Genomics Proteomics Bioinformatics 12 (2014) 276–283

HOSTED BY

ELSEVIER

Genomics Proteomics Bioinformatics

www.elsevier.com/locate/gpb
www.sciencedirect.com

ORIGINAL RESEARCH

Identification of Differentially-expressed Genes in Intestinal Gastric Cancer by Microarray Analysis

Shizhu Zang [1,#,§], Ruifang Guo [1,#,‡], Rui Xing [1,#], Liang Zhang [2], Wenmei Li [1], Min Zhao [1], Jingyuan Fang [3], Fulian Hu [4], Bin Kang [1], Yonghong Ren [2], Yonglong Zhuang [2], Siqi Liu [5], Rong Wang [6], Xianghong Li [7], Yingyan Yu [8], Jing Cheng [2], Youyong Lu [1,*]

[1] *Laboratory of Molecular Oncology, MOE Key Laboratory of Carcinogenesis and Translational Research, Peking University Cancer Hospital & Institute, Beijing 100142, China*
[2] *National Engineering Research Center for Beijing Biochip Technology and Tsinghua University School of Medicine, Beijing 102206, China*
[3] *Department of Gastroenterology, Renji Hospital of Shanghai Second Medical University, Shanghai 200001, China*
[4] *Department of Gastroenterology, Peking University First Hospital, Beijing 100034, China*
[5] *Beijing Institutes of Genomics, Chinese Academy of Sciences, Beijing 100101, China*
[6] *Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York 10029, USA*
[7] *Department of Pathology, MOE Key Laboratory of Carcinogenesis and Translational Research, Peking University Cancer Hospital & Institute, Beijing 100142, China*
[8] *Shanghai Ruijin Hospital, Shanghai Jiao Tong University, School of Medicine, Shanghai 200025, China*

Received 25 June 2014; revised 18 September 2014; accepted 30 September 2014
Available online 11 December 2014

Handled by Xiangdong Fang

KEYWORDS

Gastric cancer development;
Microarray;
Gene expression profile

Abstract Gastric cancer (GC) is one of the most frequent malignant tumors. In order to systematically characterize the cellular and molecular mechanisms of intestinal GC development, in this study, we used 22 K oligonucleotide microarrays and bioinformatics analysis to evaluate the gene expression profiles of GC in 45 tissue samples, including 20 intestinal GC tissue samples,

* Corresponding author.
E-mail: youyonglu@bjmu.edu.cn (Lu Y).
# Equal contribution.
§ Current address: Biotechnology Department, Dalian Medical University, Dalian 116044, China.
‡ Current address: Gastroenterology Department, Inner Mongolian People's Hospital, Hohhot 010017, China.

Peer review under responsibility of Beijing Institute of Genomics, Chinese Academy of Sciences and Genetics Society of China.

http://dx.doi.org/10.1016/j.gpb.2014.09.004

20 normal appearing tissues (NATs) adjacent to tumors and 5 noncancerous gastric mucosa tissue samples. These profiles allowed us to explore the transcriptional characteristics of GC and determine the change patterns in gene expression that may be of clinical significance. 1519 and 1255 differentially-expressed genes (DEGs) were identified in intestinal GC tissues and NATs, respectively, as determined by Bayesian analysis ($P < 0.001$). These genes were associated with diverse functions such as mucosa secretion, metabolism, proliferation, signaling and development, which occur at different stages of GC development.

Introduction

Gastric cancer (GC) is one of the most common malignant tumors. It is estimated that one million new cases are reported worldwide each year [1], with around two-thirds of GC occurring in developing countries. Although considerable effort has been directed toward the development of surgical and chemotherapeutic interventions, the prognosis for patients with advanced stages of GC remains poor. Thus, a major challenge toward assessing and, perhaps, improving the clinical outcome of the treatment of GC patients is to better understand the molecular basis of the disease and its development, *i.e.*, the key changes of gene expression patterns in gastric tumorigenesis. The relationship between specific gene expression patterns and certain properties of GC have been previously described, including resistance to chemotherapeutics [2], metastatic potential [3,4] and prognosis following a particular treatment [5,6]. However, to understand the underlying mechanisms of gastric tumorigenesis, it is essential to characterize the biological processes that initiate the development of GC and its subsequent progression, especially, to document the gene expression pattern from a high-risk population of GC. In this study, we have characterized the transcriptional profiles of GC in Chinese patients by using 22 K oligonucleotide microarrays and have identified differentially-expressed genes (DEGs) within GC, GC adjacent and normal tissues. These expression patterns were further examined by identifying molecular pathways associated with GC development.

Results

Genes differentially expressed between intestinal GC and normal gastric mucosa tissues

To understand the expression profile of intestinal GC, we collected 45 tissue samples, including 20 intestinal GC tissue samples, 20 normal appearing tissues (NATs) adjacent to tumors and five noncancerous gastric mucosa tissue samples, and performed microarray study to evaluate the gene expression profile. After normalization, a Bayesian analysis of gene expression level (BAGEL) was used to characterize differential gene expression between intestinal GC tissue samples and noncancerous gastric mucosa tissue samples with a significance cutoff of $P < 0.001$. A total of 1519 genes were recognized to be differentially expressed in intestinal GC when compared to normal gastric mucosa tissue (Figure S1). These included 593 upregulated and 926 downregulated genes. Hierarchical clustering of these DEGs demonstrated a dramatic variation in gene expression in tumors compared with normal gastric mucosa tissue. Tables S1 and S2 list P value, fold change and name of each DEG. Among these DEGs, 11 and 29 genes were upregulated and downregulated with fold change >10, respectively. The annotation analysis from GoMiner indicated that some of these genes were related with gastric physiological function, such as ATPase, somatostatin and gastrin.

Prediction of tumor-specific biological characteristics associated with DEGs in intestinal GC

Based on gene expression profile of GC, we were able to identify tumor-specific biological characteristics that correlate with DEGs. Using currently available chip annotation tools, including DAVID, SOURCE and the high-throughput GoMiner, we obtained the functional description, classification and location of the DEGs. Annotation results showed that these 1183 DEGs were known genes associated with a diverse set of biological pathways and functions in both cell- and organ-specific physiological processes (**Table 1**).

Prediction of tumor-specific pathways associated with gene expression profiling in intestinal GC

Signal transduction pathways associated with gene expression changes were analyzed and defined using the bioresource for array genes (BioRag, www.biorag.org). A total of 143 signal transduction pathways contained genes that were differentially expressed. Among them, 14 pathways showed altered expression of at least three up-regulated genes within each (**Table 2**). These pathways include the MAPK signaling pathway, inflammatory response pathway, TGF-β pathway and pathways associated with extracellular matrix synthesis and regulation of gluconeogenesis. Six pathways were changed with at least three downregulated genes (**Table 3**).

Genes differentially expressed between NATs and normal gastric mucosa tissues

Although NATs appear morphologically normal, Figure S2 demonstrated that the gene expression pattern of these tissues is very different from that of normal gastric mucosa tissues. A total of 1255 DEGs, including 561 upregulated and 694 downregulated genes, were identified with a $P < 0.001$ significance cutoff. The detailed information describing the upregulated and downregulated genes is presented in Table S3 and Table S4.

Comparison of gene expression patterns between GC tumors and NATs

The number of DEGs in GC and NATs is shown and clustered according to different fold changes in **Figure 1**. Our data above

Table 1 Functional alterations due to DEGs in intestinal GC samples (FDR < 0.05)

Level	Functional category	Change status	FDR	Number of DEGs
Cell level	Cell motility	↑	0.045192	52
	Cell proliferation	↑	0.005	73
	Cell communication	↑	< 0.0001	73
	Glycoprotein metabolism	↓	< 0.0001	27
	Cellular lipid metabolism	↓	0.011739	90
	Cellular carbohydrate metabolism	↓	0.012273	41
	Protein amino acid glycosylation	↓	< 0.0001	8
Organ level	Development	↑	0.001429	173
	Morphogenesis	↑	0.001667	45
	Immune response	↑	0.001333	44
	Neurophysiological process	↑	0.025385	34
	Digestion	↓	< 0.0001	15

Note: False discovery rate (FDR) is used to assess the significance of a particular category in GoMiner. DEG, differentially-expressed gene.

Table 2 List of signaling pathways affected with at least three upregulated genes

Pathway	Number	Genes
MAPK signaling pathway	12	*STK3*, *DUSP6*, *RASA2*, *MYC*, *NR4A1*, *GADD45B*, *CDC25B*, *FOS*, *PDGFRB*, *PLA2G2A*, *DUSP1*, *ARAF1*
Inflammatory response pathway	6	*COL1A2*, *LAMC1*, *FN1*, *COL3A1*, *COL1A1*, *LAMB1*
Complement and coagulation cascades	7	*DAF*, *PLAU*, *TFPI*, *SERPINE1*, *C1R*, *PROS1*, *C1S*
Role of EGF receptor transactivation by GPCRs in cardiac hypertrophy	5	*ADAM12*, *FOS*, *MYC*, *EDNRA*, *EDN1*
Glycolysis and gluconeogenesis	3	*TGIF*, *SERPINE1*, *THBS1*
Inhibition of matrix metalloproteinases	4	*TIMP3*, *TIMP2*, *TIMP1*, *MMP2*
TGF-β signaling pathway	4	*INHBB*, *MYC*, *FST*, *THBS1*
Wnt signaling	5	*WNT16*, *PLAU*, *SFRP4*, *MYC*, *SFRP2*
Toll-like receptor signaling pathway	4	*CXCL9*, *IL8*, *CCL3*, *FOS*
Prostaglandin and leukotriene metabolism	4	*PTGS2*, *PTGIS*, *PLA2G2A*
Rac 1 cell motility signaling pathway	4	*CHN1*, *NCF2*, *MYLK*, *TRIO*
Acute myocardial infarction	3	*COL4A1*, *PROS1*, *TFPI*
Cell cycle	3	*GADD45B*, *E2F3*, *CDC25B*
Mechanism of gene regulation by peroxisome proliferators via PPAR α	3	*PTGS2*, *MYC*, *DUSP1*

Table 3 List of signaling pathways affected with at least three downregulated genes

Pathway	Number	Gene
Electron transport chain	11	*COX17*, *NDUFC2*, *COX5B*, *GPRC5B*, *COX7A2L*, *PHKB*, *NDUFB8*, *NDUFA8*, *NDUFA10*, *ATP5O*, *ADRA2A*
Glycolysis and gluconeogenesis	10	*GSN*, *RHOA*, *FGFR2*, *ALDOC*, *FBP1*, *FBP2*, *SLC9A1*, *HK1*, *CFL1*, *PIP5K1B*
Rho cell motility signaling pathway	5	*GSN*, *RHOA*, *CFL1*, *PIP5K1B*, *ARPC2*
HIV-I Nef: negative effector of Fas and TNF	4	*GSN*, *CASP7*, *CASP6*, *PRKCD*
Y branching of actin filaments	3	*PIR*, *ARPC2*, *WASL*
Role of PI3 K subunit p85 in regulation of actin organization and cell migration	3	*RHOA*, *ARPC2*, *WASL*

indicated that we identified 1519 and 1255 DEGs for GC tumors (*vs.* normal) and NATs (*vs.* normal), respectively. Among these DEGs, 550 genes were shared between GC tumors (*vs.* normal) and NATs (*vs.* normal), including 169 upregulated and 381 downregulated genes (Figure S3, Figure 2). The annotation results showed that these genes are involved in a number of different pathways and processes, such as amino acid glycosylation, biopolymer glycosylation and glycoprotein biosynthesis. Among the different genes between NATs and GC, nine genes were reversely expressed in two groups (Table 4), including *ATP4B*, *ATP4A*, *CCKBR*, *KCNJ16*, *SIAT1*, *DUOX1*, *WDR37* and *APOBEC2*.

Characterization of the altered expression of genes associated with mucosal barrier function and secretion of GC tissues and NATs

Disruption of the mucosal barrier is associated with some primary gastric disorders [7]. Therefore, we analyzed the

Table 4 Opposite pattern of gene expression in NAT and GC samples

Accession	Symbol	GC			NAT			Protein name
		Expression	FC	*P* value	Expression	FC	*P* value	
NM_000705	*ATP4B*	Down	31.47	< 0.0001	Up	9.38	< 0.0001	ATPase, H + /K + exchanging, beta polypeptide
NM_000704	*ATP4A*	Down	15.96	< 0.0001	Up	3.05	< 0.0001	ATPase, H + /K + exchanging, alpha polypeptide
NM_000731	*CCKBR*	Down	4.49	< 0.0001	Up	2.34	< 0.0001	Cholecystokinin B receptor
NM_018658	*KCNJ16*	Down	3.79	< 0.0001	Up	4.49	< 0.0001	Potassium inwardly-rectifying channel, subfamily J, member 16
NM_003032	*SIAT1*	Down	2.86	< 0.0001	Up	5.04	< 0.0001	ST6 beta-galactosamide alpha-2,6-sialyltranferase 1
NM_017434	*DUOX1*	Down	2.34	< 0.0001	Up	3.2	< 0.0001	Dual oxidase 1
NM_014023	*WDR37*	Down	1.59	< 0.0001	Up	2.23	< 0.0001	WD repeat domain 37
NM_006789	*APOBEC2*	Down	1.47	0.0002	Up	3.24	< 0.0001	Apolipoprotein B mRNA editing enzyme, catalytic polypeptide-like 2

Note: Fold changes (FC) are the absolute value of $\log_2$ (GC expression/normal expression) or the $\log_2$ (NAT expression/normal expression).

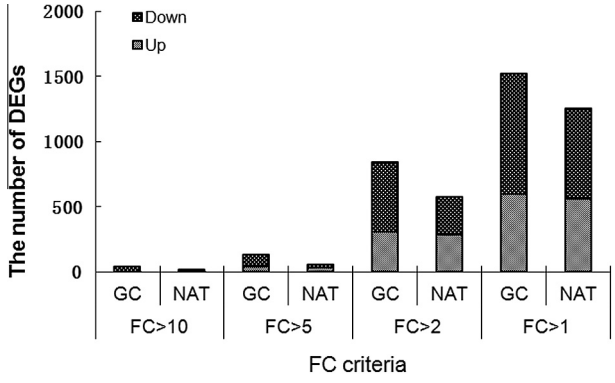

Figure 1 Number of DEGs in GC samples and NAT under various FC criteria

More DEGs were detected in GC than NAT, when compared to the normal tissues. DEG, differentially-expressed gene; NAT, normal appearing tissue adjacent to tumor; FC, fold change.

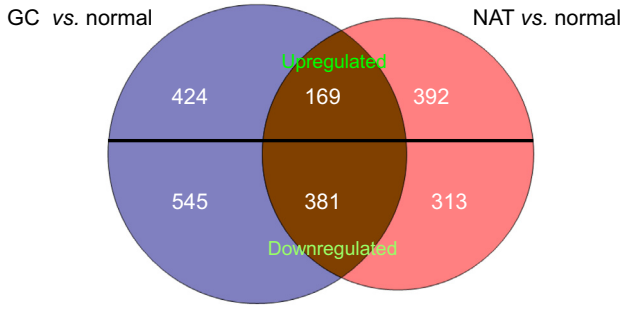

Figure 2 Distribution of DGEs between NAT and intestinal GC samples

Number of DEGs in NAT and intestinal GC samples. The overlapping regions show the shared numbers of DEGs in two tissues. Blue represents GC tissue and red represents NAT.

expression of all mucosal barrier-related genes. As a result, we identified 18 candidates that were differentially expressed, including 6 genes that were downregulated in NATs (Table 5) and 12 genes that were dramatically downregulated in GC tissues (Table 6). This result indicated that both NATs and GC tissues lose some normal physiological functions, with the alteration in the expression of barrier function-related genes being more frequent in NATs, followed by the alteration in the expression of secretion-related genes in the GC tissues.

A tumorigenic model based on selected genes and their networks in GC

In this study, we integrated the inference of gene expression profile with the known clinical features of stomach cancer over time to predict cancer progression (Figure 3). The information we could infer from the gene expression profiles in this study is listed below. (1) Characterization of altered gene expression of genes involved in barrier and secretion of GC tumors and NATs. (2) Cellular metabolism and cell function alterations due to DEGs. (3) The changes of signal transduction pathways due to DEGs. The known clinical features of stomach cancer over time are listed below. (1) Before cancer develops, pre-cancerous changes often occur in the inner lining (mucosa) of the stomach [8,9]. These early changes rarely cause symptoms and therefore often go undetected. (2) Stomach cancers tend to develop slowly over time. (3) Stomach cancers can spread (metastasize) in different ways. They can grow through the wall of the stomach and invade nearby organs. Finally, abnormalities that are related to cell proliferation, tissue architecture and remodeling, as well as differentiation and development emerge as cancer progresses.

Discussion

A key goal of cancer studies is to systematically characterize the cellular and molecular mechanisms involved in disease progression, and consequently, to identify potential biomarkers for predicting the prognosis of patients [10]. Genomic and proteomic studies on GCs have predominantly been reported from the Western countries and Japan [11–13]; only a few studies on GC have been reported from China [14,15]. Molecular pathology of GC may vary among populations, which is likely due to differential exposure to disease risk factors including diet, *Helicobacter pylori* variants, smoking, alcohol consumption and other underlying medical conditions. In this study, our objective was to establish a comprehensive gene expression profile for GC in Chinese patients, a known high-risk population, and to investigate the mechanisms underlying GC carcinogenesis.

Table 5 DEGs involved in normal biological functions of gastric mucosa in NAT

Gene description	Accession	Symbol	Fold change	*P* value
Intrinsic barrier				
Occludin	NM_002538	*OCLN*	1.812	0.0001
Catenin	NM_001903	*CTNNA1*	1.999	< 0.0001
Extrinsic barrier				
Mucin 7, salivary	L13283	*MUC7*	1.687	0.0001
Defensin, beta 1	NM_005218	*DEFB1*	5.568	< 0.0001
Gastric secretion-related genes				
Gastrin	NM_000805	*GAS*	9.916	< 0.0001
Carbonic anhydrase II	NM_000067	*CA2*	2.424	< 0.0001

Note: Fold changes are the absolute value of $\log_2$ (NAT expression/normal expression).

Table 6 DEGs involved in normal biological functions of gastric mucosa in intestinal GC samples

Gene description	Accession	Symbol	Fold change	*P* value
Intrinsic barrier				
Occludin	NM_002538	*OCLN*	2.63301	< 0.0001
Claudin 18	NM_016369	*CLDN18*	4.96291	< 0.0001
Claudin 23	BC016047	*CLDN23*	2.88944	< 0.0001
Catenin	NM_001903	*CTNNA1*	2.00327	< 0.0001
Extrinsic barrier				
Mucin 1, transmembrane	J05582	*MUC1*	2.27442	< 0.0001
Mucin 7, salivary	L13283	*MUC7*	1.57293	< 0.0001
Trefoil factor 2	NM_005423	*TFF2*	9.65021	< 0.0001
Defensin, beta 1	NM_005218	*DEFB1*	7.9556	< 0.0001
Gastric secretion-related genes				
Carbonic anhydrase II	NM_000067	*CA2*	13.20503	< 0.0001
Progastricsin	NM_002630	*PGC*	6.94681	< 0.0001
Gastrin	NM_000805	*GAS*	13.5022	< 0.0001
Cholecystokinin B receptor	NM_000731	*CCKBR*	4.49452	< 0.0001

Note: Fold changes are the absolute value of $\log_2$ (GC expression/normal expression).

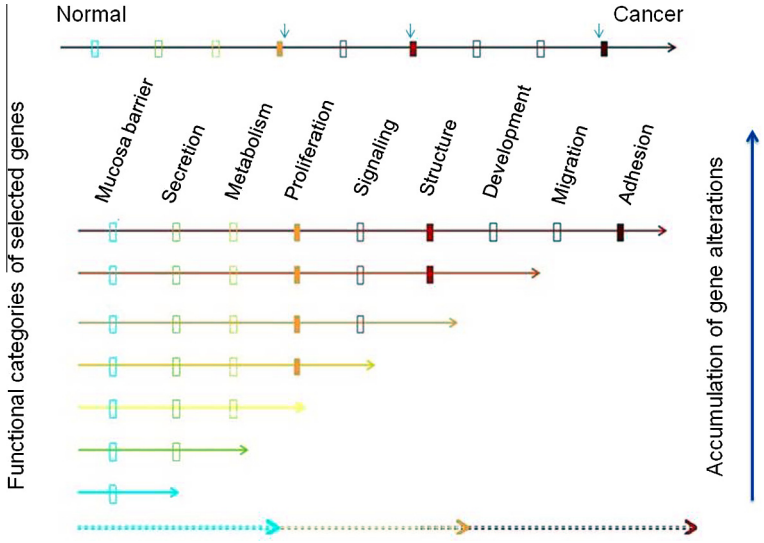

Figure 3 A tumorigenic model for GC development and progression

A tumorigenic model depicts GC development and progression. Dysfunction of the mucosal barrier appears to occur as an early event. Subsequently, metabolic abnormalities appear, following excessive DNA replication. Finally, abnormalities related to cell proliferation, tissue architecture and remodeling, as well as differentiation and development emerge as the tumor progresses. The vertical arrows mean accumulation of gene alterations. Each functional category was shown as a color code.

In this study, we performed a 22 K oligonucleotide microarray analysis on 20 intestinal GC tissue samples and 20 NATs adjacent to tumors of intestinal GC in comparison with the five normal samples, respectively. Gene expression profiles were generated for samples from intestinal GC, NATs and normal tissues. Following a systemic analysis, we integrated the gene expression patterns, functional annotations and clinical characteristics that were associated with GC. Furthermore, we identified the expression profiles of genes involved in functional pathways associated with gastric mucosal barrier function and gastric secretion. Altogether, our data provided insights into gastric carcinogenesis and proposed new leads toward the discovery of possible biomarkers for the early diagnosis of GC.

Our study has generated comprehensive gene expression profiles of GC. Our data showed that a number of DEGs may be important in the development of intestinal GC. We found that nine genes were reversely expressed between NATs and GC, including *ATP4B*, *ATP4A*, *CCKBR*, *KCNJ16*, *SIAT1*, *DUOX1*, *WDR37* and *APOBEC2*. These genes were involved in oxidative phosphorylation, N-glycan biosynthesis, atrazine degradation and glycan structures biosynthesis. This result provides some new clues to explore the mechanisms of GC development.

To our knowledge, this report, for the first time, comprehensively described the alterations in the gene expression patterns of NATs and intestinal GC tumors compared to normal tissues. DEGs shared by GC tumors (*vs*. normal tissues) and NATs (*vs*. normal tissues) might be essential for tumorigenesis. These differences may reflect the significant roles of tissue architecture and remodeling that occurs during GC development. Similarly, we have also identified a group of DEGs, including *EGR1*, *CYR61* and *ADAMTS1*, which are differentially expressed between intestinal metaplasia/dysplasia and normal mucosa [16]. However, the dramatic changes in extracellular matrix and cellular communication pathways that were noticed in tumor tissues were rarely observed in NATs. These data suggest that tissue remodeling may be essential for tumor development, possibly by maintaining the environment around the tumor cells and supporting cellular survival.

Importantly, we have also identified distinct gene expression patterns related to abnormal physiological functions of the gastric mucosa, including gastric barrier and secretion. The gastrointestinal barrier has two components, the intrinsic and extrinsic barriers [17]. The intrinsic barrier is composed of epithelial cells lining the digestive tract and the tight junctions that tie them together. The extrinsic barrier comprises secretions and other factors that are not physically part of the epithelium, but affect the epithelial cells and maintain their barrier function, such as mucus, trefoil proteins and defensins. Our data show that normal gastric physiological functions are lost in both the gastric tumor and NATs. In NATs, genes related to barrier function are almost all downregulated; while in the GC samples, mainly the secretion-related genes are downregulated. Alterations in gene expression in NATs indicate that lesions of the gastric mucosal barrier might aggravate the mucosal damage insulted by various cancer-related risk factors.

A tumorigenic model was created to depict GC development and progression. At the initial stage, dysfunction of the mucosal barrier appears to occur as an early event. Then, metabolic abnormalities occur following excessive DNA replication, which would result in genomic instabilities. Finally, abnormalities that are related to cell proliferation, as well as tissue architecture, remodeling, differentiation and development emerge as the cancer progresses. In future study, we will collect pre-cancer and gastric cancer samples for gene expression analysis to examine this model.

Materials and methods

Specimens and RNA isolation

All specimens were obtained with the written consent of the patients and were histologically graded, according to the guidelines put forth by the Union for International Cancer Control (UICC). The study was approved by the Committee on Human Studies, School of Oncology, Peking University. Specimens of GC (verified by histopathology) were obtained from the Tissue Bank of Peking University, School of Oncology. We selected 20 intestinal GC tissue samples, 20 tumor-adjacent NAT samples, and 5 normal gastric mucosa tissue samples for microarray experiments. RNA was extracted with TRIzol reagent (Life Technologies, Inc., Gaithersburg, MD) and was further purified using a Nucleospin RNA Clean-up Kit (Clontech Laboratories, Inc., Mountain View, CA), according to the manufacturer's instructions. The RNA quality was assessed by formaldehyde agarose gel electrophoresis and was quantitated spectrophotometrically.

Microarray fabrication

The human genome oligonucleotide microarray was prepared in CapitalBio Corporation (Beijing, China). A Human Genome Oligo Set Version 2.1, consisting of approximately 22,000 human genes, was purchased from Qiagen Operon (Germantown, MD, USA). As external controls, an additional eight oligonucleotides that recognize yeast intergenic sequences were included together with the internal controls provided by the manufacturer. All nucleotides were dissolved in 50% DMSO to obtain a final concentration of 40 μM and were printed on in-house manufactured aminosilanated glass slides. Arrays were fabricated using a SmartArrayTM microarrayer (CapitalBio Corp., Beijing, China). After printing, the slides were baked at 80 °C for 1 h and stored dry at room temperature until use. Prior to hybridization, the slides were rehydrated at 65 °C for 10 s, snap dried on a 100 °C heating block for 5 s, and UV cross-linked at 250 mJ/cm^2. Unbound oligonucleotides were washed off with 0.5% SDS for 15 min at room temperature, and SDS was removed by dipping the slides in anhydrous ethanol for 30 s. The slides were spin-dried at 1000 rpm for 2 min.

Probe labeling, hybridization, image analysis and normalization

The microarray experimental procedures were described in the previous published paper [18]. DNA in hybridization solution was denatured at 95 °C for 3 min prior to loading on to the microarray. The array was hybridized at 42 °C overnight. After hybridization, microarrays were rinsed consecutively with two types of washing solutions (with one solution containing 0.2% SDS and 2 × SSC at 42 °C for 5 min and with another solution containing 0.2% SSC at room temperature

for 5 min). Arrays were scanned with a confocal LuxScan scanner (CapitalBio Corporation) and resulting images were analyzed with Spot Data software (CapitalBio Corporation), using a space- and intensity-dependent normalization of fluorescence, based on a locally-weighted scatterplot smoothing (LOWESS) algorithm [19].

Statistics, annotation, relative gene expression and biological pathway analyses

Relative gene expression levels were determined using a Bayesian method (Bayesian analysis of gene expression levels, BAGEL) [20] from the normalized ratio data. This approach estimates the relative expression level for each gene based on the fluorescence ratio of Cy5/Cy3 fluorophores. For all comparisons, unless otherwise stated, difference was considered to be significant at $P < 0.001$. DEGs were identified with non-overlapping 95% confidence intervals from the Bayesian analysis. The fold change criteria were identified as mentioned in our previous report [21]. BAGEL has the advantage of detecting small yet statistically significant differences, which might be of biological significance. Hierarchical clustering of DEGs, as determined by BAGEL, was performed with Cluster [22,23] and Genesis [23], using the average linkage algorithm. Gene annotation in terms of the associated pathways was performed using DAVID [24], SOURCE [25] and GoMiner [26]. Signal transduction pathways were analyzed by BioRag [27]. Biological signature analysis was performed to correlate gene expression patterns with clinical characteristics, including age, gender, pathology, differentiation, lymph node metastasis and survival time, which were commonly used to predict clinical outcomes and prognosis.

Authors' contributions

SZ performed the bioinformatics and statistical analysis. RG prepared RNA for microarray studies. RX performed the model analysis and drafted the manuscript. LZ, YR, YZ and JC performed the microarray experiments. WL and MZ provided technical support for the experiments, BK did gene expression analysis. JF, FH, XL and YY collected the human subject samples and provided the required clinical data of the participants. SL and RW revised the manuscript. YL conceived the project, supervised the experiments and data analysis, and revised the manuscript. All authors read and approved the final manuscript.

Competing interests

The authors have declared that no competing interest exists.

Acknowledgements

This study was supported by grants from the National Program on Key Basic Research Project of China (Grant No. 2004CB518708) and the National High-Tech R&D Program of China (863 Program; Grant Nos. 2001AA233061, 2012AA02A203 and 2012AA02A504). We are thankful to the specimen bank of Beijing Cancer Hospital for providing specimens. We also thank Dr. Richard Ascione, Dr. Guanbin Luo from Case University, USA for helpful discussion, critical reading and insightful comments on the manuscript.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.gpb.2014.09.004.

References

[1] Yang L, Parkin DM, Whelan S, Zhang S, Chen Y, Lu F, et al. Statistics on cancer in China: cancer registration in 2002. Eur J Cancer Prev 2005;14:329–35.
[2] Ahn MJ, Yoo YD, Lee KH, Ahn JI, Yu DH, Lee HS, et al. cDNA microarray analysis of differential gene expression in gastric cancer cells sensitive and resistant to 5-fluorouracil and cisplatin. Cancer Res Treat 2005;37:54–62.
[3] Jin X, Zhu Z, Shi Y. Metastasis mechanism and gene/protein expression in gastric cancer with distant organs metastasis. Bull Cancer 2014;101:E1–12.
[4] Wang Y, Feng X, Jia R, Liu G, Zhang M, Fan D, et al. Microarray expression profile analysis of long non-coding RNAs of advanced stage human gastric cardia adenocarcinoma. Mol Genet Genomics 2014;289:291–302.
[5] Alevizos L, Gomatos IP, Smparounis S, Konstadoulakis MM, Zografos G. Review of the molecular profile and modern prognostic markers for gastric lymphoma: how do they affect clinical practice? Can J Surg 2012;55:117–24.
[6] Su L, Liu X, Chai N, Lv L, Wang R, Li X, et al. The transcription factor FOXO4 is down-regulated and inhibits tumor proliferation and metastasis in gastric cancer. BMC Cancer 2014;14:378.
[7] Ji R, Zuo XL, Yu T, Gu XM, Li Z, Zhou CJ, et al. Mucosal barrier defects in gastric intestinal metaplasia: *in vivo* evaluation by confocal endomicroscopy. Gastrointest Endosc 2012;75:980–7.
[8] Pandey A, Tripathi SC, Mahata S, Vishnoi K, Shukla S, Misra SP, et al. Carcinogenic *Helicobacter pylori* in gastric pre-cancer and cancer lesions: association with tobacco-chewing. World J Gastroenterol 2014;14:6860–8.
[9] Fan YF, Wu YM, Liu H, Yu Y, Jiang YY, Xue YZ, et al. TLR4 polymorphisms associated with developing gastric pre-cancer lesions in a Chinese Han population. Hum Immunol 2014;75:176–81.
[10] Cho JY. Molecular diagnosis for personalized target therapy in gastric cancer. J Gastric Cancer 2013;13:129–35.
[11] Konishi H, Ichikawa D, Komatsu S, Shiozaki A, Tsujiura M, Takeshita H. Detection of gastric cancer-associated microRNAs on microRNA microarray comparing pre- and post-operative plasma. Br J Cancer 2012;106:740–7.
[12] Matsumura N, Zembutsu H, Yamaguchi K, Sasaki K, Tsuruma T, Nishidate T, et al. Identification of novel molecular markers for detection of gastric cancer cells in the peripheral blood circulation using genome-wide microarray analysis. Exp Ther Med 2011;2:705–13.
[13] Myllykangas S, Junnila S, Kokkola A, Autio R, Scheinin I, Kiviluoto T, et al. Integrated gene copy number and expression microarray analysis of gastric cancer highlights potential target genes. Int J Cancer 2008;123:817–25.

[14] Wang YY, Ye ZY, Zhao ZS, Tao HQ, Li SG. Systems biology approach to identification of biomarkers for metastatic progression in gastric cancer. J Cancer Res Clin Oncol 2010;136:135–41.
[15] Boussioutas A, Li H, Liu J, Waring P, Lade S, Holloway AJ, et al. Distinctive patterns of gene expression in premalignant gastric mucosa and gastric cancer. Cancer Res 2003;63:2569–77.
[16] Guo RF, Zang SZ, Fang JY, Zhang L, Hu FL, Li WM, et al. Identification of biomarkers for early detection in gastric cancer and its clinical biological significance. Beijing Da Xue Xue Bao Zi Ran Ke Xue Bao 2009;41:353–60.
[17] Cunningham AB, Lennox JE, Ross RJ. Biofilms: the hypertextbook. <http://www.hypertextbookshop.com/biofilmbook/v004/r003/>; 2011 [accessed 1.04.2011].
[18] Patterson TA, Lobenhofer EK, Fulmer-Smentek SB, Collins PJ, Chu TM, Bao W, et al. Performance comparison of one-color and two-color platforms within the Microarray Quality Control (MAQC) project. Nat Biotech 2006;24:1140–50.
[19] Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J, et al. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. Nucleic Acids Res 2002;30:e15.
[20] Townsend JP, Hartl DL. Bayesian analysis of gene expression levels: statistical quantification of relative mRNA level across multiple strains or treatments. Genome Biol 2002;3, research0071.
[21] Zang S, Guo R, Zhang L, Lu Y. Integration of statistical inference methods and a novel control measure to improve sensitivity and specificity of data analysis in expression profiling studies. J Biomed Inform 2007;40:552–60.
[22] Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95:14863–8.
[23] Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. Bioinformatics 2002;18:207–8.
[24] Sherman BT, Huang da W, Tan Q, Guo Y, Bour S, Liu D, et al. DAVID Knowledgebase: a gene-centered database integrating heterogeneous gene annotation resources to facilitate high-throughput gene functional analysis. BMC Bioinformatics 2007;8:426.
[25] Diehn M, Sherlock G, Binkley G, Jin H, Matese JC, Hernandez-Boussard T, et al. SOURCE: a unified genomic resource of functional annotations, ontologies and gene expression data. Nucleic Acids Res 2003;31:219–23.
[26] Zeeberg BR, Qin H, Narasimhan S, Sunshine M, Cao H, Kane DW, et al. High-Throughput GoMiner, an 'industrial-strength' integrative gene ontology tool for interpretation of multiple-microarray experiments, with application to studies of Common Variable Immune Deficiency (CVID). BMC Bioinformatics 2005;6:168.
[27] Pandey R, Guru RK, Mount DW. Pathway Miner: extracting gene association networks from molecular pathways for predicting the biological significance of gene expression microarray data. Bioinformatics 2004;20:2156–8.