Exploring the landscape of androgen receptor binding at tRNA genes in prostate cancer

Charlotte Ballantyne Garton

MSc by Research

University of York

Biology

December 2025

Abstract

Prostate cancer (PCa) presents a major global health burden, with 1.47 million new diagnoses and approximately 397,000 deaths each year. Dysregulated androgen receptor (AR) signalling is a primary driver of PCa progression, and consequently a key therapeutic target through androgen deprivation therapies and AR antagonists. Abnormal tRNA expression has been observed in different cancer types, including breast and ovarian, and there is evidence of estrogen and progesterone receptors binding tRNA genes. However, potential AR targeting of tRNA genes remains unexplored in the literature. Previous unpublished work (White group, University of York) found preliminary evidence of AR and H3K27ac enrichment (measured by ChIP-sequencing) at a subset of tRNA genes in the androgen-responsive and AR-expressing LNCaP cell line. The FOXA1 pioneer factor (a transcription factor capable of binding and opening heterochromatin), frequently co-localised with AR-binding, suggesting a possible interaction. However, depletion of the SWI/SNF chromatin remodeler significantly increased AR and H3K27ac signal intensity, indicating a potential tumour suppressive role. The current study built on these findings by first validating AR and FOXA1 binding at tRNA genes in tumour biopsies taken from patients with PCa, highlighting the biological relevance in disease. Next, the effect of different co-factors and chromatin remodelers were examined. FOXA1 overexpression enhanced existing AR recruitment to tRNA genes in addition to promoting de-novo binding events, consistent with its role as a pioneer factor. Analysis of local chromatin accessibility revealed that tRNA genes enriched for AR, FOXA1 and H3K27ac were highly accessible. Loss of the SWI/SNF and CHD1 chromatin remodelers resulted in increased accessibility and AR binding, respectively, suggesting tumour suppressive roles in AR recruitment to tRNA genes. Moreover, c-Myc frequently co-localised with AR at tRNA genes. Together, these findings strengthen evidence for AR binding at a subset of tRNA genes in PCa and highlight the relevance to human disease.

Keywords: androgen receptor, tRNA, prostate cancer, FOXA1, CHD1, SWI/SNF, c-Myc

Table of Contents

Abbreviations

ADT	Androgen deprivation therapy
AR	Androgen receptor
ARBS	AR Binding Site(s)
ARE	Androgen Response Element
AR-V7	Androgen Receptor splice variant 7
ATAC-seq	Assay for Transposase-Accessible Chromatin with sequencing
BAM	Binary Alignment Map
BDP1	B-double prime 1
BED	Browser Extensible Data
BRF1	B-related factor 1
ChIP-qPCR	Chromatin immunoprecipitation-quantitative polymerase chain reaction
ChIP-seq	Chromatin immunoprecipitation followed by sequencing
CHD1	Chromodomain Helicase DNA-binding protein 1
Co-IP	Co-immunoprecipitation
CRPCa	Castration-resistant prostate cancer
DBD	DNA-binding domain
DHT	Dihydrotestosterone
DMSO	Dimethyl sulfoxide
ENA	European Nucleotide Archive
EP400	E1A-binding protein p400
ER	Estrogen receptor
ERE	Estrogen response element
EZH2	Enhancer of zeste homolog 2
FDR	False Discovery Rate
FKH DBD	Forkhead DNA-binding domain
FOXA1	Forkhead Box A1
FOXA2	Forkhead Box A2
GFP	Green fluorescent protein
GR	Glucocorticoid receptor
H3K27ac	Histone H3 lysine 27 acetylation

H3K27me3	Histone H3 lysine 27 trimethylation
H3K4me3	Histone H3 lysine 4 trimethylation
HAT	Histone Acetyltransferase
HDAC	Histone Deacetylase
HER2	Human Epidermal Growth Factor Receptor 2
LBD	Ligand-binding domain
LHRH	Luteinizing hormone-releasing hormone
LNCaP	Lymph Node Carcinoma of the Prostate
mCRPCa	Metastatic castration resistant prostate cancer
MR	Mineralocorticoid receptor
mRNA	Messenger ribonucleic acid
NCBI	National Centre for Biotechnology Information
NLS	Nuclear localisation signal
nmt-tRNA	Nuclear-encoded mitochondrial tRNA
PCa	Prostate cancer
PCA	Principal component analysis
Pol II	Ribonucleic Acid Polymerase II
Pol III	Ribonucleic Acid Polymerase III
PR	Progesterone receptor
PROTAC	Proteolysis Targeting Chimera
PSA	Prostate specific antigen
Q-Value/Q	Depth normalised value quantified in EASeq to reads per million
q-value	False detection rate corrected p value
RNA	Ribonucleic acid
RPM	Reads Per Million
sno-miRNA	Small nucleolar RNA/microRNA
SRA	Sequence read archive
SWI/SNF	Switch/sucrose non-fermentable complex
TBP	TATA box binding protein
TFIIIB	Transcription factor IIIB
TFIIIC	Transcription factor IIIC
TIP60	Tat-interactive protein 60
tiRNA	tRNA-derived stress-induced RNA

TNBC	Triple-negative breast cancer
tRF	tRNA-derived fragment
tRNA	Transfer Ribonucleic Acid
tsRNA	tRNA-derived small RNA
TSS	Transcription Start Site
UCSC	University of California, Santa Cruz
WSL2	Windows Subsystem for Linux version 2

List of tables

List of figures

Acknowledgments

Firstly, I would like to give a special thank you to my supervisor, Professor Bob White, not only for his academic support and guidance, but also for providing encouragement throughout. I would also like to thank Sienna for her very helpful advice. Thank you also to my TAP member, Professor William Brackenbury for providing useful feedback. Thank you to Sandy MacDonald and Ben Durkan for bioinformatic and statistics guidance. I would like to thank Lewis, Lucy, Lola and Kora for providing emotional support and motivation. Finally, thank you mum and dad for supporting me all the way and believing in me.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for a degree or other qualification at this University or elsewhere. All sources are acknowledged as references

1.Introduction

1.1 Prostate cancer

Prostate cancer (PCa) presents a major global health burden. GLOBOCAN reported 1.47 million new diagnoses and 397,430 deaths in 2022, representing the second and fifth leading rates of cancer incidence and mortality in men worldwide, respectively[1]. Known risk factors include older age (≥ 65 years), African ethnicity and family history[2,3]. The majority of new cases are localised disease, with symptoms including urinary hesitancy, nocturia (waking up at night to pass urine) and haematuria (the presence of blood in the urine), although they can be asymptomatic. Treatment options vary by stage and grade, with active-surveillance for disease progression monitoring often preferred for early localised PCa, due to the slow-growing nature of these tumours. However, advanced localised and metastatic hormone sensitive PCa often benefit from therapies that hijack the androgen receptor (AR) signalling pathway, to slow excessive cancer cell growth and proliferation[4]. Early evidence for this emerged in 1941, where surgical castration, or orchidectomy, resulted in significant PCa tumour regression[5]. Currently, androgen deprivation therapy (ADT) is primarily achieved through hormonal intervention to lower serum androgen levels. While traditional medical castration agents such as luteinizing hormone-releasing hormone (LHRH) agonists block androgen synthesis in the testes, newer generation drugs target the adrenal glands and prostate[6]. Second-generation anti-androgens such as enzalutamide and apalutamide are commonly used in combination with ADT to antagonistically disrupt AR ligand binding. However, despite initial efficacy, approximately 20% of patients treated with hormone therapies develop the incurable castration- resistant PCa (CRPCa), an androgen insensitive cancer, driven by AR mutations and splice variants such as AR-V7[4].

1.2 Androgen receptor signalling

Androgens exert their influence via the AR signalling pathway, and are a critical part of normal male sexual development. The AR is a steroid hormone-sensitive transcription factor, expressed in the epithelial cells lining the prostate lumen[7]. Part of the nuclear receptor superfamily, the AR is classified into subgroup 3, which also includes the estrogen receptor (ER), progesterone receptor (PR), mineralocorticoid receptor (MR) and glucocorticoid receptor (GR)[8]. Located on the X chromosome, the AR gene encodes a 110 kD protein, with three primary domains and a flexible hinge (H) region (Figure 1A). The DNA-binding domain (DBD) recognises androgen-response element (ARE) motifs at enhancers and promoters of target genes. While testosterone can directly bind the

ligand-binding domain (LBD), dihydrotestosterone (DHT), which is converted from testosterone by 5-α reductase, has a much stronger binding affinity. In the absence of available ligands, the AR is located cytoplasmically in a complex with chaperone proteins such as heat-shock proteins. Upon ligand binding, the AR undergoes a conformational change, exposing the nuclear localisation signal (NLS), resulting in nuclear translocation and homodimerization[7,9] (Figure 1B).

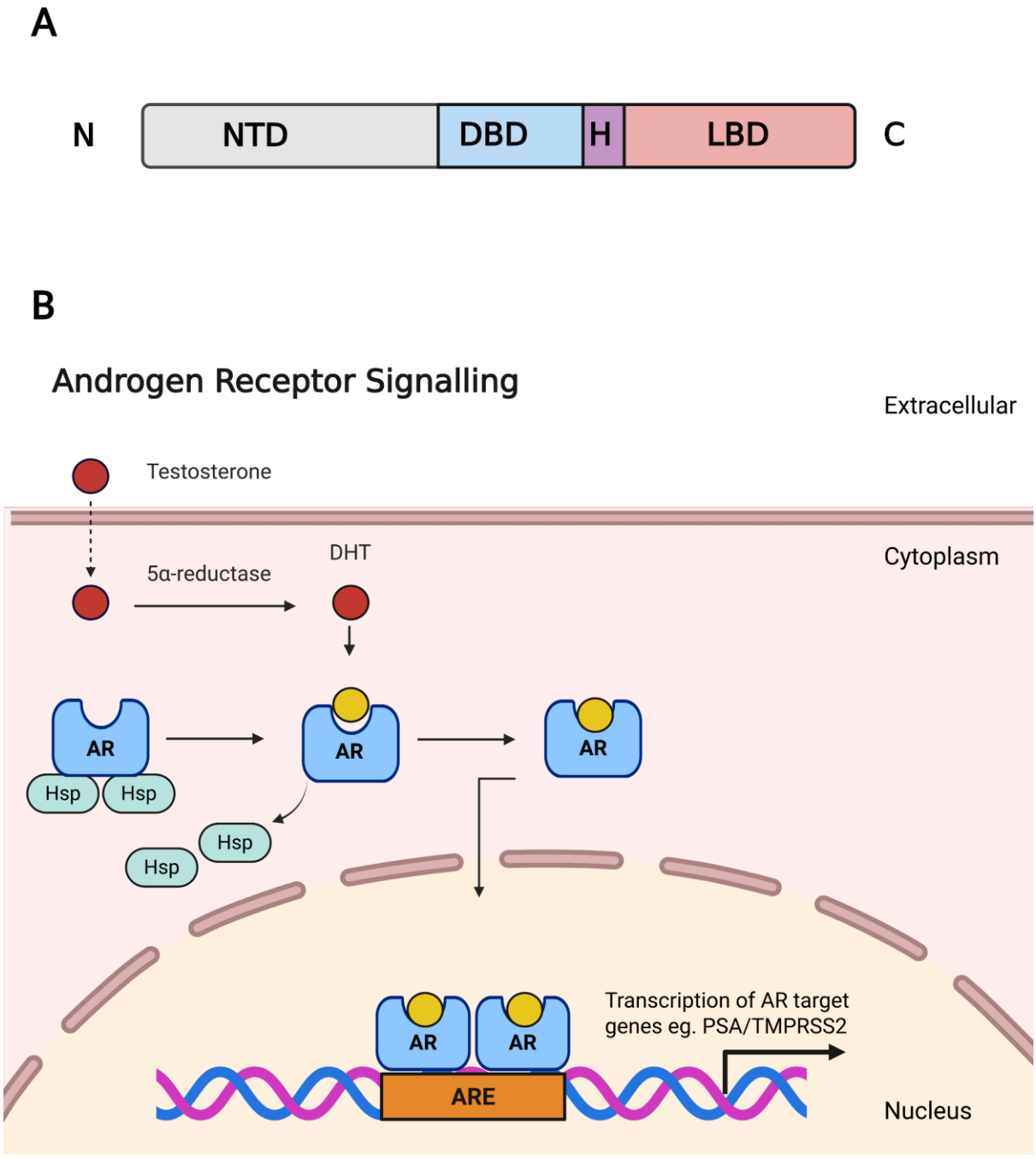


Figure 1. Schematic diagram of androgen receptor (AR) signalling in prostate cells. (A) The key functional domains in the androgen receptor protein structure: N-terminal domain (NTD), DNA-binding domain (DBD), hinge region (H) and ligand-binding domain (LBD). N = N-terminus, C = C-terminus. (B) The AR signalling cascade. Testosterone diffuses into the cytoplasm of prostate cells and is converted to dihydrotestosterone (DHT) by 5-α reductase. In the absence of an appropriate ligand such as DHT, the AR is anchored to the cytoskeletal elements, part of a complex of heat shock proteins (hsp). Upon ligand binding, the AR is released from its complex, and translocates to the nucleus. In the nucleus, the AR homodimerizes, binds to androgen response elements (AREs) and stimulates the transcription of target genes such as *PSA* (prostate-specific antigen) and *TMPRSS2* (transmembrane serine protease 2). Created with BioRender.com

1.3 AR coregulators

Although an open and accessible chromatin structure is necessary for AR to bind and stimulate transcription of target genes, euchromatin alone is insufficient to recruit AR. Several proteins have been designated as AR co-regulators, acting through mechanisms to modify the local chromatin state[10].

1.3.1 FOXA1

Forkhead box A1 (FOXA1) is a member of the FOX family of transcription factors. These proteins share a highly conserved winged forkhead (FKH) DBD, and are important regulators of gene expression. FOXA1 belongs to the FOXA subclass, alongside FOXA2 and FOXA3, characterised by their roles as pioneer factors, meaning they can bind compact heterochromatin via DNA motifs, and displace linker histones to open local chromatin structure[11]. The FOXA1 monomer is predominately recruited to enhancers enriched with mono- and di-methylation of H3K4 (H3K4me/me2)[12]. Here, FOXA1 recognises and binds discrete regions of condensed chromatin that contain the consensus motif, displacing the H1 linker histone and recruiting factors such as histone acetyltransferases (HATs) to open the local chromatin[13].

FOXA1 is a direct binding partner of AR, essential for driving normal prostate epithelial differentiation in male development[14,15]. Studies have shown that FOXA1, in addition to other pioneer factors GATA2 and HOXB13, actively recruits and stabilizes binding of AR to preferential AR binding sites (ARBS), promoting tumorigenesis[10]. However, evidence of FOXA1 inhibiting AR binding in PCa has also been observed, reflecting a complex and context-dependent relationship[16].

The *FOXA1* gene is among the most frequently mutated in PCa at approximately 11%[17]. Recurrent mutational hotspots across all stages of disease progression have been identified in the FKHD domain, representing a substantial proportion of all FOXA1 mutations (~55%). Specifically, these were missense and indel mutations, and were largely found to be clustered in the wing2 region, conferring an increase in pioneering activity by altering DNA-binding affinities[18].

1.3.2 SWI/SNF

The dynamic process of chromatin remodeling is important in regulating gene expression. Heterochromatin, where DNA is tightly wrapped around histone octamers (nucleosomes) prevents access by DNA-binding factors. However, the open euchromatin structure enables access to regions of DNA, permitting binding of transcription machinery. ATP-dependent chromatin remodelers

actively manipulate chromatin architecture, by repositioning nucleosomes at non-coding regulatory regions, thus modulating gene transcription[19,20].

One such chromatin remodeler is the switch/sucrose non-fermentable (SWI/SNF), or BRG1/BRM1-associated factor (BAF) complex. SWI/SNF is a large (~2 MDa) multi subunit complex, with three mammalian isoforms: canonical BAF (cBAF), non-canonical BAF (ncBAF) and polybromo-associated BAF (pBAF)[21]. 29 genes encode the subunits, with 16 common to all three SWI/SNF isoforms[22]. All complexes include either the SMARCA2 or SMARCA4 subunits, responsible for the enzymatic nucleosome remodeling activity[23].

Genes encoding SWI/SNF subunits are commonly mutated (~20%) across many cancer types. Emerging evidence in recent years has gained SWI/SNF a reputation as a tumour suppressor. Many studies show loss-of-function mutations in various subunits such as SMARCA4, ARID1A and SMARCD3[24].

Various SWI/SNF subunits have been identified as AR cofactors or to frequently co-localise at shared sites, including SMARCD1, SMARCA4 and SMARC1[25–27]. Gene alterations, predominately amplifications and upregulation of mRNA transcripts, of either *SMARCD1*, *SMARCD2* and *SMARCD3* are found in 6% of primary PCa and 5-11% of metastatic PCa cases, respectively[28]. Upregulation of SMARCC1 was found in PCa, positively correlating with clinical signs of disease progression such as Gleason score and time to recurrence[29]. In addition, SMARCA4 upregulation was associated with increased tumour mass[29,30]. Furthermore, SMARCAD proteins were found to both promote and antagonise AR-signalling[28].

1.3.3 CHD1

Chromodomain helicase DNA-binding (CHD) proteins are another class of ATPase dependent chromatin remodelers that have been implicated in PCa. The family consists of nine members, all containing two distinct functional domains. First, two chromodomains situated towards the N-terminal, which recognise tri-methylated H3K4 (H3K4me3), a mark associated with active gene transcription. Second, the SNF2-like ATPase domain, in the central region, responsible for nucleosome sliding through enzymatic activity. The CHD family members are further divided into three subfamilies (I-III), with CHD1 and CHD2, belonging to subfamily I. Both proteins contain the SANT-SLIDE DBD, which binds flanking DNA either side of the nucleosome to direct nucleosome movement[31,32].

Like SWI/SNF, *CHD* genes are altered in many cancer types. In PCa, *CHD1* gene deletions are found in approximately 8% of cases, and are mutually exclusive with inactivation of *PTEN* (an established

tumour suppressor in PCa)[33,34]. Consequently, CHD1 has gained recognition as a putative tumour suppressor in PCa[32]. Despite CHD1 not physically associating with the AR itself, many studies have demonstrated loss of CHD1 to drive prostate tumorigenesis through chromatin dysregulation, lineage plasticity and epigenetic reprogramming. Interestingly, one study found CHD1 to localise at enhancers containing ARE motifs. Loss of CHD1 resulted in a redistribution of AR binding, to sites enriched for HOXB13[34]. HOXB13, along with FOXA1 have been shown to co-localise at reprogrammed ARBS (AR binding sites gained or lost in PCa compared to the canonical AR binding pattern in normal prostate cells)[35].

1.4 tRNA genes

1.4.1 Transcriptional regulation

The human genome (GRCh38) contains over 600 annotated nuclear-encoded tRNA genes, of which 428 have been predicted with high-confidence[36]. In addition, there are 22 mitochondrial tRNA genes, however nuclear tRNA genes are the focus of this thesis. They are short, non-coding genes. For each of the 20 standard amino acids, tRNAs with different anticodons exist, such as tRNA-Ala-AGC-1-1 and tRNA-Ala-CGC-1-1. Genes with the same anticodon with different nucleotide sequences are termed isodecoders such as tRNA-Gly-GCC-1-1 and tRNA-Gly-GCC-2-1. There are also different copies of the same genes, such as tRNA-Gly-GCC-1-1 and tRNA-Gly-GCC-1-2. Exact copy number variations vary between individuals; however, a canonical set is annotated in the human reference genome[37–39].

RNA polymerase III (Pol III) is the largest of the three RNA polymerases, responsible for transcription of non-coding genes, including tRNA, 5S rRNA, 7SL RNA and U6 snRNA[40]. Pol III target genes can be categorised into three classes, defined by their structure and recruitment of basal and accessory transcriptional machinery. Type I and II genes are characterised by internal promoter elements, whereas type III genes contain upstream regulatory elements[41,42]. tRNA genes belong to the type II class, with internal promoter sequences termed the A and B boxes, which are recognised and bound by transcription factor IIIC (TFIIIC). TFIIIC directly interacts with and recruits transcription factor IIIB (TFIIIB), upstream of the transcription start site (TSS). TFIIIB comprises three subunits: TATA box binding protein (TBP), B-double prime 1 (BDP1) and, in the case of type I and II promoters, B-related factor 1 (BRF1). TFIIIB is responsible for recruiting Pol III to the TSS to initiate transcription[40,43].

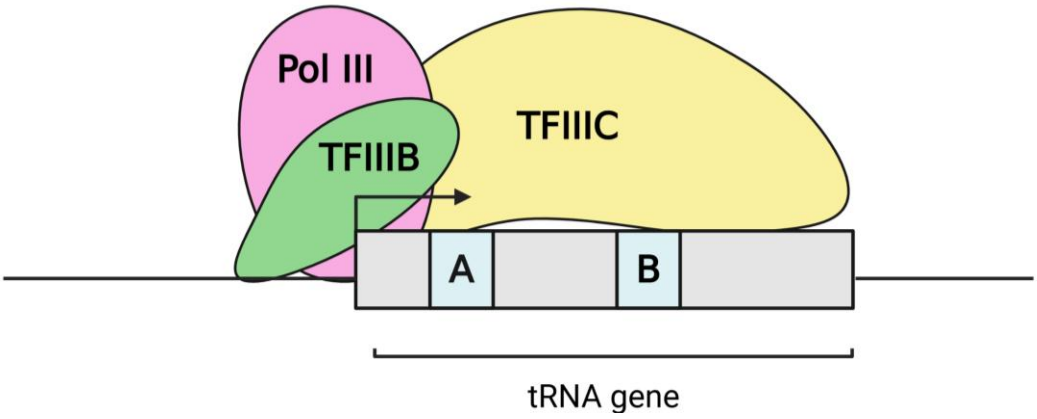


Figure 2. RNA polymerase III (Pol III) transcription machinery at tRNA genes. The assembly of the transcription machinery begins when transcription factor IIIC (TFIIIC) recognises and binds to the internal promoter elements of tRNA genes (A and B boxes). TFIIIC recruits transcription factor IIIB (TFIIIB) upstream of the transcription start site (black arrow), which in turn recruits Pol III. Created with BioRender.com

Synthesis of tRNAs requires a high energy expenditure by the cell, therefore transcription of tRNA genes is tightly regulated in normal cells. Mechanistically, tumour suppressors such as p53, RB and ARF inhibit Pol III-mediated transcription by binding TFIIIB, preventing interactions with Pol III or TFIIIC[44].

Conversely, Myc and ERK directly bind TFIIIB, stimulating transcription. In 2003, c-Myc was discovered to stimulate transcription of tRNA genes, by directly associating with TFIIIB[45]. Further work investigating the mechanism behind Pol III enhancement found c-Myc recruits the histone acetyltransferase (HAT) GCN5 and cofactor TRRAP[46].

1.4.2 Upregulation of tRNA gene expression in human cancers

Historically, cancer researchers neglected studying tRNAs, and Pol III transcription more broadly, in favour of Pol II. Their function was regarded as limited to cellular house-keeping, rather than relevant to oncogenic transformation. Uncontrolled cell growth and proliferation is a key characteristic of cancer cells. Sufficient protein production enables cells to reach a critical mass, prior to cell proliferation. Significant elevation of tRNAs in human cancers were first observed in ovarian cancer, with subsequent studies confirming these findings in breast and multiple myeloma[47–49]. Tumour samples from nine patients with luminal, basal or HER2 enriched breast cancers exhibited an average 10-fold higher tRNA expression, relative to healthy breast tissue[48]. A subsequent case-control study of 104 women with luminal A/B or triple negative breast cancers (TNBC), validated these findings using next-generation sequencing. They reported a global upregulation of 76 tRNAs, of which three were associated with reduced overall, but not disease-free, survival, after adjustment for age at diagnosis, tumour stage/grade and TNBC status[50]. Another study analysed miRNA-seq data from

TCGA cohorts spanning 31 distinct cancer types and approximately 10,000 patient samples, finding consistent tRNA overexpression[51].

Various underlying mechanisms have been shown to cause transcriptional dysregulation, such as the observed TFIIIC overexpression in ovarian tumour biopsies.[47].

Non-coding RNAs such as microRNAs (miRNAs) and long noncoding RNAs (lncRNAs) have been implicated in tumorigenesis. Interest in tRNA-derived fragments (tRFs) as potential PCa biomarkers has emerged in recent years, after the initial discovery of tRFs in PCa cell lines in 2009[52]. Abnormal expression of many (tRFs) have been found in both PCa cell lines and patient tissue compared to control, including those derived from tRNA-Glu, tRNA-Lys and tRNA-Gly[53–55]. Although whether tRNA gene transcription directly affects the production of tRFs is unknown, the differential expression profiles between tumour and normal states are worth noting.

In addition, dysregulated amino acid metabolism is commonly observed in many cancer types, including PCa. Increased dependency on amino acids such as glutamine, methionine and glycine in PCa enable altered metabolism that facilitates excessive cell growth and proliferation[56].

1.4.3 Association of the estrogen and progesterone receptors at tRNA genes in breast cancer

In recent years, tRNA genes have been identified as targets of the ER and PR steroid receptors in breast cancer (BCa). Early evidence of a direct estrogen-mediated response on tRNA gene transcription came from a rapid transcription within a 10-minute period of estradiol induction in MCF-7 cells. More than 90% of the total tRNA gene pool were upregulated[57]. More recently, ER-α was found to bind tRNA genes in the same cell line, in addition to BCa patient tumours[58].

Progesterone targeting of tRNA genes is less-well studied. One study found PR binding events at 40% of tRNA gene loci in BCa patient xenografts. The authors also discovered PR associated with subunits of the Pol III complex[59].

Despite growing interest in tRNA biology in PCa, and evidence of other nuclear receptors binding tRNA genes, whether tRNA genes are a target of the AR has yet to be explored.

1.5 Summary of previous work

To begin addressing whether AR can bind tRNA genes, ChIP-sequencing (chromatin immunoprecipitation followed by sequencing) data from the LNCaP cell line (an androgen responsive

cell line), was interrogated [60,61]. AR was found to preferentially bind a subset of tRNA genes, compared to sno-miRNA genes. FOXA1 binding co-localised with the majority of AR binding events. Transcriptional activity was inferred by H3K27ac enrichment, revealing a strong correlation with AR and FOXA1 signal intensity. Together, this suggested that FOXA1 may facilitate AR recruitment to tRNA genes, although this was a tentative conclusion as AR binding in response to FOXA1 modulation had not been studied in this context.

In addition, the effect of SWI/SNF depletion on AR, FOXA1 and H3K27ac signal intensity was evaluated. AR and H3K27ac Q values (depth normalised signal quantified ± 500 bp from the centre of tRNA genes) were found to be significantly increased upon targeted SWI/SNF inactivation compared to the control. This was indicative of a tumour suppressive role of SWI/SNF on AR recruitment to tRNA genes. FOXA1 signal was not globally significantly altered between conditions. One limitation of this investigation was that chromatin state was not analysed, which would enable a complimentary understanding of these effects.

These preliminary findings were of interest, demonstrating the need for validation and further exploration into the landscape of AR binding at tRNA genes in prostate cancer.

1.6 Project aims and objectives

This project aims to build upon the work carried out previously, by validating AR binding events in patient biopsy tissues, and investigating binding of AR co-factors at tRNA genes.

To achieve this, the following objectives were set:

- To validate AR, FOXA1 and H3K27ac enrichment at tRNA genes in PCa patient biopsies, prior to carrying out further bioinformatic analysis of AR binding events in the LNCaP cell line
- To investigate the relationship between AR and FOXA1 binding at tRNA genes by examining how FOXA1 overexpression affects AR signal intensity
- To evaluate the effect of SWI/SNF depletion on chromatin accessibility at tRNA genes
- To examine AR, c-Myc and CHD1 binding events at tRNA genes and infer transcriptional activity by histone marks
- To evaluate the effect of CHD1 loss on AR binding at tRNA genes

2. Materials and Methods

2.1 Data acquisition and processing

2.1.1 Sequencing data

Publicly accessible ChIP-seq and ATAC-seq datasets were obtained from the National Centre for Biotechnology Information (NCBI) sequence read archive (SRA), and the European Bioinformatics Institute (EMBL-EBI) European Nucleotide Archive (ENA) (Tables 1 and 2). Within the Galaxy web-based platform (v.24.2.3) (https://usegalaxy.org/), SRA and FASTQ files were downloaded via 'Faster Download and Extract Reads in FASTQ format from NCBI SRA' (v.3.1.1), and fastq-dl (v.3.0.1), respectively[62–65]. Prior to processing, unmapped read quality was assessed (FastQC v.0.74) and aggregated into reports using MultiQC (v.1.27)[66,67]. Low-quality bases (Phred score <20) were discarded and adapter sequences were trimmed using Trim Galore! (v.0.6.10)[68]. Reads were mapped to the hg38 (GRCh38) human reference genome (BowTie2 v.2.5.3) and resultant BAM files were downloaded[69].

Subsequent data processing was performed locally within the Windows Subsystem for Linux version 2 (WSL2) environment, running Ubuntu (v.22.04). Bioinformatic tools and dependencies were managed through Conda (v.25.3.1). Alignment statistics were checked using SAMtools (v.1.21) flagstat (Supplementary Table 1). Using SAMtools sort, BAM files derived from single and paired end reads were coordinate sorted and name sorted, respectively. Mate coordinates were annotated using SAMtools fixmate, prior to sorting by coordinates. PCR and optical duplicates were marked using SAMtools markdup. A final filtering step was applied to BAM files with SAMtools view, removing unmapped reads, mates and duplicates, and verified by SAMtools stats[70]. Due to the repetitive nature of tRNA genes, multi-mapped reads were deliberately retained to preserve signal enrichment at these loci.

2.1.2 tRNA and sno-miRNA genes

Tracked hg38 tRNA gene (n = 631) and sno-miRNA gene (n = 2320) tables were downloaded from the UCSC table browser as .txt files and copied into Excel for filtering[71]. Genes not present in a high confidence list of hg38 tRNA genes were removed, including pseudogenes and nuclear-encoded mitochondrial tRNAs (nmt-tRNAs), leaving 409 high confidence tRNA genes[36].

Table 1A. Publicly available ChIP-sequencing data downloaded from NCBI SRA

Source	Accession	Antibody	Condition	Cell type
GSE114737[72]	SRR7207011	AR	Patient 1	Biopsy
	SRR7207017	AR	Patient 2	
	SRR7207023	AR	Patient 3	
	SRR7207030	AR	Patient 4	
	SRR7207013	FOXA1	Patient 1	
	SRR7207019	FOXA1	Patient 2	
	SRR7207026	FOXA1	Patient 3	
	SRR7207032	FOXA1	Patient 4	
	SRR7207015	H3K27ac	Patient 1	
	SRR7207021	H3K27ac	Patient 2	
	SRR7207028	H3K27ac	Patient 3	
	SRR7207034	H3K27ac	Patient 4	
GSE171589[61]	SRR14159840	AR	Control	LNCaP
	SRR14159842	FOXA1	Control	
	SRR14159846	H3K27ac	Control	
GSE117430[34]	SRR7545370	AR	Control	
	SRR7545371	AR	Control	
	SRR7545372	CHD1	Control	
	SRR7545373	CHD1	Control	
	SRR7545374	c-Myc	Control	
	SRR7545375	c-Myc	Control	
	SRR7545376	H3K4me3	Control	
	SRR7545377	H3K4me3	Control	
	SRR7545378	H3K27me3	Control	
	SRR7545379	H3K27me3	Control	
	SRR7545380	AR	sgCHD1	
	SRR7545381	AR	sgCHD1	
	SRR7545382	AR	sgCtrl	
	SRR7545383	AR	sgCtrl	

sg = short guide RNA (sgRNA)-mediated CRISPR-Cas9 knockout.

Table 1B. Publicly available ChIP-sequencing dataset downloaded from EMBL EBI

Source	Accession	Antibody	Condition	Cell type
PRJEB4235[73]	ERR307344	-	Input DNA	LNCaP
	ERR307350	AR	Control	
	ERR307345	AR	Control	
	ERR307347	AR	FOXA1 overexpression	
	ERR307352	AR	FOXA1 overexpression	

Table 2. Publicly available ATAC-sequencing dataset downloaded from NCBI SRA

Source	Accession	Condition	Cell type
GSE171584[61]	SRR14156915	Control (sgNC + sgNC)	LNCaP
	SRR14156916	Control (sgNC + sgNC)	
	SRR14156917	sgSMARCA2	
	SRR14156918	sgSMARCA2	
	SRR14156919	sgSMARCA4	
	SRR14156920	sgSMARCA4	
	SRR14156921	sgSMARCA2 + shSMARCA4	
	SRR14156922	sgSMARCA2 + shSMARCA4	

sg = short guide RNA (sgRNA)-mediated CRISPR-Cas9 knockout; sh = short hairpin RNA (shRNA)-mediated RNA interference knockdown.

2.2 Quantification of signal intensity and gene visualisation in EaSeq

2.2.1 Signal quantification (Q values)

Filtered BAM files were imported into EaSeq (v.1.12) as 'Datasets', using default import settings. Although, the option 'only allow reads with unique positions' was deselected to ensure retention of multi-mapped reads[74]. Replicates were merged using the 'pool with' function. The hg38 reference genome was downloaded as a 'Geneset', and sno-miRNA and tRNA gene tables were imported as 'Regionsets'.

Using the 'Quantify' tool, ATAC-seq and ChIP-seq signal intensities were quantified ± 500 bp from the centre of each tRNA gene. Default parameters were applied, except for the deselection of 'normalize signal to a size of 1000bp'. Counts were normalised to DNA fragments and reads normalised to reads per million (RPM). Resultant quantified and depth-normalised or 'Q values' were downloaded as .txt files and copied into Excel for downstream analysis in R and Excel.

2.2.2 Heatmaps

Heatmaps were visualised using the 'HeatMap' plot function within EaSeq. Signal enrichment was viewed across a 20,000 bp window, with tRNA and sno-miRNA genes aligned at the centre and sorted according to increasing signal intensity (Q value).

2.2.3 Gene enrichment tracks

As positive controls, AR, FOXA1 and H3K27ac enrichment at the *PSA* gene (*KLK3*), and c-Myc enrichment at the *CDK4* gene were examined. The 'LineTrack' function in EaSeq was used to plot signal across a 10,000-20,000 bp window from the start of the gene and smoothing to 5 bins. Averaged signal intensities at tRNA genes or sno-miRNA genes were plotted in EaSeq with the 'Average' tool $\pm$ 10,000 bp. Tracks were superimposed using the 'Overlay' tool. Individual tRNA genes of interest were isolated from the tRNA Geneset by name using the 'Gate' tool. Enrichment patterns were interrogated using the 'FillTrack' plot function, setting window sizes between 4000 – 10,000 bp from the centre of the tRNA gene, and smoothing to 5 bins.

2.3 Data visualisation and statistical analysis in Microsoft Excel and R Studio

All statistical analyses were carried out within R Studio v4.4.3 [75,76]. Boxplots, bar plots and scatter graphs were created using ggplot2 (v.3.5.2) as part of the tidyverse package (v.2.0.0). Where normality assumptions were not met, non-parametric tests were applied instead. Linear regression lines were calculated using the 'lm' model, with the purpose to aid visualisation of overall trends rather than to model linearity where Spearman's rank correlation coefficients were calculated. Paired gene-level students t-tests were used to calculate Q value changes between conditions. Statistical significance was indicated by: * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$. All values were rounded to two decimal places, unless otherwise stated.

Where input DNA was available and of sufficient quality, statistical comparisons were repeated on input-normalised Q values and documented in the supplementary material (Supplementary Figure 1). This was to validate the use of sequencing depth-normalised Q values in the main text, where background signal had not been deducted. ChIP-seq Q values were divided by input, with a pseudocount of +1 and log transformed, as shown in the equation below:

$$\log_2\left(\frac{Sample\ Q\ value + 1}{Input\ Q\ value + 1}\right)$$

The relative signal intensity heatmap was created in Excel, utilising the conditional formatting function and normalizing Q values for each transcription factor or histone modification.

2.4 MACS2 peak calling

Where DNA input sequencing data was available, peak calling was performed in Galaxy using MACS2 callpeak (v.2.2.9.1)[77]. Default settings were applied, using narrow for transcription factors, FDR q < 0.05 and effective genome size of 2.7e9 (*H.sapiens*). Consensus peaks in NarrowPeak format between biological replicates were generated using bedtools Intersect intervals v2.31.1. The option 'Write the original A entry_once_if_any_overlaps found in B' (corresponding to the command line -u flag) was selected to count each tRNA gene once. To extract tRNA genes from each consensus list, a tRNA gene list was exported from the UCSC table browser in BED format using the 'send output to Galaxy' tool, and intersected with the same parameters as above to generate consensus tRNA gene lists[78]. Finally, Venn diagrams were created in PowerPoint using the consensus tRNA gene lists to identify shared and unique peaks between conditions.

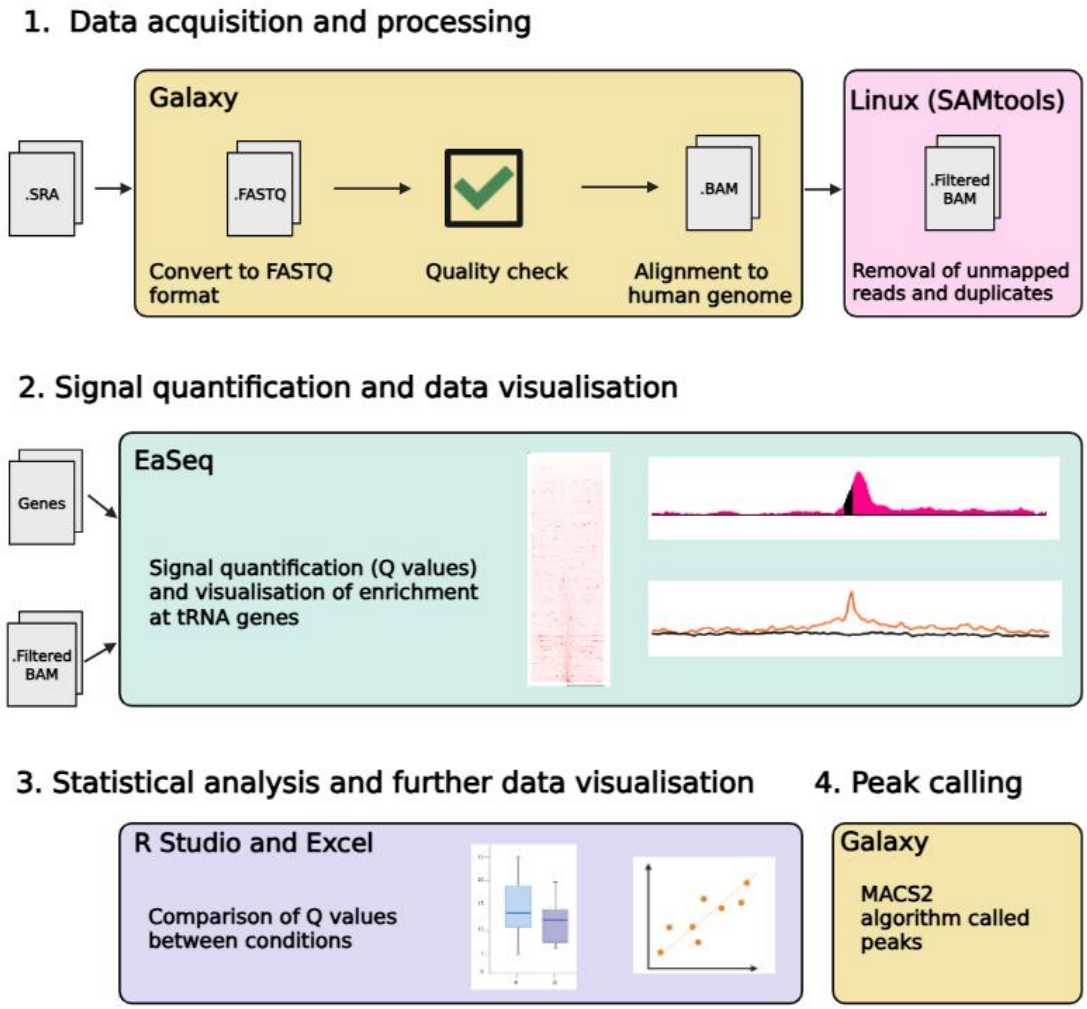


Figure 3. Schematic overview of the bioinformatic pipeline used to analyse ChIP- and ATAC-sequencing datasets. Created with BioRender.com

3.Results

3.1 Validation of AR binding events, FOXA1 co-localisation and H3K27ac enrichment at tRNA genes in prostate cancer patient biopsies

Previous unpublished findings from the White group (University of York) showed AR enrichment at a subset of tRNA genes in the LNCaP cell line. Recruitment of the FOXA1 pioneer factor was observed at a substantial proportion of these AR bound loci, and transcriptional activity was inferred by H3K27ac enrichment. Prior to further work, AR recruitment to tRNA genes were validated in human tissue. This was important to confirm that such binding events were not artefacts of the LNCaP cell line and held biological relevance. To achieve this, AR, FOXA1 and H3K27ac enrichment at tRNA genes were examined in fresh frozen tissue biopsies from four prostate cancer patients. Importantly, samples were selected from patients where neoadjuvant hormone therapy was omitted prior to surgery, thus ensuring retention of physiologic androgen levels[72].

First, the *KLK3* target gene was examined as a positive control for AR binding. As expected, the pooled input DNA control showed no AR enrichment. In contrast, clear peaks of AR signal were identified approximately 4000 bp upstream of the *KLK3* transcription start site, at the distal enhancer in patients 1, 2 and 4. Peak strength was comparable between patients 1 and 4, although substantially stronger for patient 2. An additional smaller peak was observed at the *KLK3* gene promoter in patient 2. However, AR enrichment was absent in patient 3, thus producing an unsatisfactory positive control for AR binding and consequently excluded from further analysis (Figure 4A).

Next, AR enrichment at tRNA genes (± 500 bp) was compared to snoRNA and miRNA (sno-miRNA) genes (± 500 bp) to provide a negative control. Average signal intensity plots revealed enrichment of AR at tRNA genes (n = 409) but minimal association at sno-miRNA genes (n = 2230), demonstrating specificity in targeting tRNA genes compared to other small non-coding RNA genes (Figure 4B). Mean Q values at tRNA genes were 0.76 (patient 1), 0.78 (patient 2), 0.47 (patient 3) and 0.72 (patient 4). Heatmaps revealed AR occupancy localised to a subset of tRNA genes, relative to surrounding DNA ± 10 kb from the centre of each locus (Figure 4C). Minimal association of AR was observed at sno-miRNA genes (Supplementary Figure 2A). All heatmaps were sorted by increasing signal intensity (bottom to top) in patient 1. This showed considerable overlap, suggesting a conservation of AR binding occupancy across different patients.

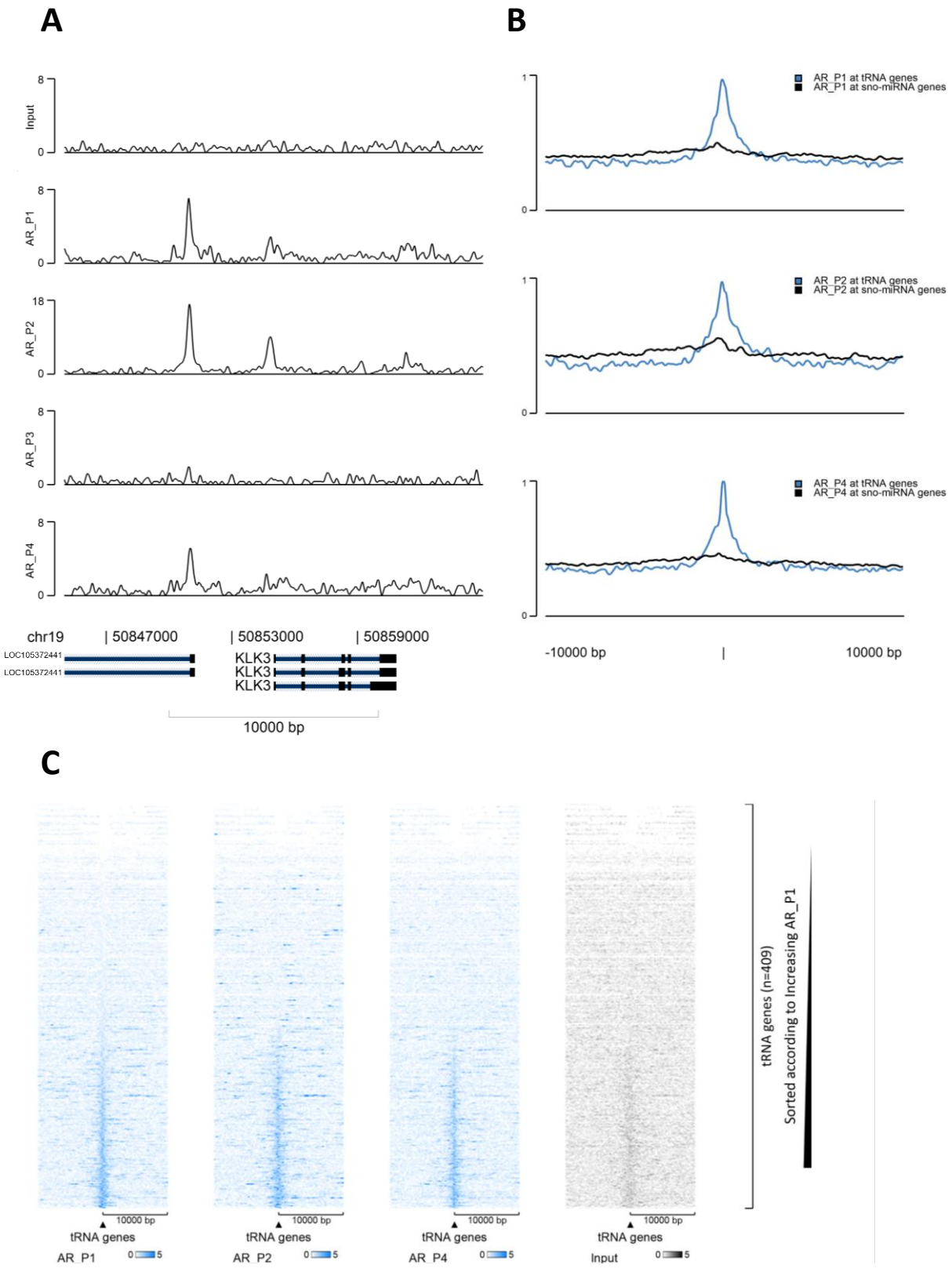


Figure 4. Validation of androgen receptor binding events at tRNA genes in individual prostate cancer patient biopsies. (**A**) Positive control for AR binding at the KLK3 promoter and upstream enhancer in the input control and patients 1-4 (**B**) Averaged signal intensities of AR at tRNA genes (blue) compared to sno-miRNA genes as a negative control (black), across patients 1, 2 and 4. (**C**) Heatmap of AR binding, sorted by increasing signal intensity in patient 1, at tRNA genes in patients 1, 2 and 4 (blue) and input control (grey).

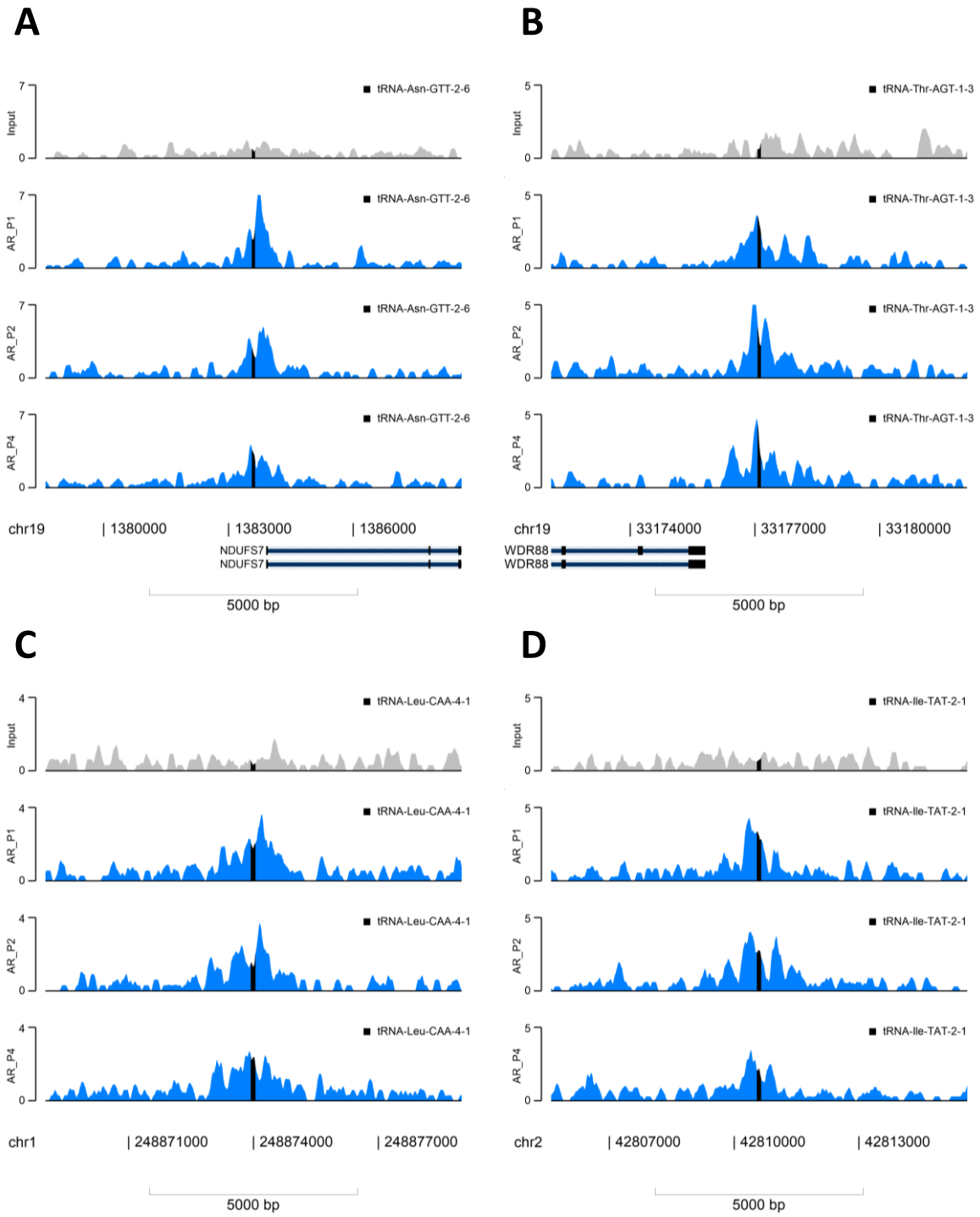


Figure 5. Representative tracks of AR binding events at tRNA genes ± 2500 bp from the centre, in patient biopsy samples (blue) compared to background input (grey). (**A**) tRNA-Asn-GTT-2-6. (**B**) tRNA-Thr-AGT-1-3. (**C**) tRNA-Leu-CAA-4-1 (**D**) tRNA Ile-TAT-2-1.

Binding patterns at individual tRNA genes were explored across patients 1, 2 and 4, relative to input control. Many examples with peaks at or near tRNA genes were found such as tRNA-Asn-GTT-2-6, tRNA-Ile-TAT-2-1, tRNA-Leu-CAA-4-1 and tRNA-Thr-AGT-1-3 (Figure 5).

In summary, the data shows that AR binding occurs at a subset of tRNA genes in patient data, suggesting biological relevance and confirming that these events are not artefacts of the LNCaP cell line.

In previous analysis, FOXA1 signal overlapped with the majority of AR-bound tRNA genes. Here, enrichment of FOXA1 at tRNA genes in the patient biopsy dataset was examined and compared to that of AR.

FOXA1 peaks at the *KLK3* gene enhancer were established in all four patients (Figure 6A). Average signal plots showed FOXA1 enrichment at tRNA genes relative to sno-miRNA genes (Figure 6B). Although present in all patient samples, mean FOXA1 signal was lower in patient 3 (Q = 0.55) compared to the others (patient 1 Q = 0.77; patient 2 Q = 0.91; patient 4 Q = 0.83). Mean Q values at sno-miRNA genes were 0.12 (patient 1), 0.18 (patient 2), 0.19 (patient 3) and 0.21 (patient 4). FOXA1 occupancy at tRNA genes relative to ± 10 kb flanking DNA upstream and downstream from the centre of each locus was sorted by increasing signal in patient 1. Signal was detected at approximately 40% of tRNA genes, again with a similar distribution of bound genes across the patient samples. In contrast, minimal signal was observed at sno-miRNA genes (Supplementary Figure 2B). However, FOXA1 signal in patient 3 was weaker than the other patients. Interestingly, signal intensity appeared visibly similar compared to input, suggesting that true peaks were not present (Figure 6C).

Examples of binding events at or near tRNA genes were observed, such as tRNA-Pro-CGG-1-1, tRNA-Thr-CGT-4-1, tRNA-Ile-TAT-2-1 and tRNA-Ala-TGC-4-1. However, as expected, at these loci, peaks were not present in patient 3 and showed similar tracks to input (grey).

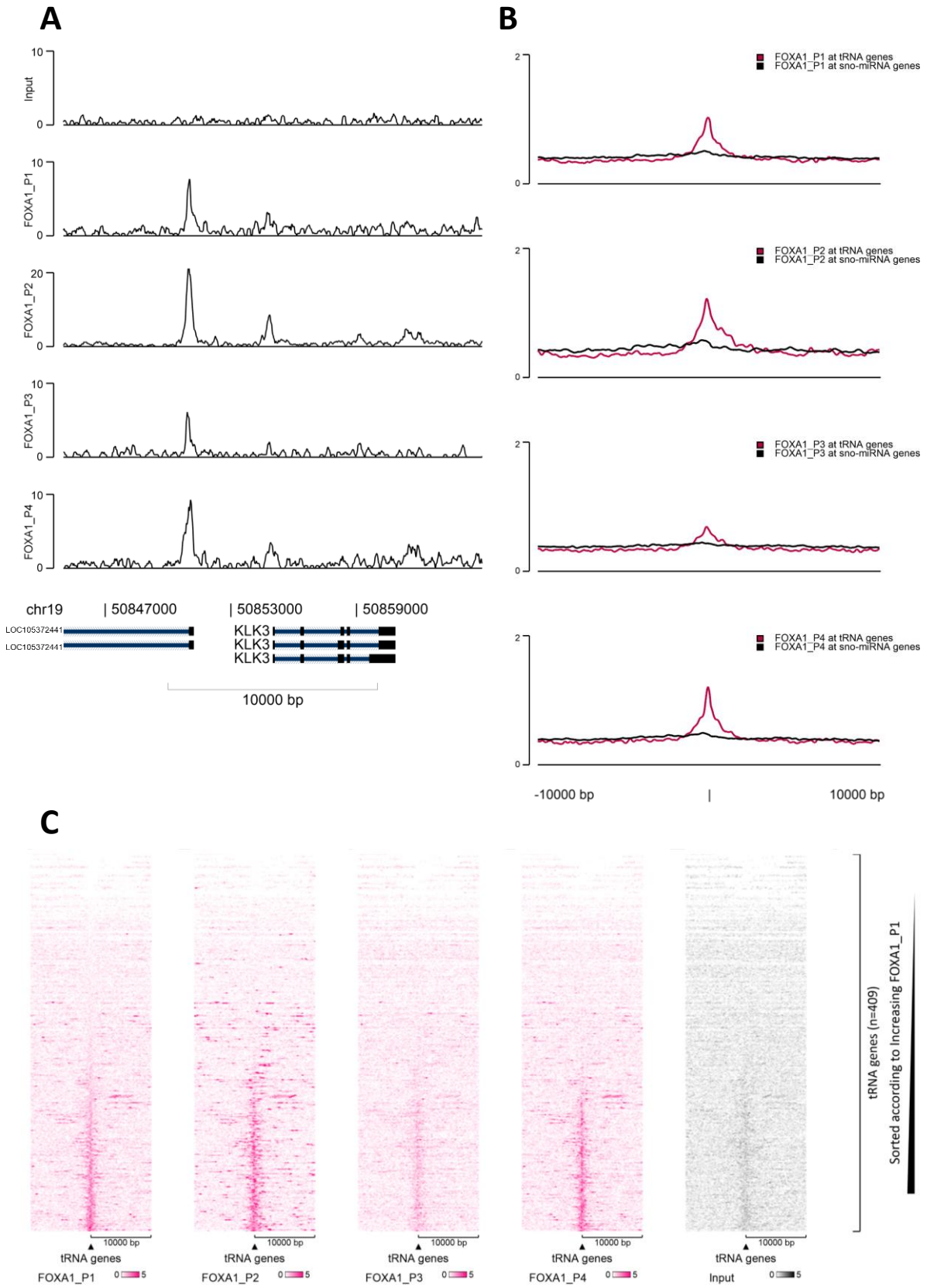


Figure 6. FOXA1 binding events at tRNA genes in prostate cancer patient biopsies. (**A**) Positive control for FOXA1 binding at the *KLK3* promoter and upstream enhancer. (**B**) Average FOXA1 signal intensities at tRNA genes (pink) compared to sno-miRNA genes (black) (**C**) Heatmap of FOXA1 binding, sorted by increasing signal intensity in patient 1, at tRNA genes in patients 1-4 (pink) and input control (grey).

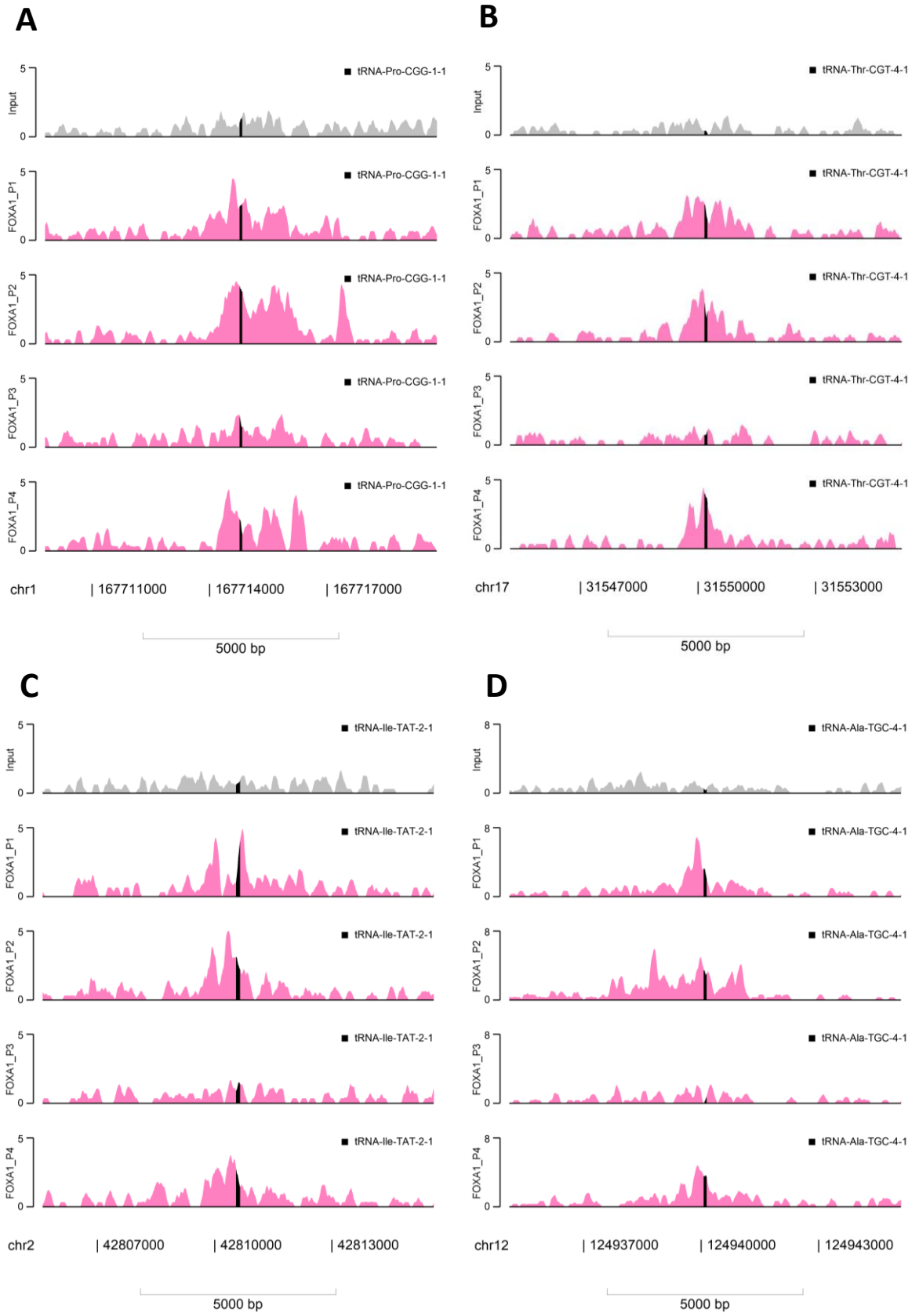


Figure 7. Representative tracks of FOXA1 binding events at tRNA genes ± 2500 bp from the centre, in patient biopsy samples (pink) compared to background input (grey). (**A**) tRNA-Pro-CGG-1-1 (**B**) tRNA-Thr-CGT-4-1 (**C**) tRNA-Ile-TAT-2-1 (**D**) tRNA-Ala-TGC-4-1.

AR and FOXA1 enrichment exhibited a very strong positive correlation in all cases (Patient 1: rho = 0.93, S = 813666, $p < 0.01$; Patient 2: rho = 0.92, S = 898542, $p < 0.01$; Patient 4: rho = 0.92, S = 884280, $p < 0.01$). This co-localisation of AR and FOXA1 in many cases may suggest that FOXA1 facilitates AR binding at tRNA genes (Figure 8).

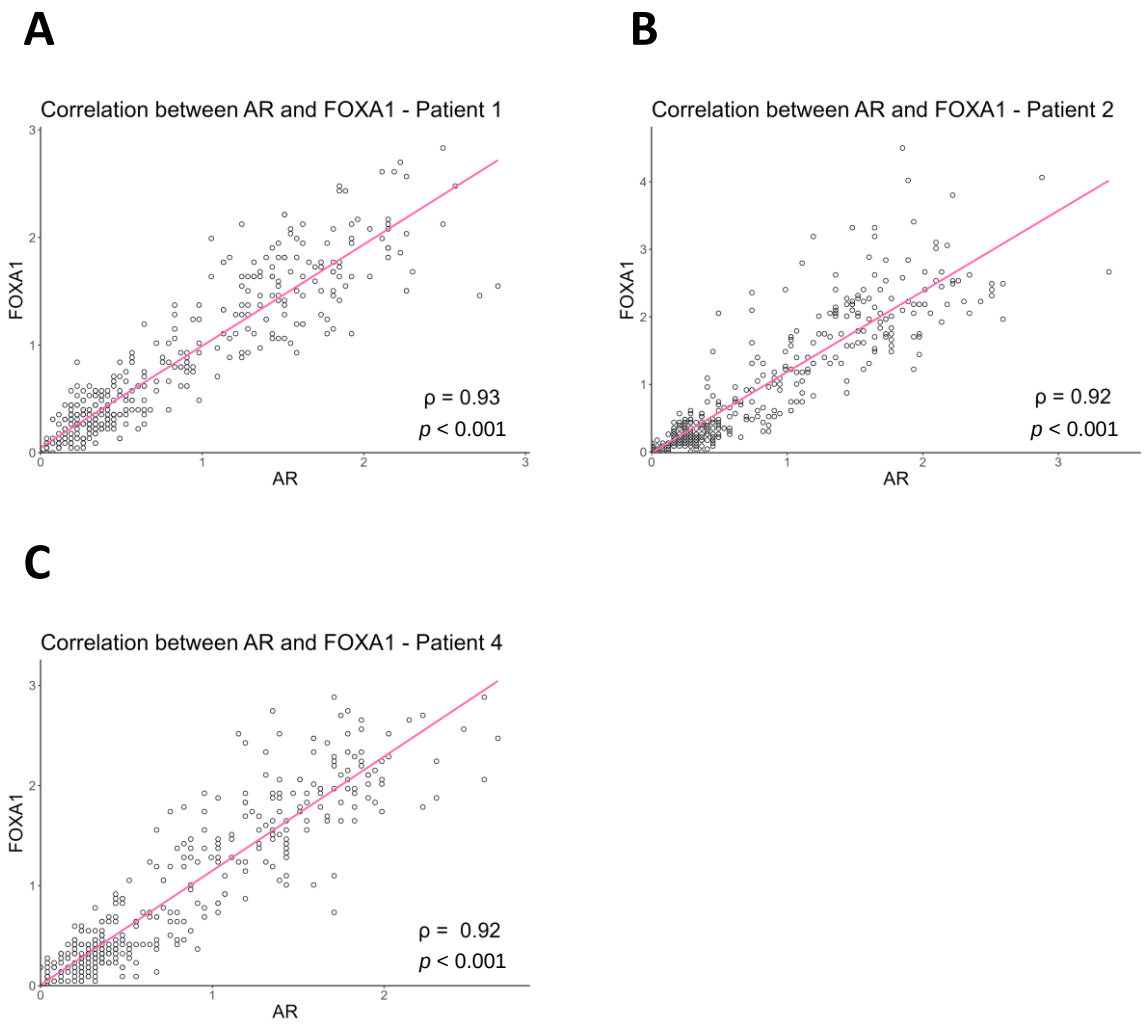


Figure 8. Spearman's rank correlation of AR and FOXA1 signal intensities at tRNA genes in patient biopsies. Linear relationship line displayed for visualisation purposes. (**A**) Patient 1 (**B**) Patient 2 (**C**) Patient 4.

In the absence of ChIP-seq data with antibodies against Pol III, H3K27ac enrichment at tRNA genes was used as a substitute measure for active gene transcription. Previously, H3K27ac was found to positively correlate with AR and FOXA1 signal, suggesting AR and FOXA1 bound tRNA genes were transcriptionally active. Here, H3K27ac signal was confirmed at the *KLK3* promoter, relative to input (Figure 9A). Average signal intensity plots revealed enrichment at tRNA genes, relative to sno-miRNA genes (Figure 9B). As tRNA genes are nucleosome depleted, this signature profile of peaks either side

of these loci are well-documented and described as peak-valley-peak patterns. Mean Q values at tRNA genes were 1.78 (patient 1), 2.04 (patient 2), 1.61 (patient 3) and 1.31 (patient 4). Mean Q values at sno-miRNA genes were 0.93 (patient 1), 1.08 (patient 2), 0.93 (patient 3) and 0.80 (patient 4). Next, heatmaps sorted by increasing H3K27ac signal in patient 1 were displayed, showing strong enrichment at just under half of tRNA genes, relative to adjacent DNA ± 10 kb, which was not observed at sno-miRNA genes (Figure 9C; Supplementary Figure 2C).

After strong H3K27ac signals were detected at a subset of tRNA genes, correlation analyses were performed with AR and FOXA1.

In all cases, very strong positive correlations were observed between AR and H3K27ac (Patient 1: rho = 0.91, S = 970137, $p < 0.001$; Patient 2: rho = 0.91, S = 1071131, $p < 0.001$; Patient 4: rho = 0.92, S = 951132, $p < 0.001$), suggestive of active transcription at AR bound tRNA genes.

Similar strong correlations were found between FOXA1 and H3K27ac in all samples (Patient 1: rho = 0.91, S = 988139, $p < 0.001$; Patient 2: rho = 0.90, S = 1141827, $p < 0.001$; Patient 3: rho = 0.88, S = 1387108, $p < 0.001$; Patient 4: rho = 0.91, S = 970415, $p < 0.001$) (Figures 10A-G). These suggest FOXA1 bound tRNA genes are largely transcriptionally active.

Taken together, these data suggest that the previously identified AR and FOXA1 binding events at tRNA genes in LNCaP cell lines are not artefacts, and occur in real patient tissue. In addition, these binding events are strongly correlated with H3K27ac, a mark of actively transcribed genes, as shown in a representative filltrack at tRNA-Ile-TAT-2-1 (Figure 10H).

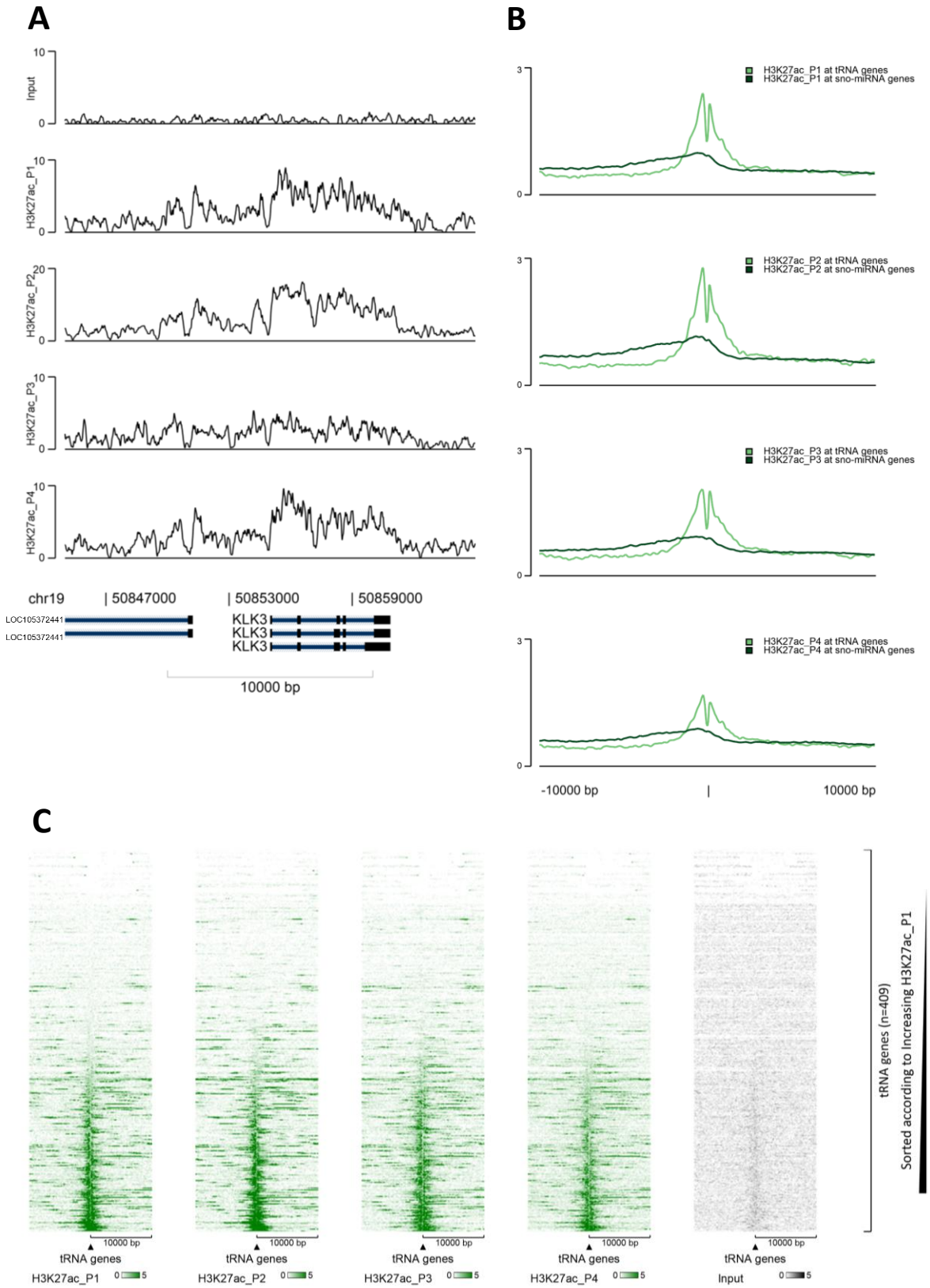


Figure 9. H3K27ac signal at tRNA genes in individual prostate cancer patient biopsies. (**A**) Positive control for H3K27ac signal at the KLK3 promoter and upstream enhancer, in the input control and biopsy data. (**B**) Averaged signal intensities of H3K27ac at tRNA genes (green) compared to sno-miRNA genes as a negative control (black), across patients 1- 4. (**C**) Heatmap of H3K27ac, sorted by increasing signal intensity in patient 1, at tRNA genes in patients 1-4 (green) and input control (grey).

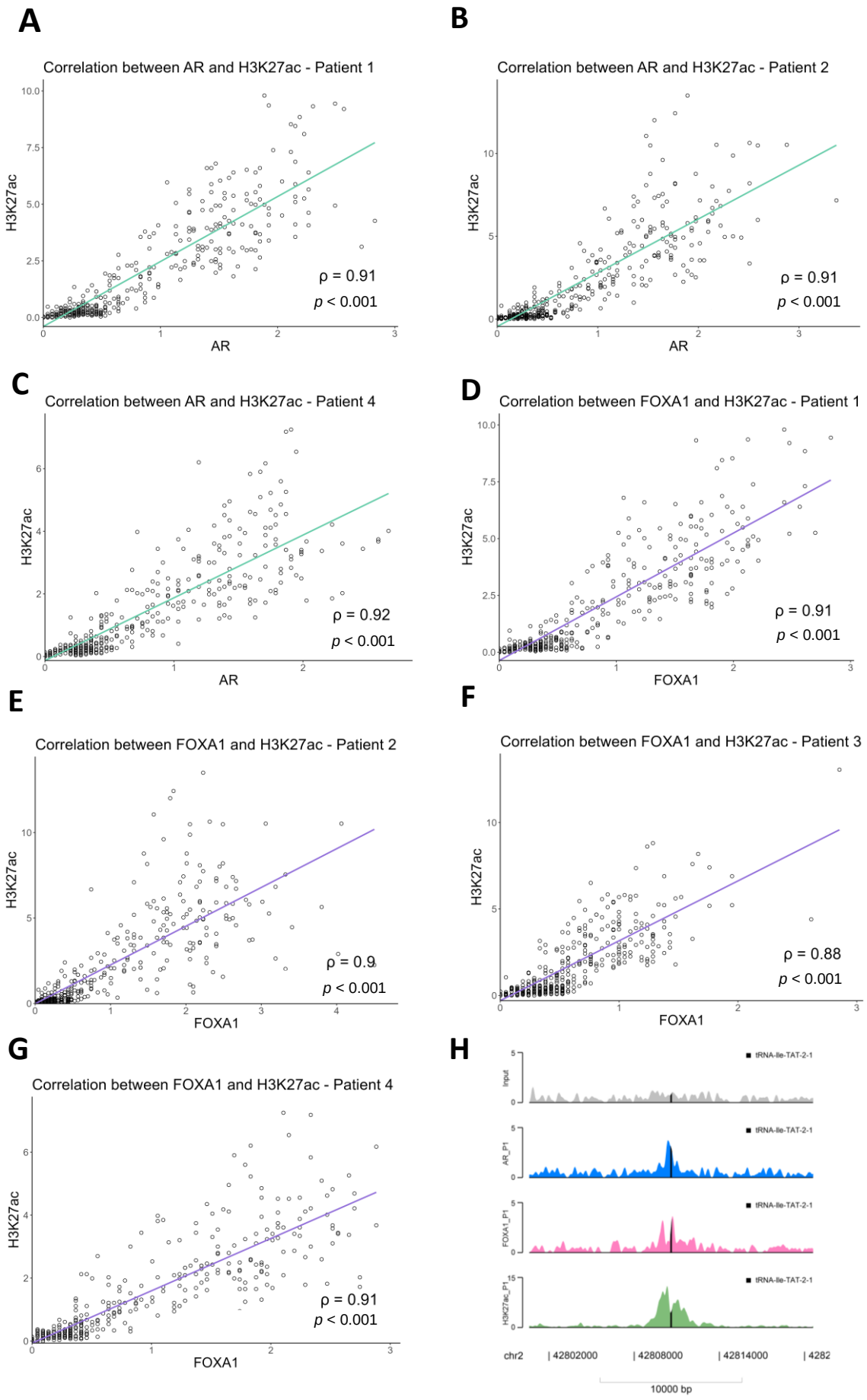

A
Correlation between AR and H3K27ac - Patient 1
H3K27ac
AR
ρ = 0.91
p < 0.001
B
Correlation between AR and H3K27ac - Patient 2
H3K27ac
AR
ρ = 0.91
p < 0.001
C
Correlation between AR and H3K27ac - Patient 4
H3K27ac
AR
ρ = 0.92
p < 0.001
D
Correlation between FOXA1 and H3K27ac - Patient 1
H3K27ac
FOXA1
ρ = 0.91
p < 0.001
E
Correlation between FOXA1 and H3K27ac - Patient 2
H3K27ac
FOXA1
ρ = 0.9
p < 0.001
F
Correlation between FOXA1 and H3K27ac - Patient 3
H3K27ac
FOXA1
ρ = 0.88
p < 0.001
G
Correlation between FOXA1 and H3K27ac - Patient 4
H3K27ac
FOXA1
ρ = 0.91
p < 0.001
H
tRNA-Ile-TAT-2-1
Input
AR_P1
FOXA1_P1
H3K27ac_P1
chr2
42802000
42808000
42814000
4282
10000 bp

Figure 10. Relationships between H3K27ac, AR and FOXA1 in patient biopsies. (**A**-**C**) Correlations between AR and H3K27ac. (**D**-**G**) Correlations between FOXA1 and H3K27ac. (**H**) Representative filltrack at tRNA-Ile-TAT-2-1. Linear relationship line displayed for visualisation purpose.

3.2 FOXA1 overexpression enhances AR recruitment to tRNA genes

FOXA1 is an established direct binding partner of AR, normally acting to facilitate recruitment to target genes. In previous unpublished work within the White group (University of York), AR and FOXA1 were found to co-localise at a substantial proportion of tRNA genes in the LNCaP cell line. Since AR and FOXA1 have been confirmed to bind tRNA genes in human tissue, the relationship between them were explored.

To evaluate this, AR binding events in response to FOXA1 overexpression in the LNCaP cell line were examined. In this dataset, cells were transfected with plasmids carrying full-length FOXA1 cDNA or GFP vector controls. First, detection of clear peaks at both the *KLK3* gene enhancer and promoter positively confirmed AR binding events (Figure 11A). Average signal plots showed AR enrichment at tRNA genes relative to sno-miRNA genes, demonstrating specificity of AR occupancy at tRNA genes (Figure 11B). Next, overlaying averaged signal intensities at tRNA genes showed an increase in the mean signal after FOXA1 overexpression (Q = 1.32), compared to control (Q = 0.95) (Figure 11C).

Heatmaps sorted by increasing AR signal in the control condition showed further enrichment at existing tRNA gene targets, after FOXA1 overexpression (Figure 11D). Minimal association was detected at sno-miRNA genes (Supplementary Figure 3A). Next, individual tRNA genes were inspected to observe different AR binding patterns. Many tracks showed either enhanced AR recruitment to existing loci, or the formation of entirely new peaks after FOXA1 overexpression (de-novo). Representative filltracks of such binding patterns can be observed in tRNA-Asn-GTT-2-2, tRNA-Ser-AGA-2-2, tRNA-iMet-CAT-1-6 and tRNA-Thr-CGT-1-1 (Figure 11E-H). However, there were cases where only minimal changes were detected, such as tRNA-Val-CAC-1-4 (Supplementary Figure 3B).

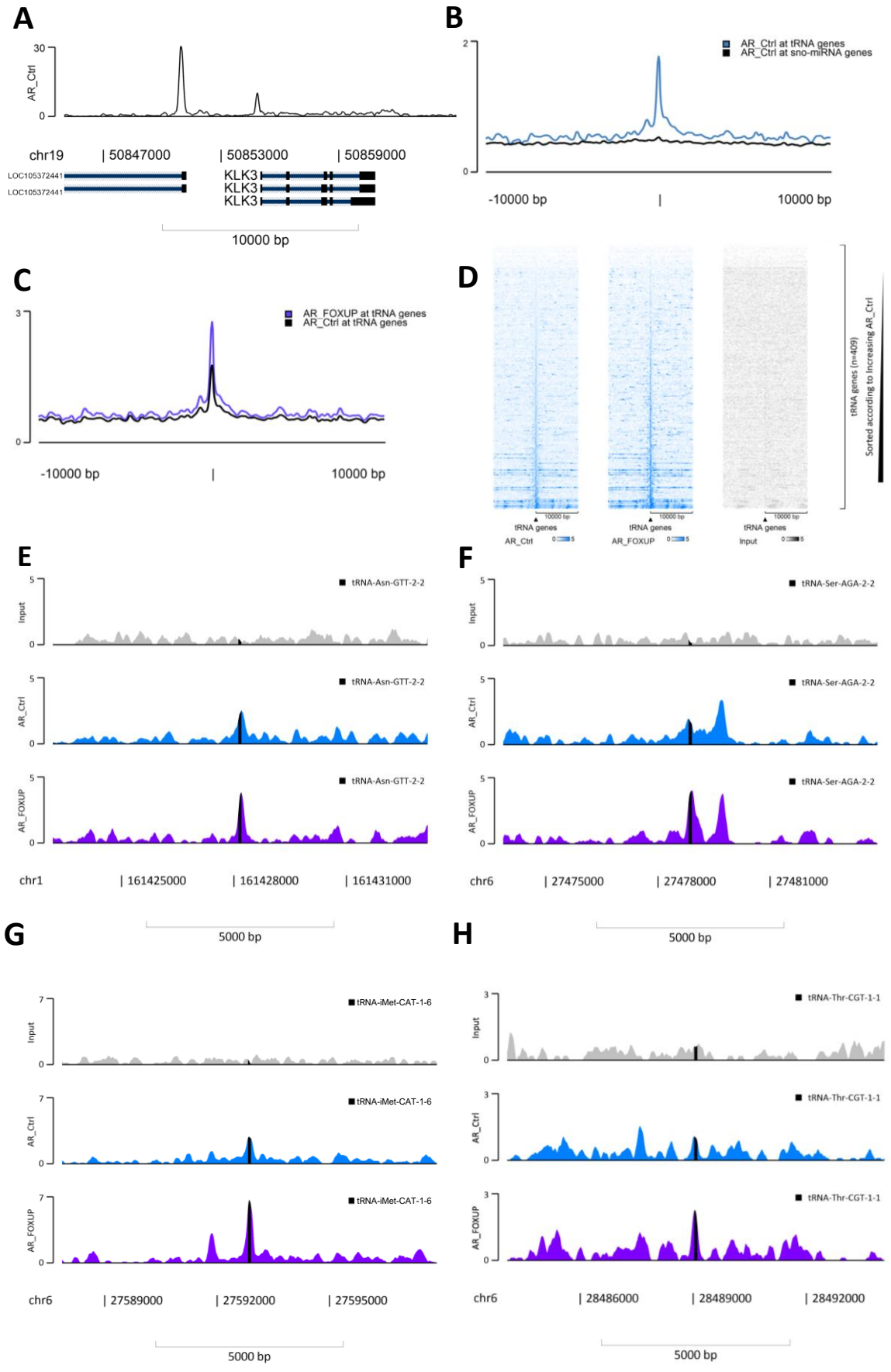

A
AR_Ctrl
30
0
chr19
| 50847000
| 50853000
| 50859000
LOC105372441
LOC105372441
KLK3
KLK3
KLK3
10000 bp
B
AR_Ctrl at tRNA genes
AR_Ctrl at sno-miRNA genes
2
0
-10000 bp
|
10000 bp
C
AR_FOXUP at tRNA genes
AR_Ctrl at tRNA genes
3
0
-10000 bp
|
10000 bp
D
tRNA genes
AR_Ctrl
tRNA genes
AR_FOXUP
tRNA genes
Input
tRNA genes (n=409)
Sorted according to Increasing AR_Ctrl
E
Input
tRNA-Asn-GTT-2-2
AR_Ctrl
tRNA-Asn-GTT-2-2
AR_FOXUP
tRNA-Asn-GTT-2-2
chr1
| 161425000
| 161428000
| 161431000
5000 bp
F
Input
tRNA-Ser-AGA-2-2
AR_Ctrl
tRNA-Ser-AGA-2-2
AR_FOXUP
tRNA-Ser-AGA-2-2
chr6
| 27475000
| 27478000
| 27481000
5000 bp
G
Input
tRNA-iMet-CAT-1-6
AR_Ctrl
tRNA-iMet-CAT-1-6
AR_FOXUP
tRNA-iMet-CAT-1-6
chr6
| 27589000
| 27592000
| 27595000
5000 bp
H
Input
tRNA-Thr-CGT-1-1
AR_Ctrl
tRNA-Thr-CGT-1-1
AR_FOXUP
tRNA-Thr-CGT-1-1
chr6
| 28486000
| 28489000
| 28492000
5000 bp

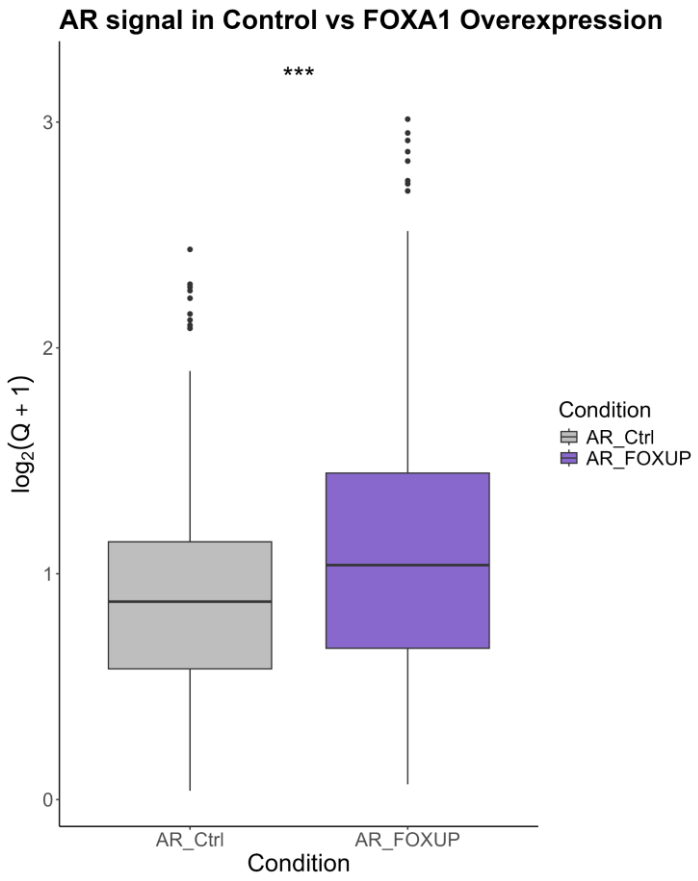
Figure 11. The effect of FOXA1 overexpression on AR binding events at tRNA genes in the LNCaP cell line. (**A**) Positive control for AR binding at the KLK3 promoter and upstream enhancer (**B**) AR binding at tRNA genes and negative control sno-miRNA genes. (**C**) AR binding at tRNA genes in control (black) and FOXA1 overexpression (purple). (**D**) AR enrichment heatmaps at tRNA genes in Control (left), FOXA1 overexpression (middle) and input (right) sorted by increasing signal intensity of AR in Control (**E**-**H**) Representative fill tracks of AR binding in control and FOXA1 overexpression in tRNA-Asn-GTT-2-2 (**E**), tRNA-Ser-AGA-2-2 (**F**), tRNA-iMet-CAT-1-6 (**G**) and tRNA-Thr-CGT-1-1 (**H**).

To test whether these observations were statistically significant, paired gene-level changes to Q values were evaluated. Across all tRNA genes, a significant median paired difference of 0.20 was found between FOXA1 overexpression and control, corresponding to a ~1.15-fold increase (Wilcoxon signed rank test, V = 79441, $p < 0.001$) (Figure 12A).

Next, to investigate these differences more closely, genes were stratified into three groups of low, medium and high, based on the 33rd and 66th percentiles in the control Q values. Paired gene-level tests were performed within each group, and the Benjamini-Hochberg correction was applied to control the false discovery rate (FDR). In all three groups, there was a statistically significant increase in the FOXA1 overexpression group (Low: median difference = 0.09, p-value = 1.14e-10; Medium: median difference = 0.19, p = 1.62e-19; High: median difference = 0.30, p = 3.00e-24) (Figure 12B). This may suggest that FOXA1 overexpression is capable of enhancing existing AR recruitment in addition to opening up chromatin to facilitate new AR binding events.

Further interrogation of this dataset revealed individual tRNA genes with the largest percentage increases in Q values after FOXA1 overexpression (Table 3).

A


AR signal in Control vs FOXA1 Overexpression
***
log2(Q + 1)
0
1
2
3
AR_Ctrl
AR_FOXUP
Condition
Condition
AR_Ctrl
AR_FOXUP

B

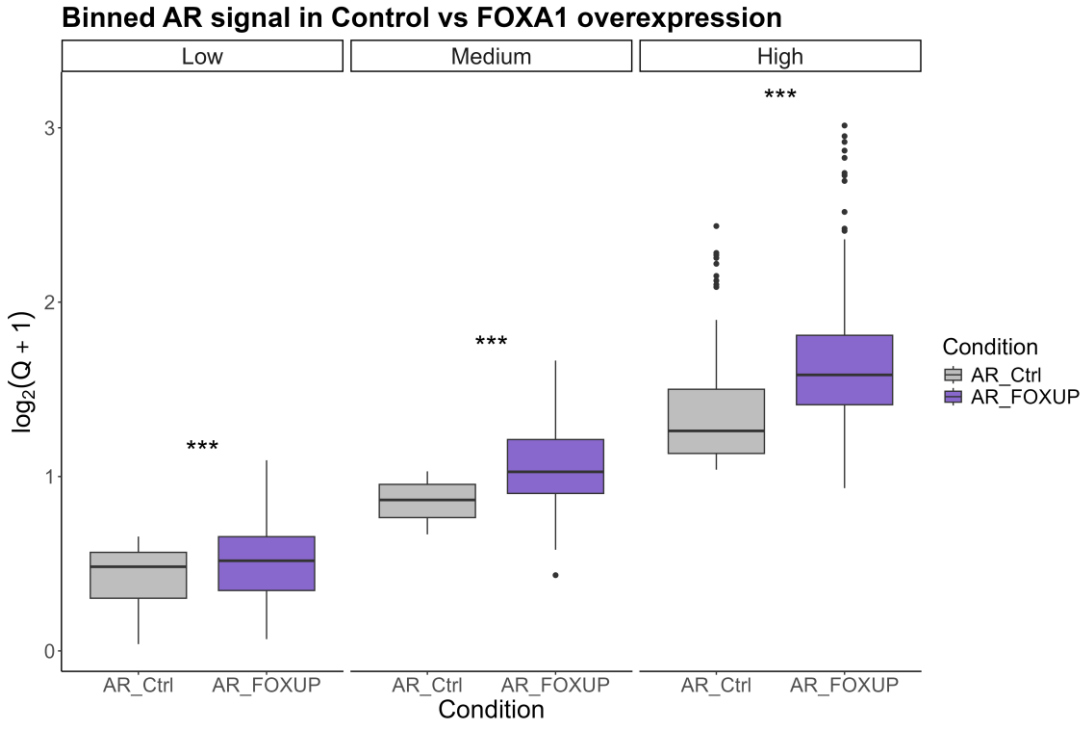

Binned AR signal in Control vs FOXA1 overexpression
Low
Medium
High
***
***
***
log2(Q + 1)
0
1
2
3
AR_Ctrl
AR_FOXUP
AR_Ctrl
AR_FOXUP
AR_Ctrl
AR_FOXUP
Condition
Condition
AR_Ctrl
AR_FOXUP

Figure 12. Comparison of Q values in the control and FOXA1 overexpression condition, in the LNCaP cell line. (**A**) Global overview of AR Q values in the control (grey) and FOXA1 overexpression dataset (purple). (**B**) AR signal intensity stratified into 3 groups based on low, medium and high control Q values, and corresponding values in FOXA1 overexpression dataset.

Table 3. Top 30 genes with the highest percentage increase between 'Control' and 'FOXA1 overexpression' Q values

Gene name	Control	FOXA1 overexpression	Percentage increase between conditions (%)
tRNA-Cys-GCA-17-1	0.34	1.07	212
tRNA-Phe-GAA-6-1	0.04	0.13	211
tRNA-Arg-CCG-1-2	0.07	0.19	180
tRNA-Val-TAC-1-2	0.41	0.93	125
tRNA-Pro-AGG-1-1	0.64	1.42	121
tRNA-Arg-CCG-1-3	1.00	2.17	117
tRNA-iMet-CAT-1-3	1.01	2.14	111
tRNA-Thr-CGT-4-1	1.18	2.47	110
tRNA-Leu-AAG-2-2	0.18	0.37	106
tRNA-Asn-GTT-7-1	0.84	1.71	104
tRNA-Phe-GAA-1-2	0.16	0.34	104
tRNA-Asn-GTT-2-1	0.97	1.98	103
tRNA-Ile-TAT-2-3	0.58	1.13	97
tRNA-Met-CAT-6-1	0.47	0.91	95
tRNA-Leu-AAG-1-1	1.01	1.98	95
tRNA-Lys-CTT-4-1	0.85	1.64	94
tRNA-Tyr-GTA-5-3	0.81	1.56	94
tRNA-iMet-CAT-1-4	0.48	0.93	93
tRNA-Arg-ACG-1-1	0.67	1.29	93
tRNA-Asn-GTT-3-1	1.06	2.03	92
tRNA-Ser-CGA-3-1	0.82	1.56	90
tRNA-Asp-GTC-2-4	1.51	2.86	90
tRNA-Glu-TTC-1-2	2.27	4.31	89
tRNA-Ala-TGC-4-1	1.43	2.70	89
tRNA-Gln-CTG-1-3	0.11	0.21	89
tRNA-Leu-TAA-4-1	0.69	1.29	89
tRNA-Ala-AGC-8-2	1.26	2.38	89
tRNA-Gly-TCC-2-5	3.36	6.31	88
tRNA-Leu-AAG-2-4	1.22	2.28	87
tRNA-Tyr-GTA-5-2	1.10	2.04	86

Finally, MACS2 peak calling was performed to identify unique and shared peaks between conditions, distinct from background signal. FOXA1 overexpression resulted in 109 new tRNA genes (>25% of the total tRNA gene pool) displaying peaks, that were absent in the control group (Figure 13). 78 tRNA genes were shared between the groups, and taken together with the heatmap distribution visualisation, this would suggest that at least a portion of these are existing binding events that increase under FOXA1 overexpression. 9 tRNA genes are unique to the control.

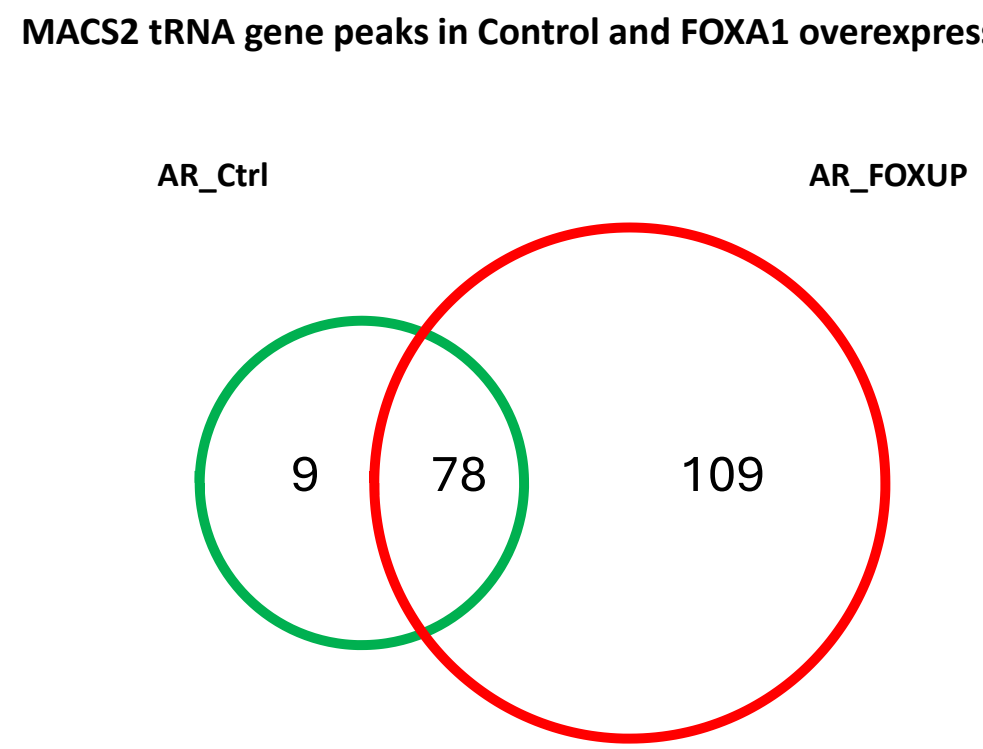


Figure 13. Venn diagram of shared and unique MACS2 called peaks in the control and FOXA1 overexpressed conditions. (control = 9; FOXA1 overexpression = 109; shared peaks = 78).

Taken together, these data suggest that FOXA1 positively influences AR recruitment to tRNA genes. Overexpression may drive increased AR binding to sites with existing AR occupancy, or drive recruitment to previously unavailable and unoccupied sites.

3.3 SWI/SNF negatively influences chromatin accessibility at tRNA genes

Building on the findings that AR, FOXA1 and H3K27ac enrichment at tRNA genes showed a strong positive correlation, ATAC-seq data from the same paper as the unpublished LNCaP findings was interrogated. Examining this data was important to establish the chromatin state of AR-bound tRNA genes.

First, accessibility in the control condition was evaluated. Heatmaps revealed a localised signal at approximately 60% of tRNA genes, relative to 10 kb upstream and downstream flanking DNA. Of all the tRNA genes, a subset of approximately 20% exhibited a particularly strong signal of highly accessible chromatin. In contrast, minimal signal was detected at most sno-miRNA genes (Figure 14A). Notably, hotspot regions of signal either side of the most accessible tRNA genes were observed. To investigate whether these accessible regions were clusters of tRNA genes, the heatmap was sorted by genomic position, revealing enrichment at chromosome 1 tRNA genes (Supplementary Figure 4). Further interrogation revealed that 12 of the top 20 genes belonged to the chr1 q23.3 'Glu-Gly-Asp-Leu' cluster repeats (Figure 14B). Accessibility signal was visualised at these clusters, revealing peaks of ATAC-seq signal (Figure 14C).

Next, comparison of ATAC-seq signal with AR, FOXA1 and H3K27ac was conducted to examine gene-level patterns of signal. To achieve this, tRNA genes were sorted by chromosome and start position and relative signal intensity for each of the four factors were visualised as a heatmap from high (red) to low (blue). As there was only 1 tRNA gene on each of chromosomes 4 and 21, the width was adjusted to accommodate the text (Figure 15). In general, expected patterns of genes with high accessibility and corresponding enrichment of AR, FOXA1 and H3K27ac were observed, including the chromosome 1 cluster described above. This suggested that AR and FOXA1 bound tRNA genes were both highly accessible and transcriptionally active. Conversely, genes with low accessibility often had low AR, FOXA1 and H3K27ac enrichment. Moreover, some of the most highly accessible genes attracted FOXA1, although there were also accessible genes without FOXA1 enrichment. Other unexpected binding patterns and chromatin states were observed, such as genes with high accessibility and H3K27ac enrichment but low AR and FOXA1 signal (e.g. tRNA-Gly-TCC-4-1 and tRNA-Gly-CCC-2-2). In addition, genes with relatively strong AR, H3K27ac and accessibility signal but low FOXA1 were noted, such as tRNA-Glu-TTC-2-2 and tRNA-Leu-TAG-3-1. In summary, the accessibility corresponded with AR, FOXA1 and H3K27ac enrichment.

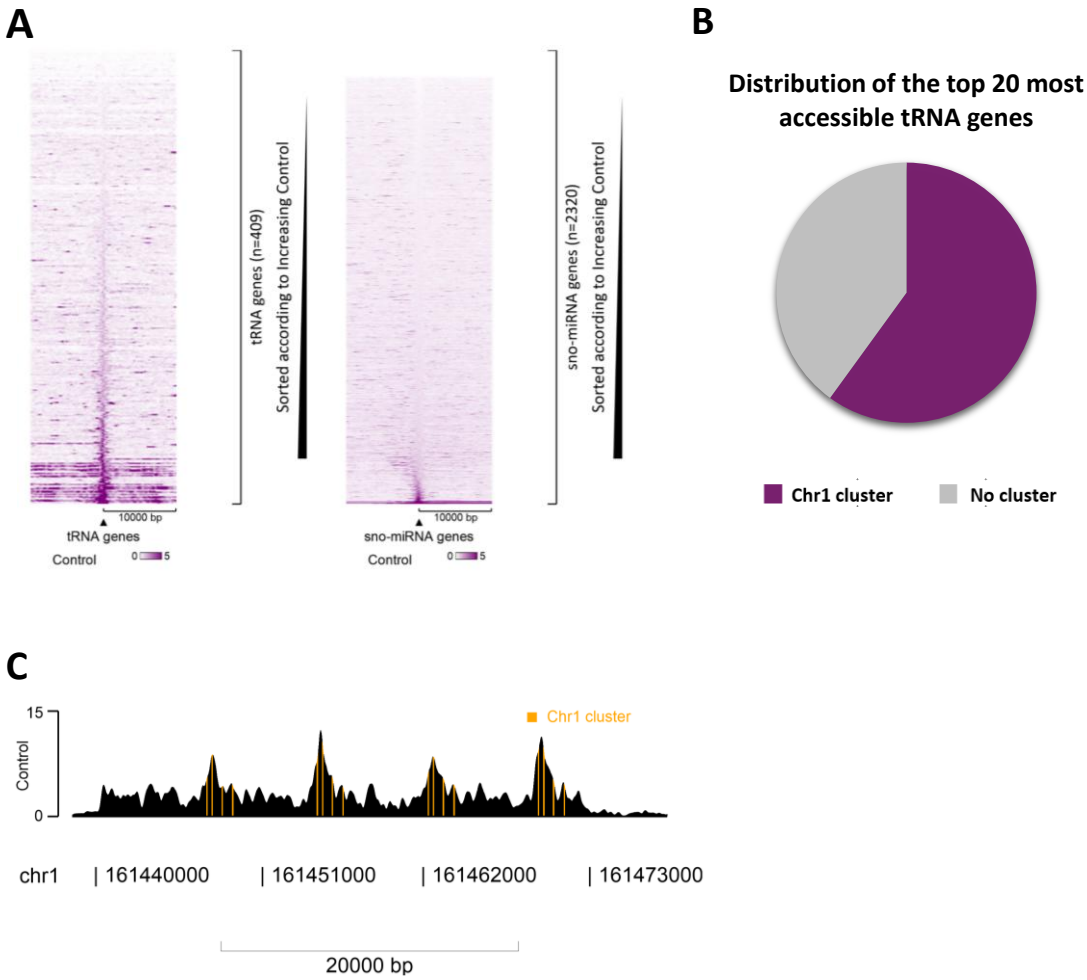


Figure 14. Chromatin accessibility at tRNA genes and the impact of SWI/SNF depletion. (**A**) Heatmaps of ATAC-seq signal sorted by increasing intensity at tRNA genes and sno-miRNA genes. (**B**) Proportion of the top 20 most accessible tRNA genes that belong to clusters of tRNA genes. (**C**) Gene track of peaks at the chromosome 1 cluster of tRNA genes.

Previously, SWI/SNF depletion was found to significantly increase AR binding at tRNA genes, indicative of a tumour-suppressor role in this context. In the ATAC-seq dataset, either one or both of the catalytic ATPase SWI/SNF subunits, SMARCA2 and SMARCA4 was depleted. For single subunit inactivation, single-guide RNA (sgRNA) targeting one subunit was used. In the dual inactivation approach, sgRNA targeting SMARCA2 was followed by single hairpin RNA targeting SMARCA4. Here, the aim was to assess whether SWI/SNF depletion also had a stimulatory effect on chromatin accessibility at tRNA genes.

Average accessibility signal across a 20kb window was higher in the single subunit knockouts compared to control at tRNA genes. Accessibility was further increased after dual subunit inactivation (Figure 16A). As sno-miRNA genes exhibited very low signal, protein-coding genes were used as a control for accessibility changes in response to SWI/SNF depletion. The same pattern was

observed at these loci, suggestive of a partial redundancy in ATPase activity between the subunits. Individual tRNA genes were examined for accessibility changes in response to single or dual subunit depletion, across a window size of 10kb from the centre of each locus. Representative filltracks show the trend of increasing accessibility, particularly between control (purple) and dual inactivation (yellow) at tRNA-Asn-GTT-6-1 and tRNA-Ser-TGA-4-1 (Figure 16C-D).

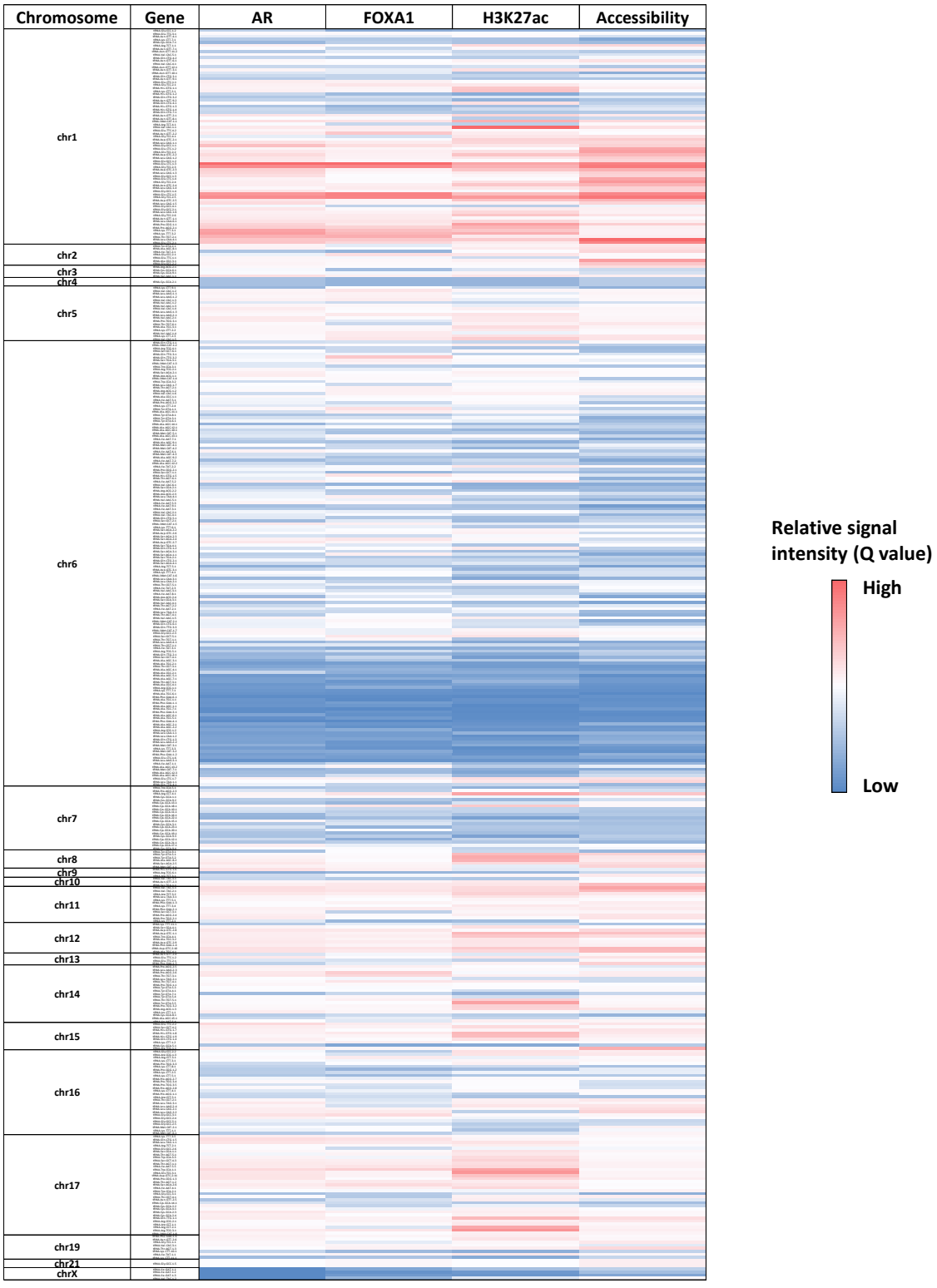


Figure 15. Relative (percentage of total) AR, FOXA1, H3K27ac and ATAC-seq enrichment at tRNA genes (n = 409) in the LNCaP cell line. Heatmap showing the relative Q values of AR, FOXA1 and H3K27ac ChIP-seq data, and ATAC-seq data, under control conditions. Q values are depicted on the colour scale, from blue (low relative signal intensity) to red (high relative signal intensity).

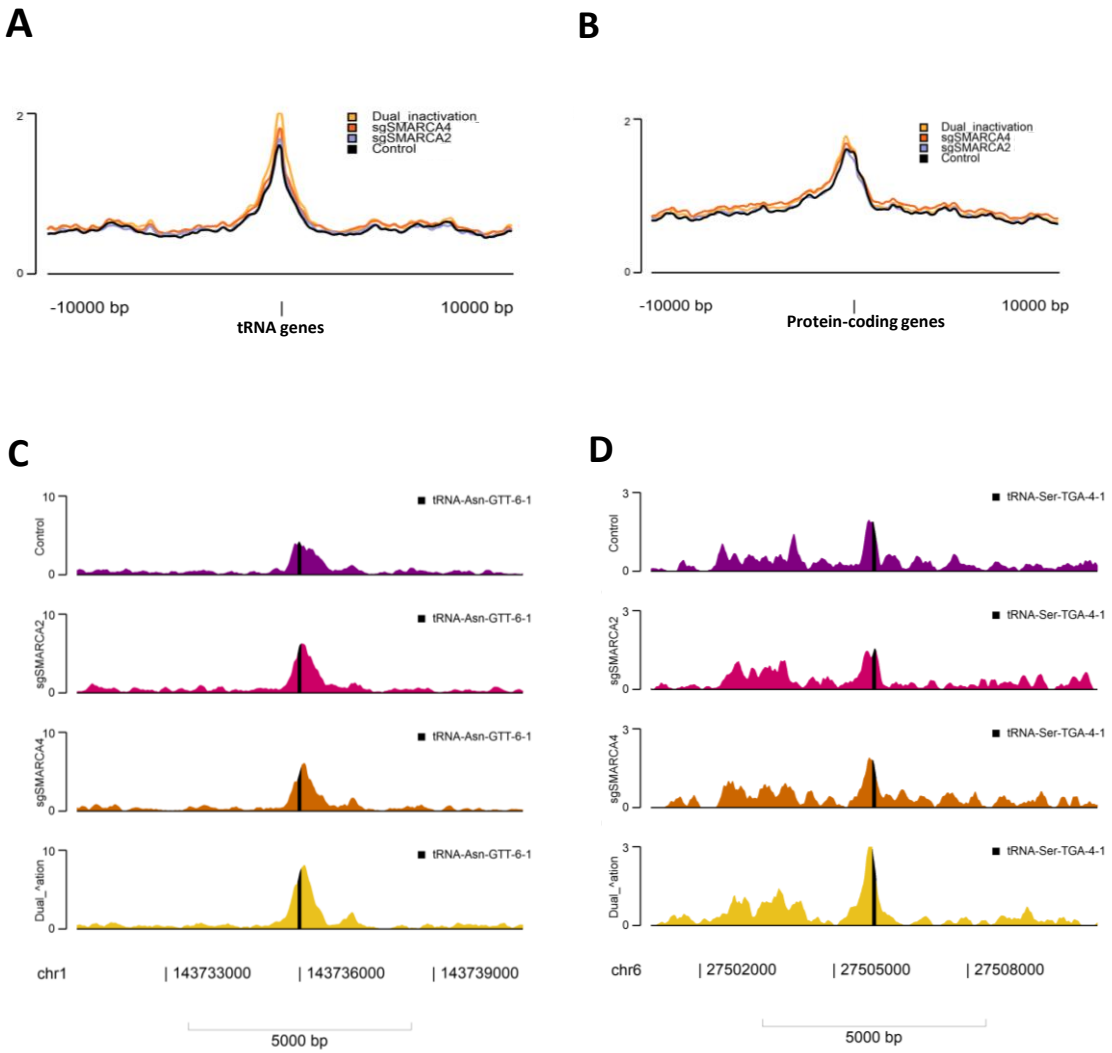
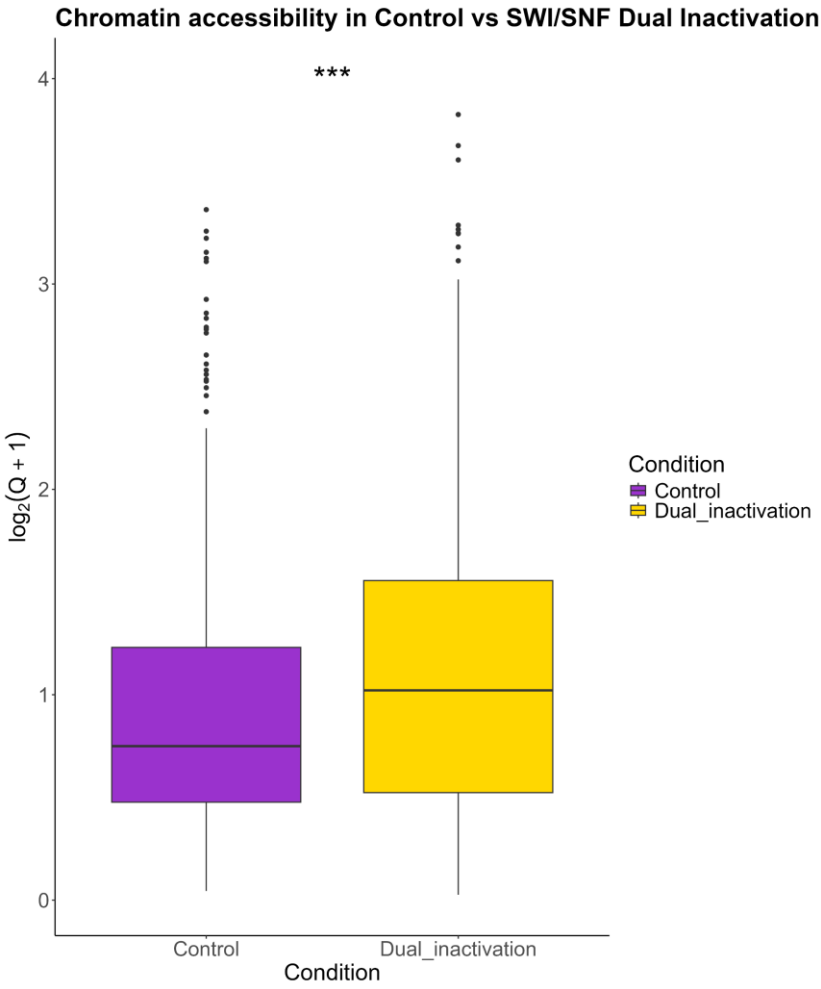


Figure 16. The impact of SWI/SNF depletion on accessibility at tRNA genes (**A**) Averaged accessibility signal intensity at tRNA genes in control, single SWI/SNF subunit inactivation or dual SWI/SNF subunit inactivation. (**B**) Averaged accessibility signal intensity at protein-coding genes in control, single SWI/SNF subunit inactivation or dual SWI/SNF subunit inactivation. (**C**-**D**) Representative filltrack of peaks at tRNA-Asn-GTT-6-1 and tRNA-Ser-TGA-4-1, in control, single subunit inactivation or dual subunit inactivation.

To further interrogate this effect, global and stratified paired gene-level accessibility signal changes were measured. Across all tRNA genes, dual SWI/SNF inactivation (median Q = 1.03) resulted in a statistically significant increase in accessibility (Wilcoxon signed rank test: V = 8021, $p < 0.001$) compared to control (median Q = 0.68) (Figure 17A). Next, genes were stratified based on their accessibility in the control condition (Figure 17B). Across all three groups, there was a significant increase in accessibility after dual inactivation (Benjamini-Hochberg corrected p-values: low = 1.21e-3; medium = 9.01e-20; high = 1.09e-21).

A


Chromatin accessibility in Control vs SWI/SNF Dual Inactivation
***
log2(Q + 1)
0
1
2
3
4
Control
Dual_inactivation
Condition
Condition
Control
Dual_inactivation

B

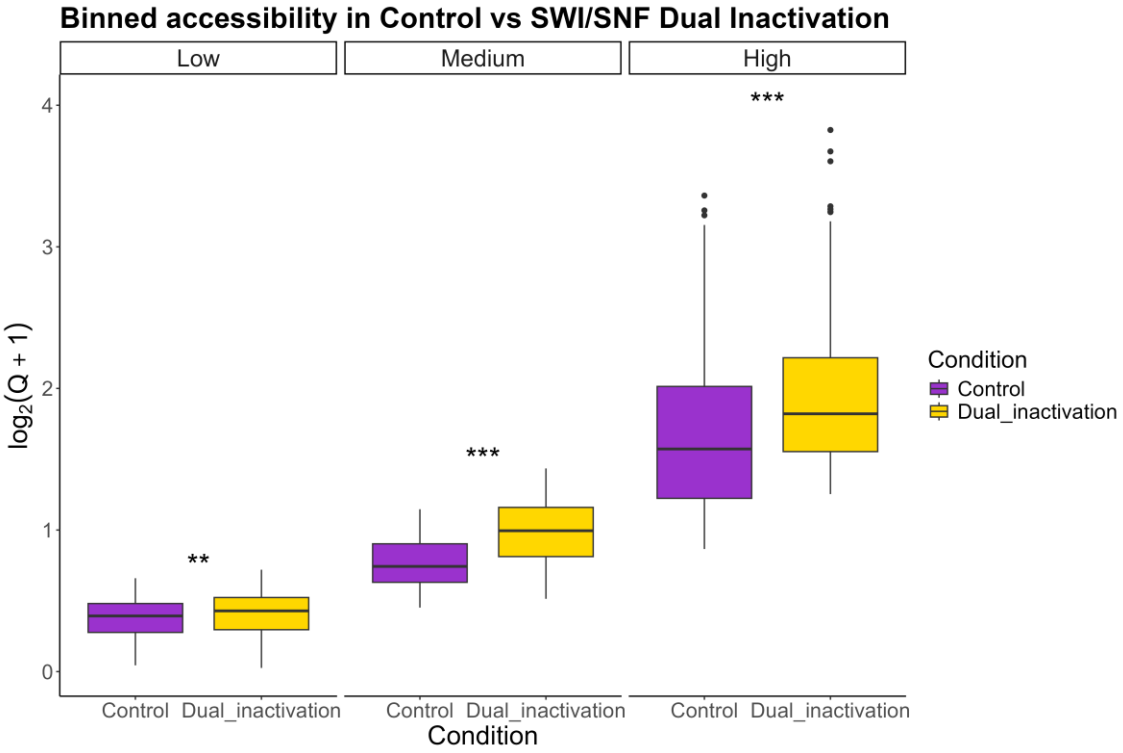

Binned accessibility in Control vs SWI/SNF Dual Inactivation
Low
Medium
High
***
***
**
log2(Q + 1)
0
1
2
3
4
Control
Dual_inactivation
Control
Dual_inactivation
Control
Dual_inactivation
Condition
Condition
Control
Dual_inactivation

Figure 17. Comparison of Q values in the control and SWI/SNF dual subunit inactivation conditions, in the LNCaP cell line. (**A**) Global overview of accessibility Q values in the control (yellow) and dual inactivation condition (purple). (**B**) Accessibility signal intensity stratified into 3 groups based on low, medium and high control Q values, and corresponding values in dual inactivation condition.

Finally, the tRNA genes with the largest percentage increases after dual inactivation were recorded (Table 4).

Table 4. Top 30 genes with the highest percentage increase between 'Control' and 'Dual inactivation' Q values

Gene name	Control	Dual inactivation	Percentage increase between conditions (%)
tRNA-Ala-AGC-5-1	0.06	0.34	438
tRNA-Ala-TGC-6-1	0.03	0.13	307
tRNA-Ile-AAT-9-1	0.14	0.37	165
tRNA-Val-CAC-3-1	0.58	1.50	159
tRNA-Ala-TGC-5-1	0.07	0.18	158
tRNA-iMet-CAT-1-7	0.67	1.70	153
tRNA-Ile-AAT-5-1	0.55	1.38	153
tRNA-Ala-AGC-6-1	0.05	0.12	152
tRNA-Arg-TCG-5-1	0.45	1.10	147
tRNA-iMet-CAT-1-6	0.54	1.33	146
tRNA-Ala-CGC-4-1	0.09	0.23	142
tRNA-Thr-AGT-6-1	0.24	0.58	140
tRNA-Ala-CGC-1-1	0.48	1.12	135
tRNA-Pro-CGG-2-1	0.38	0.87	133
tRNA-Ile-TAT-1-1	1.14	2.58	125
tRNA-Arg-TCT-2-1	0.82	1.82	122
tRNA-Leu-TAG-1-1	1.02	2.21	117
tRNA-Arg-TCT-4-1	1.00	2.17	116
tRNA-Gln-CTG-6-1	0.37	0.79	115
tRNA-Gln-CTG-1-5	0.75	1.61	114
tRNA-Leu-TAA-4-1	0.38	0.80	109
tRNA-Ala-CGC-2-1	0.22	0.46	108
tRNA-Gln-CTG-5-1	0.42	0.87	105
tRNA-Ala-AGC-11-1	0.59	1.19	103
tRNA-Gly-GCC-2-2	3.79	7.65	102
tRNA-Gln-CTG-1-2	0.43	0.87	101
tRNA-Gly-GCC-2-3	1.14	2.29	100
tRNA-Cys-GCA-2-2	0.46	0.92	99
tRNA-Leu-AAG-2-1	1.01	2.00	97
tRNA-Leu-TAA-2-1	0.31	0.61	95

3.4 CHD1 and c-Myc co-localise at tRNA genes

C-Myc can negatively influence AR recruitment to target genes, and drive an androgen independent mechanism of prostate cancer development[79–81]. Since c-Myc can also bind tRNA genes, its relationship with AR was explored[46].

AR signal peaked at the *KLK3* enhancer, provided a positive control (Figure 18A). To validate c-Myc signal detection, a peak at the *CDK4* promoter was observed (Figure 18C). Next, averaged signal intensities of AR and c-Myc were viewed across a 20,000 bp window from the centre of each tRNA gene and sno-miRNA gene control (Figure 18B and D). AR signal peaked at tRNA genes, with a smaller side peak approximately 1000bp upstream. Similarly, c-Myc signal was enriched at tRNA genes compared to sno-miRNA genes, demonstrating specificity of binding. Next, signal at tRNA genes sorted by increasing intensity, was displayed as heatmaps +/- 10kb from the centre of each tRNA gene (Figure 18G-I). Heatmaps at sno-miRNA genes showed minimal binding of AR or c-Myc (Supplementary Figure 5A-C). The heatmaps revealed localised signal at the tRNA genes for both transcription factors, with horizontal enriched signal observed. These were again confirmed to be explained by clusters of tRNA genes, and belong to the chromosome 1 region. At the tRNA-Asp-GTC-2-7, signal peaked for both AR and c-Myc.

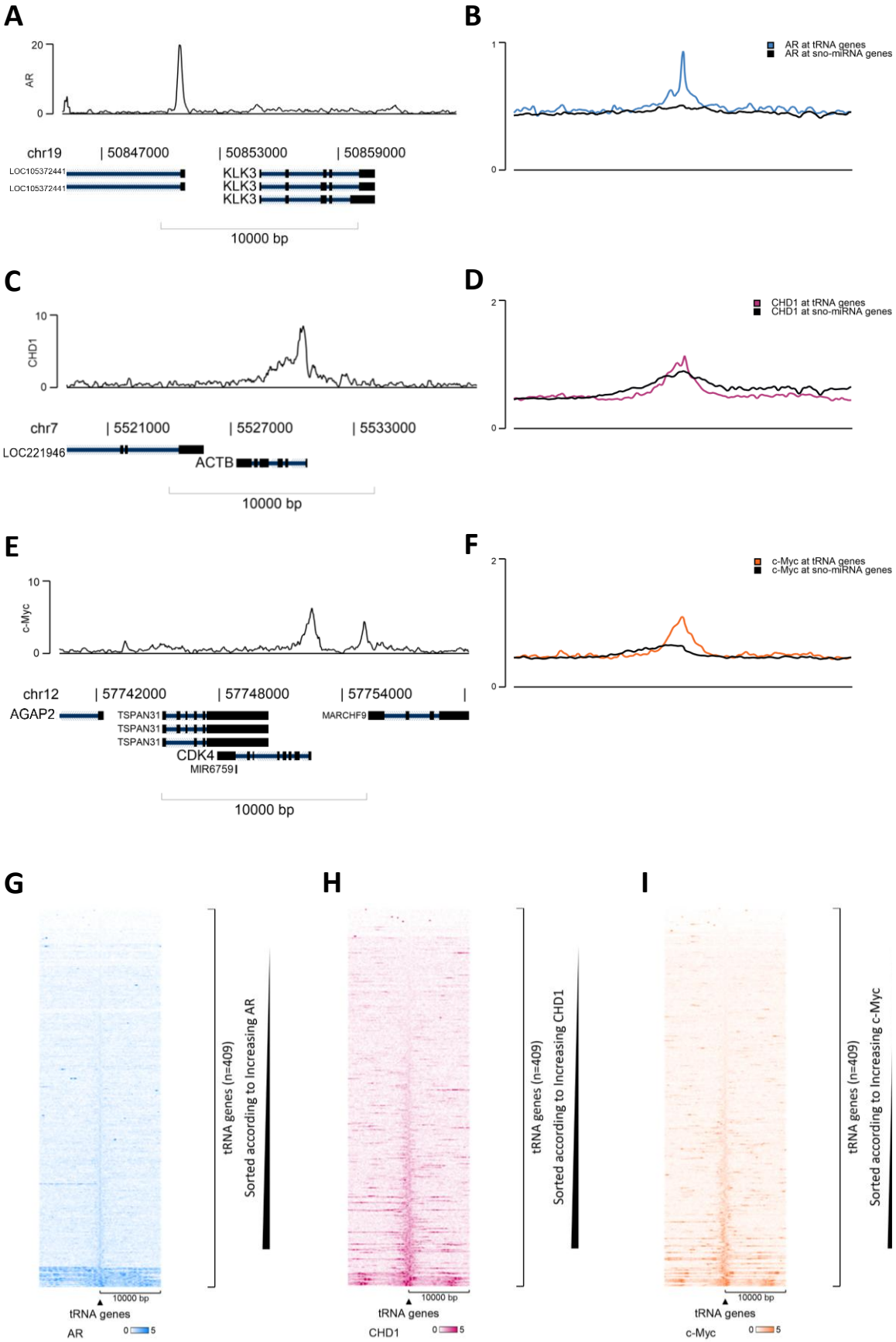

A
AR
20
0
chr19
| 50847000
| 50853000
| 50859000
LOC105372441
LOC105372441
KLK3
KLK3
KLK3
10000 bp
B
AR at tRNA genes
AR at sno-miRNA genes
1
0
C
CHD1
10
0
chr7
| 5521000
| 5527000
| 5533000
LOC221946
ACTB
10000 bp
D
CHD1 at tRNA genes
CHD1 at sno-miRNA genes
2
0
E
c-Myc
10
0
chr12
| 57742000
| 57748000
| 57754000
AGAP2
TSPAN31
TSPAN31
TSPAN31
MARCHF9
CDK4
MIR6759
10000 bp
F
c-Myc at tRNA genes
c-Myc at sno-miRNA genes
2
0
G
tRNA genes (n=409)
Sorted according to Increasing AR
10000 bp
tRNA genes
AR
0
5
H
tRNA genes (n=409)
Sorted according to Increasing CHD1
10000 bp
tRNA genes
CHD1
0
5
I
tRNA genes (n=409)
Sorted according to Increasing c-Myc
10000 bp
tRNA genes
c-Myc
0
5

Figure 18. AR, CHD1 and c-Myc binding events at tRNA genes in the LNCaP cell line. (A) Positive control for AR binding at the KLK3 promoter and upstream enhancer. **(B)** Averaged AR signal intensities at tRNA genes (blue) compared to negative control sno-miRNA genes (black). **(C)** Positive control for CHD1 binding at the ACTB promoter. **(D)** Averaged CHD1 signal intensities at tRNA genes (pink) compared to sno-miRNA genes (black). **(E)** Positive control for c-Myc binding at the CDK4 gene promoter. **(F)** Averaged c-Myc signal intensities at tRNA genes (orange) compared to sno-miRNA genes (black). **(G)** Heatmap of AR signal intensity at tRNA genes. **(H)** Heatmap of CHD1 signal intensity at tRNA genes. **(I)** Heatmap of c-Myc signal intensity at tRNA genes.

Since CHD1 and c-Myc were found to bind tRNA genes, their relationship with AR Q values was explored. Positive correlations were found between AR and c-Myc (rho = 0.84) and AR and CHD1 (0.84) Q values (Figure 19). Representative filltracks showed frequent co-localisation of these factors at tRNA genes (Supplementary Figure 6), which may suggest that c-Myc promotes rather than represses AR binding at tRNA genes in this context. However, causal relationships cannot be established by this analysis, and further work modulating c-Myc expression may provide clearer insights into its effect on AR recruitment to tRNA genes.

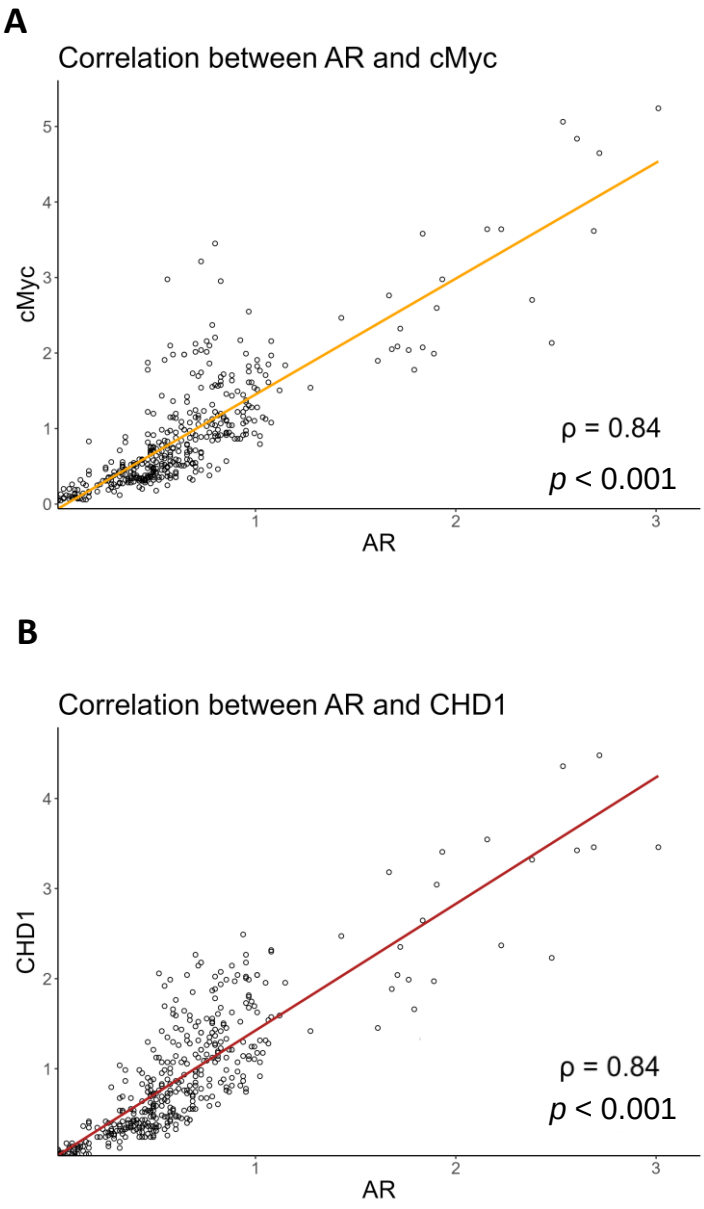


Figure 19. Spearman's rank correlations between AR and c-Myc Q values (A) and AR and CHD1 Q values (B).

Next, the epigenetic state of tRNA genes was analysed. H3K4me3 was enriched at tRNA genes relative to sno-miRNA genes (Figure 20A-B), suggesting AR-bound tRNA genes were transcriptionally active. In contrast, H3K27me3 signal was generally depleted at tRNA genes, relative to sno-miRNA genes, suggesting these loci were largely transcriptionally inactive (Figure 20C-D; Supplementary Figure 5D-E). However, a small subset of tRNA genes (largely comprising genes within the chr1 q23.3 'Glu-Gly-Asp-Leu' cluster described previously) exhibited some enrichment of H3K27me3 (Figure 20C).

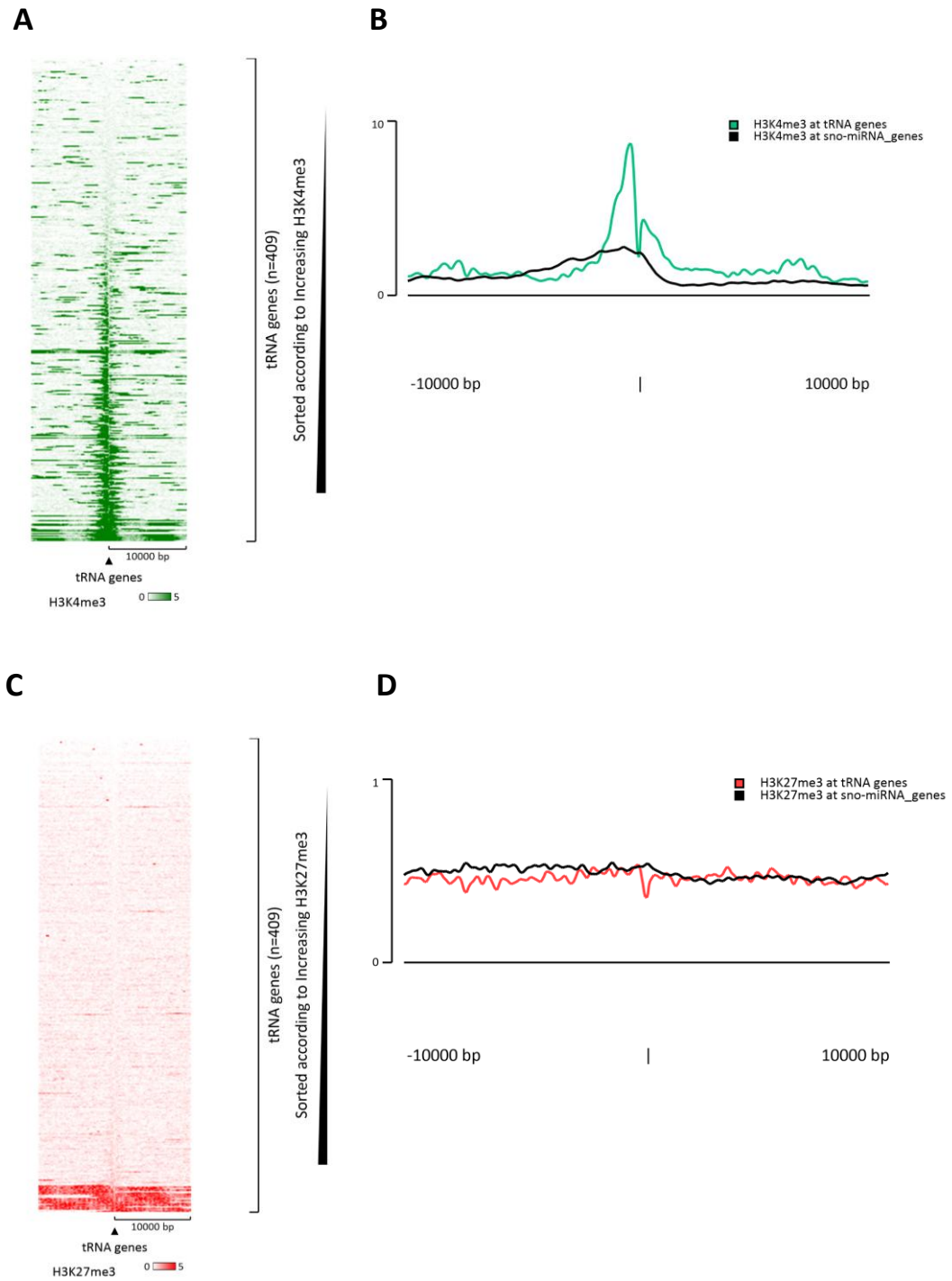


Figure 20. Epigenetic modifications at tRNA genes. (**A**) Heatmap of H3K4me3 signal. (**B**) Averaged H3K4me3 signal intensity at tRNA genes compared to sno-miRNA genes. (**C**) Heatmap of H3K27me3 signal. (**D**) Averaged H3K27me3 signal intensity at tRNA genes compared to sno-miRNA genes.

Together, these results demonstrate that CHD1 and c-Myc co-localise with AR binding at tRNA genes. H3K4me3 is enriched at these loci, and H3K27me3 is generally depleted.

3.5 CHD1 inhibits low to medium levels of AR binding at tRNA genes

CHD1 gene deletion occurs in approximately 15% of prostate cancer cases[34]. To investigate whether loss of CHD1 impacts AR binding at tRNA genes, a ChIP-seq dataset of scrambled sgCtrl and CHD1 targeting sgRNA was analysed.

First, AR peaks at the *KLK3* gene upstream enhancer and promoter were established, providing a positive control (Figure 21A). Averaged sgCtrl AR signal intensity at tRNA genes was found to be stronger and defined, compared to sno-miRNA genes (Figure 21B). Averaged AR signal in the sgCHD1 condition was slightly increased compared to control (Figure 21C). Heatmaps of AR enrichment at tRNA genes across a 20 kb window showed binding specificity relative to surrounding DNA. However, visible differences in binding intensity were harder to distinguish here (Figure 21D). Again, minimal association of AR was found at sno-miRNA genes (Supplementary Figure 7). Filltracks of AR signal across a 10kb region from the centre of a tRNA gene locus revealed stronger enrichment at or near many loci after CHD1 loss. For example, tRNA-Asp-GTC-2-7 (sgCtrl Q value = 1.63; sgCHD1 Q value = 2.44) and tRNA-Cys-GCA-18-1 (sgCtrl Q value = 1.13; sgCHD1 Q value = 1.22). However, examples of minimal impact on AR signal were also observed, such as tRNA-Lys-TTT-3-2 (sgCtrl Q value = 1.07; sgCHD1 Q value = 1.10) (Figure 21E-H).

Globally, there was a significant increase in AR signal in the CHD1 depletion condition, compared to control (Wilcoxon signed rank test: V = 50120, p-value = 0.00061) (Figure 22A). Next, tRNA genes were stratified into low, medium and high groups based on thresholds set by Q values in the control (33rd percentile Q = 0.41; 66th percentile Q = 0.65) (Figure 22B). Gene-level tests with Benjamini Hochberg correction were performed in each group, revealing a significant increase in the low and medium groups (p = 8.43e-03, 9.86e-08, respectively). However, there was no significant impact on the high group, (p = 1.15e-01). This suggests that CHD1 has a modest inhibitory effect on AR binding, but stronger levels of AR binding can counteract this effect.

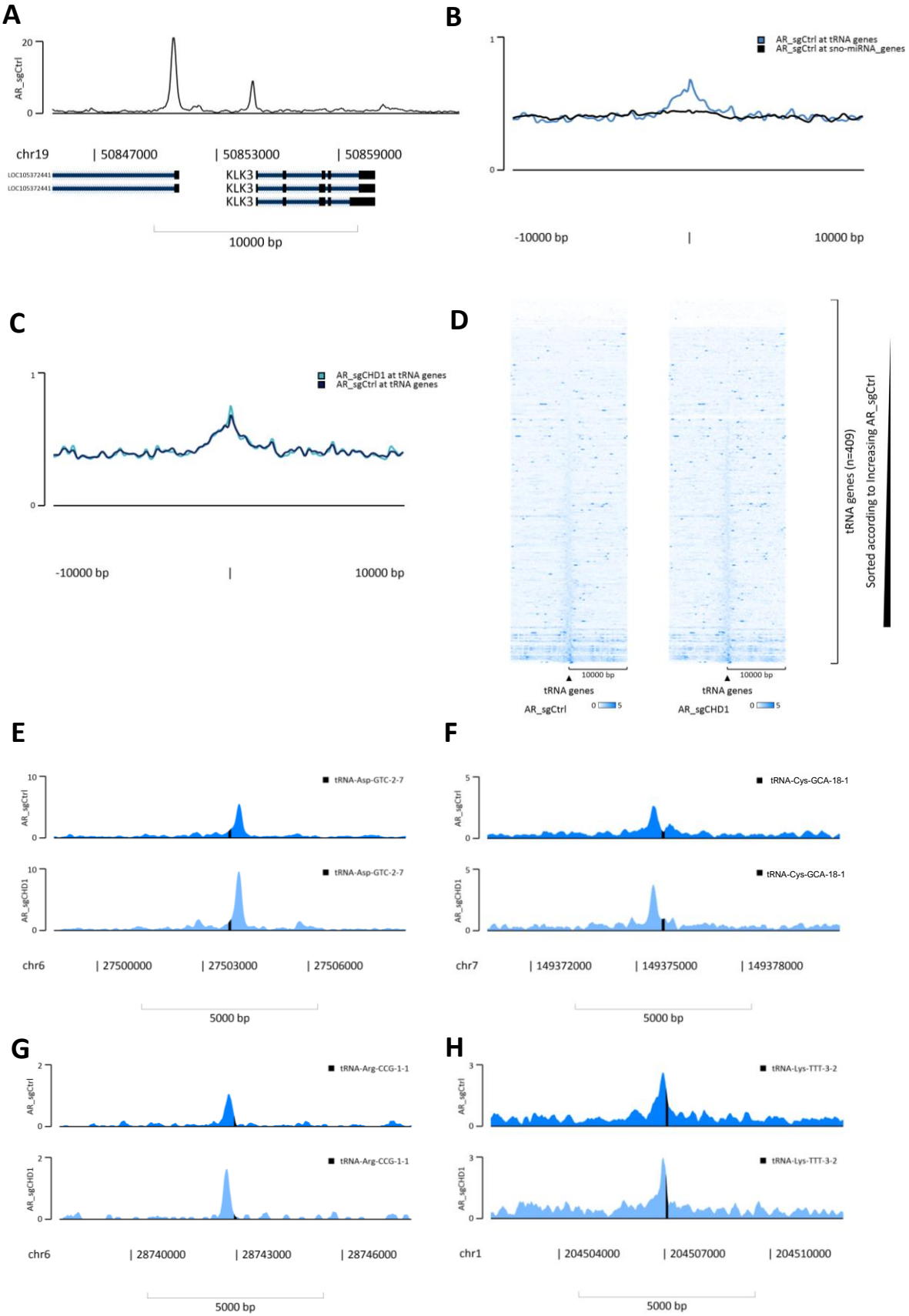

A
AR_sgCtrl
20
0
chr19
| 50847000
| 50853000
| 50859000
LOC105372441
LOC105372441
KLK3
KLK3
KLK3
10000 bp
B
AR_sgCtrl at tRNA genes
AR_sgCtrl at sno-miRNA_genes
1
0
-10000 bp
|
10000 bp
C
AR_sgCHD1 at tRNA genes
AR_sgCtrl at tRNA genes
1
0
-10000 bp
|
10000 bp
D
10000 bp
tRNA genes
AR_sgCtrl
0
5
10000 bp
tRNA genes
AR_sgCHD1
0
5
tRNA genes (n=409)
Sorted according to Increasing AR_sgCtrl
E
tRNA-Asp-GTC-2-7
AR_sgCtrl
10
0
tRNA-Asp-GTC-2-7
AR_sgCHD1
10
0
chr6
| 27500000
| 27503000
| 27506000
5000 bp
F
tRNA-Cys-GCA-18-1
AR_sgCtrl
5
0
tRNA-Cys-GCA-18-1
AR_sgCHD1
5
0
chr7
| 149372000
| 149375000
| 149378000
5000 bp
G
tRNA-Arg-CCG-1-1
AR_sgCtrl
2
0
tRNA-Arg-CCG-1-1
AR_sgCHD1
2
0
chr6
| 28740000
| 28743000
| 28746000
5000 bp
H
tRNA-Lys-TTT-3-2
AR_sgCtrl
3
0
tRNA-Lys-TTT-3-2
AR_sgCHD1
3
0
chr1
| 204504000
| 204507000
| 204510000
5000 bp

Figure 21. The impact of CHD1 knockdown on AR binding events at tRNA genes in the LNCaP cell line. (**A**) positive control at *KLK3* promoter and upstream enhancer. (**B**) Averaged AR signal intensities at tRNA genes (blue) compared to negative control sno-miRNA genes (black). (**C**) Averaged AR signal intensities at tRNA genes in control vs CHD1 knockdown. (**D**) Heatmap of AR signal intensity at tRNA genes in control vs CHD1 knockdown. (**E**) Gene track of AR binding at tRNA-Arg-CCG-1-1. (**F**) Gene track of AR binding at tRNA-Lys-TTT-3-2.

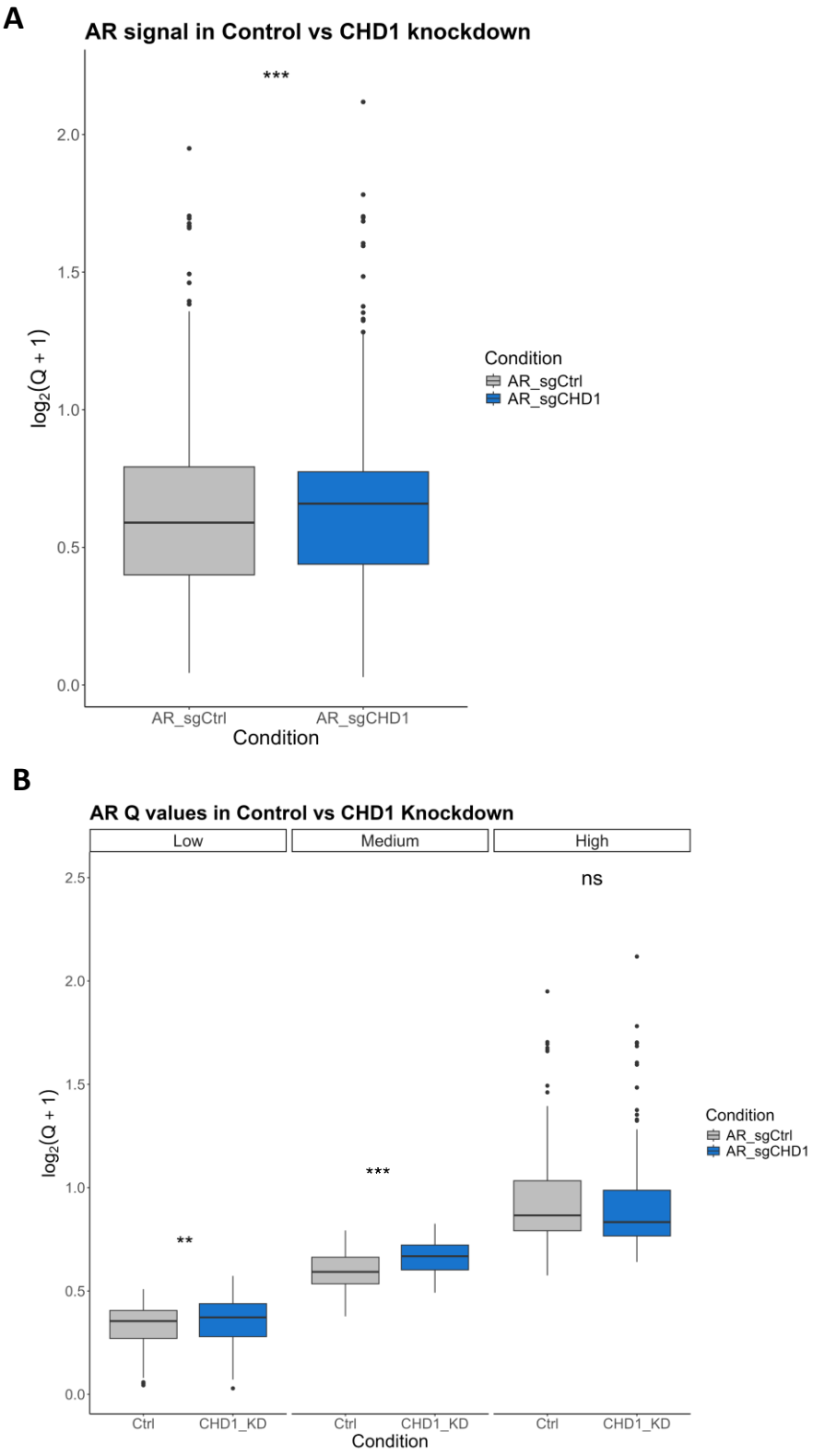


Figure 22. Comparison of Q values in the control and CHD1 knockdown conditions, in the LNCaP cell line. (**A**) Global overview of accessibility Q values in the control (grey) and knockdown (blue). (**B**) Accessibility signal intensity stratified into 3 groups based on low, medium and high control Q values, and corresponding values in knockdown.

Finally, the top 30 tRNA genes by percentage increase were listed in Table 5.

Table 5. Top 30 genes with the highest percentage increase between Control and CHD1 knockdown Q values

Gene name	Control	CHD1 knockdown	Percentage increase between conditions (%)
tRNA-Leu-CAA-1-1	0.06	0.17	204
tRNA-Ala-TGC-5-1	0.04	0.08	125
tRNA-Ser-AGA-4-1	0.31	0.67	113
tRNA-Thr-AGT-4-1	0.24	0.49	106
tRNA-Gln-CTG-1-3	0.09	0.17	97
tRNA-Leu-AAG-3-1	0.04	0.08	97
tRNA-Ile-TAT-3-1	0.09	0.18	97
tRNA-iMet-CAT-1-4	0.49	0.92	89
tRNA-Arg-TCG-5-1	0.25	0.46	81
tRNA-Arg-CCG-1-2	0.06	0.11	81
tRNA-Arg-ACG-2-2	0.30	0.54	80
tRNA-Gln-CTG-2-1	0.33	0.59	78
tRNA-Pro-AGG-1-1	0.38	0.68	78
tRNA-Trp-CCA-3-2	0.37	0.65	75
tRNA-Phe-GAA-1-1	0.08	0.14	72
tRNA-Phe-GAA-6-1	0.04	0.07	72
tRNA-Gly-CCC-3-1	0.21	0.36	72
tRNA-Thr-CGT-3-1	0.21	0.36	72
tRNA-Val-AAC-1-5	0.42	0.72	71
tRNA-Gln-TTG-2-1	0.27	0.46	70
tRNA-Ser-TGA-4-1	0.37	0.63	70
tRNA-Val-TAC-4-1	0.29	0.49	69
tRNA-Phe-GAA-3-1	0.04	0.06	69
tRNA-Lys-TTT-4-1	0.39	0.65	68
tRNA-Thr-AGT-2-2	0.39	0.65	68
tRNA-Ile-AAT-3-1	0.34	0.56	67
tRNA-Ser-GCT-2-1	0.27	0.45	67
tRNA-Cys-GCA-17-1	0.46	0.75	64
tRNA-Ala-AGC-2-2	0.08	0.13	60
tRNA-Val-AAC-5-1	0.32	0.52	59

In summary, these show that AR binding increases at a subset of tRNA genes when CHD1 is deleted.

4. Discussion

Dysregulated AR signalling is a key driver of tumorigenesis and progression in PC[9]. Targeted disruption of this pathway, either by lowering serum androgen levels or by antagonistically binding the AR with high affinity, aims to prevent AR nuclear translocation and consequently aberrant DNA binding. While these strategies are often initially clinically effective, many patients later develop castration-resistant PCa, driven by extensive reprogramming of the AR cistrome[4,35,82]. Therefore, understanding the extent of AR binding in PCa, and which co-factors modulate its interactions is important.

Abnormal tRNA expression has been observed in many cancer types, and recent studies have provided evidence for ER and PR binding at tRNA genes in breast cancer[47,49,58,59,83,84]. However, despite evidence for altered amino acid metabolism and the detection of tRNA derived fragments in PCa, whether AR can bind tRNA genes remains unexplored in the literature.

Previous unpublished work by the White group provided preliminary evidence for AR binding events at a subset of tRNA genes in the LNCaP cell line. FOXA1 co-localisation was found at a majority of these tRNA ARBS, and signal intensities of AR, FOXA1 and H3K27ac were strongly correlated. Furthermore, depletion of the SWI/SNF chromatin remodeler via targeted PROTAC-mediated degradation, resulted in significant increases to AR, but not FOXA1, binding intensities at tRNA genes. This suggested a potential inhibitory role of SWI/SNF in modulating AR recruitment to tRNA genes, consistent with its role as a tumour suppressor.

This current project aimed to build on these findings, by strengthening evidence for AR binding at tRNA genes in PCa, and exploring the role of different co-factors and chromatin remodelers in AR recruitment. This study provided evidence for AR binding at tRNA genes in four PCa ChIP-seq datasets – including validation in patient biopsy tissues. The overexpression of FOXA1 was found to promote AR recruitment to tRNA genes, by enhancing binding to existing ARBS and by facilitating de-novo binding events. Local chromatin accessibility at tRNA genes was found to broadly correspond with AR, FOXA1 and H3K27ac enrichment, and was sensitive to SWI/SNF depletion. Moreover, the CHD1 chromatin remodeler co-localises with AR and c-Myc at tRNA genes. CHD1 depletion resulted in an increase in AR binding events at low and medium affinity AR sites.

4.1 Patient biopsy validation

Prior to carrying out further bioinformatic analysis, it was important to confirm whether these binding events could be observed at tRNA genes in PCa patient tissue and were not artefacts of the LNCaP cell line.

To achieve this, ChIP-seq data from four patient tumour biopsies (with localised or locally advanced disease) were analysed. One advantage to this dataset, was the retention of physiologic androgen levels (although androgen sensitivity data was unavailable), especially since the previous analysis in LNCaP cells used an artificial DHT which is reported to have a stronger affinity for the AR. Although all samples had previously been classified as AR+ by the study authors, no noticeable enrichment of AR at either the *KLK3* enhancer or promoter could be detected in patient 3. This is consistent with the authors findings, where patient 3 exhibited the lowest number of genome-wide ARBS (91), compared to 559, 3123 and 25,475 for patients 1, 2 and 4[72]. Therefore, this sample was excluded from further analysis. Interestingly, the clinical grading of patient 3's tumour cell pathology represented the most aggressive, with a Gleason score of 10, indicating poor differentiation. Since extensive AR reprogramming during cancer progression has been shown, it could be hypothesised that in this case AR is being directed away from the classical targets. However, as expected, the remaining patients (1, 2 and 4) showed AR enrichment was specific to tRNA genes and present in a subset of genes representing approximately half of all tRNA genes in the genome.

In contrast, all four patient samples had strong FOXA1 peaks at the *KLK3* enhancer. Genome-wide FOXA1 binding-sites were similar for patients 1, 3 and 4 (12860, 7543 and 10615). However, patient 2 had a substantially large number of peaks (61152). H3K27ac enrichment was present at tRNA genes, with the characteristic peaks adjacent to the nucleosome depleted loci.

Together, these data show that AR binding at tRNA genes is not an artefact of the LNCaP cell line, and represents a novel finding in the field.

4.2 FOXA1

FOXA1 is a well-established direct binding partner of AR, and frequently mutated in PCa - generally considered to promote AR recruitment to target genes via its pioneering activity[14,35,85,86]. In addition, studies have found FOXA1 to correlate with poor PCa outcomes, in direct contrast to breast cancer[87,88]. *FOXA1* gene mutations occur frequently in all stages of PCa, and hotspots regions of missense point mutations and indels have been identified in the FKHD DBD domain (including the

Wing2 residues). Given their location, studies have characterised the functional consequences of some of the most common *FOXA1* mutations on DNA binding affinity, nuclear mobility and cell growth, and shown that they confer a gain of function[18,89,90].

This current study found that FOXA1 overexpression (by cDNA transfection) resulted in a significant increase in AR binding at tRNA genes, compared to control. Interestingly, both enhanced AR binding at sites with existing AR occupancy and de-novo binding events were identified. Only a small proportion of AR-bound tRNA genes in the control were lost after FOXA1 overexpression, suggesting that FOXA1 largely exhibits a positive influence on AR recruitment to tRNA genes in this context.

This effect is partially consistent with the genome-wide findings reported in the source paper. Globally, the average AR read count was significantly higher under FOXA1 overexpression, and a small percentage of ARBS were unique to the control. However, the distribution of shared and gained sites after FOXA1 overexpression differed between the two analyses. Globally, the authors detected a similar distribution of sites that were either conserved between conditions (29,055 - many of which exhibited stronger binding intensity after overexpression) or gained after FOXA1 overexpression (28,100) – representing 51% and 49%, respectively[73]. In contrast, the peak calling performed in this study revealed a shift in the distribution towards gained sites (58%) compared to shared (42%). This may indicate that in this context, elevated FOXA1 levels promote AR binding at tRNA genes with low affinity AR sites preferentially to enhancing occupancy at existing sites.

There is also evidence that FOXA1 can suppress AR binding events across the genome. FOXA1 knockdown by siRNA in the LNCaP cell line resulted in a total of 12,103 ARBS after treatment (with 10,869 'gained'), compared to 3,115 total ARBS in the control [91]. This was supported by another study demonstrating 17,022 total ARBS under FOXA1 knockdown (with 13,505 gained), compared to 6,215 ARBS in the control[16]. These gained ARBS in the absence of FOXA1 represent a different subset of genes from those that are gained upon FOXA1 elevation[73]. Therefore, whilst FOXA1 overexpression facilitated AR binding at a subset of tRNA genes in this analysis, the possibility of FOXA1 acting to suppress AR binding to other tRNA genes cannot be ruled out.

4.3 ATAC-seq

This study found a subset of tRNA genes to be accessible, broadly corresponding with AR, FOXA1 and H3K27ac enrichment. In particular, a cluster of tandem tRNA gene repeats on chromosome 1 (Glu-Gly-Asp-Leu) were among the most accessible. Evidence of tRNA gene clusters serving as barriers to maintain the separation of heterochromatic and euchromatic regions first came from *Saccharomyces*

cerevisiae[92]. Later, the mouse equivalent of the human chromosome 1 cluster was experimentally validated to prevent the spread of heterochromatin [93]. ChIP-seq analysis in K562 cells found enrichment of euchromatic features such as H3K27ac and H3K4me3, and Pol III machinery subunits Brf1 (TFIIIB), GTF3C2 (TFIIIC) and POLR3A (Pol III) at these gene clusters. The heterochromatic marks H3K27me3 and H3K9me3 were detected at the boundaries of these clusters, suggesting that they may also function as barriers in human cells[94]. Therefore, this study's findings that some of the most accessible tRNA genes belong to the chromosome 1 clusters attracting AR and FOXA1 is noteworthy.

The SWI/SNF chromatin remodeler plays an important role in the dynamic nature of chromatin state, by remodelling nucleosomes to influence local chromatin accessibility. In many cancer types, SWI/SNF is believed to act as a tumour suppressor, and it is frequently mutated in PCa[22,95]. The previous work carried out by the White group showed that SWI/SNF depletion resulted in a significant increase in AR and H3K27ac binding at tRNA genes, consistent with its putative role as a tumour suppressor. In the present study, inactivation of either SWI/SNF ATPase subunit, SMARCA2 and SMARCA4, resulted in an increase in chromatin accessibility signal. This effect was enhanced by dual subunit inactivation. Since each SWI/SNF molecule contains one of the two subunits, this effect is consistent with a model of residual activity delivered by SWI/SNF molecules with the unaffected ATPase.

Interestingly, these findings contrast with the genome-wide patterns reported in the source paper. Rapid losses in chromatin accessibility within 1 hour of targeted PROTAC degradation of SWI/SNF were observed. This effect was also observed in the genetic models of SWI/SNF inactivation, targeting the SWI/SNF subunits. A large proportion of distal regulatory elements including enhancer elements were compacted after SWI/SNF inactivation, whereas promoters represented a majority of retained sites. The retention of promoters in the global analysis may provide insight into a possible explanation for the discrepancies between altered chromatin state in the source paper and the present study[61]. Since tRNA genes contain internal promoter elements, and accessibility was quantified ± 500 bp from the centre of each gene, potential distal regulatory elements were not measured. Stable CRISPR-cas9 inactivation of either subunit, and Tet-ON doxycycline inducible shRNA targeting of SMARCA4 (in the case of dual inactivation), was followed by ATAC-sequencing 72 hours after the addition of doxycycline. This time period may have been sufficient to allow a promoter-specific compensatory mechanism to arise. For example, there is evidence of recovering accessibility at promoters, but not enhancers, by Tip60/EP400 after SWI/SNF loss. Although immediate loss of promoter accessibility was observed, accessibility started to recover after 4 hours, and peaked at 8 hours[96]. Since ATAC-seq was performed over a considerably longer time than this period, after SWI/SNF inactivation, it cannot be concluded whether SWI/SNF loss immediately

resulted in an increase in promoter accessibility at tRNA genes, or whether an initial decrease followed by a large recovery effect was observed.

4.4 c-Myc and CHD1 co-localisation with AR at tRNA genes

The proto-oncogene *c-Myc*, is commonly upregulated in PCa, and is a driver of both PCa initiation and progression. C-Myc can promote Pol III transcription of tRNA genes by directly interacting with TFIIIB and recruiting the GCN5 histone acetyltransferase alongside its cofactor (TRRAP) [45,46,97]. Although a direct physical binding of AR and c-Myc has not been proven, many studies have investigated their interactions [79,80,98–100].

The CHD1 chromatin remodeler is disproportionately altered in PCa. Although there is substantial evidence of c-Myc binding SWI/SNF directly (via the INI1/hSNF5 or BAF155 subunits), it is unknown whether c-Myc and CHD1 directly interact [101,102].

This study found AR, c-Myc and CHD1 enrichment and frequent colocalization at tRNA genes. In addition, tRNA genes were enriched for H3K4me3, a mark of active gene transcription. Although the heatmap (Figure 20C) showed high H3K27me3 intensity (a mark of gene repression) at DNA flanking a small proportion of tRNA genes, this coincided with H3K4me3 signal. This likely reflects bivalent chromatin – which can also be observed in the source paper at protein-coding genes. Given the strong positive correlations of c-Myc and CHD1 ChIP-seq signal with AR signal, it could be hypothesised that these are interacting with AR, although a causal link cannot be established. As AR is recruited by ARE motifs and or FOXA1, and CHD1 is recruited by H3K4me3, they may not interact directly.

The relationship between AR and c-Myc in PCa appears complex. There is some evidence to suggest a bi-directional action, with androgen deprivation increasing c-Myc expression, and c-Myc downregulation increasing AR transcriptional activity at target genes. Overexpression of c-Myc has been shown to repress AR target genes by causing increased promoter-proximal pausing of RNA Pol II[80]. However, co-localisation of c-Myc with AR has been reported, with enrichment of H3K27ac, suggestive of a promotional role in AR recruitment[79,98]. Moreover, AR protein levels were significantly reduced after c-Myc knockdown[99]. There is conflicting evidence of c-Myc driving progression to an androgen-independent CRPC disease state. Patients with low AR and high c-Myc gene expression are less likely to respond effectively to next generation AR antagonists like enzalutamide[80]. However, another study found a strong correlation between AR and c-Myc protein levels in patients with

CRPC[100]. In this context, investigating c-Myc modulation, such as by overexpression or knockdown, may provide insight into the effect on AR binding.

This study also examined the effect of CHD1 knockout on AR binding at tRNA genes. AR binding significantly increased at a subset of tRNA genes when CHD1 was lost. This may suggest a tumour suppressive role and is consistent with the effect observed in the source paper at protein-coding genes. In some cases, CHD1 can evict low affinity binding by AR, although this is a minor effect. The modest effect of losing *CHD1* is consistent with evidence that *CHD1* deregulation alone is insufficient to drive tumour growth in mouse models[34,103].

4.5 Cancer biomarkers

In recent years, interest in tRNA-derived fragments (tRFs) as potential biomarkers in different cancer types has emerged. These fragments are cleaved from mature and precursor tRNA and are approximately 14-30 nt[104]. In PCa, the overexpression of tRFs derived from tRNA-Ser, tRNA-Lys and tRNA-Glu have been detected[52–54]. Although there is no evidence to date that increasing tRNA gene transcription affects tRNA fragment processing, this study found genes encoding Ser, Lys and Glu-accepting tRNAs to be among the most upregulated.

4.6 Study limitations and bioinformatic considerations

One key limitation of ChIP-seq data analysis, is the inability to differentiate between direct DNA binding of a transcription factor, or indirect binding via protein-tethering mechanisms. Instead, what can be concluded is transcription factor occupancy - the term ‘binding’ has been widely adopted to mean occupancy in the literature. Validation of a selection of tRNA gene peaks by chromatin immunoprecipitation-quantitative polymerase chain reaction (ChIP-qPCR) and DNase footprinting could ascertain direct binding events at tRNA genes.

Peak calling is a common tool used in bioinformatic pipelines, to identify enriched chromatin regions relative to input control (DNA that has not been immunoprecipitated). Two datasets, FOXA1 overexpression (PRJEB4235) and CHD1 (GSE117430), included appropriate DNA input samples. However, an unsatisfactory quality check on the CHD1 input controls meant that they were excluded from analysis. Therefore, MACS2 peak calling was performed only on the FOXA1 dataset.

Input normalisation (ChIP/input) often improves data visualisation by mitigating for sequencing bias and background noise, leaving true enrichment patterns clearer in genome browser tracks. There was one pooled input in the patient dataset (GSE114737), which was not ideal as the input signal would be expected to be different between individual patients. To maintain consistency throughout the different analyses where input was either unavailable or of low-quality, the decision was taken to use inputs qualitatively. This was achieved through including input signal above tRNA gene tracks for a visual reference for background enrichment without disrupting and potentially biasing the complete ChIP-seq enrichment patterns.

The FOXA1 overexpression dataset was the only dataset with an appropriate input control, enabling background signal to be controlled for. The inability to distinguish true enrichment from background signal in the other datasets meant that direct comparisons between datasets was not possible. Since background signals can vary greatly between different experiments, it may have been misleading to directly compare quantified ChIP-seq or ATAC-seq signal at tRNA genes between datasets. In addition, whilst arbitrary thresholds for enriched tRNA genes could be set in other datasets such as the ATAC-seq dataset, comparisons would be inconsistent and unreliable.

Another technical limitation relating to the datasets was the number of replicates per condition. To run differential binding analysis (e.g. DiffBind in R), three biological replicates are strongly recommended to provide adequate statistical power. This would then enable plotting of differential peaks as volcano plots or heatmaps. The same limitation is true for running principal component analysis (PCA), an exploratory tool to visualise clustering of samples between conditions or replicates. Therefore, in the absence of sufficient replicates to identify differentially bound tRNA genes, tRNA genes with the largest Q value percentage changes between conditions were listed in tables. This enabled the identification of enrichment changes at individual tRNA genes.

The presence of ATAC-seq data with corresponding ChIP-seq data facilitated analysis of the correlations between chromatin accessibility and enrichment of AR, FOXA1 and H3K27ac at tRNA genes. Unfortunately, ATAC-seq data was not available for the other datasets, meaning that transcriptional activity could only be inferred by quantifying signal of specific histone modifications such as H3K27ac or H3K4me3.

4.7 Future work

Future work would benefit from directly addressing the bioinformatic limitations presented above. In addition, further laboratory investigations to explore unanswered questions, including the potential involvement of c-Myc in AR recruitment to tRNA genes, would enhance understanding.

From a bioinformatics perspective, repeating the experimental designs from the Results chapter with satisfactory input controls would enable true signal enrichment (with background signal accounted for) at tRNA genes to be detected via MACS2 peak-calling. This would also prevent the need for creating arbitrary thresholds for whether a protein is 'bound' or not, as the peak calling algorithm would ensure consistency between analyses. Moreover, adequate input controls would also facilitate comparative analyses between experiments, such as determining how local chromatin accessibility correlates with tRNA genes enriched for AR binding under control or FOXA1 overexpression. Furthermore, future experiments may benefit from containing at least three replicates per condition, to facilitate running principal component analyses and differential binding to visualise sample clustering and identify consistent binding patterns at tRNA genes between different conditions.

In this study, tRNA gene transcriptional activity is inferred by the signal quantification of various histone modifications. Future work could use an antibody against Pol III instead, to determine its occupancy at specific genomic locations. Alternatively, techniques such as tRNA-sequencing (tRNA-seq) and northern blotting would enable quantification of tRNA abundance. These techniques could be used to quantify tRNA expression in response to changing conditions, and strengthen this study's findings.

Whilst this study found evidence of AR and c-Myc co-localisation at a subset of tRNA genes, it remains unclear whether c-Myc promotes or represses AR binding in this context. Further ChIP-seq experiments using antibodies against AR and Pol III under c-Myc overexpression or knockdown conditions may help determine the role of c-Myc in AR recruitment to tRNA genes. Performing ATAC-seq and tRNA-seq experiments would support findings by investigating how chromatin accessibility and tRNA expression change in response to c-Myc modulation. In addition, co-immunoprecipitation (co-IP) experiments would be useful in determining whether there is a physical association between these two proteins.

To further investigate the effect of *CHD1* deletion on AR binding at tRNA genes, future ChIP-seq experiments could use antibodies against Pol III and histone modifications such as H3K4me3 and H3K27ac. This, in combination with ATAC-seq experiments, would facilitate investigations into

whether increased AR binding in response to loss of CHD1 correlates with increased tRNA gene transcription.

Although global AR binding intensity at tRNA genes significantly increased when FOXA1 was overexpressed, there was a small subset of tRNA genes for which AR binding was only observed in the control condition, suggestive of an inhibitory effect at select loci. To validate these findings, ChIP-qPCR could be performed to quantify AR enrichment differences in FOXA1 overexpression compared to control. Moreover, evaluating AR binding in response to FOXA1 knockdown would reveal whether the same loci that are upregulated in the FOXA1 overexpression condition, are downregulated in the knockdown, and therefore most sensitive to FOXA1 modulation.

Since global AR binding patterns shift during PCa disease progression, it would be beneficial to explore whether similar changes occur at tRNA genes. For example, studying AR binding events in tumour samples obtained from PCa patients with different grades and stages may reveal distinct differences in the selection of tRNA genes that AR targets across disease states. In addition, quantifying the binding of AR mutants, such as AR-V7, in PCa cell lines such as VCap or 22Rv1 may further enhance understanding.

4.8 Conclusion

In conclusion, this study provides evidence of AR binding at a subset of tRNA genes in PCa, and explores the effect of co-factor and chromatin remodeler modulation. FOXA1 overexpression was found to increase AR binding, suggestive of a role in AR recruitment to these loci. Loss of the SWI/SNF and CHD1 chromatin remodelers increased chromatin accessibility and AR binding, respectively, consistent with their suggested roles as tumour suppressors. Whilst AR and c-Myc co-localised, it remains unknown whether c-Myc might promote or repress AR binding at tRNA genes. The validation of AR binding events in PCa patient biopsies, in addition to confirmation in various LNCaP datasets, provides evidence of the relevance to disease of AR binding at tRNA genes.

5. Supplementary

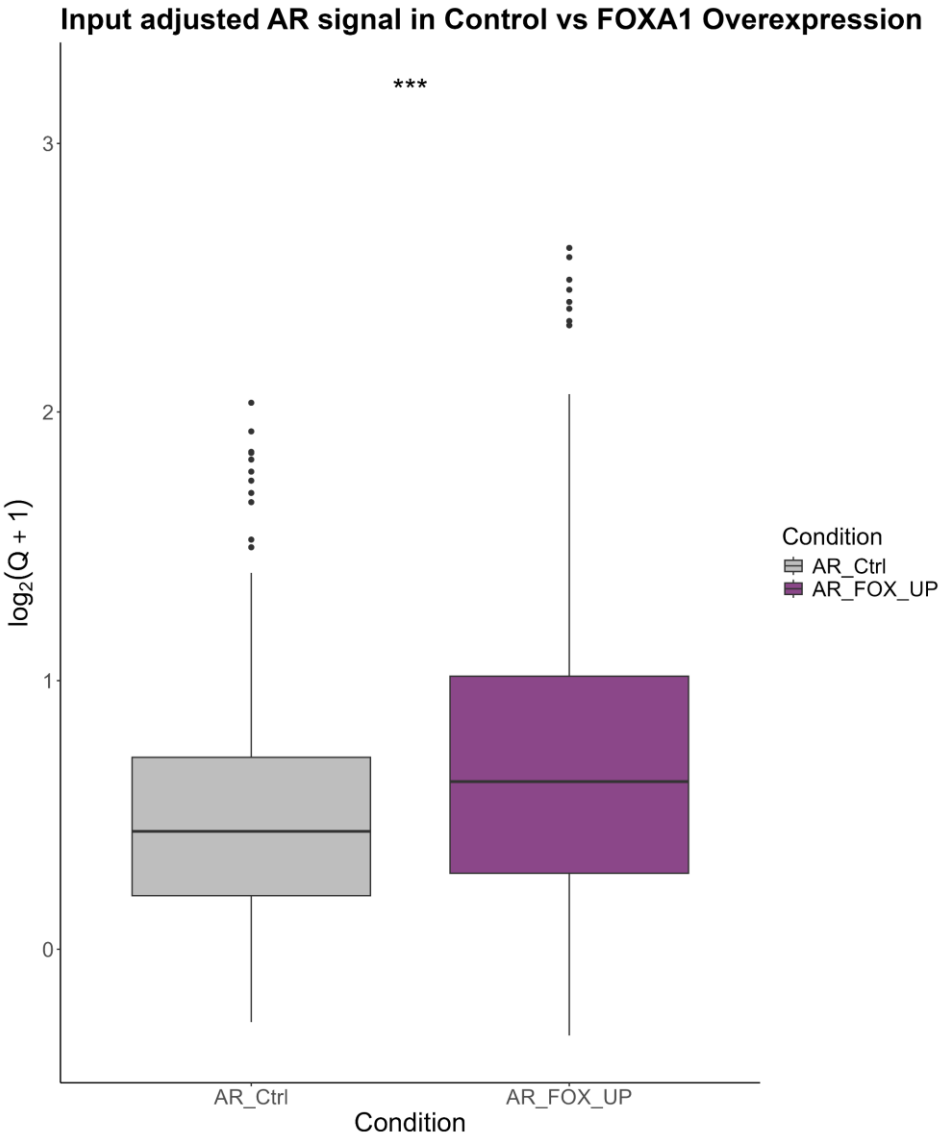


Figure S1. Input adjusted AR Q values in Control versus FOXA1 overexpression. Global AR Q values in the control (grey) compared to the FOXA1 overexpression condition (purple).

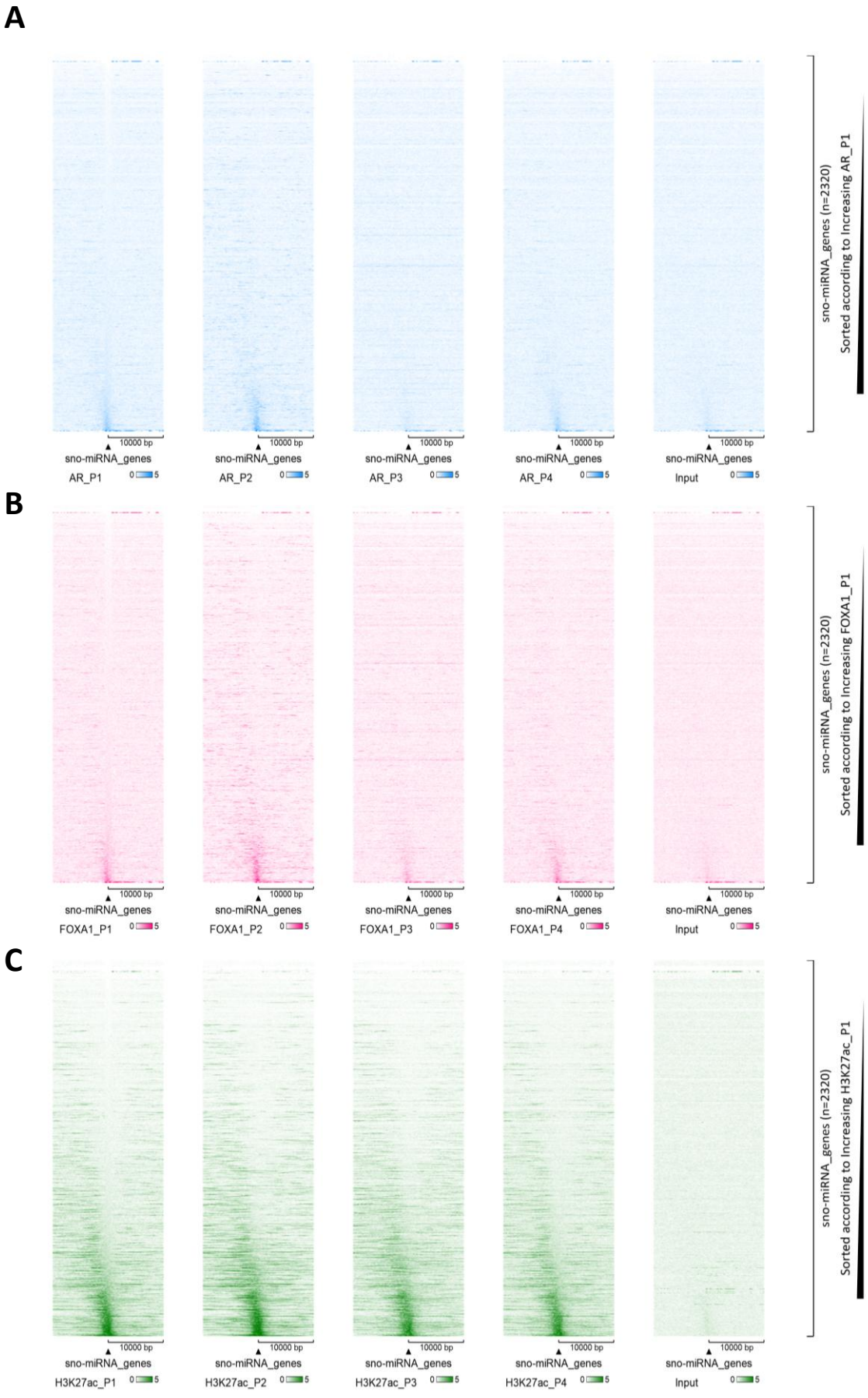

A
10000 bp
sno-miRNA_genes
AR_P1
0 5
AR_P2
AR_P3
AR_P4
Input
sno-miRNA_genes (n=2320)
Sorted according to Increasing AR_P1
B
FOXA1_P1
FOXA1_P2
FOXA1_P3
FOXA1_P4
sno-miRNA_genes (n=2320)
Sorted according to Increasing FOXA1_P1
C
H3K27ac_P1
H3K27ac_P2
H3K27ac_P3
H3K27ac_P4
sno-miRNA_genes (n=2320)
Sorted according to Increasing H3K27ac_P1

Figure S2. Heatmaps at sno-miRNA genes in patient biopsy data. (**A**) AR heatmaps (blue), sorted by increasing signal intensity in patient 1, at sno-miRNA genes in patients 1-4 and pooled input. (**B**) FOXA1 heatmaps (pink), sorted by increasing signal intensity in patient 1, at sno-miRNA genes in patients 1-4 and pooled input. (**C**) H3K27ac heatmaps (green), sorted by increasing signal intensity in patient 1, at sno-miRNA genes in patients 1-4 and pooled input.

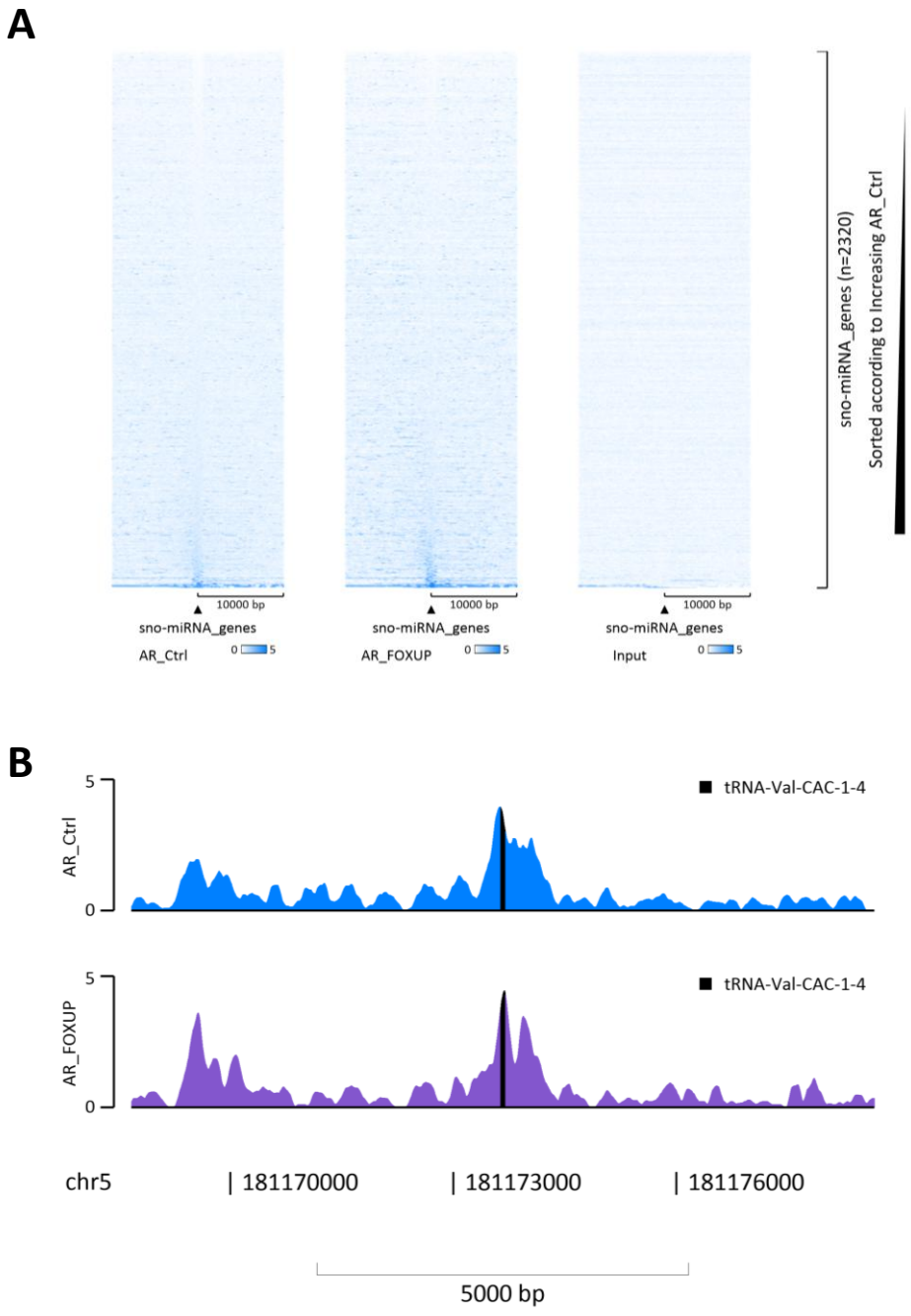


Figure S3. AR binding at sno-miRNA and tRNA genes in control and FOXA1 overexpression. (**A**) Heatmaps of AR binding, sorted by increasing AR signal intensity in the control. (**B**) Representative filltrack of minimal changes in AR binding intensity at tRNA-Val-CAC-1-4 between conditions.

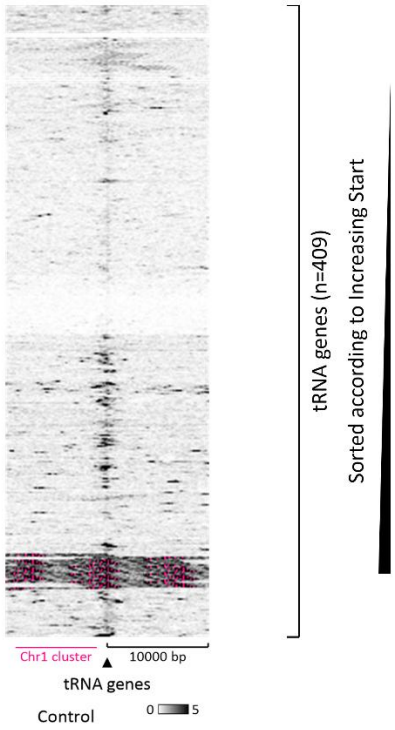


Figure S4. Highlighted chromosome 1 cluster of highly accessible tRNA genes. ATAC-seq signal sorted by genomic start position. Chromosome 1 cluster highlighted in pink.

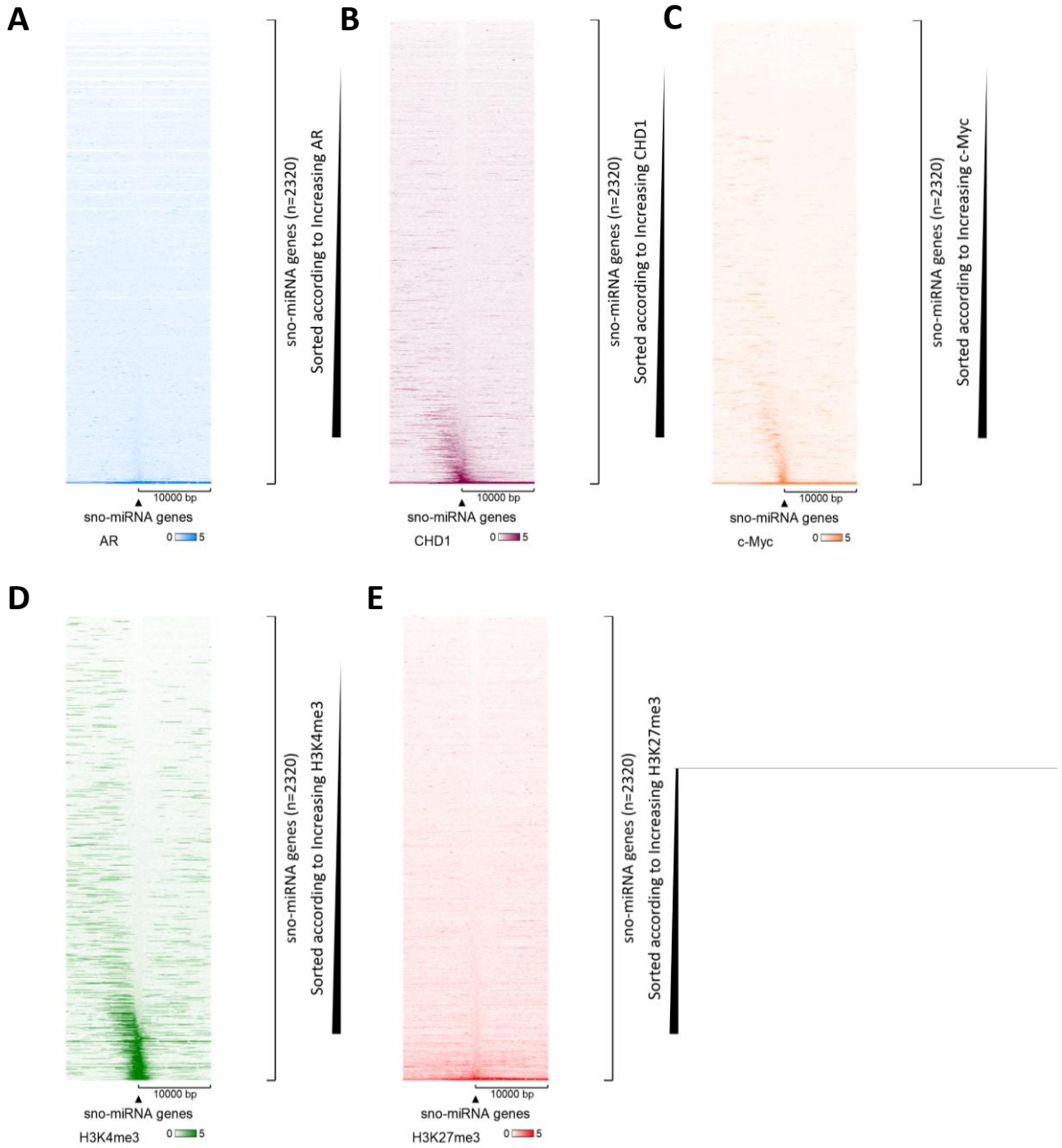


Figure S5. Heatmaps of AR, CHD1, c-Myc, H3K4me3 and H3K27me3 signal at sno-miRNA genes. (**A**) AR (**B**) CHD1 (**C**) c-Myc (**D**) H3K4me3 (**E**) H3K27me3.

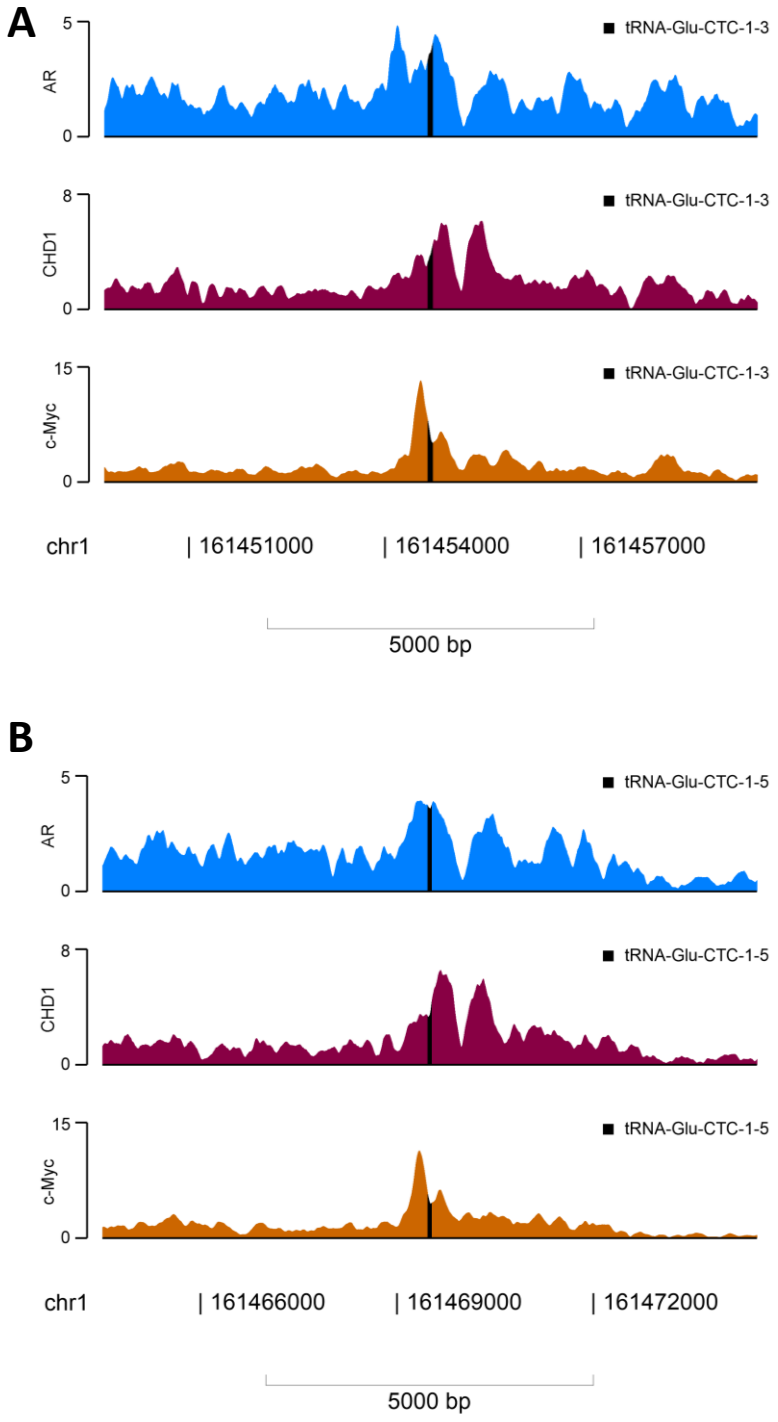


Figure S6. Representative filltracks of AR, CHD1 and c-Myc co-localisation at tRNA genes. (**A**) tRNA-Glu-CTC-1-3 (**B**) tRNA-Glu-CTC-1-5.

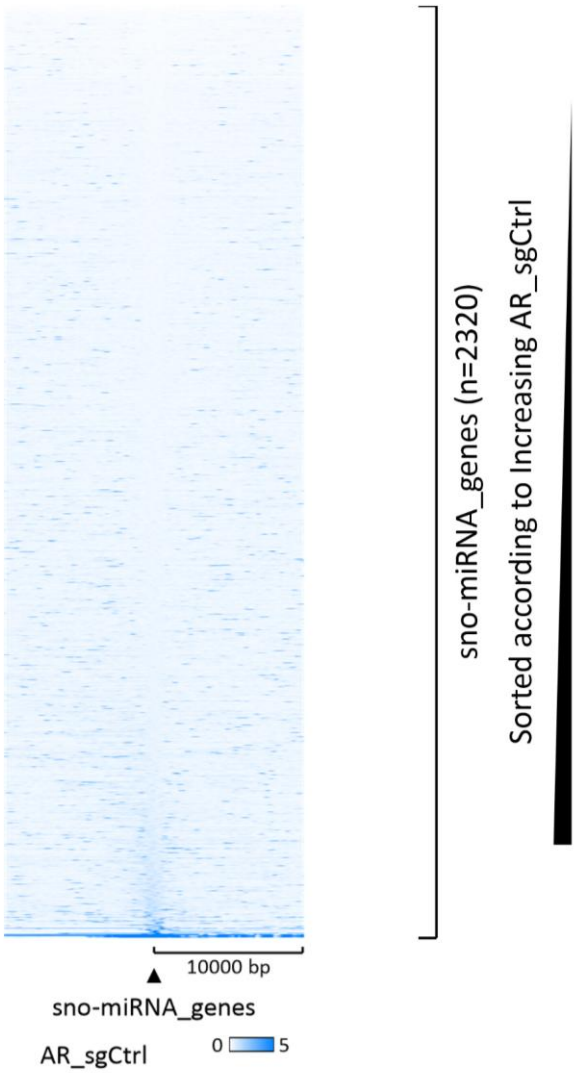


Figure S7. Heatmap of AR signal at sno-miRNA genes in the control condition. AR signal sorted by increasing signal intensity.

Table S1. Genome alignment statistics

Accession	Sample metadata	Total raw reads	Total mapped reads (%)	Unique mapped reads (%)	Multimapped reads (%)
SRR7207089	Input	21,130,210	99.90	83.23	16.68
SRR7207011	AR Patient 1	23,423,743	99.90	83.64	16.26
SRR7207017	AR Patient 2	21,498,068	99.91	83.94	15.97
SRR7207023	AR Patient 3	17,831,828	99.90	82.77	17.13
SRR7207030	AR Patient 4	22,257,334	99.90	83.56	16.34
SRR7207013	FOXA1 Patient 1	19,988,567	99.90	83.33	16.57
SRR7207019	FOXA1 Patient 2	20,417,070	99.91	84.98	14.93
SRR7207026	FOXA1 Patient 3	18,694,999	99.89	82.91	16.98
SRR7207032	FOXA1 Patient 4	19,485,854	99.90	83.48	16.42
SRR7207015	H3K27ac Patient 1	22,324,970	99.91	85.83	14.08
SRR7207021	H3K27ac Patient 2	23,241,929	99.91	86.33	13.59
SRR7207028	H3K27ac Patient 3	22,617,547	99.91	85.55	14.36
SRR7207034	H3K27ac Patient 4	26,136,062	99.91	85.05	14.85
SRR7207086	Input	21,424,614	97.30	81.82	15.48
SRR7206995	AR Control	21,464,624	99.93	84.65	15.28
SRR7206997	FOXA1 Control	17,455,117	99.91	84.58	15.33
SRR7207001	H3K27ac Control	22,585,398	99.91	84.83	15.09
SRR7206999	H3K4me3 Control	29,014,312	99.93	84.56	15.37

Table S1. Genome alignment statistics

Accession	Sample metadata	Total raw reads	Total mapped reads (%)	Unique mapped reads (%)	Multimapped reads (%)
SRR14159840	AR Control	70,702,678	79.57	65.16	14.46
SRR14159842	FOXA1 Control	64,464,591	83.30	68.92	14.46
SRR14159846	H3K27ac Control	55,047,935	85.91	74.57	11.34
SRR7545370	AR Control	38,243,360	96.47	65.87	30.51
SRR7545371	AR Control	28,050,451	95.78	64.76	31.02
SRR7545372	CHD1 Control	31,289,414	96.25	62.98	33.28
SRR7545373	CHD1 Control	23,399,221	96.63	66.59	30.04
SRR7545374	c-Myc Control	60,854,806	95.95	66.48	29.47
SRR7545375	c-Myc Control	27,202,462	92.89	45.38	47.51
SRR7545376	H3K4me3 Control	48,131,633	97.76	76.30	21.45
SRR7545377	H3K4me3 Control	27,455,815	97.32	70.18	27.14
SRR7545378	H3K27me3 Control	33,818,024	97.96	68.29	29.67
SRR7545379	H3K27me3 Control	14,541,489	92.94	53.58	39.36
SRR7545380	AR sgCHD1	46,976,421	98.37	71.81	26.56
SRR7545381	AR sgCHD1	45,922,547	98.46	75.50	22.96
SRR7545382	AR sgCtrl	103,018,660	98.93	74.16	24.78
SRR7545383	AR sgCtrl	107,848,387	99.89	74.99	23.90
ERR307344	Input DNA	33,836,466	99.20	71.31	27.89

Table S1. Genome alignment statistics

Accession	Sample metadata	Total raw reads	Total mapped reads (%)	Unique mapped reads (%)	Multimapped reads (%)
ERR307350	AR GFP (control)	32,438,496	97.60	67.30	30.30
ERR307345	AR GFP (control)	40,293,069	97.39	65.41	31.98
ERR307347	FOXA1 overexpression	30,230,792	96.89	66.20	30.69
ERR307352	FOXA1 overexpression	37,092,036	96.41	64.86	31.55
SRR14156915	ATAC-seq Control	40,975,563	73.46	42.00	31.46
SRR14156917	ATAC-seq sgSMARCA2	35,326,725	78.29	44.32	33.97
SRR14156918	ATAC-seq sgSMARCA2	39,111,947	77.92	44.96	32.96
SRR14156919	ATAC-seq sgSMARCA4	37,513,557	79.05	40.89	38.16
SRR14156920	ATAC-seq sgSMARCA4	33,609,471	78.17	40.92	37.25
SRR14156921	ATAC-seq SMARCA2 and SMARCA4 inhibited	31,463,378	78.78	44.09	34.69
SRR14156922	ATAC-seq SMARCA2 and SMARCA4 inhibited	34,362,370	79.08	44.88	34.20

Table S2. Clinical grading of prostate adenocarcinoma patient biopsies, adapted from Singh et al. [72]

Patient	Gleason score	AR status
P1	9	AR+
P2	7	AR+
P3	10	AR+
P4	9	AR+

6. References

1. Filho, A. M. *et al.* The GLOBOCAN 2022 cancer estimates: Data sources, methods, and a snapshot of the cancer burden worldwide. *Int. J. Cancer* **156**, 1336–1346 (2025).

2. Pernar, C. H., Ebot, E. M., Wilson, K. M. & Mucci, L. A. The Epidemiology of Prostate Cancer. *Cold Spring Harb. Perspect. Med.* **8**, (2018).

3. Lloyd, T. *et al.* Lifetime risk of being diagnosed with, or dying from, prostate cancer by major ethnic group in England 2008-2010. *BMC Med.* **13**, 171 (2015).

4. Sekhoacha, M. *et al.* Prostate Cancer Review: Genetics, Diagnosis, Treatment Options, and Alternative Approaches. *Molecules* **27**, (2022).

5. Huggins, C. & Hodges, C. V. 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.* **167**, 948–51; discussion 952 (2002).

6. Almeeri, M. N. E., Awies, M. & Constantinou, C. Prostate cancer, pathophysiology and recent developments in management: A narrative review. *Curr. Oncol. Rep.* **26**, 1511–1519 (2024).

7. Aurilio, G. *et al.* Androgen Receptor Signaling Pathway in Prostate Cancer: From Genetics to Clinical Applications. *Cells* **9**, (2020).

8. Weikum, E. R., Liu, X. & Ortlund, E. A. The nuclear receptor superfamily: A structural perspective: The Nuclear Receptor Superfamily. *Protein Sci.* **27**, 1876–1892 (2018).

9. Culig, Z. & Santer, F. R. Androgen receptor signaling in prostate cancer. *Cancer Metastasis Rev.* **33**, 413–427 (2014).

10. Yavuz, K. & Lack, N. A. Coregulators determine androgen receptor activity in prostate cancer. *Biosci. Rep.* **45**, (2025).

11. Hannenhalli, S. & Kaestner, K. H. The evolution of Fox genes and their role in development and disease. *Nat. Rev. Genet.* **10**, 233–240 (2009).

12. Lupien, M. *et al.* FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell* **132**, 958–970 (2008).

13. Utley, R. T. *et al.* Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. *Nature* **394**, 498–502 (1998).

14. Gao, N. *et al.* The role of hepatocyte nuclear factor-3 alpha (Forkhead Box A1) and androgen receptor in transcriptional regulation of prostatic genes. *Mol. Endocrinol.* **17**, 1484–1507 (2003).

15. Gao, N. *et al.* Forkhead box A1 regulates prostate ductal morphogenesis and promotes epithelial cell maturation. *Development* **132**, 3431–3443 (2005).

16. Sahu, B. *et al.* Dual role of FoxA1 in androgen receptor binding to chromatin, androgen signalling and prostate cancer. *EMBO J.* **30**, 3962–3976 (2011).

17. Yoon, C. E. *et al.* Genetic alterations of prostate cancer: in localized and metastatic prostate cancer. *BMC Urol.* **25**, 166 (2025).

18. Adams, E. J. *et al.* FOXA1 mutations alter pioneering activity, differentiation and prostate cancer phenotypes. *Nature* **571**, 408–412 (2019).

19. Montecino, M. *et al.* Nucleosome organization and targeting of SWI/SNF chromatin-remodeling complexes: contributions of the DNA sequence. *Biochem. Cell Biol.* **85**, 419–425 (2007).

20. Giles, K. A. *et al.* Integrated epigenomic analysis stratifies chromatin remodellers into distinct functional groups. *Epigenetics Chromatin* **12**, 12 (2019).

21. Wang, W. *et al.* Purification and biochemical heterogeneity of the mammalian SWI-SNF complex. *EMBO J.* **15**, 5370–5382 (1996).

22. Mittal, P. & Roberts, C. W. M. The SWI/SNF complex in cancer - biology, biomarkers and therapy. *Nat. Rev. Clin. Oncol.* **17**, 435–448 (2020).

23. Alver, B. H. *et al.* The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. *Nat. Commun.* **8**, 14648 (2017).

24. Kadoch, C. *et al.* Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* **45**, 592–601 (2013).

25. van de Wijngaart, D. J. *et al.* Functional screening of FxxLF-like peptide motifs identifies SMARCD1/BAF60a as an androgen receptor cofactor that modulates TMPRSS2 expression. *Mol. Endocrinol.* **23**, 1776–1786 (2009).

26. Launonen, K.-M. *et al.* Chromatin-directed proteomics-identified network of endogenous androgen receptor in prostate cancer cells. *Oncogene* **40**, 4567–4579 (2021).

27. Link, K. A. *et al.* BAF57 governs androgen receptor action and androgen-dependent proliferation through SWI/SNF. *Mol. Cell. Biol.* **25**, 2200–2215 (2005).

28. Ertl, I. E. *et al.* The SMARCD Family of SWI/SNF Accessory Proteins Is Involved in the Transcriptional Regulation of Androgen Receptor-Driven Genes and Plays a Role in Various Essential Processes of Prostate Cancer. *Cells* **12**, (2022).

29. Heebøll, S. *et al.* SMARCC1 expression is upregulated in prostate cancer and positively correlated with tumour recurrence and dedifferentiation. *Histol. Histopathol.* **23**, 1069–1076 (2008).

30. Sun, A. *et al.* Aberrant expression of SWI/SNF catalytic subunits BRG1/BRM is associated with tumor development and increased invasiveness in prostate cancers. *Prostate* **67**, 203–213 (2007).

31. Marfella, C. G. A. & Imbalzano, A. N. The Chd family of chromatin remodelers. *Mutat. Res.* **618**, 30–40 (2007).

32. Li, H., Gigi, L. & Zhao, D. CHD1, a multifaceted epigenetic remodeler in prostate cancer. *Front. Oncol.* **13**, 1123362 (2023).

33. Zhao, D. *et al.* Synthetic essentiality of chromatin remodelling factor CHD1 in PTEN-deficient cancer. *Nature* **542**, 484–488 (2017).

34. Augello, M. A. *et al.* CHD1 loss alters AR binding at lineage-specific enhancers and modulates distinct transcriptional programs to drive prostate tumorigenesis. *Cancer Cell* **35**, 603-617.e8 (2019).

35. Pomerantz, M. M. *et al.* The androgen receptor cistrome is extensively reprogrammed in human prostate tumorigenesis. *Nat. Genet.* **47**, 1346–1351 (2015).

36. Thornlow, B. P. *et al.* Predicting transfer RNA gene activity from sequence and genome context. *Genome Res.* **30**, 85–94 (2020).

37. Geslain, R. & Pan, T. Functional analysis of human tRNA isodecoders. *J. Mol. Biol.* **396**, 821–831 (2010).

38. Sizer, R. E., Butterfield, S. P., Hancocks, L. A., Gato De Sousa, L. & White, R. J. Selective Occupation by E2F and RB of Loci Expressed by RNA Polymerase III. *Cancers* **16**, (2024).

39. Hughes, L. A. *et al.* Copy number variation in tRNA isodecoder genes impairs mammalian development and balanced translation. *Nat. Commun.* **14**, 2210 (2023).

40. White, R. J. RNA Polymerase I and RNA Polymerase III in Eukaryotes. 159–161 (2013).

41. Arimbasseri, A. G. & Maraia, R. J. RNA Polymerase III Advances: Structural and tRNA Functional Views. *Trends Biochem. Sci.* **41**, 546–559 (2016).

42. Canella, D., Praz, V., Reina, J. H., Cousin, P. & Hernandez, N. Defining the RNA polymerase III transcriptome: Genome-wide localization of the RNA polymerase III transcription machinery in human cells. *Genome Res.* **20**, 710–721 (2010).

43. Kor, S. D. *et al.* RNA Pol III promoters-key players in precisely targeted plant genome editing. *Front. Genet.* **13**, 989199 (2022).

44. White, R. J. RNA polymerases I and III, growth control and cancer. *Nat. Rev. Mol. Cell Biol.* **6**, 69–78 (2005).

45. Gomez-Roman, N., Grandori, C., Eisenman, R. N. & White, R. J. Direct activation of RNA polymerase III transcription by c-Myc. *Nature* **421**, 290–294 (2003).

46. Kenneth, N. S. *et al.* TRRAP and GCN5 are used by c-Myc to activate RNA polymerase III transcription. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 14917–14922 (2007).

47. Winter, A. G. *et al.* RNA polymerase III transcription factor TFIIIC2 is overexpressed in ovarian tumors. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 12619–12624 (2000).

48. Pavon-Eternod, M. *et al.* tRNA over-expression in breast cancer and functional consequences. *Nucleic Acids Res.* **37**, 7268–7280 (2009).

49. Zhou, Y., Goodenbour, J. M., Godley, L. A., Wickrema, A. & Pan, T. High levels of tRNA abundance and alteration of tRNA charging by bortezomib in multiple myeloma. *Biochem. Biophys. Res. Commun.* **385**, 160–164 (2009).

50. Krishnan, P. *et al.* Genome-wide profiling of transfer RNAs and their role as novel prognostic markers for breast cancer. *Sci. Rep.* **6**, 32843 (2016).

51. Zhang, Z. *et al.* Global analysis of tRNA and translation factor expression reveals a dynamic landscape of translational regulation in human cancers. *Commun Biol* **1**, 234 (2018).

52. Lee, Y. S., Shibata, Y., Malhotra, A. & Dutta, A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev.* **23**, 2639–2649 (2009).

53. Wang, L. *et al.* Clinical Significance of High Expression of tRF-Glu-TTC-2 in Prostate Carcinoma and its Effect on Growth. *Am. J. Mens. Health* **16**, 15579883221135970 (2022).

54. Yang, C., Lee, M., Song, G. & Lim, W. tRNALys-Derived Fragment Alleviates Cisplatin-Induced Apoptosis in Prostate Cancer Cells. *Pharmaceutics* **13**, (2021).

55. Olvedy, M. *et al.* A comprehensive repertoire of tRNA-derived fragments in prostate cancer. *Oncotarget* **7**, 24766–24777 (2016).

56. Yao, P. *et al.* Cellular signaling of amino acid metabolism in prostate cancer. *Int. J. Mol. Sci.* **26**, 776 (2025).

57. Hah, N. *et al.* A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell* **145**, 622–634 (2011).

58. Malcolm, J. R., Leese, N. K., Lamond-Warner, P. I., Brackenbury, W. J. & White, R. J. Widespread association of ERα with RMRP and tRNA genes in MCF-7 cells and breast cancers. *Gene* **821**, 146280 (2022).

59. Finlay-Schultz, J. *et al.* Breast Cancer Suppression by Progesterone Receptors Is Mediated by Their Modulation of Estrogen Receptors and RNA Polymerase III. *Cancer Res.* **77**, 4934–4946 (2017).

60. Chung, L. W. LNCaP human prostate cancer progression model. *Urol. Oncol.* **2**, 126–128 (1996).

61. Xiao, L. *et al.* Targeting SWI/SNF ATPases in enhancer-addicted prostate cancer. *Nature* **601**, 434–439 (2022).

62. Galaxy Community. The Galaxy platform for accessible, reproducible, and collaborative data analyses: 2024 update. *Nucleic Acids Res.* **52**, W83–W94 (2024).

63. *Sra-Tools: SRA Tools*. (Github).

64. Petit, R. A., III. *Fastq-Dl: Download FASTQ Files from SRA or ENA Repositories*. (Github).

65. Leinonen, R., Sugawara, H., Shumway, M. & International Nucleotide Sequence Database Collaboration. The sequence read archive. *Nucleic Acids Res.* **39**, D19-21 (2011).

66. Andrews, S. & Others. FastQC: a quality control tool for high throughput sequence data. Preprint at (2010).

67. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32**, 3047–3048 (2016).

68. Babraham Bioinformatics - Trim Galore! https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/.

69. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

70. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10**, giab008 (2021).

71. Perez, G. *et al.* The UCSC Genome Browser database: 2025 update. *Nucleic Acids Res.* **53**, D1243–D1249 (2025).

72. Singh, A. A. *et al.* Optimized ChIP-seq method facilitates transcription factor profiling in human tumors. *Life Sci Alliance* **2**, e201800115 (2019).

73. Robinson, J. L. L. *et al.* Elevated levels of FOXA1 facilitate androgen receptor chromatin binding resulting in a CRPC-like phenotype. *Oncogene* **33**, 5666–5674 (2014).

74. Lerdrup, M., Johansen, J. V., Agrawal-Singh, S. & Hansen, K. An interactive environment for agile analysis and visualization of ChIP-sequencing data. *Nat. Struct. Mol. Biol.* **23**, 349–357 (2016).

75. Wickham, H. & Sievert, C. *Ggplot2: Elegant Graphics for Data Analysis*. vol. 10 (springer New York, 2009).

76. Wickham, H. *et al.* Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).

77. Zhang, Y. *et al.* Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

78. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

79. Barfeld, S. J. *et al.* C-Myc antagonises the transcriptional activity of the androgen receptor in prostate cancer affecting key gene networks. *EBioMedicine* **18**, 83–93 (2017).

80. Qiu, X. *et al.* MYC drives aggressive prostate cancer by disrupting transcriptional pause release at androgen receptor targets. *Nat. Commun.* **13**, 2559 (2022).

81. Crowell, P. D. *et al.* MYC is a regulator of androgen receptor inhibition-induced metabolic requirements in prostate cancer. *Cell Rep.* **42**, 113221 (2023).

82. Teo, M. Y., Rathkopf, D. E. & Kantoff, P. Treatment of Advanced Prostate Cancer. *Annu. Rev. Med.* **70**, 479–499 (2019).

83. Pavon-Eternod, M., Gomes, S., Rosner, M. R. & Pan, T. Overexpression of initiator methionine tRNA leads to global reprogramming of tRNA expression and increased proliferation in human epithelial cells. *RNA* **19**, 461–466 (2013).

84. Goodarzi, H. *et al.* Modulated Expression of Specific tRNAs Drives Gene Expression and Cancer Progression. *Cell* **165**, 1416–1427 (2016).

85. Grasso, C. S. *et al.* The mutational landscape of lethal castration-resistant prostate cancer. *Nature* **487**, 239–243 (2012).

86. Gerhardt, J. *et al.* FOXA1 promotes tumor progression in prostate cancer and represents a novel hallmark of castration-resistant prostate cancer. *Am. J. Pathol.* **180**, 848–861 (2012).

87. Dong, H.-Y. *et al.* FOXA1 in prostate cancer. *Asian J. Androl.* **25**, 287–295 (2023).

88. De Lara, S. *et al.* The prognostic relevance of FOXA1 and Nestin expression in breast cancer metastases: a retrospective study of 164 cases during a 10-year period (2004-2014). *BMC Cancer* **19**, 187 (2019).

89. Shah, N. & Brown, M. The Sly Oncogene: FOXA1 Mutations in Prostate Cancer. *Cancer cell* vol. 36 119–121 (2019).

90. Teng, M., Zhou, S., Cai, C., Lupien, M. & He, H. H. Pioneer of prostate cancer: past, present and the future of FOXA1. *Protein Cell* **12**, 29–38 (2021).

91. Wang, D. *et al.* Reprogramming transcription by distinct classes of enhancers functionally defined by eRNA. *Nature* **474**, 390–394 (2011).

92. Donze, D. & Kamakaka, R. T. RNA polymerase III and RNA polymerase II promoter complexes are heterochromatin barriers in Saccharomyces cerevisiae. *EMBO J.* **20**, 520–531 (2001).

93. Ebersole, T. *et al.* tRNA genes protect a reporter gene from epigenetic silencing in mouse cells. *Cell Cycle* **10**, 2779–2791 (2011).

94. Sizer, R. E. *et al.* TFIIIC-based chromatin insulators through eukaryotic evolution. *Gene* **835**, 146533 (2022).

95. Lu, P. & Roberts, C. W. M. The SWI/SNF tumor suppressor complex: Regulation of promoter nucleosomes and beyond: Regulation of promoter nucleosomes and beyond. *Nucleus* **4**, 374–378 (2013).

96. Martin, B. J. E. *et al.* Global identification of SWI/SNF targets reveals compensation by EP400. *Cell* **186**, 5290-5307.e26 (2023).

97. Gurel, B. *et al.* Nuclear MYC protein overexpression is an early alteration in human prostate carcinogenesis. *Mod. Pathol.* **21**, 1156–1167 (2008).

98. Guo, H. *et al.* Androgen receptor and MYC equilibration centralizes on developmental super-enhancer. *Nat. Commun.* **12**, 7308 (2021).

99. Bai, S. *et al.* A positive role of c-Myc in regulating androgen receptor and its splice variants in prostate cancer. *Oncogene* **38**, 4977–4989 (2019).

100. Gao, L. *et al.* Androgen receptor promotes ligand-independent prostate cancer progression through c-Myc upregulation. *PLoS One* **8**, e63563 (2013).

101. Cheng, S. W. *et al.* c-MYC interacts with INI1/hSNF5 and requires the SWI/SNF complex for transactivation function. *Nat. Genet.* **22**, 102–105 (1999).

102. Woodley, C. M. *et al.* Multiple interactions of the oncoprotein transcription factor MYC with the SWI/SNF chromatin remodeler. *Oncogene* **40**, 3593–3609 (2021).

103. Zhang, Z. *et al.* Loss of CHD1 promotes heterogeneous mechanisms of resistance to AR-targeted therapy via chromatin dysregulation. *Cancer Cell* **37**, 584-598.e11 (2020).

104. Wang, Y. *et al.* tRNA-derived small RNAs: Mechanisms and potential roles in cancers. *Genes Dis* **9**, 1431–1442 (2022).