SANNI JALAVA

The Role of 8q Amplification and microRNAs in Prostate Cancer

ACADEMIC DISSERTATION
To be presented, with the permission of the board of Institute of Biomedical Technology of the University of Tampere, for public discussion in the Jarmo Visakorpi Auditorium, of the Arvo Building, Lääkärinkatu 1, Tampere, on June 1st, 2012, at 12 o’clock.

UNIVERSITY OF TAMPERE
ACADEMIC DISSERTATION
University of Tampere, Institute of Biomedical Technology and BioMediTech
Tampere University Hospital, Laboratory Centre
Graduate Program in Biomedicine and Biotechnology (TGPBB)
Finland

Supervised by
Professor Tapio Visakorpi
University of Tampere
Finland
Kati Porkka, PhD
University of Tampere
Finland

Reviewed by
Professor Johanna Ivaska
University of Turku
Finland
Dr. Ian Mills
University of Oslo
Norway

Copyright ©2012 Tampere University Press and the author

Distribution
Bookshop TAJU
P.O. Box 617
33014 University of Tampere
Finland

Tel. +358 40 190 9800
Fax +358 3 3551 7685
taju@uta.fi
www.uta.fi/taju
http://granum.uta.fi

Cover design by
Mikko Reinikka

Acta Universitatis Tamperensis 1733
ISBN 978-951-44-8813-9 (print)
ISSN-L 1455-1616
ISSN 1455-1616

Acta Electronica Universitatis Tamperensis 1204
ISBN 978-951-44-8814-6 (pdf)
ISSN 1456-954X
http://acta.uta.fi

Tampereen Yliopistopaino Oy – Juvenes Print
Tampere 2012
CONTENTS

CONTENTS ....................................................................................................................... 3
LIST OF ORIGINAL COMMUNICATIONS ................................................................. 6
ABBREVIATIONS ........................................................................................................... 7
TIIVISTELMÄ ............................................................................................................... 9
ABSTRACT ................................................................................................................... 11
INTRODUCTION ......................................................................................................... 13
REVIEW OF THE LITERATURE ............................................................................... 15
  1.1 Prostate cancer carcinogenesis ................................................................. 15
    1.1.1 Inheritable factors ................................................................................. 16
    1.1.2 Chromosomal alterations ....................................................................... 16
      1.1.2.1 8q amplification ............................................................................. 17
    1.1.3 Early events .............................................................................................. 18
      1.1.3.1 NKX3.1 down-regulation ................................................................... 19
      1.1.3.2 ETS translocations .......................................................................... 20
    1.1.4 Androgen receptor signaling ............................................................... 21
      1.1.4.1 Normal functions of AR .................................................................... 21
      1.1.4.2 AR function in prostate cancer ......................................................... 22
      1.1.4.3 AR function in CRPC ....................................................................... 22
      1.1.4.4 Therapeutic strategies of targeting AR pathway ......................... 24
    1.1.5 MYC .............................................................................................................. 25
    1.1.6 PI3K pathway and PTEN ........................................................................... 26
    1.1.7 Metastatic prostate cancer ...................................................................... 28
      1.1.7.1 EZH2 ................................................................................................. 29
      1.1.7.2 TP53 ................................................................................................... 29
    1.1.8 miRNAs in prostate cancer ................................................................. 30
      1.1.8.1 microRNA biogenesis and target recognition .............................. 30
      1.1.8.2 Expression profiling of miRNAs ....................................................... 32
      1.1.8.3 Tumor suppressor miRs .............................................................. 35
      1.1.8.4 OncomiRs ....................................................................................... 36
1.1.8.5 miRNAs and AR .......................................................... 37

1.2 Experimental models of prostate cancer .................................................. 38
1.2.1 Cell lines.................................................................................. 38
1.2.2 Xenografts ............................................................................... 39
1.2.3 Mouse models ........................................................................... 39

AIMS OF THE STUDY ........................................................................ 42

MATERIALS AND METHODS .................................................................. 43
1.1 Cell lines and clinical samples (I, II, III) ............................................. 43
1.2 Microarrays (I, II, III) ................................................................. 44
1.2.1 miRNA microarray (II, III) ...................................................... 44
1.2.2 cDNA microarray (I, III) .............................................................. 45
1.3 Real-time quantitative RT-PCR .................................................... 46
1.3.1 mRNA expression analysis (I, III) ............................................ 46
1.3.2 miRNA expression analysis (II, III) ......................................... 46
1.3.3 Chip-qPCR (III) .................................................................... 46
1.4 Immunohistochemistry (III) .......................................................... 47
1.5 Western blotting (I, III) ............................................................... 47
1.6 Chromatin immunoprecipitation (III) ............................................ 48
1.7 Transfections and transductions ....................................................... 48
1.7.1 Lentiviral-mediated gene silencing and overexpression (I) .......... 48
1.7.2 Transfections of pre- and anti-miRs (III) .................................... 49
1.8 Functional assays ............................................................................ 49
1.8.1 DHT treatment prior to functional assays (III) ........................... 49
1.8.2 Proliferation assays (I, III) ......................................................... 50
1.8.3 Cell cycle analysis (I, III) .......................................................... 50
1.8.4 Apoptosis assay (III) ................................................................. 50
1.8.5 Soft agar assay (I) .................................................................. 50
1.8.6 Invasion assay (I) ................................................................... 51
1.8.7 Migration assay (I) .................................................................. 51
1.8.8 Luciferase assay (III) ............................................................... 51
1.8.9 Rescue experiment (III) ............................................................. 52
1.9 Statistical analysis (I, II, III) ......................................................... 52

RESULTS AND DISCUSSION ................................................................ 53
1.1 Functional evaluation of putative 8q target genes (I) ....................... 53
LIST OF ORIGINAL COMMUNICATIONS

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


1 Authors contributed equally to this work
# ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKT</td>
<td>v-akt murine thymoma viral oncogene homolog</td>
</tr>
<tr>
<td>AR</td>
<td>androgen receptor</td>
</tr>
<tr>
<td>ARBS</td>
<td>androgen receptor binding site</td>
</tr>
<tr>
<td>ARE</td>
<td>androgen response element</td>
</tr>
<tr>
<td>BPH</td>
<td>benign prostatic hyperplasia</td>
</tr>
<tr>
<td>BTG2</td>
<td>BTG family, member 2</td>
</tr>
<tr>
<td>ChiP</td>
<td>chromatin immunoprecipitation</td>
</tr>
<tr>
<td>CRPC</td>
<td>castration-resistant prostate cancer</td>
</tr>
<tr>
<td>DHT</td>
<td>5α-dihydrotestosterone</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>EIF3H</td>
<td>eukaryotic translation initiation factor 3, subunit H</td>
</tr>
<tr>
<td>EMT</td>
<td>epithelial-mesenchymal transition</td>
</tr>
<tr>
<td>ERG</td>
<td>v-ets erythroblastosis virus E26 oncogene homolog (avian)</td>
</tr>
<tr>
<td>ETS</td>
<td>avian erythroblastosis virus E26 homolog</td>
</tr>
<tr>
<td>EZH2</td>
<td>enhancer of zeste homolog 2 (Drosophila)</td>
</tr>
<tr>
<td>HPC</td>
<td>hereditary prostate cancer</td>
</tr>
<tr>
<td>IHC</td>
<td>immunohistochemistry</td>
</tr>
<tr>
<td>miRNA</td>
<td>microRNA</td>
</tr>
<tr>
<td>MRE</td>
<td>microRNA response element</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>MYC</td>
<td>v-myc myelocytomatosis viral oncogene homolog (avian)</td>
</tr>
<tr>
<td>NKX3.1</td>
<td>NK3 homeobox 1</td>
</tr>
<tr>
<td>PC, PCa</td>
<td>prostate cancer</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PI3K</td>
<td>phosphoinositide-3-kinase, catalytic, alpha polypeptide</td>
</tr>
<tr>
<td>PIK3IP1</td>
<td>phosphoinositide-3-kinase interacting protein 1</td>
</tr>
<tr>
<td>PIN</td>
<td>prostate intraepithelial neoplasia</td>
</tr>
<tr>
<td>PSA</td>
<td>prostate specific antigen</td>
</tr>
<tr>
<td>PTEN</td>
<td>phosphatase and tensin homolog</td>
</tr>
<tr>
<td>qRT-PCR</td>
<td>quantitative real-time PCR</td>
</tr>
<tr>
<td>RAD21</td>
<td>RAD21 homolog (S. pombe)</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>shRNA</td>
<td>shor hairpin RNA</td>
</tr>
<tr>
<td>siRNA</td>
<td>small interfering RNA</td>
</tr>
<tr>
<td>TCEB1</td>
<td>transcription elongation factor B (SIII), polypeptide 1</td>
</tr>
<tr>
<td>TMPRSS2</td>
<td>transmembrane protease, serine 2</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------</td>
</tr>
<tr>
<td>TP53</td>
<td>tumor protein 53, p53</td>
</tr>
<tr>
<td>TURP</td>
<td>transurethral resection of prostate</td>
</tr>
<tr>
<td>UTR</td>
<td>untranslated region</td>
</tr>
</tbody>
</table>
Viime vuosien ahkera tutkimustyö ei ole kyennyt täysin selvittämään eturauhassyövän syntynyyn ja etenemiseen liittyviä solutason molekulaarisia mekanismeja. Yleisimmät kromosominmuutokset tunnetaan, mutta useat kohdegeenit ovat vielä tunnistamatta. Kastraatioresistentti eturauhassyöpä, joka syntyy hormonihoidon jälkeen, on tappava tauti. Toistaiseksi ainoa geeni, jonka tiedetään liittyvän kastraatioresistentin syövän syntynyyn on androgeeniresseptori (AR).

Yksi yleisimmistä kromosomaalisista muutoksista, joka löytyy kastraatioresistentistä syövästä, on kromosomi 8 pitkän käsivarren (q) monistuma. Monistuma löytyy 60-90% kasvaimista ja sen on osoitettu olevan yhteydessä potilaiden huonoon ennusteeseen eturauhassyöpäleikkauksen jälkeen. Alueella on useita minimaalisia monistumia, mikä viittaa useampien kohdegeenien olemassaoloon. Tässä tutkimuksessa tutkittiin toiminnallisesti kokein neljää mahdollista kohdegeeniä (EIF3H, TCEB1, KIAA0196 ja RAD21). Näistä TCEB1:n, joka sijaitsee kromosomialueella 8q21, osoitettiin edistävän syöpäsoluten invaasiota ja alustasta riippumatonta kasvua. Tämän lisäksi TCEB1:n yli-ilmentäminen hiiren fibroblastisolulinjassa lisäsi solujen kasvua. Tulosten perusteella näyttää siltä, että TCEB1 on mahdollinen kohdegeeni 8q21 alueen monistumalle.


osoitettiin olevan yhteydessä lyhyempään leikkauksen jälkeiseen progressiovapaaseen aikaan.

Despite extensive research in recent years, the molecular mechanisms underlying prostate cancer initiation and progression are not fully understood. Key chromosomal aberrations have been identified, yet many of the target genes remain elusive. Castration-resistant prostate cancer (CRPC) is a lethal disease that emerges after hormonal therapy. So far, the only gene known to be involved in the formation of CRPC is the androgen receptor (AR).

A gain of the long arm of chromosome 8 is one of the most common findings in CRPC; in fact, 60-90% of these tumors harbor this particular gain. In addition, 8q gain is associated with poor prognosis in prostatectomy-treated patients. Several minimal regions have been identified, suggesting the existence of multiple target genes. Four of the putative target genes (EIF3H, TCEB1, KIAA0196 and RAD21) within the 8q gain region have been functionally evaluated. Of these, TCEB1, located at 8q21, was shown to promote invasion and to affect the anchorage-independent growth of prostate cancer cells. Furthermore, TCEB1 overexpression enhanced the growth of murine fibroblasts. These data indicate that TCEB1 is a putative target gene for gain of the minimal 8q21 region.

microRNAs (miRNA) are short, non-coding RNAs that negatively regulate gene expression and can function as tumor suppressor miRs or oncomiRs. To address whether miRNAs are involved in prostate cancer progression, their expression was investigated in two different clinical datasets (102 samples and 54 samples). A panel of 25 miRNAs was able to distinguish aggressive from less aggressive tumors.

Androgen-regulated miRNAs involved in CRPC were identified by combining information on androgen receptor binding sites (ARBS) and CRPC miRNA expression profiles. Twenty-eight miRNAs were deregulated in CRPC and contained ARBS. Exogenous overexpression of miR-32 and miR-148a enhanced the growth of an androgen-responsive cell line; miR-32 accomplished this by reducing apoptosis. The expression of these two miRNAs was demonstrated to be up-regulated by androgens and increased in CRPC. In addition, ARBS were detected in close proximity to these miRNAs. BTG2 and PIK3IP1 were identified as target genes for miR-32 and miR-148a, respectively. BTG2 expression was markedly decreased in CRPC, and the loss of BTG2 expression was associated with shorter progression-free survival in prostatectomy-treated patients.

In conclusion, TCEB1, miR-148a and miR-32 were demonstrated to possess oncogenic functions in prostate cancer. They are all overexpressed in CRPC and, in initial cell-based studies, affected invasion, proliferation and apoptosis. miR-32 and miR-148a were determined to be androgen-regulated. In addition, miRNA
expression profiling was effective in predicting the prognosis of prostatectomy-treated patients. However, additional studies are necessary to evaluate the role of miRNAs as prognostic indicators in prostate cancer.
Cancer is a disease caused by the accumulation of genetic and epigenetic alterations in the genome. Under normal conditions, cells are strictly controlled by a gene network that regulates cell growth, proliferation, death and differentiation. During tumorigenesis, alterations in these regulatory genes occur over time. The first mutation that occurs in a gatekeeper pathway provides a selective growth advantage to the cell. The developing tumor then acquires mutations in oncogenes, tumor suppressors or stability genes (for example, DNA repair genes), which eventually lead to clonal expansion and the creation of a mass of cells that forms the primary tumor (Vogelstein & Kinzler 2004). Genomic mutations that lead to cancer are inherited or caused by environmental factors or result from random errors that occur during DNA replication.

In males, prostate cancer is the most common malignancy in Finland and other Western countries and is the second most common cause of cancer-related death (Jemal et al. 2011; Finnish Cancer Register 2011). Both environmental and genetic factors have been implicated in the etiology of prostate cancer; well-established risk factors include age, ethnicity and family history (Crawford et al. 2003). The incidence of prostate cancer varies considerably among different populations (Jemal et al. 2011). Indeed, up to 42% of the risk for prostate cancer is due to heritable factors (Lichtenstein et al. 2000).

Prostate cancer is a complex and heterogeneous disease. Typically, prostate cancer is relatively indolent, and only a small minority of men with this disease eventually dies of the cancer. In contrast, a fraction of prostate tumors display an aggressive phenotype and progress rapidly (Shen & Abate-Shen, 2010; Jemal et al. 2011). A major clinical challenge is to distinguish the indolent from the aggressive tumors. Currently, outcome prediction methods are based on the Gleason score and the level of prostate specific antigen (PSA), which are considered inadequate (Walsh et al. 2007). The Gleason scoring system classifies prostate tumors based on their histology (most to least differentiated) (Mellinger et al. 1967; Epstein 2010). PSA is secreted by the normal prostate but is released into the bloodstream when normal prostate architecture is disrupted (Lilja et al. 2008). Several factors other than cancer can elevate PSA levels, such as benign prostatic hyperplasia (BPH). PSA is used for early prostate cancer detection, but a biopsy is always necessary to confirm the presence of cancerous cells in the prostate. In a large European study, PSA screening reduced prostate cancer mortality by 20%. However, overdiagnosis and overtreatment are significant adverse effects of PSA screening, as 1410 men had to be screened and 48 treated to prevent one death (Schröder et al. 2009).
Extensive research during recent years has identified common alterations and frequent sites of chromosomal gains and losses in prostate cancer. However, many of the target genes remain unknown, and attempts to develop better diagnostic and prognostic markers for prostate cancer have so far been unsuccessful. The main oncogenic drivers for prostate cancer are androgens and AR signaling. It has been known for more than half a century that castration is an effective prostate cancer treatment, and hormonal therapy remains the standard treatment for advanced disease (Huggins & Hodges, 1941).

The standard treatments for localized prostate cancer are surgery (radical prostatectomy) and radiation therapy (Klein et al. 2009). Patients who relapse, or who are initially diagnosed with advanced disease, are given androgen deprivation therapy. Initially, 70-80% of these patients respond to the therapy and experience tumor regression (Huggins & Hodges, 1941 & 1972, Knudsen & Scher, 2009). The effect can last for months to years, but eventually, castration-resistant prostate cancer (CRPC) arises, marked initially by rising PSA values (biochemical relapse). CRPC is a fatal disease with no curative treatment, but it remains driven by androgens (Goktas et al. 1999; Knudsen & Scher, 2009). Novel drugs that effectively inhibit androgen action or biosynthesis prolong the survival of CRPC patients (de Bono et al. 2011; Scher et al. 2010).

A more thorough understanding of the molecular mechanisms underlying prostate cancer tumorigenesis will hopefully reveal novel diagnostic and prognostic biomarkers and lead to the development of innovative therapeutics.
REVIEW OF THE LITERATURE

1.1 Prostate cancer carcinogenesis

Prostate cancer is a slow-growing disease. Prostatic intraepithelial neoplasia (PIN) is the pre-malignant precursor of prostate cancer. Prostate cancer progresses from early, androgen-dependent localized tumors to metastatic, invasive disease. Hormonal therapy leads to tumor regression; however, castration-resistant disease eventually emerges (Fig 1). In the following chapters, the protein-coding genes and miRNAs involved in prostate cancer progression are summarized.

Figure 1: Molecular mechanisms of prostate cancer. (Adapted from: Wright et al. 2005)
1.1.1 Inheritable factors

Family history, which reveals the inherited genetic background, is a significant risk factor for the development of prostate cancer. Epidemiological studies have shown that having a first-degree relative (brother or father) with prostate cancer increases an individual’s risk by two to three fold.

Hereditary prostate cancer (HPC) is a subtype of familial prostate cancer that is marked by a pattern of Mendelian inheritance and accounts for 5-10% of prostate cancer cases (Carter et al. 1993; Cerhan et al. 1999).

HPC is likely caused by multiple genes with varying levels of penetrance that interact with each other and with environmental factors. Highly penetrant genes are rare but increase the risk for prostate cancer by several fold. In contrast, genes with low penetrance are more common and contribute only modestly to the risk for prostate cancer (Witte et al. 2009). Genes with high penetrance that have already been identified include RNASEL, BRCA2, MSRI and ELAC, but these genes underlie only a minority of HPC cases (Xu et al. 2001 & 2002; Carpten et al. 2002; Edwards et al. 2003).

Genome-wide association studies have identified several chromosomal loci across multiple populations that are associated with prostate cancer risk, including 17q21, Xp11 and 10q21 (Witte et al. 2009). One of the most interesting recent findings is that a series of sequence variants located in 8q24 are significantly associated with prostate cancer risk. These risk variants are located upstream of the oncogene MYC (myelocytomatosis viral oncogene homolog (avian)) and do not alter the coding sequence of any known genes (Amundadottir et al. 2006; Haiman et al. 2007; Ali Olama et al. 2009; Gudmundsson et al. 2009; Yeger et al. 2009). The 8q24 region is currently considered the most important susceptibility region for prostate cancer. The combination of different SNPs in 8q24 accounts for approximately 8% of the 2-fold increased risk observed among first-degree relatives (Ali Olama et al. 2009).

1.1.2 Chromosomal alterations

During the past twenty years, extensive genomic analyses of prostate cancer have identified regions of frequent gains and losses in the prostate cancer genome. Losses are frequently located in the chromosomal areas 6q, 8p, 10q, 13q, 15q, 16q, 17p/q 18q, 19p/q and 22q, whereas gains, being more abundant, are located in on 1q, 3q, 7p/q, 8q and Xq (Saramäki & Visakorpi 2007). Several key regulatory genes have been mapped to these regions, namely, AR (Xq), NKX3.1 (NK3 homeobox 1) (8p21) and PTEN (phosphatase and tensin homolog) (10q23). However, many target genes remain to be identified (Visakorpi et al. 1995a; Cairns et al. 1997; He et al. 1997; Jenkins et al. 1997; Steck et al. 1997; Bowen et al. 2000). In addition to losses and
gains, a recurrent chromosomal rearrangement in chromosome 21 has been identified; this will be discussed later.

1.1.2.1 8q amplification

Gain of the long arm of chromosome 8 (q) is detected at very low frequencies (>8%) in early-stage prostate tumors. However, the frequency of 8q gain increase sharply with tumor stage and grade, with 60-90% of castration-resistant or metastatic prostate cancers eventually harboring this gain (Visakorpi et al. 1995b; Cher et al. 1994; Nupponen et al. 1998a; El Gammal et al. 2010). The gain of 8q is associated with poor prognosis for prostatectomy-treated patients (van Dekken et al. 2003; Ribeiro et al. 2006 & 2007; El Gammal et al. 2010; Barros-Silva et al. 2011). In addition, 8q genes have been implicated in the metastatic process in numerous cancer types (Liu et al. 2009b).

The overall 8q gain contains several minimal gain regions, such as 8q21, 8q22 and 8q24 (Figure 2), suggesting the presence of multiple target genes (van Duin et al. 2005; Saramäki et al. 2006). The MYC oncogene is located at 8q24 and is the most studied putative target gene for 8q gains. MYC amplification is present in one-third of CRPC tumors, but it is commonly overexpressed in primary tumors, suggesting that mechanisms other than amplification increase MYC expression (Nupponen et al. 1998a; Saramäki et al. 2001; Koh et al. 2010). In addition, MYC expression levels do not correlate with gene copy number (Savinainen et al. 2004; Gurel et al. 2008).

There are several genes that have been suggested as targets of 8q gains. Putative targets for the gain of 8q24 include EIF3H (eukaryotic translation initiation factor 3, subunit H, also known as EIF3S3 and p40), RAD21 (RAD21 homolog (S. pombe)), KIAA0196 and PSCA (prostate stem cell antigen) (Nupponen et al. 1999; Reiter et al. 2000; Porkka et al. 2004). Putative target genes located at 8q21 include TCEB1 (transcription elongation factor B (SIII), polypeptide 1 (15 kDa, elongin C)) and TPDS2 (tumor protein D52) (Porkka et al. 2002; Wang et al. 2004). The androgen-regulated gene NCOA2 (nuclear receptor coactivator 2) has been suggested to be an 8q target gene, but it is located at 8q13, an area that is not commonly amplified (Taylor et al. 2010).
1.1.3 Early events

High-grade PIN is thought to be the pre-malignant lesion preceding prostatic adenocarcinoma. PIN is characterized by cellular proliferation and cytologic changes that mimic cancer. However, the basal cell layer is still present in PIN, whereas it is absent in adenocarcinoma. Additionally, PIN does not cause elevated PSA levels (Bostwick & Brawer, 1987; Meiers et al. 2007).

Some of the genetic and epigenetic aberrations identified in cancer samples are present in PIN lesions; these may represent gatekeeper alterations for prostate cancer initiation.

GSTP1 (glutathione S-transferase pi 1) promoter hypermethylation is one of the earliest events identified in prostate cancer progression (Lee et al. 1994; Maruyama et al. 2002). GSTP1 is a detoxifying enzyme that inactivates electrophilic carcinogens by conjugating them with glutathione. As such, GSTP1 is thought to be a caretaker gene wherein its inactivation makes prostate cells more vulnerable to somatic aberrations. GSTP1 hypermethylation is detected in 70% of PIN and 90% of primary tumors and therefore is one of the most frequently observed aberrations (Meiers et al. 2007). Genetic instability is also caused by telomere shortening, which is detected in PIN lesions (Meeker et al. 2002).

APC (adenomatous polyposis coli) hypermethylation has also been reported to occur in 30% of PIN and 55-85% of prostate cancer samples (Kang et al. 2004; Yegnasubramanian et al. 2004). APC, a known tumor suppressor gene in colon
cancer, regulates a variety of downstream targets, including Wnt signaling, β-catenin and MYC (He et al. 1998; Fodde et al. 2001).

1.1.3.1 NKX3.1 down-regulation

Transcription factor NKX3.1 (NK3 homeobox 1) is a key regulator of the differentiation of the prostate epithelium. It is one of the earliest markers for prostate development during embryogenesis, and it continues to be expressed throughout all stages of prostate differentiation and adulthood. Studies indicate that NKX3.1 functions as transcriptional repressor, and number of its effects is determined by interacting with various proteins (Abate-Shen et al. 2008). Since the initial cloning of NKX3.1, it has been evident that NKX3.1 is positively regulated by androgens (He et al. 1997). In addition, expression levels of NKX3.1 are significantly reduced upon castration (Sciavolino et al. 1997). However, proximal promoter of NKX3.1 is non-responsive to androgens. Recent study suggests that AR is recruited to androgen responsive elements (AREs) located in 3´UTR of NKX3.1, leading to positive transcriptional activity (Thomas et al. 2010). In addition, it has been shown that NKX3.1 can associate with the AR promoter and inhibit AR activity, forming a negative feedback loop (Lei et al. 2006).

NKX3.1 maps to the 8p21 chromosomal region that frequently undergoes allelic loss during the early stage of prostate cancer and in high grade PIN (Emmet-Buck et al. 1995; Vocke et al. 1996; Håggman et al. 1997; Asatiani et al. 2005; Bowen et al. 2000). Mutations of the remaining NKX3.1 allele are not observed (Voeller et al. 1997; Ornstein et al. 2001), and it appears that some expression remains in prostate tumors (Gurel et al. 2010). In addition, there is evidence of epigenetic down-regulation of NKX3.1, perhaps through promoter methylation (Asatiani et al. 2005).

In vivo mouse studies have revealed that homo- or heterozygous deletion of Nkx3.1 leads to the development of PIN-like lesions (Kim et al. 2002a), and together with the loss of Pten or cdkn1b (cyclin-dependent kinase inhibitor 1B), these mice develop adenocarcinoma. As deletion of only one allele of Nkx3.1 leads to the phenotype, Nkx3.1 appears to function as a haploinsufficient tumor suppressor gene (Kim et al. 2002b; Abate-Shen et al. 2003; Gary et al. 2004). In addition, it has been demonstrated that NKX3.1 target genes are sensitive to NKX3.1 dosage in a stochastic manner (Magee et al. 2003).

Mouse studies demonstrated that Nk3.1 loss-of-function leads to an impaired response to oxidative damage (Ouyang et al. 2005). Studies with cell lines suggest that NKX3.1 plays a role in maintaining DNA integrity in prostate epithelial cells (Bowen & Gelmann, 2010). In conclusion, NKX3.1 down-regulation may predispose premalignant prostate tissue to the accumulation of mutations.
1.1.3.2 ETS translocations

Translocations are a relatively recent finding in prostate cancer. In 2005, Tomlins et al. described the fusion between the promoter region of the androgen-regulated gene TMPRSS2 (transmembrane protease, serine 2) and the truncated form of the ETS (avian erythroblastosis virus E26 homolog) transcription factor family member ERG (v-ets erythroblastosis virus E26 oncogene homolog (avian)) that results in androgen-driven overexpression of ERG.

After that initial discovery, several novel fusions have been reported, primarily between different ETS family members (ERG, ETV1, ETV4, ETV5 and ELK4) and a diverse set of 5´ upstream fusion partners, many of which are androgen-responsive. For example, fusions have been observed between ERG and SLC45A3 (solute carrier family 45, member 3), HERPUD1 (homocysteine-inducible, endoplasmic reticulum stress-inducible, ubiquitin-like domain member 1), FKBP5 (FK506 binding protein 5) and NDRG1 (N-myc downstream regulated 1) (Pflueger et al. 2011; Rubin et al. 2011). AR signaling is hypothesized to be mechanistically involved in the formation of the TMPRSS2:ERG fusion because it brings the TMPRSS2 and ERG genes in close proximity and potentially induces DNA double-strand breaks and facilitates the subsequent recombination event (Lin et al. 2009; Mani et al. 2009; Haffner et al. 2010).

The TMPRSS2:ERG fusion is the most common and appears to be an early event in prostate cancer tumorigenesis. It has been detected in 15% of high-grade PIN lesions and in 30-70% of localized prostate cancers (Tomlins et al. 2005; Perner et al. 2006 & 2007; Mosquera et al. 2008 & 2009; Saramäki et al. 2008). Other fusions are less studied but are likely present in fewer than 10% of prostate cancers (Mehra et al. 2007; Rubin et al. 2011). Attempts to identify an association between the TMPRSS2:ERG fusion and patient prognosis have produced conflicting data. While most of the studies report an association between the fusion and poor prognosis, other studies found no such clinical association or identified an association with favorable prognosis (Perner et al. 2006; Demichelis et al. 2007; Mehra et al. 2008; Attard et al. 2008a; Saramäki et al. 2008; Hermans et al. 2009; Leinonen et al. 2010; Toubaji et al. 2011). Translocations are present at the same frequency in advanced tumors as in primary tumors, indicating that clonal selection for the fusion does not occur at later stages of prostate cancer progression (Perner et al. 2007; Saramäki et al. 2008).

The functional significance of the TMPRSS2:ERG fusion has been analyzed in vivo and in vitro. Loss-of-function and gain-of-function studies in prostate cancer cell lines have revealed that TMPRSS2:ERG promotes an invasive phenotype, without an effect on proliferation (Klezovitch et al. 2008; Tomlins et al. 2008). ERG induces metalloproteinase and plasminogen activator pathway genes, such as MMP3 (matrix metallopeptidase 3), PLAT (plasminogen activator, tissue) and PLAU (plasminogen activator, urokinase), in prostate cancer cells. ERG silencing also increase expression of genes associated with differentiated luminal prostate epithelial cells, suggesting that ERG expression may be involved in maintaining de-
differentiated state of prostate cancer cells (Tomlins et al. 2008). This idea is supported with study by Yu et al., which demonstrated that ERG is direct activator of polycomb group protein EZH2 (Enhancer of zeste homolog 2), which induces de-differentiation program in the cells (Yu et al. 2010).

Although genetic rearrangement of ERG is considered to be early event in prostate cancer progression, it is rarely found from PIN lesions and thus it does not seem to be initiative event in prostate carcinogenesis (Mosquera et al. 2008; Carver et al. 2009; King et al. 2009). It has been shown that in human prostate tumors, PTEN loss is enriched for ERG rearrangement. It has been hypothesized, that PTEN loss and ERG rearrangement are concomitant events (Carver et al. 2009; King et al. 2009). In vivo, expression of TMPRSS2:ERG alone is insufficient to induce neoplasia in mice, resulting subtle phenotype of hyperplasia and PIN (Tomlins et al. 2008; Carver et al. 2009). Pten heterozygous mice develop high-grade PIN with increased proliferation, but no adenocarcinoma (Di Cristofano et al. 1998). However, in Pten null mice, Erg expression results in high-grade PIN and invasive carcinoma, suggesting that these two events cooperate in prostate tumorigenesis (Klezovitch et al. 2008; Tomlins et al. 2008; Carver et al. 2009; King et al. 2009).

1.1.4 Androgen receptor signaling

1.1.4.1 Normal functions of AR

The normal prostate is dependent on androgens, acting through the androgen receptor (AR), for its development, maintenance and physiological function (Heinlein & Chang, 2004).

AR belongs to a family of nuclear transcription factors that mediate the action of steroid hormones. Without a ligand, AR is primarily localized in the cytoplasm associated with heat-shock proteins (HSP), cytoskeletal proteins and chaperones. The AR protein contains four domains: the ligand-binding domain, the DNA-binding domain, the hinge region and the N-terminal domain. The ligand-binding domain of AR undergoes a conformational change upon binding to androgens, such as testosterone and the more potent dihydrotestosterone (DHT). This conformational change enables AR to dissociate from HSP and to interact with diverse co-regulators. These interactions facilitate the dimerization and nuclear localization of AR (Dehm & Tindall, 2007). In the nucleus, AR binds via its DNA-binding domain to specific DNA sequences, termed androgen responsive elements (ARE), in the promoter and regulatory regions of AR target genes. Together with coactivators and collaborating transcription factors, AR triggers the transcription of target genes (Hodgson et al. 2011).
1.1.4.2 AR function in prostate cancer

AR signaling is crucial to prostate cancer development as evidenced by the fact that men castrated early in life do not develop prostate cancer (Isaacs, 1994). AR signaling is maintained during all stages of prostate cancer including castration-resistant disease (Ruizeveld de Winter et al. 1994; Hobisch et al. 1995; Visakorpi et al. 1995a; Linja et al. 2001).

In the normal prostate, AR specifies the lineage-specific differentiation of prostate cells by inducing the expression of prostate-specific genes and maintaining the differentiated prostate epithelium phenotype (Heinlein & Chang, 2004). By contrast, cellular de-differentiation is observed in malignant cells. There is a shift in both AR activity and its target genes, which converts AR from a pro-differentiation to a pro-proliferation factor during prostate cancer progression (Nelson et al. 2002; Wang et al. 2009c).

One mechanism hypothesized to contribute to altered AR function is AR reprogramming by other transcription factors. Cell lineage-specific factors may have an important regulatory role in establishing the proper environment for AR recruitment to the chromatin; one such factor is FOXA1 (forkhead box A1), which acts as a pioneering transcription factor that engages chromatin prior to other transcription factors (Kaestner, 2010). FOXA1 binding sites are proximally located to AR binding sites, and FOXA1 interacts with the DNA-binding domain of AR (Gao et al. 2003; Lee et al. 2008). FOXA1 was reported to prevent AR from binding to specific AR binding sites, thereby revealing a mechanism by which altered FOXA1 expression may contribute to the regulation of AR-mediated signaling pathways in prostate cancer. Silencing FOXA1 in prostate cancer cells increased AR binding to genes involved in hormonal signaling and cell proliferation (Sahu et al. 2011). In addition, strong nuclear expression of FOXA1 protein is associated with increased prostate cancer mortality (Sahu et al. 2011; Wang et al. 2011a). ETS transcription factors may also modulate AR binding to DNA. ETS motifs are enriched at AR binding sites, and almost half of the areas in the genome that recruited AR also recruited ETS family member ERG (Massie et al. 2007; Yu et al. 2010). Overexpression of the ETS and FOXA1 transcription factors may be a mechanism for reprogramming AR to drive the expression of oncogenes rather than genes important for maintaining the differentiation status of the prostate epithelium (Yu et al. 2010; Sahu et al. 2011; Wang et al. 2011a).

1.1.4.3 AR function in CRPC

The molecular mechanisms responsible for the emergence of CRPC are not completely understood. The most consistent finding in CRPC is a universal upregulation of AR, and it is known that increased AR expression is necessary and sufficient to convert cells from an androgen-dependent to an androgen-independent state (Chen et al. 2004; Holzbeierlein et al. 2004; Stanbrough et al. 2006). High
expression levels, increased stability, and nuclear localization of AR sensitize prostate cancer cells to low androgen levels and enable cell survival in such environments (Gregory *et al.* 2001; Waltering *et al.* 2009). There are multiple mechanisms that sustain active AR signaling under androgen-depleted conditions. AR can be activated via mutations, alterations in coregulators, and increased intratumoral androgen production as well as via mechanisms leading to ligand-independent activity (Shen & Abate-Shen, 2010).

Copy number amplification of the AR gene is detected in one-third of CRPCs, but is undetectable in primary tumors, suggesting that this amplification is a by-product of hormone therapy (Visakorpi *et al.* 1995a; Koivisto *et al.* 1997; Linja *et al.* 2001). Gain-of-function mutations increase AR activity; mutations in AR are rare in untreated tumors, but their prevalence increases during disease progression. The highest frequency of AR mutations is detected in antiandrogen-treated patients, with 10-30% harboring mutated AR (Linja & Visakorpi, 2004).

Alternative splicing of AR was recently identified as a mechanism for AR activation. AR gene spans 180kb of DNA and contains 8 canonical exons. Recently, cryptic exons that locate within introns flanking the canonical exons have been described (Figure 3). First AR splice variants were discovered in 2002 when truncated AR was observed (Tepper *et al.* 2002). Since initial discovery, numerous AR splice variants have been described. Most variants lack the ligand-binding domain partially or entirely and majority also lacks the hinge region, which contains nuclear localization signal. In addition, many variants have unique sequences encoded by cryptic exones corresponding to introns 2, 3 and 3’ untranslated region (3’ UTR) of the full length AR (fl-AR) (Figure 3) (Haile & Sadar, 2011). Functional studies indicate that deletion of LBD results in a constitutively active AR (Jenster *et al.* 1991). Still, clinical relevance of the AR variants is not established. Many of the variants are expressed in both malignant and non-malignant cells but at higher levels in tumors and upon castration. In addition, the expression of certain variants is associated with poor prognosis after radical prostatectomy (Haile & Sadar 2011).

AR interacts with coregulatory proteins to assemble a productive transcriptional complex (Bennett *et al.* 2010). Cofactor imbalance has been suggested to be involved in the development of CRPC. The overexpression of AR coactivators has been reported, but only in a small number of samples (Fujimoto *et al.* 2001; Agoulnik *et al.* 2006; Mäki *et al.* 2006 & 2007; Zou *et al.* 2009). Genetic alterations affecting AR coregulators appear to be rare (Hodgson *et al.* 2011).

In androgen deprivation therapy, serum androgen levels are depleted, which induces *de novo* androgen production in prostate tumors. In an androgen-depleted environment, prostate cancer cells up-regulate the enzymes necessary for steroid synthesis. It has been suggested that CRPCs synthesize sufficient androgens to activate AR (Stanbrough *et al.* 2006; Locke *et al.* 2008; Montgomery *et al.* 2008).

It has been demonstrated *in vitro* and in mouse studies that AR negatively regulates itself via a negative feedback loop (Quarmby *et al.* 1990; Shan *et al.* 1990; Krongrad *et al.* 1991). AR binds to an AR binding site within the promoter of the AR gene in response to androgen stimulation and withdrawal. In CRPC, it has been
hypothesized that AR escapes AR-mediated regulation, leading to increased AR mRNA expression (Cai et al. 2011).

![Figure 3: Structure of AR mRNA with exones marked with numbers. Cryptic exones are indicated in gray. Structure of full length AR (fl-AR) mRNA after splicing and AR protein product. NTD, N-terminal domain; DBD, DNA binding domain; LBD, ligand binding domain; HR, hinge-region. (Adapted from Haile & Sadar, 2011)](image)

### 1.1.4.4 Therapeutic strategies of targeting AR pathway

Prostate cancer cells are dependent on AR pathway. This dependency has been the basis for treatment of prostate cancer for decades. AR functions are mediated by its ligands, testosterone and DHT. Testosterone is produced from the testicles, in response to luteinizing hormone (LH), which is synthesized in the anterior pituitary. For the synthesis and release of LH, pulsative stimulus of gonadotropin releasing hormone (GnRH), produced from hypothalamus, is needed (Conn & Crowley, 1994). Testosterone is the primary circulating androgen in men, but in the prostate it is converted to more potent DHT by intracellular enzyme 5α-reductase (Zhu & Imperato-McGinley, 2009).

Androgen synthesis in the testicle can be blocked with GnRH agonists, which induce regulatory changes and inhibit LH release. There are several GnRH agonists available, including goserelin and leuprolide. Also GnRH antagonists are available, which binds directly to the GnRH receptor in the pituitary blocking the release of LH. GnRH agonists and antagonist effectively block testicular testosterone production. However, testosterone is produced also the andrenal glands under independent control (Tammela, 2012).

Antiandrogens inhibit androgen action by binding to the ligand-binding pocket of AR and by that blocking the interaction with AR and DHT. There are several antiandrogens available, bicalutamide, flutamide and nilutamide are non-steroidal antiandrogens and cyproterone acetate is steroidal antiandrogen, which also inhibits the release of LH. However, these agents show relatively low affinity for AR as compared to DHT and they have shown to have agonistic effects in some circumstances (Tammela, 2012). MDV3100 is novel AR antagonist, currently in
clinical trials, which binds to AR with a 4- to 8 fold higher affinity than bicalutamide. It also inhibits AR nuclear translocation and DNA binding and do not seem to have agonistic activity (Tran et al. 2009). MDV3100 have shown to have antitumor activity in patients with CRPC in phase I-II studies (Scher et al. 2010).

Abiraterone acetate is another novel drug. It targets androgen biosynthesis by selectively inhibiting CYP17 (cytochrome P450, family 17), an enzyme which catalyzes androgen synthesis from its precursors. It inhibits both testicular-derived androgen production and tumor-derived androgen synthesis (Attard et al. 2008b). Abiraterone prolonged the survival of patients with metastatic prostate cancer (de Bono et al. 2011).

Despite of initial response to androgen deprivation therapy, advanced prostate cancer eventually progress to CRPC. CRPC often remain responsive to second line hormonal therapies, and it is still dependent on AR pathway (Tammela, 2012). However, none of these therapies are curative and median survival of patients with CRPC is 9 to 30 months (Kirby et al. 2011).

1.1.5 MYC

MYC is a transcription factor that plays a key role in regulating numerous cellular processes, including cell cycle progression, metabolism, ribosome biogenesis, protein synthesis, mitochondrial function, and stem cell renewal. Overexpression of MYC is detected in numerous tumor types, particularly those harboring genetic alterations (Albihn et al. 2010).

In prostate cancer, MYC up-regulation at the mRNA level has been confirmed by numerous mRNA microarrays and qPCR experiments. MYC up-regulation has been detected in a majority of primary prostate cancers (Koh et al. 2010). At the protein level, MYC is subject to tight posttranscriptional regulation (Adhikary & Eilers, 2005). However, studies on MYC protein expression are controversial due to the lack of a suitable antibody (Koh et al. 2010). MYC protein has been reported to be overexpressed in localized primary prostate tumors and metastatic tumors as well as in PIN lesions (Jenkins et al. 1997; Gurel et al. 2008).

MYC is located at 8q24, which is frequently gained in prostate cancer (Jenkins et al. 1997; Visakorpi et al. 1995b; Nupponen et al. 1998a; Cher et al. 1994). Gain of 8q24 is strongly associated with poor prognosis (Sato et al. 1999; El Gammal et al. 2010; Ribero et al. 2006, 2007; van Dekken et al. 2003; Barros-Silva et al. 2011). A gain of 8q is observed in 70% of CRPCs, and MYC amplification is present in one-third of these tumors (Saramäki et al. 2001; Nupponen et al. 1998a). However, amplification of MYC does not correlate with its expression level, indicating that amplification is not the primary mechanism for increasing MYC expression (Savinainen et al. 2004; Gurel et al. 2008). In addition, the gain of 8q24 is rare in localized prostate tumors, where MYC overexpression is common. Although none have been yet established, other mechanisms for MYC up-regulation have been
suggested, including APC inactivation (He et al. 1998; Yegnasubramanian et al. 2004; Kang et al. 2004) and FOXP3 (forkhead box P3) down-regulation (Wang et al. 2009a).

Prostate epithelial cells are transformed by overexpressing MYC (Gil et al. 2005; Williams et al. 2005). In addition, transgenic mice overexpressing Myc in prostate epithelial cells develop adenocarcinoma (Iwata et al. 2010; Ellwood-Yen et al. 2003).

MYC regulates several genes that are known to be involved in prostate cancer tumorigenesis. MYC has been suggested to negatively regulate NKX3.1 and to positively regulate EZH2 via direct or indirect mechanisms (Iwata et al. 2010; Koh et al. 2011). In addition, PIM1, whose overexpression in prostate cancer correlates with poor clinical outcome, synergizes with MYC in prostate cancer progression. PIM1 has been reported to enhance MYC transcriptional activity and to regulate several MYC target genes (Kim et al. 2010; Zippo et al. 2007; Wang et al. 2010 & 2011b). In addition, MYC tumorigenicity is in part due to its ability to regulate miRNAs, which will be discussed later (Dews et al. 2006).

1.1.6 PI3K pathway and PTEN

The PI3K (phosphatidylinositol 3-kinase) signaling pathway is important in prostate cancer progression because approximately 40% of primary and 70% of metastatic prostate cancers contain genomic alterations in this pathway (Taylor et al. 2010). PI3K pathway is driver of cell proliferation and cell survival. When pathway is activated, it creates a second messenger lipid phosphatidylinositol (3,4,5) trisphosphate (PIP3), which recruits downstream signaling cascade. AKT (v-akt murine thymoma viral oncogene) mediates signals downstream of PI3K activation. Substrates of AKT phosphorylation have crucial role in cell-cycle regulation. In addition, PI3K pathway interacts with other signaling networks (Cully et al. 2006).

One major player in PI3K signaling is the tumor suppressor PTEN, which negatively regulates the activity of the PI3K pathway. PTEN is one of the most frequently mutated tumor suppressors in human cancer. PTEN germ line mutations cause a group of autosomal dominant syndromes, called PTEN hamartoma tumor syndrome, which are characterized by increased risk for several cancers among other deficits (Hobert & Eng, 2009). Mouse models have shown that in some tissue types Pten deletion is sufficient to cause tumorigenesis alone but frequently Pten deletion contributes to tumorigenesis with other genetic alterations. PTEN is a lipid phosphatase which dephosphorylates its targets. Tumor suppressor function of PTEN requires functional phosphatase domain, however, other functions of PTEN may be important but yet less studied (Hollander et al. 2011).

PTEN is inactivated in a wide variety of cancers, including prostate cancer (Li et al. 1997; Steck et al. 1997; Salmena et al. 2008). The substrates for PTEN dephosphorylation are products (PIP2 and PIP3) of PI3K, which when
phosphorylated, activate the PI3K pathway (Maehama & Dixon, 1998). PTEN-null cells are shown to harbor constitutively activated AKT (Sun et al. 1999). AKT activates downstream targets involved in survival, proliferation, cell cycle progression, growth, migration, and angiogenesis. In addition, expression of PTEN induces cell cycle arrest, inhibits cell adhesion and motility and triggers apoptosis (Persad et al. 2000; Pourmand et al. 2007; Davies et al. 1999).

PTEN also has AKT-independent functions, such as increasing the protein stability, expression level and transcriptional activity of the tumor suppressor TP53 (tumor protein p53, also called p53) and regulating cell cycle arrest through cyclin D interactions (Chang et al. 2008a; Radu et al. 2003).

PTEN maps to the chromosomal region 10q23 that frequently undergoes copy number deletion (Lundgren et al. 1988; Cairns et al. 1997). PTEN loss has been reported to occur in a subset of high grade PIN, suggesting a role for PTEN in prostate cancer initiation (Bettendorf et al. 2008). It has been suggested that PTEN is a haploinsufficient tumor suppressor. Recent mouse studies demonstrated that even a subtle reduction in Pten expression promoted cancer susceptibility, thereby adding more complexity to the notion of haploinsufficiency (Alimonti et al. 2010). Primary tumors possess heterogeneous PTEN protein expression; however, disease progression is associated with an increased frequency of PTEN loss (Halvorsen et al. 2003; McMenamin et al. 1999; Fenic et al. 2004). The highest frequency of PTEN loss or mutation occurs in metastatic tumors, of which approximately 40% contain PTEN alterations (Pourmand et al. 2007; Taylor et al. 2010). Furthermore, dysregulation of the PTEN/PI3K pathway has been associated with resistance to anti-androgen therapies, and mouse studies have indicated a potential role for Pten in castration-resistance (Ham et al. 2009; Zhang et al. 2009).

Numerous key proteins have been hypothesized to cooperate with PTEN loss in prostate cancer progression, including TMPRSS2:ERG, NKX3.1, MYC and AR (Carver et al. 2009 & 2011; King et al. 2009; Lei et al. 2006; Kim et al. 2009). There is bidirectional crosstalk between the AR and PI3K pathways, illustrated in Figure 4. PI3K pathway activation by PTEN-deficiency represses AR signaling in mouse and also in human prostate cancers. On the other hand, when PI3K pathway is inhibited, the AR is activated. (Carver et al. 2011). AR-mediated AKT inhibition occurs via activation of PHLPP1 (PH domain and leucine-rich repeat protein phosphatase 1), which dephosphorylates AKT. PHLPP1 is a novel tumor suppressor gene located in chromosome 18q21 and often deleted in metastatic prostate cancer (Brognard et al. 2007; Gao et al. 2005; Chen et al. 2011). Activation of PHLPP1 by AR occurs via AR target gene FKBP5 (FK506 binding protein 5), which in turn stabilizes PHLPP1 (Carver et al. 2011). It has been hypothesized that PI3K pathway activation may be sufficient to compensate for the blockade in AR signaling during androgen ablation therapy, thereby promoting CRPC development (Mulholland et al. 2011).
1.1.7 Metastatic prostate cancer

Metastatic prostate cancer is the lethal form of the disease. The 5-year survival rate for prostate cancer patients with distant metastasis is 30% compared with approximately 100% for patients with localized cancer (Jemal et al. 2010). Prostate cancer metastasizes to distant organs, including the liver, lungs and brain, with a very high predilection for metastasizing to the bone (Bubendorf et al. 2000).

The formation of metastases is complex; cells must first disseminate from their primary location and degrade local stroma, then escape into and survive in the blood or lymphatic circulation, and finally hone to a metastatic location. At a distant site, the cells must induce neovascularization and proliferate to create a secondary mass (Arya et al. 2006). Prostate cancer cells have the ability to metastasize early during tumor formation as evidenced by the presence of disseminated cancer cells in the bone marrow of the majority of patients with low-grade cancer (Morgan et al. 2009). It has been reported that metastatic prostate cancer is clonally related, indicating a monoclonal origin that differs from the vast heterogeneity of primary tumors (Liu et al. 2009a).

The molecular mechanisms underlying prostate cancer metastasis are poorly understood. Cell adhesion molecules must be down-regulated to disrupt cell-cell interactions, and protease expression is necessary to degrade the surrounding stroma. E-cadherin down-regulation and N-cadherin up-regulation are associated with prostate cancer progression (Gravdal et al. 2007; Jaggi et al. 2005). Decreased cell-cell adhesion is involved in the epithelial mesenchymal transition (EMT) that leads to the development of a more invasive and migratory phenotype in epithelial cells. uPA (plasminogen activator, urokinase, also known as PLA2), a serine protease family member involved in extracellular matrix degradation, is overexpressed in a subset of prostate cancers and its expression has been shown to correlate with prostate cancer aggressiveness (Gaylis et al. 1989; Van Veldhuizen et
al. 1996; Cozzi et al. 2006). In addition, uPa amplification has been detected in prostate cancer cell lines and castration-resistant disease (Helenius et al. 2001). Promoter demethylation is another mechanism proposed for increasing uPA expression (Pulukuri et al. 2007).

Metastatic disease is treated by androgen deprivation therapy. These therapies can inhibit the disease for several years, but eventually the cancer progresses and CRPC arises. In CRPC, many AR-regulated genes, such as KLK3 (PSA) and TMPRSS2:ERG, are reactivated (Feldman and Feldman, 2001; Stanbrough et al. 2006; Cai et al. 2009). Certain molecular changes are found exclusively in CRPC due to the adaptation or clonal selection of cancer cells able to exist in low androgen levels. For example, the gain of chromosome Xq, which contains the AR gene, is specific for CRPC (Cher et al. 1996). The role of AR in CRPC is discussed in more detail under the title “AR pathway”. However, the genes other than AR that drive CRPC formation are unknown.

1.1.7.1 EZH2

EZH2 is a polycomb group protein that catalyzes the trimethylation of histone H3 lysine 27 (H3K27me3) and thereby represses gene transcription (Kirmizis et al. 2004). Many genes involved in development, stem cell maintenance, and differentiation are targets of H3K27 methylation (Sparmann & van Lohuizen, 2006).

EZH2 is frequently up-regulated in advanced prostate tumors, and one mechanism by which this occurs is through gene amplification (Varambally et al. 2002; Saramäki et al. 2006). Other mechanisms leading to EZH2 overexpression in tumors involve regulation by miR-101 (discussed later) and ERG (Varambally et al. 2008; Yu et al. 2010). ERG binds directly to the EZH2 promoter, which enhances its transcription and regulates the epigenetic silencing of EZH2 target genes (Yu et al. 2010).

EZH2 protein levels correlate with the aggressiveness of prostate cancer (Varambally et al. 2002; Laitinen et al. 2008). Target genes for EZH2-mediated epigenetic silencing include DAB2IP and E-cadherin (Chen et al. 2005; Cao et al. 2008). Loss of DAB2IP promotes primary tumor growth by activating Ras signaling and metastasis by activating NF-κB (Min et al. 2010). In addition, EZH2 regulates the expression of several miRNAs, which will be discussed later.

1.1.7.2 TP53

TP53 is a transcription factor that is commonly inactivated in tumors by mutation of TP53 or inactivation of the TP53 signaling pathway. TP53 is essential for protecting the cell from DNA damage and oncogene activation. TP53 activates the transcription of genes that promote apoptosis, senescence and cell cycle arrest. Mutated TP53 have dominant-negative activity over the remaining wild-type allele.
In addition to losing tumor suppressive functions, many TP53 mutants also gain new oncogenic properties. These include inhibitory interactions with p63 and p73, thereby reducing their transcriptional activity. Mutant TP53 is more stable than wild-type TP53 as evidenced by the nuclear accumulation of mutated protein (Brown et al. 2009).

In prostate cancer, a discrepancy exists in the frequency at which TP53 mutations are detected. However, recent large-scale studies reached concordant results in finding that TP53 mutations in primary tumors are rare at approximately 5% (Schlomm et al. 2008; Agell et al. 2008). Much higher frequencies of TP53 mutations were detected in metastatic (16%) and castration-resistant tumors (26%). In addition, TP53 mutations are independent predictors for biochemical relapse after radical prostatectomy (Schlomm et al. 2008; Visakorpi et al. 1992). Recent large-scale genomic profiling of primary prostate tumors confirmed the low frequency of TP53 mutations. In contrast, homo- and heterozygous deletions of TP53 were present in 24% of the tumors, indicating that copy number loss is the primary mechanism for altering TP53 expression (Taylor et al. 2010). In vitro and in vivo studies demonstrated that mutant TP53 induces androgen-independent growth, indicating that TP53 may have a role in the development of castration-resistant tumors (Nesslinger et al. 2003; Vinall et al. 2006).

1.1.8 miRNAs in prostate cancer

1.1.8.1 microRNA biogenesis and target recognition

miRNAs (miRNA, miR) are short (approximately 21 nucleotides), non-coding RNAs that function as endogenous negative regulators of gene expression. Figure 5 illustrates miRNA biogenesis pathway. miRNAs are transcribed as long primary molecules (pri-miRNAs). These transcripts are subsequently processed by the enzyme Drosha into precursor miRNAs (pre-miRNA). These hairpin-structured molecules are exported to the cytoplasm for further processing into short, double-stranded RNAs. One strand of this molecule represents the mature miRNA, which joins specific Argonaute proteins to form a complex called RNA-induced silencing complex (RISC). Within the RISC complex, miRNA guides the complex to the miRNA response element (MRE) located commonly in the 3’ UTR of the target mRNA (Krol et al. 2010). Individual miRNAs frequently target several transcripts, and individual genes can be targeted by multiple distinct miRNAs (Friedman et al. 2009b).

miRNAs regulate the expression of their targets by various mechanisms, including mRNA destabilization, translational repression and even activation of gene expression (Huntzinger & Izaurralde, 2011). Perfect pairing of miRNA to the target results endonucleotic cleavage of the mRNA by Argonaute (Yekta et al. 2004). However, this mechanism is rare in animals. Other mechanisms for target
destabilization exist. miRNA binding to the target mRNA by imperfect pairing can make mRNA susceptible to exonucleolytic cleavage, by recruiting deadenylation factors that remove the polyA tail of the mRNA. There are also cases when miRNA binding reduces the protein level without an effect on mRNA. Detailed mechanisms for this are not fully known, there is evidence for inhibition of translation initiation or elongation and direct proteolysis of synthesized peptide (Huntzinger & Izaurralde, 2011).

Identifying miRNA targets is challenging, mainly because of the limited complementary between miRNAs and their targets. Traditionally, miRNA binding sites are classified into three categories: 5´prime dominant sites have perfect pairing in the seed sequence located in positions 2-7 from the 5´ end of the miRNA and in addition, they have extensive base pairing in the 3´ end of the miRNA. Seed only sites have perfect pairing in the seed region and imperfect pairing in the 3´ end of the miRNA. 3´compensatory sites have a mismatch in the seed region but have a long stretch of base pairing in the 3´end of the miRNA (Min & Yoon, 2010). In addition to these three categories, also centered sites have been described, in which the middle region of miRNA makes perfect match with target mRNA (Pasquinelle, 2012).

*In silico* algorithms uses these criteria to predict putative miRNA targets. However, these algorithms produce divergent predictions with degrees of false positives and false negatives (Min & Yoon, 2010). Bioinformatics analyses are often coupled with genetic or biochemical methods in order to identify miRNA targets. Silencing or transfecting miRNA to a cell line followed by analysis of altered phenotype by quantitative real-time PCR (qRT-PCR) or microarray is commonly used. When miRNA is silenced target mRNA should be up-regulated and vise versa. Also putative target can be knocked out or mutated which should mimic the phenotype obtained by miRNA expression. These methods are, however, unable to distinguish direct and indirect targets. Putative interaction between target mRNA and miRNA can be tested by fusing sequences containing target site to reporter gene and assaying for regulation in the presence or absence of the cognate miRNA (Pasquinelle, 2012). Recently, genome wide mechanisms to study endogenous miRNA target sites have been developed. These methods use ultraviolet light to crosslink RISC bound RNA to the complex, followed by immunoprecipitation of RISC and high-throughput sequencing of the RNAs bound with RISC complex (CLIP experiments) (Chi *et al.* 2009; Hafner *et al.* 2010). However, detection of mRNA binding to the RISC does not guarantee that it is actually being regulated by miRNA (Pasquinelle, 2012).
MicroRNA biogenesis. miRNAs are transcribed by RNA polymerase II or III (Pol II/Pol III) into primary miRNA (pri-miRNA) transcripts. Pri-miRNAs are cleaved by enzyme Drosha to form pre-miRNAs, which are exported to cytoplasm by Exportin 5. In the cytoplasm, RNase Dicer cleaves the pre-miRNA hairpin to its mature length. The functional strand of the mature miRNA is loaded together with Ago proteins into RNA-induced silencing complex (RISC), where it guides RISC to silence target mRNAs (Adapted from Winter et al. 2009).

1.1.8.2 Expression profiling of miRNAs

The importance of miRNAs in cancer is highlighted by the observation that half of miRNAs are located in fragile sites in the genome, which are often altered in cancer (Calin et al. 2004). In addition, large-scale miRNA profiling has revealed a deregulation of miRNAs and the miRNA-processing machinery in cancer (Kumar et al. 2007; Calin & Croce 2006), and miRNA expression profile was demonstrated to better classify cancers than an mRNA expression profile (Rosenfeld et al. 2008).

There are several reports on altered miRNA expression in prostate cancer. However, discrepancy is rampant in these studies, most likely because of the use of different expression platforms. The results of seven different expression profiles of primary prostate cancer are summarized in Table 1. miRNAs that were implicated in at least two different studies are included. In addition, only studies of more than 10 primary cancer samples were taken into account (Volinia et al. 2006; Ambs et al. 2008; Ozen et al. 2008; Tong et al. 2009; Schaefer et al. 2010; Carlsson et al. 2011;
The most consistently down-regulated miRNAs in primary prostate tumors were miR-145, miR-205, miR-221, miR-222 and miR-24; these were all down-regulated in at least three independent studies. The up-regulated miRNAs included miR-182, miR-25, miR-375 and miR-93.

Fewer studies have been performed on advanced or metastatic disease, and the sample number is often limited. One of the most consistent findings is the down-regulation of miR-145 in metastatic (Porkka et al. 2007; Peng et al. 2011; Watahiki et al. 2011) or aggressive tumors (Wang et al. 2009b).

There is significant variation in the miRNAs that have been reported to have altered expression in prostate cancer. There are several possible reasons for the inconsistency. Sample selection and preparation may affect to the results. Freshly frozen samples are commonly used for miRNA studies, since in these samples RNA is least degraded. However, in some studies formalin-fixed, paraffin-embedded tissue specimen is used, in which RNA integrity is compromised (Tong et al. 2009; Carlsson et al. 2011; Watch et al. 2011; Evers et al. 2011). Sample storage time and conditions, mainly temperature, also alters the RNA integrity, as well as different RNA extraction methods may give varying results (Evers et al. 2011). Prostatectomy specimen are often used for miRNA expression studies, these are usually from untreated patients. When metastatic and castration-resistant disease is studied, patients have been given different kind of therapeutic agents, which can cause variation the miRNA profile, especially in miRNAs involved in the AR pathway. One important factor is the selection of controls. In some studies, normal adjacent tissue is used (Schaefer et al. 2010; Carlsson et al. 2011), some have used matched controls from the same patients (Tong et al. 2009; Watch et al. 2011) and in some studies BPH is used as a control (Porkka et al. 2007; Ozen et al. 2007).

There are several different platforms available for miRNA research; each has its own limitations. There is no gold-standard method for measuring miRNA expression. For the genome-wide miRNA expression profiling, array based methods are most commonly used. Short length of the miRNA causes challenges in array design, as the whole mature sequence is used as a probe. Consequently, melting temperatures of miRNA probes may vary >20º (Git et al. 2010). In addition, there are miRNA families, like let-7 family, in which different family members differ only by one nucleotide from each other (Boyerinas et al. 2010). It has been recently discovered that miRNA isoforms exist. miRNAs can have sequence heterogeneity, mainly deletion/addition of nucleotide in 3´ or in 5´end or internal modifications (Ryan et al. 2010). Different platforms detect these variants differentially; in example Agilent miRNA microarray probes have loop in 3´end, which does not allow heterogeneity in 3´end (Wang et al. 2007). It should be also noted that annotations of mature miRNA sequences are updated frequently, according to the miRBase registry (Kozomara & Griffiths-Jones, 2011). This may explain some of the inconsistencies, especially when comparing older miRNA expression studies to the newer studies.
Table 1: Combined results of 7 independent miRNA expression profiling of primary prostate tumors

<table>
<thead>
<tr>
<th>miRNA ID</th>
<th>Volinia</th>
<th>Ambs</th>
<th>Ozen</th>
<th>Tong</th>
<th>Schaefer</th>
<th>Carlsson</th>
<th>Wach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Let-7b</td>
<td>down</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>let-7d</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>let-7i</td>
<td>up</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-101</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-106a</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-106b</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-125a</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-125b</td>
<td>down</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-128a</td>
<td>down</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-141</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-143</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-145</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-149</td>
<td>down</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-16</td>
<td>up</td>
<td>down</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-17</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-181b</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-182</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-191</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-195</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-196-1</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-200b</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-200c</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-205</td>
<td>down</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-20a</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-21</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-214</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-221</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-222</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-24</td>
<td>down</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-25</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-26a</td>
<td>up</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-29a</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-30a</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-30b</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-30d</td>
<td>down</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-30e</td>
<td>down</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-31</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-32</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-345</td>
<td>down</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-34a</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-375</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-92-2</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-93</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>miR-99b</td>
<td>up</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: up, up-regulated; down, down-regulated; blank, regulation was insignificant or miRNA was not included in the study. (Volinia et al. 2006; Ambs et al. 2008; Ozen et al. 2008; Tong et al. 2009; Schaefer et al. 2010; Carlsson et al. 2011; Wach et al. 2012)
1.1.8.3 Tumor suppressor miRs

There are few miRNAs that have been confirmed to be functionally involved in prostate cancer tumorigenesis. miRNAs are classified as tumor suppressors when they target oncogenes and have decreased expression in cancer.

The miRNA that has been reported most frequently to be down-regulated in prostate cancer is miR-205. When miR-205 expression was restored, prostate cancer cells underwent morphological changes and exhibited impaired migratory and invasive capabilities. The expression levels of E-cadherin and miR-205 correlated, and the serine/threonine kinase protein kinase C epsilon (PKCε) was identified as a target of miR-205. Interestingly, the same phenotype was observed after PKCε silencing and after miR-205 restoration (Gandellini et al. 2009). PKCε have been linked to the regulation of diverse cellular functions, including cell growth, migration and malignant transformation (Griner & Kazanietz, 2007). In prostate cancer cells, PKCε expression has shown to be sufficient for transforming androgen dependent cells into androgen independent (Wu et al. 2002). In addition, PKCε expression correlates with prostate cancer aggressiveness (Aziz et al. 2007). miR-205 was also shown to up-regulate the tumor suppressor genes IL24 and IL32 by binding to their promoters (Majid et al. 2010).

miR-145 has consistently been reported to be down-regulated in primary and metastatic prostate cancer. It has been demonstrated that miR-145 is expressed in response to stress via the PI3K/AKT and TP53 pathways. TP53 induces miR-145 expression by interacting with the miR-145 promoter (Sachdeva et al. 2009; Suh et al. 2011). In prostate cancer cells that overexpress miR-145, increased cell cycle arrest and apoptosis are observed (Zaman et al. 2010). One mechanism for inactivating miR-145 in prostate cancer cell lines is promoter methylation (Suh et al. 2011; Zaman et al. 2010). In colon and breast cancer cell lines, miR-145 directly targets the oncogene MYC and is therefore the mediator of the TP53-induced post-transcriptional regulation of this particular oncogene (Sachdeva et al. 2009). In prostate cancer, the pro-apoptotic gene TNFSF10 (tumor necrosis factor (ligand) superfamily, member 10) and the actin-binding protein SWAP70 (SWAP switching B-cell complex 70kDa subunit) have been suggested as targets of miR-145 (Zaman et al. 2010; Chiyomaru et al. 2011).

The miR-34 family has demonstrated tumor suppressive function in prostate cancer. miR-34a is down-regulated in CD44-positive cells, which are considered to be prostate cancer stem cells. Reduced expression of miR-34a in prostate cancer stem cells contributes to prostate cancer development and metastasis by regulating CD44 expression and their migratory, invasive and metastatic potential (Liu et al. 2011). In addition, miR-34 family members have been identified as direct targets of TP53 that induce cell cycle arrest and apoptosis (Hermeking, 2010). Down-regulation of miR-34c has been linked to prostate cancer aggressiveness (Hagman et al. 2010).

miR-15a and miR-16 are transcribed as a cluster from genomic region 13q14, which is a region of frequent allelic loss in prostate cancer. Loss of 13q correlates
with tumor progression, and up to 90% of advanced tumors exhibit this particular loss (Hyytinen et al. 1999; Dong et al. 2001). The miR-15a-miR-16-1 cluster is commonly deleted in advanced prostate cancer; homozygous deletion has even been observed (Bonci et al. 2008; Porkka et al. 2011). In addition, down-regulation of miR-15a and miR-16 is detected in approximately 80% of tumor samples. Knockdown of miR-15a and miR-16 in prostate cancer cell lines and in mice resulted in enhanced proliferation, survival and invasion, indicative of their tumor suppressor function. BCL2 (B-cell CLL/lymphoma 2), CCND1 (encoding cyclin D1) and WNT3A (wingless-type MMTV integration site family, member 3A) have been hypothesized as target genes for miR-15a and miR-16 (Bonci et al. 2008; Cimmino et al. 2005). Recently, it was reported that the expression of miR-15 and miR-16 is decreased in cancerous stroma, and novel targets that mediate crosstalk between the stroma and the epithelium, such as FGF2 (fibroblast growth factor 2) and FGFR (fibroblast growth factor receptor), were identified (Musumeci et al. 2011).

miR-101 regulates histone methylation by targeting EZH2 (the role of EZH2 is discussed above). Overexpression of miR-101 attenuates cell migration, invasion and anchorage-independent growth. Expression of miR-101 is decreased in metastatic prostate cancers, and copy number loss of miR-101 has been identified in 66% of metastatic prostate cancers, which leads to upregulation of EZH2 (Varambally et al. 2008; Friedman et al. 2009a). In addition, EZH2 regulates miRNA expression by epigenetic silencing, repressing targets such as miR-181a, miR-181b, miR-200b, miR-200c and miR-203. Overexpression of these miRNAs in prostate cancer cells attenuates cell invasion and anchorage-independent growth (Cao et al. 2011). miR-449 has also been reported to influence epigenetic regulation by targeting histone deacetylase 1 (HDAC1), which is often up-regulated in prostate cancer (Noonan et al. 2009; Weichert et al. 2008). Decreased expression of miR-449 is observed in prostate cancer, and reintroducing this miRNA into prostate cancer cells resulted in growth arrest and apoptosis (Noonan et al. 2009).

1.1.8.4 OncomiRs

Oncogenic miRNAs (oncomiRs) target tumor suppressor genes, and their expression is up-regulated in cancer. In prostate cancer, oncomiRs have been studied less than tumor suppressor miRNAs.

The miRNA cluster encompassing miR-106b, miR-93, and miR-25 is transcribed together from the same genomic locus and targets the tumor suppressor PTEN (Poliseno et al. 2010). This miRNA cluster is located within the intron of the gene MCM7, which is overexpressed and amplified in prostate cancer (Ren et al. 2006). It has been suggested that these miRNAs cooperate with MCM7 in its tumorigenic functions. When mcm7 and the miRNA cluster were co-overexpressed, transgenic mice developed highly penetrant PIN with an onset and features similar to Pten
heterozygous mice. In contrast, MCM7 overexpression alone is not sufficient to drive oncogenesis in vitro or in vivo (Poliseno et al. 2010).

Several oncogenic functions of MYC can be explained by the activation or suppression of miRNAs (Dews et al. 2006; Chang et al. 2008b). MYC activates the miR-17-92 cluster, comprised of six miRNAs (miR-17, miR-18a, miR-19a, miR-20a, miR-19b-1, and miR-92a-1), that promotes cell proliferation and induces angiogenesis (Dews et al. 2006). Many of these miRNAs have been reported to be up-regulated in prostate cancer, including miR-17, miR-20a and miR-92 (Volinia et al. 2006; Ambs et al. 2008; Wach et al. 2012). In contrast, induction of MYC results in widespread repression of miRNA expression because MYC directly binds to the promoters or enhancers of target miRNAs. Some of these miRNAs are down-regulated in prostate cancer, including miR-15a, miR-16 and miR-34a (Bonci et al. 2008; Chang et al. 2008b; Hagman et al. 2010).

1.1.8.5 miRNAs and AR

AR transcriptionally regulates the expression of target miRNAs and is subject to miRNA-mediated gene silencing.

miR-21 is involved in pathways critical for controlling tumorigenesis and is overexpressed in several tumor types, including prostate cancer (Pan et al. 2011). For example, miR-21 is induced by AP-1 (adaptor-related protein complex 1) in response to expression of the Ras oncogene (Talotta et al. 2009). In addition, several tumor suppressors are miR-21 targets, including PTEN and BCL-2 (Pan et al. 2011). In prostate cancer, ANP32A (acidic (leucine-rich) nuclear phosphoprotein 32 family, member A) and SMARCA4 (SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily a, member 4), tumor suppressor genes involved in chromatin remodeling, are validated miR-21 targets (Schramedei et al. 2011). miR-21 is induced by androgens in androgen-responsive prostate cancer cell lines; however, the expression of miR-21 is also elevated in AR-negative cell lines. AR up-regulates miR-21 by binding directly to the miR-21 promoter. In mice, miR-21 overexpression improved tumor establishment, increased tumor growth and induced a castration-resistant phenotype (Ribas et al. 2009).

Additional miRNAs that have been suggested to be androgen-regulated are miR-125b, miR-141, miR-338 and miR-221. miR-125b was discovered to be overexpressed in AR-positive cell lines and is up-regulated in a subset of primary prostate tumors. Transfection of miR-125 enhanced cell growth in the absence of androgens, and AR was demonstrated to bind to the miR-125b promoter. The BCL-2 family member BAK1 (BCL2-antagonist/killer 1) was identified as a target of miR-125b (Shi et al. 2007). miR-141 is regulated by androgens in prostate cancer cell lines and xenografts. Transfection of miR-141 enhanced the growth of prostate cancer cells under low androgen conditions. In addition, miR-141 is overexpressed in both primary and castration-resistant tumors (Waltering et al. 2011). Androgens up-regulated miR-338 and down-regulated miR-221 in androgen-responsive prostate
cancer cells. However, these miRNAs have not been further studied (Ambs et al. 2008).

miRNAs that target AR are less studied. In a large-scale screen, miR-34a and miR-34c were reported to directly regulate AR by binding to its 3’UTR. In addition, the expression levels of miR-34a and miR-34c inversely correlate with AR expression in clinical samples (Östling et al. 2011). miR-34 family members are considered to have tumor suppressive functions (discussed above in more detail).

1.2 Experimental models of prostate cancer

1.2.1 Cell lines

Human prostate cancer derived cell lines have been used extensively to study prostate cancer, particularly in terms of genetic deregulation. Today, there are more than hundred prostate cancer cell lines available (Prostate cancer cell line database, capcelllines.ca), most of them been derivatives from other cell lines. The characteristics of most commonly used cell lines are summarized in Table 2. Only the cell lines that appear to be adenocarcinoma are included.

Two cell lines 22Rv1 and PC-346C are established from primary tumors. LAPC-4 and LNCaP are from lymph node metastasis, MDA PCa 2a/2b, PC-3 and VCaP are from bone metastasis. These sites are most common and earliest sites for prostate cancer metastasis. Prostate cancer cells metastasize less frequently to brain or dura mater, which are the sites where DU145 and DuCaP cells are derived, respectively (van Bokhoven et al. 2004).

One important feature of the cell lines are their androgen dependency. There are three types of prostate cancer cell lines: (1) Androgen dependent cell lines which require androgens for their growth, (2) androgen responsive cell lines which do not require androgens but show growth response in presence of androgens and (3) androgen independent cell lines that do not require androgens and do not response to androgens. LAPC-4 and PC-346C are most androgen dependent cell lines, of these; LAPC-4 requires the addition of androgens to its growth medium. LNCaP, 22Rv1, MDA PCa 2a and 2b, DuCaP and VCaP are androgen responsive cell lines. These cell lines express AR and many of the AR target genes, including PSA. However, many of the cell lines have mutation in AR ligand binding domain, which may alter to the response to other steroids. LAPC-4, PC-346C, DuCaP and VCaP have intact AR, VCaP and DuCaP also have amplification of the gene. DU145, MDA PCa 1 and PC-3 cells do not express AR and are androgen independent (van Bokhoven et al. 2004; Saramäki et al. 2008).

Prostate cancer is very heterogenic disease and also cell lines derived from prostate tumors show heterogenic DNA profile. Some of the cell lines exhibit some chromosomal abnormalities, which are not commonly found from prostate cancer.
For example, LAPC-4 has a loss of chromosome 16 and deletion at chromosome 12p12, and some cell lines have lost Y-chromosome (PC-3, VCaP) (Sobel & Sadar 2005a & 2005b). Cell lines often lack the common DNA alteration found from clinical tumors and that should be taken into consideration when selecting cell line for different research purposes. VCaP and DuCaP harbors common rearrangement found from prostate tumors, TMPRSS2:ERG (Tomlins et al. 2005). In addition, PC-3 is the only cell line that has 8q amplification, which is found from high percentage of advanced prostate tumors (Nupponen et al. 1998b).

1.2.2 Xenografts

In addition to cell lines, xenograft models have been widely used in prostate cancer studies. Xenografts are created by transplanting human prostate tumors into nude mice. As they are human origin, xenografts are more likely to possess molecular events involved in prostate tumorigenesis than cell lines grown in flask. Xenografts are especially useful in studying chemotherapeutic responses, as well as genetic and molecular mechanisms leading to cancer. Xenografts allow the amplification of small amount of starting material and also enrichment of homogenous cell population from heterogeneous tumor sample (Sobel & Sadar, 2005b).

Two commonly used xenograft series are called the LuCaP and PC xenografts. Some characteristics of different xenografts are summarized in Table 2. Most of the xenografts do not grow in vitro, and they are maintained by serial passage in mice. They represent heterogeneity of human prostate cancer, some are derived from primary tumors, still majority of them are derived from different metastatic sites. From some of the androgen dependent xenografts androgen-independent form is developed. This is done by implanting primary xenografts into castrated mice. These xenograft series are particularly useful of studying mechanisms to castration-resistance (Corey & Vessella 2007; van Weerden et al. 2009).

1.2.3 Mouse models

Mice do not spontaneously develop prostate carcinoma. In addition, mouse prostate anatomy is different from human prostate anatomy. Still, a lot of work has been done to manipulate mouse so that they develop prostate cancer that mimics human disease (Valkenburg & Williams, 2011). A few of the most commonly used prostate cancer mouse models are shortly summarized here.

The first generation prostate cancer models used transgenes that overexpress viral oncogenes, resulting aggressive disease that often leads to metastatic cancer. Among these models is well-studied TRAMP (transgenic adenocarcinoma of the prostate). TRAMP express both large and small SV40 antigen by prostate specific probasin promoter. Similarly, LADY model was created using probasin promoter, however,
driving only the expression of large SV40 antigen. TRAMP model develop aggressive prostate cancer very fast, in LADY model cancer is more slowly growing (Valkenburg & Williams, 2011). However, mice that overexpress SV40 large T-antigen develop cancer with features of neuroendocrine differentiation, which makes these models clinically suitable only a small population of patients which develop prostate cancer of neuroendocrine origin (Chiaverotti et al. 2008).

Second generation of prostate cancer models has used loss-of function mutations in candidate genes implicated in prostate cancer. These popular models have employed mutations in genes of interest, including Nkx3.1 and Pten. Heterozygotic Pten knock-out develops PIN. Also Nkx3.1 knockouts develop hyperplasia and PIN but no tumors are detected (Valkenburg & Williams, 2011). Several double knock-outs have been reported, in which loss of Pten with other tumor suppressors leads to development of more aggressive prostate cancer, and these tumors suppressors are thought to cooperate in prostate cancer progression. These include Nkx3.1, p53 and p27 (di Cristofano et al. 2001; Kim et al. 2002b; Couto et al. 2009).

One of the newest models for prostate cancer is TMPRSS2:ERG transgenic mouse. Fusion transcript is expressed under ARR2-Probasin promoter, which drives the expression prostate specifically. TMPRSS2:ERG mice do not develop prostate cancer, nor PIN. However, when TMPRSS2:ERG mice are crossed with Pten +/-mice, PIN and low grade adenocarcinoma is detected (Carver et al. 2009; King et al. 2009). The development of this kind of models will shed more light into function of oncogenes and tumor suppressor genes that are initially found from clinical prostate cancers, whether these genes are involved in the initiation, progression or invasion of prostate cancer. However, several limitations of mouse models should be considered when attempting to extrapolate research conclusions from mice to man.
Table 2: Summary of most commonly used cell line and xenograft models of prostate cancer

<table>
<thead>
<tr>
<th>Cell lines</th>
<th>Derived from</th>
<th>AR status</th>
<th>Important features</th>
</tr>
</thead>
<tbody>
<tr>
<td>LNCaP</td>
<td>Lymph node</td>
<td>Androgen responsive, express AR (mutated) and PSA</td>
<td>PTEN mutation</td>
</tr>
<tr>
<td>LAPC-4</td>
<td>Lymph node</td>
<td>Androgen responsive, express AR (WT) and PSA</td>
<td>p53 (mutated). Tumorigenic</td>
</tr>
<tr>
<td>PC-346C</td>
<td>primary PCa</td>
<td>Androgen responsive, express AR (WT) and PSA</td>
<td>Composed of 2 subpopulations of cells. WT p53. Low number of chromosomal abnormalities</td>
</tr>
<tr>
<td>22Rv1</td>
<td>primary PCa</td>
<td>Androgen responsive, express AR (mutated) and PSA</td>
<td>p53 (mutated). Do not form colonies in soft agar</td>
</tr>
<tr>
<td>DuCaP</td>
<td>Dura mater tissue</td>
<td>AR amplification, Androgen responsive, express AR (WT) and PSA</td>
<td>Contain mouse stromal cells. TMPRSS2:ERG translocation, p53 and Rb pos. Tumorigenic</td>
</tr>
<tr>
<td>VCaP</td>
<td>Bone metastasis from the same patient than DuCaP.</td>
<td>AR amplification, Androgen responsive, express AR (WT) and PSA</td>
<td>TMPRSS2:ERG translocation, ERG pos. Express p53 and Rb. Tumorigenic</td>
</tr>
<tr>
<td>MDA PCa 1</td>
<td>ascites fluid</td>
<td>Ai</td>
<td>Highly tumorigenic</td>
</tr>
<tr>
<td>MDA PCa 2a</td>
<td>Bone</td>
<td>Ai</td>
<td>WT p53, p21, Bcl-2, Bax and Rb. Tumorigenic.</td>
</tr>
<tr>
<td>MDA PCa 2b</td>
<td>Bone, from the same patient than MDA PC 2a</td>
<td>Androgen responsive</td>
<td>WT p53, p21, Bcl-2 and Rb Tumorigenic</td>
</tr>
<tr>
<td>PC-3</td>
<td>Lumbar vertebrae</td>
<td>Ai, AR neg</td>
<td>8q amplification. Tumorigenic. PTEN deletion</td>
</tr>
<tr>
<td>DU145</td>
<td>Brain</td>
<td>Ai, AR neg</td>
<td>Tumorigenic</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Xenografts</th>
<th>Derived from</th>
<th>AR status</th>
<th>Important features</th>
</tr>
</thead>
<tbody>
<tr>
<td>LuCaP 23 series</td>
<td>23.1 &amp; 23.8 lymph node, 23.12 liver metastasis, all from the same patient</td>
<td>All androgen sensitive, express PSA</td>
<td>Androgen independent form of LuCaP 23.1, maintained in castrated mice</td>
</tr>
<tr>
<td>LuCaP 23.1 Ai</td>
<td>LuCaP 23.1 Ai</td>
<td>All androgen sensitive, express PSA</td>
<td>Androgen independent form of LuCaP 23.1, maintained in castrated mice</td>
</tr>
<tr>
<td>LuCaP 35</td>
<td>Lymph node</td>
<td>WT AR, androgen response</td>
<td>8p deletion, PTEN neg</td>
</tr>
<tr>
<td>LuCaP 35V</td>
<td>LuCaP 35</td>
<td>Ai, express high levels of AR</td>
<td>Ai variant of LuCaP 35</td>
</tr>
<tr>
<td>LuCaP 49</td>
<td>Omentum fat</td>
<td>Ai, AR neg, PSA neg</td>
<td>neoendocrine/small cell PCa</td>
</tr>
<tr>
<td>LuCaP 58, 78, 81, 115</td>
<td>Lymph node</td>
<td>AR &amp; PSA pos.</td>
<td></td>
</tr>
<tr>
<td>LuCaP 73, 93, 96</td>
<td>primary tumor</td>
<td>LuCaP 73 &amp; 96 are androgen sensitive, AR pos. LuCaP 93 AR neg, Ai</td>
<td></td>
</tr>
<tr>
<td>LuCaP 96 Ai</td>
<td>LuCaP 96</td>
<td>Ai form of LuCaP 96</td>
<td></td>
</tr>
<tr>
<td>LuCaP 70</td>
<td>Liver</td>
<td>Express AR and PSA</td>
<td></td>
</tr>
<tr>
<td>LuCaP 77</td>
<td>Femur</td>
<td>Androgen sensitive, express AR and PSA</td>
<td></td>
</tr>
<tr>
<td>LuCaP 86.2</td>
<td>Bladder</td>
<td>Express AR and PSA</td>
<td></td>
</tr>
<tr>
<td>LuCaP 105</td>
<td>Rib metastasis</td>
<td>Express AR and PSA</td>
<td></td>
</tr>
<tr>
<td>PC-82, EW, 310, 329</td>
<td>prostate</td>
<td>Androgen dependent, express AR and PSA</td>
<td>TMPRSS2:ERG fusion, PC-82 PTEN mutation</td>
</tr>
<tr>
<td>PC-135</td>
<td>prostate</td>
<td>Ai</td>
<td></td>
</tr>
<tr>
<td>PC-133</td>
<td>bone</td>
<td>Ai, AR neg</td>
<td>PTEN deletion</td>
</tr>
<tr>
<td>PC-295</td>
<td>lymph node</td>
<td>Androgen dependent, AR pos</td>
<td>PTEN mutation</td>
</tr>
<tr>
<td>PC-324, 339, 346, 346B</td>
<td>TURP</td>
<td>PC-324 and 339 Ai, PC-346 and 346B androgen responsive</td>
<td>in PC-346 PTEN mutation</td>
</tr>
<tr>
<td>PC-346</td>
<td>PC-346</td>
<td>Ai derivative of 346</td>
<td></td>
</tr>
<tr>
<td>PC-346B</td>
<td>PC-346B</td>
<td>Ai derivative of 346B</td>
<td></td>
</tr>
<tr>
<td>PC-374</td>
<td>skin</td>
<td>androgen responsive, express AR and PSA</td>
<td>TMPRSS2:ETV1 gene fusion, PTEN mutation</td>
</tr>
</tbody>
</table>

PCa, prostate cancer; AR, androgen receptor; Ai, androgen independent; PSA, prostate specific antigen; TURP, transurethral resection of the prostate; PTEN, phosphatase and tensin homolog (Vlietstra et al. 1998; Sobel & Sadar 2005a & 2005b; Hermans et al. 2006; Corey & Vessella 2007; Saramäki et al. 2008; van Weerden et al. 2009)
AIMS OF THE STUDY

The aim of this thesis was to identify and evaluate novel protein-coding genes and miRNAs that are involved in prostate cancer progression. The specific aims were as follows:

1. To study the function of the putative 8q amplification target genes EIF3H, RAD21, KIAA0186 and TCEB1 in prostate cancer cell lines.
2. To study the miRNA expression profile in clinical prostate cancer samples to identify prognostically significant miRNAs.
3. To identify and functionally characterize androgen-regulated miRNAs that are differentially expressed in castration-resistant prostate cancer.
MATERIALS AND METHODS

1.1 Cell lines and clinical samples (I, II, III)

PC-3, DU145, NIH 3T3, HEK 293T/17 and LNCaP cell lines were obtained from American Type Culture Collection (ATCC, Manassas, VA, USA) and cultured under the recommended conditions. LNCaP-derived AR-overexpressing cells were created by transfecting LNCaP cells with empty pcDNA3.1 vector (LNCaP-pcDNA3.1) or with pcDNA3.1 containing a cDNA corresponding to wild-type AR (LNCaP-ARhi overexpresses AR by 4- to 5-fold and LNCaP-ARmo overexpresses AR 2- to 3-fold) (Waltering et al. 2009).

The clinical samples used in original communications I, II and III are summarized in Table 3. In total, 51 (BPH, PC, and CRPC) freshly frozen clinical specimens were used for quantitative RT-PCR in original communication I. Fifty-four (5 BPH, 28 PC, 7 BPH-TURP and 14 CRPC) freshly frozen clinical samples were used for the microarray in original communication III. These samples partially overlapped and were obtained from Tampere University Hospital (TAUH). The microarrays in communication II were performed on 102 samples (11 NAP, 50 PC, 3 LN-normal, 12 LN-PC, 4 BPH-TURP and 22 PC-TURP) obtained from the tissue bank of Erasmus University Medical Center. All samples were snap-frozen in liquid nitrogen and were confirmed histologically by H&E staining to contain more than 60% cancerous or hyperplastic tissue. The non-tumor samples contained 0% tumor cells. NAP samples were considered “normal” in study II, even though they were derived from prostates with adjacent tumor epithelium.

Altogether, 264 (170 PC and 94 CRPC) formalin-fixed, paraffin-embedded tissue microarray (TMA) samples were obtained from TAUH and used for immunohistochemistry (described in original communication III).

The use of clinical material was approved by the Erasmus MC Medical Ethics Committee (study II) or the Ethical Committee of the Tampere University Hospital and the National Authority for Medicolegal Affairs (study III).
Table 3: Clinical samples used in the study

<table>
<thead>
<tr>
<th>Type of sample</th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IHC</th>
</tr>
</thead>
<tbody>
<tr>
<td>BPH</td>
<td>9</td>
<td></td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>NAP</td>
<td></td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC</td>
<td>30</td>
<td>50</td>
<td>28</td>
<td>170</td>
</tr>
<tr>
<td>BPH-turp / TURP-normal</td>
<td>4</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TURP-PCa</td>
<td></td>
<td>22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRPC</td>
<td>12</td>
<td></td>
<td>14</td>
<td>94</td>
</tr>
<tr>
<td>LN-norm</td>
<td></td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LN-PC</td>
<td></td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pT1/T2</td>
<td>17</td>
<td>16</td>
<td>13</td>
<td>116</td>
</tr>
<tr>
<td>pT3/T4</td>
<td>13</td>
<td>34</td>
<td>9</td>
<td>52</td>
</tr>
<tr>
<td>not known</td>
<td>12</td>
<td></td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Gleason score &lt;7</td>
<td>15</td>
<td>35</td>
<td>4</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>17</td>
<td>10</td>
<td>81</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>32</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>not known</td>
<td>18</td>
<td>11</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>

BPH, benign prostatic hyperplasia; NAP, normal adjacent tissue; PC, primary prostate cancer; BPH-turp / TURP-normal, non-malignant transurethral resection of the prostate; TURP-PCa, malignant transurethral resection of the prostate; CRPC, castration resistant prostate cancer; LN-norm, normal lymph node; LN-PC metastatic lymph node

1.2 Microarrays (I, II, III)

1.2.1 miRNA microarray (II, III)

miRNA microarrays were purchased from Agilent Technologies (Santa Clara, CA, USA). Human miRNA V2 microarrays containing probe sets against 723 human miRNAs were utilized to analyze miRNA expression in two distinct sets of clinical samples (described in original communications II and III).

RNA was extracted from freshly frozen clinical samples by TRI-reagent (Molecular Research Center Inc., Cincinnati, OH, USA) and the integrity was assessed by RNA-nano chip run with 2100 Bioanalyzer (Agilent Techn.) (Figure 6).

miRNA microarrays were performed according to the manufacturer’s protocol. Briefly, total RNA was labeled using the miRNA Complete Labeling and Hyb Kit (Agilent Techn.) and hybridized to the microarrays. Microarrays were scanned with a DNA Microarray Scanner BA, and the data were extracted using Feature
Extraction software (v. 10.7.1.1) and analyzed using GeneSpring GX 10 (all from Agilent Techn.). Data were normalized using quantile normalization. miRNAs that were assigned a present call of at least 50% in any one of the six (original communication II) or four conditions (original communication III) were subjected to differential expression analysis. Welch’s T-test followed by correction by the Benjamini and Hochberg method with a corrected p-value cut-off of <0.05 and fold change cut-off >1.5 was used to identify differentially expressed miRNAs. Hierarchical clustering was performed using the standard uncentered correlation and average linkage methods.

The array data from the 102 clinical samples described in original communication II and from the 54 clinical samples described in original communication III were submitted to the ArrayExpress database with accession numbers E-TABM-794 and E-MTAB-408, respectively.

**Figure 6:** Two examples of RNA integrity of freshly frozen clinical samples. Prostatectomy specimen on the left, TURP specimen on the right. Ribosomal RNA peaks are marked with 18S and 28S. TURP, transurethral resection of the prostate

### 1.2.2 cDNA microarray (I, III)

Gene expression analysis was performed using Agilent’s Whole Human Genome Oligo Microarray (44k) platform.

PC-3 cells transduced with shLUC or shTCEB1 (original communication I) and LNCaP cells transfected with pre-miR-32, pre-miR-148a, or pre-miR-control (original communication III) were subjected to microarray expression analysis. Three days after transduction or transfection, the cells were lysed in Trizol (Invitrogen) or TRI-reagent (Molecular Research Center Inc.), and RNA was extracted according to the manufacturer’s instructions. RNA (500 ng) was amplified, labeled using the Low RNA Input Linear Amplification PLUS, Two-Color Kit (Agilent Techn.) (original communication I) or the Quick Amp Labeling Two-Color Kit (Agilent Techn.) (original communication III) and hybridized to the microarrays. Microarrays were scanned using Agilent’s DNA Microarray Scanner,
and the data were extracted using Agilent’s Feature Extraction software. Data were normalized either in Feature Extraction software using the “linear and lowess method” (original communication I) or in GeneSpring GX 10 (Agilent Technol.) (original communication III).

Microarray data were deposited in the ArrayExpress database with accession numbers EMEXP-1627 (original communication I) and E-MEXP-2943 (original communication III).

1.3 Real-time quantitative RT-PCR

1.3.1 mRNA expression analysis (I, III)

For mRNA expression analysis, first-strand cDNA synthesis from total RNA was performed with AMV reverse transcriptase (FinnZymes, Espoo, Finland) or Superscript II Reverse Transcriptase (Invitrogen Life Tech.) according to the manufacturer’s instructions. The expression levels of RAD21, KIAA0196, EIF3S3 and TCEB1 were analyzed using a Light-CyclerTM apparatus (Roche Diagnostics, Mannheim, Germany) with the LC Fast Start DNA SYBR Green I Kit (Roche Diagnostics). The expression of BTG2 and PIK3IP1 was analyzed using Maxima™ SYBR Green qPCR master mix (Fermentas, Burlington, Canada) and a CFX96™ qPCR machine (Bio-Rad Laboratories). The expression of each target gene was normalized to TBP expression.

1.3.2 miRNA expression analysis (II, III)

qRT-PCR analysis of miRNA expression was performed on a CFX96™ qPCR machine (Bio-Rad Laboratories, Hercules, CA, USA) (III) or an ABI Prism 7700 Sequence Detection System (Applied biosystems) (II). The TaqMan microRNA Assay (Applied Biosystems, Foster City, CA) was utilized to evaluate miRNA expression according to the manufacturer’s recommendations. Target gene expression was normalized to RNU44 (III) or to U6 (II) expression.

1.3.3 Chip-qPCR (III)

LNCaP-pcDNA3.1 and LNCaP-ARhi cells were treated with 1 nM DHT or ethanol for 2 hours. Cells were fixed (Merck KGaA, Darmstadt, Germany), pelleted and lysed. Chromatin was immunoprecipitated with 10 µg rabbit IgG (Santa Cruz Inc., Santa Cruz, California, USA) or 10 µl AR antibody (provided by Professor Olli
Jänne, Helsinki University, Finland), and qPCR was performed using Maxima™
SYBR Green qPCR master mix (Fermentas) and a CFX96™ qPCR machine (Bio-
Rad Laboratories) with primers for the putative ARBS in the proximity of miR-32
and miR-148a. The relative enrichment was calculated according to the delta Ct
method.

1.4 Immunohistochemistry (III)

In total, 264 clinical samples were utilized to evaluate BTG2 and PIK3IP1 protein
expression in prostate cancer. BTG2 and PIK3IP1 were detected in formalin-fixed,
paraffin-embedded TMA sections. First, deparaffinized TMA sections were pre-
treated by autoclaving in 10 mM sodium citrate buffer (pH 6.0) at 121°C for 2 min.
BTG2 and PIK3IP1 were detected with rabbit polyclonal antibodies (BTG2, Sigma-
Aldrich, 1:1000 dilution; PIK3IP1, S-14, Santa Cruz Biotech., 1:300 dilution) using
the PowerVision+™ Poly-HRP IHC Detection Kit (Immunovision Co.). The
sections were counterstained with hematoxylin. The staining intensity was scored in
a blinded fashion.

1.5 Western blotting (I, III)

Protein extract (20 µg) was used for Western blot analysis of TCEB1, RAD21,
BTG1, PIK3IP3 and BIM. Proteins were separated by SDS-PAGE on 8% (RAD21),
10% (PIK3IP1 and BIM) or 12% (BTG2 and TCEB1) polyacrylamide gels and
transferred to PVDF membranes (Immobilon-P, Millipore Corp., Billerica, MA,
USA). The membranes were incubated overnight with antibodies against RAD21
(1:500 dilution, Novus Biologicals, Littleton, CO, USA), TCEB1 (1:2050 dilution,
anti-SIII p15, BD Transduction Laboratories, San Jose, CA, USA), BTG2 (1:500
dilution, Sigma-Aldrich, Saint Louis, MO, USA), PIK3IP1 (1:500 dilution, H-180,
Santa Cruz Biotechnology Inc.), actin (1:400 dilution, pan AB-5 clone ACTN05,
Lab Vision Corp, Fremont, CA, USA), or BIM (1:500 dilution, Cell Signaling
Tech.). After washing and incubating with secondary antibodies (RAD21, BTG2,
PIK3IP1 and BIM: anti-rabbit IgG horseradish peroxidase-conjugated antibody,
Dako; TCEB1 and actin: anti-mouse IgG horseradish peroxidase-conjugated
antibody, Dako), protein bands were visualized by autoradiography.
1.6 Chromatin immunoprecipitation (III)

LNCaP-pcDNA3.1, LNCaP-ARmo and LNCaP-ARhi cells were analyzed for AR binding sites (ARBS). Cells were treated with various DHT concentrations, fixed (Merck KGaA, Darmstadt, Germany), pelleted and lysed. Chromatin was immunoprecipitated with 10 µg rabbit IgG (Santa Cruz Inc., Santa Cruz, California, USA) or 10 µl AR antibody (provided by Professor Olli Jänne, Helsinki University, Finland). ChIP DNA was sequenced with a Genome Analyzer II (Illumina Inc., San Diego, California, USA) according to the manufacturer’s protocol. Raw reads were aligned with Bowtie (Langmead et al. 2009) and the reads were mapped onto the human genome version 19 (hg19). Peak detection was performed with the tool MACS (Zhang et al. 2008a). ARBS were searched for upstream of a miRNA start site or, in the case of an intragenic miRNA, upstream of a host gene start site. If ARBS were identified and found to be closer to a miRNA than any other gene, they were determined to belong to the miRNA.

1.7 Transfections and transductions

1.7.1 Lentiviral-mediated gene silencing and overexpression (I)

Lentiviruses were used to generate cell lines stably expressing either shRNAs targeting RAD21, KIAA0196, EIF3S3, or TCEB1 or cDNAs of the same genes. shRNA oligos were purchased from Sigma-Proligo (The Woodlands, TX) and cloned into the lentiviral plasmid Lentilox3.7 (pLL3.7, kindly provided by Dr. Kalle Saksela, University of Helsinki, Finland). cDNA clones were obtained from Geneservice (Cambridge, UK) and the coding regions of RAD21, KIAA0196, EIF3S3 and TCEB1 were cloned into the lentiviral plasmid WPI (kindly provided by Dr. Jarmo Wahlfors, University of Tampere, Finland).

Lentiviral particles were produced by the calcium phosphate precipitation technique. The pLL3.7 vectors containing the shRNA sequences and the pWPI vectors containing the coding regions were cotransfected with the packaging plasmids pCMVD8.9 and pVSV-G (System Biosciences, Mountain View, CA) into HEK 293T/17 cells. Titers were measured by transducing PC-3 cells with serial dilutions of virus and analyzing the percentage of GFP-positive cells on an Agilent 2100 Bioanalyzer (Agilent Techn.). A multiplicity of infection (MOI) of 10 was used to obtain >90% transduction efficiency (Figure 7). To enhance viral transduction, 8µg/ml of polybrene (Sigma-Aldrich, Milwaukee, WI) was added to each transduction.
Figure 7: An example of the percentage of GFP positive PC-3 cells after transduction with MOI 10. GFP, green fluorescein protein; MOI, multiplicity of infection.

1.7.2 Transfections of pre- and anti-miRs (III)

LNCaP cells were transiently transfected either with Pre-miR™ (Applied Biosystems, Foster City, CA, USA) precursor miRNAs (miR-32, miR-590-5p, miR-148a and miR-99b) or an Anti-miR™ (Applied Biosystems) miRNA inhibitor (miR-99a) to study the functional significance of these miRNAs in prostate cancer cells. LNCaP cells were transfected with 10 nM Pre-miR™ or Anti-miR™ using INTERFERin (Polyplus-transfection, Illkirch, France) according to the manufacturer’s protocol. The scramble Pre- or Anti-miR™ negative controls (Applied Biosystems) were used as reference treatments.

1.8 Functional assays

1.8.1 DHT treatment prior to functional assays (III)

Before hormone exposure, LNCaP cells were cultured in 5% or 10% charcoal-stripped serum (CSS; HyClone, Inc., South Logan, UT, USA) for 3 days. The medium was subsequently replaced with medium containing 0, 1, 10 or 100 nM 5α-dihydrotestosterone (DHT) (Steraloids Inc., Newport, RI, USA), and RNA was extracted 6 hours later for expression analysis. When appropriate, cells were transfected with pre-miR constructs after 3 days of starvation in charcoal-stripped serum, and 15 nM DHT was added to the medium at the time of transfection.
1.8.2 Proliferation assays (I, III)

PC-3 and NIH 3T3 cells were transduced with lentiviruses and LNCaP cells were transfected with pre- or anti-miRNA constructs. Cells were plated in 24-well plates, and Alamar Blue (Trek Diagnostic Systems, Cleveland, OH, USA) or Resazurin (R&D systems, Minneapolis, MN, USA) reagents were used to assess cell growth. Luminescence was detected with a fluorometer (Wallac 1420 Victor; Perkin-Elmer, Fremont, CA, USA). The number of cells on each day was calculated relative to day one.

1.8.3 Cell cycle analysis (I, III)

Propidium iodide staining followed by flow cytometric analysis was performed to determine the number of cells in each phase of the cell cycle. PC-3 and LNCaP cells were stained according to published guidelines (Prather et al., 1999). Briefly, PC-3 cells transduced with shRNA were stained with a propidium iodide-containing hypotonic staining buffer and analyzed by flow cytometry (Coulter® EPICS XL-MCL, Beckman Coulter Inc). LNCaP cells transfected with pre-miRNAs were resuspended in GM buffer and fixed in ethanol for 24 hours. After fixation, cells were stained with propidium iodide (Sigma) and analyzed by flow cytometry (Accuri® C6). The proportion of cells in each phase of the cell cycle was determined using EXPO32 ADC or ModFitLT 3.2 software (Becton Dickinson, Mountain View, CA).

1.8.4 Apoptosis assay (III)

The number of apoptotic LNCaP cells was determined by cleaved caspase-3 staining. Three days after Pre-miR™ transfections, LNCaP cells were harvested and placed onto objective slides. Cells were fixed in ice-cold methanol and acetone and incubated with an anti-cleaved caspase-3 antibody (dilution 1:1000, Cell Signaling Technology, Beverly, MA, USA). The staining was visualized with the PowerVision+™ Poly-HRP IHC Detection Kit (ImmunoVision Technologies Corporation, Brisbane, CA). Counterstaining was performed with methyl green and cleaved caspase-3-positive cells were counted under a light microscope.

1.8.5 Soft agar assay (I)

Colony formation assays were utilized to evaluate the anchorage-independent growth of TCEB1-, EIF3S3-, RAD21- or KIAA0196-knockdown PC-3 cells and
TCEB1-overexpressing NIH 3T3 cells. In a 6-well plate, 5000 cells/well were seeded in 0.35% agarose on top of a 0.5% agarose base. The plates were incubated at 37°C and fed twice per week with culture medium. Colonies were photographed and counted after two weeks.

1.8.6 Invasion assay (I)

The invasive potential of TCEB1-, EIF3S3-, RAD21- or KIAA0196-silenced cells was determined using Matrigel™ Invasion Chambers (BD Biosciences, Bedford, MA). PC-3 or DU145 cells were plated in the chamber under serum starvation conditions (1% FBS) and incubated at 37°C for 22 hours. Culture medium containing 10% FBS was used as the chemotactrant. After incubation, cells that had invaded to the other side of the chamber were fixed with methanol, stained with hematoxylin and photographed. The relative area of invaded cells was measured using the image analysis software program ImageJ (Abramoff et al. 2004).

1.8.7 Migration assay (I)

The migration of EIF3S3-, RAD21- or KIAA0196-silenced PC-3 cells was evaluated with a wound-healing assay. Cells were plated in 6-well plates and grown to confluency. Wounds were made with a pipet tip, and movement was tracked by photographing the cells under a microscope after 0 and 15 hours. The relative size of the wound was determined by analysis with ImageJ (Abramoff et al. 2004).

1.8.8 Luciferase assay (III)

For luciferase assays, LNCaP cells were cotransfected with pSGG-3UTR (3’UTR of BTG2 or PIK3IP1) and pre-miRNA (pre-miR-control, pre-miR-32 or pre-miR-148a) using DreamFect (Oz Biosciences, Marseille, France) according to the manufacturer’s recommendations. Plasmids were purchased from Switchgear Genomics (Menlo Park, CA, USA). Cells were lysed 24 hours after transfection in reporter lysis buffer (Promega, Madison, WI, USA). Luciferase activity was analyzed using the Luciferase Assay System (Promega). For normalization, the protein concentration of each cell lysate was measured using the BioRad LC protein assay (BioRad).
1.8.9 Rescue experiment (III)

A rescue experiment was performed to determine if the growth advantage provided by Pre-miR-32™ (Applied Biosystems) was abolished by BTG2 addition. Excluding the 3’UTR, the coding sequence of human BTG2 from cDNA clone SC115914 (OriGene, Rockville, MD, USA) was cloned into the pSG5 expression vector. Cells were counted, seeded at 50,000 cells/well in 24-well plates and reverse transfected with 10 nM pre-miRNA and 150 ng of plasmid/well using Lipofectamine™ 2000 (Invitrogen) according to the manufacturer’s protocol. Growth curves were established as described above.

1.9 Statistical analysis (I, II, III)

The following tests were utilized to determine statistical significance for the data described in the Results section: in original communication I, two-tailed Mann-Whitney U-test used was to determine statistical difference of expression of TCEB1 in clinical samples and in growth curves in experimental versus control groups and unpaired t-test with Welch’s correction was used to determine statistical difference in various cell base assays (Matrigel invasion assay, soft agar assay, cell cycle analysis and in wound-healing assay. In original communication II, Welch t-test followed by Benjamini and Hochberg correction with p-value cut-off 0.05 and fold change cut of > 1.5 were used for the identification of differentially expressed miRNAs in microarray. Two-tailed nested t-test with Pearson’s correlation was used to evaluate possible correlation between clinical groups and Kaplan-Meier survival curves and log-rank tests were used to assess the predictive values of prognostic approaches. The random variance t-test was used in miR-predictor construction. In original communication III, Welch t-test followed by Benjamini and Hochberg correction with p-value cut-off 0.05 and fold change cut of > 1.5 were used for the identification of differentially expressed miRNAs in microarray. Mann-Whitney U-test was assessed to determine the statistical significance in different growth curves between experimental and control groups. Welch t-test was used to assess statistical significance of cell cycle analysis and apoptosis assay. Kaplan-Meier survival curves with Mantel-Cox tests were employed to evaluate the prognostic value of the immunostaining.
RESULTS AND DISCUSSION

1.1 Functional evaluation of putative 8q target genes

In this study, four putative target genes of 8q amplification, TCEB1, EIF3H, RAD21 and KIAA0196, were functionally examined to evaluate their influence on the tumorigenic properties of cancer cell lines. Silencing or overexpressing RAD21 or KIAA0196 in cell lines did not produce a phenotype, indicating that they are unlikely to be targets of 8q amplification. The results for TCEB1 and EIF3H are discussed below.

1.1.1 TCEB1

*TCEB1* (also known as Elongin C) was initially discovered by cDNA microarray to be highly expressed in the PC-3 cell line, which harbors an 8q gain. By fluorescent in-situ hybridization, TCEB1 was determined to be amplified in 23% of castration-resistant tumors, but not in any of the primary tumors (Porkka *et al.* 2002). Expression analysis of *TCEB1* in clinical tumors demonstrated a significantly higher expression level in CRPC compared with benign tissue (original communication I). In addition to prostate cancer, *TCEB1* amplification has been reported in breast cancer, which also exhibits 8q gains (Choschzick *et al.* 2010).

TCEB1 has two known biochemical functions. Together with Elongin A and Elongin B, TCEB1 regulates transcription via a direct interaction with RNA polymerase II (Aso *et al.* 1995). In addition, TCEB1 binds along with Elongin B (forming an Elongin B/C complex) to a number of proteins that contain a BC-box, including von Hippel-Lindau (VHL) protein and SOCS-box containing proteins. Cullin-dependent ubiquitin ligases are multisubunit enzymes, and Elongin B/C functions as an adapter to link SOCS-box proteins to the Cullin-Rbx module. These multiprotein complexes ubiquitinate their substrates and target them for degradation (Kibel *et al.* 1995; Stebbins *et al.* 1999; Kamura *et al.* 1998 & 2004; Zhang *et al.* 1999).

Lentiviral-mediated TCEB1 silencing was utilized to investigate TCEB1 function in PC-3 cells. Several cell-based assays were performed. According to growth curve and flow cytometric cell cycle analyses, TCEB1 silencing did not alter the proliferation or apoptosis of PC-3 cells compared with shLUC-transduced controls.
Anchorage-independent growth was assessed by colony formation in soft agar. TCEB1-silenced PC-3 cells formed significantly fewer colonies in soft agar than control shLUC-transduced cells (Table 4). TCEB1 silencing dramatically reduced the invasion of PC-3 cells as demonstrated by Matrigel invasion assays. A similar result was obtained when DU145 cells were evaluated (Table 4). In contrast, no effect on the migration of PC-3 cells was apparent in wound-healing assays. DU145 cells do not harbor an 8q gain and have lower TCEB1 expression than PC-3 cells (Porkka et al. 2002). These results indicate that TCEB1 promotes invasion with and without amplification. Taken together these results, TCEB1 silencing reduced cellular invasion of PC-3 and DU145 cells and anchorage-independent growth of PC-3 cells.

When TCEB1 was overexpressed in NIH 3T3 cells, cell growth was significantly increased compared with empty vector control (Table 4). However, NIH 3T3 cells overexpressing TCEB1 did not form colonies in soft agar. These data indicate that TCEB1 is not a transforming oncogene.

In this study, TCEB1 functions were only studied in vitro and in a limited number of cell lines. PC-3 is the only prostate cancer cell line that has an 8q amplification and therefore is the best model for studying 8q target genes. NIH 3T3 cells were used as a model for TCEB1 overexpression. The advantage of the use of mouse cell line is that murine cells can be transformed by single step. However, TCEB1 overexpression did not induce transformation of NIH 3T3 cells. Better model to study TCEB1 overexpression would have been a prostate cancer cell line which do no harbor an 8q gain. It seems that as 8q amplification occurs late in prostate cancer tumorigenesis, target genes for the amplification do not initiate cancer, but are rather involved in late stage changes, such as metastatic processes. Thus, normal cells may not be the best model to study 8q target genes. So far, the in vivo functions of TCEB1 in prostate cancer have not been determined, and TCEB1 has not been studied in other cancer types.

Table 4: Results of TCEB1 and EIF3H overexpression and silencing in cell lines. Significance is assessed by Welch t-test.

<table>
<thead>
<tr>
<th></th>
<th>NIH 3T3</th>
<th>pWPI</th>
<th>pWPI-TCEB1</th>
<th>pWPI-EIF3H</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative growth (relative to day 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+4d</td>
<td>4.33</td>
<td>6.15</td>
<td>5.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p=0.0055</td>
<td>p=0.0364</td>
<td></td>
<td></td>
</tr>
<tr>
<td>shLUC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shTCEB1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>shEIF3S3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of colonies in soft agar</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC-3</td>
<td>157</td>
<td>110</td>
<td>151</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p=0.0052</td>
<td></td>
<td>n.s.</td>
<td></td>
</tr>
<tr>
<td>% of cells invaded through Matrigel (relative to control)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PC-3</td>
<td>100</td>
<td>51.5</td>
<td>p=0.0317</td>
<td></td>
</tr>
<tr>
<td>DU145</td>
<td>100</td>
<td>60.8</td>
<td>p=0.0036</td>
<td></td>
</tr>
</tbody>
</table>
The gene expression profile of TCEB1-silenced PC-3 cells was determined by Agilent Whole Human Genome microarray. Genes known to be involved in invasion were among the 10 most down-regulated genes. For example, the extracellular matrix components ankyrin G (ANK3) and CHRDL2 (chordin-like 2) were down-regulated by TCEB1 silencing. In addition, proteinase domain-containing ADAMTS12 (ADAM metallopeptidase with thrombospondin type 1 motif, 12) was down-regulated. However, the role of ADAMTS12 in invasion is controversial because a recent mouse study suggested that ADAMTS12 has antiangiogenic properties and that when overexpressed in cells, it protects from *in vivo* tumor growth (El Hour *et al.* 2010). Angiopoietin like-4 (ANGPTL4) was up-regulated in TCEB1-silenced cells. ANGPTL4 appears to play a pivotal role in cancer because previous studies reported that ANGPTL4 inhibits the invasion and motility of epithelial cancer cells (Galaup *et al.* 2006). More recent studies suggest that the suppression of ANGPTL4 impairs tumor growth through enhanced apoptosis (Zhu *et al.* 2011). Because TCEB1 is in a complex with ubiquitin ligases that target substrate proteins for degradation, it is possible that transcriptional profiling will not reveal the real targets responsible for the phenotype in TCEB1-silenced cells. Thus, more comprehensive analyses of TCEB1 should be performed at the protein level. Novel proteins interacting with TCEB1 could be identified by pull-down assays and by this approach novel regulatory functions of TCEB1 could be revealed. In addition, it should be studied whether TCEB1 overexpression have an effect on transcription and especially to the transcription of known oncogenes.

To summarize, TCEB1 promotes invasion by unknown mechanisms and therefore is a potential target gene of 8q21 amplification.

*TPD52* has been suggested to be the target gene for the minimal amplification region of 8q21. It is coamplified with *TCEB1* and overexpressed in prostate cancer (Rubin *et al.* 2004; Wang *et al.* 2004). TPD52 enhances proliferation, decreases apoptosis and increases migration (Wang *et al.* 2007; Ummanni *et al.* 2008). In addition, TPD52 is androgen-regulated and enhances colony formation and proliferation under androgen-depleted conditions. One hypothesis is that TPD52 is involved in the development of castration-resistant disease by inhibiting the apoptosis caused by androgen deprivation therapy (Zhang *et al.* 2007a & 2011). In contrast to TCEB1, which is overexpressed only in advanced tumors, TPD52 is highly overexpressed in localized primary tumors and in PIN lesions. Because 8q gains occur late in prostate cancer progression, other mechanisms must exist that regulate TPD52 expression (Rubin *et al.* 2004; Wang *et al.* 2004).
1.1.2 EIF3H

High-level amplification of EIF3H (also known as EIF3S3 and EIF3-p40) was found in one-third of CRPC, where it was typically coamplified with MYC. Amplification leads to EIF3H overexpression and is associated with poor prognosis after radical prostatectomy (Nupponen et al. 1999, Saramäki et al. 2001, Savinainen et al. 2004). We previously demonstrated that overexpression of EIF3H in NIH 3T3 cells enhanced cell proliferation and that inhibition of EIF3H by siRNA in prostate cancer cells inhibited cell growth. These data provide strong evidence that EIF3H is a candidate target gene of 8q24 gains (Savinainen et al. 2006).

EIF3H is a subunit of eukaryotic translation initiation factor 3 (EIF3). Several components of EIF3 have been implicated in cancer; in one study, 5 of the 13 subunits of EIF3 caused malignant transformation of immortal NIH-3T3 cells (Zhang et al. 2007b). Chromosomal aberrations in EIF3H have been reported in breast, prostate and hepatocellular carcinoma. In addition, genome wide association studies identified EIF3H as a putative causative factor in hereditary colorectal cancer (Tomlinson et al. 2008).

When EIF3H was overexpressed in NIH 3T3 cells, the cellular growth rate was significantly enhanced (Table 4), which we have previously demonstrated (Savinainen et al. 2006). However, when EIF3H was silenced in PC-3 cells, no effects were seen on growth, invasion or anchorage-independent growth. This contradicts our previous result that EIF3H silencing by siRNA reduced the growth of several prostate cancer cell lines. Also Zhang and colleagues have shown that knockdown of EIF3H expression in PC-3 cells affects proliferation rate and colony formation on soft agar. In addition, enhanced proliferation and colony formation was seen in normal primary prostate (PrEC) cells overexpressing EIF3H (Zhang et al. 2008b). It is possible that lentiviral-mediated silencing with shRNAs was not sufficiently effective to observe a growth reduction in PC-3 cells. Silencing efficiency of EIF3H siRNA was up to 96% in PC-3 cell line, however, with shRNA using the same sequence the efficiency was approximately 80% (Savinainen et al. 2006). Off-target effects are difficult to exclude totally, siRNA sequence that was used was designed not to target any other gene. However, most efficient siRNA sequence was selected and all the experiments were done with the same siRNA/shRNA. Use of several siRNAs/shRNAs that target the same gene would diminish the fear of off-target effects.

In addition to prostate cancer cells, we studied EIF3H in two colorectal cell lines, LoVo and HT-29, of which HT-29 has high expression and gene amplification of EIF3H. In LoVo cells, EIF3H silencing reduced and overexpression enhanced cell growth. In HT-29 cells, silencing EIF3H reduced anchorage-independent growth (Pittman et al. 2010). Taken together these results, EIF3H overexpression in NIH 3T3 cells enhanced the growth rate. However, silencing EIF3H in PC-3 did not cause any phenotype.
As mentioned earlier, EIF3H is a subunit of translation initiation factor complex. Cancer cells have increased metabolic activity and enhanced proliferation, which requires enhanced protein synthesis (Rosenwald, 2004). Ratelimiting step in protein translation is the initiation step. Alteration in the expression and activity of specific translation factors are common feature of human cancer (Silvera et al. 2010). It has been shown that overexpression of EIF3H leads to enhanced rate of protein synthesis and also selectively enhances translation of several oncogenic mRNAs. Mechanism how EIF3H affects to the rate of protein synthesis is not known. EIF3H overexpression does not cause an increase in EIF3 levels, however its activity may be enhanced (Zhang et al. 2007b; Zhang et al. 2008b). These mechanisms could be studied in more detail to understand the oncogenic properties of EIF3H. According to previous studies and those in original communication I, EIF3H is considered a putative target gene for 8q gains.

No functional studies have been performed on the synergistic effects of 8q target genes. Often the entire arm of 8q is gained, triggering the up-regulation of several putative target genes. It is also possible that these genes have synergistic effects that are not observed when analyzing the genes individually.

### 1.2 miRNA expression studies on clinical prostate cancer (II, III)

Agilent miRNA microarrays were utilized to study miRNA expression in the clinical prostate cancers in original communications II and III. In original communication II, 102 clinical samples were obtained from Erasmus University Medical Center, Rotterdam, The Netherlands. The samples consisted of 11 normal adjacent tumors (NAP), 50 primary tumors (PCa), 4 non-malignant samples obtained by TURP (TURP-normal), 22 malignant TURP (TURP-PCa), 3 normal lymph nodes (LN-normal) and 12 metastatic lymph nodes (PC-lymph). The same study platform was used in original communication III to determine the miRNA expression profiles of 54 clinical samples from Tampere University Hospital, encompassing 5 benign hyperplasia (BPH), 28 primary tumors (PC), 7 benign prostatic hyperplasia obtained by TURP (BPH-TURP) and 22 castration-resistant tumors (CRPC). In both studies, microarray results were normalized by quantile normalization, and pairwise Welch t-tests were employed to obtain differentially expressed miRNAs in each group. In the Rotterdam sample set, a total of 513 (71%) miRNAs from the 723 human miRNAs included in the array were considered significantly expressed, whereas 411 (57%) miRNAs in the Tampere sample set were significantly expressed over background.

Altogether, 80 miRNAs were differentially expressed in the Rotterdam sample set that contained three types of malignant samples and the corresponding benign controls (primary tumors, TURPs and lymph nodes). In the Tampere clinical
samples, there were 65 differentially expressed miRNAs in the primary tumors and 88 differentially expressed miRNAs in the CRPCs compared with the benign controls (Figure 8).

**Figure 8:** Venn diagram of differentially expressed miRNAs in Tampere samples and in Rotterdam samples. PC, primary prostate tumor; BPH, benign prostatic hyperplasia; CRPC, castration resistant prostate cancer

When the expression profiles of the primary tumors from the Tampere and Rotterdam clinical samples were compared, 29 miRNAs were differentially expressed. The results are summarized in Table 5. In total, 19/29 miRNAs were previously reported to be deregulated in prostate cancer (Volinia et al. 2006; Ambs et al. 2008; Ozen et al. 2008; Tong et al. 2009; Schaefer et al. 2010; Carlsson et al. 2011; Wach et al. 2012). Only miR-7 had discrepant results; Ambs et al. reported down-regulation of miR-7 whereas we observed up-regulation.

Discrepancies are very common in published miRNA expression profiles; several miRNAs have been reported to be up-regulated in certain studies and down-regulated in others. One explanation stems from the use of different platforms, particularly in the earliest profiling studies when the arrays were custom-made. The array platform used by Volinia et al. also detected precursor molecules, and Carlsson et al. used a qPCR-based method. The data reported by Ozen et al. appears to be distorted because all of the miRNAs (76) with differential expression were down-regulated. The number of miRNA probes in different platforms varies. In addition, controls, normalization, sample selection and preparation can significantly affect the results.
Table 5: Combined results deregulated miRNAs in primary prostate cancer according to the two different data sets analyzed

<table>
<thead>
<tr>
<th>miRNA ID</th>
<th>Regulation</th>
<th>Implicated in previous studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-130b</td>
<td>up</td>
<td>Schaefer, up</td>
</tr>
<tr>
<td>miR-182</td>
<td>up</td>
<td>Ambs, Schaefer, Wach, up</td>
</tr>
<tr>
<td>miR-182*</td>
<td>up</td>
<td>Schaefer, up</td>
</tr>
<tr>
<td>miR-183</td>
<td>up</td>
<td>Schaefer, up</td>
</tr>
<tr>
<td>miR-183*</td>
<td>up</td>
<td></td>
</tr>
<tr>
<td>miR-200b*</td>
<td>up</td>
<td></td>
</tr>
<tr>
<td>miR-25</td>
<td>up</td>
<td>Tong, Ambs, Volinia, up</td>
</tr>
<tr>
<td>miR-375</td>
<td>up</td>
<td>Ambs, Schaefer, Wach, up</td>
</tr>
<tr>
<td>miR-425</td>
<td>up</td>
<td>Ambs, up</td>
</tr>
<tr>
<td>miR-512-3p</td>
<td>up</td>
<td></td>
</tr>
<tr>
<td>miR-615-3p</td>
<td>up</td>
<td></td>
</tr>
<tr>
<td><strong>miR-7</strong></td>
<td><strong>up</strong></td>
<td><strong>Ambs, down</strong></td>
</tr>
<tr>
<td>miR-93</td>
<td>up</td>
<td>Ambs, Volinia, Carlsson, Wach, up</td>
</tr>
<tr>
<td>miR-96</td>
<td>up</td>
<td>Schaefer, up</td>
</tr>
<tr>
<td>miR-1</td>
<td>down</td>
<td>Ambs, down</td>
</tr>
<tr>
<td>miR-133a</td>
<td>down</td>
<td>Ambs, down</td>
</tr>
<tr>
<td>miR-133b</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>miR-145</td>
<td>down</td>
<td>Tong, Ambs, Wach, down</td>
</tr>
<tr>
<td>miR-145*</td>
<td>down</td>
<td>Carlsson, down</td>
</tr>
<tr>
<td>miR-149</td>
<td>down</td>
<td>Volinia, Schaefer, down</td>
</tr>
<tr>
<td>miR-150</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>miR-204</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>miR-205</td>
<td>down</td>
<td>Tong, Ambs, Schaefer, down</td>
</tr>
<tr>
<td>miR-221</td>
<td>down</td>
<td>Tong, Ambs, Schaefer, Wach, down</td>
</tr>
<tr>
<td>miR-221*</td>
<td>down</td>
<td>Carlsson, down</td>
</tr>
<tr>
<td>miR-222</td>
<td>down</td>
<td>Tong, Schaefer, Wach, down</td>
</tr>
<tr>
<td>miR-455-3p</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>miR-455-5p</td>
<td>down</td>
<td></td>
</tr>
<tr>
<td>miR-551b</td>
<td>down</td>
<td></td>
</tr>
</tbody>
</table>

Here, when Agilent microarrays were used for two different sample sets, there was much less discrepancy. Still, many of the significantly deregulated miRNAs differed, which is most likely a function of using different data sets and controls. In the Rotterdam clinical sample set, normal adjacent tissue was used as a control, whereas BPH was the control for the Tampere clinical samples. Metastatic lymph nodes were included in the Rotterdam set. TURP-PCa covers a heterogeneous group of primarily advanced diseases; however, for half of the TURP-PCa samples, it was unknown whether they represented castration-resistant disease. It was impossible to draw any conclusions from the miRNA expression profile of the CRPC samples in the Rotterdam clinical sample set.
Many of the miRNAs that were differentially expressed in these two clinical datasets have been mechanistically implicated in prostate cancer. For example, the expression of miR-205 in prostate cancer cells impairs migration and invasion, and miR-145 expression leads to cell cycle arrest and apoptosis (Gandellini et al. 2009; Zaman et al. 2010). miR-93 and miR-25 target the tumor suppressor PTEN, and miR-221 has been reported to be androgen-regulated (Poliseno et al. 2010; Ambs et al. 2008). These data indicate that expression profiling is a useful tool to identify miRNAs that are potentially involved in cancer progression. In the case of several of the differentially expressed miRNAs, further studies are necessary to address whether they are functionally relevant in prostate cancer. In addition, the mechanisms for the deregulation of individual miRNAs are not well understood, but it appears that gene amplifications and deletions of miRNAs are rare (Porkka et al. 2010).

1.2.1 Prognostic miRNA signature (II)

The 80 miRNAs that were differentially expressed in the Rotterdam clinical samples were used for hierarchical clustering of all 102 samples. These miRNAs divided primary tumors (n= 50) into two different subgroups: group I contained 34 samples that clustered closely to the non-malignant tissues (NAP, TURP-normal and LN-normal), whereas group II, consisting of 16 samples, clustered with samples from more advanced disease (LN-PC, TURP-PCa). Group II tumors were significantly associated with an increased rate of metastasis after radical prostatectomy and with cancer-related deaths. Twenty-two miRNAs contributed to the separation of these groups, of which several were down-regulated in metastatic lymph nodes (original communication II).

Twenty miRNAs from the panel that separated the two groups together with 5 other miRNAs best separated group I from group II by the Bayesian Covariate Compound Predictor Algorithm (Table 6). This miR-predictor correctly classified 40 out of 50 samples. Ten samples were not assigned to either group. In this dataset, the miR-predictor was superior in predicting prognosis to pathological grade or stage and to preoperative PSA (original communication II).
Table 6: 25 miRNAs that were best in separating prognostic groups I and II in Rotterdam clinical data set. Regulation shown in worse prognosis group (group II)

<table>
<thead>
<tr>
<th>miRNA ID</th>
<th>Up-regulated</th>
<th>Down-regulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-25</td>
<td>miR-133a</td>
<td></td>
</tr>
<tr>
<td>miR-32</td>
<td>miR-133b</td>
<td></td>
</tr>
<tr>
<td>miR-106b</td>
<td>miR-143*</td>
<td></td>
</tr>
<tr>
<td>miR-93</td>
<td>miR-143</td>
<td></td>
</tr>
<tr>
<td>miR-96</td>
<td>miR-145</td>
<td></td>
</tr>
<tr>
<td>miR-183*</td>
<td>miR-145*</td>
<td></td>
</tr>
<tr>
<td>miR-183</td>
<td>miR-221</td>
<td></td>
</tr>
<tr>
<td>miR-425</td>
<td>miR-222</td>
<td></td>
</tr>
<tr>
<td>let-7b</td>
<td>miR-378</td>
<td></td>
</tr>
<tr>
<td>miR-663</td>
<td>miR-204</td>
<td></td>
</tr>
<tr>
<td>miR-301b</td>
<td>miR-27b</td>
<td></td>
</tr>
<tr>
<td>miR-141*</td>
<td>miR-139-5p</td>
<td></td>
</tr>
<tr>
<td>miR-1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The weakness of this predictor is that it was not tested on additional datasets. Only a few studies have reported on miRNA expression as prognostic indicators in prostate cancer. Certain studies arrived at the conclusion that there are no differences in miRNA expression between favorable and unfavorable prognoses (Leite et al. 2011). Tong et al. demonstrated that the expression of 16 miRNAs classified relapse patients from non-relapse patients. Of those 16 miRNAs, only miR-96 is on our list. In the study by Schaefer et al., miR-96 was found to be a prognostic indicator. Also miR-221 expression has been associated with tumor progression and clinical recurrence (Spahn et al. 2010). Again, there are several reasons for the discrepancies, many of them have been discussed previously (in example: different platforms, sample selection and handling). In addition to these, definition of progression is varies in different studies. In our study, group II tumors were associated with metastasis and prostate cancer related death, not with PSA progression. In many studies, PSA progression (biochemical relapse) is used to devide progressors from non-progressors (Tong et al. 2008; Schaefer et al. 2010). It is difficult to compare such results, since increase of PSA following surgery is associated with prostate cancer specific death, but still the majority of men with PSA progression die of other causes. Carver et al. reported that 36% of men with PSA progression after surgery eventually developed clinically relevant disease (Carver et al. 2006). Porter et al. observed that almost 50% of patient had PSA progression after radical prostatectomy, but prostate cancer specific death occurred in only 19% of the population (Porter et al. 2006). Against these observations, there might be too much background noise when PSA progression is used as an end-point and the real prognostic indicators are not found.
Our results indicate that miRNA expression can be used to determine the prognosis of prostate cancer patients. Still, discrepancy exists among individual studies in terms of which miRNAs have prognostic value. More detailed studies on individual miRNAs are necessary to determine if a single miRNA or a small panel miRNAs can function as a clinically relevant prognostic marker. In addition, more reliable tools to study miRNA expression are required if miRNA expression is to be used in large-scale clinical tests.

1.3 Putative AR regulated miRNAs (III)

To select interesting miRNAs for further functional analysis, miRNA expression profiling on CRPC was combined with the knowledge of AR binding sites (ARBS). AR signaling is reactivated in CRPC, but the critical target genes for the emergence of CRPC are unknown (Chen et al. 2004; Linja et al. 2001). ARBS were profiled by chromatin immunoprecipitation and sequencing (Chip-seq) in LNCaP-derived cell line models that express different levels of AR. Altogether, 414 miRNAs were identified with proximal ARBS. In a miRNA microarray, 88 miRNAs were differentially expressed in CRPC compared with benign controls. miRNAs that are deregulated in CRPC and contain ARBS are listed in Table 7. Of these, miR-125b, miR-21 and miR-221 have previously been demonstrated to be regulated by AR (Ribas et al. 2009; Shi et al. 2007; Ambs et al. 2008). From an initial screen for a growth effect of miRNA-transfected LNCaP cells, miR-32 and miR-148a were selected for further studies and will be discussed below.

miRNAs are widely distributed in the human genome, with some miRNAs located in intergenic regions, having their own transcriptional regulatory elements. Approximately 50% of miRNAs are intragenic, located in the introns of a host gene (Saini et al. 2008; Rodriguez et al. 2004). Intragenic miRNAs can be co-expressed with the host gene, or they can utilize their own promoter (Ozsolak et al. 2008; Monteys et al. 2010). ARBS were searched for upstream of the miRNA start site and, in the case of intragenic miRNAs, upstream of the host gene. If ARBS were identified without a proximal gene, the ARBS were determined to belong to the miRNA. However, it is challenging to conclusively associate gene to certain ARBS. It is possible that the identified ARBS regulate distal gene through DNA looping.
Table 7: Differentially expressed miRNAs in castration-resistant prostate cancer which have AR binding site according to the Chip-seq analysis

<table>
<thead>
<tr>
<th>miRNA ID</th>
<th>regulation</th>
<th>miRNA ID</th>
<th>regulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>let-7a</td>
<td>up</td>
<td>miR-100</td>
<td>down</td>
</tr>
<tr>
<td>miR-148a</td>
<td>up</td>
<td>miR-101*</td>
<td>down</td>
</tr>
<tr>
<td>miR-17</td>
<td>up</td>
<td>miR-125b</td>
<td>down</td>
</tr>
<tr>
<td>miR-18a</td>
<td>up</td>
<td>miR-125b-2*</td>
<td>down</td>
</tr>
<tr>
<td>miR-19a</td>
<td>up</td>
<td>miR-181a</td>
<td>down</td>
</tr>
<tr>
<td>miR-20a</td>
<td>up</td>
<td>miR-181b</td>
<td>down</td>
</tr>
<tr>
<td>miR-20a*</td>
<td>up</td>
<td>miR-221</td>
<td>down</td>
</tr>
<tr>
<td>miR-21</td>
<td>up</td>
<td>miR-222</td>
<td>down</td>
</tr>
<tr>
<td>miR-30b</td>
<td>up</td>
<td>miR-24</td>
<td>down</td>
</tr>
<tr>
<td>miR-30b*</td>
<td>up</td>
<td>miR-26a</td>
<td>down</td>
</tr>
<tr>
<td>miR-32</td>
<td>up</td>
<td>miR-28-3p</td>
<td>down</td>
</tr>
<tr>
<td>miR-340*</td>
<td>up</td>
<td>miR-498</td>
<td>down</td>
</tr>
<tr>
<td>miR-625</td>
<td>up</td>
<td>miR-584</td>
<td>down</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-873</td>
<td>down</td>
</tr>
<tr>
<td></td>
<td></td>
<td>miR-99a</td>
<td>down</td>
</tr>
</tbody>
</table>

1.3.1 miR-32 and miR148a

miR-32 and miR-148a were up-regulated in CRPC according to the microarray analysis, and both were demonstrated to possess an AR binding site (Table 7 and Figure 9). Precursor molecules for miR-32 and miR-148a and a control scramble precursor were transfected into the AR-responsive cell line, LNCaP. As demonstrated by growth curves, both miRNAs significantly enhanced cell growth compared with control-transfected cells. Cell cycle analysis by flow cytometry revealed fewer apoptotic cells in miR-32-transfected LNCaPs and more cells in S-phase in miR-148a-transfected LNCaPs. The ability of miR-32 to reduce apoptosis was verified by cleaved caspase-3 staining; miR-32 transfection significantly reduced the number of stained LNCaP cells (Table 8). The up-regulation of miR-32 and miR-148a by androgens was confirmed and the AR binding sites in close proximity to both miRNAs were validated by ChIP-PCR (original communication III).
miR-32 and miR-148a expression in clinical samples according to miRNA microarray. Significance is assessed by Mann-Whitney U-test, between samples BPH turp and CRPC. (Jalava et al. 2012)

Table 8: Functional analysis of pre-miR-32 and pre-miR-148a transfected LNCaP cells. Significance is assessed by Welch t-test.

<table>
<thead>
<tr>
<th></th>
<th>pre-miR-ctrl</th>
<th>pre-miR-32</th>
<th>pre-miR-148a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relative growth (relative to day 1)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+5d</td>
<td>2.467</td>
<td>3.363</td>
<td>3.573</td>
</tr>
<tr>
<td></td>
<td><em>p=0.0002</em></td>
<td><em>p=0.0022</em></td>
<td></td>
</tr>
<tr>
<td>Cell cycle analysis (fold change compared to pre-miR-ctrl)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G1</td>
<td>1</td>
<td>1.059</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n.s</td>
<td>n.s</td>
</tr>
<tr>
<td>G2</td>
<td>1</td>
<td>0.672</td>
<td>1.202</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n.s</td>
<td>n.s</td>
</tr>
<tr>
<td>S</td>
<td>1</td>
<td>0.856</td>
<td>4.829</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n.s</td>
<td><em>p= 0.0257</em></td>
</tr>
<tr>
<td>apoptosis</td>
<td>1</td>
<td>0.359</td>
<td>1.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>p=0.0353</em></td>
<td>n.s</td>
</tr>
<tr>
<td>Caspase 3 staining</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% of positive</td>
<td>7.17</td>
<td>4.01</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>p=0.0037</em></td>
<td></td>
</tr>
</tbody>
</table>

Agilent mRNA microarrays were utilized to identify putative targets of miR-32 and miR148a. Even though miRNAs can repress the translation of target genes, it has been shown that miRNAs predominantly act by decreasing the mRNA level of their targets (Guo et al. 2010). This makes mRNA microarrays an efficient tool for identifying putative target genes. Of the genes that were down-regulated by pre-miR-32 or pre-miR148a transfection in LNCaP cells, BTG2 (B-cell translocation gene 2) and PIK3IP1 (phosphoinositid-3-kinase interacting protein 1) were selected for further study as putative target genes of miR-32 and miR-148a, respectively. Down-regulation of BTG2 and PIK3IP1 by transfection of miR-32 and miR-148a was confirmed at both the protein and the mRNA level. Luciferase reporter assays indicated direct binding of miR-32 to the BTG2 3’UTR and of miR-148a to the PIK3IP1 3’UTR (original communication III). Taken together these results, miR-32
and miR-148a were discovered to be up-regulated in CRPC. Both of the miRNAs enhanced the growth of LNCaP cells, miR-32 by reducing apoptosis. These miRNAs were shown to be regulated by androgens, and BTG2 and PIK3IP1 were found to be the target genes for miR-32 and miR-148a, respectively.

BTG2 is a tumor suppressor, which is involved in many cellular processes, including cell cycle control and apoptosis (Hong et al., 2005; Winkler, 2010). It has been previously been reported to be down-regulated in primary prostate tumors, however in the study CRPCs were not evaluated (Ficazzola et al., 2001). PIK3IP1 interacts with PI3K and negatively regulates its activity (Zhu et al., 2007).

BTG2 and PIK3IP1 expression was analyzed by immunohistochemistry in 264 primary tumors (PC) and CRPCs. The expression of PIK3IP1 was reduced in malignant tissue compared with non-malignant tissue, but there was no significant difference in staining intensity between PC and CRPC. In contrast, BTG2 staining intensity was significantly lower in CRPC compared with PC. In prostatectomy specimens, low BTG2 staining was significantly associated with a higher pT stage (pT3 vs. pT2), but was not associated with Gleason score or PSA progression (Table 9). In addition, loss of BTG2 expression was significantly associated with a shorter progression-free period (original communication III).

Table 9: BTG2 and PIK3IP1 protein expression in clinical prostate cancer samples. Significance is assessed by chi-square test.

<table>
<thead>
<tr>
<th></th>
<th>BTG2</th>
<th>PIK3IP1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PC</td>
<td>CRPC</td>
</tr>
<tr>
<td>n=168</td>
<td>n=85</td>
<td>n=170</td>
</tr>
<tr>
<td>IHC 0 (%)</td>
<td>10</td>
<td>27</td>
</tr>
<tr>
<td>IHC 1 (%)</td>
<td>38</td>
<td>54</td>
</tr>
<tr>
<td>IHC 2 (%)</td>
<td>46</td>
<td>19</td>
</tr>
<tr>
<td>IHC 3 (%)</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>p-value</td>
<td>&lt;0.0001</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>pT2</th>
<th>pT3</th>
<th>pT2</th>
<th>pT3</th>
</tr>
</thead>
<tbody>
<tr>
<td>n=115</td>
<td>n=50</td>
<td>n=116</td>
<td>n=52</td>
<td></td>
</tr>
<tr>
<td>IHC 0 (%)</td>
<td>6</td>
<td>20</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>IHC 1 (%)</td>
<td>39</td>
<td>34</td>
<td>27</td>
<td>34</td>
</tr>
<tr>
<td>IHC 2 (%)</td>
<td>45</td>
<td>43</td>
<td>57</td>
<td>25</td>
</tr>
<tr>
<td>IHC 3 (%)</td>
<td>10</td>
<td>0</td>
<td>13</td>
<td>8</td>
</tr>
<tr>
<td>p-value</td>
<td>0.009</td>
<td>n.s.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

We have previously shown that miR-32 is up-regulated by androgens (Waltering et al., 2010). Also the androgen responsiveness of miR-148a has been suggested (Murata et al., 2010). However, several studies failed to observe androgen regulation of miR-32 and miR-148a (Ambs et al., 2008; Shi et al., 2007 and Ribas et al., 2009). There are several reasons for the contradictory results. Ambs et al., Shi et al. and Ribas et al. used the synthetic androgen R1881, whereas we treated cells with DHT.
All of the other studies treated with androgens for longer times, ranging from 12 h-72 h; we detected the best induction of miR-32 and miR-148a at 6 h (original communication III). In addition, the platforms differed in all studies (Agilent microarrays in our study, custom-made array platforms in Ambs et al. and Ribas et al. and Northern blots in Shi et al.).

There are two published targets for miR-148a in prostate cancer, CAND1 (cullin-associated and neddylation-dissociated 1) and MSK1 (antigen identified by monoclonal antibody AJ9) (Murata et al. 2010; Fujita et al. 2010). Neither of these genes was down-regulated in the mRNA array of miR-148a-overexpressing LNCaP cells, therefore they were not further studied.

In the study of Murata et al., CAND1 was found to be miR-148a target in the same cell line that we used, LNCaP. miR-148a targets were searched from Sagner miRBase Target database, and CAND1 was selected for further study. They saw CAND1 downregulation by miR-148a transfection, in both mRNA and protein level, which contradicts with our microarray data. In addition, in luciferase reporter assay where CAND1 3’UTR was used, miR-148a transfection reduced luciferase activity. These discrepancies are difficult to explain, it could be useful to study in our LNCaP cells whether CAND1 downregulation by miR-148a could be seen in protein level by Western blot analysis.

The study by Fujita et al. is one of the examples how the use of different cell lines can produce conflicting data. They selected miR-148a for further studies because in DU145 and PC-3 cells have lower expression of miR-148a than androgen responsive cell lines and in PrEC cell line. They come to conclusion that miR-148a is tumor suppressor in prostate cancer, and that it targets MSK1 in PC-3 cells. Here we show that miR-148a is up-regulated in castration-resistant prostate cancer. However, sometimes even when expression in clinical samples is studied, different research groups get opposite results. One example is miR-221 and miR-222, which have been suggested to be both, oncogenic and tumor suppressors in prostate cancer. They have been suggested to be overexpressed in androgen independent cells and possess oncogenic functions in prostate cancer (Sun et al. 2009). However, we and others have seen that these miRNAs are down-regulated in primary tumors and in CRPC (Schaefer et al. 2010; Spahn et al. 2010). In one study, up-regulation of miR-221 and 222 was demonstrated in CRPC bone metastasis. In contrast, they do see downregulation of these miRNAs in primary tumors (Sun et al. 2011). These might represent a real situation, where miRNA expression is under complex and flexible regulation, and they function differentially in different stages of prostate cancer. On the other hand, these results may be due to the differential sample preparation and different study platforms.

PIK3IP1 was identified as a target gene of miR-148a, despite the fact that none of the target prediction programs selected it as a predicted target. Luciferase assays indicated direct binding of miR-148a to the PIK3IP1 3’UTR. Target predictions are based on the idea that the six nucleotide long seed sequence of a miRNA (positions 2-7) must be fully complementary to the target sequence (Lewis et al. 2005). However, predictions suffer from both false positives and false negatives results.
There is evidence that noncanonical target sites exist, with bulges in the seed region. So far, it is not known which general rules could be used for predicting such sites (Chi et al. 2012).

In further studies, PIK3IP1 expression did not correlate with miR-148a expression because PIK3IP1 expression was not down-regulated in CRPC compared to primary tumors. It is possible that we failed to correctly identify miR-148a targets by mRNA array or that the targets differ in LNCaP cells versus clinical samples. In cell lines, mRNAs often have truncated 3’ UTRs, and therefore miRNA target genes might be different *in vitro* than *in vivo* (Sandberg et al. 2008).

Ambs and colleagues (2008) reported the up-regulation of miR-32 in prostate cancer. In the Rotterdam clinical samples, we observed miR-32 up-regulation in primary tumors (original communication II), but in the Tampere clinical sample set, miR-32 was only up-regulated in a subset of the primary tumors. However, the most dramatic up-regulation of miR-32 in the Tampere clinical sample set was detected in castration-resistant disease (Figure 9, original communication III). Ambs et al. reported that *BCL2L11* (also known as Bim) is a target gene of miR-32. We did not find down-regulation of Bim by transfecting miR-32 into LNCaP cells (original communication III). In contrast, we identified *BTG2* as a target of miR-32. It is possible, as miRNAs often have several target genes, that both genes are targeted by miR-32 in different cellular environments. BTG2 expression is reduced in CRPC, and loss of BTG2 correlates with reduced progression-free survival. In the Rotterdam clinical sample set, miR-32 was one of the 25 miRNAs that predicted the prognosis of prostatectomy-treated patients (original communication II).

Further studies are required to address whether either miR-32 or BTG2 can be used as a prognostic marker in prostate cancer. Prognostic factors in clinical use are Gleason score, pT-stage (pathological tumor stage) and the level of PSA (Partin et al. 2001; Graff et al. 2007). Careful evaluation of BTG2 staining alone and in combination with these existing markers is needed to address the question whether BTG2 staining is usable as a prognostic marker, independently or in combination with other markers. The same studies should be done with miR-32, but there are technical issues that are needed to overcome before miRNA expression can be analyzed in efficient way from clinical tumor samples. In addition, it should be evaluated if miR-32 could be detected from the blood or from the urine of the prostate cancer patients, which would make the detection easier.

Here we suggest a mechanism how the expression of tumor suppressor BTG2 is down-regulated in advanced tumors by miR-32. Expression of miR-32 on the other hand, seems to be regulated by AR binding to the miR-32 promoter. However, there might be other mechanisms for miR-32 overexpression, like epigenetic regulation that needs to be studied. If AR is responsible for miR-32 up-regulation there must be reason why this is occurring only in cancer and not in normal prostate epithelium. There might be other collaborating factors that enhance AR binding to miR-32 promoter.
The primary findings and conclusions of this thesis are as follows:

Of the four putative 8q target genes studied, EIF3H and TCEB1 had functional effects on prostate cancer cells. TCEB1 silencing reduced the invasion and colony formation of cancer cell lines. In addition, overexpression of TCEB1 enhanced the growth of murine fibroblast cells. Overexpressed EIF3H enhanced cellular growth. The increased expression of TCEB1 in castration-resistant disease was confirmed. These data indicate that TCEB1 and EIF3H are putative target genes of 8q21 and 8q24 amplifications, respectively.

miRNA expression in prostate cancer was profiled in a large clinical dataset. Altogether, 80 miRNAs were differentially expressed in malignant tissue compared with benign tissue. Of these, 25 miRNAs were able to separate primary tumors into an aggressive group and a less aggressive group. In this dataset, these miRNAs predicted prognosis better than a Gleason score, pathological stage or preoperative PSA.

miRNA expression was evaluated in castration-resistant prostate cancers. In total, 88 miRNAs were differentially expressed. The LNCaP-derived cell line was used to map ARBS in the genome, from which 414 miRNAs with ARBS were identified. Of these, 28 miRNAs were de-regulated in CRPC.

After initial screening, miR-32 and miR-148a were selected for follow-up studies. They are both up-regulated in CRPC, contain ARBS, and are up-regulated by androgens. In addition, when transfected into an androgen-responsive cell line, they each imparted a growth advantage to the cells; miR-32 accomplished this by inhibiting apoptosis.

BTG2 and PIK3IP1 were identified as target genes for miR-32 and miR-148a, respectively. By immunohistochemistry, the expression of PIK3IP1 was decreased in prostate cancer compared with non-malignant tissue, but there were no differences in PIK3IP1 expression between CRPC and primary tumors. In contrast, the expression of BTG2 was decreased in CRPC compared with primary tumors. Low BTG2 staining in prostatectomy specimens was associated with higher pT stage and a shorter progression-free period. These preliminary results indicate that either miR-32 or BTG2 could be potential therapeutic targets in advanced prostate cancer.
This study was carried out in the Molecular Biology of Prostate Cancer Group, Institute of Biomedical Technology, the University of Tampere and Tampere University Hospital, during the years 2005-2011. Former director of IBT, Professor Olli Silvennoinen, and current director of BioMediTech Dr. Hannu Hanhijärvi are acknowledged for providing the research facilities for the study.

I express my sincere gratitude to my supervisors, Professor Tapio Visakorpi and Kati Porkka, Ph.D. I am grateful for Tapio Visakorpi for taking me into his group and giving me this interesting topic to work with. It has been a privilege to learn from his expertise and work in his lab. I warmly thank Kati Porkka for taking me under her guidance especially in early years of my thesis work. She has taught me a number of laboratory techniques and always been so patient with my numerous questions.

My thesis committee members, Professor Anne Kallioniemi, Professor Johanna Schleutker and Docent Jarmo Wahlfors are thanked for their support and guidance during the years.

Professor Johanna Ivaska and Dr. Ian Mills, are acknowledged for careful reviewing of the thesis and for their valuable comments and criticism.

All my co-authors are acknowledged for their invaluable contribution to this work. I feel thankful for especially to Leena Latonen and Alfonso Urbanucci who finished some of the lab work for my last article while I was on maternity leave. Alfonso is also acknowledged for the good mood he brought to our group, you have also taught me more about olive oil that I ever wanted to know. Hanna Rauhala is thanked not only contributing to my article, but also all the help she has provided in various issues during all these years. We should continue our “SGY mammatreffit” tradition; we seem to be expanding every year.

I thank warmly all the present and former members of our group, for your help, friendship and company in and out of lab. Hanna Suikki is acknowledged for her weird sense of humor; we had fun times in the lab before you escaped into medical school. For now on, I consider you as my personal doctor. Katri Leinonen is thanked for her patience with my numerous questions concerning practical issues; I thought you must know everything since you are a member of The Postgraduate Committee. It would have probably taken another seven years to finish this thesis without our efficient and skillful technicians, Mariitta Vakkuri and Päivi Martikainen. Thank you for everything.

I thank all my dear friends for being there for me, and for all the good times we have had, NOT related to work and research.
Sydämessiset kiitokset vanhemmilleni mittaamattoman arvokkaasta tuestanne ja rakkaudestanne elämäni varrella. Kiitos anopilleni Seijalle erityisesti Jetron hoidosta, olet ollut korvaamaton apu. I thank my brothers and sister and their families. We may not see each other very often, but you are all dear to me. I appreciate of having three more sisters and their families in my life; there is always a lot of action when we all gather together!

I thank my beloved husband Tommi, for your love and support, always having faith in me and making me laugh. Your optimism is good match for my pessimism. Finally, I thank my adorable son Jetro for being the joy of my life and teaching me how to live in a moment.

This study was supported by the Tampere Graduate Program in Biomedicine and Biotechnology (TGPBB), the Medical Research Fund of Tampere University Hospital, the ProspER EU project, the Sigrid Jusélius Foundation, the Reino Lahtikari Foundation, Tampere University Foundation and the Scientific Foundation of the City of Tampere.

Tampere, April 2012

Sanni Jalava
REFERENCES


binding, transcriptional activation, and subcellular localization. Mol Endocrinol. 5:1396-404.


79


locus is homozygously deleted in a subset of prostate cancers. Genes Chromosomes Cancer. 50:499-509.
Prostate cancer cell line database, capcelllines.ca
Rosenwald IB. (2004): The role of translation in neoplastic transformation from a pathologist's point of view. Oncogene. 23:3230-47


TCEB1 promotes invasion of prostate cancer cells

Sanni E. Jalava1, Kati P. Porkka1, Hanna E. Rauhala1, Jarkko Isotalo1, Teuvo L. Tammela2 and Tapio Visakorpi1*  
1Laboratory of Cancer Genetics, Institute of Medical Technology, University of Tampere and Tampere University Hospital, Tampere, Finland  
2Department of Urology, University of Tampere and Tampere University Hospital, Tampere, Finland

Amplification of the long arm of chromosome 8 is one of the most recurrent findings in prostate cancer and it is associated with poor prognosis. Several minimal regions of amplification suggest multiple target genes which are yet to be identified. We have previously shown that TCEB1, EIF3S3, KIAA0196 and RAD21 are amplified and overexpressed in prostate cancer and they are located in the 8q area. In this study, we examined the functional effects of these genes to prostate cancer cell phenotype. We overexpressed and inhibited the genes by lentivirus mediated overexpression and RNA interference, respectively, shRNA mediated TCEB1 silencing decreased significantly cellular invasion of PC-3 and DU145 cells through Matrigel. TCEB1 silencing reduced the anchorage-independent growth of PC-3 cells. Similar effects were not seen with any other genes. When overexpressed in NIH 3T3 cells, TCEB1 and EIF3S3 increased the growth rate of the cells. Transcriptional profiling of TCEB1 silenced PC-3 cells revealed decrease of genes involved in invasion and metastasis. Finally, we also confirmed here the overexpression of TCEB1 in hormone-refractory prostate tumors. This study indicates that TCEB1 promotes invasion of prostate cancer cells, is involved in development of hormone-refractory prostate cancer and is thereby a strong candidate to be one of the target genes for the 8q gain.

Key words: prostate carcinoma; RNAi; elongin C

Prostate cancer is the most common male malignancy in Western countries and the second most common cause of cancer related deaths in males.1 Because of the intensive research in the past years main chromosomal aberrations in prostate cancer have been revealed, but hunting for many of the target genes is still in process. Gain of the long arm of chromosome 8 (8q) is one of the most recurrent findings in advanced prostate tumors and it is associated with poor prognosis.2,3 Several minimal regions of amplification of 8q have been identified by comparative genomic hybridization (CGH) and array-CGH, suggesting several target genes. Two of these minimal regions are 8q21 and 8q23-24.4-6 We have identified 4 genes that are amplified and highly expressed in prostate cancer and located in these regions. One of them, TCEB1 [transcription elongation factor B (SIII), polypeptide 1 (15 kDa, elongin C)] is located in 8q21.11 and others RAD21 [RAD21 homolog (S. pombe)], KIAA0196 and EIF3S3 (eukaryotic translation initiation factor 3, subunit H) in 8q23-24.7-9 TCEB1 was found to be amplified in 20% of hormone-refractory prostate tumors. Although in untreated primary tumors no gene amplification was found, low-level gains were found in about 30% of the tumors. In PC-3 cell line, which contains the TCEB1 gene amplification, the expression is 5 times higher compared to cell lines with no amplification.7 RAD21 is amplified in 30% of hormone-refractory carcinomas and it is expressed significantly more in untreated prostate carcinomas than in benign prostate hyperplasia (BPH). Also KIAA0196 is amplified in 30% of hormone-refractory tumors and its expression is associated with the amplification.8 High-level amplification of EIF3S3 has been found in 30–50% of hormone-refractory cancers. The expression of EIF3S3 is also higher in prostate carcinomas than in BPH.9-11 The RAD21 protein is one of the 4 subunits forming the cohesin complex which holds sister chromatids together until the onset of anaphase.12 In addition to this, RAD21 plays a role in apoptosis as it is cleaved by caspases during the apoptosis signaling cascade.13

The expression of RAD21 has been reported to be upregulated in metastatic breast cancer and it has been shown that silencing RAD21 in breast cancer cells decreases cell growth and sensitizes cells to chemotherapeutic drugs.14,15 EIF3S3 is one of the subunits of a multiprotein complex that has a central role in translation initiation.16 Amplification of EIF3S3 has been reported also in breast and hepatocellular carcinomas.17,18 We have previously shown that EIF3S3 promotes growth of prostate and breast cancer cells.18 Zhang et al.19 suggested that overexpression of any subunit of EIF3 is enough to transform immortal fibroblast cells. TCEB1 is one of the 3 subunits of the Elongin complex (SIII) that activates transcription.20 TCEB1 binds also to many proteins in cytoplasm, namely von Hippel-Lindau (VHL) protein and SOCS-box containing proteins.21,22 TCEB1 has not been studied as a putative oncogene before, KIAA0196 is a poorly known gene, and its function in cancer has not been studied before.

The aim of this study was to evaluate the roles of these 4 putative target genes (TCEB1, RAD21, KIAA0196 and EIF3S3) of the 8q amplification in prostate cancer. We overexpressed and inhibited the genes by lentivirus mediated overexpression and RNA interference, respectively. Effects of gene silencing or overexpression were evaluated using cell based assays such as proliferation assay, anchorage-independent growth and Matrigel invasion assay.

Material and methods

Cell lines and tumor samples

All cell lines (PC-3, DU145, NIH 3T3, HEK 293T/17) were purchased from ATCC (Rockville, MD) and cultured according to the recommended protocols.

Freshly frozen samples of 9 benign prostate hyperplasia (BPH), 30 untreated primary and 12 locally recurrent hormone-refractory prostate cancer specimens were obtained from Tampere University Hospital and used for real time quantitative RT-PCR. All the specimens were first histologically examined to contain more than 60% of cancerous cells. BPH samples were obtained from prostatectomy specimens from cancer patients, but the specimens were first histologically analyzed not to contain any cancer. Subsequently, cells were scratched from frozen tumor blocks using a pre-cooled, sterile scalpel for RNA extraction. The use of clinical material has been approved by Ethical Committee of the Tampere University Hospital.

Generation of the lentiviral constructs

Small interfering RNA (siRNA) sequences were designed according to the public recommendations (http://www.rockefeller.edu/). Additional Supporting Information may be found in the online version of this article.

Grant sponsors: Medical Research Fund of Tampere University Hospital, Academy of Finland, Cancer Society of Finland, Reino Lahtikari Foundation, Sigrid Juselius Foundation.

*Correspondence to: University of Tampere, Biokatu 6, FI-33520 Tampere, Finland. Fax: +358-3-355-8597. E-mail: tapio.visakorpi@uta.fi Received 15 May 2008; Accepted after revision 21 July 2008 DOI 10.1002/ijc.23916

Published online 9 October 2008 in Wiley InterScience (www.interscience.wiley.com).
ed@labheads/tuschi/sirna.html) and screened against the GenBank database by BLAST (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi). Efficiency of the siRNA target sequences of RAD21, KIAA0196, EIF3S3 and TCEB1 were first tested in PC-3 cell line using synthetic siRNA oligos (data not shown). Short hairpin RNA (shRNA) sequences were created from the tested siRNA sequence by adding XhoI restriction site and loop sequences (Table 1). shRNA oligos were purchased from Sigma-Proligo (The Woodlands, TX). shRNAs were subsequently cloned into the lentiviral plasmid WPI (kindly provided by Dr. Jarmo Wahlfors, University of Tampere, Finland), and appropriate sequence was verified by sequence analysis.

**Virus production and transductions**

VSVg pseudotyped lentiviral particles were produced in HEK 293T/17 cells by calcium phosphate precipitation technique. pLL3.7 vector containing the shRNA sequences and pWPI containing the coding regions were co-transfected with packaging plasmids pCMVΔ8.9 and pVSV-G (System Biosciences, Mountain View, CA). pL3.7 contains green fluorescence protein (GFP) as a reporter gene. Titers were measured by transducing PC-3 cells using serial dilution of viruses and analyzing the percentage of GFP positive cells by fluorescence microscopy. In most experiments, PC-3, DU145 and NIH 3T3 cells were transduced with the multiplicity of infection (MOI) 10 to obtain >90% transduction efficiency. To enhance the viral transduction, 8 μg/ml of polybrene (Sigma-Aldrich, Milwaukee, WI) was used in every transduction. The cells showed overexpression or silencing of the genes for several weeks after the transduction.

**Quantitative real-time RT-PCR (q-RT-PCR)**

Total RNA from the cell lines or tumor samples was extracted using RNAasy Mini Kit (Qiagen, Valencia, CA), and was used for the first strand cDNA synthesis with AMV Reverse Transcriptase (Finnzymes, Espoo, Finland) and Random primers (Promega, Madison, WI) (cell lines) or Superscript II Reverse Transcriptase and oligo d(T)12-18 primers (Invitrogen Life Technologies, Carlsbad, CA) (tumor samples). The expression levels of RAD21, KIAA0196, EIF3S3 and TCEB1 and a housekeeping gene TATA-box binding protein (TBP) were analyzed according to the previously described guidelines. Briefly, PCR reactions were performed using the LightCycler™ apparatus (Roche Diagnostics, Mannheim, Germany) with the LC Fast Start DNA SYBR Green I Kit (Roche Diagnostics). Melting curve analysis and agarose gel runs were performed to ensure the formation of specific PCR products.

**Western blot and antibodies**

Cells growing on a 10-cm plate were harvested using scratcher and cytoplasmic and nuclear proteins were isolated, separately. Briefly, cells were first suspended to Hypotonic buffer (10 mM Hepes, pH 7.9, 1.5 mM MgCl₂, 10 mM KCl, 0.2 mM PMSF, 0.5 mM DTT), and cytoplasmic proteins were collected from supernatant. Remaining pellet were then resuspended Low-salt buffer (20 mM Hepes, pH 7.9, 25% glycerol, 1.5 mM MgCl₂, 20 mM KCl, 0.2 mM EDTA, 0.2 mM PMSF, 0.5 mM DTT), and nuclear proteins were collected using High-salt buffer (as Low-salt buffer, but 1.2 M KCl). Protein concentration was measured using simplified Bradford method (BioRad Laboratories, Hercules, CA). 20 μg of proteins were separated by electrophoresis through 8% (RAD21) or 12% (TCEB1) SDS-PAGE gels and transferred to PVDF membranes (Immobil-P, Millipore, Billerica, MA) using semi-dry transfer technique. Membranes were incubated with primary antibody (RAD21: dilution 1:500 anti-RAD21 Novus Biologicals, Littleton, CO; TCEB1: dilution 1:250 anti-SIII p15, BD Transduction laboratories, San Jose, CA; Actin: dilution 1:400 pan AB-5 clone ACTN05, Lab Vision, Fremont, CA) for 1 hr in room temperature. After washing and incubation with secondary antibodies, protein bands were visualized on autoradiography film by using the Western Blotting Luminol Reagent (Santa Cruz, CA) according to the manufacturer’s guidelines.

**Growth curves**

Cells transduced with lentiviruses were plated on a 24-well plate at 20,000 cells/ml (PC-3) or 60,000 cells/ml (NIH 3T3) density. Alamar Blue™ (Trek Diagnostic Systems, Cleveland, OH, USA) was added to the wells to final concentration of 10%. Oxidized form of Alamar Blue enters the cytosol of the cells and is converted to reduced form by mitochondrial enzyme activity. This redox reaction changes the color of Alamar Blue from blue to pink. Fluorescence was measured after 2 hr of incubation using a 96-well fluorometer (Wallac 1420 Victor, PerkinElmer, Fremont, CA). Values were normalized against Day 1. Each experiment was done in 5 replicates.

**Cell cycle analysis by flow cytometry**

Propidium iodide-staining and flow cytometric analysis was used to determine the number of cells in S-phase fraction (SPF) and apoptosis. Cells transduced with lentiviruses were harvested and resuspended to the Hypotonic Staining Buffer (0.1% mg/ml sodium citrate dehydrate, 50 μg/ml propidium iodide, 2 μg/ml Ribonuclease A and 0.003% Triton® X-100) and the amount of propidium iodide incorporated to the genome was measured using flow cytometry ( Coulter® EPICS XL-MCL, Beckman Coulter, Fullerton, CA). Each experiment was done in triplicate. The relative proportions of SPF and apoptotic/necrotic cells were analyzed using the EXPO32 ADC software program.

**Soft agar assay**

Anchorage-independent growth was assessed by colony formation in soft agar. The 5% agarose solution (Low gelling temperature Agarose, Electran, BDH Chemicals, Poole, UK) was mixed with culture medium to make 0.5% agarose for lower layer. Cells transduced with lentiviruses were trypsinized and dilution of 5,000 cells/well was mixed with 5% agarose to form 0.35% upper layer, and allowed to solidify before incubation at 37°C. Cells were fed twice a week with culture medium. After 2 weeks, colo-
nies were photographed under UV-microscope and counted. All experiments were done in triplicate and repeated at least twice.

**Invasion assay**

Invasion assay was done using the Matrigel™ Invasion Chambers (BD Biosciences, Bedford, MA). The culture medium in the wells contained 10% FBS as chemoattractant. 1 × 10^5 of lentiviral transduced PC-3 or DU145 cells in culture medium containing 1% FBS were plated onto an insert. Plates were incubated at 37°C for 22 hr. Next, the cells inside the inserts were removed, and the invaded cells were fixed with methanol and stained with hematoxylin. Wells were photographed under microscope and the relative size of the wound was measured using image analysis software program ImageJ.24 Each experiment was done in triplicate and repeated at least twice.

**Wound-healing assay**

PC-3 cells transduced with lentiviruses were plated on a 6-well plate and allowed to grow to confluence. Three separate wounds were drawn by pipet tip into the well. Cells were rinsed with PBS and fresh medium was added to the wells. Pictures were taken by microscope in 0 and 15 hr and the relative size of the wound was analysed by ImageJ.24 Each experiment was done in triplicate.

**Microarray**

PC-3 shLUC -transduced and PC-3 shTCEB1-transduced cells were used for gene expression analysis. shLUC transduced cells were used as a control. Cells were harvested 7 days after transduction into Trizol RNA extraction reagent (Invitrogen) and total RNA was extracted. Gene expression profiling was done using Agilent Technologies’ (Santa Clara, CA) Dual-Mode Gene Expression Platform. 500 ng of RNA was labeled using Low RNA Input Linear Amplification PLUS, Two-Color Kit. Amplified and labelled samples were hybridized on Agilent’s Whole Human Genome Oligo Microarray (44k). Microarray was scanned using Agilent’s DNA Microarray Scanner and data was extracted and normalized using Agilent’s Feature Extraction software (v9.5.1). No background signals were subtracted and data was normalized using the “linear and lowess method”. All low-quality data points (low signal-to-noise ratio, saturated signals or nonuniform signals) were excluded from further analysis. Genes were considered to be differentially expressed if there was more than 2-fold difference in expression between shTCEB1 and shLUC. Microarray data have been deposited to ArrayExpress database (Accession number E-MEXP-1627)

**Gene ontology enrichment analysis**

Enrichment of differentially expressed genes in TCEB1 silenced cells into specific gene ontologies was studied using statistical computing environment R with topGO package from Bioconductor software project.25

**Statistical analysis**

Two-tailed Mann-Whitney U-test or Unpaired t-test with Welch correction was used to determine statistical difference between experimental and control groups. p-value of <0.05 was considered significant.

**Results**

**TCEB1 expression in clinical samples**

The data on the expression of EIF3S3, RAD21 and KIAA0196 in clinical prostate cancer have been published before.8,11 Here, we analyzed the TCEB1 expression in clinical samples of BPH, untreated primary, and hormone-refractory prostate carcinomas by using q-RT-PCR. As shown in Figure 1, expression of TCEB1 was significantly increased in hormone-refractory prostate carcinomas, compared to the untreated primary tumors (p = 0.0184).

**Generation of cell lines**

PC-3 cell line was used as a model to examine the role of RAD21, TCEB1, EIF3S3 and KIAA0196 in prostate cancer. PC-3 cell line contains the amplification of 8q and the genes are all expressed in PC-3.7,8,11 Lentivirus mediated shRNA system was set up for each of these genes. shRNAs were subcloned into a plasmid Lentilox3.7 which express shRNA under human U6 promoter and reporter gene EGFP under a control of cytomegalovirus (CMV) promoter. Cells were first transduced with different multiplicity of infections (MOIs) and the expression of the target genes were evaluated using the Q-RT-PCR (data not shown). Transducing cells with MOI 10 led to 65–85% specific inhibition of target gene at mRNA level (Fig. 2a). Effects of gene silencing were also seen also at the protein level with RAD21 and TCEB1 (Figs. 2c and 2d). No reliable antibodies are available against KIAA0196 or EIF3S3, which is why silencing of these genes was only demonstrated at the mRNA level. Firefly luciferase shRNA was used as a negative control in each experiment. The same shRNA system was used in DU145 cell line, where TCEB1 was 70% silenced using MOI 10 (Fig. 2b).

NIH 3T3 cells were used to set up an overexpression model of RAD21, KIAA0196, EIF3S3 and TCEB1. Coding regions of the genes were subcloned into lentiviral plasmid WPI in which the transgene and GFP are bicistronically expressed under EPI1-alpha promoter. GFP expression was used to determine the transduction efficiency by fluorescence microscope and Bioanalyzer. After transduction, transgene expression was evaluated at the mRNA level. The expression of the transgene was 30–600 times higher in transduced cell line than in empty vector control (Fig. 2e). In the case of TCEB1 and RAD21, for which specific antibodies were available, upregulation was seen also at the protein level (Figs. 2f–2g).

The effect of gene silencing and overexpression on the growth of the cells was assessed using growth curves. Silencing RAD21, KIAA0196, TCEB1 or EIF3S3 in PC-3 cell line did not cause any affects on the cell growth compared to shLUC transduced cells (Fig. 3a). Results were confirmed by determination of SPF by flow cytometry, the mean ± SD of SPF were: shLUC 15.1 ± 2.3%, shRAD21 14.9 ± 0.3%, shEIF3S3 13.8 ± 1.0%, shKIAA0196 14.4 ± 1.2% and shTCEB1 13.4 ± 0.8%. No effects were seen on the apoptotic/necrotic cell populations either in the flow cytometric analysis (data not shown).

Overexpressing RAD21 or KIAA0196 in NIH 3T3 cells did not have any effects on the cell growth, whereas overexpressing EIF3S3 or TCEB1 increased significantly cell growth compared to
cells transduced with empty vector \((p = 0.0317, p = 0.0079,\) respectively) (Fig. 3b).

**Invasion and anchorage-independent growth**

We used the highly invasive PC-3 cells to assess the influence of *RAD21, EIF3S3, KIAA0196* and *TCEB1* on invasion by using the Matrigel invasion assay. Silencing of *EIF3S3, KIAA0196* and *RAD21* had no significant effects on the ability of the cells to invade. Whereas, *TCEB1* silencing significantly reduced invasion \((p = 0.042)\) as shown in Figure 4a. Similar reduction on invasion \((p = 0.0036)\) was also seen with the other invasive prostate cancer cell line DU145 (Fig. 4b). No effect was seen on the migration of *TCEB1* silenced PC-3 cells, studied by the wound healing assay (Supp. Fig. 1).

Next we wanted to study the effects of gene silencing on the colony formation using soft agar assay. After 14 days of incubation *RAD21, EIF3S3* or *KIAA0196* silenced PC-3 cells formed colonies as much as the control cells. Whereas, PC-3 cells where *TCEB1* was silenced formed about one third less colonies than control cells \((p = 0.005)\). (Fig. 4c). However, NIH 3T3 cells over-expressing *TCEB1* did not form colonies on soft agar, indicating that the *TCEB1* is not able to transform the cells (data not shown).

**Gene expression profile of TCEB1 suppressed PC-3 cells**

To examine which genes are associated with *TCEB1* expression, we analyzed gene expression in *TCEB1* silenced PC-3 cells compared to control (shLUC transduced) cells by using Agilent’s Whole Human Genome Oligo Microarray (44k). Both copies of
TCEB1 probes on the array were among the 11 most downregulated genes, which indicate successful gene silencing. Altogether 361 genes were found to be changed at least 2-fold (Supp. Table 1). Among the 10 most downregulated genes were cytoskeleton protein ankyrin G (ANK3), putative extracellular matrix protein chordin-like 2 (CHRD12) and protease ADAM metallopeptidase with thrombospondin type 1 motif, 12 (ADAMTS12). One of the most upregulated gene was angiopoietin/angiopoietin-like gene family member, angiopoietin-like 4 (ANGPTL4). Table II shows 25 most downregulated or upregulated genes with existing annotations. All 369 at least 2-fold changed genes were included in gene ontology analysis, and of those, 181 had existing GO-annotation. Enriched pathways are found in Supp. Table 2.

Discussion

Gain of 8q is one of the most common chromosomal aberrations in advanced prostate cancers. It is also found in breast, bladder and ovarian cancers. In prostate cancer, there are several minimal regions of amplification indicating more than one target gene. Our group has identified 4 genes that are amplified in prostate cancer and whose overexpression is associated with the amplification.7,8,9 Three of the genes, RAD21, EIF3S3 and KIAA0196 are located in the minimal amplification region 8q23-24 and one of the genes, TCEB1, is in the minimal amplification region 8q21.

Originally, we identified TCEB1 as a candidate target gene as its amplification was associated with the expression in prostate...
cancer cell lines. Here, we confirmed the increased expression also in hormone-refractory tumors. We showed also that silencing of \textit{TCEB1} in PC-3 cells reduced invasion. The role of \textit{TCEB1} in invasion is in agreement with the expression as invasion and metastasis are late stage events in tumorigenesis. We also wanted to explore if the effects on invasion were seen in a cell line which does not have an amplification of \textit{TCEB1}. \textit{TCEB1} was silenced in DU145 prostate cancer cell line, which is aggressive and capable of invasion through Matrigel but does not carry 8q gain. We showed that in DU145 cells, \textit{TCEB1} silencing reduce invasion. This indicates that \textit{TCEB1} promotes invasion with or without amplification.

\textit{TCEB1} silencing also suppressed anchorage-independent growth. However, \textit{TCEB1} expression alone was not enough to transform NIH 3T3 cells. This indicates that \textit{TCEB1} is not classical transforming oncogene.

\textit{TCEB1} has several functions in cell. It is part of the Elongin complex (SIII) that works as a transcription activator. \textit{TCEB1} invasion promoting capability could be due to the enhancing expression of invasion related genes. \textit{TCEB1} has also functions in cytoplasm. There it is part of the von Hippel-Lindau (VHL) protein complex. VHL targets the oncogene HIF1\textalpha for degradation selectively in the presence of oxygen. Inherited VHL-disease is caused by inactivating mutations in VHL which causes multiple tumors in different organs. Mutations in VHL are frequently found also from kidney cancers. \textit{TCEB1} is not the only protein \textit{TCEB1} binds to. It has been shown that many SOCS-box proteins assemble with Elongin C, Elongin B and the additional components Cul-5 and Rbx2. This means that \textit{TCEB1} has a role in various complexes that involve targeting proteins for degradation.

To study the transcriptional mechanisms how \textit{TCEB1} is involved in invasion and metastasis, we profiled \textit{TCEB1} silenced PC-3 cells with whole genome microarray. More than 300 genes

\begin{table}
\centering
\caption{The 25 most downregulated or upregulated genes in the \textit{TCEB1} silenced PC-3 cells according to the microarray analysis.}
\begin{tabular}{llr}
\hline
Gene name & Description & Ratio \\
\hline
\textbf{Downregulated} & & \\
1 & PAP2D & Phosphatidic acid phosphatase type 2 & 0.15 \\
2 & ANK3 & Ankyrin 3, epithelial & 0.17 \\
3 & CHRDL2 & Chordin-like 2 & 0.18 \\
4 & LUM & Lumican & 0.19 \\
5 & TCEB1 & Transcription elongation factor B, polypeptide 1 & 0.20 \\
6 & RSHL3 & Radial spokehead-like 3 & 0.20 \\
7 & ADAMTS12 & ADAM metallopeptidase with thrombospondin type 1 motif, 12 & 0.22 \\
8 & FABP4 & Fatty acid binding protein 4, adipocyte & 0.22 \\
9 & TCEB1 & Transcription elongation factor B polypeptide 1 & 0.25 \\
10 & DSEL & Dermatan sulfate epimerase-like & 0.27 \\
11 & SPATA4 & Spermatogenesis associated 4 & 0.28 \\
12 & PAG1 & Phosphoprotein associated with glycosphingolipid microdomains 1 & 0.30 \\
13 & GDF6 & Growth differentiation factor 6 & 0.30 \\
14 & RNU12 & RNA, U12 small nuclear & 0.31 \\
15 & BC069749 & STAM binding protein-like 1 & 0.33 \\
16 & IGFBP5 & Insulin-like growth factor binding protein 5 & 0.33 \\
17 & GCK & Glucokinase & 0.34 \\
18 & AW377662 & Keratin associated protein 4–7 & 0.35 \\
19 & IL23A & Interleukin 23, alpha subunit p19 & 0.36 \\
20 & CXCL6 & Chemokine (C-X-C motif) ligand 6 & 0.36 \\
21 & TGM2 & Transglutaminase 2 & 0.36 \\
22 & ARSB & Arsulfatase B & 0.37 \\
23 & LOC204010 & Similar to 40S ribosomal protein SA (P40) & 0.37 \\
24 & IL13RA2 & Interleukin 13 receptor, alpha 2 & 0.37 \\
25 & DMD & Dystrophin & 0.38 \\
\textbf{Upregulated} & & \\
1 & THC2780053 & 50S ribosomal protein L16 & 9.32 \\
2 & ANGPTL4 & Angiopoietin-like 4 & 6.15 \\
3 & PFFIA4 & Protein tyrosine phosphatase, receptor type, f polypeptide & 4.58 \\
4 & STC1 & Stanniocalcin 1 & 3.69 \\
5 & IGF2 & Insulin-like growth factor 2 & 3.47 \\
6 & LTB & Lympothoxin beta (TNF superfamily, member 3) & 3.41 \\
7 & KRT37 & Keratin 37 & 3.32 \\
8 & CA9 & Carbonic anhydrase IX & 3.13 \\
9 & RDH12 & Retinol dehydrogenase 12 & 3.11 \\
10 & EIF4EBP1 & Eukaryotic translation initiation factor 4E binding protein 1 & 3.08 \\
11 & SEC14L1 & SEC14-like 1 & 3.06 \\
12 & NDRG1 & N-my c downstream regulated gene 1 & 2.97 \\
13 & TLE1 & Transducin-like enhancer of split 1 & 2.97 \\
14 & LCN2 & Lipocalin 2 & 2.96 \\
15 & CFLP1 & Cofilin pseudogene & 2.95 \\
16 & MDM & Midkine (neurite growth-promoting factor 2) & 2.94 \\
17 & HIF0 & H1 histone family, member 0 & 2.87 \\
18 & ZNF580 & Zinc finger protein 580 & 2.87 \\
19 & ALDH1A3 & Aldehyde dehydrogenase 1 family, member A3 & 2.87 \\
20 & WFDC5 & WAP four-disulide core domain 5 & 2.86 \\
21 & S100A3 & $\alpha$1 cement binding protein A11 & 2.82 \\
22 & VIPR1 & Vasoactive intestinal peptide receptor 1 & 2.69 \\
23 & FAM100A & Family with sequence similarity 100, member A & 2.66 \\
24 & MRPL2 & Mitochondrial ribosomal protein L2 & 2.64 \\
25 & SORBS3 & Sorbin and SH3 domain containing & 2.56 \\
\hline
\end{tabular}
\end{table}
were found to be at least 2-fold changed, and of those less than 200 had existing annotation. Profiling revealed downregulation of extracellular matrix components ANK3 and CHRDLD2, and metalloproteinase domain containing ADAMTS12, which were all among top 10 of most downregulated genes in the array. Ankyrins are a family of proteins that link the integral membrane proteins to the underlying spectrin-actin cytoskeleton. In one study ANK3 was one of the genes in 11 gene signature which associated with prognosis of several cancers, including prostate cancer. Decreased expression of ANK3 was associated with good prognosis phenotype. Another extracellular matrix component, CHRDLD2, also known as BNF-1, is found to be upregulated in tumors. Disintegrin and metalloproteinases ADAMS and ADAMTS are involved in various biological events and many of them are expressed in malignant tumors. For example ADAM12 has been shown to digest extracellular matrix compounds, gelatin, Type IV collagen and fibronectin. ADAMTS12 is known to be proteinase type, but detailed information about its substrates and biological functions is not known. One of the most upregulated gene in the array was ANGPTL4 which is suggested to inhibit invasion and motility. Thus, the association of TCEB1 with the expression of these genes could, at least partly, explain the mechanism how TCEB1 is promoting invasiveness of prostate cancer cells.

Another suggested target gene for the amplification of 8q21 in prostate cancer is a TPDS2 family member, hD52 (also known as PrL2). When expressed in NIH 3T3 cells, the murine orthologue of TPDS2 (mD52) was able to transform the cell line and form metastatic tumors in mice. However, no human cell lines and human hD52 were included in the study.

The most studied candidate target gene for 8q24 amplification is the well-known oncogene MYC. In some studies, amplification of MYC has been shown to correlate with its overexpression and poor prognosis of prostate cancer. We have also demonstrated amplification of MYC in prostate cancer. However, the expression was not associated with amplification. In addition, the MYC expression was not found to be increased in prostate cancer in general. Here, E1F3S3 located in 8q23-24 had a minor positive effect on the growth of the NIH 3T3 cells, which we have also previously reported. Whereas, RAD21 and KIAA0196 did not seem to have any significant role in growth, anchorage-independent growth or invasion in vitro. Altogether the functional data do not support the hypothesis that RAD21 or KIAA0196 would be the target genes, instead the may be coincidentally amplified. One shortcoming of this study is that the effects of the genes were analyzed one gene at the time. Since the gain of 8q most often comprises the whole long arm of the chromosome, it is possible that the target genes of the gain have synergistic effects. Thus, also the combined effects of the putative target genes should be studied in the future.

Taken together, the data suggest that TCEB1 is an invasion and metastasis promoting gene, and therefore a strong candidate target gene for the amplification of 8q21.

Acknowledgements

The authors thank Ms. Mariitta Vakkuri for her technical assistance.

References


TITLE: DIAGNOSTIC AND PROGNOSTIC SIGNATURES FROM THE SMALL non-coding RNA TRANSCRIPTOME IN PROSTATE CANCER

AUTHORS: Elena S. Martens-Uzunova, PhD1*, Sanni E. Jalava, MSc2*, Natasja F. Dits1, Dr. Geert J.L.H. van Leenders, MD, PhD3, Søren Møller, PhD4, Prof. Dr. Jan Trapman, PhD3, Prof. Dr. Chris H. Bangma MD, PhD1, Thomas Litman, PhD4, Prof. Dr. Tapio Visakorpi, MD, PhD2, Guido Jenster, PhD1

RUNNING TITLE: The small transcriptome of prostate cancer

AFFILIATIONS: 1Department of Urology and 3Pathology, Josephine Nefkens Institute, Erasmus MC, Rotterdam, The Netherlands; 2Institute of Medical Technology, University of Tampere and Tampere University Hospital, Tampere, Finland; 4Exiqon A/S, Vedbaek, Denmark

*Both Authors contributed equally

CORRESPONDING AUTHOR:

Guido Jenster, PhD
Assistant Professor of Urology
Department of Urology, Josephine Nefkens Institute, Erasmus MC,
room Be362a, P.O. Box 2040, 3000 CA Rotterdam, The Netherlands
Phone: + 31 10 7043672; Fax: + 31 10 704 4661; E-mail: g.jenster@erasmusmc.nl

FINANCIAL SUPPORT: European Community's Seventh Framework Program (FP7/2007-2013), grant agreement No HEALTH-F2-2007-201438, Academy of Finland; Cancer Society of Finland; Reino Lahtikari Foundation; Sigrid Juselius Foundation; and the Medical Research Fund of Tampere University Hospital.
ABSTRACT
Prostate cancer is the most frequent male malignancy and the second most common cause of cancer-related death in Western countries. Current clinical and pathological methods are limited in the prediction of postoperative outcome. It is becoming increasingly evident that small non-coding RNA (ncRNA) species are associated with the development and progression of this malignancy. To assess the diversity and abundance of small ncRNAs in prostate cancer we analyzed the composition of the entire small transcriptome by Illumina/Solexa deep sequencing. We further analyzed the microRNA (miRNA) expression signatures of 102 fresh-frozen patient samples during prostate cancer progression by miRNA microarrays. Both platforms were cross-validated by quantitative RT-PCR. Besides the altered expression of several miRNAs, our deep sequencing analyses revealed strong differential expression of small nucleolar RNAs and tRNAs. From microarray analysis, we derived a miRNA diagnostic classifier that accurately distinguishes normal from cancer samples. Furthermore, we were able to construct a prostate cancer prognostic predictor that independently forecasts postoperative outcome. Importantly, the majority of miRNAs included in the predictor also exhibit high sequence counts and concordant differential expression in Illumina PCa samples, supported by quantitative RT-PCR. Our findings provide miRNA expression signatures that may serve as an accurate tool for the diagnosis and prognosis of prostate cancer.

Key words: prostate cancer; microRNA; snoRNA; microarray; deep sequencing; Q-PCR
INTRODUCTION
Prostate cancer is the most common malignancy and the second cancer-related cause of death among Western men (Ferlay et al., 2007). Fatal outcome from prostate cancer is generally preceded by local tumor infiltration beyond the prostate and metastasis to lymph nodes. Although the majority of prostate cancer patients are diagnosed with curable, organ-confined disease, more than a quarter of them experience a relapse within five years after surgery (Wright and Simon, 2003). Due to their heterogeneous character, it is difficult to distinguish indolent from more-aggressive prostate neoplasms (Marberger, 2009). Accurate identification of patients with an increased risk of postsurgical prostate cancer recurrence is of major importance for the determination of adequate adjuvant therapy. Current prediction models are based on standard clinical variables like preoperative prostate specific antigen (PSA), Gleason score, and pathologic stage. However, their performance is unsatisfactory regarding preoperative evaluation of treatment choice and postsurgical clinical outcome. A better understanding of the biologic mechanisms of prostate cancer formation and progression is crucial for the discovery of new markers for this disease.

In recent years, it has become apparent that different non-coding RNAs (ncRNAs) are also implicated in prostate cancer. Several microRNAs (miRNAs) are now associated with the progression and classification of this and other malignancies (Lu et al., 2005; Ozen et al., 2008; Porkka et al., 2007; Prueitt et al., 2008; Shi et al., 2007; Tong et al., 2009; Volinia et al., 2006). Evidence is accumulating that small nuclear RNA (snRNA) and small nucleolar RNA (snoRNA) are also employed in cis and trans gene regulation and alternative splicing besides their established function in rRNA modification (Dong et al., 2008; Ender et al., 2008; Mattick and Makunin, 2006; Reis et al., 2004). Nevertheless, the assessment of the complete small ncRNA transcriptome involved in the formation and differentiation of different types and subtypes of PCa has remained a challenge.

In this study, we examined the expression of the entire small transcriptome in prostate cancer using Illumina/Solexa deep sequencing. We further analyzed the expression signature of more than 700 mature miRNA species during progression of prostate cancer disease by Agilent microarrays and were able to derive differential miRNA expression profiles strongly associated with prostate cancer diagnosis and clinical outcome. A selected subset of miRNAs was further used to cross-validate both platforms by quantitative real-time-PCR (Q-PCR).

RESULTS
Dynamics of the small transcriptome in prostate cancer

We used Illumina/Solexa deep sequencing to examine the ncRNA content in two prostate cancer small RNA libraries: organ-confined prostate cancer (PCa) and metastatic lymph node PCa (LN-PCa).

Both sequencing reactions yielded 5 547 066 unfiltered sequence reads, for PCa and 4 915 053 for LN-PCa. For PCa, 4 517 214 (81%) reads were mapped to 545 900 individual loci in the Human genome. For LN-PCa, 3 524 014 (78%) reads were mapped to 395 281 individual loci. Identified loci were further classified as known miRNA, snoRNA, snRNA, rRNA, tRNA, fragments of large ncRNA, genomic repeats, mRNA, and other hairpins originating from yet un-annotated genomic regions (Table S1). MiRNAs were the most abundant RNA detected, comprising approximately 95% of the total RNA pool. We were able to identify 725 mature miRNA species deriving from 427 known human stem-loop sequences (miRBase 12.0). Furthermore, 20 candidate miRNA sequences originating from novel genomic miRNA-loci were discovered (Table S2). These loci correspond to novel miRNAs with very low sequence counts. In addition, two miRNAs homologous to the known rat rno-mir-320 and mouse mmu-mir-449c were detected.

The overall expression of miRNA species was relatively decreased in LN-PCa as demonstrated by the noticeably lower total sequence count of miRNAs in LN-PCA (3 278 482) compared to PCa (4 232 318) (Table S1). Strikingly, the opposite tendency was observed for sequenced fragments mapped to tRNA and snoRNA loci. The total count of tRNA and snoRNA fragments was increased more than 20% in LN-PCa compared to PCa. tRNA fragment levels were elevated from 58 224 to 84 709 total sequence counts, while snoRNA fragments had corresponding total sequence counts of 16 762, and 20 569, in the PCa, and LN-PCa samples, respectively. Similarly, the diversity of detected snoRNA species was increased in the LN-PCa library. A search with the mapped sequences from both sample libraries against known snoRNAs identified 112, and 158 unique snoRNA loci encoding either C/D-box or H/ACA box snoRNAs (Fig 1A). This global pattern of differential expression was also observed on the level of individual small ncRNAs. Only 19 miRNAs were found to be up-regulated in metastatic lymph node compared to PCa, while 69 were down-regulated. While the majority of differentially expressed ncRNAs species in PCa were miRNA, in the LN-PCa the most abundant ncRNA species were sno- and tRNAs (Fig 1B, Table S3).

Mature tRNA have a size of approximately 75 nucleotides (nt), while snoRNAs range in size from 70 to 200 nt depending of their class (C/D box or H/ACA box) and cellular function
(guiding of RNA methylation or pseudourydilation). The majority of sequence reads mapped to snoRNA in both sequencing libraries had a size of 22-23 nt similarly to miRNA (Supplementary Figure S1). Most tRNA derived fragments identified in the PCa library had a size of ~18nt, while in LN-PCa library a peak in the size range of 27 nt was observed, indicating different processing of tRNA in LN-PCa (Supplementary Figure S1).

To assess whether the relative down-regulation of miRNAs observed in the LN-PCa sample is mediated by alterations in the miRNA processing pathways, we examined the expression of the miRNA processing enzymes Drosha and Dicer in each one of the samples used for the generation of our deep sequencing libraries by Q-PCR. No significant change in the expression levels of these enzymes was found between PCa and LN-PCa (Supplementary Figure S2a and S2b). This result was further confirmed by the analysis of an extended sample cohort consisting of 60 PCa and 16 LN-PCa samples (Supplementary Figure S2c and S2d).

**Transcriptional profiling of the miRNA transcriptome**

*Class discovery of miRNAs associated with prostate cancer*

To examine changes in expression of individual miRNAs that may be relevant to the occurrence and progression of clinically significant prostate cancer, we performed miRNA microarray analysis on a patient cohort consisting from 102 clinical tissue samples divided over six sample groups (Table 1).

Using a class discovery approach, we investigated whether miRNA expression profiles can distinguish normal from malignant tissue and further discriminate different prostate cancer subtypes. 513 from the 723 human microRNAs included on the chip, displaying significant hybridisation signal levels were tested for differential expression between the normal and the corresponding cancer group in a pairwise fashion. From these, 80 miRNAs were identified as significantly changing between normal and malignant tissues (Figure 2a and Table S4). These 80 miRNAs were used for the hierarchical clustering of all samples, which clearly separated non-malignant from advanced prostate cancer specimens and further divided organ-confined PCa samples into two groups (Figure 2a). Group I, clustering with normal adjacent prostate (NAP), normal lymph node (LN-normal), and trans-urethral resection of the prostate (TURP-normal), contained 34 PCa tumors. In group II, 16 PCa tumors clustered together with samples derived from patients with advanced prostate cancer (LN-PCa and TURP-PCa). Differences between both PCa groups reflected distinct prognostic characteristics, as group II PCa tumors were strongly associated with significantly increased rate of metastases occurrence after radical prostatectomy (P<0.0001), cancer-related death (P=0.0007), *ETV1* alterations (P=0.0396), and Gleason score
higher than 6 (P=0.0272) (Figure 2b). Poor outcome of group II PCa tumors was independent of the TMPRSS:ERG fusion status, the clinical stage of the tumor, or PSA progression. The average percentage cancer and percentage epithelium in tissue between both groups were significantly different (P=0.0010, and P=0.0261, respectively) however, no significant correlation was found between these two parameters (Table S5).

Within the dataset of 80 miRNAs, 22 were identified to significantly contribute to the separation of group I and II PCa tumors (Figure 2c). Subsequent clustering based on the transcriptional profiles of these 22 miRNAs accurately separated all PCa tumors from group II in one of the branches of the hierarchical tree, while 31 of the 34 (91%) PCa group I tumors fell within the other branch suggesting that miRNAs can be used to group tumors per outcome, thus have predictive power. As group II PCa tumors were associated with clinical progression and development of metastases after radical prostatectomy, we also performed a supervised analysis for miRNAs differentially expressed between organ confined PCa tumors that did form metastases and those that did not. Supervised clustering based on metastatic status was in an overall good agreement with the results obtained from unsupervised analyses. 14 miRNAs were found to be common in both analyses (Supplementary Figure S3). Thus, we were able to identify distinct sets of miRNAs which could distinguish normal from malignant samples and could be associated with disease relapse after radical prostatectomy.

MiRNAs involved in prostate cancer progression

To also examine if further alterations of miRNA expression occur during disease progression, we tested miRNAs for differential expression between PCa and LN-PCa samples, and compared the expression levels of the identified miRNAs with data obtained from primary prostate stroma cells (PrSC) and primary prostate epithelial cells PrEC cells. 70 miRNAs were found significantly differentially expressed between PCa and LN-PCa (Supplementary Figure S4). From them, 51 were down regulated in LN-PCa while 19 were up regulated. Comparison of the expression of these 70 miRNAs in stroma vs. epithelium primary cells demonstrated that the majority of miRNAs with abolished expression in metastatic lymph node were as expected from stromal origin. Nevertheless, several miRNAs with decreased expression in group II PCa tumors were further down regulated in malignant lymph node, namely miR-1, -133a, -133b, -143, -143*, -145, -145*, -204, -221, and -222. Importantly, the same set of miRNAs was identified as strongly down regulated in LN-PCa also in our deep sequencing libraries (Figure 1b). MiR-145 and miR-222 were also found to be significantly down regulated in malignant compared to normal lymph node (Supplementary Table S4).
Construction and performance of miRNA derived diagnostic classifier

We further evaluated the predictive power of miRNA expression profiles by training a Bayesian Covariate Compound Predictor Algorithm (Radmacher et al., 2002). To avoid possible over-fitting, we estimated the prediction error using Leave-One-Out Cross-Validation (LOOCV) (Simon et al., 2003). We designed a diagnostic miR-Classifier that can distinguish NAP from PCa samples. As an input, we used all 513 significantly expressed miRNAs in the entire microarray set. The best statistically significant miR-Classifier (P=0.0130, 1000 random permutations) consisted of 54 miRNAs significantly different between the classes (P =0.0005, random variance t-test) (Table S6). At a posterior probability threshold of 0.7, the diagnostic miR-Classifier correctly classified 54 out of 61 samples (89%). ROC curve analysis resulted in an AUC of 0.949, with class sensitivity of 0.636 and class specificity of 0.94 for group NAP. The positive and negative predictive values for the same group are 0.7 and 0.922, respectively.

We further validated the miR-Classifier on an independent data set recently reported by Schaefer and colleagues (Schaefer et al., 2010). All 24 paired NAP and PCa samples from that study were categorized based on the expression profiles of 43 out of the 54 miRNAs included in our miR-Classifier. The remaining 11 miRNAs were not included in the miRNA microarray platform used by the authors. Using that external dataset and the same testing parameters as described above, our miR-Classifier performed equally well in correctly predicting the class of the samples, achieving class sensitivity of 1 and class specificity of 0.667 for group NAP. The positive and negative predictive values for the same group are 0.75 and 1, respectively. ROC analysis revealed that MiR-Classifier performed equally well on both datasets with no significant difference (P=0.2534) between the derived AUCs (Figure 3).

Construction and performance of miRNA derived prognostic predictor

Similarly to the construction of the miR-Classifier we designed a diagnostic miR-Predictor that could distinguish group I (good prognosis) from group II (poor prognosis) PCa samples.

The best statistically significant predictor (P=0.0410, 1000 random permutations) with the lowest cross-validation mis-classification rate consisted of 25 miRNAs (Figure 4a) significantly different between group I and group II tumors (P =0.0005, random variance t-test). 15 of these miRNAs are also included in miR-Classifier and importantly, follow the same direction of expression change (Supplementary Figure S5). At a posterior probability threshold of 0.7, miR-Predictor correctly classified 40 out of 50 samples (11 to group II and 29 to group I). 10 samples were not assigned to either of both groups (Figure 4a and 4c). MiR-Predictor reached an AUC of 0.991 (Figure 4b) and had class sensitivity and class specificity of 0.912 and 0.750, respectively.
for group I PCa tumors with positive predictive value of 0.886 and negative predictive value of 0.800 for the same group.

To assess miR-Predictor as a prognostic factor, we compared its performance to that of other clinical parameters (Fig 5). For the described here data set, at an end point of 12 years the prognostic signature of miR-Predictor performed better than the other tested clinical parameters with lower survival probability for patients assigned to the poor outcome group (0.3463, SE=0.1600) compared to Gleason score (0.4375, SE=0.1879), preoperative PSA (0.429, SE=0.1574) or pathological stage (0.7272, SE=0.1767).

**Comparison of Solexa deep sequencing and Agilent microarray platforms and Q-RT-PCR validation**

To cross validate the results obtained from both platforms, we compared the miRNA expression data obtained from Illumina/Solexa deep sequencing and Agilent miRNA microarrays. We identified 725 miRNA sequences in both sequencing reactions. Agilent microarrays used in this study contain probes for 723 miRNAs. In all, 483 (67%) miRNAs were represented by both platforms (Figure 6a).

The Agilent microarrays contain probes for 240 miRNAs that were not detected by deep sequencing. The hybridization signal of a substantial number of these probes (approximately 53%) was below the recognized threshold levels in all 102 samples, consequently these miRNAs were assigned absent call by the Agilent Feature Extraction Software. The majority of miRNAs detected by Solexa but not by microarrays were sequences for which no probes or probes for their opposite strands have been included on the miRNA microarrays.

Comparison of log2 transformed data demonstrated that, despite the much broader dynamic range of deep sequencing reactions, the raw signals obtained from sequencing and microarrays are in good agreement since, as many as 51% of the miRNAs detected in PCa and 40% of the miRNAs detected in LN-PCa, had comparable expression levels (Figure 6b). After normalization to the mean, more than 99% of the miRNAs had a comparable scaled signal within 4 fold difference between both platforms (Figure 6c).

The good agreement between Illumina/Solexa deep sequencing and Agilent microarrays was further supported by TaqMan Q-RT-PCR assays performed on a selected set of 12 miRNAs which included ten differentially expressed miRNAs between PCa (group I and II) and LN-PCa as identified by microarrays and/or Solexa deep sequencing (miR-1, -143, -145, -205, -210, -222, -301b, -425-5p, and -451). A negative control miRNA (miR-558), which was not detected in either of both deep sequencing libraries and was assigned an absent call in more than 50% of the
microarrays derived from PCa and LN-PCa, and a non-changing control miRNA (miR-25) that was stably expressed and detected in the Solexa libraries and in all PCa and LN-PCa samples, were also included. The overall expression ratios as determined by TaqMan Q-RT-PCR assays strongly correlated with the findings from microarrays and deep sequencing (Figure 6d) since not only the direction of change (up- or down regulation), but also the relative log2 ratios of individual differentially expressed miRNAs were well comparable.
DISCUSSION

At present, consistent data regarding alterations in the expression of miRNA in prostate cancer are still limited (reviewed by (Coppola et al., 2010)) and studies addressing the genome-wide expression profiling of other small ncRNA in this malignancy are lacking. Here we utilized Solexa Illumina Deep Sequencing to examine the entire spectrum of small ncRNA in prostate cancer specimens, and further screened for changes in the expression of 723 miRNAs by microarray-profiling of 102 histologically confirmed malignant and non-malignant samples from prostatectomy, lymph node, and TURP origin.

Comparison of miRNA expression profiling results obtained by deep sequencing and microarray analyses demonstrated a high concordance in measuring miRNA expression levels between both platforms used for this study and identified similar sets of differentially expressed miRNAs. Previously, it has been reported that absolute microarray expression measures correlate better than deep-sequencing data with RNA sample-content when synthetic samples mimicking the human miRNA pool are used (Willenbrock et al., 2009). However, the correlation between absolute miRNA expression values determined by microarrays and by next-generation sequencing may differ strongly depending on the platform used for microarray profiling (Git et al., 2010), or on the algorithm used for processing of sequencing data. Therefore, to further confirm the validity of results obtained in our microarray and deep sequencing data sets we also verified the expression of a set of 12 differentially expressed miRNAs by Q-PCR.

Our deep sequencing experiments revealed that overall, many miRNAs have decreased expression in metastatic lymph node PCa compared to organ confined disease. These changes in expression were further confirmed by both microarray analysis and by Q-PCR on selected miRNAs. It has been suggested that alterations in the expression levels of Dicer may explain global changes of miRNA expression in prostate adenocarcinoma (Chiosea et al., 2006; Ambs et al., 2008). However, in this study we did not observe significant differences in the expression levels of Dicer or Drosha between the studied patient groups (Supplementary Figure S2). Alternatively, the decreased levels of many miRNAs observed here may be partially explained by their stromal origin and the loss of the stromal compartment in advanced and metastatic prostate cancer.

The independence of prior knowledge (i.e. no required probe design, which would only cover known genes) together with the very broad dynamic range of deep sequencing (over 5 orders of magnitude) allowed us to detect several novel miRNA with low expression levels in our deep
sequencing libraries. Additionally, we also identified many sequence fragments derived from tRNA and snoRNA transcripts.

We further noticed a pronounced increase of tRNA fragments in the metastatic sample pool that can be interpreted as a direct evidence of elevated RNA-polymerase III output, which in turn is a feature of transformed and cancerous cells (Marshall and White, 2008). The same was observed for several snoRNAs, providing further support for high metabolic activity and elevated protein-synthesis rate, necessary for the fast growth of tumor cells in progressing cancer.

Interestingly, in addition to their function in ribosomal maturation, snoRNA species or their derivates have been implicated in processes associated with carcinogenesis, e.g. alternative splicing events (Kishore and Stamm, 2006) and genomic imprinting (Royo et al., 2006). In prostate cancer, snoRNA U50 has been reported to be a candidate tumor-suppressor gene and a mutation in its sequence has been associated with clinically significant disease (Dong et al., 2008). Recent findings have demonstrated that discrete in size snoRNA fragments are produced from a vast majority of snoRNA loci (Taft et al., 2009) and that they can also function as miRNAs (Ender et al., 2008; Brameier et al., 2011). Additional research is needed to investigate whether this unconventional function of snoRNAs is a wide spread mechanism of gene regulation and whether this holds true for snoRNAs identified in this study.

By performing microarray analysis we were able to identify a miRNA signature that not only distinguished normal from malignant specimens, but also identified two subclasses within the group of patients with organ confined disease at the time of surgery. Comparison of this miRNA signature to known clinicopathologic parameters associated the separation of both subclasses with postoperative occurrence of clinical metastases and poor disease outcome. To evaluate whether the observed associations have a potential clinical diagnostic or prognostic value; we trained a diagnostic miR-Classifier and a prognostic miR-Predictor using the Bayesian compound covariate method followed by LOOCV analysis. The miR-Classifier consists of a set of 54 miRNAs that clearly discriminated normal adjacent from prostate cancer specimens. This miR-Classifier performed equally well on the independent microarray dataset used for validation (Schaefer et al., 2010) despite the differences in microarray platform, experimental design, and sample size, providing evidence that the included miRNAs may have a potential value as diagnostic markers for prostate cancer. Among the miRNAs included in the miR-Classifier, miR-205 is of particular interest. This miRNA was the most strongly down regulated miRNA in PCa compared to NAP and its
expression is completely abolished in metastatic lymph node and malignant TURP samples. It has been suggested that the tumor suppressive function of miR-205 that takes place through counteracting epithelial-to-mesenchymal transition and reducing cell migration and invasion in the human prostate (Gandellini et al., 2009; Majid et al., 2010). K27me3 modifications of the miR-205 locus was reported to occur in the prostate cancer cell line PC3 (Ke et al., 2009) and in muscle invasive bladder tumors and undifferentiated bladder cell lines (Wiklund et al., 2010). These findings, together with the early loss of miR-205 during prostate cancer progression observed by us and others (Ambs et al., 2008; Porkka et al., 2007; Schaefer et al., 2010) present miR-205 as an attractive diagnostic marker for prostate cancer disease.

The MiR-Predictor consists of 25 miRNAs, from which 13 are down regulated and 12 are up-regulated in prostate cancers with poor outcome. The combined expression profiles of these prognostic miRNAs are very specific for patients at high risk of clinical relapse and their combined predictive accuracy was higher compared to other clinical variables available for the patients included in the study cohort. Unfortunately, because of the limited number of radical prostatectomy samples with a clinical follow-up that includes clinical metastasis occurrence and prostate cancer-related death, and the low quantities of available material, it is difficult to obtain an independent validation set for MiR-Predictor. To our knowledge, there are no publically available data sets on genome-wide miRNA expression profiling in prostate cancer that comprise samples with sufficiently long follow-up needed for independent testing of MiR-Predictor. To partially overcome this drawback, we implied the LOOCV method to reduce over-fitting of our diagnostic and predictive models within our patient cohort. LOOCV is a validation method that iteratively chooses one sample from a given dataset to be the validation sample and uses the remaining samples to build a training set. For each round of iteration, the entire model construction is repeated starting with the selection of significant miRNAs. The predictive power of each of the miRNAs included in the final model is represented as the percentage of the iteration steps in which that miRNA has been identified as predictive. In our MiR-Predictor we included only miRNAs that had 100% cross-validation support and therefore were predictive for any of the tested combinations. We also controlled the misclassification error rate by calculating a label-permutation p-value that indicates the chance of obtaining the same or lower misclassification for two groups of samples when their sample labels are randomly assigned as compared to the real class labels. However, considering the small sample size of our patient cohort it should be noted that it is not an equivalent to a validation on an independent dataset.

37
Nevertheless, the robustness of the method was demonstrated by the validation of the identically constructed MiR-Classifier on the external dataset published by Schaefer and colleagues. Several miRNAs included in MiR-Predictor have been recently reported to be deregulated in clinical prostate cancer samples. Down regulation of let-7b, miR-1, miR-133a, miR-143, miR-145, miR-221, and miR-222 and up-regulation of miR-25, miR-93, mir-96, miR-183, miR-182 or miR-301b has been reported in prostate cancer when compared to BPH or normal prostate tissue by at least one research group (Ozen et al., 2008; Porkka et al., 2007; Schaefer et al., 2010; Spahn et al., 2010; Szczyrba et al., 2010; Tong et al., 2009; Volinia et al., 2006). In addition, over-expression of miR-96 and reduced expression of miR-221 have been associated with increased risk of biochemical recurrence and aggressive prostate cancer (Schaefer et al., 2010; Spahn et al., 2010).

The concordant expression of these miRNAs could be in part explained by their genomic organization as mir-1/133, mir-143/145, mir-221/222, and mir-96/182/183 are encoded in genomic clusters and most likely their expression levels are regulated by identical regulatory events. A recent analysis of the methylation signature in prostate cancer cell lines demonstrated that indeed histone methylation leads to the silencing of the entire mir-221/222 cluster (Ke et al., 2009). Deregulation of several of the down-regulated miRNAs identified in our prognostic profile has been demonstrated also in other human malignancies or diseases such as breast, (Wang et al., 2009) and colorectal cancer (Michael et al., 2003) for the mir-143/145 cluster; breast cancer (Zhao et al., 2008), and glioblastoma (Gillies and Lorimer, 2007) for the mir-221/222 cluster; and liver cancer (Datta et al., 2008) or cardiac hypertrophy (Care et al., 2007) for the mir-1/133 cluster.

MiRNAs function in carcinogenesis and cancer progression by modulating the expression levels of, among others, tumor-suppressor genes and oncogenes. To examine the functional relevance of miRNAs identified in this study, we performed target-gene network search for the miRNAs included in MiR-Classifier and MiR-Predictor by Ingenuity Pathways Analysis (Ingenuity® Systems) (Supplementary Data File 1). MiRNA target-genes were significantly enriched in gene networks associated with cancer and reproductive system disease affecting cellular development, cellular growth and proliferation, cell-to-cell signaling, cell morphology, and cell cycle (Supplementary Figures S6 and S7). Several up-regulated miRNAs identified in this study, e.g. miR-19a, miR130a, and the mir-20a/106/93 family, target key cancer genes involved in prostate cancer signaling. Deregulation of these miRNAs may cause defects in cell cycle checkpoint control and further promote cell cycle progression by decreasing the expression levels of tumor
suppressors PTEN, p21, and Rb1, which are commonly inactivated in prostate cancer (reviewed by Lee et al., 2008). In addition, decreased expression of miR-27, miR-143 and miR-221/222 may promote increased expression of Notch1, MAPK kinases, and c-KIT, which have been associated with epithelial-mesenchymal transition, angiogenesis and metastasis in prostate cancer (Bin Hafeez et al., 2009; Paronetto et al., 2004; Mukherjee et al., 2011).

Deregulation of miRNAs in response to altered androgen signaling in PCa has been studied previously (reviewed by Catto et al., 2011). However, based on cell line and xenograft studies, only few miRNAs have been suggested to be androgen regulated. Within the miRNAs included in MiR-Classifier and miR-Predictor, only miR-21, miR-32, miR-141, miR-221, and miR-375, demonstrate changes of expression concordant with previously published data. For example, it has been shown that elevated expression of miR-21 can enhance PCa tumor growth in vivo and, is sufficient for androgen-dependent tumors to overcome castration-mediated growth arrest. (Ribas et al., 2009). We recently identified miR-21 together with miR-32, miR-141, miR-221, and miR-375 as differentially expressed during DHT stimulation of different PCa cell lines or 13 intact-castrated pairs of PCa-xenografts (Waltering et al., 2011). It has been reported that the mir-221/222 cluster is up-regulated in matched androgen-dependent and androgen-independent PCa cell lines (Sun et al., 2009). In the patient cohort presented in this study, the expression of the diagnostic and prognostic miR-221/222 is strongly down-regulated during disease progression. Similar down-regulation was observed for PCa cell lines (Waltering et al., 2011). In the latest study, only four miRNAs showed similar regulation in both, cell lines and xenograft-pairs, and only miR-141 demonstrated concordant expression in castration resistant PCa clinical specimens when compared to PCa. The dissimilarity between the response of different miRNAs to androgen stimulation in cell lines and xenograft models and when compared to expression in clinical samples, may be explained by the direct response to androgens in in vitro studies compared to the long term effects of androgen deprivation in vivo (Waltering et al., 2011). In addition, the diagnostic and prognostic profiles obtained in this study are based on clinical samples derived early, from patients with organ confined disease, years prior clinical relapse and androgen deprivation therapy.

Taken together, our findings support the role of miRNAs, and implicate snoRNAs in prostate oncogenesis. The discovery of a subset of miRNAs, which is associated with the aggressiveness of the tumors, suggests that this regulatory system is, at least, partly responsible for the progression of the disease. Our miRNA predictor forecasts PCa recurrence after radical prostatectomy in the early organ confined stage prior clinical progression. Compared to other
prognostic models currently applied in clinical practice, miR-Predictor demonstrates enhanced accuracy and suggests novel directions into prognostic prediction of prostate cancer. It is now warranted to evaluate whether this panel of miRNA could be clinically useful in optimizing the treatment strategies for the prostate cancer patients.

MATERIALS AND METHODS

Patient samples, clinical database and cell lines

For microarray analysis 102 freshly frozen clinical samples were obtained from the tissue bank of the Erasmus University Medical Center. Clinical parameters of the patient cohort are listed in Table 1. Collection of patient samples has been performed according to national legislation concerning ethical requirements. Use of these samples has been approved by the Erasmus MC Medical Ethics Committee according to the Medical Research Involving Human Subjects Act (MEC-2004-261).

Two sample pools each comprised of four individual patient samples were used for Solexa/Illumina deep sequencing (Table S7).

Primary PCa and normal prostate samples were obtained by radical prostatectomy. Samples were snap frozen and stored in liquid nitrogen. Histological evaluation and Gleason grading were performed independently by two pathologists on hematoxylin/eosin-stained frozen sections for all analyzed material. Tumor samples in which at least 60% of the cells were cancerous and containing at least 60% ducts were selected for analysis. The non-tumor samples contained 0% tumor cells and at least 60% ducts. The latter samples were considered “normal” for this study, although they were derived from prostates that contained adjacent tumor epithelium.

Primary human prostate stromal cells (PrSC) and primary human prostate epithelial cells (PrEC) were purchased and cultured according to manufacturer’s guidelines (Clonetics Human and Animal Cell Systems, Cambrex Bio Science Walkerville Inc., USA).

RNA isolation

Total RNA was isolated from frozen tissue samples using RNABee reagent (Campro Scientific, GmbH) according to the manufacturer’s protocol.

Illumina/Solexa small RNA library preparation, genome mapping and annotation

Total RNA sample pools for small RNA library preparation and sequencing on the Illumina 1G Genome Analyzer was outsourced to ServiceXS, Leiden, The Netherlands. RNA library preparations were performed according to the “Small RNA Sample Prep v1.5.0 (Prerlease protocol, Rev-D)”, (Illumina Inc.,www.illumina.com). Shortly, total RNA pools were size fractionated on 15% TBE urea PAGE gel, and the small ncRNA fraction in the size range of 18 to
30 nt was extracted and purified. After 5’ and 3’ adapter ligation, cDNA library ready for sequencing was constructed by reverse transcription with SuperScript II Reverse Transcriptase (Invitrogen) followed by 15 cycles of PCR by Phusion DNA Polymerase (Finnzymes Oy) as previously described (Morin et al., 2008).

Raw reads obtained after sequencing were processed using the mir-Intess™ pipeline by InteRNA Genomics BV, the Netherlands (http://www.internagenomics.com), which provided annotation of miRNA reads according to Homo sapience genome (NCBI build 36) as well as identification of novel miRNA candidate genes (Berezikov E et al. 2006; Wit, et al. 2009). Mapping and annotation of aligned non-miRNA reads with a minimum sequence length of 18nt (Suplementary Figure S1) which contained at least 3 recognizable nucleotides from the 3’ adapter sequence was performed by MEGABLAST, according to genomic loci annotations retrieved from the UCSC database: hg18 (http://genome.ucsc.edu/) trak tables for sno/miRNA, tRNA, and non-coding RNA genes, and MiRBase version: 12.0 (http://www.mirbase.org/) as previously described (Morin et al., 2008).

**MiRNA microarrays and analysis**

Microarray analysis of miRNA expression was performed using Human miRNA V2 microarrays (Agilent) that contain probe sets for 723 human microRNAs from the Sanger MiRBase, v.10.1. Microarray processing and analysis are described in Supplementary data file 1. The array data were submitted using Tab2MAGE to the ArrayExpress database (http://www.ebi.ac.uk/microarray-as/ae/) with accession number E-TABM-794.

**TaqMan Q-PCR**

Quantitative real-time RT-PCR analysis was done with an ABI Prism 7700 Sequence Detection System using TaqMan® MicroRNA Reverse transcription kit and TaqMan® Universal PCR Master Mix, No AmpErase® UNG according to the manufacturer’s specifications (Applied Biosystems, Foster City, CA). Mature miRNA expression in the same total RNA pools as used for deep sequencing was assessed using specific TaqMan® MicroRNA Assays (Applied Biosystems). The amount of target gene expressed was normalized to U6 and Z30 endogenous controls by the delta Ct method.

**Statistical analysis of clinical data**

The significance of contingency for categorical clinical data was examined by Fisher’s exact test (alpha level 0.05). Continuous variables between different groups were compared using two
tailed nested t-test (alpha level 0.05). Pearson correlation coefficient was used to evaluate possible correlation between clinical parameters. Kaplan Meier survival curves and log-rank tests were used to assess the predictive values of different prognostic approaches.

**MiR-Classifier and MiR-Predictor construction and evaluation**

MiR-Classifier and MiR-Predictor were constructed using identical algorithm. Models for utilizing gene expression for class prediction of future samples were developed using the Compound Covariate predictor implemented in the software package “BRB array tools” developed by Dr. Richard Simon and BRB-ArrayTools Development Team (Wright and Simon, 2003). The models incorporated genes that were differentially expressed at the 0.001 significance level as assessed by the random variance t-test. We estimated the prediction error of each model using leave-one-out cross-validation (LOOCV) (Simon R, et al. 2003). For each LOOCV training set, the entire model building process was repeated, including the gene selection process. We also evaluated whether the cross-validated error rate estimate for a model was significantly less than one would expect from random prediction. The class labels were randomly permuted and the entire LOOCV process was repeated. The significance level was defined as the proportion of the random permutations that gave a cross-validated error rate no greater than the cross-validated error rate obtained with the real data. 1000 random permutations were used.
CONFLICT OF INTEREST
Dr Litman and Dr Møller have been employed by Exiqon A/S and own stocks in that company. Dr Martens, MSc Jalava, Mrs Dits, Dr. van Leenders, Dr Trapman, Dr Bangma Dr Tapio Visakorpi, and Dr Jenster declare no potential conflict of interest.

ACKNOWLEDGEMENTS
We thank Theo van der Kwast for histo-pathological examination of patient samples. We also thank Mark Wildhagen for providing the coded clinical data.

We acknowledge the European Community's Seventh Framework Program (FP7/2007-2013), grant agreement No HEALTH-F2-2007-201438, the Academy of Finland; Cancer Society of Finland; Reino Lahtikari Foundation; Sigrid Juselius Foundation; and the Medical Research Fund of Tampere University Hospital for their support.

SUPPLEMENTARY INFORMATION
Supplementary Information accompanies the paper on the Oncogene website (http://www.nature.com/onc)
REFERENCES


FIGURE LEGENDS

Figure 1. Deep sequencing analysis of differentially expressed small non-coding RNAs in organ confined and metastatic prostate cancer. a) Number of ncRNA species per sample pool. b) Differentially expressed small RNAs between Organ-confined prostate tumours (PCa) and metastatic lymph node (LN-PCa) libraries. We compared the fold change difference in the count number of sequences mapped to the same mature snRNA in PCa and LN-PCa libraries. Sequences were considered differentially expressed when the fold change in sequence counts between the tested conditions was more than four. Plotted are the log2 transformed count frequencies of each individual ncRNA in the LN-PCa library (log2 LN-PCa) and the PCa library (log2 PCa). Coloured are small RNAs with a fold change bigger than four. In red – miRNA; blue – snoRNA; green – t RNA; gray – ncRNAs with a fold change less than 4. Only names of miRNA tested by TaqMan Q-RT-PCR are indicated.

Figure 2. Class discovery of miRNAs associated with prostate cancer outcome. A) Hierarchical clustering of 80 miRNAs differentially expressed between normal (NAP, LNnorm, and TURPnorm) and cancer (PCa, LN-PCa, and TURP-PCa) samples. Organ confined prostate cancers (PCa) cluster in two major groups designated group I (green text) and group II (red text). Red color, up-regulation; Blue color, down regulation; Yellow color, no change. B) Clinical data is presented as a heat map: (1). Gleason score - yellow 6, orange 7, light red 8, red 9 or 10; (2) clinical stage – yellow T1, orange T2, light red T3, red T4; (3) hormone refractory tumor; (4) PCa-related death; (5) presence of metastases; (6) PSA level progression; (9) ETV1 alteration; (10) TMPRSS2:ERG fusion –yellow no, red yes; (7) % epithelium in tissue; and (8) % cancer in tissue are presented in a color range from 60% to 100%, where darker color indicates higher percentage. Grey indicates no data. C) Unsupervised hierarchical clustering of PCa tumors based on the transcriptional profiles of miRNAs differentially expressed between group I and group II PCa tumors.

Figure 3. Comparison of the performance of miR-classifier between this study (solid line) and the independent validation data-set validation (Schaefer et al., 2010) (dashed line). AUC curves do not differ significantly (P value= 0.2534). Sensitivity, specificity, positive, and negative predicted values for the validation dataset are given in parentheses.
Figure 4. Construction and performance of miR-Predictor. A) Predictions made for the 50 organ confined PCa samples in the data set. Samples are ordered by their posterior probability of membership in either of each class. The upper panel represents a heat map of the relative median centered fold change in log space for each miRNA included in miR-Predictor. Red color, up-regulation; Blue color, down regulation; Yellow color, no change. Genes are arranged by hierarchical clustering. The lower panel represents the clinical data for the corresponding samples. Green color indicates good outcome, belongs to group I. Red color indicated poor outcome, belongs to group II. Blue color indicates that miR-Predictor could not assign the sample to either of both groups based on their posterior probability when a probability cut-off of 0.7 was applied. Gray color indicates no available data. B) ROC curve for miR-Predictor applied to organ confined PCa samples. When sample’s posterior probability was greater than the cut-off of 0.7 sample was considered group I, otherwise it was considered group II. C) Kaplan-Meier curves for the three tumor groups identified by miR-Predictor.

Figure 5. Kaplan-Meier survival curves for PCa-related death based on a) MiR-Predictor, b) pathological grade (Gleason score), c) pathological stage (PT) and preoperative PSA.

Figure 6. Comparison of miRNAs detected by Solexa Deep Sequencing and/or Agilent miRNA microarrays. A) MiRNA detection between both platforms, B) Comparison of the obtained expression data between Solexa deep sequencing and Agilent microarrays. The raw log2 transformed signals of 483 miRNAs are plotted. C) Comparison of the obtained expression data between Solexa deep sequencing and Agilent microarrays. The mean scaled log2 transformed signals of 483 miRNAs are plotted. Yellow indicates fold change less or equal to 2, red indicates fold change between 2 and 4, blue indicates fold change between 4 and 8, green indicates fold change higher than 8. D) Q-RT-PCR expression analysis of selected miRNAs in comparison with microarray and deep sequencing. Data is represented in normalized log2 ratio of expression. MiR-558 was used as a negative control, which was not detected in deep sequencing libraries and was assigned an absent call in more than 50% of the microarrays derived from PCa and LN-PCa. MiR-25 was used as a non-changing control that was stably expressed and detected in both Solexa libraries as well as in all PCa and LN-PCa samples.
Table 1 Clinical parameters of patient cohort for microarray analysis

<table>
<thead>
<tr>
<th></th>
<th>NAP (n=11)</th>
<th>PCa (n=50)</th>
<th>LN-normal (n=3)</th>
<th>LN-PCa (n=12)</th>
<th>TURP-normal (n=4)</th>
<th>TURP-PCa (n=22)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age at treatment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>average (min-max)</td>
<td>62 (54-72)</td>
<td>62 (49-73)</td>
<td>66 (58-69)</td>
<td>65 (39-72)</td>
<td>75.5 (71-78)</td>
<td>73.5 (49-90)</td>
</tr>
<tr>
<td><strong>Follow-up (years)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>average (min-max)</td>
<td>11.0 (2.8-16)</td>
<td>11.3 (1-20.1)</td>
<td>12.2 (4.2-16.6)</td>
<td>10.1 (2.3-24.6)</td>
<td>14.6 (12.1-16.3)</td>
<td>10.3 (0.8-16.6)</td>
</tr>
<tr>
<td><strong>Gleason score</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>0 (0%)</td>
<td>2 (17%)</td>
<td>3 (14%)</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td>10 (20%)</td>
<td>8 (67%)</td>
<td>5 (23%)</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td>7 (14%)</td>
<td>8 (21%)</td>
<td>13 (59%)</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td>1 (2%)</td>
<td>2 (17%)</td>
<td>1 (4.5%)</td>
</tr>
<tr>
<td><strong>Clinical stage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>3 (6%)</td>
<td></td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td></td>
<td>0 (0%)</td>
</tr>
<tr>
<td>T2</td>
<td>12 (24%)</td>
<td></td>
<td>2 (17%)</td>
<td>2 (9%)</td>
<td></td>
<td>2 (9%)</td>
</tr>
<tr>
<td>T3</td>
<td>16 (32%)</td>
<td></td>
<td>1 (8%)</td>
<td>1 (8%)</td>
<td></td>
<td>0 (9%)</td>
</tr>
<tr>
<td>T4</td>
<td>3 (6%)</td>
<td></td>
<td>1 (8%)</td>
<td>1 (8%)</td>
<td></td>
<td>0 (9%)</td>
</tr>
<tr>
<td><strong>Pathological stage</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>16 (332%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>22 (44%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>12 (34%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Preoperative PSA</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=48)</td>
<td>16.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>average (min-max)</td>
<td>(0.3-181.4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PSA progression</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27 (54%)</td>
<td>6 (50%)</td>
<td>7 (32%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hormone refractory status</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hormone refractory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>non hormone refractory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>refractory unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>7 (32%)</td>
<td></td>
<td>4 (18%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11 (50%)</td>
</tr>
<tr>
<td><strong>PCA-related death</strong></td>
<td>1</td>
<td>8 (16%)</td>
<td>1 (33%)</td>
<td>6 (50%)</td>
<td>0</td>
<td>7 (32%)</td>
</tr>
<tr>
<td><strong>Occurrence of metastases</strong></td>
<td>1</td>
<td>10 (20%)</td>
<td>2 (66%)</td>
<td>6 (50%)</td>
<td>0</td>
<td>8 (36%)</td>
</tr>
<tr>
<td><strong>ETV1 alterations</strong></td>
<td>0</td>
<td>3 (6%)</td>
<td>0 (0%)</td>
<td>1 (8%)</td>
<td>0</td>
<td>3 (14%)</td>
</tr>
<tr>
<td><strong>TMPRSS-ERG fusion</strong></td>
<td>0</td>
<td>29 (58%)</td>
<td>2 (66%)</td>
<td>9 (75%)</td>
<td>2</td>
<td>11 (50%)</td>
</tr>
<tr>
<td><strong>% epithelium in tissue</strong></td>
<td>68 (60-80)</td>
<td>77 (60-100)</td>
<td>100</td>
<td>90 (60-100)</td>
<td>73 (60-80)</td>
<td>89 (70-100)</td>
</tr>
<tr>
<td>average (min-max)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>% cancer in tissue</strong></td>
<td>84 (60-100)</td>
<td>100</td>
<td>98</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>average (min-max)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Data are n (%) unless stated otherwise. Abbreviations: NAP, normal adjacent prostate; PCa, organ confined prostate cancer; LN-normal, normal lymph node; LN-PCa, metastatic lymph node; TURP-normal, non-malignant trans-urethral resection of the prostate; TURP-PCa, malignant trans-urethral resection of the prostate; n, number of patients in each group or category; min, minimum value; max, maximum value.

Figure 1

A

<table>
<thead>
<tr>
<th>snoRNA H/A box</th>
<th>22</th>
<th>38</th>
</tr>
</thead>
<tbody>
<tr>
<td>snoRNA C/D box</td>
<td>90</td>
<td>120</td>
</tr>
<tr>
<td>microRNA</td>
<td>290</td>
<td>293</td>
</tr>
<tr>
<td>tRNA</td>
<td>48</td>
<td>47</td>
</tr>
<tr>
<td>mitochondrial tRNA</td>
<td>20</td>
<td>20</td>
</tr>
<tr>
<td>ribosomal RNA</td>
<td>6</td>
<td>5</td>
</tr>
<tr>
<td>mitochondrial r RNA</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>scRNA</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>scaRNA</td>
<td>11</td>
<td>11</td>
</tr>
<tr>
<td>snRNA</td>
<td>13</td>
<td>13</td>
</tr>
<tr>
<td>rRNA</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>miscRNA</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

B

log₂ LN-PCa vs log₂ PCa

- mR-1-1
- mR-1-2
- mR-1-145
- mR-1-143
- mR-221
- mR-222
- mR-225
- mR-455
Figure 3

<table>
<thead>
<tr>
<th>Sample group</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Positive predictive value</th>
<th>Negative predictive value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAP</td>
<td>0.636 (1)</td>
<td>0.94 (0.667)</td>
<td>0.7 (0.75)</td>
<td>0.922 (1)</td>
</tr>
<tr>
<td>PCa</td>
<td>0.94 (0.667)</td>
<td>0.636 (1)</td>
<td>0.922 (1)</td>
<td>0.7 (0.75)</td>
</tr>
</tbody>
</table>
Figure 6

A

Solexa
242 miRs not on chip
483 miRs (~67%)
Agilent
240 miRs (~53% absent)

B

C

D

miR-1
miR-143
miR-145

miR-205
miR-210
miR-222

miR-451
miR-25
miR-558

Relative log2 ratio

PCaNAP
LN/LN

SC/EC