

Supplementary Material (online)

Title: The RNA-protein interactome of differentiated kidney tubular epithelial cells

Running title: The kidney epithelial RBPome

Authors: Michael Ignarski, Constantin Rill, Rainer Kaiser, Madlen Kaldirim, René Neuhaus, Reza Esmailie, Xinping Li, Corinna Klein, Kathrin Bohl, Maike Petersen, Christian K. Frese, Martin Hoehne, Ilian Atanasov, Markus M. Rinschen, Katja Höpker, Bernhard Schermer, Thomas Benzing, Christoph Dieterich, Francesca Fabretti and Roman-Ulrich Müller.

Table of contents - Supplementary Tables and Figures	page 1
Supplementary Methods	page 2
Supplementary Figure legends	page 5
References for Supplementary Material	page 7

Supplementary Tables and Figures

Supplementary Figure 1.

Heatmap of 4 replicates of RNA interactome capture experiments in mIMCD-3 cells

Supplementary Figure 2.

Characterization of the mIMCD-3 RNA interactome

Supplementary Figure 3.

Expression and ciliary localization of novel RBPs

Supplementary Figure 4.

Hypoxia signaling associated changes

Supplementary Figure 5.

RNA-binding proteins in hypoxia-induced metabolic changes

Supplementary Table.

mIMCD-3 RNA interactome and proteome

Supplementary Methods

mRNA interactome capture

Cells were washed with 1X PBS and either directly scraped off with a rubber scraper, harvested in PBS, collected by centrifugation (500 g, 5 minutes, 4 °C), and snap-frozen in liquid nitrogen (-CL samples), or crosslinked on ice before, using a Spectralinker equipped with UVA bulbs (365 nm, 150 mJ/cm²). Cell pellets were resuspended in 12,5 ml of ice-cold lysis buffer (100 mM Tris pH 7.5, 500 mM LiCl, 10 mM EDTA pH 8, 1% LiDS, 5 mM DTT supplemented with protease inhibitors) and syringed six times with a 21G hollow needle to shear genomic DNA. Subsequently, oligo(dT) pulldown was performed using 2 ml of Dynabeads Oligo(dT) beads (ThermoScientific) per sample for 1 h at room temperature. Bead-RNP-complexes were concentrated on a magnet and washed three times with one lysis volume of lysis buffer and an additional three times with one lysis volume of NP-40 wash buffer (50 mM Tris pH 7.5, 140 mM LiCl, 2 mM EDTA pH 8, 0.5% NP40, 0.5 mM DTT (stock 1 M) supplemented with protease inhibitor). RNP complexes were eluted in 10 mM Tris-HCl pH 7.5, by incubating the beads at 80°C for 2 min, and another two rounds of pulldowns were performed on the same lysate. Bound RNA was digested with 10 U/ml RNase I and 125 U/ml Benzonase at 37°C for 3 h followed by over-night precipitation with acetone at -20 °C. Proteins were pelleted and resuspended in 8M urea. PAGE and silver staining (Pierce™ Silver Stain Kit) was performed on 10% of each sample to ensure sufficient yield. In order to lower the urea concentration < 1 M, samples were diluted with 9 volumes of 20 mM Tris HCl. Samples were digested with trypsin (~1 ng/μl) at 37°C overnight, stage-tipped¹ and then submitted to mass spectrometry.

Data analysis and statistics - proteomics

Protein identification: All mass spectrometric raw data were processed with MaxQuant² (version 1.5.3.8) using default parameters. Briefly, MS2 spectra were searched against the Uniprot MOUSE.fasta (downloaded 2017) database, including a list of common contaminants. False discovery rates on protein and PSM level were estimated by the target-decoy approach to 0.01 (Protein FDR) and 0.01 (PSM FDR) respectively. The minimal peptide length was set to 7 amino acids and carbamidomethylation at cysteine residues was considered as a fixed modification. Oxidation (M) was included as variable modification. The match-between runs option was enabled. LFQ quantification was enabled using default settings³.

Statistical analysis of both the whole proteome and the RNA interactome raw data was performed with the Perseus software⁴, version 1.6.0.7. As to the RNA interactome 1 out of 4 replicates was considered an outlier and excluded from the analysis due to a strongly lower protein detection compared to the other replicates (replicate #3, Suppl. Figure 1).

RNA interactome: iBAQ intensities⁵ were used in this analysis and transformed in log₂ values. Whole Proteome: LFQ intensities were transformed into log₂. Missing values were not imputed and significantly regulated proteins were identified using a two sample t-test (Permutated FDR = 0.1, s0 = 0.1), after filtering for proteins measured at least in 2 replicates out of 3 in one condition (normoxia or hypoxia).

Gene ontology (GO), PFAM, SMART and KEGG analysis: The uniprot IDs were annotated using Perseus software. Gene set enrichment analysis of the renal RNA interactome was calculated with the mIMCD-3 whole proteome data unfiltered as background (Fisher exact test, in Perseus⁴). Level of significance was set to an FDR < 0,05 using the Benjamini and Hochberg method.

Correlation analysis hypoxia vs. normoxia: The linear regression of \log_2 foldchange of hypoxia versus normoxia was calculated with R (version 3.4.4) using the method `lm()` which applies the method of least squares. This results in the formula $y = 0.3170 + 0.7069 \cdot x$ with \log_2 foldchange of hypoxia on x-axis and values of normoxia on y-axis. This linear regression has a standard deviation of 0.9459532. To get the proteins that are most different between hypoxia and normoxia, i.e. the "outliers", we calculated the linear equations that define the upper and lower boundary of the interval in which 95% of all measurements lie. This interval is defined by adding and subtracting 1.96 standard deviation to the regression line for the upper and lower boundary, respectively. The linear equation for the upper boundary is $y = 2.171068 + 0.7069 \cdot x$, the equation for the lower boundary is $y = -1.537068 + 0.7069 \cdot x$.

Western Blotting

Cells were lysed in RIPA buffer supplemented with 2 mM Na_3VO_4 and 25 nM PMSF, homogenized using a 23G cannula syringe, centrifuged (20.000 rcf, 4 °C, 10 min) and mixed with Laemmli buffer. Equal volumes of protein solution were resolved by SDS-PAGE and blotted onto PVDF-membranes. After incubating the blots with HIF1 α (Cayman Chemical, 10006421) as well as β -Tubulin antibody (DSHB, AB2315513) proteins have been visualized with enhanced chemiluminescence. Quantitative analysis was carried out using the digital densitometry tool of Image Studio 5.2 (LI-COR) normalizing the signal to tubulin as an endogenous control.

Quantification of ciliation

Three biological replicates were analyzed independently. mIMCD-3 cells stained for ciliary markers (pericentrin, acetylated tubulin) were analyzed in a randomized, single-blinded fashion, where both DAPI-positive nuclei and cilia were counted in various view-fields. The cilia/nucleus ratio was assessed for at least 20 view fields of three biological replicates (approx. 280 cells per replicate and condition). Statistical analysis was performed using student's T-test (unpaired, untailed) in GraphPad Prism v5.0.

Plasmids used in this study

Plasmid	Reference
AAV-CAGGS GFP	AddGene (22212)
AAV-CAGGS GFP::MFAP1	this study
AAVS1 1L TALEN	AddGene (35431)
AAVS1 1R TALEN	AddGene (35432)
pcDNA6 with N-term 3xFLAG::HIC2	this study
pcDNA6 with N-term 3xFLAG::GADD45GIP1	this study
pcDNA6 with N-term 3xFLAG::MFAP1	this study

For all plasmids cloned the sequence of the insert and the region of the vector in proximity to the restriction sites were verified using Sanger Sequencing in order to rule out mutations and confirm the cDNA to be in-frame.

Primers used in this study

Target	Forward primer sequence (5' to 3')	Reverse primer sequence (5' to 3')
--------	------------------------------------	------------------------------------

MFAP1	CGTAGATGCGGCCCTAGGTAGTTTTCCGCTTCTTGGCAGATG G	AGGGAACGCGTATGTCGGTCCCAAGCGCTCTCATG
HIC2	CGAAGACTACGCGTGTGTCTGGGCCCTTGGCACTC	CGAAGACGCGGCCCGCTAGGAGGGGGAGGTGTGCATG C
GADD45GIP 1	CGAAGACTACGCGTGC GGCGTCCGTGCGACAGG	CGAAGACGCGGCCGCTCAGGAGCTGGGTGCCCCAGAGG
empty AAVS 1 locus	CGGAACTCTGCCCTCTAACG	GTGAGTTTGCCAAGCAGTCA
integrated AAVS 1 locus	TATCCGCTCACAATTCCACA	GTGAGTTTGCCAAGCAGTCA

Primers for MFAP1, HIC2 and GADD45GIP1 contain sequence overhangs for restriction digestion with MluI (forward) and NotI (reverse). All primers used were ordered at Integrated DNA Technologies, Inc. (IDT).

Secondary antibodies used in this study

Name	Manufacturer	Cat. number
anti-rabbit goat-IgG HRP-conjugated	Dianova/Jackson ImmunoResearch	111-035-003
anti-mouse goat-IgG HRP-conjugated	Dianova/Jackson ImmunoResearch	115-035-003
Cy2-AffiniPure donkey anti-mouse IgG (H+L)	Dianova/Jackson ImmunoResearch	715-225-150
Cy3-AffiniPure donkey anti-rabbit IgG (H+L)	Dianova/Jackson ImmunoResearch	711-165-152
Alexa Fluor® 488 AffiniPure Donkey Anti-Mouse IgG (H+L)	Dianova/Jackson ImmunoResearch	715-545-150

Supplementary Figure / Table Legends

Supplementary Figure 1 - Heatmap of 4 replicates of RNA interactome capture experiments in mIMCD-3 cells

Hierarchical clustering of the protein intensities clearly separates replicate #03, both with and without crosslinking (blue) from the other replicates (red), consequently, this replicate was considered an outlier and excluded from the analysis. In addition, the red cluster separates crosslinked vs non-crosslinked samples in the next branches of the tree.

Supplementary Figure 2 - Characterization of the mIMCD-3 RNA interactome

A Criteria used for the distinction of 2 classes of RBPs defined in this manuscript. Class I RBPs are significantly enriched in the crosslinked samples (compared to non-crosslinked samples; t-test – FDR < 0.1) or – in cases in which no statistical testing could be performed – measured in ≥ 4 crosslinked samples and ≤ 1 in non-crosslinked. Previous identification in RIC screens is not used for the definition of this group⁶. Class II RBPs are detected as RNA-associated in our RIC but do not fulfil the criteria defined for class I. However, they have been reported in the literature as RNA-binders before.

B Gene ontology enrichment analysis (cellular component) for class I proteins using the whole unfiltered cell proteome (n=6033) as a background shows a clear overrepresentation of RNA-associated terms (Fisher's exact test, FDR < 0.05).

C KEGG pathways enrichment analysis for class I proteins using the whole unfiltered cell proteome (n=6033) as a background shows a clear overrepresentation of RNA-associated terms (Fisher's exact test, FDR < 0.05).

Supplementary Figure 3 - Expression and ciliary localization of novel RBPs

A Expression of novel RBPs in other cell lines. *Proteomicsdb* was screened for the expression of the 25 novel RBPs identified in mIMCD-3 cells. The table shows the presence or absence of expression for each protein in cell lines that were used in the experiments underlying the RBP compendium as described by Hentze et al⁶. (for details see <https://www.proteomicsdb.org/>).

B Ciliary localization of Kif13B and Mpdz. mIMCD-3 cells were stained with anti-Kif13B and anti-Mpdz antibodies (green). Cilia were visualized by staining with anti-Arl13B antibodies (magenta). DAPI was used as a nuclear counterstain (blue). Scale bar: 20 μ m.

Supplementary Figure 4 - Hypoxia signaling associated changes

A Hypoxia does not alter the degree of ciliation. Cilia were visualized by staining with anti-acetylated tubulin and centrosomes were visualized by staining with pericentrin antibodies. The percentage of ciliated cells does not change significantly in normoxia versus hypoxia. The cilia/nucleus ratio was assessed for at least 6 view fields of three biological replicates (approx. 280 cells per replicate and condition). The statistical analysis was performed using student's T-test (unpaired, untailed) in GraphPad Prism v5.0 (p-value: 0.12).

B Incubation of mIMCD-3 cells for 6 hours at 1% O₂ (as specified in Fig. 5A) induces a strongly increased nuclear signal as observed by immunofluorescence after staining for endogenous Hif1 α indicating efficient activation of HIF-signaling.

C This is confirmed by Western Blotting and staining for endogenous Hif1 α . Normalization to β -Tubulin as a loading control and densitometric analysis shows a highly significant induction (two-tailed t-test, ** represents a p-value < 0.01, error bars represent SEM).

D Limited hypoxia does not induce major changes of the whole-cell proteome. Hierarchical clustering of samples incubated in hypoxia or normoxia (see Fig. 5A) does not separate the samples by this criterion indicating – in line with the low number of significantly altered proteins (see Fig. 5B) – that only few proteins are altered in their abundance by short-term hypoxia.

Supplementary Figure 5 - RNA-binding proteins in hypoxia-induced metabolic changes

A RBPs involved in hypoxia-associated metabolic reprogramming. A schematic representation of HIF-regulated genes involved in metabolic reprogramming (adapted from Semenza et al., 2012⁷). Proteins identified as RBPs in our screen are highlighted in blue; proteins identified as RBPs in previous screens are highlighted in green.

B Proteins involved in metabolic reprogramming shown in the correlation plot depicting RBPs modulated between hypoxic and normoxic conditions. RBPs associated with metabolic reprogramming are highlighted in blue. Aldoa and Eno1 are below the 95% prediction interval, which suggested a significantly increased RNA binding in hypoxic conditions versus normoxia. For details of the correlation plot refer to Figure 5C.

C Proteins involved in metabolic reprogramming shown in the volcano plot illustrating differential protein abundance between the proteomes of hypoxic and normoxic mIMCD-3 cells. Indicated are the six RBPs involved in metabolic reprogramming, none of these show a significant change of abundance in response to the treatment. For details of the volcano plot refer to Figure 5 A.

Supplementary Table – mIMCD-3 RNA interactome and proteome

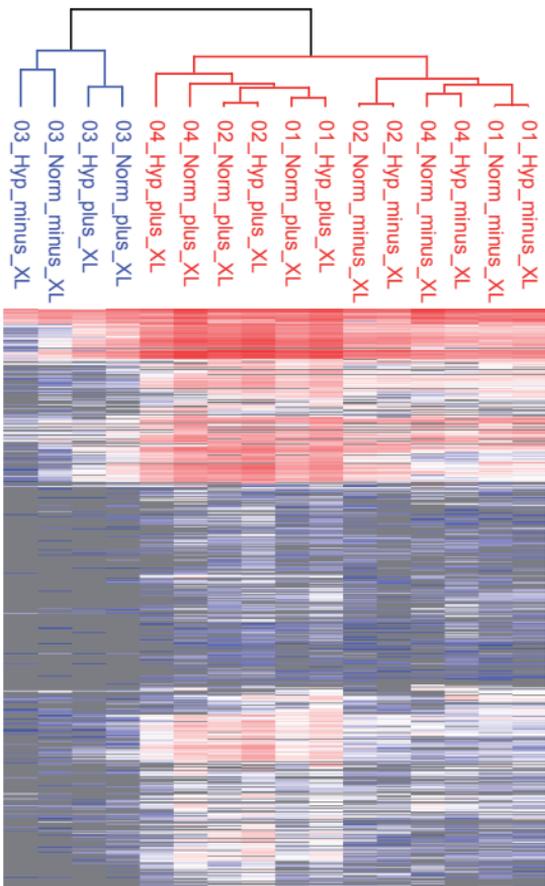
Proteome dataset (blue filling): hypoxia versus normoxia t-test results ($FDR \leq 0.1$, $s_0 = 0.1$) and mean LFQ intensities are presented.

RBPome dataset (green filling): the results for the 3 different t-tests (+CL_-CL; Hypoxia +CL_-CL and normoxia +CL_-CL) ($FDR \leq 0.1$; $s_0 = 0.1$) and the mean iBAQ intensities are presented. In addition to the dataset and the statistical analysis, we present some annotation columns for the RNA binding proteins in mIMCD-3 and in previously published datasets, Hif1a targets and interacting proteins and ciliary proteins.

References for Supplementary Material

1. Rappsilber J, Mann M, Ishihama Y: Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* 2: 1896–1906, 2007
2. Cox J, Mann M: MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26: 1367–1372, 2008
3. Cox J, Hein MY, Luber CA, Paron I, Nagaraj N, Mann M: MaxLFQ allows accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction. *Mol. Cell. Proteomics* mcp.M113.031591, 2014
4. Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J: The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat. Methods* 13: 731–740, 2016
5. Schwanhäusser B, Busse D, Li N, Dittmar G, Schuchhardt J, Wolf J, Chen W, Selbach M: Global quantification of mammalian gene expression control. *Nature* 473: 337–342, 2011
6. Hentze MW, Castello A, Schwarzl T, Preiss T: A brave new world of RNA-binding proteins. *Nat. Rev. Mol. Cell Biol.* 19: 327–341, 2018
7. Semenza GL: Hypoxia-inducible factors: mediators of cancer progression and targets for cancer therapy. *Trends Pharmacol. Sci.* 33: 207–214, 2012

Supplementary Figure 1

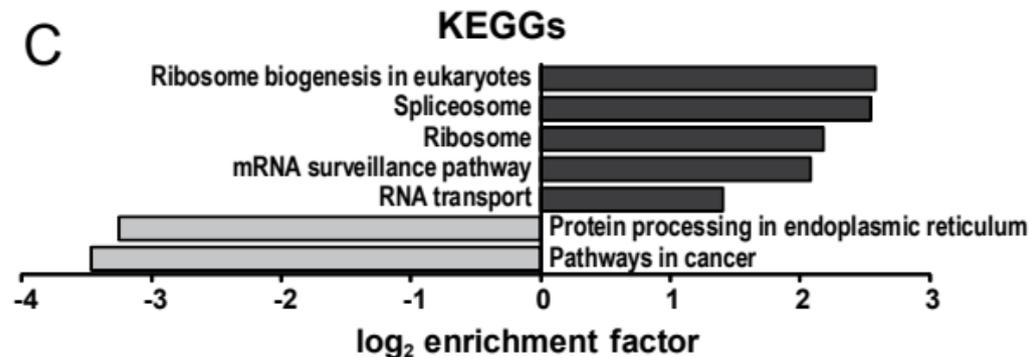


Supplementary Figure 2

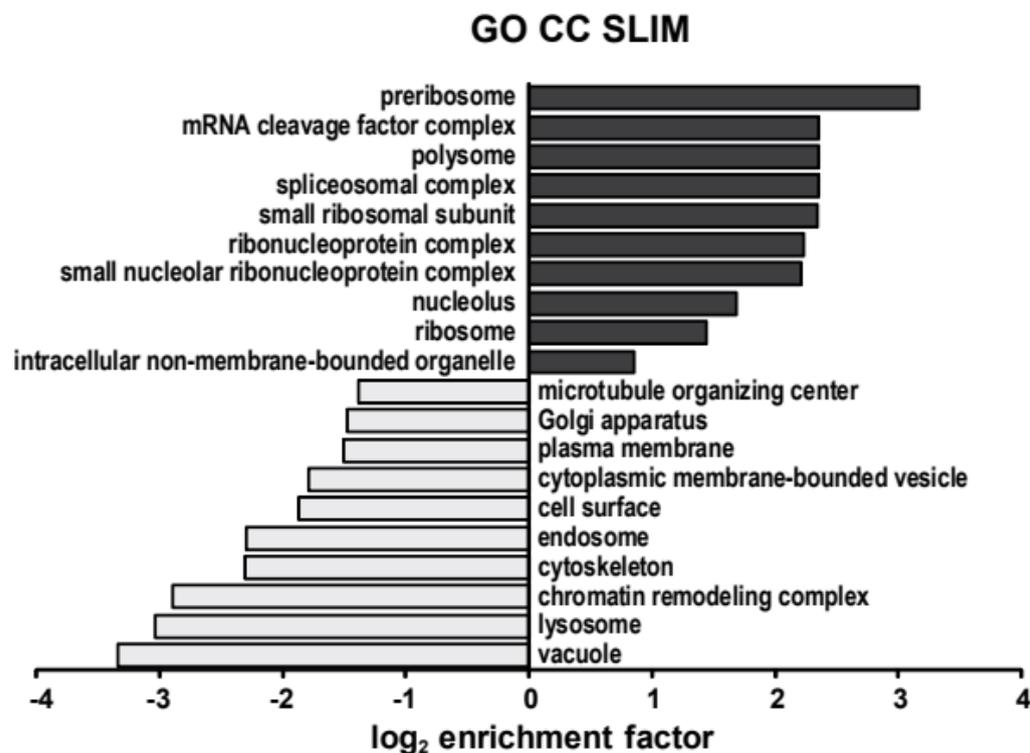
A

RBP class	Enriched in +CL vs -CL	Previously detected in RIC (Hentze et al., 2018)
class I	yes, as defined by either/or: - FDR < 0.1 in at least 1 t test - ≥ 4 measured values in +CL and ≤ 1 measured -CL	irrelevant for class I RBPs
class II	- FDR ≥ 0.1 in all t tests - < 4 measured values in +CL	yes

C



B



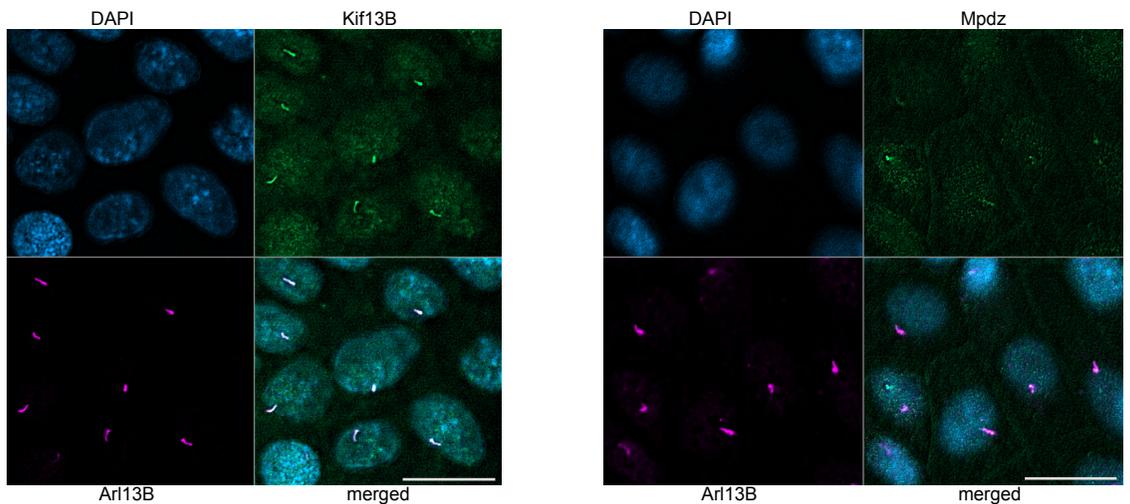
Supplementary Figure 3

A

Novel RBPs Gene name	cell lines		
	HEK-293	HeLa	K-562
<i>Gm9242;Gm6793</i>	Expressed	Expressed	Expressed
<i>Kif13b</i>	Expressed	Expressed	Expressed
<i>Mrpl47</i>	Expressed	Expressed	Expressed
<i>Usb1</i>	Expressed	Expressed	Expressed
<i>Dhx37</i>	Expressed	Expressed	Expressed
<i>Gadd45gip1</i>	Expressed	Expressed	Expressed
<i>Wasl</i>	Expressed	Expressed	Expressed
<i>Mpdz</i>	Expressed	Expressed	Expressed
<i>Nme1/2</i>	Expressed	Expressed	Expressed
<i>Tmem33</i>	Expressed	Expressed	Expressed
<i>Arhgap21</i>	Expressed	Expressed	Expressed
<i>Bclaf3</i>	Expressed	Expressed	Expressed
<i>Cdk1</i>	Expressed	Expressed	Expressed
<i>Ctc1</i>	Expressed	Expressed	Expressed
<i>Ephx1</i>	Expressed	Expressed	Expressed
<i>Hars</i>	Expressed	Expressed	Expressed
<i>Mcm5</i>	Expressed	Expressed	Expressed
<i>No66</i>	Expressed	Expressed	Expressed
<i>Ptrf</i>	Expressed	Expressed	Expressed
<i>Arid2</i>	Expressed	Expressed	Expressed
<i>Bcas2</i>	Expressed	Expressed	Expressed
<i>Ccdc94</i>	Expressed	Expressed	Expressed
<i>Oasl2</i>	Expressed	Expressed	Expressed
<i>Reps1</i>	Expressed	Expressed	Expressed
<i>Usp7</i>	Expressed	Expressed	Expressed

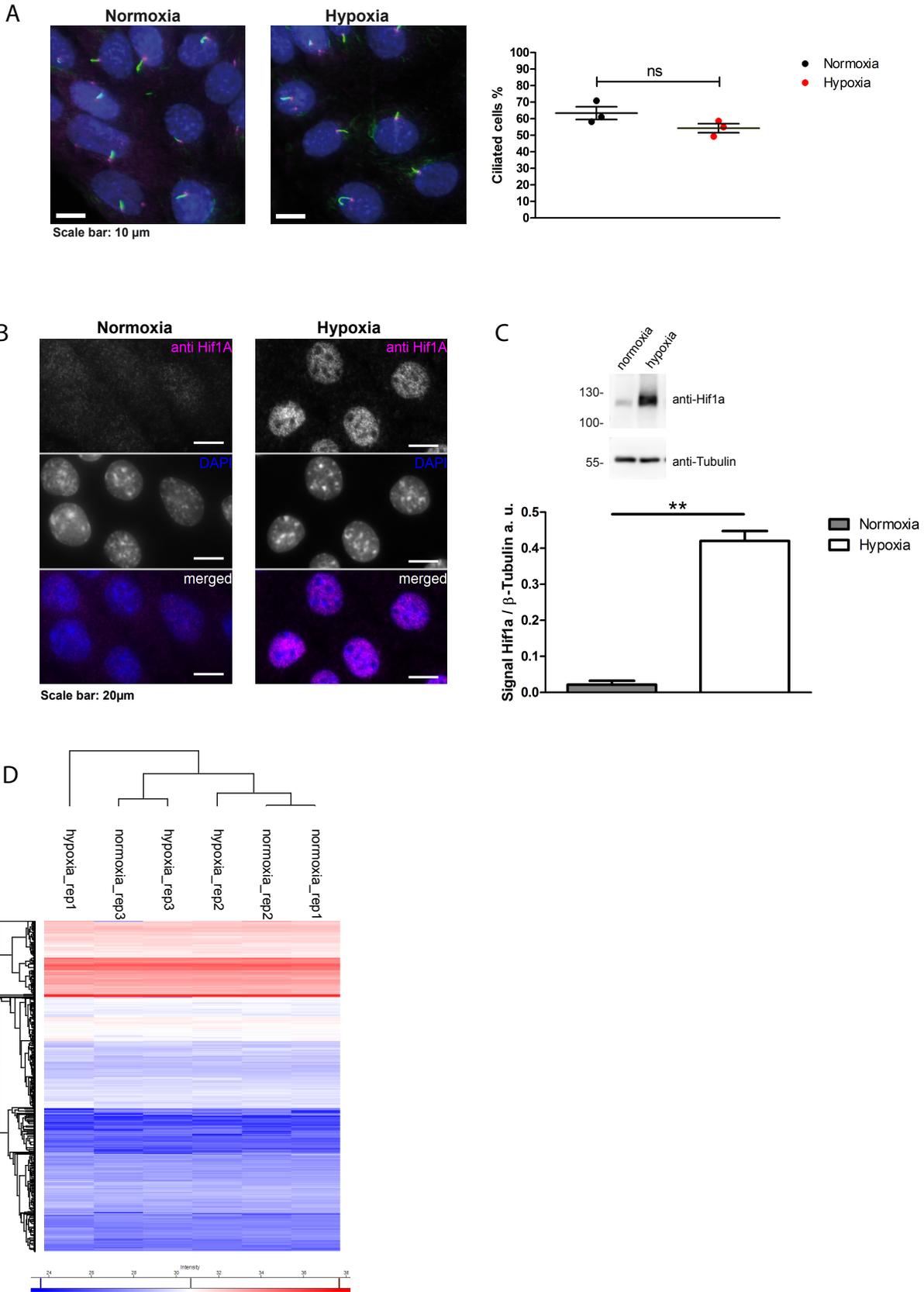
■ Expressed
■ not expressed (or listed)
■ data not available in proteomicsDB

B



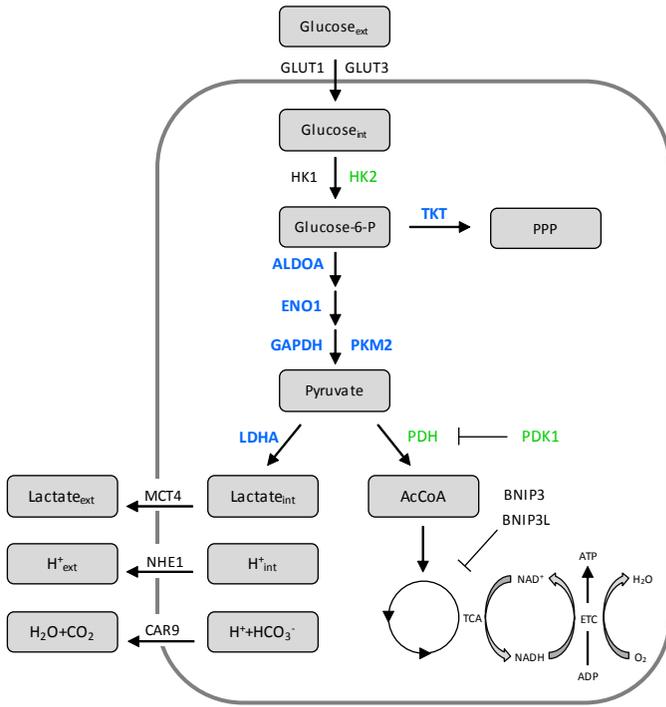
scale bar = 20 μm

Supplementary Figure 4

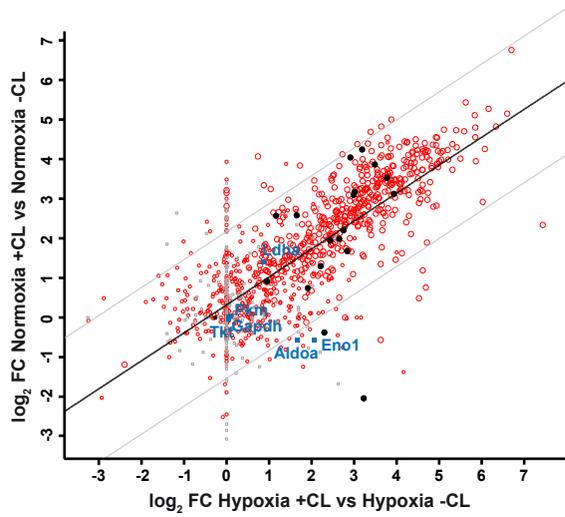


Supplementary Figure 5

A



B



C

