SUPPLEMENTARY INFORMATION

Figure S1: RT-PCR of *LINC00600* in LNCaP cells.

DNA gel after RT-PCR performed in LNCaP cells shows that the expression level *LINC00600* is undetectable by published RT-PCR primers (Gudmundsson et al. 2007).

Figure S2: The enrichment of H3K4me3 and H3K36me3 within *AR* gene

In PCa LNCaP and VCaP cells, the *AR* gene is enriched with H3K4me3 at the promoter region and H3K36me3 at the gene body.

Figure S3: H3K36me3 enrichment in 17q24.3 risk locus

Heatmap depicting H3K36me3 enrichment across 10 cell lines from the ENCODE project reveals that the 17q24.3 risk locus is deprived of transcription.

Figure S4: Enhancer chromatin signature of 17q24.3 risk locus

Heatmap depicting chromatin signatures across multiple cell lines: enhancer signature H3K4me1 from H1-hESC, HMEC, HSMM, HSMMtube, HUVEC, K562, NH-A, NHEK, NHLF, Osteobl, LNCaP and VCaP; enhancer signature H3K4me2, promoter signature H3K4me3, transcription signature H3K36me3 and repressive chromatin signature H3K9me3 from LNCaP and VCaP. Regulatory elements E1-E5 from Figure 1B-C are highlighted. PCa GWAS SNP rs1859962 resides in E1.

Figure S5: DNaseI hypersensitive sites of 17q24.3 risk locus

Heatmap depicting DNaseI hypersensitive sites across 77 distinct cell lines including PCa LNCaP cells (Highlighted by a blue box) reveals that E1 is truly a PCa specific enhancer. E2~E5 contains DNaseI hypersensitivity shared by multiple cell lines. PCa GWAS SNP rs1859962 resides in E1.

Figure S6: 3C-qPCR validating rs6983267-*MYC* interaction in LNCaP cells

Primers used in this validation are presented in Table S3.

Figure S7: Sequencing result of the 3C ligation product (between E1 and *SOX9*)

Figure S8: 3C-TaqMan assessing E1-SOX9 interaction using 3C data analysis pipeline from 'Hagege et al. 2007' (See the method section).

Figure S9: rs8072254 is in LD with rs1859962

Within the E1 enhancer, 3 SNPs (rs8071558, rs8072254 and rs984434) are included in HapMap dataset. All of the 3 are in LD with PCa risk SNP rs1859962. Note the other SNPs covered by our study are not included in the HapMap dataset.

Figure S10 and S11: DNA gel with regards to PCR amplicons of randomly ligated BAC templates.

Figure S12: Site-directed mutagenesis

Mutagenesis of each SNP was performed in the Luciferase reporter plasmids harboring the two DHS regions within the E1 enhancer. Cycle-sequencing results are presented for each SNP.

Figure S13: Luciferase assays using the E1 enhancer and the *SOX9* promoter

Luciferase assays using pGL3 plasmid the *SOX9* promoter (approximately ±500bp surrounding the *SOX9* TSS) and harboring the two DHS regions within the E1 enhancer and either the reference or variant sequence for the two functional SNPs within them. The pGL3 plasmid with the SOX9 promoter but without the enhancer region (Empty) is used as a negative control. The *y* axis represents the relative Luciferase units normalized to Renilla signal ± 1SEM. The Luciferase expression level for each variant SNP is compared to the plasmid homozygous for the reference alleles. P value is derived from a *t* test (**, $p \leq 0.01$).

Figure S14: *In vivo* ChIP-qPCR of FOXA1 and AP-1 in the E1 enhancer and the *SOX9* gene

The *y* axis indicates the ChIP enrichment normalized to the input DNA (ChIP without antibodies).

Figure S15: Luciferase assay for rs1859961 in MCF7 cells

Same as Figure 4F but done in MCF7 cells.

Table S1: LD analysis using genotype data from the 1000 genome project

Table S2: Primers used in this study

Table S3: Motif analysis for rs8072254 and rs1859961

References

Gudmundsson, J., Sulem, P., Steinthorsdottir, V., Bergthorsson, J.T., Thorleifsson, G., Manolescu, A., Rafnar, T., Gudbjartsson, D., Agnarsson, B.A., Baker, A. et al. 2007. Two variants on chromosome 17 confer prostate cancer risk, and the one in TCF2 protects against type 2 diabetes. *Nat Genet* **39**(8): 977-983.