Published online 2 June 2009

Nucleic Acids Research, 2009, Vol. 37, No. 14 ***4587–4602***
doi:10.1093/nar/gkp425

An integrative genomics approach identifies Hypoxia Inducible Factor-1 (HIF-1)-target genes that form the core response to hypoxia

Yair Benita[1], Hirotoshi Kikuchi[2], Andrew D. Smith[3], Michael Q. Zhang[3], Daniel C. Chung[2] and Ramnik J. Xavier[1,2,*]

[1]Center for Computational and Integrative Biology, [2]Gastrointestinal Unit, Center for the Study of Inflammatory Bowel Disease, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114 and [3]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

Received April 20, 2009; Revised May 6, 2009; Accepted May 8, 2009

ABSTRACT

The transcription factor Hypoxia-inducible factor 1 (HIF-1) plays a central role in the transcriptional response to oxygen flux. To gain insight into the molecular pathways regulated by HIF-1, it is essential to identify the downstream-target genes. We report here a strategy to identify HIF-1-target genes based on an integrative genomic approach combining computational strategies and experimental validation. To identify HIF-1-target genes microarrays data sets were used to rank genes based on their differential response to hypoxia. The proximal promoters of these genes were then analyzed for the presence of conserved HIF-1--binding sites. Genes were scored and ranked based on their response to hypoxia and their HIF-binding site score. Using this strategy we recovered 41% of the previously confirmed HIF-1-target genes that responded to hypoxia in the microarrays and provide a catalogue of predicted HIF-1 targets. We present experimental validation for *ANKRD37* as a novel HIF-1-target gene. Together these analyses demonstrate the potential to recover novel HIF-1-target genes and the discovery of mammalian-regulatory elements operative in the context of microarray data sets.

INTRODUCTION

Maintenance of oxygen homeostasis in mammalian cells is fundamental to the survival of the organism. One of the pivotal mediators of the cellular response to hypoxia is hypoxia-inducible factor (HIF), a transcription factor that contains a basic helix-loop-helix motif as well as PAS domain. There are three known members of the HIF family (HIF-1, HIF-2 and HIF-3) and all are α/β heterodimeric proteins. HIF-1 was the first factor to be cloned and is the best understood isoform (1). HIF-3 is a distant relative of HIF-1 and little is currently known about its function and involvement in oxygen homeostasis. HIF-2, however, is closely related to HIF-1 and both activate hypoxia-dependent gene transcription. In all three isoforms the α-subunit senses oxygen levels (2–5). In normoxia the α-subunit is ubiquitously expressed but is rapidly degraded through interaction with the *VHL* ubiquitin ligase complex (6). In hypoxic conditions, it is stabilized through multiple mechanisms that are transcriptional, post-transcriptional, as well as post-translational (6–9). The β-subunit is widely expressed and generally not regulated by hypoxia. Once stabilized HIF-1 and HIF-2 interact with co-activators such as *CREBBP*/*EP300* and bind to a consensus sequence element (5′-RCGTG-3′) in the proximal promoters of hypoxia-regulated genes. Together these genes control cellular process including a switch from oxidative to glycolytic metabolism, inhibition of cellular proliferation, and stimulation of oxygen delivery through erythropoiesis and angiogenesis (10). Metabolic deregulation is the final outcome if the adaptive response is inadequate.

Hypoxia can arise in a variety of developmental, physiological and pathological states. Solid tumors, such as breast cancer, quickly outgrow their blood supply and manipulate the hypoxia pathways to promote their own survival. The role of HIF-1 in tumor onset, progression, invasion and metastasis has been demonstrated in

*To whom correspondence should be addressed. Tel: 617 643 3331; Fax: 617 643 3328; Email: xavier@molbio.mgh.harvard.edu
Present address:
Andrew D. Smith, Department of Biology, University of Southern California, Los Angeles, CA 90089, USA

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

several solid tumors (11,12). In addition, cancer specific-regulatory roles for HIF-2 have been demonstrated in an expanding list of cancers (13). For instance, in clear renal cell carcinoma the normal predominance of *HIF1A* expression is altered in favor of *EPAS1* (HIF2α) expression and promotes prosurvival rather than proapoptotic factors (14,15). As a result HIF-target genes have emerged as potential targets for cancer specific therapy (16,17).

HIF-1 is an important metabolic regulator that allows rapid adaptation to oxygen availability. This effect is mediated through key regulators of bioenergetics and growth. For instance, *PDK1* is directly induced by HIF-1 resulting in attenuation of mitochondrial function and oxygen consumption (18) and *CDKN1A* is induced leading to growth arrest at the G1 phase (19). Identifying such HIF-target genes is central to our understanding of the HIF network and the mechanism of cellular pathways modulation during hypoxia. In addition, identifying additional binding sites in HIF-regulated genes might contribute to the paradigm of inferring function from sequence.

Currently, studies attempting to identify novel HIF-target genes rely on microarray expression profiles to detect genes that respond to changes in HIF protein levels. Typically a gene of interest is selected and experimentally validated for HIF binding to its promoter. Kim et al. identified 25 transcripts that increase by at least 4-fold during hypoxia in a B cell line. *PDK1* was selected for further validation due to its role in regulating the TCA cycle. Consequently they were able to show that HIF-1 binds to the *PDK1* promoter (18). Shen *et al.* (20) used a more rationale microarray design where tissues from *Caenorhabditis elegans* were profiled in normoxia and in hypoxia in a wild-type, HIF-1 knockout and a VHL mutant strain. This study resulted in 63 HIF-1 dependent genes that respond to hypoxia but only 12 of those had a human homolog. Elvidge *et al.* (21) compared expression profiles of human breast cancer cells in normoxia, hypoxia and in normoxia treated with a 2-OG-dependent dioxygenase inhibitor, dimethyloxalylglycine (DMOG) which activates HIF-1. In addition, they compared expression profiles in hypoxia and in hypoxia treated with siRNAs to *HIF1A* and *EPAS1* (HIF-2α). Manalo *et al.* (22) profiled arterial endothelial cells in normoxia and in hypoxia as well as in normoxia in the presence of a HIF-1α constitutively active form. Only 45 genes were identified in both studies as HIF-1 dependent. Although these approaches are more likely to filter out genes that are not dependent on HIF-1, it remains challenging to determine whether the affected transcripts were directly or indirectly modulated by HIF.

Several methods for *in-silico* identification of transcription factor-binding sites (TFBS) have been proposed (23–29). These methods fall into two categories. First, identification of TFBS based on predefined position weight matrices (PWM). Second, *de novo* prediction of sequence motifs over-represented in proximal promoters compared to control sequences [for a review, see (30,31)]. Examples for the former include STORM in CREAD (27), CORE_TF (28) and GeneAct (29) and for the latter DME (32) and the genomic identification of motifs described by Xie *et al.* (26). PWMs of HIF-1 have been previously used to identify genome-wide target genes. For instance, the Gene Set Enrichment Analysis program contains a set of HIF-1-target genes as part of its Molecular Signature Database (26,33). However, these approaches do not take into consideration promoter availability. In addition, the consensus HIF-binding sequence is very short, has low information content and is ubiquitously present in the human genome making *in-silico* genome-wide prediction unreliable. To generate a reliable prediction we propose to reduce the search space to a subset of genes expressed in specific tissues under specific conditions. This can be achieved by microarray expression profiles under the assumption that promoters of genes that respond to a stimulus are available for TF binding. Several groups previously demonstrated that microarray expression profiles can be used to improve TFBS prediction (32,34–36) and that *cis*-regulatory modules in proximal promoters correlate with differential expression of downstream-target genes (27).

In this study we present an integrative genomic approach combining both computational and experimental strategies to identify HIF-target genes. Using this approach we recovered 41% of the known HIF-1-target genes that were differentially expressed in hypoxia, defined a ranked list of HIF-target genes and experimentally validated *ANKRD37* as a novel HIF-1 target.

MATERIALS AND METHODS

Microarrays

Raw microarray data was obtained from GEO (37) and ArrayExpress (38). Microarray studies were filtered to identify those that profiled the same tissue in normoxia and hypoxia on Affymetrix platform and used at least two replicates and provided raw data. Affymetrix-based data sets were normalized using the GCRMA module (39) of bioconductor (40) and present–absent calls were calculated for each probe using the MAS5 module. All probes with no present calls were removed and from the remaining at least one sample was required to have an expression value larger than $\log_2$ (100), as recommended by Affymetrix. In all experiments, only pairwise comparisons were made with respect to normoxia. After the initial filtering was applied, the expression values were centered to the mean for each probe with replicates. Only probes for which the expression values of the replicates were all positive or all negative were further tested for differential expression. The Limma module (41) of biocoductor was used to fit a linear model for each probe and *P*-values were corrected using the false discovery rate method (42).

Promoters

Promoter sequences of human genome build 18 were obtained from UCSC genome browser (43). For each protein-coding gene, the proximal promoter was defined as 700 bases upstream and 300 bases downstream of the transcription start site. In cases where multiple transcripts were annotated all alternative promoters were obtained and analyzed individually. Due to restriction of

conservation analysis programs, for all genes the forward strand sequence of the genome was used. Background promoters for enrichment analysis were selected randomly from the promoter collection to match the CpG content and base composition of the foreground promoters. We obtained CpG sequence location from UCSC genome browser (43) and matched their location to the promoters defined. Each promoter was classified as CpG or non-CpG and for the background set the same proportion of CpG and non CpG promoters were selected randomly according to the ratio in the foreground set. Next we verified that the base composition matched within ±2% for each nucleotide with the foreground data set. For all enrichment purposes a set of 1000 promoters was used as a background.

Enrichment of binding sites

The program MOTIFCLASS v1.40 (27) was employed to identify binding sites over-represented in promoters of a foreground set (genes differentially expressed in microarrays) compared to a background set of 1000 randomly selected promoters. The relative error-rate was optimized for each motif and a *P*-value was calculated using 10 000 permutations. Finally, only motifs with an enrichment specificity of at least 0.5 were reported.

De-novo prediction of motifs

The program DME (v2.0) (27,32) was employed using the ZOOPS model to identify motifs of length 5–12 bases compared to a background set. Only the top 50 enriched motifs were reported and compared to known motifs using the program MATCOMPARE in CREAD (27).

Binding site prediction

The program STORM in CREAD (27,44) was used to match PWMs to genomic DNA and was employed using a *P*-value cutoff of 1*e*-4. This cutoff was found to be optimal since a cutoff of $P < 1e{-}3$ was not stringent enough predicting nearly every promoter in the genome as having a HIF-1 site while $P < 1e{-}5$ was too stringent predicting only a handful of promoters in the genome with a HIF-1 site. Conservation information of 28 species genome alignments (43) was used in two approaches. First, the software multiSTORM [also in CREAD (27)] was employed to identify the number of genomes in which the PWM could still be matched with the same *P*-value cutoff as used in STORM. Second, the software site-cons was employed to calculate the level of conservation at each binding site compared to the flanking 100 bases (27). The site-cons *P*-value was obtained by extracting 100 bases upstream and downstream of the binding site and sampling randomly a number of columns equal to the length of the predicted binding site. A conservation score was then calculated for these columns and the process was iterated 10 000 times to obtain a distribution. Finally, the *P*-value was calculated by comparing the original score for the binding site to the distribution. Conservation information was obtained from the 28-genome alignments available at UCSC genome browser (43). Namely, human (hg18), chimpanzee (panTro2), rhesus (rheMac2), bushbaby (otoGar1), tree shrew (tupBel1), mouse (mm8), rat (rn4), guinea pig (cavPor2), rabbit (oryCun1), shrew (sorAra1), hedgehog (eriEur1), dog (canFam2), cat (felCat3), horse (equCab1), cow (bosTau3), armadillo (dasNov1), elephant (loxAfr1), tenrec (echTel1), opossum (monDom4), platypus (ornAna1), lizard (anoCar1), chicken (galGal3), frog (xenTro2), fugu (fr2), tetraodon (tetNig1), stickleback (gasAcu1), medaka (oryLat1) and zebrafish (danRer4).

Transcription network

For each predicted HIF-target gene, the single best scoring HIF site was selected and other binding sites were identified within ±25, ±50, ±75, ±100, ±125 and ±150 bases of that site as well as in the entire promoter. To obtain a *P*-value the frequency of each site within that window was compared to a randomly selected background set of 1000 promoters from genes that respond to hypoxia but had no detectable HIF-binding site. Sequence composition for the foreground and background set were matched as described above. In each of the background promoters a single binding site was randomly selected and binding sites were identified within the window size as selected for the HIF site. A chi-square test was performed for each TF, comparing the occurrence of binding sites in the neighborhood of HIF to their occurrence in the neighborhood of the randomly selected binding site in the background set. This process was iterated 100 times and the average *P*-value for each TFBS was used as the final *P*-value. No correction for *P*-values was performed.

Functional pathway analysis

The KEGG database (release 47) (45) was obtained and used to place predicted HIF-target genes into molecular pathways. Statistical analysis was performed using a hypergeometric distribution as described in the GOhyperGAll module of Bioconductor for gene ontology terms (40). The R program was employed (v2.6.2) with the following command: phyper(x − 1, m, n–m, k and lower.tail = FALSE), where x is the number of predicted target gene within a specific pathway; m, the number of genes in that pathway; n total number of unique genes in KEGG; and k, the number of predicted HIF-target genes that have an entry in any of the KEGG pathways. *P*-values were not corrected for multiple testing. This statistical analysis was repeated using the Reactome pathway database (46), the curated gene sets of Gene Set Enrichment Molecular Signature database (33) and the microRNA predicted targets database TargetScan (47). Enrichment for protein interactions was performed by first compiling a list of interactors for each human protein (Data obtained from NCBI Gene database). The HomoloGene database (48) was used to identify interaction in other species that can be mapped to human proteins. Finally, enrichment was calculated as described for KEGG by considering all interactors of each protein. Protein domains were obtained from Interpro (49) and enrichment was calculated for each domain.

Selection of mitochondrial genes

Genes that localize to the mitochondria were selected by identifying (i) all human Swiss-Prot proteins annotated as localized to the mitochondria (50); (ii) genes annotated with Gene Ontology terms that include the keywords: mitochondria, mitochondrion or mitochondrial (51); and (iii) A predicted catalog of genes localized to the mitochondria (52). Genes from the predicted list of mitochondrial genes that conflicted with existing knowledge and annotation were removed.

Cell culture

DLD-1, HCT116, SW480, Lovo, Panc-1, HeLa and MCF7 cells were maintained in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS). Hypoxic conditions were achieved by culturing cell lines in a sealed hypoxia chamber (Billups-Rothenberg) for 18 h after flushing with a mixture of 1% O_2, 5% CO_2 and 94% N_2 (Compressed pre-mixed gas, Airgas) (53). To minimize the effect of serum growth factors, the culture medium was switched to serum-free UltraCulture (Lonza) before the cells were subjected to hypoxia.

Quantitative polymerase chain reaction analysis (qPCR)

Total RNA was isolated from cultured cells using the RNeasy kit (Qiagen) and reverse transcription with random hexanucleotide primers and SuperScript Reverse Transcriptase III (Invitrogen) was performed. The resulting cDNA was amplified by real-time polymerase chain reaction using the iQ5 Real-Time PCR Detection System (BioRad) and Power SYBR Green PCR Master Mix (Applied Biosystems). Primer sequences were 5′-GTCGC CTGTCCACTTAGCC-3′ and 5′-GCTGTTTGCCCGTT CTTATTACA-3′ for *ANKRD37*, 5′-AGGCCAGCACAT AGGAGAGA-3′ and 5′-TTTCTTGCGCTTTCGTTT TT-3′ for *VEGFA* and 5′-CGGCTACCACATCCAAGG AA-3′ and 5′-GCTGGAATTACCGCGGCT-3′ for 18S rRNA. Transcript levels of *ANKRD37* were normalized to 18S rRNA.

Western blotting

Cells were lysed in chilled lysis buffer (20 mM Tris–HCl, pH 7.5, 150 mM NaCl, 1 mM Na_2EDTA, 1 mM EGTA, 1% Triton, 2.5 mM sodium pyrophosphate, 1 mM β-glycerophosphate, 1 mM Na_3VO_4, 1 mM leupeptin) supplemented with the Complete protease inhibitor cocktail (Roche). Total 50 μg of protein extracts were resolved on a 3–8% NuPAGE Tris–acetate SDS polyacrylamide gel (Invitrogen) and transferred onto a polyvinylidene difluoride membrane (Millipore). The blots were probed with anti-*HIF1A* (Transduction Laboratories; 1:500) and anti-β actin (Sigma; 1:100 000) antibodies. Immunoreactive proteins were visualized using the Western Lighting Chemiluminescence Reagent Plus (PerkinElmer Life Sciences).

Plasmid constructions

Fragments from −325 bp to +230 bp (promoter 1) −166 bp to +230 bp (promoter 2) or −78 bp to +230 bp (promoter 3) relative to the transcriptional start site of the human *ANKRD37* promoter were amplified by polymerase chain reaction (PCR) using 1 μg of human genomic DNA extracted from DLD-1 cells. The PCR primers were as follows: −325 bp sense primer; 5′-AACCAAAC GCGTCCAGTTTCCTGGTTACGTG-3′, −166 bp sense primer; 5′-AACCAAACGCGTTCTCACGCCCACTGA CTTA-3′, −78 bp sense primer; 5′-AACCAAACGCGTA TCCACAGCTGGCCAATCG-3′ and +230 bp antisense primer; 5′-GTGGTTGAGATCTTAAGGCAGATGAC AAGAGTCC-3′. The PCR-amplified products were digested with MluI and BglII and subcloned into the pGL3-basic vector (Promega). Site directed mutations were introduced at three potential HIF-binding sites (1, 2 and 4). To prepare site 1 and site 4 mutants, the −325 *ANKRD37*-luc pGL3 plasmid was digested with MluI/PstI or PmlI/Tth111I, respectively, and ligated with a double-stranded oligonucleotide in which the consensus HIF-binding site was disrupted. The oligonucleotide sequences were as follows: 5′-CGCGTCCAGTTTCC TGGTTAAAAGCGCGCCGGCAGAGCCAAAACC TGCA-3′ and 5′-GGTTTTGGCTCTGCCGGCGCGCT TTTAACCAGGAAACTGGA-3′ for site 1 at −310 bp and 5′-GTGCCAGTGTTTGTGTAAAAGCGGCCGT GGCGGGGCTGGACA-3′ and 5′-CTGTCCAGCCCC GCCACGGCCGCTTTTACACAAACACTGGCAC-3′ for site 4 at +14 bp. To prepare the site 2 mutant, PCR-based site-directed mutagenesis was performed using primers 5′-TACCGGGAGAAAAGTCAAACTCA-3′ and 5′-TGAGTTTGACTTTTCTCCCGGTA-3′. The nucleotide sequences of all constructs were confirmed by sequencing.

Transfections and reporter assays

Transient transfections were performed using Fugene 6 (Roche) in 24-well tissue culture plates according to the manufacturer's specifications. Total 150 ng of *ANKRD37*-luciferase pGL3 reporter constructs were co-transfected with 2 ng of pRL-CMV (Promega) as a control. Ten hours after transfection, the culture medium was switched to UltraCulture and cells were grown in normoxic or hypoxic conditions for 18 h. Luciferase activity was measured with a dual luciferase reporter assay system (Promega). The experiments were performed in triplicate wells a minimum of three times.

RNA interference

HCT116 cells were transfected with 20 nM siRNA duplex oligos directed to *HIF1A* or a non-specific control siRNA using Lipofectamine RNAiMAX (Invitrogen) according to the manufacturer's instructions. Nucleotide sequences of siRNAs were as follows: *HIF1A*, 5′-r(AUCCAGAGUC ACUGGAACU)d(TT)-3′, and control, 5′-r(GCGCGCU UUGUAGGAUUCG)d(TT)-3′. The silencing effect was confirmed by Western blotting 48 h after transfection.

ANKRD37 promoter reporter constructs were transfected 12 h after siRNA transfection.

Electrophoretic mobility shift assay (EMSA)

Nuclear extracts were prepared from MCF7 cells cultured in either normoxic or hypoxic conditions for 12 h utilizing NE-PER nuclear extraction reagent (Pierce). The *ANKRD37* promoter sequence utilized as an oligonucleotide probe was 5′-CTACCGGGAGACGTGTCAAACTC AG-3′. The 5′-end of the oligonucleotide was labeled with biotin, and a complementary oligonucleotide was annealed to generate double-stranded fragments. EMSA was performed using LightShift chemiluminescent kit (Pierce) as previously described (54). Specificity of shifts was confirmed by utilizing 200-fold molar excess of unbiotinylated oligonucleotides as a specific competitor. Mutagenesis was performed to further define the elements responsible for the specific shifts obtained. The sequence of the mutant oligonucleotide probe that includes point mutations in the HIF consensus-binding site was: 5′-CT ACCGGGAGAAAAGTCAAACTCAG-3′.

Chromatin immunoprecipitation (ChIP)

MCF7 cells were cultured in either normoxic or hypoxic conditions for 6 h and then fixed with 1% formaldehyde at 37°C for 10 min. The ChIP assays were performed using a Chromatin Immunoprecipitation Assay Kit (Upstate) according to the manufacturer's protocol. Briefly, cells were lysed and sonicated to achieve DNA sizes ranging from 200 to 1000 bp. Diluted supernatants were precleared by incubating with salmon sperm DNA/protein-A agarose, 50% slurry at 4°C for 30 min. Supernatants were then incubated with anti-HIF-1α antibody (Novus Biologicals number 100–105) or normal mouse IgG (Santa Cruz) at 4°C overnight, or directly processed to the next step without antibody. Immunocomplexes were collected with salmon sperm DNA/protein A-agarose, 50% slurry, washed, eluted and cross-linkage was reversed by heating at 65°C for 4 h. The eluates were then digested with proteinase K at 45°C for 1 h. The precipitated DNAs were recovered using PCR Purification Kit (Qiagen) and amplified by PCR using PrimeSTAR HS DNA Polymerase (Takara Bio). The PCR primers used for amplification of site 2 (primers A) and a non-HIF region upstream of site 2 (primers B) were as follows: sense primer A; 5′-CCAGTTTCCTGGTTACGTGC-3′, antisense primer A; 5′-TAAGTCAGTGGGCGTGAGAG-3′, sense primer B; 5′-TGGAGGCACTTTCACCACAC-3′ and anti-sense primer B; 5′-TGTGAGGTGAGCACAC GTTG-3′.

RESULTS

HIF-1 and HIF-2 and their downstream-target genes are activated in response to hypoxia. To identify HIF-target genes we first identified genes that responded to hypoxia using publicly available microarray data sets. NCBI GEO and EBI ArrayExpress were searched for experiments where the same tissue was profiled comparing normoxia and hypoxia. Experiments that used replicates, were profiled on Affymetrix platform and provided raw data were accepted for further analysis (Supplementary Table S1). The same identical analysis protocol was applied to all experiments (see 'Materials and Methods' section). Although all microarrays were profiled on Affymetrix arrays, the number of genes differentially expressed in each cell type varied significantly and ranged from 486 genes in monocytes to 2119 genes in HeLa, with a surprisingly small overlap. The differentially expressed genes in each cell type are the potential HIF-target genes and were subject to *in-silico* analysis of HIF-binding sites.

In-silico optimization of HIF-binding site prediction

Two databases of PWMs were obtained and searched for HIF-1 PWM. The Jaspar collection (v3.0) contained no HIF-1 PWM while the TRANSFAC collection (v8.2) contained 2 HIF-1 PWMs (M00466, M00797). These PWMs are 12 and 14 bases long and were compiled based on only 12 and 23 validated binding sites, respectively. We compiled an additional HIF-1 PWM based on 104 validated binding sites in human, mouse and rat collected by Wenger *et al.* (10) (see Supplementary Figure S1; submitted to Jaspar). Despite the apparent variability in PWM length, it is clear from the alignment in Supplementary Figure S1 that most of the information content is within the HIF consensus binding site: RCGTG. The composition preference in the region flanking the core RCGTG observed in the Transfac matrices is a consequence of using a small number of validated HIF-1-binding sites.

Proximal promoters were defined from 700 bases upstream to 300 bases downstream of the transcription start site (TSS) of each transcript. Initially, HIF-1 PWMs were used to identify binding sites; however, we observed a too small overlap between the three HIF-1 PWMs for any given *P*-value and conservation filtering criteria. Differences in PWM length (12–18 bp) and in the sequences flanking the consensus binding site were found to introduce a bias in HIF-1-binding-site prediction since calculation of *P*-values and conservation consider the entire length of the matrix and all bases defined within. To avoid this bias all HIF consensus-binding sites sequences (RCGTG) were extracted from each promoter and tested for conservation in other species (see 'Materials and Methods' section). A shift from A to G and vice versa in the first position was allowed for the purpose of conservation. Next, the software site-cons in CREAD was employed to calculate the level of conservation at each binding site compared to the flanking 100 bases. A site was predicted as a HIF-binding site if it was conserved in at least 10 species (including human) or if conserved in four species with a site-cons $P < 0.05$. When the HIF consensus site is preceded by a C, it may form the E-box-binding site CACGTG which is a known binding site for several bHLH transcription factors including HIF-1. In 43% of the E-box sites other TFs were predicted to bind at the same location as HIF-1, compared to only 2% of the non-Ebox sites. The transcription factors that are most commonly predicted at the same

location as HIF are *ARNT* (HIF-1β; as homodimer), *AHR*:*ARNT*, *MYC*, *MAX*, *MYCN*, *CLOCK*:*ARNTL* and *USF1* (Supplementary Figure S2). We only accepted genes with at least one non-Ebox HIF site as HIF-target genes. Throughout this strategy our filtering criteria were very stringent in order to obtain a catalogue of high confidence at the expense of not recovering some known HIF targets. Each gene was scored as follows:

$$\text{HIF site score} = \text{Number of species} \times(-\log(\text{conservation } P\text{−value}))$$

$$\text{Hypoxia score} = \text{abs}\left(\sum\left(\frac{\text{Hypoxia}}{\text{Normoxia}}\right)\right)$$

The hypoxia/normoxia ratio was inverted and assigned a negative value for genes that were downregulated in hypoxia. Each gene was assigned two ranks, one based on the highest scoring HIF site and one on the hypoxia score. Finally, the sum of both ranks was calculated and all genes were ranked based on that value. For instance, the gene *BNIP3*, a bona fide HIF-1 target, was ranked 515 by HIF site score and 15 by hypoxia score. Using the sum of ranks BNIP3 was ranked 56 among all genes. This scoring scheme is biased towards genes that respond to hypoxia in multiple cells. While some genes could be HIF targets in only one cell, we reason that a gene that responds to hypoxia in more than one cell is more likely to be a HIF target. This strategy is summarized in Figure 1. Despite the stringent selection criteria of this approach we were able to identify a HIF-binding site in 35/86 promoters of known target genes. A list of previously identified known HIF-1-target genes is shown in Supplementary Table S2.

The majority of hypoxia responsive genes lack a HIF-binding site

HIF-target genes were predicted for each data set separately with the criteria outlined in 'Materials and Methods' section. Of the known and validated HIF-target genes that responded to hypoxia in each data set, we recovered 37–55% (Table 1). On average 70% of the hypoxia responsive genes did not contain a detectable HIF-binding site in their proximal promoter (Table 1). To verify this observation and to test that no bias was introduced by our stringent conservation filtering above, genes that responded to hypoxia in each data set were divided to two groups. Those with a predicted HIF-binding site (A) and those with no detectable HIF-binding site (B). Each group was tested for over-representation of any PWM compared to a random set of promoters with similar base composition and CpG content. This analysis was performed using the program MOTIFCLASS in CREAD and is independent from the prediction analysis and selection criteria outlined above. In addition, MOTIFCLASS does not consider conservation and is based on sequence analysis alone. As expected HIF PWMs were identified in the top 20 over-represented binding sites in group A and none of the HIF PWMs were detected in the top 20 PWMs from group B. The results for the top 20 enriched motifs in each group in MCF7 breast cancer cells are shown in Figure 2A and B. The highest scoring PWM was a HIF-1 matrix that had a classification error of 0.407. In each of the six data sets we observed HIF-1 enrichment in group A while group B varied significantly between cell types (Supplementary Table S3). We did not observe any significant HIF enrichment within the top 20 PWMs in group B throughout all data sets.

The general response to hypoxia is dependent on a small number of genes

Previous studies have suggested that the response to hypoxia varies by cell type (55–57). However, we expected a core set of genes to be shared, e.g. genes that control the shift from oxidative to glycolytic metabolism and inhibition of cellular proliferation. Yet, the cellular response to hypoxia was surprisingly selective between cells. Only 17 genes responded to hypoxia in all studied cell types. These genes include HIF targets such as *ENO1*, *BHLHB2* and *BNIP3*. We observed a positive correlation between the number of cells in which a gene responded to hypoxia and the identification of HIF-binding sites in its promoter (Figure 2C). Only 23% of the genes that responded to hypoxia in a single cell type had a predicted HIF-binding site in their promoter compared to 59% of the genes that responded in six cell types. The top scoring gene was *TMEM45A* which responded to hypoxia in all studied cell types with an average of 21.6-fold induction and a HIF-binding site conserved in 13 species with a conservation *P*-value <0.0001 compared to the flanking 100 bp. At least 5 of the top 10 genes have been previously identified as HIF-1-target genes, namely *DDIT4* (58,59), *STC2* (60), *P4HA1* (61), *JMJD1A* (62) and *ALDOC* (63). We did not find any correlation between the response to hypoxia and the HIF-binding site score for the top 100 predicted HIF-target genes. To obtain a cutoff of high confidence, the top 500 ranked genes were plotted versus the cumulative number of previously validated HIF-1 targets (Supplementary Figure S3). The resulting curve showed a logarithmic increase in validated targets for genes ranking below 200 and approached saturation between 200 and 500. Therefore, the 200th gene rank was used as a cutoff of high confidence. Supplementary Table S4 summarizes our prediction results for the top 500 HIF-target genes.

In the top 200 predicted HIF-target genes, 81 responded to hypoxia in at least three cell types (Figure 3). We hypothesize that these genes are likely to be part of the core response to hypoxia. In addition, these genes have a higher probability of being HIF-1 targets since their response to hypoxia was detected in multiple experiments. Of the 81 genes, 22 genes were previously reported to be HIF-target genes and 60 genes (74%) were upregulated in hypoxia. In addition, 80 (99%) of the genes contained a CpG island in their proximal promoter, significantly higher than the 51% CpG containing promoters in our entire promoter database. CpG islands are thought to be non-tissue specific (64), further supporting the selection of the 81 genes as a shared response to hypoxia. The Kyoto Encyclopedia of Genes and Genomes (KEGG) database

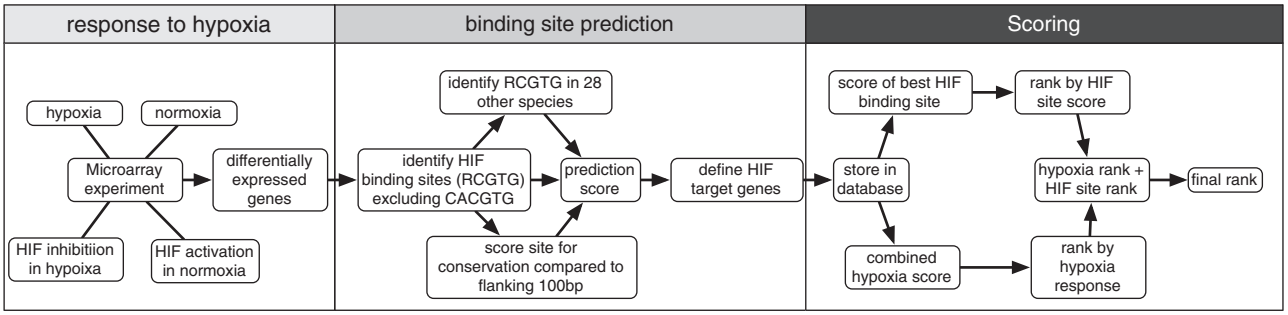

Figure 1. Prediction strategy for identifying HIF-1-target genes. Candidate genes that respond to hypoxia were first identified by microarrays. Each data set was subjected to a computational analysis in which HIF-binding sites were detected in proximal promoters. Each gene was scored for its response to hypoxia and for the best HIF-binding site. No cutoff was set for determining HIF-target genes. Finally, all genes were ranked.

Table 1. Transcriptional response to hypoxia across six different cell types as determined by microarrays

	MCF7	U251	Astrocytes	Monocytes	B cells	HeLA
Genes that respond to hypoxia	830	1702	1371	486	1920	2119
Genes that respond to hypoxia in which a HIF binding site was identified	278 (33%)	546 (32%)	380 (28%)	159 (33%)	534 (28%)	555 (26%)
Known target genes that respond to hypoxia	35	31	20	22	32	27
Known target genes recovery	19 (54%)	17 (55%)	8 (40%)	11 (50%)	13 (41%)	10 (37%)

was used to identify over-represented pathways and processes within the 81 genes. In Figure 4A a metabolic map of HIF-1-target genes was constructed from the resulting analysis (Supplementary Table S5). The metabolic map contained several known HIF-1-target genes, such as *LDHA*, *ALDOC* and *GAPDH* and also several novel targets such as *GYS1* in starch and sucrose metabolism (65) and fumarate hydratase (*FH*) in the TCA cycle (66). One of the key cellular responses to HIF-1 in hypoxia is a metabolic shift from oxidative phosphorylation to glycolytic production of ATP. To highlight direct targets that may be involved in HIF-1 regulation of mitochondria function the predicted HIF targets that localize to the mitochondria (see 'Materials and Methods' section) were used to create a mitochondrial map (Figure 4B). Eight of eleven genes were downregulated in hypoxia and are involved in critical steps of mitochondrial function including two ribosomal subunits, cytochrome *c*, the iron transporter *SLC25A28* and the iron–sulfur assembly and repair protein *ISCU*.

Classification of novel HIF targets

All previously validated HIF-1 targets (Supplementary Tables S2 and S4) were tested for functional enrichment in KEGG pathway databases, Reactome database, Gene Set Enrichment Molecular Database (GSEA), protein interactions, InterPro protein domains and microRNA-target prediction database TargetScan (see 'Materials and Methods' section). This analysis allows identification of pathways or attributes that are highly represented by the HIF targets compared to an equivalent random list of genes. We next mapped novel HIF targets to these significantly enriched attributes as demonstrated by selected examples in Table 2 (full analysis shown in Supplementary Table S6). Analysis of protein domains identified 2-oxoglutarate and Fe(II)-dependent oxygenase family as highly enriched in validated HIF targets (5 of 19 proteins, $P < 2e{-}7$). This family includes the prolyl hydroxylases known to regulate HIF-1α as part of the oxygen sensing mechanism. PLOD1, is a novel predicted HIF target that is part of the same family. Analysis of microRNA targets identified five microRNAs that target validated HIF targets. miR-1/206, for instance, is predicted to target 583 human genes, of which nine are known to be HIF-1 targets and 14 are predicted HIF targets. Analysis of gene sets from GSEA highlighted a group of 104 genes that predict a poor breast cancer prognosis and included nine validated HIF targets ($P = 1e{-}7$) and four predicted ones. A graphical representation of protein interactions and Reactome pathways is shown in Figure 4C and D and highlights genes such as PCAF. PCAF interacts with four validated HIF targets and six predicted targets.

ANKRD37 is a novel HIF-target gene

To demonstrate the utility of our method in identifying novel HIF-1-target genes, we selected five genes that were previously unknown to be HIF-1 targets, *ANKRD37* (ranked 8), *KLF10* (ranked 24) *WSB1* (ranked 32), *GYS1* (ranked 34) and *ELL2* (ranked 465). Using quantitative PCR, *ANKRD37* was induced up to 35-fold in hypoxia compared to normoxia, *KLF10* up to 2.5-fold, *WSB1* up to 4.9-fold, *GYS1* up to 7.8-fold, *ELL2* was not induced in hypoxia and *VEGFA* that was used as a positive control was induced up to 8.3-fold (Figure 5 and Supplementary Figure S4A). The gene *ANKRD37* was

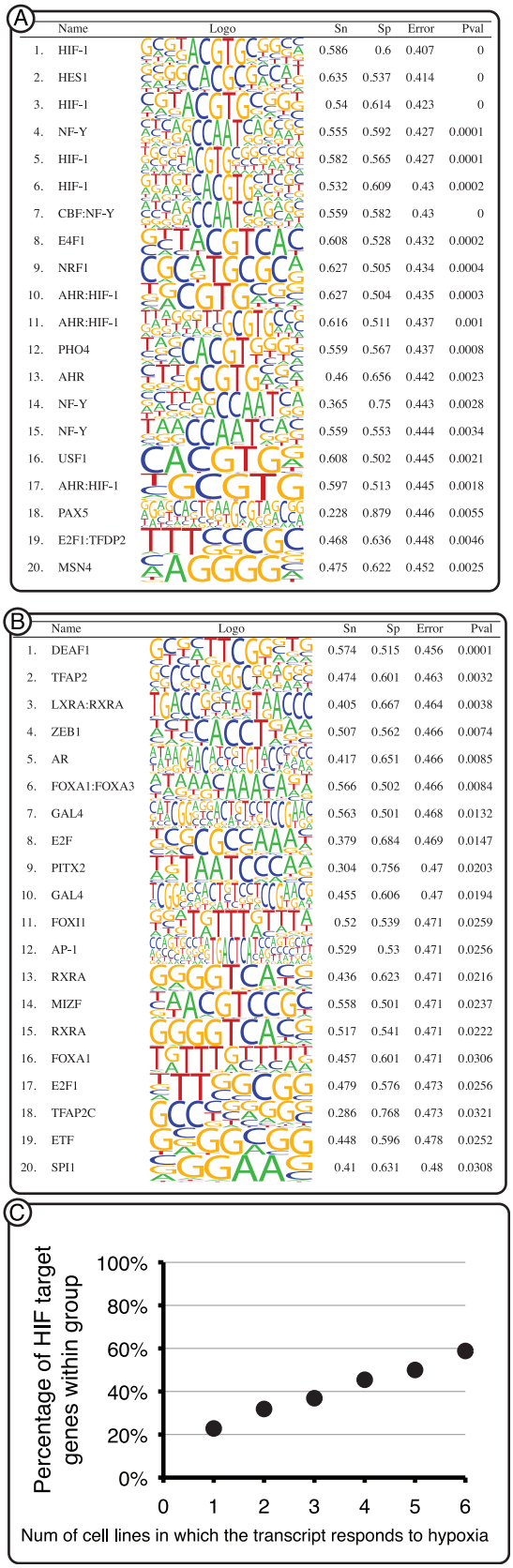

Figure 2. (**A** and **B**) Binding site enrichment analysis. The program MOTIFCLASS was employed to identify binding sites significantly over-represented in the promoters of genes that respond to hypoxia in B cells. (A) Totally 770 promoters of genes that responded to hypoxia in which a detectable HIF-1-binding site was identified. HIF-1 PWMs were identified as the most enriched. (B) Totally 1456 promoters of genes that responded to hypoxia but do not contain a detectable HIF-1-binding site. No HIF-1 PWM was identified in the top 20 matrices. Sn, sensitivity; Sp, specificity; Error, classification error rate; Pval, *P*-value. (**C**) Correlation between the number of HIF-target genes and the number of cells in which they responded to hypoxia. Only 31% of the genes that responded to hypoxia in one of six cells was predicted as a HIF-target gene, compared to 71% of the genes that responded to hypoxia in six of six cell lines.

studied further due to its extreme induction in hypoxia and its high ranking within the predicted targets. *ANKRD37* is a short (158 amino acids, 17 kDa) ankyrin repeat protein with unknown function that is conserved in mammals and zebrafish. *ANKRD37* responded to hypoxia in four of six cell types and four HIF-1-binding sites were predicted within the proximal promoter. The response to hypoxia was validated under hypoxic conditions across seven cell lines. We observed a strong induction of *ANKRD37* in all cell lines but most significantly in the colon cancer cell lines DLD-1, HCT116 and SW480 (Figure 5A). *ANKRD37* hypoxic induction was measured after 4, 8 and 12 h of incubation in hypoxic conditions in HCT116 cells and compared to *VEGFA*. HIF-1 protein levels were determined by western blot and showed a significant induction after 4 and 8 h compared to normoxia and a slight decrease after 12 h (Figure 5B). ANKRD37 and *VEGFA* mRNA levels showed a similar profile that matched the HIF-1 induction; however, ANKRD37 was induced 3-fold over the level of *VEGFA* induction (Figure 5C).

We next investigated the HIF-binding sites in the *ANKRD37* promoter. We identified four HIF-1-binding sites that matched our filtering criteria and were located at −311 to −306 bp (site 1), −228 to −223 bp (site 2), −142 to −137 bp (site 3) and +12 to +17 bp (site 4). As shown in Figure 6A, three *ANKRD37* promoters encompassing these different HIF-binding sites were cloned upstream of a luciferase reporter. Promoter 1 was induced 4.1, 7.4 and 10.7-fold in hypoxia compared to normoxia in DLD-1, HCT116 and MCF7 cells, respectively. In contrast the induction observed with promoters 2 and 3 was in the range of 1.5–3-fold (Figure 6B). To determine whether the hypoxia induction was HIF-1 dependent, luciferase activity of *ANKRD37* promoter 1 was measured in HCT116 cells in the presence of *HIF1A* siRNA. Endogenous *HIF1A* knockdown was confirmed by western blot (Supplementary Figure S4A). The 5-fold induction of promoter 1 in hypoxia was reduced to 1.3-fold in the presence of *HIF1A* siRNA (Supplementary Figure S4B). These results suggested that sites 1 and 2 are the physiologically relevant HIF-1-binding sites. Therefore, sites 1, 2 and 4 in promoter 1 were selectively mutated at the HIF consensus-binding site (ACGTG > AAAAG). As shown in Figure 6C, the hypoxic induction of promoter 1 was maximally inhibited by a mutation in site 2 while the other mutations had a minimal effect.

To establish whether endogenous HIF-1 protein binds to site 2 in the *ANKRD37* promoter, EMSA assays were performed using nuclear extracts from MCF7 cells

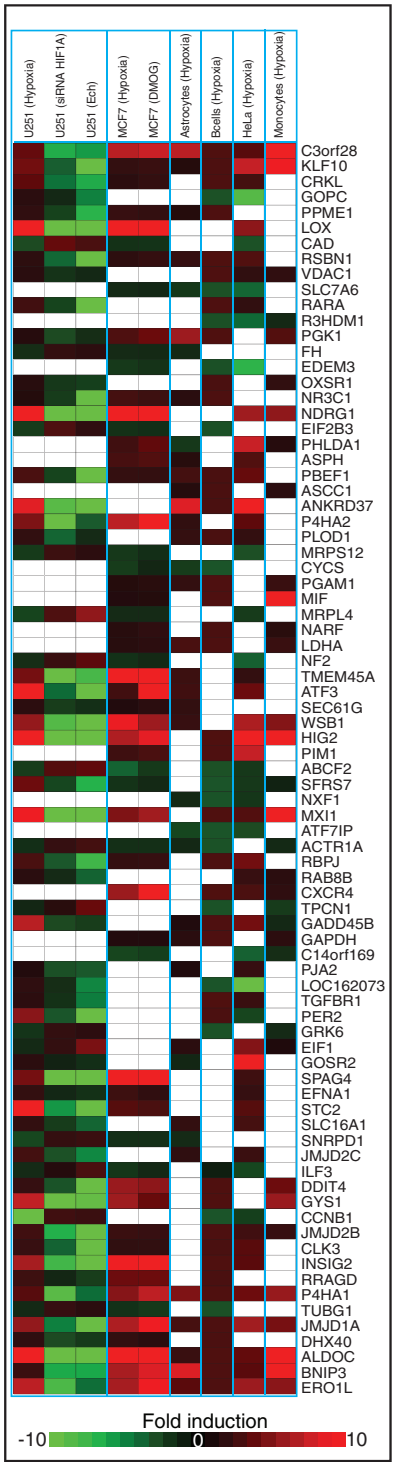

Figure 3. HIF dependent response to hypoxia across six cell types for 81 selected genes that respond to hypoxia in at least three cell types and were ranked within the top 200 HIF-target genes. The heatmap shows a fold induction compared to normoxia within each cell type. White blocks indicate the gene did not respond to hypoxia in that cell type.

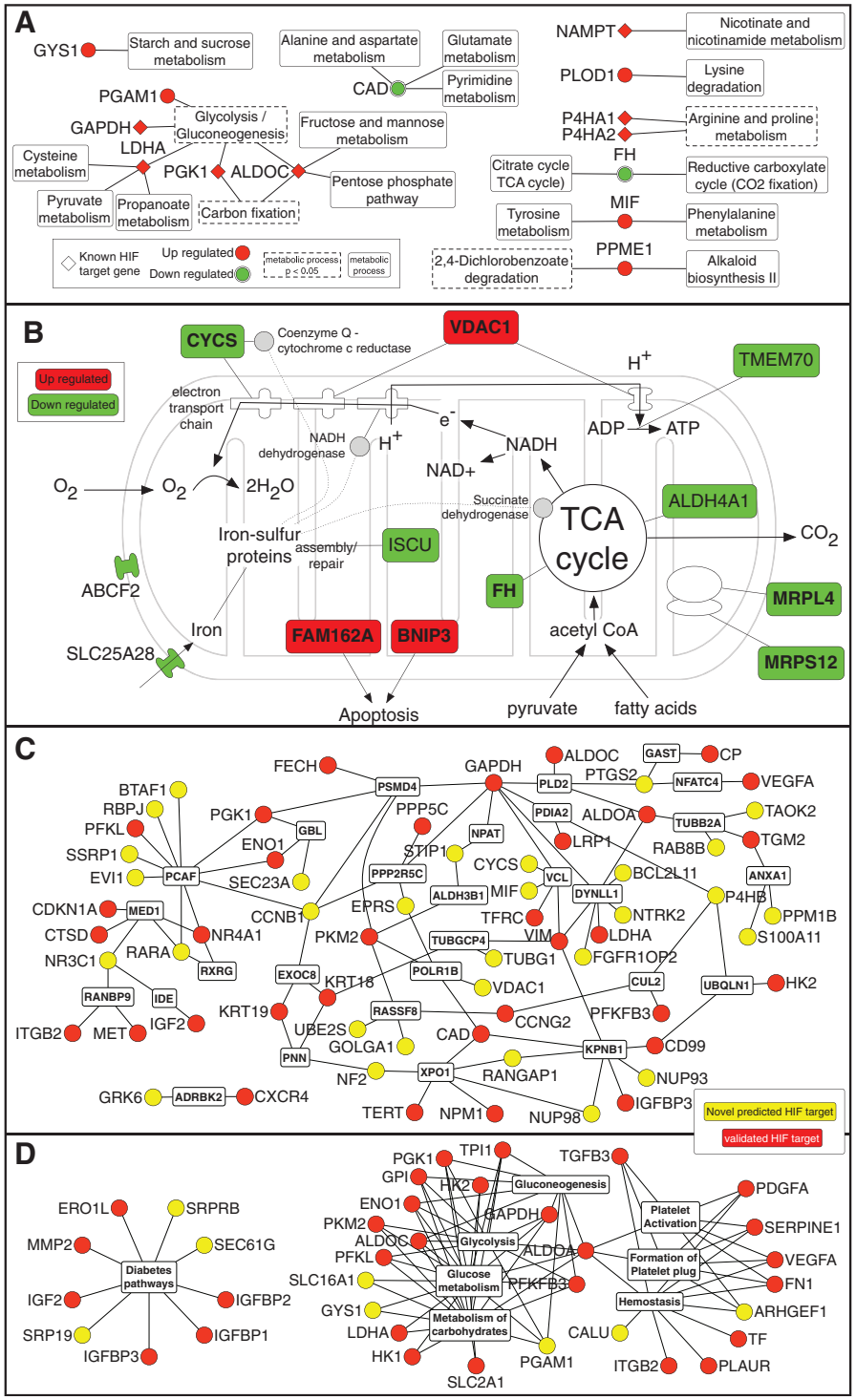

Figure 4. (**A**) Metabolic map of HIF-target genes. The KEGG pathway database was employed to map 81 HIF-target genes defined as the core response to hypoxia to pathways. Metabolism related processes are shown here (black objects) and those over-represented with a $P < 0.05$ are shown with a dotted white frame. Upregulated genes are shown as red nodes while downregulated genes as green nodes. Previously identified HIF-1-target genes are indicated by a diamond shape. (**B**) Mitochondrial map. Genes within the top 200 predicted HIF targets that localize to the mitochondria were mapped according to their functional annotation. Upregulated genes are shown as red nodes and downregulated as green nodes. Genes indicated in bold responded to hypoxia in at least three cell types. Dotted lines from iron–sulfur protein assembly and repair illustrate a few of the iron–sulfur proteins that are key to mitochondrial function, such as NADH dehydrogenase. (**C**) Graphical representation of protein-protein interaction network and Reactome pathways (**D**) enrichment for previously validated HIF-1 targets (red) and novel predicted targets (yellow). Only proteins/pathways that were significantly enriched ($P < 0.05$) are shown. See Supplementary Table S6 for analysis details.

Table 2. Selected examples of functional enrichment of 101 previously validated HIF-1 targets and mapping of novel targets to these categories

Source	Description	Number of genes	Previously validated HIF-1 targets		Novel HIF-1 targets
			P-value	Gene symbols	
Protein domains	IPR005123; 2-oxoglutarate (2OG) and Fe(II)-dependent oxogenase	6/19	2.E-07	P4HA2, EGLN3, EGLN1, P4HA1, PH-4	PLOD1
GSEA	Poor progonosis marker genes in Breast Cancer	13/104	1.E-07	CA9, CP, PGK1, EGLN1, TFRC, VEGFA, NDRG1, ADM, BNIP3	TMEFF1, IVNS1ABP, TMEM45A, RRAGD
KEGG	Arginine and proline metabolism	6/35	4.E-05	P4HA2, NOS3, P4HA1, NOS2A	EPRS, ALDH4A1
Reactome	Platelet activation	7/86	1.E-02	PDGFA, TGFB3, FN1, VEGFA, ALDOA, SERPINE1	ARHGEF1
TargetScan	MiR-1/206	23/583	3.E-03	PDGFA, NAMPT, HSP90B1, MET, ETS1, VEGFA, EDN1, CITED2, STC2	RNF165, GRK6, SLC7A6, PTPLAD1, ASPH, MYLK, RSBN1, SLC31A1, EVI1, PGAM1, SOX6, BTAF1, MXD1, GLCCI1
Protein interactions	PCAF	10/106	9.E-05	ENO1, NR4A1, PFKL, PGK1	RARA, CCNB1, SSRP1, RBPJ, EVI1, BTAF1

The number of genes reflects the total number of known and predicted targets within the functional group. *P*-values for enrichment of validated HIF-1 targets were obtained using the hypergeometric distribution as described in 'Materials and Methods' section.

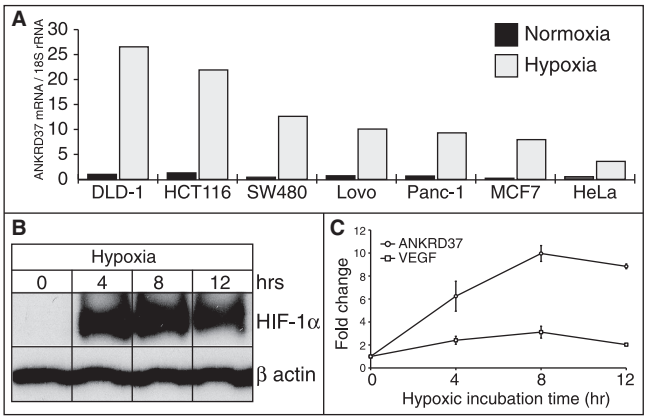

Figure 5. *ANKRD37* response to hypoxia. (**A**) *ANKRD37* mRNA levels were monitored by qPCR across seven cell lines and normalized to 18S rRNA. (**B**, **C**) *ANKRD37* hypoxic induction over time during hypoxia. (B) Endogenous *HIF1A* protein levels determined after 0, 4, 8 and 12 h of hypoxic incubation by western blot. (C) mRNA levels of *ANKRD37* and *VEGFA* measured by qPCR after 0, 4, 8 and 12 h of hypoxic incubation.

incubated in either normoxia or hypoxia. A unique band shift was obtained only when nuclear extracts from cells grown in hypoxic conditions were incubated with the probe corresponding to site 2 (Figure 6D, lane 3). When non-biotinylated probe was co-incubated as a competitor, or a probe with specific mutations in the HIF consensus-binding site (ACGTG > AAAAG) was utilized, the band obtained in hypoxic conditions was lost (Figure 6D, lanes 4 and 5). Nuclear extracts from cells grown in hypoxia were then incubated with an antibody specific to HIF-1α or control mouse IgG prior to performing EMSA assays. Incubation with an antibody against HIF-1α resulted in a supershift of this specific band (Figure 6E). To further address whether HIF-1α physiologically binds to site 2, chromatin immunoprecipitation assays were performed using MCF7 cells incubated in either normoxia or hypoxia. Cross-linked cell lysates were immunoprecipitated with or without anti- HIF-1α antibody or control mouse IgG and PCR was performed to amplify a 178 bp or 220 bp fragment of the *ANKRD37* promoter using primers corresponding to site 2 (primers A) and a region upstream of site 2 (primers B), respectively. DNA bound to HIF-1α was detected when DNA–protein complexes from hypoxic cells were precipitated with an anti-HIF-1α antibody and amplified with primers A, whereas no binding was observed at an upstream promoter region with primers B (Figure 6F). These data indicate that site 2, the most conserved and highest scoring HIF site, is indeed a HIF-1-binding site in DLD-1, HCT116 and MCF7 cells.

The HIF-transcriptional network

To identify transcription factors that are part of the HIF network we analyzed the presence of other binding sites in the 81 genes that ranked in top 200 and responded to hypoxia in at least three cell types. Matrices for validated transcription factors from the TRANSFAC and Jaspar collections were used to identify binding sites within 150 bases away from the highest scoring HIF-binding site in each of the promoters as well as within the entire promoter. For this analysis we used STORM with $P < 1e{-}4$ and accepted only binding sites predicted in at least four species and significantly ($P < 0.05$) more conserved compared to the flanking 100 bases (see 'Materials and Methods' section). To obtain a *P*-value for co-occurring binding sites, a chi-square analysis was performed compared to TFBS occurrence in a background set of 1000 promoters of genes that responded to hypoxia in at least one cell type but had no detectable HIF binding (see

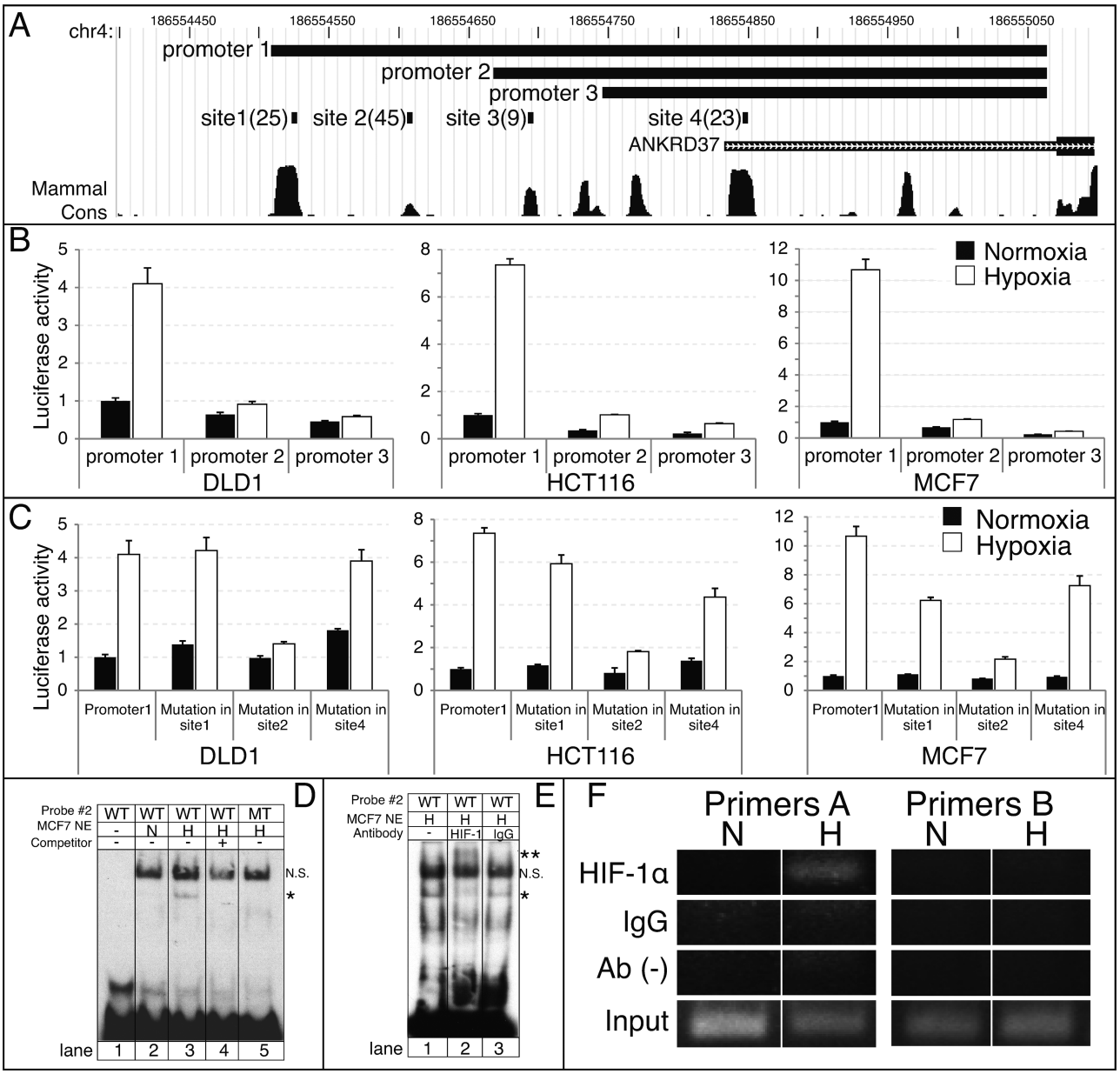

Figure 6. Mapping of HIF-1 site in *ANKRD37* promoter. (**A**) Genome browser view of the four HIF-binding sites identified in the *ANKRD37* promoters and the three promoter sequences that were cloned to identify the dominant HIF-1 site. The site score is shown in parenthesis for each site, higher is better. The larger box of the *ANKRD37* transcript (green) indicates the coding sequence. Note that the conservation shown on the plot is computed by PhastCons which scores conservation based on the evolutionary distance and does not necessarily reflect the number of species (95). (**B**) DLD1, HCT116 and MCF7 cells were transiently transfected and incubated under normoxic or hypoxic conditions. Luciferase activity of the three promoters was measured and normalized to promoter 1 in normoxic conditions. All studies were performed in triplicate and data represent the mean ± SD for three different wells. (**C**) Luciferase activity of promoter 1 in normoxia versus hypoxia with mutations in sites 1, 2 and 4 in DLD1, HCT116 and MCF7 cells. Experimental procedure and data analysis was performed as described for panel B. (**D**) *In vitro* binding of HIF-1 protein to site 2 in the *ANKRD37* promoter. EMSAs were performed using biotinylated probe corresponding to HIF-binding site 2. Nuclear extracts (NE) from MCF7 cells cultured in either normoxia (N) or hypoxia (H) were utilized. The asterisk indicates the specific shift obtained in hypoxia. 200× M excess of non-labeled probe as a specific competitor or probe with a specific mutation in the HIF-binding site2 (ACGTG > AAAAG) was utilized to confirm specificity. WT = wild-type probe. MT = mutant probe. (**E**) Nuclear extracts from MCF7 cells grown in hypoxia were incubated with antibody specific to HIF-1α or normal mouse IgG. The double asterisk indicates the super-shift obtained only with a HIF-1α antibody. N.S.; nonspecific band. (**F**) *In-vivo* binding of HIF-1α protein to site 2 in the *ANKRD37* promoter. ChIP assays were performed using primer sets corresponding to HIF-binding site 2 (primer A) or a region upstream of site 2 (primer B). DNA–protein complexes from MCF7 cells cultured in either normoxia (N) or hypoxia (H) were immunoprecipitated with or without antibody specific to HIF-1α or normal mouse IgG. Precipitated DNA or input samples were amplified by PCR.

'Materials and Methods' section). The analysis was performed in increments of 25 bases from the highest scoring HIF site and the most significant *P*-value for each PWM was selected. Only binding sites that were not predicted at the HIF consensus binding site were considered. We identified 10 TFBS that were significantly enriched within a specific distance from the highest scoring HIF site including *ATF2*, *CREB1*, *SP1*, *WT1* and *JUN* which have been previously linked to hypoxia and HIF-1 (Supplementary Table S6). When considering the entire promoter, we observed enrichment for P53:P73-binding sites and multiple members of the forkhead family, including FOXO4, FOXD1 and FOXF2. To test for potential mechanism in up or down regulation of HIF-target genes, a similar analysis was employed to upregulated and downregulated transcripts separately. Most of the sites co-occurring for all HIF targets above were also enriched for the upregulated genes, albeit with more significant *P*-values. SP-1 and NF-Y were the only TFs significantly co-occurring within a specific distance from the HIF site in downregulated genes. Surprisingly, four of five predicted NF-Y targets (*ABCF2*, *CYCS*, *FH*, *MRPL4*) were genes localized to mitochondria (Supplementary Table S7).

HIF-1 motifs are identified *de-novo* in predicted HIF-1-target genes

To identify *de-novo* motifs over-represented in the promoters of the 81 HIF-target genes selected above, DME2 was employed. For each motif length in the range of 5–12 bases we identified the top 50 over-represented motifs compared to a suitable background set (see 'Materials and Methods' section). These motifs were compared to the TRANSFAC and Jaspar collections of PWMs using the CREAD program MATCOMPARE. Comparison was estimated using a divergence score in the range of 0 to 1, where a divergence of 0 indicates the motif predicted *de-novo* is identical to a PWM. HIF-1 PWMs were matched best to six bases long *de-novo* motifs. For motifs longer than 7 bases, no HIF-1 PWMs were ranked within the top 50 (Supplementary Table S8). Comparing all available PWMs to *de-novo* predicted motifs with a divergence of 0.25 or lower identified *ARNT* (HIF-1β), *ATF2*, NF-Y and *ATF6* as the most enriched in the data set (Supplementary Table S9). Other motifs that were predicted *de-novo* include *CREB1*, AP-1 and *EVI1* and *MIZF* (Supplementary Table S9). Some of these are also identified above as part of the HIF-transcriptional network.

DISCUSSION

In this study we presented a computational strategy to identify HIF-1-target genes. This strategy integrates gene expression profiles and computational-binding site analysis of proximal promoters. The utility of our strategy in detecting novel HIF-target genes was demonstrated by recovering known HIF-1-target genes and experimentally validating *ANKRD37*. *ANKRD37* is a previously uncharacterized protein that contains four ankyrin repeats. Ankyrin repeats are thought to mediate protein–protein interactions (67) and are known to be involved in regulation of key transcription factors such as NFκB and *TP53* (68–70). *ANKRD37* induction was 3-fold higher than *VEGFA* (Figure 5E), the validated HIF-1-binding site was conserved in 10 species and was significantly induced in hypoxia in multiple cell types analyzed here. Taken together, these data suggest *ANKRD37* is likely to have an important role in the hypoxia response.

The cellular response to hypoxia was different between cell types, consistent with previous reports (55,56). The union of all differentially expressed genes in the data sets presented here contained over 6000 unique genes. Even when comparing studies that attempted to detect a HIF dependent response to hypoxia (21,22,71), only 45 genes were differentially expressed in all three. Importantly the analysis presented here indicates that 60% of the transcripts that responded to hypoxia did not have a detectable conserved HIF-binding site in their proximal promoter. This observation was also supported by MOTIFCLASS enrichment, which is an independent analysis that does not consider conservation and did not detect any enrichment for HIF-binding sites in those promoters. In addition, transcripts with no detectable HIF site were enriched for other TFBS that were previously shown to participate in the cellular response to hypoxia including, *DEAF1*, *ELK1*, *POU2F1* and *IRF1* (72,73). Moreover, genes that responded to hypoxia and did not have a detectable HIF-binding site were used as a background for the HIF-1-transcriptional network. One possible mechanism that may explain the different hypoxic response between cells is that genes directly regulated by HIF are involved in key processes and their regulation mediates a robust effect. As illustrated in the mitochondrial map, where several predicted downregulated genes affect key steps in the mitochondria. The predicted HIF-target genes *SLC25A28*, an iron transporter in the mitochondria and *ISCU*, an iron–sulfur scaffold protein are both downregulated in hypoxia. These two genes could affect several of the mitochondria iron–sulfur proteins including, NADH dehydrogenease, Coenzyme Q—cytochrome *c* reductase and Succinate—coenzyme Q reductase are critical to the function of the electron transport chain and the TCA cycle. In addition, the predicted HIF targets *MRPL4* and *MRPS12* are components of the mitochondrial ribosome that are required for mitochondrial protein translation. Additional examples include cell growth arrest by inhibition of *CAD*, a component of pyrimidine synthesis (74); *RRP9*, a component of snoRNP predicted to participate in processing and modification of pre-ribosomal RNA (75); *EIF2B3*, a translation initiation factor subunit; and *CCNB1*, a regulatory protein involved in mitosis (74). One novel predicted HIF target that is highlighted by the HIF-1 KEGG metabolic pathways map is *GYS1*, a gene involved in glycogen biosynthesis (65,76) that was validated as responsive to hypoxia by quantitative PCR in multiple cell types.

To highlight predicted HIF targets of interest and to suggest potential functional roles for these targets, a bioinformatics analysis of validated HIF targets was performed and predicted targets that were not previously

validated were mapped to the significantly enriched functional categories. This analysis is summarized in Supplementary Table S6 and identifies several attractive predicted HIF targets. The family of 2-oxoglutarate and Fe(II)-dependent oxygenase has 19 members in the human genome of which five are bona fide HIF targets, including the HIF-1α regulators *EGLN1* and *EGLN3*. The predicted HIF target *PLOD1*, is a member of that family and a lysyl hydroxylase involved in collagen synthesis and localized to the endoplasmic reticulum (77). Mutations in *PLOD1* have been associated with type VI Ehlers–Danlos syndrome, a genetic disorder caused by defects in collagen synthesis (78). In another example, microRNA-target predictions from TargetScan were analyzed and validated HIF targets were enriched for five microRNAs including miR-1/206 and miR-130/301. miR-1/206 were shown to be upregulated in a rat model of myocardial infarction (79) and miR-130 was shown to regulate angiogenesis (80). Both are relevant to the HIF pathway and would be attractive candidate microRNAs to study further with respect to the validated and predicted targets. Finally, protein interaction data presented in Figure 4C can be used to identify potential regulators, co-factors or adapter proteins that play a role in the HIF pathway. Finally, P300/CBP-associated factor (PCAF) interacts with four validated HIF targets and six predicted targets. PCAF was previously shown to be a HIF-1α co-factor that regulates p53-transcriptional activity in hypoxia (81). The specific HIF targets that interact with PCAF are attractive candidates for functional studies.

The method presented here is not without limitation. First, the microarray itself presents a limitation since only genes represented on the microarray are considered. However, this could be improved by using multiple platforms that offer complete coverage of annotated genes. Of the 86 previously validated HIF-1 targets (Supplementary Table S2), only 56 responded to hypoxia in at least one cell type. One example of a bona fide HIF-target gene that was not recovered by our analysis is *EPO* (erythropoietin). A close examination of the data for *EPO* revealed that the probe did not respond to hypoxia in any of the cells that were profiled. *EPO* was previously reported to be highly induced in the kidney (82). Second, HIF-binding sites may be located outside the promoter region we defined or may not be conserved in other species. The binding site for the well-characterized HIF-1 target *VEGFA* is located at −985 bp (83) and therefore was not recovered in the present analysis. HIF-1-binding sites have been identified as far as 12-kb downstream (*PHD3*, (84) and 5-kb upstream [eNos; (85)] of the TSS. The defined promoter region was found optimal to capture most of the known sites (60% of known HIF-1-binding sites) while minimizing potential false positive prediction.

In this study we provide a ranked catalogue of genes predicted to be direct HIF-target genes. Given the different hypoxia response between cell types, it is currently difficult to estimate how many HIF-target genes exist. Our analysis across six cell types was stringent in an effort to obtain the highest specificity and the top 200 ranked genes are of high confidence. The ranking system by itself offers a level of confidence for the prediction and was shown as reliable by high ranking for previously validated targets (Supplementary Figure S3). Currently, almost all known HIF-target genes are upregulated in hypoxia. While the majority of predicted HIF targets were upregulated in hypoxia, nearly all downregulated genes are novel. The top ranking downregulated gene is *EDEM3*, an α1,2-mannosidase involved in quality control of misfolded proteins (86). In addition, a subset of downregulated genes highlighted in the mitochondrial map is consistent with the observed attenuation of oxidative phosphorylation by HIF-1 (87). We also identified NF-Y as significantly co-occurring adjacent to the highest scoring HIF-1 site in four of these downregulated mitochondrial genes. It is currently unclear whether NF-Y is part of the hypoxic regulation of those genes. However, it was previously shown in yeast that the NF-Y homolog regulates the mitochondrial enzyme COX (88), which was also shown to be regulated by HIF-1 to optimize efficiency of mitochondrial respiration (89).

To describe the transcriptional network of HIF, we identified additional TFBS over-represented within ±150 bp from the highest scoring HIF site as well as TFBS enriched in the entire promoter. We identified only a few TFs within a specific distance from the HIF site, however, almost all were previously associated with HIF-1, including *CREB1*, *ATF2* and *SP1* (90–92). The recovery of these TFs increases the confidence in the genes predicted as HIF targets and also identifies specific gene targets in Supplementary Table S6 as predicted targets with higher confidence. *WSB1*, for instance, is a novel predicted HIF-1 target that contains binding sites for *CREB1*, *ATF6*, *ATF2* and *JUN:ATF2*, a profile very similar to LDHA that is a bona fide HIF-1 target. *WSB1* was previously shown to be involved in the ubiquitination and degradation of *HIPK2* (93), a kinase that was shown to repress HIF-1α transcription (94). *WSB1* was also confirmed to be hypoxia responsive by quantitative PCR (Supplementary Figure S4). Taken together these data suggest *WSB1* is highly likely to be a direct HIF-1 target and to be part of a feedback loop that includes HIF-1α, *WSB1* and *HIPK2*. TFs that were previously associated with HIF-1 appear to be enriched only in genes upregulated during hypoxia. YY-1 and NF-Y were detected as enriched in downregulated genes. Further studies will be necessary to validate these observations and to determine their functional relevance to HIF-1.

De-novo motif prediction suggested that no significant information flanking the HIF consensus-binding site could be found in the predicted HIF-target genes. This is consistent with the observation that enhancing the number of validated HIF-binding sites reduces the information content of the flanking nucleotides (Supplementary Figure S1). Conservation was not considered in the *de-novo* prediction process. Yet, several TFs identified as part of the HIF-transcriptional network have also been identified *de-novo*, including, *CREB1*, *JUN:ATF2*, *ATF2*, *ATF6* and AP-1. The ability to recover these TFs *de-novo* confirms their significant over-representation and demonstrates the utility of this approach in recovering previously validated TFs and co-factor that are part of the HIF network. Furthermore, novel TFs

that may be part of the HIF-transcriptional network are currently supported by sequence analysis data alone and additional studies are required to validate these findings.

HIF-1 and HIF-2 are both key transcription factors with that modulate metabolic pathways, bioenergetics and processes relevant to cancer onset and progression. This study not only gives insights into the genes potentially regulated by HIF but also serves as a guide to study direct targets of other transcription factors.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

FUNDING

Harris Fellowship to Y.B.; National Institutes of Health [grant number DK043351, AI062773, DK83756] for work in RJX lab; National Institutes of Health [grant number HG001696] for work in MZQ lab; and National Institutes of Health [grant number CA92594] for work in DCC lab. Funding for open access charge: [grant number HG001696, CA45508] for MZQ lab.

Conflict of interest statement. None declared.

REFERENCES

1. Wang,G.L. and Semenza,G.L. (1995) Purification and characterization of hypoxia-inducible factor 1. *J. Biol. Chem.*, **270**, 1230–1237.
2. Cockman,M.E., Masson,N., Mole,D.R., Jaakkola,P., Chang,G.W., Clifford,S.C., Maher,E.R., Pugh,C.W., Ratcliffe,P.J. and Maxwell,P.H. (2000) Hypoxia inducible factor-alpha binding and ubiquitylation by the von Hippel-Lindau tumor suppressor protein. *J. Biol. Chem.*, **275**, 25733–25741.
3. Ohh,M., Yauch,R.L., Lonergan,K.M., Whaley,J.M., Stemmer-Rachamimov,A.O., Louis,D.N., Gavin,B.J., Kley,N., Kaelin,W.G. and Iliopoulos,O. (1998) The von Hippel-Lindau tumor suppressor protein is required for proper assembly of an extracellular fibronectin matrix. *Mol. Cell*, **1**, 959–968.
4. Kamura,T., Sato,S., Iwai,K., Czyzyk-Krzeska,M., Conaway,R.C. and Conaway,J.W. (2000) Activation of HIF1alpha ubiquitination by a reconstituted von Hippel-Lindau (VHL) tumor suppressor complex. *Proc. Natl Acad. Sci. USA*, **97**, 10430–10435.
5. Tanimoto,K., Makino,Y., Pereira,T. and Poellinger,L. (2000) Mechanism of regulation of the hypoxia-inducible factor-1 alpha by the von Hippel-Lindau tumor suppressor protein. *EMBO J.*, **19**, 4298–4309.
6. 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.
7. Gustafsson,M.V., Zheng,X., Pereira,T., Gradin,K., Jin,S., Lundkvist,J., Ruas,J.L., Poellinger,L., Lendahl,U. and Bondesson,M. (2005) Hypoxia requires notch signaling to maintain the undifferentiated cell state. *Dev. Cell*, **9**, 617–628.
8. Jaakkola,P., Mole,D.R., Tian,Y.M., Wilson,M.I., Gielbert,J., Gaskell,S.J., Kriegsheim,A., Hebestreit,H.F., Mukherji,M., Schofield,C.J. *et al.* (2001) Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science*, **292**, 468–472.
9. Ivan,M., Kondo,K., Yang,H., Kim,W., Valiando,J., Ohh,M., Salic,A., Asara,J.M., Lane,W.S. and Kaelin,W.G. Jr. (2001) HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. *Science*, **292**, 464–468.
10. Wenger,R.H., Stiehl,D.P. and Camenisch,G. (2005) Integration of oxygen signaling at the consensus HRE. *Sci. STKE*, **2005**, re12.
11. Krishnamachary,B. and Semenza,G.L. (2007) Analysis of hypoxia-inducible factor 1alpha expression and its effects on invasion and metastasis. *Methods Enzymol.*, **435**, 347–354.
12. Liao,D., Corle,C., Seagroves,T.N. and Johnson,R.S. (2007) Hypoxia-inducible factor-1alpha is a key regulator of metastasis in a transgenic model of cancer initiation and progression. *Cancer Res.*, **67**, 563–572.
13. Rankin,E.B. and Giaccia,A.J. (2008) The role of hypoxia-inducible factors in tumorigenesis. *Cell Death Differ.*, **15**, 678–685.
14. Raval,R.R., Lau,K.W., Tran,M.G., Sowter,H.M., Mandriota,S.J., Li,J.L., Pugh,C.W., Maxwell,P.H., Harris,A.L. and Ratcliffe,P.J. (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.
15. Kondo,K., Klco,J., Nakamura,E., Lechpammer,M. and Kaelin,W.G. Jr. (2002) Inhibition of HIF is necessary for tumor suppression by the von Hippel-Lindau protein. *Cancer Cell*, **1**, 237–246.
16. Semenza,G.L. (2003) Targeting HIF-1 for cancer therapy. *Nat. Rev. Cancer*, **3**, 721–732.
17. Patiar,S. and Harris,A.L. (2006) Role of hypoxia-inducible factor-1alpha as a cancer therapy target. *Endocr. Relat. Cancer*, **13(Suppl. 1)**, S61–S75.
18. Kim,J.W., Tchernyshyov,I., Semenza,G.L. and Dang,C.V. (2006) HIF-1-mediated expression of pyruvate dehydrogenase kinase: a metabolic switch required for cellular adaptation to hypoxia. *Cell Metab.*, **3**, 177–185.
19. Cho,Y.S., Bae,J.M., Chun,Y.S., Chung,J.H., Jeon,Y.K., Kim,I.S., Kim,M.S. and Park,J.W. (2008) HIF-1alpha controls keratinocyte proliferation by up-regulating p21(WAF1/Cip1). *Biochim. Biophys. Acta*, **1783**, 323–333.
20. Shen,C., Nettleton,D., Jiang,M., Kim,S.K. and Powell-Coffman,J.A. (2005) Roles of the HIF-1 hypoxia-inducible factor during hypoxia response in Caenorhabditis elegans. *J. Biol. Chem.*, **280**, 20580–20588.
21. Elvidge,G.P., Glenny,L., Appelhoff,R.J., Ratcliffe,P.J., Ragoussis,J. and Gleadle,J.M. (2006) Concordant regulation of gene expression by hypoxia and 2-oxoglutarate-dependent dioxygenase inhibition: the role of HIF-1alpha, HIF-2alpha, and other pathways. *J. Biol. Chem.*, **281**, 15215–15226.
22. Manalo,D.J., Rowan,A., Lavoie,T., Natarajan,L., Kelly,B.D., Ye,S.Q., Garcia,J.G. and Semenza,G.L. (2005) Transcriptional regulation of vascular endothelial cell responses to hypoxia by HIF-1. *Blood*, **105**, 659–669.
23. Zhou,Q. and Wong,W.H. (2004) CisModule: de novo discovery of cis-regulatory modules by hierarchical mixture modeling. *Proc. Natl Acad. Sci. USA*, **101**, 12114–12119.
24. Gupta,M. and Liu,J.S. (2005) De novo cis-regulatory module elicitation for eukaryotic genomes. *Proc. Natl Acad. Sci. USA*, **102**, 7079–7084.
25. Zhu,Z., Shendure,J. and Church,G.M. (2005) Discovering functional transcription-factor combinations in the human cell cycle. *Genome Res.*, **15**, 848–855.
26. Xie,X., Lu,J., Kulbokas,E.J., Golub,T.R., Mootha,V., Lindblad-Toh,K., Lander,E.S. and Kellis,M. (2005) Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. *Nature*, **434**, 338–345.
27. Smith,A.D., Sumazin,P., Xuan,Z. and Zhang,M.Q. (2006) DNA motifs in human and mouse proximal promoters predict tissue-specific expression. *Proc. Natl Acad. Sci. USA*, **103**, 6275–6280.
28. Hestand,M.S., van Galen,M., Villerius,M.P., van Ommen,G.J., den Dunnen,J.T. and t Hoen,P.A. (2008) CORE_TF: a user-friendly interface to identify evolutionary conserved transcription factor binding sites in sets of co-regulated genes. *BMC bioinformatics*, **9**, 495.
29. Cheung,T.H., Kwan,Y.L., Hamady,M. and Liu,X. (2006) Unraveling transcriptional control and cis-regulatory codes using the software suite GeneACT. *Genome Biol.*, **7**, R97.
30. Bulyk,M.L. (2003) Computational prediction of transcription-factor binding site locations. *Genome Biol.*, **5**, 201.
31. Brown,C.T. (2008) Computational approaches to finding and analyzing cis-regulatory elements. *Methods Cell Biol.*, **87**, 337–365.

32. Smith,A.D., Sumazin,P. and Zhang,M.Q. (2005) Identifying tissue-selective transcription factor binding sites in vertebrate promoters. *Proc. Natl Acad. Sci. USA*, **102**, 1560–1565.
33. Subramanian,A., Tamayo,P., Mootha,V.K., Mukherjee,S., Ebert,B.L., Gillette,M.A., Paulovich,A., Pomeroy,S.L., Golub,T.R., Lander,E.S. *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA*, **102**, 15545–15550.
34. Bussemaker,H.J., Li,H. and Siggia,E.D. (2001) Regulatory element detection using correlation with expression. *Nature Genet.*, **27**, 167–171.
35. Conlon,E.M., Liu,X.S., Lieb,J.D. and Liu,J.S. (2003) Integrating regulatory motif discovery and genome-wide expression analysis. *Proc. Natl Acad. Sci. USA*, **100**, 3339–3344.
36. Das,D., Banerjee,N. and Zhang,M.Q. (2004) Interacting models of cooperative gene regulation. *Proc. Natl Acad. Sci. USA*, **101**, 16234–16239.
37. Barrett,T., Suzek,T.O., Troup,D.B., Wilhite,S.E., Ngau,W.C., Ledoux,P., Rudnev,D., Lash,A.E., Fujibuchi,W. and Edgar,R. (2005) NCBI GEO: mining millions of expression profiles – database and tools. *Nucleic Acids Res.*, **33**, D562–D566.
38. Parkinson,H., Kapushesky,M., Shojatalab,M., Abeygunawardena,N., Coulson,R., Farne,A., Holloway,E., Kolesnykov,N., Lilja,P., Lukk,M. *et al.* (2007) ArrayExpress – a public database of microarray experiments and gene expression profiles. *Nucleic Acids Res.*, **35**, D747–D750.
39. Wu,Z., Irizarry,R.A., Gentleman,R., Martinez-Murillo,F. and Spencer,F. (2004) A model based background adjustment for oligonucleotide expression arrays. *J. Am. Stat. Association*, **99**, 909–917.
40. Gentleman,R.C., Carey,V.J., Bates,D.M., Bolstad,B., Dettling,M., Dudoit,S., Ellis,B., Gautier,L., Ge,Y., Gentry,J. *et al.* (2004) Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.*, **5**, R80.
41. Smyth,G.K., Michaud,J. and Scott,H.S. (2005) Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics*, **21**, 2067–2075.
42. Benjamini,Y. and Hochberg,Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc.*, **57**, 289–300.
43. Karolchik,D., Kuhn,R.M., Baertsch,R., Barber,G.P., Clawson,H., Diekhans,M., Giardine,B., Harte,R.A., Hinrichs,A.S., Hsu,F. *et al.* (2008) The UCSC Genome Browser Database: 2008 update. *Nucleic Acids Res.*, **36**, D773–D779.
44. Schones,D.E., Smith,A.D. and Zhang,M.Q. (2007) Statistical significance of cis-regulatory modules. *BMC Bioinformatics*, **8**, 19.
45. Kanehisa,M., Goto,S., Kawashima,S., Okuno,Y. and Hattori,M. (2004) The KEGG resource for deciphering the genome. *Nucleic Acids Res.*, **32**, D277–D280.
46. Vastrik,I., D'Eustachio,P., Schmidt,E., Joshi-Tope,G., Gopinath,G., Croft,D., de Bono,B., Gillespie,M., Jassal,B., Lewis,S. *et al.* (2007) Reactome: a knowledge base of biologic pathways and processes. *Genome Biol.*, **8**, R39.
47. Lewis,B.P., Burge,C.B. and Bartel,D.P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, **120**, 15–20.
48. Wheeler,D.L., Barrett,T., Benson,D.A., Bryant,S.H., Canese,K., Church,D.M., DiCuccio,M., Edgar,R., Federhen,S., Helmberg,W. *et al.* (2005) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **33**, D39–D45.
49. Mulder,N.J., Apweiler,R., Attwood,T.K., Bairoch,A., Bateman,A., Binns,D., Bradley,P., Bork,P., Bucher,P., Cerutti,L. *et al.* (2005) InterPro, progress and status in 2005. *Nucleic Acids Res.*, **33**, D201–D205.
50. Bairoch,A., Apweiler,R., Wu,C.H., Barker,W.C., Boeckmann,B., Ferro,S., Gasteiger,E., Huang,H., Lopez,R., Magrane,M. *et al.* (2005) The Universal Protein Resource (UniProt). *Nucleic Acids Res.*, **33**, D154–D159.
51. Camon,E., Magrane,M., Barrell,D., Lee,V., Dimmer,E., Maslen,J., Binns,D., Harte,N., Lopez,R. and Apweiler,R. (2004) The Gene Ontology Annotation (GOA) Database: sharing knowledge in Uniprot with Gene Ontology. *Nucleic Acids Res.*, **32**, D262–D266.
52. Calvo,S., Jain,M., Xie,X., Sheth,S.A., Chang,B., Goldberger,O.A., Spinazzola,A., Zeviani,M., Carr,S.A. and Mootha,V.K. (2006) Systematic identification of human mitochondrial disease genes through integrative genomics. *Nature Genet.*, **38**, 576–582.
53. Mizukami,Y., Li,J., Zhang,X., Zimmer,M.A., Iliopoulos,O. and Chung,D.C. (2004) Hypoxia-inducible factor-1-independent regulation of vascular endothelial growth factor by hypoxia in colon cancer. *Cancer Res.*, **64**, 1765–1772.
54. Mizukami,Y., Fujiki,K., Duerr,E.M., Gala,M., Jo,W.S., Zhang,X. and Chung,D.C. (2006) Hypoxic regulation of vascular endothelial growth factor through the induction of phosphatidylinositol 3-kinase/Rho/ROCK and c-Myc. *J. Biol. Chem.*, **281**, 13957–13963.
55. Chi,J.T., Wang,Z., Nuyten,D.S., Rodriguez,E.H., Schaner,M.E., Salim,A., Wang,Y., Kristensen,G.B., Helland,A., Borresen-Dale,A.L. *et al.* (2006) Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. *PLoS Med.*, **3**, e47.
56. Denko,N.C., Fontana,L.A., Hudson,K.M., Sutphin,P.D., Raychaudhuri,S., Altman,R. and Giaccia,A.J. (2003) Investigating hypoxic tumor physiology through gene expression patterns. *Oncogene*, **22**, 5907–5914.
57. Gardner,L.B. and Corn,P.G. (2008) Hypoxic regulation of mRNA expression. *Cell Cycle*, **7**, 1916–1924.
58. Jin,H.O., An,S., Lee,H.C., Woo,S.H., Seo,S.K., Choe,T.B., Yoo,D.H., Lee,S.B., Um,H.D., Lee,S.J. *et al.* (2007) Hypoxic condition- and high cell density-induced expression of Redd1 is regulated by activation of hypoxia-inducible factor-1alpha and Sp1 through the phosphatidylinositol 3-kinase/Akt signaling pathway. *Cell Signal.*, **19**, 1393–1403.
59. Shoshani,T., Faerman,A., Mett,I., Zelin,E., Tenne,T., Gorodin,S., Moshel,Y., Elbaz,S., Budanov,A., Chajut,A. *et al.* (2002) Identification of a novel hypoxia-inducible factor 1-responsive gene, RTP801, involved in apoptosis. *Mol. Cell Biol.*, **22**, 2283–2293.
60. Leonard,M.O., Cottell,D.C., Godson,C., Brady,H.R. and Taylor,C.T. (2003) The role of HIF-1 alpha in transcriptional regulation of the proximal tubular epithelial cell response to hypoxia. *J. Biol. Chem.*, **278**, 40296–40304.
61. Takahashi,Y., Takahashi,S., Shiga,Y., Yoshimi,T. and Miura,T. (2000) Hypoxic induction of prolyl 4-hydroxylase alpha (I) in cultured cells. *J. Biol. Chem.*, **275**, 14139–14146.
62. Wellmann,S., Bettkober,M., Zelmer,A., Seeger,K., Faigle,M., Eltzschig,H.K. and Buhrer,C. (2008) Hypoxia upregulates the histone demethylase JMJD1A via HIF-1. *Biochem. Biophys. Res. Commun.*, **372**, 892–897.
63. Jean,J.C., Rich,C.B. and Joyce-Brady,M. (2006) Hypoxia results in an HIF-1-dependent induction of brain-specific aldolase C in lung epithelial cells. *Am. J. Physiol. Lung Cell Mol. Physiol.*, **291**, L950–L956.
64. Barrera,L.O., Li,Z., Smith,A.D., Arden,K.C., Cavenee,W.K., Zhang,M.Q., Green,R.D. and Ren,B. (2008) Genome-wide mapping and analysis of active promoters in mouse embryonic stem cells and adult organs. *Genome Res.*, **18**, 46–59.
65. Cid,E., Gomis,R.R., Geremia,R.A., Guinovart,J.J. and Ferrer,J.C. (2000) Identification of two essential glutamic acid residues in glycogen synthase. *J. Biol. Chem.*, **275**, 33614–33621.
66. Bourgeron,T., Chretien,D., Poggi-Bach,J., Doonan,S., Rabier,D., Letouze,P., Munnich,A., Rotig,A., Landrieu,P. and Rustin,P. (1994) Mutation of the fumarase gene in two siblings with progressive encephalopathy and fumarase deficiency. *J. Clin. Invest.*, **93**, 2514–2518.
67. Li,J., Mahajan,A. and Tsai,M.D. (2006) Ankyrin repeat: a unique motif mediating protein-protein interactions. *Biochemistry*, **45**, 15168–15178.
68. Cockman,M.E., Lancaster,D.E., Stolze,I.P., Hewitson,K.S., McDonough,M.A., Coleman,M.L., Coles,C.H., Yu,X., Hay,R.T., Ley,S.C. *et al.* (2006) Posttranslational hydroxylation of ankyrin repeats in IkappaB proteins by the hypoxia-inducible factor (HIF) asparaginyl hydroxylase, factor inhibiting HIF (FIH). *Proc. Natl Acad. Sci. USA*, **103**, 14767–14772.
69. Li,Q. and Verma,I.M. (2002) NF-kappaB regulation in the immune system. *Nat. Rev. Immunol.*, **2**, 725–734.
70. Dreyfus,D.H., Nagasawa,M., Gelfand,E.W. and Ghoda,L.Y. (2005) Modulation of p53 activity by IkappaBalpha: evidence suggesting a common phylogeny between NF-kappaB and p53 transcription factors. *BMC Immunol.*, **6**, 12.

71. Nickols,N.G., Jacobs,C.S., Farkas,M.E. and Dervan,P.B. (2007) Modulating hypoxia-inducible transcription by disrupting the HIF-1-DNA interface. *ACS Chem. Biol.*, **2**, 561–571.
72. Choi,K.O., Lee,T., Lee,N., Kim,J.H., Yang,E.G., Yoon,J.M., Lee,T.G. and Park,H. (2005) Inhibition of the catalytic activity of hypoxia-inducible factor-1alpha-prolyl-hydroxylase 2 by a MYND-type zinc finger. *Mol. Pharmacol.*, **68**, 1803–1809.
73. Yan,S.F., Lu,J., Zou,Y.S., Soh-Won,J., Cohen,D.M., Buttrick,P.M., Cooper,D.R., Steinberg,S.F., Mackman,N., Pinsky,D.J. *et al.* (1999) Hypoxia-associated induction of early growth response-1 gene expression. *J. Biol. Chem.*, **274**, 15030–15040.
74. Sigoillot,F.D., Berkowski,J.A., Sigoillot,S.M., Kotsis,D.H. and Guy,H.I. (2003) Cell cycle-dependent regulation of pyrimidine biosynthesis. *J. Biol. Chem.*, **278**, 3403–3409.
75. Lukowiak,A.A., Granneman,S., Mattox,S.A., Speckmann,W.A., Jones,K., Pluk,H., Venrooij,W.J., Terns,R.M. and Terns,M.P. (2000) Interaction of the U3-55k protein with U3 snoRNA is mediated by the box B/C motif of U3 and the WD repeats of U3-55k. *Nucleic Acids Res.*, **28**, 3462–3471.
76. Kollberg,G., Tulinius,M., Gilljam,T., Ostman-Smith,I., Forsander,G., Jotorp,P., Oldfors,A. and Holme,E. (2007) Cardiomyopathy and exercise intolerance in muscle glycogen storage disease 0. *N. Engl. J. Med.*, **357**, 1507–1514.
77. Takaluoma,K., Hyry,M., Lantto,J., Sormunen,R., Bank,R.A., Kivirikko,K.I., Myllyharju,J. and Soininen,R. (2007) Tissue-specific changes in the hydroxylysine content and cross-links of collagens and alterations in fibril morphology in lysyl hydroxylase 1 knock-out mice. *J. Biol. Chem.*, **282**, 6588–6596.
78. Beighton,P., De Paepe,A., Steinmann,B., Tsipouras,P. and Wenstrup,R.J. (1998) Ehlers-Danlos syndromes: revised nosology, Villefranche, 1997. Ehlers-Danlos National Foundation (USA) and Ehlers-Danlos Support Group (UK). *Am. J. Med. Genet.*, **77**, 31–37.
79. Shan,Z.X., Lin,Q.X., Fu,Y.H., Deng,C.Y., Zhou,Z.L., Zhu,J.N., Liu,X.Y., Zhang,Y.Y., Li,Y., Lin,S.G. *et al.* (2009) Upregulated expression of miR-1/miR-206 in a rat model of myocardial infarction. *Biochem. Biophys. Res. Commun.*, **381**, 597–601.
80. Chen,Y. and Gorski,D.H. (2008) Regulation of angiogenesis through a microRNA (miR-130a) that down-regulates antiangiogenic homeobox genes GAX and HOXA5. *Blood*, **111**, 1217–1226.
81. Xenaki,G., Ontikatze,T., Rajendran,R., Stratford,I.J., Dive,C., Krstic-Demonacos,M. and Demonacos,C. (2008) PCAF is an HIF-1alpha cofactor that regulates p53 transcriptional activity in hypoxia. *Oncogene*, **27**, 5785–5796.
82. Westenfelder,C., Biddle,D.L. and Baranowski,R.L. (1999) Human, rat, and mouse kidney cells express functional erythropoietin receptors. *Kidney Int.*, **55**, 808–820.
83. Forsythe,J.A., Jiang,B.H., Iyer,N.V., Agani,F., Leung,S.W., Koos,R.D. and Semenza,G.L. (1996) Activation of vascular endothelial growth factor gene transcription by hypoxia-inducible factor 1. *Mol. Cell Biol.*, **16**, 4604–4613.
84. Pescador,N., Cuevas,Y., Naranjo,S., Alcaide,M., Villar,D., Landazuri,M.O. and Del Peso,L. (2005) Identification of a functional hypoxia-responsive element that regulates the expression of the egl nine homologue 3 (egln3/phd3) gene. *Biochem. J.*, **390**, 189–197.
85. Coulet,F., Nadaud,S., Agrapart,M. and Soubrier,F. (2003) Identification of hypoxia-response element in the human endothelial nitric-oxide synthase gene promoter. *J. Biol. Chem.*, **278**, 46230–46240.
86. Hirao,K., Natsuka,Y., Tamura,T., Wada,I., Morito,D., Natsuka,S., Romero,P., Sleno,B., Tremblay,L.O., Herscovics,A. *et al.* (2006) EDEM3, a soluble EDEM homolog, enhances glycoprotein endoplasmic reticulum-associated degradation and mannose trimming. *J. Biol. Chem.*, **281**, 9650–9658.
87. Semenza,G.L. (2007) Oxygen-dependent regulation of mitochondrial respiration by hypoxia-inducible factor 1. *Biochem. J.*, **405**, 1–9.
88. Fontanesi,F., Jin,C., Tzagoloff,A. and Barrientos,A. (2008) Transcriptional activators HAP/NF-Y rescue a cytochrome c oxidase defect in yeast and human cells. *Hum. Mol. Genet.*, **17**, 775–788.
89. Fukuda,R., Zhang,H., Kim,J.W., Shimoda,L., Dang,C.V. and Semenza,G.L. (2007) HIF-1 regulates cytochrome oxidase subunits to optimize efficiency of respiration in hypoxic cells. *Cell*, **129**, 111–122.
90. Dimova,E.Y., Jakubowska,M.M. and Kietzmann,T. (2007) CREB binding to the hypoxia-inducible factor-1 responsive elements in the plasminogen activator inhibitor-1 promoter mediates the glucagon effect. *Thromb. Haemost.*, **98**, 296–303.
91. Choi,J.H., Park,M.J., Kim,K.W., Choi,Y.H., Park,S.H., An,W.G., Yang,U.S. and Cheong,J. (2005) Molecular mechanism of hypoxia-mediated hepatic gluconeogenesis by transcriptional regulation. *FEBS Lett.*, **579**, 2795–2801.
92. Sanchez-Elsner,T., Botella,L.M., Velasco,B., Langa,C. and Bernabeu,C. (2002) Endoglin expression is regulated by transcriptional cooperation between the hypoxia and transforming growth factor-beta pathways. *J. Biol. Chem.*, **277**, 43799–43808.
93. Choi,D.W., Seo,Y.M., Kim,E.A., Sung,K.S., Ahn,J.W., Park,S.J., Lee,S.R. and Choi,C.Y. (2008) Ubiquitination and degradation of homeodomain-interacting protein kinase 2 by WD40 repeat/SOCS box protein WSB-1. *J. Biol. Chem.*, **283**, 4682–4689.
94. Nardinocchi,L., Puca,R., Guidolin,D., Belloni,A.S., Bossi,G., Michiels,C., Sacchi,A., Onisto,M. and D'Orazi,G. (2009) Transcriptional regulation of hypoxia-inducible factor 1alpha by HIPK2 suggests a novel mechanism to restrain tumor growth. *Biochim. Biophys. Acta*, **1793**, 368–377.
95. Siepel,A., Bejerano,G., Pedersen,J.S., Hinrichs,A.S., Hou,M., Rosenbloom,K., Clawson,H., Spieth,J., Hillier,L.W., Richards,S. *et al.* (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. *Genome Res.*, **15**, 1034–1050.