universität wien

DIPLOMARBEIT

Titel der Diplomarbeit

Association of AURKA polymorphisms with prostate cancer risk

angestrebter akademischer Grad

Magistra der Naturwissenschaften (Mag. rer. nat.)

Verfasserin:	Doris Hummel
Studienrichtung/Studienzweig (lt. Studienblatt):	Anthropologie
Betreuerin:	Ao. Univ.-Prof. Mag. Dr. Andrea Gsur

Wien, im Mai 2010

Danksagung

Ich möchte mich bei Herrn Univ.-Prof. Dr. Michael Micksche für die Möglichkeit bedanken, meine Diplomarbeit am Institut für Krebsforschung der Medizinischen Universität zu schreiben.

Ein herzlicher Dank gilt meiner Betreuerin Frau Ao. Univ.-Prof. Mag. Dr. Andrea Gsur (Arbeitsgruppe Molekulare Epidemiologie, Institut für Krebsforschung, Medizinische Universität Wien) für ihre kompetente Unterstützung.

Weiters bedanke ich mich bei Dr. Andreas Baierl (Institut für Statistik und Decision Support Systems) für die statistische Auswertung der Daten.

Außerdem möchte ich mich bei der gesamten Arbeitsgruppe für die Unterstützung und das angenehme Arbeitsklima sehr herzlich bedanken.

Table of contents

Abbreviations

AFB_1	Aflatoxin B_1
BPH	Benign prostatic hyperplasia
BTAK	Breast tumor-amplified kinase
CI	Confidence interval
CNV	Copy number variation
DNA	Desoxyribonucleic acid
DRE	Digital rectal examination
FRET	Fluorescence resonance energy transfer
hARK1	Aurora related kinase
HWE	Hardy-Weinberg equilibrium
ISUP	International Society of Urological Pathology
LD	Linkage disequilibrium
MAF	Minor allele frequency
MGB	Minor groove binder
NFQ	Nonfluorescent quencher
NTC	No template control
OR	Odds ratio
PAHs	Polycyclic aromatic hydrocarbons
PIN	Prostatic intraepithelial neoplasia
PSA	Prostate specific antigen
R_n	Normalized reporter
RNA	Ribonucleic acid
SD	Standard deviation
SDS	Sequence detection software
SNP	Single nucleotide polymorphism
STK 15	Serine/threonine kinase 15
STR	Short tandem repeats
TNM	Tumor-node-metastasis
TRUS	Transrectal ultrasound
T_m	Melting temperature
tSNP	Tagging SNP
VACURG	Veterans Administration Cooperative Urological Research Group

List of tables and figures

Tables

Figures

1. Introduction

1.1. Anatomy of the prostate gland

The prostate gland consists of three types of cells. The secretory or epithelial cells produce a wide range of substances, mostly enzymes, which help in the mobility and feeding of the sperm. The stroma is composed of smooth muscle cells and fibroblasts. The prostate has a blood supply and somatic, parasympathetic and sympathetic nerve fibres. The perineural spaces around these nerve fibres are unique to the prostate gland and cancer cells can migrate there.

The exocrine secretion of the prostate gland consists of many different enzymes and proteins. Some are concerned with the liquefaction of the sperm, while others facilitate the penetration of the sperm into the zona pellucida of the egg.

Circulating hormones produced by the testicle and the adrenal glands influence the prostate. Luteinizing-hormone releasing hormone, growth hormone and prolactin may also have influence on the prostate gland.

Through a simple diffusion, androgens and other steroid hormones enter the prostate epithelial cell. This event activates a pathway and results to the stimulation of cell proliferation, division, growth and secretion of the prostate gland.

The most primitive cell in the prostate gland is a stem cell which, - under the influence of androgens - develops into an intermediate cell type and finally into a mature epithelial cell. These mature cells are strongly androgen-dependent, a lack of androgen causes apoptosis. This event is important for the maintenance and the integrity of organs. A balance between cellular proliferation and apoptosis is required for the healthy growth of the prostate gland (Newling 1999).

1.2. Prostate cancer

1.2.1. Epidemiology

Prostate cancer is the most common malign neoplasm in male Austrians, followed by lung cancer and colorectal cancer (Statistik Austria a). In 2006, 26% of the diagnosed tumors of men were located in the prostate (Statistik Austria b). In 2008, prostate cancer ranked third in cancer mortality, exceed by lung and colorectal cancer. 1184 Austrian men died that year from this disease (Statistik Austria c).

The incidence of prostate cancer is intensely increasing since 1993. One reason for this development is the prostate specific antigen (PSA) screening, which started in Austria for healthy men in the beginning of the 1990s. Nowadays, half of the carcinomas are diagnosed at an early stage, even in asymptomatic men, because of PSA-screening. However, the mortality rate stays nearly the same (Hackl 2004; Vutuc et al. 2005).

Another reason for the increase of the incidence is a change in the age structure towards higher proportions of older people (Alexander 1999) because an established risk factor for prostate cancer is advanced age, besides ethnic origin, country of residence and family history (Alexander 1999; Gallagher and Fleshner 1998; Johns and Houlston 2003; Hackl 2004). The frequency of this disease rises extremely from 55 years of age onwards. One out of 400 men aged between 55 and 65 suffers from a diagnosed prostate cancer (Hackl 2004).

The difference in incidence between US-American men and Chinese men demonstrates the importance of the ethnic origin. The incidence among US-American men is 90-fold higher than among Chinese men, which hold the lowest incidence and mortality rate together with the Japanese population.

The highest incidence and mortality rate in the world is among the black population in the USA, which is about 35% higher than in whites (Parkin, Bray, and Devesa 2001; Alexander 1999).

The rates of Japanese immigrants to the US are intermediate between their original and host communities. They rise with the number of generations spent in the US

(Alexander 1999), but the rates reach only half of those for white men living in the same areas (Gallagher and Fleshner 1998).

1.2.2. Diagnosis, PIN, grading systems and therapy

The methods to diagnose prostate cancer are: elevated serum PSA level, digital rectal examination, transrectal ultrasound imaging and biopsy. However, an elevated PSA level is not necessarily a sign of prostate cancer. The increase may be caused by other circumstances (Selley et al. 1997).

As a histological precursor of cancer, prostatic intraepithelial neoplasia (PIN) is verified. It can be divided into low-grade and high-grade (the pre-invasive stage of invasive carcinoma) PIN. It mostly starts in the third decade of life and turns into a carcinoma ten or more years later. It is characterized by cytological and architectural abnormalities.
Most prostatic carcinomas are located in the peripheral zone of the prostate. This is also the hotspot for PIN. In high-grade PIN, the basal layer cells are disrupted (Bostwick 1999).

The tumor tissue from a biopsy or a prostatectomy can be classified with the Gleason grading system, which is a numeric system used for prostate cancer grading based on a study of the Veterans Administration Cooperative Urological Research Group (VACURG). It depends only on the architectural characteristics of the tumor (Bostwick 1999; Köllermann and Sauter 2009; Gandellini, Folini, and Zaffaroni 2009). The malignancy correlates significantly with the histological pattern of prostate cancer.
Gleason found in 1966 that more than 50% of the carcinomas contained two or more patterns and therefore showed heterogeneity. The determinant factor is the degree of glandular differentiation (see Figure 1), reflecting tumor heterogeneity by assigning a primary pattern for the dominant (or most prevalent) grade and a secondary pattern (at least 5% of the tumor) for the non-dominant grade. The histological score is derived by adding these two patterns.
Two is the lowest possible Gleason score. The primary and the secondary pattern Gleason grade is one (well-differentiated). When both grades are 5 (undifferentiated),

the highest score of 10 is the result. The grade of the first pattern is doubled if no secondary pattern exists. For example, the total Gleason score is 6 (3+3) when more than 95% of a tumor show Gleason grade 3.

Figure 1. Original Gleason System

Because only a small part of the tumor can be examined with a biopsy, the Gleason score in a prostatectomy preparation is often higher than the score obtained in the biopsy.

The Gleason grade can predict the biologic behaviour of a prostate cancer very well, including invasiveness and metastatic potential (Bostwick 1999; Epstein 2009; Gandellini, Folini, and Zaffaroni 2009; Gleason and Mellinger 1974; Mellinger, Gleason, and Bailar 1967).

Many new methods have extended the knowledge about prostate cancer since the introduction of the Gleason grading system in 1966. Examples are the prostate

specific antigen testing, the transrectal ultrasound-guided prostate needle biopsy and the immunohistochemistry for basal cells. These achievements led to a change in the classification of prostate cancer and to the detection of new prostate cancer variants.
The Gleason grading system was updated in 2005 at a consensus conference of international experts in urological pathology, under the auspices of the International Society of Urological Pathology (ISUP; Epstein 2009).

The second grading system that is often used is the TNM classification for cancer in general. It describes the anatomical extent of a cancer and is based on the determination of three components. T is the abbreviation for tumor. The subsequent number describes the extent of the primary tumor. TX implies that a primary tumor cannot be assessed and T0 indicates no evidence of a primary tumor.
In T1 classification level, the tumor is not clinically apparent (neither palpable nor visible by imaging). Further subdivisions (a-c) describe how the tumor got detected (a: incidental tumor histological finding in 5% or less of resected tissue; b: incidental tumor histological finding in more than 5% of resected tissue; c: tumor diagnosed through a needle biopsy, for example because of an elevated PSA-level).
T2 represents a tumor which is confined within the prostate. The subdivisions (a-c) refer to the extent of the tumor (a: tumor involves one half of one lobe or less, b: tumor involves more than half of one lobe, but not both lobes, c: tumor involves both lobes).
In T3 classification level, the tumor extends through the prostatic capsule. T3a indicates an extracapsular extension (uni- or bilateral), in T3b the tumor infiltrates the seminal vesicles.
The T4 classification indicates that the tumor is fixed or has invaded adjacent structures other than seminal vesicles: bladder neck, external sphincter, rectum and/or levator muscles and/or is fixed to the pelvic wall.
N is the abbreviation for node (regional lymph nodes). NX implies that the regional lymph nodes cannot be assessed. N0 indicates that there are no regional lymph node metastases and N1 represents the fact that regional lymph node metastases exist.
Distant metastases are abbreviated with M. If distant metastases cannot be assessed, it is classified as MX. M0 describes the absence of distant metastases and

M1 describes the presence. M1a classifies non-regional lymph node metastases, M1b osseous and M1c other metastases.
The use of the C-factor is optional and reflects the validity of classification according to the diagnostic methods employed.
This classification method is not very reliable due to low sensitivity and low specificity (Bostwick 1999; Wittekind, Sobin, and Klimpfinger 2005; Malayeri and Steger 1996).

There are three kinds of therapy approaches for prostate cancer: radical prostatectomy, radiotherapy and anti-androgen therapy (Selley et al. 1997; Chen, Clegg, and Scher 2009).
The treatment for patients with hormone-sensitive metastatic disease consists of the depletion of testosterone concentration or orchiectomy, often in combination with an anti-androgen, which keeps the endogenous androgen from docking to the receptor (Chen, Clegg, and Scher 2009).
Another treatment option suitable especially for older men with low grade disease is conservative management (Selley et al. 1997). It is also known as expectant management or “watchful waiting” and implies no active treatment until the patient shows symptoms from outlet obstruction or metastatic disease. Patients are checked up regularly to assess disease development and to discuss treatment options (Chodak 1994).

1.3. Molecular epidemiology

The term „molecular cancer epidemiology“ was defined by Perera and Weinstein in 1982 as an approach in cancer research, in which advanced laboratory methods are used in combination with analytical epidemiology to determine at the biochemical or molecular level specific exogenous agents or host factors playing a role in human cancer pathogenesis (Perera and Weinstein 1982).
Molecular epidemiology can facilitate cancer prevention through the early identification of “at risk” populations with a higher susceptibility and through the improvement of the exposure assessment. A wide range of biomarkers should be incorporated into epidemiologic studies to assess individual exposure, dose, preclinical effects and susceptibility to carcinogens (Perera and Weinstein 2000).

One of the achievements of molecular epidemiology is the insight it has provided into inter-individual variations in human cancer risk and the extensive interactions between environmental factors and inherited or acquired susceptibility factors in the development of cancer (Perera 1997).

The biomarkers have been generally categorized into four categories: (1) internal dose, (2) biologically effective dose, (3) preclinical biologic effects and (4) biomarkers of susceptibility (Perera and Weinstein 1982; Perera and Weinstein 2000).
Biomarkers of internal dose consider individual differences in absorption, metabolism, bioaccumulation or excretion of a substance or its metabolites and indicate the actual level within the body or specific tissues. Examples are: cotinine, a metabolite of nicotine in serum or urine, resulting from cigarette smoke exposure; and urinary levels of 1-hydroxypyrene deriving from exposure to polycyclic aromatic hydrocarbons (PAHs) or aflatoxin B_1 (AFB_1) levels in urine from reflecting dietary sources.
Markers of internal dose do not indicate the amount to which a given compound has interacted with critical cellular targets. To solve this problem, assays have been developed to measure the biologically effective dose of a substance, i. e. the amount of activated substance that has actually reacted with the critical cellular macromolecules, for example DNA, RNA or proteins in the blood.
The biomarker of early or preclinical biologic effects resulting from exposure reflects following events in the multi-step progress of carcinogenesis. It is measured directly in target tissues or for example in white blood cells. Representative biomarkers determine various types of genotoxicity, including chromosomal aberrations, small deletions and point mutations. For example, mutations of the p53 tumor suppressor gene or the ras oncogene are common events in human cancerogenesis.
The fourth and last category of biomarkers is that of inherited or acquired variations in host susceptibility, modulating the individual response to environmental carcinogens. Many different factors, in addition to exposure to a carcinogen, influence the probability that a tumor will develop in a certain individual. Examples of susceptibility factors are interindividual differences in the metabolism of carcinogenic chemicals, DNA repair or nutritional, hormonal or immunologic factors (Perera and Weinstein 2000). Genetic factors that influence cancerogenesis are extremely variable. Some mutations are rare, dominant and highly penetrant. Examples for

cancer developments affected by such a mutation are Retinoblastoma or Wilms' tumors. More common genetic traits are often in combination with a low individual risk and influence like the metabolic activation or detoxification of carcinogenic chemicals (Perera 1997). The detection of these low-penetrance genes is important for cancers that are associated with lifestyle factors like tobacco, alcohol or diet (Kotnis, Sarin, and Mulherkar 2005).

Latest revolutionary events for the practice of molecular epidemiology were the completion of the Human Genome Project and the HapMap Project, together with advances in high-throughput genotyping ability.
The focus of molecular epidemiology in the future will lie on systems biology (integration of interactions in biological systems from diverse experimental sources); DNA, RNA or tissue-based microarrays; the investigation of epigenetic mechanisms and the role of MicroRNAs in cancer risk.
The basic questions in the science of molecular epidemiology stay the same: Why did this person get this disease at this particular time? Was it preventable? Was it predictable (Spitz and Bondy 2010)?

1.3.1. Single nucleotide polymorphism

A single nucleotide polymorphism (SNP) is a stable inherited substitution of a single nucleotide. It has a minor allele frequency (MAF) of 1% or more in at least one population. Compared to polymorphisms like insertion or deletion of a nucleotide, a SNP is the most common sequence variation in the human genome (Kruglyak and Nickerson 2001; Risch 2000; Knippers 1995). Other polymorphisms are variable numbers of short tandem repeats (STR, 2-6 bp long) or tandem repeats. The variability in the number of a whole functional gene unit (exon, gene) is called CNV - copy number variation. This term also includes deletion (Kubistova, Mrazek, and Petrek 2009).
Extrapolation of the existing data leads to a total number of over 15 million SNPs in the human genome, even more when many populations and their specific
SNPs are included (Salisbury et al. 2003). SNPs are spread throughout the whole genome and can be found at least, on average, every 0.3–1 kilobases (kb; Schork, Fallin, and Lanchbury 2000).

While most SNPs are silent and do not change the function or expression of a gene, an estimated number of 50,000 to 250,000 of them will actually confer small to moderate biological effects (Chanock 2001). A SNP in a coding region may directly influence a relevant protein while a SNP in a promotor region can impact gene expression (Krawczak, Reiss, and Cooper 1992; Drazen et al. 1999). SNPs with functional consequences are expected to have lower allele frequencies. In fact, the majority of SNPs in a coding region causing an amino acid change have an allele frequency of below 5% (Kruglyak and Nickerson 2001).

Persons who carry a particular allele at one site often predictably carry specific alleles at other proximate variant sites, known as linkage disequilibrium (LD). Blocks of SNPs in the same chromosomal region are not inherited randomly, rather as a combination of polymorphic alleles, known as haplotype blocks. Evaluation of markers inherited on a haplotype reduces the complexity of analysing SNPs in a gene or locus (International HapMap Consortium 2005; Erichsen and Chanock 2004).

This block-like structure of LD and the existence of areas of low or high recombination rate lead to the identification of tagging SNPs (tSNPs), which can be used to predict with high probability the alleles at other co-segregating SNPs (Beckmann, Estivill, and Antonarakis 2007; International HapMap Consortium 2005).

The continuing discovery of SNPs and the characterization of haplotype blocks in human populations are having an elementary impact on the molecular epidemiology of cancer. It represents a notable tool to investigate inter-individual differences in treatment responses and outcomes of specific cancers. Genetic variation studies can help reveal critical determinants in environmental exposure and cancer, which in the future can help to develop strategies for prevention and early intervention (Erichsen and Chanock 2004; Rebbeck et al. 2004).

1.3.2. Molecular epidemiology of prostate cancer

The suspicion that genetics play a key role in prostate cancer has been verified over the past twenty years due to family-based linkage studies, twin studies and molecular epidemiologic studies. It was found that prostate cancer has a much more complex genetic basis than initially supposed (Schaid 2004).

Familiy-based linkage studies confirmed familiy history as a well-established risk factor for the development of prostate cancer. A man who has a first-degree family member suffering from prostate cancer has 2.5 times higher risk developing the cancer himself than a man without family history.
The highest risk exists, if more than one first-degree family member is affected and the cancer was diagnosed before 60 years of age (Johns and Houlston 2003).
If at all, rare dominant susceptibility genes with high-penetrance account only for a small fraction of the observed genetic predisposition to prostate cancer. Linkage studies raise the possibility that prostate cancer is more likely influenced by a combination of common genetic polymorphisms of low-penetrance.
The pattern of this inheritance follows Mendelian segregation as an autosomal dominant susceptibility trait (Shand and Gelmann 2006).
Several prostate cancer susceptibility loci like HPC1, HPCX, PCAP and HPC20 have been identified. The AURKA gene (which encodes for AURKA) is located in the HPC20 prostate cancer susceptibility locus at 20q13 (Berry et al. 2000).

1.3.3. Aurora kinase A

The Aurora kinase family (A, B and C) consists of highly conserved serine/threonine protein kinases. They are involved in several mitotic events: regulation of spindle assembly checkpoint pathway, function of centrosomes and cytoskeleton, and cytokinesis. An aberrant expression of these kinases may lead to the development of cancer (Kollareddy et al. 2008).
Aurora kinase A is also known as AURKA, aurora2, STK 15 (serine/threonine kinase 15), BTAK (breast tumor-amplified kinase) or hARK1 (aurora related kinase; Feik et al. 2009; Zhou et al. 1998; Sakakura et al. 2001). It consists of 403 amino acids and has a molecular weight of approximately 46 kD (Zhou et al. 1998).
Aurora kinase A is located in metaphase and anaphase of the mitosis at the centrosome, spindle poles and the spindle itself. In telophase, it is located primarily at the spindle poles (Bischoff et al. 1998). It is required for multiple crucial events during mitosis, such as mitotic entry, separation of centriole pairs, accurate bipolar spindle assembly, alignment of chromosomes to the metaphase plate and completion of cytokinesis (Marumoto et al. 2003). Defects in these processes can lead to centrosome abnormalities (larger diameter, elevated number and abnormal structure)

and aneuploidy, which contributes to genomic instability. In the vast majority of metastatic and invasive prostate carcinomas, centrosomes are structurally and numerically abnormal. These defects increase relatively with Gleason grade and with genomic instability (Pihan et al. 2001).

Aurora kinase A is over-expressed in prostate tumors and in PINs. Overexpression in primary clinical prostate cancers correlates with tumorigenicity, clinical staging, surgical margin status and seminal vesicle invasion. A targeted inhibition of Aurora kinase A represents a possible approach toward prostate cancer therapy (Lee et al. 2006).

AURKA was also found to be over-expressed in breast tumors (Zhou et al. 1998), in human tumor cell lines of lung, colon and melanoma origin (Bischoff et al. 1998), in pancreatic tumors and cell lines, and in gastric cancers (Sakakura et al. 2001).

1.4. Aim of the study

The hospital based case-control study “Molecular Epidemiology of Prostate Cancer” was arranged at the Division of Applied and Experimental Oncology (Institute of Cancer Research, KIM1, Medical University of Vienna) in cooperation with the Departments of Urology of the Medical University of Vienna, Sozialmedizinisches Zentrum Ost and Sozialmedizinisches Zentrum Süd. The study population of this ongoing study consists of 550 BPH controls and 1027 prostate cancer cases recruited from these three hospitals in Vienna.

The aim of the present study was to explore whether any of the four genotyped SNPs (rs2180691, rs8117896, rs1468055 and rs1476394) within the AURKA gene shows an association with prostate cancer risk.

The ultimate goal is to create a polygenic model of prostate cancer susceptibility to identify high-risk individuals. The identification of these persons may furnish a basis for individualized therapy and chemoprevention.

2. Material and Methods

2.1. Study population

550 age-matched control patients and 1027 patients with histologically verified and previously untreated prostate cancer participated in this study.
The controls suffered from lower urinary tract symptoms due to benign prostatic hyperplasia (BPH). In the control group, prostate cancer was excluded either clinically by negative digital rectal examination (DRE) and negative serum PSA or histologically by transrectal ultrasound (TRUS) guided biopsies or transurethral resection of the prostate. BPH is a very common non-malignant alteration of the prostate and is not a precursor of prostate cancer. The case group consisted of patients with newly diagnosed, histologically verified prostate cancer. Prostate carcinoma was diagnosed by TRUS-guided biopsies in all case patients. Suspicious findings on DRE and/or an elevated serum PSA-level were indications for prostate cancer.
The blood for DNA extraction was collected from the study participants before any other treatments.
A written consent was obtained from each participant and the study was approved by the institutional review boards.

2.2. Genomic DNA isolation

Human genomic DNA was purified according to the QIAamp® Blood Midi Kit Spin control (QIAGEN, Hilden, Germany). 2 ml of each whole blood sample were added to 200 µl QIAGEN Protease respectively and mixed briefly. 2.4 ml lysis Buffer AL were added and this mixture was incubated in the water bath at 70°C for about thirty minutes. After adding 2 ml ethanol for molecular biology (99.8%) and mixing to obtain a homogenous solution, half of the solution was transferred onto the QIAamp Midi column and centrifuged at 1850 x g for three minutes. This step was repeated with the second half of the solution. The DNA bound to the column was washed with 2 ml Buffer AW1 (centrifuged at 4500 x g for two minutes) and 2 ml Buffer AW2 (centrifuged at 4500 x g for eighteen minutes). The purified DNA was eluted with 300

µL Elution Buffer AE, incubated at room temperature for five minutes and then centrifuged at 4500 x g for seven minutes.
DNA concentration was measured spectrophotometrically at a wavelength of 230 nm using the NanoDrop ND-1000 Sprectrophotometer (PEQLAB Biotechnologie GMBH, Erlangen, Germany). DNA samples were diluted with AE Buffer to a working dilution of 10 ng/µl and stored at –80°C.

2.3. SNP selection

The four single nucleotide polymorphisms rs2180691, rs8117896, rs1468055 and rs1476394 are tagging SNPs picked out from the HapMap database (CEU population; www.hapmap.org).

2.4. Genotyping of AURKA polymorphisms

Genotyping was conducted on a 7500 Fast Real Time PCR System (Applied Biosystems, California, US) using standard TaqMan® SNP Genotyping Assays and TaqMan® Genotyping PCR Master Mix (containing all components except primers and probes).
The TaqMan® pre-designed SNP Genotyping Assay (Applied Biosystems, California, US) was performed using MicroAmp® Fast Optical 96-Well Reaction Plates (Applied Biosystems, California, US) optimized for 10 µl reactions.
The genotyping reaction for 10 µl volume mix contained 20 ng DNA (Table 1).

Table 1. Components of the genotyping reaction mix

Amount	Component
5 µl	TaqMan® Genotyping PCR Master Mix
0.125 µl	40 x TaqMan® SNP Genotyping Assay
2.875 µl	Aqua bidestillata sterilis (Fresenius Kabi, Graz, Austria)
2 µl	DNA

Standard cycling conditions were two minutes of heating at 50°C, followed by ten minutes at 95°C for the activation of AmpliTaq Gold® DNA Polymerase. Forty cycles were performed under the conditions of fifteen seconds at 95°C for denaturation and one minute at 60°C for annealing and extending of primers.

2.4.1. TaqMan® MGB probes and primers

The TaqMan® SNP Genotyping assay consists of unlabeled pre-designed PCR primers (36 µM) and TaqMan® minor groove binder (MGB) probes (8 µM). The final concentration of primers was 450 nM and 100 nM of probes in the reaction (Table 2)

Table 2. Pre-designed TaqMan® probes for genotyping AURKA polymorphisms

SNP	Chr. Position	Component	5' dye	Context sequence	3' dye	Allele
rs2180691	54397768	Probe 1	VIC®	AAAGTGC[**A**/G]TCTTCTG	NFQ/MGB	T
		Probe 2	FAM™	AAAGTGC[A/**G**]TCTTCTG	NFQ/MGB	C
rs8117896	54389471	Probe 1	VIC®	AGCCAGG[**C**/T]ACAATGG	NFQ/MGB	G
		Probe 2	FAM™	AGCCAGG[C/**T**]ACAATGG	NFQ/MGB	A
rs1468055	54399308	Probe 1	VIC®	ACACATG[**A**/C]TCTAATG	NFQ/MGB	T
		Probe 2	FAM™	ACACATG[A/**C**]TCTAATG	NFQ/MGB	G
rs1476394	54380550	Probe 1	VIC®	GCAGACA[**C**/G]CAAGCCA	NFQ/MGB	G
		Probe 2	FAM™	GCAGACA[C/**G**]CAAGCCA	NFQ/MGB	C

All four SNPs are located in introns in the AURKA gene on 20q13.2d.

2.4.2. Allelic Discrimination with TaqMan® MGB probes

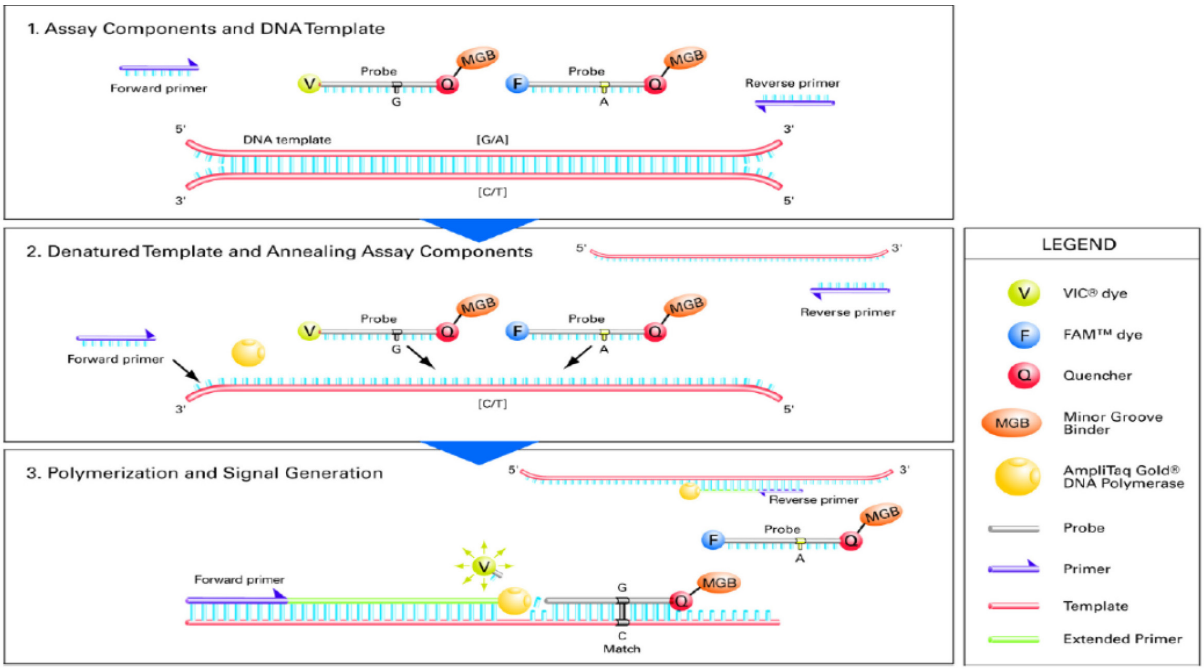
The Allelic Discrimination assay is a multiplexed end-point assay, which is able to detect variants of a single nucleic acid sequence using the 7500 Fast Real Time PCR System (Applied Biosystems, California, US; Applied Biosystems 2006a).

The TaqMan® SNP Genotyping Assay consists of two primers for the amplification of the polymorphic sequence of interest and of two TaqMan® minor groove binder (MGB) probes for the discrimination between the two alleles of interest. Each probe has a reporter dye at the 5' end (VIC® on allele 1 probe, FAM™ on allele 2 probe), which allows biallelic genotyping (Applied Biosystems 2006b). The minor groove binder at the 3' end of the probes causes an increased melting temperature (T_m), which allows the use of shorter probes (Afonina et al. 1997; Kutyavin et al. 1997).

A non-fluorescent quencher (NFQ) is located at the 3' end of each probe. During the PCR, each TaqMan® MGB probe anneals to its complementary sequence between

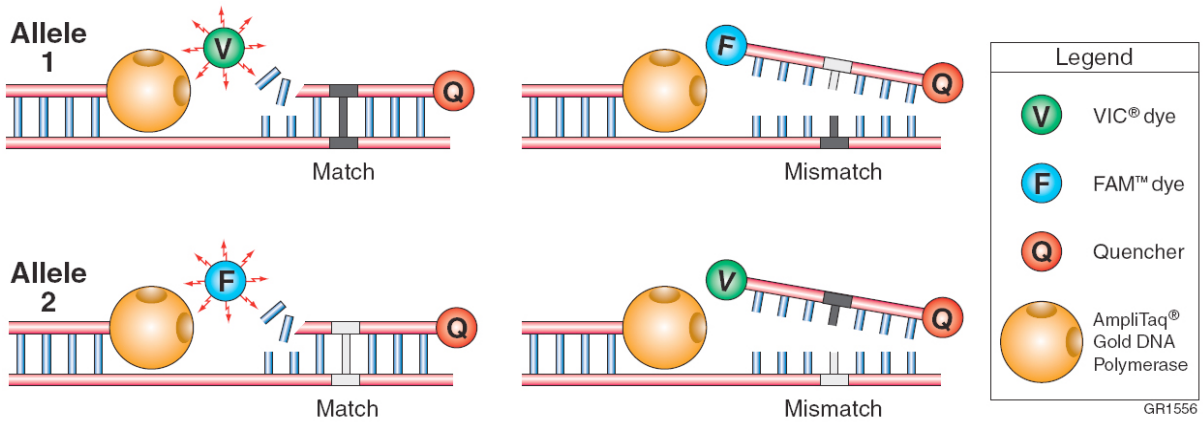
the both primers. The close proximity of the NFQ to the reporter dye results in the absorption of the emission of VIC® or FAM™ by fluorescence resonance energy transfer (FRET). In the next step, the AmpliTaq Gold DNA polymerase extends the primers and cuts off the probe from the target DNA when arriving at a probe hybridized to the template. This causes a separation of the reporter dye and the non-fluorescent quencher (Figure 2), thereby interrupting FRET (Applied Biosystems 2006b).

Figure 2. 5' Nuclease assay process (Applied Biosystems 2006b)



In Figure 3, the two possibilities of match and mismatch are shown. Even a single nucleotide mismatch between a probe and the sequence of interest may diminish the efficiency of probe hybridization, thus disabling the cleavage of the reporter dye from the probe. A mismatched probe is therefore displaced without cleavage from the AmpliTaq Gold DNA polymerase (Applied Biosystems 2006b).

Figure 3. Mismatch discrimination with 5' Nuclease Assay (Applied Biosystems 2006a)



The separation of the reporter dye from the non-fluorescent quencher leads to an increase in fluorescence signals of VIC® and/or FAM™ during each PCR-cycle. Allelic discrimination represents an end-point determination with three possible results: (1) an increase of VIC® fluorescence signal represents homozygousity for allele 1, (2) an increase of FAM™ homozygousity for allele 2 or (3) a heterozygous genotype correlates with both signals present at the end-point determination (Table 3; Applied Biosystems 2006b).

Table 3. Correlation between fluorescence signals and genotypes (based on Applied Biosystems, 2006b)

Substantial increase in:	Indication:
Only VIC fluorescence	Homozygousity for allele 1
Only FAM fluorescence	Homozygousity for allele 2
Both fluorescence signals	Heterozygousity for both alleles

The increase in the fluorescence of VIC® or FAM™ associated with the TaqMan® probes is determined by the sequence detection software (SDS). This software produces a scatter plot, in which corresponding normalized reporter (R_n) values are

grouped into three genotype clusters: homozygous for allele 1, homozygous for allele 2 and heterozygous. No template controls (NTCs) are used to measure the background signal (Applied Biosystems 2006a).

2.5. Statistical analysis

Genotypic counts of controls were tested for Hardy-Weinberg equilibrium (HWE) using a χ^2-test. Haplotype estimates were determined and the frequencies of the most common haplotypes were derived using the program FASTPHASE (Scheet and Stephens 2006).

Multiple logistic regression was applied to compare individuals of the case group against the control group. Separate models were estimated where each of the four polymorphisms described in Table 2 was included as three-level factor (homozygous wild type, heterozygous, homozygous polymorph) and each haplotype was included as explanatory variable. Age was identified as confounding variable. Odds ratios (ORs) and 95% confidence intervals (CIs) were estimated for each polymorphism and haplotype; reference categories were wild type and the most frequent haplotype, respectively.

Analysis of data was performed using the software R Version 2.9.0 (R Development Core Team 2009). All p-values are 2-sided; p-values <0.05 were considered to be statistically significant.

3. Results

3.1. Study population

In a total study population of 1577 male subjects, demographic variables and smoking status were analysed. The control group consisted of 550 BPH patients, the case group of 1027 prostate cancer patients.

The age indication is divided into four groups: subjects younger than 65, younger than 70, younger than 75 and 75 or older (Table 4). The control group had a statistically significant ($p<0.0001$) older age (67.01 years of age) than the case group (63.7).

Table 4. Age distribution

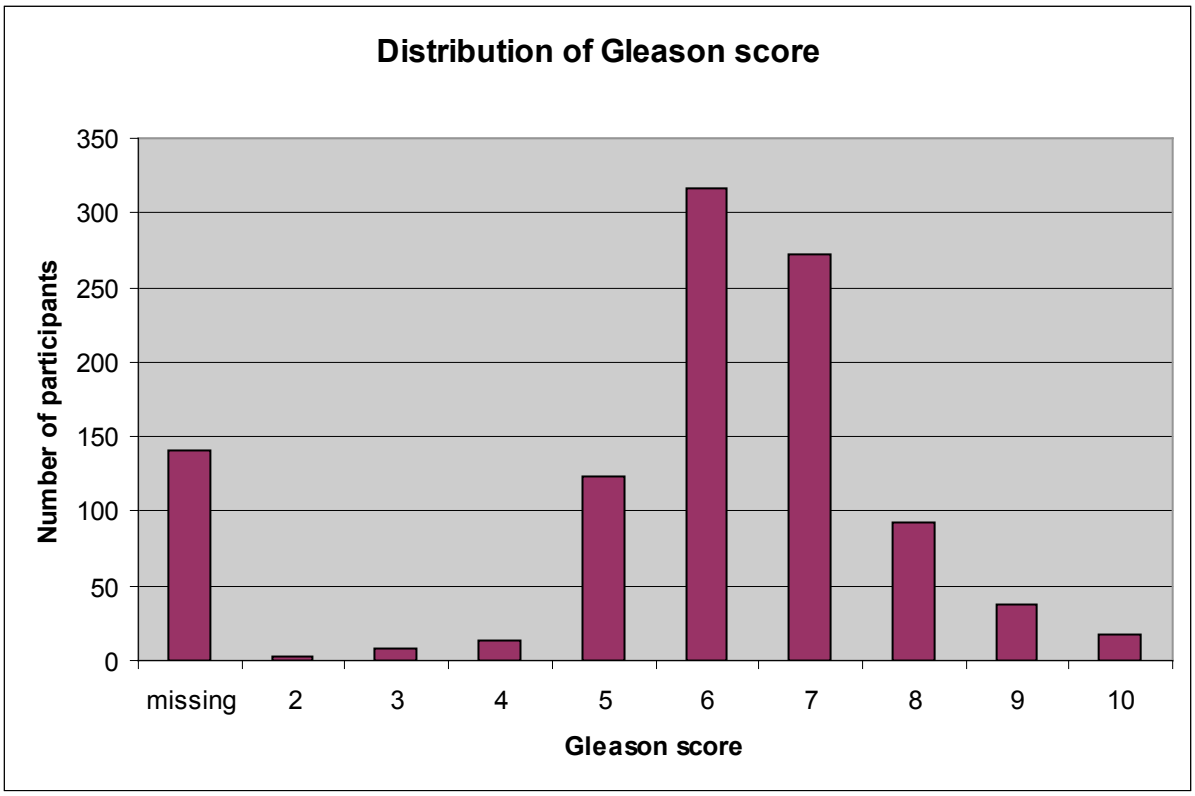
Age	Control (absolute value)	Control (%)	Cases (absolute value)	Cases (%)	Total
< 65	223	40.55	555	54.04	778
< 70	118	21.45	248	24.15	366
< 75	88	16	161	15.68	249
>= 75	121	22	63	6.13	184
Mean	67.01		63.7		
Sd	8.93		7.63		
95% CI	49.5-84.52		48.75-78.66		

In Table 5 and Figure 4 the distribution of the Gleason scores is shown. The most frequent Gleason score in the case group is 6 (30.87%), followed by 7 (26.48%) and 5 (12.07%). 141 Gleason score indications were missing (13.73%). The mean Gleason score is 6.51.

Table 5. Distribution of the Gleason scores

		Gleason score								
	Missing	2	3	4	5	6	7	8	9	10
Cases (absolute value)	141	3	8	13	124	317	272	93	38	18
Cases (%)	13.73	0.29	0.78	1.27	12.07	30.87	26.48	9.06	3.7	1.75
Mean						6.51				

Figure 4. Distribution of the Gleason scores

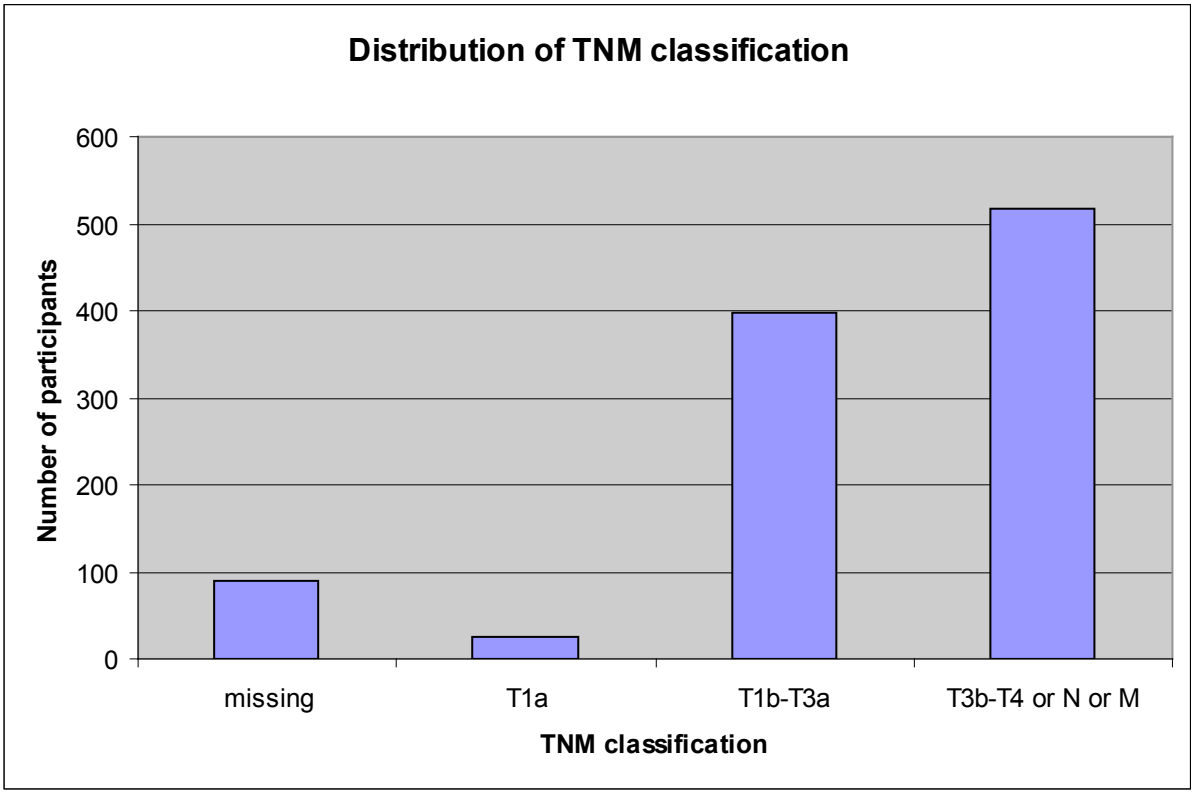


Discrimination was also made according to the TNM grading system (Table 6; Figure 5). More than half (50.15%) of the case group had a stage of T3b to T4 or regional lymph node metastases or distant metastases, and only 2.43% had a T1a classification.

Table 6. Distribution of TNM classification

		Stage		
	Missing	T1a	T1b-T3a	T3b-T4 or N or M
Cases (absolute value)	90	25	397	515
Cases (%)	8.76	2.43	38.66	50.15

Figure 5. Distribution of TNM classification

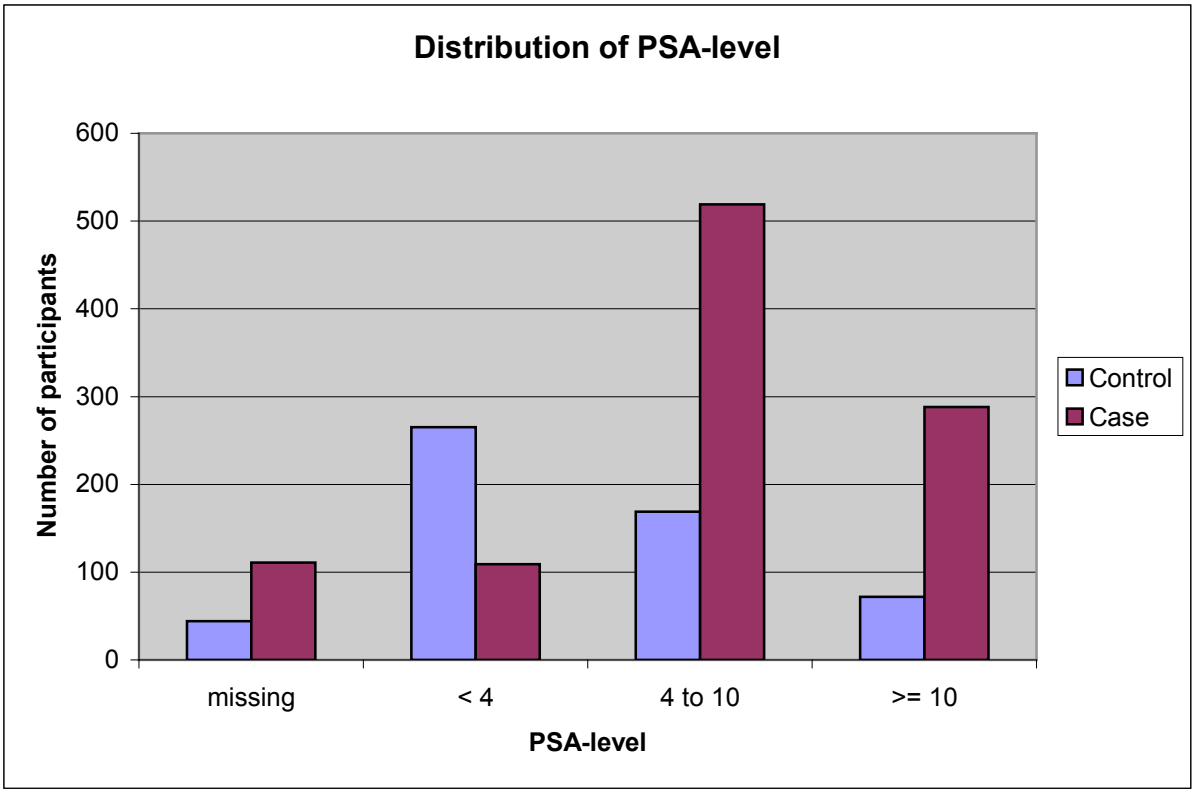


A PSA level of up to 4.0 ng/mL is regarded as unremarkable. The majority of the control group (48.18%) had a PSA level under this threshold, while more than half of the case group (50.54%) had a level between 4 and 10 ng/mL (Table 7; Figure 6). These differences are statistically significant ($p<0.0001$).

Table 7. Distribution of PSA-level

PSA (ng/mL)	Control (absolute value)	Control (%)	Cases (absolute value)	Cases (%)
Missing	44	8	111	10.81
<4	265	48.18	109	10.61
4 to 10	169	30.73	519	50.54
>=10	72	13.09	288	28.04

Figure 6. Distribution of PSA-level



The prostate weight of the study population was determined using TRUS. The control group did not have statistically significant (p=0.36) higher value (44.07 ml) than the case group (42.69 ml). The valid values, the mean prostate weight and the standard deviation are shown in Table 8.

Table 8. Distribution, mean and standard deviation of prostate weight

	Control (absolute value)	Control (%)	Case (absolute value)	Case (%)
Valid	302	54.91	534	52
Mean	44.07g		42.69g	
Sd	21.93		19.52	

Each participant of the study was interviewed about his smoking habits (Table 9). The majority of both groups (control: 44%, case: 40.7%) were former smokers. 13.64% of the controls and 12.56% of the cases were current smokers. The results were not statistically significant.

Table 9. Distribution of smoking habits

	Control (absolute value)	Control (%)	Case (absolute value)	Case (%)
Missing	30	5.45	88	8.57
Never	203	36.91	392	38.17
Former	242	44	418	40.7
Current	75	13.64	129	12.56

Furthermore, the average consumed cigarette packs per year (Table 10) and the average starting age (Table 11) were evaluated. The case group smoked in average 2.69 packs more in one year (31.18 packs) than the control group (28.49 packs). No differences could be detected considering the starting age. Both results showed no statistical significance.

Table 10. Distribution, mean and standard deviation of consumed cigarette packs per year

	Control (absolute value)	Control (%)	Case (absolute value)	Case (%)
Valid	283	51.45	450	43.82
Mean	28.49		31.18	
Sd	26.14		28.42	

Table 11. Distribution, mean and standard deviation of starting age

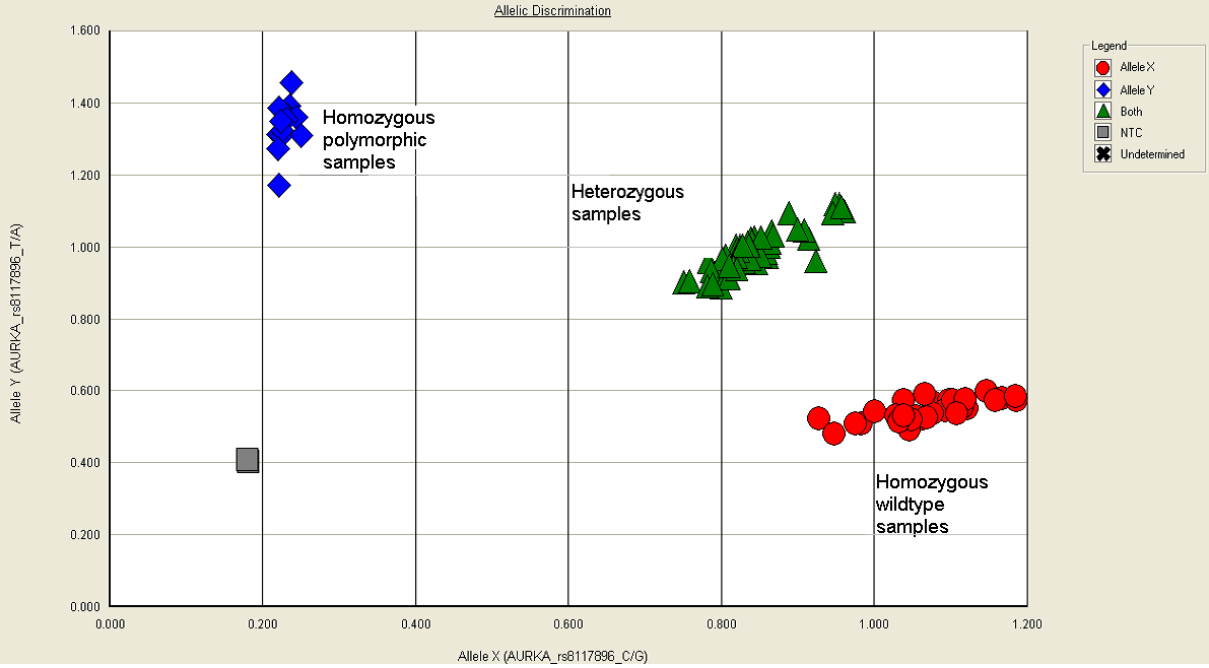
	Control (absolute value)	Control (%)	Case (absolute value)	Case (%)
Valid	293	53.27	468	45.57
Mean	18.79		18.94	
Sd	5.72		6.18	

3.2. AURKA genotypes

AURKA genotypes (rs2180691, rs8117896, rs1468055, rs1476394) were determined using TaqMan® MGB probes (Applied Biosystems, California, US). An end-point allelic determination was performed on the PCR products. Corresponding normalized reporter (R_n) values were grouped into four different clusters (homozygous for allele 1/wild type, homozygous for allele 2/polymorphic, heterozygous and no template controls/NTCs) and shown as points on a scatter plot. The graph location of the points is corresponding to the three different specific genotypes.

Figure 7 shows a representative example of the allelic discrimination scatter plot for rs8117896. The PCR reaction mix contained a VIC-labelled probe for the C/G allele on the x-axis and a FAM-labelled probe for the T/A allele on the y-axis.

Table 7. SNP rs8117896 allelic discrimination assay – example of a scatter plot



The genotype distributions of the four analysed SNPs in AURKA were all found to be in HWE among controls.

In Table 12, the frequency distribution of all four SNPs in controls and cases is shown. The homozygous wild type was applied as a reference category. Carriers of the heterozygous polymorphic genotype of rs2180691 had an OR of 0.92 (95% CI= 0.74-1.15) and carriers of the homozygous polymorphic genotype had 1.06 (95% CI= 0.73-1.53). For the rs8117896 heterozygous polymorphic genotype, the OR calculated relative to patients with the homozygous genotype was 0.89 (95% CI= 0.71-1.13). For the polymorphic homozygous variant, the OR was 1.06 (95% CI= 0.77-1.46). For rs1468055 heterozygous polymorphic genotype, the OR was 0.93 (95% CI= 0.74-1.16) and for homozygous polymorphic genotype was 0.84 (95% CI= 0.51-1.39). The fourth SNP rs1476394 had an OR of 1.03 (95% CI= 0.81-1.3) for the heterozygous polymorphic genotype and an OR of 0.96 (95% CI= 0.51-1.8) for the homozygous polymorphic genotype.

No statistically significant association between the four SNPs and prostate cancer risk was found.

Table 12. Distribution of AURKA genotypes in association with prostate cancer risk (confounder: age)

SNP	Genotype	Controls (%)	Cases (%)	OR	95% CI	p-value
rs2180691	G/G	271 (49.27)	525 (51.12)	1		
	G/A	227 (41.27)	401 (39.05)	0.92	0.74-1.15	0.47
	A/A	52 (9.46)	101 (9.83)	1.06	0.73-1.53	0.77
rs8117896	C/C	192 (34.91)	375 (36.51)	1		
	C/T	275 (50)	486 (47.32)	0.89	0.71-1.13	0.35
	T/T	83 (15.09)	166 (16.17)	1.06	0.77-1.46	0.73
rs1468055	C/C	328 (59.64)	636 (61.93)	1		
	C/A	193 (35.09)	347 (33.79)	0.93	0.74-1.16	0.53
	A/A	29 (5.27)	44 (4.28)	0.84	0.51-1.39	0.5
rs1476394	C/C	381 (69.27)	706 (68.74)	1		
	C/G	152 (27.64)	292 (28.43)	1.03	0.81-1.3	0.83
	G/G	17 (3.09)	29 (2.83)	0.96	0.51-1.8	0.9

Stratified analysis according to Gleason score was calculated for patients with a Gleason score <7 and >=7 (Table 13). No statistically significant association was found.

Table 13. Distribution of AURKA genotypes in association with prostate cancer risk and discrimination between Gleason score <7 and >=7 (confounder: age)

			Gleason <7				Gleason >=7			
SNP	Genotype	Controls (%)	Cases (%)	OR	95% CI	p-value	Cases (%)	OR	95% CI	p-value
rs2180691	G/G	271 (49.27)	307 (29.89)	1			218 (21.23)	1		
	G/A	227 (41.27)	239 (23.27)	0.94	0.73-1.2	0.61	162 (15.78)	0.9	0.68-1.18	0.45
	A/A	52 (9.46)	60 (5.84)	1.04	0-69-1.57	0.85	41 (3.99)	1.07	0.68-1.69	0.77
rs8117896	C/C	192 (34.91)	228 (22.2)	1			147 (14.31)	1		
	C/T	275 (50)	283 (27.56)	0.85	0.66-1.1	0.22	203 (19.77)	0.95	0.71-1.26	0.71
	T/T	83 (15.09)	95 (9.25)	0.98	0.68-1.4	0.91	71 (6.91)	1.15	0.78-1.7	0.49
rs1468055	C/C	328 (59.64)	364 (35.44)	1			272 (26.49)	1		
	C/A	193 (35.09)	213 (20.74)	1	0.78-1.29	0.98	134 (13.05)	0.81	0.61-1.07	0.13
	A/A	29 (5.27)	29 (2.82)	0.99	0.57-1.71	0.96	15 (1.46)	0.63	0.33-1.22	0.17
rs1476394	C/C	381 (69.27)	420 (40.89)	1			286 (27.85)	1		
	C/G	152 (27.64)	169 (16.46)	1.02	0.78-1.33	0.87	123 (11.98)	1.03	0.78-1.38	0.82
	G/G	17 (3.09)	17 (1.65)	0.9	0.45-1.83	0.78	12 (1.17)	0.95	0.44-2.08	0.9

Stratified analysis according to age also showed no statistically significant association between the polymorphic alleles of the investigated SNPs and prostate cancer (Table 14).

Table 14. Distribution of AURKA genotypes in association with prostate cancer risk and discrimination between age <=64 and >64

		Age <=64					Age >64				
SNP	Genotype	Controls (%)	Cases (%)	OR	95% CI	p-value	Controls (%)	Cases (%)	OR	95% CI	p-value
rs2180691	G/G	120 (21.82)	292 (28.43)	1			151 (27.45)	233 (22.69)	1		
	G/A	82 (14.91)	217 (21.13)	1.09	0.78-1.51	0.62	145 (26.36)	184 (17.92)	0.77	0.56-1.04	0.09
	A/A	21 (3.82)	46 (4.48)	0.9	0.52-1.57	0.71	31 (5.64)	55 (5.35)	1.16	0.7-1.91	0.57
rs8117896	C/C	80 (14.55)	196 (19.08)	1			112 (20.36)	179 (17.43)	1		
	C/T	113 (20.55)	278 (27.07)	1	0.71-1.41	0.98	162 (29.45)	208 (20.25)	0.82	0.59-1.14	0.23
	T/T	30 (5.45)	81 (7.89)	1.1	0.67-1.8	0.71	53 (9.64)	85 (8.28)	1.06	0.69-1.62	0.81
rs1468055	C/C	126 (22.91)	353 (34.37)	1			202 (36.73)	283 (27.56)	1		
	C/A	87 (15.82)	182 (17.72)	0.75	0.54-1.04	0.08	106 (19.27)	165 (16.07)	1.14	0.84-1.57	0.4
	A/A	10 (1.82)	20 (1.94)	0.71	0.32-1.56	0.39	19 (3.45)	24 (2.34)	1	0.52-1.95	0.98
rs1476394	C/C	157 (28.55)	375 (36.51)	1			224 (40.73)	331 (32.23)	1		
	C/G	58 (10.55)	164 (15.97)	1.18	0.83-1.68	0.36	94 (17.09)	128 (12.46)	0.94	0.68-1.31	0.72
	G/G	8 (1.45)	16 (1.56)	0.83	0.35-1.99	0.68	9 (1.63)	13 (1.27)	1.34	0.54-3.33	0.53

Overall, the analysed data showed no statistically significant association between the four investigated AURKA SNPs and prostate cancer risk.

The minor allele frequencies (MAFs) in the control group of this study population were in concordance (rs8117896, rs1468055, rs1476394) with those reported in the HapMap database for the Caucasian subpopulation. SNP rs2180691 shows a small aberration from the database (Table 15).

Table 15. MAF distribution of the AURKA SNPs in control group and in the HapMap database

SNP	Chr. Position	MAF (Control Group)	MAF (HapMap database [Caucasian subpopulation])
rs2180691	54397768 G → A	0.301	0.239
rs8117896	54389471 C → T	0.401	0.367
rs1468055	54399308 C → A	0.228	0.243
rs1476394	54380550 C → G	0.169	0.208

To explore multilocus associations, haplotype analyses were conducted for the four tagging SNPs (rs2180691, rs8117896, rs1468055, rs1476394) within AURKA. Again, corresponding adjustment for age was performed (Table 16).
The most common haplotype (homozygous for wild type in all four SNPs: 1111) was set as a reference category. The results showed no statistical significance.

Table 16. Distribution of AURKA haplotypes (confounder: age)

Haplotype	Controls	Cases	OR	95% CI	p-value	Percentage cases of total number
1111	444	852	1			65.7
3311	240	429	0.94	0.77-1.15	0.53	64.1
1131	135	229	0.89	0.7-1.14	0.36	62.9
1313	91	187	1.08	0.82-1.43	0.58	67.3
1333	92	147	0.85	0.64-1.14	0.27	61.5
3111	68	124	0.99	0.72-1.37	0.96	64.6

A haplotype analysis according to Gleason score was performed (Table 17). The major aberration from the reference in the group of Gleason score <7 was found in haplotypes 1333 (0.86; 95% CI= 0.62-1.19) and 3111 (1.1; 95% CI= 0.77-1.57); and in haplotypes 1131 (OR= 0.73; 95% CI= 0.54-1.0) and 1333 (OR=0.81; 95% CI= 0.57-1.16) in the group of Gleason score >=7.

No statistically significant association was found.

Table 17. Distribution of AURKA haplotypes and discrimination between Gleason score <7 and >=7 (confounder: age)

		Gleason <7				Gleason >=7			
Haplotype	Controls	Cases	OR	95% CI	p-value	Cases	OR	95% CI	p-value
1111	444	494	1			358	1		
3311	240	248	0.92	0.74-1.15	0.46	181	0.95	0.74-1.21	0.66
1131	135	147	0.99	0.76-1.3	0.91	82	0.73	0.54-1.0	0.05
1313	91	105	1.06	0.77-1.45	0.73	82	1.1	0.79-1.55	0.57
1333	92	86	0.86	0.62-1.19	0.36	61	0.81	0.57-1.16	0.25
3111	68	79	1.1	0.77-1.57	0.6	45	0.86	0.57-1.3	0.47

Furthermore, stratified analysis according to age was performed. It revealed that carriers of the haplotype 1131 have a statistically significant lower prostate cancer risk (OR= 0.7021, 95% CI= 0.49599-0.99384, p-value 0.04606) before or at the age of 64 years (Table 18).

Table 18. Distribution of AURKA haplotypes and discrimination between age <=64 and >64

	Age <=64					Age >64				
Haplotype	Controls	Cases	OR	95% CI	p-value	Controls	Cases	OR	95% CI	p-value
1111	181	472	1			263	380	1		
3311	98	228	0.89	0.67-1.19	0.44	142	201	0.99	0.75-1.3	0.92
1131	65	119	0.7	0.5-0.99	**0.05**	70	110	1.16	0.82-1.65	0.41
1313	36	110	1.17	0.77-1-77	0.46	55	77	1.08	0.73-1.6	0.71
1333	36	77	0.82	0.53-1.26	0.36	56	70	0.92	0.61-1.37	0.67
3111	21	62	1.13	0.67-1.91	0.64	47	62	0.91	0.6-1.39	0.66

4. Discussion

In this diploma thesis four intronic single nucleotide polymorphisms within the AURKA gene and their association with prostate cancer risk in an Austrian population were investigated. AURKA is located on HPC20, a region which is known to be a prostate cancer susceptibility locus (Berry et al. 2000). Therefore, AURKA is a target gene for molecular epidemiologic prostate cancer studies.

The vast majority of metastatic and invasive prostate tumors show centrosome abnormalities. Because centrosomes control the mitotic chromosome segregation, they play an important role in the maintenance of genomic stability (Pihan et al. 2001).

AURKA is required for multiple crucial events during mitosis, such as mitotic entry, separation of centriole pairs, accurate bipolar spindle assembly, and alignment of chromosomes to the metaphase plate and completion of cytokinesis (Marumoto et al. 2003). Centrosome abnormalities and aneuploidy resulting in genomic instability are the results of defects in these processes that may be due to over-expression of AURKA (Pihan et al. 2001). Two studies (Lee et al. 2006; Buschhorn et al. 2005) found that AURKA is over-expressed in PIN, in prostate tumors, as well as in prostate cancer cell lines. Buschhorn et al. found an over-expression in 29% of normal prostate glands, suggesting that it is an early event in carcinogenesis. The level of AURKA expression increases as prostate cells transform from normal to high-grade PIN to carcinoma. As a possible therapeutic approach, AURKA inhibitor VX680 reduces prostate cancer cell survival. Combination with doxorubicin intensifies this effect (Lee et al. 2006).

Until now, molecular epidemiologic studies concerning SNPs in AURKA gene and prostate cancer focused on polymorphism T31A (Phe31Ile; rs2273535).

A functional analysis of the T31A polymorphism was conducted by Matarasso et al. This study showed an association of the heterozygous and the polymorphic homozygous genotype of rs2273535 with significant AURKA over-expression (Matarasso et al. 2007).

Ewart-Toland et al. conducted a meta-analysis of fifteen case-control studies, mostly unpublished. Two prostate cancer case-control sets were included in this study. One nested case-control study population of the Physician's Health Study, a cohort study with 501 cases and 501 controls (Chan et al. 2002), showed a statistically

insignificant increased prostate cancer risk (OR= 1.76, 95% CI= 1.01-3.04) for the homozygous AA genotype of rs2273535. In the second population consisting of 559 cases and 534 population-based controls, no altered risk (OR= 0.91, 95% CI= 0.54 – 1.55) was observed. Overall, this meta-analysis showed that carriers of the polymorphic AA genotype had a statistically significant increased risk of multiple cancers such as colon, breast, lung, prostate, non-melanoma skin, esophageal and ovarian cancer with an OR of 1.4, 95% CI= 1.22-1.59, $p < 0.001$ (Ewart-Toland et al. 2005). It must be take into account that this kind of study has limitations. Publication bias cannot be excluded because positive results are more likely to be published. Ewart-Toland et al. tried to avoid this problem by using mainly unpublished studies. Another limitation is the different study-design of studies used in a meta-analysis. Eligibility criteria for inclusions of study participants were different among the studies used in the meta-analysis of Ewart-Toland et al. Furthermore, studies with population-based and hospital-based control groups were included. In general, the selection of a control group is a critical point in the design of molecular epidemiologic studies. Especially in prostate cancer studies, a population-based study design would be hampered by the uncertainty of underlying prostate cancer as this disease occurs at high frequency in the seventh to eighth decade of life. To adress this, we used hospital-based BPH controls, in which the absence of prostate cancer was confirmed either histologically or clinically. One can consider that using a control group consisting of men with BPH could be a limitation of the study because it is also a hormone-dependent disease. We consider BPH patients as an appropriate control group because it is generally accepted that this disease is neither a premalignant lesion nor a precursor carcinoma. Furthermore, the incidence of histological evidence of BPH is in the range of 70 – 80% in the seventh to eighth decade of life, so it is almost impossible to identify a population without BPH in this age group (Berry et al. 1984, Gsur et al. 2002, Gsur et al 2000). We considered the use of hospital-based BPH controls as a strength of our studies, as well as the histological confirmation of prostate cancer in cases and the large sample size of 1027 cases and 550 controls.

The present study is a supplement of a previously published molecular epidemiologic prostate cancer study investigating SNPs rs2273535 and rs8173 and their association with prostate cancer risk (Feik et al. 2009). Four additional intronic tagging SNPs (rs2180691, rs8117896, rs1468055 and rs1476394) and their association with prostate cancer risk were investigated in this study to cover the

whole AURKA gene with six different tag SNPs. To our knowledge, this is the first study to investigate these four SNPs and their association with prostate cancer risk. However, no statistically significant association was found. Stratified analysis according to age revealed that carriers of a certain haplotype (G/G genotype in rs2180691, C/C genotype in rs8117896, A/A genotype in rs1468055 and C/C genotype in rs1476394) had a statistically significant lower prostate cancer risk before or at the age of 64 years. Other results of the haplotype analysis showed no statistical significance. A statistical analysis combining the six SNPs investigated in both studies is planned to be conducted and published.

In this study, MAFs were found to be in concordance with those reported in the HapMap database for the Caucasian population. The small aberration in MAF of rs2180691 from the HapMap database can be explained by the smaller sample size (n=226) compared to the present study (n=550).

The aim of this molecular epidemiologic study is to find a reliable biomarker for prostate cancer risk. Because prostate cancer is a multi-factorial disease, many genetic variations are involved (Schaid 2004). To adress this, genome-wide association studies using SNP Arrays have been conducted recently or are currently under the way. Through this method, it is possible to investigate the effect of inherited genetic variation on cancer risk and to search for cancer risk loci without concentrating on single candidate genes. Furthermore small to modest effects of SNPs can be detected. Identification of genetic variants, which are strongly associated with prostate cancer, may form a basis for improved screening and chemoprevention.

5. References

Afonina I, Zivarts M, Kutyavin I, Lukhtanov E, Gamper H, Meyer RB. (1997) Efficient priming of PCR with short oligonucleotides conjugated to a minor groove binder. *Nucleic Acids Research* 25: 2657-2660

Alexander F. (1999) *International Handbook of prostate cancer edited by David Kirk*: 1-12. Haslemere: Euromed Communications Ltd.

Applied Biosystems. (2006a) Allelic Discrimination Getting Started Guide. www.appliedbiosystems.com

Applied Biosystems (2006b) TaqMan® SNP Genotyping Assays Protocol. www.appliedbiosystems.com

Beckmann JS, Estivill X, Antonarakis SE. (2007) Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nature Reviews. Genetics* 8: 639-646

Berry R, Schroeder JJ, French AJ, McDonnell SK, Peterson BJ, Cunningham JM, Thibodeau SN, Schaid DJ. (2000) Evidence for a prostate cancer-susceptibility locus on chromosome 20. *American Journal of Human Genetics* 67: 82-91

Berry SJ, Coffey DS, Walsh PC, Ewing LL. (1984) The development of human benign prostatic hyperplasia with age. *The Journal of Urology* 132: 474-479. Quoted in Feik E, Baierl A, Madersbacher S, Schatzl G, Maj-Hes A, Berges R, Micksche M, Gsur A. (2009) Common genetic polymorphisms of AURKA and prostate cancer risk. *Cancer Causes & Control: CCC* 20: 147-152

Bischoff JR, Anderson L, Zhu Y, Mossie K, Ng L, Souza B, Schryver B, Flanagan P, Clairvoyant F, Ginther C, Chan CS, Novotny M, Slamon DJ, Plowman GD. (1998) A homologue of Drosophila aurora kinase is oncogenic and amplified in human colorectal cancers. *The EMBO Journal* 17: 3052-3065

Bostwick D. (1999) *International Handbook of prostate cancer edited by David Kirk*: 13-42. Haslemere: Euromed Communications Ltd.

Buschhorn HM, Klein RR, Chambers SM, Hardy MC, Green S, Bearss D, Nagle RB. (2005) Aurora-A over-expression in high-grade PIN lesions and prostate cancer. *The Prostate* 64: 341-346

Chan JM, Stampfer MJ, Ma J, Gann P, Gaziano JM, Pollak M, Giovannucci E. (2002) Insulin-like growth factor-I (IGF-I) and IGF binding protein-3 as predictors of advanced-stage prostate cancer.
Journal of the National Cancer Institute 94: 1099-1106

Chanock S. (2001) Candidate genes and single nucleotide polymorphisms (SNPs) in the study of human disease. *Disease Markers* 17: 89-98. Quoted in Rebbeck TR, Ambrosone CB, Bell DA, Chanock SJ, Hayes RB, Kadlubar FF,

Chen Y, Clegg NJ, Scher HI. (2009) Anti-androgens and androgen-depleting therapies in prostate cancer: new agents for an established target. *The Lancet Oncology* 10: 981-991

Chodak GW. (1994) The role of watchful waiting in the management of localized prostate cancer. *The Journal of Urology* 152: 1766-1768. Quoted in Scheet P, Stephens M. (2006) A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *American Journal of Human Genetics* 78: 629-644

Drazen JM, Yandava CN, Dubé L, Szczerback N, Hippensteel R, Pillari A, Israel E, Schork N, Silverman ES, Katz DA, Drajesk J. (1999) Pharmacogenetic association between ALOX5 promoter genotype and the response to anti-asthma treatment. *Nature Genetics* 22: 168-170

Epstein JI. (2009) An Update of the Gleason Grading System. *The Journal of Urology* 183: 433-440

Erichsen HC, Chanock SJ. (2004) SNPs in cancer research and treatment. *British Journal of Cancer* 90: 747-751

Ewart-Toland A, Dai Q, Gao YT, Nagase H, Dunlop MG, Farrington SM, Barnetson RA, Anton-Culver H, Peel D, Ziogas A, Lin D, Miao X, Sun T, Ostrander EA, StanfordJL, Langlois M, Chan JM, Yuan J, Harris CC, Bowman ED, ClaymanGL, Lippman SM, Lee JJ, Zheng W, Balmain A. (2005) Aurora-A/STK15 T+91A is a general low penetrance cancer susceptibility gene: a meta-analysis of multiple cancer types. *Carcinogenesis* 26: 1368-1373

Feik E, Baierl A, Madersbacher S, Schatzl G, Maj-Hes A, Berges R, Micksche M, Gsur A. (2009) Common genetic polymorphisms of AURKA and prostate cancer risk. *Cancer Causes & Control: CCC* 20: 147-152

Gallagher RP, Fleshner N. (1998) Prostate cancer: 3. Individual risk factors. *CMAJ: Canadian Medical Association Journal = Journal De l'Association Medicale Canadienne* 159: 807-813

Gandellini P, Folini M, Zaffaroni N. (2009) Towards the definition of prostate cancer related microRNAs: where are we now? *Trends in Molecular Medicine* 15: 381-390

Gleason DF, and Mellinger GT. (1974) Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. *The Journal of Urology* 111: 58-64. Quoted in Epstein JI. (2009) An Update of the Gleason Grading System. *The Journal of Urology* 183: 433-440

Gsur A, Bernhofer G, Hinteregger S, Haidinger G, Schatzl G, Madersbacher S, Marberger M, Vutuc C, Micksche M. (2000) A polymorphism in CYP17 gene is associated with prostate cancer risk. *International Journal of Cancer* 87: 434-437
Gsur A, Preyer M, Haidinger G, Zidek T, Madersbacher S, Schatzl G, Marberger M, Vutuc C, Micksche M. (2002) Polymorphic CAG repeats in the androgen receptor gene, prostate-specific antigen polymorphism and prostate cancer risk. *Carcinogenesis* 23: 1647-1651

Hackl M. (2004) *Krebsinzidenz und Krebsmortalität in Österreich*. Wien: Statistik Austria.

International HapMap Consortium. (2005) A haplotype map of the human genome. *Nature* 437: 1299-1320

Johns LE, Houlston RS. (2003) A systematic review and meta-analysis of familial prostate cancer risk. *BJU International* 91: 789-794

Knippers, R. (1995) *Molekulare Genetik*. 6. Aufl. Stuttgart: Georg Thieme Verlag.

Kollareddy M, Dzubak P, Zheleva D, Hajduch M. (2008) Aurora kinases: structure, functions and their association with cancer. *Biomedical Papers of the Medical Faculty of the University Palacký, Olomouc, Czechoslovakia* 152: 27-33

Köllermann J, Sauter G. (2009) Trends in prostate biopsy interpretation. *Der Urologe* 48: 305-312

Kotnis A, Sarin R, Mulherkar R. (2005) Genotype, phenotype and cancer: role of low penetrance genes and environment in tumour susceptibility. *Journal of Biosciences* 30: 93-102

Krawczak M, Reiss J, Cooper DN. (1992) The mutational spectrum of single base pair substitutions in mRNA splice junctions of human genes: causes and consequences. *Human Genetics* 90: 41-54. Quoted in Schork NJ, Fallin D, Lanchbury JS. (2000) Single nucleotide polymorphisms and the future of genetic epidemiology. *Clinical Genetics* 58: 250-264

Kruglyak L, Nickerson DA. (2001) Variation is the spice of life. *Nature Genetics* 27: 234-236

Kubistova Z, Mrazek F, Petrek M. (2009) Polymorphisms of the immune response genes: selected biological, methodical and medical aspects. *Biomedical Papers of the Medical Faculty of the University Palacký, Olomouc, Czechoslovakia* 153: 93-102

Kutyavin IV, Lukhtanov EA, Gamper HB, Meyer RB. (1997) Oligonucleotides with conjugated dihydropyrroloindole tripeptides: base composition and backbone effects on hybridization. *Nucleic Acids Research* 25: 3718-3723.

Lee EC, Frolov A, Li R, Ayala G, Greenberg NM. (2006) Targeting Aurora kinases for the treatment of prostate cancer. *Cancer Research* 66: 4996-5002

Malayeri R, Steger G. (1996) Prostatakarzinom. *Klinische Onkologie* edited by , Pirker R, Fiegl M, and Huber H., 327-330. Wien: Facultas Universitätsverlag.

Marumoto T, Honda S, Hara T, Nitta M, Hirota T, Kohmura E, Saya H. (2003) Aurora A kinase maintains the fidelity of early and late mitotic events in HeLa cells. *The Journal of Biological Chemistry* 278: 51786-51795

Matarasso N, Bar-Shira A, Rozovski U, Rosner S, Orr-Urtreger A. (2007) Functional analysis of the Aurora Kinase A Ile31 allelic variant in human prostate. *Neoplasia* 9: 707-715.

Mellinger, G T, D Gleason, and J Bailar. (1967) The histology and prognosis of prostatic cancer. *The Journal of Urology* 97: 331-337. Quoted in Epstein JI. (2009) An Update of the Gleason Grading System. *The Journal of Urology* 183: 433-440

Newling D. (1999) *International Handbook of prostate cancer edited by David Kirk*: 61-20. Haslemere: Euromed Communications Ltd.

Parkin DM, Bray FI, Devesa SS. (2001) Cancer burden in the year 2000. The global picture. *European Journal of Cancer (Oxford, England: 1990)* 37: 4-66.

Perera FP. (1997) Environment and cancer: who are susceptible? *Science* 278: 1068-1073

Perera FP, Weinstein IB. (1982) Molecular epidemiology and carcinogen-DNAadduct detection: new approaches to studies of human cancer causation. *Journal of Chronic Diseases* 35: 581-600

Perera FP, Weinstein IB. (2000) Molecular epidemiology: recent advances and future directions. *Carcinogenesis* 21: 517-524

Pihan GA, Purohit A, Wallace J, Malhotra R, Liotta L, Doxsey SJ. (2001) Centrosome defects can account for cellular and genetic changes that characterize prostate cancer progression. *Cancer Research* 61: 2212-2219.

R Development Core Team. (2009) *R: A Language and Environment for Statistical Computing*. http://www.R-project.org.

Rebbeck TR, Ambrosone CB, Bell DA, Chanock SJ, Hayes RB, Kadlubar FF, Thomas DC. (2004) SNPs, haplotypes, and cancer: applications in molecular epidemiology. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology* 13: 681-687

Risch NJ. (2000) Searching for genetic determinants in the new millennium. *Nature* 405: 847-856

Sakakura C, Hagiwara A, Yasuoka R, Fujita Y, Nakanishi M, Masuda K, Shimomura K, Nakamura Y, Inazawa J, Abe T, Yamagishi H. (2001) Tumour-amplified kinase BTAK is amplified and overexpressed in gastric cancers with possible involvement in aneuploid formation. *British Journal of Cancer* 84: 824-831

Salisbury BA, Pungliya M, Choi JY, Jiang R, Sun XJ, Stephens JC. (2003) SNP and haplotype variation in the human genome. *Mutation Research* 526: 53-61

Schaid, DJ. (2004) The complex genetic epidemiology of prostate cancer. *Human Molecular Genetics* 13: 103-121

Scheet P, Stephens M. (2006) A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *American Journal of Human Genetics* 78: 629-644

Schork NJ, Fallin D, Lanchbury JS. (2000) Single nucleotide polymorphisms and the future of genetic epidemiology. *Clinical Genetics* 58: 250-264

Selley S, Donovan J, Faulkner A, Coast J, Gillatt D. (1997) Diagnosis, management and screening of early localised prostate cancer.
Health Technology Assessment 1: 1-96

Shand RL, Gelmann EP. (2006) Molecular biology of prostate-cancer pathogenesis. *Current Opinion in Urology* 16: 123-131

Spitz MR, Bondy ML. (2010) The evolving discipline of molecular epidemiology of cancer. *Carcinogenesis* 31: 127-134

Statistik Austria a. Österreichisches Krebsregister (Stand 24. 02. 2009, erstellt am 28. 04. 2009).
http://www.statistik.at/web_de/static/krebsinzidenz_neuerkankungen_pro_jahr_nach_lokalisationen_und_geschlecht_o_021806.xls

Statistik Austria b. Österreichisches Krebsregister (Stand 24. 02. 2009, erstellt am 9. 03. 2009).
http://www.statistik.at/web_de/statistiken/gesundheit/krebserkrankungen/krebsinzidenz_im_ueberblick/020524.html

Statistik Austria c. Todesursachenstatistik (Erstellt am 16. 06. 2009)
http://www.statistik.at/web_de/static/gestorbene_2008_nach_todesursachen_alter_und_geschlecht_021987.xls

Vutuc C, Schernhammer ES, Haidinger G, Waldhör T. (2005) Prostate cancer and prostate-specific antigen (PSA) screening in Austria.
Wiener Klinische Wochenschrift 117: 457-461

Wittekind C, Sobin LH, Klimpfinger M. (2005) *TNM-Atlas*. Berlin/Heidelberg: Springer Verlag

Zhou H, Kuang J, Zhong L, Kuo WL, Gray JW, Sahin A, Brinkley BR, Sen S. (1998) Tumour amplified kinase STK15/BTAK induces centrosome amplification, aneuploidy and transformation. *Nature Genetics* 20: 189-193

6. Appendix

6.1. Abstract

The present diploma thesis is part of the project “Molecular Epidemiology of Prostate Cancer”, a collaboration of the Department of Applied and Experimental Oncology of the Institute of Cancer Research and the Departments of Urology of the Medical University Vienna, Sozialmedizinisches Zentrum Süd and Sozialmedizinisches Zentrum Ost. Aim of this study is to create a polygenetic model to identify high-risk patients and to open new perspectives in prevention and therapy of prostate cancer. Prostate cancer is the most common malign neoplasm in men in Austria.
In this diploma thesis four single nucleotide polymorphisms (SNPs) within the AURKA gene were genotyped. AURKA (Aurora Kinase A), a serin/threonin protein kinase, is located on chromosome 20q13, a known prostate cancer susceptibility locus, and is involved in some crucial events during mitosis. Genetic polymorphisms in AURKA gene may contribute to interindividual differences in chromosomal stability and therefore influence the risk of prostate cancer.
The study population of this ongoing study consists of 1027 prostate cancer cases and of 550 benign prostatic hyperplasia (BPH) controls. Four selected tagging SNPs (rs2180691, rs8117896, rs1468055, rs1476394) were determined using -T5’ nuclease TaqMan MGB Assay. Overall, no significant association could be found between the investigated SNPs and prostate cancer risk.
However, a haplotype analysis revealed a reduced prostate cancer risk (OR: 0.7021; 95% CI: 0.49599-0.99384) for men below 64 years of age with the GCAC haplotype. Further investigations in larger study populations are required to assess the influence of polymorphisms in the AURKA gene on prostate cancer risk.

6.2. Zusammenfassung

Die vorliegende Diplomarbeit ist ein Teil des Projekts „Molekulare Epidemiologie von Prostatakrebs“, einer Zusammenarbeit der Abteilung für angewandte und experimentelle Onkologie des Instituts für Krebsforschung sowie der urologischen Abteilungen der Medizinischen Universität Wien, des Sozialmedizinischen Zentrums Süd und des Sozialmedizinischen Zentrums Ost.

Ziel dieser Studie ist es, ein polygenetisches Modell zu entwickeln, um Hochrisikopatienten identifizieren zu können und neue Perspektiven in der Prävention und der Therapie des Prostatakrebses zu eröffnen.

Prostatakrebs stellt die häufigste bösartige Neubildung bei Männern in Österreich dar.

In dieser Diplomarbeit wurden vier Single Nucleotide Polymorphisms (SNPs), welche im AURKA-Gen lokalisiert sind, genotypisiert. Das Gen für AURKA (Aurora Kinase A), eine Serin/Threonin-Proteinkinase, befindet sich am Chromosom 20q13, einem bekannten Prostatakrebs-Suszeptibilitätslokus. AURKA ist an einigen ausschlaggebenden Vorgängen während der Mitose beteiligt. Genetische Polymorphismen im AURKA-Gen tragen zu interindividuellen Unterschieden in der Chromosomenstabilität bei und beeinflussen dadurch das Prostatakrebsrisiko.

Die Studienpopulation dieser fortlaufenden Studie besteht aus 1027 Prostatakrebsfällen und 550 Kontrollpatienten mit benigner Prostatahypoplasie (BPH). Vier ausgewählte tagging SNPs (rs2180691, rs8117896, rs1468055, rs1476394) wurden mit Hilfe des T5’ Nuclease TaqMan MGB Assay bestimmt. Zusammenfassend wurde kein signifikanter Zusammenhang zwischen den untersuchten SNPs und dem Prostatakrebsrisiko gefunden.

Jedoch zeigte eine Haplotypenanalyse ein reduziertes Prostatakrebsrisiko (OR: 0,7021; 95% CI: 0,49599-0,99384) für Männer unter 64 Lebensjahren mit einem GCAC Haplotyp.

Weitere Untersuchungen in größeren Studienpopulationen sind nötig, um den Einfluss von Polymorphismen im AURKA-Gen auf das Prostatakrebsrisiko festzustellen.

6.3. Curriculum vitae

Kontaktdaten

Doris Hummel
Blattgasse 8/18
1030 Wien
hummel_doris@gmx.at

Persönliche Daten

geboren am 18. 8. 1986 in Wels
ledig

Schule, Studium

2009–2010	Diplomarbeit am Institut für Krebsforschung
seit 2006	Spezialisierung auf den Studienzweig Anthropologie (Schwerpunkt Humangenetik)
seit 2004	Studium der Biologie an der Universität Wien
1996–2004	Gymnasium Dachsberg, Prambachkirchen
1992–1996	Volksschule Nord, Eferding

Ausbildungsbegleitende Tätigkeiten

09/2008	Neurogenetisches Labor des Wiener AKH (dreiwöchiges Praktikum)
11/2006	Ärztekammer Wien (Internetrecherche)