

Linköping University Medical Dissertations No. 1041

MOLECULAR GENETIC STUDIES  
ON PROSTATE AND PENILE CANCER

PATIYAN ANDERSSON



**Linköping University**  
**FACULTY OF HEALTH SCIENCES**

Division of Cell Biology  
Department of Clinical and Experimental Medicine  
Faculty of Health Sciences, Linköping University  
SE-581 85 Linköping

Linköping 2008

© 2008 Patiyān Andersson  
All rights reserved.

ISBN 978-91-7393-984-3  
ISSN 0345-0082

Printed in Sweden by LiU-Tryck, Linköping, 2008

“Set your goals in concrete  
and your plans in sand”

To Lauren

## SUPERVISOR

*Peter Söderkvist*, Professor  
Division of Cell Biology  
Faculty of Health Sciences  
Linköping University, Linköping

## OPPONENT

*Mattias Höglund*, Associate Professor  
Division of Clinical Genetics  
Lund University Hospital  
Lund University, Lund

## EXAMINATION BOARD

*Ulf Bergerheim*, Professor  
Division of Urology  
Danderyd Hospital  
Karolinska Institute, Stockholm

*Xiao-Feng Sun*, Professor  
Division of Oncology  
Faculty of Health Sciences  
Linköping University, Linköping

*Karin Roberg*, Associate Professor  
Division of Experimental Pathology  
Faculty of Health Sciences  
Linköping University, Linköping

## ABSTRACT

This thesis is comprised of two parts. In the first part we study the influence of four frequently disputed genes on the susceptibility for developing prostate cancer, and in the second part we attempt to establish a basic understanding of the molecular genetic events in penile cancer.

In a prostate cancer cohort we have investigated the relation of prostate cancer risk and single nucleotide polymorphisms (SNPs) in four different genes coding for the androgen receptor (*AR*), the vitamin D receptor (*VDR*), insulin (*INS*) and insulin receptor substrate 1 (*IRS1*). Despite strong biological indications of an involvement of these genes in prostate carcinogenesis, the results from different studies are contradictory and inconclusive.

The action of the *AR* varies between individuals in part owing to a repetitive CAG sequence (polyglutamine) in the first exon of the *AR* gene. The results presented in this thesis shows that in our cohort of prostate cancer patients the average number of repeats is 20.1, which is significantly ( $p < 0.001$ ) fewer repeats compared to healthy control individuals, where the average is 22.5 repeats. We find a 4.94 fold ( $p = 0.00003$ ) increased risk of developing prostate cancer associated with having short repeat lengths ( $\leq 19$  repeats), compared with long repeats ( $\geq 23$  repeats). In paper I we also study the *TaqI* polymorphism in the *VDR* gene, and find that it does not modify the risk of prostate cancer.

In the *INS* gene we study the +1127 *PstI* polymorphism and find no overall effect on the risk of prostate cancer. However, we do find that the CC genotype is associated with low grade disease defined as having a Gleason score  $\leq 6$  (OR=1.46;  $p = 0.018$ ). In the *IRS1* gene we study the G972R polymorphism and observe that the R allele is significantly associated with a 2.44 fold increased prostate cancer risk ( $p = 0.010$ ).

The knowledge of molecular genetic events in penile cancer is very scarce and to date very few genes have been identified to be involved in penile carcinogenesis. We chose therefore to analyse the penile cancer samples using genome-wide high-density SNP arrays. We find major regions of frequent copy number gain in chromosome arms 3q, 5p and 8q, and slightly less frequent in 1p, 16q and 20q. The chromosomal regions of most frequent copy number losses are 3p, 4q, 11p and 13q. We suggest four candidate genes residing in these areas, the *PIK3CA* gene (3q26.32), the *hTERT* gene (5p15.33), the *MYC* gene (8q24.21) and the *FHIT* gene (3p14.2).

The mutational status of the *PIK3CA* and *PTEN* genes in the PI3K/AKT pathway and the *HRAS*, *KRAS*, *NRAS* and *BRAF* genes in the RAS/MAPK pathway was assessed in the penile cancer samples. We find the *PIK3CA*, *HRAS* and *KRAS* genes to be mutated in 29%, 7% and 3% of the cases, respectively. All mutations were mutually exclusive. In total the PI3K/AKT and RAS/MAPK pathways were found to be activated through mutation or amplification in 64% of the cases, indicating the significance of these pathways in the aetiology of penile cancer.

## TABLE OF CONTENTS

ABSTRACT .....	5
ABBREVIATIONS.....	8
LIST OF PAPERS.....	10
INTRODUCTION .....	11
The process of cancer .....	11
Prostate cancer, the most common cancer in men .....	12
Penile cancer, a rare but psychologically devastating tumour form.....	14
GENES OF INTEREST FOR THIS THESIS.....	16
Androgen receptor gene.....	16
Vitamin D receptor gene.....	17
Insulin gene .....	18
Two central signalling pathways relaying growth signals .....	19
PI3K/AKT pathway.....	20
RAS/MAPK pathway .....	21
IRS1 gene.....	23
PIK3CA gene.....	23
PTEN gene.....	24
RAS genes .....	24
RAF genes .....	24
AIMS.....	26
STUDY POPULATIONS.....	27
Prostate cancer patients (I, II) .....	27
Penile cancer patients (III, IV) .....	27
Normal control subjects (I, II).....	27
METHODS.....	28
DNA isolation (I, II, III, IV).....	28
Whole genome amplification (III) .....	28
Single nucleotide polymorphism (I, II).....	28

Single stranded conformation analysis (III).....	29
Sanger chain termination DNA sequencing (III).....	29
Analysis of AR CAG microsatellite (I).....	30
Microarray (IV).....	30
Statistics (I, II).....	32
RESULTS AND DISCUSSION.....	33
Short CAG repeat length in the AR gene increases risk of prostate cancer .....	33
The <i>TaqI</i> polymorphism in the <i>VDR</i> gene does not modify risk of prostate cancer ..	35
The +1127 <i>PstI</i> polymorphism in the insulin gene does not modify risk of prostate cancer .....	36
The G972R polymorphism in the <i>IRS1</i> gene increases risk of prostate cancer .....	38
Mutational status of the <i>PIK3CA</i> gene in prostate cancer.....	39
The <i>PIK3CA</i> gene is frequently mutated in penile cancer.....	40
RAS mutations in penile cancer.....	41
Frequent copy number changes in penile cancer.....	42
Candidate genes for penile cancer.....	44
Relation of mutations in the <i>PIK3CA</i> and <i>RAS</i> genes and amplification of the <i>PIK3CA</i> gene .....	45
CONCLUSIONS .....	47
ACKNOWLEDGEMENTS.....	49
REFERENCES .....	52

## ABBREVIATIONS

AKT	a protein kinase
AP1	activator protein 1
AR	androgen receptor
BAD	Bcl2 agonist of cell death
BPH	benign prostatic hyperplasia
BRLMM	bayesian robust linear model with mahalanobis distance classifier
CKI	cyclin dependent kinase inhibitor
CNAT	copy number analysis tool
ddNTP	dideoxynucleotide
EGF	epidermal growth factor
ELK1	member of ETS oncogene family
ERK	extracellular signal-regulated kinase
FHIT	fragile histidine triad
FKHR	forkhead receptor
FOS	subunit of AP1
GAP	GTPase activating protein
GRB2	growth factor receptor bound protein 2
GSK3 $\beta$	glycogen synthase kinase-3 $\beta$
GTYPE	genotyping analysis software
HMM	hidden markov model
HPV	human papilloma virus
hTERT	catalytical subunit of human telomerase
IKK	inhibitor of NF $\kappa$ B
INS	insulin
IRS	insulin receptor substrate
JUN	subunit of AP1
MAPK	mitogen activated protein kinase
MDM2	mouse double minute 2
MEK	mitogen-activated and extracellular signal-regulated-kinase kinase
mTOR	mammalian target of rapamycin
MYC	viral oncogene homolog
NF $\kappa$ B	nuclear factor $\kappa$ B
OR	odds ratio
PCR	polymerase chain reaction
PDK	3-phosphoinositol-dependent kinase

PH-domain	pleckstring homology domain
PI3K	phosphatidylinositol-3 kinase
PIK3CA	catalytical subunit of PI3K
PIP <sub>2</sub> /PIP <sub>3</sub>	phosphatidylinositol di/tri-phosphate
PSA	prostate specific antigen
PTEN	phosphatase and tensin homolog deleted on chromosome 10
RAF	murine sarcoma viral oncogene homolog
RAS	rat sarcoma viral oncogene homolog
RHEB	RAS homologue enriched in brain
RR	relative risk
SD	standard deviation
SHC	SH2-containing protein
SNP	single nucleotide polymorphism
SOS1	son of sevenless 1, a guanine nucleotide exchange factor
SSCA	single stranded conformation analysis
TSC2	tuberous sclerosis complex 2
TURP	transurethral resection of the prostate
UTR	untranslated region
VDR	vitamin D receptor
VNTR	variable number of tandem repeat
WGA	whole genome amplification

## LIST OF PAPERS

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.

- I            **Patiyan Andersson**, Eberhard Varenhorst, Peter Söderkvist (2006)  
*Androgen receptor and vitamin D receptor gene polymorphisms and prostate cancer risk.*  
European Journal of Cancer, 42(16), 2833-2837
  
- II           **Patiyan Andersson**, Eberhard Varenhorst, Peter Söderkvist  
*Association studies on INS and IRS1 polymorphisms: IRS1 G972R is associated with increased prostate cancer risk.*  
Manuscript submitted to Prostate cancer and prostatic diseases
  
- III          **Patiyan Andersson**, Aleksandra Kolaric, Torgny Windahl, Peter Kirrander, Peter Söderkvist, Mats G Karlsson (2007)  
*PIK3CA, HRAS and KRAS gene mutations in human penile cancer*  
Journal of Urology, in press
  
- IV          **Patiyan Andersson**, Aleksandra Kolaric, Torgny Windahl, Peter Kirrander, Ove Andrén, Jon Jonasson, Mats G Karlsson, Peter Söderkvist  
*Genome-wide analysis of penile cancer using high-density single nucleotide polymorphism arrays*  
Manuscript

# INTRODUCTION

## The process of cancer

Cancer is a process where the accumulation of several genetic alterations leads to a malignant transformation. Genetic alterations occur randomly at any location in the genome and will most often damage parts that are not of significant importance for the propagation of the cell. However, should it affect a position that will give the cell a replicative advantage, which can be passed on to the daughter cells, it may initiate the formation of a tumour. A number of key features are required for a tumour to develop; the tumour needs to secure continuous growth signals, block growth inhibitory signals, evade apoptosis, create a blood supply to obtain nutrients, acquire the ability to replicate indefinitely and infiltrate surrounding tissue to enable metastatic spread (HANAHAN and WEINBERG 2000).

Each cell in the body has a specific role and cells will replicate only when new cells are needed to maintain the homeostatic balance (HANAHAN and WEINBERG 2000). Each replication represents a risk where even small mistakes such as the insertion of a different base resulting in an altered sequence can lead to severe consequences. Therefore the entry into the cell cycle is tightly controlled. There are additional check points within the cycle which will initiate termination of the cell should something go wrong during the replication. When a disabled control feature remains unchecked and the cell continues through the cell cycle despite flaws, the cell acquires the first step in carcinogenesis. The homeostatic balance can be disrupted both through increased proliferation as well as decreased cell death.

The replicative potential of a cell is limited to maintain optimal performance of the cell throughout its lifetime (HANAHAN and WEINBERG 2000). The number of divisions which a cell has gone through is counted by a molecular clock. The chromosome ends, telomeres, are shortened in each round of division due to an inability of the DNA polymerase to replicate the outermost parts of the 3' ends of the DNA strand. After approximately 50 divisions the loss starts to affect important chromosomal material and challenges the integrity of the chromosome upon which the cell enters apoptosis and is terminated. Human stem cells express telomerase, an enzyme complex that add the missing sequences and thus prevent shortening of the telomeres. Most human cancer cells have activated this dormant function and thus ensuring the maintenance of the tumour through indefinite replication.

In order for a tumour to grow to a size beyond a few millimetres it needs to recruit and establish a blood supply that can provide the cells with nourishment and oxygen. The tumour will form an internal vascular network and connect it to existing adjacent vessels. Tumours that are unable to form a capillary network may enter a dormant state at which balance exists between formation of new cells and cell death due to lack of nutrients and oxygen.

The final stage of the carcinogenic process is when the tumour cells acquire the ability to ignore the constraints of the neighbouring cells, detach and infiltrate the surrounding tissue (HANAHAN and WEINBERG 2000). At this point a migrating tumour cell may enter the blood stream and be transported to other locations in the body where it can start to form a new tumour, a metastase.

### **Prostate cancer, the most common cancer in men**

Prostate cancer is one of the major medical problems facing the male population in Europe, North America and Australia. In Sweden 2006, the incidence was 206/100,000 males, equivalent of 9500 new cases, representing approximately 35% of the cancers diagnosed in males per year (Cancer Incidence and Mortality report for 2006, National Board of Health and Welfare, Sweden). Annually approximately 2300 Swedish men will die from their prostate cancer. The causes of prostate cancer remain an enigma, where few major risk factors can be identified. To date the only established risk factors are increased age and a positive family history. It is considered a cancer of the elderly men, usually affecting men over 60 years of age and is rare in those under 40 years. A meta-analysis of 32 population-based studies showed a 2.46 fold risk of prostate cancer for family members of a prostate cancer patient (ZEEGERS *et al.* 2003). It has been speculated how much of this can be attributed shared genetic traits and how much of it reflects a shared environment. However, hereditary prostate cancer accounts for about 9% and the majority of prostate cancers are sporadic (CARTER *et al.* 1992). While the prevalence of microscopic cancer foci detected at autopsy is similar and high among men of different ethnic groups, the clinical incidence is low among Asians and highest in Scandinavians and African-Americans (HAAS and SAKR 1997).

Two classifications are used to describe prostate cancer. The Union International Contra Cancer (UICC) 2002 tumour, node, metastasis (TNM) classification is a

common classification used for malignant tumours. The second classification system, Gleason score, is specific for grading of adenocarcinoma of the prostate (GLEASON and MELLINGER 1974). The system is based on the degree of differentiation, and describes the two most common patterns (grade 1-5) of tumour growth seen, yielding a score between 2 and 10, where 2 being the least aggressive and 10 the most aggressive.

Transurethral resection of the prostate (TURP) is a treatment used to ease some lower urinary tract symptoms more commonly caused by benign prostatic hyperplasia (BPH) rather than prostate cancer. Prostate cancer detected incidentally at TURP are designated T1, and can be subclassified as T1a or T1b, depending on whether less than 5% (T1a) or more than 5% (T1b) of the resected tissue contains tumour cells. While T1a tumours are fairly indolent with a median time to progression of disease of 15 years, T1b tumours have a much quicker median time to progression of 5 years. Tumours that can be detected by rectal palpation are designated T2 and 90-100% of these will have progressed within 15 years if left untreated (ALBERTSEN *et al.* 1995). Tumours that can be determined by palpation to be extending outside the prostatic capsule are designated T3.

Prostate cancer is an androgen driven tumour form, a trait that can be used both for screening purposes and as a target for early therapies. Prostate specific antigen (PSA) is a glycoprotein that is produced by the prostatic epithelium (BRAWER 1999). An increase in PSA can be an indication of abnormal prostatic behaviour and is used in some countries as a non-invasive means of screening for prostate cancer, normally a threshold of 4 ng/ml is used. What level of PSA that can be considered normal will depend on the age of the patient. There are several controversies being discussed regarding screening for prostate cancer using PSA, and the screening programs vary greatly from country to country. There are two major risks with PSA screening. Firstly it has been shown that 15% of men diagnosed with prostate cancer had a PSA level of <3 ng/ml, and that 15% of these had a Gleason score of  $\geq 7$  (THOMPSON *et al.* 2004). Secondly, many of the diagnosed tumours would remain indolent for the remaining lifespan of the patient, and that higher morbidity may be associated with the treatment than with the tumour itself. Such overtreatment would not only reduce the quality of life for the patient but may also divert significant amounts of funding from other areas of health care where it could be better used. For diagnostic purposes, regular PSA measurements are used in conjunction with digital rectal examination and transrectal ultrasound guided biopsy of the prostate.

The introduction of PSA screening programs in the late 1980s led to a dramatic increase in the number of cases diagnosed annually, especially of patients with a low-risk cancer (COOPERBERG *et al.* 2003). This has meant a shift from diagnosing mainly high-risk, late stage prostate cancer to detecting low-risk disease, which of course improves the chances of treatment. The only concern is the risk of over diagnosis, which might lead to over treatment of cancer that would not become symptomatic. A puzzling observation is that despite the definite shift towards diagnosing early stage cancer the mortality rates for men with prostate cancer does not decrease, but remains at the same level.

Treatment of prostate cancer will heavily depend on the stage of the diagnosed tumour and of the age and life expectancy of the patient. Generally if the patient is asymptomatic, with a tumour that has a Gleason score of  $\leq 7$  and  $< 10$  year life-expectancy, watchful-waiting is adopted, where the patient is closely monitored but no treatment is given (HEIDENREICH *et al.* 2008). Should the tumour develop a new assessment is made. Treatment for prostate cancer can be: radical prostatectomy, radiotherapy, hormonal therapy or a combination of them.

A large number of individuals would benefit if risk factors that increase the susceptibility to develop prostate cancer could be established, which could aid in the early detection of the disease which is crucial for successful treatment. Numerous studies have been conducted on the molecular genetic aetiology of the disease. The suggested involvement of certain genes in the development and progression of prostate cancer has been contradictory, making interpretation of the significance of these genes for prostate cancer difficult. In this thesis we study and discuss the role of four susceptibility genes commonly debated within the field of prostate cancer research.

### **Penile cancer, a rare but psychologically devastating tumour form**

Penile cancer, in contrast to prostate cancer, is among the rarest cancer forms in men in Europe. In Sweden the annual incidence is 2.2/100,000 males, equivalent of approximately 100 cases (PERSSON *et al.* 2007). Much higher incidence can be observed in some African, Asian and South American countries (MICALI *et al.* 2006). Early diagnosis is essential for implementation of life saving efforts, and will also enable the use of functionally and cosmetically acceptable treatments.

Penile cancer will in 95% of the cases manifest in the form of squamous cell carcinoma (MICALI *et al.* 2004). The most established risk factor is lack of circumcision, as many studies have reported a very low incidence of penile cancer among circumcised men. The preventive effect is however only evident in men that were circumcised at birth or early in life, whereas late or adult circumcisions seemed to be ineffective (TSEN *et al.* 2001). Infection with human papilloma virus (HPV) has also been suggested as a risk factor. However, in contrast to cervical carcinoma where HPV is found in approximately 95% of the cases, penile cancers show HPV infection in 30-60% of the cases (RUBIN *et al.* 2001). Among 25% of patients with penile cancer, a history of phimosis has been reported (TSEN *et al.* 2001). Chronic inflammatory conditions have also been suggested to be a risk factor (DILLNER *et al.* 2000). Collectively these results point towards poor hygiene being a major determinant for the risk and development of penile cancer.

Penile cancer starts as a lesion on the surface of the glans, foreskin or penile shaft. It can remain localized for long periods, but if left untreated will grow, with eventual corporal or urethral invasion. Tumour cells may spread via the penile lymphatics that drain into the superficial and deep inguinal lymph nodes and then to the iliac nodes (DEWIRE and LEPOR 1992). In late stages of the disease metastatic spread to liver and lung is common (KHANDPUR *et al.* 2002).

The classification systems used for describing penile cancers are; the Broder's classification system (tumours are designated I through to IV) (LUCIA and MILLER 1992) and the UICC-TNM system for staging of the tumour. The stage at diagnosis appears to be the most important prognostic indicator of survival, which emphasizes the essentiality of early detection and diagnosis.

The mainstay of treatment for penile carcinomas is surgical intervention, with focus on the primary tumour as well as regional lymph nodes. The extent of surgical removal may range from local excision to partial or total penectomy. The ideal surgical outcome is to eliminate the disease while preserving sexual and urinary function, although this is not always possible due to the extent of the tumour. Radical procedures, such as total penectomy, may be devastating for patients, especially younger, sexually active individuals.

Little research effort has gone into assessing the molecular genetic background of penile cancer. In part this is likely due to the relative rarity of this tumour form, and following with that the lack of sufficiently large cohorts to conduct conclusive studies.

## GENES OF INTEREST FOR THIS THESIS

### Androgen receptor gene

The normal growth and development of the prostate is stimulated by androgens such as testosterone and 5 $\alpha$ -dihydrotestosterone. These hormones exert their growth stimulatory effect through binding to the intracellular androgen receptor (AR) (KOKONTIS and LIAO 1999). The AR is a member of the nuclear steroid hormone receptor family, which following homodimerization will regulate the transcription of many downstream genes, for example PSA. The AR gene is located on Xq11-12 and is composed of 8 exons. The protein contains a 5' DNA-binding domain and a 3' ligand-binding domain separated by a hinge region (MCEWAN 2004). The 5' part of the protein contains the activation factor 1 (AF1), which is the major transactivation domain. The normal transcriptional activity of the AR is androgen dependent, where the ligand binding domain inhibits the DNA binding domain in absence of a ligand. Deletions in the ligand binding domain can abolish this control function and result in a constitutively active receptor (JENSTER *et al.* 1991).

Located 5' is a highly variable polyglutamine sequence (coded by CAG repeats), normally ranging from 14-35 repeats, which has been shown to alter the activity of the AR (SARTOR *et al.* 1999). Abnormally long sequences of 40-62 repeats, results in spinal and bulbar muscular atrophy also known as Kennedy's disease (LA SPADA *et al.* 1991). Within the normal range of repeats there is also a difference in the transcriptional activity and shorter repeats have been associated with higher activity (CHAMBERLAIN *et al.* 1994; IRVINE *et al.* 2000; KAZEMI-ESFARJANI *et al.* 1995). In addition to this there has been an association between the number of repeats and several androgen related clinical conditions: long repeat sequences adversely affect fertility and spermatogenesis (YOSHIDA *et al.* 1999), while fewer repeats is associated with increased risk of baldness (SAWAYA and SHALITA 1998) and benign prostatic hyperplasia (GIOVANNUCCI *et al.* 1999; MITSUMORI *et al.* 1999; SHIBATA *et al.* 2001). Furthermore, the observation of racial difference in CAG repeat number, where African-American men tend to have fewer repeats and Asian men tend to have more repeats, correlate well with the risk of prostate cancer observed in these groups (IRVINE *et al.* 1995; PRICE *et al.* 2004; ROSS *et al.* 1998).

The extent of involvement of the CAG repeat in development of prostate cancer has been investigated in many studies. However, no conclusive evidence for such an

association has yet been presented, since many studies contradict each other. The main difficulty in the research field studying the CAG repeat in the AR gene is the lack of uniform cut-offs for subdivision of the number of repeats into relevant groups. There might be a detection bias associated with this problem, with a potential tendency to report associations based of subdivisions that give the most attractive result. A meta-analysis of the to-that-date available results was published in 2004 by Zeegers and colleagues who reported a modest summary 1.19 fold risk of prostate cancer associated with shorter CAG repeats (ZEEGERS *et al.* 2004). We sought therefore to assess the effect of CAG repeat length on the risk of prostate cancer, using different cut-offs and also comparing mean lengths in cases and controls.

### **Vitamin D receptor gene**

The first epidemiological indication that vitamin D might reduce the risk of prostate cancer was the observation that mortality rates in prostate cancer had an inverse correlation to ultraviolet radiation exposure, which is essential for the production of vitamin D (HANCHETTE and SCHWARTZ 1992; SCHWARTZ and HULKA 1990). Corder *et al* further reported that men with prostate cancer had lower levels of the active form  $1\alpha,25$ -dihydroxyvitamin D<sub>3</sub> ( $1,25(\text{OH})_2\text{D}_3$ ) than age and race matched healthy controls (CORDER *et al.* 1993).  $1,25(\text{OH})_2\text{D}_3$  has been shown to have an anti-proliferative and differentiating effect on prostate cells *in vitro* (HANCHETTE and SCHWARTZ 1992; MILLER *et al.* 1995; PEEHL *et al.* 1994; SKOWRONSKI *et al.* 1993).

The actions of  $1,25(\text{OH})_2\text{D}_3$  are mediated through the vitamin D receptor (VDR), a member of the nuclear steroid hormone receptor family. Upon activation the VDR will heterodimerize with the retinoid X receptor and drive transcription of several target genes in many tissues, including the prostate (MINGHETTI and NORMAN 1988). It has been postulated that disruptions leading to decreased receptor function could increase the susceptibility to prostate cancer, due to decreased effectiveness of vitamin D signalling.

The VDR gene is located on chromosome 12q12-q14 and is composed of 9 exons. In VDR there are 4 single nucleotide polymorphisms (SNPs) which can be detected using restriction enzymes, *FokI* in exon 2, *BsmI* and *ApaI* in intron 8 and *TaqI* in exon 9, where presence of the restriction enzyme recognition sites is designated with a minor letter (*f*, *b*, *a*, *t*) and absence of sites designated with capital letters (*F*, *B*, *A*, *T*). Additionally there is a poly(A) repetitive sequence (which can either be Long or Short)

located in the 3' untranslated region, accounting for a fifth polymorphic site. The 3' located *BsmI*, *TaqI*, *ApaI* and poly(A) polymorphisms are in tight linkage disequilibrium and form the major haplotypes, *baTL* and *BAtS* (MORRISON *et al.* 1994). In a luciferase reporter assay the *BAt* haplotype was shown to have greater transcriptional activity than the *baT* haplotype. The *FokI* polymorphism located more 5' in the gene is not part of this haplotype and presence of the site will result in a 9 bp longer protein through creation of an alternative starting codon located 5' of the original start (GROSS *et al.* 1996). The *f* allele has been shown to result in decreased VDR transcriptional activity compared to the *F* allele (ARAI *et al.* 1997).

The effect of these polymorphisms on the risk of developing prostate cancer has been investigated in many studies with inconsistent results, possibly due to heterogeneous study designs (BLAZER *et al.* 2000; CORREA-CERRO *et al.* 1999; FURUYA *et al.* 1999; HABUCHI *et al.* 2000; KIBEL *et al.* 1998; MA *et al.* 1998; TAYLOR *et al.* 1996; WATANABE *et al.* 1999). Given the strong indications, at the time of the initiation of the project (in 2001) leading up to paper I, of an involvement of the vitamin D system in prostate cancer we sought to investigate the influence of the *TaqI* polymorphism, and thereby the effect of the whole 3' haplotype *BsmI/TaqI/poly(A)*, on prostate cancer risk.

## Insulin gene

Insulin is an important factor in cell growth and development, and is involved in processes regulating cell proliferation, differentiation, apoptosis and transformation. Deregulation of these characteristics are hallmarks of cancer and therefore alterations in the insulin signalling pathways been implicated in carcinogenesis (GRIMBERG and COHEN 2000). A possible association of insulin physiology with increased prostate cancer risk was first suggested by Hsing and colleagues in 2000, who observed that abdominal adiposity, as measured by waist-hip-ratio, was a risk factor for prostate cancer (HSING *et al.* 2000a). The following year the group published data showing that higher levels of circulating insulin were associated with an increased prostate cancer risk (HSING *et al.* 2001). Men in the highest tertile of insulin levels had a 2.56 fold (95% CI = 1.38-4.75) risk of prostate cancer compared to men in the lowest tertile. Studies have also observed a relation between serum insulin levels and advanced tumour stage and risk of prostate cancer recurrence (LEHRER *et al.* 2002a; LEHRER *et al.* 2002b).

An elevated serum insulin level might be caused by genetic variations in the insulin gene. The insulin (*INS*) gene is located on 11p15.5 and consists of 3 exons. Adjacent to

the 5' promoter region, lie a variable number of tandem repeat (VNTR) region, which is believed to have a direct effect on the regulation of the insulin gene (KENNEDY *et al.* 1995). The VNTR polymorphism can be classified into two main groups, short *class I* alleles (28-60 repeats) and long *class III* alleles (138-159 repeats), occurring at frequencies of 70% and 30% respectively (STEAD and JEFFREYS 2000). Intermediate *class II* alleles are rare. *Class I* alleles have been associated with increased insulin levels and risk of type 1 diabetes (BENNETT and TODD 1996; LE STUNFF *et al.* 2000; LUCASSEN *et al.* 1995). The VNTR is part of a 4.1 kb region, spanning the entire *INS* gene and flanking regions, that includes 10 SNPs which are all in tight linkage disequilibrium, such as they form two major haplotypes (COX *et al.* 1988; LUCASSEN *et al.* 1993). One of the SNP markers is +1127 *Pst*I (C/T), and can in Caucasians act as a surrogate for analysis of the VNTR region due to the tight linkage disequilibrium. The +1127 *Pst*I C allele is linked to the VNTR *class I* allele, and hence also result in increased insulin levels (LUCASSEN *et al.* 1993). The SNP is located in the 3' UTR which also might have an effect on the mRNA stability of the transcript. Our study was based on the observation of Ho *et al* that the CC genotype at the +1127 *Pst*I site was associated with a 3.14 fold ( $p = 0.008$ ) increased risk of prostate cancer (HO *et al.* 2003). Additionally the study showed that the CC genotype was associated with late age of onset and low grade tumours, characteristic for majority of prostate cancers with a high prevalence of indolent cancers and a drastic elevated incidence with age.

## Two central signalling pathways relaying growth signals

Two central signalling pathways are situated downstream of membrane bound receptor tyrosine kinases such as insulin- or growth factor receptors, and relay growth signals from extracellular ligands and orchestrates the cellular response to these. These two signaling pathways are; the phosphatidylinositol-3 kinase (PI3K) and AKT pathway, and the RAS and mitogen activated protein kinase (MAPK) pathway (Figure 1). The importance of these two pathways in carcinogenesis has been shown to be significant in numerous *in vitro* and *in vivo* studies as well as epidemiological studies. Disruptions of components in the pathways have been shown to be frequent in a wide variety of cancer forms. In the light of recent studies it may be argued that these two pathways may be treated as one, due to complex interactions at multiple points between the pathways. However, for the sake of familiarity and to easier relate to previously published studies, the pathways will be presented as separate entities where points of interaction will be stated. Some of the components included in these pathways are of special interest for this thesis and will hence be presented in more detail.

## PI3K/AKT pathway

There are three major classes of phosphatidylinositol-3 kinases (PI3K), but to date only the class 1A PI3K subgroup has been implicated in carcinogenesis. Class 1A PI3K are heterodimeric proteins composed of a catalytic subunit (p110 $\alpha$ ) and a regulatory subunit (p85) (VANHAESEBROECK and WATERFIELD 1999). The regulatory subunit p85 interacts either directly with the phosphotyrosine residues on activated RTKs or via the intermediate proteins termed insulin receptor substrates, such as IRS1 and IRS2. The catalytic subunit p110 $\alpha$  also contains a RAS binding domain, and can be directly activated by active RAS proteins (RODRIGUEZ-VICIANA *et al.* 1994; RODRIGUEZ-VICIANA *et al.* 1996). The substrate for PI3K is the membrane bound phosphatidylinositol 4,5 diphosphate (PIP<sub>2</sub>) which is converted to the active second messenger, phosphatidylinositol 3,4,5 triphosphate (PIP<sub>3</sub>) (CORVERA and CZECH 1998). Dephosphorylation of PIP<sub>3</sub> by the 3-phosphatase PTEN will result in the second messenger being returned to its inactive state, PIP<sub>2</sub> (MAEHAMA and DIXON 1998). Active PIP<sub>3</sub> will recruit proteins with a pleckstrin homology (PH) domain to the cell membrane and lead to their activation (CORVERA and CZECH 1998). The foremost protein recruited by PIP<sub>3</sub> is the AKT protein, a serine/threonine kinase, which is also known as protein kinase B (PKB). At the cell membrane AKT is activated by phosphorylation on Thr308 by 3-phosphoinositide-dependent protein kinase-1 (PDK1), another PH-domain containing serine/threonine kinase. Maximal activation of AKT is accomplished by additional phosphorylation on Ser473 by PDK2. AKT is a very central protein and will regulate a wide range of downstream targets controlling cell survival, cell growth, cell cycle progression and cell metabolism.

Promoted cell survival can be achieved by AKT through direct effects by inactivating phosphorylation of pro-apoptotic proteins such as BAD, caspase-9 and FKHR (member of the Forkhead family of transcription factors) (BRUNET *et al.* 1999; CARDONE *et al.* 1998; DATTA *et al.* 1997). Cell survival can also be achieved more indirectly, leading to increased levels of free NF $\kappa$ B or decreased levels of the tumour suppressor p53 (MAYO and DONNER 2001; ROMASHKOVA and MAKAROV 1999).

AKT can drive the cell cycle forward through phosphorylating inactivation of the kinase activity of glycogen synthase kinase-3 $\beta$  (GSK3 $\beta$ ), which leads to a decreased degradation of Cyclin D1 and allowing it to accumulate (DIEHL *et al.* 1998). AKT also negatively regulates the activity of cyclin dependent kinase inhibitors (CKIs), such as p21 and p27 (GRAFF *et al.* 2000; LAWLOR and ROTWEIN 2000).

In addition to increasing proliferation, AKT can also affect cell growth, leading to increased cell mass or size. A central protein in this aspect is mTOR (the mammalian target of rapamycin), a serine/threonine kinase that is sensitive to and will react on the availability of nutrients (NAVE *et al.* 1999). AKT initiates activation of mTOR pathway by inactivating phosphorylation of the TSC2 (tuberous sclerosis complex 2, also known as tuberin) and thereby liberating RHEB (RAS homologue enriched in brain) leading to activation of the mTORC1-raptor protein complex (MANNING and CANTLEY 2003). The mTORC1-raptor complex will regulate cell growth by activation of p70 S6 kinase and inactivation of 4EBP1, leading to the translation of mRNAs and transcription of critical growth genes. mTOR is also an important mediator of metabolic signals in the cell reacting on the levels of ATP (DENNIS *et al.* 2001).

### RAS/MAPK pathway

Upon binding of a growth stimulating ligand to a membrane bound RTK, a protein complex will form on the inside containing adaptors such as SHC (SH2-containing protein), GRB2 (growth-factor-receptor bound protein 2) and SOS1 (a guanine nucleotide exchange factor) (BOGUSKI and MCCORMICK 1993). SOS1 will initiate the change of RAS from its inactive GDP bound state (RAS-GDP) to the active GTP bound state (RAS-GTP). RAS-GTP can be returned to the inactive RAS-GDP form by GTPase activating proteins (GAPs) (DONOVAN *et al.* 2002). Active RAS can subsequently interact with and activate a number of downstream proteins, among these is RAF (CHONG *et al.* 2003). Downstream of RAF is the mitogen-activated and extracellular-signal-regulated-kinase kinase (MEK) which in turn activates the extracellular signal-regulated kinase (ERK) (MERCER and PRITCHARD 2003). This cascade of signaling, RAS-RAF-MEK-ERK, will result in a proliferative cell response. As previously mentioned RAS-GTP can also bind to and activate the catalytic p110 $\alpha$  subunit of PI3K, and thereby initiate that signal cascade (RODRIGUEZ-VICIANA *et al.* 1994; RODRIGUEZ-VICIANA *et al.* 1996).

ERK phosphorylates cytosolic and nuclear proteins, such as JUN and ELK1 (a E26 transformation specific sequence, ETS), the latter which drives the transcription of FOS. JUN and FOS will together create the activator protein 1 (AP1) transcription factor leading to the expression of proteins that control cell cycle progression, such as cyclin D1 (YORDY and MUISE-HELMERICKS 2000). ERK can also inactivate TSC2 by phosphorylation similarly to the action previously described for AKT (MA *et al.* 2005).

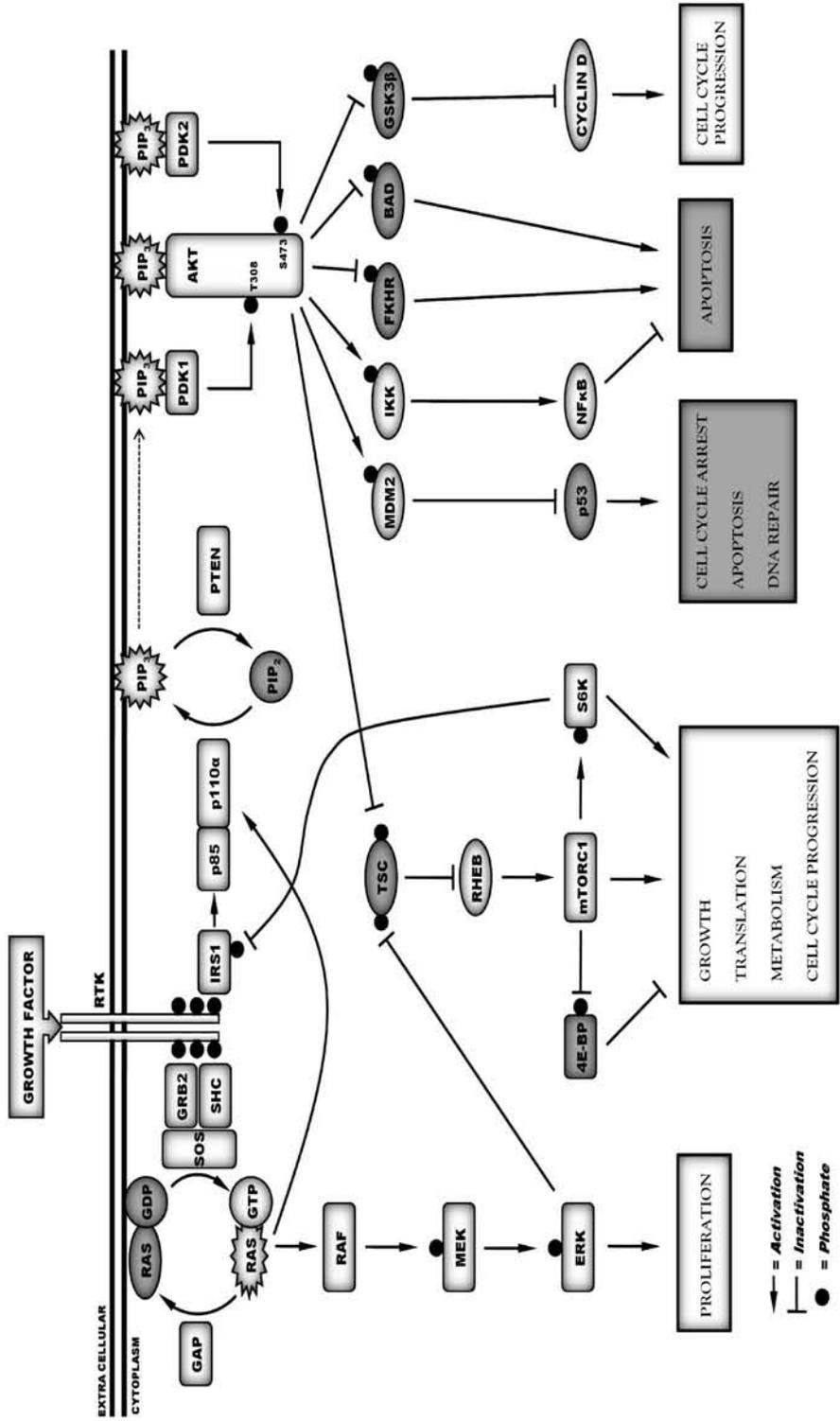


Figure 1. Overview of the PI3K/AKT and RAS/MAPK pathways

I will in the next sections go into more detail on genes coding for certain proteins in these signalling cascades and the research linking dysregulation of these to carcinogenesis.

### **IRS1 gene**

Insulin receptor substrate 1 and 2 (IRS1 and 2) can be found expressed in almost all tissues and cells in the body. IRS1 controls body growth and peripheral insulin action, while IRS2 regulate body weight control and glucose homeostasis (SCHUBERT *et al.* 2003). Polymorphisms in these two genes can affect the level of activity of the proteins. The G972R (amino acids) polymorphism in the *IRS1* gene has been associated with insulin resistance and type 2 diabetes (FEDERICI *et al.* 2003; JELLEMA *et al.* 2003), and additionally also with cancer. Slattery *et al* observed that having at least one copy of the R allele was associated with a 1.4 fold (95% CI = 1.1-1.9) increase in risk of developing colon cancer (SLATTERY *et al.* 2004). Following this study, Neuhausen *et al* showed that the *IRS1* G972R GR/RR genotypes were associated with a 2.8 fold (95% CI = 1.5-5.1) increased risk for prostate cancer (NEUHAUSEN *et al.* 2005). The study also showed a significant association of the GR/RR genotypes with a more advanced Gleason score ( $p=0.001$ ).

### **PIK3CA gene**

The *PIK3CA* gene, located on 3q26.32 and consisting of 20 exons, codes for the catalytic p110 $\alpha$  subunit of PI3K and has emerged to be one of the most frequently mutated genes in cancer. The *PIK3CA* gene was known to be overactive due to amplification (MA *et al.* 2000; RACZ *et al.* 1999; REDON *et al.* 2001; SHAYESTEH *et al.* 1999), but it was the pioneering work of Samuels and colleagues in 2004 that turned the attention to the mutational status of this gene (SAMUELS *et al.* 2004). Following the initial study there has been a constant flow of reports of mutations identified in this gene in many cancer forms, with the highest frequencies observed in cancers of the colon, breast and liver [reviewed in (KARAKAS *et al.* 2006)]. Samuels *et al* observed in their study that mutations tended to cluster to specific codons located in exon 9 and 20, and the studies following upon theirs shows that this observation hold true. The most commonly mutated codons are 542 and 545 in exon 9 and codon 1047 in exon 20.

## **PTEN gene**

The *PTEN* (phosphatase and tensin homolog deleted on chromosome 10) gene is located on 10p23 and is composed of 9 exons. *PTEN* was originally discovered as a tumour suppressor protein in breast, prostate cancer and glioblastomas (HAAS-KOGAN *et al.* 1998; LI *et al.* 1997), and as mutations in this protein subsequently was observed in other cancer forms the importance of this gene for the carcinogenic process became undisputable. Germline mutations in the gene are associated with hereditary cancer syndromes, including Bannayan-Zonana and Cowden disease (LIAW *et al.* 1997; MARSH *et al.* 1997). Mutations in the *PTEN* gene are usually found in exons 5 through to 8 and has been observed in glioblastomas, ovarian cancer, endometrial cancer, hepatocellular cancer, melanoma, thyroid cancer, prostate cancer and lymphoid cancer (CAIRNS *et al.* 1997; CELEBI *et al.* 2000; DAHIA *et al.* 1997; HALACHMI *et al.* 1998; HSIEH *et al.* 2000; KAWAMURA *et al.* 1999; SAITO *et al.* 2000; SAKAI *et al.* 1998; WANG *et al.* 1997; YOKOYAMA *et al.* 2000).

## **RAS genes**

The *RAS* (Rat sarcoma virus oncogene homologue) family of proto-oncogenes consists of three members, *HRAS*, *KRAS* and *NRAS*. The *HRAS*, *KRAS* and *NRAS* genes are located on 11p15.5, 12p12.1 and 1p13.2 respectively. Point mutations in *RAS* genes can be found in approximately 30% of human cancers. Specific *RAS* genes are mutated in different cancer forms: *HRAS* mutations are common in bladder cancer, *NRAS* mutations frequently occur in melanoma and myeloid disorders whereas in tumours from the exocrine pancreas more than 80% carry a mutated *KRAS* gene [reviewed in (BOS 1989)]. In most cases, the somatic mutations observed in the *RAS* genes affect codons 12, 13 or 61.

## **RAF genes**

Similarly to the *RAS* (murine sarcoma viral oncogene homolog) gene family, the *RAF* family is composed of three members, *ARAF*, *BRAF* and *CRAF* (which is also known as *RAF-1*). The *RAF* genes code for serine/threonine kinases that are regulated by the binding of activated *RAS*. Of the three family members, only *BRAF* is found to be mutated in cancers (DAVIES *et al.* 2002; FRANSEN *et al.* 2004). Mutations in the *BRAF*

gene are thought to be present in approximately 7% of human cancers (DAVIES *et al.* 2002; FRANSEN *et al.* 2004). The highest frequency of *BRAF* mutations can be found in malignant melanoma, papillary thyroid cancer, colorectal cancer and serous ovarian cancer [reviewed in (GARNETT and MARAIS 2004)].

## AIMS

The aims of the first two papers in this thesis were to study the influence of four much-disputed genes on the susceptibility for prostate cancer. The aims of the third and fourth paper were to establish a basic understanding of molecular genetic events in the poorly studied field of penile cancer.

The specific aims for this thesis were:

- To assess whether the length of the CAG repetitive sequence in exon 1 of the *AR* gene has an influence on the risk of developing prostate cancer.
- To assess whether the *TaqI* polymorphism in the *VDR* gene increase the susceptibility for prostate cancer.
- To assess whether the +1127 *PstI* polymorphism in the insulin gene affect the susceptibility for prostate cancer.
- To assess whether the G972R polymorphism in the *IRS1* gene affect the susceptibility for prostate cancer.
- To screen the mutational hotspots in the *PIK3CA*, *PTEN*, *HRAS*, *KRAS*, *NRAS* and *BRAF* genes and study the significance of these in penile squamous carcinomas.
- To determine chromosomal regions frequently affected by copy number aberrations in penile squamous cell carcinoma using a genome wide approach.
- To attempt to identify plausible candidate genes located in the observed regions of frequent copy number change.

## STUDY POPULATIONS

### **Prostate cancer patients (I, II)**

In 1987 the National Cancer Registry in Sweden was expanded with a separate registry for prostate carcinoma in the South-East Region of Sweden (AUS *et al.* 2005). Beside date of diagnosis, municipality, county of residence and tumour type, the extended registry included the collection of data on tumour characteristics, treatment and survival.

The prostate tumours used in this thesis were discovered and histologically verified by examination of tissue chips collected through TURP during the time period of 1987-1996. These patients are part of and representative of a random selection of a larger cohort of approximately 1300 cases.

### **Penile cancer patients (III, IV)**

Despite penile cancer being a rare tumour form, the Urology department at Örebro university hospital has managed to collect a cohort consisting of 200 tumours that have been surgically removed. We have used 28 of these tumours in the pilot studies included in this thesis. These tumours were collected between the years 2000-2006.

### **Normal control subjects (I, II)**

Healthy blood donors have been randomly chosen from the population register to establish a DNA bank comprising 800 subjects, with an even distribution of ages and sex. This cohort is used to determine allele frequencies of polymorphisms in a normal Swedish population. For PAPER I and II part of this population was used consisting of men older than 54 years, corresponding as closely as possible to the age distribution in the prostate cancer population.

## METHODS

### DNA isolation (I, II, III, IV)

DNA was extracted from formalin-fixed paraffin-embedded prostatic tissue. Tissue sections were deparaffinized with xylene followed by degradation of cell membranes and proteins by multiple additions of proteinase K over a period of 48-72 hours. Metal ions were removed by boiling in Chelex 100, since these can interfere with downstream PCR reactions. DNA was purified with phenol-chloroform extractions, precipitated with sodium acetate and ice-cold ethanol, dried and diluted to a final concentration of 50ng/ $\mu$ l.

DNA was isolated from snap frozen penile tumour tissue using proteinase K digestion followed by the Wizard™ Genomic DNA Purification kit from Promega.

### Whole genome amplification (III)

Due to limited amounts of penile cancer tissue, and to secure availability of sufficient DNA amounts, all penile samples were subjected to a whole genome amplification (WGA) using the GenomiPhi™ V2 kit from Amersham Biosciences. Here 10-20 ng of the original DNA is subjected to a 16h, one-step amplification using short random primers, giving a representative amplification of the whole genome and resulting in an average DNA yield of 10 $\mu$ g. This DNA was used for mutational analysis and in samples where mutations were identified; original non-WGA DNA was used to verify the finding.

### Single nucleotide polymorphism (I, II)

The most abundant genetic variation in the human genome is in the form of single nucleotide polymorphisms (SNPs). SNPs are locations in the genome where 2 alleles can be found in the population and it is estimated that there are more than  $10 \times 10^6$  SNPs in the genome. Given the abundance of these genetic variations and their even distribution across the genome, most SNPs will be located to areas where no genes can be found, thus in locations where the variation will have little or no phenotypic impact. However when the SNPs are located to coding or regulating areas of the genome they can have

profound effects, possibly modifying the function of the protein and the susceptibility to diseases.

A classical way of detecting SNPs is by the use of restriction enzymes which cleave DNA at specific recognition sequences, usually 4-8 nucleotides in length. A fragment containing the SNP is amplified using PCR, with a following digestion with a restriction enzyme which will recognize and cleave the wildtype allele and leave sequences containing the mutated allele intact or vice versa. The resulting product is separated and visualized on an agarose gel where genotypes can easily be scored for each sample.

### **Single stranded conformation analysis (III)**

A rapid cooling of heat-denatured DNA will cause the single stranded DNA to adopt a secondary structure determined by the specific nucleotide sequence of the DNA strand. Even though two DNA strands only differ by a single base, this will lead to the formation of two different secondary structures. This characteristic can be used to identify mutations in a heterogenous mass of tumour and normal cells. In a single stranded conformation analysis (SSCA) samples are marked in a PCR reaction by the inclusion of radioactively labelled nucleotides and separated on a polyacrylamide gel. The pattern of separated bands can be visualized using an x-ray film. Samples containing only one set of DNA strands will result in two bands on the film, while a sample containing both DNA from normal cells and from tumour cells carrying a mutation, will result in four bands. The bands representing mutated DNA can be excised, eluted and used as template in a secondary PCR and then sequenced to determine the exact sequence of bases. SSCA provides an efficient screening method for mutations in multiple samples. The method also provides a sensitive way to identify a small population of mutated cells in a big population of normal cells, even down to a ratio of 1:10.

### **Sanger chain termination DNA sequencing (III)**

In a PCR reaction the existence of a free 3' hydroxyl group on the target sequence is crucial for the polymerase to continue incorporating nucleotides. Dideoxynucleotides (ddNTP) lack this hydroxyl group and can be used to terminate the chain. In a sequencing PCR reaction there are both ordinary dNTPs and ddNTPs and as the chain is extended a dNTP or a ddNTP can be incorporated at each position. An inclusion of a ddNTP will terminate the chain, but since the ordinary dNTPs are present in a greater

concentration the majority of fragments will continue to be extended. By the end of the reaction there will be fragments of all lengths, ranging from the primer sequence with one added ddNTP to the full target sequence completed with a final ddNTP. The ddNTP can be radioactively or fluorescently labelled. The generated fragments can be separated on a denaturing sequencing gel or in a capillary gel electrophoresis system. In this thesis a MEGABASE™ capillary gel electrophoresis system was used where each of the ddNTP types have a specific fluorophore. Each of the four fluorophores can be excited by laser light and emits fluorescence of a specific wavelength that can be detected when run through the system and the sequence and potential mutations can be determined.

### **Analysis of AR CAG microsatellite (I)**

A radioactive labelled target sequence containing the CAG trinucleotide repetitive sequence was generated using PCR with the inclusion of radioactive <sup>32</sup>P-labelled dATP nucleotides. The labelled PCR products were separated on a denaturing sequencing gel and the lengths of the PCR products were determined using a sequenced known sample as reference and size marker. The length of the CAG trinucleotide repeat will be directly proportional to the length of the PCR product since the flanking sequences for the target sequence will be the same for each sample.

### **Microarray (IV)**

The microarray technique provides a powerful method to obtain genetic information on a genome-wide scale from a sample. The technique enables the simultaneous interrogation of 100,000s of SNPs. Briefly; each sample was digested using an array type specific restriction enzyme after which all fragments have adaptor sequences ligated to the ends. A PCR with specific conditions was used to selectively amplify fragments with lengths ranging from 200-1100 bases, using a generic primer recognizing the adaptor sequences. The amplified DNA was then fragmented, labelled and hybridized to short oligonucleotides, each containing a SNP site, attached to the surface of an array. In this thesis the Affymetrix GeneChip Human Mapping 250k *NspI* Array has been used, which is one part of the Affymetrix GeneChip Human Mapping 500k Array Set. The GeneChip Human Mapping 250k *NspI* Array interrogates approximately 262,000 SNPs and uses the *NspI* restriction enzyme to generate specific fragments.

The amount of genetic information obtained is breathtaking and the most time consuming part of the process is the post-experiment data-analysis. Genotype calls for each SNP were generated in the GeneChip Genotyping Analysis Software (GTYPE) provided by Affymetrix using the Bayesian Robust Linear Model with Mahalanobis distance classifier (BRLMM) algorithm. The resulting data was analyzed for copy number data for each sample using two programs: dChip (LIN *et al.* 2004) a software package freely available at [www.dchip.org](http://www.dchip.org); and GeneChip Chromosome Copy Number Analysis Tool version 4.0 (CNAT 4.0), an analysis feature included in the GTYPE software.

In dChip, following normalization the signal values for each SNP on each array were obtained with a model based (perfect-match/mismatch) method. The signal intensities were compared with data from a set of 16 normal reference samples from the HapMap project, freely available on the Affymetrix website. From the raw signal data, inferred copy number state at each SNP locus was estimated by applying a Median smoothing algorithm with a sliding window of 15 SNPs. With an average distance between SNPs of 5.8 kb, a window of 15 SNPs will show areas of aberrations larger than approximately 87 kb. Copy number gain was defined as  $\geq 2.8$  and copy number loss as  $\leq 1.2$  DNA copies. We chose these thresholds instead of the more stringent of  $\geq 3$  for gains and  $\leq 1$  for losses, to compensate for the possible presence of up to 25% normal cells in the samples. Mapping information of SNP locations and cytogenetic bands were obtained from Affymetrix and University of California Santa Cruz (<http://genome.ucsc.edu>).

Copy number data was also calculated using the CNAT feature included in the GTYPE software from Affymetrix. Copy numbers for each SNP in each sample were generated by comparing with 30 normal male reference samples from the HapMap project obtained from the Affymetrix website. The raw copy number data was smoothed by applying a Hidden Markov Model (HMM) using default values in the software. Copy number information for each sample was exported from CNAT to the UCSC genome browser to assess the exact copy number state for genes residing in the identified chromosomal regions of aberration.

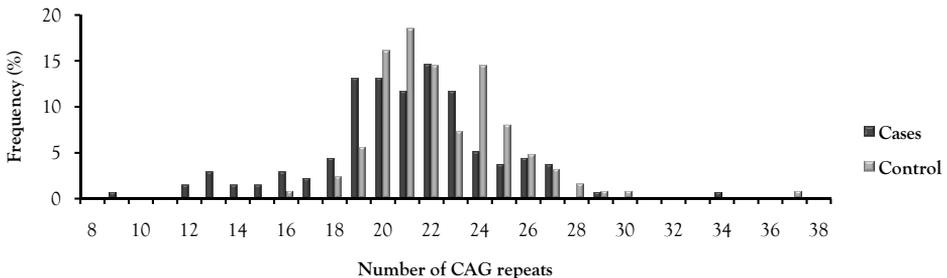
## Statistics (I, II)

The t-test is a parametric test which was used to identify differences in numbers of CAG repeats between the case and control populations in paper I. A prerequisite for using the t-test is that the variables are normally distributed and therefore the distribution was tested using Kolmogorov-Smirnov test before applying the t-test. A power calculation was also made to assess the statistical strength of the study, which takes into account the size of the populations, the magnitude of difference to be detected and the statistical threshold chosen for significance. In papers I and II  $\chi^2$ -analysis were performed to compare genotype and allele frequencies. Comparisons were expressed as odds ratio (OR) if comparing cases and controls and as relative risk (RR) if comparisons of subgroups within the case population (such as Gleason grade). A  $\chi^2$ -analysis was also used to test whether the control populations deviated from the Hardy-Weinberg equilibrium. Results were considered significant if  $p < 0.05$

## RESULTS AND DISCUSSION

### Short CAG repeat length in the AR gene increases risk of prostate cancer

The frequency distribution of CAG repeats among cases and controls in our study are presented in Figure 2. Firstly the data was analyzed by comparing the mean number of repeats in the case and control population. This may be the crudest of comparisons but has the advantage of being free from *a priori* assumptions. The mean number of CAG repeats in the case population was 20.1 (SD=3.73) and in the control population 22.5 (SD=2.85). Since both populations follow Gaussian distribution, we performed an unpaired t-test which shows a significant ( $p<0.001$ ) difference in mean number of repeats between the cases and controls. In this study we have a power of 99.9% to discover a difference of 2.5 repeats, using a threshold of  $p<0.001$ . The meta-analysis by Zeegers and colleagues showed that by statistical pooling of data from 19 studies, prostate cancer cases had on average 0.26 fewer CAG repeats than controls (ZEEGERS *et al.* 2004). We find this to give strong support to the hypothesis of fewer CAG repeats being associated with increased risk of prostate cancer.



**Figure 2.** The frequency distribution of number of CAG repeats among cases and controls.

In order to compare our results with results from previous studies we have to divide the populations into subgroups based on the number of CAG repeats. Previous studies have chosen to either dichotomize or use tertiles, generating 2 and 3 groups respectively in each population. Studies have defined short CAG repeat lengths with the lowest cut-off at  $\leq 17$  repeats (HAKIMI *et al.* 1997) and the highest at  $\leq 22$  repeats (HARDY *et al.* 1996; HSING *et al.* 2000b). Applying the  $\leq 17$  repeats cut-off on our data results in an OR=18.76 (95% CI= 2.39-787.62;  $p=0.0001$ ), which is very high due to the fact that only one sample in the control group had a repeat length this short, while 18 samples could be found in the case group. On the other hand, applying the  $\leq 22$  repeats cut-off

on our data showed a modest OR=1.72 (95% CI= 1,00-2.96;  $p=0.036$ ). In the meta-analysis by Zeeger and colleagues a cut-off at  $\leq 21$  repeats was chosen showing a summary OR=1.19 (95% CI 1.07-1.31) (ZEEGERS *et al.* 2004). Our data shows an OR=1.64 (95% CI=0.98-2.75;  $p=0.048$ ) using this cut-off, which is borderline significant. This shows that the cut-off chosen for the dichotomization of the populations will greatly determine the outcome. However, our data consistently show an association between fewer CAG repeats and increased risk of prostate cancer (Table 1).

**Table 1.** An overview of different cut-offs for dichotomization of our populations based on CAG repeat length and the risks obtained when comparing cases and controls.

Dichotomizing cut-off used	OR	95% CI	<i>p</i> -value
$\leq 17$ CAG	18.76	2.39 - 787.62	0.0001
$\leq 18$ CAG	6.42	2.03 - 26.11	0.0002
$\leq 19$ CAG	4.58	2.13 - 10.03	0.00001
$\leq 20$ CAG	2.39	1.35 - 4.15	0.001
$\leq 21$ CAG	1.64	0.98 - 2.75	0.048
$\leq 22$ CAG	1.72	1.00 - 2.96	0.036

Since the frequency of CAG repeat lengths in a population follow a Gaussian distribution an attempt to enhance visualization of the potential effects of shortest and longest repeat lengths can be to divide the population into three evenly sized groups, removing the diluting effect of the frequently occurring middle ranged repeat lengths (GIOVANNUCCI *et al.* 1999; GSUR *et al.* 2002; HSING *et al.* 2000b; INGLES *et al.* 1997). In our study we defined  $\geq 23$  repeats, 20-22 repeats and  $\leq 19$  repeats as being long, intermediate and short respectively. Using the long subgroup as reference the intermediate group showed an OR=1.14 ( $p=0.630$ ) and the short group an OR=4.94 ( $p=0.00003$ ).

As showed in the meta-analysis by Zeeger and colleagues there are many contradicting results regarding the CAG repeat length and its influence on the prostate cancer risk (ZEEGERS *et al.* 2004). As pointed out by Giovannucci in 2002, one explanation of discordant results may be that many studies that find an association have used study populations collected prior to the era of PSA screening and thus contain more tumours with a high grade (GIOVANNUCCI 2002). It seems likely that our tumours have similar characteristics to those studied by Giovannucci, Ingles and Hsing, studies that also find a positive correlation with short CAG repeats (GIOVANNUCCI *et al.* 1999; HSING *et al.* 2000b; INGLES *et al.* 1997). The CAG repeat length was suggested to be most important

for cancer diagnosed in younger patients (<60 years) (GIOVANNUCCI 2002). However, we do not observe an association with Gleason score or age at diagnosis. The lack of association concerning the age at diagnosis might be due to the fact that we have no cases that were under 60 years of age.

We acknowledge that we have a fairly small cohort and that the rather strong association we observe might be exaggerated due to this. Polymorphisms with a minor allele occurring at a high frequency will tend to have subtle phenotypic effects due to evolutionary pressure. However, we do believe that there is a true association between fewer CAG repeats and an elevated risk of prostate cancer. As pointed out by Nelson and Witte, even a modest increase in risk could be of importance from a public health perspective because of the high frequency of short CAG repeat lengths in the population (NELSON and WITTE 2002).

### **The *TaqI* polymorphism in the *VDR* gene does not modify risk of prostate cancer**

The epidemiological evidence of the involvement of the *VDR* gene in prostate carcinogenesis is unclear. The *TaqI* polymorphism is the most commonly studied SNP in the *VDR* gene. Our research group previously observed an association between the *TaqI* polymorphism in the *VDR* gene and increase risk of lymph node metastasis in a breast cancer material (LUNDIN *et al.* 1999), however we do not find an association between the *TaqI* polymorphism and risk of prostate cancer (Table 2). Comparing cases and controls regarding frequencies of the *t*- and *T*-allele result in an OR=0.98 ( $p=0.925$ ). Neither the *Tt* genotype (OR=1.03,  $p=0.935$ ) nor the *TT* genotype (OR=0.98,  $p=0.959$ ) was significantly differing from the *tt* genotype when comparing cases and controls. The *TaqI* polymorphism was not associated with stage, Gleason grade or cause of death either. We therefore believe that the *TaqI* polymorphism does not have an influence on the risk of prostate cancer.

A meta-analysis of this research field was published by Ntais and colleagues, unfortunately two years following the initiation of our *TaqI* study. In this study a summary OR of 0.95 ( $p=0.3$ ) was presented for the *t*-allele versus *T*-allele, based on 14 studies, indicative of a lack of association (NTAIS *et al.* 2003). In 2006 another meta-analysis was published by Berndt and colleagues, including an additional 4 studies on the *TaqI* polymorphism. Collectively the 18 studies generated a combined OR=1.00 (95% CI=0.85-1.18) for the [*Tt* vs *TT*] comparison and combined OR=0.94 (95% CI=0.78-1.13) for the [*tt* vs *TT*] comparison (BERNDT *et al.* 2006). These two studies

provide strong evidence that the *VDR* *TaqI* polymorphism does not affect the susceptibility for prostate cancer. Both meta-analyses investigated the other three SNPs in the *VDR* gene, *FokI*, *BsmI* and *ApaI* as well as the poly(A) polymorphism in the 3' UTR, and found none to be associated with an increased risk of prostate cancer (BERNDT *et al.* 2006; NTAIS *et al.* 2003).

There can be many reasons to the inconclusive results regarding the role of the *VDR* in prostate carcinogenesis. Firstly as pointed out by both Ntais *et al* and Berndt *et al*, individual studies may be too small to faithfully assess small differences in risk (BERNDT *et al.* 2006; NTAIS *et al.* 2003). Further, there is heterogeneity in study designs regarding the ethnical distribution in the cohorts and the choices of control population. Three of the SNPs in the 3' end of the *VDR* gene and the poly(A) polymorphism belong to the same haplotype, but the degree of linkage disequilibrium within this haplotype differ between different ethnical groups (NEJENTSEV *et al.* 2004), which may be a factor when studying the effects of single SNPs in the gene. Additionally there are reports that the SNPs in the *VDR* gene actually will not alter the level of activity of the protein (DURRIN *et al.* 1999; GROSS *et al.* 1998; VERBEEK *et al.* 1997).

A hypothesis put forward by many is that there is a susceptibility site in the vicinity of the *VDR* gene that is in partial linkage with the haplotypes in the *VDR* gene, thus explaining the sporadic nature of the associations observed. Speaking against such a hypothesis is that studies based on genome scans fail to identify significant linkage with the chromosomal region of 12q12-q14, where the *VDR* gene is located (OSTRANDER and STANFORD 2000; STANFORD and OSTRANDER 2001).

The biochemical and epidemiological evidence of a link between vitamin D and prostate cancer is strong, and alterations in the physiology surrounding the vitamin D will likely have an effect on the susceptibility for prostate cancer. However, the *VDR* gene, although an obvious candidate, might not be the right component to study and possibly the research in this field should be concentrated to other components in the pathway.

### **The +1127 *PstI* polymorphism in the insulin gene does not modify risk of prostate cancer**

In the second paper we study the influence of the +1127 *PstI* polymorphism in the *INS* gene on prostate cancer risk; with the background that the polymorphism is thought to result in altered mRNA levels of the *INS* gene and variability in mRNA stability. The

common CC genotype is additionally linked with the regulatory VNTR located adjacent 5' of the *INS* gene. We do not find any association between the +1127 PstI polymorphism and prostate cancer, neither on an allelic level (C vs T; OR=1.02,  $p=0.929$ ) nor on a genotype level (CT vs TT; OR=0.73,  $p=0.503$ ) (CC vs TT; OR=0.84,  $p=0.698$ ) (Table 2). Unfortunately we do not have data on circulating insulin levels or whether individuals have diabetes, which would facilitate the interpretation and enable further stratification. Not being able to compensate for such diluting factors will make identification of risks more challenging, especially for a frequently occurring genotype like the CC where only a modest risk can be expected.

Our results are contradictory to two previous studies conducted, the first on a multi-ethnic population and the other on an African-American population (CLAEYS *et al.* 2005; HO *et al.* 2003). The studies report that subjects with CC genotype were 3.14 times and 1.59 times, respectively, as likely to be diagnosed with prostate cancer. A noticeable difference between the present study and the previously conducted ones is that we have cohorts with significantly older subjects, especially regarding cases. Our results are in agreement with the study of Neuhausen *et al.*, where no association with the C allele was seen, which is conducted on a Caucasian population, similar to the present study but in contrast to the studies of Ho and Claeys (CLAEYS *et al.* 2005; HO *et al.* 2003; NEUHAUSEN *et al.* 2005). At a first glance one might attribute the difference observed to the fact of ethnical origin, however the strongest association was in fact among the Caucasian subpopulation in the study by Ho and co-workers. Unfortunately this study is underpowered with limited population sizes of 40 cases and 36 controls.

Similar to our study, Neuhausen and co-workers have a case population assembled prior to the era of active PSA screening, meaning that cases are more likely to be of higher grade due to later detection. In the paper of Neuhausen *et al.* 48% cases have a Gleason score of  $\geq 7$  (NEUHAUSEN *et al.* 2005), and in our study 63% cases have Gleason scores of  $\geq 7$ . Ho *et al.* on the other hand study a case population where 77% were diagnosed due to abnormal PSA-values and only 31% had Gleason scores  $\geq 7$  (HO *et al.* 2003). Since neither Neuhausen nor our study finds any association a possibility is that the polymorphism is only significant in low grade tumours. Unfortunately the data concerning Gleason score is not shown in the study by Claeys *et al.* and thus we are unable to use that data to further support such a hypothesis (CLAEYS *et al.* 2005). One may speculate that since the cases in Claeys *et al.* study are collected between 1999 and 2002, after the introduction of PSA screening, they could represent similar characteristics to those in Ho *et al.* study (HO *et al.* 2003).

When comparing genotype frequencies for each Gleason group with each other we find a significant increase in risk associated with the CC genotype in the  $\geq 7$  Gleason group. Given that we only have Gleason scores for 114 subjects together with the fact that we did not find any overall difference for this polymorphism it cannot be excluded that this may be a chance finding. These data are contradictory to previous studies where no association with Gleason score have been found (CLAEYS *et al.* 2005) or where an increased risk was associated with low Gleason scores ( $\leq 6$ ) and the CC genotype (HO *et al.* 2003).

### **The G972R polymorphism in the *IRS1* gene increases risk of prostate cancer**

In this study we find an association between the R allele of the *IRS1* G972R polymorphism and prostate cancer, conferring a 2.44 fold increased risk of disease ( $p=0.01$ ) (Table 2). We can also observe a significant  $\chi^2$  trend for an increase in risk with increasing numbers of R alleles ( $p<0.001$ ). This supports the earlier findings by Neuhausen and co-workers in 2005 where a 2.8 fold increase in prostate cancer risk was observed associated with the combined GR/RR genotypes (NEUHAUSEN *et al.* 2005). The same group has found similar associations with colon cancer and breast cancer (SLATTERY *et al.* 2004; SLATTERY *et al.* 2007).

The R allele results in decreased insulin stimulated signalling, due to decreased binding of the p85 subunit of PI3K to IRS1 and impaired PI3K/AKT signalling (LEE *et al.* 2000). In cell lines it has been shown that IRS1 degradation can be blocked by PI3K inhibitors which indicate the existence of a negative feedback mechanism, which has been shown to be mediated through S6 kinase, a downstream target of mTOR and AKT. This is somewhat surprising since one would expect that, if the R allele confers increased risk of cancer, it would increase signalling through the pathway rather than decreasing it. It has been suggested that in the case of the R allele the signalling through the PI3K pathway would be weaker, giving rise to less feedback inhibition and hence a longer and more sustained signalling (NEUHAUSEN *et al.* 2005). Speaking against this is that the sites that are phosphorylated by the S6 kinase are located 5' in the protein while codon 972 is in the 3' end. An alternative explanation could be that in response to a decreased efficiency of the insulin signalling via IRS1 through the PI3K/AKT axis, the signal diverts into the alternative pathway from the RTK comprised of activation of RAS/MAPK, resulting in a strong mitotic signal (SHAW and CANTLEY 2006). A third alternative is that the 3' located G972R polymorphism merely acts as a marker for a chromosomally linked disease causing gene.

**Table 2.** A summary of the results obtained in polymorphism studies of the VDR, INS and IRS1 genes.

Gene Allele/genotype comparison	OR	95% CI	p-value
<i>VDR TaqI</i>			
T vs t	0.98	0.70 – 1.38	0.925
Tt vs tt	1.03	0.52 – 2.02	0.935
TT vs tt	0.98	0.49 – 1.97	0.959
<i>INS +1127 PstI</i>			
C vs T	1.02	0.69 – 1.51	0.929
TC vs TT	0.73	0.26 – 2.03	0.503
CC vs TT	0.84	0.32 – 2.19	0.698
<i>IRS G972R</i>			
R vs G	2.44	1.16 – 5.18	0.01
RG vs GG	1.81	0.74 – 4.42	0.153
RR vs GG	6.68	0.65 – 330.82	0.072

### Mutational status of the *PIK3CA* gene in prostate cancer

With the publication of Samules *et al* on frequent mutations in *PIK3CA* observed in a variety of cancer forms, an attempt was made to assess the mutational status of the *PIK3CA* in the prostate cancer cohort used in papers I and II (SAMUELS *et al.* 2004). The PI3K/AKT pathway has previously been observed to be overactivated in prostate cancer through frequent loss of *PTEN* (DI CRISTOFANO and PANDOLFI 2000; LI *et al.* 1997) or through amplification of *PIK3CA* (BYUN *et al.* 2003), indicating that this pathway is of importance for the carcinogenic process in this tissue. This supports a hypothesis for an additional mutational activating mechanism for the *PIK3CA* gene. The hotspot regions, exon 9 and 20, were subjected to mutational analysis in the available 120 prostate cancer samples. We present preliminary data that 20 non-synonymous mutations were found in this material, of which 5 are known mutations (E542K, E545K, Q546R), previously observed in other cancer forms. We also observe a previously described silent mutation in exon 20 which does not alter the amino acid (C→T on nucleotide level, resulting in T1025T on amino acid level) (LEVINE *et al.* 2005; MULLER *et al.* 2007; WANG *et al.* 2005). We would like to emphasize that these are preliminary data which need to be verified in a new material.

Müller *et al* performed mutational analysis of the *PIK3CA* gene in prostate cancer and found no mutations (MULLER *et al.* 2007). However, only 12 samples were analyzed in the study, which is too few to exclude the possibility of mutations in the gene. It is puzzling that only one study has been produced on the mutational status of the *PIK3CA* gene in prostate cancer since the pioneering study by Samuels *et al* in 2004 (SAMUELS *et al.* 2004). One would expect that given the high incidence and the strong evidence of an involvement of the PI3K/AKT pathway in this cancer form that mutational analysis of the *PIK3CA* gene would receive high priority among prostate cancer researchers.

### **The *PIK3CA* gene is frequently mutated in penile cancer**

In paper III we perform mutation analysis of the *PIK3CA* and *PTEN* genes in the PI3K/AKT pathway and *HRAS*, *KRAS*, *NRAS* and *BRAF* genes in the RAS/MAPK pathway. A significant number of mutations were identified in the *PIK3CA* gene in 8/28 cases (29%), all of which cluster in the hot-spot region of exon 9 (Table 3). We chose to focus our analysis primarily on the mutational hot-spot regions in *PIK3CA*, exons 9 and 20, where 85% of mutations were initially found by Samuels *et al* which may give rise to an underestimate the mutational frequency of *PIK3CA* in penile cancer (SAMUELS *et al.* 2004). Unlike most previous studies we only find somatic mutations in exon 9 and none in exon 20, although the limited number of samples may account for this.

The mutations identified in this study, E542K and E545K are known to increase the activity of the protein. We show in our study through immunohistochemical staining that samples carrying *PIK3CA* mutations have an increased phosphorylation of the AKT protein and hence an active signalling pathway. Recent results show that substitution of the glutamic acid residues at positions 542 and 545 to lysine will render the p110 $\alpha$  subunit insensitive to the inhibition normally exerted by the regulatory p85 subunit in the resting state of the protein complex (HUANG *et al.* 2007). This means that mutated p110 $\alpha$  subunits would be active even in the absence of growth stimulatory signals from the RTKs. Previous studies have shown that overexpression of *PIK3CA* mutants (E542K, E545K and H1047R) in chicken embryo fibroblasts and NIH3T3 cells induce the phosphorylation of AKT and oncogenic transformation in comparison to wildtype *PIK3CA* (IKENOUE *et al.* 2005; KANG *et al.* 2005). Samuels and colleagues showed that in two colorectal cell lines harbouring each of the two hot-spot mutations (E545K and H1047R), and expressing either the wild-type allele or the mutated allele, the phosphorylation of AKT was dramatically increased in the cells expressing the mutant

allele compared to cells expressing the wild-type allele (SAMUELS *et al.* 2005). Zhang *et al* studied the effects of 9 common PIK3CA mutants in a breast cancer cell model and showed that the prevalent mutations, E542K, E545K and H1047R rendered the cells the strongest ability to grow, form colonies in and to invade 3D agaros gel matrix (ZHANG *et al.* 2007). The study also showed that cells expressing mutated PIK3CA continued to grow in the absence of epidermal growth factor (EGF). In addition to these *in vitro* data, Bader and colleagues show that the three most prevalent mutants (E542K, E545K and H1047R) are oncogenic also *in vivo* (BADER *et al.* 2006). The study showed that PIK3CA mutants induced angiogenesis and malignant growth in the chorioallantoic membrane of chicken embryo and caused hemangiosarcomas in young chickens. Collectively, these data provide a line of strong evidence for the oncogenic potential of mutated PIK3CA.

### **RAS mutations in penile cancer**

In addition to PIK3CA mutations, mutations in the HRAS gene were identified in two tumours, one tumour carried a G12S mutation and one a Q61L mutation, and yet another tumour carried a mutation in the KRAS gene being a G12S substitution (Table 3). Generally, somatic missense RAS mutations found in cancer cells introduces amino acid substitutions in codons, 12, 13 or 61. These mutations confer resistance to GAPs thereby causing the RAS protein to accumulate in its active GTP bound form (TRAHEY and MCCORMICK 1987). Dysregulated Ras proteins induce a constitutive activation of downstream kinase cascades, resulting in continuous mitogenic signalling, immortalizing the cell (KHOSRAVI-FAR *et al.* 1996). HRAS mutations have previously been found in bladder tumours (CZERNIAK *et al.* 1992; KHOSRAVI-FAR *et al.* 1996; PRZYBOJEWSKA *et al.* 2000), where as KRAS mutations have predominantly been identified in pancreatic and colonic tumours (ALMOGUERA *et al.* 1988; FORRESTER *et al.* 1987).

No mutations were observed in the PTEN, NRAS or BRAF genes in this material. However, due to the limited size of this cohort we cannot with certainty exclude the possibility of mutations in these genes.

## Frequent copy number changes in penile cancer

The fourth paper was based on 19 tumours, which were chosen from the larger cohort of 28 tumours on the basis of their tumour cell content (Table 3). The 75% tumour cell content was chosen as a minimum requirement for the samples, to ensure faithful detection of gene copy number changes and loss of heterozygosity.

In this study the major aberrations present in more than 25% of the tumours, are found as copy number gains in chromosome arms 1p, 3q, 5p, 8q, 16q, 20q and losses at 3p, 4q, 11q, 13q. In relation to the results of Alves *et al*, we find similar observations of copy number gains on 8q and 20q, and losses on 4q and 13q (ALVES *et al*. 2001). The observed patterns of alterations are further supported by findings in squamous cell carcinoma in other locations, such as the lung, cervix, oral and oesophagus (CHUJO *et al*. 2002; COE *et al*. 2005; MICCI *et al*. 2003; TONON *et al*. 2005; WILTING *et al*. 2006; YAKUT *et al*. 2006; YEN *et al*. 2001; ZHOU *et al*. 2004).

Seven of the tumours showed very few copy number aberrations indicating a diploid character and chromosomal stability. Interestingly, five of these tumours carry mutations in the *PIK3CA* gene while no such mutations were observed in the remaining 14 samples. This suggests that *PIK3CA* mutations might obviate the need for gross genomic changes, and that the oncogenic properties given by these mutations might be sufficient to drive the malignancy without further disruptive genetic events seen in 12/14 of the remaining samples. Two tumours with few copy number changes, and lacking activating *PIK3CA* mutations remain unexplained. Since the initial *PIK3CA* mutation analysis focused on two hotspot regions in exons 9 and 20, it is possible that these two samples may carry activating mutations in other parts of the gene.

**Table 3.** The table displays pathological data and results for the 28 penile cancer samples. Grey shaded samples have been chosen for SNP array analysis. Immunohistochemical data on phosphorylated AKT (p-AKT) is presented such as 1= weak nuclear and cytoplasmic staining, 2 = intermediate cytoplasmic staining, 3 = strong cytoplasmic staining. Black markings indicate samples with copy number change at the gene locus.

#	Age	Surgical method	Grade	pT-stage	T-stage	N-stage	M-stage	HPV	p-AKT	PIK3CA mut	HRAS mut	KRAS mut	FHIT loss	PIK3CA gain	hTERT gain	MYC gain
1	62	Total amp.	2	3	3	1	0	-	1	E545K	-	-	-	-	•	•
2	76	Partial amp.	2	2	3	0	0	16	3	-	-	-	•	•	•	•
3	83	Partial amp.	1	1	2	x	0	16	2	E542K	-	-	-	-	-	-
4	66	Total amp.	3	2	3	2	0	-	2	G12S	-	-	•	•	•	•
5	79	Total amp.	3	1	3	0	0	16	2	-	-	-	-	-	•	•
6	64	Partial amp.	2	2	3	3	0	-	2	-	-	-	•	•	•	•
7	62	Total amp.	3	2	2	2	1	35	3	-	-	-	-	-	-	-
8	57	Total amp.	3	2	3	0	0	-	3	E545K	-	-	-	-	-	-
9	78	Partial amp.	3	3	3	0	0	6B,11	2	Q61L	-	-	-	-	-	-
10	44	CO <sub>2</sub> + NdYAG	3	2	2	3	0	6B,11	1	E542K	-	-	-	-	-	-
11	66	CO <sub>2</sub> + NdYAG	2	2	2	0	0	6B,11	1	E542K	-	-	-	-	-	-
12	91	Partial amp.	2	2	pT2	x	0	•	3	-	-	-	•	•	•	•
13	70	CO <sub>2</sub> + NdYAG	2	2	2	0	0	•	2	-	-	-	-	-	-	-
14	83	Partial amp.	2	1	2	0	0	16	3	E545K	-	-	-	-	-	•
15	61	Partial amp.	2	3	3	3	0	16	2	-	-	G12S	-	-	-	•
16	54	CO <sub>2</sub> + NdYAG	1	1	1	0	0	35	2	-	-	-	-	-	-	-
17	58	CO <sub>2</sub> + NdYAG	2	1	1	x	0	16,31,33	2	-	-	-	-	-	-	-
18	76	CO <sub>2</sub> + NdYAG	2	2	2	2	0	•	2	-	-	-	-	-	-	-
19	38	CO <sub>2</sub> + NdYAG	2	2	2	2	0	16,31,33	2	-	-	-	•	•	•	•
20	48	CO <sub>2</sub> + NdYAG	N/A	1	1	x	0	16	3	-	-	-	-	-	-	-
21	65	CO <sub>2</sub> + NdYAG	2	1	2	0	0	16	2	-	-	-	•	•	•	•
22	64	CO <sub>2</sub> + NdYAG	CIS	CIS	CIS	x	0	18	2	-	-	-	•	•	•	•
23	70	NdYAG	CIS	CIS	CIS	x	0	-	2	-	-	-	•	•	•	•
24	53	Total amp.	1	2	3	3	0	-	1	E542K	-	-	-	-	-	-
25	62	Partial amp.	2	2	3	2	0	-	1	-	-	-	-	-	-	-
26	69	CO <sub>2</sub> + NdYAG	2	1	1	2	0	-	1	-	-	-	-	-	-	-
27	64	CO <sub>2</sub> + NdYAG	2	2	2	2	0	-	3	E545K	-	-	-	-	-	-
28	51	Partial amp.	2	2	1	0	0	-	1	-	-	-	-	-	-	-

pT stage = pathological stage

CIS = carcinoma in situ

• = infection with unknown HPV type

NdYAG = neodymium-doped yttrium aluminium garnet laser

CO<sub>2</sub> = carbon dioxide laser

## Candidate genes for penile cancer

The chromosomal regions identified are fairly large and given the abundance of genes located within these areas we can only tentatively make suggestions based on previous knowledge of genes known to be involved in oncogenesis. We have identified four plausible candidate genes located within four of the ten chromosomal regions frequently displaying gains or losses, namely *PIK3CA* at 3q26.32, *FHIT* at 3p14.2, *hTERT* at 5p15.33 and *MYC* at 8q24.21.

Amplification of the distal part of 3q in half of our samples in addition to the mutation analysis results in PAPER III, suggest that *PIK3CA* at 3q26.32 represents a natural candidate gene in the 3q copy number gain region. Eight samples of the 10 identified to have a 3q gain, had increased copy number at the locus for the *PIK3CA* gene, and none of these samples carried mutations in this gene (Table 3). This is not surprising since there may not be any selective pressure for amplification in addition to an existing mutation in *PIK3CA*, which already is a strong driving force for the malignancy.

In nine samples with frequent copy number loss on chromosome 3p, we found eight with only one intact copy of the tumour suppressor gene *Fragile histidine triad (FHIT)* located on 3p14.2 (Table 3). The mechanism by which *FHIT* exerts its tumour suppressive function remains unclear, but has been suggested to be through apoptosis control and cell cycle regulation (HUEBNER and CROCE 2003). *FHIT* is commonly disrupted through loss of heterozygosity or promoter-hypermethylation in many cancer forms such as oral, lung, head and neck, breast, gastric, colon and cervical cancer. The study further shows that *FHIT* seems to be susceptible to environmental damage in the early stages of carcinogenesis, which an observation that may be interesting in the context of penile cancer. A very interesting observation in our material is that *FHIT* copy number loss and *PIK3CA* copy number gain were found exclusively in the same tumours. Maisson *et al* observed in a study on non-small cell lung cancer a relationship where deletion on 3p (including the *FHIT* gene locus) was associated with amplification of 3q (including the *PIK3CA* gene locus), in analogy with our findings (MASSION *et al.* 2002). A similar observation was recently made by Bianchi *et al* who suggested a synergistic function of disruptions in *FHIT* and *HER2* in breast carcinogenesis (BIANCHI *et al.* 2007).

Of the 11 samples observed to have 5p gain, seven samples contained a copy number gain of the *hTERT* gene (5p15.33), a gene that codes for the catalytic part of the telomerase protein and is commonly activated in the carcinogenesis (Table 3). Telomere

maintenance is a hallmark in virtually all malignant cells by upregulating hTERT expression and keeping the telomeres above a critical threshold to allow unlimited replication.

The fourth commonly activated candidate gene that we suggest in this study is the *MYC* gene, which resides in the common area of copy number gain on 8q, specifically in band 8q24.21. An increased copy number at the *MYC* gene locus, was observed in seven of the 11 samples that had 8q gain (Table 3). Activation of this gene through amplification has been seen in breast cancer and oesophageal cancer (BODVARSDOTTIR *et al.* 2007; IOANNIDIS *et al.* 2003; SARBIA *et al.* 2001). In 1999 Wu *et al* showed that *hTERT* was a direct transcriptional target of *MYC*, implying that not only will *MYC* decrease apoptosis and propel the cell cycle but that it also can prevent cells from entering senescence (WU *et al.* 1999).

These genes represent likely candidates for further studies although other genes located in these chromosomal regions may still have an important role in the carcinogenesis. The level of *MYC* and *hTERT* proteins in the penile cancer tissue is currently being assessed using immunohistochemical staining and will provide a first confirmation of the involvement of these proteins. In the continuation of this project more samples will be added which hopefully can further narrow down the common areas of aberrations. The increasing number of samples will also enable us to stratify the material and study if there are any correlations with clinical parameters. The cohort of penile cancers used for this project was collected between 2000 and 2006 which means that there is a limited follow-up time for these patients. It is expected that the results obtained in this study will become even more significant as the outcome of all patients eventually can be observed.

### **Relation of mutations in the *PIK3CA* and *RAS* genes and amplification of the *PIK3CA* gene**

Mutations in the *PIK3CA*, *HRAS* and *KRAS* genes display mutual exclusivity throughout this sample collection, suggesting that mutational inactivation in either of the two pathways is sufficient for cancerous propagation. Copy number gain of the *PIK3CA* gene and mutation of the gene were also found to be mutually exclusive. In one sample both amplification of the *PIK3CA* gene and a mutation in the *HRAS* gene were seen. In this study we observe an even distribution of activation of the *PIK3CA* gene through mutation and amplification. Mutations and amplifications of *PIK3CA* have previously

rarely been observed to occur in the same cancer form, for example in ovarian and gastric cancers amplifications of *PIK3CA* are very common but mutations are rarely seen and vice versa in colon and breast cancer (BYUN *et al.* 2003; CAMPBELL *et al.* 2004; LEVINE *et al.* 2005; LI *et al.* 2005; SHAYESTEH *et al.* 1999; WANG *et al.* 2005). However, one recent study on lung cancer showed that *PIK3CA* mutations and 3q amplification can be found in the same tumour form, but not in the same tumour, in analogy to our results (OKUDELA *et al.* 2007). Should both mutational activation and activation through amplification result in the same level of increased signalling through this pathway it seems unlikely that there would be a selective pressure for just one type of aberration to appear.

We acknowledge that we have a limited population size and that these results need to be confirmed in a larger study population, but believe that the pathways involving *PIK3CA* and *RAS* genes are of significant importance in penile cancer, seeing that we find alterations in these proteins in 18/28 (64%) of the cases. Since whole genome microarray analyses could only be performed on 19 of these samples and the mutation analysis was limited to known hot-spot regions we believe this number even represents an underestimate of the extent of involvement of these pathways in penile squamous cell carcinoma.

## CONCLUSIONS

I finally would like to give a brief summary of our main findings in this thesis. Based on the results in papers I-IV, we draw the following conclusions:

- A low number of CAG repeats in the first exon of the *AR* gene significantly increases the risk of developing prostate cancer. The elevated risk is due to an increased transactivation ability of the androgen receptor protein.
- The R allele in the G972R polymorphism in the *IRS1* gene is associated with a 2.44 fold increased risk of prostate cancer. However, the mechanism by which this risk increase is mediated still remains unknown.
- The *TaqI* polymorphism in the *VDR* gene and the +1127 *PstI* polymorphism in the *INS* gene, do not modify the susceptibility to develop prostate cancer.
- Gain on chromosomal regions 1p, 3q, 5p, 8q, 16q, 20q and loss on chromosomal region 3p, 4q, 11q, 13q represents the major areas of aberration in penile cancer. The *PIK3CA*, *FHIT*, *hTERT* and *MYC* genes are suggested as plausible candidate genes, warranting further studies.
- The *PIK3CA* gene is frequently over-activated in penile squamous cell carcinoma. We find the *PIK3CA* gene to be activated through mutations in 29% of the cases and through copy number gain in yet another 29% of the cases. We also observe mutations in the *HRAS* and *KRAS* genes in 11% of the samples. In total we find disruption of the PI3K/AKT and RAS/MAPK pathways in 64% of our samples indicating a clear significance of these pathways in the aetiology of prostate cancer.



## ACKNOWLEDGEMENTS

Jag vill rikta ett stort tack till alla som på olika sätt har gett mig stöd, hjälp, uppmuntring och underhållning som möjliggjort genomförandet av denna avhandling.

Särskilt vill jag tacka:

**Peter Söderkvist**, för ditt engagemang, fantastiska sätt att entusiasmera och stimulera nyfikenheten för det okända. Tack för det förtroende som du visat under åren och friheten att prova olika idéer, som lett till att jag blivit en självständig forskare. Jag vill också tacka dig som en vän, för alla de trevliga och ibland sena pratstunderna om allt som inte är jobbrelaterat. Närhelst du vill göra din sabbatical är du välkommen till Australien.

Mina vänner och arbetskolligor i forskargruppen, för att ni skapat ett varmt, trevligt och hjälpsamt klimat som gjort det till ett dagligt nöje att gå till jobbet. **Åsa Schippert** och **Annette Molbaek**, utan er skulle verksamheten stanna, såväl forskningsmässigt som socialt. Tack för att ni alltid tagit er tid att hjälpa och ge råd. **Lena Thunell**, min stora syster på labbet, för hjälpen i forskningen, skratten och för lånet av BHN. **Anneli Karlsson**, för diskussioner om både roliga och seriösa ämnen, och för gott sällskap nu i slutet av våra projekt. **Deepti Verma**, du är ett glatt och trevligt tillskott i gruppen. **Nils Elander**, min långa och kloka kontorskamrat, för alla insiktsfulla svar på forskningsrelaterade frågor och för din underbara humor som glatt kontoret. **Jonas Ungerbäck**, min medbrottsling i hårdrocksdjungeln, för glada och engagerade diskussioner, ibland om jobb, oftast om allt annat. Tack för att ha delat udda jobbtider, god matlagning och en strävan efter att hitta det ultimata "party". **Lotta Harnevik**, min gamla vän, vi har strävat tillsammans mot detta mål alltsedan ht -97. Stundtals har det gått åt mycket godis, ffa vid C-uppsatsen, men nu är målet äntligen nått. Tack för allt stöd och hjälp, och att du "visade vägen" med din disputation. Jag kommer sakna våra diskussioner, så skriv nu upp dig på Facebook, så vi kan hålla kontakten.

Alla i vår korridor på Cellbiologen plan 9 för trevliga stunder i fikarummet och på fester. **Birgitta Persson**, **Gunnel Carlsson**, **Lena Tudegård**, **Catarina Trofast**, **Sigrun Liedgren**, **Cecilia Gunnarsson** och **Jon Jonasson**, för fantastiskt gott fika. Catarina och Jon, ni ska ha speciellt tack för hjälpen med microarrayförsöken. Alla forna medarbetare i korridoren, **Karin Fransén**, **Ahmad Ahmadi**, **Helen Ekman**, **Linda Vainikka**, **Sten Wingren**, **Malin Bergman**, **Pia Wegman**, **Pernilla Eliasson**, **Amanda Nordigården**, **Pia Druid** och **Jan-Ingvar Jönsson**.

**Eberhard Varenhorst**, för att ha givit mig tillfället och förtroendet att jobba med det prostata cancer material du samlat in. Mina kollegor i **Örebro**, för ett utvecklande och intressant samarbete, som jag hoppas fortsätter resultera i spännande upptäckter.

Alla kära vänner och doktorandkollegor som gjort sjukhuset till mer än bara en arbetsplats och förgyllt tillvaron med daglig glädje. **Johan Junker**, för en stor vänskap som jag hoppas blir livslång. Du har en glädje som smittar av sig till alla och jag vet ingen som kan "pimp my story" som du! Jag är verkligen glad att du accepterade "uppdraget". **Anders Carlsson**, för att du är en trogen vän som alltid ställer upp, och för alla blöta fester i er sommarstuga. **Alex Persson, Ana Vujic, Anders Nilsson, Anna Rydén, Calle Jonsson, Faisal Omar, Femke Lutgendorff, Henrik Blomkvist, Henrik Åberg, Jerker Linné, Jocke Henricsson, Johan Olausson, Jonna Fridfeldt, Karolin Eriksson, Lisa Karlsson, Martin Tinnerfeldt Winberg, Micke Pihl, Pernilla Eliasson, Petter Isaksson, Sa'ad Salim, Sebastian Schultz och Ulrika Nilsson**, det är ett sant nöje och en trygghet att ha vänner som er!

Mina gamla kursare, **Cecilia Bivik, Helena M Linge, Jocke Strandberg, Lotta Harnevik, Mimi Westerlund, Niclas Grahn, Sara Mattsson, Stina Lundin och Toffe Ström** för att ni gjorde studenttiden till det glada minne det är. Mina gamla sambos på Rekrytgatan (vilket är nästan löjligt många av er ovan) för en mycket bekymmerslös och minnesvär tid. Jocke, Toffe och Niclas, jag lovar att återvända till landet i framtiden så att vi tillsammans kan genomföra alla innestående 30 års presenter.

Alla eldsjälar som startade och har arbetat hårt med utbildningen Medicinsk Biologi. **Peter Söderkvist, Olle Stål, Anders Ljungman, Magnus Grenegård, Peter Pahlsson, Susanne Andersson** och många andra lärare, ni har bevisat att en biomedicinsk utbildning med problembaserat lärande är en framgångsrik kombination för en framtida karriär inom forskning.

**Håkan Wiktander**, för att du ställt upp och backat alla våra idéer och upptåg, inte bara på jobbet, utan även också på fritiden.

I would also like to express my sincere gratitude and appreciation to:

The **Brotherton family, Maria, Michael, Duncan and Tom**, for the warm welcoming I have received from you. I look forward to become part of your family and to spend more time with you all in the years to come.

Mina syskon och deras familjer, **Mia** och **Rickard Hogner, Jonas** och **Anna Dartman, Lotta** och **Berra Alm**, för ert stöd och er kärlek. **Inger Andersson**, för att du varit en vänlig och omtänksam extramamma åt mig. Ni har alla fått Linköping att verkligen kännas som hemma.

**Mamma** och **pappa**, för ni alltid tror på mig, uppmuntrar och stödjer mig på alla sätt vad jag än väljer att göra. Ni har givit mig trygghet, självförtroende och en vilja att undersöka vad som finns där ute i världen. Ni är fantastiska föräldrar och det är er förtjänst att jag är den jag är, och att jag är där jag är!

**Lauren**, my soul mate and the love of my life, thank you for your endless love and support! We have finally reached the point towards which we have been working so hard and we can now properly start our life together. I look forward to an exciting life with you and to see where the adventure will take us next...

## REFERENCES

- ALBERTSEN, P. C., D. G. FRYBACK, B. E. STORER, T. F. KOLON and J. FINE, 1995 Long-term survival among men with conservatively treated localized prostate cancer. *Jama* **274**: 626-631.
- ALMOGUERA, C., D. SHIBATA, K. FORRESTER, J. MARTIN, N. ARNHEIM *et al.*, 1988 Most human carcinomas of the exocrine pancreas contain mutant c-K-ras genes. *Cell* **53**: 549-554.
- ALVES, G., A. HELLER, W. FIEDLER, M. M. CAMPOS, U. CLAUSSEN *et al.*, 2001 Genetic imbalances in 26 cases of penile squamous cell carcinoma. *Genes Chromosomes Cancer* **31**: 48-53.
- ARAI, H., K. MIYAMOTO, Y. TAKETANI, H. YAMAMOTO, Y. IEMORI *et al.*, 1997 A vitamin D receptor gene polymorphism in the translation initiation codon: effect on protein activity and relation to bone mineral density in Japanese women. *J Bone Miner Res* **12**: 915-921.
- AUS, G., D. ROBINSON, J. ROSELL, G. SANDBLOM and E. VARENHORST, 2005 Survival in prostate carcinoma--outcomes from a prospective, population-based cohort of 8887 men with up to 15 years of follow-up: results from three countries in the population-based National Prostate Cancer Registry of Sweden. *Cancer* **103**: 943-951.
- BADER, A. G., S. KANG and P. K. VOGT, 2006 Cancer-specific mutations in PIK3CA are oncogenic in vivo. *Proc Natl Acad Sci U S A* **103**: 1475-1479.
- BENNETT, S. T., and J. A. TODD, 1996 Human type 1 diabetes and the insulin gene: principles of mapping polygenes. *Annu Rev Genet* **30**: 343-370.
- BERNDT, S. I., J. L. DODSON, W. Y. HUANG and K. K. NICODEMUS, 2006 A systematic review of vitamin D receptor gene polymorphisms and prostate cancer risk. *J Urol* **175**: 1613-1623.
- BIANCHI, F., E. TAGLIABUE, S. MENARD and M. CAMPIGLIO, 2007 Fhit expression protects against HER2-driven breast tumor development: unraveling the molecular interconnections. *Cell Cycle* **6**: 643-646.
- BLAZER, D. G., 3RD, D. M. UMBACH, R. M. BOSTICK and J. A. TAYLOR, 2000 Vitamin D receptor polymorphisms and prostate cancer. *Mol Carcinog* **27**: 18-23.
- BODVARSDOTTIR, S. K., M. STEINARSDOTTIR, H. HILMARSDDOTTIR, J. G. JONASSON and J. E. EYFIORD, 2007 MYC amplification and TERT expression in breast tumor progression. *Cancer Genet Cytogenet* **176**: 93-99.
- BOGUSKI, M. S., and F. MCCORMICK, 1993 Proteins regulating Ras and its relatives. *Nature* **366**: 643-654.
- BOS, J. L., 1989 ras oncogenes in human cancer: a review. *Cancer Res* **49**: 4682-4689.
- BRAWER, M. K., 1999 Prostate-specific antigen: current status. *CA Cancer J Clin* **49**: 264-281.
- BRUNET, A., A. BONNI, M. J. ZIGMOND, M. Z. LIN, P. JUO *et al.*, 1999 Akt promotes cell survival by phosphorylating and inhibiting a Forkhead transcription factor. *Cell* **96**: 857-868.
- BYUN, D. S., K. CHO, B. K. RYU, M. G. LEE, J. I. PARK *et al.*, 2003 Frequent monoallelic deletion of PTEN and its reciprocal association with PIK3CA amplification in gastric carcinoma. *Int J Cancer* **104**: 318-327.
- CAIRNS, P., K. OKAMI, S. HALACHMI, N. HALACHMI, M. ESTELLER *et al.*, 1997 Frequent inactivation of PTEN/MMAC1 in primary prostate cancer. *Cancer Res* **57**: 4997-5000.
- CAMPBELL, I. G., S. E. RUSSELL, D. Y. CHOONG, K. G. MONTGOMERY, M. L. CIAVARELLA *et al.*, 2004 Mutation of the PIK3CA gene in ovarian and breast cancer. *Cancer Res* **64**: 7678-7681.

- CARDONE, M. H., N. ROY, H. R. STENNICKE, G. S. SALVESEN, T. F. FRANKE *et al.*, 1998 Regulation of cell death protease caspase-9 by phosphorylation. *Science* **282**: 1318-1321.
- CARTER, B. S., T. H. BEATY, G. D. STEINBERG, B. CHILDS and P. C. WALSH, 1992 Mendelian inheritance of familial prostate cancer. *Proc Natl Acad Sci U S A* **89**: 3367-3371.
- CELEBI, J. T., I. SHENDRIK, D. N. SILVERS and M. PEACOCKE, 2000 Identification of PTEN mutations in metastatic melanoma specimens. *J Med Genet* **37**: 653-657.
- CHAMBERLAIN, N. L., E. D. DRIVER and R. L. MIESFELD, 1994 The length and location of CAG trinucleotide repeats in the androgen receptor N-terminal domain affect transactivation function. *Nucleic Acids Res* **22**: 3181-3186.
- CHONG, H., H. G. VIKIS and K. L. GUAN, 2003 Mechanisms of regulating the Raf kinase family. *Cell Signal* **15**: 463-469.
- CHUJO, M., T. NOGUCHI, T. MIURA, M. ARINAGA, Y. UCHIDA *et al.*, 2002 Comparative genomic hybridization analysis detected frequent overrepresentation of chromosome 3q in squamous cell carcinoma of the lung. *Lung Cancer* **38**: 23-29.
- CLAEYS, G. B., A. V. SARMA, R. L. DUNN, K. A. ZUHLKE, J. BEEBE-DIMMER *et al.*, 2005 INSPstI polymorphism and prostate cancer in African-American men. *Prostate* **65**: 83-87.
- COE, B. P., L. J. HENDERSON, C. GARNIS, M. S. TSAO, A. F. GAZDAR *et al.*, 2005 High-resolution chromosome arm 5p array CGH analysis of small cell lung carcinoma cell lines. *Genes Chromosomes Cancer* **42**: 308-313.
- COOPERBERG, M. R., D. P. LUBECK, S. S. MEHTA and P. R. CARROLL, 2003 Time trends in clinical risk stratification for prostate cancer: implications for outcomes (data from CaPSURE). *J Urol* **170**: S21-25; discussion S26-27.
- CORDER, E. H., H. A. GUESS, B. S. HULKA, G. D. FRIEDMAN, M. SADLER *et al.*, 1993 Vitamin D and prostate cancer: a prediagnostic study with stored sera. *Cancer Epidemiol Biomarkers Prev* **2**: 467-472.
- CORREA-CERRO, L., P. BERTHON, J. HAUSSLER, S. BOCHUM, E. DRELON *et al.*, 1999 Vitamin D receptor polymorphisms as markers in prostate cancer. *Hum Genet* **105**: 281-287.
- CORVERA, S., and M. P. CZECH, 1998 Direct targets of phosphoinositide 3-kinase products in membrane traffic and signal transduction. *Trends Cell Biol* **8**: 442-446.
- COX, N. J., G. I. BELL and K. S. XIANG, 1988 Linkage disequilibrium in the human insulin/insulin-like growth factor II region of human chromosome II. *Am J Hum Genet* **43**: 495-501.
- CZERNIAK, B., G. L. COHEN, P. ETKIND, D. DEITCH, H. SIMMONS *et al.*, 1992 Concurrent mutations of coding and regulatory sequences of the Ha-ras gene in urinary bladder carcinomas. *Hum Pathol* **23**: 1199-1204.
- DAHIA, P. L., D. J. MARSH, Z. ZHENG, J. ZEDENIUS, P. KOMMINOTH *et al.*, 1997 Somatic deletions and mutations in the Cowden disease gene, PTEN, in sporadic thyroid tumors. *Cancer Res* **57**: 4710-4713.
- DATTA, S. R., H. DUDEK, X. TAO, S. MASTERS, H. FU *et al.*, 1997 Akt phosphorylation of BAD couples survival signals to the cell-intrinsic death machinery. *Cell* **91**: 231-241.
- DAVIES, H., G. R. BIGNELL, C. COX, P. STEPHENS, S. EDKINS *et al.*, 2002 Mutations of the BRAF gene in human cancer. *Nature* **417**: 949-954.
- DENNIS, P. B., A. JAESCHKE, M. SAITOH, B. FOWLER, S. C. KOZMA *et al.*, 2001 Mammalian TOR: a homeostatic ATP sensor. *Science* **294**: 1102-1105.
- DEWIRE, D., and H. LEPOR, 1992 Anatomic considerations of the penis and its lymphatic drainage. *Urol Clin North Am* **19**: 211-219.
- DI CRISTOFANO, A., and P. P. PANDOLFI, 2000 The multiple roles of PTEN in tumor suppression. *Cell* **100**: 387-390.

- DIEHL, J. A., M. CHENG, M. F. ROUSSEL and C. J. SHERR, 1998 Glycogen synthase kinase-3beta regulates cyclin D1 proteolysis and subcellular localization. *Genes Dev* **12**: 3499-3511.
- DILLNER, J., G. VON KROGH, S. HORENBLAS and C. J. MEIJER, 2000 Etiology of squamous cell carcinoma of the penis. *Scand J Urol Nephrol Suppl*: 189-193.
- DONOVAN, S., K. M. SHANNON and G. BOLLAG, 2002 GTPase activating proteins: critical regulators of intracellular signaling. *Biochim Biophys Acta* **1602**: 23-45.
- DURRIN, L. K., R. W. HAILE, S. A. INGLES and G. A. COETZEE, 1999 Vitamin D receptor 3'-untranslated region polymorphisms: lack of effect on mRNA stability. *Biochim Biophys Acta* **1453**: 311-320.
- FEDERICI, M., A. PETRONE, O. PORZIO, C. BIZZARRI, D. LAURO *et al.*, 2003 The Gly972-->Arg IRS-1 variant is associated with type 1 diabetes in continental Italy. *Diabetes* **52**: 887-890.
- FORRESTER, K., C. ALMOGUERA, K. HAN, W. E. GRIZZLE and M. PERUCHO, 1987 Detection of high incidence of K-ras oncogenes during human colon tumorigenesis. *Nature* **327**: 298-303.
- FRANSEN, K., M. KLINTENAS, A. OSTERSTROM, J. DIMBERG, H. J. MONSTEIN *et al.*, 2004 Mutation analysis of the BRAF, ARAF and RAF-1 genes in human colorectal adenocarcinomas. *Carcinogenesis* **25**: 527-533.
- FURUYA, Y., K. AKAKURA, M. MASAI and H. ITO, 1999 Vitamin D receptor gene polymorphism in Japanese patients with prostate cancer. *Endocr J* **46**: 467-470.
- GARNETT, M. J., and R. MARAIS, 2004 Guilty as charged: B-RAF is a human oncogene. *Cancer Cell* **6**: 313-319.
- GIOVANNUCCI, E., 2002 Is the androgen receptor CAG repeat length significant for prostate cancer? *Cancer Epidemiol Biomarkers Prev* **11**: 985-986.
- GIOVANNUCCI, E., E. A. PLATZ, M. J. STAMPFER, A. CHAN, K. KRITHIVAS *et al.*, 1999 The CAG repeat within the androgen receptor gene and benign prostatic hyperplasia. *Urology* **53**: 121-125.
- GLEASON, D. F., and G. T. MELLINGER, 1974 Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. *J Urol* **111**: 58-64.
- GRAFF, J. R., B. W. KONICEK, A. M. MCNULTY, Z. WANG, K. HOUCK *et al.*, 2000 Increased AKT activity contributes to prostate cancer progression by dramatically accelerating prostate tumor growth and diminishing p27Kip1 expression. *J Biol Chem* **275**: 24500-24505.
- GRIMBERG, A., and P. COHEN, 2000 Role of insulin-like growth factors and their binding proteins in growth control and carcinogenesis. *J Cell Physiol* **183**: 1-9.
- GROSS, C., T. R. ECCLESHALL, P. J. MALLOY, M. L. VILLA, R. MARCUS *et al.*, 1996 The presence of a polymorphism at the translation initiation site of the vitamin D receptor gene is associated with low bone mineral density in postmenopausal Mexican-American women. *J Bone Miner Res* **11**: 1850-1855.
- GROSS, C., I. M. MUSIOL, T. R. ECCLESHALL, P. J. MALLOY and D. FELDMAN, 1998 Vitamin D receptor gene polymorphisms: analysis of ligand binding and hormone responsiveness in cultured skin fibroblasts. *Biochem Biophys Res Commun* **242**: 467-473.
- GSUR, A., M. PREYER, G. HAIDINGER, T. ZIDEK, S. MADERSBACHER *et al.*, 2002 Polymorphic CAG repeats in the androgen receptor gene, prostate-specific antigen polymorphism and prostate cancer risk. *Carcinogenesis* **23**: 1647-1651.
- HAAS-KOGAN, D., N. SHALEV, M. WONG, G. MILLS, G. YOUNT *et al.*, 1998 Protein kinase B (PKB/Akt) activity is elevated in glioblastoma cells due to mutation of the tumor suppressor PTEN/MMAC. *Curr Biol* **8**: 1195-1198.

- HAAS, G. P., and W. A. SAKR, 1997 Epidemiology of prostate cancer. *CA Cancer J Clin* **47**: 273-287.
- HABUCHI, T., T. SUZUKI, R. SASAKI, L. WANG, K. SATO *et al.*, 2000 Association of vitamin D receptor gene polymorphism with prostate cancer and benign prostatic hyperplasia in a Japanese population. *Cancer Res* **60**: 305-308.
- HAKIMI, J. M., M. P. SCHOENBERG, R. H. RONDINELLI, S. PIANTADOSI and E. R. BARRACK, 1997 Androgen receptor variants with short glutamine or glycine repeats may identify unique subpopulations of men with prostate cancer. *Clin Cancer Res* **3**: 1599-1608.
- HALACHMI, N., S. HALACHMI, E. EVRON, P. CAIRNS, K. OKAMI *et al.*, 1998 Somatic mutations of the PTEN tumor suppressor gene in sporadic follicular thyroid tumors. *Genes Chromosomes Cancer* **23**: 239-243.
- HANAHAN, D., and R. A. WEINBERG, 2000 The hallmarks of cancer. *Cell* **100**: 57-70.
- HANCHETTE, C. L., and G. G. SCHWARTZ, 1992 Geographic patterns of prostate cancer mortality. Evidence for a protective effect of ultraviolet radiation. *Cancer* **70**: 2861-2869.
- HARDY, D. O., H. I. SCHER, T. BOGENREIDER, P. SABBATINI, Z. F. ZHANG *et al.*, 1996 Androgen receptor CAG repeat lengths in prostate cancer: correlation with age of onset. *J Clin Endocrinol Metab* **81**: 4400-4405.
- HEIDENREICH, A., G. AUS, M. BOLLA, S. JONIAU, V. B. MATVEEV *et al.*, 2008 EAU Guidelines on Prostate Cancer. *Eur Urol* **53**: 68-80.
- HO, G. Y., A. MELMAN, S. M. LIU, M. LI, H. YU *et al.*, 2003 Polymorphism of the insulin gene is associated with increased prostate cancer risk. *Br J Cancer* **88**: 263-269.
- HSIEH, M. C., S. F. LIN, S. J. SHIN, T. C. LIU, J. G. CHANG *et al.*, 2000 Mutation analysis of PTEN/MMAC 1 in sporadic thyroid tumors. *Kaohsiung J Med Sci* **16**: 9-12.
- HSING, A. W., S. CHUA, JR., Y. T. GAO, E. GENTZSCHEIN, L. CHANG *et al.*, 2001 Prostate cancer risk and serum levels of insulin and leptin: a population-based study. *J Natl Cancer Inst* **93**: 783-789.
- HSING, A. W., J. DENG, I. A. SESTERHENN, F. K. MOSTOFI, F. Z. STANCZYK *et al.*, 2000a Body size and prostate cancer: a population-based case-control study in China. *Cancer Epidemiol Biomarkers Prev* **9**: 1335-1341.
- HSING, A. W., Y. T. GAO, G. WU, X. WANG, J. DENG *et al.*, 2000b Polymorphic CAG and GGN repeat lengths in the androgen receptor gene and prostate cancer risk: a population-based case-control study in China. *Cancer Res* **60**: 5111-5116.
- HUANG, C. H., D. MANDELKER, O. SCHMIDT-KITTLER, Y. SAMUELS, V. E. VELCULESCU *et al.*, 2007 The structure of a human p110alpha/p85alpha complex elucidates the effects of oncogenic PI3Kalpha mutations. *Science* **318**: 1744-1748.
- HUEBNER, K., and C. M. CROCE, 2003 Cancer and the FRA3B/FHIT fragile locus: it's a HIT. *Br J Cancer* **88**: 1501-1506.
- IKENOUE, T., F. KANAI, Y. HIKIBA, T. OBATA, Y. TANAKA *et al.*, 2005 Functional analysis of PIK3CA gene mutations in human colorectal cancer. *Cancer Res* **65**: 4562-4567.
- INGLES, S. A., R. K. ROSS, M. C. YU, R. A. IRVINE, G. LA PERA *et al.*, 1997 Association of prostate cancer risk with genetic polymorphisms in vitamin D receptor and androgen receptor. *J Natl Cancer Inst* **89**: 166-170.
- IOANNIDIS, P., L. MAHAIRA, A. PAPAPOPOULOU, M. R. TEIXEIRA, S. HEIM *et al.*, 2003 8q24 Copy number gains and expression of the c-myc mRNA stabilizing protein CRD-BP in primary breast carcinomas. *Int J Cancer* **104**: 54-59.
- IRVINE, R. A., H. MA, M. C. YU, R. K. ROSS, M. R. STALLCUP *et al.*, 2000 Inhibition of p160-mediated coactivation with increasing androgen receptor polyglutamine length. *Hum Mol Genet* **9**: 267-274.

- IRVINE, R. A., M. C. YU, R. K. ROSS and G. A. COETZEE, 1995 The CAG and GGC microsatellites of the androgen receptor gene are in linkage disequilibrium in men with prostate cancer. *Cancer Res* **55**: 1937-1940.
- JELLEMA, A., M. P. ZEEGERS, E. J. FESKENS, P. C. DAGNELIE and R. P. MENSINK, 2003 Gly972Arg variant in the insulin receptor substrate-1 gene and association with Type 2 diabetes: a meta-analysis of 27 studies. *Diabetologia* **46**: 990-995.
- JENSTER, G., H. A. VAN DER KORPUT, C. VAN VROONHOVEN, T. H. VAN DER KWAST, J. TRAPMAN *et al.*, 1991 Domains of the human androgen receptor involved in steroid binding, transcriptional activation, and subcellular localization. *Mol Endocrinol* **5**: 1396-1404.
- KANG, S., A. G. BADER and P. K. VOGT, 2005 Phosphatidylinositol 3-kinase mutations identified in human cancer are oncogenic. *Proc Natl Acad Sci U S A* **102**: 802-807.
- KARAKAS, B., K. E. BACHMAN and B. H. PARK, 2006 Mutation of the PIK3CA oncogene in human cancers. *Br J Cancer* **94**: 455-459.
- KAWAMURA, N., H. NAGAI, K. BANDO, M. KOYAMA, S. MATSUMOTO *et al.*, 1999 PTEN/MMAC1 mutations in hepatocellular carcinomas: somatic inactivation of both alleles in tumors. *Jpn J Cancer Res* **90**: 413-418.
- KAZEMI-ESFARJANI, P., M. A. TRIFIRO and L. PINSKY, 1995 Evidence for a repressive function of the long polyglutamine tract in the human androgen receptor: possible pathogenetic relevance for the (CAG)<sub>n</sub>-expanded neuropathies. *Hum Mol Genet* **4**: 523-527.
- KENNEDY, G. C., M. S. GERMAN and W. J. RUTTER, 1995 The minisatellite in the diabetes susceptibility locus IDDM2 regulates insulin transcription. *Nat Genet* **9**: 293-298.
- KHANDPUR, S., B. S. REDDY and H. KAUR, 2002 Multiple cutaneous metastases from carcinoma of the penis. *J Dermatol* **29**: 296-299.
- KHOSRAVI-FAR, R., M. A. WHITE, J. K. WESTWICK, P. A. SOLSKI, M. CHRZANOWSKA-WODNICKA *et al.*, 1996 Oncogenic Ras activation of Raf/mitogen-activated protein kinase-independent pathways is sufficient to cause tumorigenic transformation. *Mol Cell Biol* **16**: 3923-3933.
- KIBEL, A. S., S. D. ISAACS, W. B. ISAACS and G. S. BOVA, 1998 Vitamin D receptor polymorphisms and lethal prostate cancer. *J Urol* **160**: 1405-1409.
- KOKONTIS, J. M., and S. LIAO, 1999 Molecular action of androgen in the normal and neoplastic prostate. *Vitam Horm* **55**: 219-307.
- LA SPADA, A. R., E. M. WILSON, D. B. LUBAHN, A. E. HARDING and K. H. FISCHBECK, 1991 Androgen receptor gene mutations in X-linked spinal and bulbar muscular atrophy. *Nature* **352**: 77-79.
- LAWLOR, M. A., and P. ROTWEIN, 2000 Insulin-like growth factor-mediated muscle cell survival: central roles for Akt and cyclin-dependent kinase inhibitor p21. *Mol Cell Biol* **20**: 8983-8995.
- LE STUNFF, C., D. FALLIN, N. J. SCHORK and P. BOUGNERES, 2000 The insulin gene VNTR is associated with fasting insulin levels and development of juvenile obesity. *Nat Genet* **26**: 444-446.
- LEE, A. V., J. L. GOOCH, S. OESTERREICH, R. L. GULER and D. YEE, 2000 Insulin-like growth factor I-induced degradation of insulin receptor substrate 1 is mediated by the 26S proteasome and blocked by phosphatidylinositol 3'-kinase inhibition. *Mol Cell Biol* **20**: 1489-1496.
- LEHRER, S., E. J. DIAMOND, S. STAGGER, N. N. STONE and R. G. STOCK, 2002a Increased serum insulin associated with increased risk of prostate cancer recurrence. *Prostate* **50**: 1-3.

- LEHRER, S., E. J. DIAMOND, S. STAGGER, N. N. STONE and R. G. STOCK, 2002b Serum insulin level, disease stage, prostate specific antigen (PSA) and Gleason score in prostate cancer. *Br J Cancer* **87**: 726-728.
- LEVINE, D. A., F. BOGOMOLNIY, C. J. YEE, A. LASH, R. R. BARAKAT *et al.*, 2005 Frequent mutation of the PIK3CA gene in ovarian and breast cancers. *Clin Cancer Res* **11**: 2875-2878.
- LI, J., C. YEN, D. LIAW, K. PODSYPANINA, S. BOSE *et al.*, 1997 PTEN, a putative protein tyrosine phosphatase gene mutated in human brain, breast, and prostate cancer. *Science* **275**: 1943-1947.
- LI, V. S., C. W. WONG, T. L. CHAN, A. S. CHAN, W. ZHAO *et al.*, 2005 Mutations of PIK3CA in gastric adenocarcinoma. *BMC Cancer* **5**: 29.
- LIAW, D., D. J. MARSH, J. LI, P. L. DAHIA, S. I. WANG *et al.*, 1997 Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. *Nat Genet* **16**: 64-67.
- LIN, M., L. J. WEI, W. R. SELLERS, M. LIEBERFARB, W. H. WONG *et al.*, 2004 dChipSNP: significance curve and clustering of SNP-array-based loss-of-heterozygosity data. *Bioinformatics* **20**: 1233-1240.
- LUCASSEN, A. M., C. JULIER, J. P. BERESSI, C. BOITARD, P. FROGUEL *et al.*, 1993 Susceptibility to insulin dependent diabetes mellitus maps to a 4.1 kb segment of DNA spanning the insulin gene and associated VNTR. *Nat Genet* **4**: 305-310.
- LUCASSEN, A. M., G. R. SCREATON, C. JULIER, T. J. ELLIOTT, M. LATHROP *et al.*, 1995 Regulation of insulin gene expression by the IDDM associated, insulin locus haplotype. *Hum Mol Genet* **4**: 501-506.
- LUCIA, M. S., and G. J. MILLER, 1992 Histopathology of malignant lesions of the penis. *Urol Clin North Am* **19**: 227-246.
- LUNDIN, A. C., P. SODERKVIST, B. ERIKSSON, M. BERGMAN-JUNGESTROM and S. WINGREN, 1999 Association of breast cancer progression with a vitamin D receptor gene polymorphism. South-East Sweden Breast Cancer Group. *Cancer Res* **59**: 2332-2334.
- MA, J., M. J. STAMPFER, P. H. GANN, H. L. HOUGH, E. GIOVANNUCCI *et al.*, 1998 Vitamin D receptor polymorphisms, circulating vitamin D metabolites, and risk of prostate cancer in United States physicians. *Cancer Epidemiol Biomarkers Prev* **7**: 385-390.
- MA, L., Z. CHEN, H. ERDJUMENT-BROMAGE, P. TEMPST and P. P. PANDOLFI, 2005 Phosphorylation and functional inactivation of TSC2 by Erk implications for tuberous sclerosis and cancer pathogenesis. *Cell* **121**: 179-193.
- MA, Y. Y., S. J. WEI, Y. C. LIN, J. C. LUNG, T. C. CHANG *et al.*, 2000 PIK3CA as an oncogene in cervical cancer. *Oncogene* **19**: 2739-2744.
- MAEHAMA, T., and J. E. DIXON, 1998 The tumor suppressor, PTEN/MMAC1, dephosphorylates the lipid second messenger, phosphatidylinositol 3,4,5-trisphosphate. *J Biol Chem* **273**: 13375-13378.
- MANNING, B. D., and L. C. CANTLEY, 2003 United at last: the tuberous sclerosis complex gene products connect the phosphoinositide 3-kinase/Akt pathway to mammalian target of rapamycin (mTOR) signalling. *Biochem Soc Trans* **31**: 573-578.
- MARSH, D. J., P. L. DAHIA, Z. ZHENG, D. LIAW, R. PARSONS *et al.*, 1997 Germline mutations in PTEN are present in Bannayan-Zonana syndrome. *Nat Genet* **16**: 333-334.
- MASSION, P. P., W. L. KUO, D. STOKOE, A. B. OLSHEN, P. A. TRESELER *et al.*, 2002 Genomic copy number analysis of non-small cell lung cancer using array comparative genomic hybridization: implications of the phosphatidylinositol 3-kinase pathway. *Cancer Res* **62**: 3636-3640.

- MAYO, L. D., and D. B. DONNER, 2001 A phosphatidylinositol 3-kinase/Akt pathway promotes translocation of Mdm2 from the cytoplasm to the nucleus. *Proc Natl Acad Sci U S A* **98**: 11598-11603.
- MCEWAN, I. J., 2004 Molecular mechanisms of androgen receptor-mediated gene regulation: structure-function analysis of the AF-1 domain. *Endocr Relat Cancer* **11**: 281-293.
- MERCER, K. E., and C. A. PRITCHARD, 2003 Raf proteins and cancer: B-Raf is identified as a mutational target. *Biochim Biophys Acta* **1653**: 25-40.
- MICALI, G., M. R. NASCA, D. INNOCENZI and R. A. SCHWARTZ, 2004 Invasive penile carcinoma: a review. *Dermatol Surg* **30**: 311-320.
- MICALI, G., M. R. NASCA, D. INNOCENZI and R. A. SCHWARTZ, 2006 Penile cancer. *J Am Acad Dermatol* **54**: 369-391; quiz 391-364.
- MICCI, F., M. R. TEIXEIRA, M. SCHEISTROEN, V. M. ABELER and S. HEIM, 2003 Cytogenetic characterization of tumors of the vulva and vagina. *Genes Chromosomes Cancer* **38**: 137-148.
- MILLER, G. J., G. E. STAPLETON, T. E. HEDLUND and K. A. MOFFAT, 1995 Vitamin D receptor expression, 24-hydroxylase activity, and inhibition of growth by 1 $\alpha$ ,25-dihydroxyvitamin D<sub>3</sub> in seven human prostatic carcinoma cell lines. *Clin Cancer Res* **1**: 997-1003.
- MINGHETTI, P. P., and A. W. NORMAN, 1988 1,25(OH)<sub>2</sub>-vitamin D<sub>3</sub> receptors: gene regulation and genetic circuitry. *Faseb J* **2**: 3043-3053.
- MITSUMORI, K., A. TERAI, H. OKA, T. SEGAWA, K. OGURA *et al.*, 1999 Androgen receptor CAG repeat length polymorphism in benign prostatic hyperplasia (BPH): correlation with adenoma growth. *Prostate* **41**: 253-257.
- MORRISON, N. A., J. C. QI, A. TOKITA, P. J. KELLY, L. CROFTS *et al.*, 1994 Prediction of bone density from vitamin D receptor alleles. *Nature* **367**: 284-287.
- MULLER, C. I., C. W. MILLER, W. K. HOFMANN, M. E. GROSS, C. S. WALSH *et al.*, 2007 Rare mutations of the PIK3CA gene in malignancies of the hematopoietic system as well as endometrium, ovary, prostate and osteosarcomas, and discovery of a PIK3CA pseudogene. *Leuk Res* **31**: 27-32.
- NAVE, B. T., M. OUWENS, D. J. WITHERS, D. R. ALESSI and P. R. SHEPHERD, 1999 Mammalian target of rapamycin is a direct target for protein kinase B: identification of a convergence point for opposing effects of insulin and amino-acid deficiency on protein translation. *Biochem J* **344 Pt 2**: 427-431.
- NEJENTSEV, S., L. GODFREY, H. SNOOK, H. RANCE, S. NUTLAND *et al.*, 2004 Comparative high-resolution analysis of linkage disequilibrium and tag single nucleotide polymorphisms between populations in the vitamin D receptor gene. *Hum Mol Genet* **13**: 1633-1639.
- NELSON, K. A., and J. S. WITTE, 2002 Androgen receptor CAG repeats and prostate cancer. *Am J Epidemiol* **155**: 883-890.
- NEUHAUSEN, S. L., M. L. SLATTERY, C. P. GARNER, Y. C. DING, M. HOFFMAN *et al.*, 2005 Prostate cancer risk and IRS1, IRS2, IGF1, and INS polymorphisms: strong association of IRS1 G972R variant and cancer risk. *Prostate* **64**: 168-174.
- NTAIS, C., A. POLYCARPOU and J. P. IOANNIDIS, 2003 Vitamin D receptor gene polymorphisms and risk of prostate cancer: a meta-analysis. *Cancer Epidemiol Biomarkers Prev* **12**: 1395-1402.
- OKUDELA, K., M. SUZUKI, S. KAGEYAMA, T. BUNAI, K. NAGURA *et al.*, 2007 PIK3CA mutation and amplification in human lung cancer. *Pathol Int* **57**: 664-671.
- OSTRANDER, E. A., and J. L. STANFORD, 2000 Genetics of prostate cancer: too many loci, too few genes. *Am J Hum Genet* **67**: 1367-1375.

- PEEHL, D. M., R. J. SKOWRONSKI, G. K. LEUNG, S. T. WONG, T. A. STAMEY *et al.*, 1994 Antiproliferative effects of 1,25-dihydroxyvitamin D3 on primary cultures of human prostatic cells. *Cancer Res* **54**: 805-810.
- PERSSON, B., J. G. SJODIN, L. HOLMBERG and T. WINDAHL, 2007 The National Penile Cancer Register in Sweden 2000-2003. *Scand J Urol Nephrol* **41**: 278-282.
- PRICE, D. K., M. E. FRANKS and W. D. FIGG, 2004 Genetic variations in the vitamin D receptor, androgen receptor and enzymes that regulate androgen metabolism. *J Urol* **171**: S45-49; discussion S49.
- PRZYBOJEWSKA, B., A. JAGIELLO and P. JALMUZNA, 2000 H-RAS, K-RAS, and N-RAS gene activation in human bladder cancers. *Cancer Genet Cytogenet* **121**: 73-77.
- RACZ, A., N. BRASS, D. HECKEL, S. PAHL, K. REMBERGER *et al.*, 1999 Expression analysis of genes at 3q26-q27 involved in frequent amplification in squamous cell lung carcinoma. *Eur J Cancer* **35**: 641-646.
- REDON, R., D. MULLER, K. CAULEE, K. WANHERDRICK, J. ABECASSIS *et al.*, 2001 A simple specific pattern of chromosomal aberrations at early stages of head and neck squamous cell carcinomas: PIK3CA but not p63 gene as a likely target of 3q26-qter gains. *Cancer Res* **61**: 4122-4129.
- RODRIGUEZ-VICIANA, P., P. H. WARNE, R. DHAND, B. VANHAESEBROECK, I. GOUT *et al.*, 1994 Phosphatidylinositol-3-OH kinase as a direct target of Ras. *Nature* **370**: 527-532.
- RODRIGUEZ-VICIANA, P., P. H. WARNE, B. VANHAESEBROECK, M. D. WATERFIELD and J. DOWNWARD, 1996 Activation of phosphoinositide 3-kinase by interaction with Ras and by point mutation. *Embo J* **15**: 2442-2451.
- ROMASHKOVA, J. A., and S. S. MAKAROV, 1999 NF-kappaB is a target of AKT in anti-apoptotic PDGF signalling. *Nature* **401**: 86-90.
- ROSS, R. K., M. C. PIKE, G. A. COETZEE, J. K. REICHARDT, M. C. YU *et al.*, 1998 Androgen metabolism and prostate cancer: establishing a model of genetic susceptibility. *Cancer Res* **58**: 4497-4504.
- RUBIN, M. A., B. KLETER, M. ZHOU, G. AYALA, A. L. CUBILLA *et al.*, 2001 Detection and typing of human papillomavirus DNA in penile carcinoma: evidence for multiple independent pathways of penile carcinogenesis. *Am J Pathol* **159**: 1211-1218.
- SAITO, M., A. OKAMOTO, T. KOHNO, S. TAKAKURA, H. SHINOZAKI *et al.*, 2000 Allelic imbalance and mutations of the PTEN gene in ovarian cancer. *Int J Cancer* **85**: 160-165.
- SAKAI, A., C. THIEBLEMONT, A. WELLMANN, E. S. JAFFE and M. RAFFELD, 1998 PTEN gene alterations in lymphoid neoplasms. *Blood* **92**: 3410-3415.
- SAMUELS, Y., L. A. DIAZ, JR., O. SCHMIDT-KITTLER, J. M. CUMMINS, L. DELONG *et al.*, 2005 Mutant PIK3CA promotes cell growth and invasion of human cancer cells. *Cancer Cell* **7**: 561-573.
- SAMUELS, Y., Z. WANG, A. BARDELLI, N. SILLIMAN, J. PTAK *et al.*, 2004 High frequency of mutations of the PIK3CA gene in human cancers. *Science* **304**: 554.
- SARBIA, M., J. ARJUMAND, M. WOLTER, G. REIFENBERGER, H. HEEP *et al.*, 2001 Frequent c-myc amplification in high-grade dysplasia and adenocarcinoma in Barrett esophagus. *Am J Clin Pathol* **115**: 835-840.
- SARTOR, O., Q. ZHENG and J. A. EASTHAM, 1999 Androgen receptor gene CAG repeat length varies in a race-specific fashion in men without prostate cancer. *Urology* **53**: 378-380.
- SAWAYA, M. E., and A. R. SHALITA, 1998 Androgen receptor polymorphisms (CAG repeat lengths) in androgenetic alopecia, hirsutism, and acne. *J Cutan Med Surg* **3**: 9-15.
- SCHUBERT, M., D. P. BRAZIL, D. J. BURKS, J. A. KUSHNER, J. YE *et al.*, 2003 Insulin receptor substrate-2 deficiency impairs brain growth and promotes tau phosphorylation. *J Neurosci* **23**: 7084-7092.

- SCHWARTZ, G. G., and B. S. HULKA, 1990 Is vitamin D deficiency a risk factor for prostate cancer? (Hypothesis). *Anticancer Res* **10**: 1307-1311.
- SHAW, R. J., and L. C. CANTLEY, 2006 Ras, PI(3)K and mTOR signalling controls tumour cell growth. *Nature* **441**: 424-430.
- SHAYESTEH, L., Y. LU, W. L. KUO, R. BALDOCCHI, T. GODFREY *et al.*, 1999 PIK3CA is implicated as an oncogene in ovarian cancer. *Nat Genet* **21**: 99-102.
- SHIBATA, A., T. A. STAMEY, J. E. MCNEAL, I. CHENG and D. M. PEEHL, 2001 Genetic polymorphisms in the androgen receptor and type II 5 alpha-reductase genes in prostate enlargement. *J Urol* **166**: 1560-1564.
- SKOWRONSKI, R. J., D. M. PEEHL and D. FELDMAN, 1993 Vitamin D and prostate cancer: 1,25 dihydroxyvitamin D3 receptors and actions in human prostate cancer cell lines. *Endocrinology* **132**: 1952-1960.
- SLATTERY, M. L., W. SAMOWITZ, K. CURTIN, K. N. MA, M. HOFFMAN *et al.*, 2004 Associations among IRS1, IRS2, IGF1, and IGFBP3 genetic polymorphisms and colorectal cancer. *Cancer Epidemiol Biomarkers Prev* **13**: 1206-1214.
- SLATTERY, M. L., C. SWEENEY, R. WOLFF, J. HERRICK, K. BAUMGARTNER *et al.*, 2007 Genetic variation in IGF1, IGFBP3, IRS1, IRS2 and risk of breast cancer in women living in Southwestern United States. *Breast Cancer Res Treat* **104**: 197-209.
- STANFORD, J. L., and E. A. OSTRANDER, 2001 Familial prostate cancer. *Epidemiol Rev* **23**: 19-23.
- STEAD, J. D., and A. J. JEFFREYS, 2000 Allele diversity and germline mutation at the insulin minisatellite. *Hum Mol Genet* **9**: 713-723.
- TAYLOR, J. A., A. HIRVONEN, M. WATSON, G. PITTMAN, J. L. MOHLER *et al.*, 1996 Association of prostate cancer with vitamin D receptor gene polymorphism. *Cancer Res* **56**: 4108-4110.
- THOMPSON, I. M., D. K. PAULER, P. J. GOODMAN, C. M. TANGEN, M. S. LUCIA *et al.*, 2004 Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. *N Engl J Med* **350**: 2239-2246.
- TONON, G., K. K. WONG, G. MAULIK, C. BRENNAN, B. FENG *et al.*, 2005 High-resolution genomic profiles of human lung cancer. *Proc Natl Acad Sci U S A* **102**: 9625-9630.
- TRAHEY, M., and F. MCCORMICK, 1987 A cytoplasmic protein stimulates normal N-ras p21 GTPase, but does not affect oncogenic mutants. *Science* **238**: 542-545.
- TSEN, H. F., H. MORGENSTERN, T. MACK and R. K. PETERS, 2001 Risk factors for penile cancer: results of a population-based case-control study in Los Angeles County (United States). *Cancer Causes Control* **12**: 267-277.
- WANG, S. I., J. PUC, J. LI, J. N. BRUCE, P. CAIRNS *et al.*, 1997 Somatic mutations of PTEN in glioblastoma multiforme. *Cancer Res* **57**: 4183-4186.
- WANG, Y., A. HELLAND, R. HOLM, G. B. KRISTENSEN and A. L. BORRESEN-DALE, 2005 PIK3CA mutations in advanced ovarian carcinomas. *Hum Mutat* **25**: 322.
- VANHAESEBROECK, B., and M. D. WATERFIELD, 1999 Signaling by distinct classes of phosphoinositide 3-kinases. *Exp Cell Res* **253**: 239-254.
- WATANABE, M., K. FUKUTOME, M. MURATA, H. UEMURA, Y. KUBOTA *et al.*, 1999 Significance of vitamin D receptor gene polymorphism for prostate cancer risk in Japanese. *Anticancer Res* **19**: 4511-4514.
- VERBEEK, W., A. F. GOMBART, M. SHIOHARA, M. CAMPBELL and H. P. KOEFFLER, 1997 Vitamin D receptor: no evidence for allele-specific mRNA stability in cells which are heterozygous for the Taq I restriction enzyme polymorphism. *Biochem Biophys Res Commun* **238**: 77-80.

- WILTING, S. M., P. J. SNIJDERS, G. A. MEIJER, B. YLSTRA, P. R. VAN DEN ISSSEL *et al.*, 2006 Increased gene copy numbers at chromosome 20q are frequent in both squamous cell carcinomas and adenocarcinomas of the cervix. *J Pathol* **209**: 220-230.
- WU, K. J., C. GRANDORI, M. AMACKER, N. SIMON-VERMOT, A. POLACK *et al.*, 1999 Direct activation of TERT transcription by c-MYC. *Nat Genet* **21**: 220-224.
- YAKUT, T., H. J. SCHULTEN, A. DEMIR, D. FRANK, B. DANNER *et al.*, 2006 Assessment of molecular events in squamous and non-squamous cell lung carcinoma. *Lung Cancer* **54**: 293-301.
- YEN, C. C., Y. J. CHEN, J. T. CHEN, J. Y. HSIA, P. M. CHEN *et al.*, 2001 Comparative genomic hybridization of esophageal squamous cell carcinoma: correlations between chromosomal aberrations and disease progression/prognosis. *Cancer* **92**: 2769-2777.
- YOKOYAMA, Y., X. WAN, A. SHINOHARA, S. TAKAHASHI, Y. TAKAHASHI *et al.*, 2000 Expression of PTEN and PTEN pseudogene in endometrial carcinoma. *Int J Mol Med* **6**: 47-50.
- YORDY, J. S., and R. C. MUISE-HELMERICKS, 2000 Signal transduction and the Ets family of transcription factors. *Oncogene* **19**: 6503-6513.
- YOSHIDA, K. I., M. YANO, K. CHIBA, M. HONDA and S. KITAHARA, 1999 CAG repeat length in the androgen receptor gene is enhanced in patients with idiopathic azoospermia. *Urology* **54**: 1078-1081.
- ZEEGERS, M. P., A. JELLEMA and H. OSTRER, 2003 Empiric risk of prostate carcinoma for relatives of patients with prostate carcinoma: a meta-analysis. *Cancer* **97**: 1894-1903.
- ZEEGERS, M. P., L. A. KIEMENEY, A. M. NIEDER and H. OSTRER, 2004 How strong is the association between CAG and GGN repeat length polymorphisms in the androgen receptor gene and prostate cancer risk? *Cancer Epidemiol Biomarkers Prev* **13**: 1765-1771.
- ZHANG, H., G. LIU, M. DZIUBINSKI, Z. YANG, S. P. ETHIER *et al.*, 2007 Comprehensive analysis of oncogenic effects of PIK3CA mutations in human mammary epithelial cells. *Breast Cancer Res Treat.*
- ZHOU, X., S. C. MOK, Z. CHEN, Y. LI and D. T. WONG, 2004 Concurrent analysis of loss of heterozygosity (LOH) and copy number abnormality (CNA) for oral premalignancy progression using the Affymetrix 10K SNP mapping array. *Hum Genet* **115**: 327-330.