Tumor Heterogeneity and Molecular Characteristics of Glioblastoma Revealed by Single-Cell RNA-Seq Data Analysis

Dhanusha Yesudhas [1], S. Akila Parvathy Dharshini [1], Y-h. Taguchi [2] and M. Michael Gromiha [1,*]

1   Department of Biotechnology, Bhupat and Jyoti Mehta School of Biosciences,
    Indian Institute of Technology Madras, Chennai 600036, India; dhanusha2504@gmail.com (D.Y.);
    akilabioinfo@gmail.com (S.A.P.D.)
2   Department of Physics, Chuo University, Bunkyo-ku, Tokyo 112-8551, Japan; tag@granular.com
*   Correspondence: gromiha@iitm.ac.in

Abstract: Glioblastoma multiforme (GBM) is the most common infiltrating lethal tumor of the brain. Tumor heterogeneity and the precise characterization of GBM remain challenging, and the disease-specific and effective biomarkers are not available at present. To understand GBM heterogeneity and the disease prognosis mechanism, we carried out a single-cell transcriptome data analysis of 3389 cells from four primary IDH-WT (isocitrate dehydrogenase wild type) glioblastoma patients and compared the characteristic features of the tumor and periphery cells. We observed that the marker gene expression profiles of different cell types and the copy number variations (CNVs) are heterogeneous in the GBM samples. Further, we have identified 94 differentially expressed genes (DEGs) between tumor and periphery cells. We constructed a tissue-specific co-expression network and protein–protein interaction network for the DEGs and identified several hub genes, including *CX3CR1, GAPDH, FN1, PDGFRA, HTRA1, ANXA2 THBS1, GFAP, PTN, TNC*, and *VIM*. The DEGs were significantly enriched with proliferation and migration pathways related to glioblastoma. Additionally, we were able to identify the differentiation state of microglia and changes in the transcriptome in the presence of glioblastoma that might support tumor growth. This study provides insights into GBM heterogeneity and suggests novel potential disease-specific biomarkers which could help to identify the therapeutic targets in GBM.

Keywords: glioblastoma; transcriptome analysis; tumor heterogeneity; biomarkers; network

1. Introduction

Glioblastoma multiforme (GBM) is a highly heterogeneous tumor, with diverse co-existing cell types that include tumor cells, endothelial cells, fibroblasts, and different cell types from the immune system [1,2]. Recently, it has been shown that the GBM subtypes can co-exist in different regions and cells within the same tumor [3]. Variability is found across tumor tissues, at different stages, and in different gender and age proportions. The invasive and metastatic ability of the tumor cells contributes to its high heterogeneity. The cells from the same tumor tissue can also have different mutations, which results in different phenotypic and epigenetic changes [4]. For example, the genes *SETD2, PTEN*, and *KDM5C* encountered multiple distinct and spatially separated inactivating mutations within a single tumor, and caused phenotypic evolution [5]. Similarly, Liu et al. [6] reported that the hypermethylation phenotype in the IDH1 mutant is involved in silencing of the α-KG-dependent DNA-modifying enzyme (Tet methylcytosine dioxygenase 2 (TET2)) and eventually increased the tumor formation. Tumor heterogeneity in patients and the characterization of its invasive nature remains a significant challenge for research and targeted therapeutic approaches [3,7]

Glioblastic cells affect stromal cells as well as the central nervous system (CNS) immune cells, including microglia, astrocytes, oligodendrocytes, neurons, and monocytes [8].

Maas et al. [9] reported that glioblastoma tumor cells communicate with microglial cells by releasing extracellular vesicles and these cells hijack the immune system. Similarly, cancer cells use tunneling nanotubes (TNT) as an efficient cell-to-cell communication system to adapt the microenvironment, which is also responsible for the invasive nature of the GBM tumors [10]. Neftel et al. [11] classified glioblastomas based on the intra and inter-tumoral cell state and genetic diversity of glioblastomas by comparing pediatric and adult glioblastomas. Macrophages, monocytes, and microglial cells are reported to be important in glioblastoma tumor growth, and glioblastoma invasion has been reduced with the depletion of these cells [12]. However, the exact pathways involved in tumor supportive process have not been characterized.

Although traditional bulk RNA-seq approaches have helped to identify key genes and pathways that drive GBM cells [13–16], they provide limited insights into the tumor heterogeneity and molecular mechanisms underlying GBM invasion. Darmanis et al. [17] reported the nature of infiltrating GBM cells and characterized neoplastic and non-neoplastic cells. However, other cell types, such as neural, glial, immune, and vascular cells, and the dynamic transition state of GBM cells remain unexplored. Hence, a better understanding of GBM heterogeneity between tumor and periphery cells and the molecular mechanism behind the transition of distinct cell types are necessary for further investigations.

In this work, we carried out a single-cell transcriptome data analysis of 3389 cells from four primary IDH-WT glioblastoma patients and provided a detailed analysis of the heterogeneity between the tumor and periphery cells. Further, we have carried out transcriptome studies based on the hg38 genome and with recent scRNA-seq analysis pipelines for preprocessing and downstream analysis. From copy number variation (CNV) analysis, distinct amplifications or deletions in different chromosomes have been observed for the patient samples. The novel and potent cell type-specific differentially expressed genes (DEGs) screened from different approaches, their functional enrichment and the co-expression networks, which have not previously been reported, strengthen our analysis regarding the disease-specific target genes. The potential DEGs are enriched with glial cell differentiation and mononuclear cell migration. Our study also explains the interplay between immune microglial cells and neoplastic cells and reveals the transition state of microglial cells.

2. Method

2.1. Data Collection and Quality Check

Darmanis et al. [17] reported high-depth single-cell RNA sequencing for a cohort of four primary GBM patients (IDH1-negative, grade IV GBMs confirmed via pathological examination). Two separate tissue samples were collected from each patient (one originating from the tumor core and another from the peritumoral space immediately adjacent to the tumor core, also termed the periphery). The details of the dataset are provided in Supplementary Table S1. Out of 3589 cells, 2343 cells are from tumor cores and 1246 cells are from the peripheral region [17].

The scRNA-seq raw reads were analyzed using FASTQC for a quality check and the short reads (less than 30) were discarded. The adapters were removed using cutadapt [18]. The processed reads were mapped against the hg38 human transcriptome using quasi-mapper "SALMON" (version 1.1.1) [19]. The transcripts were quantified and obtained as transcripts per million (TPM) values. The workflow of the present work is illustrated in Figure 1.
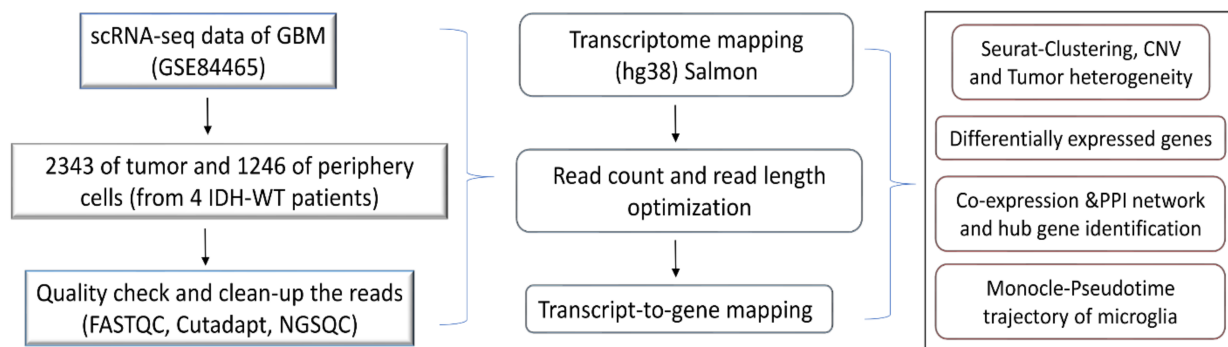
Figure 1. Schematic diagram for the study design.

2.2. Dataset Preprocessing by Seurat

The TPM-based count matrix obtained from SALMON was taken as an input for the cell preprocessing, which was performed using the Seurat package [20]. We excluded the low-quality cells based on the following quality measures: (i) the genes showed expression in at least three cells and (ii) the minimum length of the RNA read should be 200 bps (Supplementary Figure S1). Mitochondrial genes were removed from the dataset as they represent contamination in the sequencing technique.

2.3. Dimensionality Reduction and Cluster Identification

The dimensionality of the dataset was determined by principal component analysis (PCA), and the number of PCs which captured the highest variance was obtained from the elbow plot (Supplementary Figure S1). Furthermore, the dataset was reduced to two dimensions with UMAP (Uniform Manifold Approximation and Projection). Clustering the groups of similar cells was performed with the FindClusters function based on the KNN algorithm. This is a graph-based clustering algorithm with edges drawn between cells with similar gene expression patterns. The "resolution" argument will set the "granularity" of the downstream clustering, which will be needed to be optimized for the experiments. In addition, we computed the average expression and dispersion of each gene using the module "FindVariableGenes" in Seurat and selected the top 1000 over-dispersed genes (outliers) for our study.

2.4. Determination of Copy Number Variations (CNVs)

InferCNV was used to explore the tumor single-cell RNA-seq data to identify large-scale chromosomal CNVs [21]. All genes were ranked by their chromosomal location, and the copy number of each gene was calculated as the sliding average of log2-transformed TPM values with a window of 100 flanking genes within each chromosome, which was then centered across all cells. Furthermore, we performed hierarchical clustering and removed the nontumor cells, which showed a few CNVs, similar to the normal cells. A hidden Markov model was used to predict the CNV states, and the gene location data were obtained from the Biomart database.

2.5. Differential Gene Expression and Functional Annotation

Tximport was used to import the transcript level estimates (TPM values) from SALMON and summarize them as gene abundance. Subsequently, DESeq2 software [22] was used to identify the genes, which were differentially expressed in tumor core and periphery cells. The genes with log2FC > 1 and log2FC < 1 with adj. *p*-value < 0.05 were considered as significant differentially expressed genes. Benjamini and Hochberg's method was used to get the adj. *p*-value.

"FindMarkers" and "FindAllMarkers" from Seurat were used to identify the DEGs, and a gene expression level with log2FC > 1 and an adj. *p*-value < 0.05 were used as filtering criteria. The commonly observed DEGs from DESeq2 and Seurat were filtered for further

analysis. The ontology and functional annotations of DEGs were analyzed using ClueGO from Cytoscape.

2.6. Network Analysis

The significantly differentially expressed genes were analyzed using the STRING and Humanbase databases in order to analyze the interacting partners [23]. The protein–protein interaction network was generated from STRING, and the tissue (human cortex) specific co-expression network was generated from Humanbase. Cytoscape plugin MCODE was used to find the cluster and hub genes.

2.7. Monocle Pseudotime Trajectory Reconstruction and Analysis

Single-cell pseudotime trajectories were constructed with MONOCLE 2. This method uses the reverse graph embedding (a machine learning technique) method to reduce the given high-dimension expression profile to low-dimensional space [24]. Single-cell data points were projected onto this low-dimensional space and ordered into a trajectory with branch points. This method tests whether differences in gene expression are associated with particular branching events on the trajectory.

3. Results

3.1. Cell Clusters and Tumor Heterogeneity in GBM

A schematic diagram of the study design and the principal findings are shown in Figure 1. A total of 2343 tumor cells and 1246 periphery cells from four IDH-WT patients were adapted for this study (Supplementary Table S1). After stringent quality control and normalization, we analyzed these 3389 cells with 42,970 genes. In order to identify and characterize GBM cellular heterogeneity, we clustered the similar cells using UMAP (see Materials and Method). UMAP computed a total of 16 clusters for the 3389 human GBM cells (Figure 2). PCA variance analysis captured the highly variable genes across the PCs (Supplementary Figure S1b) and the top 1000 cell-to-cell variable genes are reported in Supplementary File S1. The positive and negative correlation of these variable genes along with their PCs are represented in Supplementary Figure S1c.

3.1.1. Cell Type Identifications of the Clusters

Based on metadata and marker gene identification, the UMAP identified clusters were grouped into seven major cell types (astrocyte, oligodendrocyte, vascular, OPC, neuron, immune, and neoplastic), as shown in Figure 2b. From each cluster, genes showing significant expression changes (log2FC > 1 and adj. *p*-value < 0.05) were filtered out and annotated with ScCATCH and PanglaoDB databases and subsequently labeled as potent markers for the cell types.

Clustering based on cluster ID: As seen from Figure 2b, the immune cluster is the largest cluster comprising 0, 1, 4, 5, 8, and 9 subclusters. Cluster 0 contains 540 cells and is annotated as microglial cells; clusters 1, 4, 5, 8, and 9, containing 1290 cells, are annotated as Schwann cells (a type of glial cell). The second largest cluster is the neoplastic cluster, comprising 3, 6, 7, 10, 13, and 14 subclusters. Clusters 3, 6, 7, 10, and 14, containing 945 cells, are annotated as microglial cells, and subcluster 13 contains 49 cells and is annotated as macrophages. Clusters 2, 11, 12, and 15 are annotated as OPCs, oligodendrocytes, astrocytes, and vascular cells, respectively. The number of cells in each cluster is represented in Supplementary Figure S1e.

Clustering based on patient ID: Figure 2c shows the clusters which are grouped based on patient ID. The patient BT_S2 and BT_S4 samples are highly crowded with immune and neoplastic cells. However, the patient ID BT_S1 and BT_S6 samples are mainly occupied with neoplastic cells (Figure 2c).

Clustering based on tissue: The clusters are grouped based on tissue characteristics and the data are shown in Figure 2d. The neoplastic cells are mainly from tumor core cells. However, the immune clusters contain both tumor and periphery core cells.
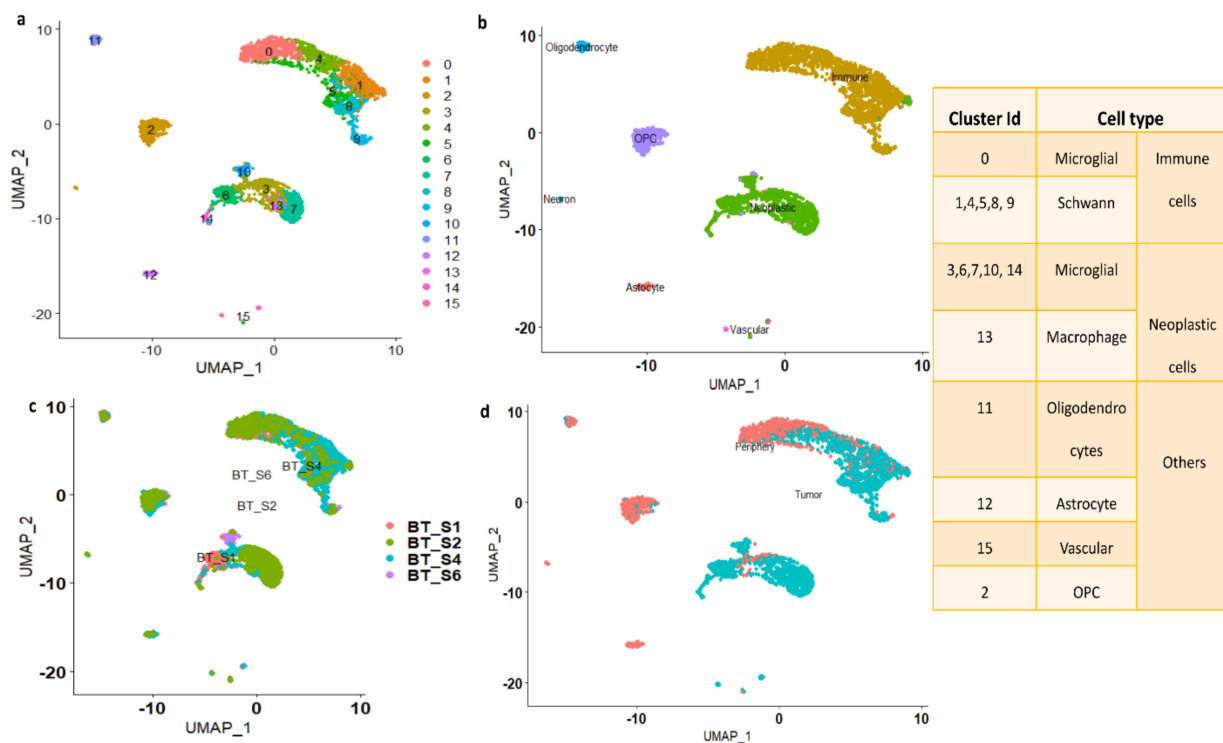
Figure 2. (**a**) UMAP plot showing the distribution of cells with cluster ID. (**b**) UMAP clusters are labeled based on major cell types. (**c**) Clustering of the cells based on patient-ID. (**d**) Clustering of the cells based on tissue. The table displays the cluster ID matches with cell type.

3.1.2. Marker Genes in Each Cluster

Genes that are significantly expressed in each cell type (astrocytes, immune cells, neurons, neoplastic cells, OPCs, oligodendrocytes, and vascular cells) have been reported in Supplementary File S1. Furthermore, we have reported the genes, which are differentially expressed between tumor and periphery cells and unique to each cluster, shown as a dot plot in Figure 3. The cytokine genes *CCL4, CCL4L2,* and *CCL3L1* are down-regulated in tumors and are up-regulated in periphery cells, and these genes are found to be markers for the immune cell cluster. The genes *TNR, OLIG1,* and *PDGFRA* are specific to the OPC cluster and are differentially expressed in tumor and periphery cells. These genes are up-regulated in tumors and down-regulated in periphery cells. It has been reported that co-expression of *EGFR* and *PDGFRA* is a driver event early in GBM [25,26]. The heat-shock genes *HSPA1A* and *HSPA1B* show down-regulation in tumor and up-regulation in the periphery, and act as marker genes for neuron clusters. *TUBA1A, DBI,* and *TUBB* are up-regulated in tumor cells and down-regulated in periphery cells, and these are marker genes for the neoplastic cluster. The tubulin proteins *TUBA1A* and *TUBB* and their heterogeneity has been associated with GBM, whereas *DBI* maintains high proliferation rates, promoting tumor growth [27]. The oligodendrocyte marker genes *OPALIN, MAG,* and *KLK6* are up-regulated in tumor cells and down-regulated in the periphery. The highly variable genes *MAG* (Supplementary Figure S1b) and *OPALIN* encode glycoproteins involved in myelinating oligodendrocytes. The genes *COL3A1, ISLR,* and *IFITM1* are specific for vascular cells and show up-regulation in the tumor and down-regulation in the periphery. The genes *COL3A1* and *IFITM1* are involved in cell migration, and *ISLR* is the contemporary gene identified from this current study.
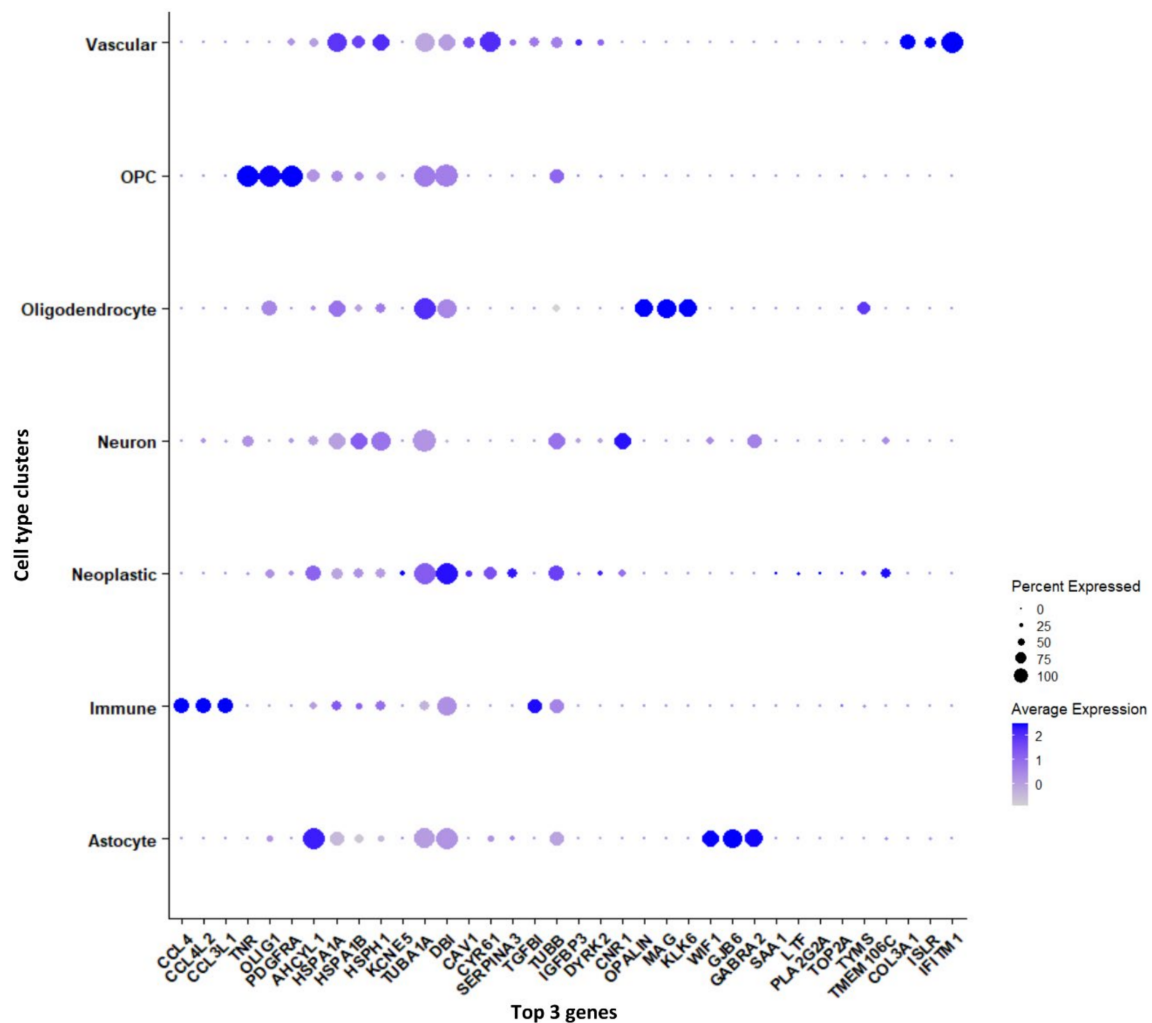
**Figure 3.** Top three marker genes and their expression unique to each cluster represented in a dot plot.

3.2. Differentially Expressed Genes (DEGs) between Tumor and Periphery Cells

Considering the DEGs from each cluster, we were interested to infer the differentially expressed genes in the tumor and periphery cell populations, so we used DESeq2 and Seurat to analyze the differentially expressed genes between the tumor and periphery cells. We compared the results from DESeq2 and Seurat and the consensus results are discussed below.

Using DESeq2, the DEGs from tumor and periphery cells were obtained. With 3389 cells, a nonzero read count of 21,250 was obtained. We have applied the filters adj. *p*-value < 0.05, log2FC > 1 for up- and log2FC < 1 for down-regulated genes. We obtained 2.1% up-regulated (454 genes) and 3.2% down-regulated genes (680 genes). The outliers and the low counts had been removed. Figure 4 shows the clear separation of up- and down-regulated genes, with significant genes marked in red.

DEGs from Seurat were filtered with log2FC > 1 and adj. *p*-value < 0.05 and 100 genes (33 up-regulated and 67 down-regulated) were obtained. The complete list of differentially expressed (up- and down-regulated) genes from DESeq2 and Seurat are presented in Supplementary Table S2. We considered the common DEGs from DESeq2 and Seurat, resulting in 94 significant differentially expressed genes. Among them, 22 DEGs are novel in the GBMs found in this study. Furthermore, *DHRS9, IPCEF1, TNR, MEGF11, EDIL3, PDZD2, ATP1A2, PDGFRA*, and *MEG3* are novel down-regulated genes that are involved in focal adhesion, cell adhesion to the extracellular matrix (ECM), thereby allowing the cells to crawl during migration. Similarly, the genes *CHI3L1, FN1, IGFBP2, TNC, FCGBP,*

CYR61, F13A1, ANXA2, AC006064.4, and *MIF-AS1* are novel up-regulated genes which are involved in ECM–receptor interactions and collagen cross-linking. Integrins interact with ECM components, such as collagen, brevican, tenascin-C, fibronectin, and thrombospondin, which leads to the adhesion and migration of glioma cells [28].
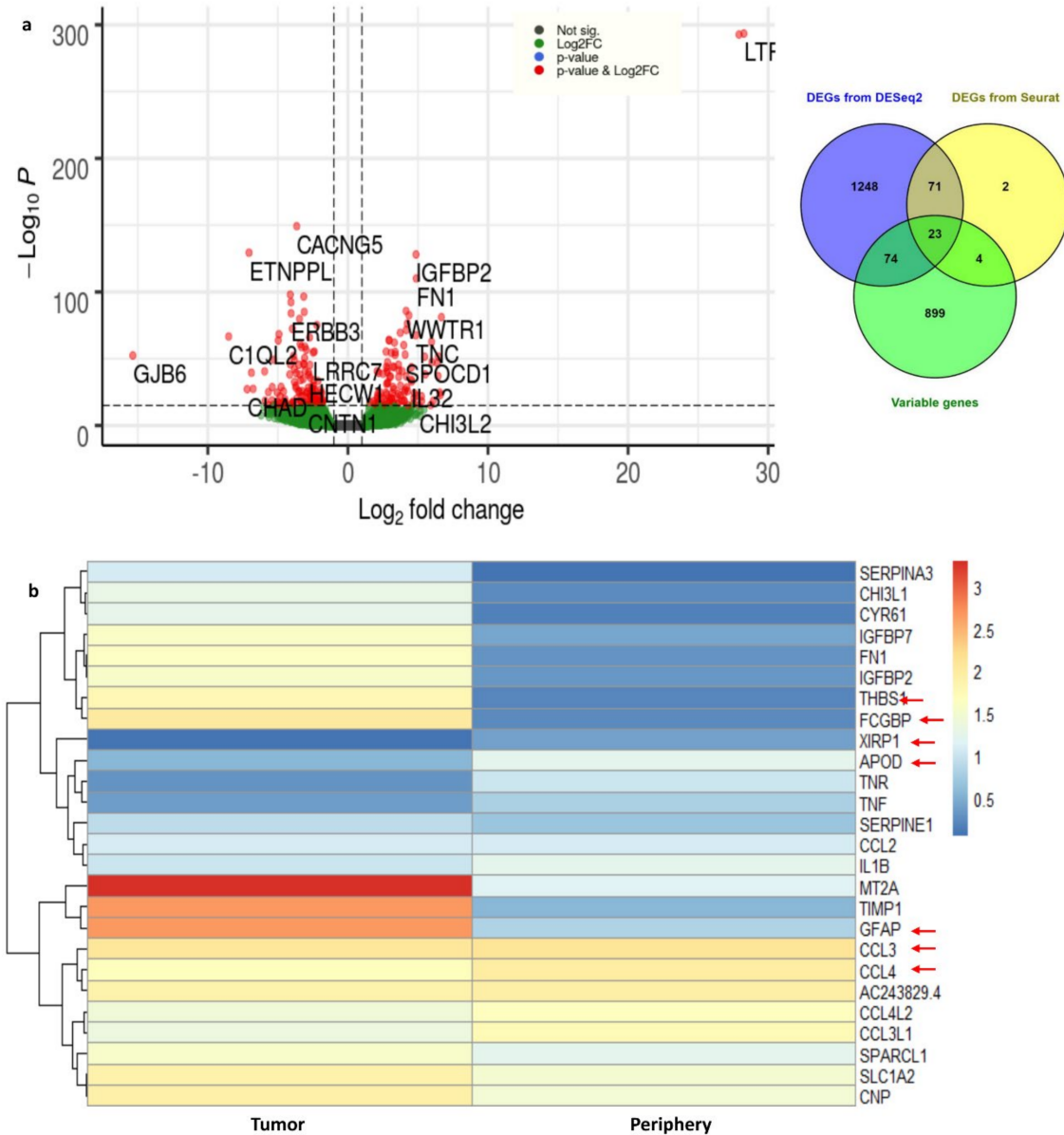
Figure 4. (**a**) Volcano plot for the differentially expressed genes from DESeq2. The up- and down-regulated genes with the cut-off of log2FC > 1 and with adj. *p*-value < 0.05 are shown in the figure. (**b**) The heat map representation for the important DEGs and their expression rates in tumor and periphery core tissues.

These 94 genes were again checked for overlap with highly variable genes which were obtained from Seurat (top 1000 genes). We ended up with 23 common genes (Supplementary Table S3); the expression of these important genes in tumor and periphery conditions are shown in Figure 4b. The heatmap shows the differential expression of these key genes in tumor and periphery tissues. Mainly, the cytokine genes (*CCL3, CCL4, CCL3L1*, and *CCL4L2*) show moderate expression in tumor core cells and show less expression in periphery cells, which correlates with the results from Darmanis et al. [17]. *MT2A, TIMP1*, and *GFAP* are more highly expressed in tumor cells and are down-regulated in the

periphery. The change in the expression level of the novel gene *AC243829.4* is not distinct; however, this lncRNA function is not reported. Similarly, the genes *FCGBP, THBS1, IGFBP2, IGFBP7,* and *FN1* are down-regulated in periphery cells and up-regulated in a disease condition. These 23 common genes are mainly involved in ECM–receptor interactions and epithelial–mesenchymal transitions (EMTs). EMT is a dynamic process of converting epithelial cells to mesenchymal phenotypes. Hypoxia, cytokines, growth factors secreted from the tumor environment, and treatment with anticancer drugs act as inducers for the EMT process [29]. Eventually, EMTs can create tumor cells with stem cell properties that are more aggressive and can increase their metastatic activity [30].

3.3. Tumor Heterogeneity with CNV Profiles

Genomic CNVs are commonly associated with tumor initiation and progression [31]. Thus, we attempted to infer large-scale copy-number alterations for each cell by averaging its relative expression levels over large genomic regions. This allowed us to suppress the individual gene-specific expression patterns and emphasize the signal of large-scale CNVs. InferCNV package was used to calculate CNV vectors for each cell and then clustered cells based on their respective profile CNV vectors (the details are given in the section on Method).

Overall, the non-neoplastic cells did not show any chromosomal abnormalities, only the neoplastic cells from all the patients having aberrations in their chromosomes (Figure 5). The CNV profiles revealed the coherent chromosomal aberration in each tumor cell. From the CNV profile, the gain of chromosome 7 expression and the loss of chromosomes 10 and 13 were constantly observed from the cells from all patient samples.

We mapped the DEGs which are located at chromosomes 7 and 10, which exhibited the copy number alternations in these chromosomes. Further, we filtered the genes based on transcript levels (gene expression) and observed that 10 potential DEGs are involved in the CNV of chromosome 7. The copy number gain of chromosome 7 is associated with the up-regulation of genes *ANLN, SRI, PON2, ITGB2, PTN, CAV1, PTPRZ1, MEST, CALD1, RAPGEF5,* and *NDUFA4*, which are mainly involved in epithelial–mesenchymal transitions. The down-regulated genes *NPTX2, PTPRZ1, EGFR,* and *COL1A2* also showed associations with the copy number gain of chromosome 7 and are involved in epithelial cell signaling. *EGFR* is the most frequently amplified oncogene in astrocytic tumors; expression of genes *SRI, NPTX2, MEST, RARRES2,* and *SEPTIN7* in association with GBM is reported in this study.

Similarly, the genes *SCD, EGR2, HTRA1, PIP4K2A, PHYHIPL, PSAP,* and *LINC00844* showed significant down-regulation with the loss of chromosome 10. These genes are mainly involved in Schwann cell differentiation and among them *SCD, OPALIN,* and *PHYHIPL* are novel genes. Wang et al. [32] reported that the Linc RNA (*LINC00844*) is associated with glioblastoma prognosis, cell proliferation, invasion, cell cycle, and metastasis [32]. However, genes *VIM* and *PPA1* showed up-regulation with the loss of chromosome 10. The up-regulation of *VIM* protein and its role in the formation of lamellipodia and invadopodia during cell invasion and migration have already been reported [33]. Prosaposin (*PSAP*) is highly expressed and secreted in gliomas and can promote glioma invasion and epithelial–mesenchymal transitions. Significant expression changes of these genes might also be the reason for the CNV change. A list of genes responsible for the chromosome alteration along with its expression changes is tabulated in Table 1.
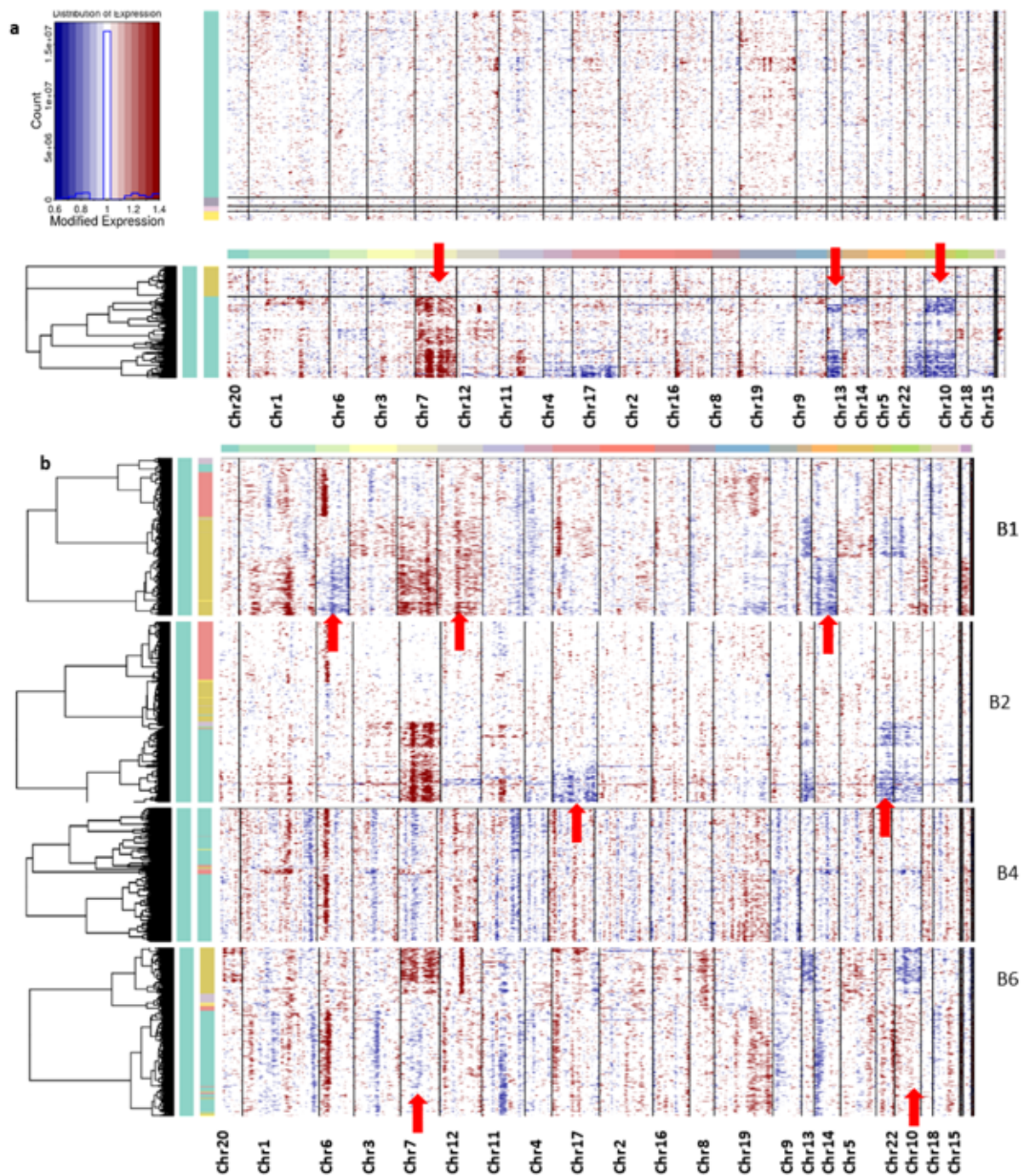
Figure 5. (**a**) Inferred chromosomal CNV profiles of complete cells based on average relative expression in a window of 100 genes. Oligodendrocytes that lack CNVs are shown as a reference group at the top. (**b**) The CNV profile of each patient's sample cells. Red indicates amplifications and blue indicates deletions.

Table 1. Candidate DEGs located on the chromosomes and their expression changes.

Chromosome	Gene Symbol	Start Position	End Position	Description	Expression
Chr 10	*VIM*	17228241	17237593	Vimentin	UP
	HTRA1	1.22×10^8	1.23×10^8	HtrA serine peptidase 1	DOWN
	SCD *	1×10^8	1×10^8	Stearoyl-CoA desaturase	DOWN
	PSAP	71816298	71851251	Prosaposin	DOWN
	EGR2	62811996	62819167	Early growth response 2	DOWN
	PIP4K2A	22534854	22714578	Phosphatidylinositol-5-phosphate 4-kinase type 2 α	DOWN
	SRGN	69057533	69104811	Serglycin	-
	PPA1	70202831	70233429	Inorganic pyrophosphatase 1	DOWN
	OPALIN *	96343221	96359002	Oligodendrocytic myelin paranodal and inner loop protein	NI
	PHYHIPL *	59175872	59247770	Phytanoyl-CoA 2-hydroxylase interacting protein like	DOWN
Chr 7	*ANLN*	36389806	36453791	Anillin actin binding protein	UP
	SRI *	88205118	88226993	Sorcin	UP
	PON2	95404863	95435329	Paraoxonase 2	UP
	ITGB8	20330702	20415754	Integrin subunit β 8	UP
	PTN	1.37×10^8	1.37×10^8	Pleiotrophin	UP
	CAV1	1.17×10^8	1.17×10^8	Caveolin 1	NI
	NPTX2 *	98617297	98629868	Neuronal pentraxin 2	DOWN
	PTPRZ1	1.22×10^8	1.22×10^8	Protein tyrosine phosphatase receptor type Z1	DOWN
	MEST *	1.30×10^8	1.31×10^8	Mesoderm specific transcript	UP
	RARRES2 *	1.50×10^8	1.50×10^8	Retinoic acid receptor responder 2	NI
	SEPTIN7 *	35800932	35907105	Septin 7	NI
	CALD1	1.35×10^8	1.35×10^8	Caldesmon 1	UP
	GNAI1	80133955	80219402	G protein subunit α I1	-
	RAPGEF5	22118238	22357144	Rap guanine nucleotide exchange factor 5	UP
	EGFR	55019021	55256620	Epidermal growth factor receptor	DOWN
	IGFBP3	45912245	45921874	Insulin-like growth factor binding protein 3	NI
	COL1A2	94394561	94431232	Collagen type I α 2 chain	DOWN
	GPR37	1.25E+08	1.25×10^8	G protein-coupled receptor 37	NI
	NDUFA4	10931951	10940256	NDUFA4 mitochondrial complex associated	UP
	GRM3	86643914	86864884	Glutamate metabotropic receptor 3	NI

Table 1. *Cont.*

Chromosome	Gene Symbol	Start Position	End Position	Description	Expression
Chr 13	*TSC22D1*	44432143	44577147	TSC22 domain family member 1	-
	*HSPH1 **	31134974	31162388	Heat shock protein family H (Hsp110) member 1	NI
	COL4A2	1.10E+08	1.11×10^8	Collagen type IV α 2 chain	NI
	*SLAIN1 **	77697854	77764242	SLAIN motif family member 1	DOWN
	*AMER2 **	25161684	25172288	APC membrane recruitment protein 2	DOWN
	*GPR183 **	99294530	99307405	G protein-coupled receptor 183	DOWN
	*PCDH9 **	66302834	67230445	Protocadherin 9	DOWN
	COL4A1	1.10×10^8	1.10×10^8	Collagen type IV α 1 chain	NI
	HMGB1	30456704	30617597	High mobility group box 1	NI
Chr6	*TNF*	31575565	31578336	Tumor necrosis factor	UP
	*F13A1 **	6144084	6320662	Coagulation factor XIII A chain	DOWN
	MYO6	75749203	75919537	Myosin VI	DOWN
	*AKAP12 **	1.51×10^8	151358559	A-kinase anchoring protein 12	DOWN
	*CD109 **	73696203	73828313	CD109 molecule	UP
	SLC16A10	1.11×10^8	111231194	Solute carrier family 16-member 10	DOWN
	*UST **	1.49×10^8	149076990	Uronyl 2-sulfotransferase	DOWN
	*IPCEF1 **	1.54×10^8	154356803	Interaction protein for cytohesin exchange factors 1	UP
	*TSPYL4 **	1.16×10^8	116254075	TSPY like 4	DOWN
	*SELPLG **	1.09×10^8	108633894	Selectin P ligand	UP
	ENO2	6914580	6923697	Enolase 2	NI
	DUSP6	89347235	89352501	Dual specificity phosphatase 6	UP
Chr 12	*C3AR1 **	8056844	8066359	Complement C3a receptor 1	DOWN
	*FAIM2 **	49866896	49903900	Fas apoptotic inhibitory molecule 2	UP
	FMNL3	49636499	49707405	Formin like 3	DOWN
	NAV3	77571856	78213010	Neuron navigator 3	DOWN
	GPN3	1.1×10^8	110468721	GPN-loop GTPase 3	DOWN
	PRPF40B	49622717	49645129	Pre-mRNA processing factor 40 homolog B	NI
	LGALS3	55129252	55145430	Complement C3a receptor 1	UP
	NDRG2	21016763	21070872	Fas apoptotic inhibitory molecule 2	DOWN
	HSPA2	64535905	64543237	Formin like 3	DOWN
	RTN1	59595976	59871288	Neuron navigator 3	UP
	SLC22A17	23346304	23354991	GPN-loop GTPase 3	UP
	PLD4	1.05×10^8	104937789	Pre-mRNA processing factor 40 homolog B	UP

Chromosome	Gene Symbol	Start Position	End Position	Description	Expression
Chr 17	*CCL2*	34255285	34257203	C–C motif chemokine ligand 2	UP
	SOX9	72121020	72126416	SRY-box transcription factor 9	DOWN
	CCL3	36088256	36090143	C–C motif chemokine ligand 3	UP
	CCL4	36103827	36105614	C–C motif chemokine ligand 4	UP
	ABCC3	50634881	50692253	ATP-binding cassette subfamily C member 3	UP
	CCL4L2	36211063	36212873	C–C motif chemokine ligand 4 like 2	UP
Chr 22	*MIF*	23894383	23895223	Macrophage migration inhibitory factor	DOWN
	LGALS1	37675636	37679802	Galectin 1	DOWN
	PDGFB	39223359	39244982	Platelet-derived growth factor subunit B	DOWN
	TNFRSF13C *	41922032	41926806	TNF receptor superfamily member 13C	DOWN
	CECR2	17359949	17558151	CECR2 histone acetyl-lysine reader	DOWN

* Novel genes reported in this study.

Interpatient heterogeneity is also studied from this CNV profile; the unique CNV abnormalities of each patient sample cell are described below. The CNV profile from each patient sample has been captured for deeper understanding (Figure 5b). It can be seen in the patient B1 sample that the neoplastic cells show the major changes, especially in the loss of chromosome 12 and the gain of chromosomes 6 and 14, whereas B2 samples show a gain of chromosomes 17 and 22. However, samples from patient B4 did not show significant changes in any chromosomes. Neoplastic cells from patient B6 samples display the disorder copy number changes in chromosomes 7, 10, and 13 in the neoplastic cells. The DEGs are located on chromosomes 6, 7, 10, 12, 13, 17, and 22, and their expression change is tabulated in Table 1.

3.4. Pathway/Function Enrichment Analysis for DEGs

The significant 94 DEGs and their gene ontology terms were calculated using ClueGO analysis. Gene ontology (GO) functional enrichment analyses [34] were used to determine the potential molecular mechanisms employed by these 94 important genes (33 up-regulated and 61 down-regulated). The pathways in which these genes are involved are discussed below.

ERK1 and ERK2 cascade: ClueGO analysis revealed that regulation of the ERK1 and ERK2 cascade is the major pathway experienced by the DEGs (Figure 6). Extracellular signal-regulated kinase plays a central role in transmitting extracellular signals to intracellular targets, including proliferation, differentiation, and survival. It has been reported that the ERK pathway is aberrantly activated in malignant gliomas [35]. In total, 16 genes actively participate in the ERK signaling cascade, of which 5 genes are up-regulated and 11 genes are down-regulated. The genes *CCL2, CCL3, CCL3L1, CCL4, CCN1, CD44, CHI3L1, CSF1R, DUSP6, FN1, IL1B, MIF, PDGFRA, PDZD2, TIMP1*, and *TNF* from our study are shown to be involved in the ERK1 and ERK2 pathway. Among them, *CD44, CHI3L1, FN1, MIF*, and *TIMP1* are up-regulated, whereas *CCL2, CCL3, CCL3L1, CCL4, DUSP6, CSF1R, IL1B, PDGFRA, PDZD2*, and *TNF* are down-regulated. *CHI3L1* binding to *CD44v3* activates ERK protein, *CD44* and *MIF* are required to activate the MEK–ERK–MAP kinase pathway [36]. The binding of *FN1* to integrins induces conformational changes and activates the recruitment of focal adhesion kinase (FAK) [37]. FAK signaling enhances the

activation of metalloproteinase and stimulates the diffuse nature of cells through ECM. ERK cascade signaling is essential to the production of inflammatory cytokines, and this signaling pathway is hyperactive in malignant gliomas due to overexpression of *EGFR* and *PDGFR* [38].

Figure 6. Functional annotations of the important DEGs (94 overlapping genes) obtained using ClueGO. (**a**) Go enrichment analysis; (**b**) Expression regulation network of DEGs.

Glial cell differentiation: The main function of glial cells is to provide support to neurons and maintain hemostasis. The DEGs identified from our study involved in glial cell differentiation are *CNP, DNER, EGR2, GFAP, GPR37L1, GSN, OLIG1, PTN, TNF,* and *VIM*. Among them, *GFAP, PTN*, and *VIM* are up-regulated and *CNP, DNER, EGR2, GPR37L1, GSN, OLIG1,* and *TNF* are down-regulated. The gene DNER is an epigenetically modified gene that enhances GBM progression. The genes *GFAP* and *VIM* are reported as glial cell markers, whereas *CNP* and *OLIG1* are novel genes from our study for glial cell differentiation. Generally, glial cells differentiate into neuron-rich and neuron-free regions which include astrocytes and oligodendrocytes. GFAP-positive radial glial cells transform into astrocytes, whereas *OLIG2* is involved in glial cell transformation into oligodendrocytes [39].

Mononuclear cell migration: Mononuclear cells are mainly monocytes and lymphocytes and these cells are critical components of the innate and adaptive immune system. The genes *ANXA1, APOD, C3AR1, CCL2, CCL3, CCL3L1, CCL4, CH25H, CSF1R, CX3CR1, THBS1,* and *TNF* are identified as important genes and are involved in mononuclear cell migration. We noticed that *ANXA1* and *CH25H*, which are involved in mononuclear cell migration, are reported here for the first time. The genes *ANXA1* and *THBS1* are up-regulated and *CCL2, CCL3, CCL3L1, CCL4, APOD, CX3CR1,* and *C3AR1* are down-regulated. The mononuclear cell migration pathway is associated with cell adhesion, leukocyte migration, chemotaxis, and inflammatory response. *CCL2, CCL4*, and *TNF* are mainly involved in all the pathways except cell adhesion, whereas *FN1* and *TNF* are involved in the cell adhesion pathway.

Fibroblast migration: Cytokines and growth factors, which are released as an inflammatory response, attract the fibroblast into the wound site, consequently starting the repair process. The genes *CCL2, CD44, EGR3, NR4A1, SULF2, THBS1,* and *TNC* are identified as potent genes for fibroblast migration. The genes *DE44, THBS1*, and *TNC* are up-regulated; *CCL2, EGR3, NR4A1*, and *SULF2* are down-regulated. Regulation of *NR4A1* and *TNC* expression in glioblastomas remains unreported; however, *TNC* is an ECM receptor protein

that involves ECM receptor interactions. The PI3K–Akt–mTOR, FGF, and ERK signaling pathways are involved in fibroblast migration. *EGR3*, *CD44*, and *CCL2* are known to be involved in the ERK signaling cascade.

Active migration is in the nature of glioma cells; it is a dynamic process of interaction between tumor cells and their microenvironment. Further, tumor cell attachment to the ECM plays an important role in tumor progression. The common feature of all GBM types is that aberrant kinase signaling includes the PI3K–Akt, MAPK, and ERK1/2 signaling pathways. Hannen et al. [40] suggested that ERK1/2 signaling is activated by phosphorylation in GBM (mainly in the mesenchymal subtype). A hypoxic microenvironment is created due to signaling disruptions which will lead to the recruitment of macrophages and microglia and to the release of growth factors (EGR, TGF, HGF, PDGF, HIF, and IGF1). The potential growth factors will be the inducers for epithelial–mesenchymal transitions (EMT) and numerous proteases that increase invasiveness into the surrounding normal brain [41–43].

3.5. The Transition of Microglial Immune Cells to Neoplastic Cells: Pseudotime Analysis

The main function of most of the DEGs is related to glial cell differentiation. Cell migration is mainly governed by the differentiation of microglial cells present in the brain. Usually, for healthy people, microglial cells in the brain regulate tissue homeostasis by surveying their environment. Their functions include phagocytizing the synaptic elements, living/dying cells, and apoptotic cells. In GBM disease conditions, where microglia are in close interaction with neurons, astrocytes and oligodendrocytes can shift into different functional states, modifying GBM proliferation and morphology [44]. During the response to inflammation or tumor growth, these microglial cells alter their morphological appearances and sometimes retract their ameboid appearance. These phenotypic changes in the immune cells have been associated with patient prognosis, though the detailed mechanism/crosstalk between GBMs and microglia is poorly understood [45–47]. Chen et al. [45] reported that the transition state of microglia significantly altered disease prognosis. Glioblastomas recruit neighboring resident microglia through the secretion of various chemokines and cytokines. *CCL2* and *CCL3* are associated with monocyte and macrophage recruitment and may act as chemoattractants to attract other microglial cells, causing disease invasion [48,49]. Bachiller et al. [44] showed that aging causes changes in gene expression as well as in the occurrence of dystrophic microglia. These changes related to aging might have an impact on the progression of neurodegenerative disorders [44].

We studied the differentiation state of microglial cells using pseudotime trajectory analysis. The immune cell landscape differs strongly between infiltrating and central regions of glioblastomas and changes from one cell type to another. Monocle 2 constructs the single-cell trajectories in pseudotime, which consist of two branch points (five branches and five states). The trajectory roots 1 and 2 (based on branch points 1 and 2) are more populated with immune and neoplastic cells, respectively (Figure 7). The trajectory starting state (state 1, dark blue) is crowded with immune cells. The highly differentiated cells (light blue) are populated with neoplastic cells (end-state 5). The transition paths from immune to neoplastic cells are represented with black arrows. The proposition of cells in each state also represents that the starting immune and neoplastic cell states have the highest proposition of cells (state 1 has 1829 cells and states 4 and 5 have 577 and 703 cells, respectively). The immune cells, especially those from cluster 0, are mainly microglial cells, which have migration potential, differentiating from states 1 to state 5 through state 3. To confirm this assumption, we used a heatmap to confirm the transition of microglial genes from immune cluster marker genes. The DEGs from the immune cluster which show significant expression changes are filtered out and used for the heatmap. The branch-dependent trajectory analysis of these microglial marker genes is shown in Figure 7b. As seen from Figure 7b, the marker genes *FN1*, *SLC1A2*, *SPARCL1*, *IGFBP7*, *MT2A*, *CYR61*, *CHI3L1*, and *TIMP1* have expression changes while traveling from root to branches B1 and B2, whereas *TNR*, *CNP*, and *APOD* show expression changes towards immune cells. Branches 2 and 3

are enriched with neoplastic cells. We set the root as branch point 2 and the genes show significant expression change while traveling towards B3 (neoplastic cells). The microglial immune cells are experiencing the transition state in pseudotime, which is observed with the heatmap analysis.
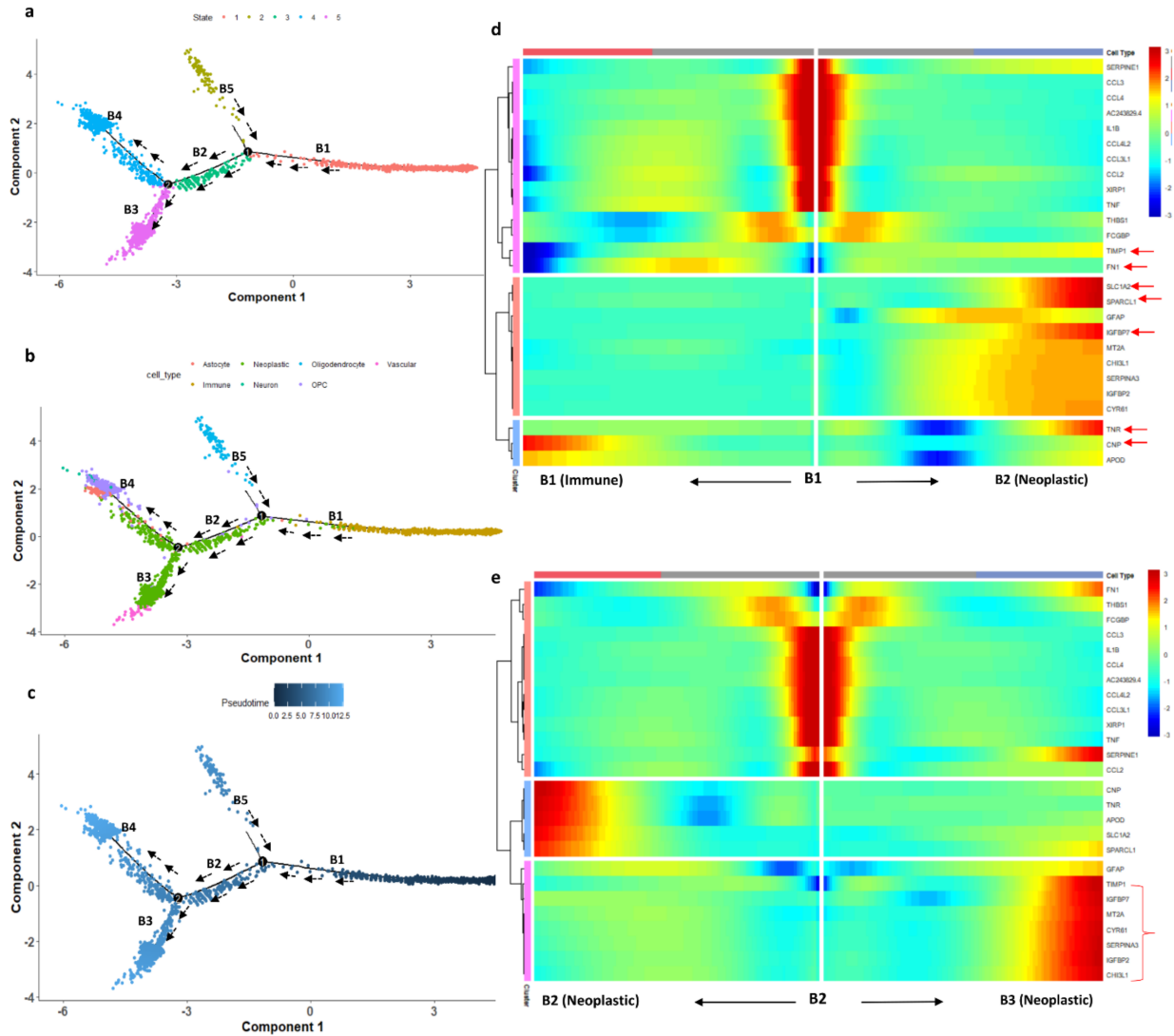

Figure 7. (**a**) DDR three-dimensional reduction of the cells computed using Monocle 2 contains five main branches; the cells took the path based on the pseudotime represented with the black arrow. (**b**) Trajectory states based on cell type. (**c**) Trajectory based on pseudotime. Dark blue represents the initial state (less differentiated) of the immune cells and light blue (more differentiated) represents the differentiated cell state (neoplastic cells). (**d**) Heatmap depicting the transition state of microglial marker genes in a branch-dependent manner for root points 1 and 2. Each row represents the dynamic expression of a gene. (**e**) Branch point 2 set as the root; the left arrow shows the root to the trajectory path 2; the right arrow shows the root to trajectory path 3.

3.6. Protein–Protein Interaction Networks (PPI) and Tissue-Specific Co-Expression Networks for DEGs

3.6.1. PPI Network in DEGs

The protein–protein interaction network and the tissue-specific gene co-expression network for the significant DEGs were obtained from the STRING and HumanBase databases, respectively. The 94 DEGs obtained from DESeq2 were provided as inputs to the STRING database for PPI networks with a confidence score of 0.8 (Figure 8). The network was

obtained from the STRING database with 103 nodes and 149 edges (PPI enrichment *p*-value < 1.0×10^{-16}).
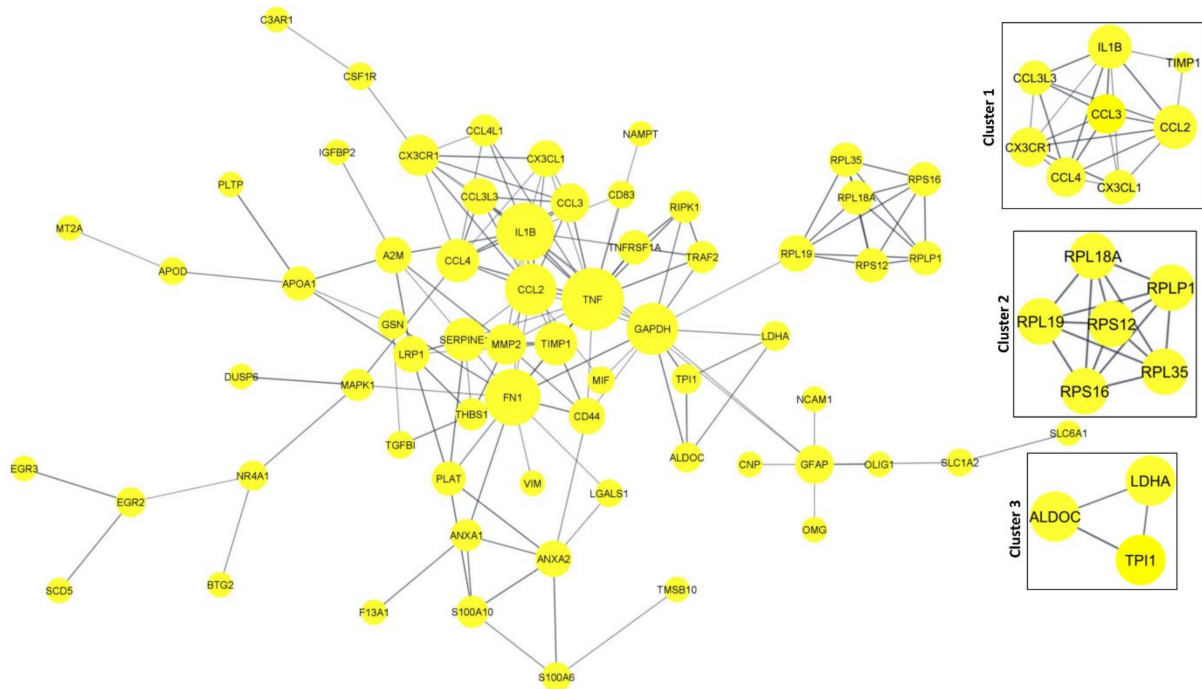
Figure 8. PPI network for the significant DEGs. Cluster and Hub gene screening from the network were performed using MCODE of Cytoscape.

The Cytoscape plugin MCODE was used for clustering the network and filtering out the hub genes. From MCODE we obtained cluster 1 with a score of 6.286 (eight DEGs), cluster 2 with a score of 6 (six DEGs), and cluster 3 with a score of 3.3 (three DEGs) (Figure 8). Network node sizes are represented based on the degrees of connectivity. The PPI network shows that *TNF, CCL2, IL1B, FN1, CCL4*, and *GAPDH* are the important genes having a greater number of interactions with the high degree of 6. All other genes except *GAPDH* are down-regulated. Based on the degrees of connectivity and the cluster analysis, *RPL19, CX3CR1, LDHA, TPI1, RIPK1, TRAF2, GFAP, APOA1, PLAT, TNF, CCL2, IL1B, FN1, CCL4, GAPDH*, and *ALDOC* are termed hub genes for the interactions. These hub genes are mainly involved in the vitamin B12 metabolism pathway. The important biological function of vitamin B12 is to accomplish DNA synthesis, which is necessary for cell division. All the living cells require vitamin B12 for survival and it strongly promotes glioblastoma cell proliferation; therefore, B12 deficiency is not favorable to GBM prognosis [50].

3.6.2. Tissue-Specific Co-Expression Network of DEGs

The human base tissue-specific network displays the co-expression network specific to the 94 DEGs. The reference network was obtained from human brain cortex tissue; the DEGs were mapped against the reference network and our co-expression network was constructed (Figure 9). The genes *CCL4L2, HTRA1, GFAP, PLTP, PDGFRA*, and *GSN* constitute the central part of the network with high closeness centrality (~0.6) and high-degree parameters (49). *GFAP* protein-expressing neural stem cells (NSCs) are responsible for the activation of PDGF receptors and their stimulation, which eventually form the glioma-like mass [51]. *PDGFRA* is also responsible for proinflammatory cytokines, including *CCL4L2*. We have identified the genes co-expressed together: gene *CCL4L2* is co-expressed with *PLTP, SELPLG, ANXA2, RPLP1*, and *P2RY12*. *CCL4L2, SELPLG*, and *P2RY12* are down-regulated, whereas *PLTP, RPLP1*, and *ANXA2* are up-regulated. The genes *SULF2, PTN, FTL, TPI1, LDHA, THBS1*, and *TGFBI* are co-expressed together. *CCL3L1* interaction with *ATP1A2* has the highest

centrality; *HTRA1* interaction with *B3GNT5, GFAP, PLTP,* and *GSN* is shown to form the highest interaction network (Figure 9).
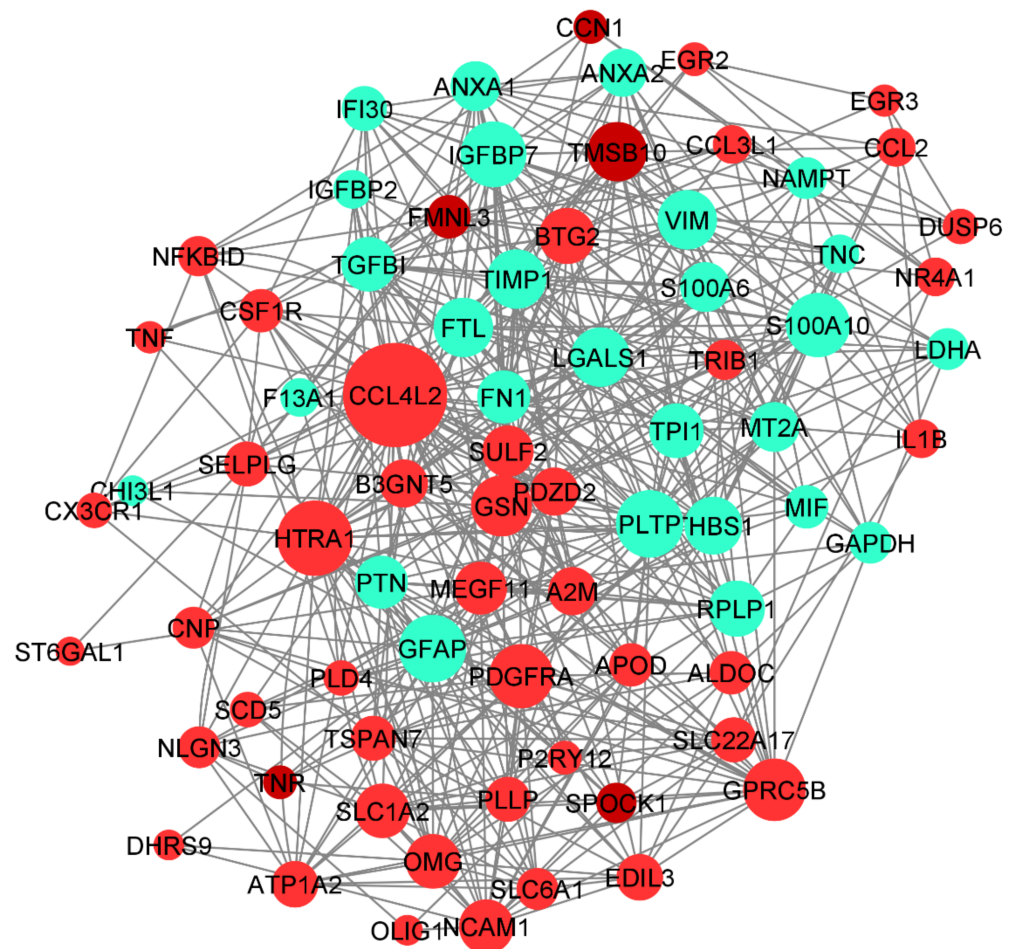
Figure 9. Co-expression network of DEGs constructed using Humanbase and visualized with Cytoscape. Proteins are represented with color nodes; the size of the nodes represents the degree parameter. The networks for up- and down-regulated genes are represented in red and cyan, respectively.

Increased expression of *B3GNT5* strongly correlated with the progression of breast cancer, lung cancer, and ovarian cancer. However, *B3GNT5* is identified as a novel gene for GBM progression. Copy number alteration of *HTRA1* is strongly associated with poor prognosis of GBM; however, there has been no clear study of the expression level of *HTRA1*. Therefore, the interaction of *HTRA1* with *B3GNT5, GFAP, PLTP*, and *GSN* could be a novel identification in the GBM disease mechanism.

4. Discussion

In this study, we utilized scRNA-data to explore the molecular cascade of GBM disease progression. Initially, we performed a single-cell transcriptome analysis to characterize tumor heterogeneity and the molecular mechanism of GBM invasion. It has been reported that GBM comprises morphologically and phenotypically diverse cells and cell types [4]. We have identified 7 major clusters and 16 subclusters from the population of 3389 cells. The functional annotation of the cluster-specific marker genes displays distinct gene ontology, representing the heterogeneity of the cells. The CNV profile explains the intra-patient tumor heterogeneity and the data shown in Figure 5 demonstrate the distinct large-scale chromosomal variation across individual patient samples. The gain of chromosome 7 and the loss of chromosomes 10 and 13 are common to all the patient sample cells, and the chromosomal aberrations are mainly caused by neoplastic cells.

According to experimental reports, the gain of chromosome 7 has been associated with the simultaneous loss of chromosome 10 and suggests that *EGFR* amplification and the deletion mutation of *PTEN* is associated with the abnormalities of chromosomes 7 and 10, respectively [52]. Along with chromosomes 7 and 10, abnormalities in chromosomes 12, 13, and 6 have also been reported for patients with human gliomas, Alzheimer's disease, and Parkinson's disease [53]. Transcriptional downregulation of the *NDRG2* gene on chromosome 14 has been observed in primary glioblastoma patients [54]. All the above-reported genes are also observed to be the same in our study. Loss of chromosomes 17 and 22 have been reported in gliomas; however, we have observed a gain in chromosomal abnormalities [55,56]. This might be because, except for the *SOX9* oncogene, the chemokine genes (*CCL2, CCL3, CCL4,* and *CCL4L2*) of chromosome 17 are up-regulated. Apart from these reported genes, other novel DEGs contributing to chromosomal abnormalities are tabulated in Table 1. However, further experimental results and analysis are needed for these genes.

We have identified the DEGs from tumor and periphery cells, suggesting the functionally important genes. Darmanis et al. [17] identified 30 potent differentially expressed genes across neoplastic and non-neoplastic cells, and among them *CA2, GAP43, PMP2, CRYAB, SOX9, EGFR,* and *ATP1A2* are reported to be DEGs between tumor and periphery cells in our analysis. The filtered DEGs are further used for the PPI and co-expression network construction. The DEGs *DHRS9, IPCEF1, TNR, MEGF11, EDIL3, PDZD2, PDGFRA, SPOCK1, CHI3L1, FN1, IGFBP2, TNC, FCGBP, CYR61, F13A1, ANXA2 NCAM1, RPL19, SLC1A12, CNP, MT2A, CHI3L1, POSTIN, LTF, MPDZ, CPZ, LRRC32, CTNNA3, LRFN5,* and *SLc22A17* reported in our study are not explored by many other studies related to GBM. The novel genes identified in our study have been tabulated in Table 2. The main functions of the DEGs are mostly involved in the ERK1/2 signaling cascade, glial cell differentiation, and mononuclear migration. These main pathways are interlinked with ECM cell adhesion and adhesion to surrounding cells, and the ECM is important for GBM cancer cell maintenance. The DEGs are involved in cell adhesion and ECM receptor interactions. It is evident that the above-mentioned potent DEGs and the novel DEGs play an important role in GBM progression. Targeting these DEGs and their pathway might help for therapeutic applications.

Table 2. Novel genes identified from our study.

Analysis Type	Novel Genes	Total
Filtered DEGs	*DHRS9, IPCEF1, TNR, MEGF11, EDIL3, PDZD2, ATP1A2, PDGFRA, LINC00632, AC243829.4, AC024909.2, MEG3, CHI3L1, FN1, IGFBP2, TNC, FCGBP, CYR61, F13A1, ANXA2, AC006064.4, ANXA1, CH25H and MIF-AS1*	24
CNV detection	*SCD, OPALIN, PHYHIPL, PSAP, SRI, NPTX2, MEST, RARRES2, SEPTIN7, HSPH1, SLAIN1, AMER2, GPR183, PCDH9, F13A1, AKAP12, CD109, UST, IPCEF1, TSPYL4, SELPLG, C3AR1, FAIM2 and TNFRSF13C*	24
Network construction	*B3GNT5, SELPLG and TPI1*	3

We have also explored glial cell differentiation by identifying the microglial cell transition state by single-cell trajectory reconstruction. The pseudotime analysis revealed that the microglial cells differentiate into neoplastic cells. The trajectory analysis identified the pseudotime starting state as immune cells and the ending state as neoplastic cells. We determined the path along which cells traveled from root to branch, representing the transition of the cells (Figure 7). Some of the marker genes from microglial cells, such as

FN1, SLC1A2, SPARCL1, IGFBP7, MT2A, CYR61, CHI3L1, and *TIMP1*, showed a gradual increase in expression while they traveled from root to branch. A glioblastoma takes control over the microglial immune cells of periphery tissue by various kinds of cytokine and chemokine signaling [9]. The invasive microglial cells from periphery cells are a possible target for the GBM disease breakthrough.

Comparison with Other Related Works

Several investigations have been carried out to understand the tumor heterogeneity of glioblastomas using RNA-seq analysis [57]. Patel et al. [4] showed that the glioblastoma subtype classifiers are variably expressed across individual cells and discussed the intra-tumoral heterogeneity within a tumor. Neftel et al. [11] compared the scRNA-seq of 20 adults and 8 pediatric glioblastomas (24,131 cells in total) and identified four main cellular states. Darmanis et al. [17] reported the characterization of neoplastic and non-neoplastic cells and determined that the neoplastic cells share common characteristics regardless of the patient of origin. However, GBM heterogeneity across tumor and periphery cells and the molecular mechanism underlying the transition of distinct cell types are not completely explored.

We have identified cell type-specific DEGs and DEGs between tumor and periphery cells to characterize tumor and periphery cells. The filtered cell type-specific DEGs are used for hub gene identification using PPI and co-expression network analysis, and 51 novel genes were identified in our study (Table 2). Among the novel differentially expressed genes, *IPCEF1, F13A1, TPI1, B3GNT5, ATP1A2, FN1, PDGFRA*, and *GSN* are found to be hub genes, and we suggest that these genes could be potential therapeutic targets in GBM disease prognosis. In addition, we compared the significantly expressed genes identified in our analysis with other experimental studies reported in the literature [11,17,58–61] and the results are presented in Supplementary File S2. We observed that 422 genes can be matched with other studies.

We have also explored the transition state of microglial immune cells during the invasion process with pseudotime analysis. The trajectory path starts with immune cells and ends with neoplastic cells. The differentiation state of immune microglial genes (*FN1, SLC1A2, SPARCL1, IGFBP7, MT2A, CYR61, CHI3L1*, and *TIMP1*) showed the expression changes while traveling from the root to the end state (Figure 7). In addition, we were able to capture the patient-wise chromosome alterations in chromosomes 13, 14, 17, and 22, and the genes responsible for the alterations.

5. Conclusions

We have performed single-cell transcriptome data analysis of 3389 cells from four primary glioblastoma patients and provided a detailed description of cellular heterogeneity in GBM. This study emphasized the tumor heterogeneity of GBM samples and the chromosomal aberration associated with it. We studied the cell differentiation state of glial cells, suggesting that the microglial cells are the possible target cells for GBM to invade. The filtered differentially expressed genes were used to construct the co-expression and PPI (protein-protein interaction) networks and subsequently identified the novel potential genes, which may act as therapeutic targets.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/genes13030428/s1, Figure S1: (a) The cell quality filtering with Seurat. (b) The genes show the higher cell to cell variation (Red dots), black dots represent housekeeping genes. (c) The correlation of highly variable genes from first two PCs. (d) represent the elbow plot for the number of PCs and its standard deviation. (e) Histogram representing the number of cells in each cluster. Table S1: Particulars of the dataset obtained from GSE84465. Table S2: List of DEGs analyzed from Seurat and DESeq2. Table S3: Overlapped DEGs between Seurat, DESeq2 and highly variable genes. File S1: List of cell-type markers and variable genes. File S2: Comparison of identified potent genes in the present study with those reported in the literature.

References

1. Grégoire, H.; Roncali, L.; Rousseau, A.; Chérel, M.; Delneste, Y.; Jeannin, P.; Hindré, F.; Garcion, E. Targeting tumor associated macrophages to overcome conventional treatment resistance in glioblastoma. *Front. Pharmacol.* **2020**, *11*, 368. [CrossRef] [PubMed]
2. Parsons, D.W.; Jones, S.; Zhang, X.; Lin, J.C.H.; Leary, R.J.; Angenendt, P.; Mankoo, P.; Carter, H.; Siu, I.M.; Gallia, G.L.; et al. An integrated genomic analysis of human glioblastoma multiforme. *Science* **2008**, *321*, 1807–1812. [CrossRef] [PubMed]
3. Couturier, C.P.; Ayyadhury, S.; Le, P.U.; Nadaf, J.; Monlong, J.; Riva, G.; Allache, R.; Baig, S.; Yan, X.; Bourgey, M.; et al. Single-cell RNA-seq reveals that glioblastoma recapitulates a normal neurodevelopmental hierarchy. *Nat. Commun.* **2020**, *11*, 3406. [CrossRef] [PubMed]
4. Patel, A.P.; Tirosh, I.; Trombetta, J.J.; Shalek, A.K.; Gillespie, S.M.; Wakimoto, H.; Cahill, D.P.; Nahed, B.V.; Curry, W.T.; Martuza, R.L.; et al. Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science* **2014**, *344*, 1396–1401. [CrossRef] [PubMed]
5. Gerlinger, M.; Rowan, A.J.; Horswell, S.; Larkin, J.; Endesfelder, D.; Gronroos, E.; Martinez, P.; Matthews, N.; Stewart, A.; Tarpey, P.; et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **2012**, *366*, 883–892. [CrossRef]
6. Liu, A.; Hou, C.; Chen, H.; Zong, X.; Zong, P. Genetics and epigenetics of glioblastoma: Applications and overall incidence of IDH1 mutation. *Front. Oncol.* **2016**, *6*, 16. [CrossRef] [PubMed]
7. Cloughesy, T.F.; Cavenee, W.K.; Mischel, P.S. Glioblastoma: From molecular pathology to targeted treatment. *Annu. Rev. Pathol. Mech. Dis.* **2014**, *9*, 1–25. [CrossRef]
8. Brown, N.F.; Carter, T.J.; Ottaviani, D.; Mulholland, P. Harnessing the immune system in glioblastoma. *Br. J. Cancer* **2018**, *119*, 1171–1181. [CrossRef] [PubMed]
9. Maas, S.L.; Abels, E.R.; Van De Haar, L.L.; Zhang, X.; Morsett, L.; Sil, S.; Guedes, J.; Sen, P.; Prabhakar, S.; Hickman, S.E.; et al. Glioblastoma hijacks microglial gene expression to support tumor growth. *J. Neuroinflamm.* **2020**, *17*, 120. [CrossRef] [PubMed]
10. Valdebenito, S.; Malik, S.; Luu, R.; Loudig, O.; Mitchell, M.; Okafo, G.; Bhat, K.; Prideaux, B.; Eugenin, E.A. Tunneling nanotubes, TNT, communicate glioblastoma with surrounding non-tumor astrocytes to adapt them to hypoxic and metabolic tumor conditions. *Sci. Rep.* **2021**, *11*, 14556. [CrossRef] [PubMed]
11. Neftel, C.; Laffy, J.; Filbin, M.G.; Hara, T.; Shore, M.E.; Rahme, G.J.; Richman, A.R.; Silverbush, D.; Shaw, M.L.; Hebert, C.M.; et al. An integrative model of cellular states, plasticity, and genetics for glioblastoma. *Cell* **2019**, *178*, 835–849. [CrossRef] [PubMed]
12. Prionisti, I.; Bühler, L.H.; Walker, P.R.; Jolivet, R.B. Harnessing microglia and macrophages for the treatment of glioblastoma. *Front. Pharmacol.* **2019**, *10*, 506. [CrossRef] [PubMed]
13. Galavotti, S.; Bartesaghi, S.; Faccenda, D.; Shaked-Rabi, M.; Sanzone, S.; McEvoy, A.; Dinsdale, D.; Condorelli, F.; Brandner, S.; Campanella, M.; et al. The autophagy-associated factors DRAM1 and p62 regulate cell migration and invasion in glioblastoma stem cells. *Oncogene* **2013**, *32*, 699–712. [CrossRef] [PubMed]
14. Wang, K.; Chen, D.; Qian, Z.; Cui, D.; Gao, L.; Lou, M. Hedgehog/Gli1 signaling pathway regulates MGMT expression and chemoresistance to temozolomide in human glioblastoma. *Cancer Cell Int.* **2017**, *17*, 117. [CrossRef] [PubMed]
15. Waghela, B.N.; Vaidya, F.U.; Ranjan, K.; Chhipa, A.S.; Tiwari, B.S.; Pathak, C. AGE-RAGE synergy influences programmed cell death signaling to promote cancer. *Mol. Cell. Biochem.* **2021**, *476*, 585–598. [CrossRef] [PubMed]

16. Tang, J.; Shao, W.; Dorak, M.T.; Li, Y.; Miike, R.; Lobashevsky, E.; Wiencke, J.K.; Wrensch, M.; Kaslow, R.A.; Cobbs, C.S. Positive and negative associations of human leukocyte antigen variants with the onset and prognosis of adult glioblastoma multiforme. *Cancer Epidemiol. Prev. Biomark.* **2005**, *14*, 2040–2044. [CrossRef] [PubMed]

17. Darmanis, S.; Sloan, S.A.; Croote, D.; Mignardi, M.; Chernikova, S.; Samghababi, P.; Zhang, Y.; Neff, N.; Kowarsky, M.; Caneda, C.; et al. Single-Cell RNA-seq analysis of infiltrating neoplastic cells at the migrating front of human glioblastoma. *Cell Rep.* **2017**, *21*, 1399–1410. [CrossRef] [PubMed]

18. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **2011**, *17*, 10–12. [CrossRef]

19. Patro, R.; Duggal, G.; Love, M.I.; Irizarry, R.A.; Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **2017**, *14*, 417–419. [CrossRef] [PubMed]

20. Satija, R.; Farrell, J.A.; Gennert, D.; Schier, A.F.; Regev, A. Spatial reconstruction of single-cell gene expression data. *Nat. Biotechnol.* **2015**, *33*, 495–502. [CrossRef] [PubMed]

21. Tickle, T.; Tirosh, I.; Georgescu, C.; Brown, M.; Haas, B. *InferCNV of the Trinity CTAT Project. Klarman Cell Observatory*; Broad Institute of MIT and Harvard: Cambridge, MA, USA, 2019. Available online: https://github.com/broadinstitute/inferCNV (accessed on 1 March 2021).

22. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef] [PubMed]

23. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [CrossRef] [PubMed]

24. Trapnell, C.; Cacchiarelli, D.; Grimsby, J.; Pokharel, P.; Li, S.; Morse, M.; Lennon, N.J.; Livak, K.J.; Mikkelsen, T.S.; Rinn, J.L. The dynamics and regulators of cell fate decisions are revealed by pseudotemporal ordering of single cells. *Nat. Biotechnol.* **2014**, *32*, 381–386. [CrossRef]

25. Alentorn, A.; Marie, Y.; Carpentier, C.; Boisselier, B.; Giry, M.; Labussiere, M.; Mokhtari, K.; Hoang-Xuan, K.; Sanson, M.; Delattre, J.Y.; et al. Prevalence, clinico-pathological value, and co-occurrence of PDGFRA abnormalities in diffuse gliomas. *Neuro-oncol.* **2012**, *14*, 1393–1403. [CrossRef] [PubMed]

26. Chakravarty, D.; Pedraza, A.M.; Cotari, J.; Liu, A.H.; Punko, D.; Kokroo, A.; Huse, J.T.; Altan-Bonnet, G.; Brennan, C.W. EGFR and PDGFRA co-expression and heterodimerization in glioblastoma tumor sphere lines. *Sci. Rep.* **2017**, *7*, 9043. [CrossRef]

27. Duman, C.; Yaqubi, K.; Hoffmann, A.; Acikgöz, A.A.; Korshunov, A.; Bendszus, M.; Herold-Mende, C.; Liu, H.K.; Alfonso, J. Acyl-CoA-binding protein drives glioblastoma tumorigenesis by sustaining fatty acid oxidation. *Cell Metab.* **2019**, *30*, 274–289. [CrossRef] [PubMed]

28. Mohiuddin, E.; Wakimoto, H. Extracellular matrix in glioblastoma: Opportunities for emerging therapeutic approaches. *Am. J. Cancer Res.* **2021**, *11*, 3742.

29. Monteiro, A.R.; Hill, R.; Pilkington, G.J.; Madureira, P.A. The role of hypoxia in glioblastoma invasion. *Cells* **2017**, *6*, 45. [CrossRef]

30. Roche, J. The epithelial-to-mesenchymal transition in cancer. *Cancers* **2018**, *10*, 52. [CrossRef]

31. Kaneshiro, D.; Kobayashi, T.; Chao, S.T.; Suh, J.; Prayson, R.A. Chromosome 1p and 19q deletions in glioblastoma multiforme. *Appl. Immunohistochem. Mol. Morphol.* **2009**, *17*, 512–516. [CrossRef]

32. Wang, Y.; Zhou, W.; Ma, S.; Guan, X.; Zhang, D.; Peng, J.; Wang, X.; Yuan, L.; Li, P.; Mao, B.; et al. Identification of a Glycolysis-Related LncRNA Signature to Predict Survival in Diffuse Glioma Patients. *Front. Oncol.* **2021**, *10*, 3316. [CrossRef] [PubMed]

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

34. Kanehisa, M.; Goto, S.; Sato, Y.; Kawashima, M.; Furumichi, M.; Tanabe, M. Data, information, knowledge and principle: Back to metabolism in KEGG. *Nucleic Acids Res.* **2014**, *42*, D199–D205. [CrossRef]

35. Guo, Y.J.; Pan, W.W.; Liu, S.B.; Shen, Z.F.; Xu, Y.; Hu, L.L. ERK/MAPK signalling pathway and tumorigenesis. *Exp. Ther. Med.* **2020**, *19*, 1997–2007. [CrossRef] [PubMed]

36. Mooney, K.L.; Choy, W.; Sidhu, S.; Pelargos, P.; Bui, T.T.; Voth, B.; Barnette, N.; Yang, I. The role of CD44 in glioblastoma multiforme. *J. Clin. Neurosci.* **2016**, *34*, 1–5. [CrossRef] [PubMed]

37. Spada, S.; Tocci, A.; Di Modugno, F.; Nisticò, P. Fibronectin as a multiregulatory molecule crucial in tumor matrisome: From structural and functional features to clinical practice in oncology. *J. Exp. Clin. Cancer Res.* **2021**, *40*, 102. [CrossRef] [PubMed]

38. Lo, H.W. Targeting Ras-RAF-ERK and its interactive pathways as a novel therapy for malignant gliomas. *Curr. Cancer Drug Targets* **2010**, *10*, 840–848. [CrossRef] [PubMed]

39. Janas, M.S.; Nowakowski, R.S.; Møllgård, K. Glial cell differentiation in neuron-free and neuron-rich regions. *Anat. Embryol.* **1991**, *184*, 559–569. [CrossRef]

40. Hannen, R.; Hauswald, M.; Bartsch, J.W. A rationale for targeting extracellular regulated kinases ERK1 and ERK2 in glioblastoma. *J. Neuropathol. Exp. Neurol.* **2017**, *76*, 838–847. [CrossRef]

41. Iwadate, Y. Epithelial-Mesenchymal transition in glioblastoma progression. *Oncol. Lett.* **2016**, *11*, 1615–1620. [CrossRef]

42. Tao, C.; Huang, K.; Shi, J.; Hu, Q.; Li, K.; Zhu, X. Genomics and prognosis analysis of epithelial-mesenchymal transition in glioma. *Front. Oncol.* **2020**, *10*, 183. [CrossRef] [PubMed]

43. Yu, S.; Yu, X.; Sun, L.; Zheng, Y.; Chen, L.; Xu, H.; Jin, J.; Lan, Q.; Chen, C.C.; Li, M. GBP2 enhances glioblastoma invasion through Stat3/fibronectin pathway. *Oncogene* **2020**, *39*, 5042–5055. [CrossRef]

44. Bachiller, S.; Jiménez-Ferrer, I.; Paulus, A.; Yang, Y.; Swanberg, M.; Deierborg, T.; Boza-Serrano, A. Microglia in neurological diseases: A road map to brain-disease dependent-inflammatory response. *Front. Cell. Neurosci.* **2018**, *12*, 488. [CrossRef] [PubMed]

45. Chen, J.W.E.; Lumibao, J.; Leary, S.; Sarkaria, J.N.; Steelman, A.J.; Gaskins, H.R.; Harley, B.A. Crosstalk between microglia and patient-derived glioblastoma cells inhibit invasion in a three-dimensional gelatin hydrogel model. *J. Neuroinflamm.* **2020**, *17*, 346. [CrossRef] [PubMed]

46. Noy, R.; Pollard, J.W. Tumor-Associated macrophages: From mechanisms to therapy. *Immunity* **2014**, *41*, 49–61. [CrossRef] [PubMed]

47. Matias, D.; Balça-Silva, J.; da Graça, G.C.; Wanjiru, C.M.; Macharia, L.W.; Nascimento, C.P.; Roque, N.R.; Coelho-Aguiar, J.M.; Pereira, C.M.; Dos Santos, M.F.; et al. Microglia/astrocytes–glioblastoma crosstalk: Crucial molecular mechanisms and microenvironmental factors. *Front. Cell. Neurosci.* **2018**, *12*, 235. [CrossRef] [PubMed]

48. Deshmane, S.L.; Kremlev, S.; Amini, S.; Sawaya, B.E. Monocyte chemoattractant protein-1 (MCP-1): An overview. *J. Interferon Cytokine Res.* **2009**, *29*, 313–326. [CrossRef]

49. Sciumè, G.; Santoni, A.; Bernardini, G. Chemokines and glioma: Invasion and more. *J. Neuroimmunol.* **2010**, *224*, 8–12. [CrossRef] [PubMed]

50. Rzepka, Z.; Maszczyk, M.; Beberok, A.; Hermanowicz, J.M.; Pawlak, D.; Gryko, D.; Wrześniok, D. Response of human glioblastoma cells to vitamin B12 deficiency: A study using the non-toxic cobalamin antagonist. *Biology* **2021**, *10*, 69. [CrossRef]

51. Chojnacki, A.; Mak, G.; Weiss, S. PDGFRα expression distinguishes GFAP-expressing neural stem cells from PDGF-responsive neural precursors in the adult periventricular area. *J. Neurosci.* **2011**, *31*, 9503–9512. [CrossRef]

52. Del Mar Inda, M.; Fan, X.; Muñoz, J.; Perot, C.; Fauvet, D.; Danglot, G.; Palacio, A.; Madero, P.; Zazpe, I.; Portillo, E.; et al. Chromosomal abnormalities in human glioblastomas: Gain in chromosome 7p correlating with loss in chromosome 10q. *Mol. Carcinog. Publ. Coop. Univ. Tex. MD Anderson Cancer Cent.* **2003**, *36*, 6–14. [CrossRef] [PubMed]

53. Gupta, M.K.; Jayaram, S.; Madugundu, A.K.; Chavan, S.; Advani, J.; Pandey, A.; Thongboonkerd, V.; Sirdeshmukh, R. Chromosome-centric human proteome project: Deciphering proteins associated with glioma and neurodegenerative disorders on chromosome 12. *J. Proteome Res.* **2014**, *13*, 3178–3190. [CrossRef] [PubMed]

54. Tepel, M.; Roerig, P.; Wolter, M.; Gutmann, D.H.; Perry, A.; Reifenberger, G.; Riemenschneider, M.J. Frequent promoter hypermethylation and transcriptional downregulation of the NDRG2 gene at 14q11. 2 in primary glioblastoma. *Int. J. Cancer* **2008**, *123*, 2080–2086. [CrossRef] [PubMed]

55. Hu, J.; Jiang, C.C.; Ho-Keung, N.; Pang, J.C.; Tong, C.Y. Chromosome 17P may harbor multiple tumor suppressor genes associated with primary glioblastoma multiforme. *Chin. J. Cancer Res.* **2002**, *14*, 60–63. [CrossRef]

56. Laigle-Donadey, F.; Crinière, E.; Benouaich, A.; Lesueur, E.; Mokhtari, K.; Hoang-Xuan, K.; Sanson, M. Loss of 22q chromosome is related to glioma progression and loss of 10q. *J. Neuro-oncol.* **2006**, *76*, 265–268. [CrossRef] [PubMed]

57. Martínez, A.H.; Madurga, R.; García-Romero, N.; Ayuso-Sacido, Á. Unravelling glioblastoma heterogeneity by means of single-cell RNA sequencing. *Cancer Lett.* **2022**, *527*, 66–79. [CrossRef]

58. Kalderimis, A.; Lyne, R.; Butano, D.; Contrino, S.; Lyne, M.; Heimbach, J.; Hu, F.; Smith, R.; Štěpán, R.; Sullivan, J.; et al. InterMine: Extensive web services for modern biology. *Nucleic Acids Res.* **2014**, *42*, W468–W472. [CrossRef] [PubMed]

59. Sankowski, R.; Böttcher, C.; Masuda, T.; Geirsdottir, L.; Sindram, E.; Seredenina, T.; Muhs, A.; Scheiwe, C.; Shah, M.J.; Heiland, D.H.; et al. Mapping microglia states in the human brain through the integration of high-dimensional techniques. *Nat. Neurosci.* **2019**, *22*, 2098–2110. [CrossRef] [PubMed]

60. Richards, L.M.; Whitley, O.K.; MacLeod, G.; Cavalli, F.M.; Coutinho, F.J.; Jaramillo, J.E.; Svergun, N.; Riverin, M.; Croucher, D.C.; Kushida, M.; et al. Gradient of developmental and injury response transcriptional states defines functional vulnerabilities underpinning glioblastoma heterogeneity. *Nat. Cancer* **2021**, *2*, 157–173. [CrossRef] [PubMed]

61. Xie, Y.; He, L.; Lugano, R.; Zhang, Y.; Cao, H.; He, Q.; Chao, M.; Liu, B.; Cao, Q.; Wang, J.; et al. Key molecular alterations in endothelial cells in human glioblastoma uncovered through single-cell RNA sequencing. *JCI Insight* **2021**, *6*, e150861. [CrossRef] [PubMed]