# Identifying novel glioma associated pathways based on integrated 'omics' data

Yangfan Hu[†1], Jinquan Li[†2], Jiajia Chen[1,3], Guang Hu[1,*], Bairong Shen[1,4*]
[1]Center for Systems Biology, Soochow University, Suzhou, China 215006
[2]The First Affiliated Hospital of Soochow University
[3]School of Chemistry and Biological Engineering, Suzhou University of Science and Technology, 215009, China
[4]Department of Bioinformatics, School of Medicine, Soochow University, Suzhou 215123, China,
*Corresponding authors: bairong.shen@suda.edu.cn, huguang@suda.edu.cn
†Equal contributors

*Abstract*—**Microarray represents a high throughput technology for analyzing expression profiles, and thus it has been widely applied in the study of pathogenesis of glioma. However, most of the analyses focused on detecting the differentially expressed genes in glioma. Although it is well accepted that the pathway-derived signatures is more reproducible than that at gene level, few meta-analyses of multiple microarray datasets at system level have been previously performed. In this article, we performed meta-analysis on different published glioma expression profiles and compared the overlapping of signature at gene and pathway level. Pathway enrichment analysis result of GeneGO database and Gene Set Enrichment Analysis (GSEA) showed that 100% and 64% of the similarity was higher than that of genes respectively. Moreover, we integrated other omics data on glioma, such as MicroRNA expression profiles and Chip-Seq data, for further verification. The results showed that the significant signatures of different data sets are more similar at pathway level than at gene level. 12 pathways found by GeneGO database were shared by four stages among several datasets. 5 of these pathways, such as Regulation of epithelial-to-mesenchymal transition (EMT), TGF-beta-dependent induction of EMT via SMADs, were putative novel pathways on glioma and need further experimental verification.**

***Keywords-glioma; omics data; meta-analysis; pathway enrichment analysis***

## I. INTRODUCTION

Microarray technology has been widely used in molecular biology studies, especially in cancer research. With the wide application of microarray, various genome-wide gene expression profiling studies were performed recently. However, the noisy nature of microarray data, together with the relatively small size in each study, often results in inconsistent biological conclusions [1, 2]. Therefore, it is necessary to apply meta-analysis approaches combining several studies to obtain a more reliable and robust result. Meta-analysis, a set of statistical techniques that combine results from several studies, has been recently applied to microarray analysis for increasing the reliability and robustness of results from individual studies [3].

Glioma is the most common type of primary brain tumor in adults [4,5], which starts in the brain or spine and arises from glial cells. Based on their histological appearance, gliomas can be divided into two major subtypes according to the 2007 WHO classification [1]: astrocytic tumors, including pilocytic astrocytomas (PA), astrocytomas and glioblastomas (GBM), and oligodendroglial (OD) tumors, including pure OD tumors and mixed oligoastrocytic (MOA) tumors.

Gene expression profiling of glioma have been performed in the past few years, and some differentially expressed genes have been identified, such as the genes CDKN2A, PTEN, RB1 and TP53. Some specific genetic changes (amplification of EGFR, IDH1 mutation, and 1p19q LOH) were also observed in glioma groups. Though there has been success in identifying the genes and underlying genetic changes correlated with glioma, the complex molecular mechanism of cancer remains unclear. It is well known that cancer is a system disease [6, 7]. The analysis of microarray data at the pathway level has an inherent advantage for researches in cancer. It has been previously reported that—the differentially expressed genes often have little overlap, whereas pathway analysis often generates improved consistency [8]. Our group has also verified the hypothesis in the more recent study [9].

Based on these conditions, we integrated and analyzed omics data sets of glioma, such as gene expression microarray, MicroRNA and Chip-Seq data sets. In this article, we applied Cancer Outlier Profile Analysis (COPA) [10] to detect the significant differentially expressed genes in R environment. We then used GeneGo' Metacore for pathway enrichment analysis. Gene Set Enrichment Analysis (GSEA) [11] and MAPE [3] approach were also implemented in this study.

## II. MATERIALS AND METHODS

We collected four gene expression profiling datasets on glioma including tumors and normal brain tissues from the Gene Expression Omnibus (GEO) database. All the datasets were performed using Affymetrix oligonucleotide microarray [12-14]. According to Rhodes et al. [15], meta-analysis on the basis of two types of sample, normal brain and glioma tissues, were comparable. Before the analysis, we processed the whole datasets by MAS5.0 algorithm in R platform. Then Median Absolute Deviation (MAD) [16] was used for between-chip


978-1-4673-4398-5/12/$31.00 ©2012 IEEE

normalization, and data filtering was then applied for eliminating bad spots.

Cancer Outlier Profile Analysis (COPA) method was performed for detecting different expressed genes between normal and cancer samples. The COPA [10] package was performed in R environments. According to the algorithm, the COPA statistic is defined: all measurements for a gene are standardized by the overall median and median absolute deviation for that gene, and then the COPA statistic is the percentile of the data in the disease group. The pre-filtration threshold was set as defaulted 95th percentile, which means that the genes with a number of outlier samples less than the 95th percentile were removed from further consideration. A threshold cut-off for 'outlier' status was set and applied to all genes.

Furthermore, the interested genes were mapped to GeneGo' MetaCore for pathway enrichment analysis. In MetaCore, the p-value represents the probability to randomly obtain the intersection of certain size between two gene/protein datasets following hypergeometric distribution. Gene Set Enrichment Analysis (GSEA) was also applied in this study to assess which geneset or pathway is significant in this study. Additionally, MAPE, a powerful tool improved by Shen [3] for meta-analysis among multiple studies, was performed for pathway analysis. Moreover, we applied it to other omics data sets, such as MicroRNA expression datasets and Chip-seq data sets to further analysis.

## III. RESULTS

### *A. Data Collcetion*

The raw gene expression data sets on glioma were downloaded from Gene Expression Omnius (GEO) public database at NCBI. The data included patient tumor and normal brain samples. The gliomas were pathologically diagnosed to subtypes according to WHO standard. Table 1 summarizes the detailed information of the four datasets.

### *B. Microarray statistical analysis for glioma datasets*

Currently, many statistical methods have been used for the identification of differentially expressed genes [17,18], such as t-statistic and SAM. The t-statistic is the most commonly used method that based on the assumption that all disease samples are over-expressed. However, these traditional analytical methods are not suitable for detecting genes which only over-expressed in a small number of cancer samples [19]. More recently, some novel methods have been developed. As a novel and powerful approach for the identification of significant genes, COPA was derived from the t-statistic by replacing the mean by the median. Through applications to public cancer microarray data sets, it has shown that COPA can perform better than the usual t-statistic in these cases. Accordingly, this method was applied in this study to meta-analyze the datasets.

According to the algorithm of COPA [10], we classified the analysis detection in 2-class samples as Normal VS Glioma. Since the glioma can be classified into several subgroups, we got 11 groups of two-class in all for the COPA analysis. The significant genes numbers of all datasets were close at the value of 1.8 that was used as the COPA threshold to define the 'outlier' status in the cancer samples. The text-mining searches in the Entrez PubMed database showed that 853 out of 6306 (14%) genes were associated with glioma.

Then we mapped these differentially expressed genes to GeneGO, a manually curated and comprehensive commercial database for pathway enrichment analysis. We found that a total of 213 pathways in GeneGO database have p value of 0.05. As shown in Fig. 1, the pathways could be divided into several GeneGO's Ontology categories. For example, 48 pathways were associated with development procedure, 41 pathways were related to immune response and 19 pathways were relevant to apoptosis and survival.

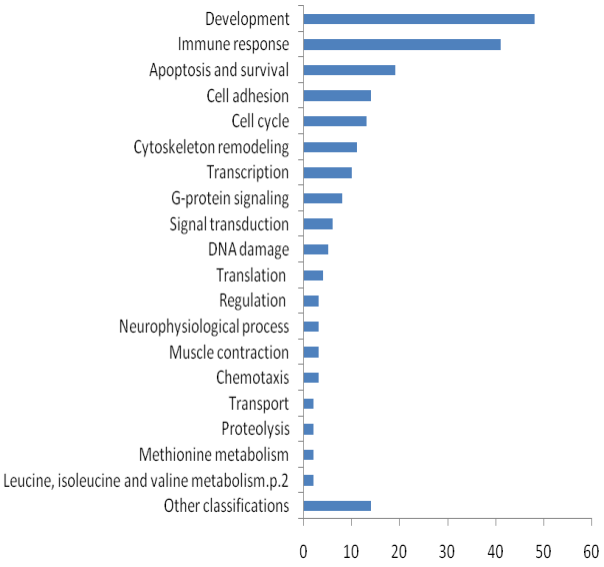


Figure 1. GeneGO Ontology classification of 213 enriched pathways

TABLE 1. Information on Microarray Expression Profiling Data of Glioma

| Dataset | Platform | Sample Number | Sample Information | | | | | | Gene Number |
|---|---|---|---|---|---|---|---|---|---|
| | | | Normal | Tumor | | | | | |
| | | | | Astrocytic | | Glioblastomas | Oligodendrogliomas | | |
| | | | | PA | A | | OD | OA | |
| Data1 | HG-U95Av2 | 25 | 5 | 6 | / | 7 | 7 | / | 12625 |
| Data 2 | U133-Plus2.0 Array | 284 | 8 | 8 | 29 | 159 | 52 | 28 | 54675 |
| Data3 | U133-Plus 2.0 Array | 15 | 6 | / | / | 8 | 1 | / | 54675 |
| Data4 | U133-Plus 2.0 Array | 180 | 23 | / | 26 | 81 | 50 | / | 54675 |


978-1-4673-4398-5/12/$31.00 ©2012 IEEE

Moreover, we performed the Gene Set Enrichment Analysis (GSEA) to assess which gene set/pathway is significant among the datasets [11]. In this study, we took the C2 curated gene sets from the Molecular Signature Database (MSigDB) as the gene set annotations, and then 513 outlier gene sets with p value of 0.05 were obtained.

### *C. Comparison of the signature similarity at different levels*

As we supposed, the similarity of signature at the system level is higher than the gene level. In order to validate the hypothesis, we performed overlapping analysis based on the gene level and pathway/gene set level. For the four datasets, 11 pairs of datasets could be compared according to the different stages of the glioma. Fig. 2 shows the comparison of the overlapping percentage among differentially expressed genes, pathways enriched by GeneGo's database, and gene sets enriched by GSEA. The result showed that the consistency across studies was higher at the pathway/Geneset level than at the gene level. The p-value for the difference of overlapping between outlier genes and GeneGO's enriched pathways were 2.77e-07 by paired t-test. We also evaluated the overlap of gene sets found by GSEA software. The result indicated that 64% of the two-pair data sets are more overlapped at the gene set level than that at the gene level. From the analysis with GeneGO and GSEA software, we verified that the signature similarities at the pathway/Geneset level are higher than that at the gene level.

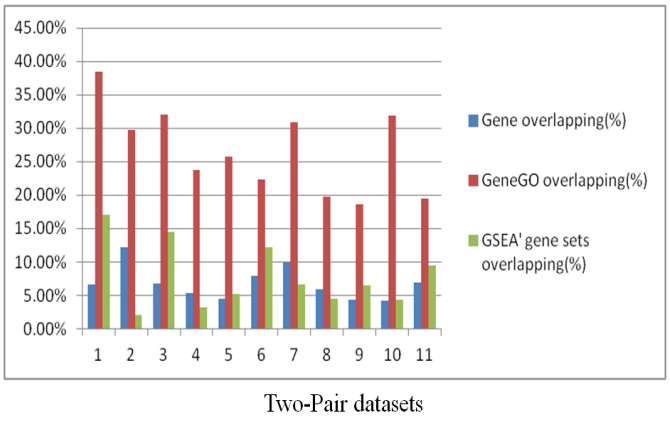


Figure 2. Comparisons of overlapping analysis among differentially expressed genes, enriched pathways in GeneGO database and enriched gene sets in GSEA

### *D. Identification of novel glioma pathways by meta-analysis at pathway level*

From the meta-analysis result, we knew that the overlapping of the enriched pathways was much higher than that at the gene level. It also obviously revealed that the identified pathways were predominantly more robust and closer to the phenotype of interest than that at the gene level. In order to identify novel glioma related pathways, we compared the number of GeneGO's enriched pathways in the four datasets classified by grades (Fig. 3). 12 common pathways shared in at least four stages were listed in Table 2. Among them, the top 6 pathways have been confirmed that related to glioma in PubMed.

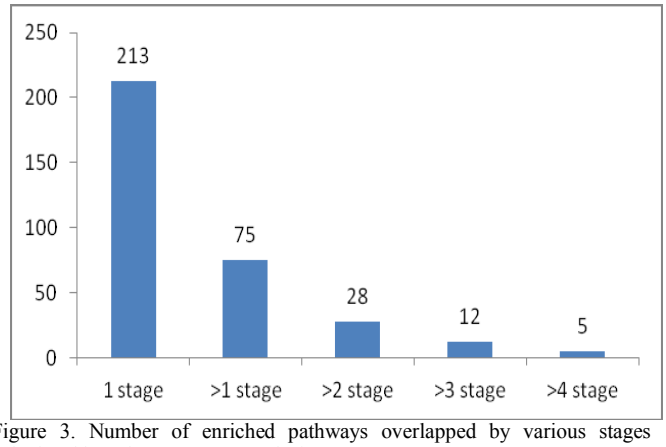


Figure 3. Number of enriched pathways overlapped by various stages calculated based on GeneGO analysis

TABLE 2. The 12 GeneGO's pathways overlapped by four stages among datasets

| Pathway Name | Pubmed Count* |
|---|---|
| Chemokines and adhesion | 633 |
| Cell cycle (generic schema) | 617 |
| TGF, WNT and cytoskeletal remodeling | 344 |
| WNT signaling pathway. Part 1. Degradation of beta-catenin in the absence WNT signaling | 5 |
| WNT signaling pathway. Part 2 | 5 |
| Cytoskeleton remodeling | 1 |
| Role of IAP-proteins in apoptosis | 0 |
| Regulation of G1/S transition (part 1) | 0 |
| NOTCH1-mediated pathway for NF-KB activity modulation | 0 |
| Regulation of epithelial-to-mesenchymal transition (EMT) | 0 |
| TGF-beta-dependent induction of EMT via SMADs | 0 |
| Non-genomic (rapid) action of Androgen Receptor | 0 |

TABLE 3.The top 6 potential novel pathways (GeneGO) found from 4 datasets

| Pathway Name | Object Count | Pubmed Count | Percentage (%) |
|---|---|---|---|
| Role of IAP-proteins in apoptosis | 31 | 23 | 74.19% |
| Regulation of G1/S transition (part 1) | 38 | 25 | 65.79% |
| NOTCH1-mediated pathway for NF-KB activity modulation | 34 | 17 | 50.00% |
| Regulation of epithelial-to-mesenchymal transition (EMT) | 64 | 51 | 79.69% |
| TGF-beta-dependent induction of EMT via SMADs | 35 | 29 | 82.86% |
| Non-genomic (rapid) action of Androgen Receptor | 40 | 24 | 60.00% |

The other six pathways (listed in Table 3) have not been reported as glioma associated pathway. So we calculated the number of identified/all genes in these pathways, respectively,


978-1-4673-4398-5/12/$31.00 ©2012 IEEE

and some indirect evidence was found to support our findings. The data in Table 3 indicate that most of the expressed genes in these pathways were reported to be related with the glioma in published papers.

### *E. Pathway enrichment analysis by MAPE software*

MAPE is an integrated approach that meta-analyzes pathway enrichment by combining statistic significance across studies at the pathway level [3]. Compared with other pathway analysis methods, MAPE has more robust and better performance among multiple microarray studies. In this study, we applied this approach to analyze the four gene expression datasets mentioned above to further verify our hypothesis. In order to understand the mechanism more accurately, we analyzed the data according to grades. As a result, 91 pathways correlated with the glioma were obtained.

By comparison with the results from the gene expression data, 27 overlapped pathways were found in the two results, where the GeneGO's pathway Ontology categories are in accordance with the previous results.

### *F. Cross-validation by analyzing other omics data*

Other omics data were also analyzed to verify our hypothesis. MicroRNAs (miRNAs) are small, non-coding RNAs consisting of 20-25 nucleotides that regulate target gene expression at the post-transcriptional level. MicroRNAs have been demonstrated to play important roles in the tumor development, prognosis and metastasis [20-22]. In recent years, several miRNAs have been reported to be involved in modulation of glioma development, such as microRNA-21 (miR-21), which has been demonstrated to be an oncogene in cultured glioblastoma multiforme cells [23]. Here, we downloaded three miRNAs expression profiles from the GEO database (Table 4) and conducted differential expression analysis using the same COPA package to detect significant miRNAs between the normal and tumor samples. The four widely web-based databases, TargetScan, miRanda [24], RNAhybrid [25] and TargetSpy [26], were used to predict target genes for selected miRNAs. To better study the biological functions of target genes, we subsequently retrieved the enriched biological pathways predicted by GeneGO database.

TABLE 4. Information on MicroRNA Expression Profiling Data of Glioma

| Country | Platform | Number (all) | Sample information | | MicroRNA Number | Publication Year |
|---|---|---|---|---|---|---|
| | | | Normal | Tumor | | |
| Italy | DiSteBa_Homo sapiens_Glioblastoma miRNA 340_v1.0 | 74 | 37 | 37 | 340 | 2011.01 |
| Italy | TJU-Human-Mouse-MicroRNA-1.6k-v1.1 | 35 | 13 | 22 | 353 | 2005.09 |
| USA | Agilent 8 x 15K Human miRNA-specific microarray | 34 | 10 | 24 | 1510 | 2009.12 |

Totally we found 187 pathways correlated with glioma shared by the three datasets with p-value < 0.05. 5 out of the top 6 potential novel glioma pathways mentioned above in the study could be found in results (Table 5). So these 5 pathways were considered to be putative novel glioma pathways. The GeneGO's pathway Ontology categories are in accordance with the results obtained by the gene expression datasets.

TABLE 5. The top 5 novel GeneGO's pathways overlapped by four stages

| Pathway |
|---|
| Regulation of G1/S transition (part 1) |
| NOTCH1-mediated pathway for NF-KB activity modulation |
| Regulation of epithelial-to-mesenchymal transition (EMT) |
| TGF-beta-dependent induction of EMT via SMADs |
| Non-genomic (rapid) action of Androgen Receptor |

Next-generation sequencing rapidly transforms our ability to profile the transcriptional, genetic, and epigenetic states of a cell [27], such as ChIP-seq is becoming the main approach to the genome-wide study of protein–DNA interactions [28]. We downloaded one chip-seq dataset (accession number GSM575227) from the study conducted by Fang [29] in the GEO database. MACS [30] and SISSRs [31] were used for statistically significant peaks determination, while PeakAnalyzer were applied for target genes annotation analysis. Then we mapped the genes to GeneGO for pathway analysis and got 76 glioma pathways with the 0.05 p-value. As one of the five pathways, TGF-beta-dependent induction of EMT via SMADs, was also be verified in the chip-seq analysis.

Moreover, we compared the pathways from gene expression data, MicroRNA expression data and Chip-seq data and found 14 overlapped pathways by the three different omics data.

## IV. DISCUSSION

It is well known that cancer is a complex systems biology disease [7], which is correlated with a lot of factors such as genetic information, environment effect and personal behaviors. Among them, genetic information (e.g. gene mutation) was thought as the crucial factor in the cancer transformation and progression; environment effect (e.g. carcinogens) and personal behavior (e.g. smoking) will also cause and accelerate the cancer disease. Therefore, we proposed that we should find the significant signatures at a system level like gene sets, dynamic network or pathway level. To prove our hypothesis, we meta-analyzed four gene expression profiling datasets on glioma, trying find several potential novel pathways for the future experimental validation.

We also used COPA, a novel method, to identify the differentially expressed genes between glioma and normal samples in this study and then detected enriched gene sets and pathways via GESA tool and GeneGO's MetaCore software. In the previous study, Katara et al [32] found that the genes CDKN2A, PTEN, RB1 and TP53 are expressing at a lower level than the normal and are also an important cause of cancers. The genes CDK4, PDGFA and PDGFB,PDGFRA


978-1-4673-4398-5/12/$31.00 ©2012 IEEE

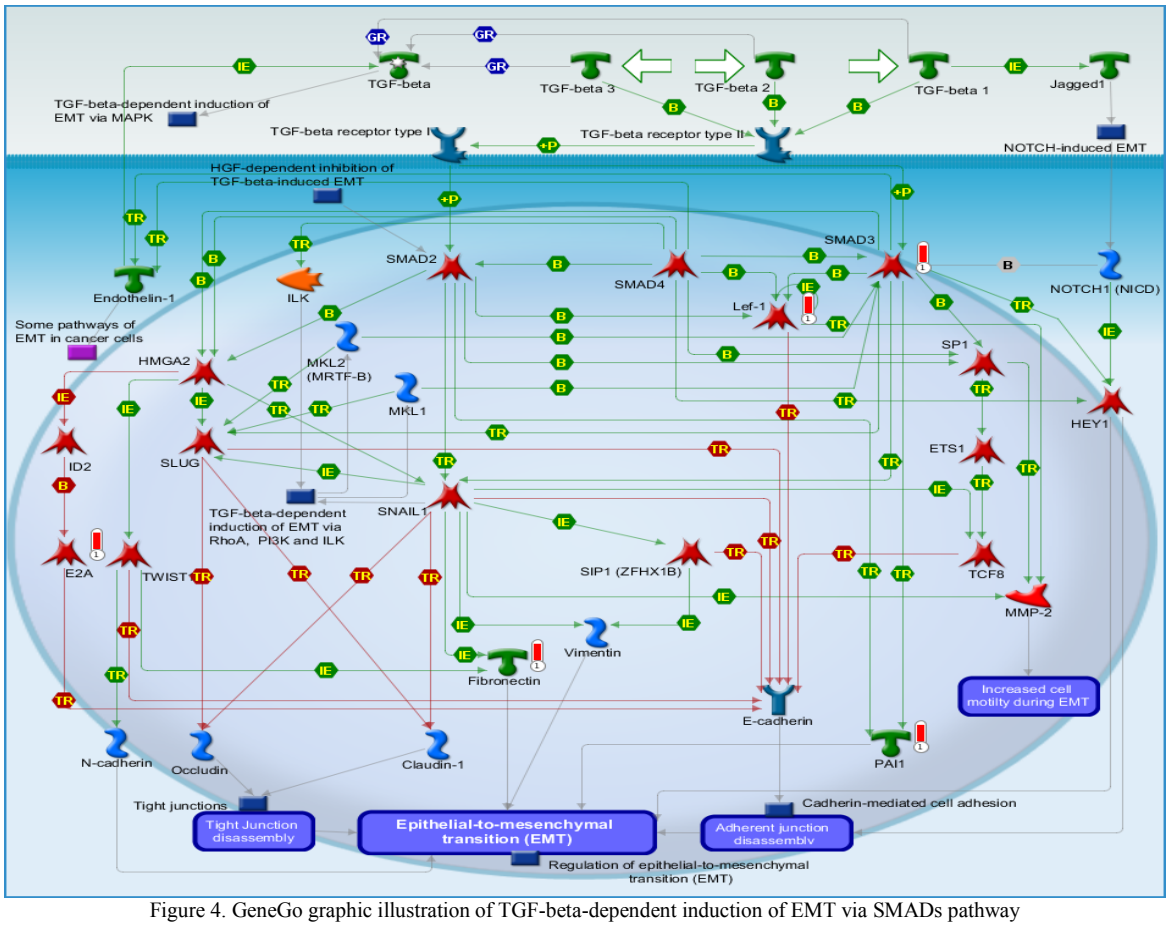


Figure 4. GeneGo graphic illustration of TGF-beta-dependent induction of EMT via SMADs pathway

and PDGFRB, MDM2 and EGFR are over expressing at a rate higher than the normal and are acting as oncogenes in these cancers. Tanwar et al [33] found that the YKL-40 is the most differentially expressed gene in the series of GBMs versus normal brain. The data suggest YKL-40 may be involved in extracellular matrix degradation and/or angiogenesis. Specific genetic changes ( EGFRamplification, IDH1 mutation, and 1p19q LOH) were also observed in glioma group [34]. In the meta-analysis of glioma gene expression datasets, Dreyfuss et al [35] collected four independent gliomaresearch datasets including anaplastic astrocytoma (AA) and glioblastoma multiforme (GBM) samples. They combined the statistics across studies using the nonparametric rank sum method to identify several differentially expressed genes associated with glioma. In these studies, the common significant gene analysis based on t-test or t-test like statistics method calculated the average expression of genes without considering the individual differences. In comparison with them, we performed a novel method, Cancer Outlier Profile Analysis (COPA), to detect the outlier genes in the subsets of the cancer samples with the consideration of individual differences. Moreover, we proposed that the similarity at pathway level is higher than that at gene level, and then applied a powerful and manually curated GeneGO database and GSEA, a well-known statistics computing approach for path way enrichment analysis, to compare the overlapping at pathway level. Fig. 4 showed the pathway map for one of the novel pathways in GeneGO. It illustrated that the differentially expressed genes may reside at the different location of the same pathway and further supported our hypothesis. TGF-beta-dependent induction of EMT via SMADs, as the potential novel pathways related to glioma, need further biological experiments. The important role of Smad interacting protein 1 (SIP1) in glioma has been identified in Xia's study [36]. Penuelas et al [37] identified a molecular mechanism that TGF-beta and LIF have an essential role in the regulation of Glioma-initiating cells (GICs) in human glioblastoma. Therefore, the novel pathway may play an active part in glioma.

In addition, MicroRNA expression profiles and Chip-seq data were analyzed for further verification. Compared with the results from gene expression datasets, the five novel glioma related pathways were also found in this study. Recent reports have shown that FOXO3a inhibits cell-cycle progression at the G1/S transition by controlling transcription of the cyclin-dependent kinase inhibitor p27 (kip1), which is frequently down-regulated in human cancers, including human glioma. NF-kB is a transcription factor that plays a key role in carcinogenesis by controlling expression of several oncogenes, growth factors and cell adhesion molecules [38-40]. Li et al [41] previously reported that ECRG4 serves as a tumor suppressor in glioma, which were speculated to be involved in glioma cell growth suppression by regulating the NF- B pathway.


978-1-4673-4398-5/12/$31.00 ©2012 IEEE

## V. CONCLUSION

Compared with the previous analyses, we verified the hypothesis that the overlapping of signatures is higher at the pathway/gene set level than that at the gene level. GeneGo database, GSEA and MAPE software were used for pathway enrichment analysis, which showed several potential but novel glioma pathways. Moreover, this method was applied to miRNAs expression profiles and Chip-seq data sets for further verification. In the future, we will develop some robust methods for interested genes detection and meta-analyze the similarity of significant signatures at a system level, such as dynamic network, to better understand the molecular mechanisms of cancer occurrence and development.

## ACKNOWLEDGMENT

This work was supported by the National Nature Science Foundation of China (31170795, 91029703), the Specialized Research Fund for the Doctoral Program of Higher Education of China (20113201110015). International S&T Cooperation Program of Suzhou (SH201120) and the National 973 Programs of China (2010CB945600).

## REFERENCES

[1] L. Ein-Dor, I. Kela, G. Getz, D. Givol, and E. Domany, "Outcome signature genes in breast cancer: is there a unique set?" Bioinformatics, vol. 21, (no. 2), pp. 171-8, Jan 2005.

[2] L. Tian, S. A. Greenberg, S. W. Kong, J. Altschuler, I. S. Kohane, and P. J. Park, "Discovering statistically significant pathways in expression profiling studies," Proc Natl Acad Sci U S A, vol. 102, (no.38), pp. 13544-9, Sep 2005.

[3] K. Shen, and G.C. Tseng, "Meta-analysis for pathway enrichment analysis when combining multiple genomic studies," Bioinformatics, vol. 26, (no.10), pp. 1316-23, Apr 2010.

[4] O. Martinho, S. Granja, T. Jaraquemada, C. Caeiro, V. Miranda-Goncalves, M. Honavar, et al., "Downregulation of RKIP is associated with poor outcome and malignant progression in gliomas," PLoS One, vol. 7, (no. 1), e30769, Jan 2012

[5] H. Ohgaki, P. Dessen, B. Jourde, S. Horstmann, T. Nishikawa, P. L. Di Patre, et al., "Genetic pathways to glioblastoma: a population-based study," Cancer Res, vol. 64. (no. 19), pp. 6892-9, Oct 2004.

[6] D. Faratian, A. Goltsov, G. Lebedeva, A. Sorokin, S. Moodie, P. Mullen, et al., "Systems biology reveals new strategies for personalizing cancer medicine and confirms the role of PTEN in resistance to trastuzumab," Cancer Res, vol. 69, (no. 16), pp. 6713-20, Aug 2009.

[7] I.G. Khalil, and C. Hill, "Systems biology for cancer," Curr Opin Oncol, vol. 17, (no. 1), pp. 44-8, Jan 2005.

[8] T. Manoli, N. Gretz, H. J. Grone, M. Kenzelmann, R. Eils, and B. Brors, "Group testing for pathway analysis improves comparability of different microarray datasets," Bioinformatics, vol. 22, (no. 20), pp. 2500-6, Oct 2006.

[9] Y. Wang, J. J. Chen, Q. H. Li, H. Y. Wang, G. Q. Liu, Q. Jing, et al., "Identifying novel prostate cancer associated pathways based on integrative microarray data analysis," Comput Biol Chem, vol. 35, (no. 3), pp. 151-8, Jun 2011.

[10] J.W. MacDonald, and D. Ghosh, "COPA--cancer outlier profile analysis," Bioinformatics, vol. 22, (no. 23), pp. 2950-1, Dec 2006.

[11] A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, et al., "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles," Proc Natl Acad Sci U S A, vol. 102, (no.43), pp.15545-50, Oct 2005.

[12] S. M. Pollard, K. Yoshikawa, I. D. Clarke, D. Danovi, S. Stricker, R. Russell, et al., "Glioma stem cell lines expanded in adherent culture have tumor-specific phenotypes and are suitable for chemical and genetic screens," Cell Stem Cell, vol. 4, (no. 6), pp. 568-80, Jun 2009.

[13] L. A. M. Gravendeel, M. C. M. Kouwenhoven, O. Gevaert, J. J. de Rooi, A. P. Stubbs, J. E. Duijm, et al., "Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology," Cancer Res, vol. 69, (no. 23), pp. 9065-72, Dec 2009.

[14] Y. Kotliarov, S. Kotliarova, N. Charong, A. G. Li, J. Walling, E. Aquilanti, et al., "Correlation analysis between single-nucleotide polymorphism and expression arrays in gliomas identifies potentially relevant target genes," Cancer Res, vol. 69, (no. 4), pp. 1596-603, Feb 2009.

[15] D. R. Rhodes, J. J. Yu, K. Shanker, N. Deshpande, R. Varambally, D. Ghosh, et al., "Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression," Proc Natl Acad Sci U S A, vol. 101, (no. 25), pp. 9309-14, Jun 2004.

[16] N. J. Chung, X. D. Zhang, A. Kreamer, L. Locco, P. F. Kuan, S. Bartz, et al., "Median absolute deviation to improve hit selection for genome-scale RNAi screens," J Biomol Screen, vol. 13, (no. 2), pp. 149-58, Feb 2008.

[17] K. A. Baggerly, K. R. Coombes, K. R. Hess, D. N. Stivers, L. V. Abruzzo, and W. Zhang, "Identifying differentially expressed genes in cDNA microarray experiments," J Comput Biol, vol. 8, (no. 6), pp. 639-59, 2001.

[18] B. Wu, "Cancer outlier differential gene expression detection," Biostatistics, vol. 8, (no. 3), pp. 566-75, Jul 2007.

[19] H. Lian, "MOST: detecting cancer differential gene expression," Biostatistics, vol. 9, (no. 3), pp. 411-8, Jul 2008.

[20] R. I. Gregory, and R. Shiekhattar, "MicroRNA biogenesis and cancer," Cancer Res, vol. 65, (no. 9), pp.3509-3512, May 2005.

[21] G. A. Calin, and C. M. Croce, "MicroRNA signatures in human cancers," Nat Rev Cancer. vol. 6, (no.11), pp.857-866, Nov 2006.

[22] A. Esquela-Kerscher, and F. J. Slack, "Oncomirs — microRNAs with a role in cancer," Nat Rev Cancer. vol. 6, (no. 4), pp.259–269, Apr 2006.

[23] Y. Ren, X. Zhou, M. Mei, X. B. Yuan, L. Han, G. X. Wang, et al., "MicroRNA-21 inhibitor sensitizes human glioblastoma cells U251 (PTEN-mutant) and LN229 (PTEN-wild type) to taxol," BMC Cancer, vol. 10, pp. 27, Jan 2010.

[24] B. John, A. J. Enright, A. Aravin, T. Tuschl, C. Sander, and D. S. Marks, "Human MicroRNA targets," PLoS Biol, vol. 2, (no. 11), 1862-1879, Nov 2004.

[25] M. Rehmsmeier, P. Steffen, M. Hochsmann, and R. Giegerich, "Fast and effective prediction of microRNA/target duplexes," RNA, vol. 10, (no. 10), pp. 1507-17, Oct 2004.

[26] M. Sturm, M. Hackenberg, D. Langenberger, and D. Frishman, "TargetSpy: a supervised machine learning approach for microRNA target prediction," BMC Bioinformatics, vol. 11, pp. 292, May 2010.

[27] A. Diaz, K. Park, D. A. Lim, and J. S. Song, "Normalization, bias correction, and peak calling for ChIP-seq," Stat Appl Genet Mol Biol, vol. 11, (no. 3), Mar 2012.

[28] D. S. Johnson, A. Mortazavi, R. M. Myers, and B. Wold, "Genome-wide mapping of in vivo protein-DNA interactions," Science, 316, pp.1497-1502, Jun 2007.

[29] X. F. Fang, J. G. Yoon, L. S. Li, W. Yu, J. F. Shao, D. S. Hua, et al., "The SOX2 response program in glioblastoma multiforme: an integrated ChIP-seq, expression microarray, and microRNA analysis," BMC Genomics, vol. 12, pp. 11, Jan 2011.

[30] Y. Zhang, T. Liu, C. A. Meyer, J. Eeckhoute, D. S. Johnson, B. E. Bernstein, et al., "Model-based analysis of ChIP-Seq (MACS)," Genome Biol, vol. 9, (no. 9), pp. R137, 2008.

[31] R. Jothi, S. Cuddapah, A. Barski, K. Cui, K. Zhao, "Genome-wide identification of in vivo protein-DNA binding sites from ChIP-Seq data," Nucleic Acids Res, vol. 36, (no. 16), pp. 5221-31, 2008.

[32] P. Katara, S. Neeru, S. Sugandha, K. Indu, K. Akansha, K. Lalima, and S. Vinay, "Comparative microarray data analysis for the expression of genes in the pathway of glioma," Bioinformation, vol. 5, (no. 1), pp. 31-4, Jun 2010.


978-1-4673-4398-5/12/$31.00 ©2012 IEEE

[33] A. S. Abdoon, N. Ghanem, O. M. Kandil, A. Gad, K. Schellander, and D. Tesfaye, "cDNA microarray analysis of gene expression in parthenotes and in vitro produced buffalo embryos," Theriogenology, vol. 77, (no. 6), pp. 1240-51, Apr 2012.

[34] L. A. M. Gravendeel, M. C. M. Kouwenhoven, O. Gevaert, J. J. de Rooi, A. P. Stubbs, J. E. Duijm, et al., "Intrinsic gene expression profiles of gliomas are a better predictor of survival than histology," Cancer Res, vol. 69, (no. 23), pp. 9065-72, Dec 2009.

[35] J. M. Dreyfuss, M. D. Johnson, and P. J. Park, "Meta-analysis of glioblastoma multiforme versus anaplastic astrocytoma identifies robust gene markers," Mol Cancer, vol. 8, pp. 71, Sep 2009.

[36] M. Xia, M. H. Hu, J. Wang, Y. J. Xu, X. B. Chen, Y. D. Ma, et al., "Identification of the role of Smad interacting protein 1 (SIP1) in glioma," J Neurooncol, vol. 97, (no. 2), pp. 225-32, Apr 2010.

[37] S. Penuelas, J. Anido, R. M. Prieto-Sanchez, G. Folch, I. Barba, I. Cuartas, et al., "TGF-beta increases glioma-initiating cell self- renewal through the induction of LIF in human glioblastoma," Cancer Cell, vol. 15, (no. 4), pp. 315-27, Apr 2009.

[38] T. D. Gilmore, M. Koedood, K. A. Piffat, and D. W. White, "Rel/NF-kappaB/IkappaB proteins and cancer," Oncogene, vol. 13, (no. 7), pp. 1367-78, Oct 1996.

[39] C. H. Lee, Y. T. Jeon, S. H. Kim, and Y. S. Song, "NF-kappaB as a potential molecular target for cancer therapy," Biofactors, vol. 29, (no. 1), pp. 19-35, 2007.

[40] F. Lerebours, S. Vacher, C. Andrieu, M. Espie, M. Marty, R. Lidereau, et al., "NF-kappa B genes have a major role in inflammatory breast cancer," BMC Cancer, vol. 8, pp. 41, Feb 2008.

[41] W. Li, X. R. Liu, B. Zhang, D. X. Qi, L. H. Zhang, Y. H. Jin, et al., "Overexpression of candidate tumor suppressor ECRG4 inhibits glioma proliferation and invasion," J Exp Clin Cancer Res, vol. 29, pp. 89, Jul 2010.


978-1-4673-4398-5/12/$31.00 ©2012 IEEE