A CASE-CONTROL STUDY OF ATM AND SUSCEPTIBILITY TO SQUAMOUS CELL CARCINOMA OF THE HEAD AND NECK

by

Sarah Grams

BS, University of Wisconsin-Madison, 1998

Submitted to Graduate Faculty of

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science

University of Pittsburgh

2007

UNIVERSITY OF PITTSBURGH

Graduate School of Public Health

This thesis was presented

by

Sarah Grams

It was defended on

June 1, 2007

and approved by

Thesis Advisor: Susanne Gollin, PhD
Professor of Human Genetics
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Robert Ferrell, PhD
Professor of Human Genetics
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Eleanor Feingold, PhD
Professor of Human Genetics
Department of Human Genetics
Graduate School of Public Health
University of Pittsburgh

Emanuela Taioli, MD, PhD
Professor of Epidemiology, Hematology and Oncology
Department of Epidemiology
Graduate School of Public Health
University of Pittsburgh

Copyright © by Sarah Grams

2007

Susanne Gollin, PhD

A CASE-CONTROL STUDY OF ATM AND SUSCEPTIBILITY TO SQUAMOUS CELL CARCINOMA OF THE HEAD AND NECK

Sarah Grams, MS

University of Pittsburgh, 2007

The American Cancer Society estimates that >500,000 new cases of squamous cell carcinomas of the head and neck (SCCHN) are diagnosed each year. Although the incidence varies widely around the world, it is especially high in developing countries and is positively associated with higher rates of exogenous risk factors including, smoking, alcohol use, and viral infection. But, only a fraction of the people in these high-risk groups will develop the disease. Treatment times tend to be long and costly with survival rates averaging 50%, one of the lowest for the major cancers. Therefore, further work is needed to aid in our understanding of genetic and environmental risk factors, as well as the underlying biology of SCCHN. To further assess the etiology of SCCHN, the study used the approach of examining one candidate gene, the Ataxia Telangiectasia Mutated gene (*ATM*), located at 11q22.3 and functioning in the DNA damage response/repair pathway. Compared to other cancers, SCCHN tumors exhibit a very high rate of chromosomal instability, often showing amplification of chromosome 11q13 and loss of 11q22-qter. We hypothesized that SCCHN patients (cases) have a higher incidence of germline alterations in *ATM* than controls. Three hundred cases and 360 controls were genotyped for nine ATM tag-SNPs and supplemented with one splice-site SNP. Logistic regression analysis showed that two SNPs (rs611646 and rs373759) were associated with increased risk of developing SCCHN: P = 0.012 and P = 0.025, respectively. In SNP rs611646, the TT genotype was more common in cases than controls (45 versus 32%) while the AA and

AT genotypes were less common in cases than controls (18 versus 21% and 37 versus 46%, respectively). In SNP rs373759, the CC genotype was more common in cases than controls (56 versus 48%) with the CT type being less common in cases than controls (29 versus 40%). Genetic studies such as this could have a public health impact by identifying markers for early detection of SCCHN, for predicting prognosis, for therapy and drug development, and aiding in the development of new, individualized treatment strategies.

TABLE OF CONTENTS

PREFACE .... **X**

1.0 INTRODUCTION **1**

1.1 CANCER ETIOLOGY **3**

1.1.1 Mechanisms of DNA Repair **4**

1.1.2 Hereditary Cancer Syndromes and Related DNA Repair Defects **5**

2.0 REVIEW OF THE RELEVANT LITERATURE **6**

2.1 EPIDEMIOLOGY OF SCCHN **6**

2.1.1 Environmental Risk Factors **6**

2.1.2 Genetic Risk Factors **9**

2.1.2.1 Tumor studies. **9**

2.1.2.2 Molecular studies: **12**

2.2 ATM AND DNA REPAIR **17**

2.2.1 Function of *ATM* **18**

2.2.2 *ATM* and Disease **22**

2.2.2.1 *ATM* and breast cancer. **22**

2.2.2.2 *ATM* and lung cancer **25**

2.2.2.3 *ATM* and squamous cell carcinoma of the head and neck (SCCHN)

25

3.0 METHODS **27**

3.1.1 Subject Recruitment **27**

3.1.2 Inclusion and Exclusion Criteria **28**

3.1.3 Epidemiological variables **31**

3.1.4 Sample Processing **31**

3.1.5 Identification of the SNP Set **32**

3.1.6 Genotyping **36**

3.1.7 Statistical Analysis **37**

4.0 RESULTS **39**

4.1.1 Epidemiological variables and SCCHN risk **39**

4.1.2 Association of ATM SNPs and SCCHN risk **40**

5.0 DISCUSSION **45**

5.1.1 Study design **45**

5.1.2 Genetic associations **47**

5.1.3 Epidemiology **48**

5.1.4 Future directions **49**

BIBLIOGRAPHY **51**

LIST OF TABLES

Table 1. Cancer incidence in 1160 relatives of 132 AT patients, (60)..23

Table 2. Genotyped SNPs...35

Table 3. Epidemiological variables..39

Table 4. Epidemiological variables..40

Table 5. Allele frequencies...41

Table 6. Sex, BMI, alcohol, and pack-years-adjusted odds ratios (ORs) for ATM SNP genotype calls. 43

Table 7. Sex, BMI, alcohol, and pack-years-adjusted odds ratios (ORs) for ATM SNP slope-calculation...43

Table 8. Joint genotype analysis of rs611646 and rs373759..50

LIST OF FIGURES

Figure 1. Proposed mechanism for SCCHN carcinogenesis, (8)..10

Figure 2. Proposed mechanism for breakage-fusion-cycle in 11q amplification, (33)................12

Figure 3. Model of DNA double-strand-break response pathway, (53)......................................18

Figure 4. The *ATM* signaling pathway, (53, 54)...20

Figure 5. Model of DNA double-strand-break response genes, (53)...21

Figure 6. Genotyped SNPs relative to the ATM gene. Adapted from the ATM gene model by the Seattle Mouse Comparative Genomics Center (77)..35

Figure 7. Graph of rs3092991 genotype data, labeled with case/control status. 44

PREFACE

I would like to extend warm thanks to: Susanne Gollin PhD, Marjorie Romkes PhD, Jennifer Grandis, MD PhD, Robert Ferrell PhD, Eleanor Feingold, PhD, Emanuela Taioli MD, PhD, Yan Lin, PhD, Dale Lewis, Brian Henson PhD, and the University of Pittsburgh Genomics and Proteomics Core Lab.

This work was funded by:

The Department of Health and Human Services, Public Health Service, National Institutes of Health, National Cancer Institute: Grant P50CA097190 (HN-SPORE, PI: Jennifer Grandis)

The Department of Health and Human Services, Public Health Service, National Institutes of Health, National Cancer Institute: Grant P60DE13059 (Oral Cancer Center, PI: Eugene N. Myers)

The Department of Health and Human Services, Public Health Service, National Institutes of Health, National Institute of Dental and Craniofacial Research: Grant R01DE14729 (PI: Susanne Gollin)

1.0 INTRODUCTION

Head and neck cancer is a broad term used to describe malignancies in the upper aero-digestive tract, including the pharynx, oral cavity and larynx. Although implied by “head and neck”, conventionally, the term does not include cancers of the brain, thyroid and melanoma (1). Ninety percent of head and neck cancer begins in the squamous cells that line the passages of the upper respiratory tract. Named for the types of affected cells and location of the cancer, this type of tumor is referred to as squamous cell carcinoma of the head and neck (SCCHN) or head and neck squamous cell carcinoma (HNSCC). The American Cancer Society estimates that cancer of the oral cavity and pharynx is the ninth most common cancer in Caucasian men and the sixth most common cancer in African American men, accounting for 3% of cancers in both ethnicities (3). These cancers are about twice as common in men as compared to women, making them the 14th most common cancer in women. The lifetime risk for Caucasian males in the United States is approximately 1 in 72 (3). To put this in perspective, it is estimated that about 34,360 new cases (24,180 in men and 10,180 in women) of cancer of the oral cavity and pharynx will be diagnosed in the United States in 2007. Worldwide, the yearly incidence is expected to be >500,000, with the majority occurring in developing countries, including India, South Africa, Eastern Europe and tropical South America (1). This makes head and neck cancer a significant cause of morbidity and mortality worldwide.

In the US, cancers of the oral cavity and pharynx are estimated to be responsible for approximately 7,550 deaths in 2007. Historically, death rates have been higher in both African American men (1.8 times higher) and women (1.2 times higher) as compared to Caucasian men and women (3). This corresponds to an average death rate of 61% in Caucasians and 40% in African Americans. Although five-year survival rates have improved over time for advanced carcinomas of the larynx and pharynx, there has been little improvement for oral cancers or early laryngeal tumors. Overall, survival is one of the lowest for the major cancers, highlighting the importance of studies that work to improve our understanding of the biology of SCCHN.

SCCHN is largely preventable and positively associated with higher rates of exogenous risk factors including, smoking and alcohol use (1) betel nut chewing and bidi smoking in Southeast Asia; but, as many as 25% of cases are not attributable to these exposures (4). A number of observations suggest that there may be a different mechanism of disease in non-smokers/non-drinkers with SCCHN cases more likely to be females (73-78%) (5, 6). In addition, a number of studies have reported differences in the anatomical site of tumors in smokers and non-smokers, with the majority of smoking-related tumors occurring in the oral cavity (5, 6). SCCHN has also been associated with Human Papillomavirus (HPV) with estimates of virus integration rates in tumors range from 21-43% (7). Although smoking, alcohol and HPV are significant risk factors for SCCHN, only a fraction of people in high-risk groups will develop cancer.

Head and neck cancer is thought to be a complex, heterogeneous disease involving multiple genetic abnormalities. It appears that a variety of mechanisms may play a role in the shift to malignancy, including the inactivation of tumor suppressor genes, activation of proto-oncogenes, promoter methylation, gene deletion, gene amplification and point mutations (8).

These genetic abnormalities could arise as a result of a combination exogenous factors and inherited susceptibility to DNA damage. With this in mind, this study aimed to examine the influence of the Ataxia Telengiectasia Mutated gene (*ATM*) on susceptibility to SCCHN. *ATM*, predicted to influence SCCHN, is a DNA damage response kinase that functions by phosphorylating key substrates involved in DNA repair and cell cycle checkpoints. Specifically, a case-control based SNP association study was performed using constitutional DNA prepared from blood. This work had the ultimate goal of further elucidating the etiology of SCCHN, as well as identifying important genetic risk factors for the disease.

1.1 CANCER ETIOLOGY

Cancer has been described as a ”genetic disease of the somatic cell” (9). The formation of cancer can occur as a series of somatic events involving any number of possible genetic changes, from simple point mutations to more complex chromosomal rearrangements. The common etiology of cancer consists of somatic changes, resulting in the inability of cells to maintain a stable genome. Cells contain a number of safeguards to protect against somatic cell changes; however, a breakdown in these systems has the potential to deregulate cell growth and lead to malignancy.

Throughout their lives, cells may be exposed to many different types of damage, either caused by endogenous or exogenous sources. When DNA damage occurs, it triggers a complex cascade of events and the cell responds by attempting DNA repair or inducing programmed cell death (apoptosis). The attempt to repair can result in success (ie., a viable, healthy cell), apoptosis, or in further genomic instability and/or cancer (9). Therefore, when studying the

etiology of cancer, it is important to consider the exogenous risk factors that can lead to DNA damage, as well as provide a discussion of DNA repair mechanisms and the endogenous factors that may affect this process.

1.1.1 Mechanisms of DNA Repair

Generally, the mechanisms responsible for DNA repair depend on the type of damage/lesion. The Nucleotide Excision Repair pathway (NER) works to remove and replace bulky, DNA-distorting lesions, resulting from exposure to various environmental agents. The damage is detected, DNA is unwound at the site, the nucleotide is excised and, finally, replaced using the opposite strand as a template (10). The Base-Excision Repair pathway (BER) works to repair bases that have been modified by deamination or alkylation. In general, these are lesions that do not distort the DNA backbone. Several DNA glycosylases recognize the lesion and work with other proteins to repair the damaged base (10). Double strand breaks (DSBs) can occur as a result of ionizing radiation (IR), chemical exposure, meiotic (or mitotic) crossing-over, a stalled replication fork, or immunoglobulin V(D)J recombination. There are two main pathways for repairing DSBs; Non-Homologous End Joining (NHEJ), the predominant pathway in mammals, involves direct ligation of the two ends of broken DNA, without use of a template and can result in small deletions. Homologous Recombination (HR) makes use of template DNA as a guide to repair and is essentially error-free (10).

1.1.2 Hereditary Cancer Syndromes and Related DNA Repair Defects

Mutations in a number of genes that function in these DNA repair pathways have been shown to be associated with various hereditary cancer syndromes. For example, mutations in the xeroderma pigmentosum (*XP*) gene, involved in NER, results in high UV light sensitivity and an increased risk for skin cancer, melanoma and other tumors (11). In addition, NER genes are associated with Cockayne syndrome and trichothiodystrophy. Germline mutations in the MutY human homologue (*MUTYH*), a gene that is involved in repairing oxidative damage to DNA (BER pathway), and have been found to be associated with familial colorectal cancer (12). The NHEJ pathway also contains genes implicated in Hereditary Breast and Ovarian Cancer (*BRCA1/2*), Fanconi Anemia (*BRCA2*), Li Fraumeni syndrome (*TP53*), Ataxia Telangiectasia (*ATM*), Bloom Syndrome *(BLM*) and others.

In general, a discussion of the particulars of these gene functions and carcinogenesis becomes extremely complex and reveals that one gene may function in different pathways, interacting with multiple proteins, and performing many cellular roles (13). Adding to this complexity, studies examining the levels of DNA repair pathway activity, have disagreed on whether an over-active or under-active pathway is associated with increased risks (14).

2.0 REVIEW OF THE RELEVANT LITERATURE

2.1 EPIDEMIOLOGY OF SCCHN

Head and neck cancer is a significant cause of morbidity and mortality worldwide with overall survival being one of the lowest for the major cancers. Treatment times tend to be long and costly with survival rates averaging 50%. Therefore, much work is needed to further our understanding of genetic and environmental risk factors, as well as the underlying biology of SCCHN. This would serve several purposes, including: identifying markers for detection of premalignant tumors, predicting clinical outcome (prognosis), identifying targets for therapy and drug development, developing individualized treatments, and developing new strategies to predict and prevent resistance to treatment.

2.1.1 Environmental Risk Factors

Squamous cell carcinoma of the head and neck (SCCHN) is thought to arise due to a number of factors, including damage to DNA as well as deregulation of the cell cycle checkpoints and subsequent genomic instability. Known environmental risk factors, and possible exogenous sources of DNA damage in SCCHN include:

1) Tobacco. Higher risks of SCCHN are associated with smoking cigarettes, pipes, and cigars, and chewing tobacco (15). About 85.4% of people with SCCHN use or have used tobacco, compared to 69.9% among controls (16). Higher risks are associated with quantities smoked/chewed, duration and type of tobacco used. Smokeless tobacco is especially harmful and increases risk by about 50 times that of a non-user (17).

2) Alcohol use. A high level of use is seen in about 75% of patients and infers a risk of oral cancer that is six times higher than that of non-drinkers. The combination of both tobacco and alcohol use is especially harmful, with significantly increased odd ratios (OR=12.7 [95%CI = 5.5-29.1; p-value = 0.008]) (16). Both alcohol and tobacco have been shown to be independent risk factors and, together, act synergistically to increase risk (18). Most likely this is due to the ability of alcohol to increase the permeability of oral mucosa to tobacco carcinogens (1, 3).

3) Human papillomavirus (HPV). Commonly seen in cervical cancers, HPV is also found in about 25% of head and neck cancers (19). HPV is not necessary, nor sufficient to cause SCCHN, but has consistently been found more frequently in SCCHN cells than in normal mucosa (10%) (7). Patients with SCCHN and an associated HPV infection appear to have a better prognosis and are less likely to be smokers and drinkers. This suggests a slightly different disease etiology in these cases (1, 3). Studies have also examined the relationship between smoking, HPV and SCCHN (20). These researchers found that HPV seropositivity and tobacco appeared have an additive effect on the risk of SCCHN with an OR of 8.5 (95%CI = 5.1-14.4). HPV seronegative smokers had an OR of 3.2 (95%CI = 2.0-5.2) and HPV VLP-seropositive non-

smokers had an OR of 1.7 (95%CI = 1.1-2.6). Although studies have also disagreed in their findings concerning the relationship between alcohol, smoking and HPV, it appears that the increased risks seen in smokers/drinkers/HPV seropositive individuals is likely due to the additive effect of HPV and smoking (7).

4) Age. The incidence of SCCHN increases with age, with the majority of cases diagnosed over 40 years of age (1). The association of tobacco and alcohol is well documented in older patients (1, 3); however, younger patients may make up a distinct subset of SCCHN cases. Younger patients are more often women than men, and tumor sites are more often located on the tongue or in the oral cavity (21).

5) Gender. The disease is twice as common in men then in women. In both men and women, tobacco and alcohol use are the most striking risk factors, with men having higher rates of use compared to women (3, 22). A recent study (22) analyzed data on 195 cases and 1113 controls from Italy and Switzerland, found slightly lower odds ratios for alcohol consumption in women compared to men. The authors suggested that these lower ORs in women may represent underreporting of alcohol consumption or that women may be more susceptible to alcohol carcinogenesis. These researchers also confirmed previous studies showing that women with oral cancer had lower Body Mass Indexes (BMI) than men, although this could be a consequence of disease rather than a cause.

6) Plummer-Vinson syndrome. A less important risk factor, Plummer-Vinson is a very rare syndrome with less than 30 cases reported between 1999 and 2005 (23). The syndrome is

characterized by a combination of iron deficiency with abnormalities of the tongue, fingernails, esophagus, and red blood cells and is associated with an increased risk (3-15%) of esophageal or pharyngeal cancer (23).

7) Ultraviolet light. Prolonged exposure to sunlight is more common in people with cancers of the lip (1, 3).

8) Diet. Dietary deficiencies may be a risk factor for SCCHN in as many as 10-15% of cases, whereas diet is rich green vegetables, fresh fruit, B-carotene and wholegrain foods have a protective effect (3, 22, 24). Increased risks of oral cancer have also been associated with greater intake of total fats and saturated fats (24).

2.1.2 Genetic Risk Factors

2.1.2.1 Tumor studies.

Microsatellite and cytogenetic studies have suggested a model for the progression of SCCHN. Specifically, an examination of SCCHN tumors reveals that they are often surrounded by a "field" of abnormal cells containing many genetic alterations. In 1953, Slaughter *et al*. examined 783 cases of SCCHN and found multiple carcinomas in 11.2% of these cases. It was postulated that these fields arose in areas of mucosa that had undergone prolonged exposure to carcinogens, causing cells to undergo numerous independent genetic changes. This first led to the Slaughter *et al*. theory that SCCHN tumors arose from multiple primary cells (25). More recent tumor studies have suggested that although the surrounding mucosa contains complex alterations, these changes are often related to each other and most likely arose from a "common

preneoplastic progenitor" (26). These progenitor cells may migrate and expand at different sites, acquiring further genetic changes, which make them appear to be separate primary tumors. As a result, distinguishing between a second primary tumor and a recurrence can often be difficult and is usually defined by the genetic similarity of the tumors (8).

Compared to other cancers, SCCHN exhibit a very high rate of chromosomal instability (2) with 70% of cases showing loss of the 9p21 region. This event is often associated with a loss of heterozygosity of 9p21. This region contains genes involved in cell cycle regulation, including *CDKN2A* (8). It is estimated that approximately 90% of SCCHN show inactivation of *CDKN2A* through deletion, promoter methylation or point mutations (27). Also thought to occur early in tumor formation is the over expression of epidermal growth factor receptor, *EGFR*, (7p12). This has been reported in approximately 90% of SCCHN and is correlated with a poor prognosis (28). (Figure 1).

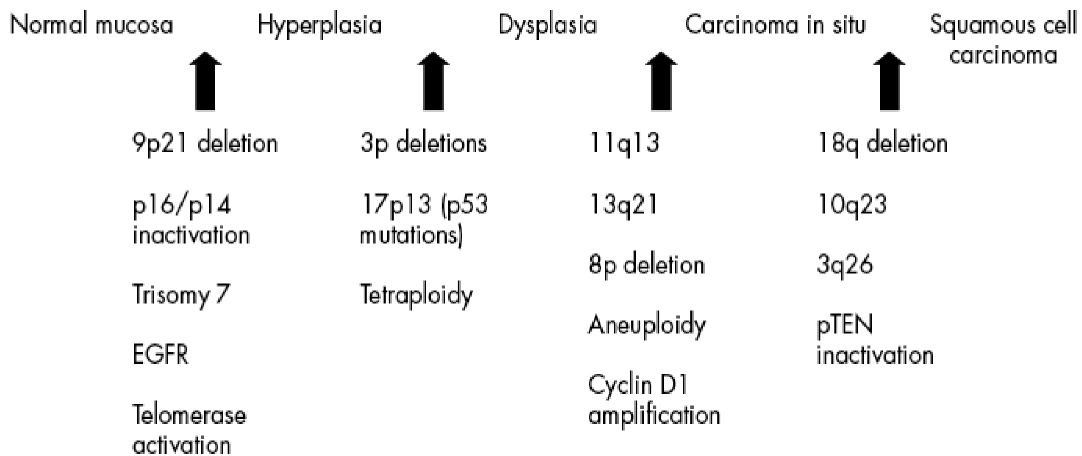

Figure 1. Proposed mechanism for SCCHN carcinogenesis, (8).

Other early events in tumor formation include loss of 3p14, 3p21, 3p22, 3p24, and 3p26. In these regions, three candidate tumor suppressor genes have been proposed to explain these observations. The fragile histidine triad gene (*FHIT*) (3p14) is found to be inactive in some SCCHN cases as well as other tumors. *RSSFIA* on 3p21 is also inactive in a set of SCCHN tumors (8). Germline polymorphisms in the glutathione peroxidase I (*GPX1*) gene, on 3p21, have been associated with susceptibility to SCCHN (29).

Later in tumor formation, loss of 17p13, has been observed, followed by loss of heterozygosity. This region includes *TP53*, which responds to exogenous and endogenous stressors, playing a major role in the regulation of the cell cycle. *TP53* is mutated in 50-80% of SCCHN tumors (8, 27). Interestingly the site of *TP53* tumor mutations may be different in non-smoking/non-drinking patients. This may indicate that certain mutations are characteristic of exogenous damage, while others are characteristic of endogenous damage (30). Although the stage at which *TP53* loss occurs has been debated, a number of studies indicate that incidence increases with the shift from non-invasive to invasive cancer (31). Although *TP53* mutation is common in SCCHN tumors, a number of tumors form in the absence of *TP53* mutations. These tumors are more likely to be HPV-positive, again indicating a distinct etiology for HPV-positive tumor formation (8).

About 30-76% of SCCHN tumors show loss of the distal end of chromosome 11q and subsequent amplification of 11q13 (32). The amplified 11q13 region contains cyclin D1 (*CCND1*) and is associated with poor prognosis as well as metastatic spread to the lymph nodes. *CCND1* is a key protein in the cell cycle, phosphorylating *RB1* and enabling the transition from G1 to S phase (8). Reshmi et al. have proposed that this amplification occurs as a result of cigarette smoke-induced double strand breaks and subsequent breakage-fusion bridge cycles in

chromosome 11 (33). The original breakpoint may occur at the common fragile site, *FRA11F,* located at 11q14. This mechanism also implies loss of the distal end of 11q, containing *ATM*, prior to the 11q13 amplification (Figure 2).

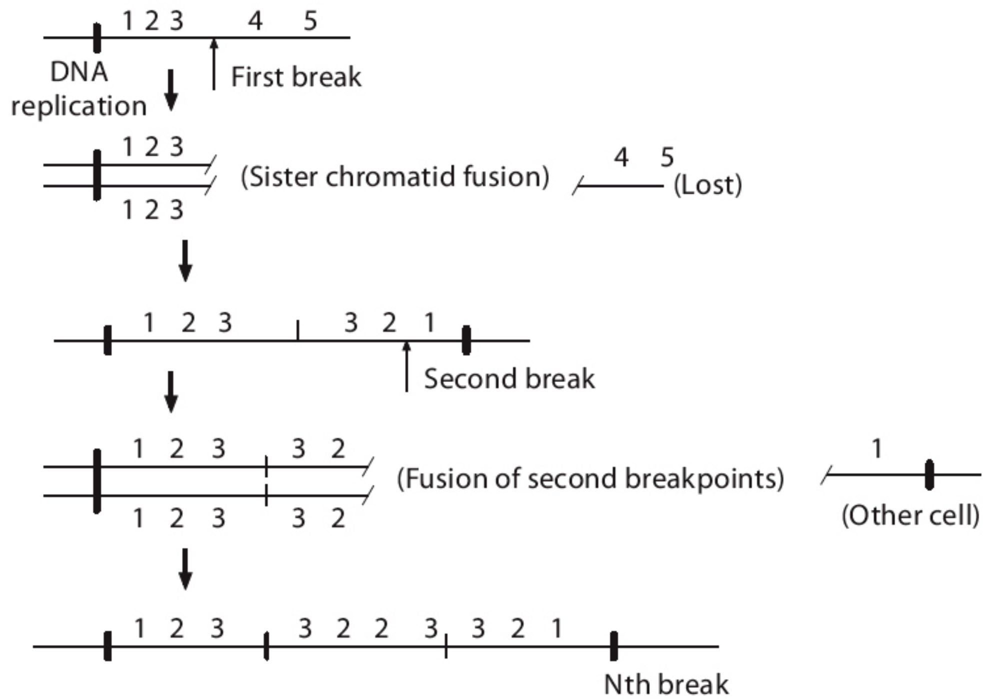

Figure 2. Proposed mechanism for breakage-fusion-cycle in 11q amplification, (33).

2.1.2.2 Molecular studies:

Although no major SCCHN susceptibility loci have been found, many modest associations have been reported.

Oncogenes:

CCND1 is a regulatory protein, which plays an important role in the G1 to S transition in the cell cycle. Mutations in the G870A splice variant lack a normally occurring

destruction box, causing overexpression of the protein (34). Several studies have implicated polymorphisms in *CCND1* with SCCHN susceptibility (35). Zheng et al. suggested that the G870A splice-site AA polymorphism may contribute to early onset SCCHN (36). Rydzanicz et al. found that the same AA allele was positively associated with higher gene expression and lymph node metastases (37). In addition, a recent study of 273 cases and 269 controls found a positive association of germline variations in the same *CCND1* allele, with higher OR's when occuring in combination with certain *ERCC2 (XPD)* variants [OR 7.09 (95% CI 4.03 – 12.5)] (34).

Tumor suppressors/DNA repair genes:

CDKN1A (P21): regulates cell-cycle progression at the G1 phase and is controlled by *TP53.* A significant association was found between risk of SCCHN and two polymorphisms (C70T and C98A). This was a case-control study of 712 patients and 1222 controls with OR's approximately 1.3-1.5 for each SNP (38).

XRCC1 facilitates DNA DSB and BE repair. A large number of studies have examined polymorphisms in *XRCC1* with differing results (39, 40).

ERCC2 (XPD) is involved in the nucleotide excision repair pathway with polymorphisms resulting in the possible reduction of DNA repair. Sturgis et al. reported nonsignificantly increased risks associated with germline polymorphisms in the *ERCC2/XPD* 23591A allele [OR 1.28 (95% CI, 0.93-1.76)] in a study of 313 cases and 313 controls (40). Risk was higher in combination with the *ERCC1* 8092CC genotype: [OR 1.78 (95% CI, 0.99-

3.17)]. In addition, Buch *et al.* showed an increased risk associated with both the *ERCC2/XPD* Asp312Asn variant [OR 1.3 (95% CI 1.0-1.8)] and *ERCC2/XPD* Lys751Gln variant [OR 2.2 (95% CI 1.5-3.2)] (34). Also higher OR's were observed in combination with the *CCND1* G870A risk allele [OR 7.09 (95% CI 4.03 – 12.5)] (34).

Detoxifying genes:

Researchers have examined the role of interindividual genetic differences in modifying the risks of alcohol and tobacco exposure (18, 41) and found that variations in enzymes responsible for metabolism of carcinogens may be associated with SCCHN risk.

GST's (Glutathione S-transferases) make up a family of enzymes involved with detoxifying tobacco smoke. Over 20 studies have focused on variations in the *GSTM1* variations in SCCHN patients with OR's ranging from nonsignificant to OR's of 3.9 (42). A recent study (692 cases, 753 controls) examined the synergy of *GST* genes with tobacco and smoking exposure (18). The researchers found that individuals who were homozygous for deletions in the *GSTM1* gene and who were heavy-drinkers and heavy-smokers, had significantly increased risks to develop SCCHN (OR, 12.6; 95% CI, 4.0-40.2). These individuals have a complete lack of enzyme function and are thought to have a reduced ability to detoxify carcinogens.

SULT1A1: Mutations in sulfotransferase *(SULT) 1A1,* also a detoxifying agent, have been found to be associated with SCCHN (OR=3.60; 95% CI=1.01-12.88) in older people who had high alcohol and low fruit intake (43).

ADH1C (alcohol dehydrogenase 1C) is known to play a role in the elimination of alcohol from the body. Peters et al. investigated *ADH1C* genotypes in 521 cases and 599 controls from the Boston area. They found a significant increase in cancer risk in those heavy alcohol users who were *ADH1C* homozygote variants: OR of 7.1 (95% CI, 2.3-22.0) (44).

Nutrient metabolism genes:

MTR (Methionine Synthase) and *MTRR* (Methionine Synthase Reductase) in the folate metabolic pathway have been associated with moderate changes in susceptibility to SCCHN, although confirmatory studies are needed (45).

MTHFR (Methylenetetrahydrofolate Reductase): Variations in the *MTHFR* gene have been shown to influence folate metabolism. Three polymorphisms have been shown to be associated with increased risk to SCCHN. Neumann et al. genotyped the C677T, A1298C and G1793A alleles in cases and controls and found OR's = 1.85 (1.3-2.5) for any 2 risk alleles and OR = 1.93 (1.4-2.7) for any three risk alleles (46).

TYMS (Thymidylate Synthase) also known to be involved in folate metabolism was examined by Zhang et al. The group found a number of polymorphisms which were protective or associated with decreased stage of cancer (47).

GPX1 (Glutathione peroxidase I) codes for a selenium-dependent enzyme that helps to protect cells against oxidative damage. Located in the 3p21 region (frequently lost in

SCCHN tumors), polymorphisms in *GPX1* have been associated with SCCHN (OR=1.6) (29). It should be noted that this study was limited by small sample sizes as well as lack of smoking/alcohol status.

VDR (Vitamin D Receptor): Liu et al. showed that polymorphisms in the VDR gene may alter risk to SCCHN. The examined two polymorphisms (FokI and TaqI restriction site polymorphisms) in 719 cases and 821 controls and found that homozygous variant genotypes for both were associated with decreased risk for SCCHN (ORs between 0.64 and 0.72) (48).

Other:

IL4 (interleukin 4) is a cytokine produced by activated T-cells. Vairaktaris et al. examined 156 cases and 162 controls and recently reported an association of the (*IL4*) C-590>T polymorphism and increased risk for oral squamous cell carcinoma (OSCC) (49). The association was significant in early stages of this malignancy (P < .0001; OR 3.17, 95% CI 1.31-7.65). The authors propose that *IL4* has a growth-promoting effect and show that this polymorphism increases expression of the gene.

Response to chemotherapy:

Studies have examined the relationship between SNPs in DNA repair genes and response to cisplatin, which kills cells by DNA crosslinking (50). The researchers analyzed four SNPs, *ERCC2 (XPD)*-Asp312Asn, *ERCC2 (XPD)*-Lys751Gln, *ERCC1*-C8092A, and *XRCC1*-Arg399Gln, in peripheral lymphocytes DNA from 103 stage IV SCCHN patients. They found a significant association between survival and polymorphic

variants. This work had built on previous findings of similar effects of these SNPs in non-small-cell lung cancer (51) and highlights the possibility of individualized chemotherapy regimens based on genotype.

2.2 ATM AND DNA REPAIR

Genes in regions implicated in tumor progression (Figure 1) may provide targets for susceptibility studies. Knudsen's two-hit hypothesis states that dominantly inherited cancers occur as a result of two mutation events that work to decrease or eliminate gene function. The first mutational event or "first hit" may occur as an inherited mutation, effectively increasing susceptibility to tumors. The second hit may occur somatically, resulting in cancer formation (52). It follows that if expression changes in genes are related to SCCHN progression, germline mutations in these genes may predispose an individual to carcinogenesis by providing the "first hit" in somatic cells. A second hit could then occur by endogenous or exogenous DNA damage, causing complete loss of heterozygosity. This is believed to be the mechanism for cancer development in a number of autosomal dominant cancer syndromes, including hereditary breast and ovarian cancer, where loss of heterozygosity of the *BRCA1* or *BRCA2* tumor suppressor function causes the observed cancers. Similarly, loss of ATM in tumors has been associated with carcinogenesis, and it is theoretically possible that germline mutations could increase susceptibility to SCCHN.

2.2.1 Function of *ATM*

Double strand breaks in DNA can occur as a result of exogenous factors such as ionization radiation (IR) or common endogenous activities such as recombination and the nucleotide excision repair pathway (53). In response to double strand DNA breaks (DSBs), the cell utilizes a number of proteins that sense the damage (sensors), transducers that amplify this signal and effectors that induce the response and repair (Figure 3).

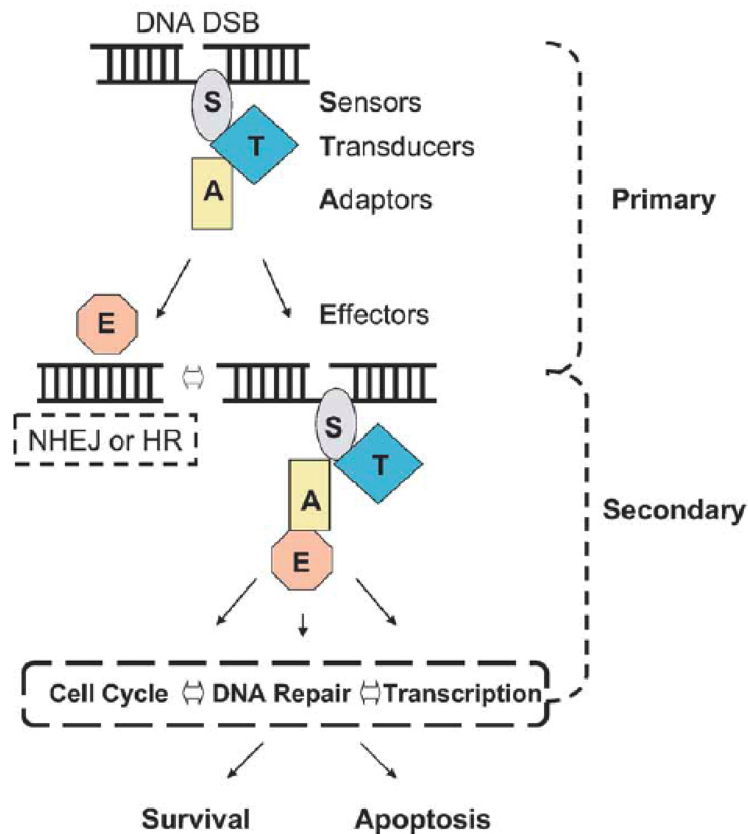

Figure 3. Model of DNA double-strand-break response pathway, (53).

These pathways are not completely understood, but it is clear that members of the PI3 kinase-related kinases (PI3KKs) play a role in the initial response. The most important PI3KKs

involved in the damage response are: Ataxia Telangiectasia Mutated (*ATM*) and Rad3 related (*ATR*). These kinases respond to double strand breaks by phosphorylating other key proteins involved with cell cycle control and DNA damage repair. In general, *ATM* responds in the G1 phase in response to IR-induced DSBs, while *ATR* function predominates in S/G2 phase in response to replication-associated DSBs or stalled replication forks. *ATM* interacts with a large number of other proteins such as checkpoint factors (*CHEK*'s), tumor suppressors *(TP53, BRCA1*), DNA repair factors (*RAD50*, *RAD51, GADD45*) and other signaling molecules (*ABL1*, *NFKB1*) (Figure 4).

Figure 4. The *ATM* signaling pathway, (53, 54).

This response can be broken down into the primary and secondary phases. In the primary response, sensor proteins detect DNA double strand breaks, transducer and adaptor proteins recruit effector proteins and repair occurs by NHEJ or homologous recombination (HR). A secondary response occurs when DSB's are abundant and severe. This results in delays in cell

cycle, transcription of certain genes, recruitment of other DNA repair proteins and if unsuccessful, apoptosis (53).

The *ATM* and *ATR* kinases appear to respond to damage both as primary sensors and as secondary transducers of the response signal. In the sensor capacity, *ATM* and *ATR* cooperate with *MRN* to phosphorylate *H2AFX*. *H2AFX*, in turn, functions to recruit a large DNA repairisome, including *BRCA1* to the site of damage (53). As transducers, *ATM* and *ATR* initiate a phosphorylation cascade through *BRCA1*, effecting *TP53*, *NBN* (*NBS1)* and *CHEK2.* This results in the initiation of cell cycle checkpoints and a delay in cell cycle progression (Figure 5).

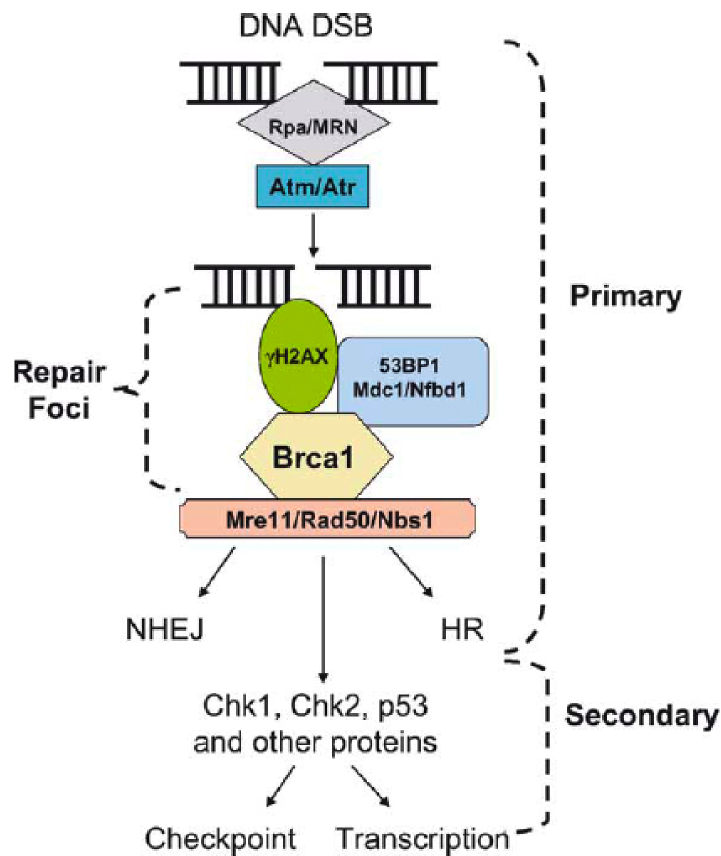

Figure 5. Model of DNA double-strand-break response genes, (53).

2.2.2 *ATM* and Disease

The *ATM* gene was first identified as the genetic cause of ataxia-telangiectasia (AT) in 1995 (55). The *ATM* gene is located at 11q22.3 and consists of 66 exons and is approximately 3Kb in length. Over 300 AT-causing mutations have been reported in the *ATM* gene all of which are either missense, nonsense, splicing or truncating mutations; the truncated proteins are almost completely inactive (56). AT is an autosomal recessive genomic instability disorder, resulting from an inability of cells to respond to double strand breaks. The disease is characterized by sensitivity to ionizing radiation and DSB-inducing agents, severe and progressive neuromotor dysfunction, immunodeficiency, genomic instability, and a predisposition to non-Hodgkin's lymphoma (40%), acute lymphocytic leukemia (20%) and Hodgkin's lymphoma (5%) (9, 57) in young adults and epithelial cancers (eg. breast cancer) in adults. Most people with AT are compound heterozygotes, but occasionally, homozygotes have been observed. Although AT is a rare disease (1 in 40,000), it is estimated that AT heterozygote frequency might fall between 0.68% and 7.7%, with 2.8% being a likely estimate (58). Therefore, even a modest increase in cancer risks associated with AT carrier status would have a significant public health impact. As a result, an intensive effort has been made to determine what risks are associated with AT heterozygosity.

2.2.2.1 *ATM* and breast cancer.

An association between AT heterozygotes and breast cancer, based on epidemiological evidence, was first reported 20 years ago (59). This study was based on AT families and

calculated the relative risk for (any) cancer to be 2.3 for men and 3.1 for women. When stratified for cancer type, the highest significance was found for breast cancer. Since then, a number of studies have attempted to confirm this epidemiological finding as well as identify *ATM* polymorphisms associated with increased risk for breast cancer. A recent epidemiological study by Thompson et al. (60) characterized the cancers observed in the relatives of AT patients (Table 1) and found the overall RR for breast cancer to be 2.23 overall and higher in women under 50 years old (RR=4.9).

Table 1. Cancer incidence in 1160 relatives of 132 AT patients, (60)**.**

Cancer site	ICD 9	Obs	Exp	RR	(95% CI)
Buccal cavity and pharynx	140–149	2	1.78	1.59	(0.15 to 16.8)
Esophagus	150	3	2.17	2.34	(0.47 to 11.6)
Stomach	151	10	4.74	3.39	(0.86 to 13.4)
Colorectal	152–154	20	12.1	2.54	(1.06 to 6.09)
Gallbladder	156	3	0.53	12.2	(1.26 to 118)
Pancreas	157	4	2.63	2.41	(0.34 to 17.1)
Lung	162	21	18.2	1.38	(0.64 to 2.97)
Breast (female)	174	23	14.6	2.23	(1.16 to 4.28)
Uterus	179	2	2.15	1.38	(0.09 to 22.4)
Ovary	183	3	2.67	1.90	(0.20 to 18.2)
Prostate	185	6	5.34	1.29	(0.30 to 5.48)
Bladder	188	5	5.22	1.41	(0.41 to 4.82)
Brain	191	2	1.93	0.06	(0.01 to 0.33)
Unknown	199	4	5.19	0.70	(0.10 to 4.92)
Myeloma	203	3	1.09	4.49	(0.32 to 62.2)
Other female genital	184	2	0.43	10.2	(0.30 to 345)
All sites except breast		95	82.1	1.47	(1.00 to 2.16)
Male: all sites		54	50.4	1.23	(0.76 to 2.00)
Female: all sites except breast		41	31.8	2.05	(1.09 to 3.84)

In general, studies whose design involved case/control screening for *ATM* variants in a breast cancer population have failed to show an association with breast cancer risk (61). This may be a result of small study size, screening modalities that were more likely to detect truncating mutations (rather than missense mutations), an inability to control for *BRCA1/2* status, or a focus on sporadic rather than familial cases (61). One recent study overcame these difficulties by studying *BRCA*-negative familial breast cancer cases and controls and screening the entire *ATM* gene. The group found that the relative risk of breast cancer associated with *ATM* mutations was estimated to be 2.37 (95% CI = 1.51–3.78, P = 0.0003), which agrees with previous epidemiological findings (56). Another study of similar design focused on a population of high-risk, non-Ashkenazi Jewish women who tested negative for *BRCA1/2* (61). The group found a significant association between one *ATM* SNP (rs228589) and risk for breast cancer OR =12.61 (95% CI =5.91 – 26.92, P = 4×10^{-9}). Additionally, other *ATM* alterations that have been found to be associated with breast cancer include: Val2424Gly (associated with increased risk of breast cancer in AT families) (62), and IVS10-6T>G (associated with early onset breast cancer) (63). Teraoka et al. examined the mutation type and found an association between missense mutations and breast cancer risk but with no association between truncating mutations and risk (64). This finding was supported by Dork et al. (65), but not in Thompson et al. (60). This leads to the possibility of two distinct *ATM* mutation types that confer different risks, however, more studies are needed to confirm this observation.

2.2.2.2 *ATM* and lung cancer

An association between *ATM* and lung cancer was first reported in a case-control study of 616 Korean lung cancer patients and 616 controls (66). This study also used a tag-SNP approach to select SNPs of interest and found an association between the IVS62 +60G>A (rs664143) allele and lung cancer risk [odds ratio = 1.6, 95% confidence interval (CI) 1.1–2.1]. Following haplotype analysis, they found and association between the ATTA haplotype at sites -4518A>G (rs189037), IVS21 -77C>T (rs664677), IVS61 -55T>C (rs664982) and IVS62 +60G>A (rs664143), and increased risk of lung cancer (OR = 7.6, 95% CI 1.7–33.5). Additionally, the (NN)TA haplotype was associated with a highly significant increased risk of lung cancer (OR = 13.2, 95% CI 3.1–56.1).

Landi et al. examined a different set of 4 *ATM* SNPs in a recent case-control study from Central and Eastern Europe. They found an association between the IVS48+238 C>G alteration (rs609429) and decreased risk for lung cancer (OR = 0.55, CI = 0.30-0.98, P = 0.03) (67).

2.2.2.3 *ATM* and squamous cell carcinoma of the head and neck (SCCHN)

A number of studies have shown loss of heterozygosity (LOH) in the 11q22-23 region (68-73) in SCCHN tumors. Lazar et al. found that 25% of SCCHN tumors showed LOH at 11q23 and upon follow-up, found that LOH in this region was associated with persistent or recurrent disease in patients receiving radiotherapy (RT). They suggest that this may indicate a resistance to RT, related to genes in this region (68). These findings have lead to efforts to examine possible mechanisms related to this observation.

Although it is unclear if *ATM* is lost in the above studies, it is located in or near the regions described and is therefore a candidate for further studies. Lingbao et al. built on previous work that demonstrated that hypermethylation *ATM* promoter results in reduced expression of

the gene. When examining *ATM* methylation in SCCHN, the group showed that approximately 25% of cases showed aberrant methylation and that this was associated with decreased overall patient survival (74).

As far as we know, this is the first study to examine germline mutations in *ATM* in patients with SCCHN.

3.0 METHODS

Research subjects were previously recruited as part of the Head and Neck SPORE and the Oral Cancer study at the University of Pittsburgh. Specifically, cases were individuals diagnosed with primary SCCHN within one year of enrollment of the study. Cases could have a history of another primary tumor but controls did not have any personal history of cancer, regardless of type. Controls were recruited from community settings and clinical settings (patients of the UPMC ENT Clinic or University of Pittsburgh Dental School). The racial, gender, and ethnic characteristics of the proposed subject population reflect the demographics of Pittsburgh and the surrounding area and/or the subject population of the University of Pittsburgh Medical Center. No exclusion criteria were based on race, ethnicity, gender, or HIV status.

3.1.1 Subject Recruitment

At the University of Pittsburgh Medical Center Ear, Nose and Throat Clinic and University of Pittsburgh Cancer Institute Hematology/Oncology Clinics, potential case and control research subjects were first identified by their primary care physician or clinical care team. If the patient had signed the Ear, Nose and Throat (ENT) Patient Data Registry (IRB # 010356), the subject was identified by a study coordinator. The research project was then discussed with the patient to assess interest in participation. If interested, the caregiver instructed

the patient to either contact the research team directly for additional information; or, if the patient had not signed the ENT Patient Data Registry, he/she was provided an IRB approved written Authorization for the Sharing of Health Information Related to Possible Participation in a Research Study. This allowed the caregiver to inform the researcher of the patient's interest, share health information (related to study inclusion), and allowed the involved researchers to contact the patient for additional discussion related to participation. The Clinical Research Coordinators conducted the verbal screening process and obtained written informed consent for potential case subjects, prior to the blood draw and interview.

At the University of Pittsburgh Dental School, Dental School students informed their patients of the research study and, if interested, the potential subject was directed to the location of the research team in the clinic. The research team conducted a verbal screening to determine eligibility and interest. If interested in participation, the researcher obtained written informed consent, obtained a blood sample, and conducted an interview.

3.1.2 Inclusion and Exclusion Criteria

Inclusion criteria (cases):

- Age 18-79 years on date of first diagnosis of qualifying head and neck cancer.
- Biopsy proven primary squamous cell cancer at a head and neck site. For the purposes of this research, head and neck cancer sites include primary tumors coding to Chapter 3 (Lip and Oral Cavity), Chapter 4 (Pharynx, including base of tongue, soft palate, and uvula), Chapter 5 (Larynx), or Chapter 6 (Nasal Cavity and Paranasal Sinuses) of the AJCC Cancer Staging Manual (6th edition).

- Study enrollment (acquisition of informed consent and collection of questionnaire data) on or before the one-year anniversary of the first diagnosis of the qualifying head and neck cancer.

Exclusion criteria (cases):

- Age less than 18 years on date of first diagnosis of qualifying head and neck cancer.
- Age more than 79 years on date of first diagnosis of qualifying head and neck cancer.
- Absence of a clinical pathology report documenting invasive cancer involving a head and neck site.
- Histopathologic diagnosis other than squamous cell carcinoma. Diagnoses were subject to verification by a Pathologist. The investigators removed a participant from this research study if the final pathology report did not confirm the provisional diagnosis used for study enrollment purposes. If the final pathology report did not demonstrate primary squamous cell carcinoma of the head and neck, the investigators removed the participant from the study and his/her data and blood sample was rendered anonymous and destroyed.
- More than one year elapsed time since date of first biopsy diagnosis of most recent primary squamous cell carcinoma at a head and neck cancer site.

Inclusion criteria (controls):

In order to be included the study, controls needed to meet the following inclusion criteria:

- Age: 18-80 years old on date of enrollment.

- Cancer history: No personal history of cancer at a head and neck site (based on eligibility screening interview and/or review of ENT or Dental Clinic medical record).
- Clinical examination (clinic controls): Clinical examination (by personal ENT physician or dentist), without clinical suspicion of head and neck cancer, based on testimony of clinician or review of primary medical records.
- County of residence (If a neighborhood control): Residence in Allegheny (including greater Pittsburgh), Butler, Green, Cambria, Jefferson, Fayette, Westmoreland, Washington, Beaver, Lawrence, Mercer, Crawford, Erie, Venango, Clarion, Armstrong, or Indiana county (in Pennsylvania); Jefferson (including Steubenville), Columbiana, or Mahoning (including Youngstown) county (in Ohio); or Ohio (including Wheeling), Brook, or Hancock county (in West Virginia).

Exclusion criteria (controls):

Controls were excluded from the study if they met any of the following criteria:

- Age less than 18 years or more than 80 years on date of enrollment.
- Self-report of personal history of cancer at a head and neck cancer site.
- (If a clinic control) Indication in ENT or dental clinic record of personal history of cancer at a head and neck cancer site.
- (If a clinic control) Physical findings, on head and neck clinical examination, that creates a suspicion of cancer at a head and neck cancer site.
- (If a neighborhood control) Residence in a location other than Allegheny (including greater Pittsburgh), Butler, Green, Cambria, Jefferson, Fayette, Westmoreland, Washington, Beaver, Lawrence, Mercer, Crawford, Erie, Venango, Clarion, Armstrong,

or Indiana county (in Pennsylvania); Jefferson (including Steubenville), Columbiana, or Mahoning (including Youngstown) county (in Ohio); or Ohio (including Wheeling), Brook, or Hancock county (in West Virginia).

3.1.3 Epidemiological variables

For the purposes of this study, the epidemiological variables collected were: age, gender, race (Caucasian or African American), family history of cancer (yes, no), and height and weight (BMI was calculated), smoking history (yes = at least one cigarette a day for six months or longer, no), alcohol consumption (yes = one or more drinks per month for one year or longer, no) and household smoking (yes = until the age of 18, your father, mother, or anyone else in your household smoked cigarettes, no), pack-years (number of years smoking history, multiplied by the number of cigarettes/day divided by 20).

3.1.4 Sample Processing

For both the OCC and HN-SPORE sample sets, whole blood was drawn into purple top EDTA vacutainer tubes and promptly frozen at -80 degrees C.

Oral Cancer Center samples:

This sample set was originally delivered to Dr. Robert Ferrell's lab at the University of Pittsburgh for genotyping, related to previous genotyping studies. Technicians in Dr. Marjorie Romkes' lab at the University of Pittsburgh had previously extracted the DNA using the Gentra System, Inc. (Minneapolis, MN, USA) Puregene® kit. When delivered to the Ferrell lab, these samples had initially performed poorly in a PCR reaction and were re-extracted (in Dr. Ferrell's

lab) using the Qiagen (Valencia, CA, USA) QIAamp DNA Blood Mini Kit. Following this second extraction, sample performances were greatly improved. Due to limited resources, the samples were whole-genome amplified by technicians in Dr. Ferrell's lab using the GenomiPhiTM DNA Amplification Kit by GE Healthcare (Piscataway, NJ, USA). These samples were further diluted to a final concentration of 1:1000 for PCR use.

HN-SPORE samples:

Technicians in Dr. Marjorie Romkes' lab at the University of Pittsburgh had previously extracted the DNA using the Gentra System, Inc. (Minneapolis, MN, USA) Puregene® kit. The samples were delivered directly to Dr. Susanne Gollin's lab and tested for PCR performance. Following successful amplification, samples were whole-genome amplified using the Qiagen (Valencia, CA, USA) Repli-g kit. Again, performance was tested and samples were diluted to optimal concentrations for PCR, depending on sample quality.

3.1.5 Identification of the SNP Set

Currently, a query in dbSNP (build 127), a part of the National Center of Bioinformatics Institute, returns over 800 publicly reported SNPs in the *ATM* gene. Due to limited resources, sequencing or genotyping of all SNPs was not possible. With this in mind, it is advisable to take advantage of known regions of linkage (haplotypes) within a gene to avoid redundancy in genotyping and to identify a subset of informative SNPs. This idea stems from the fact that individual alleles physically located near each other in a gene are often inherited together. This means that often there is also a correlation between genotypes at one SNP allele and its neighboring SNP allele. Genotyping SNPs that mark known haplotypes (tagSNPs) can allow

inference of SNP genotypes at other sites in the linkage block. This correlation is referred to as Linkage Disequalibrium (LD). The combination of alleles along a chromosome is described as a haplotype.

The publicly available data and software from the International HapMap project has made it possible to search for tagSNPs. The project was launched in 2002 to "create a public database of common human sequence variation, providing information to guide genetic studies of clinical phenotypes" (75). Tagger (Haploview), a tool for selection of tagSNPs, was created by Paul de Bakker in the labs of David Altshuler and Mark Daly at the Center for Human Genetic Research of Massachusetts General Hospital and Harvard Medical School, and the Broad Institute for the HapMap project. The tool is publicly available and described in (76) "Efficiency and power in Genetic Association Studies". This software allows the researcher to query the HapMap data based on a number of variables:

Ethnicity: HapMap data were collected from four populations, the Yorubas in Nigeria, Japanese, Han Chinese and U.S. residents with ancestry from northern and western Europe (CEU). Because LD occurs due to shared ancestry, haplotypes vary between these populations. For the purpose of this study, the data queries were restricted to "CEU" which most closely matches the ethnic background of our patient population.

Statistical TagSNP-Picking Method: The software allows tags to be picked using either a pairwise or multimarker method. Descriptions of these two methods can be found in de Bakker et al. (76). Multimarker methods aggressively search for a haplotype of tags that can be used to infer other SNPs. This method decreases the overall number of tagSNPs needed (as compared to

the pairwise method) while maintaining power. Therefore, multimarker tag picking was used for tagSNP-picking.

Coefficient of Determination: The Tagger software also lets users pick a minimal coefficient of determination (r2) at which all alleles are to be captured. By default, this was set to 0.8.

Minimum Allele Frequency: The minimum allele frequency allows the researcher specify a minimum minor allele frequency to be considered by the algorithm. The minimum allele frequency was set to 5% due to power limitations at lower allele frequencies.

Using these settings, a total number of eleven tag SNPs were picked for the initial genotyping. These tag SNPs were supplemented with one additional SNP, predicted to function as a splice-site. Two rounds of genotyping were necessary to obtain nine functional tag SNP assays and one splice-site assay (Table 2, Figure 6). Two tag SNP assays failed repeatedly and no suitable replacement assay was available. In total, the nine tag SNPs captured 60 of 64 ATM alleles (with a frequency over 5%). The mean R^2 was 0.998 with the lowest R^2 being 0.891.

Table 2. Genotyped SNPs.

ATM SNP	SNP location	dbSNP Position	SNP type	Alias
rs4987886	Intron 1	11658469	Tag	
rs4987889	Intron 1	11658701	Tag	
rs228591	Intron 1	11659749	Tag	
rs11212570	Intron 16	11693128	Tag	
rs3092991	Intron 18	11702932	Tag	
rs1800889	Exon 30	11725903	Tag	Pro1526Pro
rs611646	Intron 37	11739513	Tag	
rs17503908	Intron 57	11777813	Tag	
rs373759	Intron 59	11783073	Tag	
rs17174393	Intron 60	11787024	Splice site	

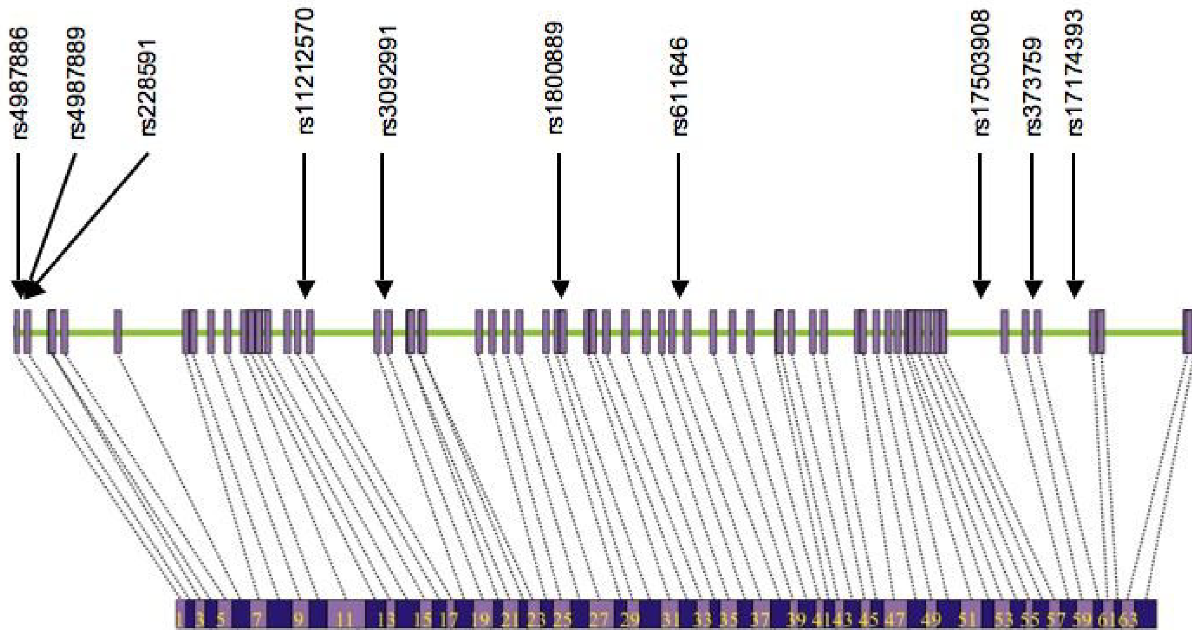

Figure 6. Genotyped SNPs relative to the ATM gene. Adapted from the ATM gene model by the Seattle Mouse Comparative Genomics Center (77)

3.1.6 Genotyping

TaqMan® SNP Genotyping Assays were purchased from Applied Biosystems (Foster City, CA, USA). The TaqMan® genotyping assays are based on real-time PCR technology, using allele-specific probes and primers. The probes are labeled with either VIC or FAM dye and selectively anneal to the corresponding DNA sequence (SNP). The detection is made possible by exonuclease activity of the Taq polymerase and the cleavage of the dye-label. Following the arraying of DNAs in 96-well plate format, the plates and assays were delivered to the University of Pittsburgh Genomics and Proteomics Core Laboratory (GPCL) for genotyping. The 96-well DNA plates were then rearrayed into a 384-well plate format. PCR was performed according to Applied Biosystems specifications using the Applied Biosystems 7900HT instrument. (AppliedBiosystems.com)

Each well contained:

2.5uL TaqMan® Universal PCR Master Mix (2), No AmpErase® UNG

0.25uL SNP Genotyping Assay Mix (20)

2.25uL sample DNA (diluted 1:20 in nuclease-free water)

5uL Total

One "No Template Control" (NTC) containing TE or Nuclease-free water was added by the GPCL to the last well of each plate. The PCR profile used was: Hold 2 minutes at 50 degrees C, hold for 10 minutes at 95 degrees C, followed by 40 cycles of: 15 second hold at 95 degrees C and hold 1 minute at 60 degrees C.

Results were returned from the GPCL as .SDS files, which contain all raw fluorescent values. Due to data quality concerns, the raw spectral data were analyzed in two separate ways and all statistical analyses were performed on both types of analyses. First, genotype calls were made by hand, by calling according to cluster. In some SNPs, there was concern that about the accuracy of these genotype calls. Therefore, a second, non-standard method was developed to supplement these calls. This method consisted of converting the raw allele X and allele Y spectral data to a single value, approximating the slope from the origin (relative to each SNP). This was done using the following formula where Y= the raw Y fluorescence value for a particular data point (representing the Y allele), and X= raw X fluorescence for a particular data point (representing the X allele): Y/(X+Y). To adjust for background (different points of origin for each SNP) the working equation became: (Y-Ymin)/[(X-Xmin) + (Y-Ymin)] where Ymin and X min represented the lowest value observed for each axis.

Blinded quadruplicate positive control samples of (unknown genotype) were included for each assay to insure proper genotype assignment and data integrity. Three to four positive controls of known genotype were included per 384-well plate. These were samples from the CEPH collection and ordered through the National Institute of General Medical Sciences (NIGMS). Researchers were blinded to case-control status until data were generated and genotype calls were made. All SOPs regarding laboratory safety were followed as well as separation of pre and post-PCR products to avoid contamination.

3.1.7 Statistical Analysis

Allele frequencies were calculated for both cases and controls and goodness of fit to the expected Hardy-Weinberg frequencies was measured in controls. The χ^2 test was used to

examine genetic differences between cases and controls and P-values were generated for each SNP. The χ^2 test was also used to examine the significance of each epidemiological risk factor (perceived to play a role in SCCHN based on the literature or the biology of the disease) relative to case-control status. For continuous variables "age" and "pack-years of smoking", data were binned according to the mean value of the controls. Body mass index (BMI) was binned into biologically relevant categories (<18.5, 18.6-25, 25.1-30, and >30 kg/m^2) and not according to means. For those variables that were found to be significant, a logistic regression model was created to test for SNP significance, controlling for these factors. The significant factors included in the regression model were: age (as a continuous variable), sex, pack-years (continuous; included instead of smoking history), alcohol history and BMI (as a continuous variable). This binary logistic regression model was used to test for disease associations in both result sets: genotype calls and the calculated slope value. In summary, the data were examined in a number of ways: genotype calls were examined for their association with case/control status, in a logistic regression (unadjusted and adjusted). The slope-calculation was also examined in a logistic regression (unadjusted and adjusted). All analyses were carried out using the SPSS statistical 15.0 software.

4.0 RESULTS

4.1.1 Epidemiological variables and SCCHN risk

Of the major epidemiological variables examined in the study, those found to be associated with case/control status were: sex [OR=1.7, (CI: 1.2-2.3)], with the higher risk in males; positive smoking history [OR=1.5 (CI: 1.02-2.1)]; pack-years [OR=2.3 (CI: 1.7-3.2)], with the highest risk in those who reported >20 pack-years; alcohol use [OR=2.0 (CI: 1.2-3.3)], with the higher risk in drinkers; and BMI of 18.6-25.0 [OR=1.7 (CI: 1.24-2.38) (Table 3).

Table 3. Epidemiological variables.

Variables		Case (N=300)	Control (N=360)	OR	95% CI
Age: >57		161 (52.8)	180 (49.9)	1.1	0.8-1.5
Sex (number male) (%)		232 (76.1)	237 (65.7)	**1.7**	**1.2-2.3**
Race: number caucasian (%)		302 (99)	358 (99.2)	0.8	0.2-4.2
History of cigarette smoking: number "yes" (%)		237 (78.5)	257 (71.4)	**1.5**	**1.0-2.1**
Packyears: >20		185 (61.7)	148 (41.1)	**2.3**	**1.7-3.2**
Household cigarette use: number "yes" (%)		223 (74.6)	248 (68.7)	1.3	0.9-1.9
Family history of cancer: number "yes" (%)		180 (59.8)	208 (57.8)	1.1	0.8-1.5
Alcohol use: number "yes" (%		278 (91.7)	305 (84.7)	**2.0**	**1.2-3.3**
BMI (kg/m^2):	**</=18.5**	11 (3.7)	7 (1.9)	1.9	0.74-5.03
	18.6-25.0	119 (39.8)	100 (27.8)	**1.7**	**1.24-2.38**
	25.1-30.0	99 (33.1)	147 (40.8)	**0.7**	**0.52-0.99**
	>30	70 (23.4)	106 (29.4)	0.7	0.52-1.04

The means of continuous variables were calculated in cases and controls. The mean age of the controls was 57.1 years and 58.0 years in cases. The mean BMI in controls was 27.8 compared to a mean BMI of 26.7 in cases. The mean pack-years smoked in controls was 20.5 compared to 37.5 pack-years in cases (Table 4).

Table 4. Epidemiological variables.

Variables	Case (N=300)	Control (N=360)
Age (mean+/- SD) (years)	58.0 +/- 10.5	57.1 +/- 11.5
BMI (mean+/- SD) (kg/m^2)	26.7 +/- 6.1	27.8 +/- 5.2
Packyears (mean =/- SD)	37.5 +/- 35.9	20.5 +/- 23.8

4.1.2 Association of ATM SNPs and SCCHN risk

Of a total of 300 cases and 360 controls genotyped for 10 ATM SNPs, the genotyping success rate ranged between 79 and 94% for all SNPs with a mean success rate of 89%. Each SNP was examined for deviations from Hardy-Weinberg and all controls were found to be in Hardy-Weinberg equilibrium. Allele frequencies and P-values for SCCHN cases and controls for each of the ten SNPs are shown in Table 5. Significant differences in genotype frequencies were observed between SCCHN cases and controls for rs611646 (P = 0.012) and rs373759 (P = 0.025). For rs611646, the TT genotype was more common in cases than controls (45 versus 32%) while the AA and AT genotypes were less common in cases than controls (18 versus 21% and 37 versus 46%, respectively). In SNP rs373759, the CC genotype was more common in

cases than controls (56 versus 48%) with the CT type being less common in cases than controls (29 versus 40%).

Table 5. Allele frequencies**.**

SNP	Total	Genotype (%)			Allele Frequency		P-value
Intron 1 (rs4987886)		AA	AT	TT	A	T	
SCCHN cases	279	253 (90.7)	24 (8.6)	2 (0.7)	0.95	0.05	
Controls	322	305 (94.7)	16 (5.0)	1 (0.3)	0.97	0.03	0.156
Intron 1(rs4987889)		CC	CT	TT	C	T	
SCCHN cases	279	269 (96.4)	10 (3.6)	0 (0)	0.98	0.02	
Controls	328	317 (96.6)	11 (3.4)	0 (0)	0.98	0.02	0.877
Intron 1(rs228591)		AA	AG	GG	A	G	
SCCHN cases	272	66 (24.3)	113 (41.5)	93 (34.2)	0.45	0.55	
Controls	310	64 (20.6)	136 (43.9)	110 (35.5)	0.43	0.57	0.576
Intron 16 (rs11212570)		AA	AG	GG	A	G	
SCCHN cases	261	6 (2.3)	37 (14.2)	218 (83.5)	0.09	0.91	
Controls	316	6 (1.9)	46 (14.6)	264 (83.5)	0.09	0.91	0.94
Intron 18 (rs3092991)		AA	AG	GG	A	G	
SCCHN cases	269	199 (74.0)	62 (23.0)	8 (3.0)	0.86	0.14	
Controls	289	228 (78.9)	53 (18.3)	8 (2.8)	0.88	0.12	0.375
Exon 30 (rs1800889)		CC	CT	TT	C	T	
SCCHN cases	284	257 (90.5)	27 (9.5)	0 (0)	0.95	0.05	
Controls	328	294 (89.6)	34 (10.4)	0 (0)	0.95	0.05	0.724
Intron 37 (rs611646)		AA	AT	TT	A	T	
SCCHN cases	250	44 (17.6)	93 (37.2)	113 (45.2)	0.36	0.64	
Controls	279	59 (21.1)	129 (46.2)	91 (32.6)	0.44	0.56	**0.012**
Intron 57 (rs17503908)		GG	GT	TT	G	T	
SCCHN cases	286	3 (1.0)	42 (14.7)	241 (84.3)	0.08	0.92	
Controls	342	3 (0.9)	60 (17.5)	279 (81.6)	0.10	0.90	0.616
Intron 59 (rs373759)		CC	CT	TT	C	T	
SCCHN cases	266	149 (56.0)	77 (28.9)	40 (15.0)	0.70	0.30	
Controls	320	152 (47.5)	127 (39.7)	41 (12.8)	0.67	0.33	**0.025**
Intron 60 (rs17174393)		AA	AG	GG	A	G	
SCCHN cases	290	0 (0)	3 (1.0)	287 (99.0)	0.01	0.99	
Controls	336	0 (0)	6 (1.8)	330 (98.2)	0.01	0.99	0.431

A logistic regression analysis was performed for each SNP with and without covariates: sex, BMI, alcohol, and pack-years. Although smoking history was significantly related to case/control status, pack-years was used in the model because it is a more specific measure of smoking. The logistic regression included one *ATM* SNP at a time and separate analyses were performed for genotype calls and the slope-calculation. There were no significant differences between the adjusted and unadjusted regression analyses; adjusted data are shown in Tables 6 and 7. SNP rs611646 remained significant when adjusted for risk factors in both the genotype call analysis [OR = 1.38, (CI: 1.07-1.77] and the slope-calculation analysis [OR= 4.04, (CI: 1.53-10.67)]. SNP rs373759 appeared to be significant when examined before logistic regression, but not after the regression analysis. This can be explained by the apparent over-dominant effect observed in this SNP. The logistic regression, which codes the genotypes as 0, 1, and 2 is specifically looking for a linear effect in the data.

Rs3092991, on the other hand, was only found to be associated with SCCHN in the slope-calculation analysis. Theoretically, this could be explained by any unexpected grouping of cases or controls within genotype clusters, which would cause the slope-calculation to perceive differences within genotype groups. To test this hypothesis, case/control status was graphed along with raw data reads (Figure 7). There were no obvious visual differences within genotype clusters. Further examination of the genotype data is needed to explain the difference in significance observed in these two analyses.

Table 6. Sex, BMI, alcohol, and pack-years-adjusted odds ratios (ORs) for ATM SNP genotype calls.

ATM SNP	OR	95% Confidence Interval	P-value
Intron 1 (rs4987886)	1.773	0.959-3.275	0.068
Intron 1 (rs4987889)	0.928	0.368-2.338	0.874
Intron 1 (rs228591)	0.883	0.698-1.117	0.299
Intron 16 (rs11212570)	1.072	0.712-1.613	0.739
Intron 18 (rs3092991)	1.152	0.816-1.627	0.422
Exon 30 (rs1800889)	0.923	0.527-1.618	0.780
Intron 37 (rs611646)	**1.378**	**1.073-1.769**	**0.012**
Intron 57 (rs17503908)	1.235	0.820-1.860	0.312
Intron 59 (rs373759)	0.914	0.718-1.164	0.466
Intron 60 (rs17174393)	1.97	0.476-8.159	0.349

Table 7. Sex, BMI, alcohol, and pack-years-adjusted odds ratios (ORs) for ATM SNP slope-calculation.

ATM SNP	OR	95% Confidence Interval	P-value
Intron 1 (rs4987886)	0.868	0.230-3.273	0.835
Intron 1 (rs4987889)	2.471	0.450-13.563	0.298
Intron 1 (rs228591)	0.77	0.387-1.531	0.456
Intron 16 (rs11212570)	1.191	0.500-2.833	0.693
Intron 18 (rs3092991)	**4.116**	**1.593-10.632**	**0.003**
Exon 30 (rs1800889)	0.634	0.157-2.563	0.523
Intron 37 (rs611646)	**4.04**	**1.533-10.67**	**0.005**
Intron 57 (rs17503908)	1.116	0.487-2.554	0.796
Intron 59 (rs373759)	0.954	0.544-1.673	0.869
Intron 60 (rs17174393)	1.41	0.482-4.120	0.530

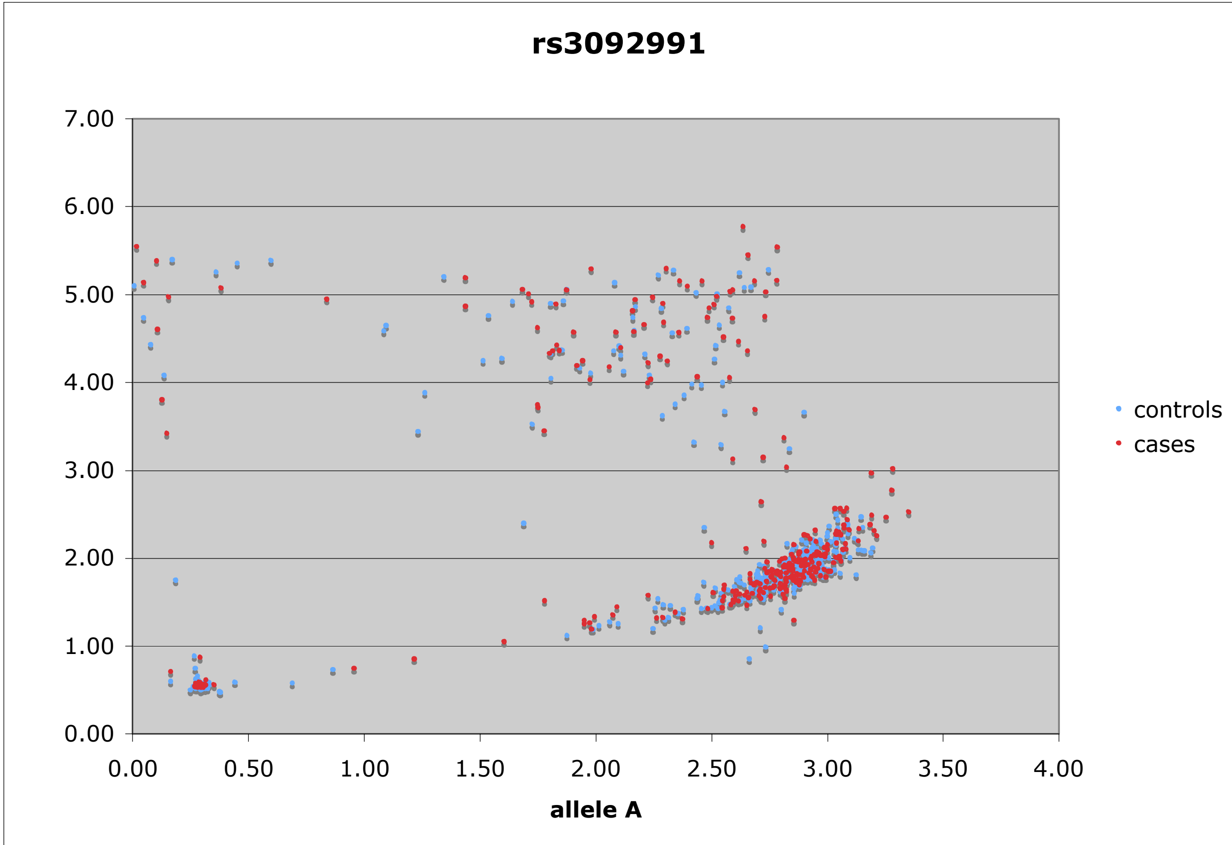

Figure 7. Graph of rs3092991 genotype data, labeled with case/control status.

5.0 DISCUSSION

The *ATM* gene, located near the distal end of chromosome 11q at 11q22.3, is part of a region implicated in SCCHN tumor progression from dysplasia to carcinoma in situ. About 30-76% of SCCHN tumors show loss of the distal end of chromosome 11q and subsequent amplification of 11q13 (32). The protein coded by the *ATM* gene is a kinase that responds to double strand breaks by phosphorylating other key proteins involved in cell cycle control and DNA damage repair. *ATM* interacts with a large number of other proteins, such as checkpoint factors (*CHEK*s), tumor suppressors *(TP53, BRCA1*), DNA repair factors (*RAD50*, *RAD51, GADD45*) and other signaling molecules (*ABL1*, *NFKB1*). This provides both a positional (chromosome 11q) and functional argument for the role of *ATM* in SCCHN.

5.1.1 Study design

ATM has long been implicated in increased risk for breast cancer through epidemiological studies (59). For many years, molecular biologists were unable to uncover the *ATM* variations that conferred odds ratios such as those seen in the epidemiological studies. Recently, however, a number of studies have successfully found associations that finally seem to explain the early epidemiological studies (56). The success of these studies highlights the importance of an emphasis on study design and the careful controlling of major genetic and environmental risk

factors of disease. Because *BRCA1/2* play such a major role in breast cancer susceptibility, it was suggested that study power would be increased if *ATM* variation were studied in sporadic breast cancer cases (61). Conversely, a study of SCCHN, which has no known major genetic risk factors may instead benefit by focusing on individuals from families with a history of related cancer (for example oral cancer or lung cancer). If germline variations are playing a role in SCCHN susceptibility, we may expect to see higher rates within a family carrying one of these gene variants. This type of design may therefore increase power to detecting genetic associations with SCCHN.

In addition, some associations between ATM SNPs breast cancer have been found in AT families only, assumingly because these *ATM* polymorphisms are known to be deleterious (62). This observation seems to indicate that a study design which examines SCCHN in AT carriers may have greater power to detect an association; however, in a study of 1160 relatives of AT patients, Thompson et al. (60) only observed two cases of buccal/pharangeal cancer and three cases of esophageal cancer. Therefore, although this study design may be desirable, it is likely to be difficult to obtain enough samples to make the data informative. This type of design would also miss the effects of non-AT-causing variants in *ATM*.

It may also be helpful to be able to control for HPV status in both cases and controls. Because HPV is a significant risk factor, with a unique etiology of disease, it would be helpful include HPV as a covariate when performing a regression analysis. It would also be interesting to examine the role of *ATM* variants in relation to HPV status, as there may be variants that affect disease susceptibility in the presence or absence of HPV infection.

5.1.2 Genetic associations

ATM has not been previously studied in SCCHN; however, a number of other cancers have been associated with *ATM,* including breast cancer and lung cancer. Specifically, breast cancer studies have implicated rs228589 (61), rs36017433 (62), and IVS10-6T>G (out-of-frame deletion, truncation at amino acid 419) (63). At the time of this study, these SNPs were not in the Hapmap database and it is unknown whether the SNPs used in this study are in LD with these variants.

Lung cancer studies (66) have found associations between a number of *ATM* variants and increased risk for disease: IVS62 +60G>A (rs664143) and the ATTA haplotype at sites: -4518A>G (rs189037), IVS21 -77C>T (rs664677), IVS61 -55T>C (rs664982) and IVS62 +60G>A (rs664143). The previously associated SNPs in the haplotype (above) are in high LD with each other (84-98%). In addition, the variant, IVS48+238 C>G (rs609429) has been associated with decreased risk for lung cancer (67). This SNP is also not in the Hapmap database and it is unknown whether any of the SNPs included in this study are in LD with rs609429.

This study analyzed both genotype calls and the slope-calculation for each SNP and found significant associations with three ATM SNPs and risk for SCCHN: rs611646, rs373759, and rs3092991.

Rs611646: Located in intron 37, this SNP was included in the study as a tag-SNP. This SNP is part of a large haplotype block, containing many SNPs with an R^2 value of 1.0. This makes it difficult to locate a functional SNP responsible for this observation.

Rs373759: This SNP was also picked as a tag-SNP and is located in intron 59. Together with rs611646, rs373759 is in the same large haplotype block as rs664143 and confirms previous findings that associated this SNP with lung cancer (66).

Rs3092991 was only found to be associated with disease in the slope-calculation analysis. This analysis was performed on every SNP due to data quality concerns and effectively assigned a single slope-value to each data point. This finding was particularly puzzling in that the data for this SNP was one of the highest quality and genotype calls were thought to be very reliable. It was hypothesized that differences between the slope-calculation and genotype analysis would be possible if there were subgroups of cases or controls that clustered within the larger genotype groups. Upon visual inspection of the data distribution, there did not appear to be any such sub-clustering. In addition, when genotypes were examined separately, no associations were found in the genotype data. This association may be an artifact of the distribution of data in that the major allele shows higher background for this SNP as raw intensities rise. This causes a slight curving of the distribution and may make the slope-calculation less accurate than genotype calls, themselves.

5.1.3 Epidemiology

Basic epidemiological analyses replicated findings of previous studies, showing that alcohol [OR=2.0, (CI:1.2-3.3)], cigarette smoking, in terms of "pack-years" [OR=2.3, (CI:1.7-3.2); OR=1.5], and male gender [OR=1.7, (CI:1.2-2.3)] appear to be major risk factors for disease. Age was not found to be associated with disease, however, these analyses are not meaningful given that cases and controls were age-matched as part of the study design. A

comprehensive analysis of interactions between these variables was not performed, as this was not the focus of the study. However, examination of SNP associations and ORs before regression and after regression did not reveal any significant interaction between the SNPs and epidemiological risk factors. Further analyses that stratify according to risk factors, simultaneously (eg. smoking and alcohol status), may provide further insight into the mechanism of SNP associations. Lower BMIs (18.6-25.0 kg/m^2) were also associated with disease [OR=1.7, (CI:1.24-2.38)], although this may be a result of disease, rather than a cause.

5.1.4 Future directions

The significant associations found in the study warrant further haplotype analyses in the region of rs611646 and rs373759. Preliminary results of the multi-locus genotype analysis can be seen in Table 8. Future analyses should also examine interactions between genotypes and covariates, especially in the presence/absence of smoking and alcohol. This may give further insight into the role of ATM in disease and possible biological mechanisms of association.

Table 8. Joint genotype analysis of rs611646 and rs373759.

rs611646/rs373759	Cases	%	Controls	%
AA/CC	3	1.3	4	1.5
AA/CT	8	3.5	22	8.4
AA/TT	30	13.0	28	10.7
AT/CC	26	11.3	29	11.1
AT/CT	59	25.7	87	33.3
AT/TT	1	0.4	2	0.8
TT/CC	101	43.9	84	32.2
TT/CT	1	0.4	5	1.9
TT/TT	1	0.4	0	0.0

Finally, sample quality was a major issue in this study. Because very little genomic DNA was available, it was necessary to perform a -genome amplification. This greatly reduced the data quality and made a number of assays difficult to interpret. In an attempt to correct for the difficulty in making accurate genotype calls, a new unbiased calculation was performed (the slope-calculation). This calculation effectively gave a value to the slope of each data-point from the origin and could be used similar to other variables in the regression analyses. For nine out of ten SNPs analysis of this slope-calculation agreed with the conclusions of the genotype calls. In one SNP, however (rs3092291), the analysis of the slope-calculation showed significance whereas the genotype call data did not. An attempt was made to explain this observation, and it is this author's opinion that the unusual distribution of the data may have caused the discrepancy. Further examination of genotype data for SNP rs3092991 is necessary to tease apart differences in the two methods of data-calling, and interpret results of this SNP.

BIBLIOGRAPHY

1. Dobrossy L. Epidemiology of head and neck cancer: magnitude of the problem. Cancer Metastasis Rev 2005;24(1):9-17.

2. Reshmi SC, Gollin SM. Chromosomal instability in oral cancer cells. J Dent Res 2005;84(2):107-17.

3. Jemal A, Siegel R, Ward E, Murray T, Xu J, Thun MJ. Cancer statistics, 2007. CA Cancer J Clin 2007;57(1):43-66.

4. Llewellyn CD, Linklater K, Bell J, Johnson NW, Warnakulasuriya KA. Squamous cell carcinoma of the oral cavity in patients aged 45 years and under: a descriptive analysis of 116 cases diagnosed in the South East of England from 1990 to 1997. Oral Oncol 2003;39(2):106-14.

5. Wiseman SM, Swede H, Stoler DL, Anderson GR, Rigual NR, Hicks WL, Jr., et al. Squamous cell carcinoma of the head and neck in nonsmokers and nondrinkers: an analysis of clinicopathologic characteristics and treatment outcomes. Ann Surg Oncol 2003;10(5):551-7.

6. Farshadpour F, Hordijk GJ, Koole R, Slootweg PJ. Non-smoking and non-drinking patients with head and neck squamous cell carcinoma: a distinct population. Oral Dis 2007;13(2):239-43.

7. Ragin CC, Modugno F, Gollin SM. The epidemiology and risk factors of head and neck cancer: a focus on human papillomavirus. J Dent Res 2007;86(2):104-14.

8. Perez-Ordonez B, Beauchemin M, Jordan RCK. Molecular biology of squamous cell carcinoma of the head and neck. J Clin Pathol 2006;59(5):445-453.

9. Shiloh Y. ATM and related protein kinases: safeguarding genome integrity. Nat Rev Cancer 2003;3(3):155-68.

10. Peterson CL, Cote J. Cellular machineries for chromosomal DNA repair. Genes Dev 2004;18(6):602-16.

11. Wei Q, Wang LE, Sturgis EM, Mao L. Expression of nucleotide excision repair proteins in lymphocytes as a marker of susceptibility to squamous cell carcinomas of the head and neck. Cancer Epidemiol Biomarkers Prev 2005;14(8):1961-6.

12. Al-Tassan N, Chmiel NH, Maynard J, Fleming N, Livingston AL, Williams GT, et al. Inherited variants of MYH associated with somatic G:C-->T:A mutations in colorectal tumors. Nat Genet 2002;30(2):227-32.

13. Risinger MA, Groden J. Crosslinks and crosstalk: human cancer syndromes and DNA repair defects. Cancer Cell 2004;6(6):539-45.

14. Brady N, Gaymes TJ, Cheung M, Mufti GJ, Rassool FV. Increased error-prone NHEJ activity in myeloid leukemias is associated with DNA damage at sites that recruit key nonhomologous end-joining proteins. Cancer Res 2003;63(8):1798-805.

15. Choi SY, Kahyo H. Effect of cigarette smoking and alcohol consumption in the aetiology of cancer of the oral cavity, pharynx and larynx. Int J Epidemiol 1991;20(4):878-85.

16. Castellsague X, Quintana MJ, Martinez MC, Nieto A, Sanchez MJ, Juan A, et al. The role of type of tobacco and type of alcoholic beverage in oral carcinogenesis. Int J Cancer 2004;108(5):741-9.

17. IRAC. Tobacco Smoking. IRAC Monograph on the Evaluation of the Carcinogenic Risk of Chemicals to Humans, Vol. 38. In: Lyon, IRAC Press; 1986.

18. Peters ES, McClean MD, Marsit CJ, Luckett B, Kelsey KT. Glutathione s-transferase polymorphisms and the synergy of alcohol and tobacco in oral, pharyngeal, and laryngeal carcinoma. Cancer Epidemiol Biomarkers Prev 2006;15(11):2196-202.

19. van Houten VM, Snijders PJ, van den Brekel MW, Kummer JA, Meijer CJ, van Leeuwen B, et al. Biological evidence that human papillomaviruses are etiologically involved in a subgroup of head and neck squamous cell carcinomas. Int J Cancer 2001;93(2):232-5.

20. Schwartz SM, Daling JR, Doody DR, Wipf GC, Carter JJ, Madeleine MM, et al. Oral cancer risk in relation to sexual history and evidence of human papillomavirus infection. J Natl Cancer Inst 1998;90(21):1626-36.

21. Martin-Granizo R, Rodriguez-Campo F, Naval L, Diaz Gonzalez FJ. Squamous cell carcinoma of the oral cavity in patients younger than 40 years. Otolaryngol Head Neck Surg 1997;117(3 Pt 1):268-75.

22. Bosetti C, Negri E, Franceschi S, Conti E, Levi F, Tomei F, et al. Risk factors for oral and pharyngeal cancer in women: a study from Italy and Switzerland. Br J Cancer 2000;82(1):204-7.

23. Novacek G. Plummer-Vinson syndrome. Orphanet Journal of Rare Diseases 2006;1(36):4.

24. McLaughlin JK, Gridley G, Block G, Winn DM, Preston-Martin S, Schoenberg JB, et al. Dietary factors in oral and pharyngeal cancer. J Natl Cancer Inst 1988;80(15):1237-43.

25. Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. Cancer 1953;6(5):963-8.

26. Bedi GC, Westra WH, Gabrielson E, Koch W, Sidransky D. Multiple head and neck tumors: evidence for a common clonal origin. Cancer Res 1996;56(11):2484-7.

27. Partridge M, Gaballah K, Huang X. Molecular markers for diagnosis and prognosis. Cancer Metastasis Rev 2005;24(1):71-85.

28. Rogers SJ, Harrington KJ, Rhys-Evans P, P OC, Eccles SA. Biological significance of c-erbB family oncogenes in head and neck cancer. Cancer Metastasis Rev 2005;24(1):47-69.

29. Jefferies, Kote J, Goldgar, Houlston, Frazer W, A'Hern, et al. Association between polymorphisms of the GPX1 gene and second primary tumours after index squamous cell cancer of the head and neck. Oral Oncology 2005;41(5):455-461.

30. Brennan JA, Boyle JO, Koch WM, Goodman SN, Hruban RH, Eby YJ, et al. Association between cigarette smoking and mutation of the p53 gene in squamous-cell carcinoma of the head and neck. N Engl J Med 1995;332(11):712-7.

31. Boyle JO, Hakim J, Koch W, van der Riet P, Hruban RH, Roa RA, et al. The incidence of p53 mutations increases with progression of head and neck cancer. Cancer Res 1993;53(19):4477-80.

32. Muller D, Millon R, Velten M, Bronner G, Jung G, Engelmann A, et al. Amplification of 11q13 DNA markers in head and neck squamous cell carcinomas: Correlation with clinical outcome. European Journal of Cancer 1997;33(13):2203-2210.

33. Reshmi SC, Roychoudhury S, Yu Z, Feingold E, Potter D, Saunders WS, et al. Inverted duplication pattern in anaphase bridges confirms the breakage-fusion-bridge (BFB) cycle model for 11q13 amplification. Cytogenetic and Genome Research 2007;116(1-2):46-52.

34. Shama Buch BZ, Autumn Gaither Davis, Dominic Odom, Jill M. Siegfried, Jennifer R. Grandis, Marjorie Romkes. Association of polymorphisms in the cyclin D1 and XPD genes and susceptibility to cancers of the upper aero-digestive tract. Molecular Carcinogenesis 2005;42(4):222-228.

35. Izzo JG, Papadimitrakopoulou VA, Liu DD, den Hollander PLC, Babenko IM, Keck J, et al. Cyclin D1 Genotype, Response to Biochemoprevention, and Progression Rate to Upper Aerodigestive Tract Cancer. J. Natl. Cancer Inst. 2003;95(3):198-205.

36. Zheng Y, Shen H, Sturgis EM, Wang L-E, Eicher SA, Strom SS, et al. Cyclin D1 polymorphism and risk for squamous cell carcinoma of the head and neck: a case-control study. Carcinogenesis 2001;22(8):1195-1199.

37. Malgorzata R, Pawel G, Daniela M-K, Wojciech G, Krzysztof S. Cyclin D1 gene (CCND1) polymorphism and the risk of squamous cell carcinoma of the larynx. European Archives of Oto-Rhino-Laryngology 2006;263(1):43-48.

38. Li G, Liu Z, Sturgis EM, Shi Q, Chamberlain RM, Spitz MR, et al. Genetic polymorphisms of p21 are associated with risk of squamous cell carcinoma of the head and neck. Carcinogenesis 2005;26(9):1596-1602.

39. Olshan AF, Watson MA, Weissler MC, Bell DA. XRCC1 polymorphisms and head and neck cancer. Cancer Letters 2002;178(2):181-186.

40. Sturgis EM, Castillo EJ, Li L, Zheng R, Eicher SA, Clayman GL, et al. Polymorphisms of DNA repair gene XRCC1 in squamous cell carcinoma of the head and neck. Carcinogenesis 1999;20(11):2125-2129.

41. Olshan AF, Weissler MC, Watson MA, Bell DA. GSTM1, GSTT1, GSTP1, CYP1A1, and NAT1 polymorphisms, tobacco use, and the risk of head and neck cancer. Cancer Epidemiol Biomarkers Prev 2000;9(2):185-91.

42. Geisler SA, Olshan AF. GSTM1, GSTT1, and the Risk of Squamous Cell Carcinoma of the Head and Neck: A Mini-HuGE Review. Am. J. Epidemiol. 2001;154(2):95-105.

43. Stefania B, Gabriella C, Giuseppe La T, Dario A, Mariangela V, Caterina C, et al. A case–control study investigating the role of sulfotransferase 1A1 polymorphism in head and neck cancer. Journal of Cancer Research and Clinical Oncology 2006;132(7):466-472.

44. Peters ES, McClean MD, Liu M, Eisen EA, Mueller N, Kelsey KT. The ADH1C Polymorphism Modifies the Risk of Squamous Cell Carcinoma of the Head and Neck Associated with Alcohol and Tobacco Use. Cancer Epidemiol Biomarkers Prev 2005;14(2):476-482.

45. Zhang Z, Shi Q, Liu Z, Sturgis EM, Spitz MR, Wei Q. Polymorphisms of Methionine Synthase and Methionine Synthase Reductase and Risk of Squamous Cell Carcinoma of the Head and Neck: a Case-Control Analysis. Cancer Epidemiol Biomarkers Prev 2005;14(5):1188-1193.

46. Neumann AS LH, Shen H, Liu Z, Shi Q, Sturgis EM, Shete S, Spitz MR, El-Naggar A, Hong WK, Wei Q. Methylenetetrahydrofolate reductase polymorphisms and risk of squamous cell carcinoma of the head and neck: A case-control analysis. International Journal of Cancer 2005;115(1):131-136.

47. Zhang Z, Shi Q, Sturgis EM, Spitz MR, Hong WK, Wei Q. Thymidylate Synthase 5'- and 3'-Untranslated Region Polymorphisms Associated with Risk and Progression of Squamous Cell Carcinoma of the Head and Neck. Clin Cancer Res 2004;10(23):7903-7910.

48. Liu Z, Calderon JI, Zhang Z, Sturgis EM, Spitz MR, Wei Q. Polymorphisms of vitamin D receptor gene protect against the risk of head and neck cancer. Pharmacogenet Genomics 2005;15(3):159-65.

49. Vairaktaris E, Yannopoulos A, Vassiliou S, Serefoglou Z, Vylliotis A, Nkenke E, et al. Strong association of interleukin-4 (-590 C/T) polymorphism with increased risk for oral squamous cell carcinoma in Europeans. Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology;In Press, Corrected Proof.

50. Quintela-Fandino M, Hitt R, Medina PP, Gamarra S, Manso L, Cortes-Funes H, et al. DNA-repair gene polymorphisms predict favorable clinical outcome among patients with advanced squamous cell carcinoma of the head and neck treated with cisplatin-based induction chemotherapy. J Clin Oncol 2006;24(26):4333-9.

51. Isla D, Sarries C, Rosell R, Alonso G, Domine M, Taron M, et al. Single nucleotide polymorphisms and outcome in docetaxel-cisplatin-treated advanced non-small-cell lung cancer. Ann Oncol 2004;15(8):1194-1203.

52. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. Proc Natl Acad Sci U S A 1971;68(4):820-3.

53. Ting NS, Lee WH. The DNA double-strand break response pathway: becoming more BRCAish than ever. DNA Repair (Amst) 2004;3(8-9):935-44.

54. AppliedBiosystems. GeneAssist Pathway Analysis: ATM Pathway. Online; 2007.

55. Savitsky K, Bar-Shira A, Gilad S, Rotman G, Ziv Y, Vanagaite L, et al. A single ataxia telangiectasia gene with a product similar to PI-3 kinase. Science 1995;268(5218):1749-53.

56. Ahmed M, Rahman N. ATM and breast cancer susceptibility. Oncogene 2006;25(43):5906-11.

57. Meyn MS. Ataxia-telangiectasia, cancer and the pathobiology of the ATM gene. Clinical Genetics 1999;55(5):289-304.

58. Swift M, Morrell D, Cromartie E, Chamberlin AR, Skolnick MH, Bishop DT. The incidence and gene frequency of ataxia-telangiectasia in the United States. Am J Hum Genet 1986;39(5):573-83.

59. Swift M, Reitnauer PJ, Morrell D, Chase CL. Breast and other cancers in families with ataxia-telangiectasia. N Engl J Med 1987;316(21):1289-94.

60. Thompson D, Duedal S, Kirner J, McGuffog L, Last J, Reiman A, et al. Cancer risks and mortality in heterozygous ATM mutation carriers. J Natl Cancer Inst 2005;97(11):813-22.

61. Koren M, Kimmel G, Ben-Asher E, Gal I, Papa MZ, Beckmann JS, et al. ATM haplotypes and breast cancer risk in Jewish high-risk women. Br J Cancer 2006;94(10):1537-43.

62. Stankovic T, Kidd AM, Sutcliffe A, McGuire GM, Robinson P, Weber P, et al. ATM mutations and phenotypes in ataxia-telangiectasia families in the British Isles: expression of mutant ATM and the risk of leukemia, lymphoma, and breast cancer. Am J Hum Genet 1998;62(2):334-45.

63. Broeks A, Urbanus JH, Floore AN, Dahler EC, Klijn JG, Rutgers EJ, et al. ATM-heterozygous germline mutations contribute to breast cancer-susceptibility. Am J Hum Genet 2000;66(2):494-500.

64. Teraoka SN, Malone KE, Doody DR, Suter NM, Ostrander EA, Daling JR, et al. Increased frequency of ATM mutations in breast carcinoma patients with early onset disease and positive family history. Cancer 2001;92(3):479-87.

65. Dork T, Bendix R, Bremer M, Rades D, Klopper K, Nicke M, et al. Spectrum of ATM gene mutations in a hospital-based series of unselected breast cancer patients. Cancer Res 2001;61(20):7608-15.

66. Kim JH, Kim H, Lee KY, Choe K-H, Ryu J-S, Yoon HI, et al. Genetic polymorphisms of ataxia telangiectasia mutated affect lung cancer risk. Hum. Mol. Genet. 2006;15(7):1181-1186.

67. Landi S, Gemignani F, Canzian F, Gaborieau V, Barale R, Landi D, et al. DNA Repair and Cell Cycle Control Genes and the Risk of Young-Onset Lung Cancer. Cancer Res 2006;66(22):11062-11069.

68. Lazar AD, Winter MR, Nogueira CP, Larson PS, Finnemore EM, Dolan RW, et al. Loss of heterozygosity at 11q23 in squamous cell carcinoma of the head and neck is associated with recurrent disease. Clin Cancer Res 1998;4(11):2787-2793.

69. Lese CM, Rossie KM, Appel BN, Reddy JK, Johnson JT, Myers EN, et al. Visualization of INT2 and HST1 amplification in oral squamous cell carcinomas. Genes Chromosomes Cancer 1995;12(4):288-95.

70. Uzawa K, Suzuki H, Komiya A, Nakanishi H, Ogawara K, Tanzawa H, et al. Evidence for two distinct tumor-suppressor gene loci on the long arm of chromosome 11 in human oral cancer. Int J Cancer 1996;67(4):510-4.

71. Tan D, Wiseman S, Zhou Y, Li Q, Ward P, Slocum HK, et al. Definition of a region of loss of heterozygosity at chromosome 11q23.3-25 in head and neck squamous cell carcinoma using laser capture microdissection technique. Diagn Mol Pathol 2004;13(1):33-40.

72. Mondal G, Tripathi A, Bhattacharya N, Sikdar N, Roy A, Sengupta A, et al. Allelic imbalance at chromosome 11 in head and neck squamous cell carcinoma in an Indian patient population. J Clin Pathol 2003;56(7):512-518.

73. Nunn J, Scholes AGM, Liloglou T, Nagini S, Jones AS, Vaughan ED, et al. Fractional allele loss indicates distinct genetic populations in the development of squamous cell carcinoma of the head and neck (SCCHN). Carcinogenesis 1999;20(12):2219-2228.

74. Ai L, Vo QN, Zuo C, Li L, Ling W, Suen JY, et al. Ataxia-Telangiectasia-Mutated (ATM) gene in head and neck squamous cell carcinoma: promoter hypermethylation with clinical correlation in 100 cases. Cancer Epidemiol Biomarkers Prev 2004;13(1):150-156.

75. The International HapMap C. A haplotype map of the human genome. Nature 2005;437(7063):1299-1320.

76. de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. Nat Genet 2005;37(11):1217-23.

77. (SCMGC) Seattle Comparative Mouse Genomics Center. ATM gene model. University of Washington Online; 2007.