Unravelling the role of Exosomes as early non-invasive predictors of disease state in Glioblastoma

Vinal Upadhyay
Gujarat University
Kinjal Bhadresha
Gujarat University
Rakesh Rawal (rakeshrawal@gujaratuniversity.ac.in)
Gujarat University https://orcid.org/0000-0002-7985-1187

Research Article

Keywords: Glioblastoma, qRT-PCR, Markers, Liquid Biopsy, Exosomes

Posted Date: June 7th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2988199/v1

Abstract

Background

Gliobastoma are the malignant grade 4 astrocytic tumours, which accounts for the major cause of recurrence or death amongst all the brain tumours. The biology to pathophysiology of the tumour being complex, the search for specific differential markers always remains the need of an hour for disease treatment and monitoring. Liquid biopsy approach has open newer avenues to study and regulate the disease in a non-invasive manner. The study focuses to find the some newer regulatory genes by a liquid biopsy approach.

Methodology:

Meta-analysis was carried for the already published datasets pertaining to Glioblastoma by Gene Spring software. The top listed genes were validated in tissue and exosomes of the patients. Sixty tissue samples and 30 blood (serum) samples were collected from the glioma patients. Expression analysis was carried out by quantitative real time PCR. The results were statistically analyzed using SPSS 16.0 and PRISM software.

Results

Meta-analysis fetched the top 9 highly upregulated genes inclusive of CD44, VEGFA, TGFβ1, THBS1, SERPINE1, TAGLN2, ATF3, FOSL2, FABP5. Amongst these genes, CD44, VEGFA, TGFβ1, THBS1 and SERPINE1 showed the significant differential expression between low grade and high grade gliomas tissue samples. These five significant genes when analysed in the exosomal RNA, prominent differential expression was observed.

Conclusion

The study conclusively shows that CD44, VEGFA, TGFβ1, THBS1 and SERPINE1 could collectively work as hub genes to differentiate high grade gliomas from low grade tumours. Moreover, blood serum could serve as a better liquid biopsy marker in case of high infiltrating GBM tumours.

INTRODUCTION

Glioblastoma (GBM) is the most frequent and severe malignant brain tumor [1] Glioblastoma has been linked with an extremely low five-year survival rate for more than three decades, owing mostly to its invasive qualities, and regardless of treatment, it entails surgical resection followed by radiation plus concurrent and adjuvant temozolomide[2]. Predictable diagnostic approaches include molecular investigations of tumor tissue, which are ineffective, difficult to access deeper areas, and incapable of predicting tumor heterogeneity[3].Even with the current advancements in molecular diagnostics, there are still no clinically useful biomarkers for early risk prediction. In order to help in early identification, it is crucial to establish reliable biomarkers that may be found utilizing a non-invasive liquid biopsy method in order to understand the progression and development of GBM. Clinical oncology research has been greatly influenced by gene transcriptome profiling analyses, notably in the areas of finding major tumor-related gene cohorts, developing molecular diagnoses, and determining cancer prognosis and recurrence. High-throughput sequencing technology has been applied recently to help understand the molecular pathways involved in the genesis and development of disease. The clinical care of GBM may benefit from the identification of several biomarkers that may aid in the diagnosis, prognosis, and prediction of therapeutic response. These biomarkers may also work in conjunction with imaging modalities[4].Some of the well studied alterations in GBM include gene amplification of epidermal growth factor receptor (EGFR), mutations in the tumor suppressors TP53 and PTEN, and genetic losses on chromosome 10 [5]. However, there are debatable studies that contradict the role of these markers in patients withglioblastoma[6]. Subsequently with the need to find newer stronger differential markers, one more raising concern is about repetitively using the tumour tissue biopsy for treatment decisions or disease monitoring whichonly provides a static snapshot of heterogeneous tumors that might undergo longitudinal changes over time, especially under selective pressure of ongoing therapy [3]. As a result, liquid biopsy is one such novel idea that prevents the comprehensive use of tumor tissue from disease monitoring to therapy screening and, furthermore, also aids in researching tumor microenvironment at a more precise level. Although brain cancer liquid biopsy analyzes appear indeed challenging, advances have been made in their use for detection of clinically relevant biomarkers in GBM, to aid non invasive and real time diagnosis to prognostication[7]. Exosomes are small membrane-like vesicles with a size range of 30–150 nm secreted by a variety of cells that are enclosed by a lipid bilayer membrane[8]. The exosomal cargo contains mRNA, microRNA (miRNA), DNA, proteins, lipids, and is involved in systemic intercellular communication & trafficking[9] .Thus, exosomal cargo categorization is at the center

of cancer research with abundant potential for non-invasive diagnosis, prognosis, and therapeutic monitoring. In this study, we used gene-spring software to conduct a thorough comparative meta-analysis of online accessible microarray debased data. In order to demonstrate the advantages of exosomes as liquid biopsy-based techniques, we investigated this hypothesis by using a cross-platform and cross-study meta-analysis methodology on several microarray datasets. We subsequently confirmed them in clinical tissue and blood samples. To the best of our knowledge, this is the first study to identify the relationship between tumor-derived gene regulatory networks, which may ultimately help with the early detection and prediction of GBM.

MATERIAL AND METHOD

Integrative Meta-analysis using Gene Spring software

The investigational search to selection criteria was explored considering all the PRISMA guidelines[10] were used to obtain raw data related to microarray studies underlying the malignancy of glioblastoma. A set of criteria's followed to narrow down the huge number of datasets obtained were (a) all the studies should have been carried out with human samples (homosapiens) (b) datasets encasing only tissue based study (c) studies underlying only transcriptome profiling data and (d) datasets with more than two sample size. Majorly the selection norms also only considered those datasets that were run in following platforms: Affymetrix, Agilent Technologies and Illumina Technologies. The brief selection protocol for the study is as follows (Fig. 1).

Raw data files in form of either .CEL or .TXT type was extracted from GEO tools to be further taken to be analyzed in GENE SPRING Software. These raw files were then uploaded in the GENE SPRING software by selecting platforms which was set onto a same baseline and normalized by Robust Multi-array Analysis (RMA). The isolated sample files were then classified into "Glioblastoma" and 'Normal' and re-analyzed as a single experiment. The experimental results obtained after analysis were subsequently passed on from quality control carried out by Principal Component Analysis (PCA). Furthermore, an adjusted P-value was used instead of a nonadjusted P-value to restrict the false positive rate. The cut-off criteria ≥ 2 for LogFC and < 0.05 for adjusted P-value were considered statistically significant. The results obtained as gene list was exported to excel for further study.

Functional Annotation using Bioinformatics Tools

The DEGs were further subjected to multigene STRING interaction analysis (https://stringdb.

org/) to evaluate the interactive map and identify the protein hub node based on a minimum required interaction score = 0.4. Cytoscape software (version 3.9.1) was used to visualise and analyse the PPI network. Further, GO enrichment analysis was performed using the FUNRICH tool (link), which included biological process (BP), cellular component (CC), molecular function (MF) and KEGG pathway.

Study design and participant consent

This study was approved by the ethics committee of Gujarat University, Gujarat, and all research was carried out according to the guidelines and written informed consent was obtained from all patients. The sample collection and treatment were carried out by the approved guidelines at the Vedant hospital, Ahmedabad, Gujarat. The validation of the predicted selected genes was performed in tissue and exosomal RNA derived from serum samples using quantitative real-time PCR. Ten healthy person blood samples and a total of 60 astrocytic brain tumor samples, and blood samples were collected with prior consent.All collected samples were pre-therapeutic and histologically proven astrocytoma inclusive of all grades. Patients with grade I and II were considered low-grade astrocytomas, while those with grade III and IV were classified as high-grade gliomas. The median age of the patients was 60 years at diagnosis, ranging from 30 to 85 years. Clinical-Pathological details include tumor location, histopathology, age, gender, habit, and the stage noted in each case.The tissue specimens were immediately snap frozen in liquid nitrogen after the biopsy or surgery and stored in liquid nitrogen containers until further use. Blood in containers without anticoagulant or coagulant was kept at room temperature for 30min and then at 4°C till separation (less than 4 h) to ensure serum separation. Serum samples were centrifuged at 5,000 rpm for 10 min and then at 3,000 rpm for 10 min and stored at -80°C before use.

Exosome Isolation

The serum from patients with glioblastoma and healthy controls was centrifuged at 3000 × g for 10 min at 37 ° C to remove cells and debris. The supernatant was transferred to a sterile tube and the exosomes were precipitated using a commercially available kit according to the manufacturer's protocol (miRCURY exosome kit for serum / plasma (Cat. No. / ID: 76603). The exosome pellet was resuspended in 50µL of sterile phosphate-buffered saline (PBS) and stored at -20 ° C until further analysis

Exosome Characterization

The concentration and size distribution of the exosomes were measured using a NanoSight 300 instrument (NanoSight Ltd., Amesbury, UK). All the parameters of the analysis were set at the same values for all samples and three 60-sec videos were recorded in all cases. The background was measured by testing filtered PBS, which revealed no signal.

Next, the expression of tetraspanins on isolated exosomes was investigated by flow cytometry. Isolated exosomes were stained with anti-human CD9 allophycocyanin (CD9-APC), anti-human CD63 Alexa Fluor 488 (CD63-Alexa 488) and anti-human CD81 phycoerythrin (CD81-PE antibodies (Thermo Fisher Scientific, Waltham, MA, USA). The exosomes were diluted in filtered PBS and incubated for 30 min at 37 ^{0}C antibody solutions with optimalpredetermined concentrations. The samples were then analysed using the FACS Calibur™ Flow Cytometer (Becton Dickinson Biosciences, San Jose, CA). Particle coincidence was assessed by acquiring different dilutions of stained exosomes. Events and seconds were recorded at each dilution using the high sample flow rate on the instrument. The lowest dilution with which a linear correlation could be observed between the dilution and events/second was considered optimal[11].

RNA extraction from tissue and exosomes

Total RNA was extracted from 30mg of tumor tissue using the RNeasy tissue kit (Qiagen 74104) based on silica gel membrane technology by selective binding, stepwise washing, and elution of RNA following the manufacturer's instructions. A digestion step on the spin column was performed using the RNase-free DNase set (Qiagen 79254). The concentration of isolated RNA was quantified with a Qubit 3.0 Fluorometer (Invitrogen by Life Technologies, CA, USA). The extracted RNA was then stored at -80 ° C until further analysis.

Exosomal RNA was extracted using Ardia Total DNA and RNA extraction kit according to the manufacturer’s instructions by magnetic bead extraction procedure. 200µl of the exosomes was taken in the defined sample plate and 20µl of the Proteinase K was added to the exosomes. The sample plate was further kept in the automated DNA /RNA extraction instrument for further extraction steps. The final RNA product obtained was measured and quality assured by 260/280 ratio by Nanodrop (Epoch BioTek System)

Quantitative real-time PCR

Quantitative real-time polymerase chain reaction assays based on SYBR green were used for gene expression analysis using gene-specific primers for CD44v6, SERPNE1,TGFβ1, VEGFA, THBS1, ATF3, TAGLN2, FABP5, FOSL2 and 18sRNA as an endogeneous control. The reaction mixtures (20 µl) consisted of 10 µl Brilliant III Ultra-Fast SYBR® Green QPCR Master Mix (Agilent Technologies, USA), 0.4 µl (400 nM) each of the forward primer & reverse primer, and 2 µl cDNA. Amplification was carried out using QuantStudio 5 (Applied Biosciences) under the following cycling conditions: 1 cycle of 3 min at 95 ° C for the initial denaturation step and 40 cycles of 5 s at 95 ° C for the denaturation step, 20 s at 60 ° C for the annealing and extension step. Melting curve analysis was performed after the amplification to distinguish the accumulation of specific reaction products from nonspecific products or primer dimmers. 18sRNA was used as an endogenous control. The relative expression of individual mRNAs in each case was calculated using the 2 –ΔΔCt method. All PCR reactions were performed in triplicate. Statistical analyses were performed using the GraphPad Prism 5.0 software. The list of all primers used for the study is shown in **Table 1**

Table 1 : PRIMER SEQUENCE				
Sr	GENE		PRIMER	Bases
1	18sRNA	F	5′-GGAGTATGGTTGCAAAGCTGA-3′	21
		R	5′-ATCTGTCAATCCTGTCCGTGT-3′	21
2	SERPINE1	F	5′-ATCGAGGTGAACGAGAGTGG- 3	20
		R	5′ -ACTGTTCCTGTGGGGTTGTG – 3′	20
3	CD44	F	5'-GGAGCAGCACTTCAGGAGGTTAC-3'	23
		R	5'-GGAATGTGTCTTGGTCTCTGGTAGC-3'	25
4	ATF3	F	5'-CCTCGGAAGTGAGTGCTTCT-3'	20
		R	5'-ATGGCAAACCTCAGCTCTTC-3'	20
5	THBS1	F	5'-TTGTCTTTGGAACCACACCA-3'	20
		R	5'-CTGGACAGCTCATCACAGG-3'	19
6	VEGFA	F	5′-CCTTGCTGCTCTACCTCCAC-3′	20
		R	5′-ATCTGCATGGTGATGTTGGA-3′	20
7	FABP5	F	5′-GCTGATGGCAGAAAAACTCAGA-3′	22
		R	5′-CCTGATGCTGAACCAATGCA-3′	20
8	FOSL2	F	5'-GAGAGGAACAAGCTGGCTGC-3'	20
		R	5'-GCTTCTCCTTCTCCTTCTGC-3'	20
9	TAGLN2	F	5'-CTACCTGAAGCCGGTGTCC-3'	19
		R	5'-ATCCCCAGAGAAGAGCCCAT-3'	20
10	TGFβ1	F	5′-ACTGAACCTGACCGTACACTGCAGCCTCCAGCCAAC-3′	36
		R	5′-GTCTGCAAGTTCATCCCCTCTT-3′	22

.

Hierarchical Clustering and ROC curve analysis

The expression profile in low-grade and high-grade gliomas could be differentiated specifically by the clustering technique. Unverified hierarchical clustering was achieved using the hcluster method of the R package ‘a map’ and the plot was created using the heatmap.2 function of the package ‘gplots’. Absolute Pearson and Pearson distances were used to calculate gene and sample distances, respectively, and gene linkages were done using the Ward algorithm. Inter-study normalization was completed with the bioconductor package ‘In-Silico merging’ using an empirical Bayes method. Correlation diagrams were obtained using the “corrplot” library in R-project, of tissues and exosomal mRNA expression levels in low grade and high grade tumours.

Receiver operator characteristics curves (ROC) were generated and AUCs of each classifier were calculated using MedCalc (Belgium, Europe). To understand the false positives and/or weaknesses of our classifiers, images frequently misclassified by the classifiers were also reviewed.

Statistical Analysis

The GraphPad Prism software, version 8.0.1(www.graphpad.com), was used to conduct statistical analysis. Three sample groups were compared using one-way analysis of variance (ANOVA). $P < 0.05$ was used as the significance level.

RESULTS

Identification of significant DEGs using GENESPRING

The meta-analysis workflow in the present study was shown in Fig. 1. To identify the gene expression signature between healthy donors and brain tumor patients, a total of 22 data sets were simultaneously analyzed by Gene Spring software. Detail of the 22 datasets that

summarised in **Table 2**.

Table 2 : LIST OF DATASETS INCLUDED IN THE STUDY

Sr#	Accession ID	Name	Sample	Tumour	Control	Platform
1	GSE139380	Aberrant Expression Of RSK1 Characterizes High-Grade Gliomas With Immune Infiltration	30	30		[HTA-2_0] Affymetrix Human Transcriptome Array 2.0 [transcript (gene) version]
2	GSE131837	Gene Expression Profiles Of Non-Recurrent Human Glioblastoma Tissues	52	52		Illumina HumanHT-12 V4.0 expression beadchip
3	GSE116520	Transcriptome Profiling Reveals PDZ Binding Kinase As A Novel Biomarker In Peritumoural Brain Zone Of Glioblastoma	42	42		Illumina HumanHT-12 V4.0 expression beadchip
4	GSE122498	Gene Expression Data Of Glioblastoma Patients From GAPVAC Trial (Glioma Actively Personalized Vaccine Consortium)	17	16	1	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
5	GSE108476	The REMBRANDT Study – A Large Collection Of Genomic Data From Brain Cancer Patients	2056	380	28	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
6	GSE90886	Expression Data From Glioblastoma Tissue And Normal Brain Tissue Samples (From Epilepsy Surgery)	18	9	9	[PrimeView] Affymetrix Human Gene Expression Array
7	GSE85033	A Machine Learning Classifier Trained On Cancer Transcriptomes Detects NF1 Inactivation Signal In Glioblastoma	12	12		[HTA-2_0] Affymetrix Human Transcriptome Array 2.0 [transcript (gene) version]
8	GSE83537	Spatial Transcriptome Analysis Reveals Notch Pathway-Associated Prognostic Markers In IDH1 Wild-Type Glioblastoma Involving The Subventricular Zone	36	36		Illumina HumanHT-12 V4.0 expression beadchip
9	GSE62802	Relative Spatial Heterogeneity Revealed By Transcriptional Profiling Of Multi-Region High-Grade Glioma Samples	20	20		[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
10	GSE62153	Implications Of The Heterogeneous Gene Expression Traits In Recurrent Glioblastoma	43	43		Illumina HumanHT-12 V4.0 expression beadchip
11	GSE51062	Expression Data From Human Gbms	52	52		[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array

Table 2 : LIST OF DATASETS INCLUDED IN THE STUDY						
12	GSE65626	Expression Data From Glioblastoma Tissue And Matched Adjacent Normal Tissue	12	3	3	[HTA-2_0] Affymetrix Human Transcriptome Array 2.0 [transcript (gene) version]
13	GSE53733	Expression Data From Primary Glioblastoma In Adults	70	70		[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
14	GSE52009	Whole Genome Expression Profile Of 120 Human Glioma Samples	120	24		Agilent-014850 Whole Human Genome Microarray 4x44K G4112F (Probe Name version)
15	GSE36245	Gene Expression Data From Glioblastoma Tumor Samples	46	46		[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
16	GSE25630	Genome-Wide Analysis Of Gene Expression In Surgical Specimens Of Primary Glioblastoma Multiform	21	21		Illumina HumanWG-6 v3.0 expression beadchip
17	GSE30563	Gene Expression Data From Human Brain Tumor Or Normal Brain	6	1	3	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
18	GSE15824	Primary And Secondary Brain Tumors: Glioblastomas, Astrocytomas And Oligodendrogliomas	45	23	5	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
19	GSE19728	Expression Data From Different Grades (WHO) Of Astrocytomas (ACM)	21	5	4	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
20	GSE7696	Glioblastoma From A Homogenous Cohort Of Patients Treated Within Clinical Trial	84	80	4	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
21	GSE4290	Expression Data Of Glioma Samples From Henry Ford Hospital	180	81	23	[HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array
22	GSE4412	Gliomas Of Grades III And IV (HG-U133B)	85	50		[HG-U133B] Affymetrix Human Genome U133B Array

The 22 datasets were uploaded in succession and each dataset was processed by matching gene ID, annotation of the sample (control / patient) and individual DEGs identification. To remove batch effects between various datasets, 'ComBat'-based batch effects adjustment based on 'ComBat' was performed and the results with and without adjustment were visualized by PCA plots, respectively. After that all the 22 gene expressional microarray data were then combined and merged. Meta-analysis was conducted following the Robust Multi-array Analysis (RMA) and Benjamini Hochberg FDR statistical methods, which facilitated to reveal the DEGs between healthy donors and brain tumor patients across different microarray datasets by permitting variable true effect size and integrating unknown cross-study heterogeneities. We found a total of 13,477 genes across the 22 datasets with significance threshold of adjusted *P*-value < 0.05. A total number of 537 DEGs from GBs were identified in which primary tumor samples were compared with normal tissue cells. From the list of these genes, we narrowed the top nine genes based on expression and role in the GBM which were further used as panel to differentiate high grade primary GBM from the low grade disease. The list included SERPINE1, THBS1, TGFβ1, VEGFA, CD44, ATF3, FOSL2, FABP5, and TAGLN2, which were further reviewed for validation.

Gene Panel Investigation by Bio-informatic Approach

Initially, the gene set narrowed down was studied for its protein level interaction by STRING Database. The results showed a strong bond interaction of SERPINE1, THBS1, TGFβ1, VEGFA, CD44, making a reasonable network among each other. ATF3 and FOSL2 again showed a strong bonding with each other, but show no link with other genes. FABP5 and TAGLN2 had a good connection between each other but a weak connection with other genes. **(Fig. 2A)**

Functional annotations were also studied for the same set of genes to understand its molecular functions, the biological process they are involved in, to which cellular components they are present in and their major site of expression. All the data was generated using FUNRICH database.

The annotations show a significant expression of these genes in cerebrospinal fluid and malignant glioma samples. A decrease in expression is seen in the exosomes, which might be good signal to support the liquid biopsy approach. **(Fig. 2C)**

This CIRCOS plot signifies the exact locations of the genes on respective chromosome and the major functions they are involved in. All the genes in some or other way are either related to proliferation, angiogenesis, migration or EMT process or act as chemo-resistance genes which need major modulation for the drugs to function efficiently **(Fig. 2B)**.

These gene sets were thoroughly investigated in order to track the survival outcomes of patients, particularly Glioblastoma patients with expression variations in these genes. Thus, we use the UALCAN program to perform the Kaplan Meir survival data for these genes. The survival graph analysis reveals that patients with greater levels of CD44, SERPINE1, and FOSL2 expression have a significantly lower chance of surviving. However, as seen in Fig. 3, the remainder of the genes showed a rapid decline in patient survival with greater expression but this was not significant.

Characterization of serum exosomes derived from brain tumor patients

Serum exosomes were isolated from patients diagnosed with Brain tumor. These exosomes were characterized by NTA and flow cytometry. The NTA results showed a single peak for exosome concentration in the size range of 30–50 nm **(**Fig. 4A**)**. Serum exosome-derived brain tumor patients had a mean size of 37+/- 0.0 nm and a concentration of 1.36×10^8 particles/ml. The exosome isolated form brain tumor patients contained higher number of exosomes and smaller in size. The AFM result indicated that a sphere-shaped vesicle with a mean radius of 50–80 nm, which was consistent with the NTA profiles **(**Fig. 4A**).**

Furthermore, flow cytometry analysis shown the existence of tetraspanins CD9, CD63 and CD81 (86–88% exosomes expressed these markers) in brain tumor serum exosomes, therefore results confirm that isolated vesicles comprise pure exosomes (Fig. 4B**)**

Gene expression pattern in low grade and high grade glioma tissue and exosomes

We looked at the overall expression of these genes in our samples based on a cohort of 60 patients **(**Fig. 5**)**. We gathered samples from patients with high-grade and low-grade gliomas to support the expression of these genes in tissue samples.TGFβ1 had the highest total expression relative to the other genes, according to total gene expression. However, the expression levels of CD44, THBS1, SERPINE1, and VEGFA were likewise respectable and reached values of over 10 times. ATF3, FOSL2, FABP5, and TAGLN2 demonstrated a modest amount of expression above the required range of 2 times. The differential gene list refined by meta-analysis observed to over express in high grade gliomas, was used to establish a gene expression pattern that may distinguish High Grade Gliomas from Low Grades. In GBM samples, it was shown that the expression of certain genes (THBS1, SERPINE1, VEGFA, TGF-1, CD44, FOSL2, FABP5, ATF3, and TAGLN2) was

upregulated (above 2 fold). We observed some substantial over-expression of CD44, THBS1, TGF1, VEGFA, SERPINE1, FABP5, and ATF3 in high grade glioma samples. As shown in Fig. 30, expression levels of CD44, THBS1, TGF1, VEGFA, SERPINE1, FOSL2, TAGLN2, and FABP5 were downregulated in low grade samples. Genes CD44, THBS1, TGF1, VEGFA, and SERPINE1 were equivalent to the other genes in their significant differential expression pattern. The expression of FOSL2, FABP5, ATF3, and TAGLN2 was close to normal and they were unable to distinguish between high-grade and low-grade gliomas.

The expression of the top 5 hub genes in exosomes was also examined because these genes hadhigh expression in tissue samples. The expression of the following genes was examined in exosomes: CD44, TGF1, THBS1, SERPINE1, and VEGFA. According to the histological grade, the samples were divided into low grade (n = 15) and high grade (n = 15) gliomas. In both grade samples, CD44, TGF-1, and SERPINE1 had significant (p0.001) values. However, in both patient classes, CD44 and SERPINE1 had the same expression pattern. The results were extremely significant since TGF-1 differently displayed an opposing pattern of expression, with low expression levels seen in low-grade gliomas and strong upregulation in high grade gliomas. However, in both sets of patients, VEGFA has low levels of expression and THBS1 exhibits high levels, although the findings are not statistically significant.

Hierarchical Clustering

It was necessary to conduct the confirmation test to see whether the nine chosen genes have the capacity to categorize high-grade gliomas and low-grade gliomas into distinct groups. As a result, hierarchical clustering method was utilized to categorize a group of 60 patients with gliomas based on the expression of these 9 genes. In order to determine if the tumors are strongly correlated or not with one another, the tumors were divided into two unique groups based on their pathological profiles, i.e. high grades and low grades .The study's main hypothesis was that it would find important genes that might specifically distinguish glioblastoma samples from low grade glial samples. The THBS1, TGF1, SERPINE1, VEGFA, and CD44 gene clustering demonstrates a distinct separation of high- and low-grade gliomas. (Fig. 6A)

Next, Hierarchical Clustering was used to examine the relationships between the samples as well as the inter-correlation of the genes in exosome marker expression. The expression of TGF-1 differed noticeably between low-grade and high-grade gliomas, as shown by the genes' strong clustering together. All other genes express themselves to some extent, but they do not distinguish between classes.

Measurements of Inter-correlation amongst the genes

As previously noted, we compared the ratios between the levels of expression of different genes in several cohorts to see whether differences in gene expression are linked. In tissue samples, ATF3, FOSL2, and VEGFA displayed substantial intercorrelation ($p < 0.001$), whereas THBS1 was not associated with either of the aforementioned markers. The correlation between VEGFA and ATF3, FOSL2, and FABP5 was also shown to be significant ($p < 0.001$). The association between TGF1 and SERPINE1 in the tissue sample was substantial ($p < 0.001$), but the correlation between CD44 and SERPINE1 was favorable.

The exosome cohorts' gene relationships were highly correlated, and the correlation pattern matched the glioblastoma pattern. According to our research, tumor tissue and exosomes exhibit different patterns of connection. Additionally, VEGFA, CD44, SERPINE1, and TGF-1 were not only re-expressed but also significantly connected with one another, indicating the development and proliferation of glioblastoma cells, which offers practical natural models for researching the glioblastoma process and mechanism. While SERPINE1 and CD44 and VEGFA exhibited substantial correlations, TGF1 also showed a link with SERPINE1, CD44, and VEGFA (Fig. 6B), indicating their potential participation in the mechanism of glioblastoma development.

Analysis of Sensitivity and Specificity of the Marker by Reciever's Operative Curve (ROC)

In order to assess the accuracy of gene expression in identifying high-grade tumors and low-grade tumors in the glioma samples, ROC curve analysis was carried out. Figure 7A shows the results of our ROC analysis for individual genes, and lists their sensitivity, specificity, and AUC. According to the ROC curve analysis of the various genes, CD44, TGF1, THBS1, SERPINE1, and VEGFA, in comparison, shown higher accuracy with the significance of the research and proved to be more specific and sensitive towards illness grade categorization. These genes demonstrated sensitivity and specificity of 100% and an AUC of 1, indicating their importance in differentiating between high-grade and low-grade gliomas.

We examined the specificity and sensitivity of the genes for high grade gliomas and the low grade gliomas from exosomes to confirm the effectiveness of these hub genes for their expression in exosomal RNA. Of all the genes, TGF1, SERPINE1, and CD44 displayed substantial data with high sensitivity and specificity. The more effective marker for grade distinction, TGF-1, on the other hand, demonstrated 100% sensitivity and around 93.3% specificity with an Area under curve of roughly 0.969. (Fig. 7B)

DISCUSSION

Multiple clinical and molecular research over the last decade have proven a predictive significance for various signal transduction molecules involved in tumor development, invasion, and homing, which eventually leads to glioblastoma. Despite breakthroughs in GBM therapy, no positive results have been reported; individuals diagnosed with these tumors often have a poor prognosis and low quality of life as the disease progresses[12]. A new generation of anti-GBM medicines, such as vaccinations, antibody-based drug conjugates, and more recently have all been made possible by the precise characterization of molecular signatures. These therapies allow for a more individualized therapeutic strategy. Additionally, a lack of knowledge on the biology of GBM tumors and the processes behind the development of treatment resistance in recurrent GBMs may be the cause of the absence of a major progress in GBM therapy[13]. Due to the unavailability of some highly important markers for the study of the aggressive cancer, the need to find novel differentially expressed markers that may be further transformed into treatment modalities and become key predictors for the diagnosis of the disease develops. As a result, we performed extensive meta-analysis for GBM tissue data using publicly accessible microarray datasets in this work. Additionally, these genes cohort were examined for their underlying functional mechanism for early GBM prediction utilizing the liquid biopsy method and verified on patient tumor tissue and serum exosomes.

Expending the commercial pathway knowledge we established pathways that were developed or over-represented in the common glioblastoma signature.

Bioinformatic analyses have been carried out in several prior research to study the involvement of differentially expressed genes (DEGs) in various pathways, molecular functions, and biological processes in patients with GBM [14–16]. The meta-analysis-based approach is a better method for future research gathering the results of microarray experiments done globally and submitted to NCBI. In order to find new glioblastoma supported biomarkers, we performed meta-analysis on microarray gene expression datasets taken from patients with high-grade tissue samples. We retrieved the archived gene expression data sets derived from high grade gliomas analyzed using GeneSpring software. The primary goal was to identify genes that are highly expressed in GBM samples. We identified nine important genes that were often overexpressed in glioblastoma patients **(**Fig. 8**).** These gene sets were then tested in 60 different patient tissue samples. Because exosomes are typically shed from primary tumors and reach the bloodstream, we discover some of these prioritized genes in actual patient blood samples, which may potentially lead to the development of novel biomarkers to predict glioblastoma. Based on the tissue expression data, we discovered a substantial differential gene expression change between low and high grade gliomas in particular genes, including CD44, SERPINE1, VEGFA, TGF1, and THBS1, hence we focused our exosomal RNA expression study to these genes. We found that not all of the 9 genes identified by meta-analysis and tissue expression are significantly overexpressed in exosomal cargo. This suggested that tumor-derived transcription may be packaged differently during exosome formation.

TGFβ1, VEGFA, THBS1, CD44, and SERPINE1 were the most significant sets of genes expressed in tissues and exosomes in high grade gliomas. TGFβ1 is known to stimulate cell proliferation[17], invasion[18], angiogenesis[19], immunological suppression [20], and glioma stem cell activity. Furthermore, SERPINE1 has been shown to modulate GBM cell-substrate adhesion and directional motility, and its expression is regulated by TGFβ signaling [21]. Data demonstrate that VEGFA, FLT1, and KDR have higher levels of expression in almost all brain tumors when compared to normal brain vasculature [22]. In particular, higher grade gliomas were shown to have more VEGFA than low grade gliomas [23]. It is widely established that THBS1 influences immune responses and GBM vascularization [24, 25]. High levels of THBS1 expression have been seen in high-grade glioma patients, and inhibiting the gene prevented these cells from growing and invading[26]. CD44 has been linked to angiogenesis, proliferation, invasion, and migration, and it may enhance EMT in glioma [27]. We also conducted a ROC analysis of these genes in combination with the function of these markers to assess how well they outperformed the other markers in terms of disease prediction accuracy. The other genes' specificity and sensitivity were less than 80% for this group of 5 genes. These five markers (THBS1, TGF1, SERPINE1, VEGFA, and CD44) showed high interaction when their interaction data was examined, but the other markers had weaker connections. These indicators have the potential to be employed as distinguishing markers to research high grade and low grade glioma in a more specialized pattern and may also be beneficial for the disease's treatment module.

In conclusion, utilizing appropriate statistical tools to do meta-analyses on publically accessible datasets provides a cost-effective method for developing innovative hypotheses that may then be verified using more specific molecular approaches. The gene panel of TGF-1, THBS1, SERPINE1, VEGFA, CD44 shown a great degree of similarity in their expression pattern with an accuracy of > 90%, which makes it simple to distinguish gliomas according to their grade. Further, by employing a similar methodology, the results of the current study have discovered and validated a panel of five possible biomarkers predict and identify glioblastoma using liquid biopsy as a component of treatment monitoring. As a result, this multi-gene classifier panel and the pathways that they are linked to may be valuable in predicting outcomes and serving as possible treatment targets in glioblastoma.

Declarations

Ethical Approval:

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the Ethics Committee of Gujarat University (GUJIEC/18/2019)

Consent to Participate:

Informed consent was obtained from all individual participants/legal guardians included in the study.

Consent for Publication: NA

Availability of data and materials:

The gene expression microarray datasets analyzed in this study were downloaded from Gene expression omnibus (GEO) repository (https://www.ncbi.nlm.nih.gov/geo/).

Competing Interests:

The authors have no relevant financial or non-financial interests to disclose

Funding:

This work has been supported by Indian Council of Medical Research (Grant number: 45/2/2019/PHY-BMS. Vinal Upadhyay has got the research support as Senior Research Fellowship by ICMR.

Authors Contribution:

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Vinal Upadhyay and Kinjal Bhadresha. The first draft of the manuscript was written by Vinal Upadhyay and Kinjal Bhadresha and Rakesh Rawal commented on previous versions of the manuscript and helped with further editing. All authors read and approved the final manuscript

Conflict of Interest

The authors declare no conflict of Interest underlying the manuscript.

References

1. Cohen AL, Colman H (2015) Glioma biology and molecular markers. Cancer Treat Res 163:15–30. https://doi.org/10.1007/978-3-319-12048-5_2
2. Carlo DTD, Cagnazzo F, Benedetto N, et al (2019) Multiple high-grade gliomas: epidemiology, management, and outcome. A systematic review and meta-analysis. Neurosurgical Review. https://doi.org/10.1007/s10143-017-0928-7
3. Hygino da Cruz LC, Rodriguez I, Domingues RC, et al (2011) Pseudoprogression and Pseudoresponse: Imaging Challenges in the Assessment of Posttreatment Glioma. AJNR Am J Neuroradiol 32:1978–1985. https://doi.org/10.3174/ajnr.A2397
4. Müller Bark J, Kulasinghe A, Chua B, et al (2020) Circulating biomarkers in patients with glioblastoma. Br J Cancer 122:295–305. https://doi.org/10.1038/s41416-019-0603-6
5. Szopa W, Burley TA, Kramer-Marek G, Kaspera W (2017) Diagnostic and Therapeutic Biomarkers in Glioblastoma: Current Status and Future Perspectives. Biomed Res Int 2017:8013575. https://doi.org/10.1155/2017/8013575
6. Schmidt M, Antweiler S, Urban N, et al (2002) Impact of Genotype and Morphology on the Prognosis of Glioblastoma. Journal of neuropathology and experimental neurology 61:321–8. https://doi.org/10.1093/jnen/61.4.321
7. Sareen H, Garrett C, Lynch D, et al (2020) The Role of Liquid Biopsies in Detecting Molecular Tumor Biomarkers in Brain Cancer Patients. Cancers (Basel) 12:1831. https://doi.org/10.3390/cancers12071831
8. Kalluri R, LeBleu VS (2020) The biology, function, and biomedical applications of exosomes. Science 367:eaau6977. https://doi.org/10.1126/science.aau6977
9. Mashouri L, Yousefi H, Aref AR, et al (2019) Exosomes: composition, biogenesis, and mechanisms in cancer metastasis and drug resistance. Mol Cancer 18:75. https://doi.org/10.1186/s12943-019-0991-5
10. Moher D, Altman DG, Liberati A, Tetzlaff J (2011) PRISMA statement. Epidemiology 22:128; author reply 128. https://doi.org/10.1097/EDE.0b013e3181fe7825

11. Bhadresha K, Upadhyay V, Brahmbhatt J, et al (2022) In vitro model of predicting metastatic ability using tumor derived extracellular vesicles; beyond seed soil hypothesis. Sci Rep 12:20258. https://doi.org/10.1038/s41598-022-24443-8
12. Alexander BM, Cloughesy TF (2017) Adult Glioblastoma. J Clin Oncol 35:2402–2409. https://doi.org/10.1200/JCO.2017.73.0119
13. Nguyen HS, Shabani S, Awad AJ, et al (2018) Molecular Markers of Therapy-Resistant Glioblastoma and Potential Strategy to Combat Resistance. Int J Mol Sci 19:1765. https://doi.org/10.3390/ijms19061765
14. Wang Z, Bao Z, Yan W, et al (2013) Isocitrate dehydrogenase 1 (IDH1) mutation-specific microRNA signature predicts favorable prognosis in glioblastoma patients with IDH1 wild type. J Exp Clin Cancer Res 32:59. https://doi.org/10.1186/1756-9966-32-59
15. Srinivasan S, Patric IRP, Somasundaram K (2011) A Ten-microRNA Expression Signature Predicts Survival in Glioblastoma. PLOS ONE 6:e17438. https://doi.org/10.1371/journal.pone.0017438
16. Wang W, Zhang L, Wang Z, et al (2016) A three-gene signature for prognosis in patients with MGMT promoter-methylated glioblastoma. Oncotarget 7:69991–69999. https://doi.org/10.18632/oncotarget.11726
17. Alexandrow MG, Moses HL (1995) Transforming growth factor β and cell cycle regulation. Cancer research 55:1452–1457
18. Wesolowska A, Kwiatkowska A, Slomnicki L, et al (2008) Microglia-derived TGF-beta as an important regulator of glioblastoma invasion--an inhibition of TGF-beta-dependent effects by shRNA against human TGF-beta type II receptor. Oncogene 27:918–930. https://doi.org/10.1038/sj.onc.1210683
19. Noboru Ueki, Masaru Nakazato, Toshihisa Ohkawa, et al (1992) Excessive production of transforming growth-factor β1 can play an important role in the development of tumorigenesis by its action for angiogenesis: validity of neutralizing antibodies to block tumor growth. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1137:189–196. https://doi.org/10.1016/0167-4889(92)90201-L
20. Platten M, Wick W, Weller M (2001) Malignant glioma biology: Role for TGF-β in growth, motility, angiogenesis, and immune escape. Microscopy Research and Technique 52:401–410. https://doi.org/10.1002/1097-0029(20010215)52:4<401::AID-JEMT1025>3.0.CO;2-C
21. Seker F, Cingoz A, Sur-Erdem İ, et al (2019) Identification of SERPINE1 as a Regulator of Glioblastoma Cell Dispersal with Transcriptome Profiling. Cancers 11:1651. https://doi.org/10.3390/cancers11111651
22. Zhang S-D, Leung KL, McCrudden CM, Kwok HF (2015) The Prognostic Significance of Combining VEGFA, FLT1 and KDR mRNA Expressions in Brain Tumors. Journal of Cancer 6:812–818. https://doi.org/10.7150/jca.11975
23. Hlobilkova A, Ehrmann J, Knizetova P, et al (2009) Analysis of VEGF, Flt-1, Flk-1, nestin and MMP-9 in relation to astrocytoma pathogenesis and progression. Neoplasma 56:284–290. https://doi.org/10.4149/neo_2009_04_284
24. Ma Y, Qu B, Xia X, et al (2015) Glioma-Derived Thrombospondin-1 Modulates Cd14+ Cell Tolerogenic Properties. Cancer Investigation 33:152–157. https://doi.org/10.3109/07357907.2015.1010089
25. de Fraipont F, Keramidas M, El Atifi M, et al (2004) Expression of the thrombospondin 1 fragment 167–569 in C6 glioma cells stimulates tumorigenicity despite reduced neovascularization. Oncogene 23:3642–3649. https://doi.org/10.1038/sj.onc.1207438
26. Tsutsui T, Kawahara H, Kimura R, et al (2020) Glioma-derived extracellular vesicles promote tumor progression by conveying WT1. Carcinogenesis 41:1238–1245. https://doi.org/10.1093/carcin/bgaa052
27. Orian-Rousseau V (2010) CD44, a therapeutic target for metastasising tumours. Eur J Cancer 46:1271–1277. https://doi.org/10.1016/j.ejca.2010.02.024

Figures

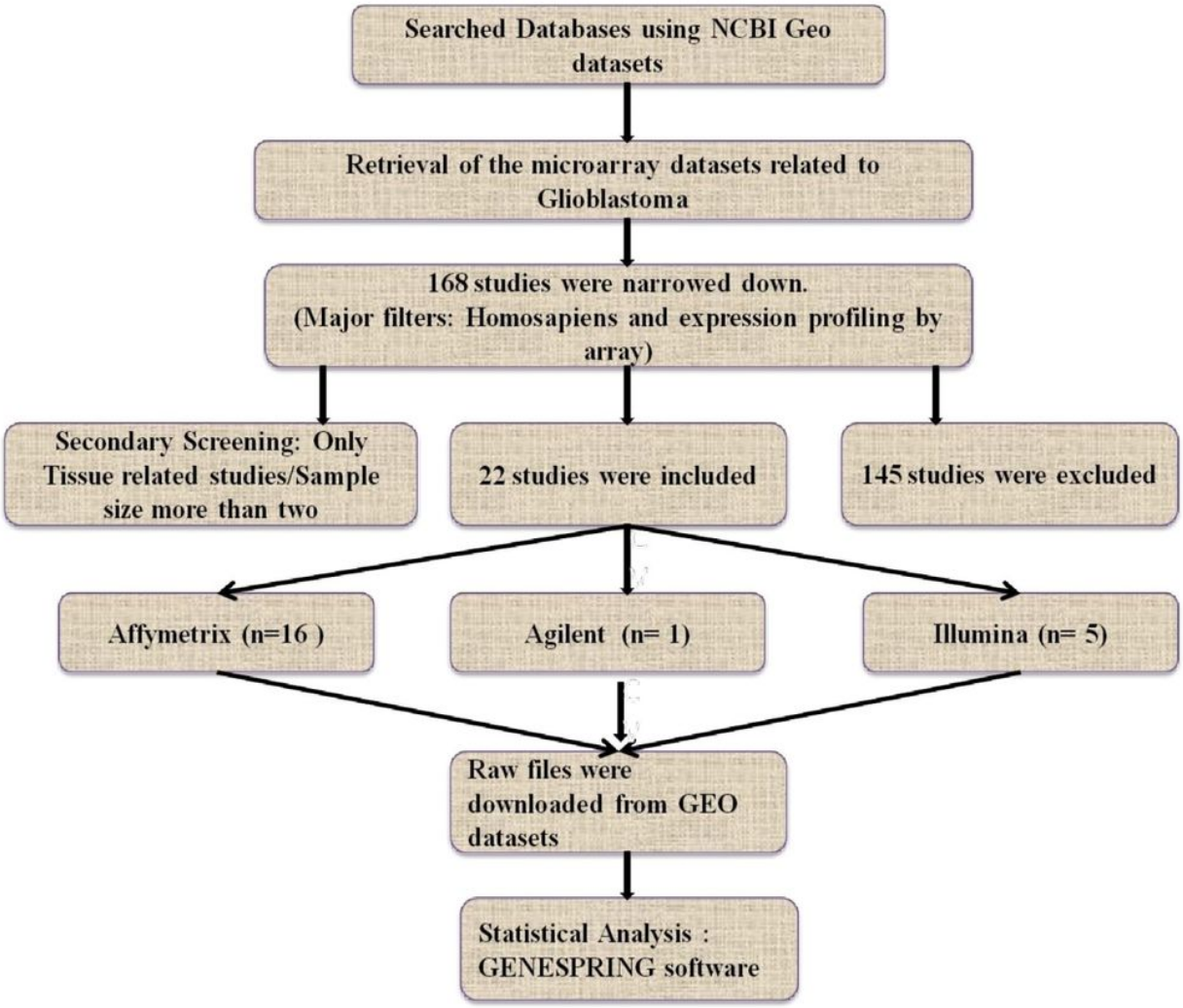

Figure 1

Representative schematic diagram of the workflow integrated in the study

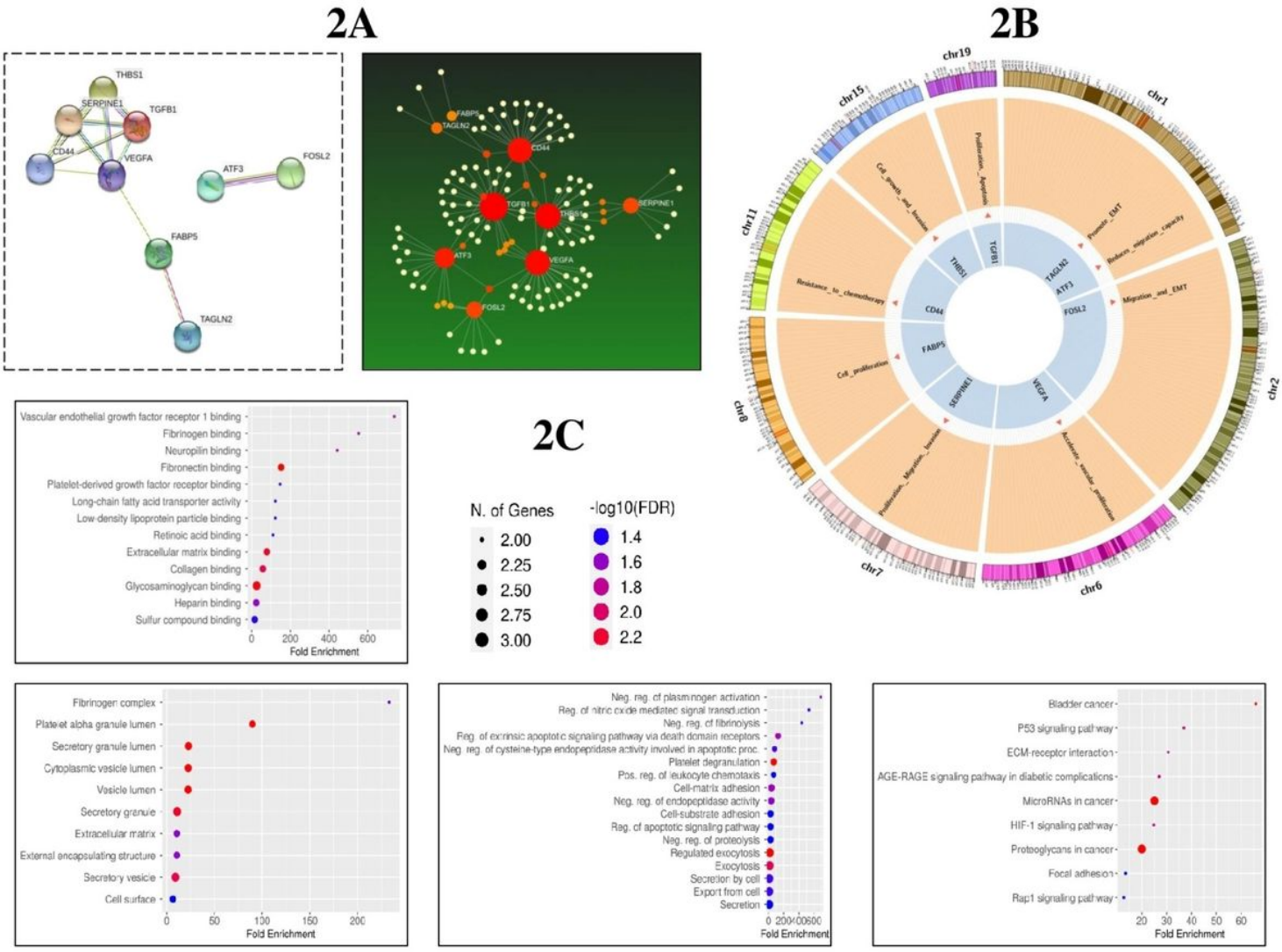

Figure 2

Bioinformatics based data analysis **(2A)** String output showing interaction of 9 genes that are specific for Glioblastoma. Visual analytics of hub genes by Network analyst **(2B)** CIRCOS Plot indication Gene Location and Function **(2C)** Functional significance of Gene Set by Bioinformatic tools (a) Shows Biological Processes (b) Cellular Components (c) Molecular Functions (d) Site of Expression related to the narrowed genes

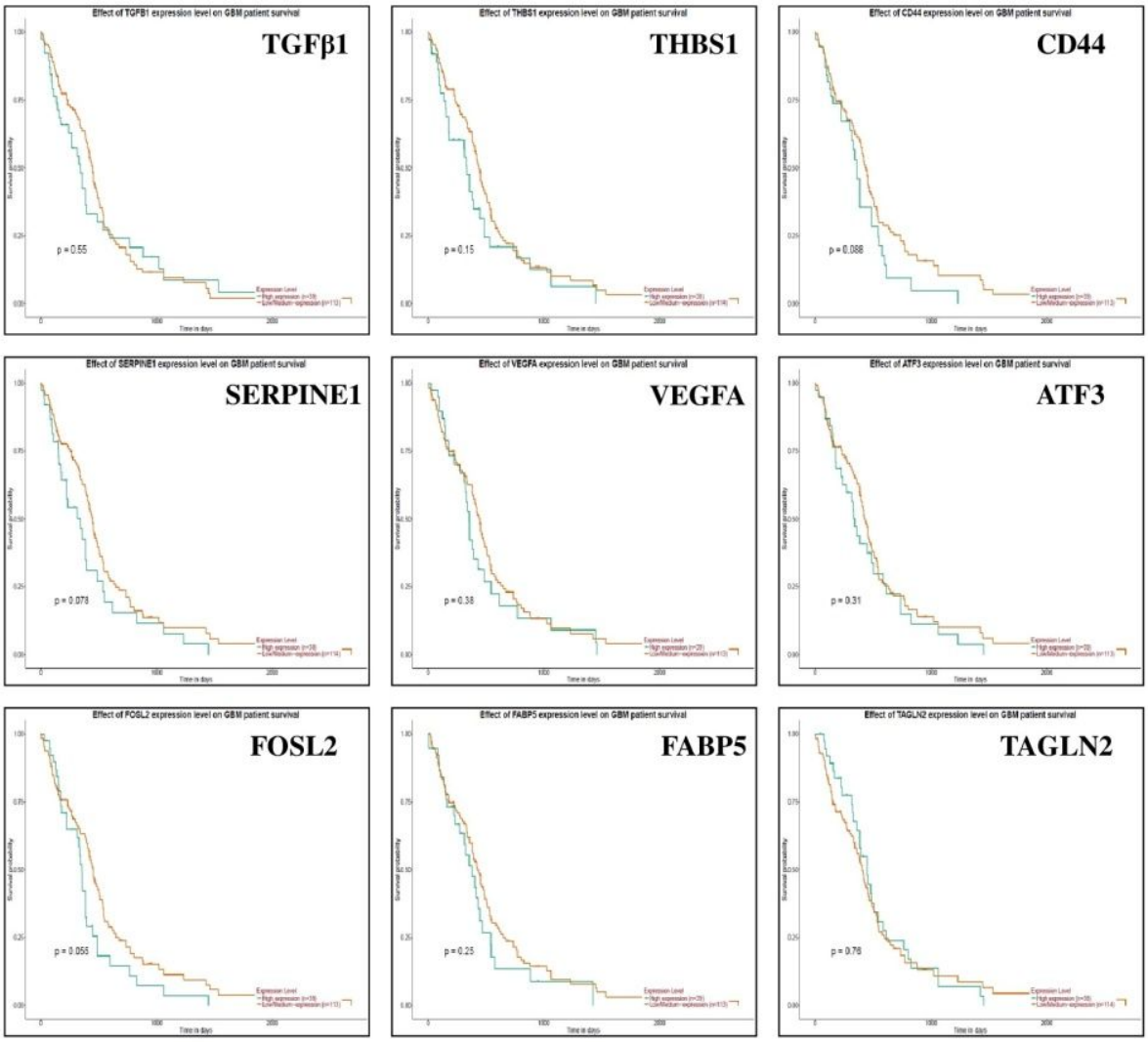

Figure 3

Software based Kaplan Meir Survival analysis gene wise. The respective graphs shows the survival based curves depending on the expression of gene

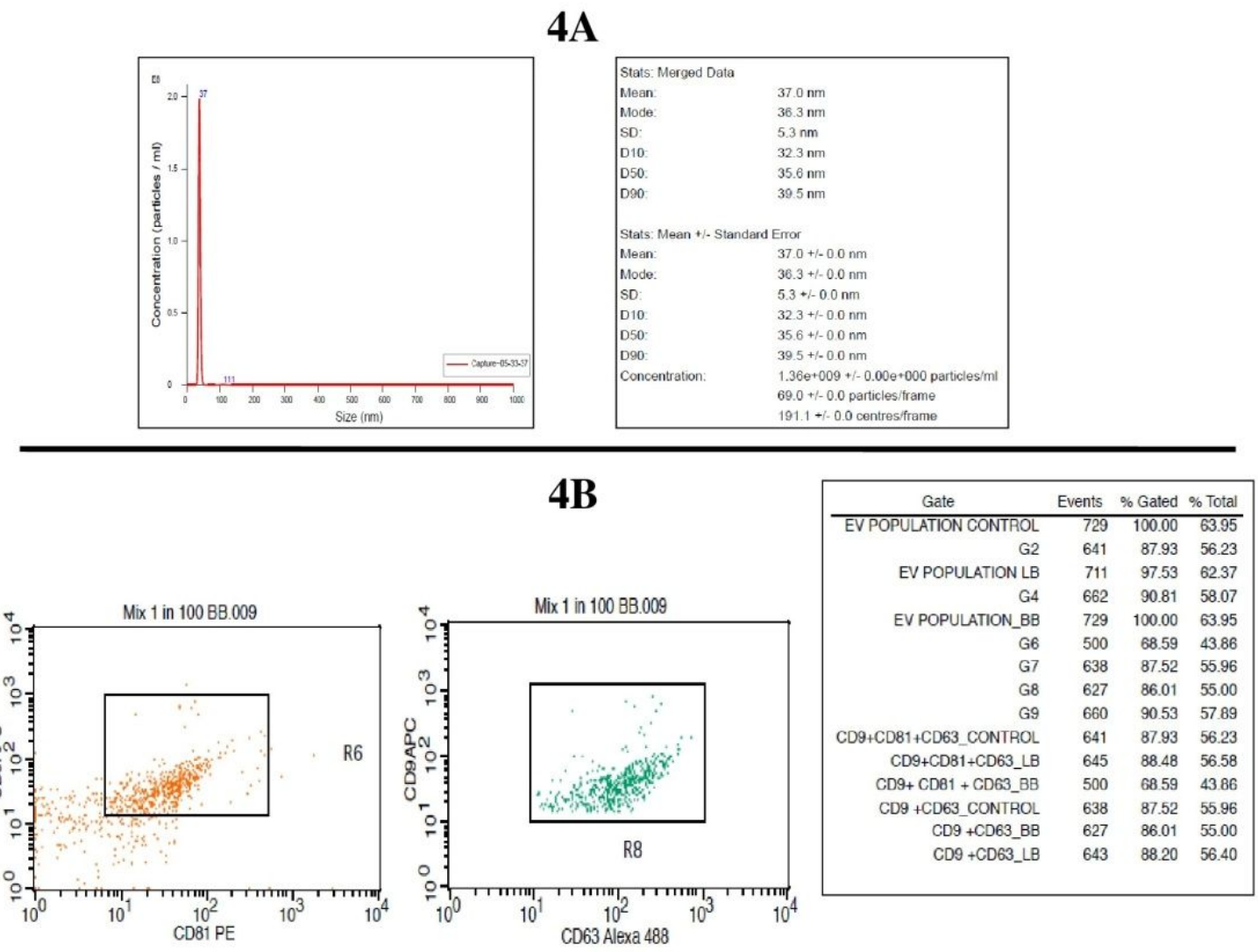

Figure 4

Characterization of exosomes. **(4A)**Size and concentration of exosomes by nanosight. **(4B)** Exosomes markers (CD63, and CD81) were analyzed using Flow-cytometer. The data demonstrated that extracts were enriched with exosomal marker protein CD81 and CD63.

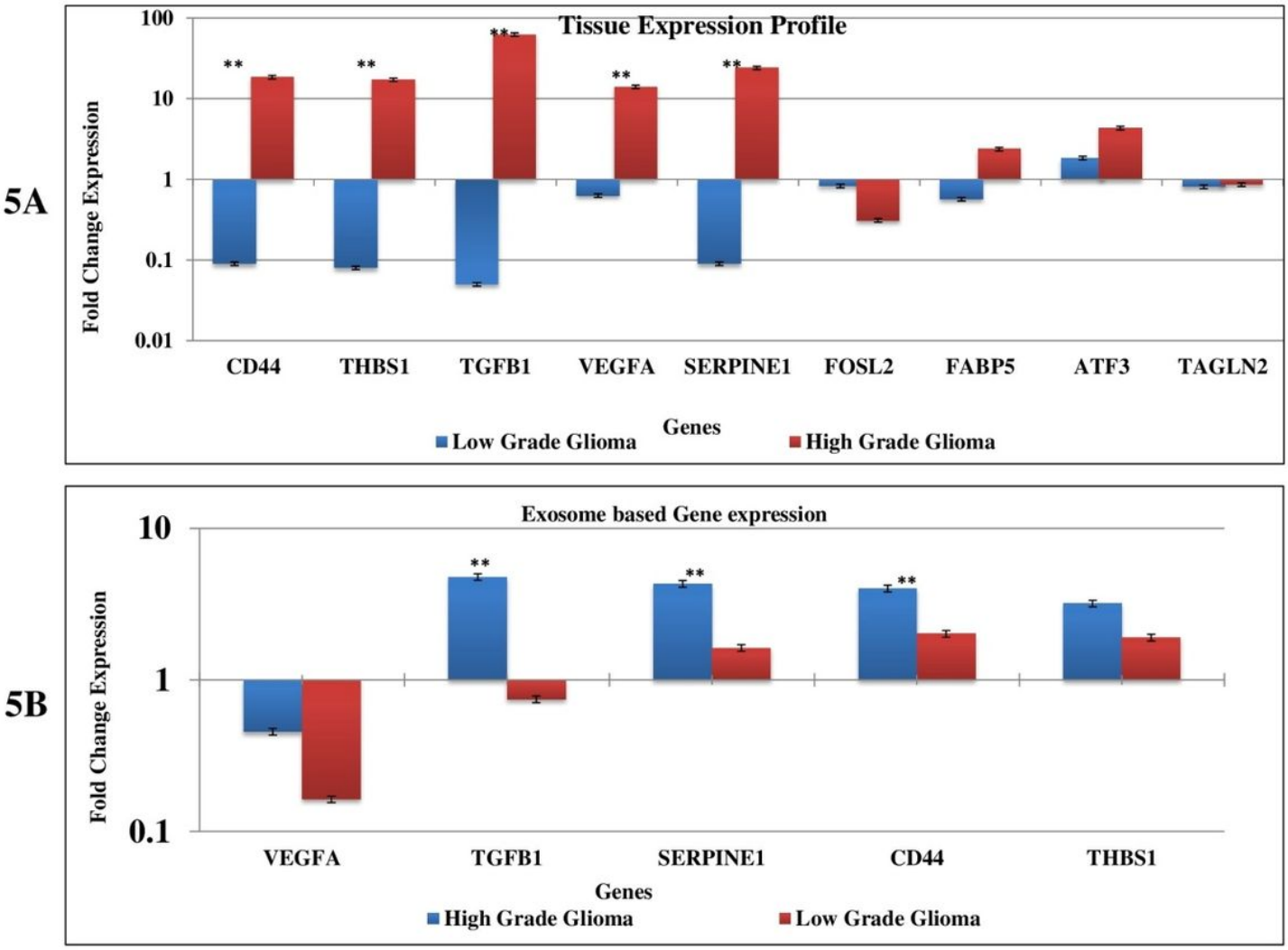

Figure 5

Gene expression analysis. (5A)Tissues based expression of genes in High grade (n=30) and Low grade Gliomas(n=30). (5B) Exosomal expression of the genes in High grade (n=15) and Low grade (n=15) gliomas analysed with qRT-PCR.

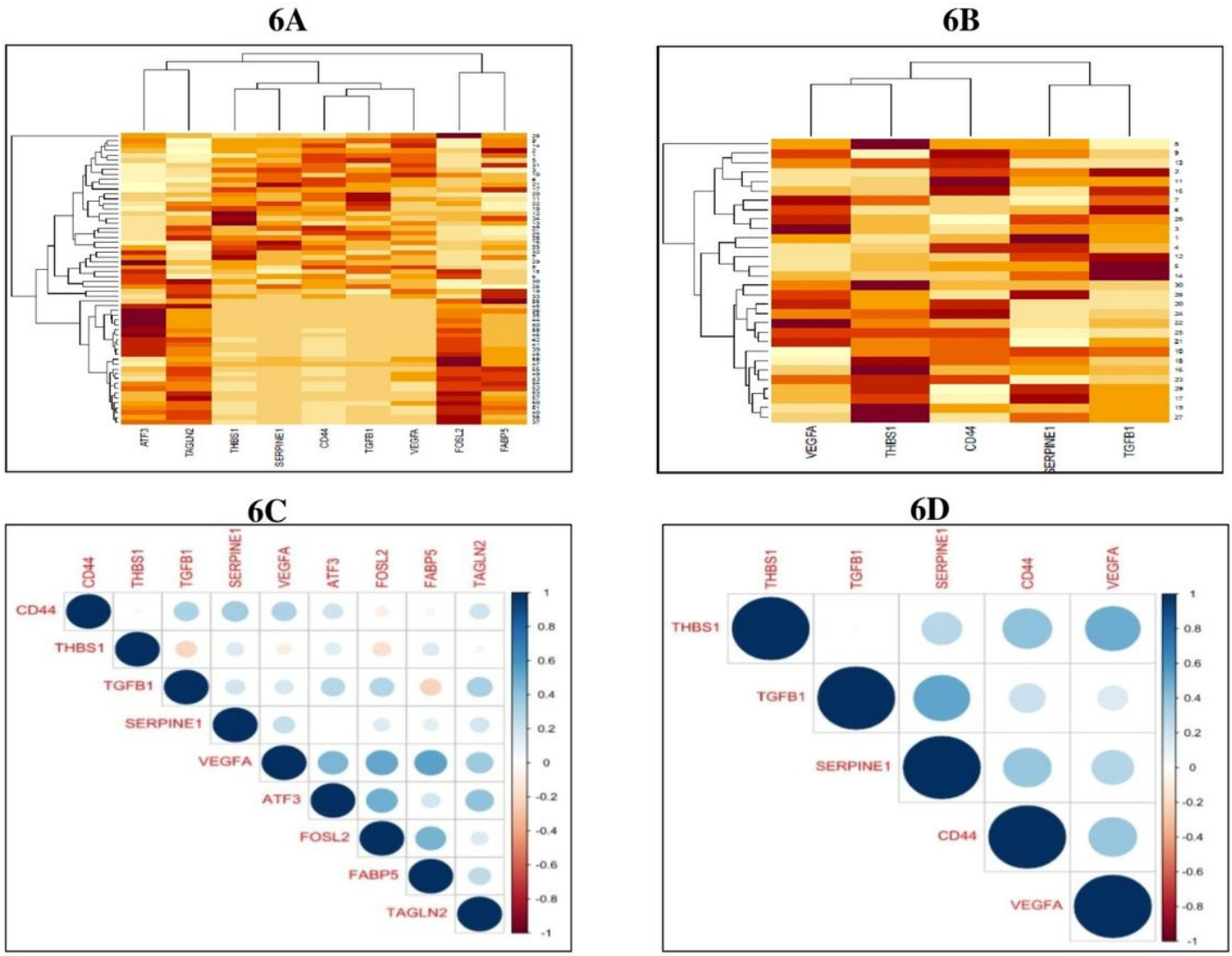

Figure 6

Unsupervised hierarchical clustering of High grade gliomas and low grade gliomas **(6A)** Tissue expression and **(6B)** Exosomal expression. Clustering was based on 9 differentially expressed genes at a false discovery ratio level of 0.05. Tumor identification at the top of the figure and each column represents gene expression of a single tumor. The colored bar specifies the variation in gene expression in target samples as compared to reference cells i.e., red, more expressed and cream, less expressed in target samples. Further, the black lines of the dendrogram stand for the support for each clustering. The metric performed was Euclidean distance, with complete linkage for distance between clusters. **(6C)**Correlation matrices of different cohorts by Pearson and Spearmen analysis. In a color coded scale from blue (positive correlation) to red (inverse correlation) are represented the correlation coefficients for each pair of targets included in the analysis for glioblastoma tissues and **(6C)** exosomes

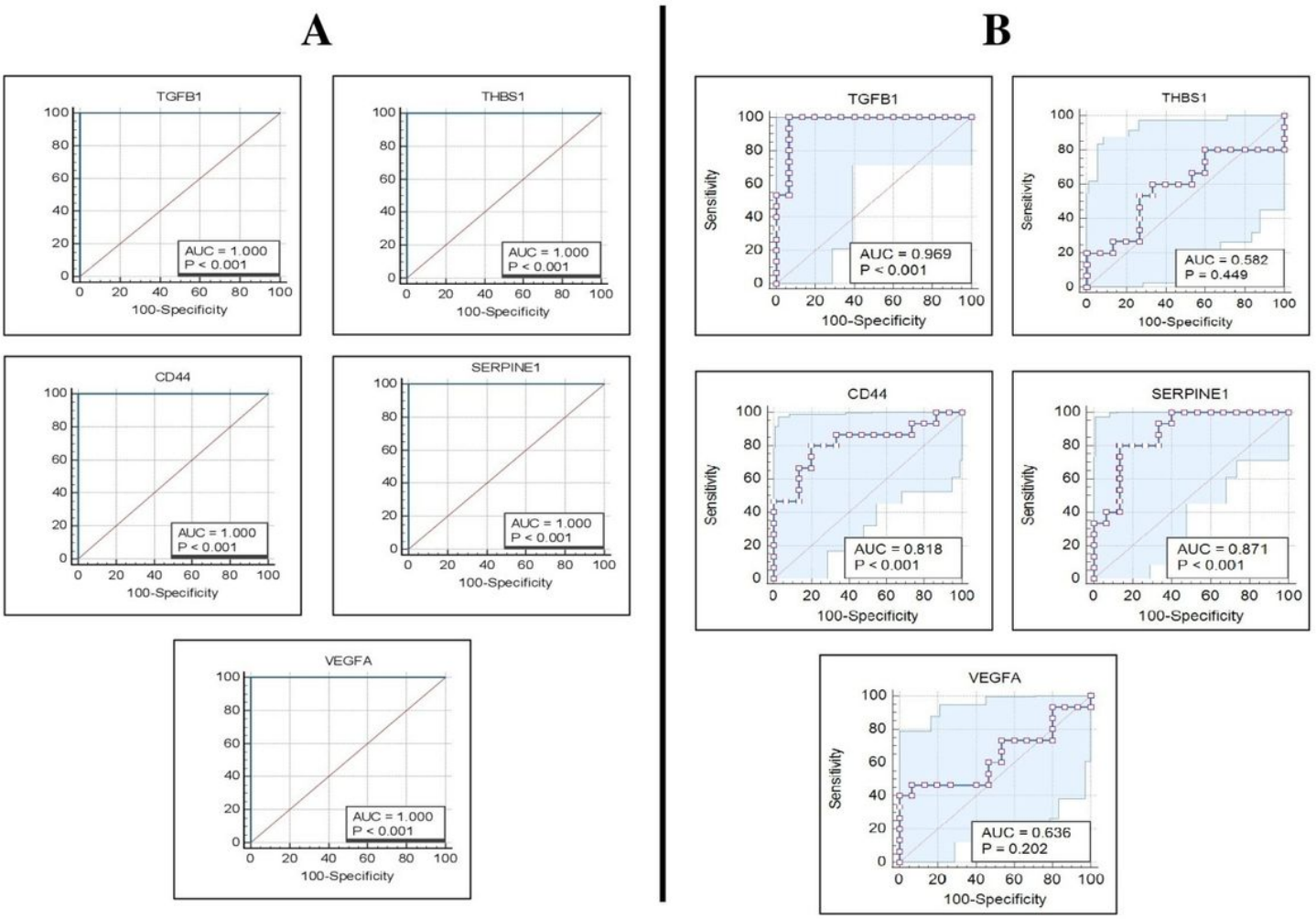

Figure 7

Receiver Operating Characteristic (ROC) Curve analysis representing marker specificity and sensitivity for 5 genes in tissues **(7A)** and exosomes **(7B)**

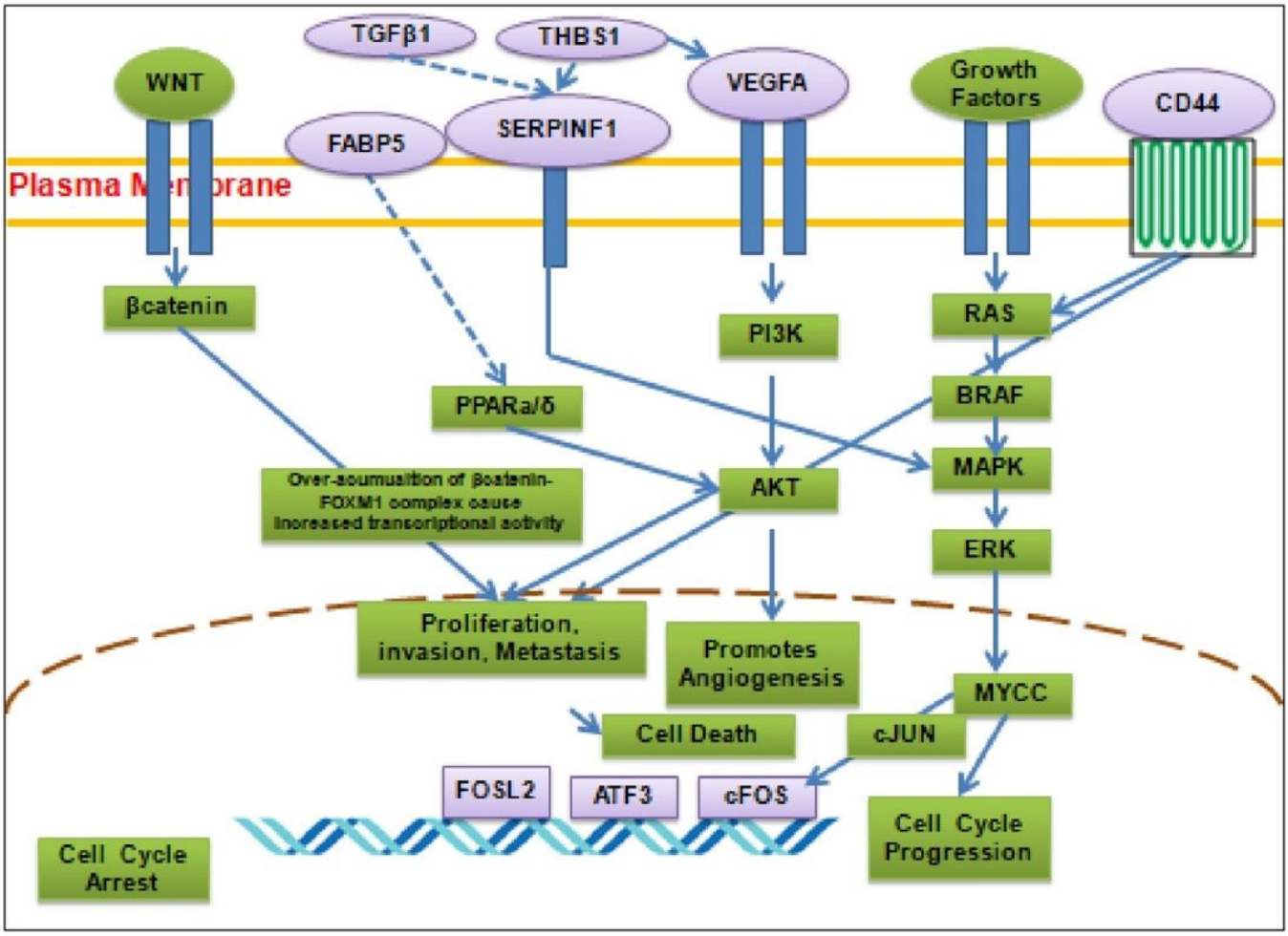

Figure 8

Molecular mechanisms connected with disease progression.