Multi-Omics Analysis of Brain Metastasis Outcomes Following Craniotomy

Jing Su1,2, Qianqian Song1, Shadi Qasem3, Stacey O’Neill3, Jingyun Lee4, Cristina M. Furdui4,5, Boris Pasche1, Linda Metheny-Barlow6, Adrianna H. Masters6, Hui-Wen Lo1, Fei Xing1, Kounosuke Watabe1, Lance D. Miller7, Stephen B. Tatter7, Adrian W. Laxton7, Christopher T. Whitlow8, Michael D. Chan6,9 and Jimmy Ruiz10,11*

1 Department of Cancer Biology, Wake Forest School of Medicine, Winston-Salem, NC, United States, 2 Department of Biostatistics, Indiana University School of Medicine, Indianapolis, IN, United States, 3 Department of Pathology, Wake Forest School of Medicine, Winston-Salem, NC, United States, 4 Proteomics and Metabolomics Shared Resource, Comprehensive Cancer Center, Wake Forest University School of Medicine, Winston-Salem, NC, United States, 5 Department of Internal Medicine, Section on Molecular Medicine, Wake Forest School of Medicine, Winston-Salem, NC, United States, 6 Department of Radiation Oncology, Wake Forest School of Medicine, Winston-Salem, NC, United States, 7 Department of Neurosurgery, Wake Forest School of Medicine, Winston-Salem, NC, United States, 8 Department of Radiology, Wake Forest School of Medicine, Winston-Salem, NC, United States, 9 Department of Radiation Oncology, University of Alabama-Birmingham, Birmingham, AL, United States, 10 Department of Medicine (Hematology & Oncology), Wake Forest School of Medicine, Winston-Salem, NC, United States, 11 Section of Hematology & Oncology, W.G. (Bill) Hefner Veterans Affairs Medical Center (VAMC), Salisbury, NC, United States

Background: The incidence of brain metastasis continues to increase as therapeutic strategies have improved for a number of solid tumors. The presence of brain metastasis is associated with worse prognosis but it is unclear if distinctive biomarkers can separate patients at risk for CNS related death.

Methods: We executed a single institution retrospective collection of brain metastasis from patients who were diagnosed with lung, breast, and other primary tumors. The brain metastatic samples were sent for RNA sequencing, proteomic and metabolomic analysis of brain metastasis. The primary outcome was distant brain failure after definitive therapies that included craniotomy resection and radiation to surgical bed. Novel prognostic subtypes were discovered using transcriptomic data and sparse non-negative matrix factorization.

Results: We discovered two molecular subtypes showing statistically significant differential prognosis irrespective of tumor subtype. The median survival time of the good and the poor prognostic subtypes were 7.89 and 42.27 months, respectively. Further integrated characterization and analysis of these two distinctive prognostic subtypes using transcriptomic, proteomic, and metabolomic molecular profiles of patients identified key pathways and metabolites. The analysis suggested that immune microenvironment landscape as well as proliferation and migration signaling pathways may be responsible to the observed survival difference.

Conclusion: A multi-omics approach to characterization of brain metastasis provides an opportunity to identify clinically impactful biomarkers and associated prognostic subtypes and generate provocative integrative understanding of disease.

Keywords: bioinformatics analysis, brain metastases, craniotomy, distant brain failure, RNA-Seq - RNA sequencing, proteomics, metabolomics, non-negative matrix factorization
INTRODUCTION

There are approximately 180,000 new cases of brain metastases in the US each year (1). However, patients with brain metastases represent a heterogeneous population and the prognosis may vary widely depending on primary tumor origin (2), molecular subset (3), histology (4), status of extracranial disease, number of lesions (5), brain metastasis volume (6), and the overall health status of the patient (7). Due to the heterogeneity in the presentation of these patients, several treatment options including surgery, stereotactic radiosurgery (SRS), and whole brain radiotherapy (WBRT) have been adopted.

The use of the various treatment options for brain metastases has evolved over time based on the relative strengths and toxicities related to each treatment regimen. Surgical resection is often reserved for larger or symptomatic brain metastases (1). SRS has proven most effective with fewer brain metastases, while WBRT is most often applied in patients with diffuse metastases (8). There is also the use of primary systemic targeted therapies for cancer that harbor actionable mutations like EGFR (9) or ALK (10). In spite of these general guidelines, a large proportion of metastases to guide management (12). There is a growing need to identify markers for disease response to systemic therapies, and there are many questions on how to use radiation and systemic therapies to treat patients and obtain optimal quality of life and survival endpoints.

There has been a number of studies that utilize single assay characterization of metastatic lesions (13, 14). This includes evaluation of brain metastasis at the gene level (15, 16). Few studies have integrated multiple levels of query to not only include the genome, but to evaluate the proteome and metabolome from the brain metastasis itself. Brain metastasis, largely from lung and breast cancers, poses a significant detriment as it relates to prognosis and quality of life. A limited number of studies have incorporated a proteomic or metabolomic based approach on human brain metastasis. There is a need to map out the signaling cascades of these tumors and if there is concordance regardless of tumor type.

Recent studies have suggested that in spite of the heterogeneity in the brain metastasis population, that there may be brain metastasis-specific mutations even across multiple histologies and primary cancer types (15, 16). Because brain metastasis patients as a population can segregate into various phenotypes of clinical behavior, the question has arisen as to whether the clinical behavior of brain metastases can be predicted. To address these questions, we profiled global proteomes, genomes, and the metabolome of resected brain metastases from a number of tumor types. We provided both individual and integrated analyses that revealed brain metastasis with similar RNA expression provided different post-transcriptional and post-translational levels.

METHODS

Patient Population
The Wake Forest Brain Tumor Tissue bank was searched for samples between 2005 and 2016. This tumor bank included fresh frozen tissue, and included samples of patients who signed an informed consent to have a portion of their tumor tissue banked. Inclusion criteria for the study included brain metastasis samples from solid tumors in which clinical follow-up data were available. After craniotomy, patients with treated with post-operative radiotherapy (either cavity-directed SRS, WBRT) or placement of breast cancer patients treated with carbamustine (BCNU)-containing wafers as previously described (17, 18). SRS was performed using the Leksell Gamma Knife B, C, or Perfexion units. Treatment planning was performed via the Leksell GammaPlan treatment planning system.

Data Acquisition
This study was approved by the Wake Forest Institutional Review Board. Electronic medical records were used to determine patient clinical characteristics as well as clinical endpoints such as survival, local failure, distant brain failure, and the likelihood of neurologic death. In general, patients were imaged every three months for the first two years after craniotomy and then every 4–6 months thereafter. Distant brain failure was defined as the development of a new brain metastasis that was not present at the time of adjuvant therapy. Neurologic death was defined as per McTyre et al. (19).

Proteomic Analysis
Prior to analysis, frozen tumor blocks were assessed by a board-certified pathologist (SQ) to assess for adequate and representative tissue. Approximately 20 mg of tissue was lysed in 1 ml of radioimmunoprecipitation assay (RIPA) buffer containing protease inhibitor using a bead mill homogenizer (Bead Ruptor, Omni International, Kennesaw, GA). RIPA lysate was then incubated sequentially with 10 mM dithiothreitol at 55°C for 30 min, and with 30 mM iodoacetamide at room temperature in the dark for another 30 min. A purified protein pellet was acquired from acetone precipitation. The pellet was subsequently treated with sequencing grade modified trypsin. The resultant peptides were de-salted using a C18 spin column, dried and then re-suspended in 5% (v/v) ACN containing 1% (v/v) formic acid for liquid chromatography-tandem mass spectroscopy (LC-MS/MS) analysis.

The LC-MS/MS analysis was performed utilizing a Q Exactive HF Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Scientific, Rockford, IL) interfaced with a Dionex Ultimate-3000 nano-UPLC system (Thermo Scientific, Rockford, IL) and a Nanospray Flex Ion Source (Thermo Scientific, Rockford, IL). An Acclaim PepMap 100 (C18, 5 μm, 100 Å, 100 μm x 2 cm) trap column and an Acclaim PepMap RSLC (C18, 2 μm, 100 Å, 75 μm x 15 cm) analytical column were used for the stationary phase. Chromatographic separation was achieved with a linear gradient consisting of mobile phases A (water with 0.1% formic acid) and B (acetonitrile with 0.1% formic acid) where the
gradient was from 5% B at 0 min to 40% B at 80 min. MS/MS analysis was performed in data dependent mode for the twenty most intense ions from the full MS scan with dynamic exclusion option for 10 s enabled. Mass spectra were searched with the Sequest HT algorithm within the Proteome Discoverer v2.1 (Thermo Scientific), in combination with the human UniProt protein FASTA database (annotated 20,193 entries, December 2015).

**Genomic Analysis**

Prior to analysis, frozen tumor block was assessed by a board-certified pathologist (SQ) to assess for adequate tumor content. Total RNA was purified from tumor specimens using the RNeasy Plus Micro Kit (Qiagen) with genomic DNA removal. RNA integrity (RIN) was determined by electrophoretic tracing using an Agilent Bioanalyzer. RNASeq libraries were constructed from samples (RIN >7.0) using the Illumina TruSeq Stranded Total RNA kit with Ribo-Zero rRNA depletion. Indexed libraries were sequenced on an Illumina NextSeq 500 DNA sequencer using 150x150-nt paired end reads, generating >40 million reads per sample (12 samples per flow cell) with >80% of sequences achieving >Q30 Phred quality scores. RNA-Seq analysis was performed following the standard pipeline (20) established by The Cancer Genome Atlas (TCGA) and the National Cancer Institute Genomic Data Commons (GDC) (21, 22). Briefly, quality of raw sequencing reads was assessed by FASTQC analysis (Babraham Bioinformatics). Sequence reads were aligned using the Spliced Transcripts Alignment to a Reference (STAR) sequence aligner (23) with a two-pass approach and gene counts determined using HTSeq software (24). The RNA-Seq data was reported as raw counts, FPKM (Fragments Per Kilobase of transcript per Million mapped reads), and FPKM-UQ (Upper 75% Quantile).

**Metabolomic Analysis**

Prior to analysis, frozen tumor block was assessed by a board-certified pathologist (SQ) to assess for adequate tumor content. The metabolomic data were generated using the AbsoluteIDQ p180 Kit (Biocrates Life Sciences, Innsbruck, Austria).

**Patient Subtyping**

The patient subtypes were determined by bi-clustering genes and samples using the RNA-Seq data and signed non-negative matrix factorization (sNMF)—an algorithm we previously developed (25, 26) to more reliably discern subpopulations defined by differential gene expression. The FPKM-UQ data, which was more robust than the FPKM one, were used for subtyping analysis. Low count inflation was controlled by discarding genes that were not detectable in more than half the samples. The FPKM-UQ expression data of each gene were then log-transformed, centered by its mean across all samples, and scaled by its root mean square. The normalized expression data were then bi-clustered using sNMF. The optimal subtype number were determined by screening the cluster number from 2 to 6 and evaluating the performance using seven different metrics (cophenetic coefficient (27), dispersion coefficient (28), explained variance (evar), residuals, residual sum of squares (RSS), silhouette, and sparseness (29)) with the randomized data as the negative control. The biclustering reached best performance when the subtype number was 4 (Supplementary Figure 1). The two largest subtypes were further analyzed.

**Survival Analysis**

Kaplan-Meier estimator and log-rank test were used for the non-parametric survival analysis (R survival package version 2.44-1.1).

**Multi-Omics Differential Analysis**

The proteomics and metabolomics data were used in enrichment analysis to identify differentially expressed proteins as well as metabolites showing differential abundance in the two largest subtypes. Briefly, low expressed proteins and low abundant metabolites which were not detectable in more than half samples were discarded. The data of each protein or metabolite were then log-transformed, centered by its mean across all samples, and scaled by its root mean square. Differential analysis between the two subtypes were performed using the R (version: 3.3.1) package “limma” (version: 3.30.13) (30), and significantly different proteins and metabolites were determined with false discovery rate (31) FDR <=0.05.

**Subtype Annotation**

The expression data of mRNAs associated with the two largest subtypes as well as proteins differentially expressed in these two subtypes were used for enrichment analysis using Ingenuity Pathway Analysis (IPA) (32). Enriched canonical pathway was scored by its -log10(p-value). The overall score of a pathway was defined as the sum of the scores generated from the transcriptomic and the proteomics data, respectively.

**RESULTS**

**Patient Characteristics**

A total of 78 patients were included in this analysis. Figure 1 depicts a diagram of the patients in this study who had brain metastasis samples analyzed for either genomics, metabolomics or proteomics. Demographic and clinical characteristics are summarized in Table 1. The 6-, 12-, and 24-month overall survival and corresponding 95% confidence intervals were in the entire cohort were 74% (65–86%), 60% (49–72%), and 45% (35–58%), respectively. Median overall survival was 17.2 months (11.1–26.9 months).

**Development of New Brain Metastases**

The distal brain failure (DBF) free survival was 10.0 month (95% confidence interval: 7.3–27.4 month). The 6-, 12-, and 24-month freedom from distant brain failure for the entire cohort and corresponding 95% confidence intervals were 68% (57–81%), 45% (33–60%), and 37% (26–52%), respectively.

**Impact of Primary Tumor Sites on Survival**

Patients with brain metastasis from breast cancers showed better survival than those from lung cancer or melanoma, with a
median survival of 32.5 months versus 13.0 and 15.9 months, respectively (Figure 2A). However, the statistical significance (p-value <= 0.08) is limited due to the modest sample sizes of the lung- and skin- originated cases (n = 15 and 12, respectively). No significant difference of DBF survival was observed among patients with different primary tumor sites (Figure 2B).

Patient Subtyping
The genomic bi-clustering analysis using sparse non-negative matrix factorization (sNMF) approach were able to characterize patients into four distinct transcriptomic molecular clusters (C.1 through C.4, Figure 3, top). Genes associated with each cluster were listed in Supplementary Table 1.

We selected the C.4 cluster as the Good Prognosis Subtype, and C.3 cluster as the Poor Prognosis Subtype for further multi-omics differential analysis. The C.3 was selected for two reasons: better overlapping with proteomics and metabolomics data, and richer transcriptomic features. The survival curves of the Good Prognosis and Poor Prognosis subtypes with regards to the likelihood of death (Figure 3, middle) and distant brain failure after surgical resection (Figure 3, bottom) of a brain metastasis showed that the C.4 subtype was associated with better prognosis, while the rest three subtypes showed poor survival. Both subtypes were mapped to the proteomics cohort with patients who had both transcriptomics and proteomics data. The similar survival difference between the Good and the Poor Prognosis Subtypes in the proteomics cohort (Figure 4) for both overall survival (Figure 4A) and distant brain failure-free survival (Figure 4B).

Multi-Omics Analysis
To comprehensively reveal the molecular profiles of the Good Prognosis and the Poor Prognosis subtypes, we further associated the proteomics and the metabolomics data and performed differential analysis. The enriched canonical signaling pathways were identified using the associated mRNAs and differentially expressed proteins, respectively, ranked and scored according to p-values, summarized together, and listed in Table 2. These pathways were further categorized

| TABLE 1 | Patient Characteristics. |
|----------|--------------------------|
| Patient Characteristics | Total (N = 78) |
| Demographics | |
| Age – yr (onset*), Mean (SD) | 59.3 (11.1) |
| Age Range | 28–85 |
| Male sex – no. (%) | 34 (44%) |
| Female sex – no. (%) | 43 (66%) |
| Race | |
| African – no. (%) | 4 (6%) |
| Caucasian – no. (%) | 73 (94%) |
| Others – no. (%) | 1 (1%) |
| Histology of Primary Tumor | |
| Lung Cancer – no. (%) | 37 (47%) |
| Breast Cancer – no. (%) | 16 (21%) |
| Melanoma – no. (%) | 14 (18%) |
| Others – no. (%) | 11 (14%) |
| Overall Survival: no. | 74 |
| Events – no. (%) | 58 (78%) |
| Followup (yr), Mean (SD) | 27.7 (31.0) |
| Time of Event (yr), Mean (SD) | 17.5 (18.9) |
| Distant Brain Failure-free Survival: no. | 75 |
| Events – no. (%) | 42 (56%) |
| Followup (yr), Mean (SD) | 13.7 (20.0) |
| Time of Event (yr), Mean (SD) | 10.7 (10.3) |
| Disease Burden | |
| Widespread – no. (%) | 36 (47%) |
| Oligometastatic – no. (%) | 33 (45%) |
| None – no. (%) | 5 (6%) |
| Unknown – no. (%) | 1 (1%) |
| Disease Status | |
| Stable – no. (%) | 8 (10%) |
| Progressive – no. (%) | 39 (50%) |
| Unknown – no. (%) | 31 (40%) |
| Karnofsky Performance Scale | |
| 50 – no. (%) | 1 (1%) |
| 60 – no. (%) | 10 (12%) |
| 70 – no. (%) | 13 (17%) |
| 80 – no. (%) | 28 (36%) |
| 90 – no. (%) | 26 (33%) |
| Unknown – no. (%) | 1 (1%) |
| Brain Metastasis at Time of Diagnosis (yr) | |
| Mean (SD) | 2.7 (3.1) |
| Range | 1–23 |
| Type of Adjuvant Local Therapy | |
| None – no. (%) | 9 (12%) |
| Gama Knife – no. (%) | 44 (56%) |
| Gliasite – no. (%) | 4 (5%) |
| WBRT – no. (%) | 10 (13%) |
| Gliadel Wafer – no. (%) | 4 (5%) |
| WBRT + Planned SRS – no. (%) | 1 (1%) |
| Fractionation IMRT – no. (%) | 1 (1%) |
| Unknown – no. (%) | 6 (8%) |

Demographics, histology of primary tumor, survival, disease burden and disease status; patients' performance, detection time of brain metastasis, and treatment information were comprehensively summarized in this table. WBRT, Whole Brain Radiation Therapy; SRS, stereotactic radiosurgery; IMRT, intensity modulated radiation therapy.
into three groups: growth, immune, and migration. The mixed pathway ranking of these three categories suggested the complexity of the tumor progress.

Differentially abundant metabolites or metabolite-based indicators were shown in Table 3, demonstrating the strength of association between metabolites and the likelihood of distant brain failure and death. Significant metabolic biomarkers include glycerophospholipids such as lysophosphatidylcholines (lysoPC a C18:2, lysoPC a C20:4, lysoPC a C20:3 and lysoPC a C20:4), phosphatidylcholines (PC ae C36:0 and PC ae C44:6), amino acids (arginine, ornithine, serine, and valine), acylcarnitines (C3, C4, and C5), a sphingomyelin (SM-OH), and a biogenic amine (carnosine). Among them, half were single metabolites and the other half were metabolite-based indicators.

**DISCUSSION**

Should the multi-omics signature for distant brain failure identified in the present study be validated, it would represent a major advance in the search for brain metastasis biomarkers. First of all, patients who, based on these findings, would be biologically at high risk of distant brain failure would certainly require post-treatment surveillance imaging. Surveillance studies have suggested that patient outcomes are improved when distant brain failures are caught prior to becoming symptomatic (33). Patients at lower risk of distant brain failure could be treated more aggressively with SRS instead of WBRT. In addition, the findings from the present analysis can potentially be used to study patient primary tumors (prior to becoming metastatic) to determine if such a signature can be applied to tumors that may be at risk of ultimately developing brain metastases.

Several previous attempts have been made to discover biomarkers for brain metastasis behavior (34–37). Dohm et al. evaluated RNA-seq data from patients who underwent fine needle aspiration for newly diagnosed non-small cell lung cancer (36). These patients were subsequently followed over the natural history of their disease and the genomics of patients who developed brain metastases were compared to those who did not develop brain metastases. Two genes were identified that with an association to development of brain metastases, but a false discovery analysis was unable to confirm this association.

The CE.7 trial is a presently accruing randomized trial assessing the efficacy of WBRT vs GKRS in the population with 5–15 brain metastases (11). This trial represents an example of a modern perspective brain metastasis study in which collected serum will be submitted for genomic analysis, and then correlated to patterns of failure. Moving forward, prospective trials with strong correlative science will likely be the next step in the evolution in the elucidation of genomic biomarkers for brain metastasis behavior. Particularly because of the heterogeneity of the brain metastasis population, large trials that bank tissue or serum for genomic analysis will be crucial in such biomarker discovery.

While the genomic and proteomic signature suggest pathways involved in metastasis development, the mechanisms to explain the metabolomic associations with distant brain failure are less clear. Early data suggests that metabolome re-programming leading to a phenotype that is more conducive for metastases is possible (38).

Lysophosphatidylcholines (lysoPCs) impact a wide range of biological and physiological functions, modulating inflammation, regulate angiogenesis, and interfere the integrity of mitochondrial membrane. Many lysophosphatidylcholines...
that showed differential abundance across the discovered prognostic subtypes are potential metabolomic molecular markers in cancers (39–42) as well as other chronic diseases such as diabetes and cardiovascular disease. For example, the abundance of lysoPC a C18:2, lysoPC a C20:3, lysoPC a C20:4, and PC ae C42:5 in plasma of cancer patients were found significantly different from healthy controls (39). The lysoPC a C18:2 abundance in serum is also a significant biomarker for
myocardial infarction (43). The increased lysoPC a C18:2 abundance in the poor prognostic subtype may reflect specific molecular traits in the immune microenvironment responsible for the disease progression. Among potential molecular mechanisms, interleukin 6 (IL-6) signaling, dendritic cell maturation, acute phase response signaling, and CDC42 signaling enriched in the poor prognostic subtype (Table 2) are known to be associated with lysoPC a C18:2.

Sphingomyelins (SM) play an essential role in brain, supporting myelination of neurons and regulating brain inflammatory responses (44). Sphingomyelin hydroxylation patterns serve as promising biomarkers for a wide range of inflammation symptoms as well as the underlying signaling events such as the activity of sphingosine kinases and the associated inflammatory signaling pathways including tumor necrosis factor-alpha (TNF-α) and interleukin1-β (IL1-β) (45). It has also been previously reported that the hydroxylation patterns on sphingomyelin backbones could be leveraged to treat cancers (46). We observed significantly decreased hydroxylation of sphingomyelins (Total SM-OH/Total SM-non OH) in the discovered poor prognostic brain metastasis subtype, probably associated with the enriched inflammation pathways in this subtype.

Some tumors exhibit a hallmark metabolic pattern known as the Warburg effect, featured by the aerobic glycolysis, the increased usage of lactate, and the up-regulated activity of corresponding enzymes including lactate dehydrogenases (LDHs) and monocarboxylate transporters (MCTs) (47, 48). In our cohort, at the mRNA level, some Warburg effect genes are negatively associated with the risk of distant brain failure, with the log fold changes of MCT genes SLC16A7 and SLC6A20 as -2.89 (FDR <= 0.0047) and -3.92 (FDR <= 0.014), respectively. However, at protein level, proteins related with Warburg effect show no significant association with distant brain failure. Furthermore, the aerobic glycolysis in the discovered Poor Prognosis group is insignificant, according to the Biocrates’ AbsoluteIDQ glycolysis metabolic indicator (log fold change: 0.10, FDR: 0.90). In summary, our results suggest that the Warburg effect is not responsible for the increased distant brain failure risk in the Poor Prognosis group.

All patients recruited in this study underwent surgical resection of brain metastasis, through which tumor tissues were retrieved. Surgical resection of brain metastasis allows for acute reduction of tumor burden and mass effect, as well as improvement in cerebral edema. This often equates to improved survival, local recurrence, in up to three metastases (49, 50). Surgical resection also allows for the retrieval of tumor tissue for diagnosis and molecular characterization. Improvements in surgical technique and radiation demonstrate improvement in survival and function.

It is still far from identifying the driver mutations that can be used to stratify patients with brain metastases and improve their prognosis. Some genetic abnormalities such as RAS/PIK3CA
(51) and EGFR (52, 53) mutations are common in the brain metastatic cases. However, the prognostic value of such biomarkers is understudied comparing with that in the primitive cancers. This work, one of the first in its kind, provides a comprehensive landscape of the molecular traits for the prognosis of brain metastasis patients, paves way to further understanding of such common but less studied mutations, and leads to mechanistic studies on the molecular underpins and possible targeted therapies.

There are several limitations to the present series. First of all, the study is limited in power based on the small population size. A larger population will be necessary in order to validate the findings. The study was also limited by patient selection bias given a single institution and the potential practice pattern biases that exist. We did not have access to patient primary tumors, and thus were unable to determine if the genomic signature for distant brain failure was inherent in the primary tumor, or evolved over the course of time to be present in the metastatic sample.

**DATA AVAILABILITY STATEMENT**

The data presented in the study are deposited in the GEO (Gene Expression Omnibus) repository, accession number GSE164150.

**ETHICS STATEMENT**

The studies involving human participants were reviewed and approved by Internal Review Board of Wake Forest School of Medicine. The patients/participants provided their written informed consent to participate in this study.

**AUTHOR CONTRIBUTIONS**

JR, MS, and MC designed the experiments. JS, SQ, and MS analyzed the data. JS, JR, MS, and MC prepared the manuscript. All authors contributed to the article and approved the submitted version.

**FUNDING**

Research reported in this publication was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR001420. The work is partial financial support from the Indiana University Precision Health Initiative to JS.

**ACKNOWLEDGMENTS**

We would like to acknowledge the support provided by the Proteomics and Metabolomics Shared Resource of the Wake Forest Baptist Comprehensive Cancer Center (NIH/NCI P30 CA12197), CF and JL.

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.615472/full#supplementary-material

**FIGURE 4** | Clinical outcomes of the corresponding proteomics cohort. (A) Overall survival. (B) Distant brain failure (DBF) free survival.

| Subtype       | n  | events | Median (mon) |
|---------------|----|--------|--------------|
| Good Prognosis (C-4) | 6  | 2      | 34.2         |
| Poor Prognosis (C-3)  | 10 | 6      | 8.1          |

Log-rank p-value: 0.005

| Subtype       | n  | events | Median (mon) |
|---------------|----|--------|--------------|
| Good Prognosis (C-4) | 6  | 2      | 42.3         |
| Poor Prognosis (C-3)  | 10 | 10     | 7.9          |

Log-rank p-value: 4e-04

**SUPPLEMENTARY MATERIAL**

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fonc.2020.615472/full#supplementary-material

**Supplementary Figure 1** | The optimization of subtype number. The optimal number of clusters, also known as the optimal factorization rank in an NMF model, was determined from an overall of many statistical metrics, including the surveys of cophenetic coefficient decrease, dispersion, explained variance (evar), residues, residual sum of squares (RSS), silhouette, and sparseness.
REFERENCES

1. Ellis TL, Neal MT, Chan MD. The role of surgery, radiosurgery and whole brain radiation therapy in the management of patients with metastatic brain tumors. Int J Surg Oncol (2012) 2012:952345. doi: 10.1155/2012/952345
2. Ayala-Peacock DN, Peiffer AM, Lucas JT, Isom S, Kuremsky JG, Urbanic JF, et al. A nomogram for predicting distant brain failure in patients treated with gamma knife stereotactic radiosurgery without whole brain radiotherapy. Neuro Oncol (2014) 16:1283–8. doi: 10.1093/neuonc/nou018
3. Vern-Gross TZ, Lawrence JA, Case LD, McMullen KP, Bourland JD, Metheny-Barlow LJ, et al. Breast cancer subtype affects patterns of failure of brain metastases after treatment with stereotactic radiosurgery. J Neuro-Oncol (2012) 110:381–8. doi: 10.1007/s11060-012-9967-3
4. Kuremsky JG, Urbanic JF, Petty WJ, Lovato JF, Bourland JD, Tatter SB, et al. Tumor histology predicts patterns of failure and survival in patients with brain metastases from lung cancer treated with gamma knife radiosurgery. Neurosurgery (2013) 73:641–7; discussion 647. doi: 10.1227/NEU.0000000000000072
5. McIntyre ER, Johnson AG, Ruiz J, Isom S, Lucas JT Jr, Hinson WH, et al. Predictors of neurologic and neuroenocologic death in patients with brain metastasis initially treated with upfront stereotactic radiosurgery without whole-brain radiation therapy. Neuro Oncol (2017) 19:558–66. doi: 10.1093/neuonc/nou184
6. Baschnagel AM, Meyer KD, Chen PY, Krauss DJ, Olson NE, Pieper DR, et al. Tumor volume as a predictor of survival and control in patients treated with gamma knife radiosurgery. J Neurosurgery (2013) 119:1139–44. doi: 10.3171/2013.7.JNS13341
7. Sperduto PW, Kased N, Roberge D, Xu Z, Shanley R, Luo X, et al. Summary of non-small-cell lung cancer: CNS ef- fects of current therapies and potential therapeutic targets. J Natl Cancer Inst (2015) 107:2561–5. doi: 10.1093/jnci/djv410
8. Sperduto PW, Kased N, Roberge D, Xu Z, Shanley R, Luo X, et al. Summary of non-small-cell lung cancer: CNS ef- fects of current therapies and potential therapeutic targets. J Natl Cancer Inst (2015) 107:2561–5. doi: 10.1093/jnci/djv410
9. McTyre E, Ayala-Peacock D, Contessa J, Corso C, Chiang V, Chung C, et al. Multi-institutional competing risks analysis of distant brain failure and salvage patterns after upfront radiosurgery without whole brain radiotherapy for brain metastasis. Ann Oncol (2018) 29:497–503. doi: 10.1016/j.annonc.2018.08.020
10. Jensen MA, Ferretti V, Grossman RL, Staudt LM. The NCI Genomic Data Commons as an engine for precision medicine. Blood (2017) 130:453–9. doi: 10.1182/blood-2017-03-736564
11. Grossman RL, Heath AP, Ferretti V, Varmus HE, Lowy DR, Kabbe WA, et al. Toward a Shared Vision for Cancer Genomic Data. N Engl J Med (2016) 375:1109–12. doi: 10.1056/NEJMmp1607591
12. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinf (Oxford England) (2013) 29:15–21. doi: 10.1093/bioinformatics/bts389
13. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinf (Oxford England) (2013) 30:1161–9. doi: 10.1093/bioinformatics/bt368
14. Liu C, Su J, Yang F, Wei K, Ma J, Zhou X. Compound signature detection on LINCS L1000 big data. Mol Biosyst (2015) 11:714–22. doi: 10.1039/c4mb00677a
15. Su J, Jin G, Votanopoulos KL, Craddock L, Bourland JD, deGuzman AF, Ellis TL, et al. Predictive genomic classifications of appendicular mucinous Neoplasms Treated with Cytoreductive Surgery and Hyperthermic Intraoperative Chemotherapy. Ann Surg Oncol (2020) 27:5:1439–47. doi: 10.1245/s10434-020-08210-5
16. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. Proc Natl Acad Sci U S A (2004) 101:4164–9. doi: 10.1073/pnas.0308311010
17. Kim H, Park H. Sparse non-negative matrix factorizations via alternating non-negative-constrained least squares for microarray data analysis. Bioinformatics (Oxford England) (2007) 23:1495–502. doi: 10.1093/bioinformatics/btm134
18. Hoyer PO. Non-negative matrix factorization with sparseness constraints. J Mach Learn Res (2004) 5:1457–69.
19. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res (2015) 43:e47. doi: 10.1093/nar/gkv007
20. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. J Roy Stat Soc B Met (1995) 57:289–300. doi: 10.1111/j.2517-6161.1995.tb20313.x
21. Kramer A, Green J, Pollard Jr, Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinf (Oxford England) (2014) 30:523–30. doi: 10.1093/bioinformatics/btu703
22. Lester SC, Takoser GB, Kuremsky JG, Lucas JT Jr, Ayala-Peacock DN, Randolph, 2nd DM, et al. Clinical and economic outcomes of patients with brain metastases based on symptomatic: an argument for routine brain screening of those treated with upfront radiosurgery. Cancer (2014) 5:1532–33. doi: 10.1002/cncr.28422
23. Li BT, Lou E, Hsu M, Yu HA, Naidoo J, Zauderer MG, et al. Serum Biomarkers Associated with Clinical Outcomes Fail to Predict Brain Metastases in Patients with Stage IV Non-Small Cell Lung Cancers. PloS One (2016) 11:e0146063. doi: 10.1371/journal.pone.0146063
24. Lee DS, Kim YS, Jung SL, Lee KY, Kang JH, Park S, et al. Identification of CD37, cystatin A, and IL-23A gene expression in association with brain metastasis: detection in advanced non-small cell lung cancer. Tumour Biol (2012) 33:1065–73. doi: 10.1007/s13277-012-0344-0
25. Shih DJH, Nayyar N, Bihun I, Daggo-Jack L, Gill CM, Aquilanti E, et al. Genomic characterization of human brain metastases identifies drivers of metastatic lung adenocarcinoma. Nat Genet (2020) 52:571–7. doi: 10.1038/s41588-020-0952-7
26. Brastianos PK, Carter SL, Santagata S, Cahill DP, Taylor-Weiner A, Jones RT, et al. Genomic Characterization of Brain Metastases Reveals Branched Evolution and Potential Therapeutic Targets. Cancer Discovery (2015) 5:1164–77. doi: 10.1158/2159-8290.CD-15-0369
37. Arora S, Ranade AR, Tran NL, Nasser S, Sridhar S, Korn RL, et al. MicroRNA-328 is associated with (non-small) cell lung cancer (NSCLC) brain metastasis and mediates NSCLC migration. *Int J Cancer* (2011) 129:2621–31. doi: 10.1002/ijc.29399

38. Shi X, Hu Z, Piskounova E, Agathocleous M, Morrison S, DeBerardinis R. Identifying metabolomic features that predict metastasis of melanoma from a primary site. *Cancer Metab* (2014) 2:67–7. doi: 10.1186/2049-3002-2-S1-P67

39. Liu C, Guo X, Sun X, Li J, Wang X, Zhou Y, et al. Mass spectrometry-based quantitative metabolomics revealed a distinct lipid profile in breast cancer patients. *Int J Mol Sci* (2013) 14:8047–61. doi: 10.3390/ijms14048047

40. Kuhn T, Floegel A, Sookthai D, Johnson T, Rolle-Kampczyk U, Otto W, et al. Higher plasma levels of lysophosphatidylcholine 18:0 are related to a lower risk of common cancers in a prospective metabolomics study. *BMC Med* (2016) 14:13. doi: 10.1186/s12916-016-0552-3

41. Klein MS, Shearer J. Metabolomics and Type 2 Diabetes: Translating Basic Research into Clinical Application. *J Diabetes Res* (2016) 2016:3898502. doi: 10.1155/2016/3898502

42. Zhang Q, Xu H, Liu R, Gao P, Yang X, Jin W, et al. A Novel Strategy for Targeted Lipidomics Based on LC-Tandem-MS Parameters Prediction, Quantification, and Multiple Statistical Data Mining: Evaluation of Lysophosphatidylcholines as Potential Cancer Biomarkers. *Anal Chem* (2019) 91:3389–96. doi: 10.1021/acs.analchem.8b04715

43. Ward-Caviness CK, Xu T, Aspeldin T, Rolle-Kampczyk U, Otto W, et al. Improvement of myocardial infarction risk prediction via in heartjnl-2016-310789

44. Schneider N, Hauser J, Oliveira M, Cazabubon E, Metzlatt SC, O'Neill BV, et al. Sphingomyelin in Brain and Cognitive Development: Preliminary Data. *eNeuro* (2019) 6:ENEURO.0421-18.2019 1–13. doi: 10.1523/ENEURO.0421-18.2019

45. Snider AJ, Orr Gandy KA, Obeid LM. Sphingosine kinase: Role in regulation of bioactive sphingolipid mediators in inflammation. *Biochimica* (2010) 92:707–15. doi: 10.1016/j.biocchi.2010.02.008

46. Jia Z, Song Y, Tao S, Cong P, Wang X, Xue C, et al. Structure of Sphingolipids From Sea Cucumber Cucumaria frondosa and Structure-Specific Cytotoxicity Against Human HepG2 Cells. *Lipids* (2016) 51:321–34. doi: 10.1007/s11745-016-4128-y

47. de la Cruz-Lopez KG, Castro-Munoz LJ, Reyes-Hernandez DO, Garcia-Carranca A, Manzo-Merino J. Lactate in the Regulation of Tumor Microenvironment and Therapeutic Approaches. *Front Oncol* (2019) 9:1143. doi: 10.3389/fonc.2019.01143

48. Reznik E, Luna A, Aksouy BA, Liu EM, Li K, Ostoenvay I, et al. A Landscape of Metabolic Variation across Tumor Types. *Cell Syst* (2018) 6:301–13.e3. doi: 10.1016/j.cels.2017.12.014

49. Ewend MG, Morris DE, Carey LA, Ladha AM, Brem S. Guidelines for the initial management of metastatic brain tumors: role of surgery, radiosurgery, and radiation therapy. *J Navel Compr Canc Netw* (2008) 6:505–13; quiz 514. doi: 10.6004/jnccn.2008.0038

50. Wang JL, Elder JB. Techniques for Open Surgical Resection of Brain Metastases. *Neurosurg Clin N Am* (2020) 31:527–36. doi: 10.1016/j.nec.2020.06.003

51. Yaeger R, Corwel E, Chou JF, Gweiritz AN, Borsu L, Vakiani E, et al. RAS mutations affect pattern of metastatic spread and increase propensity for brain metastasis in colorectal cancer. *Cancer* (2015) 121:1195–203. doi: 10.1002/cncr.29196

52. Shin DY, Na II, Kim CH, Park S, Baek H, Yang SH. EGFR mutation and brain metastasis in pulmonary adenocarcinomas. *J Thorac Oncol* (2014) 9:195–9. doi: 10.1097/JTO.0000000000000069

53. Eichler AF, Kahle KT, Wang DL, Joshi VA, Willers H, Engelman JA, et al. EGFR mutation status and survival after diagnosis of brain metastasis in nonsmall cell lung cancer. *Neuro Oncol* (2010) 12:1193–9. doi: 10.1093/neo/nos076

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.