**Supplemental Materials: Supplemental Figures and Legends, and Supplemental Tables for "The evolutionarily conserved Arginyltransferase1 mediates a pVHL-independent oxygen-sensing pathway in mammalian cells"**



**Supplemental Figure S1. Ate1 controls glycolysis by regulating HIF1α, related to Figure 1 and 2.**

**A)** The color of culture media for actively growing WT and *ATE1*-KO MEF grown for two days. The cells were still sub-confluent. The media contains phenol red and yellow color indicates acidic pH.

**B)** Concentration of ATP in WT and *ATE1*-KO MEF measured by luciferase assay and normalized by cell number and average cell volume. The measurements were repeated 5 times (n=5) and the value of WT cells at one measurement was set as 1.0 and the other data were normalized to it.

**C)** The protein level of ATE1 in two human prostate cancer cell lines: PC3 and PC3-ML. PC3-ML is derived from PC3 but naturally has a lower ATE1 level and behaves much more aggressively in most cancer phenotypes<sup>1</sup>. In this case, the signal of the immunoblot was documented in the film while, elsewhere in this study, the acquisitions were done by a digital imaging platform.

**D)** Antibody array showing the changes of protein levels associated the manipulation of ATE1 levels. Considering that arginylation often induces degradation, we aimed to identify proteins whose levels are negatively correlated with ATE1 level. The downregulation of ATE1 was performed in human prostate cancer cell PC3 with specific shRNA in comparison to nonsilencing (NS) shRNA. The upregulation of ATE1 was accomplished with stable expression of GFP-fused mouse ATE1 isoform 1 (a ubiquitously expressed splice variant<sup>2</sup>) in PC3-ML by a moderate promoter at a level close to endogenous ATE1 as previously reported<sup>3,4</sup>. GFP was used as a control. An antibody array for 86 proteins commonly associated with metabolism and oncology was used to measure the changes of protein levels compared to the corresponding controls. The blue color indicates a reduction while red indicates an elevation. The box with a green dotted line indicates the proteins whose level is in an inverse relationship to ATE1 (i.e., increased in ATE1-downregulation and decreased with ATE1-upregulation). A red arrow indicates the position of HIF1 $\alpha$  – one of the positive hits in this screening – showing an inverse correlation with ATE1 levels.



## **Supplemental Figure S2**. **Characterization of HIF1α (mRNA and protein) and additional examples of mass spectrometry data showing the N-terminus arginylation of HIF1α, related to Figure 3 and 4.**

**A)** Graphs showing correlation of mRNA expression between *ATE1* and *HIF1A* genes in different tumor tissue collections in The Cancer Genome Atlas (TCGA) including: testicular germ cell tumors (TGCT), lowgrade glioma (LGG), sarcoma (SARC), Pheochromocytoma and Paraganglioma (PCPG), and skin cutaneous melanoma (SKCM). The mRNA data (Pan-Cancer normalized RSEM values) were obtained from https://pancanatlas.xenahubs.net. Spearman's correlation coefficients were indicated in red fonts at the bottom of each plot. The number of tissue samples for each type of tumor were: TGCT: 134; LGG: 515; SARC: 259; PCPG: 179; SKCM: 103. The significance (p-value) was assessed for each analysis and shown in parentheses. The symbols \*, \*\*, and \*\*\* indicate p-values of <0.05, <0.005, and <0.0005, respectively.

**B)** Representative immunoblots showing the levels of HIF1 $\alpha$  in WT and ATE1-KO MEF treated with DMSO over a time course as a reagent control for the cycloheximide treatment (Fig. 4B). The β-actin is a loading control.

**C)** Representative immunoblots and the corresponding quantifications (n=3) showing the HIF1α level, with β-actin as a loading control, in WT and ATE1-KO MEF treated with proteasome inhibitor MG132 for 1 hour. The short duration of the MG132 treatment is intended to minimize potential impacts from transcriptional changes.

**D)** A representative image of SDS-PAGE stained with Coomassie blue for proteins purified by NTA-Ni resin from WT and *ATE1*-KO MEF expressing recombinant mouse HIF1 $\alpha$  fused with a C-terminal 6xHis tag. The boxes with the dotted line highlights the gel areas for the anticipated size of the full-length HIF1α-6xHis, which were cut and submitted separately for analysis by Mass Spectrometry.

**E)** In addition to the peptide sequence corresponding to REGAGENEK (shown in Fig 4C), we also detected peptide peaks corresponding to REGAGENEKK, which is a predicted alternative cleavage pattern of trypsin in arginylated HIF1α. The shown example is obtained from the proteomics core facility in USF, in which the submitted sample (SDS-PAGE slice) was analyzed within 2 days. The assignment of the b-ions (blue) and y-ions (red), as well as the PEP and Andromeda scores are indicated on the spectra.

**F)** An example of detected peptide peaks corresponding to REGAGENEK from the predicted arginylated N-terminus of HIF1α, obtained from the laboratory of Dr. Aldrin Gomes at the University of California, Davis. The mass of arginine (R) is predicted as 156.10. The assignment of b-ions (blue) and y-ions (red) are indicated on the spectra. The fragment error tolerance for the arginylated peptides was < 0.05 daltons. In this facility, the submitted samples (SDS-PAGE slices) were analyzed within one week.



Mass Spec spectra of HIF1 $\alpha$  purified from ATE1-KO MEF showing non-arginylated N-termini

**Supplemental Figure S3. Examples of Mass Spec results on HIF1**α **purified from** *ATE1***-KO MEF showing peptide peaks of the non-arginylated N-terminus, related to Figure 4.** 

In HIF1α purified from arginylation-deficient *ATE1*-KO MEF, we did not detect any peptide peaks corresponding to N-terminal arginylated HIF1α. The two displayed examples are both acquired in the proteomics core facility of USF, in which the samples were analyzed within 2 days. In **A)**, a spectrum corresponding to N-terminal acetylated HIF1α is shown. In **B)**, a spectrum corresponding to the Nterminal sequence of HIF1α, after the initial methionine is removed, is shown. The assigned bions (blue) and y-ions (red), as well as the PEP and Andromeda scores are indicated.



## **Supplemental Figure S4. Additional evidence showing the relationship between UBRs and HIF1**α**, related to Figure 5.**

**A)** Left: steady-state levels of recombinant HIF1α expressed in the pVHL-deficient human renal carcinoma cell line UOK111, after treatment of proteasome inhibitor MG132 for 3 hours. The different recombinant HIF1 $\alpha$  include the non-arginylated (M-), arginylation-retarding (G-), constitutively arginylated (R-), or arginylation-eligible (E-) forms. Antibody against HA tag was used to detect the level of HIF1 $\alpha$  while the level of GFP is used as loading controls to normalize the difference in expression efficiency of the vector. Right: the graph showing the quantification (n=3).

**B)** Coexpression correlations of *UBR2-5* with 50 HIF1A target genes in human tumor tissues. These samples include testicular germ cell tumors (TGCT), low-grade glioma (LGG), sarcoma (SARC), Pheochromocytoma, and Paraganglioma (PCPG), and skin cutaneous melanoma (SKCM) in The Cancer Genome Atlas (TCGA). The number of tissue samples for each type of tumor were indicated in the parenthesis. Note that not all genes were available from all tumor types in these databases. The mRNA data were Pan-Cancer normalized and obtained from [https://pancanatlas.xenahubs.net.](https://pancanatlas.xenahubs.net/) Spearman's correlation was used to calculate the co-expression correlation between *ATE1* and 50 HIF1α-target genes or a random set of 500 genes (Random-500, similarly as done in published studies $8-10$ ). The significance (p-value) was calculated by Mann−Whitney U test. The signs \*, \*\*, and \*\*\* indicate p-values of <0.05, <0.005, and <0.0005, respectively. See also Suppl. Table S1 for the list of the 50 HIF1 $\alpha$  target genes that are known to be transcriptionally activated by HIF1 $\alpha$  to be used for this co-expression analysis.

**C)** Representative immunoblots (left) showing the level of HIF1α with β-actin as loading control in *ATE1*- KO cells with *UBR1* knockdown, in comparison to non-silencing (NS) shRNA. The quantification on the right was based on three repeats (n=3).

**D)** Representative immunoblots and quantification based on five repeats (n=5) showing the knockdown efficiency of *UBR1* in HFF cells treated with non-silencing (ns) shRNA or specific shRNA against *UBR1*. The β-actin was used as a loading control.

**E)** The mRNA levels of HIF1A in HFF cells stably expressing non-silencing (ns) shRNA or shRNA against *UBR1*, measured by quantitative PCR. The human β-actin gene (ACTB) was used as a housekeeping gene for normalization. The quantification was based on two independent experiments with triplicates (n=3). **F)** Representative images of an immunoblot showing the steady-state protein levels of HIF1 $\alpha$  in HFF cells with non-silencing (ns) shRNA or specific shRNA against *UBR1* (sh*UBR1*). The β-actin was used as a loading control. The quantification shown on the right side was calculated from four independent repeats (n=4). Error bars represent SEM.



**Supplemental Figure S5. The signal of hydroxylated-HIF1α and total HIF1α under different conditions, related to Figure 6**

**A)** As a validation to the specificity of the anti-hydroxylated(Hx) HIF1 $\alpha$  (from Cell Signaling Inc) as shown in Fig. 6C, an immunoblot showing the detection signals of this antibody in WT MEF incubated under normoxia or hypoxia (0.5%  $O_2$ ) and treated with MG132 for 9 hours. To facilitate the specific comparison of the hydroxylation level of HIF1 $\alpha$ , the loading of these samples was adjusted so that the two samples have comparable loading amounts of full-length HIF1 $\alpha$  (probed by anti-HIF1 $\alpha$ ). As demonstrated by the immunoblots, the concerned anti-HxHIF1 $\alpha$  antibody showed reduced signal with cells that were incubated with hypoxia – a condition expected to minimize the hydroxylation of HIF1 $\alpha$ .

**B)** Immunoblots showing the levels of Hx-HIF1 $\alpha$ , probed with rabbit HIF1 $\alpha$  hydroxyl-P564(human)/P577(mouse) antibody (Rockland, Limerick, PA, Cat# 100-401-A25), in WT and *ATE1*-KO MEF treated with proteasome inhibitor MG132 (20 $\mu$ M) for 6 hours. The level of full-length HIF1 $\alpha$  shown in the bottom, probed with anti-HIF1 $\alpha$ , was pre-adjusted by the rationales as mentioned in (A).

**C)** Immunoblots showing the levels of endogenous HIF1 $\alpha$  (with  $\beta$ -actin as loading controls) in WT and *ATE1*-KO MEF, treated with PHD inhibitor DMOG (20µM) for 6 hours. DMSO was used as a reagent control.

**D)** On the left are representative Immunoblots showing the levels of recombinant E-HIF1 $\alpha$  fused with a C-terminal HA tag and carrying the PAPG mutations to the two critical proline residues (as described in

Fig.5A). This protein was stably expressed in arginylation-deficient *ATE1*-KO MEF, which were treated with PHD inhibitor DMOG (20µM), or the reagent control DMSO, for 6 hours. The level of the recombinant HIF1α was probed with anti-HA. The level of β-actin was used as a loading control. On the right is the quantification of the level of E-HIF1 $\alpha$ -HA (n=3).



**Supplemental Figure S6. Mutation burden on** *UBR1-5, VHL and ATE1* **in different types of solid tumors, related to Figure 5.** The mutation burdens were analyzed by specific gene queries in cBioPortal for Cancer Genomics (https://www.cbioportal.org/), which contains genomic sequencing data from multiple published sources. The results are separated by cancer studies. The data corresponding to the skin,

esophagus/stomach, uterus and kidney cancer are shown as examples. In most examined cases, collective mutation burdens in the five *UBR* family genes (*UBR1-5*) are nearly 10-fold higher than *VHL*. Particularly, individual mutation burdens in *UBR4* or *UBR5* are 3-5 folds higher than in *VHL*.

The mutation burdens in *ATE1* is still higher than *VHL* in most cases, albeit lower than the *UBR* genes. This discrepancy may be explained by the "moonlighting" function of ATE1 in mitochondria<sup>11</sup>, which is likely separate from its role in mediating protein degradation through the ubiquitin-proteasome system. Considering that mitochondrial functions are essential for certain types of cancer<sup>12</sup>, this may become a limiting factor for mutations on *ATE1*.

Only in kidney cancer, we observed a higher mutation burden on *VHL* than *UBRs* or *ATE1*.

compared with human HIF1 $\alpha$  protein (UniProtKB - Q16665) MEPSNRFDTTEPYSVLNRKERSRELAQKRRTTYKGLMKDLADELPFSKDVVSQVDYNSRLRLALCF XP 011403284.1  $\mathbf{1}$  $Q16665$ \_hHIF1A MEGAGGANDKKKISSERRKEKSRDAARSRRSKESEVFYELAHQLPLPHNVSSHLDKASVMRLTISY  $\mathbf{1}$ 1、10、1000,木。 1、木木木(木木) 木(1、木木(1、1、1、1) (木木(1木木(1):1)木。木(1木木(1木、木。(木木(1):1 cons  $\overline{**}$  :  $\cdots$ XP 011403284.1 358 Q16665\_hHIF1A TSSLFDKLKKEPDALTLLAPAAGDTIISLDFGSNDTETDDQQLEEVPLYNDVMLPSPNEKLQNINL 383 \*.\*: .. . :\*  $: * * : : : : : : : :$  $\ldots$ XP 011403284.1 TPTAVSPTPSFAESGITT------------DYGTYSPQISEEGFIKHELLSPTF-----------556  $Q16665$  hHIF1A 512 EPNSPSEYCFYVDSDMVNEFKLELVEKLFAEDTEAKNPFSTQDTDLDLEMLAPYIPMDDDFQLRSF cons \*  $: *$  :::  $: *$  \*:\*:\* :  $*$ .: \*  $: . :$ \*.:.. В Essential regions of HIF1a-like protein, NCBI Reference Sequence: XP 023346906.1 [Eurytemora affinis (Copepods)] compared with human HIF1 $\alpha$  protein (UniProtKB - Q16665) XP 023346906.1 MCGGVAGAQEPSQRFLKEFLKGSKRKRINSLEGLDGKAENQQEIQDLGALQQDNRKEKSRDAARNR  $\mathbf{1}$  $Q16665$ \_hHIF1A  ${\bf 1}$  $*5***$  $: . . . .$ \*\*\*\*\*\*\*\*\*\*\*  $\cdot$  \* cons . . . XP\_023346906.1 DMGVDETFNLENPENMNEILEMLEQDAEELKMMDEEK------------KVKEELEKMEEVEKVEL 403 DMKMTQLFTKVESEDTSSLFDKLKKEPDALTLLAPAAGDTIISLDFGSNDTETDDQQLEEVPLYND  $Q16665$  hHIF1A 368 \*\* : : \*. :.\*: ..::: \*::..: \*.::  $...: : : : : ***$ cons  $\sim$  1 . . .

A Essential regions of HIF1 $\alpha$ -like protein, NCBI accession no. XP\_011403284.1 [Amphimedon queenslandica (sponge)],

XP 023346906.1 555 -------LGTGRRLSSDLENLSPFIGDECVPLNKTETLFQPLNL-----------------DSFDD  $Q16665$ \_hHIF1A 542 EDTEAKNPFSTQDTDLDLEMLAPYIPMDDDFQLRSFDQLSPLESSSASPESASPQSTVTVFQQTQI  $: :$   $****...*$  $\cdot$   $\cdot$  \*\*: cons  $\overline{\phantom{a}}$ :  $\mathbf{H}$  $\mathbb{R}^2$  . The  $\mathbb{R}^2$ 

## HIF2 $\alpha$  levels in WT and ATE1-KO MEF

C



**Supplemental Figure S7. Sequence alignment of HIF1α-like proteins in different species, and expression level of HIF2α in WT and** *ATE1***-KO MEF, related to Figure 7**

**A)** and **B)** Representative HIF1α-like proteins in sponge (A) and Copepods (B), compared with canonical HIF1 $\alpha$  (UniprotKB – Q16665, human). Three segments are shown: the N-terminus and the two regions corresponding to the pVHL recognition sites in human HIF1 $\alpha$ . The pink boxes indicate arginylationeligible residues (D, E, N, Q, or C) at the  $2^{nd}$  residue of the N-terminus. The blue and the green box highlight the locations of the hydroxylation-eligible proline residue (P402 and P564) in human HIF1 $\alpha$ . Both HIF1 $\alpha$ -like proteins appear to have arginylation-eligible residues (D/E/N/Q/C) on the N-terminus, but only one hydroxylation-eligible proline (corresponding to P564 in human HIF1 $\alpha$ ). The signs "\*", ":", and "." Indicates high to low consensus (cons). The sign "&" indicates Cysteine (C) on the N-terminus, which can be oxidized or nitrosylated to mimic an acidic residue $^{13,14}$ .

No proteins containing the conserved functional domain (Pfam ID: 17211) of pVHL was found in Protista (Naegleria gruberi as example), Porifera (Amphimedon queenslandica), or Copepods (Eurytemora affinis) by searching the phylogenetic tree generated by pfam.xfam.org, or by performing BLASTp with the VHL-C domain or the full-length human pVHL (UniProtKB- P40337) at blast.ncbi.nlm.nih.gov with specific taxonomy. The presence of VHL-like protein was speculated in two species within the Porifera class Demosponges (Stylissa – taxid 237127 and Scopalina – taxid 85794)<sup>15</sup>. However, we did not find any meaningful match upon performing BLASTp by using the queries of VHL beta domain (Pfam ID: 01847), VHL box domain (Pfam ID: 17211), or full-length protein of human pVHL in these species.

**(C)** Representative immunoblots (left) and corresponding quantification graph (right, n=3) showing the levels of HIF2α in WT and *ATE1*-KO MEF with β-actin as loading controls.

## **Supplemental materials: Tables, Figures and Legends**





**Supplemental Table S1. The list of the 50 genes used as direct HIF1**α**-activating targets in this study, related to Figure 3.**

These include genes that were demonstrated with experimental data for their transcription activation by HIF1 $\alpha$ , and genes that are upregulated under hypoxic conditions and are predicted to bind HIF1 $\alpha$  in their promoter regions. The table is adapted from published studies<sup>5,6</sup>.

Supplemental Table S2. Oligonucleotides used in this study, related to STAR Methods and the Key **Resources Table** 



