Oncogene (2012) 31, 2153–2163

www.nature.com/onc

ORIGINAL ARTICLE

Overexpression of androgen receptor enhances the binding of the receptor to the chromatin in prostate cancer

A Urbanucci[1], B Sahu[2], J Seppälä[3], A Larjo[3], LM Latonen[1], KK Waltering[1], TLJ Tammela[4], RL Vessella[5], H Lähdesmäki[6], OA Jänne[2] and T Visakorpi[1]

[1]*Institute of Biomedical Technology, University of Tampere and Tampere University Hospital, Tampere, Finland;* [2]*University of Helsinki, Institute of Biomedicine Physiology, Biomedicum Helsinki, Helsinki, Finland;* [3]*Department of Signal Processing, Tampere University of Technology, Tampere, Finland;* [4]*Department of Urology, University of Tampere and Tampere University Hospital, Tampere, Finland;* [5]*Department of Urology, University of Washington Medical Center, NE Pacific, Seattle, WA, USA and* [6]*Department of Information and Computer Science, Aalto University School of Science and Technology, Helsinki, Finland*

Androgen receptor (AR) is overexpressed in the majority of castration-resistant prostate cancers (CRPCs). Our goal was to study the effect of AR overexpression on the chromatin binding of the receptor and to identify AR target genes that may be important in the emergence of CRPC. We have established two sublines of LNCaP prostate cancer (PC) cell line, one overexpressing AR 2–3-fold and the other 4–5-fold compared with the control cells. We used chromatin immunoprecipitation (ChIP) and deep-sequencing (seq) to identify AR-binding sites (ARBSs). We found that the number of ARBSs and the AR-binding strength were positively associated with the level of AR when cells were stimulated with low concentrations of androgens. In cells overexpressing AR, the chromatin binding of the receptor took place in 100-fold lower concentration of the ligand than in control cells. We confirmed the association of AR level and chromatin binding in two PC xenografts, one containing *AR* gene amplification with high AR expression, and the other with low expression. By combining the ChIP-seq and expression profiling, we identified AR target genes that are upregulated in PC. Of them, the expression of *ZWINT*, *SKP2* (S-phase kinase-associated protein 2 (p45)) and *FEN1* (flap structure-specific endonuclease 1) was demonstrated to be increased in CRPC, while the expression of *SNAI2* was decreased in both PC and CRPC. FEN1 protein expression was also associated with poor prognosis in prostatectomy-treated patients. Finally, the knock-down of *FEN1* with small interfering RNA inhibited the growth of LNCaP cells. Our data demonstrate that the overexpression of AR sensitizes the receptor binding to chromatin, thus, explaining how AR signaling pathway is reactivated in CRPC cells.
Oncogene (2012) **31,** 2153–2163; doi:10.1038/onc.2011.401; published online 12 September 2011

Keywords: prostatic neoplasia; AR; ChIP-seq; FEN1; SKP2; ZWINT

Correspondence: Professor T Visakorpi, Institute of Biomedical Technology, University of Tampere, Biokatu 6, Tampere, Finland FIN-33014, Finland.
E-mail: tapio.visakorpi@uta.fi
Received 3 April 2011; revised 30 July 2011; accepted 2 August 2011; published online 12 September 2011

Introduction

The development of prostate cancer (PC) is strongly dependent on androgens as evidenced by the finding that men castrated early in their life will not develop PC (Isaacs, 1994), and by trials indicating that lowering tissue levels of 5α-dihydrotestosterone (DHT) with 5α-reductase inhibitors, reduces the risk of PC (Thompson *et al*., 2003; Andriole *et al*., 2010). The efficacy of androgen deprivation in the treatment of PC was demonstrated >50 years ago (Huggins and Hodges, 2002) and castration still remains the main form of treatment for advanced PC. Despite the initial positive response, the castration-resistant PC (CRPC) phenotype will eventually emerge during the therapy. Earlier it was believed that PCs progressing during castration are androgen-independent (Thompson *et al*., 2003). Subsequently, the emergence of CRPC has been associated with increased expression of androgen receptor (AR), partly due to the amplification of the *AR* gene (Linja *et al*., 2001; Chen *et al*., 2004). Recently, it has been suggested that also a loss of *RB* gene could lead to AR overexpression (Sharma *et al*., 2010). In addition, mutations in *AR* altering transactivation properties of the receptor, expression of constitutively active AR splice variants and re-expression of androgen-regulated genes have been demonstrated in CRPC (Seruga *et al*., 2011). It has also been suggested that CRPC cells could themselves synthesize low levels of androgens from cholesterol (Seruga *et al*., 2011). Finally, recent phase II trials of CRPC with novel superantiandrogen, MDV3100, and CYP17 inhibitor, abiraterone, have directly demonstrated that CRPC cells are actually still androgen sensitive (Tran *et al*., 2009; Reid *et al*., 2010).

We showed more than a decade ago that one-third of CRPCs contain amplification of *AR* (Visakorpi *et al*., 1995). In addition, we have demonstrated by quantitative reverse transcriptase (qRT)–PCR that almost all CRPCs overexpress *AR* compared with hormone-naïve PC (Linja *et al*., 2001). However, expression of AR protein by immunohistochemistry seems to be variable in CRPC (Roudier *et al*., 2004). Later, Chen and co-workers (2004) showed that in a xenograft model system overexpression of *AR* is necessary and sufficient to

transform the androgen-dependent growth to an independent
one. To study the consequences of *AR* overexpression in PC cells, we have a stable transfected androgen-sensitive LNCaP PC cell line with wild-type *AR* and established two sublines. LNCaP-ARmo expresses 2–4 and LNCaP-ARhi 5–6 times higher level of AR protein than the control cells, LNCaP-pcDNA3.1 (Waltering *et al*., 2009). The LNCaP-ARhi cells grow faster in the presence of low levels of androgens than the control cells and the androgen-regulated genes are induced, on average, at 10-fold lower concentrations of DHT in the *AR* overexpressing compared with control cells. As the sublines share the same genomic background, the model is especially suitable for studying how AR promotes, maintains and drives the PC progression.

AR is a transcription factor that regulates the expression of hundreds of genes. Nevertheless, only one AR target gene that is commonly involved in the development of the disease, *TMPRSS2:ERG* fusion gene, has so far been identified (Tomlins *et al*., 2005). However, the fusion seems not to explain the phenotypical heterogeneity, including hormone responsiveness, of PC (Leinonen *et al*., 2010). The identification of such downstream genes could potentially provide new biomarkers and means to develop novel therapies.

Here, we utilized the LNCaP-based model as well as LuCaP xenografts to study the effect of the AR overexpression on the chromatin binding of AR by using chromatin immunoprecipitation and deep sequencing (ChIP-seq). In addition, we combined the ChIP-seq data with expression profiling to identify AR downstream genes that could be important in the progression of PC, and demonstrated overexpression of *ZWINT*, *SKP2* (S-phase kinase-associated protein 2 (p45)) and *FEN1* (flap structure-specific endonuclease 1) and reduced expression of SNAI2 in clinical samples of CRPC.

Results

In order to map AR-binding sites (ARBSs) across the genome, we first performed ChIP-seq for a total of nine samples: LNCaP-pcDNA3.1, -ARmo and -ARhi, treated for 2 h with 0, 1 and 100 nM of DHT (Supplementary Table S1). We found higher number of ARBSs in LNCaP-ARhi and -ARmo compared with control cells on stimulation with low concentration (1 nM) of DHT (Figure 1a). To confirm the association between AR level and the number of ARBSs in another model system, we utilized two PC xenografts, LuCaP69 and LuCaP73. They derive from castration-resistant tumors and have been grown in intact mice. We have previously demonstrated that LuCaP69 contains *AR* amplification, whereas LuCaP73 cells do not (Linja *et al*., 2001). The expression of AR is about 10-fold higher in LuCaP69 compared with LuCaP73 according to qRT–PCR. We found approximately 19 000 and 7000 ARBSs in LuCaP69 and LuCaP73, respectively, verifying that the level of AR is associated with the number of ARBSs.

To confirm the ChIP-seq data, we used traditional ChIP–qPCR. We studied the well-characterized enhancer and promoter regions of a known AR target gene, *PSA* (Shuur *et al*., 1996; Cleutjens *et al*., 1997), in the cell lines (Figure 1b) and xenografts (Supplementary Figure S1). The ChIP–qPCR data reproduced the ChIP-seq data. Furthermore, we were able to confirm previously identified ARBSs in AR target genes, such as *TMPRSS2* (Wang *et al*., 2007) (Supplementary Figure S1).

In the further analysis of the LNCaP-model, LNCaP-ARmo and -ARhi displayed high overlap of ARBSs in cells exposed to 1 nM DHT. On the contrary, in cells exposed to 100 nM DHT the LNCaP-ARmo and -ARhi cells showed less ARBSs overlap than -ARmo and control cells (Figure 1c and Supplementary Table S2). We then investigated the average peak heights that we assumed to represent the strength of the AR binding (Supplementary Table S2). The peak height was defined as the number of tags present in the specified loci of the ARBSs. The average peak height was higher in -ARhi and -ARmo cells compared with control cells at 1 nM DHT and lower at 100 nM DHT (Figure 1d and Supplementary Table S2).

As we used cells with various levels of AR and ligand concentrations, we were able to make comparison of the effect of AR and ligand on the binding profiles and on the chromatin loading of AR at the ARBSs (Supplementary Table S2). LNCaP-ARmo and -ARhi grown in 1 nM DHT had 985 common ARBSs, whereas pcDNA3.1 at 100 nM and ARmo at 1 nM had 1209, and pcDNA3.1 at 100 nM and ARhi at 1 nM 1323 ARBSs (Supplementary Table S2). By combining these ARBS maps, we constructed a high-confidence ARBS map of 1833 binding sites. The high-confidence ARBSs map includes all the binding sites with high reproducibility in the three samples mentioned above. We then re-analyzed the chromatin binding at these high-confidence ARBSs by computing the binding strength at each of the 1833 ARBSs. In this way, we were able to obtain normalized (against vehicle-treated cells) data on the AR-binding strength (that is, peak height) (Figure 1e). In cells stimulated with 1 nM DHT, the average peak height was greater in LNCaP-ARmo and -ARhi than in control cells. Whereas in cells stimulated with 100 nM DHT there seems to be a slight decrease in the peak height in LNCaP-ARhi cells.

Next, we analyzed the genomic localizations of ARBSs. About 40% of the high-confidence ARBSs were located in intronic regions and about 50% in distal intergenic regions. Thus, those regions probably include most of the enhancer elements (Figure 1f). Similar results were obtained if the ARBSs maps in every individual samples were taken into account.

We then performed a motif analysis by searching for the motifs deposited in TRANSFAC database (Matys *et al*., 2006). We found AR (canonical and 6-bp half site), and HNF3A (alias FOXA1) motifs to be the most significantly overrepresented ($P<10^{-6}$). Furthermore, by categorizing the ARBSs according to their genomic locations, also ETS family of transcription factor motifs

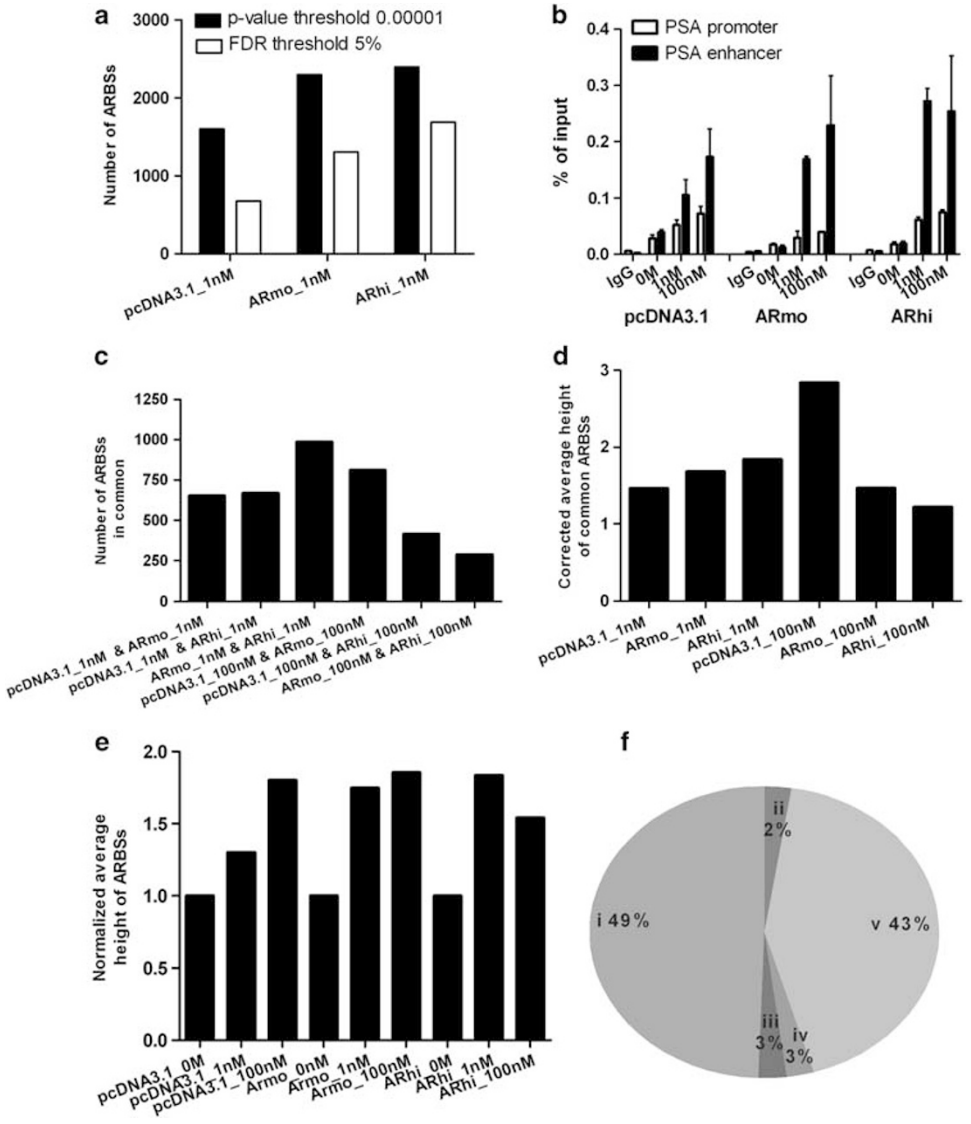


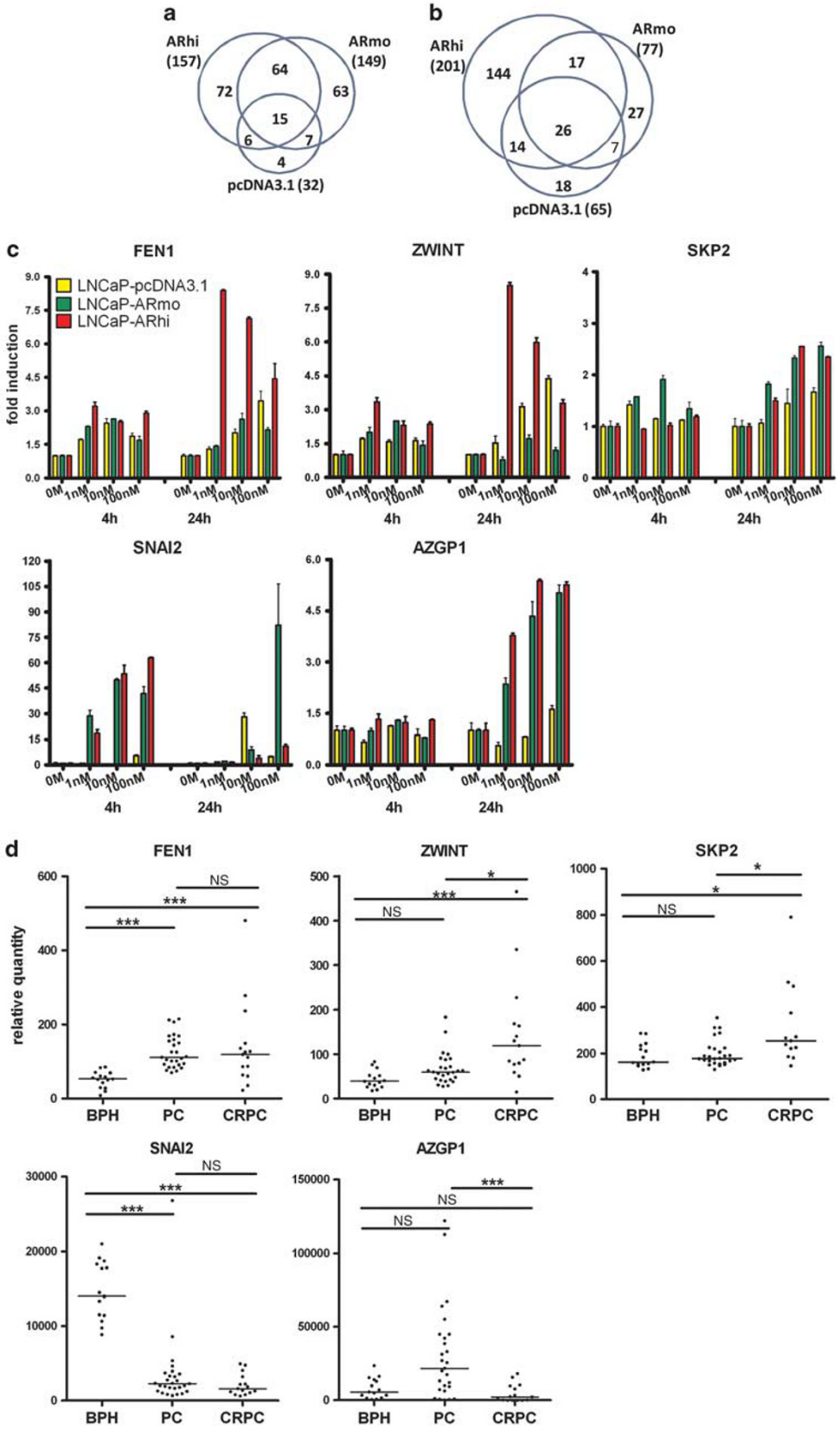
Figure 1 ChIP-seq data analyses. (**a**) Comparison of number of ARBSs between cell lines treated with 1 nM DHT according to peak detection with a *P*-value threshold of 0.00001 (black bar) and controlling also for false discovery rate (FDR) (white bar) at 5%. (**b**) AR binding to PSA promoter and enhancer in LNCaP-model. LNCaP-pcDNA3.1, -ARmo and -ARhi cells were hormone-starved for 4 days and treated for 2 h with DHT or ethanol (0 M). ChIP–qPCR was performed to assess the AR recruitment. Mean and s.e.m. are shown. (**c**) Number of ARBSs in common between the ChIP-seq samples. (**d**) Average background subtracted height of ARBSs in common between samples treated with 1 nM DHT (526 ARBSs in total) and between samples treated with 100 nM DHT (274 ARBSs in total) corrected according to the corresponding amount of raw reads (in millions) obtained in the sequencing (see Supplementary Table S1). (**e**) Normalized (against no DHT) average peak height of all high-confidence ARBSs in cell lines treated with different concentrations of DHT. (**f**) The genomic location of the 1833 high-confidence ARBSs. The high-confidence ARBSs were divided according to their location in distal intergenic regions (i), in exons (ii), within 1.5 kb upstream of transcription start site (TSS) (iii), within 1.5 kb downstream of 3′UTR (iv) and in introns (v).

were found to be enriched in promoters ($P = 3.3 \times 10^{-3}$), and within 1500-bp downstream of transcription start site ($P = 4 \times 10^{-4}$) as well as in exons ($P = 1.1 \times 10^{-2}$).

Next, we performed Gene Ontology (GO) enrichment analysis of genes located within a 25 kb window in the ARBS maps generated by pcDNA3.1 grown at 100 and 1 nM, -ARmo at 1 nM, as well as -ARhi at 1 nM DHT, by using GeneTrail (Keller *et al*., 2008). In 1 nM DHT-treated cells, the number of overrepresented GO (biological processes) categories increased from 5 in pcDNA3.1 to 22 in -ARmo, and 31 in ARhi. The processes that showed enrichment exclusively in -ARhi included, for example, cell–cell adhesion (GO:0016337), and regulation of locomotion (GO:0040012). Generally, in cells overexpressing AR, the same category of genes was enriched than in control cells, except in lower androgen concentration (Supplementary Table S3).

We then investigated the overlap of the high-confidence ARBSs map between the cell lines and the ARBSs maps of the xenografts. The overlap was surprisingly low. LuCaP69 showed 31% overlap with the high-confidence ARBS map in the cell lines, whereas

LuCaP73 only about 4%. When comparing individual samples, LNCaP-ARhi at 1 nM, -ARmo at 100 nM and -pcDNA3.1 at 100 nM showed the highest percentage of overlap with the LuCaP69 ARBS map. In order to confirm our finding, we compared LuCaP69 and LuCaP73 ARBS maps with the publicly available



LNCaP ARBS map from the work by Yu *et al*. (2010). We found even less overlap. The overlap with LuCaP69 was 6.5% and with LuCaP73 0.4%, whereas we found 86.8% overlap between our high-confidence ARBSs in LNCaP-model and the ARBSs of LNCaP cell line published by (Yu *et al*., 2010).

Finally, we integrated the high-confidence ARBS maps and expression profiles of mRNAs (Waltering *et al*., 2009) obtained from the LNCaP-model.

We used first hypergeometric distribution to compute enrichment *P*-values for differentially expressed genes and AR bound genes. When DHT concentrations were used to stimulate the cells, we observed that the differential expression, which was controlled directly by AR at the 4-h time point, still continued at the 24-h time point ($P<0.05$). Thus, we used both time points to identify direct targets of AR. Since we have previously shown that LNCaP-ARhi cells grow significantly faster in 1 nM DHT than control cells (Waltering *et al*., 2009), we focused on genes that showed AR binding and differential expression in that DHT concentration. The Venn diagrams in Figures 2a and b show that there are more such genes in LNCaP-ARhi and -ARmo than in control cells. The lists of 346 genes that were AR bound and androgen-regulated only in LNCaP-ARhi and -ARmo at 4- and/or 24-h time points are given in Supplementary Tables S4 and S5, respectively. Next, we interrogated the expression of these genes in clinical PC specimens by retrieving data from 14 independent array-based studies (see Supplementary Table S6 for references). We found that 38 out of these 346 genes (Supplementary Table S6) were overexpressed in PC according to, at least, one of the studies. Subsequently, of these we selected five putative target genes (*FEN1*, *ZWINT*, *SKP2*, *SNAI2* and *AZGP1*) based on information retrieved from the literature, and confirmed their androgen regulation (Figure 2c). Furthermore, in order to test the hypothesis that not only the androgens but also the amount of AR in the cells has an effect on such regulation, we performed a two-way analysis of variance in which the interaction between the effect on the variance of both AR amount and concentration of DHT was tested. The analysis indicated highly significant ($P<0.0001$) interaction and that both AR amount and the DHT concentration affect the gene expression significantly ($P<0.0001$).

Next, we used qRT–PCR to measure the expression of the genes in benign prostate hyperplasia (BPH), untreated PC and CRPC (Figure 2d). *SKP2*, *ZWINT* and *FEN1* transcripts were significantly overexpressed in CRPC when compared with PC and/or BPH, whereas the expression of *SNAI2* was reduced in CRPC and PC compared with BPH. The expression of *AZGP1* was significantly lower in CRPC than in PC. For *SKP2*, *ZWINT* and *FEN1*, we also confirmed the closest high-confidence ARBSs using ChIP–qPCR (Figure 3), and showed that AR binding is stronger in *AR* overexpressing than in control cells. To investigate the function of *FEN1*, *ZWINT* and *SNAI2* in PC, we suppressed them using small interfering RNAs (siRNAs) in both LNCaP-pcDNA3.1 and LNCaP-ARhi cells. *FEN1* depletion reduced significantly cell growth of both cell lines (Figure 4), while *SNAI2* and *ZWINT* depletions seem to give a growth advantage in control cells, but not in LNCaP-ARhi cells (Supplementary Figure S2). To study protein expression of FEN1, we first used western blotting in the cell line model and demonstrated increased protein expression in LNCaP-ARhi

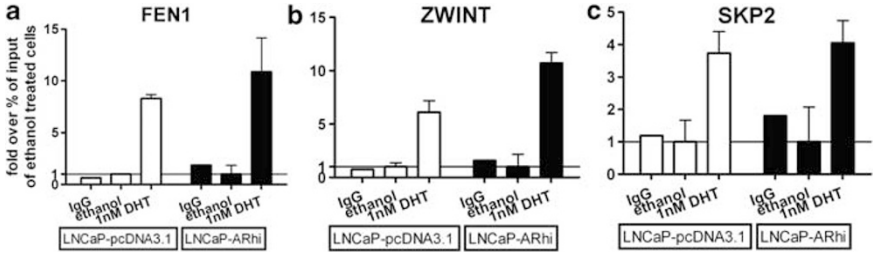


Figure 3 AR binding on putative enhancers of AR target genes. ChIP–qPCR on LNCaP-pcDNA3.1 and -ARhi cells, hormone starved for 4 days and treated for 2 h with 1 nM DHT or ethanol (0 M), was performed to assess the AR recruitment on the putative enhancers of *FEN1* (**a**), *ZWINT* (**b**) and *SKP2* (**c**) genes located 100 kb upstream (in the first intron of the gene DAGLA), 23 kb upstream and 120 kb upstream the above mentioned genes, respectively, according to our ChIPseq data. The data are presented as fold over percentage of input of the ethanol-treated sample. Mean ± s.e.m. are shown.

Figure 2 Identification of androgen-regulated AR target genes that are overexpressed in CRPC. The Venn diagrams showing the number of genes that are located in a window of 250 kb around the high-confidence ARBSs in LNCaP-pcDNA3.1, -ARmo and -ARhi cells and that showed at least 1.5-fold differential expression on 1 nM DHT stimulation for 4 h (**a**) and 24 h (**b**) (see Supplementary Tables S4 and S5). (**c**) Androgen regulation of AR target genes and effect of AR overexpression on their expression. LNCaP-pcDNA3.1, -ARmo and -ARhi cells were hormone starved for 4 days and subsequently treated with the indicated concentration of DHT or with vehicle (0 M). The expression of the genes was measured with qRT–PCR. Mean and s.e.m. of each gene against TBP values, normalized against the 0 M of each time point are shown. (**d**) Expression of the indicated AR target genes relative to average of three housekeeping genes (*TBP*, *β-actin* and *G3PDH2*) in BPH ($n=15$), PC ($n=27$) and CRPC ($n=13$) according to qRT–PCR. Kruskall–Wallis with Dunn *post-test* results are shown (***<0.001; *0.01 to 0.05; NS, not significant).

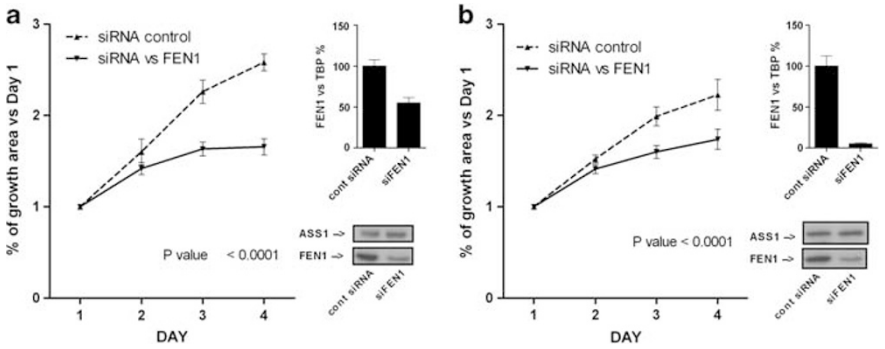


Figure 4 Functional significance of FEN1. Growth curve of siFEN1-transfected LNCaP-pcDNA3.1 (**a**) and LNCaP-ARhi (**b**) cells. Mean and ± s.d. are shown on different days. Statistical significance against control siRNA-transfected cells growth was assessed at day 4 by *t*-test. qRT–PCR at day 2.5 and western blot analysis at day 3 after transfection are also shown in each experiment confirming the FEN1 knockdown.

and -ARmo compared with control cells (Figure 5a). Next, we immunostained 185 untreated prostatectomy specimens as well as 92 CRPC samples (Figures 5b–d). Although cytoplasmic FEN1 staining was equal in PC and CRPC, the strong nuclear staining was observed significantly ($P<0.0001$) more often in CRPC than PC samples (Figure 5e). Only 5/185 (3%) prostatectomy samples showed nuclear staining in >10% of malignant cells. Interestingly, these cases had short time for biochemical recurrence (Figure 5f).

Discussion

Overexpression of AR is a common feature in CRPC (Linja *et al*., 2001), and it has been shown to sensitize cells to low levels of androgens (Kokontis *et al*., 1998; Chen *et al*., 2004; Waltering *et al*., 2009). Here, we utilized our previously established LNCaP-based model (Waltering *et al*., 2009), expressing different levels of AR to interrogate the effect of both ligand and the receptor on chromatin binding of AR. The data here indicated that both the ligand concentration and the amount of receptor affect together the chromatin binding of AR. A modest overexpression of AR enhances the chromatin binding of the receptor by sensitizing the cells to 100-fold lower ligand concentration. The majority of the previously reported ChIP-seq and ChIP-chip experiments (Jia *et al*., 2008; Wang *et al*., 2009; Takayama *et al*., 2010; Yu *et al*., 2010) have compared the binding of AR at the saturating concentration of androgens, and thus, missed the dynamics of AR binding. However, our data are in line with a recent work by Massie *et al*. (2011), in which they found almost five times more ARBSs in the strongly AR-overexpressing cell line VCaP compared with LNCaP cells. The data here are also consistent with our previous finding (Waltering *et al*., 2009) that androgen-regulated genes are induced in lower ligand concentrations in cells overexpressing AR. Thus, the increased chromatin-binding capacity of the receptor because of the overexpression of the receptor provides also a mechanistic explanation to the progression of PC in the presence of only low levels of androgens.

We confirmed the effect of AR levels on chromatin binding of the receptor also in another independent model system. There were almost three times more ARBSs in LuCaP69 than in LuCaP73 PC xenografts verifying that the level of AR is associated with the number of ARBSs. LuCaP69 contains *AR* gene amplification and 10-fold higher expression of *AR* than LuCaP73 (Linja *et al*., 2001). The reliability of the ChIP-seq data was, on the other hand, confirmed by traditional ChIP–qPCR of *PSA*, and our ARBS maps were able to confirm previously reported ARBS like for *TMPRSS2* enhancer. Furthermore, we obtained a high degree of ARBSs overlap with previously published data set on the same cell line.

Although the ChIP–qPCR data confirmed increased binding in the *PSA* enhancer in AR-overexpressing compared with control cell, no such loading difference was seen in the promoter region (Figure 1b). Previous studies (Kang *et al*., 2004; Wang *et al*., 2005) have pinpointed the importance of the dynamics of the AR recruitment in the *PSA* regulatory regions. Here, we studied the AR recruitment only 2 h after DHT stimulation. Thus, we cannot exclude the possibility that there are differences in the AR recruitment at the promoter of *PSA* between the cell lines at later time points.

Unlike most of the previous AR ChIP-chip and ChIP-seq studies (Massie *et al*., 2007; Wang *et al*., 2007, 2009; Jia *et al*., 2008; Takayama *et al*., 2010; Yu *et al*., 2010), we used several ligand concentrations and several LNCaP derivative cell lines. Thus, we were able to produce highly reproducible ARBSs data, which we called high-confidence ARBSs map. We utilized that map in comparison of the cell line model and the xenograft as well as in ARBS localizations, motif and ontology analyses, and also in identification of the critical AR target genes. The poor overlap (from 4 to 31%) of the high-confidence ARBSs map between the cell lines and the xenografts emphasizes that AR binding to chromatin varies significantly between tissue samples suggesting that genetic or other intrinsic differences,

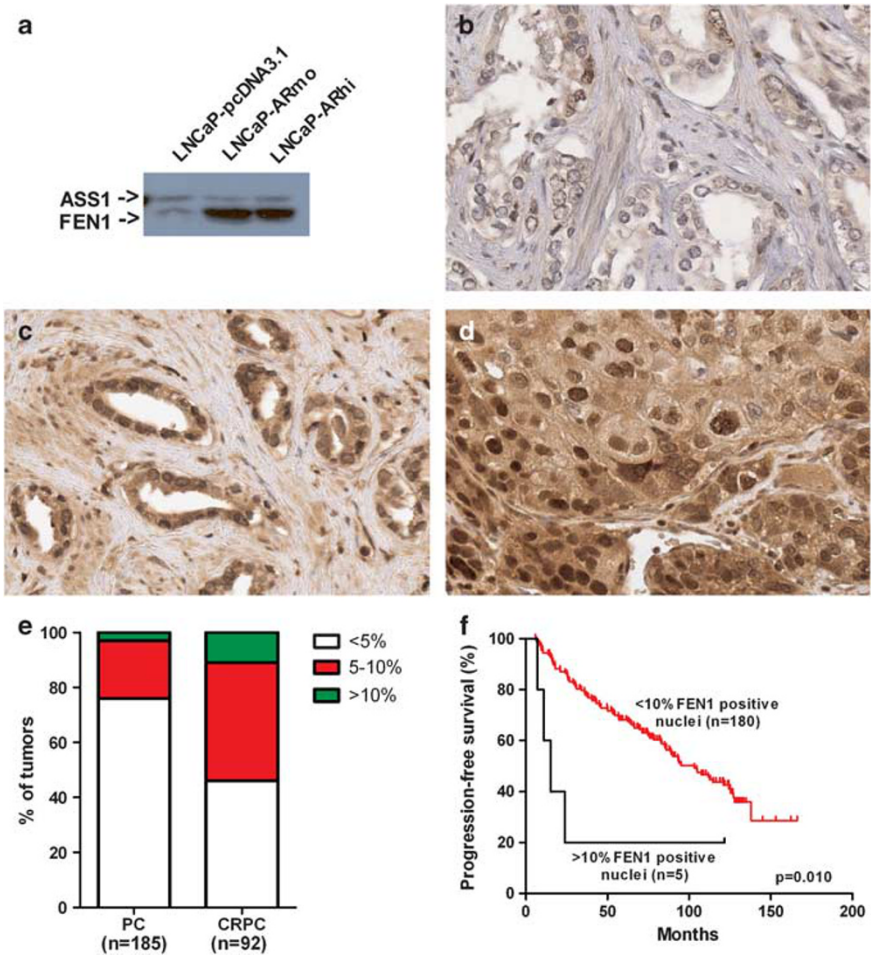


Figure 5 AR overexpression increases FEN1 protein. (**a**) AR overexpression increases FEN1 protein production. Western blot analysis of LNCaP-pcDNA3.1, -ARmo and -ARhi cells grown in normal medium showing FEN1 protein being overexpressed in AR-overexpressing cells compared with control cells. Anti-ATP synthase subunit alpha (ASS1) antibody was used as loading control. Immunohistochemical staining of (**b**) untreated PC with no cytoplasmic or nuclear staining, (**c**) untreated PC, and (**d**) CRPC specimen with strong nuclear staining in almost all malignant cells with monoclonal anti-FEN1 antibody. (**e**) The percentage of tumors according to percentage of positive nuclei in PC ($n = 185$) and CRPC ($n = 92$) specimens ($P < 0.0001$ according to χ^2 test). (**f**) Kaplan–Meier analysis of biochemical progression-free survival in prostatectomy-treated patients according to the percentage of FEN1 positive nuclei. Five patients with high frequency of FEN1-positive nuclei had very short progression-free time. The *P*-value was calculated with Mantel–Cox test.

such as binding of other transcription factors, in the cells could contribute strongly to the AR binding (Zinzen *et al*., 2009; Kasowski *et al*., 2010). This notion is also supported by our re-analysis of the publicly available data from the work by Yu and co-authors (2010). According to the re-analysis, Yu *et al*. obtained 58.1% overlap of ARBSs between LNCaP and only 28.7% overlap between VCaP and a tumor sample. The same tumor tissue sample showed 44.1% overlap with our high-confidence ARBSs, while the overlap with the LuCaP69 ARBSs was only 17.9%. Although the two xenografts overlapped poorly within each other and also with the high-confidence ARBSs map of LNCaP, the ARBSs in the xenografts, localized often close to the genes that showed ARBSs and androgen regulation in the LNCaP-model. This suggest that androgen-regulated genes may have alternative ARBSs. The poor overlap between the cell lines and xenografts may obviously also be due to the microenvironmental (cell culture versus mouse) differences.

The genomic localization of the ARBSs indicated that most of the androgen regulation is mediated by binding of AR to the distal intergenic elements and intronic regions as previously suggested (Jia *et al*., 2008; Lupien *et al*., 2008; Yu *et al*., 2010), instead of, for example, to promoter regions. Also, the motif analyses were in concordance with previously published findings (Massie *et al*., 2007; Lupien *et al*., 2008; Wang *et al*., 2009; Wei *et al*., 2010; Yu *et al*., 2010) showing that binding sites of, especially, FOXA1 and ETS family of transcription factors are enriched in the vicinity of ARBSs. FOXA1 has previously been suggested to be a pioneering factor for binding of other transcription factors to chromatin (Lupien *et al*., 2008, Wang *et al*., 2011), which is in line

with our data. Also, the ontology analysis of genes located within a 25 kb window of the ARBS maps was consistent with our previous findings based on the expression profiling of the LNCaP model (Waltering *et al*., 2009). Cell–cell adhesion and regulation of locomotion were among the most enriched ontologies.

Finally, to identify AR target genes that could be important in the progression of PC, we combined the high-confidence ARBS map with expression profiling of the LNCaP model. The analyses indicated that there are more genes, which show ARBSs in LNCaP-ARhi and -ARmo than in control cells. The data are consistent with our previously published findings on the expression of AR target genes in this model (Waltering *et al*., 2009). Thus, in cells overexpressing AR, less ligand is needed for induction of target gene expression. We then further studied the expression of *FEN1*, *ZWINT*, *SKP2*, *SNAI2* and *AZGP1* in clinical PC first by data mining publicly available microarray data, and subsequently by using qRT–PCR to measure the expression in BPH, PC and CRPC. Of those, the expression of *SKP2*, *ZWINT* and *FEN1* transcripts were significantly overexpressed in CRPC when compared with PC and/or BPH while the expression of *SNAI2* was reduced in cancer compared with BPH. Also, recently published data in advanced PC confirm the overexpression of these genes (Taylor *et al*., 2010). In addition, we confirmed the androgen regulation and the ARBSs of these genes by qRT–PCR, and ChIP–qPCR, respectively.

SKP2 is known to have oncogenic properties and to be overexpressed in many cancers (Nakayama and Nakayama, 2006). It has previously been shown to be androgen regulated (Waltregny *et al*., 2001) and the expression being associated with a short biochemical recurrence following prostatectomy (Nguyen *et al*., 2011). However, to our knowledge, this is the first report showing elevated levels of *SKP2* transcripts in CRPCs. A recent study by Lin and co-authors (2010) suggested *SKP2* as a potential target in cancer treatment and prevention, because inhibition of SKP2 triggered cellular senescence in p53/PTEN-deficient PC-3 cells and tumor regression in mice. *ZWINT* (ZW10 interactor), on the other hand, encodes a kinetokore protein (Starr *et al*., 2000) that has been suggested to have a role in the development of some malignancies (Obuse *et al*., 2004; Kops *et al*., 2005; Lin *et al*., 2006). We have previously shown *ZWINT* to be an androgen-regulated gene (Waltering *et al*., 2009). Depletion of ZWINT in LNCaP-pcDNA3.1 cells resulted in faster growth, whereas there was no significant effect on LNCaP-ARhi cells. Thus, the functional data are discordant with the finding of overexpression of the gene in cancer. SNAI2 (snail homolog 2 (*Drosophila*)) is a zinc-finger transcriptional repressor involved in epithelium to mesenchyme transition (Thiery, 2002). Depletion of also SNAI2 in LNCaP-pcDNA3.1, but not in LNCaP-ARhi, cells resulted in faster growth, which is concordant with the reduced expression in cancer.

FEN1 encodes a structure-specific metallonuclease that interacts with several other proteins involved in DNA replication, apoptosis, DNA repair and telomere stability (Zheng *et al*., 2010). Somatic mutations of *FEN1* have been reported in several common cancers (Zheng *et al*., 2007). These mutations abolish the exonuclease activity but retain the flap endonuclease activity (Zheng *et al*., 2007), which is consistent with the finding that mice carrying mutations show higher chemically induced cancer incidence (Xu *et al*., 2011). The mutations of *FEN1* in PC are yet to be investigated, however, FEN1 has been reported to be overexpressed in PC, especially in high Gleason score tumors (Lam *et al*., 2006). Here, we showed that the high frequency of nuclear staining was found significantly ($P<0.0001$) more often in CRPC than in PC, and that the staining was associated with poor prognosis in prostatectomy-treated patients. Thus, FEN1, as an androgen-regulated, overexpressed and associated with aggressive phenotype of the disease, could be an important AR downstream gene and, therefore, a putative drug target. Also, our preliminary functional data further suggest that FEN1 could promote the growth of PC cells, since the knockdown of FEN1 significantly reduced the growth of both control and AR-overexpressing cells.

In conclusion, we demonstrated that chromatin binding of AR is dependent not just on the ligand concentration, but also on the level of the receptor. Thus, the overexpression of *AR* in CRPC cells allows these cells to activate the AR signaling even in low androgen concentrations. By combining ChIP-seq and expression data, we were able to identify genes whose expression is directly regulated by AR, and are transcriptionally upregulated in PC and CRPC. These genes could be important in the progression of PC.

Materials and methods

Cell line and cell culture procedure

The establishment of LNCaP cells overexpressing AR has been described previously (Waltering *et al*., 2009). The cells were maintained under geneticin 250 μg/ml (Invitrogen Inc., Carlsbad, CA, USA). The hormone treatments and RNA extractions were performed as previously described (Waltering *et al*., 2011). LNCaP, VCaP and LAPC4 cells were purchased from ATCC (LGC/ATCC, Rockville, MD, USA) and maintained according to the manufacturer's instructions.

Xenografts material

Two PC xenografts, LuCaP69 and LuCaP73, grown in intact male mice, were provided by one of the investigators (RLV).

Clinical samples

Freshly frozen 8 BPH and 27 untreated primary PC samples from prostatectomies, as well as 7 BPH and 15 CRPC specimens from transurethral resection of the prostate-treated patients were used in the study. The samples were snap frozen in liquid nitrogen and total RNA was isolated with Trizol-Reagent (Invitrogen Inc.) according to the manufacturer's instructions. Tumor samples contained, at least, 70% of cancer cells. The use of the clinical material has been approved by the ethical committee of the Tampere University Hospital.

Tissue microarrays contained 185 formalin-fixed paraffin-embedded prostatectomy and 92 CRPC (transurethral resection of the prostate) specimens obtained from Tampere University Hospital. For the prostatectomy-treated patients, detectable PSA values ($\geqslant$0.5 ng/ml) in two consecutive measurements or the emergence of metastases were considered as signs of progression. The use of tissue microarrays has been approved by the ethical committee of Tampere University Hospital and the National Authority for Medicolegal Affairs.

ChIP, ChIP-seq assays and data analysis
Four million cells were plated and hormone-deprived for 4 days and treated with DHT at different concentrations for 2 h. Cells were fixed by adding formaldehyde (Merck KGaA, Darmstadt, Germany) in 1% final concentration for 10 min at room temperature and lysed in 1% sodium dodecyl sulfate, 10 mM EDTA, 50 mM Tris–HCl containing 2X protease inhibitor (Roche Inc., Mannheim, Germany). To perform tissue ChIP, 3 ml of phosphate-buffered saline containing 2X protease inhibitor (Roche Inc.) were added to 40×20 µm sections of freshly frozen xenograft specimens. They were first vigorously mixed three times with syringe and 14G needle, then four times with 25G needle. The cells were fixed for 10 min in room temperature by adding 1/10 volume of fixation solution (11% formaldehyde, 0.1 M NaCl, 1 mM EDTA, 0.5 mM EGTA, 50 mM HEPES). Fixation was stopped by adding 1/20 volume of 2.5 M glycine for 5 min at room temperature. The cells were pelleted, washed twice in phosphate-buffered saline containing 2X protease inhibitor (Roche Inc.) and lysed as above. The chromatin was immunoprecipitated with 10 µg of normal rabbit immunoglobulin G (Santa Cruz Inc., Santa Cruz, CA, USA) or 10 µl of anti-AR polyclonal antibody (AR3) (provided by one of the investigators: OAJ) (Karvonen *et al*., 1997, Thompson *et al*., 2006). Supplementary Figure S3 shows a validation of the AR3 antibody in western blot and ChIP assay. The libraries of ChIP DNA were prepared and sequenced with Genome Analyzer II (Illumina Inc., San Diego, CA, USA) according to the manufacturer's protocol. Detailed descriptions of the ChIP procedure, sequencing and detailed data analysis are included in Supplementary Information.

mRNA expression profiling
mRNA expression data with Illumina platform (including RefSeq genes) (Illumina Inc.) were retrieved from the studies by Waltering *et al*. (2009). A detailed description of the raw data analysis is provided in the Supplementary Information.

Quantitative PCR assays
For mRNA expression analyses, first-strand complementary DNA synthesis was performed from total RNA using AMV reverse transcriptase (Finnzymes Inc., Espoo, Finland) according to the manufacturer's instructions. The relative expression of each gene against the average value of *TBP, G3PDH2* and *β-actin* reference genes was measured with Maxima SYBR Green (Fermentas Inc., Burlington, Ontario, Canada) and CFX96 real-time PCR detection system (Bio-Rad Laboratories Inc., Hercules, CA, USA) essentially as previously described (Urbanucci *et al*., 2008). For the ChIP–qPCR analysis, the enrichment relative to input chromatin was calculated according to the delta Ct method with the percentages been calculated using the formula $2^{-\Delta Ct}$, where ΔCt is Ct(ChIP-template)-Ct(Input). A standard curve from one of the diluted input was included in the run to control that the efficiency of the reaction would be maintained in the range between 95 and 105%. A qPCR on a control region in which AR is not supposed to bind, between PSA enhancer and promoter (middle region) was performed for each ChIP assay. The ChIP assay was considered specific if in the control region the enrichment was not above the enrichment of the non-specific immunoprecipitated sample made with normal rabbit immunoglobulin G. The primers used are listed in Supplementary Table S7.

Western blot
Western Blot was performed from total cell lysates. The primary antibodies used were anti-FEN1 (clone 4E7; LifeSpan Biosciences Inc., Seattle, WA, USA) and anti-ATP synthase subunit alpha (clone 15H4C4; MitoSciences Inc., Eugene, OR, USA) monoclonal antibodies. A detailed description of the western blot procedure is provided in the Supplementary Information.

Immunohistochemistry
Mouse anti-FEN1 (mAb clone 4E7; LifeSpan Biosciences Inc.) was used with Power Vision + Poly-HRP IHC kit (ImmunoVision Technologies Co., Burlingame, CA, USA) according to the manufacturer's instructions. The protocol has previously been described (Leinonen *et al*., 2010).

siRNA transfections
Silencer selected siRNAs from Ambion (Applied Biosystems/Ambion, Austin, TX, USA) were used. Cells were transfected with INTERFERin transfection reagent (Polyplus-transfection, Illkirch, France) according to the manufacturer's protocol. Briefly, 20 000 cells/24-well plate were seeded and transfected with 20 nM of siFEN1 (s5103), 20 nM each of siZWINT (s21949 and s21951), 20 nM each of siSNAI2 (s13127 and s13128) or equal concentration of *Silencer* negative control siRNA #1. Expression levels of *FEN1*, *ZWINT* and *SNAI2* relative to *TBP* were measured by qRT–PCR (2.5 days after transfection) and protein levels by Western blot analysis (3 days after transfection).

Growth curves
Growth curve measurements were started 1 day after siRNA transfection and marked as day 1. Images of the same growth area in each well were acquired every day using a Retiga-2000R FAST Cooled Mono 12-bit camera (QImaging Inc., Surrey, BC Canada) mounted on a Motorized Inverted Research Microscope IXT1 (Olympus America Inc., Center Valley, PA, USA) and a 10X objective. The total growth area occupied by cells (area percent) in each well was determined each day of measurement with ImageJ software (Abramoff *et al*., 2004) and normalized against the growth area of the relative well at day 1. Four replicates were used in each siRNA experiment. *T*-test was used to assess significance of differences in growth curves at day 4.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We wish to thank Ms Mariitta Vakkuri, Ms Päivi Martikainen and Mr Rolle Rahikainen for the skillful technical assistance. The research leading to these results has received funding from the European Union FP6, CANCURE Programme (contract number: MEST-CT-2005-020970). In addition, grant support has been received from Academy of Finland, Cancer Society

of Finland, Reino Lahtikari Foundation, Sigrid Juselius Foundation, and the Medical Research Fund of Tampere University Hospital.

Author contributions: TV and AU designed the study, analyzed the data and wrote the paper. AU and BS performed experiments and edited the paper. JS, AL, HL performed the bioinformatic analyses and edited the paper. LML performed experiments and edited the paper. KKW provided the expression data and edited the paper. OAJ contributed new reagents and edited the paper. TLT provided clinical material and edited the paper. RLV provided the xenografts material and edited the paper.

References

Abramoff MD, Magelhaes PJ, Ram SJ. (2004). Image Processing with ImageJ. *Biophotonics Int* **11**: 36–42.

Andriole GL, Bostwick DG, Brawley OW, Gomella LG, Marberger M, Montorsi F *et al*. (2010). Effect of dutasteride on the risk of prostate cancer. *N Engl J Med* **362**: 1192–1202.

Chen CD, Welsbie DS, Tran C, Baek SH, Chen R, Vessella R *et al*. (2004). Molecular determinants of resistance to antiandrogen therapy. *Nat Med* **10**: 33–39.

Cleutjens KBJM, van der Korput HAGM, van Eekelen CCEM, van Rooij HCJ, Faber PW, Trapman J. (1997). An androgen response element in a far upstream enhancer region is essential for high, androgen-regulated activity of the prostate-specific antigen promoter. *Mol Endocrinol* **11**: 148–161.

Huggins C, Hodges CV. (2002). Studies on prostatic cancer, I: the effect of castration, of estrogen and of androgen injection on serum phosphatases in metastatic carcinoma of the prostate 1941. *J Urol* **168**: 9–12.

Isaacs JT. (1994). Role of androgens in prostatic cancer. *Vitam Horm* **49**: 433–502.

Jia L, Berman BP, Jariwala U, Yan X, Cogan JP, Walters A *et al*. (2008). Genomic androgen receptor-occupied regions with different functions, defined by histone acetylation, coregulators and transcriptional capacity. *PLoS One* **3**: e3645.

Kang Z, Jänne OA, Palvimo JJ. (2004). Coregulator recruitment and histone modifications in transcriptional regulation by the androgen receptor. *Mol Endocrinol* **11**: 2633–2648.

Karvonen U, Kallio PJ, Jänne OA, Palvimo JJ. (1997). Interaction of androgen receptors with androgen response element in intact cells. *J Biol Chem* **272**: 15973–15979.

Kasowski M, Grubert F, Heffelfinger C, Hariharan M, Asabere A, Waszak SM *et al*. (2010). Variation in transcription factor binding among humans. *Science* **328**: 232–235.

Keller A, Backes C, Al-Awadhi M, Gerasch A, Küntzer J, Kohlbacher O *et al*. (2008). GeneTrailExpress: a web-based pipeline for the statistical evaluation of microarray experiments. *BMC Bioinformatics* **9**: 552.

Kokontis JM, Hay N, Liao S. (1998). Progression of LNCaP prostate tumor cells during androgen deprivation: hormone-independent growth, repression of proliferation by androgen, and role for p27kip1 in androgen-induced cell cycle arrest. *Mol Endocrinol* **12**: 941–953.

Kops GJ, Kim Y, Weaver BA, Mao Y, McLeod I, Yates III JR *et al*. (2005). ZW10 links mitotic checkpoint signaling to the structural kinetochore. *J Cell Biol* **169**: 49–60.

Lam JS, Seligson DB, Yu H, Li A, Eeva M, Pantuck AJ *et al*. (2006). Flap endonuclease 1 is overexpressed in prostate cancer and is associated with a high Gleason score. *BJU Int* **98**: 445–451.

Leinonen KA, Tolonen TT, Bracken H, Stenman UH, Tammela TL, Saramäki OR *et al*. (2010). Association of SPINK1 expression and TMPRSS2:ERG fusion with prognosis in endocrine-treated prostate cancer. *Clin Cancer Res* **16**: 2845–2851.

Lin HK, Chen Z, Wang G, Nardella C, Lee SW, Chan CH *et al*. (2010). Skp2 targeting suppresses tumorigenesis by Arf-p53-independent cellular senescence. *Nature* **464**: 374–379.

Lin YT, Chen Y, Wu G, Lee WH. (2006). Hec1 sequentially recruits Zwint-1 and ZW10 to kinetochores for faithful chromosome segregation and spindle checkpoint control. *Oncogene* **25**: 6901–6914.

Linja MJ, Savinainen KJ, Saramäki OR, Tammela TL, Vessella RL, Visakorpi T. (2001). Amplification and overexpression of androgen receptor gene in hormone-refractory prostate cancer. *Cancer Res* **61**: 3550–3555.

Lupien M, Eeckhoute J, Meyer CA, Wang Q, Zhang Y, Li W *et al*. (2008). FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132**: 958–970.

Massie CE, Adryan B, Barbosa-Morais NL, Lynch AG, Tran MG, Neal DE *et al*. (2007). New androgen receptor genomic targets show an interaction with the ETS1 transcription factor. *EMBO Rep* **8**: 871–878.

Massie CE, Lynch A, Ramos-Montoya A, Boren J, Stark R, Fazli L *et al*. (2011). The androgen receptor fuels prostate cancer by regulating central metabolism and biosynthesis. *EMBO J* **30**: 2719.

Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A *et al*. (2006). TRANSFAC(r) and its module TRANSCompel(r): transcriptional gene regulation in eukaryotes. *Nucleic Acids Res* **34**: D108–D110.

Nakayama KI, Nakayama K. (2006). Ubiquitin ligases: cell-cycle control and cancer. *Nat Rev Cancer* **6**: 369–381.

Nguyen PL, Lin DI, Lei J, Fiorentino M, Mueller E, Weinstein MH *et al*. (2011). The impact of Skp2 overexpression on recurrence-free survival following radical prostatectomy. *Urol Oncol* **29**: 302–308.

Obuse C, Iwasaki O, Kiyomitsu T, Goshima G, Toyoda Y, Yanagida M. (2004). A conserved Mis12 centromere complex is linked to heterochromatic HP1 and outer kinetochore protein Zwint-1. *Nat Cell Biol* **6**: 1135–1141.

Reid AH, Attard G, Danila DC, Oommen NB, Olmos D, Fong PC *et al*. (2010). Significant and sustained antitumor activity in post-docetaxel, castration-resistant prostate cancer with the CYP17 inhibitor abiraterone acetate. *J Clin Oncol* **28**: 1489–1495.

Roudier MP, Corey E, True LD, Higano CS, Ott SM, Vessella RL. (2004). Histological, immunophenotypic and histomorphometric characterization of prostate cancer bone metastases. In: Keller E, Chung L (eds). *The Biology of Skeletal Metastases* vol. 118. Kluwer Academic Publishers: Boston, MA, USA, pp 311–339.

Schuur ER, Henderson GA, Kmetec LA, Miller JD, Lamparski HG, Henderson DR. (1996). Prostate-specific antigen expression is regulated by an upstream enhancer. *J Biol Chem* **271**: 7043–7051.

Seruga B, Ocana A, Tannock IF. (2011). Drug resistance in metastatic castration-resistant prostate cancer. *Nat Rev Clin Oncol* **8**: 12–23.

Sharma A, Yeow WS, Ertel A, Coleman I, Clegg N, Thangavel C *et al*. (2010). The retinoblastoma tumor suppressor controls androgen signaling and human prostate cancer progression. *J Clin Invest* **120**: 4478–4492.

Starr DA, Saffery R, Li Z, Simpson AE, Choo KH, Yen TJ *et al*. (2000). HZwint-1, a novel human kinetochore component that interacts with HZW10. *J Cell Sci* **113**: 1939–1950.

Takayama K, Tsutsumi S, Katayama S, Okayama T, Horie-Inoue K, Ikeda K *et al*. (2010). Integration of cap analysis of gene expression and chromatin immunoprecipitation analysis on array reveals genome-wide androgen receptor signaling in prostate cancer cells. *Oncogene* **30**: 619–630.

Taylor BS, Schultz N, Hieronymus H, Gopalan A, Xiao Y, Carver BS *et al*. (2010). Integrative genomic profiling of human prostate cancer. *Cancer Cell* **18**: 11–22.

Thiery JP. (2002). Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer* **6**: 442–454.

Thompson IM, Goodman PJ, Tangen CM, Lucia MS, Miller GJ, Ford LG *et al*. (2003). The influence of finasteride on the development of prostate cancer. *N Engl J Med* **349**: 215–224.

Thompson J, Lepikhova T, Teixido-Travesa N, Whitehead MA, Palvimo JJ, Jänne OA. (2006). Small carboxyl-terminal domain phosphatase 2 attenuates androgen-dependent transcription. *EMBO J* **25**: 2757–2767.

Tomlins SA, Rhodes DR, Perner S, Dhanasekaran SM, Mehra R, Sun XW *et al*. (2005). Recurrent fusion of TMPRSS2 and ETS transcription factor genes in prostate cancer. *Science* **310**: 644–648.

Tran C, Ouk S, Clegg NJ, Chen Y, Watson PA, Arora V *et al*. (2009). Development of a second-generation antiandrogen for treatment of advanced prostate cancer. *Science* **324**: 787–790.

Urbanucci A, Waltering KK, Suikki HE, Helenius MA, Visakorpi T. (2008). Androgen regulation of the androgen receptor coregulators. *BMC Cancer* **8**: 219.

Visakorpi T, Hyytinen E, Koivisto P, Tanner M, Keinänen R, Palmberg C *et al*. (1995). *In vivo* amplification of the androgen receptor gene and progression of human prostate cancer. *Nat Genet* **9**: 401–406.

Waltering KK, Helenius MA, Sahu B, Manni V, Linja MJ, Jänne OA *et al*. (2009). Increased expression of androgen receptor sensitizes prostate cancer cells to low levels of androgens. *Cancer Res* **69**: 8141–8149.

Waltering KK, Porkka KP, Jalava SE, Urbanucci A, Kohonen P, Latonen L *et al*. (2011). Androgen regulation of microRNAs in prostate cancer. *Prostate* **71**: 604–614.

Waltregny D, Leav I, Signoretti S, Soung P, Lin D, Merk F *et al*. (2001). Androgen-driven prostate epithelial cell proliferation and differentiation *in vivo* involve the regulation of p27. *Mol Endocrinol* **15**: 765–782.

Wang D, Garcia-Bassets I, Benner C, Li W, Su X, Zhou Y *et al*. (2011). Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature* **474**: 390–394.

Wang Q, Carroll JS, Brown M. (2005). Spatial and temporal recruitment of androgen receptor and its coactivatos involves chromosomal looping and polymerase tracking. *Mol Cell* **19**: 631–642.

Wang Q, Li W, Liu XS, Carroll JS, Jänne OA, Keeton EK *et al*. (2007). A hierarchical network of transcription factors governs androgen receptor-dependent prostate cancer growth. *Mol Cell* **27**: 380–392.

Wang Q, Li W, Zhang Y, Yuan X, Xu K, Yu J *et al*. (2009). Androgen receptor regulates a distinct transcription program in androgen-independent prostate cancer. *Cell* **138**: 245–256.

Wei GH, Badis G, Berger MF, Kivioja T, Palin K, Enge M *et al*. (2010). Genome-wide analysis of ETS-family DNA-binding *in vitro* and *in vivo*. *EMBO J* **29**: 2147–2160.

Xu H, Zheng L, Dai H, Zhou M, Hua Y, Shen B. (2011). Chemical-induced cancer incidence and underlying mechanisms in Fen1 mutant mice. *Oncogene* **30**: 1072–1081.

Yu J, Yu J, Mani RS, Cao Q, Brenner CJ, Cao X *et al*. (2010). An integrated network of androgen receptor, polycomb, and TMPRSS2-ERG gene fusions in prostate cancer progression. *Cancer Cell* **17**: 443–445.

Zheng L, Dai H, Zhou M, Li M, Singh P, Qiu J *et al*. (2007). Fen1 mutations result in autoimmunity, chronic inflammation and cancers. *Nat Med* **13**: 812–819.

Zheng L, Jia J, David Finger L, Guo Z, Zer C, Shen B. (2010). Functional regulation of FEN1 nuclease and its link to cancer. *Nucleic Acids Res* **39**: 781–794.

Zinzen RP, Girardot C, Gagneur J, Braun M, Furlong EE. (2009). Combinatorial binding predicts spatio-temporal cis-regulatory activity. *Nature* **462**: 65–70.

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