818	Nucleobindin-1	NUCB1
O15061	Synemin	SYNM
Q9H4A4	Aminopeptidase B	RNPEP
Q96CX2	BTB/POZ domain-containing protein KCTD1KCTD12	KCTD12
Q96A65	Exocyst complex component 4	EXOC4
O15031	Plexin-B2	PLXNB2
P09960	Leukotriene A-4 hydrolase	LTA4H
Q9P2R3	Rabankyrin-5	ANKFY1
P55039	Developmentally-regulated GTP-binding protein prDRG2	prDRG2
P08754	Guanine nucleotide-binding protein G(k) subunit GNAI3	GNAI3
AOAVT1	E3 UFM1-protein ligase 1	UFL1
O60443	Non-syndromic hearing impairment protein DFNA5	DFNA5
Q5VW32-2	BRO1 domain-containing protein BROX	BROX
Q15165-3	Serum paraoxonase/arylesterase 2	PON2
Q15631	Tranlin	TSN
Q7L775	EPM2A-interacting protein 1	EPM2AIP1
Q14974	Importin subunit beta-1	KPNB1
Q5T6V5	UPF0553 protein C9orf64	C9orf64
Q8WVM8	Sec 1 family domain-containing protein 1	SCFD1
O94874	E3 UFM1-protein ligase 1	UFL1
Q9Y678	Coatomer subunit gamma-1	COPG1
Q96K76-2	Ubiquitin carboxyl-terminal hydrolase 47	USP47
Q13131	5-AMP-activated protein kinase catalytic subunit PRKAA1	PRKAA1
Q9NZB2-4	Constitutive coactivator of PPAR-gamma-li FAM120A	FAM120A
Q96L92-3	Sorting nexin-27	SNX27
Q96D46	60S ribosomal export protein NMD3	NMD3
Q92616	Translational activator GCN1	GCN1L1
Q8TCU6	Phosphatidylinositol 3,4,5-trisphosphate-dPREX1	dPREX1
Q14160	Protein scribble homolog	SCRIB
P19174-2	1-phosphatidylinositol 4,5-bisphosphate p PLCG1	PLCG1
P78344	Eukaryotic translation initiation factor 4 gammaEIF4G2	EIF4G2
P22059	Oxysterol-binding protein 1	OSBP
Q7Z6Z7-2	E3 ubiquitin-protein ligase HUWE1	HUWE1
Q9BXB4	Oxysterol-binding protein-related protein 1	IOSBPL1
Q96AC1	Fermitin family homolog 2	FERMT2
P46939	Utrophin	UTRN
P21964-2	Catechol O-methyltransferase	COMT
P07384	Calpain-1 catalytic subunit	CAPN1
P17655	Calpain-2 catalytic subunit	CAPN2
Q9UBI6	Guanine nucleotide-binding protein G(I)/G GNG12	GNG12
Q14195-2		

Q9Y2X7	ARF GTPase-activating protein GIT1	GIT1
Q9UBC2	Epidermal growth factor receptor substrate EPS15L1	EPS15L1
Q14155-2	Rho guanine nucleotide exchange factor 7	ARHGEF7
Q9NVT9	Armadillo repeat-containing protein 1	ARMC1
Q9UD71	Protein phosphatase 1 regulatory subunit 1 PPP1R1B	PPP1R1B
Q92974-3	Rho guanine nucleotide exchange factor 2 ARHGEF2	ARHGEF2
Q60861-1	Growth arrest-specific protein 7	GAS7
Q14289-2	Protein-tyrosine kinase 2-beta	PTK2B
Q9BZV1	UBX domain-containing protein 6	UBXN6
Q05639	Elongation factor 1-alpha 2	EEF1A2
Q14141-2	Septin-6	SEPT6
Q9BPX5	Actin-related protein 2/3 complex subunit 5 SARPC5L	SARPC5L
P11137-3	Microtubule-associated protein 2	MAP2
P17677	Neuromodulin	GAP43
P8356	Phosphatidylinositol 5-phosphate 4-kinase PIP4K2B	PIP4K2B
Q8NC96	Adaptin ear-binding coat-associated proteinNECAP1	NECAP1
Q9NY65	Tubulin domain and HEAT repeat-containing protein TUBA8	TUBA8
Q99963-4	Endophilin-A3	SH3GL3
Q8N126	Cell adhesion molecule 3	CADM3
P54289	Voltage-dependent calcium channel subunit CACNA2D1	CACNA2D1
Q16623	Syntaxin-1A	STX1A
Q9UPY6	Wiskott-Aldrich syndrome protein family member WASF3	WASF3
Q92558	Wiskott-Aldrich syndrome protein family member WASF1	WASF1
Q6NV74	Uncharacterized protein KIAA1211-like KIAA1211L	KIAA1211L
Q5TF21	Protein SOGA3	SOGA3
Q02156	Protein kinase C epsilon type	PRKCE
O75145-2	Liprin-alpha-3	PPFIA3
P05771-2	Protein kinase C beta type	PRKCB
Q96P47	Arf-GAP with GTPase, ANK repeat and PH domainAGAP3	AGAP3
Q9P260-2	LisH domain and HEAT repeat-containing protein KIAA1468	KIAA1468
Q9Y2J0	Rabphilin-3A	RPH3A
Q9L109	NADH dehydrogenase [ubiquinone] 1 alpha NDUFAS12	NDUFAS12
Q8TC22-2	CD99 antigen-like protein 2	CD99L2
Q13367	AP-3 complex subunit beta-2	AP3B2
P56385	ATP synthase subunit e, mitochondrial	ATP5E1
P22695	Cytochrome b-c1 complex subunit 2, mitochondrial ATP6C2	ATP6C2
P20020-4	Plasma membrane calcium-transporting ATPase 2B1	ATP2B1
P18859	ATP synthase-coupling factor 6, mitochondrial ATP5J	ATP5J
P14927	Cytochrome b-c1 complex subunit 7	ATP6B7
P10606	Cytochrome c oxidase subunit 5B, mitochondrial COX5B	COX5B
O95169-3	NADH dehydrogenase [ubiquinone] 1 beta subunit NDUFB8	NDUFB8
O75489	NADH dehydrogenase [ubiquinone] iron-sulfur NDUF53	NDUF53
O43920	NADH dehydrogenase [ubiquinone] iron-sulfur NDUF55	NDUF55
O43181	NADH dehydrogenase [ubiquinone] iron-sulfur NDUF54	NDUF54
O43678	NADH dehydrogenase [ubiquinone] 1 alpha NDUFAS2	NDUFAS2
Q9UPR5	Sodium/calcium exchanger 2	SLC8A2

Q9BQI5	SH3-containing GRB2-like protein 3-interact	SGIP1	P02675	Fibrinogen beta chain	FGB	P43686	26S protease regulatory subunit 6B	PSMC4
Q8WXD9	Caskin-1	CASKIN1	P02671-2	Fibrinogen alpha chain	FGA	P62195	26S protease regulatory subunit 8	PSMC5
Q13825	Methylglutaconyl-CoA hydratase, mitochondr	AUH	P02656	Apolipoprotein C-III	APOC3	Q9Y265	RuvB-like 1	RUVBL1
P30044	Peroxisredoxin-5, mitochondrial	PRDX5	P02652	Apolipoprotein A-II	APOA2	Q9Y230	RuvB-like 2	RUVBL2
P09622-3	Dihydropolyl dehydrogenase, mitochondria	DLDD	P02647	Apolipoprotein A-I	APOA1	Q9H488	GDP-fucose protein O-fucosyltransferase	IPOFUT1
P17600-2	Synapsin-1	SYN1	P01024	Complement C3	C3	P47897	Glutamine-tRNA ligase	QARS
P62760	Visinin-like protein 1	VSNL1	P01023	Alpha-2-macroglobulin	A2M	P35606-2	Coatomer subunit beta	COPB2
P00492	Hypoxanthine-guanine phosphoribosyltran	HPRT1	P01011	Alpha-1-antichymotrypsin	SERPINA3	O60763	General vesicular transport factor p115	USO1
P15104	Glutamine synthetase	GLUL	P01009	Alpha-1-antitrypsin	SERPINA1	O60762	Dolichol-phosphate mannosyltransferase	sDPM1
P80723	Brain acid soluble protein 1	BASP1	P01008	Antithrombin-III	SERPINC1	O00203-3	AP-3 complex subunit beta-1	AP3B1
P49840	Glycogen synthase kinase-3 alpha	GSK3A	P00734	Prothrombin	F2	O14579	Coatomer subunit epsilon	COPE
Q6ZVM7	TOM1-like protein 2	TOM1L2	P00751	Complement factor B	CFB	Q9NYU2-2	UDP-glucose:glycoprotein glucosyltransferase	UGGT1
P05026-2	Sodium/potassium-transporting ATPase sub	ATP1B1	P30043	Flavin reductase (NADPH)	BLVRB	P61923	Coatomer subunit zeta-1	COPZ1
Q12860	Contactin-1	CNTN1	P0C0L5	Complement C4-B	C4B	P53618	Coatomer subunit beta	COPB1
P09493-5	Tropomyosin alpha-1 chain	TPM1	P00450	Ceruloplasmin	CP	O00116	Alkyldihydroxyacetonephosphate synthase	AGPS
Q9NZN3	EH domain-containing protein 3	EHD3	P06702	Protein S100-A9	S100A9	P04844-2	Dolichyl-diphosphooligosaccharide--protei	RPN2
Q9Y5K8	V-type proton ATPase subunit D	ATP6V1D	Q14108	Lysosome membrane protein 2	SCARB2	Q9Y262-2	Eukaryotic translation initiation factor 3 su	EIF3L
P63010	AP-2 complex subunit beta	AP2B1	P19827	Inter-alpha-trypsin inhibitor heavy chain	HITTH1	Q92973-2	Transpirtin-1	TNPO1
P42025	Beta-actin	ACTR1B	P02743	Serum amyloid P-component	APCS	P78527	DNA-dependent protein kinase catalytic su	PRKDC
P38606	V-type proton ATPase catalytic subunit A	ATP6V1A	Q43169	Cytochrome b5 type B	CYB5B	Q14980	Exportin-1	XPO1
P36543	V-type proton ATPase subunit E 1	ATP6V1E1	P01619	Ig kappa chain V-III region B6		P62495	Eukaryotic peptide chain release factor sub	ETP1
P21283	V-type proton ATPase subunit C 1	ATP6V1C1	P36222	Chitinase-3-like protein 1	CH3L1	Q7L576	Cytoplasmic FMR1-interacting protein 1	CYFIP1
P19404	NADH dehydrogenase [ubiquinone] flavopr	NDUFV2	P28676	Grancalcin	GCA	Q16401-2	26S proteasome non-ATPase regulatory su	PSMD5
P19367	Hexokinase-1	HK1	P04196	Histidine-rich glycoprotein	HRG	P60228	Eukaryotic translation initiation factor 3 su	EIF3E
O95782-2	AP-2 complex subunit alpha-1	AP2A1	P40404	Vitronectin	VTN	P54136-2	Arginine-tRNA ligase, cytoplasmic	RARS
P17174	Aspartate aminotransferase, cytoplasmic	GOT1	P02748	Complement component C9	C9	P53621	Coatomer subunit alpha	COPA
Q9UL15	Transgelin-3	TAGLN3	P02679-2	Fibrinogen gamma chain	FGG	O00303	Eukaryotic translation initiation factor 3 su	EIF3F
Q96GD0	Pyridoxal phosphate phosphatase	PDXP	P02654	Apolipoprotein C-I	APOC1	P16435	NADPH-cytochrome P450 reductase	POR
O76070	Gamma-synuclain	SNCG	P01031	Complement C5	C5	Q9Y3Z3	Deoxynucleoside triphosphate triphospho	SAMHD1
Q14894	Ketimine reductase mu-crystallin	CRYM	O00560	Syntenin-1	SDCBP	Q7KZF4	Staphylococcal nuclease domain-containin	SND1
Q9NQ66	1-phosphatidylinositol 4,5-bisphosphate	PLCB1	P00747	Plasminogen	PLG	Q14152	Eukaryotic translation initiation factor 3 su	EIF3A
Q96F85	CB1 cannabinoid receptor-interacting prote	CNRIP1	P02787	Serotransferrin	TF	P62701	40S ribosomal protein S4, X isoform	RPS4X
Q13555-5	Calcium/calmodulin-dependent protein kin	CAMK2G	P07738	Bisphosphoglycerate mutase	BPGM	P61313	60S ribosomal protein L15	RPL15
P42658	Dipeptidyl aminopeptidase-like protein 6	DPPE	P02766	Transthyretin	TTR	P15880	40S ribosomal protein S2	RPS2
P59768	Guanine nucleotide-binding protein G(I)/G	GNG2	P02763	Alpha-1-acid glycoprotein 1	ORM1	P60842	Eukaryotic initiation factor 4A-1	EIF4A1
Q9UPY8	Microtubule-associated protein RP/EB fami	MAPRE3	P01834	Ig kappa chain C region	IGKC	Q8WVY7	Ubiquitin-like domain-containing CTD phos	UBCLP1
Q9COC9	E2/E3 hybrid ubiquitin-protein ligase	UBE2OUBE2O	P01857	Ig gamma-1 chain C region	IGHG1	P51648	Fatty aldehyde dehydrogenase	ALDH3A2
Q12955	Ankyrin-3	ANK3	P08603	Complement factor H	CFH	P40763-2	Signal transducer and activator of transcri	STAT3
P05413	Fatty acid-binding protein, heart	FABP3	P04217	Alpha-1B-glycoprotein	A1BG	P27708	CAD protein	CAD
O95197-2	Reticulon-3	RTN3	P02774	Vitamin D-binding protein	GC	P11172	Uridine 5-monophosphate synthase	UMPS
O15020	Spectrin beta chain, non-erythrocytic 2	SPTBN2	P01876	Ig alpha-1 chain C region	IGHA1	O95373	Importin-7	IPO7
Q75061-4	Putative tyrosine-protein phosphatase auxi	DNAJC6	P01871	Ig mu chain C region	IGHM	P10155-3	60 kDa SS-A/Ro ribonucleoprotein	TROVE2
Q91UB6-2	Neurochordin	NCDN	P01860	Ig gamma-3 chain C region	IGHG3	Q9Y224	UPF0568 protein C14orf166	C14orf166
Q8IUL8-2	Calcium/calmodulin-dependent protein kin	CAMK1D	P01859	Ig gamma-2 chain C region	IGHG2	Q9UKM9	RNA-binding protein Raly	RALY
O94856-6	Neurofascin	NFASC	P0D0Y3	Ig lambdaB-6 chain C region	IGLC6	Q9UBT2	SUMO-activating enzyme subunit 2	UBA2
Q13554	Calcium/calmodulin-dependent protein kin	CAMK2B	P01042-2	Kinogenin-1	KNG1	Q2TAY7	WD40 repeat-containing protein SMU1	SMU1
Q8TB36	Ganglioside-induced differentiation associa	GDAP1	P00738	Haptoglobin	HP	Q13185	Chromobox protein homolog 3	CBX3
Q9NY10-3	PH and SEC7 domain-containing protein 3	PSD3	P02790	Hemopexin	HPX	Q12905	Interleukin enhancer-binding factor 2	ILF2
Q9BY11	Protein kinase C and casein kinase substrate	PACSIN1	Q08380	Galectin-3-binding protein	LGALS3BP	P42785	Lysosomal Pro-X carboxypeptidase	PRCP
Q7ZL60-3	Proline-rich transmembrane protein 2	PRRT2	Q9Y4L1	Hypoxia up-regulated protein 1	HYOU1	P26599	Polypyrimidine tract-binding protein 1	PTBPI
Q08495	Dematin	DMTN	P35998	26S protease regulatory subunit 7	PSMC7	P22102	Trifunctional purine biosynthetic protein a	GART
Q01814-6	Plasma membrane calcium-transporting AT	ATP2B2	P26196	Probable ATP-dependent RNA helicase	DDXDDX6	O60506	Heterogeneous nuclear ribonucleoprotein	SYNCRIP
P19086	Guanine nucleotide-binding protein G(z) su	GNAZ	Q9NRH3	Tubulin gamma-2 chain	TUBG2	P09874	Poly[ADP-ribose] polymerase 1	PARP1
O60641-4	Clathrin coat assembly protein AP180	SNAP19	O94979-3	Protein transport protein Sec31A	SEC31A	Q14103-4	Heterogeneous nuclear ribonucleoprotein	HNRNPD
P04075	Fructose-bisphosphate aldolase A	ALDOA	O00231	26S proteasome non-ATPase regulatory sub	PSMD11	Q13435	Splicing factor 3B subunit 2	SF3B2
Q91PA5	Protein bassoon	BSN	O00232	26S proteasome non-ATPase regulatory sub	PSMD12	P05455	Lap5 La protein	SSB
Q9ULHG2	ProSAAS	PCSK1N	Q9BT78	COP9 signalosome complex subunit 4	COP94	P51858-2	Hepatoma-derived growth factor	HDFG
O60268-3	Uncharacterized protein KIAA0513	KIAA0513	Q92878	DNA repair protein RAD50	RAD50	P62995-3	Transformer-2 protein homolog beta	TRA2B
Q92777	Synapsin-2	SYN2	Q13200	26S proteasome non-ATPase regulatory sub	PSMD2	P46783	40S ribosomal protein S10	RPS10
O95670	V-type proton ATPase subunit G 2	ATP6V1G2	O43242	26S proteasome non-ATPase regulatory sub	PSMD3	Q61AA8	Regulator complex protein LAMTOR1	LAMTOR1
P37840-2	Alpha-synuclain	SNCA	O00487	26S proteasome non-ATPase regulatory sub	PSMD14	P83731	60S ribosomal protein L24	RPL24
Q99962	Endophilin-A1	SH3GL2	O00505	Importin subunit alpha-4	KPNA3	P62913	60S ribosomal protein L11	RPL11
Q96J9E9	Microtubule-associated protein 6	MAP6	Q9UNM6	26S proteasome non-ATPase regulatory sub	PSMD13	P62857	40S ribosomal protein S28	RPS28
P49418	Amphiphysin	AMPH	Q99623	Inhibitor-2	PHB2	P62750	60S ribosomal protein L23a	RPL23A
Q16143	Beta-synuclain	SNCB	Q15436	Protein transport protein Sec23A	SEC23A	P62263	40S ribosomal protein S14	RPS14
P61764-2	Syntaxin-binding protein 1	STXBP1	P35580	Myosin-10	MYH10	P62249	40S ribosomal protein S16	RPS16
P61266-2	Syntaxin-1B	STXB1B	P48556	26S proteasome non-ATPase regulatory sub	PSMD8	P62241	40S ribosomal protein S8	RPS8
P60880	Synaptosomal-associated protein 25	SNAP25	Q15008	26S proteasome non-ATPase regulatory sub	PSMD6	P57772	Selenocysteine-specific elongation factor	EEFSEC
P13637	Sodium/potassium-transporting ATPase sub	ATP1A3	Q8N766-4	ER membrane protein complex subunit 1	EMC1	P46778	60S ribosomal protein L21	RPL21
P17600	Synapsin-1	SYN1	O60613	15 kDa selenoprotein		P39019	40S ribosomal protein S19	RPS19
Q9UQB8-2	Brain-specific angiogenesis inhibitor 1-asc	BIAIP2	Q8NE71-2	ATP-binding cassette sub-family F member	ABCF1	P20042	Eukaryotic translation initiation factor 2 su	EIF2S2
Q8N573-8	Oxidation resistance protein 1	ORX1	Q9P012	ER membrane protein complex subunit 3	EMC3	O75822-3	Eukaryotic translation initiation factor 3 su	EIF3J
Q13526	Peptidyl-prolyl cis-trans isomerase NIMA-in	PINI	P02786	Transferrin receptor protein 1	TFR1	P18621	60S ribosomal protein L17	RPL17
Q13153	Serine/threonine-protein kinase PAK 1	PAK1	Q96A33-2	Coiled-coil domain-containing protein 47	CCDC47	Q9Y653-2	G-protein coupled receptor 56	GPR56
P55327-4	Tumor protein D52	TPD52	Q9UHB9-4	Signal recognition particle subunit SRP68	SRP68	Q99459	Cell division cycle 5-like protein	CDC5L
O00154-4	Cytosolic acyl coenzyme A thioester hydro	ACOT7	O60610-2	Protein diaphanous homolog 1	DIAPH1	Q92945	Fat upstream element-binding protein 2	KHSRP
P09936	Ubiquitin carboxyl-terminal hydrolase isoz	UCHL1	O76094	Signal recognition particle subunit SRP72	SRP72	Q55835	Heterochromatin protein 1-binding protein	HP1BP3
Q9Y411-2	Unconventional myosin-Va	MYO5A	Q01469	Fatty acid-binding protein, epidermal	FABP5	Q13442	28 kDa heat- and acid-stable phosphoprot	PDAP1
Q9NR46-2	Endophilin-B2	SH3GLB2	P12814-3	Alpha-actinin-1	ACTN1	Q13177	Serine/threonine-protein kinase PAK 2	PAK2
Q9H115	Beta-soluble NSF attachment protein	NAPB	P05362	Intercellular adhesion molecule 1	ICAM1	P22626	Heterogeneous nuclear ribonucleoproteins	HNRNP2B1
Q05193-2	Dynamitin-1	DNM1	P02792	Ferritin light chain	FTL	Q00839	Heterogeneous nuclear ribonucleoprotein	HNRNPV
P68366-2	Tubulin alpha-4A chain	TUBA4A	P04179	Superoxide dismutase [Mn], mitochondrial	SOD2	Q9BZZ5-2	Apoptosis inhibitor 5	API5
P12532	Creatine kinase U-type, mitochondrial	CKMT1A	Q96HE7	ERO1-like protein alpha	ERO1L	Q92841	Probable ATP-dependent RNA helicase	DDXDDX17
O94925	Glutaminase kidney isoform, mitochondrial	GLS	P00338	L-lactate dehydrogenase A chain	LDHA	Q8N163-2	Cell cycle and apoptosis regulator protein	2CCAR2
O95248	Myotubularin-related protein 5	SBF1	P00558	Phosphoglycerate kinase 1	PGK1	Q86V81	THO complex subunit 4	ALYRF8
Q9Y697-2	Cysteine desulfurase, mitochondrial	NFS1	Q16851	UTP--glucose-1-phosphate uridylyltransfe	UGP2	Q52LJ0-2	Protein FAM98B	FAM98B
Q9ULQ6	Calcium/calmodulin-dependent protein kin	CAMK2A	Q04446	1,4-alpha-glucan-branching enzyme	GBE1	Q1KMD3	Heterogeneous nuclear ribonucleoprotein	HNRNPUL2
Q9ULQ16-2	Dynamitin-3	DNM3	P40121	Macrophage-capping protein	CAPG	Q15717	ELAV-like protein 1	ELAVL1
Q9ULP0-3	Protein NDRG4	NDRG4	P16070-16	CD44 antigen	CD44	Q15233	Non-POU domain-containing octamer-bind	NONO
Q9UL12-2	V-type proton ATPase subunit H	ATP6V1H	P06737-2	Glycogen phosphorylase, liver form	PYGL	Q13263	Transcription intermediary factor 1-beta	TRIM28
Q9NZR1	Tropomodulin-2	TMOD2	P13807-2	Glycogen [starch] synthase, muscle	GYSI	Q12906-7	Interleukin enhancer-binding factor 3	ILF3
Q99747	Gamma-soluble NSF attachment protein	NAPG	Q9NY33	Dipeptidyl peptidase 3	PPP3	Q08211	ATP-dependent RNA helicase A	DHX9
Q8WXF7-2	Atlastin-1	ATL1	Q8WX93-3	Palladin	PALLD	Q01844-6	RNA-binding protein EWS	EWSR1
Q14168-6	MAGUK p55 subfamily member 2	MPP2	Q14258	E3 ubiquitin/ISG15 ligase TRIM25	TRIM25	P61978-3	Heterogeneous nuclear ribonucleoprotein	HNRNPK
Q12979-4	Active breakpoint cluster region-related	proABR	P48681	Nestin	NES	P52272-2	Heterogeneous nuclear ribonucleoprotein	HNRNPM
Q08209	Serine/threonine-protein phosphatase 2B c	PPP3CA	P37802	Transgelin-2	TAGLN2	P38159	RNA-binding motif protein, X chromosome	RBMX
Q02750-2	Dual specificity mitogen-activated protein	kin MAP2K1	P26038	Moessin	MSN	P35637-2	RNA-binding protein FUS	FUS
Q01813	ATP-dependent 6-phosphofructokinase, pla	PFKP	P08670	Vimentin	VIM	P31943	Heterogeneous nuclear ribonucleoprotein	HNRNPH1
P63027	Vesicle-associated membrane protein 2	VAMP2	P04083	Annexin A1	ANXA1	P23246-2	Splicing factor, proline- and glutamine-rich	SFPQ
P62873	Guanine nucleotide-binding protein G(I)/G	GNB1	P07355	Annexin A2	ANXA2	P19338	Nucleolin	NCL
P46459	Vesicle-fusing ATPase	NSF	Q9UL46	Proteasome activator complex subunit 2	PSME2	P13010	X-ray repair cross-complementing protein	XRCC5

P35612	Beta-adducin	ADD2	Q06323	Proteasome activator complex subunit 1	PSME1	P09651	Heterogeneous nuclear ribonucleoprotein HNRNPA1
P21281	V-type proton ATPase subunit B, brain isoform	ATP6V1B2	P07203	Glutathione peroxidase 1	GPX1	P05643	U5 small nuclear ribonucleoprotein 200 kD SNRNP200
P09471	Guanine nucleotide-binding protein G(o) subunit	GNAO1	O00592-2	Podocalyxin	PODXL	O43390	Heterogeneous nuclear ribonucleoprotein HNRNPR
O43426-2	Synaptotagmin-1	SYNJ1	P01892	HLA class I histocompatibility antigen, A-2 alpha	A-2 alpha	O75533	Splicing factor 3B subunit 1 SF3B1
P00505	Aspartate aminotransferase, mitochondrial	GOT2	Q9Y394-2	Dehydrogenase/reductase SDR family member	DHRS7	Q13151	Heterogeneous nuclear ribonucleoprotein HNRNPAO
Q9Y613-3	Epsin-1	EPN1	Q9UBS4	DnaJ homolog subfamily B member 11	DNAJB11	Q13148	TAR DNA-binding protein 43 TARDBP
Q9Y4G8	Rap guanine nucleotide exchange factor 2	RAPGEF2	Q9BVK6	Transmembrane emp24 domain-containing TMED9		P51991	Heterogeneous nuclear ribonucleoprotein HNRNPA3
Q99719	Septin-5	SEPT5	Q13724-2	Mannosyl-oligosaccharide glucosidase	MOGS	P43243	Matrin-3 MATR3
Q2M218	AP2-associated protein kinase 1	AAK1	Q5JPE7-2	Nodal modulator 2	NOMO2	P19525	Interferon-induced, double-stranded RNA-EIF2AK2
Q16799	Reticulon-1	RTN1	P51571	Translocin-associated protein subunit delta	SSR4	P17844	Probable ATP-dependent RNA helicase DDX3X
Q04917	14-3-3 protein eta	YWHAH	P24821	Tenascin	TNC	O14979-3	Heterogeneous nuclear ribonucleoprotein HNRNPDL
P78324-4	Tyrosine-protein phosphatase non-receptor	SIRPA	P11021	78 kDa glucose-regulated protein	HSPA5	O43615	Mitochondrial import inner membrane translocator
P45985	Dual specificity mitogen-activated protein kinase	MAP2K4	O75369-6	Flamin-B	FLNB	Q9P2E9	Ribosome-binding protein 1 RRBP1
P40925	Malate dehydrogenase, cytoplasmic	MDH1	P02751	Fibronectin	FN1	Q9C0C2	182 kDa tankyrase-1-binding protein TNKS1BP1
O75781	Paralemmin-1	PALM	Q8NB59	Thioredoxin domain-containing protein 5	TXNDC5	Q00341	Vigilin HDLBP
O60331-4	Phosphatidylinositol 4-phosphate 5-kinase	PIP5K1C	Q15084-3	Protein disulfide-isomerase A6	PDI6A	O95793-2	Double-stranded RNA-binding protein Stau1
O75382-4	Tripartite motif-containing protein 3	TRIM3	Q12797-10	Aspartyl/asparaginyl beta-hydroxylase	ASPH	O60493	Sorting nexin-3 SNX3
P30084	Enoyl-CoA hydratase, mitochondrial	ECHS1	P30101	Protein disulfide-isomerase A3	PDI3A	O75351	Vacuolar protein sorting-associated protein VPS4B
Q16204	Coiled-coil domain-containing protein 6	CDC6	P14314-2	Glucosidase 2 subunit beta	PRKCSH	Q9NUU7	ATP-dependent RNA helicase DDX19A
Q9Y2J2	Band 4.1-like protein 3	EPB41L3	P13667	Protein disulfide-isomerase A4	PDI4A	Q9H6R3	Acyl-CoA synthetase short-chain family member
Q8IZP0-10	Abl interactor 1	ABI1	P13639	Elongation factor 2	EEF2	Q9H3P7	Golgi resident protein GCP60
Q9Y2HI	Serine/threonine-protein kinase 38-like	STK38L	P07237	Protein disulfide-isomerase	P4HB	Q96PU8-9	Protein quaking
Q96F07-2	Cytoplasmic FMRI-interacting protein 2	CYFIP2	O43852	Calumenin	CALU	Q96I24	Far upstream element-binding protein 3 FUBP3
P17252	Protein kinase C alpha type	PKCA	O95881	Thioredoxin domain-containing protein 12	TXNDC12	Q13283	Ras GTPase-activating protein-binding protein
Q9HDQ0	Protein FAM49A	FAM49A	Q9Y3E5	Peptidyl-L-tyrosine hydrolase 2, mitochondrial	PTRH2	P68104	Elongation factor 1-alpha 1 EEF1A1
Q14194-2	Dihydropyrimidinase-related protein 1	CRMP1	Q9UBF8-2	Phosphatidylinositol 4-kinase beta	P4KB	P67809	Nuclease-sensitive element-binding protein YBX1
P56211	cAMP-regulated phosphoprotein 19	ARPP19	Q9BVC6	Transmembrane protein 109	TMEM109	P17980	26S protease regulatory subunit 6A
P10636-5	Microtubule-associated protein tau	MAPT	Q96J77	Protein disulfide-isomerase TMX3	TMX3	O60841	Eukaryotic translation initiation factor 5B
O14531	Dihydropyrimidinase-related protein 4	DPYSL4	Q96A73	Peptidyl-prolyl cis-trans isomerase FKBP10	FKBP10	P11586	C-1-tetrahydrofolate synthase, cytoplasmic
P10636-2	Microtubule-associated protein tau	MAPT	Q8TCJ2	Dolichyl-diphosphooligosaccharide-protein	STT3B	Q86UP2	Kinecin
Q13885	Tubulin beta-2A chain	TUBB2A	Q07065	Cytoskeleton-associated protein 4	CKAP4	Q13596	Sorting nexin-1 SNX1
Q9LUW0	Dynactin subunit 4	DCTN4	P54709	Sodium/potassium-transporting ATPase subunit	ATP1B3	Q07020-2	60S ribosomal protein L18
Q9BPU6	Dihydropyrimidinase-related protein 5	DPYSL5	P51665	26S proteasome non-ATPase regulatory subunit	PSMD7	Q04637-5	Eukaryotic translation initiation factor 4 gamma
Q92599-4	Septin-8	SEPT8	P49755	Transmembrane emp24 domain-containing TMED10		P62333	26S proteasome regulatory subunit 10B
Q16181-2	Septin-7	SEPT7	P49257	Protein ERGIC-53	LMAN1	P61353	60S ribosomal protein L27
Q14203-6	Dynactin subunit 1	DCTN1	P39656-3	Dolichyl-diphosphooligosaccharide-protein	DDOST	P60866	40S ribosomal protein S20
Q13561	Dynactin subunit 2	DCTN2	P14625	Endoplasmic reticulum chaperone	HSP90B1	P55209-2	Nucleosome assembly protein 1-like 1
P78559	Microtubule-associated protein 1A	MAP1A	P13674	Prolyl 4-hydroxylase subunit alpha-1	P4HA1	P53992	Protein transport protein Sec24C
O14964-2	Hepatocyte growth factor-regulated tyrosin	HGS	O04843	Dolichyl-diphosphooligosaccharide-protein	RPN1	P39023	60S ribosomal protein L3
P55011-3	Solute carrier family 12 member 2	SLC12A2	P11117	Lysosomal acid phosphatase	ACP2	P36578	60S ribosomal protein L4
Q98FW1	Ubiquitin thioesterase	OTUB1	Q9Y490	Talin-1	TLN1	P26373	60S ribosomal protein L13
O6NBX0	Saccharopine dehydrogenase-like oxidoreduct	SCCPDH	Q9Y315	Deoxyribose-phosphate aldolase	DERA	P05388	60S acidic ribosomal protein P0
P68363	Tubulin alpha-1B chain	TUBA1B	P50454	Serpin H1	SERPINH1	O00571	ATP-dependent RNA helicase DDX3X
P36915	Guanine nucleotide-binding protein-like 1	GNL1	P46940	Ras GTPase-activating-like protein IQGAP1	IQGAP1	O60716-8	Catenin delta-1
P30153	Serine/threonine-protein phosphatase 2A	PPP2R1A	P43490	Nicotinamide phosphoribosyltransferase	NAMPT	Q99613	Eukaryotic translation initiation factor 3 subunit
P27361	Mitogen-activated protein kinase 3	MAPK3	P21333-2	Filamin-A	FLNA	Q14444-2	Caprin-1
O75122-3	CLIP-associated protein 2	CLASP2	P18206-2	Vinculin	VCL	Q02878	60S ribosomal protein L6
P07437	Tubulin beta chain	TUBB	P07737	Profilin-1	PFN1	P62851	40S ribosomal protein S25
Q9C040	Tripartite motif-containing protein 2	TRIM2	P04275	von Willebrand factor	VWF	P62424	60S ribosomal protein L7a
Q7L099-3	Protein RUFY3	RUFY3	O00299	Chloride intracellular channel protein 1	CLIC1	P30050	60S ribosomal protein L12
Q13509	Tubulin beta-3 chain	TUBB3	P00488	Coagulation factor XIII A chain	F13A1	P22087	rRNA 2-0-methyltransferase fibrillarin
P63208	S-phase kinase-associated protein 1	SKP1	Q9HDC9	Adipocyte plasma membrane-associated	prAPMAP	O75821	Eukaryotic translation initiation factor 3 subunit
P68371	Tubulin beta-4B chain	TUBB4B	Q99829	Copine-1	CPNE1	P11940	Polyadenylate-binding protein 1
Q9Y285	Phenylalanine-tRNA ligase alpha subunit	FARSA	P06600-2	Myosin light polypeptide 6	MYL6	P62277	40S ribosomal protein S13
Q9LUHY1	Nuclear receptor-binding protein	NKBP1	P52566	Rho GDP-dissociation inhibitor 2	ARHGDB2	P62269	40S ribosomal protein S18
Q9NSD9	Phenylalanine-tRNA ligase beta subunit	FARSB	P35579	Myosin-9	MYH9	P46781	40S ribosomal protein S9
Q12931	Heat shock protein 75 kDa, mitochondrial	TRAP1	O75165	DnaJ homolog subfamily C member 13	DNAJC13	P40429	60S ribosomal protein L13a
Q8N111	Cell cycle exit and neuronal differentiation	pCEND1	P27797	Calreticulin	CALR	P26641	Elongation factor 1-gamma
P12235	ADP/ATP translocase 1	SLC25A4	P30040	Endoplasmic reticulum resident protein 29	ERP29	P18124	60S ribosomal protein L7
Q13423	NAD(P) transhydrogenase, mitochondrial	NNT	P07686	Beta-hexosaminidase subunit beta	HEXB	P23396	40S ribosomal protein S3
P50213	Isocitrate dehydrogenase [NAD] subunit	alpIDH3A	P17931	Galectin-3	LGALS3	P62847-2	40S ribosomal protein S24
P51553	Isocitrate dehydrogenase [NAD] subunit	gamIDH3G	Q01518	Adenylyl cyclase-associated protein 1	CAP1	P98179	RNA-binding protein 3
Q02218	2-oxoglutarate dehydrogenase, mitochondri	OGDH	P61158	Actin-related protein 3	ACTR3	Q02543	60S ribosomal protein L18a
P49821-2	NADH dehydrogenase [ubiquinone] flavop	NDUFV1	O15511	Actin-related protein 2/3 complex subunit	SARPC5	P62753	40S ribosomal protein S6
P28331-4	NADH-ubiquinone oxidoreductase 75 kDa	NDUFS1	P52907	F-actin-capping protein subunit alpha-1	CAPZA1	P27635	60S ribosomal protein L10
P08559	Pyruvate dehydrogenase E1 component	subPDHA1	Q15404	Ras suppressor protein 1	RSP1	P46777	60S ribosomal protein L5
O95299	NADH dehydrogenase [ubiquinone] 1 alpha	NDUFA10	P50552	Vasodilator-stimulated phosphoprotein	VASP	Q6P209	General transcription-spllicing factor 8
O75947	ATP synthase subunit d, mitochondrial	ATP5H	P09601	Heme oxygenase 1	HMOX1	P78347-2	General transcription factor II-I
O94826	Mitochondrial import receptor subunit	TOMTM70A	P13796	Plastin-2	LCPI	Q02880-2	DNA topoisomerase 2-beta
Q9Y4G6	Talin-2	TLN2	Q15293	Reticulocalbin-1	RCN1	P08621-2	U1 small nuclear ribonucleoprotein 70 kDa
Q9LH03-2	Neuronal-specific septin-3	SEPT3	O75131	Copine-3	CPNE3	P12956	X-ray repair cross-complementing protein
Q8N4C8-3	Misshappen-like kinase 1	MINK1	P27824	Calnexin	CANX	Q5TIV8	Torsin-1A-interacting protein 1
Q13814-2	Spectrin alpha chain, non-erythrocytic 1	SPTAN1	Q07954	Protein disulfide-isomerase	PDIA6	Q15459	Splicing factor 3A subunit 1
O14184	Ankyrin-2	ANK2	P06756-3	Integrin alpha-V	ITGAV	P51114-2	Fragile X mental retardation syndrome-rela
O01082	Spectrin beta chain, non-erythrocytic 1	SPTBN1	P08758	Annexin A5	ANXA5	P49756	RNA-binding protein 25
O00429	Dynamitin-like protein	DNM1L	P54819-2	Adenylate kinase 2, mitochondrial	AK2	P26583	High mobility group protein B2
O94819	Kelch repeat and BTB domain-containing	prKBTBD11	P41567	Eukaryotic translation initiation factor 1	EIF1	P16401	Histone H1.5
Q92598-2	Heat shock protein 105 kDa	HSPH1	P09110	3-ketoacyl-CoA thiolase, peroxisomal	ACAA1	O75367-2	Core histone macro-H2A.1
Q6VY07	Phosphofurin acidic cluster sorting protein	IPACS1	P22307-7	Non-specific lipid-transfer protein	SCP2	P12270	Nucleoprotein TPR
P55809	Succinyl-CoA:3-ketoacid coenzyme A trans	OXCT1	Q9Y3A5	Ribosome maturation protein SBDS	SBDS	P09661	U2 small nuclear ribonucleoprotein A
P24752	Acetyl-CoA acetyltransferase, mitochondria	ACAT1	Q96DB5-2	Regulator of microtubule dynamics protein	RMDN1	Q15393	Splicing factor 3B subunit 3
P23634-7	Plasma membrane calcium-transporting AT	ATP2B4	P53004	Biliverdin reductase A	BLVRA	P62191	26S protease regulatory subunit 4
P11177-2	Pyruvate dehydrogenase E1 component	subPDHB	Q7LQ6-2	Basic leucine zipper and W2 domain-contai	BZWI	Q14019	Coactosin-like protein
O00139-4	Kinesin-like protein KIF2A	KIF2A	Q9Y2Q3	Glutathione S-transferase kappa 1	GSTK1	P62081	40S ribosomal protein S7
O43857	Isocitrate dehydrogenase [NAD] subunit	beta IDH3B	Q86VP6	Cullin-associated NEDD8-dissociated protei	CAND1	P08708	40S ribosomal protein S17
Q9UGV2-2	Protein NDRG3	NDRG3	Q6DD88	Atlastin-3	ATL3	P32969	60S ribosomal protein L9
Q9P2R7-2	Succinyl-CoA ligase [ADP-forming] subunit	beta UCL1A2	P51659	Peroxisomal multifunctional enzyme type 2	HSD17B4		
Q00059	Transcription factor A, mitochondrial	TFAM	P49748-2	Very long-chain specific acyl-CoA dehydrog	ACADVL		
P48047	ATP synthase subunit O, mitochondrial	ATP5O	P28288	ATP-binding cassette sub-family D member	ABCD3		
P36542	ATP synthase subunit gamma, mitochondria	ATP5C1	P13797	Plastin-3	PLS3		
P25705	ATP synthase subunit alpha, mitochondria	ATP5A1	O95747	Serine/threonine-protein kinase OSR1	OSR1		
P06576	ATP synthase subunit beta, mitochondria	ATP5B	O43776	Asparagine-tRNA ligase, cytoplasmic	NARS		
O60313-2	Dynamitin-like 120 kDa protein, mitochondri	OPA1	O00410	Importin-5	IPO5		
O75390	Citrate synthase, mitochondrial	CS	O15498	Synaptobrevin homolog YKT6	YKT6		
O95757	Heat shock 70 kDa protein 4L	HSPA4L	Q9BY32	Inosine triphosphate pyrophosphatase	ITPA		
Q9H4G0	Band 4.1-like protein 1	EPB41L1	Q96G03	Phosphoglucomutase-2	PGM2		
O99447-4	Ethanolamine-phosphate cytidyltransfera	PCYT2	Q8NBJ7	Sulfatase-modifying factor 2	SUMF2		
Q14738-3	Serine/threonine-protein phosphatase 2A	PPP2R5D	P52788	Spermine synthase	SMS		
P35611	Alpha-adducin	ADD1	P28838-2	Cytosol aminopeptidase	LAP5		
P62879	Guanine nucleotide-binding protein G(I)/G	GNB2	P07741	Adenine phosphoribosyltransferase	APRT		
Q9Y2Z0	Suppressor of G2 allele of SKP1 homolog	SUGT1	O75874	Isocitrate dehydrogenase [NADP] cytoplasm	IDH1		
Q9NZH0	G-protein coupled receptor family C group	5 GPRCSB	O95394	Phosphoacetylglucosaminase mutase	PGM3		
Q7Z4S6-2	Kinesin-like protein KIF21A	KIF21A	Q9BVG4	Protein PBDC1	PBDC1		
O14639-2	Actin-binding LIM protein 1	ABLIM1	Q9BRA2	Thioredoxin domain-containing protein 17	TXNDC17		
Q15599	Na(+)/H(+) exchange regulatory cofactor	NHSLC9A3R2	Q16881	Thioredoxin reductase 1, cytoplasmic	TXNRD1		
Q9UN36-3	Protein NDRG2	NDRG2	O95336	6-phosphogluconolactonase	PGLS		
P50993	Sodium/potassium-transporting ATPase sub	ATP1A2	P52209-2	6-phosphogluconate dehydrogenase, decar	PGD		
P14415	Sodium/potassium-transporting ATPase sub	ATP1B2	Q9BS26	Endoplasmic reticulum resident protein 44	ERP44		
P12277	Creatine kinase B-type	CKB	Q15067-2	Peroxisomal acyl-coenzyme A oxidase 1	ACOX1		

O94760	N(G),N(G)-dimethylarginine dimethylamino DDAH1		Q14697-2	Neutral alpha-glucosidase AB	GANAB
P09972	Fructose-bisphosphate aldolase C	ALDOC	Q01105	Protein SET	SET
P00367	Glutamate dehydrogenase 1, mitochondria	GLUD1	P41091	Eukaryotic translation initiation factor 2 subEIF2S3	
P51649	Succinate-semialdehyde dehydrogenase, m	ALDH5A1	P28066	Proteasome subunit alpha type-5	PSMA5
Q9UEY8-2	Gamma-adducin	ADD3	P25787	Proteasome subunit alpha type-2	PSMA2
Q9H479	Fructosamine-3-kinase	FNSK	P05198	Eukaryotic translation initiation factor 2 subEIF2S1	
Q99798	Aconitate hydratase, mitochondrial	ACO2	P00491	Purine nucleoside phosphorylase	PNP
Q16555	Dihydropyrimidinase-related protein 2	DPYSL2	O14744-2	Protein arginine N-methyltransferase 5	PRMT5
Q15555	Microtubule-associated protein RP/EB fam1	MAPRE2	O14818	Proteasome subunit alpha type-7	PSMA7
P30086	Phosphatidylethanolamine-binding protein	PEBP1	Q8N8S7-3	Protein enabled homolog	ENAH
P14621	Acylphosphatase-2	ACYP2	Q15046	Lysine-tRNA ligase	KARS
P09417	Dihydropteridine reductase	QDPR	Q13155	Aminoacyl tRNA synthase complex-interact AIMP2	
P05091	Aldehyde dehydrogenase, mitochondrial	ALDH2	P39687	Acidic leucine-rich nuclear phosphoprotein ANP32A	
P00568	Adenylate kinase isoenzyme 1	AK1	P29692-2	Elongation factor 1-delta	EEF1D
O14745	Na(+)/H(+) exchange regulatory cofactor	NHSLC9A3R1	P26368-2	Splicing factor U2AF 65 kDa subunit	U2AF2
O43813	LanC-like protein 1	LANCL1	P23284	Peptidyl-prolyl cis-trans isomerase B	PPIB
Q15435	Protein phosphatase 1 regulatory subunit 7	PPP1R7	E9PAV3-2	Nascent polypeptide-associated complex suNACA	
P43034	Platelet-activating factor acetylhydrolase IB	PAFAH1B1	O15173	Membrane-associated progesterone receptor	PMR1
P11216	Glycogen phosphorylase, brain form	PYGB			
P36405	ADP-ribosylation factor-like protein 3	ARL3			
Q16836	Hydroxyacyl-coenzyme A dehydrogenase, m	HADH			
Q02252	Methylmalonate-semialdehyde dehydrogenase	ALDH6A1			
P27144	Adenylate kinase 4, mitochondrial	AK4			
P12694	2-oxoisovalerate dehydrogenase subunit alp	BCKDHA			
P06744	Glucose-6-phosphate isomerase	GPI			
P08237-2	ATP-dependent 6-phosphofructokinase, mu	PFKM			
P62258	14-3-3 protein epsilon	YWHAE			
P17858	ATP-dependent 6-phosphofructokinase, live	PFKL			
P31946-2	14-3-3 protein beta/alpha	YWHAAB			
Q6YN16	Hydroxysteroid dehydrogenase-like protein	HSDL2			
Q13057	Bifunctional coenzyme A synthase	COASY			
Q06124-2	Tyrosine-protein phosphatase non-recepto	PTPN11			
P35222	Catenin beta-1	CTNNB1			
Q03001	Dystonin	DST			
Q16658	Fascin	FSCN1			
P61088	Ubiquitin-conjugating enzyme E2 N	UBE2N			
P55786	Puromycin-sensitive aminopeptidase	NPEPPS			
P27348	14-3-3 protein theta	YWHAQ			
P50395	Rab GDP dissociation inhibitor beta	GDI2			
Q9UBQ0	Vacuolar protein sorting-associated protein	VPS29			
Q9NTK5	Obg-like ATPase 1	OLA1			
Q96KP4	Cytosolic non-specific dipeptidase	CNDP2			
Q8NCW5-2	NAD(P)H-hydrate epimerase	APOA1BP			
Q12765	Secernin-1	SCRN1			
Q00169	Phosphatidylinositol transfer protein alpha	PITPNAA			
P49591	Serine-tRNA ligase, cytoplasmic	SARS			
P48735	Isocitrate dehydrogenase [NADP], mitochon	IDH2			
P34932	Heat shock 70 kDa protein 4	HSPA4			
P31939	Bifunctional purine biosynthesis protein	PUAT1C			
P22314-2	Ubiquitin-like modifier-activating enzyme	1UBA1			
P28161	Glutathione S-transferase Mu 2	GSTM2			
Q9Y2A7	Nck-associated protein 1	NCKAP1			
P31150	Rab GDP dissociation inhibitor alpha	GDI1			
Q15121	Astrocytic phosphoprotein PEA-15	PEA15			
P61018	Ras-related protein Rab-4B	RAB4B			
Q9GZT8-2	NIF3-like protein 1	NIF3L1			
P14618-2	Pyruvate kinase PKM	PKM			
P52565	Rho GDP-dissociation inhibitor 1	ARHGDI1A			
P49789	Bis(5'-adenosyl)-triphosphatase	FHIT			
Q8N7J2	APC membrane recruitment protein 2	AMER2			
Q9LUP00-10	LIM and calponin homology domains-conta	LIMCH1			
Q9Y2J8	Protein-arginine deiminase type-2	PADI2			
Q8TAM6	Ermin	ERMN			
Q94832	Unconventional myosin-1d	MYO1D			
P60201	Myelin proteolipid protein	PLP1			
Q16352	Alpha-internexin	INA			
P12036-2	Neurofilament heavy polypeptide	NEFH			
P07197	Neurofilament medium polypeptide	NEFM			
P02686-5	Myelin basic protein	MBP			
P07196	Neurofilament light polypeptide	NEFL			
Q94811	Tubulin polymerization-promoting protein	TPPP			
Q96PY5-3	Formin-like protein 2	FMNL2			
Q8IXJ6-2	NAD-dependent protein deacetylase sirtuin	SIRT2			
P54652	Heat shock-related 70 kDa protein 2	HSPA2			
P50148	Guanine nucleotide-binding protein G(i) su	GNAQ			
P48426	Phosphatidylinositol 5-phosphate 4-kinase	PIP4K2A			
P09543	2,3-cyclic-nucleotide 3-phosphodiesterase	CNP			
A0MZ66	Shootin-1	KIAA1598			
O43236	Septin-4	SEPT4			
Q9NUQ9	Protein FAM49B	FAM49B			
Q9NP72	Ras-related protein Rab-18	RAB18			
Q9H492	Microtubule-associated proteins	1A/1B high			
	MAP1LC3A				
Q92614-3	Unconventional myosin-XVIIIa	MYO18A			
P78357	Contactin-associated protein 1	CNTNAP1			
P63151	Serine/threonine-protein phosphatase 2A	5PPP2R2A			
P61163	Alpha-centractin	ACTR1A			
P52306	Rap1 GTPase-GDP dissociation stimulator	1RAP1GDS1			
P05023-4	Sodium/potassium-transporting ATPase sub	ATP1A1			
P04350	Tubulin beta-4A chain	TUBB4A			
O43301	Heat shock 70 kDa protein 12A	HSPA12A			
O00499-3	Myc box-dependent-interacting protein 1	BIN1			
O15400-2	Syntaxin-7	STX7			
Q969P0	Immunoglobulin superfamily member 8	IGSF8			
Q8IVF2-3	Protein AHNAK2	AHNAK2			
P63096	Guanine nucleotide-binding protein G(i) su	GNAI1			
Q5TH69	Brefeldin A-inhibited guanine nucleotide-ex	ARFGEF3			
P45381	Aspartoacylase	ASPA			
Q9Y281	Cofilin-2	CFL2			
P20916-3	Myelin-associated glycoprotein	MAG			

CHAPITRE 1 : ARTICLE SOUMIS

New glioma classification for precise therapeutic decision from spatially-resolved proteomics guided by MALDI-MSI and clinical data integration

Lauranne Drelich^{1 §}, Marie Duhamel^{1 §}, Maxence Wisztorski^{1§}, Soulimane Aboulouard¹, Jean-Pascal Gimeno¹, Pierre -Damien Caux¹, Nina Ogrinc¹, Patrick Devos², Tristan Cardon¹, Michael Weller⁶, Fabienne Escande³, Fahed Zairi⁴, Claude-Alain Maurage⁴, Isabelle Fournier^{1,7*}, Emilie Le Rhun^{1,5,6*} and Michel Salzet^{1,7*}

¹ Univ.Lille, Inserm, CHU Lille, U1192, Laboratoire Protéomique, Réponse Inflammatoire et Spectrométrie de Masse (PRISM), F-59000 Lille, France

² Univ. Lille, CHU Lille, ULR 2694 - METRICS : Évaluation des technologies de santé et des pratiques médicales, F-59000 Lille, France

³ CHU Lille, Service de biochimie et biologie moléculaire, CHU Lille, F-59000 Lille, France

⁴ Univ.Lille, Inserm, CHU Lille, U1172 - LiNCog - Lille Neuroscience & Cognition, F-59000 Lille, France

⁵ CHU Lille, Service de neurochirurgie, F-59000 Lille, France

⁶ Department of Neurology & Clinical Neuroscience Center, University Hospital and University of Zurich, Zurich, Switzerland

⁷Institut Universitaire de France, 75000 Paris, France

[§]These authors share first authorship

Corresponding authors:

* To whom correspondence should be addressed. Michel Salzet, Isabelle Fournier, Emilie Le Rhun, Phone: +33 320 43 41 94; Fax: +33 320 43 40 54; Email: michel.salzet@univ-lille.fr , isabelle.fournier@univ-lille.fr, emilie.lerhun@usz.ch

Lead Contact:

Michel Salzet Phone: +33 320 43 41 94; Fax: +33 320 43 40 54; Email: michel.salzet@univ-lille.fr

SUMMARY

Molecular heterogeneities are a key feature of glioblastoma pathology leading to high discrepancies between patients mean survivals. Based on multidimensional analysis on a cohort of 50 glioblastoma tumors, three proteomic subtypes have been uncovered. Mass spectrometry imaging resulted in the definition of these three groups. More than 95% of correct classification rates for these 3 groups was obtained using SpiderMass technology. In order to correlate this classification with functional biology, spatially resolved proteomics was performed and systemic biology established that these groups are related to neurogenesis, immune infiltration and tumorigenesis, respectively. Association with overall survival showed a signature of 6 proteins associated with prognosis. We validate these new prognostic markers of glioblastomas which allows to stratify patients and help the therapeutic decision. We also demonstrate a high disruption between proteomic and genomic classification and gave now real tools for pathologist to guide the oncologist decision.

Key Words: classification, glioma, imaging, prognosis, proteins, mass spectrometry, SpiderMass, microproteomics

Significance:

Glioblastoma are very heterogeneous tumors with median survivals usually inferior to 20 months. We conducted a comprehensive proteomics analysis of glioblastoma in order to identify proteomic subgroups of glioblastoma. We stratified the 50 glioblastoma cases into three groups with specific proteomic profiles. We find that the first group has a neurogenesis signature and is associated with a better prognosis while the second group of patients has an immune profile with a bad prognosis. The third group is more associated to tumorigenesis. We correlated these results with the TCGA data. Finally, we have identified new prognostic markers of glioblastomas which may help to stratify the patients and help the therapeutic decision.

Highlights

- A novel stratification of glioblastoma based on mass spectrometry was established.
- Three groups with different molecular features and survival were identified.
- This new classification could improve prognostication and may help therapeutic options.

Introduction

Glioblastoma represent the main malignant primary brain tumor with an incidence of 3.21 per 100,000 population (Ostrom et al., 2018a). The prognosis is poor with a median survival estimated at 16 months in clinical studies (Chinot et al., 2014; Gilbert et al., 2014; Stupp et al., 2009; Stupp et al., 2017; Weller et al., 2017a; Weller et al., 2017b) and around 12 months in contemporary population-based studies (Gramatzki et al., 2016). Approximately 5% of patients survive more than 5 years (Ostrom et al., 2018b). Favorable therapy-independent prognostic factors include lower age and higher neurological performance status at diagnosis. Furthermore, low postoperative residual tumor volume has been associated with improved outcome. In a cohort of 232 patients with centrally confirmed glioblastoma who survived at least 5 years, the median age at diagnosis was 52 years (range 21-77 years) and most patients had a gross total resection initially (Weller et al., 2019).

Morphological criteria for the diagnosis of glioblastoma according to the World Health Organization (WHO) central nervous system tumor classification of 2016 (Louis et al., 2016) include mitotic activity, anaplastic nuclear features, microvascular proliferation and necrosis. This definition comprised three groups of tumors, (i) isocitrate dehydrogenase (IDH) wild-type glioblastoma which represents at least 90% of glioblastomas and includes as morphological variants giant cell glioblastoma, gliosarcoma and epithelioid glioblastoma; (ii) IDH-mutant glioblastoma which represents less than 10% of glioblastomas and (iii) glioblastoma *not otherwise specified* (NOS) when the IDH status is unknown. These diagnostic concepts are very likely to be revised in the upcoming revision of the WHO classification based on cIMPACT-NOW (update 3 (Brat et al., 2018a) and 5 (Brat et al., 2020a)) recommendations to allow a molecular definition of glioblastoma based on the detection of EGFR amplification, a +7/-10 genotype and TERT promoter mutation in a IDH wildtype astrocytic tumor (Brat et al. 2018) and by renaming IDH-mutant glioblastoma as IDH-mutant astrocytoma, WHO grade 4 (Brat et al., 2020b; Brat et al., 2018b; Hu et al., 2020). Standard treatment of glioblastoma includes maximum safe resection followed by radiotherapy with concomitant and maintenance temozolomide (Weller et al., 2017c).

Efforts to further subclassify glioblastoma have focused on the genome, methylome and transcriptome level. In 2008, the Cancer Genome Atlas (TCGA) delineated three main signaling pathways affected by genetic alterations in glioblastoma, RTK/RAS/PI-3K - altered in 88% of tumors, p53 - altered in 87% of tumors, and RB - altered in 78% of tumors, based on DNA copy number, gene expression and DNA methylation aberrations (Brennan et al., 2013; Cancer Genome Atlas Research Network 2018). Genome methylation profiling in adult patients with IDH wildtype glioblastoma allowed the definition of three epigenetic subtypes, (i)

receptor tyrosine kinase (RTK) I often with PDGFR amplification, (ii) RTK II or classic often with EGFR amplification, CDKN2A/B deletion, and PTEN mutation, and (iii) mesenchymal (Sturm et al., 2012). Any clinical relevance of the methylation classes remains controversial. The DNA methylation-based classification of CNS tumors has meanwhile evolved to a comprehensive machine-learning approach that will shape the next revision of the WHO classification (Capper et al., 2018), resulting also in the delineation of further rare methylation classes of glioblastoma. Prior to the introduction of methylation profiling, a classification of glioblastoma based on transcriptional profiling had revealed four subtypes of glioblastoma: proneural, neural, classic and mesenchymal (Verhaak et al., 2010). The neural subtype is no longer maintained, and it appears now that transcriptomic profiles are less homogeneous and stable than genome or methylome classifiers. Proteomic approaches have been less frequently explored, although they can identify and quantify the final product of altered genomics and transcriptomics and may better characterize the activation of specific pathways (Deighton et al., 2010; Dilillo et al., 2017; Goplen et al., 2010; Kalinina et al., 2011). The determination of specific proteomic signatures could help to improve the distinction between the different glioblastoma subtypes and to guide the therapeutic management.

Here we explored 50 glioblastoma samples using MALDI MS imaging (MALDI MSI) and spatially-resolved microproteomics to determine a proteomic classification of glioblastoma. In this work, a proteomic strategy was used to classify glioblastoma with preservation of tumor tissue architecture which is of particular importance for these highly heterogeneous tumors.

Materials & Methods

Patient samples and consent

Patients with newly diagnosed glioblastoma were prospectively enrolled between September 2014 and November 2018 at Lille University Hospital, France. Patients were adult, had no medical history of other cancers or previous cancer treatment, no known genetic disease potentially leading to cancer and no neurodegenerative disease. Tumors samples were processed within 2 hours after sample extraction in the surgery room to limit the risk of degradation of proteins.

Immunohistochemistry analyses, deoxyribonucleic acid (DNA) extraction and quantification of tumor samples are reported in Note S1.

Approval was obtained from the research ethics committee (ID-RCB 2014-A00185-42) before initiation of the study. The study adhered to the principles of the Declaration of Helsinki and

the Guidelines for Good Clinical Practice and is registered at NCT02473484. Informed consent was obtained from patients.

MALDI mass spectrometry imaging (MALDI MSI)

A Leica CM1510S cryostat (Leica Microsystems, Nanterre, France) was used to cut twelve micrometer sections in order to perform the MALDI MSI analysis (Fournier et al., 2003; Lemaire et al., 2007; Lemaire et al., 2006a; Lemaire et al., 2006b; Wisztorski et al., 2007). These tissue sections were deposited on ITO-coated glass slides (LaserBio Labs, Valbonne, France) and vacuum-dried during 15 min. Tissue sections were then soaked in different solutions to remove abundant lipids: (1) 1 min in 70% ethanol, (2) 1 min in 100% ethanol, (3) 1 min in acetone and (4) 30 s in chloroform with concomitant drying between washings. An electrospray nebulizer connected to a syringe pump (flow rate 180 nL/min) was used to uniformly spray a trypsin solution (60 µg/mL in NH₄HCO₃ 50 mM) on the tissue surface for 15 min. ImagePrep (Bruker Daltonics, Bremen, Germany) was used as an incubation chamber by microspraying water heated to 37 °C for 2 h (60 cycles with 2 s spraying, 180 s incubation and 60 s drying using the nitrogen flow). For optimal digestion, a constant humidity atmosphere was maintained inside the spray chamber by filling a small container with 95°C water. After digestion, HCCA/ANI (Lemaire et al., 2006) a solid ionic matrix was deposited using ImagePrep. Briefly, 36 µL of aniline were added to 5 mL of a solution of 10 mg/mL HCCA dissolved in ACN/0.1% TFA aqueous (7:3, v/v). A real-time control of the deposition is performed by monitoring scattered light to obtain a uniform layer of matrix. MALDI MSI experiments were done on an Ultraflex II MALDI-TOF/TOF instrument (Bruker) with a smartbeam II solid state laser. Mass spectra were acquired in positive reflector mode between 800–4000 m/z range. Recorded spectra were averaged from 400 laser shots per pixel acquired at 200Hz laser repetition rate and with a 70 µm spatial resolution raster.

MALDI MSI data processing and analysis

The MALDI-MSI data were analyzed using SCiLS Lab software (SCiLS Lab 2019, SCiLS GmbH). Common processing methods for MALDI MSI were applied with a baseline removal using a convolution method and data were normalized using Total Ion Count (TIC) method (Klein et al., 2014; Trede et al., 2012). Then, the resulting pre-processing data were clustered to obtain a spatial segmentation using the bisecting k means algorithm (Alexandrov et al., 2010). Different spatial segmentations were performed. First, an individual segmentation was applied to each tissue separately. Then, the data from all tissues were clustered together to obtain a global segmentation. Briefly, the spatial segmentation consists of grouping all spectra

according to their similarity using a clustering algorithm and all pixels of a same cluster are colour coded. To limit the pixel-to-pixel variability, edge-preserving image denoising was applied. Note that a color is arbitrary assigned to a cluster and that several disconnected regions can have the same color, i.e. the same molecular content. The results of segmentation are represented on a dendrogram resulting from a hierarchical clustering. The branches of the dendrogram were defined based on a distance calculation between each cluster. The selection of different branches of the dendrogram will give a segmentation map where regions of distinct molecular composition were differentially color-coded. The individual segmentation provides information concerning the heterogeneity of the tissue section and the global segmentation is used to group patients with a similar molecular signature. For comparison, global segmentation was also performed using the Ward clustering method with IMAGEREVEAL MS Ver.1.1 (Shimadzu). The global spatial segmentation allowed to determine regions of interest (ROIs) which were then be subjected to on-tissue microdigestion followed by microextraction for protein identifications.

SpiderMass analyses

The global design of the instrument setup has been described (Saudemont et al., 2018). Briefly, the system is composed of three parts including a laser system for micro-sampling of tissues set remotely, a transfer line allowing for transfer of the micro-sampled material to the third part, which is the mass spectrometer itself (Fatou et al., 2016). The first part is composed of a tunable wavelength OPO which is tunable from 2.8 μm to 3.1 μm (Radiant version 1.0.1, OPOTEK Inc., Carlsbad, CA, USA) pumped by a pulsed Nd:YAG laser (pulse duration: 5 ns, $\lambda=1064$ nm, Quantel, Les Ulis, France). A biocompatible laser fiber (450 μm inner diameter; length of 1 m; Infrared Fiber Systems, Silver Spring, CO, USA) is connected to the laser system output and a handpiece including a 4 cm focusing lens is attached to the end of the laser fiber. The handpiece with a 4cm focusing lens allows the user to hold the system and screen the surface of raw tissues at a resolution of 400 μm . In these experiments the irradiation time was fixed to 10 sec at 4 mJ/pulse laser energy corresponding to a laser fluence of ~ 3 J/cm². The laser energy was measured at the focal point of the focusing lens using a power meter (ThorLabs, Maisons-Laffitte, France). The second part of the system corresponds to a 3 meter length transfer line made from a Tygon ND 100-65 tubing (2.4 mm inner diameter, 4 mm outer diameter, Akron, USA). The transfer line is attached on one side onto the laser hand piece at the end of the laser fiber and on its other side directly connected to the mass spectrometer (Xevo, Waters, Manchester, United Kingdom) from which the conventional electrospray source was removed and replaced by an atmospheric pressure interface (Fatou et al., 2016). Each

acquisition was accompanied by a 150 $\mu\text{L}/\text{min}$ isopropanol infusion. Spectral acquisition was performed both in positive and negative ion resolution mode with a scan time of 1 sec. Prior to SpiderMass analysis, the samples were taken out of the -20°C freezer and thawed to RT for 30 s. The spectral acquisition sequence was composed of 2 or 3 acquisitions using 1-sec irradiation periods. The ROI were selected using the morphological controls and acquired peptide MALDI-MSI data prior to each SpiderMass acquisition in order to ensure that each acquisition was performed on the same histological area (Ogrinc et al., 2019).

Classification model construction

For data analysis, all raw data files produced with the SpiderMass instrument were imported into the Abstract Model Builder (AMX v.0.9.2092.0) software. After importation, spectra were first pre-processed. The pre-processing steps include background subtraction, total ion count normalization, lockmass correction and re-binning to a 0.1 or 0.2 Da window. All processed MS spectra obtained from the 30 histologically validated samples were then used to build a principal component analysis and linear discriminant analysis (PCA-LDA) classification model (Ogrinc et al., 2019). The first step consisted of PCA to reduce data multidimensionality by generating features that explain most of the variance observed. These features were then subjected to supervised analysis using LDA by setting the classes that the model will be based upon. LDA attempts to classify the sample spectra and assess the model by cross validation. Cross-validation was carried out by either using the “20% out” or the “leave one patient out” methods. For the first method, 20% of MS spectra are randomly taken from the total spectra and the model is reconstructed from the remaining 80%. The remaining 20% of spectra are used to interrogate the reconstructed model. The permutation is automatically reiterated for 5 cycles before reporting the cross-validation results. For the second method, the spectra are grouped by patient and left out one by one; at each step the model without the patient is interrogated against this model.

Spatially-resolved proteomics

On-tissue digestion

A total of 147 ROIs were selected from MALDI-MSI. Spatially resolved microproteomics was performed on the predefined ROIs according to the previously published protocol (Quanico et al., 2013). Briefly, tissue sections of $20\mu\text{m}$ thick were cut and subjected to different washes to remove lipids. Then, on-tissue digestion is performed using a LysC-trypsin solution ($40\mu\text{g}/\text{mL}$ in Tris-HCl 50mM , pH 8.0). This solution was deposited using a piezoelectric microspotter (CHIP-1000, Shimadzu, CO, Kyoto, Japan) on each ROIs with a total area of 1 mm^2 (4×4 spots of

200 µm. Enzyme droplet was maintained for a total of 2 h digestion. After enzyme deposition 0.1% TFA was spotted for 25 cycles with 100 pL on each spot/cycle.

Microextraction by liquid microjunction

After tissue microdigestion, the tryptic peptides were extracted using an automated platform, the TriVersa Nanomate platform (Advion Biosciences Inc., Ithaca, NY, USA) with Liquid Extraction Surface Analysis (LESA) option (Quanico et al., 2013). Briefly, a volume of solvent was aspirated onto a tip and dispensed onto the digested region. The droplet formed was maintained between the tip and the tissue and then aspirated after 15s. the recovery solution is finally pooled in a low binding tube. Three extraction steps were performed per region using different solution: (1) 0.1% TFA, (2) ACN/0.1% TFA (8:2, v/v), and (3) MeOH: 0.1% TFA (7:3, v/v). Two extraction cycles per point were performed to increase the amount of material collected.

NanoLC-MS & MS/MS analysis

Prior to MS analysis, the reconstituted samples were desalted using C-18 Ziptip (Millipore, Saint-Quentin-en-Yvelines, France), eluted with 80% ACN and vacuum-dried. The dried samples were resuspended in 0.1% FA aqueous/ACN (98:2, v/v). Peptides separation was performed by reverse phase chromatography, using a NanoAcquity UPLC system (Waters) coupled to a Q-Exactive Orbitrap mass spectrometer (Thermo Scientific) via a nano electrospray source. A pre-concentration column (nanoAcquity Symmetry C18, 5 µm, 180 µm x 20 mm) and an analytical column (nanoAcquity BEH C18, 1.7 µm, 75 µm x 250 mm) were used. A 2 h linear gradient of acetonitrile in 0.1% formic acid (5%-35%) was applied, at the flow rate of 300 nl/min. For MS and MS/MS analysis, a data dependent mode was defined to analyze the 10 most intense ions of MS analysis (Top 10). The MS analysis was performed with an m/z mass range between 300 to 1600, a resolution of 70,000 FWHM, an AGC of 3e6 ions and a maximum injection time of 120 ms. The MS/MS analysis was performed with an m/z mass range between 200 to 2000, an AGC of 5e4 ions, a maximum injection time of 60 ms and the resolution was set at 17,500 FWHM. To avoid any batch effect during the analysis, the extractions were chosen at random to create analysis sequences.

Data analysis

All MS data were searched with MaxQuant software (Cox and Mann, 2008; Tyanova et al., 2015) (Version 1.5.3.30) using Andromeda search engine (Cox et al., 2011) against the complete proteome for *Homo sapiens* (UniProt, release July 2018, 20 412 entries). Trypsin was

selected as enzyme and two missed cleavages were allowed, with N-terminal acetylation and methionine oxidation as variable modifications. The mass accuracies were set to 6 ppm and 20 ppm respectively for MS and MS/MS spectra. False discovery rate (FDR) at the peptide spectrum matches (PSM) and protein levels was estimated using a decoy version of the previously defined databases (reverse construction) and set to 1%. A minimum of 2 peptides with at least one unique is necessary to complete the identification of a protein. The MaxLFQ algorithm (Cox et al., 2014) was used to performed label-free quantification of the proteins. The resulting file was analyzed using Perseus software (version 1.6.0.7). First, hits from the reverse database, proteins with only modified peptides and potential contaminants were removed. Statistical analyses were performed using ANOVA with a truncation value based on “Benjamini Hocheberg FDR” of 5%. Three categorical annotation groups were used for the ANOVA, i.e. (1) the color group based on the three colors from Scils global segmentation of the 50 samples (Red; Yellow and Blue), (2) the patient groups which are determined by the majority color present in each tumor sample (Groups A, B, C) and (3) the patients survival time (patients with an OS > to the third quartile, patients with an OS between the first and the third quartile and patients with an OS < to the first quartile). Proteins significantly different were selected and normalized by a Z-score with matrix access by rows. For representation, a hierarchical clustering was performed using the Euclidean parameter for the distance calculation, and the average option for linkage in the rows and columns of the trees with a maximum of 300 clusters.

System biology analyses

An annotation analysis of gene ontology terms for the identified proteins were performed using PANTHER Classification System (version 14.1, <http://www.pantherdb.org>), FunRich (Version 3.1.3) (Pathan et al., 2017) and the STRING database (version 11.0, www.string-db.org) (Szklarczyk et al., 2019). Potential interaction network was then loaded into Cytoscape 3.7.2 with relative expression data using Idmapper (Otasek et al., 2019). The Reactome FI plugging was used to select a subnetwork of gene ontology terms and NCI database-associated disease-specific proteins.

The relationships between the differentially expressed proteins among all conditions were also depicted based on the Ariadne ResNet database (Yuryev et al., 2009) using Elseviers' Pathway Studio (version 11.0, Elsevier). The subnetwork Enrichment Analysis (SNEA) algorithm was used to detect the statistically significant altered biological pathways in which the identified proteins are involved.

Human Pathology Atlas

The glioma data contained in the Human pathology atlas (Uhlen et al., 2017) were used. Based on TCGA transcriptomics and antibody-based protein data from 153 patients, this database identified 268 potentially prognostic genes (201 unfavorable and 67 favorable prognoses). These data were compared to the proteins identified in our study.

Alternative Proteins identification

RAW data obtained by nanoLC-MS/MS analysis were analyzed using Proteome Discoverer V2.3 (Thermo Scientific) LFQ quantification with the following parameters: trypsin as enzyme, 2 missed cleavages, methionine oxidation as variable modification and carbamidomethylation of cysteines as static modification, Precursor Mass Tolerance: 10 ppm and fragment mass tolerance: 0.6 Da. The validation was performed using Percolator with a FDR set to 0.001%. A consensus workflow was then applied for the statistical arrangement, using the high confidence protein identification. The protein database was uploaded from Openprot (<https://openprot.org/>) and included RefProt, novel isoforms and AltProts predicted from both Ensembl and RefSeq annotations (GRCh38.83, GRCh38.p7) (Brunet et al., 2018; Delcourt et al., 2018; Vanderperre et al., 2013; Vanderperre et al., 2012) for a total of 658263 entries. The identified abundance was extracted to PD2.3 and loaded in Perseus to performed statistical analysis and graphical representation.

Statistical analyses

First, descriptive analyses were performed on clinical data. Patients were divided into 3 groups according to the quartiles of the overall survival (<Q1, Q1-Q3, > Q3). The Cox model was used to determine which proteins were most associated with overall survival. Stepwise analysis and bootstrap methods (500 samples) were used to guarantee the robustness of the results. The proteins selected after this step were used to carry out a hierarchical classification (Euclidean distance and Ward's method) on the 50 patients to determine if there were any subgroups (clusters). Finally, the clinical variables were analyzed according to the different clusters in order to provide a clinical description of the clusters obtained. Statistical analyses were performed using the SAS Software, V9.4.

Confirmatory immunohistochemistry analyses

Survival group validation was performed using antibodies directed against MAOB, IGHM, HSPD1, LASP1, CFH, and ANXA6. The tissues were incubated with a primary antibody at 4°C overnight, followed by application of a secondary antibody (alexa fluor conjugated antibody, 1/1 000 dilution) for 1 hour at RT. We used the following primary antibodies: MAOB (Abxexa; 1/100 dilution), IGHM (Abcam, 1/50 dilution), HSPD1 (Abcam, 1/200 dilution), LASP1 (Santa Cruz, 1/50 dilution), CFH (Abcam, 1/200 dilution) and ANXA6 (Abcam, 1/50 dilution). All slides were imaged on the Zeiss LSM700 confocal microscope. Three to four pictures were taken for each tumor section. Processing of the images and fluorescence intensity quantification was performed using ImageJ software.

Results

Clinical characteristics

Fifty glioblastoma samples from a prospective cohort were collected (**Table 1, Supp. Figure 1A**). Thirty five patients (70%) were male, the median age at diagnosis was 59 (interquartile range (IQR), 50-65), the median Karnofsky performance score at diagnosis was 90 (80-100). Twenty-nine (58%) patients had a complete resection. Forty-six tumors were IDH wild-type (92%). A methylation of MGMT promoter was noted in 17 tumors (34%), an EGFR amplification was noted in 24 cases (48%) and a homozygous CDKN2A deletion in 29 cases (59%). A standard treatment was initiated in 45 patients (90%). At the time of the analysis, 40 patients (80%) had a progression with a median PFS of 11 months. After a median follow-up of 21 months (IQR 14-34), 41 (82%) had died. The median overall survival was 20.5 months. The pathologist (CAM) defined regions of interest for each tumor sample: tumor, necrosis, and endothelial proliferation, after hematoxylin eosin safran (HES) staining (**Supp. Figure 1B**).

MALDI MSI allows patient grouping based on molecular features

Considering the heterogeneity of glioblastoma, we conducted spatially resolved proteomic studies guided by mass spectrometry imaging (MSI) (**Figure 1A**). A comparison between the pathologist annotations and the MSI molecular images showed discrepancies for many samples (**Figure 1B, Supp. Data 1**). A global clustering was then performed by subjecting spectra from all tissue samples to spatial segmentation (**Figure 1C**). Three main regions were identified *i.e.* red, blue and yellow areas according to the segmentation map

(**Figure 1C**, **Supp. Figure 1C**). We then assigned each tumor sample to a group. If a colored area was greater than 50% of the surface area, the tumor was included into the group associated to this color (**Figure 1D**, **Supp. Figure 1C**). Therefore, 13 samples mainly represented by the red region (96,8% of the tissue surface area) were assigned to group A. Nine samples mostly represented by the yellow region (75,9% of the tissue surface area) were assigned to group B and 23 samples mostly represented by the blue region (83,9% of the tissue surface area) to group C. The 5 remaining tumors could not be classified due to the presence of many different molecular regions (**Supp. Figure 1C**). These tumors were excluded from the further analyses. Within a group, all patients shared common molecular characteristics, meaning that the spectra in each of these colored regions were similar. Some specific ions can be attributed to each region: m/z 967,621 and 1492,916 were specifically present in group A, m/z 1914,591, 2375,074 and 2376,274 were specific to group B and m/z 1473,312, 2045,815, 2046,615 and 2237,849 were specific to group C. Images of some group-specific ions are shown in **Figure S1D**. The Ward clustering method using IMAGEREVEAL MS Ver.1.1 software confirmed the segmentation of the 50 tumors into 3 groups with similar specific ions (**Supp. Figure 1E**). We also confirmed, with this method, the heterogeneity of the 5 outliers.

In order to validate the classification obtained by MALDI MSI, we analysed 30 samples by SpiderMass technology (Ogrinc et al., 2019; Saudemont et al., 2018). Following the acquisition of the MS spectra in positive ion mode, a PCA analysis of the generated spectra acquired from 30 tumor tissues was performed. The features of the PCA were subjected to a supervised analysis using linear discriminant analysis (LDA) (Balog et al., 2013; Schafer et al., 2009) which yielded 3 groups (**Figure 1Ea**). According to **Figure 1Ea**, LDA 1 discriminated group A from group C and the LDA 2 separated group B from group A and C. The LDA analysis of the SpiderMass data therefore allowed the samples to be grouped in the same way as the MALDI-MSI classification. Some examples of discriminant ions (m/z) between the three groups, corresponding to lipids, are presented as their normalized intensities in **Figure 1Eb**. The most discriminating peaks for group A in LD2+ correspond to m/z 746.75, 810.65, and 916.65; for group B in LD2- correspond to m/z 718.55, 724.65, 751.55, and 778.55; for group C in LD1- correspond to m/z 725.4, 754.6, 788.65, and 850.85. To consolidate the classification, validation was performed using either 20% randomly patients taken out or the one-patient-out method (**Inset table in Figure 1E**). Excellent cross validation results were obtained using 20% randomly patient taken out method with 99.20% and 87.94% correct classification rates with and without outliers respectively and good classification using the one-patient-out method with 93.22% and 78.01% including or not outliers respectively (**Inset table**

in Figure 1E). These results of outliers and misclassifications (mainly group B) reflect the fact that each group is not represented by only one colored region.

We then correlated our classification to the clinical data (**Table 2, Supp. Figure 2A & 2B**). Group A patients were younger, and were rarely methylated, group B patients had less often frontal tumor (**Table 2**).

Identification of specific signaling pathway signatures for each group

In order to understand the molecular differences between the three groups, spatially-resolved tissue proteomic was undertaken on the 45 tissue samples (Wisztorski et al., 2016). On each tissue, 2 to 5 specific micro extraction points were selected according to the molecular regions identified by spatial segmentation of MALDI MSI data (**Supp. Data 1**) in order to analyse the tumor heterogeneity and micro-environment presenting with specific protein signatures in each group. This resulted in a total of 147 micro-extraction points. Each extraction point was associated with one of the three regions identified by MALDI-MSI (red, yellow and blue). In all tumor samples, 32 extractions were performed in the red region, 30 in the yellow region and 85 in the blue region. From these shotgun proteomic experiments, a total of 4936 proteins were identified (**Supp. Data 2**). Some proteins were unique to a patient, *i.e.* 382 unique proteins distributed among the patients (**Supp. Data 3**). From these proteins, several transcription factors (STP2, CD2 antigen cytoplasmic tail-binding protein 2), enzymes (MMP12, AKT1) but also some proteins related to the immune response (HLA-A3, HLA-B07), extracellular vesicles initiating factors (RAL-A1), tumor suppressors or specific tumor antigens like Melanoma antigen were identified (**Supp. Table 1**). Besides proteins related to HLA-A and the ones related to exosomes, none were found in common in several tumors.

To better understand the differences between each identified group, ANOVA tests with a Benjamini Hochberg FDR of 0.05 was performed. A total of 1231 proteins showed a significant difference in expression between the three groups (**Figure 2A, Supp. Data 4**). Two main branches were identified in the heatmap. The first branch was composed of 100% of samples extracted from the yellow region and correspond to the group B. The second branch separates group A from group C. This branch is then separated into two sub-branches corresponding to group A and regrouping 77% of the samples extracted from the red region and 23% of the samples extracted from the blue region. The second sub-branch corresponds to the largest cluster, group C and contains 77% of the samples extracted from the blue region, 13% of the samples extracted from the red region and 10% of the samples extracted from the

yellow region. These 3 regions are consistent with the regions identified by MALDI MSI spatial segmentation. We confirmed that each sample from the same colored region has the same proteomic profile (**Supp. Data 4**). Three specific clusters of over-expressed proteins for each group were identified using the heatmap (**Figure 2A**) *i.e.* cluster 1 corresponds to proteins overexpressed in group B; cluster 2, to proteins overexpressed in group A and cluster 3, to proteins overexpressed in group C. The lists of overexpressed proteins per group are presented in **Supplementary Data 4**.

In group A (mainly represented in cluster 2), the proteins are associated with neuro-developmental genes, that are characteristic of neuronal/glial lineages or progenitor cells. Most proteins were related to neurogenesis and axon guidance (Dihydropyrimidinase-related protein 1 (CRMP1), Misshapen-like kinase 1 (MINK1), Neuromodulin (GAP43), Dihydropyrimidinase-related protein 5 (DPYSL5), Dihydropyrimidinase-related protein 4 (DPYSL4), Microtubule-associated protein tau (MAPT), Kinesin-like protein KIF2A (KIF2A), Neurofilament heavy polypeptide (NEFH), Unconventional myosin-XVIIIa (MYO18A), MAGUK p55 subfamily member 2 (MPP2), Alpha-internexin (INA), CLIP-associating protein 2 (CLASP2) (**Supp. Data 4**). Using the functional enrichments analysis tool of String database, the most representative Reactome pathway was devoted to axon guidance. Nine of the 16 proteins identified in this pathway are involved in neuron development projection, morphogenesis, and guidance (**Supp. Figure 3Aa**). System biology analyses using SNEA and Cytoscape confirmed that the proteins in group A (Cluster 2) are involved in neurite outgrowth, synaptogenesis, synaptic plasticity (**Figures 2Ba & 2Bb, Supp. Data 5**). Interestingly, among the identified proteins some are known to be involved in tumorigenesis like Mitogen-activated protein kinase 3 (MAPK3), Protein kinase C alpha type (PRKCA) and some were already identified in glioblastoma *e.g.* CRMP1, DPYSL2 (*i.e.* CRMP2), (Jovčevska et al., 2017), DPYSL5 (*i.e.* CRMP5) (Moutal et al., 2015), GAP43 (Gutmann et al., 2002; Huang et al., 2003; Voigt et al., 2017), as well as Tau protein encoded by MAPT in low-grade glioma (Zaman et al., 2019).

Proteins overexpressed in group B (mainly represented in cluster 1) were linked to microglia activation and more generally immune system activation. Indeed, among the proteins identified, 10 proteins are linked to the immune response such as complement C1q subcomponent subunit C and B (C1QC and C1QB), complement factor H (CFH), haptoglobin (HP), kininogen 1 (KNG1), histidine-rich glycoprotein (HRG), transthyretin (TTR), grancalcin (GCA), proteins S100-A9 (S100A9) & S100-A12 (S100A12), erythrocyte band 7 integral membrane protein (STOM) and galectin-3-binding protein (LGALS3). Immunoglobulin heavy and light chains (IGHG2; IGKC; IGHG1; IGLC6; IGHM and IGHA1) and macrophage markers, Macrophage-capping protein (CAPG) were also detected (**Supp. Data 4**). Moreover, some

proteins are related to iron transporters like ceruloplasmin (CP), serotransferrin (TF), hemopexin (HPX) and haptoglobin, and other proteins are associated to coagulation e.g. Transthyretin, Kininogen-1 (KNG1), Plasminogen (PLG). Most of these proteins are known to be present in human plasma (Uhlen et al., 2017). These results are in accordance with histological annotations reflecting that most of the tissues belonging to group B present intense proliferation of capillary endothelial cells with inflammation and hemorrhage (**Supp. Figure 1B**). The cytoscape and SNEA analysis, (**Figure 2Ba and b, Supp. Data 5**), clearly confirmed that most of the proteins are involved in complements, coagulation cascade, inflammation, ischemia, vascularization, wound healing, and cancer. Moreover, 11% of patients of group B present more thromboembolic event (**Table 2**). The same pathways were found in Reactome (**Supp. Figure 3Ab**). Some of these proteins have already been detected in TCGA glioma database (see below) and are mostly unfavorable prognosis indicators for patient survival e.g. Grancalcin and CAPG (**Supp. Figure 3B**).

The overexpressed proteins in the group C (mainly represented in cluster 3) are mainly involved in tumor growth (Hepatoma-derived growth factor (HDGF), Developmentally-regulated GTP-binding protein 2 (DRG2)), but also in virus infection (Eukaryotic translation initiation factor 3 subunit L (EIF3L), Double-stranded RNA-binding protein Staufen homolog 1 (STAU1) and Interferon-induced double-stranded RNA-activated protein kinase (EIF2AK2)) (**Supp. Data 4**). KEGGS analyses confirmed a network of proteins involved in Epstein barr infection (**Supp. Figure 3Ac**). Cytoscape pathway analyses established that this group is linked to viral infection and antiviral immune response (**Figure 2Da**). System biology analyses confirmed the involvement of proteins in virus infection (transfection, reproduction) and transcriptomic modification at the RNA level (RNA splicing, metabolism, replication) (**Figure 2Db, Supp. Data 5**). Some other markers of the group C are known to be bad prognosis indicators such as EIF2AK2 and ZC3HAV1.

Identification of alternative proteins

Using OpenProt alternative proteins database (Brunet et al., 2018), 257 AltProts were identified and 170 were quantified. After ANOVA tests with a FDR of 0.05, 41 were differentially expressed between the groups (**Figure 2E**). In group A, two AltProts are over-expressed, IP_2390879 issued from the non-coding RNA (ncRNA) LOC107985743 and IP_244732 from the 5'UTR of the mRNA coding for KIFC3, involved in cell adhesion (**Supp. Table 2**). In group B, we found a cluster of nine over-expressed AltProts. Four are transcribed from ncRNA, two are located at the 5'UTR of mRNA, one at the 3'UTR and two result from the frame shift in the CDS (**Supp. Table 2**). In Group C, 30 AltProts are over-expressed: 16 from ncRNA, two from

the 5'UTR, eight from the 3'UTR and four result from the frame shift in the CDS (**Supp. Table 2**). Taken together, we identified several AltProts issued from ncRNA (51%) which is in line with our previous work on glioma cell line (NCH82) (Cardon et al., 2020b). Like our previous work, we also identified some AltProts in the present study such as IP_2323408 coming from the uncharacterized LOC105376126 gene and is currently described as a ncRNA, IP_774693, IP_572422 and IP_671464. These last three AltProts are pseudogenes: TUBAP2, TUBBP1, and TPI1P1 respectively. These pseudogenes, for which no protein has been observed yet, have an expression of their transcripts in glioma cell lines (Expression Atlas) (Petryszak et al., 2016). Interestingly, a fifth AltProt identified is IP_079312, from the mRNA encoding EDARADD. This protein was correlated with a low survival rate in ovarian cancer patients (Cardon et al., 2020b). In another study, we found interaction partners of AltProts by XL-MS in NCH82 cells in physiological condition and during a phenotypic change (Cardon et al., 2020a). Five significantly variable AltProts were found in this study, one of them is also overexpressed in group B: IP_156671 which originates from the 3'UTR of the transcript coding for SLC13A1. The four others are overexpressed in group C: IP_261897 coming from a ncRNA, IP_063564, IP_256988 both issued from the 3'UTR region of the CLDN19 and TBX21 genes respectively and IP_073718 originating from a shift in the reading frame of the CCDC181 gene.

Correlation between TCGA and proteomic data

We then compared our almost 5000 identified proteins to the TCGA database, on which 682 genes show an elevated expression in glioma; 282 of these 682 genes were found in our samples (**Supp. Table 3**). In these 682 genes, 268 genes are suggested as prognostic based on transcriptomics data from 153 patients; 201 genes associated with an unfavourable prognosis and 67 genes associated with a favourable prognosis. In our proteomics data, we found 10 proteins associated with an unfavourable prognosis: 1 protein is over-expressed in group C (POR), 2 proteins are over-expressed in group B (AEBP1 and PDIA4) and 7 proteins are over-expressed in group A (CEND1, DMTN, PAK1, MAP2K1, THY1, VSNL1 and ME1) (**Table 3**). We also found 6 proteins associated with a favourable prognosis: 4 proteins are over-expressed in group A (GLUD1, GDI2, SARS, and ARHGEF7), 1 protein is over-expressed in group B (KPNA3) and 1 protein is overexpressed in group C (PABPC1) (**Table 4**).

In all groups we found prognosis proteins. We also established that among the proteins identified in the 3 different groups, those in group A are closely correlated with the TCGA glioma data. GBM hijacks mechanisms of neural development and contains subsets of glioblastoma stem cells (GSCs) which are cells with tumor-propagating potential. This cell state is mainly represented in tumors of group A. In fact, group A presents an overexpression of

Synaptotagmin. 62% of patients of group A have no EGFR amplification. The median survival ranged between 13.5 to 49.5 months (**Table 2**). Group B has the highest level of methylated MGMT promoter and EGFR amplification (89%). The survival median ranged between 15 to 29 months (**Table 2**). For Group C, 57% of patients have an EGFR amplification. The survival median ranged between 17-33 months (**Table 2**).

Clinical data integrated to OMICs Data

After statistical analyses, no significant difference in OS were found between the 3 groups (**Supp. Figure 3C**) whereas different physio-pathological processes were identified (**Figure 2**). In the whole cohort, MGMT status (**Supp. Figure 2Ac**) and KPS (**Supp. Figure 2Ad**) were correlated with overall survival, but not the extent of resection (**Supp. Data 6**). In order to find new prognostic proteins, we performed an ANOVA test on the entire proteomic dataset (n=50) according to patients' OS. The cohort of 50 patients was divided into 3 groups: 12 patients (24%) with OS > to the third quartile, 26 patients (52%) with an OS between the first and the third quartiles and 12 patients (24%) with an OS < to the first quartile were included in this analysis. 126 proteins have shown significance between these 3 groups of patients defined by their OS (**Supp. Data 7**). Then, using Cox model, 30 proteins were significant with a p=0.05 and 9 with a p= 0.01 (**Table 6**). After a step by step analysis and bootstrap procedure, 6 proteins remained significantly correlated with survival: MAOB, IGHM, HSPD1, LASP1, CFH and ANXA6 (**Figure 3A**). Based on the expression of these 6 proteins, 3 clusters of patients were identified (respectively groups 1, 2 and 3) (**Figure 3B**) and a decisional tree has been built (**Figure 3C**). The OS of the patients from the 3 clusters was determined (**Figure 3D**). Patients of cluster 2 (n=15) had the worst OS, whereas patients of cluster 3 (n=14), the highest and cluster 1 (n=21) is intermediate (**Figure 3D**). We then wanted to know whether these 3 clusters could be associated to the 3 color groups we have previously described. For that, we compared the expression of the 30 prognostic markers identified between the 3 clusters and the 3 groups from MALDI-MSI (**Figure 3E**). We found that 17 of them are over-expressed in group B and in cluster 2 (the worst OS) (**Figure 3E**), among which we can find IGHM, ANXA6 and MAOB. These 3 proteins are bad prognosis indicators. The hazard ratio of these 3 proteins is superior to 1 and more the value is high, more the OS is low. Therefore, MAOB is the protein with the worst prognostic value, followed by ANXA6 and IGHM (**Figure 3A**). These proteins are mainly related to macrophages and microglia activation. Another cluster of proteins is overexpressed in cluster 3 (the most favorable OS) and highly represented in group A, including LASP1 and HSPD1 which are therefore good prognosis indicators. The hazard ratio of these 2 proteins is inferior to 1 (**Figure 3A**), with HSPD1 the marker associated to the better prognosis, followed by LASP1. Regarding CFH, we find it more expressed in cluster 2 (**Figure**

3E) although it is a marker of good prognosis. In fact, CFH has a hazard ratio value of 0.77 (**Figure 3A**), however this is the marker with the lower strength compared to the two other good prognosis markers (HSPD1, LASP1). We confirmed the expression of these markers by box plot analysis in the three clusters (**Figure 3F**).

The analysis of the proteomic data, with the OpenProt alternative database, also allowed the detection of 29 AltProts in relation to the OS (**Supp. Figure 5, Supp. Table 2**). Among these 29 AltProts, 22 were overexpressed in the cluster of patients having the worst OS (cluster 2). The seven others, despite a significant variation, do not allow to differentiate the survival groups. From the 29 AltProts, 15 are issued from ncRNA (51%), emphasizing once again the importance of ncRNA in cancer development.

We then rebuild the classification model of the SpiderMass data by taking into account the 3 clusters of OS. The LDA gave a good separation of the 3 clusters (**Figure 3G**). Validation using 20% patient out method gave 92% of good classification excluding outliers and 90.64% including outliers, confirming the specificity of the OS classification (**Figure 3G See Table Inset**).

Finally, considered the 6 specific markers (MAOB, IGHM, HSPD1, LASP1, CFH and ANX6) linked to the OS clusters, we validated them by immunohistochemistry (**Figure 4A**). We confirmed in patients from cluster 2 (worst OS), higher expression of ANXA6, IGHM and MAOB, which correlate well with the proteomic data (**Figure 4B**). HSPD1 and LASP1 are higher expressed in the tumors of cluster 3 (longer OS), validating again our previous results. CFH is slightly more expressed in tumors of cluster 3, which is expected since it is a marker of good prognosis. Finally, patients belonging to the intermediate OS cluster (cluster 1) present a mixed signature with a higher expression of both good (HSPD1, LASP1) and bad prognosis markers (ANXA6, IGHM).

Taken together, we established a new classification of glioblastoma based on molecular and cellular features which can be very useful for the development of new therapeutic options and personalized medicine for patients.

Discussion

Here, we investigated glioblastoma biology by a proteomic approach at a low spatial resolution in order to capture the tumor microenvironment. A non-targeted MALDI-MSI analysis followed by spatial segmentation using different algorithms allowed to derive a new subclassification of glioblastoma. We successfully validated these observations with

SpiderMass technology with a 93% good classification. Three subgroups were identified (A-Red, B-Yellow and C-Blue regions). In order to decode the biological pathways involved in these three groups, we performed a spatially resolved proteomic that confirmed the data. Molecular signatures of different tumor subtypes were identified among the groups.

Group A is associated with neuro-developmental genes, characteristic of neuronal/glia lineages or neural progenitor cells (NPC) (**Figure 2B**). These included nervous system development markers (like CRMP family, GAP43, MAPT), oligodendrocyte development and differentiation markers (like ABI1, ASPA, CNP, CNTNAP1), stem and progenitor cell signatures (like TRIM2). However, NPC-like state is correlated with presence of markers for immature neurons (beta-3-tubulin), markers for mature neurons (NeuN) and markers indicative for synapses (synaptophysin, SV2A) (Beier et al., 2018). In our data, we found Stathmin 1, NEFH, NEFM and NEFL (Neftel et al., 2019) which are also markers of the NPC-like state of the GSC. Patients of group B present the lowest OS and proteins enriched in this group are linked to immune status with macrophages infiltration, (**Figure 2C**) such as complement factors, immunoglobulin heavy and light chains (IGHG2; IGKC; IGHM; IGHG1; IGLC6 and IGHA1), macrophage markers (CAPG) and coagulation cascade proteins (HP, KNG1, HRG, TTR, GCA, S100A9, STOM). These results are in line with a study of (Cheng et al., 2016), in which eight immune related genes (FOXO3, IL6, IL10, ZBTB16, CCL18, AIMP1, FCGR2B, and MMP9) were identified and used as unfavorable prognostic markers in GBM. High-risk patients conferred an enhanced intensity of local immune response compared to low-risk ones. From the 8 signature genes, AIMP2 has been identified in group B also. Patients of group B also express GSC markers but with a “stem-to-invasion” path. CD44, NES and VIM, enriched in group B, are markers of the mesenchymal like state. The presence of class I self-antigen HLA proteins (HLA-A3 and HLA-B07) in group B is interesting since a positive correlation between HLA and some cancers has been demonstrated, such as cervical or nasopharyngeal carcinomas (Machulla et al., 2001). In a previous study based on HLA antigen frequencies in patients with glioma compared with the control population, patients positive for HLA-A*25 had a 3.0-fold increased risk of glioma ($p = 0.04$), patients positive for HLA-B*27, a 2.7-fold risk ($p = 0.03$). In contrast, the relationship between HLA-B*07 and glioma is very rare (Tang et al., 2005), as well as for HLA-A*3 (Zhang et al., 2007). Taken together, these data clearly confirmed that there is interpatient molecular heterogeneity that may be related to tumor phenotype and cellular plasticity (Neftel et al., 2019) but not directly with transcriptional classification of glioblastoma (proneural, neural, classic and mesenchymal) (Verhaak et al., 2010). Finally, systemic biology analyses revealed that group C is linked to an anti-viral immune response and viral infection. Recent studies have reported a link between

glioblastoma and perinatal viral exposure (Akhtar et al., 2018; Dickinson et al., 2002; Limam et al., 2019; Strojnik et al., 2017). It has recently been shown that Epstein-Barr virus can be implicated in GBM etiology (Zavala-Vega et al., 2019). Moreover, some recent works have also reported that Cytomegalovirus (CMV) promotes murine glioblastoma growth via pericytes recruitment and angiogenesis (Krenzlin et al., 2019). In human, CMV nucleic acids and proteins have been observed within GBM tumor tissue (Rahman et al., 2019).

The comparison with the TCGA specific glioma genes signature showed that 28 of them were statistically different among the 3 groups identified in our study. Most of the proteins were identified in group A and are related to nervous system development, neuron differentiation axon guidance, 3 proteins were identified in group B and are linked to cytokine secretion and 5 both in groups A & C related to Notch signalling. Notch signaling is an evolutionarily conserved pathway that regulates important biological processes, such as cell proliferation, apoptosis, migration, self-renewal, and differentiation. Growing evidence reveals that Notch signaling is highly active in glioma stem cells, in which it suppresses differentiation and maintains stem-like properties, contributing to glioblastoma tumorigenesis and conventional-treatment resistance (Bazzoni and Bentivegna, 2019).

Our study includes both IDH-wildtype and IDH-mutant glioblastomas. The WHO 2016 classification of glioma did not distinguish IDH-wildtype from IDH-mutant tumors. However, recent studies suggest that they represent distinct clinical and genetic entities (Brat et al., 2020a). The present study was designed before this recent knowledge. Although, the cohort includes only 4 IDH-mutant tumors, their proteome did not impact the classification. Moreover, the OS of the IDH-wildtype and IDH-mutant patients of the cohort is also similar. No marked differences are observed. These results are first observations and future studies including a bigger cohort of IDH-mutant tumors are needed to confirm our observations.

Taken together, we established that glioblastomas can be divided into 3 groups. Each group has a distinct molecular pattern, reflecting a specific molecular phenotype of the tumors. These different groups may be explained by an early differentiation due to the presence, in primary tumours, of subpopulation of cells with a distinct functional profile as well as the existence of cells with a high invasive potency. A recent study (Pang et al., 2019) proposed that glioblastoma stem cells (GSCs) acquire a high invasive activity through a mechanism called the 'stem-to-invasion path' and that long noncoding RNAs are one of the key factors. It has been demonstrated that these non-coding genomic regions can result to the synthesis of proteins, so called alternative proteins forming an unexplored ghost proteome with function in cancer (Delcourt et al., 2018). 170 alternative proteins (AltProts) were found significantly

variable in the three groups identified above. If the function of these AltProts are still poorly understood, it has been shown that they can have a role in regulation of transcription and can also be present in extracellular vesicles (Murgoci et al., 2020). Finally, more than 50% of the AltProts identified in the present study come from the translation of ncRNAs transcribed from pseudogenes. Recently it has been demonstrating that pseudogenes can also be used as signatures for glioma prognosis. 6 pseudogenes (SP3P, ANXA2P3, PTTG3P, LPAL2, CLCA3P, and TDH) have been found to be associated with overall survival in glioma (Gao et al., 2015). Nine other pseudogenes (TP73-AS1, AC078883.3, RP11-944L7.4, HAR1B, RP4-635E18.7, HOTAIR, SAPCD1-AS1, AC104653.1, and RP5-1172N10.2.) constitute a set of prognosis markers to predict survival of patients with glioma (Lei et al., 2018). All these results provide novel insights into the biological role of pseudogenes in cancer and especially in glioma. Additionally, the novel identified AltProts translated from ncRNAs add additional information to the already known pseudogenes in glioma.

Finally, this study allowed us to highlight the presence of 6 new prognostic proteins for glioblastoma: LASP1, HSPD1 and CFH are favourable prognostic markers while IGHM, ANXA6 and MAOB are unfavourable prognostic markers when over-expressed. We found that the expression of these unfavourable protein markers is higher in the tumors belonging to the cluster 2 (worst OS) and group B. Moreover, 2 known unfavourable prognostic proteins, coming from the Human Pathology Atlas, are also overexpressed in the group B (AEBP1 and PDIA4). The 2 other clusters are associated to a better overall survival, the associated markers are of good prognosis.

If we come back to the molecular signature of the patients of group B, we have demonstrated that the proteins are mainly involved in inflammation, phagocytosis, oxidative stress, and ischemia and are related to macrophages infiltration. The brain neoplasm cells could influence the polarization state of the tumor associated macrophages (Guadagno et al., 2018). In turn, innate immunity cells have a decisive role through regulation of the acquired immune response, but also through humoral cross-talking with cancer cells in the tumor microenvironment. Thus, we can assume that immunotherapeutic strategies would be much more suited for patients of group B. Further studies need to be conducted to see the correlation between the proteomic signatures we highlight and response to treatment. In this context, our study could be an interesting starting point to guide the development of new personalized therapeutic strategies and better treatment decisions.

Acknowledgments

This research was supported by grants from the Ministère de L'Éducation Nationale, de L'Enseignement Supérieur et de la Recherche, ANR (IF), SIRIC ONCOLille (MS), Grant INCa-DGOS-Inserm 6041aa (IF, MS), and INSERM, Ligue contre le Cancer (EL) .

AUTHORS CONTRIBUTION

Conceptualization : MS, IF, ELR, Methodology : MD, MW, LD, MS, IF, JP, SA, NO, PDC
Software : LD, MW, TC, FZ; Validation: MD, LD, CAM,FE Formal Analysis : MS, MD, LD, TC,
MW; Investigation : MD, MW, LD, ELR, IF, MS, Resources : Data curation : MW, MS, LD, MD,
Writing : MS, MD, ELR, MW, TC, LD, MW Original Draft : MS, LD, MD, MW Supervision, Project
: MW, ELR, IF, MS Administration : ELR, MS, IF Funding Acquisition IF, MS; MW

The authors declare no competing interests.

Figures Legends

Figure 1 : Histological, MALDI MSI and SpiderMass data

- A. General workflow of the MALDI-MS imaging combined with microproteomics used for glioma classification.
- B. Representative annotated histopathology images of three glioblastoma samples and their corresponding segmentation map obtained from MALDI-MSI data. Colors represent molecularly different regions. Note that for 2 different tissues, similar colors are not equivalent to similar molecular groups. The segmentation map shows different clusters for each case and non-observable with HES coloration.
- C. Global segmentation maps of all tissues together after MALDI-MSI analysis. Colors represent molecularly different regions as shown in the corresponding dendrogram. The segmentation map gives 3 main clusters.
- D. Cases are grouped together based on their molecular profiles: group A (n=13) contains cases predominantly red, group B (n=9) contains cases predominantly yellow and group C (n=23) contains cases predominantly blue
- E. The built PCA-LDA classification model based on 3 glioma groups; Group A (red), Group B (yellow), Group C (blue). a) LDA representation of the 3-class PCA-LDA (right). The table (right) represents the “20% out” and “leave-one-patient-out” cross-validation results of the built classification model. b) LD2 loading spectra (top) indicate the discrimination between Group A (red) and Group B (yellow). The 10 most discriminatory lipid peaks are indicated by the blue dash line. LD1 loading spectra (bottom) indicate the discrimination between Group A (red) and Group C (blue). The 10 most discriminatory lipid peaks are indicated by the blue dash line.

Figure 2. Shotgun microproteomics analysis

- A. Heatmap of proteins with different regulation profiles as determined after label free quantification in the three groups highlighting the presence of 3 clusters. Shotgun proteomics was performed after on-tissue trypsin digestion followed by microextraction at the spots determined from MALDI MSI data.
- B. Pathway analysis of proteins overexpressed in group A reveals that a large majority of protein is involved in (a) neurogenesis, brain development, axon development and (b) more generally neurodegenerative disease and nerve degeneration.

- C. Pathway analysis of proteins overexpressed in group B reveals that majority of protein are involved in (a) injuries, inflammation and more generally immune system response but also in (b) blood coating and vascularization.
- D. Pathway analysis of proteins overexpressed in group C shows implication in (a) cancer, cell growth, DNA repair and viral reproduction with protein involved in (b) RNA splicing, metabolism and replication.
- E. Heatmap of alternatives proteins with different regulation profiles as determined after label free quantification in the three groups highlighting the presence of 3 clusters.

Figure 3. Proteomic and lifetime analysis

- A. Analysis of maximum likelihood estimates of the 6 proteins significantly correlated with survival (MAOB, IGHM, HSPD1, LASP1, CFH and ANXA6) identified after a step by step analysis and bootstrap procedure and B. patient clustering based on these proteins
- C. Decisional tree
- D. Overall survival of each identified group reveal that cluster 2 observe the better prognosis whereas cluster 3 has the lower survival
- E. Heatpmap of the 30 proteins significant after Cox model ($p=0.05$) between the 3 groups of patients defined by their OS (left) and expression of theses 30 proteins in each imaging group group A B and C (right)
- F. Boxplot of the 6 proteins used for clustering according to the OS cluster
- G. The built PCA-LDA classification model based on 3 OS clusters. The table (left) represents the “20% out” cross-validation rates of the built model. LDA representation (right) of the 3-class OS cluster; Cluster 1 (red), Cluster 2 (cyan) and Cluster 3 (blue). LDA 1 separates Cluster 1 and Cluster 2, while LD2 separates Cluster 1 and Cluster 3.

Figure 4. Immunohistochemistry of the 6 survival markers identified.

- A. Representative fluorescence images of the 6 proteins in the three OS clusters of patients. MAOB, IGHM and ANXA6 are associated to a bad prognosis while CFH, LASP1 and HSPD1 are related to a good prognosis. Images were acquired with a confocal microscope at 40x magnification.
- B. Quantification of fluorescence intensity of the 6 proteins in the three OS clusters. Images taken from 15 tumors of cluster 1, 14 tumors of cluster 2 and 10 tumors of cluster 3 were quantified. For each tumor, 3 to 4 images were acquired and quantified. Significant differences were identified using unpaired t test with **** $p<0.0001$; *** $p<0.001$; ** $p<0,01$ and * $p<0.05$.

Supplementary Figures

Supp. Figure 1. **A.** Scanned pictures after hematoxylin-eosin staining of the 50 glioblastoma and **B.** anathomo-pathologist annotations. The tissues are classified according to the three groups obtained after MALDI-MSI data segmentation (groups Red, Blue and Yellow). **C.** Spatial segmentation of all tissues and grouping according to the largest molecular area (+50%) represented in each tissue and **D.** MALDI MSI images of characteristics m/z ions for each group. **E.** Ward clustering method give 3 main branches with same characteristic ions. **F.** Principal component analysis (PCA) of each individual spectra reveals separation between the three groups.

Supp Figure 2. **A.** Global survival curve of all patients according the Karnofsky indice, MGMT statut and resection quality and **B.** survival without progression. 0-1 = Total resection, subtotal 2-3 = partial resection biopsies

Supp. Figure 3. **A. a)** Analysis of proteins overexpressed in group A shows an involvement in axon guidance. **b)** Proteins overexpressed in group B and mainly involved in complements, coagulation cascade and inflammation **c)** Analysis of overexpressed proteins in group C shows a network of proteins involved in Epstein barr infection. **B.** Correlation between Grancalcin and CAPG expression and glioma patient survival according the TGCA data. Patients were divided based on level of expression into "low" or "high". **C.** Analysis of overall survival of patients reveals no significant difference between the 3 proteomic groups.

Supp. Figure 4 : 1. Heatmap of alternative proteins with different regulation profiles as determined after label free quantification in the three OS groups highlighting the presence of a cluster of overexpressed alternative proteins in group 2 which is the group having the worst OS.

DATA AVAILIBILITY

Proteomic datasets including MaxQuant files and annotated MS/MS datasets were uploaded to the ProteomeXchange Consortium via the PRIDE database, and then assigned the dataset identifier PXD016165".

References

- Akhtar, S., Vranic, S., Cyprian, F. S., and Al Moustafa, A.-E. J. F. i. o. (2018). Epstein–Barr virus in gliomas: cause, association, or artifact? 8, 123.
- Alexandrov, T., Becker, M., Deininger, S. O., Ernst, G., Wehder, L., Grasmair, M., von Eggeling, F., Thiele, H., and Maass, P. (2010). Spatial segmentation of imaging mass spectrometry data with edge-preserving image denoising and clustering. *Journal of proteome research* 9, 6535-6546.
- Balog, J., Sasi-Szabo, L., Kinross, J., Lewis, M. R., Muirhead, L. J., Veselkov, K., Mirnezami, R., Dezso, B., Damjanovich, L., Darzi, A., et al. (2013). Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. *Sci Transl Med* 5, 194ra193.
- Bazzoni, R., and Bentivegna, A. J. C. (2019). Role of notch signaling pathway in glioblastoma pathogenesis. 11, 292.
- Beier, C. P., Rasmussen, T., Dahlrot, R. H., Tenstad, H. B., Aarø, J. S., Sørensen, M. F., Heimisdóttir, S. B., Sørensen, M. D., Svenningsen, P., and Riemenschneider, M. J. J. S. r. (2018). Aberrant neuronal differentiation is common in glioma but is associated neither with epileptic seizures nor with better survival. 8, 1-12.
- Brat, D. J., Aldape, K., Colman, H., Figarella-Branger, D., Fuller, G. N., Giannini, C., Holland, E. C., Jenkins, R. B., Kleinschmidt-DeMasters, B., and Komori, T. (2020a). cIMPACT-NOW update 5: recommended grading criteria and terminologies for IDH-mutant astrocytomas. *Acta neuropathologica* 139, 603-608.
- Brat, D. J., Aldape, K., Colman, H., Figarella-Branger, D., Fuller, G. N., Giannini, C., Holland, E. C., Jenkins, R. B., Kleinschmidt-DeMasters, B., and Komori, T. J. A. n. (2020b). cIMPACT-NOW update 5: recommended grading criteria and terminologies for IDH-mutant astrocytomas. 139, 603-608.
- Brat, D. J., Aldape, K., Colman, H., Holland, E. C., Louis, D. N., Jenkins, R. B., Kleinschmidt-DeMasters, B., Perry, A., Reifenberger, G., and Stupp, R. (2018a). cIMPACT-NOW update 3: recommended diagnostic criteria for “Diffuse astrocytic glioma, IDH-wildtype, with molecular features of glioblastoma, WHO grade IV”. *Acta neuropathologica* 136, 805-810.
- Brat, D. J., Aldape, K., Colman, H., Holland, E. C., Louis, D. N., Jenkins, R. B., Kleinschmidt-DeMasters, B., Perry, A., Reifenberger, G., and Stupp, R. J. A. n. (2018b). cIMPACT-NOW update 3: recommended diagnostic criteria for “Diffuse astrocytic glioma, IDH-wildtype, with molecular features of glioblastoma, WHO grade IV”. 136, 805-810.
- Brennan, C. W., Verhaak, R. G., McKenna, A., Campos, B., Noushmehr, H., Salama, S. R., Zheng, S., Chakravarty, D., Sanborn, J. Z., and Berman, S. H. J. C. (2013). The somatic genomic landscape of glioblastoma. 155, 462-477.

- Brunet, M. A., Brunelle, M., Lucier, J. F., Delcourt, V., Levesque, M., Grenier, F., Samandi, S., Leblanc, S., Aguilar, J. D., Dufour, P., et al. (2018). OpenProt: a more comprehensive guide to explore eukaryotic coding potential and proteomes. *Nucleic acids research*.
- Capper, D., Jones, D. T., Sill, M., Hovestadt, V., Schrimpf, D., Sturm, D., Koelsche, C., Sahm, F., Chavez, L., and Reuss, D. E. J. N. (2018). DNA methylation-based classification of central nervous system tumours. *555*, 469-474.
- Cardon, T., Franck, J., Coyaud, E., Laurent, E. M. N., Damato, M., Maffia, M., Vergara, D., Fournier, I., and Salzet, M. (2020a). Alternative proteins are functional regulators in cell reprogramming by PKA activation. *Nucleic acids research*.
- Cardon, T., Herve, F., Delcourt, V., Roucou, X., Salzet, M., Franck, J., and Fournier, I. (2020b). Optimized Sample Preparation Workflow for Improved Identification of Ghost Proteins. *Anal Chem* 92, 1122-1129.
- Cheng, W., Ren, X., Zhang, C., Cai, J., Liu, Y., Han, S., and Wu, A. J. N. (2016). Bioinformatic profiling identifies an immune-related risk signature for glioblastoma. *86*, 2226-2234.
- Chinot, O. L., Wick, W., Mason, W., Henriksson, R., Saran, F., Nishikawa, R., Carpentier, A. F., Hoang-Xuan, K., Kavan, P., Cernea, D., et al. (2014). Bevacizumab plus radiotherapy-temozolomide for newly diagnosed glioblastoma. *N Engl J Med* 370, 709-722.
- Cox, J., Hein, M. Y., Lubner, C. A., Paron, I., Nagaraj, N., and Mann, M. (2014). Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Molecular & cellular proteomics : MCP* 13, 2513-2526.
- Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol* 26, 1367-1372.
- Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., and Mann, M. (2011). Andromeda: a peptide search engine integrated into the MaxQuant environment. *Journal of proteome research* 10, 1794-1805.
- Deighton, R. F., McGregor, R., Kemp, J., McCulloch, J., and Whittle, I. R. J. B. P. (2010). Glioma pathophysiology: insights emerging from proteomics. *20*, 691-703.
- Delcourt, V., Staskevicius, A., Salzet, M., Fournier, I., and Roucou, X. (2018). Small Proteins Encoded by Unannotated ORFs are Rising Stars of the Proteome, Confirming Shortcomings in Genome Annotations and Current Vision of an mRNA. *Proteomics* 18, e1700058.
- Dickinson, H., Nyari, T., and Parker, L. J. B. j. o. c. (2002). Childhood solid tumours in relation to infections in the community in Cumbria during pregnancy and around the time of birth. *87*, 746-750.

- Dilillo, M., Ait-Belkacem, R., Esteve, C., Pellegrini, D., Nicolardi, S., Costa, M., Vannini, E., De Graaf, E., Caleo, M., and McDonnell, L. J. S. r. (2017). Ultra-high mass resolution MALDI imaging mass spectrometry of proteins and metabolites in a mouse model of glioblastoma. *7*, 1-11.
- Fatou, B., Saudemont, P., Leblanc, E., Vinatier, D., Mesdag, V., Wisztorski, M., Focsa, C., Salzet, M., Ziskind, M., and Fournier, I. (2016). In vivo Real-Time Mass Spectrometry for Guided Surgery Application. *Scientific reports* 6, 25919.
- Fournier, I., Day, R., and Salzet, M. (2003). Direct analysis of neuropeptides by in situ MALDI-TOF mass spectrometry in the rat brain. *Neuro Endocrinol Lett* 24, 9-14.
- Gao, K.-M., Chen, X.-c., Zhang, J.-x., Wang, Y., Yan, W., You, Y.-P. J. J. o. e., and research, c. c. (2015). A pseudogene-signature in glioma predicts survival. *34*, 23.
- Gilbert, M. R., Dignam, J. J., Armstrong, T. S., Wefel, J. S., Blumenthal, D. T., Vogelbaum, M. A., Colman, H., Chakravarti, A., Pugh, S., and Won, M. (2014). A randomized trial of bevacizumab for newly diagnosed glioblastoma. *New England Journal of Medicine* 370, 699-708.
- Goplen, D., Bougnaud, S., Rajcevic, U., Bøe, S. O., Skaftnesmo, K. O., Voges, J., Enger, P. Ø., Wang, J., Tysnes, B. B., and Laerum, O. D. J. T. A. j. o. p. (2010). α B-crystallin is elevated in highly infiltrative apoptosis-resistant glioblastoma cells. *177*, 1618-1628.
- Gramatzki, D., Dehler, S., Rushing, E. J., Zaugg, K., Hofer, S., Yonekawa, Y., Bertalanffy, H., Valavanis, A., Korol, D., and Rohrmann, S. J. C. (2016). Glioblastoma in the Canton of Zurich, Switzerland revisited: 2005 to 2009. *122*, 2206-2215.
- Guadagno, E., Presta, I., Maisano, D., Donato, A., Pirrone, C. K., Cardillo, G., Corrado, S. D., Mignogna, C., Mancuso, T., and Donato, G. J. I. j. o. m. s. (2018). Role of macrophages in brain tumor growth and progression. *19*, 1005.
- Gutmann, D. H., Huang, Z. Y., Hedrick, N. M., Ding, H., Guha, A., Watson, M. A. J. A. o. N. O. J. o. t. A. N. A., and Society, t. C. N. (2002). Mouse glioma gene expression profiling identifies novel human glioma-associated genes. *51*, 393-405.
- Hu, L. S., Brat, D. J., Bloch, O., Ramkissoon, S., and Lesser, G. J. J. A. S. o. C. O. E. B. (2020). The Practical Application of Emerging Technologies Influencing the Diagnosis and Care of Patients With Primary Brain Tumors. *40*, 1-12.
- Huang, Z.-y., Wu, Y., Burke, S. P., and Gutmann, D. H. J. C. r. (2003). The 43,000 growth-associated protein functions as a negative growth regulator in glioma. *63*, 2933-2939.
- Jovčevska, I., Zupanec, N., Urlep, Ž., Vranič, A., Matos, B., Stokin, C. L., Muyldermans, S., Myers, M. P., Buzdin, A. A., and Petrov, I. J. O. (2017). Differentially expressed proteins in glioblastoma multiforme identified with a nanobody-based anti-proteome approach and

- confirmed by OncoFinder as possible tumor-class predictive biomarker candidates. 8, 44141.
- Kalinina, J., Peng, J., Ritchie, J. C., and Van Meir, E. G. J. N.-o. (2011). Proteomics of gliomas: initial biomarker discovery and evolution of technology. 13, 926-942.
- Klein, O., Strohschein, K., Nebrich, G., Oetjen, J., Trede, D., Thiele, H., Alexandrov, T., Giavalisco, P., Duda, G. N., and von Roth, P. J. P. (2014). MALDI imaging mass spectrometry: discrimination of pathophysiological regions in traumatized skeletal muscle by characteristic peptide signatures. 14, 2249-2260.
- Krenzlin, H., Behera, P., Lorenz, V., Passaro, C., Zdioruk, M., Nowicki, M. O., Grauwet, K., Zhang, H., Skubal, M., and Ito, H. J. T. J. o. c. i. (2019). Cytomegalovirus promotes murine glioblastoma growth via pericyte recruitment and angiogenesis. 129.
- Lei, B., Yu, L., Jung, T. A., Deng, Y., Xiang, W., Liu, Y., and Qi, S. J. J. o. N. S. P. A. C. E. N. (2018). Prospective series of nine long noncoding RNAs associated with survival of patients with glioblastoma. 79, 471-478.
- Lemaire, R., Menguellet, S. A., Stauber, J., Marchaudon, V., Lucot, J. P., Collinet, P., Farine, M. O., Vinatier, D., Day, R., Ducoroy, P., et al. (2007). Specific MALDI imaging and profiling for biomarker hunting and validation: fragment of the 11S proteasome activator complex, Reg alpha fragment, is a new potential ovary cancer biomarker. *Journal of proteome research* 6, 4127-4134.
- Lemaire, R., Tabet, J. C., Ducoroy, P., Hendra, J. B., Salzet, M., and Fournier, I. (2006a). Solid ionic matrixes for direct tissue analysis and MALDI imaging. *Anal Chem* 78, 809-819.
- Lemaire, R., Wisztorski, M., Desmons, A., Tabet, J. C., Day, R., Salzet, M., and Fournier, I. (2006b). MALDI-MS direct tissue analysis of proteins: Improving signal sensitivity using organic treatments. *Anal Chem* 78, 7145-7153.
- Limam, S., Missaoui, N., Mestiri, S., Yacoubi, M., Krifa, H., Selmi, B., and Mokni, M. J. C. r. i. t. m. (2019). Epstein-Barr virus infection in gliomas. 67, 129-133.
- Louis, D., Suva, M., Burger, P., Kleihues, P., Aldape, K., Hegi, M., and von Deimling, A. J. W. c. o. t. o. t. c. n. s. (2016). Glioblastoma, IDH wild-type. 28-45.
- Machulla, H. K., Steinborn, F., Schaaf, A., Heidecke, V., and Rainov, N. G. J. J. o. n.-o. (2001). Brain glioma and human leukocyte antigens (HLA)—is there an association. 52, 253-261.
- Moutal, A., Honnorat, J., Massoma, P., Désormeaux, P., Bertrand, C., Malleval, C., Watrin, C., Chounlamountri, N., Mayeur, M.-E., and Besançon, R. J. C. r. (2015). CRMP5 controls glioblastoma cell proliferation and survival through notch-dependent signaling. 75, 3519-3528.

- Murgoci, A.-N., Cardon, T., Aboulouard, S., Duhamel, M., Fournier, I., Cizkova, D., and Salzet, M. J. I. (2020). Reference and Ghost proteins identification in Rat C6 glioma extracellular vesicles. 101045.
- Nature, C. G. A. R. N. J. (2008). Comprehensive genomic characterization defines human glioblastoma genes and core pathways. 455, 1061.
- Nettel, C., Laffy, J., Filbin, M. G., Hara, T., Shore, M. E., Rahme, G. J., Richman, A. R., Silverbush, D., Shaw, M. L., and Hebert, C. M. J. C. (2019). An integrative model of cellular states, plasticity, and genetics for glioblastoma. 178, 835-849. e821.
- Ogrinc, N., Saudemont, P., Balog, J., Robin, Y. M., Gimeno, J. P., Pascal, Q., Tierny, D., Takats, Z., Salzet, M., and Fournier, I. (2019). Water-assisted laser desorption/ionization mass spectrometry for minimally invasive in vivo and real-time surface analysis using SpiderMass. *Nature protocols* 14, 3162-3182.
- Ostrom, Q. T., Gittleman, H., Truitt, G., Boscia, A., Kruchko, C., and Barnholtz-Sloan, J. S. J. N.-o. (2018a). CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2011–2015. 20, iv1-iv86.
- Ostrom, Q. T., Kinnersley, B., Armstrong, G., Rice, T., Chen, Y., Wiencke, J. K., McCoy, L. S., Hansen, H. M., Amos, C. I., Bernstein, J. L., et al. (2018b). Age-specific genome-wide association study in glioblastoma identifies increased proportion of 'lower grade glioma'-like features associated with younger age. *Int J Cancer* 143, 2359-2366.
- Otasek, D., Morris, J. H., Bouças, J., Pico, A. R., and Demchak, B. J. G. b. (2019). Cytoscape Automation: empowering workflow-based network analysis. 20, 1-15.
- Pang, B., Xu, J., Hu, J., Guo, F., Wan, L., Cheng, M., and Pang, L. J. M. o. (2019). Single-cell RNA-seq reveals the invasive trajectory and molecular cascades underlying glioblastoma progression. 13, 2588-2603.
- Pathan, M., Keerthikumar, S., Chisanga, D., Alessandro, R., Ang, C. S., Askenase, P., Batagov, A. O., Benito-Martin, A., Camussi, G., Clayton, A., et al. (2017). A novel community driven software for functional enrichment analysis of extracellular vesicles data. *J Extracell Vesicles* 6, 1321455.
- Petryszak, R., Keays, M., Tang, Y. A., Fonseca, N. A., Barrera, E., Burdett, T., Füllgrabe, A., Fuentes, A. M.-P., Jupp, S., and Koskinen, S. J. N. a. r. (2016). Expression Atlas update—an integrated database of gene and protein expression in humans, animals and plants. 44, D746-D752.
- Quanico, J., Franck, J., Daully, C., Strupat, K., Dupuy, J., Day, R., Salzet, M., Fournier, I., and Wisztorski, M. (2013). Development of liquid microjunction extraction strategy for improving protein identification from tissue sections. *Journal of proteomics* 79, 200-218.

- Rahman, M., Dastmalchi, F., Karachi, A., and Mitchell, D. J. O. (2019). The role of CMV in glioblastoma and implications for immunotherapeutic strategies. *8*, e1514921.
- Saudemont, P., Quanico, J., Robin, Y. M., Baud, A., Balog, J., Fatou, B., Tierny, D., Pascal, Q., Minier, K., Pottier, M., et al. (2018). Real-Time Molecular Diagnosis of Tumors Using Water-Assisted Laser Desorption/Ionization Mass Spectrometry Technology. *Cancer Cell* *34*, 840-851 e844.
- Schafer, K. C., Denes, J., Albrecht, K., Szaniszló, T., Balog, J., Skoumal, R., Katona, M., Toth, M., Balogh, L., and Takats, Z. (2009). In vivo, in situ tissue analysis using rapid evaporative ionization mass spectrometry. *Angew Chem Int Ed Engl* *48*, 8240-8242.
- Strojnik, T., Duh, D., and Lah, T. T. J. i. v. (2017). Prevalence of neurotropic viruses in malignant glioma and their onco-modulatory potential. *31*, 221-229.
- Stupp, R., Hegi, M. E., Mason, W. P., van den Bent, M. J., Taphoorn, M. J., Janzer, R. C., Ludwin, S. K., Allgeier, A., Fisher, B., Belanger, K., et al. (2009). Effects of radiotherapy with concomitant and adjuvant temozolomide versus radiotherapy alone on survival in glioblastoma in a randomised phase III study: 5-year analysis of the EORTC-NCIC trial. *The Lancet Oncology* *10*, 459-466.
- Stupp, R., Taillibert, S., Kanner, A., Read, W., Steinberg, D., Lhermitte, B., Toms, S., Idbaih, A., Ahluwalia, M. S., Fink, K., et al. (2017). Effect of Tumor-Treating Fields Plus Maintenance Temozolomide vs Maintenance Temozolomide Alone on Survival in Patients With Glioblastoma: A Randomized Clinical Trial. *JAMA* *318*, 2306-2316.
- Sturm, D., Witt, H., Hovestadt, V., Khuong-Quang, D.-A., Jones, D. T., Konermann, C., Pfaff, E., Tönjes, M., Sill, M., and Bender, S. J. C. c. (2012). Hotspot mutations in H3F3A and IDH1 define distinct epigenetic and biological subgroups of glioblastoma. *22*, 425-437.
- Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., and Bork, P. J. N. a. r. (2019). STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *47*, D607-D613.
- Tang, J., Shao, W., Dorak, M. T., Li, Y., Miike, R., Lobashevsky, E., Wiencke, J. K., Wensch, M., Kaslow, R. A., Cobbs, C. S. J. C. E., and Biomarkers, P. (2005). Positive and negative associations of human leukocyte antigen variants with the onset and prognosis of adult glioblastoma multiforme. *14*, 2040-2044.
- Trede, D., Kobarg, J. H., Oetjen, J., Thiele, H., Maass, P., and Alexandrov, T. J. J. o. i. b. (2012). On the importance of mathematical methods for analysis of MALDI-imaging mass spectrometry data. *9*, 1-11.

- Tyanova, S., Temu, T., Carlson, A., Sinitcyn, P., Mann, M., and Cox, J. (2015). Visualization of LC-MS/MS proteomics data in MaxQuant. *Proteomics* 15, 1453-1456.
- Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhor, G., Benfeitas, R., Arif, M., Liu, Z., and Edfors, F. J. S. (2017). A pathology atlas of the human cancer transcriptome. *357*, eaan2507.
- Vanderperre, B., Lucier, J. F., Bissonnette, C., Motard, J., Tremblay, G., Vanderperre, S., Wisztorski, M., Salzet, M., Boisvert, F. M., and Roucou, X. (2013). Direct detection of alternative open reading frames translation products in human significantly expands the proteome. *PLoS One* 8, e70698.
- Vanderperre, B., Lucier, J. F., and Roucou, X. (2012). HALtORF: a database of predicted out-of-frame alternative open reading frames in human. *Database : the journal of biological databases and curation* 2012, bas025.
- Verhaak, R. G., Hoadley, K. A., Purdom, E., Wang, V., Qi, Y., Wilkerson, M. D., Miller, C. R., Ding, L., Golub, T., and Mesirov, J. P. J. C. c. (2010). Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1. *17*, 98-110.
- Voigt, A., Nowick, K., and Almaas, E. J. P. c. b. (2017). A composite network of conserved and tissue specific gene interactions reveals possible genetic interactions in glioma. *13*, e1005739.
- Weathers, S. P., and Gilbert, M. R. (2014). Advances in treating glioblastoma. *F1000Prime Rep* 6, 46.
- Weller, M., Butowski, N., Tran, D. D., Recht, L. D., Lim, M., Hirte, H., Ashby, L., Mechtler, L., Goldlust, S. A., Iwamoto, F., et al. (2017a). Rindopepimut with temozolomide for patients with newly diagnosed, EGFRvIII-expressing glioblastoma (ACT IV): a randomised, double-blind, international phase 3 trial. *The Lancet Oncology* 18, 1373-1385.
- Weller, M., Butowski, N., Tran, D. D., Recht, L. D., Lim, M., Hirte, H., Ashby, L., Mechtler, L., Goldlust, S. A., and Iwamoto, F. J. T. L. O. (2017b). Rindopepimut with temozolomide for patients with newly diagnosed, EGFRvIII-expressing glioblastoma (ACT IV): a randomised, double-blind, international phase 3 trial. *18*, 1373-1385.
- Weller, M., Reifenberger, G., Le Rhun, E., Clarke, J. L., Soffiatti, R., Wick, A., Chinot, O. L., Ducray, F., Hau, P., and McDonald, K. L. (2019). Molecular genetic, host-derived and clinical determinants of long-term survival in glioblastoma: First results from the ETERNITY study (EORTC 1419). In, (American Society of Clinical Oncology).
- Weller, M., Van Den Bent, M., Tonn, J. C., Stupp, R., Preusser, M., Cohen-Jonathan-Moyal, E., Henriksson, R., Le Rhun, E., Balana, C., and Chinot, O. J. T. I. o. (2017c). *European*

- Association for Neuro-Oncology (EANO) guideline on the diagnosis and treatment of adult astrocytic and oligodendroglial gliomas. 18, e315-e329.
- Wisztorski, M., Desmons, A., Quanico, J., Fatou, B., Gimeno, J. P., Franck, J., Salzet, M., and Fournier, I. (2016). Spatially-resolved protein surface microsampling from tissue sections using liquid extraction surface analysis. *Proteomics*.
- Wisztorski, M., Lemaire, R., Stauber, J., Menguelet, S. A., Croix, D., Mathe, O. J., Day, R., Salzet, M., and Fournier, I. (2007). New developments in MALDI imaging for pathology proteomic studies. *Curr Pharm Des* 13, 3317-3324.
- Yuryev, A., Kotelnikova, E., and Daraselia, N. J. E. o. o. d. d. (2009). Ariadne's ChemEffect and Pathway Studio knowledge base. 4, 1307-1318.
- Zaman, S., Chobrutskiy, B. I., Sikaria, D., and Blanck, G. J. O. r. (2019). MAPT (Tau) expression is a biomarker for an increased rate of survival for low-grade glioma. 41, 1359-1366.
- Zavala-Vega, S., Palma-Lara, I., Ortega-Soto, E., Trejo-Solis, C., de Arellano, I. T.-R., Ucharima-Corona, L. E., Garcia-Chacón, G., Ochoa, S. A., Xicohtencatl-Cortes, J., and Cruz-Córdova, A. J. C. R. i. O. (2019). Role of Epstein-Barr Virus in Glioblastoma. 24.
- Zhang, J. G., Eguchi, J., Kruse, C. A., Gomez, G. G., Fakhrai, H., Schroter, S., Ma, W., Hoa, N., Minev, B., and Delgado, C. J. C. C. R. (2007). Antigenic profiling of glioma cells to generate allogeneic vaccines or dendritic cell-based therapeutics. 13, 566-575.

Tables

Table 1. Clinical characteristics

	Total population (n=50)
Gender	
female, n (%)	15 (30)
male, n (%)	35 (70)
Age at diagnosis (years)	
median (IQR)	59 (50-65)
Karnofsky performance score at diagnosis	
median (IQ)	90 (80-100)
0-80, n (%)	15 (30)
90-100, n (%)	35 (70)
Main location of the tumor	
frontal, n (%)	15 (30)
occipital, n (%)	3 (6)
parietal, n (%)	12 (24)
temporal, n (%)	20 (40)
Type of surgical resection	
complete, n (%)	29 (58)
partial, n (%)	20 (40)
biopsies, n (%)	1 (2)
IDH1/2 mutation status*	
Wild-type, n (%)	46 (92)
mutant, n (%)	4 (8)
Nuclear expression of ATRX	
lost, n (%)	7 (14)
maintained, n (%)	43 (86)
MGMT promoter methylation status	
not methylated, n (%)	33 (66)
methylated, n (%)	17 (34)
EGFR amplification	
no, n (%)	26 (52)
yes, n (%)	24 (48)
Chromosome 7 gain combined with chromosome 10 loss (+7/-10)	
no, n (%)	15 (30)
yes, n (%)	35 (70)
EGFR amplification combined with 7 gain / 10 loss	
EGFR amplification or gain 7 / lost 10	42 (84)
EGFR amplification without gain 7 / lost 10	7 (14)
EGFR amplification and gain 7 / lost 10	17 (34)
gain 7 / lost 10 without EGFR amplification	18 (36)
Homozygous CDKN2A deletion	
no, n (%)	19 (39)
yes, n (%)	29 (59)
other: CDKN2A amplification, n (%)	1 (2)
Median follow-up (months)	
median (IQR)	20.5 (13.5-34)

Initial treatment	
RT/TMZ followed by 6 cycles of TMZ, n (%)	20 (40)
RT/TMZ then more than 6 months TMZ, n (%)	4 (8)
RT/TMZ followed by less than 6 cycles of TMZ, n (%)	21 (42)
other treatment**, n (%)	2 (4)
clinical study, n (%)	2 (4)
no treatment, n (%)	1 (2)
Progression	
yes, n (%)	40 (80)
no, n (%)	5 (10)
unknown, n (%)	5 (10)
Progression-free survival (months)	
median (IQR)	10.6 (6.9-16.8)
Treatment at first progression	
yes, n (%)	35*** (87.5)
no, n (%)	5 (12.5)
Death	
yes, n (%)	41 (82)
no, n (%)	9 (18)
Survival from surgery (months)	
median (IQR)	20.5 (13.5-34.5)
Survival	
upper IQR, n (%)	12 (24)
intermediate IQR, n (%)	26 (52)
lower IQR, n (%)	12 (24)

* according to WHO 2016 recommendation

** one patient: RT only, one patient: 6 cycles TMZ then SRT

*** including surgery in 5 cases

Abbreviations:

EGFR, epidermal growth factor receptor

IDH: isocitrate dehydrogenase

IQR: interquartile range

MGMT: O⁶-methylguanine DNA methyltransferase

RT: radiotherapy

SRT: stereotactic radiotherapy

TMZ: temozolomide

Table 2. Clinical characteristics of each groups

COLOR GROUPS (n=45)	A (n=13) (red)	B (n=9) (yellow)	C (n=23) (blue)
Gender			
female, n (%)	4 (31)	2 (22)	7 (30)
male, n (%)	9 (69)	7 (78)	16 (70)
Age at diagnosis (years)			
median (IQR)	48 (43-55)	60 (59-64)	62 (51-68)
Karnofsky performance score at diagnosis			
median (IQ)	90 (70-100)	90 (90-100)	90 (80-95)
0-80, n (%)	5 (38)	0 (0)	8 (35)
90-100, n (%)	8 (62)	9 (100)	15 (65)
Main location of the tumor			
frontal, n (%)	4 (31)	1 (11)	9 (39)
	1 (8)	0 (0)	2 (9)
occipital, n (%)	4 (31)	3 (33)	4 (17)
	4 (31)	5 (55)	7 (30)
parietal, n (%)			
temporal, n (%)			
Type of surgical resection			
complete, n (%)	6 (46)	8 (89)	10 (43)
	7 (54)	1 (11)	12 (52)
partial, n (%)	0	0	1 (4)
biopsies, n (%)			
IDH1/2 mutation status*			
wild-type, n (%)	11 (85)	9 (100)	21 (91)
	2 (15)	0 (0)	2 (9)
mutant, n (%)			
Nuclear expression of ATRX			
lost, n (%)	2 (15)	0 (0)	4 (17)
	11 (85)	9 (100)	19 (83)
maintained, n (%)			
MGMT promoter methylation status			
not methylated, n (%)	11 (85)	4 (44)	13 (57)
	2 (15)	5 (56)	10 (43)
methylated, n (%)			
EGFR amplification			
no, n (%)	8 (62)	5 (56)	10 (43)
	5 (38)	4 (44)	13 (57)
yes, n (%)			

Chromosome 7 gain combined with chromosome 10 loss (+7/-10)

no, n (%)	5 (38)	1 (11)	8 (35)
yes, n (%)	8 (62)	8 (89)	15 (65)

EGFR amplification combined with 7 gain / 10 loss

EGFR amplification or gain 7 / lost 10	8 (62)	8 (89)	20 (87)
EGFR amplification without gain 7 / lost 10	0 (0)	0 (0)	6 (26)
EGFR amplification and gain 7 / lost 10	5 (38)	4 (44)	7 (30)
gain 7 / lost 10 without EGFR amplification	3 (2)	4 (44)	7 (30)

Homozygous CDKN2A deletion

no, n (%)	6 (46)	3 (33)	8 (35)
yes, n (%)	6 (46)	6 (67)	15 (65)
other: CDKN2A amplification, n (%)	1 (8)	0 (0)	0 (0)

Median follow-up (months)

median (IQR)	28.5 (12.5-51)	21 (15-29)	19.5 (16-32)
--------------	----------------	------------	--------------

Initial treatment

RT/TMZ followed by 6 cycles of TMZ, n (%)	5 (38)	6 (67)	8 (35)
RT/TMZ followed by less than 6 cycles of TMZ, n (%)	1 (8)	0 (0)	3 (13)
RT/TMZ followed by less than 6 cycles of TMZ, n (%)	4 (31)	2 (22)	11 (48)
other treatment**, n (%)	2 (15)	0 (0)	0 (0)
clinical study, n (%)	0 (0)	0 (0)	1 (4)
no treatment, n (%)	1 (8)	1 (11)	0 (0)

Progression

yes, n (%)	12 (92)	7 (78)	17 (74)
no, n (%)	1 (8)	1 (11)	3 (13)
unknown, n (%)	0 (0)	1 (11)	3 (13)

Progression-free survival (months)

median (IQR)	10.5 (6.5-32.5)	12.5 (10-19)	12 (6-14.5)
--------------	-----------------	--------------	-------------

Treatment at first progression

yes, n (%)	11 (92)	7 (100)	14 (82)
no, n (%)	1 (8)	0 (0)	3 (18)

Death

yes, n (%)	11 (85)	7 (78)	18 (78)
	2 (15)	2 (22)	5 (22)

no, n (%)			
Survival from surgery (months)			
median (IQR)	24.5 (13.5-49.5)	23 (15-29)	20 (17-33)
Survival			
upper IQR, n (%)	4 (31)	2 (22)	4 (17)
intermediate IQR, n (%)	4 (31)	6 (67)	14 (61)
lower IQR, n (%)	5 (38)	1 (11)	5 (22)
Epilepsy			
no	6 (46)	2 (22)	9 (39)
before diagnosis or at diagnosis	6 (46)	7 (77)	14 (60)
during follow-up	1 (8)	0 (0)	0 (0)
Thromboembolic event			
no	13 (100)	8 (89)	21 (91)
before diagnosis	0 (0)	1 (11)	1 (4)
during follow-up	0 (0)	0 (0)	1 (4)

* according to WHO 2016 recommendation

Abbreviations:

EGFR, epidermal growth factor receptor

IDH: isocitrate dehydrogenase

IQR: interquartile range

MGMT: O⁶-methylguanine DNA methyltransferase

RT: radiotherapy

TMZ: temozolomide

Table 3 : Proteins associated with unfavorable prognostic in glioma and identified in groups A, B and C

Uniprot	Gene description	Gene	Group
Q8N111	Cell cycle exit and neuronal differentiation 1	CEND1	A
Q08495	Dematin actin binding protein	DMTN	A
Q13153	P21 (RAC1) activated kinase 1	PAK1	A
Q02750	Mitogen-activated protein kinase kinase 1	MAP2K1	A
P04216	Thy-1 cell surface antigen	THY1	A
P62760	Visinin like 1	VSNL1	A
P48163	Malic enzyme 1	ME1	A
Q8IU7	AE binding protein 1	AEBP1	B
P13667	Protein disulfide isomerase family A member 4	PDIA4	B
P16435	Cytochrome p450 oxidoreductase	POR	C

Table 4 : Proteins associated with favorable prognostic in glioma and identified in groups A, B and C

Uniprot	Gene description	Gene	Group
P00367	Glutamate dehydrogenase 1	GLUD1	A
Q14155	Rho guanine nucleotide exchange factor 7	ARHGEF7	A
P50395	GDP dissociation inhibitor 2	GDI2	A
P49591	Seryl-tRNA synthetase	SARS	A
O00505	Karyopherin subunit alpha 3	KPNA3	B
P11940	Poly(A) binding protein cytoplasmic 1	PABPC1	C

1 Uhlen, M. *et al.* A pathology atlas of the human cancer transcriptome. **357**, eaan2507 (2017).

Table 6. Proteins associated with survival after Cox model p = 0.05

Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Confidence Limits	
MAOB*	1	0.41898	0.13045	10.3155	0.0013	1.52	1.177	1.963
IGHM*	1	0.33461	0.10481	10.1918	0.0014	1.397	1.138	1.716
HSPD1*	1	-1.19189	0.39077	9.303	0.0023	0.304	0.141	0.653
CPSF7	1	-0.49726	0.18099	7.5487	0.006	0.608	0.427	0.867
PLXNB2	1	0.58661	0.21583	7.3875	0.0066	1.798	1.178	2.745
ANXA11	1	0.26235	0.09736	7.2604	0.007	1.3	1.074	1.573
LASP1*	1	-0.72992	0.27234	7.1836	0.0074	0.482	0.283	0.822
CFH*	1	0.20856	0.07907	6.957	0.0083	1.232	1.055	1.438
ANXA6*	1	0.38089	0.1452	6.8818	0.0087	1.464	1.101	1.945
CHCHD3	1	-1.01314	0.39913	6.4433	0.0111	0.363	0.166	0.794
C3	1	0.31568	0.12696	6.1828	0.0129	1.371	1.069	1.759
A2M	1	0.27035	0.10906	6.1444	0.0132	1.31	1.058	1.623
ITIH2	1	0.18318	0.07451	6.0444	0.014	1.201	1.038	1.39
SLC25A5	1	-0.39375	0.16102	5.9795	0.0145	0.675	0.492	0.925
CD44	1	0.30417	0.12591	5.8359	0.0157	1.356	1.059	1.735
IMMT	1	-0.88068	0.36659	5.7713	0.0163	0.415	0.202	0.85
SERPINF1	1	0.24845	0.1048	5.6204	0.0178	1.282	1.044	1.574
SERPINA1	1	0.25497	0.10929	5.4422	0.0197	1.29	1.042	1.599
CP	1	0.1988	0.08547	5.4095	0.02	1.22	1.032	1.442
HRG	1	0.20116	0.08657	5.3997	0.0201	1.223	1.032	1.449
AIFM1	1	-0.68714	0.29594	5.3912	0.0202	0.503	0.282	0.898
ATXN2L	1	-0.68322	0.30086	5.1569	0.0232	0.505	0.28	0.911
LDHA	1	0.43625	0.19324	5.0965	0.024	1.547	1.059	2.259
SOD2	1	0.26316	0.11711	5.0495	0.0246	1.301	1.034	1.637
SERPING1	1	0.22661	0.1038	4.7662	0.029	1.254	1.023	1.537
ITIH1	1	0.18455	0.08499	4.7153	0.0299	1.203	1.018	1.421
SERPINA3	1	0.21643	0.10036	4.6508	0.031	1.242	1.02	1.512

CALM3	1	0.14066	0.06903	4.1523	0.0416	1.151	1.005	1.318
EPS15	1	0.45937	0.22611	4.1277	0.0422	1.583	1.016	2.466
SERPIND1	1	0.19904	0.098	4.1251	0.0423	1.22	1.007	1.479

* proteins that remained significantly correlated to survival after step by step and bootstrap analyses.

Figure 1: Histological, MALDI MSI and SpiderMass data

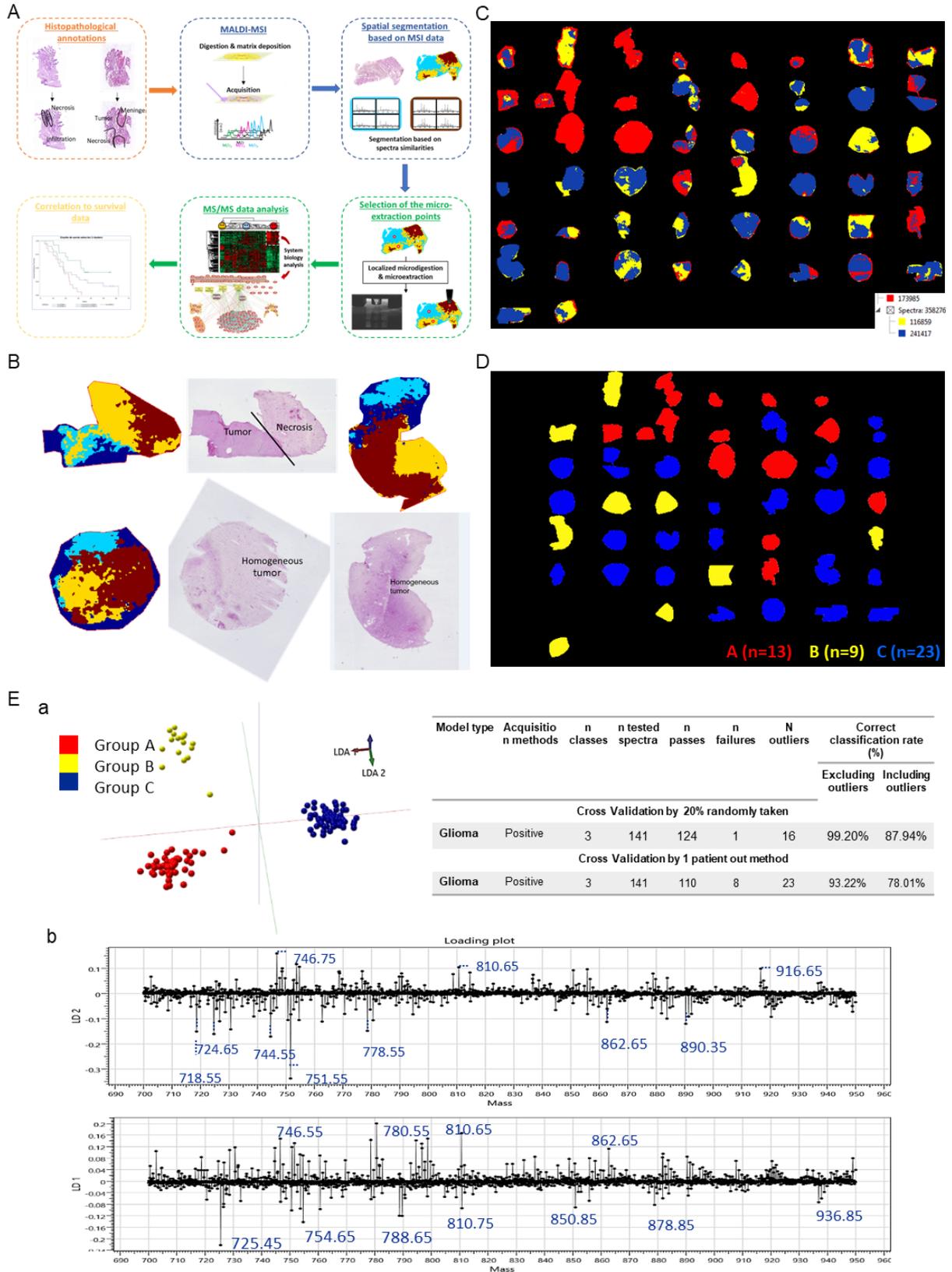


Figure 2. Shotgun microproteomics analysis

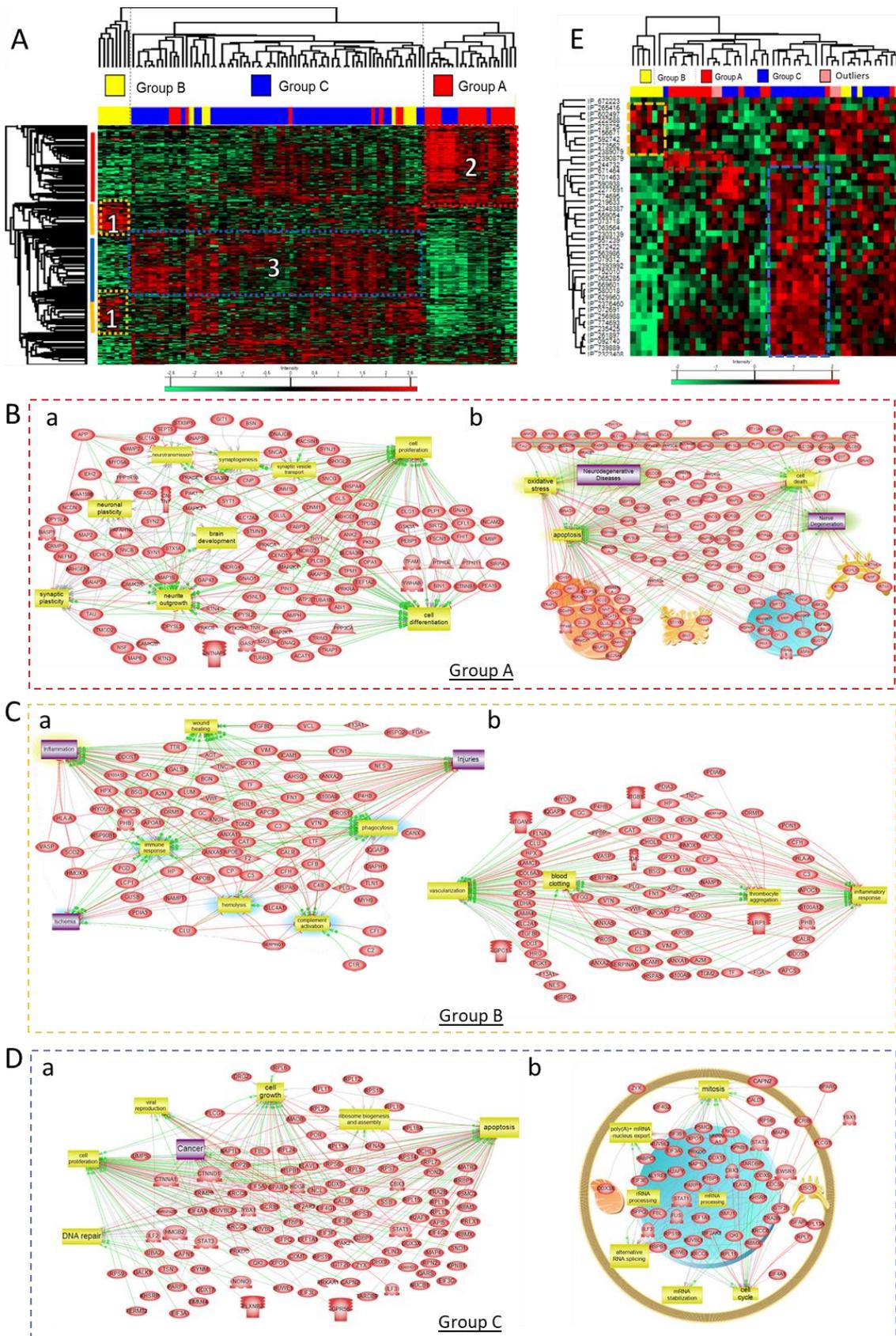
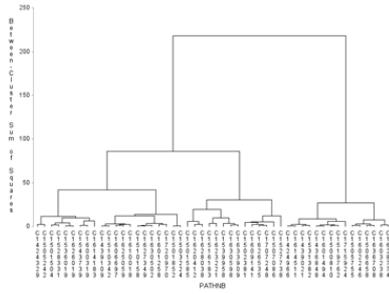


Figure 3. Proteomic and lifetime analysis

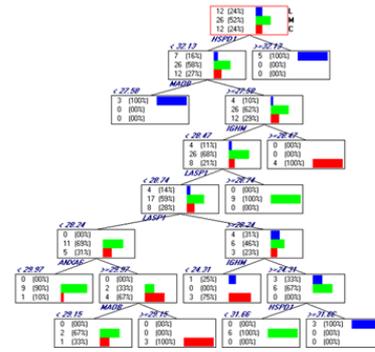
A

Analysis of Maximum Likelihood Estimates								
Parameter	DF	Parameter Estimate	Standard Error	Chi-Square	Pr > ChiSq	Hazard Ratio	95% Hazard Ratio Limits	
MAOB	1	0.626	0.181	12.002	0.001	1.870	1.312	2.665
IGHM	1	0.321	0.136	5.533	0.019	1.378	1.055	1.800
HSPD1	1	-1.204	0.568	4.493	0.034	0.300	0.099	0.913
LASP1	1	-0.582	0.282	4.257	0.039	0.559	0.321	0.971
CFH	1	-0.257	0.123	4.367	0.037	0.774	0.608	0.984
ANXA6	1	0.458	0.191	5.735	0.017	1.580	1.087	2.298

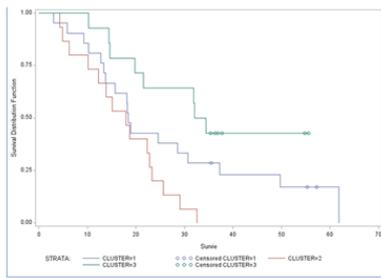
B



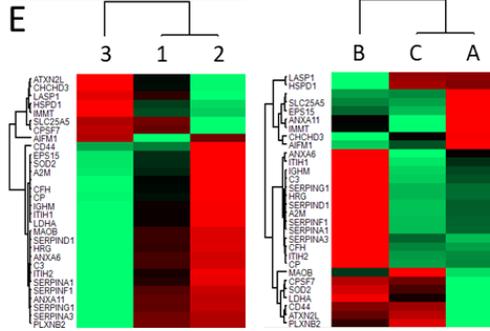
C



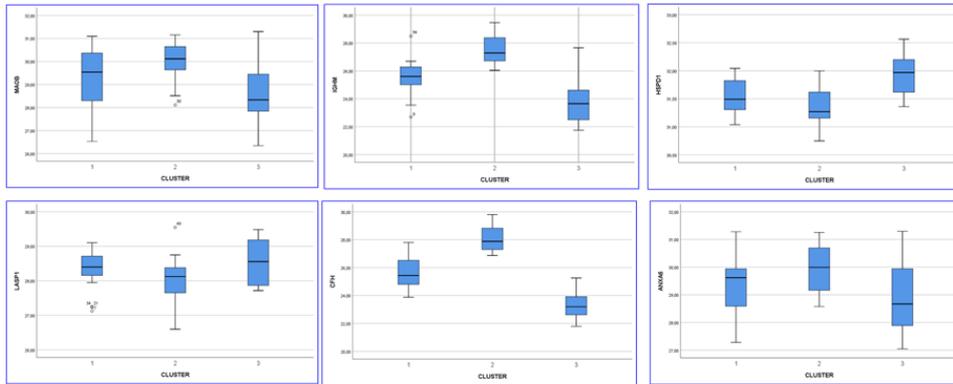
D



E



F



G

Group	Number of spectra	Number of passes	Number of failures	Number of outliers	Correct Classification Rate	
					Excluding outliers	Including outliers
Total	203	184	16	3	92.00%	90.64%

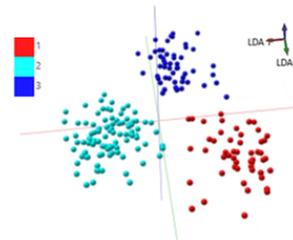
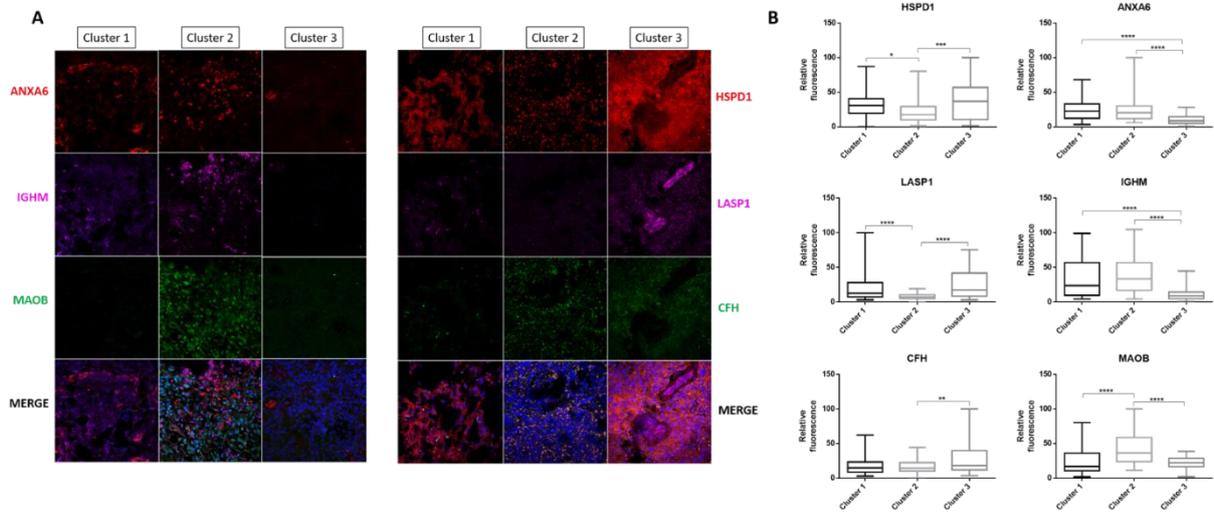


Figure 4. Immunohistochemistry of the 6 survival markers identified.



Manuscript #	NMED-A106851
Current Revision #	0
Submission Date	16th Jul 20
Current Stage	Manuscript received
Title	New glioma classification for precise therapeutic decision from spatially-resolved proteomics guided by MALDI-MSI and clinical data integration
Manuscript Type	Article
Manuscript Comment	<p>SUGGESTED REVIEWERS =====</p> <p>Pr. Massimiliano Bonafè, Department of Experimental, Diagnostic and Specialty Medicine, AlmaMater Studiorum, Università di Bologna, Bologna, Italy. email : massimiliano.bonafe@unibo.it</p> <p>Pr. R. N. Caprioli, Department of Medicine, Vanderbilt University, 465 21st Ave S #9160, Nashville, TN 37235, USA. Electronic address: richard.m.caprioli@Vanderbilt.Edu</p> <p>Pr. A. THoley, Institute for Experimental Medicine , Christian-Albrechts-Universität zu Kiel, 24105 Kiel, Germany, a.tholey@iem.uni-kiel.de</p> <p>Pr. David N Louis, epartment of Pathology, Massachusetts General Hospital, Harvard Medical School, WRN225, 55 Fruit Street, Boston, MA, 02114, USA. dlouis@mgh.harvard.edu</p> <p>Pr. Roel G W Verhaak, epartment of Genomic Medicine, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. Department of Bioinformatics and Computational Biology, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. rverhaak@mdanderson.org.</p>
Corresponding Author	Prof. Michel Salzet (michel.salzet@univ-lille.fr) (Univ.Lille, Inserm, CHU Lille, U1192,

	Laboratoire Protéomique, Réponse Inflammatoire et Spectrométrie de Masse (PRISM), F-59000 Lille)
Contributing Authors	Dr. Lauranne Drelich , Marie Duhamel , Dr. Maxence Wisztorski , Soulaïmane Aboulouard , Jean-pascal Gimeno , Pierre-Damien Caux , Dr. Nina Ogrinc , Mr. Patrick Devos , Tristan Cardon , Prof. Michael Weller , Fabienne Escande , Fahed Zairi , Claude-Alain Maurage , Prof. Isabelle Fournier , Emilie Le Rhun
Authorship	Yes
Abstract	Molecular heterogeneities are a key feature of glioblastoma pathology leading to high discrepancies between patients mean survivals. Based on multidimensional analysis on a cohort of 50 glioblastoma tumors, three proteomic subtypes have been uncovered. Mass spectrometry imaging resulted in the definition of these three groups. More than 95% of correct classification rates for these 3 groups was obtained using SpiderMass technology. In order to correlate this classification with functional biology, spatially resolved proteomics was performed and systemic biology established that these groups are related to neurogenesis, immune infiltration and tumorigenesis, respectively. Association with overall survival showed a signature of 6 proteins associated with prognosis. We validate these new prognostic markers of glioblastomas which allows to stratify patients and help the therapeutic decision. We also demonstrate a high disruption between proteomic and genomic classification and gave now real tools for pathologist to guide the oncologist decision.
Subject Terms	Biological sciences/Cancer/CNS cancer Biological sciences/Biochemistry/Proteomics Health sciences/Biomarkers Health sciences/Medical research/Translational research
Show Author Information	Allow Reviewers to see Author information.
Competing interests policy	There is NO Competing Interest.
Applicable Funding Source	No Applicable Funding

Manuscript Items

1. Author Cover Letter [PDF \(58KB\)](#)
2. Article File [PDF \(3490KB\)](#)
3. Supp. Data 4_cluster_GROUP_RougeAJauneBBleuC.xlsx [PDF \(81KB\)](#)
4. Supp. Material.docx [PDF \(651KB\)](#)
5. Supp. data 3_ unique proteins_identif.xlsx [PDF \(51KB\)](#)
6. Supp. data 5_ Subnetwork per group.xlsx [PDF \(42KB\)](#)
7. Supp. data 6 OS for Groups and MGMT.pdf [PDF \(467KB\)](#)
8. Supp. data 7 OS versus cluster.xlsx [PDF \(51KB\)](#)
9. Supp. data1_Segmentation_individual.pdf [PDF \(3624KB\)](#)
10. Supp_Data_2_Total_matrix.xlsx [PDF \(3301KB\)](#)
11. Supp_Figure_1.pdf [PDF \(609KB\)](#) 12. Supp_Figure_2.pdf [PDF \(474KB\)](#) 13. Supp_Figure_3.pdf [PDF \(412KB\)](#)
14. Supp_Figure_4.pdf [PDF \(115KB\)](#)

Manuscript Tasks

[Send Manuscript Correspondence](#)

[Check Status](#)

[tracking system home](#) | [privacy policy](#) | [cookie policy](#) | [manage cookies](#)

Q9Z7L0	Voltage-dependent anion-selective channel protein 1	Vdac1
Q9Z7Q3	Protein transport protein Sec31A	Sec31a
Q9Z7S9	Foslin-2	Fos2

CHAPITRE 2: ARTICLE SOUMIS

Spatially resolved proteomic based on expansion microscopy on FFPE and frozen tissues sections

Lauranne Drelich, Soulaymane Aboulouard, Julien Franck, Michel Salzet, Isabelle Fournier, Maxence Wisztorski

Abstract:

Expansion microscopy is an emerging approach for morphological examination of biological specimens at nanoscale resolution using conventional optical microscopy. To achieve physical separation of cell structures, tissues are embedded in a swellable polymer and isotropically expanded several folds.

We propose a new method for spatially resolved proteomics of magnified tissue obtain using expansion microscopy method. This method aims to enlarge the tissue sections to perform manual microdissection on regions of interest to then carry out proteomic analysis. Finally, the original size of the sample is calculated by dividing the area cut by the expansion factor.

For isotropic expansion, different homogenization agents are investigated both to maximize protein identification and minimize protein migration. Better results are obtained with SDS and finally, we are able to identify up to 655 proteins from a region of 460 μ m in diameter, which correspond to an average of 940 cells. This approach doesn't require expensive instrument and can be performed easily. Also, it represents a new interesting feature for imaging as it enables to map a large number of proteins with a spatial resolution of less than 350 μ m. This strategy is therefore highly relevant as a new strategy for proteomic study.

Introduction

Today, the study of tumor heterogeneity is an emergent issue in clinical oncology. The spatial context then becomes the principal key determinant of the cellular identity in the tissue. In fact, the precise organization or interaction and communication processes between cells, are involved in their physiology and their molecular identity ¹. Approaching the analysis of the tumor microenvironment (TME) or intra-tumor heterogeneity (ITH) to reveal this complexity in a qualitative and quantitative manner should facilitate better characterization and therapeutic assignments. There is a real need to precisely map cells/tissue in order to put information in a spatial context. This field is essentially approached by the way of genomics or transcriptomics

analysis¹⁻³. Performing proteomics analysis remains a challenge to perform this type of spatial approach due to the limited tools available. In fact, traditionally, organs or tissues analysis relies on an enzymatic or chemical homogenization to obtain cellular population that could be analyzed. These procedures would inevitably lead to a loss of information concerning the spatial localization making impossible the understanding of the architecture of the tissue.

Mass spectrometry imaging (MSI) is a powerful tool to visualize cellular heterogeneity on a tissue and could be used to determine a distinction between samples different^{4,5}. Information obtained served to guide spatially resolved analysis to obtain identification of molecules of highlighted regions. In this context, several micro-sampling approaches were developed including e.g. Laser Capture Microdissection (LCM)^{6,7} and Liquid MicroJunction (LMJ) micro-extraction⁸⁻¹⁰.

LCM is a cell enrichment technique allowing the isolation of specific cell types from heterogeneous biological samples for subsequent proteomics analysis^{11,12}. This procedure can be time-consuming depending on the number of cells to collect from individual tissue section. Recent studies show that this method cannot provide reproducible results without optimization in the handling and processing of the laser microdissected tissue samples¹³. More recently, LCM has been combined with the NanoPOTS method (Nanodroplet Processing in One pot for Trace Samples) allowing the identification of up to 180 proteins by analyzing a region of 50µm corresponding to 2-6 cells¹⁴.

Our team developed several years ago an LMJ strategy derived from LESA (Liquid Extraction Surface Analysis) to perform spatially resolved microproteomics on specific regions of interest (ROIs) of fresh frozen or formalin fixed paraffin-embedded (FFPE) tissue section^{8,15}. On this procedure, small volume droplets of enzyme were applying onto the ROIs using a chemical inkjet printer and generated tryptic peptides are extracted using LESA prior to proteomics analysis. It was possible to achieve the identification and quantification of up to 1500 proteins in a limited size region of 650µm in diameter⁸. These strategies have physical limitations such as droplet size and can currently be applied to regions of 200 µm in diameter. However, droplet-based liquid extraction for spatially-resolved proteomics has been demonstrated up to a diameter of 150 µm, which corresponds to approximately 300 cells¹⁶ and can also be applied for extraction of intact proteins^{9,10}.

More recently, hydrogel discs containing trypsin were developed with a size down to hundreds of micrometer^{17,18}. Deposited directly onto tissue surface, protein digestion is performed after passive proteins diffusion within the gel. The resulting peptides are then extracted from the hydrogel and analyzed by LC-MS/MS. Manual manipulation of the hydrogel

is one of the main limitations in this approach to precisely target a region of interest, but the identification of over 600 proteins in a 260 μm diameter hydrogel has been demonstrated¹⁸.

Another way is to physically modify the tissue section to attain smallest regions. It has been employed for example to increase the spatial resolution that could be obtain in MALDI-MSI by stretching the sample. In this method samples are prepared by adhering tissue section to glass beads array fixed to a stretchable membrane¹⁹. When the membrane is stretched, it separates the tissue section into thousands of cell-sized pieces for tissue. The stretched tissue is then prepared and analyzed for MALDI-MSI experiments.

To assess the spatial variability, we present in this paper, the development of a new way to perform spatially resolved proteomics. The originality of our approach will be to expand a piece of tissue or a tissue section in 3 dimensions to easily select regions of interest to analyze. This will be performed without any sophisticated or cost expansive instrument. We adapt the protocol developed by Boyden et al. of expansion microscopy²⁰⁻²² in which a tissue section is magnified by 3- to 100-fold. This technique was developed to observe sample at a nanoscale precision with conventional microscopes. To obtain this expansion, a dense and uniform mesh of swellable polymer is introduced in the tissue. Proteins are anchored to the polymer network and the resulting hydrogel-tissue hybrid will expand after immersion in water, resulting in physical magnification of the tissue. From this point, we demonstrate that is possible to cut off a specific expanded region and treat the sample to perform proteomics analysis after some adjustments. Here, we report a protocol to achieve a 3-fold volumetric expansion and obtain the identification of over 650 proteins for a region of 460 μm original diameter. By doing so, we have overcome some limitation of the existing spatially resolved proteomics methods and we will be able to focus our analysis on small regions of interest on tissue section and with a limited quantity of material/cells.

Material & Methods

The complete protocols are detailed in supporting information

Reagents and Chemicals

For the different experiments, High purity chemicals were purchased from various supplier (complete list in the supporting information) and used without further modification.

Tissue expansion

Tissue preparation

Adult Wistar male rats (University of Lille) were sacrificed according to animal use protocols approved by the university of Lille Animal Ethics committee. FFPE and Fresh frozen tissue sections of 12µm thickness were used.

Tissue gelation

Tissue expansion was realized in accordance with the protocol published by Tillberg ²⁰. Complete workflow is represented in **Figure 1A**. Briefly, acryloyl-X (AcX) is deposited on tissue and incubated overnight in a humid chamber at room temperature. Gelling solution is freshly prepared by adding 8µL of 4-Hydroxy TEMPO (0.625%, acrylamid polymerization inhibitor), 8µL d'APS, and 8µL of TEMED to 376µL of monomere solution (sodium acrylate 85,5mg/mL, 30% acrylamid/bis-acrylamid 30% ; 37.5 : 1, sodium chloride 1,02g/mL and 10% PBS 1X). This solution is deposited on the tissue and spread inside a gelation chamber with a glass cover and placed in a humid chamber a 37°C for 2hours.

Mechanical homogenization

Original protocol uses proteinase K (8 units/mL) diluted in a homogenization buffer (50 mM Tris/HCl (pH8), 1 mM EDTA, 0,5% Triton X-100, 1 M NaCl). Gel containing tissue is submerged in 2mL of proteinase K for 3 hours at 60°C. Several homogenization processes are tested to replace proteinase K. Different enzymes are used and to minimize the total amount of enzyme, the homogenization solution is deposited directly on the tissue section instead of a complete immersion. Our first experiment was done using a proteinase K solution diluted to 4 units/mL. In a second time, we tested several other enzymes or homogenization agent.. LysC (20µg/mL in 6M Urea, 1mM EDTA) is deposited with a sufficient volume to cover the surface of the tissue section (typically 500 µl) and placed at 37°C overnight. As a second enzyme, 1mL of Trypsin-EDTA solution used for cell culture dissociation (0.025% trypsin and 0.01% EDTA in Phosphate Buffered Saline (PBS)) is deposited and placed at 37°C for 20 minutes. Finally, we also tested SDS as an homogenization agent, realized in first instance by adding 5% SDS diluted in water and placed 30 minutes at 95°C and in a second test by decreasing the temperature at 58°C in a humidity chamber overnight.

Expansion and cutting

Tissue sections embedded in the gel were separated from the glass slide by submersion in HPLC water for one hour. The water is replaced 3 times every 15 minutes. The expansion factor is typically about 3-fold. Once maximal expansion size has been reached, square areas of 1x1mm² and 5x5mm² are cut manually with a scalpel blade and transferred individually into a microcentrifuge tube (LoBind 1.5mL microcentrifuge tubes, Eppendorf). A control region,

located outside the tissue, is also cut to evaluate possible diffusion of proteins and peptides in the gel. The expansion factor, determined from the pre-expansion measurement (**Figure 1Ba**) divided by the post-expansion measurement (**Figure 1Bc**), is usually between 2.9- and 3.4-fold. Thus, an area of 5x5mm² after expansion corresponds to 1.6x1.6mm² as a real size and around 300x300µm² for an expanded area of 1x1mm². Thereafter, results will be expressed considering the region size post-expansion: 5x5mm² and 1x1mm². We also used some biopsy punch with a diameter of 1.5mm for sampling regions after tissue expansion. For imaging-like analysis, we sampled 15 consecutive regions of 1x1mm² each along a line through the tissue in order to assimilate each piece of gel to an image pixel.

Proteomics analysis

Pieces of gel in the microcentrifuge tube are covered with 50µL of NH₄HCO₃ (50mM). reduction process is achieved by incubation of the pieces in 10µL of DTT (45mM in NH₄HCO₃ 50mM) 15min at 50°C. For alkylation process, gel pieces are incubated with 10µL of IAA (100mM in NH₄HCO₃ 50mM) 15min at room temperature in obscurity. Finally, 20µL of trypsin (20µg/mL in NH₄HCO₃ 50mM) is added in each sample and incubated overnight at 37°C. Digestion is stopped by adding TFA 1% final volume.

Liquid Microjunction experiment

To compare to another spatially resolved proteomic method, experiment of droplet-based liquid microextraction were performed according the protocol previously published^{15,23}. Briefly, region of interest was first digested using trypsin deposited by a piezoelectric microspotter Chemical Inkjet Printer (CHIP-1000, Shimadzu, CO, Kyoto, Japan) and peptides were extracted using the TriVersa Nanomate platform (Advion Biosciences Inc., Ithaca, NY, USA) with a Liquid Extraction Surface Analysis (LESA) option.

NanoLC-MS & MS/MS analysis

Desalting samples were analyzed by nanoLC-MS/MS (Q Exactive mass spectrometer, Thermo Scientific) (see Support Information for complete description). Proteins identification and label free quantification were performed with MaxQuant^{24,25} (Version 1.6.1). The search was done against a database containing reviewed proteome for *Rattus norvegicus* from UniprotKB/Swiss-Prot (8,168 sequences, July 2019). False discovery rates lower than 1% were set at peptides and proteins level. Relative label-free quantification of proteins was conducted into Max-Quant using the MaxLFQ algorithm²⁷ with default parameters. The match between run (MBR) feature, with a match window of 0.7 min and an alignment window of 20 min, was activated to increase peptide/protein identification of small samples.

Analysis of identified proteins was performed using Perseus software (<http://www.perseus-framework.org/>) (version 1.6.0.7). The file containing the information from the identification were used and hits from reverse database, proteins with only modified peptides and potential contaminants were removed. The different methods were evaluated in terms of overlap of protein identifications (Venn diagrams) and also expressed as Pearson correlation coefficients (dot plots and r value). Visualization in Venn Diagram was performed with BioVenn (<http://www.biovenn.nl/>)³⁰. For quantification-based mass spectrometry imaging, LFQ values of proteins were used to construct images with TIGR Multiexperiment viewer (MEV v 4.9).

The datasets used for analysis were deposited at the ProteomeXchange Consortium²⁸ (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository²⁹ with the dataset identifier PXD00XXX (for review: Username: XXX Password: XXX).

Results and discussion

Proteinase K homogenization does not allow proteins identification

On this article, we propose to develop a new methodology to perform Spatially Resolved Proteomics. To achieve this goal, we want to magnify a tissue section using a swellable polymer allowing an isotropic expansion. In first instance, we based our protocol of tissue expansion according to the publication of Tillberg²⁰ concerning the protein-retention expansion microscopy (proEXM). The expansion process is divided into several phases (**Figure 1A**). Briefly, FFPE tissue sections were deparaffinized and then cover by a solution of AcX. AcX is a linker that bind the primary amine groups on proteins and is then incorporated into the swellable hydrogel polymer during the process of gelation. The link created by AcX on the proteins will be tethered to the hydrogel polymer chains allowing the biomolecules to retain their spatial organization relative to one other. Thereafter, mechanical homogenization step consists to suppress mechanical properties in the sample in order to keep structural integrity and sample organization during expansion. Disruption of the embedded sample is generally realized by using an enzymatic digestion with the proteinase K, and the gel is finally submerged in water until maximal expansion (**Figure 1B**). In this configuration, we obtained an expansion factor around 2.3-fold from the original size.

From the expanded FFPE tissue section, we proceeded with a scalpel blade to a manual cutting of small squares of 5x5 mm² size on the embedded tissue and outside the tissue (hydrogel alone) as a control to assess possible molecule migration within the gel. The gel

pieces were then submitted to a conventional proteomics digestion protocol and retrieved peptides were analyzed by mass spectrometry. On the tissue, 14 proteins were identified and 9 proteins in the control region (**Supplementary Figure 1, Supplementary Data 1**). The number of proteins identified is low compared to what is expected for this tissue surface and an equivalent number of proteins is identified outside the tissue section. The detection of peptides outside the tissue could indicate a loss of some peptides produced during the homogenization step using the proteinase K. In fact, as FFPE tissue was used, proteins are normally crosslinked by the paraformaldehyde, preventing any diffusion within the gel. The AcX will also create some link between proteins and the hydrogel. Indeed, the proteinase K preferentially cleaves at aliphatic or aromatic amino acid residues with low specificity and leads to the formation of small peptides non-anchored to the gel which can diffuse through the reticulation of the hydrogel during the different processing steps of expansion or proteomics analysis. This diffusion could lead to a loss of proteins localization and significantly reduce protein identification.

Nevertheless, some proteins could be identified from expanded tissue. These results incite us to investigate the replacement of the proteinase k during the mechanical homogenization step to reduce possible loss of proteins.

Comparative analysis of different alternatives for Proteinase K

As other homogenization strategies, we tested two enzymes with specific cleavage sites, *i.e.* the trypsin and the LysC, and an anionic detergent, the SDS, which allows to disrupt protein bonds inducing protein linearization. Using the trypsin or LysC, the expansion factor was estimated to 2.1-fold; 2.3-fold with SDS and 2.4-fold with proteinase K (**Figure 2A, B, C and D**). The correlation between the size of the sampled regions and their real size according to the expansion factor is presented in the **Table 1** with an estimation of the corresponding number of cells. For trypsin and LysC, an incomplete homogenization is observed with a lot of cracks and tissue distortion appearing during the expansion process (**Figure 2B and C**). Using SDS, the expansion is more homogeneous (**Figure 2D**) and the expansion factor is quite similar than the one obtained with the proteinase K.

For proteomics analysis, same results as above are obtained using Proteinase K with a low number of proteins identified (**Table 1**) For the other conditions, the number of identified proteins increased significantly (**Table 1 and Supplementary Data 2**). In the 5x5mm² regions, up to 536 proteins were quantified using trypsin as homogenizing agent, 488 for SDS and 495 for Lys C (**Table 1**). For 1x1mm² regions, 35 proteins were quantified for trypsin and 32 proteins for SDS. Identification increase up to 65 protein with Lys-C homogenization. In control, only 2

proteins were quantified for LysC, 1 for trypsin and none for SDS, revealing no or only very minimal migration within the gel. Identified protein are presented in the **Table 1**.

However, as homogenization with trypsin and LysC was not homogeneous, it is difficult to precisely estimate the real size of the digested region (**Figure 2B and C**). Based on the measured expansion factor, there is a variation of about 20% in the number of cells analyzed between the different tests (**Table 1**). This number of cells is also difficult to estimate when the expansion is not isotropic as for trypsin and LysC. In these cases, a high number of cracks and apparent deformations were observed, which means that some groups of cells retain their original size, resulting in a higher than expected cell density in the sampled region (**Figure 2B and C; Table 1**). This variability was also observed by comparing the proteins identified in each condition for the region of 1x1mm² (**Figure 2E**). However, for the region of 5x5 mm², a high protein identification overlap (approximately 73%) was observed between Trypsin, Lys C and SDS (**Figure 2 F**).

Considering these results, we decided to focus our development on SDS to replace proteinase K in the homogenization process. As a homogenized agent, SDS allows lipid removal, keep protein integrity and affect only their conformation and charge without disturbing their anchoring to the hydrogel^{20,31,32}. Moreover, SDS is cheap and easy to use, contrary to enzyme which need specific temperature to conserve their activity.

Method optimization and reproducibility for proteomics analysis using the SDS homogenization protocol

As a first step, we increased the homogenization time up to one night in a humid chamber with a temperature maintained to 58°C instead of 95°C to avoid liquid evaporation and hydrogel degradation. These improvements result in a greater expansion factor up to 3 times the initial tissue size.

An antigen retrieval step was also added to the protocol after dewaxing in order to improve protein digestion efficiency and therefore protein identification. No clear increase was observed, and the tissue was rather more brittle to handle. A protocol of “in-gel digestion”³³ of the gel pieces was evaluated to improve peptides extraction from the acrylamide network, but no significant difference was observed. As the digestion in solution is easier to perform, we decide to continue with this method.

Reproducibility was assessed by comparing triplicate sampled regions on consecutive FFPE tissue sections (S1, S2 and S3) of the midbrain (**Figure 3A**). Only 1 protein was identified in control area from S3, no protein was identified in S1 and S2 control areas which means that

proteins keep their location within the tissue despite the longer preparation time (**Supplementary Data 3**). This is in accordance with our expectation, proteins are anchored to the acrylamide network and consequently do not migrate outside the tissue. Overlapping protein identification (Venn diagrams) was used to evaluate qualitative reproducibility and Pearson's correlation coefficients (dot plots and r value) for quantitative reproducibility. The number of proteins identified in all replicates, regardless of their abundance, is presented in the Venn diagrams, while the quantitative values (LFQ values) were used to calculate the scatter plots.

For 5x5mm² regions, 1259 different proteins are identified with 898, 945 and 994 proteins respectively for replicate 1; 2 and 3 (**Table 2; Supplementary Data 3**). The expansion factor and the estimated number of cells were presented in the **Table 2**. Considering the expansion factor, the real analyzed area is about 1,6x1,6mm². If we assume that cells present an average size of 15 µm and are round-shaped, we can estimate at less than 15 000 the number of cells. The optimization allows doubling the number of protein identifications compared to previous experiments. In **Figure 3B**, the Venn diagram shows that 675 proteins are shared by all replicates corresponding to 53.6% of common proteins with a small individual variation (about 5%). Looking at the direct side-by-side comparison (**Supplementary Figure 2 A**), replicates S2 and S3 are closed with a high overlap in protein identification (approximately 80%). The protein content of S1 is slightly different from that of the two other replicates (around 60% of overlapping). But for the quantification, the Pearson's correlation coefficient values show that the identified proteins gave the same quantification value between the replicates (mean of 0.956) (**Figure 3C**).

Reduction of the sampling region at 1x1mm² after expansion correspond to a real analyzed region of 330x330µm². This surface corresponds to approximately 600 cells and from this region, 335 groups of proteins were identified in section 1; 392 in section 2 and 454 in section 3 for a total of 542 different proteins (**Table 2; Supplementary Data 3**). A higher variability than for the 5x5mm² region, was observed with 260 protein shared by all replicates representing 47.97% of common protein (**Figure 3D**) (around 60% for pairwise comparison (**Supplementary Figure 2B**)). Same observations could be done for the quantitative reproducibility (variation of Pearson's correlation coefficient from 0.83 to 0.89 (**Figure 3E**)). We observed an average of 50% of the proteins identified by the match between run feature with 5x5mm² samples. It should be noted that more than 98% of the proteins are also identified in the 5x5mm² regions, suggesting that using a reference sample containing a larger region or more cells for the MBR, more proteins could be identified and quantified from a smaller expanded region.

Good reproducibility both qualitatively and quantitatively is observed even though variations exist. The variation could be explained by the evolution of the histological features on the consecutive tissue sections and by the difficulties of precisely sampling the same region due to the transparency of the expanded tissue. Reproducibility of smallest regions is lower than for bigger piece of tissue, probably due to manual cutting and expansion factor variation. Sampling regions with a scalpel blade may result in differences in gel size. In this sense, an instrument capable of cutting a specific region with reproducible size was subsequently used.

Potentially more proteins could be identified. Formalin fixation is advantageous for preservation and conservation of cellular and architectural morphologic detail in tissue sections but results in formaldehyde cross-links and many proteins are still not detectable using these approaches³⁴. For this purpose, we applied also the strategy developed to fresh frozen tissue section with encouraging results but a poor reproducibility and great variation. Indeed, we identify a total of 1 038 proteins in 5x5mm² samples in all replicates with respectively 136 proteins in section 1; 78 in section 2 and 788 in section 3 (**Table 3**). As the tissue was not fixed, the different steps, in particular the mechanical homogenization, lead to significant migration of proteins within the gel resulting in a loss of the actual localization (*i.e.* 10 proteins, 26 and 0 respectively in the control region of sections 1; 2 and 3 - **Table 3**). Post fixation with paraformaldehyde could be interesting to stabilize protein, also a shorter incubation time with SDS could be applied in order to limit protein migration. Another issue with fresh frozen tissue is the clarification, it is impossible to distinguish the tissue inside the gel and consequently to perform a precise and reproducible extraction. To overcome this, it is nevertheless possible to use, for both FFPE or Frozen tissues, dye like Fast green which allows to see tissue border or fluorescent dye like Hoechst which specifically stain nucleus **Supplementary Figure 3**. Preliminary results show that the presence of the dyes does not altered identification proteins (Data not shown)

Comparison with on tissue microdigestion and microjunction using Liquid extraction surface analysis

We then performed a comparison with another direct surface sampling method using in situ digestion and extraction by liquid microextraction. This micro extraction strategy was used in particular for tissue microenvironment characterization^{5,8,10,15,35}. The results obtained after tissue expansion were compared to one obtained with the best setup using this method.

On-tissue enzymatic digestion combined with liquid microextraction performed on a fresh frozen tissue section allows identification of 1157 proteins in a region of 900µm in diameter. Compared to the total of proteins identified in the replica of the same region from an expanded

FFPE tissue section, we observed 688 common proteins (**Figure 3F**). Equivalent number of proteins are identified in the two techniques but interestingly 469 proteins are unique to microextraction extract whereas 215 are unique to the replicate of tissue expansion proteomics (**Supplementary Data 4**).

The quantitative reproducibility shows an average Pearson's correlation of 0.678 (**Supplementary Figure 4A**). Although size of region and digestion methods are different it appears that the identified proteins gave equivalent quantification values. The observed variation is undoubtedly due to FFPE fixation of the tissue used for expansion, whilst microextraction was performed on fresh frozen tissue which gave better accessibility to proteins than on fixed tissue.

Mass ranges are similar between the two techniques and are comprised between 5 kDa and 620 kDa (**Supplementary Figure 4B**). High molecular weight proteins like Serine/threonine-protein kinase mTOR (288.79kDa) and Ankyrin-3 (284.44kDa) were identified in both methods and Huntingtin (343.76kDa) and Fibronectin (272.51kDa) were identified respectively with in situ digestion and extraction by liquid microextraction and tissue expansion. Smaller proteins were also identified, Myotrophin (12.861kDa) and Thioredoxin (11.677kDa) respectively with in situ digestion and extraction by liquid microextraction and tissue expansion and Complexin-2 (15.394kDa) for both methods (**Supplementary Data 4**).

Use of a biopsy punch for small regions sampling

To improve reproducibility, a biopsy punch was used to cut and remove a tissue disc. This allowed to precisely cut a 1400 μm diameter gel disc corresponding to ~ 460 μm diameter region of tissue before expansion (around 940 cells) (**Table 2**). For three consecutive regions, we obtained punch diameter of $1400 \pm 37 \mu\text{m}$ (**Figure 4A**). A total of 808 distinct groups of protein were obtained in the replicates (582, 682 and 700 in replicate S1, S2 and S3 respectively) (**Figure 4B; Supplementary Data 3**). Venn diagram shows a good reproducibility between the three replicates with 62.13% of common protein. The quantitative reproducibility is high with an average Pearson's correlation of 0.949 (**Figure 4C**). For individual variability between the replicates, the same observations as above are made with greater differences between replicate S1 and S3 (**Supplementary Figure 2C**). Using Punch Biopsy, Pearson's correlation coefficient increases up to 0.97, showing reproducible cut size allows to decrease proteome variation.

These results highlight that variation observed in the proteomics content is mainly due to the manual cut of the region of interest. Indeed, it is difficult to sample gel pieces of the same size

manually. Reproducibility is significantly improved with the use of a punch biopsy. This circular blade allows to cut region of same size easily and with precision. Consequently, we observed a reduction of individual variation at 6.3% and considerable increase of common protein at 62.13%. The quantitative reproducibility significantly increases with an average Pearson correlation factor of 0.95. Therefore, for a sample corresponding to a region of 460 μ m in diameter, we were able to identify an average of 655 proteins from approximately 940 cells. A variation is still observed certainly due to minor variation in the position of the sampling due to the transparency of the tissue section. Evolution of the histological features between the tissue sections or variation of the expansion factor could also be a source of variations.

Decreasing the punch diameter up to 300 μ m could significantly decrease the spatial resolution at 100 μ m. Recently, enzyme delivering hydrogels were used as a tool for localized analyte extraction directly on the tissue sample^{17,18}. They were able to identify around 700 proteins with a punch diameter of 357 μ m¹⁸.

Quantification-based mass spectrometry profiling using tissue expansion

Next step consisted of analyzing consecutive adjacent points to perform a quantification-based profiling, as we have done in other publications^{9,36,37}. To achieve this, we sampled 15 consecutive regions of 1x1mm² each along a line through a rat brain tissue section in order to assimilate each piece of gel to an image pixel (**Figure 5**). It is then possible to use MS-based quantification data for an image reconstruction³⁶.

All extracts were digested and analyzed by LC-MS/MS. A total of 511 proteins were identified (**Supp Data 5**). Images were constructed and each square represents a piece of gel with its position on the tissue. The differences in expressions are color coded, red represents protein overexpression whereas dark blue represent a low detection (**Figure 5**). Despite a spatial resolution of 300 μ m, expressions of proteins are different depending on the localization in the tissue. Expression of proteins specific in each region of the rat brain was then examined. The validation of this spatial mapping is confirmed by the overexpression of Creatine kinase and Calcium/calmodulin-dependent protein kinase in cortex and fimbria of the hippocampus which are known to contain these proteins^{38,39}. In these regions we also observed overexpression of Alpha-internexin relative to thalamus (**Figure 5, Supp Data 5**). This protein is an intermediate filament involved in neuron morphogenesis and neurite outgrowth⁴⁰. Cofilin-1 appears to be preferentially expressed in hippocampus region. This specific protein, as a regulator of actin dynamics, may contribute to degenerative processes⁴¹. Preferential distribution of GFAP and Pro SAAS is observed in thalamus, which is consistent with previous publications^{36,38} (**Supplementary Data 5**). In particular, MBP quantification is in accordance with

immunohistology data (**Figure 5**)⁹. Housekeeping genes such as Tubulin beta-4 chain and 14-3-3 protein epsilon were also identified. These proteins are known to have low variation in expression in tissues, here we observed a coefficient of variation of 2% and 1% respectively for these 2 proteins (**Figure 5**).

This strategy provides indirect molecular imaging based on the identification and subsequent quantification of proteins with a spatial resolution of less than 350 μm . It represents a new interesting feature for MSI as it enables to map a large number of high and low abundance proteins. Recently, methods have been developed to increase the number of proteins identification from small number of cells. For example, the nanoPOTS⁴² method improve the identification of thousands of proteins from a reduced number of cells. The combination of the tissue expansion with this type of methods would certainly allow an increase of the number of proteins identified from small regions.

Conclusion

In conclusion, we demonstrated for the first time that expansion microscopy is plainly compatible with conventional large-scale mass spectrometry-based proteomics. Substituting the proteinase K used in the mechanical homogenization step by SDS reduces protein loss due to the generation of small peptides that diffuse through the hydrogel. Using our protocol, a physical expansion of FFPE tissue section by more than 3-fold was achieved, makes it possible to easily and manually withdraw a well-controlled area with an original size down to 330 μm on side and analyze its proteins content. We successfully used this protocol to perform spatially resolved proteomics analysis of regions of interest and to identify more than 655 proteins for an equivalent of less than 940 cells. Identification of proteins from expanded tissue showed a good reproducibility both qualitatively and quantitatively. This strategy is particularly useful for mapping the protein content of closely related regions on a tissue section which has been challenging in previous spatially resolved proteomics studies. We performed a quantification-based mass spectrometry imaging of more than 500 proteins with a lateral resolution close to 300 μm . This resolution can be further reduced by investigating some new expansion microscopy protocols like the x10 Expansion Microscopy⁴⁸, ZOOM (Zoom by hydrOgel cOnversion Microscopy)⁴⁹ or iterative expansion microscopy (iExM)⁵⁰ to expand biological sample up to 20- to 100-fold. Interestingly as our protocol is compatible with conventional expansion microscopy it will be possible to target a protein using fluorescent antibodies²⁰ and retrieve all the proteomic related to this protein. Tissue expansion can also be combined with ExFISH which involve the use of linker that enables RNA to be covalently

attached to the gel^{43,44}. Then, fluorescent in situ hybridization (FISH) imaging of mRNA can be done⁴³ and the region containing fluorescence signal can be delimited, cut and analyzed to observe the presence of the translated protein by mass spectrometry. Thus, our strategy could be used as a simple single cell tool for proteomic analysis in order to reveal single-cell variations and understand specific disease mechanisms.

Bibliography

- (1) Moor, A. E.; Itzkovitz, S. Spatial Transcriptomics: Paving the Way for Tissue-Level Systems Biology. *Current Opinion in Biotechnology*. Elsevier Ltd August 1, 2017, pp 126–133. <https://doi.org/10.1016/j.copbio.2017.02.004>.
- (2) Burrell, R. A.; McGranahan, N.; Bartek, J.; Swanton, C. The Causes and Consequences of Genetic Heterogeneity in Cancer Evolution. *Nature* **2013**, *501* (7467), 338–345. <https://doi.org/10.1038/nature12625>.
- (3) Graham, T. A.; Sottoriva, A. Measuring Cancer Evolution from the Genome. *J. Pathol.* **2017**, *241* (2), 183–191. <https://doi.org/10.1002/path.4821>.
- (4) Quanico, J.; Franck, J.; Wisztorski, M.; Salzet, M.; Fournier, I. Progress and Potential of Imaging Mass Spectrometry Applied to Biomarker Discovery. In *Methods in Molecular Biology*; Methods Mol Biol, 2017; Vol. 1598, pp 21–43. https://doi.org/10.1007/978-1-4939-6952-4_2.
- (5) Le Rhun, E.; Duhamel, M.; Wisztorski, M.; Gimeno, J.-P.; Zairi, F.; Escande, F.; Reyns, N.; Kobeissy, F.; Maurage, C.-A.; Salzet, M.; Fournier, I. Evaluation of Non-Supervised MALDI Mass Spectrometry Imaging Combined with Microproteomics for Glioma Grade III Classification. *Biochim. Biophys. Acta - Proteins Proteomics* **2016**. <https://doi.org/10.1016/j.bbapap.2016.11.012>.
- (6) De Marchi, T.; Braakman, R. B. H.; Stingl, C.; van Duijn, M. M.; Smid, M.; Foekens, J. A.; Luider, T. M.; Martens, J. W. M.; Umar, A. The Advantage of Laser-Capture Microdissection over Whole Tissue Analysis in Proteomic Profiling Studies. *Proteomics* **2016**. <https://doi.org/10.1002/pmic.201600004>.
- (7) Longuespée, R.; Baiwir, D.; Mazzucchelli, G.; Smargiasso, N.; De Pauw, E. Laser Microdissection-Based Microproteomics of Formalin-Fixed and Paraffin-Embedded (FFPE) Tissues. In *Methods in Molecular Biology*; 2018; Vol. 1723, pp 19–31. https://doi.org/10.1007/978-1-4939-7558-7_2.

- (8) Quanico, J.; Franck, J.; Dauly, C.; Strupat, K.; Dupuy, J.; Day, R.; Salzet, M.; Fournier, I.; Wisztorski, M. Development of Liquid Microjunction Extraction Strategy for Improving Protein Identification from Tissue Sections. *J. Proteomics* **2013**, *79*, 200–218. <https://doi.org/10.1016/j.jprot.2012.11.025>.
- (9) Wisztorski, M.; Desmons, A.; Quanico, J.; Fatou, B.; Gimeno, J.-P.; Franck, J.; Salzet, M.; Fournier, I. Spatially-Resolved Protein Surface Microsampling from Tissue Sections Using Liquid Extraction Surface Analysis. *Proteomics* **2016**, *16* (11–12), 1622–1632. <https://doi.org/10.1002/pmic.201500508>.
- (10) Delcourt, V.; Franck, J.; Leblanc, E.; Narducci, F.; Robin, Y. M.; Gimeno, J. P.; Quanico, J.; Wisztorski, M.; Kobeissy, F.; Jacques, J. F.; Roucou, X.; Salzet, M.; Fournier, I. Combined Mass Spectrometry Imaging and Top-down Microproteomics Reveals Evidence of a Hidden Proteome in Ovarian Cancer. *EBioMedicine* **2017**, *21*, 55–64. <https://doi.org/10.1016/j.ebiom.2017.06.001>.
- (11) Banks, R. E.; Dunn, M. J.; Forbes, M. A.; Stanley, A.; Pappin, D.; Naven, T.; Gough, M.; Harnden, P.; Selby, P. J. The Potential Use of Laser Capture Microdissection to Selectively Obtain Distinct Populations of Cells for Proteomic Analysis — Preliminary Findings. **1999**, *20*, 689–700. [https://doi.org/10.1002/\(SICI\)1522-2683\(19990101\)20:4/5<689::AID-ELPS689>3.0.CO;2-J](https://doi.org/10.1002/(SICI)1522-2683(19990101)20:4/5<689::AID-ELPS689>3.0.CO;2-J).
- (12) Liu, N. Q.; Braakman, R. B. H.; Stingl, C.; Luider, T. M.; Martens, J. W. M.; Foekens, J. A.; Umar, A. Proteomics Pipeline for Biomarker Discovery of Laser Capture Microdissected Breast Cancer Tissue. *J. Mammary Gland Biol. Neoplasia* **2012**, *17* (2), 155–164. <https://doi.org/10.1007/s10911-012-9252-6>.
- (13) Longuespée, R.; Alberts, D.; Pottier, C.; Smargiasso, N.; Mazzucchelli, G.; Baiwir, D.; Kriegsmann, M.; Herfs, M.; Kriegsmann, J.; Delvenne, P.; De Pauw, E. A Laser Microdissection-Based Workflow for FFPE Tissue Microproteomics: Important Considerations for Small Sample Processing. *Methods* **2016**, *104*, 154–162. <https://doi.org/10.1016/j.ymeth.2015.12.008>.
- (14) Zhu, Y.; Dou, M.; Piehowski, P. D.; Liang, Y.; Wang, F.; Chu, R. K.; Chrisler, W. B.; Smith, J. N.; Schwarz, K. C.; Shen, Y.; Shukla, A. K.; Moore, R. J.; Smith, R. D.; Qian, W.-J.; Kelly, R. T. Spatially Resolved Proteome Mapping of Laser Capture Microdissected Tissue with Automated Sample Transfer to Nanodroplets. *Mol. Cell. Proteomics* **2018**, *17* (9), 1864–1874. <https://doi.org/10.1074/mcp.TIR118.000686>.
- (15) Wisztorski, M.; Fatou, B.; Franck, J.; Desmons, A.; Farré, I.; Leblanc, E.; Fournier, I.;

- Salzet, M. Microproteomics by Liquid Extraction Surface Analysis: Application to FFPE Tissue to Study the Fimbria Region of Tubo-Ovarian Cancer. *Proteomics - Clin. Appl.* **2013**, 7 (3–4), 234–240. <https://doi.org/10.1002/prca.201200070>.
- (16) Quanico, J.; Franck, J.; Cardon, T.; Leblanc, E.; Wisztorski, M.; Salzet, M.; Fournier, I. NanoLC-MS Coupling of Liquid Microjunction Microextraction for on-Tissue Proteomic Analysis. *Biochim. Biophys. Acta - Proteins Proteomics* **2017**. <https://doi.org/10.1016/j.bbapap.2016.11.002>.
- (17) Harris, G. A.; Nicklay, J. J.; Caprioli, R. M. Localized in Situ Hydrogel-Mediated Protein Digestion and Extraction Technique for on-Tissue Analysis. *Anal. Chem.* **2013**. <https://doi.org/10.1021/ac3031493>.
- (18) Rizzo, D. G.; Prentice, B. M.; Moore, J. L.; Norris, J. L.; Caprioli, R. M. Enhanced Spatially Resolved Proteomics Using On-Tissue Hydrogel-Mediated Protein Digestion. *Anal. Chem.* **2017**, 89 (5), 2948–2955. <https://doi.org/10.1021/acs.analchem.6b04395>.
- (19) Monroe, E. B.; Jurchen, J. C.; Koszczuk, B. A.; Losh, J. L.; Rubakhin, S. S.; Sweedler, J. V. Massively Parallel Sample Preparation for the MALDI MS Analyses of Tissues. *Anal. Chem.* **2006**, 78 (19), 6826–6832. <https://doi.org/10.1021/ac060652r>.
- (20) Tillberg, P. W.; Chen, F.; Piatkevich, K. D.; Zhao, Y.; Yu, C. C.; English, B. P.; Gao, L.; Martorell, A.; Suk, H. J.; Yoshida, F.; Degennaro, E. M.; Roossien, D. H.; Gong, G.; Seneviratne, U.; Tannenbaum, S. R.; Desimone, R.; Cai, D.; Boyden, E. S. Protein-Retention Expansion Microscopy of Cells and Tissues Labeled Using Standard Fluorescent Proteins and Antibodies. *Nat. Biotechnol.* **2016**. <https://doi.org/10.1038/nbt.3625>.
- (21) Chen, F.; Tillberg, P. W.; Boyden, E. S. Expansion Microscopy. *Science (80-)*. **2015**. <https://doi.org/10.1126/science.1260088>.
- (22) Zhao, Y.; Bucur, O.; Irshad, H.; Chen, F.; Weins, A.; Stancu, A. L.; Oh, E.-Y.; DiStasio, M.; Torous, V.; Glass, B.; Stillman, I. E.; Schnitt, S. J.; Beck, A. H.; Boyden, E. S. Nanoscale Imaging of Clinical Specimens Using Pathology-Optimized Expansion Microscopy. *Nat. Biotechnol.* **2017**, 35 (8), 757–764. <https://doi.org/10.1038/nbt.3892>.
- (23) Wisztorski, M.; Quanico, J.; Franck, J.; Fatou, B.; Salzet, M.; Fournier, I. Droplet-Based Liquid Extraction for Spatially-Resolved Microproteomics Analysis of Tissue Sections. In *Methods in molecular biology (Clifton, N.J.)*; 2017; Vol. 1618, pp 49–63. https://doi.org/10.1007/978-1-4939-7051-3_6.

- (24) Cox, J.; Mann, M. MaxQuant Enables High Peptide Identification Rates, Individualized p.p.b.-Range Mass Accuracies and Proteome-Wide Protein Quantification. *Nat. Biotechnol.* **2008**, *26* (12), 1367–1372. <https://doi.org/10.1038/nbt.1511>.
- (25) Tyanova, S.; Temu, T.; Carlson, A.; Sinitcyn, P.; Mann, M.; Cox, J. Visualization of LC-MS/MS Proteomics Data in MaxQuant. *Proteomics* **2015**, *15* (8), 1453–1456. <https://doi.org/10.1002/pmic.201400449>.
- (26) Cox, J.; Neuhauser, N.; Michalski, A.; Scheltema, R. A.; Olsen, J. V.; Mann, M. Andromeda: A Peptide Search Engine Integrated into the MaxQuant Environment. *J. Proteome Res.* **2011**, *10* (4), 1794–1805. <https://doi.org/10.1021/pr101065j>.
- (27) Cox, J.; Hein, M. Y.; Lubner, C. A.; Paron, I.; Nagaraj, N.; Mann, M. Accurate Proteome-Wide Label-Free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ. *Mol. Cell. Proteomics* **2014**, *13* (9), 2513–2526. <https://doi.org/10.1074/mcp.M113.031591>.
- (28) Vizcaíno, J. A.; Deutsch, E. W.; Wang, R.; Csordas, A.; Reisinger, F.; Ríos, D.; Dianes, J. A.; Sun, Z.; Farrah, T.; Bandeira, N.; Binz, P.-A.; Xenarios, I.; Eisenacher, M.; Mayer, G.; Gatto, L.; Campos, A.; Chalkley, R. J.; Kraus, H.-J.; Albar, J. P.; Martinez-Bartolomé, S.; Apweiler, R.; Omenn, G. S.; Martens, L.; Jones, A. R.; Hermjakob, H. ProteomeXchange Provides Globally Coordinated Proteomics Data Submission and Dissemination. *Nat. Biotechnol.* **2014**, *32* (3), 223–226. <https://doi.org/10.1038/nbt.2839>.
- (29) Vizcaíno, J. A.; Côté, R. G.; Csordas, A.; Dianes, J. A.; Fabregat, A.; Foster, J. M.; Griss, J.; Alpi, E.; Birim, M.; Contell, J.; O’Kelly, G.; Schoenegger, A.; Ovelheiro, D.; Pérez-Riverol, Y.; Reisinger, F.; Ríos, D.; Wang, R.; Hermjakob, H. The PRoteomics IDentifications (PRIDE) Database and Associated Tools: Status in 2013. *Nucleic Acids Res.* **2013**, *41* (Database issue), D1063-9. <https://doi.org/10.1093/nar/gks1262>.
- (30) Hulsen, T.; de Vlieg, J.; Alkema, W. BioVenn - A Web Application for the Comparison and Visualization of Biological Lists Using Area-Proportional Venn Diagrams. *BMC Genomics* **2008**, *9* (1), 488. <https://doi.org/10.1186/1471-2164-9-488>.
- (31) Treweek, J. B.; Chan, K. Y.; Flytzanis, N. C.; Yang, B.; Deverman, B. E.; Greenbaum, A.; Lignell, A.; Xiao, C.; Cai, L.; Ladinsky, M. S.; Bjorkman, P. J.; Fowlkes, C. C.; Gradinaru, V. Whole-Body Tissue Stabilization and Selective Extractions via Tissue-Hydrogel Hybrids for High-Resolution Intact Circuit Mapping and Phenotyping. *Nat. Protoc.* **2015**, *10* (11), 1860–1896. <https://doi.org/10.1038/nprot.2015.122>.

- (32) Ku, T.; Swaney, J.; Park, J. Y.; Albanese, A.; Murray, E.; Hun Cho, J.; Park, Y. G.; Mangena, V.; Chen, J.; Chung, K. Multiplexed and Scalable Super-Resolution Imaging of Three-Dimensional Protein Localization in Size-Adjustable Tissues. *Nat. Biotechnol.* **2016**, *34* (9), 973–981. <https://doi.org/10.1038/nbt.3641>.
- (33) Shevchenko, A.; Tomas, H.; Havliš, J.; Olsen, J. V.; Mann, M. In-Gel Digestion for Mass Spectrometric Characterization of Proteins and Proteomes. *Nat. Protoc.* **2007**, *1* (6), 2856–2860. <https://doi.org/10.1038/nprot.2006.468>.
- (34) Wiśniewski, J. R. Proteomic Sample Preparation from Formalin Fixed and Paraffin Embedded Tissue. *J. Vis. Exp.* **2013**, No. 79. <https://doi.org/10.3791/50589>.
- (35) Delcourt, V.; Franck, J.; Quanico, J.; Gimeno, J.-P.; Wisztorski, M.; Raffo-Romero, A.; Kobeissy, F.; Roucou, X.; Salzet, M.; Fournier, I. Spatially-Resolved Top-down Proteomics Bridged to MALDI MS Imaging Reveals the Molecular Physiome of Brain Regions. *Mol. Cell. Proteomics* **2018**, *17* (2), 357–372. <https://doi.org/10.1074/mcp.M116.065755>.
- (36) Franck, J.; Quanico, J.; Wisztorski, M.; Day, R.; Salzet, M.; Fournier, I. Quantification-Based Mass Spectrometry Imaging of Proteins by Parafilm Assisted Microdissection. *Anal. Chem.* **2013**. <https://doi.org/10.1021/ac4009397>.
- (37) Quanico, J.; Franck, J.; Gimeno, J. P.; Sabbagh, R.; Salzet, M.; Day, R.; Fournier, I. Parafilm-Assisted Microdissection: A Sampling Method for Mass Spectrometry-Based Identification of Differentially Expressed Prostate Cancer Protein Biomarkers. *Chem. Commun.* **2015**. <https://doi.org/10.1039/c4cc08331h>.
- (38) Katagiri, T.; Hatano, N.; Aihara, M.; Kawano, H.; Okamoto, M.; Liu, Y.; Izumi, T.; Maekawa, T.; Nakamura, S.; Ishihara, T.; Shirai, M.; Mizukami, Y. Proteomic Analysis of Proteins Expressing in Regions of Rat Brain by a Combination of SDS-PAGE with Nano-Liquid Chromatography-Quadrupole-Time of Flight Tandem Mass Spectrometry. *Proteome Sci.* **2010**, *8* (1), 41. <https://doi.org/10.1186/1477-5956-8-41>.
- (39) Erondy, N.; Kennedy, M. Regional Distribution of Type II Ca²⁺/Calmodulin-Dependent Protein Kinase in Rat Brain. *J. Neurosci.* **1985**, *5* (12), 3270–3277. <https://doi.org/10.1523/JNEUROSCI.05-12-03270.1985>.
- (40) Bott, C. J.; Winckler, B. Intermediate Filaments in Developing Neurons: Beyond Structure. *Cytoskeleton.* 2020. <https://doi.org/10.1002/cm.21597>.
- (41) Bamberg, J. R.; Bernstein, B. W.; Davis, R. .; Flynn, K. C.; Goldsbury, C.; Jensen, J. R.;

- Maloney, M. T.; Marsden, I. T.; Minamide, L. S.; Pak, C. W.; Shaw, A. E.; Whiteman, I.; Wiggan, O. ADF/Cofilin-Actin Rods in Neurodegenerative Diseases. *Curr. Alzheimer Res.* **2010**. <https://doi.org/10.2174/156720510791050902>.
- (42) Piehowski, P. D.; Zhu, Y.; Bramer, L. M.; Stratton, K. G.; Zhao, R.; Orton, D. J.; Moore, R. J.; Yuan, J.; Mitchell, H. D.; Gao, Y.; Webb-Robertson, B. J. M.; Dey, S. K.; Kelly, R. T.; Burnum-Johnson, K. E. Automated Mass Spectrometry Imaging of over 2000 Proteins from Tissue Sections at 100-Mm Spatial Resolution. *Nat. Commun.* **2020**, *11* (1). <https://doi.org/10.1038/s41467-019-13858-z>.
- (43) Chen, F.; Wassie, A. T.; Cote, A. J.; Sinha, A.; Alon, S.; Asano, S.; Daugharthy, E. R.; Chang, J. B.; Marblestone, A.; Church, G. M.; Raj, A.; Boyden, E. S. Nanoscale Imaging of RNA with Expansion Microscopy. *Nat. Methods* **2016**, *13* (8), 679–684. <https://doi.org/10.1038/nmeth.3899>.
- (44) Asano, S. M.; Gao, R.; Wassie, A. T.; Tillberg, P. W.; Chen, F.; Boyden, E. S. Expansion Microscopy: Protocols for Imaging Proteins and RNA in Cells and Tissues. *Curr. Protoc. Cell Biol.* **2018**, *80* (1), e56. <https://doi.org/10.1002/cpcb.56>.
- (45) Budnik, B.; Levy, E.; Harmange, G.; Slavov, N. SCoPE-MS: Mass Spectrometry of Single Mammalian Cells Quantifies Proteome Heterogeneity during Cell Differentiation. *Genome Biol.* **2018**, *19* (1). <https://doi.org/10.1186/s13059-018-1547-5>.
- (46) Cardon, T.; Salzet, M.; Franck, J.; Fournier, I. Nuclei of HeLa Cells Interactomes Unravel a Network of Ghost Proteins Involved in Proteins Translation. *Biochim. Biophys. Acta - Gen. Subj.* **2019**, *1863* (10), 1458–1470. <https://doi.org/10.1016/j.bbagen.2019.05.009>.
- (47) Cardon, T.; Hervé, F.; Delcourt, V.; Roucou, X.; Salzet, M.; Franck, J.; Fournier, I. Optimized Sample Preparation Workflow for Improved Identification of Ghost Proteins. *Anal. Chem.* **2020**, *92* (1), 1122–1129. <https://doi.org/10.1021/acs.analchem.9b04188>.
- (48) Truckenbrodt, S.; Maidorn, M.; Crzan, D.; Wildhagen, H.; Kabatas, S.; Rizzoli, S. O. X10 Expansion Microscopy Enables 25-nm Resolution on Conventional Microscopes. *EMBO Rep.* **2018**, *19* (9). <https://doi.org/10.15252/embr.201845836>.
- (49) Park, H.; Choi, D.; Park, J. S.; Sim, C.; Park, S.; Kang, S.; Yim, H.; Lee, M.; Kim, J.; Pac, J.; Rhee, K.; Lee, J.; Lee, Y.; Lee, Y.; Kim, S. Scalable and Isotropic Expansion of Tissues with Simply Tunable Expansion Ratio. *Adv. Sci.* **2019**, *6* (22), 1901673. <https://doi.org/10.1002/advs.201901673>.
- (50) Chang, J. B.; Chen, F.; Yoon, Y. G.; Jung, E. E.; Babcock, H.; Kang, J. S.; Asano, S.;

Suk, H. J.; Pak, N.; Tillberg, P. W.; Wassie, A. T.; Cai, D.; Boyden, E. S. Iterative Expansion Microscopy. *Nat. Methods* **2017**, *14* (6), 593–599. <https://doi.org/10.1038/nmeth.4261>.

Table 1 Expansion factor and number of proteins identified according to the homogenizing agent

	<i>Homogenization</i>	<i>Expansion Factor</i>	<i>Real size (Number of cells)</i>	<i>Number of proteins ID (control)</i>
<i>5x5mm²</i>	Proteinase K	2.4	2.08x2.08mm ² (~24 500)	2 (2)
	Trypsin	2.1	2.38x2.38mm ² (~32 000)	536 (2)
	LysC	2.1	2.38x2.38mm ² (~32 000)	495 (0)
	SDS	2.3	2.17x2.17mm ² (~26 600)	488 (0)
<i>1x1mm²</i>	Proteinase K	2.4	416x416μm ² (~980)	0 (0)
	Trypsin	2.1	476x476μm ² (~1 280)	35 (2)
	LysC	2.1	476x476μm ² (~1 280)	65 (0)
	SDS	2.3	435x435μm ² (~1 070)	32 (0)

Table 2 Expansion factor, estimation of number of cells and number of proteins identified after homogenization with SDS in each triplicate

	Section (1,2cm)	Size after expansion	Expansion Factor	Real size (Number of cells)	Number of proteins ID (control)
<i>5x5mm²</i>	S1	3,7 cm	3,08	1.62x1.62mm ² (~14 860)	898 (0)
	S2	3,65 cm	3,04	1.64x1.64mm ² (~15 229)	945 (0)
	S3	3,65 cm	3,04	1.64x1.64mm ² (~15 229)	994 (1)
<i>1x1mm²</i>	S1	3,7 cm	3,08	325x325μm ² (~598)	335 (0)
	S2	3,65 cm	3,04	329x329μm ² (~613)	392 (0)
	S3	3,65 cm	3,04	329x329μm ² (~613)	454 (1)
<i>punch biopsy</i>	S1	3,7 cm	3,08	∅ 454 μm ² (~916)	582 (0)
	S2	3,65 cm	3,04	∅ 460 μm ² (~941)	682 (0)
	S3	3,65 cm	3,04	∅ 460 μm ² (~941)	700 (1)

Table 3 Expansion factor and number of proteins identified in fresh frozen tissue after homogenization with SDS in each triplicate

Gel + tissue (1,5cm)	Size after expansion	Expansion Factor	Real size (Number of cells)	Number of proteins ID (control)
S1	4,9 cm	3,26	1.53x1.53mm ² (~13 248)	136 (10)
S3	5 cm	3,33	1.50x1.50mm ² (~12 733)	78 (26)
S2	4,8 cm	3,2	1.56x1.56mm ² (~13 772)	788 (0)

Figure 1. General workflow of the micro-proteomic protocol based on gel expansion

A. Micro-proteomic analysis is performed by cutting area directly in the gel using a scalpel. Gel pieces are then digested and finally peptides are extracted and analyzed by tandem mass spectrometry. B. Rat brain tissue section in its original size of 1,2cm, before expansion and B. the same tissue section after expansion with a front size of 3,8cm. Expansion factor is calculated by dividing the measure after expansion (c) by the measure before expansion (a). In this case we can estimate that the expansion factor is around 3.

Figure 2. Comparative analysis of different alternatives for Proteinase K

A/B/C/D. Expanded rat brain section after different homogenization process. A. Proteinase K. B. Trypsin. C. Lys C and D. SDS. Squares correspond to the different area cut: 5x5mm² and 1x1mm² inside the tissue and a control area (5x5mm²) in the gel only. The red dotted line delimits the contour of the tissue section. Proteins in these areas were digested and identified by mass spectrometry. E/F. Number of proteins identified in each area is represented in Venn diagrams. E. For 5x5mm, 536 proteins were quantified using trypsin as homogenizing agent (B), 495 for Lys C (C) and 488 for SDS (D). F. For 1x1mm² regions, 35 proteins were quantified for trypsin (B), 65 protein with Lys-C (C) and 32 proteins for SDS (D).

Figure 3. Analysis of proteomic reproducibility using the SDS homogenization protocol

A. Expanded rat brain section after SDS homogenization, red square stand for the analyzed region. B/D. Experiment were performed in triplicates on 3 consecutive tissue sections Venn diagram representation of the number of proteins identified in extraction of (B) 5x5mm in size and (D) 1x1mm. C/E. Pearson correlation of label free quantification of proteins identified in extraction of (C) 5x5mm in size and (E) 1x1mm. F. Comparison of proteins identified in expansion and with digestion in situ and liquid micro junction extraction.

Figure 4. Utilization of punch biopsy improve proteomic reproducibility

A. Disk of gel cut using a punch biopsy of 1350 µm in diameter. B. Venn diagram representation of the number of proteins identified in punch biopsy extraction. Experiment were performed in triplicates on 3 consecutive tissue sections C. Pearson correlation of LFQ in each replicate.

Figure 5. Quantification-based mass spectrometry profiling using tissue expansion

Image of the expanded rat brain tissue section, squares represents extractions for microproteomic, and reconstructed distribution of representative proteins based on label free quantification provided by nanoLC-MS/MS analysis. ec: external capsule; fi: fimbria of the hippocampus; ic: internal capsule;

Cx: cerebral cortex; st: stria terminalis; opt: optic tract; eml: external medullary lamina; ml: medial lemniscus

Supplementary Figure 1. Protein Identification after homogenization with Proteinase K

Few proteins are identified (14) in the tissue and significant migration is observed in the control sample (9 proteins).

Supplementary Figure 2. Side by side comparison

Direct comparison shows greater similarity between S2 and S3 replicates than with S1 replicate **A.** 5x5mm² **B.** 1x1mm² **C.** Punch biopsy

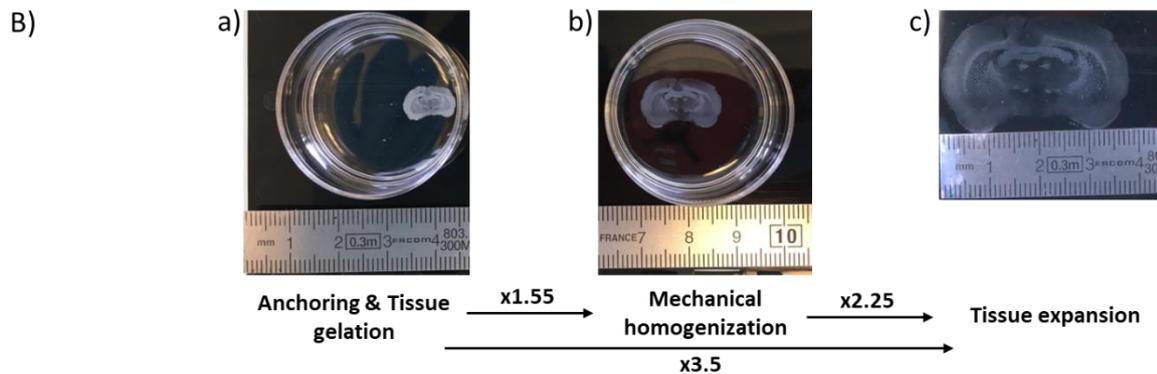
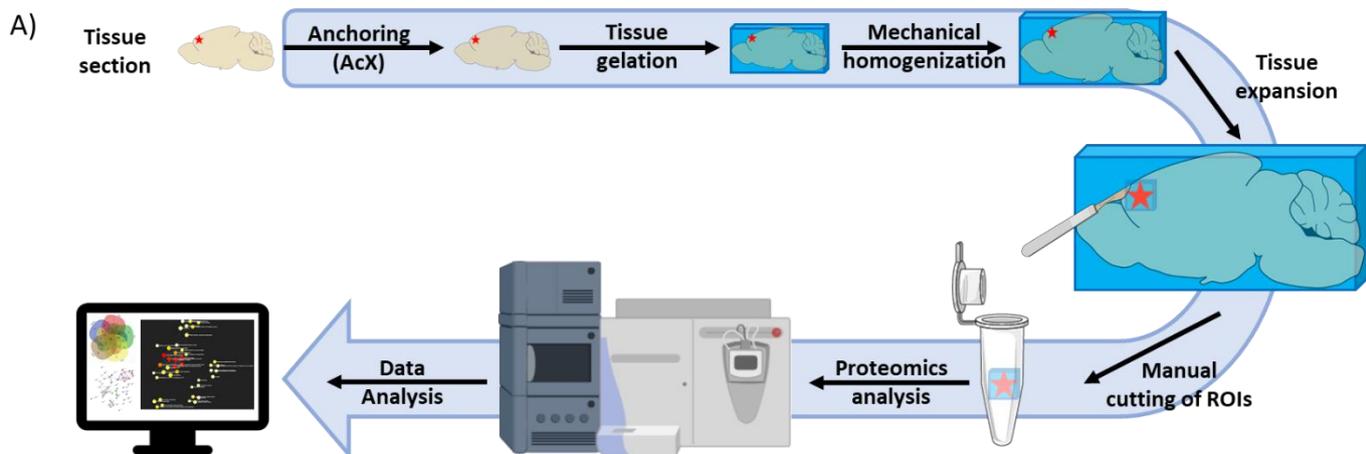
Supplementary Figure 3. Section of fresh frozen rat brain after expansion

Section of fresh frozen rat brain after expansion and staining with Hoescht observed under an ultraviolet lamp

Supplementary Figure 4. Qualitative comparison of proteins identified after expansion and with in situ digestion and LESA extraction

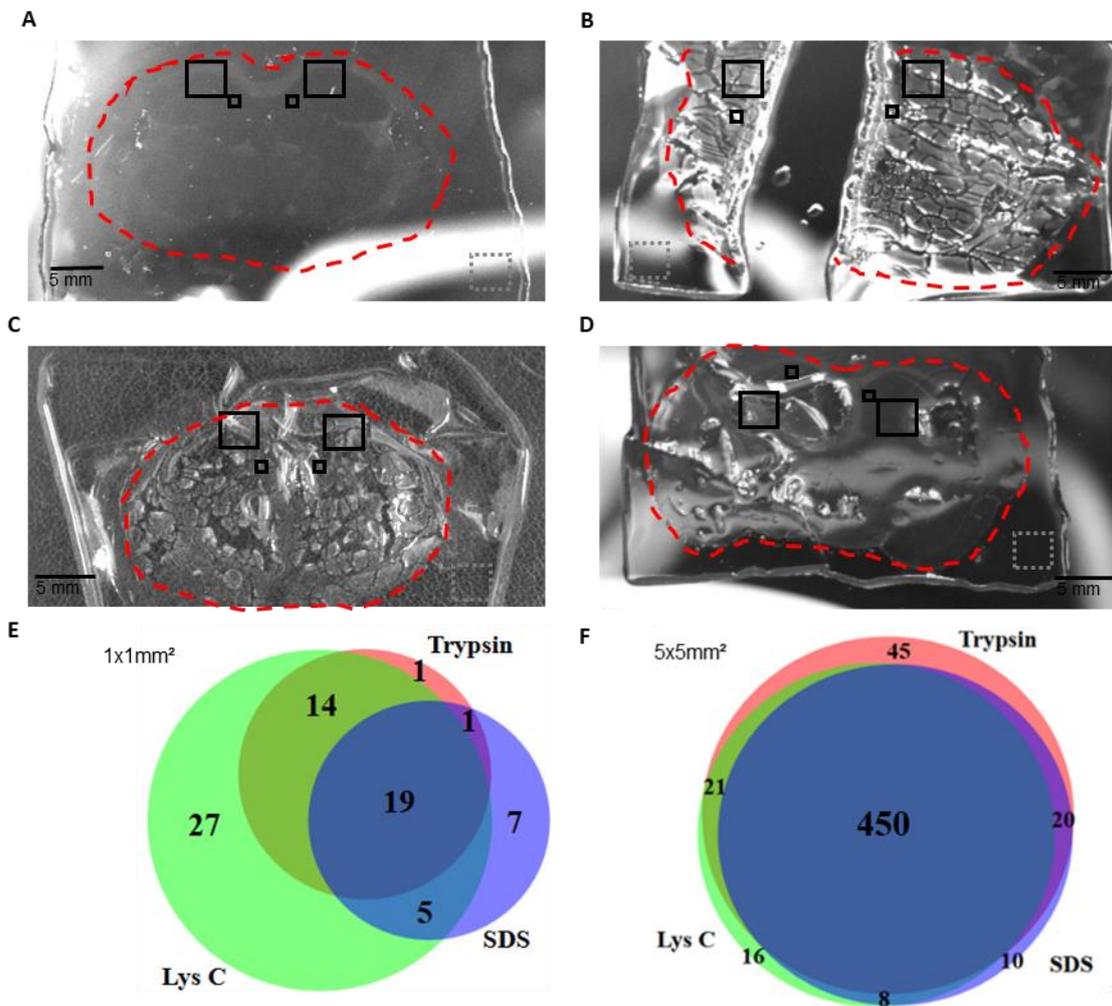
A. Pearson's correlation between LFQ and **B.** Estimated molecular weight of proteins detected with each condition

Figure 1. General workflow of the micro-proteomic protocol based on gel expansion



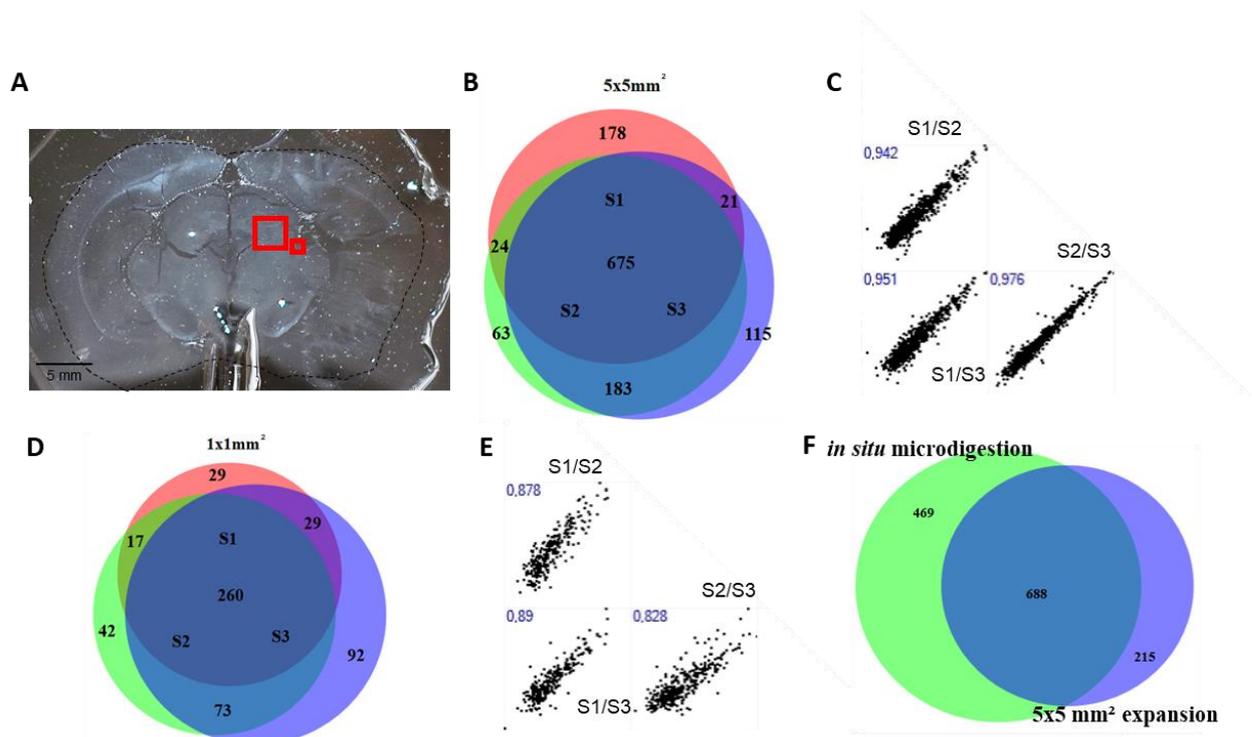
A. Micro-proteomic analysis is performed by cutting area directly in the gel using a scalpel. Gel pieces are then digested and finally peptides are extracted and analyzed by tandem mass spectrometry. B. Rat brain tissue section in its original size of 1,2cm, before expansion and B. the same tissue section after expansion with a front size of 3,8cm. Expansion factor is calculated by dividing the measure after expansion (c) by the measure before expansion (a). In this case we can estimate that the expansion factor is around 3.

Figure 2. Comparative analysis of different alternatives for Proteinase K



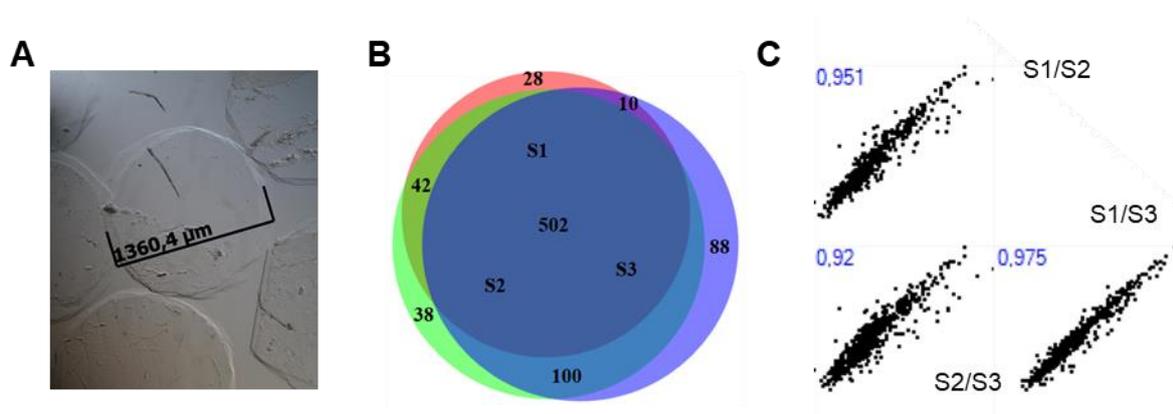
A/B/C/D. Expanded rat brain section after different homogenization process. A. Proteinase K. B. Trypsin. C. Lys C and D. SDS. Squares correspond to the different area cut: 5x5mm² and 1x1mm² inside the tissue and a control area (5x5mm²) in the gel only. The red dotted line delimits the contour of the tissue section. Proteins in these area were digested and identified by mass spectrometry. E/F. Number of proteins identified in each area is represented in Venn diagrams. E. For 5x5mm, 536 proteins were quantified using trypsin as homogenizing agent (B), 495 for Lys C (C) and 488 for SDS (D). F. For 1x1mm² regions, 35 proteins were quantified for trypsin (B), 65 protein with Lys-C (C) and 32 proteins for SDS (D).

Figure 3. Analysis of proteomic reproducibility using the SDS homogenization protocol



A. Expanded rat brain section after SDS homogenization, red square stand for the analyzed region. B/D. Experiment were performed in triplicates on 3 consecutives tissue sections Venn diagram representation of the number of proteins identified in extraction of (B) 5x5mm in size and (D) 1x1mm. C/E. Pearson correlation of label free quantification of proteins identified in extraction of (C) 5x5mm in size and (E) 1x1mm. F. Comparison of proteins identified in expansion and with digestion in situ and liquid micro junction extraction.

Figure 4. Utilization of punch biopsy improve proteomic reproducibility



A. Disk of gel cut using a punch biopsy of 1350 µm in diameter. B. Venn diagram representation of the number of proteins identified in punch biopsy extraction. Experiment were performed in triplicates on 3 consecutives tissue sections C. Pearson correlation of LFQ in each replicate.

Figure 5. Quantification-based mass spectrometry profiling using tissue expansion

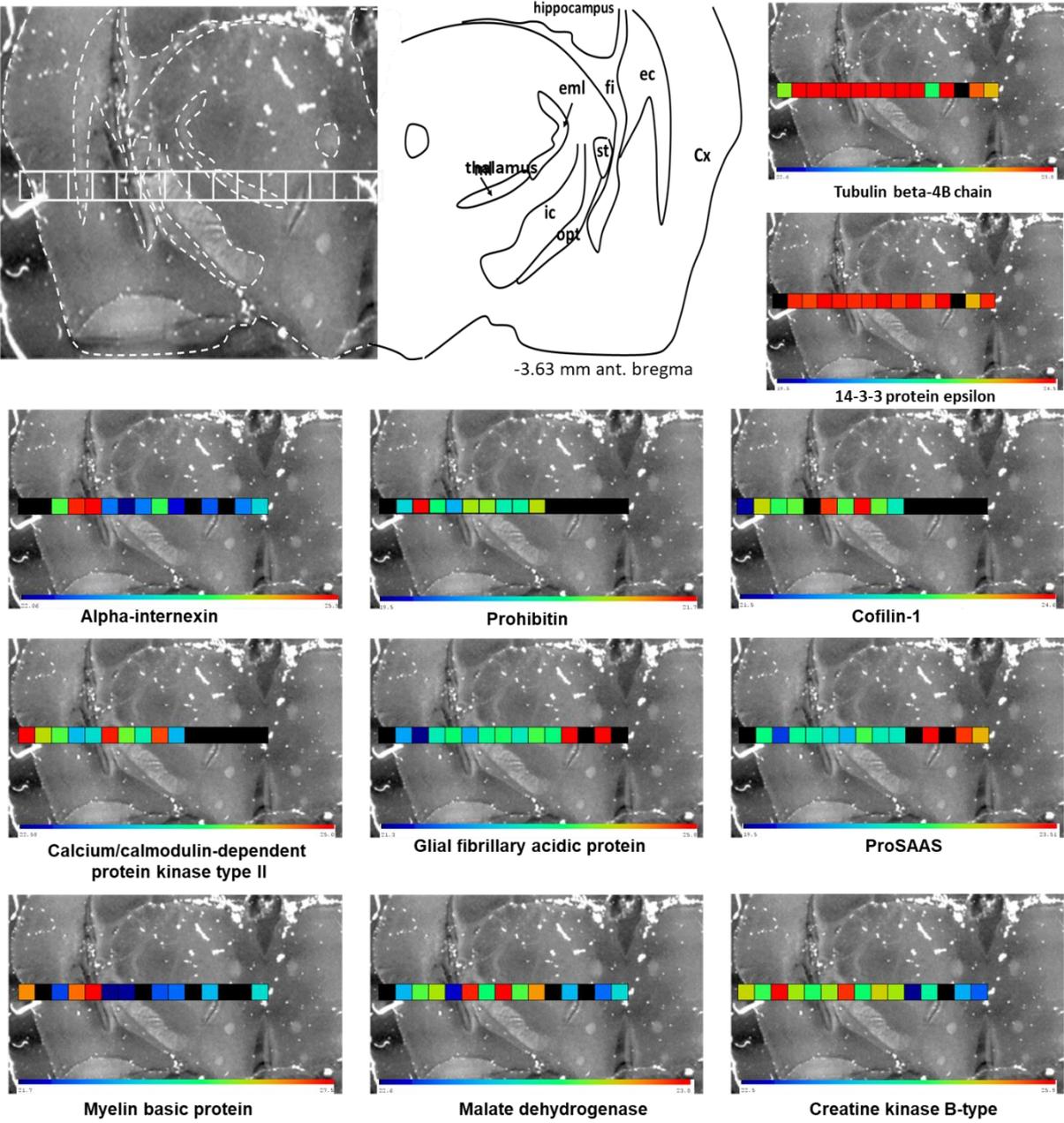
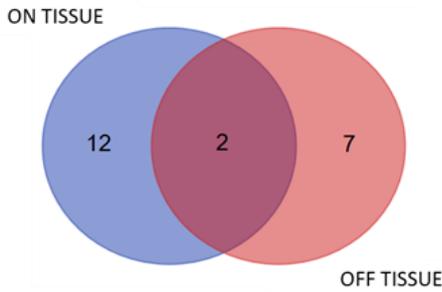


Image of the expanded rat brain tissue section, squares represents extractions for microproteomic, and reconstructed distribution of representative proteins based on label free quantification provided by nanoLC-MS/MS analysis. ec: external capsule; fi: fimbria of the hippocampus; ic: internal capsule; Cx: cerebral cortex; st: stria terminalis; opt: optic tract; eml: external medullary lamina; ml: medial lemniscus

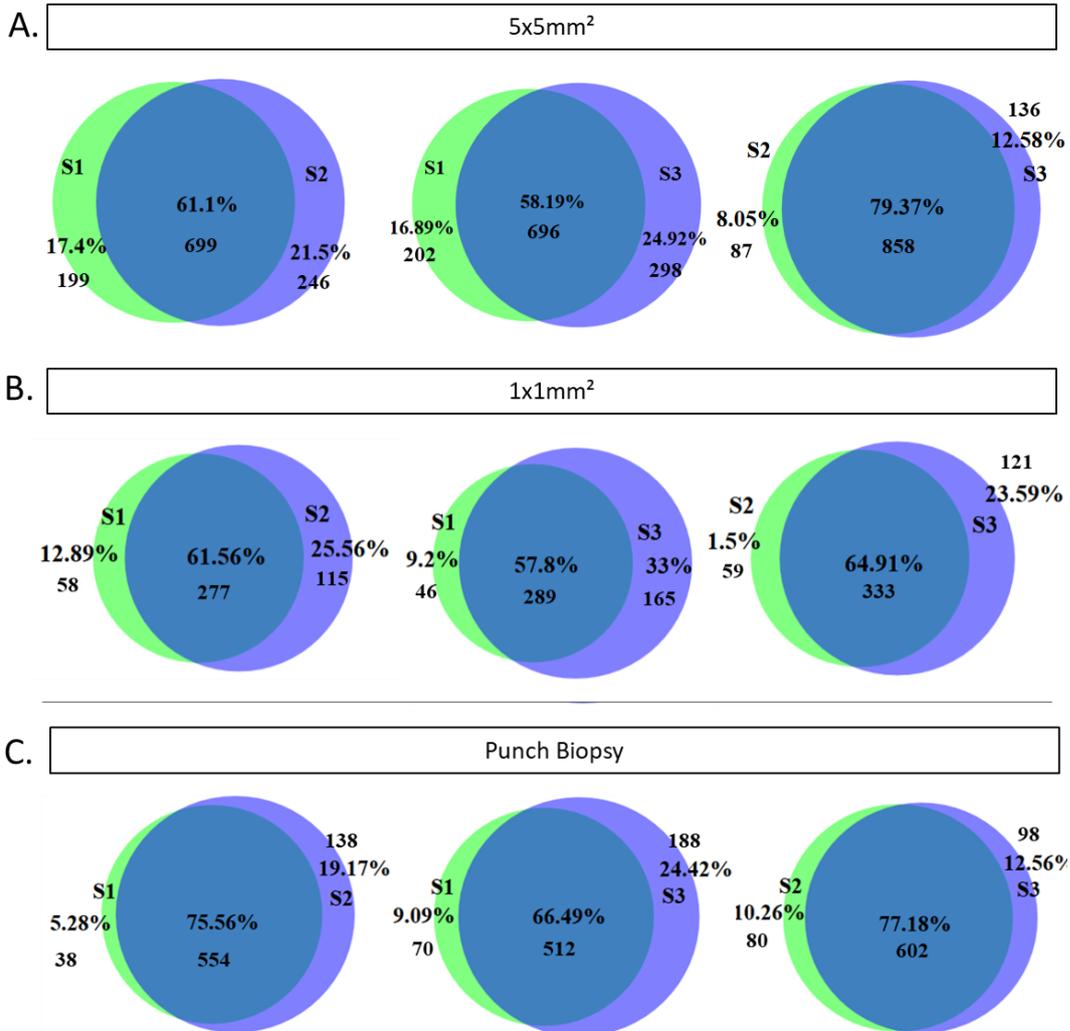
Supplementary. Figure

Supplementary Figure 1. Protein Identification after homogenization with Proteinase K



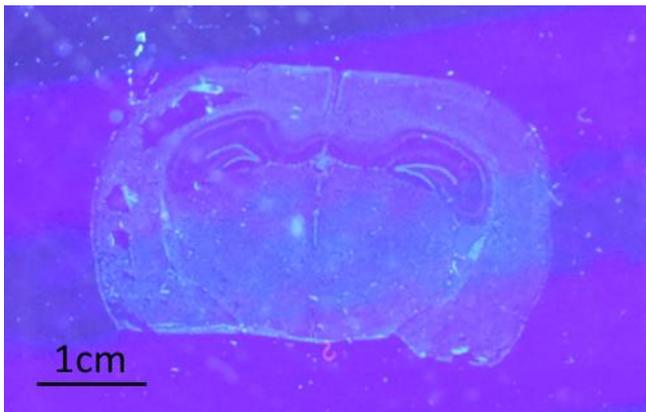
Few proteins are identified (14) in the tissue and significant migration is observed in the control sample (9 proteins).

Supplementary Figure 2. Side by side comparison



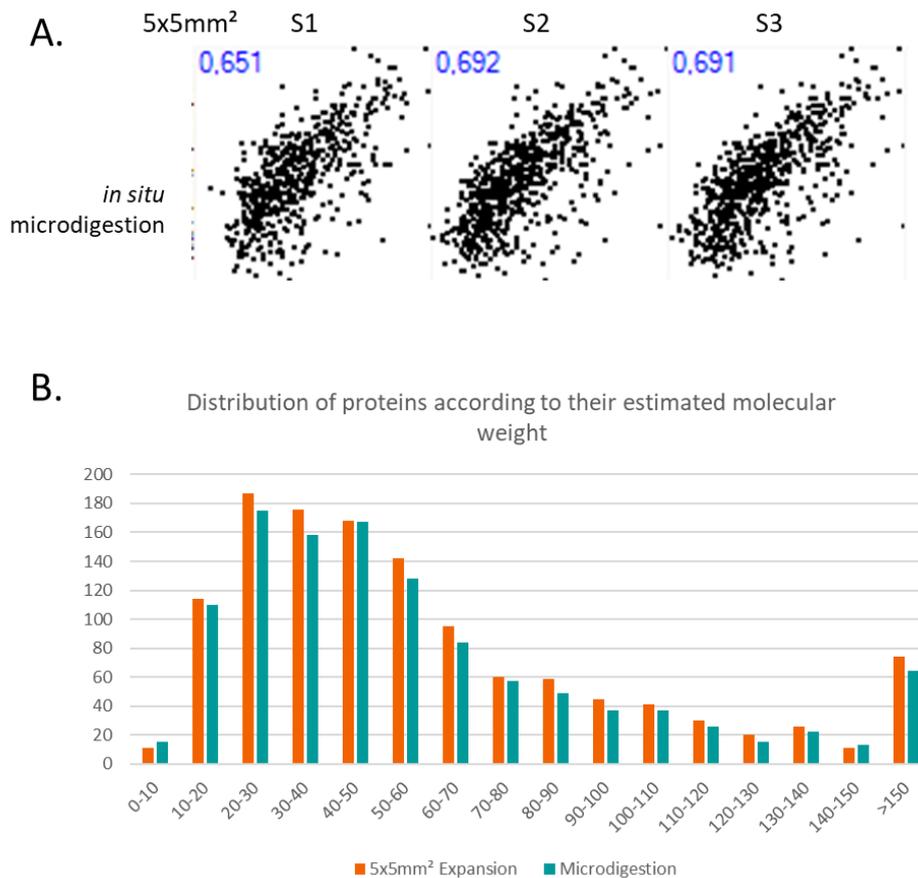
Direct comparison shows greater similarity between S2 and S3 replicates than with S1 replicate A. 5x5mm² B. 1x1mm² C. Punch biopsy

Supplementary Figure 3. Section of fresh frozen rat brain after expansion



Section of fresh frozen rat brain after expansion and staining with Hoescht observed under an ultraviolet lamp

Supplementary Figure 4. Qualitative comparison of proteins identified after expansion and with in situ digestion and LESA extraction



A. Pearson's correlation between LQF and **B.** Estimated molecular weight of proteins detected with each condition

Au sein des gliomes de haut-grades, les glioblastomes sont les tumeurs primitives du système nerveux central les plus communes et les plus agressives. L'imagerie par spectrométrie de masse MALDI est une technique intéressante pour l'étude de ces tumeurs, car elle permet de prendre en considération l'hétérogénéité intra-tumorale. Dans cette étude, l'imagerie MALDI MS est couplée à la microprotéomique localisée sur tissu dans l'objectif d'identifier des sous-groupes de glioblastomes afin d'aider au diagnostic et au pronostic. Les images moléculaires ont été générées à partir de coupes fines de tissus dans le but de déterminer la localisation de peptides digérés. Ensuite, des analyses statistiques non supervisées nous ont permis de générer un regroupement hiérarchique de régions moléculaires homogènes au sein du tissu. Le regroupement hiérarchique a révélé trois régions moléculaires dans l'ensemble des échantillons. À partir de ces régions, 3 groupes de patients ont été identifiés, cinq patients sont restés non classables. Chaque groupe présentait des signatures moléculaires spécifiques. Les analyses microprotéomiques ont montré un panel de protéines spécifiquement surexprimées dans chaque groupe : les protéines surexprimées dans le groupe A sont associées à la neurogenèse ; celles du groupe B sont associées à une activation du système immunitaire et celles du groupe C sont impliquées dans des processus viraux. Enfin, nous avons identifié 6 nouveaux marqueurs pronostiques des glioblastomes qui pourraient aider à stratifier les patients et aider à la décision thérapeutique. En parallèle, nous nous sommes intéressés au développement d'une nouvelle technique de microprotéomique localisée basée sur l'expansion de tissu. Jusqu'à 655 protéines ont pu être identifiées dans une région de 460 μm de diamètre, ce qui correspond à une moyenne de 940 cellules. De plus, cette stratégie est intéressante pour l'imagerie MS, car elle donne la possibilité de cartographier facilement un grand nombre de protéines sur la base de leur quantification avec une résolution spatiale inférieure à 350 μm . Finalement, étudier les vésicules extracellulaires circulant dans le sang pourrait à terme permettre d'aider au dépistage, au diagnostic ainsi qu'au suivi des glioblastomes. De plus, nous avons été capables d'isoler spécifiquement les vésicules contenues dans une coupe fine de cerveau de rat.

Among high-grade gliomas, glioblastomas are the most common and aggressive tumors. To study gliomas, MALDI mass spectrometry imaging is a technique of choice, allowing for tumor heterogeneity analysis. In this study, MALDI-MSI is coupled with spatially-resolved microproteomic aiming at identifying subgroups of glioblastomas patients in order to help for diagnosis and prognosis. Molecular images are generated from thin tissue sections to determine digested peptides spatial localizations. Based on unsupervised statistical analysis, we generated hierarchical clustering of homogeneous molecular regions. The hierarchical clustering reveals three molecular regions within all the tumor samples. Based on these regions, three groups of patients can be determined, each group with specific molecular signatures. Microproteomic analyzes show a panel of proteins specifically overexpressed in each group: proteins overexpressed in group A are associated with neurogenesis; those of group B are linked to immune system activation; and those of group C are involved in viral processes. Finally, we have identified 6 new prognosis markers for glioblastomas that could help stratifying patients and orient clinician in choosing therapeutics. In parallel, we were interested in developing a novel spatially resolved microproteomic technique based on tissue expansion. Up to 655 proteins were identified within a region of 460 μm diameter, which corresponds to an average of 940 cells. In addition, this strategy is relevant for MS imaging as it gives the possibility to easily map a large number of proteins based on their quantification within a spatial resolution of less than 350 μm . Finally, studying circulating extracellular vesicles in blood may ultimately lead to an earlier diagnosis of glioblastomas. Preliminary results from control plasma show that it is possible to isolate and perform proteomic analyzes on circulating vesicles. In addition, we have been able to specifically isolate the vesicles embedded in a thin section of rat brain. These results are encouraging and, in the long term, will make it possible to identify vesicles originating from the tumor.