

Systematic Proteomic Profiling and Sub-Classification Glioblastoma

Background

It is my pleasure to provide my progress report for the Richard Motyka Research Fellowship that I received from Brain Tumour Foundation of Canada in July 2017. In my project proposal “Systematic proteomic profiling and sub-classification of glioblastoma” I set out to use state of the art mass spectrometry (MS) proteomic analysis of clinically and molecularly annotated glioma cohorts to define novel molecular predictors of either prognostic or therapeutic relevance. The conception of this project was based on similar experimental approaches I used to spatiotemporally define proteomic landscapes of human fetal brain development (1). In that study I optimized specialized sample preparation techniques to make use of plentiful archival tissue material stored in formalin-fixed paraffin-embedded (FFPE) blocks. I established that sectioning of this material enables us to macro-dissect and isolate different cellular layers of the brain to resolve intra-tissue heterogeneity issues for molecular profiling. With the emergence of mass spectrometry approaches to reliably quantitate global proteomic profiles, I decided to translate my developed tools to sub-classification of glioblastomas (GBMs).

GBM is the most common primary brain tumour with a dismal prognosis of <12 months, despite spirited multimodal therapy. Given that GBMs have highly variable clinical outcomes with respect to patient survival and therapeutic responsiveness, we reasoned that proteomic profiling could provide an additional layer of molecular subtyping for more refined clinical patient management. While GBMs with *isocitrate dehydrogenase* mutations (IDH-*mut*) have been shown to exhibit a superior prognosis, this subtype is rare and only found in ~5-10% of GBMs. The vast majority (~90-95%) of GBMs are IDH wild-type (IDH-*wt*) and remain without actionable or prognostic biomarkers. Interestingly, while most IDH-*wt* GBMs follow an aggressive course (baseline survival (BS) <12 months), ~20% of patients survive beyond 3 years (defined as long-term survival, LTS). So far, this biological variability cannot be explained by clinical, treatment or other genomic parameters (e.g. MGMT promoter methylation). To assess the ability of MS-based approaches to define proteomic differences between various brain tumour subtypes, and GBMs in particular, we assembled several clinically-stratified cohorts and tissue culture models (GBM stem cell-like cell lines, GSCs) of GBM development. Parallel profiling of GSCs and primary patient biopsies was designed to establish molecular subtype-specific *in vitro* models for downstream predictive chemical screen experimentation. As a complimentary project with my proteomic profiling, I have been involved in using deep neural networks (DNNs) to develop artificial intelligence (AI) classification algorithms of digitized brain tumour pathology tissue sections. Our prototype DNN is able to accurately discern and highlight tumour lesions within surgically removed material from surrounding necrotic and normal tissues and, thus, provide a workflow for macrodissecting and isolating pure tumour tissues (2, 3). Ultimately, combinatorial approaches of such computational tools with proteomic profiling of GBM tumours would increase the likelihood of identifying *bona fide* tumour-associated biomarkers for downstream validation.

Progress

MS-based glioma tumour and cell line proteomic profiling and subtyping

I am happy to report that our FFPE-based mass spectrometry analysis of glioma subtypes was recently published (4), in addition to a similar project investigating proteomic markers of meningioma development (5). To achieve this goal, we assembled three cohorts containing IDH-*wt* and IDH-*mut* GBMs, low grade oligodendrogliomas and pilocytic astrocytomas, as well as

control samples of meningiomas and medulloblastomas. From the three cohorts, our first one was assembled to include the full range of World Health Organization (WHO) grade gliomas (frozen tissues obtained from Brain Tumour Foundation of Canada, n=15), the second one includes solely higher grade glioma tumours, including IDH-*mt* and *wt* GBMs (FFPE tissues obtained locally from UHN, n=15), and the third one contains IDH-*mt* and -*wt* GSC *in vitro* models of GBMs, either in their undifferentiated state (in presence of FGF/EGF) or upon growth factor withdrawal (n=18). By growing cells in tissue culture in these differentiation states enabled us to enrich and define proteomic biomarkers of proliferating GSCs that are responsible for therapy resistance commonly observed in GBM patients.

My findings of analyzing these initial cohorts were presented at four conferences: Proteome Organization conference (Sep, 2017 in Dublin, Ireland), Advancing Precision Medicine Conference (January, 2018 in Toronto, Ontario), Human Proteome Organization conference (Sep, 2018 in Orlando, USA) and Keystone Translational Proteomics Conference (Apr, 2019 in Stockholm, Sweden). The findings in the MCP publication highlight significant outcomes from this FFPE glioma proteomic characterization. Firstly, our abbreviated MS profiling approach was highly successful in quantitatively detecting upwards of ~2,500 proteins per sample and defining proteomic signatures of gliomas of different WHO grade status, even without performing tumour lesion enrichment through macrodissection. Secondly, we found that GBM microdissection of FFPE tumours increases the likelihood of defining IDH-*mt* and -*wt* GBM tumour proteomic signatures. Thirdly, a subset of differentially abundant proteins in IDH-*wt* and -*mt* GBMs are found in similar levels in *in vitro* GSCs, specifically cultured in undifferentiated conditions. There are several candidate markers that I propose could be important in Grade IV glioma aggressiveness. The functional relevance of these proteins (ie. CLIC1 and PLOD3) is a focus of my follow-up experiments where I use genetic and chemical inhibitors to interrogate how GSC growth is affected.

In the future, I think that in order to define therapy responsiveness and GBM survival-related proteomic signatures in our cohorts of clinical outcomes it may be necessary to perform a more comprehensive proteomic profiling method, using sample fractionation that ensures quantification of even the lower abundance proteins beyond the 5,496 total proteins we have already detected. With our “shallow” proteomic coverage I identify 98 proteins that distinguish LS (>36 month survival) and BS (<13 months survival) GBMs. I expect that applying these “deep” proteome MS approaches will further expand our list of biomarkers of long-term survival and therapy sensitivity. In this long-term survival-enriched cohort, we have performed comprehensive OMIC analysis to more accurately define molecular events that guide tumour aggressiveness. These profiles were generated in collaboration with OICR and include global DNA methylation analysis, RNAseq and exome sequencing. This approach will allow us to interrogate RNA/DNA and proteomic relationships in different classes of GBM. I am in the process of optimizing proteomic fractionation-based MS approaches to more comprehensively quantitate the translational outputs of tumour tissues. This approach enables us to reliably quantitate >8,000 proteins per sample and to measure global phosphorylation levels of proteins which will enhance my ability to identify signaling cascades that are perturbed in GBMs. I believe that completion of this high impact project will result in a fantastic manuscript in the near future.

Future Directions:

I am greatly expanding our proteomic profiling toolbox by performing “deep” proteomic coverage and assessing the “phospho” proteome of our clinically stratified GBM cohorts. Once these tools

have been optimized, I will process our greatly expanded GBM cohorts to generate large datasets for inquiry into more refined biomarker identification of GBM survival and therapy response. These profiles will then be further overlaid onto datasets from GSC studies using drug screens to further determine predictive abilities of proteomic profiles to respond to different chemotherapeutic pharmaceutical agents.

References and Publications:

1. Djuric, U., Rodrigues, D. C., Batruch, I., Ellis, J., Shannon, P., and Diamandis, P. (2017) Spatiotemporal Proteomic Profiling of Human Cerebral Development. *Mol. Cell. Proteomics* 16, 1548–1562
2. Faust, K., Xie, Q., Han, D., Goyle, K., Volynskaya, Z., Djuric, U., and Diamandis, P. (2018) Visualizing histopathologic deep learning classification and anomaly detection using nonlinear feature space dimensionality reduction. *BMC Bioinformatics* 19, 173
3. Faust, K., Bala, S., van Ommeren, R., Portante, A., Al Qawahmed, R., Djuric, U., and Diamandis, P. (2019) Intelligent feature engineering and ontological mapping of brain tumour histomorphologies by deep learning. *Nat. Mach. Intell.* 1, 316–321
4. Djuric, U., Lam, K. H. B., Kao, J., Batruch, I., Jevtic, S., Papaioannou, M.-D., and Diamandis, P. E. (2019) Defining protein pattern differences among molecular subtypes of diffuse gliomas using mass spectrometry. *Mol. & Cell. Proteomics*, mcp.RA119.001521
5. Papaioannou, M.-D., Djuric, U., Kao, J., Karimi, S., Zadeh, G., Aldape, K., and Diamandis, P. (2019) Proteomic analysis of meningiomas reveals clinically-distinct molecular patterns. *Neuro. Oncol.*,

COMBINATORIAL MASS SPECTROMETRY (MS) AND ARTIFICIAL INTELLIGENCE (AI) SUBCLASSIFICATION OF DIFFUSE GLIOMA

Ugljesa Djuric¹, Ihor Batruch², Laszlo Bognar³, Tibor Hortobagyi³, Ken Aldape¹, Almos Klekner³, Phedias Diamandis¹¹ University Health Network, Princess Margaret Hospital, MacFeeters Hamilton Center for Neuro-Oncology, Toronto, Canada. ² Mount Sinai Hospital, Toronto, Canada. ³ Departments of Neurosurgery and Neuropathology, University of Debrecen, Debrecen, Hungary.

Summary

BACKGROUND: Diffuse gliomas are the most common primary brain tumors with variable prognosis. While genomic profiling efforts have identified modest genomic predictors of glioma patient survivals in ~8% of cases, to date large-scale proteomic profiles have not been performed. Similarly, little progress has been made to refine histologic classification and risk stratification of diffuse gliomas and would benefit from artificial intelligence (AI)-based image analysis of glioma biopsies using convolutional neural networks (CNNs). We hypothesize that this combinatorial analysis will improve understanding and prognosis of diffuse gliomas.

METHODS: We utilize MS and CNNs to establish (1) protein and (2) morphometry-based ("phenotypic") predictive diffuse glioma clinical subgroups. Towards AIM1, we apply a developed pipeline utilizing Q Exactive high resolution label-free quantification (LFQ) MS to characterize proteomic signatures in a cohort of diverse clinically well-annotated brain tumor specimens (n=50). Towards AIM2, we utilize a CNN-based image analysis for automated brain tumor diagnosis. We are, thus, in a position to leverage histological analytical outcomes with glioma proteomic profiles.

RESULTS: Our LFQ MS analytical method is well validated with ~2,500 protein quantifications per tumour sample identifying distinct proteomic-based glioma subtypes (ie. oligodendroglioma, astrocytoma and glioblastoma) based on 92 changes in protein abundance (p<0.005). GO term based pathway analysis demonstrates that glioma-associated molecular pathways are perturbed in correct tumour types, providing validation that MS-based proteomic measurements are identifying unbiased proteomic signatures of glioma subtypes. Similarly, training our CNN using tumour images produces a tumour identification tool enabling further glioma subtype classifications.

CONCLUSIONS: Our combinatorial approach identifies molecular- and image-based glioma subtypes and, thus, has the potential to provide precise and cost-effective clinical prognosis with faster turn-around times than classical neuropathology workflows

Results

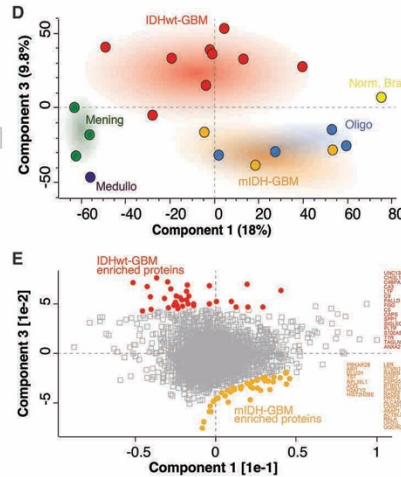
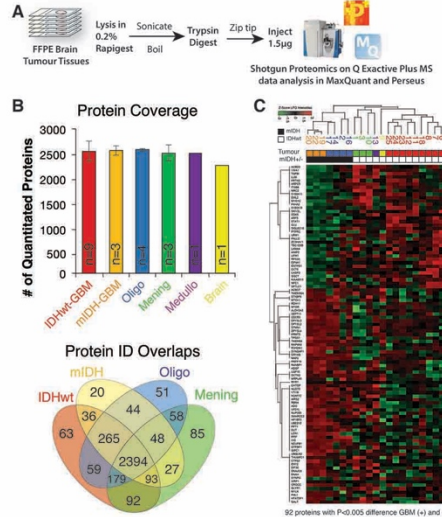


Figure 3 | Pathway analysis in Cytoscape of proteins with abundance changes between IDHwt-GBMs and IDHwt-GBMs. The more aggressive wt-GBMs exhibit an increase in epigenetic proteins involved in stem cell replication and telomere maintenance as well as deacetylase activity. mRNA processing proteins are increased in IDHwt-GBMs indicating a higher rate of RNA modifications in this less aggressive GBM subtype. MAP2K and BRAF signaling, which mediates cellular responses to cell growth signals also show protein abundance changes and demonstrates that our shotgun proteomic profiles of FFPE tissues are capturing proteins involved in cell growth.

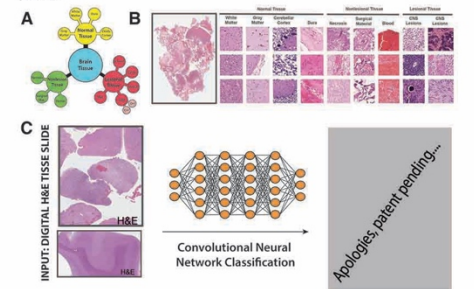


Figure 4 | Inter- and Intra-slide tissue class variability in surgical neuropathology. **A.** Examples of trained convolutional neural network training classes of tissue. **B.** WS of a glioblastoma containing a heterogeneous mixture of tumor, necrosis, normal brain tissue, blood and surgical material. The tumor comprises less than 30% of the surface area on the slide. This diversity, if not accounted for, can result in erroneously classification errors (e.g. mistaking dura for schwannoma or surgical material for calcification). **C.** Inclusion of these classes allow accurate detailed annotation of slide constituents and more accurate delineation of true lesion for future isolation or classification tasks. Example of digital slide classification of our current CNN tissue classifier.

Rationale and Aims

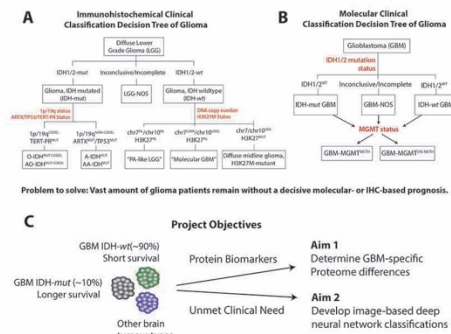


Figure 1 | Integrated histologic and molecular approaches to the classification of diffuse (A) lower grade gliomas (LGGs) and (B) glioblastoma (GBMs). **A.** Performed by immunohistochemical (IHC) assessment of the most common IDH1-R132H mutation found in ~90% of IDH-mut gliomas. If IHC is negative, sequencing for rare non-canonical IDH1/2 mutations is necessary. When IDH1/2 testing is indeterminate or incomplete, the "not otherwise specified" (Glioma-NOS) designation is used. In the setting of IDH-mutant (IDH1/2-mut) LGGs, 1p/19q codeletion (1p/19q^{codelet}) status is also assessed to characterize "molecular" WHO grade II & III (anaplastic) oligodendrogliomas (O-IDH^{mut}CODEL & AO-IDH^{mut}CODEL). IDH1/2^{mut} LGG without 1p/19q^{codelet} are categorized as astrocytomas WHO grade II & III (A/AA-IDH1/2^{mut}). The right portion of the figure addresses IDH-wildtype (IDH1/2-wt) LGGs. These often represent aggressive tumors and carry gain of chromosome 7 and loss of chromosome 10 (chr7^{gain}/chr10^{loss}) and are considered "molecular GBMs". In the absence of chromosome 7 and 10 alterations (chr7NL/chr10NL), a subset of IDH-wt LGGs have been proposed as "PA-like LGGs" based on molecular characteristics similar to pilocytic astrocytoma (e.g. BRAF/NF1 alterations). Finally, mutations in histone H3 family members (H3K27M^{mut}), can also occur. **B.** Molecular workup of a diffuse glioma with a GBM morphology is often abbreviated and most centers focus only on prognostically significant IDH mutation and MGMT promoter methylation status. **C.** Project objective is to develop robust protein-based biomarkers and deep neural network image-based classifications of glioma with improved survival and risk stratification schemes.

Conclusions and Future

Discovery of genomic changes (i.e. IDH1/2 mutations) in a small subset of diffuse gliomas has revolutionized clinical practice of modern neuro-oncology. However, additional discovery of protein biomarkers in larger, molecularly-undefined, subgroups of GBMs (e.g. IDH-wt) would provide further prognostic significance for risk stratification. Our optimized FFPE-based LC-MS/MS workflow aims to translate this promising technology to clinically stratified cohorts of diffuse gliomas. Shotgun LFQ LC-MS/MS of FFPE tissues achieves sufficient proteome resolving power to discriminate between aggressive (glioma) and benign (meningioma) brain tumour types and, importantly, between IDHwt- and IDHmt-GBMs. Candidate proteins are currently being confirmed in larger clinical cohorts. Convolutional neural networks being developed in our lab successfully stratify digital images of H&E tissue slides based on gross tissue morphology.

Future Directions

- Proteomics:**
- Profiling of a larger clinical cohort and glioma cell lines is being conducted.
 - Fractionated peptide LC-MS/MS will be performed to increase coverage and determine whether additional low abundance proteins change in glioma subtypes.
 - Post-translational modifications will be assessed in selected cases by global phosphoproteome analysis
- CNN:**
- Further training with sufficiently large image dataset from online image libraries (TCGA) and the UHN slide digitization service
 - Train CNN using well-annotated subgroups of glioma with patient metadata.

Institutions and Funding

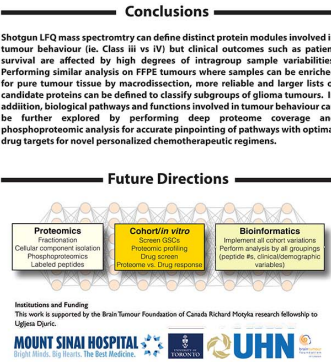
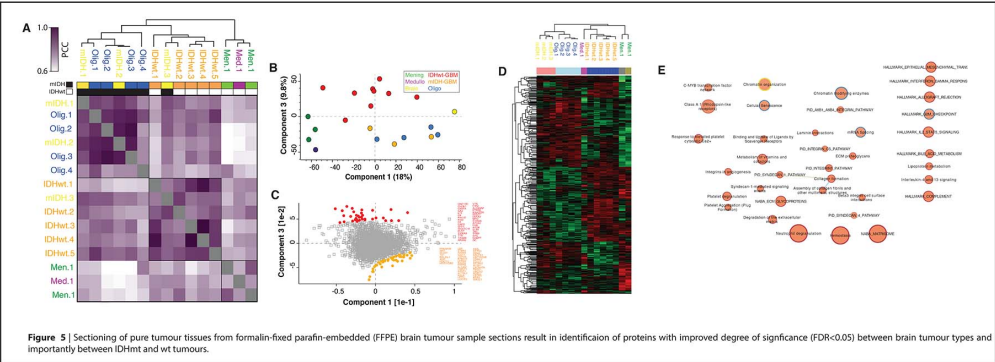
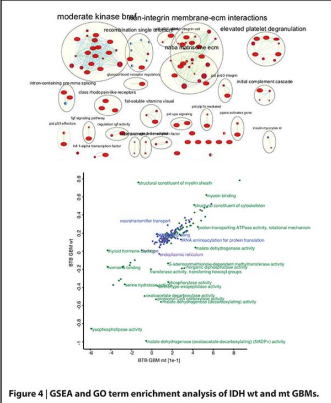
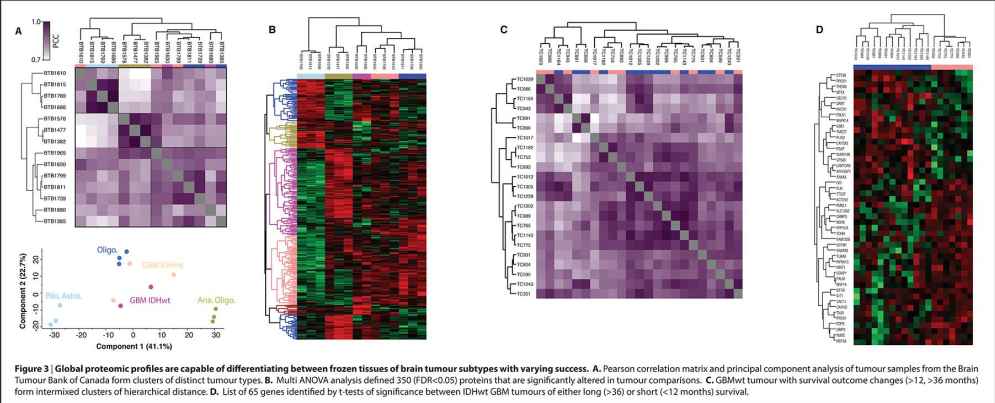
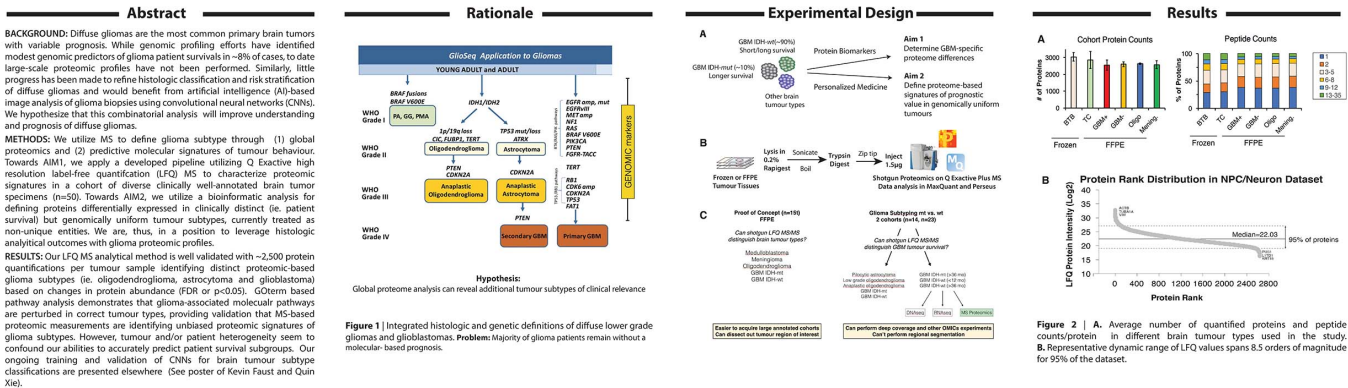
This work is supported by the Brain Tumour Foundation of Canada Richard Motyka research fellowship to Ugljesa Djuric.

MOUNT SINAI HOSPITAL **UHN**

MASS SPECTROMETRY-BASED PROTEOMIC SUBCLASSIFICATION OF DIFFUSE GLIOMAS

Ugljesa Djuric¹, Ihor Batruch², Laszlo Bogнар³, Tibor Hortobagyi¹, Ken Aldape¹, Almos Klekner³, Phedias Diamandis¹

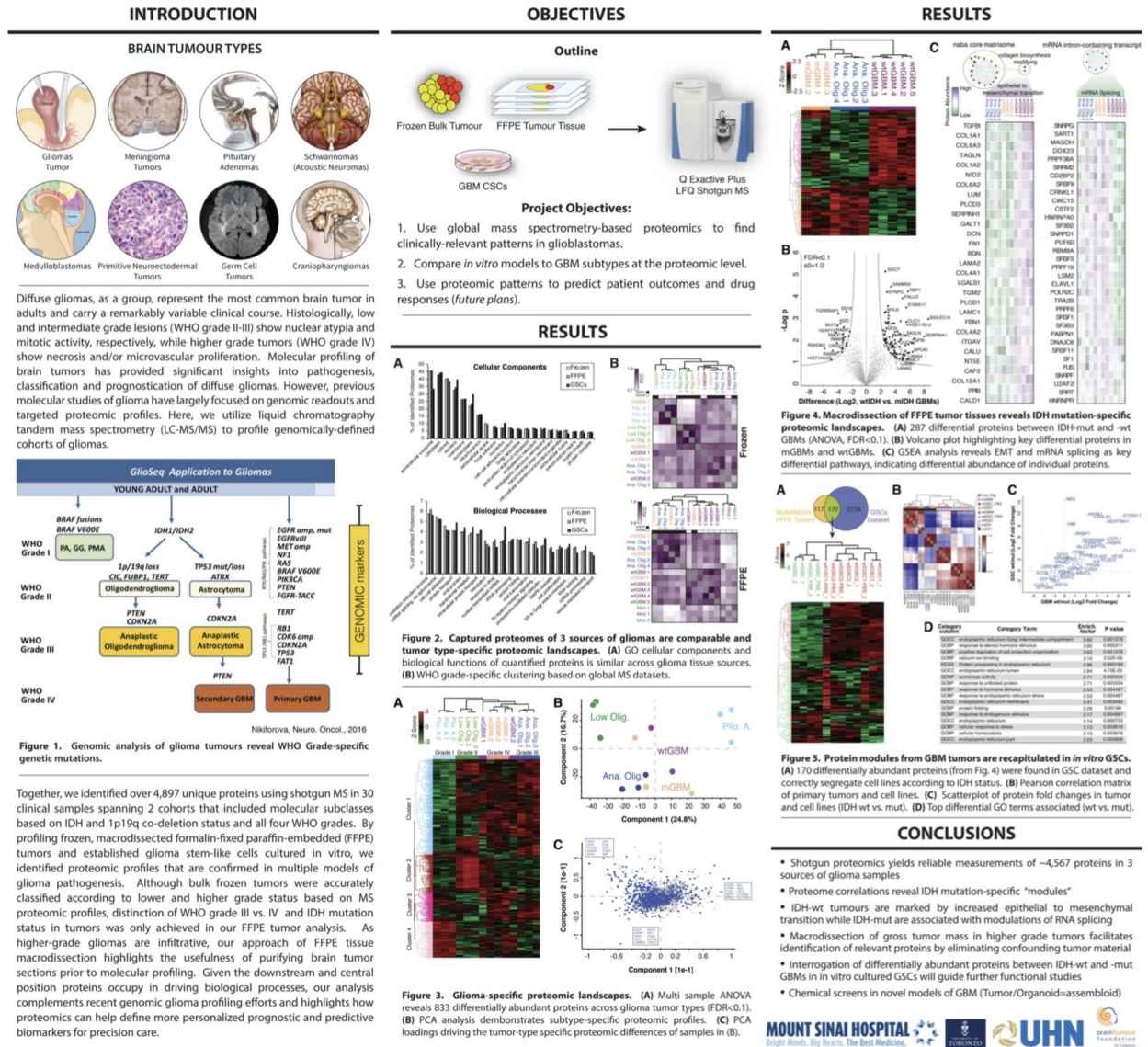
1. University Health Network, Princess Margaret Hospital, MacFeeters Hamilton Center for Neuro-Oncology, Toronto, Canada. 2. Mount Sinai Hospital, Toronto, Canada. 3. Departments of Neurosurgery and Neuropathology, University of Debrecen, Debrecen, Hungary.



MS Analysis of Glioma Tumors and Glioblastoma Stem Cells Reveal Conserved Profiles According to IDH Mutation Status

Ugljesa Djuric¹, Jennifer Kao¹, Ihor Batruch², Stefan Jevtic¹, Phedias Diamandis¹

1. University Health Network, Princess Margaret Hospital, MacFeeters Hamilton Center for Neuro-Oncology, Toronto, Canada. 2. Mount Sinai Hospital, Toronto, Canada.



- Shogun proteins yields reliable measurements of ~4,567 proteins in 3 sources of glioma samples
 - Proteome correlations reveal IDH1 mutation- and recurrence-specific biomarkers
 - IDH1-mut tumors are marked by increased epithelial to mesenchymal transition while IDH1-mut are associated with modulations of RNA splicing
 - Macrodissection of gross tumor mass in higher grade tumors facilitates identification of relevant proteins by eliminating confounding tumor heterogeneity
 - Glioma cell lines reflect tumor sub-type-specific proteomic profiles
 - Chemical screens in these novel models of GBM will utilize cell-based profiles to predict drug sensitivity
 - MS-characterized cell lines with variable proteomes
 - Chemical Screen:
 - Kinase inhibitor library (48 compounds)
 - Epigenetic inhibitor library (48 compounds)
 - Tumor-specific inhibitors