Original Article

Tumor profiling of co-regulated receptor tyrosine kinase and chemoresistant genes reveal different targeting options for lung and gastroesophageal cancers

Jianzhong Wu*, Shuchun Li*, Rong Ma, Ashok Sharma, Shan Bai, Boying Dun, Haixia Cao, Changwen Jing, Jinxiong She, Jifeng Feng

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, Augusta University, Augusta, GA, USA. *Equal contributors.

Received August 16, 2016; Accepted September 10, 2016; Epub December 15, 2016; Published December 30, 2016

Abstract: The expression of a number of genes can influence the response rates to chemotherapy while genes encoding receptor tyrosine kinases (RTKs) determine the response to most targeted cancer therapies currently used in clinics. In this study, we evaluated seven genes known to influence chemotherapeutic response (ERCC1, BRCA1, RRM1, TUBB3, STMN1, TYMS, and TOP2A) and five RTKs (EGFR, ERBB2, PDGFRB, VEGFR1 and VEGFR2) in non-small cell lung cancer (NSCLC) and esophagus cancer (EC) and the data are compared to gastric cancer (GC) data reported previously. We demonstrate significant differences in the expression profiles between different cancer types as well as heterogeneity among patients within the same cancer type. In all three cancer types, five chemoresistant genes (TOP2A, STMN1, TYMS, BRCA1 and RRM1) are coordinately up-regulated in almost all EC, approximately 90% of NSCLC and one third of GC patients. Most EC and nearly half of GC patients have increased expression of the three RTKs critical to angiogenesis (PDGFR, VEGFR1 and VEGFR2), while almost none of the NSCLC patients have elevated expression of angiogenic RTKs. A variable percentage of patients in the three cancer types show upregulation of the EGFR family RTKs, EGFR and/or ERBB2. It is of interest to note that approximately 10% of the NSCLC and GC patients are triple-negative for the chemosensitivity genes, angiogenic and EGFR RTK genes. The results suggest significant gene expression differences between different cancer types as well as heterogeneity within each cancer type and therefore different molecules should be targeted for future drug development and clinical trials.

Keywords: Co-regulation of gene expression, chemoresistant genes, drug targets, normalization, reference genes, cancer heterogeneity

Introduction

Lung cancer (LC) is one of the leading causes of all cancer-related deaths worldwide, with a 5-year survival rate of approximately 15% [1], while it is the most common cancer, followed by gastric cancer, in China [2]. Despite some advances in early detection and recent improvements in its treatment, the prognosis of patients with lung cancer remains poor [3, 4] because it exhibits high resistance to anticancer chemotherapy. Biomarkers that allow early diagnosis, guidance of therapeutic selection and/or early assessment of therapeutic outcome should improve care for lung cancer patients. Several widely known cancer anti-gens including carcinoembryonic antigen (CEA), cytokeratin 19 fragment (CYFRA21-1), neuron-specific enolase (NSE) have been extensively studied and are found elevated in some lung cancer patients [5-8], however, none of these biomarkers is satisfactory for diagnosis at an early stage because of their low sensitivity and specificity even for late stage lung cancer.

Esophagus cancer (EC) is also one of the most common malignant cancers in the world and especially in China, where it is the fourth most common cancer-related deaths [2]. EC is much less extensively studied and the therapeutic outcome remains unchanged for the last several decades, with a five-year survival rate
Tumor profiling of chemotherapeutic response genes

between 15-25% [9]. A few studies have investigated the genomic abnormalities including copy number variations, mutations in specific oncogenes and tumor suppressor genes. More recently, whole exomes and targeted sequencing have identified a large number of mutations in EC [10]. Differential gene expression in EC has also been investigated in a number of studies using real-time RT-PCR and/or microarray technology [11-13].

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 by many studies [14-17]. 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 [16, 17]. One study also found that the expression levels of ERCC1 and BRCA1 and TYMS have an impact on the survival of EC patients after chemotherapy [18].

In this study, we selected twelve genes that are currently used in clinical practice and determined their expression level in non-small cell lung cancer (NSCLC) and EC patients. 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 RTKs that have been targeted for cancer therapy (EGFR, ERBB2, PDGFRB, VEGFR1 and VEGFR2). Together with our published data on gastric cancer [19], our results suggest significant heterogeneity in gene expression among different types of cancer and among different patients of the same cancer type and that the heterogeneity can be delineated by coordinate-ly regulated expression patterns of chemotherapeutic genes and RTKs.

Materials and methods

Patients and tissue samples

A total of 52 patients who underwent curative surgery for NSCLC and 22 patients with EC were enrolled into this study. Tumor tissues and their adjacent normal tissues were obtained for gene expression studies. 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 Biotech, LLC, Augusta, USA) according to the manufacturer’s instructions. 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 [20], NormFinder [21], Delta [22] and best keeper [23]). 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.

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 chang-
es 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 normal tissues, tumor tissues, normal and tumor tissues combined, and the tumor versus normal ratios, using Pearson correlation coefficient. Clustering and visualization of correlation matrix was performed using hierarchical clustering method and heatmap. A separate hierarchical clustering was also performed to identify the cancer subtypes by grouping the individuals exhibiting similar expression patterns. All statistical analyses were performed using the R language and environment for statistical computing (R version 3.1.1; R Foundation for Statistical Computing; www.r-project.org).

**Results**

**Evaluation of reference genes for normalization**

Since the choice of reference genes for normalizing mRNA concentration is the most critical factor that determines the accuracy of gene expression levels, we first conducted experiments to determine the reference genes that should be used for lung and esophagus tissues. 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. We evaluated the performance of nine candidate reference genes that were selected from the literature. All nine genes were analyzed in the entire set of 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 highest 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 tissues from NSCLC and esophagus cancer patients (Figure 1B) as well as gastric cancer [19], suggesting

![Figure 1. Stability analysis of reference genes in NSCLC and EC. A. Distribution of CT (cycle threshold) values for each of the nine candidate reference genes; B. Relative stability values and ranking of the nine candidate reference genes based on four different methods.](image-url)
Tumor profiling of chemotherapeutic response genes

Figure 2. Gene expression differences between tumor and adjacent normal in NSCLC (A) and EC (B). Expression differences are expressed as fold change (FC) between tumor and adjacent normal. Each dot represents fold-change value of one patient and box plots represent the distribution of fold-change values. The p-value and percentages of patients with FC ≥ 2.0 are shown on the top to the chart.

that GAPDH is not an appropriate internal reference control for normalizing gene expression in 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 different programs.
Any combination of 3-4 of these genes should be excellent choices for gene expression studies. These results are very similar to what have been found in a previous study with gastric cancer [19].

**Figure 3.** Heatmaps showing pairwise correlations of expression levels between genes in NSCLC (A), EC (B) and GC (C). The pairwise correlations were computed using the expression values in adjacent normal tissues, tumor tissues and tumor/normal ratios (log2 fold change values) separately.

**Gene expression changes in lung cancer patients**

The expression of twelve genes known to influence therapeutic outcomes of cancer was ana-
lyzed in this study. Five of the twelve genes (EGFR, ERBB2, PDGFRB, VEGFR1 and VEGFR2) belong to the RTK signaling pathway, which plays a critical role in tumor cell proliferation and/or angiogenesis. The remaining seven genes are chemosensitivity 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.

The twelve genes were analyzed for NSCLC 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 2A. The most interesting finding from this dataset is the highly increased expression of chemosensitivity genes and lack of up-regulation of RTKs in NSCLC patients. Indeed, none of the NSCLC patients have very high upregulation of PDGFR, VEGFR1 or VEGFR2, while very few patients have highly increased expression of EGFR and/or ERBB2. Among the chemosensitivity genes, TOP2A and TYMS are increased by more than 2-fold in 87% and 81% of NSCLC patients (P = 2.4×10^{-15} and 5.1×10^{-15} respectively). BRCA1 and STMN1 are also highly increased in NSCLC patients (P = 1.9×10^{-10} and 2.2×10^{-6}).

**Gene expression changes in esophagus cancer patients**

The same twelve genes were also analyzed for EC 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 2B. Similar to NSCLC, most chemosensitivity genes are highly increased in EC patients. TOP2A, TYMS, BRCA1, TUBB3 and STMN1 are increased by more than 2-fold in more than 70% of EC patients. In contrary to NSCLC, a higher proportion of EC patients also have increased expression of several RTK genes, including PDGFR, EGFR and ERBB2, although the proportion of patients with RTK up-regulation is much lower than the proportion of patients with up-regulation of chemosensitivity genes.

**Co-regulation of gene expression in NSCLC and EC**

Pair-wise Pearson correlation was computed for each pair of the genes using data in adjacent normal tissues, tumor tissues and the tumor versus normal ratios in each patient for NSCLC (Figure 3A) and EC (Figure 3B). In normal lung tissue, all twelve genes showed some degree of coordinated regulation. However, the co-regulation appears to be broken between RTK and chemosensitivity genes in the lung cancer tissues, while co-regulation becomes stronger among some RTK genes and among some chemosensitivity genes. In the normal esophagus tissues, two distinct groups of genes are strongly correlated within the groups while the two groups of genes show strong negative correlation. The first group of genes that show strong positive correlation include three RTK genes that are critical to angiogenesis (PDGFR, VEGFR1 and VEGFR2) and the second group of genes includes six chemosensitivity genes (STMN1, TOP2A, TYMS, BRCA1 and RRM1) and one RTK gene (ERBB2). Interestingly, one RTK (EGFR) and one chemosensitivity gene (TUBB3) are not correlated with any other genes. In the EC tissue, the overall positive and negative correlations have reduced. The correlation between the three RTK genes is reduced and remains relatively positive; however, ERCC1 and ERBB2 are no longer correlated with the other five chemosensitivity genes.

The correlation between the tumor/normal ratios is an indication of coordination of changes in tumors. Examination of the data for the three cancer types (Figure 3) suggests that the change in cancer tissues for five chemosensitivity genes (STMN1, TOP2A, TYMS, BRCA1 and RRM1) is highly coordinated in all three cancer types, while the change of three angiogenic RTK genes (PDGFR, VEGFR1 and VEGFR2) is also coordinated in these three tumors. 

**Heterogeneity between and within tumor types**

Gene expression differences (tumor/normal fold change) of the twelve genes are graphically presented as a heatmap for NSCLC and EC from this study as well as GC data from a previous study [19] (Figure 4). Patients in each cancer type were clustered based on gene expression ratios. Almost all EC patients have elevated expression for six of the seven chemosensi-
Tumor profiling of chemotherapeutic response genes

Figure 4. Heatmap of gene expression changes (log2 fold change values) in EC (A), NSCLC (B) and GC (C). Patients in each cancer type were clustered based on expression changes of 12 genes.

...tivity genes (except ERCC1), while approximately one third of the EC patients have down-regulated expression for the three angiogenic RTKs and two thirds of the EC patients have up-regulated angiogenic RTKs (Figure 4A). ERBB2 is up-regulated in some EC patients and down-regulated in a different subset of patients. In contrast to EC patients, almost all NSCLC patients have severely down-regulated expression of the three angiogenic RTKs and approximately 10% of NSCLC patients also have down-regulated EGFR family RTKs (EGFR and ERBB2). Therefore, these are triple-negative NSCLC patients. Almost all remaining 90% of NSCLC patients have coordinated up-regulation of five chemosensitivity genes, while TUBB3 is down-regulated in approximately 50% of the NSCLC patients and ERCC1 is unchanged in most patients (Figure 4B). Some of these patients also have increased expression of the EGFR RTKs. Similarly, about 10% of the GC patients are also triple-negative for angiogenic RTKs, EGFR RTKs and chemosensitivity genes, while the other GC patients are very heterogeneous with different combinations of positivity for the chemosensitivity genes and/or different RTKs. The three angiogenic RTKs and six of the seven chemosensitivity (except ERCC1) are highly co-regulated, while two of the EGFR family members (EGFR and ERBB2) are also co-regulated in GC (Figure 4C).

Discussion

This study was designed to compare the expression of genes involved in determining responses to chemotherapies and target therapies between lung cancer and gastroesophageal cancers. At first, we carefully evaluate the performance of reference genes that can be used for mRNA normalization. A previous 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 [24]. The stability of these two genes has been questioned in some tissue types [25, 26] and the suitability of selected reference genes must be unconditionally validated prior to each study [27]. In one of our previous studies [19], the stability of nine candidate reference genes (ESD, MRPL19 [28], TBP [29], RPLP0 [29], PPIA [29], IPO8 [30], GUSB [31], GAPDH and HPRT1 [32]) was investigated with qRT-PCR in gastric cancer. Our data suggested four excellent genes (ESD, MRP19, IPO8 and PIAA) that could be used in combination to obtain reliable normalization of RT-PCR data. In this study, our data suggest that these genes are also excellent reference genes for lung and esophagus cancers.

A major finding of this study is the up-regulation of chemosensitivity genes in both NSCLC and EC patients. Five genes (TOP2A, TYMS, STMN1, BRCA1 and RRM1) are coordinately up-regulated in all EC patients and approximately 90% of NSCLC patients. The same five genes are also coordinately up-regulated in approximately 70% of GC patients based on data in our previous study [19], suggesting that they are commonly up-regulated genes in cancer. Another chemosensitivity gene (TUBB3) is also up-regulated in all EC patients and a large proportion of GC patients but a smaller portion of the NSCLC patients.

TOP2A is increased by more than 2-fold in 87% of NSCLC, 86% of EC and 54% of the GC patients. A previous immunohistochemistry (IHC) study found that 37% of NSCLC patients stained positive for TOP2A [33]. In a large IHC study, it was shown that 86% and 67% of EC and GC patients stained positive for TOP2A [34], while another smaller IHC study found positive TOP2A in 100% of the GC cases [35]. These results together suggest that TOP2A over-expression may be related to drug insensitivity and could be an excellent therapeutic target for most gastroesophageal cancers.

TYMS encodes the thymidylate synthase which catalyzes the methylation of deoxyuridylate to deoxythymidylate, a function that maintains the dTMP (thymidine-5-prime monophosphate) pool critical for DNA replication and repair. TYMS is over-expressed by > 2-fold in 81% of NSCLC and 77% of EC as shown in this study but only 44% of GC [19]. 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 ana-
logs. Our studies suggest that TYMS may be an excellent target for NSCLC, EC and some GC patients.

Stathmin 1 (STMN1) is an important cytosolic protein associated with microtubule stability reported to be involved in tumorigenesis. STMN1 gene expression is increased in 46% of NSCLC, 77% of EC and 48% of GC. 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 [36]. STMN1 siRNA also regresses gastric tumors in xenograft models [37], gallbladder cancer [38] and esophagus cancer [39]. STMN1 also increases radioresistance in NSCLC patients [40]. These combined evidences suggest that STMN1 is potentially an excellent target for various cancers.

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 cancers including breast and ovarian cancers [41]. Among individuals with NSCLC, low expression of BRCA1 in the primary tumor correlated with improved survival after platinum-containing chemotherapy [42, 43]. 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. 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 [44]. 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 [42, 43]. In our studies, BRCA1 expression is high in about 58% of NSCLC, 73% of EC and only 35% of GC patients. Analyses of these genes may be potentially important tool for tailoring chemotherapy in EC, NSCLC and especially GC management.

RRM1 is the fifth chemotherapy resistant genes that have co-regulated expression in the three cancers in our studies. RRM1 encodes the large subunit of ribonucleoside-diphosphate reductase that is essential for the production of deoxyribonucleotides prior to DNA synthesis in S phase of dividing cells. Network analysis of our microarray data revealed that many of the cell cycle genes altered by mycophenolic acid (MPA) treatment are connected to the RRM1 gene and the RRM1 protein is drastically reduced by MPA treatment although the gene expression change was not altered [45]. We found that RRM1 gene expression is increased in 29% of NSLC, 41% of EC and 30% of GC patients. However, RRM1 protein is detected in 87% of NSCLC patients in one study [46]. Cancer cells that survive chemotherapy have increased expression of RRM1 and silencing of RRM1 can increase the sensitivity of certain anticancer drugs such as gemcitabine [47].

A second major finding in this study is that the five RTK genes studied here can be grouped into two subsets based on coordinated expression that correlate with their function. The first group are RTKs involved in angiogenesis and include three genes (PDGFR, VEGFR1 and VEGFR2), while the second group corresponds to the EGFR family (EGFR and ERBB2) critical to cell proliferation. Surprisingly, almost no NSCLC patients have up-regulated angiogenic RTKs, suggesting that NSCLC tumors are likely under hypoxic conditions and targeting these three angiogenic RTKs is unlikely successful for NSCLC. In contrast to NSCLC, angiogenic RTKs are coordinately overexpressed in about two thirds of EC and GC. Selective targeting of these factors for this subset of EC and GC patients may be more effective. Our data also indicate that various percentages of EC, GC and NSCLC patients have elevated expression of EGFR and or ERBB2 (Figure 4). A number of studies also reported varying numbers of positivity for these RTKs. For example, it was found that ERBB2 gene was only overexpressed in less than 10% of the patients while EGFR is overexpressed in 62% of the patients [48]. At the protein level, 44-90% of the NSCLC patients were found to be positive for EGFR [49, 50]. Over 90% of the EC patients were found to be positive for various RTKs through IHC analysis [51]. It appears that positivity for IHC is generally higher than gene expression. The discrepancies between gene expression and IHC data need to be further investigated. Possibilities include differences in study patient population, non-specific binding of antibodies in IHC and true discrep-
cies between gene and protein expression. Answer to this question has important clinical implications as it may guide future therapeutic development and clinical trials.

Finally, the third and perhaps the most important finding of this study is the genomic heterogeneity between different cancer types as well as within each cancer type. This study suggests for the first time that a small subset (approximately 10%) of NSCLC and GC patients are triple-negative for angiogenic RTKs, EGFR RTKs and chemosensitivity genes, while the vast majority of patients show positivity for various combinations of the three classes of genes. These triple-negative patients most likely require other targets for therapeutic intervention. The genomic heterogeneity also suggests different targeting strategies for different subsets of patients. Five coordinate up-regulated chemosensitivity genes (TOP2A, TYMS, STMN1, BRCA1 and RRM1) may be excellent targets for almost all EC patients and most NSCLC patients and a good subset of GC patients. The three coordinate regulated angiogenic RTKs (PDGFR, VEGFR1 and VEGFR2) may serve as good targets for a subset of EC and GC patients but are probably poor targets for all NSCLC patients.

Acknowledgements

This research is supported by the National Science Foundation of China (grant number 81272244 and 81372396) and general program from Jiangsu province health department (grant number H2014110).

Disclosure of conflict of interest

None.

Address correspondence to: Jifeng Feng, Clinical Oncology Research Center, Jiangsu Cancer Hospital, Nanjing, Jiangsu Province, People’s Republic of China. E-mail: fjif@vip.sina.com

References


[33] Huang H, Liu J, Meng Q and Niu G. Multidrug resistance protein and topoisoamerase 2 alpha expression in non-small cell lung cancer are


