universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Association of a polymorphic tandem repeats minisatellite of human telomerase gene with risk of prostate cancer and lung cancer

Verfasserin

Julia-Stephanie Gertraud Zerelles

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag.rer.nat.)

Wien, 2012

Studienkennzahl lt. Studienblatt:	A 442
Studienrichtung lt. Studienblatt:	Diplomstudium Anthropologie
Betreuerin / Betreuer:	Ao. Univ.-Prof. Mag. Dr. Andrea Gsur

Danksagung

Ich bedanke mich herzlich bei meiner Betreuerin Ao. Univ.-Prof. Mag. Dr. Andrea Gsur (Arbeitsgruppe Molekulare Epidemiologie, Institut für Krebsforschung, Medizinische Universität Wien), für die kompetente und freundliche Betreuung.

Weiters bedanke ich mich bei Philipp Hofer, Stefanie Brezina und Vanita Vanas für die tolle Atmosphäre und die geduldige Unterstützung während der Arbeit im Labor.

Außerdem bedanke ich mich bei Mag. Dr. Andreas Baierl für die statistische Auswertung der Daten.

Ich möchte mich ebenfalls bei meinen Freunden und Familie bedanken, die mich bei der Entstehung dieser Arbeit unterstützt und ermutigt haben.

Danke an meine Eltern, ohne die dieses Studium nicht möglich gewesen wäre.

Abstract

Telomerase is a complex holoenzyme able to elongate telomere ends that have been shortened during successive cycles of cell division. Expression of telomerase reverse transcriptase gene (TERT), which is fundamental for the maintenance of telomeres length, is usually absent in somatic cells, but present in the majority of cancer cells and immortalized cells. MNS16A, a polymorphic tandem repeats minisatellite is located downstream of exon 16 of the TERT gene, within a promoter region of a TERT antisense transcript. MNS16A genotypes were determined by PCR amplification followed by electrophoretical separation in 2.5% agarose gels. Multiple logistic regression analysis was performed in the R programming environment. The aim of my diploma thesis was to assess the role of MNS16A as a potential biomarker for the risk of developing prostate cancer for the first time as well as for lung cancer. The prostate cancer study population comprised 1787 participants, composed of 1137 newly diagnosed prostate cancer cases and 650 controls with benign prostate hyperplasia (BPH). A strength of our prostate cancer study was, that controls underwent biopsy of the prostate and were therefore attested to be free of prostate cancer. No significant associations were observed in the overall study population, however, in the stratified analysis of the prostate cancer cohort, the MNS16A variable number of tandem repeats (VNTR)-274 was associated with a decreased risk of prostate cancer in men aged 70 or older (p= 0.023).

Additionally, MNS16A was investigated as a biomarker for the risk of lung cancer based on a study population of 3167 participants. The lung cancer patients (n= 619) were classified according to their histological subtype in a small cell lung cancer (SCLC=125) and a non-small cell lung cancer (NSCLC=476) subgroup. No significant associations were observed in the overall lung cancer study population. In the stratified analysis, genotype 243/302 was associated with a reduced risk of lung cancer in subjects aged 70 or older (p= 0.045). However, in subjects aged 55 or younger the genotypes 243/302 and 274/302 and the single MNS16A VNTR-243 were associated with an elevated risk of lung cancer (p= 0.007-0.038). In the subgroup of NSCLC patients, genotypes 243/302 and 302/333 and single VNTR-243 were associated with a reduced NSCLC risk in subjects aged 70 and older (p= 0.012-0.046), which is in accordance with the results of the very first MNS16A study, published by Wang et al. in 2003. In subjects aged 55 and younger, the genotypes 243/302 and 274/302 and the single VNTR-243 were associated with an increased risk of NSCLC (p= 0.006-0.03).

Five VNTRs of MNS16A are known so far and termed according to their PCR fragment size. In the course of my diploma thesis a novel MNS16A allele was identified, namelyVNTR-212, the shortest VNTR that has been identified so far.

Although further investigations in larger cohorts are required to establish MNS16A as a possible biomarker for the risk assessment of prostate cancer and lung cancer, this diploma thesis has given several significant associations of MNS16A variants and the risk of prostate cancer and lung cancer.

Zusammenfassung

Telomerase ist ein komplexes Holoenzym, das in der Lage ist, Telomere, welche sich durch Zellteilung stetig verkürzen, zu verlängern. Das Gen Telomerase reverse Transkriptase (TERT), wird normalerweise in somatischen Zellen nicht exprimiert, wohl aber in der Mehrheit von Krebs- bzw. immortalen Zellen.

MNS16A, ein polymorpher „tandem repeats“ Minisatellit, ist „downstream“ von Exon 16 lokalisiert und liegt innerhalb einer Promoter Region eines TERT Antisense- Transkripts. Die MNS16A Genotypen wurden durch PCR Amplifikation und anschließender Separation in 2,5%igen Agarose Gelen bestimmt. Multiple logistische Regressionsanalysen wurden im Programm „R“ durchgeführt.

Ziel dieser Arbeit war es, zu untersuchen, ob MNS16A ein geeigneter Biomarker für Prostatakarzinomrisiko als auch Lungenkarzinomrisiko sein könnte. Die Studienpopulation der Prostatakarzinomstudie bestand insgesamt aus 1787 Teilnehmern, davon 1137 Patienten mit neu diagnostiziertem Prostatakarzinom und 650 Kontrollpatienten mit benigner Prostatahyperplasie (BPH). In der gesamten Studienpopulation konnten keine statistisch signifikanten Assoziationen beobachtet werden. Jedoch wurde in einer stratifizierten Analyse das MNS16A „variable number of tandem repeats“ (VNTR)-274 mit einem reduzierten Prostatakrebsrisiko bei Männern in einer Altersklasse von 70 Jahren oder älter signifikant assoziiert (p=0,023).

Die Lungenkarzinom Studienpopulation bestand aus 3167 Teilnehmern. Patienten mit Lungenkarzinom (n=619) bestanden ihrem histologischen Subtyp entsprechend, aus 125 Patienten mit kleinzelligem Lungenkarzinom (SCLC) und 476 Patienten mit nicht-kleinzelligem Lungenkarzinom (NSCLC). In der gesamten Studienpopulation konnten keine statistisch signifikanten Assoziationen beobachtet werden. In einer stratifizierten Analyse wurde bei Teilnehmern von 70 Jahren oder älter ein signifikanter Zusammenhang zwischen dem Genotyp 243/302 und einem reduzierten Lungenkrebsrisiko beobachtet (p= 0,045). Bei Männern von 55 Jahren oder jünger konnten die Genotypen 243/302 und 274/302 und das MNS16A VNTR-243 mit einem erhöhten Lungenkrebsrisiko assoziiert werden (p= 0,007-0,038). In der NSCLC Untergruppe der ≥70 Jährigen wurden die Genotypen 243/302 und 302/333 und VNTR-243 mit einem reduziertem NSCLC Risiko assoziiert (p= 0,012-0,046), was mit den Ergebnissen der ersten MNS16A Studie von Wang et al. im Jahre 2003 übereinstimmt. In der Untergruppe der ≤55 Jährigen konnten die Genotypen 243/302 und 274/302 und VNTR-243 mit einem erhöhten NSCLC Risiko assoziiert werden (p= 0,006-0,03).

Bisher waren fünf VNTRs, die nach der Größe ihrer PCR Fragmente benannt wurden, bekannt. Im Rahmen meiner Diplomarbeit wurde eine weitere MNS16A Variante, VNTR-212, die bislang kürzeste Variante entdeckt.
Obwohl noch weitere Untersuchungen in größeren Studienpopulationen nötig sind, um MNS16A als möglichen Biomarker für die Risikobeurteilung von Prostatakrebs und Lungenkrebs zu etablieren, konnten in dieser Diplomarbeit signifikante Assoziationen zwischen MNS16A Varianten und dem Risiko von Prostata- und Lungenkrebs beobachtet werden.

Table of contents

Abbrevations

ALT	Alternative lengthening of telomeres
APC	Adenomatous polyposis coli
BPH	Benign prostate hyperplasia
BRCA1/2	Breast cancer gene 1/2
CNV	Copy number variation
CLPTM1L	Cleft lip and palate transmembrane protein 1 gene
EtBr	Ethidium Bromid
GWAS	Genome wide association study
GSTM1	Glutathione S transferase M1
HWE	Hardy-Weinberg equilibrium
KLK3	Kallikrein related peptidase 3
LD	Linkage disequilibrium
NSCLC	Non-small cell lung cancer
OR	Odds ratio
PC	Prostate cancer
PCR	Polymerase Chain Reaction
PIN	Prostatic intraepithelial neoplasma
PSA	Prostate specific antigen
RB1	Retinoblastoma 1 gene
SCLC	Small cell lung cancer
SD	Standard deviation
SHBG	Sex hormone-binding globulin
SNP	Single nucleotide polymorphism
TERC	Telomere template-containing RNA component
TERT	Telomerase reverse transcriptase
TP53	Tumor protein p53 gene
tSNP	Tagging SNP
VNTR	Variable number of tandem repeats

List of tables and figures

Tables

Figures

1. Introduction

1.1. Epidemiology

1.1.1. Epidemiology of prostate cancer

1.1.1.1. Epidemiology of prostate cancer worldwide

Prostate cancer is the second most frequently diagnosed cancer in male patients worldwide, comprising 13.6% of the total cancer cases and 6.1% of the total cancer deaths in men. Incidence rates for prostate cancer can vary up to 25fold worldwide, with Australia/New Zealand having the highest incidence rate for prostate cancer, followed by western and northern Europe and northern America. Particularly high incidence rates occur among African Americans, the lowest incidence rates worldwide are recorded in south-east Asia. In 2008, around 900.000 men were diagnosed with prostate cancer worldwide ,(Globocan/a, 2008) 382.300 of them were recorded in Europe (Ferlay et al., 2010). The high prostate cancer incidence in Western countries is associated with the implementation of large-scale prostate cancer screening programs (Sakr et al., 1996). By this measure, mortality due to prostate cancer was significantly reduced (Hugosson et al., 2010; Jemal et al., 2005; Schroder et al., 2009), since a greater proportion of men were diagnosed with prostate cancer in an early stage of their disease, referred to as “stage migration” (Cooperberg et al., 2003). The amount of patients presented with distant metastases at time of diagnosis was reduced from 16 to 4% with screening, as reviewed in (Tosoian and Loeb, 2010). Anyway, screening for prostate cancer leads to the key aspect of over-diagnosis, followed by over-treatment of clinically insignificant cancers, which is assumed to be as much as 50% (Draisma et al., 2003). Men with an indolent prostate cancer would not require radical treatment, like surgery, and may suffer from side effects, like erectile dysfunction and incontinence, wherefore the cost-benefit ratio of surgery versus active surveillance must be considered (Marberger et al., 2012).
The mortality ratio for prostate cancer varies about 10fold worldwide, the highest mortality occurs in primarily black populations, the lowest mortality rates occur in Asia. In 2008, roughly 260.000 deaths due to prostate cancer were recorded worldwide, around 70.000 of them occurred in Europe (Globocan/a, 2008).

1.1.1.2. Epidemiology of prostate cancer in Austria

In Austria, prostate cancer is the leading cancer site in males, accounting for 26.4% of the total new cancer cases. It is the third leading cause of death among Austrian men with 11.3% of all cancer deaths following lung and colorectal cancer deaths.
In 2009 almost 4.900 patients have been diagnosed of prostate cancer, which is in strong

contrast to 1989, when 2.317 cases of prostate cancer were reported only . The increased incidence rate is associated with the widespread prostate cancer screening programs, conducted particularly in Western regions of Austria. Although a gradual decline in prostate cancer incidence was observed in Austria since 2003 this trend might largely be attributed to the earlier identification of prostate cancer, including non-symptomatic indolent prostate cancer.

Prostate cancer related mortality lessened from 21.9/100.000 men in 1989 to 15/100.000 men in 2009, due to the introduction of prostate cancer detection programs and improved curative treatment (Hugosson et al., 2010; Jemal et al., 2005; Schroder et al., 2009). The cumulative risk to develop prostate cancer before the age of 75 was in 2009 reduced to 9%, the cumulative risk to die from prostate cancer dropped to 0.7% (Statistik-Austria/a, 2009).

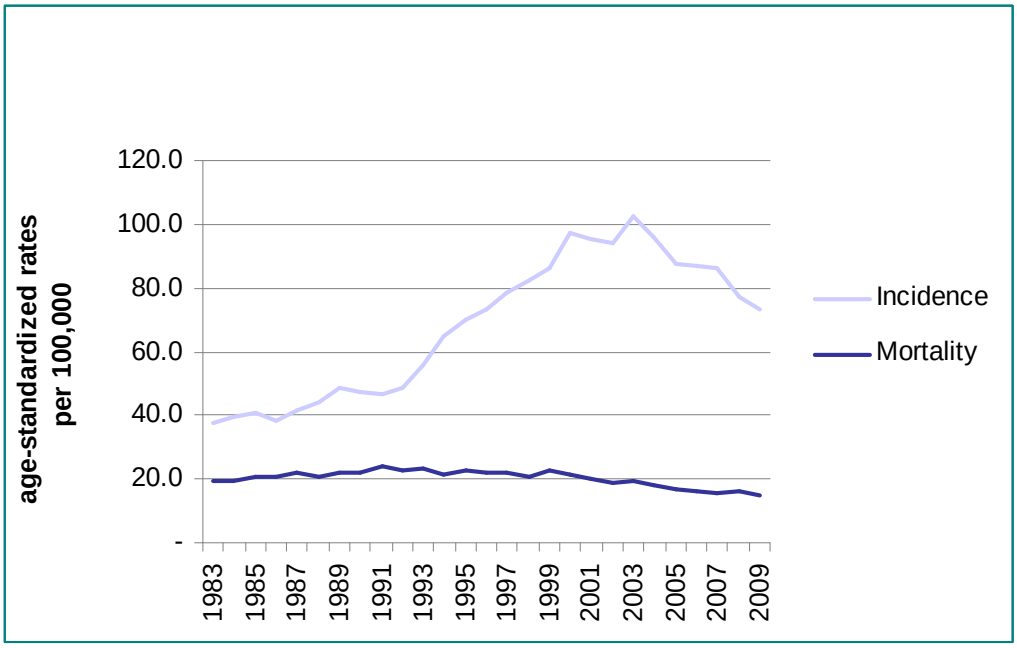

Figure 1: Statistik Austria a.; Incidence and mortality of prostate cancer in Austria

1.1.1.3. Risk factors for prostate cancer

One of the most significant risk factors for the development of prostate cancer is advanced age. Only 0.6% of all prostate cancer cases arise in men younger than 45 years, whereas it is between 62% and 85% after the age of 65 (Fournier et al., 2004). According to the US Surveillance, Epidemiology and End Results Program (SEER), the incidence rate of prostate cancer, recorded between 2000 and 2008, rises from 9.2/100.000 for men aged 40-44 to 984.8/100.000 for men aged 70-74, as reviewed in (Leitzmann and Rohrmann, 2012). Since

androgens affect the development and growth of the prostate gland, testosterone is considered to be an important determinant for the occurrence of prostate cancer. Increased levels of circulating testosterone combined with low levels of sex hormone-binding globulin (SHBG) (both within an endogenous range) were associated with an elevated risk of prostate cancer (Gann et al., 1996).

Furthermore, ethnicity varies the risk of developing prostate cancer. Men of African descent have a 1.6fold higher incidence rate of prostate cancer compared to the Caucasian population (Leitzmann and Rohrmann, 2012). A positive family history of prostate cancer is another risk factor for its occurrence. According to the Swedish Family Cancer Database, 20.2% of men diagnosed with cancer before the age of 72 have a paternal or fraternal cancer history. The risk of developing prostate cancer rises with the amount of first degree relatives, diagnosed with prostate cancer (Figure 2) (Hemminki, 2011). Family-based linkage studies, twin studies and molecular epidemiology studies provided evidence that genetics play a key role in the development of prostate cancer. It is estimated, that 42% of prostate cancer risk is possibly explained by genetic determinants (Lichtenstein et al., 2000).

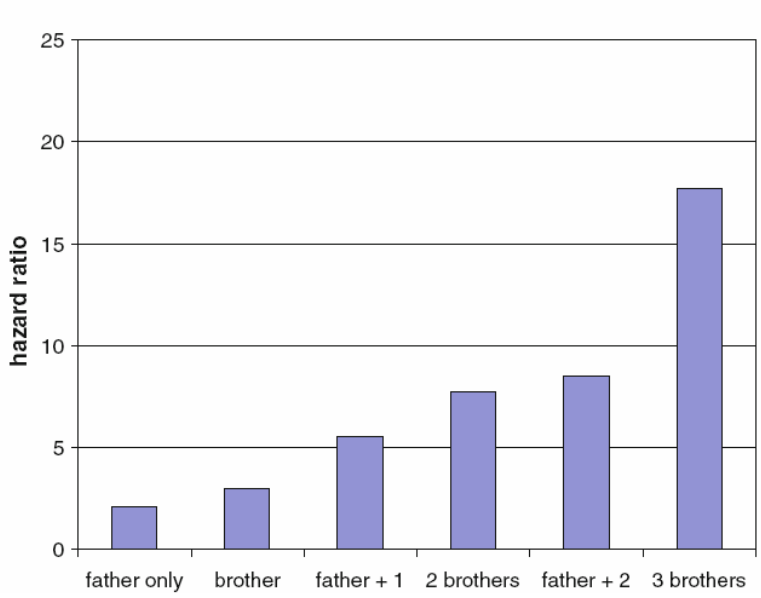

Figure 2: Hemminki 2011; prostate cancer risk and affected relatives

1.1.1.4. Screening for prostate cancer

Prostate cancer is indicated by increased levels of prostate specific antigen (PSA) in blood plasma. The glycoprotein, that is encoded by the KLK3 (kallikrein related peptidase 3) gene, is present at a low concentration in the serum of healthy men, but elevated in men with prostate disorders like prostate cancer (Kuriyama et al., 1980; Papsidero et al., 1980). PSA

level as a biomarker, was introduced throughout the 1990s and was considered to be a tool for early diagnosis and monitoring of prostate cancer (Catalona et al., 1991). Unfortunately, the PSA test has a high false-positive rate (Etzioni, 2008). PSA screening may lead to over-diagnosis and over-treatment of indolent tumors, possibly followed by preventable side effects (Marberger et al., 2012). However, PSA is the only available marker so far. Therefore, there is a great demand for a biomarker that enables to distinguish the aggressive from the indolent form of prostate cancer at the time of diagnosis. While men with aggressive tumors will benefit more from radical treatment, men with the indolent form profit more from „watchful waiting“ and can be spared from side effects of surgery.

1.1.2. Epidemiology of lung cancer

1.1.2.1. Epidemiology of lung cancer worldwide

Before the 20th century lung cancer was a rare disease, but since then tobacco consumption rose rapidly and lung cancer incidences increased subsequently (Brennan et al., 2011). In 2008 lung cancer was the third most common malignancy worldwide, accounting for 12.5% of all new cancer cases, and the most common cause of cancer related death with a share of 18.2%. The highest incidence rates occur in Central- Eastern- and Southern Europe and Northern America, while Middle- Western- and Eastern Africa have the lowest incidence rates. There is also a discrepancy between lung cancer incidence and mortality in men and women, in ways that female cases are generally less frequent. In men lung cancer is the most common diagnosed kind of cancer, while in women it is placed fourth, following breast, cervix uteri and colorectal cancer. Regarding mortality, however, it is placed second in females and first in males. In Northern America, females have the highest incidence and mortality rates, for males it is Central and Eastern Europe. The ratio of mortality to incidence is 0.86, illustrating the high death rate, which is similar in every country or region (Globocan/b, 2008).

1.1.2.2. Epidemiology of lung cancer in Austria

In Austria, 4239 newly diagnosed lung cancer cases, accounting for 11% of all newly diagnosed neoplasms, were observed in 2009. Of 29 per 100.000 habitants 24 patients will die from lung cancer. In men, lung cancer is the second most frequently diagnosed cancer, following prostate cancer, but the leading cause of cancer related death. In Austrian women,

lung cancer is placed third regarding the incidence, only surpassed by breast and colorectal cancer, and placed second regarding the mortality.
Concerning both, incidence and mortality, men account for two thirds of all cases, largely caused by the higher rate of male smokers. However, age standardized incidence and mortality rates in women increased by about 20% (23% and 20%) during the last ten years. Although the mortality in men slightly decreased during the last years, the overall mortality of lung cancer remained more or less unchanged (Figure 3). The life time risk for developing lung cancer, before the age of 75, accounts to 4.8% for men and 2.2% for women (Statistik-Austria/b, 2009).

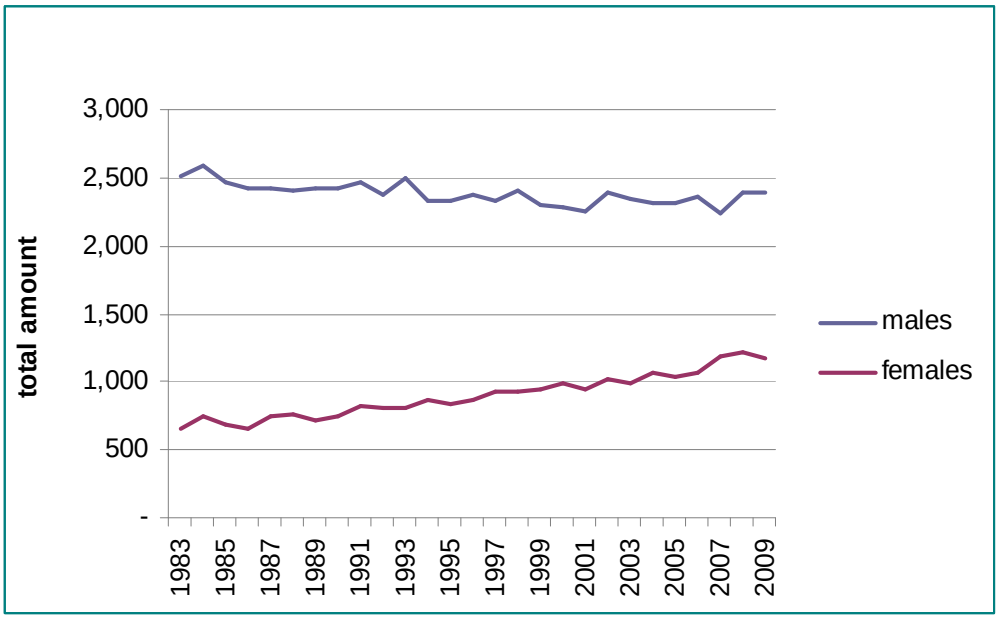

Figure 3: Statistik Austria b.; Lung cancer mortality in Austria

1.1.2.3. Risk factors

The carcinogenic impact of tobacco was first demonstrated in the 1950s, followed by epidemiologic studies, showing, that tobacco is the main exogenic cause of lung cancer (Brennan et al., 2011). Between 80 and 90% of lung cancer cases are attributed to smoking (Hammond and Seidman, 1980), however, only 10-15% of smokers develop lung cancer (Mattson et al., 1987), indicating inter-individual differences in carcinogen susceptibility (Schwartz et al., 2007). Except smoking and second hand smoking, exposure of environmental carcinogens, like asbestos or radon (Goeckenjan et al., 2011) and differences in susceptibility to lung cancer may contribute to the remaining lung cancer risk (Schwartz et al., 2007).

A positive family history of lung cancer elevates the personal risk of developing the disease, suggesting a genetic component in lung carcinogenesis. With an affected first-degree relative the risk of lung cancer increases up to 9fold (Dong and Hemminki, 2001; Risch, 2001), determined by genetic susceptibility and/or familial clustering of environmental risk factors (Lorenzo Bermejo and Hemminki, 2005).
Further, ethnicity varies the incidence rate of lung cancer, with above-average rates in Afro-Americans and among women in certain regions of China. Again, the increased prevalence of lung cancer among certain ethnicities may be explained by genetic variation in susceptibility to lung cancer and/or by a certain behavior in dealing with carcinogens, like carcinogen generating cooking in certain regions of China (Brennan et al., 2011; DeMarini et al., 2001).

1.1.2.4. Screening for lung cancer

Lung cancer is one of most aggressive kinds of cancer with a strong tendency to metastasize (Petersen, 2011). In many patients, the cancer is already in an advanced state by the time they show characteristic symptoms (Wood et al., 2012). At time of diagnosis less than 0.1% of lung cancer cases are staged carcinoma *in situ* while the majority is already disseminated (27%). Lung cancer is usually diagnosed by chest X-ray and confirmed by lung biopsy. The 5-year survival rate for males amounts 15%, for females 18%, reflecting only limited advances in screening and treatment (Petersen, 2011). Therefore, it is an essential issue to find a minimal invasive marker for early detection of lung cancer, when treatment is more likely to be successful (Aberle et al., 2011).

1.2. Anatomy and Pathology of the prostate

The prostate is an exocrine glandular organ located in the small pelvis, just under the bladder and in front of the rectum. The prostate secretes a fluid that nourishes and protects sperm and is squeezed into the urethra, by smooth muscle cell contraction, during ejaculation. Four anatomic regions of the prostate gland are described: 70% of the prostate tissue is the peripheral zone. It surrounds the distal urethra and prostate tumors usually originate here. The central zone constitutes about 25% of the gland and surrounds the ejaculation ducts. The remaining 5% of the prostate gland are divided in transition zone, which grows lifelong and is therefore typically the origin of benign prostatic enlargements, and the anterior fibromuscular stroma, which forms the anterior surface of prostate as a non-glandular apron (Amin et al., 2010). The prostate comprises an epithelial and a stromal compartment, which are separated by the basement membrane, a barrier consisting of collagen fibers (Bonkhoff et al., 1991). The epithelial cell compartment is built up by two kinds of cell layers. The luminal epithelial

cell layer produces prostatic secretion and is completely separated from the stroma by the second kind of epithelial cells, the basal epithelial cells (Brawer et al., 1985). The stromal cell compartment consists of smooth muscle cells and myofibroblasts, which are interspersed with inflammatory cells, cells of the nerve system, blood vessels and lymphatics (Bartsch et al., 1979).

Between 70-80% of men in their 7^{th}-8^{th} life decade, develop a hyperplastic prostate epithelium, which is called benign prostate hyperplasia. It is well accepted that BPH is neither a premalignant lesion nor a precursor carcinoma (Gsur et al., 2000). In BPH the luminal epithelial cells accumulate, but stay morphologic and cytologic inconspicuously. About 10% of men with BPH require treatment (Bushman, 2009). In contrast to BPH, prostatic intraepithelial neoplasms (PIN) is considered to be a precursor to invasive prostate carcinomas. In PIN the luminal cells feature enlarged cell nuclei with striking nucleoli, further, integrity of basal epithelial cells decrease (Ayala and Ro, 2007). Usually, prostate carcinomas are classified as multifocal adenocarcinomas. In 80% of all cases, prostate carcinomas arise in the peripheral zone of the prostate gland. Characteristics of prostate cancerous tissues may be ramification of the glandular morphology, gland fusion, inability to form a lumen, attrition of the basal cell layer and nucleus enlargement attended by aneuploidy. If prostate cancer is in an advanced stage, metastases in bone and lymph nodes occur frequently (Knudsen and Vasioukhin, 2010). Tumors are classified based on the Gleason Score, which was introduced by the pathologist D. Gleason in 1974. This Grading relies on the most prevalent architectural pattern of the tumor (Gleason and Mellinger, 1974).

1.3. Anatomy and Pathology of the lung

The lung is divided in the right lung (pulmo dexter) and the left lung (pulmo sinister), which are further segmented in pulmonary lobes (lobi pulmonis). The left lung is more limited in space and consists of two lobes, whereas the right lung consists of three lobes. Lobes in turn are composed of bronchopulmonary segments, ten for the right side and nine for the left side. Air enters the lung via the bronchial system, finally the gaseous exchange happens in the alveoles, which are located at the end of the bronchial tree. The bronchial tubes are built up by ciliated epithelial cells and a subjacent layer of smooth muscles, including nestled elastic fibers and glands (Fanghänel J., 2003).

Lung cancer is classified in two main types, namely small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). NSCLC can be further divided in Adenocarcinoma, Squamous cell carcinoma and Large cell carcinoma. SCLC cases account for about 20%,

Adenocarcinomas for 40%, Squamous cell carcinomas for 25-30% and Large cell carcinomas for 10-15% of all lung cancer cases. Metastases of lung cancer are primarily sited at lymph nodes, liver, bones, suprarenal gland and brain. Bronchial carcinomas are staged according to the TNM-classification system, which is based on the size of the primary tumor (T in situ- 4), location of afflicted lymph nodes (N0-N3) and the existence of distant metastases (M0-M1). Dependent on the TNM-categories, lung cancer staging is set (stage 1-stage 4), with different prognosis of therapeutic success (Pirker R., 1996).

1.4. Molecular Epidemiology

The term “molecular epidemiology” was introduced by Perera and Weinstein as a multidisciplinary approach in cancer research in 1982 (Perera and Weinstein, 1982). The discipline of molecular cancer epidemiology integrates advanced molecular biology into traditional epidemiology, in order to unveil possible biochemical and molecular causes for cancer. Its aim is to elucidate the molecular events occurring between the causative exposure of a carcinogen and the cancer outcome, considering an individual predisposition (Vineis and Perera, 2007). A range of Biomarkers were integrated in molecular epidemiologic studies, classified as (1) markers of internal dose, (2) markers of biologically effective dose, (3) markers of early response/effect and (4) markers of susceptibility. The biomarker of the internal dose was used to measure the concentration of a suspected carcinogen to assess the associated personal risk. The carcinogen or its metabolites can be detected in cells, tissues or body fluids, even at a very low concentration, but is not predictive for any genetic alteration. Therefore assays have been developed to detect the biologically effective dose of an agent, which is interacting with critical cellular macromolecules, like DNA. The carcinogen-DNA-adduct extent, respectively the DNA damage, mirrors the amount of incorporated carcinogens. It must be considered, that there is an inter-individual disparity in absorption and metabolism of carcinogens, as well as in repair events, regarding DNA-adducts. Beside the external pathway, several endogenous determinants are known, such as inflammatory processes, contributing to non-specific DNA damage. Markers of early biologic effects resulting from exposure reflect the subsequent events in carcinogenesis. They are the third kind of biomarkers and determine chromosomal and genetic alterations, like aberrations, deletions and point mutations. Some rearrangements are characteristic for an incipient, specific cancer, and may therefore operate as an early reporter. Biomarkers of susceptibility, the fourth category of biomarkers, address inherited or acquired genomic patterns, modulating the personal risk of developing cancer (Perera and Weinstein, 2000).

Biomarkers of epigenetic modification were integrated in the field of molecular epidemiology. The functional regulation of gene expression by epigenetic events, such as DNA methylation, plays a major role in carcinogenesis, induced from exposure (Kim, 2004). Through advanced techniques, new approaches in the establishment of biomarkers are available. They are fielded for identification of intermediate changes in the molecular profile caused by exposure, additionally they are used as an early marker for cancer onset. In proteomics, particular proteins or proteomic patterns within the whole protein output are identified and utilize as a tool to distinguish between normal and cancerous cells (Kikuchi and Carbone, 2007). In the field of metabolomics, the total of low-molecular weight metabolites present in a cell is determined, to design a personal metabolomic profile in order to evaluate the individual risk of cancer (Lindon et al., 2004).

1.4.1. Genetic susceptibility

Genetic influence in cancer development is an issue since linkage studies based on high-risk families, identified rare cancer susceptibility loci for common cancers, such as adenomatous polyposis coli (APC) for colorectal cancer and breast cancer gene 1/2 (BRCA1/2) in familial breast cancer (Bodmer et al., 1987; Hall et al., 1990; Wooster et al., 1994). Mutations in high-risk genes, usually tumor suppressor genes, are of low frequency (<1/1000), but of high penetrance (Fletcher and Houlston, 2010; Hindorff et al., 2011). Mutations in high penetrance genes account for just a small proportion of inherited susceptibility. The remaining familial risk is very likely correlated with combinations of common, low to moderate penetrance susceptibility alleles (Pasche and Yi, 2010).

The first approach to identify common variants of low-penetrance cancer susceptibility loci, was through candidate gene association studies (Gianfagna et al., 2012). Candidate genes are selected based on a priori hypothesis of their involvement in carcinogenic-related events, such as DNA repair or apoptosis, referred to as “functional candidate gene approach”.

Additionally, genes that are often deleted or amplified in cancerous cells are investigated by the “positional candidate gene approach” (Imreh et al., 2003; Oldenburg et al., 2007). Initial candidate gene studies investigated on one polymorphism within one candidate gene, based on small case-control cohorts. Subsequent studies focused on more than one polymorphism in one single candidate gene, considering that 10-100 functional polymorphisms, especially single nucleotide plymorphisms (SNPs) and copy number variations (CNVs), are located within a 30Kb gene. Additionally, polymorphisms sited thousands of base pairs away from the target gene or within its introns interfere with gene expression and/or modify the splicing

process of the transcribed mRNA (Gianfagna et al., 2012). The identification of all polymorphic sites within one gene enabled to identify linkage disequilibrium (LD) relations between certain SNPs, condensed in Haplotypes. SNPs within one Haplotype are mutually correlated and are therefore characterized by tagging SNPs (tSNP). The gene-wide variation is more exact determined by just the merest subset of tSNPs, which are exemplary for the whole Haplotype, than by independent SNPs (Fallin and Schork, 2000). In 2003, researchers from the HapMap project came up with a validated set of haplotypes resp. tSNPs encompassing the whole human genome, also considering different ethnicities, that enabled to extend association studies from single genes to the whole genome (HapMap, 2003). In the course of the HapMap project, about 5 Mio SNPs were identified and further clustered according to their LD state. Based on this new knowledge and advanced technologies, genome wide association studies (GWAS) followed. The genome wide approach enabled to genotype up to 1 Mio SNPs at once, about 10 SNPs per gene, resulting in a huge dataset. The big advantage of this new approach is the ability to detect functional polymorphisms, sited in unsuspected genomic regions of the entire genome, wherefore association studies are no longer dependent on a priori hypothesis of genes, potentially involved in carcinogenic events (Gianfagna et al., 2012).

GWAS are conducted concerning cancer risk, clinical outcome and treatment response and require some essentials, such as a sufficient amount of cases and controls, approved statistical significant levels and stratification for confounders (Miyagawa et al., 2008; Ziegler et al., 2008). So far, about 100 cancer susceptibility polymorphisms were mapped by GWAS, with a low solitary effect, but due to their prevalence, with a strong common impact on cancer susceptibility (Figure 4) (Chanock and Hunter, 2008). Next generation sequencing, which supports genotyping enormous quantities of genomic DNA up to entire genomes, is the new, cost-effective approach in cancer research (Shendure and Ji, 2008). One such implementation is the sequencing of the complete human genome, as published recently (Levy et al., 2007; Wang et al., 2008a; Wheeler et al., 2008). Furthermore, DNA of normal and cancerous cells of one individual can be matched, in order to gain insights on the total genetic changes in carcinogenesis (Ley et al., 2008).

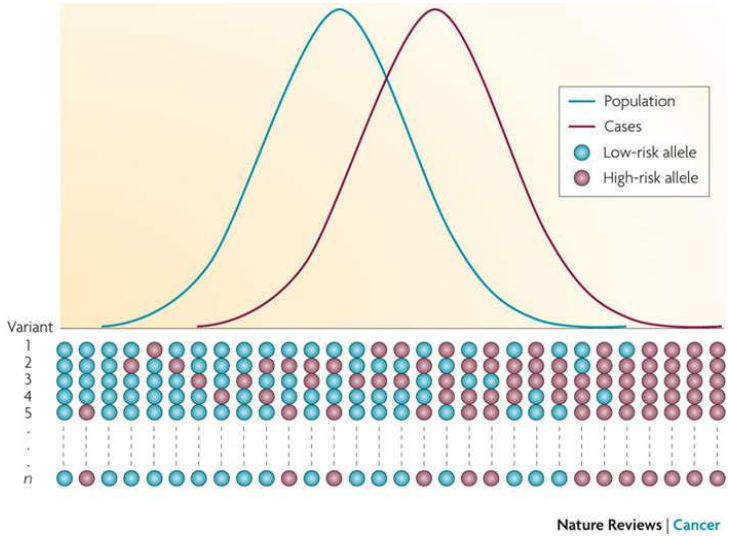

Figure 4: Fletcher et al. 2010; Dosage effect of susceptibility alleles

1.4.1.1.Genetic Susceptibility of prostate cancer

Familial clustering is among the most important risk factors for prostate cancer (Simard et al., 2002), however, unlike other malignancies, no classical high-penetrance genes have been identified yet (Hindorff et al., 2011). The closest candidate is BRCA2, which confers a 20fold increase in prostate cancer risk, compared to the general population (Edwards et al., 2003). Recent GWAS identified at least 40 common, low-risk (OR<2) variants on several chromosomes associated with prostate cancer risk, e.g. on 2p15, 8q24, 17q, 19q13, and Xp11 (Nakagawa et al., 2012). This is the highest number among common cancers, suggesting that these common alleles generate the genetic influence in prostate carcinogenesis (Hindorff et al., 2011). The high number of functional polymorphisms may be explained by the weaker selection pressure, since prostate cancer occurs mostly in elder men. The cumulative impact of these identified loci explain at the minimum 20% of the familial risk of prostate cancer (Varghese and Easton, 2010). It is assumed, that the identified loci take part in the initiation of cancer, rather than in the progression or aggressiveness, since associations between genetic variants and the severity of prostate cancer could not be determined (Witte, 2009).

1.4.1.2.Genetic Susceptibility of lung cancer

Although environmental exposure exerts the primary influence on the etiology of lung cancer, several genetic predispositions, modifying the risk of developing lung cancer are known. Lung cancer related genes are classified in rare high-risk genes (effect size >10; 1% of

population), moderate-risk variants (effect size=2-5; <5% of population) and common low-risk variants (effect size =1.1-1.5; >5% of population) (Brennan et al., 2011). Family and linkage studies detected high-penetrance alleles of the tumor protein p53 gene (TP53), retinoblastoma 1 gene (RB1) and 6q23-25, which are associated with lung cancer risk (Fletcher et al., 2004; Hwang et al., 2003; You et al., 2009). TP53 mutation carrier have just a slightly higher incidence of lung cancer, compared to the general population, but the mean age of onset is noticeable lower (Birch et al., 2001; Olivier et al., 2003). Two genes of low to moderate penetrance, glutathione S transferase M1 (GSTM1) and CHEK2, have been identified through candidate gene studies. Individuals with a deletion in the GSTM1 gene, have a diminished ability to detox environmental carcinogens, whereby the risk of lung cancer is elevated (Shi et al., 2008; Vineis et al., 2007). Cancer of the bladder is, like lung cancer, strongly influenced by tobacco exposure and also associated with GSTM1-null carriers, supporting the role of GSTM1 as a susceptibility marker for lung cancer (Garcia-Closas et al., 2005). However, confirmation in larger cohorts is required (Brennan et al., 2011). Rare variants of CHEK2, an important cell-cycle control gene with key role in the maintenance of cell integrity, are significantly associated with a decreased risk of lung cancer (Brennan et al., 2007), whereas they have a contrary association with several other cancers (Cybulski et al., 2004). The underlying mechanisms of this opposed functionality is unclear (Brennan et al., 2011). GWAS identified three low-penetrance loci, 15q25.1, 6p21 and 5p15.33, associated with lung cancer, accounting for 7% of the familial risk (Varghese and Easton, 2010). The polymorphism at 15q25.1 varies lung cancer susceptibility, either direct (Amos et al., 2008; Hung et al., 2008; Thorgeirsson et al., 2008) or implicit by an enhanced nicotine dependence (Saccone et al., 2007). The second susceptibility locus, located at 6p21, resides within HLA class II region, but the functionality is little understood (Wang et al., 2008b). Furthermore, polymorphisms at 5p15.33, a chromosomal region containing the cleft lip and palate transmembrane protein 1-like protein gene (CLPTM1L) and the telomerase reverse transcriptase (TERT) gene, were found to be associated with lung cancer, especially with adenocarcinoma. All identified SNPs are intronic and residue within a region of LD containing the CLPTM1L gene and the 5` end of the telomerase gene, including the TERT promoter. Besides SNPs, non coding minisatellites, like MNS16A, were found to modulate lung cancer risk dependent on the number of tandem repeats, as reviewed in (Baird, 2010). MNS16A is located near the telomerase gene, which is a promising hotspot in cancer research.

1.5. Telomerase

Telomeres are stabilizing ribonucleoprotein structures at chromosome ends, preventing them from being recognized as DNA double strand breaks by the DNA repair machinery (Baird, 2010). Telomerase is composed of TERT, the catalytic subunit, telomere template-containing RNA component (TERC), the RNA template for telomeric repeats synthesis and telomerase-associated proteins (Cong et al., 2002).

During cell division, synthesis of the lagging strand remains incomplete, resulting in a progressive shortening of telomeres, termed as end-replication problem. Each cell division leads to a telomere abbreviation of 25-200bp. Once telomeres become critically short, the cell enters p53 mediated replicative senescence, a viable but quiescent state, referred as the Hayflick limit (Dhaene et al., 2000). The limited replicative potential is an important tumor suppressor mechanism through preserving the cell from accumulation of multiple oncogenic mutations, cluttered during life span. It functions as a barrier against tumorigenesis (Cong et al., 2002). Due to mutations within the p53 pathway, critically short telomeres escape from the cell cycle arrest, whereby the life span is extended by around 30-40 doublings (Dhaene et al., 2000). Finally telomeres get ultra-short and genomic instability, telomere fusion, dicentric chromosomes and genomic rearrangements are the consequences (Baird, 2010). This state is called crisis and usually it is associated with cell death. A few cell bypass crisis through re-activation of telomerase expression, leading to cell immortalization and unlimited proliferation, a necessity for cancer. Reactivated and up-regulated telomerase expression is a critical event in tumorigenesis, at least some immortal cells are present in most malignant tumors (Dhaene et al., 2000). Besides mutations in tumor-suppressor genes and proto-oncogenes, telomerase expression is an essential requirement for a tumorigenic cell conversion. Around 90% of tumor samples show telomerase activity (Cong et al., 2002). The presence of telomerase is a useful and common tumor marker for diagnostic and prognostic purposes. Somatic cells, except for germ cells and cycling stem cells of self-renewing tissues, do not show telomerase expression (Shay and Bacchetti, 1997). Besides telomerase dependent lengthening of telomeres, alternative lengthening of telomeres (ALT) is capable of telomere stabilization (Dhaene et al., 2000).

The limiting factor of telomerase activity is its catalytic subunit, telomerase reverse transcriptase, which is encoded by the TERT gene, located at 5p15.33. Up-regulation of telomerase in tumors is sometimes caused by the amplification of the TERT locus. In 13% of cancer samples the short arm of the fifth chromosome, where TERT gene is located, is amplified, the 5p15.33 segment is amplified frequently in many cancer types, for example in

early-stages of SCLC. The TERT gene is subject to a complex regulation, including transcriptional activators and repressors and epigenetic events, indicating that little variation within this area may result in an altered telomerase expression level. It is also highly conserved with only a slight inter-individual genetic variance, supporting the thesis, that variation affects telomerase activity and therefore telomere length. The 5p15.33 region also contains the CLPTM1L gene, which is involved in cell apoptosis, induced by cisplatin treatment. Multiple GWAS identified strong associations between polymorphisms within the TERT/CLPTM1L locus and the risk of developing basal cell cancer, pancreatic cancer and lung cancer. Furthermore, risk alleles, associated with bladder cancer, cervical cancer, prostate cancer and glioma were identified by GWAS (Baird, 2010).

1.5.1. MNS16A-Polymorphism

The polymorphic tandem repeats minisatellite MNS16A is located downstream of exon 16 of telomerase gene, within the promoter region of a TERT antisense transcript (indicated by red circle). It was first identified in 2003 by Wang et al. in the course of a lung cancer study. The region of the TERT gene is rich in micro- and minisatellites, which are indicated by asterisks (Figure 5) (Wang et al., 2003).

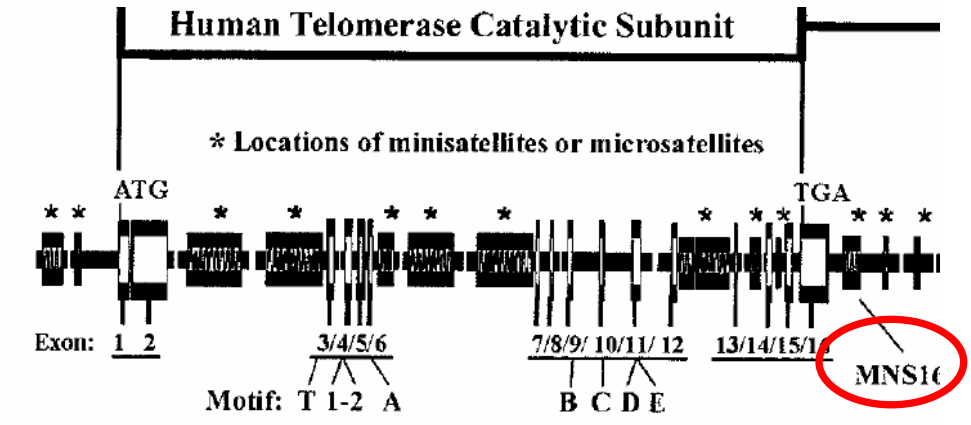

Figure 5: Wang et al. 2003; TERT gene

The repetitive core element is composed of either a 23bp or a 26bp sequence, containing CAT trinucleotide insertions. These insertions operate as binding sites for the transcription factor GATA-1. Four variable number of tandem repeats (VNTRs) of MNS16A are known and

termed according to their Polymerase Chain Reaction (PCR) fragment size (VNTR-243,-274,-302,-333) (Wang et al., 2003). Recently a fifth rare MNS16A variant was identified, VNTR-364 (Hofer et al., 2011).

Figure 6 illustrates the wild-type allele, VNTR-302, containing five tandem repeats, indicated by boxes, separated by CAT inserts. VNTRs are often classified in short (S) variants (VNTR-243,-274) and long (L) variants (VNTR-302,-333,-364).

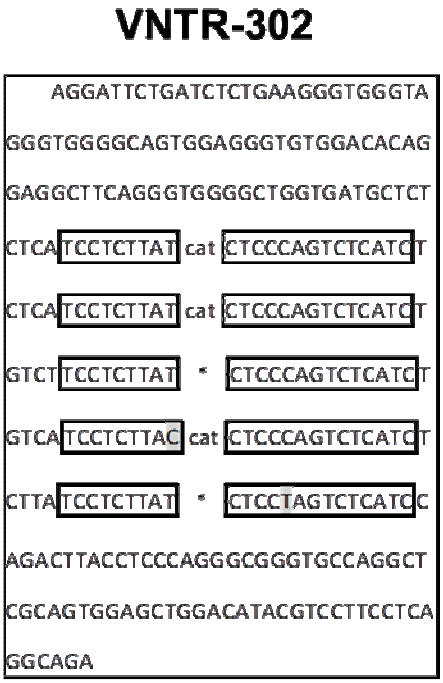

Figure 6: Hofer et al. 2011; Nucleotide sequence of wild-type VNTR-302

It has been shown that this functional minisatellite has a promoter activity for an hTERT antisense transcript that is dependent on the number tandem repeats. Shorter variants were shown to have a higher promoter activity suggesting a role in the repression of telomerase expression. This minisatellite has been investigated in the context of lung, brain, breast, colorectal and nasopharyngeal cancer with controversial results (Andersson et al., 2009; Carpentier et al., 2007; Hofer et al., 2011; Jin et al., 2011; Wang et al., 2003; Wang et al., 2010; Wang et al., 2006; Wang et al., 2008c; Zhang et al., 2011).

1.6. Aim of the study

The aim of the present diploma thesis was to investigate whether the polymorphic tandem repeats minisatellite MNS16A is associated with the risk of prostate cancer and/or lung cancer.

Based on a prostate cancer study population of 1774 Austrian participants, the role of MNS16A-single alleles, genotypes and SL-genotypes as potential biomarkers to distinguish between the aggressive and the indolent form of prostate cancer was assessed. A marker to estimate the aggressiveness of prostate cancer helps to adapt cancer treatment. In addition, MNS16A variants were investigated as potential biomarkers for lung cancer risk within a study population of 3134 Austrian participants. Finding a minimal invasive marker for lung cancer is an essential issue, particularly because early detection of lung cancer is fundamental for lung cancer therapy.

2. Material and Methods

2.1. Material

2.1.1. Study population

2.1.1.1. Prostate cancer

Blood samples of the prostate cancer cohort were collected in the general hospital, the Sozialmedizinisches Zentrum Süd (or Kaiser Franz-Josef hospital), and the Sozialmedizinisches Zentrum Ost. Together with the written informed consent, a short questionnaire on the smoking habits was obtained.

The study population comprised 1137 newly diagnosed prostate cancer cases and 637 hospital-based controls with BPH. Case subjects were classified according to the aggressiveness of their tumor in 505 patients with Gleason Score 2-6 and 490 patients with Gleason Score 7-10. For all patients, histological records were available. All controls underwent biopsy of the prostate and were therefore known to be free of prostate cancer.

2.1.1.2. Lung cancer

Blood samples of the lung cancer cohort were collected from the LHK Grimmenstein Hochegg and from the Sozialmedizinisches Zentrum Baumgartner Höhe. The study population comprised 665 lung cancer patients that can be further classified in 132 SCLC patients and 533 NSCLC patients. The control group comprised 2469 participants and consisted of 757 hospital-based controls including 120 orthopedics controls and 637 BPH controls and additional 1712 population-based controls.

2.2.Methods

2.2.1. DNA Isolation

Isolation of genomic DNA from peripheral blood was performed with a spin-column based nucleic acid purification kit, namely QIAamp DNA blood midi kit (QIAGEN, Hilden Germany), according to manufacturer's instructions.

200µl QIAGEN Protease solution were provided in a 15ml tube, 1-2ml peripheral blood were added and mixed briefly. 2.4ml lysis Buffer AL were added followed by mixing for one minute, before the tubes were placed in a shaking water bath for 30min at 70°C. After the samples cooled down for a few minutes, 2ml 100% Ethanol were added to each sample and

mixed for one minute. Half of the lysate was transferred onto the midi spin column and centrifuged for 5min at 1850xg (Eppendorf Centrifuge 5810 R, Eppendorf Germany). The remaining half was also transferred onto the same midi spin column and centrifuged on equal terms. The membrane bound DNA was rinsed with 2ml wash Buffer AW1 and centrifuged for 2min at 4000rpm. Afterwards 2ml wash Buffer AW2 were pipetted onto the spin column and tubes were centrifuged for 18min at 4000rpm. Purified DNA was eluted with 300µl elution Buffer AE. After 5min of incubation at room temperature, tubes were centrifuged for 7min at 4000rpm.

The concentration of genomic DNA was quantified by Nanodrop ND-1000 Spectrophotometer (PEQLAB Biotechnologie GmbH, Erlangen Germany) as provided in manufacturer's instructions. Purified DNA was stored at -80°C. Working dilutions of 10ng/µl were obtained by dilution with Aqua bidest. "Fresenius" (Fresenius Kabi, Austria) and stored at -20°C.

2.2.2. Genotyping of MNS16A

MNS16A genotypes were determined by PCR amplification followed by electrophoretical separation in 2.5% agarose gels and visualization of PCR products using UV light excitation. Standard PCR was done in 20µl reactions containing 40ng of genomic DNA and 16µl master mix set up in 0.2ml PCR SOFTTUBES (Biozym, Hessen Germany).

Master mix consisted of 2µl Gene Amp 10xPCR Buffer II (Applied Biosystems, California US), 1.2µl 25mM $MgCl_2$ Solution (Applied Biosystems, California US), 0.3 µl dNTP Mix, 10mM each (Applied Biosystems, California, US), each 3.5µl 5µM forward (5`-AGG ATT CTG ATC TCT GAA GGG TG-3`) and reverse primer (5`-TCT GCC TGA GGA AGG ACG TAT G-3`) (VBC-BIOTECH Service, Vienna Austria), 0.06µl 5 U/µl AmpliTaq DNA Polymerase (Applied Biosystems, California US) and 5.44µl Aqua bidest. "Fresenius" (Fresenius Kabi, Austria).The whole procedure was carried out on ice.

PCR amplification was performed on a 2720 Thermal Cycler (Applied Biosystem, California US) by the following thermal profile: initial denaturing step for 5min at 95°C followed by 35 cycles of 95°C (30s), 65°C (30s), 72°C (30s) and a final extension step of 10min at 72°C.

For visualization of PCR amplicons agarose gel electrophoresis coupled with UV detection of the Ethidium Bromid (EtBr) stained DNA molecules was accomplished. Hence, 50ml 2.5% agarose gels were set up after following formula: 1.25g agarose (Biozym LE Agarose [Biozym, Hessen Germany]) were disbanded in 50ml 1xTBE buffer (10.8g/L Tris(hydroxymethyl)-aminomethane [Merck, Darmstadt Germany], 5.5 g/L boracid acid [Alfa

Aesar, Karlsruhe Germany], 4ml/L 0.5M EDTA pH8 [Merck, Darmstadt Germany]) and mixed with 25ng EtBr (Amresco, Ohio US).
Electrophoretical separation was carried out in Hoefer HE 33 Mini Horizontal Submarine Units (Pharmacia Biotech, Zurich Switzerland) at 130V for 50min. 20µl PCR product and additional 4µl 6x Loading Dye (Fermentas Int. Inc., Ontario Canada) were loaded. On each gel 0.2ng GeneRuler 50bp DNA Ladder (Fermentas Int. Inc., Ontario Canada) was loaded as length standard. Genotype patterns were made visible by UV light exposure of EtBr stained agarose gels with GelDoc detection system (Bio-Rad Laboratories, California US).

2.2.3. Cloning of MNS16A PCR products

Cloning of MNS16A PCR products was performed with TOPO TA Cloning Kit Dual Promoter, containing pCRII-TOPO plasmid (Figure 7) utilizing chemically competent TOP10 bacterial cells for transformation and blue/white screening, without IPTG as an induction molecule, following manufacturer´s instructions (Invitrogen, California US).
TOPO TA Cloning kit provides direct insertion of Taq polymerase amplified, sticky end products into plasmid vector. Taq polymerase is supplied with a nontemplate-dependent terminal transferase activity that adds a single deoxyadenosine to the 3´ends of the amplified sequence. The linearized plasmid pCRII-TOPO is equipped with compatible deoxythymedine overhangs. Furthermore, the supplied plasmid is activated by the covalently bound enzyme topoisomerase I. Topoisomerase I cleaves duplex DNA after the sequence 5`-CCCTT in one strand. The resulting energy is conserved in a covalent bond between the 3`phosphate of the cleaved strand and a tyrosyl residue of topoisomerase I. This bond can be attacked by a 5`hydroxyl group, and consequently topoisomerase is released. For this reasons, insertion can happen directly, without any post PCR modifications nor the use of ligase.

Cloning Reaction

The cloning reaction was made of 0.5µl salt solution (1.2 M NaCl, 0.06M $MgCl_2$), 0.5µl vector (10ng/µl plasmid DNA provided in 50% glycerol, 1mM EDTA, 1mM DTT, 0.1% Triton X-100, 100µg/ml BSA, 50mM Tris-HCL pH 7.4 and phenol red) and 2µl PCR product. After 20min of incubation at room temperature, the mixture was placed on ice.

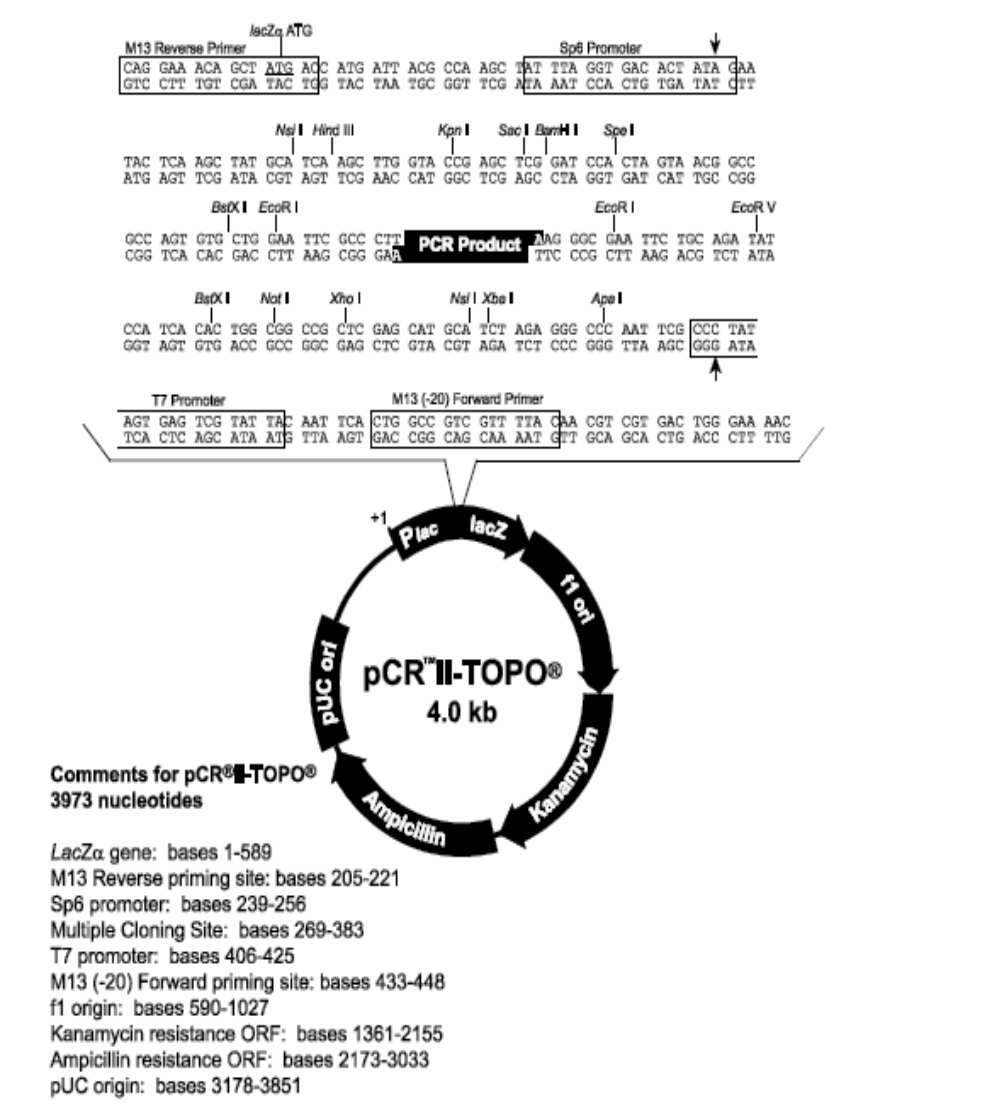

Figure 7: Invitrogen, Cloning site of PCRII- TOPO vector

Transformation

Two vials of One Shot chemically competent *E.coli*, each with 200µl volume, were thawed on ice. In one vial 4µl of the prepared TOPO cloning mixture were added and mixed gently. To determine the transformation efficiency 4µl 10pg/µlpUC 19 DNA (Invitrogen, California US) were added to the second vial of chemically competent cells. The following transformation steps were performed equally for cloning and control reaction. After an incubation on ice (20min), a heat shock of 30sec at 42°C was applied by water bath, whereby cell membranes get temporary porous and plasmids are able to enter.

Instantly, vials were placed on ice and 250ml room tempered SOC medium (6ml SOB medium, 60µl 1M MgCl2 [Merck, Darmstadt Germany] , 60µl 2M glucose [Sigma-Aldrich, Inc., Missouri US]) were added. SOB medium was made of 20g/L Tryptone (Roth, Karlsruhe Germany), 5g/L yeast extract (Sigma-Aldrich Inc., Missouri US), 0.5g/L NaCl (Merck,

Darmstadt Germany), 10ml 250mM KCL (Merck, Darmstadt Germany) and was finally adjusted to pH 7 with NaOH (Merck, Darmstadt Germany). Cell suspensions were then incubated at 37°C and 200rpm for 1 hour to induce the expression of antibiotic resistances. 50µl of each transformation suspension were plated on pre-warmed selective LB agar plates (20g/L LB, 15g/L agar [Sigma-Aldrich Inc., Missouri US] and 100µg/ml ampicillin [AppliChem, Darmstadt Germany]), previously plated with 40µl X-gal (AppliChem, Darmstadt Germany). The plates were incubated over night at 37°C.

Next day, white colonies were picked and solitary transferred to 14ml tubes, containing 5ml ampicillin-dosed LB-medium. For picking, sterile single-use loops (Nunc, Wiesbaden Germany) were utilized.

Cell suspensions were incubated over night at 37°C and 230rpm.

To assure, that the right PCR product inserted into plasmid vector while TOPO cloning reaction, a PCR was accomplished. Therefore, 4µl of overnight culture were diluted with 95µl "Fresenius" aqua bidest. (Fresenius kabi, Austria) and heated for 15min at 95°C. This way, cell membrane and wall disrupt and release the plasmid. 1µl plasmid DNA suspension was then used as template for MNS16A-specific PCR, followed by an electrophoretical separation in 2.5% agarose gels. Expected PCR product lengths confirm a successful cloning.

Positive overnight cultures underwent the procedure of plasmid DNA isolation utilizing Wizard Plus SV Minipreps DNA Purification System (Promega, California US). The production of a cleared lysate was carried out following the provided protocol, comprising harvest and lysis of bacterial cells.

Plasmid DNA was then precipitated and purified according to the provided centrifugation protocol, and finally eluted in 30µl nuclease free water. For sequencing, 10µl at a concentration of 80ng/µl was sent to Microsynth (Balgach, Switzerland).

2.3. Statistical analysis

Statistical analysis was performed in the R programming environment and carried out by Mr. Andreas Baierl. Control groups were tested for Hardy-Weinberg Equilibrium (HWE) using a Chi^2 -test, which is important as a quality control. By multiple logistic regression, odds ratios and 95% confidence intervals were estimated for the genotype, the single VNTRs and the SL classification. Covariants age and smoking were used in all models as confounders; additionally, gender was used in the lung cancer population.

3. Results

3.1. Studypopulation

3.1.1. Prostate Cancer Cohort

The prostate cancer study population comprised 1787 participants, composed of 1137 newly diagnosed prostate cancer cases and 650 controls with BPH. The age distribution was significantly different between cases and controls, with a mean age of 63.8 years for the cases and 67.4 years for the BPH controls. Therefore all statistical models were adjusted for age. Smoking habits did not significantly differ between carcinoma patients and control subjects. 58.1% of all cases were classified as former or current smokers, in the control group it was slightly more. For 87 cases and 31 controls no data regarding smoking habits were available (Table 1).

Table 1. Age distribution and smoking status of prostate cancer study population

	PC patients	Gleason Score <7	Gleason Score >=7	BPH Controls
	N (%)	N (%)	N (%)	N (%)
Age (years)				
<50	55 (4.8)	32 (6.3)	21 (4.3)	18 (2.8)
<60	308 (27.1)	136 (26.9)	127 (25.9)	133 (20.5)
<70	563 (49.5)	259 (51.3)	244 (49.8)	270 (41.5)
<80	190 (16.7)	73 (14.5)	90 (18.4)	181 (27.8)
>80	21 (1.8)	5 (1.0)	8 (1.6)	48 (7.4)
Total	1137 (100)	505 (100)	490 (100)	650 (100)
Mean Age (SD a)	63.8 (7.7)	63.4 (7.6)	63.9 (7.5)	67.4 (8.9)
Smoking				
never	435 (41.4)	188 (40.7)	189 (41.7)	238 (38.4)
former	465 (44.3)	203 (43.9)	199 (43.9)	291 (47.0)
current	150 (14.3)	71 (15.4)	65 (14.3)	90 (14.5)
Smoking rate b	615 (58.6)	274 (59.3)	264 (58.2)	381 (61.5)
Total	1050 (100)	462 (100)	453 (100)	619 (100)

a Standard deviation; b Former and current smokers

3.1.2. Lung Cancer Cohort

Demographic factors and smoking history were evaluated in the lung cancer study population consisting of 3167 participants. The lung cancer patients (n= 619) were classified according to their histological subtype in a SCLC (n=125) and a NSCLC (n=476) subgroup. The entire control group encompassing 2548 subjects consisted of 118 orthopedics controls, 650 BPH

controls and 1780 population-based controls. Men were significantly overrepresented in the case group, composing 69.1% of all cases. Due to the BPH control subgroup, men were also more prevalent in the total control population. The mean ages of cases and controls were significantly different, with a mean age of 63.4 years for the cases, 62.0 years for the orthopedics controls, 67.4 years for the BPH controls and 61.3 years for the population-based controls. The average mean age was slightly lower in the control group, compared to the case group. Smokers were significantly overrepresented in the case group, with 47.1% former and 42.4% current smokers, while it was 32.4% former and 14.9% current smokers in the overall control population. For 165 participants no information on their smoking status was available. All lung models were adjusted for gender, age and smoking (Table 2). The MNS16A genotype distribution of all control groups were found to be within the HWE.

Table 2. Sex/Age distribution and smoking status of lung cancer study population

	Lung cancer patients	SCLC	NSCLC	Ortho. Controls	BPH Controls	Pop.-based Controls	Total Controls
	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)	N (%)
Sex							
Male	428 (69.1)	86 (68.8)	331 (69.5)	44 (37.3)	650 (100)	829 (46.6)	1523 (59.8)
Female	191 (30.9)	39 (31.2)	145 (30.5)	74 (62.7)	0	951 (53.4)	1025 (40.2)
Total	619 (100)	125 (100)	476 (100)	118 (100)	650 (100)	1780 (100)	2548 (100)
Age (years)							
<50	68 (11.0)	11 (8.8)	55 (11.6)	21 (17.8)	18 (2.8)	339 (19.0)	378 (14.8)
<60	188 (30.4)	40 (32.0)	142 (29.8)	29 (24.6)	133 (20.5)	428 (24.0)	590 (23.2)
<70	196 (31.7)	43 (34.4)	149 (31.3)	36 (30.5)	270 (41.5)	554 (31.1)	860 (33.8)
<80	141 (22.8)	27 (21.6)	108 (22.7)	25 (21.2)	181 (27.8)	428 (24.0)	634 (24.9)
>80	26 (4.2)	4 (3.2)	22 (4.6)	7 (5.9)	48 (7.4)	31 (1.7)	86 (3.4)
Total	619 (100)	125 (100)	476 (100)	118 (100)	650 (100)	1780 (100)	2548 (100)
Mean Age (SD a)	63.4 (10.1)	63.2 (9.9)	63.4 (10.3)	62.0 (12.9)	67.4 (8.9)	61.3 (11.2)	62.9 (11.1)
Smoking							
Never	63 (10.6)	7 (5.9)	52 (11.3)	46 (39.7)	238 (38.4)	1017 (58.7)	1301 (52.7)
Former	281 (47.1)	56 (47.5)	216 (46.9)	41 (35.3)	291 (47.0)	467 (26.9)	799 (32.4)
Current	253 (42.4)	55 (46.6)	193 (41.9)	29 (25.0)	90 (14.5)	249 (14.4)	368 (14.9)
Smoking rate b	534 (89.5)	111 (94.1)	409 (88.8)	70 (60.3)	381 (61.5)	716 (41.3)	1167 (47.3)
Total	597 (100)	118 (100)	461 (100)	116 (100)	619 (100)	1733 (100)	2468 (100)

a Standard deviation; b Former and current smokers

3.2. MNS16A genotyping

MNS16A genotyping was successful in 98.1% of prostate samples and in 98.6% of lung samples. Six different VNTRs of MNS16A were identified (VNTR-364,-333,-302,-274,-243,-212), including the newly found variant VNTR-212. VNTRs are termed according to their PCR fragment size and are illustrated in Figure 8. The shortest variant, VNTR-212, was discovered for the first time in this study. MNS16A alleles appeared in 12 different genotype patterns, which are summarized in Table 3. Two genotypes, namely the homozygote 274/274 genotype and the genotype 212/302, containing the novel allele, were also found for the first time. Alleles can be classified in short (S) (VNTR-212,-243,-274) and long (L) (VNTR-302,-333,-364) variants, resulting in SL-genotypes (table 3). The rare genotypes 212/302 and 274/274 were found only once and were therefore excluded from statistical models.

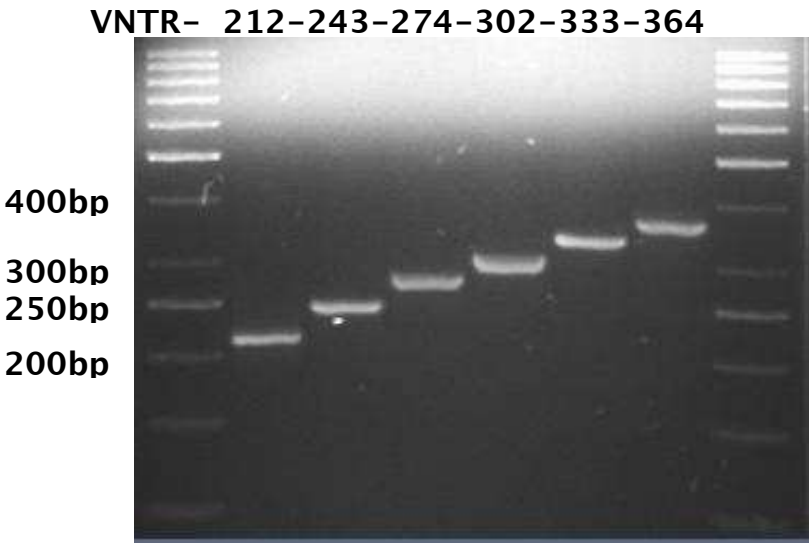

Figure 8. Six different VNTRs of MNS16A

Table 3. Genotype patterns

Genotype	SL-classification
302/302	LL
243/243	SS
243/302	SL
302/333	LL
243/274	SS
274/302	SL
243/333	SL
333/333	LL
302/364	LL
274/333	SL
274/274	SS
212/302	SL

The identity of the newly found variant was verified by cloning and sequencing and the nucleotide sequence is illustrated in Figure 9. The five previously known VNTRs are shown in Figure 10. VNTR-212 is the shortest VNTR that has been reported so far and contains only two of the characteristic tandem repeats. It is 31 base pairs shorter than the VNTR-243 and contains only one CAT trinucleotide insert, thereby carrying also the smallest number in transcription factor binding sites.

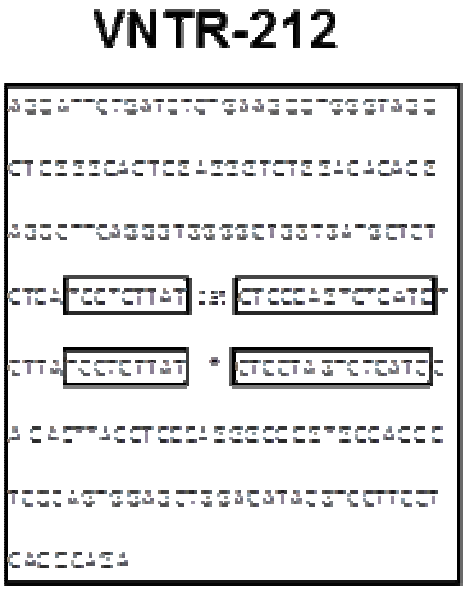

Figure 9. Base sequence of VNTR-212

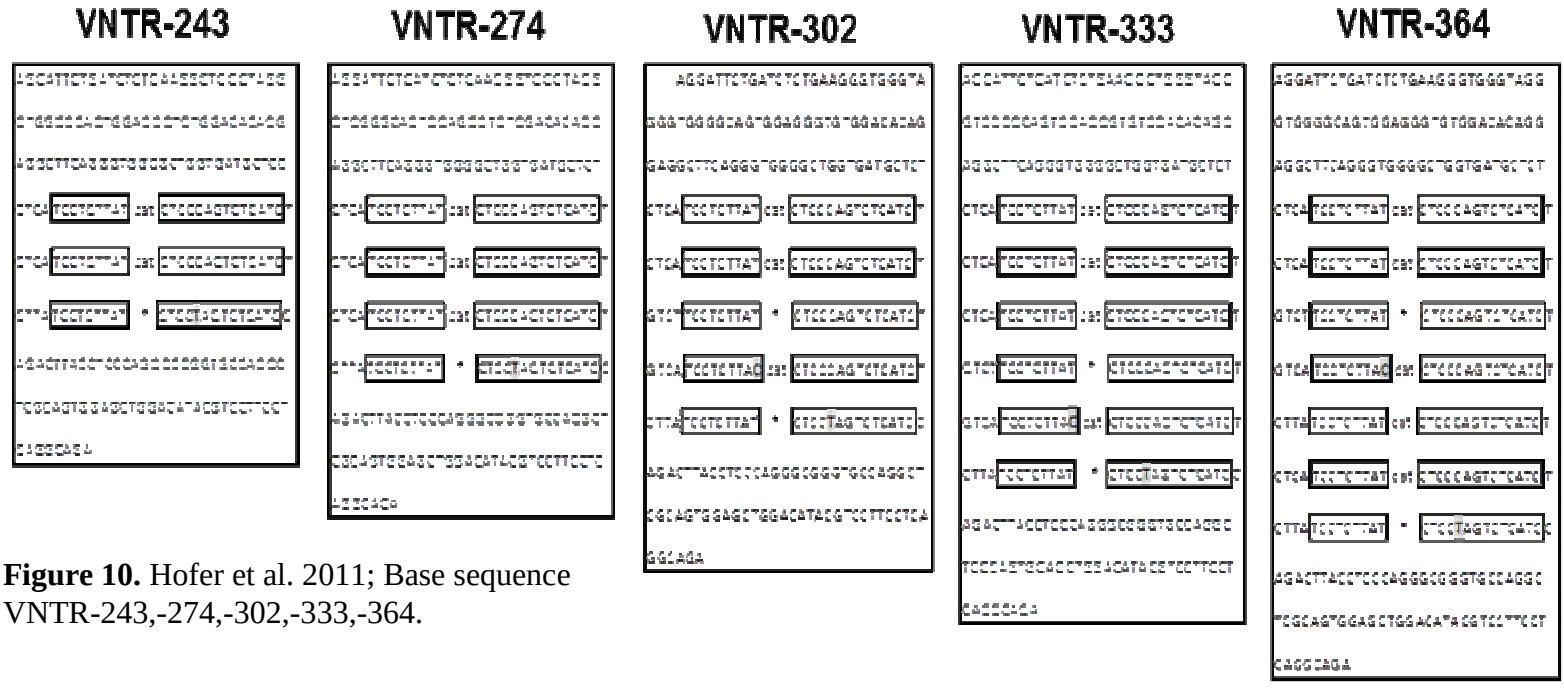

Figure 10. Hofer et al. 2011; Base sequence VNTR-243,-274,-302,-333,-364.

3.2.1. Prostate cancer

3.2.1.1. Overall study population

By multiple logistic regression analysis, no significant effects were observed in the overall prostate cancer population. The genotype distribution of cases and controls are illustrated in Figure 11, genotype, allele and SL-genotype distribution, containing odds ratio (OR), 95% confidence interval and p-value, are summarized in Table 4.

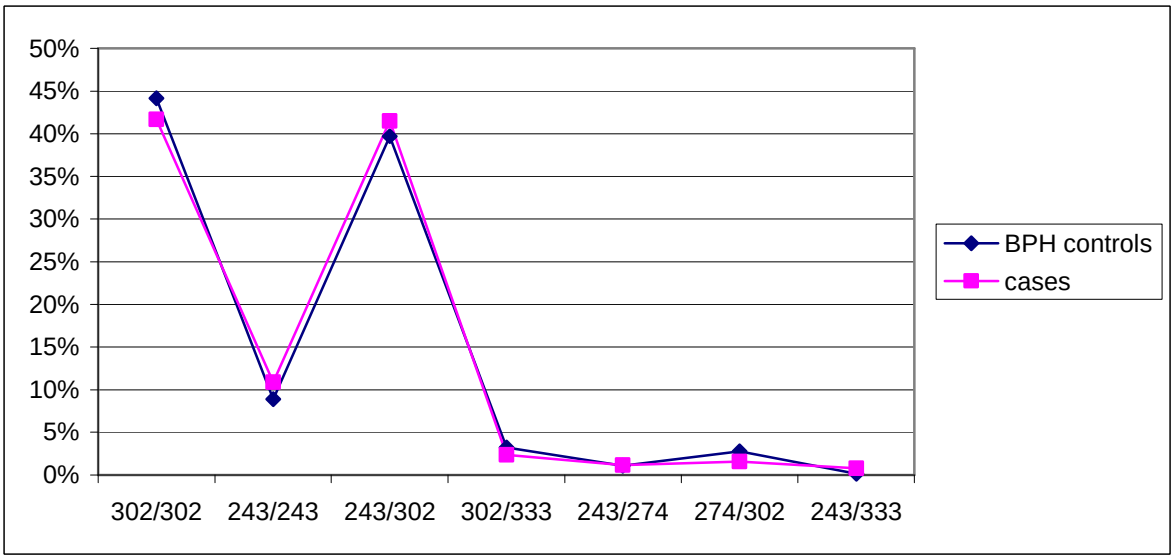

Figure 11. Genotype distribution of the prostate cancer study population

Table 4. Genotype distribution of the prostate cancer study population

	BPH-Controls	PC patients		
Genotype	N (%)	N (%)	OR a (95% CI b)	P-value
302/302	287 (44.2)	474 (41.7)	1.00	
243/243	58 (8.9)	124 (10.9)	1.27 (0-89-1.81)	0.180
243/302	258 (39.7)	472 (41.5)	1.13 (0.91-1.41)	0.261
302/333	21 (3.2)	27 (2.4)	0.75 (0.41-1.36)	0.341
243/274	7 (1.1)	13 (1.1)	1.03 (0.40-2.66)	0.958
274/302	18 (2.8)	18 (1.6)	0.67 (0.33-1.34)	0.252
243/333	1 (0.2)	9 (0.8)	4.43 (0.56-35.26)	0.160
Total	650 (100)	1137 (100)		
Alleles				
302	871 (67.0)	1465 (64.4)	1.00	
243	382 (29.4)	742 (32.6)	1.15 (0.99-1.34)	0.068
274	25 (1.9)	31 (1.4)	0.76 (0.44-1.31)	0.321
333	22 (1.7)	36 (1.6)	0.90 (0.52-1.56)	0.719
SL-Genotype				
LL	308 (47.4)	501 (44.1)	1.00	
SL	277 (42.6)	499 (43.9)	1.14 (0.92-1.40)	0.229
SS	65 (10.0)	137 (12.0)	1.27 (0.91-1.77)	0.165
SL+SS	342 (52.6)	636 (55.9)	1.16 (0.95-1.42)	0.140

a Odds ratio; b 95% confidence interval

3.2.1.2. Subpopulation ≥ 70 years

In a subgroup analysis of men aged 70 and older the genotype 274/302 was associated with a significantly reduced risk of prostate cancer, when compared to the wild-type genotype (OR=0.09) Also, regarding the single alleles model this effect for VNTR-274 remained significant with an OR of 0.23. All results are summarized in Table 5, significant associations are given in bold.

Table 5. Genotype distribution of the prostate cancer study population ≥ 70

	BPH-Controls	PC patients ≥ 70		
Genotype	N (%)	N (%)	OR a (95% CI b)	P-value
302/302	94 (41.0)	91 (43.1)	1.00	
243/243	22 (9.6)	21 (10.0)	0.85 (0.42-1.71)	0.653
243/302	92 (40.2)	92 (43.6)	1.10 (0.72-1.70)	0.658
302/333	8 (3.5)	4 (1.9)	0.48 (0.14-1.70)	0.257
243/274	2 (0.9)	2 (0.9)	0.99 (0.12-8.45)	0.998
274/302	**11 (4.8)**	**1 (0.5)**	**0.09 (0.01-0.72)**	**0.023**
Total Alleles	229 (100)	211 (100)		
302	299 (65.3)	279 (66.1)	1.00	
243	138 (30.1)	136 (32.2)	1.04 (0.76-1.40)	0.821
274	**13 (2.8)**	**3 (0.7)**	**0.23 (0.06-0.83)**	**0.026**
333	8 (1.7)	4 (0.9)	0.49 (0.14-1.70)	0.262
SL-Genotype				
LL	102 (44.5)	95 (45.0)	1.00	
SL	103 (45.0)	93 (44.1)	1.03 (0.68-1.56)	0.895
SS	24 (10.5)	23 (10.9)	0.90 (0.46-1.77)	0.768
SL+SS	127 (55.5)	116 (55.0)	1.00 (0.67-1.49)	0.990

a Odds ratio; b 95% confidence interval

3.2.2. Lung cancer

3.2.2.1. Overall study population

In our overall lung cancer population, no evidence of a significant association of MNS16A genotypes with risk of lung cancer was identified. By analysing the genotype distribution of all control groups compared to the case group no significant effects could be observed, as summarized in Table 6 and visualized in Figure12. Same results were obtained by investigation of associations between the control subgroups and the total control group with the NSCLC subpopulation (Table 7).

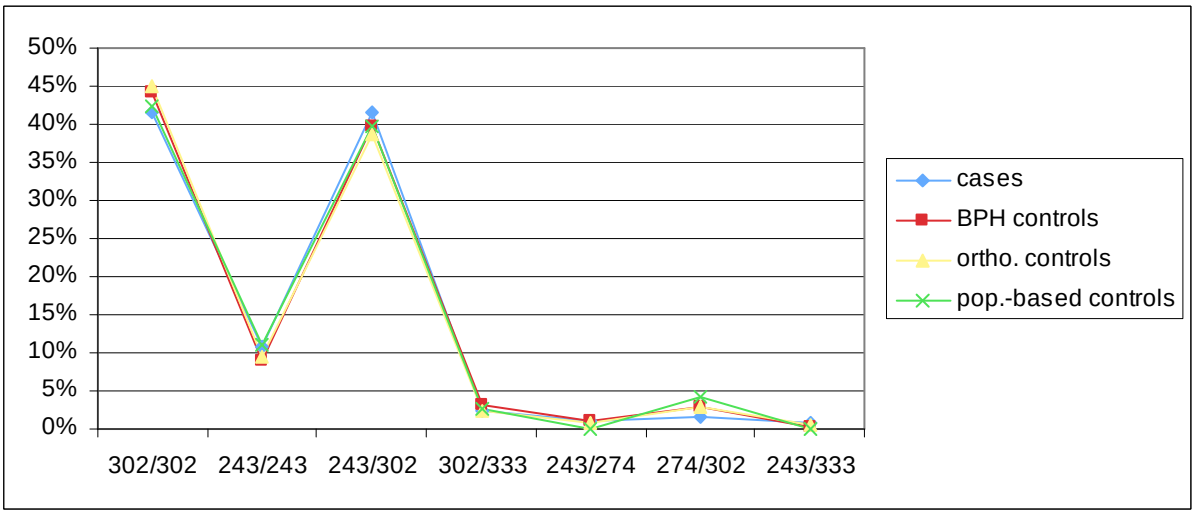

Figure 12. Genotype distribution of the lung cancer study population

Table 6. Genotype distribution of lung cancer study population

	Lung cancer patients	Population based Controls			Hospital based Controls			Total Controls		
Genotype	N (%)	N (%)	OR a (95% CI b)	P-value	N (%)	OR a (95% CI b)	P-value	N (%)	OR a (95% CI b)	P-value
302/302	270 (45.2)	740 (42.7)	1		317 (43.1)	1		1057 (42.8)	1	
243/243	56 (9.4)	186 (10.7)	0.83 (0.57-1.21)	0.335	68 (9.3)	0.88 (0.57-1.36)	0.57	254 (10.3)	0.86 (0.61-1.21)	0.379
243/302	230 (38.5)	701 (40.5)	0.92 (0.73-1.16)	0.489	298 (40.5)	0.92 (0.70-1.20)	0.543	999 (40.5)	0.94 (0.76-1.16)	0.567
302/333	15 (2.5)	35 (2.0)	1.16 (0.56-2.41)	0.685	23 (3.1)	0.76 (0.36-1.61)	0.48	58 (2.4)	0.88 (0.47-1.66)	0.704
243/274	4 (0.7)	14 (0.8)	0.96 (0.25-3.70)	0.957	7 (1.0)	0.50 (0.13-1.91)	0.309	21 (0.9)	0.71 (0.22-2.31)	0.572
274/302	18 (3.0)	41 (2.4)	1.48 (0.75-2.90)	0.257	21 (2.9)	0.95 (0.44-2.03)	0.894	62 (2.5)	1.18 (0.65-2.16)	0.582
243/333	4 (0.7)	16 (0.9)	0.74 (0.22-2.47)	0.624	1 (0.1)	6.54 (0.55-77.40)	0.136	17 (0.7)	1.09 (0.34-3.56)	0.884
Total	597 (100)	1733 (100)			735 (100)			2468 (100)		
Alleles										
302	803 (67.3)	2257 (65.1)	1		976 (66.4)	1		3233 (65.5)	1	
243	350 (29.3)	1103 (31.8)	0.90 (0.76-1.06)	0.217	442 (30.1)	0.94 (0.77-1.13)	0.506	1545 (31.3)	0.93 (0.80-1.08)	0.318
274	22 (1.8)	55 (1.6)	1.38 (0.76-2.49)	0.294	28 (1.9)	0.84 (0.44-1.61)	0.6	83 (1.7)	1.08 (0.64-1.83)	0.778
333	19 (1.6)	51 (1.5)	1.04 (0.56-1.92)	0.896	24 (1.6)	0.98 (0.50-1.93)	0.954	75 (1.5)	0.94 (0.54-1.63)	0.827
SL-Genotyp										
LL	285 (47.7)	775 (44.7)	1		340 (46.3)	1		1115 (45.2)	1	
SL	252 (42.2)	758 (43.7)	0.94 (0.75-1.17)	0.566	320 (43.5)	0.95 (0.73-1.23)	0.704	1078 (43.7)	0.96 (0.78-1.18)	0.712
SS	60 (10.1)	200 (11.5)	0.83 (0.58-1.20)	0.324	75 (10.2)	0.85 (0.56-1.30)	0.457	275 (11.1)	0.85 (0.61-1.19)	0.344
SL+SS	312 (52.3)	958 (55.3)	0.91 (0.74-1.13)	0.412	395 (53.7)	0.93 (0.73-1.19)	0.566	1353 (54.8)	0.94 (0.77-1.14)	0.524

a Odds ratio; b 95% confidence interval

Table 7. Genotype distribution of the NSCLC study population

	NSCLC patients	Population based controls			Hospital based controls			Total Controls		
Genotype	N (%)	N (%)	OR a (95% CI b)	P-value	N(%)	OR a (95% CI b)	P-value	N(%)	OR a (95% CI b)	P-value
302/302	216 (46.9)	740 (42.7)	1		317 (43.1)	1		1057 (42.8)	1	
243/243	43 (9.3)	186 (10.7)	0.78 (0.52-1.17)	0.236	68 (9.3)	0.83 (0.52-1.33)	0.45	254 (10.3)	0.81 (0.56-1.19)	0.279
243/302	169 (36.7)	701 (40.5)	0.86 (0.66-1.11)	0.236	298 (40.5)	0.87 (0.66-1.16)	0.36	999 (40.5)	0.88 (0.69-1.11)	0.268
302/333	12 (2.6)	35 (2.0)	1.24 (0.57-2.69)	0.584	23 (3.1)	0.74 (0.33-1.64)	0.457	58 (2.4)	0.89 (0.45-1.76)	0.739
243/274	2 (0.4)	14 (0.8)	0.73 (0.14-3.84)	0.712	7 (1.0)	0.30 (0.06-1.64)	0.166	21 (0.9)	0.47 (0.10-2.17)	0.332
274/302	16 (3.5)	41 (2.4)	1.70 (0.85-3.40)	0.134	21 (2.9)	1.01 (0.46-2.20)	0.977	62 (2.5)	1.33 (0.72-2.49)	0.365
243/333	3 (0.7)	16 (0.9)	0.68 (0.18-2.57)	0.565	1 (0.1)	5.71 (0.39-82.73)	0.202	17 (0.7)	0.99 (0.27-3.69)	0.989
Total	461 (100)	1733 (100)			735 (100)			2468 (100)		
Alleles										
302	629 (68.2)	2257 (65.1)	1		976 (66.4)	1		3233 (65.5)	1	
243	260 (28.2)	1103 (31.8)	0.85 (0.71-1.02)	0.083	442 (30.1)	0.90 (0.73-1.10)	0.297	1545 (31.3)	0.88 (0.74-1.04)	0.13
274	18 (2.0)	55 (1.6)	1.52 (0.82-2.85)	0.186	28 (1.9)	0.84 (0.42-1.67)	0.616	83 (1.7)	1.15 (0.65-2.02)	0.628
333	15 (1.6)	51 (1.5)	1.09 (0.56-2.09)	0.806	24 (1.6)	0.93 (0.45-1.93)	0.847	75 (1.5)	0.94 (0.52-1.71)	0.842
SL-Genotype		3466 (100)								
LL	228 (49.5)	775 (44.7)	1		340 (46.3)	1		1115 (45.2)	1	
SL	188 (40.8)	758 (43.7)	0.88 (0.69-1.13)	0.319	320 (43.5)	0.91 (0.69-1.20)	0.511	1078 (43.7)	0.91 (0.73-1.14)	0.409
SS	45 (9.8)	200 (11.5)	0.77 (0.52-1.15)	0.198	75 (10.2)	0.79 (0.50-1.24)	0.309	275 (11.1)	0.79 (0.55-1.14)	0.213
SL+SS	233 (50.5)	958 (55.3)	0.86 (0.68-1.08)	0.198	395 (53.7)	0.89 (0.68-1.15)	0.365	1353 (54.8)	0.88 (0.71-1.09)	0.258

a Odds ratio; b 95% confidence interval

3.2.2.2. Subpopulation ≥ 70 years

For the age group of 70 or above, the genotype 243/302 was found to be associated with a significant decreased lung cancer risk when cases were compared to the hospital based control population. This reduced risk is reflected by an OR of 0.61 as illustrated in Table 8. If cases were limited to NSCLC, the genotype 243/302, which contains the short VNTR-243, was found to be associated with a decreased risk of NSCLC, with an OR of 0.53 for the hospital based controls and an OR of 0.62 for the total control group. A significant protective influence, regarding the single VNTR-243, was also observed in each control group, with ORs ranging from 0.63 to 0.66. This effect remained significant in the SL group, that contains the short VNTR-243 (OR=0.59;0.65).The SL+SS group was also found to be significantly associated with an reduced NSCLC risk in all control populations, with ORs ranging from 0.58 to 0.64. Moreover, the genotype 302/333 conferred a significant decrease in NSCLC risk if cases were compared to the hospital based controls (OR= 0.15; Table 9)

Table 8. Genotype distribution of lung cancer study population ≥ 70

	Lung cancer patients	Population based Controls			Hospital based Controls			Total Controls		
Genotype	N (%)	N (%)	OR a (95% CI b)	P-value	N (%)	OR a (95% CI b)	P-value	N (%)	OR a (95% CI b)	P-value
302/302	79 (49.4)	186 (42.2)	1		97 (39.1)	1		283 (41.1)	1	
243/243	13 (8.1)	49 (11.1)	0.63 (0.30-1.35)	0.236	22 (8.9)	0.67 (0.30-1.54)	0.349	71 (10.3)	0.65 (0.33-1.30)	0.224
243/302	**58 (36.3)**	183 (41.5)	0.80 (0.50-1.26)	0.326	**107 (43.1)**	**0.61 (0.38-0.99)**	**0.045**	290 (42.1)	0.74 (0.49-1.10)	0.135
302/333	4 (2.5)	6 (1.4)	1.00 (0.23-4.25)	0.998	8 (3.2)	0.38 (0.09-1.58)	0.185	14 (2.0)	0.67 (0.20-2.25)	0.515
243/274	1 (0.6)	3 (0.7)	0.83 (0.04-17.5)	0.905	2 (0.8)	0.38 (0.03-4.58)	0.443	5 (0.7)	0.44 (0.04-4.61)	0.491
274/302	5 (3.1)	10 (2.3)	2.05 (0.54-7.70)	0.29	12 (4.8)	0.66 (0.20-2.15)	0.487	22 (3.2)	0.92 (0.32-2.71)	0.886
243/333	0 (0.0)	4 (0.9)	x	x	0 (0.0)	x	x	4 (0.6)	X	x
Total Alleles	160 (100)	441 (100)			248 (100)			689 (100)		
302	225 (70.3)	571 (64.7)	1		321 (64.7)	1		892 (64.7)	1	
243	85 (26.6)	288 (32.7)	0.76 (0.55-1.06)	0.107	153 (30.8)	0.74 (0.53-1.06)	0.097	441 (32.0)	0.77 (0.57-1.03)	0.079
274	6 (1.9)	13 (1.5)	1.85 (0.56-6.10)	0.315	14 (2.8)	0.71 (0.25-2.02)	0.519	27 (2.0)	0.88 (0.34-2.33)	0.804
333	4 (1.3)	10 (1.1)	0.67 (0.19-2.43)	0.542	8 (1.6)	0.46 (0.12-1.83)	0.27	18 (1.3)	0.61 (0.19-1.94)	0.403
SL-Genotype										
LL	83 (51.9)	192 (43.5)	1		105 (42.3)	1		297 (43.1)	1	
SL	63 (39.4)	197 (44.7)	0.81 (0.52-1.27)	0.363	119 (48.0)	0.65 (0.41-1.04)	0.07	316 (45.9)	0.75 (0.51-1.11)	0.154
SS	14 (8.8)	52 (11.8)	0.64 (0.30-1.34)	0.234	24 (9.7)	0.68 (0.31-1.49)	0.331	76 (11.0)	0.65 (0.33-1.26)	0.201
SL+SS	77 (48.1)	249 (56.5)	0.78 (0.51-1.18)	0.237	143 (57.7)	0.66 (0.42-1.02)	0.06	392 (56.9)	0.73 (0.51-1.06)	0.098

a Odds ratio; b 95% confidence interval

Table 9. Genotype distribution of the NSCLC study population ≥ 70

	NSCLC patients	Population based controls			Hospital based controls			Total Controls		
Genotype	N (%)	N (%)	OR a (95% CI b)	P-value	N (%)	OR a (95% CI b)	P-value	N(%)	OR a (95% CI b)	P-value
302/302	68 (54.4)	186 (42.2)	1		97 (39.1)	1		283 (41.1)	1	
243/243	9 (7.2)	49 (11.1)	0.45 (0.19-1.06)	0.068	22 (8.9)	0.54 (0.21-1.36)	0.189	71 (10.3)	0.50 (0.23-1.09)	0.08
243/302	**42 (33.6)**	183 (41.5)	0.66 (0.40-1.08)	0.1	**107 (43.1)**	**0.53 (0.31-0.89)**	**0.017**	**290 (42.1)**	**0.62 (0.40-0.97)**	**0.035**
302/333	**2 (1.6)**	6 (1.4)	0.58 (0.09-3.54)	0.554	**8 (3.2)**	**0.15 (0.02-0.97)**	**0.046**	14 (2.0)	0.36 (0.07-1.77)	0.211
243/274	0 (0.0)	3 (0.7)	x	x	2 (0.8)	x	x	5 (0.7)	x	x
274/302	4 (3.2)	10 (2.3)	1.87 (0.44-7.86)	0.394	12 (4.8)	0.60 (0.17-2.18)	0.44	22 (3.2)	0.84 (0.26-2.71)	0.768
243/333	0 (0.0)	4 (0.9)	x	x	0 (0.0)	x	x	4 (0.6)	x	x
Total	125 (100)	441 (100)			248 (100)			689 (100)		
Alleles										
302	184 (73.6)	571 (64.7)	1		321 (64.7)	1		892 (64.7)	1	
243	**60 (24.0)**	**288 (32.7)**	**0.63 (0.44-0.91)**	**0.014**	**153 (30.8)**	**0.65 (0.45-0.96)**	**0.03**	**441 (32.0)**	**0.66 (0.47-0.91)**	**0.012**
274	4 (1.6)	13 (1.5)	1.69 (0.44-6.45)	0.446	14 (2.8)	0.59 (0.18-1.98)	0.392	27 (2.0)	0.74 (0.24-2.28)	0.597
333	2 (0.8)	10 (1.1)	0.41 (0.08-2.14)	0.292	8 (1.6)	0.20 (0.03-1.19)	0.076	18 (1.3)	0.36 (0.08-1.65)	0.188
SL-Genotype										
LL	70 (56.0)	192 (43.5)	1		105 (42.3)	1		297 (43.1)	1	
SL	**46 (36.8)**	197 (44.7)	0.69 (0.43-1.12)	0.131	**119 (48.0)**	**0.59 (0.36-0.97)**	**0.037**	**316 (45.9)**	**0.65 (0.43-1.00)**	**0.052**
SS	9 (7.2)	52 (11.8)	0.45 (0.19-1.05)	0.064	24 (9.7)	0.52 (0.21-1.29)	0.157	76 (11.0)	0.48 (0.22-1.04)	0.062
SL+SS	**55 (44.0)**	**249 (56.5)**	**0.64 (0.40-1.00)**	**0.052**	**143 (57.7)**	**0.58 (0.36-0.93)**	**0.023**	**392 (56.9)**	**0.62 (0.41-0.93)**	**0.02**

a Odds ratio; b 95% confidence interval

3.2.2.3. Subpopulation ≤ 55 years

In the stratified analysis for the age group ≤ 55 years, several associations were identified and the genotype, allele and SL-genotype distribution for the lung cancer cases and the population based, hospital based and total control group are summarized in Table 9. The genotype 243/302 was found to be significantly associated with an increase in lung cancer risk when cases were compared to the hospital based controls. This elevated risk is illustrated by an OR of 2.58. The genotype 243/302 contains the short VNTR-243 that in turn is also associated with a 1.75fold increase in lung cancer risk within the same study population. The genotype containing the other short VNTR, genotype 274/302 was also found to be associated with an elevated lung cancer risk, reflected by an OR of 3.73, if the total control group is used, and an OR of 3.6, if controls are population based. Again these effects remained significant in the SL-classification system with ORs of 2.66, 2.46 and 1.57 for the SL and SS+SL genotypes, with either the control group containing hospital based controls or the total control group (Table 10).

Furthermore, the genotype 243/302 was found to be associated with a significant increase in NSCLC risk in a study population aged 55 or younger, if controls were hospital based (OR= 2.75). The enclosed VNTR-243 is also associated with a 1.75fold increase in NSCLC risk within the same study population. Genotype 274/302 is associated with a substantially elevated NSCLC risk, when either population based controls (OR= 4.54) or the total control group were considered (OR= 4.76). The SL-genotype, t includes the 243/302 genotype and the 274/302 genotype, was also associated with a raised risk of NSCLC, illustrated by an OR of 2.85 for the hospital based controls and 1.67 for the total controls. Moreover, the SL+SS group is associated with a 2.56fold increase in NSCLC risk, using the hospital based controls (Table 11).

Table 10. Genotype distribution of the lung cancer study population ≤ 55

	Lung cancer patients	Population based Controls			Hospital based Controls			Total Controls		
Genotype	N (%)	N (%)	OR a (95% CI b)	P-value	N (%)	OR a (95% CI b)	P-value	N (%)	OR a (95% CI b)	P-value
302/302	51 (38.1)	239 (43.1)	1		49 (55.7)	1		288 (44.8)	1	
243/243	12 (9.0)	72 (13.0)	0.70 (0.33-1.48)	0.349	8 (9.1)	2.15 (0.68-6.82)	0.193	80 (12.4)	0.87 (0.42-1.79)	0.703
243/302	**57 (42.5)**	205 (36.9)	1.26 (0.79-2.02)	0.338	**28 (31.8)**	**2.58 (1.29-5.14)**	**0.007**	233 (36.2)	1.49 (0.95-2.35)	0.083
302/333	4 (3.0)	16 (2.9)	1.39 (0.40-4.89)	0.606	1 (1.1)	2.94 (0.31-27.81)	0.347	17 (2.6)	1.58 (0.47-5.38)	0.463
243/274	1 (0.7)	7 (1.3)	0.67 (0.06-7.25)	0.744	1 (1.1)	0.68 (0.04-11.45)	0.786	8 (1.2)	0.74 (0.08-7.03)	0.792
274/302	**7 (5.2)**	**11 (2.0)**	**3.60 (1.08-12.06)**	**0.038**	1 (1.1)	5.58 (0.61-51.40)	0.129	**12 (1.9)**	**3.73 (1.19-11.69)**	**0.024**
243/333	2 (1.5)	5 (0.9)	1.85 (0.32-10.67)	0.492	0 (0.0)	x	x	5 (0.8)	2.38 (0.42-13.58)	0.328
Total	134 (100)	555 (100)			88 (100)			643 (100)		
Alleles										
302	170 (63.4)	710 (64.0)	1		128 (72.7)	1		838 (65.2)	1	
243	**84 (31.3)**	361 (32.5)	0.92 (0.67-1.27)	0.604	**45 (25.6)**	**1.75 (1.07-2.86)**	**0.025**	406 (31.6)	1.05 (0.77-1.43)	0.759
274	8 (3.0)	18 (1.6)	2.17 (0.79-6.01)	0.134	2 (1.1)	2.12 (0.43-10.59)	0.358	20 (1.6)	2.12 (0.82-5.52)	0.123
333	6 (2.2)	21 (1.9)	1.35 (0.49-3.70)	0.556	1 (0.6)	3.59 (0.42-30.56)	0.243	22 (1.7)	1.51 (0.56-4.05)	0.414
SL-Genotype										
LL	55 (41.0)	225 (42.9)	1		50 (56.8)	1		305 (47.4)	1	
SL	**66 (49.3)**	221 (42.1)	1.34 (0.86-2.10)	0.201	**29 (33.0)**	**2.66 (1.36-5.20)**	**0.004**	**250 (38.9)**	**1.57 (1.02-2.41)**	**0.041**
SS	13 (9.7)	79 (15.0)	0.68 (0.33-1.40)	0.298	9 (10.2)	1.79 (0.61-5.25)	0.292	88 (13.7)	0.83 (0.42-1.67)	0.607
SL+SS	**79 (59.0)**	300 (57.1)	1.16 (0.76-1.78)	0.491	**38 (43.2)**	**2.46 (1.31-4.60)**	**0.005**	338 (52.6)	1.37 (0.91-2.07)	0.132

a Odds ratio; b 95% confidence interval

Table 11. Genotype distribution of the NSCLC study population ≤ 55

	NSCLC patients	Population based Controls			Hospital based Controls			All Controls		
Genotype	N (%)	N (%)	OR a (95% CI b)	P-value	N (%)	OR a (95% CI b)	P-value	N (%)	OR a (95% CI b)	P-value
302/302	39 (35.8)	239 (43.1)	1		49 (55.7)	1		288 (44.8)	1	
243/243	10 (9.2)	72 (13.0)	0.74 (0.33-1.64)	0.454	8 (9.1)	2.11 (0.65-6.91)	0.216	80 (12.4)	0.92 (0.42-2.01)	0.837
243/302	**47 (43.1)**	205 (36.9)	1.32 (0.80-2.20)	0.281	**28 (31.8)**	**2.75 (1.34-5.63)**	**0.006**	233 (36.2)	1.58 (0.97-2.59)	0.067
302/333	4 (3.7)	16 (2.9)	1.76 (0.50-6.22)	0.377	1 (1.1)	3.84 (0.40-36.56)	0.242	17 (2.6)	2.02 (0.59-6.91)	0.261
243/274	0 (0.0)	7 (1.3)	x	x	1 (1.1)	x	x	8 (1.2)	X	x
274/302	**7 (6.4)**	**11 (2.0)**	**4.54 (1.36-15.19)**	**0.014**	1 (1.1)	7.07 (0.77-65.17)	0.084	**12 (1.9)**	**4.76 (1.52-14.95)**	**0.008**
243/333	2 (1.8)	5 (0.9)	2.35 (0.41-13.61)	0.34	0 (0.0)	x	x	5 (0.8)	3.05 (0.53-17-42)	0.21
Total	109 (100)	555 (100)			88 (100)			643 (100)		
Alleles										
302	136 (62.4)	710 (64.0)	1		128 (72.7)	1		838 (65.2)	1	
243	**69 (31.7)**	361 (32.5)	0.93 (0.66-1.31)	0.668	**45 (25.6)**	**1.75 (1.06-2.90)**	**0.03**	406 (31.6)	1.07 (0.76-1.49)	0.712
274	7 (3.2)	18 (1.6)	2.46 (0.87-7.00)	0.09	2 (1.1)	2.30 (0.45-11.82)	0.318	20 (1.6)	2.38 (0.89-6.41)	0.085
333	6 (2.8)	21 (1.9)	1.66 (0.61-4.52)	0.326	1 (0.6)	4.49 (0.53-38.33)	0.17	22 (1.7)	1.85 (0.69-4.97)	0.22
SL-Genotype										
LL	43 (39.4)	225 (42.9)	1		50 (56.8)	1		305 (47.4)	1	
SL	**56 (51.4)**	221 (42.1)	1.42 (0.88-2.30)	0.152	**29 (33.0)**	**2.85 (1.43-5.69)**	**0.003**	**250 (38.9)**	**1.67 (1.05-2.66)**	**0.03**
SS	10 (9.2)	79 (15.0)	0.66 (0.30-1.46)	0.305	9 (10.2)	1.63 (0.53-5.04)	0.397	88 (13.7)	0.81 (0.38-1.75)	0.591
SL+SS	**66 (60.6)**	300 (57.1)	1.21 (0.77-1.92)	0.406	**38 (43.2)**	**2.56 (1.34-4.91)**	**0.005**	338 (52.6)	1.44 (0.93-2.25)	0.106

a Odds ratio; b 95% confidence interval

3. Discussion

Telomerase is a complex holoenzyme, able to elongate telomere ends that have been shortened during successive cycles of cell division. Somatic cells lose about 25-200 telomeric nucleotides per cell division, unless telomeres reach a critical length and trigger replicative senescence or crisis. For immortal cells, the maintenance of telomeres length is essential and is provided by telomerase (Cong et al., 2002). Expression of hTERT gene, the limiting factor of telomerase activity, is usually absent in somatic cells, but present in the majority of cancer cells and immortalized cells (~90%) (Cong et al., 2002; Counter et al., 1995). Therefore, hTERT is one of the universal candidate genes in cancer research, and polymorphisms within hTERT gene or its regulatory elements may influence the personal risk of developing cancer. To date, two polymorphisms, located at the hTERT locus were found to modulate the risk of developing prostate cancer (Rafnar et al., 2009). For lung cancer 8 SNPs at 5p15.33 have been identified in several GWAS. Besides SNPs, some variable non-coding miniatellites like the VNTR MNS16A are located within the region of hTERT and may influence lung cancer risk (Baird, 2010). MNS16A is sited downstream of hTERT within a promoter region for an hTERT antisense transcript and comprises binding sites for transcription factor GATA1. Its number of tandem repeats was shown to modulate the promoter activity, alter the level of antisense transcript and therefore potentially alter the hTERT expression level. Several GWAS, conducted in diverse ethnicities, considered associations between the risk of developing cancer and MNS16A variants (Baird, 2010).

The present diploma thesis investigated associations between MNS16A in the context of prostate cancer and lung cancer in an Austrian population. Furthermore a novel MNS16A variant, VNTR-212, was identified.

MNS16A variants were for the first time investigated as a potential biomarker for the risk of developing prostate cancer. One major advantage of this study set up was, that the control group underwent biopsy of the prostate and was therefore known to be free of prostate cancer. Prostate cancer cases were histologically confirmed. Controls were diagnosed with BPH, an enlargement of the prostate, which is not considered as precursor to cancer (Kristal et al., 2010). Between 70% to 80% of men in their 7^{th} to 8^{th} life decade show indication of BPH, wherefore it is hardly achievable to find a male control population without prostate alterations, at least in this age group (Berry et al., 1984). Moreover this study encompassed a large number (1137) of prostate cancer patients, which is the third highest number of cases in all published sudies on MNS16A variants so far. Another positive issue of our study design

was the successful genotyping of MNS16A variants, with 98.1% genotyping efficiency of prostate cancer samples and 98.6% genotyping efficiency of lung cancer samples.

The initial aim of the study was to identify a genetic biomarker, which enables to differentiate between aggressive and indolent forms of prostate cancer. Currently the common treatment of advanced prostate cancer is radical prostatectomy, external beam radiation therapy (with and without hormonal therapy) and primary androgen deprivation (Ponholzer et al., 2011). However, the majority of prostate cancer cases are indolent and would not require radical treatment, which is often accompanied by harmful side effects (Tosoian and Loeb, 2010). Anyway, no significant associations were observed in the stratified analysis of patients with Gleason score 2-6 and 7-10 However, our study revealed a significant association between MNS16A VNTR-274 and a diminished risk to develop prostate cancer at least in men ≥70 years (OR=0.09). MNS16A VNTR-274 may therefore be utilized as a biomarker for prostate cancer risk assessment in men ≥70 years.

Another aim was the investigation of MNS16A alleles as a potential biomarker for the risk of developing lung cancer in our Caucasian study population. The lung cancer study population comprised 665 lung cancer cases, which is rather small compared to other MNS16A surveys. A strength of our control group was, that it comprised a population-based as well as a hospital-based subgroup. Three studies on MNS16A alleles and risk of lung cancer or lung cancer survival were published so far. The very first association study on MNS16A was made in the context of lung cancer by L. Wang et al. in a Texan study population (79% ≥ 60 years) in 2003. The Texan study was rather preliminary and suboptimal, with a small hospital-based study-population, comprising 53 NSCLC patients and 72 cancer-free controls. Moreover, Wang et al. isolated genomic DNA from blood cells as well as from the tumor samples of the NSCLC patients. Their investigations obtained no significant results, but a tendency towards the LL genotype as risk factor was observed (Wang et al., 2003). This trend is in accordance with the significant associations we observed for the NSCLC patients ≥70 years. This diploma thesis demonstrated, that, single VNTR-243, genotype 243/302 and SL and SS+SL genotypes were associated with a reduced risk of NSCLC in patients of 70 years and above. Wang et al. published another study on MNS16A alleles affecting the survival duration of patients with NSCLC in 2010. The hospital-based, Texan study population comprised 808 NSCLC patients. The aforementioned authors showed that the SS genotype was associated with a shorter survival time (OR=2.34) compared to the LL genotype among 221 NSCLC patients stage I and II. These findings were not evident in the overall study-population(Wang et al., 2010). So far, this thesis did not consider associations of MNS16A alleles and the survival of lung

cancer, but they may be involved in subsequent studies, for which this study is a heplful prerequisite. The most recent study on MNS16A variants and lung cancer risk was conducted in a Korean study-population in 2010 by Jin et al.. The study-population consisted of 937 lung cancer patients, 703 NSCLC patients and 943 healthy controls. They demonstrated that NSCLC patients with genotypes comprising at least one VNTR-243 were significantly associated with a better overall survival duration compared to genotypes without VNTR-243. On the other hand, single VNTR-243 and genotype 243/302 were identified as risk markers for lung cancer onset (Jin et al., 2011). In the current study, the single VNTR-243 and the genotype 243/302 were also found to be associated with an increased lung cancer risk, in subjects ≤55 years. Furthermore, this survey was able to demonstrated, that genotype 274/302 (containing the other short allele) and SL and SL+SS genotypes were associated with an increased lung cancer risk in patients ≤55 years. Jin et al.`s study population is rather young (61.7+- 9.2 years) and may therefore be comparable with our study population ≤55 years. Considering the similar age distribution, equal results are even more indicative. VNTR-243 may be a potential genetic biomarker for lung cancer risk at an early age, at least in Caucasian and Korean populations. Since the 5-year survival rate for lung cancer is poor and early detection of lung cancer is therefore crucial for a better prognosis (Petersen, 2011), the identification of genetic risk factors may help to develop an improved risk assessment.

Besides lung cancer, MNS16A variants were also investigated as a possible biomarker in the context of several other cancer sites. In 2007, a study on MNS16A variants, the occurrence of malignant gliomas and the survival time of glioma patients was published. The study population consisted of 352 Caucasian patients with glioblastoma or anaplastic gliomas and 302 volunteer controls. No associations between MNS16A genotypes and the survival time were observed, however the SS genotype was significantly overrepresented in patients, when the LL genotype was taken as reference (Carpentier et al., 2007). In 2008, results from a study on MNS16A, conducted in a Chinese study-population, comprising 1029 breast cancer cases and 1107 matching controls, were presented. A significantly elevated risk of developing breast cancer was observed for the genotypes 274/302, 243/302 and 243/243 when compared to the wild-type genotype. Furthermore, it was reported, that the genotype 274/302 was associated with a significant increase in axillary lymph nodes metastasis (Wang et al., 2008c). A survey in 2009 compared the MNS16A genotype distribution of 1391 glioma and meningnoma patients with 1359 population based controls and did not detect an association between MNS16A alleles and the risk of glioma. Anyways, a significantly better survival in

glioma patients, carrying the LL genotype, compared to patients carrying the SL or SS genotype was observed (Andersson et al., 2009). In 2011, our group conducted a study based on 90 colorectal cancer patients, 308 patients with high-risk polyps, 1022 patients with low-risk polyps and 1822 cancer- and polyp free controls (all Caucasians) and showed that the VNTR-274 significantly increased the risk of developing colorectal cancer, Furthermore, a novel MNS16A variant, VNTR-364, was identified (Hofer et al., 2011). In the course of the present study another novel MNS16A allele, the short VNTR-212, was identified. Also in 2011, a Chinese study on MNS16A variants and the risk of nasopharyngeal cancer, based on 855 Chinese patients and 1036 matching controls was published. It concluded, that subjects carrying the S-allele had a significantly reduced risk of nasopharyngeal cancer, compared to the LL-genotype carrier (Zhang et al., 2011). So far, studies on MNS16A as cancer biomarker yielded in opposing results, since cancer type and characteristics of the study populations varies among them.

To establish MNS16A variants as biomarker for prostate cancer and lung cancer, further confirmation in larger cohorts and various ethnicities are required.

5. References

Aberle, D. R., Adams, A. M., Berg, C. D., Black, W. C., Clapp, J. D., Fagerstrom, R. M., Gareen, I. F., Gatsonis, C., Marcus, P. M., and Sicks, J. D. (2011). Reduced lung-cancer mortality with low-dose computed tomographic screening. The New England journal of medicine 365, 395-409.

Amin, M., Khalid, A., Tazeen, N., and Yasoob, M. (2010). Zonal Anatomy of Prostate Modern Pathology.

Amos, C. I., Wu, X., Broderick, P., Gorlov, I. P., Gu, J., Eisen, T., Dong, Q., Zhang, Q., Gu, X., Vijayakrishnan, J., et al. (2008). Genome-wide association scan of tag SNPs identifies a susceptibility locus for lung cancer at 15q25.1. Nature genetics 40, 616-622.

Andersson, U., Osterman, P., Sjostrom, S., Johansen, C., Henriksson, R., Brannstrom, T., Broholm, H., Christensen, H. C., Ahlbom, A., Auvinen, A., et al. (2009). MNS16A minisatellite genotypes in relation to risk of glioma and meningioma and to glioblastoma outcome. International journal of cancer 125, 968-972.

Ayala, A. G., and Ro, J. Y. (2007). Prostatic intraepithelial neoplasia: recent advances. Archives of pathology & laboratory medicine 131, 1257-1266.

Baird, D. M. (2010). Variation at the TERT locus and predisposition for cancer. Expert reviews in molecular medicine 12, e16.

Bartsch, G., Frick, J., Ruegg, I., Bucher, M., Holliger, O., Oberholzer, M., and Rohr, H. P. (1979). Electron microscopic stereological analysis of the normal human prostate and of benign prostatic hyperplasia. The Journal of urology 122, 481-486.

Berry, S. J., Coffey, D. S., Walsh, P. C., and Ewing, L. L. (1984). The development of human benign prostatic hyperplasia with age. The Journal of urology 132, 474-479.

Birch, J. M., Alston, R. D., McNally, R. J., Evans, D. G., Kelsey, A. M., Harris, M., Eden, O. B., and Varley, J. M. (2001). Relative frequency and morphology of cancers in carriers of germline TP53 mutations. Oncogene 20, 4621-4628.

Bodmer, W. F., Bailey, C. J., Bodmer, J., Bussey, H. J., Ellis, A., Gorman, P., Lucibello, F. C., Murday, V. A., Rider, S. H., Scambler, P., and et al. (1987). Localization of the gene for familial adenomatous polyposis on chromosome 5. Nature 328, 614-616.

Bonkhoff, H., Wernert, N., Dhom, G., and Remberger, K. (1991). Basement membranes in fetal, adult normal, hyperplastic and neoplastic human prostate. Virchows Archiv 418, 375-381.

Brawer, M. K., Peehl, D. M., Stamey, T. A., and Bostwick, D. G. (1985). Keratin immunoreactivity in the benign and neoplastic human prostate. Cancer research 45, 3663-3667.

Brennan, P., Hainaut, P., and Boffetta, P. (2011). Genetics of lung-cancer susceptibility. The lancet oncology 12, 399-408.

Brennan, P., McKay, J., Moore, L., Zaridze, D., Mukeria, A., Szeszenia-Dabrowska, N., Lissowska, J., Rudnai, P., Fabianova, E., Mates, D., et al. (2007). Uncommon CHEK2 missense variant and reduced risk of tobacco-related cancers: case control study. Human molecular genetics 16, 1794-1801.

Bushman, W. (2009). Etiology, epidemiology, and natural history of benign prostatic hyperplasia. The Urologic clinics of North America 36, 403-415, v.

Carpentier, C., Lejeune, J., Gros, F., Everhard, S., Marie, Y., Kaloshi, G., Laigle-Donadey, F., Hoang-Xuan, K., Delattre, J. Y., and Sanson, M. (2007). Association of telomerase gene hTERT polymorphism and malignant gliomas. Journal of neuro-oncology 84, 249-253.

Catalona, W. J., Smith, D. S., Ratliff, T. L., Dodds, K. M., Coplen, D. E., Yuan, J. J., Petros, J. A., and Andriole, G. L. (1991). Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. The New England journal of medicine 324, 1156-1161.

Chanock, S. J., and Hunter, D. J. (2008). Genomics: when the smoke clears. Nature 452, 537-538.

Cong, Y. S., Wright, W. E., and Shay, J. W. (2002). Human telomerase and its regulation. Microbiol Mol Biol Rev 66, 407-425, table of contents.

Cooperberg, M. R., Lubeck, D. P., Mehta, S. S., and Carroll, P. R. (2003). Time trends in clinical risk stratification for prostate cancer: implications for outcomes (data from CaPSURE). The Journal of urology 170, S21-25; discussion S26-27.

Counter, C. M., Gupta, J., Harley, C. B., Leber, B., and Bacchetti, S. (1995). Telomerase activity in normal leukocytes and in hematologic malignancies. Blood 85, 2315-2320.

Cybulski, C., Gorski, B., Huzarski, T., Masojc, B., Mierzejewski, M., Debniak, T., Teodorczyk, U., Byrski, T., Gronwald, J., Matyjasik, J., et al. (2004). CHEK2 is a multiorgan cancer susceptibility gene. American journal of human genetics 75, 1131-1135.

DeMarini, D. M., Landi, S., Tian, D., Hanley, N. M., Li, X., Hu, F., Roop, B. C., Mass, M. J., Keohavong, P., Gao, W., et al. (2001). Lung tumor KRAS and TP53 mutations in nonsmokers reflect exposure to PAH-rich coal combustion emissions. Cancer research 61, 6679-6681.
Dhaene, K., Van Marck, E., and Parwaresch, R. (2000). Telomeres, telomerase and cancer: an up-date. Virchows Arch 437, 1-16.

Dong, C., and Hemminki, K. (2001). Modification of cancer risks in offspring by sibling and parental cancers from 2,112,616 nuclear families. International journal of cancer 92, 144-150.

Draisma, G., Boer, R., Otto, S. J., van der Cruijsen, I. W., Damhuis, R. A., Schroder, F. H., and de Koning, H. J. (2003). Lead times and overdetection due to prostate-specific antigen screening: estimates from the European Randomized Study of Screening for Prostate Cancer. Journal of the National Cancer Institute 95, 868-878.

Edwards, S. M., Kote-Jarai, Z., Meitz, J., Hamoudi, R., Hope, Q., Osin, P., Jackson, R., Southgate, C., Singh, R., Falconer, A., et al. (2003). Two percent of men with early-onset prostate cancer harbor germline mutations in the BRCA2 gene. American journal of human genetics 72, 1-12.

Etzioni, R. (2008). Statistical issues in the evaluation of screening and early detection modalities. Urologic oncology 26, 308-315.

Fallin, D., and Schork, N. J. (2000). Accuracy of haplotype frequency estimation for biallelic loci, via the expectation-maximization algorithm for unphased diploid genotype data. American journal of human genetics 67, 947-959.

Fanghänel J., P. F., Anderhuber F., Nitsch R. (2003). Waldeyer Anatomie des Menschen 17. Auflage).

Ferlay, J., Shin, H. R., Bray, F., Forman, D., Mathers, C., and Parkin, D. M. (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International journal of cancer 127, 2893-2917.

Fletcher, O., Easton, D., Anderson, K., Gilham, C., Jay, M., and Peto, J. (2004). Lifetime risks of common cancers among retinoblastoma survivors. Journal of the National Cancer Institute 96, 357-363.

Fletcher, O., and Houlston, R. S. (2010). Architecture of inherited susceptibility to common cancer. Nature reviews 10, 353-361.

Fournier, G., Valeri, A., Mangin, P., and Cussenot, O. (2004). [Prostate cancer. Epidemiology. Risk factors. Pathology]. Annales d'urologie 38, 187-206.

Gann, P. H., Hennekens, C. H., Ma, J., Longcope, C., and Stampfer, M. J. (1996). Prospective study of sex hormone levels and risk of prostate cancer. Journal of the National Cancer Institute 88, 1118-1126.

Garcia-Closas, M., Malats, N., Silverman, D., Dosemeci, M., Kogevinas, M., Hein, D. W., Tardon, A., Serra, C., Carrato, A., Garcia-Closas, R., et al. (2005). NAT2 slow acetylation, GSTM1 null genotype, and risk of bladder cancer: results from the Spanish Bladder Cancer Study and meta-analyses. Lancet 366, 649-659.

Gianfagna, F., Cugino, D., Santimone, I., and Iacoviello, L. (2012). From candidate gene to genome-wide association studies in cardiovascular disease. Thrombosis research 129, 320-324.

Gleason, D. F., and Mellinger, G. T. (1974). Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. The Journal of urology 111, 58-64.

Globocan/a (2008). a. Prostate Cancer Incidence, Mortality and Prevalence Worldwide in 2008. In.

Globocan/b (2008). b. Lung Cancer Incidence, Mortality and Prevalence Worldwide in 2008. In.

Goeckenjan, G., Sitter, H., Thomas, M., Branscheid, D., Flentje, M., Griesinger, F., Niederle, N., Stuschke, M., Blum, T., Deppermann, K. M., et al. (2011). [Prevention, diagnosis,

therapy, and follow-up of lung cancer. Interdisciplinary guideline of the German Respiratory Society and the German Cancer Society--abridged version]. Pneumologie (Stuttgart, Germany) 65, e51-75.

Gsur, A., Bernhofer, G., Hinteregger, S., Haidinger, G., Schatzl, G., Madersbacher, S., Marberger, M., Vutuc, C., and Micksche, M. (2000). A polymorphism in the CYP17 gene is associated with prostate cancer risk. International journal of cancer 87, 434-437.

Hall, J. M., Lee, M. K., Newman, B., Morrow, J. E., Anderson, L. A., Huey, B., and King, M. C. (1990). Linkage of early-onset familial breast cancer to chromosome 17q21. Science (New York, NY 250, 1684-1689.

Hammond, E. C., and Seidman, H. (1980). Smoking and cancer in the United States. Preventive medicine 9, 169-174.

HapMap (2003). The International HapMap Project. Nature 426, 789-796.

Hemminki, K. (2011). Familial risk and familial survival in prostate cancer. World journal of urology 30, 143-148.

Hindorff, L. A., Gillanders, E. M., and Manolio, T. A. (2011). Genetic architecture of cancer and other complex diseases: lessons learned and future directions. Carcinogenesis 32, 945-954.

Hofer, P., Baierl, A., Feik, E., Fuhrlinger, G., Leeb, G., Mach, K., Holzmann, K., Micksche, M., and Gsur, A. (2011). MNS16A tandem repeats minisatellite of human telomerase gene: a risk factor for colorectal cancer. Carcinogenesis 32, 866-871.

Hugosson, J., Carlsson, S., Aus, G., Bergdahl, S., Khatami, A., Lodding, P., Pihl, C. G., Stranne, J., Holmberg, E., and Lilja, H. (2010). Mortality results from the Goteborg randomised population-based prostate-cancer screening trial. The lancet oncology 11, 725-732.

Hung, R. J., McKay, J. D., Gaborieau, V., Boffetta, P., Hashibe, M., Zaridze, D., Mukeria, A., Szeszenia-Dabrowska, N., Lissowska, J., Rudnai, P., et al. (2008). A susceptibility locus for lung cancer maps to nicotinic acetylcholine receptor subunit genes on 15q25. Nature 452, 633-637.

Hwang, S. J., Cheng, L. S., Lozano, G., Amos, C. I., Gu, X., and Strong, L. C. (2003). Lung cancer risk in germline p53 mutation carriers: association between an inherited cancer predisposition, cigarette smoking, and cancer risk. Human genetics 113, 238-243.
Imreh, S., Klein, G., and Zabarovsky, E. R. (2003). Search for unknown tumor-antagonizing genes. Genes, chromosomes & cancer 38, 307-321.

Jemal, A., Ward, E., Wu, X., Martin, H. J., McLaughlin, C. C., and Thun, M. J. (2005). Geographic patterns of prostate cancer mortality and variations in access to medical care in the United States. Cancer Epidemiol Biomarkers Prev 14, 590-595.

Jin, G., Yoo, S. S., Cho, S., Jeon, H. S., Lee, W. K., Kang, H. G., Choi, Y. Y., Choi, J. E., Cha, S. I., Lee, E. B., et al. (2011). Dual roles of a variable number of tandem repeat polymorphism in the TERT gene in lung cancer. Cancer science 102, 144-149.

Kikuchi, T., and Carbone, D. P. (2007). Proteomics analysis in lung cancer: challenges and opportunities. Respirology (Carlton, Vic 12, 22-28.

Kim, Y. I. (2004). Folate and DNA methylation: a mechanistic link between folate deficiency and colorectal cancer? Cancer Epidemiol Biomarkers Prev 13, 511-519.

Knudsen, B. S., and Vasioukhin, V. (2010). Mechanisms of prostate cancer initiation and progression. Advances in cancer research 109, 1-50.

Kristal, A. R., Price, D. K., Till, C., Schenk, J. M., Neuhouser, M. L., Ockers, S., Lin, D. W., Thompson, I. M., and Figg, W. D. (2010). Androgen receptor CAG repeat length is not associated with the risk of incident symptomatic benign prostatic hyperplasia: results from the Prostate Cancer Prevention Trial. The Prostate 70, 584-590.

Kuriyama, M., Wang, M. C., Papsidero, L. D., Killian, C. S., Shimano, T., Valenzuela, L., Nishiura, T., Murphy, G. P., and Chu, T. M. (1980). Quantitation of prostate-specific antigen in serum by a sensitive enzyme immunoassay. Cancer research 40, 4658-4662.

Leitzmann, M. F., and Rohrmann, S. (2012). Risk factors for the onset of prostatic cancer: age, location, and behavioral correlates. Clinical epidemiology 4, 1-11.

Levy, S., Sutton, G., Ng, P. C., Feuk, L., Halpern, A. L., Walenz, B. P., Axelrod, N., Huang, J., Kirkness, E. F., Denisov, G., et al. (2007). The diploid genome sequence of an individual human. PLoS biology 5, e254.

Ley, T. J., Mardis, E. R., Ding, L., Fulton, B., McLellan, M. D., Chen, K., Dooling, D., Dunford-Shore, B. H., McGrath, S., Hickenbotham, M., et al. (2008). DNA sequencing of a cytogenetically normal acute myeloid leukaemia genome. Nature 456, 66-72.

Lichtenstein, P., Holm, N. V., Verkasalo, P. K., Iliadou, A., Kaprio, J., Koskenvuo, M., Pukkala, E., Skytthe, A., and Hemminki, K. (2000). Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland. The New England journal of medicine 343, 78-85.

Lindon, J. C., Holmes, E., Bollard, M. E., Stanley, E. G., and Nicholson, J. K. (2004). Metabonomics technologies and their applications in physiological monitoring, drug safety assessment and disease diagnosis. Biomarkers 9, 1-31.

Lorenzo Bermejo, J., and Hemminki, K. (2005). Familial lung cancer and aggregation of smoking habits: a simulation of the effect of shared environmental factors on the familial risk of cancer. Cancer Epidemiol Biomarkers Prev 14, 1738-1740.

Marberger, M., Barentsz, J., Emberton, M., Hugosson, J., Loeb, S., Klotz, L., Koch, M., Shariat, S. F., and Vickers, A. (2012). Novel approaches to improve prostate cancer diagnosis and management in early-stage disease. BJU international 109 Suppl 2, 1-7.

Mattson, M. E., Pollack, E. S., and Cullen, J. W. (1987). What are the odds that smoking will kill you? American journal of public health 77, 425-431.

Miyagawa, T., Nishida, N., Ohashi, J., Kimura, R., Fujimoto, A., Kawashima, M., Koike, A., Sasaki, T., Tanii, H., Otowa, T., et al. (2008). Appropriate data cleaning methods for genome-wide association study. Journal of human genetics 53, 886-893.

Nakagawa, H., Akamatsu, S., Takata, R., Takahashi, A., Kubo, M., and Nakamura, Y. (2012). Prostate cancer genomics, biology, and risk assessment through genome-wide association studies. Cancer science 103, 607-613.

Oldenburg, R. A., Meijers-Heijboer, H., Cornelisse, C. J., and Devilee, P. (2007). Genetic susceptibility for breast cancer: how many more genes to be found? Critical reviews in oncology/hematology 63, 125-149.

Olivier, M., Goldgar, D. E., Sodha, N., Ohgaki, H., Kleihues, P., Hainaut, P., and Eeles, R. A. (2003). Li-Fraumeni and related syndromes: correlation between tumor type, family structure, and TP53 genotype. Cancer research 63, 6643-6650.

Papsidero, L. D., Wang, M. C., Valenzuela, L. A., Murphy, G. P., and Chu, T. M. (1980). A prostate antigen in sera of prostatic cancer patients. Cancer research 40, 2428-2432.
Pasche, B., and Yi, N. (2010). Candidate gene association studies: successes and failures. Current opinion in genetics & development 20, 257-261.

Perera, F. P., and Weinstein, I. B. (1982). Molecular epidemiology and carcinogen-DNA adduct detection: new approaches to studies of human cancer causation. Journal of chronic diseases 35, 581-600.

Perera, F. P., and Weinstein, I. B. (2000). Molecular epidemiology: recent advances and future directions. Carcinogenesis 21, 517-524.

Petersen, I. (2011). The morphological and molecular diagnosis of lung cancer. Deutsches Arzteblatt international 108, 525-531.

Pirker R., F. M., Huber H. (1996). Klinische Onkologie).

Ponholzer, A., Steinbacher, F., Madersbacher, S., and Schramek, P. (2011). [Current treatment of locally advanced and metastatic prostate cancer]. Wiener medizinische Wochenschrift (1946) 161, 377-381.

Rafnar, T., Sulem, P., Stacey, S. N., Geller, F., Gudmundsson, J., Sigurdsson, A., Jakobsdottir, M., Helgadottir, H., Thorlacius, S., Aben, K. K., et al. (2009). Sequence variants at the TERT-CLPTM1L locus associate with many cancer types. Nature genetics 41, 221-227.
Risch, N. (2001). The genetic epidemiology of cancer: interpreting family and twin studies and their implications for molecular genetic approaches. Cancer Epidemiol Biomarkers Prev 10, 733-741.

Saccone, S. F., Hinrichs, A. L., Saccone, N. L., Chase, G. A., Konvicka, K., Madden, P. A., Breslau, N., Johnson, E. O., Hatsukami, D., Pomerleau, O., et al. (2007). Cholinergic nicotinic receptor genes implicated in a nicotine dependence association study targeting 348 candidate genes with 3713 SNPs. Human molecular genetics 16, 36-49.

Sakr, W. A., Grignon, D. J., Haas, G. P., Heilbrun, L. K., Pontes, J. E., and Crissman, J. D. (1996). Age and racial distribution of prostatic intraepithelial neoplasia. European urology 30, 138-144.

Schroder, F. H., Hugosson, J., Roobol, M. J., Tammela, T. L., Ciatto, S., Nelen, V., Kwiatkowski, M., Lujan, M., Lilja, H., Zappa, M., et al. (2009). Screening and prostate-cancer mortality in a randomized European study. The New England journal of medicine 360, 1320-1328.

Schwartz, A. G., Prysak, G. M., Bock, C. H., and Cote, M. L. (2007). The molecular epidemiology of lung cancer. Carcinogenesis 28, 507-518.

Shay, J. W., and Bacchetti, S. (1997). A survey of telomerase activity in human cancer. Eur J Cancer 33, 787-791.

Shendure, J., and Ji, H. (2008). Next-generation DNA sequencing. Nature biotechnology 26, 1135-1145.

Shi, X., Zhou, S., Wang, Z., Zhou, Z., and Wang, Z. (2008). CYP1A1 and GSTM1 polymorphisms and lung cancer risk in Chinese populations: a meta-analysis. Lung cancer (Amsterdam, Netherlands) 59, 155-163.

Simard, J., Dumont, M., Soucy, P., and Labrie, F. (2002). Perspective: prostate cancer susceptibility genes. Endocrinology 143, 2029-2040.

Statistik-Austria/a (2009). Krebsregister a. Prostata. In.

Statistik-Austria/b (2009). Krebsregister b. Lunge. In.

Thorgeirsson, T. E., Geller, F., Sulem, P., Rafnar, T., Wiste, A., Magnusson, K. P., Manolescu, A., Thorleifsson, G., Stefansson, H., Ingason, A., et al. (2008). A variant associated with nicotine dependence, lung cancer and peripheral arterial disease. Nature 452, 638-642.

Tosoian, J., and Loeb, S. (2010). PSA and beyond: the past, present, and future of investigative biomarkers for prostate cancer. TheScientificWorldJournal 10, 1919-1931.

Varghese, J. S., and Easton, D. F. (2010). Genome-wide association studies in common cancers--what have we learnt? Current opinion in genetics & development 20, 201-209.

Vineis, P., Anttila, S., Benhamou, S., Spinola, M., Hirvonen, A., Kiyohara, C., Garte, S. J., Puntoni, R., Rannug, A., Strange, R. C., and Taioli, E. (2007). Evidence of gene gene interactions in lung carcinogenesis in a large pooled analysis. Carcinogenesis 28, 1902-1905.

Vineis, P., and Perera, F. (2007). Molecular epidemiology and biomarkers in etiologic cancer research: the new in light of the old. Cancer Epidemiol Biomarkers Prev 16, 1954-1965.

Wang, J., Wang, W., Li, R., Li, Y., Tian, G., Goodman, L., Fan, W., Zhang, J., Li, J., Zhang, J., et al. (2008a). The diploid genome sequence of an Asian individual. Nature 456, 60-65.

Wang, L., Soria, J. C., Chang, Y. S., Lee, H. Y., Wei, Q., and Mao, L. (2003). Association of a functional tandem repeats in the downstream of human telomerase gene and lung cancer. Oncogene 22, 7123-7129.

Wang, L., Wang, L. E., Mao, L., Spitz, M. R., and Wei, Q. (2010). A functional variant of tandem repeats in human telomerase gene was associated with survival of patients with early stages of non-small cell lung cancer. Clin Cancer Res 16, 3779-3785.

Wang, L., Wei, Q., Wang, L. E., Aldape, K. D., Cao, Y., Okcu, M. F., Hess, K. R., El-Zein, R., Gilbert, M. R., Woo, S. Y., et al. (2006). Survival prediction in patients with glioblastoma multiforme by human telomerase genetic variation. J Clin Oncol 24, 1627-1632.

Wang, Y., Broderick, P., Webb, E., Wu, X., Vijayakrishnan, J., Matakidou, A., Qureshi, M., Dong, Q., Gu, X., Chen, W. V., et al. (2008b). Common 5p15.33 and 6p21.33 variants influence lung cancer risk. Nature genetics 40, 1407-1409.

Wang, Y., Hu, Z., Liang, J., Wang, Z., Tang, J., Wang, S., Wang, X., Qin, J., Wang, X., and Shen, H. (2008c). A tandem repeat of human telomerase reverse transcriptase (hTERT) and risk of breast cancer development and metastasis in Chinese women. Carcinogenesis 29, 1197-1201.

Wheeler, D. A., Srinivasan, M., Egholm, M., Shen, Y., Chen, L., McGuire, A., He, W., Chen, Y. J., Makhijani, V., Roth, G. T., et al. (2008). The complete genome of an individual by massively parallel DNA sequencing. Nature 452, 872-876.

Witte, J. S. (2009). Prostate cancer genomics: towards a new understanding. Nat Rev Genet 10, 77-82.

Wood, D. E., Eapen, G. A., Ettinger, D. S., Hou, L., Jackman, D., Kazerooni, E., Klippenstein, D., Lackner, R. P., Leard, L., Leung, A. N., et al. (2012). Lung cancer screening. J Natl Compr Canc Netw 10, 240-265.

Wooster, R., Neuhausen, S. L., Mangion, J., Quirk, Y., Ford, D., Collins, N., Nguyen, K., Seal, S., Tran, T., Averill, D., and et al. (1994). Localization of a breast cancer susceptibility gene, BRCA2, to chromosome 13q12-13. Science (New York, NY 265, 2088-2090.

You, M., Wang, D., Liu, P., Vikis, H., James, M., Lu, Y., Wang, Y., Wang, M., Chen, Q., Jia, D., et al. (2009). Fine mapping of chromosome 6q23-25 region in familial lung cancer families reveals RGS17 as a likely candidate gene. Clin Cancer Res 15, 2666-2674.

Zhang, Y., Zhang, H., Zhai, Y., Wang, Z., Ma, F., Wang, H., Li, P., Zhang, Y., Yu, L., Cui, Y., et al. (2011). A functional tandem-repeats polymorphism in the downstream of TERT is associated with the risk of nasopharyngeal carcinoma in Chinese population. BMC medicine 9, 106.

Ziegler, A., Konig, I. R., and Thompson, J. R. (2008). Biostatistical aspects of genome-wide association studies. Biometrical journal 50, 8-28.

6. Curriculum Vitae

Name: Julia-Stephanie Gertraud Zerelles
Adresse: Pfeilgasse 28/24
1080 Wien
julia.zerelles@gmx.de

Persönliche Daten

Geburtstag: 24. Juni 1985

Geburtsort: Pfarrkirchen

Staatsbürgerschaft: Deutschland

Ausbildung

1991-95 Grundschule Triftern (Deutschland)

1995-96 Hauptschule Triftern

1997-2003 Gymnasium Pfarrkirchen

2003-05 Gymnasium Simbach (Abschluss Abitur);Leistungskurs Biologie/Deutsch

2005-06 Auslandsreise nach Thailand und Australien

2006-08 Diplomstudium Biologie

Seit 2008 Diplomstudium Anthropologie (Fokus Humangenetik)

2011-12 Diplomarbeit am Institut für Krebsforschung

Kenntnisse:
Windows Office, Excel, Word, Labor Erfahrungen, verschiedene Praktika etc.

Sprachen:
Englisch Schrift und Sprache, Latein