

Meta-Analysis of Brain and Central Nervous System Microarray Datasets

Icxa Khandelwal¹, Aditi Sharma², Jayashree Ramana^{1*}

Department of Biotechnology & Bioinformatics, Jaypee University of Information Technology, Waknaghat, Solan (H.P),
India.

Article Info

Article history:

Received Jul 31st, 2017

Revised Dec 12th, 2017

Accepted Apr 4th, 2018

Keyword:

Brain and CNS cancer
Microarray Datasets
Meta-Analysis
Up-regulated genes
Significant Pathways

ABSTRACT (10 PT)

Brain and CNS cancer are rare in comparison to other types of cancer. Currently there are no effective therapies for their treatment. In this study, meta-analysis of microarray datasets of Brain and CNS cancer was done to obtain significantly upregulated genes with increased statistical power and generalizability. A total of 130 significantly up-regulated genes were obtained. Some of the genes found during analysis have not yet been associated with this cancer. Different biological networks were created and analyzed using the significantly up-regulated genes as input. For each network, the most significant pathways have also been identified computationally.

Copyright © 2018 *International Journal for Computational Biology*,
<http://www.ijcb.in>. All rights reserved.

Corresponding Author:

Jayashree Ramana,
Department of Biotechnology &
Bioinformatics,
Jaypee University of Information
Technology,
Waknaghat, Solan (H.P), India
Email: jayashree.ramana@juit.ac.in



How to Cite:

First Author name *et. al.* Title of the Manuscript.
IJCB. 2018; Volume 7 (Issue 1): Page 03-15.

1. INTRODUCTION

Brain and CNS cancer are classified as heterogenous tumors on genetic and biological basis [1]. These cancers comprise for approximately 3% of worldwide cancer cases. Though these are rare, there are no effective therapies for their treatment because of the relatively inaccessible location. They have been reported to be more commonly observed in men as compared to women [2]. With the variation in age as well as histological type, the prognoses of brain and CNS cancers differs significantly. Prognosis is significantly poor for old aged people and people suffering from glioblastomas [3]. The increased rate of survival in higher income countries can be directly linked to improvement in medical care as well as availability of new therapies [4][5]. Various studies have reported an increase in occurrence of brain and CNS cancers in elderly population of Western countries [6][7][8]. The risk of this cancer increases with certain genetic factors and ionizing radiation exposure, while allergic conditions seem to decrease the risk [9] [10]. European countries have the highest rates of brain and CNS cancer while Asian countries have the lowest. This difference can be partially linked to the difference in genetic background of the population in these continents [11]. The recent World Health Organization report (2016) includes classification of brain tumors and mentions molecular markers to determine subclasses of gliomas and medulloblastomas. But still very few markers are sufficiently characterized to impact the clinical practice in patients with CNS cancers. Through this paper we aim to predict the genes that are significantly up-regulated and the corresponding pathways that play a significant role in CNS cancers.

2. RESEARCH METHOD

2.1. Brain and CNS cancer microarray datasets

In this study, four brain and CNS cancer datasets (Table 1.) were selected from Oncomine [12] database. These datasets contained a differential analysis of tumor and normal samples, experiment type was mRNA and number of samples in both tumor and normal category was more than one. Oncomine [12] currently contains 715 datasets (Oncomine _Research Edition) and it is one of the most comprehensive cancer-specific database. The major advantage of using this database is that prior to inclusion in Oncomine database, the microarray datasets obtained from public resources such as Stanford Microarray Database and the NCBI Gene Expression Omnibus or literature sources are reviewed by a panel of experts to ensure that they meet certain quality standards [13].

Table 1. Brain and CNS cancer microarray datasets included in the study

Dataset name*	Accession id	Genes**	Platform	Dataset summary*
Bredel Brain 2	GSE2223	14386	Platform not pre-defined in Oncomine	Fifty (50) brain CNS carcinoma samples and 4 normal brain samples were analyzed on cDNA microarrays. Sample data includes disease type.
Lee Brain	GSE4536	19574	Human Genome U133 Plus 2.0 Array	Ninety-eight (98) brain and CNS cancer (26 of which are cell lines) and 3 normal neural stem cell samples were analyzed on Affymetrix U133 Plus 2.0 microarrays. Sample data includes type, growth media, cell line name, and passage.
Liang Brain	GSE4058	9957	Platform not pre-defined in Oncomine	Thirty (30) glioblastoma, 3 mixed astrocytoma-oligodendroglioma, 2 oligodendroglioma, 2 normal brain, and 1 normal cerebellum sample were analyzed on cDNA microarrays. Sample data includes type, age, location, primary/recurrent, sex, and survival.
Murat Brain	GSE7696	19574	Human Genome U133 Plus 2.0 Array	Eighty (80) glioblastoma samples and 4 normal brain tissue samples were analyzed. Sample data includes sample recurrence status, age, sex, MGMT methylation status, and survival status.

*As identified by the Oncomine database

**Number of genes probed

2.2. Initial screening of microarray datasets

Each dataset obtained from the Oncomine database contained more than one type of brain and CNS cancer. Hence for each dataset, sub-datasets were created based on the type of cancers it contained, location or treatments.

2.3. Creation of sub-datasets

Bredel Brain 2 dataset consists specimens belonging to astrocytic, glioblastomas and oligodendroglial types of brain tumors. Hence this dataset was divided on the basis of these tumors and their corresponding subparts. Six sub-datasets were created and the list is provided in Bredel Brain 2 sheet of Supplementary File1. Lee Brain dataset was divided based on the region from where the tumor was extracted. For instance, a separate sub-dataset was created for each tumor cell line. Thirteen sub-datasets were created and the list is provided in Lee Brain sheet of Supplementary File1. Liang Brain dataset contained data from 5 different platforms. The normal samples belonged to GEO Platform GPL2935 and on comparison we found that platforms GPL182, GPL2778 and GPL2935 have identical genes whereas GPL2648 and GPL3010 have different genes. So, the samples belonging to the latter two platforms were removed from further analysis. The sub-datasets were created on the basis of different types of tumors. The detailed information is provided in Liang Brain sheet of Supplementary File1. In Murat Brain, the dataset was divided to three sub-datasets based on the various treatments provided. The sub-datasets are listed in Murat Brain sheet of Supplementary File1.

2.4. Identification of up-regulated genes in each sub dataset

For each sub-dataset, Significance Analysis of Microarray (SAM) [14] was performed using the software Multiple experiment Viewer (MeV) [15]. The details of the output generated are provided in the Bredel Brain 2 sheet, Lee Brain sheet, Liang Brain sheet and Murat Brain sheet of Supplementary File2 for the sub-datasets of Bredel Brain 2, Lee Brain, Liang Brain and Murat Brain respectively.

2.5. Z score calculation

Z score is calculated for all possible pairs of sub-datasets. The formula used is

$$z = \frac{R_{observed} - R_{expected}}{\sigma} = \frac{R_{observed} - n_B P_A}{\sqrt{n_B P_A (1 - P_A)}}$$

where R_{obs} is the number of significant genes in both datasets A and B, n_B is the number of genes in dataset B and P_A is the probability of gene being significantly upregulated in A [16].

2.6. Obtaining ranked list of up-regulated genes

For the identification of differentially expressed genes across multiple datasets, 'rank product' method was used. It is a non-parametric method implemented in the RankProd package [17] [18]. RankProd is a biologically intuitive algorithm and statistically rigorous, which has been shown to be robust against noise in microarray data [19] [20]. This algorithm is shown to have higher specificity and sensitivity as compared to other types of meta-analytic tools for microarrays [17]. Based on the conservative estimation of the percentage of false positive predictions (pfp), a list of up-regulated genes is created. As recommended, a pfp value of 0.15 was used to set the threshold for genes that are significantly up-regulated [16].

3. RESULTS AND ANALYSIS

3.1. Congruency between microarray datasets

Since the dataset MBSP3 contains 0 significantly up-regulated genes hence it has not been considered in Z score calculation. The datasets which had pairwise z score > 1.96 were considered for further analysis. The selected datasets are: BB2SP1, BB2SP2, BB2SP3, BB2SP4, BB2SP5, BB2SP6, LBSP1, LBSP2, LBSP3, LBSP4, LBSP5, LBSP6, LBSP7, LBSP8, LBSP9, LBSP10, LBSP11, LBSP12, LBSP13. The sub-datasets belonging to Liang Brain and Murat Brain were thus excluded from further analysis. In Figure 1., the cells in red color represent values of z-score > 1.96 and in brown represent values < 1.96 .

3.2. Identification of upregulated genes

Combining the result of SAM for all the selected sub-datasets, 3861 significant genes were obtained. To improve the result of SAM, the significant genes were further analyzed using RankProd with pfp threshold of < 0.15. 271 genes were found to be significantly up-regulated using RankProd program and after removing duplicates, there were 130 significant genes. A complete ranked list of significantly up-regulated genes has been provided in Supplementary File3. Top 26 significantly up-regulated genes are listed in Table2.

Table 2. List of top 26 significantly up-regulated genes

Gene symbol	Description
ALDH1A3	aldehyde dehydrogenase 1 family, member A3
ANXA2	annexin A2
C1QC	complement component 1, q subcomponent, C chain
C3	complement component 3
CHI3L2	Chitinase 3-like 2: chitinase precursor
COL1A1	collagen, type I, alpha 1
COL6A2	collagen, type VI, alpha 2
CTGF	connective tissue growth factor
CXCL14	chemokine (C-X-C motif) ligand 14
DDX3Y	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked
GBP1	guanylate binding protein 1, interferon-inducible, 67kDa
GFAP	glial fibrillary acidic protein
HBA1 /// HBA2	hemoglobin, alpha 1 /// hemoglobin, alpha 2
HBB	hemoglobin, beta
HLA-DPA1	major histocompatibility complex, class II, DP alpha 1
HLA-DRA	major histocompatibility complex, class II, DR alpha
IGFBP7	insulin-like growth factor binding protein 7
IL13RA2	interleukin 13 receptor, alpha 2
MAGEA6	melanoma antigen family A, 6
MEOX2	mesenchyme homeobox 2
MGST1	microsomal glutathione S-transferase 1
RBMS1	RNA binding motif, single stranded interacting protein 1
RGS1	regulator of G-protein signaling 1
SERPINA3	Serine (or cysteine) proteinase inhibitor, clade A (alpha-1 antiproteinase, antitrypsin), member 3
SRGN	Serglycin
TGFBI	transforming growth factor, beta-induced, 68kDa

To enhance the validity of our findings, we searched for experimental works which prove the involvement of these genes in Brain and CNS cancer. It has been reported that the prognosis in primary Glioblastoma multiformes is highly correlated with ALDH1A3 promoter methylation [21] [22]. Studies have shown that ANXA2 is significantly over-expressed in glioma samples as compared to normal brain samples [23]. This gene regulates angiogenesis and invasion of malignant gliomas as well [24]. C3 gene expression is enhanced in time as well as dose dependent manner through interleukin-1 (IL-1 beta) as well as tumor necrosis factor-alpha (TNF-alpha) in the astrogloma cell line D54-MG [25]. CHI3L2 gene encodes a protein YKL-40 which serves as a prognosticator for cancer and this gene is found to be highly up-regulated in glioma [26]. COL1A1 is over-expressed in pilomyxoid astrocytomas as well as pilocytic astrocytomas [27]. COL1A1 and COL6A2 are over-expressed in both primary as well as meta-static brain tumors [28]. CTGF plays a significant role in gliomas. Studies have suggested that gene expression level of this gene can have prognostic significance [29]. GBP1 gene

over-expression enhances glioma cell invasion [30]. GFAP is highly expressed in tumor cells and hence can be used in the detection of tumor recurrence [31]. HLA-DPA1 and HLA-DRA are the important genes involved in Glioblastoma multiforme [32]. IGFBP7 is highly over-expressed in Glioblastoma multiforme [33] and Pilocytic astrocytomas [34]. IL13RA2 is over-expressed in highly invasive glioblastoma multiforme [35]. SERPINA3 gene has been found to be involved in brain metastasis [36]. TGFBI is methylated in neuroblastoma [37]. The remaining top significant genes which have not been associated with brain and CNS cancer till now are: C1QC, CXCL14, DDX3Y, HBA1 or HBA2, HBB, MAGEA6, MEOX2, MGST1, RBMS1, RGS1, SRGN and TGFBI.

3.3. Functional analysis of up-regulated genes

The significant genes obtained through RankProd were entered as an input to the tool “Database for Annotation, Visualization and Integrated Discovery”, DAVID [38] [39] which is a functional annotation tool. 123 genes matched out of 130 genes given as an input. Functional Annotation for these 123 genes was obtained using the tool. Here we present the results of some of the functional annotation categories: Reactome Pathway, OMIM Diseases, KEGG Pathway and Genetic Association DB Disease Class. (Figures2-5) The complete results of Functional Annotation can be found in the Supplementary File-DAVID: Category-based sheet contains the information on the basis of various categories that were selected by us during the execution, Cluster-based sheet contains the information for the different annotation clusters generated by DAVID [38] [39] and Gene-based sheet contains the functional annotation information for each gene.

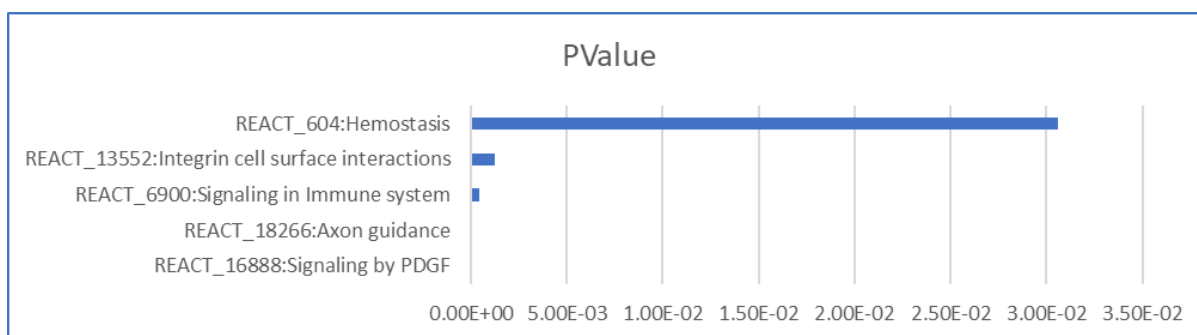


Figure 2. Enriched Reactome Pathway obtained through DAVID [38] [39]

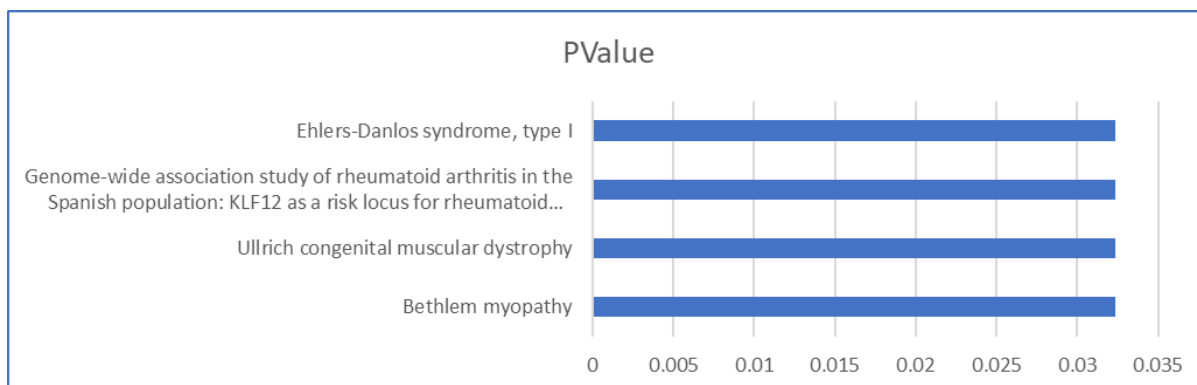


Figure 3. Enriched OMIM Diseases obtained through DAVID [38] [39].

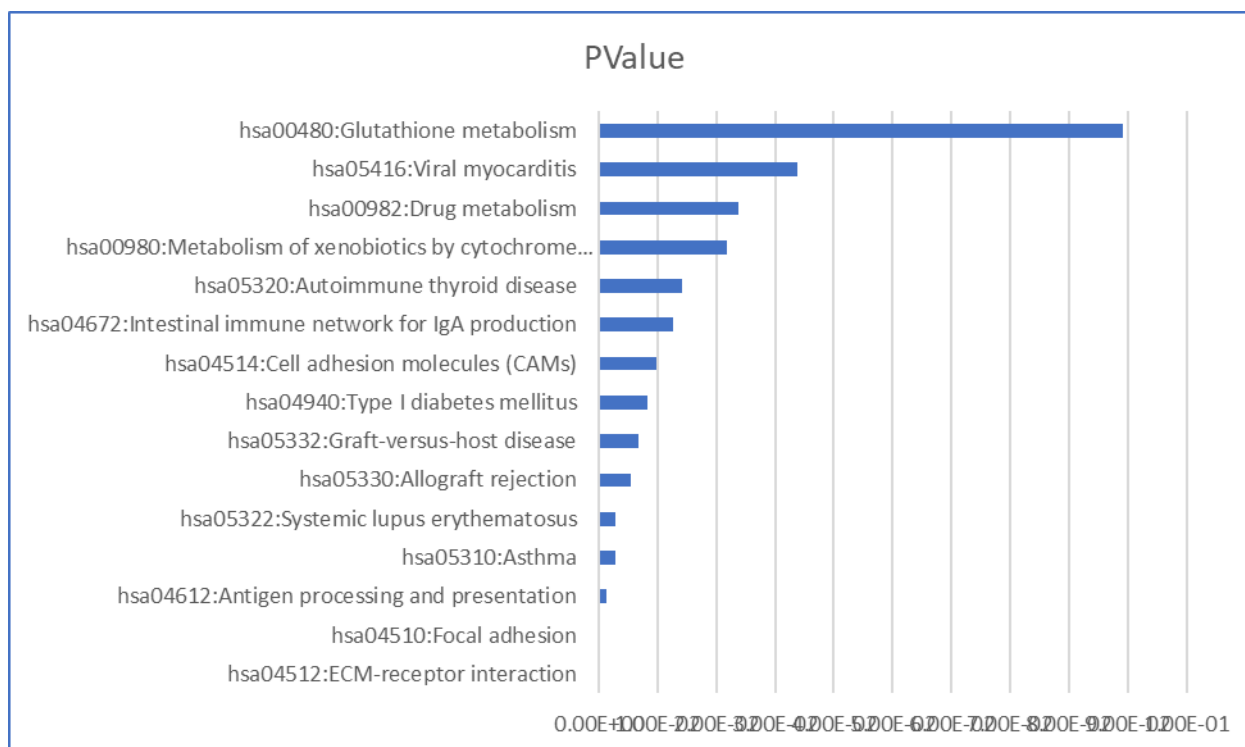


Figure 4. Enriched KEGG PATHWAY obtained through DAVID [38] [39].

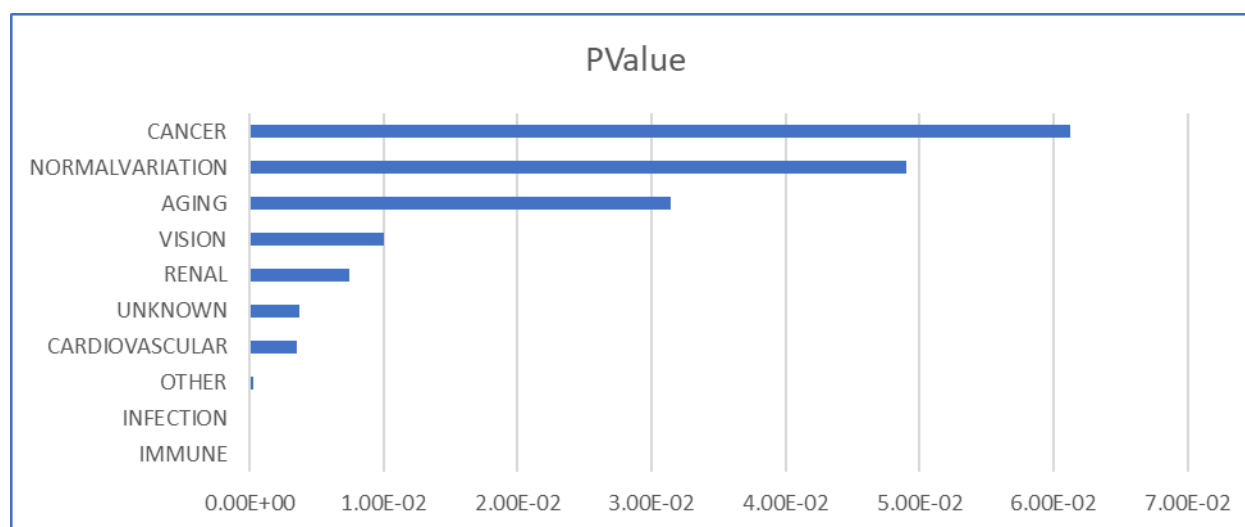


Figure 5. Enriched GENETIC ASSOCIATION DB DISEASE CLASS obtained through DAVID [38] [39]

3.4. Identification of significant pathways

Using online tool of Genemania [40] a network was created with 127 genes as an input. The seven types of network which were created are: Co-expression, Co-localization, Genetic Interaction, Pathway, Physical Interaction, Predicted and Shared Protein Domains network. The predicted pathway contains only one interaction and hence was removed from further analysis. The network properties for each network was determined using Network analyzer tool of Cytoscape [41]. They are mentioned in the tables below. (Tables 3-8)

Table 3. Network Properties of Co-expression network

Properties	Co-expression network
Clustering Coefficient	0.177
Connected components	2
Network diameter	6
Network radius	1
Shortest paths	4288 (29%)
Characteristic path length	2.091
Average number of neighbors	16.639
Number of nodes	122
Network density	0
Isolated nodes	0
Number of self-loops	0
Multi-edge node pairs	322
Analysis time (sec)	0.085

Table 4. Network Properties of Co-localization network

Properties	Co-localization network
Clustering Coefficient	0.208
Connected components	3
Network diameter	4
Network radius	1
Shortest paths	319 (5%)
Characteristic path length	1.639
Average number of neighbors	4.184
Number of nodes	76
Network density	0
Isolated nodes	0
Number of self-loops	0
Multi-edge node pairs	7
Analysis time (sec)	0.303

Table 5. Network Properties of Genetic Interaction network

Properties	Genetic interaction
Clustering Coefficient	0.08
Connected components	1
Network diameter	6
Network radius	1
Shortest paths	1951 (20%)
Characteristic path length	2.317
Average number of neighbors	7.204
Number of nodes	98
Network density	0
Isolated nodes	0
Number of self-loops	0
Multi-edge node pairs	0
Analysis time (sec)	0.067

Table 6. Network Properties of Pathway network

Properties	Pathway
Clustering Coefficient	0.141
Connected components	5
Network diameter	2
Network radius	1
Shortest paths	30 (4%)
Characteristic path length	1.133
Average number of neighbors	2
Number of nodes	26
Network density	0
Isolated nodes	0
Number of self-loops	0
Multi-edge node pairs	1
Analysis time (sec)	0.035

Table7: Network Properties of Physical Interaction network

Properties	Physical interaction
Clustering Coefficient	0.024
Connected components	8
Network diameter	4
Network radius	1
Shortest paths	87 (6%)
Characteristic path length	1.77
Average number of neighbors	2.222
Number of nodes	36
Network density	0
Isolated nodes	0
Number of self-loops	0
Multi-edge node pairs	35
Analysis time (sec)	0.301

Table8: Network Properties of Shared Protein Domains network

Properties	Shared protein domains
Clustering Coefficient	0.236
Connected components	18
Network diameter	3
Network radius	1
Shortest paths	81 (2%)
Characteristic path length	1.16
Average number of neighbors	2.1875
Number of nodes	64
Network density	0
Isolated nodes	0
Number of self-loops	0
Multi-edge node pairs	0
Analysis time (sec)	0.584

The edge details of Co-expression, Co-localization, Genetic Interaction, Pathway, Physical Interaction, Predicted and Shared Protein Domains network are provided in the Co-expression sheet, Co-localization sheet, Genetic Interaction sheet, Pathway sheet, Physical Interaction sheet, Predicted sheet and Shared Protein Domains sheet of Supplementary File4 respectively. The node details of Co-expression, Co-localization, Genetic Interaction, Pathway, Physical Interaction, Predicted and Shared Protein Domains network are provided in the Co-expression sheet, Co-localization sheet, Genetic Interaction sheet, Pathway sheet, Physical Interaction sheet, Predicted sheet and Shared Protein Domains sheet of Supplementary File5 respectively.

For each network, significant pathways were determined. The methodology used is as follows: Firstly, for each gene present in the network, we determined the KEGG pathways that are associated with that gene. The obtained pathways are assigned the Genemania [40] score of the corresponding gene. Now, for each pathway present in the network, combined score and total count of the pathway in the network is calculated. The enriched pathways are then classified on the basis of the pathway class of KEGG. The Supplementary File-SIGNIFICANT PATHWAYS contains the graphical representation for enriched pathways present in each network. The complete list of the pathways can be found in the Supplementary File-Pathways. Here we present the topmost significant pathways relevant to our discussion.

In co-expression network for brain and CNS cancer, top five significant pathways are Retinol metabolism; cAMP signaling pathway; Notch signaling pathway; Prion diseases and Signaling pathways regulating pluripotency of stem cells. For co-localization network, top five significant pathways are: Signaling pathways regulating pluripotency of stem cells; Notch signaling pathway; Intestinal immune network for IgA production; RIG-I-like receptor signaling pathway and Sphingolipid signaling pathway. In genetic interaction network top five significant pathways are: Prion diseases; Retinol metabolism; Notch signaling pathway; cAMP signaling pathway and Signaling pathways regulating pluripotency of stem cells. For pathway network, top five significant pathways are: Regulation of actin cytoskeleton; Gap junction; MAPK signaling pathway; Ras signaling pathway and Phospholipase D signaling pathway. In physical interaction network, top five significant pathways are: Gap junction; Phospholipase D signaling pathway; Glioma; Melanoma and Choline metabolism in cancer. In shared protein domains network, the top five significant pathways are: cAMP signaling pathway; Notch signaling pathway; Prion diseases; Regulation of actin cytoskeleton and Gap junction.

4. CONCLUSION

It can be concluded that meta-analysis of microarray datasets yields more comprehensive and reliable results as compared to a single dataset because it has generalizability and increased statistical power. On creating different types of networks of significantly up-regulated genes, various pathways that are possibly enriched in Brain and CNS cancer have been obtained.

SUPPLEMENTARY FILES

1. Supplementary File1
2. Supplementary File2
3. Supplementary File3
4. Supplementary File4
5. Supplementary File5
6. Supplementary File-DAVID
7. Supplementary File-Pathways
8. Supplementary File-SIGNIFICANT PATHWAYS
9. ZScore-ExcelSheet

REFERENCES

- [1] Louis DN, Perry A, Reifenberger G, Von Deimling A, Figarella-Branger D, Cavenee WK, Ohgaki H, Wiestler OD, Kleihues P, Ellison DW. The 2016 World Health Organization classification of tumors of the central nervous system: a summary. *Acta neuropathologica*. 2016 Jun 1;131(6):803-20.
- [2] Ferlay J, Soerjomataram I, Ervik M, et al. GLOBOCAN 2012 v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11 [Internet]. Lyon, France: International Agency for Research on Cancer; 2013. Available from: <http://globocan.iarc.fr>. Accessed May 30, 2016.
- [3] Miranda-Filho A, Piñeros M, Soerjomataram I, Deltour I, Bray F. Cancers of the brain and CNS: global patterns and trends in incidence. *Neuro-oncology*. 2016 Aug 29;19(2):270-80.
- [4] Fisher JL, Schwartzbaum JA, Wrensch M, Wiemels JL. Epidemiology of brain tumors. *Neurologic clinics*. 2007 Nov 30;25(4):867-90.
- [5] Ho VK, Reijneveld JC, Enting RH, Bienfait HP, Robe P, Baumert BG, Visser O. Changing incidence and improved survival of gliomas. *European journal of cancer*. 2014 Sep 30;50(13):2309-18.
- [6] Davis D, Hoel D, Percy C, Ahlbom A, Schwartz J. Is brain cancer mortality increasing in industrial countries?. *Annals of the New York Academy of Sciences*. 1990 Nov 1;609(1):191-204.
- [7] Grieg NH, Ries LG, Yancik R, Rapoport SI. Increasing annual incidence of primary malignant brain tumors in the elderly. *JNCI: Journal of the National Cancer Institute*. 1990 Oct 17;82(20):1621-4.
- [8] Hoffman S, Propp JM, McCarthy BJ. Temporal trends in incidence of primary brain tumors in the United States, 1985-1999. *Neuro-oncology*. 2006 Jan 1;8(1):27-37.
- [9] Wigertz A, Lönn S, Schwartzbaum J, Hall P, Auvinen A, Christensen HC, Johansen C, Klæboe L, Salminen T, Schoemaker MJ, Swerdlow AJ. Allergic conditions and brain tumor risk. *American journal of epidemiology*. 2007 Jul 23;166(8):941-50.
- [10] Cahoon EK, Inskip PD, Gridley G, Brenner AV. Immune-related conditions and subsequent risk of brain cancer in a cohort of 4.5 million male US veterans. *British journal of cancer*. 2014 Apr 1;110(7):1825.
- [11] Miranda-Filho A, Piñeros M, Soerjomataram I, Deltour I, Bray F. Cancers of the brain and CNS: global patterns and trends in incidence. *Neuro-oncology*. 2016 Aug 29;19(2):270-80.

- [12] Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D, Barrette T, Pander A, Chinnaiyan AM. ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia*. 2004 Jan 1;6(1):1-6.
- [13] Rhodes DR, Kalyana-Sundaram S, Mahavisno V, Varambally R, Yu J, Briggs BB, Barrette TR, Anstet MJ, Kincaid-Beal C, Kulkarni P, Varambally S. Oncomine 3.0: genes, pathways, and networks in a collection of 18,000 cancer gene expression profiles. *Neoplasia*. 2007 Feb 1;9(2):166-80.
- [14] Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proceedings of the National Academy of Sciences*. 2001 Apr 24;98(9):5116-21.
- [15] Saeed AI, Sharov V, White J, Li J, Liang W, Bhagabati N, Braisted J, Klapa M, Currier T, Thiagarajan M, Sturn A. TM4: a free, open-source system for microarray data management and analysis. *Biotechniques*. 2003 Feb;34(2):374.
- [16] Goonesekere NC, Wang X, Ludwig L, Guda C. A meta analysis of pancreatic microarray datasets yields new targets as cancer genes and biomarkers. *PloS one*. 2014 Apr 16;9(4):e93046.
- [17] Hong F, Breitling R, McEntee CW, Wittner BS, Nemhauser JL, Chory J. RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics*. 2006 Sep 18;22(22):2825-7.
- [18] Laing E, Smith CP. RankProdIt: A web-interactive Rank Products analysis tool. *BMC research notes*. 2010 Aug 6;3(1):221.
- [19] Breitling R, Armengaud P, Amtmann A, Herzyk P. Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. *FEBS letters*. 2004 Aug 27;573(1-3):83-92.
- [20] Dudley JT, Tibshirani R, Deshpande T, Butte AJ. Disease signatures are robust across tissues and experiments. *Molecular systems biology*. 2009 Jan 1;5(1):307.
- [21] Ni W, Luo L, Ping Z, Yuan HP, Zhao XD, Xu W. Prognostic value of ALDH1A3 promoter methylation in glioblastoma: a single center experience in Western China. *Asian Pacific Journal of Cancer Prevention*. 2015;16(2):591-4.
- [22] Han J, Yang L, Puri RK. Analysis of target genes induced by IL-13 cytotoxin in human glioblastoma cells. *Journal of neuro-oncology*. 2005 Mar 1;72(1):35-46.
- [23] Barnum SR, Jones JL, Benveniste EN. Interleukin-1 and tumor necrosis factor- α mediated regulation of C3 gene expression in human astrogloma cells. *Glia*. 1993 Mar 1;7(3):225-36.
- [24] Onishi M, Ichikawa T, Kurozumi K, Inoue S, Maruo T, Otani Y, Fujii K, Ishida J, Shimazu Y, Yoshida K, Michiue H. Annexin A2 regulates angiogenesis and invasion phenotypes of malignant glioma. *Brain tumor pathology*. 2015 Jul 1;32(3):184-94.
- [25] Barnum SR, Jones JL, Benveniste EN. Interleukin-1 and tumor necrosis factor- α mediated regulation of C3 gene expression in human astrogloma cells. *Glia*. 1993 Mar 1;7(3):225-36.
- [26] Saidi A, Javerzat S, Bellahcene A, De Vos J, Bello L, Castronovo V, Deprez M, Loiseau H, Bikfalvi A, Hagedorn M. Experimental anti-angiogenesis causes upregulation of genes associated with poor survival in glioblastoma. *International journal of cancer*. 2008 May 15;122(10):2187-98.
- [27] Kleinschmidt-DeMasters BK, Donson AM, Vogel H, Foreman NK. Pilomyxoid astrocytoma (PMA) shows significant differences in gene expression vs. pilocytic astrocytoma (PA) and variable tendency toward maturation to PA. *Brain Pathology*. 2015 Jul 1;25(4):429-40.
- [28] Liu Y, Carson-Walter EB, Cooper A, Winans BN, Johnson MD, Walter KA. Vascular gene expression patterns are conserved in primary and metastatic brain tumors. *Journal of neuro-oncology*. 2010 Aug 1;99(1):13-24.
- [29] Xie D, Yin D, Wang HJ, Liu GT, Elashoff R, Black K, Koeffler HP. Levels of expression of CYR61 and CTGF are prognostic for tumor progression and survival of individuals with gliomas. *Clinical Cancer Research*. 2004 Mar 15;10(6):2072-81.
- [30] Li M, Mukasa A, Inda MD, Zhang J, Chin L, Cavenee W, Furnari F. Guanylate binding protein 1 is a novel effector of EGFR-driven invasion in glioblastoma. *Journal of Experimental Medicine*. 2011 Dec 19;208(13):2657-73.
- [31] Kiviniemi A, Gardberg M, Frantzén J, Parkkola R, Vuorinen V, Pesola M, Minn H. Serum levels of GFAP and EGFR in primary and recurrent high-grade gliomas: correlation to tumor volume, molecular markers, and progression-free survival. *Journal of neuro-oncology*. 2015 Sep 1;124(2):237-45.
- [32] Wang L, Wei B, Hu G, Wang L, Bi M, Sun Z, Jin Y. Screening of differentially expressed genes associated with human glioblastoma and functional analysis using a DNA microarray. *Molecular medicine reports*. 2015 Aug 1;12(2):1991-6.
- [33] Pen A, Moreno MJ, Martin J, Stanimirovic DB. Molecular markers of extracellular matrix remodeling in glioblastoma vessels: Microarray study of laser-captured glioblastoma vessels. *Glia*. 2007 Apr 15;55(6):559-72.

- [34] Jones TA, Jeyapalan JN, Forsheew T, Tatevossian RG, Lawson AR, Patel SN, Doctor GT, Mumin MA, Picker SR, Phipps KP, Michalski A. Molecular analysis of pediatric brain tumors identifies microRNAs in pilocytic astrocytomas that target the MAPK and NF- κ B pathways. *Acta neuropathologica communications*. 2015 Dec 18;3(1):86.
- [35] Marie SK, Okamoto OK, Uno M, Hasegawa AP, Oba- Shinjo SM, Cohen T, Camargo AA, Kosoy A, Carlotti CG, Toledo S, Moreira- Filho CA. Maternal embryonic leucine zipper kinase transcript abundance correlates with malignancy grade in human astrocytomas. *International journal of cancer*. 2008 Feb 15;122(4):807-15.
- [36] Sevenich L, Bowman RL, Mason SD, Quail DF, Rapaport F, Elie BT, Brogi E, Brastianos PK, Hahn WC, Holsinger LJ, Massagué J. Analysis of tumor-and stroma-supplied proteolytic networks reveals a brain metastasis-promoting role for cathepsin S. *Nature cell biology*. 2014 Sep;16(9):876.
- [37] Carén H, Djos A, Nethander M, Sjöberg RM, Kogner P, Enström C, Nilsson S, Martinsson T. Identification of epigenetically regulated genes that predict patient outcome in neuroblastoma. *BMC cancer*. 2011 Feb 11;11(1):66.
- [38] Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic acids research*. 2008 Nov 25;37(1):1-3.
- [39] Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, Franz M, Grouios C, Kazi F, Lopes CT, Maitland A. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic acids research*. 2010 Jun 21;38(suppl_2):W214-20.
- [40] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*. 2003 Nov 1;13(11):2498-504.