Oncogenomics/Proteomics of Head and Neck Cancers

4

Jason I. Kass, Howard S. Moskowitz, and Jennifer R. Grandis

Abstract

Over the past 20 years, we have seen a rapid development in our understanding of the molecular mechanisms of head and neck squamous cell cancer (HNSCC). These discoveries are in large part the result of advances in technology. Here, we first review the major technologies in both oncogenomics and proteomics as they apply to HNSCC and then review the results of seminal studies. These data lay the foundation for the next decade of discovery in HNSCC and are relevant to the clinician and scientist.

Keywords

Oncogene • Proteomics • Sequencing technology • Head and neck cancer • The Cancer Genome Atlas (TCGA) • Tissue microarray (TMA) • RPPA • Mass spectroscopy • MALDI • SELDI • Exome

4.1 Introduction

Head and neck squamous cell carcinoma (HNSCC) is the most common histology of cancers arising from the upper aerodigestive tract, comprising approximately 90 % of all tumors in this region. HNSCC encompasses a variety of anatomic subsites. Despite possessing similar histologic characteristics, the clinical behavior, including metastatic rate and response to therapy, varies between subsites and even within an individual subsite, indicating biologic heterogeneity in the setting of common histology. Current treatment strategies rely on traditional, clinical, radiologic, and histopathologic parameters to determine the stage of disease using the T (tumor), N (node), and M (metastasis) classification system. This system allows for estimation of disease burden, which is presumed to predict clinical outcomes and assist the clinician in making the most appropriate decision for patient management. However, the biologic heterogeneity of HNSCC is reflected by the dysregulation of multiple pathways including cellular differentiation, angiogenesis, and apoptosis. Apparently, identical histologic tumors may have similar phenotypic characteristics but develop through dysregulation of different pathways and can have different clinical courses.

Despite their intrinsic differences, all HNSCCs are treated similarly. Standard therapy for stage I/II tumors is surgical resection and/or radiation therapy. By contrast, treatment for advanced stage III/IV tumor requires the combination of chemotherapy, radiation therapy, and/or surgery. Given this relatively uniform treatment, clinical outcome after curative therapy varies greatly. The advent of new surgical techniques, radiation therapy, and chemotherapy have improved local control and overall quality of life, but survival rates for head and neck cancer have not increased significantly. It is likely that the diversity in outcome reflects intrinsic heterogeneity in the molecular components of individual tumors.

Clinical outcome is not accurately predicted by clinical, radiographic, or histologic characteristics. A limited number

J.I. Kass, MD, PhD
Department of Otolaryngology – Head and Neck Surgery, Boston University School of Medicine, Veterans Administration Boston Health Care System, Boston, MA, USA

H.S. Moskowitz, MD, PhD
Department of Otorhinolaryngology - Head and Neck Surgery, Montefiore Medical Center Albert Einstein College of Medicine Bronx, New York, NY USA

J.R. Grandis, MD (✉)
University of California, San Francisco, Box 0558, 550 16th Street, 6th Floor, San Francisco, CA 94143, USA
e-mail: jennifer.grandis@ucsf.edu

J. Bernier (ed.), *Head and Neck Cancer*, DOI 10.1007/978-3-319-27601-4_4

of histologic features such as perineural, perivascular, or nodal extracapsular spread are associated with increased tumor aggressiveness and may influence management decisions. Unfortunately, currently recognized individual markers associated with tumor development generally lack sensitivity or specificity, and there is currently no single molecular marker that is used for patient management in HNSCC. Human papillomavirus (HPV) has emerged as a viral-mediated driver of oropharyngeal HNSCC. Although patients with HPV-positive HNSCC generally have better survival outcomes compared with individuals with HPV-negative disease, HPV status is not a part of current evidence-based NCCN (National Comprehensive Cancer Network) guidelines. Ongoing clinical trials include both surgical and nonsurgical phase II studies that are determining eligibility based, in part, on HPV (or its surrogate p16) tumor status.

Given the heterogeneity of genetic alterations found in these tumors, a greater understanding of the molecular basis of the biochemical pathways involved in carcinogenesis potentially can facilitate diagnosis, drug discovery, and therapy for affected patients. These molecular changes involve interacting networks that operate at the transcriptional, translational, and posttranslational levels. Traditional approaches have generally not been useful due to the complexity of interactions, the difficulty of finding the proper combinations of genes and proteins to investigate, and the reliance on techniques that examine one or only several genes or proteins at a time.

The application of novel unbiased discovery technologies offers the opportunity for comprehensive and systematic molecular analysis to capture the complex cascade of events underpinning the clinical behavior of tumors. Tumors are believed to harbor molecular signatures that can be identified through the combined application of high-throughput profiling techniques and sophisticated bioinformatics tools for complex data analysis and pattern recognition. The main underlying goal is the identification of new targets that may provide insights into the underlying mechanisms of cancer biology, which in turn can potentially lead to novel approaches to cancer diagnosis, prediction of clinical outcomes, and development of new therapeutic strategies.

4.2 Oncogenomic Technologies

Cancer can be simplistically thought of as the overexpression of oncogenes and/or the silencing of tumor suppressor genes. However, in most cancers, including HNSCC, cancer development and progression is likely due to numerous genetic alterations involving a variety of different pathways. Although common alterations underlie many types of cancer, an individual cancer often develops due to an accumulation of specific mutations in DNA. Since these mutations accumulate randomly, different combinations of mutations exist between different individuals with the same type of cancer. Cytogenetic analysis of cells has evolved from the gross visual analysis of chromosomes to a detailed study of the regions of chromosomal gain, loss, and translocation. Techniques used include comparative genomic hybridization (CGH) where normal and tumor DNA is labeled and hybridized to normal metaphase chromosomes and the fluorescence pattern is then analyzed for increased or decreased intensity, representing copy number differences between genomes. Similarly, fluorescent in situ hybridization (FISH) utilizes labeled sequence specific probes, allowing for the detection of particular genes of interest as well as visualization of copy number per cell.

More localized and specific analysis has been made possible through the advent of high-throughput DNA-sequencing facilities as well as novel approaches to examine genomic variability. Single nucleotide polymorphisms (SNPs) are areas in the genome with an altered DNA sequence that may represent markers for disease predisposition or may be used to genetically identify patients. Microsatellites are tandem nucleotide repeats that are generally located in noncoding areas of the genome. They can have variable length and have been mapped to specific chromosomal regions, allowing for detection of adjacent genes of interest. In addition, microRNAs are a noncoding family of genes involved in posttranscriptional gene regulation that are associated with cell proliferation, cell differentiation, cell death, and carcinogenesis. Each of these can be investigated through the use of array technologies.

Another commonly utilized platform for oncogenomic analysis is DNA microarray technology, which offers the capacity for parallel measurement of relative gene expression levels (Fig. 4.1). These technologies are based on the selective mRNA or cDNA hybridization to DNA probes on the array surface. There are two general categories of microarrays, commercially available microarrays with defined content or microarrays produced with variable and customizable content. Microarray technology involves DNA sequence hybridization onto microscopic surfaces, which can be read by a laser able to detect the signal of minute fluorophores. These studies can incorporate nearly the entire known genome in a single experiment.

Advances in DNA-sequencing technology now allow for large-scale whole-genome sequencing with high fidelity and low cost in a timely fashion. Collectively referred to as next-generation sequencing, these technologies can sequence upward of three billion bases in a single run [1]. There are currently over ten different strategies being applied to whole-genome sequencing. They employ technologies that vary from amplifying DNA fragments inside water droplets immersed in oil to the detection of electric currents created from the chemical reaction during DNA synthesis. One of

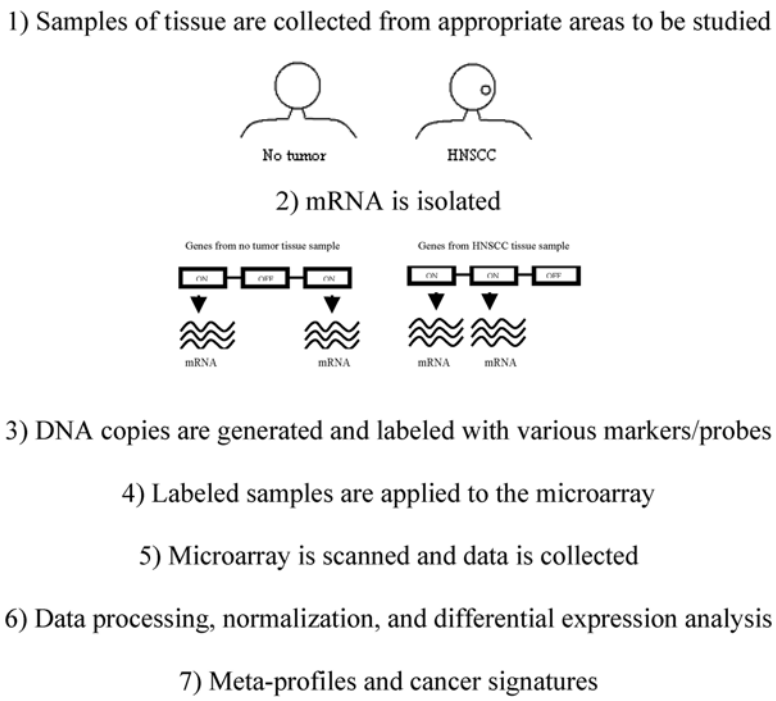

Fig. 4.1 Algorithm for using DNA microarray analysis to identify altered expression levels in HNSCC. After careful selection of patients, tissue samples are collected from study participants and mRNA is isolated. The mRNA represents the expression profile of the isolated cells as only active genes will produce mRNA. Microarray data from various tissues can be compared to generate differential expression patterns reflective of variations in gene expression between subjects. This data can be combined to define cancer signatures reflective of specific steps in tumorigenesis

the fastest strategies includes engineered polymerases with reversible fluorescent nucleotides that can rapidly terminate and restart DNA synthesis. This provides nucleotide readouts of over one million nucleotides/second. Accuracy varies between 98 % and 99.9 % depending on the sequencing strategy. Additionally, by looking at only the coding regions of a genome, the exome, one can rapidly identify all the expressed mutations in an individual tumor.

Each of these technologies generates large amounts of data from a single sample, particularly from tumor lysates or serum. Bioinformatics technologies enable the statistical analysis of the data and generate prediction algorithms to shortcut the experimental process. These data can be examined via unsupervised analysis using data based only on gene expression patterns regardless of the specific characteristics of the tissue being examined. This approach offers the potential to segregate different tumor types and allows identification of tumor subtypes that are not distinguishable by clinical, radiologic, or histologic characteristics. By contrast, supervised approaches select genes with parameters or conditions, and the analysis is dependent on the supervising parameter to discriminate the groups or categories with highest prediction accuracy. A predictive gene list is generated from a training set and the results are then confirmed by cross validation and analysis by an independent cohort of patient samples. Importantly for many cancers, including HNSCC, the molecular data has been collated and organized into a readily available online database that can be accessed by researchers worldwide (Table 4.1). The Cancer Genome Atlas (TCGA) provided by the National Cancer Institute has sequenced 528 head and neck tumors to date and represents a valuable translational research resource [8].

4.3 Proteomic Technologies

Proteome analysis is complementary to DNA microarray and sequencing technologies. Some techniques of proteomic analysis are widely used and clinically applicable such as enzyme-linked immunosorbent assay and immunohistochemistry, while others are used primarily as research tools such as immunoblotting and immunoprecipitation. Most of these techniques are limited to the study of only one or a few proteins at a given time. More comprehensive screening is permitted through 2D gel electrophoresis (2-DE). 2-DE is the method with the highest resolution for separation of protein mixtures and is believed to be superior for pattern analysis of complex samples. However, 2-DE may be difficult to use with certain proteins such as membrane proteins and basic proteins and has limited resolution of proteins in the low molecular weight spectrum. 2-DE separates proteins according to isoelectric points (isoelectric focusing) followed by separation according to molecular mass (SDS-PAGE). Peptide mass fingerprinting permits in-gel digestion of the protein spot of interest with a specific enzyme and resulting peptides are extracted from the gel and molecular weights of these peptides are measured. Alternatively, the peptides can be fragmented in a mass spectrometer yielding partial amino acid sequences from the peptides, which act as sequence tags.

Table 4.1 Publically available HNSCC microarray gene expression datasets

Authors	Tissue samples	Platform
Cohen et al. [2]	10 primary HNSCC	Affymetrix HG-U133 Plus 2
Rickman et al. [3]	186 primary HNSCC	Affymetrix HG-U133 Plus 2
Thurlow et al. [4]	71 primary HNSCC, 14 normal oral epithelium	Affymetrix HG-U133 Plus 2
Chung et al. [5]	40 primary HNSCC from 29 patients	Affymetrix X3P
Chung et al. [6]	55 primary HNSCC, 5 recurrent	Agilent Human 1
Walter et al. [7]	138 tumors from larynx, oral cavity, oropharynx and hypopharynx	Agilent 44 K microarray
TCGA [8]	528 primary HNSCC from 508 patients (as of 11/2014)	>20 platforms (https://tcga-data.nci.nih.gov/tcga/tcgaPlatformDesign.jsp)
Toruner et al. [9]	16 primary oral cavity SCC and 4 adjacent normal tissue from 16 patients	Affymetrix HG-U133A
Ye et al. [10]	26 primary oral cavity SCC and 12 adjacent normal tissues from 26 patients	Affymetrix HG-U133A
Kuriakose et al. [11]	22 primary HNSCC and 22 adjacent normal tissues from 22 patients	Human Genome U95A (Affymetrix)
Sticht et al. [12]	35 primary oral cavity SCC from 35 patients and 6 normal oral tissue from normal controls	Human Oligo Set 4.0 (Operon)
Pyeon et al. [13]	42 primary HNSCC from 42 patients and 14 normal oral tissue from normal controls	Affymetrix HG-U133A

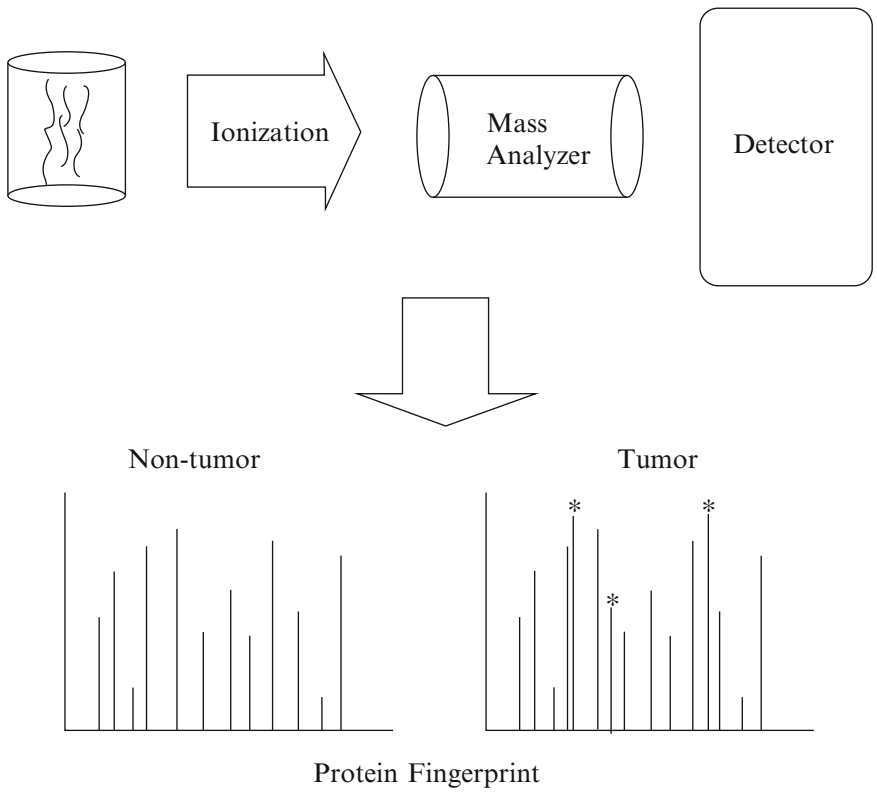

Fig. 4.2 Mass spectrometry approaches to biomarker analysis. Analysis begins with a protein or peptide mixture that is processed to maximize the number of detected differentially expressed proteins. The sample is subsequently ionized by a variety of instruments such as a laser and separated by a mass analyzer (time-of-flight or ion trap) based on mass and charge. The resulting spectra are representative of the ionized proteins within the initial sample. Bioinformatics approaches are then utilized to compare the spectra to identify unique and differing protein components (*asterisk* indicates differentially expressed *m*/*z* species)

Fundamentally important to recent advances in proteomics have been improvements in the speed, accuracy, and sensitivity of mass spectrometry (MS) instruments for the analysis of complex protein mixtures or tissues (Fig. 4.2). MS analyzes proteins or peptides as ions, which can be distinguished based on mass to charge ratio (*m*/*z*). Basic components of the instrument are the ion source that volatilizes and ionizes the proteins, the mass analyzer which separates proteins based on *m*/*z* values, and the detector which detects the sample after separation. The two most commonly used MS approaches are matrix-assisted laser desorption ionization (MALDI) and surface-enhanced laser desorption ionization (SELDI). These high-throughput methodologies have the ability to observe large numbers of protein events. Furthermore, as compared to 2-DE, they permit improved speed, high-throughput capability,

lower amounts of protein sample, effective resolution of low mass proteins, and direct application to assay development. Furthermore, sample loading and processing can be fully automated.

MALDI is commonly used for bioanalysis and employs laser energy to ionize and volatize proteins. A matrix such as a UV-absorbing organic acid is mixed with the sample to absorb laser energy and transfer it to the proteins to generate ions, which are then transferred to the mass analyzer. Ionization is not uniform and depends on relative protein abundance and intrinsic chemical characteristics. MALDI is generally coupled with a time-of-flight (TOF) mass analyzer, which separates proteins based on time to traverse a flight tube and strike a detector. MALDI-TOF-MS is a particle-counting method that relies on molar abundance. It requires minimal sample preparation, can distinguish hundreds to thousands of proteins from a complex mixture, and can detect subtle protein modifications. However, MALDI has a limited mass range and limited sensitivity for low-abundance proteins, and proteins with extremely high concentration can interfere with detection of proteins with similar m/z ratios.

SELDI utilizes a surface to capture and partially purify proteins from a complex sample based on physical and biochemical properties and is dependent on protein conformational stability for reliable detection. A variety of coated surfaces are presently available that bind proteins based on hydrophobicity, anionic or cationic charge, or binding to metals. SELDI also partially purifies the protein sample, making it less complex than the similar unfractionated sample for MALDI. This partial purification may lose critical proteins, but theoretically generates fewer problems with highly abundant proteins. When the process is expanded to many hundreds of samples, population-specific protein expression profiles can be deduced that are characteristic of the assayed group. However, the identified mass spectrum does not enable protein identification and none of the interactions are specific.

Reverse phase protein array (RPPA) is another high-throughput platform for marker screening. RPPA utilizes lysed histopathologically relevant pure cell populations. The lysate is immobilized in an array configuration via a pin-based microarray onto nitrocellulose slides with each spot containing the whole cellular protein contents. Each slide is then probed with an antibody that can be detected by a variety of assays. Protein samples are arrayed in miniature dilution curves to ensure that the analyte of interest remains in the linear range of detection. A subset of HNSCC TCGA samples have been analyzed by RPPA and the proteomic information is publically available.

Tissue microarray (TMA) technology applies advanced array-based approaches to data gathering with standardized medical pathology laboratory practices. A TMA block is loaded with freshly sectioned core biopsies from paraffin-embedded tissues derived from cohorts of cancer patients on a single slide. Automated digital image capture is followed by pathologist scoring of the image. Further evolution in the analysis of stained TMA sections involves automated scoring of staining intensities and features on TMA slides using image analysis software. TMA provides the capability to perform rapid analysis of comprehensive panels of normal and disease specimens. TMA allows visualization of molecular targets in up to thousands of tissue specimens at a time and reveals cellular localization, prevalence, and clinical significance of candidate genes and gene products. However, TMA is limited by the availability of antisera, only provides a semiquantitative estimation of protein levels, and may miss important histologic areas due to the small size of the core biopsies utilized in these arrays.

4.4 Oncogenomics of HNSCC

4.4.1 Genomic Changes Underlying Malignant Transformation

Cancer develops from the accumulation of various genetic alterations. DNA microarrays and whole-exome sequencing have emerged as powerful tools for the parallel measurement of relative gene expression levels in HNSCC (Table 4.1). The usage of DNA microarrays and genome sequencing to generate clinically relevant molecular signatures has grown in its acceptance. Early studies showed the heterogeneous nature of HNSCC tumors at the molecular level. However, direct comparison between studies has often proved difficult due to the variety of gene expression arrays, platforms, and data analysis algorithms used.

HNSCC cell line studies have provided initial insights into the genetic variations that may underlie the cancer phenotype using these preclinical models. Cell lines offer relative homogeneity of samples for investigation but may suffer from artifacts of immortalization and passage in vitro compared with human tumors. One microarray study analyzed 25 HNSCC cell lines and one immortalized human oral keratinocyte cell line and found wide alteration in the gene expression in cell cycle regulation, oncogenesis, cell proliferation, differentiation, and apoptosis [14]. This study revealed two distinctive subtypes of gene expression patterns, but these patterns did not seem to correlate with the clinical staging or differentiation grade of the original tumors. Another study used SNP array-based loss of heterozygosity (LOH) profiling on whole-genome loss of 41 HNSCC cell lines and found several frequent LOH regions [15]. This report identified a region on chromosome 8 that exhibited the most frequent LOH (87.9 %) and found that the mitochondrial tumor suppressor gene 1, a candidate tumor suppressor gene residing in this area, was consistently down-

regulated in expression, suggesting that it may be a tumor suppressor in HNSCC.

Another report utilized genome-wide comparative genomic hybridization and expression microarray analyses to reveal known and novel amplicons that showed concomitant increase of copy number and expression of target genes for both laryngeal SCC cell lines and primary tumors [16]. They found that the overexpression of 739 genes could be attributed to gene copy number alteration in cell lines, of which 325 genes showed the same phenomenon in primary tumors. Subsequently, this group analyzed oral tongue SCC cell lines and found that these cell lines exhibited similar genomic alterations as had been previously found in their laryngeal SCC cell lines despite the differences in clinicopathologic features between these anatomic subsites [17]. A wide variety of genes were found to be altered including deletions of known tumor suppressor genes including *FHIT*, *CSMD1*, and *CDKN2A*.

Other studies have attempted to provide a framework for improving our understanding of the molecular events underpinning various aspects of these tumors. The progression of normal epithelia through premalignancy to HNSCC is a multistep process that has been associated with distinct histologic characteristics at each stage. An early study analyzed invasive SCC lesions from the oropharynx and oral cavity, and using hierarchical clustering analysis, they were able to show that oral SCC was distinguishable from normal oral tissue, but there was heterogeneity among the tumors even of a particular histopathologic grade and stage [18]. This study identified 239 genes that were overexpressed and 75 genes that were downregulated, but could not find statistically significant differences in gene expression between metastatic and nonmetastatic tumors. Later, another group established a transcriptional progression model of HNSCC in the progression from normal mucosa to dysplastic epithelium to invasive HNSCC [14]. Matched samples were analyzed using gene expression arrays, significance analysis of microarrays, hierarchical clustering, and principal components analysis to identify genes with differential expression patterns between the tissue groups. The progression from normal to premalignant was associated with altered expression of 334 genes (108 upregulated and 226 downregulated), while the progression of premalignant to malignant was only associated with altered expression of 18 genes (5 upregulated and 13 downregulated). This transcriptional model suggested that the majority of alterations occurred before the development of invasive cancer.

An alternative strategy was used in another study employing forward and stepwise logistic regression analyses to identify potential biomarkers for the early detection of oral SCC by comparing gene expression of primary oral SCC, oral dysplasia, and clinically normal oral tissue [15]. They identified combinations of genes, which differentiated oral SCC from controls that included laminin-gamma 2 chain, collagen type IV alpha 1 chain, collagen type I alpha 1 chain, and peptidyl arginine deiminase type 1. Another group analyzed 41 HNSCC tumors from various anatomic sites and compared them with normal oral mucosa with gene expression arrays [16]. They used statistical and data-filtering criteria to identify 2890 genes differentially expressed between the two groups and revealed functional gene expression signatures that were highly represented in HNSCC including those involved in inflammatory response, epidermal differentiation, cell adhesion, and extracellular matrix functions. They suggested that the disease signature is an intrinsic feature of a HNSCC and may function as a predictor of early local treatment failure.

Several studies have attempted to build on the growing lists of putative biomarkers by generating gene sets, which may be able to lead to useful predictions regarding the propensity for a given lesion to be or develop into a cancerous lesion. One study matched tumor and normal specimens from the oral cavity and analyzed microarray gene expression data with a supervised learning algorithm [17]. This study generated a 25-gene signature that could classify normal and tumor specimen that was highly accurate on independent validation test sets but failed to predict non-oral tumors. Many of the genes in the predictor set had been previously implicated in oral SCC. The predictor set comprised several epithelial marker genes that had categories of potential interest including extracellular matrix components and cell adhesion molecules. Similarly, a different group attempted to generate a classifier set for oral SCC and leuokoplakias and found differential expression of 118 marker gene candidates by complementary DNA microarray [18]. Further evaluation demonstrated an 11-gene predictor set that could distinguish the two groups with greater than 97 % accuracy.

Most recently, there have been collaborative efforts to sequence the exome of HNSCC tumors. In 2011, two groups reported on genome data from 125 tumors [19, 20]. The whole-exome sequencing of 92 tumors provided a snapshot of the commonly mutated genes and signaling pathways for individual tumors [19]. Not surprisingly, this data validated many of the smaller sequencing efforts in the literature. The most commonly mutated genes were involved in cell death (TP53, PTEN, PIK3CA) and proliferation (CDKN2A, HRAS). This work revealed a previously unknown contribution from genes involved in terminal differentiation of squamous cells (NOTCH1, IRF6, TP63). Additionally, there were clear mechanistic differences between patients whose cancers were driven by tobacco and alcohol exposure as compared to HPV-positive tumors in the oropharynx. HPV-positive oropharyngeal tumors had two- to fourfold fewer mutations, did not have TP53 mutations, and were more likely to have PIK3CA-activating mutations. Interestingly, the impact of HPV on HNSCC outside the oropharynx does carry the same impact on mutation rate [21]. Overall, there were few activating mutations observed in the sequence data,

and the majority of mutations were tumor suppressor genes. The phosphoinositol 3-kinase (PI3K) pathway is one of the few pathways that have activating mutations. Inhibitors of PIK3CA are in phase I and II clinical trials.

4.4.2 Genomic Changes Underlying Metastases

Metastasis is the principal cause of death in patients suffering from cancer, but the underlying molecular mechanisms are poorly understood. It is widely believed that the accumulation of genetic damage leads to the expression of a malignant phenotype that precedes metastasis formation. In order for a tumor to metastasize, it has to gain a number of functions distinct from the primary tumor. These include the ability to adhere to and then traverse the basement membrane, pass through an extracellular matrix, enter and exit the blood stream, and finally invade a new microenvironment to replicate. To do so requires a number of molecular changes distinct from cellular division [22].

Several groups have investigated differences in gene expression between head and neck primary tumors that had or had not metastasized. In one analysis of tumors from the oral cavity and oropharynx, 101 genes demonstrated significant expression differences between the metastatic and nonmetastatic tumors [23]. These genes included a variety of cellular functions putatively associated with cancer behavior, and the gene with the greatest differential expression between the metastatic and nonmetastatic tumors was collagen type 11 alpha 1. A different study used microarray analysis to measure gene expression changes associated with tumor progression in patients with stage III or stage IV untreated oral SCC [24]. They identified 140 genes that consistently increased in expression during progression from normal tissue to invasive tumor to metastatic node as well as 94 genes that decreased in expression in a similar progression, which revealed a distinct pattern of gene expression during the progression from histologically normal tissue to primary carcinoma to nodal metastasis.

In another study, 82 primary tumors located in the oropharynx or oral cavity regions were analyzed using DNA microarray gene expression profiling [25]. This study established a set of 102 predictor genes for determining the presence of lymph node metastases. Many of the predictor genes they found were previously implicated in metastasis. The application of this gene set to a validation group gave an overall predictive accuracy of 86 % as compared with 68 % based solely on clinical diagnosis. A subsequent study implemented this dataset as a reference dataset and an independent gene expression dataset of metastasized and nonmetastasized HNSCC tumors as validation dataset [26]. They utilized supervised gene-based and pathway-based analysis to evaluate differences in gene expression to enhance the understanding of the biological context of the results. The identified gene sets were involved in extracellular matrix remodeling (including matrix metalloproteinases (MMPs) and their regulatory pathways) as well as hypoxia and angiogenesis.

Another group looked at 186 primary tumors and analyzed the samples with respect to whether the development of metastasis was the first recurrent event [3]. They collected transcriptome and array-comparative genomic hybridization data followed by non-supervised hierarchical clustering to distinguish tumors differing in pathological differentiation. They were able to identify associated functional changes and created a four-gene model (*PSMD10*, *HSD17B12*, *FLOT2*, and *KRT17*) which predicted metastatic status with 77 % success in a separate validation group, and the prediction was independent of clinical criteria. Similarly, another study revealed that gene expression patterns in 60 primary and previously untreated HNSCC allowed the tumors to be categorized into four distinct subtypes with statistically different recurrence-free survival [6]. Clinical nodal staging resulted in low prediction accuracy when used as the supervising parameter. However, supervised analyses using pathological staging to predict lymph node metastasis status improved the prediction accuracy of gene expression from the primary tumor, which was further improved by analysis based on anatomic subsites leading to a prediction accuracy of 83 %.

A large-scale gene expression analysis of the hypopharynx, a location associated with particularly aggressive behavior, found 119 genes that were highly differentially expressed between early and late tumors [27]. Furthermore, 164 differentially expressed genes were found that differentiated between relatively non-aggressive and aggressive tumors. Clustering of the associated probe sets defined the two groups of samples and correctly assigned 92 % of the tumors. In a separate study, genome-wide analysis was performed looking for LOH and allelic imbalance (AI) on specimens of tumor stroma and tumor epithelium isolated by laser capture microdissection on 122 patients with HNSCC and a history of smoking [28]. They found nearly twice as many areas of LOH/AI within the stroma as was found in the epithelium, more than 40 areas in total. Furthermore, they found three stroma-specific loci that were significantly associated with tumor size and cervical lymph node metastasis, highlighting the importance of examining stromal and epithelial elements and suggesting that stromal alterations play an important role in HNSCC behavior.

In the last 7–10 years, a new concept has been solidified regarding metastases of squamous cell carcinoma. This reflects recognition of dedifferentiation of squamous cells from an epithelial molecular profile to a more primal mesenchymal phenotype, normally present in embryonic development, but lost in mature tissues. This change in gene

expression has collectively been referred to as "epithelial to mesenchymal transition" (EMT). During EMT, many of the genetic changes documented above occur including a loss of cell attachment via changes in integrins (E-cadherin), activation of matrix metalloproteinases, and activation of genes involved with cell migration. In metastatic HNSCC, these tumors have altered patterns of expression, reflecting these changes with changes in EMT-related genes including *snail* and *twist* [29].

Finally, although head and neck cancer cells may become mesenchymal as they leave their primary site and metastasize, they switch back to an epithelial expression pattern once they arrive in a new distant site. It is currently unknown how this occurs, but recent evidence regarding a chloride ion channel may provide insight regarding this behavior. *TMEM16A (Ano1)*, a calcium-activated chloride channel, is frequently overexpressed in HNSCC and has been shown to behave as an oncogene [30]. When *TMEM16A* is inhibited in stable cell lines, these cells became more motile and were able to metastasize in a mouse model [31]. On the expression level, inhibition of the ion channel correlated with an expression pattern of a mesenchymal phenotype, while overexpression correlated with an epithelial phenotype. Dynamic changes in ion channel function may play a role in how cells transition between these phenotypes.

4.4.3 Genomic Changes Underlying Variable Responses to Treatment

Treatment protocols often involve the use of chemotherapy and/or radiation. Several recent studies have directed their attention toward the identification of genetic alterations that would give prognostic information regarding a given tumor's likelihood of response to various treatment protocols. Cetuximab, the *EGFR* antagonist, was initially developed as a potential radiosensitizer, when it was observed that tumors with high *EGFR* expression were radioresistant.

One study on HNSCC cell lines that exhibited relative radioresistance and radiosensitivity identified 167 genes that were significantly overexpressed in radioresistant cells, 25 of which included cancer-related genes involved in growth, proliferation, apoptosis, and adhesion [32]. Another study used significance analysis of microarrays for gene selection and a multivariate linear regression model for prediction of radiosensitivity [33]. They identified three novel genes whose expression values correlated with radiation sensitivity, and the overexpression of one of these genes, *RbAp48*, in a cancer cell line induced radiosensitization.

The use of tissue microarrays has also been used clinically to find genes that may help predict a response to therapy. Recently, 38 patients who received radiation were analyzed using a cDNA tissue microarray, and five candidate genes were identified (*VEGF*, *BCL-2*, *CLAUDIN-4*, *YAP-1*, and *c-MET*) as predictors for response to therapy. Protein expression of these five genes was then prospectively evaluated in 86 patients who underwent radiation. All five biomarkers were predictive of a poor response to therapy and two (*YAP-1* and *c-MET*) were synergistic [34].

In another study, 92 biopsies were obtained from untreated HNSCC patients prior to treatment with cisplatin-based chemoradiation for advanced HNSCC [35]. This group utilized supervised analyses to predict locoregional control and disease recurrence and found several gene sets that were enriched in recurrences. They utilized a signature established by Chung et al. [6] for HNSCC defining a high-risk group and found it to be predictive for locoregional control and disease-free survival in their dataset. A more targeted analysis utilized a cDNA array consisting of genes associated with angiogenesis and/or metastasis [36]. Seventeen genes were correlated with locoregional failure, of which *MDM2* and *erbB2* were found to be predictors of locoregional failure in their population of patients treated with CRT.

The biomarker *ERCC1*, a DNA repair enzyme, has been a strong biomarker for response to cisplatin. Cisplatin acts as an alkylating agent inducing lethal mutations in cell and preferentially targets dividing cells. Tumors that express high levels of *ERCC1* can repair these DNA mutations and are resistant to cisplatin. Patients whose tumors have low *ERCC1* expression are more susceptible to treatment. The usefulness of *ERCC1* to predict response to cisplatin has been documented prospectively in the treatment of non-small-cell lung cancer [37] as well as a recent randomized phase II clinical trial of HSNCC [38].

Other targeted chemotherapeutics are in development. *TP53*, the most widely mutated gene in HNSCC, leads to loss of apoptosis and oncogenesis, and strategies to restore *TP53* function could be promising in the treatment of HNSCC. Recently, a small molecule 17-(allylamino)-17-demethoxygelanamycin (17AAG) has been shown to restore p53 function and induce increased cell death in HNSCC cell lines [39]. It remains to be seen what effect it can have in animal and human models.

4.4.4 Genomic Changes Found in Surrogate Tissues

An evolving area of investigation involves the use of surrogate tissues in the investigation of HNSCC. Using saliva from patients with primary T1/T2 oral SCC with matched control patients in terms of age, gender, and smoking history, one group used microarrays to profile the human salivary transcriptome [40]. They found 1679 genes that were signifi-

cantly differentially expressed between the groups including seven cancer-related mRNA biomarkers that exhibited at least a 3.5-fold elevation in oral SCC saliva (*IL8*, *IL1B*, *DUSP1*, *HA3*, *OAZ1*, *S100P*, *SAT*). The combination of four of these biomarkers had a discriminatory power of 91 % sensitivity and specificity for oral cancer detection. A subsequent study compared the clinical accuracy of saliva with that of blood by using RNA biomarkers for oral cancer detection [41]. Using four serum mRNA markers, a sensitivity of 91 % and a specificity of 71 % were obtained for distinguishing oral cancer. However, the four salivary mRNA markers had a higher receiver operating characteristic curve value, demonstrating that for oral cancer detection, salivary transcriptome diagnostics may demonstrate a slight advantage as compared with serum.

MicroRNAs (miRNAs) are small noncoding molecules of RNA, often 20 nucleotides in length, that act at the posttranscriptional level to change the expression of key genes and have emerged as a mechanism for transcriptional control of tumors, including HNSCC [42, 43]. As of June 2013, there were over 1600 human miRs documented. Many are specific to squamous epithelium and associated with all aspects of cellular function including cell cycle regulation, apoptosis, cell proliferation, migration, and growth [44]. There are even miRNA expression profiles that are different between HPV-infected and non-infected cells. As with mRNA, miRs can be detected in saliva. A recent study of panel of three miRs, which are differentially expressed in HNSCC, was tested in 112 subjects (56 with HNSCC and 56 normal controls) [45]. These three miRs (miR-9, miR-134, and miR-191) were able to discriminate with good reliability HNSCC from normal controls. These miRs were also validated by TCGA miR data.

In the current era of HPV-associated oropharyngeal HNSCC, both plasma and saliva have been investigated as markers for response to treatment. HPV-16 DNA can be detected in both saliva and plasma samples. The presence of HPV-16 DNA in both saliva and plasma has been noted to be 100 % specific with a 100 % positive predictive value [46]. In this study however it had poor sensitivity (76 %) and very low negative predictive value (42 %). Interestingly, its presence after treatment was 91 % specific in predicting recurrence within 3 years.

4.4.5 Meta-analyses of HNSCC Microarray Studies

A cumulative analysis looked at studies incorporating DNA microarray analysis to examine genetic expression changes associated with the development of HNSCC [47]. Eighty-four genes were identified with common alterations in transcriptional expression across multiple studies. Many of these had been reported to be involved with HNSCC including MMPs, integrins, collagens, fibronectin, tenascin C, and cathepsin L, as well as many genes with less characterized roles in HNSCC. Only one gene, transglutaminase 3, was common to at least three of the reviewed studies. Overall, they found that genes encoding extracellular matrix and integral membrane proteins, cell adhesion molecules, and proteins involved in epidermal development and differentiation were most frequently identified in these studies. Furthermore, their results suggested a global downregulation of genes encoding ribosomal proteins and cholesterol biosynthesis enzymes and an upregulation of MMPs and inflammatory response genes.

Another study looked at 63 HNSCC transcriptomic studies in three categories of comparisons, premalignant vs. normal (Pre), primary tumors vs. normal (TvN), and metastatic or invasive vs. primary tumors (Meta) [48]. They used a systems biology approach via network-based meta-analysis and verified that 82 genes, 1260 genes, and 321 genes in the Pre, TvN, and Meta comparisons, respectively, were found reported at least twice. Overall, 1442 unique genes were reported at least twice in the studies that they analyzed. In terms of the direction of fold changes of the verified genes, the least contradiction was found in the TvN group and the most contradiction was found in the Pre group. Furthermore, they found that few genes overlapped between the Pre and Meta groups, although many genes overlapped between the other pairs of comparisons. Genes that were highly reported in prior studies across all three stages were *ECM1*, *EMP1*, *CXCL10*, and *POSTN*. Subsequently, they constructed knowledge-based networks, which revealed that integrin signaling and antigen presentation pathways were highly enriched in the dataset, and they found that chromosomal regions of 6p21, 19p13, and 19q13 had genomic alterations that were correlated with the nodal status of HNSCC.

There are currently 12 published gene expression datasets of HNSCC publically available with full clinical annotation [49]. Three of them were obtained using the same Affymetrix platform (U133 plus 2.0) and contain nearly 21,000 gene transcripts. These three datasets were recently used to generate a 172-gene profile to risk-stratify patients as either high or low for disease recurrence and then validated against six other datasets. This most recent gene profile compares well to the four other genetic signatures also generated by microarray data (radiosensitivity index, 13-gene SCCA signature, hypoxia metagene, and 42-gene high-risk signature). These genetic signatures are working their way toward a clinical-grade assay for detecting HSNCC and determining the severity of disease.

4.5 HNSCC Proteomics

4.5.1 Tumor Tissue Studies

High-throughput proteomic technologies have been utilized to detect biologically significant differences in protein expression of HNSCC in the same types of samples utilized in gene expression analysis. These studies have used a variety of techniques as outlined earlier in the chapter. One study utilized SELDI-TOF-MS to generate proteomic spectra and used the "Lasso algorithm" to extrapolate proteomic patterns that can best discriminate HNSCC patients from non-cancer controls which identified 65 significant data points to be used for discrimination [50]. Testing of these points yielded moderate sensitivity of 68 % and specificity of 73 % indicating that with further improvement and validation, it may be useful as a screening test for HNSCC in the future. More recently, another study analyzed 113 HNSCC, 73 healthy, 99 tumor-distant, and 18 samples of tumor-adjacent squamous mucosa by SELDI-TOF-MS [51]. They found 48 protein peaks differentially expressed between healthy mucosa and HNSCC. A supervised prediction analysis revealed greater than 90 % classification of healthy mucosa and tumor samples, and 72 % of the tumor-adjacent mucosa samples were predicted as aberrant, providing evidence for the existence of genetically altered fields with inconspicuous histology.

MALDI-TOF has also been successfully used in HNSCC proteomic studies. In one such investigation, MALDI-TOF was coupled with magnetic bead fractionation to analyze an HNSCC cohort consisting of matched pretreatment and 6–12 month posttreatment samples for analysis [52]. A set of approximately 200 spectral peaks was used and was able to largely correctly classify normal from pretreatment HNSCC samples, pretreatment from posttreatment, and normal from posttreatment samples. This showed the potential for use of this technology as a discovery platform in order to generate biomarker panels that potentially could be used for more accurate prediction of prognosis and treatment efficacies for HNSCC.

Another study used multidimensional LC-MS/MS to identify proteins that are differentially expressed in HNSCC for cancer biomarker discovery [53]. More than 811 proteins were identified which included structural proteins, signaling components, and transcription factors. They utilized a panel of the three best performing biomarkers, YWHAZ, stratifin, and S100-A7, to discriminate cancerous from noncancerous head and neck tissue. Their differential expression was verified by immunohistochemistry, immunoblotting, and RT-PCR and achieved a sensitivity of 92 % and specificity of 87 % in an independent set of HNSCC in discriminating tissue types. More recently, an analysis of samples from HNSCC patients with 2-DE and MALDI-TOF-MS revealed 181 proteins with differential expression between pretreatment and posttreatment samples [54]. Classification by disease status revealed significant differential expression of 16 proteins including several protease inhibitors and other molecules with direct implications on tumor survival. Another study attempted to validate DNA microarray results on a subset of genes that could potentially serve as biomarkers of oral SCC [55]. This group identified six potential biomarkers and used Q-RT-PCR to examine expression changes in oral SCC and normal control tissues, five of which were validated by this technique. TMA analysis then revealed that four of the six biomarkers (*SPARC*, *POSTN*, *TNC*, and *TGM3*) had differential expression and localization.

Biomarker clinical results from the EGOG 2303 phase II trial, where locally advanced stage III/IV resectable HNSCC was treated with induction chemo followed by CRT, were recently published [56]. Forty-two of the 63 patients had tissue samples available. A TMA was constructed and probed for the following proteins: EGFR, ERK ½, Met, Akt, STAT3, beta-catenin, E-cadherin, EGFR vIII, IGFR-1, NF-kB, p53, PI3Kp85, PI3Kp110a, PTEN, NRAS, and pRB. These protein biomarkers highlighted the important role that the MAPK and PI3K pathways play in HNSCC. Consistent with mutational analysis, overexpression of the peptides in these pathways were associated with inferior overall survival and inferior progression-free survival. Protein expression of ERK 1/2 had the most promising correlation with outcomes.

4.5.2 Surrogate Tissue Studies

Serum studies have been widely used in investigations of HNSCC given the challenges in obtaining repeat tumor samples. One study used MALDI on sera from 99 HNSCC and 143 controls to obtain serum protein patterns [57]. The mass spectra and linear discriminant analysis were used to select the top 45 spectral features. The subsequent spectral profiles from the sera of the HNSCC patients statistically significantly differed from the sera of control subjects. In a separate study, samples were analyzed by SELDI-TOF, and 80 common peaks or clusters were generated from the training set and used to create classification trees [58]. This algorithm correctly identified 91 % of HNSCC sera in the training set and 83 % of HNSCC samples in the test set, yielding an overall sensitivity of 83 % and an overall specificity of 90 %. Furthermore, they were able to identify a particular peak as the known biomarker metallopanstimulin-1 based on mass and whose relative intensity consistently correlated with levels detected by radioimmunoassay.

More recent research has sought novel surrogate tissue sources, which may be convenient for investigation. Alterations in the levels of biomarkers have been investigated in other body fluids that are near or bathe tumor sites. Accordingly, saliva is an ideal complementary resource for

developing HNSCC diagnostics, and more recent study attempts have focused on the use of salivary proteomics for oral cancer biomarker discovery. One analysis collected saliva from 64 oral SCC and 64 healthy subjects and utilized subtractive proteomics to find that several salivary proteins were differentially expressed [59]. Five candidate biomarkers were validated and demonstrated high sensitivity (90 %) and specificity (83 %) in detecting oral SCC. Another recent study found two proteins, alpha-1-B-glycoprotein and complement factor B proteins, to be present in patients with HNSCC but not in normal specimens, while cystatin S, parotid secretory factor, and poly-4-hydrolase beta-subunit proteins were detected in most normal saliva samples but not in HNSCC [60]. These results suggest that certain proteins are differentially found in patients and normal saliva and a small set of proteins may be useful for future validation for clinical investigation. Finally, another study built on prior data indicating that the expression of IL-6 and IL-8 are uniquely associated with oral SCC. They analyzed patients with newly diagnosed T1 or T2 oral cavity or histologically confirmed oropharyngeal SCC. Their analysis revealed that IL-8 was detected at higher concentrations in saliva and IL-6 was detected at higher concentrations in serum of patients with oral SCC, indicating that these markers and tissues hold promise for biomarker analysis in oral SCC [61].

Tandem mass spectrometry has also been used to identify proteins that may serve as biomarkers for neck disease. A recent study used serum from 40 patients, 18 without neck disease, and quantified 282 serum proteins [62]. Four candidate biomarkers (gelsolin, fibronectin, angiotensinogen, and haptoglobin) were identified, and the best one, gelsolin, had high validity for identifying node-positive HNSCC. Gelsolin is a cytosolic protein that regulates cytoskeleton assembly and disassembly. It is a protein that has been implicated in epithelial to mesenchymal transitions.

Table 4.2 Key advantages and limitations of DNA microarrays

Advantages
– Provide insight into fluctuations in gene transcription
– Capable of generating large amounts of expression data quickly
– Current microarrays give expression data from essentially the entire genome
– Technological advances have generated microarrays that can be implemented using automated, high-throughput strategies at reduced costs
Limitations
– High-quality RNA is required for the generation of good expression data
– Changes in RNA expression may not correlate with changes in protein levels
– Advanced biostatistics are necessary to process vast amounts of data generated

4.6 Challenges of HNSCC Oncogenomics/Proteomics

The application of these novel technologies offers many opportunities for advanced analyses of HNSCC (Table 4.2). With the completion of the Human Genome Project and advances in array technology, gene expression studies offer an opportunity to look at the full complement of genes expressed by a tumor. Gene expression profiling experiments have generated a tremendous amount of information regarding concomitant genetic events during disease. However, the functional consequences of disease are also regulated by the deregulation of protein products and protein networks so the information flow cannot be ascertained from gene analysis alone.

Furthermore, there are a variety of potential pitfalls in microarray analysis that may obscure the quantification of genes of interest. One of the most important variables relates to the quality of the transcripts utilized for the microarray, which may relate to initial and long-term tissue handling as well as processing of the transcripts for use in the microarray studies. A recent report indicated that there may be a storage time decrease in the predictive performance of tissue samples. There may be a decrease in the predictive performance of tissue samples based on their storage time. Other common causes of signal variations include errors with fluidics protocols, spoiled or omitted hybridization cocktail reagents, and inaccurate quantification of labeled samples. There are also a variety of factors inherent to the microarray technology such as intensity-dependent dye effect and spatial-dependent dye effect that can influence the quantification process. In addition, studies vary in the heterogeneity of the cell types included in the samples from 50 % tumor cells to the pure isolation of single tumor cells.

By contrast, proteins are dependent on highly regulated processes at the transcriptional, translational, and posttranslational level (Table 4.3). Many of the standard proteomic approaches rely on the usage of complex protein mixtures and the indirect assignment of spectra to identify target pro-

Table 4.3 Key advantages and limitations of proteomic approaches

Advantages
– Provide insight into fluctuations in transcribed and translated gene products as well as posttranslational modifications
– Capable of using a variety of tissue sources with minimal processing to analyze variations
– Increasingly offering high-throughput technologies
Limitations
– High-abundance proteins may obscure data
– Generally only analyze a minority of proteins within the entire sample
– Difficult to correlate individual spectral peaks/signatures with actual proteins

teins. These approaches are often hampered by the presence of large quantity proteins that may obscure quantification of the proteins of interest. Accordingly, there has been increasing interest in developing protein microarrays capable of identifying hundreds of protein events simultaneously; however, these arrays have a set of unique problems. Protein interactions are governed by complex associations between the target protein and the antigen-binding site on the antibody. Furthermore, proteins tend to denature with changes in pH or temperature, and antibodies must exhibit strong affinities and specificity to each of their respective substrates especially in the analysis of specific protein states such as phosphorylation or proteolytic cleavage. In addition, the variation in protein concentration in cells may vary widely, so detection methods must exist that can quantify protein concentration over many orders of magnitude.

These studies also require careful experiment planning starting with the selection of appropriate controls. Many studies use matched "normal" epithelium, but this may confound interpretations of gene expression changes occurring in HNSCC tumorigenesis. Although logistically difficult to achieve, the theoretically ideal control tissue would match for patient age, gender, smoking and drinking history, and other variables to minimize further confounding factors.

4.7 Conclusion and Future Directions

The goals of oncogenomics and proteomics are to improve diagnosis, therapy, and cure rates for cancer patients. A patient's genomic signature of a cancer may serve as the basis for choosing the most effective therapy for the individual patient to improve their chances of recovery and their quality of life. Oncogenomics and proteomics have progressed from molecular profiling to model systems, cancer pharmacology, and clinical trials. With whole-genome sequencing, personalized tumor profiles are now possible. Although it is unlikely that a single biomarker will accurately predict response to therapy, analyses that can detect multiple markers may have improved predictive value when used in combination. Imperfect biomarkers may still be clinically useful for serial testing of single individuals because acute changes in biomarker levels may signal the need for an aggressive search for the cause. An important challenge for biomarker validation is the considerable molecular heterogeneity of individual cancers and the low overall incidence of the disease in general population, making it difficult to validate the true prognostic potential of a biomarker or panel of biomarkers. 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.

Furthermore, although current oncogenomic and proteomic approaches may yield valuable information in the identification of novel diagnostic markers, gene and protein expression profiles may not be able to provide an alternative method of diagnosis on their own. It may become necessary to include other technologies such as metabolomics, peptidomics, glycomics, and lipidomics for better isolation and identification of molecular targets. In order to obtain reliable prognostic markers, these technologies will need to be combined with advanced bioinformatics tools to integrate and mine the data from basic and clinical research. Once molecular signatures are successfully validated, it will also be important to perform long-term clinical studies to determine the validity of using these signatures in independent cohorts of patients for the prediction of patient response to therapeutic options.

References

1. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008; 456(7218):53–9.
2. Cohen EEW, Zhu H, Lingen MW, Martin LE, Kuo WL, Choi EA, et al. A feed-forward loop involving protein kinase Cα and microRNAs regulates tumor cell cycle. Cancer Res. 2009;69:65–74.
3. Rickman DS, Millon R, De Reynies A, Thomas E, Wasylyk C, Muller D, et al. Prediction of future metastasis and molecular characterization of head and neck squamous-cell carcinoma based on transcriptome and genome analysis by microarrays. Oncogene. 2008;27:6607–22.
4. Thurlow JK, Murillo CLP, Hunter KD, Buffa FM, Patiar S, Betts G, et al. Spectral clustering of microarray data elucidates the roles of microenvironment remodeling and immune responses in survival of head and neck squamous cell carcinoma. J Clin Oncol. 2010;28:2881–8.
5. Chung CH, Parker JS, Ely K, Carter J, Yi Y, Murphy BA, et al. Gene expression profiles identify epithelial-to-mesenchymal transition and activation of nuclear factor-kappaB signaling as characteristics of a high-risk head and neck squamous cell carcinoma. Cancer Res. 2006;66:8210–8.
6. Chung CH, Parker JS, Karaca G, Wu J, Funkhouser WK, Moore D, et al. Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. Cancer Cell. 2004;5:489–500.
7. Walter V, Yin X, Wilkerson MD, Cabanski CR, Zhao N, Du Y, et al. Molecular subtypes in head and neck cancer exhibit distinct patterns of chromosomal gain and loss of canonical cancer genes. PLoS One. 2013;8(2), e56823.
8. Collins FS. The Cancer Genome Atlas (TCGA). Online. 2007;1–17.
9. Toruner GA, Ulger C, Alkan M, Galante AT, Rinaggio J, Wilk R, Tian B, Soteropoulos P, Hameed MR, Schwalb MN, et al. Association between gene expression profile and tumor invasion in oral squamous cell carcinoma. Cancer Genet Cytogenet. 2004;154(1):27–35.
10. Ye H, Yu T, Temam S, Ziober BL, Wang J, Schwartz JL, Mao L, Wong DT, Zhou X. Transcriptomic dissection of tongue squamous cell carcinoma. BMC Genomics. 2008;9:69.

11. Kuriakose MA, Chen WT, He ZM, Sikora AG, Zhang P, Zhang ZY, Qiu WL, Hsu DF, McMunn-Coffran C, Brown SM, et al. Selection and validation of differentially expressed genes in head and neck cancer. Cell Mol Life Sci. 2004;61(11):1372–83.
12. Sticht C, Freier K, Knopfle K, Flechtenmacher C, Pungs S, Hofele C, Hahn M, Joos S, Lichter P. Activation of MAP kinase signaling through ERK5 but not ERK1 expression is associated with lymph node metastases in oral squamous cell carcinoma (OSCC). Neoplasia. 2008;10(5):462–70.
13. Pyeon D, Newton MA, Lambert PF, den Boon JA, Sengupta S, Marsit CJ, Woodworth CD, Connor JP, Haugen TH, Smith EM, et al. Fundamental differences in cell cycle deregulation in human papillomavirus-positive and human papillomavirus-negative head/neck and cervical cancers. Cancer Res. 2007;67(10):4605–19.
14. Ha PK, Benoit NE, Yochem R, Sciubba J, Zahurak M, Sidransky D, et al. A transcriptional progression model for head and neck cancer. Clin Cancer Res. 2003;9(8):3058–64.
15. Chen C, Méndez E, Houck J, Fan W, Lohavanichbutr P, Doody D, et al. Gene expression profiling identifies genes predictive of oral squamous cell carcinoma. Cancer Epidemiol Biomarkers Prev. 2008;17:2152–62.
16. Ginos MA, Page GP, Michalowicz BS, Patel KJ, Volker SE, Pambuccian SE, et al. Identification of a gene expression signature associated with recurrent disease in squamous cell carcinoma of the head and neck. Cancer Res. 2004;64:55–63.
17. Ziober AF, Patel KR, Alawi F, Gimotty P, Weber RS, Feldman MM, et al. Identification of a gene signature for rapid screening of oral squamous cell carcinoma. Clin Cancer Res. 2006;12:5960–71.
18. Kondoh N, Ohkura S, Arai M, Hada A, Ishikawa T, Yamazaki Y, et al. Gene expression signatures that can discriminate oral leukoplakia subtypes and squamous cell carcinoma. Oral Oncol. 2007;43:455–62.
19. Stransky N, Egloff AM, Tward AD, Kostic AD, Cibulskis K, Sivachenko A, et al. The mutational landscape of head and neck squamous cell carcinoma. Science. 2011;333(6046):1157–60.
20. Agrawal N, Frederick MJ, Pickering CR, Bettegowda C, Chang K, Li RJ, et al. Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science. 2011;333(6046):1154–7.
21. International Cancer Genome Consortium [Internet]. ICGC Cancer Genome Projects. 2014. https://icgc.org/. Accessed 22 Dec 2014.
22. Howell GMS, Grandis JR. Molecular mediators of metastasis in head and neck squamous cell carcinoma. Head Neck. 2005;27(8):710–7.
23. Schmalbach CE, Chepeha DB, Giordano TJ, Rubin MA, Teknos TN, Bradford CR, et al. Molecular profiling and the identification of genes associated with metastatic oral cavity/pharynx squamous cell carcinoma. Arch Otolaryngol Head Neck Surg. 2004;130:295–302.
24. Belbin TJ, Singh B, Smith RV, Socci ND, Wreesmann VB, Sanchez-Carbayo M, et al. Molecular profiling of tumor progression in head and neck cancer. Arch Otolaryngol Head Neck Surg. 2005;131:10–8.
25. Roepman P, Wessels LFA, Kettelarij N, Kemmeren P, Miles AJ, Lijnzaad P, et al. An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas. Nat Genet. 2005;37:182–6.
26. Hensen EF, De Herdt MJ, Goeman JJ, Oosting J, Smit VT, Cornelisse CJ, et al. Gene-expression of metastasized versus non-metastasized primary head and neck squamous cell carcinomas: a pathway-based analysis. BMC Cancer. 2008;8:168.
27. Cromer A, Carles A, Millon R, Ganguli G, Chalmel F, Lemaire F, et al. Identification of genes associated with tumorigenesis and metastatic potential of hypopharyngeal cancer by microarray analysis. Oncogene. 2004;23:2484–98.
28. Weber F, Xu Y, Zhang L, Patocs A, Shen L, Platzer P, et al. Microenvironmental genomic alterations and clinicopathological behavior in head and neck squamous cell carcinoma. JAMA. 2007;297:187–95.
29. Chen C, Wei Y, Hummel M, Hoffmann TK, Gross M, Kaufmann AM, et al. Evidence for epithelial-mesenchymal transition in cancer stem cells of head and neck squamous cell carcinoma. PLoS One. 2011;6(1), e16466.
30. Duvvuri U, Shiwarski DJ, Xiao D, Bertrand C, Huang X, Edinger RS, et al. TMEM16A, induces MAPK and contributes directly to tumorigenesis and cancer progression. Cancer Res. 2012;72:3270–81.
31. Shiwarski DJ, Shao C, Bill A, Kim J, Xiao D, Bertrand C, et al. To "grow" or "go": TMEM16A expression as a switch between tumor growth and metastasis in SCCHN. Clin Cancer Res. 2014;20(17):4673–88.
32. Ishigami T, Uzawa K, Higo M, Nomura H, Saito K, Kato Y, et al. Genes and molecular pathways related to radioresistance of oral squamous cell carcinoma cells. Int J Cancer. 2007;120:2262–70.
33. Torres-Roca JF, Eschrich S, Zhao H, Bloom G, Sung J, McCarthy S, et al. Prediction of radiation sensitivity using a gene expression classifier. Cancer Res. 2005;65:7169–76.
34. Akervall J, Nandalur S, Zhang J, Qian C-N, Goldstein N, Gyllerup P, et al. A novel panel of biomarkers predicts radioresistance in patients with squamous cell carcinoma of the head and neck. Eur J Cancer. 2014;50(3):570–81.
35. Pramana J, Van den Brekel MWM, van Velthuysen M-LF, Wessels LFA, Nuyten DS, Hofland I, et al. Gene expression profiling to predict outcome after chemoradiation in head and neck cancer. Int J Radiat Oncol Biol Phys. 2007;69(5):1544–52.
36. Ganly I, Talbot S, Carlson D, Viale A, Maghami E, Osman I, et al. Identification of angiogenesis/metastases genes predicting chemoradiotherapy response in patients with laryngopharyngeal carcinoma. J Clin Oncol. 2007;25:1369–76.
37. Olaussen KA, et al. DNA repair by ERCC1 in non-small-cell lung cancer and cisplatin-based adjuvant chemotherapy. N Engl J Med. 2006;355(10):983–91.
38. Bauman JE, Austin MC, Schmidt R, Kurland BF, Vaezi A, Hayes DN, et al. ERCC1 is a prognostic biomarker in locally advanced head and neck cancer: results from a randomised, phase II trial. Br J Cancer. 2013;109(8):2096–105.
39. Roh J-L, Kim EH, Park HB, Park JY. The Hsp90 inhibitor 17-(allylamino)-17-demethoxygeldanamycin increases cisplatin antitumor activity by inducing p53-mediated apoptosis in head and neck cancer. Cell Death Dis. 2013;4(12), e956.
40. Li Y, St. John MAR, Zhou X, Kim Y, Sinha U, Jordan RCK, et al. Salivary transcriptome diagnostics for oral cancer detection. Clin Cancer Res. 2004;10:8442–50.
41. Li Y, Elashoff D, Oh M, Sinha U, St John MAR, Zhou X, et al. Serum circulating human mRNA profiling and its utility for oral cancer detection. J Clin Oncol. 2006;24:1754–60.
42. John K, Wu J, Lee B, Farah CS. MicroRNAs in head and neck cancer. Int J Dent. 2013;2013:650218. doi:10.1155/2013/650218.
43. Davalos V, Esteller M. MicroRNAs and cancer epigenetics: a macrorevolution. Curr Opin Oncol. 2010;22:35–45.
44. Nagadia R, Pandit P, Coman WB, Cooper-White J, Punyadeera C. miRNAs in head and neck cancer revisited. Cell Oncol (Dordr). 2013;36(1):1–7.
45. Salazar C, Nagadia R, Pandit P, Cooper-White J, Banerjee N, Dimitrova N, et al. A novel saliva-based microRNA biomarker panel to detect head and neck cancers. Cell Oncol (Dordr). 2014;37(5):331–8.
46. Ahn SM, Chan JYK, Zhang Z, Wang H, Khan Z, Bishop JA, et al. Saliva and plasma quantitative polymerase chain reaction-based detection and surveillance of human papillomavirus-related head and neck cancer. JAMA Otolaryngol Head Neck Surg. 2014;140(9):846–54.

47. Choi P, Chen C. Genetic expression profiles and biologic pathway alterations in head and neck squamous cell carcinoma. Cancer. 2005;104:1113–28.
48. Yu Y-H, Kuo H-K, Chang K-W. The evolving transcriptome of head and neck squamous cell carcinoma: a systematic review. PLoS One. 2008;3, e3215.
49. De Cecco L, Bossi P, Locati L, Canevari S, Licitra L. Comprehensive gene expression meta-analysis of head and neck squamous cell carcinoma microarray data defines a robust survival predictor. Ann Oncol. 2014;25(8):1628–35.
50. Soltys SG, Le QT, Shi G, Tibshirani R, Giaccia AJ, Koong AC. The use of plasma surface-enhanced laser desorption/ionization time-of-flight mass spectrometry proteomic patterns for detection of head and neck squamous cell cancers. Clin Cancer Res. 2004;10:4806–12.
51. Roesch-Ely M, Nees M, Karsai S, Ruess A, Bogumil R, Warnken U, et al. Proteomic analysis reveals successive aberrations in protein expression from healthy mucosa to invasive head and neck cancer. Oncogene. 2007;26:54–64.
52. Freed GL, Cazares LH, Fichandler CE, Fuller TW, Sawyer CA, Stack BC, et al. Differential capture of serum proteins for expression profiling and biomarker discovery in pre- and posttreatment head and neck cancer samples. Laryngoscope. 2008;118:61–8.
53. Ralhan R, Desouza LV, Matta A, Chandra Tripathi S, Ghanny S, Datta Gupta S, et al. Discovery and verification of head-and-neck cancer biomarkers by differential protein expression analysis using iTRAQ labeling, multidimensional liquid chromatography, and tandem mass spectrometry. Mol Cell Proteomics. 2008;7:1162–73.
54. Gourin CG, Zhi W, Adam B-L. Proteomic identification of serum biomarkers for head and neck cancer surveillance. Laryngoscope. 2009;119:1291–302.
55. Choi P, Jordan CD, Mendez E, Houck J, Yueh B, Farwell DG, et al. Examination of oral cancer biomarkers by tissue microarray analysis. Arch Otolaryngol Head Neck Surg. 2008;134:539–46.
56. Psyrri A, Lee J-W, Pectasides E, Vassilakopoulou M, Kosmidis EK, Burtness BA, et al. Prognostic biomarkers in phase II trial of cetuximab-containing induction and chemoradiation in resectable HNSCC: Eastern cooperative oncology group E2303. Clin Cancer Res. 2014;20(11):3023–32.
57. Sidransky D, Sidransky D, Irizarry R, Califano JA, Li X, Ren H, et al. Serum protein MALDI profiling to distinguish upper aerodigestive tract cancer patients from control subjects. J Natl Cancer Inst. 2003;95:1711–7.
58. Wadsworth JT, Somers KD, Stack Jr BC, Cazares L, Malik G, Adam BL, et al. Identification of patients with head and neck cancer using serum protein profiles. Arch Otolaryngol Head Neck Surg. 2004;130:98–104.
59. Hu S, Arellano M, Boontheung P, Wang J, Zhou H, Jiang J, et al. Salivary proteomics for oral cancer biomarker discovery. Clin Cancer Res. 2008;14:6246–52.
60. Ohshiro K, Rosenthal DI, Koomen JM, Streckfus CF, Chambers M, Kobayashi R, et al. Pre-analytic saliva processing affect proteomic results and biomarker screening of head and neck squamous carcinoma. Int J Oncol. 2007;30:743–9.
61. St John MA, Li Y, Zhou X, Denny P, Ho CM, Montemagno C, et al. Interleukin 6 and interleukin 8 as potential biomarkers for oral cavity and oropharyngeal squamous cell carcinoma. Arch Otolaryngol Head Neck Surg. 2004;130:929–35.
62. Chai YD, Zhang L, Yang Y, Su T, et al. Discovery of potential serum protein biomarkers for lymph-node metastasis in oral cancer. Head Neck. 2016;38:118–25.