UCSF
UC San Francisco Electronic Theses and Dissertations

Title
Genomic and Genetic Profiling of Oral Squamous Cell Carcinoma Progression and Metastasis

Permalink
https://escholarship.org/uc/item/3wx601qb

Author
Bhattacharya, Aditi

Publication Date
2010

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Genomic and Genetic Profiling of Oral Squamous Cell Carcinoma Progression and Metastasis

by

Aditi Bhattacharya

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Oral and Craniofacial Sciences

in the

Copyright (2010)

by

Aditi Bhattacharya

Acknowledgements

I would like to express my heartfelt gratitude to all the individuals who have contributed to the successful completion of this project.

I would first like to express my sincere gratitude to my thesis advisor and dissertation chair Donna Albertson. Thank you Donna, for your patience, guidance and support. I have come to value immensely our discussions on science and so many other things. Thank you for teaching me to deal with problems by identifying the cause, to find workable solutions and above all never stopping in the pursuit of knowledge. I truly admire your ability to make science fun and make me look forward to each new day in the lab. I am indeed lucky to have a mentor like you!!

I would also like to thank Brian Schmidt my thesis co-mentor and clinical advisor. Brian, thanks for giving me the opportunity of working on this project and letting me have a free hand. Thank you for your constructive feedback, positive attitude and tremendous support during challenging times.

Thanks to my committee members Richard Jordan, Fred Waldman and Joe Costello for being readily available to view my data and provide constructive guidance regarding the science and writing. Richard, special thanks to you for your constant support as my program director.

My faculty advisor John Greenspan, for giving me the opportunity to pursue graduate education at UCSF. I am thankful for the fortuitous meeting many years ago in India!!

Dan Pinkel for the discussions on array CGH and FISH and instilling confidence to face my qualifying exam. Rick Segraves, Antoine Snijders and Bing Huey for teaching me about science, and life in the lab. Bing thank you also for the introduction to American popular music and culture during the hours spent in cell culture!! A big thank you to my labmates Hua, Thad, Aruna, Mau-ting and Eui-Chang for continuous support and for sharing equally my elation and disappointment with my data.

Special thanks to many many people: Adam Olshen, Taku Tokuyasu and Ritu Roy (Bioinformatics Core); Greg Hamilton (Array Core); Ann Griffin (UCSF Cancer Registry); Hubert Stoppler and Tatiana Rogut (Cancer Center Tissue Core); Linda Prentice (Orofacial Sciences Tissue Core); Dave Lam, Stacy Achdijan and Jenny Zhang (Schmidt lab); Ricardo Salinas, Katie Colino, Michael Brown and Nusi Dekker (Orofacial Sciences) for all their help in various aspects of this project.

Thanks to my brother Ashmit for always being supportive and urging me to follow the unbeaten track. Thanks also to him and his wife Sunita for welcoming me to their home when I moved to the United States and making the transition so easy for me. Profound gratitude and love to my parents Sailesh and Krishna for teaching me that nothing is impossible……..once you set your mind on it.

Genomic and Genetic Profiling of Oral Squamous Cell Carcinoma Progression and Metastasis

Abstract

Oral squamous cell carcinoma (SCC) is the 6th most common malignancy in the US and is often associated with a history of tobacco and alcohol use. Five year survival and prognosis remain poor at 60-80% with the outcome becoming bleaker upon metastasis to the cervical (neck) lymph nodes (25-40%). Oral SCCs are preceded by oral dysplasias (precancers), but not all oral dysplasias turn malignant. Thus, challenges facing oral cancer surgeons and oncologists include prediction of which dysplasias will undergo malignant transformation and which tumors will metastasize to the neck. Therefore there is an urgent need to identify accurate biomarkers to aid in the treatment and management of patients with oral cancers or pre-cancers. This thesis research addressed these problems by assessing whether genomic DNA copy number measurements could provide such markers. As described below, the work has culminated in the identification of two oral SCC subtypes that differ in clinical behavior, specifically the risk of metastasis. The genomic copy number biomarker is suitable for stratifying patients for risk of metastasis prior to surgery and so can guide surgical planning for one of the most challenging treatment decisions – how to treat patients with clinically node-negative (N0) necks. Current clinical practice recommends removal of the cervical lymph nodes at the time of surgery to resect the tumor, because salvage of patients who subsequently develop lymph node metastasis has a poor success rate. Identification of patients with low risk for metastasis would spare them additional major surgery with its risks and morbidity, as well as reduce medical costs. A molecular marker for metastasis risk will also inform follow up management of oral cancer patients.

Table of contents

Chapter 1: Introduction, Background and Specific Aims

Introduction……………………………………………………………………………….2

Oral squamous cell carcinoma………………………………………………………......3

Current management of oral SCC………………………………………………………4

Metastasis…………………………………………………………………………………5

Molecular events well established to contribute to oral carcinogenesis and metastasis…..7

Role of stroma in metastasis……………………………………………………………..10

Summary and significance……………………………………………………………….11

Chapter 2: Results and Discussion

Genomic copy number measurements of oral dysplasia and SCC………………………14

DNA isolation and array CGH…………………………………………………………...15

Dysplasia cohort…………………………………………………………………………15

Genomic analysis of Oral Dysplasia and Oral SCCs……………………………………15

Amplicons in oral dysplasia and SCCs……………………………………………......21

Two routes to oral cancer………………………………………………………………...22

Association of 3q8pq20 status with clinical parameters…………………………………22

Establishment of CAF and NF cell cultures……………………………………………..27

Chapter 3: Conclusion and Future Directions

Conclusions………………………………………………………………………………30

Future directions…………………………………………………………………………32

Chapter 4: Materials and Methods

Identification of Cohort of Metastatic and Non-metastatic Oral Squamous Cell Carcinomas (cohort#2)…………………………………………………………………...36

DNA Isolation and Array CGH on Oral Squamous Cell Carcinoma Cohort…………....38

Cell Culture and Reagents…………………………………………………………………41

Immunofluorescence……………………………………………………………………..43

TP53 sequencing…………………………………………………………………………44

References……………………………………………………………………………..58

List of Tables

Table 1: Clinical characteristics of cohort#2……………………………………………….46

Table 2: Dysplasia cases…………………………………………………………………...48

Table 3: Recurrent regions of aberration at ≥ 20% frequency in dysplasia with no known association with cancer………………………………………………………………..50

Table 4: Frequency in SCC cohorts and dysplasia associated with cancer of copy number changes occurring in ≥ 20% of dysplasia with no known association with cancer……...51

Table 5: Amplicons in 29 dysplasia samples from patients with no known history of oral cancer………………………………………………………………………………….52

Table 6: Patient and tumor characteristics relative to tumor subtype…………………...53

Table 7: Patient and tumor characteristics relative to cervical node status……………...54

Table 8: Clinical features of established cell cultures…………………………………...55

Table 9: Cell culture characterization…………………………………………………...57

List of figures

Figure 1: Copy number aberrations in dysplasias and oral SCC cohort#1……...............17

Figure 2: Copy number aberrations in oral SCC cohort#2……………………………………...20

Figure 3: Non-uniform distribution of low level gains and losses among 3q8pq20 and non-3q8pq20 oral SCC cases……………………………………………………………….24

Figure 4: Copy number aberrations frequent in oral dysplasia with cancer…………………….25

Figure 5: Two routes to oral cancer……………………………………………………..26

Chapter 1: Introduction, Background and Specific Aims

Introduction

The prognosis of patients with oral squamous cell carcinoma (SCC) has remained relatively unchanged for several decades. While the 5 year survival of metastasis free patients remains at 60-80%, survival reduces to 25-40% when metastatic tumor is present in cervical lymph nodes. The biologic behavior of oral SCC is unpredictable. Currently treatment is guided by the UICC TNM system, but this has not been very successful or accurate at predicting survival or prognosis. The mainstay of treatment is surgical removal of the tumor and cervical lymph node dissection. If there is evidence of residual tumor or neck involvement, postoperative radiation and chemotherapy are used to reduce the chance for local recurrence and/or regional failure in the lymph nodes. Surgery and radiation therapy are both associated with severe, progressive complications. Therefore there is an acute need to identify biomarkers that can accurately detect metastasis. The goal of this project was to identify predictive markers of oral SCC metastasis. We undertook two approaches. In the first, we took advantage of the fact that oral SCCs are commonly characterized by recurrent DNA copy number aberrations (CNAs), which often result from unbalanced translocations, amplifications and deletions. (Akervall, 2006). Since tumor progression is associated with multi-step genetic, genomic and/or epigenetic changes, we hypothesize that the spectra of genomic alterations present in metastatic compared to non metastatic oral SCCs differs. These differences were used to identify genomic biomarkers that could detect metastatic potential in an oral SCC. In the second approach, we investigated whether the stroma of metastatic and non-metastatic tumors differs and therefore whether these differences can be exploited to develop biomarkers of metastasis.

Oral squamous cell carcinoma

In 2008, an estimated 47,500 people were diagnosed with head and neck cancer in the United States, representing approximately 3% of new cancer diagnoses, and an estimated 11,260 people died from this disease (Jemal et al., 2008). Oral squamous cell carcinoma (SCC) is mainly a disease of older men, 90% being in the over 45-year-age group who are exposed to the known risk factors of tobacco and/or alcohol (IARC, 2004). Although there has been a reduction in the overall incidence of head and neck SCC (HNSCC), a recent analysis of SEER data from 1973-2001 showed an annual increase in the incidence of oral tongue, base of tongue and tonsil cancer of 2.1% in 20- to 44-year-old white patients, while the overall incidence of SCC at other oral cavity and pharyngeal sites declined (Shiboski et al., 2005).

Head and neck SCC includes cancers of the oral cavity, larynx and pharynx. (IARC-WHO 2002) (Mignogna et al., 2004). Oral SCCs are distinct from other HNSCCs such as tonsillar, laryngeal or base of tongue in their embryologic origin, etiology, clinical presentation and behavior, response to treatment, bio-environment, survival and prognosis (Dobrossy, 2005, Timar et al., 2005). In particular, the prevalence of human papillomavirus (HPV) in oral SCC is relatively low about 2-5% (Smith et al., 2004, Ritchie et al., 2003) in comparison to tonsillar SCC where the prevalence of HPV 16 and 18 is up to 60% of cases. (Mork et al., 2001, Sugerman and Shillitoe, 1997, Fakhry et al., 2008). Thus, various studies of HNSCC as a group are likely to have been confounded by inclusion of at least two distinct disease entities. Indeed, it has been proposed that oral cavity and oropharyngeal tumors are initiated by separate molecular pathways, exposure to environmental carcinogens (modifiable life-style risk factors such as tobacco and alcohol use) and HPV infection, respectively (Smeets et al., 2009).On the other hand,

there are 25 % of patients who present with oral cavity cancer who are non-smokers and non-drinkers (Blot et al., 1988), and it is possible that oral SCC in these patients develop along yet a third molecular pathway.

Current management of oral SCC

Once a diagnosis of SCC is established based on the histopathologic evaluation of the biopsy specimen, plans are made for clinical management. The mainstay of treatment is surgical removal of the tumor coupled with cervical lymph node dissection. Oral SCC patients are evaluated for possible cervical metastasis by clinical and radiographic examination. Because of the high and unpredictable rate of cervical metastasis and the possibility of occult metastasis, even when the radiographic studies are negative, a neck dissection is often performed to evaluate histologically the lymph nodes for metastasis and to remove the nodes in the case of occult metastasis (Baatenburg de Jong et al., 1993). In addition to adding time and cost to the operation, some neck dissections that compromise the accessory nerve are associated with a number of complications including shoulder dysfunction (Genden et al., 2003). Tumor margins are evaluated for residual disease and dissected lymph nodes are evaluated for histopathologic evidence of metastasis. If there is residual tumor involvement or positive lymph nodes the patients are then treated with radiation and/or chemotherapy.

While radiation therapy can improve tumor control and survival, radiation therapy to the oral cavity and neck produces profound side effects that are both acute and chronic. Additionally radiation therapy also adds significant cost to the treatment of oral cancer (Konski and Watkins-Bruner, 2004). Following surgery and radiation therapy patients with oral SCC are seen and examined on a regular basis to identify recurrence or

metastasis. Patients are typically followed once every month for the first year, every two months for the second year, and once every three months for the third and subsequent years for the life of the patient.

The incidence of metastasis tends to increase with the size of the tumor (Shah, 1990, Remmler et al., 1986), and tumor thickness (O'Brien et al., 1986, Kowalski and Sanabria, 2007, Spiro et al., 1986); however, small oral SCCs can metastasize early and some large oral SCCs never metastasize (Remmler et al., 1986, Mignogna et al., 2001). While it is clear that not all oral SCCs metastasize there is no clinical, pathologic or molecular marker that can be used to discriminate between oral SCCs that will metastasize and those that will not.

Metastasis

When oral SCCs metastasize, they typically spread to the cervical lymph nodes. Death for oral cancer patients is often due to either local or regional failure. Distant metastasis occurs in late stage disease, and is a rare event compared to cancers of the breast, lung, colon and prostate (Ferlito et al., 2001, Puri et al., 2003). More than 50% of patients diagnosed with oral SCC present with lymph node metastasis, with about 60% of the metastasis free cases reporting lymph node involvement within two years of diagnosis (Kowalski and Sanabria, 2007).

Metastasis requires that primary tumor cells lose adherence to adjacent normal and tumor epithelial cells and stroma and migrate toward the lymphatic, or more rarely, the blood system. Entry into the blood or lymphatic system (intravasation) is followed by arrest in and subsequently exit from the capillary or lymphatic vessels (extravasation). At

the metastatic site, angiogenesis and interaction with the host stroma is necessary for cell proliferation and the formation of macro metastases. Otherwise, cells may remain in a non-proliferative, dormant state. Thus, it is likely that tissue- and tumor-specific characteristics determine whether malignant cells will successfully colonize a distant site. As discussed above, oral SCCs typically metastasize locally to the cervical lymph nodes; distant, systemic metastasis is less common.

It is clear that neoplastic cells acquire both genetic and epigenetic changes that underlie the altered behavior of the cells. It has also been appreciated for some time that reciprocal communication between the stroma and tumor (or normal epithelial) cells can modify behavior of cells in both compartments resulting in normalization or promotion of the tumor cells. Thus, stromal cells, including activated carcinoma associated fibroblasts (CAFs), endothelial cells, infiltrating leukocytes and other mesenchymal cells unique to each organ site are likely contributors to, or regulators of, development of pre-neoplastic lesions, promotion to malignancy and neoplastic progression to metastases (Mueller and Fusenig, 2004).

In breast cancer, expression profiling has distinguished favorable and unfavorable prognosis based on expression from the entire tumor cell population, suggesting that genetic changes in the tumor dictate capability to metastasize. Breast cell populations have been selected that metastasize to particular organs and organ-specific functional involvement of some genes identified. If tumor and stroma have co-evolved then these studies suggest (1) that genomic alterations might be discovered that distinguish oral SCC likely to metastasize, and (2) that the primary and metastatic microenvironments should be similar so that examination of the tumor associated stroma of metastatic and non-

metastatic tumors could both reveal biomarkers of metastasis as well as predict likely sites of colonization. (Chambers et al., 2000, Mueller and Fusenig, 2004).

Molecular events well established to contribute to oral carcinogenesis and metastasis

Genomic alterations in oral SCC

Genomic alterations in oral SCC and oral dysplasia have been studied extensively using cytogenetics, chromosomal CGH and loss of heterozygosity (LOH) approaches. Sidransky and colleagues have proposed a genetic progression model for head and neck SCC (Califano et al., 1996, Sidransky, 1995). They reported that LOH of chromosomal region 9p21 was the most common genomic aberration in 70-80% dysplastic lesions of the oral mucosa- suggesting this was an early even in oral carcinogenesis (Califano et al., 1996, Mao et al., 1996, van der Riet et al., 1994). LOH of 3p is another common early event in epithelial carcinogenesis (Garnis et al., 2003, Masayesva et al., 2004). On the other hand LOH at 13q and 8p are observed more frequently in carcinomas than dysplasias and are considered later events in carcinogenesis (Califano et al., 1996). LOH of 9p21 is believed to be an early event in squamous neoplasia of the head and neck and has been found in preneoplastic lesions, including 30% of cases of squamous hyperplasia (Mao et al., 1996, Rosin et al., 2000, van der Riet et al., 1994). LOH of chromosome region 17p and mutation of the *TP53* gene are genetic alterations that are found in the later stage of progression from dysplasia to invasive SCC. *TP53* alterations including deletion and mutation are associated with increased genomic instability in oral cancer and may accelerate the rate of genetic alterations in oral SCC carcinogenesis. Amplification

at 11q13 and overexpression of cyclin D1 occur in about 40% cases of squamous dysplasia (Rousseau et al, 2001). Recently, Snijders et al. (2005) applied array comparative genomic hybridization to the analysis of 89 oral SCC cases using DNA extracted from routinely processed clinical formalin fixed paraffin embedded (FFPE) specimens. Common recurrent copy number alterations (CNAs) included deletion of chromosome arm 3p and 8p and gains of 3q, 8q and 20q. The regions less commonly affected were losses of 4q, 9p, 4p and gains of 5p, 7p, 9q and 14q (Snijders et al., 2005a). This spectrum of copy number alterations was consistent with earlier lower resolution studies carried out using chromosome CGH or low density genomic arrays; however they offered new evidence of the presence of recurrent focal amplicons in these tumors (Snijders et al., 2005a). Some earlier studies (using chromosome CGH, low density arrays) attempted to identify genomic profiles that defined metastatic oral SCCs, but these were conducted on primary oral SCCs and their matched lymph node metastases. These studies did not detect many significant differences in CNAs between the primary tumor and its metastasis, results which are not unexpected as both are likely to arise from the same clonal progenitor and therefore might not be expected to be very different genomically (Liu et al., 2006b, Patmore et al., 2004).

It is also important to note that most studies of oral cancer have been carried out on head and neck SCC (HNSCC). As discussed above, oral SCCs are distinct from other HNSCCs such as tonsillar, laryngeal or base of tongue tumors. Therefore, it is likely that the array CGH profiles obtained were not specific to oral SCCs. Thus, at the present time, the power of array CGH does not appear to have been appropriately applied to investigate whether genomic copy number profiles can distinguish metastatic and non-

metastatic oral SCC. This thesis research addressed this question by identifying and profiling non-metastatic (N0) oral SCCs and metastatic (N+) oral SCCs. N0 and N+ oral SCCs are distinct in their clinical behavior, response to treatment and prognosis; therefore it is reasonable to assume that their genomic profiles will also be dissimilar (Rosin et al., 2000, Liu et al., 2006b).

Gene expression profiles of oral SCC

Another approach to identifying the differences between metastatic and non metastatic oral SCCs has utilized microarray technology focusing on the gene expression profiles of N0 and N+ tumors (Nagata et al., 2003, Nguyen et al., 2007, O'Donnell et al., 2004, Schmalbach et al., 2004). These studies are challenging because ideally RNA isolated from fresh tissue is required, whereas most banked oral SCC specimens are formalin fixed, paraffin embedded (FFPE) and not known to yield RNA suitable for expression profiling. Perhaps, for this reason gene expression profiling studies have been conducted using the heteregenous group of head and neck SCCs with oral SCCs being a subset of the tumors analyzed. Since oral SCCs differ from other HNSCCs as discussed above it is likely that the sets of metastasis associated genes would differ between tumors in the oral cavity and other sites in the head and neck. Indeed, Chung *et al*. (2004) reported that the predictive accuracy of their gene signature improved from 63% to 83% if oral cavity tumors were removed from the analysis (Chung et al., 2004).

In a recent series of papers focusing on a large cohort comprised only of oral SCCs, Roepman and colleagues identified a 102 gene signature of oral SCC metastasis

(Roepman et al., 2005). Subsequently, these authors determined that these genes were part of a larger set of 825 genes that predicted lymph node metastasis (Roepman et al., 2006a, Roepman et al., 2006b). Moreover, analyses involving 685 of these genes found that 12% of the genes were expressed exclusively in the stroma, while 25% were expressed only in the tumor. The metastasis associated stromal genes were up-regulated, whereas the tumor genes were down regulated. The skewed distribution in expression impacted ability to predict nodal status in samples with lower percentages of tumor cells. Additional studies have found that the expression profiles of primary tumors and their lymph node metastases are very similar (Mendez et al., 2007, Roepman et al., 2006a, Weigelt et al., 2005). Taken together these observations suggest that the microenvironments of the primary and metastatic sites are similar and draw attention to the stroma as a source of biomarkers for detecting lymph node metastasis in oral SCC.

Role of stroma in metastasis

Tumor progression has been recognized as the product of an evolving crosstalk between different cell types within the tumor and its surrounding supporting tissue or tumor stroma (Mueller and Fusenig, 2004). The tumor stroma consists of the extracellular matrix (ECM), as well as cellular components such as fibroblasts, immune cells, inflammatory cells and blood vessel cells. Cancer cells can alter their adjacent stroma to form a permissive and supportive environment for tumor progression known as the 'reactive tumor stroma'. Among the stromal cells the carcinoma associated fibroblasts (CAFs) are of special interest since carcinoma cells induce normal fibroblasts to the CAF phenotype. These CAFs in the reactive stroma synthesize smooth muscle actin, ECM

components collagen I, III, fibronectin, tenascin and versican. They also express proteases including uPA, MMPs and FAP causing ECM remodeling leading to stimulation of cancer cell growth and migration (Brown et al., 1999, Lagace et al., 1985). Additionally CAFs secrete growth factors that promote angiogenesis and lymphangiogenesis (Koyama et al., 2008, Orimo et al., 2005). CAFs have been implicated in the progression of breast, prostate and pancreatic cancers (Iacobuzio-Donahue et al., 2002a, Iacobuzio-Donahue et al., 2002b, Tuxhorn et al., 2002). CAFs have been cultured from the stroma of oral SCCs and shown to express enhanced levels of TGFβ1 and MMP14 (Liu et al., 2006a, Rosenthal et al., 2004a, Rosenthal et al., 2004b). Thus, this component of the primary tumor stroma may be informative for discriminating between N0 and N+ tumors.

Summary and significance

Since tumor progression is associated with multi-step genetic, genomic and/or epigenetic changes, it is highly probable that the spectra of genomic alterations present in metastatic compared to non metastatic oral SCCs differ. On the other hand, the importance of the tumor microenvironment is becoming increasingly more apparent and the stroma is likely to participate in defining the metastatic niches that tumor cells might occupy. Identification of either genomic or stromal differences between N0 and N+ tumors could be employed to diagnose and detect metastasis in the clinical setting, as well as suggesting treatment protocols.

We have outlined a dual approach to identifying candidate biomarkers capable of distinguishing between N0 and N+ oral SCC. The first approach is motivated by our

hypothesis that the spectra of genomic alterations present in metastatic compared to non metastatic oral SCCs differs. Therefore, we undertook a retrospective study and evaluated N0 and N+ tumors for genomic copy number differences. The second approach is based on the hypothesis that the tumor stroma, in particular the CAFs associated with N0 and N+ tumors differ. To investigate this hypothesis we prospectively collected and characterized N0 and N+ associated CAFs and NFs.

Chapter 2: Results and Discussion

1. Genomic copy number measurements of oral dysplasia and SCC – assembly of cohorts.

We applied array comparative genomic hybridization (array CGH) to determine the genome-wide spectrum of copy number gains and losses in dysplasia and SCC cohorts to identify CNAs that might discriminate dysplasia from SCC and SCC with differing risk for metastasis. Prior work in the laboratory had profiled a group of 89 oral SCCs (cohort#1) lacking clinical information (Snijders et al., 2005b) and a smaller number of oral dysplasia. To address the objectives of the thesis, a second cohort of oral SCCs (cohort#2) was assembled with complete clinical information by identifying cases with five year follow up from the oral cancer tissue bank (OCTB) at UCSF. Inclusion criteria were oral cavity sites only with no chemotherapy or radiation prior to surgery, and tumor size ≥ 1.0 cm in order to maximize DNA yield for performing array CGH. From about 2500 cases in the OCTB, we identified 58 non-metastatic (N0, metastasis free for 5 years after diagnosis) and 96 metstatic (N+, metastatic at diagnosis or within 5 year follow up) SCCs diagnosed from 1998 – 2005. We also queried clinical records in the departments of Oral and Maxillofacial Surgery and Head and Neck Surgery and identified 14 N0 and 7 N+ tumors and 12 N0 and 21 N+ tumors from these sources, respectively. Our final tally of identified cases was 84 N0 and 149 N+ cases. We successfully accessioned these tissue blocks from 71 N0 and 82 N+ cases from the Department of Pathology. Upon review of H&E sections from the blocks we found 40 N0 and 48 N+ cases to have sufficient tissue for performing the array CGH protocol.

2. DNA isolation and array CGH

DNA was extracted from formalin fixed paraffin embedded (FFPE) archival blocks for 88 oral SCCs using established methods (Snijders 2003). 64 oral SCCs were found to have sufficient DNA concentrations for genomic analysis and we obtained array CGH profiles for these cases. Subsequent verification of date of metastasis from the Cancer Registry revealed that 14 profiled tumors had been incorrectly classified as N0 progressing to N+, as the patients had developed recurrent primary tumors, not metastases. Thus, on reclassifying, CGH profiles were available from 40 N0 and 23 N+ cases (Table 1). One case was eliminated due to poor CGH data.

3. Dysplasia cohort

Dysplasia cohorts for this project included 39 oral dysplasia samples, 29 with no known association with cancer and 10 that either subsequently progressed to cancer or appeared at the site of a previous cancer (Table 2). Dysplasia progression data was obtained from the OCTB and clinical records in the Department of Orofacial Sciences. DNA isolation, array CGH and data analysis were performed using the same techniques described for cohort#2

4. Genomic analysis of oral dysplasia and oral SCCs.

Dysplasias not associated with cancer: In dysplasias not associated with cancer, we found four regions of low level aberration (*e.g*. single copy gain and loss) to be present in

≥ 20% of cases (Figure 1A), including gains at 3q24-qter, 8q12-q24.2 and chromosome 20, and loss at 8pter-p23.1 (Table 3). The majority of the dysplasia cases (79%) harbored one or more of these recurrent aberrations, suggesting that these cases comprise a group, the 3q8pq20 subgroup. We observed a low frequency of additional aberrations in these dyplasia cases in addition to +3q, -8p, +8q and +20 (Figure 1C). The remaining 21% of cases lack +3q, -8p, +8q and +20 and only rarely had other regions of aberration. We term these the non-3q8pq20 subgroup. Dysplasia grade and *TP53* mutation status were not associated with one subgroup or the other (Figure 1C).

Figure 1: Copy number aberrations involving 3q, 8p, 8q and chromosome 20 are frequent in oral dysplasia and occur at similar frequency in oral SCC. (Legend on next page)

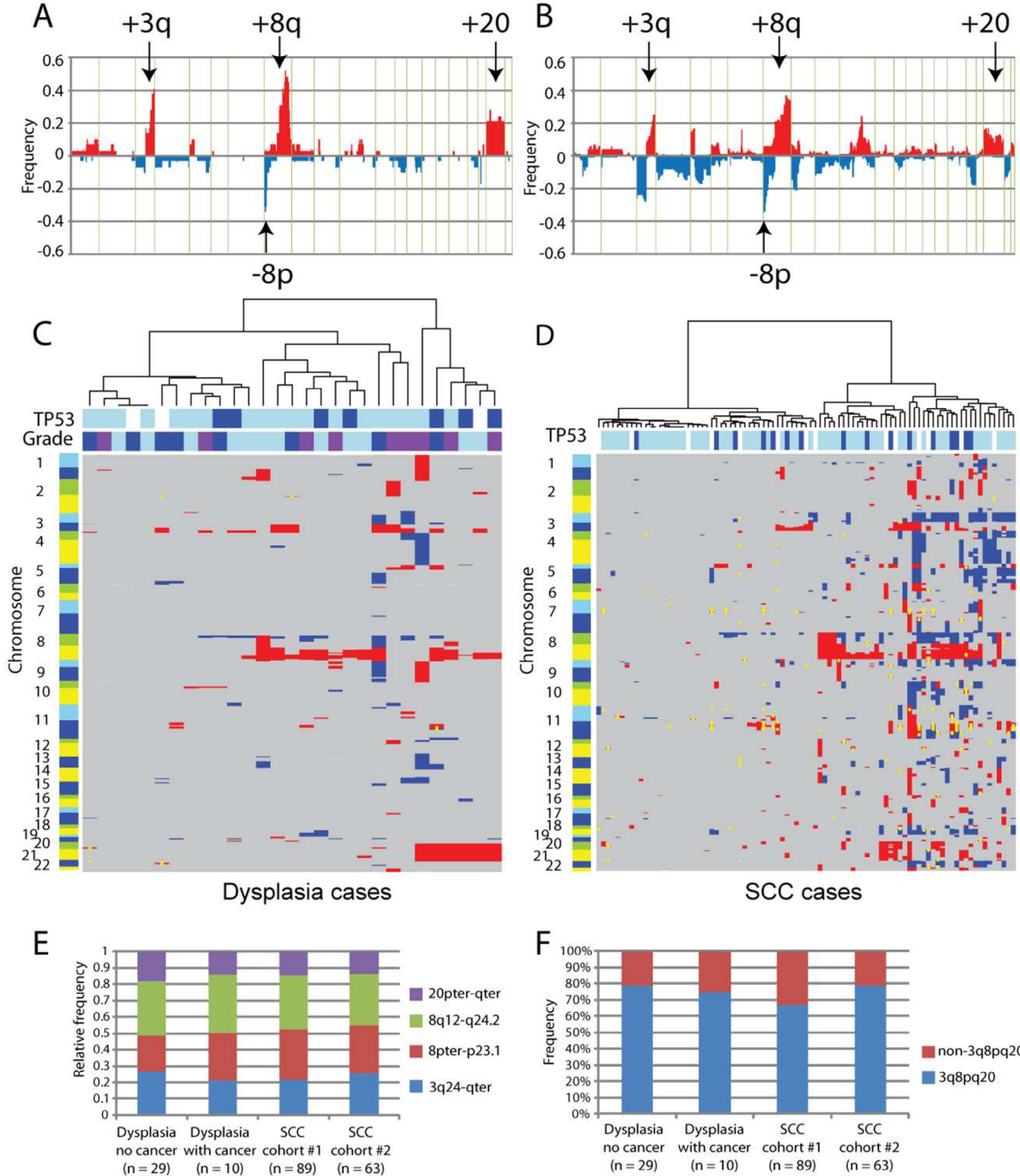


Figure 1 legend

(A and B) Frequency of copy number aberrations shown in genome order for each clone in 29 oral dysplasia samples with no known association with cancer (A) and oral SCC cohort#1 (B). We indicate gains by the red bars ranging from 0 to 0.6, and losses, by blue bars ranging from 0 to -0.06. Chromosome boundaries are indicated by vertical lines.

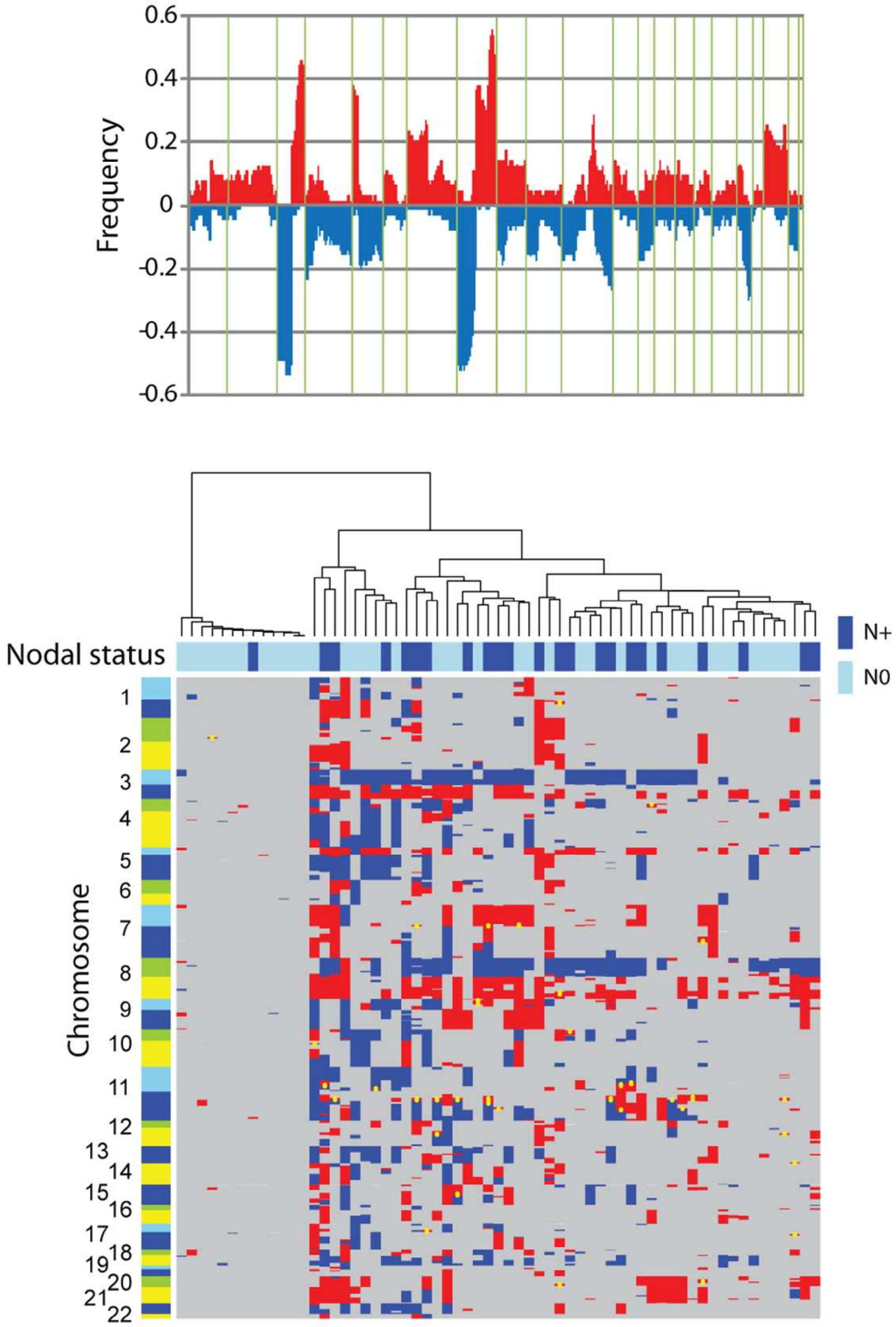
(C and D) Hierarchical clustering based on genome-wide DNA copy number profile of 29 oral dysplasia samples with no known association with cancer (C) and oral SCC cohort#1 (D). We generated heatmaps by unsupervised clustering of samples on trichotomous gain/loss/normal data for the autosomes. We used Euclidean distance d as the distance metric and Ward's linkage as the agglomeration method. We represented individual clones as rows and ordered them by chromosome and genome position according to the May 2004 freeze of the human genome (hg17). Clones on the p-arm are indicated either in light blue or yellow, and clones on the q-arm in dark blue or green. We show acrocentric chromosomes in green or dark blue. Columns represent individual tumor samples. We colored gains and losses red and blue, respectively and focal amplifications yellow. We indicated dysplasia grade (mild, light blue; moderate, dark blue; severe, purple) and the *TP53* mutation status of cases (*TP53* mutant, dark blue; no detected mutation, light blue; *TP53* status unknown, white).

(E) Frequencies of gains of 3q, 8q, 20 and loss of 8p in oral dysplasia and SCC normalized to the total number of aberrations at these loci in each cohort.

(F) Frequency of 3q8pq20 and non-3q8pq20 cases in oral dysplasia and SCC.

Oral SCCs: We noted that gains of 3q, 8q, 20 and loss of 8p were also frequent aberrations in the SCC cohort#1 (Figure 1B), as well as in the independent set of 63 oral SCCs in cohort#2 (Figure 2). Each of these aberrations occurred at a similar frequency in the two SCC cohorts as in the dysplasia cases (Figure 1E, Table 4). Moreover, the frequency of tumors harboring one or more of these aberrations was similar to the frequency in dysplasia (Figure 1F, 67% and 79% SCC cohort#1 and cohort#2, respectively). In the SCCs, we also observed additional recurrent gains and losses. For example, in SCC cohort#1, 11 loci that occurred in ≥15% of cases included -3p, -4p, -4q, +5p, -5q, +7p, -9p, +11q13, -18q, -21q and a loss at 8p12 that maps proximal to the region of loss at 8p shared by dysplasia and SCC (Figure 1B). Hierarchical clustering of the SCC cohorts revealed that recurrent low level gains and losses were not uniformly distributed among the SCC cases (Figures 1D and 2). For example, the median number of these recurrent copy number aberrations in SCC cohort#1 as a whole was 2 (range 0-13), 26 tumors (29%) harbored ≥5 of these copy number changes or 69% of the total number of these recurrent aberrations (203/295). Separating the oral SCCs into 3q8pq20 and non-3q8pq20 types (Figure 3) revealed that the recurrent aberrations were more frequent in the 3q8pq20 group (median = 4, range 1-13) than the non-3q8pq20 tumors (median = 0, range 0-7). In cohort#1 for which the mutational status of *TP53* had been assessed, mutations in exons 5-8 of *TP53* were more frequent in the 3q8pq20 group compared to the non-3q8pq20, but the difference was not significant (Fisher's exact test p-value = 0.12).

Figure 2. Copy number aberrations in oral SCC cohort#2. Frequency plot (top) and hierarchical clustering of cases showing nodal status (bottom) as in Figure 1.



Dysplasias associated with cancer: Analysis of a limited number of dysplasia cases (n=10) that progressed to cancer or arose at the site of a previously treated cancer revealed that 3q8pq20 and non-3q8pq20 types were present in similar proportions as in the dysplasia cohort with no known association with cancer and in the SCC cohorts, 70% and 30%, respectively (Figures 1E, F and 4). Although the median number of recurrent copy number changes in this group was only modestly larger than in the dysplasia with no known association with cancer (median = 2.5 compared to median = 2), additional losses at 3p, 5q, 9p and 13q were present in ≥ 30% of cases.

5. Amplicons in oral dysplasia and SCC.

Dysplasia genomes harbored amplifications, defined as focal regions of higher level increased copy number. Previously, the laboratory had reported that oral SCCs typically amplify narrow regions of the genome (< 3 Mb) and identified 18 such recurrent amplicons (Snijders et al., 2005b). In the 29 dysplasia cases with no known association with cancer, we found two of these amplicons at 11q13 (*CCND1*, *PAK1*) and 20p12.2 (*JAG1*) to be present, as well as amplification at 2q11.2 in two dysplasia cases and two non-recurrent amplicons at 20q13.33 and 21q21.3 (Figure 1C and Table 5). The amplification at 21q21.3, however, spans a region that is gained in ≥15% of SCC cases and a likely driver gene for this amplicon is *MIR155*. Although the 2q11.2 amplicon had not been observed previously in SCC cohort#1 as a recurrent amplicon (Snijders et al., 2005b), it has been reported in an oral SCC cell line (Hermsen et al., 2005) and was recently reported by others in dysplasia (Garnis et al., 2009). The recurrent amplicons

are present in both 3q8pq20 and non-3q8pq20 dysplasia and SCC genomes (Figures 1-4), and thus their formation appears to be mediated by processes independent of those driving low level gains and losses.

6. Two routes to oral cancer.

Considered together the distribution of copy number aberrations in dysplasia and SCC suggest that there are two routes to oral cancer, one associated with greater genome instability and acquisition of +3q, -8p, +8q and/or +20 in pre-malignant stages and the other lacking chromosomal level instability that can be detected by CGH (Figure 5). Gains of 3q, 8q and 20 or losses of 8p appear to contribute to enhanced proliferation and/or disruption of differentiation that characterize the histopathologic diagnosis of dysplasia, whereas additional recurrent aberrations are necessary for tumor formation. Nevertheless, although these additional aberrations are rare in dysplasia, we found only the region +7pter-p11.2 (FDR adjusted p-value = 0.036) to be significantly more frequent in SCCs compared to the 39 dysplasia cases, suggesting that up-regulation of genes in this region including *EGFR* is a late event in tumor progression.

7. Association of 3q8pq20 status with clinical parameters

We investigated whether the 3q8pq20 and non-3q8pq20 subtypes differed in their clinical behavior. For this purpose, we assigned cases from cohort#2 to the 3q8pq20 subtype if one or more of the aberrant regions at 3q, 8p, 8q and 20 as defined by occurrence at

≥20% frequency in the dysplasia cohort with no association with cancer was present. Analysis of 3q8pq20 status and clinical characteristics in cohort#2 revealed a highly significant association of 3q8pq20 status with nodal status (2-sided Fisher's exact test $p<0.006$; odds ratio 11.5), *i.e*. neck metastasis was present in 46% (22/48) of 3q8pq20 tumors and in only 7% (1 of 15) of non-3q8pq20 tumors. We observed a more modest association with age, non-3q8pq20 tumors were more frequent in older patients, but there was no association with other features such as tobacco or alcohol use, or tumor thickness (Table 6). Considering cohort#2 as one group, *i.e*. not separating cases according to 3q8pq20 status, we observed that tumors in patients with positive nodes were significantly associated with increased numbers of whole chromosome copy number changes (p=0.04), fraction of the genome gained (p=0.007) and fraction of the genome altered (p=0.02). This observed enhanced genome instability associated with nodal status is consistent with and likely reflects the significantly higher metastatic rate of the 3q8pq20 tumor subtype with its higher frequency of copy number aberrations. Fraction of the genome gained and fraction of the genome altered were also associated with alcohol use (current vs. previous vs. ever, p=0.035, p=0.02, respectively) and tumor thickness (p=0.03, p=0.04, respectively). We also analyzed cohort#2 based on lymph node status and found a significant association between positive nodes and tumor size ≥ 2.7 cm (p=0.018) and tumor thickness ≥ 1.3cm (p=0.01) (Table 7). Other clinical parameters tested such as age, gender, tobacco and alcohol use etc. were not significant.

Figure 3. Non-uniform distribution of low level gains and losses among 3q8pq20 and non-3q8pq20 oral SCC cases.

(A and B) Frequency of copy number aberrations for each clone in non-3q8pq20 (A) and 3q8pq20 (B) cases in SCC cohort#1 as in Figures 1A and B (excluding regions on 3q, 8p, 8q and 20 used to assign cases to each group).

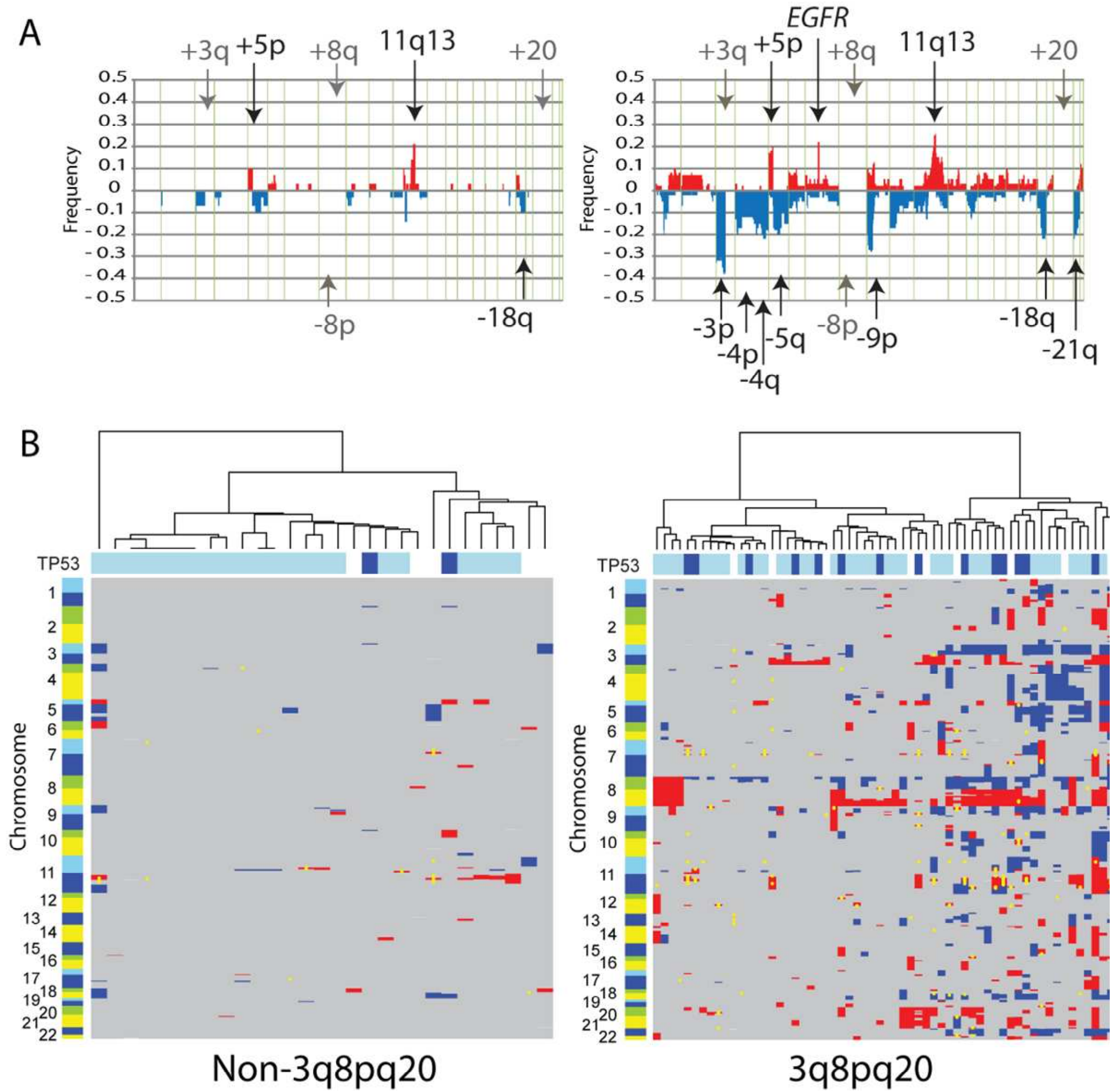


Figure 4. Copy number aberrations involving 3q, 8p, 8q and chromosome 20 are frequent in oral dysplasia with cancer.

(A) Frequency of aberrations plotted as in Figures 1A and B.

(B) Hierarchical clustering based on genome-wide DNA copy number profiles of oral dysplasia samples associated with a previous and/or subsequent cancer as in Figures 1C. Dysplasia grade and *TP53* mutation status are indicated as in Figure 1C. Cases with a previous cancer are indicated in light blue, a subsequent cancer in dark blue and both a previous and a subsequent cancer in pink.

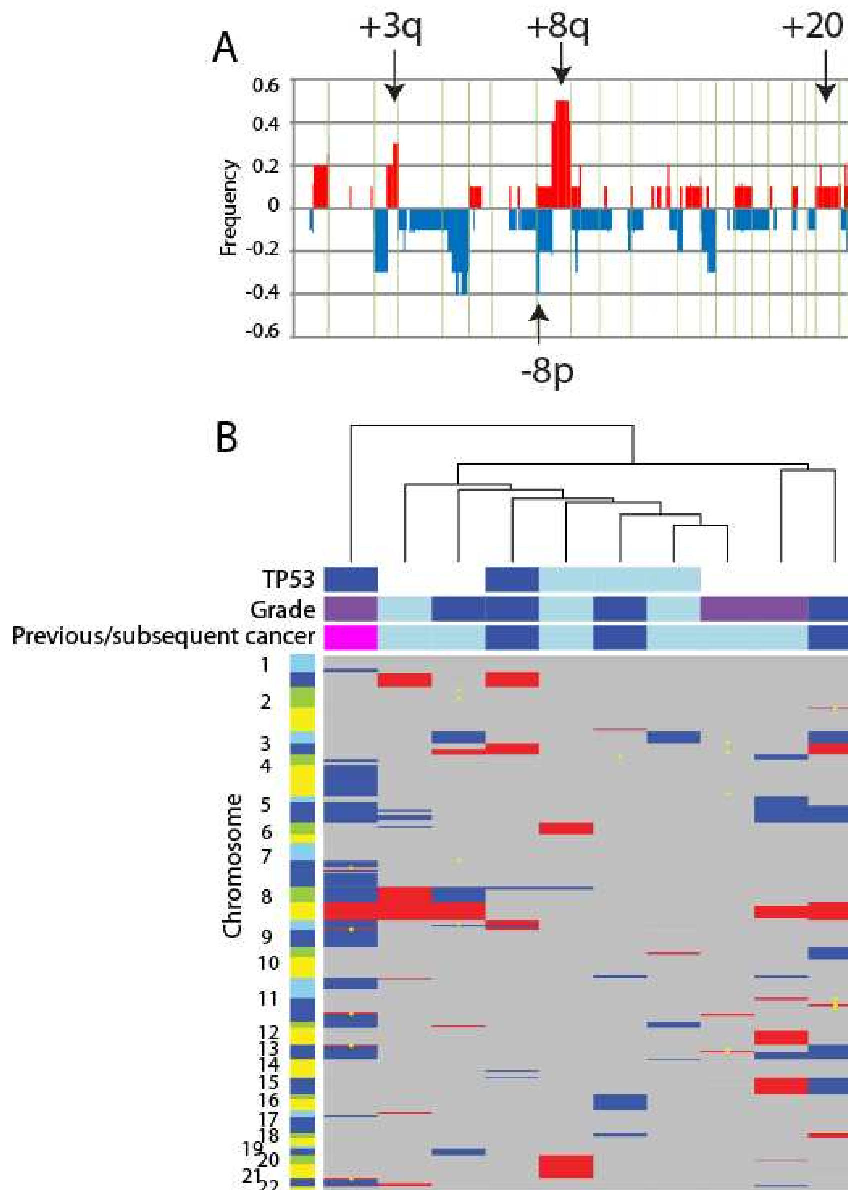
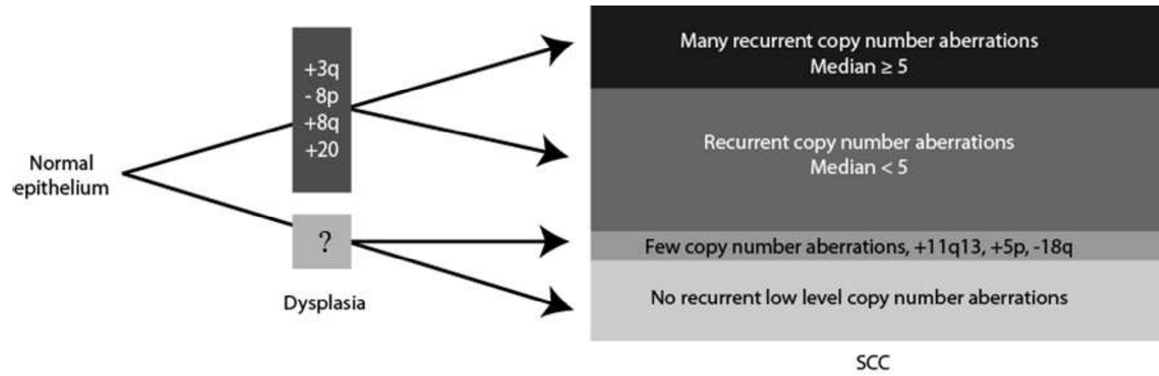


Figure 5: Two routes to cancer progression model

Possible origin and progression of dysplastic lesions to cancers are differentiated by acquisition of +3q, -8p, +8q and/or +20 in dysplasia, which subsequently progress to 3q8pq20 oral SCC. Other lesions lacking these aberrations progress to non-3q8pq20 SCC. The 3q8pq20 and non-3q8pq20 cancers may arise from different cell types, a stem cell vs. a transit amplifying cell, for example.



Other lines of investigation provide support for the existence of two different types of dysplasia and SCCs, which progress to cancer via different routes. For example, cytogenetic studies have found cultured oral SCCs with simple karyotypes (< 3 numerical and/or structural aberrations) in 36% of cases, while the remainder display complex karyotypes with mostly unbalanced structural rearrangements and numerical changes (Jin et al., 2006) consistent with the distribution of non-3q8pq20 and 3q8pq20, which would be expected to correspond to the simple and complex karyotype groups, respectively. Primary cultures of oral SCC and dysplasia have also been reported to separate into two classes – immortal and mortal (Hunter et al., 2006, Hunter et al., 2005) . The two groups differ in gene expression patterns and genomic instability, and it has been suggested that the two types may arise from different cell types.

8. Establishment of CAF and NF cell cultures.

We established 12 CAF and 25 NF cell lines from prospectively collected primary oral SCCs (Table 8). Characterization of these cultures is on-going (Table 9). The CAFs stained strongly positive for smooth muscle actin (SMA) and vimentin (10 CAF cultures studied). All CAF cultures except one were negative for cytokeratin (CK), which may indicate that the exceptional culture contained tumor cells undergoing epithelial to mesenchymal transition or had a small admixed population of tumor cells or keratinocytes. On the other hand NFs (17 studied) stained strongly positive for vimentin and only mildly positive for SMA. All NFs stained negative for CK. These findings are consistent with earlier reports regarding characterization of CAFs and NFs derived from

primary tumors (Lebret et al., 2007, Wang et al.). Because there is controversy over the genomic stability of CAFs (Zheng et al., 2009), we will carry out DNA copy number and allelic imbalance analyses of the CAF genomes, as well as mRNA and miRNA expression profiling to begin to investigate the phenotypes of oral CAFs. We anticipate the results will aid in clarifying and distinguishing the molecular characteristics of the tumor stroma in N0 and N+ tumors. Recent studies report that changes can be detected in the stroma of progressing lesions such as Barrett's esophagus (Gisselsson et al., 2005), suggesting that future studies may identify biomarkers of dysplasia progression or metastasis in the stroma.

Chapter 3: Conclusions and Future Directions

Conclusions

This thesis research has identified two routes to oral cancer in which the tumors are distinguished by the presence of aberrations involving chromosomes 3q, 8p, 8q and 20. Much attention has focused on regions of genomic imbalance, which are present at greater frequency in oral SCCs compared to pre-cancers, as biomarkers of progression (Bremmer et al., 2008). While such markers may be able to report on likely progression of 75-80% of lesions that are of the 3q8pq20 type, they cannot provide information on progression of the 20-25% of non-3q8pq20 lesions that progress to cancers with stable genomes. Further, the gene expression profiling studies of mortal/immortal dysplasia and SCC cultures (Hunter et al., 2006) suggest that routes to oral cancer differ not only in genomic instability, but also in pathways that are altered, which will impact response to therapy.

Importantly, the copy number status of chromosome 3q, 8p, 8q and 20, which defines the two oral SCC subtypes, is a potential biomarker for risk of associated neck metastasis. Management of the N0 neck presents a clinical dilemma. Because of the poor prognosis when neck metastasis is discovered after surgical resection of the tumor, prophylactic removal of the lymph nodes is recommended for clinically N0 patients. Identifying which of these patients have non-3q8pq20 tumors associated with low risk of metastasis could spare them the additional major surgery to remove the lymph nodes, as well as reducing morbidity and medical costs.

Early identification of oral SCCs with metastatic potential would profoundly improve clinical management of oral SCC. Oral SCCs that are likely to metastasize would be

treated with surgical resection of the primary SCC and a neck dissection to remove the regional lymphatics, followed by radiation therapy. Patients would then be followed by an intense tumor surveillance protocol to identify an early recurrence either in the oral cavity or the neck. On the other hand, oral SCCs with minimal or no metastatic potential could be managed with significantly less morbidity, decreased cost and improved quality of life.

Future directions

1. Validation of 3q8pq20 status as a biomarker of metastasis

The highly significant association of non 3q8pq20 tumors with lymph node negative tumors in this study has identified a possible subset of 'low risk' oral SCC that may have a lower risk for developing subsequent lymph node metastasis. However there was one node positive tumor in this subset of 15 tumors, and therefore the sensitivity and specificity of the 3q8pq20 biomarkers is yet to be determined. We have begun the process of identifying a new cohort (cohort#3) of oral SCCs to validate the accuracy of the above mentioned biomarker. Cohort#3 will be developed in an identical fashion to cohort#2 and array CGH will be performed. We are also trying to obtain the clinical and metastasis information for cohort#1 in order to test whether non-3q8pq20 status is significantly associated with node negative tumors in this dataset as well.

2. DNA methylation profiling of tumors lacking genomic aberrations

This study has identified two subsets of oral SCC and dysplasias: one harboring genomic aberrations and the other without, we are interested in further investigating the likely drivers of progression in the subset of tumors and dysplasias that lack genomic aberrations. DNA methylation (Ha, 2006) and microsatellite instability (MSI) (Koy 2008) are two well established modalities of tumor progression other than DNA copy number aberration. However MSI has a low prevalence in head and neck SCC (De Schutter etal., 2009). Moreover a recent array CGH analysis of oral SCC has identified a subset of oral SCCs lacking genomic aberrations that are associated with better survival (Smeets,

2009). Therefore investigation into methylation profiles of genomically silent tumors and dysplasias may reveal genetic pathways affected in this subset, eventually leading to biomarker identification to identify ‘good prognosis’ oral cancer patients.

3. Gene expression profiling of cohort#2

We are currently isolating RNA from oral SCC FFPE blocks to perform mRNA and microRNA expression profiling analyses for cohort#2. As discussed earlier, a number of gene expression ‘metastasis’ signatures have been identified in HNSCC (Chung et al., 2004, Roepman et al., 2006a, Roepman et al., 2006b), and we are in a position to develop such a signature specific to oral SCC. Our ultimate goal is to utilize multiple molecular techniques (array CGH, expression profiling, RT-PCR and immunohistochemistry) to develop a comprehensive biomarker of oral SCC metastasis for clinical use.

4. Stromal role in oral SCC progression and metastasis

Genomic and genetic characterization of CAF and NF cohorts: we are currently profiling the established cell cultures for allelic imbalance, DNA copy number alterations, mRNA and microRNA expression profiles in order to determine the differences between the CAFs and NFs in oral SCC. This will aid in clarifying the stromal distinctions between N0 and N+ tumors and ultimately aid in development of a stromal biomarker of metastasis.

5. Dysplasia and carcinoma in situ associated with oral SCC

We have identified 24 dysplasias/ carcinoma in situ (CIS) associated with the 63 oral SCCs with extensive clinical data (cohort#2). These are regions of dysplasia/CIS that are separated from the invasive carcinoma by $\geq$ 2mm. To address the question of whether dysplasias lacking genomic aberrations progress to oral SCCs similarly lacking copy number aberrations, and dysplasias having genomic aberrations progress to carcinomas with genomic aberrations. We will determine the genomic profiles of these dysplasia/CIS cases using array CGH. Copy number profiles of the dysplasias/CIS will be compared to the CGH profiles of their adjacent tumors to determine the similarities and differences between the two groups. Moreover as these are matched dysplasia/CIS and tumor samples, we anticipate that the genomic profiles will provide direct evidence whether dysplasias lacking genomic aberrations indeed do progress to similar tumors and vice versa.

Chapter 4: Materials and Methods

Identification of a Cohort of Metastatic and Non-metastatic Oral Squamous Cell Carcinomas (cohort #2): The oral cancer tissue bank at UCSF has been collecting and storing oral cancers and precancers diagnosed and/or treated at UCSF since 1995. The bank maintains a comprehensive database that includes clinical, demographic and follow-up data including risk factors such as tobacco usage and alcohol history. This tissue bank provides us with an excellent pre-existing cohort to study the genomic profiles of metastatic and non metastatic oral SCCs in a nested case control setting. This design affords the advantage of not having to enroll and subsequently track large populations of subjects over lengthy periods of time until they develop metastasis.

1. We conducted a search for all formalin fixed and paraffin embedded (FFPE) oral SCCs diagnosed from 1998 - 2005. We restricted the initial search to cases diagnosed from 1998 onwards as DNA quality of formalin fixed paraffin embedded (FFPE) tissue samples for genomic analysis was found to be optimum for approximately 7-8 years after formalin fixation. We then accessioned the detailed pathologic reports of all the oral SCCs identified thus far in order to obtain the corresponding demographic data: age and sex of patient, location, size and extent of tumor, clinical and pathologic staging, histological grading, tumor type (primary, second primary or recurrence) and distant organ involvement if any. We subsequently excluded all base of tongue, tonsil, oropharyngeal and nasal SCCs as we restricted our study to oral cavity sites only. We also searched for tumors of size $\geq$ 0.8 cm, as smaller tumors did not yield sufficient quantity of DNA for the genomic analysis (array CGH) protocol.

2. After our preliminary search of the UCSF Oral Cancer tissue bank we identified 2500 formalin fixed paraffin embedded (FFPE) oral SCCs. We defined non metastatic (N0) tumors as remaining metastasis free for at least 5 years after diagnosis, metastatic (N+) tumors as metastatic at diagnosis or developing metastasis within five years of diagnosis and tumors with unknown metastatic status (Nx). Review of the detailed pathology reports of these cases identified 57 non-metastatic tumors, 96 metastatic tumors and 95 tumors with unknown lymph node status based on our inclusion and exclusion criteria as described above. We next verified the lymph node status of these tumors with records from the UCSF Cancer Registry. We had *a priori* defined N0 tumors as those that remained non metastatic for at least 5 years after diagnosis – we found only 33 of our original 57 non metastatic tumors fitted this criterion. Nevertheless, the registry was able to confirm that 25 tumors in the unknown lymph node status group were actually non metastatic tumors and fitted our N0 criteria. Additionally we have identified 14 N0 and 7 N+ cases from Dr Brian Schmidt's recent patient records and 12 N0 and 21 N+ cases from the head and neck surgery clinic. We identified a total of 84 N0 tumors, 124 N+ tumors and 70 Nx tumors, tumors in the Nx group had to be discarded from the study as the lymph node status could not be ascertained. Our final tally of accessioned blocks was 71 N0 and 82 N+ blocks.

3. I verified tumor content of these tissue blocks by review of hematoxylin and eosin stained sections of each block together with a trained pathologist. We selected 40 N0 and 48N+ blocks which had sufficient areas of tumor and normal tissue for performing the array CGH protocol.

4. Tobacco and alcohol use data: We categorized tobacco use as current, previous and never smokers. Similarly alcohol use was categorized as current, previous and never alcohol drinkers. For some analyses we collapsed the categories of alcohol and tobacco users as ever smokers (current and previous users) and drinkers (current and previous) versus never smokers and never drinkers respectively.

We quantified tobacco use in pack years (number of packs per day X number of years smoked) where available. We could not quantify alcohol use as amount of alcohol consumed was not recorded for the majority of cases. We adopted the same search strategy in identifying the tobacco and alcohol use information as we did in identifying the lymph node status of the cohort. In many cases we were unable to obtain the tobacco and alcohol history of the cases from the pathology reports or the Cancer Registry data. We also searched the medical records of these patients as well as their clinical case records in the head and neck oncology clinics to determine this information. We were able to retrieve the tobacco use information for 80% of the tumor samples in our cohort, thus putting us in a position to investigate the association of tobacco with oral cancer and metastasis in this cohort.

DNA Isolation and Array CGH on Oral Squamous Cell Carcinoma Cohort

1. *Tissue sectioning and microdissection*: the de-identified tissue blocks selected for array CGH were sectioned into 25 serial sections 10 µm thick. The 1st, 13th and 25th sections from each block were stained with hematoxylin and eosin for review. I verified the location of the tumor and the tumor-stromal interface together with a trained pathologist.

The tumor tissue was outlined on the reverse surface of the glass slide with an indelible marker and scraped off with a scalpel under a dissection microscope ensuring that stromal tissue did not contaminate the sample of tumor tissue. The dissected tumor samples were transferred to labeled Eppendorf tubes for DNA isolation. We also dissected areas of normal non epithelial tissue such as muscle proximal to but not continuous with the tumor to obtain DNA from normal tissues

2. *DNA isolation*: DNA was isolated from the tumor samples using a protocol established by Snijders (2005) (Snijders et al., 2005b) . Briefly twenty 10µm tissue sections were collected in an Eppendorf tube mixed with xylene, centrifuged and supernatant discarded. This step was repeated twice and all remaining xylene was discarded. 100% ethanol was added, mixed on a nutator, centrifuged and supernatant discarded. The pellet was air dried; digestion buffer added and incubated overnight in a water bath (55° C). Proteinase K was added the next day and further incubated overnight in the same water bath, this step was repeated for two more days. Phenol chloroform isoamyl alcohol (PCI) was added to the proteinase K digested aqueous solution shaken and spun in the centrifuge. The upper aqueous phase was removed to a new Eppendorf tube; ammonium acetate, ethanol and glycogen were added and incubated at -20°C for a few hours. Finally the tubes were spun and centrifuged and supernatant discarded. The pellets were air dried, re-suspended in double distilled water and allowed to dissolve overnight. The DNA concentration was measured using a fluorometer.

3. *DNA labeling and array CGH hybridization*: DNA was random primed and labeled as described by Snijders et al. (2005). We incorporated Cy3 or Cy5 dCTP in a 25µL reaction to label genomic DNA. Labeled test (600ng) and reference (600ng) DNA was hybridized with human Cot-1 DNA for ~ 48 hours at 37°C to arrays of 2464 BAC clones spotted in triplicate (HumArray3.2, UCSF Array Core). The slides were washed with 50% formamide, 2X SSC for 15 minutes at 45°C followed by a 15 minute wash in PN buffer (0.1M phosphate buffer in 0.1% Nonidet P40) at room temperature. The slides were mounted in a DAPI solution to stain the array spots (90% glycerol, 10% PBS, 1µM DAPI) and imaged.

4. *Array Imaging and analysis*: we acquired DAPI, Cy3 and Cy5 images using a custom CCD camera system (Hamilton et al., 2006). Image and data analysis was done with UCSF SPOT software (Jain et al., 2002); SPROC software was used to filter data based on low DAPI intensity, low correlation between Cy3 and Cy5 within each segmented spot and low reference/DAPI signal intensity. Array clones with no replicate spots or those with standard deviation (SD) of the replicates greater than 0.2 were treated as missing observations. Clones were screened out if: data was missing in more than 15% of the samples, clones did not map on the genome sequence or they were known to harbor common copy number polymorphisms. We plotted the data in genome order as the mean $\log_2$ratio of the replicate spots for each clone normalized to the genome median $\log_2$ratio.

5. *Data processing for cohort#2 and dysplasia cohort*: Novel array data pre-processing techniques were used to improve the typically noisy CGH profiles obtained with DNA from FFPE tissue. For cohort#2, two CGH profiles were obtained, one from DNA extracted from the tumor and the other from adjacent non-epithelial tissue such as muscle. The normal profile was then used to correct for the systematic enhanced noise in the profiles. The method was also extended to correct a similar consistent noise pattern in a subset of the dyplasia CGH profiles. In this case, the data from these samples was clustered, revealing two different patterns. For each of these a template was calculated as the median log2 ratio per array probe across samples in each population. After appropriate scaling (equal to the dot product amplitude of the tumor profile on the template), the template was subtracted from the dysplasia or tumor profile. Copy number analysis followed published procedures (Snijders et al., 2005b) and used circular binary segmentation (Olshen et al., 2004). Statistical analysis of genomic aberrations with nodal status or among other clinical features used Fisher's exact test and log rank test as appropriate.

Cell Culture and Reagents

We prospectively identified oral SCC cases from Dr Brian Schmidt's clinical records and confirmed oral SCC diagnosis from the preliminary histopathologic reports. We obtained surgical discard material with all identifiers removed at the time of surgical excision or resection. Representative samples of tissue were obtained from the tumor proper and normal areas as identified by Dr Schmidt at the time of surgery. The samples were

transported in transport medium (Dulbecco's Modified Eagle Medium, DMEM Low Glucose (DME H-16) with 10% fetal calf serum for tumor tissue; Hank's buffered saline solution (HBSS) for normal tissue) supplemented with antibiotics (gentamicin (50mg/ml), fungizone (250ug/mL) and penicillin-streptomycin (100X). The tissue pieces were first transferred to a petri dish and blood clots and visible pieces of fat were trimmed off using scalpels. The tissues were then transferred to 15 ml Falcon tubes and washed thoroughly in 70% ethanol for 2 minutes followed by HBSS supplemented with antibiotics for 15 minutes.

1. Carcinoma associated fibroblast culture: after the HBSS wash the tumor tissue pieces were transferred to a fresh petri dish. The pieces were minced finely using crossed scalpels and placed in 6 well tissue culture plates in 1ml of DME H-16 with 10% fetal calf serum (FCS) supplemented with antibiotics. The cultures were left undisturbed for 7 days. When fibroblasts were seen growing out of the explants, medium was changed twice weekly until the cultures were ~70% confluent. Cells were passaged twice at 70% confluence and frozen for subsequent DNA and RNA extraction at passage 2.

2. Normal fibroblast culture: normal tissue was chopped into 5mm^2 pieces and incubated overnight in dispase grade II (Roche Diagnostics GmbH, Germany) at 4°C. The tissue pieces in dispase were then incubated at 37°C for 5-7 minutes following which the epithelium was teased off from the underlying connective tissue using a scalpel and forceps. Residual connective tissue was scraped on the surface to remove any adherent epithelium, cut into small pieces and placed in 3 ml of collagenase type 1A (Sigma-Aldrich, USA) solution (1 mg/ml in DMEM) supplemented with antibiotics. After

incubation overnight at room temperature, 20 ml of DME H-16 supplemented with 10% fetal calf serum was added and the mixture was pipetted up and down a number of times to further dissociate the cells. Fibroblasts were collected by centrifugation at 1500 rpm for 3 minutes in a 15 ml tube, and approximately 5-200µl of cells were placed in a T75 flask containing 10 ml of DME H-16 supplemented with 10% fetal calf serum supplemented with antibiotics.Medium was changed twice weekly and cells were passaged at ~70% confluence. Cells were frozen for subsequent DNA and RNA extraction at passage 2.

Immunofluorescence

Each NF and CAF cell culture established from primary oral squamous cell carcinomas was characterized using pan Cytokeratin (mouse anti-human AE1/AE3, Millipore, USA; mouse anti-human CAM 5.2, Becton-Dickinson, USA), alpha smooth muscle actin (α-SMA) (mouse anti-human SMA, Dako Denmark A/S) and Vimentin (rabbit anti-human vimentin, Abcam, USA) antibodies.

$1x10^4$ cells at passage 2 were plated in DME H-16 with 10% FCS on sterile glass cover slips in 6 well plates and allowed to attach overnight. Medium was aspirated and cells were washed with PBS for 3 minutes. Cells were fixed with acetone-methanol-formaldehyde (19:19:2) for 3 minutes and fixative was aspirated. Cells were washed twice with PBS Tween (1Liter PBS with 0.5 mL Tween 20) for 3 minutes each, and wash solution aspirated. Cells were incubated in 3% bovine serum albumin (BSA) for 30 minutes to block non specific binding sites following which BSA was aspirated off. Cells

were incubated in 50ul of primary antibody (1:100 dilution) at 4°C overnight. Primary antibody was washed off with 3 washes of PBS Tween with agitation for 5 minutes. Cells were incubated with 50 uL secondary antibody (1:500 dilution) (goat anti-mouse IgG FITC, Santa Cruz Biotechnology, USA; sheep anti-rabbit TR, Abcam) for 30 minutes in the dark. Secondary antibody was washed off with 3 washes of PBS Tween with agitation for 5 minutes. Cover slips with attached cells were mounted in Vectastain DAPI (Vector Labs, Burlingame, CA), transferred to slides and imaged.

All cell culture reagents were obtained from the UCSF Cell Culture Facility (San Francisco, CA) except where otherwise specified.

***TP53* sequencing**.

We amplified exons 5-8 of *TP53* from genomic DNA and carried out cycle sequencing, as described previously (Snijders et al., 2005b)

Tables

Table: 1 Clinical characteristics of cohort #2 (63 oral SCCs)

ID#	Final Node status	Age at Diagnosis	Sex	Site	Tumor size (cm)	Tumor thickness (cm)	Node status (clinical)	Histology*	Recurrence, Type	Total Follow-up (Months)	Vital Status	Tumor Status	Node status (path)	Alcohol	Tobacco
AB003	N0	70	M	Retromolar Region	1	1	0	MD	Local	20	DEAD	Unknown	0	Unknown	Previous
AB004	N0	68	M	Gingiva	5	1.7	X	MD	None	85	DEAD	Free	0	Past	Current
AB007	N+	48	M	Floor of Mouth	10	5.5	2c	WD	Regional	19	DEAD	Not free	2C	None	None
AB010	N0	61	M	Tongue	1.5	1.5	X	MD	None	126	ALIVE	Free	0	None	None
AB011	N0	47	M	Tongue	2	1.4	0	MD	None	86	ALIVE	Free	0	Current	Previous
AB014	N+	59	M	Retromolar Region	3.2	1.7	2	MD	None	34	ALIVE	Free	1	None	None
AB015	N+	68	F	Tongue	5.2	3.9	X	MD	Distant mets	22	DEAD	Not free	2C	None	None
AB016	N+	74	M	Buccal Mucosa	2.3	0.9	X	PD	None	18	DEAD	Unknown	2B	Past	Current
AB017	N+	59	M	Buccal Mucosa	5	NR	0	WD	None	5	DEAD	Free	2B	Current	Current
AB018	N+	50	M	Floor of Mouth	3.6	0.6	1	MD	None	60	ALIVE	Free	2C	Current	Previous
AB019	N+	60	F	Floor of Mouth, tongue	4.5	3.5	2c	MD	None	8	DEAD	Free	2C	Current	Current
AB020	N+	86	F	Hard Palate	0.9	NR	1	WD	None	4	DEAD	Not free	N1	None	None
AB021	N0	68	F	Tongue	3.2	NR	X	MD	Local	36	DEAD	Not free	0	Unknown	Unknown
AB022	N0	60	M	Floor of Mouth	0.9	0.6	X	MD	None	100	ALIVE	Free	0	Current	Current
AB023	N0	41	M	Tongue	3.2	2.2	0	WD	Local	80	ALIVE	Free	0	Current	Current
AB025	N0	69	F	Gingiva	2.4	0.2	0	MD	None	44	ALIVE	Free	0	Current	Unknown
AB026	N0	65	F	Retromolar Region	2	0.8	0	MD to PD	None	15	DEAD	Free	0	Current	Current
AB027	N0	80	M	tongue	4.4	NR	0	MD to PD	None	2	DEAD	Not free	0	None	None
AB029	N0	64	F	Floor of Mouth, tongue, buccal mucosa	3	0.8	2	PD	Local	105	ALIVE	Free	0	Current	Current
AB030	N0	58	M	Floor of Mouth	0.9	NR	X	MD	None	94	ALIVE	Free	0	None	None
AB031	N0	78	F	Tongue	5.2	NR	0	WD	None	2	DEAD	Free	0	Current	Current
AB032	N0	46	M	Buccal Mucosa	8	NR	NR	MD	None	15	DEAD	Unknown	0	Current	Current
AB033	N+	77	F	Retromolar Region	6.3	2.2	1	MD	Local	7	DEAD	Not free	2B	Current	None
AB034	N0	69	M	Buccal Mucosa	6.4	NR	0	WD	None	42	DEAD	Free	0	Unknown	Unknown
AB035	N0	83	M	Tongue	4.7	1.3	1	MD	None	48	DEAD	Free	0	Unknown	None
AB037	N+	64	M	Floor of Mouth	2.7	NR	NR	MD	None	27	DEAD	Free	1	None	Current
AB038	N+	66	M	Tongue	6	3.2	2C	MD	None	13	DEAD	Free	2	Past	Previous
AB039	N0	75	M	Gingiva	3	1.5	0	MD	Local	34	ALIVE	Not free	0	Past	Current
AB040	N+	49	F	Tongue	9.2	4.6	X	MD	Unknown	31	ALIVE	Free	2B	Current	Current
AB041	N0	51	M	Tongue	NR	NR	0	MD	None	97	ALIVE	Free	0	Current	None
AB042	N0	65	M	Tongue	3	1.5	0	MD	None	88	ALIVE	Free	0	Past	Previous
AB045	N0	76	M	Gingiva	2.7	1	2c	WD	None	37	DEAD	Free	0	None	None
AB047	N0	46	F	Buccal Mucosa	1	0.5	0	MD	Never	39	DEAD	Not free	0	Unknown	Previous
AB048	N0	46	M	Tongue	1.5	0.3	X	MD	None	79	ALIVE	Free	N0	None	None
AB049	N0	46	M	Tongue	1.1	0.7	X	MD	None	70	ALIVE	Free	0	Unknown	Unknown
AB051	N+	43	M	Tongue	3	0.9	0	MD	Local	32	DEAD	Not free	1	Current	Current

ID#	Final Node status	Age at Diagnosis	Sex	Site	Tumor size (cm)	Tumor thickness (cm)	Node status (clinical)	Histology*	Recurrence, Type	Total Follow-up (Months)	Vital Status	Tumor Status	Node status (path)	Alcohol	Tobacco
AB052	N+	51	M	Tongue	6	3.5	X	MD to WD	Local	15	DEAD	Not free	1	None	None
AB054	N+	76	F	Buccal Mucosa	4.2	2.9	X	PD	None	11	DEAD	Free	1	Current	None
AB055	N0->N+	57	M	Floor of Mouth	1.6	1.6	X	MD	Lymph node met	35	DEAD	Not free	0	Past	Current
AB056	N+	83	M	Retromolar Region	4	3.6	2b	MD	None	7	DEAD	Free	2B	Current	Current
AB059	N+	68	F	Tongue	1.2	0.35	0	MD	None	76	ALIVE	Free	1	Unknown	Unknown
AB060	N+	57	F	Tongue, Floor of Mouth	1.7	1.5	0	MD	Unknown	33	ALIVE	Free	2C	Current	Previous
AB061	N0	39	M	Buccal Mucosa	1.3	0.5	0	WD	Unknown	99	ALIVE	Unknown	0	Unknown	Current
AB062	N0	56	M	Gingiva	1	0.4	NR	MD	None	22	DEAD	Free	0	Current	Current
AB063	N0	49	M	Tongue	1	0.8	0	MD	None	74	ALIVE	Free	0	Current	Previous
AB064	N0	39	M	Buccal Mucosa	1.5	0.5	0	MD	None	66	DEAD	Free	0	None	Current
AB065	N0	85	F	Gingiva	2.5	1	0	MD	None	7	DEAD	Free	0	None	None
AB066	N0	57	M	Tongue	1.7	1.4	0	MD	None	56	ALIVE	Free	0	Current	None
AB067	N0	81	F	Floor of Mouth	8	NR	NR	MD to PD	None	2	DEAD	Free	0	None	Previous
AB068	N0	70	F	Gingiva	2	0.5	X	WD	Distant met	130	ALIVE	Not free	0	None	None
AB069	N0	56	M	Buccal Mucosa	1	0.6	NR	WD	None	79	ALIVE	Free	0	Unknown	Unknown
AB070	N0	71	M	Floor of Mouth	2.5	0.3	NR	WD	None	3	DEAD	Not free	0	Past	Previous
AB071	N0	90	F	Hard Palate	3.5	3	0	MD	None	1	DEAD	Free	0	Unknown	Unknown
AB073	N+	96	F	Gingiva	2.7	1.7	x		None	5	DEAD	Not free	2	Unknown	Unknown
AB076	N0	81	F	Gingiva	2.5	NR		MD	None	50	DEAD	Free	0	None	Previous
AB077	N0	61	F	Floor of Mouth	2.3	NR	0	MD	None	64	ALIVE	Free	0	Current	Current
AB079	N0	77	F	Tongue	4.8	1.4	2B	WD	None	8	DEAD	Free	0	None	None
AB080	N0	81	F	Tongue	3.3	2.4	0	PD	None	37	ALIVE	Free	0	Current	Previous
AB081	N+	83	F	Gingiva	4.4	1.8	X	MD	None	46	DEAD	Free	2C	Current	Previous
AB082	N+	70	M	Floor of Mouth	2.9	1.3	1	MD	None	31	DEAD	Free	1	Unknown	Unknown
AB083	N+	54	F	Floor of Mouth, tongue, gingiva	7	5	X	MD	None	13	DEAD	Not free	N1	Current	Current
AB084	N+	57	F	Gingiva	1.1	0.5	1	MD	None	60	ALIVE	Free	2B	None	None
AB085	N0	79	F	Tongue	0.8	0.8	X	MD	None	11	DEAD	Free	0	Unknown	Unknown
AB086	N0	77	M	Floor of Mouth	1.8	1.2	0	WD	None	43	ALIVE	Free	0	Current	Previous

* WD: well differentiated, PD: poorly differentiated, MD: moderately differentiated; NR: not reported

Table 2. Dysplasia cases

OCRC#	Patient ID	Grade	Site	Sex	Age	*TP53* Sequence exons 5-8	Prior Cancencer progress (months)	(months)	Amplification
5707	3297	mild	tongue	M	49	NA	Unknown	Unknown	
5724	2860	mild	tongue	F	53	no mutation	Unknown	Unknown	
5749	3329	severe	tongue	M	64	no mutation	Unknown	Unknown	
5769	3346	mild	tongue	M	82	exon 5, H179R (CAT > CGT)	Unknown	Unknown	
5779	3354	moderate	tongue	F	23	NA	Unknown	Unknown	2q11.2, 21q21.3
5807	3377	mild	tongue	M	52	exon 6, H193L (CAT > CTT)	Unknown	Unknown	
5824	2215	severe	tongue	M	62	no mutation (exons 5, 7, 8)	Unknown	Unknown	
5905	3436	moderate	tongue	M	42	NA	Unknown	Unknown	
5914	1921	moderate	tongue	M	56	no mutation	Unknown	Unknown	2q11.2
5952	3470	moderate	FOM	M	77	exon 6, I195S (ATC > AGT)	Unknown	Unknown	CCND1, PAK1
6162	3539	mild	tongue	F	66	no mutation	Unknown	Unknown	
6201	3665	moderate	gingiva	F	71	exon 8, P278T (CCT > ACT)	Unknown	Unknown	
6390	287	moderate	tongue	F	60	no mutation	Unknown	Unknown	JAG1
6402	3784	mild	tongue	M	57	no mutation	Unknown	Unknown	
6419	2734	severe	tongue	M	70	no mutation	Unknown	Unknown	
6427	3801	mild	tongue	F	73	exon 6, T211I (ACT > ATT)	Unknown	Unknown	
6463	3832	severe	tongue	M	50	no mutation	Unknown	Unknown	
6475	3839	severe	palate	F	52	no mutation	Unknown	Unknown	
6486	2578	mild	FOM	M	50	no mutation	Unknown	Unknown	
6686	3981	mild	uccal mucos	M	61	no mutation	Unknown	Unknown	
6689	3983	severe	uccal mucos	M	81	no mutation	Unknown	Unknown	
6690	3984	mild	tongue	M	60	no mutation	Unknown	Unknown	
6695	3989	mild	uccal mucos	M	41	exon 5, S127F (TCC > TTC)	Unknown	Unknown	
6756	4036	mild	tongue	F	50	no mutation	Unknown	Unknown	
7453	3467	moderate	tongue	M	56	no mutation	Unknown	Unknown	
7618	4865	severe	tongue	M	82	exon 5, H179R (CAT > AAT)	Unknown	Unknown	
7646	4622	severe	tongue	M	57	no mutation (exons 5, 7, 8)	Unknown	Unknown	
7678	4649	severe	uccal mucos	F	74	no mutation	Unknown	Unknown	
7694	4662	moderate	tongue	M	42	no mutation	Unknown	Unknown	
5653	3223	severe	tongue	F	81	NA	Yes, –1	Unknown	
5809	3332	mild	lower lip	M	64	no mutation	Yes, –2	Unknown	
8292	5097	severe	tongue	M	71	NA	Yes, –2	Unknown	

OCRC#	Patient ID	Grade	Site	Sex	Age	*TP53* Sequence exons 5-8	Prior Cancencer (months)	progress (months)	Amplification
8444	769	mild	tongue	M	51	NA	Yes, –348	Unknown	
6889	4127	Severe	Tongue	M	62	het. deleletion exon 5 (TTTGCCAACTGGCCAA)	Yes, ~140	Yes, +33	7q21.12, 7q21.12, 9p13.3, 11q22, 13q, 21q
149	74	moderate	tongue	F	83	NA	Unknown	Yes, +41	
5681	3271	moderate	gingiva	F	86	exon 8, R282W (CGG > TGG)	Unknown	Yes, +21	
5922	3450	mild	tongue	M	47	no mutation	Unknown	Yes, +49	
6367	2071	moderate	omolar trig	M	56	no mutation	Unknown	Yes, +3	
8417	5234	moderate	tongue	F	49	NA	Unknown	Yes, +20	2q11.2, CCND1, PAK1

FOM: floor of mouth

Table 3. Recurrent regions of aberration at $\geq 20\%$ frequency in dysplasia with no known association with cancer

Region	Aberration	Start kb	End kb	Proximal clone	Marker	Distal clone	Marker	Maximum Frequency
3q24-qter	Gain	146541.9	199505.74	RP11-72E23	AFM210VE7	GS1-56H22		0.41
8pter-p23.1	Loss	0.001	10893.274	GS1-77L23		RP11-252K12	SHGC-1962	0.34
8q12-q24.2	Gain	61264.08	134076.075	RP11-258B14	SHGC-32354	RP11-184M21	SHGC-1948	0.52
20pter-qter	Gain	0.001	62435.964	RP1-82O2		RP1-81F12		0.28

Table 4. Frequency in SCC cohorts and dysplasia associated with cancer of copy number changes occurring in ≥20% of dysplasia with no known association with cancer.

Aberration	Dysplasia[a] (n = 39)	Dysplasia w/cancer[b] (n = 10)	SCC cohort#1[c] (n = 89)	SCC cohort#2[d] (n = 63)
3q24-qter	0.41	0.3	0.25	0.46
8pter-p23.1	0.34	0.4	0.35	0.52
8q12-q24.2	0.52	0.5	0.37	0.56
20pter-qter	0.28	0.2	0.17	0.25

[a]Dysplasia with no known association with cancer

[b]Dysplasia associated with cancer

[c]SCC cohort#1, Snijders *et al* . (2005)

[d]SCC cohort#2, this study

Table 5. Amplicons in 29 dysplasia samples from patients with no known history of oral cancer

Dysplasia case no.	Cyto-Band	Size (Mb)	SCC[1] (%)	Proximal flanking clone	STS	Distal flanking clone	STS	Candidate oncogenes
5779, 5914	2q11.2	3.7	0%	RP11-327M19		RP11-629A22	AFMB355ZG1	*CIAO1, CNNM3*
5952	11q13.3	1.6	11%	CTD-2080I19	RH7839	RP11-120P20	SHGC-4518	*CCND1, EMS1*
5952	11q13.5	0.9	2%	CTC-352E23	RH52308	RP11-98G24	SHGC-31540	*PAK1*
6390	20p12.2	1.2	3%	RMC20P160	WI-7829	RMC20P178	D20S186	*JAG1*
6390	20q13.33	3.2	0%	RP11-94A18	AFM218XE7	RP11-358D14	X70940	*CDH4, PSMA7* *ADRM1, LAMA5,NTSR1, BIRC7*
5779	21q21.3	4.8	0%	RP11-86J21	AFMA081WF1	RP11-115H17	SHGC-11277	*MIR155*

[1]Frequency reported in oral SCC cohort#1 by Snijders *et al* . (2005)

Table 6. Patient and tumor characteristics relative to tumor subtype

	n	3q8pq20	non-3q8pq20	p-value[a]	Odds Ratio	95% confidence interval lower	upper
Nodal status	63	48	15	0.006	11.494	1.516	521.823
N0		26	14				
N+		22	1				
Age	63			0.018	0.199	0.032	0.87
<65		27	3				
≥65		21	12				
Tumor thickness	50			0.314	2.228	0.473	12.171
<1.3 cm		17	7				
≥1.3 cm		22	4				
Tobacco use	46						
never		9	3	0.355	2.443	0.302	17.653
ever		30	4				
Tobacco use excluding snuff	45			0.362	2.363	0.291	17.099
never		9	3				
ever		29	4				
Tobacco use	46			0.25	NA	NA	NA
never		9	3				
previous		11	3				
current		19	1				
Tobacco use excluding snuff	45			0.286	NA	NA	NA
never		9	3				
previous		11	3				
current		18	1				
Alcohol use	43			0.347	2.558	0.31	18.918
never		8	3				
ever		28	4				
Alcohol use	43			0.376	NA	NA	NA
never		8	3				
previous		7	0				
current		21	4				

[a] Two –sided Fisher's exact test

Table 7. Patient and tumor characteristics relative to cervical node status

	n	N0	N+	p-value	Odds Ratio	95% confidence interval lower	upper
3q8pq20 status	63			0.006	11.494	1.516	521.823
3q8pq20		26	22				
non-3q8pq20		1	14				
Age	63			0.611	1.327	0.422	4.225
<65		18	12				
≥65		22	11				
Gender	63			0.44	1.517	0.475	4.879
Female		15	11				
Male		25	12				
Tumor size	62			0.018	0.251	0.066	0.852
<2.7 cm		23	6				
≥2.7 cm		16	17				
Tumor thickness	50			0.01	0.2	0.044	0.78
<1.3 cm		19	5				
≥1.3 cm		11	15				
Tobacco use	46			1	0.918	0.167	4.365
never		8	4				
ever		22	12				
Tobacco use excluding snuff	45			1	1	0.18	4.826
never		8	4				
ever		22	11				
Tobacco use	46						
never		8	4	0.923	NA	NA	NA
previous		10	4				
current		12	8				
Tobacco use excluding snuff	45			0.922	NA	NA	NA
never		8	4				
previous		10	4				
current		12	7				
Alcohol use	43			0.494	0.556	0.08	2.91
never		8	3				
ever		19	13				
Alcohol use	43			0.751	NA	NA	NA
never		8	3				
previous		4	3				
current		15	10				
Site	57			0.496	NA	NA	NA
buccal mucosa		6	3				
floor of mouth		6	5				
gingiva		8	3				
retromolar region		2	3				
tongue		16	5				

Table 8. Clinical features of established cell cultures

Cell Type	ID	Age	Sex	Site	Donor node status
CAF	OF52	56	M	Tongue	N0
CAF	OF77	60	F	Tongue	N+
CAF	OF81	66	F	Tongue	NR
CAF	OF83	72	F	Tongue	N+
CAF	OF99	65	F	Tongue	N0
CAF	OF68	84	F	Gingiva	N0
CAF	OF6 +	63	M	Gingiva	N0
CAF	HN112	71	F	Gingiva	N+
CAF	OF79	62	M	Gingiva	N+
CAF	OF108	58	F	Gingiva	N0
CAF	OF112	85	F	Tongue	N0
CAF	OF97	84	F	Gingiva	N0
NF	OF51	56	M	Tongue	N0
NF	OF78	60	F	Tongue	N+
NF	OF82	66	F	Tongue	NR
NF	OF84	72	F	Tongue	N+
NF	OF100	65	F	Tongue	N0
NF	OF1	43	M	Tongue	NA
NF	OF6i	63	M	Tongue	N0
NF	OF8	37	M	Tongue	NA
NF	OF59	24	M	Tongue	NR
NF	OF66	84	F	Tongue	N0
NF	OF6ii	63	M	Gingiva	NO
NF	OF40	71	F	Gingiva	N+
NF	OF50	62	M	Gingiva	N+
NF	OF49	84	F	Gingiva	N0
NF	OF76	60	F	Gingiva	N+
NF	OF47	26	M	Gingiva	NA
NF	OF44	34	F	Gingiva	NA
NF	OF2	27	M	Maxilla	NA
NF	OF35	29	F	Maxilla	NA
NF	OF36	17	M	BM	NA
NF	OF37	37	M	Mandible	NA
NF	OF42	85	M	Maxilla	NA
NF	OF48	47	M	BM	NA
NF	OF48A	NR	NR	BM	NA
NF	OF8	37	M	Tongue	NA
NF	OF57+	83	M	FOM	NA
NF	OF58+	83	M	BM	NA
NF	OF61	62	M	FOM	NA
NF	OF63	78	M	Gingiva	NA
NF	OF64	78	M	BM	NA

Cell Type	ID	Age	Sex	Site	Donor node status
NF	OF74	76	F	BM	NA
NF	OF75	60	F	BM	N+
NF	OF80	63	M	BM	N+
NF	OF88	66	F	Tongue	NR
NF	OF98	85	F	BM	NR
NF	OF102	66	M	BM	N0
NF	OF104	53	F	BM	N+
NF	OF105	64	M	Mandible	NA
NF	OF106	64	M	Maxilla	NA
NF	OF107	61	M	Tongue	NA
NF	OF110	72	M	Mandible	NA
NF	OF111	72	M	BM	NA
NF	OF115	54	M	Tongue	N+

CAF: carcinoma associated fibroblasts; NF: normal fibroblasts ; NA: not associated with cancer; NR: not reported, BM: buccal mucosa, FOM: floor of mouth

Table 9. Cell culture characterization

Cell Type	ID	Age	Sex	Site	Node status	Immunofluoresence, %+ α SMA	Vimentin	Pan CK
CAF	OF52	56	M	TONGUE	N0	50-75	100	0
CAF	OF77	60	F	TONGUE	N+	5	100	0
CAF	OF81	66	F	TONGUE	NR	25-40	100	0
CAF	OF83	72	F	TONGUE	N+	25-40	100	0
CAF	OF99	65	F	TONGUE	N0	40	100	0
CAF	OF68	84	F	GINGIVA	N0	not done	not done	not done
CAF	OF6 +	63	M	GINGIVA	N0	20-40	100	0
CAF	HN112	71	F	GINGIVA	N+	30	100	0
CAF	OF79	62	M	GINGIVA	N+	25-30	>90	0
CAF	OF108	58	F	GINGIVA	N0	not done	not done	not done
CAF	OF112	85	F	TONGUE	N0	not done	not done	not done
CAF	OF97	84	F	GINGIVA	N0	75	100	50
NF	OF51	56	M	TONGUE	N0	0-5	100	0
NF	OF78	60	F	TONGUE	N+	2	100	0
NF	OF82	66	F	TONGUE	NR	0	100	0
NF	OF84	72	F	TONGUE	N+	0	0	0
NF	OF100	65	F	TONGUE	N0	5	100	0
NF	OF01	43	M	TONGUE	NA	1	100	0
NF	OF6i	63	M	TONGUE	N0	5-10	100	<1
NF	OF8	37	M	TONGUE	NA	25	100	0
NF	OF59*	24	M	TONGUE	NR	0	100	0
NF	OF66*	84	F	TONGUE	N0	not done	not done	not done
NF	OF6ii	63	M	GINGIVA	NO	5	100	0
NF	OF40	71	F	GINGIVA	N+	1	100	0
NF	OF50	62	M	GINGIVA	N+	0	100	0
NF	OF49	84	F	GINGIVA	N0	0	100	0
NF	OF76*	60	F	GINGIVA	N+	2	100	0
NF	OF47	26	M	GINGIVA	NA	2	100	0
NF	OF44	34	F	GINGIVA	NA	15	100	0

α-SMA: α-smooth muscle actin Pan CK: pan cytokeratin; CAF: carcinoma associated fibroblasts ; NF: normal fibroblasts ; NA: not associated with cancer; %+: percent positive

References

AKERVALL, J. (2006) Genomic screening of head and neck cancer and its implications for therapy planning. *Eur Arch Otorhinolaryngol,* 263**,** 297-304.

BAATENBURG DE JONG, R. J., GULLANE, P. G. & FREEMAN, J. L. (1993) The significance of false-positive nodes in patients with oral cancer. *J Otolaryngol,* 22**,** 154-9.

BLOT, W. J., MCLAUGHLIN, J. K., WINN, D. M., AUSTIN, D. F., GREENBERG, R. S., PRESTON-MARTIN, S., BERNSTEIN, L., SCHOENBERG, J. B., STEMHAGEN, A. & FRAUMENI, J. F., JR. (1988) Smoking and drinking in relation to oral and pharyngeal cancer. *Cancer Res,* 48**,** 3282-7.

BREMMER, J. F., BRAAKHUIS, B. J., BRINK, A., BROECKAERT, M. A., BELIEN, J. A., MEIJER, G. A., KUIK, D. J., LEEMANS, C. R., BLOEMENA, E., VAN DER WAAL, I. & BRAKENHOFF, R. H. (2008) Comparative evaluation of genetic assays to identify oral pre-cancerous fields. *J Oral Pathol Med,* 37**,** 599-606.

BROWN, L. F., GUIDI, A. J., SCHNITT, S. J., VAN DE WATER, L., IRUELA-ARISPE, M. L., YEO, T. K., TOGNAZZI, K. & DVORAK, H. F. (1999) Vascular stroma formation in carcinoma in situ, invasive carcinoma, and metastatic carcinoma of the breast. *Clin Cancer Res,* 5**,** 1041-56.

CALIFANO, J., VAN DER RIET, P., WESTRA, W., NAWROZ, H., CLAYMAN, G., PIANTADOSI, S., CORIO, R., LEE, D., GREENBERG, B., KOCH, W. & SIDRANSKY, D. (1996) Genetic progression model for head and neck cancer: implications for field cancerization. *Cancer Res,* 56**,** 2488-92.

CHAMBERS, A. F., NAUMOV, G. N., VANTYGHEM, S. A. & TUCK, A. B. (2000) Molecular biology of breast cancer metastasis. Clinical implications of experimental studies on metastatic inefficiency. *Breast Cancer Res,* 2**,** 400-7.

CHUNG, C. H., PARKER, J. S., KARACA, G., WU, J., FUNKHOUSER, W. K., MOORE, D., BUTTERFOSS, D., XIANG, D., ZANATION, A., YIN, X., SHOCKLEY, W. W., WEISSLER, M. C., DRESSLER, L. G., SHORES, C. G., YARBROUGH, W. G. & PEROU, C. M. (2004) Molecular classification of head and neck squamous cell carcinomas using patterns of gene expression. *Cancer Cell,* 5**,** 489-500.

DOBROSSY, L. (2005) Epidemiology of head and neck cancer: magnitude of the problem. *Cancer Metastasis Rev,* 24**,** 9-17.

FAKHRY, C., WESTRA, W. H., LI, S., CMELAK, A., RIDGE, J. A., PINTO, H., FORASTIERE, A. & GILLISON, M. L. (2008) Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. *J Natl Cancer Inst,* 100**,** 261-9.

FERLITO, A., SHAHA, A. R., SILVER, C. E., RINALDO, A. & MONDIN, V. (2001) Incidence and sites of distant metastases from head and neck cancer. *ORL J Otorhinolaryngol Relat Spec,* 63**,** 202-7.

GARNIS, C., BALDWIN, C., ZHANG, L., ROSIN, M. P. & LAM, W. L. (2003) Use of complete coverage array comparative genomic hybridization to define copy number alterations on chromosome 3p in oral squamous cell carcinomas. *Cancer Res,* 63**,** 8582-5.

GARNIS, C., CHARI, R., BUYS, T. P., ZHANG, L., NG, R. T., ROSIN, M. P. & LAM, W. L. (2009) Genomic imbalances in precancerous tissues signal oral cancer risk. *Mol Cancer,* 8**,** 50.

GENDEN, E. M., FERLITO, A., SHAHA, A. R., TALMI, Y. P., ROBBINS, K. T., RHYS-EVANS, P. H. & RINALDO, A. (2003) Complications of neck dissection. *Acta Otolaryngol,* 123**,** 795-801.

GISSELSSON, D., LV, M., TSAO, S. W., MAN, C., JIN, C., HOGLUND, M., KWONG, Y. L. & JIN, Y. (2005) Telomere-mediated mitotic disturbances in immortalized ovarian

epithelial cells reproduce chromosomal losses and breakpoints from ovarian carcinoma. *Genes Chromosomes Cancer,* 42**,** 22-33.
HAMILTON, G., BROWN, N., OSEROFF, V., HUEY, B., SEGRAVES, R., SUDAR, D., KUMLER, J., ALBERTSON, D. & PINKEL, D. (2006) A large field CCD system for quantitative imaging of microarrays. *Nucleic Acids Res,* 34**,** e58.
HERMSEN, M., SNIJDERS, A., GUERVOS, M. A., TAENZER, S., KOERNER, U., BAAK, J., PINKEL, D., ALBERTSON, D., VAN DIEST, P., MEIJER, G. & SCHROCK, E. (2005) Centromeric chromosomal translocations show tissue-specific differences between squamous cell carcinomas and adenocarcinomas. *Oncogene,* 24**,** 1571-9.
HUNTER, K. D., PARKINSON, E. K. & HARRISON, P. R. (2005) Profiling early head and neck cancer. *Nat Rev Cancer,* 5**,** 127-35.
HUNTER, K. D., THURLOW, J. K., FLEMING, J., DRAKE, P. J., VASS, J. K., KALNA, G., HIGHAM, D. J., HERZYK, P., MACDONALD, D. G., PARKINSON, E. K. & HARRISON, P. R. (2006) Divergent routes to oral cancer. *Cancer Res,* 66**,** 7405-7413.
IACOBUZIO-DONAHUE, C., ARGANI, P., HEMPEN, P., JONES, J. & KERN, S. (2002a) The Desmoplastic Response to Infiltrating Breast Carcinoma: Gene Expression at the Site of Primary Invasion and Implications for Comparisons between Tumor Types. *Cancer Res,* 62**,** 5351-5357.
IACOBUZIO-DONAHUE, C. A., MAITRA, A., SHEN-ONG, G. L., VAN HEEK, T., ASHFAQ, R., MEYER, R., WALTER, K., BERG, K., HOLLINGSWORTH, M. A., CAMERON, J. L., YEO, C. J., KERN, S. E., GOGGINS, M. & HRUBAN, R. H. (2002b) Discovery of novel tumor markers of pancreatic cancer using global gene expression technology. *Am J Pathol,* 160**,** 1239-1249.
IARC (2004) Tobacco smoke and involuntary smoking. *IARC Monogr Eval Carcinog Risks Hum,* 83**,** 1-1438.
JAIN, A. N., TOKUYASU, T. A., SNIJDERS, A. M., SEGRAVES, R., ALBERTSON, D. G. & PINKEL, D. (2002) Fully automatic quantification of microarray image data. *Genome Res,* 12**,** 325-32.
JEMAL, A., SIEGEL, R., WARD, E., HAO, Y., XU, J., MURRAY, T. & THUN, M. J. (2008) Cancer statistics, 2008. *CA Cancer J Clin,* 58**,** 71-96.
JIN, C., JIN, Y., WENNERBERG, J., ANNERTZ, K., ENOKSSON, J. & MERTENS, F. (2006) Cytogenetic abnormalities in 106 oral squamous cell carcinomas. *Cancer Genet Cytogenet,* 164**,** 44-53.
KONSKI, A. & WATKINS-BRUNER, D. (2004) The RTOG Outcomes Model: economic end points and measures. *Expert Opin Pharmacother,* 5**,** 513-9.
KOWALSKI, L. P. & SANABRIA, A. (2007) Elective neck dissection in oral carcinoma: a critical review of the evidence. *Acta Otorhinolaryngol Ital,* 27**,** 113-7.
KOYAMA, H., KOBAYASHI, N., HARADA, M., TAKEOKA, M., KAWAI, Y., SANO, K., FUJIMORI, M., AMANO, J., OHHASHI, T., KANNAGI, R., KIMATA, K., TANIGUCHI, S. & ITANO, N. (2008) Significance of tumor-associated stroma in promotion of intratumoral lymphangiogenesis: pivotal role of a hyaluronan-rich tumor microenvironment. *Am J Pathol,* 172**,** 179-93.
LAGACE, R., GRIMAUD, J. A., SCHURCH, W. & SEEMAYER, T. A. (1985) Myofibroblastic stromal reaction in carcinoma of the breast: variations of collagenous matrix and structural glycoproteins. *Virchows Arch A Pathol Anat Histopathol,* 408**,** 49-59.
LEBRET, S. C., NEWGREEN, D. F., THOMPSON, E. W. & ACKLAND, M. L. (2007) Induction of epithelial to mesenchymal transition in PMC42-LA human breast carcinoma cells by carcinoma-associated fibroblast secreted factors. *Breast Cancer Res,* 9**,** R19.
LIU, HU, SHEN, LI, SF, LIN, JW, ZHENG, XH, GAO, QH, ZHOU & HM (2006a) Separation, cultivation and biological characteristics of oral carcinoma-associated fibroblasts. *Oral Diseases,* 12**,** 375-380.

LIU, C. J., LIN, S. C., CHEN, Y. J., CHANG, K. M. & CHANG, K. W. (2006b) Array Comparative Genomic Hybridization to Detect Genomewide changes in Microdissected Primary and Metastatic Oral Squamous Cell Carcinoma. *Molecular Carcinogenesis,* 45**,** 721-731.

MAO, L., LEE, J. S., FAN, Y. H., RO, J. Y., BATSAKIS, J. G., LIPPMAN, S., HITTELMAN, W. & HONG, W. K. (1996) Frequent microsatellite alterations at chromosomes 9p21 and 3p14 in oral premalignant lesions and their value in cancer risk assessment. *Nat Med,* 2**,** 682-5.

MASAYESVA, B. G., HA, P., GARRETT-MAYER, E., PILKINGTON, T., MAO, R., PEVSNER, J., SPEED, T., BENOIT, N., MOON, C. S., SIDRANSKY, D., WESTRA, W. H. & CALIFANO, J. (2004) Gene expression alterations over large chromosomal regions in cancers include multiple genes unrelated to malignant progression. *Proc Natl Acad Sci U S A,* 101**,** 8715-20.

MENDEZ, E., FAN, W., CHOI, P., AGOFF, S. N., WHIPPLE, M., FARWELL, D. G., FUTRAN, N. D., WEYMULLER, E. A., JR., ZHAO, L. P. & CHEN, C. (2007) Tumor-specific genetic expression profile of metastatic oral squamous cell carcinoma. *Head Neck,* 29**,** 803-14.

MIGNOGNA, M. D., FEDELE, S. & LO RUSSO, L. (2004) The World Cancer Report and the burden of oral cancer. *Eur J Cancer Prev,* 13**,** 139-42.

MIGNOGNA, M. D., LO MUZIO, L., LO RUSSO, L., FEDELE, S., RUOPPO, E. & BUCCI, E. (2001) Metastases in small thickness oral squamous-cell carcinoma arising in oral lichen planus. *Med Oncol,* 18**,** 159-63.

MORK, J., LIE, A. K., GLATTRE, E., HALLMANS, G., JELLUM, E., KOSKELA, P., MOLLER, B., PUKKALA, E., SCHILLER, J. T., YOUNGMAN, L., LEHTINEN, M. & DILLNER, J. (2001) Human papillomavirus infection as a risk factor for squamous-cell carcinoma of the head and neck. *N Engl J Med,* 344**,** 1125-31.

MUELLER, M. M. & FUSENIG, N. E. (2004) Friends or foes - bipolar effects of the tumour stroma in cancer. *Nat Rev Cancer,* 4**,** 839-49.

NAGATA, M., FUJITA, H., IDA, H., HOSHINA, H., INOUE, T., SEKI, Y., OHNISHI, M., OHYAMA, T., SHINGAKI, S., KAJI, M., SAKU, T. & TAKAGI, R. (2003) Identification of potential biomarkers of lymph node metastasis in oral squamous cell carcinoma by cDNA microarray analysis. *International Journal of Cancer,* 106**,** 683-689.

NGUYEN, S. T., HASEGAWA, S., TSUDA, H., TOMIOKA, H., USHIJIMA, M., NODA, M., OMURA, K. & MIKI, Y. (2007) Identification of a predictive gene expression signature of cervical lymph node metastasis in oral squamous cell carcinoma. *Cancer Sci,* 98**,** 740-6.

O'BRIEN, C. J., SMITH, J. W., SOONG, S. J., URIST, M. M. & MADDOX, W. A. (1986) Neck dissection with and without radiotherapy: prognostic factors, patterns of recurrence, and survival. *Am J Surg,* 152**,** 456-63.

O'DONNELL, R., KUPFERMAN, M., WEI, J., SINGHAL, S., WEBER, R., O'MALLEY, B., CHENG, Y., PUTT, M., FELDMAN, M., ZIOBER, B. & MUSCHEL, R. (2004) Gene expression signature predicts lymphatic metastasis in squamous cell carcinoma of the oral cavity. *Oncogene,* 24**,** 1244-1251.

OLSHEN, A. B., VENKATRAMAN, E. S., LUCITO, R. & WIGLER, M. (2004) Circular binary segmentation for the analysis of array-based DNA copy number data. *Biostatistics,* 5**,** 557-72.

ORIMO, A., GUPTA, P. B., SGROI, D. C., ARENZANA-SEISDEDOS, F., DELAUNAY, T., NAEEM, R., CAREY, V. J., RICHARDSON, A. L. & WEINBERG, R. A. (2005) Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell,* 121**,** 335-48.

PATMORE, H. S., ASHMAN, J. N., CAWKWELL, L., MACDONALD, A., STAFFORD, N. D. & GREENMAN, J. (2004) Can a genetic signature for metastatic head and neck squamous cell carcinoma be characterised by comparative genomic hybridisation? *British Journal of Cancer,* 90**,** 1976-1982.

PURI, S. K., FAN, C. Y. & HANNA, E. (2003) Significance of extracapsular lymph node metastases in patients with head and neck squamous cell carcinoma. *Curr Opin Otolaryngol Head Neck Surg,* 11**,** 119-23.

REMMLER, D., BYERS, R., SCHEETZ, J., SHELL, B., WHITE, G., ZIMMERMAN, S. & GOEPFERT, H. (1986) A prospective study of shoulder disability resulting from radical and modified neck dissections. *Head Neck Surg,* 8**,** 280-6.

RITCHIE, J. M., SMITH, E. M., SUMMERSGILL, K. F., HOFFMAN, H. T., WANG, D., KLUSSMANN, J. P., TUREK, L. P. & HAUGEN, T. H. (2003) Human papillomavirus infection as a prognostic factor in carcinomas of the oral cavity and oropharynx. *Int J Cancer,* 104**,** 336-44.

ROEPMAN, P., DE JAGER, A., GROOT KOERKAMP, M. J., KUMMER, J. A., SLOOTWEG, P. J. & HOLSTEGE, F. C. (2006a) Maintenance of head and neck tumor gene expression profiles upon lymph node metastasis. *Cancer Res,* 66**,** 11110-4.

ROEPMAN, P., DE KONING, E., VAN LEENEN, D., DE WEGER, R., KUMMER, A., SLOOTWEG, P. & HOLSTEGE, F. (2006b) Dissection of a metastatic gene expression signature into distinct components. *Genome Biology,* 7**,** R117.

ROEPMAN, P., WESSELS, L. F., KETTELARIJ, N., KEMMEREN, P., MILES, A. J., LIJNZAAD, P., TILANUS, M. G., KOOLE, R., HORDIJK, G. J., VAN DER VLIET, P. C., REINDERS, M. J., SLOOTWEG, P. J. & HOLSTEGE, F. C. (2005) An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas. *Nat Genet,* 37**,** 182-6.

ROSENTHAL, E., MCCRORY, A., TALBERT, M., YOUNG, G., MURPHY-ULLRICH, J. & GLADSON, C. (2004a) Elevated expression of TGF-beta1 in head and neck cancer-associated fibroblasts. *Mol Carcinog,* 40**,** 116-21.

ROSENTHAL, E. L., MCCRORY, A., TALBERT, M., CARROLL, W., MAGNUSON, J. S. & PETERS, G. E. (2004b) Expression of proteolytic enzymes in head and neck cancer-associated fibroblasts. *Arch Otolaryngol Head Neck Surg,* 130**,** 943-7.

ROSIN, M. P., CHENG, X., POH, C., LAM, W. L., HUANG, Y., LOVAS, J., BEREAN, K., EPSTEIN, J. B., PRIDDY, R., LE, N. D. & ZHANG, L. (2000) Use of allelic loss to predict malignant risk for low-grade oral epithelial dysplasia. *Clin Cancer Res,* 6**,** 357-62.

SCHMALBACH, C. E., CHEPEHA, D. B., GIORDANO, T. J., RUBIN, M. A., TEKNOS, T. N., BRADFORD, C. R., WOLF, G. T., KUICK, R., MISEK, D. E., TRASK, D. K. & HANASH, S. (2004) Molecular profiling and the identification of genes associated with metastatic oral cavity/pharynx squamous cell carcinoma. *Arch Otolaryngol Head Neck Surg,* 130**,** 295-302.

SHAH, J. P. (1990) Patterns of cervical lymph node metastasis from squamous carcinomas of the upper aerodigestive tract. *Am J Surg,* 160**,** 405-9.

SHIBOSKI, C. H., SCHMIDT, B. L. & JORDAN, R. C. (2005) Tongue and tonsil carcinoma: increasing trends in the U.S. population ages 20-44 years. *Cancer,* 103**,** 1843-9.

SIDRANSKY, D. (1995) Molecular genetics of head and neck cancer. *Curr Opin Oncol,* 7**,** 229-33.

SMEETS, S. J., BRAKENHOFF, R. H., YLSTRA, B., VAN WIERINGEN, W. N., VAN DE WIEL, M. A., LEEMANS, C. R. & BRAAKHUIS, B. J. (2009) Genetic classification of oral and oropharyngeal carcinomas identifies subgroups with a different prognosis. *Cell Oncol,* 31**,** 291-300.

SMITH, E. M., RITCHIE, J. M., SUMMERSGILL, K. F., KLUSSMANN, J. P., LEE, J. H., WANG, D., HAUGEN, T. H. & TUREK, L. P. (2004) Age, sexual behavior and human

papillomavirus infection in oral cavity and oropharyngeal cancers. *Int J Cancer,* 108**,** 766-72.

SNIJDERS, A. M., SCHMIDT, B. L., FRIDLYAND, J., DEKKER, N., PINKEL, D., JORDAN, R. C. & ALBERTSON, D. G. (2005a) Oncogene - Abstract of article: Rare amplicons implicate frequent deregulation of cell fate specification pathways in oral squamous cell carcinoma. *Oncogene,* 24**,** 4232-4242.

SNIJDERS, A. M., SCHMIDT, B. L., FRIDLYAND, J., DEKKER, N., PINKEL, D., JORDAN, R. C. & ALBERTSON, D. G. (2005b) Rare amplicons implicate frequent deregulation of cell fate specification pathways in oral squamous cell carcinoma. *Oncogene,* 24**,** 4232-42.

SPIRO, R. H., HUVOS, A. G., WONG, G. Y., SPIRO, J. D., GNECCO, C. A. & STRONG, E. W. (1986) Predictive value of tumor thickness in squamous carcinoma confined to the tongue and floor of the mouth. *Am J Surg,* 152**,** 345-50.

SUGERMAN, P. B. & SHILLITOE, E. J. (1997) The high risk human papillomaviruses and oral cancer: evidence for and against a causal relationship. *Oral Dis,* 3**,** 130-47.

TIMAR, J., CSUKA, O., REMENAR, E., REPASSY, G. & KASLER, M. (2005) Progression of head and neck squamous cell cancer. *Cancer Metastasis Rev,* 24**,** 107-27.

TUXHORN, J. A., AYALA, G. E., SMITH, M. J., SMITH, V. C., DANG, T. D. & ROWLEY, D. R. (2002) Reactive stroma in human prostate cancer: induction of myofibroblast phenotype and extracellular matrix remodeling. *Clin Cancer Res,* 8**,** 2912-23.

VAN DER RIET, P., NAWROZ, H., HRUBAN, R. H., CORIO, R., TOKINO, K., KOCH, W. & SIDRANSKY, D. (1994) Frequent loss of chromosome 9p21-22 early in head and neck cancer progression. *Cancer Res,* 54**,** 1156-8.

WANG, J., YING, G., WANG, J., JUNG, Y., LU, J., ZHU, J., PIENTA, K. J. & TAICHMAN, R. S. Characterization of phosphoglycerate kinase-1 expression of stromal cells derived from tumor microenvironment in prostate cancer progression. *Cancer Res,* 70**,** 471-80.

WEIGELT, B., HU, Z., HE, X., LIVASY, C., CAREY, L. A., EWEND, M. G., GLAS, A. M., PEROU, C. M. & VAN'T VEER, L. J. (2005) Molecular portraits and 70-gene prognosis signature are preserved throughout the metastatic process of breast cancer. *Cancer Res,* 65**,** 9155-8.

ZHENG, Y. L., HU, N., SUN, Q., WANG, C. & TAYLOR, P. R. (2009) Telomere attrition in cancer cells and telomere length in tumor stroma cells predict chromosome instability in esophageal squamous cell carcinoma: a genome-wide analysis. *Cancer Res,* 69**,** 1604-14.

Publishing Agreement

It is the policy of the University to encourage the distribution of all theses, dissertations, and manuscripts. Copies of all UCSF theses, dissertations, and manuscripts will be routed to the library via the Graduate Division. The library will make all theses, dissertations, and manuscripts accessible to the public and will preserve these to the best of their abilities, in perpetuity.

Please sign the following statement:

I hereby grant permission to the Graduate Division of the University of California, San Francisco to release copies of my thesis, dissertation, or manuscript to the Campus Library to provide access and preservation, in whole or in part, in perpetuity.

Aditi Bhattacharya
Author Signature

6/8/10
Date