Differential gene expression profiling of esophageal adenocarcinoma

Zane T. Hammoud, MD,[a] Sunil Badve, MD,[b,c,d] Qianqian Zhao, MS,[e] Lang Li, PhD,[e] Romil Saxena, MD,[b] Mangesh A. Thorat, MD,[b] Akira Morimiya, BS,[b] Karen M. Rieger, MD,[a] and Kenneth A. Kesler, MD[a]

GTS

Background: Differential gene expression offers an attractive means by which to study genes that may be involved in disease development and/or progression. We performed quantitative gene expression in various stages of esophageal adenocarcinoma, treated exclusively by surgery with complete 2-field lymphadenectomy, in an attempt to discern genes involved in disease progression as well as genes that may predict survival.

Methods: Gene expression profiling was accomplished by cDNA-mediated annealing, selection, extension, and ligation (DASL) assay. RNA was extracted from 89 archived formalin-fixed, paraffin-embedded esophageal adenocarcinoma tissues. DASL assay was performed with the Sentrix Universal Array (Illumina Corp, San Diego, Calif) of 502 known cancer-related genes. Bioinformatics tools were used to determine significant differential gene expression in T1-2 versus T3-4 tumors and tumors without lymph node involvement (N0) versus tumors with lymph node involvement (N+). Gene expression was also correlated with overall survival.

Results: Twenty-one genes were overexpressed in T1-2 compared with T3-4 tumors (false discovery rate of 0). Underexpression of 1 gene was seen in N+ compared with N0 tumors (false discovery rate of 0). For overall survival, underexpression of 9 genes correlated with long survival.

Conclusions: Using differential gene expression of 502 known cancer genes, we identified genes that may be involved at various stages in the progression of esophageal adenocarcinoma. We also identified genes that may correlate with prolonged survival and, thus, may serve as prognostic markers. These findings may provide further insight into the mechanisms of development and/or progression of esophageal adenocarcinoma. Prospective studies are needed to verify the prognostic value of these genes.

The incidence of esophageal adenocarcinoma has increased markedly over the past 30 years. In the Western world, adenocarcinoma is now more common than esophageal squamous cell carcinoma.[1] Surgical resection, with or without the addition of chemotherapy and/or radiation therapy, remains the cornerstone of therapy for esophageal adenocarcinoma and represents the best curative treatment option. However, despite improvements in preoperative staging and in operative morbidity and mortality, the prognosis of patients with adenocarcinoma remains relatively poor. A variety of factors have been explored to determine the biologic behavior of esophageal adenocarcinoma. Widely known prognostic factors, such as stage of tumor and lymph node involvement, are included in the TNM staging system. In addition, other factors such as the number of positive lymph nodes, grade of differentiation, and response to neoadjuvant chemoradiation have been shown to have additional prognostic value.[2,3] The description of molecular and/or genetic changes involved in carcinogenesis has led to opportunities to explore the impact of such changes on clinical behavior. A better understanding of these changes may identify additional factors that have the potential to improve prognostication and to influence additional therapy.

Carcinogenesis is a complex process that involves multiple genetic alterations. Many methods have been developed to study these changes and to delineate their potential impact on clinical behavior. The sequentiation of the human genome, combined with high throughput technologies, has led to the ability to describe these genetic alterations in a quantitative manner, thus allowing the development of tumor ''profiles'' that can distinguish subsets of disease, predict response to therapy, and possibly outcome. In esophageal adenocarcinoma, gene expression profiling using DNA microarrays has been used to compare adenocarcinoma with squamous cell carcinoma and with Barrett esophagus, establishing the presence of unique gene expression profiles capable of discriminating between these diseases.[4]

We performed gene expression profiling on formalin-fixed, paraffin-embedded (FFPE) esophageal adenocarcinomas using the cDNA-mediated annealing, selection, extension and ligation (DASL) assay. Quantitative gene expression was performed in various stages of adenocarcinoma, treated exclusively by surgery with complete 2-field lymphadenectomy, in an attempt to discern genes involved in disease progression as well as genes that may predict survival.

PATIENTS AND METHODS

Patients and Samples

The study was approved by the Indiana University Institutional Review Board. From 1990 to 2005, 138 patients with the diagnosis of esophageal

From the Departments of Surgery, Thoracic Surgery Division,[a] Pathology and Laboratory Medicine,[b] Internal Medicine,[c] Translational Genomics Core of Indiana University Cancer Center,[d] and Bioinformatics,[e] Indiana University School of Medicine, Indianapolis, Ind.

Received for publication June 2, 2008; revisions received July 24, 2008; accepted for publication Aug 31, 2008.

Address for reprints: Zane T. Hammoud, MD, Department of Surgery, Thoracic Surgery Division, Indiana University School of Medicine, 545 Barnhill Dr, EH 215, Indianapolis, IN 46202 (E-mail: zhammoud@iupui.edu).

J Thorac Cardiovasc Surg 2009;137:829-34
0022-5223/$36.00

doi:10.1016/j.jtcvs.2008.08.052

GTS

Abbreviations and Acronyms

DASL = cDNA-mediated annealing, selection, extension, and ligation
DNA = deoxyribonucleic acid
FFPE = formalin-fixed, paraffin-embedded
RNA = ribonucleic acid

adenocarcinoma, treated exclusively by surgery with complete 2-field lymphadenectomy as initial therapy at the Indiana University Medical Center, were identified from a prospectively maintained database. Representative hematoxylin and eosin–stained slides from all patients were reviewed by two pathologists (S.B. and R.S.). A section with a minimum of 1 cm^2 tumor-bearing area with at least 70% tumor was the selection criterion for inclusion in the study. Forty-one specimens did not fulfill this criterion and were excluded. Specimens with the minimum 1 cm^2 tumor-bearing area but with adjacent or surrounding stromal/granulation tissue were included; these specimens were subjected to manual macrodissection with hematoxylin and eosin slides used to facilitate macrodissection of the tumor-bearing area. The corresponding paraffin blocks were obtained from the Indiana University Department of Pathology. Three 10-μm sections were obtained from each selected paraffin block (one block per case) on noncharged glass slides taking due precautions to avoid nucleic acid contamination. Sections were deparaffinized with CitriSolv cleaning agent (Fisher Scientific Company, Fair Lawn, NJ) and scraped off from the slide into a microcentrifuge tube. For sections requiring macrodissection, only the tumor-bearing areas were scraped off using visual matching with marked hematoxylin and eosin slides. RNA was extracted from a total of 97 cases using High Pure RNA Paraffin Kit (Roche Applied Bioscience, Indianapolis, Ind). Eighty-nine cases fulfilled the RNA requirement of 200 ng/5 μL for the assay and served as the basis for the study. RNA was pre-qualified using iScript (Bio-Rad Laboratories Inc, Hercules, Calif) to reverse transcribe and SYBR Green Master Mix (Applied Biosystems, Foster City, Calif) to perform quantitative polymerase chain reaction for *RPL13a* gene. DASL assay was performed with the Sentrix Universal Array (Illumina Corp., San Diego, Calif) of 502 known cancer genes as per the manufacturer's instructions.[5] Technical duplicates for four samples were also included in the assay.

Clinical follow-up was obtained from office visits and through telephone contact. The pathologic TNM stage, date of surgery, date of death where applicable, and the date of last follow-up were recorded. Specimens were grouped into T1 and T2 tumors (T1-2) or T3 and T4 tumors (T3-4) to delineate genes that may be involved in advancing T stage as well as into tumors without lymph node involvement (N0) or with lymph node involvement (N+) to delineate genes that may be involved in lymph node metastasis. On the basis of survival from the date of surgery, Kaplan–Meier survival curves were calculated. A comparison was made on the basis of patient overall survival to delineate genes that may confer a survival advantage.

Statistical Analysis

Gene expression data were normalized at the median level. Hierarchical clustering and singular value decomposition methods were applied to detect the outliers for quality control purposes. The gene expression data was then correlated with T stage (1-2 vs 3-4) and with lymph node status (negative vs positive) through significant analysis of microarray.[6] T stage, lymph node status, and gene expression were correlated with patient overall survival through log–rank tests. Patients who died perioperatively as well as patients who died of noncancer causes were censored in the overall survival analysis. The false discovery rate,[7] an estimate of the proportion of errors committed by falsely rejecting null hypotheses and widely used in genome wide correlative studies, was calculated for each gene. Top-ranked genes were selected by a prespecified false discovery rate (q value) of 0.20. Gene set analysis was performed among the top ranked genes.[8]

TABLE 1. Pathologic TNM stage of 89 specimens used for analysis

TNM stage	No.
T1 N0	10
T1 N1	5
T2 N0	4
T2 N1	7
T3 N0	9
T3 N1	51
T4 N1	2
Tx N1	1

RESULTS

Table 1 lists the TNM pathologic stage of all 89 specimens analyzed. There were 26 T1-2 and 62 T3-4 specimens whereas there were 23 N0 and 66 N+ specimens. The median follow-up was 25 months (range 2–132 months) and 4 patients were lost to follow-up. The median survival for the entire group was 2.18 years (Figure 1). Analysis of T1-2 specimens compared with T3-4 specimens demonstrated up-regulation (overexpression) of a total of 63 genes, 21 of which had a false discovery rate (q value) of 0. Tables 2A and 2B demonstrate the results of this analysis and the genes identified. Table 3 demonstrates results of the analysis of N0 specimens compared with N+ specimens. A total of 16 genes were overexpressed and 1 gene *(MYB)* was underexpressed (down-regulated); only the underexpressed gene had a false discovery rate of 0. For the survival analysis, no overexpressed genes correlated with prolonged survival. However, underexpression of 9 genes correlated with prolonged survival (Table 4). Figure 2 demonstrates a typical survival curve for one of these genes (*CSPG2* or chondroitin sulfate proteoglycan core protein 2); the other 8 genes demonstrated similar survival curves. Figures 3 and 4

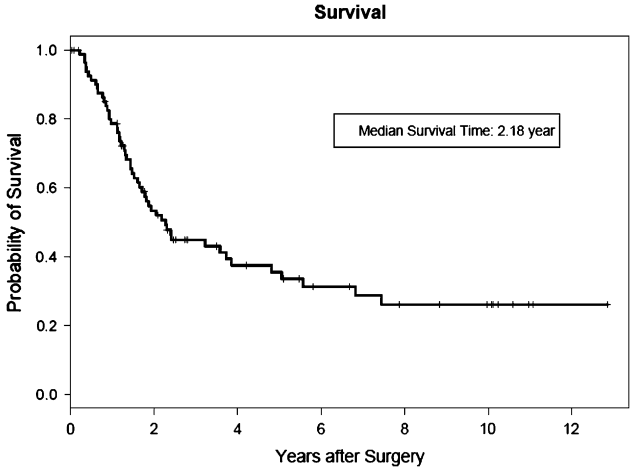

FIGURE 1. Kaplan–Meier survival curve of all 89 patients.