TOBACCO CARCINOGEN-INDUCED 3p14.2 ALTERATIONS IN EXFOLIATED CELLS COLLECTED FROM THE ORAL CAVITY OF SMOKERS

by

Vicki Maria Fleming
BSc., Simon Fraser University, 2002

PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ENVIRONMENTAL TOXICOLOGY

In the Department of
Biological Sciences

SIMON FRASER UNIVERSITY

Spring 2005

APPROVAL

Name: Vicki Maria Fleming
Degree: Masters of Environmental Toxicology

Title of Project: Tobacco carcinogen-induced 3p14.2 alterations in exfoliated cells collected from the oral cavity of smokers

Examining Committee:

Chair: **Dr. Lawrence Albright**
Professor/ Department of Biological Sciences

Dr. Miriam Rosin
Senior Supervisor
Professor/ Department of Kinesiology

Dr. Margo Moore
Associate Professor / Department of Biological Sciences

Dr. Chris Kennedy
Associate Professor / Department of Biological Sciences

Dr. Russell Nicholson
External Examiner
Associate Professor / Department of Biological Sciences

Date Approved: February 23, 2005

SIMON FRASER UNIVERSITY

PARTIAL COPYRIGHT LICENCE

Simon Fraser University

Ethics Approval

The author, whose name appears on the title page of this work, has obtained human research ethics approval from the Simon Fraser University Office of Research Ethics for the research described in this work, or has conducted the research as a member of a project or course approved by the Ethics Office.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for ethics approval and letter of approval is filed with the Office of Research Ethics. Inquiries may be directed to that Office.

Bennett Library
Simon Fraser University
Burnaby, BC, Canada

ABSTRACT

Oral cancer is usually diagnosed late in the development of the disease when prognosis is poor. This study explored the feasibility of using Fluorescence *in situ* hybridization (FISH) analysis of exfoliated cells collected from smokers, non-smokers and oral squamous cell carcinoma (OSCC) patients to identify infrequent, but critical alterations to tissue that might predict cancer risk. We focused on the Fragile Histidine Triad (FHIT) locus located on 3p14.2 as it is commonly altered early in the development of the disease. Five signal patterns observed in the tumor margins were not found in samples from non-smokers; of these patterns, 3 were found in smokers. In addition, 16.5% of smokers showed an elevated number of cells with alterations to the FHIT locus. In conclusion, the data showed that FHIT alterations are present in exfoliated cells of smokers, and specific patterns and frequencies of such alterations could be used in screening smokers to identify early changes associated with cancer risk.

DEDICATION

To my mother Annette for all her support and encouragement, and for always showing me that hard work and perseverance lead to personal success.

ACKNOWLEDGEMENTS

I would like to extend my greatest thanks and appreciation to my senior supervisor, Dr. Miriam Rosin for challenging me to rise to both personal and professional challenges, and for giving me the opportunity to expand my knowledge in the multidisciplinary area of cancer research. Dr. Rosin's remarkable insight and dedication to her efforts in cancer prevention are truly admirable, and inspirational. I would also like to thank Dr. Lewei Zhang for her continued support throughout my research project, Dr. Margo Moore for her knowledge of toxicology, and Dr. Chris Kennedy for his helpful guidance throughout my time in the MET program. I also extend my thanks to my colleagues in the Cancer Prevention Lab at S.F.U. for all their support during the last two years. In addition, I thank Dr. Soon and staff of the Abacus Dental Centre for their support during the sample collection period of the study.

Finally, a heartfelt thank-you to my family and Steve for their endless support and encouragement, and who have shared in all my frustrations and achievements.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

AA	Aromatic Amines
ADH	Alcohol Dehydrogenase
BaP	Benzo[a]pyrene
BCCA	British Columbia Cancer Agency
BPDE	Benzo[a]pyrene Diol Epoxide
CDK	Cyclin Dependant Kinase
CIS	Carcinoma *in situ*
EGFR	Epidermal Growth Factor Receptor
FHIT	Fragile Histidine Triad Gene
FISH	Fluorescence *in situ* Hybridization
GST	Glutathione-S-Transferase
HNSC	Head and Neck Squamous Cell Carcinoma
HPV	Human Papilloma Virus
LOH	Loss of Heterozygosity
NER	Nucleotide Excision Repair
OPL	Oral Premalignant Lesion
OR	Odds Ratio
OSCC	Oral Squamous Cell Carcinoma
PAH	Polycyclic Aromatic Hydrocarbon
PML	Premalignant Lesion
SCC	Squamous Cell Carcinoma
SNP	Single Nucleotide Polymorphism
TSG	Tumor Suppressor Gene
TSNA	Tobacco-Specific N-Nitrosamines
UV	Ultraviolet
XP	Xeroderma Pigmentosum

1. INTRODUCTION

1.1. Overview

Oral Squamous Cell Carcinoma (OSCC) is the sixth most common malignancy worldwide and is a major health problem resulting in widespread morbidity and mortality. Globally, about 300,000 new cases of oral and pharyngeal cancers are diagnosed annually, three quarters of them in developing countries such as India and Southeast Asia (Forastiere *et al*, 2001). In Canada and the United States, oral cancer accounts for 3% of all new cancers (Greenlee *et al*, 2001); however, in developing countries such as India, oral cancers represent up to 40% of all new cancers (Saranath *et al*, 1992). Tobacco exposure and alcohol consumption have been identified as independent risk factors of oral cancer (Znaor *et al*, 2003, Dal Maso *et al*, 2002, Zhu *et al*, 2002, De Flora *et al*, 2003), with a synergistic effect on risk associated with exposure to both agents. Differences in type and duration of these exposures can partially explain the different rates of oral cancers observed between various countries (Andre *et al*, 1995).

Despite improvements in surgical, chemotherapeutic, and radiation therapies over the last five decades, the prognosis for oral cancer patients has not improved and remains poor with the 5-year survival rate not exceeding 50% (Franceschi *et al*, 1993, Vokes *et al*, 1993). This is one of the lowest survival rates among human cancers and can mainly be attributed to detection of the disease at a late stage, and a high frequency of recurrence (20-30%) after

treatment of the primary tumor (Kowalski *et al*, 1993, Close *et al*, 1989).

These unfavorable statistics have led to an increased interest in developing new approaches to prevent the development of the disease through the identification of individuals at an elevated risk. Current strategies to improve the prognosis of OSCC rely heavily on the identification and appropriate management of high-risk oral premalignant lesions (OPLs) before they progress into invasive cancer.

Early OPLs are primarily identified by dentists during a dental examination where they appear as either white or red patches located in the oral cavity. The majority of these lesions are benign in nature and will not progress to cancer (Lumerman *et al*, 1995), but a small percentage (5-10%) will progress (Liu *et al*, 1998). Unfortunately, there is currently no way to discriminate between those OPLs that will progress into cancer and their benign counterparts (Neville *et al*, 2002). This inability to accurately identify high-risk premalignant lesions is a major barrier to cancer prevention, since it results in a lack of consensus on which lesions should be treated and in what fashion. Surgical removal of early OPLs might be an effective prevention strategy; however, removing all OPLs (many of which have little or no risk) would not only unnecessarily overload the health care system, but also result in undue stress to the patient and functional compromise to the oral cavity (e.g., could affect the patient's ability to speak or swallow).

Presently, the histology of biopsies taken from OPLs is used as the "Gold Standard" for risk prediction. Unfortunately, although this approach is well recognized to be a good predictor of risk for late-stage OPLs (histologically classified as severe dysplasia or carcinoma *in situ*), it has only a limited predictive value for the earliest lesions (those with minimal or no histological dysplasia). Most OPLs that are biopsied are categorized at this latter stage (Pindborg *et al.*, WHO, 1968).

Recent approaches to risk prediction have incorporated molecular markers in an effort to identify tissue with genetic alterations that are associated with an increased risk of malignant transformation (Grandis *et al*, 1993, Overgaard *et al*, 1998, Raybaud-Diogene *et al*, 1996, Lane *et al*, 1992). One approach currently being evaluated involves the use of Fluorescence *in situ* Hybridization (FISH). This procedure uses Fluorescence probes to hybridize to specific chromosomal regions within interphase cells, allowing one to identify small foci of cells with altered copy numbers of targeted regions. Copy number gain and loss is strongly associated with cancer development (Sudbo *et al*, 2001, Rosin *et al*, 2000, Mao *et al*, 1996, Zhang *et al*, 2001). Thus this approach may play a significant role as a risk predictor for early stages of the disease, even in the absence of clinical lesions.

The focus of this study was to investigate the potential application of FISH to identify cells in exfoliated samples from the oral cavities of smokers (with no

clinical lesions) that have copy number alterations to a specific gene, *Fragile Histidine Triad* (*FHIT*). This gene, located at 3p14.2, is known to be lost early in the molecular development of oral cancer and may play a significant role in the genesis of the disease (Virgilio *et al*, 1996, Kuroki *et al*, 2003). The results obtained in the study support this possibility and also suggest that the use of FISH on exfoliated cell samples might be a valuable tool to assist clinicians in the identification of individuals at risk at a very early stage, greatly improving the chance of a successful intervention.

1.2. Etiology of Oral Cancer

It is generally accepted that cancer develops as a result of the interaction of numerous factors, some environmental and others genetic. Tobacco exposure, alcohol consumption, papilloma virus infection, diet, and the presence of specific genetic polymorphisms have all been identified as risk factors for the development of oral cancer (Zavras *et al*, 2001, Znaor *et al*, 2003). The potential roles of each of these factors will be discussed below.

1.2.1. Tobacco and Oral Cancer

The first epidemiological study on the effects of smoking on human health was completed in 1950 when smoking was documented as an inducer of lung cancer in a dose-response relationship (Doll *et al*, 1950). Since then smoking has been shown to be associated with increased risk of several diseases, including chronic obstructive disease and cancers of the lung, upper aerodigestive system,

bladder, cervix, and nasal cavity (Ellard *et al*, 1995, Hoffman *et al*, 1995, Doll *et al*, 1950). Despite this knowledge, millions continue to smoke, thus increasing their risk for a multitude of tobacco related diseases.

Tobacco smoke is composed of over 4000 different constituents including some that are toxic, mutagenic, or carcinogenic. More than 50 carcinogens have been identified, consisting mainly of polycyclic aromatic hydrocarbons (PAH's), aromatic amines (AA), N-nitrosamines, and tobacco specific N-nitrosamines (TSNA) (Hoffman *et al*, 1997). Cigarette smoke also contains various promoters that can enhance the carcinogenic properties of cigarette smoke and promote the development of the disease (Hoffman *et al*, 1997).

The mutagenicity of cigarette smoke varies with several factors including the use of filters, pH, tobacco species, nitrate concentration, and method of tobacco harvesting (Hoffman *et al*, 1997). Cigarette smoke mutagenicity has also been shown to be dependant on the temperature of pyrolysis (White *et al*, 2001). Little or no tumorigenic activity is observed in tobacco smoke condensates if the tobacco is pyrolyzed below 400°C within *in vitro* systems; the mutagenicity level of tobacco pyrolysates thereafter has been found to increase with increasing temperature (White *et al*, 2001).

Of the numerous carcinogenic compounds found in cigarette smoke, nine compounds have been classified by IARC as Group 1 human carcinogens: benzene, cadmium, arsenic, vinyl chloride, 4-aminobiphenyl, nickel, chromium, 2-naphthylamine, and beryllium (Smith *et al*, 2000). Group 1 carcinogens are

those for which strong evidence of human carcinogenicity has been documented. Nine additional compounds in cigarette smoke have been classified as Group 2A (suspected to be probable human carcinogens) and 40 more as Group 2B carcinogens (possible human carcinogens) (Smith *et al*, 2001). These known and suspected carcinogens interact with each other (and other components) resulting in synergistic, additive, or subtractive effects on the carcinogenicity that are poorly understood (Smith *et al*, 2000).

1.2.2. Alcohol and Oral Cancer

Alcohol consumption has been identified as an independent risk factor in the development of oral cancer (Znaor *et al*, 2003, Andre *et al*, 1995), although the mechanism by which it acts is not well understood. It is not clear whether a metabolite of ethanol such as acetaldehyde is responsible (Zavras *et al*, 2002). Nor is the mechanism by which alcohol interacts with tobacco to increase cancer risk understood, although it is speculated that it may act as a co-carcinogen or promoter of tobacco-induced carcinogenesis (Bouchardy *et al*, 2000). Alcohol may increase the carcinogenic effect of cigarette smoke by creating a hydrophobic medium covering the surfaces of the oral cavity providing an ideal solvent for the hydrophobic carcinogens found in tobacco smoke, thus increasing the exposure of oral cavity surfaces to tobacco carcinogens and elevating the risk of tobacco-induced carcinogenesis (Zavras *et al*, 2002). In addition, the pooling of saliva on the floor of the oral cavity may increase the exposure time of alcohol, and dissolved tobacco carcinogens to this area, thus resulting in an increased prevalence of OSCC developing at the ventrolateral tongue and floor

of mouth. Several epidemiological studies have demonstrated a synergistic effect between smoking and drinking and the risk of oral cancer (Andre *et al*, 1995, Znaor et *al*, 2003).

Different drinking habits also affect the carcinogenic effect of alcohol. All three forms of alcohol (beer, wine, and hard liquor) have been associated with oral cancer, however the consumption of alcoholic beverages with higher ethanol concentrations (40% ethanol) appears to have a greater effect on risk compared with the consumption of the other 2 types of alcoholic beverages which have lower ethanol concentrations (Zavras *et al*, 2001). In addition, the pattern of alcohol consumption can affect the associated risk of oral cancer development. Consuming alcohol outside of meals or in an episodic fashion has a greater risk than the consumption of alcohol during mealtimes (Dal Maso *et al*, 2002).

1.2.3. Human Papilloma Virus Infection and Oral Cancer

Extensive epidemiologic and laboratory evidence exists in support of the involvement of Human Papilloma Virus (HPV) in the etiology of oral cancers. For developed Western countries, HPV-DNA has been detected in many head and neck tumor types including oral cancers, with the frequency of the infection dependent upon the site. Thus, approximately 18% of cancers of the oral cavity carry HPV-DNA compared with 50-70% of cancers in the oropharyngeal region (back of the tongue and tonsillar area (Gillison *et al,* 1999; Gillison and Shah, 2001). The prevalence of infection also appears to be dependent on the geographic local of the patient. For example, HPV prevalence in oral cancers is

much higher in southern India, with up to 67% of oral cancers positive for HPV-DNA. Healthy controls from that country display an HPV infection rate of 27% (Fenech *et al*, 2002).

Two strains of HPV have been associated with increased cancer risk, HPV 16 and HPV 18. These HPV types produce viral proteins E6 and E7 which inactivate two critical control proteins in the cell: *p53* and retinoblastoma (*Rb*) respectively (Herrero *et al*, 2003, Nagpal *et al*, 2002). *p53* plays a central role in the control of cell proliferation, DNA repair, apoptosis and angiogenesis, and the *Rb* protein controls the entry of cells into the cell cycle (Alberts *et al*, 1994, Munger *et al*, 1989). Removal of these key negative regulatory activities is felt to accelerate the progression of cells to cancer.

1.2.4. Diet and Oral Cancer

Several epidemiological studies have demonstrated an association between diet and risk of developing oral cancer (Petridou *et al*, 2002, Soler *et al*, 2001, Franceschi *et al*, 2001). Most reports describe a protective effect against oral cancer for consumption of fruits and to some extent vegetables with increased risk associated with a diet high in meat products (Petridou *et al*, 2002). These associations have been reported for a variety of cancers including colon, esophageal, breast, prostate, and lung (Michels *et al*, 2004, Franceschi *et al*, 2001, Cho *et al*, 2003, Brouwer *et al*, 2004, Bostick *et al*, 1994). Of interest, attention has recently focused specifically on oral cancers in Greece, as the incidence or oral cancer in this population is surprisingly low given reported

alcohol intake and smoking habits. The Greek diet is high in cereals with starchy roots, olive oil, and fruits, which may help explain the unexpected low incidence (Petridou *et al,* 2002). High fiber intake, characteristic of this population, has also shown to be protective against oral cancer (Petridou *et al,* 2002, Soler *et al,* 2001). These different food components are thought to work through diverse mechanisms to reduce the action of carcinogens including physical binding to carcinogens (or activated electrophiles) to reduce DNA damage and alteration to the metabolism of carcinogens (Soler *et al*, 2001).

1.3 Chemoprevention of Oral Cancer

The genetic heterogeneity, aggressive behavior, and invasiveness of cancers make these diseases extremely difficult to cure and explain why no significant increase in survival rates has occurred for many cancers over the past five decades. Therefore, instead of focusing on cancer treatment, new approaches for cancer management are now targeting the precursors of invasive cancer to prevent or delay the onset of the clinical disease using chemopreventive agents (Lippman *et al*, 2002). Oral premalignant lesions are established risk factors of oral cancer since most cancers develop after their appearance. Thus treatment of high-risk PMLs may have a large impact in preventing oral cancer development (Neville *et al*, 2002). Chemopreventive agents are compounds that can prevent, inhibit, or reverse carcinogenesis by intervention with chemical substances that can be administered as individual

drugs or as naturally occurring substances within the diet (Rinaldi *et al*, 2002).

One of the first chemopreventive agents shown to be effective in preventing cancer was tamoxifen, used in the reduction of breast cancer risk (Lippman *et al*, 1999). Tamoxifen is an estrogen-receptor modulator that competitively binds to estrogen receptors (O'Shaughnessy *et al*, 2002). Binding of tamoxifen inhibits the binding of estrogen to its receptors and therefore prevents the proliferative effect which estrogen has on the cell. The short-term use of tamoxifen in patients at high-risk of developing breast cancer has been shown to reduce the risk of breast cancer by ~ 49% (Hong *et al*, 2000). The ability of tamoxifen to prevent breast cancer over long durations (greater than 10 years) has not been established because of the serious side effects of the drug, but animal models have shown that the protective effect of tamoxifen wears off when treatment is stopped (Lippman *et al*, 1999). Tamoxifen therefore is not a cure for breast cancer; it only prevents the onset of the disease through the inhibition of molecular pathways required for cancer progression to occur. Once this proliferative pathway is restored, cancer progression can continue (Jordan *et al*, 1989). The administration of tamoxifen as a chemopreventative agent demonstrates the ability of chemical substances to prevent cancer development through involvement within pathways known to be involved in cancer development. Chemoprevention may not be a cure for cancer, but it has the potential to delay the onset of malignancy.

Due to the enormous potential of chemopreventative compounds on the incidence of cancer, numerous new compounds are under study in animal

models and clinical settings that are aimed at investigating their ability to target specific molecular pathways involved in carcinogenesis (Stoner *et al*, 1997). Many of these compounds are synthetic; however, some of these compounds are naturally occurring. Two examples of synthetic chemopreventative agents that are currently being evaluated for ability to prevent progression of oral premalignant lesions to cancer are celecoxib, a COX-2 inhibitor, and a monoclonal antibody targeted towards the Epidermal Growth Factor Receptor (EGFR). The latter was designed to bind to and inhibit the proliferative signaling transduction pathway induced by the activation of EGFR (Sudbo *et al*, 2003).

One naturally-occurring compound that is currently being investigated for its chemopreventative effects for oral cancer is curcumin. Curcumin is a major antioxidant and anti-inflammatory agent found in the spice tumeric and has been demonstrated to have anti-carcinogenic activities in animal models (Rinaldi *et al*, 2002). One of the major PAHs found in cigarette smoke is benzo[a]pyrene (BaP). This compound, like many other xenobiotics found in cigarette smoke, requires bioactivation before it can cause genotoxic damage. BaP is metabolized by Cyp1A1 which is a Phase I enzyme abundantly expressed in the oral cavity (Rinaldi *et al,* 2002). Curcumin has been demonstrated to reduce the bioactivation of BaP to its reactive metabolite benzo[a]pyrene diol epoxide (BPDE). Thus curcumin could potentially reduce the risk of genotoxic damage in the oral cavity induced by BPDE damage (Rinaldi *et al*, 2002). Curcumin chemoprevention could be effective in smokers who refuse to quit by decreasing the rate of BaP bioactivation, thus presumably decreasing the amount of DNA

adducts formed from those reactive metabolites. A second compound demonstrated to have a chemopreventative effect for oral cancer development is the dietary flavenoid morin (Kawabata *et al,* 1999). In animal models, morin has been identified as an inducer of the phase II detoxifying enzyme GST where the overall result of morin administration was a powerful chemopreventative effect on the induction of tongue carcinogenesis in rats (Kawabata *et al*, 1999).

1.4. Histological Progression of Oral Premalignant Lesions

Premalignancy is defined as "morphologically altered tissue in which cancer is more likely to occur than its apparently normal counterpart" (Kreyberg *et al*, 1978). Clinically, premalignant lesions appear as either leukoplakia or erythroplakia.

Leukoplakia, the more common form, are generally defined as a white patch that can not be rubbed off and can not be diagnosed clinically or pathologically as any other disease (WHO, 1978). Leukoplakia are categorized by appearance into 2 main types, homogeneous and non-homogenous. Homogenous leukoplakia appear homogenously white, flat, and thin, with a smooth surface. Such lesions are associated with a low risk of malignant transformation (Neville *et al*, 2002). In contrast, non-homogenous leukoplakia, the less common form (~10% of all leukoplakias) has a speckled red and white appearance, often with a verrucous, nodular composition that may have papilla extending above the lesion's surface. Generally this form of leukoplakia have a

higher risk of malignant transformation that its homogenous counterpart (Pindborg *et al*, 1968).

Erythroplakia are more rare than leukoplakia. Such lesions appear as homogenous bright red plaques that cannot be characterized clinically or pathologically as being due to any other condition (WHO, 1978). Unfortunately, diagnosis for OPLs is frequently derived by exclusion of other entities. It is often difficult to discriminate leukoplakia from reactive non-premalignant hyperplasic lesions that also present as clinically white lesions. Also, erythroplakia can be indistinguishable from inflammatory lesions which also present as red plaques (Neville *et al*, 2002). The gold standard for determining risk for OPLs requires a biopsy and evaluation by a pathologist for the presence and degree of dysplasia, a term used to describe the histological changes associated with increased cancer risk.

The World Health Organization has established criteria for dysplasia. They include the following: loss of basal cell polarity, the presence of more than one layer of basaloid cells, increased nuclear to cytoplasmic ratio, drop-shaped rete-ridges, irregular stratification, increase and/or abnormal mitosis, the presence of mitotic figures in the superficial half of the epithelium, cellular pleomorphism, nuclear hyperchromatism, enlarged nuclei, and reduction in cellular cohesion (WHO, 1978). The degree of dysplasia is dependent upon the severity of cellular change and the extent of tissue involvement in the biopsy. Lesions identified as mildly dysplastic display changes that are confined to the lower third of the epithelium. Moderately dysplastic lesions exhibit a greater

degree of dysplastic changes that penetrate further into the epithelium involving 2/3 of the tissue. Severe dysplastic lesions contain an even greater severity of dysplastic changes found within the upper third of the epithelium. The most severe form of OPL, before invasive cancer develops, is known as "carcinoma *in situ*" (*CIS*). In this case, the dysplastic changes are found throughout the entire epithelium, from the basal lamina to the tissue surface. The risk of an OPL progressing to cancer increases with the severity of the dysplasia found within the tissue biopsy (Silverman *et al*, 1984). OSCC develops when the dysplastic cells invade through the basement membrane into the underlying supporting structures and enter the blood and lymphatic systems as metastases.

1.5. Molecular Progression Model Of OSCC

The first molecular progression model for cancer was developed by Vogelstein *et al* (1988). This model recognized that at least 4 different genetic events were required for colorectal cancer development, with these alterations underlying the change in histology as a tissue went from normal through premalignant lesions to cancer. Since then progression models have been proposed for other sites, including the head and neck, where it is thought that at least 6 – 12 genetic events are required for cancer development (Lippman *et al*, 2001, Forastiere *et al*, 2001).

The process of acquisition of these genetic changes is thought to involve sequential processes of mutation and clonal expansion. The initial event is

considered to involve a critical mutation in a stem cell that provides it with a growth advantage over neighboring cells. This results in outgrowth of a cluster of cells (termed clonal expansion) each containing the mutation that provided its predecessor its growth advantage. Initially all cells in the clone are genetically identical. However, subsequent mutations and clonal expansion generates a heterogeneous cell mass, with the eventual production of an aggressive cell cluster that is capable of invasion and metastasis. This phenomenon is defined in Nowell's Hypothesis, which states that all tumor cells arise from a monoclonal origin, however due to genetic instability, random mutations are accumulated in subsequent daughter cells (Califano *et al*, 1996, Tabor *et al*, 2001, Braakhuis *et al*, 2003). Some accumulated mutations are lethal while others provide an additional selective advantage over the neighboring clones resulting in an outgrowth of autonomous clones with a more aggressive behavior that now differ from their predecessors by an additional mutation. Therefore, all daughter cells share the early genetic events, however subsequent subclones have additional genetic changes that give them a more aggressive growth advantage. Overall, it is the accumulation of mutations in several critical genes that are ultimately responsible for the initiation, development and progression of cancer (Scully *et al*, 2000).

1.5.1 Types of Genes Altered in Tumorigenesis

Two major types of genes are critical to tumorigenesis: proto-oncogenes and tumor suppressor genes (TSGs). There is a general consensus that alterations to both forms of genes are required for cancer to develop.

1.5.1.1 Role of Oncogenes in Cancer Development

During the development of cancer, normal cellular genes known as 'proto-oncogenes accumulate mutations through one of several possible mechanisms including: point mutation, gene amplification, chromosome translocation, and possibly through viral infections such as HPV. The normal functions of proto-oncogenes are to positively regulate cell division and differentiation. These genes encode for proteins such as growth factors, growth factor receptors, protein kinases and transcription factors. Examples of such genes known to be altered in oral carcinogenesis include *ras, EGFR, TGF-α, c-erb B-2, cyclin D1, and int-2 (FGF-3)* (Xia *et al*, 1997, Saraneth *et al*, 1992, Grandis *et al*, 1993, Merritt *et al*, 1990, Somers *et al*, 1990, Saraneth *et al*, 1989). When these genes become mutated, resulting in altered gene expression, and often increased cellular proliferation, these proto-oncogenes are then referred to as oncogenes. For example, the *ras* protein is a GTPase involved in relaying signals from receptor tyrosine kinases to the nucleus to stimulate cell proliferation. Mutation of the *ras* gene results in a protein that lacks the ability to hydrolyze GTP rendering it constitutively active, resulting in a continuous signal to the cell to divide (Alberts *et al*, 1994).

1.5.1.2 Tumor Suppressor Genes in Cancer Development

Unlike proto-oncogenes, tumor suppressors are negative regulators of the cell cycle and differentiation. They are responsible for inhibiting the cell cycle, thus ensuring that any mutations accumulated throughout the genome are repaired before the cell cycle proceeds. This ensures that the integrity of the genome is maintained and that mutations are not passed on to daughter cells. The function of these genes in cancer development is recessive in nature meaning that both copies of the gene (or its product) must be altered before the critical cell function is lost. In most cases, this requires two independent events, which can be diverse, including direct genetic damage (e.g. point mutations, chromosome breaks, chromosome loss), promoter hypermethylation (blocks transcription of gene) and even downstream inactivation of the protein by complexing to viral proteins (e.g. inactivation of *p53* and *Rb* proteins by human papilloma virus, see section 1.2.3). Several TSG's are lost in the development of OSCC including *INK4a/ARF*, and *p53* (Jares *et al*, 1999, Kresty *et al*, 2002, van Houten *et al*, 2002). Other candidate genes such as *FHIT* and *APC* are also possible TSGs involved in OSCC development (Nagpal *et al*, 2003). Unlike oncogenes, the function of a TSG is lost; however, the overall result of either gain of oncogene function or loss of a TSG can be increased cellular proliferation.

In a non-cancerous cell, there is a balance between proto-oncogene and TSG function resulting in a normal rate of cellular proliferation and turnover while maintaining genomic integrity. In cancer this balance is lost resulting in

increased proliferation and the inability to arrest the cell cycle resulting in the increased expansion of initiated cells.

1.5.2 Molecular Alterations of OSCC Development

Recently evidence suggests that specific molecular changes underlie the histological alterations that occur as oral cancer develops. A basic molecular progression model has been developed by Califano *et al.* (1996) to summarize these data. This model incorporates data on both the loss of specific tumor suppressor genes and the gain of oncogenes. Described below are the main early molecular events that occur in OSCC development.

1.5.2.1 Early Molecular Alterations of OSCC Development

The initiating event that is believed to occur in OSCC is the loss of the TSG located at 9p21 (Califano *et al*, 1996, Califano *et al*, 2000). The gene located at this locus actually consists of two genes produced from alternate reading frames. The first gene product is *p16*, a protein that binds and inhibits the action of Cyclin Dependant Kinase-4 (*CDK-4*), thus inhibiting the progression of the cell cycle past the R-Point (Rocco *et al*, 2001); a restriction point in the cell cycle where the decision is made to progress from the G1 to S phase (Forastiere *et al*, 2001). The protein produced by the alternate reading frame is *p19*, a protein that binds and destabilizes *mdm-2* (a gene involved in positive cell cycle regulation) (Sharpless *et al*, 1999).

The second genetic event in the development of OSCC involves loss of the short arm of chromosome 3 (3p) (Scully *et al*, 2000). At least three loci on

this chromosome arm are thought to display TSG activity, with most attention focusing on the *FHIT* located at 3p14.2 (Kisielewski *et al*, 1998). Loss at 17p13 occurs frequently with later stages in progression, commonly associated with alteration at the *p53* locus. *p53* is one of the most frequently altered genes in cancer development with mutation occurring in many diverse forms of human cancers. The protein product of this gene plays a multifunctional role in maintaining genomic integrity. *p53* is a transcription factor that regulates other gene products that either arrest the cell cycle (e.g. *p21*), stimulate DNA repair (e.g. DNA Polymerase β), induce apoptosis (e.g. *bcl-2*) or inhibit angiogenesis (e.g. *VEGF*), all of which are characteristics that are lost in cancer development (Zhou *et al*, 2001, Weinberg *et al*, 1991). Loss of 3p and 17p13 correlate histologically with the development and progression from mild to severe dysplasia (Gollin *et al*, 2001). Table 1 summarizes the progression molecular events correlated with OSCC development.

These early molecular changes that occur during the development of OSCC are prime candidates for molecular markers that could indicate early development of OSCC. Identification of such early genetic changes could be used to discriminate between those benign OPLs with a low risk of progressing towards OSCC and those that are histologically similar to their benign counterparts, but may possibly be at an elevated risk of OSCC, thus facilitating early intervention for high-risk individuals and therefore decreasing OSCC incidence.

Table 1 . Cytogenetic alterations in the genetic progression of HNSCC*

Histopathologic Description	Cytogenetic Change	% Frequency	Candidate Gene
Precursor Lesion or Benign Squamous Hyperplasia	9p21 loss	73	*INK4a/ARF*
Dysplasia	3p loss 17p13 loss	67 55	*FHIT* *TP53*
Carcinoma *in situ*	11q13 amplification 13q21 loss 14q24 loss	61 52 44	*CCND1* Novel TS Gene Not Reported
Invasive Carcinoma	6p loss 8p23 loss 4q26-q28 loss	38 40 47	Novel TS gene Novel TS gene Not Reported

**Adapted from Nagpal et al, 2001*

1.6. Fluorescence in situ Hybridization

The study of genetic alterations involved in the development and progression of OSCC is crucial both for understanding the sequential order of genetic events responsible for the disease progression and for the identification of genetic changes within individuals that could be used to indicate the risk of disease development or progression. Fluorescence *in situ* Hybridization (FISH) has been developed to work on interphase cells and has the ability to allow the enumeration of specifically targeted genomic regions of interest within single cells.

FISH is a molecular technique that can be used to visually detect the presence of loss or gain of whole chromosomes (called aneuploidy) or specific targeted chromosomal regions (Werner *et al*, 1997). FISH has successfully been

utilized to delineate the sequence of genetic alterations that occur during the progression of several different cancers including OSCC (Gunawan *et al*, 2001, Jin *et al*, 2001). The ability to detect such change with FISH in the future may have important clinical implications, with use to identify the presence of alterations known to be associated with risk of disease progression or recurrence. Currently, no molecular markers identified through FISH have been validated as an independent molecular marker of cancer risk; however cytogenetic evaluation of tumor material by FISH is playing an increasing role in clinical pathology where the results obtained from FISH can be used to complement the histological analysis of biopsies where the conclusions are unclear (Fiegl *et al*, 1999, Wolman *et al*, 1992). FISH involves the use of Fluorescently labeled DNA probes which hybridize to specific chromosomal regions complementary to the probe sequence (Awata *et al*, 2000). FISH probes can therefore be targeted to detect the loss of specific Tumor Suppressor Genes (TSG) or the gain of oncogenes providing molecular evidence for various cancer progression models. The procedure has successfully been applied to detect chromosomal alterations in exfoliated cells collected via bronchial lavage or voided in the urine of cancer patients and more recently brushings of the oral cavity (Awata *et al*, 2000, Inoue *et al*, 2000, Ishiwata *et al*, 2001, Barnes, R., 2003). It has also been applied to paraffin-embedded tissues, where it has successfully identified chromosomal alterations in several cancers including uveal melanoma (Patel *et al*, 2001), HNSCC (Jin *et al*, 2001), lung cancer (Gunawan *et al*, 2001), and oropharyngeal SCC (Khan *et al*, 2002).

1.7. Biomarkers of Oral Cancer Risk

Because individuals poses different genetic backgrounds, some individuals have an increased risk of cancer development. For instance, some individuals are at an increased risk of breast cancer development as a result of inherited mutations within the *BCRA1* or *BCRA2* genes (Futreal *et al*, 1994, Tavtigian *et al*, 1996). Other individuals inherit far less penetrable mutations within genes involved in various cellular processes such as DNA repair and metabolism. These mutations may elevate their risk of acquiring random DNA mutations that may possibly lead to the development of initiated cells. Identifying biomarkers indicative of early cancer progression or of a potential future risk of cancer development has a huge potential value for cancer prevention. The following sections will discuss the types of biomarkers that are currently being assessed as potential indicators of elevated genetic predisposition to oral cancer.

1.7.1. Genetic Susceptibility

Among the processes that are already known to play a significant role in susceptibility to oral cancer are those that affect DNA repair processes and metabolism. The following is a brief description of each.

1.7.1.1. Polymorphisms in DNA Repair

The association between defective DNA repair systems and a predisposition to cancer is best documented for rare autosomal recessive

diseases such as Xeroderma Pigmentosum (XP) (where there is an enzyme deficiency in the nucleotide excision repair (NER) pathway), and Hereditary Non-Polyposis Colorectal Cancer (HNPCC) where patients are at a high risk of developing colon cancer due to a deficiency in the MLH1 or MSH2 enzymes required for the Mismatch Repair pathway (Cleaver *et al*, 2001, King, 2000). The inability to repair or recognize DNA adducts or mismatches facilitates the rapid accumulation of mutations throughout the genome thus resulting in an increased cancer risk. Highly penetrant diseases such as XP and HNPCC only account for ~5% of all human cancers (Coughlin *et al*, 1999). The remaining cancers may be due to low penetrant cancer susceptibility genes such as those involved in DNA repair (Shen *et al*, 2003). It is these less penetrant polymorphisms that are far more common in the general population, and the data increasingly suggests that many cancers are a result of genetically-susceptible individuals that are exposed to low levels of carcinogens (Roberts *et al*, 1999). Genetic polymorphisms may modulate an individual's response to DNA damaging agents found in diet, smoking, alcohol, and the environment. The identification of these less penetrant polymorphisms that confer an increased cancer risk may therefore have a large impact on the identification of at risk individuals in the general population (Hu *et al*, 2002).

There are greater that 130 DNA repair genes that have been associated with single nucleotide polymorphisms (SNPs) where the change in a single codon may or may not affect the enzyme activity (Wood *et al*, 2001). It has been suggested that in the general population, many healthy individuals exhibit

reduced DNA repair capacity (60-85% of normal repair capacities) (Hu *et al*, 2002). Several recent experiments have used biomarkers to quantify DNA adducts and other forms of DNA damage in lymphocytes collected from various individuals, to evaluate the efficiencies of repair pathways in response to carcinogen exposure (see section 1.6.1.2. for more details). Such studies have reported a 20-50% difference in the DNA damage at a given exposure level, presumably due to differences in repair capacity. Whether this difference indicates a similar range in susceptibility to cancer is yet to be determined (Hu *et al*, 2002, Cheng *et al*, 2002).

The nucleotide excision repair (NER) pathways of DNA repair are involved in the repair of bulky DNA adducts and pyrimidine dimers. NER is a complex pathway that involves the interaction of at least 30 enzymes in order to recognize and repair DNA damage (Goode *et al*, 2002). Within several of the NER enzymes many polymorphisms have been identified, including the gene encoding for *XPC* (Shen *et al*, 2001). *XPC* is involved early in the process of NER where it is partially responsible for detecting the DNA damage and recruiting enzymes required for nucleotide excision repair (Shen *et al*, 2001). One polymorphism within the *XPC* gene has been identified as a risk factor for HNSCC. This polymorphism is found within intron 9 and involves the deletion of a five base-pair sequence (GTAAG) and an 83bp insertion consisting of only adenosine and thymine (Shen *et al*, 2001). The mechanism by which this polymorphism affects XPC function cannot easily be explained as it is located within an intron, but it may induce aberrant splicing resulting in an altered protein product, or the

polymorphism may be in linkage with another susceptibility gene for HNSCC (Shen *et al*, 2001).

Polymorphisms within double stranded (DS) DNA repair pathways have also been associated with an increased risk of cancer. DS DNA breaks are caused by normal metabolic processes, ionizing radiation, or exposure to environmental agents such as tobacco carcinogens (Wang *et al*, 2003). *XRCC3* is an enzyme which functions in promoting homologous recombination of DS DNA breaks through binding and stabilizing *Rad51* at the site of DNA damage (Shen *et al*, 2002, Wang *et al*, 2003). One SNP in *XRCC3* located in exon 7 at position 18067 results in a change from cytosine to thymine. This polymorphism has been identified as a risk factor for skin melanoma, bladder cancer, HNSCC, and lung cancer (Matullo *et al*, 2001, Shen *et al*, 2002, Wang *et al*, 2003, Winsey *et al*, 2000). Of interest, the increase in risk in HNSCC and lung cancer is only observed among heavy smokers, possibly as a result of a gene-environment interaction where no risk is conferred with low levels of carcinogen exposure (Shen *et al*, 2002, Wang *et al*, 2003).

Overall, polymorphisms in several cooperating enzymes in such complex DNA repair pathways may interact and affect an individual's overall ability to repair DNA damage. Since the polymorphisms may affect different types of DNA repair this may result in some specificity with respect to an individual's sensitivity to specific agents and the DNA damage that they produce.

1.7.1.2. Mutagen Sensitivity

Due to the complexity of the various DNA repair systems, it is difficult to link a reduced DNA repair capacity of an entire system to a single genetic change. To overcome this obstacle, mutagen sensitivity assays have been developed to determine an individual's overall DNA repair capacity for various repair systems. The results of these assays have been promising, suggesting that they may play a role in identifying individuals more susceptible to environmentally induced carcinogenesis, including those susceptible to oral cancer (Wu *et al*, 2002).

Most of the data that support a defect in DNA repair for oral cancer patients comes from work pioneered by Hsu *et al* (1989). Hsu developed an *in vitro* assay that could indirectly measure an individual's DNA repair capacity. This test involves the collection of blood samples from patients, followed by a cell culture of the blood lymphocytes. The lymphocytes are then incubated with bleomycin for 5 hours, after which they are treated with colcemid to arrest all mitosis. The cells are then harvested and prepared for chromosome analysis, which involves an assessment of the number of chromatid breaks for each sample (Zhang *et al*, 2000). The assay assesses the ability of an individual's base excision repair pathway to repair DNA damage since bleomycin induces both double-strand and single-strand DNA breaks in cells (Gu *et al*, 1999). The test has been modified to detect other forms of DNA repair by changing the mutagen used. For example, as mentioned earlier, benzo[a]pyrene diol epoxide (BPDE) is the reactive intermediate formed after phase I metabolism of BaP, a

chief carcinogen found in tobacco smoke (Cheng *et al*, 2002). BPDE binds to DNA resulting in a bulky adduct which must be repaired through the nucleotide excision repair (NER) pathway. The ability of an individual's NER pathways to repair damaged DNA is determined through the application of BPDE in the mutagen sensitivity assay (Wei *et al*, 1996).

Several case-control studies have examined the significance of mutagen sensitivity phenotypes with respect to risk of oral cancer (Yu *et al*, 1999, Zhang *et al*, 2000, Wu *et al*, 2002, Zhu *et al*, 2002). These studies have shown that chromosome breaks by both BPDE and bleomycin are significantly elevated in individuals with oral cancer, or in individuals with a family history of oral cancer. These data suggest that a reduced capacity may genetically predispose individuals to oral cancer development (Wu *et al*, 2002, Zhu *et al*, 2002). Of further interest is the effect of second hand tobacco smoke and its interaction with mutagen sensitivity on the risk of head and neck cancer. In a case control study by Zhang *et al* (2000) an odds ratio (OR) of 2.8 was found for the association between second hand smoke and the risk for head and neck cancer. However this relationship was not significant, probably due to its small sample size and a larger study is required to confirm this finding.

Mutagen sensitivity assays have also been used to demonstrate that DNA damage is not found randomly distributed throughout the genome, but rather it is concentrated in specific regions (Zhu *et al*, 2002). As mentioned previously, loss of 3p occurs early in the molecular progression of oral cancer and is one of the most frequently identified genetic aberrations. Three specific regions are

commonly deleted: 3p24-ter, 3p21.3 and 3p14-centromere (Zhu *et al*, 2002). Zhu *et al* (2002) used a combination of the mutagen sensitivity assay with BPDE and FISH analysis to demonstrate that 3p may be a molecular target of BPDE in HNSCC patients. Their data showed an increased number of genetic aberrations at 3p21 among cases. The odds ratio (OR) for HNSCC risk associated with BPDE-induced 3p21 aberrations was 4.8. This suggests that these genetic aberrations may be etiologically related to oral cancer and that defects in their repair might partially explain why cancer developed in these individuals.

The results of these mutagen sensitivity assays cannot directly be linked to the risk of oral cancer as these studies are all done on surrogate tissues (lymphocytes); however, other studies have demonstrated a strong relationship between genetic relatedness and mutagen sensitivity suggesting that DNA repair capacity is highly heritable (Yu *et al*, 1999, Zhu *et al*, 2002). For example, the OR for increased mutagen sensitivity of an individual with siblings affected by HNSCC is 14.6 (Yu *et al*, 1999). These studies on the heritability of the mutagen sensitivity phenotype suggest that using lymphocytes as a surrogate tissue may be relevant to predicting oral cancer risk.

1.7.2. Metabolic Polymorphisms

There is an increasing body of literature suggesting that metabolic polymorphisms also have an impact on an individual's cancer risk (Laffon *et al*, 2003, Morita *et al*, 1999, Cabelguenne *et al*, 2001). Biotransformation of compounds by metabolic enzymes plays an important role in carcinogenic

activity and organ specificity of environmental carcinogens, and a large inter-individual variation in the production of metabolic metabolites can partially be explained by genetic polymorphisms in metabolic enzymes (Autrup, 2000).

The amount of an ultimate carcinogen that reacts with DNA is determined by both its activation and detoxification. An increase in the rate of formation, or a decrease in the rate of detoxification theoretically results in an increased concentration of ultimate carcinogens. Thus, more DNA adducts are formed, consequently increasing the individual's risk of accumulating mutations in critical genes required to prevent cancer development (Autrup, 2000). An example of a decreased metabolic capacity and an increased level of DNA adducts is found in bladder tissues, where a decreased activity of N-acetyltransferase (a phase II enzyme involved in the detoxification of several compounds) results in a generally increased level of aromatic amine DNA adducts in the bladder tissue (Autrup, 2000). Several polymorphisms in metabolic enzymes have been identified and found to be associated with an increased cancer risk. These are described below.

1.7.2.1. Phase I Enzymes

Cytochromes p450s are a superfamily of enzymes that are involved in the metabolism of a large number of exogenous and endogenous compounds. The family consists of over 50 genes, coding for enzymes that all function as mixed function mono-oxygenases that require NADPH as a co-factor (Autrup, 2000). Substrates for these enzymes show a wide range from fatty acids to steroids and

a large number of xenobiotics. All tissues in the body express these enzymes in a tissue specific fashion although the liver is the site of highest activity (Autrup, 2000). Of the cytochrome families, Cyp 1, Cyp 2 and Cyp3 are the predominant ones involved in xenobiotic metabolism. Many genes in these 3 families have been found to be polymorphically expressed and hence, could impact on enzymes activities (Bartsch *et al*, 2000).

The Cyp 1 family consists of three genes: *CYP1A1*, *CYP1A2*, and *CYP1B1*. *CYP1A1* metabolizes a range of PAHs including benzo[a]pyrene and others in tobacco smoke, *CYP1A2* metabolizes arylamines, nitrosamines, and aromatic hydrocarbons, and *CYP1B1* metabolizes a number of PAHs and endogenous compounds such as estrogen (Autrup, 2000).

Two different genetic polymorphisms within the *CYP1A1* gene are thought to be associated with the large inter-individual differences in this enzyme's activity (Ko *et al*, 2001). The first polymorphism involves an Ile / Val substitution due to a base change at A4889G located within the heme binding region in exon 7. The variant form has been associated with a 2-fold increase in the enzyme's catalytic activity; however, several *in-vitro* studies have concluded that this variant is not functionally important (Persson *et al*, 1997). The second polymorphism found within CYP1A1 also involves a point mutation where thymine is substituted for cytosine and is known as the MspI variant allele; however, no change in the resultant protein amino acid sequence occurs as this polymorphism is found in the 3'-noncoding region at the nucleotide T6235C (Ko *et al*, 2001). Nevertheless, this variant allele has been associated with an

increased susceptibility to smoking-induced cancers in the Japanese population (Kiyohara *et al*, 1998).

Four polymorphisms of *CYP1B1* have also been identified at the following amino acid residues: 48 (Arg:Gly), 119 (Ala:Ser), 432 (Val:Leu), and 453 (Asn:Leu) (Ko *et al*, 2001). The Val:Leu variant found at codon 432 has been implicated in alterations to the catalytic ability of *CYP1B1* and has been found to be associated with colorectal cancer (Fritsche *et al*, 1999) and HNSCC (Ko *et al*, 2001). Ko *et al* identified the variant allele at codon 432 as a susceptibility factor in smoking-related cancer cases. Smoking cases were found to be 4.5 times more likely to carry the variant than non-smoking cases of HNSCC. In addition, an interactive effect between the variant *CYP1B1* allele and the phase II GSTM Null variant was found to further increase an individual's risk of HNSCC.

1.7.2.2. Phase II Enzymes

There are several phase II enzyme systems that detoxify the reactive metabolites produced from phase I reactions. The Glutathione-S-Transferases (GSTs) belong to one major class of phase II enzymes. This family of detoxification enzymes conjugates hydrophobic and electrophillic compounds with reduced glutathione in order to make the compounds more water soluble and therefore more easily excreted. Most compounds conjugated with glutathione are less toxic; however, a small number of compounds are actually bioactivated (Autrup, 2000).

GSTM is one subclass of the GST family. It consists of 5 genes (M1-M5), located at 1p13 (Autrup, 2000). The catalytic activity of proteins from these 5 genes is very similar; however there is some variation in expression patterns among tissues for the different classes of GSTM where, for example, *GSTM1* is highly expressed in the liver but not in the lung, and *GSTM3* is expressed in the lung. Among the general population there is a wide range of inter-individual GSTM activities due to gene deletions (which result in null genotypes) and allelic variation (Autrup, 2000). Individuals homozygous for GSTM null (no enzyme activity) have been reported to have an increased risk of lung, bladder, colon, and breast cancer, presumably due to their inability to detoxify reactive intermediate metabolites (d'Errico *et al*, 1996). A gene interaction has also been found to exist between GSTM alleles, smoking habits, and the risk of oral cancer. A significant risk in oral cancer (Odds Ratio 2.6) was found in individuals who smoked and carried the GSTM1 null allele (Ko *et al*, 2001).

1.7.2.3. Alcohol Metabolism

Alcohol is an independent risk factor for oral cancer; therefore polymorphisms in the enzymes responsible for the metabolism of ethanol may have an impact on an individual's risk. Recent reports have revealed that acetaldehyde, a metabolic intermediate of ethanol, may be the main carcinogen related to increased cancer risk. It has been found to cause mutations, form DNA adducts, and inhibit DNA repair (Zavras *et al*, 2002). Acetaldehyde levels are regulated by three enzyme systems: (1) Alcohol dehydrogenase (*ADH*) which produces acetaldehyde from ethanol; (2) Acetaldehyde dehydrogenase which

converts acetaldehyde into acetate; and (3) the cytochrome p450 phase 1 enzyme *CYP2E1*.

Alcohol dehydrogenase is primarily responsible for the metabolism of alcohol. Seven genes encoding *ADH* have been identified in humans and variations within these genes have been associated with altered enzyme activity (Zavras *et al*, 2002). One SNP found within *ADH3* has been found to significantly affect the V_{max} for ethanol oxidation. Individuals with $ADH3^{1-1}$ genotype have a 2.5 fold increase in their V_{max} compared with individuals with the homozygous genotype $ADH3^{2-2}$ (Bosron *et al*, 1986). A study on French alcoholics identified the variant $ADH3^{1-1}$ as a potential risk factor for laryngeal cancer; cancer patients had elevated frequencies of the $ADH3^{1-1}$ allele over controls (Coutelle *et al*, 1997). However, a second study analyzing the effect of *ADH3* polymorphisms and oral cancer risk due to alcohol consumption, found exactly the opposite relationship between drinking and *ADH3* genotype. The results from this study indicated that individuals with an $ADH3^{1-1}$ allele were at a decreased risk of oral cancer compared with individuals carrying the $ADH3^{2-2}$ genotype (Zavras *et al*, 2002). Several other studies have also investigated the effect of *ADH3* polymorphisms and oral cancer risk, and have still not resolved any definitive relationship between *ADH3* and oral cancer risk due to conflicting results. All of these studies were carried out on a small scale and only produced barely significant results. To determine the true relationship between *AHD3* and oral cancer risk, studies will need to be carried out using much larger sample sizes in order to produce repeatable, convincingly significant results.

Alcohol is also metabolized to acetaldehyde to a lesser extent by *CYP2E1*. Polymorphisms in *CYP1E2* may therefore also affect an individual's cancer risk due to alcohol consumption. *CYP2E1* and *ADH* are both expressed in the liver where the majority of ethanol metabolism occurs; however, both of these enzymes are also expressed in the upper aerodigestive tract where differences in enzyme activities may also directly influence oral cancer risk. In addition to alcohol, *CYP2E1* is also involved in the metabolism of several xenobiotics, including a range of low-molecular-weight organic compounds such as benzene and N-nitrosamines (both tobacco carcinogens). One genetic variation found within this gene is known as the Rsa1/PstI allele, found within the 5' flanking region of the gene. This polymorphism results in differential transcription regulation (Bouchardy *et al*, 2000). The variant allele results in an increased metabolizer phenotype for benzene and has been associated with an increased risk of lung cancer in smokers (Wu *et al*, 1997). In one study by Bouchardy *et al* (2000), the relationship between *CYP2E1* polymorphisms and cancers of the upper aerodigestive tract was investigated. This study showed a borderline significant association between the variant *CYP2E1* allele and increased oral, pharyngeal, and laryngeal cancer risk, with a 2.6-fold increase in risk of oral and pharyngeal cancer.

1.7.2.4. Additional Polymorphisms Conferring Risk

Further complicating the overall evaluation of an individual's risk based on genetic polymorphism are polymorphisms that have been identified in enzymes regulating the cell cycle such as *p21*. *p21* is a universal inhibitor of cyclin

dependant kinases (CDK's), responsible for progression through the cell cycle. This protein plays a central role in arresting cellular growth, inducing terminal differentiation, and apoptosis (Xiong *et al*, 1993). The *p21* protein has several domains that interact directly with *p53*, CDKs, and cyclins and changes in its amino acid sequence could significantly alter binding of these proteins (Ralhan *et al*, 2000). One polymorphism is located at codon 149 in exon 2, resulting in a substitution of aspartic acid for glycine. This polymorphism has been reported to be at a significantly higher frequency in patients with either oral premalignant lesions (involving hyperplasia or dysplasia) or oral cancer suggesting that this change in amino acid sequence within *p21* may a risk factor for oral cancer development (Ralhan *et al*, 2000).

1.7.3. *FHIT* as a Molecular Marker of Oral Cancer Risk

As previously discussed, loss of 3p is one of the early events in several cancers (Mao *et al*, 1996). Although at least three different regions are lost and are proposed to be the sites of putative tumor suppressor genes, 3p14 is the region that is most strongly linked to risk of progression (Rosin *et al.*, 2000). The putative tumor suppressor gene found in this region is thought to be *Fragile Histidine Triad* gene or *FHIT*, located at 3p14.2. Interestingly, 3p14 may be targeted by carcinogens, as the *FHIT* gene lies within the fragile site *FRA3B* (Corbin *et al*, 2002).

Several fragile sites are found throughout the human genome including *FRA3B* located at 3p14, *FRA7G* on chromosome 7, and *FRA16D* on

chromosome 16. These sites are described as highly unstable regions of the genome that show gaps, breaks, rearrangements, and mutated chromosomes when cultured under conditions that inhibit DNA replication (Glover *et al*, 1998). Due to their unstable sequences, they have also been implicated in the initiation and perpetuation of breakage-fusion cycles leading to chromosome rearrangements (Corbin *et al*, 2002). *FRA7G* and *FRA16D* have both been shown to be deleted in tumor cells.

Although the location of such common fragile sites has been determined, the underlying mechanism that gives rise to their increased susceptibility to DNA damage has not been determined. Sequencing the areas of deletion within *FRA3B* has not identified specific sequences that are more susceptible to fragmentation (Corbin *et al*, 2002). It is not known if these fragile sites are composed of a single spanning unstable sequence, or if they are composed of multiple "hotspots" for DNA recombination or breakage (Corbin *et al,* 2002).

FRA3B is one of the most highly expressed common fragile sites within the genome (Wilke *et al*, 1996). This has great implications for the stability of *FHIT*, increasing its susceptibility to instability and carcinogen-induced damage. Rearrangements within *FHIT* are common in carcinomas of the breast, ovary, pancreas, colon, kidney, and esophagus (Negrini *et al*, 1996, Virgilio *et al,* 1996, Sozzi *et al*, 1996). In breast cancer, loss of *FHIT* is believed to be associated with the loss of *BCRA2* or *BCRA1*, with the latter both involved in the repair of DS DNA breaks (Pandis *et al*, 1997). Significantly higher frequencies of LOH at 3p14.2 have been observed in breast cancer tumors with *BCRA2* deleterious

mutations compared with sporadic breast carcinomas not harboring *BCRA2* mutations (Turner *et al,* 2002), suggesting that loss of *BCRA2* affects the stability of *FHIT/FRA3B* within breast carcinomas, resulting in an increased loss of 3p14.2 that has been associated with a reduced expression of *FHIT* (Turner et *al,* 2002).

The function of *FHIT* in the development of carcinomas has also been investigated in Hereditary Non-Polyposis Colorectal Cancer (HNPCC). *FHIT* knock-out mice have been shown to be more susceptible to the development of "Muir-Torre Syndrome", characterized by one or more sebaceous tumors and the coexistence of one or more visceral carcinomas (Turner *et al,* 2002). This Muir-Torre Syndrome in mice is histologically similar to HNPCC in humans. A large sub-group of HNPCC patients are predisposed to HNPCC through germ-line mutations in *MSH2* or *MLH1*, genes responsible for mismatch repair. It has been hypothesized that the absence of *FHIT* in a large fraction of these diseases may be due to unrepaired damage at *FRA3B* (Turner *et al,* 2002). *FHIT* may therefore be a molecular target in such repair-deficient cancers leading to *FHIT* loss and clonal expansion of *FHIT -/-* tumors. This supposition is supported by other studies that suggest that genes located within fragile sites such as the *WWOX* gene (located within *FRA16D*) exhibit hallmarks of a TSG (Huebner *et al,* 2001) may be directly involved in tumorigenesis of repair-deficient cancers, as these areas are more susceptible to DNA damage (Turner *et al,* 2002).

Loss of *FHIT* expression has also been identified in esophageal SCC, where hypermethylation of the *FHIT* promoter has been associated with *FHIT*

transcriptional inactivation (Tanaka *et al* 1998). Like the molecular progression model of HNSCC (see section 1.5), loss of *FHIT* has been demonstrated to occur early in the process of esophageal tumorigenesis, present in precancerous lesions (Mori *et al*, 2000). Moreover, a high frequency of *FHIT* hypermethylation has also been identified in non-cancerous tissues of patients with esophageal SCC and may represent early predisposing changes in that tissue (Kuroki *et al*, 2003).

Due to the early appearance of alterations to *FHIT* in the progression of tumorigenesis and its apparent role in the development of several types of carcinomas, *FHIT* may be an excellent early biomarker of HNSCC development and may be indicative of an elevated risk of premalignant lesions progressing to cancer. Detection of such high-risk premalignant lesions may allow early intervention in the disease, resulting in a decrease in the incidence, morbidity and mortality of the disease.

2. STATEMENT OF THE PROBLEM

Despite improvements in surgical, chemotherapeutic, and radiation therapies over the last five decades, the prognosis for oral cancer patients has not improved and remains poor with a 5-year survival rate not exceeding 50%. An alternative approach to reducing the number of deaths due to oral cancer is to reduce the rate of cancer incidence through early detection resulting in cancer delay or prevention. Oral premalignant lesions (OPL's) are indicative of increased oral cancer risk, however due to their low rates of malignant transformation, clinicians are unable to accurately identify those OPL's that will progress to cancer. In addition, not all oral cancers arise at the same site of OPL's and some cancers arise without the presence of OPL's. The identification of molecular markers that are mechanistically linked to the genetic progression of oral cancer may provide more accurate risk estimates of individuals with an increased risk of oral cancer development. Loss of *FHIT* occurs early in oral cancer, and the identification of changes to *FHIT* associated with increased cancer risk within individuals practicing high-risk behaviors for oral squamous cell carcinoma may be able to identify those at risk of malignant transformation and thus facilitate early intervention and cancer prevention.

3. SIGNIFICANCE OF THE STUDY

This study is unique as its research focus is not limited to individuals with clinical signs of either pre-cancer or malignant transformation, rather the goal of this study was to determine if genetic alterations associated with oral cancer could be identified within individuals before any clinical onset of the disease presents itself. This study demonstrated that non-invasive methods of sample collection such as the cytobrush are capable of collecting sufficient sample sizes from individuals of the general population, and provided evidence that such sampling procedures would be applicable in the development of high-throughput oral cancer prevention and screening programs.

The results of this study provides strong evidence that genetic alterations associated with oral cancer development can be detected in individuals with no clinical signs of the disease and supports future research into the development of molecular markers such as *FHIT* associated with early cancer development. Such molecular markers have a great potential for early cancer detection and intervention that may reduce the morbidity and mortality currently associated with oral cancer.

4. OBJECTIVES

1. To optimize a protocol that would produce high quality nick translation probes targeted towards 3p14.2 and centromere 3 to apply in *in situ* Hybridization using exfoliated cells.

2. To determine if *in situ* Hybridization with probes targeted towards 3p14.2 and centromere 3 could be used to detect the presence of abnormal signal patterns within these genomic regions in samples obtained from smokers.

3. To develop cut-off values that could be used to discriminate between naturally occurring background levels of genetic alterations and elevated levels of damage at 3p14.2 and centromere 3 due to prolonged exposure to tobacco carcinogens.

5. HYPOTHESIS

Due to prolonged exposure to tobacco carcinogens, exfoliated cells collected from high-risk regions (ventrolateral tongue and floor of mouth) from smokers will contain rare genetic alterations to centromere 3 and 3p14.2 that are consistent with genetic alterations occurring in tumor margins of oral cancer patients, and secondly that smokers will exhibit significantly increased levels of genetic damage at these loci when compared with age-matched and gender-matched nonsmokers.

6. MATERIALS AND METHODS

6.1. Study Groups

Ethical approval was obtained from the research ethics committees at Simon Fraser University (SFU) (Appendix 1), University of British Columbia (UBC), and the British Columbia Cancer Agency (BCCA). Participation in the study was voluntary, and signed consent was required before the participants entered into the study. To ensure confidentiality, each study subject was assigned a study number that was solely used throughout the entire study. There were three groups of individuals included in this study:

1) Oral cancer patients (19 individuals): diagnosed with oral carcinoma *in situ* or oral squamous cell carcinoma at time of collection.

2) Smokers (30 individuals): exposure of at least 1 pack year of cigarettes and no history of head and neck cancer.

3) Non-smoker controls (29 individuals): no history of smoking or head and neck cancer.

6.1.1. Oral Cancer Patients

Samples from individuals with cancer were collected from the patients attending the Oral Dysplasia Clinic at the BC Cancer Agency. This clinic has been accruing oral cancer and pre-cancer patients into a province-wide, ongoing

longitudinal study referred to at the Oral Health Study (OHS) since November 1999. Inclusion of participants into the study required histological confirmation of either CIS or OSCC by the study pathologist Dr. Lewei Zhang. CIS was confirmed by the presence of dysplastic changes found throughout the entire epithelium, while OSCC was characterized by invasion of the underlying basal lamina.

6.1.2. Smokers and Non-smokers

Samples from individuals with no clinical signs of OSCC development were collected from smokers and non-smokers from patients attending the Abacus Dental Centre, owned and operated by Dr. Stanly Soon. To be incorporated into the study, all patients had to be a minimum of 19 years old and provide information regarding their demographics and tobacco exposure through the completion of a questionnaire (Appendix 2).

6.2. Demographic Data and Sample Collection

6.2.1. Collection of Demographic Information

All participants in this study filled out a detailed questionnaire upon entry into the study (Appendix 2). However, for the purpose of this study, only the data providing information on age, gender, smoking habits, and smoking exposure were used as they were the factors required to investigate the effects of tobacco carcinogens on *FHIT* and centromere 3. To quantitate the amount of tobacco usage by both former and current smokers, the pack-years of cigarette exposure

was calculated (refer to Appendix 3 for an example of pack-year calculation).

6.2.2. Collection of Exfoliated Cells

Prior to sample collection, all oral cancer subjects were instructed to rinse their mouths with 20ml of tap water to remove food debris. All participants (smokers and non-smokers) from the Abacus Dental Centre located in Mission British Columbia were instructed to brush their teeth before sample collection if the sample was collected before their teeth were cleaned. Patient whose samples were collected following teeth cleaning were not required to brush their teeth prior to sample collection.

Samples were collected with a sterile cytobrush (Henry Schein Arcone Inc., Delta BC) applied to the oral mucosa and repeatedly passed over the surface for at least 12 strokes. Following sample collection, the cytobrush was immediately immersed in a sterile cryovial containing 1ml of TE-9 buffer solution (0.03 M Tris, 3.0 mM EDTA, pH 8.9) and was spun in the solution vigorously to ensure effective transfer of the cells from the cytobrush into the buffer solution. The cryovials were then immediately transferred into a liquid nitrogen tank where they were snap frozen and subsequently stored until use.

One sample was collected from each oral cancer patient. This sample was taken from the tumor margin mucosa 5 mm-15 mm outside of the clinically-identifiable lesion. Two clinically normal samples were collected from both the smokers and nonsmoker patients of the Abacus dental clinic. The first sample was taken by brushing the entire surface of both buccal mucosae, and the

second sample was then collected by brushing both sides of the ventrolateral tongue and floor of mouth as these locations are at a higher risk of oral cancer development in western countries (Parkin *et al*, 2001).

6.3. Laboratory Procedures

The data collection of this study was done blinded. All samples were randomly coded so that the patient type was not known. This study used two Fluorescently labeled DNA probes that were made within our laboratory using the Nick Translation Kit supplied by Vysis (Downers Grove, IL). The first probe from the BAC 91A15 was made to target chromosome 3 (CEP3), while the second probe made using the BAC 170K19 was targeted to bind to 3p14.2. The CEP3 probe was labeled with Spectrum Green fluorophores and the 3p14.2 probe was labeled with Spectrum Orange fluorophores. Figure 4 provides an overview of the location along chromosome 3 where the BACS used to make these probes are located.

6.3.1. Nick Translation

6.3.1.1. Preparing the Reagents for Nick Translation

The overall goal of the nick translation procedure is to produce Fluorescently labeled DNA probes targeted towards specific genomic regions with maximal , optimal for sample scoring. Due to the sensitivity of the reagents to light, all reagents used in Nick Translation were handled in a low light

environment. Nick Translation is supplied as a kit from Vysis (Downers Grove, IL); therefore all reagents in the procedure are supplied through Vysis. Three different Nick Translation reaction mixtures were individually prepared as per the manufacturer's directions, directly before their use as described by Vysis. The solution A contained the Fluorescently labeled dUTP, and this solution was prepared by adding 1µl of 1mM dUTP to 4µl of PCR water. The dUTP is Fluorescently labeled and when applied in FISH, the dUTP provides the source of used to enumerate the copy number of the targeted genomic region;, therefore the selection of the color of the dUTP applied in Nick Translation determines what color the DNA probe will fluoresce when applied in FISH. For the 3p14.2 locus, dUTP labeled with Fluorescence orange fluorophores was selected, and for CEP3, dUTP Fluorescently labeled with green fluorophores was selected. Solution B included dTTP where 2µl of 0.3mM dTTP was added to 4µl of PCR water. Solution C contained a mixture of dNTP's where 4µl of 0.3mM dATP, dCTP, and dGTP were combined together in a microcentrifuge tube.

6.3.1.2. Nick Translation Assay Procedure

Again, due to the sensitivity of the reagents to light, the Nick Translation Assay was carried out in a low light environment. The reaction was carried out in a 1.5ml blue centrifuge tube to reduce the penetrance of light into the reaction mixture. First, 1 µg of the specific extracted bacterial artificial chromosome (BAC) DNA was added to the 1.5ml centrifuge tube. The BAC 91A15 contains the genomic region encompassing the 3p14.2 locus therefore 91A15 BAC DNA

was used to make the probes targeted towards 3p14.2, whereas the BAC170K19 contains a genomic region encompassing the centromeric region of centromere 3 and was therefore applied in the Nick Translation Assay when preparing probes targeted for CEP 3. After the addition of the appropriate BAC DNA, nuclease free water was added to the reaction mixture to bring the total volume up to 17.5µl. Following the addition of the nuclease free water, 2.5µl of solution A was added, followed by addition of 5µl of solution B and 10µl of solution C. The Nick Translation 10X buffer was then removed from the –20°C freezer and immediately 5µl of the buffer was added to the reaction mixture. Finally, 10µl of the nick translation enzyme was added, bringing the total volume of the reaction mixture to 50µl. The time immediately after the enzyme was added was recorded and used as time zero. The nick translation reaction was then incubated in the dark at 15°C for 2-4 hours or until the final average DNA fragment size was approximately 200-350 base pairs long. Fluorescence DNA probes within this size range were found to produce maximal signals when used in FISH, optimal for sample scoring.

At regular intervals, 9µl of the incubating Nick Translation reaction mixture was removed and added to 1µl DNA loading buffer where it and a 100 base pair DNA ladder were then loaded into a 2% agarose gel and run at 120V/cm to determine the average size of the DNA fragments. When the average size of the DNA fragments in the sample were approximately 500 base pairs, the Nick Translation reaction was stopped by immersing the centrifuge tube containing the reaction mixture in a 70°C water bath for ten minutes to denature the enzyme

and stop the further digestion of the DNA. The reaction mixture was stopped when the latest sample size was determined to be 500 base pairs, as one hour had elapsed since obtaining the sample and running it on the agarose gel, therefore during that hour, the nick translation reaction continued, resulting in an average fragment size within the desired optimized size range. The final DNA fragment length was then determined electrophoretically, and the Nick Translation reaction was stored at -20°C until applied in FISH.

6.3.1.3. Precipitating the Probe

Dual probes were applied to each slide, and each of the probes (3p14.2 and CEP3) were precipitated separately prior to application in FISH. 10ul of nick translation probe DNA for four FISH slides was precipitated out of solution and re-suspended in an appropriate hybridization buffer before being applied to the FISH slides. For each slide processed through FISH, 2.5µl of each nick translation reaction mixture was required for sufficient probe binding, and four slides were processed through FISH at the same time. Therefore 10µl of each nick translation reaction mixture was precipitated each time slides were processed through FISH. To the 10µl of nick translation reaction mixture, 2µg of COT-1 DNA was added, followed by the addition of 5 µg of salmon sperm DNA and 8µl of double distilled water. To this precipitation solution, 2.4 µl of 3M sodium acetate and 60µl 100% ethanol were added. The precipitation solution was then placed at -20°C for 15-20 minutes after which it was centrifuged at 8000g for thirty minutes at 4°C to collect the precipitated DNA as a pellet.

Following centrifugation, the supernatant was removed and discarded, and the DNA pellet was air dried at room temperature for 10-15 minutes. The DNA was then re-suspended in 16.2µl LSI hybridization buffer supplied by Vysis (Downers Grove, IL) and stored at -20°C until required for FISH.

6.3.2. Fluorescence *in situ* Hybridization

6.3.2.1. Slide Preparation

Samples were removed from liquid nitrogen and left to thaw on the bench. A sufficient amount of cells was then removed from the cryovial (100-300 µl) and transferred into a 1.5 ml centrifuge tube. Exfoliated cells were collected from the sample by centrifugation at 5000g for 5 minutes. Centrifugation was repeated for an additional 5 minutes if the pellets were still too diffuse. The supernatant was then removed from the pellet of cells, followed by resuspension in Carnoy's solution (100% methanol : glacial acetic acid 3:1 (v/v)). The volume of Carnoy's solution added to resuspend the cells depended on the pellet size. If the pellet was not visible, 20 µl of Carnoy's was added, if a pellet was visible, up to 35 µl of Carnoy's was added depending on the pellet size. The suspended exfoliated cells were then pipetted onto a silanized glass slide from a height of 5-7cm to ensure proper dispersion of the cells. The slides were then left at room temperature for approximately 15 minutes in the fume hood to allow the Carnoy's solution to evaporate resulting in fixation of the exfoliated cells to the slide. The location of the cells was then marked on the bottom side of the slide using a diamond pen to show the area of the slide containing the sample. By marking

the area of the cells, this ensured that later in the procedure, the probes were added onto the area of the slide containing the sample.

6.3.2.2. Pretreatment of FISH Slides

Before probes were hybridized to the samples, the sample material was pretreated to reduce the non-specific probe hybridization to non-target nucleic acids, and to reduce the interaction of the probe with proteins or other cellular components. Pretreatment of the samples also facilitated the penetration of the probe into the cell nuclei in order for probe binding to occur. As soon as the samples dropped onto the glass slides were dry (10-15 minutes at room temperature) the slides were immediately placed in a 2X saline sodium citrate (SSC) solution (pH 7.2-7.4) at 37°C and aged for twenty minutes. The purpose of aging is to improve hybridization efficiency and signal brightness. Next, the slides were immersed in a pepsin solution (Sigma-Aldrich Corp., St. Louis, MO; prepared from 49.5ml purified water, 0.5ml of 1M HCL, and 25 µl of 10% pepsin in distilled water) at 37°C for four minutes. Immersion in the pepsin solution increases the permeability of the nuclei and unmasks nucleic acids from associated proteins which overall increases probe detection and accessibility. Following immersion in pepsin, slides were then washed with 1X phosphate-buffered saline (PBS) at room temperature for two minutes. Next, slides were immersed in the fixation solution (41.1ml H_20, 2.5ml 1M $MgCl_2$, 5.0ml of 10X PBS, and 1.35ml of 37% formaldehyde) for two minutes at room temperature. Fixation stabilized the cellular structure by cross-linking the cellular proteins

resulting in the reduction of diffusion and loss of DNA during denaturation. Slides were then washed twice in1X PBS for two minutes. The slides were then dehydrated in a series of increasing concentrations of ethanol (70%, 85%, 100%) at room temperature for one minute each, then placed in the fume hood for 1-2 minutes to dry.

6.3.2.3. Slide and Probe Denaturation

Denaturation of the cellular and probe DNA is essential as it separates the double stranded DNA and thus facilitates the single stranded probes to bind to their complementary DNA sequence within the cell. Following the dehydration step, the slides were placed in the denaturation solution (35ml 70% formamide, 5ml 20X SSC, and 10ml H_2O; pH 7.2-7.4) at 73°C for five minutes. Immediately following denaturation, the slides were dehydrated in a series of increasing concentrations of ethanol (70%, 85%, and 100%) at room temperature for one minute each. Following dehydration, the slides were left at room temperature to dry. At this point, the target cellular DNA was ready to receive the probe. After denaturation of the samples, both of the probes (3p14.2 and CEP3) suspended in the LSI buffer as described earlier were removed from -20°C and placed in a 73°C water bath for 5 minutes to ensure complete denaturation of the probe. Following denaturation, the two probes were combined and mixed.

6.3.2.4. Probe Hybridization

To each of the prepared FISH slides, 8µl of the mixed probe hybridization solution was applied to the center area where the sample was located. A sterile

25 mm^2 coverslip (Corning, Acton, MA) was placed over the droplet of hybridization mixture and slight pressure was applied to the coverslip to remove air bubbles and to ensure complete coverage of the sample area with the hybridization mixture. Rubber cement was then applied on all edges of the coverslip to ensure the probes did not escape from under the coverslip. Slides were placed in a humidified hybridization box (Figure 1) and this box was stored at 37°C for 16-24 hours to ensure adequate time for hybridization of the probe to the cellular DNA. Again, as the probes applied to the slides were light sensitive, the steps involved in probe denaturation and hybridization were done in a low-light environment.

6.3.2.5. Washing and Counterstaining

After hybridization, the non-specifically bound and weakly bound probe was removed by repeated washing, leaving probe bound only to areas of perfectly matched nucleotides and thereby reducing background . To help facilitate the removal of non-specifically bound probe the detergent Nonidet P-40 (NP-40) was used (Sigma-Aldrich Corp.) which is anionic, DNAase-free detergent. The slides were removed from the hybridization box and the rubber cement and coverslips were removed. The slides were then immediately immersed in the first washing solution (0.4% SSC / 0.3% NP-40; pH 7.5) at 73°C for two minutes. Next, the slides were immersed in the second washing solution (2X SSC, 0.1% NP-40; pH 7.5) for two minutes at room temperature. The slides were then allowed to dry by placing them upright in a dark chamber at room

temperature for 30 minutes.

After the slides had dried, the non-hybridized DNA was counterstained with 4,6 diamidino-2-phenylindole (DAPI II, Vysis, Downers Grove, IL) to identify the nuclear regions of each cell within which the Fluorescence signals could be counted. DAPI II is an effective nuclear counterstain as it binds to DNA and fluoresces bright blue when exposed to UV light. To each target area on each slide, 6µl of DAPI II was added, and a 25 mm^2 glass coverslip was then placed over the target area. Slight pressure was applied to the coverslip to ensure no bubbles existed and to ensure the DAPI II covered the entire target area. To permanently seal the coverslip to the slide, nail polish was applied to all edges of the coverslip. Slides were then kept in the dark at room temperature while the nail polish dried, and afterwards all slides were stored in a dark container at -20°C until they were later analyzed by microscopy.

6.3.4. Signal Enumeration

Signal enumeration was carried out using an Olympus BX51 microscope. To ensure that the observer was scoring FISH samples with accuracy, a FISH slide accompanied with average FISH scores supplied by Vyvis was analyzed to ensure the FISH scores obtained were consistent with the results supplied by Vysis. The criteria used to determine whether a cell was evaluable followed those recommended by Vysis and those validated by Zhang *et al* (1993). Only those cells that satisfied the following criteria were evaluated and included: (1) no overlapping cells; (2) cells have an intact nucleus; and (3) non-specific

hybridization signals were not counted (recognized by lower intensity and different shape). The following guidelines were followed for signal enumeration: (1) split signals characterized by two small overlapping signals were scored as 1 if they did not separate when finely adjusting fine focus; (2) to be scored as separate, all signals within each cells had same size and intensity of; (3) diffuse signals were counted if they were completely separate from all other signals; (4) two signals connected by a strand of were counted as one signal; and (5) only nuclei in which enumeration could confidently be determined were counted.

Using a 15X ocular lens and a 100X oil immersion objective lens, slides were brought into focus using the fine adjustment. The slide was then moved so the field of view located at the uppermost left area of the hybridization area. The procedure used to enumerate the slide is demonstrated by figure 2. Slides were scored by beginning to scan the slide from left to right, scoring all enumerable nuclei within the field of view; once the visible boundary of the hybridization area was reached, the field of view was moved down until the next field of view was brought into focus, and the slide was then scanned from right to left and so on. For the samples obtained from the tumor margins of oral cancers, 200 nuclei were counted, or if less than 200 enumerable cells existed on the slide, the entire hybridization area was scored. For the smoker and non-smoker samples, 500 nuclei were scored, and as was done in the tumor margin samples, if 500 enumerable cells were not found on the slide, all enumerable nuclei on the slide were scored. Figure 3 provides an example of signal patterns that were observed among the samples.

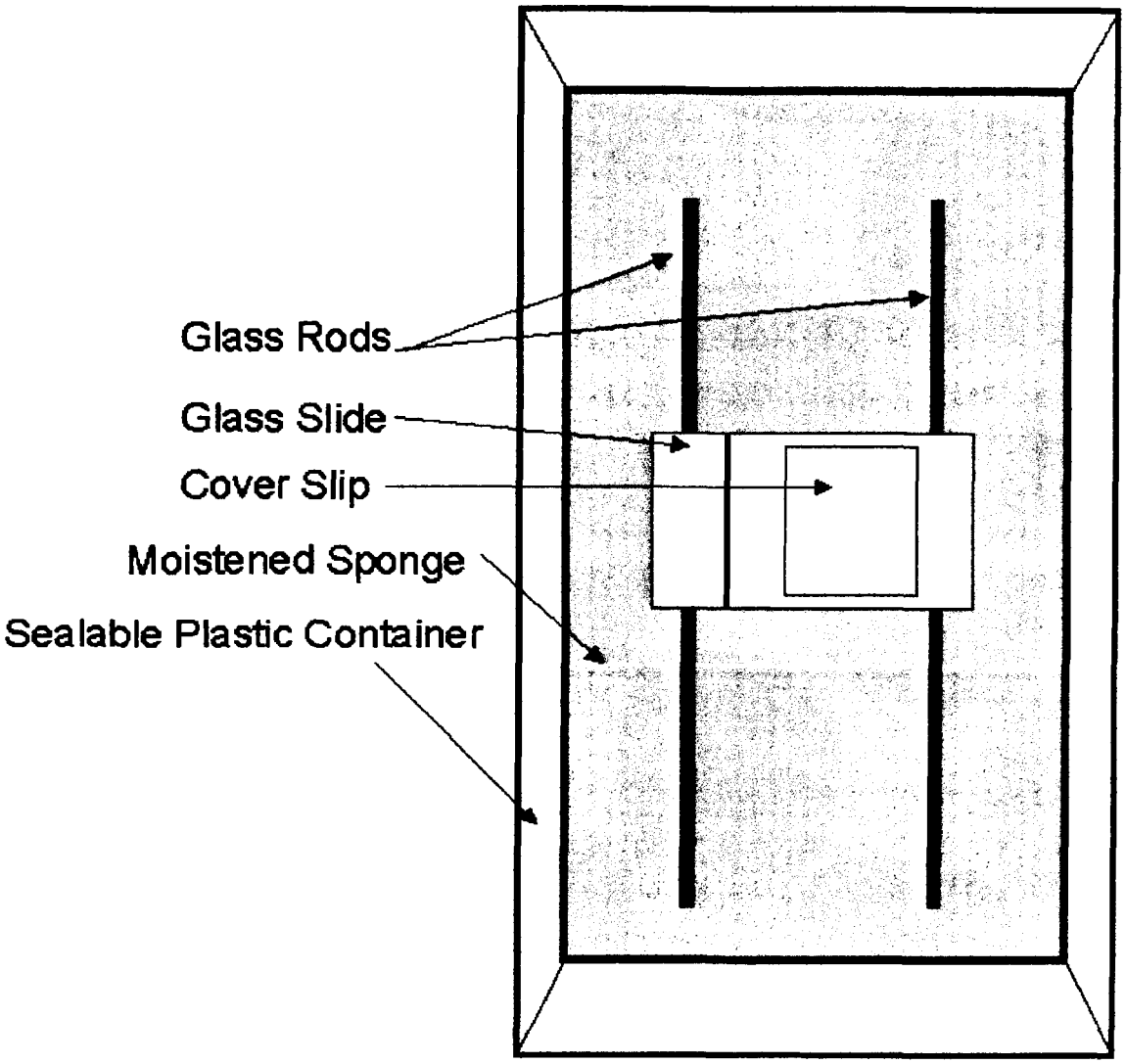


Figure 1. The Hybridization Chamber.

An illustration of the sealable plastic container that is used for the hybridization step of the FISH protocol. A moistened sponge at the bottom of the container is used to maintain humidity during probe binding.

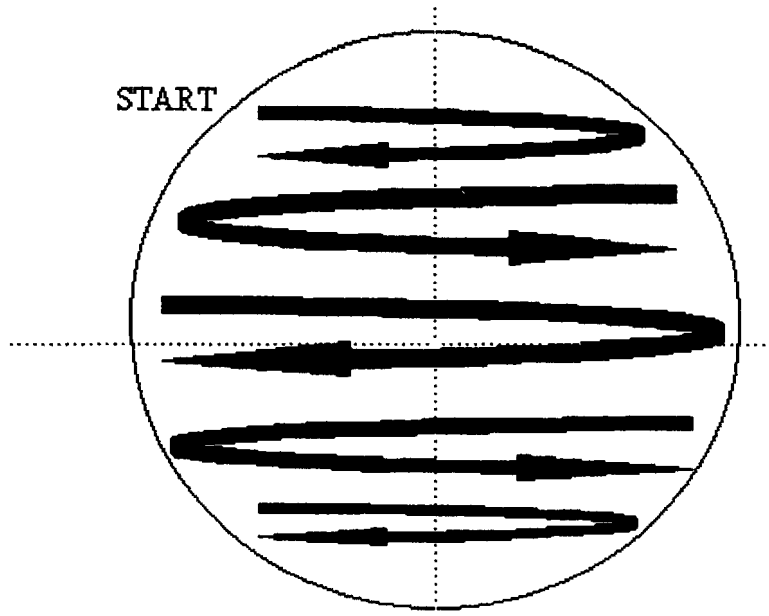
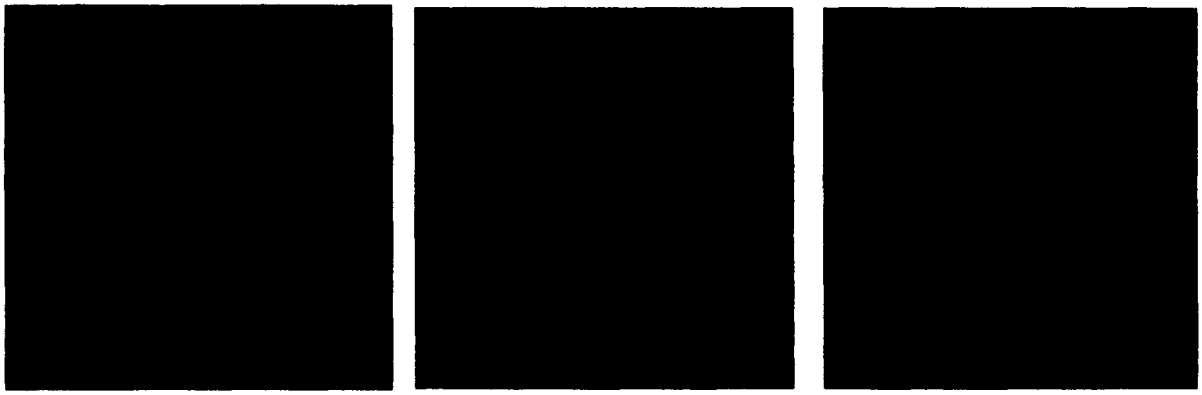


Figure 2. Scoring Pattern of FISH Slides.

An illustration of the approach used to score cells for FISH patterns. Scoring begins at the left uppermost corner of the slide, with the microscope objective moving across the slide to the right. All cells that met the criteria described in section 6.3.4. were evaluated. Once the edge of the sample on the right hand side of the slide was reached, the view of the microscope was adjusted down one field of view, and all scorable cells were counted scanning the sample right to left. This process was repeated until the slide was covered.



3A 3B 3C

Figure 3. FISH signal patterns.

DAPI nuclear stain results in a blue stain that facilitates the identification of single nuclei. The 3p14.2 probes are observed as red signals, and the CEP3 probes are observed as green signals. 3A is an example of cells with a normal (2.2) signal pattern. Figure 3B is an example of aneuploidy, where there are three copies of chromosome 3 (3 green signals) and only one copy of the 3p14.2 locus. Figure 3C is an example of loss of one copy of 3p14.2 observed as only one red signal while two copies of chromosome 3 are still retained.

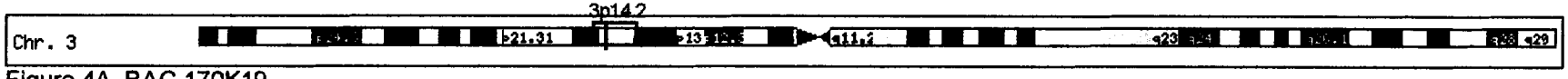


Figure 4A, BAC 170K19

Figure 4B, BAC 91A15

Figure 4. Location of Bacterial Artificial Chromosomes
Figure 4A displays the location of the BAC 170K19 along the short arm of chromosome 3 (3p). The 170K19 clone observed as the red band in figure 4A is 165721 bp and located within 3p14.2. Figure 4B displays the location of the 91A15 clone along chromosome 3. The 91A15 clone observed as the red band in figure 4B is 164150 bp long and located within the centromeric regions of the short arm of chromosome 3.

6.4. Statistical Analysis

6.4.1. Analysis of Demographic Data

When the gender of the individuals among study groups was compared, the Pearson chi-square test was applied. Comparisons of age and mean pack-years (for smokers only) were performed using t-tests. To incorporate the tobacco exposure due to cigars and pipes into the analysis of cigarette pack years, pipe and cigar exposure was converted into pack-years equivalents (1 pipe = 1.5 cigars = 3 cigarettes) (Jourenkova-Mironova *et al*, 1999). All statistical analysis was performed using JMP software version 5.0 (SAS Institute Inc., Cary INC, USA). The limit for significance for all comparisons was set at $P = 0.05$.

6.4.2. Analysis of FISH Results

Normalization was required to account for the different sample sizes of the study groups, and the non-normal distribution of percent data. The percent data was normalized using the equation $\sin^{-1}$(square root (% cells)) for all comparisons between the percent of cells that fell within the five abnormality groups among the cancer patients, smokers, and non-smokers: (1) alterations limited to the number of centromere 3 signals, (2) alterations limited to the number of 3p14.2, (3) any alteration in the number of centromere 3 signals, (4) any alteration in the number of 3p14.2 signals, and (5) alterations to both centromere 3 and 3p14.2.

To analyze the FISH data from the three study groups (tumor margins, smokers, nonsmokers) for a difference in the number of cells within the five

abnormality groups of alterations, Oneway Analysis of Variance comparisons were used to determine if there was a significant difference among any of the five abnormality group. Significance for the one-way analysis of variance was set at $P = 0.05$. If a significant difference was found to exist, Tukey's multiple comparison test was applied to determine where among the means of the five abnormality groups a significant difference existed where the level of significance was set *at* $P = .05$.

7. RESULTS

7.1. FISH Patterns in Cancer Patients

Table 2 presents the demographic and smoking information for the 19 cancer patients included in this study as positive controls. The average age for these patients was 62 years (range, 38 to 90), with 58% being male. Sixty-three percent (12 cases) had a history of tobacco use and four individuals reported a current cigarette habit. Of those who had used tobacco, the average number of pack-years was 44.8 (range, 0.3 to 110).

Two hundred cells were scored per sample if the sample size was sufficient; otherwise effort was made to score all evaluable cells on the slide. Of the 19 cancer cases, 8 samples had at least 200 cells scored. The average number of cells scored per slide among the tumor samples was 136 (range, 14 to 221). There are several explanations for the variation of cell numbers in these samples. Since they were obtained by brushing the circumference of the lesion, the total area brushed varied considerably. The location of the lesion site could have also affected the sample size because not all areas of the oral mucosa yield the same amount of epithelial cells. In addition, variation in sample collection over time could have affected cell numbers. Three different individuals collected these samples over a period of 2 years. A total of 20 different patterns were observed among all cases (cancer patients, smokers, and nonsmokers) and 17 of these patterns were observed in the tumor margin samples (Table 3). These

17 patterns included the normal diploid pattern (i.e., 2 centromeres and 2 3p14.2 signals, scored as 2,2) and 16 abnormal patterns.

Figure 5A is a graphic representation of the mean frequency of each of the abnormal patterns in the cancer group, providing a quick view of the relative prevalence of these patterns in this group. The data was also graphed to show the proportion of cases that had each of the abnormal patterns (Figure 6).

Table 2. Demographics, smoking history, and sample cell number for cancer cases

Patient #	Age (yrs)	Gender	Smoking history[a]	Smoking habit[b]	Pack-years	Number of cells scored[c]
1031	90	F	1	FS	18	107
1053	71	M	1	S	93	14
1054	61	M	1	FS	13	157
1185	55	F	1	S	9	217
1262	47	F	0	NS	0	200
1379	69	M	0	NS	0	214
1384	51	F	1	FS	0.3	201
1388	86	M	1	FS	110	70
1390	57	M	1	S	67	221
1396	68	F	0	NS	0	80
1397	56	F	0	NS	0	22
1407	71	M	0	NS	0	133
1408	55	M	1	FS	18	92
1410	55	F	1	FS	61	118
1412	60	M	1	S	81	105
1421	57	F	0	NS	0	17
1424	38	M	0	NS	0	201
1432	59	M	1	FS	38	203
1440	69	M	1	FS	30	212

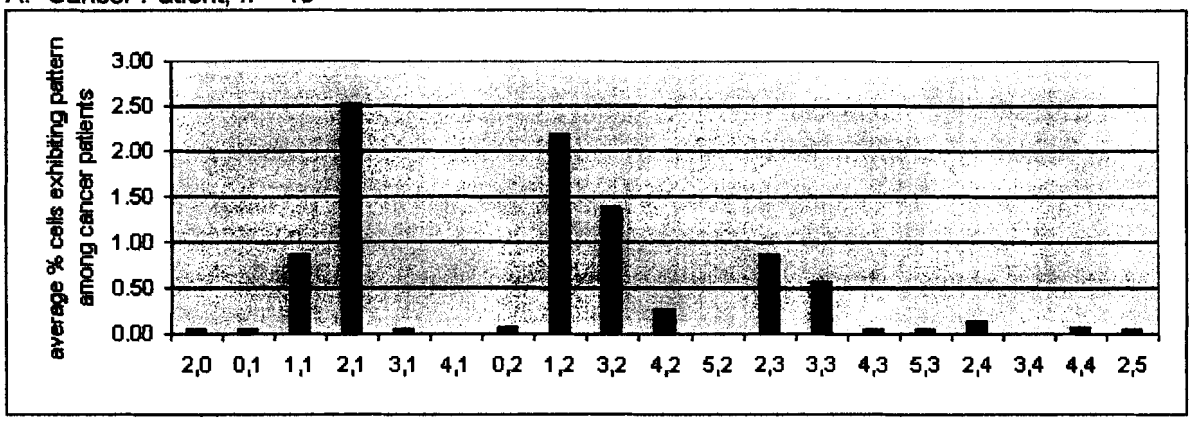
[a] 1=smoker (ever), 0=Nonsmoker
[b] S,current smoker; FS former smoker, NS nonsmokers
[c] Includes only cells that match criteria described in section 6.3.4.
M indicates male gender of study participant
F indicates female gender of study participant

Table 3. FISH patterns for cancer patients

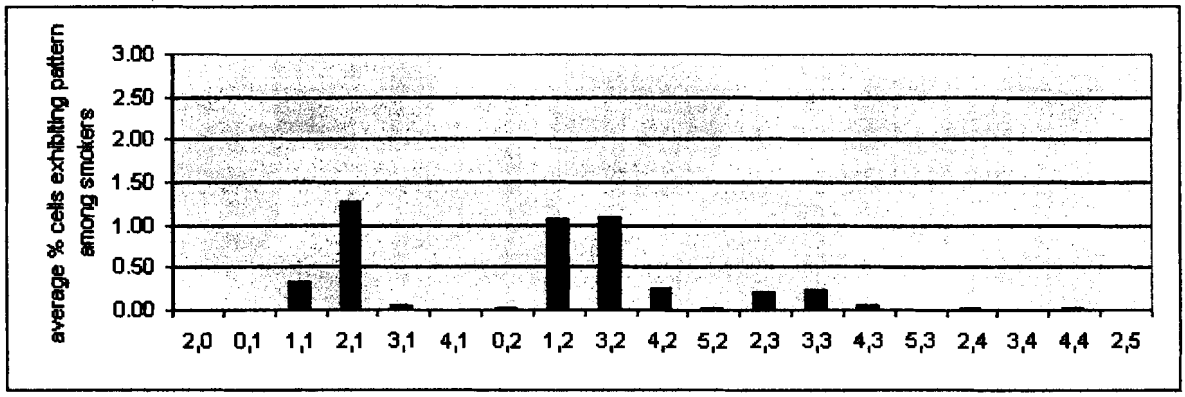
Patient #	Number of cells scored	Percent of cells exhibiting indicated FISH pattern[a]																			
		2,2	2,0	0,1	1,1	2,1	3,1	4,1	0,2	1,2	3,2	4,2	5,2	2,3	3,3	4,3	5,3	2,4	3,4	4,4	2,5
1031	107	87.85	0.00	0.93	2.80	1.87	0.00	0.00	0.00	3.74	0.93	0.00	0.00	0.93	0.93	0.00	0.00	0.00	0.00	0.00	0.00
1053	14	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1054	157	90.45	0.00	0.00	1.27	3.18	0.00	0.00	0.00	3.18	0.64	0.00	0.00	0.64	0.64	0.00	0.00	0.00	0.00	0.00	0.00
1185	217	90.32	0.00	0.00	1.38	2.30	0.00	0.00	0.00	2.30	0.92	0.00	0.00	0.92	1.38	0.00	0.00	0.00	0.00	0.46	0.00
1262	200	88.50	0.50	0.00	1.50	3.00	0.00	0.00	0.50	4.00	1.50	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1379	214	86.92	0.00	0.00	1.87	3.27	0.93	0.00	0.00	2.34	2.34	0.93	0.00	0.47	0.93	0.00	0.00	0.00	0.00	0.00	0.00
1384	201	92.04	0.00	0.00	0.00	3.48	0.00	0.00	0.00	1.00	1.99	1.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1388	70	90.00	0.00	0.00	1.43	1.43	0.00	0.00	0.00	2.86	2.86	0.00	0.00	0.00	0.00	0.00	0.00	1.43	0.00	0.00	0.00
1390	221	90.50	0.00	0.00	1.81	1.36	0.00	0.00	0.00	3.62	1.36	0.45	0.00	0.45	0.45	0.00	0.00	0.00	0.00	0.00	0.00
1396	80	88.75	0.00	0.00	0.00	2.50	0.00	0.00	0.00	2.50	3.75	0.00	0.00	0.00	2.50	0.00	0.00	0.00	0.00	0.00	0.00
1397	22	90.91	0.00	0.00	0.00	0.00	0.00	0.00	0.00	4.55	4.55	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1407	133	86.47	0.00	0.00	0.75	2.26	0.00	0.00	0.75	3.01	2.26	0.00	0.00	1.50	0.75	0.75	0.75	0.00	0.00	0.00	0.75
1408	92	95.65	0.00	0.00	0.00	2.17	0.00	0.00	0.00	1.09	0.00	1.09	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1410	118	88.98	0.00	0.00	0.00	3.39	0.00	0.00	0.00	1.69	0.00	0.00	0.00	4.24	0.85	0.00	0.00	0.00	0.00	0.85	0.00
1412	105	88.57	0.00	0.00	1.90	0.00	0.00	0.00	0.00	1.90	0.95	0.95	0.00	2.86	1.90	0.00	0.00	0.95	0.00	0.00	0.00
1421	17	88.24	0.00	0.00	0.00	11.76	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1424	201	93.03	0.00	0.00	1.00	1.49	0.00	0.00	0.00	2.49	0.00	0.00	0.00	1.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1432	203	93.60	0.00	0.00	0.00	2.46	0.00	0.00	0.00	0.00	1.48	0.49	0.00	1.48	0.49	0.00	0.00	0.00	0.00	0.00	0.00
1440	212	94.34	0.47	0.00	0.94	1.89	0.00	0.00	0.00	1.42	0.94	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average %	136	90.80	0.05	0.05	0.88	2.52	0.05	0.00	0.07	2.19	1.39	0.26	0.00	0.87	0.57	0.04	0.04	0.13	0.00	0.07	0.04
# samples with pattern		19	2	1	11	16	1	0	2	16	14	6	0	12	10	1	1	2	0	2	1
% samples with pattern		100	10.5	5.8	57.9	84.2	5.8	0	10.5	84.2	73.7	31.6	0	63.2	52.6	5.8	5.8	10.5	0	10.5	5.8

[a] *FISH patterns shown represent number of hybridization signals visualized with probes using the BAC 170K19 to target centromere 3 (first number of ratio (green)) and the BAC 91A15 to target 3p14.2 (second number of ratio (red)).*

A. Cancer Patient, n = 19



B. Smokers, n = 30



C. Nonsmokers, n = 29

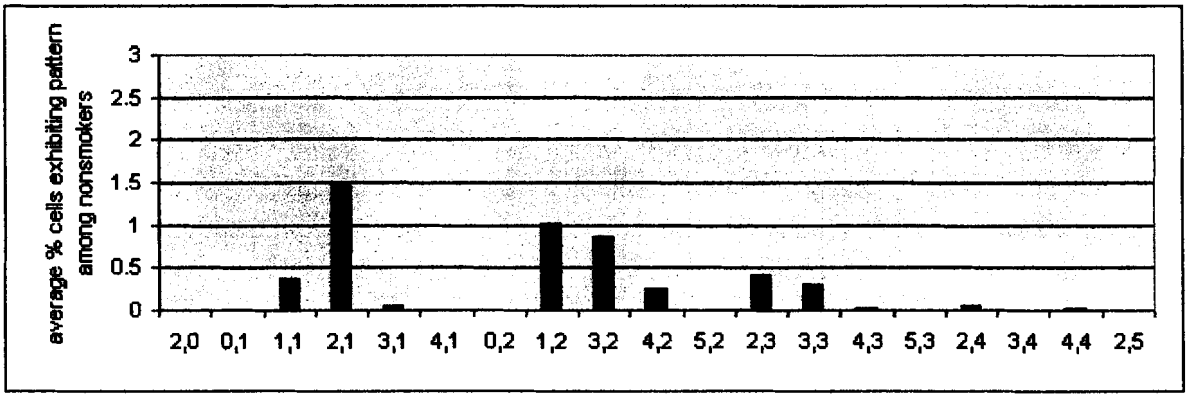
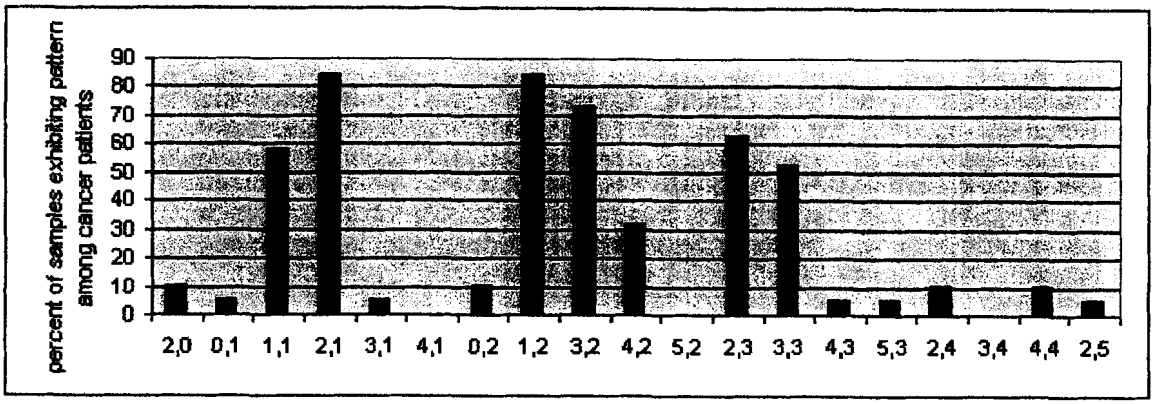


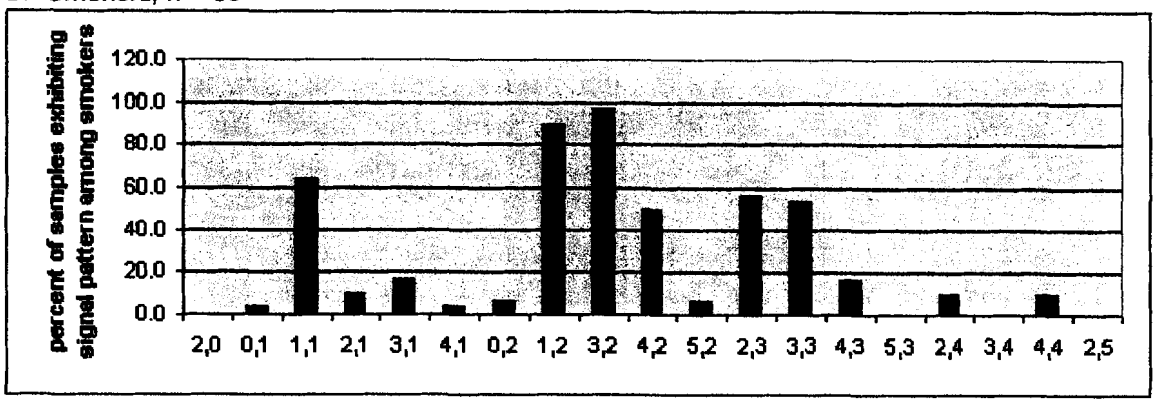
Figure 5. The average percentage of cells with abnormal signal patterns from sample groups.

For each signal pattern there is a single bar representing the percent of cells exhibiting the abnormal signal pattern among study group

A. Cancer Patients, n = 19



B. Smokers, n = 30



C. Nonsmokers, n = 29

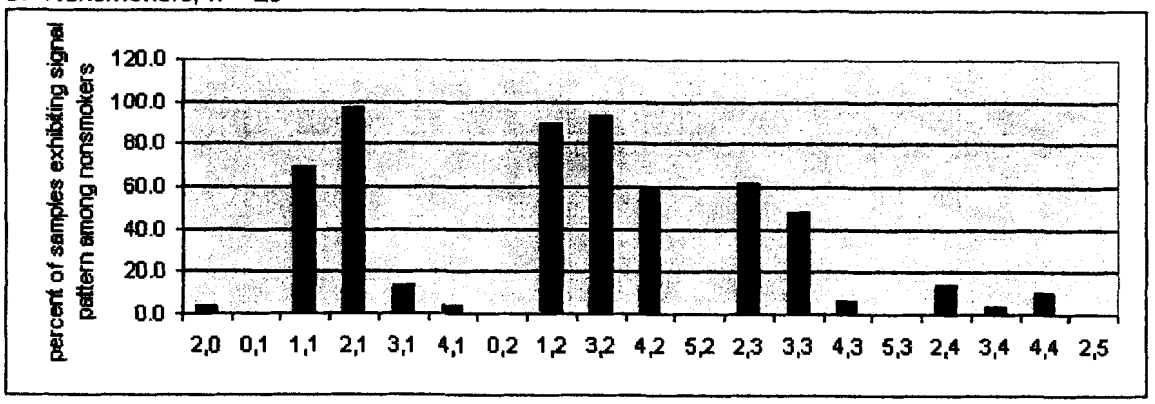


Figure 7. The average percent of samples exhibiting abnormal signal patterns among study groups.

For each signal pattern there is a single bar representing the percent of samples exhibiting that pattern among all samples evaluated for that group

As shown in Table 3, the most frequent pattern observed for cancer samples was the normal diploid pattern (called 2,2 for two *FHIT* signals and two CEP3 signals). The mean frequency for this pattern among the cancer samples was 90.8%, leaving 9.2% of cells with abnormal patterns. The most frequent abnormal signal patterns were 2,1 and 1,2 both found in 16 of 19 patients. The mean frequencies for these 2 patterns were 1.47% and 1.01% respectively. Other abnormal patterns, in decreasing order were: 2,3 (present in 12 patients); 1,1 (present in 11 patients); 3,3 (10 patients); 4,2 (6 patients); 4,4, 2,4, 0,2, and 2,0 each found in 2 patients; 3,1 (1 patient); and finally 0,1, 4,3, 5,3, and 2,5 each only found in one patient. The mean percentages of cells exhibiting these abnormal patterns was: 1,1 (0.88%); 3,2 (1.39%); 2,3 (0.87%); 3,3 (0.57%); 4,2 (0.26%); 2,4 (0.13%); 4,4 and 0,2 (0.07% each); 3,1 (0.05%); 0,1(.05%); and 4,3, 5,3, and 2,5 (0.04%) each.

Copy number alteration to *FHIT* can occur by different mechanisms. Alteration to centromere number is generally equated with whole chromosome change or aneuploidy. Such alteration indirectly causes a parallel change in *FHIT* copy number. Other mechanisms produce alterations at the sub-chromosomal level (e.g., breaks, deletions, recombinations) resulting in an increase or decrease in copy number that is restricted to chromosome segments that contain the *FHIT* region (3p14.2). In Table 4, five different patterns of copy number alterations were considered: Group 1, cells with alterations limited to the number of CEP 3 signals, i.e., abnormal number of centromere signals but

normal 3p14.2 signals (2,0; 2,1; 2,3, 2,4, and 2,5); Group 2, cells with alterations to the 3p14.2 region only, i.e. abnormal number of *FHIT* signals but 2 centromere signals (0,2; 1,2; 3,2; 4,2; 5,2); Group 3, any alteration in the number of CEP 3 signals, i.e., abnormal number of centromere signals with the 3p14.2 number not considered (2,0; 0,1; 1,1; 2,1; 3,1;4,1; 2,3; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5); Group 4, any alteration to 3p14.2, i.e., abnormal number of 3p14.2 signals with CEP 3 not considered (0,1; 1,1; 3,1; 4,1; 0,2; 1,2; 3,2; 4,2; 5,2; 3,3; 4,3; 5,3; 3,4; 4,4); and Group 5 involving alterations to both the number of CEP 3 and 3p14.2 signals in the same cells (0,1; 1,1; 3,1; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5. The most common group of abnormalities was Group 4 (any alteration to 3p14.2). Seventeen of the 19 cancer samples had such change, with an average frequency of 5.6%. In contrast, Group 5 (alteration to both probes) was very infrequent, present in only 0.63% with only 11 of the samples showing such change. The remaining groups showed an intermediate level of alteration with 5.3% (Group 3, any alteration to CEP 3), 3.9% (Group 2, 3p14.2 only) and 3.6% (Group 1, CEP 3 only) of cells showing such change.

Table 4. Frequencies of aneuploidy and 3p14.2 (*FHIT* locus) alterations in cancer patients

OHS#	Total # Cells Counted	% Cells with normal pattern (2,2)	% Cells with abnormal pattern	Group1: Alterations limited to CEP 3[a]	Group 2: Alterations limited to 3p14[b]	Group 3: Any alteration to CEP 3[c]	Group 4: Any alteration to 3p14[d]	Group 5: Alteration to CEP 3 and 3p14[e]
1031	107	87.85	12.15	2.80	4.67	7.48	9.35	0.93
1053	14	100.00	0.00	0.00	0.00	0.00	0.00	0.00
1054	157	90.45	9.55	3.82	3.82	5.73	5.73	0.00
1185	217	90.32	9.68	3.23	3.23	6.45	6.45	0.46
1262	200	88.50	11.50	4.00	6.00	5.50	7.50	0.00
1379	214	86.92	13.08	3.74	5.61	7.48	9.35	1.87
1384	201	92.04	7.96	3.98	3.98	3.98	3.98	1.00
1388	70	90.00	10.00	2.86	5.71	4.29	7.14	1.43
1390	221	90.50	9.50	1.81	5.43	4.07	7.69	0.45
1396	80	88.75	11.25	2.50	6.25	5.00	8.75	0.00
1397	22	90.91	9.09	0.00	9.09	0.00	9.09	0.00
1407	133	86.47	13.53	3.76	6.02	7.52	9.02	1.50
1408	92	95.65	4.35	2.17	2.17	2.17	2.17	1.09
1410	118	88.98	11.02	7.63	1.69	9.32	3.39	0.85
1412	105	88.57	11.43	3.81	3.81	7.62	7.62	1.90
1421	17	88.24	11.76	11.76	0.00	11.76	0.00	0.00
1424	201	93.03	6.97	3.48	2.49	4.48	3.48	0.00
1432	203	93.60	6.40	3.94	1.97	4.43	2.46	0.49
1440	212	94.34	5.66	2.36	2.36	3.30	3.30	0.00
Average	136	90.8	9.2	3.56	3.91	5.29	5.6	0.63

[a] *The following FISH patterns were included: 2,0; 2,1; 2,3; 2,4; 2,5*

[b] *Patterns: 0,2; 1,2; 3,2; 4,2; 5,2*

[c] *Patterns: 2,0; 0,1; 1,1; 2,1; 3,1;4,1; 2,3; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

[d] *Patterns: 0,1; 1,1; 3,1; 4,1; 0,2; 1,2; 3,2; 4,2; 5,2; 3,3; 4,3; 5,3; 3,4; 4,4*

[e] *Patterns: 0,1; 1,1; 3,1; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

7.2. FISH Patterns in Smokers

Thirty samples were collected from individuals with tobacco exposure, but no history or clinical signs of oral cancer or precancer. Table 6 shows the demographic information and smoking history for these cases. The average age of the smokers was 49 years (range, 21 to 80 years) with 40% being male. Although all individuals had a smoking history, only 20% of the cases were current smokers. The average pack-year exposure was 25.6 (range, 1.5-81 pack years).

Table 7 shows a comparison of gender distribution, age, and tobacco usage for the cancer and smoker groups. Gender distributions were similar (P = 0.22), but age and tobacco exposure were significantly different. Cancer patients were significantly older than individuals in the smoker group ($P = 0.0007$). The average age of the cancer patients was 61.8 years (range, 38 to 90), while that of the smokers was 48.4 years (range, 21 to 80). Cancer patients also reported a significantly higher level of tobacco exposure when compared to the smokers ($P = 0.0366$), with an average of 44.7 pack years for the former compared to 25.6 pack-years among the latter.

Table 5. Demographics, smoking history and sample quality data for smokers

Patient #	Age (yrs)	Gender	Smoking History[a]	Smoking Habit[b]	Pack-years	Number of cells counted[c]
8528	74	F	1	FS	53	294
8529	51	F	1	FS	7	225
8532	56	M	1	FS	11.9	504
8533	31	F	1	S	11.3	509
8542	80	F	1	FS	28	319
8549	23	F	1	S	3.3	387
8550	21	M	1	S	7.8	514
8555	48	M	1	S	33.5	505
8556	48	M	1	S	81	500
8557	58	F	1	FS	30.8	518
8562	53	M	1	FS	50	500
8565	62	F	1	S	41.5	522
8566	34	M	1	S	24.5	400
8567	31	F	1	S	50.6	497
8572	52	F	1	FS	1.5	201
8573	53	M	1	FS	22.5	505
8575	30	F	1	FS	3.3	521
8577	47	F	1	FS	8.4	497
8578	53	M	1	FS	60	484
8579	42	F	1	FS	15.4	401
8580	49	F	1	S	25	500
8582	55	M	1	FS	40	491
8584	47	M	1	S	40	500
8586	63	F	1	FS	25.4	500
8587	41	F	1	FS	4	156
8589	57	F	1	FS	25.5	123
8591	52	M	1	FS	15	469
8595	66	M	1	FS	21.5	500
8596	22	F	1	S	1.9	500
8598	53	F	1	S	35.8	500

[a] *1 indicates a history of smoking, 0 indicates no history of smoking*
[b] *S, current smoker; FS, former smoker*
[c] *Includes only cells that match criteria described in section 6.3.4.*
M indicates male gender of study participant
F indicates female gender of study participant

Table 6. Comparison of demographic data between study groups

		Cancer Patients	Smokers	non-smoking controls
Total	# Cases	19	30	29
Gender	% Male	57.89	40.00	41.00
Age	Mean yrs	61.84*	48.4	45.14
	SD	12.53	14.49	13.85
	Range	38 - 90	21 - 80	20 - 78
Tobacco Use	Ever smokers %	63.16	100	0
	Current smokers %	21.05	20.34	NA
	Mean, pack-yrs	44.78*	25.64	NA
	SD, pack-yrs	36.44	19.43	NA
	Range, pack-yrs	0.25 - 110	1.5 - 81	NA

** all P values determined with a T-test*

**indicates a significant difference from other study groups*

A larger number of cells were available for scoring for smoker samples compared to the previously discussed cancer samples. This was primarily due to the larger surface area that was brushed in smokers. A decision was made to increase the number of cells evaluated for smokers to 500 cells whenever the sample size would allow it, in order to better identify cells with abnormal patterns since frequencies of abnormal cells were low in these specimens. Of the 30 samples, 16 contained a sufficient sample size to allow at least 500 cells to be counted on the hybridized slides; for others all evaluable cells were scored, with the number ranging from 123 to 497 cells.

5B shows the mean frequency of each of the abnormality patterns in the smoker group with data presented in Table 7. As expected, smokers had a lower frequency of cells with abnormal patterns, present in an average of 5 % of cells compared with 9% for cancer cases. Strikingly, one of the patterns 2,1 was present in each of the smoker samples albeit at a low frequency (mean of 1% of cells). Other abnormal patterns, in order of decreasing frequency were: 3,2 (29 samples, 1.1% cells); 1,2 (27 samples, 1.1% cells exhibited this pattern on average); 1,1 (19 samples, 0.33% cells); 4,2 (15 samples, 0.25% cells), 2,3 (17 individuals, 0.22% cells); 3,3 (16 samples, 0.23% cells); 3,1 and 4,3 (0.04% cells each); 4,4 (0.03% cells); 5,2 (0.3%) ; 0,2 and 2,4 (0.02% cells each); and finally 0,1 and 4,1 FISH signal were found in 0.01% of cells each.

Table 7. FISH patterns in smokers

Patient #	Number of cells scored	Percent of cells exhibiting indicated FISH pattern[a]																			
		2,2	2,0	0,1	1,1	2,1	3,1	4,1	0,2	1,2	3,2	4,2	5,2	2,3	3,3	4,3	5,3	2,4	3,4	4,4	2,5
8528	294	94.90	0.00	0.00	0.34	2.38	0.00	0.00	0.00	1.36	0.68	0.00	0.00	0.00	0.34	0.00	0.00	0.00	0.00	0.00	0.00
8529	225	94.22	0.00	0.00	0.89	1.78	0.00	0.00	0.00	2.67	0.44	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8532	504	96.03	0.00	0.00	0.00	1.19	0.00	0.00	0.00	0.99	1.39	0.00	0.00	0.20	0.00	0.20	0.00	0.00	0.00	0.00	0.00
8533	509	94.30	0.00	0.00	0.39	0.59	0.20	0.00	0.00	2.16	0.98	0.59	0.00	0.00	0.39	0.00	0.00	0.00	0.00	0.39	0.00
8542	319	96.24	0.00	0.00	0.00	0.94	0.00	0.00	0.00	0.94	1.57	0.00	0.00	0.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8549	387	94.06	0.00	0.00	0.26	0.78	0.26	0.00	0.00	0.78	1.29	1.55	0.00	0.78	0.00	0.26	0.00	0.00	0.00	0.00	0.00
8550	514	94.16	0.00	0.00	0.19	0.97	0.00	0.00	0.00	1.17	1.17	0.78	0.00	0.78	0.19	0.39	0.00	0.19	0.00	0.00	0.00
8555	505	95.05	0.00	0.00	0.20	0.99	0.00	0.00	0.00	2.18	0.99	0.20	0.00	0.40	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8556	500	94.20	0.00	0.00	0.00	1.80	0.00	0.00	0.00	1.00	1.20	0.40	0.00	0.20	1.20	0.00	0.00	0.00	0.00	0.00	0.00
8557	163	95.71	0.00	0.00	0.00	1.84	0.00	0.00	0.00	0.00	0.61	0.00	0.61	0.00	1.23	0.00	0.00	0.00	0.00	0.00	0.00
8562	500	94.40	0.00	0.00	1.60	1.20	0.20	0.00	0.00	1.40	1.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00
8565	522	95.79	0.00	0.00	0.19	1.34	0.00	0.00	0.00	0.38	1.72	0.19	0.00	0.19	0.00	0.19	0.00	0.00	0.00	0.00	0.00
8566	400	96.00	0.00	0.25	0.75	1.25	0.00	0.00	0.00	1.00	0.50	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8567	497	94.77	0.00	0.00	0.20	0.80	0.00	0.00	0.00	1.01	2.01	0.60	0.00	0.00	0.40	0.00	0.00	0.00	0.00	0.00	0.00
8572	201	96.52	0.00	0.00	0.00	1.99	0.00	0.00	0.00	0.50	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8573	505	95.25	0.00	0.00	0.79	0.59	0.00	0.00	0.00	1.58	0.59	0.20	0.00	0.59	0.20	0.00	0.00	0.00	0.00	0.20	0.00
8575	521	95.01	0.00	0.00	0.58	1.34	0.00	0.00	0.00	1.15	0.96	0.00	0.38	0.00	0.38	0.19	0.00	0.00	0.00	0.00	0.00
8577	497	93.56	0.00	0.00	0.40	1.61	0.00	0.20	0.00	1.41	1.01	0.20	0.00	1.01	0.20	0.00	0.00	0.00	0.00	0.40	0.00
8578	484	95.45	0.00	0.00	0.21	1.24	0.00	0.00	0.00	0.41	1.86	0.41	0.00	0.00	0.41	0.00	0.00	0.00	0.00	0.00	0.00
8579	401	96.01	0.00	0.00	1.00	1.25	0.00	0.00	0.00	1.00	0.75	0.00	0.00	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8580	500	95.40	0.00	0.00	0.00	1.00	0.40	0.00	0.00	1.20	0.80	0.60	0.00	0.20	0.40	0.00	0.00	0.00	0.00	0.00	0.00
8582	491	94.70	0.00	0.00	1.43	1.22	0.00	0.00	0.00	1.43	0.61	0.00	0.00	0.20	0.20	0.00	0.00	0.20	0.00	0.00	0.00
8584	500	96.80	0.00	0.00	0.20	0.60	0.00	0.00	0.00	0.40	0.80	0.00	0.00	0.60	0.20	0.00	0.00	0.20	0.00	0.00	0.00
8586	500	96.80	0.00	0.00	0.20	1.20	0.00	0.00	0.00	0.40	1.00	0.20	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00
8587	156	96.79	0.00	0.00	0.00	1.92	0.00	0.00	0.00	1.28	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8589	123	96.75	0.00	0.00	0.00	0.81	0.00	0.00	0.00	0.00	1.63	0.81	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8591	469	95.31	0.00	0.00	0.00	1.71	0.00	0.00	0.21	0.85	1.49	0.00	0.00	0.21	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8595	500	96.60	0.00	0.00	0.00	1.20	0.00	0.00	0.00	0.40	1.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8596	500	95.60	0.00	0.00	0.00	1.20	0.00	0.00	0.00	0.00	1.80	0.60	0.00	0.20	0.60	0.00	0.00	0.00	0.00	0.00	0.00
8598	500	93.20	0.00	0.00	0.20	1.40	0.20	0.00	0.40	3.00	1.20	0.20	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average %	422.00	95.32	0.00	0.01	0.33	1.27	0.04	0.01	0.02	1.07	1.10	0.25	0.03	0.22	0.23	0.04	0.00	0.02	0.00	0.03	0.00
# samples with pattern		30	0	1	19	30	5.00	1	2	27	29	15	2	17	16	5	0	3	0	3	0
% samples with pattern		100.0	0.0	3.3	63.3	10.0	16.70	3.3	6.7	90.0	96.7	50.0	6.7	56.7	53.3	16.7	0.0	10.0	0.0	10.0	0.0

[a] *FISH patterns shown represent number of hybridization signals visualized with probes using the BAC 170K19 to target centromere 3 (first number of ratio) and the BAC 91A15 to target 3p14.2 (second number of ratio).*

Next, the different abnormal FISH patterns were categorized into the same five groups used in the analysis of the cancer samples (Table 8). A significantly fewer number of cells were present in Groups 1 through 4 for smokers compared to cancer cases. Group1 alterations (CEP 3 only) were present in only 1.5% of the cells compared to 3.5% of cells in the cancer group ($P < 0.0001$). Group 2 (3p14 only) and group 3 (any alteration to 3p14.2) alteration were present in 2.47% and 2.23% of smoker samples respectively compared to 3.9% and 5.3% of cancer cases ($P = 0.0019$ and $P < 0.0001$ respectively). Finally, Group 4 alterations were present in 3.1% of cells in smokers compared to 5.6% in cancer cases, again significantly less ($P < 0.0001$). In contrast, similar frequencies were observed for Group 5 (alterations to both probes), most likely due to the low frequency of occurrence in both groups (0.7% and 0.63% of cases).

Table 8. Frequencies of aneuploidy and 3p14 alterations in smokers

OHS#	Total # Cells Counted	% Cells with normal pattern (2,2)	% Cells with abnormal pattern	Group1: Alterations limited to CEP 3[a]	Group 2: Alterations limited to 3p14[b]	Group 3: Any alteration to CEP 3[c]	Group 4: Any alteration to 3p14[d]	Group 5: Alteration to CEP 3 and 3p14[e]
8530	91	95.6	4.4	1.1	1.1	3.3	2.2	1.1
8531	500	95.4	4.6	1.4	2	2.6	2.2	0.2
8536	260	95.4	4.6	1.9	2.3	2.3	2.7	0.4
8538	191	91.1	8.9	2.6	3.1	5.8	4.7	1.6
8540	506	92.7	7.3	1.8	3.2	4.2	4.3	1.4
8541	508	96.3	3.7	1.4	1	2.8	1.8	0.8
8543	502	96	4	1.4	2	2	2.4	0.4
8544	101	97	3	0	1	2	3	2
8545	317	95	5	1.9	1.9	3.2	2.8	1.3
8546	350	96.9	3.1	1.7	1.4	1.7	1.4	0
8548	517	96.5	3.5	1.4	1.9	1.5	1.9	0
8551	258	95.7	4.3	1.9	1.6	2.7	1.9	0.4
8553	471	95.3	4.7	1.7	2.5	2.1	3	0.4
8558	510	94.5	5.5	1	1.8	3.7	3.7	2
8559	500	94.8	5.2	1.6	2.4	2.8	3	0.6
8560	501	95.8	4.2	1	2.2	2	2.8	0.6
8561	406	94.8	5.2	1.5	3.7	1.5	3.7	0
8563	234	96.6	3.4	0.9	1.3	2.1	2.6	1.3
8564	118	92.4	7.6	3.4	2.5	5.1	3.4	0.8
8568	501	95.8	4.2	1.2	2.6	1.6	2.8	0.2
8569	270	97.8	2.2	0.7	1.1	1.1	1.5	0.4
8576	502	95.8	4.2	0.6	2.2	1.8	2.8	0.6
8585	318	95	5	1.6	2.5	2.5	3.5	0.9
8590	494	94.7	5.3	1.8	2.4	2.8	3.2	0.8
8592	407	93.4	6.6	1.2	3.4	3.2	4.7	1.5
8593	346	95.1	4.9	2.3	2	2.9	2.6	0.6
8599	314	93.6	6.4	1.9	2.2	4.1	3.5	1.6
8600	153	95.4	4.6	2	1.3	3.3	1.3	0
8601	178	95.5	4.5	1.1	2.8	1.7	3.4	0.6
Average	356	95.17	4.83	1.52	2.12	2.7	2.85	0.77
# Patients with pattern		29	29	28	29	29	29	25

[a] *The following FISH patterns were included: 2,0; 2,1; 2,3; 2,4; 2,5*

[b] *Patterns: 0,2; 1,2; 3,2; 4,2; 5,2*

[c] *Patterns: 2,0; 0,1; 1,1; 2,1; 3,1;4,1; 2,3; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

[d] *Patterns: 0,1; 1,1; 3,1; 4,1; 0,2; 1,2; 3,2; 4,2; 5,2; 3,3; 4,3; 5,3; 3,4; 4,4*

[e] *Patterns: 0,1; 1,1; 3,1; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

7.3. FISH Patterns in Non-smokers

The final sample group included in this study was nonsmokers whom had never regularly smoked or chewed any tobacco product. The 29 non-smoking samples were matched for age (+/- 5 years) and gender to the smoker samples so there was no significant difference in age or gender among the smokers and nonsmokers. Non-smokers were included in this study for two major reasons: (1) to provide baseline values for naturally occurring alterations at centromere 3 and 3p14.2; and (2) to statistically determine the range of naturally occurring alterations at centromere 3 and 3p14.2 that could then be used to identify smokers with significantly elevated levels of genetic alterations at these genetic loci.

The demographic data for the non-smoker study group is summarized in Table 9. The average age was 45 years (range, 20 to 78 years) with 41.4% of the cases male. Gender distributions were similar among the non-smokers and cancer patients (Table 10, $P = 0.26$); however, the cancer patients were significantly older with an average age 61 years (range, 38 to 90 years) compared with 45 years (range, 20 to 78 years) for non-smokers ($P < 0.0001$).

Table 9. Demographics, smoking history and sample quality data for non-smokers

Patient #	Age (yrs)	Gender	Number of cells counted[a]
8530	29	F	91
8531	22	M	500
8536	37	M	260
8538	37	F	191
8540	55	F	506
8541	22	F	508
8543	35	F	502
8544	78	M	101
8545	54	M	317
8546	22	F	350
8548	45	F	517
8551	50	F	258
8553	55	F	471
8558	43	M	510
8559	33	F	500
8560	20	F	501
8561	39	M	406
8563	59	M	242
8564	60	F	118
8568	48	M	501
8569	53	F	270
8576	49	M	502
8585	44	M	318
8590	52	F	494
8592	47	M	407
8593	53	F	346
8599	58	F	314
8600	62	F	153
8601	48	M	178

[a] *Includes only cells that match criteria described in section 6.3.4.*
M indicates male gender of study participant
F indicates female gender of study participant

The number of cells counted was significantly greater among non-smoking samples compared to cancer cases although numbers did not significantly differ from those scored for smokers. An average of 365 cells (range, 91 to 517) was scored per sample. Figure 5C shows the mean frequencies for abnormal patterns among non-smokers, while figure 6C shows the percent of samples exhibiting the abnormal FISH patterns. The data used to create these figures is presented in Table 10.

Table 10. FISH patterns for nonsmokers

Patient #	Number of cells scored	Percent of cells exhibiting FISH pattern[a]																		
		2,2	2,0	0,1	1,1	2,1	3,1	4,1	0,2	1,2	3,2	4,2	5,2	2,3	3,3	4,3	5,3	2,4	3,4	4,4	2,5
8530	91	95.60	0.00	0.00	0.00	1.10	0.00	0.00	0.00	0.00	1.10	0.00	0.00	1.10	1.10	0.00	0.00	0.00	0.00	0.00	0.00
8531	500	95.40	0.00	0.00	0.20	1.40	0.00	0.00	0.00	0.40	0.80	0.80	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8536	260	95.38	0.00	0.00	0.38	1.92	0.00	0.00	0.00	1.15	0.77	0.38	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8538	191	91.10	0.00	0.00	0.00	2.09	0.00	0.00	0.00	2.09	0.52	0.52	0.00	1.57	1.57	0.00	0.00	0.52	0.00	0.00	0.00
8540	506	92.69	0.00	0.00	0.59	1.78	0.00	0.00	0.00	1.98	0.99	0.20	0.00	0.99	0.59	0.00	0.00	0.00	0.00	0.20	0.00
8541	508	96.26	0.00	0.00	0.20	1.38	0.20	0.00	0.00	0.59	0.39	0.00	0.00	0.59	0.39	0.00	0.00	0.00	0.00	0.00	0.00
8543	502	96.02	0.00	0.00	0.40	1.39	0.00	0.00	0.00	1.20	0.40	0.40	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8544	101	97.03	0.00	0.00	1.98	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8545	317	94.95	0.00	0.00	0.63	1.89	0.00	0.00	0.00	1.58	0.32	0.00	0.00	0.00	0.32	0.00	0.00	0.00	0.00	0.32	0.00
8546	350	96.86	0.00	0.00	0.00	1.71	0.00	0.00	0.00	0.00	1.43	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8548	517	96.52	0.00	0.00	0.00	1.35	0.00	0.00	0.00	0.39	1.35	0.19	0.00	0.19	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8551	258	95.74	0.00	0.00	0.39	1.94	0.00	0.00	0.00	1.16	0.39	0.00	0.00	0.39	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8553	471	95.33	0.00	0.00	0.21	1.70	0.21	0.00	0.00	1.06	0.85	0.64	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8558	510	94.51	0.00	0.00	0.98	0.98	0.00	0.00	0.00	0.39	0.98	0.39	0.00	0.78	0.98	0.00	0.00	0.00	0.00	0.00	0.00
8559	500	94.80	0.00	0.00	0.20	1.40	0.00	0.00	0.00	1.20	0.80	0.40	0.00	0.60	0.40	0.00	0.00	0.20	0.00	0.00	0.00
8560	501	95.81	0.20	0.00	0.40	0.80	0.00	0.00	0.00	1.00	1.00	0.20	0.00	0.40	0.20	0.00	0.00	0.00	0.00	0.00	0.00
8561	406	94.83	0.00	0.00	0.00	1.48	0.00	0.00	0.00	1.97	1.48	0.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8563	234	96.58	0.00	0.00	0.00	0.85	0.00	0.00	0.00	0.43	0.85	0.00	0.00	0.00	1.28	0.00	0.00	0.00	0.00	0.00	0.00
8564	118	92.37	0.00	0.00	0.85	3.39	0.00	0.00	0.00	1.69	0.00	0.85	0.00	0.85	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8568	501	95.81	0.00	0.00	0.00	1.20	0.00	0.00	0.00	1.20	1.00	0.40	0.00	0.20	0.20	0.00	0.00	0.00	0.00	0.00	0.00
8569	270	97.78	0.00	0.00	0.37	0.74	0.00	0.00	0.00	1.11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8576	502	95.82	0.00	0.00	0.40	0.60	0.00	0.00	0.00	1.59	0.60	0.00	0.00	0.60	0.20	0.00	0.00	0.00	0.00	0.00	0.00
8585	318	94.97	0.00	0.00	0.63	1.26	0.00	0.00	0.00	0.94	0.94	0.63	0.00	0.00	0.31	0.00	0.00	0.31	0.00	0.00	0.00
8590	494	94.74	0.00	0.00	0.20	1.62	0.00	0.00	0.00	0.81	1.42	0.20	0.00	0.20	0.61	0.00	0.00	0.20	0.00	0.00	0.00
8592	407	93.37	0.00	0.00	0.49	1.23	0.00	0.25	0.00	1.97	1.23	0.25	0.00	0.49	0.25	0.25	0.00	0.00	0.25	0.00	0.00
8593	346	95.09	0.00	0.00	0.58	2.31	0.00	0.00	0.00	0.29	1.73	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8599	314	93.63	0.00	0.00	0.64	1.91	0.32	0.00	0.00	0.64	0.96	0.64	0.00	0.64	0.00	0.32	0.00	0.00	0.00	0.32	0.00
8600	153	95.42	0.00	0.00	0.00	1.96	0.00	0.00	0.00	0.65	0.65	0.00	0.00	1.31	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8601	178	95.51	0.00	0.00	0.00	1.12	0.56	0.00	0.00	1.69	1.12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Average % Cells		95.17	0.01	0.00	0.37	1.47	0.04	0.01	0.00	1.01	0.86	0.25	0.00	0.42	0.29	0.02	0.00	0.04	0.01	0.03	0.00
# samples with pattern		29	1	0	20	28	4	1	0	26	27	17	0	18	14	2	0	4	1	3	0
% samples with pattern		100.0	3.5	0.0	69.0	96.6	13.8	3.5	0.0	89.7	93.1	58.6	0.0	62.1	48.3	6.9	0.0	13.8	3.5	10.3	0.0

[a] *FISH patterns shown represent number of hybridization signals visualized with probes using the BAC 170K19 to target centromere 3 (first number of ratio) and the BAC 91A15 to target 3p14.2 (second number of ratio).*

The average number of cells with abnormal FISH patterns among nonsmokers was 4.8% (range, 2.2% to 8.9%). Overall, this was not significantly different from smokers ($P = 0.65$), however when comparing nonsmokers to cancer cases, the cancer patients had significantly elevated levels of abnormal signal patterns ($P < 0.0001$). Like the smoking cases, the most common abnormal FISH pattern among the nonsmoker samples was 2,1. This pattern was present in 28 of the 29 samples with an average of 1.47% of the cells showing this change. Other abnormal patterns observed among nonsmokers, in order of decreasing frequency were: 1,2 (26 samples, 1.01% cells); 3,2 (27 samples, 0.9% cells); 2,3 (18 samples, 0.4% cells); 1,1 (20 samples, 0.37% cells); 3,3(0.29%); 4,2 (0.25% cells); 3,1 and 2,4 (0.039% cells); 4,4 (0.29% cells); 4,3 (0.02% cells); and finally 2,0, 3,4, and 4,1 which were all present in 0.01% of the cells counted. As a comparison, Figure 7 shows the mean frequencies of different abnormal signal patterns in cancer patients, smokers, and nonsmokers.

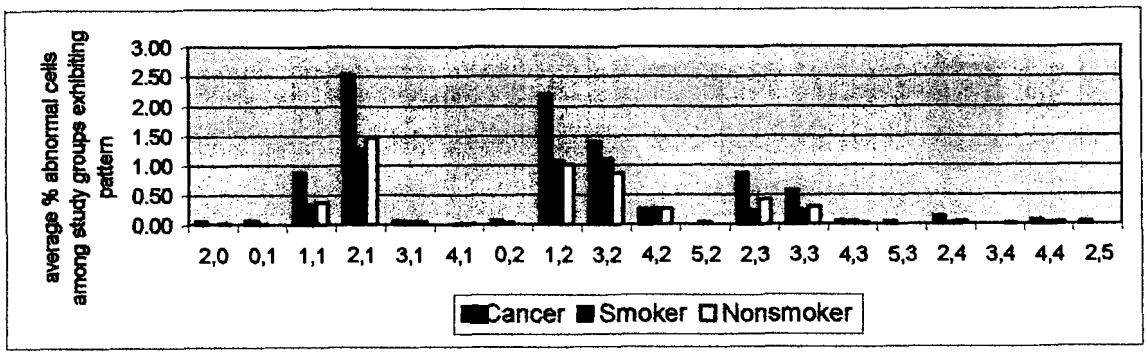


Figure 7. Percent abnormal FISH patterns among study groups.

For each signal pattern there are three bars each representing the average percent of cells exhibiting that signal pattern from its respective study group.

Finally we assessed the overall frequencies of aneuploidy and 3p14.2 alterations in the nonsmoking samples by categorizing the various FISH patterns into the five abnormality groups previously discussed (section 7.1). These results are summarized in Table 12 with comparisons of frequencies for cancer and smoking cases in Table 13. The results were similar to those described for smokers. Groups 1 through 4 had significantly fewer cells when compared with cancer values ($P < 0.0001$). When comparing the frequency of alterations between smokers and nonsmokers, no significant differences existed between the two study groups (Table 12). Again, as observed for both cancer and smoker samples, Group 5 (alterations to both CEP 3 and 3p14.2) contained the fewest cells, and no significant difference between the frequencies of cells that fell into this category among nonsmokers and cancer patients was found ($P= 0.5770$).

Table 11. Frequencies of aneuploidy and 3p14 alterations in non-smokers

OHS#	Total # Cells Counted	% Cells with normal pattern (2,2)	% Cells with abnormal pattern	Group1: Alterations limited to CEP 3[a]	Group 2: Alterations limited to 3p14[b]	Group 3: Any alteration to CEP 3[c]	Group 4: Any alteration to 3p14[d]	Group 5: Alteration to CEP 3 and 3p14[e]
8530	91	95.6	4.4	1.1	1.1	3.3	2.2	1.1
8531	500	95.4	4.6	1.4	2	2.6	2.2	0.2
8536	260	95.4	4.6	1.9	2.3	2.3	2.7	0.4
8538	191	91.1	8.9	2.6	3.1	5.8	4.7	1.6
8540	506	92.7	7.3	1.8	3.2	4.2	4.3	1.4
8541	508	96.3	3.7	1.4	1	2.8	1.8	0.8
8543	502	96	4	1.4	2	2	2.4	0.4
8544	101	97	3	0	1	2	3	2
8545	317	95	5	1.9	1.9	3.2	2.8	1.3
8546	350	96.9	3.1	1.7	1.4	1.7	1.4	0
8548	517	96.5	3.5	1.4	1.9	1.5	1.9	0
8551	258	95.7	4.3	1.9	1.6	2.7	1.9	0.4
8553	471	95.3	4.7	1.7	2.5	2.1	3	0.4
8558	510	94.5	5.5	1	1.8	3.7	3.7	2
8559	500	94.8	5.2	1.6	2.4	2.8	3	0.6
8560	501	95.8	4.2	1	2.2	2	2.8	0.6
8561	406	94.8	5.2	1.5	3.7	1.5	3.7	0
8563	234	96.6	3.4	0.9	1.3	2.1	2.6	1.3
8564	118	92.4	7.6	3.4	2.5	5.1	3.4	0.8
8568	501	95.8	4.2	1.2	2.6	1.6	2.8	0.2
8569	270	97.8	2.2	0.7	1.1	1.1	1.5	0.4
8576	502	95.8	4.2	0.6	2.2	1.8	2.8	0.6
8585	318	95	5	1.6	2.5	2.5	3.5	0.9
8590	494	94.7	5.3	1.8	2.4	2.8	3.2	0.8
8592	407	93.4	6.6	1.2	3.4	3.2	4.7	1.5
8593	346	95.1	4.9	2.3	2	2.9	2.6	0.6
8599	314	93.6	6.4	1.9	2.2	4.1	3.5	1.6
8600	153	95.4	4.6	2	1.3	3.3	1.3	0
8601	178	95.5	4.5	1.1	2.8	1.7	3.4	0.6
Average	356	95.17	4.83	1.52	2.12	2.7	2.85	0.77
# Patients with pattern		29	29	28	29	29	29	25

[a] *The following FISH patterns were included: 2,0; 2,1; 2,3; 2,4; 2,5*

[b] *Patterns: 0,2; 1,2; 3,2; 4,2; 5,2*

[c] *Patterns: 2,0; 0,1; 1,1; 2,1; 3,1;4,1; 2,3; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

[d] *Patterns: 0,1; 1,1; 3,1; 4,1; 0,2; 1,2; 3,2; 4,2; 5,2; 3,3; 4,3; 5,3; 3,4; 4,4*

[e] *Patterns: 0,1; 1,1; 3,1; 3,3; 4,3; 5,3; 3,4; 4,4;*

Table 12. Statistical Comparison between tumor margins, smokers, and non- smokers, for the average proportion of cells in the 5 abnormality groups

	Total # cells counted	Normal Pattern (2,2)	Total Altered Pattern	Group 1: Alterations limited to CEP 3[a]	Group 2: Alterations limited to 3p14[b]	Group 3: Any alteration to CEP 3[c]	Group 4: Any alteration to 3p14[d]	Group 5: Alterations to CEP 3 and 3p14[e]
Cancer Patients	2445	90.4700	9.5300	3.6000	3.9700	5.5500	5.8900	0.6300
Smokers	12687	95.3200	4.6800	1.2900	2.4700	2.2300	3.0900	0.6900
P[f]	<0.0001	<0.0001	<0.0001	<0.0001	0.0481	0.0003	0.0148	0.3595
Controls	10324	95.1700	4.8300	1.5200	2.1200	2.7000	2.8500	0.7700
P[g]	<0.0001	<0.0001	<0.0001	0.0014	0.0149	0.0049	0.0062	0.1699
P[h]	0.0656	0.6500	0.6500	0.3095	0.0869	0.0657	0.3770	0.5771

[a] *The following FISH patterns were included: 2,0; 2,1; 2,3; 2,4; 2,5*

[b] *Patterns: 0,2; 1,2; 3,2; 4,2; 5,2*

[c] *Patterns: 2,0; 0,1; 1,1; 2,1; 3,1;4,1; 2,3; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

[d] *Patterns: 0,1; 1,1; 3,1; 4,1; 0,2; 1,2; 3,2; 4,2; 5,2; 3,3; 4,3; 5,3; 3,4; 4,4*

[e] *Patterns: 0,1; 1,1; 3,1; 3,3; 4,3; 5,3; 3,4; 4,4;*

[f] *Mean % cells compared between tumor margins and smokers*

[g] *Mean % cells compared between tumor margins and nonsmokers*

[h] *Mean % cells compared between smokers and nonsmokers*

7.4. Determination of Pathonomic FISH Patterns

Alteration to *FHIT* is associated with the early stages of histological and molecular progression of OSCC, so the identification of abnormal, pathonomic FISH patterns within individuals practicing high-risk behaviors for OSCC development may be an effective method of early disease identification. The underlying rationale for this approach is that some hybridization patterns may be more important than others in the identification of individuals at risk of developing OSCC. For example, patterns that result from technical artifacts like signal overlap, or the presence of dividing cells will have no value as indicators of risk. High-risk patterns should occur exclusively within cancer patients or individuals practicing high-risk behaviors and not be present among non-smoking controls.

Of the twenty FISH patterns found, four were present in cancer samples but not within the nonsmoker samples. For the purpose of identifying hybridization patterns associated with OSCC, these four signals were classified as potentially pathonomic. The first potentially pathonomic pattern was 0,2 (loss of both copies of 3p14.2 with retention of 2 copies of centromere 3). The 0,2 pattern was very rare, occurring in only 2 of 19 cancer patients and representing on average 0.07% of the tumor margin cells evaluated. The second pattern was 2,5; only one sample was found to exhibit this pattern. The third and fourth patterns were 0,1 and 5,3, with each found only once among all tumor margin cells. As these patterns were not observed among the nonsmoking samples, we hypothesized that they represented alterations that may indicate high-risk of

OSCC development even when observed at low frequencies. Table 13 shows the number of cancer patient samples exhibiting these potentially pathonomic patterns. Pathonomic patterns were observed in 15.8% (3 of 19) of cancer samples.

Next, we assessed the frequency of these potentially pathonomic patterns to the smoker's samples to determine if such signal patterns were also observed among smokers who have no clinical signs of the OSCC. We proposed that the identification of such potentially pathonomic signal patterns among smokers might indicate an elevated risk of developing OSCC. Two of the four potentially pathonomic signal patterns were also observed in a very small percentage of the smoker samples (Table 14). The first pathonomic pattern observed among smokers was 0,2 where this signal patterns was counted in three cells from two samples. The second pathonomic pattern counted among the smoker samples was 0,1, which was detected in only one sample. An additional abnormal pattern 5,2 was observed exclusively in the smokers samples and may be pathonomic as it was not found among the nonsmoking samples, and due to the limited availability of tumor margin samples and small samples sizes, this rare abnormal pattern may not have been counted among the tumor margin samples due to the small sample size. The signal pattern 5,2 was therefore included within the potentially pathonomic signal patterns detected in this study.

Table 13. Total pathonomic FISH patterns in cancer patients

Patient #	# of Pathonomic Cells	% Pathonomic cells[a]
1031	1	0.04
1053	0	0.00
1054	0	0.00
1185	0	0.00
1262	1	0.04
1379	0	0.00
1384	0	0.00
1388	0	0.00
1390	0	0.00
1396	0	0.00
1397	0	0.00
1407	3	0.12
1408	0	0.00
1410	0	0.00
1412	0	0.00
1421	0	0.00
1424	0	0.00
1432	0	0.00
1440	0	0.00
TOTAL	5	0.20

[a] *The following patterns were included: 0,1; 0,2; 5,3; 2,5,*

Table 14. Potentially pathonomic FISH patterns identified among smokers

	% patients with any occurrence of the pattern		
Pattern	**Cancer patients[a]**	**Smokers[a]**	**Controls[a]**
2,2	100.00	100.00	100.00
2,0	10.50	0.00	3.50
0,1	5.80	3.30	0.00
1,1	57.00	63.33	69.00
2,1	84.21	100.00	96.60
3,1	5.80	16.70	13.80
4,1	0.00	3.30	3.50
0,2	10.50	6.70	0.00
1,2	16.00	90.00	89.70
3,2	14.00	96.67	93.10
4,2	31.60	50.00	58.60
5,2	0.00	6.67	0.00
2,3	63.20	56.67	62.10
3,3	52.60	53.33	48.30
4,3	5.80	16.70	6.90
5,3	5.80	0.00	0.00
2,4	10.50	10.00	13.80
3,4	0.00	0.00	3.50
4,4	10.50	10.00	10.30
2,5	5.80	0.00	0.00

[a] *values represent percent of patients with indicated pattern among study group*

7.5. Cut-Off Values for FISH Patterns

A second approach that could be used to detect high-risk individuals is to examine the level of alteration in non-smokers and determine cut-off values or thresholds for 3p14.2 and CEP 3 change that can be used as baseline indicators. Such values could be used to determine the extent to which tobacco smoke may have increased genetic alterations at these two loci.

Two different parameters were proposed for cut-off values. The first was the mean for each of the five abnormality groups for non-smoking cases plus two units of standard deviation. Cut-off values were also derived using the normalized data accounting for the differences in sample size and non-normal distribution of percent data, from this data analysis, the same results were obtained regarding individuals exceeding upper cut-off values (see appendix 4). The resulting cut-off values for Groups 1 to 5 using the percent data were: 2.82%, 3.65%, 4.88%, 4.63%, and 1.95% respectively. Table 15 presents the number of tumor margin and smoker samples with significantly elevated levels of DNA damage as determined by the upper limit cut-off values. Figure 8 presents this information graphically.

Table 15. Number of cancer and smoking cases with group frequencies exceeding cut-offs derived from mean plus 2 S.D. values of non-smoking controls

	Cancer Cases Above Cut-off					Smokers Above Cut-off				
	Group 1: Alterations limited to CEP 3[a]	Group 2: Alteration limited to 3p14[b]	Group 3: Any alteration to CEP 3[c]	Group 4: Any alteration to 3p14[d]	Group 5: Alteration to CEP 3 and 3p14[e]	Group 1: Alterations limited to CEP 3[a]	Group 2: Alteration limited to 3p14[b]	Group 3: Any alteration to CEP 3[c]	Group 4: Any alteration to 3p14[d]	Group 5: Alteration to CEP 3 and 3p14[e]
Number of samples	12**	11**	11**	12**	0	0	3	0	2	1
Percent of samples	63.20%	57.90%	57.90%	63.20%	0.00%	0.00%	10.00%	0.00%	6.70%	3.30%

[a] *The Following FISH patterns were included: 2,4;2.3; 2,1; 2;0*

[b] *The following FISH patterns were included: 1,2; 0,2; 3,2; 4,2; 5;2*

[c] *The following FISH patterns were included: 2;3; 2,4; 1,1; 3,4; 0,1; 2,1; 4,3; 3,1; 3,3; 4,4; 2,0; 4,1*

[d] *The following FISH patterns were included: 1,2; 0,2; 1,1; 3,4; 0,1; 3,2; 4,2; 4,3; 3,1; 3,3; 4,4; 4,1; 5,2*

[e] *The following FISH patterns were included: 1,1; 3,4; 0,1; 4,3; 3,1; 3,3; 4,4; 0,1; 5,3*

* *Statistically significant (P<0.01) comparison between tumor margins and smokers*

** *Statistically significant (P<0.001) comparison between tumor margins and smokers*

For Groups 1 through 4, a large percentage of the tumor margin samples exceeded the cut-off values. This included 63% for Group 1, with 57.9%, 57.9% and 63.2% of Groups 2 through 4, respectively. In contrast, only a few of the smoker cases had samples that exceeded the cut-off values. No smokers were found to contain significantly elevated levels of Group 1 or Group 3 signal patterns. However, three smokers (10% of the smoker samples) had significantly increased levels of Group 2 alterations, 2 cases (6.7%) had increased Group 4 alterations and 1 case (3.3%) had increased Group 5 alterations. Overall, 5 (16.7%) of the smokers contained elevated genetic changes that exceeded the cut-off values.

As discussed previously, damage to 3p14.2 and centromere 3 represent different types of genetic events. Of interest, as clearly shown in Figure 8, the smoker samples exceeding the cut-off values appear only in Groups 2, 4 and 5. These are groups with alteration to 3p14.2. None of the smoker samples exceed the cut-offs for Groups 1 and 3, which involve CEP3 alteration. These data suggest a preferential change at the 3p14.2 locus among smokers that does not involve aneuploidy when analyzing the general type of DNA damage being accumulated in smokers from tobacco carcinogens before cancer develops. CEP3 change (Group 1) is only significant for cancer cases, with over 60% of these samples exceeding cut-off. These results regarding general elevated levels of DNA damage among smokers in conjunction with the identification of rare pathonomic signal patterns may have a future use in identifying individuals with an increased risk of OSCC development.

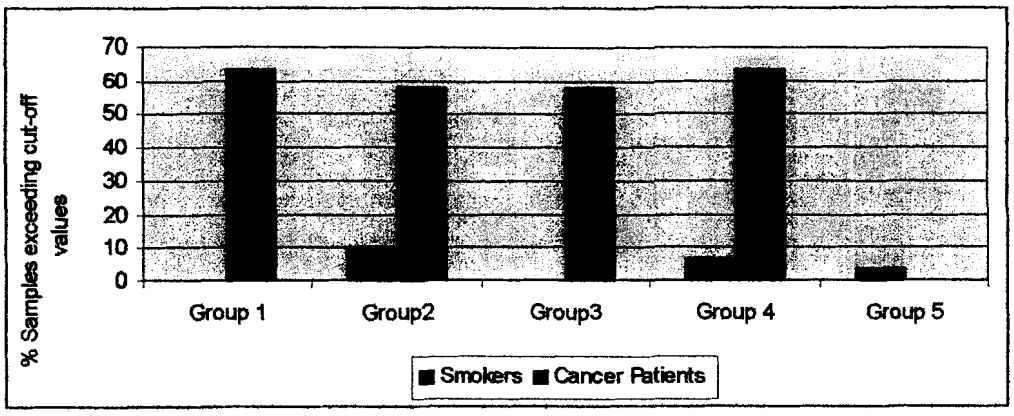


Figure 8. Percentage of samples from cancer cases and smokers that exceed cut-off values among the five abnormality groups.
For each of the 5 abnormality groups there is a single horizontal bar representing the percent of samples from the respective study group that exceeded cut-off values derived from non-smoking data using the mean percent data + 2SD.

Table 16. Demographics of smoking cases with alteration frequencies that exceed cut-off values

Age	Gender[a]	Habit	Pack-years
23	F	Current	3.3
31	F	Current	11.3
31	F	Current	50.6
53	M	Former	50.0
53	F	Current	35.8

[a]*Gender, F = female, M = Male*

Table 16 shows the characteristics of the 5 smokers that had alterations which exceeded cut-off values. Of interest, 4 of the 5 individuals were current rather than former smokers, supporting the possibility of there being more damage among individuals still receiving an exposure. Also, 4 of the 5 cases were female. Whether this indicates a greater susceptibility for females in the accumulation of genetic damage compared with males is not known. There was a large range of ages and pack-years among these 5 individuals. The ages ranged from 23 years to 53 years, while the pack-years ranged from 3.3 to 50.

As a final analysis, we used a second parameter to determine cut-off values for the non-smoking controls and redid the evaluation. We looked at different possibilities, including the mean, the median, one standard deviation, and 2 times the mean. The latter was the most conservative and it was chosen for the analysis. As shown in Table17, the data obtained was similar to the results produced for cancer cases using the mean plus 2 units of standard deviation as the cut-off determinant. The exception for cancer cases was for Group 5, where three samples were found to exceed the cut-off limit compared to none with the previous parameter. Among the smoker samples, the number of positive samples dropped to three samples, as the cut-off values are higher for 4 of the 5 groups, however, even so, samples exceeding the cut-off value still belonged only to those groups (2 and 5) which had 3p14.2 alterations in them.

Table 17. Potential cut-off values from non-smoking data using different parameters

	Group 1: Alterations limited to CEP 3[a]	Group 2: Alteration limited to 3p14[b]	Group 3: Any alteration to CEP 3[c]	Group 4: Any alteration to 3p14[d]	Group 5: Alteration to CEP 3 and 3p14[e]
Mean	1.52	2.19	2.7	2.85	0.77
Median	1.48	2.19	2.6	2.79	0.6
SD	0.65	0.73	1.09	0.89	0.59
Mean + 2SD	2.82	3.65	4.88	4.63	1.95
2XMean	3.4	4.38	5.4	5.7	1.54

[a] *The following FISH patterns were included: 2,0; 2,1; 2,3; 2,4; 2,5*

[b] *Patterns: 0,2; 1,2; 3,2; 4,2; 5,2*

[c] *Patterns: 2,0; 0,1; 1,1; 2,1; 3,1;4,1; 2,3; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

[d] *Patterns: 0,1; 1,1; 3,1; 4,1; 0,2; 1,2; 3,2; 4,2; 5,2; 3,3; 4,3; 5,3; 3,4; 4,4*

[e] *Patterns: 0,1; 1,1; 3,1; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

Table 18. Samples exceeding cut-off values of twice the mean number for abnormal FISH patterns in non-smokers

	Tumor Margins % cells with indicated pattern					Smokers % cells with indicated pattern				
	Group 1: Alterations limited to CEP 3[a]	Group 2: Alteration limited to 3p14[b]	Group 3: Any alteration to CEP 3[c]	Group 4: Any alteration to 3p14[d]	Group 5: Alteration to CEP 3 and 3p14[e]	Group 1: Alterations limited to CEP 3[a]	Group 2: Alteration limited to 3p14[b]	Group 3: Any alteration to CEP 3[c]	Group 4: Any alteration to 3p14[d]	Group 5: Alteration to CEP 3 and 3p14[e]
# of samples > cutoff	11**	8**	10**	12**	3	0**	1**	0**	0**	2
% of samples > cutoff	57.89%	42.11%	52.63%	63.16%	15.79%	0.00%	3.33%	0.00%	0.00%	6.67%

[a] *The following FISH patterns were included: 2,0; 2,1; 2,3; 2,4; 2,5*

[b] *Patterns: 0,2; 1,2; 3,2; 4,2; 5,2*

[c] *Patterns: 2,0; 0,1; 1,1; 2,1; 3,1;4,1; 2,3; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

[d] *Patterns: 0,1; 1,1; 3,1; 4,1; 0,2; 1,2; 3,2; 4,2; 5,2; 3,3; 4,3; 5,3; 3,4; 4,4*

[e] *Patterns: 0,1; 1,1; 3,1; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

** Statistically significant (P<0.01) comparison between tumor margins and smokers*

*** Statistically significant (P<0.001) comparison between tumor margins and smokers*

8. DISCUSSION

The identification of biomarkers for oral cancer risk could have major implications in both disease prevention and treatment. This thesis describes a pilot study aimed at exploring the potential value of using 3p14.2 copy number change, as assessed by FISH in exfoliated cells, to determine risk in seemingly healthy individuals. This approach has not been previously investigated. Of interest, the loss of 3p14, along with another locus at 9p21, has been previously shown by microsatellite analysis to be a strong predictor of the risk of OSCC for individuals that already had OPLs with low-grade dysplasia (Rosin *et al*, 2000). In that study, these 2 loci were used as an initial triage step to identify lesions with a 33-fold increased risk of developing into oral cancer. Unfortunately, microsatellite analysis is time-consuming and invasive, with an accurate assessment requiring biopsies of the lesions and the microdissection of altered cells from these biopsies. The data in this thesis suggests that FISH may provide a non-invasive approach to collecting such information, which would be more suitable for individuals with no clinical signs of the disease. It may be a powerful method of identifying extremely rare, premalignant cells within such individuals.

This study had three objectives: 1) To set up a protocol to produce high

quality nick translation probes targeted towards 3p14.2 and centromere 3; 2) To determine if FISH with probes targeted towards these genomic regions could be used to detect cells with abnormal signal patterns within in exfoliated cell samples obtained from smokers; and 3) To develop cut-off values that could be used to discriminate between naturally occurring background levels of genetic alterations and elevated levels of damage due to prolonged exposure to tobacco carcinogens. All three of these study objectives were accomplished. First, the nick translation protocol was optimized to produce bright, easily observable probes targeted towards the regions of interest. Second, twenty different FISH patterns were observed among all the samples analyzed in this study, with four of these signal patterns present in tumor samples but not in the nonsmoker control samples. Two of these four potentially pathonomic signal patterns also occurred within smoker samples. These patterns may be an indication of premalignant changes within a small number of cells among these individuals. Finally, using the data collected from the nonsmoker control samples, cut-off values were established to identify individuals with elevated levels of genomic damage at 3p14.2 and centromere 3. As expected, the tumor margin samples had greatly elevated levels of genomic damage at these loci, and interestingly, 16.5% of smokers were also found to exhibit elevated levels of damage within 3p14.2.

The following sections will further describe the significant findings of this study. The study limitations and future directions will also be discussed.

8.1. The Role Of the Fragile Histidine Triad Gene in Oral Cancer

The function of *FHIT* in cancer development is as yet not known, however many studies have demonstrated that it this gene is likely to play a role as a tumor suppressor gene. Loss of 3p14.2 is linked to the early progression of many carcinomas, including breast and HNPCC (Turner *et al*, 2002), as well as esophagus, (Kuroki *et al*, 2003), cervix, kidney and lung (Virgilio *et al*, 1996). Loss of *FHIT* is a common event among HNSCC cases in which 45-50% of the cases present loss of heterozygosity at 3p14.2 in both precancerous (dysplastic) and cancerous stages (Virgilio *et al*, 1996). In addition, alterations to *FHIT* resulting in the expression of aberrant *FHIT* transcripts have been demonstrated to occur in 55% of HNSCC cell lines in which one or both alleles have been shown to be at least partially deleted (Virgilio *et al*, 1996). Overall, there is a large body of evidence indicating that loss of *FHIT* function occurs early in oral cancer development; is histologically related to dysplastic changes; and is linked to an increased risk of progression to malignancy (Virgilio *et al*, 1996, Rosin *et al*, 2000, Rosin *et al*, 2002).

8.2. Biological Significance of Alterations to 3p14.2 and Centromere 3 in Oral Cancer

When analyzing the genetic alterations sustained at 3p14.2 and centromere 3, an important distinction between the two genomic regions must be made regarding the type of genetic damage that is represented by each type of change. Alterations to genomic loci such as 3p14.2 located along a

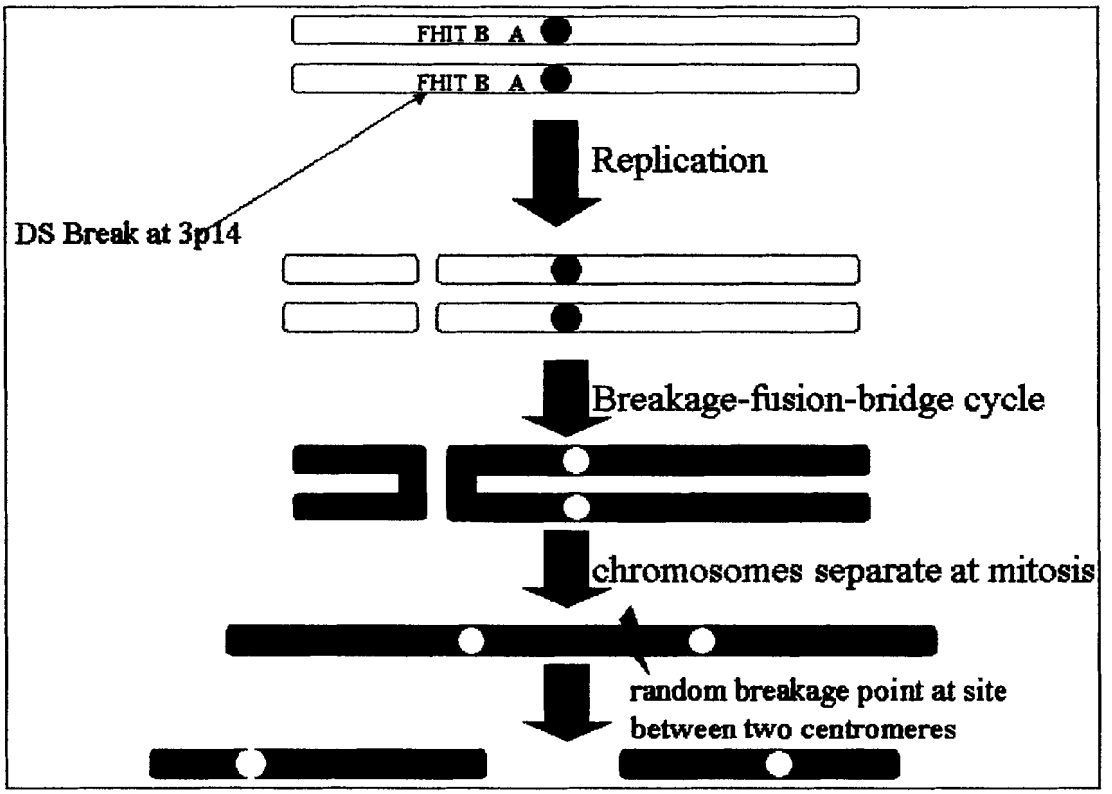
chromosomal arm represent localized damage to that genomic region that can arise in the form of deletions, translocations, breaks, or rearrangements. The *FHIT* gene is located within a fragile site *FRA3B* that is highly unstable, recombinogenic and shows increased levels of DNA gaps, breaks, rearrangements, and sister chromatid exchanges (Corbin *et al*, 2002). Genetic alterations to the *FHIT* region are therefore a result of damage specific to the region spanning 3p14.2.

Altered FISH patterns that involve the centromere of chromosome 3 represent a mechanistically different type of genomic damage. Abnormal centromere 3 FISH signals represent a copy number change of the entire chromosome 3, signaling an aneuploid state rather than specific damage to a genetic locus located along the chromosomal arm. The five different abnormality groups that were analyzed in this study each represent different types of genetic alterations involving centromere 3 and/or 3p14.2. The groups represented cells with abnormal signal patterns limited to centromere 3 only (Group 1) or to 3p14.2 only (Group 2), or with alterations to centromere 3 regardless of the alterations to 3p14.2 (Group 3), with alterations to 3p14.2 regardless of copy changes of chromosome 3 (Group 4) or, finally, with cells that had sustained alterations to both 3p14.2 and centromere 3 (Group 5). To determine if significant increases in genetic damage among the five abnormality groups were present among the cancer patients and smokers, upper limit cut-off values were applied to each of the five abnormality groups using the mean plus 2SD of the nonsmoker data. As shown in Figure 8, cancer samples showed elevated levels of genetic alterations

for abnormality groups one through four, indicating that they had sustained both types of changes investigated in this study, aneuploidy, and alterations to the genomic loci 3p14.2. In contrast to the results obtained for the cancer samples, smokers had elevated levels of genetic damage that was limited to groups 2 and 4, those that involved 3p14.2, with no elevation in groups 1 and 3, those associated with aneuploidy. These results may have important implications, adding to our understanding of the sequence of genetic events that leads to the progression to OSCC.

It has been hypothesized that the increased instability of common fragile sites such as *FRA3B* may play a mechanistic role in the recurring chromosomal rearrangements and genetic changes observed in tumor cells (Corbin *et al*, 2002). Fragile sites have been implicated in the initiation and perpetuation of breakage-fusion-bridge cycles that can lead to chromosome rearrangements and increased genomic instability, generating intratumor heterogeneity, which is a common feature of neoplastic tissues (Gisselsson *et al*, 2000). The initial event in a breakage-fusion-bridge cycle is a chromosomal break that generates an acentric fragment (Figure 9). Due to its lack of a centromere, the acentric fragment is lost at mitosis resulting in a daughter cell with two copies of the incomplete chromosome. Sister chromatid fusion of the two incomplete chromosomal arms results in a dicentric chromosome. During a subsequent cell division, the dicentric chromosome attempts to segregate on the mitotic spindle and the centromeres are pulled to opposite ends of the cell, creating tension on the genomic regions between the two centromeres. This tension breaks the

dicentric chromosome at a random site between the two centromeres and results in two genetically different daughter cells. Continuation of such breakage-fusion-bridge cycles throughout further cell divisions allows genetic changes to continue in the descendants, thus generating intratumor heterogeneity (Lewin, 2000). Early loss or damage to *FHIT* located within *FRA3B* therefore may be an important genetic event for the later development of tobacco-related malignancies. Targeting of the 3p14.2 by tobacco carcinogens and damage to this fragile region can initiate breakage-fusion-bridge events, leading to the acquisition of an instability phenotype which in an inherent feature of malignancy (Ban *et al*, 1995, Gisselsson *et al*, 2000, Fenech, 2002).



Adapted from Alberts et al, 1994

Figure 9. Breakage-fusion-bridge cycle.
Double stranded breaks that remain un-repaired before entry into mitosis are lost due to the absence of a centromere. The two incomplete chromosomal arms then fuse producing a chromosome with two centromeres. During chromosome segregation along the mitotic spindle, each centromere is pulled to opposite ends of the cell, creating tension on the genomic regions between the two centromeres, ultimately resulting in a random beak point between the two centromeres. Overall the result is genetic heterogeneity among daughter cells.

In summary, the presence of genetic damage limited to 3p14.2 and / or the presence of rare, potentially pathonomic signal patterns among smokers may be an indication of early disease development, where the genetic damage is limited to specific genetic loci, with large-scale gains or losses of entire chromosomes that have major impacts on gene dosage yet to occur. The tumor samples exhibited a much greater extent of genetic damage involving alterations to both the 3p14.2 region and aneuploidy. These progressive genetic changes are most likely due to the acquisition of the genomic instability phenotype, a crucial event in the evolution of cancer (Fenech, 2002).

8.3. Importance of Aneuploidy in Cancer Development

Alteration to chromosome number is a frequent event in cancers and pre-cancers with reports of this association in the literature for decades. More recently, chromosomal instability has been linked to mutations in genes controlling chromosome segregation and data from cell culture models and human tissue support a causal role for such change in cell transformation (Duesberg *et al* 2000, Vogelstein *et al* 2004).

Aneuploidy is felt to contribute to risk of malignant transformation by altering gene dosage and hence gene expression of critical regulatory genes (Fenech, 2002, Duesberg *et al*, 2000). Of interest, the acquisition of aneuploidy has been strongly correlated with immortality of cell lines (Duesberg *et al*, 2000). This association is supported by the fact that to date no permanent cell line with a strictly euploid chromosome constitution has been established (Harris, 1991).

The extent to which aneuploidy is present within cells has also been shown to correlate histologically with cancer progression, present at low frequencies in early preneoplastic lesions but increasingly apparent in later lesions and cancer (Duesberg *et al*, 2000).

Over the last decade several articles have been published describing an association of "polysomy", with histological progression. Polysomy is defined as an increase in centromere number. One laboratory, that of Hittelman and co-workers has led this field (Yamal *et al*, 2004, Vassiliki *et al*, 2002). This group has employed FISH on tissue sections of archival blocks to demonstrate that the frequency of polysomy increases with the appearance of dysplasia and progression to cancer. Hittelman's group has also reported that leukoplakia with polysomy in non-cancer patients is more likely to progress to cancer than that lacking polysomy (Lee *et al*, 2000). These data suggest that the identification of cells with increased levels of polysomy may be indicative of increased cancer risk.

Clinical interest in this relationship has been further fueled by a recent series of studies by Subdo and co-workers that support a strong prognostic role for DNA ploidy assessments (Sudbo *et al* 2001, Sudbo *et al* 2004). Sudbo employed computer imaging (rather than FISH) to place cells isolated from tissue sections into 3 categories depending upon DNA content: diploid (normal), tetraploid or aneuploid. With this approach they were able to identify patients with oral leukoplakia who had a very high risk for subsequent development of carcinomas, even if they were histologically defined as being without (or with

minimal) malignant potential. In an initial study of 150 patients with oral leukoplakia and verified dysplasia, the risk of developing a carcinoma was shown to be 27.6 for patients with DNA aneuploid lesions. Three of 105 diploid cases (3%), as opposed to 21 of 25 aneuploid cases (84%), developed a carcinoma during follow-up, yielding a negative predictive value of 97% for diploid lesions and a positive predictive value of 84% for aneuploid lesions (Sudbo *et al*, 2004). These lesions tended to be very aggressive, very rapidly developing into cancer, and even when treated, frequently recurring. The mean time from initial assessment of DNA content to cancer development in aneuploid cases was 35 months (range, 4 to 57). Eighty-five percent of aneuploid cases recurred, compared with 0% of diploid cases and 25% of tetraploid cases (Sudbo *et al*, 2004).

The data presented in this thesis suggest a less invasive approach that could be done without biopsy as a screen for aneuploidy and elevated cancer risk. A larger samples set is required to determine whether or not aneuploidy also occurs in a sub-set of smokers without clinical lesions, perhaps signaling a higher risk of cancer. Such evidence would benefit not only smokers without clinical lesions but also patients that already have leukoplakia. This is an important possibility as reports in the literature suggest that although the majority of oral cancers develop at the site of leukoplakia, others occur distant to the leukoplakia, either due to a field effect (independent mutations by smoke) or possibly due to lateral spread of premalignant clones from the OPL (Lippman *et al*, 2001).

8.4. Pathonomic patterns as independent risk predictors

As discussed previously, many studies have demonstrated that loss of 3p is associated with oral cancer development (Lippman *et al*, 2001, Jin *et al*, 2002, Cheng *et al*, 1998, Gollin, 2001, Scully *et al*, 2000, Nagpal *et al*, 2003, Rosin *et al*, 2002, Virgilio *et al*, 1996, Kuroki *et al*, 1996). Moreover, 3p14 is one of the important sites of alteration on this chromosome arm that appears to be targeted by tobacco carcinogens. Therefore identifying rare, pathonomic FISH patterns involving loss of 3p14.2 may have potential as a biomarker of early OSCC development. Two pathonomic FISH patterns involving loss of 3p14.2 were identified among the tumor margin samples (Table 13). These two pathonomic patterns were 0,1 and 0,2. Both of these FISH patterns were also identified among the smokers but not among the nonsmoking samples, possibly indicating that the signal patterns 0,1, and 0,2 are associated with tobacco carcinogen insult and OSCC development.

Overall, the pathonomic FISH patterns identified among the tumor margins and smokers provides evidence that supports the application of FISH in characterizing chromosomal aberrations on a single-cell basis.

8.4. Potential Susceptibility Factors for Alterations to FHIT

Although the sample size is small, the five smokers who were identified with elevated levels of damage to the *FHIT*-containing region had some similar characteristics and these may be an indication of potential susceptibility factors

for damage to the *FHIT* locus. First, 4 of the 5 individuals were female. This result may be due to chance. However, several studies have concluded that females are at a higher risk than males of developing tobacco-related malignancies at comparable exposure levels (Zavras *et al*, 2001, Cheng *et al*, 2001, Muscat *et al*, 1996). Several explanations have been suggested for this preference including an effect of estrogen on tumor promotion, slower plasma clearance of nicotine, greater activity of p450 enzymes, enhanced formation of DNA adducts, exposure to environmental tobacco smoke, and certain cooking habits (Cheng *et al*, 2001). Even more speculative are suggestions that there may be female-specific nutritional variables or endogenous factors produced/controlled by genes on the X or Y chromosomes that influence a female's risk of oral cancer development (Zavras *et al*, 2001). Although the current study also provides some indication that females are at an increased risk of accumulating damage to the FHIT locus, a much larger study is required to confirm these findings.

The second common feature among these five smokers is that four of five of the individuals are current smokers with only one individual, a former smoker with a heavy history of tobacco use (50 pack years) (Table 16). These data suggest that current smokers may be more likely to exhibit elevated levels of DNA damage at the *FHIT* locus compared with former smokers although a history of heavy tobacco may still result in elevated levels of DNA damage at this locus. A larger study focusing on former versus current smokers is required to confirm this difference.

8.6. Study Limitations and Biases

There were also several sources of potential bias in this study. These include: 1) cancer patients were significantly older than both the smokers and nonsmokers; 2) smaller exfoliated cell samples sizes were obtained for cancer patients compared with smokers and nonsmokers; and 3) the self-reported smoking history may be inaccurate. These potential sources of bias will now be discussed.

The oral cancer patients were significantly older than both the smokers and nonsmokers. The risk of developing oral cancer increases with increasing age. In the present work, cancer patients had an average age of 62 and an age range of 38-90. The only selection criteria used when collecting exfoliated cell samples from individuals at the Abacus Dental Centre was that the individual be at least nineteen years old and have no history of HNSCC. Unfortunately, the age range seen at the Dental Centre was considerably lower than that observed among the cancer patients. Other sites for patient accrual would have to be identified to obtain the required age demographics in future studies. Age may therefore be a confounding factor in this study since the increased levels of genetic damage or the potentially pathonomic patterns identified within the tumor margin samples may be partially due to the effect of age and not HNSCC progression. In the future, gender and age matching among all three study groups would eliminate this potential source of age bias.

The second source of potential bias in this study is the smaller sample size of the tumor margin samples when compared to both the smokers and

nonsmokers. As a result, the number of exfoliated cells counted from the tumor margin samples was significantly less than the number of cells counted for either the smokers or nonsmokers. This difference in sample size most likely reflects the area of oral mucosa sampled among the study groups. The samples from both the smokers and nonsmokers were collected by brushing the entire surface of each ventrolateral tongue and floor of mouth epithelium, whereas the samples collected from the oral cancer patients were taken from a 5 mm wide perimeter surrounding the tumor (within the region 5 to 15 mm from the clinically-identifiable tumor). The tumor margin samples were used in this study as they were readily available; however they were not the ideal sample type to identify genetic changes occurring in oral cancer as they contained small sample sizes, and they were collected from the region outside of the clinically identifiable lesion. To improve this study design in the future, samples should be collected from the cancer itself. This will hopefully increase the sample size and the proportion of cells containing genetic alterations characteristic of the molecular progression of oral cancer.

A further potential area of bias involves the assessment of tobacco history. The report of ever smoker versus non-smoker is probably fairly accurate. However, assessment of tobacco exposure may be less reliable. This is mainly due to a recall bias, or the inability to accurately recall history of tobacco use. Furthermore, there may be inaccuracy in the report by individuals of whether or not they have quit tobacco use. A definitive study around this issue would require the use of a bioindicator for tobacco use, such as the assessment of

cotinine in urine, blood or saliva. Finally, there is increasing interest in the possible involvement of marijuana usage as a source of smoking risk. Collection of information on this habit is difficult and is an ethical concern since the use of the drug is illegal and protection of identity of participants is uncertain.

Marijuana is the most commonly used illegal drug in the United States (Zhang *et al*, 1999, Rosenblatt *et al*, 2004). It is estimated that 31% of the United States population 12 years of age or older have used this drug (Zhang *et al*, 1999). Marijuana smoke is known to contain carcinogens, many of which are similar to those found within tobacco smoke such as phenols and PAHs like benzo[a]pyrene. The latter carcinogen is present at 50% higher concentrations in marijuana tar compared with unfiltered tobacco smoke (Zhang *et al*, 1999). Smoking marijuana cigarettes also deposits four times as much tar in the respiratory tract as that deposited from unfiltered tobacco, again increasing carcinogen exposure (Zhang *et al*, 1999). The smoke produced from marijuana cigarettes has tested positive in the Ames test. It has also been shown to cause molecular and cellular damage similar to that of tobacco smoke in bronchial tissues of humans (Rosenblatt *et al*, 2004). It is therefore feasible that chronic use of marijuana could have a major impact on this study.

A study by Zhang *et al* (1999) investigated the risk of OSCC associated with chronic use of marijuana and found an odds ratio of 2.6 (95% CI, 1.1 – 6.6) compared to never users. However, a second study by Rosenblatt *et al* (2004) did not find any significant increase in risk with marijuana use. However the authors of that study did acknowledge that there was a very limited percentage of

participants that reported chronic marijuana use and that users were mostly under age 50. Due to the increased use of marijuana among teenagers and young adults beginning in the 1960's, and with the latency period of 20-30 years, this first cohort that will be the earliest group to experience and clinically manifest the elevated risk of OSCC is only now approaching the time when this risk can be more accurately assessed.

Marijuana use among the smokers and nonsmokers could have impacted the results of this current study if the levels of marijuana use were not equal among the study groups. Zhang *et al* (1999) found that tobacco cigarette smoking was generally independent of marijuana use in both cases and controls, whereas Rosenblatt *et al* (2004) found that marijuana users more often also smoked tobacco. Consistent results among studies regarding population trends of smoking and marijuana use have yet to be well described. If marijuana use was independent of smoking tobacco in this study, equal levels of carcinogen exposure from marijuana would be expected among the smokers and nonsmokers. If so, each of the sample groups would have equal levels of genetic damage as a result of marijuana use and therefore the levels of genomic damage counted and the cut-off values applied to identify smokers with elevated genomic damage in this study would reflect tobacco carcinogen-induced DNA damage. On the other hand, if marijuana smokers were also more likely to smoke tobacco cigarettes, as observed in the study by Rosenblatt *et al* (2004), then overall the smokers would have an increased exposure to marijuana and the levels of genomic damage identified among the smokers could not be directly

attributed to tobacco exposure alone. However, since similar carcinogens are found within marijuana and tobacco cigarettes, the identification of genomic damage as a result of the combined carcinogen exposure would still provide an indication of the damage sustained at the *FHIT* locus, which could still be applied in the molecular analysis of risk for OSCC development.

8.7. Future Research

The results from this current study have shown that FISH is an effective tool in detecting genetic aberrations within individuals. However the importance of these findings in assessing an seemingly healthy individual's risk of OSCC could not be adequately demonstrated in this current study design due to the few participants and difficulties in sample collection (number of cells, age of participants) and tobacco exposure estimates. A large-scale prospective study is required that involves the participation of several hundred participants including both those with and without premalignant lesions of the oral cavity. Exfoliated cells would be collected from the oral cavity and analyzed for alterations to both *FHIT* and centromere 3. Additional prospective molecular markers for OSCC development such as 9p.21 could also be included in this study to provide additional information on the predictive power of early molecular changes in OSCC development. At fixed intervals in time, the oral cavity of all participants would be examined for the occurrence or progression of oral PMLs, and exfoliated cell samples would be collected and analyzed for molecular alterations to the specified genomic loci. Correlations between genomic changes and

OSCC development could be identified that would facilitate the understanding of the important genetic changes that are associated with oral cancer development. The results from this study design could determine if *FHIT* or other molecular markers do in fact have a predictive power for OSCC development.

A second approach to identify molecular markers that have a predictive power in OSCC development is to look at differences in FISH patterns between 4 different groups of people with different stages of OSCC development including: 1) smokers with no clinical signs of premalignant changes, 2) patients with PMLs displaying mild to moderate dysplasia, 3) patients with severe dysplasia, and 4) patients with confirmed oral squamous cell carcinoma. This study design would sample various stages of OSCC simultaneously and therefore provide information regarding the significant FISH patterns associated with each step of oral cancer progression.

The identification of significant FISH patterns could form the basis of an intervention study, that would determine if oral cancer can be prevented or significantly delayed with the application of various intervention strategies. The latter could include smoking cessation strategies and/or drug therapy targeted towards specific molecular markers. One of the major difficulties in performing such studies involves the large amount of heterogeneity among subjects with respect to risk. Biomarkers that establish an elevated risk for clinically normal individuals would greatly facilitate such efforts. Such a study would involve very large numbers of subjects (several thousand) in a placebo-controlled study design with individuals identified as high-risk for OSCC (i.e., smokers and heavy

drinkers with specific molecular alterations), but show no clinical signs of the disease. These studies are extremely complicated and very expensive. A lot of pilot work, aimed at the issues described above, is necessary before such a study could be conducted

8.8. Summary

In summary, our results support the hypothesis that significantly elevated levels of DNA damage to the *FHIT* region can be detected in exfoliated cells collected from the ventrolateral tongue and floor of mouth of individuals with prolonged tobacco exposure. In addition, potentially pathonomic signal patterns identified within oral cancer tumor margin samples can also be identified within exfoliated cells obtained from smokers, possibly indicating rare, premalignant changes that may be associated with early cancer development. These data provide evidence supporting the molecular classification of oral cancer risk, suggesting that genetic alterations associated with oral cancer can be identified within individuals with no clinical signs of disease. The use of such potential early molecular markers of risk may contribute significantly to the early detection of the disease, allowing time for intervention to prevent cancer development. Finally, and most significantly, this study also determined the ability of this non-invasive, inexpensive, easy, sampling procedure as a screen for individuals at high risk of oral cancer development.

9. APPENDICES

9.1. Oral Health Study Questionnaire

1. In addition to being Canadian or a landed immigrant, what is your ethnic or cultural heritage?
(Check one box only):
   - ☐ White
   - ☐ East or South-east Asian (e.g., China, Japan, Indonesia, Philippines, Vietnam)
   - ☐ South Asian (e.g., India Pakistan, Sri Lanka)
   - ☐ First Nations
   - ☐ Black
   - ☐ Other (Please Specify)_____

2. a) What is the highest grade (or year) of high school or elementary school that you have completed?
Grade ___ Never attended school ____

   b) How many years of post-secondary school have you completed (college, university)?
Years ___ None __

3. a) Have you ever used chewing tobacco?
Yes ☐ No ☐

   b) Have you ever used betel nut?
Yes ☐ No ☐

4. Have you ever regularly smoked cigarettes, cigars or pipes more than once per week for one year or longer? Yes ☐ No ☐
If Yes, please specify:

   a) At what age did you begin smoking:
   Cigarettes? ____
   Cigars? ____
   Pipes? ____

   b) Do you currently smoke:

Cigarettes?	Yes	☐	No	☐
Cigars?	Yes	☐	No	☐
Pipes?	Yes	☐	No	☐

   c) If you have quit smoking, at what age did you permanently stop:
   Cigarettes? ____
   Cigars? ____
   Pipes? ____

d) Looking back over your entire life, on average, how many did you usually smoke *per day*?

	Before Age 20 years	In your 20's	In your 30's	In your 40's	In your 50's	60's & older
Cigarettes	_____	_____	_____	_____	_____	_____
Cigars	_____	_____	_____	_____	_____	_____
Pipes	_____	_____	_____	_____	_____	_____

5. Looking back over the last year, please think about your exposure to the smoke of others, either at home, at work, and in public places (such as restaurants, recreational facilities).

Are you regularly exposed to smoke of others:

At home?	Yes	☐	No	☐
At work?	Yes	☐	No	☐
In public places?	Yes	☐	No	☐

If Yes, to any of the above, please specify:
How often are you regularly exposed to smoke of others:

	Never	Less than once a month	More than once a month but less than once a week	At least once a week	Daily
At home?	☐	☐	☐	☐	☐
At work?	☐	☐	☐	☐	☐
In Public Places?	☐	☐	☐	☐	☐

6. Looking back over your entire life, please check the age periods in which you were daily exposed to the smoke of others.

Before Age 20 years	In your 20's	In your 30's	In your 40's	In your 50's	60's & older
☐	☐	☐	☐	☐	☐

7. Have you ever regularly consumed alcoholic beverages more than once per month for one year or longer? Yes ☐ No ☐

If Yes, please specify:

a) At what age did you begin drinking:

Beer?	____
Wine?	____
Spirits (liquor)?	____

b) Do you currently drink:

Beer?	Yes	☐	No	☐
Wine?	Yes	☐	No	☐
Spirits (liquor)?	Yes	☐	No	☐

c) If you have quit drinking, at what age did you permanently stop:

Beer?	____
Wine?	____
Spirits (liquor)?	____

d) On average, how much did you usually drink *per week*:

Beer	____	bottles
Wine	____	glasses
Spirits (liquor)	____	(shots – 1 oz.)

8. Have any of your immediate family members (parents, brothers/sisters, daughters/sons, grandparents, aunts/uncles related by birth not marriage) had cancer in the head and neck region (excluding skin cancer)?
Yes ☐ No ☐

If Yes, please specify all who had head and neck cancer:

☐ Parents
☐ Brothers/sisters
☐ Daughters/sons
☐ Grandparents
☐ Aunts/uncles related by birth not marriage

9.2. Calculation Of Pack-years

Pack-years are calculated by multiplying smoking duration by daily tobacco consumption (number of cigarettes per day). One pack-year is equal to one pack of cigarettes per day, for one year (Prignot, 1987). To convert other tobacco products to cigarette equivalents, the following conversion factors are applied: 1 pipe = 1.5 cigars = 3 cigarettes (Jourenkova-Mironova *et al*, 1999).

For Example

Age Start Smoking: 20
Smoking Status: Current
Age at sample collection: 30

Tobacco Consumption Over Decades
Age 20-29: half a pack per day (10 cig cigarettes/day)
Age 30-39: one pack per day (20cigarettes/day)

Calculation:
Age 20-29: 10 years • 10 cigarettes/day • 365 day/yr = 36500 cigarettes
Age 30-39: 10 years * 20 cigarettes/day • 365 day/yr = 73000 cigarettes

Overall cigarette consumption = 36500 + 73000 = 109500 cigarettes
1 pack year = 20 cig/day * 365 day/yr = 7300 cigarettes
Therefore overall pack-years = 109500 / 7300 = 15 pack-years

9.3.1. Normalized data of frequencies of aneuploidy and 3p14.2 (FHIT locus) alterations in cancer patients

OHS #	Group1: Alterations limited to CEP 3[a]	Group 2: Alterations limited to 3p14[b]	Group 3: Any alteration to CEP 3[c]	Group 4: Any alteration to 3p14[d]	Group 5: Alteration to CEP 3 and 3p14[e]
1031	0.17	0.22	0.28	0.31	0.10
1053	0.00	0.00	0.00	0.00	0.00
1054	0.20	0.20	0.24	0.24	0.00
1185	0.18	0.18	0.26	0.26	0.07
1262	0.20	0.25	0.24	0.28	0.00
1379	0.19	0.24	0.28	0.31	0.14
1384	0.20	0.20	0.20	0.20	0.10
1388	0.17	0.24	0.21	0.27	0.12
1390	0.13	0.24	0.20	0.28	0.07
1396	0.16	0.25	0.23	0.30	0.00
1397	0.00	0.31	0.00	0.31	0.00
1407	0.20	0.25	0.28	0.31	0.12
1408	0.15	0.15	0.15	0.15	0.10
1410	0.28	0.13	0.31	0.19	0.09
1412	0.20	0.20	0.28	0.28	0.14
1421	0.35	0.00	0.35	0.00	0.00
1424	0.19	0.16	0.21	0.19	0.00
1432	0.20	0.14	0.21	0.16	0.07
1440	0.15	0.15	0.18	0.18	0.00

[a] *The following FISH patterns were included: 2,0; 2,1; 2,3; 2,4; 2,5*

[b] *Patterns: 0,2; 1,2; 3,2; 4,2; 5,2*

[c] *Patterns: 2,0; 0,1; 1,1; 2,1; 3,1;4,1; 2,3; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

[d] *Patterns: 0,1; 1,1; 3,1; 4,1; 0,2; 1,2; 3,2; 4,2; 5,2; 3,3; 4,3; 5,3; 3,4; 4,4*

[e] *Patterns: 0,1; 1,1; 3,1; 3,3; 4,3; 5,3; 3,4; 4,4;*

9.3.2. Normalized data of frequencies of aneuploidy and 3p14.2 (FHIT locus) alterations in smokers

OHS #	Group1: Alterations limited to CEP 3[a]	Group 2: Alterations limited to 3p14[b]	Group 3: Any alteration to CEP 3[c]	Group 4: Any alteration to 3p14[d]	Group 5: Alteration to CEP 3 and 3p14[e]
8528	0.08	0.19	0.14	0.22	0.12
8529	0.14	0.11	0.14	0.11	0.00
8532	0.14	0.12	0.14	0.12	0.00
8533	0.11	0.15	0.11	0.15	0.00
8542	0.09	0.16	0.09	0.16	0.00
8549	0.10	0.16	0.11	0.16	0.00
8550	0.13	0.16	0.14	0.16	0.00
8555	0.10	0.18	0.13	0.19	0.04
8556	0.11	0.15	0.13	0.16	0.04
8557	0.12	0.15	0.14	0.16	0.06
8562	0.09	0.11	0.13	0.13	0.06
8565	0.11	0.13	0.13	0.14	0.06
8566	0.11	0.16	0.14	0.17	0.08
8567	0.09	0.19	0.12	0.21	0.08
8572	0.11	0.16	0.14	0.18	0.08
8573	0.15	0.14	0.18	0.17	0.08
8575	0.09	0.19	0.15	0.21	0.09
8577	0.11	0.18	0.17	0.20	0.09
8578	0.10	0.16	0.14	0.19	0.09
8579	0.13	0.18	0.16	0.20	0.09
8580	0.11	0.13	0.16	0.17	0.10
8582	0.11	0.12	0.16	0.16	0.10
8584	0.12	0.16	0.17	0.18	0.11
8586	0.08	0.15	0.15	0.18	0.11
8587	0.13	0.16	0.18	0.20	0.11
8589	0.13	0.16	0.20	0.18	0.11
8591	0.14	0.11	0.19	0.14	0.11
8595	0.12	0.14	0.18	0.19	0.13
8596	0.11	0.16	0.18	0.21	0.14
8598	0.12	0.22	0.14	0.23	0.06

[a] *The following FISH patterns were included: 2,0; 2,1; 2,3; 2,4; 2,5*

[b] *Patterns: 0,2; 1,2; 3,2; 4,2; 5,2*

[c] *Patterns: 2,0; 0,1; 1,1; 2,1; 3,1;4,1; 2,3; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

[d] *Patterns: 0,1; 1,1; 3,1; 4,1; 0,2; 1,2; 3,2; 4,2; 5,2; 3,3; 4,3; 5,3; 3,4; 4,4*

[e] *Patterns: 0,1; 1,1; 3,1; 3,3; 4,3; 5,3; 3,4; 4,4;*

9.3.3. Normalized data of frequencies of aneuploidy and 3p14.2 (FHIT locus) alterations in non-smokers

OHS #	Group1: Alterations limited to CEP 3[a]	Group 2: Alterations limited to 3p14[b]	Group 3: Any alteration to CEP 3[c]	Group 4: Any alteration to 3p14[d]	Group 5: Alteration to CEP 3 and 3p14[e]
8530	0.11	0.11	0.18	0.15	0.11
8531	0.12	0.14	0.16	0.15	0.04
8536	0.14	0.15	0.15	0.17	0.06
8538	0.16	0.18	0.24	0.22	0.13
8540	0.13	0.18	0.21	0.21	0.12
8541	0.12	0.10	0.17	0.13	0.09
8543	0.12	0.14	0.14	0.16	0.06
8544	0.00	0.10	0.14	0.17	0.14
8545	0.14	0.14	0.18	0.17	0.11
8546	0.13	0.12	0.13	0.12	0.00
8548	0.12	0.14	0.12	0.14	0.00
8551	0.14	0.13	0.17	0.14	0.06
8553	0.13	0.16	0.15	0.17	0.06
8558	0.10	0.13	0.19	0.19	0.14
8559	0.13	0.16	0.17	0.17	0.08
8560	0.10	0.15	0.14	0.17	0.08
8561	0.12	0.19	0.12	0.19	0.00
8563	0.10	0.11	0.15	0.16	0.11
8564	0.19	0.16	0.23	0.19	0.09
8568	0.11	0.16	0.13	0.17	0.04
8569	0.08	0.11	0.11	0.12	0.06
8576	0.08	0.15	0.13	0.17	0.08
8585	0.13	0.16	0.16	0.19	0.10
8590	0.13	0.16	0.17	0.18	0.09
8592	0.11	0.19	0.18	0.22	0.12
8593	0.15	0.14	0.17	0.16	0.08
8599	0.14	0.15	0.20	0.19	0.13
8600	0.14	0.11	0.18	0.11	0.00
8601	0.11	0.17	0.13	0.19	0.08

[a] *The following FISH patterns were included: 2,0; 2,1; 2,3; 2,4; 2,5*

[b] *Patterns: 0,2; 1,2; 3,2; 4,2; 5,2*

[c] *Patterns: 2,0; 0,1; 1,1; 2,1; 3,1;4,1; 2,3; 3,3; 4,3; 5,3; 2,4; 3,4; 4,4; 2,5*

[d] *Patterns: 0,1; 1,1; 3,1; 4,1; 0,2; 1,2; 3,2; 4,2; 5,2; 3,3; 4,3; 5,3; 3,4; 4,4*

[e] *Patterns: 0,1; 1,1; 3,1; 3,3; 4,3; 5,3; 3,4; 4,4;*

10. REFERENCES

Ai H, Barrera JE, Pan Z, Meyers AD, Varella-Garcia M. 1999. Identification of individuals at high risk for head and neck carcinogenesis using chromosome aneuploidy detected by in situ hybridization. Mutat Res 439(2):223-32.

Alberts B. 1994. Molecular biology of the cell. New York: Garland Pub. xliii, 1294 p.

Almendingen K, Xie X, Hofstad B, Vatn MH. 2004. Dietary habits and growth and recurrence of colorectal adenomas: results from a three-year endoscopic follow-up study. Nutr Cancer 49(2):131-8.

Andre K, Schraub S, Mercier M, Bontemps P. 1995. Role of alcohol and tobacco in the aetiology of head and neck cancer: a case-control study in the Doubs region of France. Eur J Cancer B Oral Oncol 31B(5):301-9.

Autrup H. 2000. Genetic polymorphisms in human xenobiotica metabolizing enzymes as susceptibility factors in toxic response. Mutat Res 464(1):65-76.

Awata S, Sakagami H, Tozawa K, Sasaki S, Ueda K, Kohri K. 2000. Aberration of chromosomes 8 and 11 in bladder cancer as detected by in situ hybridization. Urol Res 28(3):185-90.

Ban S, Cologne JB, Neriishi K. 1995. Effect of radiation and cigarette smoking on expression of FUdR-inducible common fragile sites in human peripheral lymphocytes. Mutat Res 334(2):197-203.

Barnes RO. 2003. Numerical chromosome alterations in mucosa adjacent to oral cancers

Bartsch H, Nair U, Risch A, Rojas M, Wikman H, Alexandrov K. 2000. Genetic polymorphism of CYP genes, alone or in combination, as a risk modifier of tobacco-related cancers. Cancer Epidemiol Biomarkers Prev 9(1):3-28.

Benowitz NL. 1999. Biomarkers of environmental tobacco smoke exposure. Environ Health Perspect 107 Suppl 2:349-55.

Bolt HM. 2003. Genotoxicity--threshold or not? Introduction of cases of industrial chemicals. Toxicol Lett 140-141:43-51.

Bosron WF, Li TK. 1986. Genetic polymorphism of human liver alcohol and aldehyde dehydrogenases, and their relationship to alcohol metabolism and alcoholism. Hepatology 6(3):502-10.

Bosron WF, Li TK. 1986. Genetic polymorphism of human liver alcohol and aldehyde dehydrogenases, and their relationship to alcohol metabolism and alcoholism. Hepatology 6(3):502-10.

Bostick RM, Potter JD, Kushi LH, Sellers TA, Steinmetz KA, McKenzie DR, Gapstur SM, Folsom AR. 1994. Sugar, meat, and fat intake, and non-dietary risk factors for colon cancer incidence in Iowa women (United States). Cancer Causes Control 5(1):38-52.

Bouchardy C, Hirvonen A, Coutelle C, Ward PJ, Dayer P, Benhamou S. 2000. Role of alcohol dehydrogenase 3 and cytochrome P-4502E1 genotypes in susceptibility to cancers of the upper aerodigestive tract. Int J Cancer 87(5):734-40.

Braakhuis BJ, Tabor MP, Kummer JA, Leemans CR, Brakenhoff RH. 2003. A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications. Cancer Res 63(8):1727-30.

Brennan JA, Mao L, Hruban RH, Boyle JO, Eby YJ, Koch WM, Goodman SN, Sidransky D. 1995. Molecular assessment of histopathological staging in squamous-cell carcinoma of the head and neck. N Engl J Med 332(7):429-35.

Brouwer IA, Katan MB, Zock PL. 2004. Dietary alpha-linolenic acid is associated with reduced risk of fatal coronary heart disease, but increased prostate cancer risk: a meta-analysis. J Nutr 134(4):919-22.

Cabelguenne A, Loriot MA, Stucker I, Blons H, Koum-Besson E, Brasnu D, Beaune P, Laccourreye O, Laurent-Puig P, De Waziers I. 2001. Glutathione-associated enzymes in head and neck squamous cell carcinoma and response to cisplatin-based neoadjuvant chemotherapy. Int J Cancer 93(5):725-30.

Califano J, van der Riet P, Westra W, Nawroz H, Clayman G, Piantadosi S, Corio R, Lee D, Greenberg B, Koch W and others. 1996. Genetic progression model for head and neck cancer: implications for field cancerization. Cancer Res 56(11):2488-92.

Califano J, Westra WH, Meininger G, Corio R, Koch WM, Sidransky D. 2000. Genetic progression and clonal relationship of recurrent premalignant head and neck lesions. Clin Cancer Res 6(2):347-52.

Callender T, el-Naggar AK, Lee MS, Frankenthaler R, Luna MA, Batsakis JG. 1994. PRAD-1 (CCND1)/cyclin D1 oncogene amplification in primary head and neck squamous cell carcinoma. Cancer 74(1):152-8.

Cheng L, Sturgis EM, Eicher SA, Spitz MR, Wei Q. 2002. Expression of nucleotide excision repair genes and the risk for squamous cell carcinoma of the head and neck. Cancer 94(2):393-7.

Cheng Y, Poulos NE, Lung ML, Hampton G, Ou B, Lerman MI, Stanbridge EJ. 1998. Functional evidence for a nasopharyngeal carcinoma tumor suppressor gene that maps at chromosome 3p21.3. Proc Natl Acad Sci U S A 95(6):3042-7.

Cheng YW, Hsieh LL, Lin PP, Chen CP, Chen CY, Lin TS, Su JM, Lee H. 2001. Gender difference in DNA adduct levels among nonsmoking lung cancer patients. Environ Mol Mutagen 37(4):304-10.

Cho E, Spiegelman D, Hunter DJ, Chen WY, Colditz GA, Willett WC. 2003. Premenopausal dietary carbohydrate, glycemic index, glycemic load, and fiber in relation to risk of breast cancer. Cancer Epidemiol Biomarkers Prev 12(11 Pt 1):1153-8.

Cleaver JE, Karplus K, Kashani-Sabet M, Limoli CL. 2001. Nucleotide excision repair "a legacy of creativity". Mutat Res 485(1):23-36.

Close LG, Brown PM, Vuitch MF, Reisch J, Schaefer SD. 1989. Microvascular invasion and survival in cancer of the oral cavity and oropharynx. Arch Otolaryngol Head Neck Surg 115(11):1304-9.

Corbin S, Neilly ME, Espinosa R, 3rd, Davis EM, McKeithan TW, Le Beau MM. 2002. Identification of unstable sequences within the common fragile site at 3p14.2: implications for the mechanism of deletions within fragile histidine triad gene/common fragile site at 3p14.2 in tumors. Cancer Res 62(12):3477-84.

Coughlin SS, Khoury MJ, Steinberg KK. 1999. BRCA1 and BRCA2 gene mutations and risk of breast cancer. Public health perspectives. Am J Prev Med 16(2):91-8.

Coutelle C, Ward PJ, Fleury B, Quattrocchi P, Chambrin H, Iron A, Couzigou P, Cassaigne A. 1997. Laryngeal and oropharyngeal cancer, and alcohol dehydrogenase 3 and glutathione S-transferase M1 polymorphisms. Hum Genet 99(3):319-25.

Dal Maso L, La Vecchia C, Polesel J, Talamini R, Levi F, Conti E, Zambon P, Negri E, Franceschi S. 2002. Alcohol drinking outside meals and cancers of the upper aero-digestive tract. Int J Cancer 102(4):435-7.

De Flora S, D'Agostini F, Balansky R, Camoirano A, Bennicelli C, Bagnasco M, Cartiglia C, Tampa E, Longobardi MG, Lubet RA and others. 2003. Modulation of cigarette smoke-related end-points in mutagenesis and carcinogenesis. Mutat Res 523-524:237-52.

Doll R, Hill AB. 1950. Smoking and carcinoma of the lung; preliminary report. Br Med J 2(4682):739-48.

Druck T, Hadaczek P, Fu TB, Ohta M, Siprashvili Z, Baffa R, Negrini M, Kastury K, Veronese ML, Rosen D and others. 1997. Structure and expression of the human *FHIT* gene in normal and tumor cells. Cancer Res 57(3):504-12.

Duesberg P, Rasnick D. 2000. Aneuploidy, the somatic mutation that makes cancer a species of its own. Cell Motil Cytoskeleton 47(2):81-107.

Ellard GA, de Waard F, Kemmeren JM. 1995. Urinary nicotine metabolite excretion and lung cancer risk in a female cohort. Br J Cancer 72(3):788-91.

Fenech M. 2002. Chromosomal biomarkers of genomic instability relevant to cancer. Drug Discov Today 7(22):1128-37.

Fiegl M, Zojer N, Kaufmann H, Mullauer L, Schuster R, Huber H, Drach J. 1999. Hyperdiploidy and apparent aneusomy in mesothelial cells from non-malignant effusions as detected by in situ hybridization (FISH). Cytometry 38(1):15-23.

Forastiere A, Koch W, Trotti A, Sidransky D. 2001. Head and neck cancer. N Engl J Med 345(26):1890-900.

Forastiere A, Koch W, Trotti A, Sidransky D. 2001. Head and neck cancer. N Engl J Med 345(26):1890-900.

Franceschi D, Gupta R, Spiro RH, Shah JP. 1993. Improved survival in the treatment of squamous carcinoma of the oral tongue. Am J Surg 166(4):360-5.

Franceschi S, Dal Maso L, Augustin L, Negri E, Parpinel M, Boyle P, Jenkins DJ, La Vecchia C. 2001. Dietary glycemic load and colorectal cancer risk. Ann Oncol 12(2):173-8.

Fritsche E, Bruning T, Jonkmanns C, Ko Y, Bolt HM, Abel J. 1999. Detection of cytochrome P450 1B1 Bfr I polymorphism: genotype distribution in healthy German individuals and in patients with colorectal carcinoma. Pharmacogenetics 9(3):405-8.

Gisselsson D, Pettersson L, Hoglund M, Heidenblad M, Gorunova L, Wiegant J, Mertens F, Dal Cin P, Mitelman F, Mandahl N. 2000. Chromosomal breakage-fusion-bridge events cause genetic intratumor heterogeneity. Proc Natl Acad Sci U S A 97(10):5357-62.

Glover TW, Hoge AW, Miller DE, Ascara-Wilke JE, Adam AN, Dagenais SL, Wilke CM, Dierick HA, Beer DG. 1998. The murine FHIT gene is highly similar to its human orthologue and maps to a common fragile site region. Cancer Res 58(15):3409-14.

Gollin SM. 2001. Chromosomal alterations in squamous cell carcinomas of the head and neck: window to the biology of disease. Head Neck 23(3):238-53.

Goode EL, Ulrich CM, Potter JD. 2002. Polymorphisms in DNA repair genes and associations with cancer risk. Cancer Epidemiol Biomarkers Prev 11(12):1513-30.

Grandis JR, Tweardy DJ. 1993. Elevated levels of transforming growth factor alpha and epidermal growth factor receptor messenger RNA are early markers of carcinogenesis in head and neck cancer. Cancer Res 53(15):3579-84.

Greenlee RT, Hill-Harmon MB, Murray T, Thun M. 2001. Cancer statistics, 2001. CA Cancer J Clin 51(1):15-36.

Gu J, Bondy ML, Sigurdson A, Spitz MR, Hsu TC, Wu X. 1999. Three measures of mutagen sensitivity in a cancer-free population. Cancer Genet Cytogenet 110(1):65-9.

Gunawan B, Mirzaie M, Schulten HJ, Heidrich B, Fuzesi L. 2001. Molecular cytogenetic analysis of two primary squamous cell carcinomas of the lung using multicolor in situ hybridization. Virchows Arch 439(1):85-9.

Haley NJ, Sepkovic DW, Hoffmann D, Wynder EL. 1985. Cigarette smoking as a risk for cardiovascular disease. Part VI. Compensation with nicotine availability as a single variable. Clin Pharmacol Ther 38(2):164-70.

Harris CC. 1991. Chemical and physical carcinogenesis: advances and perspectives for the 1990s. Cancer Res 51(18 Suppl):5023s-5044s.

Hecht SS. 1999. DNA adduct formation from tobacco-specific N-nitrosamines. Mutat Res 424(1-2):127-42.

Herrero R, Castellsague X, Pawlita M, Lissowska J, Kee F, Balaram P, Rajkumar T, Sridhar H, Rose B, Pintos J and others. 2003. Human papillomavirus and oral cancer: the International Agency for Research on Cancer multicenter study. J Natl Cancer Inst 95(23):1772-83.

Hoffmann D, Brunnemann KD, Prokopczyk B, Djordjevic MV. 1994. Tobacco-specific N-nitrosamines and Areca-derived N-nitrosamines: chemistry, biochemistry, carcinogenicity, and relevance to humans. J Toxicol Environ Health 41(1):1-52.

Hoffmann D, Hoffmann I. 1995. Tobacco consumption and lung cancer. Cancer Treat Res 72:1-42.

Hoffmann D, Hoffmann I. 1997. The changing cigarette, 1950-1995. J Toxicol Environ Health 50(4):307-64.

Hong WK, Spitz MR, Lippman SM. 2000. Cancer chemoprevention in the 21st century: genetics, risk modeling, and molecular targets. J Clin Oncol 18(21 Suppl):9S-18S.

Hsu TC, Johnston DA, Cherry LM, Ramkissoon D, Schantz SP, Jessup JM, Winn RJ, Shirley L, Furlong C. 1989. Sensitivity to genotoxic effects of bleomycin in humans: possible relationship to environmental carcinogenesis. Int J Cancer 43(3):403-9.

Hu JJ, Mohrenweiser HW, Bell DA, Leadon SA, Miller MS. 2002. Symposium overview: genetic polymorphisms in DNA repair and cancer risk. Toxicol Appl Pharmacol 185(1):64-73.

Inoue T, Nasu Y, Tsushima T, Miyaji Y, Murakami T, Kumon H. 2000. Chromosomal numerical aberrations of exfoliated cells in the urine detected by *in situ* hybridization: clinical implication for the detection of bladder cancer. Urol Res 28(1):57-61.

Ishiwata S, Takahashi S, Homma Y, Tanaka Y, Kameyama S, Hosaka Y, Kitamura T. 2001. Noninvasive detection and prediction of bladder cancer by *in situ* hybridization analysis of exfoliated urothelial cells in voided urine. Urology 57(4):811-5.

Jares P, Nadal A, Fernandez PL, Pinyol M, Hernandez L, Cazorla M, Hernandez S, Bea S, Cardesa A, Campo E. 1999. Disregulation of p16MTS1/CDK4I protein and mRNA expression is associated with gene alterations in squamous-cell carcinoma of the larynx. Int J Cancer 81(5):705-11.

Jin C, Jin Y, Wennerberg J, Akervall J, Dictor M, Mertens F. 2002. Karyotypic heterogeneity and clonal evolution in squamous cell carcinomas of the head and neck. Cancer Genet Cytogenet 132(2):85-96.

Jin Y, Jin C, Wennerberg J, Hoglund M, Mertens F. 2001. Cytogenetic and *in situ* hybridization characterization of chromosome 8 rearrangements in head and neck squamous cell carcinomas. Cancer Genet Cytogenet 130(2):111-7.

Jordan VC. 1989. Long-term tamoxifen therapy for breast cancer. Important Adv Oncol:179-92.

Jourenkova-Mironova N, Voho A, Bouchardy C, Wikman H, Dayer P, Benhamou S, Hirvonen A. 1999. Glutathione S-transferase GSTM3 and GSTP1 genotypes and larynx cancer risk. Cancer Epidemiol Biomarkers Prev 8(2):185-8.

Kawabata K, Tanaka T, Honjo S, Kakumoto M, Hara A, Makita H, Tatematsu N, Ushida J, Tsuda H, Mori H. 1999. Chemopreventive effect of dietary flavonoid morin on chemically induced rat tongue carcinogenesis. Int J Cancer 83(3):381-6.

Khan AJ, King BL, Smith BD, Smith GL, DiGiovanna MP, Carter D, Haffty BG. 2002. Characterization of the HER-2/neu oncogene by immunohistochemical and *in situ* hybridization analysis in oral and oropharyngeal squamous cell carcinoma. Clin Cancer Res 8(2):540-8.

Kim HR, Christensen R, Park NH, Sapp P, Kang MK. 2001. Elevated expression of hTERT is associated with dysplastic cell transformation during human oral carcinogenesis *in situ*. Clin Cancer Res 7(10):3079-86.

Kim J, Shin DM, El-Naggar A, Lee JS, Corrales C, Lippman SM, Hong WK, Hittelman WN. 2001. Chromosome polysomy and histological characteristics in oral premalignant lesions. Cancer Epidemiol Biomarkers Prev 10(4):319-25.

King RJB. 2000. Cancer biology. Harlow, England ; New York: Pearson Education Ltd. xv, 308 p.

Kisielewski AE, Xiao GH, Liu SC, Klein-Szanto AJ, Novara M, Sina J, Bleicher K, Yeung RS, Goodrow TL. 1998. Analysis of the FHIT gene and its product in squamous cell carcinomas of the head and neck. Oncogene 17(1):83-91.

Kiyohara C, Nakanishi Y, Inutsuka S, Takayama K, Hara N, Motohiro A, Tanaka K, Kono S, Hirohata T. 1998. The relationship between CYP1A1 aryl hydrocarbon hydroxylase activity and lung cancer in a Japanese population. Pharmacogenetics 8(4):315-23.

Ko Y, Abel J, Harth V, Brode P, Antony C, Donat S, Fischer HP, Ortiz-Pallardo ME, Thier R, Sachinidis A and others. 2001. Association of CYP1B1 codon 432 mutant allele in head and neck squamous cell cancer is reflected by somatic mutations of p53 in tumor tissue. Cancer Res 61(11):4398-404.

Kowalski LP, Magrin J, Waksman G, Santo GF, Lopes ME, de Paula RP, Pereira RN, Torloni H. 1993. Supraomohyoid neck dissection in the treatment of head and neck tumors. Survival results in 212 cases. Arch Otolaryngol Head Neck Surg 119(9):958-63.

Kramer IR, Lucas RB, Pindborg JJ, Sobin LH. 1978. Definition of leukoplakia and related lesions: an aid to studies on oral precancer. Oral Surg Oral Med Oral Pathol 46(4):518-39.

Kuroki T. 1996. [Study of cancers as genetic diseases]. Nippon Naika Gakkai Zasshi 85(3):318-21.

Kuroki T, Trapasso F, Yendamuri S, Matsuyama A, Alder H, Mori M, Croce CM. 2003. Allele loss and promoter hypermethylation of VHL, RAR-beta, RASSF1A, and FHIT tumor suppressor genes on chromosome 3p in esophageal squamous cell carcinoma. Cancer Res 63(13):3724-8.

Kuroki T, Trapasso F, Yendamuri S, Matsuyama A, Alder H, Williams NN, Kaiser LR, Croce CM. 2003. Allelic loss on chromosome 3p21.3 and promoter hypermethylation of semaphorin 3B in non-small cell lung cancer. Cancer Res 63(12):3352-5.

Laffon B, Perez-Cadahia B, Pasaro E, Mendez J. 2003. Individual sensitivity to DNA damage induced by styrene in vitro: influence of cytochrome p450, epoxide hydrolase and glutathione S-transferase genotypes. Toxicology 186(1-2):131-41.

Lane DP. 1992. Cancer. p53, guardian of the genome. Nature 358(6381):15-6.

Lee JJ, Hong WK, Hittelman WN, Mao L, Lotan R, Shin DM, Benner SE, Xu XC, Lee JS, Papadimitrakopoulou VM and others. 2000. Predicting cancer development in oral leukoplakia: ten years of translational research. Clin Cancer Res 6(5):1702-10.

Lippman SM, Brown PH. 1999. Tamoxifen prevention of breast cancer: an instance of the fingerpost. J Natl Cancer Inst 91(21):1809-19.

Lippman SM, Hong WK. 2001. Molecular markers of the risk of oral cancer. N Engl J Med 344(17):1323-6.

Lippman SM, Hong WK. 2001. Molecular markers of the risk of oral cancer. N Engl J Med 344(17):1323-6.

Lippman SM, Hong WK. 2002. Cancer prevention by delay. Commentary re: J. A. O'Shaughnessy et al., Treatment and Prevention of Intraepithelial Neoplasia: An Important Target for Accelerated New Agent Development. Clin. Cancer Res., 8: 314-346, 2002. Clin Cancer Res 8(2):305-13.

Liu SC, Sauter ER, Clapper ML, Feldman RS, Levin L, Chen SY, Yen TJ, Ross E, Engstrom PF, Klein-Szanto AJ. 1998. Markers of cell proliferation in normal epithelia and dysplastic leukoplakias of the oral cavity. Cancer Epidemiol Biomarkers Prev 7(7):597-603.

Lumerman H, Freedman P, Kerpel S. 1995. Oral epithelial dysplasia and the development of invasive squamous cell carcinoma. Oral Surg Oral Med Oral Pathol Oral Radiol Endod 79(3):321-9.

Mao L. 2000. Can molecular assessment improve classification of head and neck premalignancy? Clin Cancer Res 6(2):321-2.

Mao L, Lee JS, Fan YH, Ro JY, Batsakis JG, Lippman S, Hittelman W, Hong WK. 1996. Frequent microsatellite alterations at chromosomes 9p21 and 3p14 in oral premalignant lesions and their value in cancer risk assessment. Nat Med 2(6):682-5.

Matullo G, Guarrera S, Carturan S, Peluso M, Malaveille C, Davico L, Piazza A, Vineis P. 2001. DNA repair gene polymorphisms, bulky DNA adducts in white blood cells and bladder cancer in a case-control study. Int J Cancer 92(4):562-7.

McGregor F, Muntoni A, Fleming J, Brown J, Felix DH, MacDonald DG, Parkinson EK, Harrison PR. 2002. Molecular changes associated with oral dysplasia progression and acquisition of immortality: potential for its reversal by 5-azacytidine. Cancer Res 62(16):4757-66.

Merritt WD, Weissler MC, Turk BF, Gilmer TM. 1990. Oncogene amplification in squamous cell carcinoma of the head and neck. Arch Otolaryngol Head Neck Surg 116(12):1394-8.

Michels KB, Ekbom A. 2004. Caloric restriction and incidence of breast cancer. Jama 291(10):1226-30.

Mori M, Mimori K, Shiraishi T, Alder H, Inoue H, Tanaka Y, Sugimachi K, Huebner K, Croce CM. 2000. Altered expression of FHIT in carcinoma and precarcinomatous lesions of the esophagus. Cancer Res 60(5):1177-82.

Morita S, Yano M, Tsujinaka T, Akiyama Y, Taniguchi M, Kaneko K, Miki H, Fujii T, Yoshino K, Kusuoka H and others. 1999. Genetic polymorphisms of drug-metabolizing enzymes and susceptibility to head-and-neck squamous-cell carcinoma. Int J Cancer 80(5):685-8.

Mork J, Lie AK, Glattre E, Hallmans G, Jellum E, Koskela P, Moller B, Pukkala E, Schiller JT, Youngman L and others. 2001. Human papillomavirus infection as a risk factor for squamous-cell carcinoma of the head and neck. N Engl J Med 344(15):1125-31.

Mukherjee B, Morgenbesser SD, DePinho RA. 1992. Myc family oncoproteins function through a common pathway to transform normal cells in culture: cross-interference by Max and trans-acting dominant mutants. Genes Dev 6(8):1480-92.

Munger K, Werness BA, Dyson N, Phelps WC, Harlow E, Howley PM. 1989. Complex formation of human papillomavirus E7 proteins with the retinoblastoma tumor suppressor gene product. Embo J 8(13):4099-105.

Muscat JE, Richie JP, Jr., Thompson S, Wynder EL. 1996. Gender differences in smoking and risk for oral cancer. Cancer Res 56(22):5192-7.

Nagpal JK, Das BR. 2003. Oral cancer: reviewing the present understanding of its molecular mechanism and exploring the future directions for its effective management. Oral Oncol 39(3):213-21.

Nagpal JK, Mishra R, Das BR. 2002. Activation of Stat-3 as one of the early events in tobacco chewing-mediated oral carcinogenesis. Cancer 94(9):2393-400.

Nagpal JK, Patnaik S, Das BR. 2002. Prevalence of high-risk human papilloma virus types and its association with P53 codon 72 polymorphism in tobacco addicted oral squamous cell carcinoma (OSCC) patients of Eastern India. Int J Cancer 97(5):649-53.

Negrini M, Monaco C, Vorechovsky I, Ohta M, Druck T, Baffa R, Huebner K, Croce CM. 1996. The FHIT gene at 3p14.2 is abnormal in breast carcinomas. Cancer Res 56(14):3173-9.

Neville BW, Day TA. 2002. Oral cancer and precancerous lesions. CA Cancer J Clin 52(4):195-215.

O'Shaughnessy JA, Kelloff GJ, Gordon GB, Dannenberg AJ, Hong WK, Fabian CJ, Sigman CC, Bertagnolli MM, Stratton SP, Lam S and others. 2002. Treatment and prevention of intraepithelial neoplasia: an important target for accelerated new agent development. Clin Cancer Res 8(2):314-46.

Pandis N, Bardi G, Mitelman F, Heim S. 1997. Deletion of the short arm of chromosome 3 in breast tumors. Genes Chromosomes Cancer 18(4):241-5.

Parkin DM. 2001. Global cancer statistics in the year 2000. Lancet Oncol 2(9):533-43.

Partridge M, Li SR, Pateromichelakis S, Francis R, Phillips E, Huang XH, Tesfa-Selase F, Langdon JD. 2000. Detection of minimal residual cancer to investigate why oral tumors recur despite seemingly adequate treatment. Clin Cancer Res 6(7):2718-25.

Partridge M, Pateromichelakis S, Phillips E, Emilion GG, A'Hern RP, Langdon JD. 2000. A case-control study confirms that microsatellite assay can identify patients at risk of developing oral squamous cell carcinoma within a field of cancerization. Cancer Res 60(14):3893-8.

Patel KA, Edmondson ND, Talbot F, Parsons MA, Rennie IG, Sisley K. 2001. Prediction of prognosis in patients with uveal melanoma using *in situ* hybridisation. Br J Ophthalmol 85(12):1440-4.

Persson I, Johansson I, Ingelman-Sundberg M. 1997. In vitro kinetics of two human CYP1A1 variant enzymes suggested to be associated with interindividual differences in cancer susceptibility. Biochem Biophys Res Commun 231(1):227-30.

Petridou E, Zavras AI, Lefatzis D, Dessypris N, Laskaris G, Dokianakis G, Segas J, Douglas CW, Diehl SR, Trichopoulos D. 2002. The role of diet and specific micronutrients in the etiology of oral carcinoma. Cancer 94(11):2981-8.

Pindborg JJ, Jolst O, Renstrup G, Roed-Petersen B. 1968. Studies in oral leukoplakia: a preliminary report on the period pervalence of malignant transformation in leukoplakia based on a follow-up study of 248 patients. J Am Dent Assoc 76(4):767-71.

Prignot J. 1987. Quantification and chemical markers of tobacco-exposure. Eur J Respir Dis 70(1):1-7.

Quon H, Liu FF, Cummings BJ. 2001. Potential molecular prognostic markers in head and neck squamous cell carcinomas. Head Neck 23(2):147-59.

Ralhan R, Agarwal S, Mathur M, Wasylyk B, Srivastava A. 2000. Association between polymorphism in p21(Waf1/Cip1) cyclin-dependent kinase inhibitor gene and human oral cancer. Clin Cancer Res 6(6):2440-7.

Raybaud-Diogene H, Tetu B, Morency R, Fortin A, Monteil RA. 1996. p53 overexpression in head and neck squamous cell carcinoma: review of the literature. Eur J Cancer B Oral Oncol 32B(3):143-9.

Rinaldi AL, Morse MA, Fields HW, Rothas DA, Pei P, Rodrigo KA, Renner RJ, Mallery SR. 2002. Curcumin activates the aryl hydrocarbon receptor yet significantly inhibits (-)-benzo(a)pyrene-7R-trans-7,8-dihydrodiol bioactivation in oral squamous cell carcinoma cells and oral mucosa. Cancer Res 62(19):5451-6.

Roberts SA, Spreadborough AR, Bulman B, Barber JB, Evans DG, Scott D. 1999. Heritability of cellular radiosensitivity: a marker of low-penetrance predisposition genes in breast cancer? Am J Hum Genet 65(3):784-94.

Rosenblatt KA, Daling JR, Chen C, Sherman KJ, Schwartz SM. 2004. Marijuana use and risk of oral squamous cell carcinoma. Cancer Res 64(11):4049-54.

Rosin MP, Cheng X, Poh C, Lam WL, Huang Y, Lovas J, Berean K, Epstein JB, Priddy R, Le ND and others. 2000. Use of allelic loss to predict malignant risk for low-grade oral epithelial dysplasia. Clin Cancer Res 6(2):357-62.

Rosin MP, Lam WL, Poh C, Le ND, Li RJ, Zeng T, Priddy R, Zhang L. 2002. 3p14 and 9p21 loss is a simple tool for predicting second oral malignancy at previously treated oral cancer sites. Cancer Res 62(22):6447-50.

Rubin Grandis J, Melhem MF, Gooding WE, Day R, Holst VA, Wagener MM, Drenning SD, Tweardy DJ. 1998. Levels of TGF-alpha and EGFR protein in head and neck squamous cell carcinoma and patient survival. J Natl Cancer Inst 90(11):824-32.

Saranath D, Panchal RG, Nair R, Mehta AR, Sanghavi V, Sumegi J, Klein G, Deo MG. 1989. Oncogene amplification in squamous cell carcinoma of the oral cavity. Jpn J Cancer Res 80(5):430-7.

Saranath D, Panchal RG, Nair R, Mehta AR, Sanghavi V, Sumegi J, Klein G, Deo MG. 1989. Oncogene amplification in squamous cell carcinoma of the oral cavity. Jpn J Cancer Res 80(5):430-7.

Saranath D, Panchal RG, Nair R, Mehta AR, Sanghavi VD, Deo MG. 1992. Amplification and overexpression of epidermal growth factor receptor gene in human oropharyngeal cancer. Eur J Cancer B Oral Oncol 28B(2):139-43.

Saranath D, Panchal RG, Nair R, Mehta AR, Sanghavi VD, Deo MG. 1992. Amplification and overexpression of epidermal growth factor receptor gene in human oropharyngeal cancer. Eur J Cancer B Oral Oncol 28B(2):139-43.

Saunders WS, Shuster M, Huang X, Gharaibeh B, Enyenihi AH, Petersen I, Gollin SM. 2000. Chromosomal instability and cytoskeletal defects in oral cancer cells. Proc Natl Acad Sci U S A 97(1):303-8.

Schoumacher F, Glaus A, Mueller H, Eppenberger U, Bolliger B, Senn HJ. 2001. BRCA1/2 mutations in Swiss patients with familial or early-onset breast and ovarian cancer. Swiss Med Wkly 131(15-16):223-6.

Scully C, Bedi R. 2000. Ethnicity and oral cancer. Lancet Oncol 1(1):37-42.

Scully C, Field JK, Tanzawa H. 2000. Genetic aberrations in oral or head and neck squamous cell carcinoma 2: chromosomal aberrations. Oral Oncol 36(4):311-27.

Shen H, Spitz MR, Qiao Y, Guo Z, Wang LE, Bosken CH, Amos CI, Wei Q. 2003. Smoking, DNA repair capacity and risk of nonsmall cell lung cancer. Int J Cancer 107(1):84-8.

Shen H, Sturgis EM, Dahlstrom KR, Zheng Y, Spitz MR, Wei Q. 2002. A variant of the DNA repair gene XRCC3 and risk of squamous cell carcinoma of the head and neck: a case-control analysis. Int J Cancer 99(6):869-72.

Shen H, Sturgis EM, Khan SG, Qiao Y, Shahlavi T, Eicher SA, Xu Y, Wang X, Strom SS, Spitz MR and others. 2001. An intronic poly (AT) polymorphism of the DNA repair gene XPC and risk of squamous cell carcinoma of the head and neck: a case-control study. Cancer Res 61(8):3321-5.

Shen H, Zheng Y, Sturgis EM, Spitz MR, Wei Q. 2002. P53 codon 72 polymorphism and risk of squamous cell carcinoma of the head and neck: a case-control study. Cancer Lett 183(2):123-30.

Silverman S, Jr., Gorsky M, Lozada F. 1984. Oral leukoplakia and malignant transformation. A follow-up study of 257 patients. Cancer 53(3):563-8.

Silverman S, Jr., Gorsky M, Lozada F. 1984. Oral leukoplakia and malignant transformation. A follow-up study of 257 patients. Cancer 53(3):563-8.

Smith CJ, Perfetti TA, Rumple MA, Rodgman A, Doolittle DJ. 2000. "IARC group 2A Carcinogens" reported in cigarette mainstream smoke. Food Chem Toxicol 38(4):371-83.

Smith CJ, Perfetti TA, Rumple MA, Rodgman A, Doolittle DJ. 2001. "IARC Group 2B carcinogens" reported in cigarette mainstream smoke. Food Chem Toxicol 39(2):183-205.

Soler M, Bosetti C, Franceschi S, Negri E, Zambon P, Talamini R, Conti E, La Vecchia C. 2001. Fiber intake and the risk of oral, pharyngeal and esophageal cancer. Int J Cancer 91(3):283-7.

Somers KD, Cartwright SL, Schechter GL. 1990. Amplification of the int-2 gene in human head and neck squamous cell carcinomas. Oncogene 5(6):915-20.

Sozzi G, Veronese ML, Negrini M, Baffa R, Cotticelli MG, Inoue H, Tornielli S, Pilotti S, De Gregorio L, Pastorino U and others. 1996. The FHIT gene 3p14.2 is abnormal in lung cancer. Cell 85(1):17-26.

Sriuranpong V, Park JI, Amornphimoltham P, Patel V, Nelkin BD, Gutkind JS. 2003. Epidermal growth factor receptor-independent constitutive activation of STAT3 in head and neck squamous cell carcinoma is mediated by the autocrine/paracrine stimulation of the interleukin 6/gp130 cytokine system. Cancer Res 63(11):2948-56.

Stoner GD, Morse MA, Kelloff GJ. 1997. Perspectives in cancer chemoprevention. Environ Health Perspect 105 Suppl 4:945-54.

Sudbo J, Bryne M, Mao L, Lotan R, Reith A, Kildal W, Davidson B, Soland TM, Lippman SM. 2003. Molecular based treatment of oral cancer. Oral Oncol 39(8):749-58.

Sudbo J, Kildal W, Risberg B, Koppang HS, Danielsen HE, Reith A. 2001. DNA content as a prognostic marker in patients with oral leukoplakia. N Engl J Med 344(17):1270-8.

Sudbo J, Reith A. 2002. When is an oral leukoplakia premalignant? Oral Oncol 38(8):813-4; author reply 811-2.

Sudbo J, Ried T, Bryne M, Kildal W, Danielsen H, Reith A. 2001. Abnormal DNA content predicts the occurrence of carcinomas in non-dysplastic oral white patches. Oral Oncol 37(7):558-65.

Tabor MP, Brakenhoff RH, van Houten VM, Kummer JA, Snel MH, Snijders PJ, Snow GB, Leemans CR, Braakhuis BJ. 2001. Persistence of genetically altered fields in head and neck cancer patients: biological and clinical implications. Clin Cancer Res 7(6):1523-32.

Tanaka H, Shimada Y, Harada H, Shinoda M, Hatooka S, Imamura M, Ishizaki K. 1998. Methylation of the 5' CpG island of the FHIT gene is closely associated with transcriptional inactivation in esophageal squamous cell carcinomas. Cancer Res 58(15):3429-34.

Tradati N, Grigolat R, Calabrese L, Costa L, Giugliano G, Morelli F, Scully C, Boyle P, Chiesa F. 1997. Oral leukoplakias: to treat or not? Oral Oncol 33(5):317-21.

Turner BC, Ottey M, Zimonjic DB, Potoczek M, Hauck WW, Pequignot E, Keck-Waggoner CL, Sevignani C, Aldaz CM, McCue PA and others. 2002. The fragile histidine triad/common chromosome fragile site 3B locus and repair-deficient cancers. Cancer Res 62(14):4054-60.

van Houten VM, Tabor MP, van den Brekel MW, Kummer JA, Denkers F, Dijkstra J, Leemans R, van der Waal I, Snow GB, Brakenhoff RH. 2002. Mutated p53 as a molecular marker for the diagnosis of head and neck cancer. J Pathol 198(4):476-86.

Virgilio L, Shuster M, Gollin SM, Veronese ML, Ohta M, Huebner K, Croce CM. 1996. FHIT gene alterations in head and neck squamous cell carcinomas. Proc Natl Acad Sci U S A 93(18):9770-5.

Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, Nakamura Y, White R, Smits AM, Bos JL. 1988. Genetic alterations during colorectal-tumor development. N Engl J Med 319(9):525-32.

Vokes EE, Weichselbaum RR, Lippman SM, Hong WK. 1993. Head and neck cancer. N Engl J Med 328(3):184-94.

Wang Y, Liang D, Spitz MR, Zhang K, Dong Q, Amos CI, Wu X. 2003. XRCC3 genetic polymorphism, smoking, and lung carcinoma risk in minority populations. Cancer 98(8):1701-6.

Wang Y, Spitz MR, Zhu Y, Dong Q, Shete S, Wu X. 2003. From genotype to phenotype: correlating XRCC1 polymorphisms with mutagen sensitivity. DNA Repair (Amst) 2(8):901-8.

Wei Q, Gu J, Cheng L, Bondy ML, Jiang H, Hong WK, Spitz MR. 1996. Benzo(a)pyrene diol epoxide-induced chromosomal aberrations and risk of lung cancer. Cancer Res 56(17):3975-9.

Weinberg RA. 1991. Tumor suppressor genes. Science 254(5035):1138-46.

Welcsh PL, Owens KN, King MC. 2000. Insights into the functions of BRCA1 and BRCA2. Trends Genet 16(2):69-74.

Werner M, Wilkens L, Aubele M, Nolte M, Zitzelsberger H, Komminoth P. 1997. Interphase cytogenetics in pathology: principles, methods, and applications of *in situ* hybridization (FISH). Histochem Cell Biol 108(4-5):381-90.

White JL, Conner BT, Perfetti TA, Bombick BR, Avalos JT, Fowler KW, Smith CJ, Doolittle DJ. 2001. Effect of pyrolysis temperature on the mutagenicity of tobacco smoke condensate. Food Chem Toxicol 39(5):499-505.

Winsey SL, Haldar NA, Marsh HP, Bunce M, Marshall SE, Harris AL, Wojnarowska F, Welsh KI. 2000. A variant within the DNA repair gene XRCC3 is associated with the development of melanoma skin cancer. Cancer Res 60(20):5612-6.

Wood RD, Mitchell M, Sgouros J, Lindahl T. 2001. Human DNA repair genes. Science 291(5507):1284-9.

Wu X, Lippman SM, Lee JJ, Zhu Y, Wei QV, Thomas M, Hong WK, Spitz MR. 2002. Chromosome instability in lymphocytes: a potential indicator of predisposition to oral premalignant lesions. Cancer Res 62(10):2813-8.

Wu X, Shi H, Jiang H, Kemp B, Hong WK, Delclos GL, Spitz MR. 1997. Associations between cytochrome P4502E1 genotype, mutagen sensitivity, cigarette smoking and susceptibility to lung cancer. Carcinogenesis 18(5):967-73.

Xia W, Lau YK, Zhang HZ, Liu AR, Li L, Kiyokawa N, Clayman GL, Katz RL, Hung MC. 1997. Strong correlation between c-erbB-2 overexpression and overall survival of patients with oral squamous cell carcinoma. Clin Cancer Res 3(1):3-9.

Yu GP, Zhang ZF, Hsu TC, Spitz MR, Schantz SP. 1999. Family history of cancer, mutagen sensitivity, and increased risk of head and neck cancer. Cancer Lett 146(1):93-101.

Yuan A, Yu CJ, Luh KT, Kuo SH, Lee YC, Yang PC. 2002. Aberrant p53 expression correlates with expression of vascular endothelial growth factor mRNA and interleukin-8 mRNA and neoangiogenesis in non-small-cell lung cancer. J Clin Oncol 20(4):900-10.

Zavras AI, Douglass CW, Joshipura K, Wu T, Laskaris G, Petridou E, Dokianakis G, Segas J, Lefantzis D, Nomikos P and others. 2001. Smoking and alcohol in the etiology of oral cancer: gender-specific risk profiles in the south of Greece. Oral Oncol 37(1):28-35.

Zavras AI, Wu T, Laskaris G, Wang YF, Cartsos V, Segas J, Lefantzis D, Joshipura K, Douglass CW, Diehl SR. 2002. Interaction between a single nucleotide polymorphism in the alcohol dehydrogenase 3 gene, alcohol consumption and oral cancer risk. Int J Cancer 97(4):526-30.

Zhang L, Poh CF, Lam WL, Epstein JB, Cheng X, Zhang X, Priddy R, Lovas J, Le ND, Rosin MP. 2001. Impact of localized treatment in reducing risk of progression of low-grade oral dysplasia: molecular evidence of incomplete resection. Oral Oncol 37(6):505-12.

Zhang L, Robertson ML, Kolachana P, Davison AJ, Smith MT. 1993. Benzene metabolite, 1,2,4-benzenetriol, induces micronuclei and oxidative DNA damage in human lymphocytes and HL60 cells. Environ Mol Mutagen 21(4):339-48.

Zhang ZF, Morgenstern H, Spitz MR, Tashkin DP, Yu GP, Hsu TC, Schantz SP. 2000. Environmental tobacco smoking, mutagen sensitivity, and head and neck squamous cell carcinoma. Cancer Epidemiol Biomarkers Prev 9(10):1043-9.

Zhang ZF, Morgenstern H, Spitz MR, Tashkin DP, Yu GP, Marshall JR, Hsu TC, Schantz SP. 1999. Marijuana use and increased risk of squamous cell carcinoma of the head and neck. Cancer Epidemiol Biomarkers Prev 8(12):1071-8.

Zhou J, Ahn J, Wilson SH, Prives C. 2001. A role for p53 in base excision repair. Embo J 20(4):914-23.

Zhu Y, Spitz MR, Hsu TC, Wu X. 2002. Genetic instability of specific chromosomes associated with a family history of cancer. Cancer Genet Cytogenet 136(1):73-7.

Zhu Y, Spitz MR, Strom S, Tomlinson GE, Amos CI, Minna JD, Wu X. 2002. A case-control analysis of lymphocytic chromosome 9 aberrations in lung cancer. Int J Cancer 102(5):536-40.

Znaor A, Brennan P, Gajalakshmi V, Mathew A, Shanta V, Varghese C, Boffetta P. 2003. Independent and combined effects of tobacco smoking, chewing and alcohol drinking on the risk of oral, pharyngeal and esophageal cancers in Indian men. Int J Cancer 105(5):681-6.