

# HINCUTs in cancer: hypoxia-induced noncoding ultraconserved transcripts

J Ferdin<sup>1,2,12</sup>, N Nishida<sup>1,3,12</sup>, X Wu<sup>4,12</sup>, MS Nicoloso<sup>1</sup>, MY Shah<sup>1</sup>, C Devlin<sup>4</sup>, H Ling<sup>1</sup>, M Shimizu<sup>1</sup>, K Kumar<sup>4</sup>, MA Cortez<sup>1</sup>, M Ferracin<sup>5</sup>, Y Bi<sup>6</sup>, D Yang<sup>7</sup>, B Czerniak<sup>7</sup>, W Zhang<sup>7</sup>, TD Schmittgen<sup>8</sup>, MP Voorhoeve<sup>9,10</sup>, MJ Reginato<sup>11</sup>, M Negrini<sup>5</sup>, RV Davuluri<sup>6</sup>, T Kunej<sup>\*,2</sup>, M Ivan<sup>\*,4</sup> and GA Calin<sup>\*,1</sup>

Recent data have linked hypoxia, a classic feature of the tumor microenvironment, to the function of specific microRNAs (miRNAs); however, whether hypoxia affects other types of noncoding transcripts is currently unknown. Starting from a genome-wide expression profiling, we demonstrate for the first time a functional link between oxygen deprivation and the modulation of long noncoding transcripts from ultraconserved regions, termed transcribed-ultraconserved regions (T-UCRs). Interestingly, several hypoxia-upregulated T-UCRs, henceforth named 'hypoxia-induced noncoding ultraconserved transcripts' (HINCUTs), are also overexpressed in clinical samples from colon cancer patients. We show that these T-UCRs are predominantly nuclear and that the hypoxia-inducible factor (HIF) is at least partly responsible for the induction of several members of this group. One specific HINCUT, uc.475 (or HINCUT-1) is part of a retained intron of the host protein-coding gene, O-linked N-acetylglucosamine transferase, which is overexpressed in epithelial cancer types. Consistent with the hypothesis that T-UCRs have important function in tumor formation, HINCUT-1 supports cell proliferation specifically under hypoxic conditions and may be critical for optimal O-GlcNAcylation of proteins when oxygen tension is limiting. Our data gives a first glimpse of a novel functional hypoxic network comprising protein-coding transcripts and noncoding RNAs (ncRNAs) from the T-UCRs category. *Cell Death and Differentiation* (2013) 20, 1675–1687; doi:10.1038/cdd.2013.119; published online 13 September 2013

Hypoxia is one of the classic features of the neoplastic microenvironment and is causally associated with resistance to conventional anticancer therapy.<sup>1</sup> Under hypoxic conditions, cells undergo a complex adaptive process involving genes regulating angiogenesis, cell invasion, metabolism and survival pathways. Novel insights into the pathways regulated by low oxygen concentration have shown that in addition to protein-encoding genes, the noncoding transcriptome is also affected by hypoxia. We and others previously reported that hypoxia regulates specific microRNAs (miRNAs), which in turn regulate signals of relevance for tumor biology,<sup>2–5</sup> including angiogenesis, cell cycle, DNA repair and energy metabolism.<sup>5,6</sup>

During the past few years, it has become clear that long noncoding RNAs (ncRNAs) (>200 bp in length) also participate in complex genetic events in eukaryotes.<sup>7</sup> One intriguing family of long ncRNAs<sup>8</sup> is transcribed from regions that exhibit extremely high conservation between the orthologous regions of human, rat and mouse genomes,<sup>9,10</sup> hence the term 'transcribed-ultraconserved regions' (T-UCRs) or ultraconserved genes.<sup>8</sup> Their striking evolutionary retention strongly suggests profound biological roles in a wide variety of physiologic responses. Recent studies have identified alterations in T-UCR-expression patterns that are associated with specific tumor phenotypes, pointing towards a mechanistic involvement of T-UCRs in cancer development. In particular,

<sup>1</sup>Department of Experimental Therapeutics, The Center for RNA Interference and Non-Coding RNAs, MD Anderson Cancer Center, The University of Texas, So Campus Research Building 3, 1881 East Road, Houston, TX 77030, USA; <sup>2</sup>Chair of Genetics, Animal Biotechnology and Immunology, Department of Animal Science, Biotechnical Faculty, University of Ljubljana, Groblje 3 1230, Domzale, Slovenia; <sup>3</sup>Department of Gastroenterological Surgery, Osaka University Graduate School of Medicine, 2-2 Yamada-oka, Suita City, Osaka, Japan; <sup>4</sup>Department of Medicine, Indiana University School of Medicine, 980W Walnut Street, Indianapolis, IN 46202, USA; <sup>5</sup>Department of Morphology, Surgery and Experimental Medicine, Laboratory for Technologies of Advanced Therapies (LTTA), University of Ferrara, Via Fossato di Mortara, 70, Ferrara, Italy; <sup>6</sup>Center for Systems and Computational Biology, Molecular and Cellular Oncogenesis Program, The Wistar Institute, 3601 Spruce Street, Philadelphia, PA, USA; <sup>7</sup>Department of Pathology, MD Anderson Cancer Center, The University of Texas 1515, Holcombe Boulevard, Houston, TX, USA; <sup>8</sup>College of Pharmacy, Ohio State University, 500 West 12th Avenue, Columbus, OH, USA; <sup>9</sup>Cancer and Stem Cell Biology Program, Duke-NUS Graduate Medical School, 8 College Road, Singapore; <sup>10</sup>Department of Biochemistry, Yong Loo Lin School of Medicine, 8 Medical Drive, National University of Singapore, Singapore and <sup>11</sup>Department of Biochemistry and Molecular Biology, Drexel University College of Medicine, 245 North 15th Street, Philadelphia, PA, USA

\*Corresponding authors: T Kunej, Chair of Genetics, Animal Biotechnology and Immunology, Department of Animal Science, Biotechnical Faculty, University of Ljubljana, Groblje 3 1230, Domzale, Slovenia. Tel: +386 1320 3890; Fax: +386 17241 005; E-mail: tanja.kunej@bf.uni-lj.si or M Ivan, Department of Medicine, Indiana University School of Medicine, 980W Walnut Street, Indianapolis, IN 46202, USA. Tel: +317 274 0322; Fax: +317 274 0396. E-mail: mivan@iu.edu

or GA Calin, Department of Experimental Therapeutics, The Center for RNA Interference and Non-Coding RNAs, MD Anderson Cancer Center, The University of Texas, So Campus Research Building 3, 1881 East Road, Houston, TX 77030, USA. Tel: +713 792 5461; Fax: +713 745 4528; E-mail: gcalin@mdanderson.org

<sup>12</sup>These authors contributed equally to this work.

**Keywords:** ultraconserved genes; colorectal cancer; glioblastoma; hypoxia; OGT

**Abbreviations:** miRNA, microRNA; ncRNA, noncoding RNA; T-UCR, transcribed-ultraconserved gene; PCR, polymerase chain reaction; RT-qPCR, quantitative reverse transcription PCR; DMOG, dimethylloxalylglycine; DMSO, dimethyl sulfoxide; CA9, carbonic anhydrase IX; HIF, hypoxia-inducible factor; CRC, colorectal carcinoma; Cq, quantification cycle; OGT, O-linked N-acetylglucosamine (GlcNAc) transferase; HINCUTs, hypoxia-induced noncoding ultraconserved transcripts; VHL, Von Hippel–Lindau tumor suppressor; WT, wild type; FACS, fluorescence activated cell sorting

Received 27.2.13; revised 19.7.13; accepted 29.7.13; Edited by G Melino; published online 13.9.13

alterations in T-UCR expression have been described in adult chronic lymphocytic leukemia, colorectal cancer (CRC),<sup>8</sup> hepatocellular carcinomas<sup>8,11</sup> and neuroblastomas.<sup>12,13</sup> In the human genome, T-UCRs are often found to overlap with exons of genes involved in RNA splicing or are located within host gene introns or in close proximity to genes that are involved in transcription and development regulation.<sup>9,14</sup> Whereas ultraconserved sequences appear to act as tissue-specific regulators of gene expression during development,<sup>7,9,15</sup> very little, if anything, is currently understood about the mechanisms underlying the cancer-associated profile alterations of the T-UCR. To address this limitation, we hypothesized that decreased oxygenation contributes to the deregulation of specific T-UCRs in cancer, pursuing a similar rationale to our original work on hypoxia-regulated miRNAs.<sup>2</sup> We initiated the search by exposing several cancer cell lines to low oxygen followed by quantification of changes in T-UCR expression *versus* normoxic controls using a genome-wide microarray profiling. Subsequently, we interrogated the contribution of hypoxia-inducible factor (HIF) activity to the upregulation of a set of hypoxia-induced noncoding ultraconserved transcripts (HINCUTs). Finally, the leading candidate of this group was functionally dissected in detail.

## Results

**T-UCR-expression changes under hypoxic conditions in malignant cells.** We re-interrogated our previous microarray data (Array Expression accession: E-TABM-184) obtained from CRC samples (78 CRCs compared with 21 normal colonic mucosas)<sup>8</sup> to identify T-UCRs overexpressed in CRC that could be potentially induced by hypoxia (for a detailed workflow, see Figure 1a and Supplementary Figure 1). The same type of microarrays was used for the discovery of hypoxia-induced miRNAs.<sup>2</sup> We identified 61 T-UCRs that had a significantly ( $P$ -value  $< 0.05$  Welch  $t$ -test with Benjamini–Hochberg correction) altered expression in CRC when compared with normal colon tissue (59 overexpressed and two with lower expression). Twenty-seven T-UCRs were upregulated compared with normal colonic mucosa at a more stringent  $P$ -value  $< 0.005$ , and further analyzed (Supplementary Table 1).

In order to validate the microarray data, we exposed two cancer cell lines (HCT-116 CRC and MCF-7 breast cancer cells) to 1% oxygen or 500  $\mu$ M DMOG (dimethylxalylglycine) – a widely used hypoxia mimetic – for 24 and 48 h and measured T-UCR expressions using quantitative reverse transcriptase PCR (RT-qPCR). Comparisons were performed with normoxic controls (21% oxygen) or dimethyl sulfoxide (0.5% DMSO). We identified a set of T-UCRs (uc.63, uc.73, uc.106, uc.134 and uc.475) that was induced more than two-fold ( $P$ -value  $< 0.05$  by Student's  $t$ -test) after 48 h both of hypoxia and DMOG exposures (Figure 1b and Table 1). All of these transcripts were predominantly present in the nucleus, as identified by nuclear and cytoplasmic fractionation experiments (Supplementary Figure 2). Robust activation of hypoxia signaling was confirmed by HIF-1 $\alpha$  (or HIF-2 $\alpha$ ) protein overexpression (Figures 1b and c), carbonic anhydrase IX (CA9) (data not shown) and miR-210

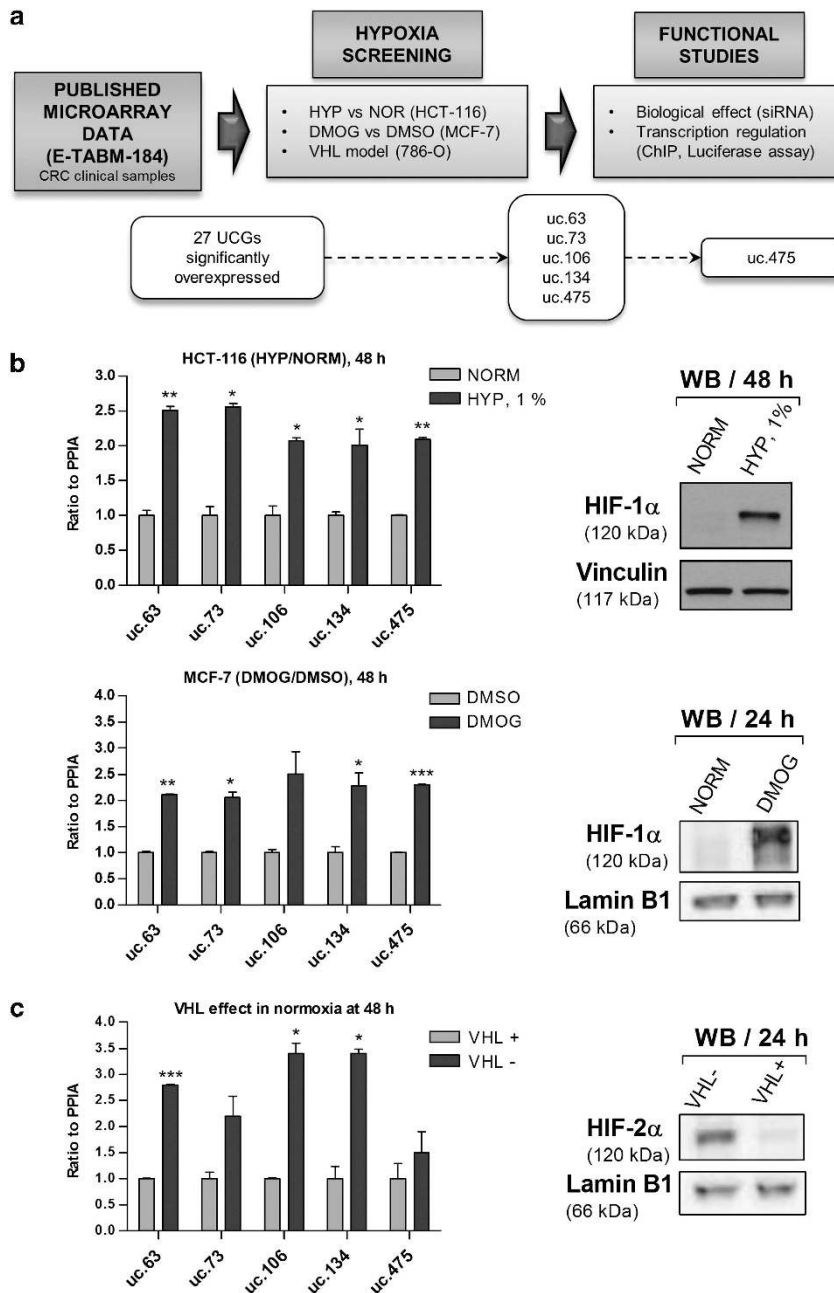
(Supplementary Figure 3) gene expression levels, which are both well-established coding and noncoding HIF targets, respectively.<sup>2</sup> Henceforth, these T-UCRs will be referred to as HINCUTs; their functional dissection under low-oxygen conditions is described below.

**Evidence for HIF as HINCUT's regulator.** The effect of DMOG on the expression of HINCUTs suggests a central role for HIFs as an upstream regulator. To further address the role of HIF, we used the renal cell carcinoma (RCC) cell line 786-O, which exhibits homozygous inactivation of the von Hippel–Lindau tumor suppressor gene (*VHL*), the well-established E3 ubiquitin ligase for all HIF family members. 786-O cells exhibit constitutive activation of the HIF pathway.<sup>16</sup> We hypothesized that HIF-2 $\alpha$ -induced T-UCR expression would have a similar pattern. To this end, when we compared normoxic 786-O cells expressing an empty vector with cells expressing wild-type *VHL* to restore the defects, and consistent with being *bona fide* HIF-2 $\alpha$  targets, all HINCUTs were upregulated at least 1.5-fold in the *VHL*-deficient cell line (*VHL*–) compared with the *VHL*-restored (*VHL*+ ) cells (Figure 1c). Furthermore, time course experiments in hypoxic models revealed that uc.475 was the most consistently induced (Supplementary Figure 4).

As no T-UCR promoter has been characterized to date, we performed an *in silico* search for HIF-binding sites within the 15-kb genomic region surrounding HINCUTs. For uc.63, uc.73, uc.134 and uc.475, we found multiple putative HIF-binding sites per UCR (Supplementary Table 2). Then T-UCRs with yet unknown functional role and more than one binding sites identified were also included in these studies. In the uc.63 promoter region, we identified the following three candidate HIF-binding sites: Q3\_uc.63A 5'-GGCAACGTG CATG-3' (63A.1), Q3\_uc.63A 5'-GAGTGCCTGCTCCT-3' (63A.2) and Q6\_uc.63A 5'-CGCGTGTGC-3' (63A.3), where as for uc.475 the following three candidate sites were found: Q3\_uc.475 5'-GCTCGCATGCGGG-3' (475.1) and two in close proximity with each other, Q5\_uc.475 5'-CACACGT GCGCG-3' and Q6\_uc.475 5'-CACGTGCGC-3' (475.2) (Figure 2a).

To further substantiate the role of HIF in the regulation of uc.475 and uc.63, we performed chromatin-immunoprecipitation assay (ChIP) of HIF-1 $\alpha$  in MCF-7 cells under normoxic conditions or following treatment with the hypoxia mimetic, DMOG. We confirmed direct recruitment of endogenous HIF-1 $\alpha$  at the promoters of uc.63 (data not shown) and uc.475, which was more evident during hypoxic conditions (Figure 2a; genomic regions from The University of California, Santa Cruz (UCSC) Genome Browser (<http://genome.ucsc.edu/>)<sup>17</sup> NCBI36/hg18). For each ChIP reaction, we used the promoter of *CA9*, a well-characterized HIF target, as the positive control and hemoglobin beta (*HBB*) as the negative control.

We further confirmed the transcriptional activation of uc.475 by hypoxia and HIF-1 $\alpha$  using promoter fragments subcloned into the pGL4-10 (*luc2*) basic expression vector. These constructs were co-expressed with the pRL-TK (Renilla) control vector and constitutively active HIF-1 $\alpha$  constructs (pcDNA 3.1-HIF-1 $\alpha$  P/A (proline/alanine mutant) and pcDNA 3.1-HIF-1 $\alpha$  3M (proline/alanine/asparagine mutant)<sup>18</sup> in two CRC cell lines, HCT-116 and HT-29 cells (Figures 2b and c).



**Figure 1** T-UCR-expression induction by hypoxia. (a) Experiment flowchart. (b) RT-qPCR confirmation of T-UCR-expression induction after 48 h under hypoxic conditions in HCT-116 colon cancer cells and hypoxia-simulation conditions by DMOG in the MCF-7 breast cancer cell line. Expression was normalized *versus* normal conditions (normoxia in HCT-116 and DMSO in MCF-7). Bars indicate the standard deviation from two sample duplicates. Western blot assay performed with HIF-1 $\alpha$ -specific antibodies. (c) T-UCR induction measured using RT-qPCR in the VHL-deficient (VHL -) cell line was compared with that in the VHL-proficient (VHL +) 786-O cell line under normoxic conditions. Western blot analysis was performed using HIF-2 $\alpha$ -specific antibodies. \* $P < 0.05$ ; \*\* $P < 0.01$ ; \*\*\* $P < 0.001$

We found that hypoxia and the active HIF-1 $\alpha$ s induced an increase (up to five-fold) in luciferase activity (Figure 2b), which was correlated with the uc.475 expression in these cells (Supplementary Figure 5). In contrast, luciferase activity of the mutant clones that lack HIF-1 $\alpha$ -binding sites (PT1 and PT2) was significantly reduced (Figure 2c) ( $P$ -value  $< 0.005$ ), suggesting that the two predicted binding sites are functional HIF-recognition sites that mediate hypoxia effects on uc.475 transcription.

**Biological effects of uc.475 downregulation.** To determine the biological significance of HINCUTs, we performed knockdown experiments for uc.63 and uc.475 using custom-designed siRNAs and found that these effected cell proliferation using the general strategy described initially for uc.73.<sup>8</sup> siRNA-specific oligonucleotides for HINCUTs were designed after evaluating by strand-specific RT-qPCR if the sense or antisense transcript was transcribed at higher levels (Supplementary Figures 5 and 6). No major biological effect

**Table 1** Genomic information for the five HINCUTs identified in the present study

T-UCR	Genomic location	Symbol of the host gene	Gene location	Host gene orientation <sup>a</sup>	Length (nt)	HCT-116 HYP/ NORM (48 h)	MCF-7 DMOG/ DMSO (48 h)
uc.475 +	Xq13.1	<i>OGT</i> (O-linked N-acetylglucosamine (GlcNAc) transferase)	Intron 2	S	397	2.02	2.30
uc.63 +/+ A	2p15	<i>XPO1</i> (Exportin 1)	Intron 3	AS	278	2.51	2.11
uc.73 + A	2q22.3	<i>GTDC1</i> (Glycosyltransferase-like domain containing 1)	Intron 6	AS	201	2.56	2.05
uc.106 +	2q31.1	<i>OLA1</i> (Obg-like ATPase 1)	Intron 4	AS	206	2.07	2.51
uc.134 + A	3q25.32	<i>RSRC1</i> (Arginine/serine-rich coiled-coil 1)	Intron 7	S	211	2.01	2.28

Abbreviations: DMOG, dimethylallylglycine; DMSO, dimethyl sulfoxide; HINCUT, Hypoxia-induced noncoding ultraconserved transcripts. Location data are according to the UCSC Genome Browser (assembly GRCh37/hg19). Host gene names are according to the HUGO nomenclature. AS or +A, antisense; S or +, sense; nt, nucleotide; we used the same name for the T-UCR as for the array probes found in E-TABM-184

<sup>a</sup>Host gene orientation is reported *versus* HINCUT orientation considered as reference

on the dynamics of COLO-205 culture was identified for uc.63 (Supplementary Figure 7) and henceforth, we focused our efforts on uc.475. Notably, the primary uc.475 transcript has the same orientation as its host gene, O-linked N-acetylglucosamine (GlcNAc) transferase (*OGT*) (Supplementary Figure 6). The effect of uc.475 downregulation on cell survival was tested in HT-29 cells under hypoxic and normoxic conditions. These cells were chosen based on their high endogenous uc.475 expression level (Supplementary Figure 5). We employed a custom-designed pool of siRNAs to knockdown uc.475; each of the three siRNAs tested to be efficient reduced the basal expression of uc.475 by more than 65% as confirmed by RT-qPCR (Supplementary Figures 7 and 8 and Supplementary Table 3).

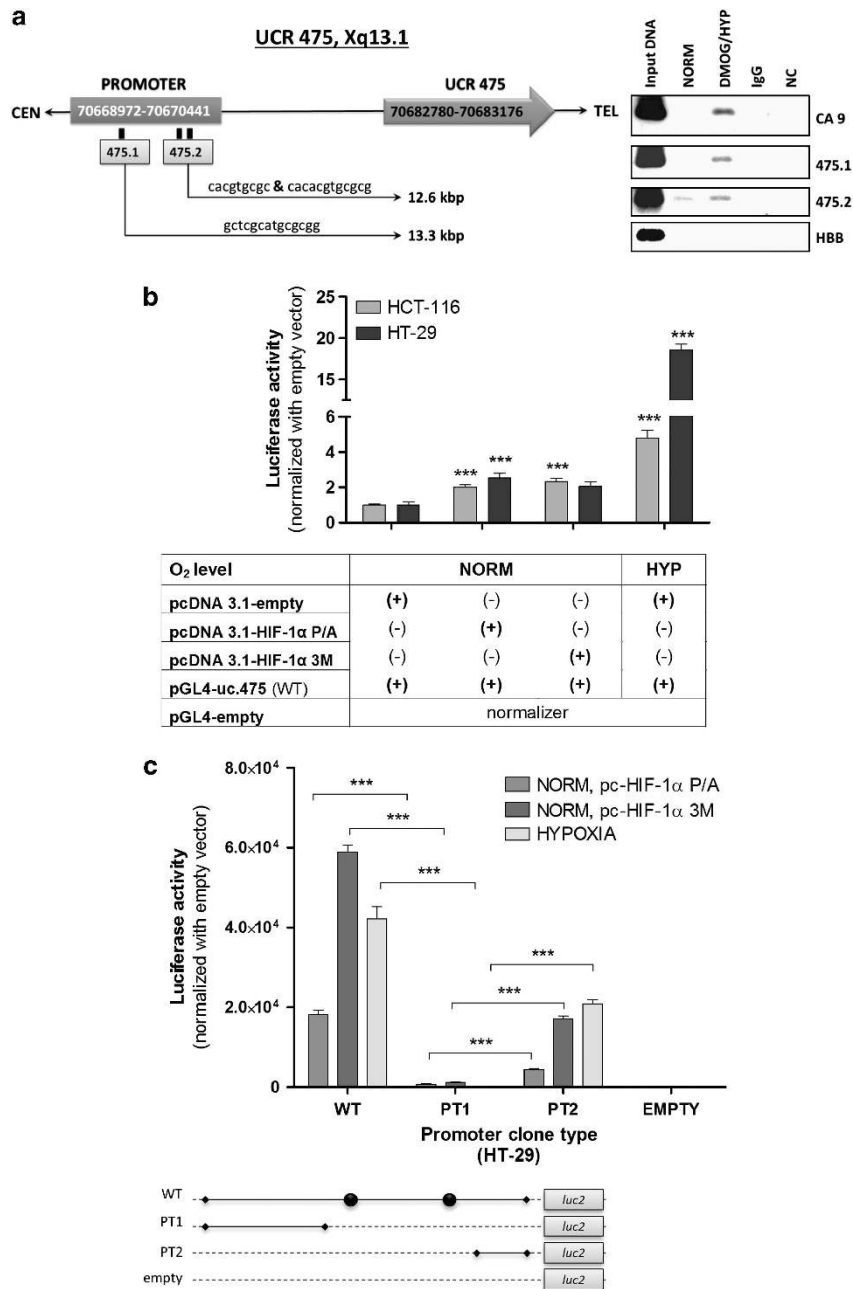
Under normoxic conditions, downregulation of uc.475 had no significant impact on the total cell number *versus* control (siRNA-NEG) cells for up to 120 h (Figure 3a, left). In contrast, under hypoxic conditions (1% O<sub>2</sub>), uc.475 knockdown significantly decreased cell proliferation (Figure 3a, right). This effect was not the result of increased apoptosis, as poly ADP-ribose polymerase, a critical component of programmed cell death, remained uncleaved (data not shown). Instead, through the analysis of the cell cycle, using fluorescence-activated cell sorting (FACS), we observed that uc.475 knockdown under hypoxic condition enriched the G2/M fraction more than two-fold at 72 and 96 h (from 10.3 and 8.2% with the negative control to 21.7 and 18.8% with siRNA, respectively), suggesting a G2/M blockade. No similar consistent variations in the percentage of cells in the various phases of the cell cycle under normoxic conditions were identified (Figure 3b and Supplementary Figure 9). Cell cycle distribution was also checked after treatment with siRNA against *OGT* (Supplementary Table 3 and Supplementary Figure 8) under normoxia and hypoxia, and only minimal effects on cell cycle progression were observed (Supplementary Figure 9b). To exclude the possibility that siRNA against uc.475 downregulates pre-mRNA of *OGT* and leads to a phenotypic effect, we designed another siRNA targeting another intron of *OGT*. We selected intron 12 to knockdown the pre-mRNA, as this siRNA exhibited the highest efficiency among several siRNAs we have designed. As a result, although *OGT* protein expression was moderately downregulated by *OGT*-intron 12 knockdown, more

significant reduction was observed with uc.475 knockdown (Figure 3c), and no proliferation inhibition or altered cell cycle distribution was observed after the knockdown of *OGT*-intron 12 (Supplementary Figure 11). These results indicate that the alteration of *OGT* expression and proliferation inhibition by uc.475 knockdown is a distinct effect of the *OGT* pre-mRNA knockdown.

**OGT locus: a mix of *OGT* and HINCUT-1 transcripts.** As mentioned previously, and in line with earlier reports,<sup>19–21</sup> the uc.475 sequence (396 bp in length) is found inside the fourth intron of the *OGT* gene (UCSC Genome Browser; GRCh/hg19, Figure 4a). *OGT* has two known isoforms (*OGT1* of 5497 bp and *OGT2* of 5467 bp) because of a variant splice site, which is known to encode proteins of 1046 or 1036 amino acids (116 and 103 kDa), respectively, and that catalyze polypeptide glycosylation by the addition of N-acetylglucosamine (O-GlcNAc) to serine and threonine residues (O-GlcNAcylation) on many nuclear and cytosolic proteins.<sup>19</sup> Recently, this modification was proposed to be a new link between glucose metabolism and cancer growth and invasion in breast,<sup>22</sup> colon and lung cancer types.<sup>23</sup> In addition, *OGT* was found to be highly expressed in prostate cancer compared with the normal epithelium.<sup>24</sup> Taking into account the link between *OGT* and cancer, we hypothesized that this, at least in part, could be explained by the associated uc.475 transcript with a potential role as the regulator of its host gene expression during hypoxia.

By using a multiple-primer amplifications approach, we cloned a transcript of ~2.7 kb that included the uc.475 region but excluded the flanking *OGT* exons (for details see Supplementary Figure 12). In contrast, amplification levels were much lower when primers in the 'flanking' *OGT* exons (Exons 4 and 5) and intron region containing uc.475 were used in combination; the *OGT* and uc.475 transcripts were 2–10 times more abundant than the amplicons obtained with the mixed combinations of primers between flanking sequences of uc.475 and exons of *OGT* (Figure 4b). These data support two possibilities: (a) the existence of an *OGT*-independent transcript that contains the fourth intron or (b) the existence of a primary transcript that contains both *OGT* and uc.475 but is further spliced into two transcripts, one noncoding (the intronic uc.475) and one coding (the *OGT* isoforms).

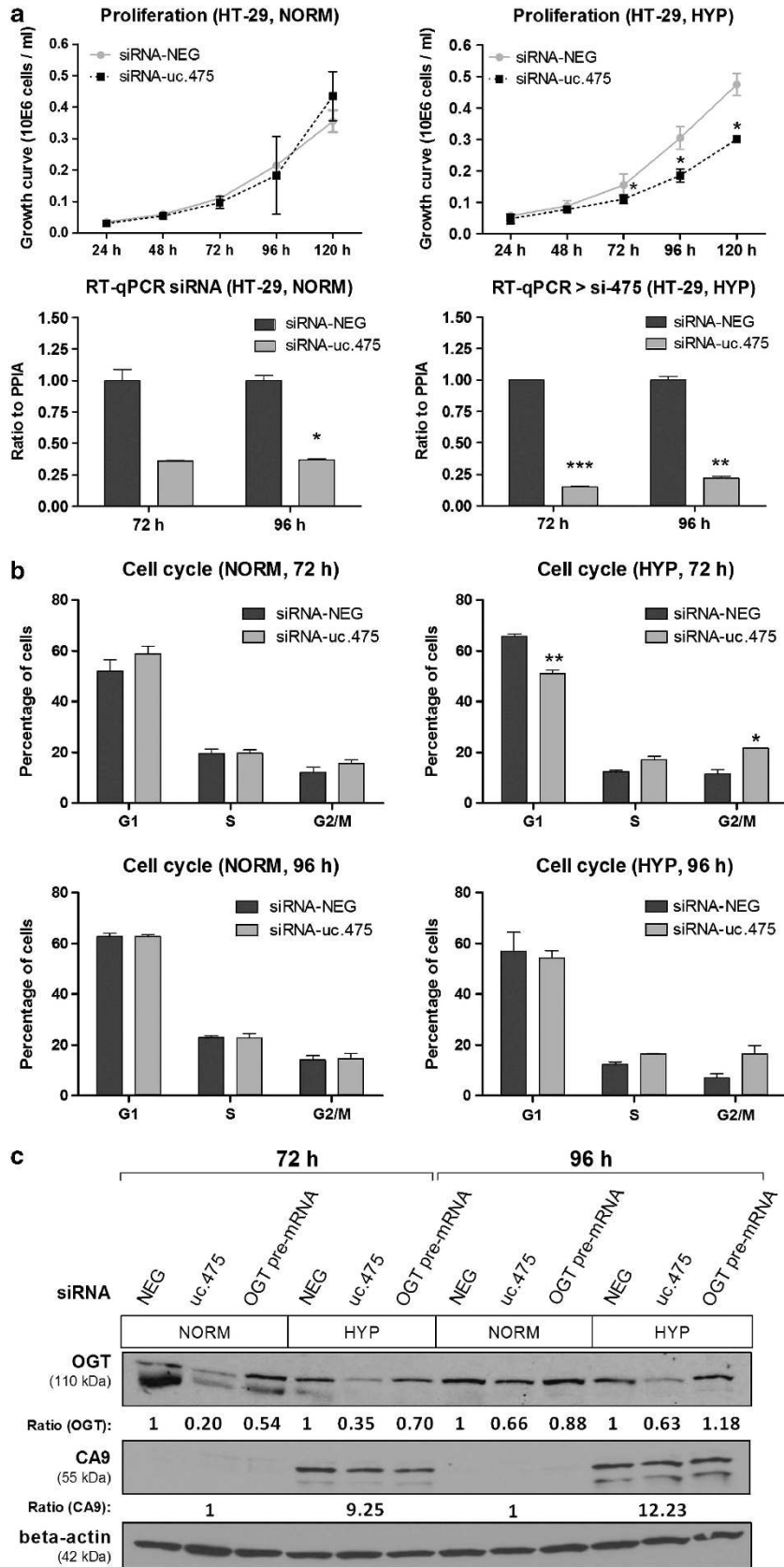




**Figure 2** Confirmation of a direct association between HIF and T-UCR using ChIP and luciferase assay. (a) Direct recruitment of HIF on T-UCR-expression induction in hypoxia-simulated conditions by DMOG in the MCF-7 breast cancer cell line. We confirmed a direct connection between HIF and T-UCR (*uc.475*) using ChIP method with HIF-1 $\alpha$  antibodies or an IgG control. To enrich genomic fragments that contain hypoxia-response elements (HIF-interacting sites), we amplified the regions represented on the genome graphic using specific primers for pointed locations. The ChIP figures are inverted for a better view of expression induction. NC, negative control; HBB, hemoglobin globin beta, which is supposed to be transcriptionally inactive; and carbonic anhydrase IX (CA9), which was used as a positive control in the ChIP experiment. (b) The direct effect of HIF on the *uc.475* promoter reporter construct was confirmed by using luciferase assay and was shown as an effect of exogenous pcDNA 3.1-HIF-1 $\alpha$  vector constructs (HIF-1 $\alpha$  P/A and HIF-1 $\alpha$  3M) on cloned promoter activity. A stronger effect was detected by using luciferase assay normalized using empty pcDNA 3.1 vector. The error bars represent the standard deviation from two independent experiments, each performed in triplicate and measured in duplicate with a luciferase assay. (c) Luciferase activity in HCT-116 and HT-29 cells was measured after transfection with truncated promoter reporter constructs (PT1 and PT2) with a complete deletion of the HIF-1 $\alpha$ -binding sites that were present in the originally designed *uc.475* promoter reporter construct (WT). Filled circles show the position of the HIF-1 $\alpha$ -predicted binding site. \*\*\**P*-value < 0.001 compared with luciferase activity of an empty vector; WT, wild type; PT1, truncated promoter 1; PT2, truncated promoter 2

To further substantiate the tight link between *OGT* and *uc.475*, we performed experiments under stress conditions for which *OGT* function is relevant.<sup>25</sup> As *OGT*-mediated O-GlcNAcylation is proportional to substrate (glucose)

availability,<sup>26</sup> we performed experiments under glucose- as well as oxygen-deprivation conditions (OGD experiments) – two tightly associated characteristics of the tumor micro-environment. In HT-29 cells, the expression levels of *uc.475*



and *OGT* mRNA significantly increased during 12 and 24 h of normoxia in glucose-free media. After 48-h exposure under these conditions, the expression of both decreased dramatically to the same or lower level observed under hypoxia in the presence or absence of glucose (Figure 4c). These data show that the responses of uc.475 and *OGT* mRNA expression to oxygen and glucose deprivation are similar.

In parallel, we attempted to visualize the sizes of the transcripts from this locus by using Northern blot analysis. This would allow us to determine whether uc.475 exists as an independent transcript, or only part of a hybrid transcript with *OGT*, or as both. We used RNA isolated from MCF-7 cells after exposure to hypoxia or normoxia for 24 h. Using a probe specific to uc.475, we identified two transcripts (~8 and ~9.5 kb) with higher expression under hypoxia in MCF-7 cells, whereas the *OGT*-coding region transcripts (4.2 kb long) had barely detectable signal (Figure 5, left). Hybridization with an *OGT*-specific probe detected two transcripts (1.7 and 4.2 kb) with similar intensity under both normoxia and hypoxia conditions (Figure 5, right). Therefore, we believe that uc.475 is the part of a larger *OGT* primary transcript, which retains intron 4 and which is expressed in a hypoxia-dependent manner.

When we screened for potential open reading frames (ORF) in the 'retained intron', we detected no ORF > 100 bp. The 9-kb transcript was not predicted to encode for either of the two *OGT* isoforms, as the predicted ORF corresponds to 665 amino acids, potentially representing the smaller 78-kDa version of the *OGT* reported recently<sup>19,20,27</sup> (Supplementary Figure 13). In addition, we have searched for miRNAs and other putative ncRNAs in the uc.475 region using the UCSC database, in which current integrated RNA deep-sequencing experiment data (RNA-seq and miRNA-seq) are available, and found no annotated miRNAs or other ncRNAs to be located within the uc.475 region.

**Uc.475 inhibition has an impact on *OGT* expression and function.** We first investigated whether *OGT* expression at RNA and protein levels was affected by uc.475 knockdown in HT-29 cells. Expression of *OGT* mRNA levels was decreased by knocking down uc.475 (Figure 6a), as well as *OGT* protein levels, which was more evident in hypoxia (Supplementary Figure 10). Silencing *OGT* with siRNAs (siRNA-*OGT*) resulted in decreased levels of uc.475 and *OGT* mRNAs (Figure 6b) and *OGT* protein expression (Figure 6c). Interestingly, and contrary to the mRNA, the *OGT* protein level does not increase in hypoxia, as evidenced by measurements performed in three different cell lines – HCT-116 colon cancer, MCF-7 breast cancer and U-87 glioblastoma cells (Supplementary Figure 14).

To provide functional support for the impact of uc.475 downregulation on *OGT* function, we quantitated the O-GlcNAcylation rate following siRNA-mediated knockdown

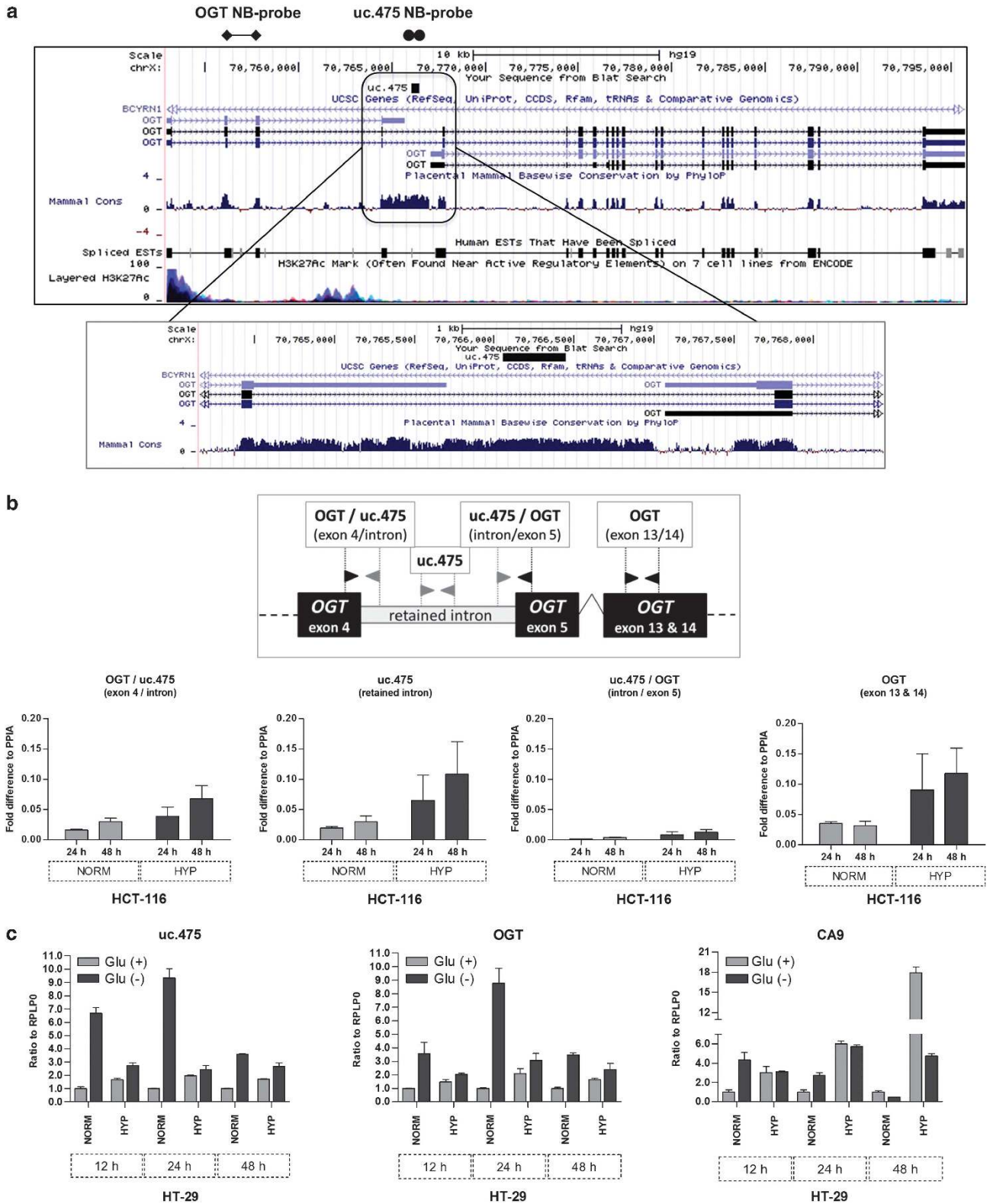
of the transcript containing the ultraconserved sequence (Supplementary Figure 15). This was performed in HT-29 cells treated with PUGNAC (1,5-hydroxymethylolactone), an inhibitor of the N-acetylhexosaminidases (*OGA* or *MGEA5*), in order to prevent the removal of O-GlcNAc from proteins.<sup>28</sup> Downregulation of uc.475 led to a significant decrease in the global O-GlcNAcylation rate, an effect similar to treatment with *OGT*-siRNA (Figure 6d), whereas the inhibition of *OGT* protein and global O-GlcNAcylation was relatively modest after treatment with siRNA against *OGT*-intron 12, which targets *OGT* pre-mRNA.

**The enhancer-like function of the DNA region containing uc.475 ncRNA.** It has been reported that ultraconserved regions can act as developmental enhancers, so we determined whether, in addition to having a transcriptional activity, the uc.475 could also have an enhancer function.<sup>29</sup> We used sequence-transcription stimulation by an adjacent heterologous promoter region as previously described.<sup>30</sup> The 2.7-kb long region (–1 kb upstream and +1.3 kb downstream of uc.475) (Supplementary Figure 12) was subcloned into the pGL3-TK vector (Figure 7a). This cloned insert revealed 2–3 fold luciferase activity enhancement, under both normoxia and hypoxia, in U-87 cells (Figure 7b). These findings demonstrate that the uc.475 DNA region also has an enhancer activity both under hypoxic and normoxic conditions.

## Discussion

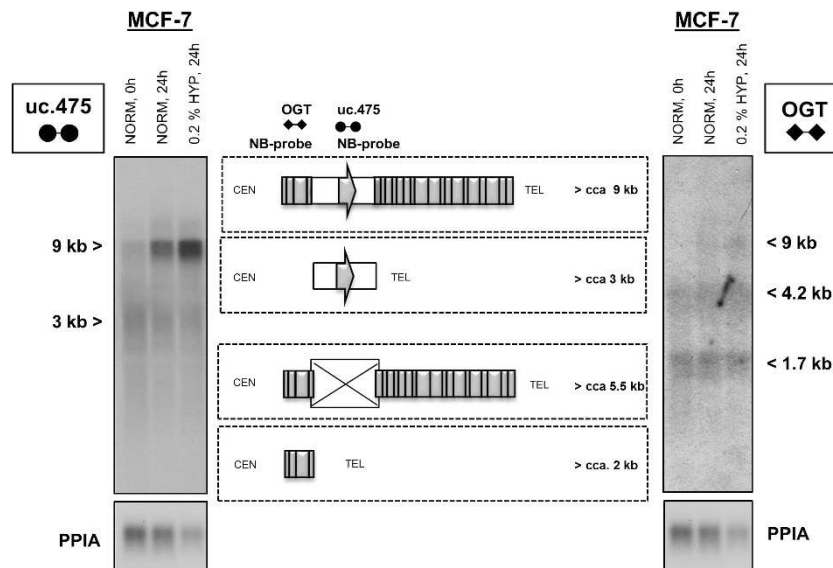
We aimed to investigate whether hypoxia induces the expression of noncoding ultraconserved transcripts in a similar manner to the conventional hypoxia-regulated genes or more recently studied miRNAs. To this end, we employed a strategy that integrates tumor data with a microarray-based screen.<sup>2</sup> We identified in multiple cell types the following hypoxia-induced, noncoding ultraconserved transcripts: uc.63, uc.73, uc.106, uc.134, and uc.475, and termed these HINCUTs. Similar to the earlier work on miRNAs,<sup>2</sup> we are providing evidence that HIF has a critical role for their hypoxic regulation, based on induction with chemical inactivators of the HIF prolyl hydroxylases, effect of the VHL tumor-suppressor inactivation and effect of exogenous HIF-1 $\alpha$ s on luciferase reporters containing fragments from the candidate promoters of HINCUTs. We identified HIF-1 $\alpha$ -binding sites upstream of uc.475. This region in fact overlaps the *OGT* gene promoter region; therefore, the transcriptional control of uc.475 is not considered to be completely distinct from *OGT* transcription. Our *in silico* prediction detects no putative binding site for HIF-1 $\alpha$  very close to *OGT* intron 4, where uc.475 is located. Recent reports have demonstrated that many intronic miRNAs possess their own promoter, which is independent of the host gene transcription.<sup>31</sup>

**Figure 3** Uc.475 biological significance under normoxia and hypoxia. (a) The effect of siRNA-targeting uc.475 on cell proliferation was determined under normoxic (left) and hypoxic (right) conditions by using a pool of three siRNAs against uc.475 (siRNA-uc.475). Three independent experiments were performed, and as control scrambled siRNA sequence, siRNA-NEG, was used. (b) FACS analysis was performed 72 and 96 h post transfection with either siRNA-uc.475 or siRNA-NEG in normoxia (left) and hypoxia (right). (c) Western blot analysis to detect differences of uc.475 host gene (*OGT*) protein expression levels. Hypoxic conditions were confirmed by CA9 induction. Protein-specific bands were normalized to beta-actin. siRNA against *OGT* pre-mRNA was designed to target *OGT*-intron 12. \**P*-value < 0.05, \*\**P*-value < 0.01, \*\*\**P*-value < 0.001



**Figure 4** Transcripts produced at the uc.475 genomic locus in an environment deprived of oxygen and/or glucose. (a) The genomic region of *OGT* from the UCSC Genomic Browser, GRCh37/hg19 with the known transcripts of *OGT* isoforms 1 and 2. The rounded rectangle represents region containing uc.475. Northern blot-probe locations are indicated. The intron 4 region where uc.475 is located is zoomed in for better view of region of conservation in mammals. (b) RT-qPCR data showing that the uc.475 noncoding transcript and the coding *OGT* transcript are expressed at similar levels, whereas the 'hybrid' amplicons between the intron and the two adjacent exons are barely expressed. (c) Glucose- and oxygen-deprivation effects on mRNA expression level of uc.475 and *OGT*. *CA9* mRNA expression was analyzed as a hypoxia-induced gene





**Figure 5** Uc.475-transcript induction by hypoxia. Northern blot analysis of MCF-7 RNA identified transcripts of different sizes detected with uc.475 (left) or *OGT* probes (right). Transcript variants identified by Northern blot signals in MCF-7 cells under normoxia and hypoxia are marked by arrows and visualized by a drawing (middle). The panel PPIA was used as a loading control

Further investigation is required to reveal whether there is an additional transcriptional factor-binding site closer to the uc.475 region.

Our data suggest that uc.475 is part of a primary (potentially retained intron) transcript that also includes the coding sequences of its 'host gene' *OGT*. The role of *OGT* is to catalyze the addition of O-GlcNAc on proteins, which are involved in transcription, cellular signaling pathways and processes, and the cell cycle.<sup>27,32</sup> *OGT* was also recently reported to be overexpressed in breast cancer and primary and metastatic prostate cancer,<sup>22,24</sup> further supporting *OGT* as a therapeutic target in solid malignancies.

Our results may expand the understanding of the complexity of the *OGT* genomic locus, as it reveals the presence of a noncoding ultraconserved sequence, part of a retained intron, at the core of a hybrid nuclear transcript. We provide evidence that the *OGT* locus generates multiple coding and noncoding transcripts. The retained intron that contains uc.475 does not possess a coding ORF but does include a second start codon (ATG) or it could be part of a long 5'-untranslated region (3.27 kb).<sup>21</sup>

We also noticed a contrast between the hypoxia-inducible nature of *OGT*/uc.475 at RNA and protein levels. Indeed, measurements in multiple cell lines indicate that *OGT* is either unchanged (Supplementary Figure 14) or even decreased during oxygen deprivation (data not shown). This is likely part of the wide translational suppression consistently reported under hypoxia. Higher transcript levels may serve to compensate this decrease, and it is entirely conceivable that in the absence of mRNA induction the level of O-GlcNAcylation would be compromised below the level compatible with survival.

What is therefore the role of uc.475-containing noncoding region in the context of *OGT*-hypoxic regulation? The luciferase experiments (Figure 7) suggest that on one hand this sequence may serve as an enhancer, which could

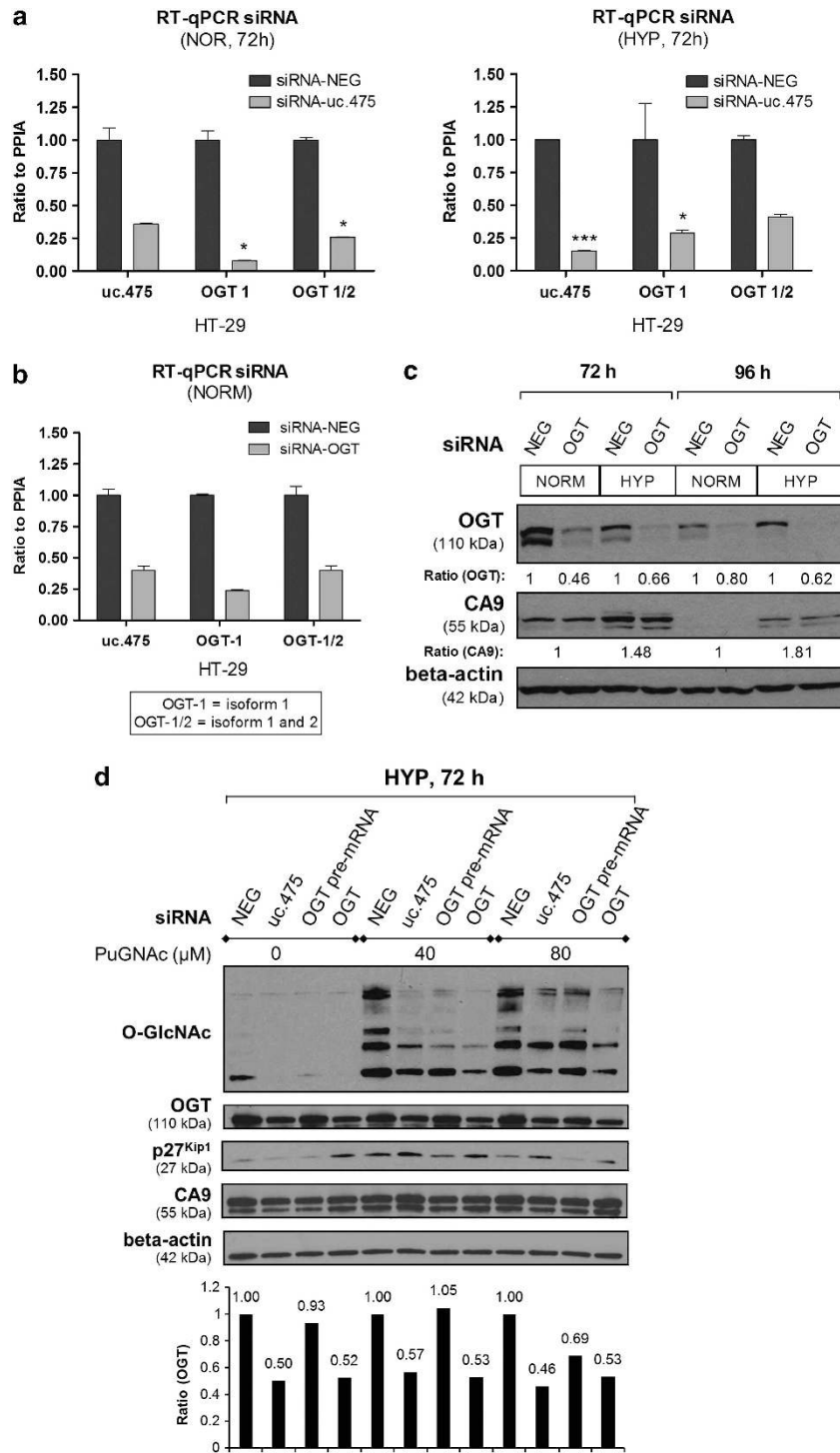
cooperate with HIF for elevated transcription. The predominantly nuclear localization of the hypoxia-inducible hybrid transcript may point also to a role in RNA stability. To date, retained introns within several genes were found to enhance almost every step of gene expression and also affects mRNA stability. However, some studies showed that intron retention has little or no effects on mRNA half-life,<sup>33–35</sup> whereas others confirmed that it is important for mRNA stability and has an effect on prolonged mRNA half-life.<sup>36–38</sup> In addition, cells under hypoxia were even found to inhibit nonsense-mediated RNA decay (NMD) mechanisms to enable the expression of genes critical for cell adaptation and response to a cellular stress.<sup>39</sup> Moreover, many studies found that constitutively spliced introns are required for optimal gene expression,<sup>33,40,41</sup> and one possible way through which introns can affect transcription is by acting as repositories for transcriptional regulatory elements such as enhancers and repressors.<sup>41</sup>

Although further mechanistic studies as well as an understanding of the specific roles of the many transcripts at the *OGT* locus are needed, this is the first study that purposely investigates T-UCRs as a part of the hypoxic response and extends the knowledge of the interplay between coding and noncoding transcripts under stress conditions.

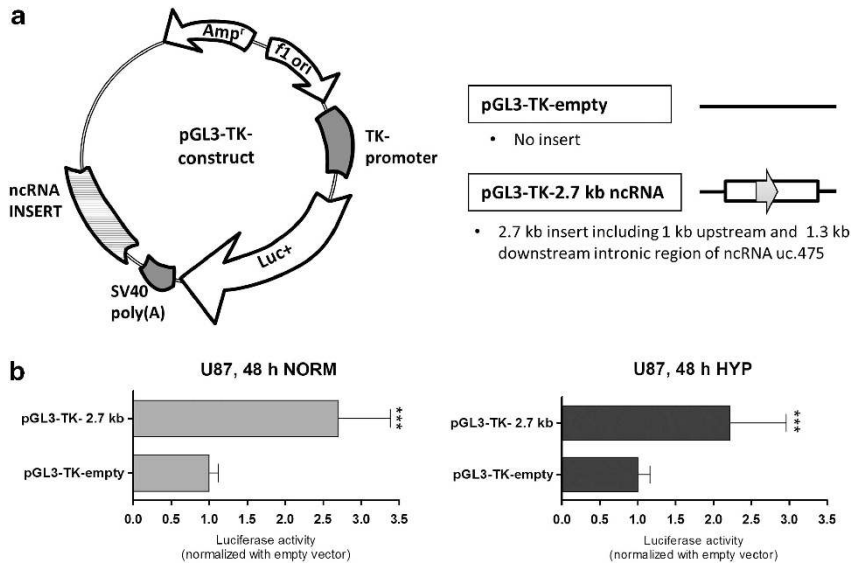
## Materials and Methods

The complete version is included in the Supplementary Material.

**Cell cultures and growth conditions.** Colon cancer (HCT-116, HT-29 and COLO-205), breast cancer (MCF-7), bladder cancer (UC-2) and glioblastoma (U-87) cell lines were obtained from ATCC (Manassas, VA, USA). RCC cell lines<sup>42</sup> were obtained from Dr. Maria Czyzyk-Krzeska (University of Cincinnati, Cincinnati, OH, USA). HCT-116 and HT-29 were maintained in McCoy's 5A (enriched with L-glutamine and sodium bicarbonate; Thermo Scientific, Logan, UT, USA); UC-2 was maintained in minimum essential Eagle medium (MEM; Cellgro Mediatech, Manassas, VA, USA); MCF-7 was cultured in Dulbecco's modified Eagle's



**Figure 6** Impacts on OGT expression and function after uc.475 inhibition and O-GlcNAcylation change after knocking down uc.475 or OGT. (a) RT-qPCR expression levels of *OGT* and uc.475 after uc.475 knockdown in HT-29 cells under normoxia and hypoxia. Primers were designed to detect both OGT1 and OGT2 isoforms. Error bars indicate standard deviation. UCR or uc, ultraconserved gene; OGT, O-linked N-acetylglucosamine (GlcNAc) transferase; siRNA, small interfering RNA; HT-29, colon cancer cell line; NOR, normoxia; HYP, hypoxia; PPIA, reference gene cyclophilin A; \**P*-value < 0.05, \*\*\**P*-value < 0.001. (b) RT-qPCR confirmation of *OGT* gene knockdown using siRNA against *OGT*. Primers specific for uc.475 and each of the *OGT* isoforms were used. (c) Western blot performed to detect differences of uc.475 host gene (*OGT*) protein expression levels after *OGT* silencing. Hypoxic conditions were confirmed by CA9 induction. Protein-specific bands were normalized to beta-actin as a reference protein. (d) O-GlcNAc protein levels were detected in samples exposed to hypoxic conditions for 72 h or simultaneously transfected with siRNA-targeting either uc.475, *OGT* pre-mRNA or *OGT*. All levels were compared with sample transfected with siRNA-NEG. Different concentrations of PUGNAc were added in cell media (40 and 80  $\mu$ M) to prevent O-GlcNAc from being removed from the proteins. siRNA against *OGT* pre-mRNA was designed to target *OGT*-intron 12



**Figure 7** Testing the enhancer activity for cloned uc.475-containing region. Uc.475 enhancer activity experiments were performed in normoxia and hypoxia for 48 h. **(a)** Insert representing a transcript of 2.7-kb region (including uc.475 region) was cloned downstream of the luciferase reporter. **(b)** Dual-Luciferase reporter assays were performed on U-87 cell line samples exposed to normoxia or hypoxia for 48 h. x axis represents the relative Firefly to Renilla luciferase activity, whereas y axis presents the plasmids. All measurements generated were normalized to the pGL3-TK empty vector. All data shown are the mean of two biological repetitions (at least eight luciferase measurements per experiment). \*\*\**P*-value < 0.001

high-glucose medium (DMEM; Cellgro Mediatech); U-87 was cultured in special DMEM supplemented with F12 50/50 mix (Cellgro Mediatech); COLO-205 was maintained in RPMI 1640 (Cellgro Mediatech); and 786-O VEC (VHL <sup>-</sup>) and 786-O WT (VHL <sup>+</sup>)<sup>42</sup> were cultured in DMEM:Ham's F12 supplemented with G418 (200  $\mu$ g/ml). All cell lines used were *Mycoplasma*-free, which was confirmed by using the MycoAlert mycoplasma detection kit (Lonza, Rockland, ME, USA).

Before being used, cell lines were validated by STR DNA fingerprinting using the AmpF/STR Identifier kit, according to the manufacturer's instructions (Applied Biosystems, cat. no. 4322288). The STR profiles were compared with known ATCC fingerprints (ATCC.org), the Cell Line Integrated Molecular Authentication database version 0.1.200808 (<http://bioinformatics.istge.it/clima/>)<sup>43</sup> and The University of Texas MD Anderson Cancer Center fingerprint database. The STR profiles of all the cell lines matched the known DNA fingerprints.

Clinical samples information is detailed in Supplementary Material.

**Hypoxia experiments.** Hypoxic experiments were performed in an *InVivo*<sub>2</sub> 200 hypoxia workstation (Ruskinn, Inc., Cincinnati, OH, USA), where oxygen concentration was maintained at 1 or 0.2% (in 5% CO<sub>2</sub> and 95% N<sub>2</sub>) for 24, 48, 72 and 96 h. RNA and proteins were collected for further analysis using RT-qPCR and western blot, respectively.

Chemical induction of HIF-1 $\alpha$  expression under normoxic conditions was performed using DMOG (Frontier Scientific, Logan, UT, USA). DMOG was added to the growth medium to a final concentration of 500  $\mu$ M versus of 0.5% DMSO vehicle as control.

RNA extractions, RT-qPCR, real-time PCR controls and reference genes are detailed in Supplementary Material.

**Data analysis for RT-qPCR.** Data analysis was performed using the Bio-Rad CFX Manager Software (Bio-Rad, Hercules, CA, USA). Only samples with a single clear peak in melting temperature and no expression in nontargeted control cells were considered as positive and specific to be further analyzed. All RT-qPCR data of expressed genes were represented as *C<sub>q</sub>* values for each sample of duplicates. Relative expression, as the fold change or difference,<sup>44</sup> was calculated using the  $2^{(-\Delta\Delta C_q)}$  formula, in which  $\Delta C_q$  represents the difference in *C<sub>q</sub>* values between the target gene and the mean *C<sub>q</sub>* of the reference gene.<sup>45</sup> Each sample, amplified using RT-qPCR, was tested in duplicate. Standard deviation (S.D.) of each target gene was calculated.

Western blot analysis, siRNA transfection and FACS are detailed in Supplementary Material.

**HIF-1 $\alpha$ -binding site predictions.** We scanned the putative promoter sequence (-10 kb upstream and +10 kb downstream of the 5' end of the annotated T-UCR) for predicted HIF-1 $\alpha$ -binding sites generated using the MATCH program and the V\$HIF1\_Q3 (5'-gnnkACGTGcggnn-3'), V\$HIF1\_Q5 (5'-ngtACGTGcngb-3') and V\$HIF1\_Q6 (5'-nRCGTGngn-3') position weight matrices from the TRANSFAC database, version 9.1<sup>46</sup> (<http://www.generegulation.com>). The position weight matrices describe the position preferences of different nucleotides in the HIF-binding site.<sup>2</sup>

ChIP and the promoter reporter assays are detailed in Supplementary Material.

**Detection of global O-GlcNAc cellular-protein modifications.**

To detect O-GlcNAc-modified proteins, we chemically (PUGNAc as a 1,5-hydroxymilactone is an inhibitor of various N-acetylhexosaminidases prevented the removal of O-GlcNAc modifications by the glycosidase MGEA5 (meningioma-expressed antigen 5 (hyaluronidase)).<sup>47</sup> PUGNAc treatment was initiated 24 h after siRNA transfection. Fresh complete media containing PUGNAc was added to final concentrations of 0, 40 and 80  $\mu$ M, respectively, and in duplicate wells. The stock PUGNAc solution (SIGMA; cat. no. 132489-69-1) was prepared following the manufacturer's instructions and stored at -20 °C. After 72 h under normoxic or hypoxic condition, cells were washed with phosphate-buffered saline (PBS) and later harvested for RNA and protein extractions to determine siRNA knockdown efficiency and to detect changes in O-GlcNAc, respectively. Primary and mouse-specific secondary antibodies for detecting O-GlcNAc were reported by Caldwell *et al.*<sup>22</sup>

**ORF detection.** *OGT* isoforms 1 and 2, and the retained intron sequence (intron 4) of the *OGT* gene were all tested for ORF using the NCBI ORF finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>).

**Cell fractionation.** Cells (HCT-116, 2  $\times$  10<sup>6</sup>) were pelleted, and cytoplasmic and nuclear fractions were obtained using the PARIS kit following the manufacturer's protocol (Ambion RNA, Life Technologies, Carlsbad, CA, USA). Oxygen and Glucose deprivation assays are described in Supplementary Materials.

**ncRNA array data analyses.** Microarray chips were developed with a total of 481 human UCR sequences as in <http://www.so.eucsc.edu/~jill/ultra.html>. For each UCR, two 40-mer probes were designed, one corresponding to the sense genomic sequence (named '+') and the other to the complementary sequence

(named 'A +'). The design criteria were as described<sup>48</sup> and the detailed array description is in Calin *et al.*<sup>9</sup> Images were quantified using the GenePix Pro 6.0 (Axon Instruments). Raw data were normalized and analyzed in GeneSpring GX 7.3 (Agilent Technologies, Santa Clara, CA, USA). Tumors were normalized using the on-chip and on-gene median normalization of the GeneSpring (Agilent Technologies). Statistical comparisons of tumors and normal tissues were performed by filtering on fold change and then by using the Welch *t*-test with Benjamini–Hochberg correction for reduction of false-positives. All data were submitted to the ArrayExpress database (accession number E-TABM-184).

**Statistics.** All results are based on at least two independent experimental replicates, with error bars representing the standard error of the mean unless stated otherwise. Unpaired two-tailed Student's *t*-tests were used to address differences between groups. Additionally, where indicated *R*-statistic was used (version 2.14.2).

### Conflict of Interest

The authors declare no conflict of interest.

**Acknowledgements.** Dr. Calin is The Alan M. Gewirtz Leukemia and Lymphoma Society Scholar. He is supported also as a Fellow at The University of Texas, MD Anderson Research Trust, as a University of Texas System Regents Research Scholar and by the CLL Global Research Foundation. Work in Dr. Calin's laboratory is supported in part by the NIH/NCI (CA135444), a Department of Defense Breast Cancer Idea Award, Developmental Research Awards in Breast Cancer, Ovarian Cancer, Brain Cancer, Prostate Cancer, Multiple Myeloma, Leukemia (P50 CA100632) and Head and Neck (P50 CA097007) SPORES, a SINF MDACC\_DKFZ grant in CLL, the Laura and John Arnold Foundation, the RGK Foundation and the Estate of CG Johnson, Jr. Dr. Ling is an Odyssey Fellow, and his work is supported in part by the Odyssey Program and The Estate of CG Johnson, JR at The University of Texas, MD Anderson Cancer Center. Additional support for the research was provided by the Slovenian Research Agency (ARRS) through the Research program P4-0220 to TK and project 30767 to JF. Dr. Ivan is an American Cancer Society Research Scholar (RSG-09-244-01-TBG). Work in Dr. Ivan's laboratory is also supported by the NIH/NCI (R01 CA155332-01A1). We thank Mei Koh for assistance with the hypoxia experiments. We thank Ashley R. Heath (Sigma-Aldrich, Woodlands, TX, USA) for the siRNA design. STR DNA fingerprinting was performed by the Cancer Center Support grant-funded Characterized Cell Line core, NCI no. CA16672. We thank Ann Sutton (Department of Scientific Publications, MD Anderson Cancer Center) for her help with the editing of this manuscript.

### Authors Contributions

**Study concept and design:** Calin, Ivan and Kunej; **development of methodology:** Calin, Ivan, Kunej and Davuluri; **acquisition of data:** Ferdin, Wu, Nishida, Nicoloso, Shah, Devlin, Ling, Shimizu, Kumar, Cortez and Schmittgen; **analysis and interpretation of data:** Ferdin, Nishida, Voorhoeve, Reginato, Negrini, Davuluri, Kunej, Ivan and Calin; **statistical analysis:** Ferracin, Bi, Yang, Czerniak, Zhang and Davuluri; **administrative, technical or material support:** Shimizu, Schmittgen and Reginato; **drafting of the manuscript:** Ferdin, Nishida, Kunej, Ivan and Calin. All authors critically reviewed the manuscript and approved the final version.

- Harris AL. Hypoxia—a key regulatory factor in tumour growth. *Nat Rev Cancer* 2002; **2**: 38–47.
- Kulshreshtha R, Ferracin M, Wojcik SE, Garzon R, Alder H, Agosto-Perez FJ *et al.* A microRNA signature of hypoxia. *Mol Cell Biol* 2007; **27**: 1859–1867.
- Chan SY, Zhang YY, Hemann C, Mahoney CE, Zweier JL, Loscalzo J. MicroRNA-210 controls mitochondrial metabolism during hypoxia by repressing the iron-sulfur cluster assembly proteins ISCU1/2. *Cell Metab* 2009; **10**: 273–284.
- Favaro E, Ramachandran A, McCormick R, Gee H, Blancher C, Crosby M *et al.* MicroRNA-210 regulates mitochondrial free radical response to hypoxia and krebs cycle in cancer cells by targeting iron sulfur cluster protein ISCU. *PLoS One* 2010; **5**: e10345.
- Devlin C, Greco S, Martelli F, Ivan M. miR-210: more than a silent player in hypoxia. *JUBMB Life* 2011; **63**: 94–100.
- Huang X, Le QT, Giaccia AJ. miR-210—micromanager of the hypoxia pathway. *Trends Mol Med* 2010; **16**: 230–237.
- Mattick JS. The genetic signatures of noncoding RNAs. *PLoS Genet* 2009; **5**: e1000459.
- Calin G, Liu C, Ferracin M, Hyslop T, Spizzo R, Sevignani C *et al.* Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. *Cancer Cell* 2007; **12**: 215–229.
- Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, Mattick JS *et al.* Ultraconserved elements in the human genome. *Science* 2004; **304**: 1321–1325.
- Bejerano G, Haussler D, Blanchette M. Into the heart of darkness: large-scale clustering of human non-coding DNA. *Bioinformatics* 2004; **20**(Suppl 1): i40–i48.
- Braconi C, Valeri N, Kogure T, Gasparini P, Huang N, Nuovo GJ *et al.* Expression and functional role of a transcribed noncoding RNA with an ultraconserved element in hepatocellular carcinoma. *Proc Natl Acad Sci USA* 2011; **108**: 786–791.
- Scaruffi P, Stigliani S, Coco S, Valdora F, De Vecchi C, Bonassi S *et al.* Transcribed-ultra conserved region expression profiling from low-input total RNA. *BMC Genomics* 2010; **11**: 149.
- Mestdagh P, Fredlund E, Pattyn F, Rihani A, Van Maerken T, Vermeulen J *et al.* An integrative genomics screen uncovers ncRNA T-UCR functions in neuroblastoma tumours. *Oncogene* 2010; **29**: 3583–3592.
- Bejerano G, Lowe CB, Ahituv N, King B, Siepel A, Salama SR *et al.* A distal enhancer and an ultraconserved exon are derived from a novel retroposon. *Nature* 2006; **441**: 87–90.
- Pennacchio LA, Ahituv N, Moses AM, Prabhakar S, Nobrega MA, Shoukry M *et al.* *In vivo* enhancer analysis of human conserved non-coding sequences. *Nature* 2006; **444**: 499–502.
- Xu J, Wang B, Xu Y, Sun L, Tian W, Shukla D *et al.* Epigenetic regulation of HIF-1 $\alpha$  in renal cancer cells involves HIF-1 $\alpha$ /2 $\alpha$  binding to a reverse hypoxia-response element. *Oncogene* 2012; **31**: 1065–1072.
- Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM *et al.* The human genome browser at UCSC. *Genome Res* 2002; **12**: 996–1006.
- Ivan M, Kondo K, Yang H, Kim W, Valiando J, Ohh M *et al.* HIF1 $\alpha$  targeted for VHL-mediated destruction by proline hydroxylation: implications for O<sub>2</sub> sensing. *Science* 2001; **292**: 464–468.
- Nolte D, Muller U. Human O-GlcNAc transferase (OGT): genomic structure, analysis of splice variants, fine mapping in Xq13.1. *Mamm Genome* 2002; **13**: 62–64.
- Kreppel LK, Blomberg MA, Hart GW. Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats. *J Biol Chem* 1997; **272**: 9308–9315.
- Lubas WA, Frank DW, Krause M, Hanover JA. O-Linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats. *J Biol Chem* 1997; **272**: 9316–9324.
- Caldwell SA, Jackson SR, Shahriari KS, Lynch TP, Sethi G, Walker S *et al.* Nutrient sensor O-GlcNAc transferase regulates breast cancer tumorigenesis through targeting of the oncogenic transcription factor FoxM1. *Oncogene* 2010; **29**: 2831–2842.
- Mi W, Gu Y, Han C, Liu H, Fan Q, Zhang X *et al.* O-GlcNAcylation is a novel regulator of lung and colon cancer malignancy. *Biochim Biophys Acta* 2011; **1812**: 514–519.
- Lynch TP, Ferrer CM, Jackson SR, Shahriari KS, Vosseller K, Reginato MJ. Critical role of O-GlcNAc transferase in prostate cancer invasion, angiogenesis and metastasis. *J Biol Chem* 2012; **287**: 11070–11081.
- Taylor RP, Parker GJ, Hazel MW, Soesanto Y, Fuller W, Yazzie MJ *et al.* Glucose deprivation stimulates O-GlcNAc modification of proteins through up-regulation of O-linked N-acetylglucosaminyltransferase. *J Biol Chem* 2008; **283**: 6050–6057.
- Robinson KA, Weinstein ML, Lindenmayer GE, Buse MG. Effects of diabetes and hyperglycemia on the hexosamine synthesis pathway in rat muscle and liver. *Diabetes* 1995; **44**: 1438–1446.
- Butkinaree C, Park K, Hart GW. O-linked beta-N-acetylglucosamine (O-GlcNAc): extensive crosstalk with phosphorylation to regulate signaling and transcription in response to nutrients and stress. *Biochim Biophys Acta* 2010; **1800**: 96–106.
- Dong DL, Hart GW. Purification and characterization of an O-GlcNAc selective N-acetyl-beta-D-glucosaminidase from rat spleen cytosol. *J Biol Chem* 1994; **269**: 19321–19330.
- Visel A, Prabhakar S, Akiyama JA, Shoukry M, Lewis KD, Holt A *et al.* Ultraconservation identifies a small subset of extremely constrained developmental enhancers. *Nat Genet* 2008; **40**: 158–160.
- Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G *et al.* Long noncoding RNAs with enhancer-like function in human cells. *Cell* 2010; **143**: 46–58.
- Ozsolak F, Poling LL, Wang Z, Liu H, Liu XS, Roeder RG *et al.* Chromatin structure analyses identify miRNA promoters. *Genes Develop* 2008; **22**: 3172–3183.
- Love DC, Ghosh S, Mondoux MA, Fukushige T, Wang P, Wilson MA *et al.* Dynamic O-GlcNAc cycling at promoters of *Caenorhabditis elegans* genes regulating longevity, stress, and immunity. *Proc Natl Acad Sci USA* 2010; **107**: 7413–7418.
- Nott A, Meislin SH, Moore MJ. A quantitative analysis of intron effects on mammalian gene expression. *RNA* 2003; **9**: 607–617.
- Ryu WS, Mertz JE. Simian virus 40 late transcripts lacking excisable intervening sequences are defective in both stability in the nucleus and transport to the cytoplasm. *J Virol* 1989; **63**: 4386–4394.
- Chang CY, Hong WW, Chong P, Wu SC. Influence of intron and exon splicing enhancers on mammalian cell expression of a truncated spike protein of SARS-CoV and its implication for subunit vaccine development. *Vaccine* 2006; **24**: 1132–1141.



36. Zhao C, Hamilton T. Introns regulate the rate of unstable mRNA decay. *J Biol Chem* 2007; **282**: 20230–20237.
37. Wang HF, Feng L, Niu DK. Relationship between mRNA stability and intron presence. *Biochem Biophys Res Commun* 2007; **354**: 203–208.
38. Gencheva M, Lin TY, Wu X, Yang L, Richard C, Jones M *et al*. Nuclear retention of unspliced pre-mRNAs by mutant DHX16/hPRP2, a spliceosomal DEAH-box protein. *J Biol Chem* 2010; **285**: 35624–35632.
39. Gardner LB. Hypoxic inhibition of nonsense-mediated RNA decay regulates gene expression and the integrated stress response. *Mol Cell Biol* 2008; **28**: 3729–3741.
40. Buchman AR, Berg P. Comparison of intron-dependent and intron-independent gene expression. *Mol Cell Biol* 1988; **8**: 4395–4405.
41. Le Hir H, Nott A, Moore MJ. How introns influence and enhance eukaryotic gene expression. *Trends Biochem Sci* 2003; **28**: 215–220.
42. Kibel A, Iliopoulos O, DeCaprio JA, Kaelin WG Jr. Binding of the von Hippel-Lindau tumor suppressor protein to Elongin B and C. *Science* 1995; **269**: 1444–1446.
43. Romano P, Manniello A, Aresu O, Armento M, Cesaro M, Parodi B. Cell Line Data Base: structure and recent improvements towards molecular authentication of human cell lines. *Nucleic Acids Res* 2009; **37**, Database issue D925–D932.
44. Safdar A, Abadi A, Akhtar M, Hettinga BP, Tarnopolsky MA. miRNA in the regulation of skeletal muscle adaptation to acute endurance exercise in C57Bl/6J male mice. *PLoS One* 2009; **4**: e5610.
45. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001; **25**: 402–408.
46. Matys V, Fricke E, Geffers R, Gossling E, Haubrock M, Hehl R *et al*. TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* 2003; **31**: 374–378.
47. Gao Y, Wells L, Comer FI, Parker GJ, Hart GW. Dynamic O-glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neutral, cytosolic beta-N-acetylglucosaminidase from human brain. *J Biol Chem* 2001; **276**: 9838–9845.
48. Liu CG, Calin GA, Meloon B, Gamlieil N, Sevignani C, Ferracin M *et al*. An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. *Proc Natl Acad Sci USA* 2004; **101**: 9740–9744.

Supplementary Information accompanies this paper on Cell Death and Differentiation website (<http://www.nature.com/cdd>)