Expression and Function of ETS Genes in Prostate Cancer

Delila Gasi Tandefelt

The studies described in this thesis were performed at the Department of Pathology of the Josephine Nefkens Institute, Erasmus MC, Rotterdam, The Netherlands, and were financially supported by the Marie Curie grant "Cancure" of the European Union.

Cover design: Proefschriftmaken.nl || Uitgeverij BOXPress
Printed & Lay Out by: Proefschriftmaken.nl || Uitgeverij BOXPress
Published by: Uitgeverij BOXPress, 's-Hertogenbosch

The printing of this thesis was financially supported by:
Department of Pathology, Erasmus MC
Erasmus University Rotterdam (EUR)
Stichting Wetenschappelijk Onderzoek Prostaatkanker (SWOP)

Expression and Function of ETS Genes in Prostate Cancer

Expressie en functie van ETS genen in prostaatkanker

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. H.G. Schmidt

and in accordance with the decision of the Doctorate Board.
The public defense shall be held on

Wednesday 6 March 2013 at 11.30 hours

by

Delila Gasi Tandefelt
born in Belgrade, former Yugoslavia

Doctoral Committee:

Promotors:	Prof.dr.ir. J. Trapman
	Prof.dr.ir. G. W. Jenster
Other members:	Prof.dr. C.H. Bangma
	Prof.dr. J.A. Foekens
	Dr. W.N.M. Dinjens

Contents

List of abbreviations

AR	androgen receptor
BCL2A1	BCL2-related protein A1
BRFS	biochemical recurrence-free survival
CDH11	cadherin 11
CR	castration resistant
CRISP3	cysteine-rich secretory protein 3
CRPC	castration resistant prostate cancer
CSS	cancer specific
CV	coefficient of variation
ERG	v-ets erythroblastosis virus E26 oncogene homolog (avian)
ETS	E26 transforming sequence
ETV1	ETS variant gene 1
ETV4	ETS variant gene 4
ETV5	ETS variant gene 5
EST	expressed sequence tag
FDR	false discovery rate
FISH	fluorescent *in situ* hybridization
FKBP10	FK506 binding protein 10
GFP	green fluorescent protein
GP	good prognosis
GPR19	G protein-coupled receptor 19
IHC	immunohistochemistry
INHBA	inhibin ß A
LN	lymph node
MMP1	matrix metallopeptidase 1 interstitial collagenase
NP	normal prostate
OS	overall survival
PAEP	progestagen-associated endometrial protein
PAPPA	pregnancy-associated plasma protein A, pappalysin 1
PBGD	porphobilinogen deaminase
PCR	polymerase chain reaction
PIN	prostatic intraepithelial neoplasia
PLAUR	plasminogen activator, urokinase receptor
PP	poor prognosis

PSA	prostate specific antigen
QPCR	quantitative polymerase chain reaction
RLM-RACE	RNA ligase mediated rapid amplification of cDNA ends
RP	radical prostatectomy
SRGN	serglycin
RT-PCR	reverse transcriptase polymerase chain reaction
TAD	transactivation domain
TMPRSS2	transmembrane protease, serine 2
TURP	transurethral resection of the prostate

CHAPTER 1

Introduction and outline of the thesis

Chapter 1.1

TMPRSS2-ERG and other ETS gene fusions in prostate cancer

Delila Gasi Tandefelt[1], Joost Boormans[2], Karin Hermans[1] and Jan Trapman[1]

1 Department of Pathology, Erasmus University Medical Centre, Rotterdam, The Netherlands 2 Department of Urology, Erasmus University Medical Centre, Rotterdam, The Netherlands

Submitted for publication

Abstract

Prostate cancer is a heterogeneous disease that is very common in elderly men in developed countries. Understanding the molecular and biological processes that contribute to tumor development and progressive growth is a challenging task. The fusion of the genes *ERG* and *TMPRSS2* is the most frequent genomic alteration in prostate cancer. *ERG* is an oncogene that belongs to the family of ETS transcription factors. At lower frequency other members of this gene family are rearranged and overexpressed in prostate cancer. *TMPRSS2* is an androgen-regulated gene that is preferentially expressed in the prostate. Most other ETS fusion partners are similarly regulated and prostate specific. During the last few years novel concepts about the process of gene fusion have emerged and initial experimental results explaining the function of the ETS genes *ERG* and *ETV1* in prostate cancer have been published. In this review we focus on describing most relevant ETS gene fusions and summarize the current knowledge of the role of ETS genes in prostate cancer. Finally, we discuss the clinical relevance of *TMRPSS2-ERG* and other ETS gene fusions in prostate cancer.

Introduction

Prostate cancer is the most frequent malignancy of men in the United States and in other countries with a Western lifestyle [1]. Almost all prostate cancers are adenocarcinomas and it is generally accepted that most prostate cancers develop from a precursor stage denoted as prostate intraepithelial neoplasia (PIN) [2]. Growth patterns of tumors can be very different and heterogeneous, reflected in the so-called Gleason grade [3]. Like other tumors, prostate cancer growth is driven by the accumulation of genetic and epigenetic alterations. One of the earliest molecular alterations in prostate cancer is overexpression of the *ERG* oncogene that occurs in over 50% of prostate cancers [4-6]. ERG overexpression is already detected in tumor adjacent PIN [7-10]. ERG is a member of a large family of ETS transcription factors [11].

Localized prostate cancer can be cured by surgery or radiotherapy, but metastatic disease cannot be cured. It is well established that the growth of prostate cancer depends on male steroid hormones, androgens. Therefore, the treatment of choice of metastasized prostate cancer is one of various types of endocrine therapy, all aiming at the inhibition of the function of the androgen receptor, the intracellular molecular target of androgens [12-14]. The androgen receptor is a member of the family of ligand-dependent nuclear receptor transcription factors. Although almost all prostate tumors show an initial response to endocrine therapy, within 1-3 years essentially all become resistant to the therapy and patients develop Castration Resistant Prostate Cancer (CRPC). Remarkably, in CRPC the androgen receptor still plays a prominent role.

In this review the role of ERG and other ETS transcription factors in prostate cancer is described. Focus is on the mechanism of ETS overexpression and on the clinical relevance of ETS fusions.

Functions of ETS transcription factors

The founding member of the ETS family of transcription factors, v-ets*,* was originally discovered as part of the gag-myb-ets fusion protein of the transforming virus E26 that induces leukemia in chickens. The ETS family is composed of approx. 27 members, that all share high homology in their evolutionally conserved ETS domain in the C-terminal part of the protein (Figure 1) [11, 15, 16]. This 85 amino acid region forms a helix-turn-helix DNA-binding domain that recognizes a GGAA/T core consensus sequence, ETS binding site, in promoter/enhancer regions of target genes. Small differences in binding sites contribute to specificity of ETS binding [11, 17]. A second conserved domain present in a

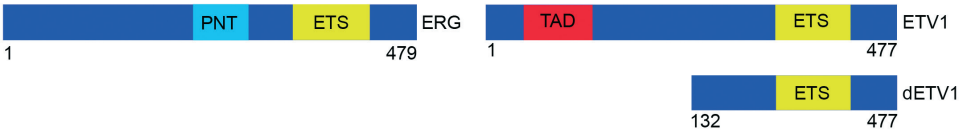


Figure 1:

Schematic presentation of the ETS transcription factors ERG, ETV1 and truncated ETV1. ETS: ETS DNA-binding domain; PNT: pointed protein-protein interaction domain; TAD: acidic transactivation domain.

subset of ETS factors is the pointed domain (PNT). This 65-85 amino acid helix-loop-helix domain functions in protein-protein interactions. Based on their structural composition and the homology in the ETS domain, ETS transcription factors can be separated in approx. 11 subfamilies [11, 15]. Additionally to the ETS- and PNT-domains, activation and repression domains have been postulated in most ETS factors. *ERG* is together with *FLI1* member of the ERG-subfamily and *ETV1*, *ETV4* and *ETV5* are members of PEA3-subfamily that contain an N-terminal acidic transactivation domain (TAD) [11, 15, 18]. ETS proteins can function as transcription activators and repressors. Many directly or indirectly ETS-regulated genes have been defined. Recently, the first Chip-chip and Chip-seq data have been published that identify and compare binding sites for different ETS transcription factors in different cell lines and tissues [11, 17]. A wide variety of overlapping and more specific binding sites have been documented.

ETS proteins are able to play crucial roles in many biological processes, including cellular proliferation, differentiation, apoptosis, tissue remodeling, angiogenesis, metastasis and transformation. Deregulated expression of ETS genes has been described in leukemia's and solid tumors [16]. Moreover, overexpression of ETS, most commonly ETS1, ETS2, ETV1 and ETV4 was found in breast, colon, lung and prostate cancer. In general, overexpression of ETS genes was associated with advanced stage of the disease. As mentioned above, *ERG* is frequently overexpressed in prostate cancer. This overexpression is detected in both early and late stage castration resistant (CR) prostate cancer [4-6].

TMPRSS2 fusion to ERG and other ETS-gene fusions in prostate cancer

ETS genes are frequently involved in gene fusions, resulting in the synthesis of chimerical proteins or in altered expression of the ETS protein. First, ETS fusion genes have been detected in Ewing's sarcoma and in leukemia [11, 19, 20]. Fusion of the 5'-region of the Ewing's sarcoma gene (*EWS*) to the 3'-region of *FLI1* occurs in over 90% of Ewing's

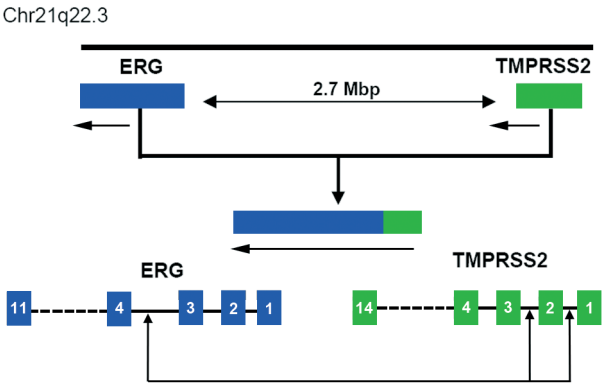


Figure 2:
Schematic presentation of the *TMPRSS2-ERG* fusion on chromosome band 21q22.

sarcoma. This gene fusion leads to the production of a chimerical protein, linking the N-terminal region of EWS to the ETS-domain of FLI1. *EWS-ERG* fusions are detected in approx. 5% of Ewing's sarcoma. In rare cases *EWS* is linked to other ETS genes. The first exons of *EWS* encode a strong transactivation domain. The chimerical protein produced not only modulates the expression of ETS target genes, but probably also induces expression of novel genes. In leukemia many different fusion genes involving the ETS gene *TEL (ETV6)* have been described [11].

In 2005, Petrovics *et al*. observed frequent overexpression of *ERG* in prostate cancer [21]. Later during that year, the mechanism underlying this overexpression, caused by recurrent genomic rearrangement between the first exon(s) of the transmembrane protease, serine 2 (*TMPRSS2*) gene and the *ERG* oncogene was discovered (Figure 2) [4]. This finding was rapidly confirmed and extended, and it is now generally accepted that over half of prostate cancers harbor the *TMPRSS2-ERG* gene fusion as the most frequent genomic alteration in all prostate cancers [22] (Table 1).

TMPRSS2 is an androgen-regulated gene that is preferentially expressed in the prostate [23]. The gene is located on chromosome band 21q22. *ERG* maps also on 21q22 in the same orientation as *TMPRSS2* and at a distance of approx. 3 megabases (Mb) (Figure 2). The fusion of the androgen-regulated and prostate-specific regulating sequences and first exon(s) of *TMPRSS2* to the coding sequences of *ERG* resulted in the androgen-regulated overexpression of *ERG*. Fusion of *TMPRSS2* to *ERG* can occur via two mechanisms: the genomic region between both genes can be lost by interstitial deletion, which is the case in approximately 60% of the fusion-positive tumours, or it can be the result of more complex genomic rearrangements involving chromosome 21q22 and presumably other chromosomes [23, 24]. *TMPRSS2-ERG* fusion is cancer-specific, it is not present in benign

Table 1. ETS gene fusions in prostate cancer

5' fusion partner	Prostate specific	Androgen regulated	ETS partner	freq
TMPRRS2 (chr21q)	+	+	ERG (chr 21q)	~60%
SLC45A3 (chr 1q)	+	+		<1%
NDRG1 (chr 8)	+/-	+		<1%
SLC45A3 (chr 1q)	+	+	FLI1	<1%
TMPRRS2 (chr21q)	+	+	ETV1 (chr 7p)	<1%
SLC45A3 (chr 1q)	+	+		<1%
FOXP1 (chr 3p)	ND	ND		<1%
EST14 (chr 14q)	+	+		<1%
HERVK17 (chr 17p)	+			<1%
HERV-K_22q11.23	+	+		<1%
C15orf21 (chr 15q)	+	+ (down)		<1%
HNRPA2B1 (chr 7p)	-	-		<1%
ACSL3 (chr 2q)	+/-	+		<1%
TMPRRS2 (chr21q)	+	+	ETV4 (chr 17q)	<1%
KLK2 (chr 19p)	+	+		<1%
CANT (chr17q)	+	+		<1%
DDX5 (chr 17q)	-	-		<1%
TMPRRS2 (chr21q)	+	+	ETV5 (chr 3q)	<1%
SLC45A3 (chr 1q)	+	+		<1%
*SLC45A3 (chr 1q)	+	+	ELK4 (chr 1q)	<1%
* read-through transcript				
ND: not determined				

prostatic hyperplasia. *TMPRSS2-ERG* has been detected in approx. 20% of PIN lesions [7, 8]. More recently, ERG overexpression has been detected by immunohistochemistry in a much higher percentage of PIN [9,10]. So, *TMPRSS2-ERG* is an early event in prostate carcinogenesis. It remains to be established whether this fusion plays a role in PIN to

cancer progression or even earlier in prostate cancer development. At low frequency, ERG overexpression is not caused by fusion to *TMPRSS2*, but by fusion to *SLC45A3* or *NDRG1*, also androgen-regulated genes that are preferentially expressed in the prostate (Table 1) [25]. These two fusion partners do not map to 21q22, indicating that chromosomal proximity is not essential for the fusion event.

ETV1 is overexpressed in 5-10% of prostate cancers [4, 26]. *ETV1* gene fusions lead to overexpression of a truncated ETV1 protein, which lacks the TAD domain (Figure 1). In a low percentage of tumors, *ETV4* or *ETV5* is overexpressed due to gene fusion. ETS gene fusions in prostate cancer seem mutually exclusive, although in multifocal disease more than one fusion event can be found. *ERG* is predominantly fused to *TMPRSS2*, but *ETV1, ETV4* and *ETV5* have multiple fusion partners that all are located on different chromosomes (Table 1) (see also refs [26-34]). Remarkably, two of the fusion partners are apparently insignificant retroviral *HERV-K* sequences.

Most *ETV1,4,5* fusion partners are androgen-regulated and prostate-specific expressed (Table 1). There are, however, exceptions. The *ETV1* fusion partner *C15orf21* is down-regulated by androgens and *HNRPA2B1* and *DDX5* are housekeeping genes. Remarkably, expression of *HNRPA2B1* is regulated by a dual-specific CG-rich promoter that cannot be methylated and always maintains an open chromatin structure [35, 36].

ETV1 is not only overexpressed as a fusion gene, but can also be overexpressed as a full length mRNA due to translocation of the complete *ETV1* gene [26, 27, 37]. Some wild type *ETV1* translocations are to a specific chromosome 14 region, but for most wild type *ETV1* translocations the chromosomal region of translocation is not yet known.

In a small percentage of tumors, overexpression of the ETS gene *ELK4* is caused by cis-splicing of the adjacent gene *SLC45A3* [38, 39]. In another proportion of ETS negative samples overexpression of *SPINK1* and more recently, *SPOP* mutation have been identified [40,41]. However, a direct association between ETS genes and *SPINK1* or *SPOP* has not been found.

Mechanism of gene fusion

Probably, the genomic proximity of *TMPRSS2* and *ERG* is an important factor in explaining the high frequency of *TMPRSS2-ERG* fusion as compared to other ETS gene fusions. Although all other ETS gene fusion events are between genes that are on different chromosomes or at a long distance on the same chromosome, it has been postulated that other ETS fusions might also be facilitated by nuclear proximity of the fusion

partners. This is presumed to be accomplished by looping out of genomic regions under certain cell growth conditions, including expression of the fusion partners [42-44].
As described above, most ETS-fusion partners share the properties of androgen up-regulated and prostate-specific expression. So, mechanisms of regulation of expression of the fusion partner seem a second important determinant in the fusion event [42-44]. It might even be that regulation of expression contributes to nuclear proximity. It has also been postulated that binding of an activated androgen receptor to genes encoding fusion partners plays an active role in the fusion process. Evidence has been provided that the androgen receptor induces genomic breaks by recruiting enzymes like topoisomerase II beta (TOP2B), or cytidine deaminase (AID) and ORF2 endonuclease [42, 44-46] . However, the experimental conditions used to investigate the mechanisms of chromosomal proximity and of induced DNA damage are rather complex, and follow-up studies of the data presented thus far are needed.

Biological and molecular functions of ETS proteins in prostate cancer

The main function of ERG and other ETS proteins in prostate cancer is not yet understood. Sole ETS overexpression seems insufficient to induce prostate cancer. *In vitro* studies showed that overexpression of *ERG* or *ETV1* in immortalized, non-tumorigenic epithelial prostate cells increased cell migration and invasion [26, 27, 47-49] and that knockdown of *ERG* or *ETV1* in prostate cancer cell lines slowed invasion [27, 47, 49, 50].
In genetically modified mice (GEMMs), overexpression of *ERG* or *ETV1* resulted in the development of PIN but not of invasive cancer [48, 49]. However, in other studies, *TMPRSS2-ERG* GEMMs did not even develop PIN. Only after crossing with tumor-suppressor *Pten* knockout mice, PIN and micro invasive cancer were observed [51, 52]. The group of Witte provided additional evidence that ERG can cooperate with several different oncogenes or tumor suppressor genes in the development of mouse prostate tumors [53].
In clinical prostate cancer samples gene expression profiles have been used to identify genes and pathways associated with *ERG* overexpression. Iljin *et al*. pointed to the importance of the WNT pathway in prostate cancer and high expression of *HDAC1* in *ERG* overexpressing tumor samples [54]. Also, activation of the TGFß pathway has been associated with *ERG* overexpression [55]. Although reported data are variable, a dozen genes are in several studies associated with ERG overexpression, including *HDAC1, CACNA1D, TDRD1* and *NCALD* [54-56]. It remains to be investigated whether genes

associated with *ERG* overexpression are direct ERG targets or that their expression is indirectly regulated by ERG.

Recently, important initial knowledge of the molecular effect of ERG overexpression in prostate cancer has been published [57]. It has been shown by whole genome ChipSeq analysis that there is overlap between genomic regions that bind the androgen receptor and that bind ERG. As a result ERG overexpression could interfere with and modify the expression of androgen receptor regulated genes. A model has been proposed that ERG overexpression inhibits androgen receptor regulated differentiation and stimulates dedifferentiation mediated by the H3K27 methyltransferase polycomb gene *EZH2* [57]. Further experiments on androgen receptor and ERG association are of high importance.

Our knowledge of the biological and molecular effects of *ETV1* overexpression in prostate cancer is more limited. *ETV1* can be overexpressed as a truncated protein lacking the N-terminal TAD domain (dETV1) due to gene fusion (Figure 1), or as a full-length protein due to translocation of the complete gene to a different genomic region [26, 35]. In *in vitro* studies, full length ETV1 is a strong transcriptional activator, but dETV1 is much less active. Although both gene-variants similarly induce migration and invasion in non-tumor prostate cells, only the full-length protein seems capable of inducing anchorage-independent growth. Although MMPs and the UPA/UPAR system are known as *ETV1* associated genes, a clear global picture of ETV1-regulated gene expression is lacking [26, 58].

Heterogeneity of prostate cancer

ETS gene fusions in prostate cancer are mutually exclusive. Because localized prostate cancer is often a multifocal disease, tumours have been tested for ETS gene fusions in different cancerous foci within one prostate. In approx. half of the cases, individual tumour foci differed according to the presence of ETS rearrangements or fusion mechanism (deletion or translocation) [59, 60]. Because ERG gene fusion is an early event, this finding most likely implies that the majority of men develop multiple cancers in their prostate. Metastatic prostate cancer foci in one individual, however, displayed identical ETS rearrangement showing that only one tumor focus seeded metastatic deposits [61-63].

ETS-fusions as diagnostic and prognostic marker of prostate cancer

Because *ERG* fusing transcripts are present in approx. 50% of prostate tumours, it is clear that the presence of *ERG* fusion transcripts in prostate tissue or in urine or overexpression of ERG protein in prostate biopsies is an important robust diagnostic marker of prostate cancer in a large subgroup of patients. The prognostic significance of *TMPRSS2-ERG* gene fusion is controversial and contradicting results have been reported (Table 2). The discrepancies might be a result of the differences in the populations studied, the techniques used to detect gene fusions and the effect of treatment on the examined tumor samples. Originally, Petrovics *et al.* found that patients with high expression levels of *ERG* showed a slower progression than patients with tumors without *ERG* overexpression [21]. After the discovery of recurrent gene fusions in prostate cancer *TMPRSS2-ERG* was more frequently found correlated with poor clinical outcome (Table 2) [65, 69, 71-73]. However, this was not confirmed in later studies [66-68, 70]. It also was suggested that a subgroup of patients who had gene fusion with an interstitial deletion between *TMPRSS2* and *ERG* (so called “class Edel”) had poorer clinical outcome than gene fusion-negative patients or than patients with *TMPRSS2-ERG* gene fusion without loss of the genomic region between the two genes [65]. Alternatively, it is not impossible that the poor survival associated with copy number increase of chromosome 21 reflected generalized aneuploidy and genomic instability. We showed that patients expressing *TMPRSS2-ERG* fusion transcripts starting at an alternative first exon had better outcome after radical prostatectomy than patients carrying tumors that only expressed *TMPRSS2(exon1)-ERG* [69].

ERG positive tumors in CRPC patients who were treated with the novel FDA-approved inhibitor of testosterone synthesis abiraterone acetate were more frequent in patients who responded well to the therapy than in patients who did not show a good response [74]. These results suggest that ETS gene fusions remain dependent on androgen signalling, despite the castration-resistant stage of the disease. Whether gene fusion status of the tumour has implications for the timing and the choice of endocrine therapy remain to be clarified. In ETS negative samples *SPINK1* overexpression was an independent predictor of clinical progression [40].

Table 2. Studies on clinical relevance of *TMPRSS2-ERG* fusion gene in prostate cancer

A. Cancer-specific (CSS) and overall survival (OS)						
Reference	**N**	**Tissue**	**Technique**	***TMPRSS2-ERG***	**Follow-up (median)**	**Fusion-positive versus fusion-negative cases**
Gopalan *et al.* [67]	521	RP	FISH	42%	7.9 yr	No difference in OS (univariate level)
Fitzgerald *et al.* [66]	214	RP / TURP	FISH	35.5%	12.3 yr	No difference in CSS (multivariate level)
Attard *et al.* [65]	445	TURP / biopsies	FISH	30%	7.5 yr	Poorer CSS and OS (multivariate level)
Demichelis *et al.* [64]	111	TURP / Millin	FISH	15%	9.1 yr	Poorer CSS (univariate level)
B. Biochemical recurrence-free survival (BRFS)						
Petrovics *et al.* [21]	114	RP	Quantitative RT-PCR	62%*	n.s.	Longer BRFS (univariate level)
Saramäki *et al.* [70]	150	RP	FISH	33%	5.5 yr	Longer BRFS (multivariate level)
Nam *et al.* [71]	165	RP	RT-PCR	42%	1.7 yr	Poorer BRFS (multivariate level)
Perner *et al.* [72]	118	RP	FISH	49%	n.s.	**
Wang *et al.* [73]	59	RP	RT-PCR	59%	n.s.	***
Lapointe *et al.* [68]	63	RP LN	RT-PCR	70%	2.0 yr	No difference in BRFS
Hermans *et al.* [69]	67	RP	Quantitative RT-PCR	66%	10.2 yr	No difference in BRFS ****

***** *ERG* overexpression

****** Higher recurrence rate, no survival analysis

******* More early recurrences**,** no survival analysis

******** Longer BRFS for *TMPRSS2(exon0)-ERG* (multivariate level)

Concluding remarks

The finding of ETS gene fusions in prostate cancer has been a major factor in increasing our knowledge of the molecular and biological mechanisms of development and progressive growth of the disease. The postulated mechanisms of gene fusion and molecular function of ETS genes are of high general interest. Further exploration of proposed mechanisms will contribute to understanding the processes of genomic rearrangements in general. The gene fusions are also of utmost importance for clinical prostate cancer. At the moment, ERG overexpression is already instrumental in diagnosis

of the disease. Moreover, elucidation of the mechanisms of ETS gene expression and function increase the opportunity to find new therapeutic targets for early and late stage prostate cancer.

References

1. Siegal R, Naishadham D & Jemal D (2012) Cancer statistics, 2012. *CA Cancer J Clin* **62**, 10-29.
2. DeMarzo AM, Nelson WG, Isaacs WB & Epstein JI (2003) Pathological and molecular aspects of prostate cancer. *The Lancet* **361**, 955-964.
3. Lotan TL & Epstein JI (2010) Clinical implications of changing definitions within the Gleason grading system. *Nat Rev Urol* **7**, 136-142.
4. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW, Varambally S, Cao X, Tchinda J, Kuefer R, Lee C, Montie JE, Shah RB, Pienta KJ, Rubin MA & Chinnaiyan AM (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* **310**, 644-648.
5. Soller MJ, Isaksson M, Elfving P, Soller W, Lundgren R & Panagopoulos I (2006) Confirmation of the high frequency of the TMPRSS2/ERG fusion gene in prostate cancer. *Genes Chromosomes Cancer* **45**, 717-719.
6. Hermans KG, Boormans JL, Gasi D, van Leenders GJ, Jenster G, Verhagen PC & Trapman J (2009) Overexpression of prostate-specific TMPRSS2(exon 0)-ERG fusion transcripts corresponds with favorable prognosis of prostate cancer. *Clin Cancer Res* **15**, 6398-6403.
7. Cerveira N, Ribeiro FR, Peixoto A, Costa V, Henrique R, Jeronimo C & Teixeira MR (2006) TMPRSS2-ERG gene fusion causing ERG overexpression precedes chromosome copy number changes in prostate carcinomas and paired HGPIN lesions. *Neoplasia* **8**, 826-832.
8. Mosquera JM, Perner S, Genega EM, Sanda M, Hofer MD, Mertz KD, Paris PL, Simko J, Bismar TA, Ayala G, Shah RB, Loda M & Rubin MA (2008) Characterization of TMPRSS2-ERG fusion in high-grade prostatic intraepithelial neoplasia and potential clinical implications. *Clin Cancer Res* **14**, 3380-3385.
9. Park K, Tomlins SA, Mudaliar KM, Chiu YL, Esqueva R, Mehra R, Suleman K, Varembally S, Brenner JC, MacDonald T, Srivasta A, Tewari A, Sathyanarayana U, Nagy D, Pestano G, Kunju LP, Demichelis F, Chinnaiyan AM & Rubin MA (2010) Antibody-based detection of ERG rearrangement-positive prostate cancer. *Neoplasia* **12**, 590-598.
10. Van Leenders GJ, Boormans JL, Vissers CJ, Hoogland AM, Bressers AA, Furusato B & Trapman J (2011) Antibody EPR3864 is specific for ERG genomic fusions in prostate cancer: implications for pathological practice. *Mod Pathol* **24**, 1128-1138.
11. Hollenhorst PC McIntosh LP & Grave BJ (2011) Genomic and biochemical insights into the specificity of ETS transcription factors. *Annu Rev Biochem* 8**0**, 437-471.
12. Feldman BJ & Feldman D (2001) The development of androgen-independent prostate cancer. *Nature Rev Cancer* **1**, 34-45.
13. Scher HI & Sawyers CL (2005) Biology of progressive, castration-resistant prostate cancer: Directed therapies targeting the androgen receptor signalling axis *J Clin Oncol* **23**, 8253-8261.
14. Lonergan PE & Tindall DJ (2011) Androgen receptor signalling in prostate cancer development and progression. *J Carcinog* **10**, 20.
15. Oikinawa T & Yamada T (2003) Molecular biology of the Ets family of transcription factors. *Gene* **303**, 11-34.
16. Seth A. and Watson DK (2005) ETS transcription factors and their emerging roles in human cancer. *Eur J Cancer* **41**, 2462-2478.

17. Wei G-H, Badis G, Berger MF, Kivioja T, Palin K, Enge M, Bonke M, Jolma A, Varjosalo M, Gehrke AR, Yan J, Talukder S, Turunen M, Taipale M, Stunnenberg HG, Ukkonen E, Hughes TR, Bulyk ML & Taipale J (2010) Genome-wide analysis of ETS-family DNA-binding in vitro and in vivo. *EMBO J* **29**, 2147-2160.
18. Oh S, Shin S & Janknecht R (2012) ETV1, 4 and 5: An oncogenic subfamily of ETS transcription factors. *Biochim Biophys Acta* **1826**, 1-12.
19. Khoury JD (2005) Ewing sarcoma family of tumors. *Adv Anat Pathol* **12**, 212-220.
20. Bohlander SK (2005) ETV6: a versatile player in leukemogenesis. *Semin cancer biol* **15**, 162-174.
21. Petrovics G, Liu A, Shaheduzzaman S, Furusato B, Sun C, Chen Y, Nau M, Ravindranath L, Chen Y, Dobi A, Srikantan V, Sesterhenn IA, McLeod DG, Vahey M, Moul JW & Srivastava S (2005) Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome. *Oncogene* **24**, 3847-3852.
22. Kumar-Sinha C, Tomlins SA & Chinnaiyan AM (2008) Recurrent gene fusions in prostate cancer. *Nat Rev Cancer* **8**, 497-511.
23. Hermans KG, van Marion R, van Dekken H, Jenster G, van Weerden WM & Trapman J (2006) TMPRSS2:ERG fusion by translocation or interstitial deletion is highly relevant in androgen-dependent prostate cancer, but is bypassed in late-stage androgen receptor-negative prostate cancer. *Cancer Res* **66**, 10658-10663.
24. Perner S, Demichelis F, Beroukhim R, Schmidt FH, Mosquera JM, Setlur S, Tchinda J, Tomlins SA, Hofer MD, Pienta KG, Kuefer R, Vessella R, Sun XW, Meyerson M, Lee C, Sellers WR, Chinnaiyan AM & Rubin MA. (2006) TMPRSS2:ERG fusion-associated deletions provide insight into the heterogeneity of prostate cancer. *Cancer Res* **66**, 8337-8341.
25. Esgueva R, Perner S, J LaFargue C, Scheble V, Stephan C, Lein M, Fritzsche FR, Dietel M, Kristiansen G & Rubin MA (2010) Prevalence of TMPRSS2-ERG and SLC45A3-ERG gene fusions in a large prostatectomy cohort. *Mod Pathol* **23**, 539-546.
26. Hermans KG, van der Korput HA, van Marion R, van de Wijngaart DJ, Ziel-van der Made A, Dits NF, Boormans JL, van der Kwast TH, van Dekken H, Bangma CH, Korsten H, Kraaij R, Jenster G & Trapman J (2008) Truncated ETV1, fused to novel tissue-specific genes, and full-length ETV1 in prostate cancer. *Cancer Res* **68**, 7541-7549.
27. Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS, Menon A, Jing X, Cao Q, Han B, Yu J, Wang L, Montie JE, Rubin MA, Pienta KJ, Roulston D, Shah RB, Varambally S, Mehra R & Chinnaiyan AM (2007) Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature* **448**, 595-599.
28. Tomlins SA, Mehra R, Rhodes DR, Smith LR, Roulston D, Helgeson BE, Cao X, Wei JT, Rubin MA, Shah RB & Chinnaiyan AM (2006) TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. *Cancer Res* **66**, 3396-3400.
29. Helgeson BE, Tomlins SA, Shah N, Laxman B, Cao Q, Prensner JR, Cao X, Singla N, Montie JE, Varambally S, Mehra R & Chinnaiyan AM (2008) Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. *Cancer Res* **68**, 73-80.
30. Attard G, Clark J, Ambroisine L, Mills IG, Fisher G, Flohr P, Reid A, Edwards S, Kovacs G, Berney D, Foster C, Massie CE, Fletcher A, De Bono JS, Scardino P, Cuzick J & Cooper CS (2008) Heterogeneity and clinical significance of ETV1 translocations in human prostate cancer. *Br J Cancer* **99**, 314-320.
31. Han B, Mehra R, Dhanasekaran SM, Yu J, Menon A, Lonigro RJ, Wang X, Gong Y, Wang L, Shankar S, Laxman B, Shah RB, Varambally S, Palanisamy N, Tomlins SA, Kumar-Sinha C & Chinnaiyan AM (2008) A

fluorescence in situ hybridization screen for E26 transformation-specific aberrations: identification of DDX5-ETV4 fusion protein in prostate cancer. *Cancer Res* **68**, 7629-7637.

32. Hermans KG, Bressers AA, van der Korput HA, Dits NF, Jenster G & Trapman J (2008) Two unique novel prostate-specific and androgen-regulated fusion partners of ETV4 in prostate cancer. *Cancer Res* **68**, 3094-3098.
33. Clark JP & Cooper CS (2009) ETS gene fusions in prostate cancer. Nat Rev Urol **6**, 429-439.
34. Rubin MA, Maher CA & Chinnaiyan AM (2011) Common gene rearrangements in prostate cancer. *J Clin Oncol* **27**, 3659-3668.
35. Antoniou M, Harland L, Mustoe T, Williams S, Holdstock J, Yague E, Mulcahy T, Griffiths M, Edwards S, Ioannou PA, Mountain A & Crombie R (2003) Transgenes encompassing dual-promotor CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing. *Genomics* **82**, 269-279.
36. Lindahl-Allen M & Antoniou M (2007) Correlation of DNA Methylation with Histone Modifications across the HNRPA2B1-CBX3 Ubiquitously-Acting Chromatin Open Element (UCOE). *Epigenetics* **2**, 227-236.
37. Gasi D, van der Korput HA, Douben HC, de Klein A, de Ridder CM, van Weerden WM & Trapman J (2011) Overexpression of full-length ETV1 transcripts in clinical prostate cancer due to gene translocation. *PLoS One* **26**, 6:e16332.
38. Rickman DS, Pflueger D, Moss B, VanDoren VE, Chen CX, de la Taille A, Kuefer R, Tewari AK, Setlur SR, Demichelis F& Rubin MA (2009) SLC45A3-ELK4 is a novel and frequent erythroblast transformation-specific fusion transcript in prostate cancer. *Cancer Res* **69**, 2734-2738.
39. Zhang Y, Gong M, Yuan H, Park HG, Frierson HF & Li H (2012) Chimeric transcript generated by cis-splicing of adjacent genes regulates prostate cancer cell proliferation. *Cancer Discov* **2**, 598-607.
40. Tomlins SA, Rhodes DR, Yu J, Varambally S, Mehra R, Perner S, Demichelis F, Helgeson BE, Laxman B, Morris DS, Cao Q, Cao X, Andrén O, Fall K, Johnson L, Wei JT, Shah RB, Al-Ahmadie H, Eastham JA, Eggener SE, Fine SW, Hotakainen K, Stenman UH, Tsodikov A, Gerald WL, Lilja H, Reuter VE, Kantoff PW, Scardino PT, Rubin MA, Bjartell AS & Chinnaiyan AM (2008) The role of SPINK1 in ETS rearrangement negative prostate cancers. *Cancer Cell* **13**, 519-528.
41. Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blater M, Theurillat J-P, White TA,Stojanov P, Van Allen E, Stransky N *et al* (2012) Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. *Nat Genet* **44**, 685-689.
42. Lin C, Yang L, Tanasa B, Hutt K, Ju B-G, Ohgi KA, Zang J, Rose DW, Fu X-D, Glass CK & Rosenfeld MG (2009) Nuclear receptor-induced chromosomal proximity and DNA breaks underlie specific translcations in cancer. *Cell* **139**, 1069-1083.
43. Mani R-S, Tomlins SA, Callahan K, Ghosh A, Nyati MK, Varambally S, Palanisamy N & Cinnaiyan AM (2009) Induced chromosomal proximity and gene fusions in prostate cancer *Science* **326**, 1230.
44. Lin C, Yang L & Rosenfeld MG (2012) Molecular logic underlying chromosomal translocations, random or non-random? *Adv Cancer Res* **113**, 241-279.
45. Haffner MC, Aryee MJ, Toubaji A, Esopi DM, Albadine R, Gurel B, Isaacs WB, Bova GS, Liu J, Meeker AK, Netto G, DeMarzo AM, Nelson WG & Yegnasubramanian S (2010) Androgen-induced TOP2B-mediated double-strand breaks and prostate cancer gene rearrangements *Nat Genet* **42**, 668-675.

46. Haffner MC, De Marzo AM, Meeker AK, Nelson WG & Yegnasubramanian S (2011) Transcription-induced DNA double strand breaks: Both oncogenic force and potential therapeutic target? *Clin Cancer Res* **17**, 3858-3864.
47. Wang J, Cai Y, Yu W, Ren C, Spencer DM & Ittmann M (2008) Pleiotropic biological activities of alternatively spliced TMPRSS2/ERG fusion gene transcripts. *Cancer Res* **68**, 8516-8524.
48. Klezovitch O, Risk M, Coleman I, Lucas JM, Null M, True LD, Nelson PS & Vasioukhin V (2008) A causal role for ERG in neoplastic transformation of prostate epithelium. Proc *Natl Acad Sci USA* **105**, 2105-2110.
49. Tomlins SA, Laxman B, Varambally S, Cao X, Yu J, Helgeson BE, Cao Q, Prensner JR, Rubin MA, Shah RB, Mehra R & Chinnaiyan AM (2008) Role of the TMPRSS2-ERG gene fusion in prostate cancer. *Neoplasia* **10**, 177-188.
50. Sun C, Dobi A, Mohamed A, Li H, Thangapazham RL, Furusato B, Shaheduzzaman S, Tan SH, Vaidyanathan G, Whitman E, Hawksworth DJ, Chen Y, Nau M, Patel V, Vahey M, Gutkind JS, Sreenath T, Petrovics G, Sesterhenn IA, McLeod DG & Srivastava S (2008) TMPRSS2-ERG fusion, a common genomic alteration in prostate cancer activates C-MYC and abrogates prostate epithelial differentiation. *Oncogene* **27**, 5348-5353.
51. Carver BS, Tran J, Gopalan A, Chen Z, Shaikh S, Carracedo A, Alimonti A, Nardella C, Varmeh S, Scardino PT, Cordon-Cardo C, Gerald W % Pandolfi PP (2009) Aberrant ERG expression cooperates with loss of PTEN to promote cancer progression in the prostate. *Nat Genet* **41**, 619-624.
52. King JC, Xu J, Wongvipat J, Hieronymus H, Carver BS, Leung DH, Taylor BS, Sander C, Cardiff RD, Couto SS, Gerald WL & Sawyers CL (2009) Cooperativity of TMPRSS2-ERG with PI3-kinase pathway activation in prostate oncogenesis. *Nat Genet* **41**, 524-526.
53. Zong Y Goldstein AS, Lawson DA, Teitell MA & Witte ON (2009) ETS family transcription factors collaborate with alternative signaling pathways to induced carcinoma from adult murine prostate cells. *Proc Natl Acad Sci USA* **106**, 12465-12470.
54. Iljin K, Wolf M, Edgren H, Gupta S, Kilpinen S, Skotheim RI, Peltola M, Smit F, Verhaegh G, Schalken J, Nees M, & Kallioniemi O (2006) TMPRSS2 fusions with oncogenic ETS factors in prostate cancer involve unbalanced genomic rearrangements and are associated with HDAC1 and epigenetic reprogramming *Cancer Res* **66**, 10242-10246.
55. Brase JC, Johannes M, Mannsperger H, Fälth M, Metzger J, Kacprzyk LA, Andrasiuk T, Gade S, Meister M, Sirma H, Sauter G, Simon R, Schlomm T, Beissbarth T, Korf U, Kuner R & Sültmann H (2011) TMPRSS2-ERG -specific transcriptional modulation is associated with prostate cancer biomarkers and TGF-β signaling. *BMC Cancer* **11**, 507.
56. Jhavar S, Reid A, Clark J, Kote-Jarai Z, Christmas T, Thompson A, Woodhouse C, Ogden C, Fisher C, Corbishley C, De-Bono J, Eeles R, Brewer D & Cooper C (2008) Association of TMPRSS2-ERG gene fusion in human prostate cancer by expression profiling using GeneChip Human Exon 1.0 ST arrays. *J Mol Diagn* **10**, 50-57.
57. Yu J, Yu J, Mani R-S, Cao Q, Brenner CJ, Cao X, Wang X, Wu L, Li J, Hu M, Gong Y, Cheng H, Laxman B, Vellaichamy A, Shankar S, Li Y *et al* (2010) An integrated netwerk of androgen receptor, polycomp, and TMPRSS2-ERG gene fusions in prostate cancer progression *Cancer Cell* **17**, 443-454.
58. de Launoit Y, Baert JL, Chotteau-Lelievre A, Monte D, Coutte L, Mauen S, Firlej V, Degerny C & Verreman K (2006) The Ets transcription factors of the PEA3 group: Transcriptional regulators in metastasis. *Biochim. Biophys Acta* **1766**, 79-87.

59. Barry M, Perner S, Demichelis F & Rubin MA (2007) TMPRSS2-ERG fusion heterogeneity in multifocal prostate cancer: clinical and biologic implications. *Urology* **70**, 630-633.

60. Mehra R, Han B, Tomlins SA, Wang L, Menon A, Wasco MJ, Shen R, Montie JE, Chinnaiyan AM & Shah RB (2007) Heterogeneity of TMPRSS2 gene rearrangements in multifocal prostate adenocarcinoma: molecular evidence for an independent group of diseases. *Cancer Res* **67**, 7991-7995.

61. Mehra R, Tomlins SA, Yu J, Cao X, Wang L, Menon A, Rubin MA, Pienta KJ, Shah RB & Chinnaiyan AM (2008) Characterization of TMPRSS2-ETS Gene Aberrations in Androgen-Independent Metastatic Prostate Cancer. *Cancer Res* **68**, 3584-3590.

62. Liu W, Laitinen S, Khan S, Vihinen M, Kowalski J, Yu G, Chen L, Ewing CM, Eisenberger MA, Carducci MA, Nelson WG, Yegnasubramanian S, Luo J, Wang Y, Xu J, Isaacs WB, Visakorpi T & Bova GS (2009) Copy number analysis indicates monoclonal origin of Lethal Metastatic prostate cancer. *Nat Med* **15**, 559-565.

63. Guo C.C., Wang Y., Xiao L., Troncoso P & Czerniak B (2012) The relationship of TMPRSS2-ERG gene fusion between primary and metastatic prostate cancers. *Human Pathol* **43**, 644-649.

64. Demichelis F, Fall K, Perner S, Andrén O, Schmidt F, Setlur SR, Hoshida Y, Mosquera JM, Pawitan Y, Lee C, Adami HO, Mucci LA, Kantoff PW, Andersson SO, Chinnaiyan AM, Johansson JE & Rubin MA (2007) TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. *Oncogene* **26**, 4596-4599.

65. Attard G, Clark J, Ambroisine L, Fisher G, Kovacs G, Flohr P, Berney D, Foster CS, Fletcher A, Gerald WL, Moller H, Reuter V, De Bono JS, Scardino P, Cuzick J & Cooper CS (2008) Duplication of the fusion of TMPRSS2 to ERG sequences identifies fatal human prostate cancer. *Oncogene* **27**, 253-263.

66. FitzGerald LM, Agalliu I, Johnson K, Miller MA, Kwon EM, Hurtado-Coll A, Fazli L, Rajput AB, Gleave ME, Cox ME, Ostrander EA, Stanford JL & Huntsman DG (2008) Association of TMPRSS2-ERG gene fusion with clinical characteristics and outcomes: Results from a population-based study of prostate cancer. *BMC Cancer* **8**, 230.

67. Gopalan A, Leversha MA, Satagopan JM, Zhou Q, Al-Ahmadie HA, Fine SW, Eastham JA, Scardino PT, Scher HI, Tickoo SK, Reuter VE & Gerald WL (2009) TMPRSS2-ERG gene fusion is not associated with outcome in patients treated by prostatectomy. *Cancer Res* **69**, 1400-1406.

68. Lapointe J, Kim YH, Miller MA, Li C, Kaygusuz G, van de Rijn M, Huntsman DG, Brooks JD & Pollack JR (2007) A variant TMPRSS2 isoform and ERG fusion product in prostate cancer with implications for molecular diagnosis. *Mod Pathol* **20**, 467-73.

69. Hermans KG, Boormans JL, Gasi D, van Leenders GJ, Jenster G, Verhagen PC & Trapman J (2009) Overexpression of prostate-specific TMPRSS2(exon 0)-ERG fusion transcripts corresponds with favorable prognosis of prostate cancer. *Clin Cancer Res* **15**, 6398-6403.

70. Saramaki OR, Harjula AE, Martikainen PM, Vessella RL, Tammela TL & Visakorpi T (2008) TMPRSS2-ERG fusion identified a subgroup of prostate cancers with a favorable prognosis. *Clin Cancer Res* **14**, 3395-4000.

71. Nam RK, Sugar L, Yang W, Srivastava S, Klotz LH, Yang LY, Staminirovic A, Encioiu E, Neill M, Loblaw DA, Trachtenberg J, Narod SA & Set A (2007) Expression of the TMPRSS2:ERG fusion gene predicts cancer recurrence after surgery for localised prostate cancer. *Br J Cancer* **97**, 1690-1695.

72. Perner S, Demichelis F, Beroukhim R, Schmidt FH, Mosquera JM, Setlur S, Tchinda J, Tomlins SA, Hofer MD, Pienta KG, Kuefer R, Vessella R, Sun XW, Meyerson M, Lee C, Sellers WR, Chinnaiyan AM & Rubin

MA (2006) TMPRSS2-ERG fusion-associated deletions provide insight into the hetrogeneity of prostate cancer. *Cancer Res* **66**, 8337-8341.

73. Wang J, Cai Y., Ren C & Ittman M (2006) Expression of variant TMPRSS2-ERG mRNAs is associated with aggressive prostate cancer. *Cancer Res* **66**, 8347-8351.

74. Attard G, Swennenhuis JF, Olmos D, Reid AH, Vickers E, A'Hern R, Levink R, Coumans F, Moreira J, Riisnaes R, Oommen NB, Hawche G, Jameson C, Thompson E, Sipkema R, Carden CP, Parker C, Dearnaley D, Kaye SB, Cooper CS, Molina A, Cox ME, Terstappen LW, & de Bono JS (2009) Characterization of ERG, AR and PTEN gene status in circulating tumor cells from patients with castration-resistant prostate cancer. *Cancer Res* **69**, 2912-2918.

Chapter 1.2

Outline of the thesis

Outline of the thesis

The aim of this thesis is to investigate the function and expression of the most commonly overexpressed ETS-transcription factors in prostate cancer, ERG and ETV1. In chapters 2 and 3 of the thesis the focus is on ETV1, in chapters 4, 5 and 6 on ERG.

ETV1 is overexpressed as fusion gene (dETV1) or as full length gene in 5-10% of prostate cancers. In **Chapter 2** clinical prostate cancer samples with *ETV1* overexpression are studied. Gene fusions involving *ETV1* are identified by 5'RACE of mRNA combined with sequencing. FISH and multicolor FISH are used to identify translocations of the full-length gene. Moreover, novel ETV1 isoforms are identified and their function characterized in anchorage independent growth assays. In **Chapter 3** genes associated with upregulation of ETV1 and dETV1 are identified in expression arrays of PNT2C2-cells that overexpress *ETV1* or *dETV1*. The regulation of a selection of identified genes is studied in promoter assays. Furthermore the effect of different growth conditions on the gene expression and composition are studied in the ETV1 or dETV1 overexpressing PNT2C2 cells grown *in vitro* in 2D and as *in vivo* transplanted tumors. Data obtained are compared to genes expressed in clinical prostate cancer overexpressing *ETV1*.

TMPRSS2-ERG is the most common fusion gene in prostate cancer. The transcription of the fusion gene can start at two alternative first exons of *TMPRSS2*, here defined exon 0 and exon 1. In **Chapter 4** the expression of *TMPRSS2-ERG* starting at these two alternative exons is studied in prostate cancer cell lines, xenografts and clinical prostate cancer samples. Furthermore, the expression of the transcripts is correlated with progression of primary prostate cancers. **Chapter 5** describes the methods used to identify novel gene fusions involving the ETS genes and the approach to characterize the properties of their fusion partners. Gene expression profiles of primary prostate cancer samples from patients with a long clinical follow-up are investigated in **Chapter 6.** ERG positive and ERG negative subgroups are identified by unsupervised hierarchical clustering and correlated with clinical progression. Expression data from the subgroups are used to define a 36-gene classifier for poor prognosis. The classifier is validated in an independent patient cohort and selected genes are tested for biological properties. The results of this thesis are discussed in **Chapter 7** together with suggestions of continuation of this work.

CHAPTER 2

Overexpression of full-length *ETV1* transcripts in clinical prostate cancer due to gene translocation

Delila Gasi[1], Hetty A. van der Korput[1], Hannie C. Douben[2], Annelies de Klein[2], Corrina M. de Ridder[3], Wytske M. van Weerden[3] and Jan Trapman[1]

Departments of Pathology[1], Clinical Genetics[2] and Urology[3], Erasmus University Medical Centre, Rotterdam, The Netherlands

PLoS ONE 2011;6:e16332

Abstract

ETV1 is overexpressed in a subset of clinical prostate cancers as a fusion transcript with many different partners. However, *ETV1* can also be overexpressed as a full-length transcript. Full-length ETV1 protein functions different from truncated ETV1 produced by fusion genes. In this study we describe the genetic background of full-length *ETV1* overexpression and the biological properties of different full-length ETV1 isoforms in prostate cancer. Break-apart FISH showed in five out of six patient samples with overexpression of full-length *ETV1* a genomic rearrangement of the gene, indicating frequent translocation. We were able to study the rearrangements in more detail in two tumors. In the first tumor 5'-RACE on cDNA showed linkage of the complete *ETV1* transcript to the first exon of a prostate-specific two exon ncRNA gene that maps on chromosome 14 (*EST14*). This resulted in the expression of both full-length *ETV1* transcripts and *EST14-ETV1* fusion transcripts. In chromosome spreads of a xenograft derived from the second prostate cancer we observed a complex *ETV1* translocation involving a chromosome 7 fragment that harbors *ETV1* and fragments of chromosomes 4 and 10. Further studies revealed the overexpression of several different full-length transcripts, giving rise to four protein isoforms with different N-terminal regions. Even the shortest isoform synthesized by full-length *ETV1* stimulated *in vitro* anchorage-independent growth of PNT2C2 prostate cells. This contrasts the lack of activity of even shorter N-truncated ETV1 produced by fusion transcripts. Our finding that in clinical prostate cancer overexpression of full-length *ETV1* is due to genomic rearrangements involving different chromosomes and the identification of a shortened biologically active ETV1 isoform are highly relevant for understanding the mechanism of ETV1 function in prostate cancer.

Introduction

Gene fusions are important in the development of many hematological malignancies and sarcomas, but are rare in most other tumor types (1). However, the frequent fusion between *TMPRSS2* and the ETS gene *ERG* showed that gene fusion is a highly relevant event in prostate cancer (2,3). Overexpression of the *TMPRSS2-ERG* fusion gene has been reported in 40-70% of prostate cancer cases (2-5). Fusions of *TMPRSS2* and three other genes encoding ETS transcription factors, *ETV1, ETV4* and *ETV5*, which are located on different chromosomes, occur at low frequency in prostate cancer (2,6,7). However, for *ETV1* at least 10 different fusion partners have been described (8-10). Most of these have in common that, like *TMPRSS2,* they are prostate-specific and androgen-regulated expressed. The properties of fusion partners are key elements in explaining the androgen-regulated overexpression of an ETS oncogene in prostate cancer. However, a unique characteristic of *ETV1* is that it can not only be overexpressed in prostate cancer as a fusion transcript but also as a full-length wild-type transcript.

The large family of ETS transcription factors is composed of 27 members (11-13). All members have in common a highly homologous DNA binding domain, the ETS domain. The remaining regions of most ETS proteins show limited structural homology. ETV1, ETV4 and ETV5 are the members of a small subfamily of structurally related ETS proteins. These proteins contain in the N-terminal region a conserved short acidic transactivation domain (TAD) that is absent in ERG. ETS proteins regulate many target genes that modulate biological processes like cell growth, angiogenesis, migration, proliferation and differentiation. However, which of the many molecular and biological functions of ETS proteins are most important in prostate cancer is not known.

Following *ERG, ETV1* is the most frequently overexpressed ETS gene in prostate cancer (~10% of the tumors) (10). The ETV1 protein translated from most fusion transcripts is truncated, lacking the 131 N-terminal amino acids (dETV1). Approx. half of the tumors with *ETV1* overexpression express a fusion transcript, the others show a high level of full-length *ETV1* expression. The *in vitro* biological and molecular properties of dETV1 seem different from those of the full-length 477 amino acid protein (10). This observation suggests that tumors with overexpression of full-length ETV1 are different from tumors expressing dETV1.

Little is known about the mechanism of full-length *ETV1* overexpression and its function in clinical prostate cancer. Our present results show that overexpression of full-length *ETV1* is correlated with rearrangement of the *ETV1* chromosomal region. Moreover, we identified a novel, N-truncated *ETV1* isoform with the same activity as full-length *ETV1*.

Results and Discussion

Previously, we reported *ETV1* overexpression in eight out of 84 prostate cancer samples. In four cases this was caused by a gene fusion, in the other four a full-length *ETV1* transcript was overexpressed (10). In the present study we investigated overexpression of *ETV1* by quantitative reverse transcriptase reaction (QPCR) in a novel cohort of 66 prostate cancers. In six RNAs *ETV1* overexpression was detected. Samples G51, G59, G233, G270 and G268 were derived from primary tumors and G210 was derived from a recurrent tumor. The six samples were further studied by QPCR with primer pairs at the 5'-end of *ETV1* mRNA (exon 1F and exon 4R) and at the 3'-end of the mRNA (exon 11F and exon 12R) (Figure 1A). A high exon 11-12 to exon 1-4 ratio is indicative for a fusion gene; a 1:1 ratio indicated overexpression of full-length *ETV1* mRNA. Based on these criteria, tumors G51, G59, G233 and G270 expressed full-length *ETV1* whereas G210 and G268 expressed a fusion transcript (see also control PC374 that expresses *TMPRSS2-ETV1*). *HNRPA2B1-ETV1* was found as the fusion transcript in sample G210 (data not shown); the fusion transcript in G268 has not been identified as yet. These two samples were not investigated further in this study.

In the next experiments we focused on the elucidation of the mechanism of overexpression of full-length *ETV1.* Recently, it has been shown that in two prostate cancer cell lines overexpressing full-length *ETV1*, LNCaP and MDA PCa2b, *ETV1* is translocated (8). We now addressed the question whether in clinical samples translocation of the complete *ETV1* gene might occur. To detect genomic rearrangements tissue slides of all four prostate cancer samples with full-length *ETV1* overexpression were analyzed by break-apart interphase FISH with two labeled BAC probes, one BAC recognized *ETV1* and the second one the flanking gene *DGKB* (Figure 1B). The series was supplemented with two samples harboring full-length *ETV1* overexpression from our previous study, G89 and G308 (10), for which frozen tissues of sufficient morphological quality were available. Figure 1B shows the results of the FISH experiments. Interestingly, we found split signals in all samples except for G59, indicating frequent *ETV1* rearrangements in prostate cancers that overexpress full-length *ETV1*. Absence of a split signal in G59 suggests absence of translocation, although we cannot exclude a breakpoint outside the investigated region. Its observed high frequency indicates genomic rearrangement as an important mechanism of full-length *ETV1* overexpression in clinical prostate cancer. Obviously, the chromosomal region to which *ETV1* is translocated can contribute to elucidation of the mechanism of *ETV1* overexpression. We were able to identify more details of the rearrangements in samples G270 and G89.

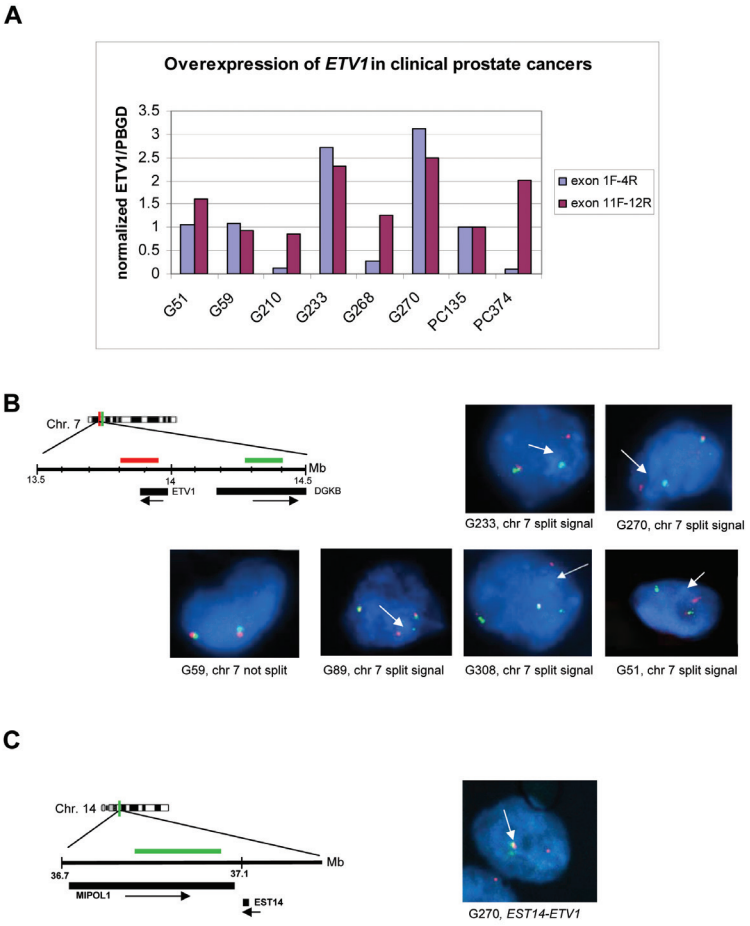


Figure 1. Full-length *ETV1* overexpression in clinical prostate cancers correlates with genomic rearrangement of the complete gene. (A) QPCR on RNA from clinical samples that show *ETV1* overexpression. Two primer-pairs were used to determine *ETV1* expression: *ETV1* exon 1 forward and exon 4 reverse, and exon 11 forward and exon 12 reverse, respectively. Primer sequences are given in Supplementary Table S1. Amplified products were quantified relative to the expression of the porphobilinogen deaminase (*PBGD*) housekeeping gene. Data were normalized to full length *ETV1* overexpressed in the xenograft PC135. A high exon 11/12 to exon 1/4 ratio indicates an *ETV1* fusion event, a 1 to 1 ratio indicates overexpression of a full length *ETV1* transcript. PC374 is a control xenograft that expresses a *TMPRSS2-ETV1* fusion gene. A representative experiment of the six samples that show *ETV1* overexpression is depicted. (B) Interphase FISH on fresh-frozen prostate cancer tissue sections. BACs used are indicated below the chromosome 7 region investigated. BAC RP11-124L22 (red) spans *ETV1* and RP11-1149J13 (green) overlaps *DGKB* (left panel). Positions of genes from the top of chromosome 7 are indicated in Mbp. A split signal representing an *ETV1* translocation is indicated by an arrow. (C) Translocation of *ETV1* to chromosome 14 in tumor G270. Tissue sections were hybridized with BAC RP11-460G19 (green) that overlaps *MIPOL1* and flanks *ETS14* and with the *ETV1* BAC RP11-124L22 (red) (see B for details). In the left panel the position of *MIPOL1* BAC on chromosome 14 is indicated. In the right panel a merging signal (yellow) shows co-localization of *ETV1* and *MIPOL1/EST14* in G270, as indicated by the arrow.

It has been shown in LNCaP cells that *ETV1* is translocated to 14q13.3-q21.1. The whole gene is integrated in the last intron of *MIPOL1*. In MDA PCa2b, *ETV1* is translocated to the same region, although the precise position is unknown (8). Moreover, we previously described insertion of truncated *ETV1* in the intron of a two exon gene encoding a ncRNA, denoted *EST14*, giving rise to an *EST14-ETV1* fusion transcript that contains *ETV1* exon 5-12 sequences (sample G342 in ref. 10; Figure 2A). Importantly, *EST14* maps directly adjacent to *MIPOL1* on 14q. Like most *ETV1* fusion partners, *EST14* is an androgen-regulated prostate-specific gene. To investigate whether the same chromosomal region was involved in full-length *ETV1* translocation in our novel cohort, interphase FISH was performed with the *ETV1* BAC (Figure 1B) in combination with a *MIPOL1* BAC (Figure 1C). A merging yellow signal was detected in sample G270 (Figure 1C) but in none of the other tumors. These data indicate that although there seems a preference for chromosome 14q13.3-21.1, other genomic regions will also contribute to rearrangement and overexpression of full-length *ETV1*.

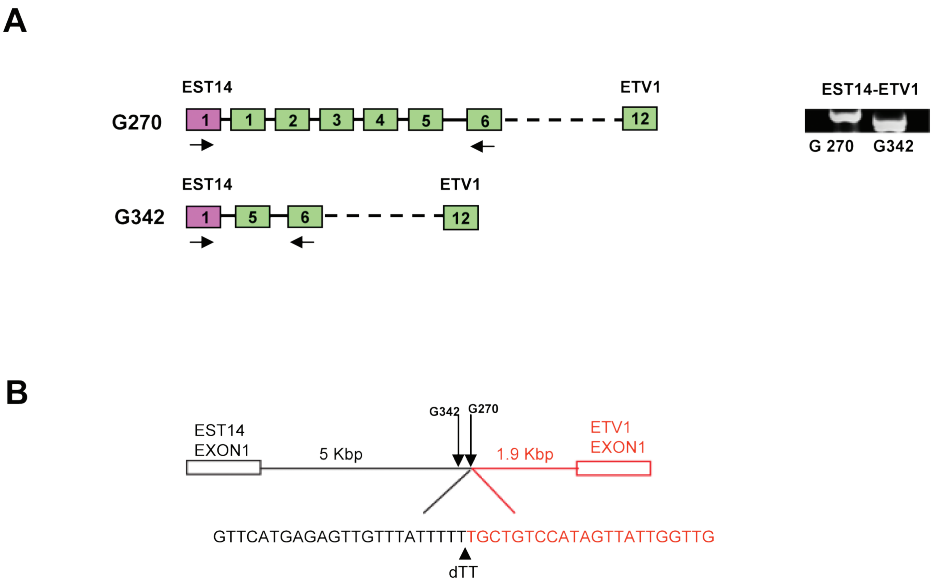


Figure 2. Characterization of the *EST14* to *ETV1* gene fusion in prostate cancer G270. (A) Schematic representation of *ETV1-EST14* fusion transcripts in prostate tumors G270 and G342 (sample G342 is from ref. 10). Arrows indicate positions of primers used in the RT-PCR experiment. Primer sequences are shown in Supplementary Table SI. (B) Sequence of the fusion point of *EST14* and *ETV1* in sample G270. The position of the fusion point in tumor G270 was mapped by long-range PCR on genomic DNA with a forward primer in the *EST14* intron and reverse primer upstream of *ETV1* exon 1. At the fusion point two T residues were lost. The breakpoint in *EST14* in G270 and G342 are indicated by arrows.

Additional information of *ETV1* rearrangement in G270 came from 5'-RACE of tumor cDNA (data not shown). Remarkably, we did not only detect as expected the full-length *ETV1* transcript but also a fusion transcript (Figure 2A). Such a result was not found for any of the other tumors overexpressing full-length *ETV1*. Sequencing showed that the fusion transcript in G270 was composed of *ETV1* exon 1-12 preceded by the first exon of *EST14* (Figure 2A). Scanning the *EST14* intron and *ETV1* flanking region by long-range PCR and sequencing mapped the breakpoints in G270 ~1.9 Kbp upstream of *ETV1* exon 1 and ~5 Kbp downstream of *EST14* exon 1. The breakpoint in *EST14* is only 180 bp apart from the breakpoint in G342 (Figure 2B and ref. 10).

Further information of *ETV1* rearrangement was also collected for sample G89. Previously, a xenograft propagated on male nude mice had been generated from this tumor (PC135). Like tumor G89, PC135 overexpressed full-length *ETV1* (3). The availability of the xenograft allowed the preparation of metaphase chromosome spreads. In multicolor FISH a complex chromosomal rearrangement pattern was found (data not shown). Individual chromosome paints were used to validate the multicolor FISH data. Painting of chromosome 7 indicated the presence of multiple chromosome 7 fragments (Figure 3). Hybridization with an *ETV1* BAC identified the presence of three gene copies: two in apparently normal chromosomes 7 and one in a complex marker chromosome. Follow-up experiments showed that the marker chromosome contained fragments of chromosomes 4, 7 and 10, as first indicated by multicolor FISH (Figure 3). The *ETV1* BAC hybridized at the junction of the chromosome 7 and the chromosome 4 fragment, strongly suggesting that the 4;7 translocation was instrumental in overexpression of *ETV1*. The precise positions of the breakpoints in 4 and 7 remain to be determined. Our data predict that multiple chromosomal regions are involved in overexpression of full-length *ETV1*. At least one of these regions is on chromosome 14 and a second one on chromosome 4. The chromosome 14 region is also involved in *ETV1* gene fusion. Deep sequencing technology could be instrumental in identification of other *ETV1* rearrangements.

Detailed characterization of the full-length *ETV1* transcripts in the various tumors by 5'-RACE and sequencing showed that not only *ETV1* exon 1- exon 12 transcripts were present but also various other full-length *ETV1* transcripts, resulting from alternative promoter usage. These transcripts were denoted as *ETV1*, *ETV1-1a*, *ETV1-1b* and *ETV1-1c.* Figure 4A and Supplementary Figure S1 show the positions of the different first exons in the gene and indicate the various ATG start codons. Of both *ETV1-1a* and *ETV1-1b* two splice variants were found (data not shown). QPCR experiments using transcript-specific primers on RNA from all six clinical prostate cancer samples that overexpressed

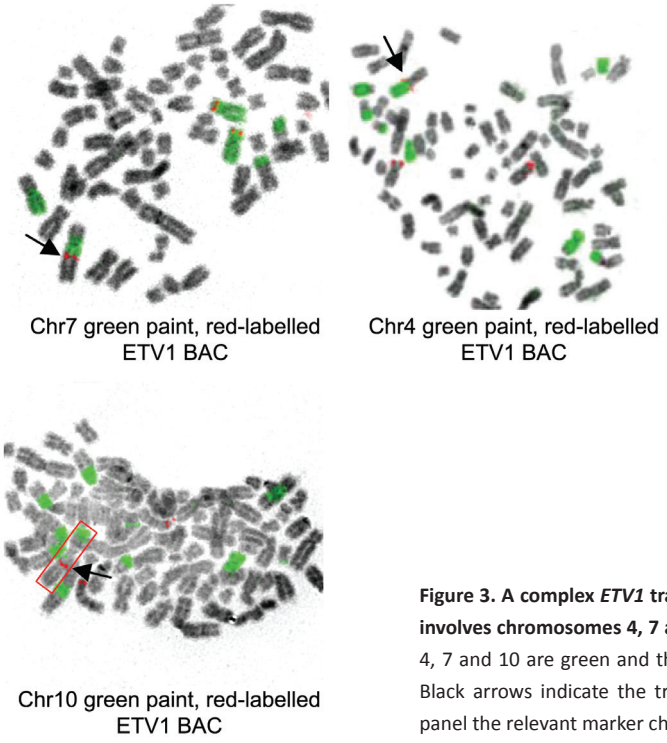


Figure 3. A complex *ETV1* translocation in xenograft PC135 involves chromosomes 4, 7 and 10. Paints of chromosomes 4, 7 and 10 are green and the *ETV1* BAC (Figure 1b) is red. Black arrows indicate the translocated *ETV1*. In the lower panel the relevant marker chromosome is boxed in red.

ETV1 showed that the level of expression of the different transcripts was variable in the various tumors (Supplementary Figure S2). *ETV1* was hardly expressed in control benign prostate hyperplasia sample G277. Figure 4B schematically represents the predicted composition of the various ETV1 protein isoforms that will be produced. Note that ETV1-1c is by far the shortest, lacking the N-terminal 60 amino acids, including the major part of the conserved acidic TAD. In dETV1 that is expressed by most fusion genes, the N-termimal 131 amino acids are absent. ETV1, ETV1-1b1 and -1b2 were of similar size, as shown in Western blots of lysates from transfected HEK293T cells (Figure 4B).

Previously, we have shown that ETV1 and dETV1 differed in stimulation of *in vitro* anchorage-independent growth (10). PNT2C2 cells infected with all novel ETV1 constructs induced anchorage-independent growth in a similar manner as ETV1 (Figure 4C). Remarkably, ETV1-1c, although expressed at a lower level and much smaller, is as active as the longer ETV1 isoforms. Thus, the full N-terminal TAD was not needed but amino acids 61-131 seem essential for biological activity of ETV1 (Figure 4 B, C).

In summary, the data presented reveal two important novel aspects of the role of ETV1 in prostate cancer. First, it is shown that in clinical prostate cancers a subgroup

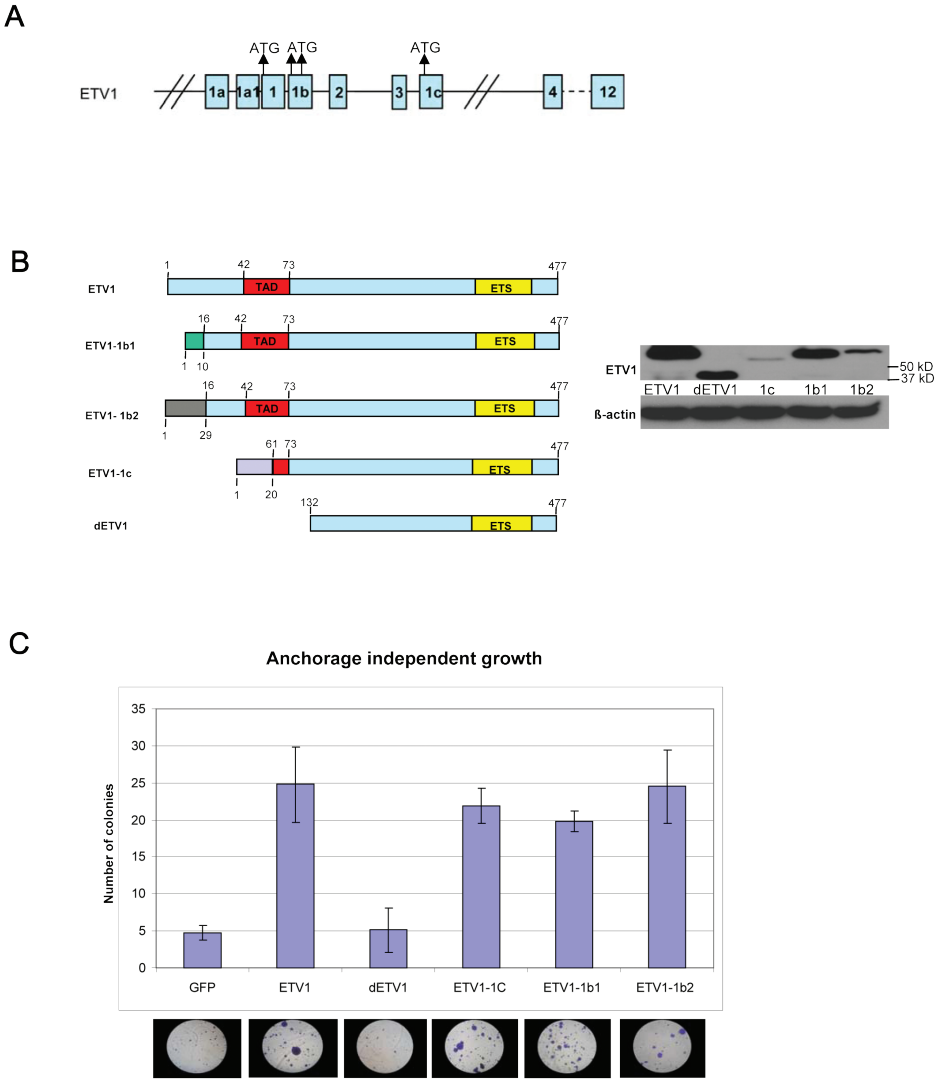


Figure 4. Alternative full-length *ETV1* transcripts give rise to proteins that all induce *in vitro* anchorage-independent growth. (A) Schematic representation of the organization of the *ETV1* gene. Alternative first exons are 1a, 1, 1b and 1c. The positions of the four different ATG start codons are indicated. More details are given in Supplementary Figure S1. (B) Schematic representation of the composition of different ETV1 proteins and their expression in transfected HEK293T cells. Different colors of N-terminal regions indicate a different amino acid composition. dETV1 is the truncated ETV1 protein produced by fusion genes. This protein is unable to induce anchorage independent growth. In the right panel a Western blot of ETV1 isoforms produced by transiently transfected HEK293T cells is shown. β-actin was used as loading control. (C) Soft-agar assay showing the anchorage-independent growth of PNT2C2 cells infected with lentiviruses expressing the various ETV1 isoforms or control GFP. The bars represent the average number of colonies per microscope field of three independent experiments (±SD). Representative images of the stained colonies are shown.

of ETV1 positive patients show full-length *ETV1* overexpression due to translocations of the whole gene to different chromosomes. This novel observation complements the well-described mechanism of overexpression of truncated ETV1 caused by gene fusions where expression regulation is determined by the promoter and enhancers of the fusion partners. Secondly, in contrast to dETV1 produced by gene fusions, a short isoform of full-length ETV1, ETV1-1c, lacking most of the N-terminal TAD, is as active as longer ETV1 isoforms, containing the complete N-terminal acidic TAD. This finding pinpoints the anchorage-independent growth to a small region that is absent in truncated ETV1 expressed by fusion genes.

It is highly relevant to extend the number of clinical samples in order to be able to compare tumor progression in the two subgroups of prostate cancers showing overexpression of truncated vs. full-length ETV1, and to determine the molecular mechanisms involved in their different biological behavior.

Materials and Methods

Ethics Statement

Use of the samples was approved by the Erasmus MC Medical Ethics Committee according to the Medical Research Involving Human Subjects Act in protocol MEC-2004-261, entitled "The use of human normal and cancer residual tissue from a tissue bank for characterization of DNA, RNA and protein".
Mice were housed according to guidelines of the Erasmus Medical Centre, and procedures were carried out in compliance with standards for use of laboratory animals. Animal experiments performed in this manuscript have been approved by the animal experimental committee of the Erasmus Medical Centre (DEC-consult Erasmus MC project 102-10-01).

Tissue samples, RNA and DNA isolation

Snap-frozen prostate cancers were obtained by radical prostatectomy or transurethral resection. Hematoxilin/eosin (HE) stained tissue sections were histologically evaluated by a pathologist (G. van Leenders). All samples contained at least 50% tumor cells.
RNA from clinical specimens was isolated using RNA-Bee (Campro Scientific, Berlin, Germany). DNA was isolated using the DNeasy DNA Extraction kit (Qiagen, Valencia, CA, USA). RNA from cell lines was isolated using the RNeasy RNA Extraction kit (Qiagen).

Breakpoint mapping

The position of the fusion point in tumor 270 was mapped by long-range PCR on genomic DNA using a forward primer in the *EST14* intron and reverse primer upstream of *ETV1* exon 1. PCR products were separated on a 1% agarose gel and sequenced in an ABI 3100 genetic analyzer (Applied Biosystems, Carlsbad, CA, USA).

Q-PCR

mRNA expression was analyzed by QPCR. cDNA was prepared with MMLV-RT (Invitrogen, Carlsbad, CA, USA) and oligo(dT)12 primer. QPCR was performed in Power SYBR Green PCR Master Mix on an ABI Prism 7700 Sequence Detection System (Applied Biosystems).

Amplified products were quantified relative to porphobilinogen deaminase (*PBGD*) by the standard curve method. For primers see supplementary Table SI.

RNA ligase-mediated rapid amplification of cDNA ends

5'-RLM-RACE was performed using the GeneRacer kit of Invitrogen. cDNA was amplified using the GeneRacer 5'-primer and a reverse gene-specific primer (ETV1 exon 6). PCR products were analyzed on a 1.5% agarose gel, and bands were excised, purified, and sequenced.

Fluorescence in situ hybridization

FISH was done on 5-µm frozen tissue sections according to standard protocols with minor modifications (14). BAC clones RP11-124L22 (*ETV1*), RP11-460G19 (*MIPOL1*), RP11-1149J13 (*DGKB*) were purchased from BacPac Resources (bacpac.chori.org). BACs were either digoxigenin-11-dUTP or biotin-16-dUTP (Roche, Basel, Schweiz) labeled and visualized with anti-digoxigenin FITC (Roche) or streptavidin-Alexa 594 (Invitrogen). Tissue sections were counterstained with DAPI. Images were collected on an epifluorescence microscope (Leica DM, Wetzlar, Germany) equipped with a charge-coupled device cooled camera (Photometrics, Tuscon, AZ, USA).
For preparation of metaphase spreads, xenograft PC135 was propagated on male nude mice. Single cells were collected by mincing and filtration. Metaphase preparation and hybridization were essentially as described (14). Chromosome paints 4, 7 and 10 were from Euro-Diagnostica (Malmö, Sweden). Metaphases were analyzed with an Axioplan 2 Imaging microscope (Carl Zeiss, Oberkochen, Germany) and images were captured using Isis software (MetaSystems, Altiussheim, Germany).

Expression plasmids

cDNAs of the different ETV1 isoforms were PCR amplified and cloned into pGEM-TEasy (Promega). Inserts were sequence verified and cloned into the pcDNA3 expression vector (Invitrogen) or the lenti-viral vector pWPXLd (Didier Trono, University of Geneva).

Western blot analysis

For Western blot analysis, HEK293T cells were transfected with the different pcDNA3-ETV1 expression constructs using the calcium phosphate precipitation method. Cells

were harvested 48 h after transfection. Western blot analysis was carried out using standard procedures with antibody directed to the ETV1 C-terminus (Santa Cruz, Santa Cruz, CA, USA). ß-actin was used as loading control (Sigma, St Louis, MO, USA). Proteins were visualized by chemiluminescence (Pierce, Rockford, IL, USA).

Lentiviral infections

HEK293T cells were cotransfected with pWPXLd-ETV1 expression vectors, or pWPXLd-GFP (control), and pPAX2 and pMD2.G (Trono) using the calcium phosphate precipitation method. Virus was harvested from the supernatant and used for infection of PNT2C2 cells. Pools of infected cells were propagated.

Soft agar assay

A layer of 0.6% low-melting agarose in standard culture medium was prepared in six-well plates. On top, a layer of 0.3% agarose containing $1x10^4$ PNT2C2 cells infected with various ETV1 expressing viruses or control PNT2C2-GFP cells were plated. At day 14, cells were stained with crystal violet and colonies were counted.

References

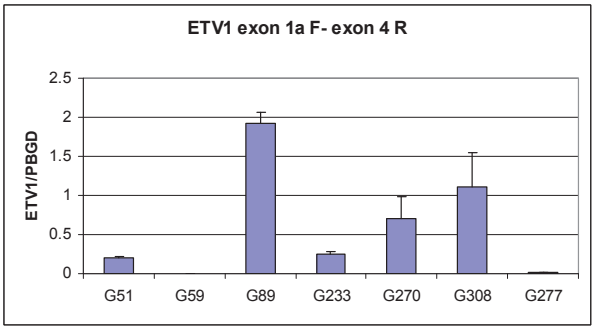
1. Futreal PA, Coin L, Marshall M, Down T, Hubbard T *et al.* (2004) A census of human cancer genes. Nat Rev Cancer 4: 177-183.
2. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R *et al.* (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 310: 644-648.
3. Hermans KG, van Marion R, van Dekken H, Jenster G, van Weerden WM *et al.* (2006) TMPRSS2:ERG fusion by translocation or interstitial deletion is highly relevant in androgen-dependent prostate cancer, but is bypassed in late-stage androgen receptor-negative prostate cancer. Cancer Res 66: 10658-10663.
4. Perner S, Demichelis F, Beroukhim R, Schmidt FH, Mosquera JM *et al.* (2006) TMPRSS2-ERG fusion-associated deletions provide insight into the heterogeneity of prostate cancer. Canser Res 66: 8337-8341.
5. Wang J, Cai Y, Ren C, Ittmann M. (2006) Expression of variant TMPRSS2/ERG fusion messenger RNAs is associated with aggressive prostate cancer. Cancer Res 66: 8347-8351.
6. Tomlins SA, Mehra R, Rhodes DR, Smith LR, Roulstron D *et al.* (2006) TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. Cancer Res 66: 3396-3400.
7. Helgeson BE, Tomlins SA, Shah N, Laxman B, Cao Q *et al.* (2008) Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. Cancer Res 68 :73-80.
8. Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X *et al.* (2007) Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature 448: 595-599.
9. Attard G, Clark J, Ambroisine L, Mills IG, Fisher G *et al.* (2008) Heterogeneity and clinical significance of ETV1 translocations in human prostate cancer. Br J Cancer 99: 314-320.
10. Hermans KG, van der Korput HA, van Marion R, van de Wijngaart DS, Ziel-van der Made A *et al.* (2008) Truncated ETV1, fused to novel tissue-specific genes, and full-length ETV1 in prostate cancer. Cancer Res 68: 7541-7549.
11. Wasylyk B, Hahn SL, Giovane A. (1993) The Ets family of transcription factors. Eur J Biochem 211: 7-18.
12. Oikawa T, Yamada T. (2003) Molecular biology of the Ets family of transcription factors. Gene 303 :11-34.
13. Seth A, Watson DK. (2005) ETS transcription factors and their emerging roles in human cancer. Eur J Cancer 41: 2462-2478.
14. Eussen BH, van de Laar I, Douben H, van Kempen L, Hochstenbach R *et al*. (2007) A familial inverted duplication 2q33-q34 identified and delinieated by multiple cytogenetic techniques. Eur J Med. Genet 50: 112-119.

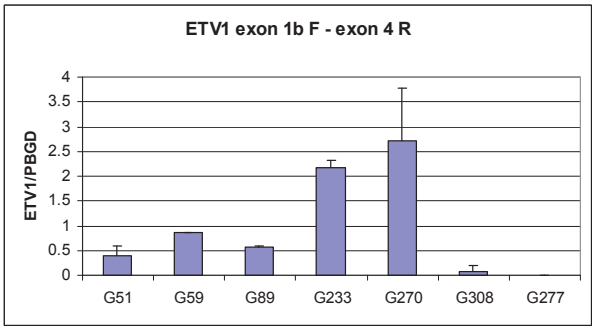
Primers Fig. 1A		
ETV1 E11-12	CATACCAACGGCGAGGATCA	TGGAGAAAAGGGCTTCTGGA
ETV1 E1-4	AGCTGAGATTTGCGAAGAGC	CTGCTCATCATTGTCAGGTAC
PBGD	CATGTCTGGTAACGGCAATG	GTACGAGGCTTTCAATGTTG
Primers Fig. 2A		
EST14 F	AGAGGAGAAAGAGTGCTCTA	
ETV1 E6 R		TACAAGACGACAGCTCAG
Primers Fig. 2B		
EST14 I1	TCTCCAGGCTTTTTCATCTC	
ETV1 upstream E1		AAGAGCGCACAGCACTCAAC
Supplementary Fig. 2		
ETV1 E1a F	GTTTTTGCAGCCTTCGCCT	
ETV1 E1b F	TAGAAGCCCAGTCTTCA	
ETV1 E1c F	TCTCTGGGTTCATCAGGA	
ETV1 E4 R	CTGCTCATCATTGTCAGGTAC	
PBGD	CATGTCTGGTAACGGCAATG	GTACGAGGCTTTCAATGTTG

Supplementary Table S1,
Gasi et al.

```
AGTAGAGGAA AAACCACCGA ATTGAGCTCT GTCAGAGGCG CTTTCGGCTT CCAAGGGGGA
AGTGCTGGGC TATAATTAAT GTTTTTATTA AATTTGGAGG GAAGTTTTTG CAGCCTTTCG Exon 1a
CCTAGCGTGG CCTTCAGGTG GGTCTTTCCA ACAACTTTTT ATGTCTCTCC AGATAATTGC
ATGGTTGTGG GTCGCAAAAA CTATCCTGAT GAAAGAGGTG TCTCCGTCTC TTTAGTCCCG
TTAGGTGCAA AGCAAGTCTC GTTGATCGCC ATTGCTAGTT TTGCACACGT TTGCGAATCA
GAGCTGCCCG GGGTACACCG ACCGCGCAGG GAAACATCGA GAGTGTAAAT AAATACATCG Exon 1a1
CCTCTTGTTC GGATTTTTGC TACTACCGAA AATATGTAAA TTGTGAACTC TGTTGGCTCT
CTTTGGATCC AGGTGAAGGA GGCGGTGTAG GGAGGGTTAT TTTTGTGAAT GGGACTGTCG
GAGGGTAATT AGCTATGCAA ATTCAAGAGC TTCATTCATG ATTTTTTTTT AAGCCTAAAA
GCCATCTTGT TCCCCTCTAG GTTGATAGAA GTCCAGATCC TGAGGAAATC TCCAGCTAAA
TGCTCAAAAT ATAAAATACT GAGCTGAGAT TTGCGAAGAG CAGCAGCATG GATGGATTTT Exon 1
ATGACCAGCA AGTGCCTTAC ATGGTCACCA ATGTGAGTGA TCAGTTTGAA AGTTGCTGTT
TATAAACTTG ACTCCGTCGG GGGTGGGGAG AGGGAGAAAA TGAGAAGGGA GGGGGCACGG
GGGTTGGGAT TGCAGATACT TATCTGCTTT GTTGCCACTG TAGGGCGACT CTGCTTCTAG
AAGCCCAGTC TTCAAAATGA GCTTACCTTT CAGTGATTTG GATAAGGCAT AGTTTTGTTT Exon 1b1/1b2
TTAAGACCCC TTTTTCTGAT TAAAGTGCCC AACATGAGTG GAAGAGGAGA TGGAGGGCAG
CAGCAGCTGC TGCACTCAAA GTTTTTGGCT GGGTTTGTCT GCCACATTGA AAAGAATGAA
GTTGAGACAA ATGCTGACAC TTTTTGTTTA TATGGGGTGT TTTTGTTTCT TTTTTTTTTC
TCATTTTCTC TTTTTTTTTT TTTTCATTTC TCTACTCTTT TTGCTTTTTT TTTTTGTGGG
GGTGGGGTGG GGAGGAGGGG CTGTCCAAGA GAGAATAAAA CCCACTGTTT TAATCTAGCA
TTGAACCAGC CTAAAAACAA CTTTAAGTGT AAATGCTTCA GGTTGTTTTG ATGCTGAGAT
ACCATGAATA TATGATTGAT AAATAGTAAT GTGCTAAGAT CGGCACTGGG AAGCAACTCT
CCCTGCCTGC TTAACATAAG CTTTCACTGT CAAATAAGTG AGAAAATGAA TATTAGAGTT
GATACTATGG ACTTCTATTT TTCTATTCAG AGAAGATCCA TTAGACTTAT GTATGCATGT
CTGTGTGTGT GTGTGTGCGC GCGCGTGTGC GTGTGTGTGA TCTGAGGTTT ACATTCTTTT
AAAGGATTTT ATCTTTCCCT TTGTAGAGTC AGCGTGGGAG AAATTGTAAC GAGAAACCAA Exon 2
CAAATGTCAG GAAAAGAAAA TTCATTAACA GAGATCTGGC TCATGATTCA GAAGGTGAGG
TTTGATTTTG GGCTGAATCC CCATTTTTTT ACCTACACTC TCCACCCTGC TAAAAGAAAA
GGAGCACTTC AGTCTTATCT TAAAAATAAT AAACTTTGAA ATTACTGAAT CTAACTAAGA
CTTTAACTTG GAACGATTTA CATAATGCGG CAGTTTGCAT GGGAGACTAT TTTATGGTCC
TATCACAAAG TAATGAAAGA AGGAAATAAA AACTTCTCAA TGACAACCTT CTTTTAATCA
TTGTTTTGTG TTTAAATACC TTTGTCCATG GGGCATATAA CTCACCCCAT TCACTAAGAT
AACAATTTAA AAATATCCAA TACAGTGGGA CAAATTCAGC AGCTATGTAC ACTGCTTACT
TTTTTCTTTC TCCTTTATTA CAAGTTAAAC AGTATTAGTG ATTGTTTGGT TTCCCACCTC
CTTTACACTA AAACCATCTA CTATTAACAG AATTTTCAAG GGTGATTTTG ATGACTAAAT
TGTGAGGTCA TAATACCTTG TTTATCGACC ATAATGTGTT CACTGAATCA TACTTAAGGC
CTCCTTAATA ATTATAATAC TATTTGGGAA ATTAGTTTGC AGATTCTTTT TATGAACTGT
TGCTAGGATG TCAGTATGTT ACAAATTATT TTAGCCATAT GAGGACATTC ATAATAAACT
TAATATACTT AGATTTAATT ATCAAAGTTT AGTTGGCAAT TTTAATTAGA TATACTGAAT
TCTGAAACTT CCTCAAATCT TCATACCATT TAAATACCTC CCTCTCTCCC TCCCTAACAT
GACCATATAG TTTAGAACTT GATGAGAACC TGGTGCCAAA GAATTATCTT TCTGACTTTT
ATAAACCTAT TAATGTAGAA ATTTTCCTCT AAAAACTTGT ATTTGAAGTG GTTTCTTTTA
GTTTCGAGGT ATTGAAGTTG AAAACATAAT ACTTTGACAG TTGTACCATG TACCATTATA
TATTTTATGT AAGATTCAGA AGATATAAAT AAAATGCTTT TCAGTCTTTC TAATCTTCCC
CCTACTACTA TTTTAGCAAG AGTTAAATAT ATTTAAGGAA GATTTCAGTG AACAGACACA
TATTCACCTC AGCCTTTTTG TGAATTCTGG TATCCCTGAG CTTATTTATT AAGCTGTTTG
TTAAAAAGAA GAAAAGATGA TTGACTTTAT TCTTTAGTTG ACTAAAATAA TTTCTTAAAG
AGTTGCTTAA TAGGCATGAA AAGTAGGCAG TATTATTGAG ATATTTAAGT TATGCCTCTT
TTTTTTATTT TTATGTTGAC TGCGTCTCTT GAAATTAAAT TTATATTTAA TACCAATTTC
ATGACAAATA TTTCCATTTA ATGAAGATTA AAGATTTGCA AATCTTATGT TTATGTCTTT
GACTTGTTTT TTAGAACTCT TTCAAGATCT AAGTCAATTA CAGGAAACAT GGCTTGCAGA
AGGTAAGGCA AAATTTGCTT TAAAAGGGGG GAAAGCAACT CTAGAAGGAG AAAGAAAAGA Exon 3
AGTCCTGAAC TTGCTGTCTT AATATTCAGC CCAATTAATT GAGCTCTAAA AGAGCCACCT
CATGTGTCAT GCATACATTA GAGCCTCTGA TGAGTTTGTC TTTGGGGACT CTGGGGCTGT
ACTACCCAGT GTCATCACAA ATTACCAGGA GAAATTGCTT CCAGCTCACA ATCACATCTG
CTTTTGGCAA GAACTAATGC ACCAAGACTT CAAGTTCTAA GCCTCTGTTC AGATTTTAAT
TGCAATTGAT CAGGTTTATA TTATTGTACC TCCAGAGACC TCCTAGAGCC AGAACCCGGC
TGGCTTGCTG TTTCCTTTAG AGCAGCGCAT ATCATTATTT GGTGTTCTGG TGGAGGACTT
TTCTGATGGC AGAAATTAGT TTCTCTGGGT TCATCAGGAC GGGATGCTTC AAGATTTAAG Exon 1c
TGCAAGTGTC TTCTTTCCAC CTTGTTCACA ACACAGAACG TTAGGTGTGT ATCTAATGCT
TAGGAAATCT
```

Figure S1. Part of genomic sequence of *ETV1*. The different exons are highlighted in yellow. Translation start codons are underlined. Note that the transcript starting at exon 1a may or may not include exon 1a1 but the translation start is the same as that of the transcript starting at exon 1. The end of exon 1b1 is indicated in red. *ETV1* transcripts starting at exon 1c lack exons 1, 2 and 3 which results in a truncated TAD in the translated protein.





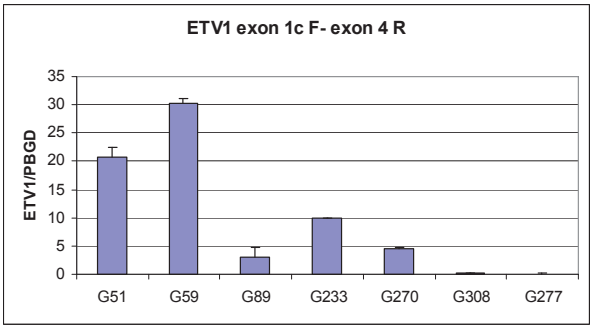


Figure S2. Expression of the different *ETV1* transcripts was determined by QPCR in the 7900HT Fast Real-Time PCR system from Applied Biosystems using the power SYBR-green master mix (Applied Biosystems). Expression levels are relative to the housekeeping gene *PBGD*. *ETV1* and *PBGD* primers are listed in Supplementary Table SI. Sample G277 is a BPH. It has very low or no expression of all of the different *ETV1* transcripts. Samples G51, G59, G89, G233, G270 and G308 all overexpress *ETV1*. The different transcripts are expressed in variable levels.

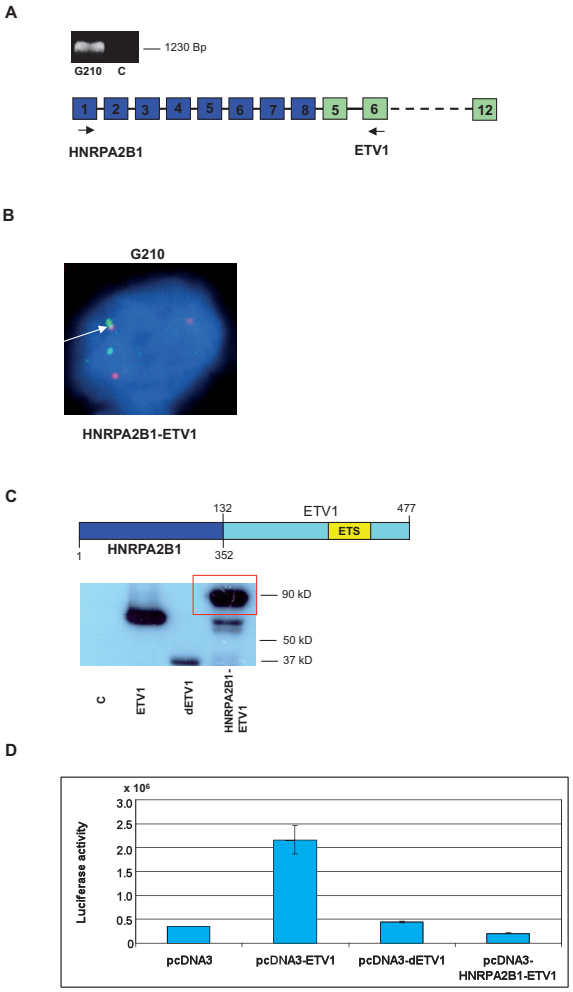
Appendix 1

The *HNRPA2B1-ETV1* fusion gene in clinical prostate cancer sample G210 gives rise to a fusion protein

As described in this chapter, novel *ETV1* fusion transcripts where identified by primarily measuring the outlier expression of *ETV1* by Q-PCR. The samples with overexpression where selected and tested for known fusion partners using RT-PCR. In one of the samples, G210, we found that *ETV1* was fused to *HNRPA2B1* (Fig 1A). The positions of the primers used are indicated by arrows in the schematic figure of the fusion event. The forward primer is in the first exon of *HNRPA2B1* and the reverse primer is in exon 6 of *ETV1*. The gene fusion was confirmed by FISH using BAC's spanning *ETV1* (red) and *HNRPA2B1* (green). The arrow indicates overlap of the red and green staining, showing the gene fusion (Fig 1B). By sequencing we discovered that the fusion coupled *HNRPA2B1* exon 8 to *ETV1* exon 5 sequence, retaining the open reading frame. The fusion gave rise to a chimeric protein composed of the N-terminal 352 residues of HNRPA2B1 and the C-terminal 345 residues of ETV1 (Fig. 1C). Previously, three other *HNRPA2B1-ETV1* fusion genes were described (1,2). In all cases a truncated ETV1 protein was expressed. To our knowledge, the gene described here is the first *ETV1* fusion gene giving rise to a chimeric protein.

Next, we investigated whether the HNRPA2B1-ETV1 protein was able to activate the ETV1-regulated PAL8x-TK promoter coupled to the Luciferase reporter gene, compared to full-length *ETV1* and the truncated dETV1 and found that, like dETV1, HNRPA2B1-ETV1 is not able to activate this promoter (Fig. 1D).

Fusion partners of *ETV1* are commonly prostate-specific and androgen-regulated (1,3,4 and this thesis). *HNRPA2B1* is a housekeeping gene and thus an exception to this rule. Nevertheless, there seems to be a preference for *HNRPA2B1*, since *HNRPA2B1-ETV1* fusion has been reported independently at several occasions. *HNRPA2B1* is located on chromosome 7 closely linked to *CBX3*. It was earlier described that the CpG islands in the *HNRPA2B1/CBX3* dual promoter region negates heterochromatin formation and prevents transcriptional silencing (5,6). It can be speculated that the mechanism of gene fusion between *ETV1* and its fusion partners is at least partly dependent on the chromatin structure. In the prostate, the prostate-specific and androgen-regulated genes might have the correct, open conformation. In addition, genes that have an ubiquitously-acting open chromatin structure might also be good fusion candidates, and possibly contribute to androgen-independent prostate cancer.



Appendix 1: Legend to the Figure

HNRPA2B1-ETV1 fusion in prostate cancer sample G210.

A. Identification of the *HNRPA2B1-ETV1* fusion by agarose gel electroforesis of an RT-PCR fusion fragment and schematic representation of the fusion transcript. Arrows indicate positions of primers used in RT-PCR (HNRPA2B1: GAGGTGCTCCTCGCAGAGTT, ETV1: GGTTTCGGTGTATGAGTTGA).

B. Interphase FISH of a fresh-frozen prostate cancer tissue section of G210 confirms gene fusion. BAC RP11-124L22 (red) spans *ETV1* and RP11-634J7 (green) overlaps *HNRPA2B1*. A merging signal (yellow) shows co-localization of *ETV1* and *HNRPA2B1*, as indicated by the arrow.

C. Schematic representation of the HNRPA2B1-ETV1 protein and expression of HNRPA2B1-ETV1 compared to ETV1 and dETV1 as analyzed on a Western blot. ETV1, dETV1 and HNRPA2B1-ETV1 are transiently overexpressed in HEK293T cells.

D. The HNRPA2B1-ETV1 fusion protein does not activate the PALx8-TK-LUC reporter. HEK293T cells were co-transfected with pcDNA3-ETV1, pcDNA3-dETV1, pcDNA3-HNRPA2B1-ETV1 or control pcDNA3 and the reporter gene construct.

Appendix 2:

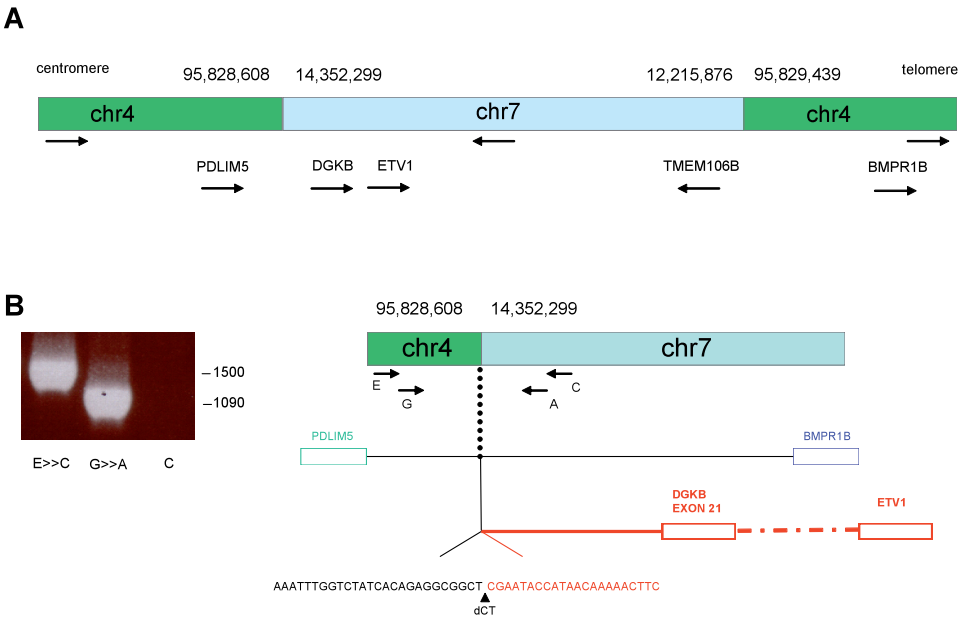
Mapping of the breakpoint of the ETV1 translocation in prostate cancer sample G89

In Chapter 2 we found that in all except one of the clinical prostate cancer samples with overexpression of full-length *ETV1,* the complete gene was translocated. In one sample, xenograft PC135, we could perform multicolor-FISH on chromosome spreads prepared from the tumor material. PC135 is derived from patient sample G89 harboring a complex chromosome 4;7;10 translocation encompassing *ETV1*, as described in Chapter 2. We used individual chromosome paints to show that *ETV1* mapped at the chromosome 4/7 border, but could not in more detail pinpoint the breakpoint.

Next, sample G89 DNA was sent to Complete Genomics (Mountain View, Cal, USA) for whole genome sequencing. In Figure 1A a schematic representation of the translocation event as indicated by the whole-genome sequencing of sample G89 is shown. It turned out that the *ETV1* translocation was even more complex than suggested by the chromosome painting experiments. Probably following or during the chromosome 4;7 translocation, an approx. 2 Mbp chromosome 7 fragment, containing *ETV1*, was integrated in the indicated region on chromosome 4. The integration site was at 95.8 Mbp from the top of chromosome 4. One of the breakpoints in chromosome 7 was in intron 20 of the *ETV1* flanking gene *DGKB* (Fig 1A,B).

PCR's were performed to confirm the translocation. Figure 1B (left panel) shows the PCR product on genomic DNA from G89 and from a control sample, using appropriate chromosome 4 and chromosome 7 primers. In the right panel, in Figure 1B, the primers used are represented by arrows. The breakpoint mapped approximately 20 Kbp from the gene *PDLIM5* and 70 Kbp from *BMPR1B.* The sequence of the breakpoint is given in the lower right panel of Figure 1B with black font for chromosome 4 sequence and red font for chromosome 7 sequence. *PDLIM5* is a scaffold protein that adheres protein kinases to the Z-disk in muscle tissue. It is highly expressed in heart and skeletal muscle, but also in the prostate. No reports were found of *PDLIM5* and prostate cancer.

In summary, we identified a novel insertion region of translocated full-length *ETV1*. The common feature of *EST14* (Chapter 2) and *PDLIM5* is that both genes are highly expressed in prostate. *EST14* is also androgen-regulated. The exact mechanism of overexpression remains to be investigated in both cases. Obviously, it would be interesting to investigate the position of the other translocated *ETV1* cases and to find a possible common mechanism.



Appendix 2: Legend to the Figure

Fine-mapping of the integration of *ETV1* to chromosome 4 in prostate cancer sample G89. A) The size, gene content and position of the translocated chromosome 7 fragment in chromosome 4. Arrows represent the positions and direction of the genes in this region. B) The translocation of ETV1 is confirmed by PCR on G89 genomic DNA. The positions of the primers E (GAAGCTCTGAGCTTTGGACA), C (AGGAAGAAGCACCTTAGGCA) and G (CGTTCAGGCGAAGCAAACAT), A (CTAGCAGACTCAAGTACACC) are indicated by arrows in the schematic figure of the breakpoint region. Insertion of the chromosome 7 fragment was between PDLIM5 and BMPR1B. The exact sequence of the breakpoint on chromosome 4 (black) and on chromosome 7 (red) is shown.

References:

1. Tomlins SA, Laxman B, Dhanasekaran SM et al: Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature 448:595-599, 2007
2. Miyagi Y, Sasaki T, Fujinami K et al. ETS family-associated gene fusions in Japanese prostate cancer: analysis of 194 radical prostatectomy samples. Mod Pathol 23:1492-8, 2010
3. Hermans KG, van der Korput HA, van Marion R et al. Truncated ETV1, fused to novel tissue specific genes and full-length ETV1 in prostate cancer. Cancer Res 68:7541-7549, 2008
4. Gasi D, van der Korput HA, Douben HC et al. Overexpression of full-length ETV1 transcripts in clinical prostate cancer due to gene translocation. PLoS One 6:e16332, 2011
5. Antoniou M, Harland L, Mustoe T et al. Transgenes encompassing dual-promotor CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing. Genomics 82:269-279, 2003
6. Williams S, Mustoe T, Mulcahy T et al: CpG-island fragments from the HNRPA2B1/CBX3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells. BMC Biotechnology 5:17, 2005

Chapter 3

Identification of ETV1 and dETV1 associated genes in PNT2C2-cells grown in 2D or *in vivo* and in clinical prostate cancers

Delila Gasi Tandefelt, Hetty van der Korput, Carola de Vos and Jan Trapman
Department of Pathology, Erasmus University Medical Centre, Rotterdam, The Netherlands

Abstract

ETV1 is upregulated in 5-10% of prostate cancers due to gene fusion or to translocation of the complete gene. Gene fusion causes the overexpression of an N-truncated ETV1 protein (dETV1), lacking an acidic transactivation domain. ETV1 and dETV1 differ in the capacity of inducing *in vitro* growth in soft-agar if overexpressed in the immortalized normal prostate cell line PNT2C2. They also differ in activation of an artificial ETV1-sensitive promoter.

We studied gene expression profiles of PNT2C2 cells overexpressing ETV1 or dETV1 grown *in vitro* in 2D-culture and grown *in vivo* as subcutaneous transplant on nude mice. Moreover, expression array data of clinical prostate cancers that overexpress ETV1 or dETV1 were collected.

In PNT2C2 cells grown in 2D-culture, ETV1 and dETV1 induced the expression of 149 and 182 genes 2-fold or more. Most of these genes were identical, with some exceptions. *BCL2A1* shows highest induction by both ETV1 and dETV1; *PAPPA* expression was preferentially induced in PNT2C2-ETV1 cells. But, in transient transfection promoter assays, both the *BCL2A1* and *PAPPA* promoter was stronger activated by ETV1 than by dETV1. In both cases this activation could be direct or indirect.

PNT2C2-ETV1 and PNT2C2-dETV1 cells grown *in vivo* resulted in bigger and faster growing tumors, as compared to PNT2C2-GFP control tumors. Gene expression experiments showed 73 more genes that were 2-fold or more induced in PNT2C2-ETV1 tumors than in PNT2C2-dETV1 tumors (245 and 172, respectively), of which approx. 40% were identical. *PAEP* was the highest preferentially induced gene in PNT2C2-ETV1 cells (30-fold). *MMP1* showed high induced expression by both ETV1 and dETV1 (approx. 10-fold).

Many genes were similarly regulated in 2D- and *in vivo* culture, but differences were also detected. Interesting examples were *MMP1*, *CRISP3 and PAEP,* which were much stronger induced *in vivo* than *in vitro*.

We found *GPR19* and *FKBP10* as most prominent *ETV1*/dETV1-associated genes expressed in clinical prostate cancers. Remarkably, almost no similarities were observed between clinical data and PNTC2-ETV1 and PNT2C2-dETV1 data. In clinical prostate cancers almost no overlap was found between *ETV1* and *ERG* associated genes.

Introduction

Fusion of the prostate-specific and androgen-regulated gene *TMPRSS2* and the ETS transcription factor gene *ERG* is the most frequent genomic alteration in prostate cancer, occurring in approximately 50% of all tumors (1). Upregulation of the ETS transcription factor gene *ETV1* occurs in 5-10% of prostate cancers (1-3).

Compared to *ERG*, *ETV1* has many more fusion partners (1,2,4-6). Most of these partner genes are androgen-stimulated and prostate-specific expressed (2, 4). *ETV1* can also be fused to the *HNRPA2B1* housekeeping gene, or to the androgen-downregulated gene *C15orf21* (4,5). Overexpression of *ETV1* can be driven by fusion to other genes or by translocation of the complete *ETV1* gene to novel genomic sites (3).

All ETS transcription factors share high homology in their DNA-binding domain but substantially differ in other regions of the proteins (7). ETV1 is a member of the PEA3 subfamily of ETS transcription factors (8). Additional to their DNA-binding domain the PEA3 subfamily members have a transactivation domain (TAD) in their N-terminus. If *ETV1* is involved in fusions to other genes a truncated protein (dETV1) is produced, lacking a fragment of 131 amino acid residues, including TAD. In the non-tumorigenic PNT2C2 prostate cell line overexpression of ETV1 induces *in vitro* anchorage independent growth. This function was not shared by PNT2C2 cells overexpressing dETV1 (2). In transient transfection experiments, ETV1 and not dETV1, had transactivating properties (2,3).

Because in clinical prostate cancers overexpression of ERG and ETV1 is mutually exclusive, it has been suggested that all ETS gene fusions define one common subclass of prostate cancers (4,9). However, in clinical prostate cancers, *ETV1* overexpression is correlated with higher Gleason score, PSA level at diagnosis and with higher clinical stage (5), which indicates a subgroup of prostate cancers with poor prognosis. But it is not clear whether ERG and the PEA3 family of ETS transcription factors have the same or different function. So, it is important to clarify the function of ETV1 and dETV1.

Several studies showed that overexpression of *ETV1* induced PIN-like lesions in mouse prostate cancer models (10,11). One study described ETV1-associated genes in clinical prostate cancer samples (12).

In this study we used gene expression arrays of PNT2C2 cells overexpressing ETV1 or dETV1 to identify genes associated with these two variants of ETV1. Furthermore, we describe differences and similarities of the ETV1 and dETV1 regulated genes in different growth conditions. Finally, we describe *ETV1*-associated genes in gene expression arrays from clinical prostate cancer samples and compare these to *ERG*-associated genes and the cell line data.

Materials and Methods

Ethics Statements

Use of the clinical prostate cancer samples was approved by the Erasmus MC Medical Ethics Committee according to the Medical Research Involving Human Subjects Act in protocol MEC-2004-261, entitled “The use of human normal and cancer residual tissue from a tissue bank for characterization of DNA, RNA and protein”.
Mice used for tumor transplantation were housed according to guidelines of the Erasmus Medical Centre, and procedures were carried out in compliance with standards for use of laboratory animals. Animal experiments performed in this manuscript are filed and approved by the Dutch committee for experimental animals in research (DEC) to comply with the “Code of Practice for use and handling of experimental animals in cancer research” as introduced by Dutch law, and by the animal experimental committee of the Erasmus Medical Centre (DEC-consult Erasmus MC project 102-10-01).

Tissue samples

The initial treatment of clinical prostate cancer patients was radical prostatectomy in all cases. Samples were snap-frozen and stored in liquid nitrogen. Hematoxilin/eosin stained tissue-sections were histologically evaluated by two pathologists (T. van der Kwast; G.J.H.L. van Leenders). Only samples that contained at least 50% tumor cells were selected for analysis.

RNA and DNA isolation

RNA from clinical specimens was isolated from frozen tissue sections using RNA-Bee (Campro Scientific, Berlin, Germany). RNA from PNT2C2 cells overexpressing ETV1 or dETV1 or GFP control, and from frozen *in vivo* grown tumors of these cell lines were isolated using the RNeasy RNA Extraction kit (Qiagen, Valencia, CA).
DNA was isolated from the prostate tumor cell line VCaP, using the DNeasy DNA Extraction kit (Qiagen).

In vivo tumor growth of infected PNT2C2 cells in male nude mice

For each PNT2C2 cell line overexpressing ETV1 or dETV1 and GFP control, four nude mice (NMRI, Taconic, M&B, Ry, Denmark) were subcutaneously injected with these cells; $5x10^{6}$/*100µl each.* At different time points, tumor volumes were determined. Four weeks after injecting cells, tumors were harvested. Samples were snap-frozen and stored at -80°C.

Expression plasmids

cDNAs of ETV1 and dETV1 were PCR amplified and cloned into pGEM-TEasy vector (Promega, Madison, WI). For primers see Supplementary Table S10. Inserts were verified by sequencing and cloned into the EcoRI site in pcDNA3 expression vector (Invitrogen, Carlsbad, CA) or in the lentiviral vector pWPXLd (Didier Trono, University of Geneva). As a controle we used the same lentiviral vector expressing GFP.

Reporter plasmids

BCL2A1 and PAPPA promoter constructs were generated by PCR amplification with different primer combinations (see Supplementary Table S10 for primers) on genomic DNA from VCaP cells sequenced and cloned in front of the LUC reporter into the promoterless vector pLuc. The BCL2A1 (2) construct was made by restriction digestion from BCL2A1 (1) by deletion of a SalI/XhoI fragment.

Lentiviral infections

For lentivirus production HEK293T cells were cotransfected with different pWPXLd-ETV1 expression vectors, or pWPXLd-GFP (control), and pPAX2 and pMD2.G (Trono) using the calcium phosphate precipitation method. Viruses were harvested from the supernatant and used for infection of PNT2C2 cells. Pools of infected cells were propagated and used in all assays.

Affymetrix micro-array hybridization

PNT2C2 cells

Labelled antisense RNAs from different PNT2C2 cells grown in vitro or in vivo (two tumours per cell line) were hybridized to Affymetrix GeneChip Human Genome U133 plus 2.0 arrays containing 54 614 probe sets. Antisense biotinylated RNA was prepared from 1 µg total RNA according to the Affymetrix GeneChip eukaryotic one-cycle target preparation protocol (Affymetrix, Santa Clara, CA). Briefly, single-stranded cDNA was synthesized using a T7-Oligo(dT) Promoter Primer, followed by RNase H-facilitated second-strand cDNA synthesis. The cDNA served as a template in the subsequent in vitro transcription (IVT) reaction utilizing T7 RNA polymerase and a biotinylated nucleotide analogue/ribonucleotide mix for cRNA. The biotinylated cRNA targets were then cleaned up and fragmented. The amount of fragmented-labeld cRNA was determined by standard spectrophotometric analysis, and the quality was checked on an Agilent 2100 Bioanalyzer (Agilent, Amstelveen, the Netherlands), using an RNA 6000 Nano assay. Affymetrix GeneChip was hybridized with 15 µg of fragmented biotinylated cRNA (45 °C for 16 h) and subsequently washed, stained with streptavidin phycoerythrin and imaged according to the GeneChip Expression Analysis Technical Manual (Affymetrix). All GeneChips were visually inspected for irregularities.

Clinical prostate cancer samples

RNA samples for microarray analysis were processed according to the protocol of the array supplier (Affymetrix). Expression profiles were determined using GeneChip Human Exon 1.0 ST (Affymetrix) at the center of Biomics, Erasmus Medical Centre, Rotterdam and at ServiceXS, Leiden, The Netherlands. Microarray data were processed and robust multichip average (RMA) quantile normalized using Partek Genomics Suite (St. Louis, MO, USA). For further normalization on gene level, values below 30 were set to 30. The samples were normalized to geometric mean and then log2 values of the samples were calculated. Gene Cluster from Stanford microarray resources and TreeView were used for unsupervised hierarchical average linkage clustering. Significance of microarrays (SAM) analysis was used to identify differentially expressed genes between subgroups.

Quantitative PCR (Q-PCR)

For Q-PCR analysis cDNA samples were prepared with M-MLV reverse transcriptase (Invitrogen) and oligo(dT)12 primer. QPCR was performed using Power SYBR Green

PCR Master Mix (25 µl), containing 0.33 µM forward and reverse primer on an ABI 7900 HT Fast Real-Time PCR system (Applied Biosystems, Foster City, CA). Amplified products were quantified relative to porphobilinogen deaminase (*PBGD*) by the standard curve method. For primers see Supplementary Table S10.

Western blot analysis

For Western blot analysis, PNT2C2 cells overexpressing ETV1, dETV1 and GFP were grown in T25 flasks and lysed directly in the cell culture flasks. Western blot analysis was carried out using standard procedures with antibody directed to the ETV1 C-terminus ER81 (Abcam, Cambridge, UK). ß-actin was used as loading control (Sigma, St Louis, MO, USA). Proteins were visualized by chemiluminescence (Pierce, Rockford, IL, USA).

Transfection/Reporter assays

HEK293T cells and PNT2C2 cells were grown in DMEM supplemented with 5% FCS and antibiotics. Cells were co-transfected with ETV1 or dETV1 expression constructs and BCL2A1 or PAPPA LUC promoter constructs using the calcium phosphate precipitation method. Cells were harvested after 24h and luciferase activity was measured in a Glomax 96 Microplate Luminometer (Promega).

Results

Identification of ETV1- and dETV1-associated genes in PNT2C2-cells grown in 2D-cultures
To identify genes that are regulated by ETV1 and dETV1 we generated PNT2C2 cells infected with lentiviruses expressing ETV1 or dETV1, or GFP as the control vector. In Figure 1A, the expression of the two ETV1 variants in the PNT2C2 cells is shown with dETV1 slightly higher expressed than ETV1.
RNA from PNT2C2, PNT2C2-GFP, PNT2C2-ETV1 and PNT2C2-dETV1 was isolated, cRNA was synthesized and hybridized to the Affymetrix GenChip Human Genome U133 plus 2.0 array. The gene expression profiles of the different cell lines were compared. All expression data obtained were normalized to the values of the PNT2C2-GFP control sample. Gene probes with more than two-fold induction by *ETV1* or *dETV1* in PNT2C2 cells compared to PNT2C2-GFP were identified. The Venn diagram in Figure 1B shows the number of genes that are more than two-fold up-regulated by ETV1 or dETV1 overexpression. An almost identical number of genes showed more than 2-fold overexpressed in PNT2C2-ETV1 and PNT2C2-dETV1 cells: 159 genes were higher in PNT2C2-ETV1 cells than in PNT2C2-GFP cells and 182 were higher in the PNT2C2-dETV1 cells. Sixty-three genes were two-fold higher induced in PNT2C2-ETV1 cells but not in PNT2C2- dETV1. Similarly, 86 genes were more than 2-fold induced in PNT2C2-dETV1 cells and not in PNT2C2-ETV1 cells (Supplementary Table S2). Remarkably, most of the genes overexpressed in PNTC2-ETV1 and PNT2C2-dETV1 cells (96) were identical, and the level of overexpression was quite similar for most genes (Supplementary Table S3).
BCL2-related protein A1 (BCL2A1) was the highest upregulated gene in both ETV1 and dETV1 overexpressing PNT2C2 cells (20- to 30-fold) (Table 1). Pappalysin 1 (*PAPPA*) and serglycin (*SRGN*) seemed stronger regulated by ETV1 than by dETV1 (Table 1). To confirm the expression array data, expression of *ETV1, BCL2A1, PAPPA* and *SRGN* were measured by Q-PCR (Figure 1C). Indeed, all genes were induced by both ETV1 and dETV1 compared to PNT2C2-GFP cells and parental PNT2C2 cells, although the relative expression levels showed some differences. By Q-PCR, the expression of *ETV1* in PNT2C2 –ETV1 cells seemed slightly lower than that of *dETV1* in PNT2C2-dETV1 cells. Higher expression of *PAPPA* and *SRGN* in PNT2C2-ETV1 cells as measured in the expression arrays was confirmed by Q-PCR.

Promoter analyses of selected ETV1-induced genes in HEK293T and PNT2C2 cells
To obtain evidence for direct regulation of *BCL2A1* and *PAPPA* by ETV1 and dETV1 we studied the induction of activation of *BCL2A1* and *PAPPA* promoter constructs co-

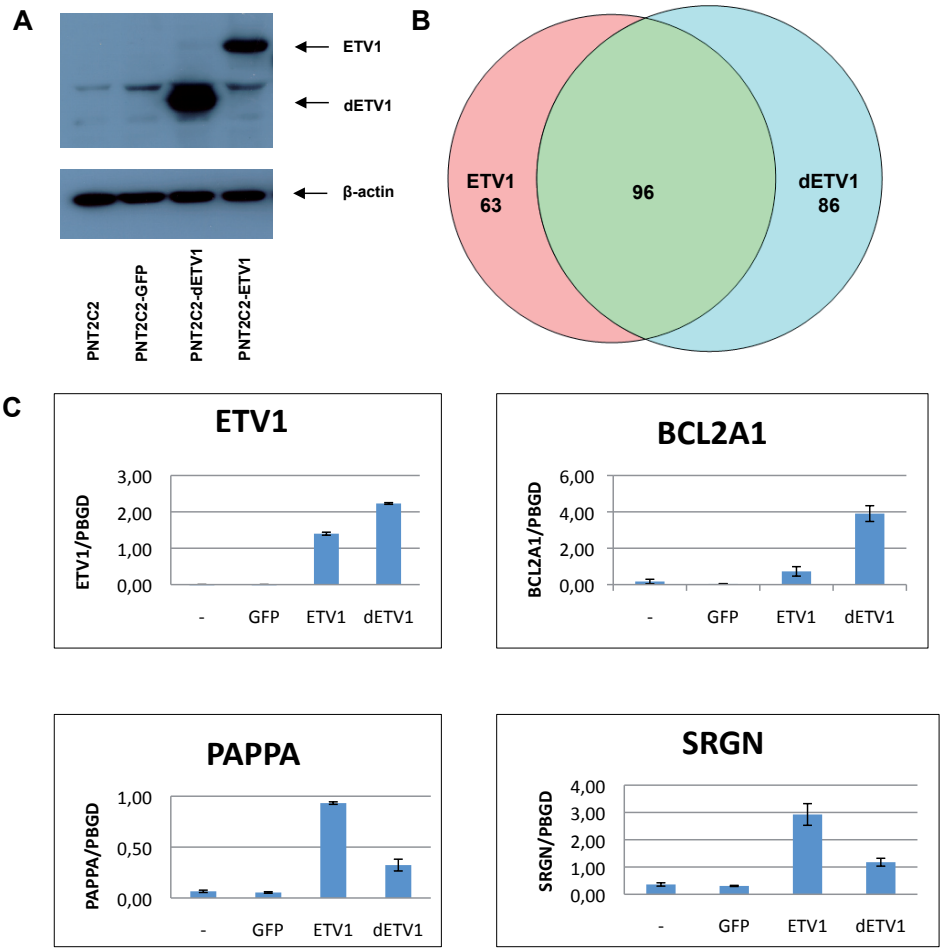


Figure 1. Genes associated with ETV1 and dETV1 overexpression in PNT2C2 cells grown in 2D-culture. A) Western blot showing the expression of ETV1 and dETV1 in PNT2C2-ETV1 and PNT2C2-dETV1 cells. Lysates of PNT2C2-GFP cells and parental PNT2C2 cells were used as control samples. ETV1 antibody recognizing the C-terminus was used for ETV1 detection and ß-actin antibody was used as a loading control. **B)** Venn-diagram showing two-fold ETV1 and dETV1- induced genes. The number of overlapping genes is shown in the green overlapping part of the circles. **C)** Expression of ETV1, BCL2A1, PAPPA and SRGN mRNA relative to PBGD expression in the indicated PNT2C2 cell lines as measured by Q-PCR, n=2.

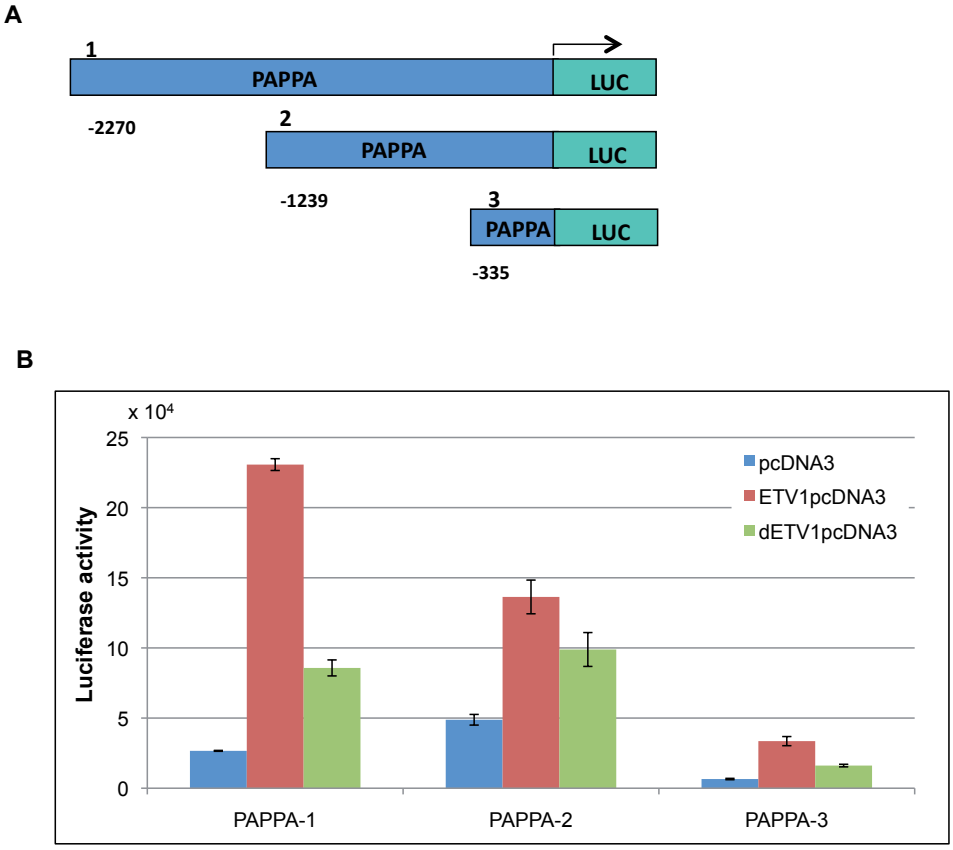


Figure 2. Activity of the PAPPA promoter in HEK293T cells transiently transfected with ETV1 and dETV1 expression plasmid. A) Schematical representation of the different PAPPA-promoter constructs **B)** Activity of PAPPA promoters indicated in A as measured by luciferase activity, n=2.

transfected with ETV1 or dETV1 expression plasmids in HEK293T cells. First, we searched for candidate consensus ETS binding sequences GGAA/T or A/TTCC in the promoter regions of the two genes, and then generated plasmids in which promoter fragments of different size were coupled to the luciferase reporter gene. The longest *PAPPA* promoter fragment was 2.2 Kbp; additionally two shorter promoter fragments (1.2 Kbp and 335 bp, respectively) were generated (Figure 2A). As measured by luciferase activity we found that all three promoter fragments were activated by ETV1 and to a lower extent by dETV1 in HEK293T cells (Figure 2B). The promoter activity decreased in shorter fragments. However, this decrease was not related to the loss of predicted candidate ETS binding sites.

Similarly, we generated five promoter constructs of *BCL2A1*, ranging from 1.8 Kbp to 232 bp (Figure 3A) and used these in promoter assays. Although the *BCL2A1* expression data indicated up-regulation by *ETV1* and *dETV1* to a similar level, the promoter experiments in HEK293T cells demonstrated that ETV1 activated the promoter to a much higher extent than dETV1 (Figure 3B). Again, luciferase reporter activity was lower using smaller promoter fragments. Moreover, although ETV1 and dETV1 clearly increased promoter activity, no evidence was found for an obvious direct role of both transcription factors.
Using the longest *BCL2A1* promoter fragment, we could also measure luciferase activity in transient transfected PNT2C2 cells (Figure 3C). We confirmed that ETV1 and dETV1 were capable of activating the *BCL2A1* promoter although dETV1 was clearly less active than ETV1.

Identification of ETV1- and dETV1-associated genes in PNT2C2 cells grown in vivo as transplanted tumors

Next we investigated whether growth conditions affected expression of ETV1- and dETV1-associated genes. To study different conditions we attempted to isolate RNA from PNT2C2-cells grown in soft-agar and grown *in vivo* as transplanted tumor cells. For the *in vivo* growth the different PNT2C2 cells were subcutaneously injected in nude mice. Unfortunately, the quality and quantity of RNA from cells grown in soft-agar was not sufficient for gene expression experiments and thus we limited our study to analysis of RNAs expressed in tumors transplanted on nude mice.
All mice injected with PNT2C2-ETV1, PNT2C2-dETV1 and PNT2C2-GFP cells developed tumors during a period of 4 weeks. In Figure 4A we show the volumes of PNT2C2-ETV1, PNT2C2-dETV1 and control PNT2C2-GFP tumors at one week intervals. The PNT2C2-GFP tumors were clearly smaller and had a slower growth rate compared to PNT2C2-ETV1 and PNT2C2-dETV1 tumors. The growth pattern of PNT2C2-ETV1 and PNT2C2-dETV1 tumors was very similar, although at 4 weeks PNT2C2-ETV1 tumors were slightly larger than PNT2C2-dETV1 tumors.
To distinguish the genes that were associated with ETV1 or dETV1 expression in *in vivo* transplanted cells, genes with more than two-fold induction in PNT2C2-ETV1 and in PNT2C2-dETV1 tumors compared to PNT2C2-GFP tumors were selected. In the Venn diagram (Figure 4B) it is shown that 245 genes were expressed in PNT2C2-ETV1 and 172 were induced in PNT2C2-dETV1 tumor samples (see also Suppl. Tables S4-S6). Sixty-seven genes were more than two-fold upregulated in both PNT2C2-ETV1 and PNT2C2-dETV1 grown as transplants (Suppl. Table S5). Most of these genes were preferentially induced by ETV1 (Table 2 and Suppl. Tables S4 and S6). However, there are exceptions.

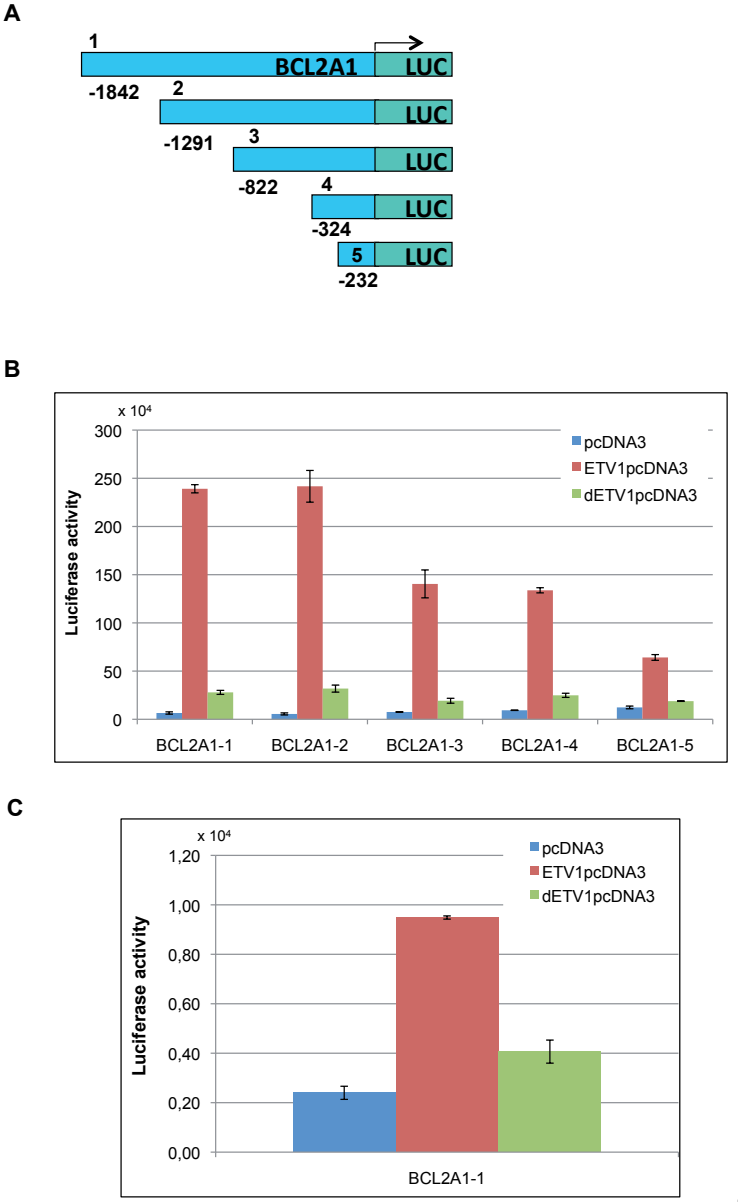


Figure 3. Activity of the BCL2A1 promoter in HEK293T cells and in PNT2C2 cells transiently transfected with ETV1 and dETV1 expression constructs. A) Schematical representation of the different BCL2A1-promoter constructs. **B)** Activity of BCL2A1 promoters indicated in A as measured by luciferase activity in HEK293T-cells. **C)** Activity of the BCL2A1-1 promoter as measured by luciferase activity in PNT2C2 cells, n=2.

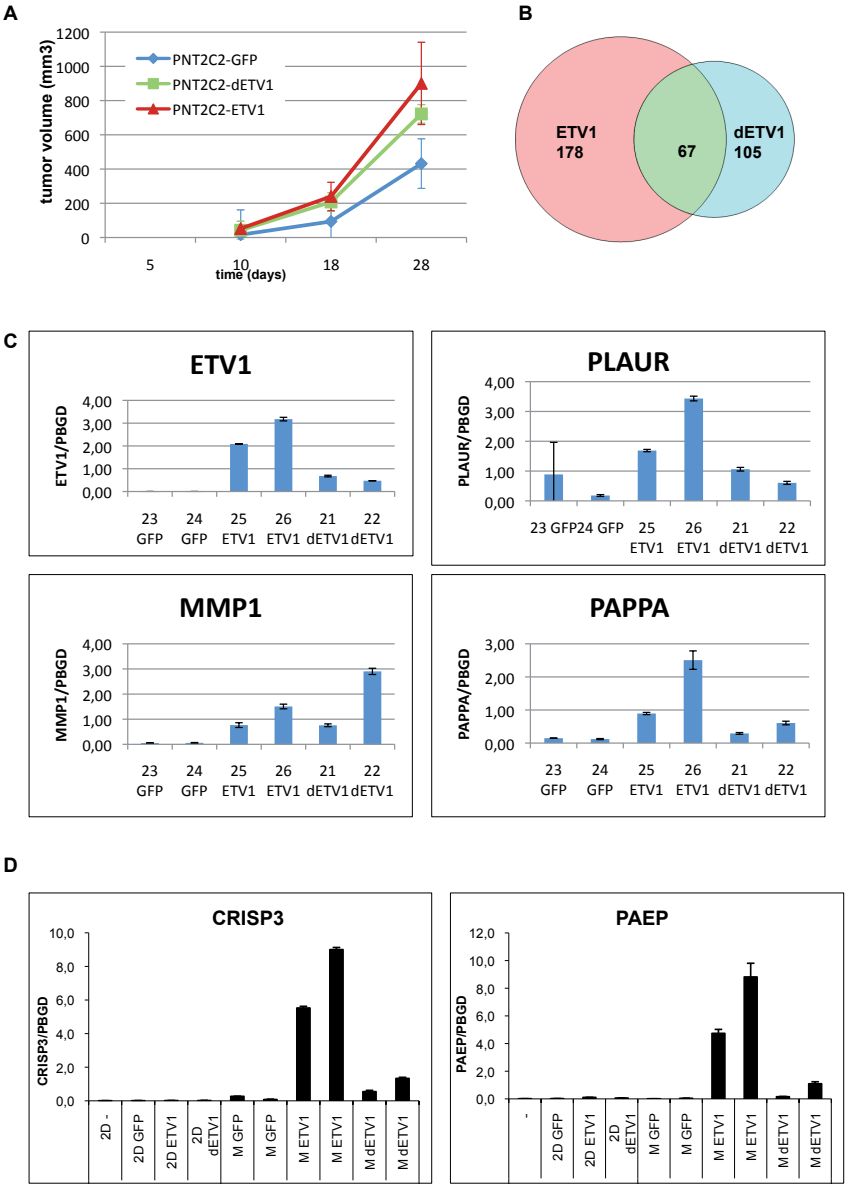


Figure 4. Genes associated with ETV1 and dETV1 overexpression in PNT2C2 cells grown as transplanted tumors on nude mice. A) Volumes of PNT2C2-ETV1, PNT2C2-dETV1 and control PNT2C2-GFP tumors as measures at different time points after inoculation of indicated cells. **B)** Venn-diagram showing two-fold ETV1- and dETV1-induced genes. The number of overlapping genes is shown in the green overlapping part of the circles. **C)** Expression of ETV1, PLAUR, MMP1 and PAPPA mRNA measured by Q-PCR relative to PBGD expression in the different tumors harvested at 28 days. 23 and 24: PNT2C2-GFP; 25 and 26: PNT2C2-ETV1; 21 and 22: PNT2C2-dETV1, n=2. **D)** Expression of CRISP3 and PAEP mRNA measured by Q-PCR relative to PBGD expression in different PNT2C2 cells grown in 2D and grown as tumors (M), n=2.

As shown in Suppl. Table S6 there are genes that are strongly induced by both ETV1 and dETV1, and probably even stronger by dETV1. *MMP1* and *PLAT* are interesting examples. We confirmed *MMP1* expression data by Q-PCR (Figure 4C).
Next, we compared the gene expression array data of PNT2C2-GFP, PNT2C2-ETV1 and PNT2C2-dETV1 cells grown *in vitro* in 2D-culture with the genes expressed in tumors in mice. Comparison of the 2D and tumor data indicated more genes preferentially induced by ETV1 than dETV1 *in vivo*, which was not the case in 2D grown cells. Still, there are similarities independently of growth condition. *PAPPA, THBD* and *GPR116* are examples of genes that are induced by *ETV1* both in 2D cultures and *in vivo* grown tumors (see Tables 1 and 2). Q-PCRs of *PLAUR, MMP1* and *PAPPA* together with *ETV1* expression are shown as examples (Figure 4C). *BCL2A1,* the gene showing highest induction by both *ETV1* and *dETV1* in 2D-grown PNT2C2 cells (Table 1) was not present in the gene list of highest induced genes in the *in vivo* grown cells. By Q-PCR, similarly to in 2D-grown cells, the expression and induction of *PAPPA* is *in vivo* clearly higher in PNT2C2-ETV1 and PNT2C2-dETV1 compared to the PNT2C2-GFP control and slightly higher in ETV1 than in dETV1 tumor samples. There are examples of genes that are highly induced by ETV1 and dETV1 *in vivo*, but not in 2D-grown cells. By Q-PCR we confirmed that *CRISP3* and *PAEP* were hardly expressed in 2D culture but clearly induced *in vivo* by both ETV1 and dETV1. Both genes are preferentially induced by ETV1 (Figure 4D).

ETV1-associated genes in clinical prostate cancers are not identical to ETV1-associated genes in PNT2C2 cells

To identify genes that are associated with *ETV1* overexpression in clinical prostate cancer samples we performed a SAM comparison of 10 *ETV1* overexpressing primary prostate cancer samples with 11 ETS-negative primary prostate cancers (Figure 5A and Suppl. Table S7). The ETV1 group was composed of 7 tumors that overexpressed full-length ETV1 and 3 dETV1 expressing samples. *GPR19* and *FKBP10* were the two strongest ETV1-associated genes and *TARP* was the most prominent downregulated gene in the *ETV1* overexpressing samples. We compared the data obtained in the PNT2C2 cells with clinical prostate cancers. Remarkably, there is no difference in expression of *GPR19*, *FKBP10* and *TARP* between PNT2C2-GFP-control cells and the PNT2C2-ETV1 or PNT2C2-dETV1 cells grown in 2D or *in vivo* (data not shown). Moreover, genes that are highly induced by *ETV1* in PNT2C2 cells have no correlation with genes regulated in *ETV1* positive clinical prostate cancer samples (data not shown). *CRISP3* is the only gene that shows some differential expression in clinical samples. *CRISP3* is expressed in tumor and not normal

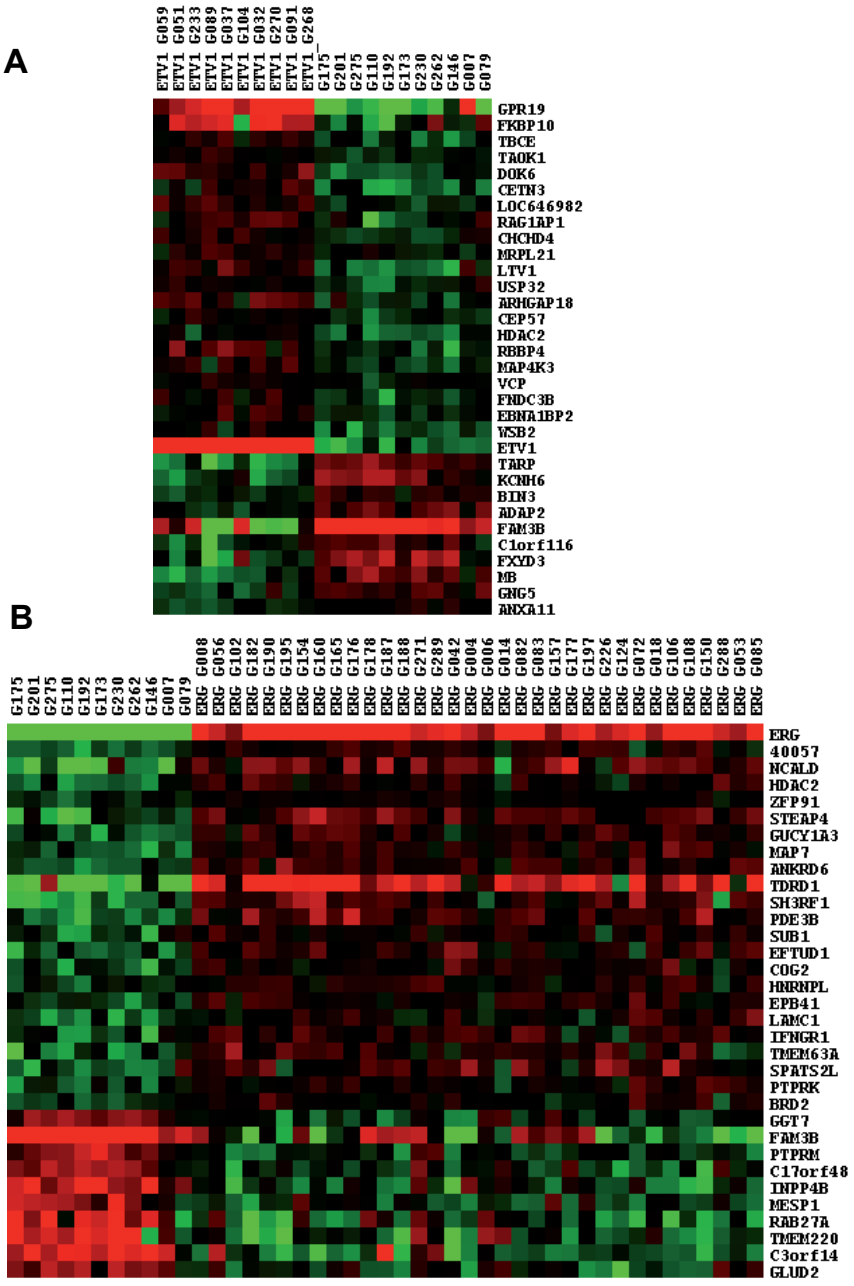


Figure 5. ETV1 and ERG associated genes in clinical prostate cancer samples. A) A heat-map of the highest positively and negatively regulated genes associated with ETV1 overexpression (G091, G032 and G104 overexpressing dETV1) as measured by SAM of gene expression array data of primary prostate cancer samples. **B)** A heat-map of the highest positively and negatively regulated genes associated with ERG overexpression as measured by SAM of expression array data of primary prostate cancer samples.

prostate samples but its expression is very variable. Only one *ETV1* positive prostate cancer showed an outlier expression of *CRISP3.*

To investigate possible similarities between *ETV1* and *ERG* regulated genes in clinical samples we performed a SAM experiment on *ERG* overexpressing samples with the same ETS negative group as used above (Figure 5B and Suppl. Table S8). We found induction of *HDAC2*, *LAMC1* and *PRDX5* correlated with both *ERG* and *ETV1* overexpressing cancers. *FAM3B* is the highest of ten inverse-correlated genes affected by both *ETV1* and *ERG* overexpression (for complete lists of the overlapping genes see Suppl. Table S9). However, further no obvious similarly regulated genes were detected.

Discussion

In this chapter we presented preliminary data on the function of ETV1 and N-truncated dETV1 in prostate cancer. PNT2C2 normal prostate cells overexpressing ETV1 or dETV1 were generated, and genes induced were identified and compared under different growth conditions. Moreover, genes associated with ETV1 expression in clinical prostate cancer were identified. Several major conclusions, counting for most of our results, can be drawn from the gene expression patterns.

PNT2C2 cell model systems: (I) Genes associated with ETV1- and dETV1- overexpression are largely overlapping. (II) Growth conditions affect expression of ETV1- and dETV1-associated genes.

Clinical prostate cancer: (III) Genes associated with ETV1- and dETV1-overexpression in clinical prostate cancer samples are not identical to those in the PNT2C2-ETV1 and dETV1-PNT2C2 model systems. (IV) Genes associated with ETV1-overexpression are different from genes associated with ERG-overexpression in clinical prostate cancer.

According to gene expression array data, 50-60% of genes associated with ETV1 and dETV1 in PNT2C2 cells are identical, but there are exceptions. One of our experiments is shown in Figure 1, but largely similar results were obtained in a duplicate experiment, with independently generated PNT2C2-ETV1 and PNT2C2-dETV1 cells (data not shown). *BCL2A1* shows high overexpression in both PNT2C2-ETV1 and PNT2C2-dETV1 cells. BCL2A1 is a member of the BCL-2 protein family as regulation of apoptosis and known as a direct NF-kappaB target gene (13). *BCL2A1* is overexpressed in several cancer types, but so far a role in prostate cancer is unknown. *PAPPA* and *SRGN* are examples of genes that are more prominently induced by ETV1 than by dETV1 overexpression. PAPPA is a metalloproteinase that can cleave IGFBPs, SRGN is a proteoglycan. So far, no function of these genes in prostate cancer has been published, but *PAPPA* can be overexpressed in lung cancer (14).

Promoter activation studies (Figures 2 and 3) indicated that both *PAPPA* and *BCL2A1* can be regulated by ETV1 and dETV1 in transient transfection experiments. However, in both cases ETV1 is much more active than dETV1. This fits with the expression of endogenous *PAPPA*, but not with expression of *BCL2A1* mRNA (Figure 1 and Table 1). One explanation might be that the acidic transactivation region TAD that is present in ETV1 and absent in dETV1 is less important in regulation of expression of endogenous genes than in promoter analysis. It is also possible that the promoter studies were incomplete, and that important enhancer regions were missing. ChipSeq analysis might be helpful in clarifying this point. Alternatively, higher dETV1 expression than ETV1 expression in

stable cell lines (Figure 1) can be caused by increased stability of dETV1, because of the deletion of important binding sites for the E3 ubiquitin ligase COP1, that triggers proteasomal degradation (15). The absence of a correlation between the presence of canonical ETV1 binding sites and activity of the various promoter fragments indicates a complex mechanism of ETV1 activity, which might include interaction with weak binding sites or activation by protein-protein interaction.

PNT2C2-GFP, PNT2C2-ETV1 and PNT2C2-dETV1 cells all grow as tumors if subcutanuously transplanted on nude mice. There was no substantial difference between growth of PNT2C2-ETV1 and PNT2C2-dETV1 tumors, but both grew faster than the control PNT2C2-GFP cells (Figure 4). ETV1 seems more active than dETV1 under *in vivo* growth conditions. As indicated in the Results section, there are genes that are similarly induced *in vitro* in 2D-culture and *in vivo*, but there are also interesting differences. MMPs, including *MMP1*, are earlier described targets of ETV1, and overexpressed in both breast and prostate cancer (4,16). *PLAUR* was correlated with invasiveness and progression of prostate cancer (17). *CRISP3* is known as a prostate tumor marker. So, data collected suggest that some of the genes found to be associated with ETV1 or dETV1 overexpression and that can play a role in tumor progression can be detected in the tumor transplants and not in cells grown *in vitro*.

So far, we were not able yet to answer satisfactory the question of molecular and biological differences between ETV1 and dETV1. An obvious aspect that can affect our results are the properties of the parental cell line, PNT2C2 cells, in which ETV1 and dETV1 were overexpressed. We realize that the PNT2C2 cell might not be an ideal cell line for overexpression studies and does not fully reflect a normal prostate cell that can function as a model for tumor initiation or model of prostate cancer progression. However, none of other described “normal” prostate cells has such properties.

Another approach to study the role of ETV1 in prostate cancer is by downregulation of ETV1 in an ETV1-positive prostate cancer cell. Two prostate cancer cell lines that overexpress ETV1 are available, LNCaP cells and MDA PCa 2B cells. Although we were able to downregulate ETV1 in LNCaP cells, the effects on gene expression were very limited (data not shown). Probably, because ETV1 expression in LNCaP cells is variable and rather low, it has to be questioned whether ETV1 in LNCaP cells still is of importance for processes that regulate cell growth and other functions essential for oncogenesis. MDA PCa 2B cells were very difficult to manipulate, and the appropriate siRNA and shRNA experiments to decrease ETV1 expression, are still ongoing.

An obvious complementary approach that we took to address the function of ETV1 in prostate cancer is to study gene expression in clinical prostate cancer samples. We

focused on RNAs from the 10 ETV1-positive primary prostate cancers that were present in our collection. As described in the Results section, the genes that showed the strongest association with ETV1 overexpression were not expressed in the PNT2C2 cell line and *vice versa*. Obviously, the tumors are rather heterogeneous and the number of clinical samples studied is quite small. However, in our opinion, our preliminary results already highlight the difficulty to mimic in an experimental setup clinical prostate cancer. Important aspects in this regard are no doubt not only the properties of tumor initiating cells, but also the different microenvironment and the differences in time frame of tumor development and progression following ETV1 overexpression, which spans many years in the clinical situation.

Another remarkable result from our study of ETV1 overexpression in clinical prostate cancer concerned the difference in gene expression profiles of ETV1-positive tumors versus ERG-positive tumors. Although ETV1 and ERG both are members of the ETS family of transcription factors, and ERG and ETV1 overexpression in prostate cancers are mutually exclusive, there was hardly any overlap in the associated gene expression patterns (see Figure 5). Recently, in a different patient group some overlap was found, however, this might also be explained by multifocality of heterogeneous tumors (12). In fact; we know that the strongest ERG-associated upregulated gene, *TDRD1*, is one of the genes that is downregulated by *ETV1* in clinical prostate cancer (Boormans et al, in press and unpublished data). A recent review (18) describes in more detail differences between ETV1 and ERG function as transcription factor. However, these differences are difficult to extrapolate to our findings. Many of the genes detected in this study will not be directly, but indirectly regulated by ERG or ETV1.

ERG is already overexpressed in PIN and seems important in both early and late prostate cancer (19). So far, ETV1 is not detected in PIN and together with data presented in this chapter it might indicate that ETV1 has different functions from ERG and is important only in later stages of prostate cancer. Obviously, this hypothesis needs further experimental validation.

References

1. Tomlins SA, Rhodes DR, Perner S et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644-8.
2. Hermans KG, van der Korput HA, van Marion R et al. Truncated ETV1, fused to novel tissue-specific genes, and full-length ETV1 in prostate cancer. Clinical Res 2008;68:7541-9.
3. Gasi D, van der Korput HA, Douben HC et al. Overexpression of full-length ETV1 transcripts in clinical prostate cancer due to gene translocation. PLoS One 2011;26:6:e16332.
4. Tomlins SA, Laxman B, Dhanasekaran SM et al. Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature 2007;448:595-9.
5. Attard G, Clark J, Ambroisine L et al. Heterogeneity and clinical significance of ETV1 translocations in human prostate cancer. Br J Cancer 2008;99:314-20.
6. Clark JP and Cooper CS. ETS gene fusions in prostate cancer. Nat Rev Urol 2009;6:429-39.
7. Hollenhorst PC, McIntiosh LP and Graves BJ. Genomic and biochemical insights into the specificity of ETS transcription factors. Annu Rev Biochem 2011;80:437-71.
8. Seth A and Watson DK. ETS transcription factors and their emerging roles in human cancer. Eur J Cancer 2005;42:2462-78.
9. Rubin MA, Maher CA and Chinnaiyan AM. Common gene rearrangements in prostate cancer. J Clin Oncol 2011;27:3659-68.
10. Zong Y, Xin L, Goldstein AS, Lawson DA, Tetell MA and Witte ON. ETS family transcription factors collaborate with alternative signaling pathways to induce carcinoma from adult murine prostate cells. Proc Natl Acad Sci U S A 2009;106:12465-70.
11. Shin S, Kin TD, van Deursen JM et al. Induction of prostatic intraepithelial neoplasia and modulation of androgen receptor by ETS variant 1/ETS-related protein 81. Cancer Res 2009;68:8102-10.
12. Paulo P, Ribeiro FR, Santos J et al. Molecular Subtyping of Primary Prostate Cancer Reveals Specific and Shared Target Genes of Different ETS Rearrangements. Neoplasia 2012;14:600-11.
13. Vogler M. BCL2A1: the underdog in the BCL2 family. Cell Death Differ 2012;19:67-74.
14. Pan H, Hanada S, Zhao J, Mao L and Ma MZ. Protein secretion is required for pregnancy-associated plasma protein-A to promote lung cancer growth in vivo. PLoS One 2012;7:e48799.
15. Vitari AC, Leong KG, Newton K et al. COP1 is a tumor suppressor that causes degradation of ETS-transcription factors. Nature 2011;474:403-6.
16. Bosc DG, Goueli BS, Janknecht R. HER2/Neu-mediated activation of the ETS transcription factor ER81 and its target gene MMP-1. Oncogene 2001;20:6215-24.
17. Gavrilov D, Kenzior O, Evans M, Calaluce R and Folk WR. Expression of urokinase plasminogen activator and receptor in conjunction with the ets family and AP-1 complex transcription factors in high grade prostate cancers. Eur J Cancer 2001;37:1033-40.
18. Oh S, Sook S and Janknecht J. ETV1, 4 and 5: An oncogenic subfamily of ETS transcription factors. Biochim Biophys Acta 2012;1826:1-12.
19. Clark J, Attard G, Jhavar S et al. Complex patterns of ETS gene alteration arise during cancer development in the human prostate. Oncogene 2008;27:1993-2000.

Supplementary Table S1 Genes that are specifically induced in PNT2C2-ETV1 cells grown in 2D**

GeneID	ETV1	ETV1/dETV1	GeneID	ETV1	ETV1/dETV1
LOC100302650	6.2	4.3	OGFRL1	2.2	1.2
PAPPA *	4.2	2.7	FAM84B	2.2	1.2
KIAA1199	4.0	4.6	TMPRSS2	2.2	1.3
TMPRSS2	3.7	2.2	B3GNT7 *	2.2	1.1
GPR116 *	3.7	2.0	PCSK5 *	2.1	1.2
ADAMTS9	3.4	2.1	KYNU	2.1	1.5
H19	3.4	2.4	ANKRD29	2.1	1.1
PDE4B	3.1	2.1	OGFRL1	2.1	1.1
CYP24A1	3.1	2.8	SCARA5	2.1	2.9
BMP6	3.0	2.1	FGA	2.1	3.1
H19	3.0	1.7	GLRX	2.1	1.2
SLC4A7 *	3.0	1.5	ALS2	2.1	1.6
IL13RA2	3.0	2.3	FNDC3B	2.1	1.7
IL15	2.7	1.4	RFK *	2.1	1.1
SLC6A14	2.7	1.4	EFNB2	2.1	1.7
FNDC3B	2.6	1.6	COL21A1	2.1	1.6
C21orf84	2.6	1.6	CFH *	2.1	1.6
PLAT	2.6	1.5	CORO2B	2.1	1.1
S100A8	2.6	2.6	IGFBP5	2.1	2.0
MTRR	2.5	1.5	TFPI	2.1	1.0
ESM1	2.5	1.6	NLRP11	2.0	1.2
PDLIM5 *	2.4	1.5	MTRR	2.0	1.2
JAG1 *	2.4	1.3	GRAMD1B	2.0	1.0
KCTD12	2.4	1.5	MPZL2 *	2.0	1.1
FGA	2.4	2.4	CEP70	2.0	1.4
CFI	2.4	1.4	ZNF75D	2.0	1.3
KCTD12	2.4	1.4	PTX3	2.0	1.1
AOX1	2.4	1.9	LPGAT1	2.0	1.7
C5	2.4	1.2			
ZBTB47	2.4	1.5			
CHRNA9	2.3	2.1			
RGNEF *	2.3	1.3			
PLA2G4A	2.3	1.3			
TFPI2 *	2.3	1.2			
COL13A1 *	2.3	1.3			
LOC645431	2.3	1.5			
SCG5	2.3	2.5			
OGFRL1	2.3	1.2			
ELMOD1	2.3	1.2			
PDE4B	2.2	2.7			
CXCL2	2.2	1.2			
EYS	2.2	2.5			
FGG	2.2	2.2			
CTH	2.2	1.2			

* genes induced in both PNT2C2-ETV1 and PNT2C2-dETV1 cells

**2-fold or higher induced compared to PNT2C2-GFP

Supplementary Table S2 Genes that are specifically induced in PNT2C2-dETV1 cells grown in 2D**

GeneID	dETV1	dETV1/ETV1	GeneID	dETV1	dETV1/ETV1
CEACAM1	9.4	5.3	PLCXD2	2.3	1.2
ANKRD1	3.8	2.2	SP3	2.2	1.6
F2RL2	3.6	2.0	PRR5L	2.2	1.8
TMC8	3.3	1.7	C2orf63	2.2	1.2
PRKCA	3.1	2.1	HRCT1	2.2	1.4
PRKCA	3.1	1.7	LPIN2	2.2	1.2
SAMSN1	3.1	1.7	SH3BP5	2.2	1.1
TMEM40	3.0	1.9	ARRDC2	2.2	1.4
B4GALT6 *	2.9	1.6	CHSY3	2.2	1.5
ERAP1	2.8	1.5	TMEM40	2.2	2.0
SMAD1 *	2.8	1.4	DPP4	2.2	1.3
NCRNA00173	2.7	1.6	KLHDC1	2.2	1.2
ABHD5 *	2.6	1.5	CCPG1 *	2.2	1.2
FRMD3 *	2.6	1.6	FCGR2A	2.2	1.2
ZHX3	2.6	1.9	GBP1	2.2	1.4
STIM2	2.6	1.3	RUNX2	2.2	1.8
ADAMTSL3	2.6	1.5	COL5A2	2.2	1.4
HOXB3	2.6	1.6	NOG	2.1	1.2
PDE2A	2.6	1.9	GGCX	2.1	1.5
IGFBP6	2.6	4.2	CARD6	2.1	1.6
CD59 *	2.6	1.5	WWC3	2.1	1.2
SH3PXD2A	2.6	1.7	GUCY1B3	2.1	1.5
TSPAN14	2.5	1.8	NEBL	2.1	1.9
MAML2	2.5	2.1	TSPAN14	2.1	1.4
DENND3	2.5	1.4	TGFBR3	2.1	1.4
EGLN3	2.5	1.6	LPIN2	2.1	1.1
ROBO4	2.5	2.6	SPTB	2.1	1.5
NR2F1	2.5	2.3	EHBP1L1	2.1	1.4
RFESD	2.4	1.5	STIM2	2.1	1.2
FLJ42709	2.4	1.9	DAB2	2.1	1.2
PRR5L	2.4	1.4	SERTAD4	2.1	2.0
DAB2	2.4	1.7	NR0B1	2.1	1.7
FAM65C	2.4	2.4	EIF3B	2.0	1.3
ABCG1	2.4	1.4	LAMA3	2.0	1.1
CLEC2B	2.4	1.2	LOC344887	2.0	1.8
SH3BP5	2.4	1.2	DPP4	2.0	1.6
S100A13	2.4	1.4	C15orf37	2.0	1.5
ARHGAP23	2.4	1.9	PLD1	2.0	1.6
TENC1	2.3	1.8	KAZALD1	2.0	1.3
CTSO	2.3	1.3	ST8SIA4 *	2.0	1.4
LOC439938	2.3	2.1	CCPG1 *	2.0	1.2
FAM101B *	2.3	1.2	KIF9	2.0	1.5
C4BPA	2.3	1.4	C10orf116	2.0	1.6
GNA14 *	2.3	1.3	LOC285943	2.0	1.4
CMTM7	2.3	1.4	LIMK2	2.0	1.2
EP400	2.3	1.4	ULK4	2.0	1.7
PCSK5 *	2.3	1.2	PDXDC1	2.0	1.7
ZEB1	2.3	1.2	CGNL1	2.0	1.5

* genes induced in both PNT2C2-ETV1 and PNT2C2-dETV1 cells

**2-fold or higher induced compared to PNT2C2-GFP

Supplementary Table S3 2-fold or higher induced genes in both PNT2C2-ETV1 and PNT2C2-dETV1 compared to PNT2C2-GFP grown in 2D

GeneID	ETV1	dETV1	GeneID	ETV1	dETV1	GeneID	ETV1	dETV1
ETV1	53.5	90.6	CD55	3.7	3.9	CFHR4	2.6	2.7
BCL2A1	20.0	30.8	TMEM155	3.7	4.5	PRKAB1	2.6	3.3
PLXNA2	17.8	19.3	CREM	3.6	3.5	G0S2	2.6	2.0
GPR116 *	15.2	3.1	RGNEF	3.6	2.8	ANTXR2	2.6	3.3
ETV1	14.8	22.0	ANKRD55 ^	3.6	8.4	MAML3	2.6	2.3
ETV1	14.7	27.3	MPZL2	3.6	2.4	PCSK6	2.6	2.5
PAPPA *	12.4	4.2	INHBB	3.5	2.0	ITGB3	2.6	2.3
THBD	10.0	6.3	SLCO4A1	3.5	3.8	ADAM23	2.4	2.8
LTBP1	9.9	5.1	PLAUR	3.5	2.9	B3GNT7	2.4	2.0
PAPPA *	9.1	2.9	CREM	3.5	3.4	PDLIM5	2.4	2.1
PAPPA *	9.0	2.8	SFTPB	3.5	2.6	SMAGP	2.4	3.4
LTBP1	8.9	7.5	ADAM23	3.5	4.3	RFK	2.4	2.5
THBD	8.7	6.5	ENC1	3.4	2.1	PCSK5	2.4	2.4
ANKRD22	7.9	7.2	TMEM150C	3.4	3.5	ABCA1	2.3	2.6
THBD	7.7	5.6	CD55	3.4	3.7	GALNT6	2.3	2.6
PELO	7.5	6.9	ANTXR2	3.4	5.1	ABHD5	2.3	2.3
SERPINA3	7.0	12.4	CEACAM6	3.4	3.9	RTN1	2.3	2.4
COL13A1	6.8	6.2	PLAUR	3.3	2.7	LAT2	2.3	2.3
SRGN *	6.6	3.3	LY96 ^	3.3	8.0	CCPG1	2.3	2.2
SFTPB	5.7	4.4	SRGN	3.3	2.1	GALNT6	2.3	2.8
ST3GAL6	5.6	4.5	JAG1	3.3	2.5	HPGD	2.3	2.8
PAPPA *	5.5	3.1	JAG1	3.3	2.2	SLC26A2	2.3	2.0
ST3GAL6	5.4	3.3	CLDN11	3.3	3.8	ABHD5	2.2	2.4
S100A9 *	5.3	2.3	RGNEF	3.3	3.1	CTGF	2.2	2.7
C8orf4 *	5.3	2.1	ENC1	3.2	2.1	SLC26A2	2.2	2.2
PXDN	5.2	4.7	DCUN1D3	3.2	3.5	SMAD1	2.2	2.7
SFTPB	5.2	4.0	PLAUR	3.1	2.5	ST8SIA4	2.2	3.0
PELO	5.1	4.8	JAG1	3.1	2.2	PDLIM5	2.2	2.3
ITGA10 ^	5.0	10.4	HS3ST1	3.1	3.7	CD59	2.2	2.8
PXDN	5.0	4.6	LYRM1	3.1	3.2	NR5A2	2.2	2.4
RDH10 ^	4.9	9.8	GNA14	3.0	3.1	PTPRO	2.2	2.6
RDH10 ^	4.7	9.3	CADPS2	3.0	3.5	SNRK	2.2	3.7
CA2	4.7	3.1	ABCA1	3.0	3.7	AGXT2L1	2.1	2.0
RDH10	4.6	6.9	MPZL2	3.0	2.5	PDLIM5	2.1	2.1
CREM	4.5	3.8	RGNEF	2.9	2.8	B4GALT6	2.1	2.8
PELO	4.4	4.6	CFH /// CFHR1	2.9	2.2	CD59	2.1	2.5
CD59	4.4	5.4	ANKRD22	2.9	2.1	F2RL1	2.1	2.2
FCRLB	4.3	3.1	BATF3	2.9	4.5	ST3GAL5	2.1	2.3
PLXNA2	4.3	4.6	FZD4 ^	2.9	13.3	PLN	2.1	2.4
TMEM150C	4.2	3.6	COL13A1	2.9	3.2	TMC5	2.1	2.2
GJA1	4.1	3.3	TFPI2	2.9	2.1	HOXB8	2.1	2.7
KLF9	4.0	5.0	LHFP	2.8	2.5	FAM101B	2.1	2.5
ITGB3	4.0	3.8	DOCK4	2.8	2.4	IPCEF1	2.1	2.7
SLC4A7	4.0	3.3	SLC4A7	2.8	2.1	PRKAB1	2.0	2.8
PCDH1	3.8	3.7	LHFP	2.8	2.2	KCNQ5	2.0	2.0
CREM	3.7	3.8	CACHD1	2.8	3.2	CD59	2.0	2.5
CD55	3.7	3.2	FAM43A	2.7	3.5	CNTNAP3B	2.0	2.2
CEACAM6	3.7	4.4	ARSI	2.7	2.0	TNFRSF10D	2.0	2.0
DPYD	3.7	2.4	FRMD3	2.7	4.1			

* genes specifically induced in PNT2C2-ETV1 cells

^ genes specifically induced in PNT2C2-dETV1 cells

Supplementary Table S4 Genes that are specifically induced in PNT2C2-ETV1 cells grown *in vivo***

GeneID	ETV1	ETV1/ dETV1	GeneID	ETV1	ETV1/ dETV1	GeneID	ETV1	ETV1/ dETV1
RTN1	12.8	11.1	BMP5	3.4	2.5	CREM	2.6	1.6
GPR116	8.2	7.3	HP	3.4	2.8	OGFRL1	2.6	2.2
KISS1	7.7	7.7	PHACTR3	3.3	2.5	SYNJ1	2.6	2.1
PLUNC	7.4	11.0	PLN	3.3	3.5	SLC25A45	2.6	1.8
PAPPA	7.2	3.6	PITPNC1 *	3.3	1.8	SLC26A2	2.6	1.7
PRR16	7.0	8.0	IL1RAP	3.3	2.1	SIPA1L2	2.6	2.7
PELO	6.3	5.9	CREM	3.3	2.0	CD36	2.6	2.4
RTN1	6.1	5.3	B3GNT7	3.3	3.9	CREM	2.6	1.8
LOC100302650	5.8	4.5	PLN	3.3	4.6	PAG1	2.6	2.6
ZEB1	5.6	2.9	RAB27A	3.3	1.9	ELL2	2.5	2.0
ALB	5.6	5.2	FAM20A	3.2	3.4	ADSSL1	2.5	1.8
OGFRL1	5.5	3.8	GPR37	3.2	2.7	KLHL5	2.5	3.2
PAPPA	5.5	3.1	KLF9	3.2	2.1	TNFSF10	2.5	1.6
PLXNA2	5.4	2.8	SHISA3	3.2	3.4	PRTG	2.5	2.0
PLN	5.2	7.6	INHBB	3.2	5.6	IKBIP	2.5	1.5
MMP10	5.0	4.6	A2M	3.2	2.9	ADAMTS9	2.5	1.7
AG2	5.0	3.8	TNIP3	3.1	2.5	SLC39A8	2.5	1.5
TREM1	4.9	2.7	TMEM150C	3.1	2.6	CPVL	2.5	1.9
RIMKLB	4.9	2.6	DMBT1	3.1	3.8	CP	2.5	5.0
RGNEF	4.9	3.2	PDE4B	3.1	2.8	SCG5	2.5	2.6
PLXNA2	4.9	3.7	RORA	3.0	2.6	TNFSF10	2.5	1.8
DUSP6	4.9	2.6	TMEM156 *	3.0	1.5	CHSY3	2.5	2.2
PAPPA	4.7	2.6	B3GNT7	3.0	3.6	FAM46C	2.5	1.9
DUSP6	4.6	2.8	LIX1	3.0	2.6	IKBIP	2.5	1.7
MFSD4	4.6	5.6	ZBTB10	3.0	4.0	CREM	2.5	1.8
CA2	4.6	2.4	CP	3.0	5.0	SLC26A2	2.5	1.9
PDE4B	4.6	6.4	GPR116	3.0	2.8	KLHL5	2.5	2.7
ANTXR1	4.4	2.4	PELO	3.0	3.0	ANKRD37	2.5	1.8
ITGB3	4.4	3.2	MMP13	3.0	6.1	PNP	2.4	2.5
OGFRL1	4.3	3.7	WSB1	2.9	3.5	RGNEF	2.4	1.9
RECK	4.3	4.2	KLF9	2.9	2.2	DOCK4	2.4	1.8
PLN	4.2	8.2	ST3GAL6	2.9	1.7	CD36	2.4	2.4
DUSP6	4.2	2.5	CHST2	2.9	2.7	DNTTIP1	2.4	1.5
CTGF	4.1	6.2	CP	2.9	4.7	SNX21	2.4	2.4
ELMOD1	3.9	2.8	CD302	2.9	2.8	SLC4A7	2.4	2.3
LTBP1	3.9	3.0	RGNEF	2.9	2.1	IFITM2	2.4	2.8
CD36	3.9	4.4	RORA	2.8	2.2	PRTG	2.4	2.0
LOC375295	3.9	5.1	ANGPTL4 *	2.8	1.8	SLC7A11	2.4	1.3
TMEM150C	3.8	2.8	CD274	2.8	2.2	IRS1	2.4	2.0
ST3GAL5	3.7	2.3	PTPRB	2.8	1.4	LRP2	2.4	2.2
LTBP1	3.6	2.3	LOC344595	2.8	2.1	ZBTB10	2.4	2.4
CXCL2	3.6	3.3	GPR37	2.7	2.4	PTGS2	2.4	1.3
ST3GAL6	3.6	1.8	ZIC2	2.7	1.6	KIF13A	2.4	1.7
ANTXR1	3.6	3.0	ANTXR1	2.7	2.7	RORA	2.4	2.3
ESR1	3.6	3.8	PRDM1	2.7	2.3	ATP6V0A4	2.4	1.2
WSB1	3.5	4.7	LOX	2.7	5.5	SIPA1L2	2.4	2.4
HP /// HPR	3.5	3.0	MFSD4	2.7	2.4	MFSD4	2.4	2.4
C8orf4	3.4	4.3	IGFBP5	2.7	6.2	GPR126 *	2.4	1.3
IRS1	3.4	2.4	SAMSN1	2.7	1.9	PDE8A	2.4	1.4

Supplementary Table S4 (*continued*)

GeneID	ETV1	ETV1/ dETV1	GeneID	ETV1	ETV1/ dETV1	GeneID	ETV1	ETV1/ dETV1
LIFR	2.4	1.9	NPAS2	2.2	1.2	CXorf24	2.0	1.2
SLC39A8	2.4	1.7	TMTC2	2.2	3.1	FHL1	2.0	2.0
NEK7	2.3	1.9	ISG20	2.2	2.1	CXCR7	2.0	3.5
IL18R1	2.3	1.5	UMODL1	2.2	2.9	CFH	2.0	2.7
DNTTIP1	2.3	1.5	KCNMB4	2.2	1.2	F2R	2.0	1.3
TNFSF10	2.3	1.7	CD46	2.1	1.2	CCDC68	2.0	1.2
LIX1L	2.3	1.2	PCDH1	2.1	1.9	AOX1	2.0	2.1
ZBTB10	2.3	2.9	JAM3	2.1	2.6	C9orf150	2.0	3.2
STEAP4	2.3	2.8	ADCY9	2.1	3.3	HIP1	2.0	2.0
NDE1	2.3	1.3	FNDC3B	2.1	1.7	GNA14 *	2.0	1.2
COL12A1	2.3	2.6	LIX1L	2.1	1.1	HAS3	2.0	1.3
SLC26A2	2.3	1.5	GLRX	2.1	1.2	TMEM30A	2.0	1.4
GPR109B	2.3	3.2	KLF9	2.1	2.3	FAM102B	2.0	1.8
WSB1	2.3	2.0	ATP11A	2.1	2.3	FAM84B	2.0	1.6
FAM43A	2.3	1.5	GALNT6	2.1	1.3	LOC401022	2.0	1.7
PDE8A	2.3	1.3	SEMA4B	2.1	1.9			
CTPS	2.3	1.4	ABCC6 /// LOC100292715	2.1	2.3			
FAM20A	2.3	2.4	ENC1	2.1	2.1			
RAB27A	2.3	1.2	FAM13A	2.1	2.9			
GPR133	2.3	2.3	SLC4A7	2.1	2.1			
NPTX1	2.3	2.2	DAB2	2.1	3.8			
SUCNR1	2.3	2.1	PAG1	2.1	2.2			
KLHDC8A	2.3	2.4	GNG11	2.1	1.2			
KIF13A	2.3	1.2	LOC344595	2.1	1.4			
WFDC2	2.3	3.8	COL12A1	2.1	2.1			
ZBTB10	2.3	2.4	CEBPD	2.1	2.9			
CD46	2.3	1.2	KCNQ2	2.1	2.3			
NOX1	2.2	2.0	NPAS2	2.1	1.1			
PRKCA	2.2	1.5	PPBP	2.1	1.6			
FNDC3B	2.2	1.7	CPD	2.1	1.4			
SFTPA2	2.2	2.0	WSB1	2.1	2.1			
LOX	2.2	2.6	FZD4	2.1	1.3			
CP	2.2	3.9	ANO1	2.1	3.2			
MYBL1	2.2	2.8	ISG20	2.1	2.0			
SMOC1	2.2	1.2	ATP11C	2.1	1.3			
SLC7A11	2.2	1.2	LPAR3	2.1	2.3			
CALCR	2.2	1.3	WSB1	2.1	2.0			
COL4A5	2.2	1.2	PDE8A	2.1	1.2			
ZBTB10	2.2	2.1	KIF26A	2.1	1.4			
SEZ6L2	2.2	1.6	IKBIP	2.1	1.3			
SPIRE1	2.2	1.5	CD59	2.1	1.6			
MAML3	2.2	1.9	HIP1	2.1	2.3			
RECK	2.2	2.2	KIAA0090	2.1	1.3			
PDLIM5 *	2.2	1.2	FAM20A	2.1	2.0			
RAB27A	2.2	1.2	IL1RAP	2.1	2.1			
SLC34A2	2.2	1.8	PAG1	2.1	2.4			
FAM13A	2.2	3.0	FNDC3B	2.0	1.4			
FAM13A	2.2	3.0	IL15	2.0	1.2			
KIF26A	2.2	1.5	PTPRJ	2.0	1.5			

* genes induced in both PNT2C2-ETV1 and PNT2C2-dETV1 cells

**2-fold or higher induced compared to PNT2C2-GFP

Supplementary Table S5 Genes that are specifically induced in PNT2C2-dETV1 cells grown *in vivo***

GeneID	dETV1	dETV1/ ETV1	GeneID	dETV1	dETV1/ ETV1	GeneID	dETV1	dETV1/ ETV1
AREG	7.0	7.1	GCNT3	2.4	1.6	KYNU	2.1	2.1
CD38	6.7	3.6	ASPH *	2.4	1.4	TNNC2	2.1	1.7
SERPINA1	6.5	3.3	GAL	2.4	2.1	AGT	2.1	1.2
CRYAB	6.2	14.9	DUSP5	2.4	1.6	KLRC3	2.0	2.1
NT5E	5.2	3.6	AGR3	2.4	1.4	MERTK	2.0	1.6
ABRA	5.1	5.4	CTSL1	2.3	1.4	CAMK2N1	2.0	2.4
MYLPF	4.9	4.6	PSG3	2.3	1.8	RHOF	2.0	1.6
NT5E	4.9	3.3	LRRC19	2.3	2.1	KYNU	2.0	2.0
PHLDA1	4.6	3.0	LOC100289208 /// PHLDA1	2.3	2.4	PECAM1 *	2.0	1.2
FHL1	4.6	8.5	TM4SF1	2.3	1.3	AQP3	2.0	1.5
PDLIM5	4.5	4.4	ASPH *	2.3	1.3	CRIP1	2.0	2.5
NT5E	4.4	4.5	GPR87	2.3	1.6	LOC100289632	2.0	1.7
TPM1	4.4	4.3	FOXO1	2.3	1.3	FST	2.0	1.8
THBS4	4.3	6.3	CENPV	2.3	1.3	PHLDA1	2.0	2.3
PHLDA1	3.9	3.4	CENPV	2.3	1.2	SLC16A1	2.0	1.2
PHLDA1	3.6	3.2	LBP	2.3	1.2	SLC16A1	2.0	1.1
PHLDA1	3.6	2.9	ATP6V0D2	2.3	2.7			
TFF1	3.4	1.8	AREGB	2.2	1.9			
ASPH *	3.4	1.7	SLC16A1	2.2	1.2			
PKIB	3.3	1.8	DIRAS3	2.2	3.5			
NOSTRIN	3.3	2.2	MAN1A1	2.2	1.3			
EREG	3.3	6.7	PTPRM	2.2	1.6			
FGFBP1	3.2	2.8	LRP8	2.2	1.6			
FHL1	3.2	3.3	PDLIM7	2.2	2.8			
AQP3	3.1	2.0	SEMA4G	2.2	2.1			
CHI3L1	3.0	3.8	DCLK1	2.2	2.2			
AQP3	3.0	1.9	TM4SF1	2.2	1.3			
SERPIND1	3.0	2.6	LRP8	2.2	1.7			
KRT13	2.9	5.2	LONRF2	2.2	2.0			
KLRC1 /// KLRC2	2.9	2.7	IL27RA	2.2	1.1			
GKN2	2.9	1.8	MET	2.2	3.9			
ACTA1	2.9	2.0	BZW1 /// BZW1L1	2.2	1.6			
LOC1518	2.8	1.6	RHOF	2.2	1.5			
PKIB	2.8	2.0	PVALB	2.2	1.8			
A2BP1	2.8	3.6	ITPR1	2.1	1.8			
FHL1	2.7	2.8	FAS	2.1	1.7			
CCL14-CCL15 /// CCL15	2.7	2.2	FAS	2.1	1.9			
ANXA10	2.7	2.1	RAMP1	2.1	1.3			
MAN1A1	2.6	1.6	PPTC7	2.1	1.3			
LOC144481	2.6	1.6	AGPAT9	2.1	2.5			
PCSK9	2.5	1.6	LOC645722	2.1	1.2			
COL4A6	2.5	1.9	ALDH3A1	2.1	2.2			
TIMP1	2.5	1.7	CRABP2	2.1	2.6			
SMOC1	2.5	1.3	TM4SF1	2.1	1.5			

* genes induced in both PNT2C2-ETV1 and PNT2C2-dETV1 cells

**2-fold or higher induced compared to PNT2C2-GFP

Supplementary Table S6 **2-fold or higher induced genes in both PNT2C2-ETV1 and PNT2C2-dETV1 compared to PNT2C2-GFP grown *in vivo***

GeneID	ETV1	dETV1	GeneID	ETV1	dETV1
ETV1 *	67.8	24.9	COL13A1	3.2	2.2
ETV1 *	58.3	20.9	SLC2A13	3.2	2.3
PAEP *	30.5	4.4	UBE2QL1	3.1	2.1
SLCO4A1	22.6	13.3	GNA14	3.1	2.2
AQP1 *	20.3	4.4	RDH10	3.1	3.0
ETV1 *	12.5	5.0	PDLIM5	3.1	2.5
CRISP3 *	11.1	2.5	PDK4	3.0	2.8
PLAUR *	10.8	4.4	NR4A2	3.0	3.2
PLAUR *	10.4	4.2	KIAA1199	3.0	3.2
ENPEP *	10.4	2.9	RDH10	3.0	3.0
ENPEP *	10.4	3.0	TFPI2	3.0	3.9
THBD *	10.4	4.2	NR4A2	3.0	3.3
ANKRD1 *	10.0	2.3	SOCS2	2.9	5.4
AQP1 *	9.5	2.2	PDLIM5	2.9	2.6
THBD *	9.5	3.8	PECAM1	2.9	3.5
PAPPA *	8.9	2.3	GPR126	2.9	2.3
THBD *	8.6	3.1	TFPI	2.8	2.5
CXCR4	8.3	6.2	NR4A2	2.8	2.8
MMP1	8.3	11.8	RDH10	2.8	2.8
IL13RA2 *	7.8	2.4	SERPINA1 ^	2.8	6.4
RIMKLB *	7.8	3.6	FRMD3	2.8	3.3
SLC6A14	7.5	6.0	HYAL1	2.8	2.8
RIMKLB *	7.5	3.6	SOCS2	2.8	4.5
S100A8	7.4	4.1	C21orf84	2.7	2.3
S100A9	7.4	3.8	PDLIM5	2.7	2.4
RIMKLB *	6.8	3.2	SFRP1	2.7	4.9
PAPPA *	6.8	2.3	SLC2A13	2.7	2.0
GJA1	6.7	3.8	MPZL2	2.7	2.4
SRGN	6.5	4.6	TFPI2	2.7	3.3
TMEM156 *	6.3	2.1	SFRP1	2.6	3.9
PPIL4 /// ZC3H12D *	6.1	2.6	MUC5B	2.6	2.5
RGNEF *	6.0	2.0	SLC16A3	2.6	2.1
RIMKLB *	5.9	2.8	SFTA1P	2.5	2.9
PLAUR *	5.8	3.1	CD55	2.5	3.2
SCEL	5.7	6.5	SERPINA1 ^	2.4	5.5
PLXNA2 *	5.6	2.2	CD55	2.4	2.9
PITPNC1	5.5	3.0	SFRP1	2.4	4.1
RGNEF *	5.5	2.0	SRGN	2.4	2.1
ANGPTL4 *	5.2	2.3	GRAMD1B	2.3	2.6
PLAT ^	5.1	13.0	HBEGF	2.3	3.1
SCEL	5.0	6.3	SLC16A3	2.3	2.0
CXCR4	4.6	3.6	HBEGF	2.3	2.9
DPYD	4.5	2.3	ITGA2	2.2	2.5
CLEC2B	4.4	2.4	CD55	2.2	3.1
SCEL	4.3	5.9	SLC16A1	2.2	2.4
MPZL2	4.0	3.0	SAT1	2.2	2.4
PLAU	4.0	3.0	PDLIM5	2.1	2.2
PITPNC1	4.0	3.0	TFPI	2.1	2.0
CXCR4	3.9	3.2	VEGFC	2.1	2.4
MPZL2	3.8	2.8	ASPH	2.1	2.5
PLAU	3.6	2.6	F2RL1	2.1	2.0
FRMD3	3.4	3.0	CXorf61	2.1	2.1
PDLIM5	3.3	2.5	LOC572558	2.0	2.1
LOC401097	3.3	2.4	SLC16A1	2.0	2.3
			CDC42EP5	2.0	2.3

* genes specifically induced in PNT2C2-ETV1 cells

^ genes specifically induced in PNT2C2-dETV1 cells

Supplementary Table S7A ETV1-associated upregulated genes in primary prostate tumors

Gene Name	Score(d)	Fold Change	Gene Name	Score(d)	Fold Change
GPR19	5.61	7.23	UBE2I	3.90	1.23
FKBP10	5.31	3.31	LAMC1	3.88	2.10
TBCE	4.80	1.63	ZFAND2B	3.86	1.41
TAOK1	4.72	1.34	CALR	3.84	1.26
DOK6	4.69	1.89	KRTCAP2	3.82	1.35
CETN3	4.69	1.81	MAP3K2	3.78	1.58
LOC646982	4.60	1.52	MAP2K3	3.76	1.30
RAG1AP1	4.47	1.69	C6orf153	3.74	1.64
CHCHD4	4.43	1.47	COPS7B	3.73	1.31
MRPL21	4.41	1.35	CCDC6	3.69	1.66
LTV1	4.39	1.90	PRDX5	3.68	1.57
USP32	4.32	1.40	CCT4	3.67	1.29
ARHGAP18	4.31	1.73	LEPRE1	3.66	1.42
CEP57	4.27	1.36	TMEM206	3.65	1.54
HDAC2	4.27	1.52	NRBP1	3.64	1.23
RBBP4	4.24	1.88	RAB43	3.63	1.68
MAP4K3	4.23	1.56	XRCC5	3.62	1.24
VCP	4.19	1.25	PTGES2	3.62	1.37
FNDC3B	4.13	1.49	CTAGE5	3.61	1.64
EBNA1BP2	4.11	1.48	STK3	3.59	1.35
WSB2	4.04	1.53	SSBP3	3.58	1.62
ETV1	4.03	27.24	RDX	3.58	1.50
JMJD4	4.02	1.54	CDK2AP1	3.58	1.41
ZCCHC13	4.00	1.60	PTK7	3.58	2.07
KCNN4	3.98	3.23	SNRNP40	3.57	1.51
MRPL55	3.94	1.40	EIF5B	3.56	1.38
FAM71E2	3.92	1.46	MRS2	3.55	1.20
RPN1	3.91	1.45	HSP90B1	3.53	1.28

Supplementary Table S7B ETV1-associated downregulated genes in primary prostate tumors

Gene Name	Score(d)	Fold Change	Gene Name	Score(d)	Fold Change
TARP	-8.64	0.35	NFE2	-4.06	0.50
TARP	-6.76	0.22	CASP1	-4.03	0.59
KCNH6	-5.65	0.41	GGT7	-4.02	0.55
BIN3	-5.20	0.65	HOXB4	-3.99	0.58
ADAP2	-5.07	0.56	IFI6	-3.98	0.53
FAM3B	-4.91	0.22	CPEB3	-3.95	0.50
C1orf116	-4.87	0.54	BTBD8	-3.92	0.62
FXYD3	-4.77	0.38	ANKRD49	-3.92	0.49
MB	-4.76	0.44	CNTNAP2	-3.92	0.29
GNG5	-4.58	0.61	WBSCR17	-3.87	0.54
ANXA11	-4.47	0.74	PLD6	-3.85	0.68
C17orf48	-4.45	0.53	SLC45A3	-3.85	0.54
IMMP2L	-4.45	0.64	SNAPC2	-3.83	0.67
PPFIBP1	-4.42	0.57	RBX1	-3.80	0.46
NHLRC2	-4.41	0.56	TBC1D10A	-3.78	0.69
NRAS	-4.39	0.59	TLE4	-3.76	0.51
RPL18	-4.37	0.64	CIRBP	-3.74	0.75
PIK3IP1	-4.36	0.55	SYNGR1	-3.73	0.65
POM121L12	-4.34	0.62	TXNRD2	-3.73	0.65
RPS16	-4.32	0.73	FAM35A	-3.72	0.63
MARVELD3	-4.30	0.56	KIAA1191	-3.70	0.78
CATSPER3	-4.28	0.64	NPPB	-3.70	0.39
CDIPT	-4.23	0.63	ABHD2	-3.69	0.28
TSPAN9	-4.12	0.68	SCGB2A2	-3.67	0.39
GLUD2	-4.07	0.54	TASP1	-3.66	0.57
ACVRL1	-4.06	0.66			

Supplementary Table S8A ERG-associated upregulated genes in primary prostate tumors

Gene Name	Score(d)	Fold Change
ERG	7.97	30.27
40057	7.22	1.80
NCALD	6.08	3.16
HDAC2	5.96	1.72
ZFP91	5.79	1.28
STEAP4	5.78	2.20
GUCY1A3	5.75	1.72
MAP7	5.72	1.69
ANKRD6	5.52	1.79
TDRD1	5.34	6.00
SH3RF1	5.28	2.41
PDE3B	5.22	2.23
SUB1	5.16	1.49
EFTUD1	5.04	1.79
COG2	5.04	1.57
HNRNPL	5.00	1.44
EPB41	4.95	1.48
LAMC1	4.88	1.67
IFNGR1	4.87	1.59
TMEM63A	4.82	1.84
SPATS2L	4.81	2.04
PTPRK	4.81	1.47
BRD2	4.80	1.43
SAMD1	4.79	1.57
CSGALNACT1	4.76	2.73
PACS1	4.73	1.57
ANXA5	4.72	1.36
NECAP2	4.71	1.40
CDC42SE2	4.70	1.75
PRDX5	4.68	1.65
PLA2G7	4.68	4.50
ZFAND5	4.67	1.43
CPSF7	4.65	1.30
NKAIN1	4.61	6.84

Supplementary Table S8B ERG-associated downregulated genes in primary prostate tumors

Gene Name	Score(d)	Fold Change	Gene Name	Score(d)	Fold Change
GGT7	-9.06	0.49	SGCD	-6.45	0.42
FAM3B	-8.58	0.23	LIPH	-6.42	0.51
PTPRM	-8.52	0.46	C6orf192	-6.40	0.55
C17orf48	-8.41	0.45	PPM1K	-6.37	0.50
INPP4B	-8.17	0.31	FXYD3	-6.37	0.49
MESP1	-8.12	0.42	GSTM3	-6.33	0.36
RAB27A	-7.97	0.32	MYL6B	-6.33	0.74
TMEM220	-7.85	0.33	PPFIBP1	-6.32	0.65
C3orf14	-7.75	0.26	LCOR	-6.31	0.70
GLUD2	-7.73	0.48	RWDD1	-6.25	0.46
C1orf115	-7.46	0.36	ABHD2	-6.21	0.41
DLG2	-7.40	0.41	DBI	-6.14	0.51
LAT2	-7.29	0.51	NCAPD3	-6.12	0.28
CPEB3	-7.00	0.50	HPGDS	-6.07	0.33
LOC200261	-6.98	0.53	MPST	-5.97	0.70
EXTL2	-6.97	0.48	ACAD8	-5.96	0.21
SLC13A3	-6.91	0.41	SMUG1	-5.95	0.57
SLC16A9	-6.89	0.53	SLC4A4	-5.94	0.37
MIPEP	-6.78	0.23	DIRAS1	-5.93	0.42
RAB2B	-6.74	0.49	DPY19L2	-5.92	0.39
USP46	-6.73	0.69	CDIPT	-5.90	0.66
FAM115A	-6.66	0.59	SCGB2A2	-5.90	0.42
PRPS2	-6.64	0.57	POLD4	-5.89	0.68
ABHD2	-6.50	0.36	TFF3	-5.89	0.17
SUN2	-6.47	0.59	AFF3	-5.89	0.31
FAM122B	-6.46	0.65			

Supplementary Table S9A Genes associated with ETV1 and ERG overexpression in primary prostate cancer

Gene Name	ETV1 Score(d)	position	ERG Score(d)	position
HDAC2	4.27	15	5.96	4
LAMC1	3.88	30	4.88	18
PRDX5	3.68	39	4.68	30

Supplementary Table S9B Genes inversely correlated with ETV1 and ERG overexpression in primary prostate cancer

Gene Name	ETV1 Score(d)	position	ERG Score(d)	position
FAM3B	-4.91	6	-8.58	2
FXYD3	-4.77	8	-6.37	31
C17orf48	-4.45	12	-8.41	4
PPFIBP1	-4.42	14	-6.32	34
CDIPT	-4.23	23	-5.9	47
GLUD2	-4.07	25	-7.73	10
GGT7	-4.02	29	-9.06	1
CPEB3	-3.95	33	-7	14
ABHD2	-3.69	49	-6.5	24
SCGB2A2	-3.67	50	-5.9	48

Supplementary Table S10 Primer sequences

Primer name	Forward (5'->3')	Reverse (5'->3'
primers expression constructs		
ETV1	AGATTTGCGAAGAGCAGCAG	CCCTGCTTGACTGTCACTTG
dETV1	ATCAGAAGCCACAAGTGGGA	CCCTGCTTGACTGTCACTTG
primers fig 1C		
ETV1	CATACCAACGGCGAGGATCA	TGGAGAAAAGGGCTTCTGGA
BCL2A1	CACAGGAGAATGGATAAGGC	CAGGAGAGATAGCATTTCAC
PAPPA	TGCTGGTCTCAAGTGGTATC	GCACAGTCACCTTGTAGATC
SRGN	AGGACCAATGTTCGAACTAC	GCAGATTCCTGTCAAGAGAC
primers fig 2A		
PAPPA (1)	GTCGACGTCCAACCTCTATTGGAGAG	GCTAGCTCTTCGCTTCTTCACCCTTC
PAPPA (2)	GTCGACGTACTAGGACTACAGCTTCT	GCTAGCTCTTCGCTTCTTCACCCTTC
PAPPA (3)	GTCGACGCTTGTGTTATAAAGGACGC	GCTAGCTCTTCGCTTCTTCACCCTTC
primers fig 3A		
BCL2A1 (1)	GTCGACGACTGTAAGT CCAGTGCTCT	AAGCTTAGTC TTGAGCTGGC TCACCT
BCL2A1 (3)	GTCGACTGCTAGGATTACAGGCGTGA	AAGCTTAGTC TTGAGCTGGC TCACCT
BCL2A1 (4)	GTCGACCATTTCACATTCCCACCAGC	AAGCTTAGTC TTGAGCTGGC TCACCT
BCL2A1 (5)	GTCGACGGAAGTGGCTTCTCTGAAAC	AAGCTTAGTC TTGAGCTGGC TCACCT
primers fig 4C		
ETV1	CATACCAACGGCGAGGATCA	TGGAGAAAAGGGCTTCTGGA
PLAUR	GAAGAACAGTGCCTGGATGT	CGGCAGATTTTCAAGCTCCA
MMP1	TCCTGGAATTGGCCACAAAG	GCAGTTGAACCAGCTATTAGC
PAPPA	TGCTGGTCTCAAGTGGTATC	GCACAGTCACCTTGTAGATC
primers fig 4D		
CRISP3	GTTTGCCAATATTGTCCTGC	CTCTCTACATAGCCCTACTC
PAEP	TGCTCGATACTGACTACGAC	CCATCTGTTTCAAGTCCAGC
PBGD	CATGTCTGGTAACGGCAATG	GTACGAGGCTTTCAATGTTG

CHAPTER 4

Overexpression of prostate specific *TMPRSS2(exon 0)-ERG* fusion transcripts corresponds with favorable prognosis of prostate cancer

Karin G. Hermans[1*], Joost L. Boormans[2*], Delila Gasi[1], Geert J.H.L. van Leenders[1], Guido Jenster[2], Paul C.M.S. Verhagen[2] and Jan Trapman[1]

[1] Department of Pathology, Josephine Nefkens Institute, Erasmus University Medical Center, Rotterdam, the Netherlands.
[2] Department of Urology, Erasmus University Medical Center, Rotterdam, the Netherlands.
* Equal contribution

Clin Cancer Res 2009; 15:6398-6403

Statement of Translational Relevance

The *TMPRSS2-ERG* gene fusion is the major genetic alteration in prostate cancer. The prognostic value of *TMPRSS2-ERG* in prostate cancer is still a subject to debate. We describe novel findings of a *TMPRSS2* transcript starting at an alternative exon, denoted exon 0. We show that this transcript is much more prostate-specific than the generally studied transcript starting at *TMPRSS2* exon 1. Furthermore, we provide important evidence that the expression of the *TMPRSS2(exon 0)-ERG* fusion transcript correlates with a good prognosis of prostate cancer, whereas *TMPRSS2(exon 1)-ERG* transcripts do not show such a correlation. Our findings urge further investigation of the heterogeneity of *TMPRSS2-ERG* in prostate cancer. More systematic identification of specific fusion transcripts like *TMPRSS2(exon 0)-ERG* or alternatively spliced mRNAs might assist in a molecular classification of prostate cancers, which can be integrated in therapeutic decision making in the future.

Abstract

Purpose: To gain insight in the mechanism and clinical relevance of *TMPRSS2-ERG* expression in prostate cancer, we determined the specific characteristics of fusion transcripts starting at *TMPRSS2* exon 1 and at a more upstream and less characterized exon 0.

Experimental design: We used quantitative PCR analysis to investigate expression of wild-type *TMPRSS2(exon 0)* and *TMPRSS2(exon 1)* and of *ERG* fusion transcripts. Expression was tested in normal tissue samples, in prostate cancer cell lines and xenografts, and in fresh-frozen clinical prostate cancer samples (primary tumors and recurrences). Expression in clinical samples was correlated with disease progression.

Results: *TMPRSS2(exon 0)* and *TMPRSS2(exon 1)* transcripts were similarly androgen regulated in prostate cancer cell lines, but the expression levels of *TMPRSS2(exon 1)* were much higher. Comparison of expression in different tissues showed *TMPRSS2(exon 0)* expression to be much more prostate specific. In androgen receptor-positive prostate cancer xenografts, *TMPRSS2(exon 1)* transcripts were expressed at similar levels, but *TMPRSS2(exon 0)* transcripts were expressed at very variable levels. The same phenomenon was observed for *TMPRSS2-ERG* fusion transcripts. In clinical prostate cancers, the expression of *TMPRSS2(exon 0)-ERG* was even more variable. Expression of *TMPRSS2(exon 0)-ERG* transcripts was detected in 55% (24 of 44) of gene fusion-positive primary tumors but only in 15% (4 of 27) of gene fusion-positive recurrences and at much lower levels. Furthermore, in primary tumors, expression of *TMPRSS2(exon 0)-ERG* transcripts was an independent predictor of biochemical progression-free survival.

Conclusion: The expression of *TMPRSS2(exon 0)-ERG* fusion transcripts in prostate cancer is associated with a less-aggressive biological behavior.

Introduction

Recently, recurrent fusions of prostate-specific and androgen-regulated *TMPRSS2* to the ETS genes *ERG*, *ETV1*, *ETV4,* and *ETV5* have been reported as the most frequent genetic alterations in clinical prostate cancer (1-7). *TMPRSS2-ERG* fusion is detected in 40-70% of clinical prostate cancers. Fusion of *ETV1*, *ETV4,* and *ETV5* to *TMPRSS2* are much less frequent, but *ETV1*, *ETV4,* and *ETV5* have multiple fusion partners. Expression of most of these partner genes is prostate specific and androgen regulated (1-5). Some clinical studies have shown *TMPRSS2-ERG* to be associated with a more aggressive prostate cancer phenotype (8-12). However, other studies did not find an association with outcome in patients treated by radical prostatectomy (13,14) or even described *TMPRSS2-ERG* to be correlated with a more favorable outcome (15,16).

TMPRSS2 has more than one first exon (UCSC Genome Browser, genome.ucsc.edu). Not only fusion transcripts starting at the well known *TMPRSS2* exon 1 but also transcripts that start from a more upstream and less well characterized alternative first exon, here denoted exon 0, have been identified (14; Hermans, unpublished).

In the present study, we determined the specific characteristics of *TMPRSS2* transcripts starting at exons 1 and 0 in benign prostatic tissue and in prostate cancer. Moreover, we investigated clinical prostate cancer samples (primary tumors and recurrences) for expression of *TMPRSS2(exon 0)-ERG* and *TMPRSS2(exon 1)-ERG* fusion transcripts. In the primary tumors, we correlated fusion gene expression with time to biochemical progression after radical prostatectomy. Our data show different expression patterns of *TMPRSS2(exon 0)* and *TMPRSS2(exon 1)* transcripts. Furthermore, our findings indicate a more favorable prognosis of tumors with *TMPRSS2(exon 0)-ERG* expression.

Materials and Methods

Prostate cancer cell lines and xenografts. Prostate cancer cell lines LNCaP and DuCaP were grown in RPMI-1640 supplemented with 5% fetal calf serum and antibiotics. Androgen receptor-positive prostate cancer xenografts PCEW, PC82, PC295, PC310, PC329, PC346, and PC374, and androgen receptor-negative xenografts PC133, PC135, PC324, and PC339, were propagated by serial transplantation on male nude mice, as described (17,18).

Clinical samples. Primary prostate tumors were obtained by radical prostatectomy and recurrent tumors by transurethral resection of the prostate. Hematoxilin/eosin stained tissue sections were histologically evaluated by two pathologists (T. van der Kwast; G.J.H.L. van Leenders). Although introducing a bias, only samples that contained at least 70% tumor cells were selected for analysis. The clinical and pathological demographics of the patients with primary prostate tumors included in the statistical analysis (N=67) are given in Supplementary Table S1. Tissues were snap frozen and stored in liquid nitrogen. Use of the samples for research purposes was approved by the Erasmus MC Medical Ethics Committee according to the Medical Research Involving Human Subjects Act (MEC-2004-261).

RNA isolation. RNAs from the prostate cancer cell lines LNCaP and DuCaP cultured in the absence or the presence of 10^{-9} mol/L R1881 were isolated using the RNeasy RNA extraction kit (Qiagen). RNA from clinical prostate cancer samples was isolated from frozen tissue sections using RNA-Bee (Campro Scientific). Xenograft RNA was isolated according to the LiCl protocol.

Quantitative PCR (Q-RT-PCR). Total RNA was reverse transcribed using M-MLV reverse transcriptase (Invitrogen) and an oligo dT12 primer. cDNAs of 16 different tissues were purchased from Clontech. Quantitative RT-PCRs were done in Power SYBR Green PCR Master Mix (25 µl), containing 0.33 µmol/L forward and reverse primer in an ABI Prism 7700 Sequence Detection System (Applied Biosystems). Amplified products were quantified relative to porphobilinogen deaminase (*PBGD*). Primer sequences of the primers used are summarized in Supplementary Table S2.

Statistical analysis. Associations between clinical and histopathological variables and expression of *TMPRSS2-ERG* transcripts were evaluated by the Pearson's χ^2 test, the Mann-Whitney U test, or Kruskal-Wallis test, where appropriate. Expression of *TMPRSS2-ERG* transcripts was correlated with the primary end point, biochemical progression-free survival, defined as time from radical prostatectomy to date of biochemical recurrence.

Biochemical recurrence was defined as (*a*) a prostate-specific antigen level of >0.2 ng/ml at two consecutive measurements with a 3-month interval if the prostate-specific antigen nadir was <0.1 ng/ml or (*b*) a prostate-specific antigen nadir of ≥0.2 ng/ml. Patients that died from causes other than prostate cancer or that were lost to follow-up were censored at the date of last prostate-specific antigen test. Patients were routinely followed thrice monthly during the first year after radical prostatectomy, semi annually during the second year, and subsequently at 12-month intervals. In case of progression, patients were followed every 3 months. Kaplan-Meier curves were constructed to assess the probability of remaining free of biochemical recurrence as a function of time after surgery. The differences between the survival curves of the groups were tested using the log-rank test or Breslow

method if appropriate. A Cox proportional regression analysis with forward stepwise elimination was done to assess the impact of various parameters on time to recurrence. In the multivariate analysis, the model included pathological T stage, surgical margin status, the Gleason score of the primary tumor, and expression of indicated *TMPRSS2-ERG* fusion transcripts. Patients with unknown parameters were excluded from the analysis. Statistical analyses were done using the Statistical Package for Social Sciences, version 15.0 (SPSS), with a significance level of 0.05 (two-tailed probability).

Results

TMPRSS2-ERG gene fusion is present in 40-70% of primary prostate tumors. *ERG* and *TMPRSS2* are located ~3 Mbp apart on chromosome 21q22.3 in the same orientation (Fig. 1A). The most common *TMPRSS2-ERG* fusion transcripts are composed of *TMPRSS2* exon 1 or exons 1 and 2 linked to exon 4 of *ERG*. Less frequently, fusion of *TMPRSS2* exon 1 or 2 to other *ERG* exons have been detected (12). Genomic databases describe that *TMPRSS2* transcripts might also contain an alternative first exon, here denoted exon 0, which maps ~4-kbp upstream of exon 1 (Fig. 1A). *TMPRSS-ERG* fusion transcripts might also contain *TMPRSS2* exon 0 (14; Hermans, unpublished).

We determined the specific characteristics of transcripts starting at either exon 0 or exon 1 of *TMPRSS2*. First, the expression of both transcripts was investigated on a panel of cDNA from 16 different normal tissues. Although expression of *TMPRSS2(exon 1)* mRNA was highest in the prostate, these transcripts were also detected in lung, kidney, pancreas and colon samples. In contrast, *TMPRSS2(exon 0)* mRNA had a completely prostate-specific expression pattern (Fig. 1B). Subsequently, we did a quantitative RT-PCR analysis on RNA from prostate cancer cell lines DuCaP and LNCaP cultured in absence or presence of the synthetic androgen R1881. Both *TMPRSS2* transcripts were induced by androgens (Fig. 1C). Notably, in the cell lines, expression of *TMPRSS2(exon 0)* transcripts was much lower than expression of *TMPRSS2(exon 1)*.

Testing of RNAs from eleven human prostate cancer xenografts for expression of *TMPRSS2* starting at either exon 0 or 1 showed that six androgen receptor-positive xenografts, PCEW, PC82, PC295, PC329, PC346, and PC374, expressed

TMPRSS2(exon 1) at less-variable levels (Fig. 1D). Xenograft PC310 showed no expression of *TMPRSS2* because of a deletion of the wild-type allele (17). Five xenografts expressed *TMPRSS2(exon 0)* with a much more variable level of expression. None of the androgen receptor-negative xenografts expressed *TMPRSS2*.

Previously, we have shown that five androgen-dependent xenografts, PCEW, PC82, PC295, PC310, and PC329, contained *TMPRSS2(exon 1)-ERG* mRNA (17). Quantitative RT-PCR analysis for [*TMPRSS2(exon 1)-ERG* and *TMPRSS2(exon 0)-ERG*] fusion transcripts showed that three xenografts, PC82, PC295, and PC329, expressed *TMPRSS2(exon 0)-ERG* at different levels (Fig. 1D). However, the other two xenografts that expressed *TMPRSS2-ERG* transcripts, PCEW and PC310, did not express the *TMPRSS2(exon 0)-ERG* fusion transcript at all. The expression levels of *TMPRSS2(exon 1)-ERG* transcripts in the five xenografts did not show a large variation (Fig. 1D). The difference in *TMPRSS2(exon*

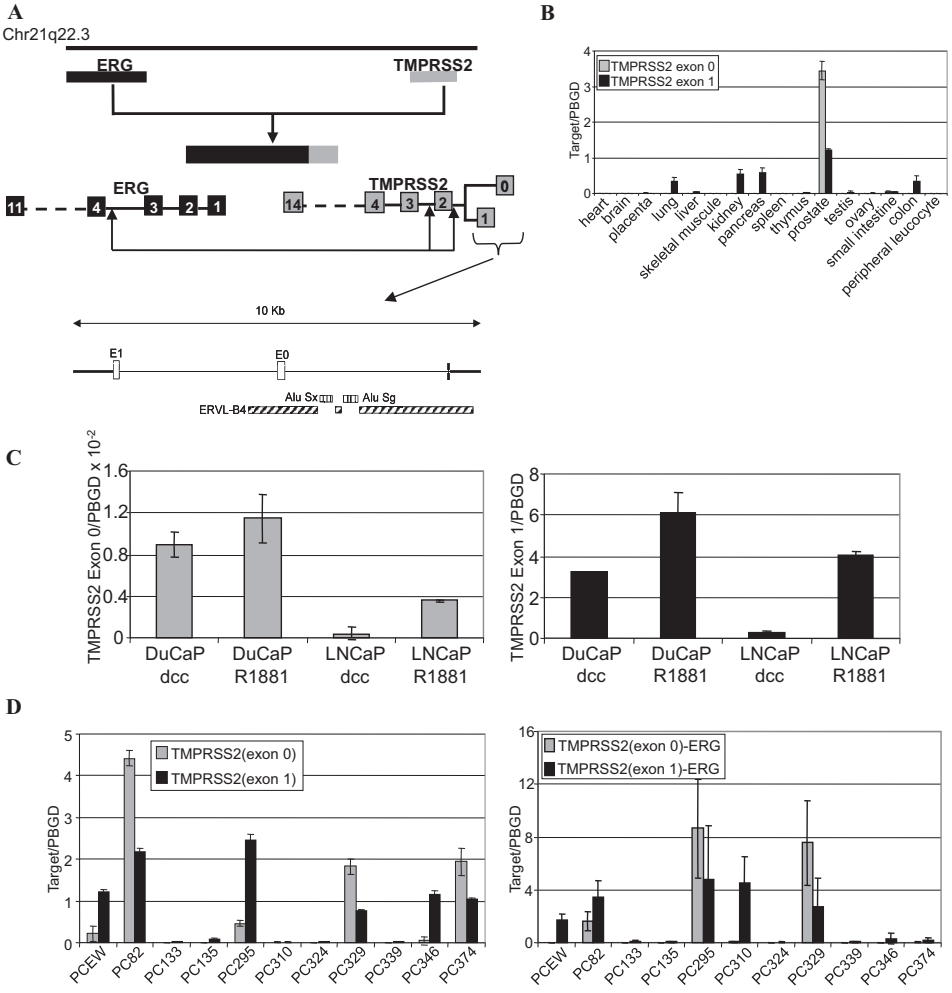


Fig. 1. Characterization of *TMPRSS2* and *TMPRSS2-ERG* transcripts starting at exon 0 or exon 1. *A,* schematic representation of the *TMPRSS2-ERG* locus on chromosome band 21q22.3. The most frequent gene fusion events are indicated. The enlarged genomic region containing *TMPRSS2* shows exon 0 and 1 and repeat sequences. *B,* tissue-specific expression of *TMPRSS2(exon 0)* and *TMPRSS2(exon 1)* mRNA assessed by quantitative RT-PCR analysis on a cDNA panel from 16 different normal tissues. Mean of duplicate experiment relative to *PBGD* with SD. *C,* androgen-regulated expression of *TMPRSS2(exon 0)* (*left*) and *TMPRSS2(exon 1)* (*right*) mRNA in androgen receptor-positive prostate cancer cell lines LNCaP and DuCaP. LNCaP and DuCaP cells were grown in the absence and presence of synthetic androgen R1881 (10^{-9} mol/L) for 24h. Mean of duplicate experiments relative to *PBGD* with SD are depicted. Note that the level of *TMPRSS2(exon 0*) expression is much lower in the cell lines than in the normal prostatic tissue (*B*). *D,* quantitative RT-PCR analysis of *TMPRSS2(exon 0)* and *TMPRSS2(exon 1)* (*left*) and *TMPRSS2(exon 0)-ERG* and *TMPRSS2(exon 1)-ERG* (*right*) transcripts in 11 human prostate cancer xenografts. Mean of duplicate experiments relative to *PBGD* with SD.

0)-ERG and *TMPRSS2(exon 1)-ERG* expression was similar to that observed for wild-type *TMPRSS2(exon 0)* and *(exon 1)* transcripts.

Next, we determined the expression of *TMPRSS2(exon 0)-ERG* and *TMPRSS2(exon 1)-ERG* transcripts in a cohort of 126 fresh-frozen clinical prostate cancer samples (81 primary tumors and 45 recurrent tumors; Fig. 2A). *TMPRSS2-ERG* transcripts were detected in 54% (44 of81) of the primary tumors and in 60% (27 of 45) of the recurrences. In the primary tumors, 20 (25%) of 81 of the cases exclusively expressed *TMPRSS2(exon 1)-ERG* transcripts, three samples (4%) exclusively expressed *TMPRSS2(exon 0)-ERG* and 21 (26%) expressed both transcripts. Analysis of wild-type *TMPRSS2* in benign prostatic tissue excluded

preferential expression of *TMPRSS2(exon 1)* or *TMPRSS2(exon 0)* transcripts in one of the different prostate zones (data not shown). In the recurrent tumors, exclusive expression of *TMPRSS2(exon 1)-ERG* was detected in 23 (51%) of 45 of the cases, whereas none expressed exclusively *TMPRRS2(exon 0)-ERG* and only four cases (9%) expressed both transcripts. Expression levels of *TMPRSS2(exon 0)-ERG* transcripts were significantly higher in primary tumors than in recurrent tumors (p=0.015), and variation in expression was much larger in the primary tumors than in the recurrences (Fig. 2B). In contrast, the percentage of tumors expressing *TMPRSS2(exon 1)-ERG* transcripts was in the same range for primary and recurrent tumors, and the expression levels of these transcripts did not differ between both tumor types (p=0.74).

We correlated expression of *TMPRSS2(exon 0)-ERG* with clinical outcome in the primary prostate cancer cohort (n = 81) to see whether it was of prognostic value. We excluded from the analysis 10 patients that were known to harbor fusion or overexpression of other *ETS* genes and four patients whose primary treatment was not a radical prostatectomy. Despite the very long follow-up available (median, >10 years), only 11 of the remaining 67 patients died from prostate cancer, precluding statistical analysis. Instead, we used time to prostate-specific antigen recurrence after radical prostatectomy as an end point. The patients' demographics are summarized in Supplementary Table S1. No differences were seen in clinical and histopathological characteristics between patients expressing *TMPRSS2-ERG* and gene fusion-negative patients, although *TMPRSS2-ERG*-negative patients had higher Gleason scores with borderline significance (P = 0.053; Supplementary Table S3). The median time to prostate-specific antigen progression was not significantly different between the two groups: 73.2 versus 122.1 months [95% confidence interval (95% CI), 32.7-113.7 versus 70.6-173.6; P = 0.45; Fig. 3A).

Within the *TMPRSS2-ERG*-positive population the only difference between patients that exclusively expressed *TMPRSS2(exon 1)-ERG* transcripts and patients that expressed the

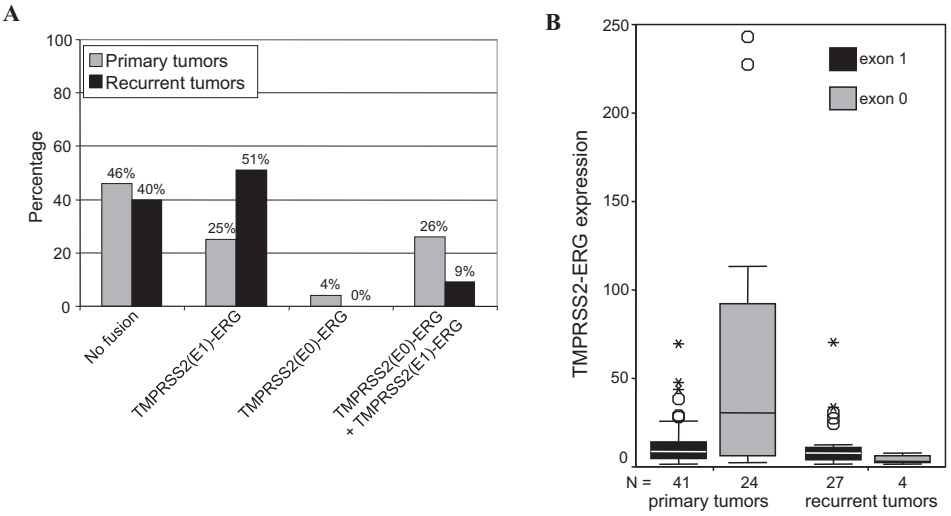


Fig. 2. Expression of *TMPRSS2(exon 0)-ERG* and *TMPRSS2(exon 1)-ERG* transcripts in clinical prostate cancer samples. *A,* distribution of both *TMPRSS2-ERG* fusion transcripts in primary and recurrent tumors. Primary tumors, n = 81; recurrent tumors, n = 45. *B,* box plot of *TMPRSS2(exon 0)-ERG* and *TMPRSS2(exon 1)-ERG* mRNA expression levels in primary tumors and recurrences. Open circles, outliers; asterisk, extremes.

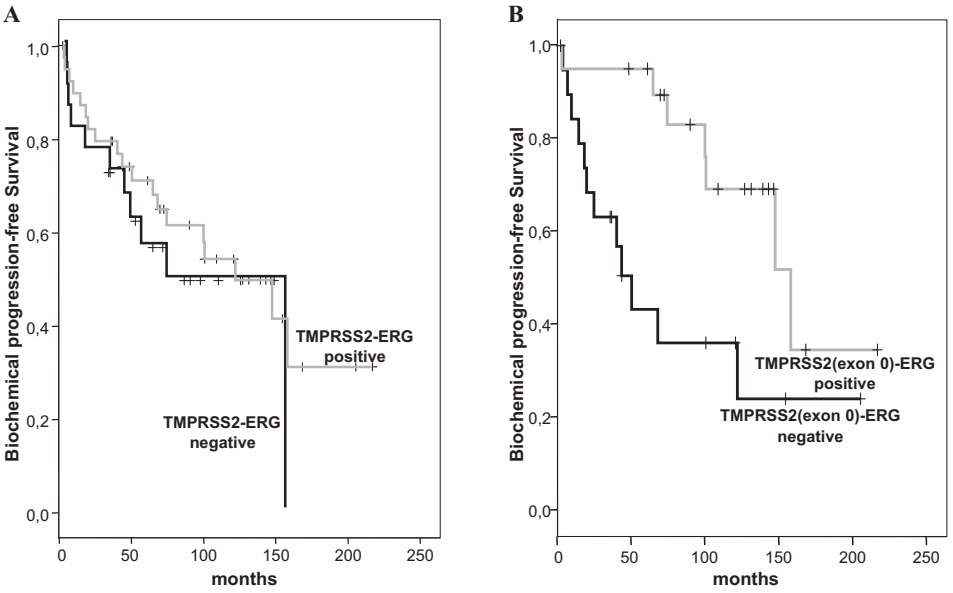


Fig. 3. Kaplan-Meier curves for time to prostate-specific antigen recurrence after radical prostatectomy defined by *TMPRSS2-ERG* fusion transcript status. *A,* biochemical progression-free survival curves for prostate cancer patients with or without expression of *TMPRSS2-ERG* transcripts. *B,* biochemical progression-free survival curves for the *TMPRSS2-ERG*-positive group, which was stratified in patients with and without expression of *TMPRSS2(exon 0)-ERG* transcripts.

TMPRSS2(exon 0)-ERG subtype was that the former had higher pathological stages than the latter (P = 0.009; Supplementary Table S4). The median time to prostate-specific antigen progression for patients expressing *TMPRSS2(exon 0)-ERG* transcripts was significantly longer than for patients that exclusively expressed *TMPRSS2(exon 1)-ERG* transcripts: 158.2 versus 50.5 months (95% CI: 98.9-217.5 versus 32.6-68.4; P = 0.012; Fig. 3B).

Using a Cox proportional hazards model, positive surgical margins, Gleason score of ≥7, pathological stage of ≥pT3a, and absence of *TMPRSS2(exon 0)-ERG* transcripts were all associated with a worse biochemical progression-free survival. Importantly, multivariate analysis with forward stepwise selection showed expression of *TMPRSS2(exon 0)-ERG* fusion transcripts to be an independent predictor of progression-free survival (hazard ratio, 0.34; 95% CI, 0.14-0.84; P = 0.019; Table 1).

Table 1. Results of univariate and multivariate analyses

		Univariate			Multivariate		
Variable	**N**	**Median time to PSA recurrence (months)**	**95% CI**	**p-value**	**HR (95% CI)**	**p-value**	**HR (95% CI)**	**p-value**
Surgical margins								
positive	30	43.9	31.0-56.8	< 0.001	7.5 (3.0-18.5)	< 0.001	7.7 (3.0-19.4)	0.001
negative	31	122.1	49.5-194.7		1.0			
Gleason Score								
≥ 7	26	68.2	41.5-94.9	0.037	2.1 (1.0-4.5)	0.041		
< 7	35	155.4	117.7-192.2		1.0			
pT-stage								
extra prostatic	43	65.0	37.4-92.5	< 0.001	5.7 (1.9-16.8)	0.002		
organ confined	18	158.2	146.7-169.7		1.0			
TMPRSS2(exon 0)-ERG expression								
Yes	2140	158.2	98.9-217.5	0.015	0.36 (0.15-0.85)	0.02	0.34 (0.14-0.84)	0.019
No		68.2	36.8-99.6		1.0			

Abbreviations: HR, hazard ratio; pT-stage, pathological T-stage; PSA, prostate-specific antigen; CI, confidence interval

Discussion

This study addresses two important aspects of *TMPRSS2-ERG* expression in prostate cancer. First of all, a remarkable difference in expression characteristics was detected between *TMPRSS2(exon 1)* and *TMPRSS2(exon 1)-ERG* transcripts on the one hand and *TMPRSS2(exon 0)* and *TMPRSS2(exon 0)-ERG* transcripts on the other hand. Secondly, the clinical data indicated a more favorable prognosis for prostate cancer patients expressing *TMPRSS2(exon 0)-ERG* transcripts.

It is estimated that almost half of all genes in the human genome contain more than one first exon, as an important mechanism to regulate gene expression (19). Here, we showed that *TMPRSS2* transcripts starting at exon 0 were much more prostate specific than those starting at exon 1 (Fig. 1B) and that the expression level of transcripts containing exon 0 was much more variable (Fig. 1D). *TMPRSS2* exon 0 is located in a retroviral repeat element, ERVL-B4 (Fig. 1A). This repeat does not contain a standard long terminal repeat promoter element; however, other retroviral repeat sequences might function as cryptic promoters (19,20). Within the same retroviral repeat, the *TMPRSS2* sequence present in a *TMPRSS2-ETV4* fusion transcript is located (6). Although a different 5'-untranslated region might affect translation efficacy, the major proteins translated from the fusion transcripts seem identical N-truncated ERG proteins, which are translated from an ATG codon in the *ERG* exon 4 part of the fusion transcripts.

It could be speculated that the prostate-specific *TMPRSS2(exon 0)* transcripts are expressed in tumors with a more differentiated phenotype. Recurrent tumors represent late-stage prostate cancer that normally has a less differentiated phenotype. In our study, 55% of the recurrences had a Gleason score of ≥8
compared with 15% of the primary tumors. An alternative explanation is that expression from exon 0 is stimulated by the stromal compartment, which will be different in primary tumors and recurrences. Obviously, there are no stromal cells present in *in vitro* cultures of prostate cell lines, which showed very low expression levels of *TMPRSS2(exon 0)* (Fig. 1).

The prognostic significance of *TMPRSS2-ERG* gene fusion remains subject of debate, although a growing number of studies has been published on this matter (8-13, 16). Because technology used to investigate *TMPRSS2-ERG* varies (FISH or quantitative RT-PCR) and compositions of patient cohorts differ considerably, it is difficult to draw general conclusions from available data. In the present study on a selected and rather limited group of well-defined patients with a very long median follow-up, there was no difference in time to prostate-specific antigen recurrence after radical prostatectomy

between patients that expressed *TMPRSS2-ERG* and patients without expression of the fusion gene.

In two studies on watchful waiting cohorts, it was shown that patients having *TMPRSS2-ERG* fusion had a higher incidence of metastases or cancer-specific death than gene fusion-negative patients (8,9); however, this was not confirmed in a large surgically treated cohort (13). Attard et al. (8) showed by FISH analysis that patients with an interstitial deletion of genomic sequences between *TMPRSS2* and *ERG* (so called "class Edel") had poorer cancer-specific and overall survival than gene fusion-negative patients or than patients with *TMPRSS2-ERG* fusion without loss of the genomic region between the two genes. Other studies have correlated *TMPRSS2-ERG* with biochemical progression after radical prostatectomy, like in the

present study. Before the identification of *TMPRSS2-ERG*, Petrovics et al. (15) found that patients with high expression levels of *ERG* had longer prostate-specific antigen recurrence-free survival than patients without *ERG* overexpression. Recently, similar results were reported by Saramaki et al. (16), using FISH analysis of *TMPRSS2-ERG*. However, other studies claimed a negative correlation between *TMPRSS2-ERG* and prostate-specific antigen recurrence (10-12). Perner et al. indicated that patients with *TMPRSS2-ERG* rearrangement through deletion showed a trend for higher prostate-specific antigen recurrence rate than patients without fusion (11). Wang et al. (12) provided evidence that specific *TMPRSS2-ERG* splice variants were associated with early prostate-specific antigen recurrence.

Information on *TMPRSS2(exon 0)-ERG* transcripts in prostate cancer is still scarce. Thus far, this transcript was only identified by Lapointe et al. (14). However, the clinical implications of this specific fusion transcript subtype were not investigated. The low expression frequency of *TMPRSS2(exon 0)-ERG* transcripts in late-stage prostate cancer and the favorable prognosis for patients expressing this fusion transcript in primary tumors, as shown in our study, urges further investigation of the heterogeneity of *TMPRSS2-ERG* in prostate cancer. More systematic identification of specific fusion transcripts like *TMPRSS2(exon 0)-ERG* or alternatively spliced mRNAs (12) might assist in a molecular classification of prostate cancers, which can be integrated in therapeutic decision making in the future.

Disclosure of Potential Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgments

The authors thank Theo van der Kwast for pathology, Wilma Teubel for collection of clinical samples, Wytske van Weerden for the xenograft tissues, Anieta Siewerts for the RNA isolation, Natasja Dits for the cDNA samples of clinical prostate tumors, and Mark Wildhagen for support with statistical analysis.

References

1. Attard G, Clark J, Ambroisine L, et al. Heterogeneity and clinical significance of ETV1 translocations in human prostate cancer. Br J Cancer 2008;99:314-20.
2. Helgeson BE, Tomlins SA, Shah N, et al. Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. Cancer Res 2008;68:73-80.
3. Hermans KG, Bressers AA, van der Korput HA, Dits NF, Jenster G, Trapman J. Two unique novel prostate-specific and androgen-regulated fusion partners of ETV4 in prostate cancer. Cancer Res 2008;68:3094-8.
4. Hermans KG, van der Korput HA, van Marion R, et al. Truncated ETV1, fused to novel tissue-specific genes, and full length ETV1 in prostate cancer. Cancer Res 2008;68:7541-9.
5. Tomlins SA, Laxman B, Dhanasekaran SM, et al. Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature 2007;448:595-9.
6. Tomlins SA, Mehra R, Rhodes DR, et al. TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. Cancer Res 2006;66:3396-400.
7. Tomlins SA, Rhodes DR, Perner S, et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644-8.
8. Attard G, Clark J, Ambroisine L, et al. Duplication of the fusion of TMPRSS2 to ERG sequences identifies fatal human prostate cancer. Oncogene 2008;27:253-63.
9. Demichelis F, Fall K, Perner S, et al. TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. Oncogene 2007;26:4596-9.
10. Nam RK, Sugar L, Yang W, et al. Expression of the TMPRSS2:ERG fusion gene predicts cancer recurrence after surgery for localised prostate cancer. Br J Cancer 2007;97:1690-5.
11. Perner S, Mosquera JM, Demichelis F, et al. TMPRSS2-ERG fusion prostate cancer: an early molecular event associated with invasion. Am J Surg Pathol 2007;31:882-8.
12. Wang J, Cai Y, Ren C, Ittmann M. Expression of variant TMPRSS2/ERG fusion messenger RNAs is associated with aggressive prostate cancer. Cancer Res 2006;66:8347-51.
13. Gopalan A, Leversha MA, Satagopan JM, Zhou Q, et al. *TMPRSS2-ERG* gene fusion is not associated with outcome in patients treated by prostatectomy. Cancer Res 2009;69:1400-06.
14. Lapointe J, Kim YH, Miller MA, et al. A variant TMPRSS2 isoform and ERG fusion product in prostate cancer with implications for molecular diagnosis. Mod Pathol 2007;20:467-73.
15. Petrovics G, Liu A, Shaheduzzaman S, et al. Frequent overexpression of ETS-related gene-1 (ERG1) in prostate cancer transcriptome. Oncogene 2005;24:3847-52.
16. Saramaki OR, Harjula AE, Martikainen PM, Vessella RL, Tammela TL, Visakorpi T. TMPRSS2:ERG fusion identifies a subgroup of prostate cancers with a favorable prognosis. Clin Cancer Res 2008;14:3395-400.
17. Hermans KG, van Marion R, van Dekken H, Jenster G, van Weerden WM, Trapman J. TMPRSS2:ERG fusion by translocation or interstitial deletion is highly relevant in androgen-dependent prostate cancer, but is bypassed in late-stage androgen receptor-negative prostate cancer. Cancer Res 2006;66:10658-63.
18. van Weerden WM, de Ridder CM, Verdaasdonk CL, et al. Development of seven new human prostate tumor xenograft models and their histopathological characterization. Am J Pathol 1996;149:1055-62.

19. Landry JR, Mager DL, Wilhelm BT. Complex controls: the role of alternative promoters in mammalian genomes. Trends Genet 2003;19:640-8.
20. Gerlo S, Davis JR, Mager DL, Kooijman R. Prolactin in man: a tale of two promoters. Bioessays 2006;28:1051-5.

Supplemental Data

Supplementary Table S1. Clinical and pathological characteristics of 67 patients with primary prostate cancer.

Characteristic		
Mean age (±SD)	62.5	± 5.4 years
Median follow-up (±SD)	127.0	± 48.3 months
Mean PSA (±SD)	15.6	± 22.8 ng/ml
cT-stage		
organ confined	45	67.2%
extra prostatic	13	19.4%
unknown	9	13.4%
pT-stage		
pT2a	2	3.0%
pT2b	3	4.5%
pT2c	13	19.4%
pT2x	1	1.5%
pT3a	24	35.8%
pT3b	16	23.9%
pT4	5	7.5%
unknown	3	4.5%
Surgical margins		
positive	31	46.3%
negative	32	47.8%
unknown	4	6.0%
Gleason score		
< 7	38	56.7%
= 7	19	28.3%
> 7	10	15.0%
Occult metastases at RP		
Yes	7	10.4%
No	56	83.6%
unknown	4	6.0%
PSA recurrence		
Yes	33	49.3%
No	33	49.3%
Unknown	1	1.5%
Local recurrence		
Yes	9	13.4%
No	55	82.1%
Unknown	3	4.5%
Distant metastases at follow-up		
Yes	11	16.4%
No	55	82.1%
Unknown	1	1.5%
Death		
Yes	27	40.3%
No	40	59.7%
Prostate cancer death		
Yes	11	16.4%
No	54	80.6%
Unknown	2	3.0%

Abbreviations: SD = standard deviation, cT-stage = clinical T-stage, pT-stage = pathological T-stage, RP = radical prostatectomy

Supplementary Table S2. Primer sequences of primers used for QPCR analysis

Target	Forward 5'->3'	Reverse 5'->3'
PBGD	catgtctggtaacggcaatg	gtacgaggctttcaatgttg
TMPRSS2 Exon 1-Exon 3	gagctaagcaggaggcgga	aggggttttccggttggtatc
TMPRSS2 Exon 0-Exon 3	gactacttctactccaccag	aggggttttccggttggtatc
TMPRSS2(exon 0)-ERG	gactacttctactccaccag	catcaggagagttccttgag
TMPRSS2(exon 1)-ERG	gagctaagcaggaggcgga	catcaggagagttccttgag

Supplementary Table S3. Clinical and pathological characteristics of patients expressing ***TMPRSS2-ERG*** versus patients not expressing ***TMPRSS2-ERG***

	TMPRSS2-ERG positive (N=44)	TMPRSS2-ERG negative (N=23)	p-value	Test
Mean age (±SD)	62.2 (± 5.6)	63.0 (± 5.1)	0.57	MWU
Mean PSA (±SD)	16.1 (± 26.7)	14.8 (±12.5)	0.81	MWU
Gleason Score				
< 7	29 (65.9%)	9 (40.9%)	0.053	Chi²
≥ 7	15 (34.1%)	13 (59.1%)		
pT-stage				
extra prostatic	13 (31.0%)	6 (27.3%)	0.76	Chi²
organ confined	29 (69.0%)	16 (72.7%)		
Surgical margins				
Positive	21 (50.0%)	10 (47.6%)	0.86	Chi²
Negative	21 (50.0%)	11 (52.4%)		

Abbreviations: SD= standard deviation, MWU = Mann-Whitney U test, pT-stage=pathological T-stage

*Patients with unknown parameters were not included in the analysis.

Supplementary Table S4. Clinical and pathological characteristics of patients expressing *TMPRSS2(exon 0)-ERG* fusion transcripts, versus patients exclusively expressing *TMPRSS2(exon 1)-ERG* fusion transcripts.

	***TMPRSS2(exon 0)-ERG* expression (N=24)**	**Exclusive *TMPRSS2(exon 1)-ERG* expression (N=20)**	**p-value**	**Test**
Mean age (±SD)	61.5 (±4.8)	63.1 (±6.4)	0.40	MWU
Mean PSA (±SD)	19.7 (± 35.2)	11.9 (± 10.3)	0.63	MWU
Gleason score				
< 7	15 (62.5%)	14 (70.0%)	0.60	Chi^2
≥ 7	9 (37.5%)	6 (30.0%)		
pT-stage				
extra prostatic	11 (47.8%)	2 (10.5%)	0.009	Chi^2
organ confined	12 (52.2%)	17 (89.5%)		
Surgical margins				
Positive	9 (39.1%)	12 (63.2%)	0.12	Chi^2
Negative	14 (60.9%)	7 (36.8%)		

Abbreviations: SD = standard deviation, MWU = Mann-Whitney U test, pT-stage=pathological T-stage

*Patients with unknown parameters were not included in the analysis.

CHAPTER 5

Androgen-regulation of ETS-gene fusion transcripts in prostate cancer

Delila Gasi and Jan Trapman, Department of Pathology, Josephine Nefkens Institute, Erasmus University Medical Centre, Rotterdam, The Netherlands

Methods Mol Biol. 2011;776:335-48.

Abstract

Fusion between androgen-regulated *TMPRSS2* and ETS-transcription factor gene *ERG* is the most frequent genetic alteration that occurs in 40-70% of prostate cancers.
Not only *ERG* but also other ETS-transcription factor genes are involved in gene fusions. *ETV1, ETV4* and *ETV5* have all several fusion partners. One common feature shared by the majority of these partners is androgen-regulated expression.
Despite its high frequency, the biological and molecular effects of ETS gene fusion in prostate cancer development and progression are unknown.
In this chapter quantitative polymerase chain-reaction (Q-PCR) is used for detection and further studying the incidence and properties of these fusion transcripts. The focus is on the expression of *TMPRSS2-ERG* transcripts in clinical prostate samples.
Androgen regulation of *TMPRSS2* is measured in commonly used LNCaP prostate cancer cells grown with and without the synthetic androgen R1881. Furthermore, combining Q-PCR with 5' RLM-RACE and sequencing is described for the identification of novel ETS fusion partners.

1. Introduction

Gene fusions are common genetic alterations in prostate cancer. The fusion between *TMPRSS2* and the ETS transcription factor gene *ERG* has been reported in 40–70% of prostate cancers (1-4). *TMPRSS2* is an androgen-regulated gene that is preferentially expressed in the prostate. *ERG* is a known oncogene and due to the fusion with *TMPRSS2* it is now androgen-regulated overexpressed in the prostate.

The high frequency of *TMPRSS2-ERG* fusion is probably due to the close proximity of the two genes on chromosome 21 (see Fig. 1). Most breakpoints in *TMPRSS2* are in intron 1 or 2 and most breakpoints in *ERG* are in intron 3, however, occasionally other fusions have been found. Moreover, alternative splicing and differential promoter usage have been described. It is assumed that the major ERG protein that is translated from *TMPRSS2-ERG* fusion transcripts is a slightly N-terminal truncated version of the protein. In addition to *TMPRSS2*, two other *ERG* fusion partners, *SLC45A3* and *NDRG1* have been described (see Table 1). Fusions with these genes occur at much lower frequencies. Both

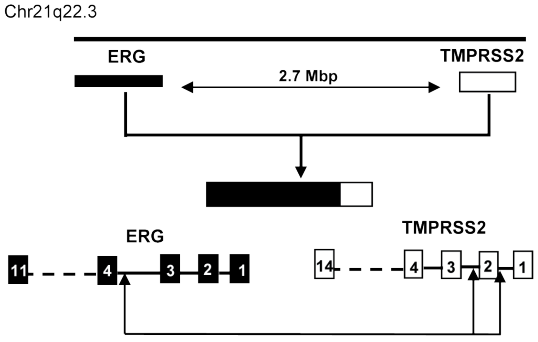


Figure 1: Schematic representation of the fusion between *TMPRSS2* and *ERG*. Both genes are located in the same orientation on chromosome 21, separated by 2.7 Mbp. The black arrows indicate the most frequently occurring fusion events, which are either in intron 1 or 2 of *TMPRSS2* to intron 3 *of ERG*.

Table 1. Fusion partners of ETS-genes in prostate cancer

5' fusion partner	Prostate-specific	Androgen-regulated	3' fusion partner	Present
TMPRSS2 (chr 21q)	+	+	ERG (chr 21q)	40-70%
SLC45A3 (chr 1q)	+	+		< 1%
NDRG1 (chr 8)	-	+		~1%
TMPRSS2 (chr 21q)	+	+	ETV1 (chr 7p)	< 1%
FOXP1 (chr 3p)	ND	ND		< 1%
EST14 (chr 14q)	+	+		< 2%
HERVK17 (chr 17p)	+	+		< 2%
SLC45A3 (chr 1q)	+	+		< 1%
HERV-K_22q11.23	+	+		< 1%
C15orf21 (chr 15q)	+	+ (down)		< 1%
HNRPA2B1 (chr 7p)	-	-		< 1%
ACSL3 (chr 2q)	-	+		< 1%
CANT1 (chr 17)	+	+		< 1%

fusion partners share with *TMPRSS2* the property of androgen-regulated expression (5-6).

Fusion of ETS factor genes *ETV1, ETV4* and *ETV5* occur at low frequencies in prostate cancer (7-9). *ETV1* has at least 10 fusion partners, but again, almost all of its partners are androgen-regulated and most are also prostate-specific expressed (see Table 1). Different from *ERG, ETV1* can be overexpressed also as the full-length gene.

ETS transcription factors are important mediators of many biological processes including proliferation, metastasis, angiogenesis and growth, all also important in tumor development (10). The precise role that the fusion transcripts play in prostate cancer is still unclear. To understand the molecular mechanisms that determine the specific gene fusions, it is important to elucidate their common regulatory expression mechanisms and to distinguish in which stages of prostate cancer they occur.

Because the gene fusions are unique, they are promising functional biomarkers for diagnosis of prostate cancer (11). However, it is unclear whether they are also prognostic markers. So far, conflicting data have been published on this subject (12-14).

Several methods to detect gene fusions in clinical samples can be discriminated. For known fusion events, these include as classical methods interphase Fluorescence In Situ Hybridization (FISH) and specific Q-PCR. More recently, genome wide methods like genomic or transcript paired-end sequencing have been described as global approaches to find novel fusion events (15).

Affymetrix exon microarrays can also function as a first step to identify genes overexpressed by gene fusion (16). Here we describe Q-PCR to identify and to monitor *TMPRSS2-ERG* expression in prostate cancer. Moreover, we describe 5' RACE-PCR to discover novel fusion partners of ETS genes.

2. Materials

2.1 Cell-culture and clinical prostate cancer samples

1. LNCaP prostate cancer cells (ATCC).
2. Dulbecco's modified Eagle's medium (DMEM) (BioWhittaker/Lonza) supplemented with 5 % fetal bovine serum (FBS), 100 U penicillin and 100 U streptomycin (PS) stored at 4-8°C.
3. Trypsin (170.000 U/L)/EDTA (200 mg/L) (BioWhittaker/Lonza) stored at 4-8°C.
4. Phosphate-buffered saline (PBS) (BioWhittaker/Lonza) stored at room temperature (RT).
5. Synthetic androgen R1881 (NEN Chemicals) 10-6 M stock solution in 100% EtOH stored at -20°C in the dark (*see* **Note 1**).
6. 100% EtOH (Sigma Aldrich) stored at RT.
7. Clinical prostate cancer samples, fresh-frozen and stored in liquid nitrogen.

2.2 RNA isolation and cDNA preparation

1. RNeasy Mini kit (Qiagen) for RNA extraction from LNCaP cells.
2. QIAshredder columns (Qiagen) for homogenization of the sample.
3. RNA-Bee solution (BIO Connect) for RNA isolation from clinical prostate samples.
4. 100% and 70 % EtOH (Sigma Aldrich) for RNA isolation from *in vitro* cell cultures and from clinical samples. 100% EtOH is diluted to 70 % with distilled and autoclaved water (dH2O).
5. Chloroform (Fluka) for RNA isolation with RNA-Bee solution.
6. 2-Propanol (Fluka) for RNA precipitation with the RNA-Bee method.
7. RNase free DNase (Qiagen) to eliminate genomic DNA.
8. Spectrophotometer NanoDrop, ND-1000 with software v3.3.1 to measure RNA concentration at 260nm.
9. SuperScript™ II Reverse Transcriptase kit (Invitrogen) for cDNA preparation. The kit contains 0.1M DTT, 5X First-Strand Buffer (250 mM Tris-HCl, pH 8.3; 375 mM KCl; 15 mM MgCl2) and Moloney murine leukemia virus reverse transcriptase (MMLV RT 200 U/µl), stored at -20°C.
10. 10 mM dNTP's 1:1:1:1 (Roche) for cDNA synthesis.
11. T12 primer (Invitrogen) for cDNA synthesis.
12. RiboLock RNase Inhibitor (Fermentas) for cDNA synthesis.
13. DNA Thermal cycler, Perkin Elmer 480.

2.3 Quantitative PCR

1. Power SYBR-green master mix (Applied Biosystems) for Q-PCR and for product detection.
2. Gene specific and control gene primers (Invitrogen) for Q-PCR (manually designed).
3. MicroAmp fast optical 96-well reaction plate with barcode (0.1 ml) (Applied Biosystems).
4. MicroAmp optical adhesive film, PCR compatible, DNA/RNA/RNase free (Applied Biosystems) to cover the plate in order to avoid evaporation of the samples.
5. 7900HT Fast Real-Time PCR system with SDS 2.3 software (Applied Biosystems).

2.4 5' RLM-RACE

1. GeneRacer Kit, for full-length RNA ligase-mediated rapid amplification of 5' and 3' cDNA ends (RML-RACE), (Invitrogen).
2. 95%, 70 % EtOH and dry ice.
3. Gene specific primers to perform 5' RLM-RACE, both for reverse transcribing mRNA and for the different PCR steps.
4. Biometra T1 Thermocycler (Westburg) for PCR.
5. Instead of Platinum Taq DNA polymerase high fidelity as suggested in the kit, QiaTaq DNA polymerase (Qiagen) can be used in the first PCR reaction "amplifying cDNA ends".
6. GoTaq flexi DNA polymerase (Promega) in case of nested PCR.

2.5 Sequencing

1. Bigdye terminator (BDT) v3.1 cycle sequencing kit (Applied Biosystems).
2. Biometra, T1 Thermocycler (Westburg).
3. 3130xl genetic analyzer and software for data collection v3.0 (Applied Biosystems).
4. MicroAmpTM optical 96-well reaction plates for sequence reactions (Applied Biosystems).
5. Gene specific primers (3.2 pmol).
6. Sodium Acetate (3M stock solution; pH 4.6) for DNA precipitation.
7. 95% and 70% EtOH for DNA precipitation and washing.
8. Formamide (Applied Biosystems) to dissolve the precipitated DNA.
9. DNAMAN and ChromasLite software for sequence analysis.

3. Methods

mRNA expression of *ERG* or another ETS-factor gene can be measured by Q-PCR in clinical prostate cancer samples to identify overexpression of the gene. In this chapter product detection with SYBR-green dye is described. A fusion event is indicated by low expression of the first exons of the gene compared to high expression of the last exons of the same gene. This is because the gene fusion results in overexpression of the essential protein coding part of the ETS oncogene, regulated by the fusion partner, *TMPRSS2* or other.

If overexpression of (part of) an ETS mRNA is detected, the sample can be tested for fusion transcripts with previously described fusion partners by standard RT-PCR, using a forward primer in the ETS part and a reverse primer in the candidate fusion partner. If none of the known fusion partners is detected, this can be an indication of a novel fusion partner.

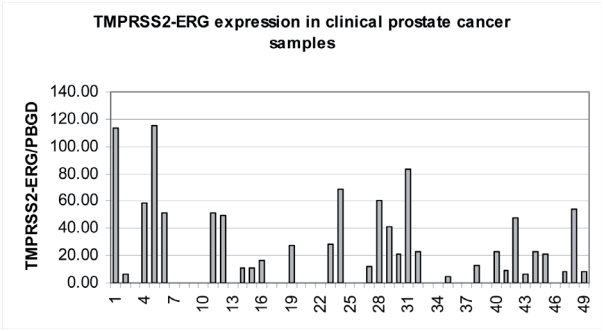
Taking the advantage of the fact that the 3' end of the fusion transcript is a known ETS mRNA, it is possible to identify the unknown fusion partners by 5' RLM-RACE, followed by sequencing of an amplified product. If a novel fusion partner is detected its frequency of expression can be studied by Q-PCR in a clinical sample cohort with fusion specific primers. Figure 2A shows an example of the analysis of *TMPRSS2- ERG* fusion transcript expression in prostate cancers.

Q-PCR can also be used to study changes in expression of a fusion gene in *in vitro* propagated cell lines under varying experimental conditions. This is important for studying the properties of a novel found fusion partner. In this chapter we show as an example the upregulation of *TMPRSS2* mRNA expression by incubating LNCaP prostate cancer cells in the presence of R1881 (Fig. 2B). Note that LNCaP cells donot harbor the*TMPRSS2-ERG* gene fusion.

3.1 Cell-culture and clinical prostate cancer samples

1. To study androgen regulation of *TMPRSS2*, LNCaP cells are grown in two T25 cell-culture flasks to 50% confluence in Dulbecco's modified Eagle's medium (DMEM) supplemented with 5 % FBS, and PS in a humidified incubator at 5% CO2 and at 37°C .
2. After reaching 50 % confluence the medium is replaced by DMEM supplemented with 5 % dextran-coated charcoal (DCC; dextran 0.1%, charcoal 1%) treated FBS (i.e. steroid-depleted).

A



B

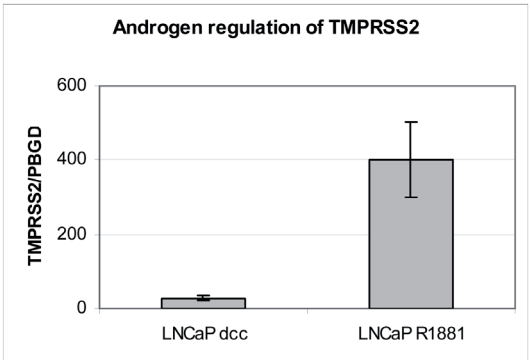


Figure 2A: Expression of *TMPRSS2-ERG* fusion transcripts in 50 clinical prostate cancer samples measured by Q-PCR relative to the housekeeping gene *PBGD*.

Figure 2B: Expression of *TMPRSS2* in LNCaP cells incubated with our without the synthetic androgen R1881 in DCC-supplemented medium for 24 h. The mRNA levels are depicted relative to *PBGD*.

3. 4 µl R1881 (10-6 M in EtOH) is added to one of the two T25 flasks in 4 ml medium, giving a final R1881 concentration of 10-9 M (*see* **Note 1**). To the second, control, flask 4 µl 100% EtOH is added (*see* **Note 2**). Next, the cell cultures are incubated for 24 h (see above).

4. Medium is removed and cells are washed with PBS and harvested by trypsinization. Trypsinized cells are transferred to 15 ml tubes and pelleted by centrifugation for 5 min at 1000g. RNA is isolated from the pellet as described in sections 3.2.1-3.2.4.

5. Clinical prostate cancer samples contain at least 50% tumor cells, as determined by a pathologist. To this end 5 µm tissue sections are cut and stained by routine hematoxiline/eosine (HE) staining. RNA is isolated from 20 consecutive 100 µm sections. Finally again a 5 µm section is cut and HE stained. RNA is isolated from the pooled tissue sections as described in sections 3.2.5-3.2.11.

3.2 RNA-isolation and cDNA preparation

1. For RNA-isolation from LNCaP cells the RNeasy Mini kit (Qiagen) is used and the RNeasy Mini Handbook followed under the section "Purification of total RNA from Animal Cells" (*see* **Note3**).
2. The pellet contains approx. 10^7 cells and the recommended buffer volumes for this number of cells are used.
3. Qiashredder tubes are used for homogenization of the lysed cells as the RNA yield is generally higher than by pipetting or passing the sample through a needle with a syringe.
4. A DNase treatment step is performed to eliminate possible DNA contamination.
5. For RNA-isolation of fresh-frozen clinical prostate cancer samples RNA-Bee (*see* **Note 4)** (BIO Connect) is used according to the protocol provided by the supplier.
6. The tissue is homogenized in a Polytron homogenizer in RNA-Bee solution with 1ml solution per 50 mg tissue.
7. 0.2 ml chloroform (*see* **Note 5)** is added to every 1 ml sample (homogenized tissue in RNA-Bee solution), vigorously shaken for 30 sec and kept on ice for 5 min.
8. The solution is then centrifuged at 12,000g for 15 min at 4° C. This separates the RNA to an aqueous top phase, which is transferred to a new tube.
9. RNA is precipitated by adding one sample volume 2-propanol and incubating at RT for 15 min.
10. Next, centrifuge to pellet the RNA at 12,000g for 5 min at 4° C. Wash the pellet with 1 ml 75% EtOH, centrifuge at 7,500g for 5 min at 4° C.
11. The pellet is air-dried and finally dissolved in 50 µl distilled and autoclaved water (dH2O) by pipetting.
12. cDNA is prepared from 2 µg total RNA.
13. dH2O is added to the RNA to a volume of 18 µl. 2 µl 100 ng/µl T12 primer is added and the sample is denaturated at 70°C for 10 min in the thermal cycler.
14. The sample is cooled on ice for 2 min and spinned down by a short centrifugation to collect the fluid at the bottom of the tube. Following reagents are added on ice: 8 µl 5x first strand buffer, 4 µl 0.1 M DTT, 2 µl dNTP's, 0.5 µl RiboLock RNase Inhibitor and 2 µl M-MLV RT (400 U).
15. The mixture is then incubated for 1 h at 37°C followed by 10 min at 95°C in a DNA thermal cycler. After the incubation samples are cooled on ice for at least 2 min, centrifuged shortly to collect fluid and then stored at -20°C.

3.3 Quantitative PCR

1. To facilitate pipetting and the analysis of a larger scale Q-PCR experiment, a scheme of a 96-well plate where samples are pipetted is made before the start of the experiment. In addition to the samples that will be analyzed, a serial dilution of (mixture of) standard samples is added to each plate for preparation of a standard line. Other controls are a minus RT sample and a genomic DNA sample. Expression of each gene is calculated relative to the expression of a house keeping gene. So, all samples are tested also for the expression of this gene.
2. A standard line is composed of five serial four-fold dilutions of a mix of cDNA samples. Five µl of four different samples expressing the gene of interest is mixed and diluted 10 times to 200 µl, this is standard 1. Values for standard line are set to 25600 for standard 1, 6400 for standard 2 and fourfold decreased till 100 for standard 5. These pre-set values in the standard line are later used to quantify the samples.
3. A primer mix is prepared from two 100 µM primer stock solutions by adding 10 µl of each primer to 580 µl dH20. For primer design see **Notes 6, 7 and 8**.
4. It is advisable to run a melting curve program with each primer set and the standard line. This gives peaks at the melting temperature of the samples for each different set of primers. Primers that anneal to each other (primer-dimers) can be detected because they will give a peak at a lower temperature than the product of interest. Normally in this experiment the standard line is run together with a water control and a DNA control as to make sure that the primers form a proper standard line (*see* **Note 9**), do not amplify DNA and do not form primer-dimers.
5. The housekeeping gene porphobilinogen deaminase (PBGD) is chosen as an internal control. Thus, PCR reactions are carried out with all cDNA samples and the standard line using PBGD primers (*see* **Note 10**). The PCR values of all samples are divided by their corresponding PBGD value to adjust for small differences in cDNA levels present between the samples.
6. cDNA is diluted 20 times and 5 µl diluted cDNA is added per well in the 96-well plate.
7. 5 µl of the five serially diluted standard samples is added to the plate for every primer pair.
8. 5 µl of the primer mix prepared in section 3.3.3 is added to each well.
9. 12.5 µl Power SYBR-green master mix (Applied Biosystems) and 2.5 µl dH2O is added to each well. cDNA, primer mix and SYBR-green PCR mix together give a total volume of 25 µl. (*see* **Note 11**)

10. The 7900HT Fast Real-Time PCR system from Applied Biosystems with accompanying software SDS 2.3 is used for amplification. The exact settings are variable depending on the primer annealing temperature (*see* **Note 12**). A PCR program can be designed as follows:
Step 1. 5 min at 95 °C
Step 2. 15 sec at 95 °C
Step 3. 30 sec at 60 °C
Step 4. 30 sec at 72 °C, steps 2 to 4 is repeated 39 cycles.
11. Analysis is performed using SDS 2.3 software. First the standard line and control points are defined. A threshold that will define the Ct values is determined (*see* **Note 13**). The quantification of the sample is automatically calculated by the program during analysis by comparing the Ct values (*see* **Note 14**) with those of the standard line and its pre-set values.

3.4 5'-RACE

1. GeneRacer Kit, for full-length, RNA ligase-mediated rapid amplification of 5'- and 3'-cDNA ends (RML-RACE), (Invitrogen) is used and the manual provided by the supplier is followed with minor modifications.
2. In the first step, "dephosphorylating RNA", 5µg total RNA is used.
3. For the reverse transcription step a gene specific primer is used (2 pmol).
4. The first PCR step "Amplifying cDNA ends" is done according to the protocol, but using QiaTaq DNA Polymerase and 10x PCR buffer (Qiagen), dH2O is added to a final volume of 50 µl.
5. The amplified products are separated over a 1 % agarose gel by electrophoresis. Clear bands can be isolated for sequencing (see paragraph 3.5 Sequencing). Often the product of this step appears as a smear and a second, nested PCR, is needed to increase the specificity and the amount of product.
6. A nested PCR is performed as described in the protocol but using GoTaq flexi DNA polymerase (Promega) and 5x GoTaq flexi buffer. The buffer volume is then increased two times and dH2O added to a final volume of 50 µl (*see* **Note 15**).

3.5 Sequencing

1. 5'-RACE PCR products for sequencing are isolated from the agarose gel using QIAquick Gel Extraction kit 250 (Qiagen).
2. Check the yield and quality of DNA after isolation by measuring the absorbance at 260 nm and analyzing part of the sample on an agarose gel. The DNA is added to the reaction mix described in step 3.5.3 as final step.
3. Add on ice 1 µl BDT and 4 µl 5x sequencing buffer provided by the Bigdye terminator v3.1 cycle sequencing kit (Applied Biosystems). 1 µl of the gene specific sequence primer (3.2 pmol) is added, dH2O is added to a final volume of 20 µl.
4. The Biometra T1 Thermocycler (Westburg) is programmed for the sequencing reaction as follows:
Step 1. 1 min at 96°C
Step 2. 10 sec at 96°C
Step 3. 5 sec at 50°C
Step 4. 4 min at 60°C, step 2 to 4 is repeated 24 cycles
The samples are kept at 4-8 °C until continuing with precipitation.
5. In 1.5 ml eppendorf tubes 13 µl dH2O, 3 µl 3 M NaAc (pH 4.6) and 64 µl 95 % EtOH is added to the 20 µl sequence reaction and product is precipitated at RT for 15 min up to max 24 h.
6. The orientation of the tube is marked (since pellet might be hard to visualize) and tubes are spinned at maximum speed in an Eppendorf centrifuge for 20 min at RT. Discard the supernatant and wash the pellets with 250 µl ice-cold 70 % EtOH, place the tubes in the same orientation as before and spin for 10 min at maximum speed in the Eppendorf centrifuge. Aspirate the supernatant carefully and leave the tubes to dry a few min at RT (*see* **Note 16**). Resuspend the pellets in 20 µl formamide.
7. Add the samples to a MicroAmpTM optical 96-well reaction plate (Applied Biosystems) and analyze in an Applied Biosystems 3130xl genetic analyzer.
8. The sequence can be analyzed using different available software programs. Here ChromasLite and DNAMAN are used (*see* **Note 17**).

Notes

1. R1881 is light sensitive and is stored in the dark at – 20 °C.

2. Do not use more than 0.3% EtOH in the cell culture, because this might be toxic to the cells.

3. It is highly important to keep all tubes and samples on ice and work as efficiently as possible to preserve RNA from degrading. During RNA isolation and handling it is very important to work RNase-free, wear gloves and use autoclaved/sterilized materials and solutions.

4. When working with RNA-Bee, use gloves and eye protection. The RNA-Bee solution contains guanidine thiocyanate and phenol that is an irritant and causes skin burns.

5. Chloroform can dissolve disposables. Most grades polythene and polypropylene are resistant but most other disposables are not (e.g. PET, perspex, PVC, polystyrene). However, it is good practise to test all disposables before use in RNA isolation.

6. There are commercial programs for design of primers. These programs take into consideration most important parameters. It is also possible to design primers yourself. The two primers of a PCR are complementary to one of each DNA strands of the amplified fragment. Long primers increase the specificity but also increase annealing time. Most appropriate PCR primers have a length of 18-20 nucleotides.

The GC content of the primers should be 40-60% because it affects the annealing temperature and specificity of the primers. Interaction between the two primers should also be avoided (forming of primer-dimers).

7. If the product yield is unexpectedly low and no primer-dimer is seen in the optimization step, the cause can still be in primer design. Primers might form secondary structures. How prone primers are to form secondary structures is difficult to predict, but there is software that can predict and calculate the ΔG, G standing for the Gibbs free energy. The ΔG represents the energy needed to break a secondary structure, the acceptable values are between 2-6 kcal/mol dependent on the type of secondary structure. The GC content, size and structure of the amplified fragment are also important aspects of PCR efficacy.

8. Primer sequences used:

TMPRSS2 exon 1 forward: GAGCTAAGCAGGAGGCGGA

TMPRSS2 exon 3 reverse: AGGGGTTTTCCGGTTGGTATC

ERG exon 4 reverse: CATCAGGAGAGTTCCTTGAG

PBGD forward: CATGTCTGGTAACGGCAATG

PBGD reverse: GTACGAGGCTTTCAATGTTG

9. Check the slope of the standard line and the R^2 value. The slope should not deviate much from -3.0 and R^2 should be close to1.0.

10. Although a housekeeping gene by definition is supposed to be expressed at equal levels and by all cells this might not always be completely correct. Therefore, more than one housekeeping gene should be tested for every tissue or cell line studied to distinguish that the expression levels do not vary significantly over the sample cohort.

11. Mix the PCR reaction by pipetting and spin the plate before performing the QPCR, this doesn't only collect the liquid to the bottom of the well but also helps to decrease air bubbles in the sample.

12. The design of the Q-PCR will be dependent on the design of primers, the equipment, the software and manufacturer's advice.

13. Threshold is the allowed background level. It is important to use the same threshold if several plates/runs are needed for an experiment. This allows comparing the results. In the experiments described in this chapter the threshold is always manually set to 0.02.

14. Ct value is a value defining at which of the amplification cycles the fluorescence signal crosses the pre-set threshold value.

15. If there are several fragments (bands) from the (nested) PCR reaction, sequence all bands. The bands can be specific and represent different transcripts present in the sample e.g. different splice-forms or alternative first exons.

16. Do not over-dry the sample because this can cause dye-blobs and disturb reading of the sequence.

17. ChromasLite is free software available on internet. DNAMAN can be purchased from several different providers.

References:

1. Tomlins SA, Rhodes DR, Perner S*, et al.* (2005) Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* **310**:644-8
2. Hermans KG, van Marion R, van Dekken H, Jenster G, van Weerden WM, Trapman J. (2006) TMPRSS2:ERG fusion by translocation or interstitial deletion is highly relevant in androgen-dependent prostate cancer, but is bypassed in late stage androgen receptor-negative prostate cancer. *Cancer Res* **66**:10658-63.
3. Perner S, Demichelis F, Beroukhim R*, et al.* (2006) TMPRSS2-ERG fusion-associated deletions provide insight into the heterogeneity of prostate cancer. *Cancer Res* **66**:8337-41.
4. Wang J, Cai Y, Ren C, Ittmann M. (2006) Expression of variant TMPRSS2/ERG fusion messenger RNAs is associated with aggressive prostate cancer. *Cancer Res* **66**:8347-51.
5. Pflueger D, Rickman DS, Sboner A, *et al.* (2009) N-myc downstream regulated gene 1 (NDRG1) is fused to ERG in prostate cancer. *Neoplasia* **11**:804-11.
6. Han B, Mehra R, Dhanasekaran SM, *et al.* (2008) A fluorescence in situ hybridization screen for E26 transformation-specific aberrations: identification of DDX5-ETV4 fusion protein in prostate cancer. *Cancer Res* **68**:7629-37.
7. Tomlins SA, Mehra R, Rhodes DR*, et al.* (2006) TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. *Cancer Res* **66**:3396- 400.
8. Helgeson BE, Tomlins SA, Shah N*, et al.* (2008) Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. *Cancer Res* **68**:73-80.
9. Tomlins SA, Laxman B, Dhanasekaran SM*, et al.* (2007) Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. *Nature* **448**:595-9.
10. Seth A, Watson DK. (2005) ETS transcription factors and their emerging roles in human cancer. *Eur J Cancer* **41**:2462-78.
11. Turner DP and Watson DK. (2008) ETS transcription factors: oncogenes and tumor suppressor genes as therapeutic targets for prostate cancer. Expert. Rev. *Anticancer Ther* **8**:33-42.
12. Lapointe J, Kim YH, Miller MA, *et al.* (2007) A variant TMPRSS2 isoformand ERG fusion product in prostate cancer with implication for molecular diagnosis. *Mod Pathol* **20**:497-73.
13. Saramaki OR, Harjula AE, Martikainen PM, *et al.* (2008) TMPRSS2:ERG fusion identifies a subgroup of prostate cancers with a favourable prognosis. Clin *Cancer Res* **14**:3395-400.
14. Demichelis F, Fall K, Perner S, *et al.* (2007) TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. *Oncogene* **26**:4596-9.
15. Maher CA, Palanisamy N, Brenner JC *et al.* (2009) Chimeric transcript discovery by paired-end transcriptome sequencing. *Proc Natl Acad Sci USA* **106**: 12353–8.
16. Jhavar S, Reid A, Clark J *et al*. (2008) Detection of *TMPRSS2-ERG* translocations in human prostate cancer by expression profiling using GeneChip human exon 1.0 ST Arrays. *J Mol Diagn* **10**:50–7.

CHAPTER 6

A 36 gene-signature of primary prostate tumor samples with ERG overexpression selects for subgroups with poor clinical progression

Delila Gasi Tandefelt[1], Joost L. Boormans[2], Hetty A. van der Korput[1], Guido W. Jenster[2] and Jan Trapman[1]

Departments of Pathology[1] and Urology[2], Josephine Nefkens Institute, Erasmus University Medical Centre, Rotterdam, The Netherlands

Submitted for publication

Statement of translational relevance

Overtreatment of primary prostate cancer is a major clinical problem. Selection of patients that should be treated and patients that do not need therapy yet, is of utmost importance. Knowledge of genes that are associated with tumor progression might be very helpful in selection of patients that should receive therapy. Results of gene expression profile experiments of prostate cancer published so far have been disappointing. In the present manuscript we identified within the group of TMPRSS2-ERG positive tumors a significant subgroup of patients with poor clinical outcome. By further bioinformatics approaches we were able to define a 36-gene classifier that discriminates a patient group with rapid clinical progression from a group with slow progression. These data were validated in a completely independent patient cohort.

Abstract

Purpose: The molecular basis of the clinical heterogeneity of prostate cancer is not well understood. Our purpose is to identify and characterize genes in a clinically relevant gene expression signature in a *TMPRSS2-ERG* positive subgroup of primary prostate cancers.

Experimental Design: We studied gene expression profiles in 48 primary prostate cancers from patients with a long clinical follow-up by unsupervised hierarchical clustering. Results were correlated with clinical outcome and validated in an independent patient cohort. Selected genes from a defined classifier were *in vitro* tested for biological properties.

Results: Most tumors that overexpressed *ERG* clustered separately from other primary prostate cancers. No differences in any clinical end points between *ERG*-positive and *ERG*-negative cancers were detected. Importantly, within the *ERG*-positive samples, two subgroups were identified, which differed significantly in PSA recurrence-free survival, cancer-specific and overall survival. From our findings we defined a gene expression classifier of 36 genes. In a second completely independent tumor set the classifier also distinguished *ERG*-positive subgroups with different clinical outcome. In both patient cohorts the classifier was not predictive in *ERG*-negative tumors. Biological processes regulated by genes in the classifier included cell adhesion and bone remodeling. TGF-β signaling was indicated as main differing signaling pathway between the two *ERG*-subgroups. *In vitro* biological assays of two selected genes from the classifier (*INHBA* and *CDH11*) supported a functional role in prostate cancer progression.

Conclusion: The classifier identified can contribute to prediction of tumor progression in *ERG*-positive primary prostate tumors and might be instrumental in therapy decisions and decrease overtreatment.

Introduction

Prostate cancer is the most frequent cancer in men in most developed countries and the second cause of cancer death (1). A substantial proportion of diagnosed primary tumors will not or slowly progress from a localized tumor to a metastatic, lethal disease. These patients might benefit more from active surveillance and delayed treatment rather than from instant invasive treatments (2). To prevent overtreatment it is crucial to find biomarkers that are predictive for clinical progression. Currently used Gleason score in combination with clinical stage and serum PSA level have limited selectivity.

Molecular classification is instrumental in elucidation of the mechanisms by which cancers acquire the biological properties responsible for progression. In this way novel molecular patterns that are associated with clinical outcome or biological function in prostate cancer have been postulated (3-6). However, the composition of tissue samples, platforms and methods used for analysis and interpretation of data contribute to differences in signatures presented. One of the major challenges is the clinical and molecular heterogeneity that is observed in prostate cancer. No doubt part of this heterogeneity will be due to differences in genetic alterations in subsets of prostate cancer. The rapidly increasing knowledge of genomic alterations in prostate cancer (7-8) allows selection of tumor subgroups carrying specific, combinations of, genetic alterations. In this regard it is a major advantage if a specific genetic alteration is frequent and already present in early stages of tumor development, preferentially in a large patient cohort with long clinical follow-up.

Recurrent genomic rearrangements of the ETS transcription factor family occur in up to 70 % of all prostate cancers. Fusion of *TMRSS2* to *ERG* is responsible for approximately 50% of ETS gene fusions, representing the most frequent genomic alteration (9-10). Due to the gene fusion, *ERG* is positioned under regulation of the prostate-specific and androgen-regulated gene *TMPRSS2*. This leads to androgen-regulated overexpression of the ERG oncoprotein. The findings that *TMPRSS2-ERG* fusions and ERG overexpression can be found in PIN-lesions indicate that the gene fusion is an early event in prostate cancer development (11-13). Moreover, *ERG* is still overexpressed in late-stage castration-resistant prostate cancer (CRPC), indicative of its continuous important role. However, the function of *ERG* overexpression is not well understood and conflicting studies have been published on the relevance of *ERG* overexpression for clinical outcome (14-19).

In mouse prostate cancer models *ERG* overexpression induces development of precursor lesions of prostate cancer but is not sufficient for transformation to prostate cancer (20-21) pointing to additional genetic and epigenetic molecular alterations that cooperate

with *ERG* overexpression in prostate cancer development. In fact, such cooperation has been shown in mouse models (22-24). It is a challenge to identify the pathways that combined with *ERG* overexpression determine the progressive growth of clinical prostate cancer.

In the present study we defined in a subgroup of *ERG*-positive prostate cancer patients a signature gene expression profile that was associated with tumor progression. This classifier is composed of 36 genes. It was validated in a completely independent *ERG*-positive patient cohort. Interestingly, the gene set does not predict progressive disease in patients with *ERG*-negative prostate cancer.

Materials and Methods

Ethics statement

Use of clinical samples was approved by the Erasmus MC Medical Ethics Committee according to the Medical Research Involving Human Subjects Act in protocol MEC-2004-261, entitled “The use of human normal and cancer residual tissue from a tissue bank for characterization of DNA, RNA and protein”.

Tissue samples, RNA isolation and cDNA preparation

The initial treatment of primary tumors was radical prostatectomy in all the cases. Hematoxilin/eosin stained tissue sections were histologically evaluated by two pathologists (T. van der Kwast; G.J.H.L. van Leenders). Only samples that contained at least 50% tumor cells were selected for analysis.
RNA from clinical specimens was isolated using RNA-Bee (Campro Scientific, Berlin, Germany). DNA was isolated using the DNeasy DNA Extraction kit (Qiagen, Valencia, CA, USA). RNA from cell lines was isolated using the RNeasy RNA Extraction kit (Qiagen).

Gene-expression arrays

RNA samples for microarray analysis were processed according to the protocol of the array supplier (Affymetrix, Santa Clara, CA, USA). Expression profiles were determined using GeneChip Human Exon 1.0 ST (Affymetrix) at the center of Biomics, Erasmus Medical Centre, Rotterdam and at ServiceXS, Leiden, The Netherlands. Microarray data were processed and RMA quantile normalized using Partek Genomics Suite (St. Louis, MO, USA). For further normalization on gene level values below 30 were set to 30. The samples were normalized to geometric mean and then log2 values of the samples were calculated. Gene Cluster from Stanford microarray resources and TreeView were used for unsupervised hierarchical average linkage clustering. Significance of microarrays (SAM) was used to identify differentially expressed genes between subgroups A and B (Fig. 1) and subgroups BI and BII (Fig. 2). False discovery rate (FDR) was set to zero. BRB array tool Class Prediction analysis was used for selecting the classifier and to validate the classifier in an independent patient cohort.

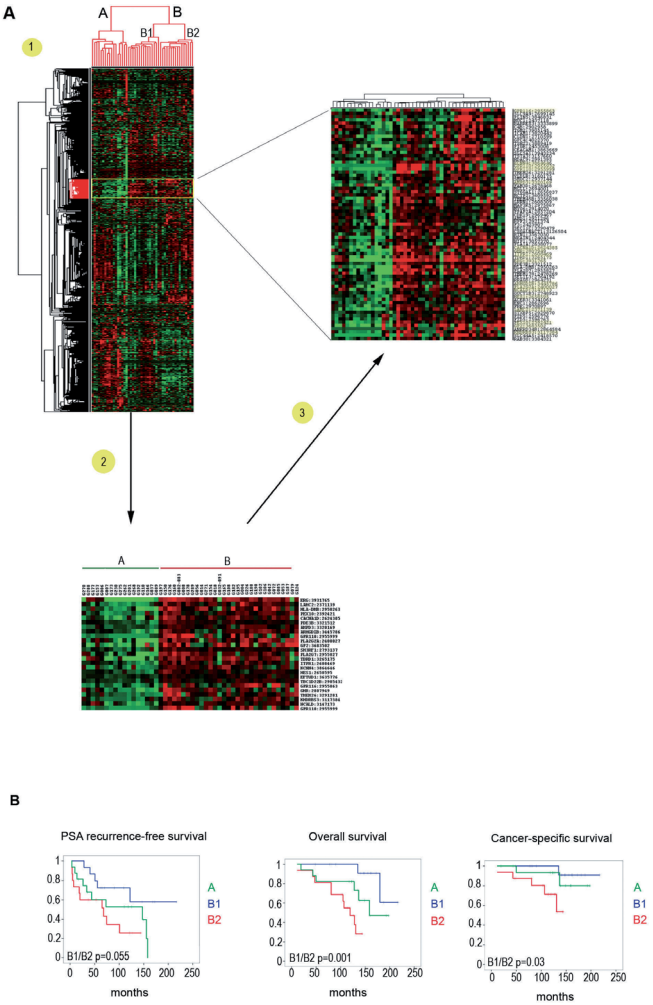


Fig 1. Gene expression and survival characteristics of primary prostate cancer patients. **A.1.** Gene Cluster (Stanford microarray resources) and TreeView were used for unsupervised hierarchical average linkage clustering of microarray data determined by Affymetrix GeneChip Human Exon 1.0 ST. The data of 48 primary prostate cancer samples were normalized to geometric mean and log2 ratios for the samples are represented in the heat-map. Two main clusters A and B are separated, furthermore group B contains the distinct subgroups B1 and B2. **A.2.** Significance of microarrays (SAM) was used to identify differentially expressed genes between group A and B with false discovery rate (FDR) set to zero. A heat-map of the 25 top differentially expressed genes was generated in TreeView and Gene Cluster. *ERG* was the most significantly differentiating gene between groups A and B. **A.3.** A zoomed-in view of the *ERG*-cluster. The top 25 genes calculated by SAM are marked in yellow. **B.** Kaplan-Meier curves assessing the probability of PSA recurrence free survival (p=0.055), overall survival (p=0.001) or cancer specific survival (p=0.03) of subgroups A, B1 and B2 as a function of time after surgery . The differences between the survival curves of the groups B1 and B2 were tested using the log-rank test. Statistical analyses were done using the Statistical Package for Social Sciences, version 17.0 (SPSS), with a significance level of 0.05 (two-tailed probability).

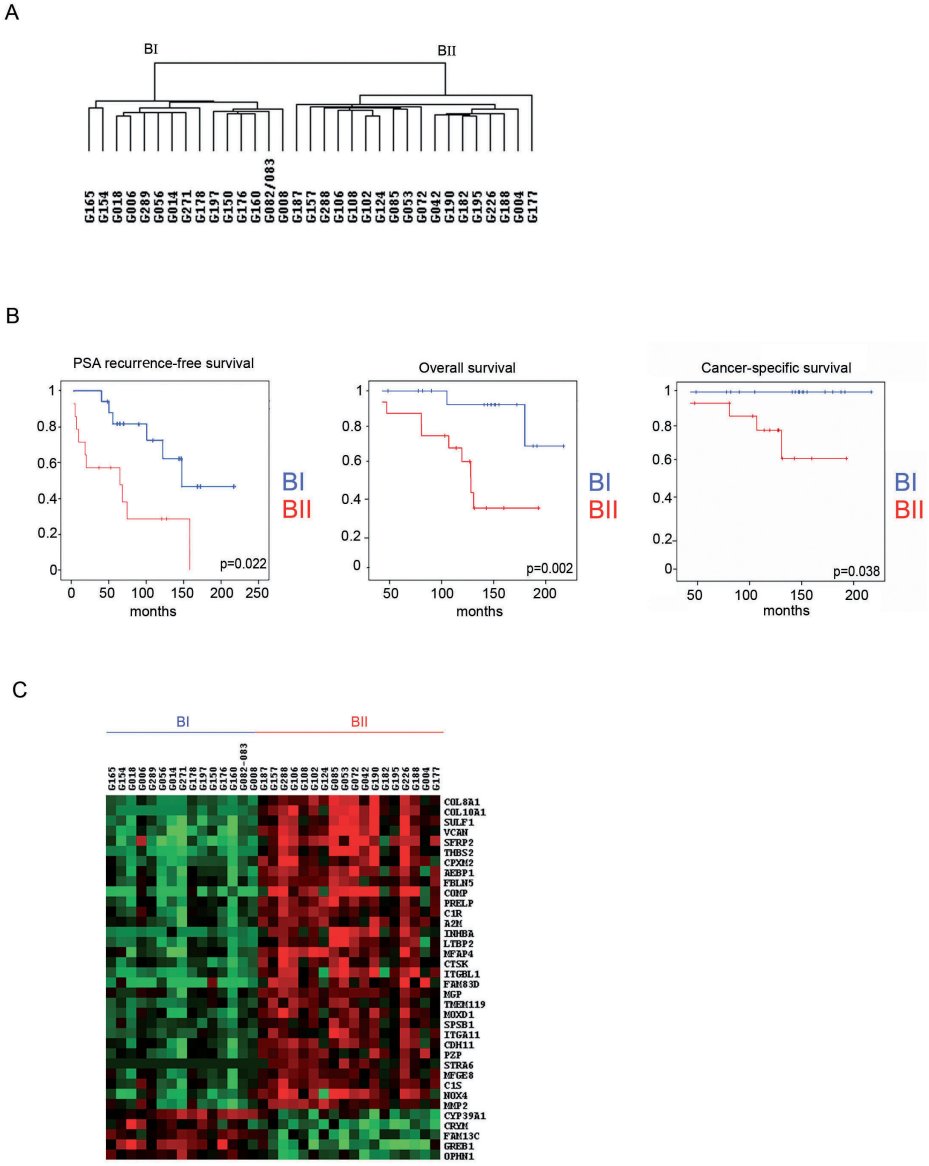


Fig 2. Gene expression and survival characteristics of prostate cancer patients overexpressing ERG. **A.** Unsupervised hierarchical average linkage clustering of all 33 *ERG*-positive prostate cancer samples separate into groups BI and BII. The four *ERG*-positive samples from group A (Fig. 1) are in group BI (G006) and group BII (G157, G177 and G288), respectively. **B.** Kaplan-Meier curves assessing the probability of PSA recurrence free survival (p=0.022), overall survival (p=0.002) or cancer specific survival (p=0.038) as a function of time after surgery between groups BI and BII. BI is correlated with good prognosis (GP) and BII with poor prognosis (PP). **C.** Brb array tool Class Prediction analysis was used for selecting the genes that predict groups GP and PP as described in Patients and Methods. Expression of the 36 genes in the resulting classifier is presented in a heat-map.

Lentiviral production and infection

To study CDH11 and INHBA function by overexpression, lentiviruses producing these proteins where generated. By cotransfection of HEK293T cells with pLenti-CMV-h-CDH11 (ABM Inc., Richmond, Canada), or Precision LentiORF INHBA (Abgene, Epsom, UK), and pPAX2 and pMD2.G (Didier Trono) using the calcium phosphate precipitation method. Similarly, control lentivirus, producing GFP was generated from pWPXLd-GFP (Didier Trono). Virus was harvested from the supernatant and used for infection of PNT2C2 cells. Pools of infected cells were propagated and used in biological assays.
To study CDH11 and INHBA function by selective downregulation, lentiviruses producing shRNA directed against these mRNAs were generated in HEK293T cells by cotransfection of pLKO-shCDH11 (sh5700, TRCN0000054335), pLKO-shINHBA (sh5704, TRCN0000059264) or pLKO control vector (Sigma,St. Louis,USA) and pPAX2 and pMD2.G. Viruses where collected and used for infection of PC3 cells.

Western blot analysis

For Western blot analysis, PNT2C2 cells were infected with CDH11 lentivirus, precision LentiORF INHBA lentivirus, or control GFP lentivirus. Similary, PC3 cells were infected with CDH11-shRNA, INHBA-shRNA or control lentivirus. Cells were harvested after 48 h. Western blot analysis was carried out using standard procedures using INHBA- (Novus Biologicals, Littleton, USA) and CDH11- (Life Technologies, Paisley, UK) and loading control β-actin (Sigma) antibodies. Protein bands were visualized by chemiluminescence (Pierce, Rockford, USA).

Migration assay

Migration assays of PC3-sh5700, PC3-sh5704, and control PC3-pLKO cells (1×10^5 cells per well) were performed according to the instructions of the manufacturer of the QCM colorimetric cell Migration Assay kit (Millipore, Billerica, USA). The assay was terminated after 24 h incubation.

Soft-agar assay

A layer of 0.6% low-melting agarose in standard culture medium was prepared in six-well plates. On top, a layer of 0.3% agarose containing $1x10^4$ PNT2C2 cells overexpressing CDH11 or INHBA or control PNT2C2-GFP cells was plated. Similarly, PC3 cells infected

with sh5700, sh5704 or control PC3 cells in 0.3% agarose were plated with the same approach. At day 14, cells were stained with crystal violet and colonies were counted.

Statistical analysis

Associations between clinical and histopathological variables and different clusters of prostate cancer patients were evaluated by the Pearson's χ^2 test, the Mann-Whitney U test, or Kruskal-Wallis test, where appropriate. Gene expression of the subgroups were correlated with the end points, biochemical progression-free survival, which was defined as time from radical prostatectomy to date of biochemical recurrence, and cancer-specific and overall survival. Biochemical recurrence was defined as (*a*) a prostate-specific antigen level of >0.2 ng/ml at two consecutive measurements with a 3-month interval if the prostate-specific antigen nadir was <0.1 ng/ml or (*b*) a prostate-specific antigen nadir of ≥0.2 ng/ml. Patients who died from causes other than prostate cancer or were lost to follow-up were censored at the date of last prostate-specific antigen test. Kaplan-Meier curves were constructed to visualize the probability of remaining free of biochemical recurrence or death as a function of time after surgery. The differences between the survival curves of the groups were tested using the log-rank test or Breslow method if appropriate. Statistical analyses were done using the Statistical Package for Social Sciences, version 17.0 (SPSS), with a significance level of 0.05 (two-tailed probability).

Results

Unsupervised clustering of gene expression profiles of primary prostate cancer samples identified a subgroup with poor prognosis

To distinguish potentially important patterns in progression of prostate cancer, gene expression profiles of 48 primary tumors from a patient cohort with long clinical follow-up (mean follow-up over 10 years) were analyzed. First, the normalized micro-array data was used in unsupervised hierarchical clustering. By this approach two main clusters could be separated, denoted A and B (Fig. 1A). Within the B-group two distinct gene-expression profiles generated a sub-grouping of the tumors (B1 and B2, Fig. 1A).

To gain knowledge of the genes that were differentially expressed SAM was performed, comparing the main clusters A and B (Fig. 1A and Suppl. Table S1). Remarkably, *ERG* turned out to be the strongest differentially expressed gene between the two clusters, indicative of a dominant role of *ERG* in a subgroup of primary prostate cancers. In addition, most of the top 20 genes in the SAM experiment were in the same cluster as *ERG* (Fig. 1 and Suppl. Table S1). This cluster not only contained many genes that are known to be associated with *ERG* overexpression from previous studies, including *CACNA1D, LAMC2, TDRD1* and *NCALD* (24-26).

ERG overexpression was detected in almost 70% of the prostate cancer samples in the cohort studied. QPCR confirmed that with two exceptions (G32/91 and G79) all samples in cluster B showed *ERG* over-expression. In all cases *ERG* over-expression was due to *TMPRSS2-ERG* fusion transcripts (data not shown). Cluster A included four samples showing *TMPRSS2-ERG* over-expression (G006, G157, G177 and G288). The array data indicated that these four samples also had high expression of other genes that were strongly correlated with *ERG*. Overall, however, the expression pattern of these four samples seemed to be more related to the samples in group A. Intra-tumor heterogeneity or a mixture of two or more independent tumor foci containing *ERG*-positive and *ERG*-negative cells might be the cause of the apparent discrepancies.

To investigate differences in clinical outcome between patients in A and B, Kaplan-Meier analysis was performed for biochemical progression-free survival (serum PSA levels), overall survival and cancer-specific survival. No significant differences were detected between both groups (Supplementary Fig. S1). In addition, no significant differences in clinicopathological characteristics were observed between clusters A and B (Suppl. Table S2). However, stratifying cluster B into B1 and B2, revealed that patients in cluster B2 had significantly poorer outcome in terms of overall and cancer-specific survival than cluster B1 patients ($p = 0.001$ and $p = 0.03$, respectively; Fig 1B). A borderline significant

difference in biochemical (serum PSA level) progression was detected between B1 and B2 (Fig. 1B). No significant differences in clinicopathological characteristics were observed between clusters B1 and B2 (Suppl. Table S3).

Gene expression profiling of ERG-positive prostate cancer

Next, we focused on all 33 *ERG*-positive tumors (29 tumors from group B and four from group A). We repeated the unsupervised hierarchical clustering for these ERG-positive samples. As expected, again two clusters (BI and BII) were identified (Fig. 2A). The four samples originating from group A were now in cluster BI (G006) and BII (G157, G177 and G288) (Fig. 2A). Again, Kaplan-Meier analysis showed cluster BII to have significantly poorer outcome than BI in terms of overall survival (p=0.002) and cancer specific survival (p=0.038) and now also in biochemical progression-free survival (p=0.022) (Fig. 2B). In this comparison, tumor samples of cluster BII had borderline significantly higher Gleason scores (p = 0.048) and higher pT-stages (p = 0.06) (Suppl. Table S4).

Identification of a molecular signature that predicts poor prognosis in ERG positive samples

The clear different clinical progression of patients in *ERG*-subgroups BI and BII suggested that gene expression data could be applied for the identification of a prognostic signature within this at least initially genetically homogeneous *ERG* positive group.

The BRB-Stanford array tool was used for further statistical analysis of the expression data. The BII and BI groups were renamed in PP (poor prognosis) and GP (good prognosis), respectively. A class prediction analysis was performed to identify the classifier for the samples. Genes different between the two classes at 1e-05 significance level were used for class prediction and the leave-one-out cross validation method was used to predict the misclassification rate based on 1000 permutations. The full classifier was composed of 45 genes, 9 of these genes had a CV-support lower than 100% and were removed from the classifier. So, 36 genes remained that combined could separate the PP group from the GP group in samples showing *ERG* overexpression (Fig. 2C, Table 1).

Validation of the classifier in an independent cohort of primary prostate tumors

Next we tested the performance of the classifier in a recently described completely independent data cohort of primary prostate cancers (27).

The validation cohort was composed of 131 primary tumors, 19 metastases, 29 adjacent normal samples and 6 prostate cancer cell-lines. We selected expression array data derived from primary prostate cancers. Patients whose initial treatment was

Table 1: Classifier genes and their function

Gene	Name	Function
NM_001850 , COL8A1	alpha 1 type VIII collagen	cell communication, ECM-interaction, focal adhesion
NM_000493, COL10A1	type X collagen alpha 1	
NM_015170, SULF1	sulfatase 1	cell communication
NM_004385, VCAN	versican	cell adhesion, proliferation, migration and angiogenesis
NM_003013, SFRP2	secreted frizzled-related protein 2	wnt signalling pathway, development
NM_003247, THBS2	thrombospondin 2	cell communication, ECM-receptor interaction, focal adhesion, TGF-β signalling pathway, angiogenesis
NM_198148, CPXM2	carboxypeptidase X (M14 family), member 2	
NM_001129, AEBP1	Adipocyte enhancer-binding protein 1	
NM_006329, FBLN5	fibulin 5	ECM-interaction, focal adhesion, vascular development
NM_000095, COMP	cartilage oligomeric matrix protein	cell communication, focal adhesion, TGF-β signalling pathway, bone remodelling
NM_002725, PRELP	proline/arginine-rich end leucine-rich repeat protein	
NM_001733, C1R	complement component 1, r subcomponent	
NM_000014, A2M	alpha-2-macroglobulin	protease inhibitor and cytokine transporter
NM_002192, INHBA	inhibin beta A	TGF-β signalling pathway
NM_000428, LTBP2	latent transforming growth factor beta binding protein 2	TGF-β signalling pathway, cell adhesion
NM_002404, MFAP4	microfibrillar-associated protein 4	Cell adhesion
NM_000396, CTSK	cathepsin K	bone remodeling and resorption
NM_004791, ITGBL1	Integrin beta-like protein 1	angiogenesis
NM_030919, FAM83D	Spindle protein CHICA	Cell cycle
NM_001190839, MGP	matrix Gla protein	Inhibitor of bone formation
NM_181724, TMEM119	Transmembrane protein 119	
NM_015529, MOXD1	monooxygenase, DBH-like 1	
NM_025106, SPSB1	splA/ryanodine receptor domain and SOCS box containing 1	Protein modification; protein ubiquitination
NM_001004439, ITGA11	integrin, alpha 11	ECM-receptor interaction, focal adhesion, regulation of actin cytoskeleton, receptor for collagen
NM_001797, CDH11	OB-cadherin	cell-cell adhesion, bone development and maintenance
NM_002864, PZP	Pregnancy-zone protein	protease inhibitor
NM_001199041, STRA6	stimulated by retinoic acid gene 6 homolog	Retinol metabolism
NM_005928, MFGE8	milk fat globule-EGF factor 8 protein	
NM_201442, C1S	complement component 1, s subcomponent	
NM_016931, NOX4	NADPH oxidase 4	signal transduction, cell differentiation and tumor cell growth
NM_004530, MMP2	matrix metallopeptidase 2	degrading extracellular matrix, remodeling of the vasculature, angiogenesis, tissue repair, tumor invasion, inflammation, and atherosclerotic plaque rupture
NM_016593, CYP39A1	cytochrome P450, family 39, subfamily A, polypeptide 1	drug metabolism and synthesis of cholesterol, steroids and other lipids
NM_001888, CRYM	crystallin, mu	binds thyroid hormone for possible regulatory or developmental roles
NM_198215, FAM13C	family with sequence similarity 13, member C	
NM_014668, GREB1	growth regulation by estrogen in breast cancer 1	
NM_002547, OPHN1	oligophrenin 1	Rho cell motility signalling pathway

not a radical prostatectomy were removed from further analysis. This left us with 127 samples. Fiftytwo tumors showed *ERG* over-expression, 75 samples were *ERG*-negative. The median follow-up of the cohort was almost five years. The expression data of the validation cohort was normalized using the same approach as for the original data set. The original *ERG*-positive samples were selected as the training set and *ERG* samples from the validation set as the dataset to be predicted. Using the BRB-array tool, we performed a class prediction test that sorted the samples from the independent data cohort according to class PP or GP, applying the same parameters as mentioned earlier and based on the gene expression thresholds of the 36 classifier genes. Due to the relatively short follow-up period, few overall deaths and only one cancer-specific death in the validation cohort, there was no statistical power for these two end-points of analysis. We generated Kaplan Meier curves for biochemical progression-free survival of the two predicted classes. The PP group indeed showed a significantly shorter time to biochemical progression than the GP group (p=0.001, Fig. 3). Based on this finding we concluded that the set of 36 genes could be used to predict the samples showing poorer prognosis in an independent *ERG*-positive cohort.

Analysis of ERG-negative sample cohorts

To establish whether the 36-gene signature could also be predictive in *ERG*-negative samples we next performed the class prediction analysis using *ERG*-negative samples present in our own patient cohort as well as in the *ERG*-negative patient samples in the primary prostate cancers of the validation group. In the original patient group of 15 *ERG*-negative samples. Kaplan Meier analysis showed no significant differences in

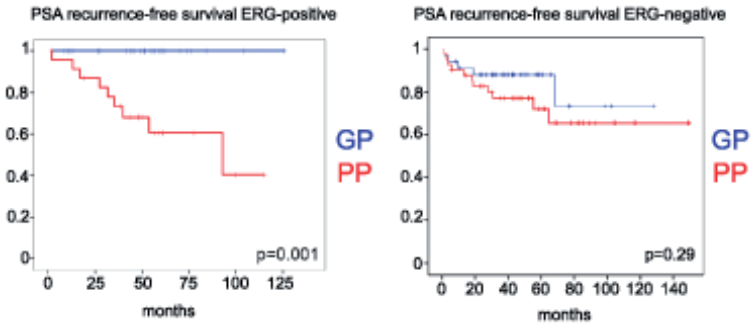


Fig 3. Validation of the classifier using primary prostate cancer samples of a completely independent patient cohort (see ref. 28). Using the 36-gene classifier, Brb array tool was applied to predict the GP and PP groups. Kaplan-Meier curves assessing PSA recurrence free survival were generated for *ERG*-positive (p=0.001; left panel) and *ERG*-negative (p=0.29; right panel) samples.

biochemical progression-free survival between PP and GP (p=0.81) (data not shown). In the second, independent cohort 75 *ERG*-negative samples were used for a class prediction analysis. Also in this set Kaplan Meier analysis with end-point biochemical progression-free survival the groups were not significantly different (p =0.29) (Fig. 3) confirming the results of the original small *ERG*-negative group.

Properties and expression of the classifier genes

A Kegg Pathway analysis was performed using 15357 genes to select for significantly different gene sets between the PP and GP group. The top three pathways detected by this approach were cell-cell communication, TGF- β signaling and focal adhesion. The majority of genes overexpressed in the classifier were involved in cell-cell communication, including cell adhesion and ECM-interaction (Table 1); examples are *COL8A1, VCAN, THBS2* and *ITGA11*. *THBS2* is also involved in the TGF-β signaling pathway, together with *COMP, INHBA and LTBP2*. Another group of genes in the classifier are involved in bone remodeling and regulation, including *COMP, CTSK, MGP* and *CDH11*.

The five genes that were down-regulated in the PP group, *CYP39A1, CRYM, FAM13C, GREB1* and *OPHN1*, seemed not to be functionally related, nor are they included in pathways described above. *GREB1* is known to be expressed in the normal prostate and in prostate cancer, and has been described to be androgen regulated (28).

Next we compared the expression of all classifier genes in PP and GP with expression in normal prostates. Many of the genes that were overexpressed in the classifier were significantly lower expressed in normal prostates as compared to the PP group. Examples are shown in Fig. 4 (*INHBA* and *CDH11*) and in Supplementary Fig. S2 (*COL8A1, COL10A1, COMP, ITGBL1, THBS2, NOX4)*. These data clearly showed that expression in normal prostate tissue (NP) is comparable with expression in the GP group. We also

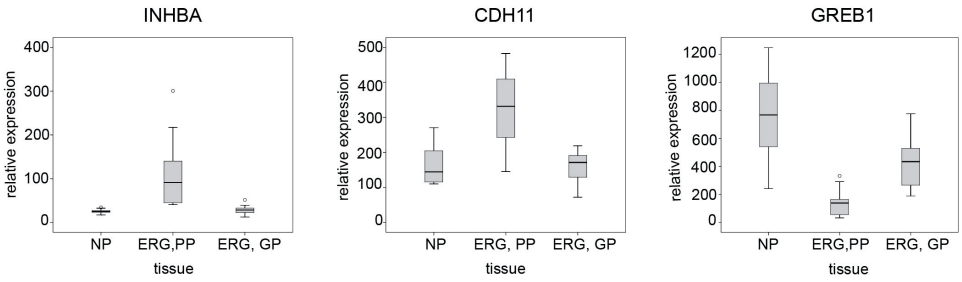


Fig 4. Box plots showing median expression of INHBA, CDH11 and GREB1 mRNA in normal prostate (NP, n=12) and groups PP and GP. For all three genes p < 0.001 (MWU test) when comparing GP with PP or NP with PP. Outliers are depicted by an open circle.

found significantly higher expression in normal prostates of classifier genes with low expression in PP. As shown in Fig. 4 for *GREB1*, the expression in the normal prostate might even be higher than in the GP samples.

In vitro biological functions of INHBA and CDH11

To obtain functional data on classifier genes we investigated expression of the genes in prostate cancer cell lines. From all available prostate cancer cell lines, PC3 cells showed the highest number of classifier genes, including expression of *INHBA, CDH11, VCAN* and *ITGBL1* mRNA (data not shown). Because antibodies directed against CDH11 and INHBA were available, we focused follow-up studies on these two proteins. In separate experiments expression of the two genes was downregulated by specific shRNAs (Fig. 5A,B). Next, properties of the two novel cell lines were compared with the control parental PC3 cells in migration and anchorage-independent growth assays (Fig. 5C). Both *CDH11* and *INHBA* shRNA were able to inhibit the migration and anchorage independent growth of PC3 cells (Fig. 5C). In a complementary experiment, overexpression of *CDH11* or *INHBA* in PNT2C2 immortalized normal prostate epithelial cells stimulated anchorage independent growth (Fig. 5D). These findings provide evidence that classifier genes are not just biomarkers, but can be of functional importance in progressive growth of prostate cancer.

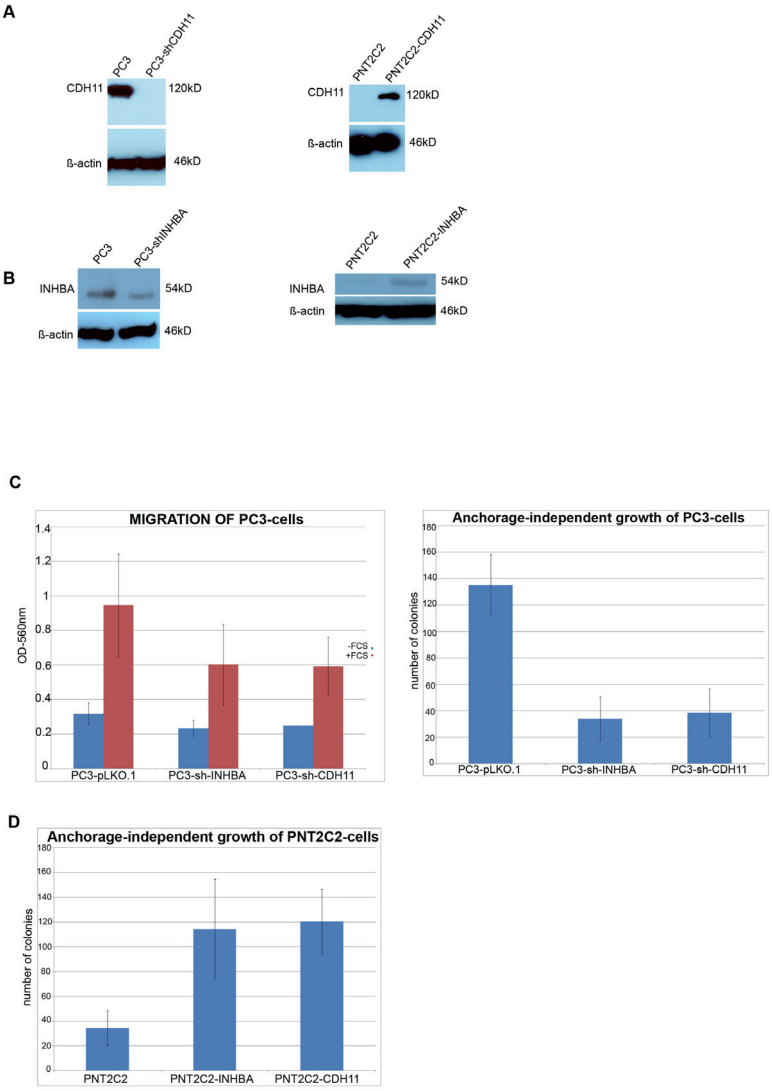


Fig 5. The effect of modulating CDH11 and INHBA expression on *in vitro* biological properties of prostate cells. **A.** Western blot analysis of CDH11 protein expression. PC3 cells were infected with short-hairpin 5700 (CDH11) lentivirus or pLKO-control virus. PNT2C2 cells were infected with pLenti-CMV-h-CDH11 lentivirus or control pWPXLd-GFP virus. Protein detection was performed with the antibodies CDH11 (Invitrogen) and β-actin (Sigma) as loading control. **B.** Western blot analysis of INHBA protein expression. PC3 cells were infected with short-hairpin 5704 (INHBA) or pLKO-control virus. PNT2C2 cells were infected with pLenti-CMV-h-INHBA lentivirus or control pWPXLd-GFP virus. Protein detection was performed with the antibodies INHBA (Novus Biologicals) and β-actin as loading control. **C.** PC3-shCDH11, PC3-INHBA, and control PC3-pLKO cells were assayed for migration and anchorage-independent growth as described in the Methods section were counted. **D.** PNT2C2 cells infected with CDH11- or INHBA-expressing viruses or control PNT2C2-GFP cells were assayed for anchorage-independent growth as described in the Methods section.

Discussion

Gleason score, clinical stage and serum PSA levels are classical parameters in predicting prostate cancer progression. Although valuable, this approach has limited specificity. Based on gene expression profiles or on preselected gene subsets with different functions, several subgroups of prostate cancer that correlated with clinical outcome and that provided insights in biological pathways involved in tumor progression have been postulated (3, 5, 6, 29-31). Differential expression of genes might be due to the properties of the tumor cell itself, but also to reactive stromal or inflammatory cells in the tumor. So far, published data are not very consistent, no doubt partly due to differences in preselected genes, but also probably due to different composition of patient cohorts and tumor heterogeneity. For some genes in our classifier additional expression data are available. *CDH11, INHBA, LTBP2* and *NOX4* were among the genes that were overexpressed in high-grade prostate cancer (3).

Recently, attempts have also been made to identify prostate cancer subgroups by combinations of genetic alterations (27, 31, 32). In the present study we combined genetic analysis with gene expression profiling of primary prostate cancers. We selected tumors that overexpressed *ERG* due to *TMPRSS2-ERG* gene fusion. The value of *ERG* overexpression by itself as a prognostic marker is still a matter of debate. We found in the present study no significant difference in clinical progression between *ERG*-positive and *ERG*-negative tumors (see also refs 18, 19, 33). However, based on gene expression profiles in the ERG-positive group a subgroup of tumors with rapid clinical progression (PP) was identified (Fig. 2). We postulated that within this subgroup other genetic or epigenetic alterations stimulated tumor progression. Importantly, the classifier that we identified also predicted progressive disease in a completely independent patient cohort of *ERG*-positive prostate cancer. The classifier was not predictive in *ERG*-negative samples.

Although the common ERG overexpression by itself might explain increased homogeneity in a tumor subgroup, it might also be explained by the similarity of tumor initiating cells in which gene fusions occurred. The ERG-negative group is genetically more heterogeneous and might select for different combinations of pathways for tumor progression. In our ERG-negative group (A) several tumors were present that overexpressed *ERG*-related *ETV1*. Apparently, these tumors show a gene expression profile that is different from the *ERG*-positive samples.

Brase et al. (26) described by Ingenuity pathway analysis an association between *TMPRSS2*-ERG expression and “factors promoting cardiogenesis”, including TGFβ, WNT

and BMP signaling. Kegg pathway analysis of the complete array-data of all our ERG positive samples indicated significant differences in cell-cell communication and TGFβ-signaling between the PP and GP subgroups. So, not unexpectedly, several of the genes in the classifier are known to be involved in TGFβ-signaling (see Table 1). Recently, *THBS2* and *INHBA*, in combination with *COL11A1* have been correlated with invasion by molecular sub-classification in ovarian and colorectal cancer (34). In prostate cancer, TGFβ-signaling has been proposed to be involved in angiogenesis and EMT transition (35) and TGFβ overexpression is associated with poor outcome and metastasis (36,37). Prostate cancer metastases to the bone. However, the process of bone metastasis is not well understood (38). It is therefore of interest that many genes in the classifier are involved in cell-cell interation and/or bone remodeling, and development.

Detection of ERG overexpression by immunohistochemistry (IHC) is now standard in prostate cancer diagnosis on biopsies (13). The classifier identified in this study can be instrumental in predicting the clinical outcome of ERG-positive cancers, or by multiplex QPCR of RNA isolated from biopsies or by IHC of selected genes when high-quality antibodies become available. The large difference in expression levels of the classifier genes between *ERG*-positive PP tumors, and *ERG*-positive GP tumors and NP are very promising in this regard.

Acknowledgments

We thank Theo van der Kwast and Arno van Leenders for pathological examination of tissue slides, Marcel Smid and Maurice Jansen for bioinformatical assistance and Frank van der Panne for assistance with the figures.

Grant Support

D. Gasi Tandefelt was supported by the Marie Curie grant “Cancure” of the European Union.

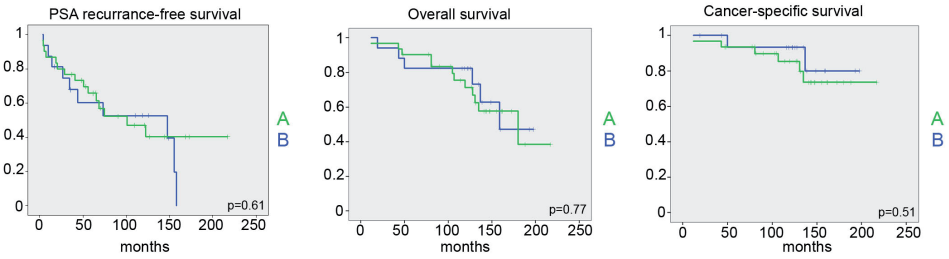
References

1. Siegel R, Naishadham D, Jemal A: Cancer Statistics, 2012. CA Cancer J Clin 2012;62:10-29.
2. Parker C: Active surveillance: towards a new paradigm in the management of early prostate cancer. Lancet Oncology 2004;5:101-6.
3. Lapointe J, Li C, Higgins JP, van de Rijn M, Bair E, Montgomery K et al: Gene expression profiling identifies clinically relevant subtypes of prostate cancer. Proc Natl Acad Sci U S A 2004;101:811-6.
4. Tomlins SA, Mehra R, Rhodes DR, Cao X, Wang L, Dhanasekaran et al: Integrative molecular concept modelling prostate cancer progression. Nat Genet 2006;39: 41-51.
5. Penney KL, Sinnott JA, Fall K, Pawitan Y, Hoshida Y, Kraft P et al: mRNA expression signature of gleason grade predicts lethal prostate cancer. J Clin Oncol 2011;29:2391-6.
6. Markert EK, Mizuno H, Vazquez A, Levine AJ: Molecular classification of prostate cancer using curated expression signatures. Proc Natl Acad Sci U S A 2011;108:21276-81.
7. Grasso CS, Wu Y, Robinson DR, Cao X, Dhanasekaran SM, Khan AP et al: The mutational landscape of lethal castration-resistant prostate cancer. Nature 2012;487:239-43.
8. Barbieri CE, Baca SC, Lawrence MS, Demichelis F, Blattner M, Theurillat JP et al. Exome sequencing identifies recurrent SPOP, FOXA1 and MED12 mutations in prostate cancer. Nat Genet 2012;44:685-9.
9. Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW et al: Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310: 644-8.
10. Hermans KG, van Marion R, van Dekken H, Jenster G, van Weerden WM, Trapman J: TMPRSS2:ERG fusion by translocation or interstitial deletion is highly relevant in androgen-dependent prostate cancer, but is bypassed in late-stage androgen receptor-negative prostate cancer. Cancer Res 2006;66:10658-63.
11. Cerveira N, Ribeiro FR, Peixoto A, Costa V, Henrique R, Jeronimo C et al: TMPRSS2-ERG Gene Fusion Causing ERG Overexpression Precedes Chromosome Copy Number Changes in Prostate Carcinomas and Paired HGPIN Lesions. Neoplasia 2006;8:826-32.
12. Clark J, Attard G, Jhavar S, Flohr P, Reid A, De-Bono J et al: Complex patterns of ETS gene alteration arise during cancer development in the human prostate. Oncogene 2008;27:1993-2003.
13. van Leenders GJ, Boormans JL, Vissers CJ, Hoogland AM, Bressers AA, Furusato B et al: Antibody EPR3864 is specific for ERG genomic fusions in prostate cancer: implications for pathological practice. Modern Pathology 2011;24:1128-38.
14. Saramäki OR, Harjula AE, Martikainen PM, Vessella RL, Tammela TL, Visakorpi T: TMPRSS2:ERG Fusion Identifies a Subgroup of Prostate Cancers with a Favorable Prognosis. Clin Cancer Res 2008;14:3395-400.
15. Demichelis F, Fall K, Perner S, Andren O, Schmidt F, Setlur SR et al: TMPRSS2:ERG gene fusion associated with lethal prostate cancer in a watchful waiting cohort. Oncogene. 2007;26:4596-9.

16. Attard G, Clark J, Ambroisine L, Fisher G, Kovacs G, Fisher P et al : Duplication of the fusion of TMPRSS2 to ERG sequences identifies fatal human prostate cancer. Oncogene 2008;27:253–63.
17. FitzGerald LM, Agalliu I, Johnson K, Miller MA, Kwon EM, Hurtado-Coll A et al: Association of TMPRSS2-ERG gene fusion with clinical characteristics and outcomes: results from a population-based study of prostate cancer. BMC Cancer 2008;8:230.
18. Hoogland AM, Jenster G, van Weerden WM, Trapman J, van der Kwast, Roobol MJ et al: ERG immunohistochemistry is not predictive for PSA recurrence, local recurrence or overall survival after radical prostatectomy for prostate cancer. Mod Pathol 2012;25:471-9.
19. Hermans KG, Boormans JL, Gasi D, van Leenders GJ, Jenster G, Verhagen PC et al: Overexpression of prostate-specific TMPRSS2(exon 0)-ERG fusion transcripts corresponds with favorable prognosis of prostate cancer. Clin Cancer Res 2009;15: 6398-403.
20. Zong Y, Li X, Goldstein AS, Lawson DA, Teitell MA, Witte ON: ETS family transcription factors collaborate with alternative signaling pathways to induce carcinoma from adult murine prostate cells. Proc Natl Acad Sci U S A 2009;106:12465-70.
21. Klezovitch O, Risk M, Coleman I, Lucas JM, Null M, True LD et al: A causal role for ERG in neoplastic transformation of prostate epithelium. Proc Natl Acad Sci U S A 2008;105:2105-10.
22. Carver BS, Tran J, Gopalan A, Chen Z, Shaikh S, Carracedo A et al. Aberrant ERG expression cooperates with loss of PTEN to promote cancer progression in the prostate. Nat Genet 2009;41:619-24.
23. King JC, Xu J, Wonqvipat J, Hieronymus H, Carver BS, Leung DH et al. Cooperativity of TMPRSS2-ERG with PI3-kinase pathway activation in prostate oncogenesis. Nat Genet 2009;41:524-6.
24. Tomlins SA, Laxman B, Varambally S, Cao X, Yu J, Helgeson BE et al: Role of the TMPRSS2-ERG gene fusion in prostate cancer. Neoplasia 2008;10:177–8.
25. Jhavar S, Brewer D, Edwards S, Kote-Jarai Z, Attard G, Clark J et al. Integration of ERG gene mapping and gene-expression profiling identifies distinct categories of human prostate cancer. BJU Int 2008;103:1256-69.
26. Brase JC, Johannes M, Mannsperger H, Fälth M, Metzger J, Kacprzyk LA et al: TMPRSS2-ERG -specific transcriptional modulation is associated with prostate cancer biomarkers and TGF-β signaling. BMC Cancer 2011;11:507.
27. Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS et al: Integrative Genomic Profiling of Human Prostate Cancer. Cancer Cell 2010;18:11–22.
28. Rae JM, Johnson MD, Cordero KE, Scheys JO, Larios JM, Gottardis MM et al: GREB1 is a Novel Androgen-Regulated Gene Required for Prostate Cancer Growth. The Prostate 2006;66:889-94.
29. Heemers HV, Schmidt LJ, Sun Z, Regan KM, Anderson SK, Duncan K et al: Identification of a Clinically Relevant Androgen-Dependant Gene Signature in Prostate Cancer. Cancer Res 2011;71:1978-88.
30. Cuzick J, Swanson GP, Fisher G, Brothman AR, Barney DM, Reid JE et al: Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospective study. Lancet Oncol 2011;12:245-55.

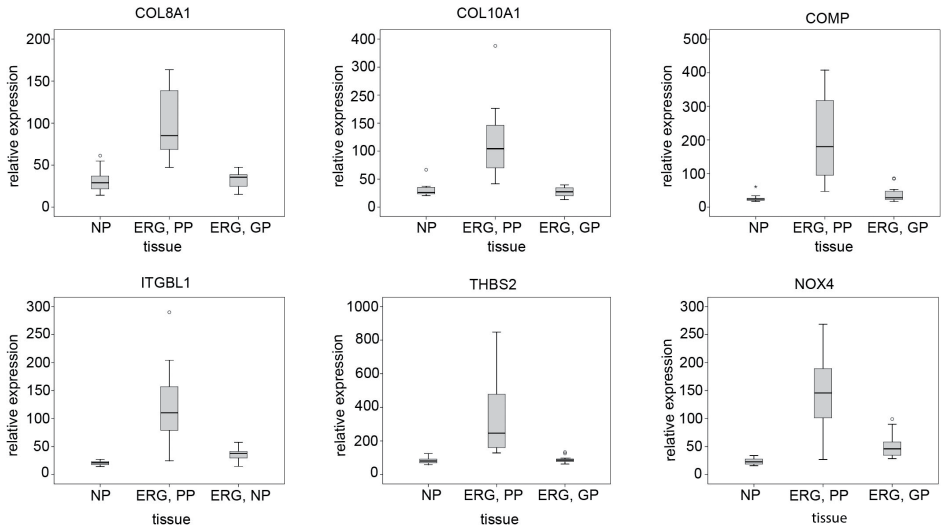
31. Tomlins SA, Laxman B, Dhanasekaran SM, Helgeson BE, Cao X, Morris DS et al: Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature 2007;448:595-9.
32. Lapointe J, Li C, Giacomini CP, Salari K, Huang S, Wang P et al: Genomic profiling reveals alternative genetic pathways of prostate tumorigenesis. Cancer Res 2007;67: 8504-10.
33. Toubaji A, Albadine R, Meeker AK, Isaacs WB, Lotan T, Haffner MC et al: Increased gene copy number of ERG on chromosome 21 but not TMPRSS2–ERG fusion predicts outcome in prostatic adenocarcinomas. Mod Pathol 2011;24:1511-20.
34. Kim H, Watkinson J, Varadan V, Anastassiou D: Multi-cancer computational analysis reveals invasion-associated variant of desmoplastic reaction involving INHBA, THBS2 and COL11A1. BMC Med Genomics 2010;3:51.
35. Liu YN, Abou-Kheir W, Yin JJ, Fang L, Hynes P, Casey O et al: Critical and Reciprocal regulation of KLF4 and SLUG in Transforming Growth Factor β-Initiated Prostate Cancer Epithelial-Mesenchymal Transition. Mol Cell Biol 2012;32:941-53.
36. Wikström P, Stattin P, Franck-Lissbrant I, Damber JE, Bergh A: Transforming Growth Factor **βI** Is Associated With Angiogenesis, Metastasis, and Poor Clinical Outcome in Prostate Cancer. The Prostate 1998;37:19-29.
37. Shariat SF, Shalev M, Menesses-Diaz A, Kim IY, Kattan MW, Wheeler TM et al: Preoperative Plasma Levels of Transforming Growth Factor Beta1 (TGF-β1) Strongly Predict Progression in Patients Undergoing Radical Prostatectomy. J Clin Oncol 2001;19:2856-64.
38. Bubendorf L, Schöpfer A, Wagner U, Sauter G, Moch H, Willi N et al: Metastatic patterns of prostate cancer: an autopsy of 1.589 patients. Hum Pathol 2000;31:578-83.

Supplementary data



Gasi *et al.* Supplementary Figure 1

Supplementary Fig S1. Kaplan-Meier curves assessing the probability of PSA recurrence free survival (p=0.61), overall survival (p=0.77) or cancer specific survival (p=0.51) as a function of time after surgery of patients in group B (blue) and group A (green) separated by unsupervised hierarchical clustering of gene expression profiles (see Fig1). The differences between the curves of the groups A and B were assayed using the log-rank test. Statistical analyses were done using the Statistical Package for Social Sciences, version 17.0 (SPSS), with a significance level of 0.05 (two-tailed probability).



Supplementary Fig S2. Box plots showing median expression of COL8A1, COL10A1, COMP, ITGBL1, THBS2 and NOX4 mRNA in normal prostate (NP, n=12) and primary prostate tumors overexpressing ERG. PP: poor prognosis (n=18), and GP: good prognosis (n=15). For all genes $p < 0.001$ (MWU test) comparing PP versus GP, and PP versus NP. Outliers are depicted by an open circle and extremes are depicted by an asterisk.

Supplementary Table S1 Significantly differentially expressed genes between groups A and B (Fig 1A) at false discovery rate (FDR) of zero

Positive genes (113)						Negative genes (98)					
Gene Name	F Ch	q-val.	Gene Name	F Ch	q-val.	Gene Name	F Ch	q-val.	Gene Name	F Ch	q-val.
ERG	10.90	0	**RTN1**	2.18	0	**EXTL2**	0.52	0	**DDX28**	0.60	0
LAMC2	3.44	0	**RALGAPA2**	1.74	0	**GDEP**	0.28	0	**C7orf51**	0.67	0
HLA-DMB	5.15	0	**MFSD6**	1.66	0	**PAG1**	0.51	0	**DUSP13**	0.65	0
PEX10	3.02	0	**CYorf15B**	1.88	0	**AUTS2**	0.64	0	**FEN1**	0.65	0
CACNA1D	3.97	0	**PCDHB2**	2.52	0	**MESP1**	0.54	0	**EEF1A2**	0.45	0
PDE38	2.15	0	**RHPN2**	1.75	0	**HOXD13**	0.46	0	**ADAMTS8**	0.64	0
AMPD3	2.28	0	**PTGR1**	1.78	0	**ZDHHC4**	0.64	0	**KAZALD1**	0.58	0
ARHGDIB	3.03	0	**PIGR**	3.96	0	**GGT7**	0.63	0	**ODZ1**	0.45	0
GPR110	4.56	0	**CSGALNAC**	2.14	0	**C1orf115**	0.50	0	**PSKH2**	0.55	0
PLA2G2A	6.19	0	**PLA1A**	2.45	0	**MST4**	0.52	0	**C1orf229**	0.63	0
GP2	5.87	0	**SYNJ2BP**	1.88	0	**ITM2C**	0.56	0	**SLC12A5**	0.67	0
SH3RF1	2.38	0	**BCAM**	1.78	0	**LOC100133**	0.50	0	**SNRPF**	0.49	0
PLA2G7	4.18	0	**RRAGD**	1.83	0	**GSTM3**	0.55	0	**MGC45800**	0.70	0
TDRD1	4.17	0	**WSB2**	1.50	0	**TMEM220**	0.49	0	**ELAC1**	0.72	0
ITPR1	2.53	0	**GPR137B**	1.65	0	**SALL1**	0.61	0	**ADH5**	0.59	0
KCNN4	2.67	0	**TMEM178**	3.10	0	**RHOXF2**	0.47	0	**CCDC134**	0.73	0
HES1	2.02	0	**ENC1**	1.78	0	**C6orf192**	0.62	0	**CAPN6**	0.64	0
EFTUD1	1.65	0	**STXBP5**	1.74	0	**MT3**	0.61	0	**DAPK3**	0.71	0
TBC1D228	1.80	0	**PLIN2**	1.69	0	**WRB**	0.56	0	**SLC41A1**	0.63	0
GPR116	3.52	0	**GUCY1B3**	1.72	0	**ZMAT3**	0.62	0	**GJC1**	0.56	0
GHR	3.06	0	**TMEM63A**	1.58	0	**PKP1**	0.39	0	**ARHGEF19**	0.68	0
TMEM26	2.73	0	**STEAP4**	2.32	0	**LAT2**	0.64	0	**SUN2**	0.69	0
KHDRBS3	2.45	0	**PDSS2**	1.78	0	**DLK1**	0.64	0	**SARDH**	0.74	0
NCALD	2.44	0	**ADIPOR1**	1.40	0	**HPGD**	0.40	0	**DIRAS1**	0.65	0
GPR110	3.51	0	**GLB1**	1.53	0	**NKX2**	0.61	0	**RRS1**	0.63	0
ARHGEF38	2.15	0	**PRSS8**	1.68	0	**GLIPR2**	0.54	0	**TMEM18**	0.70	0
NKAIN1	2.42	0	**NR4A2**	2.15	0	**IGFBP2**	0.54	0	**DPY19L2P4**	0.69	0
FZD5	1.95	0	**TMEM45B**	2.47	0	**DPY19L2**	0.60	0	**GPC6**	0.55	0
FRK	2.09	0	**MAPRE3**	1.62	0	**FAM55C**	0.65	0	**IRS2**	0.59	0
VLDUR	2.04	0	**SEL1L3**	1.72	0	**SLC36A1**	0.49	0	**OSTC**	0.67	0
GPR160	1.85	0	**ZNF765**	1.95	0	**FADS6**	0.60	0	**METTL7B**	0.56	0
ACER3	2.13	0	**SIPA1L1**	1.45	0	**KITLG**	0.58	0	**LEPREL1**	0.51	0
ZMYND8	1.77	0	**CADPS**	1.92	0	**ZNF706**	0.62	0	**MAGEF1**	0.67	0
PCDHB10	1.81	0	**CDC42SE2**	1.60	0	**FAM105A**	0.61	0	**IER5**	0.71	0
REXO2	1.98	0	**40795**	1.45	0	**CNTNAP2**	0.45	0	**NPPB**	0.65	0
SPATS2L	1.94	0	**USE1**	1.52	0	**LMO4**	0.64	0	**PRPS2**	0.67	0
ANKRD6	1.56	0	**SIX1**	1.65	0	**CLEC3A**	0.68	0	**FECH**	0.60	0
KIF16B	1.83	0	**CPNE4**	1.83	0	**PIP4K2B**	0.70	0	**ZBTB10**	0.70	0
FAM3D	2.96	0	**KAT2B**	1.49	0	**C17orf48**	0.59	0	**ANAPC4**	0.66	0
SLC44A5	3.06	0	**RBM47**	1.52	0	**FAM35A**	0.64	0	**AQP11**	0.65	0
TSTA3	1.89	0	**APOD**	2.19	0	**RAB27A**	0.51	0	**PNMA1**	0.68	0
CD24	1.74	0	**CLDN1**	2.29	0	**RAMP1**	0.56	0	**PANK1**	0.69	0
MANSC1	1.77	0	**LGMN**	1.52	0	**ADAMTSL1**	0.56	0			
CA5B	1.66	0	**SLC5A1**	2.21	0	**PTPRU**	0.66	0			
LRRN1	2.91	0	**MAP3K5**	1.77	0	**GARNL3**	0.63	0			
SCYL3	1.53	0	**EPCAM**	1.45	0	**C3orf14**	0.59	0			
MAP7	1.61	0	**PACS1**	1.42	0	**CBS**	0.57	0			
AKD1	1.68	0	**TMEM49**	1.35	0	**TFF3**	0.38	0			
SAT1	1.67	0	**PUS10**	1.43	0	**BCL2L2**	0.67	0			
RNF144B	1.91	0	**FKBP1A**	1.60	0	**PAPOL8**	0.67	0			
KCNN2	2.18	0	**ATP11B**	1.42	0	**ITGAE**	0.72	0			
C8orf4	2.81	0	**RPF2**	1.81	0	**OMRT3**	0.68	0			
DSC2	1.68	0	**CRYL1**	1.72	0	**ZNF185**	0.55	0			
GJB1	1.93	0	**C11orf51**	1.48	0	**BMP8B**	0.65	0			
SH3BGRL3	1.51	0	**CGNL1**	1.61	0	**CLEC4G**	0.65	0			
PCDHB11	2.15	0	**KCNH8**	2.01	0	**ANPEP**	0.25	0			
ELMO2	1.57	0									

Supplementary Table 2. Clinical and pathological characteristics of patients in cluster A versus patients in cluster B1 and B2

	Cluster A (n = 17)	**Cluster B1/B2 (n = 31)**	***P***	Test
Mean age (± SD)	63.1 (± 5.4)	61.0 (± 6.1)	0.28	MWU
Mean PSA (± SD)	23.3 (± 38.2)	13.0 (± 14.7)	0.088	MWU
Gleason Score				
< 7	7 (41%)	16 (52%)	0.49	χ^2
≥ 7	10 (59%)	15 (48%)		
pT-stage				
organ confined	8 (47%)	8 (26%)	0.14	χ^2
extra prostatic	9 (53%)	23 (74%)		
Surgical margins				
Negative	9 (53%)	15 (48%)	0.76	χ^2
Positive	8 (47%)	16 (52%)		

Abbreviations: SD, standard deviation; MWU, Mann-Whitney U test; PSA, prostate-specific antigen; pT-stage, pathological T-stage

Supplementary Table 3. Clinical and pathological characteristics of patients in cluster B1 versus patients in cluster B2

	Cluster B1 (n = 15)	**Cluster B2 (n = 16)**	***P***	Test
Mean age (± SD)	59.5 (± 6.2)	62.5 (± 5.8)	0.17	MW
Mean PSA (± SD)	11.5 (± 12.4)	14.5 (± 16.9)	0.49	MW
Gleason Score				
< 7	10 (67%)	6 (38%)	0.10	χ^2
≥ 7	5 (33%)	10 (62%)		
pT-stage				
organ confined	6 (40%)	2 (13%)	0.076	LR
extra prostatic	9 (60%)	14 (87%)		
Surgical margins				
Negative	6 (40%)	9 (56%)	0.37	χ^2
Positive	9 (60%)	7 (44%)		

Abbreviations: SD, standard deviation; PSA, prostate-specific antigen; MWU, Mann-Whitney U test; pT-stage, pathological T-stage; LR, Likelihood Ratio

Supplementary Table 4. Clinical and pathological characteristics of patients in ***ERG*** cluster BI versus patients in ***ERG*** cluster BII

	Cluster BI (n = 15)	**Cluster BII (n = 18)**	***P***	Test
Mean age (± SD)	60.0 (± 5.6)	62.6 (± 5.6)	0.33	MW
Mean PSA (± SD)	8.4 (± 5.3)	22.9 (± 39.7)	0.15	MW
Gleason Score				
< 7	11 (73%)	7 (39%)	0.048	χ^2
≥ 7	4 (27%)	11 (61%)		
pT-stage				
organ confined	7 (47%)	3 (17%)	0.06	LR
extra prostatic	8 (53%)	15 (83%)		
Surgical margins				
Negative	9 (60%)	7 (39%)	0.23	χ^2
Positive	6 (40%)	11 (61%)		

Abbreviations: SD, standard deviation; PSA, prostate-specific antigen; MWU, Mann-Whitney U test; pT-stage, pathological T-stage; LR, Likelihood Ratio

CHAPTER 7

General Discussion

The aim of this thesis was to determine the expression and function of *ERG* and *ETV1* in prostate cancer. *ETV1* and *ERG* are members of two different sub-families of ETS transcription factors (1, 2) and are frequently deregulated in prostate cancer mainly due to gene fusions (3-5). Increasing the knowledge of the molecular mechanisms involved in these common genetic aberrations are not only important in development of biomarkers and novel therapeutic targets but as described in chapter 6, sub-classification by molecular profiles can be crucial in identifying patients with more progressive disease.

ERG in prostate cancer

The discovery of *TMPRSS2-ERG* fusion and its high frequency in clinical prostate cancer was a break-through and much effort has been given in elucidating its relevance for development and progression of prostate cancer.

TMPRSS2-ERG is expressed exclusively in prostate cancer and its precursor lesion PIN. This suggested the discovery to be a very promising biomarker in the clinic. Today *TMPRSS2-ERG* seems to be a robust diagnostic marker (6-8) and it can be detected not only on protein level in prostate biopsies but also as mRNA in the urine.

As a prognostic marker there are contradicting reports and there is still not a clear answer on whether or not *TMPRSS2-ERG* fusion can be used in identifying patients who will have an indolent disease from patients who will progress to more advanced stages of prostate cancer. As described in Chapter 1, in some studies *TMPRSS2-ERG* was correlated with poor clinical outcome (9-12) but there are other studies that do not supported this finding (13-16). One of the studies that do find a correlation with poorer prognosis focused on the different mechanisms of *TMPRSS2-ERG*

fusion. It was described that duplication of the fusion gene results in clearly poorer clinical prognosis but only in the cases where the fusion occurs via deletion of the small region between *TMPRSS2* and *ERG* (9). The explanation to this finding is that the high *ERG* overexpression is the probable driver of the cancer progression. Interestingly a study published two years later which focused on genomic-based classification of prostate cancers showed that the amount of copy number alterations alone define clusters of low- and high-risk disease (17). This might explain the correlation with duplication of *ERG* and poor progression to genomic instability and not necessarily the *ERG* overexpression itself.

In both Chapter 4 and Chapter 6 our results are in agreement with the studies that do not find any difference in clinical progression between patients having *TMPRSS2-ERG* fusion and prostate cancer patients without fusion. Although in Chapter 4 we find that

patients having a *TMPRSS2-ERG* fusion starting at an alternative first exon of *TMPRSS2* (exon 0) had a better prognosis than the rest of *TMPRSS2-ERG* positive patients (18).
Expression arrays have not only been useful in distinguishing prognostic differences within subgroups of clinical prostate cancers but also to identify molecular pathways and candidate target genes of *ERG*. Many differences have been reported between ERG-positive and ERG-negative tumor samples. Some of the discovered genes; *HDAC1*, *CACNA1D, TDRD1, NCALD* seem clearly associated with *ERG* and are reported by several independent groups (19-21). Several pathways have been associated with *ERG*, among others WNT (21, 22) and TGFβ (20).
The discrepancies in the reported results can be caused by several different factors. Mainly there are differences in the populations studied, but also the platforms used, the techniques of determining gene fusion and the possible effect of preliminary treatment of the patients. To obtain more robust and comparable data, standardized methodology should be implemented. The standardization could address sample collection and preferably use of not pre-treated samples, fusion status should be validated by at least two methods and a long clinical follow up would be preferable to correctly draw conclusions from the determined endpoints.
Clearly continued efforts are still needed in elucidating the function of this gene fusion, its targets and possible role in the prognosis and therapy decisions in the clinic.

ETV1 in prostate cancer

In addition to *TMPRSS2-ERG* fusion that occurs in approximately 50% of clinical prostate samples (3, 23, 24), several other fusion events involving ETS transcription factors were reported (25-33).
ETV1 is the second most frequent ETS factor overexpressed in prostate cancer and occurs in 5-10% of prostate tumors (5, 34). Part of that overexpression is due to fusion between *TMPRSS2* and *ETV1* (3). In contrast to *ERG, ETV1* has many more fusion partners (Chapter 1, Table 1). Out of the *ETV1* fusion partners known today only *HNRPA2B1* lacks the common feature of being prostate specific and/or androgen regulated (4). The overexpression can be as truncated ETV1 or as shown here a fusion protein encoded by the first 8 exons of HNRPA2B1 and ETV1 exon 6-13 (Chapter 2, Appendix 1). *HNRPA2B1* is a housekeeping gene and is identified independently by different groups (4,35, see also Chapter 2 Appendix 1) so there seems some preference for this housekeeping gene in particular. One theory could be that the fusion occurs because of spatial proximity of *ETV1* and *HNRPA2B1* genomic regions under specific conditions in the cell. An

alternative hypothesis could be that the selection is due to the specific properties of the promoter region of *HNRPA2B1*. *HNRPA2B1* is transcribed via a convergent dual promoter with ubiquitously open chromatin structure spanning the promoter region and a large part of the gene (36, 37). *DDX5* is another housekeeping gene with similar structure in its promoter region which is involved in fusions with *ETV4*, member of the same ETS subfamily as *ETV1* (38). It could be speculated that alternative mechanisms to the common androgen-regulated and prostate-specific fusion partners include a constitutive open chromatine structure.

High expression of *ETV1* was observed in the prostate cancer cell lines LNCaP and MDAPCa2B without identification of a fusion event (4). This feature was also observed for a subgroup of clinical prostate cancer samples (5, 34). It was described for both MDA and LNCaP that this overexpression most probably was due to the translocation of *ETV1* to the same region on chromosome 14. In Chapter 2 we investigated primary tumors that had overexpression of full-length ETV1 by fluorescence *in situ* hybridization (FISH) to determine whether translocation of the gene had occurred also in clinical prostate cancer. Indeed we could in all, except one, sample determine translocation of *ETV1*. In one sample we could locate the translocation by FISH and sequencing to the close proximity of the same *ETV1* insertion point as in LNCaP cells. The insertion point of *ETV1* was in the only intron of a prostate specific and androgen regulated spliced EST, an EST earlier identified as fusion partner of *ETV1* in another clinical prostate cancer sample (5). To determine the translocation in the tumor of a second patient we used material from xenograft PC135, generated from the same patients harboring the translocation. We combined multi-color FISH with individual chromosome paints on chromosome spreads and could roughly map the novel translocation to chromosome 4
(see Chapter 2). Genome-wide deep sequencing was performed (data not shown) and the data was confirmed by PCR and sequencing to the exact breakpoint on chromosome 4 close to *PDLIM5,* a gene highly expressed in prostate (Appendix 2). The data suggests that in case of translocation of the complete gene the insertion points of *ETV1* are in proximity of genes that are either androgen regulated or prostate specific or both, which is in concordance with the properties of the majority of ETS fusion partners. We have so far only described two regions of the translocated full-length *ETV1*.

Identifying the position of *ETV1* insertion should be applied to all other *ETV1* translocated samples. Investigating the characteristics of these regions is important in clarifying if there are common mechanisms of *ETV1* overexpression. Circulized chromosome confirmation capture (4C) seems a promising method not only for identifying the insertion points of

the translocations but can also be used for capturing the interactions that are occurring in the promoter region of *ETV1.*

ETV1 and dETV1

As described previously in prostate cancer ETV1 can be overexpressed in a truncated form when involved in fusion events (dETV1) or as the full-length gene.
To address the question of which genes are regulated by ETV1 and dETV1 we performed expression arrays of normal prostate cell line PNT2C2 infected with viruses expressing constructs of these two different forms (Chapter3). ETV1 and dETV1 have very similar biological functions but there are also differences. Anchorage-independent growth is only induced in normal prostate cells infected with *ETV1* expression constructs and not *dETV1* (5, 34). From our microarray data we could conclude that the majority of overexpressed genes are shared by both *ETV1* and not by *dETV1* (Chapter 3). *BCL2A1* is the clearest example. *BCL2A1* is member of the BCL2 protein family that are important regulators of apoptosis. *BCL2A1* is overexpressed in several cancer types (39) and could be involved in prostate tumor progression. There were also examples of genes more highly expressed in *ETV1* compared to *dETV1*, *PAPPA* as most obvious example. *PAPPA* is a metalloproteinase that cleaves IGFBP's and is known to be involved in wound healing and bone remodeling. The relevance of this gene in cancer or prostate cancer has not been determined. Because of the low frequency of *ETV1* overexpressing samples, joint and thus larger sample cohorts should be used to best determine the relevance of these novel targets in the clinic.
Since we saw clear differences between ETV1 and dETV1 in soft-agar assays but not in invasion, migration nor proliferation assays an interesting aspect was to also study how different growth conditions affect the expression of the genes selected in our studies, which were based on 2D experiments. We performed expression arrays on cells grown in 2D or grown *in vivo*. In Chapter 3 we show the results of infected PNT2C2 cells grown *in vivo* in mice. *PAPPA* is similarly induced in PNT2C2 cells grow in 2D and *in vivo* but there are several observed differences. *CRISP3* and *PEAP* are genes that appear specifically in the *in vivo* situation and not in 2D. To obtain robust data it is of high importance to repeat the experiment as to be able to determine proper controls and validate the selected genes.

ETV1 versus ERG

Expression of *ERG* and *ETV1* occur mutually exclusive in clinical prostate tumors suggesting a similar function. ETS transcription factors, ERG and ETV1 share high homology in their DNA-binding site, the so called ETS-domain. ERG and ETV1 are grouped in two different subfamilies (1, 2). The homology decreases substantially in the flanking sequences and have structural differences which compose structured regions like protein-interaction domains. Fusion events involving *ERG* are detected already in precursor leasons (PIN) which points to their role in the onset of prostate cancer (8, 40-42), but are still expressed in later stages of the disease and can thus as well play a role in progression of prostate cancer to more advanced stages. There are no reports showing expression of *ETV1* in PIN. This might suggest that *ERG* and *ETV1* have different roles in prostate cancer development where *ERG* is more important in the onset and *ETV1* in later stages of tumor progression. Neither the de-regulation of *ETV1* or *ERG* is sufficient for transformation to cancer in *in vivo* studies performed in mice (43). Several possible pathways and additional alterations have been proposed to cooperate with *ERG* overexpression in tumor development (44, 45) but the biological and clinical role of *ERG* in prostate is still far from elucidated.

Gene expression profiling has been used to identify genes and pathways associated with *ERG* in clinical samples but probably due to the much lower frequency of *ETV1* rearrangement it has been difficult to find similar associations within the *ETV1* subgroup of patients. In Chapter 6 we show by unsupervised hierarchical clustering that *ERG* samples group separately from *ETV1* and other primary tumors. This finding suggests that the two family members mainly affect different target genes and pathways, alternatively they are active in different cell types. Recently the question of the similarities and differences between *ETV1* and *ERG* expression and target genes was addressed (46). It was shown on one hand shared target genes between *ERG* and *ETV1,* on the other hand they also show that there are specific targets for respective ETS gene. In Chapter 3 we find only one gene that is both associated with ERG and ETV1.

ERG and *FLI1* are closely related ETS transcription factors as are the PEA3-group members *ETV1*, *ETV4* and *ETV5*. So far ETS-related tumors have been considered as a one specific group of prostate cancers. Most studies have been preformed on *ERG*-positive samples only. The PEA3 genes compose a separate sub-group in prostate cancer with considerable differences from *ERG* samples. How and if this difference is reflected in the clinic is still to be investigated.

Heterogeneity of prostate cancer

When different foci of the prostate were tested for their fusion status considerable heterogeneity was discovered (see Chapter 1). Some cancer foci could be positive for a fusion while others could be negative, but also different fusion events could be present in the different foci of the same primary prostate tumor (47,48). If investigating several metastases in a prostate cancer patient, however, the fusion status was identical (49-51). This strongly indicates that several different tumors can arise in the different foci independently of each other but that only one of the foci is responsible for seeding the metastasis. It would be interesting to investigate the differences between the tumor cells in the foci that do not metastasize from the focus that gives rise to the metastasis in the same individual. Are there common genetic and epigenetic pathways that trigger the metastatic potential in these foci or is it the type of cell where the fusion occurs that is the driving mechanism, and can sub-classification of tumors help in decreasing the heterogeneity of samples and thus facilitate the discovery of prognostic markers?

In Chapter 6 unsupervised hierarchical clustering was used to find molecular subgroups in primary prostate cancer samples. Although the clustering defines *ERG* as a distinct subgroup of primary tumors, as mentioned earlier, the clinical parameters did not differ between the *ERG* positive group and rest of the primary tumors. Interestingly we observed additional sub-grouping within the *ERG*-positive samples. One subgroup had significantly poorer outcome in overall-, cancer specific- and PSA recurrence-free survival. The differentially expressed genes between these two groups were used to determine a classifier that was predictive in *ERG*-positive but not *ERG*-negative samples. The findings were validated in a completely independent sample cohort and there are additional data available on some of the genes present in our classifier.

We postulate that *ERG* status selects for a more homogenous group and therefore facilitates the identification of genes responsible for tumor progression. These findings are directly important for discriminating patients that are in need of treatment from patients who can benefit from active surveillance. Obviously, a panel of discovered gene markers can be selected and tested by immunohistochemistry in tumor biopsies, with clearly identified ERG-status. Alternatively, isolated RNA from the biopsy could be used in multiplex QPCR analysis to select for the tumors with predicted progression.

With the development of improved high-throughput techniques and lower costs on large scale sequencing it is today possible to collect large amounts of data on novel genetic deviations occurring in the human genome of prostate cancer patients and study their associations. The challenge now is to find more standardized methods of analysis of

the cancer genome and of the interpretation for these extremely large amounts of data, and to distinguish the causal drivers from the irrelevant passengers.

References

1. Hollenhorst PC, McIntosh LP and Graves BJ. Genomic and Biochemical Insights into the Specificity of ETS Transcription Factors. Annu Rev Biochem 2011;80:437-71.
2. Seth A and Watson DK. ETS transcription factors and their emerging roles in human cancer. Eur J Cancer 2005;41:2462-78.
3. Tomlins SA, Rhodes DR, Perner S et al. Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. Science 2005;310:644-8.
4. Tomlins SA, Laxman B, Dhanasekaran SM et al. Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature 2007;448:595-9.
5. Hermans KG, van der Korput HA, van Marion R et al. Truncated ETV1, fused to novel tissue-specific genes, and full-length ETV1 in prostate cancer. Cancer Res. 2008;68:7541-9.
6. Hessels D, Smit FP, Verhaegh GW, Witjes JA, Cornel EB, Schalken JA. Detection of TMPRSS2-ERG fusion transcripts and prostate cancer antigen 3 in urinary sediments may improve diagnosis of prostate cancer. Clin Cancer Res 2007;13:5103-8.
7. Tomlins SA, Palanisamy N, Siddiqui J, Chinnaiyan AM, Kunju LP. Antibody-based detection of ERG rearrangements in prostate core biopsies, including diagnostically challenging cases: ERG staining in prostate core biopsies. Arch Pathol Med 2012;136:935-46.
8. van Leenders GJ, Boormans JL, Vissers CJ, Hoogland AM, Bressers AA, Furusato B, Trapman J. Antibody EPR3864 is specific for ERG genomic fusions in prostate cancer: implications for pathological practice. Mod Pathol 2011;24:1128-38.
9. Attard G, Clark J, Ambroisine L et al. Duplication of the fusion of TMPRSS2 to ERG sequences identifies fatal human prostate cancer. Oncogene 2008;27: 253-63.
10. Nam RK, Sugar L, Yang W et al. Expression of the TMPRSS2-ERG fusion gene predicts cancer recurrence after surgery for localized prostate cancer. Br J Cancer 2007;97:1690-5.
11. Perner S, Demichelis F, Beroukhim R et al. TMPRSS2-ERG fusion-associated deletions provide insight into the heterogeneity of prostate cancer. Cancer Res 2006;66:8337-41.
12. Wang J, Cai Y., Ren C and Ittman M. Expression of variant TMPRSS2-ERG mRNA is associated with aggressive prostate cancer. Cancer Res 2006;66:8347-51.
13. FitzGerald LM, Agalliu I, Johnson K et al. Association of TMPRSS2-ERG gene fusion with clinical characteristics and outcomes: Results from a population-based study of prostate cancer. BMC Cancer 2008;8,230.
14. Gopalan A, Leversha MA, Satagopan JM et al. TMPRSS2-ERG gene fusion is not associated with outcome in patients treated by prostatectomy. Cancer res 2009;69:1400-6.
15. Lapointe J, Kim YH, Miller MA et al. A variant TMPRSS2 isoform and ERG fusion product in prostate cancer with implications for molecular diagnosis. Mod Pathol 2007;20:467-73.
16. Saramaki OR, Harjula AE, Martikainen PM et al. TMPRSS2-ERG fusion identified a subgroup of prostate cancers with a favourable prognosis. Clin Cancer Res 2008;14:3395-4000.
17. Taylor BS, Schultz N, Hieronymus H et al. Integrative Genomic Profiling of Human Prostate Cancer. Cancer Cell 2010;18:11-22.

18. Hermans KG, Boormans JL, Gasi D et al. Overexpression of prostate-specific TMPRSS2(exon 0)-ERG fusion transcripts corresponds with favorable prognosis of prostate cancer. Clin Can Res 2009;15:6398-403.
19. Iljin K, Wolf M, Edgren H et al. TMPRSS2 fusions with oncogenic ETS factors in prostate cancer involve unbalanced genomic rearrangements and are associated with HDAC1 and epigenetic reprogramming. Cancer Res 2006;66:10242-6.
20. Brase JC, Johannes M, Mannsperger H et al. TMPRSS2-ERG –specific transcriptional modulation is associated with prostate cancer biomarkers and TGF-β signalling. BMC Cancer 2011;11,507.
21. Jhavar S, Reid A, Clark J et al. Association of TMPRSS2-ERG gene fusion in human prostate cancer by expression profiling using GeneChip Human Exon 1.0 ST arrays. J Mol Diagn 2008;10:50-7.
22. Mochman LH, Bock J, Ortiz-Tanches J et al. Genome-wide screen reveals WNT11, a non-canonical WNT gene, as a direct target of ETS transcription factor ERG. Oncogene 2011;30:2044-56.
23. Soller MJ, Isaksson M, Elfving P et al. Confirmation of the high frequency of the TMPRSS2/ERG fusion gene in prostate cancer. Genes Chromosomes Cancer 2006;45:717-9.
24. Hermans KG, van Marion R, van Dekken H et al. TMPRSS2:ERG fusion by translocation or intestitial deletion is highly relevant in androgen-dependent prostate cancer, but is bypassed in late-stage androgen receptor-negative prostate cancer. Cancer Res 2006;66:10658-63.
25. Hermans KG, van der Korput HA, van Marion R et al. Truncated ETV1, fused to novel tissue-specific genes, and full-length ETV1 in prostate cancer. Cancer Res 2008;68:7541-9.
26. Tomlins SA, Laxman B, Dhanasekaran SM et al. Distinct classes of chromosomal rearrangements create oncogenic ETS gene fusions in prostate cancer. Nature 2007;448:595-9.
27. Tomlins SA, Mehra R, Rhodes DR et al. TMPRSS2:ETV4 gene fusions define a third molecular subtype of prostate cancer. Cancer Res 2006;66:3396-400.
28. Helgeson BE, Tomlins SA, Shah N et al. Characterization of TMPRSS2:ETV5 and SLC45A3:ETV5 gene fusions in prostate cancer. Cancer Res 2008;68:73-80.
29. Attard G, Clark J, Ambroisine L et al. Heterogeneity and clinical significance of ETV1 translocations in human prostate cancer. Br J Cancer 2008;99:314-20.
30. Han B, Mehra R, Dhanasekaran SM et al. A fluorescence in situ hybridization screen for E26 transformation-specific aberrations: identification of DDX5-ETV4 fusion protein in prostate cancer. Cancer Res 2008;68:7629-37.
31. Hermans KG, Bressers AA, van der Korput HA, Dits NF, Jenster G and Trapman J. Two unique novel prostate-specific and androgen-regulated fusion partners of ETV4 in prostate cancer. Cancer Res 2008;68:3094-98.
32. Clark JP and Cooper CS. ETS gene fusions in prostate cancer. Nat Rev Urol 2009;6:429-39.
33. Rubin MA, Maher CA & Chinnaiyan AM. Common gene rearrangements in prostate cancer. J Clin Oncol 2011;27:3659-68.
34. Gasi D, van der Korput HA, Douben HC et al. Overexpression of full-length ETV1 transcripts in clinical prostate cancer due to gene translocation. PLoS One 2001;6:e16332.
35. Miyagi Y, Sasaki T, Fujinami K et al. ETS family-associated gene fusions in Japanese prostate cancer: analysis of 194 radical prostatectomy samples. Mod Pathol 2011;23:1492-8.

36. Antoniou M, Harland L, Mustoe T et al: Transgenes encompassing dual-promotor CpG islands from the human TBP and HNRPA2B1 loci are resistant to heterochromatin-mediated silencing. Genomics 2003;82:269-79.
37. Williams S, Mustoe T, Mulcahy T et al: CpG-island fragments from the HNRPA2B1/CBX3 genomic locus reduce silencing and enhance transgene expression from the hCMV promoter/enhancer in mammalian cells. BMC Biotechnology 2005;5:17
38. Han B, Mehra R, Dhanasekaran SM et al. A fluorescence in situ hybridization screen for E26 transformation-specific aberrations: identification of DDX5-ETV4 fusion protein in prostate cancer. Cancer Res 2008;68:7629-37.
39. Vogler M. BCL2A1: the underdog in the BCL2 family. Cell Death Differ 2012;19:67-74.
40. Cerveira N, Ribeiro FR, Peixoto A et al. TMPRSS2-ERG gene fusion causing ERG overexpression precedes chromosome copy number changes in prostate carcinomas and paired HGPIN lesions. Neoplasia 2006;8:826-32.
41. Mosquera JM, Perner S, Genega EM et al. Characterization of TMPRSS2- ERG fusion in high-grade prostatic intraepithelial neoplasia and potential clinical implications. Clin Cancer Res 2008;14:3380-5.
42. Park K, Tomlins SA, Mudaliar KM et al. Antibody-based detection of ERG rearrangement-positive prostate cancer. Neoplasia 2010;12:590-8.
43. Zong Y, Xin L, Goldstein AS et al. ETS family transcription factors collaborate with alternative signaling pathways to induce carcinoma from adult murine prostate cells. Proc Natl Acad Sci USA 2009;106:12465-70.
44. Carver BS, Tran J, Gopalan A et al. Aberrant ERG expression cooperates with loss of PTEN to promote cancer progression in the prostate. Nat Genet 2009;41:619-624.
45. King JC, Xu J, Wongvipat J et al. Cooperativity of TMPRSS2-ERG with PI3-kinase pathway activation in prostate oncogenesis. Nat Genet 2009;41:524-6.
46. Paulo P, Ribeiro FR, Santo J et al. Molecular Subtyping of Primary Prostate Cancer Reveals Specific and Shared Target Genes of Different ETS Rearrangements. Neoplasia 2012;14:600-11.
47. Barry M, Perner S, Demichelis F and Rubin MA. TMPRSS2-ERG fusion heterogeneity in multifocal prostate cancer: clinical and biologic implications. Urology 2007;70:630-3.
48. Mehra R, Han B, Tomlins SA et al. Heterogeneity of TMPRSS2 gene rearrangements in multifocal prostate adenocarcinoma: molecular evidence for an independent group of diseases. Cancer Res 2007;67:7991-5.
49. Mehra R, Tomlins SA, Yu J et al. Characterization of TMPRSS2-ETS Gene Aberrations in Androgen-Independent Metastatic Prostate Cancer. Cancer Res 2008;68:3584-90.
50. Liu W, Laitinen S, Khan S et al. Copy number analysis indicates monoclonal origin of Lethal Metastatic prostate cancer. Nat Med 2009;15:559-65.
51. Guo CC, Wang Y, Xiao L, Troncoso P and Czerniak B. The relationship of TMPRSS2-ERG gene fusion between primary and metastatic prostate cancers. Human Pathol 2012;43:644-9.

Summary

In this thesis the expression and function of two ETS transcription factors, ETV1 and ERG, in prostate cancer were studied. Prostate cancer is the most frequent cancer in men in countries with a western lifestyle. Overexpression of ETS transcription factors is a very common molecular alteration. In approx. 50% of prostate cancers the ETS gene *ERG* is overexpressed, due to fusion of *ERG* to another gene. In almost all cases the *ERG* fusion partner is the androgen-regulated, prostate-specific gene *TMPRSS2***.** In 5-10% of prostate cancers the ETS gene *ETV1* is overexpressed, or due to gene fusion or due to translocation of the complete *ETV1* gene. In many cases the *ETV1* fusion partners are, like *TMPRSS2*, androgen-regulated, prostate-specific genes. In Chapter 1, an overview of the current knowledge of *TMPRSS2-ERG* and other ETS gene fusions in prostate cancer is given.

To investigate the mechanism of *ETV1* overexpression in clinical primary prostate cancer samples without *ETV1* gene fusion, FISH experiments were performed. In Chapter 2 it was described that overexpression of full-length ETV1 was, in four out of five clinical prostate cancer samples due to translocation of the complete gene. In one case, *ETV1* was translocated to the intron of an earlier described gene encoding an ncRNA, EST14. In a second case, we identified a novel translocation site of the complete *ETV1* gene to chromosome 4. Four different isoforms of ETV1 were identified and it was shown that, in contrast to the N-truncated ETV1 protein (dETV1) that is synthesized due to gene fusion, even the shortest isoform synthesized by the complete gene, missing a large part of the TAD domain, is as active as the longer isoforms in *in vitro* anchorage independent growth. In Appendix 1, following Chapter 2, a novel fusion-event giving rise to HNRPA2B1-ETV1 fusion protein was identified. As indicated by whole genome sequencing of genomic DNA from sample G89, the exact integration site on chromosome 4, to where ETV1 was translocated, was pinpointed in Appendix 2.

Next, the effect of ETV1 overexpression was studied in an immortalized normal prostate cell line (see Chapter 3 for preliminary results). PNT2C2 cells overexpressing ETV1 or dETV1 were grown *in vitro* in 2-D culture and *in vivo* as transplanted tumors, and ETV1 regulated genes were assessed from gene expression profiles compared to control PNT2C2-GFP cells. In PNT2C2 cells grown *in vitro*, both ETV1 and dETV1 induced the expression of approx. 200 genes 2-fold or more. Most of these genes were identical. *BCL2A1* showed the highest induction by both ETV1 and dETV1. *PAPPA* was preferentially induced in PNT2C2-ETV1 cells. However, in transient promoter assays both the *BCL2A2* and the *PAPPA* promoter were stronger activated by ETV1 than by dETV1. Further, gene

expression experiments showed almost two times more genes that were 2-fold or more induced in PNT2C2-ETV1 than in PNT2C2-dETV1 tumors, of which almost half were identical. *MMP1* showed high induction by both ETV1 and dETV1. Many genes were similarly regulated *in vitro* and *in vivo*, but some genes, including *MMP1* and *CRISP3* were much stronger induced *in vivo.* In clinical primary prostate cancer samples ETV1/dETV1-associated genes were not identical to ETV1/dETV1-induced genes in PNT2C2 cells. Moreover, ETV1-associated genes in clinical prostate cancer showed almost no overlap with ERG-associated genes.

Despite of its high frequency, the role of *TMPRSS2-ERG* in clinical prostate cancer is not well understood. In Chapter 4 we describe that *TMPRSS2-ERG* does not predict prostate cancer progression. On the other hand, we showed that *TMPRSS2-ERG* positive samples expressing an upstream, alternative first exon of *TMPRSS2*, denoted exon 0, was an independent predictor of longer biochemical (PSA) progression-free survival.

In Chapter 5 the use of QPCR, 5' RLM-RACE and sequencing as methods for identifying novel gene fusion events and methods of studying the androgen-regulation of the novel ETS gene fusion partners were described.

By expression-array profiles and unsupervised clustering of primary prostate cancer tumors *TMPRSS2-ERG* positive prostate cancer samples were separated from *ETV1*-positive and ETS-negative tumors (Chapter 6). This separation did not have any differential outcome in clinical endpoints as PSA recurrence-free-, overall- and cancer specific survival supporting the earlier findings described in Chapter 4. Interestingly, a subgrouping was seen within the ERG-positive sample group. Two subgroups showed differences in overall-, cancer-specific- and PSA recurrence-free survival giving a basis to distinguish a 36-gene signature that would discriminate rapid progression from slow progression. This gene signature was validated in an independent patient cohort. The gene signature was not predictive in *ERG*-negative tumors.

In Chapter 7, the general discussion, main findings and limitation of this thesis are described together with suggestions on future perspectives for continuation of this work.

Samenvatting

In dit proefschrift worden de expressie en de functie van twee ETS transcriptiefactoren (ETV1 en ERG) in prostaatkanker bestudeerd. Prostaatkanker is de meest voorkomende vorm van kanker bij mannen in landen met een westerse levensstijl. Verhoogde expressie van ETS transcriptiefactoren wordt heel frequent waargenomen in deze tumoren. In ongeveer 50% van prostaattumoren komt het *ERG* gen verhoogd tot expressie. Deze verhoogde expressie wordt veroorzaakt door fusie van *ERG* met een tweede gen. In vrijwel alle gevallen is de *ERG* fusiepartner het *TMPRRS2* gen. De expressie van dit laatste gen wordt geregeld door androgenen en is tevens specifiek voor de prostaat. In 5-10% van prostaattumoren komt het ETS gen *ETV1* verhoogd tot expressie. Deze verhoogde expressie kan het resultaat zijn van genfusie of van chromosomale translocatie van het complete *ETV1* gen. In veel gevallen wordt de expressie van de *ETV1* fusiepartners, net zoals *ERG*, gereguleerd door androgenen en is specifiek voor de prostaat. In Hoofdstuk 1 wordt een overzicht gegeven van de huidige kennis van *TMPRSS2-ERG* en andere fusies van ETS genen in prostaatkanker.

Ten einde het mechanisme van verhoogde expressie van *ETV1* in primaire prostaattumoren zonder *ETV1* genfusie te bestuderen, werden FISH experimenten uitgevoerd. In Hoofdstuk 2 wordt beschreven dat in 4 van 5 onderzochte tumoren *ETV1* overexpressie veroorzaakt wordt door translocatie van het gehele gen. In 1 tumor werd gevonden dat het *ETV1* gen transloceerd was in het intron van het eerder beschreven gen EST14, dat codeert voor een ncRNA. In een tweede tumor werd een compleet andere positie van translocatie geconstateerd, in dit geval naar chromosoom 4. Vier verschillende iso-vormen van het complete ETV1 eiwit werden gevonden, en aangetoond werd dat, in tegenstelling tot het verkorte ETV1 eiwit (dETV1) dat gesynthetiseerd wordt als gevolg van genfusie, zelfs de kortste ETV1 iso-vorm, die een groot deel van het TAD domein mist, actief is in *in vitro* “anchorage-independent” celgroei experimenten. In Appendix 1 van Hoofdstuk 2 wordt een nieuw fusie-eiwit, HNRPA2B1-ETV1 beschreven. In Appendix 2 wordt met behulp van “whole-genome sequencing” de exacte positie bepaald van het breukpunt op chromosoom 4, waarnaar *ETV1* getransloceerd wordt.

Vervolgens werd het effect van verhoogde expressie van ETV1 in een geïmmortaliseerde normale prostaat cellijn onderzocht (zie Hoofdstuk 3 voor de voorlopige resultaten). PNT2C2 cellen die ETV1 of dETV1 verhoogd tot expressie brengen, werden *in vitro* gegroeid in 2D-cultures en *in vivo* als getransplanteerde tumoren en daarna werden de ETV1-afhankelijke profielen van genexpressie bepaald, vergeleken met die van PNT2C2-GFP controle cellen. Zowel ETV1 als dETV1 induceerde de expressie van zo’n dan 200

genen meer dan 2 keer in *in vitro* gegroeide PNT2C2 cellen. Veel van deze geïnduceerde genen waren identiek. Het gen dat het sterkst geïnduceerd werd door ETV1 en dETV1 was *BCL2A1*. *PAPPA* werd preferentieel geïnduceerd door ETV1. In experimenten waarin de activiteit van de promotoren van het *BCL2A1* gen en van *PAPPA* bepaald werden, bleek echter in alle gevallen ETV1 beter te werken dan dETV1. Verdere gen-expressie experimenten doen vermoeden dat in PNT2C2-ETV1 tumoren bijna twee keer meer genen geïnduceerd worden dan in PNT2C2-dETV1 tumoren. *MMP1* toonde sterke inductie door zowel ETV1 als dETV1. Veel genen die in 2D-cultures gereguleerd werden door ETV1/dETV1 werden ook in tumoren gereguleerd, maar er zijn ook voorbeelden van genen, zoals *MPP1* en *CHRISP3*, die *in vivo* sterker gereguleerd schijnen dan *in vitro*. Genen geassocieerd met overexpressie van ETV1/dETV1 in klinische prostaattumoren vertonen vrijwel geen overlap met genen geassocieerd met ETV1/dETV1 overexpressie in PNT2C2 cellen. Ook vertonen ETV1-geassocieerde genen in klinische tumoren vrijwel geen overlap met ERG-geassocieerde genen.

In Hoofdstuk 4 wordt beschreven dat de aanwezigheid van *TMPRSS2-ERG* in prostaatkanker geen voorspellende waarde heeft voor de snelheid van progressie van de tumor. Aanvullend werd aangetoond dat, in de groep van *TMPRSS2-ERG* positieve tumoren, start van transcriptie vanaf een alternatief *TMPSS2* eerste exon (hier exon 0 genoemd) een onafhankelijke voorspeller is van langere biochemische (PSA) progressie-vrije overleving.

Analyse van mRNA van prostaattumoren met genexpressie arrays en clustering van de resultaten geven een scheiding van *TMPRSS2-ERG* positieve tumoren van *ETV1*-positieve tumoren en ETS-negatieve tumoren (Hoofdstuk 6). Deze scheiding toonde geen verschil in klinisch relevante eindpunten, zoals PSA progressie-vrije overleving, overall overleving en kanker-specifieke overleving, zoals ook beschreven in Hoofdstuk 4. Een interessante waarneming was dat, in de groep van *TMPRSS2-ERG* positieve tumoren, twee subgroepen met verschillen in klinische verloop konden worden onderscheiden, bepaald door verschillen in overall genexpressie profielen. Gebaseerd op deze gegevens kon een 36-gen signatuur worden ontwikkeld, die discriminerend was tussen tumoren met snelle en langzame progressie. Deze signatuur kon worden gevalideerd in tumoren van een onafhankelijke groep van patiënten. De signatuur bleek echter niet voorspellend in een groep van ERG-negatieve tumoren.

In Hoofdstuk 7 (algemene discussie) worden de belangrijkste bevindingen van dit proefschrift bediscussieerd en suggesties voor verder onderzoek besproken.

Curriculum Vitae

Delila Gasi Tandefelt was born the 19th of August 1982 in Belgrade, former Yugoslavia. In 2001 she finished her studies in natural sciences at Fyrisskolan Highschool, Uppsala Sweden, and started in September of the same year her studies at Uppsala University. She obtained her Master of Science diploma in February 2006 with the final project "Coxackie virus infection influences the effects of brominated flame retardants on cytochrome P450 and thyroid hormones in mice" conducted at the Toxicology Department of The National Food Administration of Sweden. Same month she started a six months project at Georgia Institute of Technology in Atlanta, GA, USA in a multidisciplinary team lead by Prof Cyrus Aiden. She was studying cellulose nanocrystals as drug carriers in tumor targeting. Her role was to add the molecular biology background to the group. In November of 2006 she started an internship at Åmic AB in Uppsala, Sweden, followed by an employment as research engineer at the company. Åmic AB developed *in vitro* diagnostic technologies for use in Point-of-Care. Delila worked in the assay development team, generating and screening new particle-antibody conjugates and was responsible for the purification facility and the direct labeling of antibodies with Eu-chelates. In July 2007 she decided to join the group of Prof Jan Trapman to conduct her PhD-studies at Erasmus University Medical Centre, Rotterdam, The Netherlands. The employment in Rotterdam started in September of 2007 and was funded by a Marie Curie Scholarship of the European Union and by Erasmus MC. The results of the research are described in this thesis.

ACKNOWLEDGEMENTS

Dear supervisor, **Prof. Trapman** I cannot find the words to express my gratitude. Your door has always been open to me and you always took the time for many crucial discussions concerning work but sometimes also life in general. I appreciate your broad knowledge and also your sense of humor. Thank you.

I am also grateful to my second promoter **Prof. Jenster**, for always making time when needed.

I would like to acknowledge all the members of my **PhD committee** for their time regarding my thesis and for adding valuable discussion.

A huge thanks to my second home in Rotterdam and members of Prof Trapmans lab. **Angelique, Carola, Petra, Hetty, Dennis, Michel, Hanneke, Karin, Ines** and **Joost**. **Angelique**, we have been roomies for many years now and I will never forget our old-school, special-tool RNA-isolation. **Carola** thanks for enduring the many, long work-discussions and for your thorough and accurate help the last year. **Joost**, for all the discussions and statistics. Also thanks to **Kimberly** and **Marije**, sharing space and their chocolate with me.

Big thanks to the colleagues from **Urology lab**, particularly **Wilma** and **Natasja** for preparing precious RNA and DNA patient plates and **Corrina** for mice work. Thank you **Arno** for always-efficient help with pathological evaluations.

My sincere thanks go to **Prof. Fodde** for adopting me to your international lab and for picking your members so well. **Matthias, Yaser, Rosalie, Patrick, Joel, Sabrina, Joana B, Joana M, Martin and Cate,** I have had most laughs with you as well as scientific discussions and geeky lunches. Extra thanks to **Patrick**, on top of always efficiently helping at work you became my first friend in Rotterdam. **Sabrin**a, I will never forget how unconditionally you where there for me. **Joana M**, with you I shared the deepest talks and insane laughs. Thank you **Joel**, for being our lab-medical support line and the best listener. **Joana B**, who would have thought that a new colleague would turn out to be the best roommate in the world, cheers for our many laughs and some tears. **Martin**, thank you for providing the most exquisite alternatives to kapsalon and for your support. **Ulli**, although your stay was short, our friendship will not be.

I would like to thank **Kees** for pedagogic introduction to IHC, for the most pleasant incubation times and preparing many important tissue-sections. **Ronald**, thank you for all the time spent starting me up with FISH and the pre-historic IF-microscope. **Frank** thank you for pimping up my figures and making them look decent. **Alex**, but also **Bert-**

Jaap and **Martijn** for help and support during the computer-crash! **Marcel Smid** for the efficient introduction and fast tutorial in brb.
Further I want to thank my friends and family. Mainly for the understanding and for still supporting me despite my neglect toward you caused in favor to this thesis.
First I want to thank my dear paranimfs; **Hetty**, I'm so glad we started to work together and I got not only the trust worthiest colleague but also a friend, I owe my deepest gratitude to you. **Annica**, my best friend, rock and never ending source of optimism from the day we met and surely for life.
How lucky am I to deserve your friendship **Christophe, Silvia, Rob, David, Fotis, Joana O, Astrid and Friederike**. Thank you for cheering for me and my work, sharing your homes, all the dinners, checking on me regularly even the periods when you only received silence back.
Johan A for always being so optimistic and pushing me "to get it over with already". **Jonas**, all great thoughts are indeed produced by long walks, mostly when you accompany them. Thank you **Caroline** for bringing support from Sweden with your many visits. **Anna-Karin** you were there for the last and most hectic touch.
My **uncle Zlatko**, for genuine interest in any task I have ever commenced since birth.
My **brother, mother** and **fathe**r, I can't imagine anyone feeling as loved as I do thanks to you and your support for me.
Max, thank you for giving me the freedom to pursue what I want and the home I need to be able to do so. Whatever difficulties I encounter, your patience and love brings me back on track. You are the warmth and the light of my life.

List of publications

Gasi Tandefelt D, Boormans J, van der Korput H, Jenster G and Trapman J. A thirty six gene signature predicts clinical progression in a subgroup of ERG-positive prostate cancers. (submitted)

Gasi Tandefelt D, Boormans J, Hermans KG and Trapman J. The role of ETS genes in prostate cancer. (submitted)

Gasi D, van der Korput H, Douwben J et al. Overexpression of Full-Length *ETV1* Transcripts in Clinical Prostate Cancer Due to Gene Translocation. PLoS One. 2011;26:6:e16332

Gasi D and Trapman J. Androgen Regulation of ETS-gene Fusion Transcripts in Prostate Cancer. Methods Mol Biol. 2011;776:335-48

Hermans KG, Boormans J, **Gasi D** et al. Overexpression of prostate-specific TMPRSS2(exon 0)-ERG fusion transcripts corresponds with favorable prognosis of prostate cancer. Clin Cancer Res. 2009; 15:6398-403

PhD Portfolio

Name PhD student:	Delila Gasi
Department:	Pathology
Research School:	Molecular Medicine, Erasmus MC
PhD Period:	2007-2012
Promoter:	Prof.dr. J. Trapman
Supervisor:	Prof.dr. J. Trapman

Courses/workshops	**Year**	**Hours**
Workshop Applied Bioinformatics	2007	8
Course Biomedical Research Techniques	2007	40
Workshop Browsing Genes and Genomes with UCSC	2007	8
Workshop Browsing Genes and Genomes with Ensembl	2007	16
Course Basic and Translation Oncology	2007	32
Course SNP's and Human Diseases	2008	32
Workshop Basic data analysis on gene expression arrays II	2008	8
Workshop Photoshop CS3	2010	8
Local Presentations		
Research group project and journal clubs	2007-2012	80
Pathology Department Prostate meetings	2011	32
Josephine Nefkens Institute weekly meetings	2007-2012	40
Oral presentations at meetings/conferences		
CANCURE, Innsbruck	2008	32
Androgens, Rotterdam	2008	32
CANCURE, Amsterdam	2009	27
AACR Annual meeting, Washington DC	2010	40

Poster presentations at meetings/conferences		
Molecular Medicine Day, Rotterdam	2008, 2009	32
PRIMA, Innsbruck	2008	32
ESUR, Barcelona	2008	40
AACR Prostate meeting, San Diego	2009	56
European Association of Urology, Amsterdam	2009	16
Teaching activities		
Medical Students Year 1, Erasmus MC, Rotterdam	2009	16
Medical Students Year 2, Erasmus MC, Rotterdam	2011	24
Fellowships and awards		
Marie Curie Fellowship for Early Stage Research Training	2007-2010	
Total hours		**651**
Total ECTS		**23**