Abstract: Advances in gene expression analyses have allowed global assessment of expressed genes in clinical samples. Gene expression profiles derived from clinical specimens have been used to distinguish differences in tumors that are not obvious by clinical, radiographic, or histologic characteristics. Despite its common histology and presentation, head and neck squamous cell carcinoma (HNSCC) is associated with widely varying clinical behavior and response to therapy. Currently, clinicians have a dearth of tools to predict response to therapy or to identify patients at high risk of poor outcome. Recently comprehensive analyses of gene expression patterns of individual tumors have shown promise to improve discovery of biomarkers for 1) progression of premalignant lesions, 2) disease presence or absence, 3) prediction of clinical outcome, and 4) identification of targets for therapy. In this review, we will discuss advances, limitations and future directions of genomics as it applies to HNSCC.

Keywords: head and neck cancer; genomics; DNA microarray; biomarkers; predictive profiles

Squamous cell carcinoma (SCC) is the most common histology of cancers arising from the upper aerodigestive tract, comprising about 80% to 90% of all tumors of various subsites within the head and neck. Clinical behavior, such as metastatic rate and response to therapy, of head and neck squamous cell carcinoma (HNSCC) varies from subsite to subsite; however, even within a single subsite, individual tumor behavior varies greatly, with some tumors growing slowly and never metastasizing and other tumors growing rapidly or metastasizing early. Despite intrinsic differences in tumor behavior, all HNSCCs are treated similarly. Therapy of HNSCC is limited by the anatomic location of the tumors, which may involve vital structures such as the eye, brain, or those necessary for speech and swallowing. Surgical resections can cause severe cosmetic deformity in the face and neck that cannot be easily covered with clothing, leading to social stigma. Standard therapy for stage I/II tumors is surgical resection.
and/or radiation therapy, whereas treatment for advanced stage III/IV tumors requires a combination of chemotherapy, radiation therapy, and/or surgery. Despite relative uniformity of treatment, clinical outcome after curative therapy varies greatly with respect to locoregional recurrence, organ preservation rate, distant metastasis, survival, and treatment-related toxicity. Outcome after therapy of HNSCC cannot be accurately predicted by clinical, radiographic, molecular, or histologic characteristics.

The observed diversity of outcomes is an indication of the intrinsic heterogeneity in the biological properties of individual tumors. Differences in tumor behavior are likely a direct result of multiple genetic mutations and epigenetic changes within the cancer. These alterations in DNA directly affect gene expression and ultimately the amount and activity of protein produced. Because genes are transcribed into mRNA, measuring the abundance of mRNA within the tumor is reflective of genetic and epigenetic alterations that determine tumor behavior. Traditional approaches to identify genes predictive of biological and clinical tumor behavior have relied on multiple techniques designed to examine one or, at most, a handful of genes at a time. Comprehensive analyses of gene expression patterns of individual tumors can now be achieved with DNA microarray. By use of this approach, various clinical needs in HNSCC can be addressed, including discovery of biomarkers for (1) the progression of premalignant lesions, (2) disease presence or absence, (3) prediction of clinical outcome, and (4) identification of targets for therapy.

**DNA MICROARRAY TECHNOLOGIES**

Over the past several years, DNA microarrays have had a major impact on biomedical research and have emerged as a powerful tool for the parallel measurement of relative gene expression levels. The evolution of these technologies has been driven by the need for more comprehensive analytical approaches to use the enormous amount of genomic data and resources being acquired through the genome sequencing projects. DNA microarrays can be grouped into two general categories: (1) commercially available microarrays with defined content such as those produced by Affymetrix, Agilent, and a number of other manufacturers; and (2) microarrays produced with variable and customizable content, generally using either spotted cDNA or oligonucleotides. Efforts by the laboratory of Dr. Patrick Brown at Stanford University established an open-source modality in DNA microarrays that helped fuel the explosion of microarray technology, making it a predominant tool in genomic research. Custom microarrays generally involve samples of DNA with known sequence being spotted and immobilized onto a substrate, most commonly a glass microscope slide. Next, RNA isolated from samples of interest is reverse-transcribed into cDNA and labeled with one of two spectrally distinct fluorescent dyes. The use of dyes with distinct fluorescent characteristics allows the two labeled cDNA samples to be pooled and hybridized on a single microarray. Strands of cDNA in the pooled samples hybridize to their complementary sequence immobilized on the substrate, and any unbound cDNA is washed off (Figure 1A). The ratio between fluorescent signal intensities of the two dyes at a particular spot is representative of the relative abundance of the corresponding mRNA in the samples of interest. In some commercially available arrays such as Affymetrix, the target RNA is labeled with a single fluorophore and hybridized onto the immobilized DNA probes. Gene expression is determined by degree of fluorescence at each immobilized DNA probe (Figure 1B).

With either single or dual fluorophores, a moderate amount of RNA is needed for hybridization to the array. Frequently, with small or necrotic tumors, obtaining an adequate amount of RNA may be an issue. In this case, very small amounts of RNA (5–100 ng) can be linearly amplified and the amplified RNA (aRNA) subsequently used for hybridization. We favor this approach because it does not limit analyses to larger tumors. In addition, because such small amounts of RNA are needed for amplification, frozen sections are adequate for analyses. Because sections of the tumor serve as the starting material for gene expression analyses, histologic control of the amount of epithelia, stroma, inflammatory cells, and necrosis is possible. Some studies have favored the use of microdissection, so that only the epithelial component of the tumor is analyzed. Although this technique has distinct research advantages, we believe that for correlation of clinical parameters with gene expression, all parts of the tumor, including infiltrating stromal, vascular, and inflammatory cells, should be represented. Some data, including our own, suggest that these nonepithelial components of tumors may contribute significantly to the overall clinical behavior of the tumor.

Because of the large quantity of data generated by a single microarray experiment, gene
expression analyses require statistical tools to identify the genes of interest. There are two widely accepted methods of gene expression data analyses: (1) unsupervised and (2) supervised analyses. Unsupervised techniques examine the data based only on the gene expression pattern regardless of specific characteristics of the tissue being examined. These types of analyses can segregate different histologic tumor types or tumors from the normal tissue from which they arose, but more importantly these techniques allow identification of tumor subtypes that are not otherwise distinguishable by clinical, radiographic, or histologic characteristics. One example of unsupervised analysis is the molecular classification of HNSCC by hierarchical clustering based on an intrinsic gene set that identifies four major subtypes of HNSCC.

On the other hand, supervised analysis selects the genes that are associated with the “supervising” parameters or conditions such as the presence or absence of recurrence or metastasis. When using supervised techniques, the analysis is dependent on the supervising parameter to discriminate the groups or categories with the highest prediction accuracy. Once a predictive gene list is generated from a training set, the results are confirmed by leave-one-out cross validation and ultimately by analysis of an independent cohort of patient samples. A number of different methods are used in supervised approaches. We recently published an evaluation of several machine learning approaches. Expression profiling studies to predict the lymph node metastasis from primary HNSCC tumors uses the presence or absence of metastasis as the supervising parameter and specifically queries for genes that distinguish the two groups.

**CLINICAL APPLICATION OF MICROARRAY TECHNOLOGY**

**Prevention.** Progression of normal epithelia through premalignancy to HNSCC is a multistep process that has been associated with histologic characteristics of each stage. Each histologic stage has also been correlated with defined alterations in tumor suppressor genes, oncogenes, or both. Increased understanding of progression from normal to malignant lesions may provide mechanistic biological insight of use for prevention strategies. Using Affymetrix GeneChips containing >12,000 genes and sequence tags, Ha et al described a transcriptional progression model of the HNSCC. Malignant lesions (M), premalignant lesions (PM), matched normal mucosa (MN) from the patients with M or PM lesions, and normal mucosa from patients with noncancer diagnoses (N) were analyzed for gene expression. The data were analyzed using significance analysis of microarray (SAM), hierarchical clustering, and principal components analysis (PCA). Most transcriptional alterations...
were found to occur before the development of malignancy. Progression from normal to premalignant was associated with altered expression of 354 genes, whereas progression from premalignant to malignant was associated with altered expression of only 23 genes. These results suggest that there are greater transcriptional alterations associated with progression from normal to premalignant than with progression from premalignant to invasive carcinoma. The functions of the protein products of genes that were upregulated in malignant lesions compared with normal mucosa were mostly associated with angiogenesis, apoptosis, cell adhesion, cell cycle, cytokine modulation, transcriptional regulation, and cell-signaling pathways such as mitogen-activated protein (MAP) kinase and Wnt. Several of the identified genes including vascular endothelial growth factor (VEGF) and its receptors, platelet-derived growth factor receptor (PDGFR), epithelial growth factor receptor (EGFR), and transforming growth factor-β (TGF-β) are currently being evaluated as targets for anticancer drugs.

Gene expression analyses of histologically normal (N and MN), premalignant (PM), and malignant (M) tissue of the upper aerodigestive tract could accurately distinguish these histologically defined groups. Remarkably, the N and MN tissues clustered together and separate from the PM and M groups, which also clustered in two distinct branches. One PM sample clustered with N that was histologically a leukoplakia without atypia, whereas one PM sample clustered with M that was carcinoma in situ. On the basis of these preliminary results, a similar method could be further developed to identify PM lesions at high risk of progression to invasive carcinoma. Once identified, these high-risk PM lesions could be aggressively treated, whereas PM lesions identified as low risk could be treated conservatively. This type of analysis may complement our histologic diagnosis by further classifying PM lesions into those at high or low risk of progression.

**Prediction of Metastasis.** The most reliable predictor of recurrence and survival in HNSCC remains the presence or absence of neck metastases.8–11 In fact, the decreased survival and risk of recurrence portended by cervical lymphatic metastases is of such great concern that necks without clinical or radiographic signs of metastases (clinically and radiographically N0) are treated by surgery or radiation therapy if risk of cervical metastases is deemed to be greater than 20% on the basis of the size and site of the primary tumor. This treatment strategy ensures that patients at risk of metastases will be treated aggressively, but by definition mandates that up to 80% of patients at no risk of metastases are subjected to the expense and morbidity of neck therapy. Surgical treatment of the neck creates surgical scars and carries the risk of infection, damage to cranial nerves, and vascular structures, as well as the risk of fistulization if mucosal lesions are simultaneously resected. Radiation therapy for head and neck cancer is associated with lymphedema, xerostomia, and neck skin and muscle fibrosis. In addition, radiation therapy increases the long-term risk of stroke because of acceleration of arteriosclerosis in the carotid artery.12 Clearly, better markers of metastatic potential are needed so that patients at risk of metastases are aggressively treated and patients at low risk are spared morbidity.

Development of metastases after primary tumor formation depends on migration of tumor cells, availability of lymphatic or blood vessels, destruction of barriers to migration, development of growth, and antiapoptotic characteristics that allow for implantation, survival, and growth in a foreign environment without supporting cells or normal growth signals, and evasion of immune defenses. In all likelihood, the tumor that succeeds in metastasizing must alter expression of multiple genes/proteins from expected pathways such as angiogenesis, apoptosis, cytoskeletal regulation, proliferation, immortalization, immune responsiveness, and invasiveness. That multiple genes must act in concert to allow metastases provides a ready explanation as to why molecular analyses of a small number of genes has to date failed to prognosticate metastases in HNSCC, but also suggests that more powerful simultaneous examination of multiple molecular markers may succeed. In addition, unexpected biochemical pathways that contribute to metastases may be identified using more global approaches. The ability to accurately predict primary head and neck tumors with metastatic potential will appropriately guide aggressive therapy and spare the cost and morbidity of neck therapy to patients at low risk.

Traditionally, it has been assumed that metastases arise from rare cells within a primary tumor that have gained metastatic potential; however, gene expression profiling of 64 primary tumors and 12 metastatic adenocarcinomas by
Ramaswamy et al\textsuperscript{13} suggests that genetic programming of metastasis exists within most, if not all, of the primary tumor at the time of diagnosis. From these primary tumors, a “metastasis signature” defined by 17 unique genes was identified. When this metastasis signature was applied to an independent set of 279 primary solid tumors of various organ sites, patients whose tumor contained this signature had worse clinical outcome. These data suggest that heterogeneity within the primary tumor does not preclude the ability of gene expression analyses to predict metastases. Although still somewhat controversial, most data suggest that the bulk of the primary tumor carries a signature that predicts metastatic potential. It could be that metastases requires many genetic alterations, and despite the fact that only a small portion of the primary tumor will attain all necessary alterations, most of the tumor will contain cells with many of the metastatic features that can be detected by gene expression analyses.

To determine whether metastases could be predicted from gene expression within primary HNSCC, we determined gene expression within 60 primary and previously untreated HNSCCs. When we attempted to predict lymph node metastasis using clinical nodal staging as the supervising parameter, prediction accuracy of gene expression was poor (prediction accuracy rate of 53\%). When matching clinical lymph node staging (LN\textsuperscript{+} or LN\textsuperscript{−}) with pathologic staging, we determined that accuracy of clinical staging was poor (roughly 40\% false positive or false negative), suggesting that inaccuracies in the supervising parameter may have explained the poor predictive power of gene expression. When we used pathologic staging of nodal metastases as the supervising parameter, prediction accuracy of gene expression from the primary tumor improved significantly. Further improvement in predictive accuracy was obtained when the tumors were analyzed on the basis of their anatomic subsites (ie, oral cavity vs oropharynx, hypopharynx, and larynx).\textsuperscript{3} Cromer et al\textsuperscript{14} used gene expression to compare three groups of patients with hypopharyngeal SCC: (1) those who had metastasis develop, (2) those who did not have metastases develop, and (3) those who had local recurrence within 3 years of treatment. By use of 168 gene targets, metastatic prediction accuracy was 92\%,\textsuperscript{14} whereas the prediction accuracy in our study examining oropharyngeal, hypopharyngeal, and laryngeal SCC was 83\%.\textsuperscript{3} These results suggest that the metastatic prediction accuracy of gene expression analyses for HNSCC may be improved through analyses of subsites separately. Subsite analyses would incorporate anatomic differences of each site, which may result from different lymphatic drainage, etiology (eg, human papillomavirus [HPV] in oropharynx), or genetic differences.

Recently, Roepman et al\textsuperscript{6} also published an expression profile of lymph node metastases determined from 82 primary SCCs of the oral cavity and oropharynx. This profile contained 102 predictor genes with overall predictive accuracy of 86\% relative to clinical staging accuracy of 68\% in their data set. Interestingly, long-term storage of samples even without evidence of RNA degradation affected prediction accuracy. These data suggest that time from tissue procurement to gene expression analysis should be controlled or that expression data should be generated prospectively rather than retrospectively. We have compared the predictive genes discovered by Roepman et al using our data set. Although 45 of the 102 genes were present on our microarray, none of them overlapped with our predictive gene list. The non-concordance of predictive gene lists is common in many microarray studies using different platforms and data mining tools and may represent differences in experimental design or data analyses but also may represent true differences in biology based on different subsites or other unknown factors (ie, HPV infection).

One limitation of all studies to date is that results were based on relatively small sample size. Although microarray data were validated by various statistical methods to estimate the false-positive rate from large variables (ie, thousands of genes), as well as independent verification of gene expression by immunohistochemistry, quantitative real-time polymerase chain reaction (PCR), or quantitative PCR, only one study used an independent test set for validation of predictive gene sets.\textsuperscript{8} Before these types of predictors can be used as a clinical test, results must be confirmed and validated in large HNSCC patient cohorts possibly for each subsite of the head and neck. Validation is particularly critical, because there is definitive treatment for the lymph node metastases and because metastatic potential is such a strong prognostic indicator. Attaining the large number of patients needed for validation studies will require multicenter randomized clinical trials. One example recognizing this need is the large-scale study of oral cavity SCC currently accruing through the American College of Surgery Oncology Group (ACOSOG).
Molecular Classification of HNSCC for Recurrence and Survival Prediction. Clinical decisions regarding therapy of HNSCC rely on assessment of tumor size, location, and metastases. Despite aggressive treatment with curative intent, approximately 50% of patients with advanced HNSCC will have a recurrence and die of their disease. The current standard of care for both patients with advanced unresectable HNSCC and for patients who desire organ preservation is concurrent chemoradiation therapy. This approach has shown increased survival but is associated with severe toxicity. Because of the debilitating nature of aggressive therapy and the lack of adequate control of advanced HNSCC, better markers of clinical outcome are needed. Identification of low-risk patients would decrease individual patient morbidity, as well as decrease cost by avoiding unnecessary therapy. Simultaneously, identification of patients at high risk of poor outcome would target the appropriate patients for more aggressive treatment. Class prediction using gene expression analyses is a first step toward individualized therapy of HNSCC.

In our study, HNSCCs were molecularly classified into four clinically distinct groups based on unsupervised gene expression a form of unsupervised statistical analysis (Figure 2). The molecular classification was independent of the primary tumor site, histologic differentiation, and stage but had statistically significant differences in recurrence-free survival (p = .04). Other groups have analyzed gene expression data using recurrence or survival as a supervising parameter. Ginos et al analyzed 41 HNSCC tumors resected at surgery and compared them with normal oral mucosa to identify biomarkers of recurrence. They identified a gene set whose expression is associated with invasion, metastasis, and recurrence. These data suggest that recurrence, at least in part, is an intrinsic biological property of the primary tumor and that this recurrence signature can be detected before initiation of therapy.

Understanding Treatment Resistance and Identification of Novel Therapeutic Targets. Through the comprehensive analysis of tumors, novel pathways involved in treatment resistance and novel therapeutic targets can be identified. In our study, the subtype of HNSCC with the worst survival had higher expression of TGF-α and markers suggestive of activation of the EGFR pathway. These data suggest that tumors with activation of the EGFR pathway, either alone or in combination with other genetic defects, will have worse survival. EGFR activation has been associated with radiation resistance and decreased response to chemotherapy, possibly explaining the poor progression-free survival noted in this group. On the other hand, groups with the most favorable survival had higher expression of genes involved in detoxification, suggesting that these patients may detoxify carcinogens more efficiently than other patients. In other cancer types, gene expression analyses accurately predicted chemotherapy resistance and sensitivity.

Identification of patients with HNSCC responsive to standard chemotherapy will greatly advance patient care and can be accomplished with standard clinical trial design; however, the acceleration of targeted therapeutics for cancer presents special problems for trial design. Recently, the EGFR inhibitor gefitinib was found to be relatively ineffective for therapy of non-small cell lung cancer when the entire population of these patients was evaluated, but a subgroup of patients with EGFR mutations had remarkable responses. Learning from this experience, we may find that only a small subset of patients will benefit from each targeted therapy. If this is the case, traditional clinical trials of unselected patients may suggest no benefit, when in reality there is a large benefit, but only for a small percentage of patients. The challenge for trial design becomes how to identify patients appropriate for a targeted therapy. Gene expression analyses of patient groups treated for short periods of time may allow identification of expression profiles suggestive of response. Furthermore, gene expression analyses may identify deregulated signaling, apoptotic, angiogenic, or proliferative pathways, indicating that a specific novel targeted therapy may have activity. The ability to rationally select patients for trials of these new biologically active agents will enable a new paradigm of clinical trials where these agents are tested either alone or in combination with standard therapy on smaller groups of patients. Once the subset of patients with HNSCC responsive to a particular therapy is identified, clinical trials will be more effectively designed and implemented, resulting in a tremendous benefit to drug development efforts and to the overall care of patients with HNSCC.

CURRENT LIMITATIONS AND FUTURE POSSIBILITIES
Several limitations exist in using DNA microarray in the clinical setting. To date, global gene
expression analysis has required fresh frozen tissue that must be preserved, usually by freezing under liquid nitrogen, shortly after excision to prevent RNA degradation. Intrinsic chip-to-chip, hybridization, labeling, and other technical variables exist within the assay, which contribute to noise. Other limitations relate to differences in the timing of tissue procurement, tissue process-
ing, and heterogeneity of the tumor sample such as percentage necrosis, stroma, and inflammatory infiltrate. A significant bioinformatics overhead is associated with these studies and their clinical implementation. Methods used for statistical evaluation of the data, normalization, data filtering, and identification of important biomarkers are all areas of active research and development. These practical issues of microarray experiments were investigated by the members of the Toxicogenomics Research Consortium who identified the sources of error and data variability among seven different laboratories and across 12 micro-array platforms and determined that these variations could be minimized by use of a common platform and standard procedures.22

In addition, many statistical analyses are being developed to overcome the systematic bias from experimental variations. One example is the use of distance weighted discrimination (DWD) to overcome and correct for systematic bias introduced by technical variation at the array level.23 This allows chips from different institutions and experimental variations to be corrected for data comparison and integration. We recently reanalyzed 21 of the 60 HNSCCs originally analyzed by Agilent Human 1 cDNA microarrays, using Affymetrix Human U133 2.0Plus Gene-Chips. The bias related to the differences in array platforms was corrected using DWD, and the arrays were clustered using the intrinsic gene set generated from the previous study for the molecular classification. Nineteen of the 21 pairs classified in the same group, suggesting that experimental variation can be corrected by statistical data manipulation. Statistical tools such as these will expand analyses of data from various institutions and platforms and allow integration and mining of distinct data sets.

Furthermore, because of the large number of variables, small sample sizes, and lack of independent validation sets in currently published microarray data, there has been a concern that the prediction of clinical outcome is reflective of experimental noise or chance alone. When Michiels et al tested the seven largest published microarray studies that predict clinical outcome, five of the seven studies did not classify patients better than random chance and concluded that the prognostic value based on the gene expression needs to be confirmed by repeated random sampling.24,25 The overestimation of prediction accuracy is a serious concern; however, careful experimental design with strict eligibility criteria in the patient population, clearly defined outcomes, and increased sample size with independent validation of the data can overcome these difficulties in statistical analyses.

CONCLUSION
Gene expression profiling of human cancers has great promise in both research and clinical care. Global expression analyses may expand understanding of complex, but clinically relevant, tumor behavior such as metastases, recurrence, resistance to therapy, and survival. In addition, gene expression data may help to predict response to new targeted therapeutic agents. Clinical application of gene expression data has begun with pilot studies accruing patients in Europe. The Breast International Group (BIG), a European consortium of 40 partners in 21 countries, has initiated prospective randomized controlled trial (TransBIG project) in which the clinical treatment of breast cancer is in part determined by microarray gene expression profiles.26 This trial will play a large role in determining whether gene expression profiles are more effective than current clinical and pathologic prognostic criteria. Adopting this type of trial to head and neck cancer will require a multi-institutional effort not only to confirm and validate currently existing microarray data but also for clinical trial design and accrual that are sufficiently powered to evaluate the clinical safety and efficacy. Although a daunting task lies ahead, with concerted effort we will be able to integrate the knowledge gained through expression analyses with clinical practice to provide substantial and measurable benefits to patients.

REFERENCES


