Prognostication in esophageal cancer
Lagarde, S.M.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Gene expression profiling can not be used to predict hematogenous dissemination in patients with adenocarcinoma of the esophagus

Sjoerd M Lagarde
Perry D Moerland
P Emiel Ver Loren van Themaat
Lisa A Gilhuijs-Pederson
Fiebo JW ten Kate
Pieter H Reitsma
Antoine H van Kampen
Aelko H Zwinderman
Frank Baas
J Jan B van Lanschot

to be submitted
Abstract

Introduction
Many patients with esophageal adenocarcinoma have unrecognized metastatic disease at the time of first presentation. These patients will have recurrent disease after esophagectomy. The aim of this study was to discover a gene expression signature that predicts the development of hematogenous dissemination and to identify genes and pathways that provide insight in the mechanisms underlying hematogenous dissemination in patients with adenocarcinoma of the esophagus.

Patients and Methods
Patients who developed hematogenous dissemination (N=33) were compared to patients without hematogenous dissemination (N=34). Whole genome oligonucleotide microarrays were used to evaluate the genetic signature of 77 esophageal cancers. Multiple random validation was used to analyze the stability of the molecular signature and predictive power. Gene set enrichment analysis (GSEA) was applied to elucidate oncogenetic pathways.

Results
The sensitivity for predicting the presence of hematogenous dissemination of the classifier was only 56% and the specificity of the classifier was only 42%. SAM did not identify probes which were significantly differentially expressed between the two groups. Gene Set Enrichment Analysis for predefined gene sets was applied to the gene expression data. The top five of significantly enriched gene sets are all related with cancer specific immunity. The leading-edge subset consisted of a subgroup of genes consisting of CD3D, CD3G, CD247 (CD3H) and CD4 genes.

Conclusions
The results of the present study indicate that with use of gene expression profiling it is not possible to predict the development of hematogenous dissemination in our patients with adenocarcinoma of the esophagus. Moreover, differentially expressed genes could not be identified. Only with use of gene set enrichment analysis pathways could be identified that are related with hematogenous dissemination.
Introduction

The incidence of adenocarcinoma of the esophagus is rapidly rising in the Western World\textsuperscript{1-3}. It is an aggressive disease and is known for its early lymphatic and hematogenous dissemination\textsuperscript{4-6}. When distant metastasis and local irresectability are absent, surgery is the best curative treatment option. However, esophagectomy is accompanied by a high morbidity and substantial mortality rate\textsuperscript{6,7}. Moreover, many patients have unrecognized metastatic disease at the time of first presentation (comprehensive preoperative staging) and later present with recurrent disease after esophagectomy. The majority of patients presents with recurrent disease within two years after surgery and, of these patients, the majority has hematogenous recurrences\textsuperscript{5,8-10}. Even the additional value of FDG-PET scanning in detecting these distant metastases is limited\textsuperscript{11}.

Numerous genetic changes have been described that are related with prognosis and the process of hematogenous dissemination\textsuperscript{12}. These genetic changes are all part of a multistep process (tumor cell detachment, local invasion, intravasation in the circulation, survival in the circulation, extravasation and regrowth in different organs) that a primary tumor must complete before the tumor can disseminate to distant sites. The value of single genes in prognostication is limited\textsuperscript{12} and this is a reason to look at multiple genes together. Completion of the sequencing of the human genome in 2003 marked the dawn of a new era of human biology and medicine\textsuperscript{13}. With the development of DNA microarray technologies it has been made possible to analyze the expression of the whole human genome simultaneously\textsuperscript{14}. The original enthusiasm about the clinical applicability of this new technology has somewhat silenced\textsuperscript{15}, but gene expression analysis with microarrays has revealed important prognostic information in a variety of tumors\textsuperscript{14,16-18}.

In adenocarcinoma of the esophagus, microarray research has focused on the progression of Barrett’s esophagus into early cancers and discrimination of normal tissue from cancer tissue\textsuperscript{19-22}. Analyzing gene expression in advanced esophageal adenocarcinomas might lead to a prognostic profile for an individual patient. Patients with a profile for hematogenous dissemination probably can not be cured with surgery alone and might benefit from systemic (neo-) adjuvant chemotherapy. Furthermore, the analysis of gene expression can advance biological insight in molecular processes underlying the process of hematogenous dissemination by identifying candidate genes and pathways\textsuperscript{23,24}. Therefore, the aim of this study was to discover a gene expression signature that predicts the development of hematogenous dissemination and to identify genes and pathways that provide insight in the mechanisms underlying hematogenous dissemination in patients with adenocarcinoma of the esophagus.
Patients and Methods (Figure 1)

Esophageal tumor selection criteria
Between January 1993 and January 2004, a consecutive series of 385 patients underwent potentially curative esophagectomy for an adenocarcinoma of the esophagus or gastroesophageal junction. Patients did not receive preoperative chemo- and/or radiation therapy. Clinicopathological data from all operated patients were prospectively collected in a database. All patients were followed in the outpatient department every 3-6 months up to a period of five years or until death. Complete data are available concerning the pattern of locoregional and/or hematogenous cancer recurrence and long-term survival. Fresh frozen tumor material was stored in a tissue bank and available for roughly 70% of the patients. Information from the database guided the selection of samples from the tissue bank. To avoid the influence of confounding factors as much as possible, strict criteria were applied to the selection of the patients. Two groups were created; tumors that developed clinically manifest hematogenous dissemination within 4.5 years after initial surgery. To exclude the possibility of metastasis from metastases, tumors did not have evidence of simultaneously locoregional recurrence. The second group consisted of tumors without hematogenous dissemination during at least 4.5 years follow-up (also without signs for locoregional recurrences) (Figure 1A).

RNA isolation
All samples were snap-frozen in liquid nitrogen and stored at -80°C. Only samples with >70% of tumor cells were selected. Tumor percentage was estimated on 5μm hematoxylin and eosin sections by two researchers (SL and FtK). Frozen sections of 10 μm thickness were cut with a cryo-microtome and carefully transferred into a chilled 2ml tube. Total RNA was extracted using TRIzol reagent (Invitrogen, Paisly, UK). Twenty-five micrograms of total RNA was treated with DNase using the Qiagen RNase-free DNase kit and RNeasy spin columns (Qiagen, Venlo, The Netherlands). Quality control of total RNA samples was performed with the RNA 6000 pico LabChip Kit (Agilent Technologies, Palo Alto, CA, USA) and analyzed on the Agilent 2100 bioanalyzer (Agilent Technologies)(Figure 1B).

RNA amplification, labeling, hybridization and scanning
The RNA amplification, labeling, hybridization and data extraction were performed at ServiceXS (Leiden, the Netherlands). Briefly, the mRNA was double amplified using the Amino Allyl MessageAmp kit (Ambion, Austin, TX, USA) and 20 to 100 ng total RNA was reversely transcribed into cDNA and further transcribed in vitro into cRNA. Before labeling, all samples (including ten technical replications) were randomized to avoid confounding by extraneous conditions. Subsequently, RNA was labeled with Cy3 or Cy5 using Amino Allyl MessageAmp RNA Kit (CyDye Amersham Pharmacia Biotech, Diegem, Belgium) according to the manufacturer’s protocol. Five micrograms of Cy3-labelled cRNA from one esophageal tumor was mixed with the same amount of reverse color Cy5-labelled product.
from a reference pool. This so-called reference pool consisted of equal amounts of cRNA from each individual tumor (Figure 1C). Dye labeled cRNA was purified and the samples were checked on concentration and dye incorporation with the RNA 6000 pico LabChip Kit (Agilent Technologies) and analyzed on the Agilent 2100 bioanalyzer (Agilent Technologies). The fluorescently labeled cRNA was used for microarray hybridization. Before hybridization, again all samples were randomized to avoid bias. Hybridization was performed using a 60-mer oligonucleotide 44k whole genome microarray (Agilent Products, Palo Alto, CA, USA) representing over 41,000 human genes and transcripts. After hybridization, the slides were washed and scanned (Figure 1D) with a confocal laser scanner (Agilent Technologies). Default settings of the Agilent Feature Extraction pre-processing protocol were used to obtain raw intensity values from the scans. Exact protocol and parameter settings are described in the Agilent Feature Extraction Software User Manual 7.5 (http://chem.agilent.com/scripts/LiteraturePDF.asp?iWHID=37629). The intensity data were normalized by applying variance stabilization normalization (VSN). For each array and all genes the generalized log-ratio of common reference (Cy5) and tumor (Cy3) VSN transformed intensities was calculated. Log-ratios of replicate arrays were averaged. Quality control of the Agilent Feature Extraction results was performed using methods available from Bioconductor packages (marray, limma, arrayQuality) in the statistical software package R (http://www.r-project.org/)(Figure 1E). Intensities of negative and positive controls were inspected, as well as spatial effects, M/A plots, and signal-to-noise distributions. Only arrays of high quality were included. Analysis of the ten technical replicates confirmed reproducibility across days.

Class prediction
To test whether the development of hematogenous dissemination can be predicted from the gene expression profiles obtained, a naive Bayes classifier was used. The classifiers were validated with the repeated random sampling strategy as described by Michiels et al. Such a random validation strategy enables reliable estimation of the error rate of a classifier and the corresponding confidence interval. We divided the data set (N=77) into
500 training sets (size n) and 500 associated validation sets (size N–n) using resampling without replacement. Resampling was done such that the proportion of samples of tumors which developed hematogenous dissemination and those which did not in training and validation sets was comparable to the proportion in the full data set. For each training set a molecular signature was identified from the 10, 20, 30, 40 or 50 genes for which expression was most highly correlated with prognosis as determined by the t-statistics between samples which developed hematogenous dissemination and which did not. The optimal number of genes for inclusion in the classifier was selected with 5-fold cross-validation on the training set. Accuracy (proportion of correctly classified samples), specificity (proportion of correctly predicted samples without hematogenous dissemination), and sensitivity (proportion of correctly predicted samples with hematogenous dissemination) of the resulting classifier were estimated for each associated validation set. This set-up guarantees independent validation of the classifier since the validation data are not involved in gene selection and training of the classifier. To study the influence of sample size, the size of the training set n was varied from 17 to 75 in steps of two. Remaining samples were attributed to the validation set that, therefore, varies from 60 to 2 samples. The complete analysis was done with R scripts based on the Bioconductor package MCRestimate.

Differentially expressed genes
The normalized log-ratios were analyzed for differential expression between tumors which developed hematogenous dissemination and those which did not. This analysis was done using significance analysis of microarrays (SAM) with the samr package from Bioconductor. Differentially expressed genes were selected by controlling the false discovery rate at 5% (FDR-q-value) with 500 permutations.

Gene Set Enrichment Analysis (GSEA)
Gene set enrichment analysis (GSEA) was used to uncover modest but correlated changes on a gene set level and extract biological insight from the gene expression data. The R-GSEA R package (available from http://www.broad.mit.edu/gsea/index.html) was used. The gene sets used are available from the Molecular Signature Database (MsigDB, http://www.broad.mit.edu/gsea/index.html). In the present series, the C2 catalog of 522 functional sets (mainly metabolic and signaling pathways) was used.

Cluster analysis was used to identify a common biological process within all significantly up- and downregulated pathways. For this purpose, leading-edge subsets, which can be interpreted as the core of a gene set that accounts a pathway to be significantly up- or downregulated, were examined and clustered.
Results

Esophageal tumor selection criteria

Thirty-three of the 77 selected patients developed hematogenous dissemination and consequently died during follow-up. Metastasis occurred at one of the following sites; liver metastasis in 19 patients (58%), bone metastasis in 9 patients (27%), brain metastasis in 3 patients (9%), muscle metastasis in one patient (3%) and a skin metastasis in one patient (3%). Patient and tumor characteristics of patients who developed hematogenous dissemination and of patients who did not develop hematogenous metastasis were not significantly different for age, depth of tumor penetration and differentiation grade. The development of hematogenous recurrence without signs of lymph node dissemination was very rare in our patients. Only two of the 33 patients were N0 and developed only hematogenous dissemination (liver and bones) without evidence for locoregional recurrences. Consequently, patients who had hematogenous dissemination had significantly more lymphatic dissemination (Table 1).

Table 1: Clinicopathological characteristics of patients who underwent potentially curative surgery for adenocarcinoma of the esophagus with and without hematogenous dissemination during follow-up.

<table>
<thead>
<tr>
<th></th>
<th>Hematogenous dissemination present (N=33)</th>
<th>Hematogenous dissemination absent (N=44)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age* [yrs]</td>
<td></td>
<td></td>
<td>0.228</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- male</td>
<td>65 (45-79)</td>
<td>67 (44-78)</td>
<td>0.228</td>
</tr>
<tr>
<td>- female</td>
<td>3</td>
<td>10</td>
<td>0.114</td>
</tr>
<tr>
<td>Tumor length* [cm]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumor penetration</td>
<td>5.8 (1.5-11.0)</td>
<td>6.0 (1.2-10.0)</td>
<td>0.505</td>
</tr>
<tr>
<td>- pT2</td>
<td>3</td>
<td>9</td>
<td>0.216</td>
</tr>
<tr>
<td>- pT3</td>
<td>30</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>N-stage</td>
<td></td>
<td></td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>- N0</td>
<td>2</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>- N1</td>
<td>31</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Differentiation</td>
<td></td>
<td></td>
<td>0.349</td>
</tr>
<tr>
<td>- moderate</td>
<td>11</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>- poor</td>
<td>22</td>
<td>24</td>
<td></td>
</tr>
</tbody>
</table>

* = numbers are given as median (range)
N0 = lymph node dissemination absent
N1 = lymph node dissemination present
Class prediction
A feature selection and classification approach was pursued to identify a prognostic profile. To predict the presence of hematogenous dissemination, thousands of signatures were estimated (500 for every validation set size). The average classification accuracy (and 95% CI), as a function of the training set size for the naïve Bayes classifier, is shown in Figure 2. Since the mean accuracy was constant for increasing sample size, we further only discuss representative results for a training set of 61 samples and a validation set of 16 samples. The list of genes included in the signature was very unstable. The seven probes, which were selected in at least 200 of the 500 signatures, are shown in Figure 3. The naïve Bayes classifier had an average accuracy of 49±18% and therefore did not perform better than random

![Figure 2: Average classification accuracy (black line) with its 95% confidence interval (red lines) in 500 validation sets as a function of training set.](image)

![Figure 3: Gene names belonging to probes included in at least 200 of 500 molecular signatures](image)

GRAP = GRB2-related adaptor protein, DNAH11 = dynein, axonemal, heavy polypeptide 11, MYST4 = MYST histone acetyltransferase monocytic leukemia 4, ALDH161a = aldehyde dehydrogenase 16 family, member A1, DDX50 = DEAD (Asp-Glu-Ala-Asp) box polypeptide 50, JMDJ1c = jumonji domain containing 1C
Table 2: Enriched gene sets associated with hematogenous dissemination in patients with adenocarcinoma of the esophagus. The top of downregulated gene sets are given. All shown gene sets have an FDR-q-value of <0.05. Gene sets are sorted by FDR-q-value.

<table>
<thead>
<tr>
<th>Gene Set Description</th>
<th>FDR-q-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Macrophages activate NK cells by releasing IL-12, which induces NK cytotoxic activity</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Csk inhibits T-cell activation by phosphorylating Lck</td>
<td>0.0004</td>
</tr>
<tr>
<td>Antigen-presenting dendritic cells and macrophages secrete IL-12, which induces Th1 cell differentiation.</td>
<td>0.0005</td>
</tr>
<tr>
<td>Cancer related genes involved in immune function</td>
<td>0.0007</td>
</tr>
<tr>
<td>The kinases Lck and Fyn phosphorylate and activate the T cell receptor, which leads to T cell activation.</td>
<td>0.0007</td>
</tr>
<tr>
<td>Endogenous anti-thrombosis pathways are overwhelmed in plaque-narrowed blood vessels, resulting in potentially lethal myocardial infarction.</td>
<td>0.0011</td>
</tr>
<tr>
<td>T cell activation requires interaction with an antigen-MHC-I complex on an antigen-presenting cell (APC), as well as CD28 interaction with the APC’s CD80 or 86</td>
<td>0.0022</td>
</tr>
<tr>
<td>Any series of molecular signals initiated by the binding of an extracellular ligand to a receptor on the surface of the target cell.</td>
<td>0.0024</td>
</tr>
<tr>
<td>Helper T cells coordinate the actions of B cells, macrophages, and other immune cells via surface molecules such as T cell receptor/CD3 and their characteristic marker CD4.</td>
<td>0.0024</td>
</tr>
<tr>
<td>B cells express the major histocompatibility complex (class II MHC), immunoglobulins, adhesion proteins, and other factors on their cell surface</td>
<td>0.0028</td>
</tr>
<tr>
<td>Helper T subtype Th1 produces pro-inflammatory cytokines that stimulate phagocytosis, while Th2 cells promote antibody production and activate eosinophils.</td>
<td>0.0047</td>
</tr>
<tr>
<td>Cytotoxic T cells release perforin and granzyme to lyse foreign cell targets and express Fas ligand to promote Fas-induced apoptosis</td>
<td>0.0065</td>
</tr>
<tr>
<td>HIV infection upregulates Fas ligand in macrophages and CD4 in helper T cells, leading to widespread Fas-induced T cell apoptosis</td>
<td>0.0167</td>
</tr>
<tr>
<td>Cytotoxic T lymphocytes induce apoptosis in infected cells presenting antigen-MHC-I complexes via the perforin and Fas/Fas ligand pathways</td>
<td>0.0224</td>
</tr>
<tr>
<td>The interferon gamma pathway resembles the JAK-STAT pathway and activates STAT transcription factors</td>
<td>0.0234</td>
</tr>
<tr>
<td>On activation of the T cell receptor, phospholipase C is activated to produce second messengers DAG and PIP3, both required for T cell activation</td>
<td>0.0343</td>
</tr>
</tbody>
</table>

For whole table see: GSEA-DOWNREGULATEDinM1

classification. The sensitivity for predicting the presence of hematogenous dissemination of the naïve Bayes classifier was only 56% and the specificity of the classifier was only 42%.
Table 3: Enriched gene sets associated with hematogenous dissemination in patients with adenocarcinoma of the esophagus. The top 10 of upregulated gene sets are given. All gene sets are significant (p<0.05) with an FDR of <0.20. Gene sets are sorted by the FDR-q-value.

<table>
<thead>
<tr>
<th>Gene Set</th>
<th>Description</th>
<th>FDR-q-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GNF_FEMALEGENES</td>
<td>female reproductive tissue expressed genes</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>vipPathway</td>
<td>Apoptosis of activated T cells is inhibited by vasoactive intestinal peptide (VIP)</td>
<td>0.1026</td>
</tr>
<tr>
<td>insulin genes</td>
<td>genes regulated by insulin</td>
<td>0.1140</td>
</tr>
<tr>
<td>extrinsicPathway</td>
<td>The extrinsic prothrombin activation pathway requires the release of thromboplastin from damaged tissues to activate the blood clotting cascade</td>
<td>0.1597</td>
</tr>
<tr>
<td>pkcPathway</td>
<td>Gq-coupled receptors promote hydrolysis of PIP2 to DAG and IP3, which causes calcium influx and activates protein kinase C.</td>
<td>0.1882</td>
</tr>
<tr>
<td>ecmPathway</td>
<td>Extracellular matrix induces integrin-mediated FAK phosphorylation in epithelial cells</td>
<td>0.2052</td>
</tr>
<tr>
<td>leptinPathway</td>
<td>Leptin promotes fatty acid oxidation, decreases cells’ lipid content, and promotes insulin sensitivity.</td>
<td>0.2205</td>
</tr>
<tr>
<td>KRAS_TOP100_KNOCKDOWN</td>
<td>Genes upregulated in kras knockdown vs control in a human cell line</td>
<td>0.2453</td>
</tr>
</tbody>
</table>

For whole table see: GSEA-UPREGULATEDinM1

Differentially expressed genes
SAM did not identify probes which were significantly differentially expressed. All FDR-q-values were higher than 0.3. (see supplementary file DEgeneshematogenous disseminationcomplete.xls for FDR for all probes).

Gene Set Enrichment Analysis (GSEA)
To increase insight in the underlying molecular processes responsible for hematogenous dissemination, GSEA for predefined gene sets was applied to the gene expression data. Predefined pathways that were downregulated (FDR q-value<0.05) in patients who developed hematogenous dissemination are shown in Table 2. In the supplementary file GSEA-DOWNREGULATEDinM1 the complete list of downregulated pathways can be found. The top five of significantly enriched gene sets are all related with cancer specific immunity. Two of these pathways are related with the stimulating effect of IL-12 on T-helper 1 and a cytotoxic T lymphocyte response. Only one enriched gene set was upregulated with an FDR<0.05.
Predefined pathways that were up-regulated (FDR q-value<0.25) in tumors with hematogenous dissemination are shown in Table 3. In the supplementary file, GSEA-UPREGULATEDinM1 the complete list can be found.
To explore whether significant pathways reflect a common biological function, the leading-edge subset (core of a gene set that accounts for the enrichment signal of the gene set) for
each gene set was examined. A common subgroup of genes consists of CD3D, CD3G, CD247 (CD3H) and CD4 genes (Figure 4).

Discussion

Although still relatively rare, the incidence of adenocarcinoma of the esophagus is rapidly rising in the Western World. Even after potentially curative surgery the majority of patients develop recurrent disease, mainly due to the development of hematogenous dissemination. The ability of the primary tumor to develop hematogenous dissemination might be explained by the role of genetic alterations in the primary tumor. The present study analyzed the gene expression of 77 tumor samples from patients with adenocarcinoma of the esophagus. The results of the present study indicate that with use of gene expression profiling it is not possible to predict the development of hematogenous dissemination in our patients with adenocarcinoma of the esophagus. Moreover, differentially expressed genes could not be identified. Only with use of gene set enrichment analysis pathways could be identified that are related with hematogenous dissemination.

The original enthusiasm about the clinical applicability of microarrays has somewhat dampened, because of lack of reproducibility with different patient cohorts, platforms, and bioinformatic analysis methods. One reason is that standard validation with a single control group is not sufficient. In a previous study it was found that with use of a multiple random validation strategy, classification accuracy was highly unstable and its prognostic strength depended on the selection of patients in the training sets. In this study, seven microarray
cancer studies were reanalyzed and it was shown that previously published prognostic results were overoptimistic in almost all studies. Therefore, in the present study, validation by repeated random sampling was used. It was shown that, even in a carefully selected patient population, classification accuracies were also highly variable between random splits of patients. This clearly illustrates the bias that might be introduced by reporting results on only a single split in training and validation set. Although gene expression analysis was performed in a relatively large number of samples, this number is more than 500 times lower than the 42,000 probes analyzed. Theoretically, an increase in sample size may identify profiles with increased predictive power. However, since the mean classification accuracy did not increase with increased training set size; this might not lead to a better predictive power. After correcting for multiple hypotheses testing, no individual gene met the threshold for statistical significance. This may be explained by the fact that biological differences are modest relative to the noise inherent to the microarray technology. Moreover, to extract meaningful biological insight across the immense number of genes and uncover modest but coordinate changes on a gene set level (instead of looking to the most up- and down regulated genes only) GSEA\textsuperscript{23} was applied. This method seems robust since it allows a more reliable comparison of different platforms, projects, laboratories, and public access databases\textsuperscript{32,33}. With GSEA several gene sets that were associated with hematogenous dissemination were found. The leading-edge analysis showed that a downregulated subset containing CD3 and CD4 genes were mainly associated with the development of hematogenous dissemination (CD3 and CD4 are both T-helper cell surface molecules). Theoretically, it might be possible that specific antigens are erroneously expressed on the membrane of the tumor cells. However, it is more probable that the CD3 and CD4 genes are expressed by tumor-infiltrating lymphocytes. The present study was originally designed to detect genetic alterations of the primary epithelial tumor cells. However, macrodissected samples (with at least 70% tumor cells) were included and thus contain a certain amount of stromal cells, including lymphocytes. This setup was chosen since laser captive microdissection may result in degradation of RNA and results in a limited amount of RNA only\textsuperscript{34}. It is well possible that the modest changes of immunological cells, which were included in the specimen, could only be detected with GSEA, since the study did not focus on the immune infiltrates of the tumor and effects in immune cells are “overshadowed” by the RNA concentrations of tumor cells. Tumor cells can become targets for a T cell–mediated adaptive immune response. Although the role of the immune response in controlling tumor growth and cancer recurrence is controversial, recent research in patients with colorectal cancers indicated that type, density and location of immune cells within tumor tissues were a better predictor of patient survival than the current methods to stage colorectal cancers\textsuperscript{35,36}. Interestingly, colorectal tumors from patients without recurrence had higher immune cell densities of CD3 positive cells within the tumor and at the infiltrating tumor margin. Also in esophageal adenocarcinoma there is evidence, suggesting that T-cell infiltrates have a beneficial prognostic impact\textsuperscript{37}. It was shown that CD8 positive T cell infiltrations within the tumor specimen have a favorable outcome\textsuperscript{38}. Furthermore, it was shown that, the ability of T cells in serum to produce IL-2
was related to long term survival, indicating a crucial role of TH1-type cells in anti-tumor immunosurveillance\(^3\)\(^9\).

Clearly, the results of the present study must be further validated. To evaluate the effect of the T cell-mediated adaptive immune response, immunohistochemistry of diverse T-cells (CD3, CD4 and CD8) in the primary tumor will be performed. However, the results of the present study suggest that, with use of GSEA, it is possible to detect modest changes in a microarray experiment, facilitating the interpretation of a large-scale experiment by identifying pathways and processes, which tend to be more reproducible and more interpretable. Moreover, with use of GSEA it is possible to more reliably compare different platforms, projects, laboratories, and public access databases\(^3\)\(^2\);\(^3\)\(^3\).

In conclusion, in the present study it was not possible to predict hematogenous dissemination in esophageal adenocarcinoma with use of gene expression profiling. However, several pathways were identified, which seem to be important in the development of hematogenous dissemination. In future, discovered pathways may be targeted by specific therapeutic strategies.

### Reference List


