Molecular-cytogenetic characterization of head and neck cancer: Identification of novel prognostic factors and gene targets for therapy [double dissertation 2]
Wreesmann, V.B.; Singh, B.

Citation for published version (APA):
Chapter 12

Molecular Classification of Head and Neck Squamous Cell Carcinoma Using cDNA Microarrays

Thomas J. Belbin, Bhuvanesh Singh, Ilana Barber, Nicholas Socci, Bruce Wenig, Richard Smith, Michael B. Prystowsky, Geoffrey Childs

Gene Expression Profiling in Head and Neck Cancer

Abstract

Squamous cell carcinomas of the head and neck constitute an anatomically heterogeneous group of neoplasms that share in common a causal association with tobacco and alcohol exposure. The clinical course of these neoplasms is difficult to predict based on established prognostic clinicopathological criteria. Given the genetic complexity of head and neck cancers, it is not surprising that correlations with individual genetic abnormalities have also been disappointing. Several authors have suggested that global gene expression patterns can be used to subgroup patients with cancer. Here we report the use of cDNA microarrays containing 9216 clones to measure global patterns of gene expression in these neoplasms. We have used a statistical algorithm to identify 375 genes, which divide patients with head and neck tumors into two clinically distinct subgroups based on gene expression patterns. Our results demonstrate that gene expression profiling can be used as a predictor of outcome.

Head and neck squamous cell carcinoma (HNSCC) is the fifth most common malignancy worldwide, representing a major international health problem. The 5-year survival rate for this disease (~50%) has improved only marginally over the past decade; as a result, >30,100 cases and 7,800 deaths occur annually in the United States.[1] These tumors arise from diverse anatomical locations, including the oral cavity, oropharynx, hypopharynx, larynx, and nasopharynx, but have in common an etiological association with tobacco and/or alcohol exposure.[2] Treatment planning for HNSCC depends largely on anatomical staging of the disease at presentation. Conventional treatment with surgery and/or radiation therapy and/or chemotherapy is associated with significant morbidity affecting speech, swallowing, and overall quality of life. Despite these interventions, recurrence of the disease is observed in ~50% of patients, either locally, regionally, or at a distant site with high rates of associated mortality.[3]

Treatment failures for HNSCC can be attributed to multiple factors but remain difficult to predict. The majority of the patient population present advanced stages in the disease when the risk of nodal metastasis and recurrence has increased. Additionally, mechanisms of disease progression are not well understood. Finally, there are few, if any, molecular markers that can be reliably used in early detection or as indicators of prognosis. In fact, the overall genetic and molecular basis of HNSCC remains ill defined. Genetic aberrations have been investigated as markers of disease progression and/or outcome, including gains of 3q, 8q, 9q, 20q, 7p, 11q13, and 5p and losses of 3p, 9p, 21q, 5q, 13q, 18q, and 8p.[4] More recently, several studies involving cDNA microarrays have identified some genes whose expression has changed in HNSCC samples compared with normal tissue.[5, 6]

Here we report the use of cDNA microarrays to measure global patterns of gene expression in HNSCC tissue. Measurements of global gene expression have been the basis for molecular classification of other forms of cancer. Golub et al. [7] successfully used cDNA microarrays to identify molecular distinctions between two forms of acute
leukemia with very different treatment regiments: acute myeloid leukemia and acute lymphoblastic leukemia. By using those genes that best discriminate between the two forms of the disease as "class predictors," it was possible to predict the subtype of acute leukemia (acute myeloid leukemia versus acute lymphoblastic leukemia) without previous biological knowledge of the cases. Classification based on patterns of global gene expression has also been reported for diffuse large B-cell lymphoma, where cDNA microarrays were successfully used to identify two molecularly distinct forms of the disease.[8] We have used a statistical algorithm to divide patients with head and neck tumors into two clinically distinct subgroups based on gene expression patterns and identify genes whose expression correlates highly with these clinically identifiable phenotypes. Our results suggest that gene expression profiling can be used as a predictor of outcome in HNSCC.

**Materials and Methods**

**Tissue Samples and RNA Extraction**

Primary tumor tissue was obtained at time of surgery for HNSCC under the auspices of the Institutional Review Board, after obtaining informed consent. Only patients undergoing surgical treatment with curative intent, having a history of smoking and no prior nonsurgical treatment for head and neck cancer, were included in the study. To rule out gene expression alterations because of stromal cell contamination, we confirmed that each tumor specimen used in our study contained >70% cancer cells by analysis of corresponding H&E-stained sections. Tissue samples were flash frozen in liquid nitrogen and stored at -80°C before RNA extraction. Tissues were homogenized using a Brinkmann Model PT 10/35 Tissue Homogenizer. Total RNA was extracted using the RNeasy Maxi RNA purification kit following the protocol of the manufacturer (Qiagen, Valencia, CA).

Normal human epithelial keratinocytes (Adult Skin, NHEK-6168) were obtained commercially and maintained in keratinocyte growth medium according to the recommendations of the manufacturer (BioWhittaker, San Diego, CA). Although not from the head and neck mucosa, the gene expression profiles of these keratinocytes is similar to normal mucosa and close in nature to the tumor cell type (expressing largely the same genes) and provided a denominator value for accurate ratio determination for the genes on the microarray.

**Preparation of cDNA Microarrays**

A set of 9216 sequence-verified human IMAGE cDNA clones representing both known genes and expressed sequence tags were spotted onto polylysine-coated microscope slides using a custom robot designed and built at Albert Einstein College of Medicine. Before hybridization, slides were preprocessed as described previously. (Internet address: http://sequence.aecom.yu.edu/bioinf/funcgenomic.html)

**Labeling of cDNA and Hybridization to Arrays**

Labeling and hybridization of cDNA to arrays was carried out using the 3DNA Submicro Expression Array Detection Kit (Genisphere, Montvale, NJ) with some mod-
ifications. Approximately 1 μg of total RNA from each sample was used to synthesize cDNA with 200 units of Superscript II reverse transcriptase (Invitrogen/Life Technologies, Inc., Carlsbad, CA) using Cy3 and Cy5 RT primers (0.2 pmol) supplied with the kit. Reactions were carried out at 42°C for 2 h in a buffer containing 50 mM Tris-HCl (pH 8.3), 75 mM KCl, 3 mM MgCl2, 0.5 mM deoxynucleotide triphosphate, 10 mM DTT, and 40 units of RNaseOUT RNase (Invitrogen/Life Technologies, Inc.). Hybridization to cDNA arrays was carried out overnight at 50°C in a buffer containing 30% formamide, 3 X SSC, 0.75% SDS, and 100 ng of human Cot-1 DNA. After hybridization, slides were washed 10 min at 50° C with 2 X SSC, 0.2% SDS, 15 min at room temperature in 2 X SSC, and 15 min at room temperature in 0.2 X SSC. Slides were immediately dried as before and scanned using a GenePix 4000A microarray scanner (Axon Instruments, Foster City, CA).

Data Collection and Statistical Analysis

Red (Cy5) and green (Cy3) signal intensities for each element on the array were calculated using the GenePix Pro 3.0 software (Axon Instruments). This software gives an integrated intensity per spot for each channel in addition to an integrated background count. Plots of the fold change versus the average intensity were examined to look for abnormalities in single-array data. It is common to plot a red versus green channel intensity scatter plot to examine for problems; however, we found that transforming to fold change versus average intensity displayed the data in a more easily viewable form. We removed any samples with anomalous scatter plots. The remaining samples were normalized using an intensity-dependent normalization believed to be more accurate than normalization with a fixed constant.[9] Normalized fold changes in gene expression were then used to cluster the various patient samples. Data were filtered to select genes, which had both a significant average intensity and fold R:G ratio; the absolute value of the fold change had to be >2, and the average intensity had to be >300. The exact values used are largely arbitrary in that we simply wanted to raise the level of signal to noise. The average intensity cutoff was set by looking at the scatter plots from so-called "yellow tests," where the same sample is labeled both with Cy5 and Cy3. The average error in the fold change was intensity dependent and increased with decreasing average intensity. The average intensity cutoff was set higher than this divergence in fold change error. We also insisted that each gene meet these criteria for at least two samples. A total of 906 genes fit these criteria. We clustered tumor samples using the dot product (angle) metric to measure the distance between the samples. We clustered these distances using a hierarchical Wards method.

To assess the magnitude of the error in the R:G ratio data, a single tumor sample array experiment was replicated eight times. These data were used to determine the intensity-dependent error in the fold change and used both to set the intensity cutoff as described above and as input for the parametric bootstrapping used to measure cluster significance. We made the assumption based on the analysis of the replicated data that the noise was additive and normal for the logarithm of the fold change. We used a stan-
standard parametric bootstrap resampling technique to determine the relative significance of the clustering results. The basis for the parametric bootstrap method is that one has a model for the errors in the data, allowing model parameters to be estimated. Using this model and the estimated parameters, one can generate bootstrap copies of the original data. By fitting a curve to the scatter of the SD of the log fold change of the eight replicates as a function of average log intensity, we obtained a curve for the average noise as a function of intensity. Copies (100) of the data set were generated by sampling from this noise model and adding to the logarithm of the original fold change. Each of the 100 bootstrap samples was clustered to give 100 trees. A consensus was built using the Consensus program for the Phylip package. (Internet address: http://evolution.genetics.washington.edu/phylip/software.html) The program constructed a tree by finding for each node the pairing that occurred most often in the 100 separate trials. It displayed this count at each node of the tree. Nodes with values close to 100 were more significant in that this pairing was robust when noise was added to the data set.

The selection of genes for further examination in Fig. 3 was based on a ranking of genes which best discriminated between the two groups using an elemental t test. Clones for which no functional information was available (expressed sequence tags) were eliminated from the final list. The remaining genes were grouped into broad functional categories. A complete list of genes and corresponding GenBank numbers is available as Supplementary Data on our laboratory website. (Internet address: http://www.microarray.info/Childs_Lab)

**Results and Discussion**

We categorized 17 patients with HNSCC by comparison of the global gene expression patterns between primary tumor samples and normal human adult keratinocytes using cDNA microarrays containing 9216 genes. The study population was representative of the general population with HNSCC, having a median age of 65.5 years at presentation and a male predominance (56%; Table 1). It should be pointed out that non-smokers were not included in the study population to control for the influence of smoking on the gene expression profile. Furthermore, although human papillomavirus infection plays an etiological role in HNSCC development in 10-20% of cases, human papillomavirus analysis was not performed, as it was not the primary focus of our study. \[11\]

By clustering patients based on the analysis of a subgroup of 906 genes (see "Materials and Methods"), it was possible to subclassify these tumors into distinct groups. Each genotypic group contained 8 patients; 1 patient (MSK76) did not cluster with either of these two groups (Fig. 1). On further analysis, we identified 375 genes that discriminate between the genotypic subtypes of HNSCC (Fig. 1). We summarized corresponding clinical data for each of the patient groupings defined by gene expression profiling (Table 2). Although patients in Group I were younger than those in Group II and universally had poorly differentiated tumors, they had more early stage tumors and a lower prevalence of nodal metastasis at presentation. Patients in Group I had lower cause-specific \( P=0.057 \) and overall survivals relative to those in Group II (Fig. 2).
Group I patients were followed for a median of 3 years, and 38% died of their cancer. However, 58% of patients died within 2 years overall, reflecting the impact of comorbid conditions. Of the Group II patients, none died as a consequence of their cancer over a median follow-up period of 4 years, but 37% died of other causes. The impact of comorbidity on survival was evident in this study population and is likely attributable to the high rates of smoking and alcohol use in this population of patients.[12] Similar to our observations, another study showed that of patients with HNSCC, the probability of dying before 3 years was 44%; 65% died of their cancer, 18% would have died anyway without cancer, and 16% died because of the increased risk of their comorbid conditions.[13] Overall, patient segregation by gene expression profiling was a better predictor of outcome in our study population than established clinicopathological variables.

We selected a representative subset of the genes that genotypically dichotomized our study population for further investigation (Fig. 3). The absence of genes already associated with HNSCC is not surprising: (a) some genes, such as p53, cyclin D1, and p16, were not present on the arrays used in the study; and (b) because our study specifically addressed differences in gene expression between the primary tumors themselves, we

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age</th>
<th>Sex</th>
<th>Tobacco</th>
<th>Alcohol</th>
<th>Location</th>
<th>TNM stage¹</th>
<th>Pathological differentiation</th>
<th>Treatment</th>
<th>Site of recurrence</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSK8</td>
<td>68</td>
<td>F</td>
<td>+</td>
<td>+</td>
<td>Oropharynx</td>
<td>T2N2, IV</td>
<td>Poor</td>
<td>Surgery/radiation</td>
<td>None</td>
<td>DOD</td>
</tr>
<tr>
<td>MSK9</td>
<td>69</td>
<td>M</td>
<td>+</td>
<td>-</td>
<td>Oral cavity</td>
<td>Recurrent (T2N0, II)</td>
<td>Poor</td>
<td>Surgery/radiation</td>
<td>None</td>
<td>DOC</td>
</tr>
<tr>
<td>MSK53</td>
<td>71</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>Oral cavity</td>
<td>T1N0, I</td>
<td>Poor</td>
<td>Surgery</td>
<td>Neck</td>
<td>DOC</td>
</tr>
<tr>
<td>MSK57</td>
<td>58</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>Oropharynx</td>
<td>T2N0, II</td>
<td>Poor</td>
<td>Surgery/radiation</td>
<td>None</td>
<td>DOC</td>
</tr>
<tr>
<td>MSK72</td>
<td>59</td>
<td>F</td>
<td>+</td>
<td>+</td>
<td>Oral cavity</td>
<td>T2N0, II</td>
<td>Poor</td>
<td>Surgery</td>
<td>Neck</td>
<td>DOC</td>
</tr>
<tr>
<td>MSK73</td>
<td>57</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>Oral cavity</td>
<td>Recurrent (T2N1, III)</td>
<td>(?)</td>
<td>Surgery/radiation</td>
<td>Neck</td>
<td>AWD</td>
</tr>
<tr>
<td>MSK77</td>
<td>49</td>
<td>F</td>
<td>+</td>
<td>+</td>
<td>Oral cavity</td>
<td>T4N1, IV</td>
<td>Poor</td>
<td>Surgery/radiation</td>
<td>Primary</td>
<td>DOD</td>
</tr>
<tr>
<td>MSK78</td>
<td>69</td>
<td>F</td>
<td>+</td>
<td>-</td>
<td>Larynx</td>
<td>T1N0, I</td>
<td>Poor</td>
<td>Surgery</td>
<td>None</td>
<td>NED</td>
</tr>
<tr>
<td>Cluster group 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSK7</td>
<td>57</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>Larynx</td>
<td>Recurrent (T1N0, I)</td>
<td>Poor</td>
<td>Surgery/radiation</td>
<td>None</td>
<td>NED</td>
</tr>
<tr>
<td>MSK10</td>
<td>74</td>
<td>F</td>
<td>+</td>
<td>+</td>
<td>Oral cavity</td>
<td>T3N2, IV</td>
<td>Poor</td>
<td>Surgery/radiation</td>
<td>None</td>
<td>NED</td>
</tr>
<tr>
<td>MSK11</td>
<td>69</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>Oral cavity</td>
<td>T2N1, III</td>
<td>Poor</td>
<td>Surgery/radiation</td>
<td>None</td>
<td>NED</td>
</tr>
<tr>
<td>MSK12</td>
<td>82</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>Oral cavity</td>
<td>Recurrent (T2N0, II)</td>
<td>Moderate</td>
<td>Surgery/radiation</td>
<td>Primary</td>
<td>DOC</td>
</tr>
<tr>
<td>MSK13</td>
<td>64</td>
<td>F</td>
<td>+</td>
<td>-</td>
<td>Oral cavity</td>
<td>T2N0, II</td>
<td>Well</td>
<td>Surgery/radiation</td>
<td>None</td>
<td>NED</td>
</tr>
<tr>
<td>MSK14</td>
<td>73</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>Oral cavity</td>
<td>T4N1, IV</td>
<td>Poor</td>
<td>Surgery/radiation</td>
<td>None</td>
<td>NED</td>
</tr>
<tr>
<td>MSK23</td>
<td>64</td>
<td>F</td>
<td>+</td>
<td>+</td>
<td>Oropharynx</td>
<td>T2N0, II</td>
<td>Poor</td>
<td>Surgery/radiation</td>
<td>Distant</td>
<td>AWD</td>
</tr>
<tr>
<td>MSK24</td>
<td>97</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>Oral cavity</td>
<td>T3N2, IV</td>
<td>Poor</td>
<td>Surgery/radiation</td>
<td>Primary</td>
<td>DOC</td>
</tr>
<tr>
<td>MSK76</td>
<td>34</td>
<td>M</td>
<td>+</td>
<td>+</td>
<td>Oral cavity</td>
<td>Recurrent (T2N1, II)</td>
<td>Poor</td>
<td>Surgery/radiation</td>
<td>None</td>
<td>DOC</td>
</tr>
</tbody>
</table>

¹ DOD, died of disease; DOC, died other causes; NED, no evidence of disease at last follow-up; AWD, alive with disease at last follow-up; TNM, Tumor-Node-Metastasis.

* No patients at distant metastasis at presentation.
* Based on worst differentiation category reported.
* Stage at initial presentation.
* Not clustered with either group.
would not necessarily expect to see genes identified previously by comparing tumor cells with corresponding "normal" cells.

The Group I-specific genes originate from diverse functional categories and include genes already associated with neoplastic disease states. Carcinoembryonic antigen (CEA) is a highly glycosylated cell surface protein and a member of the human CEA family.[14] In HNSCC, serum CEA levels are believed to be valuable in monitoring tumor response; decline in CEA serum levels has been observed in patients after surgical resection of tumor or after regression of tumor after irradiation.[15] More recently, CEA was shown to be significantly elevated in metastatic squamous cell lung cancer.[16] In immunohistochemical studies of cervical biopsies, CEA staining intensity was significantly increased in high-grade squamous lesions compared with normal cervical mucosa and low-grade intraepithelial neoplasia, leading to the conclusion that CEA expression could be used as a marker for identifying patients at risk for more progressive forms of the disease.[17]

Another disease-related gene identified in our study is the AIB1 gene, a member of the steroid receptor coactivator 1 family. Overexpression of this gene has been observed in 60% of breast carcinomas and 40% of primary gastric cancers.[18, 19] In the case of primary gastric tumors, amplification of AIB1 correlated with a much poorer prognosis than those without AIB1 amplification.

Transporter molecules have also been observed to be up-regulated in Group I tumors, including a human glucose transporter-related protein.[20] Positron emission tomography studies have shown elevated glucose uptake in squamous cell carcinoma of the head and neck.[21] We also identified an amino acid transporter, SLC7A8, which interacts with the 4F2 heavy chain in the complex mediating neutral and/or cationic amino acid transport, as well as cystine/glutamate exchange.[22] These findings are intriguing for several reasons. It has been established that cancer cells use more glucose and amino
acids than the corresponding benign cells.[23] Additionally, there is evidence that transport of neutral amino acids can play a role in altering tumor cell growth and invasion.[24]

We have identified Group I-specific genes whose products are involved in RNA processing and translation. Two ribosomal proteins whose expression was increased in Group I are S6 and S19. Phosphorylation of S6 and the signaling pathway of S6 kinase have been implicated in the control of cellular growth and proliferation by translational up-regulation of mRNAs coding for components of the protein synthetic apparatus. Soluble protein extracts derived from squamous cell carcinomas and adenocarcinomas of the lung exhibited elevated phosphotransferase activity compared with extracts derived from nonneoplastic lung parenchyma adjacent to the tumor.[25] One of the in vitro substrates identified for this phosphotransferase activity was a peptide whose sequence corresponded to the COOH terminus of S6. Amplification of the S6 kinase gene at chromosomal region 17q23 has been shown to correlate significantly with poor prognosis in human breast tumors.[26] S19 has been identified as overexpressed in primary colon carcinoma and primary cervical carcinoma.[27, 28] In fact, increased expression of S19 in colon carcinoma was well correlated with cells of higher malignant potential.

Of the Group I-specific genes coding for metabolic enzymes, a subset is associated with oxygen metabolism. Tumor hypoxia has been shown recently to correlate with poorer treatment outcome and a poorer locoregional tumor control probability in head and neck cancer patients.[29, 30] We observed an increase in expression of enzymes, such as catalase and xanthine oxidase, as well as proteins associated with oxidative phos-

---

**Table 2. Comparison of tumor, patient, and outcome characteristics based on cDNA microarray defined subgroups**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Group 1</th>
<th>Group 2</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample size, n (%)</td>
<td>8 (50%)</td>
<td>8 (50%)</td>
<td></td>
</tr>
<tr>
<td>Male gender, n (%)</td>
<td>4 (50%)</td>
<td>5 (63%)</td>
<td>NS'</td>
</tr>
<tr>
<td>Median age</td>
<td>63.5 years</td>
<td>72.5 years</td>
<td></td>
</tr>
<tr>
<td>Anatomic location, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oral cavity</td>
<td>5 (63%)</td>
<td>6 (75%)</td>
<td></td>
</tr>
<tr>
<td>Larynx</td>
<td>1 (13%)</td>
<td>1 (13%)</td>
<td>NS'</td>
</tr>
<tr>
<td>Oropharynx</td>
<td>2 (25%)</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Hypopharynx</td>
<td>0</td>
<td>1 (13%)</td>
<td></td>
</tr>
<tr>
<td>Recurrent cancer at presentation, n (%)</td>
<td>2 (25%)</td>
<td>2 (25%)</td>
<td>NS'</td>
</tr>
<tr>
<td>Clinical TNM stage, n (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I-II</td>
<td>4 (67%)</td>
<td>2 (33%)</td>
<td>NS'</td>
</tr>
<tr>
<td>III-IV</td>
<td>2 (33%)</td>
<td>4 (67%)</td>
<td></td>
</tr>
<tr>
<td>Tumor differentiation'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Well-moderate</td>
<td>0</td>
<td>3 (38%)</td>
<td>0.12'</td>
</tr>
<tr>
<td>Poor</td>
<td>7 (100%)</td>
<td>5 (63%)</td>
<td>NS'</td>
</tr>
<tr>
<td>Nodal metastasis at presentation, n (%)</td>
<td>3 (38%)</td>
<td>5 (63%)</td>
<td>NS'</td>
</tr>
<tr>
<td>Median follow-up, months</td>
<td>36</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>Locoregional recurrence</td>
<td>4 (50%)</td>
<td>3 (38%)</td>
<td>NS'</td>
</tr>
<tr>
<td>Distant failure, n (%)</td>
<td>0</td>
<td>1 (13%)</td>
<td>NS'</td>
</tr>
<tr>
<td>Death attributable to cancer</td>
<td>3 (38%)</td>
<td>0</td>
<td>NS'</td>
</tr>
</tbody>
</table>
| 2-year cause-specific survival'                              | 56%          | 100%         | 0.057'
| 2-year overall survival'                                    | 22%          | 63%          | NS' |

'Fisher's exact test.
'Mann-Whitney U test.
'X' test.
'For new primary tumors only (n = 12); TNM, Tumor-Node-Metastasis.
'Grade information was not available for one case.
'Log-rank test for the comparison of Kaplan-Meier survival curves.
phorylation, the leading source of in vivo reactive oxygen species. Xanthine oxidase activity has been implicated in the production of free radicals by bioactivation of ethanol to acetaldehyde, a process with possible tumor-promoting effects in conjunction with heavy alcohol consumption.[31] Other metabolic enzyme genes demonstrating an increase in expression include α (1,2)-fucosyltransferase and dihydrofolate reductase. Increased activity of dihydrofolate reductase has been shown previously to be independently involved in methotrexate resistance in HNSCC cells (32). α (1,2)-fucosyltransferase catalyzes a key step in the synthesis of histoblood group antigens and is expressed in colon adenocarcinoma.[33] In oral SCC, the high expression of sialyl Lewis a antigen in primary tumors is believed to be involved in nodal metastasis and predicts a poor prognosis.[34]

The control of epithelial cell proliferation, differentiation, and apoptosis requires a balance between signaling and transcriptional activation/repression. Transforming Growth Factor-β (TGF-β) is expressed in the majority of supraglottic SCC tumors and may be part of an immune-related process to aid in tumor progression.[35] TGF-β levels could also be an important predictor for survival in patients with carcinoma of the cervix after radiation therapy and in patients with adenocarcinoma of the lung.[36, 37] We have identified several genes whose expression is known to be induced by TGF-β: (a) TGF-β inducible early response 2; and (b) TGF-β activated kinase 1b. TGF-β inducible early response 2 is part of the family of TGF-β inducible Sp1-like zinc finger transcription factors and has been shown to be involved in the regulation of cell growth and differentiation by acting as a transcriptional repressor.[38] TGF-β activated kinase 1b is another signaling molecule acting downstream of TGF-β. It functions as a mitogen-activated kinase kinase kinase in the c-Jun N-terminal kinase/stress activated protein kinase, the p38 mitogen-activated protein kinase cascades, and the NF-κB activation pathway.[39]

Additional Group I-specific genes implicate other signaling pathways in HNSCC. These include cytokine subfamily members, such as SCYA16, which is inducible by Interleukin-10 (IL-10).[40] IL-10 has been identified as a mediator of immunosuppressive effects in HNSCC.[41] In non-small cell lung carcinoma, production of cytoplasmic
Fig. 3. List of selected Group I-specific genes. Tumors are designated as Group I or Group II as indicated at the top of the figure. Genes whose expression is higher in the tumor sample relative to the reference sample are shown in red; those whose expression is lower than the reference sample are shown in green. Accession numbers and gene names for all of the data are shown at the right.
IL-10 correlates significantly with poor prognosis.\cite{42} We also observed two phosphodiesterase (PDE) genes expressed in Group I tumors: (a) cyclic AMP-specific PDE; and (b) cGMP-specific PDE 6A. The decreased cellular-mediated immune response of patients with HNSCC has been linked to an increase in the specific activity (Vmax) and affinity (Km) of lymphocyte cyclic AMP-PDE.\cite{43} The product of the ephrin B1 gene we observe in Group I has also not been identified previously in HNSCC. It belongs to the largest subfamily of receptor protein tyrosine kinases, and evidence is now emerging which links expression of ephrins and their receptors in tumor tissue and the observed degree of malignancy, most likely caused by factors such as promotion of angiogenesis and enhanced cellular motility.\cite{44}

In conclusion, we have shown that the pattern of global gene expression in a HNSCC clinical specimen can be used as a predictor of prognosis. We isolated subsets of genes showing the greatest patterns of divergence in gene expression. Although building a complete story of a signaling or metabolic pathway is not possible based on this analysis, these results highlight pathways that merit exploration for possible links to outcome in HNSCC. A subset of these genes will be studied using supplemental methodologies, such as tissue microarrays, to examine their suitability as diagnostic markers and as possible targets for drug treatment.

References
Gene Expression Profiling in Head and Neck Cancer


18. Bouras T, Southey MC, Venter DJ. Overexpression of the steroid receptor coactivator AIB1 in breast cancer correlates with the absence of estrogen and progesterone
30. Nordsmark M, Overgaard J. A confirmatory prognostic study on oxygenation status and loco-regional control in advanced head and neck squamous cell carcinoma treat-
Gen Expression Profiling in Head and Neck Cancer


