ERG AND SIM2 TRANSCRIPTION FACTORS IN PROSTATE CANCER – THEIR ROLE IN CARCINOGENESIS AND POTENTIAL CLINICAL UTILITY

Kari Rostad

Dissertation for the degree philosophiae doctor (PhD)

The Gade Institute

University of Bergen
Bergen, Norway

2011

ISBN: 978-82-308-1631-8
Bergen, Norway
2011

CONTENTS

ACKNOWLEDGEMENTS

This work was carried out at the Section for Microbiology and Immunology, The Gade Institute, University of Bergen, during the period 2003-2009. Financial support by the University of Bergen made this study possible.

I would like to express my gratitude to my main supervisor Karl-Henning Kalland and co-supervisor Lars A. Akslen, for providing me with the opportunity to join their research and guide me through my PhD-period. Karl-Henning Kalland introduced me to the field of prostate cancer research, and his enthusiasm, hard-working capacity and impressive scientific knowledge have been of great inspiration to me. I am grateful for his continuous support and encouragement during these years.

I would like to thank my co-authors for their valuable and important contributions to the papers: Anne Margrete Øyan, Ole Johan Halvorsen, Monica Mannelqvist, Hanne Puntervoll, Trond Hellem Bø, Laila Stordrange, Sue Olsen, Svein Andreas Haukaas, Biaoyang Lin, Leroy Hood, Inge Jonassen, Olaf J. C. Hellwinkel, Alexander Haese, Lars Budäus, Heiko Albrecht and Thorsten Schlomm.

Colleagues and friends in the group, especially Anne Margrete Øyan, Xisong Ke, and Yi Qu, are thanked and appreciated for their encouragement and help, and for providing me with a friendly working atmosphere. The technical support provided by Beth Johannessen and Hua My Hoang, as well as their friendly enthusiasm, has been greatly appreciated. At the Section for Pathology, I am thankful to Solrun Steine for her technical assistance in purifying DNA from the urine samples, and to Hanne Puntervoll and Monica Mannelqvist for their assistance with the sequencing of nucleic acids, as well as for many much appreciated lunches, coffee-breaks and girls-outings.

I am very grateful to all my friends who have shown me support and encouragement during these exciting, challenging and interesting, but also sometimes difficult and frustrating years. You are all greatly appreciated.

I would like to direct a warm thank you to my parents-in-law, Astrid and Bjarne, for their encouragement and for believing in me. Their support and help made busy, challenging times a little easier.

Finally, I am deeply grateful to my husband, Per, who has always been supportive and caring, and encouraged me all the way, especially during difficult and hard-working times. Per and our three wonderful boys, Magnus (11), David (5) and Erik (5), are a constant joy and make my life meaningful.

Bergen, October 2010 Kari Rostad

LIST OF PUBLICATIONS

This thesis is based on the following three papers, which are referred to by their roman numerals in the text:

I. Rostad K, Mannelqvist M, Halvorsen OJ, Øyan AM, Bø TH, Stordrange L, Olsen S, Haukaas SA, Lin B, Hood L, Jonassen I, Akslen LA, Kalland KH. *ERG* upregulation and related *ETS* transcription factors in prostate cancer. *Int J Oncol* 2007;30:19-32.

II. Halvorsen OJ, Rostad K, Øyan AM, Puntervoll H, Bø TH, Stordrange L, Olsen S, Haukaas SA, Hood L, Jonassen I, Kalland KH, Akslen LA. Increased expression of SIM2-s protein is a novel marker of aggressive prostate cancer. *Clin Cancer Res* 2007; 13(3):892-7.

III. Rostad K, Hellwinkel OJC, Haukaas SA, Halvorsen OJ, Øyan AM, Haese A, Budäus L, Albrecht H, Akslen LA, Schlomm T, Kalland KH. *TMPRSS2:ERG* fusion transcripts in urine from prostate cancer patients correlate with a less favourable prognosis. *APMIS* 2009;117:575-82.

ABBREVIATIONS

AR	androgen receptor
B	benign prostate tissue sample
BPH	benign prostatic hyperplasia
cDNA	complementary DNA
EMT	epithelial to mesenchymal transition
HGPIN	high-grade PIN
IHC	immunohistochemistry
ISH	*in situ* hybridization
mRNA	messenger RNA
PIA	proliferative inflammatory atrophy
PIN	prostatic intraepithelial neoplasia
s-PSA	serum-prostate specific antigen
SI	staining index
T	prostate tumour (cancer) tissue sample
TMA	tissue microarray
qPCR	quantitative polymerase chain reaction

INTRODUCTION

An increasing number of men are diagnosed with prostate cancer each year. Prostate cancer is now the most common form of cancer in men and the second leading cause of cancer deaths after lung cancer.[1] There is presently no efficient cure for disseminated, androgen-independent prostate cancer. Some of the major challenges of prostate cancer are to diagnose the cancer at the earliest possible stage and be able to characterize the molecular nature of the cancer tumour to discern between the patients who need active, invasive treatment versus those who will benefit from active surveillance. It would also be preferable to be able to tailor the treatment regime according to the nature of the tumour. Although of clinical diagnostic and prognostic value, the currently widely used PSA (prostate specific antigen) biomarker does not fulfil the above mentioned requirements, and there is an immense search for new more specific and sensitive biomarkers in order to provide the necessary information regarding screening, classification, prognosis and prediction. Our motivation to start genome-wide gene expression analysis of prostate cancer was the hope that this could be a very promising strategy to identify novel markers, to understand better the molecular mechanisms of prostate carcinogenesis and progression and that this might again be very useful for the discovery of potential therapeutic molecular targets.

INCIDENCE

In Norway, prostate cancer is the most frequent form of cancer in men, with 4168 out of 14000 new cases in 2008, which account for approximately 30 % of all cancers in men that year. Over the 5-year period 2004-2008 prostate cancer accounted for 29 % of all cancer cases in men.

The vast majority of men diagnosed with prostate cancer are over the age of 50 (Fig. 1).[1, 2] In Norway during the years 2004-2008, 91.3 % of prostate cancer cases occurred in men aged 50 years or older.[1] The mean age at prostate cancer diagnosis is 72-74 years.[2] Over the years the annual number of prostate cancer cases has increased. Since the five-year period 1956-60 to the five-year period 2001-2005 there has been a

five-time increase in the reported number of cases in Norway. This trend is worldwide and largely due to increased PSA testing since it became commercially available in 1989, also in men who do not have any symptoms, and a growing and aging population. The Cancer Registry of Norway predicts that there will be an increase in annually reported cases of 50 % by 2020.[3, 4] In terms of number of cancer deaths in 2007, lung cancer ranks first in men (1224 deaths) closely followed by prostate cancer (1090 deaths; 19.3%). There has been a steady increase in the mortality rates since the 1970s, but there is some evidence that recent mortality trends are more favourable, probably due to advances in early diagnosis, therapy and cancer care.[1] Norway is one of the countries with the highest number of annual cases and deaths due to prostate cancer, but there is a general increase internationally as well. Prostate cancer accounts for approximately 10% of all malignant tumours in men worldwide. Incidences of prostate cancer vary widely between ethnic populations and countries, with the lowest rates in Asia and the highest in North America and Scandinavia, especially in African American people in the USA.[2] Relative to Caucasians, prostate cancer incidence is 66% higher in African Americans and 39% lower in Asian Americans.[5]

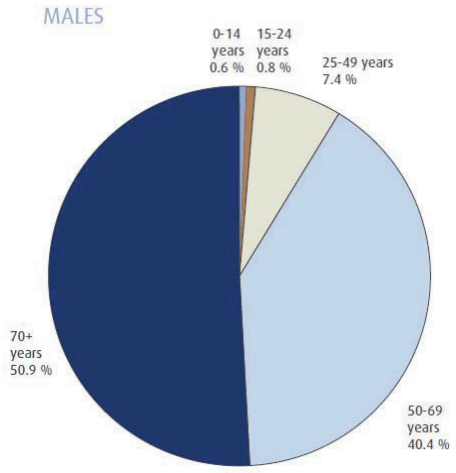

Figure 1. Percentage distribution of cancer incidence by age in males, 2004-2008. (Cancer in Norway 2008)[1]

AETIOLOGY AND RISK FACTORS

The factors that contribute to an increased risk of or predisposition for prostate cancer are complex and not entirely understood. Both genetic dispositions, environmental factors including geographical location and diet and ethnic origin play a role in the elucidation of this complex picture.

As mentioned above, there are ethnic variations concerning the risk of developing prostate cancer. Migration studies have shown that when people from low incidence areas (*i.e.* Japan) move to areas with higher incidence (*i.e.* USA), the incidence of prostate cancer increases, but the ethnic influence is still present. The increase is to about 50% of the rate for white Caucasians and to 25% of that for African-American people in the USA.[6]

While most prostate cancers are sporadic, a hereditary predisposition to prostate cancer has been identified. Familial prostate cancer is estimated to account for 10% to 20% of all cases of prostate cancer, and 5% to 10% of all cases are considered hereditary[7, 8] and associated with early onset disease. The distinction between familial and hereditary prostate cancer relates to the number of family members and generations affected.[8] Men with hereditary prostate cancer are diagnosed an average of 6-7 years earlier than sporadic prostate cancer cases.[9] Prostate cancer is genetically heterogeneous and several genes are likely to contribute to disease susceptibility. Hereditary candidate susceptibility genes with high penetrance have been identified, including *HPC1*/*RNASEL*, *HPC2*/*ELAC2*, *HPCX*, *MSR1* and *PCAP*.[10, 11] Linkage analysis based on genome-wide scans has mapped susceptibility loci for prostate cancer to chromosomes 1, X, 20, 17 and 8. Low-penetrance polymorphisms, in these genes and others, including the *androgen receptor* (*AR*), *vitamin D-receptor*, *CYP17* and *SRD5A2*, seem to play a role in the risk of developing prostate cancer.[2] Also, men with germ-line mutations in the breast/ovarian cancer susceptibility genes *BRCA1* and *BRCA2* are at greater risk of developing prostate cancer, with a higher risk for carriers of *BRCA2* than *BRCA1* mutations.[12, 13] Data from Iceland indicate that men with a mutation in *BRCA2* are at particularly high risk of developing poorly differentiated, disseminated prostate cancer.[14]

Prostate cancer has been associated with a Western lifestyle, in particular a diet with a high intake of fat, meat and dairy products.[2, 15-17] High intakes of α–linolenic acid and calcium are associated with prostate cancer[2, 18] while a high intake of phyto-oestrogens (in for example soybeans) and anti-oxidative compounds like tomato-based products high in lycopene, and the micronutrients selenium, vitamin E (α-tocopherol) and omega-3 fatty acids seem to reduce the risk of prostate cancer.[19-24]

Androgens are important for the development of the normal prostate and withdrawal of testosterone is a well known and effective treatment for prostate cancer, but studies demonstrating the importance of varying androgen concentrations in prostate cancer are few and uncertain. Stattin *et al.* showed that high levels of circulating testosterone are not associated with increased prostate cancer risk.[2, 25] High concentrations of the Insulin growth factor I (IGF-I), a peptide growth factor, increases the risk of prostate cancer and is proposed to represent a link between the western lifestyle and prostate cancer.

Viruses are known etiological agents accounting for approximately 20-25% of human cancers. Recently, the newly discovered gammaretrovirus xenotropic murine leukaemia virus-related virus (XMRV) has been identified in a percentage (0-27%) of prostate cancers with positive findings in the USA, but negative in European studies.[26-29] The incidence of the virus remains uncertain, but there seems to be a possible association between viral infection and prostate cancer, with different possible models of carcinogenesis dependent on whether epithelial or stromal cells are infected. The possible molecular mechanisms and etiological role in prostate carcinogenesis remain uncertain and needs to be studied further.

THE PROSTATE AND CHARACTERISTICS OF PROSTATE CANCER

The prostate

In 335 B.C. Herophilus of Alexandria used the word '*prohistani*' (Greek), which means 'to stand in front of', to describe the organ located 'in front of' the urinary bladder (Fig. 2). Although the existence of the prostate has been known for more than

two thousand years, accurate anatomical description of the gland did not appear until the Renaissance, with illustrations of the prostate and seminal vesicles by Regnier de Graaf around 1660.[30]

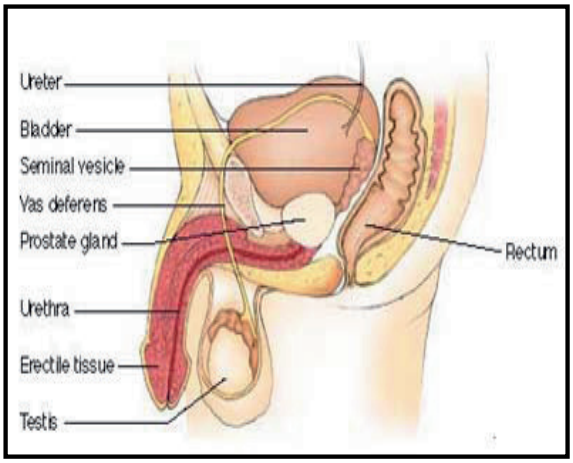

Figure 2. A sketch of the anatomical location of the prostate gland in men.

The prostate consists of three anatomical zones; the peripheral, transitional and central zone[31] (Fig. 3). Few biochemical differences between the epithelial cells of the three zones have been demonstrated.

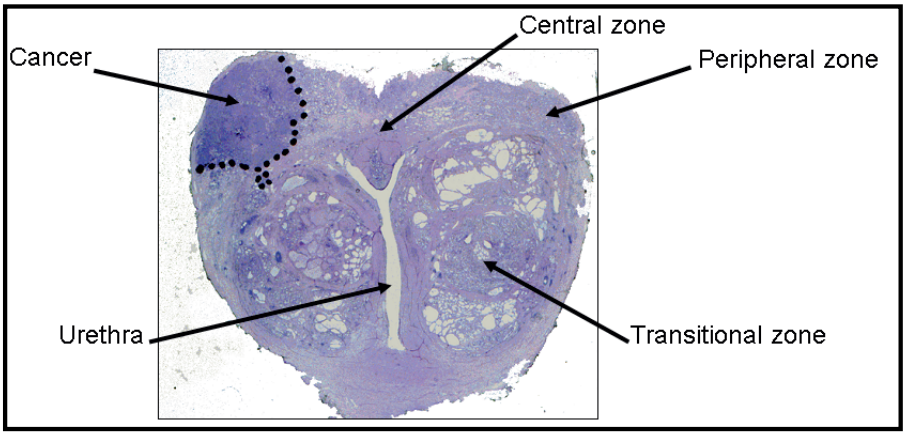

Figure 3. An anatomical horizontal cross section of the prostate, displaying the three anatomical zones as well as the presence of cancer in the peripheral zone. (Picture courtesy by Lars A. Akslen and Ole J. Halvorsen).

The central zone does differ in containing a relatively large proportion of epithelial cells containing pepsinogen 2.[32] Epidermal growth factor (EGF) receptors also seem to be present in a greater concentration in the central and transition zones than in the peripheral zone.

The prostate is an exocrine walnut-sized gland consisting of ductal-acinar structures embedded in stromal tissue.[33] The acini are lined by well-differentiated secretory or luminal epithelial cells, which are androgen dependent and secrete proteins like Prostate specific antigen (PSA) into the lumen of the duct. These cells are surrounded by an underlying layer of proliferating non-secretory basal epithelial cells, that are primarily androgen independent and rest on the basement membrane, separating the epithelial cells from the surrounding stroma (Fig. 4). The basal cells express high molecular weight cytokeratines and p63, as opposed to luminal cells. Their absence is used as a marker of prostate cancer. In addition, rare neuroendocrine cells are present and are believed to be involved in the regulation of prostatic secretory activity and cell growth. The stroma surrounding the prostate is composed of smooth muscle cells, fibroblasts, lymphocytes and neurovascular tissue in a supporting extracellular matrix.[34-37]

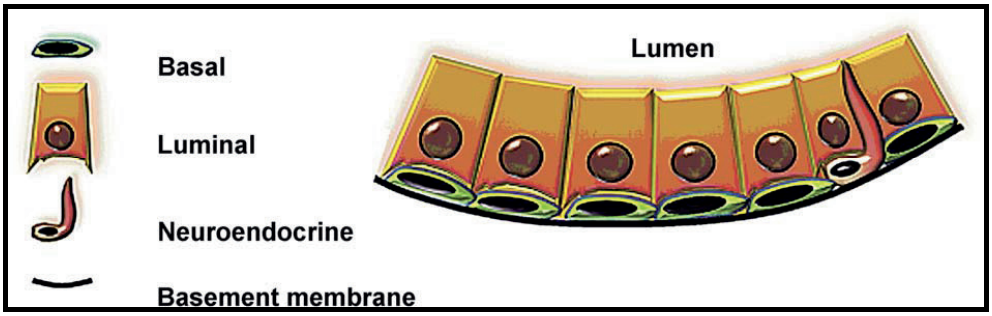

Figure 4. Three cell types in adult prostate epithelium. Basal cells (green) line the outside of the gland and reside against the basement membrane (black). Luminal cells (orange) contact the basal layer and the fluid-filled lumen. Rare neuroendocrine cells (red) are typically found in the basal layer with neurite-like extensions that can approach the luminal layer. (Goldstein *et al.*, Mol. Oncol.,2010, 1-12)[35].

Characteristics of prostate cancer

Most prostate cancers occur in the peripheral zone, less than 30% occur in the transitional zone and these have lower biochemical recurrence rates and are less malignant than tumours originating in the peripheral zone.[38] The transitional zone is the predilection site of benign prostate hyperplasia (BPH).

Prostatic intraepithelial neoplasia (PIN) which progresses to high-grade PIN (HGPIN) is considered the precursor lesion of prostate cancer.[39-41] In HGPIN, the basal layer is present, but it shares otherwise many phenotypic similarities with cancer. HGPIN is characterized by benign prostatic acini and ducts, lined by cytologically atypical cells with prominent nucleoli in many cells, nuclear enlargement, nuclear crowding, an increased density of the cytoplasm and variation in nucleolar size. The volume of HGPIN has a positive correlation with the risk of cancer, tumour stage and Gleason grade.[42, 43] Proliferative inflammatory atrophy (PIA) is described as discrete foci of proliferative glandular epithelium, with morphological appearance of simple atrophy or postatrophic hyperplasia occurring in association with inflammation.[44, 45] Etiological and pathological findings suggest that PIA may be involved in prostate carcinogenesis as maybe a very early precursor followed by HGPIN and malignant transformation,[46, 47] PIA has also been suggested to represent the intermediate luminal cell type suggested to be the target of neoplastic transformation in prostate cancer.[48] The role of PIA in prostate cancer is uncertain and needs to be studied further.

Recently, a model outlining the hierarchial relationship between the cells in the prostate epithelium was suggested[35] (Fig. 5), with implications for the presumed cell of origin for prostate cancer. The cancer initiating cell type has remained unclear[49] (Fig. 6). Pathology observations, showing that more than 95% of prostate cancers express luminal markers with absence of basal cells, have led many to propose luminal cells as the source of prostate cancer (Fig. 6 – benign prostate tissue section to the left and prostate cancer tissue section to the right). The alternative stem cell hypothesis, however, proposes that a cancer stem cell might be the cancer initiating cell.[50] A third scenario is that differentiation blockage at any intermediate developmental stage towards terminal luminal differentiation may give rise to cancer initiating cells. Very recently, strong evidence was presented that prostate cancer may originate among

basal cells[51] and that the basal compartment and the luminal compartment may be capable of proliferating independently.[52] Different possibilities are, however, not mutually exclusive, and further investigation into both normal lineage differentiation and prostate carcinogenesis is required.

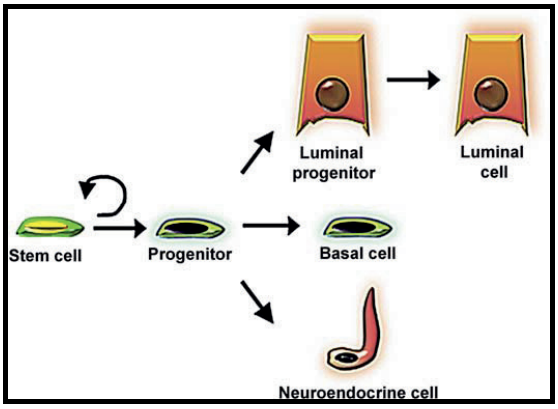

Figure 5. Proposed model for the prostate epithelial hierarchy. Stem cells within the basal layer likely give rise to multi-potent progenitor or intermediate cells that generate all three epithelial cell types. Evidence supports the existence of a luminal-restricted progenitor that can give rise to mature luminal cells. (Goldstein *et al.*, Mol. Oncol.,2010, 1-12)[35]

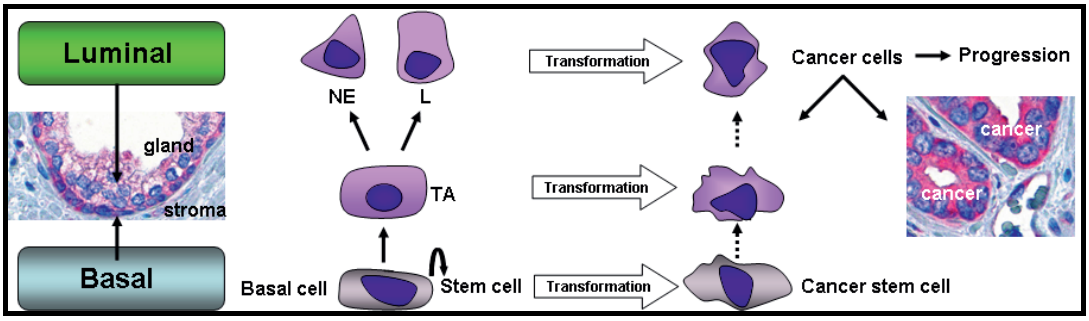

Figure 6. A schematic to illustrate alternative relationships between different epithelial cell types in the prostate gland as well as stained histological sections from prostate benign (left) and cancer (right) tissue from Rostad & Mannelqvist *et al.* (Paper I).

Clinical and histopathological factors

The initial TNM classification (before biopsy) and Gleason grading (after biopsy or surgery) are useful and widely applied prognostic tools in the assessment of prostate

cancer. The Gleason histological grading system, developed by Gleason in 1966[53] and later revised,[54, 55] is based on the histological architectural pattern of the tumour (Fig. 7).

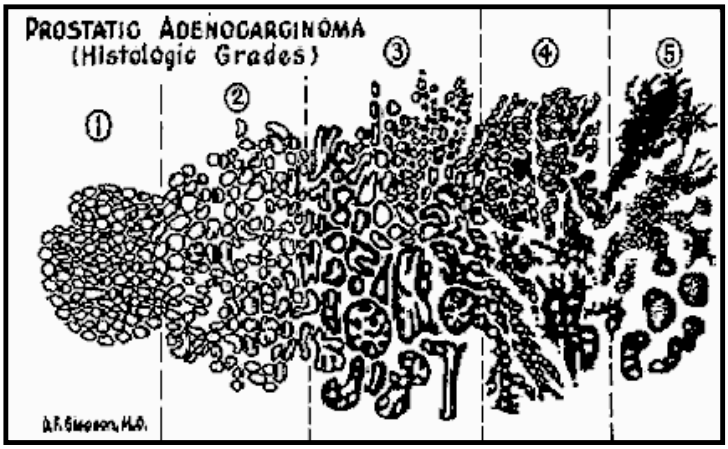

Figure 7. The Gleason grading system in which the sum of the two most prominent histological grades between one and five gives the Gleason score.[56]

The grade is defined as the sum of the two most common growth patterns among five different patterns (grade 1-5) and reported as the Gleason score, thereby taking into consideration the heterogeneity of prostate cancer. This histologic grading is a powerful predictor of progression, and the prognosis of the cancer is more adverse with higher Gleason score.[57, 58] The TNM classification (T-primary tumour; N-lymph node status; M-distant metastasis) is the most widely used system for prostate cancer clinical staging, in which stage T1 is clinically unsuspected prostate cancers, stage T2 is prostate-confined cancer and stages T3 and T4 are tumours that transgress the boundaries of the prostatic gland (extension into the periprostatic tissue and/or seminal vesicle invasion (T3) with possible metastasis to other organs (T4)).[59, 60]

CHARACTERISTICS OF THE ETS-FAMILY OF TRANSCRIPTION FACTORS

A transcription factor is any protein required to initiate or regulate gene transcription. The ETS family of nuclear transcription factors consists of approximately 30 evolutionary conserved members in mammals, of which 27 have been identified in humans. The founding member of the *ETS* (E26 transformation-specific) (E-twenty-six specific) gene family, *v-ets*, was originally identified as a *gag-myb-ets* fusion oncogene of the avian transforming retrovirus E26, which induces both erythroblastic and myeloblastic leukaemia in chickens.[61, 62] A characteristic feature of this family is that they share an evolutionary conserved winged helix-turn-helix DNA binding ETS domain of about 85 amino acid residues, which mediates binding to purine-rich DNA sequences with a central GGAA/T core consensus, the ETS binding site (EBS), and additional flanking nucleotides.[63] It is one of the largest transcription factor families and based upon their structural composition and similarities in the ETS domain they are divided into 11 subfamilies (Fig 8).

Most of the *ETS*-family members have the ETS domain in their C-terminal regions, although some have the domain in their N-terminal regions. In addition, a subset of ETS family proteins (ETS, ERG, ELG, ESE, TEL and PDEF) has another conserved domain called the Pointed domain (PNT) at their N-terminal regions, which forms a helix-loop-helix structure for protein-protein interactions. Some ETS proteins (TEL, ERF and TCF) contain a repressor domain and the majority (ETS, ERG, ELG, PEA3, ESE, SPI and TCF) contain a transcription activation domain (TAD).[64-66] The ETS family of proteins displays distinct DNA binding specificities. The ETS domain and the flanking amino acid sequences of the proteins influence the DNA binding affinity, and alterations of single amino acids in the ETS domain can change its DNA binding specificities.

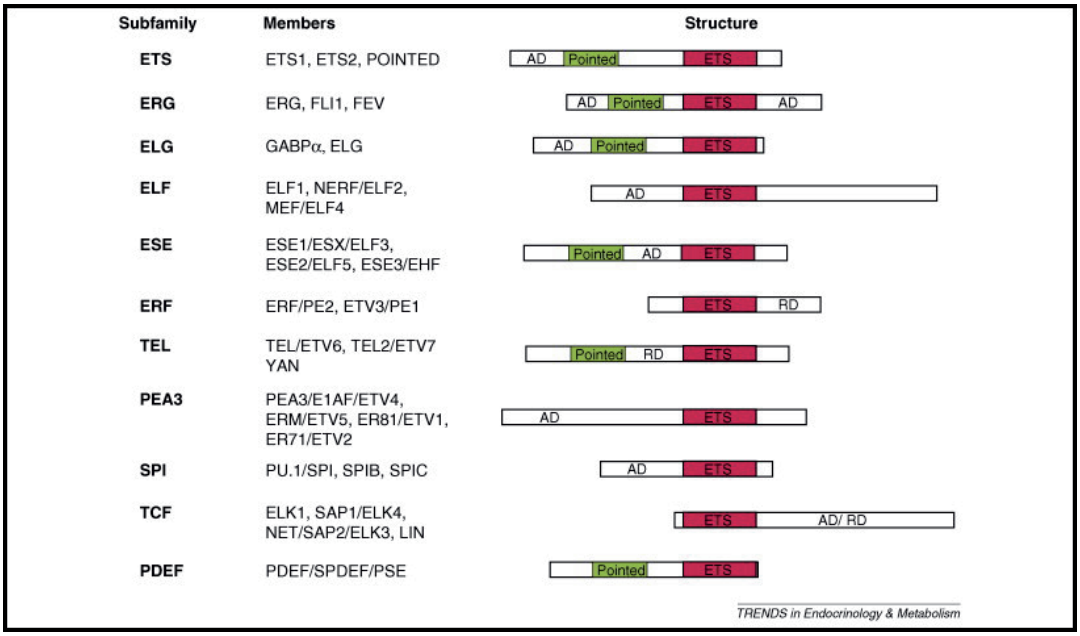

Figure 8. The ETS family of transcription factors. The main functional domains characteristic of members of each ETS sub family are depicted; alternative names for each member are given. Domains: AD, transcriptional activation domain; ETS, DNA binding domain; Pointed, basic helix–loop–helix pointed domain; RD, transcriptional repressor domain. Protein abbreviations: E1AF, E1A enhancer binding protein; EHF, ETS homologous factor; ELF, E74-like factor; ELG, ETS like gene; ER81, ETS related protein 81; ERF, ETS repressor factor; ERG, v-ets avian erythroblastosis virus E26 oncogene related; ERM, ETS related molecule; ESE, Epithelial specific ETS; ETS, v-ets erythroblastosis virus E26 oncogene homolog; ETV, ETS variant gene; FLI1, Friend leukemia virus integration 1; FEV, Fifth Ewing variant; GABP, GA repeat binding protein; LIN, abnormal cell lineage; MEF, myeloid ELF1-like factor; NERF, New ETS-related factor; PEA3, polyomavirus enhancer activator-3; PDEF, prostate derived ETS transcription factor; PSE, prostate epithelium-specific ETS; SAP, Serum response factor accessory protein; SPDEF, SAM pointed domain containing ETS transcription factor; SPI, spleen focus forming virus proviral integration oncogene; TEL, translocation, Ets, leukemia; TCF, Ternary complex factor. (Gutierrez-Hartmann, TRENDS in Endocrinology and Metabolism; 18; 150-158; 2007)[64]

ETS binding sites (EBS) have been identified in the promoter regions of viral and cellular genes, and ETS factors are involved in the regulation of expression of genes critical for proper control of cellular proliferation, differentiation, development, haematopoiesis, apoptosis, metastasis, tissue remodelling, angiogenesis, metastasis and

transformation. More than 400 ETS target genes have been postulated based upon the presence of functional EBS in their regulatory regions, 200 of which have been identified.[62, 63, 67-72]

Some ETS family proteins are expressed ubiquitously and some in a tissue-specific manner. For example, *ERG* is initially expressed in embryonic endothelial tissues and later in the kidney, urogenital tracts and hematopoietic cells, while *ETS2* is expressed ubiquitously.[71]

ETS family proteins regulate gene expression by functional interaction and complex formation with other transcription factors and co-factors on their DNA binding sites. Many ETS family proteins are downstream nuclear targets of the signal transduction cascades. Post-translational modification of ETS family proteins, for example by phosphorylation, modulates DNA-binding activities, association with co-regulatory partners, transcriptional activation capacities, and subcellular localization.[62, 67, 71] Many ETS domain transcription factors are subject to autoregulation, during which their DNA binding activity is usually masked until an appropriate trigger and interactions with co-regulatory transcription factor(s) are in place. The ability of individual ETS factors to function as activators or repressors is also dependent upon promoter, co-factors and cell context.[67] Unique combinations of protein-protein interactions direct different ETS factors to regulate the expression of specific target genes. A subset of ETS factors have repressor activity (*e.g.* ERF, YAN, TEL, NET) and may directly compete with other ETS factors for binding to EBS. For example, the transcriptional activity of ETS2 is inhibited by protein-protein interaction with ERG.[73] It has also been shown that ERG interacts with ESET (Erg-associated protein with SET domain), a histone H3-specific methyltransferase, thus participating in transcriptional repression.[74] Unique promoter interactions with specific ETS factors have been demonstrated in the case of ETS2 (or ETS1) and ERG on the *collagenase* (*MMP1*) and s*tromelysin* (*MMP3*) promoters. ERG appears to act as an activator of the *collagenase* promoter, while it inhibits stimulation of the *stromelysin* promoter by ETS2, whereas ETS2 stimulates both. In addition, interaction with other proteins can block the ability of ETS factors to activate transcription.[67, 75] Overlap between specific protein-protein interactions may provide a mechanism to control the diverse functions

of ETS family. Such combinatorial control provides a mechanism to fine-tune the networks of cellular processes.

Cellular responses to environmental stimuli are controlled by a series of signalling cascades that transduce extracellular signals from ligand activated cell surface receptors to the nucleus. There is a dynamic interplay between signalling pathways that results in the complex pattern of cell-type specific responses required for proliferation, differentiation and survival. Many of the ETS family proteins are downstream nuclear targets of the Ras-MAP kinase signalling pathway. They also interact with and influence crosstalk between specific cellular partners, which influence other signalling pathways such as the Jak/Stat, Smad and Wnt signalling pathways.[71,76] ETS family members can act as both upstream and downstream effectors of signalling pathways. As downstream effectors their activities are directly controlled by specific phosphorylations, resulting in their ability to activate or repress specific target genes. As upstream effectors they are responsible for the expression of numerous growth factor receptors.[76]

Among the first characterised interactions between ETS factors and another transcription factor, were studies demonstrating cooperativity between ETS factors and the AP1 (FOS/JUN) transcriptional complex to activate cellular responses by increasing the transcriptional activities of promoters containing AP1-EBS binding sites, including *MMP1* (*matrix metalloprotease-1* / *collagenase*), *uPA* (*urokinase plasminogen activator*), *GM-CSF* (*granulocyte-macrophage colony stimulating factor*), *maspin* (*serpinB5*) and *TIMP-1* (*tissue inhibitor of metalloproteinase-1*). In contrast, MafB, and AP1 like protein, inhibits ETS1-mediated transactivation of the AP1-EBS sites.[77]

ETS transcription factors and cancer

Following the identification of *ERG* as being highly upregulated in a subset of prostate cancer patients,[78, 79] *ETS* fusions have become one of the most common genetic markers of prostate cancer.[72, 80] The first clinically relevant candidates to dominant oncogenes in prostate cancer are *ETS* fusion genes resulting from chromosomal rearrangement of the 5' untranslated region of the prostate-specific, androgen

responsive, *Transmembrane serine protease gene (TMPSS2)* to *ERG*, *ETV1 (ER81)*, *ETV4* (*PEA3*) or *ETV5*.[80, 81] *TMPRSS2:ETS* gene fusions might be the most common genetic abnormality identified so far in human malignancies, resulting in androgen mediated induction of the respective *ETS* factors, which are then thought to activate a repertoire of ETS-responsive genes, leading towards prostate cell transformation.[80]

Multiple genetic and epigenetic events may be required for cancer development. Oncogenes and tumour suppressor genes act as modulators of cell proliferation, while the balance of apoptotic and anti-apoptotic genes controls cell death. The hallmarks of cancer cells are: 1. independence from mitogenic/growth signals; 2. loss of sensitivity to "anti-growth" signals; 3. evasion of apoptosis; 4. induction of angiogenesis; 5. release from senescence; and 6. invasiveness and metastasis[82] (Fig. 9).

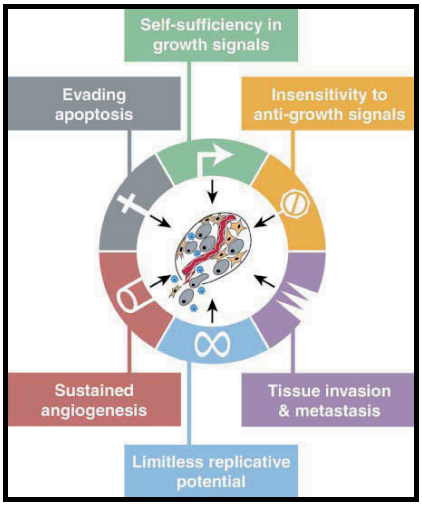

Figure 9. Acquired capabilities of cancer. (Hanahan and Weinberg., Cell; 100; 57-70, 2000)[82]

Oncogenic activation of cellular genes may occur through multiple mechanisms, including amplification and/or overexpression, activation by insertions of new regulatory sequences following retroviral integration, fusion with other proteins as a consequence of chromosomal translocations or through point mutations. *ETS* genes have altered expression patterns in both leukaemia and solid tumours, are chromosomally amplified or deleted, and are located at translocation breakpoints.[72, 80] As many ETS family transcription factors are downstream nuclear targets of Ras-MAP

kinase signalling, the deregulation of *ETS* genes may result in malignant transformation of cells.[63, 68, 71] Since some ETS family proteins affect the expression of several oncogenes and tumour suppressor genes by direct regulation of their promoters, activation and repression, respectively, or by protein-protein interactions, and it is evident that they play important roles in cell proliferation, apoptosis and differentiation in normal cells, deregulated expression of ETS family proteins could lead to disruption of these processes and contribute to development and progression of malignant tumours.[71] Several *ETS* family genes are expressed in the normal and/or cancerous prostate, including *ETS1*, *ETS2*, *ELF1*, *ESE2* (*ELF5*), *ER81* (*ETV1*), *ERG*, *PDEF* and *PEA3* (*ETV4*). Advanced stages of prostate cancer are associated with expression of *FLI1*, *ELF1*, *PDEF*, *ETS1* and *ETS2*. Transcriptional activation of *ETS* genes is essential for upregulation of extracellular matrix-degrading proteins including MMP1, MMP9, uPA, and the uPA receptor, many of which are associated with clinical features such as lymph node status and prognosis.[72]

The function of ETS family proteins has to be considered in combination with other cellular proteins, since the function of the same ETS protein sometimes differs in different types of tissues based on differences in cellular context.[71] For example, expression of *FLI1* is induced by ETS1 in endothelial cells but not in fibroblasts[83], and ETS1 is involved in angiogenesis, but overexpression of *ETS1* in umbilical vein endothelial cells induces apoptosis under serum deprived conditions.[84]

Individual ETS factors are overexpressed or downregulated in cancers. *ETS2* is overexpressed in prostate and breast cancer, and this overexpression is necessary for transformed properties of the cancer cells. *ETS1* expression is correlated with more malignant carcinomas and is a negative prognostic indicator.[71] Conversely, *PDEF* expression is lost in many epithelial cancers.[85] Among the multiple ETS target genes that are important for cancer progression are those that function in control of cell proliferation (cyclins and cdks), adhesion (cadherins, integrins, cell adhesion molecules (CAMs)), motility/migration (hepatocyte growth factor receptor c-Met, vimentin), cell survival (Bcl-2), invasion (uPA & uPAR, PAI, MMPs, TIMPs, heparanase), extravasation (MMPs, integrins), micro-metastasis (osteopontin, parathyroid hormone-related peptide (PTHrP), chemokine receptors (RANTES, MIP-

3α), CD44), and establishment and maintenance of distant site metastasis and angiogenesis (integrin β3, VEGF, Flt-1/KDR, Tie2).[67, 71] Many known stroma-modifying factors have known linkage to ETS factors. For example, ETS1 is a downstream effector of the epithelial mesenchymal transition (EMT) promoting hepatocyte growth factor (HGF), emanating from the stroma, while in tumour cells ETS1 and PEA3 can induce the expression of EMT markers such as vimentin and MMPs. ETS1 is also an activator of the HGF receptor *c-MET*, thus forming a positive feedback loop. ETS proteins can also mediate similar communication across different tumour and stroma compartments. *VEGF*, produced by tumour cells and fibroblasts, can induce *ETS1* expression in endothelial cells.[86] Concomitantly, ETS1, in cooperation with Hif-2α, activates the transcription of *VEGF receptor 2*.[87] Both *ETS1* and *FLI1* are downstream effectors of, and are differentially regulated, by *TGFβ*, and these two factors have divergent functions in both fibroblasts and endothelial cells.

Several ETS family proteins have been shown to be involved in the apoptotic process, and most members behave anti-apoptotically. For example, ETS2 and PU1 rescue apoptosis in macrophages upon deprivation of macrophage colony-stimulating factor (M-CSF), through upregulation of anti-apoptotic *Bcl-X_L* but not of apoptotic *Bcl-Xs*.[88] FLI1 and ERG inhibit apoptosis in NIH/3T3 cells induced by serum depletion or treated with a calcium ionophore.[89] Whether the ETS family proteins induce or prevent apoptotic cell death may depend on several factors such as expression levels, cellular contexts and the existence of agonistic or antagonistic signals in cells. ETS1 and ETS2 have been reported to be pro- as well as anti-apoptotic. For example, expression of the p42 splice variant of *ETS1* promotes Fas-mediated apoptosis by upregulating *caspase-1* in human colon cancer cells,[90] and overexpression of ETS1 in human umbilical vein endothelial cells induces apoptosis under serum-deprived conditions.[84] Overexpression of *ETS2* in prostate tumour cells increases apoptosis accompanied by increased levels of *$p21^{WAF1/Cip1}$*.[91] There are several reports showing that ETS family proteins directly induce expression of apoptosis related genes. Expression of the *Fas ligand* gene in vascular smooth muscle cells is controlled by cooperative activation between ETS1 and Sp1,[92] and the EBS of the 5'-flanking region of the *caspase-3* gene is necessary to achieve sustained

transcriptional activity of *caspase-3*.[93] FLI1 negatively regulates *Rb* expression by binding to an EBS in the promoter.[94] It has also been reported that FLI1 enhances the *bcl-2* promoter activity in leukaemic cells, thereby rescuing the cells from apoptosis.[95]

Several ETS transcription factors are preferentially expressed in certain lineages of hematopoietic cells and play crucial roles in their development and differentiation. Many are also aberrantly expressed, often due to chromosomal translocations, and play essential roles in the transformation and development of leukaemias. These includes *PU1*,[96, 97] *TEL,* which is often a target for chromosomal translocations,[98, 99] *TLS-ERG* in acute megakaryoblastic leukaemia (AMKL)[100, 101] and *ERG* and *ETS2* in myeloid leukaemia.[102] The *TEL* (*ETV6*) gene, for example, is juxtaposed to several tyrosine kinase genes in leukaemias, including the platelet-derived growth factor receptor β (*PDGFR β*) gene by [t(5;12)(q33;p13)] translocation in human chronic myelomonocytic leukaemias,[103] the *c-abl* gene by [t(9;12)(q34;p13)] in chronic myelogenous leukaemias (CML) and acute lymphoblastic leukaemias (ALL),[104] the *Jak2* gene by [t(9;12)(p24;p13)] in T-cell and B-cell ALL,[105] the *TrkC/NTRK3* (neutrophin-3 receptor) gene by [t(12;15)(p13q24)] in congenital fibrosarcomas,[106] and *ARG* (c-abl related gene)/*ABL2* by [t(1;12)(q25;p13)] in an acute myelogenous leukaemia (AML) line.[107] All of the above-mentioned fused proteins possess the N-terminal region including the pointed (PNT) domain for homo- and hetero-dimerization from TEL and the intact tyrosine kinase domains from the partner proteins. Self-association through the PNT domain of TEL and subsequent activation of kinase activity of the fusion protein likely contributes to transformation of the cells.[108]

The Ewing sarcoma (EWS) family of tumours share recurrent translocations that fuse the *EWS* gene from 22q12 to mainly *FLI1*, but also *ERG* (in approximately 10% of EWS)[109] *ETV1*, *E1AF* and *FEV*, all members of the ETS family of transcription factors. The N-terminal region of EWS, an RNA-binding protein, and the C-terminal region of FLI1, including the Ets domain, are fused forming *EWS:FLI1* [t(11;22)(q24;q12)] in 85% to 95% of the cases.[110] They possess increased transactivation potential in comparison with the wild-type *FLI1* gene and this activity is thought to contribute to malignant transformation of the cells. The *EWS-ETS* fusion

is causative in the development of Ewing's tumours, mainly due to the abnormal transcriptional regulation of key target genes which are involved in the regulation of cell cycle, signal transduction and migration.[111] EWS and related tumours are characterized by elevated level of *c-MYC* expression. It has been shown that EWS-FLI1 is a transactivator of the *c-MYC* promoter[112] and is often associated with poor prognosis.[113] The expression of *EWS:FLI1* also leads to a considerable downregulation of the $p57^{KIP2}$ tumour suppressor gene.[114] In some cases of Ewing's sarcoma, the *EWS* gene is fused with other *ETS* family genes including *ERG*, *ER81/ETV1, FEV* and *E1AF*.[71,115]

ERG

ERG is most often referred to as 'ets-related gene', but also as 'v-ets erythroblastosis virus E26 oncogene like', 'v-ets erythroblastosis virus E26 oncogene homolog' or 'v-ets erythroblastosis virus E26 oncogene related'. In 1987, Reddy *et al.*[116] isolated cDNA clones representing the complete coding sequence of an ets-related gene which they named *ERG1*, due to the fact that nucleotide sequence analysis of this 4.6 kb long cDNA, predicted a 363 amino acid protein, whose amino acid sequence showed a homology of approximately 40% and 70% to two domains corresponding to the 5' and 3' regions of *v-ets* oncogene, respectively. Rao *et al.*[117] identified ERG1 and another cDNA clone with alternative splicing, encoding a longer protein of 462 amino acids, named ERG2. They proposed that the various isoforms are formed by alternative sites of splicing and polyadenylation, together with alternative sites of translation initiation. The identification of other isoforms followed.[118, 119] ERG3 was characterized by a differential splicing which results in the insertion of 24 amino acids in the coding region of the ERG2 protein.[119] All ERG isoforms can bind the ETS site in a specific manner and act as transcriptional activators, although they demonstrate differential

interactions with the AP1 complex (transcription factor consisting of jun/fos family proteins).[120]

The *ERG* gene has been localized to chromosome 21q22.2[117, 121] which is part of the Down syndrome critical region (DSCR) of chromosome 21. The DSCR of chromosome 21 is abnormally triplicated in a subset of individuals with Down's syndrome. Owczarek *et al.*[121] determined that the *ERG* gene consists of at least 17 exons spanning approximately 300 kb of genomic sequence, generating at least 9 separate transcripts, of which the last 4 (*ERG6 – ERG9*) are likely of relatively low abundance. Only two of these transcripts encoded proteins that may have functions. *ERG1 – ERG5* encode five proteins of 38 to 55 kDa, all of which bind DNA at ETS sites and act as transcriptional activators. They differ in their 5' regions and the expression of two alternative exons, A81 (81 bp) and A72 (72 bp). Later we revised the exon maps of *ERG1* and *ERG2* (Paper I).[79]

During mammalian embryogenesis, *ERG* is first expressed in endothelium and later in the kidney, urogenital tract and hematopoietic cells, whereas down-regulation is observed following tissue differentiation.[122, 123] The isoforms of *ERG* may form homodimers with itself or heterodimers with other ETS proteins including FLI1, ETS2, Er81 and PU1.[120]

Isoforms *ERG3* (p55) and *ERG5* (p38) are the predominant forms expressed in endothelial cells.[124] By *in-situ* hybridization we identified expression of *ERG* in prostatic endothelial cells but not in benign epithelial cells.[79] *ERG* is involved in vascular development and angiogenesis as it regulates the expression of endothelial-specific genes including von *Willebrand factor*, *VE-cadherin*, *endoglin* and *intracellular adhesion molecule-2* (*ICAM2*).[125-127]

ERG is one of the ETS members involved in a number of chromosomal translocations in human leukaemias, including a [t(8;21)(q22;q22)] non-random translocation in patients with myelogenous leukaemia subtype M2 (AML-M2),[128] a [t(16;21)(p11;q22)] translocation in human myeloid leukaemia fusing the *ERG* gene with the *TLS/FUS* gene[129] and chromosomal rearrangement with the *EWS* gene in Ewing´s sarcoma.[130] As Petrovics *et al.*[78] and our group[79] have shown, *ERG* is highly upregulated in around 50 % of prostate cancer patients. In 2005 Tomlins *et al.*[80]

identified the mechanism for this as a chromosomal rearrangement fusing the promoter region of the highly expressed androgen responsive serine protease gene *TMPRSS2* (21q21.3) to the *ERG* (21q21.2) coding sequence (either through deletion or translocation). Although genetic rearrangements through translocations are very common in leukaemias, they had so far not been identified in epithelial adenocarcinomas until Tomlins *et al.* [80] demonstrated *ERG* gene fusions in prostate cancer. This has become one of the most common genetic markers of prostate cancer and the first clinically relevant candidate to a dominant oncogene in prostate cancer, together with *ER81 (ETV1)*, *PEA3* (*ETV4*) and *ETV5,*[81] which may alternatively be fused with *TMPRSS2* in a minority of *ETS* fusion positive cases. A number of alternative 5' and 3' fusion partners have since been identified (Fig. 10). Although there have been opposing conclusions regarding the implications of this fusion, there seems to be an association between positive fusion status and adverse prognosis (Table 1).

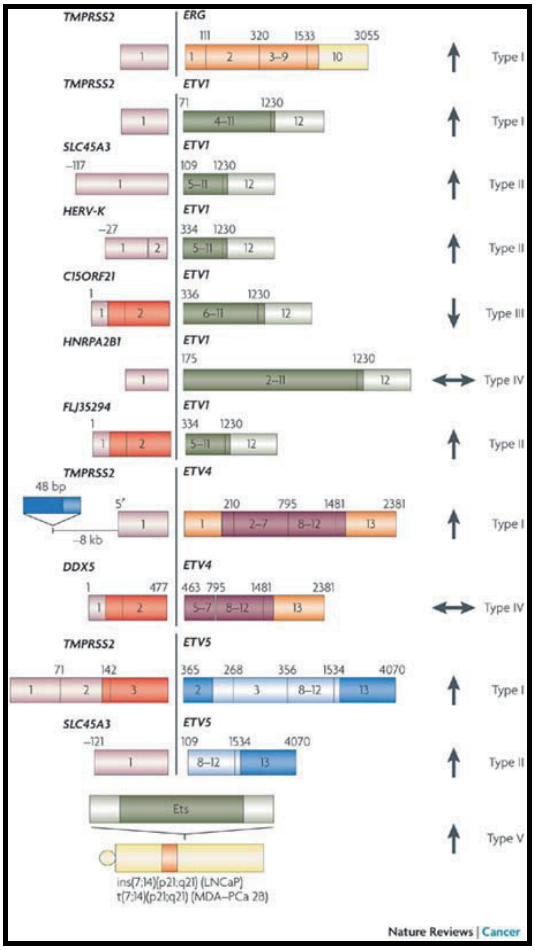

Figure 10. A sketch representation of the gene fusions characterized in prostate cancers so far. The 5' fusion partners are depicted on the left side and corresponding 3' partners on the right. Light colours at the ends of the genes depict untranslated exons. The dark-coloured boxes depict coding exons. The numbers on the boxes identify the base positions of the exons. The arrows represent androgen responsiveness of the fusion genes: arrows pointing up signify androgen-mediated upregulation; arrows pointing down represent androgen-mediated downregulation of the corresponding gene; the horizontal arrows represent absence of androgen action on the fusion genes' expression. *TMPRSS2*–ETS gene fusions have been grouped as type I; other gene fusions which are androgen-inducible have been grouped as type II, androgen-repressed fusion genes make up type III, androgen-insensitive fusion genes, type IV, and lastly, the novel situation in prostate cancer cell lines, with ETS genes rearranged to an androgen-sensitive location (without the generation of classical gene fusions), has been classified as type V. (Kumar-Sinha *et al.*, 2008. Recurrent gene fusions in prostate cancer. Vol 8 (7):497-511[131])

ETS gene status	Assay	Patient cohort	Prognostic association	Study
TMPRSS2:ERG rearrangement	Break-apart FISH	Prostate cancer, surgically treated, n = 96	High-pathologic stage.	132
TMPRSS2:ERG fusion	RT-PCR and DNA sequencing	Prostate cancer, surgically treated, n = 26, Gleason score 7	Higher rate of recurrence. Single most important prognostic factor.	133
TMPRSS2:ERG fusion	RT-PCR and DNA sequencing	Prostate cancer, surgically treated, n = 165	Higher risk of recurrence. Strong prognostic factor independent of grade, stage and PSA level.	134
ERG rearrangement	Break-apart FISH	Prostate cancer, cohort of conservatively managed patients (no hormone treatment), n = 445. TMAs of transurethral resection	Very poor cause-specific survival (25% at 8 years) (2+ Edel) compared with *ERG* rearrangement-negative cases (90% at 8 years).	135
TMPRSS2:ERG rearrangement	Multicoloured fusion FISH	Prostate cancer, population-based watchful waiting cohort, n = 111	Prostate-cancer specific death.	136
TMPRSS2:ERG rearrangement	Dual colour break-apart FISH	Prostate cancer, hormone-naive and hormone-refractory lymph node metastases, n = 136	Higher tumour stage, presence of metastatic disease involving pelvic lymph nodes.	137
TMPRSS2:ERG fusion	FISH Break-apart FISH	Prostate cancer, surgically treated, TMAs, n = 196	Moderate to poorly differentiated tumours.	138
ERG overexpression	Microarray, real-time PCR	Prostate cancer, laser capture microdissected epithelial cells *ERG*-overexpressing tumours	Longer recurrence-free survival, well and moderately differentiated stages, lower pathological stage, and negative surgical margins.	78
TMPRSS2:ERG fusion	RT-PCR and DNA sequencing	Prostate cancer, TRUS-guided needle biopsies, n = 50	Lower Gleason grade and better survival than fusion-negative tumours.	139
TMPRSS2:ERG fusion	RT-PCR and DNA sequencing	Prostate cancer, surgically treated, n = 54	No correlation with clinical outcome.	140
TMPRSS2:ERG fusion	RT-PCR and DNA sequencing FISH	Prostate cancer, surgically treated, n = 54 Hormone-naive pelvic lymph node metastases, n = 9	No association with tumour stage, Gleason grade or recurrence-free survival.	141

2+Edel, deletion of 5'*ERG* sequences, accompanied by duplication of *TMPRSS2:ERG* sequences; FISH, fluorescence *in situ* hybridization; PSA, prostate-specific antigen; RT-PCR, reverse-transcription PCR; TMA, tissue microarray; TRUS, transrectal-ultrasound.

Table 1. Prognostic associations of the TMPRSS2:ERG gene fusions. (Kumar-Sinha *et al.*, 2008. Recurrent gene fusions in prostate cancer. Vol 8 (7):497-511[131])

SIM2

The *SIM2* (Single-minded homolog 2) gene has also been identified within the Down's syndrome critical region (DSCR) on chromosome 21 (21q22.2), which is associated with trisomy 21.[142, 143] *SIM2* was originally identified in *Drosophila* where it plays an important role in development and has peak levels of expression during the period of neurogenesis. *Drosophila* single-minded acts as a positive master gene regulator in central nervous system midline formation. *SIM2* encoded proteins belong to a family of transcriptional repressors and may control brain developments and neuronal differentiation.[144-147] Chen *et al.*,[142] proposed that the human *SIM* gene is a candidate for involvement in certain dysmorphic features (particularly the facial and skull characteristics), abnormalities of brain development, and/or mental retardation of Down syndrome. Due to alternative splicing, the *SIM2* gene exists in two distinct isoforms, *SIM2-long* (*SIM2-l*) and *SIM2-short* (*SIM2-s*).[144]

SIM2 has been shown to be involved in the pathogenesis of solid tumours. Higher expression levels of *SIM2-s* have been seen in the carcinomas of colon, pancreas and prostate in comparison to the normal tissues, but not in breast, lung or ovarian carcinomas or in most normal tissues (it is expressed in the kidneys and tonsils). Elevated expression has been seen in early colon adenomas and BPH as well, raising the possibility that the SIM2-s activation may be an early event. *SIM2-s* specific immunoreactivity was detected in the majority of tumours of different Gleason scores and in prostatic intraepithelial neoplasia (PIN), but not in most stromal hyperplasia.[148] In our own gene expression profiling study, *SIM2* ranked second among highly upregulated genes in prostate cancer.[149] We also identified both *SIM2-s* and, for the first time, *SIM2-l*, as being upregulated in prostate tumour tissue compared with paired benign tissue samples.[150]

A proposed cancer-related role of the *SIM* family of genes is their ability to transcriptionally regulate key metabolic enzymes to inactivate carcinogens.[151] *SIM2* belong to a family of transcription factors containg PAS (Per/Arnt/Sim) heterodimerization domains.[148] The PAS domains are also cytosolic sensors that detect xenobiotics, redox changes, and light, oxygen and energy levels in prokaryotes and

eukaryotes.[152] *SIM2*, if dysregulated (due to mutations, amplifications or loss of repression), could suppress xenobiotic-stimulated induction of Phase II enzymes by inhibiting the dimerization of aryl hydrocarbon receptor (Ahr) and Ahr nuclear translocator (ARNT) at one of their PAS domains.[153] The resultant absence of the Ahr-ARNT-mediated protective and homeostatic pathway would render cells vulnerable to mutagenesis and other forms of oxidative damage, and would provide an environment for tumourigenesis. The Ahr-ARNT heterodimer also mediates xenobiotic-induced apoptosis in foetal ovarian cells, by binding to the xenobiotic response element in the promoter region of the pro-apoptotic *bcl-2* family member, *bax*.[154] Therefore, suppression of Ahr-ARNT activation by SIM2 might disable apoptotic checkpoints that are essential for cancer surveillance. The precise function and nature of genes regulated by SIM2 are not completely clear.

Several groups have studied the expression of *SIM2* in various cancers or cancer cell lines. DeYoung *et al.*[155] made a systematic study of the expression differences among *SIM* family members in pancreatic cancer. In APAN-1, a pancreatic cancer-derived cell line, antisense inhibition of *SIM2-s* expression caused a dose-dependent inhibition of *SIM2-s* mRNA. The targeted protein SIM2-s was also inhibited in the antisense-treated cells accompanied by growth inhibition and induction of apoptosis, providing a rationale for preclinical testing of the SIM2-s antisense drug in pancreatic cancer models. They identified both *SIM2-s* and *SIM2-l* isoform as expressed in lung, kidney, skeletal muscle, testis and tonsils. Low-level expression of *SIM2-l* was seen in the bone marrow as well. Real-time qPCR analysis of pancreatic tissues and cell lines showed expression of both *SIM2* isoforms in tumours and tumour-derived cell lines. DeYoung *et al.*[156] also found that antisense inhibition of SIM2-s in a RKO-derived colon carcinoma cell line caused growth inhibition, apoptosis and inhibition of tumour growth in a nude mouse tumorigenicity model. On the other hand, Kwak *et al.*[157] observed that *SIM2-s* expression was lost in human breast cancers, and Laffin *et al.*[158] found that loss of *SIM2-s* promotes epithelial mesenchymal transition (EMT) and tumourigenesis in breast cancer cells. Loss of *SIM2-s* caused aberrant mouse mammary gland ductal development with features associated with malignant transformation, and knockdown of *SIM2-s* in MCF-7 breast cancer cells contributed to

an EMT and increased tumourigensis. These changes were associated with increased *SLUG* (*SNAI2*) and matrix metalloprotease 2 (*MMP2*) levels. They suggested that *SIM2-s* is a key regulator of mammary-ductal development and that loss of expression is associated with an invasive, EMT-like phenotype. These results suggest that *SIM2-s* plays a key role in controlling normal EMT processes involved in mammary gland development and that loss of *SIM2-s* promotes pathological EMT associated with tumour progression. These tumour suppressor properties of *SIM2-s* in breast cancer is contradictory to its cancer promoting role in colon, pancreas and prostate cancers, and may reflect different tissue specific functions or differences in effect depending upon the cellular context. Increased expression of SIM2-s has also been identified in glioma and glioblastom cell lines,[159] in which they were suggested to play a role in invasion, which may partly be associated with increased expression of *TIMP2* and decreased expression of *MMP2*.

DeYoung *et al.*[156] showed that antisense inhibition of *SIM2-s* expression in a colon cancer cell line caused inhibition of gene expression, growth inhibition and apoptosis. Administration of the antisense, but not the control oligonucleotides, caused significant inhibition of tumour growth in nude mice with no major toxicity, establishing *SIM2-s* as a molecular target for cancer therapeutics.

DIAGNOSTIC AND PROGNOSTIC BIOMARKERS OF PROSTATE CANCER

Prostate cancer is a heterogeneous and multifocal disease and biomarkers are strongly needed to enable more accurate detection, improved prediction of tumour grade, and stage, as well as facilitated discovery of new therapeutic targets for improved treatment.

Currently, an important diagnostic and prognostic marker of prostate cancer is prostate specific antigen (PSA). Based upon initial concentration of total PSA in serum, prostate cancer is diagnosed by histological examination of prostate tissue obtained by ultrasound guided transrectal needle biopsy. This method has suboptimal sensitivity and specificity, leading to many unnecessary initial and repeat biopsies.

Biomarkers may be detected in prostatic cancerous tissue and in body fluids (blood, serum, urine). Prostate tissue sampling requires an invasive procedure (transrectal ultrasound-guided biopsy) and the chances of sampling error represent a problem. It has been known since 1869 that cancer cells break away from the primary tumour and are present in body fluids.[160] Serum and urine contain degradation products of extracellular matrix and of benign and malignant cells and their secreted products. Even in early cancer development, these cells are shed and may be detected.[161, 162] For prostate cancer both blood (serum) and urine are viewed as attractive samples for diagnostic assays, due to the less invasive procedure compared with tissue sampling (Fig. 11).

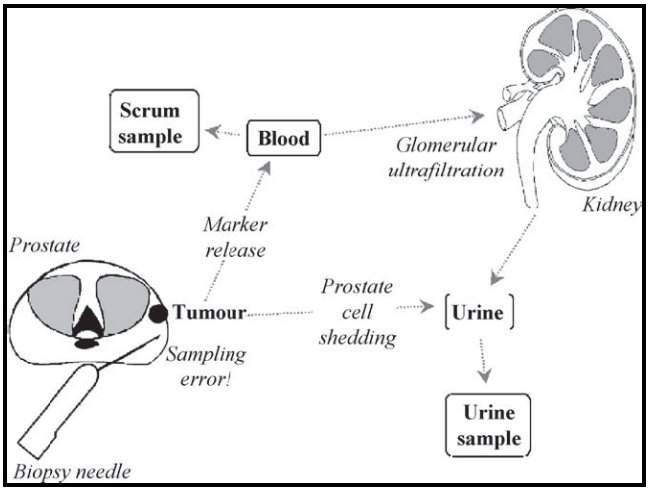

Figure 11. Blood / serum and urine prostate cancer markers have certain advantages over tissue prostate cancer markers. They may easily be obtained while prostate tissue sampling requires and invasive procedure (transrectal ultrasound-guided biopsy) (van Gils *et al.*, Eur Urol; 48(6):1031-41, 2005).[162]

Early detection of prostate cancer has proved difficult and current detection methods are inadequate. At present, one of the major challenges in prostate cancer treatment is to distinguish between patients with aggressive and clinically significant tumours who need more intense treatment, and patients with indolent tumours, who will benefit from active surveillance. Novel biomarkers are strongly needed to enable more accurate detection of prostate cancer, improved prediction of tumour

aggressiveness and facilitated discovery of new therapeutic targets. Prostate cancer specific molecules have the potential to serve as diagnostic and prognostic indicators and therapeutic targets. The challenge lies in finding potential molecular biomarkers only present in prostatic cancerous tissue and not in benign tissue, which might be detected by noninvasive techniques in blood/serum or urine. The heterogenous and multifocal nature of prostate cancer must be taken into consideration. This is a challenge most likely solved with a combinatorial test in which detection of combinations of biomarkers confer higher specificity and sensitivity than todays' PSA testing. Ideally, biomarkers of prostate cancer aggressiveness should be available at the time of diagnosis to allow optimal treatment planning.

In addition to diagnostic markers, prognostic, predictive, and therapeutic markers are needed to predict disease severity, choosing treatments, and monitoring responses to therapies, respectively. Guidelines for biomarker development have been established to aid in the validation of candidates.[163, 164] There are several existing and potentially interesting novel prostate cancer biomarkers which confer increased diagnostic and prognostic information as well as improved sensitivity and specificity compared with PSA alone.

PSA and PSA-derived forms

Prostate specific antigen (PSA) was identified by Ablin *et al.* in 1970.[165, 166] It is a seminal proteinase produced by normal and malignant prostate epithelial cells. PSA was originally used for monitoring prostate cancer patients and was subsequently implemented for screening purposes. Serum PSA testing has been used for over 20 years as an aid in the diagnosis and management of prostate cancer and PSA is the most successful and widely employed cancer serum marker in use today. The measurement of total PSA has been shown to be useful as a prognostic tool, with high preoperative values being associated with advanced disease and a poor clinical outcome. PSA is a very sensitive marker, which enables us to diagnose prostate cancer before it manifests itself symptomatically or clinically. It is unclear, however, whether PSA screening has led to a decline in mortality due to prostate cancer.

The tissue specificity of PSA is responsible for its utility as a serum marker.[167] PSA is produced almost exclusively in the prostate, but an increase in serum PSA levels is not necessarily associated with cancer, it is not cancer specific. Although highly sensitive, it suffers from a lack of specificity, showing elevated serum levels in a variety of pathological conditions in the prostate including prostatitis, benign prostate hyperplasia (BPH), and non-cancerous neoplasia. Even though prostate cancer cells make less PSA than normal cells, PSA leakage around disrupted gap junctions of cancer cells causes elevated protein in the circulation.[168-170] Many patients undergo unnecessary biopsies or treatment for benign or latent tumours. More than half of the men with a PSA over 4.0 ng/ml, which is the accepted clinical decision limit, are negative on initial biopsy.[162, 171] On the other hand, there is strong evidence that a cut-point of 4 ng/ml misses a significant number of cancers. In a prospective cohort study, designed to evaluate the preventive effect of the drug Finasteride, 15% of men enrolled in the untreated control arm of this trial, and who had an initial PSA 4 ng/ml harboured prostate cancer, with 14% of them showing high grade disease.[172] The "PSA dilemma" population of men (those with elevated PSA who are negative on initial biopsy) are frequently biopsied multiple times as they age to assess the possible development of clinically significant cancers. For those men who are diagnosed and undergo curative surgical treatment, about 20-30% will clinically relapse, revealing that for many men cancer was not detected at an early enough stage. Nevertheless, 15-40% of the patients who undergo intended curative treatment for clinically localized PC will experience biochemical recurrence (*i.e.*, a rise in serum PSA) within 5 years.[173]

Systematic PSA screening has resulted in marked overdiagnosis and overtreatment of clinically insignificant tumours.[174, 175] As an effect of PSA screening, the lifetime risk of prostate cancer diagnosis has increased to 16%, whereas the lifetime risk of dying from the disease is only 3.4 %. Further, during the last decade, a significant shift at radical prostatectomy has been observed, also called "stage migration", which is related to the widespread use of PSA for screening. Tumours detected by PSA alone are characterized by small size, low grade, and they express low levels of PSA. There is, however, a very strong evidence of a highly significant

association between long-term cancer risk and PSA-levels in the blood measured at early middle age in representative populations of healthy men.[169],[176]

Measurement of total PSA has been shown to be useful as a prognostic tool, with high preoperative values being associated with advanced disease and a poor clinical outcome. It is unclear whether PSA screening has actually led to a decline in mortality due to prostate cancer. The relationship of PSA to tumour grade is also not clear. The tissue PSA concentration has been shown to decrease with increasing Gleason score,[177] although concentrations in the serum increase because of disruption of the basement membrane surrounding the prostate epithelial cells and in the overall prostate tissue architecture. Currently used routine prognostic tools (*i.e.*, the Partin staging tables[178] and the postoperative nomograms developed by Kattan *et al.*[179] and Stephenson *et al.*[179, 180]) rely solely on pathological and clinical parameters, including serum PSA, Gleason score and tumour stage. These tools have limited utility for many patients who are mid-range, *i.e.* have serum PSA values in the range of 4-10 ng/ml. The inadequacies of PSA as a marker have created a need for novel markers of prostate cancer to prevent overtreatment of indolent tumours.

PSA alternatives. PSA circulates in a number of distinct forms, and several variations have been studied as an alternative to the original total PSA test (for example evaluation of velocity, density, levels of free *vs.* bound proisoforms).[168, 181-190] PSA processing is different in benign tissue and cancer tissue and measurement of these in addition to total PSA may significantly increase the diagnostic utility.[162]

Potential biomarkers

A large number of potentially clinically useful biomarkers in prostate cancer have been investigated, some of which have been studied by our group and collaborators, and shown to be associated with adverse pathological parameters and of prognostic value in prostate cancer. These include loss of PTEN/p27 expression,[191] increased expression of the p16 protein,[192] strong EZH2 expression,[193] high vascular proliferation[194] as well as an association between the epithelial to mesenchymal transition (EMT), characterized by reduced E-cadherin and increased N-cadherin expression, and

prostate cancer progression.[195] Table 2 provides a more comprehensive selection of a number of prostate cancer biomarkers which have been investigated, with various degrees of success. As previously mentioned, combinations of various biomarkers (multiplexed tests) are most likely to provide the necessary information needed, some of which studies are outlined in Table 3. The identification of the pathognomonic fusion between *TMPRSS2:ERG* and our ability to identify this gene product (mRNA detection) in tissue and urine of prostate cancer patients provide new hope regarding both more exact discrimination between grades of cancer and development of new therapeutic targets. This will be more closely covered under Discussion of results (Paper III).

Recently identified potential biomarkers are Sarcosine and Annexin A3. Sarcosine (*N*-methyl derivative of the amino-acid glycine) has recently been identified as a differentially expressed metabolite that is greatly elevated during prostate cancer progression to metastasis and it can be detected noninvasively in urine.[196] Sreekumar *et al.*[196] linked activation of the sarcosine pathway to AR and ETS gene fusion regulation. Both ERG- and ETV1-induced invasion were associated with a threefold sarcosine increase in benign RWPE cells. Knockdown of the *ERG* gene fusion in VCaP cells resulted in a more than threefold decrease in sarcosine with a similar decrease in the invasive phenotype. Androgen receptor and the *ERG* gene fusion product co-ordinately regulate components of the sarcosine pathway, and sarcosine is a potentially important metabolic intermediary of cancer cell invasion and aggressivity, making it a possible promising target for therapeutic interventions. *Annexin A3* (*ANXA3*) is negatively associated with prostate cancer.[197, 198] ANXA3 protein expression is reduced in cancer providing a negative staining rate, which correlated with increasing pT stage and Gleason score. ANXA3 status was shown to be an independent adverse prognostic factor and ANXA3 may be detected in urine samples with improved specificity compared with PSA.

TMPRSS2:ERG gene fusions may also be detected in circulating prostate cancer cells. Mao *et al.*[199] was unable to detect *TMPRSS2:ERG* transcripts by real-time qPCR in enriched cancer cells from peripheral blood from 15 patients with advanced androgen independent prostate cancer. However, they analyzed isolated circulating

cancer cells from 10 of these patients with FISH, and found *TMPRSS2:ERG* fusions in six of these cases. This suggests that cancer cells with the gene fusion may migrate into the blood vessel for seeding at distant sites. Analysis of circulating tumour cells may be used to monitor tumour progression and response to therapies,[200] but further investigation is required to evaluate the application of the gene fusion in monitoring early stage disease.[199]

Table 2. Overview over a number of potential prostate cancer biomarkers

BIOMARKER	SAMPLE	MARKER MEASURED	CLINICAL RELEVANCE	REFE-RENCE
Human kallikrein 2 (hK2)	Serum	Protein	Diagnostic and potential prognostic/predictive marker. Increased levels associated with more adverse grade, stage and volume. Combination studies with totPSA and freePSA.	169, 201-203
Prostate specific membrane antigen (PSMA)	Serum, tissue, seminal fluid, urine	mRNA, protein,	Diagnostic and prognostic marker. Increased expression in prostate cancer epithelial cells. Associated with more aggressive tumours. Elevated expression after androgen deprivation therapy or in hormone refractory tumours. Function unkown. Clinical usefulness uncertain. ProstaScint (radiolabelled anti-PSMA antibody) detects cancer tissue after biochemical recurrence and may identify metastasis. Urine combination test with PSA of diagnostic value.	204-206
Chromogranin A (CgA)	Serum	Protein	Diagnostic and prognostic marker. Increased serum levels correlate with adverse prognosis in metastatic hormone therapy resistant cancer.	207-209
Neuron-specific Enolase (NSE)	Serum	Protein	Neuroendocrine marker; androgen independent Predictive marker. Increased levels predict poorer survival in metastatic prostate cancer patients treated with endorcrine therapy.	210
Glutathione S-transferase-π (GSTP1)	Urine, seminal fluid	DNA	Epigenetic marker. Tumour suppressor gene. Diagnostic value and screening. Reduced expression in prostate cancer due to hypermethylation of the promoter in approximately 90% of prostate cancers.	211-215
Enhancer of zeste homolog 2 (EZH2)	Tissue	mRNA, protein,	Epigenetic marker; polycomb family of proteins. Prognostic and predictive value. Increased expression associated with progression and poor prognosis. Increased expression in metastatic versus localized cancer.	193, 216, 217
Micro-RNAs (miRNAs)	Tissue, serum	Non-coding RNA	Diagnostic and prognostic markers. Several miRNAs are either upregulated (*i.e. miR-25*, *miR-141*) or downregulated (*i.e.* miR-125b, *miR-145*, *miR-221*) in prostate cancer compared with benign tissue. Some may also provide prognostic information, *i.e. miR-221*.	218-220
Loss of heterozygosity (LOH)	Blood, urine, tissue	DNA	Potential diagnostic marker. The presence of LOH of prostatic cells was associated with prostate cancer. May help identify patients who are candidates for further prostate biopsies.	221, 222

Prostate Cancer gene 3 (PCA3 / DD3)	Urine, serum, seminal fluid, tissue	mRNA, protein	Diagnostic and prognostic value. Overexpressed in 95% of prostate cancers. Associated with organ-confinement, volume and aggressiveness. Urine tests /PROGENSA PCA3 assay (PCA3 mRNA normalized to PSA mRNA)	223-230
Hepsin	Tissue, urine	mRNA	Possible diagnostic value. Overexpression in ca 90% of prostate cancers.	149, 231-233
A-Methylacyl coenzyme A racemase (AMACR)	Urine, serum, tissue, prostatic secretion	mRNA, protein	Diagnostic and prognostic value. Upregulated expression in primary prostate cancer and decreased levels of AMACR in metastatic tumours. May predict prostate cancer biochemical recurrence and death. Measure of circulating autoantibodies to AMACR in serum may diagnose prostate cancer.	149, 234-237
Urokinase-type plasminogen activator receptor (uPAR)	Serum, tissue	Protein	Predictors of prognosis and progression. Increased levels in prostate cancer associated with aggressiveness, progression and metastasis.	238, 239
Early prostate cancer antigen (EPCA)	Tissue, serum	mRNA, protein	Possible diagnostic and prognostic value. Increased expression in cancer. Nuclear matrix protein. Clinical usefulness uncertain.	240, 241
TMPRSS2:ETS members	Tissue, serum, urine	mRNA, protein,	Diagnostic and prognostic value. Prostate cancer specific gene fusions	80, 136, 242
Transforming growth factor β_1 (TGF β_1)	Plasma, tissue	Protein	Predictive value. Increased concentrations in cancer tissue correlates with tumour grade, stage and biochemical recurrence and metastasis.	243, 244
E-cadherin	Tissue	mRNA, protein	Prognostic value. Levels of E-cadherin correlates with cancer grade and tumour stage and reduced levels associates with poorer prognosis and shorter survival time.	245, 246
Sarcosine	Urine	Protein	Prognostic value. Levels greatly elevated during cancer progression to metastasis.	196
Annexin A3 (ANXA3)	Tissue, urine	Protein	Diagnostic and adverse prognostic marker. Inverse relationship to cancer	197, 198
Prostate stem cell antigen (PSCA)	Tissue, serum	mRNA, protein,	Diagnostic and prognostic marker. Increased expression in cancer and the levels correlates with Gleason score, stage, progression and metastasis.	247-250

Table 3. Combination tests of potential biomarkers in prostate cancer

BIOLOGICAL SAMPLE	BIOMARKERS	MARKER MEASURED	CLINICAL RELEVANCE	REFERENCE
Urine	PCA3, GOLPH, SPINK1, TMPRSS2:ERG	mRNA	Detection and prediction of prostate cancer	[251]
Seminal fluid	GSTP1, hTERT	DNA, mRNA	GSTP1 methylation and hTERT expression may help predict negative biopsies for men with elevated PSA levels.	[211]
Tissue	E-cadherine, EZH2	protein	Increased EZH2:ECAD status associated with recurrence after radical prostatectomy.	[216]
Urine	PCA3, TMPRSS2:ERG	mRNA	Detection and prediction of prostate cancer	[252]

BACKGROUND AND AIMS OF THE STUDY

Background and general aim

The general aim of our study was to understand critical gene expression changes and regulatory patterns associated with prostate cancer, based upon new technological achievements and the possibilities for genome-wide analysis of gene expression. The hypothesis was that this discovery driven approach, in addition to increasing the understanding of prostate carcinogenesis, might result in novel diagnostic and prognostic markers. One long term goal of our research is to provide sufficient molecular information for individualized and tailored treatment options.

Specific aims

Paper I

Based on initial gene expression analysis of prostate cancer by Halvorsen *et al.,*[149] the aim of this study was to explore the expression profiles of prostate cancer with special focus on transcription factors. Differentially expressed genes in matched pairs of benign and malignant prostate tissue were identified and validated, with special focus on the ETS family of transcription factors, out of which *ERG* was the most consistently and highly upregulated member.

Paper II

In a previous study of gene expression profiles in prostate cancer,[149] the transcription factor *SIM2* was identified as being highly overexpressed in prostate cancer, and has also been proposed as a molecular target for cancer therapy. The aim of this study was to examine the expression status of *SIM2* at the transcriptional and protein level as related to patient outcome in prostate cancer.

Paper III

The mechanism behind the overexpression of the transcription factor *ERG* (Paper I) was found by another group to be due to a recurrent gene fusion between the promoter region of the constitutively expressed gene *TMPRSS2* and the coding region of *ERG*.

This generates a fused gene and transcript characteristic for prostate cancer. The aim of paper III was to determine the presence or absence of the nucleic acid fusions of *TMPRSS2:ERG* in urine samples from prostate cancer patients who underwent radical prostatectomy. Aspects important for optimal detection were examined. Possible correlations between fusion status and clinicopathological variables were also investigated.

MATERIALS AND METHODOLOGICAL CONSIDERATIONS

PATIENTS SERIES AND TISSUES

The patient series used in our studies include both archival paraffin embedded prostatic tumour tissue material, fresh frozen prostatic tumour and benign tissues as well as urine samples from patients prior to radical prostatectomy (Table 4).

Table 4. Patient material

Time period	Sample	No. cases	No. samples	Histology	Material	Method	Paper
1988 - 1994	RP tissue	104	104	Carcinoma	Paraffin	ISH	II
1997 - 2003	RP tissue	33	29[1]	Carcinoma	Fresh frozen	Microarray	I, II
			23	Benign			
1997 - 2003	RP tissue	49[2]	37[3]	Carcinoma	Fresh frozen	qPCR ISH	I, II
			39	Benign			
2006	Urine Hamburg	42	42	Carcinoma	Fresh frozen	qPCR	III
2006 - 2007	Urine Bergen	13	13	Carcinoma	Fresh frozen	qPCR	III

[1]19 tumour/benign pairs. [2]Includes 33 cases from 1997-2003. [3]27 tumour/benign pairs.

In 1984 radical prostatectomy was established at the Haukeland University Hospital in Bergen. Patients with localized prostate cancer were offered this treatment if they had a clinical stage T1/T2 disease, negative bone scan, general good health and 10 to 15 years life expectancy. The earliest study population of 104 patients (median age 62 years) were treated between 1988 and 1994*, i.e.* before the PSA era. PSA testing was introduced in Norway in the middle of the 1990s. The majority of the cancers in this series were clinical stage T2 (89%) and PSA detected tumours more typically seen today are clinical stage T1c.[253] These tumours were also larger (median diameter 28 mm) with more advanced pathologic stages than usually seen today.[254] Fresh frozen prostate tissue samples were collected during 1997 – 2003 from an independent series

of 49 radical prostatectomies. The majority of these cancers were clinical stage T1c and median Gleason grade 6.2, in line with other findings after the introduction of PSA screening.[253]

Urine sampling - Hamburg cohort: Through collaboration with the University Medical Centre Hamburg-Eppendorf in Hamburg, Germany, we received urine pellets and corresponding urine supernatants from 42 prostate cancer patients. The urine pellets were resuspended in a small volume of urine and TRK-β-mercaptoethanol was added before the pellets and supernatants were shipped to Norway on dry ice and frozen at -80 ^{o}C. These patients did not receive prostatic massage prior to sampling. **Bergen cohort:** In 2006 we started to collect urine samples after prostatic massage from all prostate cancer patients (who gave their consent) prior to treatment with radical prostatectomy at Haukeland University Hospital, Bergen. Each urine sample was divided into whole urine sample and pelleted urine sample with corresponding supernatant. Total RNA was isolated from both cohorts as well as genomic DNA for further studies according to protocols (Paper III). Table 5 summarizes clinico- and histopathological variables in the two cohorts. Studied together, 15 patients had Gleason scores of < 7, while 35 patients had a Gleason score of 7 and only 2 patients had a Gleason score > 7.

Table 5. Clinico- and histopathological variables of the Hamburg and Bergen urine sample cohorts.

	Hamburg cohort	Bergen cohort
No. of patients	42	13
Mean (median) age at diagnosis	62.8 (64)	62.9 (63)
Pre-treatment sPSA	7.5 (6.4)	7.4 (7.1)
After treatment Gleason score		
≤ 6	13	2
≥ 7	26	11
Pathological stage (TNM-staging)		
T2 (pT2)	29	8
T3 (pT3)	10	4

Archival tumour material

Archival radical prostatectomy specimens were retrieved from the files of Department of Pathology, Haukeland University Hospital. From 1988, formalin fixed prostatectomy specimens were totally embedded and studied by whole mount step sections at 5 mm intervals by one pathologist (Ole Johan Halvorsen), and a representative area of 1-2 cm^2 of highest tumour grade was selected retrospectively from the paraffin blocks of each specimen, and reembedded for further studies.

Fresh tissue sampling

From 1997 radical prostatectomy specimens were brought to the pathology department for fresh tissue sampling immediately following surgical removal. Guided by needle biopsy findings and palpation, the prostate was incised vertically and small 2-4 mm tissue specimens were dissected from macroscopic tumour and benign areas, snap frozen in liquid nitrogen and stored at -80 ^{o}C. Benign tissues were harvested from the contralateral zone. In support of this sampling strategy, gene expression patterns in benign tissues adjacent to tumour have been shown to be so substantially altered that it resembles a cancer field effect.[255] Histopathologic confirmation of benign or tumour tissues and evaluation of tumour content were performed on HE (hematoxylin-eosin) slides on opposing sides of the dissected area. Tissue samples from presumed benign areas were excluded if malignant or dysplastic glands (PIN) bordered on or were seen in the vicinity of the dissected area. Tumour cases were selected for RNA extraction if more than 50% carcinomatous tissue were present (mean 76%).

Clinicopathologic variables

Several clinical variables were recorded for subsequent analysis: patient age at radical prostatectomy, date of primary diagnosis, date of prostatectomy, clinical stage (TNM-classification as described in the Introduction), serum-PSA before and after surgical treatment, and complete follow-up information including date and time of eventual biochemical failure, as well as date and site of clinical recurrences, and survival.

Histopathologic variables

The following variables were studied retrospectively: largest tumour dimension in the paraffin embedded specimen (measured as height or maximum dimension on the histological sections), capsular penetration, seminal vesicle invasion, involvement of surgical margins,[59] presence of lymph node metastasis and pathologic stage. In line with recent recommendations,[55] a modified Gleason grading was applied. Mean Gleason score for the 29 carcinomas selected for fresh frozen tissue sampling: 6.2 (median 6, range 5-8). For validation purposes (qPCR), this series was expanded to a total of 37 malignant tumours and 39 benign samples, including 27 tumour/benign pairs. Although a high Gleason score is a significant determinant of prostate cancer death, there is an urgent need for additional biomarkers to increase the predictive value.[256]

MOLECULAR METHODS

Microarray and bioinformatics

Global gene expression analyses performed by DNA microarrays have revolutionized the study of global analysis of gene expression in cancer patient samples and experimental models. The technology takes advantage of the fact that a fragment of each gene, the probe, can be positioned in a dot matrix on one single glass slide or chip. Hybrids formed between the probes and the solubilised targets of enzymatically modified and fluorescently labelled mRNA from experiments or patient samples can be recorded by means of high precision microarray fluorescence laser scanners and assisting computer software. In principle, the DNA microarray technology can quantify the entire gene expression pattern in a sample at a given moment based upon one single hybridization experiment. This powerful technology might resolve the underlying gene expression and regulatory patterns of normal cell differentiation and the aberrant patterns of disease. Even though this method has proven to be highly successful and useful, the global gene expression strategy still has its limitations and pitfalls, and may at times add more noise than elucidation to the discovery process.[257]

It is important to keep this in mind when analysing and interpreting the gene expression data.

The microarray technology has been through a major development over the last decade. There are a number of commercial platforms available, for example Affymetrix, Illumina, Nimblegen and Agilent. The latter is the platform used by our group. The first microarrays developed were cDNA microarrays, and initially our group utilized a sequence-validated human cDNA library of 40 000 clones (40k) obtained from Research Genetics, originating from the IMAGE consortium.[149] There are several challenges and potential sources of noise with cDNA-arrays, including wrongly annotated sequences and errors during maintenance of the cDNA library. Maintaining and handling a library of 40 000 cDNA clones with control over which specific sequences are present is difficult, and wrong annotations of the sequences are also a problem.[257] These challenges make it difficult to compare results between different platforms (cross-platform comparisons)[258] and it is estimated that up to 10-30% of cDNA probes were wrongly annotated. A method for updating the annotation of probe sets within a platform, based on sequence alignments and specific probe selection, was proposed by Dai *et al.*[259] The platforms using synthetic oligonucleotide probes provide generally data of high-quality, with superior reproducibility compared with custom-spotted cDNA arrays,[170] and by now oligonucleotide arrays have largely replaced the more unreliable cDNA arrays for the genome-wide study of human gene expression. Agilent now applies 60-mer microarrays, consisting of probes containing 60 oligonucleotides for each gene. Around 60 nucleotide probe length provides optimal sensitivity and specificity, providing a balance between stability, sensitivity, specificity and possible problems with crosshybridization if the probe is too long, even though longer probes could be more sensitive for individual target genes.[260] Microarrays containing 44 000 (44k) reference sequences are currently available, covering all known genes and making it possible to examine the global mRNA gene expression in one sample. Therefore, due to the technological development, our first patient series was analysed using a cDNA microarray and later validated with oligonucleotide microarrays in an extended patient series. Another development is the present use of one-colour microarrays (only Cy3) instead of two-colour microarrays

(Cy3 and Cy5) as in our studies. An advantage of the two-channel system is that the reference sequence may cancel and correct for non-biological variations between samples. A disadvantage is that the ratios are dependent upon the expression levels in the reference sample. The one channel system generates absolute gene expressions which can be compared side by side. Cy3 is also a more stable fluorochrome than Cy5 and makes the target easier to handle and less prone to ozone mediated damage and fading.

Microarray. The Research Genetics human 40k cDNA microarray was printed at the Institute for Systems Biology, Seattle, and described by Halvorsen *et al.*[149] The slides were scanned in an Axon Genepix Scanner according to protocol and GenePix Pro 4.0 was used for quality control and spotfinding. For validation purposes the patient series was expanded and the 21k Agilent human 1A oligonucleotide microarray was used. The oligonucleotide microarrays were scanned and features automatically extracted, recorded and analysed using the Agilent Microarray Scanner Bundle.

Data analysis. Normalization, flooring or filtering of data was performed as described[149] and the data was then formatted in a J-Express file suitable for additional data mining (normalization, statistical analysis, gene search and visualization of gene expression using clustering and other tools) (http://www.molmine.com/)[261] (Fig. 12). The bioinformatics analysis of gene expression data remains a major challenge, and the software program J-Express was developed by the bioinformatics group at the University of Bergen, for analysis and visualisation of microarray data. To obtain consistency and reduce non-significant data, the genes were filtered before inclusion into the dataset (filtered dataset). Genes were included if the signal intensities in both channels differed by more than 2 standard deviations over background in at least 70% of samples in each class (for example T or B). Since filtering may exclude important candidate genes expressed only in subsets of samples (for example only detectable in T or B), as turned out to be the case for *i.e.* the *ERG* gene, another dataset based on flooring was generated (floored dataset). Here, a small, but constant value (in our dataset the value 20) was substituted for a missing signal in one of the two channels, ensuring that a possible strong signal in the second channel is being stored as a "floored" ratio instead of being filtered and removed. A number of bioinformatics

algorithms and statistical methods were applied to explore the gene expression signatures of our datasets and validate the results. We were interested in identifying differential gene expression in matched pairs of benign and tumour tissue. Following normalization a Cy5/Cy3 $\log_2$-ratio was calculated for each gene and an average fold change between T and B was calculated (2^d - where d is the absolute difference of the average log ratio in T and B). The t-score (two-sample t-test) was determined for each gene, to quantify the distance of the average log-ratios between the groups compared to the spread of log ratios within each group. Paired and unpaired t-tests were also performed. The t-scores will be higher if the gene is much stronger expressed in T than in B or if a gene is consistently upregulated in T compared with B tissue.

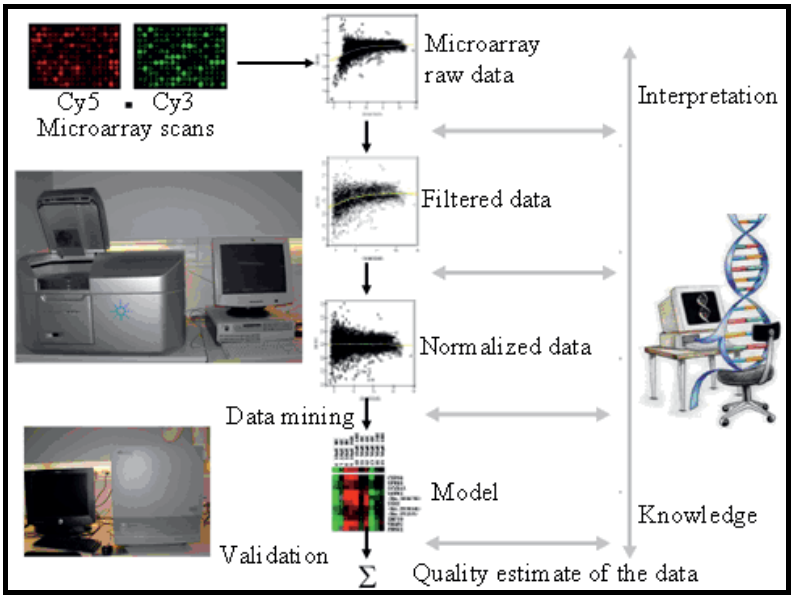

Figure 12. Pictures of the Agilent scanning equipment to the lower left, and a flowchart of the principle for data analysis and validation of gene expression data following microarray Cy5/Cy3 scanning to the right.

Polymerase chain reaction (PCR) and sequencing of PCR products

In the Polymerase chain reaction (PCR) method, a pair of DNA oligonucleotide primers specific for the gene of interest, is used to hybridize with the sample DNA. The temperature of the sample is repeatedly raised and lowered to help the DNA

polymerase transcribe the target DNA sequence. The PCR reaction can be divided into three phases; the exponential phase during which there is an exact doubling of product accumulating at every cycle; the linear phase, during which the reactions start to slow down and the PCR product is no longer doubled at each cycle; and the plateau, at which no more products are being made and eventually the PCR products will begin to degrade. The plateau phase is where traditional PCR takes its measurements, also known as end-point detection. By comparing the intensity of amplified band on a gel to standards of a known concentration, semi-quantitative results may be achieved.

We found *ERG* to be highly overexpressed in a subgroup of our prostate cancer cases (Paper I), and we therefore wanted to investigate in more detail which of the *ERG* isoforms that were responsible for this increased expression. Conventional PCR was performed using a number of *ERG* specific forward and reverse primers followed by agarose-gel visualization of the resulting amplicon. In the course of this work it became obvious to us that the published exon mapping of *ERG* and its isoforms,[121] did not match our findings. We therefore decided to sequence *ERG* ourselves. Conventional PCR and DNA-sequencing was used to characterize the exon structure of *ERG1* and *ERG2*. ss-cDNA was made using *ERG* specific primers and then primary and nested PCR were performed, followed by PCR product clean-up (Qiagen protocol) and sequencing-PCR with BigDye buffer 3.1 (BigDye Terminator v1.1, Applied Biosystems). The sequence reactions were analysed on a 3100 Genetic Analyser (Applied Biosystems). Our publication (Paper I) is now referred to by Genbank concerning the corrected and revised exon organization of *ERG1* (NM_182918) and *ERG2* (NM_004449).

Real-time quantitative PCR (qPCR)

Gene expression changes observed in the microarrays need to be validated with an independent method to ensure that the observed changes are reproducible in a larger number of samples, and to verify that the array findings are not the result of problems inherent to the array technology, but truly reflect the differential gene expression in the samples. Real-time qPCR is a powerful and sensitive gene analysis technique used for a number of applications including quantitative gene expression analysis, genotyping,

SNP analysis, pathogen detection, drug target validation and for measuring RNA interference. Combined with reverse transcription we used the method for relative quantification of messenger RNA (mRNA) in cells and prostatic tissue samples for a number of genes, like *AMACR*, *SIM2*, *ERG* and other *ETS*-family members. Real-time qPCR has a higher linear dynamic range (and accuracy) than DNA microarrays, but the capacity for number of genes studied at the same time is more limited.

Unlike traditional PCR, which measures the amount of accumulated PCR product at the end of the PCR cycles, real-time qPCR measures PCR products as they accumulate during amplification. During the qPCR run two values are calculated. The threshold line is the level of detection at which a reaction reaches a fluorescent intensity above background in the exponential phase of amplification. The PCR cycle at which the sample reaches this level is called the Cycle threshold, the Ct-value. The Ct-value is used in downstream quantification of the PCR product. By comparing the Ct-values of samples of unknown concentration with a series of standards, the amount of template DNA in an unknown reaction can be accurately determined.

A number of platforms exist for real-time qPCR analysis. Our group utilizes the Applied Biosystems platform with the ABI 7900HT Sequence Detection System and SDS2.2 software for analysis of gene expression. In order to detect the accumulation of PCR products, Applied Biosystems utilizes two types of fluorescent reporter molecules based on two types of chemistry, TaqMan chemistry and SYBR Green dye chemistry. As the quantity of target amplicon increases, so does the amount of fluorescence emitted from the fluorophore, and we detect this increase in fluorescent signal. The SYBR Green dye is a highly specific, double-stranded DNA binding dye (minor-groove binder), which detects PCR product as it accumulates during the PCR cycles. Initially our group used the SYBR Green dye, but one disadvantage of this method, which reduces the specificity, is that it detects all amplified double-stranded DNA, including non-specific reaction products which may generate false positive signals. In Papers I to III we used the TaqMan based detection, which exploits the exonuclease activity of AmpliTaq Gold DNA polymerase by using a cleavable fluorescent probe in combination with forward and reverse PCR primers. This approach requires homology for both primers and the probe for producing a

fluorescent signal and is thus much more specific than the SYBR Green method. Primers and probes were designed according to Applied Biosystems guidelines for quantitative assays (ABI Real-time PCR systems Chemistry guide, P/N 7378658 Rev A) (Assays-by-design) or ordered as ready-to-use assays (Assays-on-demand).

During the PCR reaction, specific TaqMan primers and probe anneal to complementary sequences in the target gene ss-cDNA, the probe between the primers. The probe contains a reporter dye linked to the 5'-end of the probe and a nonfluorescent quencher (NFQ) at the 3'-end. When the probe is intact, the proximity of the reporter dye to the quencher dye results in suppression of the reporter fluorescence. During polymerization (extension) of the DNA strands, AmpliTaq Gold DNA polymerase cleaves the probes that are hybridized to the target. This cleavage separates the reporter dye from the NFQ, which results in increased fluorescence by the reporter. The principle for the primer and probe chemistry is illustrated in Fig. 13. The increase in fluorescence signal occurs only if the target sequence is complementary to the probe and is amplified during PCR. Additional reporter dye molecules are cleaved from their respective probes with each cycle, resulting in an increase in fluorescence intensity proportional to the amount of amplicon produced. The higher the starting copy number of the nucleic acid target, the sooner a significant increase in fluorescence is observed above the background. The amount of the nucleic acid target is measured during each amplification cycle of the PCR. An amplification plot graphically displays the fluorescence signal versus cycle number (Fig. 14). The Ct-value for each sample is the fractional cycle number at which the fluorescence passes a set threshold level. The amount of nucleic acid and the Ct-value are proportionally inversely related variables. By using an endogenous control, we can normalize quantification of the target gene for differences in the amount of total RNA added to each reaction.

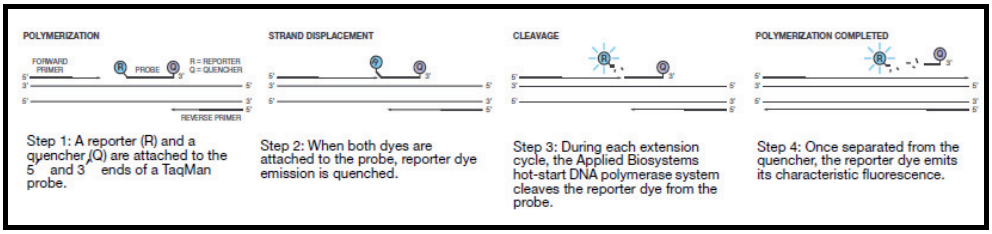

Figure 13. The principle of the TaqMan-probe based assay chemistry. An illustration of how the 5' nuclease chemistry uses a fluorigenic probe to enable detection of a specific PCR product (ABI Real-time PCR systems Chemistry guide, P/N 7378658 Rev A).

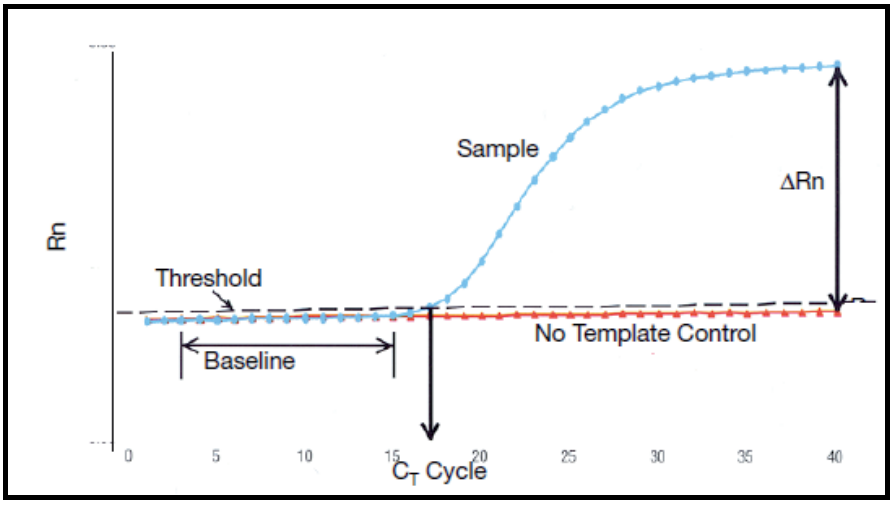

Figure 14. Illustration of a single-sample amplification plot, showing terms commonly used in quantitative analysis. The amplification plot graphically displays the fluorescence signal versus PCR cycle number. (Ct: threshold cycle – the fractional cycle number at which the fluorescence passes the threshold; R_n : normalized reporter – the ratio of the fluorescence emission intensity of the reporter dye to the fluorescence emission intensity of the passive reference dye; ΔR_n: the magnitude of the signal generated by the specified set of PCR conditions (R_n – baseline); Baseline: the initial cycles of PCR, in which there is little change in fluorescence signal; Threshold: A level of ΔR_n that is used for Ct-determination. The level is set to be above the baseline and sufficiently low to be within the exponential growth region of the amplification curve. It is the line whose intersection with the amplification plot defines the Ct.) (ABI Real-time PCR systems Chemistry guide, P/N 7378658 Rev A)

The choice of an appropriate endogenous control is of vital importance for the normalization between samples and analysis of gene expression. The endogenous control is a gene present in each experimental sample and ideally should be expressed in equal amounts in each sample to be analyzed. *β-actin* (*ACTB*), *GAPDH* and other housekeeping genes are commonly used, although expression may vary between samples and introduce bias. An evaluation of the most optimal endogenous control must be performed when designing the experimental protocol, and through this we found *ACTB* to be most optimal for our purposes.

During relative quantification, a change in gene expression in a given sample is analyzed relative to another reference sample (the calibrator), for example an untreated control sample or gene expression in tumour samples relative to benign samples. We applied two alternative methods for relative quantification:

Principle of the Standard curve method: The target and endogenous control amplifications are run in separate tubes, and this method requires the least amount of optimization and validation, but drawbacks are reduced throughput because wells are needed for the standard curve samples and errors in dilutions made in creating the standard curve which may affect the final results. All that is required of the standards is that we know their relative dilutions, and any stock RNA or DNA containing the appropriate target can be used. In papers I and II, serial dilutions of pooled prostate cDNA was found to generate good standard curves for the detection of *ERG* and *SIM2* and endogenous controls in our patient samples. Normalized gene target quantity for the sample is determined from the standard curves (the quantity of the target gene divided by the quantity of the endogenous control) and then divided by the normalized target quantity of the calibrator. The tumour samples are thus expressed as an n-fold difference relative to the calibrator (the benign sample in our experiments).

Principle of the Comparative ΔΔCT calculation method: This method also describes the change in expression of the target gene in a sample relative to a calibrator. It is similar to the standard curve method, except that it uses arithmetic formulas to achieve the results for relative quantification. The advantage of using this approach is that the need for a standard curve is eliminated. For the ΔΔCT calculations to be valid, the efficiency of the amplification of target and endogenous control must be

approximately equal. The amount of target, normalized to an endogenous control and relative to a calibrator, is calculated using the formula $2^{-\Delta\Delta Ct}$.

Total RNA from patient tissues, cell cultures and urine samples (100 µl of whole urine, urine supernatant and pellet resuspended in 100 µl of whole urine after centrifugation at 1000 xg for 8 min) were isolated using the EZNA total RNA Kit (Omega Bio-tek) according to the manufacturer's instructions. Prior to reverse transcription (ss-cDNA synthesis) the RNA was DNase treated, and ss-cDNA was synthesised according to Ambion instructions (MessageSensor™ RT Kit, catalog #1745, Instruction Manual). In the urine study (Paper III), genomic DNA was purified from 1 ml of whole urine (Bergen) or urine supernatant (Hamburg) using the Qiagen M48 Biorobot according to the manufacturers instructions for soft tissue (MagAttract DNA Mini M48 Handbook).

Specific custom TaqMan gene expression assays (primers and probe) were designed (Assays-by-design) for the detection of the endogenous control *β-actin* (*ACTB*), *ERG* isoforms (Paper I), *SIM2* short and long isoform (Paper II) and two *TMPRSS2:ERG* transcripts (Paper III). Ready-made assays (Assays-on-demand) were ordered for an assay common to both *SIM2* isoforms (Paper II) and the endogenous control *GAPDH* (Paper III). *Real-time qPCR:* Hexamer-primed ss-cDNA corresponding to 5-10 ng of prostate total RNA was used in each PCR reaction (Papers I and II). For each reaction in the urine study (Paper III), we used 2.5 µl of total RNA (as ss-cDNA) for the endogenous controls (*β-actin* and *GAPDH*) and 7.5 µl total RNA for the *TMPRSS2:ERG* fusion transcript assays, to obtain optimal Ct-values. The real-time qPCR reaction mixtures were prepared in 96-well optical microtiter plates and amplified in the ABI7900HT Sequence Detection System (SDS) (Applied Biosystems, Foster City, USA) according to protocol as described in Papers I to III. The SDS2.2 software and Excel were used for analysis of relative gene expression using the Standard curve method (Papers I and II) or the Comparative threshold cycle (Ct) method (Paper III) according to program manuals and the ABI User Bulletin #2. For the analysis of the urine samples (Paper III), normal female urine was used both as negative control and calibrator for the analysis. Samples with a real-time qPCR Ct-value above 38 were considered to show no amplification and defined as negative.

This was supported by the observation that one of the duplicates showed poor amplification and often failed above this value. Laxman *et al.* [242] applied the same limit in their urine study.

TaqMan Low-density arrays (TLDA) are customizable 384-well microfluidic cards, allowing for simultaneous real-time qPCR for a large number of genes, greatly improving the efficiency of this method. Using TLDA the gene expression of a number of vascular markers and various *ETS*-family members were validated for 10 prostate patient tumour/benign (T/B) pairs (Paper I). To validate the increased expression of *SIM2-l* isoform in tumour samples in the 96-well format, this isoform was included on a TLDA-card (Paper II). Each low-density array card was configured for 96 different genes in duplicates. ss c-DNA corresponding to 5 ng total RNA was diluted in TaqMan universal buffer and added to each loading well. The cards were prepared according to ABI guidelines and run at the same cycle parameters as 96-well plates before analysis were done by relative quantification using the SDS2.2 software.

All real-time qPCR samples were run in duplicates (TLDA) or triplicates (qPCR) for quality control and statistical verification. Negative controls (non-template controls (NTCs)) were also included on all plates and cards. Gene expression levels found by real-time qPCR corresponded well with, and validated, the microarray findings. We observed that expression ratios found between tumour samples and benign samples were compressed in the DNA microarrays compared with real-time qPCR, which is probably due to the fact that real-time qPCR has a much higher linear dynamic range than the microarray technology (10^6 *vs* 10^2-10^3) and this has previously been reported.[260]

In-situ **hybridization (ISH)**

In-situ hybridization was performed both to validate our microarray and real-time qPCR findings of *ERG* (paper I) and *SIM2* (paper II) mRNA expression in our cases, but also to provide information about their cellular locations.

In-situ hybridization, which preserves the histological morphology of the tissue sample, provides an advantage over extraction based techniques, since they are not able to take into account the heterogeneity of the prostatic tissue and distinguish

between gene expressions performed by the various cells present in the sample. Based upon the ability of nucleic acids to anneal to one another in a sequence specific complementary manner, and our ability to detect this annealing with labelled probes, this method identifies specific nucleic acids *in situ* in histological tissue samples, and therefore provides information concerning both the presence or absence of DNA or RNA and their location within the tissue sample, specifying the specific cell type responsible for gene expression and the location within the cell.[262-264] *In-situ* hybridization provides excellent qualitative information, and a semi-quantitative evaluation of expression levels can be achieved by assessing staining intensity and frequency, although other methods like real-time qPCR are more optimal for quantification of gene expression.[262]

Archival tissues from patients are an invaluable resource to study and validate genes of clinical interest. One challenge of archival tissues is the fact that stored tissues are formalin-fixed and paraffin-embedded, which preserves the histological morphology, but can severely restrict the methods applicable for gene expression analysis.[262, 263] In order for the probe to be able to reach its nucleic target, the tissue needs to be deparaffinised and the RNA unmasked by proteolytic digestion, which removes components of the cell nucleus and cytoplasm to allow probe access. The labelled probe may then hybridize with the target sequence before the visualization steps. Factors affecting hybridization efficacy and specificity are possible RNA degradation and crosslinking with proteins, degree of proteolytic digestion affecting the balance between probe penetration and destruction of cell and tissue architecture, specificity of the probe and the hybridization process and sensitivity of detection.[263]

Two types of probes are established for mRNA *in-situ* hybridization of paraffin embedded tissue, DNA oligonucleotide probes or RNA probes (riboprobes), and they may be radioisotope- (^{33}P or ^{35}S) or non-isotopic (*e.g.* Biotin or Digoxigenin) labelled.[262, 263] Riboprobes are commonly used for RNA detection because RNA-RNA hybrids are more stable against denaturation than DNA-RNA hybrids. Riboprobe vectors can be used to generate both sense and antisense probes to allow control of hybridization, and unbound probe can be digested using RNase which does not digest double stranded hybridized RNA. DNA oligonucleotides may be used for the detection

of high abundance RNA. They have the advantage that they can be custom synthesised with high specificity and may be easily and efficiently labelled. Digoxigenin-labelled-RNA probes (DIG-cRNA) may be generated by cloning (*ERG* in Paper I) or PCR (*SIM2* in Paper II). When cloning, the probe sequence is cloned into a vector containing RNA polymerase promoter sites, and probe molecules are generated using phage RNA polymerase (T7, T3 or SP6) mediated incorporation of labelled nucleotides like digoxigenin-11-UTP. In the PCR method, labelled nucleotides may be directly incorporated or by using labelled primers.[263]

For *in-situ* hybridization of *ERG* (Paper I), histological tissue sections of cases with a high expression of ERG were selected. T3- and T7-containing Bluescript SK^- *ERG*/β-actin plasmids (Invitrogen) were used for synthesis of DIG-RNA *ERG* antisense, *ERG* sense (negative control) and *β-actin* antisense probes (used as endogenous positive control), respectively. Plasmids were cut with restriction enzymes and then sequenced to verify the specificity of the sequence. Initially, short DNA oligonucleotide probes were designed, but they did not provide positive signals, possibly due to sensitivity limitations of short probes with less labelled groups than the Riboprobes. Cases expressing high levels of the *SIM2-s* isoform selected for *in-situ* hybridization as well (Paper II). *SIM2-s* antisense and sense and *β-actin* antisense fragmented DIG-cRNA probes were made by the PCR-based approach. The method was carried out as described in Papers I and II (for *ERG* and *SIM2* respectively) according to protocols. Slides of paraffin-embedded tumour tissue were deparaffinised and the RNA unmasked to allow probe access. Fragmented DIG-cRNA probes for *SIM2* and *ERG* were diluted and incubated on the slides overnight. Post-hybridization wash was done twice before the slides were RNase treated (to remove unbound probe). Re-fixation of the slides was followed by blocking before incubation with anti-DIG-AP Fab fragments (Roche). Staining was done by Liquid Permanent Red Chromogen (LPR) (Dako) and hematoxylin was used for counter staining.

IMMUNOHISTOCHEMISTRY AND TISSUE MICROARRAY (TMA)

As with *in situ* hybridization, used for studying nucleic acids present in tissue samples, immunohistochemistry (IHC) is also widely used for *in situ* studies, studying protein expression patterns in histological samples thereby preserving the cellular morphology of the tissue samples, which is an advantage considering the heterogeneous nature of the prostatic tissue. Both *in situ* hybridization and immunohistochemistry are invaluable tools in combination with other molecular methods like DNA microarray and real-time qPCR when studying gene expression and potential biomarkers for cancer.

Immunohistochemistry may be performed on full sections of paraffin-embedded cancer tissue samples or using large scale tissue microarrays (TMAs), allowing high-throughput molecular profiling of tissue specimens. Large scale expression analysis on tissues by TMA was introduced in the late 1990s, and each TMA contains a large number of cylindrical tissue cores from paraffin embedded full sections (donor blocks, each from a different patient) arrayed into a receptor paraffin block, which is then cut into potentially several hundred thin sections on a slide ready for various *in situ* methods applicable for nucleic acid detection or protein detection (ISH or IHC respectively).[262, 265] The TMA technique has been validated in a number of studies, including prostate cancer,[266] and immunohistochemistry based on TMA versus large sections of tissue show good concordance.[267-269] The principle of the TMA method is illustrated in Fig. 15.

Protein expression of SIM2-s in prostate cancer tissue samples was studied with immunohistochemistry on tissue microarrays (TMAs). Previous studies on tumour cell proliferation by Ki-67 expression in "hot spot" areas on regular tissue slides[270] as well as data regarding expression of p16 and p27[191, 192, 271] were also included in the *SIM2* study (Paper II).

After formalin fixation, radical prostatectomy specimens were totally embedded and studied by whole mount step sections. Immunohistochemistry was performed on 5 μm slides and the area of highest tumour grade was selected for tissue microarray construction, using three parallel cores (0.6 mm in diameter) from each case.[191, 265, 272]

Triplicates were used to account for both intratumour heterogeneity and problems with drop-outs. Epitope retrieval (necessary because of eventual masking or loosing of the antigen due to formalin fixation) was achieved by microwave treatment for 20 min in Tris-EDTA buffer at pH 9.0. Both monoclonal and polyclonal antibodies may be used for immunohistochemistry, in general however, monoclonal antibodies are preferred, due to their superior reproducibility and single specificity.[273] The differences in epitope specificities of various antibodies as well as differences of the detection systems may influence the results. Challenges when optimizing standard protocols for individual IHC include that antigens may be "masked" or lost during formalin fixation, requiring different methods of antigen retrieval, and the various antibodies require different dilutions, optimal incubation periods and staining procedures. The immunohistochemistry for SIM2 was carried out according to protocol as outlined in Paper II.

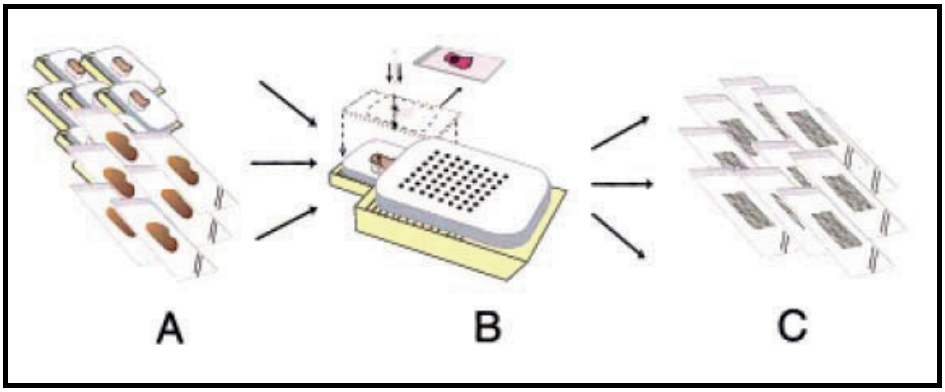

Figure 15. Tissue-microarray construction: (*a*) Slides and paraffin blocks of possible donor tissue are collected from the archive. (*b*) Tissue core biopsy of 0.6 mm in diameter is punched from a pre-selected region of a donor block using a thin-wall stainless steel tube. A hematoxylin & eosin-stained section overlaid on the surface of the donor block is used to guide sampling from representative sites in the tissue. The tissue core is transferred into a pre-made hole at defined array coordinates in the recipient block. (*c*) Sections from a tissue microarray block are ready to be used for simultaneous *in situ* analyses. (Nocito *et al*., Int J Cancer: 94, 1-5.2001)[274]

STATISTICAL METHODS

Statistical methods and analysis are invaluable, essential tools in all clinico-pathological research. In our studies, statistical software packages used were BMDP (BMDP, Los Angeles), SPSS (SPSS Inc, Chicago) and J-Express (Molmine, Bergen). The most common methods for *univariate survival analysis*, the Kaplan-Meier product-limit method as well as the log-rank test, which may be used to test for differences between survival curves from various groups, were used.[275] For *multivariate survival analysis* the most commonly used approach to regression analysis of survival data, the Cox' proportional hazard regression model (likelihood ratio test) was used, including only significant variables ($p < 0.05$) from the univariate analysis.[276]

Associations were assessed with the appropriate methods for categorical or continuous variables. The following methods were applied: Mann-Whitney U test (nonparametric rank test which examines the difference between two groups), Kruskal-Wallis test (nonparametric rank test which examines the difference between two or more groups), Pearson's χ^2 test (parametric correlation coefficient which measures the strength of the linear association between outcome and exposure variables), Fisher's exact test (the preferred method when studying categorical variables in small groups), the Chi-X^2-test (examines the association between two categorical variables in larger groups) and the paired and unpaired T-test (provides associations in small groups based on a normal distribution).[276] The Spearman rank correlation test (Spearman rho test) was used to investigate possible *correlations* between continuous variables (nonparametric method based on ranks which calculates a correlation coefficient providing a measure of the strength of the association between two variables).

Microarray gene expression analysis. In the analysis of gene expression data, a number of methods were used to compare gene expression between groups (paired and unpaired T-tests), prediction of gene profiles (Fishers linear discriminant), hierarchical cluster analysis to identify homogeneous groups based on gene expression (average linkage and Pearson's correlation), testing of the predictability of gene

expression classifiers (*i.e.* T and B) (leave one out crossvalidation) and to test prediction accuracy achieved against pure chance (random permutations).[149]

Statistical cut-off values. Sometimes continuous variables may be divided into subgroups forming categories for relevant statistical analysis. When doing this it is important to avoid cut-point selection bias, and this was done by categorization by median, quartiles or tertiles while considering the size of the subgroups, the frequency distribution and the number of events in each category.[277] This categorization method was used for the analysis of expression studies and immunohistochemistry results, survival analysis and in grouping of the patients according to preoperative s-PSA (<4, 4-10, 10-20, >20) (or median s-PSA),[278] age (two groups with the age ≤ 63 or >63), pathological stage (two groups with ≤ pT2 or ≥ pT3) or by Gleason score (two groups with scores ≤ 6 and ≥ 7) for association and correlation studies.

MAIN RESULTS

Paper I

In Paper I we identified a number of genes differentially upregulated in tumour tissue compared with benign tissue, including previously described genes like *AMACR* and *hepsin*, as well *SIM2*, which was studied in more detail in Paper II, and *ERG*. The transcription factor *ERG* was found to be highly upregulated in a subset of prostate cancer patients (more than 20-fold upregulation in tumour samples (T) compared with benign samples (B) in approximately 50% of matched T and B tumour pairs). *ERG1* was identified as the predominant isoform expressed in prostatic tissue and *ERG* was shown to be expressed mainly in epithelial malignant tissue but also in vascular endothelial cells. The exon organization of *ERG1* and *ERG2* was also revised. The upregulated expression of the *ETS*-members *ELF5*, *ETV1* and *ETS1* was found to be inversely correlated with the overexpression of *ERG*, while *ETS2* was moderately to strongly downregulated in most prostate cancers. Prostate tissue biopsies contain different cell types, including epithelial cells, stromal cells, endothelial cells and leukocytes in vessels and infiltrates. A very important result of this paper was to show by *in situ* hybridization that the abundant overexpression of *ERG* mRNA occurred in the epithelial cancer cells. Previously, ERG was known as an endothelial transcription factor, and the *in situ* hybridization was able to demonstrate *ERG* expression in endothelial cells although this was a minor contribution to the total *ERG* mRNA detected in cancer.

Paper II

In Paper II we validated the upregulation of SIM2 in prostate cancer. mRNA from both *SIM2* isoforms, *SIM2-s* and *SIM2-l*, were shown to be highly upregulated in malignant tumour tissue compared with benign tissue. The SIM2-s protein was expressed in 44 of 103 prostate carcinomas (43%) and was associated significantly with preoperative serum PSA, high histological grade, extra-prostatic extension and increased tumour cell proliferation by Ki-67 expression as well as reduced expression of the p27 protein. A univariate survival analysis of 103 prostate cancer patients

showed a significant association between positive SIM2-s expression and reduced prostate cancer-specific survival. WHO histologic grade and SIM2-s expression were significantly associated with survival in a univariate analysis and only SIM2-s remained in a multivariate survival model as a significant independent predictor of reduced cancer-specific survival.

Paper III

In Paper III we were able to detect *TMPRSS2:ERG* gene fusion mRNAs in urine pellets from 19 out of 55 patients (34,5%) treated with radical prostatectomy. Gene fusion *TMPRSS2:ERGa* was identified in the majority of patients with a positive fusion status (89.5%), while *TMPRSS2:ERGb* was identified in a minority (36.8%) (7/19 positive cases). Five of these patients were positive for both fusion isoform a and b. Prostatic massage prior to urine collection improved the sensitivity of the test (69% positive cases in samples collected after prostatic massage *vs.* 24% positive cases in samples collected without prior massage). The highest detectable levels of gene fusion transcripts were found in total RNA from urine pellets collected after prostatic massage. The presence of the *TMPRSS2:ERG* gene fusion in urine was found to correlate with adverse clinicopathological variables, such as increasing s-PSA, high pathological stage and Gleason score 7 or higher.

DISCUSSION OF RESULTS

Although there was initially (6 – 10 years ago) much scepticism to the potential, reliability and precision of the microarray technology, this methodology revolutionized global gene expression studies and is now widely applied in cancer research. Microarray studies analyze the genome-wide gene expression at a given time under certain conditions. An initial microarray gene expression study by Halvorsen *et al.*[149] provided us with a microarray data set used as a basis for the generation of gene lists of up- and downregulated genes in prostate cancer tissue compared with benign tissue, based on filtered *(SIM2)* or floored *(ERG)* datasets. Filtering of the genes before inclusion in the dataset may exclude candidate genes expressed only in subgroups of the samples. To compensate for this, the alternative method of flooring was also used. "Flooring" means to substitute a low and fixed value for either the denominator or the nominator when the hybridization signal is zero or close to zero in either channel. Both filtered and floored datasets were used for identification of differentially expressed genes. In the list of the differentially most upregulated genes based upon the floored dataset (Table II, Halvorsen *et al.*[149]) several unknown ESTs where highly ranked. ESTs (Expressed Sequence Tags) are short single-read transcript sequences of yet unidentified genes. The upregulated EST (IMAGE number 767130) (Table II, Paper I), exhibited full homology with the 3' untranslated region of the *ERG* mRNA based upon BLAST analysis. Real-time PCR assays were designed for different regions of the *ERG* reading frame (ORF) and confirmed the very high overexpression of *ERG* in prostate cancer. The ETS family of transcription factors, and its member *ERG* in particular, was studied further in Papers I and III. The transcription factor *SIM2* ranked second on the list of upregulated genes in the filtered dataset, after the already well known upregulated gene *AMACR*,[279, 280] and was characterized further in relation to prostate cancer in Paper II. The majority of microarray gene expression studies of prostate cancer[149, 232, 255, 279, 281-294] provide potential molecular signatures for prostate cancer *vs*. benign prostate, classification of moderate- *vs*. high-grade prostate cancer, prediction of PSA recurrence, tumour aggressiveness, predictive signatures, prediction of PSA recurrence *etc*. The search for clinically useful and applicable microarray

profiles and molecular biomarkers are ongoing, and the identification of *ERG* overexpression in half of all prostate cancer marks a new era in the understanding of prostate carcinogenesis.

SIM2

Based upon our initial microarray studies of gene expression in localized prostate cancers compared with benign tissue,[149] we decided to select highly differentially upregulated genes for validation and further studies. *SIM2* ranked second on the gene list based on the filtered data set after *AMACR* (Table I in Halvorsen *et al.*[149]). We found that both the short and long isoforms of the *SIM2* gene were expressed in benign prostatic tissue and that the expression was significantly upregulated in prostatic cancer tissue (Paper II). Tumour cell expression of the SIM2-s protein was shown to be associated with more aggressive clinicopathological features and reduced cancer-specific survival.

Although there are several microarray studies performed on prostate cancer, with the exception of one,[295] none of them focus on or mention *SIM2*.[255, 283-287, 295, 296] But, when analyzing these studies in the publicly available gene microarray datasets within the Oncomine database,[297] the data support our findings that the expression levels of *SIM2* were significantly higher in prostate cancer tissue than in benign samples. The fold change of expression in tumour tissue compared with benign tissue varied between 2.5 to 65.8 in these studies according to Oncomine.

SIM2 is a member of the PAS (Per/Arnt/Sim) family of transcription factors involved in regulation of key oxidative enzymes involved in carcinogen metabolism and cancer surveillance.[148, 153] Although the exact molecular role and function of SIM2 and pathways involved are not entirely clear, studies involving SIM2 in processes potentially involved in the carcinogenic processes are emerging. *BNIP3* (HIF1α-dependent, hypoxically induced pro-apoptotic Bcl-2 BH3-only family member) has recently been identified as a novel target of SIM2-s.[298] The SIM2-s mediated repression of *BNIP3*, which has recently emerged as a pro-autophagic factor, is coupled to the role of SIM2-s in increasing prostate tumour-cell survival during

prolonged hypoxic conditions. Aleman *et al.*,[299] have found a possible link between SIM2-s and differentiation. In antisense-SIM2-s treated colon-cancer derived cells (RKO cells), the expression of a key stress response gene, growth arrest and DNA damage gene (GADD) 45α, was found to be upregulated compared with normal cells, leading to subsequent apoptosis. Various apoptosis and differentiation-related genes were found to be up-regulated in the SIM2-s antisense-treated RKO cells. Key pathways, including GADD, caspase and p53 function, were identified as critical to the function of the *SIM2-s* gene. Li *et al.*[300] proposed *SIM2*, as a candidate β-catenin/TCF target gene in Wilm's tumours, since it contains at least one consensus TCF site in its promoter region, but states that additional studies will be necessary to validate *SIM2* as a β-catenin/TCF-responsive gene. Gravdal *et al.*[195] however, did not find an association between *SIM2-s* and *β-catenin* expression in the prostate cancer series.

There are tissue differences both with regard to *SIM2* up- or downregulation promoting or suppressing tumourigenesis and with regards to isoform preferences in both benign and malignant tissues.[155, 157, 158, 301] *SIM2-s* is expressed in benign breast tissue and has been found to be downregulated in breast cancer derived cell lines and in human breast cancer samples, and linked to tumour suppressor activities, including decreased expression of matrix metalloprotease 3 (MMP3) and a role in the control of epithelial to mesenchymal transition (EMT). Loss of *SIM2-s* in MCF-7 breast cancer cells correlated to cell survival through the activation of SLUG-mediated EMT.[157, 158] These tumour suppressor properties of SIM2-s in breast cancer are contradictory to its upregulation of expression and oncogenic cancer promoting role in colon, pancreas and prostate cancers, and may reflect different tissue specific functions or differences in effect depending upon the cellular context.

Previous studies found *SIM2-s*, but not *SIM2-l*, to be associated with pancreas, prostate and colon cancer, although not expressed in the normal tissues of these organs. Elevated expression of *SIM2-s* seen in early colon adenomas and in BPH,[148] raises the possibility that this activation may be an early event in tumourigenesis. In contrast, we found both *SIM2* isoforms to be expressed in both benign and tumour tissue (Paper II), and indeed the ratio between expression levels in tumour tissue and

benign tissue was much higher for *SIM2-l* compared with *SIM2-s* (fold change 6.7 *vs.* 3.9). This is the first study to identify *SIM2-l* as overexpressed in prostate cancer and the relevance of this isoform should be studied further.

Our findings that SIM2 expression might be important for clinical progression of prostate cancer and associated with reduced prostate cancer-specific survival support the proposal of *SIM2-s* as a candidate for targeted therapy in prostate cancer.[156, 302] Arredouani *et al.*,[302] found that human HLA-A2.1 restricted SIM2 epitopes induce specific T-cells *in vivo*, and that anti-SIM2 antibodies are detectable in the sera from prostate cancer patients. They suggest SIM2 as a prostate cancer associated antigen that is a potential target for prostate cancer immunotherapy.

The precise function of genes regulated by SIM2, molecular pathways involved and the role of both *SIM2-l* and *SIM2-s* in prostate cancer are not clear and needs to be further elucidated, as well as its prognostic and therapeutic potential and utility.

ERG AND OTHER ETS TANSCRIPTION FACTORS IN PROSTATE CANCER

Increased expression of *ERG* in prostate cancer

The *ERG* gene was first described by Reddy *et al.,*[116] in 1987, who suggested that *ERG* might be a member of the *ETS* oncogene family. Considering that *ETS1* and *ETS2* are translocated in certain leukaemias, they discussed the possibility of *ERG* being linked to any human malignancy either by amplification, translocation or other rearrangement. In Paper I, we identified *ERG* as one of the most highly upregulated genes in a large subset of prostate cancer patients through gene expression microarray analysis and validation with real-time qPCR. Approximately 50 % of our patient tumour tissue samples showed a high overexpression of *ERG* compared with benign tissue samples. This was a confirmation of the independent work done by Petrovics *et al.,*[78] who in 2005 published *ERG1* as being frequently overexpressed in the majority of prostate cancers. Since then several studies have confirmed the upregulation of *ERG*

in prostate cancer, ranging from 40-80% of prostate cancer patients being fusion positive.

Exon organization of *ERG*

During the course of our work it became apparent that the published exon organization of *ERG* isoform 1 (NM_182918) and isoform 2 (NM_004449),[121] did not match our experimental findings. Especially, when analyzing agarose gels run after PCR amplification across pre-mRNA splice sites, there was a discrepancy between the observed gel bands and the ones we expected according to Owczarek *et al.*.[121] A gel band of approximately 70 nucleotides present in *ERG1* but not in *ERG2*, did not correspond with the assumed exon overview. We therefore undertook extensive sequencing of the ERG cDNA from benign prostate tissue, and based upon this we revised the exon organization of these two isoforms (Fig. 1, Paper I) and the relevant Genbank Accession numbers (NM_182918 and NM_004449) were corrected accordingly. The observed gel band corresponded to the 72-nucleotide exon 9 in the revised *ERG1*.

Mechanism behind upregulated gene expression of *ERG* and other *ETS* family members

Later in 2005, Tomlins *et al.*,[80] through a bioinformatic method called cancer outlier profile analysis (COPA), which analyses microarray data for marked overexpression of genes in subsets of cases, identified the two *ETS* transcription factors *ERG* and *ETV1* as outliers and highly overexpressed in prostate cancer. COPA is useful to avoid that strongly expressed genes in subgroups go undetected, and in this respect serves a similar purpose as the "flooring" of gene expression data used in our studies. Without "flooring" both *ERG* and *ETV1* would also have been filtered away in our microarray data. Importantly, Tomlins *et al.*[80] identified the mechanism behind the observed upregulation of *ERG*, and also several other *ETS* transcription factors, in prostate cancer. The initial finding was that the promoter of the *TMPRSS2* gene was fused to the *ERG* reading frame due to a chromosomal translocation or deletion. The *TMPRSS2* promoter is androgen-dependent and highly active in prostate luminal cells. This

translocation therefore results in strong aberrant activation of the *ERG* gene, which is silenced in benign prostate epithelium. They identified recurrent gene fusions of the 5' untranslated region of *TMPRSS2* also to *ETV1*, and suggested that the androgen-responsive promoter element of TMPRSS2 drives the overexpression of *ETS* family members in prostate cancer. Although gene fusions, through translocations or interstitial deletions are very common oncogenic mechanisms in haematological tumours and sarcomas, as exemplified by the gene fusion *BCR-ABL1* (Philadelphia chromosome) in CML,[303] as well as the gene fusions between mainly *FLI1* but also *ERG* and other *ETS* transcription factor genes and *EWS* in Ewing's sarcoma,[130] this finding was the first identification of a recurrent gene fusion in solid epithelial adenocarcinomas. The two alternative isoforms of *TMPRSS2:ERG* were called *TMPRSS2:ERGa* and *TMPRSS2:ERGb* (GenBank accession numbers DQ204772 and DQ204773 respectively), of which *TMPRSS2:ERGa* is the most prevalent one, and they were further analysed in our urine study (Paper III). The previous absence of identified gene fusions in common solid epithelial tumours like prostate cancer and breast cancer has been attributed to technical difficulties associated with their cytogenetic analysis.[304, 305] Also, epithelial cancers are clonally heterogeneous, which makes it difficult to separate tissue with chromosomal aberrations from clinically irrelevant tissue. These recurrent gene fusions in prostate cancer were identified on the basis of gene expression data, thereby bypassing the technical limitations of cytogenetics in solid cancers.

The mechanisms involved in generating the double-strand break (DSB) that leads to the *TMPRSS2:ERG* gene rearrangements are not completely understood, but recently a proposed mechanism behind this was published,[306] linking the generation of *TMPRSS2:ERG* fusions through DSBs to androgen stimulation. Topoisomerase IIβ (TOP2B) is required for androgen-mediated gene expression and is recruited to the *TMPRSS2* promoter by the androgen receptor (AR) which is induced upon luminal epithelial cell differentiation. Topoisomerases catalyze transient DSB as part of their normal activity, but dysfunction may lead to aberrant rejoining of DSB and resulting translocations. Both *AR* and *TOP2B* were highly coexpressed in *TMPRSS2:ERG* fusion positive neoplastic cells comprising prostatic intraepithelial neoplasia (PIN)

lesions. It was therefore hypothesized that androgen receptor signalling might lead to TOP2B-mediated DSBs and that such breaks could be involved in the generation of *TMPRSS2:ERG* fusions.

Since the first paper on the gene fusions was published,[80] a number of different isoforms of *TMPRSS2* and *ERG* gene fusions have been identified, as well as alternative 5' partners for *ERG*, alternative *ETS* 3' partners for *TMPRSS2* and alternative 5' partners for other *ETS* factors involved in recurrent gene fusions. A classification of *ETS* gene fusions in prostate cancer have been proposed.[131, 307-309] Most of the fusion genes and transcripts characterized so far have the protein coding region derived largely from the *ETS* gene, and in all fusion variants the ETS domain appears to be retained and the DNA binding domain preserved. The most common variants involve *TMPRSS2* exon 1 or 2 fused to *ERG* exon 2, 3, 4 or 5.[80, 138, 140, 141, 307, 310-313] Less frequent combinations include *TMPRSS2* exon 4 or 5 fused to ERG exon 4 or 5[313] and, in one case, *TMPRSS2* exon 2 was found fused to inverted *ERG* exon 6-4.[312] These variant fusion transcripts most probably represent alternative splicing variants, and there are distinct phenotypic effects produced by different isoforms. The *TMPRSS2:ERG* fusion is found in approximately 90% of fusion positive cases, followed by *ETV1* and *ETV4* fusions, respectively. The reason for the observed frequencies of fusion partners with *TMPRSS2* is unclear.[314]

Differential expression of other *ETS* transcription factors

Other *ETS* transcription factors were also differentially expressed, either up- or downregulated, in a number of our prostate cancer samples (Paper I). Our Agilent microarray expression data were available for 20 different *ETS* transcription factors. Most of the *ETS* factors did not show a differential gene expression in tumour tissue compared with benign tissue, we did find however, a negative correlation between the expression of *ERG* and *ETV1*, *ETS1* and *ELF5*, *i.e.* some cases lacking *ERG* overexpression showed upregulation of these instead, suggesting that these transcription factors might substitute for *ERG* in prostate cancer. Whether genetic rearrangements are also responsible for the observed increased expression of *ETS1* and/or *ELF5* in a subset of prostate cancers remains to be determined.

ETS2 was moderately to strongly downregulated in a majority of our prostate cancer samples. Both *ERG* and *ETS2* are located on chromosome 21 (21q22.2-q22.3) only approximately 150 kb apart,[117, 121] and it is likely that the *ETS2* gene, which is located between *TMPRSS2* and *ERG*, is lost in the process when the *TMPRSS2* promoter is fused to the *ERG* reading frame. The ETS transcription factors ERG and ETS2 may have competing roles in transcription complexes,[73, 75] and the loss of ETS2 could further promote ERG activity. The interactions between different ETS factors may have stimulatory or repressive effects on transcription. As one example, together with the AP1 transcription complex (FOS/JUN), ERG promotes *collagenase-1* (*MMP1*) activation and ETS2 promotes activation of *stromeolysin-1* (*MMP3*). However, when ERG and ETS2 is coexpressed, ERG can bind to the *stromelysin-1* promoter and repress its activation by ETS2.[73, 75]

One of our patients without increased *ERG* expression had instead a marked overexpression of *ETV5*. We suspected that this *ETS* factor could also possibly be involved in a gene fusion responsible for the upregulation, maybe with *TMPRSS2*. Exon-walking quantitative PCR identified an increased expression of the exons located towards the 3'-end of the gene compared with a 5'-end exon, suggesting the presence of a possible gene fusion. We performed 5' RACE (5' RNA-ligase-mediated rapid amplification of cDNA ends) and sequenced the 5' end of the transcripts (unpublished work) in an attempt to identify a possible gene fusion partner. Unfortunately, this time consuming and challenging method revealed only the complete *ETV5* nucleotide sequence, and we were at that time unable to identify any gene fusions. The reason for this is unclear, but could be due to one intact allele of ETV5 or contamination during PCR amplification by a normal ETV5 sequence. At the end of this work, Helgeson *et al.*,[81] published the fusion between *ETV5* and *TMPRSS2* confirming our hypothesis, and we therefore did not go any further with this work. They also identified *SLC45A3* as an alternative 5' partner for *ETV5*.

Carcinogenesis

Due to its high prevalence and recurrence in prostate cancer, it seems clear that *TMPRSS2:ERG* gene fusions play an essential role in the development of prostate

carcinogenesis. ETS factors have been reported to play an important role in extra-cellular matrix remodelling and epithelial-to-mesenchymal transition (EMT) and their overexpression has been linked to increased motility, invasion and metastasis in various cancer models, as mentioned in the introduction. The actual role, and when in the malignant process the *TMPRSS2:ERG* gene fusions appear, is presently being thoroughly investigated. We suggested that the observed activation of *ERG*, or related *ETS* factors, might reactivate an embryonic proliferation program (Paper I).

The relationship between the epithelial luminal cells and basal cells in the prostate, both during development, in the adult and during cancer development has been discussed extensively, and the cancer initiating cell type has remained unclear. Pathology observations, showing that more than 95% of prostate cancers express luminal markers with absence of basal cells, have led many to propose luminal cells as the source of prostate cancer cells.[315, 316] Crum *et al.*[52] states that under normal conditions, there is little evidence that the basal cells of the prostate differentiate into secretory cells, and that both basal and luminal compartments appear capable of independent proliferation and self-renewal. This is contradictory to recent reports that prostate cancer may originate among basal cells.[49, 51] Goldstein *et al.*,[35] proposed an epithelial hierarchy of the normal prostate in which a stem cell within the basal layer of the normal prostate can give rise to multi-potent progenitor cells. This progenitor cell likely gives rise to neuroendocrine cells, mature basal cells and luminal-restricted progenitors that can generate mature luminal cells. They suggested that prostate cancer likely originates from a progenitor cell with multi-lineage differentiation potential or a mature cell that acquires this property. It could be that re-programming of basal cells and progeny causes them to differentiate in the direction of the luminal cell lineage and that prostate carcinogenesis *in vivo* is associated with differentiation towards mature luminal cells. This is consistent with the generally acknowledged loss of many basal cell markers and the presence of luminal cell markers in prostate cancer biopsies.[51] Lawson *et al.*[49] assessed the tumourigenic potential of different prostate cell subpopulations, and found that basal/stem cells were more efficient targets for transformation than luminal cells following the introduction of multiple alternative oncogenic stimuli.

Emerging evidence suggests that the *TMPRSS2:ERG* gene fusion is an early event in human prostate tumorigenesis. *ERG* associated gene rearrangements and the resulting aberrant overexpression of the transcription factor may represent critical cooperative progression events in prostate tumourigenesis. *ERG* rearrangements have been identified in approximately 50% of localized cancers and 30% of metastatic cancers, but only in approximately 16-20% of PIN and HGPIN lesions.[137, 307, 310, 317-320] Genetic rearrangement of *ERG* is infrequently found in HGPIN and with the same *ERG* fusion pattern in the adjacent invasive prostate cancer.[320, 321] However, the majority of prostate cancer specimens with *ERG* did not display this rearrangement in the associated HGPIN and there were no cases of *ERG* fusion in HGPIN associated with *ERG* negative cancers. Therefore, it seems likely that the *TMPRSS2:ERG* translocation is an early event in human prostate tumourigenesis associated with progression from HGPIN to cancer. Increased expression of *ERG* is found to increase invasion, but is not sufficient to drive the malignant transformation from HGPIN to cancer.[319] However, forced expression of the androgen receptor, ERG and activated AKT1 can induce prostate cancer from basal cells.[51] Deletion of all or part of the tumour suppressor gene *PTEN* (which is one way to activate the AKT pathway) is a frequent event in prostate cancer development (30%-70% of the cases) and *PTEN* loss and *ETS* gene rearrangements are proposed to be critically important and common molecular events in prostate carcinogenesis.[80,319,321-322] Several authors[319, 323, 324] have speculated that loss of *PTEN* or *NKX3-1* may precede and cooperate with *ETS* gene fusions to drive cancer development from PIN to malignant cancer. King *et al.*[322] found that transgenic *TMPRSS2:ERG* mice developed PIN, but only in the context of PI3-kinase pathway activation, and suggested that additional events are likely required for actual malignancy. Yoshimoto *et al.*[325] demonstrated that the co-occurrence of *PTEN* loss and *ERG* genetic rearrangement was a statistically independent predictor of biochemical failure after radical prostatectomy, and Carver *et al.*[321] suggested that *ERG* targeted therapy may be effective at preventing the transition between HGPIN and invasive cancer, but this still remains to be studied further.

Recently, Yu *et al.,*[326] provided a link between the androgen receptor (AR), epigenetic regulation by polycomb and *TMPRSS2:ERG* gene fusions in prostate

cancer. They found that ERG disrupts AR signalling by inhibiting AR expression, binding to and inhibiting AR activity at gene-specific loci, and inducing repressive epigenetic programs via direct activation of the H3K27 methyltransferase polycomb protein EZH2, which has also been shown by Kundefranco *et al.*,[327] who identified *EZH2* as a target gene of *ERG* and *ESE3*. Interestingly, they showed that the AR and ERG co-occupy target genes, that there is an extensive overlap between AR and ERG binding sites, and they suggested that *TMPRSS2:ERG* plays a central role as a ”malignant regulatory switch” that shuts down androgen signalling, thereby inhibiting normal prostate differentiation and inducing an epithelial stem cell-like dedifferentiation program that may be exploited during carcinogenesis.

Screening

One of the major challenges in cancer diagnostics and treatment is to find genes and proteins specific for cancer and preferably not present in benign tissue, which ideally are of diagnostic use, provide information regarding prognosis and predictive information for choosing treatment options, as well as represent targets for clinical therapy. There is an urgent need for biomarkers capable of distinguishing between indolent tumours and aggressive tumours. Biomarkers that are specific for cancer and only present in cancerous tissue and never in benign tissue, are called pathognomonic (“characteristic for a particular disease”) biomarkers. The current routinely used and dominating biomarker for prostate cancer is prostate specific antigen (PSA). One of the disadvantages of PSA, as discussed in the introduction, is the lack of specificity and sensitivity for prostate cancer. PSA is produced in benign as well as cancerous prostate cells, and elevated levels of PSA are present in benign conditions like BPH and prostatitis as well as in non-physiological conditions, for example following prostatic massage. The gene fusions between *TMPRSS2* and *ERG* and other *ETS* transcription factors have only been identified in prostate cancer and the precursor stages PIN and HGPIN, and therefore these gene rearrangements are pathognomonic and extremely attractive potential biomarkers for clinical use as well as potential targets for therapy. The ability to detect these pathognomonic gene fusions in body fluids without the need for invasive strategies like biopsies is also an advantage.

We analyzed our urine samples (Paper III) for the first two *TMPRSS2:ERG* isoforms identified, which are also among the most prevalent isoforms. We were able to detect *TMPRSS2:ERG* isoforms *a* and *b* in urine from prostate cancer patients taken prior to radical prostatectomy. We were very surprised when we identified the presence of more than one *TMPRSS2:ERG* gene fusion variant in one urine sample from five out of 19 fusion positive patients. Before our work on the urine samples was published, this observation was supported by the findings of others in both urine samples[252] and prostate tissue samples.[307, 328] Localized prostate cancer is typically multifocal, and various isoforms of the *TMPRSS2:ERG* gene fusion as well as other *TMPRSS2:ETS* variants may be identified in different foci of a prostate cancer,[328-331] reflecting the heterogeneity of the various foci present in a prostate cancer patient. This provides a challenge when screening for prostate cancer and deciding on the number of biopsies necessary to perform as well as how to analyse the results. Do the tumour clones from which the biopsies were taken reflect the true picture of the cancer? However, different sites of metastatic prostate cancer from the same patient are uniformly fusion-positive or fusion-negative.

Our ability to detect the *TMPRSS2:ETS* gene fusion products in body fluids like urine, provides both the possibility of easy noninvasive sampling and also circumvents the aforementioned challenge of multifocality and heterogeneity which has to be taken into consideration when performing and evaluating tissue sample biopsies. Since the first report of identification of the gene fusion in urine[242] and the completion of our work, there has been published a number of papers regarding the identification of *TMPRSS2:ERG* and other potential prostate cancer biomarkers either alone or together in urine samples from prostate cancer patients (reviewed by Jamaspishvili *et al.*[332]), regarding diagnostic, prognostic and predictive purposes. Serum and prostatic fluids are also being evaluated. Urine sample detection relies on the fact that prostate cancer tumour cells are shed into the urine, and we demonstrated (Paper III) that the sensitivity of the detection of *TMPRSS2:ERG* greatly improved with prostatic massage prior to urine sampling.

Hessels *et al.*[252] calculated a score for the prostate specific non-coding RNA *PCA3*, which is highly upregulated in most prostate cancers, and found that a

combination of *TMPRSS2:ERG* fusion status and the PCA3 score improved the sensitivity for cancer diagnosis. Later, in a multiplex study of biomarkers in urine sediments using qPCR, Laxman *et al.*[251] found that detection of *GOLPH2*, *SPINK1*, *PCA3* and *TMPRSS2:ERG* fusion transcripts were significant predictors of prostate cancer, and outperformed serum PSA or PCA3 alone in detecting prostate cancer (sensitivity 66% and specificity 76%). Mao *et al.*,[199] was able to detect the *TMPRSS2:ERG* fusion gene in circulating prostate cancer cells and suggested that it had a potential in monitoring tumour metastasis.

A recent very interesting strategy to identify the presence of *TMPRSS2:ERG* gene fusions in prostate cancers, is the antibody-based detection of a truncated ERG protein by Park *et al.*.[333] They characterized a rabbit anti-ERG monoclonal antibody (clone EPR 3864; Epitomics, Burlingame, CA) using immunoblot analysis on prostate cancer cell lines, synthetic TMPRSS2-ERG constructs, chromatin immunoprecipitation, and immunofluorescence. ERG protein expression was correlated with the presence of *ERG* gene rearrangements in prostate cancer tissues using a combined immunohistochemistry (IHC) and fluorescence in situ hybridization (FISH) analysis. ERG expression was confined to prostate cancer cells and high-grade prostatic intraepithelial neoplasia associated with ERG-positive cancer, as well as vessels and lymphocytes. They detected ERG rearrangement prostate cancer with close to 100% sensitivity (96%) and specificity (97%). This study identifies a specific anti-ERG antibody and demonstrates association between *ERG* gene rearrangement and truncated ERG protein product expression. ERG protein expression may be useful for molecularly subtyping prostate cancer based on ERG rearrangement status.

The usefulness of *TMPRSS2:ETS* gene fusions and other recently identified potential biomarkers, like Sarcosine (a metabolite)[196] and *Annexin A3* (*ANXA3*) (negatively associated with prostate cancer),[197, 198] as clinically useful tests, either alone or in combinations with other biomarkers including PSA, needs to be studied further. Most likely specific and sensitive multiplex biomarker urine tests will be developed which hopefully provides more specific information and probably outperforms PSA alone in prostate cancer diagnosis and treatment.

Prognosis

TMPRSS2:ETS fusions may represent the most common recurrent structural aberration and gene fusion among all human malignancies. There are a great number of conflicting reports regarding possible associations between gene fusion status and patient outcome, but emerging data suggest that *TMPRSS2:ERG* positive cancers represent a subclass of prostate cancers that have a more aggressive nature and poor prognosis,[131, 309] supporting our findings of a correlation between a positive *TMPRSS2:ERG* fusion status (Paper III) and adverse clinicopathological variables (increasing preoperative s-PSA, Gleason score ≥ 7 and high pathological stage), although we were the first to report an association between positive fusion status and increasing s-PSA (Paper III). These gene fusions have been variously associated with high pathological stage[132] and higher rate of recurrence[133] in independent cohorts of surgically treated localized prostate cancer cases, and the presence of gene fusion has been scored as the single most important prognostic factor.[133, 134] In an assessment of gene fusion status in a population-based 'watchful waiting' cohort of men with localized prostate cancer, the *TMPRSS2:ERG* fusion positive subset of 15 % men was found to be significantly associated with prostate cancer specific death.[136] In other studies, significant associations has been found between *TMPRSS2:ERG* rearranged tumours and higher tumour stage, as well as the presence of metastatic disease involving pelvic lymph nodes,[137] more frequent gene fusions in moderate to poorly differentiated tumours as compared with well-differentiated tumours[138] and a significant higher risk of recurrence (58.4% at 5 years) than fusion negative patients (8.1%).[134]

Fusion of *TMPRSS2* and *ERG* can occur through either translocation between both chromosome 21s or interstitial deletion (Edel) of the genomic material between *TMPRSS2* and *ERG.*[137, 140] Interestingly, observations suggest that the rearrangement through Edel represents an aggressive molecular subtype of prostate cancer.[135, 137, 317] However, many studies have reported a positive association or an absence of clinical correlation between the *TMPRSS2:ERG* fusion and prognosis.[78, 139-141] Many of the negative reports have small sample sizes though, and more studies are needed with larger patient cohorts to resolve specific prognostic associations and assess the actual

clinical usefulness of the gene fusions. Phenotypic morphological associations of the gene fusions (and their variants) have been identified,[334] but molecular associations require follow-up studies.

General conclusions

The discovery of the upregulation of the oncogene *ERG* in prostate cancer in 2005 is one of the major success stories of genome-wide microarray studies. The following discovery of the mechanism behind this upregulation due to the gene fusions between *TMPRSS2* and *ERG* (or several alternative *ETS* transcription factors) and the progress in the elucidation of the biological functions and mechanism of ERG activity, have been of major importance in the ongoing understanding of prostate carcinogenesis. Representing pathognomonic biomarkers, it is likely that these gene fusions will be useful for the screening of prostate cancer and possibly also of prognostic or predictive use. A future issue is whether and how ERG and other ETS family members, or parts of their regulatory networks, may be utilized for the development of therapeutic molecular targets.

SPECIFIC CONCLUSIONS

1. The *ETS* family transcription factor *ERG* was identified as highly and consistently upregulated in prostate cancers. Approximately 50% of the prostate cancers showed 20-fold to more than 100-fold increased expression of *ERG*. Although endothelial cells expressed *ERG*, epithelial cancer cells were the main source of *ERG* expression. *ERG1* was the predominant isoform expressed in cancerous tissue. (Paper I)

2. Several *ETS*-family members were differentially expressed in cancerous tissue compared with benign tissue. *ELF5*, *ETV1* and *ETS1* were alternatively overexpressed in patients without increased *ERG* expression. Their increased expression was inversely related to *ERG*s, suggesting that these *ETS* transcription factors might substitute for *ERG* in prostate cancer. (Paper I)

3. *ETS2* was moderately to strongly downregulated in most prostate cancers. *TMPRSS2*, *ERG* and *ETS2* are located on chromosome 21q22 and it is possible that the *ETS2* gene is lost when *TMPRSS2* is fused to *ERG*. Since ERG and ETS2 seem to have competing roles in transcription complexes, the loss of ETS2 could further promote ERG activity. (Paper I)

4. It is possible to detect *TMPRSS2:ERG* mRNA transcripts in urine from prostate cancer patients. Isoform a (89.5%) was most prevalent compared with isoform b. Prostatic massage prior to urine sampling greatly improved detection of the fusion transcripts and thereby increased the sensitivity of the test. (Paper III)

5. More than one *TMPRSS2:ERG* isoform may be detected in urine from one patient, reflecting the heterogeneity of prostate cancer. (Paper III)

6. The presence of *TMPRSS2:ERG* fusions in urine was significantly associated with adverse clinicopathological variables (preoperative s-PSA, Gleason score

and pathological stage). We were the first to show a positive association between fusion status and PSA levels. (Paper III)

7. The transcription factor *SIM2* was one of the most highly upregulated genes in prostate cancer, and both *SIM2* isoforms, *SIM2-long* and *SIM2-short*, were highly overexpressed in cancer tissue. (Paper II)

8. Expression of the SIM2-s protein in prostate cancer was significantly associated with adverse clinicopathological factors and reduced survival. SIM2 may be involved in the clinical progression of prostate cancer, and our findings support the proposal of SIM2 as a candidate for targeted therapy of prostate cancer. (Paper II)

REFERENCES

1. Cancer in Norway 2008. Cancer Registry of Norway. Institute of Populationbased Cancer Research, 2008
2. Gronberg H. Prostate cancer epidemiology. *Lancet*. 2003 Mar 8;361(9360):859-64.
3. Potosky AL, Miller BA, Albertsen PC, Kramer BS. The role of increasing detection in the rising incidence of prostate cancer. *Jama*. 1995 Feb 15;273(7):548-52.
4. Cancer in Norway 2005: Cancer Registry of Norway. Institute of Populationbased Cancer Research, 2005.
5. Weir HK, Thun MJ, Hankey BF, Ries LA, Howe HL, Wingo PA*, et al.* Annual report to the nation on the status of cancer, 1975-2000, featuring the uses of surveillance data for cancer prevention and control. *J Natl Cancer Inst*. 2003 Sep 3;95(17):1276-99.
6. Ries LAG EM, Kosary CL, et al. SEER cancer statistics review, 1973-1999. *Bethesda: National Cancer Institute*. 2002.
7. Carter BS, Bova GS, Beaty TH, Steinberg GD, Childs B, Isaacs WB*, et al.* Hereditary prostate cancer: epidemiologic and clinical features. *J Urol*. 1993 Sep;150(3):797-802.
8. Stanford JL, Ostrander EA. Familial prostate cancer. *Epidemiol Rev*. 2001;23(1):19-23.
9. Bratt O, Damber JE, Emanuelsson M, Gronberg H. Hereditary prostate cancer: clinical characteristics and survival. *J Urol*. 2002 Jun;167(6):2423-6.
10. Ostrander EA, Markianos K, Stanford JL. Finding prostate cancer susceptibility genes. *Annu Rev Genomics Hum Genet*. 2004;5:151-75.
11. Shand RL, Gelmann EP. Molecular biology of prostate-cancer pathogenesis. *Curr Opin Urol*. 2006 May;16(3):123-31.
12. Friedenson B. BRCA1 and BRCA2 pathways and the risk of cancers other than breast or ovarian. *MedGenMed*. 2005;7(2):60.
13. Rosen EM, Fan S, Goldberg ID. BRCA1 and prostate cancer. *Cancer Invest*. 2001;19(4):396-412.
14. Sigurdsson S, Thorlacius S, Tomasson J, Tryggvadottir L, Benediktsdottir K, Eyfjord JE*, et al.* BRCA2 mutation in Icelandic prostate cancer patients. *J Mol Med*. 1997 Oct;75(10):758-61.
15. Howell MA. Factor analysis of international cancer mortality data and per capita food consumption. *Br J Cancer*. 1974 Apr;29(4):328-36.
16. Armstrong B, Doll R. Environmental factors and cancer incidence and mortality in different countries, with special reference to dietary practices. *Int J Cancer*. 1975 Apr 15;15(4):617-31.
17. Hayes RB, Ziegler RG, Gridley G, Swanson C, Greenberg RS, Swanson GM*, et al.* Dietary factors and risks for prostate cancer among blacks and whites in the United States. *Cancer Epidemiol Biomarkers Prev*. 1999 Jan;8(1):25-34.
18. Chan JM, Stampfer MJ, Ma J, Gann PH, Gaziano JM, Giovannucci EL. Dairy products, calcium, and prostate cancer risk in the Physicians' Health Study. *Am J Clin Nutr*. 2001 Oct;74(4):549-54.
19. Bylund A, Zhang JX, Bergh A, Damber JE, Widmark A, Johansson A*, et al.* Rye bran and soy protein delay growth and increase apoptosis of human LNCaP prostate adenocarcinoma in nude mice. *Prostate*. 2000 Mar 1;42(4):304-14.
20. Giovannucci E, Rimm EB, Liu Y, Stampfer MJ, Willett WC. A prospective study of tomato products, lycopene, and prostate cancer risk. *J Natl Cancer Inst*. 2002 Mar 6;94(5):391-8.

21. Heinonen OP, Albanes D, Virtamo J, Taylor PR, Huttunen JK, Hartman AM*, et al.* Prostate cancer and supplementation with alpha-tocopherol and beta-carotene: incidence and mortality in a controlled trial. *J Natl Cancer Inst*. 1998 Mar 18;90(6):440-6.
22. Redman C, Scott JA, Baines AT, Basye JL, Clark LC, Calley C*, et al.* Inhibitory effect of selenomethionine on the growth of three selected human tumor cell lines. *Cancer Lett*. 1998 Mar 13;125(1-2):103-10.
23. Redman C, Xu MJ, Peng YM, Scott JA, Payne C, Clark LC*, et al.* Involvement of polyamines in selenomethionine induced apoptosis and mitotic alterations in human tumor cells. *Carcinogenesis*. 1997 Jun;18(6):1195-202.
24. Leitzmann MF, Stampfer MJ, Michaud DS, Augustsson K, Colditz GC, Willett WC*, et al.* Dietary intake of n-3 and n-6 fatty acids and the risk of prostate cancer. *Am J Clin Nutr*. 2004 Jul;80(1):204-16.
25. Stattin P, Lumme S, Tenkanen L, Alfthan H, Jellum E, Hallmans G*, et al.* High levels of circulating testosterone are not associated with increased prostate cancer risk: a pooled prospective study. *Int J Cancer*. 2004 Jan 20;108(3):418-24.
26. Arnold RS, Makarova NV, Osunkoya AO, Suppiah S, Scott TA, Johnson NA*, et al.* XMRV infection in patients with prostate cancer: novel serologic assay and correlation with PCR and FISH. *Urology*. 2010 Apr;75(4):755-61.
27. Rusmevichientong A, Chow SA. Biology and pathophysiology of the new human retrovirus XMRV and its association with human disease. *Immunol Res*. 2010 Aug 18.
28. Silverman RH, Nguyen C, Weight CJ, Klein EA. The human retrovirus XMRV in prostate cancer and chronic fatigue syndrome. *Nat Rev Urol*. 2010 Jul;7(7):392-402.
29. Urisman A, Molinaro RJ, Fischer N, Plummer SJ, Casey G, Klein EA*, et al.* Identification of a novel Gammaretrovirus in prostate tumors of patients homozygous for R462Q RNASEL variant. *PLoS Pathog*. 2006 Mar;2(3):e25.
30. Kirby RS CT, Brawer M. Prostate Cancer: Mosby, Times Mirror International Publishers; 1996.
31. McNeal JE. Regional morphology and pathology of the prostate. *Am J Clin Pathol*. 1968 Mar;49(3):347-57.
32. Reese JH, McNeal JE, Redwine EA, Samloff IM, Stamey TA. Differential distribution of pepsinogen II between the zones of the human prostate and the seminal vesicle. *J Urol*. 1986 Nov;136(5):1148-52.
33. Cunha GR, Donjacour A. Stromal-epithelial interactions in normal and abnormal prostatic development. *Prog Clin Biol Res*. 1987;239:251-72.
34. Vashchenko N, Abrahamsson PA. Neuroendocrine differentiation in prostate cancer: implications for new treatment modalities. *Eur Urol*. 2005 Feb;47(2):147-55.
35. Goldstein AS, Stoyanova T, Witte ON. Primitive origins of prostate cancer: In vivo evidence for prostate-regenerating cells and prostate cancer-initiating cells. *Mol Oncol*. 2010 Jul 14.
36. Bonkhoff H, Stein U, Remberger K. The proliferative function of basal cells in the normal and hyperplastic human prostate. *Prostate*. 1994;24(3):114-8.
37. Bonkhoff H, Stein U, Remberger K. Multidirectional differentiation in the normal, hyperplastic, and neoplastic human prostate: simultaneous demonstration of cell-specific epithelial markers. *Hum Pathol*. 1994 Jan;25(1):42-6.
38. McNeal JE, Redwine EA, Freiha FS, Stamey TA. Zonal distribution of prostatic adenocarcinoma. Correlation with histologic pattern and direction of spread. *Am J Surg Pathol*. 1988 Dec;12(12):897-906.
39. Beheshti B, Vukovic B, Marrano P, Squire JA, Park PC. Resolution of genotypic heterogeneity in prostate tumors using polymerase chain reaction and comparative

genomic hybridization on microdissected carcinoma and prostatic intraepithelial neoplasia foci. *Cancer Genet Cytogenet*. 2002 Aug;137(1):15-22.
40. Sakr WA, Grignon DJ, Crissman JD, Heilbrun LK, Cassin BJ, Pontes JJ*, et al.* High grade prostatic intraepithelial neoplasia (HGPIN) and prostatic adenocarcinoma between the ages of 20-69: an autopsy study of 249 cases. *In Vivo*. 1994 May-Jun;8(3):439-43.
41. Sakr WA, Haas GP, Cassin BF, Pontes JE, Crissman JD. The frequency of carcinoma and intraepithelial neoplasia of the prostate in young male patients. *J Urol*. 1993 Aug;150(2 Pt 1):379-85.
42. Bishara T, Ramnani DM, Epstein JI. High-grade prostatic intraepithelial neoplasia on needle biopsy: risk of cancer on repeat biopsy related to number of involved cores and morphologic pattern. *Am J Surg Pathol*. 2004 May;28(5):629-33.
43. Qian J, Wollan P, Bostwick DG. The extent and multicentricity of high-grade prostatic intraepithelial neoplasia in clinically localized prostatic adenocarcinoma. *Hum Pathol*. 1997 Feb;28(2):143-8.
44. Cheville JC, Bostwick DG. Postatrophic hyperplasia of the prostate. A histologic mimic of prostatic adenocarcinoma. *Am J Surg Pathol*. 1995 Sep;19(9):1068-76.
45. Ruska KM, Sauvageot J, Epstein JI. Histology and cellular kinetics of prostatic atrophy. *Am J Surg Pathol*. 1998 Sep;22(9):1073-7.
46. De Marzo AM, Nakai Y, Nelson WG. Inflammation, atrophy, and prostate carcinogenesis. *Urol Oncol*. 2007 Sep-Oct;25(5):398-400.
47. De Marzo AM, Platz EA, Sutcliffe S, Xu J, Gronberg H, Drake CG*, et al.* Inflammation in prostate carcinogenesis. *Nat Rev Cancer*. 2007 Apr;7(4):256-69.
48. van Leenders GJ, Gage WR, Hicks JL, van Balken B, Aalders TW, Schalken JA*, et al.* Intermediate cells in human prostate epithelium are enriched in proliferative inflammatory atrophy. *Am J Pathol*. 2003 May;162(5):1529-37.
49. Lawson DA, Zong Y, Memarzadeh S, Xin L, Huang J, Witte ON. Basal epithelial stem cells are efficient targets for prostate cancer initiation. *Proc Natl Acad Sci U S A*. 2010 Feb 9;107(6):2610-5.
50. Miki J, Rhim JS. Prostate cell cultures as in vitro models for the study of normal stem cells and cancer stem cells. *Prostate Cancer Prostatic Dis*. 2008;11(1):32-9.
51. Goldstein AS, Huang J, Guo C, Garraway IP, Witte ON. Identification of a cell of origin for human prostate cancer. *Science*. 2010 Jul 30;329(5991):568-71.
52. Crum CP, McKeon FD. p63 in epithelial survival, germ cell surveillance, and neoplasia. *Annu Rev Pathol*. 2010;5:349-71.
53. Gleason DF. Classification of prostatic carcinomas. *Cancer Chemother Rep*. 1966 Mar;50(3):125-8.
54. Epstein JI. An update of the Gleason grading system. *J Urol*. 2010 Feb;183(2):433-40.
55. Epstein JI, Allsbrook WC, Jr., Amin MB, Egevad LL. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. *Am J Surg Pathol*. 2005 Sep;29(9):1228-42.
56. Gleason DF. Histologic grading and clinical staging of prostatic carcinoma. In: Urologic pathology: The Prostate. Philadelphia: Leah & Febiger; 1977.
57. Epstein JI, Amin M, Boccon-Gibod L, Egevad L, Humphrey PA, Mikuz G*, et al.* Prognostic factors and reporting of prostate carcinoma in radical prostatectomy and pelvic lymphadenectomy specimens. *Scand J Urol Nephrol Suppl*. 2005 May(216):34-63.
58. Humphrey PA. Gleason grading and prognostic factors in carcinoma of the prostate. *Mod Pathol*. 2004 Mar;17(3):292-306.

59. Bostwick DG. Staging prostate cancer--1997: current methods and limitations. *Eur Urol*. 1997;32 Suppl 3:2-14.
60. Hoedemaeker RF, Vis AN, Van Der Kwast TH. Staging prostate cancer. *Microsc Res Tech*. 2000 Dec 1;51(5):423-9.
61. Leprince D, Gegonne A, Coll J, de Taisne C, Schneeberger A, Lagrou C*, et al.* A putative second cell-derived oncogene of the avian leukaemia retrovirus E26. *Nature*. 1983 Nov 24-30;306(5941):395-7.
62. Sharrocks AD. The ETS-domain transcription factor family. *Nat Rev Mol Cell Biol*. 2001 Nov;2(11):827-37.
63. Graves BJ, Petersen JM. Specificity within the ets family of transcription factors. *Adv Cancer Res*. 1998;75:1-55.
64. Gutierrez-Hartmann A, Duval DL, Bradford AP. ETS transcription factors in endocrine systems. *Trends Endocrinol Metab*. 2007 May-Jun;18(4):150-8.
65. Slupsky CM, Gentile LN, Donaldson LW, Mackereth CD, Seidel JJ, Graves BJ*, et al.* Structure of the Ets-1 pointed domain and mitogen-activated protein kinase phosphorylation site. *Proc Natl Acad Sci U S A*. 1998 Oct 13;95(21):12129-34.
66. Kim CA, Phillips ML, Kim W, Gingery M, Tran HH, Robinson MA*, et al.* Polymerization of the SAM domain of TEL in leukemogenesis and transcriptional repression. *EMBO J*. 2001 Aug 1;20(15):4173-82.
67. Sementchenko VI, Watson DK. Ets target genes: past, present and future. *Oncogene*. 2000 Dec 18;19(55):6533-48.
68. Bassuk AG, Leiden JM. The role of Ets transcription factors in the development and function of the mammalian immune system. *Adv Immunol*. 1997;64:65-104.
69. Dittmer J, Nordheim A. Ets transcription factors and human disease. *Biochim Biophys Acta*. 1998 Apr 17;1377(2):F1-11.
70. Wasylyk B, Hagman J, Gutierrez-Hartmann A. Ets transcription factors: nuclear effectors of the Ras-MAP-kinase signaling pathway. *Trends Biochem Sci*. 1998 Jun;23(6):213-6.
71. Oikawa T, Yamada T. Molecular biology of the Ets family of transcription factors. *Gene*. 2003 Jan 16;303:11-34.
72. Seth A, Watson DK. ETS transcription factors and their emerging roles in human cancer. *Eur J Cancer*. 2005 Nov;41(16):2462-78.
73. Basuyaux JP, Ferreira E, Stehelin D, Buttice G. The Ets transcription factors interact with each other and with the c-Fos/c-Jun complex via distinct protein domains in a DNA-dependent and -independent manner. *J Biol Chem*. 1997 Oct 17;272(42):26188-95.
74. Yang L, Mei Q, Zielinska-Kwiatkowska A, Matsui Y, Blackburn ML, Benedetti D*, et al.* An ERG (ets-related gene)-associated histone methyltransferase interacts with histone deacetylases 1/2 and transcription co-repressors mSin3A/B. *Biochem J*. 2003 Feb 1;369(Pt 3):651-7.
75. Buttice G, Duterque-Coquillaud M, Basuyaux JP, Carrere S, Kurkinen M, Stehelin D. Erg, an Ets-family member, differentially regulates human collagenase1 (MMP1) and stromelysin1 (MMP3) gene expression by physically interacting with the Fos/Jun complex. *Oncogene*. 1996 Dec 5;13(11):2297-306.
76. Yordy JS, Muise-Helmericks RC. Signal transduction and the Ets family of transcription factors. *Oncogene*. 2000 Dec 18;19(55):6503-13.
77. Sieweke MH, Tekotte H, Frampton J, Graf T. MafB is an interaction partner and repressor of Ets-1 that inhibits erythroid differentiation. *Cell*. 1996 Apr 5;85(1):49-60.

78. Petrovics G, Liu A, Shaheduzzaman S, Furasato B, Sun C, Chen Y*, et al.* Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome. *Oncogene*. 2005 May 26;24(23):3847-52.
79. Rostad K, Mannelqvist M, Halvorsen OJ, Oyan AM, Bo TH, Stordrange L*, et al.* ERG upregulation and related ETS transcription factors in prostate cancer. *Int J Oncol*. 2007 Jan;30(1):19-32.
80. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW*, et al.* Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science*. 2005 Oct 28;310(5748):644-8.
81. Helgeson BE, Tomlins SA, Shah N, Laxman B, Cao Q, Prensner JR*, et al.* Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. *Cancer Res*. 2008 Jan 1;68(1):73-80.
82. Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell*. 2000 Jan 7;100(1):57-70.
83. Lelievre E, Lionneton F, Mattot V, Spruyt N, Soncin F. Ets-1 regulates fli-1 expression in endothelial cells. Identification of ETS binding sites in the fli-1 gene promoter. *J Biol Chem*. 2002 Jul 12;277(28):25143-51.
84. Teruyama K, Abe M, Nakano T, Iwasaka-Yagi C, Takahashi S, Yamada S*, et al.* Role of transcription factor Ets-1 in the apoptosis of human vascular endothelial cells. *J Cell Physiol*. 2001 Aug;188(2):243-52.
85. Feldman RJ, Sementchenko VI, Gayed M, Fraig MM, Watson DK. Pdef expression in human breast cancer is correlated with invasive potential and altered gene expression. *Cancer Res*. 2003 Aug 1;63(15):4626-31.
86. Lavenburg KR, Ivey J, Hsu T, Muise-Helmericks RC. Coordinated functions of Akt/PKB and ETS1 in tubule formation. *FASEB J*. 2003 Dec;17(15):2278-80.
87. Elvert G, Kappel A, Heidenreich R, Englmeier U, Lanz S, Acker T*, et al.* Cooperative interaction of hypoxia-inducible factor-2alpha (HIF-2alpha) and Ets-1 in the transcriptional activation of vascular endothelial growth factor receptor-2 (Flk-1). *J Biol Chem*. 2003 Feb 28;278(9):7520-30.
88. Sevilla L, Zaldumbide A, Carlotti F, Dayem MA, Pognonec P, Boulukos KE. Bcl-XL expression correlates with primary macrophage differentiation, activation of functional competence, and survival and results from synergistic transcriptional activation by Ets2 and PU.1. *J Biol Chem*. 2001 May 25;276(21):17800-7.
89. Yi H, Fujimura Y, Ouchida M, Prasad DD, Rao VN, Reddy ES. Inhibition of apoptosis by normal and aberrant Fli-1 and erg proteins involved in human solid tumors and leukemias. *Oncogene*. 1997 Mar 20;14(11):1259-68.
90. Li R, Pei H, Papas T. The p42 variant of ETS1 protein rescues defective Fas-induced apoptosis in colon carcinoma cells. *Proc Natl Acad Sci U S A*. 1999 Mar 30;96(7):3876-81.
91. Foos G, Hauser CA. Altered Ets transcription factor activity in prostate tumor cells inhibits anchorage-independent growth, survival, and invasiveness. *Oncogene*. 2000 Nov 16;19(48):5507-16.
92. Kavurma MM, Bobryshev Y, Khachigian LM. Ets-1 positively regulates Fas ligand transcription via cooperative interactions with Sp1. *J Biol Chem*. 2002 Sep 27;277(39):36244-52.
93. Liu W, Wang G, Yakovlev AG. Identification and functional analysis of the rat caspase-3 gene promoter. *J Biol Chem*. 2002 Mar 8;277(10):8273-8.
94. Tamir A, Howard J, Higgins RR, Li YJ, Berger L, Zacksenhaus E*, et al.* Fli-1, an Ets-related transcription factor, regulates erythropoietin-induced erythroid proliferation and differentiation: evidence for direct transcriptional repression of the Rb gene during differentiation. *Mol Cell Biol*. 1999 Jun;19(6):4452-64.

95. Lesault I, Quang CT, Frampton J, Ghysdael J. Direct regulation of BCL-2 by FLI-1 is involved in the survival of FLI-1-transformed erythroblasts. *EMBO J*. 2002 Feb 15;21(4):694-703.
96. Mueller BU, Pabst T, Osato M, Asou N, Johansen LM, Minden MD*, et al.* Heterozygous PU.1 mutations are associated with acute myeloid leukemia. *Blood*. 2002 Aug 1;100(3):998-1007.
97. Vangala RK, Heiss-Neumann MS, Rangatia JS, Singh SM, Schoch C, Tenen DG*, et al.* The myeloid master regulator transcription factor PU.1 is inactivated by AML1-ETO in t(8;21) myeloid leukemia. *Blood*. 2003 Jan 1;101(1):270-7.
98. Golub TR, Barker GF, Stegmaier K, Gilliland DG. The TEL gene contributes to the pathogenesis of myeloid and lymphoid leukemias by diverse molecular genetic mechanisms. *Curr Top Microbiol Immunol*. 1997;220:67-79.
99. Zelent A, Greaves M, Enver T. Role of the TEL-AML1 fusion gene in the molecular pathogenesis of childhood acute lymphoblastic leukaemia. *Oncogene*. 2004 May 24;23(24):4275-83.
100. Dastugue N, Lafage-Pochitaloff M, Pages MP, Radford I, Bastard C, Talmant P*, et al.* Cytogenetic profile of childhood and adult megakaryoblastic leukemia (M7): a study of the Groupe Francais de Cytogenetique Hematologique (GFCH). *Blood*. 2002 Jul 15;100(2):618-26.
101. Shimizu K, Ichikawa H, Tojo A, Kaneko Y, Maseki N, Hayashi Y*, et al.* An ets-related gene, ERG, is rearranged in human myeloid leukemia with t(16;21) chromosomal translocation. *Proc Natl Acad Sci U S A*. 1993 Nov 1;90(21):10280-4.
102. Baldus CD, Liyanarachchi S, Mrozek K, Auer H, Tanner SM, Guimond M*, et al.* Acute myeloid leukemia with complex karyotypes and abnormal chromosome 21: Amplification discloses overexpression of APP, ETS2, and ERG genes. *Proc Natl Acad Sci U S A*. 2004 Mar 16;101(11):3915-20.
103. Golub TR, Barker GF, Lovett M, Gilliland DG. Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. *Cell*. 1994 Apr 22;77(2):307-16.
104. Golub TR, Goga A, Barker GF, Afar DE, McLaughlin J, Bohlander SK*, et al.* Oligomerization of the ABL tyrosine kinase by the Ets protein TEL in human leukemia. *Mol Cell Biol*. 1996 Aug;16(8):4107-16.
105. Lacronique V, Boureux A, Valle VD, Poirel H, Quang CT, Mauchauffe M*, et al.* A TEL-JAK2 fusion protein with constitutive kinase activity in human leukemia. *Science*. 1997 Nov 14;278(5341):1309-12.
106. Knezevich SR, McFadden DE, Tao W, Lim JF, Sorensen PH. A novel ETV6-NTRK3 gene fusion in congenital fibrosarcoma. *Nat Genet*. 1998 Feb;18(2):184-7.
107. Iijima Y, Ito T, Oikawa T, Eguchi M, Eguchi-Ishimae M, Kamada N*, et al.* A new ETV6/TEL partner gene, ARG (ABL-related gene or ABL2), identified in an AML-M3 cell line with a t(1;12)(q25;p13) translocation. *Blood*. 2000 Mar 15;95(6):2126-31.
108. Golub TR, McLean T, Stegmaier K, Carroll M, Tomasson M, Gilliland DG. The TEL gene and human leukemia. *Biochim Biophys Acta*. 1996 Aug 8;1288(1):M7-10.
109. Peter M, Couturier J, Pacquement H, Michon J, Thomas G, Magdelenat H*, et al.* A new member of the ETS family fused to EWS in Ewing tumors. *Oncogene*. 1997 Mar 13;14(10):1159-64.
110. Delattre O, Zucman J, Plougastel B, Desmaze C, Melot T, Peter M*, et al.* Gene fusion with an ETS DNA-binding domain caused by chromosome translocation in human tumours. *Nature*. 1992 Sep 10;359(6391):162-5.
111. Delattre O. [Ewing's tumours, genetic and cellular aspects]. *Pathol Biol (Paris)*. 2008 Jul;56(5):257-9.

112. Bailly RA, Bosselut R, Zucman J, Cormier F, Delattre O, Roussel M*, et al.* DNA-binding and transcriptional activation properties of the EWS-FLI-1 fusion protein resulting from the t(11;22) translocation in Ewing sarcoma. *Mol Cell Biol*. 1994 May;14(5):3230-41.
113. Lin PP, Brody RI, Hamelin AC, Bradner JE, Healey JH, Ladanyi M. Differential transactivation by alternative EWS-FLI1 fusion proteins correlates with clinical heterogeneity in Ewing's sarcoma. *Cancer Res*. 1999 Apr 1;59(7):1428-32.
114. Dauphinot L, De Oliveira C, Melot T, Sevenet N, Thomas V, Weissman BE*, et al.* Analysis of the expression of cell cycle regulators in Ewing cell lines: EWS-FLI-1 modulates p57KIP2and c-Myc expression. *Oncogene*. 2001 May 31;20(25):3258-65.
115. Mastrangelo T, Modena P, Tornielli S, Bullrich F, Testi MA, Mezzelani A*, et al.* A novel zinc finger gene is fused to EWS in small round cell tumor. *Oncogene*. 2000 Aug 3;19(33):3799-804.
116. Reddy ES, Rao VN, Papas TS. The erg gene: a human gene related to the ets oncogene. *Proc Natl Acad Sci U S A*. 1987 Sep;84(17):6131-5.
117. Rao VN, Papas TS, Reddy ES. erg, a human ets-related gene on chromosome 21: alternative splicing, polyadenylation, and translation. *Science*. 1987 Aug 7;237(4815):635-9.
118. Duterque-Coquillaud M, Niel C, Plaza S, Stehelin D. New human erg isoforms generated by alternative splicing are transcriptional activators. *Oncogene*. 1993 Jul;8(7):1865-73.
119. Prasad DD, Rao VN, Lee L, Reddy ES. Differentially spliced erg-3 product functions as a transcriptional activator. *Oncogene*. 1994 Feb;9(2):669-73.
120. Carrere S, Verger A, Flourens A, Stehelin D, Duterque-Coquillaud M. Erg proteins, transcription factors of the Ets family, form homo, heterodimers and ternary complexes via two distinct domains. *Oncogene*. 1998 Jun 25;16(25):3261-8.
121. Owczarek CM, Portbury KJ, Hardy MP, O'Leary DA, Kudoh J, Shibuya K*, et al.* Detailed mapping of the ERG-ETS2 interval of human chromosome 21 and comparison with the region of conserved synteny on mouse chromosome 16. *Gene*. 2004 Jan 7;324:65-77.
122. Maroulakou IG, Bowe DB. Expression and function of Ets transcription factors in mammalian development: a regulatory network. *Oncogene*. 2000 Dec 18;19(55):6432-42.
123. Vlaeminck-Guillem V, Carrere S, Dewitte F, Stehelin D, Desbiens X, Duterque-Coquillaud M. The Ets family member Erg gene is expressed in mesodermal tissues and neural crests at fundamental steps during mouse embryogenesis. *Mech Dev*. 2000 Mar 1;91(1-2):331-5.
124. Hewett PW, Nishi K, Daft EL, Clifford Murray J. Selective expression of erg isoforms in human endothelial cells. *Int J Biochem Cell Biol*. 2001 Apr;33(4):347-55.
125. Yuan L, Nikolova-Krstevski V, Zhan Y, Kondo M, Bhasin M, Varghese L*, et al.* Antiinflammatory effects of the ETS factor ERG in endothelial cells are mediated through transcriptional repression of the interleukin-8 gene. *Circ Res*. 2009 May 8;104(9):1049-57.
126. McLaughlin F, Ludbrook VJ, Cox J, von Carlowitz I, Brown S, Randi AM. Combined genomic and antisense analysis reveals that the transcription factor Erg is implicated in endothelial cell differentiation. *Blood*. 2001 Dec 1;98(12):3332-9.
127. Oettgen P. Regulation of vascular inflammation and remodeling by ETS factors. *Circ Res*. 2006 Nov 24;99(11):1159-66.

128. Rao VN, Modi WS, Drabkin HD, Patterson D, O'Brien SJ, Papas TS*, et al.* The human erg gene maps to chromosome 21, band q22: relationship to the 8; 21 translocation of acute myelogenous leukemia. *Oncogene*. 1988 Nov;3(5):497-500.
129. Ichikawa H, Shimizu K, Hayashi Y, Ohki M. An RNA-binding protein gene, TLS/FUS, is fused to ERG in human myeloid leukemia with t(16;21) chromosomal translocation. *Cancer Res*. 1994 Jun 1;54(11):2865-8.
130. Shing DC, McMullan DJ, Roberts P, Smith K, Chin SF, Nicholson J*, et al.* FUS/ERG gene fusions in Ewing's tumors. *Cancer Res*. 2003 Aug 1;63(15):4568-76.
131. Kumar-Sinha C, Tomlins SA, Chinnaiyan AM. Recurrent gene fusions in prostate cancer. *Nat Rev Cancer*. 2008 Jul;8(7):497-511.
132. Mehra R, Tomlins SA, Shen R, Nadeem O, Wang L, Wei JT*, et al.* Comprehensive assessment of TMPRSS2 and ETS family gene aberrations in clinically localized prostate cancer. *Mod Pathol*. 2007 May;20(5):538-44.
133. Nam RK, Sugar L, Wang Z, Yang W, Kitching R, Klotz LH*, et al.* Expression of TMPRSS2:ERG gene fusion in prostate cancer cells is an important prognostic factor for cancer progression. *Cancer Biol Ther*. 2007 Jan;6(1):40-5.
134. Nam RK, Sugar L, Yang W, Srivastava S, Klotz LH, Yang LY*, et al.* Expression of the TMPRSS2:ERG fusion gene predicts cancer recurrence after surgery for localised prostate cancer. *Br J Cancer*. 2007 Dec 17;97(12):1690-5.
135. Attard G, Clark J, Ambroisine L, Fisher G, Kovacs G, Flohr P*, et al.* Duplication of the fusion of TMPRSS2 to ERG sequences identifies fatal human prostate cancer. *Oncogene*. 2008 Jan 10;27(3):253-63.
136. Demichelis F, Fall K, Perner S, Andren O, Schmidt F, Setlur SR*, et al.* TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. *Oncogene*. 2007 Jul 5;26(31):4596-9.
137. Perner S, Demichelis F, Beroukhim R, Schmidt FH, Mosquera JM, Setlur S*, et al.* TMPRSS2:ERG Fusion-Associated Deletions Provide Insight into the Heterogeneity of Prostate Cancer. *Cancer Res*. 2006 Sep 1;66(17):8337-41.
138. Rajput AB, Miller MA, De Luca A, Boyd N, Leung S, Hurtado-Coll A*, et al.* Frequency of the TMPRSS2:ERG gene fusion is increased in moderate to poorly differentiated prostate cancers. *J Clin Pathol*. 2007 Nov;60(11):1238-43.
139. Winnes M, Lissbrant E, Damber JE, Stenman G. Molecular genetic analyses of the TMPRSS2-ERG and TMPRSS2-ETV1 gene fusions in 50 cases of prostate cancer. *Oncol Rep*. 2007 May;17(5):1033-6.
140. Yoshimoto M, Joshua AM, Chilton-Macneill S, Bayani J, Selvarajah S, Evans AJ*, et al.* Three-color FISH analysis of TMPRSS2/ERG fusions in prostate cancer indicates that genomic microdeletion of chromosome 21 is associated with rearrangement. *Neoplasia*. 2006 Jun;8(6):465-9.
141. Lapointe J, Kim YH, Miller MA, Li C, Kaygusuz G, van de Rijn M*, et al.* A variant TMPRSS2 isoform and ERG fusion product in prostate cancer with implications for molecular diagnosis. *Mod Pathol*. 2007 Apr;20(4):467-73.
142. Chen H, Chrast R, Rossier C, Gos A, Antonarakis SE, Kudoh J*, et al.* Single-minded and Down syndrome? *Nat Genet*. 1995 May;10(1):9-10.
143. Rahmani Z, Blouin JL, Creau-Goldberg N, Watkins PC, Mattei JF, Poissonnier M*, et al.* Critical role of the D21S55 region on chromosome 21 in the pathogenesis of Down syndrome. *Proc Natl Acad Sci U S A*. 1989 Aug;86(15):5958-62.
144. Chrast R, Scott HS, Chen H, Kudoh J, Rossier C, Minoshima S*, et al.* Cloning of two human homologs of the Drosophila single-minded gene SIM1 on chromosome 6q and SIM2 on 21q within the Down syndrome chromosomal region. *Genome Res*. 1997 Jun;7(6):615-24.

145. Goshu E, Jin H, Lovejoy J, Marion JF, Michaud JL, Fan CM. Sim2 contributes to neuroendocrine hormone gene expression in the anterior hypothalamus. *Mol Endocrinol*. 2004 May;18(5):1251-62.
146. Moffett P, Reece M, Pelletier J. The murine Sim-2 gene product inhibits transcription by active repression and functional interference. *Mol Cell Biol*. 1997 Sep;17(9):4933-47.
147. Rachidi M, Lopes C, Charron G, Delezoide AL, Paly E, Bloch B*, et al.* Spatial and temporal localization during embryonic and fetal human development of the transcription factor SIM2 in brain regions altered in Down syndrome. *Int J Dev Neurosci*. 2005 Aug;23(5):475-84.
148. Deyoung MP, Scheurle D, Damania H, Zylberberg C, Narayanan R. Down's syndrome-associated single minded gene as a novel tumor marker. *Anticancer Res*. 2002 Nov-Dec;22(6A):3149-57.
149. Halvorsen OJ, Oyan AM, Bo TH, Olsen S, Rostad K, Haukaas SA*, et al.* Gene expression profiles in prostate cancer: association with patient subgroups and tumour differentiation. *Int J Oncol*. 2005 Feb;26(2):329-36.
150. Halvorsen OJ, Rostad K, Oyan AM, Puntervoll H, Bo TH, Stordrange L*, et al.* Increased expression of SIM2-s protein is a novel marker of aggressive prostate cancer. *Clin Cancer Res*. 2007 Feb 1;13(3):892-7.
151. Hankinson O. The aryl hydrocarbon receptor complex. *Annu Rev Pharmacol Toxicol*. 1995;35:307-40.
152. Taylor BL, Zhulin IB. PAS domains: internal sensors of oxygen, redox potential, and light. *Microbiol Mol Biol Rev*. 1999 Jun;63(2):479-506.
153. Probst MR, Fan CM, Tessier-Lavigne M, Hankinson O. Two murine homologs of the Drosophila single-minded protein that interact with the mouse aryl hydrocarbon receptor nuclear translocator protein. *J Biol Chem*. 1997 Feb 14;272(7):4451-7.
154. Matikainen TM, Moriyama T, Morita Y, Perez GI, Korsmeyer SJ, Sherr DH*, et al.* Ligand activation of the aromatic hydrocarbon receptor transcription factor drives Bax-dependent apoptosis in developing fetal ovarian germ cells. *Endocrinology*. 2002 Feb;143(2):615-20.
155. DeYoung MP, Tress M, Narayanan R. Down's syndrome-associated Single Minded 2 gene as a pancreatic cancer drug therapy target. *Cancer Lett*. 2003 Oct 8;200(1):25-31.
156. DeYoung MP, Tress M, Narayanan R. Identification of Down's syndrome critical locus gene SIM2-s as a drug therapy target for solid tumors. *Proc Natl Acad Sci U S A*. 2003 Apr 15;100(8):4760-5.
157. Kwak HI, Gustafson T, Metz RP, Laffin B, Schedin P, Porter WW. Inhibition of breast cancer growth and invasion by single-minded 2s. *Carcinogenesis*. 2007 Feb;28(2):259-66.
158. Laffin B, Wellberg E, Kwak HI, Burghardt RC, Metz RP, Gustafson T*, et al.* Loss of singleminded-2s in the mouse mammary gland induces an epithelial-mesenchymal transition associated with up-regulation of slug and matrix metalloprotease 2. *Mol Cell Biol*. 2008 Mar;28(6):1936-46.
159. He Q, Li G, Su Y, Shen J, Liu Q, Ma X*, et al.* Single minded 2-s (SIM2-s) gene is expressed in human GBM cells and involved in GBM invasion. *Cancer Biol Ther*. Mar 15;9(6):430-6.
160. Ashworth A. A case of cancer in which cells similar to those in the tumours were seen in the blood after death. *Aust Med J*. 1869;14:146.
161. Duffy MJ. Can molecular markers now be used for early diagnosis of malignancy? *Clin Chem*. 1995 Oct;41(10):1410-3.

162. van Gils MP, Stenman UH, Schalken JA, Schroder FH, Luider TM, Lilja H*, et al.* Innovations in serum and urine markers in prostate cancer current European research in the P-Mark project. *Eur Urol*. 2005 Dec;48(6):1031-41.
163. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Reporting recommendations for tumor marker prognostic studies (REMARK). *J Natl Cancer Inst*. 2005 Aug 17;97(16):1180-4.
164. Pepe MS, Etzioni R, Feng Z, Potter JD, Thompson ML, Thornquist M*, et al.* Phases of biomarker development for early detection of cancer. *J Natl Cancer Inst*. 2001 Jul 18;93(14):1054-61.
165. Ablin RJ, Bronson P, Soanes WA, Witebsky E. Tissue- and species-specific antigens of normal human prostatic tissue. *J Immunol*. 1970 Jun;104(6):1329-39.
166. Ablin RJ, Soanes WA, Bronson P, Witebsky E. Precipitating antigens of the normal human prostate. *J Reprod Fertil*. 1970 Aug;22(3):573-4.
167. Reynolds MA, Kastury K, Groskopf J, Schalken JA, Rittenhouse H. Molecular markers for prostate cancer. *Cancer Lett*. 2007 Apr 28;249(1):5-13.
168. Sardana G, Dowell B, Diamandis EP. Emerging biomarkers for the diagnosis and prognosis of prostate cancer. *Clin Chem*. 2008 Dec;54(12):1951-60.
169. Steuber T, Helo P, Lilja H. Circulating biomarkers for prostate cancer. *World J Urol*. 2007 Apr;25(2):111-9.
170. Sorensen KD, Orntoft TF. Discovery of prostate cancer biomarkers by microarray gene expression profiling. *Expert Rev Mol Diagn*. 2010 Jan;10(1):49-64.
171. Stamey TA. Preoperative serum prostate-specific antigen (PSA) below 10 microg/l predicts neither the presence of prostate cancer nor the rate of postoperative PSA failure. *Clin Chem*. 2001 Apr;47(4):631-4.
172. Thompson IM, Pauler DK, Goodman PJ, Tangen CM, Lucia MS, Parnes HL*, et al.* Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter. *N Engl J Med*. 2004 May 27;350(22):2239-46.
173. Freedland SJ, Humphreys EB, Mangold LA, Eisenberger M, Dorey FJ, Walsh PC*, et al.* Risk of prostate cancer-specific mortality following biochemical recurrence after radical prostatectomy. *JAMA*. 2005 Jul 27;294(4):433-9.
174. Andriole GL, Crawford ED, Grubb RL, 3rd, Buys SS, Chia D, Church TR*, et al.* Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med*. 2009 Mar 26;360(13):1310-9.
175. Schroder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V*, et al.* Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med*. 2009 Mar 26;360(13):1320-8.
176. Lilja H, Ulmert D, Bjork T, Becker C, Serio AM, Nilsson JA*, et al.* Long-term prediction of prostate cancer up to 25 years before diagnosis of prostate cancer using prostate kallikreins measured at age 44 to 50 years. *J Clin Oncol*. 2007 Feb 1;25(4):431-6.
177. Aihara M, Lebovitz RM, Wheeler TM, Kinner BM, Ohori M, Scardino PT. Prostate specific antigen and gleason grade: an immunohistochemical study of prostate cancer. *J Urol*. 1994 Jun;151(6):1558-64.
178. Partin AW, Yoo J, Carter HB, Pearson JD, Chan DW, Epstein JI*, et al.* The use of prostate specific antigen, clinical stage and Gleason score to predict pathological stage in men with localized prostate cancer. *J Urol*. 1993 Jul;150(1):110-4.
179. Kattan MW, Eastham JA, Stapleton AM, Wheeler TM, Scardino PT. A preoperative nomogram for disease recurrence following radical prostatectomy for prostate cancer. *J Natl Cancer Inst*. 1998 May 20;90(10):766-71.

180. Stephenson AJ, Scardino PT, Eastham JA, Bianco FJ, Jr., Dotan ZA, Fearn PA*, et al.* Preoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy. *J Natl Cancer Inst*. 2006 May 17;98(10):715-7.
181. Oesterling JE. Age-specific reference ranges for serum PSA. *N Engl J Med*. 1996 Aug 1;335(5):345-6.
182. Catalona WJ, Beiser JA, Smith DS. Serum free prostate specific antigen and prostate specific antigen density measurements for predicting cancer in men with prior negative prostatic biopsies. *J Urol*. 1997 Dec;158(6):2162-7.
183. Catalona WJ, Richie JP, deKernion JB, Ahmann FR, Ratliff TL, Dalkin BL*, et al.* Comparison of prostate specific antigen concentration versus prostate specific antigen density in the early detection of prostate cancer: receiver operating characteristic curves. *J Urol*. 1994 Dec;152(6 Pt 1):2031-6.
184. D'Amico AV, Chen MH, Roehl KA, Catalona WJ. Preoperative PSA velocity and the risk of death from prostate cancer after radical prostatectomy. *N Engl J Med*. 2004 Jul 8;351(2):125-35.
185. Chun FK, Briganti A, Graefen M, Porter C, Montorsi F, Haese A*, et al.* Development and external validation of an extended repeat biopsy nomogram. *J Urol*. 2007 Feb;177(2):510-5.
186. Karakiewicz PI, Benayoun S, Kattan MW, Perrotte P, Valiquette L, Scardino PT*, et al.* Development and validation of a nomogram predicting the outcome of prostate biopsy based on patient age, digital rectal examination and serum prostate specific antigen. *J Urol*. 2005 Jun;173(6):1930-4.
187. Brawer MK, Meyer GE, Letran JL, Bankson DD, Morris DL, Yeung KK*, et al.* Measurement of complexed PSA improves specificity for early detection of prostate cancer. *Urology*. 1998 Sep;52(3):372-8.
188. Parsons JK, Brawer MK, Cheli CD, Partin AW, Djavan R. Complexed prostate specific antigen (PSA) reduces unnecessary prostate biopsies in the 2.6-4.0 ng/mL range of total PSA. *BJU Int*. 2004 Jul;94(1):47-50.
189. Zhang WM, Finne P, Leinonen J, Salo J, Stenman UH. Determination of prostate-specific antigen complexed to alpha(2)-macroglobulin in serum increases the specificity of free to total PSA for prostate cancer. *Urology*. 2000 Aug 1;56(2):267-72.
190. Zhu L, Leinonen J, Zhang WM, Finne P, Stenman UH. Dual-label immunoassay for simultaneous measurement of prostate-specific antigen (PSA)-alpha1-antichymotrypsin complex together with free or total PSA. *Clin Chem*. 2003 Jan;49(1):97-103.
191. Halvorsen OJ, Haukaas SA, Akslen LA. Combined loss of PTEN and p27 expression is associated with tumor cell proliferation by Ki-67 and increased risk of recurrent disease in localized prostate cancer. *Clin Cancer Res*. 2003 Apr;9(4):1474-9.
192. Halvorsen OJ, Hostmark J, Haukaas S, Hoisaeter PA, Akslen LA. Prognostic significance of p16 and CDK4 proteins in localized prostate carcinoma. *Cancer*. 2000 Jan 15;88(2):416-24.
193. Bachmann IM, Halvorsen OJ, Collett K, Stefansson IM, Straume O, Haukaas SA*, et al.* EZH2 expression is associated with high proliferation rate and aggressive tumor subgroups in cutaneous melanoma and cancers of the endometrium, prostate, and breast. *J Clin Oncol*. 2006 Jan 10;24(2):268-73.
194. Gravdal K, Halvorsen OJ, Haukaas SA, Akslen LA. Proliferation of immature tumor vessels is a novel marker of clinical progression in prostate cancer. *Cancer Res*. 2009 Jun 1;69(11):4708-15.
195. Gravdal K, Halvorsen OJ, Haukaas SA, Akslen LA. A switch from E-cadherin to N-cadherin expression indicates epithelial to mesenchymal transition and is of strong and

independent importance for the progress of prostate cancer. *Clin Cancer Res*. 2007 Dec 1;13(23):7003-11.
196. Sreekumar A, Poisson LM, Rajendiran TM, Khan AP, Cao Q, Yu J*, et al.* Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression. *Nature*. 2009 Feb 12;457(7231):910-4.
197. Kollermann J, Schlomm T, Bang H, Schwall GP, von Eichel-Streiber C, Simon R*, et al.* Expression and prognostic relevance of annexin A3 in prostate cancer. *Eur Urol*. 2008 Dec;54(6):1314-23.
198. Schostak M, Schwall GP, Poznanovic S, Groebe K, Muller M, Messinger D*, et al.* Annexin A3 in urine: a highly specific noninvasive marker for prostate cancer early detection. *J Urol*. 2009 Jan;181(1):343-53.
199. Mao X, Shaw G, James SY, Purkis P, Kudahetti SC, Tsigani T*, et al.* Detection of TMPRSS2:ERG fusion gene in circulating prostate cancer cells. *Asian J Androl*. 2008 May;10(3):467-73.
200. Hayes DF, Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Miller MC*, et al.* Circulating tumor cells at each follow-up time point during therapy of metastatic breast cancer patients predict progression-free and overall survival. *Clin Cancer Res*. 2006 Jul 15;12(14 Pt 1):4218-24.
201. Becker C, Piironen T, Pettersson K, Hugosson J, Lilja H. Clinical value of human glandular kallikrein 2 and free and total prostate-specific antigen in serum from a population of men with prostate-specific antigen levels 3.0 ng/mL or greater. *Urology*. 2000 May;55(5):694-9.
202. Haese A, Graefen M, Steuber T, Becker C, Noldus J, Erbersdobler A*, et al.* Total and Gleason grade 4/5 cancer volumes are major contributors of human kallikrein 2, whereas free prostate specific antigen is largely contributed by benign gland volume in serum from patients with prostate cancer or benign prostatic biopsies. *J Urol*. 2003 Dec;170(6 Pt 1):2269-73.
203. Haese A, Graefen M, Steuber T, Becker C, Pettersson K, Piironen T*, et al.* Human glandular kallikrein 2 levels in serum for discrimination of pathologically organ-confined from locally-advanced prostate cancer in total PSA-levels below 10 ng/ml. *Prostate*. 2001 Oct 1;49(2):101-9.
204. Elgamal AA, Holmes EH, Su SL, Tino WT, Simmons SJ, Peterson M*, et al.* Prostate-specific membrane antigen (PSMA): current benefits and future value. *Semin Surg Oncol*. 2000 Jan-Feb;18(1):10-6.
205. Elgamal AA, Troychak MJ, Murphy GP. ProstaScint scan may enhance identification of prostate cancer recurrences after prostatectomy, radiation, or hormone therapy: analysis of 136 scans of 100 patients. *Prostate*. 1998 Dec 1;37(4):261-9.
206. Talesa VN, Antognelli C, Del Buono C, Stracci F, Serva MR, Cottini E*, et al.* Diagnostic potential in prostate cancer of a panel of urinary molecular tumor markers. *Cancer Biomark*. 2009;5(6):241-51.
207. Fracalanza S, Prayer-Galetti T, Pinto F, Navaglia F, Sacco E, Ciaccia M*, et al.* Plasma chromogranin A in patients with prostate cancer improves the diagnostic efficacy of free/total prostate-specific antigen determination. *Urol Int*. 2005;75(1):57-61.
208. Taplin ME, George DJ, Halabi S, Sanford B, Febbo PG, Hennessy KT*, et al.* Prognostic significance of plasma chromogranin a levels in patients with hormone-refractory prostate cancer treated in Cancer and Leukemia Group B 9480 study. *Urology*. 2005 Aug;66(2):386-91.
209. Tricoli JV, Schoenfeldt M, Conley BA. Detection of prostate cancer and predicting progression: current and future diagnostic markers. *Clin Cancer Res*. 2004 Jun 15;10(12 Pt 1):3943-53.

210. Kamiya N, Akakura K, Suzuki H, Isshiki S, Komiya A, Ueda T*, et al.* Pretreatment serum level of neuron specific enolase (NSE) as a prognostic factor in metastatic prostate cancer patients treated with endocrine therapy. *Eur Urol*. 2003 Sep;44(3):309-14; discussion 14.
211. Crocitto LE, Korns D, Kretzner L, Shevchuk T, Blair SL, Wilson TG*, et al.* Prostate cancer molecular markers GSTP1 and hTERT in expressed prostatic secretions as predictors of biopsy results. *Urology*. 2004 Oct;64(4):821-5.
212. Gonzalgo ML, Nakayama M, Lee SM, De Marzo AM, Nelson WG. Detection of GSTP1 methylation in prostatic secretions using combinatorial MSP analysis. *Urology*. 2004 Feb;63(2):414-8.
213. Gonzalgo ML, Pavlovich CP, Lee SM, Nelson WG. Prostate cancer detection by GSTP1 methylation analysis of postbiopsy urine specimens. *Clin Cancer Res*. 2003 Jul;9(7):2673-7.
214. Nakayama M, Gonzalgo ML, Yegnasubramanian S, Lin X, De Marzo AM, Nelson WG. GSTP1 CpG island hypermethylation as a molecular biomarker for prostate cancer. *J Cell Biochem*. 2004 Feb 15;91(3):540-52.
215. Yegnasubramanian S, Kowalski J, Gonzalgo ML, Zahurak M, Piantadosi S, Walsh PC*, et al.* Hypermethylation of CpG islands in primary and metastatic human prostate cancer. *Cancer Res*. 2004 Mar 15;64(6):1975-86.
216. Rhodes DR, Sanda MG, Otte AP, Chinnaiyan AM, Rubin MA. Multiplex biomarker approach for determining risk of prostate-specific antigen-defined recurrence of prostate cancer. *J Natl Cancer Inst*. 2003 May 7;95(9):661-8.
217. Varambally S, Dhanasekaran SM, Zhou M, Barrette TR, Kumar-Sinha C, Sanda MG*, et al.* The polycomb group protein EZH2 is involved in progression of prostate cancer. *Nature*. 2002 Oct 10;419(6907):624-9.
218. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL*, et al.* Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A*. 2008 Jul 29;105(30):10513-8.
219. Spahn M, Kneitz S, Scholz CJ, Stenger N, Rudiger T, Strobel P*, et al.* Expression of microRNA-221 is progressively reduced in aggressive prostate cancer and metastasis and predicts clinical recurrence. *Int J Cancer*. 2010 Jul 15;127(2):394-403.
220. Tong AW, Fulgham P, Jay C, Chen P, Khalil I, Liu S*, et al.* MicroRNA profile analysis of human prostate cancers. *Cancer Gene Ther*. 2009 Mar;16(3):206-16.
221. Cussenot O, Teillac P, Berthon P, Latil A. Noninvasive detection of genetic instability in cells from prostatic secretion as a marker of prostate cancer. *Eur J Intern Med.* 2001 Feb;12(1):17-9.
222. Thuret R, Chantrel-Groussard K, Azzouzi AR, Villette JM, Guimard S, Teillac P*, et al.* Clinical relevance of genetic instability in prostatic cells obtained by prostatic massage in early prostate cancer. *Br J Cancer*. 2005 Jan 31;92(2):236-40.
223. de la Taille A. Progensa PCA3 test for prostate cancer detection. *Expert Rev Mol Diagn*. 2007 Sep;7(5):491-7.
224. Kirby RS, Fitzpatrick JM, Irani J. Prostate cancer diagnosis in the new millennium: strengths and weaknesses of prostate-specific antigen and the discovery and clinical evaluation of prostate cancer gene 3 (PCA3). *BJU Int*. 2009 Feb;103(4):441-5.
225. Schilling D, de Reijke T, Tombal B, de la Taille A, Hennenlotter J, Stenzl A. The Prostate Cancer gene 3 assay: indications for use in clinical practice. *BJU Int*. Feb;105(4):452-5.
226. Bussemakers MJ, van Bokhoven A, Verhaegh GW, Smit FP, Karthaus HF, Schalken JA*, et al.* DD3: a new prostate-specific gene, highly overexpressed in prostate cancer. *Cancer Res*. 1999 Dec 1;59(23):5975-9.

227. de Kok JB, Verhaegh GW, Roelofs RW, Hessels D, Kiemeney LA, Aalders TW*, et al.* DD3(PCA3), a very sensitive and specific marker to detect prostate tumors. *Cancer Res*. 2002 May 1;62(9):2695-8.
228. Gandini O, Santulli M, Cardillo MR, Stigliano A, Toscano V. Correspondence re: J. B. de Kok et al., DD3, A very sensitive and specific marker to detect prostate tumors. Cancer Res., 62: 2695-2698, 2002. *Cancer Res*. 2003 Aug 1;63(15):4747; author reply 8-9.
229. Roobol MJ, Schroder FH, van Leenders GL, Hessels D, van den Bergh RC, Wolters T*, et al.* Performance of Prostate Cancer Antigen 3 (PCA3) and Prostate-Specific Antigen in Prescreened Men: Reproducibility and Detection Characteristics for Prostate Cancer Patients with High PCA3 Scores (>/=100). *Eur Urol*. 2010 Sep 26.
230. Roobol MJ, Schroder FH, van Leeuwen P, Wolters T, van den Bergh RC, van Leenders GJ*, et al.* Performance of the prostate cancer antigen 3 (PCA3) gene and prostate-specific antigen in prescreened men: exploring the value of PCA3 for a first-line diagnostic test. *Eur Urol*. 2010 Oct;58(4):475-81.
231. Chen Z, Fan Z, McNeal JE, Nolley R, Caldwell MC, Mahadevappa M*, et al.* Hepsin and maspin are inversely expressed in laser capture microdissectioned prostate cancer. *J Urol*. 2003 Apr;169(4):1316-9.
232. Dhanasekaran SM, Barrette TR, Ghosh D, Shah R, Varambally S, Kurachi K*, et al.* Delineation of prognostic biomarkers in prostate cancer. *Nature*. 2001 Aug 23;412(6849):822-6.
233. Stephan C, Yousef GM, Scorilas A, Jung K, Jung M, Kristiansen G*, et al.* Hepsin is highly over expressed in and a new candidate for a prognostic indicator in prostate cancer. *J Urol*. 2004 Jan;171(1):187-91.
234. Rubin MA, Bismar TA, Andren O, Mucci L, Kim R, Shen R*, et al.* Decreased alpha-methylacyl CoA racemase expression in localized prostate cancer is associated with an increased rate of biochemical recurrence and cancer-specific death. *Cancer Epidemiol Biomarkers Prev*. 2005 Jun;14(6):1424-32.
235. Rogers CG, Yan G, Zha S, Gonzalgo ML, Isaacs WB, Luo J*, et al.* Prostate cancer detection on urinalysis for alpha methylacyl coenzyme a racemase protein. *J Urol*. 2004 Oct;172(4 Pt 1):1501-3.
236. Sreekumar A, Laxman B, Rhodes DR, Bhagavathula S, Harwood J, Giacherio D*, et al.* Humoral immune response to alpha-methylacyl-CoA racemase and prostate cancer. *J Natl Cancer Inst*. 2004 Jun 2;96(11):834-43.
237. Zielie PJ, Mobley JA, Ebb RG, Jiang Z, Blute RD, Ho SM. A novel diagnostic test for prostate cancer emerges from the determination of alpha-methylacyl-coenzyme a racemase in prostatic secretions. *J Urol*. 2004 Sep;172(3):1130-3.
238. Shariat SF, Roehrborn CG, McConnell JD, Park S, Alam N, Wheeler TM*, et al.* Association of the circulating levels of the urokinase system of plasminogen activation with the presence of prostate cancer and invasion, progression, and metastasis. *J Clin Oncol*. 2007 Feb 1;25(4):349-55.
239. Miyake H, Hara I, Yamanaka K, Gohji K, Arakawa S, Kamidono S. Elevation of serum levels of urokinase-type plasminogen activator and its receptor is associated with disease progression and prognosis in patients with prostate cancer. *Prostate*. 1999 May;39(2):123-9.
240. Dhir R, Vietmeier B, Arlotti J, Acquafondata M, Landsittel D, Masterson R*, et al.* Early identification of individuals with prostate cancer in negative biopsies. *J Urol*. 2004 Apr;171(4):1419-23.
241. Leman ES, Cannon GW, Trock BJ, Sokoll LJ, Chan DW, Mangold L*, et al.* EPCA-2: a highly specific serum marker for prostate cancer. *Urology*. 2007 Apr;69(4):714-20.

242. Laxman B, Tomlins SA, Mehra R, Morris DS, Wang L, Helgeson BE*, et al.* Noninvasive detection of TMPRSS2:ERG fusion transcripts in the urine of men with prostate cancer. *Neoplasia*. 2006 Oct;8(10):885-8.
243. Ivanovic V, Melman A, Davis-Joseph B, Valcic M, Geliebter J. Elevated plasma levels of TGF-beta 1 in patients with invasive prostate cancer. *Nat Med.* 1995 Apr;1(4):282-4.
244. Shariat SF, Walz J, Roehrborn CG, Montorsi F, Jeldres C, Saad F*, et al.* Early postoperative plasma transforming growth factor-beta1 is a strong predictor of biochemical progression after radical prostatectomy. *J Urol*. 2008 Apr;179(4):1593-7.
245. Umbas R, Isaacs WB, Bringuier PP, Schaafsma HE, Karthaus HF, Oosterhof GO*, et al.* Decreased E-cadherin expression is associated with poor prognosis in patients with prostate cancer. *Cancer Res*. 1994 Jul 15;54(14):3929-33.
246. Umbas R, Schalken JA, Aalders TW, Carter BS, Karthaus HF, Schaafsma HE*, et al.* Expression of the cellular adhesion molecule E-cadherin is reduced or absent in high-grade prostate cancer. *Cancer Res*. 1992 Sep 15;52(18):5104-9.
247. Gu Z, Thomas G, Yamashiro J, Shintaku IP, Dorey F, Raitano A*, et al.* Prostate stem cell antigen (PSCA) expression increases with high gleason score, advanced stage and bone metastasis in prostate cancer. *Oncogene*. 2000 Mar 2;19(10):1288-96.
248. Han KR, Seligson DB, Liu X, Horvath S, Shintaku PI, Thomas GV*, et al.* Prostate stem cell antigen expression is associated with gleason score, seminal vesicle invasion and capsular invasion in prostate cancer. *J Urol*. 2004 Mar;171(3):1117-21.
249. Hara N, Kasahara T, Kawasaki T, Bilim V, Obara K, Takahashi K*, et al.* Reverse transcription-polymerase chain reaction detection of prostate-specific antigen, prostate-specific membrane antigen, and prostate stem cell antigen in one milliliter of peripheral blood: value for the staging of prostate cancer. *Clin Cancer Res*. 2002 Jun;8(6):1794-9.
250. Lam JS, Yamashiro J, Shintaku IP, Vessella RL, Jenkins RB, Horvath S*, et al.* Prostate stem cell antigen is overexpressed in prostate cancer metastases. *Clin Cancer Res*. 2005 Apr 1;11(7):2591-6.
251. Laxman B, Morris DS, Yu J, Siddiqui J, Cao J, Mehra R*, et al.* A first-generation multiplex biomarker analysis of urine for the early detection of prostate cancer. *Cancer Res*. 2008 Feb 1;68(3):645-9.
252. Hessels D, Smit FP, Verhaegh GW, Witjes JA, Cornel EB, Schalken JA. Detection of TMPRSS2-ERG fusion transcripts and prostate cancer antigen 3 in urinary sediments may improve diagnosis of prostate cancer. *Clin Cancer Res*. 2007 Sep 1;13(17):5103-8.
253. Partin AW, Mangold LA, Lamm DM, Walsh PC, Epstein JI, Pearson JD. Contemporary update of prostate cancer staging nomograms (Partin Tables) for the new millennium. *Urology*. 2001 Dec;58(6):843-8.
254. Eichelberger LE, Koch MO, Eble JN, Ulbright TM, Juliar BE, Cheng L. Maximum tumor diameter is an independent predictor of prostate-specific antigen recurrence in prostate cancer. *Mod Pathol*. 2005 Jul;18(7):886-90.
255. Yu YP, Landsittel D, Jing L, Nelson J, Ren B, Liu L*, et al.* Gene expression alterations in prostate cancer predicting tumor aggression and preceding development of malignancy. *J Clin Oncol*. 2004 Jul 15;22(14):2790-9.
256. Andren O, Fall K, Franzen L, Andersson SO, Johansson JE, Rubin MA. How well does the Gleason score predict prostate cancer death? A 20-year followup of a population based cohort in Sweden. *J Urol*. 2006 Apr;175(4):1337-40.
257. Oyan AM, Bo TH, Jonassen I, Ulvestad E, Tore Gjertsen B, Bruserud O*, et al.* Global gene expression in classification, pathogenetic understanding and identification of

therapeutic targets in acute myeloid leukemia. *Curr Pharm Biotechnol*. 2007 Dec;8(6):344-54.
258. Tinker AV, Boussioutas A, Bowtell DD. The challenges of gene expression microarrays for the study of human cancer. *Cancer Cell*. 2006 May;9(5):333-9.
259. Dai M, Wang P, Boyd AD, Kostov G, Athey B, Jones EG*, et al.* Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res*. 2005;33(20):e175.
260. Petersen K, Oyan AM, Rostad K, Olsen S, Bo TH, Salvesen HB*, et al.* Comparison of nucleic acid targets prepared from total RNA or poly(A) RNA for DNA oligonucleotide microarray hybridization. *Anal Biochem*. 2007 Jul 1;366(1):46-58.
261. Dysvik B, Jonassen I. J-Express: exploring gene expression data using Java. *Bioinformatics*. 2001 Apr;17(4):369-70.
262. Henke RT, Maitra A, Paik S, Wellstein A. Gene expression analysis in sections and tissue microarrays of archival tissues by mRNA in situ hybridization. *Histol Histopathol*. 2005 Jan;20(1):225-37.
263. Herrington CS. Demystified ... in situ hybridisation. *Mol Pathol*. 1998 Feb;51(1):8-13.
264. Specht K, Richter T, Muller U, Walch A, Werner M, Hofler H. Quantitative gene expression analysis in microdissected archival formalin-fixed and paraffin-embedded tumor tissue. *Am J Pathol*. 2001 Feb;158(2):419-29.
265. Kononen J, Bubendorf L, Kallioniemi A, Barlund M, Schraml P, Leighton S*, et al.* Tissue microarrays for high-throughput molecular profiling of tumor specimens. *Nat Med*. 1998 Jul;4(7):844-7.
266. Kallioniemi OP, Wagner U, Kononen J, Sauter G. Tissue microarray technology for high-throughput molecular profiling of cancer. *Hum Mol Genet*. 2001 Apr;10(7):657-62.
267. Camp RL, Charette LA, Rimm DL. Validation of tissue microarray technology in breast carcinoma. *Lab Invest*. 2000 Dec;80(12):1943-9.
268. Hoos A, Urist MJ, Stojadinovic A, Mastorides S, Dudas ME, Leung DH*, et al.* Validation of tissue microarrays for immunohistochemical profiling of cancer specimens using the example of human fibroblastic tumors. *Am J Pathol*. 2001 Apr;158(4):1245-51.
269. Stefansson IM, Salvesen HB, Akslen LA. Prognostic impact of alterations in P-cadherin expression and related cell adhesion markers in endometrial cancer. *J Clin Oncol*. 2004 Apr 1;22(7):1242-52.
270. Halvorsen OJ, Haukaas S, Hoisaeter PA, Akslen LA. Maximum Ki-67 staining in prostate cancer provides independent prognostic information after radical prostatectomy. *Anticancer Res*. 2001 Nov-Dec;21(6A):4071-6.
271. Halvorsen OJ. Expression of p16 protein in prostatic adenocarcinomas, intraepithelial neoplasia, and benign/hyperplastic glands. *Urol Oncol*. 1997;3:59-66.
272. Rubin MA, Dunn R, Strawderman M, Pienta KJ. Tissue microarray sampling strategy for prostate cancer biomarker analysis. *Am J Surg Pathol*. 2002 Mar;26(3):312-9.
273. Nelson PN, Reynolds GM, Waldron EE, Ward E, Giannopoulos K, Murray PG. Monoclonal antibodies. *Mol Pathol*. 2000 Jun;53(3):111-7.
274. Nocito A, Kononen J, Kallioniemi OP, Sauter G. Tissue microarrays (TMAs) for high-throughput molecular pathology research. *Int J Cancer*. 2001 Oct 1;94(1):1-5.
275. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part I: basic concepts and first analyses. *Br J Cancer*. 2003 Jul 21;89(2):232-8.
276. Kirkwood B, Sterne, JAC. Essential medical statistics: Blackwell Science; 2005.
277. Clark TG, Bradburn MJ, Love SB, Altman DG. Survival analysis part IV: further concepts and methods in survival analysis. *Br J Cancer*. 2003 Sep 1;89(5):781-6.

278. Partin AW, Kattan MW, Subong EN, Walsh PC, Wojno KJ, Oesterling JE*, et al.* Combination of prostate-specific antigen, clinical stage, and Gleason score to predict pathological stage of localized prostate cancer. A multi-institutional update. *JAMA*. 1997 May 14;277(18):1445-51.
279. Luo J, Duggan DJ, Chen Y, Sauvageot J, Ewing CM, Bittner ML*, et al.* Human prostate cancer and benign prostatic hyperplasia: molecular dissection by gene expression profiling. *Cancer Res*. 2001 Jun 15;61(12):4683-8.
280. Xu J, Stolk JA, Zhang X, Silva SJ, Houghton RL, Matsumura M*, et al.* Identification of differentially expressed genes in human prostate cancer using subtraction and microarray. *Cancer Res*. 2000 Mar 15;60(6):1677-82.
281. Ashida S, Nakagawa H, Katagiri T, Furihata M, Iiizumi M, Anazawa Y*, et al.* Molecular features of the transition from prostatic intraepithelial neoplasia (PIN) to prostate cancer: genome-wide gene-expression profiles of prostate cancers and PINs. *Cancer Res*. 2004 Sep 1;64(17):5963-72.
282. Glinsky GV, Glinskii AB, Stephenson AJ, Hoffman RM, Gerald WL. Gene expression profiling predicts clinical outcome of prostate cancer. *J Clin Invest*. 2004 Mar;113(6):913-23.
283. Lapointe J, Li C, Higgins JP, van de Rijn M, Bair E, Montgomery K*, et al.* Gene expression profiling identifies clinically relevant subtypes of prostate cancer. *Proc Natl Acad Sci U S A*. 2004 Jan 20;101(3):811-6.
284. LaTulippe E, Satagopan J, Smith A, Scher H, Scardino P, Reuter V*, et al.* Comprehensive gene expression analysis of prostate cancer reveals distinct transcriptional programs associated with metastatic disease. *Cancer Res*. 2002 Aug 1;62(15):4499-506.
285. Magee JA, Araki T, Patil S, Ehrig T, True L, Humphrey PA*, et al.* Expression profiling reveals hepsin overexpression in prostate cancer. *Cancer Res*. 2001 Aug 1;61(15):5692-6.
286. Singh D, Febbo PG, Ross K, Jackson DG, Manola J, Ladd C*, et al.* Gene expression correlates of clinical prostate cancer behavior. *Cancer Cell*. 2002 Mar;1(2):203-9.
287. Welsh JB, Sapinoso LM, Su AI, Kern SG, Wang-Rodriguez J, Moskaluk CA*, et al.* Analysis of gene expression identifies candidate markers and pharmacological targets in prostate cancer. *Cancer Res*. 2001 Aug 15;61(16):5974-8.
288. Dakhova O, Ozen M, Creighton CJ, Li R, Ayala G, Rowley D*, et al.* Global gene expression analysis of reactive stroma in prostate cancer. *Clin Cancer Res*. 2009 Jun 15;15(12):3979-89.
289. Knight JF, Shepherd CJ, Rizzo S, Brewer D, Jhavar S, Dodson AR*, et al.* TEAD1 and c-Cbl are novel prostate basal cell markers that correlate with poor clinical outcome in prostate cancer. *Br J Cancer*. 2008 Dec 2;99(11):1849-58.
290. Richardson AM, Woodson K, Wang Y, Rodriguez-Canales J, Erickson HS, Tangrea MA*, et al.* Global expression analysis of prostate cancer-associated stroma and epithelia. *Diagn Mol Pathol*. 2007 Dec;16(4):189-97.
291. Tamura K, Furihata M, Tsunoda T, Ashida S, Takata R, Obara W*, et al.* Molecular features of hormone-refractory prostate cancer cells by genome-wide gene expression profiles. *Cancer Res*. 2007 Jun 1;67(11):5117-25.
292. Thorsen K, Sorensen KD, Brems-Eskildsen AS, Modin C, Gaustadnes M, Hein AM*, et al.* Alternative splicing in colon, bladder, and prostate cancer identified by exon array analysis. *Mol Cell Proteomics*. 2008 Jul;7(7):1214-24.
293. Tomlins SA, Mehra R, Rhodes DR, Cao X, Wang L, Dhanasekaran SM*, et al.* Integrative molecular concept modeling of prostate cancer progression. *Nat Genet*. 2007 Jan;39(1):41-51.

294. True L, Coleman I, Hawley S, Huang CY, Gifford D, Coleman R*, et al.* A molecular correlate to the Gleason grading system for prostate adenocarcinoma. *Proc Natl Acad Sci U S A*. 2006 Jul 18;103(29):10991-6.
295. Luo JH, Yu YP, Cieply K, Lin F, Deflavia P, Dhir R*, et al.* Gene expression analysis of prostate cancers. *Mol Carcinog*. 2002 Jan;33(1):25-35.
296. Ramaswamy S, Tamayo P, Rifkin R, Mukherjee S, Yeang CH, Angelo M*, et al.* Multiclass cancer diagnosis using tumor gene expression signatures. *Proc Natl Acad Sci U S A*. 2001 Dec 18;98(26):15149-54.
297. Rhodes DR, Yu J, Shanker K, Deshpande N, Varambally R, Ghosh D*, et al.* ONCOMINE: a cancer microarray database and integrated data-mining platform. *Neoplasia*. 2004 Jan-Feb;6(1):1-6.
298. Farrall AL, Whitelaw ML. The HIF1alpha-inducible pro-cell death gene BNIP3 is a novel target of SIM2s repression through cross-talk on the hypoxia response element. *Oncogene*. 2009 Oct 15;28(41):3671-80.
299. Aleman MJ, DeYoung MP, Tress M, Keating P, Perry GW, Narayanan R. Inhibition of Single Minded 2 gene expression mediates tumor-selective apoptosis and differentiation in human colon cancer cells. *Proc Natl Acad Sci U S A*. 2005 Sep 6;102(36):12765-70.
300. Li CM, Kim CE, Margolin AA, Guo M, Zhu J, Mason JM*, et al.* CTNNB1 mutations and overexpression of Wnt/beta-catenin target genes in WT1-mutant Wilms' tumors. *Am J Pathol*. 2004 Dec;165(6):1943-53.
301. Metz RP, Kwak HI, Gustafson T, Laffin B, Porter WW. Differential transcriptional regulation by mouse single-minded 2s. *J Biol Chem*. 2006 Apr 21;281(16):10839-48.
302. Arredouani MS, Lu B, Bhasin M, Eljanne M, Yue W, Mosquera JM*, et al.* Identification of the transcription factor single-minded homologue 2 as a potential biomarker and immunotherapy target in prostate cancer. *Clin Cancer Res*. 2009 Sep 15;15(18):5794-802.
303. Ren R. Mechanisms of BCR-ABL in the pathogenesis of chronic myelogenous leukaemia. *Nat Rev Cancer*. 2005 Mar;5(3):172-83.
304. Mitelman F, Johansson B, Mertens F. Fusion genes and rearranged genes as a linear function of chromosome aberrations in cancer. *Nat Genet*. 2004 Apr;36(4):331-4.
305. Mitelman F, Mertens F, Johansson B. Prevalence estimates of recurrent balanced cytogenetic aberrations and gene fusions in unselected patients with neoplastic disorders. *Genes Chromosomes Cancer*. 2005 Aug;43(4):350-66.
306. Haffner MC, Aryee MJ, Toubaji A, Esopi DM, Albadine R, Gurel B*, et al.* Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements. *Nat Genet*. 2010 Aug;42(8):668-75.
307. Clark J, Merson S, Jhavar S, Flohr P, Edwards S, Foster CS*, et al.* Diversity of TMPRSS2-ERG fusion transcripts in the human prostate. *Oncogene*. 2007 Apr 19;26(18):2667-73.
308. Hermans KG, Bressers AA, van der Korput HA, Dits NF, Jenster G, Trapman J. Two unique novel prostate-specific and androgen-regulated fusion partners of ETV4 in prostate cancer. *Cancer Res*. 2008 May 1;68(9):3094-8.
309. Pflueger D, Rickman DS, Sboner A, Perner S, LaFargue CJ, Svensson MA*, et al.* N-myc downstream regulated gene 1 (NDRG1) is fused to ERG in prostate cancer. *Neoplasia*. 2009 Aug;11(8):804-11.
310. Cerveira N, Ribeiro FR, Peixoto A, Costa V, Henrique R, Jeronimo C*, et al.* TMPRSS2-ERG gene fusion causing ERG overexpression precedes chromosome copy number changes in prostate carcinomas and paired HGPIN lesions. *Neoplasia*. 2006 Oct;8(10):826-32.

311. Hermans KG, van Marion R, van Dekken H, Jenster G, van Weerden WM, Trapman J. TMPRSS2:ERG fusion by translocation or interstitial deletion is highly relevant in androgen-dependent prostate cancer, but is bypassed in late-stage androgen receptor-negative prostate cancer. *Cancer Res*. 2006 Nov 15;66(22):10658-63.
312. Iljin K, Wolf M, Edgren H, Gupta S, Kilpinen S, Skotheim RI*, et al.* TMPRSS2 fusions with oncogenic ETS factors in prostate cancer involve unbalanced genomic rearrangements and are associated with HDAC1 and epigenetic reprogramming. *Cancer Res*. 2006 Nov 1;66(21):10242-6.
313. Soller MJ, Isaksson M, Elfving P, Soller W, Lundgren R, Panagopoulos I. Confirmation of the high frequency of the TMPRSS2/ERG fusion gene in prostate cancer. *Genes Chromosomes Cancer*. 2006 Jul;45(7):717-9.
314. Tomlins SA, Mehra R, Rhodes DR, Smith LR, Roulston D, Helgeson BE*, et al.* TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. *Cancer Res*. 2006 Apr 1;66(7):3396-400.
315. Okada H, Tsubura A, Okamura A, Senzaki H, Naka Y, Komatz Y*, et al.* Keratin profiles in normal/hyperplastic prostates and prostate carcinoma. *Virchows Arch A Pathol Anat Histopathol*. 1992;421(2):157-61.
316. Parsons JK, Gage WR, Nelson WG, De Marzo AM. p63 protein expression is rare in prostate adenocarcinoma: implications for cancer diagnosis and carcinogenesis. *Urology*. 2001 Oct;58(4):619-24.
317. Mehra R, Tomlins SA, Yu J, Cao X, Wang L, Menon A*, et al.* Characterization of TMPRSS2-ETS gene aberrations in androgen-independent metastatic prostate cancer. *Cancer Res*. 2008 May 15;68(10):3584-90.
318. Perner S, Mosquera JM, Demichelis F, Hofer MD, Paris PL, Simko J*, et al.* TMPRSS2-ERG fusion prostate cancer: an early molecular event associated with invasion. *Am J Surg Pathol*. 2007 Jun;31(6):882-8.
319. Tomlins SA, Laxman B, Varambally S, Cao X, Yu J, Helgeson BE*, et al.* Role of the TMPRSS2-ERG gene fusion in prostate cancer. *Neoplasia*. 2008 Feb;10(2):177-88.
320. Mosquera JM, Perner S, Genega EM, Sanda M, Hofer MD, Mertz KD*, et al.* Characterization of TMPRSS2-ERG fusion high-grade prostatic intraepithelial neoplasia and potential clinical implications. *Clin Cancer Res*. 2008 Jun 1;14(11):3380-5.
321. Carver BS, Tran J, Gopalan A, Chen Z, Shaikh S, Carracedo A*, et al.* Aberrant ERG expression cooperates with loss of PTEN to promote cancer progression in the prostate. *Nat Genet*. 2009 May;41(5):619-24.
322. King JC, Xu J, Wongvipat J, Hieronymus H, Carver BS, Leung DH*, et al.* Cooperativity of TMPRSS2-ERG with PI3-kinase pathway activation in prostate oncogenesis. *Nat Genet*. 2009 May;41(5):524-6.
323. Shaffer DR, Pandolfi PP. Breaking the rules of cancer. *Nat Med*. 2006 Jan;12(1):14-5.
324. Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS*, et al.* Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature*. 2007 Aug 2;448(7153):595-9.
325. Yoshimoto M, Joshua AM, Cunha IW, Coudry RA, Fonseca FP, Ludkovski O*, et al.* Absence of TMPRSS2:ERG fusions and PTEN losses in prostate cancer is associated with a favorable outcome. *Mod Pathol*. 2008 Dec;21(12):1451-60.
326. Yu J, Yu J, Mani RS, Cao Q, Brenner CJ, Cao X*, et al.* An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. *Cancer Cell*. 2010 May 18;17(5):443-54.

327. Kunderfranco P, Mello-Grand M, Cangemi R, Pellini S, Mensah A, Albertini V*, et al.* ETS transcription factors control transcription of EZH2 and epigenetic silencing of the tumor suppressor gene Nkx3.1 in prostate cancer. *PLoS One*. 2010;5(5):e10547.
328. Barry M PS, Demichelis F, Rubin MA. TMPRSS2-ERG Fusion Heterogeneity in Multifocal Prostate Cancer: Clinical and Biologic Implications. *Urology*. 2007;70(4):630-3.
329. Arora R, Koch MO, Eble JN, Ulbright TM, Li L, Cheng L. Heterogeneity of Gleason grade in multifocal adenocarcinoma of the prostate. *Cancer*. 2004 Jun 1;100(11):2362-6.
330. Clark J, Attard G, Jhavar S, Flohr P, Reid A, De-Bono J*, et al.* Complex patterns of ETS gene alteration arise during cancer development in the human prostate. *Oncogene*. 2008 Mar 27;27(14):1993-2003.
331. Mehra R, Han B, Tomlins SA, Wang L, Menon A, Wasco MJ*, et al.* Heterogeneity of TMPRSS2 gene rearrangements in multifocal prostate adenocarcinoma: molecular evidence for an independent group of diseases. *Cancer Res*. 2007 Sep 1;67(17):7991-5.
332. Jamaspishvili T, Kral M, Khomeriki I, Student V, Kolar Z, Bouchal J. Urine markers in monitoring for prostate cancer. *Prostate Cancer Prostatic Dis*. 2010 Mar;13(1):12-9.
333. Park K, Tomlins SA, Mudaliar KM, Chiu YL, Esgueva R, Mehra R*, et al.* Antibody-based detection of ERG rearrangement-positive prostate cancer. *Neoplasia*. 2010 Jul;12(7):590-8.
334. Mosquera JM, Perner S, Demichelis F, Kim R, Hofer MD, Mertz KD*, et al.* Morphological features of TMPRSS2-ERG gene fusion prostate cancer. *J Pathol*. 2007 May;212(1):91-101.