

PRIMARY RESEARCH

Open Access



# Integrated whole genome microarray analysis and immunohistochemical assay identifies COL11A1, GJB2 and CTRL as predictive biomarkers for pancreatic cancer

Defeng Sun<sup>1†</sup>, Haoyi Jin<sup>1†</sup>, Jun Zhang<sup>2</sup> and Xiaodong Tan<sup>3\*</sup>

## Abstract

**Background:** Pancreatic cancer is characterized by its unsatisfying early detection rate, rapid disease progression and poor prognosis. Further studies on molecular mechanism and novel predictive biomarkers for pancreatic cancer based on a large sample volume are required.

**Methods:** Multiple bioinformatic analysis tools were utilized for identification and characterization of differentially expressed genes (DEGs) from a merged microarray data (100 pancreatic cancer samples and 62 normal samples). Data from the GEO and TCGA database was utilized to validate the diagnostic and prognostic value of the top 5 upregulated/downregulated DEGs. Immunohistochemical assay (46 paired pancreatic and para-cancerous samples) was utilized to validate the expression and prognostic value of COL11A1, GJB2 and CTRL from the identified DEGs.

**Results:** A total number of 300 DEGs were identified from the merged microarray data of 100 pancreatic cancer samples and 62 normal samples. These DEGs were closely correlated with the biological characteristics of pancreatic cancer. The top 5 upregulated/downregulated DEGs showed good individual diagnostic/prognostic value and better combined diagnostic/prognostic value. Validation of COL11A1, GJB2 and CTRL with immunohistochemical assay showed consistent expression level with bioinformatics analysis and promising prognostic value.

**Conclusions:** Merged microarray data with bigger sample volume could reflect the biological characteristics of pancreatic cancer more effectively and accurately. COL11A1, GJB2 and CTRL are novel predictive biomarkers for pancreatic cancer.

**Keywords:** Bioinformatics, Pancreatic cancer, Biomarker, COL11A1, GJB2, CTRL

## Background

Pancreatic cancer is characterized by its unsatisfying early detection rate, rapid disease progression and poor prognosis [1, 2]. Multiple aberrantly expressed genes and dysregulated signaling pathways have been reported to play critical roles in the development of pancreatic cancer [3–7]. However, the underlying molecular mechanism of pancreatic cancer is still not fully understood. Whole

genome microarray is an effective way to analyze the expression profile of a human being in cell or tissue level [8]. Multiple studies and government supported projects like The Cancer Genome Atlas (TCGA) have performed whole genome microarray of patients' pancreatic cancer tissues and its paired adjacent normal pancreatic tissues to identify the differentially expressed genes (DEGs) [9–11]. These DEGs could be potential key regulators in disease progression and predictive biomarkers for pancreatic cancer [12–14]. However, the accuracy and efficacy of aforementioned studies are compromised by their limited sample volume. The results obtained from studies with limited sample volume might be unrepresentative

\*Correspondence: tanxdcmu@163.com

<sup>†</sup>Defeng Sun and Haoyi Jin contributed equally to this work

<sup>3</sup>Thyroid and Pancreatic Surgery Ward, Shengjing Hospital of China Medical University, Shenyang 110004, People's Republic of China  
Full list of author information is available at the end of the article



and biased. Comparing with other cancers like lung cancer or breast cancer, pancreatic cancer is a rather rare cancer type. Therefore, an integrated analysis of the existing whole genome microarray data is a more practical and cost effective way to overcome the aforementioned shortcomings [15–17].

In this study, we selected and further merged the data of 162 samples from 3 qualified Gene Expression Omnibus (GEO) datasets (GSE15471, GSE16515 and GSE32676) for integral analysis. Gene Set Enrichment Analysis (GSEA) was performed to evaluate the efficacy and accuracy of our merged data to reflect the biological differences between cancer and normal tissues [18]. A total number of 300 DEGs were identified from the merged data ( $|\log_2^{\text{fold change}}| > 1.5$  and adjusted  $p < 0.05$ ). Gene Ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were performed to analyze the functional classification and enriched signaling pathways of the DEGs [19–22]. Protein–protein interaction (PPI) analysis was performed to visualize the interaction network of identified DEGs. Individual and combined diagnostic value of the top 5 upregulated and downregulated genes were evaluated with receiver operating characteristic curve (ROC) analysis. We further validated the expression level and performed survival analysis of these 10 genes with the data from TCGA database. Finally, we selected and validated the clinical predictive value of COL11A1 (collagen alpha-1(XI) chain), GJB2 (gap junction beta-2 protein) and CTRL (chymotrypsin-like protease CTRL-1) with immunohistochemical (IHC) analysis of 46 paired pancreatic cancer and para-cancerous tissue sections.

## Methods

### GEO datasets selection

The following standards were applied for qualified GEO datasets selection. (1) Whole genome microarrays of pancreatic cancer tissues and paired adjacent normal pancreatic tissues were included only. Microarrays containing samples as pancreatic cancer cell lines or pancreatic organoids were excluded. (2) Whole genome microarrays with sample volume  $> 30$  samples were included only. Datasets with sample volume  $< 30$  samples were excluded to avoid the inclusion of unrepresentative data. (3) Published whole genome microarrays were included only for better quality control and repeatability. (4) Whole genome microarrays of the same platform were included only (GPL570[HG-U133\_Plus\_2]Affymetrix Human Genome U133 Plus 2.0 Array). Datasets from other platforms were excluded to avoid the potential bias caused by technological difference between different platforms. In accordance with these standards, three

qualified datasets (GSE15471, GSE16515 and GSE32676) were selected for further analysis.

### Data normalization and merging

Raw data of the selected microarrays were extracted with the R language package *affy*. Extracted expression data were normalized and transformed to log<sub>2</sub> based logarithm with the *rma* function of *affy* package. Batch effect was excluded with the R language package *sva* before the aforementioned data were merged into one dataset.

### Gene Set Enrichment Analysis

Efficacy and accuracy of the merged data to reflect the biological differences between tumor and normal groups were evaluated with Gene Set Enrichment Analysis (GSEA, Broad Institute, <http://www.broadinstitute.org/gsea/index.jsp>) in accordance with the official tutorial.

### Identification of differentially expressed genes

The R language package *limma* was utilized for the calculation of DEGs (Fold change = Pancreatic cancer sample expression/paired adjacent normal pancreatic tissue sample expression,  $|\log_2^{\text{fold change}}| > 1.5$  and adjusted  $p < 0.05$ ). For repeated gene expression data, the one with smaller *p* value was used for the downstream analysis. The expression level of DEGs was visualized with R language package *ggplot2*, *pheatmap* and GraphPad Prism 7 software.

### GO analysis and KEGG analysis

The R language package *clusterProfiler* and online database Database for Annotation, Visualization and Integrated Discovery (DAVID, <https://david.ncifcrf.gov/>) was utilized for GO analysis and KEGG analysis of the DEGs. The results were visualized with R language.

### Protein–protein interaction network analysis

The PPI network of identified DEGs was calculated with STRING database (<https://string-db.org/cgi/input.pl>) and visualized with Cytoscape software in accordance with the official tutorial [23].

### ROC analysis of selected differentially expressed genes

Diagnostic value of the top 5 upregulated and downregulated genes from the identified DEGs was calculated with ROC analysis. ROC analysis of individual genes and combined genes were performed with SPSS 19.0 and visualized with GraphPad Prism 7. Diagnostic model of combined genes was established by binary logistic regression with SPSS 19.0.

### Validation with TCGA database and survival analysis

The expression level of the top 5 upregulated and downregulated genes from the identified DEGs was validated

with microarray data from TCGA database and further visualized with GraphPad Prism 7. Prognostic value of these 10 selected genes was evaluated with survival analysis. Survival analysis was performed with OncoLnc platform (<http://www.oncolnc.org/>) utilizing clinical data from TCGA database.

### Immunohistochemical assay

IHC analysis of the patients' pancreatic cancer tissue sections and its paired para-cancerous tissue sections was performed as previously described [24]. Antibody against COL11A1, GJB2 and CTRL was used at the dilution range of 1:300 (Abcam, Cambridge, UK). The stained sections were evaluated by two different specialists. The postoperative survival rate of patients with different expression level of these two genes was visualized with GraphPad Prism 7.

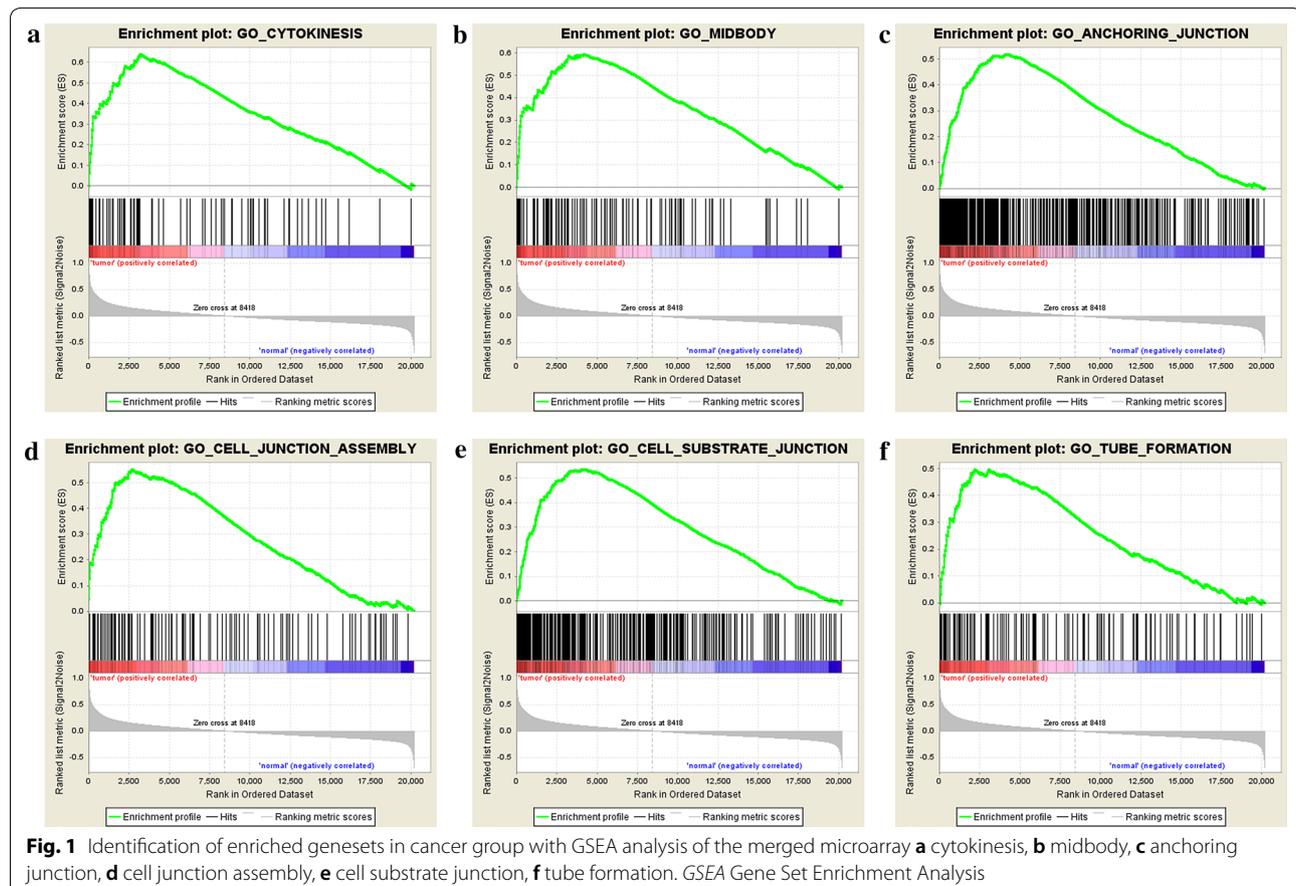
### Statistical analysis

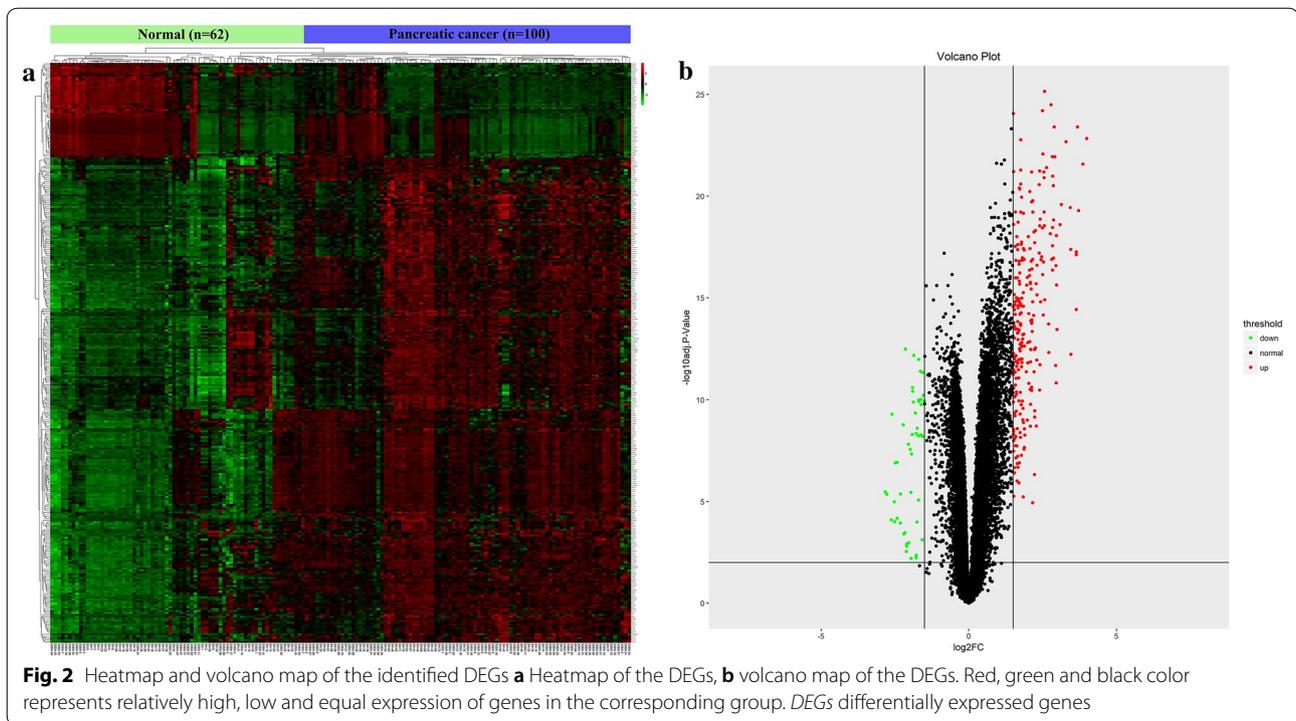
Student's t-test was performed with IBM SPSS Statistics version 19.0.  $p < 0.05$  was considered statistically significant.

## Results

### Identification of differentially expressed genes from merged microarray dataset

Three qualified GEO datasets were merged into one dataset which contains microarray data of 100 pancreatic cancer tissue samples and 62 normal pancreas tissue samples. GSEA analysis of the merged dataset showed that the tumor group was enriched in genes regulating cell division (cytokinesis, midbody), cell junction (anchoring junction, cell junction assembly, cell substrate junction) and tube formation (Fig. 1a–f). These enriched gene sets were closely correlated with cell proliferation, migration, invasion and angiogenesis. The aforementioned results suggested that our merged data was qualified to reflect the biological characteristics and expression profile of cancer samples. A total number of 300 DEGs was further identified from the merged microarray data, of which 251 were upregulated and 49 were downregulated in pancreatic cancer samples (Additional file 1). Identified DEGs were visualized with heatmap and volcano map (Fig. 2a and b).





### GO analysis and KEGG analysis of the identified differentially expressed genes

GO analysis and KEGG analysis were performed to analyze the functional classification and signaling pathway enrichment of the identified DEGs. The results showed that the DEGs were closely correlated with extracellular environment reorganization in cellular component (CC) and biological process (BP) classification (Fig. 3a and b). For molecular function (MF) classification, the DEGs were enriched in multiple peptidase activity and integrin binding (Fig. 3c). These results suggested that the DEGs were closely correlated with extracellular matrix (ECM) degradation and remodeling which is the essential step for local invasion and distant metastasis. The results from KEGG analysis showed that the DEGs were significantly enriched in signaling pathways of pancreatic secretion, fat/protein digestion and absorption (Fig. 3d). This suggested that the development of pancreatic cancer is characterized by the loss of pancreas's normal physiological functions. In accordance with GO analysis, the DEGs were closely correlated with signaling pathways regulating ECM-receptor interaction and focal adhesion.

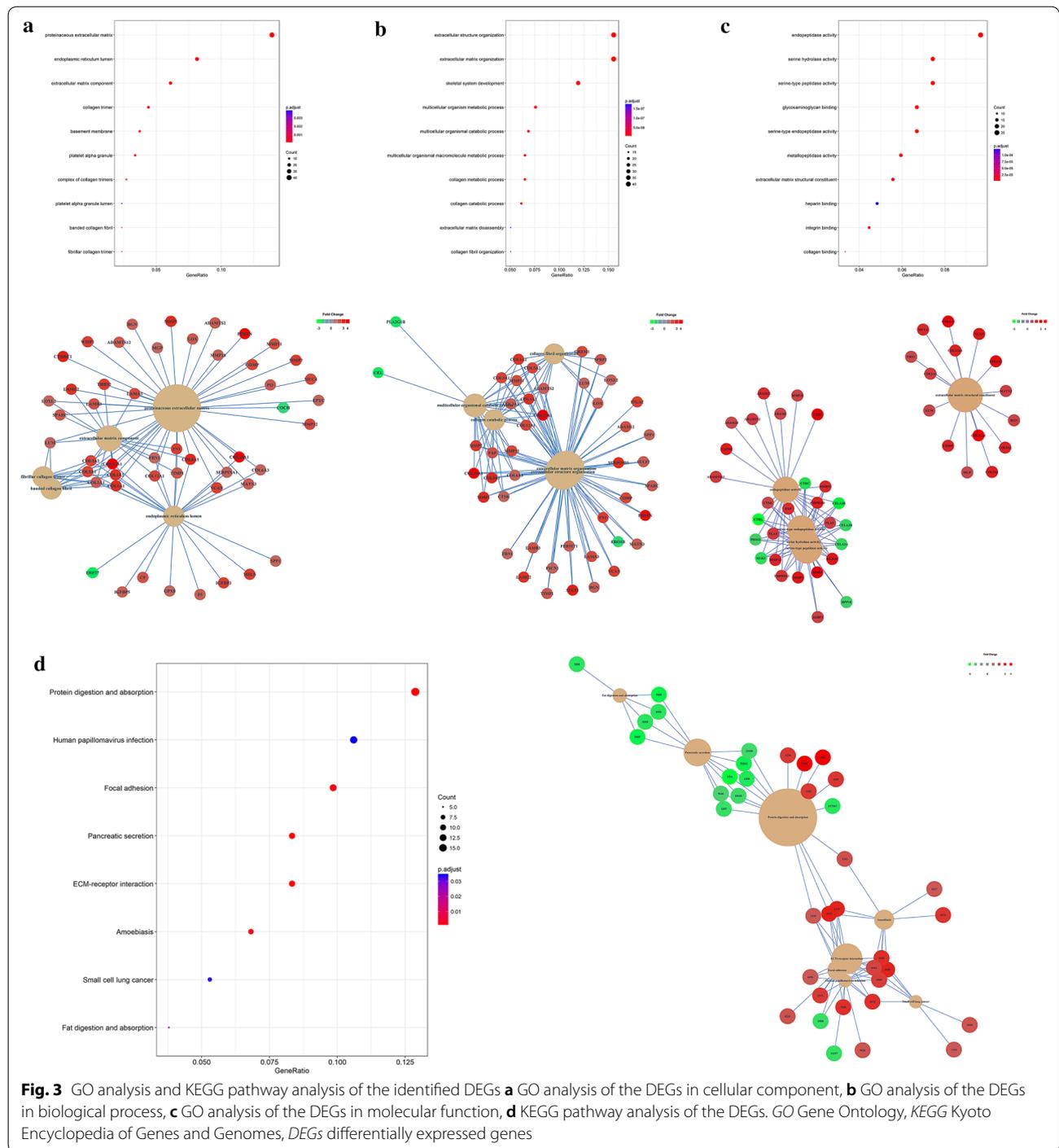
### Protein–protein interaction analysis of the identified differentially expressed genes

We utilized STRING database to analyze the PPI network of the DEGs to identify the key genes and their

interactions in pancreatic cancer progression. The visualized results showed that epidermal growth factor (EGF) was located in the core of our PPI network (Fig. 4). COL11A1, COL10A1 from the top 5 upregulated genes and CTRL, SYCN, PNLIPRP1 from the top 5 downregulated genes were also found to be the key genes in the PPI network (Table 1).

### Potential clinical value of the top 5 upregulated and downregulated differentially expressed genes

The top 5 upregulated and downregulated DEGs were selected for further validation of their potential diagnostic value (Fig. 5a and b). The results of ROC analysis indicated that these 5 upregulated DEGs possessed higher individual diagnostic value than those 5 downregulated DEGs. We established a diagnostic panel of these genes combined with binary logistic regression. These 10 genes showed better combined diagnostic value than individually (Fig. 5c). We further validated the expression level of these 10 genes with the data from TCGA database. The results showed similar expression level of these selected DEGs except no expression data of SPX was found (Fig. 6a). Survival analysis of the top 5 upregulated DEGs showed promising prognostic value (Fig. 6b). Combined analysis of these 10 genes showed that dysregulation of these 10 genes was closely correlated with poorer overall survival rate and disease free survival rate (Fig. 6c).

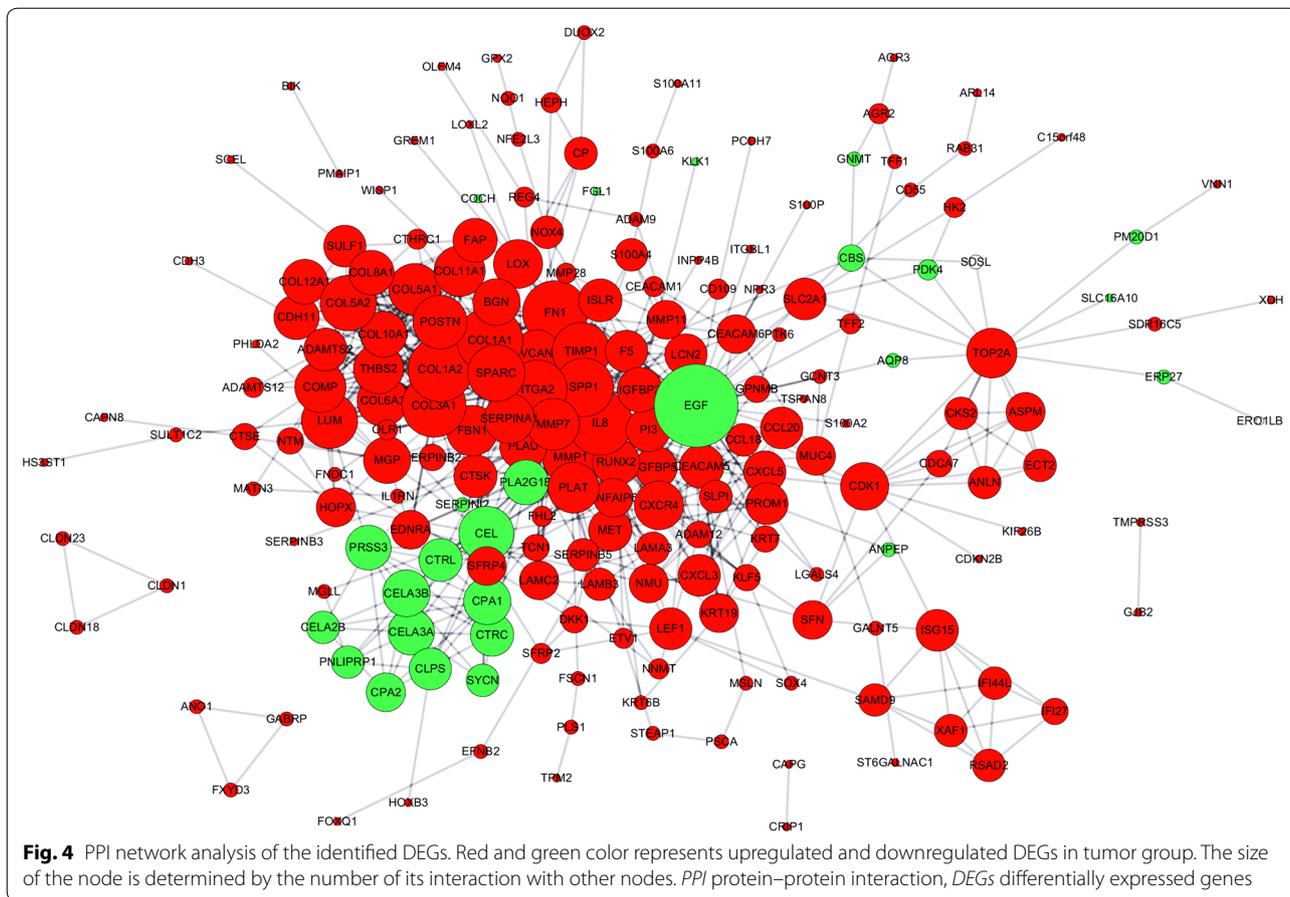


**Fig. 3** GO analysis and KEGG pathway analysis of the identified DEGs **a** GO analysis of the DEGs in cellular component, **b** GO analysis of the DEGs in biological process, **c** GO analysis of the DEGs in molecular function, **d** KEGG pathway analysis of the DEGs. *GO* Gene Ontology, *KEGG* Kyoto Encyclopedia of Genes and Genomes, *DEGs* differentially expressed genes

**Validation of COL11A1, GJB2 and CTRL with immunohistochemical assay of clinical samples**

We validated the expression level of COL11A1, GJB2 and CTRL with IHC assay of 46 paired pancreatic cancer and para-cancerous tissue sections (Table 2). COL11A1 and CTRL were chosen for further validation

from the top 5 upregulated and downregulated DEGs as they were also identified as key genes in our PPI analysis and had suitable antibodies for IHC assay. GJB2 was chosen as dysregulation of gap junction proteins is closely correlated with cancer progression and there was no published papers regarding the role of



**Table 1** Top 5 upregulated and downregulated genes of the identified differentially expressed genes

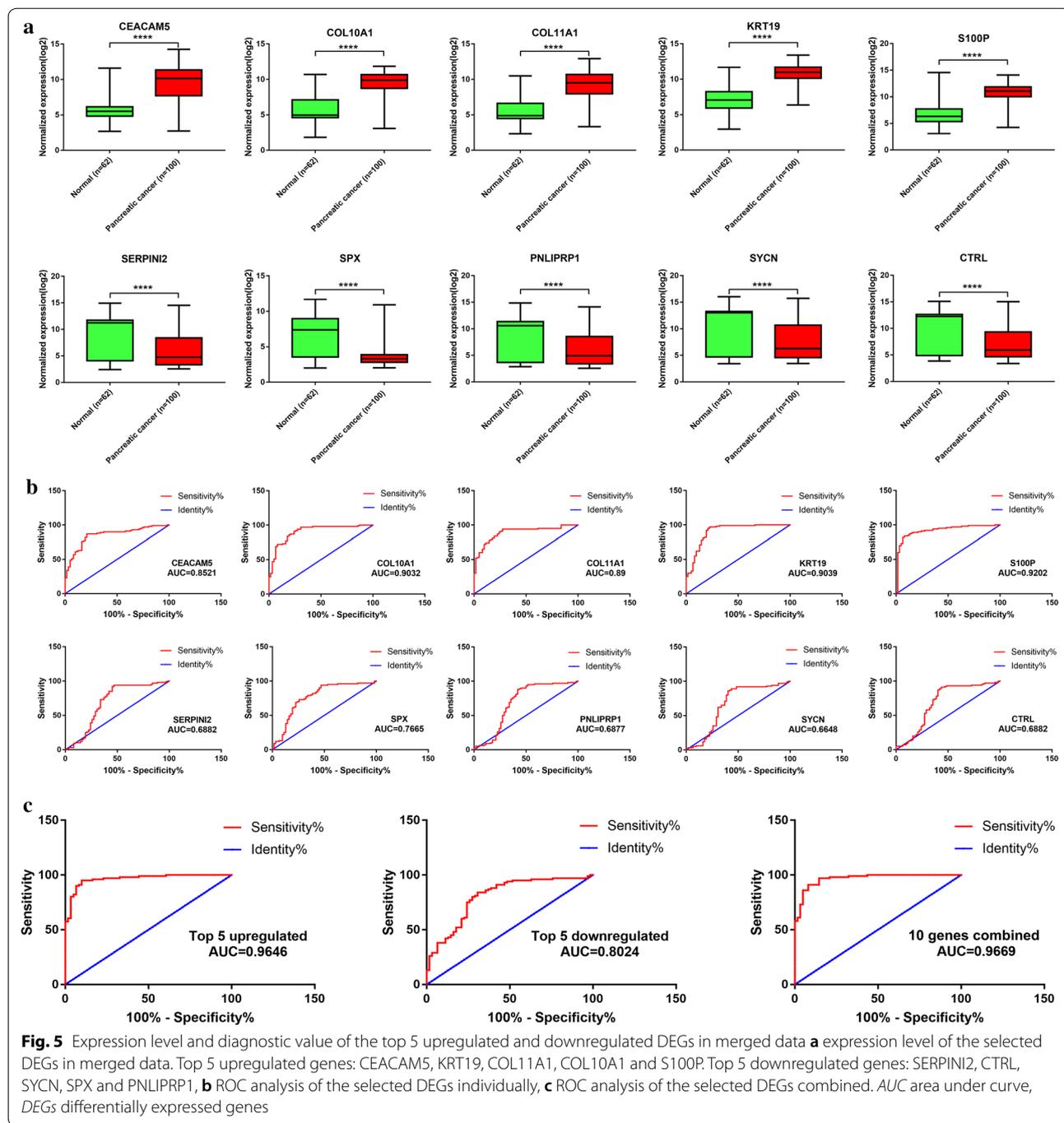
Top 5 upregulated			Top 5 downregulated		
Gene symbol	Log <sub>2</sub> Fold change	Adjusted p value	Gene symbol	Log <sub>2</sub> Fold change	Adjusted p value
S100P	3.99423324	1.47E–23	SERPINI2	–2.82817983	3.33E–06
COL10A1	3.860974743	2.67E–22	CTRL	–2.77212775	4.45E–06
COL11A1	3.71532016	5.14E–20	SYCN	–2.6228281	7.92E–05
KRT19	3.679278432	4.01E–24	SPX	–2.594566083	5.17E–10
CEACAM5	3.639738913	3.80E–15	PNLIPRP1	–2.520394006	1.02E–05

GJB2 in pancreatic cancer. IHC analysis showed consistent results with our bioinformatic analysis (Fig. 7a). Representative images of the IHC assay were uploaded in supplementary files (Additional file 2). We also analyzed the expression level of these three genes in different cancer stages (Fig. 7b). Next, we investigated the prognostic value of COL11A1, GJB2 and CTRL with our clinical and follow-up data (Fig. 7c). The result showed that high expression level of COL11A1, GJB2 or low expression level of CTRL in pancreatic cancer tissue sections indicated poorer prognosis and less

survival rate. These results suggested promising clinical predictive value of these three genes.

### Discussion

In the present study, we utilized multiple bioinformatic analysis and IHC analysis to investigate two key hypotheses. (1) Merged microarray data with bigger sample volume could reflect the biological characteristics of cancer group and normal group more effectively and accurately. We could obtain a more representative and accurate results from the merged microarray data. (2)



DEGs identified from the merged microarray data hold promising potential as key regulators and predictive biomarkers for pancreatic cancer. GSEA analysis of the merged data showed that the cancer group was enriched in genes closely correlated with cell proliferation, migration, invasion and angiogenesis. These results suggested that our merged data was qualified to reflect the characteristic expression profile of cancer and normal group.

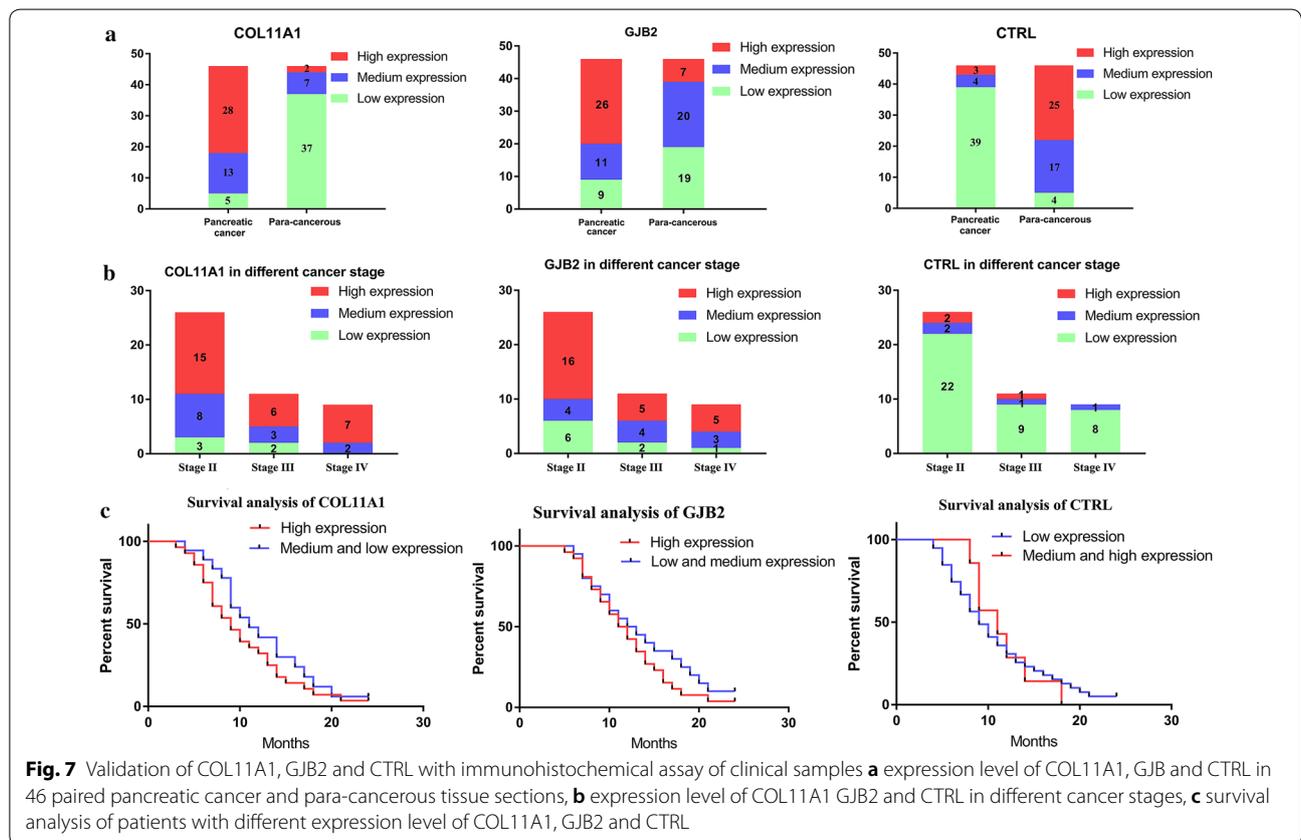
GO analysis and KEGG pathway analysis of the identified DEGs showed that they were closely correlated with ECM regulation and pancreatic secretion. These results were consistent with the clinical features of pancreatic cancer which is characterized by early local invasion/distant metastasis and loss of normal exocrine function. PPI network analysis showed that EGF was the core in the interaction network of identified DEGs. Multiple



**Table 2 Clinical characteristics of 46 paired pancreatic cancer tissue sections and para-cancerous tissue sections**

Patients characteristics	Number	Percentage
Gender		
Male	25	54%
Female	21	46%
Age		
Mean (years)	58.36 ± 9.48	
Histopathologic subtype		
Well-differentiated adenocarcinomas	19	41%
Moderately differentiated adenocarcinomas	10	22%
Poorly differentiated adenocarcinomas	17	37%
TNM staging		
T2N0M0	11	24%
T2N1M0	3	7%
T2N1M1	1	2%
T3N0M0	14	30%
T3N1M0	7	15%
T3N0M1	5	11%
T3N1M1	3	7%
T4N0M0	1	2%
T4N1M0	1	2%
Local invasion and distant metastasis		
Local invasion	14	31%
Regional lymph node metastasis	15	33%
Liver metastasis	6	13%
Ascites metastasis	2	4%

clinical trials on investigating the efficacy of combining Cetuximab (a targeted antibody against EGFR) with currently applied therapies have been reported [25–27]. All of the aforementioned results suggested that our merged microarray data could effectively and accurately reflect the characteristic biological differences between pancreatic cancer tissues and adjacent normal pancreatic tissues. These identified DEGs of our study could be potential candidates for novel predictive biomarkers or targets for chemotherapy. Therefore, we first validated the diagnostic and prognostic value of the top 5 upregulated and down regulated DEGs with public data from the GEO and TCGA database. We further validated the clinical predictive value of COL11A1, GJB2 and CTRL with IHC analysis of our own clinical derived sections. COL11A1 has been reported to be secreted by cancer associated fibroblasts and is closely correlated with the progression of multiple cancer types. The relationship between GJB2, CTRL and cancer progression was not reported. Furthermore, no published paper evaluated the expression and predictive value of COL11A1, GJB2 and CTRL in pancreatic cancer. Here, the results of IHC assay and survival analysis showed that these three genes could serve as predictive biomarkers for pancreatic cancer. Further functional analysis these three genes with in vitro and in vivo experiments are required in our future studies.



**Fig. 7** Validation of COL11A1, GJB2 and CTRL with immunohistochemical assay of clinical samples **a** expression level of COL11A1, GJB and CTRL in 46 paired pancreatic cancer and para-cancerous tissue sections, **b** expression level of COL11A1 GJB2 and CTRL in different cancer stages, **c** survival analysis of patients with different expression level of COL11A1, GJB2 and CTRL

## Conclusions

In this study, we showed that merged microarray data with bigger sample volume could reflect the biological characteristics of pancreatic cancer more effectively and accurately. Results of bioinformatic analysis and IHC analysis suggested that COL11A1, GJB2 and CTRL are novel predictive biomarkers for pancreatic cancer.

## Additional files

**Additional file 1.** Identified differentially expressed genes of the merged data.

**Additional file 2.** Representative images of IHC assay. A. IHC assay of COL11A1 B. IHC assay of GJB2 C. IHC assay of CTRL.

## Abbreviations

TCGA: The Cancer Genome Atlas; DEGs: differentially expressed genes; GEO: Gene Expression Omnibus; GSEA: Gene Set Enrichment Analysis; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; PPI: protein-protein interaction; ROC: receiver operating characteristic curve; COL11A1: collagen alpha-1(XI) chain; GJB2: gap junction beta-2 protein; CTRL: chymotrypsin-like protease CTRL-1; IHC: immunohistochemical; DAVID: Database for Annotation, Visualization and Integrated Discovery; CC: cellular component; BP: biological process; MF: molecular function; ECM: extracellular matrix; EGF: epidermal growth factor.

## Authors' contributions

DS and HJ analyzed the data and write the paper. JZ performed IHC assay and analyzed the data. XT designed the experiment and revised the paper. All authors read and approved the final manuscript.

## Author details

<sup>1</sup> Shengjing Hospital of China Medical University, Shenyang 110004, People's Republic of China. <sup>2</sup> Gastric Cancer Department, Liaoning Province Cancer Hospital & Institute (Cancer Hospital of China Medical University), Shenyang 110004, People's Republic of China. <sup>3</sup> Thyroid and Pancreatic Surgery Ward, Shengjing Hospital of China Medical University, Shenyang 110004, People's Republic of China.

## Acknowledgements

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

## Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

## Consent for publication

Not applicable.

## Ethics approval and consent to participate

Collection and usage of clinical tissue samples in this study was approved by the local ethics committee of the Shengjing Hospital (certificate number 2017PS24K). Fifty-six qualified paraffin-embedded tissue sections with paired para-cancerous tissue sections were selected from the pathological database of Shengjing Hospital between 2013 and 2016. The following standards were applied for qualified sample inclusion. (1) Patients received regular preoperative preparation and surgical operation without preoperative surgical, radical or chemical therapy. (2) Samples with complete clinical and follow-up data. (3) Sections were in intact condition with proper storage.

## Funding

The present study was funded by the Outstanding Scientific Fund of Shengjing Hospital (Grant Number M731).

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 13 September 2018 Accepted: 27 October 2018

Published online: 06 November 2018

## References

- Hidalgo M. Pancreatic cancer. *N Engl J Med*. 2010;362(17):1605–17.
- Jin H, Wu Y, Tan X. The role of pancreatic cancer-derived exosomes in cancer progress and their potential application as biomarkers. *Clin Transl Oncol*. 2017;19(8):921–30.
- Ferro R, Falasca M. Emerging role of the KRAS-PDK1 axis in pancreatic cancer. *World J Gastroenterol*. 2014;20(31):10752–7.
- Liu P, Weng Y, Sui Z, Wu Y, Meng X, Wu M, Jin H, Tan X, Zhang L, Zhang Y. Quantitative secretomic analysis of pancreatic cancer cells in serum-containing conditioned medium. *Sci Rep*. 2016;6:37606.
- Mello SS, Valente LJ, Raj N, Seoane JA, Flowers BM, McClendon J, Biegging-Rolett KT, Lee J, Ivanochko D, Kozak MM, et al. A p53 super-tumor suppressor reveals a tumor suppressive p53-Ptpn14-Yap axis in pancreatic cancer. *Cancer cell*. 2017;32(4):460 e466–473 e466.
- Zhao X, Wang X, Fang L, Lan C, Zheng X, Wang Y, Zhang Y, Han X, Liu S, Cheng K, et al. A combinatorial strategy using YAP and pan-RAF inhibitors for treating KRAS-mutant pancreatic cancer. *Cancer Lett*. 2017;402:61–70.
- Strnadel J, Choi S, Fujimura K, Wang H, Zhang W, Wyse M, Wright T, Gross E, Peinado C, Park HW, et al. eIF5A-PEAK1 signaling regulates YAP1/TAZ protein expression and pancreatic cancer cell growth. *Can Res*. 2017;77(8):1997–2007.
- Pongsuchart M, Kuchimaru T, Yonezawa S, Tran DTP, Kha NT, Hoang NTH, Kadonosono T, Kizaka-Kondoh S. Novel lymphoid enhancer-binding factor 1-cytoglobin axis promotes extravasation of osteosarcoma cells into the lungs. *Cancer Sci*. 2018;109(9):2746–56.
- Pei H, Li L, Fridley BL, Jenkins GD, Kalari KR, Lingle W, Petersen G, Lou Z, Wang L. FKBP51 affects cancer cell response to chemotherapy by negatively regulating Akt. *Cancer Cell*. 2009;16(3):259–66.
- Donahue TR, Tran LM, Hill R, Li Y, Kovochich A, Calvopina JH, Patel SG, Wu N, Hindoyan A, Farrell JJ, et al. Integrative survival-based molecular profiling of human pancreatic cancer. *Clin Cancer Res*. 2012;18(5):1352–63.
- Berger AC, Korkut A, Kanchi RS, Hegde AM, Lenoir W, Liu W, Liu Y, Fan H, Shen H, Ravikumar V, et al. A comprehensive pan-cancer molecular study of gynecologic and breast cancers. *Cancer cell*. 2018;33(4):690 e699–705 e699.
- Cheng W, Ren X, Zhang C, Cai J, Liu Y, Han S, Wu A. Bioinformatic profiling identifies an immune-related risk signature for glioblastoma. *Neurology*. 2016;86(24):2226–34.
- Tsuyoshi H, Yoshida Y. Molecular biomarkers for uterine leiomyosarcoma and endometrial stromal sarcoma. *Cancer Sci*. 2018;109(6):1743–52.
- Wang Z, Yang B, Zhang M, Guo W, Wu Z, Wang Y, Jia L, Li S, Xie W, Cancer Genome Atlas Research N, et al. lncRNA epigenetic landscape analysis identifies EPIC1 as an oncogenic lncRNA that interacts with MYC and promotes cell-cycle progression in cancer. *Cancer cell*. 2018;33(4):706 e709–720 e709.
- Cheng W, Zhang C, Ren X, Jiang Y, Han S, Liu Y, Cai J, Li M, Wang K, Liu Y, et al. Bioinformatic analyses reveal a distinct Notch activation induced by STAT3 phosphorylation in the mesenchymal subtype of glioblastoma. *J Neurosurg*. 2017;126(1):249–59.
- Cheng W, Li M, Jiang Y, Zhang C, Cai J, Wang K, Wu A. Association between small heat shock protein B11 and the prognostic value of MGMT promoter methylation in patients with high-grade glioma. *J Neurosurg*. 2016;125(1):7–16.
- Hayes J, Thygesen H, Tumilson C, Droop A, Boissinot M, Hughes TA, Westhead D, Alder JE, Shaw L, Short SC, et al. Prediction of clinical outcome in

- glioblastoma using a biologically relevant nine-microRNA signature. *Mol Oncol*. 2015;9(3):704–14.
18. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102(43):15545–50.
  19. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet*. 2000;25(1):25–9.
  20. The Gene Ontology C. Expansion of the Gene Ontology knowledgebase and resources. *Nucleic acids research*. 2017;45(D1):D331–8.
  21. Jiao X, Sherman BT, da Huang W, Stephens R, Baseler MW, Lane HC, Lempicki RA. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics*. 2012;28(13):1805–6.
  22. Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC, Lempicki RA. DAVID: database for annotation, visualization, and integrated discovery. *Genome Biol*. 2003;4(5):P3.
  23. 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 Res*. 2003;13(11):2498–504.
  24. Jin H, Liu P, Wu Y, Meng X, Wu M, Han J, Tan X. Exosomal zinc transporter ZIP4 promotes cancer growth and is a novel diagnostic biomarker for pancreatic cancer. *Cancer Sci*. 2018;109(9):2946–56.
  25. Esnaola NF, Chaudhary UB, O'Brien P, Garrett-Mayer E, Camp ER, Thomas MB, Cole DJ, Montero AJ, Hoffman BJ, Romagnuolo J, et al. Phase 2 trial of induction gemcitabine, oxaliplatin, and cetuximab followed by selective capecitabine-based chemoradiation in patients with borderline resectable or unresectable locally advanced pancreatic cancer. *Int J Radiat Oncol Biol Phys*. 2014;88(4):837–44.
  26. Hong JY, Nam EM, Lee J, Park JO, Lee SC, Song SY, Choi SH, Heo JS, Park SH, Lim HY, et al. Randomized double-blinded, placebo-controlled phase II trial of simvastatin and gemcitabine in advanced pancreatic cancer patients. *Cancer Chemother Pharmacol*. 2014;73(1):125–30.
  27. Fensterer H, Schade-Brittinger C, Muller HH, Tebbe S, Fass J, Lindig U, Settmacher U, Schmidt WE, Marten A, Ebert MP, et al. Multicenter phase II trial to investigate safety and efficacy of gemcitabine combined with cetuximab as adjuvant therapy in pancreatic cancer (ATIP). *Ann Oncol*. 2013;24(10):2576–81.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

