PLOS | ONE

RESEARCH ARTICLE

Heterogeneous Effects of Direct Hypoxia Pathway Activation in Kidney Cancer

Rafik Salama, Norma Masson, Peter Simpson, Lina Katrin Sciesielski, Min Sun, Ya-Min Tian, Peter John Ratcliffe*, David Robert Mole*

The Henry Wellcome Building for Molecular Physiology, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

* drmole@well.ox.ac.uk (DRM); pjr@well.ox.ac.uk (PJR)

OPEN ACCESS

Citation: Salama R, Masson N, Simpson P, Sciesielski LK, Sun M, Tian Y-M, et al. (2015) Heterogeneous Effects of Direct Hypoxia Pathway Activation in Kidney Cancer. PLoS ONE 10(8): e0134645. doi:10.1371/journal.pone.0134645

Editor: Jörn Karhausen, Duke University Medical Center, UNITED STATES

Received: May 27, 2015

Accepted: July 10, 2015

Published: August 11, 2015

Copyright: © 2015 Salama et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data have been deposited at the NCBI Gene Expression Omnibus database (GSE67237, http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE67237).

Funding: This work was supported by a Cancer Research UK (grant number A16016) to RS, NM and DRM, the Ludwig Institute for Cancer Research to PJR and MS, the Higher Education Funding Council for England to DRM, the Wellcome Trust (grant numbers 078333/Z/05/Z, WT091857MA) to PJR, YMT and PS, and the German Research Foundation (grant number SC132/2-1) to LKS. The funders had

Abstract

General activation of hypoxia-inducible factor (HIF) pathways is classically associated with adverse prognosis in cancer and has been proposed to contribute to oncogenic drive. In clear cell renal carcinoma (CCRC) HIF pathways are upregulated by inactivation of the von-Hippel-Lindau tumor suppressor. However HIF-1**α** and HIF-2**α** have contrasting effects on experimental tumor progression. To better understand this paradox we examined pan-genomic patterns of HIF DNA binding and associated gene expression in response to manipulation of HIF-1**α** and HIF-2**α** and related the findings to CCRC prognosis. Our findings reveal distinct pan-genomic organization of canonical and non-canonical HIF isoform-specific DNA binding at thousands of sites. Overall associations were observed between HIF-1**α**-specific binding, and genes associated with favorable prognosis and between HIF-2**α**-specific binding and adverse prognosis. However within each isoform-specific set, individual gene associations were heterogeneous in sign and magnitude, suggesting that activation of each HIF-**α** isoform contributes a highly complex mix of pro- and anti-tumorigenic effects.

Introduction

Hypoxia is strongly associated with adverse prognosis in cancer and hypoxia signaling pathways are commonly activated during cancer development[1–4]. This is exemplified by clear cell renal cancer (CCRC) in which inactivation of the von Hippel-Lindau tumor suppressor (pVHL) is a common and early event[5–7]. pVHL is the recognition component of an E3 ubiquitin ligase complex that targets hypoxia inducible factor (HIF) **α**-subunits for degradation by the ubiquitin-proteasome pathway, and inactivation of pVHL results in constitutive activation of the HIF transcriptional pathway[8]. HIF transcriptional targets include many with functions that associate with common phenotypic characteristics of cancer (e.g. enhanced angiogenesis, dysregulated energy metabolism, increased cell motility), and it is commonly assumed that oncogenesis is driven largely by the aggregation of the positive effects of HIF activation on such processes.

Transcriptional activation by HIF is principally mediated through binding of HIF **α**/**β** heterodimers to a core consensus sequence (RCGTG) in hypoxia response elements (HREs)[9]. Although the two best-characterized HIF-**α** subunits, HIF-1**α** and HIF-2**α**, manifest similar

no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

domain architecture, and indistinguishable DNA-binding sequences[10], they transactivate distinct targets and show contrasting tumorigenic roles[11,12]. Surprisingly, strong up-regulation of HIF following inactivation of VHL in CCRC is associated with an unusual bias in HIF isoform expression, towards HIF-2**α**[13]. Several lines of investigation indicate that this is functionally important in the development of CCRC. Genome-wide association studies identified polymorphic variants that affect susceptibility to CCRC at the HIF-2**α**, but not HIF-1**α**, locus and at loci within the HIF-2**α** transcriptional pathway[14]. Conversely, HIF-1**α** gene dosage is commonly reduced in CCRC by loss of chromosome 14q [15], and the HIF-1**α** gene, but not the HIF-2**α** gene, is subject to a small but significant excess of inactivating mutations[16]. Furthermore, HIF-2**α** but not HIF-1**α**, can over-ride the tumor suppressor activity of pVHL in experimental tumor systems[17,18]. Specifically, re-expression of HIF-1**α** in CCRC lines that lack wild-type HIF-1**α** slows growth, whilst overexpression of HIF-2**α** accelerates growth in tumor xenografts[11,15].

These observations challenge the paradigm that general activation of HIF signaling drives oncogenesis through activation of a small number of discrete transcriptional targets, and suggest a more complex interface. For instance, pan-genomic analyses revealed hundreds to thousands of direct HIF transcriptional targets[10,19–22]. Since HIF-1**α** and HIF-2**α** may potentially compete for binding to HIF-1β or for occupancy of HREs, or may bind distinct sets of transcriptional targets, it is unclear how the isoform specific manipulation of HIF-**α** impacts on the transcription patterns associated with CCRC.

To study this we performed detailed ChIP-seq and RNA-seq analysis of HIF-**α** isoform binding site occupancy and gene expression in the pVHL-defective CCRC 786–0 cell line, following re-expression of HIF-1**α** or overexpression of HIF-2**α**, and related these findings to prognostically associated patterns of gene expression in human CCRC tumors[6]. Our findings reveal large numbers of discrete isoform-specific HIF-**α** binding sites that manifest an isoform-specific genomic architecture, activate distinct patterns of gene expression and associate with contrasting prognostic gene expression patterns in clinical CCRC. However, although clear overall associations were observed between HIF-1**α**-associated genes and good clinical prognosis and between HIF-2**α**-associated genes and poor clinical prognosis, this dichotomy was incomplete. At the level of individual genes, the effects were heterogeneous in both sign and size of effect, suggesting that even within this defined context, each HIF-**α** isoform has potentially both pro- and anti-tumorigenic effects.

Materials and Methods

Laboratory methods

Cell Culture. 786-O and HEK293T cells were purchased from ATCC (http://www.lgcstandards-atcc.org) and grown in Dulbecco modified Eagle medium supplemented with 10% fetal calf serum, 2 mM L-glutamine, 100U penicillin and streptomycin 50 U/ml (v/v) (Sigma-Aldrich).

Production of 786–0 HIF-1α/HIF-2α cells. HIF-1**α** and HIF-2**α** cDNA sequences[11] were first cloned into pRRL.IRES.EGFP (kind gift from Kamil Kranc, Glasgow) to generate bicistronic vectors. These were then co-transfected with pCMV-dR8.2 and pCMC-VSVG into HEK293T cells and the resultant viral particles were isolated by centrifugation and ultrafiltration. 786–0 cells (ATCC) were then transduced with either control (pRRL), HIF-1**α** (pRRL-HIF-1**α**) or HIF-2**α** (pRRL-HIF-2**α**) expressing virus.

ChIP-seq. Single ChIP-seq analyses were performed as previously described[19]. Chromatin was immunoprecipitated using rabbit polyclonal antisera to HIF-1**α** (PM14), HIF-2**α** (PM9)[23], HIF-1β (NB-100-110, Novus Biologicals, UK) or pre-immune serum as control.

PolyA+ selected RNA-seq. Total RNA was prepared in triplicate using the mirVana miRNA Isolation Kit (Ambion; Life Technologies Ltd, Paisley, UK) and treated with DNaseI (TURBO DNA-free, Ambion). PolyA+ RNA libraries were then prepared using the ScriptSeq v2 RNA-Seq kit (Epicentre, Madison, WI, USA).

High-throughput sequencing. All libraries were prepared according to Illumina protocols and sequenced on the HiSeq 2000 platform (Illumina, San Diego, CA, USA).

Accession codes. ChIP-seq and RNA-seq data are available from Gene Expression Omnibus (GSE67237).

Bioinformatic analysis of ChIP-seq data

Initial Analysis. Illumina adaptor sequences were trimmed using Trimgalore (0.3.3) and reads were aligned to Genome Reference Consortium GRCh37 (hg19) using BWA (0.7.5a-r405). Low quality mapping was removed (MapQ < 15) using SAMtools (0.1.19)[24] and reads mapping to Duke Encode black list regions (http://hgwdev.cse.ucsc.edu/cgi-bin/hgFileUi?db = hg19&g = wgEncodeMapability) were excluded using BEDTools (2.17.0)[25]. Duplicate reads were marked for exclusion using Picard tools (1.106) (http://picard.sourceforge.net/). Read densities were normalized and expressed as reads per kilobase per million reads (RPKM)[26]. One million random non-overlapping regions selected from ENCODE DNase Cluster II peaks (http://hgdownload.cse.ucsc.edu/goldenPath/hg19/encodeDCC/wgEncodeRegDnaseClustered/) were used as a control.

Peak Calling. ChIP-seq peaks were identified using T-PIC (Tree shape Peak Identification for ChIP-Seq)[27] and MACS (Model-based analysis of ChIP-Seq)[28] in control mode. Peaks detected by both peak callers were filtered quantitatively using the total count under the peak to include only peaks that were above the 99.99th percentile of random background regions selected from the ENCODE DNASE II cluster (p-value < 0.0001).

De-novo Motif Analysis. Sequences flanking each peak summit (±150bp) were repeat-masked using RepeatMasker 4.0.3 (http://www.repeatmasker.org). De-novo motifs were identified using Meme-chip (4.9.1)[29] and matched, using the TomTom module, to known transcription factor motifs in the 2009 JASPAR core database[30].

Principal Component Analysis (PCA). Binding sites for all subunits included in the PCA were merged into one binding set. PCA was performed using Singular Value Decomposition (Prcomp, R 3.1.1 –stats library, http://cran.r-project.org) for both individual binding sites and for the entirety of the ChIP-seq signal for each subunit. Biplots were generated in R.

Heat maps. Binding site heat maps were generated using Ngsplot (2.08)[31] with the parameters: FL = 50, MW = 5, RZ = 1, SC = 0–1, MQ = 15.

Motif Likelihood. Each HIF-**α** binding region (summit ±150bp) was scanned for the HRE (Jaspar/MA0259.1) and the AP-1 (Jaspar/MA0491.1) binding motifs using position weight matrices (PWMs) retrieved from the 2009 JASPAR Core database[30]. The normalized likelihood ratio (NLR) for each motif was calculated according to Eq 1. The maximum normalized likelihood ratio for each binding region was reported for both the HRE and AP-1 motif as given by Eq 2.

$$NLR_j = \frac{1}{L}\sum_{i=1}^{L} \frac{PWM(b,i)}{B(b)}, \qquad \text{Equation (1)}$$

$$P_{NLR} = \max\left(\bigcup_{j=1}^{W-L} NLR_j\right), \qquad \text{Equation (2)}$$

Where ***j*** is the position in the peak, ***i*** is the position in the motif, ***L*** is the length of the motif,

PWM(b,i) is the PWM value at *i* for the corresponding base *b*, ***B(b)*** is the background weight for the corresponding base calculated over a million random DNAse sites and ***W*** is the peak width.

Bioinformatic analysis of RNA-seq data

Initial Analysis. Adapter sequences were trimmed as above. Reads were then aligned to GRCh37 using Tophat 2.0.8b (http://ccb.jhu.edu/software/tophat/index.shtml) and bowtie 1.0.0 (http://bowtie-bio.sourceforge.net/index.shtml) and non-uniquely mapping fragments excluded using SAMtools (0.1.19)[24]. Total read counts for each UCSC defined gene were extracted using HTSeq (0.5.4p3)[32] with 'intersection-strict' mode and significantly regulated genes were identified using DESeq2 (ref. [33]).

Gene Set Enrichment Analysis (GSEA). GSEA enrichment analysis used 10000 permutations, weighted enrichment score and pre-ranking of genes[34]. Both differential expression significance according to DESeq2 and fold-difference between the two conditions were used to rank genes[35] (Eq 3).

$$\pi_i = \varphi_i(-log_{10}pv_i), \qquad \text{Equation 3}$$

Where $\boldsymbol{\varphi_i}$ is the log2 fold-change and $\boldsymbol{pv_i}$ is the p-value for gene *i*.

The Cancer Genome Atlas (TCGA) GSEA. Clinical and RNA-seq V2 data for Kidney Renal Clear Cell Carcinoma (KIRC) patients (https://tcga-data.nci.nih.gov/tcga/) [6] was collated. Patients with missing clinical data or multiple RNA-seq datasets were excluded. Normalized mRNA counts were used as provided. Patients were divided into two groups using a prognostic score, adding 1 for each of; age over 60, pathologic stage 3 or 4, presence of metastasis, or patient deceased. Good prognosis patients scored 0 or 1 and bad prognosis patients scored 2, 3 or 4. The differential expression of each gene between the patient groups was assessed by the Likelihood Ratio Test for negative binomial fitted models using glm.nb in R (3.1.1). Genes were ranked for GSEA based on combined significance and fold-change as above.

Disease Annotation. Binding site enrichment was calculated using -log10 of the binomial test FDR (q-value) by GREAT (Genomic Regions Enrichment of Annotations Tool[36]) 2.0.2. Cancer subtypes with no enrichment in either ChIP-seq dataset were removed.

Binding Gene Predictor. Genes within 10-kb of sites with at least 3-fold difference between HIF-1**α** and HIF-2**α** binding were selected to design a gene predictor using Supervised Principle Component Analysis (SPCA)[37,38]. First, HIF-binding genes demonstrating univariate significance ($p<0.05$) in predicting patient prognosis (Cox proportional hazard model) were selected. For each set of genes (those binding re-introduced HIF-1**α** or overexpressed HIF-2**α**) Singular Value Decomposition (SVD) across all the TCGA patients was used to assign gene weights[37,38]. Each gene predictor was validated using 'leave-one-out' cross-validation, employing 414 patients at a time to generate the gene predictor and then predict the survival of each left-out patient. The final significance and hazard ratio of the patient's risk predictor was assessed similar to univariate analysis.

Results

Firstly, to define the HIF isoform specific activities that are associated with contrasting effects on growth in CCRC-derived VHL-defective 786–0 cells, pools of cells were infected with viruses expressing empty vector (VA), wild-type HIF-1**α**, or wild-type HIF-2**α**, and pan-genomic patterns of HIF-binding and gene expression were analyzed by ChIP-seq and RNA-seq. HIF-1**α** and HIF-2**α** infections achieved roughly equal mRNA levels that were approximately

10-fold higher than the endogenous HIF-2**α** mRNA level (S1A Fig). Similarly, HIF-2**α** protein levels were approximately 8-fold higher in the HIF-2**α** infected cells than in the control cells, whilst HIF-1**α** protein levels were 15-20-times higher in HIF-1**α** infected cells than in another CCRC cell line (RCC4), which expresses full-length HIF-1**α** (S1B Fig). In control cells, a total of 1719 peaks binding endogenous HIF-2**α** were identified. As anticipated, these sites show a high level of concordance with HIF-1**β** signal, consistent with binding of a HIF-2**α**/1**β** heterodimer (S2 Fig).

Our RNA-seq analysis confirmed previous findings[15] that, whilst 786–0 cells do not express wild-type HIF-1**α**, they do express several truncated and/or fusion transcripts encompassing HIF-1**α** exons 1–9, but not more distal sequences (S3 Fig), and are predicted to be incapable of encoding a transcriptionally active HIF-1**α**. Re-expression of full-length wild-type HIF-1**α** in 786–0 cells has a negative effect on growth as tumor xenografts in mice. We therefore began by examining pan-genomic patterns of DNA binding for HIF-1**α**, HIF-1**β** and HIF-2**α** in such cells, and compared these with control 786–0 cells.

Re-expression of HIF-1α does not antagonize HIF-2α binding

Since HIF-1**α** and HIF-2**α** both dimerize with HIF-1**β**, and recognize a similar consensus DNA sequence, but have opposing effects in tumor xenograft growth in 786–0, we first considered the possibility that re-expressed HIF-1**α** might antagonize HIF-2**α** DNA-binding, either through direct competition for binding sites or through competition for HIF-1**β**. Somewhat surprisingly, genome-wide analysis revealed that HIF-2**α** binding was little affected by re-expression of HIF-1**α** (Fig 1A, compare i and iii). Consistent with this, principal component analysis (PCA) of HIF-2**α** binding in HIF-1**α** re-expressing cells demonstrated strong co-variance with HIF-2**α** binding in control cells (Fig 1B, compare vectors HIF-2**α**(VA) and HIF-2**α** (1**α**RE)). These analyses suggest that HIF-2**α** binding is largely unaffected by HIF-1**α** re-expression. To test this more quantitatively, we compared the relative strength of HIF-2 binding signals in HIF-1**α** re-expressing, versus control cells, across the 1719 endogenous HIF-2**α** binding sites identified in the control cells. This analysis revealed a tight correlation centered on equity of binding in the two cell lines, for both HIF-2**α** binding and for HIF-1**β** binding (S4 Fig). Thus, under these conditions, HIF-1**α** re-expression does not appear to globally antagonize HIF-2**α** by disrupting HIF-2**α** binding, either through direct competition for binding sites, or by competing HIF-1**β** away from HIF-2**α**.

Extensive binding of re-expressed HIF-1α at new functional sites

In contrast to limited effects on existing HIF-2 binding, re-expression of HIF-1**α** led to extensive new binding across the genome, with specific HIF-1**α** signal identified at 5147 sites. Marked correlation was observed between HIF-1**α** and HIF-1**β** binding in these HIF-1**α** re-expressing cells (Fig 1A, compare iv and v). In keeping with this, the PCA confirmed strong covariance between HIF-1**α** and HIF-1**β** in HIF-1**α** re-expressing cells indicating that these sites are largely distinct from those binding endogenous HIF-2**α** (Fig 1B, compare vectors HIF-1**α**(1**α**RE) and HIF-1**β**(1**α**RE) and contrast with HIF-2**α**(VA) and HIF-1**β**(VA)). Taken together, these findings indicate that following its re-expression HIF-1**α** binds to a large number of sites distinct from endogenous HIF-2**α** sites, most likely as a HIF-1**α**/HIF-1**β** heterodimer. Furthermore, total HIF-1**β** chromatin immunoprecipitation signal (number of reads mapping to peak regions) increased approximately 2–3 fold following re-expression of HIF-1**α** (also see S5 Fig). This suggests that HIF-1**β** protein abundance is not limiting for endogenous HIF binding in VHL-defective 786–0 cells, and that additional HIF-1**β** can be recruited to DNA binding sites following increased expression of a dimerization partner.

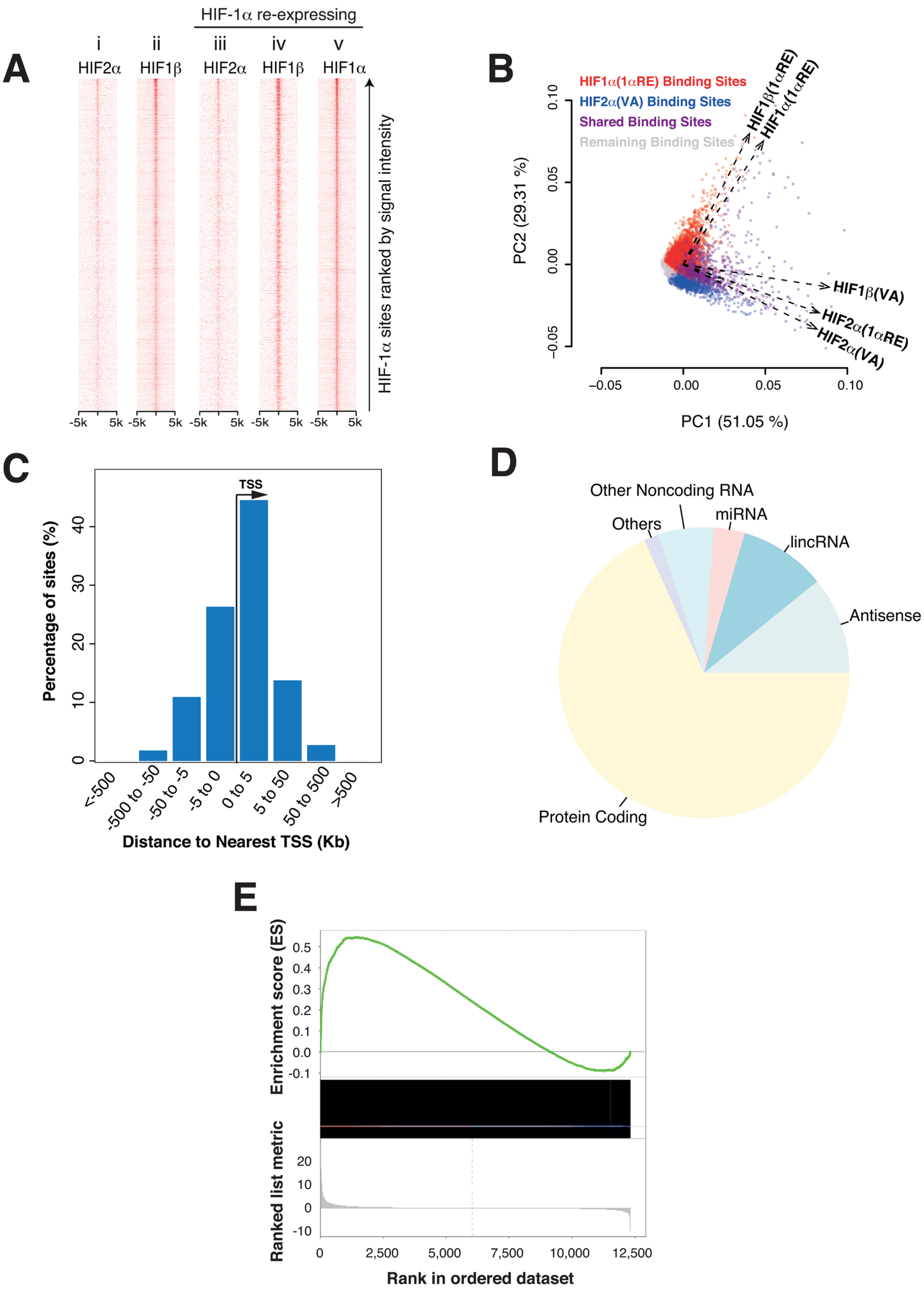
A
HIF-1α re-expressing
i
ii
iii
iv
v
HIF2α
HIF1β
HIF2α
HIF1β
HIF1α
HIF-1α sites ranked by signal intensity
-5k
5k
B
HIF1α(1αRE) Binding Sites
HIF2α(VA) Binding Sites
Shared Binding Sites
Remaining Binding Sites
HIF1β(1αRE)
HIF1α(1αRE)
HIF1β(VA)
HIF2α(1αRE)
HIF2α(VA)
PC2 (29.31 %)
PC1 (51.05 %)
−0.05
0.00
0.05
0.10
C
TSS
Percentage of sites (%)
0
10
20
30
40
<-500
-500 to -50
-50 to -5
-5 to 0
0 to 5
5 to 50
50 to 500
>500
Distance to Nearest TSS (Kb)
D
Other Noncoding RNA
Others
miRNA
lincRNA
Antisense
Protein Coding
E
Enrichment score (ES)
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
Ranked list metric
20
10
0
-10
0
2,500
5,000
7,500
10,000
12,500
Rank in ordered dataset

Fig 1. HIF-1α re-expression in 786-O cells. (A) HIF-1**α** binding sites in the HIF-1**α** re-expressing cells were identified by peak calling and ranked on the vertical axis according to signal intensity. Heat maps of these sites (±5kb on horizontal axis) showing ChIP-seq read density for the indicated HIF subunits were generated for both the control cells (i, ii) and the HIF-1**α** re-expressing cells with HIF-1**α** re-introduced (iii-v). The pattern of HIF-2**α** binding is minimally affected by the re-expression of full-length HIF-1**α** (compare i and iii). Sites binding re-expressed HIF-1**α** are largely co-occupied by HIF-1**β** (compare iv and v). (B) Biplot showing Principal Component Analysis (PCA) of ChIP-seq signal intensity (RPKM values) for both individual binding sites (dots) and HIF-subunits (vectors) across all HIF-binding sites identified in control cells and in HIF-1**α** re-expressing cells. Sites binding endogenous HIF-2**α** in control cells are shown in blue while sites binding re-expressed HIF-1**α** are shown in red, sites binding both are colored purple and the remaining sites are shown in grey. PCA for each subunit shows high co-variance between HIF-2**α** binding in the control cells and in the HIF-1**α** re-expressing cells (compare HIF2**α**(VA) and (HIF2**α**(1**α**RE)). This indicates only minimal change in the HIF-2**α** binding as a consequence of the HIF-1**α** re-expression. Conversely, the HIF-1**β** vector changes dramatically with HIF-1**α** re-expression (compare HIF1**β**(VA) with HIF1**β**(1**α**RE)) and aligns closely with the vector for re-expressed HIF-1**α** (HIF1**α** (1**α**RE)). The individual binding sites in the control and HIF-1**α** re-expressing cells (blue and red dots) aligned closely with their respective PCA vectors. (C) Histogram of the distance to nearest transcription start site (TSS) for HIF-1**α** binding sites in cells re-expressing HIF-1**α**. (D) HIF-1**α** binding sites in the re-expressing cells were categorized according to the class (Ensemble) of the nearest gene. The relative frequency of each class is show by pie chart. (E) Gene set enrichment analysis (GSEA) for the set of genes nearest to HIF-1**α** binding sites when genes are ranked according to fold-change and significance in mRNA expression following re-expression of HIF-1**α** (horizontal axis).

doi:10.1371/journal.pone.0134645.g001

Analysis of the HIF-1**α** binding indicated that its pan-genomic distribution was strikingly non-random, being strongly enriched close to particular classes of gene promoter (Fig 1C and 1D) and closely resembling HIF-1**α**-specific patterns of binding in other cell types [10,20,22,39]. Annotation of the HIF-1**α** 'nearest-neighbor' promoters, according to transcript class, revealed that 68% are associated with protein-coding genes, with the remainder largely associated with long non-coding RNAs (lncRNAs) and antisense RNAs (Fig 1D); proportions that are similar to those reported for HIF-1**α** in MCF7 breast cancer cells[39].

We next sought to define the functional effects of this new HIF-1**α** binding on gene expression. Pan-genomic gene expression was profiled in control and HIF-1**α** re-expressing cells by RNA-seq. Genes were ranked according to a combination of fold-change in transcript level following re-expression of HIF-1**α** and the significance of this change[35]. Using this ranking, gene set enrichment analysis (GSEA)[34] of the nearest gene to each HIF-1**α** binding site demonstrated a highly significant ($p<0.001$) association between HIF-1**α** binding and positive, but not negative, regulation of these transcripts (Fig 1E). This is consistent with previous findings that HIF acts predominantly as a transcriptional activator[10,19]. It also indicates that the new HIF-1**α** binding is functional, and has direct and largely positive effects on expression across large numbers of genes.

In contrast, we observed few effects of HIF-1**α** re-expression on transcript levels of (nearest neighbor) HIF-2**α** binding genes. GSEA demonstrated a positive, but non-significant ($p = 0.15$) effect of HIF-1**α** re-expression on the expression of HIF-2**α** binding genes (S6 Fig), presumably as a direct effect of HIF-1**α** binding to some of these sites (Fig 3C). Importantly, there was no evidence of large-scale down-regulation of HIF-2**α** binding genes (i.e. no enrichment of HIF-2**α** binding genes amongst the downregulated genes) as a consequence of HIF-1**α** re-expression in 786-O cells. This confirms the finding that re-expressed HIF-1**α** does not significantly antagonize HIF-2**α** activity across the genome.

Overexpression of HIF-2α further activates its endogenous targets

Since overexpression of HIF-2**α** in 786-O cells has the opposite effect on tumor xenograft growth to re-expression of wild-type HIF-1**α**[11], we next examined the effect of increasing HIF-2**α** levels on HIF binding and gene expression. We first sought to distinguish whether binding of HIF-2**α** in VHL defective 786-O cells is saturated or whether overexpression of HIF-2**α** would increase binding at sites that are already occupied in control cells. To test this, we correlated the strength of HIF-2**α** binding signals in cells overexpressing transfected HIF-2**α** with that in the control cells. This revealed an increase in both HIF-2**α** signals and HIF-1**β** signals at the majority of endogenous HIF-2**α** binding sites following HIF-2**α** overexpression

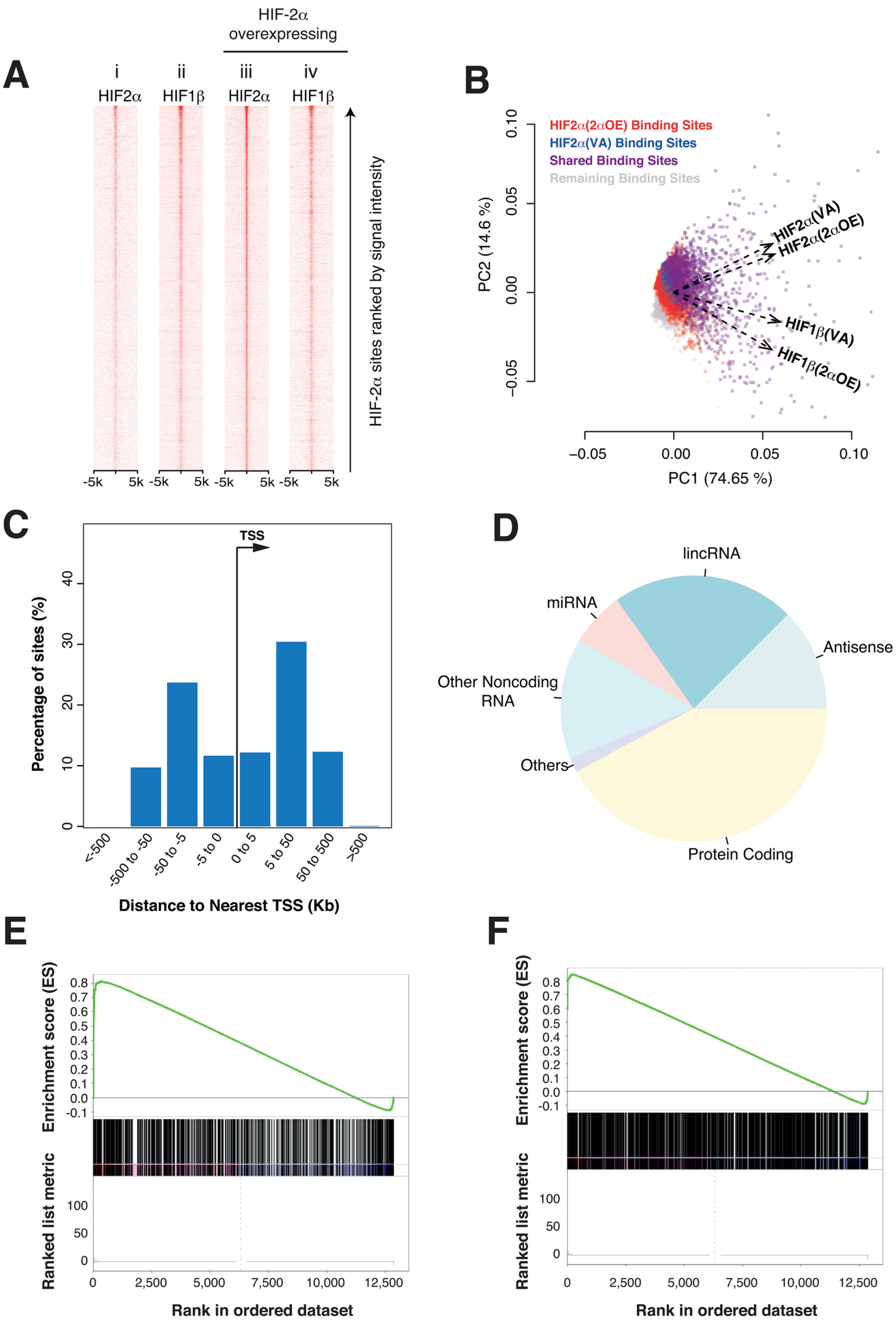
A
HIF-2α overexpressing
i
ii
iii
iv
HIF2α
HIF1β
HIF2α
HIF1β
HIF-2α sites ranked by signal intensity
-5k
5k
B
HIF2α(2αOE) Binding Sites
HIF2α(VA) Binding Sites
Shared Binding Sites
Remaining Binding Sites
HIF2α(VA)
HIF2α(2αOE)
HIF1β(VA)
HIF1β(2αOE)
PC2 (14.6 %)
PC1 (74.65 %)
−0.05
0.00
0.05
0.10
C
TSS
Percentage of sites (%)
0
10
20
30
40
<-500
-500 to -50
-50 to -5
-5 to 0
0 to 5
5 to 50
50 to 500
>500
Distance to Nearest TSS (Kb)
D
lincRNA
miRNA
Antisense
Other Noncoding RNA
Others
Protein Coding
E
F
Enrichment score (ES)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
Ranked list metric
100
50
0
0
2,500
5,000
7,500
10,000
12,500
Rank in ordered dataset

Fig 2. HIF-2α overexpression in 786-O cells. (A) HIF-2α binding sites in the HIF-2α overexpressing cells were identified by peak calling and ranked on the vertical axis according to signal intensity. Heat maps of these sites (±5kb on horizontal axis) showing ChIP-seq read density for the indicated HIF subunits were generated for both the control cells (i, ii) and the cells with HIF-2α overexpressed (iii, iv). In contrast to re-expressed HIF-1α, overexpressed HIF-2α binds to a large number of sites (compare i and iii), without HIF-1β (compare iii and iv) and has little effect on the distribution of HIF-1β (compare ii and iv). (B) Biplot showing Principal Component Analysis (PCA) of ChIP-seq signal intensity (RPKM values) for both individual binding sites (dots) and HIF-subunits (vectors) across all HIF-binding sites identified in control cells and in HIF-2α overexpressing cells. Sites binding endogenous HIF-2α in control cells are shown in blue while sites binding re-expressed HIF-1α are shown in red, sites binding both are colored purple and the remaining sites are colored grey. PCA for HIF subunits shows that HIF-2α and HIF-1β co-vary more closely in the control cells (compare HIF2α(VA) and HIF1β(VA)) than in the overexpressing cells (compare HIF2α(2αOE) and HIF1β(2αOE)). (C) Histogram of the distance to nearest transcription start site (TSS) for HIF-2α binding sites in cells overexpressing HIF-2α. (D) HIF-2α binding sites in the HIF-2α overexpressing cells were categorized according to the class (Ensemble) of the nearest gene. The relative frequency of each class is shown by pie chart. Gene set enrichment analysis (GSEA) for the set of genes nearest to (E) HIF-2α binding sites in the control cells and (F) newly identified HIF-2α binding sites in the overexpressing cells, when genes are ranked according to fold-change and significance in mRNA expression following overexpression of HIF-2α (horizontal axis).

doi:10.1371/journal.pone.0134645.g002

(S7 Fig) indicating that endogenous levels of HIF-2**α** in 786–0 cells are not sufficient to fully load these HIF-2**α** binding sites.

Extensive non-canonical binding of overexpressed HIF-2α

Remarkably however, pan-genomic analysis revealed a total of 5283 discrete HIF-2**α** binding sites in the 786–0 cells overexpressing HIF-2**α**. These sites include many of those occupied by endogenous HIF-2**α**. In contrast to the findings following re-expression of HIF-1**α**, heat map analysis of HIF-2**α** and HIF-1**β** binding indicated that many of these sites are occupied solely by HIF-2**α**, without HIF-1**β** (Fig 2A, compare iii and iv). Concordant with this, PCA analysis revealed that transfected HIF-2**α** and HIF-1**β** binding co-vary less extensively in cells overexpressing transfected HIF-2**α** than in the control cells (Fig 2B).

As with HIF-1**α** binding in re-expressing cells, the distribution of transfected HIF-2**α** binding sites was strikingly non-uniform across the genome (Fig 2C and 2D). Interestingly, despite the large increase in number of sites they conformed to a pattern similar to that observed for endogenous HIF-2**α** in both these (S8 Fig) and non-CCRC (MCF7) cells[10,39], but distinct from that for HIF-1**α**. In contrast with HIF-1**α**, the distribution was more promoter-distal (Fig 2C). Furthermore, as with the endogenous HIF-2**α** binding patterns, annotation of 'nearest neighbor' genes revealed that a substantially greater proportion were at non-coding gene loci, than was observed for re-introduced or endogenous HIF-1**α** (Fig 2D). These differences between HIF-**α** isoforms, which were observed irrespective of cell type expression level or number of sites bound, suggest that the ability to recognize promoter-distant enhancers at non-coding gene loci versus promoter proximal sites is an innate property of HIF-**α** isoforms.

Finally, we sought to identify whether increased loading of existing HIF-2**α** binding sites, or binding of HIF-2**α** at newly detected sites, or both, is associated with functional effects on gene expression. We used GSEA to test enrichment of HIF-binding loci (nearest-neighbor genes) amongst genes whose expression was changed by HIF-2**α** overexpression (as measured by RNA-seq). We performed this analysis for both the binding sites that were occupied by endogenous HIF-2**α**, and those that were newly observed following HIF-2**α** overexpression. An association with positive, but not negative, effects on gene expression was observed for both groups of binding loci (Fig 2E and 2F). Thus, constitutive activation of HIF-2**α** in 786–0 cells is submaximal and increasing HIF-2**α** leads to both quantitative and qualitative effects on HIF target gene expression.

Interestingly, close inspection of the apparently new binding sites frequently revealed low levels of HIF binding at these sites in the control cells, which were below the thresholds for identification of significant binding, but above those observed at a set of control regions (S9A Fig). This suggests that these signals are in fact 'real' and reflect uneven loading of potential HIF binding sites (S9B Fig).

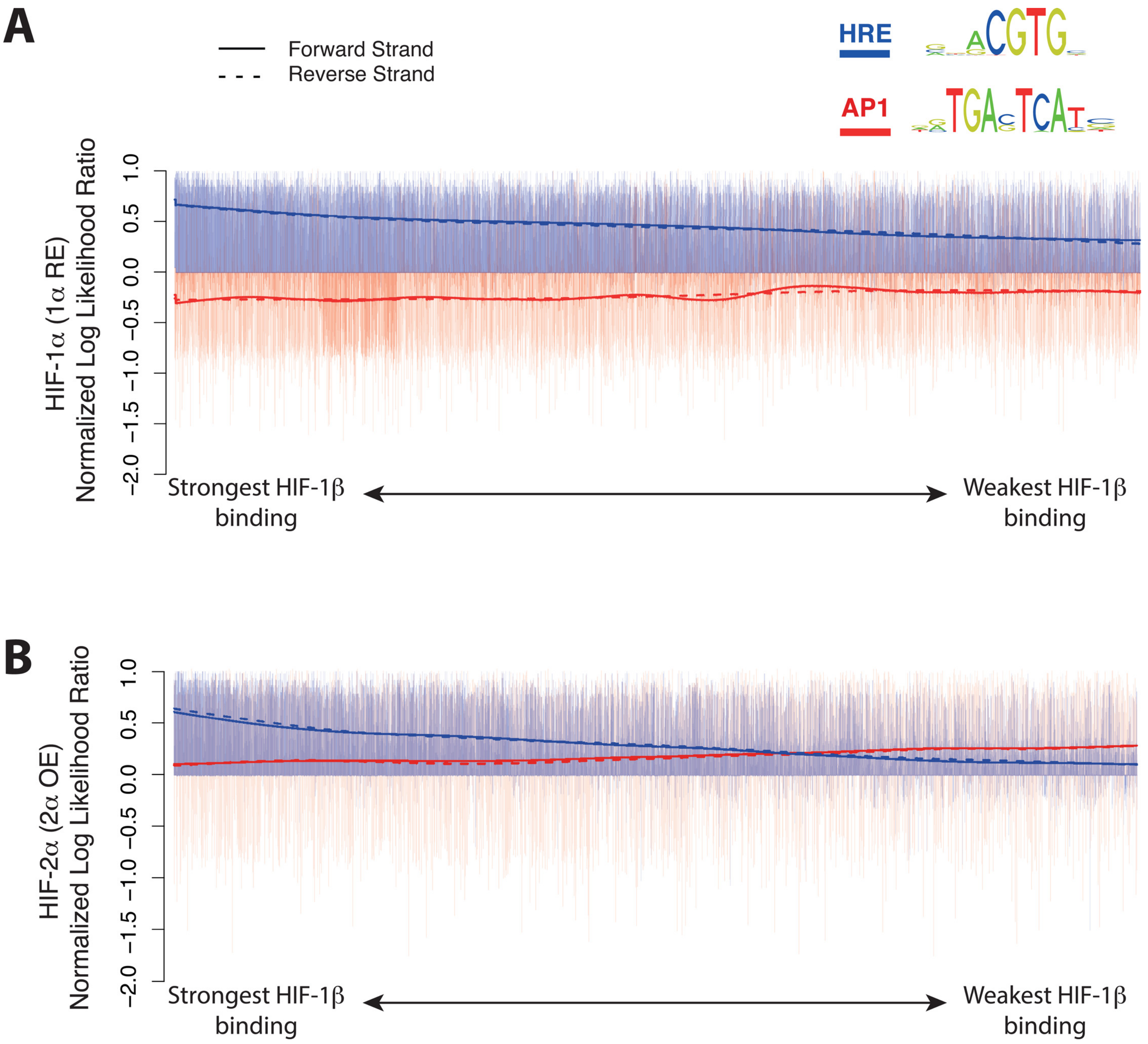

Fig 3. Preferential distribution of AP-1 binding motifs at HIF-2α versus HIF-1α binding loci. In addition to a hypoxia response element (HRE) motif, analysis of sites binding endogenous and overexpressed HIF-2α also identified an AP-1 motif. For each site, the maximum normalized log likelihood ratio for the AP-1 motif in red and the HRE motif in blue is plotted on the vertical axis as a bar chart. A smooth spline cubic fit line is overlaid to show the trend. The smoothing parameter is automatically determined using a 'leave-one-out' cross validation as implemented by the Smooth.spline function in R. Sites were categorized as binding (A) re-expressed HIF-1**α**, (B) overexpressed HIF-2**α** and ranked according to the HIF-1**β** signal at each site. Spline fit curves are overlaid (solid/dashed lines) to indicate overall trends across both forward and reverse strands. (A) Sites binding re-expressed HIF-1**α** show specific enrichment (positive score) for the HRE motif that decreases as the HIF-1**β** signal falls. In contrast, these same sites show depletion of the AP-1 motif. (B) Sites binding overexpressed HIF-2**α** show enrichment of the HRE motif that declines more steeply as the HIF-1**β** signal falls. In contrast to sites binding re-expressed HIF-1**α**, those binding overexpressed HIF-2**α** show enrichment of the AP-1 motif that increases (and exceeds that seen for the HRE) as the HIF-1**β** signal falls.

doi:10.1371/journal.pone.0134645.g003

Analysis of HIF-2α and HIF-1α binding motifs

Since HIF-1 and HIF-2 are reported to recognize an identical core DNA binding sequence[10], it is unclear why binding of HIF-1**α** and HIF-2**α** is distributed so differently. To address this we started by analyzing the immunoprecipitated sequences for transcription factor binding motifs using MEME-ChIP. As expected, the canonical HRE motif (Jaspar/MA0259.1)[40] was enriched in all sets of binding sites; those occupied by endogenous HIF-2**α**, overexpressed HIF-2**α** and re-expressed HIF-1**α** and enrichment was positively correlated with the strength of the HIF-1β signal (Fig 3A and 3B and S8C Fig, blue line). The second most enriched motif was an AP-1 (TRE) binding motif (Jaspar/MA0491.1)[40]. Interestingly, this enrichment was quite differently distributed between HIF-1**α** and HIF-2**α** binding loci. Strong enrichment was observed in sequences binding overexpressed HIF-2**α** (Fig 3B). However in marked contrast to the HRE motif, in the binding set for over-expressed HIF-2**α**, enrichment of the AP-1 motif was inversely correlated with the strength of HIF-1**β** binding (Fig 3B). Furthermore, in the set of HIF-1**α** binding loci, the AP-1 binding motif was significantly depleted (Fig 3A). Overall the differential distribution of this AP-1 motif between HIF-1**α** and HIF-2**α** binding loci was highly significant ($p < 10^{-16}$). Taken together these findings suggest that the AP-1 binding motif is important in the differential distribution of HIF-1**α** versus HIF-2**α** binding sites across the genome at least in this setting.

Direct HIF-isoform-specific transactivation is associated with opposing prognosis in human kidney cancer

Many studies reported overall associations between HIF-dependent gene expression and adverse prognosis in cancer (reviewed in[3,41]), but there is little data that relates prognosis to HIF-isoform specific transcriptional targets and to isoform-specific HIF-binding patterns. We therefore sought to explore associations between HIF isoform-specific targeting defined in 786–0 cells, and genes expression patterns associated with clinical prognosis derived from The Cancer Genome Atlas (TCGA- https://tcga-data.nci.nih.gov/tcga/) [6].

We first stratified 415 CCRC patients, for whom both clinical data and tumor gene expression data were available, into good and poor prognosis groups based on a series of commonly used clinical criteria (survival, histological grade, metastasis, age)[42], (S10A Fig). From this stratification we ranked each gene, based on the significance and fold-difference in its median expression level, between the good and poor prognosis groups. We then sought to determine how HIF-**α** isoform specific changes in gene expression in 786–0 cells associated with this rank using GSEA. These analyses revealed striking isoform specificity in the associations. Genes that were upregulated by HIF-1**α** re-expression in 786-O cells were enriched ($p = 0.01$) amongst those genes that were more highly expressed in the good prognosis tumors (S10B Fig). Conversely, genes that were downregulated by HIF-1**α** were enriched ($p<0.001$) amongst those genes more highly expressed in the poor prognosis tumors (S10C Fig). Comparable analysis for HIF-2**α** regulated genes revealed a non-significant tendency for HIF-2**α** up-regulated genes to be more highly expressed in poor prognosis tumors, and for HIF-2**α** down-regulated genes to have a lower expression in poor prognosis tumors (S10D and S10E Fig). Taken together, these results indicate that the HIF-**α** isoform specific interventions in 786-O cells are relevant to clinical CCRC, and support the existence of opposing effects of HIF-1**α** and HIF-2**α** in a clinical context.

Since the differential gene expression observed in 786-O cells is determined at least in part by disparate HIF-**α** subunit binding, we next sought to relate these differences in binding to cancer pathways. Gene ontology annotation of genes neighboring HIF-binding sites revealed a strong bias towards cancer association for HIF-binding genes, which was stronger for HIF-2**α** than HIF-1**α** associated transcripts (Fig 4A), in keeping with the pro- and anti-tumorigenic actions of HIF-2**α** and HIF-1**α** in this setting.

We next examined whether differential HIF-binding in 786-O cells could provide a predictive model of patient prognosis[37,38] in the TCGA CCRC dataset. First, genes with overlapping HIF-1**α** and HIF-2**α** binding (defined by signal intensities within a 3-fold difference) and those lacking a significant individual (univariate) association with prognosis were excluded. In total, 93 genes (44 in the HIF-1**α** binding group and 49 in the HIF-2**α** binding group) showed individual associations with prognosis that were significant. For each HIF-**α** isoform there was a mixture of genes associated with both a favorable and a poor prognosis. However, combined gene signatures generated for either the HIF-1**α** or HIF-2**α** binding genes were both significantly associated with overall patient survival (Fig 4B and 4C and S11 Fig). Specifically, the HIF-1**α** binding gene signature was associated with higher gene expression in good prognosis patients (Fig 4B and S11A Fig) and conversely that generated for the HIF-2**α** binding genes showed higher expression in the poor prognosis patients (Fig 4C and S11B Fig). These findings indicate that although HIF-1**α** and HIF-2**α** transcriptional targets may have heterogeneous effects on tumor behavior, the overall balance of HIF-1**α** transcriptional targets favors a good prognosis, while the balance of HIF-2**α** targets confers a poor prognosis.

Discussion

Our findings reveal remarkable complexity in the HIF transcriptional response and its relationship to VHL-associated CCRC. Despite an identical consensus recognition sequence[10], large numbers of sites with distinct HIF-isoform specific binding preferences were defined. Our data suggests that it is this, rather than competition for binding at the same chromatin sites, or competition for dimerization with HIF-**β**, that underlies the contrasting HIF-isoform specific effects on experimental and clinical CCRC. It also points to a role for other transcription factors such as AP-1 in directing differential pan-genomic patterns of HIF-**α** binding, at least in this setting. Therefore differences in HIF-binding patterns and/or additional transcriptional activator activity may (at least in part) explain the different effects of HIF-1 and HIF-2 in different types of cancer.

Several systematic differences were observed in patterns of HIF-**α** binding. First, although as expected, HIF-**α** binding was generally associated with HIF-1**β** binding, when overexpressed, this was more robust for HIF-1**α** than HIF-2**α**, which appeared to be capable of binding without HIF-1**β**, particularly in association with AP-1 motifs. This suggests that non-canonical interactions of HIF-2**α** with AP-1 may be important[43], as has been reported for the interaction of HIF-**α** with Myc proteins[44,45]. Further analysis will be required to understand the precise nature of such interactions and their relationship to the biology of CCRC. However it is of interest that pVHL-inactivation has been reported to activate AP-1, in part due to up-regulation of an atypical protein kinase c-JunB pathway[46] and that CCRC-specific regions of accessible chromatin are highly enriched for an AP-1 motif[47]. Second, distinct pan-genomic patterns of HIF-**α** binding were observed that conform strikingly to those previously observed in other contexts. HIF-2**α** bound more distant from promoters and was more likely to be associated with non-coding RNAs than HIF-1**α** binding, the extent of bias being similar to that recently been described for the endogenous HIF-**α** proteins in MCF7 breast cancer cells[10,39]. This suggests that these patterns reflect some intrinsic property of the protein rather than expression level or cellular context.

A somewhat surprising finding in this work was the much larger number of discrete HIF-**α** binding sites observed after re-expression of HIF-1**α** or enhanced expression of HIF-2**α**, as compared to previous reports of the pan-genomic distribution of HIF binding by our

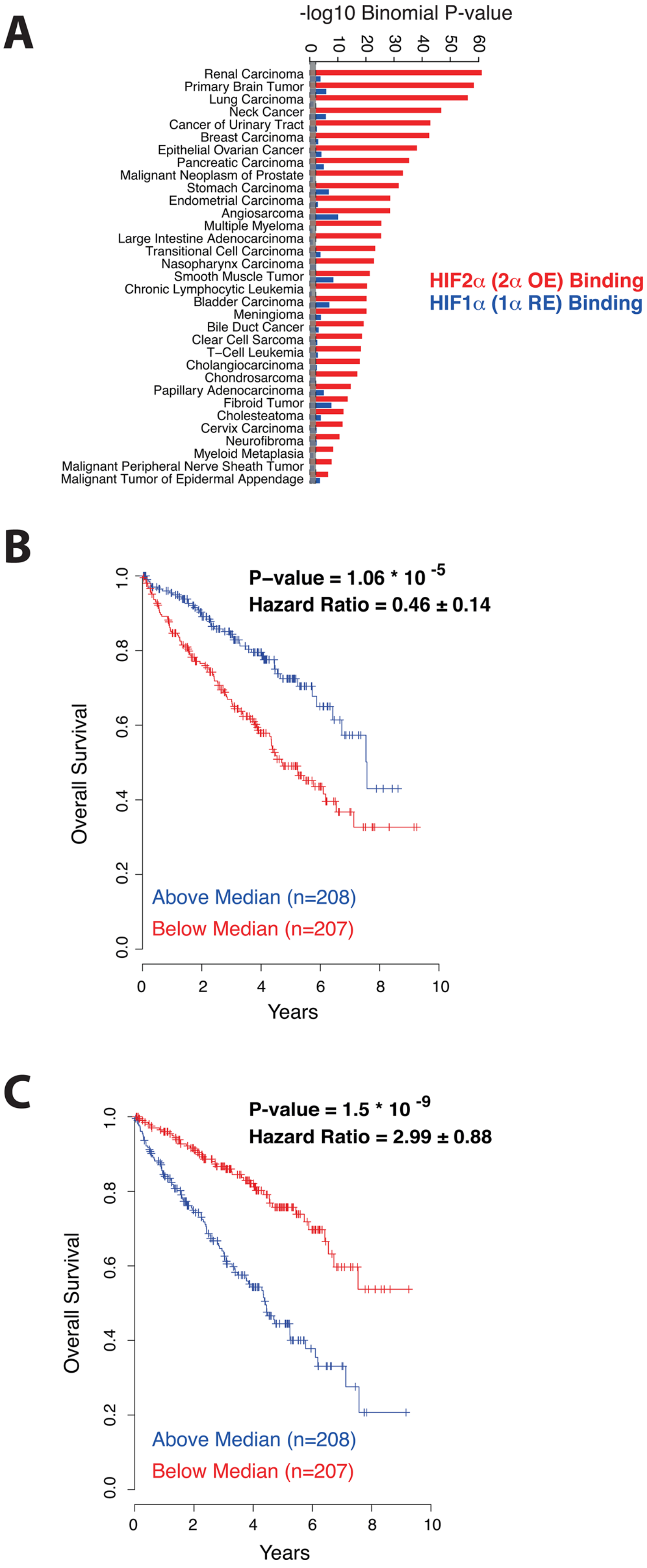
A
-log10 Binomial P-value
0
10
20
30
40
50
60
Renal Carcinoma
Primary Brain Tumor
Lung Carcinoma
Neck Cancer
Cancer of Urinary Tract
Breast Carcinoma
Epithelial Ovarian Cancer
Pancreatic Carcinoma
Malignant Neoplasm of Prostate
Stomach Carcinoma
Endometrial Carcinoma
Angiosarcoma
Multiple Myeloma
Large Intestine Adenocarcinoma
Transitional Cell Carcinoma
Nasopharynx Carcinoma
Smooth Muscle Tumor
Chronic Lymphocytic Leukemia
Bladder Carcinoma
Meningioma
Bile Duct Cancer
Clear Cell Sarcoma
T−Cell Leukemia
Cholangiocarcinoma
Chondrosarcoma
Papillary Adenocarcinoma
Fibroid Tumor
Cholesteatoma
Cervix Carcinoma
Neurofibroma
Myeloid Metaplasia
Malignant Peripheral Nerve Sheath Tumor
Malignant Tumor of Epidermal Appendage
HIF2α (2α OE) Binding
HIF1α (1α RE) Binding
B
P−value = 1.06 * 10 -5
Hazard Ratio = 0.46 ± 0.14
Overall Survival
0.0
0.2
0.4
0.6
0.8
1.0
Above Median (n=208)
Below Median (n=207)
0
2
4
6
8
10
Years
C
P-value = 1.5 * 10 -9
Hazard Ratio = 2.99 ± 0.88
Overall Survival
0.0
0.2
0.4
0.6
0.8
1.0
Above Median (n=208)
Below Median (n=207)
0
2
4
6
8
10
Years

Fig 4. HIF-1α and HIF-2α binding genes confer opposing prognosis in kidney cancer. (A) The genes nearest to re-expressed HIF-1α (blue bars) and overexpressed HIF-2α (red bars) binding sites were defined and examined for enrichment amongst genes annotated in different cancers using the Human Disease Ontology database (http://www.disease-ontology.org).–log10 Binomial p-values are plotted for each set of HIF-binding genes in each type of cancer. Grey bar denotes p = 0.05 (-log10, 1.3) level of significance. HIF-2α nearest binding genes are consistently more significantly enriched amongst cancer-associated genes than are HIF-1α binding genes. (B) Differential HIF-1α binding genes or (C) differential HIF-2α binding genes were filtered for significant associations with overall survival and used to generate a weighted gene predictor of prognosis for each set of genes. Patients were then divided into those with above or below median values for each gene predictor and subjected to Kaplan-Meier survival analysis. The Cox proportional hazard model indicated a significant protective effect for patients with above median gene predictor values based on the HIF-1α binding genes. Conversely, patients with above median values for the HIF-2α binding gene predictor had a significantly worse prognosis.

doi:10.1371/journal.pone.0134645.g004

laboratory and others[10,20,22,48,49]. Close inspection of these apparently new binding sites frequently revealed low, sub-threshold, levels of HIF ChIP-seq signal, which likely reflect uneven loading of potential HIF binding sites. This raises the interesting possibility that uneven binding site loading during graded hypoxic induction may be important in shaping the HIF response. Further work will be required to explore the physiological consequences of this in normal and cancer contexts.

The associations between binding of HIF-**α** in 786–0 cells and prognostic patterns of gene expression in human CCRC, clearly supports the relevance of these patterns to the clinical behavior of CCRC. Clear overall associations between HIF-1**α** binding and genes specifying good prognosis were mirrored by associations between specific HIF-2**α** binding and genes specifying poor prognosis, supporting the hypothesis that differential HIF-**α** binding patterns are driving contrasting effects on CCRC outcomes. Importantly however gene specific analyses revealed considerable heterogeneity in these effects. Individual gene analyses revealed that both the HIF-1**α** and the HIF-2**α**-associated gene-predictor each encompassed genes associated with both good and bad prognosis, indicating that each HIF-**α** isoform could potentially activate a spectrum of pro- and anti-tumourigenic effects. Given the prevalence of hypoxia and upregulation of the HIF system in cancer, understanding the factors that control, and potentially re-balance, HIF-**α** target gene repertoires as cancer develops will clearly be of importance.

Supporting Information

S1 Fig. Expression levels of endogenous and transfected HIF-1α and HIF-2α in 786-O cells. (A) Normalized read counts (RPKM–Reads Per Kilobase per Million reads) showing endogenous and transfected mRNA levels for HIF-1**α** (red) and HIF-2**α** (blue). Note endogenous HIF-1**α** is truncated (see S3 Fig), whilst re-expressed HIF-1**α** is full-length. (B) Immunoblot of HIF-1**α**, HIF-2**α** and **β**-actin in 786-O cells infected with control and HIF-**α** expressing viruses and in RCC4 cells, which express full-length HIF-1**α**. Quantitation of band intensities revealed HIF-2**α** protein levels to be roughly 8-times higher in the HIF-2**α** infected cells than in the control cells and HIF-1**α** protein levels to be 15-20-fold higher in the HIF-1**α** infected cells when compared to RCC4 cells.
(EPS)

S2 Fig. Binding of endogenous HIF-2α and HIF-1β in control 786-O cells. (A) Scatterplot showing good correlation between the intensity of ChIP-seq signal (RPKM count) for HIF-1**β** (y-axis) and that for HIF-2**α** (x-axis) at the endogenous HIF-2**α** peaks. (B) Spatial distribution of HIF-2**α** and HIF-1**β** ChIP-seq signal (RPKM values) centred on the endogenous HIF-2**α** peaks (maximum HIF-2**α** signal ± 1kb) showing that HIF-2**α** and HIF-1**β** signals co-localize

precisely.
(EPS)

S3 Fig. 786-O cells express truncated HIF-1α, which is able to bind DNA. (A) RNA-seq alignment (red) of HIF-1**α** transcripts in control 786-O cells showing read density across the gene (annotated in blue). The schematic (black) relates the genomic and domain structures of HIF-1**α** (bHLH–basic helix-loop-helix domain, PAS–PER, AHR, ARNT, SIM domain, N–N-terminal transactivation domain, ID–inhibitory domain, C–C-terminal transactivation domain). The RNA-seq demonstrates an absence of reads mapping to exons 13, 14 and 15 encoding the C-terminal activation domain. Exons 11 and 12 encoding the N-terminal activation domain are greatly reduced in read number. In addition, there is abnormal splicing of intron 10, and therefore any transcripts containing exons 11 and 12 will be incorrectly translated C-terminal to exon 10. Paired-end mating of reads (black dotted lines) confirms the existence of fusion transcripts containing both HIF-1**α** and SNAPC1 sequences, presumably as a result of deletion of the HIF1A gene transcriptional termination site.
(EPS)

S4 Fig. Re-expressed HIF-1α does not antagonize HIF-2α binding. (A) A scatter plot showing HIF-2**α** ChIP-seq signal (RPKM counts) at sites binding endogenous HIF-2**α** in the control cells. HIF-2**α** signal intensity in cells re-expressing HIF-1**α** (y-axis) is plotted against that for HIF-2**α** in the control cells (x-axis). The majority of HIF-2**α** signals in both the control and HIF-1**α** re-expressing cells correlate positively and are clustered around the (dotted) line representing equal HIF-2**α** binding in the two conditions. (B) A scatter plot showing HIF-1**β** signals at the same sites in both HIF-1**α** re-expressing and control cells. Similar to HIF-2**α**, HIF-1**β** binding signals show a positive correlation between the two conditions clustered around the (dotted) line representing equal HIF-1**β** binding in the endogenous and HIF-1**α** re-expressing cells.
(EPS)

S5 Fig. The effect of HIF-1α re-expression on HIF-1β binding distribution. The scatter plot shows ChIP-seq signal (RPKM count) for HIF-1**β** in the HIF-1**α** re-expressing cells (y-axis) versus HIF-1**β** in the control cells (x-axis) at sites binding endogenous HIF-2**α** in the control cells (blue) and at sites binding re-expressed HIF-1**α** (red). The HIF-1**β** signal shows a clear separation between the HIF-2**α** and HIF-1**α** binding sites. Specifically HIF-1**α** re-expression increases the HIF-1**β** signal at sites binding re-expressed HIF-1**α** (i.e. the red dots cluster above the dotted line representing equity of binding), whilst at the sites binding endogenous HIF-2**α** in the control cells, the HIF-1**β** signal was little altered by HIF-1**α** re-expression (i.e. the blue dots are clustered around the dotted line representing equal binding).
(EPS)

S6 Fig. Re-expressed HIF-1α does not globally antagonize transactivation of HIF-2α binding genes. Gene set enrichment analysis (GSEA) for the set of nearest-neighbour genes that are closest to the endogenous HIF-2**α** binding sites in control cells. All genes are ranked according to their fold-change and significance following re-expression of HIF-1**α** (horizontal axis). Genes nearest to endogenous HIF-2**α** binding sites are enriched amongst the genes upregulated by re-expressed HIF-1**α**, but there is little enrichment amongst the genes downregulated by re-expressed HIF-1**α**.
(EPS)

S7 Fig. Binding of endogenous HIF-2α in control cells is not fully saturated. (A) A scatter plot showing HIF-2**α** ChIP-seq signal (RPKM counts) at sites binding endogenous HIF-2**α**.

The signal intensity in cells expressing transfected HIF-2**α** (y-axis) is plotted against that in the control cells (x-axis). HIF-2**α** binding signal intensity is positively correlated under the two conditions, and is generally increased in cells expressing transfected HIF-2**α** (i.e. the majority of signals cluster above the dotted line representing equal binding). (B) A similar plot showing HIF-1**β** ChIP-seq signal at the same set of binding sites exhibiting a comparable increase in HIF-1**β** binding in cells expressing transfected HIF-2**α**.
(EPS)

S8 Fig. Binding distribution of endogenous HIF-2α in control 786-O cells. (A) Histogram of the distance to nearest transcription start site (TSS) for HIF-2**α** binding sites in control cells. (B) HIF-2**α** binding sites in the control cells were categorized according to the class (Ensemble) of the nearest gene. The relative frequency of each class is shown by pie chart. (C) Analysis of sites binding endogenous HIF-2**α** identified an AP-1 motif as well as the HRE motif. For each site, the maximum normalized log likelihood ratio for the AP-1 motif in red and the HRE motif in blue is plotted on the vertical axis as a bar chart. Sites were ranked according to the HIF-1**β** signal at each site. Spline fit curves are overlaid (solid/dashed lines) to indicate overall trends across both forward and reverse strands. Enrichment of the HRE motif decreases and enrichment of the AP-1 motif increases as the HIF-1**β** signal falls.
(EPS)

S9 Fig. Low-level binding of endogenous HIF-2α at sites binding over-expressed HIF-2α. (A) Sites binding HIF-2**α** in the overexpressing 786–0 cells were ranked (x-axis) according to HIF-2**α** signal intensity (displayed in green on y-axis). These same sites were then examined in the control 786–0 cells for binding of endogenous HIF-2**α** (ranked on the x-axis and displayed in blue on the y-axis). Although the endogenous HIF-2**α** signals were lower than those for transfected HIF-2**α**, signals below the peak-calling threshold were observed at these sites in the control cells. To determine whether these were above background levels, we used ENCODE Faire-seq data from 786-O cells (GSM1011120) to randomly select an identical number of open chromatin sites (excluding HIF-2**α** sites). These sites were also ranked (x-axis) according to their endogenous HIF-2**α** binding signal (displayed in red on the y-axis). The binding signal for endogenous HIF-2**α** in control cells was considerably greater at sites identified as binding HIF-2**α** in the overexpressing cells than at the randomly selected accessible sites. This indicates that the majority of newly detected HIF-2**α** binding sites in HIF-2**α** overexpressing cells are in fact weak HIF-2**α** binding sites in the control cells. (B) Sites binding any isoform across all the datasets were ranked according to HIF-2**α** RPKM counts in control cells (indicated by the red line). The signal intensity for HIF-2**α** in the overexpressing cells was then plotted (blue dots) together with a spline smoothed (blue) trend line. When the signal intensity in the overexpressing cells is compared with the control cells, HIF-2**α** overexpression is seen to lead to a global but uneven increase of the HIF-2**α** signal. Specifically, there was a greater fold-increase in binding amongst the weaker sites than was observed amongst the stronger sites.
(EPS)

S10 Fig. Expression of HIF-1α transcriptional targets correlates with clinical prognosis in CCRC. RNA-seq data from 415 patients studied by The Cancer Genome Atlas (TCGA) was analysed as described in methods. Based on clinical criteria, patients were divided into two prognostic groups illustrated in (A) a Kaplan-Meier plot of overall survival for each group (red = good prognosis, blue = poor prognosis). Genes were then ranked according to the expression fold-difference and significance (see Eq 3 in methods) between the good prognosis tumours and the poor prognosis tumours. (B) GSEA for the set of genes induced by re-expressing HIF-1**α** in 786-O cells showed that they were generally more highly expressed in the good

prognosis tumours. (C) GSEA for genes suppressed by transfected HIF-1**α** in 786-O cells revealed that they were more highly expressed in the poor prognosis tumours. The same analysis was performed for (D) genes induced and (E) genes suppressed by overexpressed HIF-2**α** showing an opposing, although non-significant association.
(EPS)

S11 Fig. HIF-1α and HIF-2α binding gene predictors. The genes within 10kb of both differential (> 3-fold difference in intensity) HIF-1**α** and HIF-2**α** binding sites were tested using Cox proportional hazard model for their ability to predict patient prognosis. Genes with a p-value of <0.05 in the univariate analyses were then used to generate a multivariate predictor for both the HIF-1**α** and HIF-2**α** as described in the methods. The genes and their respective weights as fitted by the gene predictor model are shown for (A) a HIF-1**α** predictor and (B) a HIF-2**α** predictor. The blue bars show genes upregulated in good prognosis patients and the red bars show genes upregulated in bad prognosis patients.
(EPS)

Author Contributions

Conceived and designed the experiments: RS NM PJR DRM. Performed the experiments: NM PS LKS MS. Analyzed the data: RS PJR DRM. Contributed reagents/materials/analysis tools: NM PS LKS YMT. Wrote the paper: RS PJR DRM.

References

1. Hockel M, Vaupel P (2001) Tumor hypoxia: definitions and current clinical, biologic, and molecular aspects. J Natl Cancer Inst 93: 266–276. PMID: 11181773
2. Bertout JA, Patel SA, Simon MC (2008) The impact of O2 availability on human cancer. Nat Rev Cancer 8: 967–975. doi: 10.1038/nrc2540 PMID: 18987634
3. Semenza GL (2010) Defining the role of hypoxia-inducible factor 1 in cancer biology and therapeutics. Oncogene 29: 625–634. doi: 10.1038/onc.2009.441 PMID: 19946328
4. Ratcliffe PJ (2013) Oxygen sensing and hypoxia signalling pathways in animals: the implications of physiology for cancer. J Physiol 591: 2027–2042. doi: 10.1113/jphysiol.2013.251470 PMID: 23401619
5. Kim WY, Kaelin WG (2004) Role of VHL gene mutation in human cancer. J Clin Oncol 22: 4991–5004. PMID: 15611513
6. Cancer Genome Atlas Research N (2013) Comprehensive molecular characterization of clear cell renal cell carcinoma. Nature 499: 43–49. doi: 10.1038/nature12222 PMID: 23792563
7. Gerlinger M, Rowan AJ, Horswell S, Larkin J, Endesfelder D, Gronroos E, et al. (2012) Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. N Engl J Med 366: 883–892. doi: 10.1056/NEJMoa1113205 PMID: 22397650
8. Kaelin WG Jr., Ratcliffe PJ (2008) Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. Mol Cell 30: 393–402. doi: 10.1016/j.molcel.2008.04.009 PMID: 18498744
9. Wenger RH, Stiehl DP, Camenisch G (2005) Integration of oxygen signaling at the consensus HRE. Sci STKE 2005: re12.
10. Schodel J, Oikonomopoulos S, Ragoussis J, Pugh CW, Ratcliffe PJ, Mole DR (2010) High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. Blood 117: e207–217.
11. Raval RR, Lau KW, Tran MG, Sowter HM, Mandriota SJ, Li JL, et al. (2005) Contrasting Properties of Hypoxia-Inducible Factor 1 (HIF-1) and HIF-2 in von Hippel-Lindau-Associated Renal Cell Carcinoma. Mol Cell Biol 25: 5675–5686. PMID: 15964822
12. Hu CJ, Wang LY, Chodosh LA, Keith B, Simon MC (2003) Differential roles of hypoxia-inducible factor 1alpha (HIF-1alpha) and HIF-2alpha in hypoxic gene regulation. Mol Cell Biol 23: 9361–9374. PMID: 14645546
13. Mandriota SJ, Turner KJ, Davies DR, Murray PG, Morgan NV, Sowter HM, et al. (2002) HIF activation identifies early lesions in VHL kidneys: evidence for site-specific tumor suppressor function in the nephron. Cancer Cell 1: 459–468. PMID: 12124175

14. Purdue MP, Johansson M, Zelenika D, Toro JR, Scelo G, Moore LE, et al. (2011) Genome-wide association study of renal cell carcinoma identifies two susceptibility loci on 2p21 and 11q13.3. Nat Genet 43: 60–65. doi: 10.1038/ng.723 PMID: 21131975

15. Shen C, Beroukhim R, Schumacher SE, Zhou J, Chang M, Signoretti S, et al. (2011) Genetic and Functional Studies Implicate HIF1alpha as a 14q Kidney Cancer Suppressor Gene. Cancer Discov 1: 222–235. doi: 10.1158/2159-8290.CD-11-0098 PMID: 22037472

16. Dalgliesh GL, Furge K, Greenman C, Chen L, Bignell G, Butler A, et al. (2010) Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature 463: 360–363. doi: 10.1038/nature08672 PMID: 20054297

17. Kondo K, Kim WY, Lechpammer M, Kaelin WG Jr. (2003) Inhibition of HIF2alpha is sufficient to suppress pVHL-defective tumor growth. PLoS Biol 1: E83. PMID: 14691554

18. Maranchie JK, Vasselli JR, Riss J, Bonifacino JS, Linehan WM, Klausner RD (2002) The contribution of VHL substrate binding and HIF1-alpha to the phenotype of VHL loss in renal cell carcinoma. Cancer Cell 1: 247–255. PMID: 12086861

19. Mole DR, Blancher C, Copley RR, Pollard PJ, Gleadle JM, Ragoussis J, et al. (2009) Genome-wide association of hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha DNA binding with expression profiling of hypoxia-inducible transcripts. J Biol Chem 284: 16767–16775. doi: 10.1074/jbc.M901790200 PMID: 19386601

20. Xia X, Lemieux ME, Li W, Carroll JS, Brown M, Liu XS, et al. (2009) Integrative analysis of HIF binding and transactivation reveals its role in maintaining histone methylation homeostasis. Proc Natl Acad Sci U S A 106: 4260–4265. doi: 10.1073/pnas.0810067106 PMID: 19255431

21. Xia X, Kung AL (2009) Preferential binding of HIF-1 to transcriptionally active loci determines cell-type specific response to hypoxia. Genome Biol 10: R113. doi: 10.1186/gb-2009-10-10-r113 PMID: 19828020

22. Tanimoto K, Tsuchihara K, Kanai A, Arauchi T, Esumi H, Suzuki Y, et al. (2010) Genome-wide identification and annotation of HIF-1alpha binding sites in two cell lines using massively parallel sequencing. Hugo J 4: 35–48. doi: 10.1007/s11568-011-9150-9 PMID: 22132063

23. Lau KW, Tian YM, Raval RR, Ratcliffe PJ, Pugh CW (2007) Target gene selectivity of hypoxia-inducible factor-alpha in renal cancer cells is conveyed by post-DNA-binding mechanisms. Br J Cancer 96: 1284–1292. PMID: 17387348

24. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. (2009) The Sequence Alignment/Map format and SAMtools. Bioinformatics 25: 2078–2079. doi: 10.1093/bioinformatics/btp352 PMID: 19505943

25. Dale RK, Pedersen BS, Quinlan AR (2011) Pybedtools: a flexible Python library for manipulating genomic datasets and annotations. Bioinformatics 27: 3423–3424. doi: 10.1093/bioinformatics/btr539 PMID: 21949271

26. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5: 621–628. doi: 10.1038/nmeth.1226 PMID: 18516045

27. Hower V, Evans SN, Pachter L (2011) Shape-based peak identification for ChIP-Seq. BMC Bioinformatics 12: 15. doi: 10.1186/1471-2105-12-15 PMID: 21226895

28. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, et al. (2008) Model-based analysis of ChIP-Seq (MACS). Genome Biol 9: R137. doi: 10.1186/gb-2008-9-9-r137 PMID: 18798982

29. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, et al. (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37: W202–208. doi: 10.1093/nar/gkp335 PMID: 19458158

30. Portales-Casamar E, Thongjuea S, Kwon AT, Arenillas D, Zhao X, Valen E, et al. (2010) JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. Nucleic Acids Res 38: D105–110. doi: 10.1093/nar/gkp950 PMID: 19906716

31. Shen L, Shao N, Liu X, Nestler E (2014) ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. BMC Genomics 15: 284. doi: 10.1186/1471-2164-15-284 PMID: 24735413

32. Anders S, Pyl PT, Huber W (2015) HTSeq-a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166–169. doi: 10.1093/bioinformatics/btu638 PMID: 25260700

33. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15: 550. PMID: 25516281

34. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102: 15545–15550. PMID: 16199517

35. Xiao Y, Hsiao TH, Suresh U, Chen HI, Wu X, Wolf SE, et al. (2014) A novel significance score for gene selection and ranking. Bioinformatics 30: 801–807. doi: 10.1093/bioinformatics/btr671 PMID: 22321699

36. McLean CY, Bristor D, Hiller M, Clarke SL, Schaar BT, Lowe CB, et al. (2010) GREAT improves functional interpretation of cis-regulatory regions. Nat Biotechnol 28: 495–501. doi: 10.1038/nbt.1630 PMID: 20436461

37. Bair E, Tibshirani R (2004) Semi-supervised methods to predict patient survival from gene expression data. PLoS Biol 2: E108. PMID: 15094809

38. Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, et al. (2012) Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature 481: 389–393. doi: 10.1038/nature10730 PMID: 22217937

39. Choudhry H, Schodel J, Oikonomopoulos S, Camps C, Grampp S, Harris AL, et al. (2014) Extensive regulation of the non-coding transcriptome by hypoxia: role of HIF in releasing paused RNApol2. EMBO Rep 15: 70–76. doi: 10.1002/embr.201337642 PMID: 24363272

40. Mathelier A, Zhao X, Zhang AW, Parcy F, Worsley-Hunt R, Arenillas DJ, et al. (2014) JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res 42: D142–147. doi: 10.1093/nar/gkt997 PMID: 24194598

41. Semenza GL (2003) Targeting HIF-1 for cancer therapy. Nat Rev Cancer 3: 721–732. PMID: 13130303

42. Project TIN-HsLPF (1993) A predictive model for aggressive non-Hodgkin's lymphoma. The International Non-Hodgkin's Lymphoma Prognostic Factors Project. N Engl J Med 329: 987–994. PMID: 8141877

43. Laderoute KR (2005) The interaction between HIF-1 and AP-1 transcription factors in response to low oxygen. Semin Cell Dev Biol 16: 502–513. PMID: 16144688

44. Gordan JD, Bertout JA, Hu CJ, Diehl JA, Simon MC (2007) HIF-2alpha promotes hypoxic cell proliferation by enhancing c-myc transcriptional activity. Cancer Cell 11: 335–347. PMID: 17418410

45. Gordan JD, Lal P, Dondeti VR, Letrero R, Parekh KN, Oquendo CE, et al. (2008) HIF-alpha effects on c-Myc distinguish two subtypes of sporadic VHL-deficient clear cell renal carcinoma. Cancer Cell 14: 435–446. doi: 10.1016/j.ccr.2008.10.016 PMID: 19061835

46. Kanno T, Kamba T, Yamasaki T, Shibasaki N, Saito R, Terada N, et al. (2012) JunB promotes cell invasion and angiogenesis in VHL-defective renal cell carcinoma. Oncogene 31: 3098–3110. doi: 10.1038/onc.2011.475 PMID: 22020339

47. Simon JM, Hacker KE, Singh D, Brannon AR, Parker JS, Weiser M, et al. (2014) Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. Genome Res 24: 241–250. doi: 10.1101/gr.158253.113 PMID: 24158655

48. Schodel J, Mole DR, Ratcliffe PJ (2013) Pan-genomic binding of hypoxia-inducible transcription factors. Biol Chem 394: 507–517. doi: 10.1515/hsz-2012-0351 PMID: 23324384

49. Mimura I, Nangaku M, Kanki Y, Tsutsumi S, Inoue T, Kohro T, et al. (2012) Dynamic change of chromatin conformation in response to hypoxia enhances the expression of GLUT3 (SLC2A3) by cooperative interaction of hypoxia-inducible factor 1 and KDM3A. Mol Cell Biol 32: 3018–3032. doi: 10.1128/MCB.06643-11 PMID: 22645302