ARTICLE

Altered regulation of DPF3, a member of the SWI/SNF complexes, underlies the 14q24 renal cancer susceptibility locus

Leandro M. Colli,[1,2] Lea Jessop,[1] Timothy A. Myers,[1] Sabrina Y. Camp,[1] Mitchell J. Machiela,[1] Jiyeon Choi,[1] Renato Cunha,[2,3] Olusegun Onabajo,[1] Grace C. Mills,[1] Virginia Schmid,[4] Seth A. Brodie,[1] Olivier Delattre,[5] David R. Mole,[4] Mark P. Purdue,[1] Kai Yu,[1] Kevin M. Brown,[1] and Stephen J. Chanock[1,*]

Summary

Our study investigated the underlying mechanism for the 14q24 renal cell carcinoma (RCC) susceptibility risk locus identified by a genome-wide association study (GWAS). The sentinel single-nucleotide polymorphism (SNP), rs4903064, at 14q24 confers an allele-specific effect on expression of the double PHD fingers 3 (*DPF3*) of the BAF SWI/SNF complex as assessed by massively parallel reporter assay, confirmatory luciferase assays, and eQTL analyses. Overexpression of *DPF3* in renal cell lines increases growth rates and alters chromatin accessibility and gene expression, leading to inhibition of apoptosis and activation of oncogenic pathways. siRNA interference of multiple DPF3-deregulated genes reduces growth. Our results indicate that germline variation in *DPF3*, a component of the BAF complex, part of the SWI/SNF complexes, can lead to reduced apoptosis and activation of the STAT3 pathway, both critical in RCC carcinogenesis. In addition, we show that altered *DPF3* expression in the 14q24 RCC locus could influence the effectiveness of immunotherapy treatment for RCC by regulating tumor cytokine secretion and immune cell activation.

Introduction

Annually, more than 300,000 new cases of renal cell carcinoma (RCC [MIM: 144700]) are diagnosed worldwide and approximately one-fifth are in the United States.[1] Clear cell renal cell carcinoma (ccRCC), the most common RCC subtype, is characterized by the somatic bi-allelic loss of *VHL* (MIM: 608537),[2] which increases the pleotropic effects of the hypoxia-inducible factor (*HIF*),[3,4] leading to dysregulated cellular metabolism and enhanced cell survival.[5] Although loss of *VHL* is an early ccRCC event, both mouse models[6] and tumor evolution analyses[7] indicate additional gene mutations and epigenetic events are required. Landscape analyses of somatic alterations in RCC have revealed important features of renal cell carcinogenesis,[8–12] such as driver mutations in chromatin modification genes, including *PBRM1* (MIM: 606083),[13] *BAP1* (MIM: 603089),[14] and *SETD2* (MIM: 612778).[14] *PBRM1*, a component of the SWI/SNF complexes, is mutated in 40% of ccRCC[13] and, when mutated, complements the effects of *VHL* deficiency on hypoxia as well as mTOR pathway deregulation,[15] suggesting a key role of SWI/SNF components in ccRCC development.

Genome-wide association studies (GWASs) have successfully identified common cancer susceptibility alleles;[16] to date, over 1,000 independent loci have been established for more than 30 cancers.[17] For RCC, thirteen RCC suscep-

tibility regions have achieved statistical significance at the genome-wide threshold: 1p32, 2p21, 2q22, 3p22, 3q26, 8p21, 8q24, 10q24, 11q13, 11q22, 12p12, 12q24, and 14q24.[18–23] Investigation of three RCC risk loci has revealed new insights into RCC biology. At 11q13, HIF binding to the common risk-associated haplotype increases *CCND1* (MIM: 168461) expression, leading to cell cycle deregulation.[24] At the 12p12 locus, common risk alleles increase binding of c-MYC and HIF, thus enhancing expression of *BHLHE41* (MIM: 606200)[25,26] and leading to increased tumor growth for the latter in a mouse model.[25] At 8q24, a multi-cancer risk region, the risk SNP, rs35252396, creates a HIF-binding motif leading to increased *MYC* (MIM: 190080) and *PVT1* (MIM: 165140) expression.[27]

The present study reports on the germline allele-specific effects of the 14q24 locus sentinel SNP, rs4903064, located in an enhancer of double PHD fingers 3 (*DPF3*) (MIM: 601672), a member of one of the SWI/SNF complexes. DPF3 is an important regulator of gene expression that acts through nucleosome repositioning[28] and has two common isoforms: *DPF3a*, which has a truncated PHD finger, and *DPF3b*, which has two PHD fingers[29] that can bind to acetylated histones (H3K14ac) that influence regulatory networks.[30] Somatic alteration in components of SWI/SNF, such as common *PBRM1* mutations, can be an important molecular event in ccRCC.[13,15] SWI/SNF is

comprised of at least 29 genes, arranged in three complexes: canonical BAF, polybromo-associated BAF (PBAF), and non-canonical BAF (ncBAF).[31,32] Overall, SWI/SNF genes have been reported to be somatically mutated in 20% of surveyed tumors (range of 6% to 75% depending on cancer type[28,33]); notably, somatically mutated SWI/SNF members can act as either oncogenes or tumor suppressor genes.[34] Herein, we identified germline allele-specific effects of a component of one of the SWI/SNF complexes, more specifically the BAF complex, pinpointing the role of *DPF3* on RCC oncogenesis and elucidating the biological mechanism for the 14q24 RCC risk region.

Material and methods

Screen with massively parallel reporter assay (MPRA)

As a part of a larger MPRA screening assay, we evaluated 39 variants across the 14q24 region, one of nineteen regions we evaluated via MPRA in which the lead variant had an association p value less than 10^{-6} in our published RCC GWAS.[22] For each region, we included correlated variants in linkage disequilibrium (LD) with the lead signal that were either $r^2 > 0.2$ (1000G EUR) or $D' > 0.5$ for minor allele frequency (MAF) < 0.05. We performed subsequent filtering of these correlated variants to assay the subset of variants with existing functional evidence *in silico*, namely a RegulomeDB[35] score ≤ 4 or reported as significant in a previously published unsupervised genome-wide study integrating several functional annotations into one measure of functional importance *in silico*.[36] Those variants were incorporated in the MPRA library design. In addition, the ten SNPs with the highest degree of LD with the lead SNP for each locus irrespective of this *in silico* analysis were also included, and for 14q24, five of these overlapped with the *in silico* screen above. In total, 774 variants were selected from 20 regions with the following oligonucleotide design: 201 bp sequence for each variant in which 16 bp were for vector cloning, 72 bp flanking each side of the SNP, 12 bp for insertion of the reporter open reading frame (ORF), a unique 10 bp barcode, and 17 bp for 3′ vector cloning. For each allele of each SNP, ten sequences in the forward direction and ten sequences in the reverse direction were generated; each had a unique random barcode as per MPRA protocol.[37,38] Ten unique forward and ten unique reverse scrambled sequences per SNP served as controls. We used the mean value of the scrambles for normalization, therefore reducing the chance of creating a false positive signal. In total, 47,460 sequences were synthesized by Agilent Technologies and the MPRA library construction as per published protocols.[38,39]

Transfections were performed in one renal cell cancer cell line (ACHN) and a non-cancer renal cell line (HEK293T, an embryonic renal cell line) after plating to 60%–70% confluence in 10 cm dishes via Lipofectamine 3000 (Thermo Fisher): 14,000 ng of MPRA vector/dish, P3000 28 μL/dish, and LF3000 32 μL/dish. We transfected one extra dish with GFP in parallel to confirm transfection efficiency (>70%). Cell culture media was changed 5 h after the transfection, and RNA was extracted 48 h later. We extracted RNA with the RNeasy Mini QIAcube Kit (QIAGEN) on the QIAcube with DNase treatment to remove cell DNA and MPRA vectors. The DNase treatment efficiency was checked by PCR (see below) and no sample showed DNA contamination. The samples were Poly-A enriched with the Ambion Poly(A)Purist MAG Kit, and subsequently, cDNA was generated with Invitrogen Super-

Script III Reverse Transcriptase. PCR products were resolved on a 2% agarose gel, and the 0.25 kb band was excised and purified subsequently (QIAquick Gel Extraction Kit). On the basis of power calculation (R package "pwr"), at least five transfections (replications) provided 82.2% power for detecting alleles that explain 50% of the variance.

Because of the importance of the hypoxia pathway to RCC,[40] effects on candidate variants within RCC susceptibility loci were also evaluated under hypoxia conditions in which, after media change, the cells were incubated in 1% oxygen hypoxia with an incubator chamber (Stemcell Technologies). We used a HIF-1 immunoblot with protein extract (RIPA buffer) from an extra dish to confirm induction of hypoxia. We sequenced the MPRA libraries twice on an Illumina HiSeq 2500, which had a read length of 150 nt configured for paired-end sequencing, and from the FASTQ files, we mapped the first 10 base pairs to the designed barcode sequences and the next 10 base pairs to the MPRA vector to confirm barcode expression. The analytical approach was performed as per two published reports.[38,39] Briefly, the total number of barcodes was counted and normalized per 1,000,000 total reads, followed by counts of DNA vector sequence. Each barcode was normalized to the median of specific scramble controls. Each SNP was tested with logistic regression including allele, direction, cell type, and condition. p values were corrected by Bonferroni (774 comparisons).

Cell culture

Cell lines were grown in media supplemented with 10% FBS and maintained in a 37°C incubator with 5% CO_2. Because the reported RCC GWAS[22] included a combination of affected individuals with clear cell (∼90%) and papillary subtypes (<10%), the following cell lines were investigated: HK2 (ATCC CRL-2190), HEK293T (ATCC CRL-3216), ACHN (ATCC CRL-1611), SN12C, UO-31, and UOK-121 (kindly provided by Dr. W. Marston Linehan, NCI), which were grown in DMEM, and 786-O (ATCC CRL-1932), which was grown in RPMI. For hypoxia conditions, dishes were placed in an airtight chamber and flushed with 1% O_2 and 5% CO_2 daily and placed in a 37°C incubator. Via the MycoAlert Mycoplasma Detection Kit (Lonza), all cultures tested negative every 3 months. Cells were authenticated with the AmpFLSTR Identifiler PCR Amplification Kit (Thermo Fisher) and compared with ATCC reference.

RNA extraction

RNA extraction from cell lines was done with the RNeasy Plus Mini Kit (QIAGEN) on an automatic extraction system (QIAcube, QIAGEN). RNA was quantified and quality was checked with the Fragment Analyzer Systems (Agilent).

Rapid amplification of cDNA ends (RACE)

RACE was done with the SMARTer RACE 5′/3′ Kit (Takara Bio USA). Total RNA extracted from renal cancer cell lines (SN12C and UO-31) was used for constructing cDNA for 5′-RACE and 3′-RACE with kit primers and unique primers listed in Table S1. PCR products were purified from agarose gel via In-Fusion (Takara Bio USA) and cloned into the pRACE vector. Following transfection, individual colonies were picked for Sanger sequencing and compared to known DPF3 isoforms via BLAST.

Individual luciferase reporter assays

28 bp oligos for each tested SNP for each allele in both forward and reverse orientation (sequences are provided in Table S1; IDT DNA)

were cloned into the pGL4.23[luc2/minP] vector (Promega) via In-Fusion (Takara Bio) and verified by Sanger sequencing. Transfections were performed using Lipofectamine 3000 (Life Technologies) according to the manufacturer's instructions. After 48 h, lysates were collected and luciferase activity was measured via the Dual-Luciferase Reporter Assay System (Promega). Experiments were carried out in quadruplicate and 5–18 biological replicates were done for each SNP. Linear regression was used for evaluation of allele effect in a model that included cell line and replicate.

Eletromobility shift assays (EMSAs)

Nuclear extracts were prepared with the Nuclear Extract Kit (Active Motif). Complementary 30 bp biotinylated oligos (sequences in Table S1; IDT DNA) were annealed for creation of substrates for EMSA via the LightShift Chemiluminescent EMSA Kit (Thermo Fisher). Reactions contained 240 ng of biotinylated probe, 1× binding buffer, 15 mg nuclear extract, 0.5 μg poly dIdC, 100 nM MgCl2, and 2% glycerol. Competitive assays were performed by addition of 1- to 40-fold excess of non-biotinylated probes. Images were captured on the ChemiDoc Gel Imaging System (Bio-Rad).

Dox-inducible DPF3-stable cell lines

In order to evaluate the effects of altered DPF3 levels on cell phenotype and perform molecular studies, we created *DPF3a* (GenBank: NM_012074.5), *DPF3b* (GenBank: NM_001280542.3), and empty vector stable cell lines by using lentivirus to achieve DPF3 overexpression by doxycycline (dox) treatment. We designed and cloned *DPF3a* and *DPF3b* in a vector with IRES-GFP. Cells were transduced first with neomycin-resistant Tet transactivator lentivirus and selected with antibiotics. Tet expression was checked by immunoblot (TetR Monoclonal Antibody, Clone 9G9, catalog number 631132, Takara Bio USA). *DPF3a*, *DPF3b*, or empty vector were transduced in HK2, ACHN, 786-O, and UOK-121 cell lines. After antibiotic selection, cell lines were tested for expression via qPCR (PrimeTime Predesigned qPCR Assays, catalog number Hs.PT.58.19404383, IDT DNA), immunoblot (Anti-FLAG antibody, catalog number F1804-50UG, Sigma) and GFP by flow cytometry.

Real-time quantitative PCR (qPCR)

We evaluated gene expression by using qPCR with RNA extracted by the RNeasy Plus Mini Kit (QIAGEN) with the High-Capacity cDNA Reverse Transcription Kit (Thermo Fisher). qPCR reactions were done on a QuantStudio 7 (Thermo Fisher) in a 10 μL reaction: 5 μL of MasterMix (TaqMan Fast Advanced Master Mix, Thermo Fisher), 0.5 μL probe, 2.5 μL of cDNA diluted 1:5, and 2 μL of H_2O. Reactions were done in triplicate. Quantification was done via the $2^{-\Delta\Delta CT}$ method with *ACTB* (MIM: 102630) (Hs99999903_m1, Thermo Fisher), *PPIA* (MIM: 123840) (Hs99999904_m1, Thermo Fisher), and *GAPDH* (MIM: 138400) (Hs99999905_m1, Thermo Fisher) as endogenous controls. qPCR probes were purchased from IDT DNA: *DPF3* (Hs.PT.58.19404383), *CEMIP* (MIM: 608366) (Hs.PT.58.28305095), *IL23R* (MIM: 607562) (Hs.PT.58.2858675), and *STAT3* (MIM: 102582) (Hs.PT.58.3750282).

Small interfering RNA (siRNA) gene knockdown

SMARTpool ON-TARGETplus Human siRNAs for *DPF3* (L-012798-03-0005), *CEMIP* (L-022291-00-0005), *IL23R* (L-007976-00-0005), and N-TARGETplus non-targeting controls (D-001810-01-05) were obtained from Dharmacon. Transfections were performed with Lipofectamine RNAiMAX (Life Technologies) in which 5,000 cells were seeded in each well of a 96-well dish and, on

the next day, transfected with siRNA (10 μM/0.5 μL of siRNA and 0.44 μL of RNAiMAX per well). For 6-well dishes, 300,000 cells were seeded and transfected with 5 μL of siRNA and 10 μM and 11 μL of RNAiMAX the day after. Cells were collected after 48 h for RNA extraction. Control qPCRs and/or immunoblot were done for verification of gene knockdown by siRNA.

Cell growth assays

Cell growth assays were performed in real time with impedance-based technology (xCELLigence system, ACEA Biosciences). Stable inducible DPF3 cell lines were counted and plated (5,000 cells/well for all cell lines except ACHN, which was seeded at 10,000 cells/well) in an E-Plate 16, and when indicated after 4 h, doxycycline 100–1000 ng/μL or DPBS were added to individual wells in quadruplicate. Impedance was measured every 10 min for 3 to 5 days. Each experiment was repeated at least three times. Growth rate was calibrated for the point just before adding dox or PBS to the cell lines.

PBMC (peripheral blood mononuclear cell)-cytotoxicity assay

PBMC-cytotoxicity assays were performed with impedance-based technologies (xCELLigence system, ACEA Biosciences). Stable inducible DPF3 cell lines were treated with dox for 48 h while PBMCs from healthy normal human donors were stimulated with 25 μL/mL CD3/CD28/CD2 (ImmunoCult Human CD3/CD28/CD2 T Cell Activator, catalog number 10970, StemCell Technologies) and 100 UI/mL IL-2 (catalog number 10799068001, Sigma-Aldrich) for 48 h. Afterward, PBMCs were washed in PBS and 100,000 cells were added to DPF3-overexpressing cells. Impedance was measured every 10 min for 2 to 3 days. Experiments were repeated two to three times per PBMC donor and were recovered for flow cytometry analysis while media recovered for cytokine ELISA (enzyme-linked immunosorbent assay) quantification.

Flow cytometry

Apoptosis was assessed in renal cell lines by staining with the PE Annexin V Apoptosis Detection Kit I (BD Biosciences, San Jose, California, USA, catalog number 559763). Approximately 10^5 cells (100 μL) in a 5 mL round-bottom tube with cell-strainer cap (Falcon/Corning, Corning, New York, USA, catalog number 352235) were tested with 5 μL each of PE Annexin V and 7-AAD (7-aminoactinomycin D) viability staining dye (BioLegend, San Diego, California, USA, catalog number 420404) and incubated for 15 min at room temperature, protected from light. 400 μL of 1X Binding Buffer was added, and the tube was immediately analyzed on a BD Accuri C6 flow cytometer. All experiments for each cell line tested were run in triplicate.

Transfection efficiency was evaluated in three renal cell lines (786-O, ACHN, HK2) that separately stably expressed both DFP3 isoforms, DPF3a and DPF3b. Cells were analyzed on a BD FACSAria III with BD FACSDiva Software version 7.0. We used forward scatter and side scatter parameters to eliminate debris, dead cells, and doublets from the analysis. GFP was excited with the blue 488 nm laser and fluorescence was detected via a 530/30 nm bandpass filter in the FITC channel. Untransfected cells were used as negative controls.

Motif analysis for 14q24 variants

To identify putative transcription factor (TF) binding sites, we performed motif analysis for 14q24 variants by using the *motifbreakR*[41] R package and HOCOMOCO[42] positional weight

matrices with default parameters. The p value threshold was set at 1.0×10^{-4}.

eQTL analyses and colocalization of GWAS signals
We analyzed The Cancer Genome Atlas (TCGA) KIRC (kidney renal clear cell carcinoma) RNA sequencing (RNA-seq) and genotype data by using FastQTL[43] to identify deregulated genes and associated SNPs ± 2 Mb around the GWAS tag SNP, with TCGA-KIRC tumor (n = 486) and normal tissue (n = 72) data analyzed separately, with a significance threshold corrected for multiple testing of p < 4.7×10^{-5}. Colocalization analysis for the 14q24 RCC region based on GWAS meta-analysis[22] used eCAVIAR[44] as implemented in vQTL. Both GWAS and TCGA-KIRC eQTL summary statistics of variants in that locus were extracted as the input for eCAVIAR. We selected 50 flanking SNPs of rs4903064. We computed the colocalization posterior probability (CLPP) score with a maximum number of two causal SNPs in the locus. LocusCompare[45] colocalization plots were generated with the vQTL website.

Fine-mapping with PAINTOR
Bayesian fine-mapping of the original summary data[22] for the 14q24 RCC risk region was performed with PAINTOR 3.0[46] with default parameters of window size of 100 kb and maximum number of causals set to two, filtering out all the SNPs with p value > 0.5 for computational efficiency. The pairwise LD between all SNPs in the window was computed with 1000 Genomes EUR data. No functional annotations were used in this analysis because the prioritization was done with in-house high-throughput assays. Note, although rs4903064 was only the fourth most significant variant in these summary data (p = 1.06×10^{-14}), subsequent replication genotyping and meta-analysis of select SNPs from this region in a set of 3,182 independent cases and 6,301 independent controls showed rs4903064 as the most significant SNP at the locus and similar in significance to set variants (rs4903064 p = 2.19×10^{-24}; rs2109794 p = 2.50×10^{-24}).

Immunoblotting
The following antibodies were used: anti-FLAG M2 (catalog number F1804-50UG, Sigma-Aldrich; 1:5,000), anti-HIF-1α (clone 54; BD Transduction Laboratories; 1:1,000), anti-β-actin (NB600-505, Novus Biologicals, Littleton, CO; 1:5,000), anti-IL23R (catalog number NB600-114, Novus; 1:1,000), anti-CEMIP (catalog number 21129-1-AP, Proteintech; 1:500), anti-caspase-3 (catalog number 9662, Cell Signal; 1:1,000), anti-STAT3 (catalog number 4904S, Cell Signaling; 1:2,000), anti-p-STAT3 (catalog number D3A7, Cell Signaling; 1:2,000), anti-PARP (catalog number 9531, Cell Signaling; 1:1,000), anti-β-actin (catalog number NB600-505, Novus Biologicals; 1:5,000), goat anti-mouse HRP (NB7511, Novus Biologicals; 1:5,000), goat anti-rabbit HRP (NB7160, Novus Biologicals; 1:5,000), and donkey anti-goat HRP (sc-2020, Santa Cruz Biotechnology; 1:1,000). Whole-cell extracts were prepared in RIPA buffer with protease inhibitors (cOmplete Protease Inhibitor, Roche) or phosphatase inhibitors (PhosSTOP, Roche). Samples were quantified with the DC Protein Assay (Bio-Rad) and electrophoresed on 4%–12% Bis Tris Plus Bolt gels in MES buffer (Invitrogen). Proteins were transferred to nitrocellulose via an iBlot2 (Invitrogen). Blots were blocked with 5% non-fat dry milk in Tris-buffered saline with 0.1% Tween-20 (TBST). Primary and secondary antibodies were diluted in 5% milk in TBST, and all washes were performed with TBST. Blots were rinsed briefly with PBS before the addition of ECL Prime Western Blotting

Detection Reagent (Amersham). Images were captured on Chem-iDoc Gel Imaging System (Bio-Rad).

Multiplex immunoassays
Mutliplex immunoassays were used for quantifying apoptosis proteins in stable DPF3-inducible cell lines and quantifying cytokines in media from PBMC-cytotoxic assays. For apoptosis protein quantification, cytosolic and nuclear+mitochondrial fractions were extracted from inducible DPF3 cell lines treated with and without dox, according to the Bio-Plex Pro RBM Multiplex Apoptosis Assays manual (Bio-Rad). We analyzed Bak, Bax, Lamin B, Smac, Bad, Bax/Bcl-2 dimer, Bcl-xL, Bim, Mcl-1, active caspase-3, Bcl-xL/Bak dimer, Mcl-1/Bak dimer, and Survivin. Plates were read with Bio-Plex MAGPIX Multiplex Reader (Bio-Rad). Analysis was done via a linear regression model (protein ~ cell line + DPF3 isoform + dox/control + replication), and p values were corrected for multiple comparisons via false discovery rate (FDR).

For cytokine assays, we quantified 40 cytokines in the media collected after PBMC-cytotoxic assay: IFN-, TNF-α, IL-1β, IL-2, IL-4, IL-6, IL-8, IL-10, IL-16, MIF, GM-CSF, CCL1, CCL2, CCL3, CCL7, CCL8, CCL11, CCL15, CCL17, CCL19, CCL20, CCL21, CCL22, CCL23, CCL24, CCL25, CCL26, CCL27, CXCL1, CXCL2, CXCL6, CXCL9, CXCL10, CXCL11, CXCL12, CXCL13, CXCL16, and CX3CL1. Reactions were done following manufacturer's instructions (Bio-Plex Pro Human Chemokine Assays, Bio-Rad). Analysis was done via a linear regression model (cytokine ~ cell line + dox/control + replication), and p values were corrected for multiple comparisons via FDR.

Whole-exome sequencing
200 ng genomic DNA was purified with Agencourt AMPure XP Reagent (Beckman Coulter, Brea, CA, USA), and an adaptor-ligated library was prepared with the KAPA HyperPlus Kit (KAPA Biosystems, Wilmington, MA) with Bioo Scientific NEXTflex DNA Barcoded Adapters (Bioo Scientific, Austin, TX, USA). Genomic DNA sample libraries were amplified pre-hybridization by ligation-mediated PCR consisting of one reaction containing 20 μL library DNA, 25 μL 2× KAPA HiFi HotStart ReadyMix, and 5 μL 10× Library Amplification Primer Mix with PCR cycling conditions as follows: 98°C for 45 s, followed by five cycles of 98°C for 15 s, 60°C for 30 s, 72°C for 30 s plus an extension at 72°C for 1 min and cleaned with Agencourt AMPure XP Reagent (Beckman Coulter, Brea, CA, USA). Amplified sample libraries with unique barcoded adapters were combined in equal amounts into 1.1 μg pools for multiplex sequence capture, and exome sequence capture was performed with NimbleGen's SeqCap EZ Human Exome Library v.3.0 (64 Mb targeted exonic sequence) or Exome+UTR (96 Mb targeted; Roche NimbleGen, Madison, WI, USA). Washing and recovery of captured DNA were performed as described in NimbleGen SeqCap EZ Library SR protocol. Pools of captured DNA were amplified by ligation-mediated PCR consisting of one reaction for each pool containing 20 μL captured library DNA, 25 μL 2× KAPA HiFi HotStart ReadyMix, and 5 μL 10× Library Amplification Primer Mix. PCR cycling conditions were as follows: 98°C for 45 s, followed by eight cycles of 98°C for 15 s, 60°C for 30 s, 72°C for 30 s with an extension at 72°C for 1 min, and amplified material was cleaned with Agencourt AMPure XP Reagent (Beckman Coulter, Brea, CA, USA). Pools of amplified captured DNA were quantified with Kapa's Library Quantification Kit (Kapa Biosystems, Woburn, MA, USA) on a LightCycler 480 (Roche, Indianapolis, IN, USA). The resulting post-capture enriched

multiplexed sequencing libraries were used in cluster formation on an Illumina cBOT (Illumina, San Diego, CA, USA), and paired-end sequencing was performed with an Illumina HiSeq following Illumina-provided protocols for 2 × 125 bp (HiSeq 2500) or 2 × 150 bp (HiSeq 4000) paired-end sequencing. We sequenced each exome to high depth to achieve a minimum threshold of 80% of coding sequence (CDS) bases covered with at least 15 reads, based on the UCSC hg19 "known gene" transcripts. An average CDS coverage of over 35-fold was required for analyses.

Analysis was done with a previously published pipeline.[47] The Trimmomatic[48] program was used for sequence trimming, and then these sequences were aligned to the GRCh37 reference genome with the Novoalign software. Variant calling was performed with the UnifiedGenotyper and HaplotypeCaller modules from GATK as well as the FreeBayes[48] variant caller, and all three calls were then integrated via an Ensemble variant calling pipeline. We used R package *maftools*[48] to plot TCGA-KIRC genes and SWI/SNF genes.

RNA sequencing

We performed RNA sequencing on stable DPF3-inducible cell lines to evaluate gene expression deregulation by DPF3 overexpression in two biological replicates per condition. DPF3-stable cell lines were treated with dox for 48 h, followed by RNA extraction with the RNeasy Mini QIAcube Kit (QIAGEN) with DNase treatment (QIAGEN) on a QIAGEN QIAcube. Ribosomal RNA was removed with Epicenter's Ribo-Zero Gold (human/mouse/rat), followed by RNAClean XP purification. Barcoded sequencing library preparation was performed with the Kapa Stranded RNA-Seq Library Preparation Kit (Kapa Biosystems). Libraries were quantitated for sequencing on the Agilent Bioanalyzer with the High Sensitivity DNA Kit. Sequencing was performed on the NextSeq 500 (2 × 151 bp paired-end, v.2 chemistry; eight samples were run across two sequencing runs). Analysis was done with a pseudo-alignment approach (Kallisto and Sleth pipelines).[49,50]

Capture HiC

We employed a high-throughput method, Capture-C, to quantify chromatin spatial organization interactions between genes and MPRA SNPs in the cell lines used for MPRA (ACHN and HEK293T). HiC libraries were prepared with the 4-base cutter, DpnII.[51] We designed an Agilent SureSelect library to specifically target risk-associated SNPs within restriction fragments harboring SNPs with an LD of $R^2 \geq 0.2$ to the lead SNP from each of the 20 GWAS regions. For 14q24 region, baits included 26 out 39 MPRA variants, while 13 could not be designed (rs80189309, rs185825120, rs6574106, rs145441531, rs11629006, rs17122159, rs12587742, rs148614085, rs144565870, rs75701277, rs12588830, rs68132241, and rs141451882). Two biological replicates were performed per cell line and sequenced on Illumina NextSeq or HiSeq 4000. When we performed technical replicates to increase read depth, FASTQ files were merged. We used Bowtie 2[52] to remove phiX reads, and we used Trimmomatic[48] for quality trimming. After trimming and filtering, we mapped 595 million pair-end reads for HEK293T and 681.5 million for ACHN by using HiCUP[52] (Bowtie 2 to hg19) and we analyzed HiC contacts by using CHiCAGO[52] with the following settings: minFragLen = 25, maxFragLen = 5,000, binsize = 2,500, minNPerBait = 85, tlb.minProxOEPerBin = 150, tlb.minProxB2BPerBin = 15, and techNoise.minBaitsPerBin = 150. Using a CHiCAGO signal cutoff of at least 5, we noted a total of 445,773 interactions in HEK293T and 809,142 in ACHN.

Integrated System for Motif Activity Response Analysis (ISMARA)

We used the ISMARA[53] package to identify the key TFs driving expression/chromatin changes by using RNA-seq data from *DPF3* dox-inducible cell lines. FASTQ files were submitted to the ISMARA online tool. ISMARA activity reports for each sample were aggregated and a t test was used for comparison of inferred TF activity for each cell line and *DPF3* isoform. All TFs that have the *DPF3* gene as one of the genes for activity were flagged and not further considered because their activity measurement would be artificially inflated because of manipulated *DPF3* overexpression.

Accessible chromatin with high-throughput sequencing (ATAC-seq)

We performed ATAC-seq to evaluate chromatin accessibility comparing DFP3-overexpressing cell lines to controls, following a published protocol.[54] After testing several starting cell numbers (10,000, 20,000, 30,000, 40,000, 50,000, 60,000, 70,000, 80,000, 90,000, and 100,000 cells) for fragment size distribution and PCR amplification cycle needed, we established 30,000 starting cells for the cell lines used on this study. ATAC-seq was done in duplicate for HK2 DPF3a, HK2 DPF3b, UOK-121 DPF3a, and UOK-121 DPF3b. ATAC-seq libraries were sequenced on HiSeq 4000, and peak calling was performed with the ENCODE pipeline. For enhancer analyses of ACHN and HEK293T, the ENCODE pipeline was employed (web resources). Files were rimmed and aligned with Bowtie 2. ATAC-seq signal was generated via BEDTools,[52] and peak calling was done with MACS.[55] Peaks on the replicates were evaluated with IDR (irreproducibility discovery rate).

For differential peak analysis between DPF3 and control samples, the analysis was done with the ALTRE package.[56] To compare genes deregulated in RNA-seq to peaks of ATAC-seq, we mapped peaks up to 2 Mb each direction from each deregulated gene. Statistical analysis was done via chi-square test on a 2 × 2 table of mapping DPF3 RNA-seq deregulated genes and random genes not deregulated by *DPF3*.

H3K27ac chromatin immunoprecipitation sequencing (ChIP-seq)

H3K27ac ChIP-seq was performed with ACHN and UO31 cells at 70%–90% confluence in 15 cm dishes, rinsed twice with DPBS (Quality Biological) and crosslinked with 10 mL 0.75% MeOH-free formaldehyde (Thermo Fisher) for 10 min. Formaldehyde was quenched by the addition of 600 μL of 2.5 M glycine. Cells from 2 × 15 cm dishes were scraped into a conical tube, centrifuged at 200 × *g* for 5 min at 4°C and washed twice with ice cold DPBS + protease inhibitors (Roche cOmplete tablets) and resuspended in 5 mL ice-cold cell lysis buffer (5 mM PIPES [pH 8], 85 mM KCl, 0.5% NP-40) + protease inhibitors for 10 min. To facilitate lysis, cells were passed thru a 20G needle ten times and centrifuged at 1,250 × *g* for 3 min at 4°C. Pellets were resuspended in 600 μL nuclei lysis buffer (50 mM Tris [pH 8], 10 mM EDTA, 1% SDS) + protease inhibitors and split into two microcentrifuge tubes for sonication in a Bioruptor, and cellular debris was removed by centrifugation. Sheared chromatin was confirmed on a Bioanalyzer (Agilent) and 30 μg of sheared chromatin was diluted 1:10 with dilution buffer (16.7 mM Tris [pH 8], 167 mM NaCl, 1.2 mM EDTA, 0.01% SDS, 1.1% Triton X-100) + protease inhibitors and mixed with 4 μg of anti-H3K27Ac (ab4729, lot GR254707-1) and 60 μL of prepared magnetic protein A/G beads (Thermo Fisher) and rotated overnight at 4°C. Beads were collected with a magnet
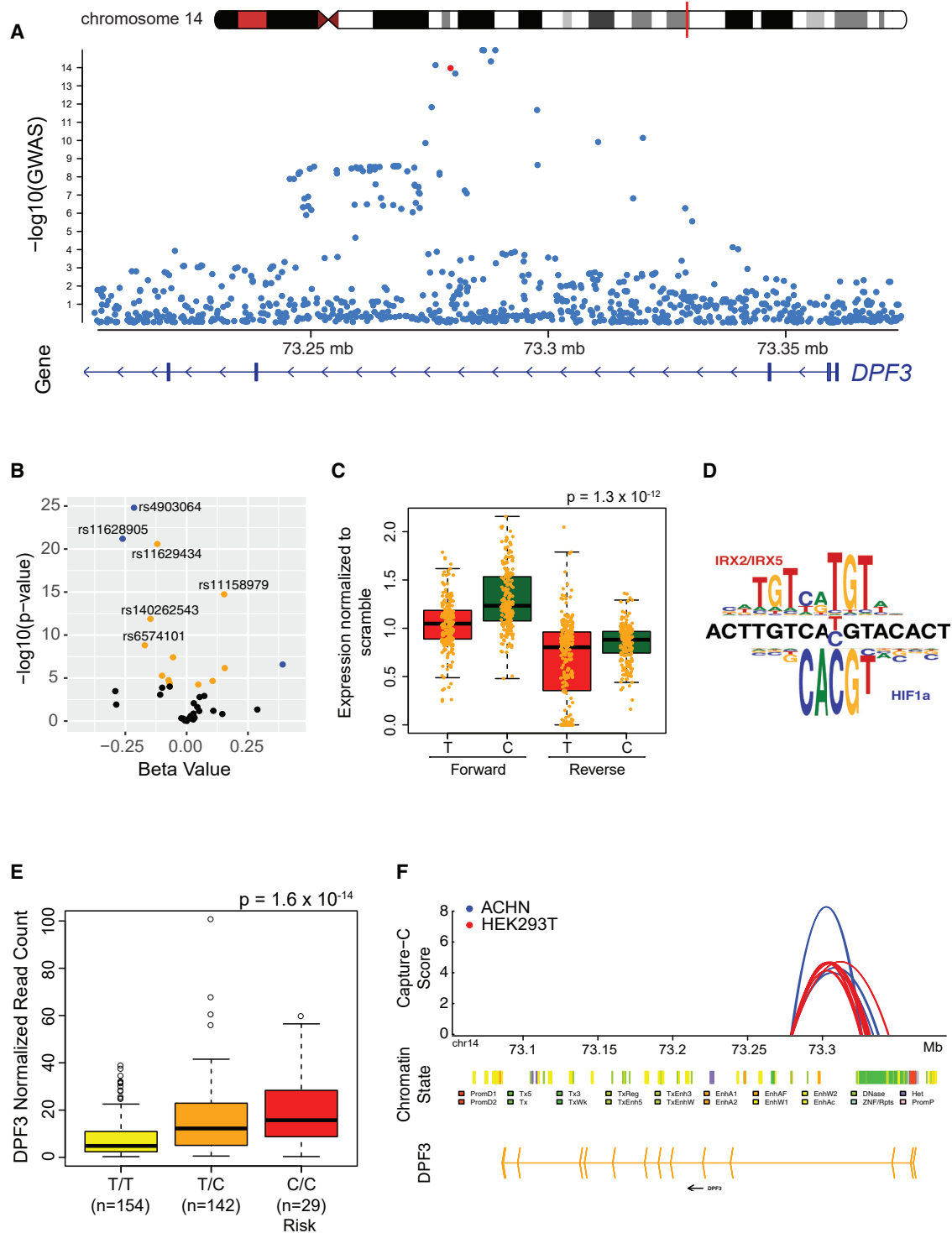
Figure 1. Functional characterization of the 14q24 renal cancer risk loci

(A) Visualization of 14q24 RCC GWAS results; each point is a SNP; the red point is the tag SNP rs4903064; the y axis is the −log10 of GWAS discovery phase p value.

(B) Volcano plot for MPRA analysis of 14q24 SNPs identifies the tag SNP rs4903064 as the top functional candidate for the region; p values and beta values are for allelic effect of MPRA signal correcting for direction, cell lines (ACHN and HEK293T), and condition (hypoxia and normoxia). Yellow circles represent variants with significant allele difference (q value < 0.05) and beta value < 0.2; blue circles represent variants with significant allele difference (q value < 0.05) and beta value > 0.2; and black circles represent non-significant variants.

(C) Boxplot for 26 MPRA biological replicates (ACHN and HEK293T; 40 data points per replication) for each allele of rs4903064 in each direction shows the C risk allele has higher signal in both directions compared to the T protective allele; the black line in the center of boxes represents the median, whiskers extend to the minimum of either the data range or 1.5 times the interquartile range, and statistical

(legend continued on next page)

and washed four times with high salt wash buffer (50 mM HEPES [pH 7.9], 500 mM NaCl, 1 mM EDTA, 1% Trix-100, 0.1% Deoxycholate) followed by two washings with Tris-EDTA (pH 8) and eluted with 200 μL of buffer (50 mM Tris [pH 8], 10 mM EDTA, 1% SDS) at 37°C for 2 min. Crosslinks were reversed by addition of RNase A and proteinase K overnight at 65°C, followed by cleaning with 1.8× AmpureXP (Beckman Coulter) and elution in 50 μL 0.5× TE. Fragmented DNA was prepared for Illumina paired-end sequencing with the NEBNext Ultra DNA Library Prep Kit for Illumina (NEB). Libraries were then quantified with the NEBNext Library Quant Kit for Illumina (NEB) and sequenced with the MiSeq Reagent Kit v.3 150 cycle PE. Reads were aligned to hg19 via BWA MEM (Burrows-Wheeler aligner) and peaks were called with MACS2[55] callpeak with the broad flag.

Results

Characterization of functional activity of RCC GWAS sentinel SNP, rs4903064, at 14q24

We used an integrated approach to prioritize putative functional *cis*-regulatory SNPs from 20 RCC GWAS regions (n = 11 loci achieving genome-wide significance and n = 9 marginally significant regions, $5 \times 10^{-8} < p_{assoc} < 5 \times 10^{-6}$) and screened with 774 variants with a massively parallel reporter assay (MPRA).[38] Herein, we focused on variants from the 14q24 RCC susceptibility region, tagged by rs4903064[22] (Figure 1A, Figure S1); 39 SNPs in LD to the sentinel SNP ($R^2 > 0.2$, 1000G EUR or $D' > 0.5$ and MAF < 0.05) and lying within putative *cis*-regulatory elements were screened via MPRA on both a renal cell cancer cell line (ACHN) and embryonic renal cell line (HEK293T) under both normoxic and hypoxic conditions. 14 displayed a significant allele-specific effect (q for allelic effect < 0.05), and three of these were not consistent across both cell lines (p for cell type < 0.05), yielding 11 high-interest variants (Figure 1B, Table S2) that were advanced. Notably, rs4903064 displayed the most significant allelic effect (q value = 1.2×10^{-22}; beta = −0.21; Figure 1B, Table S2); the C risk allele displayed higher reporter expression than the T protective allele when cloned in both forward and reverse orientations (Figure 1C). Under hypoxic conditions, there was a stronger allelic difference for rs4903064 (q value = 5.9×10^{-17}; beta = 0.33) compared to normoxic conditions (q value = 1.6×10^{-5}; beta = 0.14; Figure S2).

A series of studies confirmed rs4903064 maps to a plausible *cis*-regulatory element in kidney tumors. Nine of 39 MPRA variants also fell under ATAC-seq peaks from either of the lines used for MPRA (ACHN and HEK293T; Table S3). Kidney tumor ATAC-seq data from TCGA (KIRC and KIRP [kidney renal papillary cell carcinoma]) indicated that seven variants, including rs4903064, were in regions of chromatin accessibility (Figures S1 and S3, Table S3); 12 of 39 variants were located in regions marked by H3K27ac in either ACHN or U0131 cells (Table S4, Figure S3). Renal ENCODE data also supported our observations, as rs4903064 is located in a region of open chromatin as defined by peaks for DNase hypersensitivity (Table S3, Figure S4; ENCODE kidney samples HRGEC, RPTEC, and HEK293) and FAIRE (formaldehyde-assisted isolation of regulatory elements; Figure S4). Finally, we evaluated predicted allelic TF binding across the 14q24 MPRA SNPs in the HOCOMOCO database;[57] nine variants, including rs4903064, showed allele-specific binding sites (Figure 1D, Table S5).

rs4903064-C risk-increasing allele is associated with higher DPF3 expression

To investigate downstream effects of the allele-specific regulatory activity of rs4903064, we analyzed *cis*-expression quantitative trait loci (*cis*-eQTL) for genes within 1 Mb on either side of rs4903064 in two RCC tumor datasets (TCGA-KIRC[8] and International Agency for Research on Cancer [IARC][22]). Both sets identified *DPF3* as the strongest signal surpassing genome-wide significance and as an eQTL target gene for rs4903064.[22] In a re-analysis of both TCGA-KIRC tumor and normal data with the GTEx eQTL pipeline,[43] the rs4903064-C risk-increasing allele was associated with higher *DPF3* expression specifically in tumors ($p_{eQTL} = 1.6 \times 10^{-14}$, q value = 6×10^{-10}, and beta = 66.6; n = 486; Figure 1E, Table S6), consistent in direction with the MPRA results. Although rs4903064 was the top eQTL variant for *DPF3*, 73 total variants in total revealed a significant eQTL with *DPF3* in tumors ($p < 4.7 \times 10^{-5}$; Table S6); there was no significant eQTL for other genes in TCGA-KIRC tumors nor TCGA-KIRC normal kidney tissue (n = 72; Table S6). In an eQTL derived from normal kidney tissue,[58] the same allelic effect was observed for *DPF3* expression with a high-LD surrogate (rs4903074, $D' = 0.82$, p = 1.1×10^{-6}; Figure S5). Although kidney tissue has low representation in GTEx data, the rs4903064-C allele was associated with higher *DPF3* expression across a range of normal tissues by eQTL meta-analysis ($p_{meta-analysis} = 3.0 \times 10^{-27}$; Figure S6). The lead SNP, rs49030704, is not a splice QTL (sQTL) in GTEx, further supporting a *cis*-regulatory mechanism.

outliers are plotted as points. The p value is for analysis of both directions. Direction specific analysis: forward direction p value = 2×10^{-20} and beta = −0.22; reverse p value = 1×10^{-8} and beta −0.17.

(D) rs4903064 T and C allele sequence contexts create binding sites for IRX2/IRX5 and HIF-1a, respectively.

(E) eQTL analysis of the TCGA-KIRC dataset via FastQTL shows that the rs4903064-C risk allele is associated with higher *DPF3* expression relative to the T protective allele; the black line in the center of boxes represents the median, whiskers extend to the minimum of either the data range or 1.5 times the interquartile range, and statistical outliers are plotted as points.

(F) Plot for chromatin looping between the Capture HiC-baited restriction fragment that includes rs4903064 and a region close to the *DPF3* promoter for two cell lines (ACHC in blue and HEK293T in red). Among 14q24 MPRA SNPs, only rs4903064 was present on this plotted bait; the y axis indicates Capture-C score from CHiCAGO analysis, which is the −log(weighted p value); the region is characterized with 25 chromatin states data from Roadmap fetal kidney.
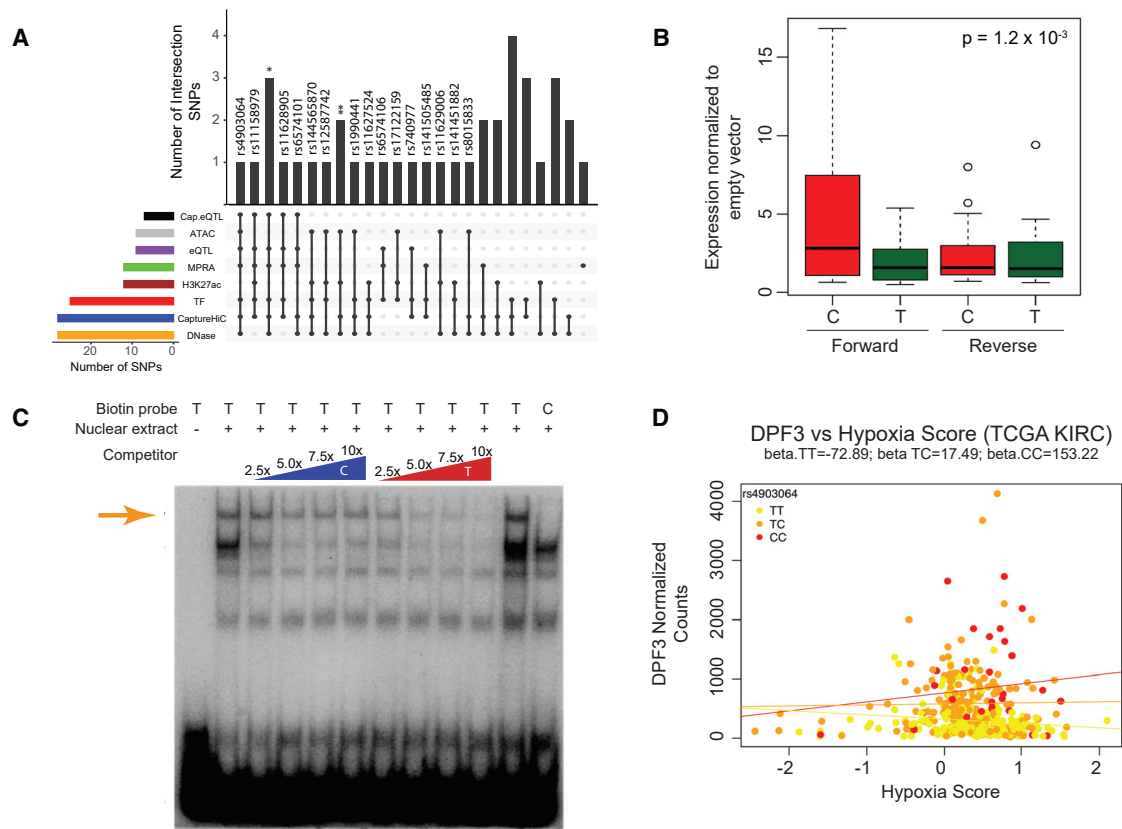
Figure 2. Summary of prioritization of 14q24 renal cancer risk loci variants

(A) UpSet plot for 14q24 variants summarizing evidence for each one's being a potential functional *cis*-regulatory variant. ATAC, variants within ATAC-seq peaks; eQTL, variants with a significant eQTL with a gene in the region; MPRA, variants presenting allele-specific MPRA signal in both tested cell lines; H3K27ac, variants located within H3K27ac ChIP-seq peaks; TF, variants with predicted allele-specific transcription factor binding; Capture HiC, presence of a loop to a gene from the restriction fragment harboring the variant; Cap.eQTL, Capture HiC loop is from the variant to an eQTL gene; DNase, variant present within a DNase accessible peak; *, rs2332922, rs72730353, and rs11629434; **, rs11158980 and rs68132241.

(B) Boxplot for individual luciferase reporter assay results (combination of ACHN, HEK293T, HK2 cell lines; six replicates each; four data points per replicate) confirms the rs4903064-C allele has higher signal compared to the T protective allele; the black line in the center of boxes represents the median, whiskers extend to the minimum of either the data range or 1.5 times the interquartile range, and statistical outliers are plotted as points. The p value is for all three cell lines and both forward and reverse orientations analyzed together. The analysis was done with linear regression for evaluation of allele effect in a model that included cell line, direction, and replicate with both forward and reverse results.

(C) EMSA for rs4903064 alleles shows a higher protein binding activity for the T protective allele compared to the C risk allele with nuclear extract from HEK293T; lanes represents different conditions (#1, T allele without nuclear protein extract; #2, T allele with nuclear protein extract; #3–6, T allele competed with increasing amounts of non-labeled C allele probe; #7–10, T allele competed with increasing amounts of non-labeled T allele probe; #11, T allele with nuclear protein extract; #12, C allele with nuclear protein extract.

(D) Comparison of *DPF3* expression and hypoxia score in TCGA-KIRC samples. Grouping samples by rs4903064 genotype shows that higher hypoxia score is associated with increasing *DPF3* expression, mainly in samples homozygous for the C risk allele. Linear regression analysis was conducted with the following model: $DPF3 \sim rs4903064.73279420.T.C + hypoxia + rs4903064.73279420.T.C \times hypoxia + IRX2 + IRX5 + IRX2 \times IRX5$. p value (hypoxia × rs4903064) = 0.07.

Finally, colocalization with GWAS summary statistics and the TCGA-KIRC tumor eQTL data was confirmed by eCAVIAR[44] (Figure S7); rs4903064 displayed the highest colocalization posterior probability (CLPP) score (CLPP = 0.13; cutoff for colocalization at CLPP > 0.01; Table S7). Because rs4903064 is intronic, we performed RACE to search for novel *DPF3* isoforms as well as confirmation of known isoforms DPF3a and DPF3b, but we identified no new isoforms (Figure S8). Our results indicate the rs4903064-C risk allele is associated with higher *DPF3* levels in both normal and tumor samples, consistent with the pattern of allelic *cis*-regulatory activity we observed for rs4903064 *in vitro*.

Physical interaction between rs4903064 and *DPF3* was investigated by Capture HiC (CHiC) assay[59] in both ACHN and HEK293T cells; we designed capture baits to specifically target DpnII restriction fragments harboring risk-associated SNPs, including 26 out 39 variants tested by MPRA. We observed evidence for physical interactions between the restriction fragment harboring rs4903064 and regions downstream of the *DPF3* transcriptional start site in both cell lines (Figure 1F, Table S8). The region interacting with rs4903064 is classified as "transcribed and regulatory (promoter/enhancer)" in fetal kidney chromatin state data of the Roadmap.[60] For the other 25 baited

MPRA variants on other baited restriction fragments, there was evidence of interaction with the body of *DPF3* and other genes (Table S8). Together with the high-throughput assays, namely MPRA, ATAC-seq, H327ac ChIP-seq, Capture HiC, and eQTL (Figure 2A, Figure S9), the observed physical link between rs4903064 and *DPF3*[61–63] transcriptional machinery further supports that rs4903064 acts as an enhancer of *DPF3* expression.

Validation of rs4903064 allelic enhancer activity by luciferase assay

Allelic enhancer activity of rs4903064 was confirmed by individual luciferase assay analysis in the same cell lines used in the MPRA (Figure S10). Further analysis in a third cell line, HK2 (kidney proximal tubule cell line), confirmed the allele-specific effect (p = 1.1 × 10^{-6} for HK2, Figure S10; p = 1.2 × 10^{-3} for all cells analyzed together for both forward and reverse orientation; Figure 2B), an effect that was more consistent in the forward orientation (Figure S10).

Six other SNPs (rs8015833, rs141505485, rs140262543, rs740977, rs6574106, and rs72730353) were investigated further on the basis of allelic difference p value, direction of effect, and cell line specificity in the MPRA. Only two, rs72730353 (D' = 0.87, R^2 = 0.49 with rs4903064; EUR,[64] Figure 2A) and rs140262543 (D' = 1, R^2 = 0.01; EUR, Figure 2A), confirmed allele-specific enhancer activity in individual reporter assays (Figures S11A–S11F). Finally, although rs4903064 is a member of the 99% credible causal variant set as assessed via PAINTOR-[46]based analysis of RCC GWAS summary data, neither of these two variants were, nor were any of the other variants assessed via MPRA (Table S9). Our subsequent work thus focused on rs4903064 but with knowledge that other variants could contribute to the net effect of the germline RCC risk locus (Figure 2A).[25]

Allele-specific protein binding to rs4903064

An allele-specific pattern of protein binding to rs4903064 (Figure 2C, Figure S12) was observed with nuclear extract from both HEK293T and ACHN cells by EMSAs. The protective rs4903064-T allele displayed stronger competition from unlabeled T-probe versus C-probe, and when unlabeled rs4903046-C is present at 10-fold excess, the specific protein-DNA complex with the labeled T protective allele remains (Figure 2C, lane 6). In motif analyses, rs4903064-T creates a binding site for IRX2/IRX5, while the C risk allele creates a binding site for HIF (Figure 1D). IRX2 and IRX5 are known transcriptional repressors,[65] while HIF can be an activator, which is consistent with the observed stronger allelic effect under hypoxia (Figure S2). We examined *IRX2* and *IRX5* gene expression against *DPF3* expression in normal and tumor samples (TCGA-KIRC, n = 538 tumor and 71 normal tissue analyzed together), and *IRX2* and *IRX5* levels correlate inversely with *DPF3* expression (p$_{IRX2}$ = 1.2 × 10^{-4} and p$_{IRX5}$ = 3.1 × 10^{-7}, respectively), suggesting transcriptional repression by *IRX2* and *IRX5* (Figure S13). HIF

expression is not necessarily a robust surrogate for hypoxia, but application of a reported hypoxia score[66] revealed a trend of correlation between higher hypoxia levels and increased *DPF3* expression with the rs4903064-C allele, particularly for CC homozygotes (Figure 2D; interaction between rs4903064 and hypoxia score on *DPF3*, p = 0.07); the effect progressively diminished for TC and TT genotypes. Previously published HIF1A ChIP-seq data provided evidence of HIF1A binding spanning rs4903064 in the T-47D breast cancer cell line (Figure S14).[67] Collectively, these results suggest that IRX2/IRX5 could bind to rs4903064-T to decrease enhancer activity, as observed in the MPRA, even under normoxic conditions. Thus, rs4903064 is located within a multifunctional binding site: in normal tissue under normoxic conditions, the T protective allele could be a binding site for a transcriptional repressor, possibly IRX2/5, whereas in tumor tissue after loss of VHL leading to stabilization of HIF and activation of HIF-controlled transcriptional programs, HIF could bind the rs4903064-C risk-increasing allele, increasing the signal difference between the two alleles.

DPF3 overexpression increases growth rate in RCC-stable cell lines

Because higher *DPF3* expression is correlated with the renal-cancer-risk-increasing C allele of rs4903064, we investigated the effects of increased *DPF3* levels on kidney cell growth. We generated doxycycline (dox)-inducible DPF3 cell lines to evaluate the effects of modulating *DPF3* levels by real-time growth by using assays that measure bio-impedance as a measure of cell density every 10 min in a 96-well format (5,000 cells/well, 4–8 replicates) treated after 4 or 22 h of plating the cells with dox. Each of the two common *DPF3* isoforms, *DPF3a* or *DPF3b*, were separately introduced into four cell lines: ACHN (papillary RCC origin), UOK-121 and 786-O (clear cell RCC origin), and HK2 (Figures S15 and S16). Cell lines representing papillary and clear cell subtypes were selected because the reported RCC GWAS[22] included both clear cell (~90% of cases) and papillary subtypes (<10%).

To characterize the genomic background of the RCC cell lines given the known importance of mutations in the SWI/SNF complexes in RCC,[13,33,34] we performed exome sequencing on UOK-121, ACHN, 786-O, and HK2 cell lines and revealed mutations in seven of the 50 most commonly mutated genes in the KIRC-TCGA database: *TP53* (UOK-121 and 786-O), *PBRM1* (UOK-121 and ACHN), *MAGEC1* (UOK-121), *ATM* (786-O), *DST* (786-O), *VHL* (786-O), and *PTEN* (786-O) (Figure S17A). The HK2 immortalized normal renal cell line lacked mutations in the top 50 mutated genes in KIRC-TCGA. In addition to *PBRM1*, mutations were observed in other mSWI/SNF genes but not *DPF3*; these included a missense mutation in *DPF2* (ACHN cell line) and nonsense mutation in *SS18* (HK2 cell line) (Figure S17B).

Increased expression of *DPF3* mRNA and protein was confirmed at 12, 24, and 48 h after dox treatment by
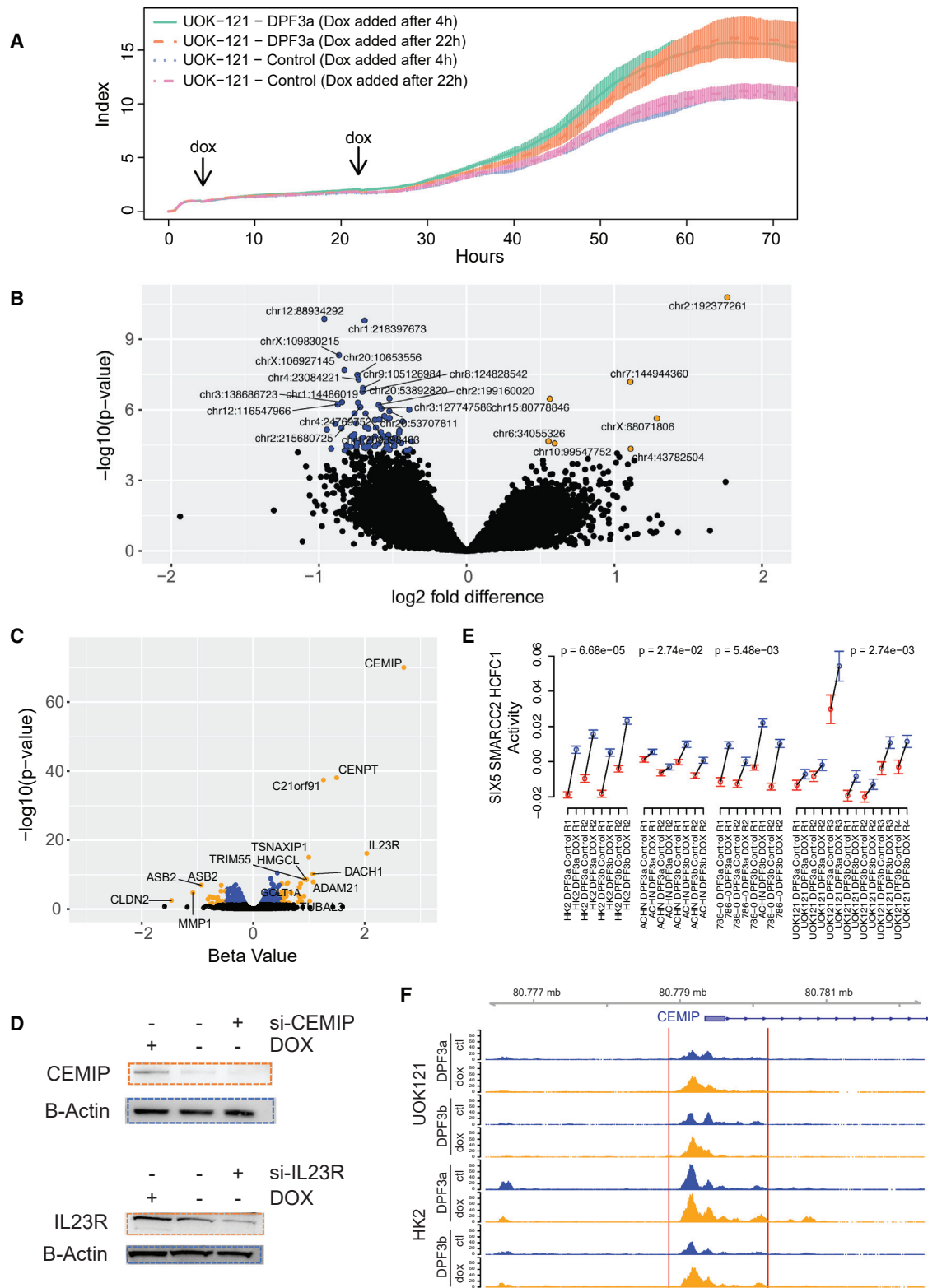
Figure 3. Effect of *DPF3* overexpression on renal cancer cell lines

(A) Real time impedance growth assay for the UOK-121-stable dox-inducible DPF3a cell line showing DPF3 increasing growth rate. 10,000 cells were plated in 96-well plates and treated with dox or control for 4 or 22 h following plating (arrows); data are combined from three technical replicates and mean and confidence interval for each sample/condition are shown.

(B) Volcano plot for genomic regions showing changes in chromatin accessibility after DPF3 overexpression via dox treatment. Data were combined for HK2 DPF3a, HK2 DPF3b, UOK-121 DPF3a, and UOK-121 DPF3b and analyzed with the R ALTRE package. DESeq2 was used for comparison of peak intensities between conditions, adjusting for cell line, *DPF3* isoform, and experimental replicate. The x axis shows the log2 fold change of chromatin accessibility due to *DPF3* overexpression.

(legend continued on next page)

quantitative PCR (qPCR) and immunoblot; more than 95% of the cells were GFP+ and expressing *DPF3* (Figures S16A–S16C). Overexpression of either *DPF3* isoform increased the growth rate in UOK-121, ACHN, and HK2 (Figure 3A, Figures S18, S19, and S20) but not in 786-O (Figure S21). There was no difference in growth curves between normoxic (Figures S19 and S20) and hypoxic conditions (Figures S22 and S23).

DPF3 overexpression leads to changes in chromatin accessibility

Mutations in mSWI/SNF complex genes have been well described, and differential gene loss results in distinct expression and cellular phenotypes.[33] To date, residual mSWI/SNF complexes have only been reported to harbor somatically mutated members, but our reported findings suggest that differential levels of *DPF3* expression, driven by germline variation, could alter chromatin accessibility. In ATAC-seq analysis of dox-inducible *DPF3* cell lines at 48 h after treatment as compared to dox naive cells,[68] 86 regions were identified with significant changes in chromatin accessibility following *DPF3* overexpression of either a and/or b isoforms: seven showed increased accessibility, while 79 showed decreased accessibility (Figure 3B, Table S10).

To identify genes deregulated by *DPF3* overexpression, we performed RNA-seq on viable cells 48 h following dox treatment in four stable cell lines (two replicates for each cell line/isoform). In an analysis of differential gene expression with a pseudo-alignment approach,[49,50] we combined data from both isoforms across cell lines because DPF3a and DPF3b displayed similar effects (R^2 = 0.99; Figure S24). Overexpression of *DPF3* led to a significant increase in expression of 263 genes, while 230 genes were downregulated (Figure 3C, Table S11). The top two deregulated genes (*CEMIP* and *IL23R*) were confirmed by qPCR (Figure S25), and protein levels were similarly confirmed by immunoblot (Figure 3D, Figure S26). Using the Reactome database for pathway analysis of differentially expressed genes,[69] we observed enrichment for seven pathways (Figure S27), including "estrogen-dependent gene expression" and "ESR-mediated signaling," which could contribute to the reported difference in RCC signal at 14q24 when stratified by sex.[22]

RNA-seq expression data was further analyzed with ISMARA,[53] which identifies upstream TFs likely to drive observed gene expression changes by comparing gene expression to regulatory circuitry to address expression/chromatin state changes.[53] TFs known to directly regulate *DPF3* were flagged following ISMARA analysis and not further considered because their activity measurement would be artificially inflated because of dox-inducible *DPF3* overexpression. The subsequent analysis identified 48 TFs likely displaying differential activity (Table S12). SMARCC2, a member of BAF and PBAF complexes, was one of the top two deregulated TFs: overexpression of either isoform resulted in higher inferred SMARCC2 transcriptional activity (Figure 3E) as measured by changes in SMARCC2-targeted genes. These results suggest the effect of *DPF3* on gene expression is at least partially mediated through the mSWI/SNF complexes (BAF and/or PBAF), perhaps through disrupting those complexes and leading to increased activity of SMARCC2. High activity was also observed for CTCF/CTCFL (Figure S28), which is important for mSWI/SNF complex recruitment to topologically associated domain (TAD) boundaries.[70] It is plausible that residual mSWI/SNF activity due to *DPF3* overexpression could reduce CTCF/CTCFL activity and deregulation of gene expression could lead to RCC development.

DPF3 changes chromatin accessibility, leading to altered gene expression

In an integrated analysis of ATAC-seq data generated from *DPF3*-overexpressing cells with corresponding RNA-seq data, ATAC-seq peaks related to *DPF3* overexpression were investigated within 2 Mb of *DPF3*-deregulated genes (as identified by RNA-seq) and compared with ATAC-seq peaks of random genes not deregulated by *DPF3* in the RNA-seq. 12% of deregulated genes have altered chromatin accessibility ($p_{chi-square}$ = 0.008, compared to random non-*DPF3* target gene ATAC-seq peaks; Figure S29), underscoring the effect of *DPF3* on chromatin accessibility and consequential gene deregulation. The *CEMIP* promoter region showed one of the highest increases in accessibility following *DPF3* overexpression (Figure 3F), which is notable because *CEMIP* was the top deregulated gene and has previously been implicated in cancer proliferation[71] and apoptosis.[72,73] The *DPF3* effect on cell growth is at least partially mediated through *CEMIP* overexpression based on siRNA knockdown of *CEMIP* (>90% *CEMIP* knockdown from baseline; Figure S30A); the growth rates

(C) Volcano plot for DPF3-deregulated genes in *DPF3*-stable cell lines (HK2 DPF3a and DPF3b, UOK-121 DPF3a and DPF3b, 786-O DPF3a and DPF3b, ACHN DPF3a and DPF3b). RNA-seq was performed for two samples of each DPF3-isoform-stable cell line and analyzed via a regression model (gene ~DPF3 (dox or control) + isoform + cell line + replication). p values reported on the plot are DPF3 overexpression across all cell lines. The x axis shows the beta value for change of gene expression due to *DPF3* overexpression.
(D) Immunoblot confirming increasing CEMP and IL-23R protein expression following *DPF3* overexpression in UOK-121 and HK2 DPF3a cell lines, respectively. Full gels are included in Figure S25.
(E) ISMARA analysis for each replicate of DPF3-stable cell line RNA-seq data comparing inferred SMARCC2 transcription factor activity in control and DPF3 overexpression states. Statistical test was performed via combination of replicates of control or doxycycline (dox) for each cell line in a t test.
(F) ATAC-seq peaks showing increasing accessibility in the promoter region of *CEMIP* after DPF3 expression. Data combined from HK2 DPF3a, HK2 DPF3b, UOK-121 DPF3a, and UOK-121 DPF3b were analyzed with the R ALTRE package. The x axis shows peak intensity for change of chromatin accessibility at the *CEMIP* promoter due to DPF3 overexpression.
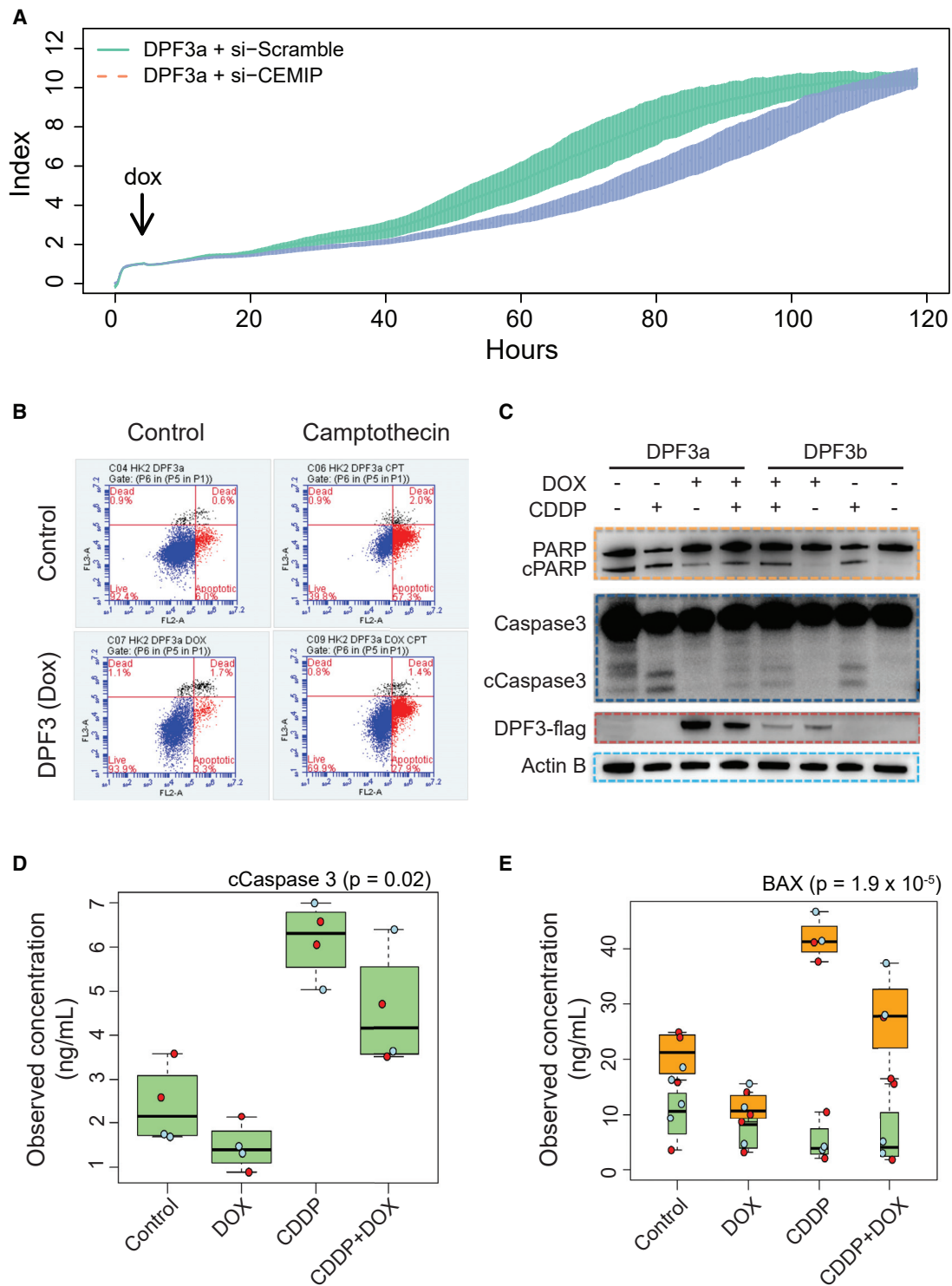
Figure 4. *DPF3* effect on renal cancer is partially mediated by apoptosis

(A) Real-time impedance growth assay for HK2-stable dox-inducible DPF3a cell line with CEMIP siRNA or control treatment. Cell lines were plated and, after 4 h, treated with doxycycline (dox) or PBS (control) ± siRNA against *CEMIP* for evaluation of the growth effect of DPF3 overexpression with *CEMIP* knockdown.

(B) Flow cytometry (Anexin V and 7-AAD) analysis of the HK2 DPF3a-stable cell line treated with dox or control for 48 h. Camptothecin (CPT) was used for induction of apoptosis in the cell lines. FL2-A was the detector for Anexin V and FL3-A for 7-AAD.

(C) Apoptosis proteins were evaluated by immunoblot: PARP, Caspase-3, DPF3-FLAG, and ActinB. HK2 DPF3a and DPF3b cell lines were treated with dox for 48 h, followed by cisplatin (CDDP) treatment for 12 h. Full gels are included in Figure S32.

(legend continued on next page)

of two *DPF3*-overexpressing cell lines (HK2 and ACHN) were substantially reduced when *CEMIP* was targeted compared to scrambled control siRNA (Figure 4A, Figure S31A). Depletion of *CEMIP* without *DPF3* overexpression (stable cell lines without dox) also reduced RCC cell line growth rates (Figures S31B and S31C).

Because *CEMIP* has an established role in apoptosis, we evaluated the effect of *DPF3* overexpression by flow cytometry (Annexin V and 7-AAD) after exposure to two agents capable of inducing apoptosis in kidney cancer cells *in vitro*: cisplatin, an alkylating agent, and camptothecin, a topoisomerase inhibitor. *DPF3* overexpression resulted in a lower fraction of apoptotic cells compared to controls when treated with camptothecin (Figure 4B): the percentage of apoptotic cells was reduced from 57% to 28% following *DPF3* overexpression. To confirm that the *DPF3* effect was independent of a specific agent, we evaluated apoptosis-related proteins by immunoblot following exposure to cisplatin. *DPF3* overexpression led to a reduction in cleavage of caspase-3 and PARP both in the presence and absence of cisplatin (Figure 4C, Figure S32). Thirteen important proteins involved in apoptosis were evaluated via multiplex immunoassays in cytosolic and nuclear/mitochondrial fractions (Figure S33). Overexpression of *DPF3* led to a reduction of cleaved caspase-3 (Figure 4D) as well as a reduction of BAX in the nuclear/mitochondrial fraction (Figure 4E), which is known to trigger intrinsic apoptosis pathways.[74]

IL23R, a subunit of the IL23 receptor and part of the JAK-STAT signaling pathway, was upregulated following *DPF3* overexpression[75] as confirmed by RNA-seq, qPCR, and immunoblot (Figures 3C and 3D, Figures S25 and S26B). Because STAT3 activation is an important pathway for RCC and is activated by *VHL* loss and/or enhanced after *PBMR1* mutation,[15] we examined whether *DPF3*-dependent IL23R overexpression could lead to STAT3 activation. Transfection of siRNA targeting *IL23R* (si-IL23R efficiency in Figure S30B) in RCC cells overexpressing *DPF3* resulted in reduced growth (Figure 5A), suggesting that, in addition to *CEMIP*, the *DPF3* growth effect can be in part attributed to *IL23R* overexpression. Consistent with the known role of IL23 binding to IL23R's resulting in phosphorylation of STAT3 through JAK,[76] a higher level of STAT3 phosphorylation was observed by immunoblot in each of the two cell lines tested overexpressing DPF3a and DPF3b (HK2 and UOK-121; Figure 5B, Figures S34, S35A, and S35B), yet total STAT3 protein and mRNA expression remained stable as measured by immunoblot (Figure 5B, Figures S34, S35A, and S35B), qPCR (Figure S36), and RNA-seq (*STAT3* p = 0.14). To confirm STAT3 pathway activation by *DPF3* overexpression, we performed gene set enrichment analysis (GSEA) pathway analysis on our RNA-seq

data and found that the IL6-JAK-STAT3 pathway was enriched in cell lines that overexpressed *DPF3* (Figure 5C). Notably, this *DPF3*-dependent STAT3 activation is similar to the IL6-JAK-STAT3 pathway activation in RCC cell lines, particularly in those cell lines with mutated *PBRM1*[15,77] (Figure S37). Together, these results suggest that the allele-specific effect of rs4903064 could be partially mediated through STAT3 activation, comparable to the effect of select *PBRM1* somatic mutations.[15]

DPF3 effect on T-cell-mediated cytotoxic and immunotherapy response

Previous studies have suggested that mutations within the SWI/SNF complexes, particularly in *PBRM1*, could contribute to response to immune checkpoint inhibitor therapy in RCC affected individuals,[77,78] but other studies could not find this association;[79–81] confirmation studies are needed.[82] *Ex vivo* data suggest that *PBRM1* mutations are associated with enhanced sensitization to killing by T cells,[83] an effect most likely mediated through mSWI/SNF complex regulation of chromatin accessibility for INF-γ target genes.[83] Accordingly, we tested whether *DPF3*, through activation of the *STAT3* pathway, could increase T-cell-mediated killing. PBMCs from healthy donors were stimulated with CD3, CD28, CD2, and IL-2 for 48 h, followed by co-culture with confluent *DPF3*-overexpressing RCC cell lines (HK2 and UOK-121) or controls. In a bio-impedance growth assay, *DPF3*-overexpressing cells showed higher T-cell-mediated cytotoxicity than controls (Figure 6A, Figure S38), mimicking the above observations in both *PBRM1*-mutated cells and affected individuals.[77,83] PBMCs were characterized by flow cytometry; PBMCs co-cultured with *DPF3*-overexpressing cells had a higher fraction of CD8+ T cells (Figure 6B) compared to PBMCs co-cultured with RCC control cells. Further, there was an increase in IFN-γ+ CD8+ T cells (Figures S39A and S39B). To test whether *DPF3*-overexpressing cells altered expression of cytokines that stimulate CD8+ T cells, we quantified 40 secreted cytokines and chemokines, which showed a significant reduction of secreted IL-6 and IFN-γ in the media following *DPF3* overexpression (Figure 6C, Figure S40). Notably, *IL6* was also previously seen to be downregulated by DPF3 in RNA-seq in renal cell lines (HK2 *DPF3a*, HK2 *DPF3b*, UOK-121 *DPF3a*, UOK-121 *DPF3b*, 786-O *DPF3a*, 786-O *DPF3b*, ACHN *DPF3a*, and ACHN *DPF3b*; Table S11), and there was a reduction of secreted IL-10 in media after co-culture of RCC cell lines and PBMC (Figure S41).

Lower levels of IL-6 and IL-10 have been associated with a positive response to immunotherapy.[84–87] Although IL-10 can limit CD8+ T cell proliferation and survival, reduction of IL-6 in mouse models has increased

---

(D and E) Immunoassay was used for quantification of cleaved caspase-3 (D) and BAX (E) in cytosolic (green) and nuclear+mitochondrial (orange) fractions of inducible HK2 DPF3a (green circle) and DPF3b (red circle) cell lines with and without DPF3 overexpression (dox). The black line in the center of boxes represents the median, whiskers extend to the minimum of either the data range or 1.5 times the interquartile range, and statistical outliers are plotted as points.

A

B

C
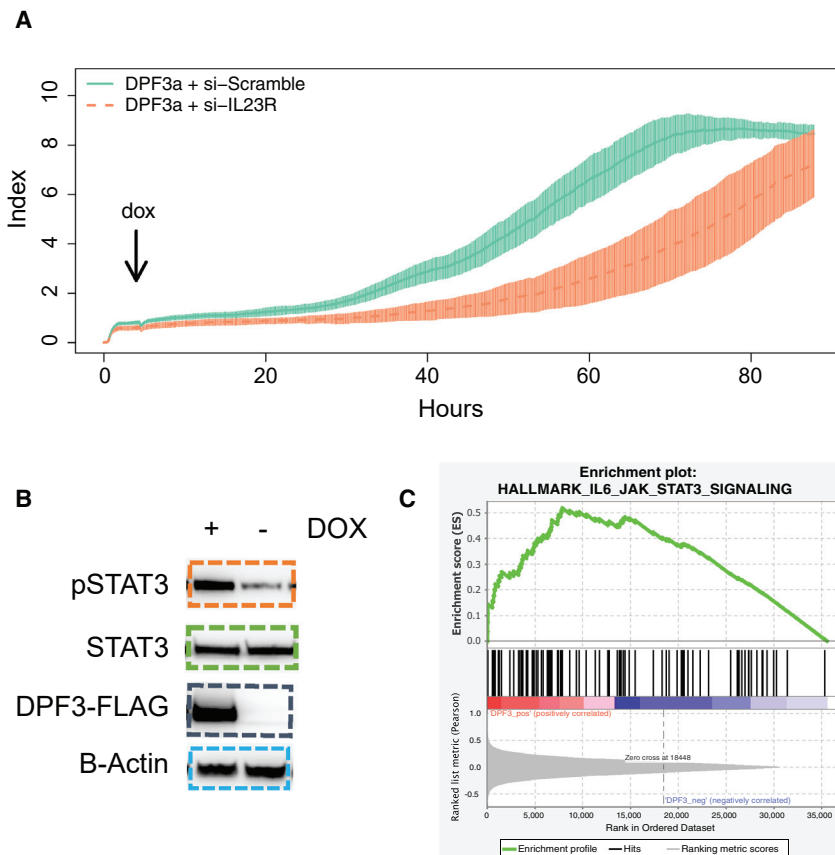
Figure 5. *DPF3* effect on renal cancer is partially mediated by *STAT3* pathway activation

(A) Real-time impedance growth assay for UOK-121-stable dox-inducible DPF3a cell line with si-IL23R or control treatment. Cell lines were plated and treated with doxycycline (dox) or PBS (control) ± siRNA against IL23R after 4 h for evaluation of the growth effect of *DPF3* overexpression without IL23R overexpression.

(B) Immunoblot showing STAT3 phosphorylation after DPF3 overexpression in the HK2 DPF3a cell line. Full gels are shown in Figure S34.

(C) Enrichment plot for pathway analysis shows activation of the STAT3 pathway in four DPF3-stable cell lines (expressing both DPF3 isoforms) with RNA-seq data on gene set enrichment analysis (GSEA).

checkpoint inhibitor efficacy by increasing cytotoxic T cells[88] through PD-1 reduction of CD8+ T cell activity.[89] Reduction of IL-6 with monoclonal antibodies has emerged as a plausible single or combination therapy with a checkpoint inhibitor.[90] In an exploratory analysis of published RNA-seq data from RCC-affected individuals treated with checkpoint inhibitors,[77] we observed a positive but non-significant trend between clinical benefit of checkpoint inhibitors and increased expression of *DPF3* (Figure S42) on *PBRM1* wild-type samples. Together, these data suggest that *DPF3*-dependent reduction of IL-10 and IL-6 could increase CD8+ T activity, enhancing PBMC-mediated cytotoxicity as a potential mechanism for improved response to checkpoint inhibitor therapy. Furthermore, it is plausible that the actual common rs4903064 allele could contribute to the discrepancies in the preliminary reports and merits further investigation.[77–81]

Discussion

Our study provides insight into the underlying mechanism for the 14q24 RCC susceptibility risk locus and offers proof-of-concept for applying high-throughput methods for acceleration of post-GWAS functional studies. We have shown that the allele-specific effect of rs4903064 is mediated through an enhancer within the 14q24 RCC susceptibility locus, altering *DPF3* expression and most likely leading to changes in chromatin accessibility and gene expression as well as inhibition of apoptosis and activation of oncogenic pathways in kidney cells. In studies of overexpression of *DPF3* in stable renal cell lines together with RNA-seq analyses, two of the top deregulated genes are *CEMIP* and *IL23R*, which are notable because *CEMIP* inhibits apoptosis by reducing BAX translocation to the mitochondrial membrane, while IL23R enhances growth by activating the oncogenic STAT3 pathway, important in RCC oncogenesis (Figure 7).

In this study, we specifically show that rs4903064, located in an intron of *DPF3*, displays an allele-specific effect on *DPF3* expression on the basis of MPRAs, luciferase assays, and EMSAs, together with a significant eQTL between rs4903064 and *DPF3* in multiple tissues. The rs4903064-C RCC risk allele creates a HIF-binding site and would be predicted to enhance gene expression, whereas the rs4903064-T protective allele is predicted to increase the affinity for the IRX2/IRX5 binding, which represses *DPF3* transcription. Increased expression of *DPF3* confers a growth advantage to cells by at least two pathways: inhibition of apoptosis via *CEMIP* and activation of STAT3 via *IL23R*. Notably, the *CEMIP* promoter is more accessible after DPF3 overexpression, suggesting that allele-differential effects in germline *DPF3* could influence this pathway, particularly because *CEMIP* overexpression is associated with reduction of apoptosis.

STAT3 pathway activation provides an important clue for connecting germline variation in 14q24 with RCC oncogenesis. *VHL* loss can activate HIF and STAT3 pathways, but *PBRM1* represses oncogenesis by reducing the activation of key pathways. For example, when *PBRM1* function is lost, together with mTORC1 activation, modified SWI/SNF can drive RCC development.[15] Here, we
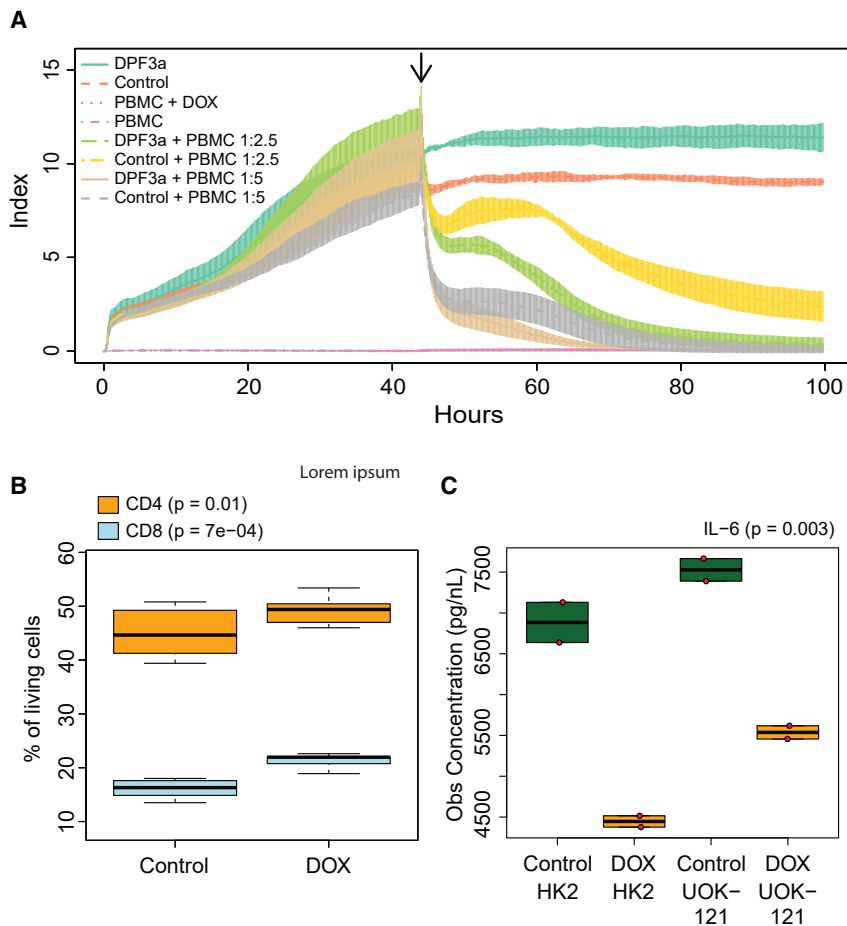
Figure 6. *DPF3* effect on T cell cytotoxicity on renal cancer cell lines

(A) Real-time impedance T cell cytotoxicity assay for the HK2-stable dox-inducible DPF3a cell line with or without activated PBMCs. Cell lines were plated and treated with doxycycline (dox) or PBS (control). After 48 h, activated PMBCs in 1:2.5 and 1:5 ratios relative to initial tumor cells were added to each well. T cell cytotoxicity is quantified as the difference between control (no PBMC) and PBMC samples.

(B) Flow cytometry analysis of PBMCs recovered after the T cell cytotoxicity assay for the HK2 DPF3a cell line for quantification of CD8+ and CD4+ cells. Data shown represent two biological replicates, and the p value is calculated from linear regression including dox/control status and replicate in the model. The black line in the center of boxes represents the median, whiskers extend to the minimum of either the data range or 1.5 times the interquartile range, and statistical outliers are plotted as points.

(C) Immunoassay shows a reduction of IL-6 in media after *DPF3* overexpression in HK2 DPF3a and UOK-121 DPF3a cells. Cells were plated in 6 wells, treated with dox or control, and after 48 h, media was recovered for immunoassay. Analysis was done via regression, including cell line and replicate in the model.

show the importance of *DPF3*-dependent activation of the STAT3 pathway, which could represent an additional mechanism by which SWI/SNF complexes could enhance the effect of *VHL* loss on RCC.[15] Our model suggests that, in part, the 14q24 susceptibility risk locus could be attributed to *DPF3*-dependent STAT3 pathway activation after *VHL* loss, simulating *PBRM1* loss consistent with a multi-step model for RCC development.[15] STAT3 activation is also an important antitumor immune response suppressor,[91,92] and thus, *DPF3*-dependent STAT3 activation could increase RCC risk through immune evasion as a second mechanism, although we did not test *DPF3* effect on immune evasion. Notably, a rare coding variant in *IL23R*, rs11209026, was one of the first GWAS loci found to be associated with inflammatory bowel disease,[75] and its effect is mediated through STAT3 signaling in T cells.[93]

Our investigation of the functional underpinnings of the 14q24 susceptibility locus points toward overexpression of *DPF3* as a disruptive contribution to mSWI/SNF complexes, leading to alterations in nucleosome positioning and downstream gene regulation. *DPF3* is a component of the BAF nucleosome-remodeling complex,[31] which plays a critical role in controlling cell homeostasis, and its impairment contributes to cancer development.[15,33,34] Loss of mSWI/SNF complex mem-

bers can be present in up to 20% of cancer cases[33] and create residual mSWI/SNF complexes that lead to gene deregulation and cancer. Approximately 40% of ccRCC cases show somatic loss of *PBRM1*,[13] a member of the PBAF SWI/SNF complex, underscoring its role in RCC biology[15] as well as treatment response.[77] Moreover, rare familial germline *PBRM1* mutations leading to renal cancer have been described.[33] Given the role of DPF3 in the BAF SWI/SNF complex, our work raises the possibility that DPF3 overexpression could potentially lead to possible enhanced activity. Consistent with this hypothesis, ATAC-seq data show that *DPF3* overexpression changes accessibility of chromatin regions, which are enriched near *DPF3*-deregulated genes. Pathway analyses with the Reactome database[69] revealed enrichment of seven pathways (Figure S4), including "estrogen-dependent gene expression" and "ESR-mediated signaling," which could contribute to the reported difference in RCC signal at 14q24 when stratified by sex.[22]

Because SWI/SNF complexes can regulate enhancer effects, perhaps through *DPF3*-driven chromatin accessibility changes, changes in gene expression patterns due to DPF3 could contribute to RCC. Our data suggest that changes in *DPF3* regulation plausibly could have an impact on BAF function, however the specifics of how this might influence function requires further investigation. We surmise that germline variation could lead to
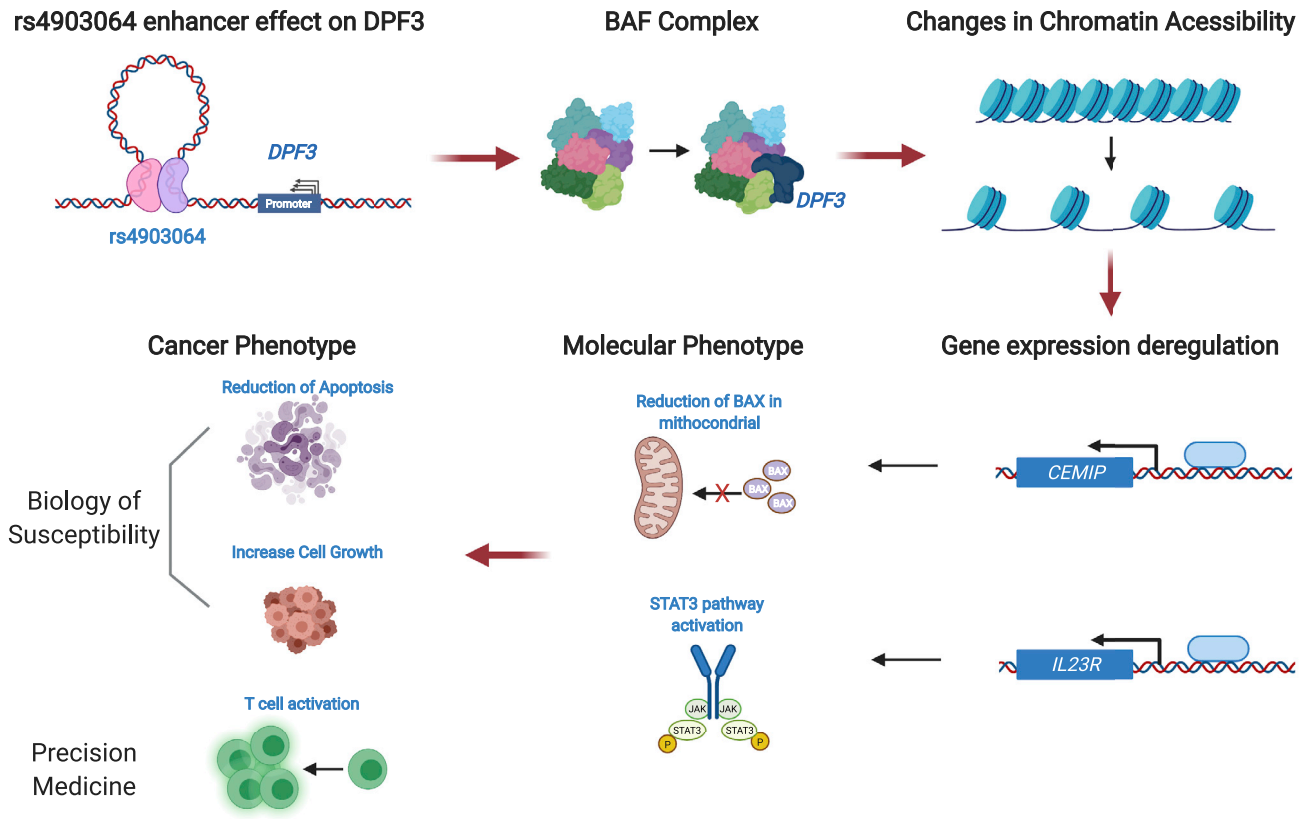
Figure 7. Model for the biology underlying the 14q24 RCC susceptibility locus
The risk-associated rs49030674-C allele creates a binding site for HIF1 and acts as an enhancer, while the protective T allele creates a binding site for IRX2/5 and acts as transcriptional repressor. Therefore, the rs4903064-C risk allele is associated with increasing expression of *DPF3*, which is a member of the BAF SWI/SNF complex. *DPF3* overexpression changes chromatin accessibility, leading to changes in gene expression profile, including overexpression of *CEMIP* and *IL-23R*. *CEMIP* is an anti-apoptosis gene and IL-23R is part of the STAT3 pathway. Downstream, the reduction of apoptosis and activation of the STAT3 pathway, a vital step on RCC oncogenesis, increases the probability of developing renal cancer.

overexpression of *DPF3*, which in turn, replaces *DPF1/2* in the BAF complex and could change both chromatin accessibility and gene expression; DPF3 has recently been shown to be important in kidney development.[94] In RCC, increased DPF3 could either displace or supplement *DPF2*, the most common DPF in the BAF complex,[31,32] and in turn reactivate kidney stem-cell characteristics, leading to RCC, possibly through STAT3 activation and reduced apoptosis. Notably. *DPF3* overexpression activates a similar gene expression pattern to that of *PBRM1* mutation (Figure S37). Nonetheless, we cannot exclude the possibility that DPF3 could work directly or through another mechanism; further work investigating the mechanistic effects of DPF3 is warranted.

We note that a large proportion of ccRCC tumors, particularly in metastatic disease, have been shown to have copy-number loss of the 14q24 region,[8] and loss is thought to drive more aggressive disease at least in part because of loss of *HIF1A* on chromosome band 14q23.2.[8] Most common germline variants influencing risk, such as rs4903064, have relatively small effect sizes, and somatic alterations have a more substantial effect on tumor progression. We hypothesize that this common germline variant

exerts its effect(s) at a relatively early stage, before cancer onset and/or initial cancer development, and additional somatic events, including loss of 14q/*HIF1A*, provide a far greater growth advantage than the effect of this variant.

Our study is not without limitations. We have provided data on the oncogenic effect of DPF3 on RCC susceptibility through changes in chromatic accessibility, but we could not fully evaluate whether this effect is directly due to only DPF3 or mediated through BAF or other changes in SWI/SNF complexes.[94] Tissue-specific analyses have shown that DPF3 is part of BAF in muscle and heart tissue and also present during embryonic kidney development.[31],[94] Consequently, how DPF3 overexpression effects BAF in renal tissue remains an important line of investigation. In addition, although the cell line 786-O showed a number of effects downstream of DPF3 gene expression upon DPF3 overexpression (CEMIP and IL-23R overexpression; activation of SMARCC2 targets), this cell line did not display the DPF3-dependent growth effect observed in three other cell lines. 786-O is known to have chromosomal loss of 14q, and we speculate that this cell line may have developed somatic alterations that activate pathways that are normally DPF3 dependent. Notably, exome sequence analysis

showed that 786-O harbors a PTEN mutation, which could provide both a mitochondria-dependent reduction of apoptosis,[95] separate from DPF3/CEMIP, as well as STAT3 pathway activation independent of DPF3/IL-23R.[96,97] Therefore, PTEN mutation in 786-O could have activated the same targets of DPF3 and limited DPF3 phenotype effect.

In pursuit of understanding the biological basis of the 14q24 RCC locus, additional studies have shed light on how this locus could have implications for RCC treatment. The combination of two checkpoint inhibitors (anti-CTL4 ipilimumab and anti-PD1 nivolumab) showed better overall survival compared to the standard sunitinib (receptor tyrosine kinase inhibitor) treatment,[98] leading to FDA approval for this immunotherapy combination as the first line of treatment for RCC. Two studies identified mutations in the mSWI/SNF complex, particularly *PBRM1*, associated with a favorable response to immunotherapy in RCC.[77,83] Herein, we observed that *DPF3* overexpression enhanced CD8+ T cytotoxic effect on tumor cells, perhaps through reduction of secreted IL-6 and IL-10 by tumor cells. Further mechanistic studies as well as larger clinical trials are required for exploration of this intriguing hypothesis. It is notable that the underlying mechanism for susceptibility risk might also be associated with response to a specific checkpoint treatment, thus opening a new avenue for evaluating *DPF3* expression as a biomarker for immunotherapy response in RCC studies.

Declaration of interests

The authors declare no competing interests.

Web resources

ATAC-seq ENCODE pipeline, https://github.com/kundajelab/atac_dnase_pipelines
BEDTools, https://bedtools.readthedocs.io/en/latest/
BioRender, https://biorender.com/
Biowulf cluster, https://hpc.nih.gov/
Bowtie 2, http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
CHiCAGO, http://regulatorygenomicsgroup.org/chicago
FreeBayes, https://github.com/freebayes/freebayes
GATK, https://gatk.broadinstitute.org/hc/en-us
ISMARA online tool, https://ismara.unibas.ch/mara/
MACS, https://github.com/macs3-project/MACS
Maftools, https://bioconductor.org/packages/release/bioc/html/maftools.html
NovoAlign, http://www.novocraft.com/products/novoalign/
R ALTRE package, https://mathelab.github.io/ALTRE/
Roadmap fetal kidney, https://egg2.wustl.edu/roadmap/data/byFileType/chromhmmSegmentations/ChmmModels/imputed12marks/jointModel/final/E086_25_imputed12marks_stateno.bed.gz
TCGA Research Network, https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga
Trimmomatic, http://www.usadellab.org/cms/?page=trimmomatic
UCSC hg19 transcripts, http://genome.ucsc.edu/
vQTL, https://analysistools.cancer.gov/vqtl/

References

1. Ferlay, J., Soerjomataram, I., Dikshit, R., Eser, S., Mathers, C., Rebelo, M., Parkin, D.M., Forman, D., and Bray, F. (2015). Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int. J. Cancer *136*, E359–E386.

2. Latif, F., Tory, K., Gnarra, J., Yao, M., Duh, F.M., Orcutt, M.L., Stackhouse, T., Kuzmin, I., Modi, W., Geil, L., et al. (1993). Identification of the von Hippel-Lindau disease tumor suppressor gene. Science *260*, 1317–1320.

3. Maxwell, P.H., Wiesener, M.S., Chang, G.-W., Clifford, S.C., Vaux, E.C., Cockman, M.E., Wykoff, C.C., Pugh, C.W., Maher, E.R., and Ratcliffe, P.J. (1999). The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. Nature *399*, 271–275.

4. Schödel, J., and Ratcliffe, P.J. (2019). Mechanisms of hypoxia signalling: new implications for nephrology. Nat. Rev. Nephrol. *15*, 641–659.

5. Harris, A.L. (2002). Hypoxia–a key regulatory factor in tumour growth. Nat. Rev. Cancer *2*, 38–47.

6. Goh, G., Scholl, U.I., Healy, J.M., Choi, M., Prasad, M.L., Nelson-Williams, C., Kunstman, J.W., Korah, R., Suttorp, A.C., Dietrich, D., et al. (2014). Recurrent activating mutation in PRKACA in cortisol-producing adrenal tumors. Nat. Genet. *46*, 613–617.

7. Mitchell, T.J., Turajlic, S., Rowan, A., Nicol, D., Farmery, J.H.R., O'Brien, T., Martincorena, I., Tarpey, P., Angelopoulos, N., Yates, L.R., et al. (2018). Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. Cell *173*, 611–623.e17.

8. Cancer Genome Atlas Research Network (2013). Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature *499*, 43–49.

9. Ricketts, C.J., De Cubas, A.A., Fan, H., Smith, C.C., Lang, M., Reznik, E., Bowlby, R., Gibb, E.A., Akbani, R., Beroukhim, R., et al. (2018). The Cancer Genome Atlas Comprehensive Molecular Characterization of Renal Cell Carcinoma. Cell Rep. *23*, 313–326.e5.

10. Davis, C.F., Ricketts, C.J., Wang, M., Yang, L., Cherniack, A.D., Shen, H., Buhay, C., Kang, H., Kim, S.C., Fahey, C.C., et al. (2014). The somatic genomic landscape of chromophobe renal cell carcinoma. Cancer Cell *26*, 319–330.

11. Linehan, W.M., and Ricketts, C.J. (2019). The Cancer Genome Atlas of renal cell carcinoma: findings and clinical implications. Nat. Rev. Urol. *16*, 539–552.

12. Linehan, W.M., Spellman, P.T., Ricketts, C.J., Creighton, C.J., Fei, S.S., Davis, C., Wheeler, D.A., Murray, B.A., Schmidt, L., Vocke, C.D., et al. (2016). Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma. N. Engl. J. Med. *374*, 135–145.

13. Varela, I., Tarpey, P., Raine, K., Huang, D., Ong, C.K., Stephens, P., Davies, H., Jones, D., Lin, M.-L., Teague, J., et al. (2011). Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. Nature *469*, 539–542.

14. Peña-Llopis, S., Vega-Rubín-de-Celis, S., Liao, A., Leng, N., Pavía-Jiménez, A., Wang, S., Yamasaki, T., Zhrebker, L., Sivanand, S., Spence, P., et al. (2012). BAP1 loss defines a new class of renal cell carcinoma. Nat. Genet. *44*, 751–759.

15. Nargund, A.M., Pham, C.G., Dong, Y., Wang, P.I., Osmangeyoglu, H.U., Xie, Y., Aras, O., Han, S., Oyama, T., Takeda, S., et al. (2017). The SWI/SNF Protein PBRM1 Restrains VHL-Loss-Driven Clear Cell Renal Cell Carcinoma. Cell Rep. *18*, 2893–2906.

16. Chung, C.C., and Chanock, S.J. (2011). Current status of genome-wide association studies in cancer. Hum. Genet. *130*, 59–78.

17. Chanock, S.J. (2020). Susceptibility and somatic patterns. In World Cancer Report, C.P. Wild, E. Weiderpass, and B.W. Stewart, eds. (International Agency for Research on Cancer), pp. 154–163.

18. Han, S.S., Yeager, M., Moore, L.E., Wei, M.-H., Pfeiffer, R., Toure, O., Purdue, M.P., Johansson, M., Scelo, G., Chung, C.C., et al. (2012). The chromosome 2p21 region harbors a complex genetic architecture for association with risk for renal cell carcinoma. Hum. Mol. Genet. *21*, 1190–1200.

19. Wu, X., Scelo, G., Purdue, M.P., Rothman, N., Johansson, M., Ye, Y., Wang, Z., Zelenika, D., Moore, L.E., Wood, C.G., et al. (2012). A genome-wide association study identifies a novel susceptibility locus for renal cell carcinoma on 12p11.23. Hum. Mol. Genet. *21*, 456–462.

20. Henrion, M., Frampton, M., Scelo, G., Purdue, M., Ye, Y., Broderick, P., Ritchie, A., Kaplan, R., Meade, A., McKay, J., et al. (2013). Common variation at 2q22.3 (ZEB2) influences the risk of renal cancer. Hum. Mol. Genet. *22*, 825–831.

21. Purdue, M.P., Ye, Y., Wang, Z., Colt, J.S., Schwartz, K.L., Davis, F.G., Rothman, N., Chow, W.-H., Wu, X., and Chanock, S.J. (2014). A genome-wide association study of renal cell carcinoma among African Americans. Cancer Epidemiol. Biomarkers Prev. *23*, 209–214.

22. Scelo, G., Purdue, M.P., Brown, K.M., Johansson, M., Wang, Z., Eckel-Passow, J.E., Ye, Y., Hofmann, J.N., Choi, J., Foll, M., et al. (2017). Genome-wide association study identifies multiple risk loci for renal cell carcinoma. Nat. Commun. *8*, 15724.

23. Gudmundsson, J., Sulem, P., Gudbjartsson, D.F., Masson, G., Petursdottir, V., Hardarson, S., Gudjonsson, S.A., Johannsdottir, H., Helgadottir, H.T., Stacey, S.N., et al. (2013). A common variant at 8q24.21 is associated with renal cell cancer. Nat. Commun. *4*, 2776.

24. Schödel, J., Bardella, C., Sciesielski, L.K., Brown, J.M., Pugh, C.W., Buckle, V., Tomlinson, I.P., Ratcliffe, P.J., and Mole, D.R. (2012). Common genetic variants at the 11q13.3 renal cancer susceptibility locus influence binding of HIF to an enhancer of cyclin D1 expression. Nat. Genet. *44*, 420–425.

25. Bigot, P., Colli, L.M., Machiela, M.J., Jessop, L., Myers, T.A., Carrouget, J., Wagner, S., Roberson, D., Eymerit, C., Henrion, D., and Chanock, S.J. (2016). Functional characterization of the 12p12.1 renal cancer-susceptibility locus implicates BHLHE41. Nat. Commun. *7*, 12098.

26. Grampp, S., Schmid, V., Salama, R., Lauer, V., Kranz, F., Platt, J.L., Smythies, J., Choudhry, H., Goppelt-Struebe, M., Ratcliffe, P.J., et al. (2017). Multiple renal cancer susceptibility polymorphisms modulate the HIF pathway. PLoS Genet. *13*, e1006872.

27. Grampp, S., Platt, J.L., Lauer, V., Salama, R., Kranz, F., Neumann, V.K., Wach, S., Stöhr, C., Hartmann, A., Eckardt, K.U., et al. (2016). Genetic variation at the 8q24.21 renal cancer susceptibility locus affects HIF binding to a MYC enhancer. Nat. Commun. *7*, 13183.

28. Kadoch, C., Hargreaves, D.C., Hodges, C., Elias, L., Ho, L., Ranish, J., and Crabtree, G.R. (2013). Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. Nat. Genet. *45*, 592–601.

29. Lange, M., Kaynak, B., Forster, U.B., Tönjes, M., Fischer, J.J., Grimm, C., Schlesinger, J., Just, S., Dunkel, I., Krueger, T., et al. (2008). Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. Genes Dev. *22*, 2370–2384.

30. Zeng, L., Zhang, Q., Li, S., Plotnikov, A.N., Walsh, M.J., and Zhou, M.-M. (2010). Mechanism and regulation of acetylated histone binding by the tandem PHD finger of DPF3b. Nature *466*, 258–262.

31. Mashtalir, N., D'Avino, A.R., Michel, B.C., Luo, J., Pan, J., Otto, J.E., Zullow, H.J., McKenzie, Z.M., Kubiak, R.L., St Pierre, R., et al. (2018). Modular Organization and Assembly of SWI/SNF Family Chromatin Remodeling Complexes. Cell *175*, 1272–1288.e20.

32. Mashtalir, N., Suzuki, H., Farrell, D.P., Sankar, A., Luo, J., Filipovski, M., D'Avino, A.R., St Pierre, R., Valencia, A.M., Onikubo, T., et al. (2020). A Structural Model of the Endogenous Human BAF Complex Informs Disease Mechanisms. Cell *183*, 802–817.e24.

33. Helming, K.C., Wang, X., and Roberts, C.W.M. (2014). Vulnerabilities of mutant SWI/SNF complexes in cancer. Cancer Cell *26*, 309–317.

34. Kadoch, C., and Crabtree, G.R. (2015). Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. Sci. Adv. *1*, e1500447–e1500447.

35. Boyle, A.P., Hong, E.L., Hariharan, M., Cheng, Y., Schaub, M.A., Kasowski, M., Karczewski, K.J., Park, J., Hitz, B.C., Weng, S., et al. (2012). Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. *22*, 1790–1797.

36. Ionita-Laza, I., McCallum, K., Xu, B., and Buxbaum, J.D. (2016). A spectral approach integrating functional genomic annotations for coding and noncoding variants. Nat. Genet. *48*, 214–220.

37. Agarwal, V., Bell, G.W., Nam, J.W., and Bartel, D.P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. eLife *4*, e05005.

38. Melnikov, A., Murugan, A., Zhang, X., Tesileanu, T., Wang, L., Rogov, P., Feizi, S., Gnirke, A., Callan, C.G., Jr., Kinney, J.B., et al. (2012). Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. Nat. Biotechnol. *30*, 271–277.

39. Melnikov, A., Zhang, X., Rogov, P., Wang, L., and Mikkelsen, T.S. (2014). Massively parallel reporter assays in cultured mammalian cells. J. Vis. Exp. *90*, 51719.

40. Schödel, J., Grampp, S., Maher, E.R., Moch, H., Ratcliffe, P.J., Russo, P., and Mole, D.R. (2016). Hypoxia, Hypoxia-inducible Transcription Factors, and Renal Cancer. Eur. Urol. *69*, 646–657.

41. Coetzee, S.G., Coetzee, G.A., and Hazelett, D.J. (2015). motifbreakR: an R/Bioconductor package for predicting variant effects at transcription factor binding sites. Bioinformatics *31*, 3847–3849.

42. Kulakovskiy, I.V., Vorontsov, I.E., Yevshin, I.S., Soboleva, A.V., Kasianov, A.S., Ashoor, H., Ba-Alawi, W., Bajic, V.B., Medvedeva, Y.A., Kolpakov, F.A., and Makeev, V.J. (2016). HOCOMOCO: expansion and enhancement of the collection of transcription factor binding sites models. Nucleic Acids Res. *44* (D1), D116–D125.

43. Ongen, H., Buil, A., Brown, A.A., Dermitzakis, E.T., and Delaneau, O. (2016). Fast and efficient QTL mapper for thousands of molecular phenotypes. Bioinformatics *32*, 1479–1485.

44. Hormozdiari, F., van de Bunt, M., Segrè, A.V., Li, X., Joo, J.W.J., Bilow, M., Sul, J.H., Sankararaman, S., Pasaniuc, B., and Eskin, E. (2016). Colocalization of GWAS and eQTL Signals Detects Target Genes. Am. J. Hum. Genet. *99*, 1245–1260.

45. Liu, B., Gloudemans, M.J., Rao, A.S., Ingelsson, E., and Montgomery, S.B. (2019). Abundant associations with gene expression complicate GWAS follow-up. Nat. Genet. *51*, 768–769.

46. Kichaev, G., Yang, W.Y., Lindstrom, S., Hormozdiari, F., Eskin, E., Price, A.L., Kraft, P., and Pasaniuc, B. (2014). Integrating functional data to prioritize causal variants in statistical fine-mapping studies. PLoS Genet. *10*, e1004722.

47. Mirabello, L., Khincha, P.P., Ellis, S.R., Giri, N., Brodie, S., Chandrasekharappa, S.C., Donovan, F.X., Zhou, W., Hicks, B.D., Boland, J.F., et al. (2017). Novel and known ribosomal causes of Diamond-Blackfan anaemia identified through comprehensive genomic characterisation. J. Med. Genet. *54*, 417–425.

48. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics *30*, 2114–2120.

49. Pimentel, H., Bray, N.L., Puente, S., Melsted, P., and Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. Nat. Methods *14*, 687–690.

50. Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nat. Biotechnol. *34*, 525–527.

51. Belaghzal, H., Dekker, J., and Gibcus, J.H. (2017). Hi-C 2.0: An optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. Methods *123*, 56–65.

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

53. Balwierz, P.J., Pachkov, M., Arnold, P., Gruber, A.J., Zavolan, M., and van Nimwegen, E. (2014). ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs. Genome Res. *24*, 869–884.

54. Buenrostro, J.D., Wu, B., Chang, H.Y., and Greenleaf, W.J. (2015). ATAC-seq: A Method for Assaying Chromatin Accessibility Genome-Wide. Curr. Protoc. Mol. Biol. *109*, 21.29.1–21.29.9.

55. Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., and Liu, X.S. (2008). Model-based analysis of ChIP-Seq (MACS). Genome Biol. *9*, R137.

56. Baskin, E., Farouni, R., and Mathé, E.A. (2017). ALTRE: workflow for defining ALTered Regulatory Elements using chromatin accessibility data. Bioinformatics *33*, 740–742.

57. Kulakovskiy, I.V., Vorontsov, I.E., Yevshin, I.S., Sharipov, R.N., Fedorova, A.D., Rumynskiy, E.I., Medvedeva, Y.A., Magana-Mora, A., Bajic, V.B., Papatsenko, D.A., et al. (2018). HOCOMOCO: towards a complete collection of transcription factor binding models for human and mouse via large-scale ChIP-Seq analysis. Nucleic Acids Res. *46* (D1), D252–D259.

58. Qiu, C., Huang, S., Park, J., Park, Y., Ko, Y.-A., Seasock, M.J., Bryer, J.S., Xu, X.-X., Song, W.-C., Palmer, M., et al. (2018). Renal compartment-specific genetic variation analyses identify new pathways in chronic kidney disease. Nat. Med. *24*, 1721–1731.

59. Davies, J.O.J., Telenius, J.M., McGowan, S.J., Roberts, N.A., Taylor, S., Higgs, D.R., and Hughes, J.R. (2016). Multiplexed analysis of chromosome conformation at vastly improved sensitivity. Nat. Methods *13*, 74–80.

60. Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A., Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J., et al. (2015). Integrative analysis of 111 reference human epigenomes. Nature *518*, 317–330.

61. Mumbach, M.R., Satpathy, A.T., Boyle, E.A., Dai, C., Gowen, B.G., Cho, S.W., Nguyen, M.L., Rubin, A.J., Granja, J.M., Kazane, K.R., et al. (2017). Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. Nat. Genet. *49*, 1602–1612.

62. Weintraub, A.S., Li, C.H., Zamudio, A.V., Sigova, A.A., Hannett, N.M., Day, D.S., Abraham, B.J., Cohen, M.A., Nabet, B., Buckley, D.L., et al. (2017). YY1 Is a Structural Regulator of Enhancer-Promoter Loops. Cell *171*, 1573–1588.e28.

63. Ong, C.-T., and Corces, V.G. (2011). Enhancer function: new insights into the regulation of tissue-specific gene expression. Nat. Rev. Genet. *12*, 283–293.

64. Machiela, M.J., and Chanock, S.J. (2015). LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics *31*, 3555–3557.

65. Bilioni, A., Craig, G., Hill, C., and McNeill, H. (2005). Iroquois transcription factors recognize a unique motif to mediate transcriptional repression in vivo. Proc. Natl. Acad. Sci. USA *102*, 14671–14676.

66. Chen, F., Zhang, Y., Şenbabaoğlu, Y., Ciriello, G., Yang, L., Reznik, E., Shuch, B., Micevic, G., De Velasco, G., Shinbrot, E.,

et al. (2016). Multilevel Genomics-Based Taxonomy of Renal Cell Carcinoma. Cell Rep. *14*, 2476–2489.

67. Zhang, J., Wang, C., Chen, X., Takada, M., Fan, C., Zheng, X., Wen, H., Liu, Y., Wang, C., Pestell, R.G., et al. (2015). EglN2 associates with the NRF1-PGC1α complex and controls mitochondrial function in breast cancer. EMBO J. *34*, 2953–2970.

68. Buenrostro, J.D., Giresi, P.G., Zaba, L.C., Chang, H.Y., and Greenleaf, W.J. (2013). Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. Nat. Methods *10*, 1213–1218.

69. Fabregat, A., Jupe, S., Matthews, L., Sidiropoulos, K., Gillespie, M., Garapati, P., Haw, R., Jassal, B., Korninger, F., May, B., et al. (2018). The Reactome Pathway Knowledgebase. Nucleic Acids Res. *46* (D1), D649–D655.

70. Barutcu, A.R., Lian, J.B., Stein, J.L., Stein, G.S., and Imbalzano, A.N. (2017). The connection between BRG1, CTCF and topoisomerases at TAD boundaries. Nucleus *8*, 150–155.

71. Jami, M.-S., Hou, J., Liu, M., Varney, M.L., Hassan, H., Dong, J., Geng, L., Wang, J., Yu, F., Huang, X., et al. (2014). Functional proteomic analysis reveals the involvement of KIAA1199 in breast cancer growth, motility and invasiveness. BMC Cancer *14*, 194.

72. Liang, G., Fang, X., Yang, Y., and Song, Y. (2018). Knockdown of CEMIP suppresses proliferation and induces apoptosis in colorectal cancer cells: downregulation of GRP78 and attenuation of unfolded protein response. Biochem. Cell Biol. *96*, 332–341.

73. Shostak, K., Zhang, X., Hubert, P., Göktuna, S.I., Jiang, Z., Klevernic, I., Hildebrand, J., Roncarati, P., Hennuy, B., Ladang, A., et al. (2014). NF-κB-induced KIAA1199 promotes survival through EGFR signalling. Nat. Commun. *5*, 5232.

74. Wolter, K.G., Hsu, Y.T., Smith, C.L., Nechushtan, A., Xi, X.G., and Youle, R.J. (1997). Movement of Bax from the cytosol to mitochondria during apoptosis. J. Cell Biol. *139*, 1281–1292.

75. Duerr, R.H., Taylor, K.D., Brant, S.R., Rioux, J.D., Silverberg, M.S., Daly, M.J., Steinhart, A.H., Abraham, C., Regueiro, M., Griffiths, A., et al. (2006). A genome-wide association study identifies IL23R as an inflammatory bowel disease gene. Science *314*, 1461–1463.

76. Yang, X.O., Panopoulos, A.D., Nurieva, R., Chang, S.H., Wang, D., Watowich, S.S., and Dong, C. (2007). STAT3 regulates cytokine-mediated generation of inflammatory helper T cells. J. Biol. Chem. *282*, 9358–9363.

77. Miao, D., Margolis, C.A., Gao, W., Voss, M.H., Li, W., Martini, D.J., Norton, C., Bossé, D., Wankowicz, S.M., Cullen, D., et al. (2018). Genomic correlates of response to immune checkpoint therapies in clear cell renal cell carcinoma. Science *359*, 801–806.

78. Braun, D.A., Ishii, Y., Walsh, A.M., Van Allen, E.M., Wu, C.J., Shukla, S.A., and Choueiri, T.K. (2019). Clinical Validation of PBRM1 Alterations as a Marker of Immune Checkpoint Inhibitor Response in Renal Cell Carcinoma. JAMA Oncol. *5*, 1631–1633.

79. Hakimi, A.A., Attalla, K., DiNatale, R.G., Ostrovnaya, I., Flynn, J., Blum, K.A., Ged, Y., Hoen, D., Kendall, S.M., Reznik, E., et al. (2020). A pan-cancer analysis of PBAF complex mutations and their association with immunotherapy response. Nat. Commun. *11*, 4168.

80. Dizman, N., Bergerot, P.G., Bergerot, C.D., Hsu, J., and Pal, S.K. (2019). Duration of treatment (DOT) with targeted therapies (TT) or immunotherapy (IO) in PBRM1 mutated metastatic renal cell carcinoma (mRCC). J. Clin. Oncol. *37*, 622.

81. Hakimi, A.A., Ged, Y., Flynn, J., Hoen, D.R., Di Natale, R.G., Blum, K.A., Makarov, V., Kuo, F., Carlo, M.I., Lee, C.-H., et al. (2019). The impact of PBRM1 mutations on overall survival in greater than 2,100 patients treated with immune checkpoint blockade (ICB). J. Clin. Oncol. *37*, 666.

82. Cormedi, M.C.V., Van Allen, E.M., and Colli, L.M. (2021). Predicting immunotherapy response through genomics. Curr. Opin. Genet. Dev. *66*, 1–9.

83. Pan, D., Kobayashi, A., Jiang, P., Ferrari de Andrade, L., Tay, R.E., Luoma, A.M., Tsoucas, D., Qiu, X., Lim, K., Rao, P., et al. (2018). A major chromatin regulator determines resistance of tumor cells to T cell–mediated killing. Science *359*, 770–775.

84. Tsukamoto, H., Fujieda, K., Miyashita, A., Fukushima, S., Ikeda, T., Kubo, Y., Senju, S., Ihn, H., Nishimura, Y., and Oshiumi, H. (2018). Combined blockade of IL6 and PD-1/PD-L1 signaling abrogates mutual regulation of their immunosuppressive effects in the tumor microenvironment. Cancer Res. *78*, 5011–5022.

85. Laino, A.S., Woods, D., Vassallo, M., Qian, X., Tang, H., Wind-Rotolo, M., and Weber, J. (2020). Serum interleukin-6 and C-reactive protein are associated with survival in melanoma patients receiving immune checkpoint inhibition. J. Immunother. Cancer *8*, 842.

86. Tarhini, A.A., Zahoor, H., Lin, Y., Malhotra, U., Sander, C., Butterfield, L.H., and Kirkwood, J.M. (2015). Baseline circulating IL-17 predicts toxicity while TGF-β1 and IL-10 are prognostic of relapse in ipilimumab neoadjuvant therapy of melanoma. J. Immunother. Cancer *3*, 39.

87. Naing, A., Papadopoulos, K.P., Autio, K.A., Ott, P.A., Patel, M.R., Wong, D.J., Falchook, G.S., Pant, S., Whiteside, M., Rasco, D.R., et al. (2016). Safety, antitumor activity, and immune activation of pegylated recombinant human interleukin-10 (AM0010) in patients with advanced solid tumors. J. Clin. Oncol. *34*, 3562–3569.

88. Ohno, Y., Toyoshima, Y., Yurino, H., Monma, N., Xiang, H., Sumida, K., Kaneumi, S., Terada, S., Hashimoto, S., Ikeo, K., et al. (2017). Lack of interleukin-6 in the tumor microenvironment augments type-1 immunity and increases the efficacy of cancer immunotherapy. Cancer Sci. *108*, 1959–1966.

89. Sun, Z., Fourcade, J., Pagliano, O., Chauvin, J.-M., Sander, C., Kirkwood, J.M., and Zarour, H.M. (2015). IL10 and PD-1 Cooperate to Limit the Activity of Tumor-Specific CD8+ T Cells. Cancer Res. *75*, 1635–1644.

90. Kampan, N.C., Xiang, S.D., McNally, O.M., Stephens, A.N., Quinn, M.A., and Plebanski, M. (2017). Immunotherapeutic Interleukin-6 or Interleukin-6 Receptor Blockade in Cancer: Challenges and Opportunities. Curr. Med. Chem. *25*, 4785–4806.

91. Yu, H., Pardoll, D., and Jove, R. (2009). STATs in cancer inflammation and immunity: a leading role for STAT3. Nat. Rev. Cancer *9*, 798–809.

92. Bromberg, J. (2002). Stat proteins and oncogenesis. J. Clin. Invest. *109*, 1139–1142.

93. Sarin, R., Wu, X., and Abraham, C. (2011). Inflammatory disease protective R381Q IL23 receptor polymorphism results in decreased primary CD4+ and CD8+ human T-cell functional responses. Proc. Natl. Acad. Sci. USA *108*, 9560–9565.

94. Basta, J.M., Singh, A.P., Robbins, L., Stout, L., Pherson, M., and Rauchman, M. (2020). The core SWI/SNF catalytic subunit Brg1 regulates nephron progenitor cell proliferation and differentiation. Dev. Biol. *464*, 176–187.

95. Zhu, Y., Hoell, P., Ahlemeyer, B., and Krieglstein, J. (2006). PTEN: a crucial mediator of mitochondria-dependent apoptosis. Apoptosis 11, 197–207.

96. Zhou, J., Wulfkuhle, J., Zhang, H., Gu, P., Yang, Y., Deng, J., Margolick, J.B., Liotta, L.A., Petricoin, E., 3rd, and Zhang, Y. (2007). Activation of the PTEN/mTOR/STAT3 pathway in breast cancer stem-like cells is required for viability and maintenance. Proc. Natl. Acad. Sci. USA 104, 16158–16163.

97. Yokogami, K., Wakisaka, S., Avruch, J., and Reeves, S.A. (2000). Serine phosphorylation and maximal activation of STAT3 during CNTF signaling is mediated by the rapamycin target mTOR. Curr. Biol. 10, 47–50.

98. Motzer, R.J., Tannir, N.M., McDermott, D.F., Arén Frontera, O., Melichar, B., Choueiri, T.K., Plimack, E.R., Barthélémy, P., Porta, C., George, S., et al. (2018). Nivolumab plus Ipilimumab versus Sunitinib in Advanced Renal-Cell Carcinoma. N. Engl. J. Med. 378, 1277–1290.