
**Investigation of extracellular vesicles from
glioblastoma multiforme and meningioma patients
for cancer liquid biopsy by differential quantitative
proteomics**

Dissertation

For the acquisition of the academic degree

Doctor rerum naturalium (Dr. rer. nat.)

Faculty of Mathematics, Informatics and Natural Sciences

Department of Chemistry

University of Hamburg

Submitted by

Marceline Manka Fuh, M. Sc.

from Bafut, Cameroon

Hamburg

2019

Supervisors during thesis:

Supervisor: Prof. Dr. Hartmut Schlüter

Co-Supervisor: Prof. Dr. Dr. Christian Betzel

Recommendation of the evaluators for the dissertation:

Prof. Dr. Hartmut Schlüter

Dr. Maria Riedner

Date of Defence: 27.09.2019

This work was done from September 2016 to June 2019 in the department of clinical chemistry at the University Medical Centre Eppendorf, Hamburg (UKE) in the mass spectrometric proteome analytics working group, under the supervision of Prof. Dr. Hartmut Schlüter.

Dedicated to my parents, siblings, nieces and nephews

Abstract

Although intracranial cancers are typically not among the most common cancers in adults, they encompass one of the deadliest and are the second most common cancers in children. Gliomas account for over 78% of malignant brain tumours, with glioblastoma multiforme (GBM) being the most invasive. Despite notable recent achievements in oncology research, GBM patients have a survival rate between 15 and 23 months and a 5-year survival less than 6%. Meningiomas (MEN) are less invasive but however, make up 36.6% of all primary brain and other CNS tumours and still possess life threatening characteristics. Like other cancers, early and non/minimally invasive diagnosis therefore is vital for better prognosis. The ability of extracellular vesicles (EVs) to transfer molecular cargo and be selectively taken up by specific cells can be exploited in multiple facets of cancer biology. The main goal of this study was therefore to employ SWATH-DIA mass spectrometric technique in identifying potential GBM and MEN EV specific biomarkers for cancer liquid biopsies, hence facilitating early detection and personalised cure.

Tumour tissues were obtained from four GBM patients and four MEN patients. Primary cells were cultured from each patient tissue and Extracellular vesicles were spun down from each cell culture supernatant. Firstly, the variable window method for SWATH-DIA was compared to the classic go-to DDA method. SWATH-DIA produced outstanding results and was therefore the method of choice throughout this study. With an average of 2500 to 3000 proteins quantified in every patient tissue and cells, proteins which were quantified in every patient's tissue, cells and EVs within the various cancer types were further analysed to obtain specific cancer-EV markers. A list of 99 proteins were identified as unique to the MEN samples and therefore postulated as potential MEN-EV specific markers. For GBM, 91 proteins were unique to these samples. When compared to existing literature of potential GBM-EV specific proteins, interestingly, 14 GBM-EV specific proteins were identified. Pathway analysis of these 14 proteins indicated that; YWHAQ, YWHAG and THBS1 are associated with the PI3K/Akt signalling pathways which have been reported to be frequently disturbed in many human cancers; VCL, TLN1, IQGAP1, VCP, A2M, F2, ANXA1, THBS1 and LGALS3BP are associated with the biological process of vesicle mediated transport while FLNB, IQGAP1 and THBS1 are categorised as proteoglycans in cancer. These proteins may serve as potential GBM-EV specific markers and are therefore subject to further validation studies. The variation between these tissues and their primary cells was also studied. Primary tumour cells however showed little variation to the original tumours, maintaining many of the important markers. This supports their current stance as a great sample source especially for research in personalised medicine. Differential proteomic studies were also carried out on tumour (GBM and MEN) and non-tumour epileptic (white matter and cortex) brain tissues. This goal was also achieved as lists of aberrantly regulated proteins were generated in

each cancer type. Specific proteins involved in tumourigenesis and metastasis were quantified in both cases; and represent panels for further validation. Finally, identified cancer-EV markers which were also significantly dysregulated in tumour tissues were studied. These therefore could be potential tumour markers which can possibly be identified from liquid biopsy hence facilitating early and non/minimally invasive diagnosis of cancers which is vital for better prognosis. This is the first time which this approach is implemented for GBM and MEN EV biomarker discovery and may contribute massively towards research in the field of tumour diagnosis from patient EVs.

Zusammenfassung

Obwohl intrakranielle Krebserkrankungen bei Erwachsenen normalerweise nicht zu den häufigsten Krebserkrankungen gehören, umfassen sie eine der tödlichsten und sind die zweithäufigste Krebserkrankung bei Kindern. Gliomen machen über 78% der malignen Gehirntumoren aus, wobei Glioblastoma Multiforme (GBM) am invasivsten sind. Trotz der bemerkenswerten aktuellen Fortschritte in der Onkologie-Forschung haben GBM-Patienten eine Überlebensrate zwischen 15 bis 23 Monaten und eine 5-Jahres-Überlebensrate haben unter 6% der Patienten. Meningeome (MEN) sind weniger invasiv, bilden jedoch 36,6% aller primären Hirn- und anderen zentralen Nervensystem-(ZNS) Tumoren und besitzen lebensbedrohliche Eigenschaften. Wie bei anderen Krebserkrankungen ist eine frühzeitige und nicht-/minimalinvasive Diagnose entscheidend für eine bessere Prognose. Die Fähigkeit der extrazellulären Vesikel (EVs), molekulare Ladung zu übertragen und von spezifischen Zellen selektiv aufgenommen zu werden, kann für die Krebs-Diagnostik genutzt werden. Daher war das Hauptziel dieser Studie, EVs nach potenziellen spezifischen Markern der GBM- und MEN-EV zu suchen. Diese Marker könnten zukünftig mittels der flüssigen Krebsbiopsien zur Früherkennung und individuellen Therapie der Patienten beitragen.

Tumorgewebe wurden von vier GBM- und vier MEN-Patienten erhalten. Primärzellen wurden aus jedem Patientengewebe kultiviert und extrazelluläre Vesikel aus jedem Zellkultur-Überstand gewonnen. Zunächst wurde die Variable-Fenster-Methode für SWATH-DIA mit der klassischen DDA-Methode verglichen. SWATH-DIA lieferte hervorragende Ergebnisse und wurde deshalb die Methode der Wahl in dieser Studie. Mit einem Durchschnittswert von 2500 Proteinen in jeder Probe wurden Proteine, die in jedem Patientengewebe, Zellen und EVs der verschiedenen Krebstypen identifiziert und quantifiziert werden konnten, weiter analysiert, um spezifische Krebs-EV-Marker zu erhalten. Eine Liste von 99 Proteinen wurde für die MEN-Proben eindeutig identifiziert und daher als potenzielle MEN-EV-spezifische Marker postuliert. Für GBM waren 91 Proteine für diese Proben spezifisch. 14 potenziellen GBM-EV spezifische Proteine stimmten mit denen in der Literatur überein. Die Pathway-Analyse dieser 14 Proteine zeigte, dass YWHAQ, YWHAG und THBS1 im Zusammenhang mit den PI3K/Akt Signalwegen stehen, von denen berichtet wurde, dass sie bei vielen Krebserkrankungen häufig gestört sind. VCL, TLN1, IQGAP1, VCP, A2M, F2, ANXA1, THBS1 und LGALS3BP sind mit dem biologischen Prozess des durch Vesikel vermittelten Transports assoziiert, während FLNB, IQGAP1 und THBS1 als Proteoglykane eingestuft werden. Diesen Proteinen sind potenzielle GBM-EV spezifische Marker und sollten zukünftig validiert werden. In der vorliegenden Arbeit wurden zudem Krebs-Gewebe und ihre kultivierten Primärzellen untersucht. Primäre Tumorzellen zeigten nur geringe Unterschiede zu den Ursprungstumoren. Dies unterstützt ihre derzeitige Stellung als geeignete Probenquelle für die Forschung in der individuellen Medizin. Als

weiterer Punkt wurden differenzielle Proteomik-Studien an Tumoren (GBM und MEN) und Nicht-Tumor-Epileptikum (weiße Substanz und Kortex) im Gehirngewebe ausgeführt. Für jeden Krebstyp wurden Listen von Marker-Proteinen identifiziert. In beiden Fällen wurden Proteine, die an Tumorgenese und Metastase beteiligt sind, quantifiziert und repräsentieren Panels für die weitere Validierung. Anschließend wurden die identifizierten Krebs-EV-Marker, die auch in Tumorgeweben signifikant fehlreguliert waren, identifiziert. Diese könnten daher potenzielle Tumormarker sein, die möglicherweise anhand einer Flüssigkeitsbiopsie identifiziert werden können, wodurch eine frühzeitige und minimale invasive Diagnose von Krebserkrankungen erleichtert wird. Dies ist das erste Mal, dass dieser Ansatz für die Feststellung von GBM- und MEN-EV-Biomarkern eingesetzt wurde.

List of Publications

1. **M.M. Fuh**, L. Heikaus, and H. Schlüter, MALDI mass spectrometry in medical research and diagnostic routine laboratories. *International Journal of Mass Spectrometry*, 2017. 416: p. 96-109.
2. Dahmen, J., Otte, C., **Fuh, M.**, Uschold, S., Schlüter, M., Antoni, S.-T., et al. (2017). Massenspektrometrische Gewebeanalyse mittels OCT-navigierter PIR-Laserablation. In 16. Jahrestagung der Deutschen Gesellschaft für Computer- und Roboterassistierte Chirurgie (pp. 112-116). Garbsen: PZH Verlag, TEWISS-Technik und Wissen GmbH.
3. Martínez Sánchez, A.H., Omid, M., Wurlitzer, M., **Fuh M.M.**, Feyerabend, F., Schlüter, H., Willumeit-Römer, R., Luthringer, J.C.B. Proteome analysis of human mesenchymal stem cells undergoing chondrogenesis when exposed to the products of various magnesium-based materials degradation. *Bioactive Materials*, 2019. 4: p. 168-188.
4. Matthias Schlüter, **Manka M. Fuh**, Stephanie Maier, Christoph Otte, Parnian Kiani, Nils-Owe Hansen, R. J. Dwayne Miller, Hartmut Schlüter, and Alexander Schläefer Towards OCT-Navigated Tissue Ablation with a PIR Laser and Mass-Spectrometric Analysis of Ablation Plume
5. **M. Fuh**, P. Steffen, and Schlüter, H., Tools for the analysis and characterization of therapeutic protein species. *Biosimilars*, Vol. 2016:6 Pages 17—24.
6. Wurlitzer M, Hessling E, Rinas K, **Fuh M**, Petersen H, Ricklefs F, Lamszus K, Regelsberger J, Maier S, Kruber S, Hansen NO, Miller R, Schlüter H. Mass Spectrometric Lipid Profiles of Picosecond Infrared Laser-Generated Tissue Aerosols Discriminate Different Brain Tissues. *Lasers Surg Med*, 2019

List of poster presentations

1. Poster presentation at the 2nd Annual Lipidomics Forum (LMS), “Picosecond infrared laser (PIRL) desorption of lipids from porcine brain areas”, 13-15 November 2016, Dortmund, Germany
2. Poster presentation at the 50th Annual meeting of the German Society of Mass Spectrometry (DGMS), “Picosecond infrared laser (PIRL) desorption of lipids from porcine brain areas”, Feb – March 2017, Kiel, Germany.
3. Poster presentation at the Symposium des Forschungszentrum Medizintechnik Hamburg (fmthh), “OCT basierte Navigation für die Gewerbeablation und Charakterisierung mit einem PIR-Laser”, 24th January 2018, Hamburg, Germany.
4. Poster presentation at the Symposium des Forschungszentrum Medizintechnik Hamburg (fmthh), “OCT basierte Navigation für die Gewerbeablation und Charakterisierung mit einem PIR-Laser”, 22nd January 2019, Hamburg, Germany.

Table of Contents

Abstract	6
Zusammenfassung	8
List of Publications	10
List of Poster Presentations	11
Table of Contents	12
List of Abbreviations	15
1. Introduction	17
1.1 Cancers.....	17
1.2 Intracranial tumours	17
1.2.1 Glioblastoma Multiforme	20
1.2.1 Meningioma	21
1.3 Biomarker discovery in cancer diagnosis and prognosis	22
1.4 Cancer liquid biopsy and Extracellular vesicles.....	24
1.5 Mass spectrometry based clinical proteomics.....	26
1.5.1 Peptide ionisation and fragmentation.....	28
1.5.1 MS Quantification methods.....	30
2. Aim	32
3. Results	33
3.1 SWATH DIA MS of tumour tissues, primary cell culture and extracellular vesicles of glioblastomas and meningiomas	33
3.1.1 Spectral library generation	34
3.1.2 Spectral library evaluation	36
3.1.3 Data Acquisition	38
3.1.4 Tumour tissue proteome	43
3.1.5 Patient primary cell culture proteome	46
3.1.6 Extracellular vesicles	48
3.1.7 Comparison of proteomes identified in all Patient tissues, cells and EVs	51

3.2 Proteome Variation between tissues and patient-derived primary cell culture	54
3.2.1 Meningioma tissues and cell culture proteome	55
3.2.2 Glioblastoma tissues and cell culture proteome	58
3.3 Differential proteomic analysis of tumour and non-tumour tissues	60
3.3.1 Library generation and evaluation.....	60
3.3.2 Non-tumour brain tissues: White matter and cortex	64
3.3.3 Comparison of cancerous and non-cancerous brain tissues	66
3.3.3.1 Meningioma tissues versus non-tumour tissues	68
3.3.3.1 Glioblastoma tissues versus non-tumour tissues	69
3.4 Differences between data dependent and data independent acquisition.....	70
4. Discussion	72
4.1 The choice of DIA analysis with SWATH over DDA for this study	73
4.2 Comprehensive versus sample specific spectral libraries.....	74
4.3 Potential biomarkers for cancer liquid biopsy	75
4.4 Cell culture for clinical research.....	84
4.5 Differential Proteome of cancerous and non-cancerous brain tissues	85
5. Materials and methods	92
5.1 Chemicals, Instruments and software	92
5.2 Workflow overview	94
5.3 Brain biopsy samples.....	94
5.4 Lysis and protein extraction.....	95
5.5 In-solution protein digestion to peptides	95
5.6 Data dependent acquisition mass spectrometry data acquisition	96
5.7 Liquid chromatography and SWATH DIA mass spectrometry analysis.....	96
5.8 Data analysis	98
5.8.1 Data dependent LC-MS/MS experiments	98
5.8.2 Spectral library generation	98
5.8.3 Protein quantification and statistical analysis	99

5.8.4 Pathway and enrichment analysis	99
6. References	100
7. Supplement	111
8. Risk and safety statements	122
9. Acknowledgement	124
10. Declaration	125

List of abbreviations

Abbreviation	Meaning
ACN	Acetonitrile
AmbiCa	ammonium bicarbonate
BCA	Bicinchoninic acid
CBTRUS	Central Brain Tumour Registry of the United States
CID	Collision induced dissociation
CNS	Central nervous system
DDA	Data dependent Acquisition
DIA	Data independent Acquisition
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
EGFR	Epidermal growth factor receptor
ESI	Electrospray ionisation
EV	Extracellular vesicles
FA	Formic acid
FASP	Filter aided sample preparation
FDR	False discovery rate
GBM	Glioblastoma Multiforme
GO	Gene ontology
HCD	Higher-energy C-trap dissociation
HPLC	High performance liquid chromatography
IAA	Iodoacetamide
iRT	Indexed retention time standard
KEGG	Kyoto Encyclopaedia of Genes and Genomes
L	Litre
LC	Liquid chromatography
LFQ	Label free quantification
m/z	Mass to charge ratio
MEN	Meningioma
μ	micro-
M	milli-
Min	Minutes
MGMT	O ⁶ -methylguanine-DNA methyltransferase
MRI	Magnetic Resonance Imaging
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS1	Mass spectrum of intact analytes
MS2	Mass spectrum of analyte fragments
%	Percent
PI3K	Phosphoinositide 3-kinases
PTM	Post translational modification
PQPs	Peptide query parameters
RNA	Ribonucleic acid

rpm	rotations per minute
sec	seconds
SILAC	stable isotope labelling with amino acids in culture
SDC	Sodium deoxycholate
SWATH	Sequential window acquisition of all theoretical mass spectra
TEAB	Triethylammonium bicarbonate
UPLC	ultra-high-performance liquid chromatography
WHO	World health organisation

1. Introduction

1.1 Cancers

Cancer is a major public health problem worldwide killing millions of people across the globe every year with varying sources of carcinogens. It is a highly complex disease without a unified explanation for its aetiology so far. It was estimated that about 1.7 million people in the United States will be newly diagnosed in 2018 with more than 600,000 deaths; charting as the second leading cause of death in the United States[1]. It is typically characterized by the genomic instability of somatic cells[2], but the development of targeted medicines, novel drug delivery technologies and clinical research data offer a glimpse towards a cancer-free future. Cancers can be classified by site of origin, tissue type, grade or stage. Tumours on the other hand, although commonly interchangeably termed as cancers, are just a mass. It could be benign/non-malignant or malignant/cancerous and either of primary origin or metastasis (secondary tumours). Benign tumours are usually localised, and do not spread to other parts of the body and generally respond well to treatment. However, if left untreated, some benign tumours can grow large and lead to serious diseases because of their size. They can also mimic malignant tumours, hence are sometimes treated. Malignant tumours are cancerous growths which are often resistant to treatment, may spread to other body parts and they sometimes recur after removal.

1.2 Intracranial tumours

The human brain is not only one of the most important organs in the human body; it is also the most complex. It is the largest brain of all vertebrates relative to body size, weighing about 3.3 lbs. (1.5 kg) and makes up about 2% of a human's body weight[3]. The brain contains above 80 billion nerve cells (neurons) called the "grey matter" and billions of nerve fibres (axons and dendrites) known as the "white matter"[4-6]. The cerebrum makes up about 82% of the brain's weight [6, 7]. The brain has three main parts: the cerebrum which performs the higher functions like interpreting touch, emotion, vision, hearing and speech; the cerebellum which coordinates muscle movements, maintain posture, and balance; and the brainstem which performs many automatic functions such as breathing, heart rate, body temperature, wake and sleep cycles, digestion, sneezing, coughing, vomiting, and swallowing. The cerebrum is divided in two parts; the right and left hemispheres which are joined together by fibres called corpus callosum. These hemispheres are further divided into lobes known as the frontal, parietal, occipital and temporal lobes. A schematic of the human brain is shown in figure 1.

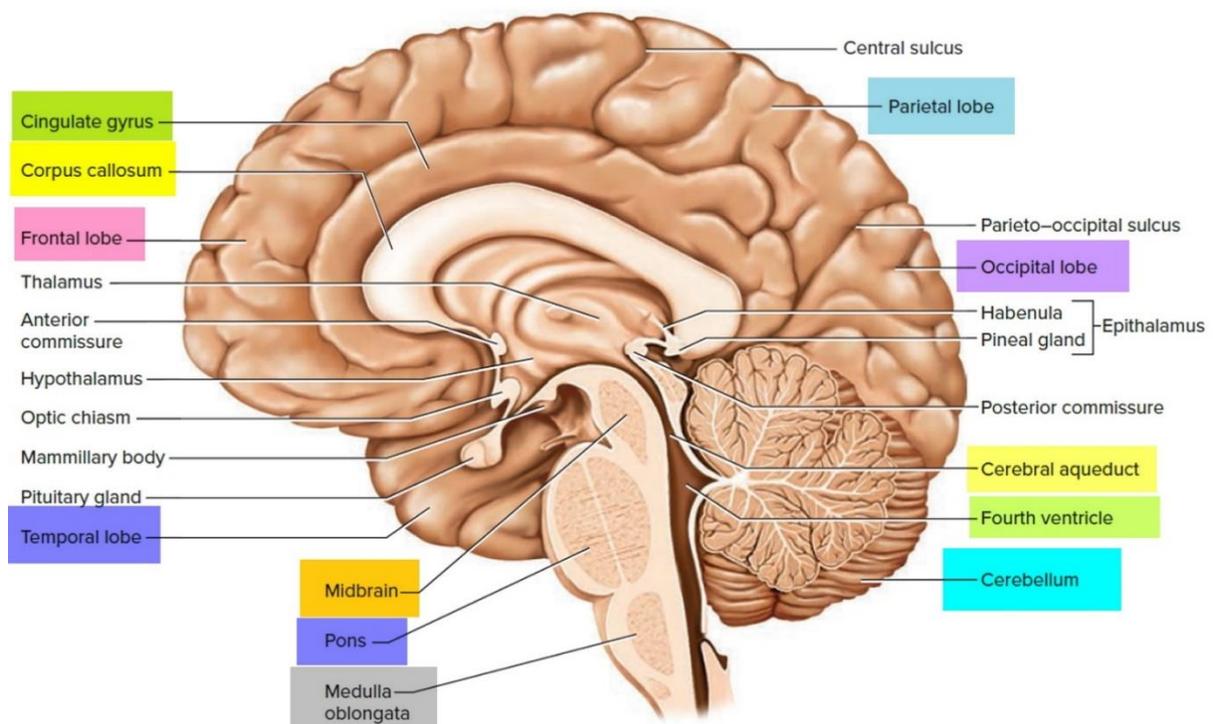


Fig 1: An annotated pictorial of the human brain, illustrating the different lobes and parts[8].

The brain has a plethora of functions, these include; receiving and processing sensory information, initiating and controlling movement, and executing cognitive (thought) processes, thereby enabling its owner to interact with other beings and objects, to flexibly adapt to ever changing situations and lots more[7, 9]. Brain malfunctions are quite known to many, with the common ones being; Alzheimer’s, dementias, brain tumours, epilepsy and Parkinson’s.

Intracranial tumours generally consist of tumours of the brain, cranial nerves, cranial meninges and pituitary and pineal glands, while brain tumours refer to tumours infiltrating the brain parenchyma. Throughout this thesis, intracranial tumour has been at times replaced by the term brain tumour to facilitate understanding as the term “intracranial tumour” is not in everyday use by clinicians. Brain tumours are typically not among the most common tumours in adults, but they are one of the deadliest. In children, brain and spinal cord tumours are the second most common cancers (second to leukaemia), accounting for 1 in 4 cancers[10]. Primary central nervous system (CNS) tumours are a heterogeneous group of tumours arising from cells within the CNS which can be benign or malignant. Malignant primary brain tumours remain among the most difficult cancers to treat, with a 5-year overall survival no greater than 35%. Treatment of malignant CNS tumours in children is an enormous challenge given that the developing CNS is highly susceptible to damage from treatment by methods such as craniospinal irradiation and/or cytotoxic chemotherapy. Even after successful treatment, these kids are still at risk for several physical and cognitive problems that may lower their

health-related quality of life [11]. Figure 2 depicts the classification of brain tumours adapted from *Shergalis et al; 2018*, as reported in 2016 by the Central Brain Tumour Registry of the United States [12, 13].

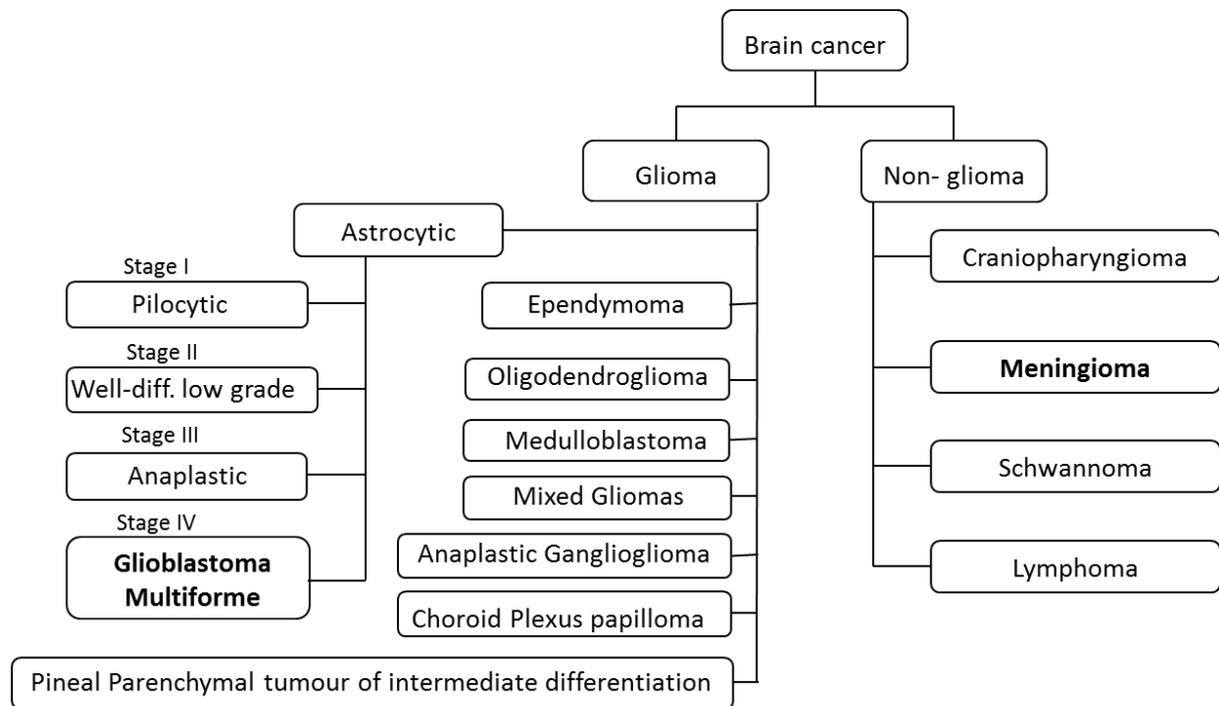


Fig 2: Classification of brain tumours as reported from the Central Brain Tumour Registry of the United States, adapted from *Shergalis et al; 2018* [12, 13].

The causes of brain tumours remain largely unknown[14], ranging from environmental risk factors such as exposure to certain chemicals or ionizing radiation to hereditary cancer predispositions causing embryonal cancers[15]. Therefore, research on uncovering subtle differences within disease categories, aiding early diagnosis and proper treatment will improve clinical management and ultimate long-term outcomes of these patients. The most common example of primary malignant brain tumour is a glioma and meningioma for primary benign brain tumours. Figure 3 indicates the distribution of all primary tumours by brain and other CNS histologies.

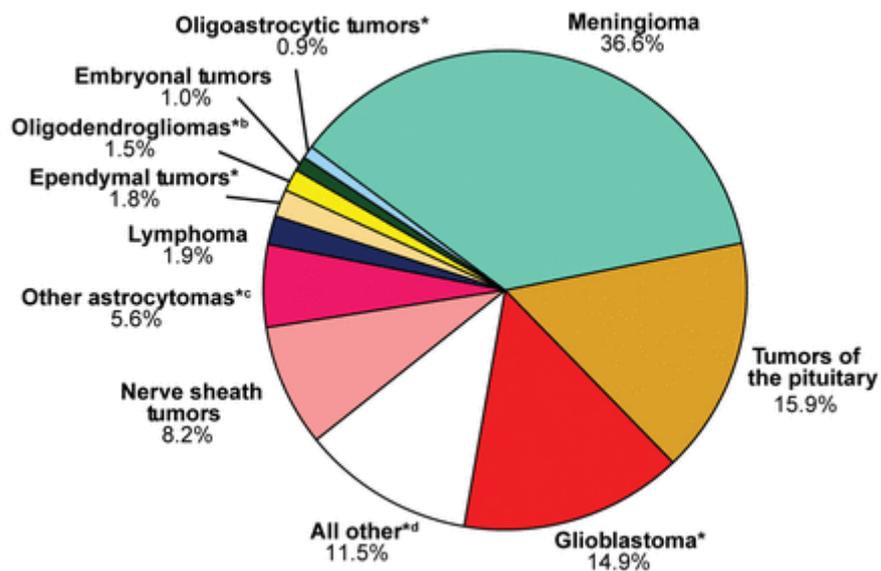


Fig 3: Distribution of All Primary Brain and Other CNS Tumours by CBTRUS Histology Groupings and Histology (N=368,117), CBTRUS Statistical Report: NPCR and SEER, 2009-2013[13].

1.2.1 Glioblastoma Multiforme (GBM)

The most common brain tumours especially in children are gliomas and they typically originate in the glial tissue. Examples of gliomas include: astrocytomas which arise from small, star-shaped cells called astrocytes; oligodendrogliomas arising in myelin producing cells; and ependymomas which usually develop in the lining of the ventricles [12, 13]. Sixty-one percent of all primary brain and CNS gliomas occur in the four lobes of the brain: frontal (25.3%), temporal (19.6%), parietal (12.7%), and occipital (3.3%)[16]. Astrocytomas account for three quarters of all gliomas and Glioblastoma multiforme are a grade IV astrocytoma. GBM is the most malignant primary central nervous system tumour, presenting some of the greatest challenges in cancer patient management worldwide. Despite notable recent achievements in oncology research, patients diagnosed with GBM have a survival rate between 15 and 23 months and a 5-year survival less than 6%, which is the lowest long-term survival rate of malignant brain tumours [13]. Patients suffering from GBM usually have progressive focal neurologic deficits, headaches, and seizures and the tumours are mostly found in the frontal lobes of the supratentorial compartments [17]. Scientific efforts have been made to relate specific characteristics of GBM with environmental or occupational exposure to no avail. However, Ionizing radiation has been linked to an increased risk of glioma development, or radiation-induced secondary GBM after therapeutic radiation as a result of other diseased states [18, 19]. Despite its highly invasive properties, GBM, like most other malignant CNS tumours, does not metastasize outside the CNS. The diagnosis and treatment of GBM still remains a challenge due to its

high heterogeneity. Many studies had detailed the overexpression of epidermal growth factor receptor (EGFR) and mutations of EGFR gene as one of the most characteristic features of glioblastoma, leading to increased cell proliferation through the receptor tyrosine kinase/Ras/PI3K/AKT signalling pathway [20, 21]. It was proposed as one of the potential therapeutic targets especially for primary GBM, unfortunately, several clinical trials with EGFR inhibitors failed hence not fulfilling this dream [22, 23]. Reasons such as intra-tumoural diversity, poor blood brain barrier permeability and the cell-biological complexity of this target have been linked to the clinical trial failure [23, 24]. The Cancer Genome Atlas (TCGA) Research Network gave new insights to GBM genomic characterisation, identifying three core pathways with biologically relevant alterations: protein p53 pathway, the receptor tyrosine kinase/Ras/phosphoinositide 3-kinase signalling pathway, and the retinoblastoma pathway [25]. Up to recently, routine GBM diagnosis still relied on magnetic resonance imaging (MRI) and a combination of Surgery and temozolomide as chemotherapy with radiotherapy has become the cornerstone in the initial treatment [12]. However, the TCGA project was able to recognise four distinct molecular subtypes of GBM; establish a new subtype of GBM that affects younger adults; identify possible mechanisms involving gene mutations that can cause some resistance to therapy after treatment with the standard temozolomide chemotherapy [25, 26]. EGFRvIII is a clinical diagnostic marker whereas MGMT promoter methylation is a prognostic and predictive marker [27]. In 2016, the world health organisation (WHO) Classification of Tumours of the Central Nervous System replaced the traditional histology-based glioma diagnostics with an integrated histological and molecular classification system that enables more-precise tumour categorization [28]. This implies that proteomic studies in understanding the molecular variation of the disease are tremendously vital in improving survival rates. Current understanding of the molecular characteristics of this disease has demonstrated that there are unlikely to be single genetic or cellular events that can be effectively targeted for all patients. Hence the shift towards individualizing treatment for each patient's tumour genotype or proteomic type (personalised medicine)

1.2.2 Meningioma (MEN)

The term meningioma was first described by the eminent neurosurgeon Harvey Cushing, referring to tumours arising from the meningeal linings of the brain and spinal cord[29]. According to the CBTRUS Statistical Report of 2013 as indicated in figure 3, meningiomas make up 36.6% of all primary brain and other CNS tumours [13]. Its incidence increases with age and it is twice as common in females than in males with females between the ages of 35 to 54 being the most affected[30]. They are mostly benign and slow growing, however; there exist more aggressive though uncommon subtypes. The WHO classification of these tumours based mainly on their histo/

cyto-morphological characteristics, groups them into 3 main tiers. About 80% are WHO grade I and are benign (benign meningioma), 17% are WHO grade II and are referred to as the atypical meningioma while only 2 % are WHO grade III, known as anaplastic meningioma [31-33]. Although most grade I tumours are considered slow growing and benign, their location can cause serious morbidity and mortality. Many meningiomas are located at the skull base or other high-risk areas that do not facilitate easy surgical access for resection or even biopsy [34]. Early diagnosis of meningioma is essential as the symptoms are often times confused with general aging symptoms by both patients and doctors due to the slow growth rate of the tumour. The state of the art treatment methods generally range from wait and watch, surgery, surgery plus radiation therapy to varying forms of radiation therapy alone[35]. For most patients however, gross total resection (GTR) remains the benchmark. Even after MRI based identification, differentiating between Grade I and higher-grade meningioma is also quite critical as this not only aids in better predicting the biological behaviour but it also gives insight to the treatment method to be applied and an overview on the recurrence possibilities of the tumour after treatment. Often times, the goal of surgery is to safely resect the tumour and obtain a diagnosis and pathological grade and subsequent radiation therapy depends on the extent of resection and pathological characteristics [36]. Another challenge is the recurrence probability of the tumour. Grade I meningioma recurrence prediction can be done with the established Simpson system which grades the extent of meningioma resection [37, 38].

Most Grade I meningiomas harbour some mutations with neurofibromatosis 2 (NF2) being very common; however, they are less than those of higher-grade meningioma [39, 40]. Genetic and proteomic markers are being studied to better diagnostic and treatment purposes as most therapeutic drugs are proteins [41]. Understanding molecular characteristics helps in the further sub classification of the tumours, bettering treatment methods and improving patient management. Hence proteomics analysis of brain tissues is an essential part of neuroscience research [42].

1.3 Biomarker discovery in cancer diagnosis and prognosis

Biomarkers have been defined as “a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention”[43]. The WHO definition of a biomarker is “any substance, structure, or process that can be measured in the body or its products and influence or predict the incidence of outcome or disease” [44, 45]. They serve as a corner stone in understanding the development of diseases like cancer, aiding from the process of diagnosis, prognosis assessment and prediction related to patient response to treatment, and in some cases disease recurrence. Biomarkers can generally be classified into six main groups: biomarkers of risk; diagnostic or trait biomarkers; state

or acuity biomarkers; stage biomarkers; treatment response biomarkers; and prognostic biomarkers[46]. Scientific research in the past decades has focused on understanding the molecular patterns behind diseases to help improve early diagnosis and discover better therapies especially for chronic diseases. For the first time in 2016, molecular parameters were used by the WHO together with histology in the classification of tumours of the CNS [32]. Initially, the idea of biomarkers revolved around individual genes, proteins or metabolites (molecular biomarkers) linked to a disease. However, a disease is rarely a consequence of an abnormality in a single gene, due to the functional interdependencies between the molecular components in a human cell[47].

Biomarkers have been identified for the prognosis of primary Glioblastoma although they are not sensitive to majority of the population. An upregulation of O⁶-methylguanine-DNA methyltransferase (MGMT) has been reported in about 64% of primary GBM[48]. Its identification has also been reported to be dependent on the site of tumour collection (concentric layers) and this is used as a prognosis and predictive marker [49]. Another prominent prognosis primary GBM marker is the amplification/mutation of the EGFR gene which has been reported in about 36–60% of primary GBM samples[50]. Other relevant biomarkers include Isocitrate dehydrogenase (IDH) [51], Neurofibromin (NF1) [52] and Vascular endothelial growth factor (VEGF) [52].

Unlike GBM, molecular parameters have not yet been incorporated into the WHO classification of meningiomas. They still largely depend on histological parameters to differentiate between the subtypes. However, certain pathways have been associated with MEN and potential biomarkers have been proposed. MEN have been associated with Neurofibromatosis type 2, hence a mutation in the Neurofibromin (NF2) has been recorded in 45% to 58% of patients [53]. However, the NF2 mutation occurs similarly between the 3 grades of meningioma, suggesting that it could be related to tumour initiation and less likely tumour progression [54].

Though the idea of a single marker is highly clinically appealing because of simplicity and low cost, it may not capture the variability of the disease across a population, hence lowering the accuracy. Recently, the theory of a single molecular marker is being debunked as it does not reflect the perturbations of the complex intracellular network. Hundreds to thousands of biomarkers are reported annually but only a few make it through to the clinics[55]. Research has now shifted to focus on identifying multi-marker panels for cancers and other diseases to salvage these limitations [56-58].

1.4 Cancer liquid biopsy and Extracellular vesicles

Biopsies have been a part of the medical terminology for over a century. The term “Biopsy” was first introduced in 1879 by Ernest Besnier [59]. It has been defined as a “complete or partial removal of a lesion for laboratory examination to aid definitive diagnosis”[60]. According to the National Health Service (NHS) of the United Kingdom, the term biopsy is often used to refer to both the act of taking the sample and the sample itself [61]. While most structural imaging techniques like magnetic resonance imaging (MRI), computed tomography (CT), ultrasound and X-ray scans are great to identify areas of concern in the medical field; they cannot give sufficient insights on the differential expression of the biomolecular species present in the area of interest. In the field of oncology, biopsies are used to investigate suspicious tissue masses, differentiate between benign and malignant tumours, understand the molecular characteristics of the tumour to aid device a treatment plan and predict its recurrence probability.

Traditionally, physicians and researchers have regarded tissue biopsy as the gold standard for providing data that produces positive health outcomes among patients with a variety of cancers [62]. For a subset of intracranial tumours like meningiomas, tumour resection can serve as both a diagnostic and a curative measure. For others like most gliomas, stereotactic or open biopsy is often performed as they are considered less invasive and resection may follow later depending on other factors. Tumours that are considered particularly radio- or chemo-sensitive, such as germ-cell tumours or lymphomas, are also excellent candidates for stereotactic biopsy. Those who favour stereotactic biopsy point out its low risk, diagnostic accuracy, and minimally invasive nature [63, 64]. However, other studies have indicated the shortcomings of this biopsy technique, pointing to the possibility of it increasing mortality and morbidity in patients with gliomas [65]. Other limitations to stereotactic biopsies include: access to limited amounts of tumour tissue for pathologic and molecular diagnoses; tissue biopsy is also subject to sampling bias; and tissue from a single tumour location may fail to capture intra-tumour heterogeneity.

In efforts to salvage these limitations, research is now focusing on the sampling of the “liquid biome”. The “liquid biome” refers to biological fluids (biofluids) including blood, urine, saliva, breast milk, bronchial lavage fluid, synovial fluid, nasal lavage fluid, semen and cerebrospinal fluid (CSF)[66]. Biofluids are believed to reflect the ensemble of tissues and cells present within a patient[67]. They may contain a diverse array of extracellular vesicles (EVs), circulating tumour cells (CTCs), circulating tumour DNA (ctDNA) and RNA (ctRNA), cell free DNA (cfDNA), microRNA (miRNA), fragmented peptides and even intact proteins alongside other biomolecules [66]. The possibility to use CTCs and ctDNA in order to monitor disease progression has been largely studied as they had

been identified in patients' blood after death as early as the 1800s [68-72]. Evs, which are one of the contents of biofluids, shed from individual cells are highly molecularly complex and heterogenous. They serve as vehicles for intracellular communication and molecular exchange as they facilitate the "packaging and shipment" process between cells as explicitly described in figure 4 [73]. Since EVs are released essentially by all cells and their role particularly in the nervous system has represented an area of considerable interest [74, 75], studies are now geared towards understanding the impact of EV release and trafficking on the pathogenesis of diseases such as brain tumours [73, 76]. The ultimate goal is the possibility of early cancer detection especially before the onset of metastasis. A recent study published in Science details a multi-analyst blood test known as CancerSEEK which not only detects cancer but indicates the location for eight common solid tumour types [77]. Unfortunately, brain tumours did not make the cut. Therefore, exploring the characteristics, composition of EVs and their location in biofluids could be revolutionary in the world of tumour biopsy; especially brain tumours, owing to their location in the brain and their degree of aggressiveness. This is a very promising domain where clinically relevant biomarkers can be discovered from the less invasive liquid biopsy techniques which will give a "snapshot" of a patient's proteome at any given time.

Modes of intercellular communication mediated by extracellular vesicles

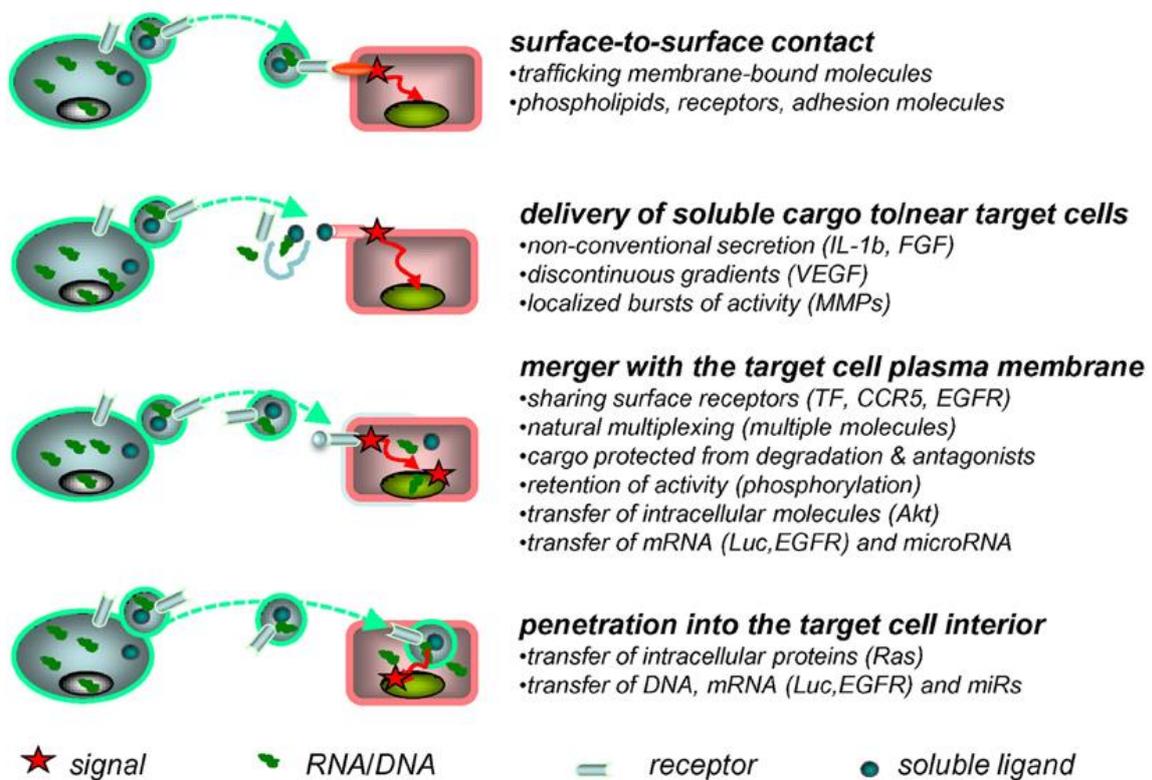


Fig 4: Extracellular vesicles as mediators of intercellular communication. Exchange of molecular information between cells mediated by EVs in four ways: surface receptors of EVs interacting directly with counter-receptors on the surface of a target cell; the latter may also come in contact with the bioactive inner cargo of EVs upon their pericellular rupture; EVs may also merge with the plasma membrane of the target cell; or penetrate into its interior via endocytosis, or other processes, to release their content of proteins and nucleic acids into the intracellular compartments [73].

1.5 Mass spectrometry based clinical proteomics

The “proteome” which was a term coined as a linguistic equivalent to the term “genome” refers to the complete set of expressed proteins and those modified after expression, by the entire genome in the lifetime of a cell [78, 79]. Proteomics in general aims to identify, characterise and quantify expressed proteins from biological systems in order to understand gene and cellular function[80]. It aids in deciphering protein-protein interactions, protein abundance, localisation and identification of post-translational modifications (PTM) and differences in expression [81]. Each human gene may encode for several proteins and these proteins may possess several post-translational modifications, generating varying protein species and proteoforms. These proteins generally function as part of an intricate network of biological pathways and never by themselves[82]. This implies that genomics alone cannot possibly account for the plethora of diseased states which exist within the human population. The genome can be considered more or less as stable when compared to the proteome

which is dynamic and ever-changing, depending on many factors. Proteins are the most diverse group of biologically important substances and are often considered to be the central compound necessary for life[83]. With this diversity comes complexity and challenges in analysing them. Traditionally, Edman degradation was used to sequence proteins by chemically cleaving amino acids one at a time from the amino terminus [84]. Although it is still a method of choice for sequencing unknown proteins/peptides, high throughput methods have been developed with the advances in mass spectrometry (MS) which has been the driving force behind progress in proteomics. MS-based proteomics is now the most commonly used technology for identifying as many proteins in a single experiment as possible either from gel spots/bands or in solution. A basic MS analysis workflow generally includes; sample collection and storage, protein and peptide separation, protein and peptide identification and quantification, statistical data analysis and biological interpretation. Modification of the workflow which may include enrichment techniques is usually done to enable unbiased proteome analysis which is necessary for identification and quantification especially for samples of clinical relevance.

Sample preparation is also one of the most crucial steps in proteome analysis, owing to the broad platform of sample origins. In comparative proteomics, very minute changes between control and experimental samples are studied. This implies an efficient experimental design and careful sample preparation is very vital for obtaining accurate and reproducible results. There are three main MS approaches which are often used for protein analysis: top-down approach where intact proteins are ionised and fragmented[85]; bottom-up approach which is very often used whereby, proteins are first enzymatically digested to peptides prior to ionisation and fragmentation[86]; and the not so popular middle-down approach which aims at generating large peptides with size approaching those of small intact proteins that are readily analysed in top-down[87]. This study is based on the bottom up approach hence that is the approach which will be discussed in detail. In bottom-up proteomics, proteins are enzymatically cleaved using proteases generating peptides which are stable under the conditions required for MS. The serine protease trypsin is often the protease of choice due to its exceptional substrate specificity, high proteolytic activity, and it is very stable under a wide variety of conditions[88]. It cleaves proteins at the C-terminal side of arginine and lysine residues. Although it was initially known to cleave next to arginine or lysine but not before proline (“Keil rule”), some studies have proven otherwise[89]. Other proteases used include Lys-C and Glu-C which cleave at the C-terminus of lysine and glutamate residues respectively and Asp-N which cleaves at the N-terminus of aspartate residues. Prior to MS measurements, the digestion peptides are often separated with chromatographic tools. The chromatography method of choice depends on the peptides of interest and also on the ionisation method to be used downstream. Most common are

liquid chromatography (LC) and gas chromatography (GC). With LC, peptides are loaded onto an ultra-high-performance liquid chromatography (UHPLC) column which often contains a reversed-phase stationary phase that resolves peptides based on hydrophobicity. These Peptides are eluted using an increasing organic solvent gradient with hydrophilic peptides eluting at the start of the gradient. LC systems can be coupled directly to the mass spectrometer, where the eluted peptides are ionised, and mass spectra recorded

1.5.1 Peptide Ionisation & fragmentation

Mass spectrometers are generally made up of an ion source, a mass analyser and a detector which are operated under high vacuum conditions. The history of the mass analyser dates back to the research years of Sir Joseph John (J. J.) Thomson in the early 19th century and its science today has greatly evolved[90, 91]. Several techniques have been discovered over the decades for ion production in the MS with each having their pros and cons. Some of these include; gas phase ionisation techniques like electron impact and chemical ionisation, ionisation from liquid phase like electrospray and atmospheric pressure ionisation and ionisation in solid phase with MALDI (Matrix-Assisted Laser Desorption/ionisation). The ionisation technique of choice in this study was electrospray ionisation (ESI). ESI-MS has been widely explored for the ionisation of large biomolecules. It dates back to the pioneering efforts of Dole and his colleagues in 1968[92], however only about 20 years later was it used as an interface for mass spectrometry by Fenn[93]. And Fenn was later awarded a Nobel Prize for Chemistry in 2002. ESI is a soft ionisation method which ionises biomolecules by multiple charging with very little residual energy retained by the analyte and generally no fragmentation occurs during the process. The transfer of ionic species from solution into the gas phase by ESI involves three steps: (1) dispersal of a fine spray of charge droplets, followed by (2) solvent evaporation and (3) ion ejection from the highly charged droplets (Figure 5) tube, which is maintained at a high voltage relative to the wall of the surrounding chamber[94, 95]. A nebulizing gas is also used which helps to direct the spray emerging from the capillary tip towards the mass spectrometer.

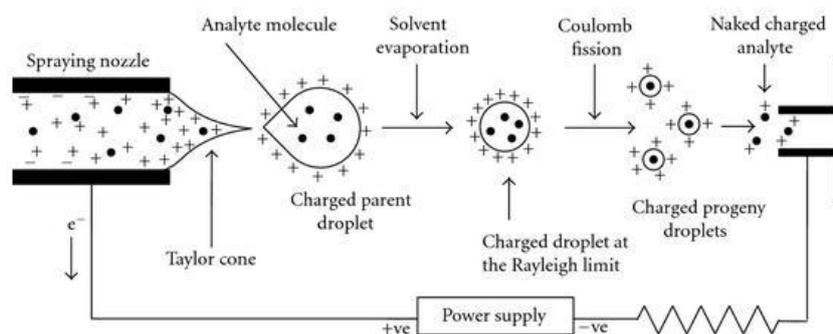


Fig 5: A schematic representation of the ESI process[96].

In tandem mass spectrometry experiments, after ionisation, the MS spectra of the ions are recorded, the ions fragmented and MS2 spectra of the fragment ions recorded. Several fragmentation methods can be used to generate tandem mass spectra, depending on the molecular size of the analyte, ionisation states required and the software available for data analysis. These methods are characterised by how energy is transferred to the precursor ion, the amount of energy transferred, and how the transferred energy is internally distributed [97]. Some of these fragmentation methods are; electron capture dissociation, electron transfer dissociation, laser induced dissociation, infrared multiphoton dissociation, collision induced dissociation (CID) and Higher-energy C-trap dissociation (HCD). CID is a slow heating technique achieved by colliding the ions with neutral species like helium or argon, cleaving the weakest bonds and generating typically b and y ions. The collision energy is then transferred to internal energy of the ion, resulting in bond breakage. HCD on the other hand is a method exclusive to orbitrap instruments. It is a CID variation that uses a higher RF voltage to retain fragment ions in the C-trap. The HCD cell is used to fragment the ions, after which they are accelerated into, cooled down and stored inside of the C-trap. Ions are then injected into and separated inside the Orbitrap based on their rotational frequency differences [98]. These fragmentation mechanisms occur by cleavage of the backbone amide bonds to produce a- and x-, b- and y- or c- and z- type ions. The precursor ions are fragmented at the amide bonds between amino acids and the product ions that retain the charge at the N-terminus are the a-, b- and c-ions and those that retain the charge at the C-terminus are x-, y- and z-ions (Figure 6). CID and HCD fragmentation produce b- and y-ions

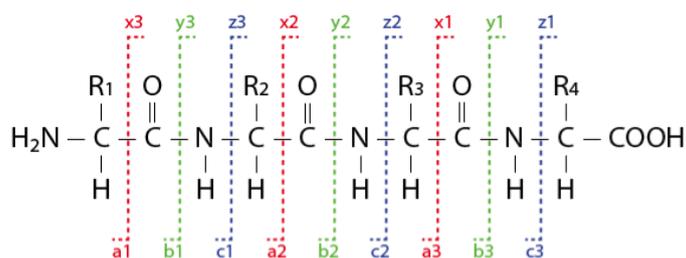


Fig 6: Nomenclature for the dissociation products of fragmented peptides. The use of low energy CID or HCD predominantly generates b-ions and y-ions, spectra obtained by using ETD contain predominantly c-ions and z-ions and a-ions and x-ions[99].

1.5.2 MS Quantification methods

Protein and peptide quantification represent an important extension to identification proteomics. Quantification information about proteins is necessary in understanding the perturbations occurring when there is a deviation from the normal physiological state of any living organism. Quantification can either be absolute; where the absolute amount of proteins in a sample are measured (copy number per cell) or relative; where the relative change in protein amounts between different states are measured. Protein quantification methods generally involve labelling or label free (LFQ) approaches. Labelling techniques range from metabolic incorporation of stable isotopes such as stable isotope labelling with amino acids in culture (SILAC) or chemical labelling of the protein mixture like isotope coded affinity tag (ICAT). Label free approaches are mostly based on spectral counting peptide chromatographic peak intensity measurements[100]. The standard for complex samples in the last years has been LC-MS/MS in data-dependent acquisition (DDA) mode and multiple reaction monitoring (MRM) for targeted proteomics with precise quantification[101]. Recently, improvements in mass spectrometer design and bioinformatics algorithms have resulted in the rediscovery and development of another sampling method: data-independent acquisition (DIA). In this study, both DDA and SWATH-DIA modes of data acquisition were employed in the bottom-up proteomics approach and a comparison generated.

In the DDA approach, peptides eluting from the LC column are ionised, scanned, and a subset (precursor ions) which most commonly are the most abundant are selected for further fragmentation to generate MS/MS data from the fragment ions. The proteins in the sample are then generated using bioinformatics tools. Over the course of the entire LC run, the subset of precursors selected in each duty cycle is therefore dependent upon the precise make-up of the mixture of compounds entering the instrument at every moment in time. It also relies on the probability that a precursor ion will be acquired at the apex of its chromatographic elution profile, which is improbable[102]. This therefore adversely affects the data quality for purposes of both identification and quantification. A dynamic exclusion window is usually implemented to ensure peptides which

have recently been analysed are not constantly being re-fragmented when newer targets are available in successive dissociation events. This approach diminishes reproducibility, prevents the measurement of low-abundance peptides and some important compounds might go unidentified as they never make the cut[101]. To circumvent these shortcomings on dynamic range and increase quantification accuracy, data independent acquisition (DIA) methods are being developed. Unlike DDA, predetermined m/z ranges are interrogated either: (1) by fragmenting all ions entering the mass spectrometer at a single point in chromatographic time (broadband DIA); or (2) by dividing the full m/z range into smaller m/z isolation windows that are each independently and consecutively analysed[103]. Many DIA techniques have been developed relying on these two principles and the m/z range of the various methods affects the inherent presence of multiplexed spectra, hence a difference in data analysis. One technique which divides the full m/z range into smaller m/z isolation windows is SWATH (sequential window acquisition of all theoretical mass spectra). In SWATH-MS measurements, all ionized peptides of a given sample that fall within a specified mass range are fragmented in a systematic and unbiased fashion using precursor isolation windows[104]. This method was reported by *Gillet et al*; in 2012 where they generated time-resolved fragment ion spectra for 32 consecutive 25-Da precursor isolation windows [103, 105]. Fragment spectral libraries are also created to facilitate data analysis. They usually consist of information on the m/z value, intensities and standardized retention times of all precursors and their corresponding fragment ions. Figure 7 illustrates the fundamental differences between DDA and DIA modes.

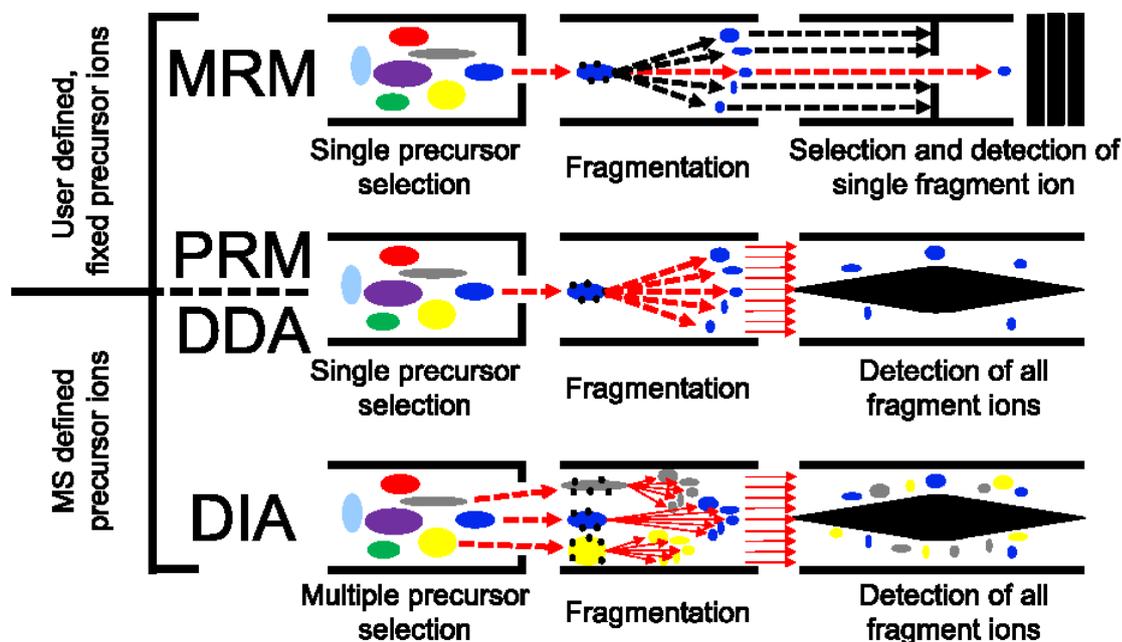


Fig 7: Fundamental differences between the MRM, DDA and DIA modes.

2. Aim

Although primary brain cancers are typically not among the most common cancers, they encompass one of the deadliest. Primary malignant brain tumours such as gliomas are the most prevalent type of adult brain tumour accounting for over 78% of malignant brain tumours, with glioblastoma multiforme (GBM) being the most invasive. Primary benign brain tumours like meningiomas (MEN) are less invasive but however, still possess life threatening characteristics. Prognosis of these tumours is very poor with only 6.4% survival rate up to two years for GBM. Early and non/minimally invasive diagnosis therefore is vital for better prognosis. Evidence is currently mounting for extracellular vesicles (EVs) involvement in multiple facets of cancer biology. They have been thought to aid metastasis and even facilitate cell drug resistance. Due to easy accessibility, they can be used as biomarkers for cancer which will enable the use of liquid biopsies for early detection and personalised cure. Their ability to transfer molecular cargo and be selectively taken up by specific cells is also being exploited for cancer therapy.

Therefore, the overall aim of this study was to explore the possibility of extracellular vesicle proteins as potential biomarkers for early and non/minimally invasive diagnosis of primary brain tumours. To achieve this goal, two primary brain tumours were studied: Glioblastoma multiforme, which is typically malignant and meningioma, which is typically benign. The first part of this study focuses on the proteome of tumour tissue samples, cells and EVs. Eight tumour tissues were surgically obtained from four GBM and four MEN patients. Patient-derived primary cells were then cultured off each patient tumour tissues. From the cell culture supernatant, extracellular vesicles were harvested. The goal is to identify potential cancer EV specific biomarker proteins which can be obtained from cancer liquid biopsy to facilitate early cancer diagnosis. The second part highlights the proteome variation between the patient tissue samples and cells obtained from primary cell culture. The goal is to understand the effects of using cells as a research tool for biomarker discovery. The third part of the study focuses on identifying potential tumour biomarkers by comparing the patient tumour samples to non-tumour brain tissue samples. These non-tumour tissues were epileptic patient brain tissues excised from either white matter or cortex. The differentially regulated proteins in either GBM or MEN tissues with respect to the non-tumour tissues were then further researched to gain insights on their role in cancer diagnosis, progression/metastasis or potential resistance to therapy and to identify potential cancer biomarkers. The last part of the study focuses more on the mass spectrometry technique of choice and highlighting its potentials when compared to the go-to method which is data dependent precursor acquisition for protein label free quantification.

3. Results

The overall aim of this study was to explore the possibility of extracellular vesicle proteins as potential biomarkers for early and non/minimally invasive diagnosis of primary brain tumours. The primary brain tumours studied were Glioblastoma multiforme (GBM) and meningioma (MEN). To identify and quantify these proteins, the mass spectrometry method used throughout this work was the variable data independent acquisition by sequential window acquisition of all theoretical mass spectra (vDIA-SWATH).

The first set of experiments deal with the proteome analysis of tumour tissue samples, cells and exosomes. The proteins identified in all three patient samples are then further studied to discover their potential as biomarkers for possible cancer liquid biopsy. The second part highlights the proteome variation between each patient tissue sample and cells obtained from primary cell culture, the third part compares the patient tumour samples to non-tumour brain tissue samples while the last part of the study highlights the differences in the results obtained from using either DDA or SWATH-DIA mass spectrometry methods. The differentially regulated proteins in either GBM or MEN tissues with respect to the non-tumour tissues were then further researched to gain insights on their role in cancer diagnosis, progression/metastasis or potential resistance to therapy.

3.1 SWATH DIA MS of tumour tissues, primary cell culture and extracellular vesicles of glioblastomas and meningiomas

Brain tumour samples excised from 8 patients with glioblastoma or meningioma were prepared for proteomic analysis (see section 5.2 figure 45). Primary cells were cultured from each patient tumour tissue and harvested. Extracellular vesicles were then spun down from the cell culture supernatant from each patient culture. The different samples in the sample cohort were prepared as described in sections 5.4 to 5.6. The proteins were then identified and quantified on MS2 level using the SWATH tool for DIA based protein quantification.

3.1.1 Spectral Library generation

For post-acquisition deconvolution and processing of SWATH MS data obtained in these experiments, spectral libraries were generated. They usually consist of information on the m/z value, intensities and standardized retention times of all precursors and their corresponding fragment ions. Therefore, these spectral libraries were generated from shotgun proteome analysis of the samples of interest. Commercially available calibration peptides based on the indexed retention time concept, were spiked into the sample prior to MS measurements. The resulting raw files were processed with MaxQuant and the libraries generated with Spectronaut Professional. There exist publicly available spectral libraries which can be obtained from the SWATHAtlas webpage (www.swathatlas.org).

Individual sample specific libraries were generated for tissues, cells and extracellular vesicles (ssTT, ssCC, ssEV). A project specific library containing combined information from all processed tissues, cells and extracellular vesicles was also generated from the existing sample specific libraries. A fifth library known as the Pan-human library was downloaded from SWATHAtlas. This spectral library was generated based on 331 measurements from fractionated samples of different cell lines, tissues and affinity enriched protein samples and is claimed to support the confident detection and quantification of 50.9% of all human proteins annotated by UniProtKB/Swiss-Prot[106]. Figure 8 shows the precursor m/z count from all the libraries individually.

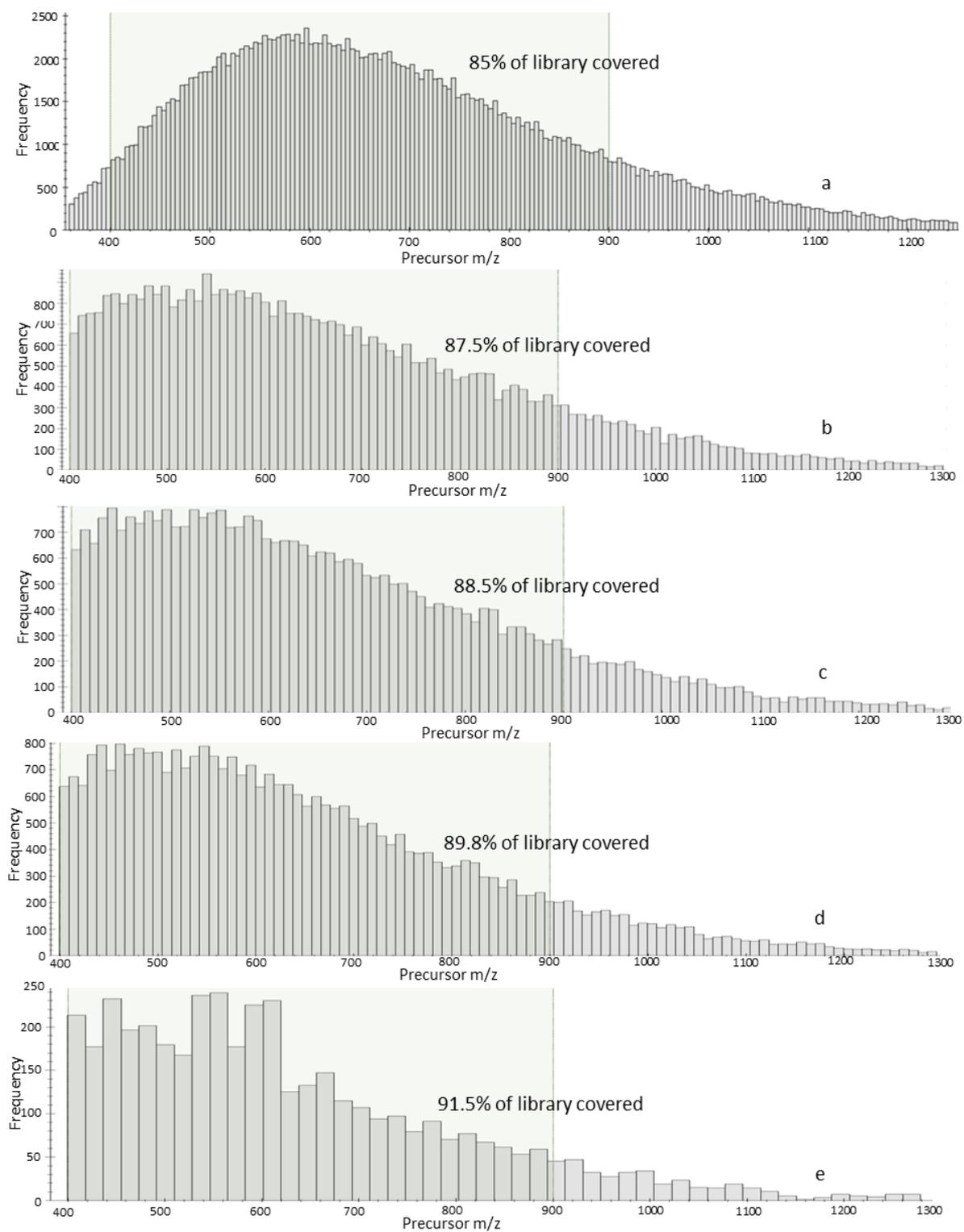


Fig 8: Precursor m/z population of the different libraries with precursor frequency plotted as ordinate and precursor m/z as abscissa. A) Pan-human b) Project specific c) Tissue d) cells e) Extracellular vesicles.

Information on the number of precursors, peptides and proteins available in all the aforementioned libraries are depicted in figure 9 below.

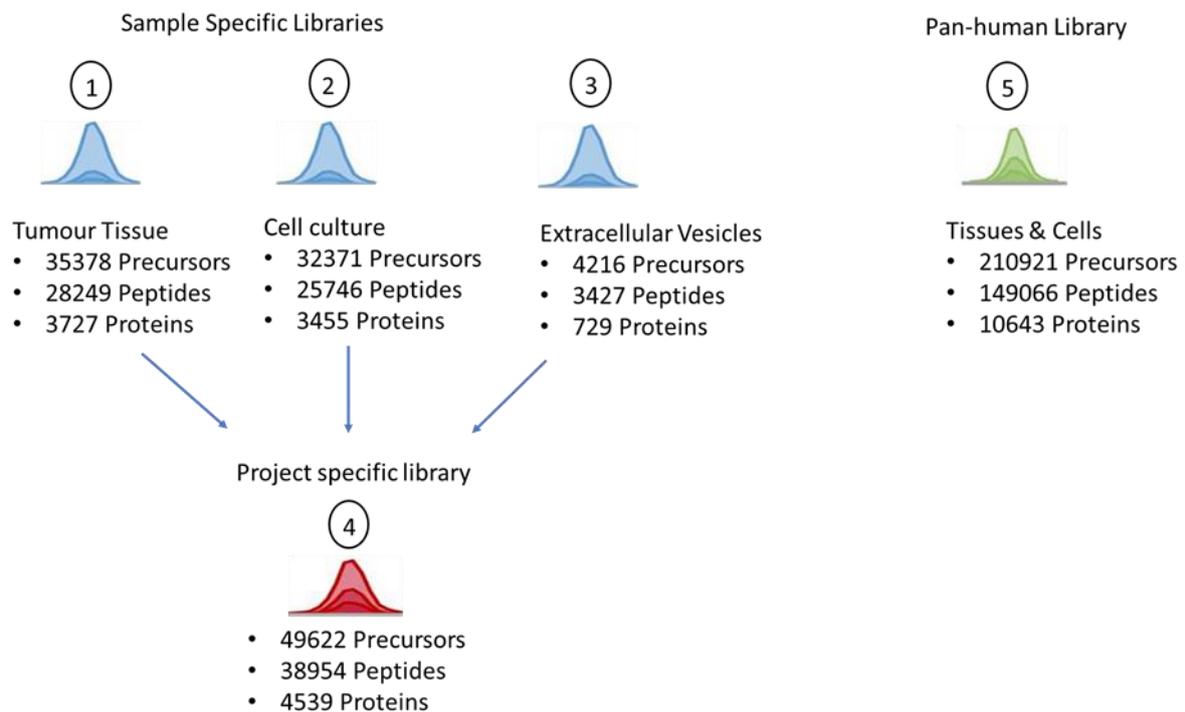


Fig 9: Number of precursors, peptides and proteins available in various spectral libraries. The sample specific libraries (1, 2, 3) and project specific library (4) were generated from sample shotgun runs using the maxquant output file in Spectronaut Pulsar X. The Pan-human library was downloaded from the SWATHAtlas webpage.

3.1.2 Spectral library evaluation

Three different spectral libraries were assessed: sample specific library, project specific library and the publicly available pan-human library. The assessment criterion was based on the number of identified peptides and proteins from each analysed DIA sample. Tumour Tissues, their corresponding cells generated from primary cell culture, and their extracellular vesicles separated from their cell culture supernatant from 4 meningioma patients and glioblastoma patients were analysed. Figures 10 & 11 depict the number of identified proteins and peptides from each sample using the 3 different spectral libraries. For the EV samples, the largest number of peptides and proteins were identified with the project specific library. For the cell culture samples, the project specific and pan-human libraries performed quite similarly however, the sample specific library identified the least number of peptides and proteins. This was also quite similar with the tumour tissue samples.

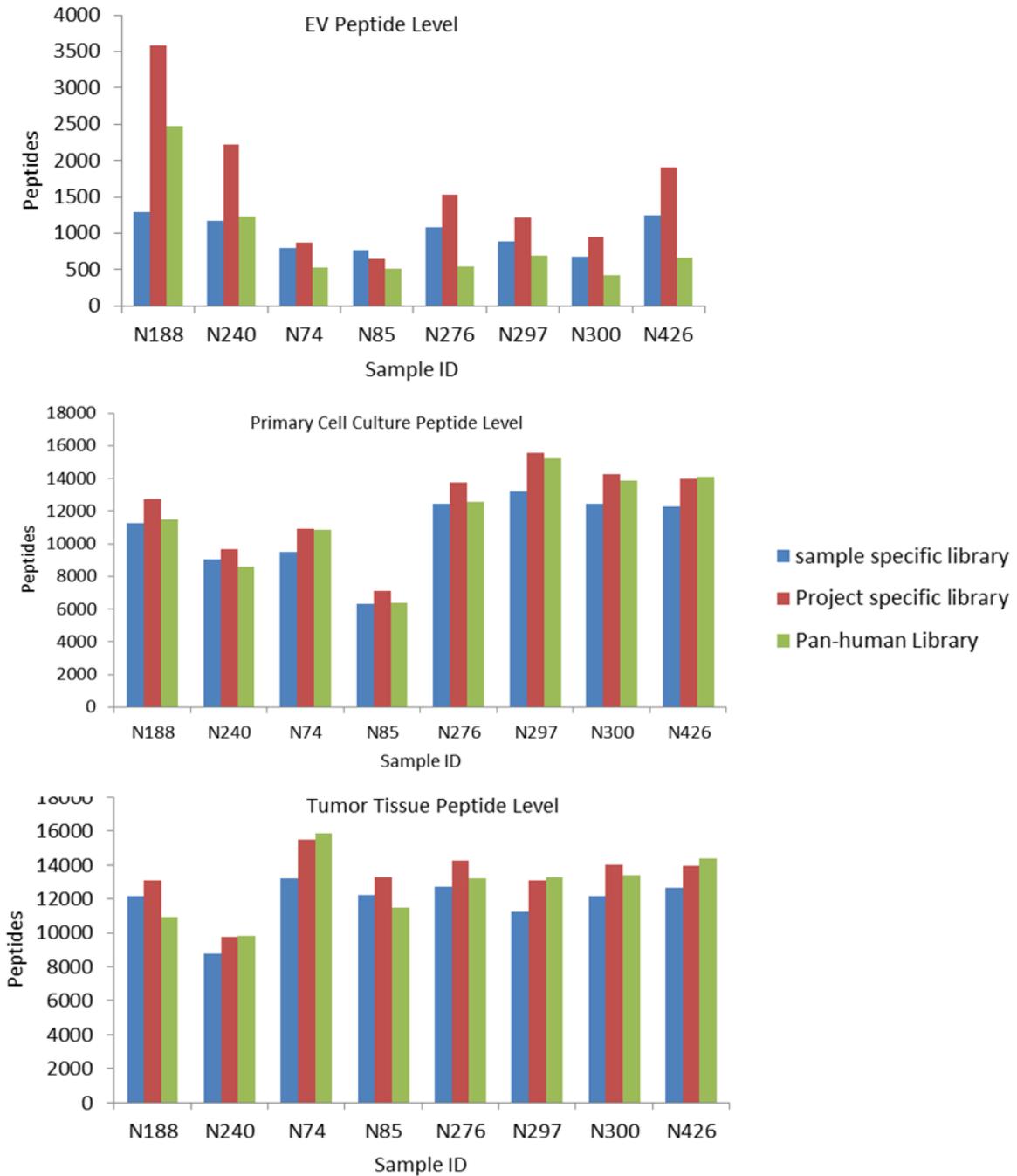


Fig 10: Spectral library influence on SWATH MS post processing. Effects on number of identified peptides. 4 meningioma patient (N188, N240, N74, N85) and 4 glioblastoma patients (N276, N297, N300, N426) tissues, cells and extracellular vesicles respectively were analysed with a DIA method using Q Exactive mass spectrometer.

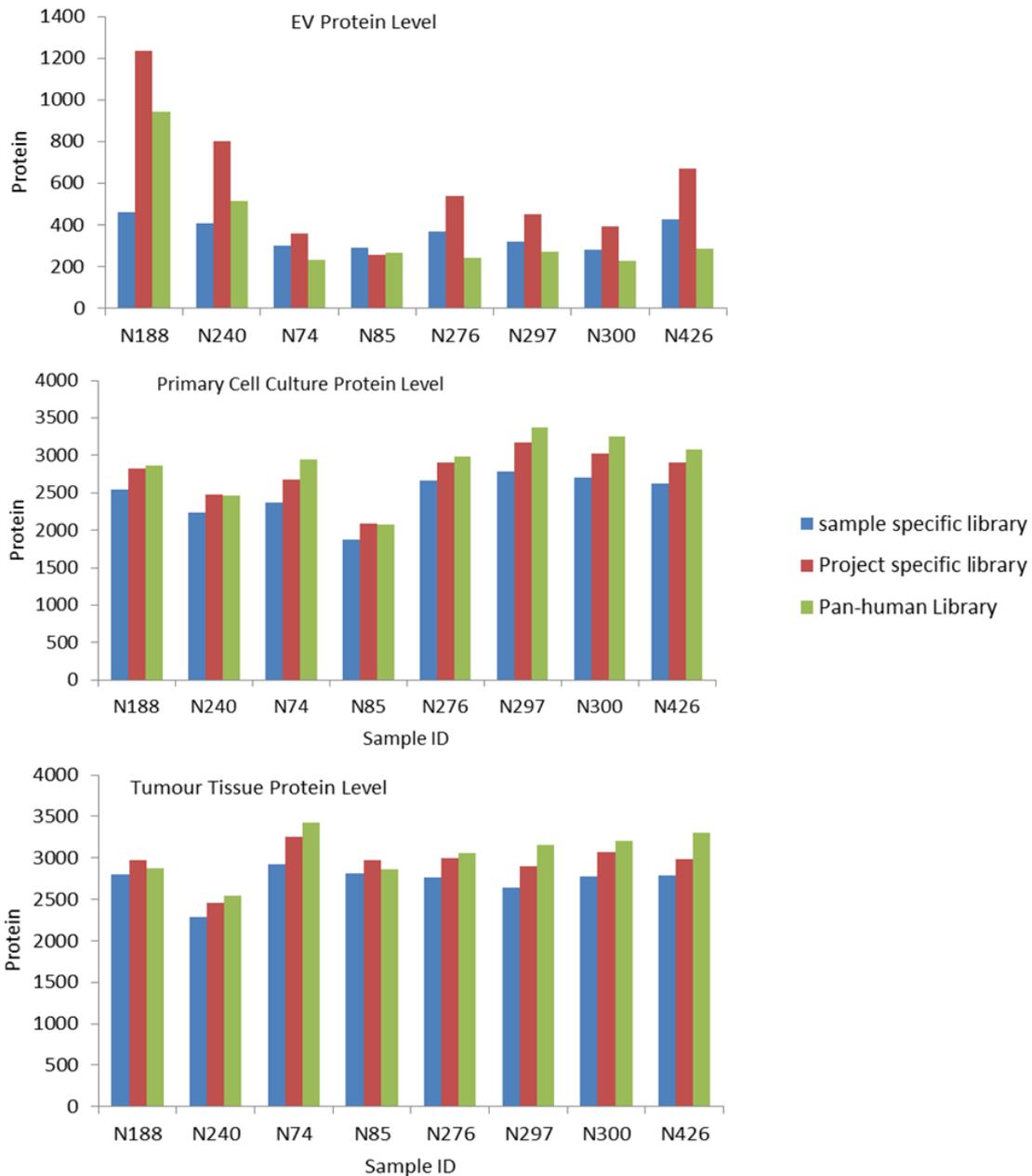


Fig 11: Spectral library influence on SWATH MS post processing. Effects on number of identified proteins. 4 meningioma patient (N188, N240, N74, N85) and 4 glioblastoma patients (N276, N297, N300, N426) tissues, cells and extracellular vesicles respectively were analysed with a DIA method using Q Exactive mass spectrometer.

3.1.3 Data Acquisition

The dynamic window method was used to acquire all data in these experiments. To achieve better specificity in complex matrices, high resolution fragment ion spectra were recorded for all precursors within a user-defined precursor ion window. Smaller Q1 windows were therefore used in dense m/z areas where higher precursors were measured whereas wider Q1 windows were implemented in less dense m/z regions. The Q1 isolation window widths used for all samples are

detailed under Section 5.8 Table 13 while an example of an LC gradient acquired for one sample is shown in supplement figure S1. Within each defined precursor window, the information of the identified precursor ions is detailed. An example of the spectrum information obtained for a glioblastoma tissue sample (N300) is as shown in the figure 12 below.

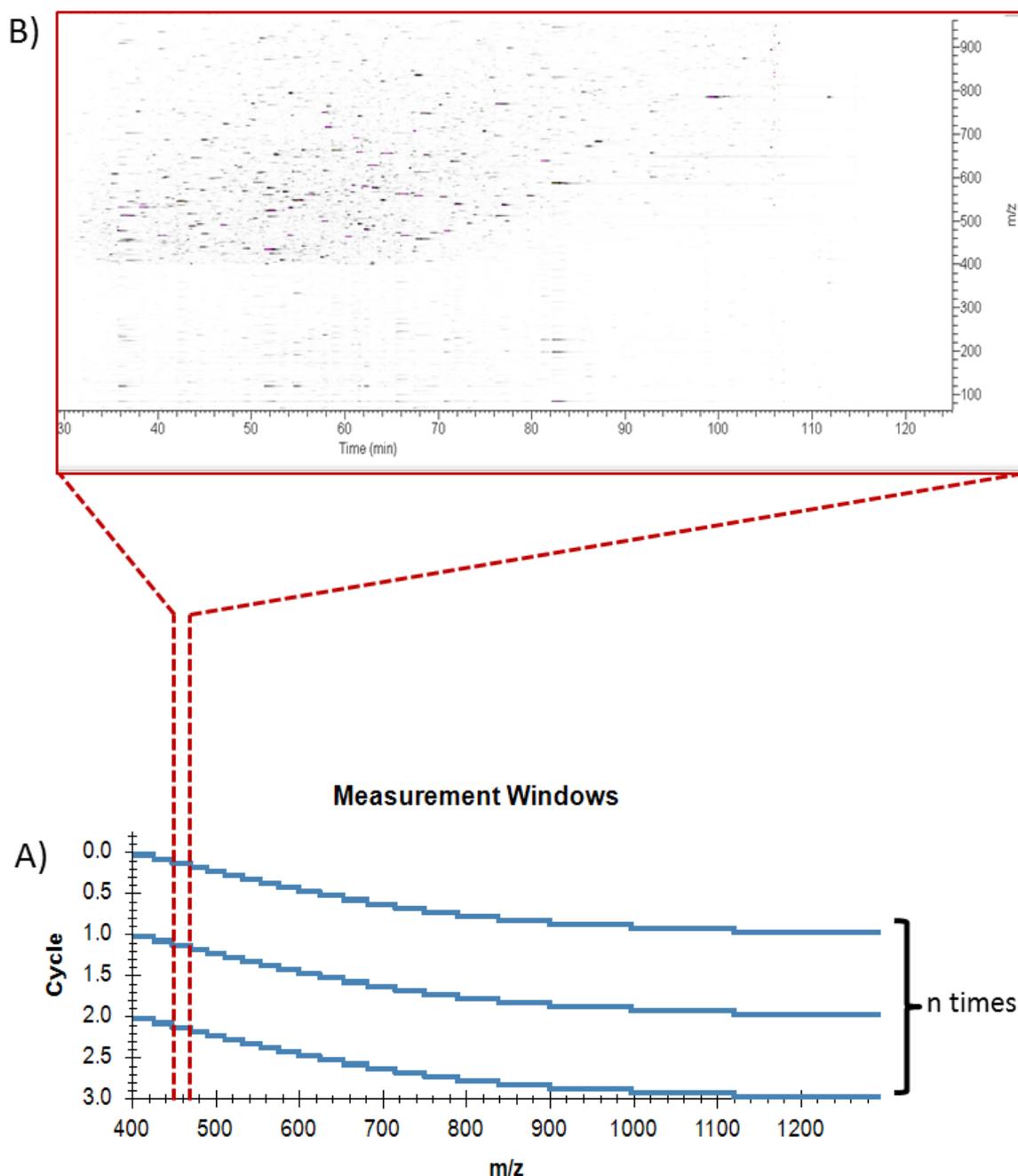


Fig 12: SWATH MS Data independent acquisition data from a tryptic digested glioblastoma tissue sample (N300): A) Depiction of variable sized 20 precursor isolation windows (blue lines) where acquisition of high-resolution fragment ion spectra occurs during the entire chromatographic run of 120 mins across a 400-1300 m/z range within each cycle. B) Representation of MS2 ion map obtained from one SWATH isolation window between 451 and 463 m/z with retention time as abscissa and fragment ion m/z as ordinate.

This precursor ion is among the 667 precursors recorded within the Q1 isolation window of 451 to 463 m/z. Figure 13 shows a typical overlay of peptide fragment extracted ion chromatograms. This peptide is doubly charged and has the amino acid sequence LEAALGEAK. It eluted at the 48.18th minute and was identified as a unique peptide to the protein Prelamin-A/C. Typically, the peak group of the fragment ions displaying the best co-eluting characteristics and matching best to the peak group of extracted reference fragment ion traces, identifies and quantifies the target peptide. The MS1 and MS2 spectrum at apex of this peptide is also illustrated in figure 14. The MS1 signal at apex retention time illustrates the isotopic envelop of the precursor while also indicating the expected relative intensities of the different isotopic forms. The MS2 spectrum at apex illustrates the full recorded MS2 spectrum corresponding to this given peptide's LC-peak apex, highlighting the fragments as provided by the library. The protein coverage of Prelamin-A/C, highlighting the peptide LEAALGEAK++ which was a unique peptide identified with very high accuracy of Q value ≤ 0.01 is as shown in figure 15. Exemplary plots related to the behaviour of the target and the decoy distribution estimation for discriminant scores (Cscores), Qvalues, and sensitivity on the precursor level are also detailed in the supplement figures S2 to S4.

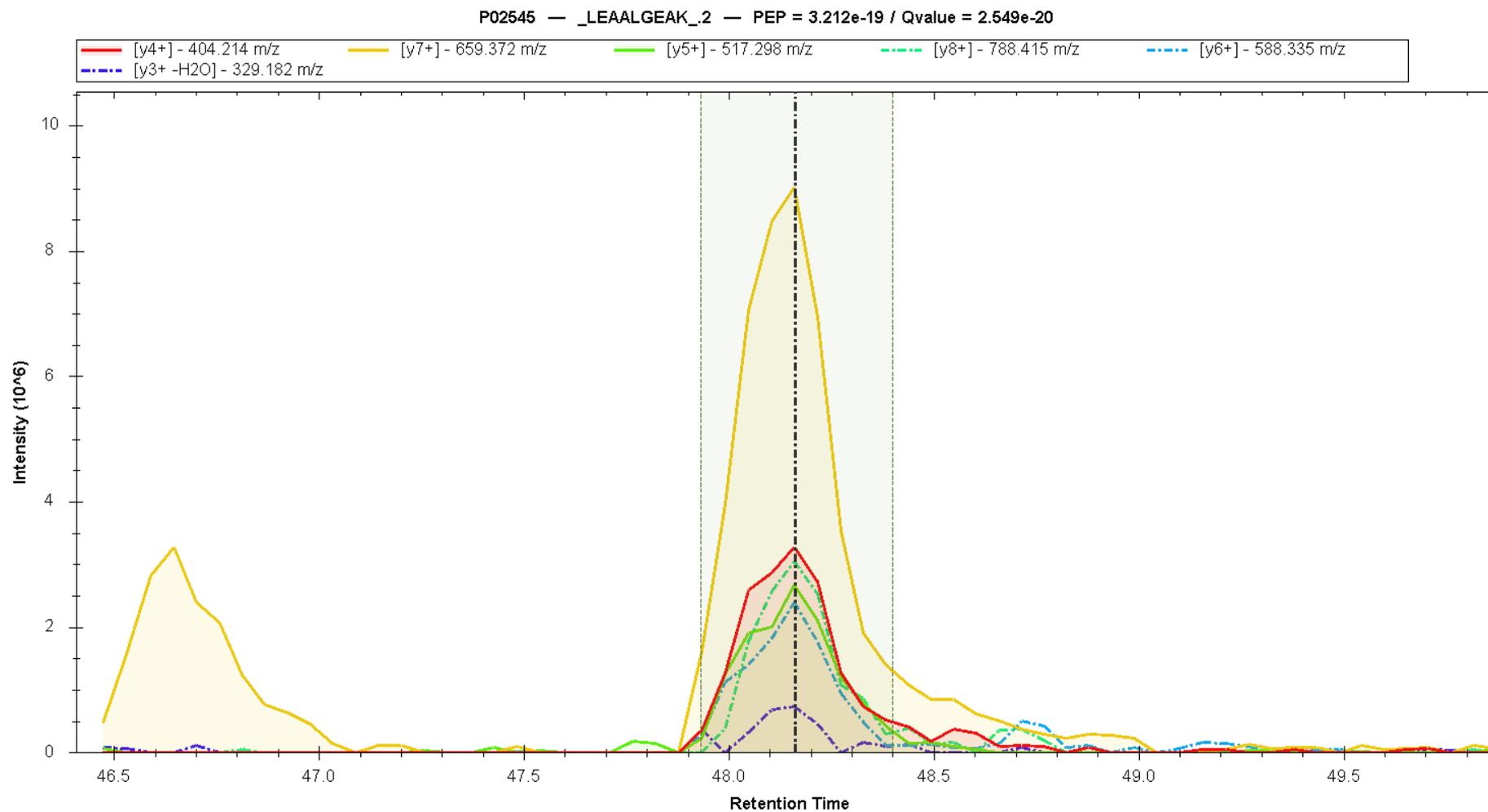


Fig. 13: Fragment ion traces plotted as overlaid extracted ion chromatograms of the peptide with amino acid sequence LEAALGEAK++, eluting at the 48.18th minute recorded within the Q1 isolation window of 451 to 463 m/z from the glioblastoma tissue sample (N300).

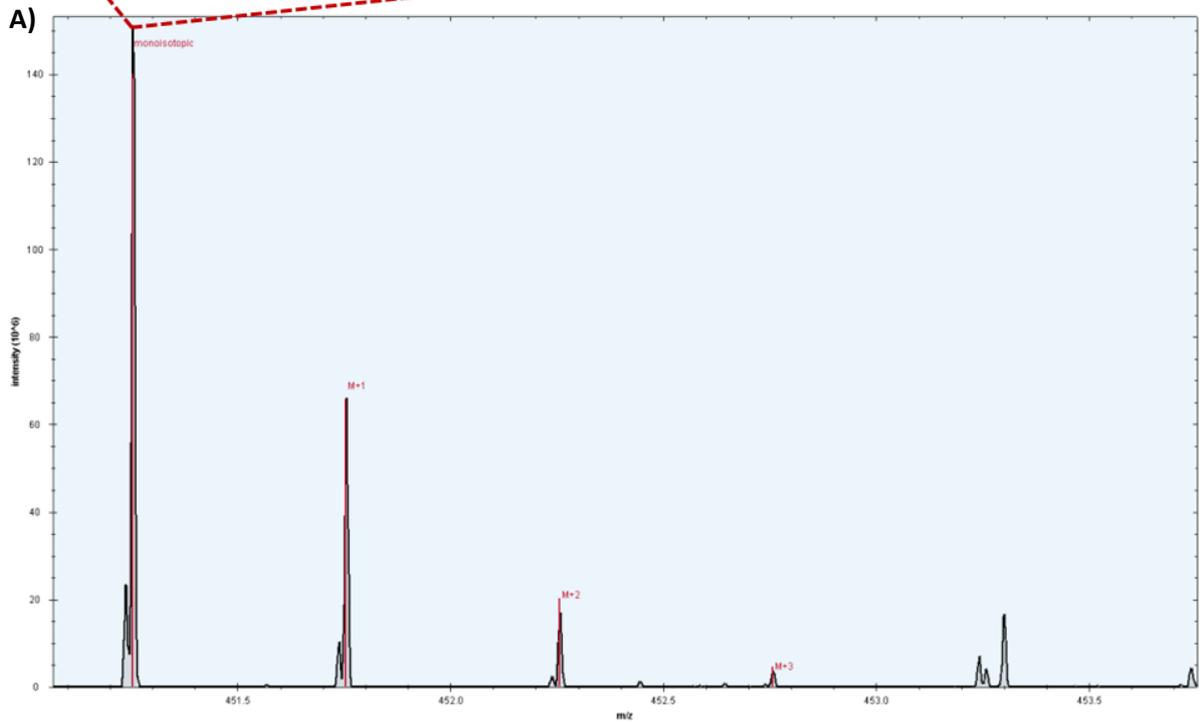
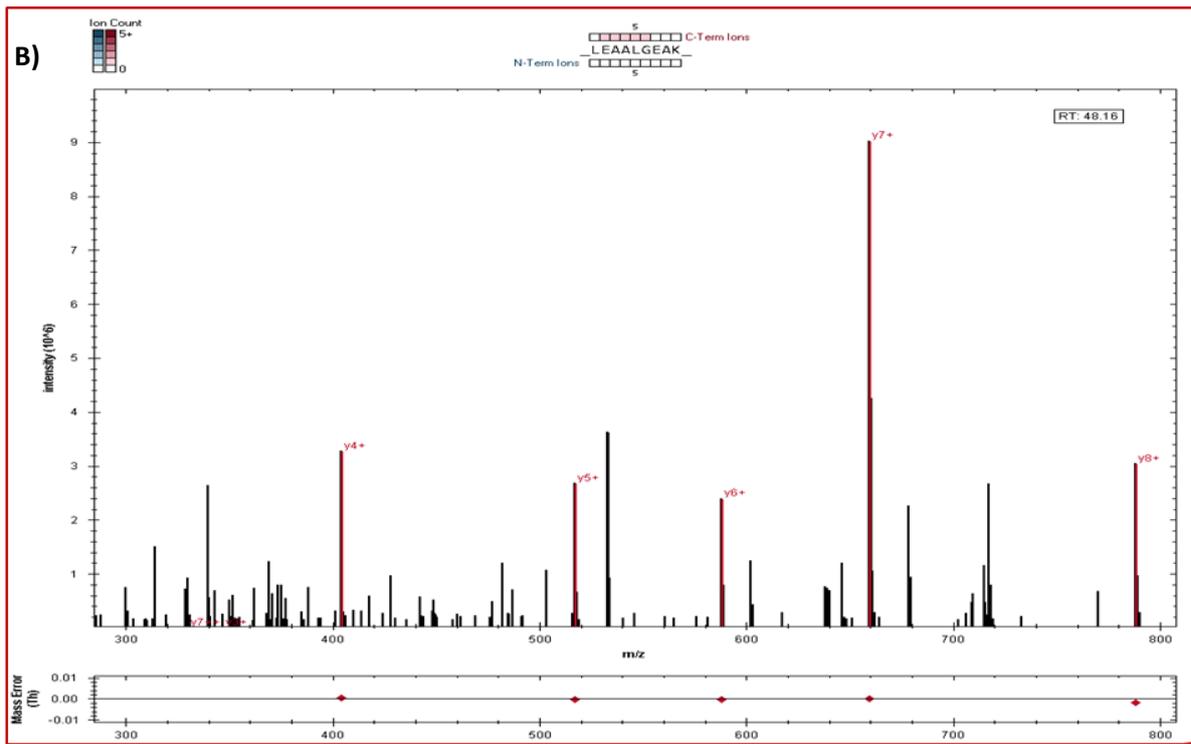


Fig 14: A) MS1 spectrum at apex for the peptide LEAALGEAK++ showing the monoisotopic precursor plus its first 3 isotopic forms. B) MS2 Spectrum at Apex for the peptide LEAALGEAK++ with the option "Show only Library Fragments" as default.

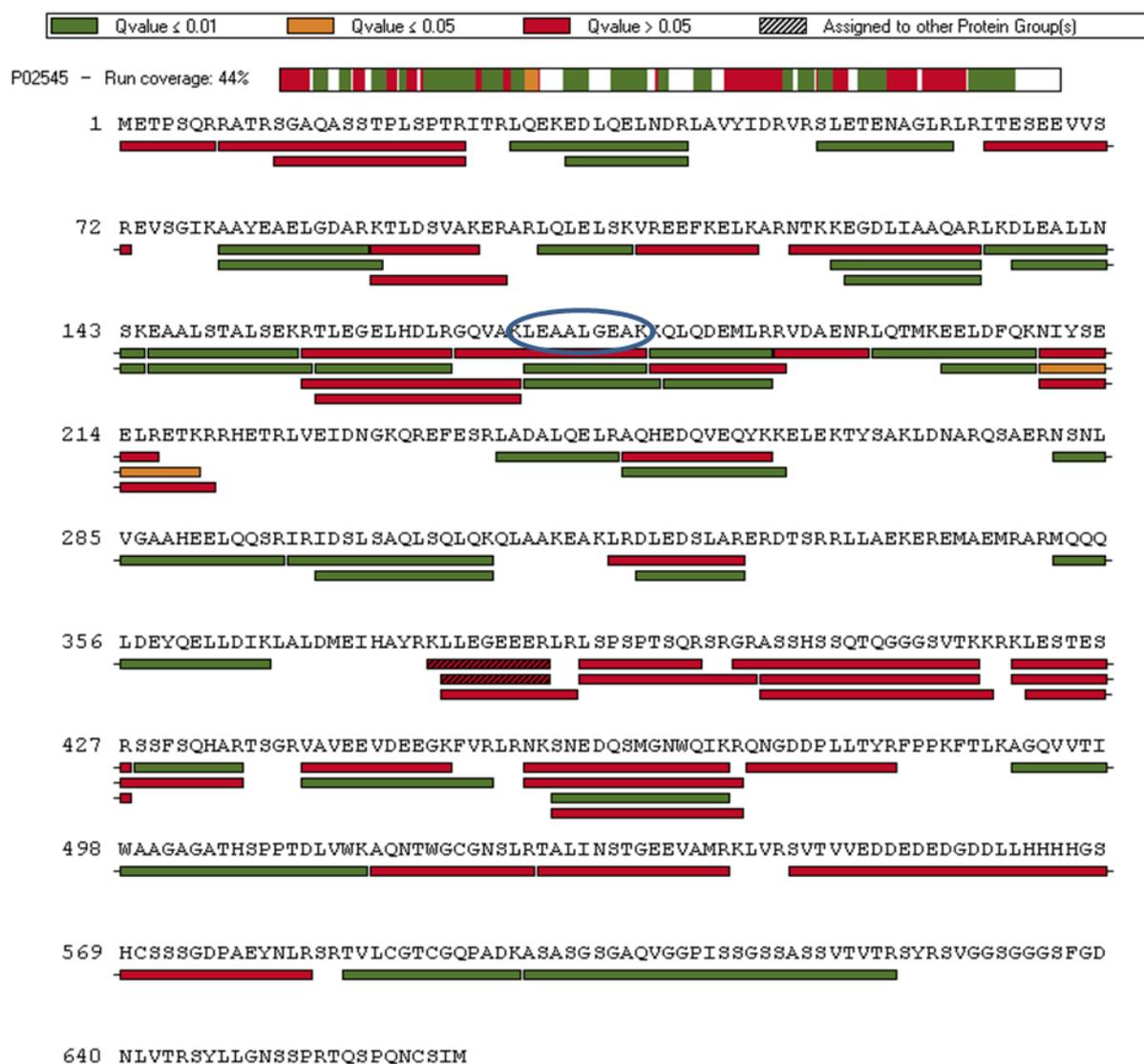


Fig 15: Protein coverage plot for Prelamin-A/C, highlighting the peptide LEAALGEAK++ which is a unique peptide identified with very high accuracy of Q value 2.549×10^{-20} , eluting at the 48.18th minute recorded within the Q1 isolation window of 451 to 463 m/z from the glioblastoma tissue sample (N300). Green represents peptides recorded with Q value ≤ 0.01, yellow for peptides recorded with Q value ≤ 0.05 and red for peptides recorded with Q value > 0.05.

3.1.4 Tumour Tissue proteome

The data obtained from the tumour tissues were processed with the project specific library. The relationship between the different acquired precursors used for identification within all the tumour tissue samples is shown in the figure 16 below.

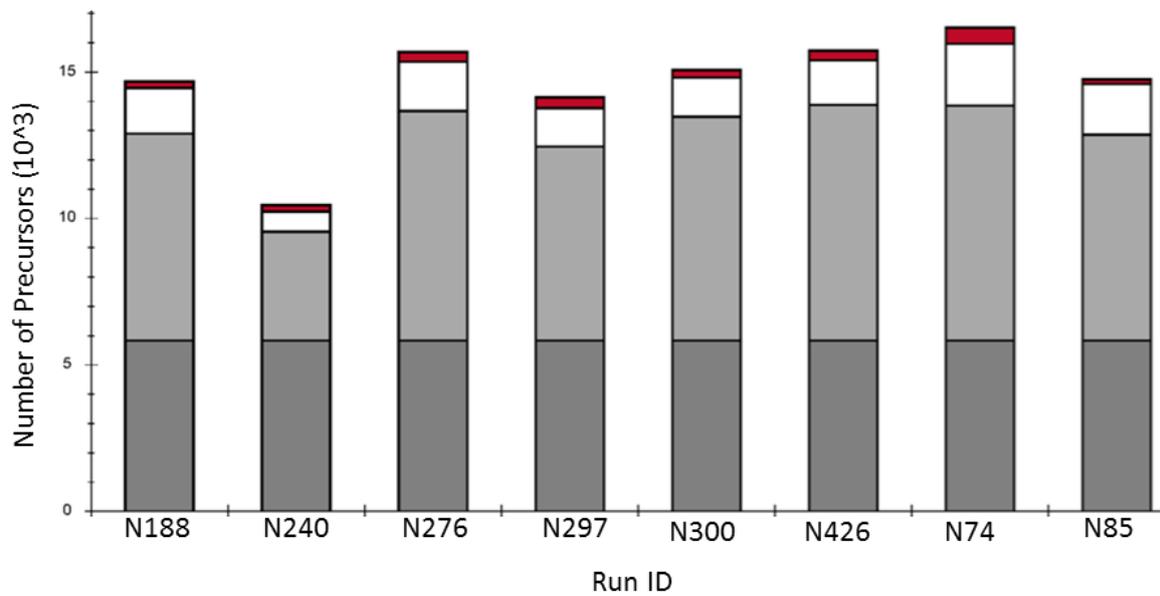


Fig 16: Relationship between the different acquired precursors used for identification within all the tumour tissue samples. Dark grey bars represent complete identifications across all samples; grey bars represent identifications shared in ≥ 50 identifications and red bars represent unique identifications in every run.

The number of identified proteins/protein groups in each patient sample is as shown in table 1. An average of 2900 proteins could be identified in the various tumour tissue samples.

Table 1: Number of identified proteins in different brain tumour tissues

Tumour type	Sample ID	Number of protein groups	Number of identified proteins
Meningioma	N188	2921	2978
	N240	2405	2463
	N74	3194	3255
	N85	2907	2972
Glioblastoma	N276	2948	2995
	N297	2850	2900
	N300	3023	3076
	N426	2942	2985

Venn diagrams were also generated to compare the variability between the two tumour groups. A slightly higher number of proteins/protein groups were identified in the glioblastoma tissues in comparison to the meningioma tissues. A PCA plot was also generated to show the variation of tissues between each tumour type. 71.9% of protein groups (2446) were identified in all four

glioblastoma patient tissue samples and 82.9% (2818) in at least 3 of 4 patients. With meningioma, 60.6% of protein groups (2100) were identified in all four meningioma patient tissue samples and 79% (2737) in at least 3 of 4 patients.

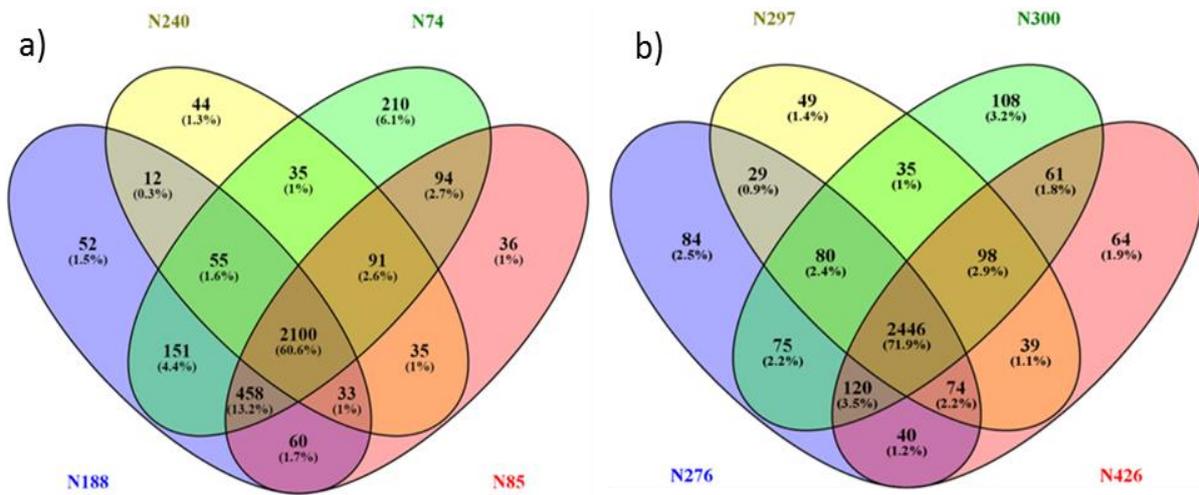
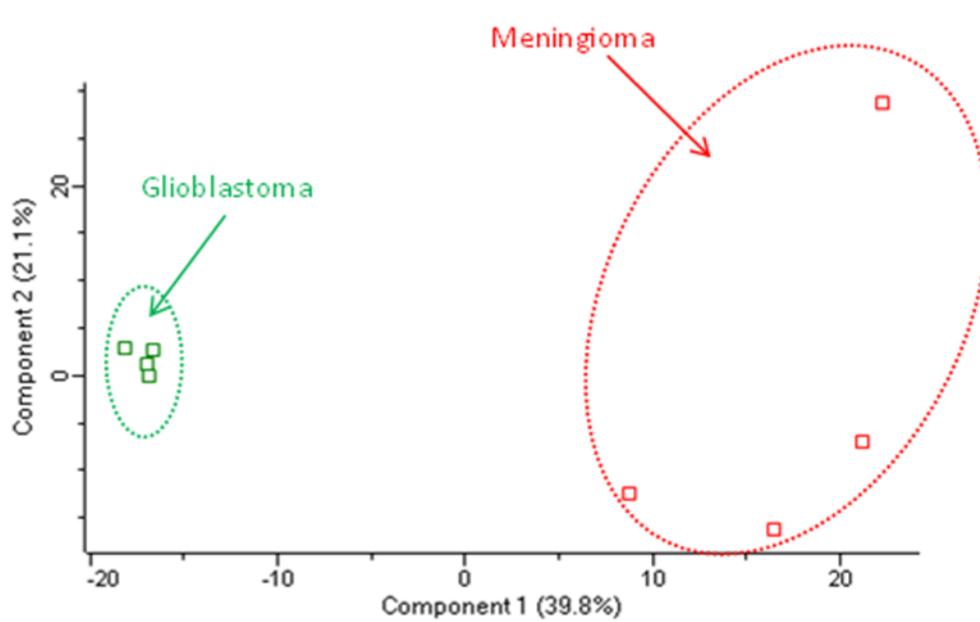


Fig 17: Venn diagrams displaying the overlap of the identified protein groups in each patient within each tumour type. A) Meningioma patients. B) Glioblastoma patients.



3.1.5 Patient Primary Cell culture proteome

From every patient tumour sample obtained, cell cultures were cultivated. The SWATH MS data obtained was processed with the project specific library. The relationship between the different acquired precursors used for identification within the primary cell cultures is shown in the figure 19 below.

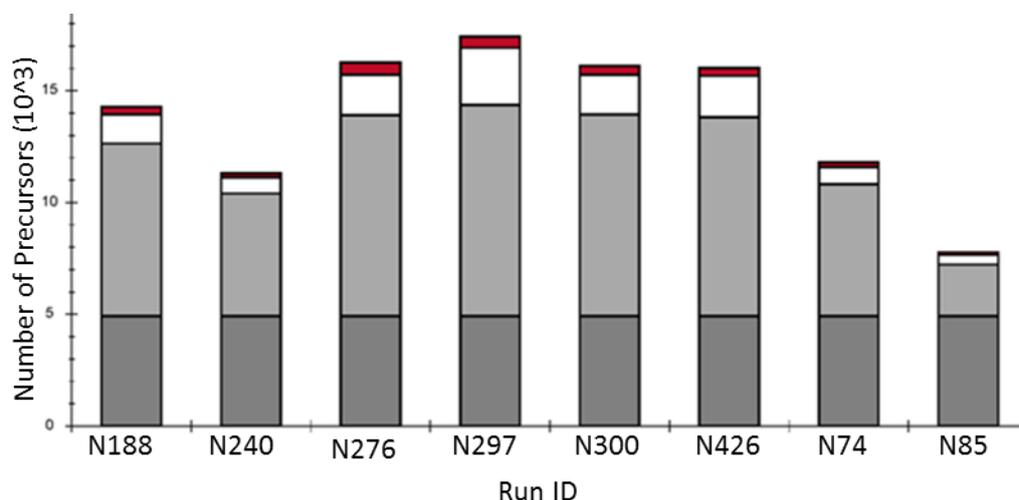


Fig 19: Relationship between the different acquired precursors used for identification within all the cell samples. Dark grey bars represent complete identifications across all samples; grey bars represent identifications shared in ≥ 50 identifications and red bars represent unique identifications in every run.

The number of identified proteins in each tumour cell sample is as shown in table 2. An average of 2600 proteins could be identified in the various cell samples.

Table 2: Number of identified proteins in different patient primary cell cultures

Tumour type	Sample ID	Number of protein groups	Number of identified proteins
Meningioma	N188	2777	2819
	N240	2439	2479
	N74	2641	2677
	N85	2063	2094
Glioblastoma	N276	2869	2910
	N297	3123	3171
	N300	2983	3030
	N426	2868	2911

Venn diagrams were also generated to compare the variability between the two groups of cell lines. On average, almost 500 more proteins/protein groups were identified in the glioblastoma cells in comparison to the meningioma cells. A PCA plot was also generated to show the variation of proteins between each cell type. 70.6% of protein groups (2421) were identified in all four glioblastoma cell lines and 83.1% (2850) in at least 3 of 4 cell lines. With meningioma, 55.6% of protein groups (1746) were identified in all four meningioma patient cell lines and 73.5% (2303) in at least 3 of 4 patient primarily cultured cells.

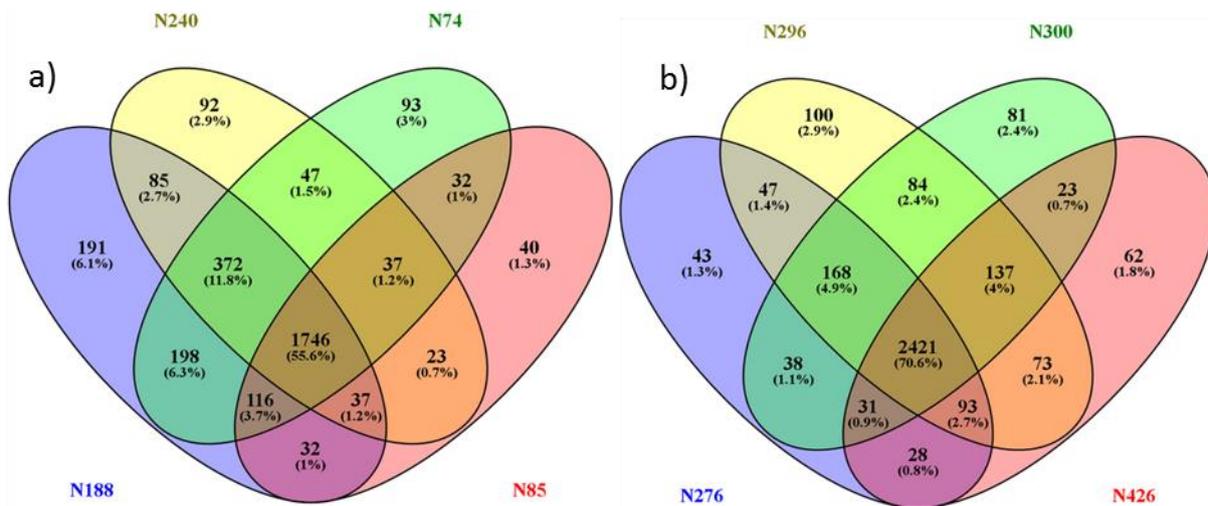


Fig 20: Venn diagrams displaying the overlap of the identified protein groups in each patient obtained from each patient primary cell culture. A) Meningioma patients. B) Glioblastoma patients.

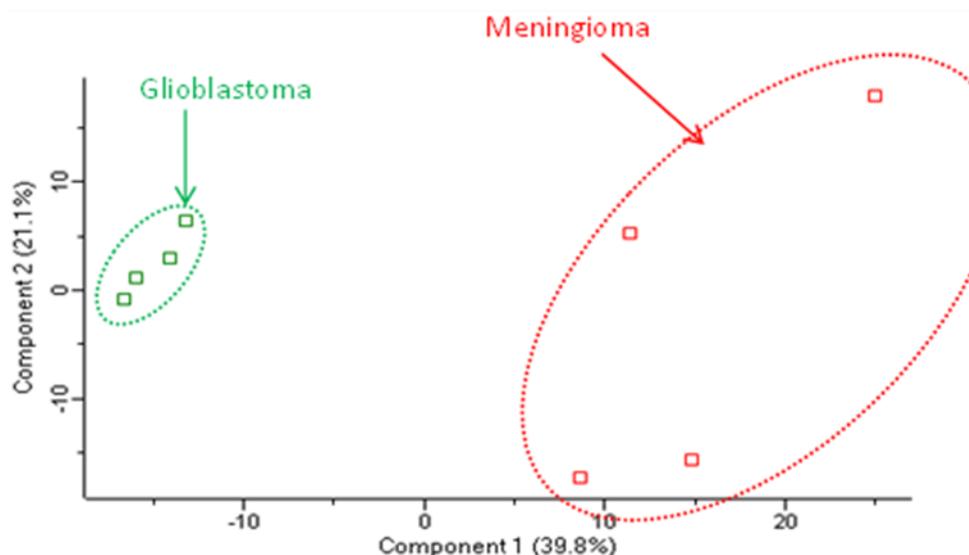


Fig 21: Principal component Analysis (PCA) clustering of glioblastoma (green) and meningioma (red) cells. Component 1 and 2 are shown for each sample.

3.1.6 Extracellular vesicles

Extracellular vesicles were harvested from cell culture supernatant by the density-based approach. The SWATH MS data obtained was processed with the project specific library. The relationship between the different acquired precursors used for identification within the extracellular vesicles is shown in the figure 22 below.

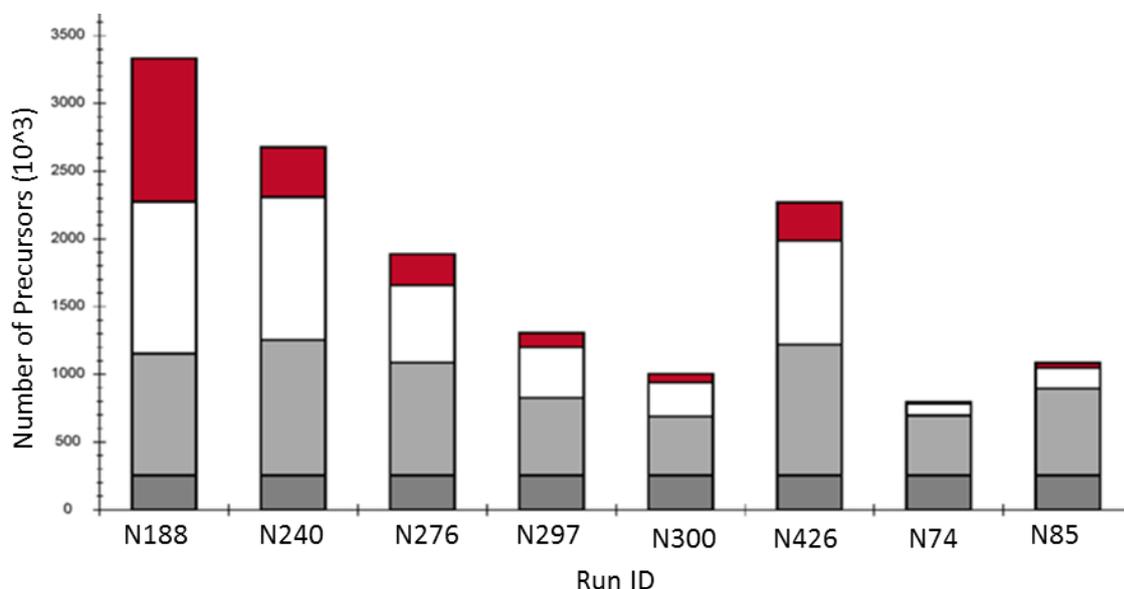


Fig 22: Relationship between the different acquired precursors used for identification within all the EV samples. Dark grey bars represent complete identifications across all samples; grey bars represent identifications shared in ≥ 50 identifications and red bars represent unique identifications in every run.

The number of identified proteins in EV sample is as shown in table 3. An average of 2600 proteins could be identified in the various extracellular vesicles.

Table 3: Number of identified proteins in different extracellular vesicles

Tumour type	Sample ID	Number of protein groups	Number of identified proteins
Meningioma	N188	1202	1235
	N240	787	802
	N74	342	359
	N85	243	254
Glioblastoma	N276	528	538
	N297	439	450
	N300	382	392
	N426	650	669

Venn diagrams were also generated to compare the variability between the two groups of cell lines. A much lower number of proteins were identified in the exosome samples compared to the cell lines and the tumour tissues. The variation of identified proteins/protein groups in both the glioblastoma and meningioma extracellular vesicles was quite vast. A PCA plot was also generated to show the variation of proteins between the EVs. 22.7% of protein groups (211) were identified in all four glioblastoma extracellular vesicles and 15.3% of protein groups (204) were identified in all four of the meningioma patient EV samples.

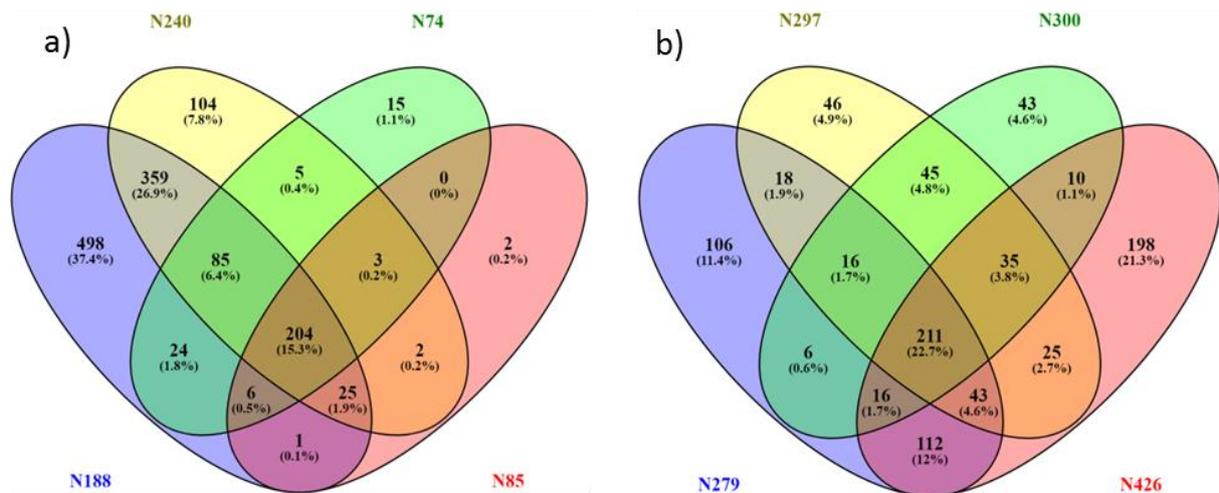


Fig 23: Venn diagram of the overlap of the identified protein groups in each patient within the extracellular vesicles obtained from extracellular vesicles. A) Meningioma patients. B) Glioblastoma patients.

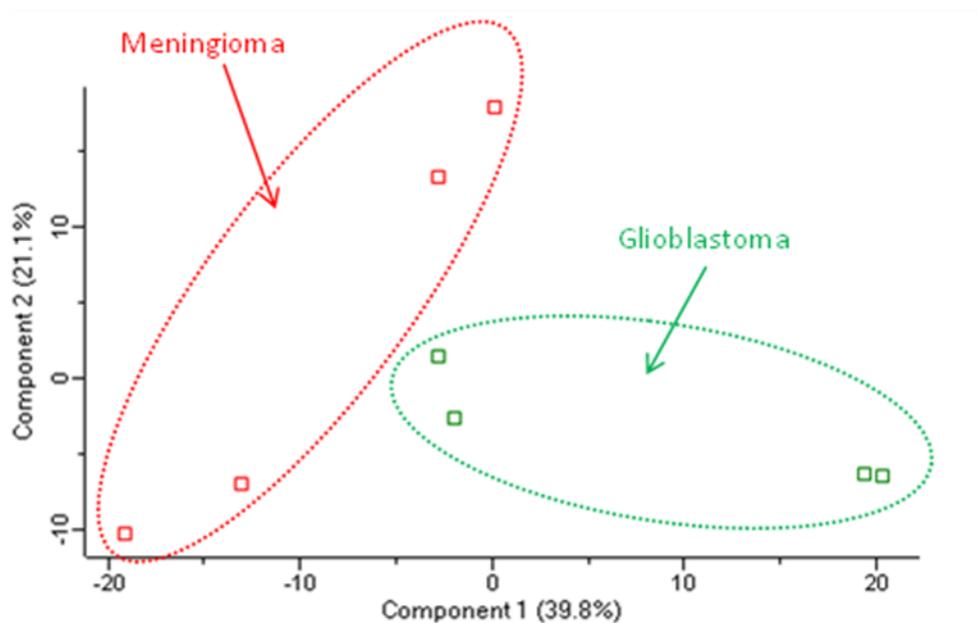


Fig 24: Principal component Analysis (PCA) clustering of glioblastoma (green) and meningioma (red) extracellular vesicles. Component 1 and 2 are shown for each sample.

These proteins identified in either GBM or MEN extracellular vesicles were compared to the list of top 100 proteins that are often identified in vesicles which is found on the Vesiclepedia webpage version 4.1 (http://microvesicles.org/extracellular_vesicle_markers)[107]. Out of the 359 proteins identified in the meningioma vesicles, 64 were identified among these 100 protein extracellular vesicle makers, including 19 of the top 20 proteins on the list (Table 4). For the glioblastoma vesicles, 353 proteins could be identified in the vesicles of at least three patients. These made up 72 of the 100 potential vesicle markers, with 19 also being identified in amongst the top 20 markers (Table 4). The protein marker which was not quantified in both the glioblastoma and meningioma patient vesicular samples was the CD63 antigen. A deeper analysis of the individual patients showed that CD63 antigen was only identified for one meningioma patient with the sample code EV_N188. Another protein of interest was the cytochrome c oxidase. This was not identified in any of the meningioma or glioblastoma EV samples.

Table 4: Extracellular vesicle proteins of both MEN and GBM identified amongst the top 20 EV protein markers listed in the Vesiclepedia database

Protein Name	Protein Accessions	Gene symbol in Vesiclepedia
Programmed cell death 6-interacting protein	Q8WUM4	PDCD6IP
Glyceraldehyde-3-phosphate dehydrogenase	P04406	GAPDH
Heat shock cognate 71 kDa protein	P11142	HSPA8
Actin, cytoplasmic 1	P60709	ACTB
Annexin A2	P07355	ANXA2
CD9 antigen	P21926	CD9
Pyruvate kinase PKM	P14618	PKM
Heat shock protein HSP 90-alpha	P07900	HSP90AA1
Alpha-enolase	P06733	ENO1
Annexin A5	P08758	ANXA5
Heat shock protein HSP 90-beta	P08238	HSP90AB1
14-3-3 protein zeta/delta	P63104	YWHAZ
14-3-3 protein epsilon	P62258	YWHAE
Elongation factor 1-alpha 1	P68104	EEF1A1
Phosphoglycerate kinase 1	P00558	PGK1
Clathrin heavy chain 1	Q00610	CLTC
Peptidyl-prolyl cis-trans isomerase A	P62937	PPIA
Syntenin-1	O00560	SDCBP
Fructose-bisphosphate aldolase A	P04075	ALDOA

3.1.7 Comparison of protein identified in all patient tissues, cells and EVs

Proteins which were identified and quantified in all the tumour tissue samples, cells and extracellular vesicles from all glioblastoma and meningioma patients were clustered in the figure 25 below.

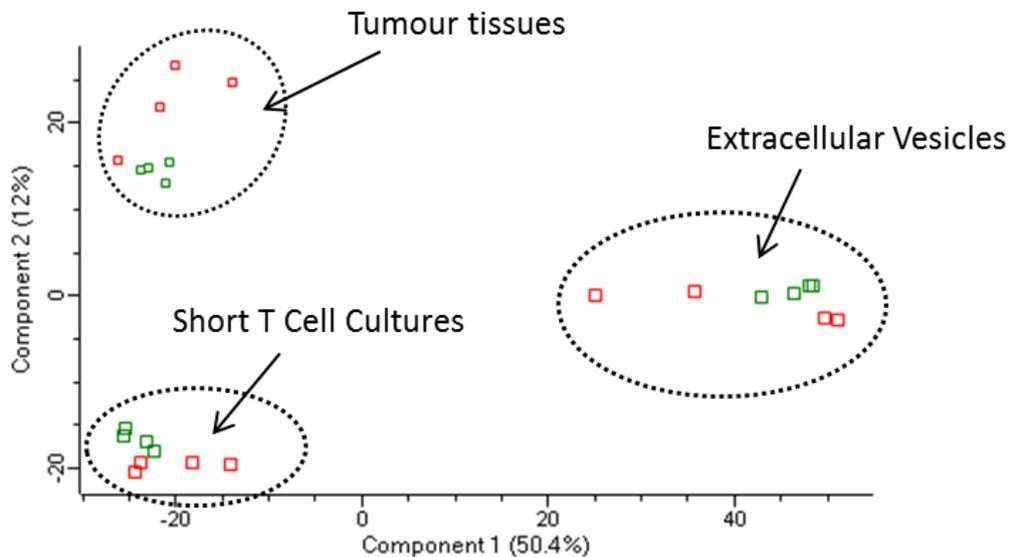


Fig 25: Principal component Analysis (PCA) clustering of all glioblastoma (green) and meningioma (red) patients' tissues, cells and extracellular vesicles. Component 1 and 2 are shown for each sample.

To identify potential EV protein biomarkers which can be used for diagnostics of both meningiomas and glioblastomas, a closer look was taken into the proteins identified and quantified in the extracellular vesicles and also in the cells and tumour tissue samples. Proteins were filtered as found in at least three of the four biological replicates in every scenario. For the meningioma samples, 270 proteins were identified and quantified in all extracellular vesicles, cells and tumour tissues and the number of proteins in all glioblastoma samples was 262. Among these proteins, 171 of them were identified in meningioma and glioblastoma extracellular vesicles, cells and tumour tissues. The figure 26 is a Venn diagram comparison of the proteins identified within the different tumour types. A list of these proteins presented is available in the supplement Table S1, S2 and S3. They are presented in a descending order of abundance in reference to extracellular vesicles only. The first protein on the list is the most abundant in the extracellular vesicles of each tumour type.

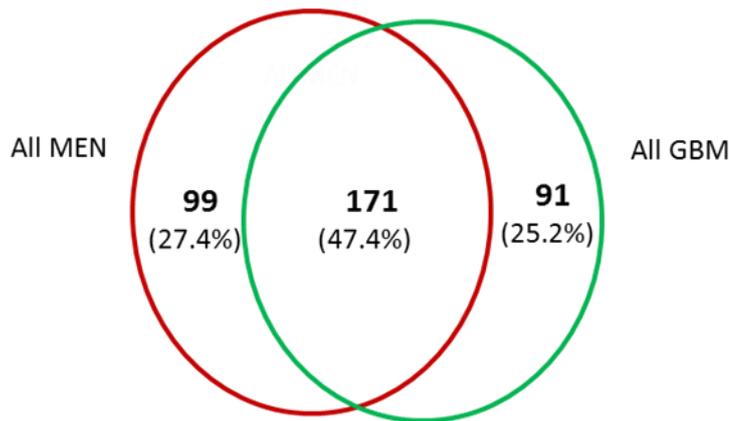


Fig 26: Venn diagram of the overlap of the identified protein groups in all Meningioma patient tissues, cells and EVs (red) and Glioblastoma patient tissues, cells and EVs (green).

These proteins which were exclusive for each tumour types were further analysed with a GO analysis tool called Panther GO Slim. This was to classify the proteins based on their molecular functions, biological process and cellular component. In both meningioma and glioblastoma exclusive proteins, the most abundant molecular functions were binding (41% and 43%) and catalytic activity (41% and 35%) respectively. Some biological processes were identified only with the proteins exclusive to glioblastomas. These include; biological adhesion, growth and reproduction. Meanwhile for the cellular component criterion, proteins responsible for one extra component (cell junction) was found only found in the meningiomas.

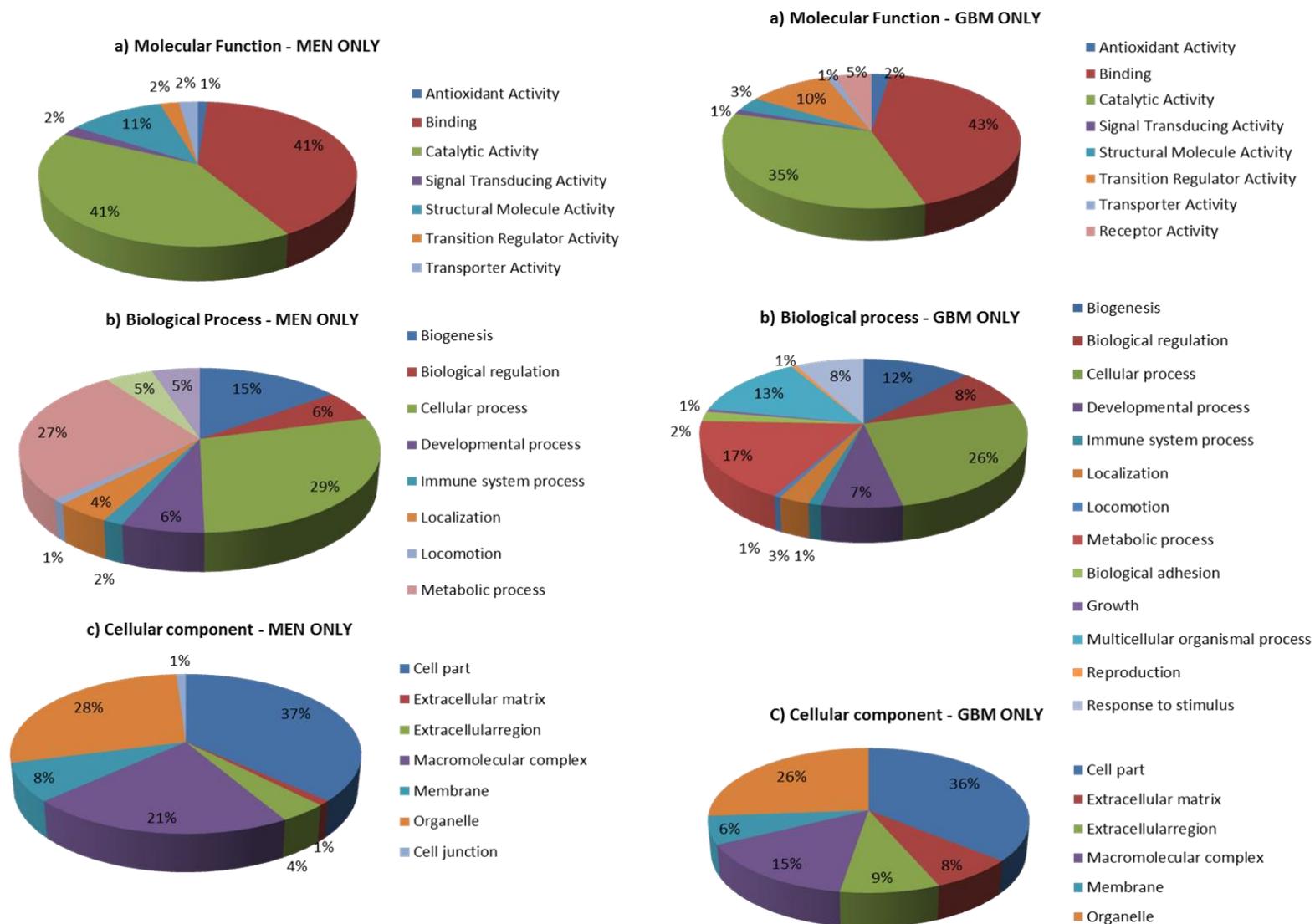


Fig 27: GO terms of the target Proteins. GO analysis according to (a) biological process, (b) cellular component, (c) molecular function of proteins exclusive to Meningioma cells, tissues and EVs (left) and those exclusive to Glioblastoma Multiforme cells, tissues and EVs (right)

3.2 Proteome variation between tissues and patient-derived primary cell culture

One of the goals of this project was to evaluate how closely primary cell cultures of meningioma and glioblastoma resemble their originator tumour on proteome level. The proteins identified and quantified from the experiments as described above in section 5.1 were further statistically analysed to study these differences. Comparing the spectral libraries which were generated from the shotgun analysis of these samples, there is a slight difference in the number of proteins identified. The tissue specific library contains 35378 Precursors, 28249 Peptides and 3727 Proteins. Meanwhile, the cell culture library contains 32371 Precursors, 25746 Peptides and 3455 Proteins. For the SWATH MS data, various filtering criteria within each tumour type were employed. Proteins identified in either one of four patients, two of four, three of four or all four were plotted in the figures 28 and 29 for meningioma and glioblastoma samples respectively. However, like most scientific experiments, the criteria used for further data analysis was for a protein to be identified in at least 3 of four patient samples each. With this in consideration, 2198 proteins were identified in the cell cultures and 2565 proteins in the tissues for meningiomas and 2694 proteins were identified in the cell cultures and 2557 proteins in the tissues for glioblastomas. The table 5 below summarises the number of identified proteins within the cells and tissues of the different tumour types.

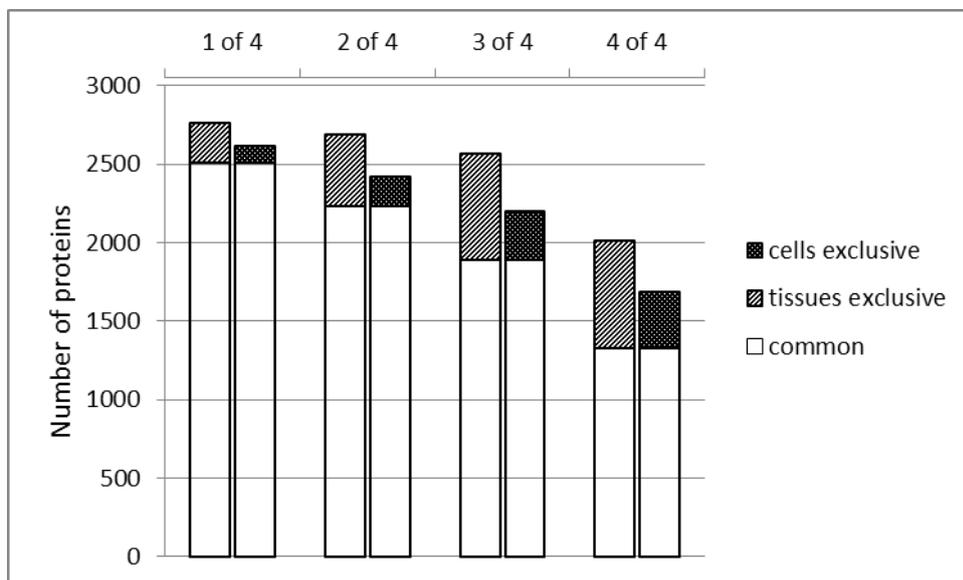


Fig 28: Common and exclusive proteins identified in various replicates of meningioma tissue and cell culture samples

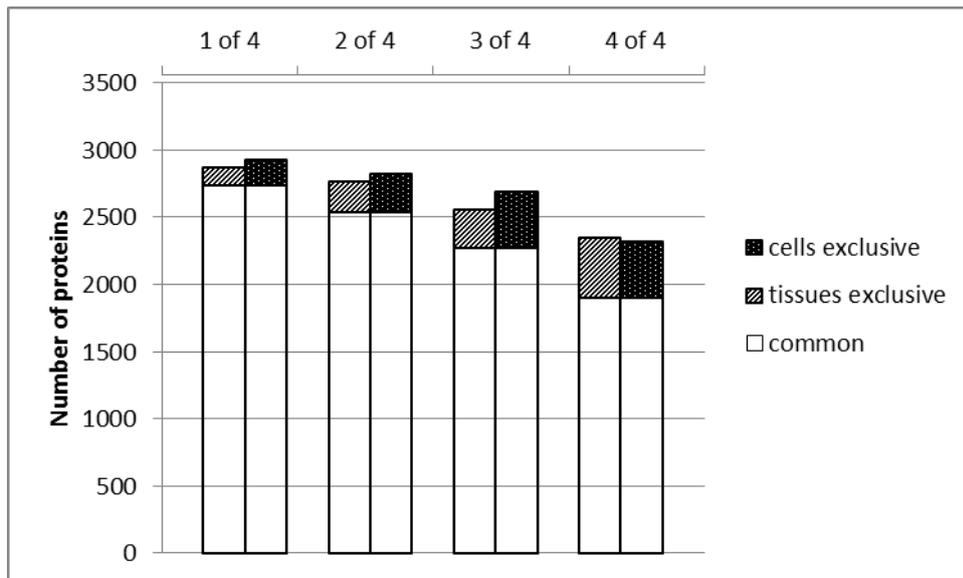


Fig 29: Common and exclusive proteins identified in various replicates of glioblastoma tissue and cell culture samples

Table 5: Number of identified proteins in various replicates of cells and tissues from GBM and MEN patients

Tumour type	Replicates	Tissues	cells	Common proteins
Meningioma	1 of 4	2763	2620	2513
	2 of 4	2688	2419	2237
	3 of 4	2565	2198	1894
	4 of 4	2010	1688	1327
Glioblastoma	1 of 4	2869	2930	2743
	2 of 4	2771	2829	2544
	3 of 4	2557	2694	2273
	4 of 4	2351	2317	1903

3.2.1 Meningioma tissues and cell culture proteome

To view these differences on a quantification basis, only proteins which were quantified in both cells and tissues of the different tumour types were further analysed. With this in consideration, 1894 proteins were found in three of four replicates of meningioma tissues and cells. 254 (13.4%) of these proteins were significantly regulated with greater than 2-fold change in cells more than tissues and 146 (7.7%) were significantly regulated in tissues with greater than 2-fold change than cells.

Therefore, 1494 proteins (79%) had no significant change in relative protein abundance. A volcano plot was generated (figure 30) to highlight these differences.

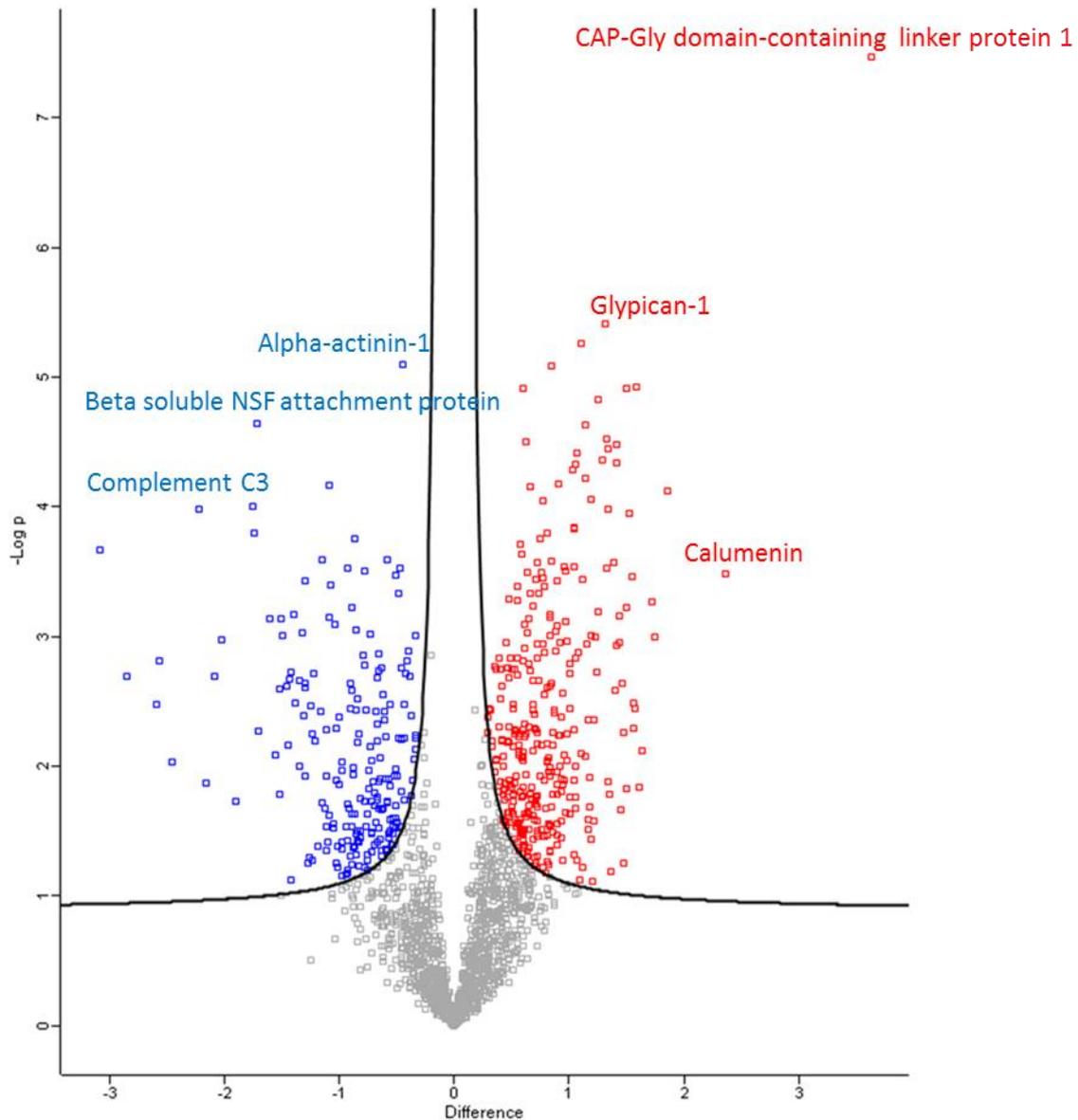


Fig 30: Levels of dysregulated proteins between meningioma tissues and primary cell culture. Proteins highlighted in red are upregulated in >2-fold change in cells compared to tissues and blue represents proteins downregulated in >2-fold change in tissues compared to cells.

The next set of proteins of interest contains the proteins which were identified in either one of the samples and not the other. 672 proteins were identified only in the MEN tissues when compared to primary cells while 304 proteins were unique to the primary cells. The GO analyses of these proteins were done using PantherDB to understand their molecular function and possibly the reason for their presence (figure 31).

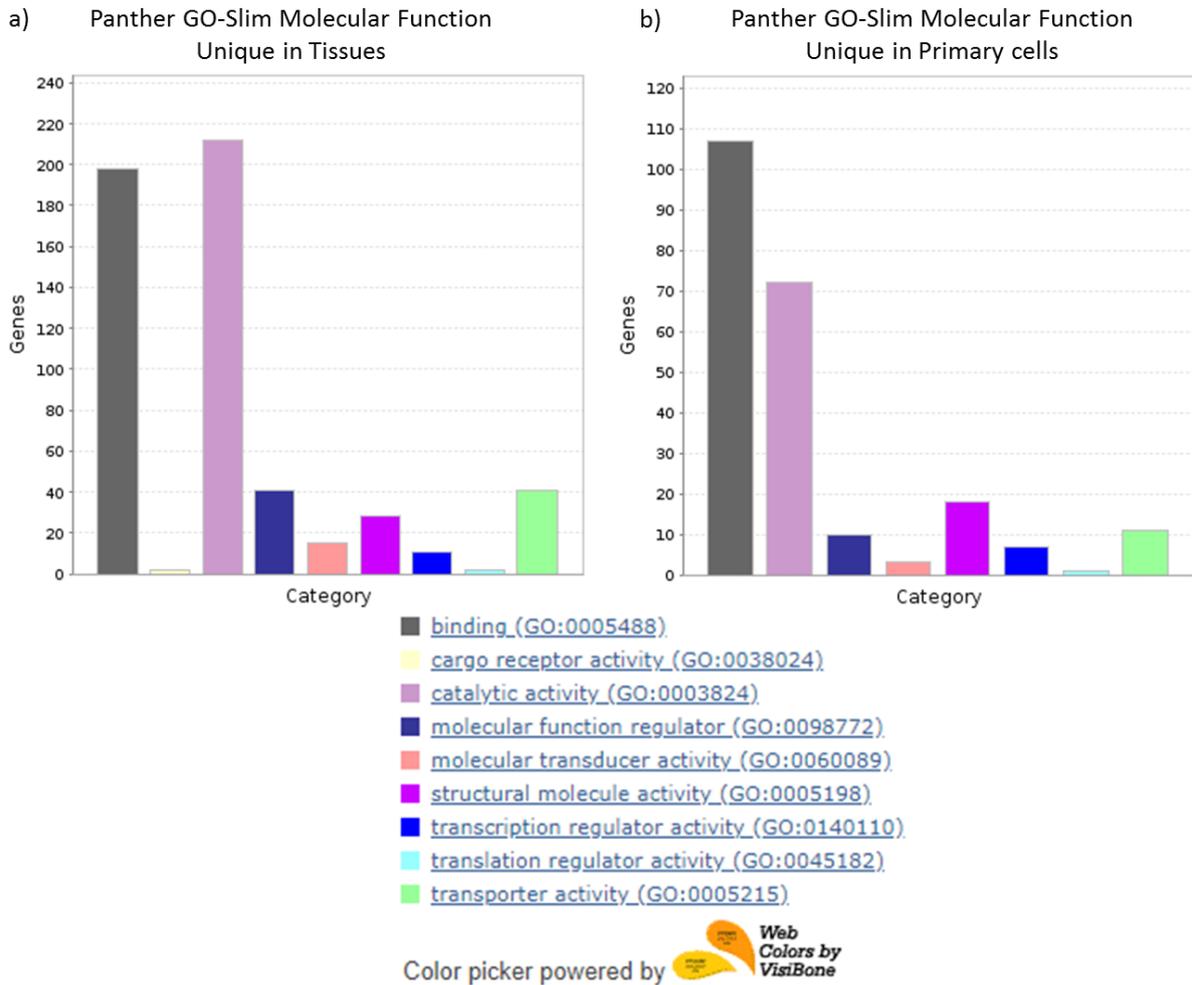


Fig 31: Panther GO-Slim analysis of protein molecular function. a) Number of genes involved in various molecular processes of 672 proteins uniquely identified in MENINGIOMA tissues when compared to their primary cells. b) Number of genes involved in various molecular processes of 304 proteins uniquely identified in primary cells when compared to their originator patient tumour tissues.

3.2.2 Glioblastoma tissues and cell culture proteome

For GBM samples, 2273 proteins were found in three of four replicates of tissues and cells. Of the 759 proteins which were differentially regulated, 421 (18.5%) of these proteins were up regulated more than 2-fold change in cells more than tissues and 338 were down regulated in tissues with greater than 2-fold change then cells (14.87%). Therefore, 1514 proteins (66.6%) had no or less than 2-fold change in relative protein abundance. A volcano plot was generated (figure 32) to highlight these differences.

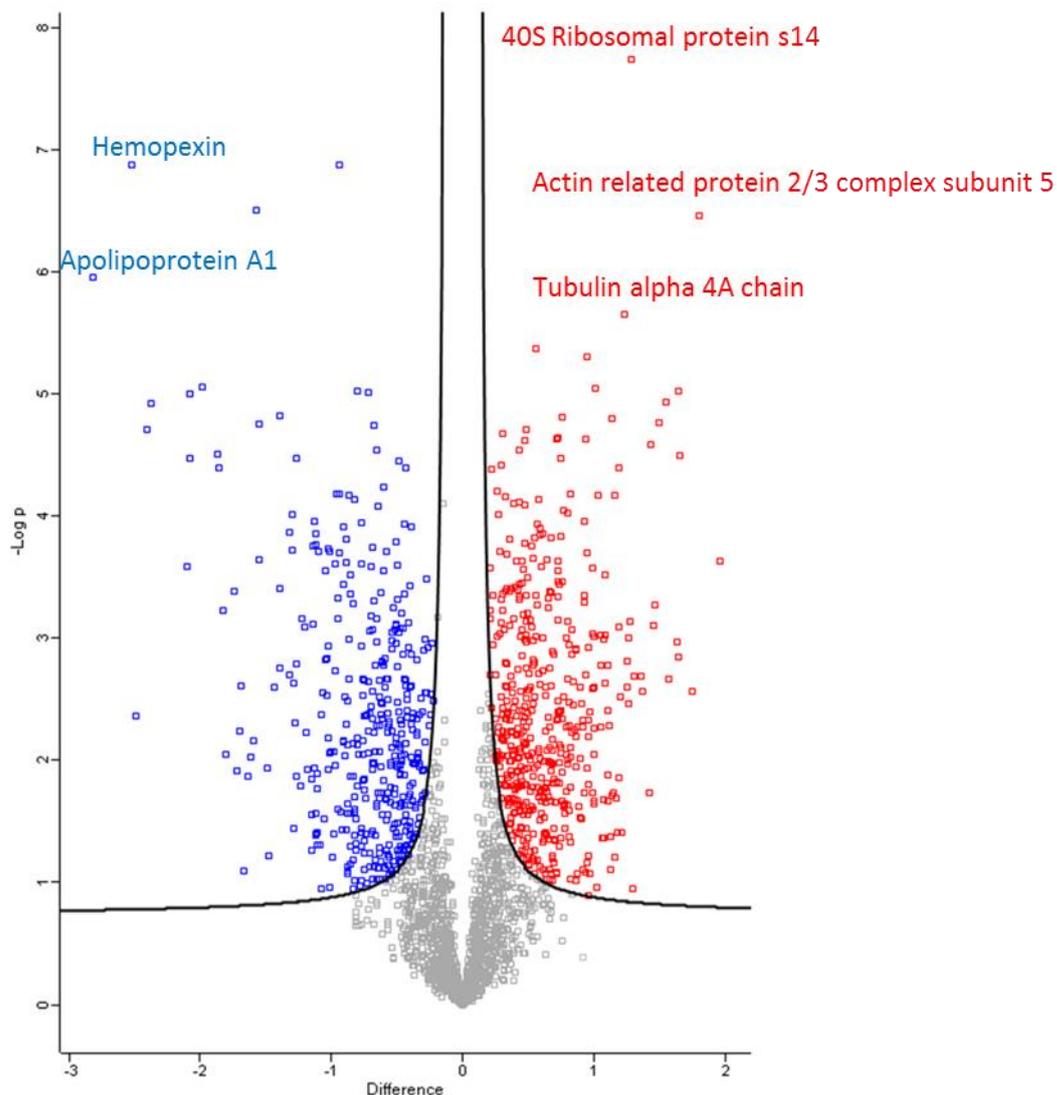


Fig 32: Levels of dysregulated proteins between glioblastoma tissues and primary cell culture. Proteins highlighted in red are up regulated in >2-fold change in cells compared to tissues and blue represents proteins downregulated in >2-fold change in tissues compared to cells.

The next set of proteins of interest contains the proteins which were identified in either one of the samples and not the other. 284 proteins were identified only in the GBM tissues when compared to primary cells while 421 proteins were unique to the primary cells. The GO analyses of these proteins were done using PantherDB to understand their molecular function and possibly the reason for their presence (figure 33).

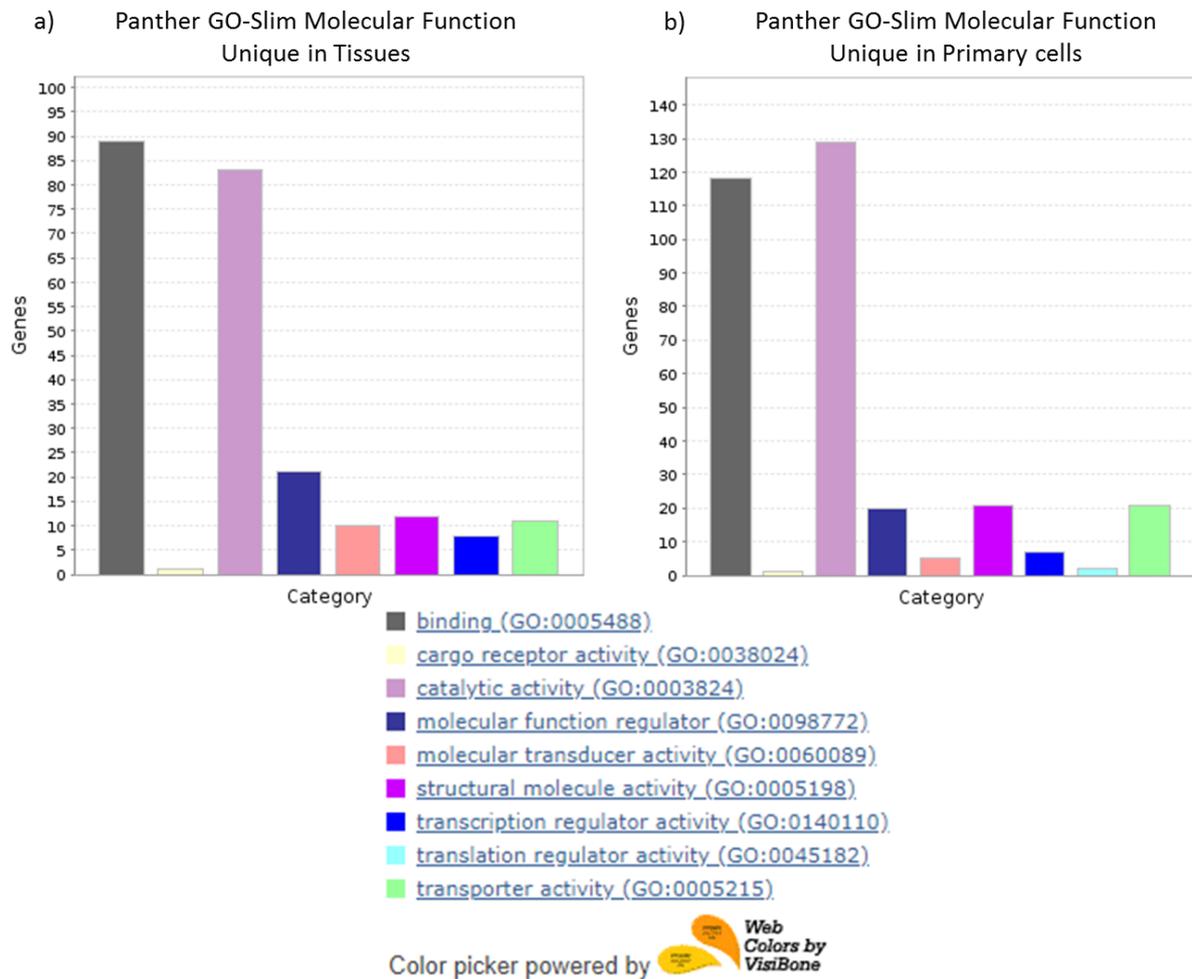


Fig 33: Panther GO-Slim analysis of protein molecular function. a) Number of genes involved in various molecular processes of 672 proteins uniquely identified in GLIOBLASTOMA tissues when compared to their primary cells. b) Number of genes involved in various molecular processes of 304 proteins uniquely identified in primary cells when compared to their originator patient tumour tissues.

3.3 Differential proteomic analysis of tumour and non-tumour tissues

To assess the proteomic differences between cancer and non-cancerous tissues, the tumour biopsies were compared to control samples. Due to ethical reasons, healthy brain tissue can generally not be excised from human brain for research purposes. The control samples used were human epileptic brain tissue excised to either stop or reduce seizures in patients. Eight patient biopsies were used in total, four obtained from the white matter of the temporal lobe and four from the cortex. These were compared against patient biopsies: four meningioma and four glioblastoma tissues. These samples were prepared for mass spectrometric analysis as described in sections 5.4 to 5.6. The proteins were identified and quantified on MS2 level using the SWATH tool for DIA based proteomics. The Q1 isolation window widths used for all samples are detailed under Section 5.8 Table 13 while an example of an LC gradient acquired for one sample is shown in supplement figure S1.

3.3.1 Library generation and evaluation

The spectral libraries required for data analysis of these samples were generated as similarly described under Section 3.1.2. A sample specific library was generated from shotgun tissue analysis of all 8 non-cancerous tissue biopsies. The tissue sample specific library (ssTT) detailed under Section 3.1.2 was merged with the sample specific library using Spectronaut Professional to generate a project specific library. A third library was downloaded from the SWATHAtlas webpage called the Pan human library. Figure 34 shows the precursor m/z count from the libraries. A summary on the number of precursors, peptides and proteins available in all the aforementioned libraries are depicted in Fig. 35 below.

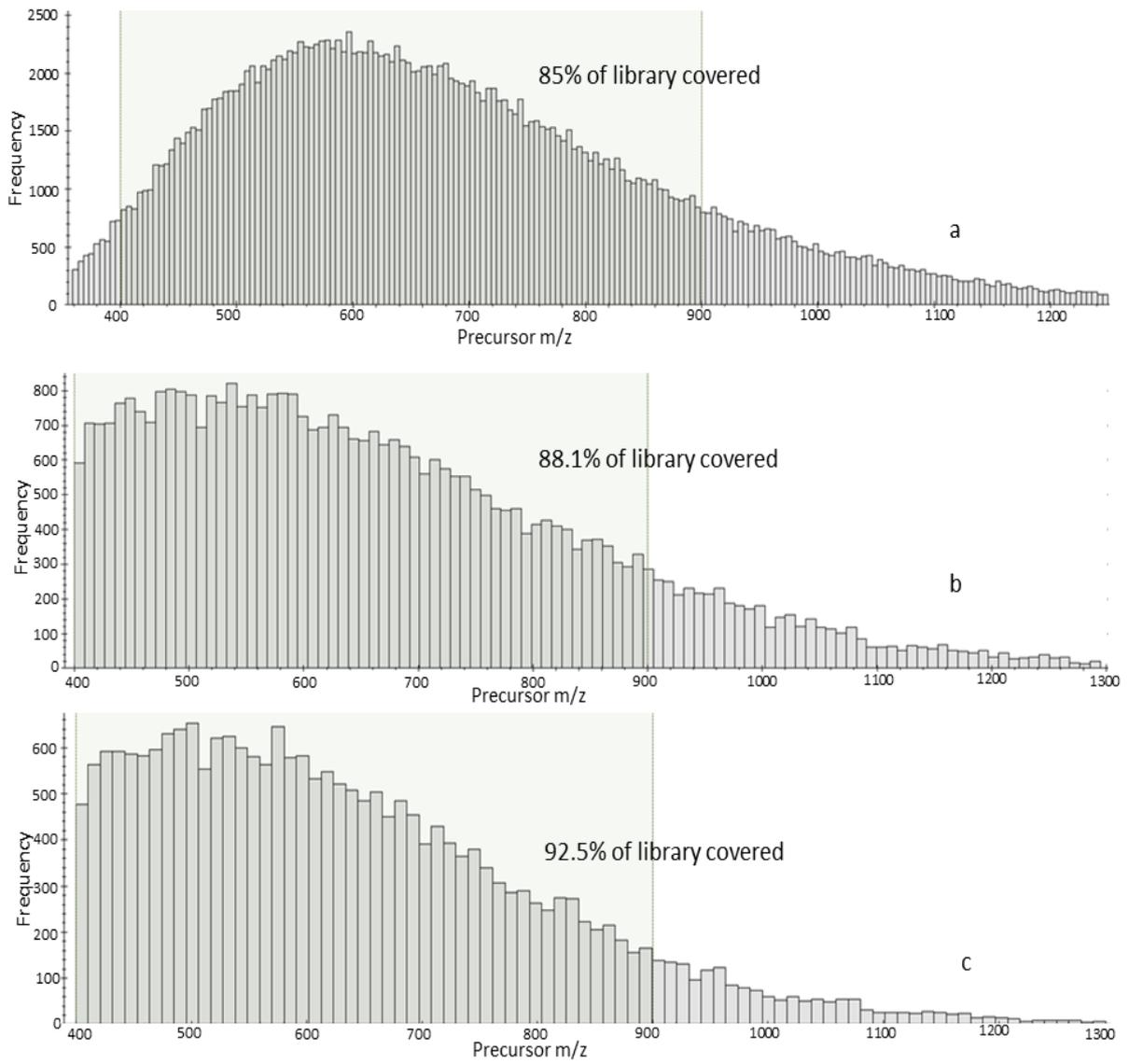
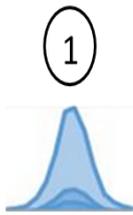


Fig 34: Precursor m/z population of different sample digests used for the generation of the various spectral libraries with precursor frequency plotted as ordinate and precursor m/z as abscissa. A) Pan-human b) Project specific c) Sample specific library

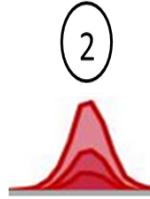
Sample Specific Library



Non-cancer Tissue

- 22793 Precursors
- 18560 Peptides
- 2575 Proteins

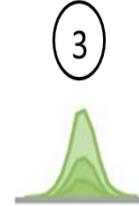
Project specific library



Non-cancer + cancer tissue

- 44023 Precursors
- 34980 Peptides
- 4265 Proteins

Pan-human Library



Tissues & Cells

- 210921 Precursors
- 149066 Peptides
- 10643 Proteins

Fig 35: Number of precursors, peptides and proteins available in the sample specific libraries, project specific library and the Pan-human library.

The three different spectral libraries were assessed based on the number of identified peptides and proteins from each analysed DIA sample. Eight patient non-tumour tissues: four from white matter (WM) and four from cortex (K) were analysed. Figure 36 depicts the number of identified proteins and peptides from each sample using the 3 different spectral libraries. For all cortex and white matter samples, the largest number of peptides and proteins were identified with the project specific library.

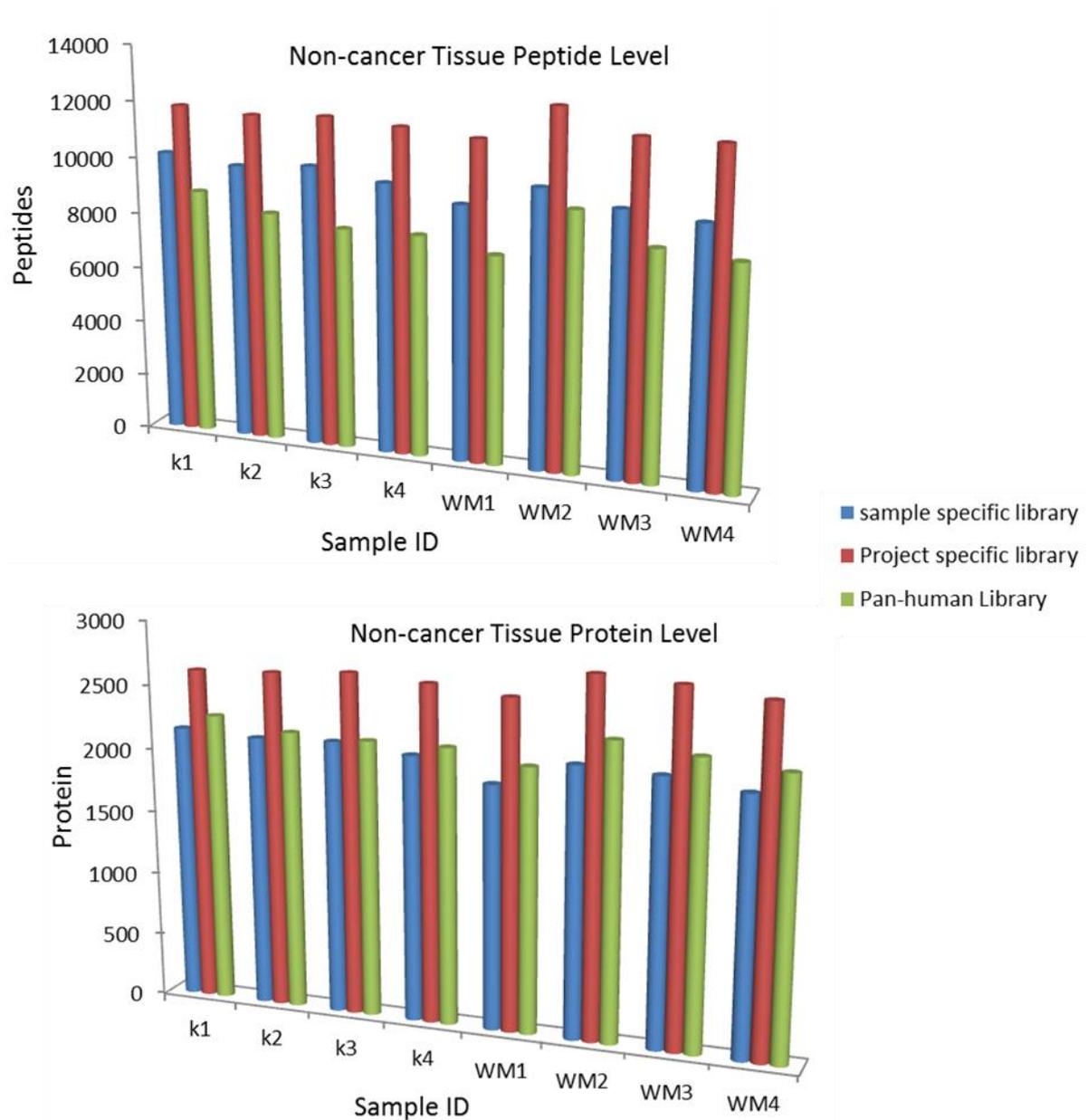


Fig 36: Spectral library influence on SWATH MS post processing. Effects on number of (a) identified peptides (b) identified proteins. 4 epileptic brain tissues from the cortex (K1, K2, K3, and K4) and 4 from white matter (WM1, WM2, WM3 and WM4) were analysed with a DIA method using Q Exactive mass spectrometer.

3.3.2 Non-tumour brain tissues: White matter and Cortex

The data obtained from the non-tumour brain tissues were processed with the project specific library. The number of identified proteins/protein groups in each sample is as shown in table 6. An average of 2900 proteins could be identified in the white matter samples and 2900 proteins in the cortex.

Table 6: Number of identified proteins in different epileptic brain tissues from different brain areas

Tissue type	Sample ID	Number of protein groups	Number of identified proteins
Cortex	K1	2621	2695
	K2	2569	2638
	K3	2598	2673
	K4	2556	2630
White Matter	WM1	2490	2568
	WM2	2693	2776
	WM3	2664	2736
	WM4	2589	2665

Venn diagrams were also generated to compare the variability between each epileptic subgroup. A slightly higher number of proteins/protein groups were identified in the cortex tissues in comparison to the white matter. A PCA plot was also generated to show the variation of tissues between each tumour type. 78% of protein groups (2261) were identified in all four cortex patient tissue samples and 85.8% (2486) in at least 3 of 4 samples. With white matter, 70% of protein groups (2141) were identified in all four tissue samples and 81.6% (2495) in at least 3 of 4 samples.

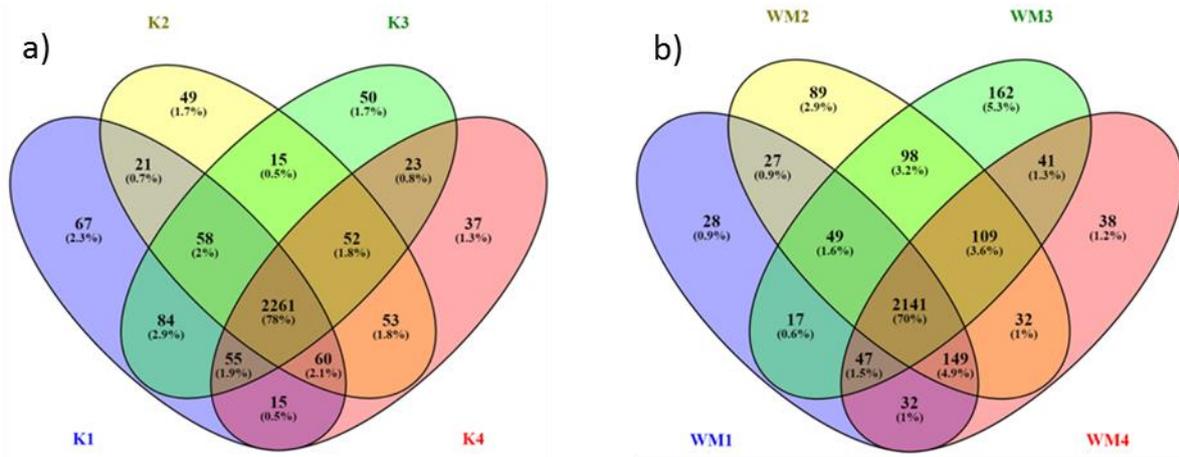


Fig 37: Venn diagram of the overlap of the identified protein groups in each patient within the extracellular vesicles obtained from exosome. A) Cortex B) White matter

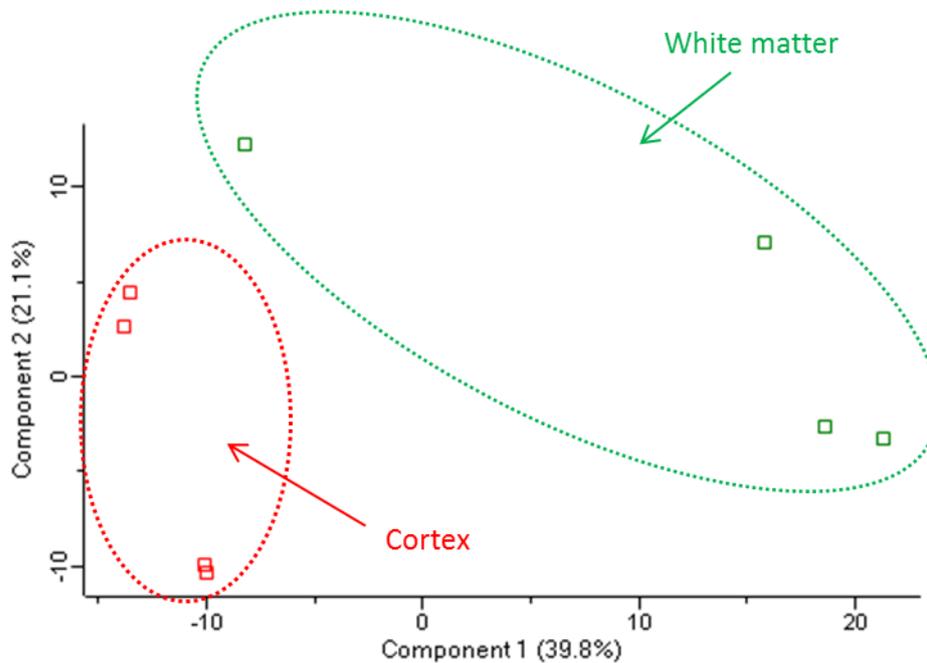


Fig 38: Principal component Analysis (PCA) clustering of all glioblastoma (green) and all meningioma (red) extracellular vesicles. Component 1 and 2 are shown for each sample.

3.3.3 Comparison of cancerous and non-cancerous brain tissues

The data obtained from white matter, cortex, GBM and MEN tissues were all processed with the project specific library detailed under section 5.3.1. Post processing was done with the Spectronaut software and the analysis report exported to Perseus via Excel for statistical analysis. The goal was to compare the changes in proteome between the tumour tissues and the non-tumour tissues. Figure 39 shows an overview of the raw intensity distribution and clustering between the different glioblastoma, meningioma, white matter and cortex samples.

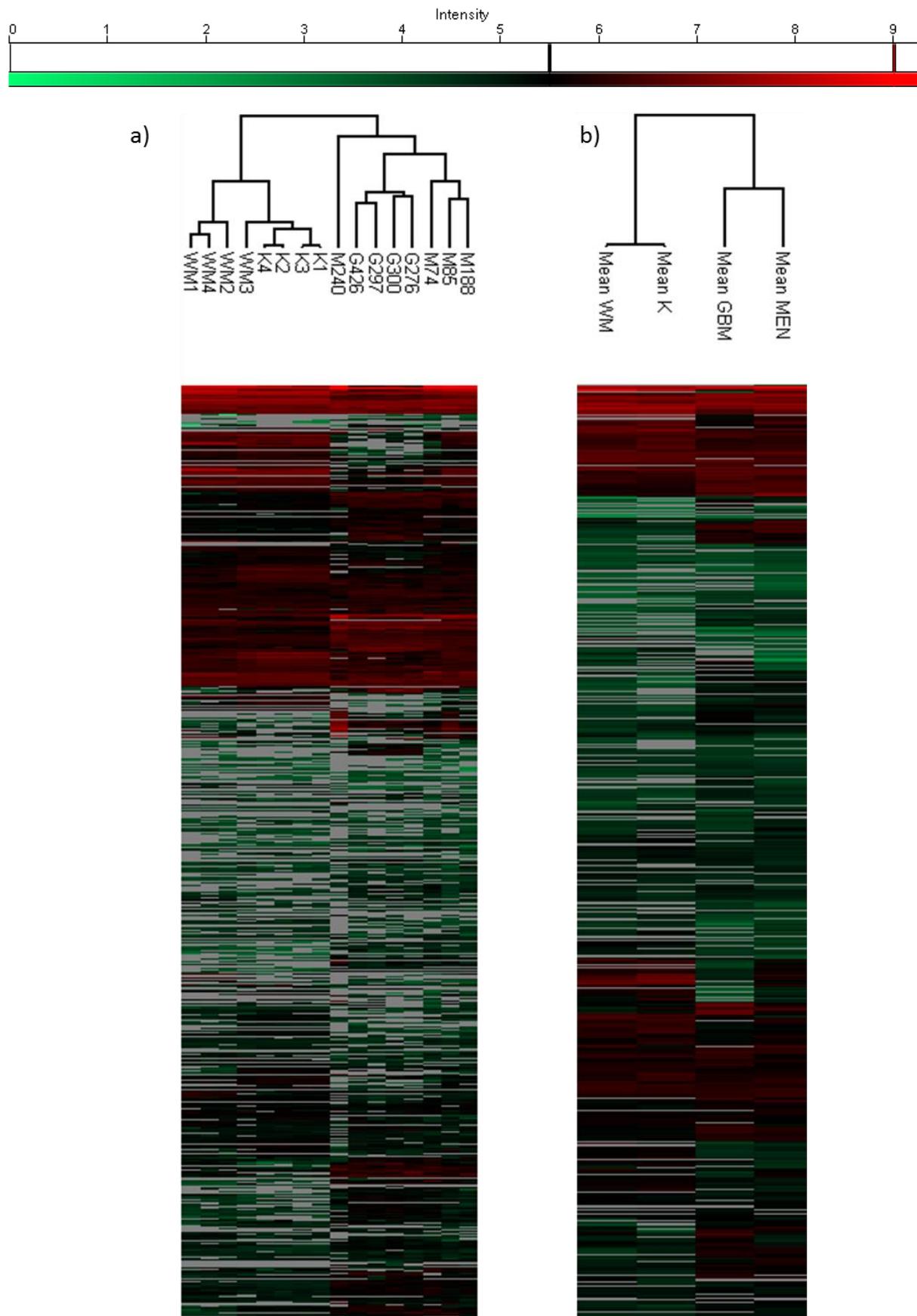


Fig 39: Log₁₀ Intensity distribution and clustering of various tumour and non-tumour replicates. a) Clustering of white matter (WM); cortex (K); Meningioma (M) and Glioblastoma (G) replicates. b) Mean intensities.

3.3.3.1 Meningioma (MEN) versus non-tumour tissues.

To view these differences on a quantification basis, only proteins which were quantified in both MEN tissues and non-tumour tissues were further analysed. With this in consideration, 2110 proteins were found in three of four replicates of both meningioma tissues and cortex tissues. 591 of these proteins were significantly up regulated with greater than 2 fold change in meningioma tissue as oppose to the cortex tissues and 613 were significantly down regulated with greater than 2 fold change. For the white matter samples, 2173 proteins were quantified in both WM and MEN samples; of which 518 were significantly up regulated and 389 were significantly down regulated. Figure 40 shows the intensity regulation between MEN and the different control samples.

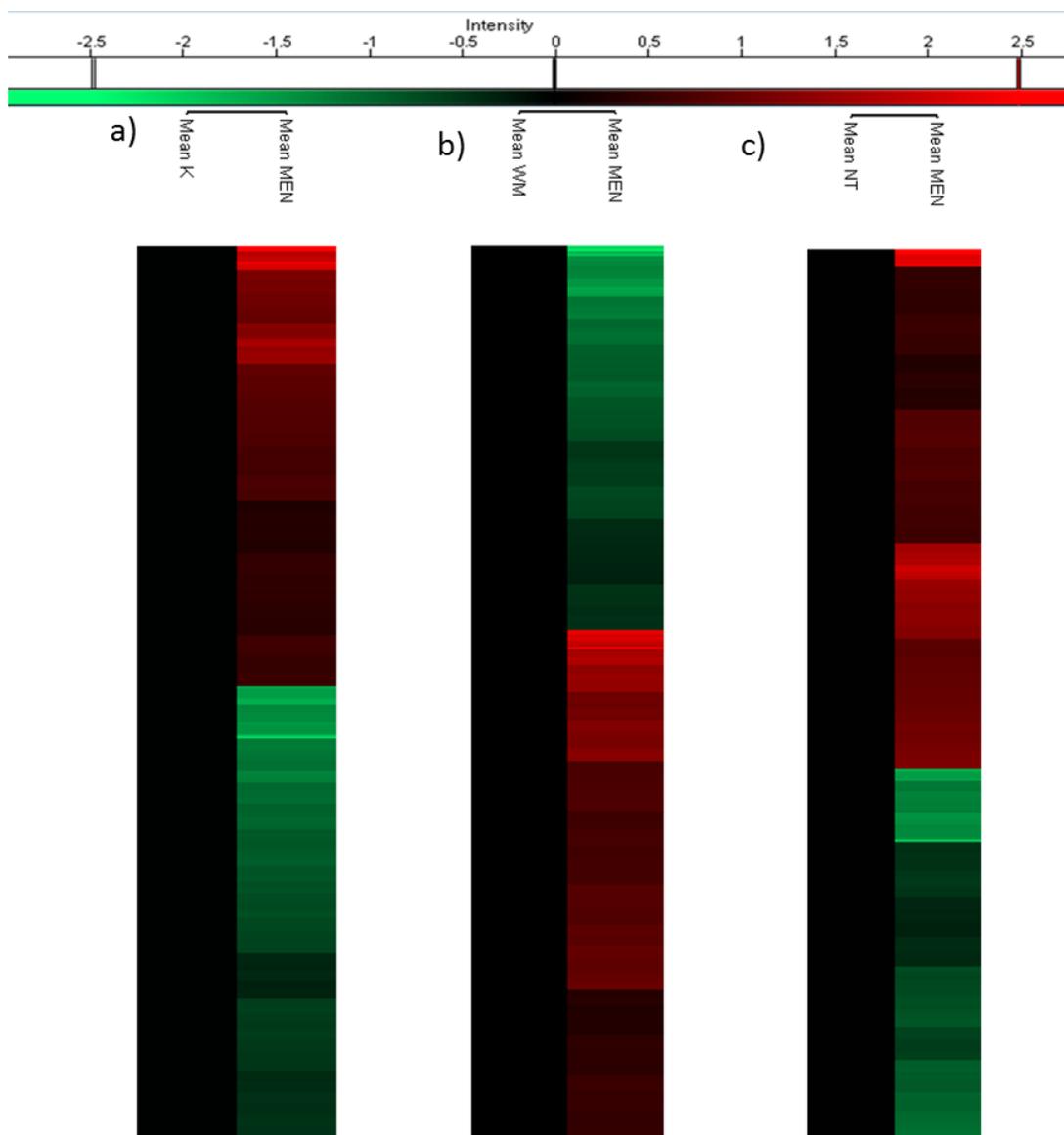


Fig 40: Log₁₀ Mean Intensity distribution and clustering after a 2 sample T-test of proteins identified in all samples of a) Meningioma (MEN) and cortex (K); b) MEN and white matter (WM); c) MEN and all non-tumour (NT) tissues.

3.3.3.2 Glioblastoma (GBM) tissues versus non-tumour tissues.

Proteins which were quantified in both GBM tissues and non-tumour tissues were further analysed. With this in consideration, 1992 proteins were found in three of four replicates of both glioblastoma tissues and cortex tissues. 724 of these proteins were significantly up regulated with greater than 2 fold change in meningioma tissue as oppose to the cortex tissues and 575 were significantly down regulated with greater than 2 fold change. For the white matter samples, 2060 proteins were quantified in both WM and GBM samples; of which 723 were significantly up regulated and 394 were significantly down regulated. Figure 41 shows the intensity regulation between GBM and the different control samples.

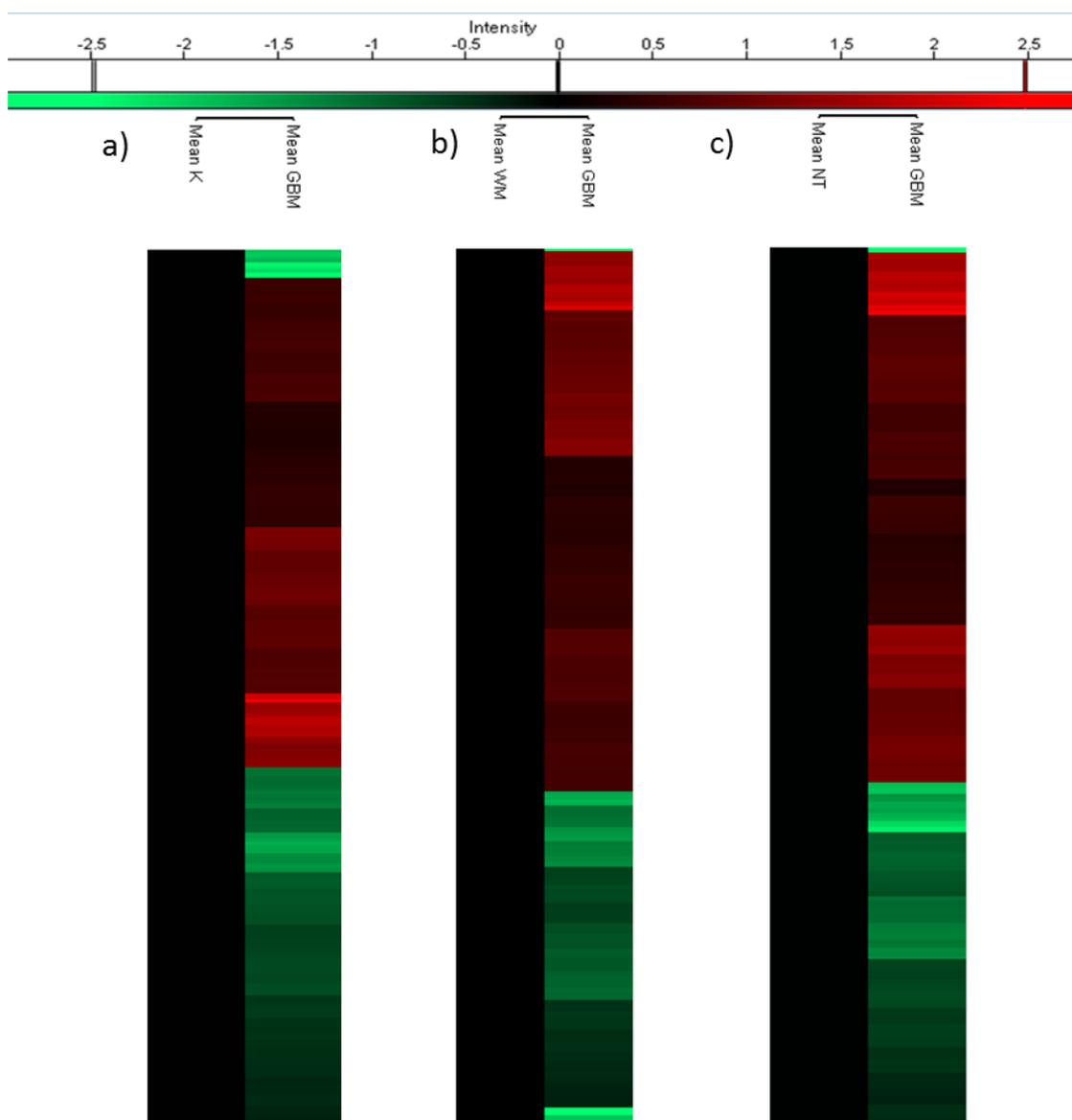


Fig 41: Log₁₀ Mean Intensity distribution and clustering after a 2 sample T-test of proteins identified in all samples of a) Glioblastoma (GBM) and cortex (K); b) GBM and white matter (WM); c) GBM and all non-tumour (NT) tissues.

3.4 Differences between Data Dependent Acquisition and Data Independent Acquisition

The differences between the data obtained from Data dependent and Data independent acquisition methods of mass spectrometry was compared based on the number of identified proteins. The raw files from DDA analysis were processed with the MaxQuant software and the filtering criterion was based on the number of identified unique peptide per protein. Only proteins with unique peptides ≥ 2 were documented. Data obtained from DIA analysis was processed as described under sections 5.1.1, 5.1.2 and 5.1.3. The identified number of proteins in each patient's tissue, cells and EVs are as listed in Table 7 and these differences can be visualised in figure 42.

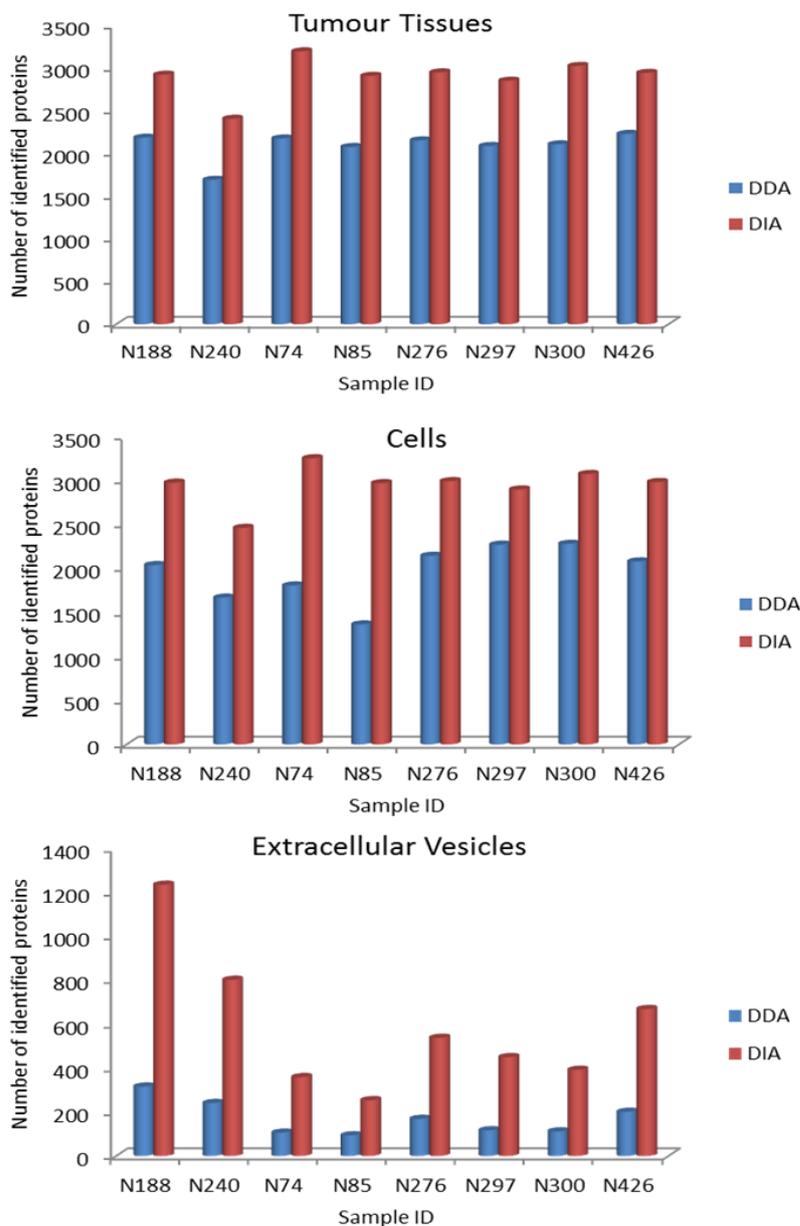


Fig 42: Differences in the number of identified proteins from tryptic digested peptides of each patient tissues, cells and extracellular vesicles between DDA MS (blue) or SWATH DIA MS (red) using Q Exactive mass spectrometer

Table 7: Number of proteins identified in each patient sample with DDA and DIA modes of Mass spectrometry

Sample type	Tumour type	Sample ID	DDA protein number	DIA protein number
Tumour Tissue	Meningioma	N188	2183	2921
		N240	1690	2405
		N74	2174	3194
		N85	2075	2907
	Glioblastoma	N276	2152	2948
		N297	2086	2850
		N300	2106	3023
		N426	2228	2942
Primary Cell Culture	Meningioma	N188	2040	2978
		N240	1669	2463
		N74	1807	3255
		N85	1364	2972
	Glioblastoma	N276	2145	2995
		N297	2272	2900
		N300	2281	3076
		N426	2081	2985
EVs	Meningioma	N188	316	1235
		N240	241	802
		N74	106	359
		N85	94	254
	Glioblastoma	N276	169	538
		N297	117	450
		N300	112	392
		N426	201	669

4. Discussion

Intracranial tumours or commonly referred to as brain tumours throughout this study, belong to the cancer disease subset whose causes remain widely unknown, prognosis very poor with a very low survival rate. Primary malignant brain tumours such as gliomas are the most prevalent type of adult brain tumour accounting for over 78% of malignant brain tumours, with glioblastoma multiforme (GBM) being the most invasive with only 6.4% survival rate of up to two years. Primary benign brain tumours like meningiomas (MEN) are less invasive but however, still possess life threatening characteristics. Therefore, uncovering subtle differences within cancer disease categories and identifying biomarkers to aid early diagnosis, prognosis and proper treatment will improve clinical management and ultimate long-term outcomes of patients. Obtaining biopsies for disease diagnosis is commonly practised in clinics worldwide today. However, traditional biopsy methods like stereotactic biopsies have demonstrated some severe shortcomings, pointing to the possibility of it increasing mortality and morbidity in patients with gliomas [65]. Other limitations to stereotactic biopsies include; access to limited amounts of tumour tissue for pathologic and molecular diagnoses; tissue biopsy is also subject to sampling bias; and tissue from a single tumour location may fail to capture intra-tumour heterogeneity. In efforts to salvage these limitations, research is now focusing on the sampling of the “liquid biome”. Evidence of extracellular vesicles (EVs) involvement in multiple facets of cancer biology has been proven. They have been thought to aid metastasis and even facilitate cell drug resistance. Due to easy accessibility, they can be used as biomarkers for cancer which will enable the use of liquid biopsies for early detection and personalised cure. Their ability to transfer molecular cargo and be selectively taken up by specific cells is also being exploited for cancer therapy. The overall aim of this study was to explore the possibility of extracellular vesicle proteins as potential biomarkers for early and non/minimally invasive diagnosis of primary brain tumours. The primary brain tumours studied were Glioblastoma multiforme (GBM) and meningioma (MEN). Therefore, this study (i) has identified and quantified proteins from GBM and MEN patients’ tissue, cells and EVs, (ii) Identified potential biomarkers for MEN and GBM which are found in EVs hence establishing the possibility of liquid biome diagnosis, (iii) Identified the differences between tumour tissues and primary cell culture of MEN and GBM samples on proteome level, (iv) Identified proteome differences between GBM and MEN tissues from non-cancerous tissues and (v) highlighted the differences between DDA and SWATH-DIA methods of protein quantification based on these samples.

4.1 The choice of DIA analysis with SWATH over DDA for this study

Data dependent acquisition mode of mass spectrometry has been used over the years for the quantitative analysis of proteomes; be it label free or in combination with labelling methods such as SILAC. Due to the extra costs incurred for labelling experiments, or the inability to carry out some labelling techniques such as metabolic labelling on certain samples especially of human origin, label free quantification has gained grounds and is often used in most large-scale experiments. Other label free methods like data independent acquisition have been developed to improve and acquire more reliable quantification data. In this study, eight patient tumour tissues (4 GBM and 4 MEN), their cells obtained from primary cell culture and Extracellular vesicles were analysed with both DDA and SWATH DIA modes. Figure 42 and table 7 details the differences between the two acquisition modes on the various samples. In all samples, the number of quantified proteins in the DIA mode was greater than in the DDA mode with some even twice as much in the EVs and some cell samples. The difference in the number of quantifiable proteins between these two methods can be attributed to the following: firstly, the MS2 spectra data obtained from data dependent acquisition mode greatly depends on the precursor ion selected for fragmentation. This selection process is mostly intensity based, whereby the highly abundant ions are selected whereas the lower abundant peptide ions do not make the cut [101]. It also relies on the probability that a precursor ion will be acquired at the apex of its chromatographic elution profile, which is improbable [102]. Unlike DDA, all peptide ions scanned in the first quadrupole during DIA are fragmented and MS2 spectra recorded [105]. Therefore, the probability of missing certain peptide ions which could be potential unique peptides for identification and quantification in DDA mode especially in complex samples is higher and this could account for the lower number of quantifiable proteins in each sample. Secondly, Label-free quantitative techniques have until recently been based entirely on integrated ion intensity measurements of precursors in the MS1 scan, or in the case of spectral counting the number of assigned MS2 spectra [108]. For DIA experiments, data analysis is based on peptide centric scoring and spectral library from which peptide query parameters (PQPs) are obtained. These PQPs contain information about: the underlying protein, peptide sequence, precursor m/z , fragment m/z , precursor and fragment charge, fragment ion type, expected relative fragment ion intensities and normalized retention time from which data acquired from SWATH-MS experiments are inferred [104]. Although there now exist options to increase the number of identified peptides in DDA experiments, by transferring peptide identifications to unsequenced or unidentified peptides by matching their mass and retention times (“match-between-runs” feature in MaxQuant) [109], however, the number of missing values in DDA data sets still remains higher than for data acquired in SWATH-MS mode, especially for peptides and proteins in the low concentration range. SWATH

data acquisition can be done either by varying the size of the Q1 window (Dynamic DIA windows) or keeping an equidistant window (static DIA windows) with the latter being the most common. Comparative studies have been done between these options and the findings were independent of these customizable window sizes as they performed similarly in terms of the number of quantified proteins and quantitative precision[110]. Figures 8 and 34 demonstrate the precursor m/z population distribution over the m/z scan range of 400 to 1330 for all samples investigated in this study. These indicate majority of the precursor ions falling in the range m/z before 900 Th, making this area to be highly dense. Therefore, a variable window SWATH method was used throughout this study whereby smaller Q1 windows were used in dense m/z areas where higher precursors were measured whereas wider Q1 windows were implemented in less dense m/z regions. Hence, the method of choice throughout this study was SWATH DIA with variable/dynamic windows.

4.2 Comprehensive versus sample specific libraries

Data generated from SWATH MS experiments are highly complex; hence the need for sophisticated bioinformatic tools to deconvolute the signals and the data mapped to peptide/protein databases for identification/quantification. Spectral libraries are essential with PQPs for SWATH data analysis as explained under section 4.1. The idea of obviating the spectral library generation step during experiments has given birth to some publicly available specie specific libraries available on the SWATHAtlas webpage [106, 111-114]. In the first part of this study, the potential of different libraries was compared based on the number of identifiable/quantifiable peptides and proteins in each sample. Figures 10 & 11 detail the outcome of three libraries (Sample specific: which was only information from either EVs or cells or tumour tissues, Project specific: which contained information from all cells, tissues and EVs and the publicly available Pan human library) on GBM and MEN EVs, cells and tissues on peptide and protein level respectively. The project specific library was best in 66.7% of the samples at peptide level and 50% at protein level. In the next part of the study, the potential of sample specific (information from non-cancerous white matter and cortex), project specific (information from GBM, MEN, white matter and cortex tissues) and the pan human libraries were also compared on tumour and non-tumour tissues (figure 36). The project specific library still performed better in 100% of the samples on both peptide and protein levels. The fact that the project specific library performed better than the Pan human library could possibly be attributed to factors such as: differences in instrument type, instrument methods (both chromatographic and mass spectrometric parameters) and even inter laboratory dependencies. The variation of chromatographic factors such as column dimensions and gradient length between laboratories can be greatly narrowed with the use of a set of reference iRT-peptides for standardisation [115]. In both

experiments for the generation of the project specific and Pan human libraries, retention time normalization was done with the commercially available peptides from the iRT Kit (Biognosys AG), adding it to all samples prior to MS injection according to vendor instruction. However, errors resulting from disparities in other factors like mobile/stationary phases or column temperature may significantly affect the peptide elution[104]. Also, the mass spectrometric methods and instruments used differed in both experiments. This could possibly generate different fragment ions as those generated on same instruments are best comparable. In 2017, Collins *et al* published a multi laboratory study based on the sensitivity, quantitative and qualitative assessment of SWATH MS[116]. Their findings indicated that acquisition of reproducible quantitative data by multiple laboratories worldwide was achievable, suggesting that organism-scale spectral libraries can effectively be used to analyse SWATH-MS data. However, all laboratories in this study used the same mass spectrometer while the nanoLCs consisted of various models from the same vendor and the chromatographic columns had the same dimensions.

4.3 Potential biomarkers for cancer liquid biopsy

Extracellular vesicles (EVs) perform a plethora of pathological functions as they aid in intracellular communication and molecular exchange by facilitating the “packaging and shipment” process between cells[73]. The correlation between tumour cell and tumour-EV proteome across multiple tumour contexts has highlighted the potential for tumour-EVs as candidate markers for disease diagnosis and monitoring[117]. In this study, proteomes of EVs, cells and tumour tissues from GBM and MEN patients were compared to identify potential cancer specific EV protein biomarkers. In the MEN EVs, 64 proteins were identified in the top 100 proteins which are often identified in EVs while 72 of these proteins were identified in GBM EVs. From the top 20 proteins, 19 were identified in both GBM and MEN EVs. This indicates the quality of the data was quite good. Cytochrome c oxidase which is a marker for mitochondria membrane and apoptotic blebs[118] was not identified in any of the cancer EV samples. This suggests the separation and purity of the EVs of all eight samples was great as non-vesicular macromolecule contamination was not identified. In the MEN patient samples, 270 proteins were identified and quantified in all extracellular vesicles, cells and tumour tissues while 262 proteins were quantified in all GBM samples. Among these proteins, 171 of them were identified in meningioma and glioblastoma extracellular vesicles, cells and tumour tissues (figure 26). The list of potential MEN EV specific proteins (99) and GBM EV specific proteins (91) can be found in supplement Table S1 and S2 presented in descending order of abundance in reference to EVs.

In 2017, Mallawaarachy *et al* published a study suggesting a list of 145 GBM EV specific proteins which were secreted by six cell lines[119]. Among these 145 potential GBM EV specific proteins, 44 were quantified in MEN and GBM tissues, cells and EVs in this study. This therefore implies that these proteins may not be limited to GBM EVs but also other tumour EVs. Among these subset of 44 common proteins reported by Mallawaarachy *et al* is cathepsin D; which showed an increasing correlation with GBM cell invasion[119]. Cathepsin D (cath-D) is a well-characterised aspartyl lysosomal protease expressed in all tissues and has been extensively studied in relation to mammary malignancies [120, 121]. Most metastatic breast cancer cell lines, unlike normal cells, aberrantly over-produced pro-cath-D due to the overexpression of the cath-D gene in both malignant and extra-tumoural cells such as macrophages and fibroblasts [122, 123]. Studies have also shown that the immunohistochemical localization of Cathepsin D expression in histopathologic sections from breast cancer patients correlated with disease progression but its amounts in the serum based on disease progression was contradictory[122].

The next protein subset of interest is 15 proteins which were identified by the Mallawaarachy *et al* study as potentially GBM EV specific and were also identified only in GBM tissues, cells and EVs. These proteins are therefore postulated here as potential GBM EV specific protein markers with each protein function discussed in detail below.

Annexin A1

Annexin A1 was first described in the late 1970s and subsequently has been referred to as macrocortin, renocortin, lipomodulin, lipocortin-1 and more commonly Annexin 1[124]. It is a 37-kDa protein, belongs to the annexin superfamily of calcium-dependent phospholipid-binding proteins[125]. It participates in a variety of important biological processes, such as cellular transduction, membrane aggregation, inflammation, phagocytosis, proliferation, differentiation and apoptosis [126, 127]. Dysregulation of ANXA1 levels and alterations to its sub-cellular localization have been associated with the development, invasion, metastasis, occurrence and drug resistance of a large number of cancers including lung [128, 129], breast [130], pancreatic [131], colorectal [132], prostate[133] and even astrocytic cancers [134]. Annexin A1 has been shown to be upregulated in some tumours but downregulated in others, making its mechanism of participation in carcinogenesis and tumour progression complex and unclear although it has been suggested as a potential therapeutic target in the treatment of malignant disease[135].

Galectin-3-binding protein

Gal-3 binding protein (Gal-3BP), also known as Mac-2 binding protein (Mac-2BP) or tumour-associated antigen 90K (TAA90K), is a glycoprotein without a Trans membrane domain [136]. It is a large hyperglycosylated protein that acts as a ligand for several galectins through glycan-dependent interactions and induces galectin-mediated tumour cell aggregation to increase cancer cell survival in the bloodstream during metastasis[137]. It has been associated with breast cancer [136], pancreatic cancer [138] and also known to be secreted by neuroblastoma cells [139].

Talin-1

Talin-1, a macromolecular cytoskeletal protein concentrated at regions of cell-substratum contact, which has been reported to interact with multiple adhesion molecules (e.g. integrin and F-actin) and to activate the integrin/focal adhesion kinase (FAK) pathway[140]. It has been associated with cancer as it reportedly promotes hepatocellular carcinoma (HCC) [141] and oral squamous cell carcinoma [142] progression. Data about its association to intracranial tumours or specifically GBM if any; is still unavailable.

Filamin-B

Filamin-B (FLNB) is an Actin-binding protein which is expressed in endothelial cells and plays an essential role during vascular development, connecting cell membrane constituents to the actin cytoskeleton [143]. It is a key player in chondrocyte progenitor differentiation for endochondral ossification and it is involved in the development of the skeleton before birth[144]. It also aids in cell migration, vascular development, extracellular signalling and activity of integrins [145]. Its role in cancer progression is still unreported, however, it has been suggested as a potential biomarker for prostate cancer [146].

Ras GTPase-activating-like protein IQGAP1

IQ-domain GTPase-activating proteins (IQGAPs) are an evolutionary conserved family of multi-domain proteins that regulate distinct cellular processes including cell adhesion, cell migration, extracellular signals and cytokinesis [147]. IQGAP1 is the most studied and the first of three human IQGAP homologues discovered and it is ubiquitously expressed[148]. It is a plasma membrane-associated scaffolding protein and an important regulator of the dynamics and assembly of the actin cytoskeleton, contributing to cell migration, polarity and adhesion [149]. As a scaffold protein, IQGAP1 interacts with multiple proteins to exert various roles in carcinogenesis. The dysregulation of IQGAP1 levels and alterations to its sub-cellular localization have been associated

with the development, invasion and metastasis number of cancers including breast [150], colorectal [151], and ovarian [152] cancers. In patients with gliomas, elevated levels of IQGAP1 have been shown to correlate with tumour grades and poor overall survival [153]. Its potential as a valuable GBM prognostic marker has also been demonstrated [154] with other studies also showing a correlation between elevated levels of IQGAP1 and increase in tumour aggressiveness and invasion [119].

Protein disulfide-isomerase

Protein disulfide isomerase (PDI) is a 57-kDa multifunctional protein belonging to the thioredoxin superfamily[155]. It has been extensively studied and reported as it is associated with many diseases. It plays a vital role in maintaining cellular homeostasis by mediating oxidative protein folding [156]. PDI predominantly resides in the endoplasmic reticulum (ER), although it can be released to function at the cell surface or extracellular[157]. It functions as a molecular chaperone and a matrix disulphide oxidoreductase/isomerase, catalysing disulphide bond formation, breakage and rearrangement in the endoplasmic reticulum [156, 157]. It is therefore a central component for protein folding, post-translational modifications, and quality control of proteins. Analysis of microarray data sets revealed that PDI (PDIA1) is significantly and universally over-expressed in a wide variety of cancer types including brain, lymphoma, kidney, ovarian, prostate and lung cancers[158]. Their oncogenic effects have been suggested to be mediated by their role in the unfolded protein response signalling pathway [156, 158]. Many robust PDI activity assays have been developed and several PDI inhibitors identified, but none of the anti-cancer inhibitors have been approved for clinical use.

Vinculin

This is 117 kDa protein which is an abundant, prominent and a well-characterised F-actin binding protein localised in focal adhesions as well as in cell-adherence junctions[159]. It is a major regulator of cell adhesion, motility, as well as cell spreading. Vinculin attaches to the cell surface by binding to specific phospholipids and several actin-organizing proteins. In mature focal adhesions, vinculin is a key component of the “molecular clutch” that mediates the transmission of force from cytoplasmic F-actin to membrane-bound integrins[160]. It has been linked to tumours as it enhances the conversion of Phosphatidylinositol 4,5-bisphosphate to phosphatidylinositol 1,4,5-trisphosphate by phosphoinositide 3-kinase, leading to enhanced Akt signalling and thereby promoting tumour progression [160, 161]. Increased cell proliferation through the receptor tyrosine kinase/Ras/PI3K/Akt signalling pathway due to the overexpression of EGFR and mutations of EGFR

gene has been detailed by many studies as one of the most characteristic features of glioblastomas [20, 21]. EGFR is currently a diagnostic marker for glioblastomas although a standard therapy is not yet available due to many failed clinical trial attempts with EGFR inhibitors [27]. Therefore, vinculin could possibly be a great potential biomarker.

14-3-3 protein theta and 14-3-3 protein gamma

14-3-3 proteins comprise a family of small acidic proteins (~30 kDa), discovered in 1967 during a systemic classification of brain proteins [162]. There exist seven isoforms in mammals (β , γ , ϵ , η , ζ , σ , and τ/θ) with each having its physiological functions as potential regulators of diverse signalling pathways, cell cycle, transcription, apoptosis and neuronal development [163, 164]. They are ubiquitously expressed in various types of tissues, but their highest expression is in the brain, where they make up approximately 1% of its total soluble proteins. They have been associated with many neurological diseases like schizophrenia [165], Parkinson's [166], Alzheimer's [167] however, no clinically significant mutations linked to brain cancer have been reported. Nonetheless, 14-3-3 theta has been reported as an antigen that induces a humoral response in lung cancer [168, 169] and also as a predictive biomarker of resistance to neoadjuvant chemotherapy in luminal breast cancer [170].

Inter-alpha-trypsin inhibitor heavy chain H2

ITI H2 belongs to the inter-alpha-trypsin inhibitors (ITI) family of structurally related plasma serine protease inhibitors involved in extracellular matrix stabilization and in prevention of tumour metastasis [171]. Studies have shown that overexpression of ITI family chains leads to inhibition of tumour development and/or metastatic spreading [172, 173]. In 2006, Werbowetski-Ogilvie *et al* isolated and identified inter α -trypsin inhibitor heavy chain 2 (ITI H2) as a strong natural inhibitor of brain tumour invasion [174]. Overexpression of ITI H2 in their U251 glioma cell line confirmed its inhibitory role in malignant glioma invasion and revealed an inhibitory effect on proliferation with a concomitant increase in cell adhesion. Therefore, ITI H2 could serve both as a potential indicator of tumour malignancy and as a novel target for therapeutic intervention.

Thrombospondin-1

Thrombospondin-1 (TSP-1) is a large glycoprotein secreted by platelets and synthesized by many cell types, including endothelial and tumour cells. It is a matricellular, calcium-binding protein that participates in cellular responses to growth factors, cytokines and injury, regulating cell proliferation, migration and apoptosis in a variety of physiological and pathological settings, including wound healing, inflammation, angiogenesis and neoplasia [175]. Overexpression of TSP1 has been reported to suppress growth or metastasis of some tumours in vivo and inhibits angiogenesis and also

increased tumour progression in others[176]. These opposing roles in both tumour progression and inhibition are secondary to the action of the various domains of TSP-1, local milieu, and tumour cell type[177]. In gliomas however, Thrombospondin-1 was found to be increased with glioma grades as TSP-1 silencing inhibits tumour cell invasion and growth, alone and in combination with anti-angiogenic therapy [178].

Transitional endoplasmic reticulum ATPase

Transitional endoplasmic reticulum ATPase commonly known as valosin-containing protein (VCP) or p97 belongs to the type II AAA (ATPases associated with diverse cellular activities) family[179]. P97 is suggested to be a ubiquitin-selective chaperone and its key function is to disassemble protein complexes [180]. No detailed reports linking VCP to brain tumours or GBM was found.

Prothrombin

Prothrombin, or coagulation factor II, is abundantly present in the blood where it circulates at a concentration of 0.1 mg/ml and a half-life of about 60 h[181]. In the coagulation cascade, prothrombin is proteolytically converted to the active protease thrombin by the prothrombinase complex and Thrombin then catalyses the conversion of fibrinogen to an insoluble fibrin clot[182]. The abnormal form of prothrombin known as Des- γ -carboxy prothrombin (DCP) induced by the absence of vitamin K₂ has been reported to increase in the serum of patients with hepatocellular carcinoma (HCC) and appears to be a useful tumour marker for the evaluation of patients with HCC [183, 184]. However, a correlation between Prothrombin and brain tumour cells was not found.

Neutral alpha-glucosidase AB

Neutral alpha-glucosidase AB is also known as Alpha-glucosidase 2 or Glucosidase II subunit alpha. It is involved in the pathway N-glycan metabolism and functions in the quality control mechanism of glycoprotein folding[185]. Glucosidase II is a resident protein of the endoplasmic reticulum and a heterodimeric enzyme that cleaves sequentially the two innermost alpha-1,3-linked glucose residues from N-linked oligosaccharides on nascent glycoproteins[186, 187]. Its relationship to brain tumours still remains vague however; it has been suggested as a potential biomarker for lung cancer [185].

Alpha-2-macroglobulin

Alpha-2 macroglobulin (A2M) is a homotetrameric protein of 720 kDa and a major extracellular protein in the blood, playing a role in maintaining homeostasis of cytokines and growth factors [188, 189]. Alpha-2-Macroglobulin controls practically all known proteinases, regulating all proteolytic processes including immune reactions, as it has the broad spectrum of protection, including antitumor effect. Malignant tumours are able to synthesize and release proteins into the environment, including α 2M that protects the tumour against immune system attacks[190]. In Glial tumours, α 2-Macroglobulin has been reported to inhibit the malignant properties of Astrocytoma cells by impeding β -Catenin signalling [191].

The functional interactions, molecular actions and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway relations between these 15 proteins which have also been previously reported as potential GBM EV specific proteins was further analysed with the STRING database version 11.0. The STRING database consolidates known and predicted protein–protein association data for a large number of organisms[192]. A diagram was generated (figure 43) detailing the type of interactions between these proteins, those associated with vesicle-mediated transport, directly linked to the PI3K-Akt signalling pathway and those belonging to the proteoglycans in cancer family.

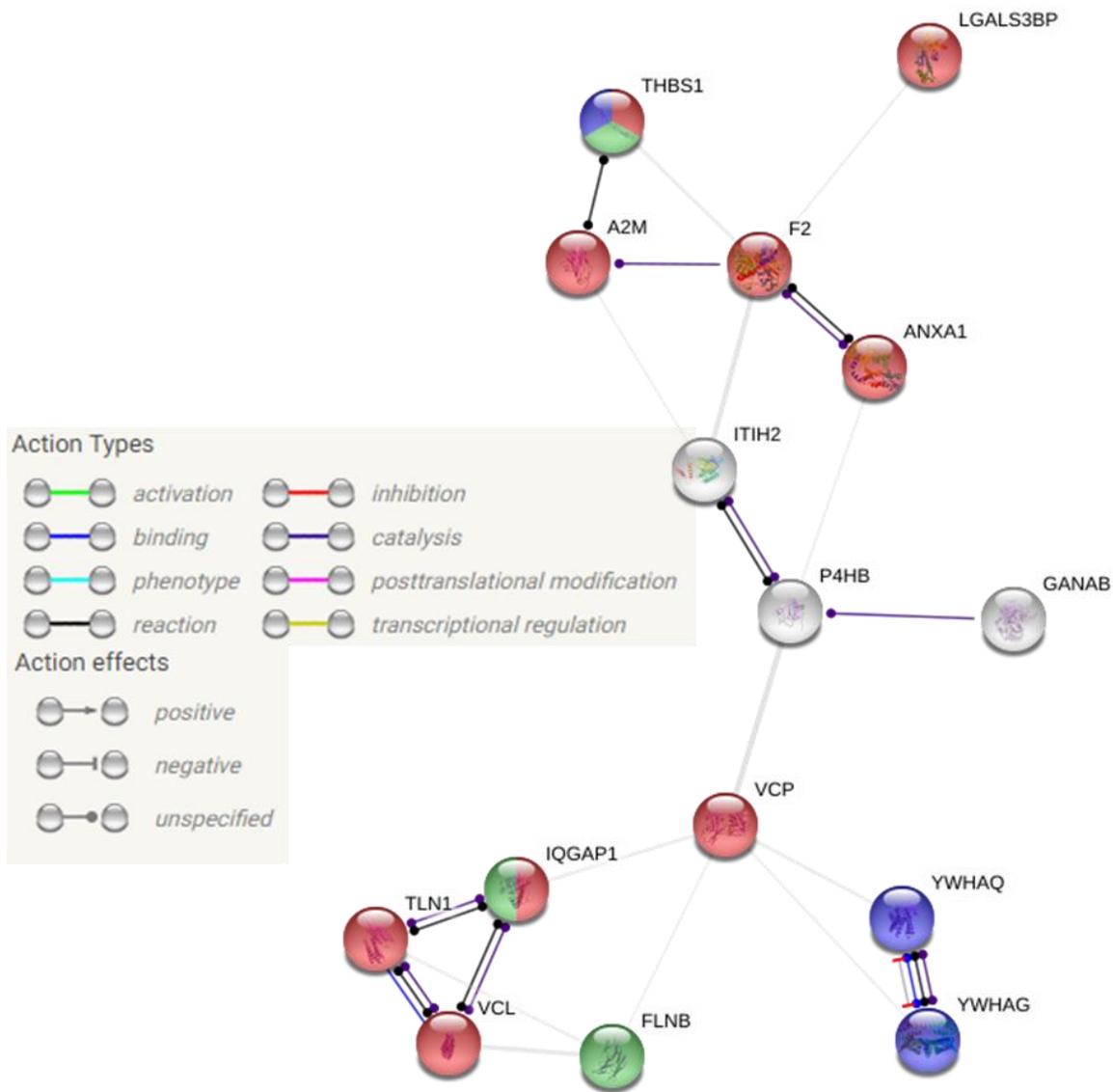


Fig 43: Predicted protein-protein interactions generated following analysis of potential GBM EV specific proteins input into the STRING database which are postulated as potential EV markers for GBM liquid biopsy. Red nodes indicate proteins associated with vesicle-mediated transport; purple nodes highlight proteins linked to the PI3K-Akt signalling pathway while the green nodes represent proteoglycans in cancer.

Proteins and signalling pathways regulating cell growth, differentiation and development typically undergo drastic changes in oncogenic states and collectively cause the cell to become malignant. Among these are the PI3K/Akt signalling pathways which have been reported to be frequently disturbed in many human cancers [193]. The proteins in purple nodes; 14-3-3 protein theta (YWHAQ), 14-3-3 protein gamma (YWHAG) and Thrombospondin-1 (THBS1) were reported to be associated with these pathways according to the KEGG pathway in STRING database. Many protein inhibitors have been developed to target these pathways and their downstream growth factor targets as they do not only induce different cell survival mechanisms, but also play a role in the tumour's potential response to cancer treatment[21]. A biological process of interest is

vesicle-mediated transport. The proteins reported here to be associated with this process are: Vinculin (VCL), Talin-1 (TLN1), Ras GTPase-activating-like protein (IQGAP1), Transitional endoplasmic reticulum ATPase commonly known as valosin-containing protein (VCP), Alpha-2 macroglobulin (A2M), Prothrombin (F2), Annexin A1 (ANXA1), Thrombospondin-1 (THBS1) and Galectin-3-binding protein (LGALS3BP). These proteins are transported in membrane-bounded vesicles; either enclosed in the vesicle lumen or located in the vesicle membrane. The presence of these proteins confirms their presence in the EVs whose formation and function is detailed in Figure 4. The last subset of proteins detailed in green in Figure 43 is made up of proteoglycans in cancer. These proteins according to the KEGG pathway incorporated in the STRING database include: Filamin-B (FLNB), Ras GTPase-activating-like protein (IQGAP1) and Thrombospondin-1 (THBS1). Proteoglycans perform multiple functions in cancer and angiogenesis by virtue of their polyhedral nature and their ability to interact with both ligands and receptors that regulate neoplastic growth and neovascularization. Although most proteoglycans promote cancer cell proliferation, adhesion, angiogenesis and metastasis, other small leucine-rich proteoglycans, such as decorin and lumican, can function as tumour repressors [194, 195]. The three aforementioned proteins have however been shown to promote tumour cell proliferation and metastasis as described in their functions above. Although Thrombospondin-1 (THBS1) either promotes or represses tumour growth depending on the tumour, in Gliomas it has been shown to increase with glioma grades and TSP-1 silencing inhibits tumour cell invasion and growth [178]. Ras GTPase-activating-like protein (IQGAP1) has been suggested as a potential GBM prognostic marker and studies have demonstrated a correlation between elevated levels of IQGAP1 and increase in tumour aggressiveness and invasion [119, 154].

The proteins are therefore suggested here as GBM EV specific or MEN EV specific markers which could possibly aid in cancer liquid biopsy and can serve as either diagnostic, prognostic or predictive markers upon further specific studies and marker validation.

4.4 Cell culture for clinical research

Cell culture has been widely used over the years as workhorses for medical research particularly in the fields of drug discovery, cancer biology, and regenerative medicine. These methods typically involve either cell lines or primary cells. Gene expression profiles in cell line models have been identified as non-identical with those for primary tissues [196]. These cells in culture are prone to genotypic and phenotypic drifting since they are subjected to very different conditions than the normal physiological state of the cells in the host organism[197]. In one part of this study, the drift between primary cell culture and patient tumour tissues was studied on proteome level (identification and quantification).

From all the proteins which were identified in both MEN tissues and primary cells, 79% had no significant change in relative protein abundance while 21% were either up regulated in cells (13.4%) or tissues (7.7%). In GBM samples, approximately 67% had no significant change in relative [195]protein abundance while 33% were either up regulated in cells (18.5%) or tissues (14.5%). The proteins up regulated in cells of both tumour types were mostly binding proteins (nucleic acid binding proteins and cytoskeletal (actin) binding proteins) while those in higher abundance in the tissues were mostly those involved in protein binding, transmembrane, ATPase and GTPase activity. Ertel *et al* studied the pathway-specific differences between tumour cell lines and normal and tumour tissue cells and observed that the most pronounced metabolic pathways up regulated in cell lines included cell nucleotide metabolism and oxidative phosphorylation[198]. Other proteins which were up regulated in primary cells include proteoglycans like Glypican-1. It is known to interact with growth factors, cytokines and enzymes through Heparan Sulphate chains, leading to tumour growth and invasion[199]. Furthermore, 672 proteins were identified only in MEN tissues but not in the primary cells and 304 proteins were identified only in the primary cells and not in the tumour tissues. For GBM, 284 proteins were identified only in GBM tissues but not in the primary cells and 421 proteins were identified only in the primary cells and not in the tumour tissues. The GO-Slim molecular functions of these unique proteins did not particularly exhibit a noticeable trend in both tumour types. Cancer cells in culture might introduce new mutations and change the cell line characteristics as the multicellular interfaces with which tumour cells interact in vivo are not replicated for cells grown in cell culture plate. They have been removed from their physiological milieu of other cell types, tissue architecture, hormonal influences, and autocrine/paracrine signals[197]. Unlike immortalised cell lines, primary tumour cells however show less variation to the original tumours, maintaining many of the important markers and functions seen in vivo as they are isolated directly from tissues, have a finite lifespan and limited expansion capacity [200, 201].

4.5 Differential proteome of cancerous and non-cancerous brain tissues

Biomarkers are a corner stone in understanding the development of diseases like cancer. Scientific research in the past decades has focused on understanding the molecular patterns behind diseases to help improve early diagnosis and discover better therapies especially for chronic diseases. Initially, the idea of biomarkers revolved around individual genes, proteins or metabolites (molecular biomarkers) linked to a disease. However, a disease is rarely a consequence of an abnormality in a single gene, due to the functional interdependencies between the molecular components in a human cell. Therefore, the idea of biomarker panels or barcodes in clinical research has taken the lead in recent years. In this section of this thesis, a comparative study was carried out between Meningioma and Glioblastoma tissues and control brain tissues (White matter and cortex). Approximately 2600 proteins could be identified in each biological replicate of either white matter or cortex tissues (Table 6) with about 85.8% reproducibility in the cortex samples and 81.6% in white matter samples for proteins identified in at least 3 of four biological replicates (Figure 37). A heatmap was generated with all intensities after a SWATH DIA experiment to understand the clustering within the samples (Figure 39). As expected, all biological replicates clustered quite close to each other with the replicates of cortex and GBM tissues having the smallest Euclidean distance between each other (Figure 39a). The White matter replicates were quite similar but for one which clustered closer to the cortex samples. One MEN replicate was equidistant between the other MEN replicates and the cluster of GBM replicates. Clustering the mean values of the intensities for the various tissue types showed a very distinguishable difference between these tissue types just as expected. The non-tumour tissues (White matter and cortex) clustered quite close to each other, while the tumour tissues clustered together though with a larger Euclidean distance (Figure 39b). This therefore implies that although there exist strong differences between both tumour tissues (GBM and MEN) there exist even stronger differences when comparing them to the non-tumour tissues. This could be backed by the fact that tumour tissues generally show an aberrant expression of a variety of genes and hence their gene products which include proteins[202].

For the Meningioma tissues, 1204 proteins were dysregulated with more than 2-fold change with respect to cortex tissues, 907 proteins with respect to white matter and 1022 when all non-tumour tissues were combined (Figure 40 a, b and c respectively). These proteins therefore suggest a global view of the differences underlying the different diseased states. Most of the proteins which were differentially regulated in the MEN tissues were proteins involved in signalling pathways regulating cell growth, differentiation and development which are quite typical for many human cancers. To further explain this point, a subset of proteins was highlighted from comparing MEN to cortex

controls (Figure 44). Proteins which are linked to cancer are indicated in the subset which was up regulated in MEN tissues (Figure 44a) and proteins linked to epilepsy where indicated in the subset which was up regulated in cortex (Figure 44b).

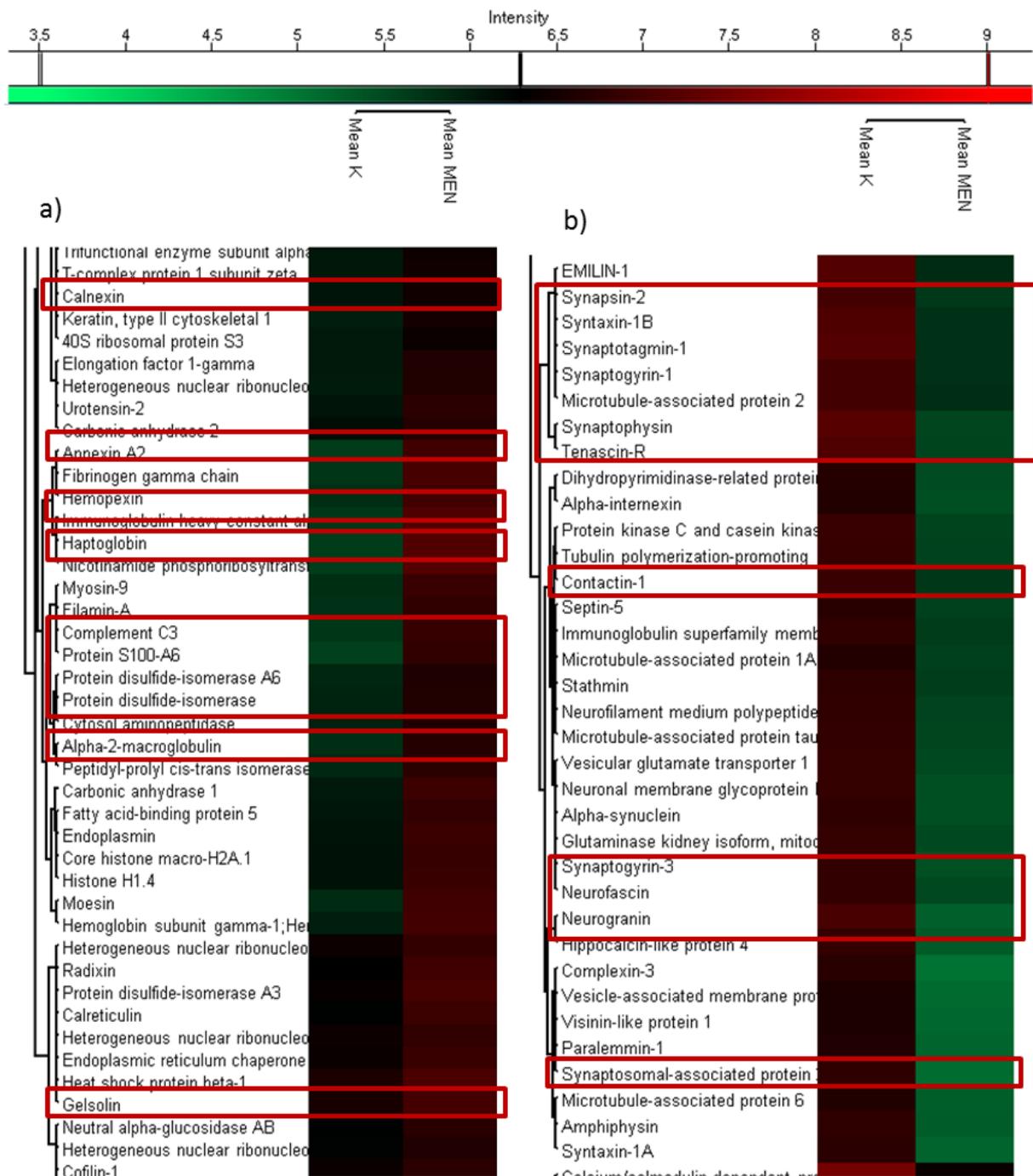


Fig 44: A subset of proteins obtained from Log 10 Mean Intensity distribution and clustering after a 2 sample T-test of proteins identified in all samples of Meningioma (MEN) and cortex (K). a) Subset of proteins up regulated in MEN tissues; b) Subset of proteins down regulated in MEN tissues. The red bars highlight interesting proteins which could possibly account for the differences underlying the different tissue health states.

Among the up regulated proteins in the MEN tissues were the following proteins which have been shown to either regulate tumour proliferation, migration, invasion or resistance to therapy: Calnexin[203], Annexin A2[204], hemopexin[205], Haptoglobin[205, 206], Complement C3[205], Protein s100-A6 also known as calcyclin[207], Protein disulphide-isomerase A6[208], Protein disulphide-isomerase[157], Alpha-2-macroglobulin[190] and Gelsolin[209]. Abbritti *et al* in 2016 carried out a bioinformatics study focusing on proteins reported in the last 10 years in PubMed as dysregulated in meningioma in tissues, serum and cerebrospinal fluid[205]. They highlighted a dysregulation in 9 proteins belonging to the pathways which showed major involvement in meningioma development and progression, plasma complement/coagulation cascades and lipoprotein particle remodelling. Six of these proteins were identified as significantly dysregulated in this study when comparing MEN tissues to both white matter and cortex tissues as controls. These six proteins include: Hemopexin, albumin, ceruloplasmin, complement C3, heptoglobulin and Apolipoprotein A1. The next subset of proteins were proteins which were down regulated in MEN tissues hence were of higher abundance in the non-tumour tissues (Figure 44b). Epilepsy in general is characterised by recurrent and unpredicted seizures, related to improper synaptic transmission between neurones[210]. Most of the proteins often involved in epilepsy are therefore synaptic proteins (Presynaptic neurotransmitter release, Postsynaptic proteins, Trans-synaptic adhesion molecules and Secreted-synaptic molecules)[211]. The highlighted proteins which have been shown to be involved in epilepsy include: Synapsin-2[212], syntaxin-1B [213], synaptotagmin-1[214], snaptophysin[215], Tenascin-R[216], contactin-1[217], neurogranin[218] and Synaptosomal-associated protein 25[219]. These proteins which were identified as dysregulated in MEN tissues therefore could serve as a guide to biomarker identification after their role in tumourigenesis and especially in meningioma is further understood.

From the work of Mallawaarachy *et al* detailed under section 4.3 above, a list of 145 GBM EV specific proteins was published which were secreted by six cell lines. Among these 145 potential GBM EV specific proteins, 44 were quantified in MEN and GBM tissues, cells and EVs in this study. These 44 proteins were analysed in this chapter to identify if they were significantly dysregulated (\geq or \leq 2 fold change) in tumour tissues in comparison to non-tumour tissues. Table 8 shows the list of these proteins with the Student's T-test Difference value after a two sample T-test. Positive values indicate proteins which are up regulated while negative values indicates the proteins down regulated in every scenario. The letter "n" indicates that this protein was identified as a potential cancer-EV specific marker but was not significantly regulated in the tumour versus non-tumour tissues. The proteins with values were identified as cancer-EV specific markers and were also regulated in the tumour tissues. These proteins therefore can serve as potential tumour markers

which can possibly be identified from liquid biopsy hence facilitating early and non/minimally invasive diagnosis of cancers which is vital for better prognosis.

Table 8: Proteins identified as cancer-EV specific markers and were also significantly dysregulated in the tumour tissues. The values represent the student's T-test Difference value after a two sample T-test, indicating the degree of regulation.

Protein name	GBM / K	GBM/ WM	MEN / K	MEN/ WM
Calreticulin	0.791929	0.811164	0.634826	0.652668
Nucleolin	0.832535	0.730259	0.621905	0.504317
Elongation factor 2	0.460792	0.413264	0.553389	0.524813
Clusterin	0.796579	0.917178	0.958129	1.07489
Serum albumin	1.17416	1.09693	1.46241	1.39541
Vimentin	1.90467	1.52052	1.21256	0.828407
Myosin-9	0.871594	0.84753	1.09752	1.05878
Histone H4	0.742805	0.459689	0.930268	0.648534
Filamin-A	1.01967	0.95846	1.02339	0.947361
Fibronectin	2.2068	1.59892	2.76377	2.1559
Collagen alpha-1(VI) chain	1.63183	1.78004	1.19854	1.34674
Complement C3	0.573912	0.405111	1.16619	1.00929
Moesin	0.930908	0.562667	1.1163	0.747714
Nucleophosmin	0.968032	0.946436	0.671795	0.640125
Fructose-bisphosphate aldolase A	-0.660135	-0.628321	-0.360867	-0.316226
Clathrin heavy chain 1	-0.791674	-0.432035	-0.468514	n
Prelamin-A/C	1.79297	1.20501	0.799185	n
14-3-3 protein zeta/delta	-0.609378	n	-0.744413	-0.485953
Adenosylhomocysteinase	-0.340375	0.430736	n	n
Glyceraldehyde-3-phosphate dehydrogenase	-0.424707	-0.536392	n	n
78 kDa glucose-regulated protein	0.639921	0.839064	n	n
4F2 cell-surface antigen heavy chain	-1.25668	-0.964626	n	n
Keratin, type I cytoskeletal 9	-0.422953	-0.826083	n	n
Keratin, type II cytoskeletal 1	n	-0.910286	0.539975	n
Gelsolin	0.649486	n	0.492905	n

Hemoglobin subunit alpha	n	n	0.907079	0.840523
Plectin	0.360993	n	n	n
Transketolase	0.308771	n	n	n
Ubiquitin-40S ribosomal protein S27a	n	n	n	-0.449411
Triosephosphate isomerase	n	n	n	-0.335056
T-complex protein 1 subunit gamma	n	n	n	n
Peptidyl-prolyl cis-trans isomerase A	n	n	n	n
Alpha-enolase	n	n	n	n
Amyloid beta A4 protein	n	n	n	n
Cullin-associated NEDD8-dissociated protein 1	n	n	n	n
Heat shock cognate 71 kDa protein	n	n	n	n
14-3-3 protein epsilon	n	n	n	n
Cathepsin D	n	n	n	n
14-3-3 protein beta/alpha	n	n	n	n
Heat shock protein HSP 90-alpha	n	n	n	n
Keratin, type I cytoskeletal 10	n	n	n	n
Programmed cell death 6-interacting protein	n	n	n	n
Heat shock protein HSP 90-beta	n	n	n	n
Phosphoglycerate kinase 1	n	n	n	n

Under section 4.3 a list of 15 proteins which were identified by the Mallawaarachy *et al* study as potentially GBM EV specific and were also identified only in GBM tissues, cells and EVs and not MEN samples were thoroughly studied. Their individual functions and involvement in critical cancer (Glioblastoma) pathways was also detailed. Here, these proteins were studied to identify if they were significantly regulated in the GBM tumours. Table 9 shows the list of these proteins with the Student's T-test Difference value after a two sample T-test. Positive values indicate proteins which are up regulated while negative values indicates the proteins down regulated in every scenario. The letter "n" indicates that this protein was identified as a potential GBM-EV specific marker but was not significantly regulated in the GBM versus cortex or white matter. The proteins with values were identified as GBM-EV specific markers and were also regulated in the GBM tissues. These proteins therefore can serve as potential GBM markers which can possibly be identified from liquid biopsy

hence facilitating early and non/minimally invasive diagnosis of Glioblastomas which is vital for better prognosis.

Table 9: Proteins identified as GBM-EV specific markers and were also significantly dysregulated in the tumour tissues. The values represent the student's T-test Difference value after a two sample T-test, indicating the degree of regulation.

Protein	GBM / K	GBM / WM
Talin-1	1.24867	0.956236
Filamin-B	2.0384	1.58991
Protein disulfide-isomerase	0.936713	1.02371
Ras GTPase-activating-like protein IQGAP1	0.693553	0.838277
Inter-alpha-trypsin inhibitor heavy chain H2	1.27048	1.22434
Vinculin	1.12205	1.06863
Annexin A1	1.99085	1.87121
Neutral alpha-glucosidase AB	0.491559	0.569994
Galectin-3-binding protein	1.51829	n
14-3-3 protein theta	n	-0.485601
14-3-3 protein gamma	n	-0.689998
Prothrombin	n	1.49964
Thrombospondin-1	n	n
Transitional endoplasmic reticulum ATPase	n	n
Alpha-2-macroglobulin	n	n

The proteins are therefore suggested here as GBM EV specific or MEN EV specific markers which are also up/down regulated when compared to non-tumour tissues. They could possibly aid as protein biomarkers, obtainable from EVs during a cancer liquid biopsy and can serve as diagnostic, prognostic or predictive markers upon further specific studies and marker validation. This study has for the first time identified and quantified EV specific markers from primary tumours of GBM and MEN patient tissues, primary cells and EVs using the variable window approach of SWATH-DIA mass spectrometry. Deeper insights to these proteins were observed after a differential proteomic study between tumour and non-tumour tissues. This study has therefore contributed massively towards research in the idea of tumour diagnosis from patient EVs. Recently, these lipid-based membrane vesicles are being widely studied to exploit their potentials as they are released both by cancer and healthy cells[220], with approximately two-thirds of current EV clinical trials relating to diagnostics,

and the rest to therapeutics[221]. Unlike other micro particles, they are highly abundant in bio-fluids[222] and their levels in the plasma of cancer patients have been studied [223]. Therefore fluids like cerebrospinal fluids and plasma of patients with brain tumours need to be studied to validate these biomarkers identified in this study.

5. Materials and Methods

5.1 Chemicals, Instruments and Software

The tables below detail the list of chemicals, instruments and software which were used in all the experiments in this study.

Table 10: List of chemicals

Name	Supplier (location)
Urea	Merck (Darmstadt, Germany)
Trypsin	Promega (Mannheim, Germany)
Trypsin Resuspension buffer	Promega (Mannheim, Germany)
Ethanol (LiChrosolv®)	Merck (Darmstadt, Germany)
Acclaim PepMap	Thermo scientific (Germany)
Ammonium bicarbonate	Merck (Darmstadt, Deutschland)
Acetonitrile (LiChrosolv®)	Merck (Darmstadt, Germany)
Methanol (LiChrosolv®)	Merck (Darmstadt, Germany)
MS Water (LiChrosolv®)	Merck (Darmstadt, Germany)
Formic acid	Merck (Darmstadt, Germany)
Dithiotrietol	Sigma-Aldrich (Steinheim, Germany)
Iodoacetamide	Sigma-Aldrich (Taufheim, Germany)
Sodium hydrogen carbonate	Merck (Darmstadt, Germany)
BCA test kit	Thermo scientific (Germany)
Sodium Deoxycholate	Sigma-Aldrich (Steinheim, Germany)
Triethylammonium bicarbonate	Thermo scientific (Germany)
iRT kit	Biognosys AG (Schlieren, Switzerland)

Table 11: List of instruments

Name	Supplier (location)
Q Exactive Hybrid-Quadrupol-Orbitrap™ mass spectrometer	Thermo Scientific (Waltham, USA)
Waters nanoACQUITY UPLC system	Waters (Manchester, UK)
Speed Vac concentrator 5301	Eppendorf AG (Hamburg, Deutschland)
Thermo mixer 5320	Eppendorf AG (Hamburg, Deutschland)
Table Centrifuge	Sigma-Aldrich (Steinheim, Deutschland)
TissueLyser II	QIAGEN (Hilden, Germany)
Sonicator HD2200	Bendelin (Berlin, Germany)

Table 12: List of software

Name	Supplier (location)
Spectronaut pulsar Professional	Biognosys Inc, (Massachusetts, USA)
HTRMS convert	Biognosys Inc (Massachusetts, USA)
Microsoft Excel, word, power point 2007	Microsoft Corporation
Thermo Xcalibur™ 4.0.27.13	Thermo Scientific (Bremen, Germany)
Perseus 1585	Max-Planck-Institut (München, Germany)
MaxQuant 1.5.8.3	Max-Planck-Institut (München, Germany)
FreeStyle™ 1.4.60.34	Thermo Fisher Scientific Inc. Bremen, Germany)

5.2 Workflow Overview

A general overview of the samples and the workflow employed in the experiments in this study is as depicted in figure 45 below. All the samples used in this study were provided by Dr. Franz Lennard Ricklefs from the Department of Neurosurgery of the Head and Neuro centre at the University Medical centre Hamburg Eppendorf (UKE). They were all processed and analysed at the laboratory of Prof. Schlüter, Institute of clinical chemistry at the UKE. All these samples were of human origin. Figure 45 shows a summary of the samples used in this study and the key steps involved.

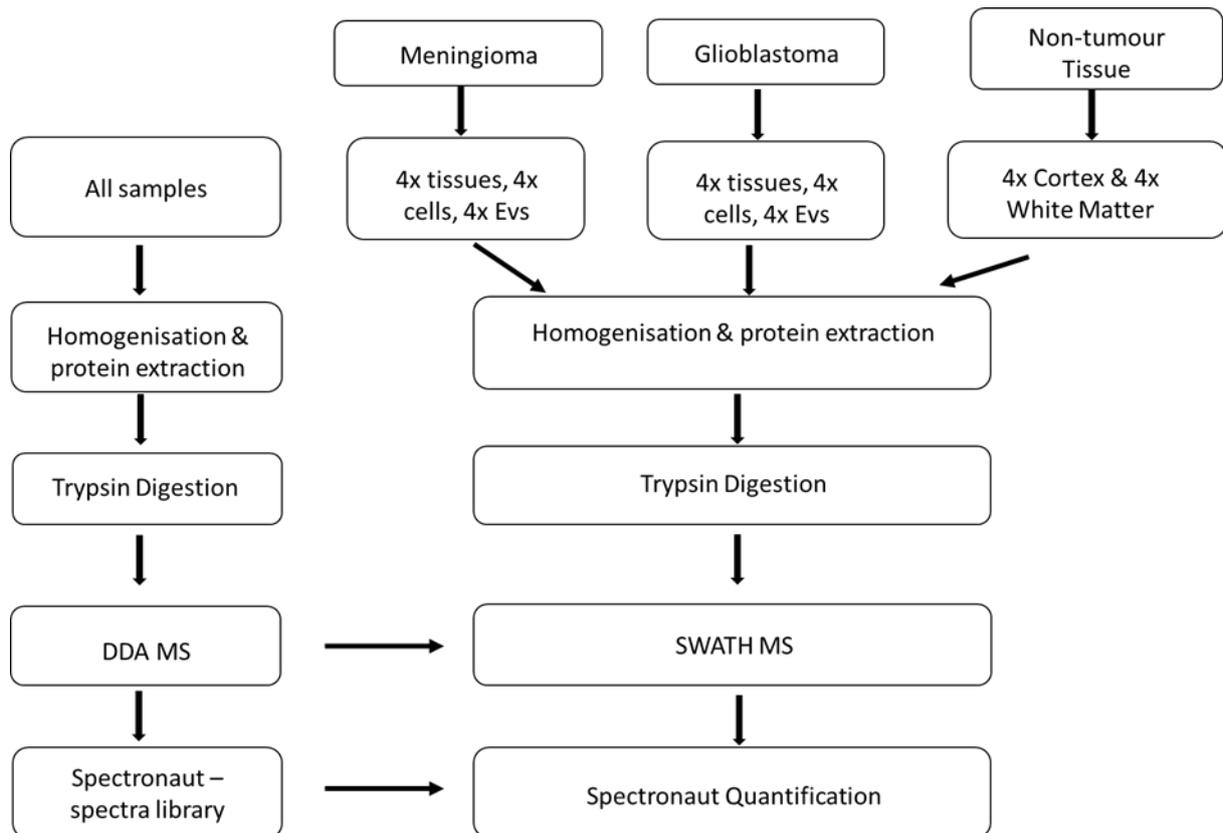


Fig 45: A schematic of the experimental workflow used in this study. These include the key steps that were undertaken in the generation of the spectral library and the protein identification and quantification from the different samples.

5.3 Brain Biopsy Samples

The brain tissue samples used in this study were of three main categories: meningioma (primary benign brain tumour), glioblastoma (primary malignant brain tumour) and non-tumour brain. The meningioma and glioblastoma samples were obtained after surgical excision from four meningioma and four glioblastoma patients. The non-tumour tissues were obtained from surgical treatment of epileptic patients. Four were obtained from the white matter of the temporal lobe and four from the cortex. All tissue samples were fresh frozen at -80 °C prior to homogenisation. Primary cell cultures

were done from each patient tissue sample and each patient EV was also harvested from cell culture supernatant.

5.4 Lysis and Protein Extraction

Tissue homogenisation was done with a Tissue lyser and the homogenization buffer used was Sodium Deoxycholate (SDC) buffer which is made up of 1% w/v SDC in 0.1M Triethylammonium bicarbonate (TEAB). One homogenisation stainless steel bead was put in each Eppendorf tube containing the tissue. 300 µL of tissue lysis buffer was added and the tissues homogenised with a bead mill (TissueLyser II, Qiagen). It was operated at 30 Hz frequency for 3:30 minutes per cycle. Due to the fact that brain tissue is relatively soft, only one cycle was needed. The whole tissue homogenate was then heated in a thermomixer at 99 °C for 5 minutes. It was then centrifuged at 14000g for 20 minutes and the supernatant collected in a new tube (Supernatant tube). To the pellet, 8M urea was added and vortexed to 30 seconds. It was then placed on a mixer at 800 rpm for 10 minutes. It was also centrifuged at 14000g for 20 minutes and the supernatant collected in the Supernatant tube. The supernatant was then stored for trypsin digestion.

To the cell and extracellular vesicle samples, 300 µL of hot SDC buffer was added to each sample. They were ultrasonicated with a probe for one sonication cycle which is 30 seconds at 25%. Immediately after sonication, the samples were heated on a thermo block at 99 °C for 5 minutes. The homogenates were then centrifuged at 14000g for 20 minutes and the supernatant collected for trypsin digestion. Protein concentration was determined using a bicinchoninic acid (BCA) test following the standard protocol provided by Thermo Fisher.

5.5 In-solution protein digestion to peptides

Protein digestion was carried out according to the filter-aided sample preparation (FASP) protocol for proteomics [224]. A buffer exchange was done for all lysates to 6M Urea. The disulphide bonds of cysteine residues were reduced by incubating them at 56 °C for 30 minutes with 20 mM dithiothreitol (DTT). To prevent re-linking of these disulphide bonds, the reduced cysteines were alkylated with 60mM Iodoacetamide (IAA) at 37 °C in the dark. Trypsin was then added to the solution at a ratio of 1:100 (Trypsin: Protein) after AmbiCa and incubated at 37 °C overnight. The resulting peptides were desalted with a reverse phase column (C18 cartridges from Waters) and lyophilised using a vacuum centrifuge concentrator.

5.6 Data Dependent Acquisition Mass Spectrometry (DDA MS) data acquisition

For liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis, the vacuumed peptides were resuspended in 0.1% formic acid (FA) and spiked with an internal retention time (iRT) calibration mix (Biognosys AG) and analysed on a Waters nanoACQUITY HPLC system coupled to a Thermo Q Exactive-orbitrap mass spectrometer via electrospray-ionization (ESI) with a nano-spray source. The peptides were loaded on a (2 cm x 75 μ m ID; Acclaim PepMap trap column packed with 3 μ m beads, Thermo Fisher Scientific) and separated on a reversed phase column (25 cm x 75 μ m ID, Acclaim PepMap, 3 μ m beads, Thermo Fisher Scientific) analytical column.

The flow rate during the whole chromatographic run was 250 nL/min. Buffer A was composed of 0.1% FA in HPLC-H₂O and buffer B; 0.1% FA in ACN. During column equilibration, sample application and the wash phase for unbound protein, the ratio of buffer A: B was 98%:2%. A gradient elution phase then followed with an increase in the concentration of buffer B to 30% in 90 minutes followed by an increase to 70% in a 5-minutes, 90% in 3 minutes, a 2-minute hold phase and back to 2% for 20 minutes for column equilibration. The column temperature was maintained at 45 °C. The eluted peptides were ionized via electrospray-ionization in positive mode. The MS data was acquired in a data dependent mode. Full scan spectra were acquired at a resolution of 70,000, an ACG target of 3×10^6 and a maximum injection time of 100ms. The scan range was set from 400 to 1300 m/z and the spectra data type was set to profile. The top 8 most intense precursor ions were selected for fragmentation with an intensity threshold of 5×10^4 . Single charged signals and those greater than 6 were excluded from fragmentation. The MS/MS scans were acquired at an m/z range was 200-2000, with a resolution of 17,500. The collision energy of was set to 30%, AGC target of 1×10^4 , an isolation width of 2 m/z. The maximum injection time was 200ms and the dynamic exclusion of fragmented ions was set to 25s. The same LC-MS/MS parameters on the same instrument were applied for all samples in this study.

5.7 Liquid chromatography and SWATH DIA mass spectrometry analysis

For all samples in this study, the same LC and full scan method was employed as described in section 7.6. However, the MS/MS scans were not acquired in a data dependent mode but rather data independent mode. The method used was the sequential window acquisition of all theoretical mass spectra (SWATH). Fragment ion spectra were recorded for all precursors within a user-defined precursor ion window. Smaller Q1 windows were therefore used in dense m/z areas where higher precursors were measured whereas wider Q1 windows were implemented in less dense m/z regions. The Q1 isolation window widths used for all samples were manually calculated after processing the

data obtained from the data dependent acquisition experiments. These data were analysed with the MaxQuant software and information on the retention time and m/z values of the peptides were obtained. The samples were grouped into the following groups and processed separately: tumour tissues, cells, extracellular vesicles and non-tumour tissues. For all samples, 20 precursor isolation windows with variable sizes were employed across the scan range of 400 to 1300 m/z. Precursor ions were fragmented with stepped collision energy of 22.5%, 25% and 27.5%. The MS/MS spectra were recorded at a resolution of 35000 and an AGC target of 2×10^5 for all samples. The difference between the samples came in at the level of the inclusion list. The inclusion list for every sample group is as listed in tables 13 below.

Table 13: Inclusion list for SWATH DIA spanning the m/z range of 400-1300 for various sample types for SWATH MS analytical method

Tissues	Cells	Extracellular vesicles	Non-tumour tissues
412.5	413.5	413.5	412.5
435.0	436.0	437.0	435.0
457.0	457.5	456.5	457.0
475.0	478.0	472.0	475.0
498.5	498.0	492.5	498.5
520.5	519.0	514.0	520.5
542.5	540.5	529.5	542.5
564.0	561.5	550.0	564.0
587.0	853.0	573.0	587.0
611.5	606.0	590.5	611.5
638.0	631.0	613.0	638.0
668.5	657.5	640.0	668.5
697.5	686.0	672.5	697.5
732.0	718.0	703.0	732.0
770.5	754.0	728.0	770.5
815.0	796.5	765.0	815.0
870.5	848.5	818.0	870.5
953.0	921.5	894.0	953.0
1072.0	1071.0	1039.5	1072.0
1216.5	1235.0	1215.0	1216.5

5.8 Data Analysis

5.8.1 Data dependent LC-MS/MS experiments

LC-MS/MS data recorded on the Q Exactive hybrid quadrupole-Orbitrap mass spectrometer was processed with MaxQuant (version 1.6.1.0)[225]. The basic standard settings were maintained with an additional option of Label free quantification (LFQ) selected when the samples were processed for analysis but was unselected when the raw data was being processed for DIA spectral library generation. The digestion enzyme was set to Trypsin/P. Variable modification settings by oxidation of methionine and acetylation (Acetyl (Protein N-term); Oxidation (M)) were included as well as a static modification setting of cysteine (Carbamidomethyl (C)). The data were searched against a Uniprot reviewed *Homo sapiens* proteome database with 20,399 entries including common contaminants for false-discovery rate filtering of peptide and protein identifications. Matches were filtered setting false peptide and protein (PSM FDR and protein FDR) hits to 1%. The minimum peptide length was set to be 7 amino acids; the minimum score for modified peptides was set to 40. For protein identification, one non-unique razor peptide was required, whereas protein quantitation was only performed if at least 2 razor peptides were associated with the protein hit. The match between run option was selected for advanced identification with a match time window of 0.7 mins, an alignment time window of 20 mins and the match unidentified features selected. The output files were text files with protein groups which were further statistically processed.

5.8.2 Spectral library generation

The spectral libraries were generated using MaxQuant and Spectronaut Pulsar. The prepare perspective in Spectronaut™ allows to generate, merge and manage spectral libraries which are needed for the analysis of DIA data in Spectronaut. The commercially available iRT peptides which were spiked into the samples prior to MS analysis were used to empirically determine accurate iRT values for all the peptides in the spectral library. The text files from MaxQuant were generated as specified in section 5.8.1. The spectronaut library generation settings were set as follows: for identification parameters, maxquant score type was PSM with an FDR of 1 % and the confidence level set to high. For protein interference, the digest type was specific, and the digest rule was Trypsin/P. The modifications included oxidation of methionine, Protein N-term acetylation and Carbamidomethylation of cysteine cysteine. The protein database used was same as that used to obtain the maxquant text files. A human gene ontology database was also used which was downloaded from the GO consortium webpage (www.geneontology.org) [226, 227]. The spectral libraries were then generated from MaxQuant output files, assigning the shotgun files. Sample

specific libraries were generated directly from shotgun files whereas one project specific library was generated using the merge library option in spectronaut.

5.8.3 Protein Quantification & statistical analysis

The generated DIA raw files were first converted from the *.raw* format to HTRMS format using the HTRMS converter provided by Biognosys Inc for faster processing. The settings were set to default and the mass spectrometer provider set to Thermo. Spectral alignment and targeted data extraction was done using Spectronaut™ software. The HTRMS files were then loaded into spectronaut in the review perspective for quantification. The previously generated spectral libraries were then assigned and the Uniprot reviewed *Homo sapiens* proteome database was selected. The settings were set similar to those described under section 5.8.2 for spectral library generation. The results were then checked under the QC and post analysis perspective. The protein list was then exported to excel for further statistical analysis. The statistical software used was Perseus version 1.5.1.6 [228] for T tests, generating heat maps and volcano plots after statistical testing and filtering for valid values.

5.8.4 Pathway and enrichment analysis

For Gene Ontology (GO) pathway analysis was done using the Panther DB. The network under section 3.1.7 was created using the proteins identified unique to GBM or MEN tissues, cells and EVs. Network enrichment analysis under sections 3.2.1 and 3.2.2 was also done using Panther DB. The proteins which unique to either primary cell culture or tissues, for GBM and MEN patients were also further analysed with Panther GO-Slim. String DB analysis was also done with KEGG pathway option incorporated in the search for the proteins under section 4.3.

6. References

1. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2018*. CA Cancer J Clin, 2018. **68**(1): p. 7-30.
2. Ben-David, U., R. Beroukhim, and T.R. Golub, *Genomic evolution of cancer models: perils and opportunities*. Nat Rev Cancer, 2018.
3. Hadders-Algra, M., *Early human brain development: Starring the subplate*. Neurosci Biobehav Rev, 2018. **92**: p. 276-290.
4. Andersen, B.B., L. Korbo, and B. Pakkenberg, *A quantitative study of the human cerebellum with unbiased stereological techniques*. J Comp Neurol, 1992. **326**(4): p. 549-60.
5. Doetsch, F., *The glial identity of neural stem cells*. Nat Neurosci, 2003. **6**(11): p. 1127-34.
6. Azevedo, F.A., et al., *Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain*. J Comp Neurol, 2009. **513**(5): p. 532-41.
7. Hofman, M.A., *Evolution of the human brain: when bigger is better*. Front Neuroanat, 2014. **8**: p. 15.
8. *Human Brain*.
9. Hari, R., et al., *Centrality of Social Interaction in Human Brain Function*. Neuron, 2015. **88**(1): p. 181-93.
10. team, T.A.C.S.m.a.e.c. *Key Statistics for Brain and Spinal Cord Tumors in Children*. 2018 June 20, 2018; Available from: <https://www.cancer.org/cancer/brain-spinal-cord-tumors-children/about/key-statistics.html>.
11. Netson, K.L., et al., *Executive dysfunction is associated with poorer health-related quality of life in pediatric brain tumor survivors*. J Neurooncol, 2016. **128**(2): p. 313-21.
12. Shergalis, A., et al., *Current Challenges and Opportunities in Treating Glioblastoma*. Pharmacol Rev, 2018. **70**(3): p. 412-445.
13. Ostrom, Q.T., et al., *CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2009-2013*. Neuro Oncol, 2016. **18**(suppl_5): p. v1-v75.
14. Bondy, M.L., et al., *Brain tumor epidemiology: consensus from the Brain Tumor Epidemiology Consortium*. Cancer, 2008. **113**(7 Suppl): p. 1953-68.
15. Ohgaki, H., *Epidemiology of brain tumors*. Methods Mol Biol, 2009. **472**: p. 323-42.
16. Blissitt, P.A., *Clinical practice guideline series update: care of the adult patient with a brain tumor*. J Neurosci Nurs, 2014. **46**(6): p. 367-8.
17. Adamson, C., et al., *Glioblastoma multiforme: a review of where we have been and where we are going*. Expert Opin Investig Drugs, 2009. **18**(8): p. 1061-83.
18. Ellor, S.V., T.A. Pagano-Young, and N.G. Avgeropoulos, *Glioblastoma: background, standard treatment paradigms, and supportive care considerations*. J Law Med Ethics, 2014. **42**(2): p. 171-82.
19. Johnson, D.R., et al., *Case-Based Review: newly diagnosed glioblastoma*. Neuro-Oncology Practice, 2015. **2**(3): p. 106-121.
20. Smith, J.S., et al., *PTEN mutation, EGFR amplification, and outcome in patients with anaplastic astrocytoma and glioblastoma multiforme*. J Natl Cancer Inst, 2001. **93**(16): p. 1246-56.
21. Huang, P.H., A.M. Xu, and F.M. White, *Oncogenic EGFR signaling networks in glioma*. Sci Signal, 2009. **2**(87): p. re6.
22. Wen, P.Y., et al., *Phase I/II study of erlotinib and temsirolimus for patients with recurrent malignant gliomas: North American Brain Tumor Consortium trial 04-02*. Neuro Oncol, 2014. **16**(4): p. 567-78.
23. Westphal, M., C.L. Maire, and K. Lamszus, *EGFR as a Target for Glioblastoma Treatment: An Unfulfilled Promise*. CNS Drugs, 2017. **31**(9): p. 723-735.
24. Taylor, T.E., F.B. Furnari, and W.K. Cavenee, *Targeting EGFR for treatment of glioblastoma: molecular basis to overcome resistance*. Curr Cancer Drug Targets, 2012. **12**(3): p. 197-209.

25. *Comprehensive genomic characterization defines human glioblastoma genes and core pathways.* Nature, 2008. **455**(7216): p. 1061-8.
26. Brennan, C.W., et al., *The somatic genomic landscape of glioblastoma.* Cell, 2013. **155**(2): p. 462-77.
27. Staedtke, V., O. Dzaye, and M. Holdhoff, *Actionable molecular biomarkers in primary brain tumors.* Trends Cancer, 2016. **2**(7): p. 338-349.
28. Reifenberger, G., et al., *Advances in the molecular genetics of gliomas - implications for classification and therapy.* Nat Rev Clin Oncol, 2017. **14**(7): p. 434-452.
29. Cushing, H. and L. Eisenhardt, *Meningiomas arising from the tuberculum sellae: With the syndrome of primary optic atrophy and bitemporal field defects combined with a normal sella turcica in a middle-aged person.* Archives of Ophthalmology, 1929. **1**(2): p. 168-206.
30. Shaikh, N., K. Dixit, and J. Raizer, *Recent advances in managing/understanding meningioma.* F1000Res, 2018. **7**.
31. Ostrom, Q.T., et al., *CBTRUS Statistical Report: Primary brain and other central nervous system tumors diagnosed in the United States in 2010-2014.* Neuro Oncol, 2017. **19**(suppl_5): p. v1-v88.
32. Louis, D.N., et al., *The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary.* Acta Neuropathol, 2016. **131**(6): p. 803-20.
33. Harter, P.N., Y. Braun, and K.H. Plate, *Classification of meningiomas-advances and controversies.* Chin Clin Oncol, 2017. **6**(Suppl 1): p. S2.
34. Varlotto, J., et al., *Distinguishing grade I meningioma from higher grade meningiomas without biopsy.* Oncotarget, 2015. **6**(35): p. 38421-8.
35. Chamberlain, M.C., *Treatment of Meningioma, Including in Cases With No Further Surgical or Radiotherapy Options.* Oncology, 2015. **29**(5).
36. Walcott, B.P., et al., *Radiation Treatment for WHO Grade II and III Meningiomas.* Front Oncol, 2013. **3**: p. 227.
37. Simpson, D., *The recurrence of intracranial meningiomas after surgical treatment.* J Neurol Neurosurg Psychiatry, 1957. **20**(1): p. 22-39.
38. Nanda, A., et al., *Relevance of Simpson grading system and recurrence-free survival after surgery for World Health Organization Grade I meningioma.* J Neurosurg, 2017. **126**(1): p. 201-211.
39. James, M.F., et al., *Modeling NF2 with human arachnoidal and meningioma cell culture systems: NF2 silencing reflects the benign character of tumor growth.* Neurobiol Dis, 2008. **29**(2): p. 278-92.
40. Riemenschneider, M.J., A. Perry, and G. Reifenberger, *Histological classification and molecular genetics of meningiomas.* Lancet Neurol, 2006. **5**(12): p. 1045-54.
41. Petricoin, E.F., et al., *Clinical proteomics: translating benchside promise into bedside reality.* Nat Rev Drug Discov, 2002. **1**(9): p. 683-95.
42. Barkhoudarian, G., et al., *Proteomics Analysis of Brain Meningiomas in Pursuit of Novel Biomarkers of the Aggressive Behavior.* J Proteomics Bioinform, 2016. **9**(2): p. 53-57.
43. *Biomarkers and surrogate endpoints: preferred definitions and conceptual framework.* Clin Pharmacol Ther, 2001. **69**(3): p. 89-95.
44. Programme on Chemical, S., *Biomarkers In Risk Assessment: Validity And Validation.* 2001.
45. Strimbu, K. and J.A. Tavel, *What are biomarkers?* Curr Opin HIV AIDS, 2010. **5**(6): p. 463-6.
46. Davis, J., et al., *Towards a classification of biomarkers of neuropsychiatric disease: from encompass to compass.* Molecular Psychiatry, 2014. **20**: p. 152.
47. Barabási, A.L., N. Gulbahce, and J. Loscalzo, *Network medicine: a network-based approach to human disease.* Nat Rev Genet, 2011. **12**(1): p. 56-68.
48. Sasmita, A.O., Y.P. Wong, and A.P.K. Ling, *Biomarkers and therapeutic advances in glioblastoma multiforme.* Asia Pac J Clin Oncol, 2018. **14**(1): p. 40-51.

49. Della Puppa, A., et al., *MGMT expression and promoter methylation status may depend on the site of surgical sample collection within glioblastoma: a possible pitfall in stratification of patients?* J Neurooncol, 2012. **106**(1): p. 33-41.
50. Szopa, W., et al., *Diagnostic and Therapeutic Biomarkers in Glioblastoma: Current Status and Future Perspectives*. BioMed Research International, 2017. **2017**: p. 13.
51. Yan, H., et al., *IDH1 and IDH2 Mutations in Gliomas*. New England Journal of Medicine, 2009. **360**(8): p. 765-773.
52. Vizcaíno, M.A., et al., *Clinicopathologic implications of NF1 gene alterations in diffuse gliomas*. Human Pathology, 2015. **46**(9): p. 1323-1330.
53. Goutagny, S., et al., *Long-term follow-up of 287 meningiomas in neurofibromatosis type 2 patients: clinical, radiological, and molecular features*. Neuro Oncol, 2012. **14**(8): p. 1090-6.
54. Wellenreuther, R., et al., *Analysis of the neurofibromatosis 2 gene reveals molecular variants of meningioma*. Am J Pathol, 1995. **146**: p. 827-832.
55. Paulovich, A.G., et al., *The interface between biomarker discovery and clinical validation: The tar pit of the protein biomarker pipeline*. Proteomics Clin Appl, 2008. **2**(10-11): p. 1386-1402.
56. Marshall, K.W., et al., *A blood-based biomarker panel for stratifying current risk for colorectal cancer*. Int J Cancer, 2010. **126**(5): p. 1177-86.
57. Barbazan, J., et al., *A multimarker panel for circulating tumor cells detection predicts patient outcome and therapy response in metastatic colorectal cancer*. Int J Cancer, 2014. **135**(11): p. 2633-43.
58. Simmons, A.R., et al., *Validation of a Biomarker Panel and Longitudinal Biomarker Performance for Early Detection of Ovarian Cancer*. Int J Gynecol Cancer, 2016. **26**(6): p. 1070-7.
59. Zerbino, D.D., *[Biopsy: its history, current and future outlook]*. Lik Sprava, 1994(3-4): p. 1-9.
60. C. G. Cowan, J.M., *Oral lesions: differential diagnosis and biopsy techniques*, in *Oral and Maxillofacial Surgery*, J.F. Jonathan Pedlar, Editor. 2007, Churchill Livingstone: United Kingdom. p. 100-114.
61. *Biopsy - Overview*. 01/06/2018 [cited 2019 24.01.2019]; Available from: <https://www.nhs.uk/conditions/biopsy/>.
62. Arneith, B., *Update on the types and usage of liquid biopsies in the clinical setting: a systematic review*. BMC Cancer, 2018. **18**(1): p. 527.
63. Apuzzo, M.L., et al., *Computed imaging stereotaxy: experience and perspective related to 500 procedures applied to brain masses*. Neurosurgery, 1987. **20**(6): p. 930-7.
64. Hamisch, C., et al., *Update on the diagnostic value and safety of stereotactic biopsy for pediatric brainstem tumors: a systematic review and meta-analysis of 735 cases*. J Neurosurg Pediatr, 2017. **20**(3): p. 261-268.
65. Jackson, R.J., et al., *Limitations of stereotactic biopsy in the initial management of gliomas*. Neuro Oncol, 2001. **3**(3): p. 193-200.
66. Bonner, E.R., et al., *Liquid biopsy for pediatric central nervous system tumors*. NPJ Precis Oncol, 2018. **2**: p. 29.
67. Veenstra, T.D., et al., *Biomarkers: Mining the Biofluid Proteome*. Molecular & Cellular Proteomics, 2005. **4**(4): p. 409.
68. Ashworth, T.R., *A Case of Cancer in Which Cells Similar to Those in the Tumors Were Seen in the Blood after Death*. Australasian Medical Journal, 1869. **14**: p. 146-149.
69. Krebs, M.G., et al., *Circulating tumour cells: their utility in cancer management and predicting outcomes*. Ther Adv Med Oncol, 2010. **2**(6): p. 351-65.
70. Ignatiadis, M., M. Lee, and S.S. Jeffrey, *Circulating Tumor Cells and Circulating Tumor DNA: Challenges and Opportunities on the Path to Clinical Utility*. Clin Cancer Res, 2015. **21**(21): p. 4786-800.
71. Chen, X.Q., et al., *Microsatellite alterations in plasma DNA of small cell lung cancer patients*. Nat Med, 1996. **2**(9): p. 1033-5.

72. Jahr, S., et al., *DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells*. *Cancer Res*, 2001. **61**(4): p. 1659-65.
73. D'Asti, E., et al., *Oncogenic extracellular vesicles in brain tumor progression*. *Front Physiol*, 2012. **3**: p. 294.
74. Fitzner, D., et al., *Selective transfer of exosomes from oligodendrocytes to microglia by macropinocytosis*. *J Cell Sci*, 2011. **124**(Pt 3): p. 447-58.
75. Rajendran, L., et al., *Emerging roles of extracellular vesicles in the nervous system*. *J Neurosci*, 2014. **34**(46): p. 15482-9.
76. D'Asti, E., et al., *Extracellular Vesicles in Brain Tumor Progression*. *Cell Mol Neurobiol*, 2016. **36**(3): p. 383-407.
77. Cohen, J.D., et al., *Detection and localization of surgically resectable cancers with a multi-analyte blood test*. 2018. **359**(6378): p. 926-930.
78. *Proteomics, transcriptomics: what's in a name?* *Nature*, 1999. **402**: p. 715.
79. Sanchez, J.-C., et al., *Progress with Proteome Projects: Why all Proteins Expressed by a Genome Should be Identified and How To Do It* AU - Wilkins, Marc R. *Biotechnology and Genetic Engineering Reviews*, 1996. **13**(1): p. 19-50.
80. Aebersold, R. and M. Mann, *Mass spectrometry-based proteomics*. *Nature*, 2003. **422**(6928): p. 198-207.
81. Gregorich, Z.R. and Y. Ge, *Top-down proteomics in health and disease: challenges and opportunities*. *Proteomics*, 2014. **14**(10): p. 1195-210.
82. Chothia, C., et al., *Evolution of the protein repertoire*. *Science*, 2003. **300**(5626): p. 1701-3.
83. Toyama, B.H. and M.W. Hetzer, *Protein homeostasis: live long, won't prosper*. *Nature Reviews Molecular Cell Biology*, 2012. **14**: p. 55.
84. Edman, P., *A method for the determination of amino acid sequence in peptides*. *Arch Biochem*, 1949. **22**(3): p. 475.
85. Patrie, S.M., *Top-Down Mass Spectrometry: Proteomics to Proteoforms*. *Adv Exp Med Biol*, 2016. **919**: p. 171-200.
86. Walther, T.C. and M. Mann, *Mass spectrometry-based proteomics in cell biology*. *The Journal of Cell Biology*, 2010. **190**(4): p. 491-500.
87. Fornelli, L., et al., *Middle-Down Analysis of Monoclonal Antibodies with Electron Transfer Dissociation Orbitrap Fourier Transform Mass Spectrometry*. *Analytical Chemistry*, 2014. **86**(6): p. 3005-3012.
88. Olsen, J.V., S.-E. Ong, and M. Mann, *Trypsin Cleaves Exclusively C-terminal to Arginine and Lysine Residues*. *Molecular & Cellular Proteomics*, 2004. **3**(6): p. 608-614.
89. Rodriguez, J., et al., *Does Trypsin Cut Before Proline?* *Journal of Proteome Research*, 2008. **7**(1): p. 300-305.
90. *LXXXIII. Rays of positive electricity* AU - Thomson, J.J. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 1910. **20**(118): p. 752-767.
91. *Rayleigh, Joseph John Thomson. 1856-1940*. *Obituary Notices of Fellows of the Royal Society*, 1941. **3**(10): p. 587-609.
92. Dole, M., et al., *Molecular Beams of Macroions*. *The Journal of Chemical Physics*, 1968. **49**(5): p. 2240-2249.
93. Yamashita, M. and J.B. Fenn, *Negative ion production with the electrospray ion source*. *The Journal of Physical Chemistry*, 1984. **88**(20): p. 4671-4675.
94. Ho, C.S., et al., *Electrospray ionisation mass spectrometry: principles and clinical applications*. *Clin Biochem Rev*, 2003. **24**(1): p. 3-12.
95. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules*. *Science*, 1989. **246**(4926): p. 64-71.

96. Banerjee, S. and S. Mazumdar, *Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte*. International Journal of Analytical Chemistry, 2012. **2012**: p. 40.
97. Bythell, B.J., C.L. Hendrickson, and A.G. Marshall, *Relative Stability of Peptide Sequence Ions Generated by Tandem Mass Spectrometry*. Journal of The American Society for Mass Spectrometry, 2012. **23**(4): p. 644-654.
98. Michalski, A., et al., *Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer*. Molecular & Cellular Proteomics, 2011: p. MCP.M111.011015.
99. Nicolas, A. *Proteomics. Follow The White Rabbit – Interpreting Individual Peaks – MS2 spectra*; Available from: <https://www.dcbiosciences.com/news/interpreting-peaks-ms2-spectra/>.
100. Brown, D.P.G., et al., *Label-free mass spectrometry-based protein quantification technologies in proteomic analysis*. Briefings in Functional Genomics, 2008. **7**(5): p. 329-339.
101. Hu, A., W.S. Noble, and A. Wolf-Yadlin, *Technical advances in proteomics: new developments in data-independent acquisition*. F1000Research, 2016. **5**: p. F1000 Faculty Rev-419.
102. Bern, M., et al., *Deconvolution of mixture spectra from ion-trap data-independent-acquisition tandem mass spectrometry*. Anal Chem, 2010. **82**(3): p. 833-41.
103. Chapman, J.D., D.R. Goodlett, and C.D. Masselon, *Multiplexed and data-independent tandem mass spectrometry for global proteome profiling*. Mass Spectrom Rev, 2014. **33**(6): p. 452-70.
104. Ludwig, C., et al., *Data-independent acquisition-based SWATH-MS for quantitative proteomics: a tutorial*. Mol Syst Biol, 2018. **14**(8): p. e8126.
105. Gillet, L.C., et al., *Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis*. Mol Cell Proteomics, 2012. **11**(6): p. O111.016717.
106. Rosenberger, G., et al., *A repository of assays to quantify 10,000 human proteins by SWATH-MS*. Sci Data, 2014. **1**: p. 140031.
107. Kalra, H., et al., *Vesiclepedia: a compendium for extracellular vesicles with continuous community annotation*. PLoS biology, 2012. **10**(12): p. e1001450-e1001450.
108. Rardin, M.J., et al., *MS1 Peptide Ion Intensity Chromatograms in MS2 (SWATH) Data Independent Acquisitions. Improving Post Acquisition Analysis of Proteomic Experiments*. Molecular & cellular proteomics : MCP, 2015. **14**(9): p. 2405-2419.
109. Cox, J., et al., *Accurate Proteome-wide Label-free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed MaxLFQ*. Molecular & Cellular Proteomics, 2014. **13**(9): p. 2513-2526.
110. INC, B., *Comparison of Static and Dynamic Segments in Data-Independent-Acquisition (DIA) Experiments*, B. INC, Editor.
111. Caron, E., et al., *An open-source computational and data resource to analyze digital maps of immunopeptidomes*. Elife, 2015. **4**.
112. Picotti, P., et al., *A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis*. Nature, 2013. **494**: p. 266.
113. Schubert, O.T., et al., *Absolute Proteome Composition and Dynamics during Dormancy and Resuscitation of Mycobacterium tuberculosis*. Cell Host Microbe, 2015. **18**(1): p. 96-108.
114. Muller, D.B., et al., *Systems-level Proteomics of Two Ubiquitous Leaf Commensals Reveals Complementary Adaptive Traits for Phyllosphere Colonization*. Mol Cell Proteomics, 2016. **15**(10): p. 3256-3269.
115. Escher, C., et al., *Using iRT, a normalized retention time for more targeted measurement of peptides*. Proteomics, 2012. **12**(8): p. 1111-21.

116. Collins, B.C., et al., *Multi-laboratory assessment of reproducibility, qualitative and quantitative performance of SWATH-mass spectrometry*. Nature Communications, 2017. **8**(1): p. 291.
117. Lane, R.E., et al., *Extracellular vesicles as circulating cancer biomarkers: opportunities and challenges*. Clinical and translational medicine, 2018. **7**(1): p. 14-14.
118. Palazzolo, G., et al., *Proteomic analysis of exosome-like vesicles derived from breast cancer cells*. Anticancer Res, 2012. **32**(3): p. 847-60.
119. Mallawaarachy, D.M., et al., *Comprehensive proteome profiling of glioblastoma-derived extracellular vesicles identifies markers for more aggressive disease*. Journal of neuro-oncology, 2017. **131**(2): p. 233-244.
120. Garcia, M., et al., *Biological and clinical significance of cathepsin D in breast cancer metastasis*. Stem Cells, 1996. **14**(6): p. 642-50.
121. Rochefort, H., *The prognostic value of cathepsin D in breast cancer. A long road to the clinic*. Eur J Cancer, 1996. **32a**(1): p. 7-8.
122. Abbott, D.E., et al., *Reevaluating cathepsin D as a biomarker for breast cancer: serum activity levels versus histopathology*. Cancer biology & therapy, 2010. **9**(1): p. 23-30.
123. Johnson, M.D., et al., *The Role of Cathepsin D in the Invasiveness of Human Breast Cancer Cells*. Cancer Research, 1993. **53**(4): p. 873-877.
124. Lim, L.H.K. and S. Pervaiz, *Annexin 1: the new face of an old molecule*. The FASEB Journal, 2007. **21**(4): p. 968-975.
125. Crumpton, M.J. and J.R. Dedman, *Protein terminology tangle*. Nature, 1990. **345**(6272): p. 212.
126. Guo, C., S. Liu, and M.-Z. Sun, *Potential role of Anxa1 in cancer*. Future Oncology, 2013. **9**(11): p. 1773-1793.
127. Sheikh, M.H. and E. Solito, *Annexin A1: Uncovering the Many Talents of an Old Protein*. International journal of molecular sciences, 2018. **19**(4): p. 1045.
128. Biaoxue, R., et al., *Upregulation of Hsp90-beta and annexin A1 correlates with poor survival and lymphatic metastasis in lung cancer patients*. J Exp Clin Cancer Res, 2012. **31**: p. 70.
129. Wang, W., et al., *Circulating Antibodies to Linear Peptide Antigens Derived from ANXA1 and FOXP3 in Lung Cancer*. Anticancer Res, 2017. **37**(6): p. 3151-3155.
130. Tu, Y., C.N. Johnstone, and A.G. Stewart, *Annexin A1 influences in breast cancer: Controversies on contributions to tumour, host and immunoediting processes*. Pharmacol Res, 2017. **119**: p. 278-288.
131. Pessolano, E., et al., *Annexin A1 May Induce Pancreatic Cancer Progression as a Key Player of Extracellular Vesicles Effects as Evidenced in the In Vitro MIA PaCa-2 Model System*. Int J Mol Sci, 2018. **19**(12).
132. Roth, U., et al., *Differential expression proteomics of human colorectal cancer based on a syngeneic cellular model for the progression of adenoma to carcinoma*. Proteomics, 2010. **10**(2): p. 194-202.
133. Patton, K.T., et al., *Decreased annexin I expression in prostatic adenocarcinoma and in high-grade prostatic intraepithelial neoplasia*. Histopathology, 2005. **47**(6): p. 597-601.
134. Tadei, M.B., et al., *Expression of the Annexin A1 and its correlation with matrix metalloproteinases and the receptor for formylated peptide-2 in diffuse astrocytic tumors*. Ann Diagn Pathol, 2018. **37**: p. 62-66.
135. Biaoxue, R., C. Xiguang, and Y. Shuanying, *Annexin A1 in malignant tumors: current opinions and controversies*. Int J Biol Markers, 2014. **29**(1): p. e8-20.
136. Lin, T.-W., et al., *Galectin-3 Binding Protein and Galectin-1 Interaction in Breast Cancer Cell Aggregation and Metastasis*. Journal of the American Chemical Society, 2015. **137**(30): p. 9685-9693.
137. Lin, T.W., et al., *Galectin-3 Binding Protein and Galectin-1 Interaction in Breast Cancer Cell Aggregation and Metastasis*. J Am Chem Soc, 2015. **137**(30): p. 9685-93.

138. Nigjeh, E.N., et al., *Spectral library-based glycopeptide analysis-detection of circulating galectin-3 binding protein in pancreatic cancer*. Proteomics. Clinical applications, 2017. **11**(9-10): p. 10.1002/prca.201700064.
139. Fukaya, Y., et al., *Identification of galectin-3-binding protein as a factor secreted by tumor cells that stimulates interleukin-6 expression in the bone marrow stroma*. J Biol Chem, 2008. **283**(27): p. 18573-81.
140. Austen, K., et al., *Extracellular rigidity sensing by talin isoform-specific mechanical linkages*. Nat Cell Biol, 2015. **17**(12): p. 1597-606.
141. Chen, P., et al., *Talin-1 interaction network promotes hepatocellular carcinoma progression*. Oncotarget, 2017. **8**(8): p. 13003-13014.
142. Lai, M.T., et al., *Talin-1 overexpression defines high risk for aggressive oral squamous cell carcinoma and promotes cancer metastasis*. J Pathol, 2011. **224**(3): p. 367-76.
143. Del Valle-Perez, B., et al., *Filamin B plays a key role in vascular endothelial growth factor-induced endothelial cell motility through its interaction with Rac-1 and Vav-2*. J Biol Chem, 2010. **285**(14): p. 10748-60.
144. Baudier, J., Z.A. Jenkins, and S.P. Robertson, *The filamin-B–refilin axis – spatiotemporal regulators of the actin-cytoskeleton in development and disease*. Journal of Cell Science, 2018. **131**(8): p. jcs213959.
145. Baldassarre, M., et al., *Filamins regulate cell spreading and initiation of cell migration*. PLoS One, 2009. **4**(11): p. e7830.
146. Narain, N.R., et al., *Identification of Filamin-A and -B as potential biomarkers for prostate cancer*. Future science OA, 2016. **3**(1): p. FSO161-FSO161.
147. Johnson, M., M. Sharma, and B.R. Henderson, *IQGAP1 regulation and roles in cancer*. Cell Signal, 2009. **21**(10): p. 1471-8.
148. Weissbach, L., et al., *Identification of a human rasGAP-related protein containing calmodulin-binding motifs*. J Biol Chem, 1994. **269**(32): p. 20517-21.
149. Johnson, M., et al., *IQGAP1 translocates to the nucleus in early S-phase and contributes to cell cycle progression after DNA replication arrest*. Int J Biochem Cell Biol, 2011. **43**(1): p. 65-73.
150. Sakurai-Yageta, M., et al., *The interaction of IQGAP1 with the exocyst complex is required for tumor cell invasion downstream of Cdc42 and RhoA*. The Journal of Cell Biology, 2008. **181**(6): p. 985-998.
151. Nabeshima, K., et al., *Immunohistochemical analysis of IQGAP1 expression in human colorectal carcinomas: its overexpression in carcinomas and association with invasion fronts*. Cancer Lett, 2002. **176**(1): p. 101-9.
152. Dong, P., et al., *Overexpression and diffuse expression pattern of IQGAP1 at invasion fronts are independent prognostic parameters in ovarian carcinomas*. Cancer Lett, 2006. **243**(1): p. 120-7.
153. Cui, X., et al., *Elevated IQGAP1 and CDC42 levels correlate with tumor malignancy of human glioma*. Oncology reports, 2017. **37**(2): p. 768-776.
154. McDonald, K.L., et al., *IQGAP1 and IGFBP2: valuable biomarkers for determining prognosis in glioma patients*. J Neuropathol Exp Neurol, 2007. **66**(5): p. 405-17.
155. Chiu, J., et al., *Protein Disulfide Isomerase in Thrombosis*. Semin Thromb Hemost, 2015. **41**(7): p. 765-73.
156. Xu, S., S. Sankar, and N. Neamati, *Protein disulfide isomerase: a promising target for cancer therapy*. Drug Discov Today, 2014. **19**(3): p. 222-40.
157. Benham, A.M., *The Protein Disulfide Isomerase Family: Key Players in Health and Disease*. Antioxidants & Redox Signaling, 2011. **16**(8): p. 781-789.
158. Lee, E. and D.H. Lee, *Emerging roles of protein disulfide isomerase in cancer*. BMB reports, 2017. **50**(8): p. 401-410.

159. Goldmann, W.H., et al., *Vinculin, cell mechanics and tumour cell invasion*. Cell Biol Int, 2013. **37**(5): p. 397-405.
160. Izard, T. and D.T. Brown, *Mechanisms and Functions of Vinculin Interactions with Phospholipids at Cell Adhesion Sites*. J Biol Chem, 2016. **291**(6): p. 2548-55.
161. Rubashkin, M.G., et al., *Force Engages Vinculin and Promotes Tumor Progression by Enhancing PI3K Activation of Phosphatidylinositol (3,4,5)-Triphosphate*. Cancer Research, 2014. **74**(17): p. 4597-4611.
162. Moore BW, P.V., *Specific acidic proteins of the nervous system*, in *Physiological and Biochemical Aspects of Nervous Integration*, C. FD, Editor. 1967: Englewood Cliffs, NJ: Prentice Hall. p. pp. 343–359.
163. Sun, S., et al., *14-3-3 and its binding partners are regulators of protein-protein interactions during spermatogenesis*. The Journal of endocrinology, 2009. **202**(3): p. 327-336.
164. Fu, H., R.R. Subramanian, and S.C. Masters, *14-3-3 proteins: structure, function, and regulation*. Annu Rev Pharmacol Toxicol, 2000. **40**: p. 617-47.
165. Keshavan, M.S., H.A. Nasrallah, and R. Tandon, *Schizophrenia, "Just the Facts" 6. Moving ahead with the schizophrenia concept: from the elephant to the mouse*. Schizophr Res, 2011. **127**(1-3): p. 3-13.
166. Chandra, S., et al., *Double-knockout mice for alpha- and beta-synucleins: effect on synaptic functions*. Proc Natl Acad Sci U S A, 2004. **101**(41): p. 14966-71.
167. Goedert, M., *Tau protein and the neurofibrillary pathology of Alzheimer's disease*. Trends Neurosci, 1993. **16**(11): p. 460-5.
168. Pereira-Faca, S.R., et al., *Identification of 14-3-3 theta as an antigen that induces a humoral response in lung cancer*. Cancer Res, 2007. **67**(24): p. 12000-6.
169. Qiu, J., et al., *Occurrence of autoantibodies to annexin I, 14-3-3 theta and LAMR1 in prediagnostic lung cancer sera*. J Clin Oncol, 2008. **26**(31): p. 5060-6.
170. Hodgkinson, V.C., et al., *Proteomic identification of predictive biomarkers of resistance to neoadjuvant chemotherapy in luminal breast cancer: a possible role for 14-3-3 theta/tau and tBID?* J Proteomics, 2012. **75**(4): p. 1276-83.
171. Himmelfarb, M., et al., *ITIH5, a novel member of the inter-alpha-trypsin inhibitor heavy chain family is downregulated in breast cancer*. Cancer Lett, 2004. **204**(1): p. 69-77.
172. Paris, S., et al., *Inhibition of tumor growth and metastatic spreading by overexpression of inter-alpha-trypsin inhibitor family chains*. International Journal of Cancer, 2002. **97**(5): p. 615-620.
173. Kobayashi, H., et al., *Inhibitory effect of a conjugate between human urokinase and urinary trypsin inhibitor on tumor cell invasion in vitro*. J Biol Chem, 1995. **270**(14): p. 8361-6.
174. Werbowetski-Ogilvie, T.E., et al., *Isolation of a natural inhibitor of human malignant glial cell invasion: inter alpha-trypsin inhibitor heavy chain 2*. Cancer Res, 2006. **66**(3): p. 1464-72.
175. Chen, H., M.E. Herndon, and J. Lawler, *The cell biology of thrombospondin-1*. Matrix Biol, 2000. **19**(7): p. 597-614.
176. Roberts, D.D., *Regulation of tumor growth and metastasis by thrombospondin-1*. Faseb j, 1996. **10**(10): p. 1183-91.
177. Esemuede, N., et al., *The role of thrombospondin-1 in human disease*. J Surg Res, 2004. **122**(1): p. 135-42.
178. Daubon, T., et al., *Deciphering the complex role of thrombospondin-1 in glioblastoma development*. Nat Commun, 2019. **10**(1): p. 1146.
179. Ogura, T. and A.J. Wilkinson, *AAA+ superfamily ATPases: common structure–diverse function*. Genes to Cells, 2001. **6**(7): p. 575-597.
180. Yamanaka, K., Y. Sasagawa, and T. Ogura, *Recent advances in p97/VCP/Cdc48 cellular functions*. Biochim Biophys Acta, 2012. **1823**(1): p. 130-7.
181. Butenas, S., C. van't Veer, and K.G. Mann, *"Normal" thrombin generation*. Blood, 1999. **94**(7): p. 2169-78.

182. Pozzi, N. and E. Di Cera, *Prothrombin structure: unanticipated features and opportunities*. Expert review of proteomics, 2014. **11**(6): p. 653-655.
183. Bertino, G., et al., *Prognostic and diagnostic value of des-gamma-carboxy prothrombin in liver cancer*. Drug News Perspect, 2010. **23**(8): p. 498-508.
184. Weitz, I.C. and H.A. Liebman, *Des-gamma-carboxy (abnormal) prothrombin and hepatocellular carcinoma: a critical review*. Hepatology, 1993. **18**(4): p. 990-7.
185. Suradej, B., et al., *Glucosidase II exhibits similarity to the p53 tumor suppressor in regards to structure and behavior in response to stress signals: a potential novel cancer biomarker*. Oncol Rep, 2013. **30**(5): p. 2511-9.
186. Pelletier, M.F., et al., *The heterodimeric structure of glucosidase II is required for its activity, solubility, and localization in vivo*. Glycobiology, 2000. **10**(8): p. 815-27.
187. Strous, G.J., et al., *Glucosidase II, a protein of the endoplasmic reticulum with high mannose oligosaccharide chains and a rapid turnover*. J Biol Chem, 1987. **262**(8): p. 3620-5.
188. Borth, W., *Alpha 2-macroglobulin, a multifunctional binding protein with targeting characteristics*. Faseb j, 1992. **6**(15): p. 3345-53.
189. Sottrup-Jensen, L., et al., *Primary structure of human alpha 2-macroglobulin. V. The complete structure*. J Biol Chem, 1984. **259**(13): p. 8318-27.
190. Gaziev, U.M., et al., *Alpha-2-macroglobulin as the marker for pancreatic cancer process direction*. Journal of Clinical Oncology, 2016. **34**(15_suppl): p. e15686-e15686.
191. Lindner, I., et al., *Alpha2-macroglobulin inhibits the malignant properties of astrocytoma cells by impeding beta-catenin signaling*. Cancer Res, 2010. **70**(1): p. 277-87.
192. Szklarczyk, D., et al., *The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible*. Nucleic acids research, 2017. **45**(D1): p. D362-D368.
193. Fresno Vara, J.A., et al., *PI3K/Akt signalling pathway and cancer*. Cancer Treat Rev, 2004. **30**(2): p. 193-204.
194. Moscatello, D.K., et al., *Decorin suppresses tumor cell growth by activating the epidermal growth factor receptor*. J Clin Invest, 1998. **101**(2): p. 406-12.
195. Vij, N., et al., *Lumican suppresses cell proliferation and aids Fas-Fas ligand mediated apoptosis: implications in the cornea*. Exp Eye Res, 2004. **78**(5): p. 957-71.
196. Niu, N. and L. Wang, *In vitro human cell line models to predict clinical response to anticancer drugs*. Pharmacogenomics, 2015. **16**(3): p. 273-285.
197. Zeeberg, B.R., et al., *Concordance of gene expression and functional correlation patterns across the NCI-60 cell lines and the Cancer Genome Atlas glioblastoma samples*. PloS one, 2012. **7**(7): p. e40062-e40062.
198. Ertel, A., et al., *Pathway-specific differences between tumor cell lines and normal and tumor tissue cells*. Molecular cancer, 2006. **5**(1): p. 55-55.
199. Iozzo, R.V. and R.D. Sanderson, *Proteoglycans in cancer biology, tumour microenvironment and angiogenesis*. J Cell Mol Med, 2011. **15**(5): p. 1013-31.
200. Ben-David, U., et al., *Genetic and transcriptional evolution alters cancer cell line drug response*. Nature, 2018. **560**(7718): p. 325-330.
201. Pan, C., et al., *Comparative proteomic phenotyping of cell lines and primary cells to assess preservation of cell type-specific functions*. Molecular & cellular proteomics : MCP, 2009. **8**(3): p. 443-450.
202. Axelsen, J.B., et al., *Genes overexpressed in different human solid cancers exhibit different tissue-specific expression profiles*. Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(32): p. 13122-13127.
203. Kobayashi, M., et al., *Calnexin is a novel sero-diagnostic marker for lung cancer*. Lung Cancer, 2015. **90**(2): p. 342-5.
204. Christensen, M.V., et al., *Annexin A2 and cancer: A systematic review*. Int J Oncol, 2018. **52**(1): p. 5-18.

205. Abbritti, R.V., et al., *Meningiomas and Proteomics: Focus on New Potential Biomarkers and Molecular Pathways*. *Cancer genomics & proteomics*, 2016. **13**(5): p. 369-379.
206. Chang, Y.K., et al., *Haptoglobin is a serological biomarker for adenocarcinoma lung cancer by using the ProteomeLab PF2D combined with mass spectrometry*. *Am J Cancer Res*, 2016. **6**(8): p. 1828-36.
207. Zhang, L., et al., *The role of calyculin gene in the proliferation and migration of gastric cancer cells*. *Oncol Rep*, 2011.
208. Samanta, S., et al., *Expression of protein disulfide isomerase family members correlates with tumor progression and patient survival in ovarian cancer*. *Oncotarget*, 2017. **8**(61): p. 103543-103556.
209. Deng, R., et al., *Gelsolin regulates proliferation, apoptosis, migration and invasion in human oral carcinoma cells*. *Oncology letters*, 2015. **9**(5): p. 2129-2134.
210. Steinlein, O.K., *Genetic mechanisms that underlie epilepsy*. *Nature Reviews Neuroscience*, 2004. **5**: p. 400.
211. Fukata, Y. and M. Fukata, *Epilepsy and synaptic proteins*. *Current Opinion in Neurobiology*, 2017. **45**: p. 1-8.
212. Lakhan, R., et al., *Association of intronic polymorphism rs3773364 A>G in synapsin-2 gene with idiopathic epilepsy*. *Synapse*, 2010. **64**(5): p. 403-8.
213. Peres, J., et al., *Sleep-related hypermotor epilepsy and peri-ictal hypotension in a patient with syntaxin-1B mutation*. *Epileptic Disord*, 2018. **20**(5): p. 413-417.
214. Xiao, Z., et al., *Altered expression of synaptotagmin I in temporal lobe tissue of patients with refractory epilepsy*. *J Mol Neurosci*, 2009. **38**(2): p. 193-200.
215. Xiao, Z., et al., *The effect of IL-1beta on synaptophysin expression and electrophysiology of hippocampal neurons through the PI3K/Akt/mTOR signaling pathway in a rat model of mesial temporal lobe epilepsy*. *Neurol Res*, 2017. **39**(7): p. 640-648.
216. Pitkanen, A., et al., *Neural ECM and epilepsy*. *Prog Brain Res*, 2014. **214**: p. 229-62.
217. Gorlewicz, A. and L. Kaczmarek, *Pathophysiology of Trans-Synaptic Adhesion Molecules: Implications for Epilepsy*. *Frontiers in cell and developmental biology*, 2018. **6**: p. 119-119.
218. Wu, L., C. Zhang, and F. Yin, *Dynamic changes in the proteomic profile of mesial temporal lobe epilepsy at different disease stages in an immature rat model*. *Protein Pept Lett*, 2015. **22**(2): p. 180-92.
219. Rohena, L., et al., *Mutation in SNAP25 as a novel genetic cause of epilepsy and intellectual disability*. *Rare diseases (Austin, Tex.)*, 2013. **1**: p. e26314-e26314.
220. Merchant, M.L., et al., *Isolation and characterization of urinary extracellular vesicles: implications for biomarker discovery*. *Nature reviews. Nephrology*, 2017. **13**(12): p. 731-749.
221. Roy, S., F.H. Hochberg, and P.S. Jones, *Extracellular vesicles: the growth as diagnostics and therapeutics; a survey*. *Journal of extracellular vesicles*, 2018. **7**(1): p. 1438720-1438720.
222. SADOVSKA, L., J. EGLĪTIS, and A. LINĒ, *Extracellular Vesicles as Biomarkers and Therapeutic Targets in Breast Cancer*. *Anticancer Research*, 2015. **35**(12): p. 6379-6390.
223. Logozzi, M., et al., *High levels of exosomes expressing CD63 and caveolin-1 in plasma of melanoma patients*. *PLoS One*, 2009. **4**(4): p. e5219.
224. Wisniewski, J.R., et al., *Universal sample preparation method for proteome analysis*. *Nat Methods*, 2009. **6**(5): p. 359-62.
225. Cox, J. and M. Mann, *MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification*. *Nat Biotechnol*, 2008. **26**(12): p. 1367-72.
226. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium*. *Nat Genet*, 2000. **25**(1): p. 25-9.
227. *The Gene Ontology Resource: 20 years and still GOing strong*. *Nucleic Acids Res*, 2019. **47**(D1): p. D330-d338.

228. Tyanova, S., et al., *The Perseus computational platform for comprehensive analysis of (prote)omics data*. Nat Methods, 2016. **13**(9): p. 731-40.

7. Supplement

Table S1: Proteins identified exclusively in all MEN samples presented in descending order of abundance in reference to extracellular vesicles only (Section 3.1.7 Fig. 26)

T: PG.ProteinDescriptions	mean
SPARC-like protein 1	6.6622525
Osteopontin	6.25541
Tenascin	6.1627325
Annexin A4	6.1530625
Testis-specific Y-encoded-like protein 2	6.0493025
HLA class I histocompatibility antigen, A-2 alpha chain	6.00231
Hexokinase-1	5.99593333
Myristoylated alanine-rich C-kinase substrate	5.92823
Urotensin-2	5.84240667
Neuromodulin	5.839985
Heterogeneous nuclear ribonucleoproteins C1/C2	5.80864
Katanin p60 ATPase-containing subunit A-like 2	5.791255
Heterogeneous nuclear ribonucleoprotein U	5.75187667
Protein disulfide-isomerase A4	5.742105
MARCKS-related protein	5.695905
Aldehyde dehydrogenase, mitochondrial	5.6883525
Glutamate dehydrogenase 1, mitochondrial	5.6695425
Microtubule-associated protein 1B	5.65245
Glycogen phosphorylase, muscle form	5.6337675
Histone H2B type 1-D;Histone H2B type 1-C/E/F/G/I;Histone H2B type 1-N;Histone H2B type 1-M;Histone H2B type 1-L	5.59872333
Pleiotrophin	5.5937375
Fatty acid-binding protein, brain	5.5437925
Neurocan core protein	5.52673333
Heat shock protein beta-1	5.4911125
Nucleoside diphosphate kinase B	5.47489333
L-lactate dehydrogenase B chain	5.46652
Extracellular superoxide dismutase [Cu-Zn]	5.44623
116 kDa U5 small nuclear ribonucleoprotein component	5.4179375
Ras-related protein Rab-7a	5.39177667
Phosphoglycerate mutase 1	5.38665667
Vigilin	5.38035667
Immunoglobulin heavy constant gamma 2	5.35122
60S ribosomal protein L14	5.33856
Importin-7	5.32172667
40S ribosomal protein S8	5.26758333
Stress-70 protein, mitochondrial	5.25379333
Protein SET	5.241605
Secretogranin-2	5.23478667
2,4-dienoyl-CoA reductase, mitochondrial	5.1983525
Multifunctional protein ADE2	5.192785
40S ribosomal protein S18	5.172835
Transmembrane glycoprotein NMB	5.165045

Elongation factor 1-gamma	5.1643925
3-ketoacyl-CoA thiolase, peroxisomal	5.15131333
Catenin beta-1	5.14728333
Cell division control protein 42 homolog	5.13222
Neuronal membrane glycoprotein M6-a	5.10526
40S ribosomal protein S25	5.08584
40S ribosomal protein S4, X isoform	5.07429
Cytosolic 10-formyltetrahydrofolate dehydrogenase	5.07118667
Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	5.07108667
3-hydroxyacyl-CoA dehydrogenase type-2	5.0705225
Guanine nucleotide-binding protein-like 1	5.03855333
Fumarate hydratase, mitochondrial	5.02267667
60S ribosomal protein L31	5.01017667
Nestin	5.007565
Eukaryotic translation initiation factor 3 subunit C	4.9985125
Cytoplasmic dynein 1 heavy chain 1	4.99346667
Probable ATP-dependent RNA helicase DDX5	4.99236667
60S ribosomal protein L13	4.98717667
Nucleosome assembly protein 1-like 4	4.95824
Glucosidase 2 subunit beta	4.95737667
40S ribosomal protein S14	4.93039333
Enoyl-CoA hydratase, mitochondrial	4.90899667
Serine/arginine-rich splicing factor 1	4.88249
Tropomyosin alpha-4 chain	4.82776
60S ribosomal protein L7	4.820365
UDP-glucose 6-dehydrogenase	4.8119375
40S ribosomal protein S9	4.80735333
Mitogen-activated protein kinase 1	4.79971333
Glycogen phosphorylase, brain form	4.79619
Bifunctional glutamate/proline--tRNA ligase	4.79343
D-3-phosphoglycerate dehydrogenase	4.77746667
Neural cell adhesion molecule 1	4.76447
Astrocytic phosphoprotein PEA-15	4.76085333
CD276 antigen	4.75560667
Transaldolase	4.7469275
Chromobox protein homolog 5	4.73242333
Poly(rC)-binding protein 1	4.71530667
Low molecular weight phosphotyrosine protein phosphatase	4.70467667
Dynein light chain 2, cytoplasmic	4.6882525
Cystatin-B	4.66680333
60S ribosomal protein L36a-like	4.6397525
Thioredoxin domain-containing protein 5	4.63313
Proteasome subunit beta type-7	4.61043
40S ribosomal protein S19	4.608895
Golgi apparatus protein 1	4.60232
Protein phosphatase 1G	4.59029667
Dehydrogenase/reductase SDR family member 4	4.56045
Importin subunit beta-1	4.51153333
Putative RNA-binding protein Luc7-like 2	4.47650667

T-complex protein 1 subunit eta	4.47087
Dihydropyrimidine dehydrogenase [NADP(+)]	4.44806333
Enhancer of rudimentary homolog	4.43102667
Very long-chain specific acyl-CoA dehydrogenase, mitochondrial	4.2862425
CAP-Gly domain-containing linker protein 2	4.14353333
N-acetylglucosamine-6-sulfatase	4.11588667
Copine-3	4.04967
Histone H1.0	3.98982667

Table S2: Proteins identified exclusively in all GBM samples presented in descending order of abundance in reference to extracellular vesicles only (Section 3.1.7 Fig. 26)

T: PG.ProteinDescriptions	mean
Titin	8.14471
Collagen alpha-1(III) chain	7.1768275
Prothrombin	7.05296667
Alpha-2-macroglobulin	6.94776333
Procollagen C-endopeptidase enhancer 1	6.92225333
Metalloproteinase inhibitor 1	6.80374
Signal transducing adapter molecule 1	6.69406333
Collagen alpha-2(I) chain	6.5653075
Propionyl-CoA carboxylase alpha chain, mitochondrial	6.55421667
Insulin-like growth factor-binding protein 5	6.3878875
Lumican	6.37849333
Thrombospondin-1	6.2765975
Hemopexin	6.15246667
Spectrin beta chain, non-erythrocytic 1	6.12844
Transforming growth factor-beta-induced protein ig-h3	5.903085
CD59 glycoprotein	5.755875
Tubulin beta chain	5.73573
Follistatin-related protein 1	5.69719
Galectin-3-binding protein	5.694625
Inter-alpha-trypsin inhibitor heavy chain H2	5.6390475
Elongation factor 1-delta	5.63509
5'-nucleotidase	5.59477
Basement membrane-specific heparan sulfate proteoglycan core protein	5.59454
Histidine triad nucleotide-binding protein 1	5.53721667
Sodium/potassium-transporting ATPase subunit alpha-1	5.526
Tetranectin	5.46089
Annexin A1	5.4481075
CD9 antigen	5.44366
Prostaglandin E synthase 3	5.4094
14-3-3 protein theta	5.34838333
Ras-related C3 botulinum toxin substrate 1	5.298205
Transgelin	5.281255
ADP/ATP translocase 3	5.25000333
EGF-containing fibulin-like extracellular matrix protein 1	5.24434667

Fibrillin-1	5.240635
Dynein heavy chain 5, axonemal	5.231
Na(+)/H(+) exchange regulatory cofactor NHE-RF1	5.20537
Aminopeptidase N	5.1547525
Ectonucleotide pyrophosphatase/phosphodiesterase family member 6	5.1409
Neutrophil defensin 1;Neutrophil defensin 3	5.13778667
Neuroblast differentiation-associated protein AHNAK	5.1215325
Heterogeneous nuclear ribonucleoprotein Q	5.1156325
Myosin regulatory light chain 12B	5.099735
Talin-1	5.076185
Tenascin-X	5.07504667
Cystatin-C	5.05472667
Vinculin	5.00102
Collagen alpha-1(XI) chain	4.98606
Chloride intracellular channel protein 1	4.98123
Nucleobindin-1	4.9659025
Protein CutA	4.9640575
Filamin-B	4.96203
Peroxiredoxin-6	4.93914
Heterogeneous nuclear ribonucleoprotein K	4.93164
Plastin-2	4.90196333
Transitional endoplasmic reticulum ATPase	4.8738325
Hsp90 co-chaperone Cdc37	4.86062333
Neutral alpha-glucosidase AB	4.85357667
F-actin-capping protein subunit beta	4.84922667
Fructose-bisphosphate aldolase C	4.84818333
14-3-3 protein gamma	4.8280975
Protein S100-A9	4.81604667
40S ribosomal protein S28	4.7972675
Prosaposin	4.7830175
40S ribosomal protein SA	4.77367
Plasma membrane calcium-transporting ATPase 4	4.75581
Antigen peptide transporter 2	4.75564667
Superoxide dismutase [Cu-Zn]	4.75235
Coronin-1A	4.74276667
Protein disulfide-isomerase	4.7368375
Filamin-C	4.73332333
T-complex protein 1 subunit theta	4.70865333
Heterogeneous nuclear ribonucleoprotein H	4.68724
Plastin-3	4.66558
Septin-7	4.64692333
Ras GTPase-activating-like protein IQGAP1	4.64201
Collagen alpha-3(VI) chain	4.60282
Coronin-1B	4.59468333
Tropomyosin beta chain	4.5431125
Rho GDP-dissociation inhibitor 1	4.5116875
14-3-3 protein sigma	4.49951667
Protein-glutamine gamma-glutamyltransferase 2	4.49799
Mitogen-activated protein kinase kinase kinase kinase 4	4.48377333

Translationally-controlled tumor protein	4.4452575
Glucose-6-phosphate 1-dehydrogenase	4.43587
C-type mannose receptor 2	4.41474
Mitogen-activated protein kinase 15	4.39463
Prohibitin	4.11731
Cathepsin L1	4.10089333
Delta(3,5)-Delta(2,4)-dienoyl-CoA isomerase, mitochondrial	3.83029333
Stathmin	3.70926

Table S3: Proteins identified in both all GBM and all MEN samples (Section 3.1.7 Fig. 26)

T: PG.ProteinDescriptions gbm	mean GBM	mean MEN
10 kDa heat shock protein, mitochondrial	5.3791825	5.49958
14-3-3 protein beta/alpha	5.761675	5.941615
14-3-3 protein epsilon	5.1908775	5.48923
14-3-3 protein zeta/delta	5.5076925	5.50997
40S ribosomal protein S12	4.528655	4.69433667
4F2 cell-surface antigen heavy chain	5.32393	5.21354667
4-trimethylaminobutyraldehyde dehydrogenase	4.47686667	4.927465
60 kDa heat shock protein, mitochondrial	5.4435725	4.906315
60S acidic ribosomal protein P0	5.31140333	5.16563
60S acidic ribosomal protein P1	5.230785	5.0658425
60S ribosomal protein L7a	4.74764667	4.79353667
60S ribosomal protein L8	4.56033	4.9044675
78 kDa glucose-regulated protein	5.8929375	5.8739175
Actin, alpha cardiac muscle 1	5.126595	4.911695
Actin, cytoplasmic 1	6.6808525	6.9409025
Actin-related protein 2/3 complex subunit 3	4.54325667	4.7639675
Adenosylhomocysteinase	5.36463	5.8964175
Alpha-enolase	6.054375	6.16565
Amyloid beta A4 protein	5.97103667	6.9015575
Annexin A11	5.0822425	5.16906667
Annexin A2	6.38214	6.4182925
Annexin A5	6.03891	6.645495
Annexin A6	5.179135	6.12958
Apolipoprotein C-III	5.96730333	6.1648025
Apolipoprotein E	6.1352925	5.84034
Arfaptin-1	5.11920667	5.5305475
ATP synthase subunit alpha, mitochondrial	4.34292333	5.14192333
ATP synthase subunit beta, mitochondrial	5.57689	5.43025
ATP-dependent RNA helicase DDX3Y	4.46458333	5.02436
Basigin	4.8016425	5.624625
Brain acid soluble protein 1	5.4312225	5.5314625
Caldesmon	5.8663375	5.60385
Calmodulin	5.7077275	6.1853075
Calnexin	4.52233667	5.1470175
Calpain-2 catalytic subunit	4.9903275	4.754735
Calreticulin	5.84045667	5.6634225

Catalase	5.33648	6.4327975
Cathepsin B	5.2233225	5.32299667
Cathepsin D	5.02241	5.2680375
CD44 antigen	5.5643425	6.16364667
Clathrin heavy chain 1	4.956875	5.20073
Clusterin	6.3251525	6.6275075
Coactosin-like protein	5.16970333	4.87514333
Cofilin-1	6.4295825	6.4168775
Collagen alpha-1(I) chain	7.178055	5.45675
Collagen alpha-1(VI) chain	6.264535	5.36781
Complement C3	5.5378925	4.6610475
Core histone macro-H2A.1	5.5192325	5.59436333
Creatine kinase B-type	5.7884125	5.5459975
Cullin-associated NEDD8-dissociated protein 1	5.09984	5.219825
Cytoskeleton-associated protein 2	6.8798175	7.15251
Cytoskeleton-associated protein 4	4.77076	4.85463
Desmoplakin	5.6797775	5.595695
Dihydropyrimidinase-related protein 2	4.8846375	5.62401
Dihydropyrimidinase-related protein 3	5.1118925	5.21998
Elongation factor 1-alpha 1	6.5179075	6.6779375
Elongation factor 1-beta	5.501345	5.1779275
Elongation factor 2	4.933545	5.52901667
Endoplasmic reticulum chaperone	6.1536975	5.9871925
Eukaryotic translation initiation factor 5A-1	5.329555	5.245915
Fascin	5.34561	5.48817667
Fatty acid-binding protein, epidermal	5.15051667	4.9061825
Fibronectin	8.0928075	6.343975
Filamin-A	5.6119925	5.692975
Fructose-bisphosphate aldolase A	5.90969	5.6800625
Fumarylacetoacetase	4.36244	4.71104
Galectin-1	5.8948475	5.866975
Gamma-enolase	5.4361	5.0747475
Gelsolin	6.427055	5.1463925
Glial fibrillary acidic protein	5.581775	5.76571
Glutathione S-transferase P	5.64429	5.3630075
Glyceraldehyde-3-phosphate dehydrogenase	6.7600875	6.812795
Golgin subfamily A member 3	5.47841	5.2993325
Guanine nucleotide-binding protein G(i) subunit alpha-2	5.8045025	6.263205
Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-1	4.93639333	5.70593667
Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2	4.98848	5.6209275
Heat shock 70 kDa protein 1A;Heat shock 70 kDa protein 1B	5.559095	4.87191333
Heat shock cognate 71 kDa protein	5.911785	5.4422525
Heat shock protein HSP 90-alpha	6.4819125	6.6376225
Heat shock protein HSP 90-beta	6.5369325	6.8719325
Heat shock-related 70 kDa protein 2	5.556005	5.67605667
Hemoglobin subunit alpha	7.3673575	6.2474
Hemoglobin subunit beta	7.47389	6.70512
Heterogeneous nuclear ribonucleoprotein A1	5.2052025	4.9129375

Heterogeneous nuclear ribonucleoprotein D-like	4.63444667	4.9584625
Heterogeneous nuclear ribonucleoproteins A2/B1	5.7932375	5.567725
High mobility group protein B1	5.3783125	5.6797275
High mobility group protein B2	5.28888333	5.44787
Histone H1.2	5.46108	5.5662075
Histone H1.4	5.990135	6.14764
Histone H1.5	5.46633	6.14913
Histone H2A type 1;Histone H2A type 1-H;Histone H2A type 1-J	6.8359525	7.1855
Histone H2B type 1-O	5.7695	6.8247525
Histone H3.2	6.349995	6.6117225
Histone H4	6.86257	7.16738
HLA class I histocompatibility antigen, B-7 alpha chain	5.32034	5.55059333
Hsc70-interacting protein	5.76532	5.8037425
Integrin beta-1	5.5760075	5.6424475
Isocitrate dehydrogenase [NADP] cytoplasmic	5.02312333	5.804035
Keratin, type I cytoskeletal 10	6.750215	6.873465
Keratin, type I cytoskeletal 14	7.0177875	7.0102375
Keratin, type I cytoskeletal 9	7.527185	7.409225
Keratin, type II cytoskeletal 1	7.726605	7.6959125
Keratin, type II cytoskeletal 2 epidermal	6.9104025	6.8608725
Keratin, type II cytoskeletal 6A	6.7710975	6.94019
Lactadherin	5.47547333	5.607875
Lactotransferrin	7.37544	6.208895
Macrophage migration inhibitory factor	5.40396	5.4298275
Malate dehydrogenase, mitochondrial	5.035965	5.2336625
Membrane-associated progesterone receptor component 1	4.91743667	5.35801667
Metallothionein-2;Metallothionein-1E;Metallothionein-1G;Metallothionein-1X;Metallothionein-1M	5.27447	4.8188725
Mimecan	5.7593125	4.56185667
Moesin	5.2775575	5.8153825
Myosin light polypeptide 6	5.5931475	5.80844
Myosin-9	5.87976	5.2727225
Nascent polypeptide-associated complex subunit alpha, muscle-specific form;Nascent polypeptide-associated complex subunit alpha	5.24909333	5.12624667
Nuclear migration protein nudC	4.15992	4.21361
Nucleolin	5.295975	5.9495625
Nucleophosmin	5.3865625	6.3835775
Parathymosin	5.5051475	5.411335
Peptidyl-prolyl cis-trans isomerase A	6.01092	5.803915
Peroxiredoxin-1	5.848685	6.114335
Peroxiredoxin-2	4.8120525	5.66473667
Phosphatidylethanolamine-binding protein 1	5.4715475	5.1021525
Phosphoglycerate kinase 1	5.6355725	5.7519725
Plectin	4.80333	5.10835667
Poly(rC)-binding protein 2	4.896735	5.26472
Pre-B-cell leukemia transcription factor-interacting protein 1	5.8072125	6.0427825
Prelamin-A/C	5.62178	5.64846
Profilin-1	6.2996975	6.1099775
Programmed cell death 6-interacting protein	5.82470333	5.084995

Prolow-density lipoprotein receptor-related protein 1	5.2986425	4.93661
Prostaglandin F2 receptor negative regulator	4.83802667	5.5183925
Protein disulfide-isomerase A3	5.43083667	5.49871667
Protein S100-A11	5.9144375	6.0286075
Protein S100-A6	5.886545	6.108425
Prothymosin alpha	6.38053	6.0361225
Pyruvate kinase PKM	6.401995	6.4278575
Rab GDP dissociation inhibitor alpha	4.69926333	5.32788
Rab GDP dissociation inhibitor beta	4.595825	5.17712
Radixin	5.84978	6.376905
Ras-related protein Rap-1b	4.89929	5.2544
Serine protease HTRA1	5.33517	5.74691
Serine/arginine-rich splicing factor 2	4.48317	5.0165675
Serotransferrin	6.309405	7.033735
Serum albumin	9.28594	9.0357175
SH3 domain-binding glutamic acid-rich-like protein 3	5.709425	5.30805
Small nuclear ribonucleoprotein Sm D3	5.3596325	5.4635825
Sodium/potassium-transporting ATPase subunit alpha-2	5.2606675	6.27805667
SPARC	5.559565	5.10529667
Spliceosome RNA helicase DDX39B	4.52387667	4.6618625
Staphylococcal nuclease domain-containing protein 1	4.74694	4.65705333
Superoxide dismutase [Mn], mitochondrial	4.1438025	5.3208525
Synaptic vesicle membrane protein VAT-1 homolog	4.72532333	5.19465667
Syntenin-1	5.737835	5.762785
T-complex protein 1 subunit gamma	6.4975425	6.82462
T-complex protein 1 subunit zeta	4.83552	5.14548333
Tetraspanin-14	7.1611975	7.37546
Thioredoxin	5.5846025	5.8375475
Thymosin beta-4	4.6811375	4.99348
Transforming protein RhoA	5.023195	5.64393
Transgelin-2	6.1219525	5.32854
Transketolase	5.3181625	5.503515
Triosephosphate isomerase	6.07249	5.7820275
Tropomyosin alpha-3 chain	5.52831	4.7526875
Tubulin alpha-1A chain	6.270645	6.6764475
Tubulin beta-2A chain	5.936335	6.3243225
Tubulin beta-4B chain	6.09906	5.96272
Ubiquitin-40S ribosomal protein S27a	5.9621925	6.034785
Villin-like protein	6.2648375	6.1174525
Vimentin	6.55442	6.6418525

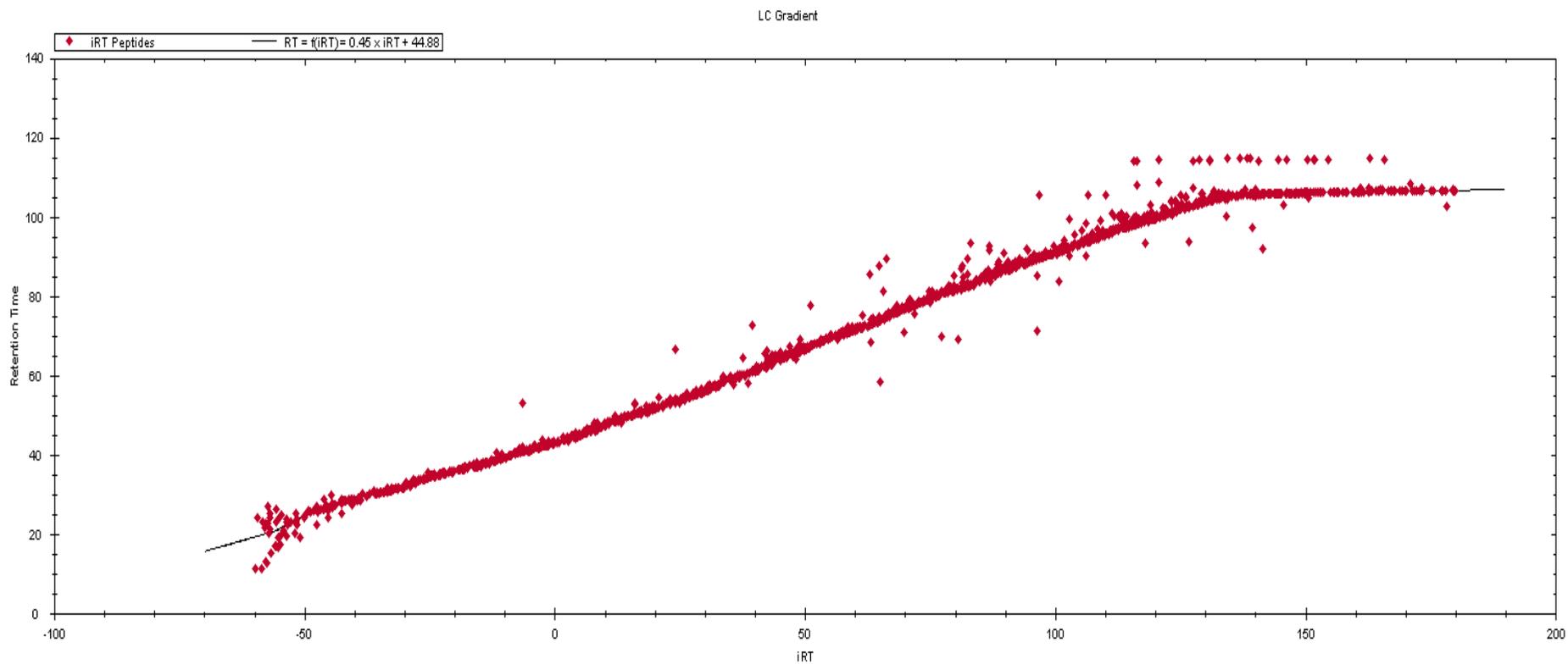


Fig S1: An example of an LC gradient from a tryptic digested glioblastoma primary cell sample (N426): iRT Calibration Chart showing the extended non-linear transformation from library iRT to actual predicted retention times.

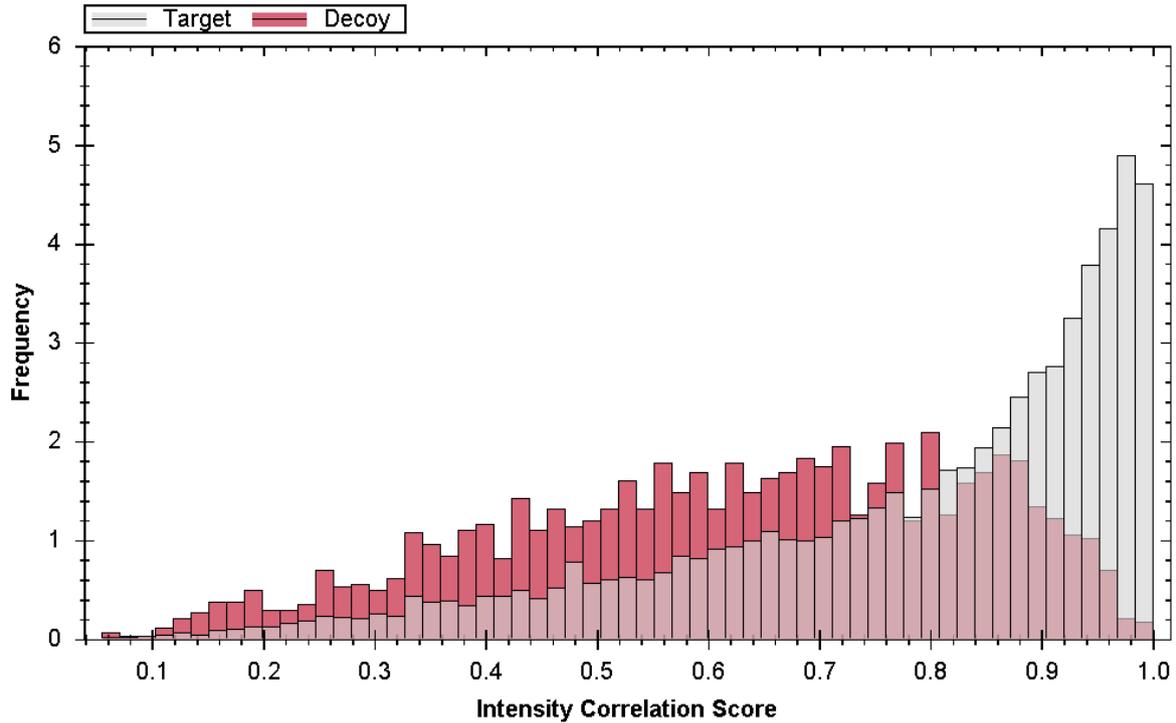


Fig S2: An example of correlation score histograms on precursor level from SWATH MS Data independent acquisition data of a tryptic digested glioblastoma primary cell sample (N426): Score for target and decoy peptides with intensity correlation score plotted as abscissa and frequency as ordinate.

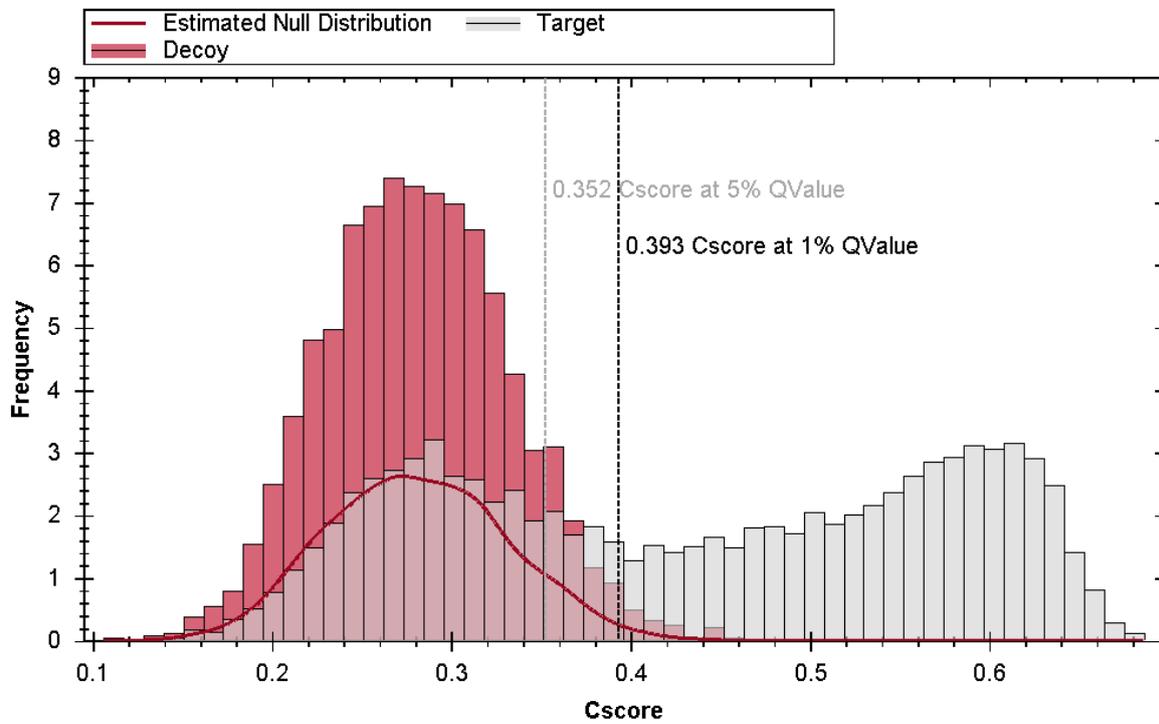


Fig S3: An example of discriminant score (Cscores) histograms on precursor level from SWATH MS Data independent acquisition data of a tryptic digested glioblastoma primary cell sample (N426): Score for target and decoy peptides.

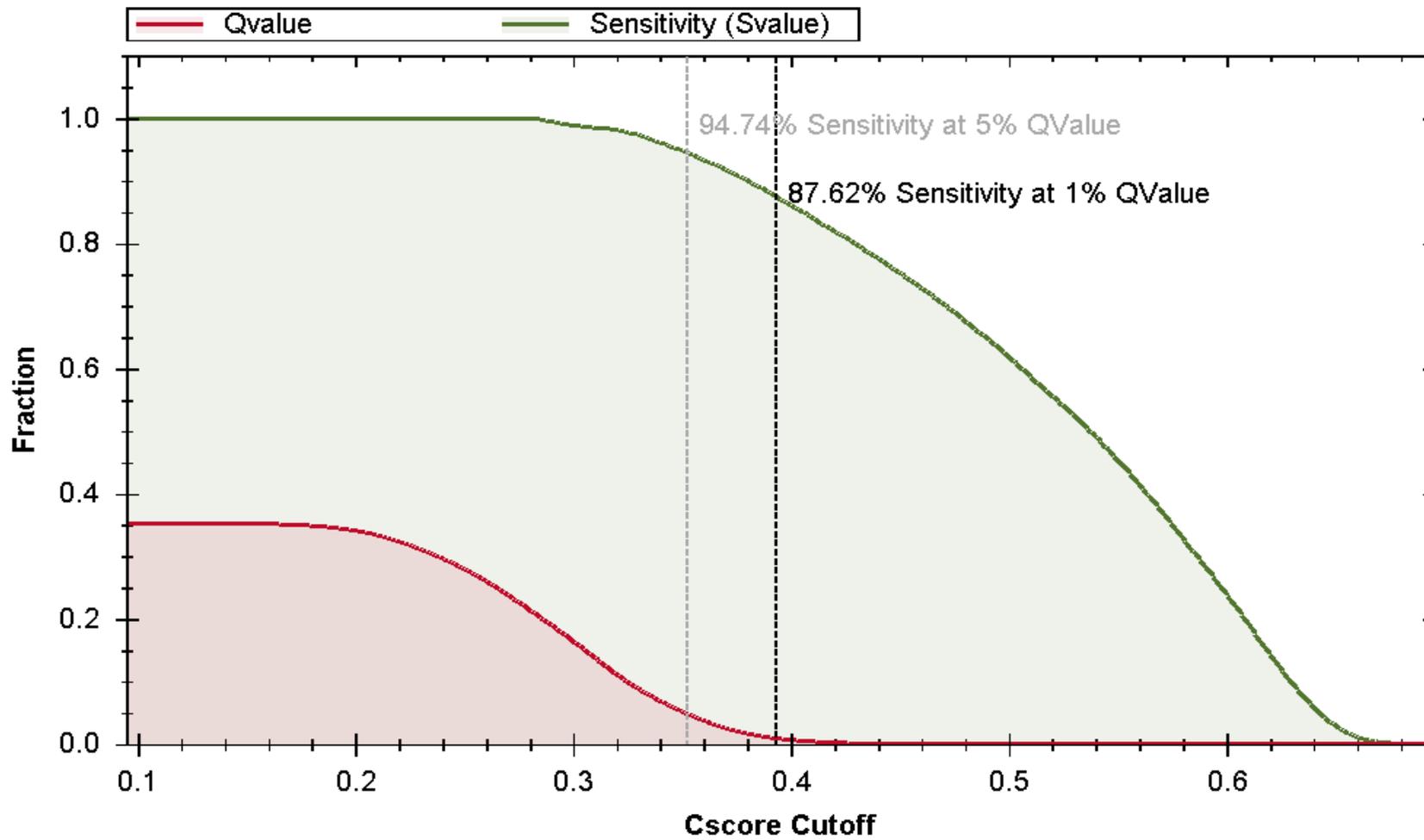


Fig S4: An example of Qvalue and sensitivity percentage plots on precursor level from SWATH MS Data independent acquisition data of a tryptic digested glioblastoma primary cell sample (N426)

8. Risk and safety statements

Table 14: Risk and safety pictograms of all chemicals used throughout this study, based on the GHS (Globally Harmonized System of Classification and Labelling of Chemicals), GHS hazard and precautionary statements.

Chemicals	GHS Symbols	GHS hazard statements	GHS precautionary statements	Disposal key
Acetonitrile (Li-Chrosolv®)	GHS02  GHS07 	H226 H314	P210 P280 P301+P330+P331 P305+P351+P338 P308+P310	4
Acetic acid	GHS02  GHS05 	H226 H314	P210 P280 P301+P330+P331 P305+P351+P338 P308+P310	4
Ammonium bicarbonate	GHS07 	H302	P301+P312 P330	2
Dithiotreitol	GHS07 	H302 H312 H332	P261 P280 P301+P310 P304+P340 P361 P501	2
Formic acid	GHS02  GHS05  GHS06 	H226 H302 H314 H331 EUH071	P210 P280 P301 + P330 + P331 P304+P340 P305+P351+P338 P308 + P310	4

Chemicals	GHS Symbols	GHS hazard statements	GHS precautionary statements	Disposal key
Iodoacetamide	GHS06  GHS08 	H301 H317 H334	P261 P280 P301 + P310 P342 + P311	2
Methanol (Li-Chrosolv®)	GHS02  GHS06  GHS08 	H225 H331 H311 H301 H370	P210 P240 P280 P302+P352 P304+P340 P308+P310 P403+P235	1
Phosphoric acid	GHS05 	H290 H314	P280 P303 + P361 + P353 P304 + P340 + P310 P305+P351+P338	4
Sodium deoxycholate	GHS07 	H302	P301+P312+P330	2
Trypsin	GHS07  GHS08 	H315 H319 H334 H335	P302 + P352 P304+P340 P305+P351 P342+P311	3
Urea	No hazardous substance or mixture according to GHS			2

9. Acknowledgement

First, I would like to thank my supervisor and first supervisor Professor Hartmut Schlüter for the subject design, the opportunity to work in a conducive workspace in his group at the University Medical Centre Hamburg-Eppendorf (UKE), his guidance and fantastic supervision throughout my PhD thesis.

I also want to thank my co-supervisor Prof. Dr. Dr. Christian Betzel of the University Hamburg for his work and opinions, insightful encouragement and all the academic support.

Special appreciation goes to Dr. Med Franz Lennard Ricklefs of the Head and Neurocenter of the department of Neurosurgery at the University Medical Centre Hamburg-Eppendorf (UKE) for his close collaboration, academic discussions and for providing all patient samples used throughout this study.

Also, I am grateful to Sönke Harder for his support in the laboratory and Core Facility affairs. I would like to also thank all our group members, Yudong, Manasi, Siti, Christoph, Hanna, Isabel, Sumera, Fritz, Laura, Dennis and Benjamin for the nice atmosphere in our lab.

To all my friends especially Beryl, Elphy and Marion, thank you for always being there.

Heartfelt gratitude to my amazing friend Parnian for all the support, question answering sessions, German translations and all the fun times we had together in and out of work. It was definitely worth it.

Special thanks to Francis Junior for editing this thesis, all the words of encouragement, motivation and support.

Finally, lots of appreciation goes to my amazing parents and siblings (Rosemary, Mirabel, Ernest and Blaise) for all their support, encouragement and endless love.

10. Declaration

Declaration on oath

I hereby declare that this thesis and the work presented in it are my own and have been generated by me as the result of my own original research. I have not used other than the acknowledged resources and aids. The submitted written version corresponds to the version on the electronic storage medium.

Eidesstattliche Versicherung

Hiermit erkläre ich, dass diese These und die darin präsentierten Arbeiten meine eigenen sind und von mir persönlich als Ergebnisse meiner eigenen Forschung erstellt wurden. Ich habe nur die anerkannten Ressourcen und Hilfsmitteln verwendet. Die eingereichte schriftliche Version entspricht der Version auf dem elektronischen Speichermedium.

Place and date

Signature