Gliosarcoma Protein - Protein Interaction Network Analysis and Gene Ontology

Mostafa Rezaei Tavirani,1, *) Vahid Mansouri,1 Sina Rezaei Tavirani,1 Saeed Hesami Tackallou,2 and Mohammad Rostami - Nejad3

1Proteomics Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran
2Islamic Azad University, Central Tehran Branch, Tehran, Iran
3Basic and Molecular Epidemiology of Gastrointestinal Disorders Research Center, Research Institute for Gastroenterology and Liver Diseases, Shahid Beheshti University of Medical Sciences, Tehran, Iran

*) Corresponding author: Mostafa Rezaei Tavirani, Proteomics Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Tel: +98-9122650447, E-mail: tavirany@yahoo.com

Received 2017 December 30; Revised 2018 January 22; Accepted 2018 March 06.

Abstract

Background: Gliosarcoma (GS) is a rare primary neoplasm of the central nervous system. It is a subtype of glioblastoma and has a biphasic pattern consisting of glial and malignant mesenchymal elements. Its onset is between the fourth and sixth decade of life.

Objectives: Since protein - protein interaction (PPI) network analysis can provide useful information about molecular aspects of diseases, the aim of this study is GS protein analysis via PPI network and gene ontology assessment.

Methods: The related genes to GS were gathered from STRING DB and organized in the interacted network by Cytoscape software version 3.6.0. The network was analyzed based on topological parameters and the central nodes were introduced. The significant clusters were identified by ClusterONE and the cluster included more key genes enriched via gene ontology by ClueGO.

Results: Nine crucial genes including TP53, EGFR, PTEN, EGR1, VEGFA, HSP90AA1, IL2, KNG1, and HSP90AB1 were introduced as related key genes to GS. Two significant clusters contain most of central genes. Twenty - one elements of cluster - 1, which included 7 key genes, were enriched via gen ontology and 115 related terms were determined and discussed.

Conclusions: The nine introduced central genes may play main roles in pathology of GS. However, experimental investigation is proposed to validate the findings.

Keywords: Gliosarcoma, Gene Ontology, Cluster

1. Background

One of the rare primary neoplasms of the central nervous system is gliosarcoma (GS), a subtype of glioblastoma (GBM) (1). GS was first reported by Strobe in 1895. It has a biphasic pattern consisting of glial (anaplastic astrocytes) and malignant mesenchymal elements. However, the monoclonal or biclonal origin of GS biphasic nature is still subject to debate (2). The onset of GS, as a rare neoplasm, is between the fourth and sixth decade of life (3), and the male/female ratio of GS is 1.8/1 (4). Treatment consists of surgical resection of the tumor followed by external radiotherapy or chemotherapy in some cases (4, 5). Authors represented different genes involved in the disease as p53 mutant expression (6, 7). Glial fibrillary acidic protein also reported as an diagnostic protein embedded in glioblastoma cells (8, 9). Scientists assessed the genetic profile of GS with mutation in P53, PTEN, and deletion of P6 with CDK4 amplification (10). Douglas et al. from Stanford compared systematic genes variations in 60 different cancer cell lines. Their article published in Genetics Nature represented common genes involved in gliosarcoma and other cancer cells (11). On the other hand, the mesenchymal component of GSs can present differentiation along several lineages as fibroblasts and chondroblasts, etc. (12). Investigation indicates that EGFR amplification is much lower in GS than GMB (13). Identifying the genes and proteins involved in the development of GS or the other types of cancer can effectively determine their treatment (14). From the perspective of systems biology, the connection between proteins involved in the disease is important (15). PPI network analysis of diseases has attracted attention of medical and biological scientist. In this approach, examination of the interaction between genes involved in the disease could lead us to improve the diagnosis and treatment of patients (16-20). In PPI network analysis, the related genes to the disease are gathered and organized in an...
integrative structure as a interactome (16, 21). The assessment of topological properties of the network, including central parameters such as degree and betweenness centrality provides useful information about molecular mechanism of disease onset and pathology (22-24). Introducing selected genes among large number of query genes can lead to specific biomarker panel related to the disease (25). In this study, we aim at identifying and analyzing genes interaction involved in GS disease. It may be conducted to introduce a biomarker panel related to GS.

2. Methods

A number of 200 genes related to gliosarcoma were searched from disease query of String data bank (String is one of Cytoscape applications), but only 106 genes were introduced. The genes were analyzed for network construction by Cytoscape software version 3.6.0 (26). The main connected component of PPI network includes 78 nodes and 269 undirected edges were layout by degree values. The top 20% of the nodes based on degree values were selected as hub genes. The bottleneck nodes were identified considering top 20% of the nodes based on betweenness centrality. Similarly, top 20% of the nodes based on closeness centrality and stress values were selected as central genes. The hub nodes, which are bottleneck node, are introduced as hub-bottleneck genes. The nodes that were hub-bottleneck genes and their closeness and stress ranks were less than 10 and selected as crucial nodes. The main connected component (the constructed network) was analyzed by Cluster ONE plugin of Cytoscape software (27). The significant clusters were introduced and discussed. Finally, the main identified clusters were analyzed via gene ontology method. Gene ontology for the nodes of cluster -1 indicates that 115 biological terms including biological processes, cellular components, molecular function, and biochemical pathways (from KEGG data base) are related to this cluster (Figure 5). The term was included if at least 2 genes and 5% gene/term contribution were considered. The analysis of a large number of the terms (115 terms) is a difficult process; therefore, as it is shown in Figure 6, the terms are organized in 5 groups.

3. Results

The 106 related genes of gliosarcoma for human species were retrieved from disease query of String database (confides 0.40 score). The genes were organized in a PPI network, including a main connected component (a network constructed with 78 nodes and 296 undirected edges) and 28 isolated nodes. The network is presented in Figure 1. The network was analyzed and for more resolution, the 20% of top nodes (equal to 16 nodes) based on degree values were selected as hub genes. Similarly, 16 top nodes based on betweenness values were identified as bottleneck genes. Two separated groups (each of them includes 16 top genes, corresponding to the values of closeness centrality and stress) were determined as central nodes. Therefore, 64 nodes were organized in 4 groups. If an individual node was not hub or bottleneck genes and was not found in the one of the other two groups, it would be omitted from study. Finally, 9 crucial genes of the analyzed network were identified and presented in Table 1. Connections between the crucial genes were illustrated in Figure 2 as a sub network. This sub network was resulted via the deletion of the other elements of gliosarcoma network. Since members of a cluster (cluster plays a role as an integrative part of a network) are involved in closed functional terms, cluster analysis of the network limited to the presence of at least 10 nodes in a cluster was performed. The cluster analysis revealed that there are 2 clusters (cluster 1 and 2) characterized by different quality and P value related to the constructed network. These clusters in interaction with the other parts of the network are presented in the Figures 3 and 4. Since cluster -1 contains more crucial genes, this cluster was analyzed via gene ontology method. Gene ontology for the nodes of cluster -1 indicates that 115 biological terms including biological processes, cellular components, molecular function, and biochemical pathways (from KEGG data base) are related to this cluster (Figure 5). The term was included if at least 2 genes and 5% gene/term contribution were considered. The analysis of a large number of the terms (115 terms) is a difficult process; therefore, as it is shown in Figure 6, the terms are organized in 5 groups.

4. Discussion

Gliosarcoma is a disease, which is studied by several molecular approaches such as genetics molecular biology and metabolite analysis (10, 29, 30). Therefore, there are a number of genes that are related to this disease. It is an essential need to rank and screen the genes for introducing the important elements among them. PPI network analysis is the right approach in this regard. The constructed network in this study contains 78 nodes and 296 edges. It seems that compared to other diseases PPI network, this network is a small network. Topology analysis of the network based on central parameters revealed that there are 9 crucial genes related to GS. Here, we tried to represent an overview of the significant roles of identified crucial genes of network in GC pathogenesis and development.

As it is known, TP53 is the most famous gene related to cancers. Rarely, it is discussed about the cancer without TP53 participation. TP53 mutations accompanied with the other genes such as CDKN2A is reported in GS and it is discussed in details (31). In the another report, the role of TP53 mutation beside CD34 in the GS patients was studied and its significant role in the disease was emphasized (32). In addition, the amplification of EGFR in correlation with...
Figure 1. PPI network of gliosarcoma is constructed with 78 nodes and 296 undirected edges. The network is layout by degree value. The bigger size of a node corresponds to bigger value of degree. Color from gray to blue refers to increment of degree value.

Table 1. The 9 crucial nodes related to gliosarcoma are presented. All nodes (except HSP90AB1) are hub – bottleneck genes. HSP90AB1 was not included in hub – genes, but is identified as a bottleneck node. D, BC, CC, S, and DS are degree, betweenness centrality, closeness centrality, stress, and disease score, respectively.

| R | Name                  | Description                                  | D   | BC  | CC  | S    | DS  |
|---|-----------------------|----------------------------------------------|-----|-----|-----|------|-----|
| 1 | TP53                  | Tumor protein p53                            | 42  | 0.31| 0.63| 4896 | 2.58|
| 2 | EGFR                  | Epidermal growth factor receptor             | 35  | 0.26| 0.61| 4104 | 1.51|
| 3 | PTEN                  | Phosphatase and tensin homolog               | 25  | 0.07| 0.53| 1448 | 1.32|
| 4 | EGR1                  | Early growth response 1                      | 22  | 0.30| 0.54| 2044 | 0.56|
| 5 | VEGFA                 | Vascular endothelial growth factor A         | 21  | 0.03| 0.53| 884  | 0.61|
| 6 | HSP90AA1              | Heat shock protein 90kDa alpha (cytosolic); class A member 1 | 21  | 0.33| 0.55| 3054 | 0.63|
| 7 | IL2                   | Interleukin 2                                | 13  | 0.02| 0.48| 566  | 1.33|
| 8 | KNG1                  | Kininogen 1                                  | 12  | 0.12| 0.48| 2036 | 0.55|
| 9 | HSP90AB1              | Heat shock protein 90kDa alpha (cytosolic); class B member 1 | 11  | 0.3  | 0.49| 566  | 1.46|
PTEN in the glioblastoma and GS patients were studied and confirmed (33). Phosphatidylinositol-3'-kinase (PI3k) - dependent activation of protein kinase B (Akt) signaling is the inhibitory role of PTEN as a tumor suppressor gene (34). Investigation indicates that EGFR overexpression, occurrence mutations in PTEN, and PI6 deletions accompanied with MDM2 amplification are the typical events in this disease (35). EGRI is responsible for synaptic plasticity and neuronal activity in both physiological and pathological conditions (36). A study indicates that increment of VEGF-A and TGFβ2 signaling invokes gene expression changes related to glioblastoma vessels (37).

The significant role of HSP90AA1 and HSP90AB1 in several cancer types such as lung cancer, ovarian cancer, and GS is detected and discussed (38-40). The attenuation role of IL2 in growth of glioma and brain edema in rats is assessed and discussed in details (41). Down regulation of KNG1 in bladder and gastric cancers is studied and confirmed (42, 43). As it was shown, there is an exact relationship between expression changes of the nine introduced key genes and cancer diseases especially GS.

Consequently, the represented crucial genes probably are right related genes to GS.

A cluster is a combination of elements with tight relationships. In this study, a cluster was introduced containing the most number of the important genes, which is a novel finding. The cluster analysis identified 2 significant clusters. Cluster-1 includes 21 nodes, in which 7 were common with the key genes. Similarly, cluster-2 includes 17 nodes and 6 common key genes. The finding indicates that cluster-1 and cluster-2 are 78% and 67% common considering key genes, respectively. Based on the findings, it can be concluded that cluster-1 is a main structural component of the network. Thus, gene ontology enrichment of elements of cluster-1 can lead to introduce the related terms of GS.
Figure 5. Gene ontology result for the nodes of cluster -1, including biological processes, cellular components, molecular function, and biochemical pathways (from KEGG database). The term was included if at least 2 genes and 5% gene/term contribution were considered.
A number of 115 terms were recognized, in which several terms are characterized by high percentage of attribution such as the VEGF involved terms. On the other hand, there are several terms, which are specified by the presence of more genes in the terms such as glioma, melanoma, cell cycle, and bladder cancers. Since the interpretation of a large number of the terms is difficult, the terms were categorized in 5 groups. The main group is determined as glioma which indicates that our analysis was conducted in an appropriate way. The second important group is p53 signaling pathway. There is evidence about correlation between p53 pathway and glioma (44). The other groups were involved in different malignancies.

4.1. Conclusion

The 9 determined key genes may be the central core of GS and play major roles in pathology of disease. The validation of findings via experimental investigation is useful for the approval, tool.

Acknowledgments

None declared.

Footnotes

Authors’ Contribution: Mostafa Rezaei Tavirani was responsible to plan the project, writing, and editing the manuscript. Vahid Mansouri, Saeed Hesami Takalou, and Mohammad Rostami Nejad were responsible for analysis and writing the manuscript. Sina Rezaei Tavirani was involved in technical and analysis parts.

Conflict of Interests: The authors declared that they have no conflict of interest.

Financial Disclosure: None declared.

Funding/Support: This project is supported by Shahid Beheshti University of Medical Sciences.

References

1. Singh G, Das KK, Sharma P, Guruprasad B, Jaiswal S, Mehrutra A, et al. Cerebral gliosarcoma: Analysis of 16 patients and review of literature. *Asian J Neurosurg.* 2015;10(3):95–202. doi: 10.4103/1793-5482.181473. [PubMed: 26396601]. [PubMed Central: PMC455373].

2. Pardo J, Murcia M, Garcia F, Alvarado A. Gliosarcoma: A rare primary CNS tumor. Presentation of two cases. *Rep Pract Oncol Radiother.* 2010;15(4):98–102. doi: 10.1016/j.rpor.2010.05.001. [PubMed: 24376912]. [PubMed Central: PMC3863303].

3. Little JA, Shanoff HM, Roe RD, Csima A, Yano R. Studies of Male Survivors of Myocardial Infarction. IV. Serum Lipids and Five-Year Survival. *Circulation.* 1965;31:854–62. [PubMed: 14297519].

4. Hattab EM, Harris BT. Molecular Pathology of the Central Nervous System. *Molecular Genetic Pathology.* Springer; 2008. p. 497–532.

5. Newton HB, Otero J. Overview of epidemiology, pathology, and treatment of primary brain tumors. *Epilepsy and Brain Tumors.* Elsevier; 2015. p. 11–28.

6. Cho SY, Park C, Na D, Han JY, Lee J, Park OK, et al. High prevalence of TP53 mutations is associated with poor survival and an EMT signature in gliosarcoma patients. *Exp Mol Med.* 2017;49(4), e37. doi: 10.1038/emm.2017.9. [PubMed: 28408749]. [PubMed Central: PMC420801].

7. Aldape K, Zadeh G, Mansouri S, Reifenberger G, von Deimling A. Glioblastoma: pathology, molecular mechanisms and markers. *Acta Neuropathol.* 2015;129(6):829–48. doi: 10.1007/s00401-015-1442-4. [PubMed: 25941888].

8. Behling F, Barrantes-Freer A, Skardelly M, Nieser M, Christians A, Stockhammer F, et al. Frequency of BRAF V600E mutations in 969 central nervous system neoplasms. *Diagn Pathol.* 2016;11(1):55. doi: 10.1186/s13000-016-0506-2. [PubMed: 27350555]. [PubMed Central: PMC4924554].

9. Nsir AB, Thai QA, Kasar AZ, Said IB, Jemel H. Primary cerebellar gliosarcoma with extracranial metastases: an orphan differential diagnosis. *World Neurosurg.* 2015;84(6):2076. e3–7. [PubMed: 25941888].

10. Reis RM, Konu-Leblebicioglu D, Lopes JM, Kleihues P, Ohgaki H. Genetic profile of gliosarcomas. *Am J Pathol.* 2000;156(2):425–32. doi: 10.1016/S0002-9440(10)64746-3. [PubMed: 10666371]. [PubMed Central: PMC1850048].

11. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, et al. Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet.* 2000;24(3):227–35. doi: 10.1038/73432. [PubMed: 10700174].

12. deCarvalho AC, Nelson K, Lemke N, Lehman NL, Arbab AS, Kalkanis S, et al. Gliosarcoma stem cells undergo glial and mesenchymal differentiation in vivo. *Stem Cells.* 2010;28(2):381–90. doi: 10.1002/stem.264. [PubMed: 19937755]. [PubMed Central: PMC2879790].

13. Miller CR, Perry A. Glioblastoma. *Arch Pathol Lab Med.* 2007;131(3):397–406. doi: 10.1043/1543-2165(2007)131[397]20.CO;2. [PubMed: 17506742].

14. Safavi-Aligharleo N, Taghizadeh M, Rezaei-Tavirani M, Goliaei B, Peyvandi AA. Protein-protein interaction networks (PPI) and complex diseases. *Gastroenterol Hepatol Bed Bench.* 2014;7(1):17–31. [PubMed: 25436094]. [PubMed Central: PMC407556].

15. Karbalaei R, Piran M, Rezaei-Tavirani M, Asadzadeh-Aghdaei H, Heidari MH. A systems biology analysis protein-protein interaction of NASH and IBD based on comprehensive gene information. *Gastroenterol Hepatol Bed Bench.* 2017;10(3):394–201. [PubMed: 29188935]. [PubMed Central: PMC5860269].

16. Zali H, Rezaei Tavirani M. Meningioma protein-protein interaction network. *Arch Iran Med.* 2014;17(4):262–72. [PubMed: 24724603].

Figure 6. The represented 115 terms in figure 4 are grouped. Term P values and group P values are less than 10⁻³.
17. Safari-Alighiarloo N, Rezaei-Tavirani M, Taghizadeh M, Tabatabaee SM, Namaki S. Network-based analysis of differentially expressed genes in cerebrospinal fluid (CSF) and blood reveals new candidate genes for multiple sclerosis. PeerJ. 2016;4, e2775. doi: 10.7717/peerj.2775. [PubMed: 28028462]. [PubMed Central: PMC5381126].

18. Peyvandi AA, Khashishar S, Safaei A, Rezaei-Tavirani M, Azodi-Zamanian M. Interactome Analysis of 11-Dehydroxytryptamine-Labeled Ovarian Carcinoma Cells Such as CaS-22 and Cal-27 and Melanoma Cell Line. Int J Canc Manag. 2017;10(7). doi: 10.5821/jicm.10096.

19. Yildirim MA, Goh KI, Cusick ME, Barabasi AL, Vidal M. Drug-target network. Nat Biotechnol. 2007;25(10):1105–26. doi: 10.1038/nbt1338. [PubMed: 17921997].

20. Vanunu O, Magger O, Ruppin E, Shlomi T, Sharan R. Associating genes and protein complexes with disease via network propagation. PLoS Comput Biol. 2010;6(1), e1000641. doi: 10.1371/journal.pcbi.1000641. [PubMed: 20090828]. [PubMed Central: PMC2970785].

21. Abbasszadeh HA, Peyvandi AA, Sadeghi Y, Safaei A, Zamanian-Azodi M, Khoramagh MS, et al. Ery:Yag Laser and Cisplatin Effect on Cell Cycle Regulation of Human Gingival Fibroblast Cells. J Lasers Med Sci. 2017;8(Supp 1):S20–1. doi: 10.15171/jlms.2017.s4. [PubMed Central: PMC5349820].

22. Safaei A, Rezaei-Tavirani M, Zamanian-Azodi M, Lashay A, Mohammadi SF, Ghasemi-Broumand M, et al. Diabetic Retinopathy and Protein Lysophosphatase in Laser Therapy: A Protein-Protein Interaction Network Analysis. J Lasers Med Sci. 2017;8(8):1–6. doi: 10.15171/jlms.2017.03. [PubMed: 2907030]. [PubMed Central: PMC642275].

23. Manzoni C, Ferrari R, Kia D, Lovering RC, Hardy J, Lewis PA. Protein Network Analysis to Prioritize Candidate Genes and Pathways for Parkinson’s Disease. J Proteomics. 2012;75(2):101–9. doi: 10.1016/j.jprot.2012.09.002. [PubMed: 22813572]. [PubMed Central: PMC3517894].

24. Tjuvajev J, Gansbacher B, Desai R, Beattie B, Kaplitt M, Matei C, et al. Gene-expression profiles in glioma cells treated with a novel matrix metalloproteinase inhibitor. J Proteome Res. 2010;9(4):715–32. doi: 10.1021/pr100507a. [PubMed: 20747117]. [PubMed Central: PMC2955519].

25. Rezaei Tavirani M, Ok HF, Zamanian Azodi M, Rezaei Tavirani M. Interactome Analysis in Parkinson and Alzheimer’s Diseases. J Proteomics. 2012;75(2):225–36. doi: 10.1016/j.jprot.2012.09.002. [PubMed: 22986152].

26. Diagl et al. Identifying novel therapeutic targets for glioma using proteomic approaches. J Proteomics. 2012;75(2):225–36. doi: 10.1016/j.jprot.2012.09.002. [PubMed: 22986152].

27. Prados MD, Chang SM, Butowski N, Deloer R, Papavant R, Carliner H, et al. Phase II study of erlotinib plus temozolomide during and after radiation therapy in patients with newly diagnosed glioblastoma multiforme or gliosarcoma. Clin Oncol. 2009;27(4):579–84. doi: 10.1002/CO.2018.100693. [PubMed: 19075262]. [PubMed Central: PMC3664589].

28. Machuca TN, Prevedello DM, Pope L, Hao J, Suh SS, Araujo JC, Torres LF. Gliosarcoma: report of four cases with immunohistochemical findings. Arq Neuropsiquiatra. 2004;62(3):508–12. [PubMed: 1533427].

29. Khayer N, Zamanian-Azodi M, Mansouri V, Ghasemi-Broumand M, et al. Phase II study of erlotinib plus temozolomide during and after radiation therapy in patients with newly diagnosed glioblastoma multiforme or gliosarcoma. Clin Oncol. 2009;27(4):579–84. doi: 10.1002/CO.2018.100693. [PubMed: 19075262]. [PubMed Central: PMC3664589].

30. Warnick RE, Pietrangelo D, McBride DQ, Flamm ES. In vivo metabolism of radiolabeled putrescine in gliomas: implications for positron emission tomography of brain tumors. Neurosurgery.

31. Rezaei-Tavirani M et al. 2018;11(5):e65701. 7