Original Article
Delineation of gastric cancer subtypes by co-regulated expression of receptor tyrosine kinases and chemosensitivity genes

Shu-Chun Li1,2,3*, Rong Ma1*, Jian-Zhong Wu4, Xia Xiao3, Wei Wu4, Gang Li1, Bo Chen2,3, Ashok Sharma2, Shan Bai2, Bo-Ying Dun2, Jin-Xiong She2, Jin-Hai Tang1

1Clinical Oncology Research Center, Jiangsu Cancer Hospital, Nanjing, Jiangsu Province, People’s Republic of China; 2Center for Biotechnology and Genomic Medicine, Medical College of Georgia, Georgia Regents University, Augusta, GA; 3Sino-American Institute of Translational Medicine, Nanjing Tech University, Nanjing, Jiangsu Province, People’s Republic of China; 4Zhenjiang Jintai Life Technologies, Zhenjiang, Jiangsu Province, People’s Republic of China. *Equal contributors.

Received June 14, 2015; Accepted June 22, 2015; Epub August 15, 2015; Published August 30, 2015

Abstract: Chemotherapy plays a key role in improving disease-free survival and overall survival of gastric cancer (GC); however, response rates are variable and a non-negligible proportion of patients undergo toxic and costly chemotherapeutic regimens without a survival benefit. Several studies have shown the existence of GC subtypes which may predict survival and respond differently to chemotherapy. It is also known that the expression level of chemotherapy-related and target therapy-related genes correlates with response to specific antitumor drugs. Nevertheless, these genes have not been considered jointly to define GC subtypes. In this study, we evaluated seven genes known to influence chemotherapeutic response (ERCC1, BRCA1, RRM1, TUBB3, STMN1, TYMS and TOP2A) and five receptor tyrosine kinases (RTKs) (EGFR, ERBB2, PDGFRB, VEGFR1 and VEGFR2). We demonstrate significant heterogeneity of gene expression among GC patients and identified four GC subtypes using the expression profiles of eight genes in two co-regulation groups: chemosensitivity (BRCA1, STMN1, TYMS and TOP2A) and RTKs (EGFR, PDGFRB, VEGFR1 and VEGFR2). The results are of immediate translational value regarding GC diagnostics and therapeutics, as many of these genes are currently widely used in relevant clinical testing.

Keywords: Gastric cancer, gene expression, co-regulation, chemotherapy

Introduction

Gastric cancer (GC) is one of the major causes of cancer-related deaths worldwide. Among the multidisciplinary treatment approaches, chemotherapy plays a key role both before and after surgery in operable GC, improving disease-free survival and overall survival. Many antitumor reagents have been shown to be effective against GC in clinical trials, but response rates are variable, with a non-negligible proportion of patients undergoing toxic and costly chemotherapeutic regimens without a survival benefit. Several studies have shown the existence of GC subtypes which may have different survival and response to chemotherapy [1-4]. For example, whole genome gene expression studies of 37 GC cell lines revealed two major genomic subtypes that are also validated in GC patients [3]. These subtypes predict survival and respond differently to chemotherapy. Another transcriptomic study identified three subtypes of gastric cancer designated as proliferative, metabolic, and mesenchymal [2]. Tumors of the metabolic subtype are more sensitive to 5-fluorouracil (5-FU) than the other subtypes, while cancer cells of the mesenchymal subtype are more sensitive to PI3K-AKT-mTOR inhibitors in vitro [2].

A variety of genes can be used to predict chemotherapeutic sensitivity or prognosis. The expression level of genes may correlate with response to specific antitumor drugs. For example, an association between TYMS expression and sensitivity to 5-FU has been demonstrated...
RTK and chemosensitive gene expression in gastric cancer

by many studies [5-8]. As only patients with low TYMS expression can respond to 5-FU, individualized chemotherapy can be selected according to tumor classification by the expression of TYMS [7, 8]. In this study, we evaluated twelve genes that are related to chemosensitivity and currently used in clinical practice. Seven of the twelve genes are known to influence the outcomes of chemotherapeutic drugs (ERCC1, BRCA1, RRM1, TUBB3, STMN1, TYMS and TOP2A), while the other five genes are receptor tyrosine kinases (RTKs) that have been targeted for cancer therapy (EGFR, ERBB2, PDGFRB, VEGFR1 and VEGFR2). Our results suggest significant heterogeneity in gene expression among GC patients and subtypes of patients can be delineated by the expression profiles.

Materials and methods

Patients and tissue samples

A total of 54 patients who underwent curative surgery for gastric cancer were enrolled into this study, including tumor tissues and their adjacent normal tissues. Among those patients, only 38 (28 males and 10 females) were integrally described with clinical characteristics. According to the WHO classification 2010, 4 patients had moderate differentiation and 34 patients had poor differentiation. Five cases were diagnosed as stage I-II and thirty-three cases as stage III-IV based on UICC/AJCC staging for gastric cancer. Thirty cases had lymph node metastasis and eight cases were negative for lymph node metastasis, while only four cases had distant metastasis. The present study was approved by the ethical committee of Jiangsu Cancer Hospital, Nanjing Medical University, China.

RNA isolation and cDNA synthesis

Total RNA samples were prepared with the Miracle isolation kit for tissues and cells (Jinfiniti Biosciences, LLC, Augusta, USA) according to the manufacture’s instruction. RNA samples were examined for concentration and purity using a Nanodrop ND-1000 spectrophotometer. cDNA synthesis was performed from total RNA using the TaqMan high capacity reverse transcription kit (Applied Biosystems). The 20 μl reverse transcriptase reaction system containing 1 μg of total RNA was incubated for 10 min at 25°C, 2 hours at 37°C and then 5 min at 85°C with the Biometer PCR System.

Quantitative real-time PCR (qRT-PCR)

Each cDNA sample was analyzed in triplicates using the Applied Biosystems 7900 with customized TaqMan low density array, containing twelve target genes and nine reference genes. qRT-PCR was carried out using TaqMan Universal PCR Master Mix (Applied Biosystems) containing ROX™ reference dye to normalize fluorescence values. For thermal cycling, the following conditions were applied: 10 min at 95°C, then 40 cycles of 15 s at 95°C and 1 min at 60°C.

Normalization of gene expression

Nine candidate reference genes were used and stability of the candidate reference genes was evaluated using four different methods (geNorm [9], NormFinder [10] Delta [11] and best keeper [12]). The three top performing genes (ESD, MRPL19 and IPO8) were selected according to the consensus from four different programs. The geometric average of these three genes was used for normalization.

GO biological processes analysis

To investigate the affected biological processes by the twelve genes, annotation and enrichment analysis was performed using DAVID (http://david.abcc.ncifcrf.gov/) functional annotation tool [13].

Statistical analysis

Gene expression Ct (cycle threshold) values were normalized using geometric average of the three selected reference genes. Paired t-test was used to compare the normalized gene expression between tumor and adjacent normal tissues. The effect of age on the changes in expression of each gene was determined using a linear regression of gene expression with age as covariate. The potential sex specific differences were examined using a t-test comparing expression changes in males and females. All p-values were two-tailed and a p<0.05 was considered statistically significant. The pairwise correlation between normalized gene expressions was computed in adjacent
RTK and chemosensitive gene expression in gastric cancer

Evaluation of reference genes for normalization

In all quantitative RT-PCR analysis, the choice of reference genes for normalizing mRNA concentration is the most critical factor that determines the accuracy of gene expression levels. In many studies, a housekeeping gene such as GAPDH or HPRT has been used as a reference gene to normalize the data. It is well known that the use of any single gene can be very problematic for quantitative RT-PCR analysis as all genes do have some expression variations in different individuals. An alternative approach is to use a small panel of genes for the normalization purpose. In this study, we evaluated the performance of nine candidate reference genes that we have selected from the literature. All nine genes were analyzed in the entire set of gastric cancer tissues and paired adjacent normal tissues. These nine genes have quite different expression levels as indicated in Figure 1A. GAPDH has the lowest Ct value (likely reflecting highest expression) while TBP has the lowest Ct value. Stability analyses were performed for all nine reference genes using four different methods (NormFinder, geNorm, Delta and Best keeper). Despite high expression levels and its common use in quantitative RT-PCR analyses, GAPDH has the lowest stability for gastric tissues from gastric cancer patients (Figure 1B), suggesting that GAPDH is not an appropriate internal reference control for normalizing gene expression in gastric cancer. Among the other eight candidate reference genes, the top four performing genes are ESD, MRPL19, IPO8 and PPIA according to the consensus from the four

Results

Evaluation of reference genes for normalization

In all quantitative RT-PCR analysis, the choice of reference genes for normalizing mRNA concentration is the most critical factor that determines the accuracy of gene expression levels. In many studies, a housekeeping gene such as GAPDH or HPRT has been used as a reference gene to normalize the data. It is well known that the use of any single gene can be very problematic for quantitative RT-PCR analysis as all genes do have some expression variations in different individuals. An alternative approach is to use a small panel of genes for the normalization purpose. In this study, we evaluated the performance of nine candidate reference genes that we have selected from the literature. All nine genes were analyzed in the entire set of gastric cancer tissues and paired adjacent normal tissues. These nine genes have quite different expression levels as indicated in Figure 1A. GAPDH has the lowest Ct value (likely reflecting highest expression) while TBP has the lowest Ct value. Stability analyses were performed for all nine reference genes using four different methods (NormFinder, geNorm, Delta and Best keeper). Despite high expression levels and its common use in quantitative RT-PCR analyses, GAPDH has the lowest stability for gastric tissues from gastric cancer patients (Figure 1B), suggesting that GAPDH is not an appropriate internal reference control for normalizing gene expression in gastric cancer. Among the other eight candidate reference genes, the top four performing genes are ESD, MRPL19, IPO8 and PPIA according to the consensus from the four
RTK and chemosensitive gene expression in gastric cancer

Gene expression changes in gastric cancer patients

In this study, we investigated the expression of twelve genes known to influence therapeutic outcomes of cancer. Using DAVID, a web-accessible program that allow investigators to understand biological meaning of genes [13], we first examined the gene ontology (GO) biological processes in which these genes are involved. The biological processes with an enrichment p-value < 0.05 are presented in Figure 2. Five of the genes (EGFR, ERBB2, PDGFRB, VEGFR1 and VEGFR2) belong to the receptor tyrosine kinase (RTK) signaling pathway, which plays a critical role in tumor cell proliferation and/or angiogenesis. The other seven genes...
are chemo-therapy associated genes. TUBB3 and STMN1 are involved in microtubule-based processes, such as cell cycle and intracellular signaling cascade, while the other five genes (BRCA1, TYMS, ERCC1, RRM1 and TOP2A) are involved in the DNA metabolic processes such as DNA repair (Figure 2).

The twelve genes were analyzed for gastric cancer samples and the expression data were normalized using geometric means of ESD, MRPL19 and IPO8. Comparison of normalized data between tumor tissue and adjacent normal, including paired t-test p-values, are shown in Figure 3. Among these twelve genes, PDGFRB, TOP2A and STMN1 showed the most significant increase in tumor tissues compared to adjacent normal tissue (p = 3.9×10^{-7}, 6.2×10^{-6}, and 2.1×10^{-5}, respectively). Other genes with significant increase in tumors include BRCA1 (p = 3×10^{-4}), TUBB3 (p = 0.0011), RRM1 (p = 0.017). The remaining five genes (ERCC1, EGFR, ERBB2, VEGFR1 and VEGFR2) did not reach significance level; however, a small percentage of patients have highly increased levels of EGFR, ERBB2, VEGFR1 and VEGFR2.
Figure 5. Heatmap of gene expression ratio (log₂ FC) between tumor and adjacent normal tissues. Hierarchical clustering of patients based on similar expression patterns revealed five subsets of GC. Patient clusters are indicated at the bottom of the figure.

Figure 6. Impact of age and sex on gene expression differences. Influence of age was examined using logistic regression, while student t test was used to examine the influence of sex on gene expression.
Co-regulation of gene expression in gastric cancer

Pair-wise Pearson correlation was computed for each pair of the genes using data in adjacent normal tissues, tumor tissues, normal and tumor tissues combined, and the tumor versus normal ratios in each patient (Figure 4). Gene expression correlation data in both adjacent normal and tumors suggest that three clusters of genes are co-regulated in the GC tissues. The first cluster consists of the DNA metabolic genes (STMN1, BRCA1, TYMS, TOP2A and RRM1). The second cluster consists of mainly three RTKs (VEGFR1, VEGFR2 and PDGFRB) and two other RTKs (ERBB2 and EGFR) are also more loosely correlated. The third correlation group includes two genes (ERCC1 and TUBB3).

Pair-wise correlations based on the expression differences between tumor and adjacent normal tissues revealed two very distinct clusters of genes, one being the RTK cluster (PDGFR, VEGFR1 and VEGFR2) and the second being DNA metabolic cluster (BRCA1, TYMS, TOP2A and STMN1). Furthermore, weak negative correlations were also observed between genes in the two different clusters.

Subtypes of GC patients defined by co-regulated genes

Eight genes in two co-regulated groups were selected to identify the subgroups based on the gene expression patterns (seven genes in the two clusters of differentially expressed genes and the EGFR gene). Gene expression differences (tumor/normal fold change) of these eight genes are graphically presented as a heat map (Figure 5), which clearly indicates the heterogeneity of GC. Each patient has a distinct pattern of gene expression. PDGFRB and TOP2A are over-expressed in a larger proportion of patients than all other genes; however, not all patients are positive for these two genes or any other gene analyzed here. Despite the tremendous heterogeneity, subsets of patients do have similar expression patterns for subsets of genes. Hierarchical clustering of all GC patients revealed several subsets of GC defined by the expression differences (Figure 5). Some GC patients have increased expression of both RTK and DNA metabolic genes (cluster 1); some patients have increased expression for RTKs only (cluster 2); a larger proportion of patients have increased expression for metabolic genes only (cluster 3), while the fourth and fifth cluster have unchanged or even lower expression for both RTK and metabolic genes.

Correlation with clinicopathologic variables

Gene expression differences were examined in subsets of patients with different demographic or clinical characteristics including age, gender, TMN stage, differentiation, lymph node metastasis and distant metastasis. Age has a significant influence on gene expression differences between tumor and adjacent normal tissues. Older patients tend to have larger overexpression in tumors for the DNA metabolic genes (TOP2A, p = 0.006; RRM1, p = 0.0016; STMN1, p = 0.0042), while younger patients tend to have larger overexpression in tumors for RTK genes (PDGFR and VEGFR2) (Figure 6A). Sex may also have a significant impact on gene expression differences in GC patients (Figure 6B). Female patients tend to have larger overexpression of RTK genes in tumors (p = 0.002 for PDGFR and p = 0.027 for VEGFR2), while male patients tend to have larger overexpression in DNA metabolic genes although the differences did not reach statistical significance (Figure 6). The other clinicopathologic variables did not have a significant impact on gene expression differences in this study.

Discussion

This study was designed to assess the expression of genes involved in determining responses to chemo- and target-therapies for cancer using real-time RT-PCR. In order to obtain accurate data, it was necessary to carefully evaluate the performance of reference genes that can be used for mRNA normalization [14]. Several studies have suggested suitable reference genes for gene expression studies in various human tissues. Among the suggested reference genes, GAPDH and ACTB (β-actin) are the most commonly used genes in clinical testing. One recent study compared different genes and suggested that the best reference genes differ by tumor tissues and the best performing genes are GAPDH and β-actin [15]. The stability of these two genes has been questioned in some tissue types [16, 17] and the suitability of selected reference genes must be unconditionally validated prior to each study [18]. In this study, the stability of nine candidate reference
RTK and chemosensitive gene expression in gastric cancer

genes (ESD, MRPL19 [19], TBP [20], RPLPO [20], PPIA [20], IPO8 [21], GUSB [22], GAPDH and HPRT1 [23]) was investigated with qRT-PCR in gastric cancer. Our data suggest that four excellent genes (ESD, MRPL19, IPO8 and PPIA) may be used in combination to obtain reliable normalization of RT-PCR data. Our unpublished data also suggest that these genes are also excellent reference genes for other cancers including lung and esophagus.

A major finding of this study is the coordinated regulation of gene expression in both normal and tumor tissues from GC patients. As expected, these co-regulated genes belong to specific functional groups. One strongly correlated group of genes is those involved in DNA metabolic processes (TOP2A, STMN1, BRCA1 and TYMS) and another group of strongly correlated genes is the RTKs (PDGFRB, VEGFR1 and VEGFR2). Surprisingly, these two groups of genes are anti-correlated or negatively correlated, suggesting that the RTKs and DNA metabolic genes are infrequently up-regulated in the same GC patients as shown in Figure 5.

A second major finding of this study confirms the tremendous heterogeneity of gene expression in GC patients. Several DNA metabolic genes well known to influence chemotherapeutic response are up-regulated in different proportions of GC patients. TOP2A is increased by more than 2-fold in 54% of the patients. In a previous study of GC using immunohistochemistry, it was found that TOP2A protein was detectable in 100% of the GC cases [24]. These results together suggest that TOP2A could be an excellent therapeutic target for GC. Three other DNA metabolic genes (STMN1, TYMS and BRCA1) are highly correlated with TOP2A and over-expressed by > 2-fold in 48%, 44% and 35% of the GC cases. STMN1 protein expression was found to be negatively correlated with recurrence-free survival in the diffuse type of GC and siRNA knockdown of STMN1 inhibits GC cell proliferation, migration and invasion [25]. STMN1 siRNA also regresses gastric tumors in xenograft models [26]. Thymidylate synthase encoded by TYMS catalyzes the methylation of deoxyuridylate to deoxymethylidylate, a function that maintains the dTMP (thymidine-5′-monophosphate) pool critical for DNA replication and repair. TYMS has been of interest as a target for cancer chemotherapeutic agents and it is considered to be the primary site of action for 5-fluorouracil, 5-fluoro-2-prime-deoxyuridine, and some folate analogs. BRCA1 is a tumor suppressor gene that encodes a protein involved in DNA repair. A large number of mutations have been found in patients with various types of cancer including breast and ovarian cancers [27]. Among individuals with non-small cell lung cancer (NSCLC), low expression of BRCA1 in the primary tumor correlated with improved survival after platinum-containing chemotherapy [28, 29]. This correlation implies that low expression of BRCA1 and the consequent low level of DNA repair may cause vulnerability of the tumor cells to treatment by the DNA cross-linking agents. High BRCA1 may protect cancer cells by acting in a pathway that removes the damages in DNA caused by platinum drugs. Therefore, BRCA1 expression level is a potentially important tool for tailoring chemotherapy in lung cancer management [28, 29]. Patients with sporadic ovarian cancer treated with platinum drugs have longer median survival times if their BRCA1 expression was low compared to patients with higher BRCA1 expression [30]. BRCA1 expression is high in about 30% of the GC patients and its expression is highly correlated with the other chemotherapeutic genes (TOP2A, STMN1 and TYMS). Analyses of these genes may be potentially important tool for tailoring chemotherapy in GC management.

EGFR and ERBB2 are two members of the epidermal growth factor receptor family, which is implicated in multiple cancers. Genetic alterations in the EGFR gene family have been shown to contribute to tumorigenesis and tumor progression in different types of cancer [31]. EGFR mutations can predict survival of lung cancer patients treated with EGFR inhibitors [32]. Trastuzumab is the only approved target agent for a subgroup of GC patients with HER2 overexpression at present, which represent about 20% of all the patients [33], based on the results of phase III ToGA trial [34]. A subset of gastric cancer patients with EGFR amplification and over expression of EGFR responds to therapy with cetuximab, an anti-EGFR monoclonal antibody [35]. Consistent with previous studies, we found that EGFR is overexpressed by 2-fold or more in only 26% of the GC patients and ERBB2 is overexpressed by 2-fold or more in 30% of the GC patients. An even lower percentage of GC patients in this study have higher
overexpression for *EGFR* or *ERBB2*. Similarly, only 25% of the GC patients have >2-fold overexpression for *VEGFR1* or *VEGFR2*, two receptors important for angiogenesis.

In contrast to *EGFR*, *ERBB2*, *VEGFR1* and *VEGFR2*, overexpression of *PDGFRB* is found in 61% of the Chinese GC patients in this study. It has been reported that 41% of Japanese GC patents are positive for PDGFR [1]. PDGFR is also mutated, amplified or overexpressed in a large number of tumors [36]. Platelet-derived growth factor (PDGF) is one of the numerous growth factors that regulate cell growth and division. In particular, it plays a significant role in angiogenesis, a characteristic of cancer. Upon activation by PDGF, PDGFR receptors dimerize, and are switched on by autophosphorylation of several sites on their cytosolic domains, which serve to mediate binding of cofactors and subsequently activate signal transduction through the PI3K pathway or through reactive oxygen species (ROS)-mediated activation of the STAT3 pathway. PDGFR has been targeted for cancer therapy with some success [36-38] and is probably a better target than the other RTKs for GC due to its overexpression in a higher percentage of GC patients.

Finally and perhaps most importantly, the expression profile of these RTK and DNA metabolic genes allows the delineation of subsets of GC patients that have elevated expression for: 1) both groups of genes, 2) only RTKs, 3) only DNA metabolic genes, and 4) neither group. Analyses of these genes should be of clinical value in selecting the most appropriate therapies for each group of patients.

Acknowledgements

This research is supported by The National Science Foundation of China (grant number 81272244), the Plan for Technology Infrastructure Construction of Jiajsu Province (BM2011118).

Disclosure of conflict of interest

None.

Address correspondence to: Jinhai Tang, Jiangsu Cancer Hospital, Nanjing, Jiangsu Province, People’s Republic of China. E-mail: jschtjh@126.com

References


RTK and chemosensitive gene expression in gastric cancer


