UNIVERSITY OF LJUBLJANA
BIOTECHNICAL FACULTY

Jana FERDIN

# INFLUENCE OF HYPOXIA ON EXPRESSION OF ULTRACONSERVED GENOMIC REGIONS IN TUMOR TISSUE

DOCTORAL DISSERTATION

# VPLIV HIPOKSIJE NA IZRAŽANJE ZELO OHRANJENIH REGIJ GENOMA V TUMORSKEM TKIVU

DOKTORSKA DISERTACIJA

Ljubljana, 2013

On the basis of the Statute of the University of Ljubljana and according to the decisions of the Senate of the Biotechnical Faculty and the University Senate dating 27th of September 2010, it has been confirmed that the candidate fulfills all the conditions for direct transmission to Doctorate Postgraduate Study of Biological and Biotechnical Sciences and acquiring a PhD in Genetics. Prof. Tanja Kunej, PhD has been appointed as candidate's supervisor and prof. George Adrian Calin, PhD MD as co-supervisor.

Na podlagi Statuta Univerze v Ljubljani ter po sklepu Senata Biotehniške fakultete in sklepa Senata Univerze z dne 27.9.2010 je bilo potrjeno, da kandidatka izpolnjuje pogoje za neposreden prehod na doktorski Podiplomski študij bioloških in biotehniških znanosti ter opravljanje doktorata znanosti s področja genetike. Za mentorico je bila imenovana prof. dr. Tanja Kunej in za somentorja prof. dr. George Adrian Calin.

The research work was carried out at the Department of Animal Science, Biotechnical Faculty, University of Ljubljana and at the Department of Experimental Therapeutics, MD Anderson Cancer Center, University of Texas, Houston, USA.

Raziskovalno delo je bilo opravljeno na Oddelku za zootehniko, Biotehniška fakulteta, Univerza v Ljubljani in Department of Experimental Therapeutics v MD Anderson Cancer Center, University of Texas, Houston, ZDA.

Commission on doctoral dissertation assessment and defense:

| | |
|---|---|
| Chairman: | Prof. Simon HORVAT, PhD<br>University of Ljubljana, Biotechnical Faculty and National Institute of Chemistry, Slovenia |
| Member and supervisor: | Prof. Tanja KUNEJ, PhD<br>University of Ljubljana, Biotechnical Faculty, Slovenia |
| Member and co-supervisor: | Prof. George Adrian CALIN, MD, PhD<br>University of Texas, MD Anderson Cancer Center, ZDA |
| Member: | Prof. Maja ČEMAŽAR, PhD<br>Institute of Oncology, Slovenia |

Date of dissertation defense:

Doctoral dissertation is a result of personal research work. As undersigned I agree with publication of my thesis in full text on the website of the Digital Library of the Biotechnical Faculty. I declare that the work that I have submitted in electronic version is identical to the printed version.

Jana Ferdin

## KEY WORDS DOCUMENTATION

DN Dd
DC UDC 575:616(043.3)=111
CX medicine/cancer/tumor tissue/molecular genetics/genome/ultraconserved regions /hypoxia/HIF1A
AU FERDIN, Jana
AA KUNEJ, Tanja (supervisor)/CALIN, George Adrian (co-supervisor)
PP SI-1000 Ljubljana, Jamnikarjeva 101
PB University of Ljubljana, Biotechnical Faculty, Postgraduate Study of Biological and Biotechnical Sciences, Field: Genetics
PY 2013
TI INFLUENCE OF HYPOXIA ON EXPRESSION OF ULTRACONSERVED GENOMIC REGIONS IN TUMOR TISSUE
DT Doctoral Dissertation
NO XVI, 150 p., 4 tab., 45 fig., 15 ann., 348 ref.
LA en
AL en/sl

AB In hypoxic cells important responses are activated for metabolic, bioenergetics, and redox demands and primary transcriptional response to hypoxia is mediated by the hypoxia-inducible factors (HIFs). Alpha subunit of HIF heterodimer is $O_2$-sensitive, and once stabilized in hypoxia, it functions as master regulator of various genes involved in hypoxia pathway. Identification of upstream *HIF1A* regulators and its downstream targets are essential for understanding the cellular and systemic response to hypoxia, however, information related to this topic are scattered among numerous publications. Even though that recent data linked hypoxia to the function of specific miRNAs, very little is known how hypoxia affects other types of noncoding transcripts. The aim of this study was to construct an Atlas of *HIF1A* gene regulatory network integrating known information from publications and results of bioinformatics analyses and to experimentally validate a novel HIF1A target. Atlas comprises: 1) functional *HIF1A* polymorphisms from 70 association studies; 2) validated and predicted transcription factor (TF) and miRNA binding sites; and 3) HIF1A downstream targets and pathways. To date, 30 *HIF1A* single nucleotide polymorphisms (SNPs) were genotyped and 15 of them were found to be associated with 32 phenotypes including 16 cancer types. Out of 62 TF binding sites identified within upstream HIF1A region, additional six were predicted to be gained/lost due to four SNPs. Five miRNAs were previously identified to regulate *HIF1A* and 150 miRNAs were predicted to target polymorphic regions residing within exons, 5'- and 3'-UTR regions. Downstream HIF1A targets were suggested to be involved in 21 enriched pathways. Additionally, we also demonstrated for the first time a functional link between $O_2$ deprivation and long transcripts of ultraconserved regions (T-UCRs), which we termed hypoxia-induced noncoding ultraconserved transcripts (HINCUTs). These T-UCRs are predominantly nuclear and HIF1A is partly responsible for the induction of at least some of them. One T-UCR, uc.475, whose functional role we studied, is part of a retained intron of its host gene, *OGT*. In conclusion, the constructed Atlas of *HIF1A* gene regulatory network, which we expanded with two new targets, presents a central location of information that will enable to accelerate research of this field. Consistent with the hypothesis we showed that T-UCRs have important function in tumor formation, since uc.475 supports cell proliferation specifically in hypoxia and may be critical for optimal O-GlcNAcylation of proteins when $O_2$ tension is limiting. Our data gives a first glimpse of a novel functional hypoxic network composed of protein-coding transcripts and noncoding RNAs.

# KLJUČNA DOKUMENTACIJSKA INFORMACIJA

ŠD Dd
DK UDK 575:616(043.3)=111
KG medicina/rak/tumorsko tkivo/molekularna genetika/genom/zelo ohranjene regije /hipoksija/HIF1A
AV FERDIN, Jana, univ. dipl. mikr.
SA KUNEJ, Tanja (mentor)/CALIN, George Adrian (somentor)
KZ SI-1000 Ljubljana, Jamnikarjeva 101
ZA Univerza v Ljubljani, Biotehniška fakulteta, Podiplomski študij bioloških in biotehniških znanosti, področje: genetika
LI 2013
IN VPLIV HIPOKSIJE NA IZRAŽANJE ZELO OHRANJENIH REGIJ GENOMA V TUMORSKEM TKIVU
TD Doktorska disertacija
OP XVI, 150 str., 4 pregl., 45 sl., 15 pril., 348 vir.
IJ en
JI en/sl
AI V hipoksičnih celicah so aktivirani ključni odzivi za zagotovitev metabolnih, bioenergetskih in redoks potreb celice. Primaren odziv na hipoksijo je sprožen s hipoksijo induciranimi transkripcijskimi dejavniki, HIF. Alfa podenota heterodimera HIF je občutljiva na $O_2$, vendar v odsotnosti le-tega postane stabilna in uravnava prepisovanje številnih genov vpletenih v hipoksične poti. Določitev regulatorjev *HIF1A* in njegovih tarčnih genov je ključno za razumevanje celičnega in sistemskega odziva na hipoksijo, vendar so te informacije zelo razpršene med številnimi publikacijami. Kljub temu, da je bila hipoksija že povezana s funkcijo miRNA, je zaenkrat malo znanega kako le-ta vpliva na druge skupine ne-kodirajočih transkriptov. Namen te študije je bil izdelati Atlas regulatorne mreže gena *HIF1A* s sintezo že znanih informacij iz publikacij in rezultatov bioinformacijskih analiz ter eksperimentalno potrjevanje nove tarče HIF1A. Izdelani Atlas sestavljajo: 1.) funkcionalni polimorfizmi *HIF1A* iz 70-ih analiz povezav genotipa s fenotipom; 2) potrjena in predvidena vezavna mesta za transkripcijske dejavnike (TF) in miRNA; 3) HIF1A tarče in biološke poti, v katere so uvrščene. Do danes je bilo genotipiziranih že 30 polimorfizmov posameznega nukleotida (SNP) gena *HIF1A* in 15 izmed njih povezanih z 32 fenotipi in boleznimi vključno s 16 tipi raka. Izmed 62 vezavnih mest za TF v območju pred začetkom gena *HIF1A*, je bilo za dodatnih šest TF-jev predvideno, da pridobijo/izgubijo svoje vezavno mesto zaradi prisotnosti štirih SNP-jev gena *HIF1A*. Za pet miRNA je bilo že dokazano, da uravnavajo izražanje *HIF1A*, medtem ko je bilo za 150 miRNA predvideno, da prepoznavajo mesta, ki se prekrivajo s preučevanimi polimorfizmi *HIF1A* znotraj eksonov, 5'- in 3'-neprevedenih regij gena. Analiza HIF1A tarč je pokazala, da so le-te uvrščene v 21 različnih bioloških poti. V tej raziskavi je bila prvič dokazana povezava med znižano vsebnostjo $O_2$ in dolgimi prepisi zelo ohranjenih regij (T-UCR), ki smo jih poimenovali HINCUTs (angl. *hypoxia-induced noncoding ultraconserved transcripts*). Te regije se izražajo pretežno v jedru celice in je HIF1A vsaj deloma odgovoren za indukcijo prepisovanja nekaterih izmed njih. Prepis uc.475, ki smo ga podrobneje preučevali, je del neizrezanega introna gostiteljskega gena *OGT*. Izdelani Atlas regulatorne mreže gena *HIF1A* smo tako dopolnili z dvema novima tarčama in predstavlja centralno mesto, ki bo omogočilo pospešitev raziskav na tem področju. V skladu s hipotezo smo potrdili, da imajo T-UCR-ji pomembno vlogo pri nastanku tumorjev, saj uc.475 predvsem v hipoksiji podpira proliferacijo celic in je v takem okolju lahko ključen za vzdrževanje ustrezne ravni O-GlcNAcilacije proteinov. Naši rezultati predstavljajo prvi vpogled v nove funkcije omrežja hipoksije, ki ga sestavljenjo protein kodirajoči in ne-kodirajoči prepisi RNA.

TABLE OF CONTENTS

# LIST OF TABLES

p.

# LIST OF FIGURES

p.

# LIST OF ANNEXES

# ABBREVIATIONS AND SYMBOLS

| | |
|---|---|
| "+" | sense genomic sequence |
| "A+" | anti-sense genomic sequence |
| "e" | exon |
| aa | amino acid |
| AAAAAA | polyA tail |
| AMPL | amplification |
| ARNT | aryl hydrocarbon receptor nuclear translocator (synonym HIF1B) |
| bHLH–PAS | basic helix-loop-helix–Per-Arnt-Sim |
| bp | base pair |
| BSA | bovine serum albumin |
| CA9 | carbonic anhydrase IX |
| CAD | coronary artery disease with stable exertional angina |
| CAGR | cancer-associated genomic region |
| CBP | CREB binding protein |
| cDNA | complementary DNA |
| CDS | coding DNA sequence |
| ChIP | chromatin immunoprecipitation |
| ChIP-seq | chromatin immunoprecipitation (ChIP) with massively parallel DNA sequencing |
| CIS | common integrative sites |
| CLL | chronic lymphocytic leukemia |
| CNS | conserved non-genic sequence |
| COPD | chronic obstructive pulmonary disease |
| CpG | -C-phosphate-G- or CG site |
| Cq | quantification cycle |
| crasiRNA | Centrosome-associated RNA |
| CRC | colorectal cancer |
| CREB1 | cAMP responsive element binding protein 1 |
| C-TAD | C-terminal transactivation domain |
| DEL | deletion |
| DMOG | dimethyloxalylglycine |
| DMSO | dimethyl sulfoxide |
| dsDNA | double-stranded DNA |
| DTT | dithiothreitol |
| EDTA | ethylenedinitrilo-tetraacetic acid |
| EGLN1-3 | egl nine homolog 1-3; prolyl hydroxylases (PHD1-3) |
| EGTA | ethylene glycol tetraacetic acid |
| EP300 | E1A binding protein p300 |
| EPAS1 | endothelial PAS protein 1 |

| | |
|---|---|
| FACS | fluorescence activated cell sorting |
| FBS | fetal bovine serum |
| FIH | factor inhibiting HIF |
| FOXA1 | forkhead box A1 |
| FOXC1 | forkhead box C1 |
| FOXI1 | forkhead box I1 |
| FRA | fragile site |
| GAPDH | glyceraldehyde-3-phosphate dehydrogenase |
| gDNA | genomic DNA |
| GlcNAc | β-D-N-acetylglucosamine |
| HBB | hemoglobin, beta |
| HCR | highly conserved region |
| HGNC | HUGO Gene Nomenclature Committee |
| HGVS | Human Genome Variation Society |
| HIF | hypoxia-inducible factor |
| HIF1AN | HIF1A subunit inhibitor |
| HINCUT | hypoxia-induced noncoding ultraconserved transcript |
| HLF | HIF1A‑like factor |
| HNF4A | hepatocyte nuclear factor 4, alpha |
| HRE | hypoxia response element |
| HRF | HIF-related factor |
| HRM | hypoxia-regulated miRNA |
| HRP | horseradish peroxidase |
| INS | insertion |
| ISH | *in situ* hybridization |
| JUN | jun proto-oncogene |
| kb | kilobase |
| KEGG | Kyoto Encyclopedia of Genes and Genomes |
| KO | knockout |
| LCNS | long conserved noncoding sequence |
| lncRNA | long noncoding RNA |
| LOH | loss of heterozygosity |
| MCS | multiple species conserved sequence |
| METH | methylation |
| MGEA5 | meningioma-expressed antigen 5 (hyaluronidase), (synonym OGA) |
| miRNA | microRNA |
| MOP2 | member of PAS family 2 |
| moRNA | microRNA-offset |
| mRNA | messenger RNA |
| MSY-RNA | MSY2-associated RNA |

| | |
|---|---|
| MUT | mutation |
| NC | negative control |
| ncRNA | noncoding RNA |
| NLS | nuclear localization signal |
| NMD | nonsense-mediated RNA decay |
| NP-40 | nonyl phenoxypolyethoxylethanol |
| nsSNP | non-synonymous SNP |
| N-TAD | N-terminal transactivation domain |
| NTC | no-template control |
| ODD domain | $O_2$-dependent degradation domain |
| OGA | O-GlcNAcase (official name is MGEA5) |
| OGD | $O_2$- and glucose- deprivation |
| OGT | O-linked N-acetylglucosamine (GlcNAc) transferase |
| ORF | open reading frame |
| OSCC | oral squamous cell carcinoma |
| p300/CBP | CREB-binding protein |
| PANTHER | Protein ANalysis THrough Evolutionary Relationships |
| PAR | promoter-associated RNA |
| PARP | poly ADP-ribose polymerase |
| PBS | phosphate-buffered saline |
| PCR | polymerase chain reaction |
| PHD | prolyl hydroxylase domain-containing enzymes (official name EGLN) |
| PI | propidium iodide |
| piRNA | piwi-interacting RNA |
| POL II | RNA polymerase II |
| PPIA | peptidylprolyl isomerase A (cyclophilin A) |
| pre-miRNA | precursor miRNA molecule |
| pri-miRNA | primary miRNA molecule |
| Pro | proline |
| PT1 | truncated promoter 1 |
| PT2 | truncated promoter 2 |
| PUGNAc | 1,5-hydroximolactone |
| pVHL | von Hippel-Lindau protein complex |
| PWMs | positional weight matrices |
| RCC | renal cell carcinoma |
| RFU | relative fluorescence units |
| RISC | RNA-induced silencing complex |
| RLU | relative light units |
| RPLP0 | ribosomal protein, large, P0 |
| rRNA | ribosomal RNA |

| | |
|---|---|
| rs | reference SNP |
| RT | reverse transcription |
| RT-qPCR | quantitative reverse transcription PCR |
| SD | standard deviation |
| sdRNA | sno-derived RNA |
| SDS | sodium dodecyl sulfate |
| SIM | single-minded protein |
| snoRNA | small nucleolar RNAs |
| SNP | single nucleotide polymorphism |
| snRNA | small nuclear RNA |
| SSC | saline-sodium citrate |
| sSNP | synonymous SNP |
| T1DM | diabetes mellitus type 1 |
| T2DM | diabetes mellitus type 2 |
| TAL1 | T-cell acute lymphocytic leukemia 1 |
| TCF3 | transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47) |
| tel-sRNA | thelomere small RNA |
| TF | transcription factor |
| TFBS | TF binding sites |
| TRANSFAC | TRANScription FACtor database |
| TRANSL | translocation |
| tRNA | transfer RNA |
| TSS | transcription start site |
| T-UCR | transcribed-ultraconserved gene |
| U6 snRNA | RNU6-1; RNA, U6 small nuclear 1 |
| UCG | ultraconserved gene |
| UCR | ultraconserved region |
| UCSC | University of California, Santa Cruz |
| UTR | untranslated region |
| WoS | Web of Science |
| WT | wild type |
| xiRNA | X-inactivation RNA |

# 1 INTRODUCTION AND RESEARCH HYPOTHESIS (UVOD IN PREDSTAVITEV HIPOTEZ)

## 1.1 INTRODUCTION AND DEFINITION OF A SCIENTIFIC PROBLEM

Molecular oxygen ($O_2$) is essential molecule that serves as a key substrate in cellular metabolism and bioenergetics as an electron acceptor in many organic and inorganic reactions. In a variety of physiological and pathological states organism encounters insufficient $O_2$ availability or hypoxia. Hypoxia, defined as reduced $O_2$ levels, occurs in a variety of pathological conditions including stroke, tissue ischemia, inflammation, and the growth of solid tumors (Bertout et al., 2008). Until few years ago, most mapping projects have focused only on protein coding sequences, which represent only 1-2% of a whole human genome. Recently more attention is given to noncoding regions of the genome previously known as "junk" DNA. These regions were later named noncoding RNA molecules (ncRNA), in group of which microRNAs (miRNAs) are the most known member of the family involved in a large number of cellular processes, including tumor formation.

In mammal cells primary response to low $O_2$ conditions (hypoxia) is made rapidly through adaptation to conditions they are being captured. In hypoxic cells important responses are activated for metabolic, bioenergetics, and redox demands (reviewed in (Majmundar et al., 2010)). Primary transcriptional response to hypoxia is mediated by the hypoxia-inducible factors (HIFs). Protein HIF is known as pivotal regulator under hypoxia stress and beside his adaptive response in cellular stress, it carries important roles in physiological and pathological processes (Majmundar et al., 2010; Semenza, 2003). It is a heterodimer consisted from an $O_2$-sensitive (alpha, A) and an $O_2$-stable (beta, B) subunit (Wang and Semenza, 1993). In mammals three different HIFA isoforms are present, among which HIF1A is expressed ubiquitously meanwhile HIF2A and HIF3A expression vary depending on type of tissue cells (Bertout et al., 2008). The life cycle of HIF is well studied, dynamic and by now quite well understood process (Wang et al., 1995). In normally oxidized conditions (normoxia) alpha-subunits are hydroxylated at conserved proline residues (*p.Pro402* and *p.Pro564*) (Ivan et al., 2001) by $O_2$ regulated prolyl hydroxylase domain-containing enzymes (EGLN1, 2, 3 or PHD1, 2, 3) (Chan et al., 2005; Kaelin and Ratcliffe, 2008; Wang et al., 1995). Marked HIFA-subunits are consequentially recognized by E3 ubiquitin ligase from the von Hippel-Lindau protein complex (pVHL) for proteasomal degradation (Majmundar et al., 2010; Wang et al., 1995). The VHL protein can target the N-terminal transactivation domain (N-TAD) within the $O_2$-dependent degradation domain (ODD domain), which controls HIF1A degradation by ubiquitin-proteasome pathway and consists of approximately 200 amino acid residues (Huang et al., 1998). The removal of ODD domain renders HIF1A stability even under normoxic conditions, consequentially

resulting in autonomous HIF1A heterodimerisation, DNA binding, and transactivation independently from hypoxic signaling (Huang et al., 1998).

The HIF1A mediates transcriptional responses to hypoxia for a high number of genes to control cellular $O_2$ supply and maintain cell viability during periods of low $O_2$ concentration (Keith et al., 2012; Vilela et al., 2008; Wang et al., 1995; Wenger et al., 2005). Role of HIF1A is reported in metabolism and redox homeostasis (glucose catabolism, function in metabolism, regulation of lipid metabolism), in vascular responses in hypoxia (ischemia-induced angiogenesis, endothelial cells), in cancer (tumorigenesis, metastasis, tumor angiogenesis, cancer stem cells, regulation by cancer metabolism), in inflammation (regulation in inflammatory cells, myeloid cell function, tumor-associated macrophages), and also as part of systemic response to hypoxia (reviewed in (Majmundar et al., 2010)). To gain insight into the molecular pathways regulated by HIF1, it is essential to identify downstream-target genes.

To date HIF1 was confirmed as a master regulator of hundreds of genes, including a panel of miRNAs. The link between hypoxia and miRNA expression (hypoxia-regulated miRNAs (HRMs); for example miR-23a and miR-23b (Kulshreshtha et al., 2007)) was first described by Kulshreshtha et al. (2007) and a subgroup of them were suggested to play a role in cell survival in a low $O_2$ environment (Crosby et al., 2009a; Devlin et al., 2011; Kulshreshtha et al., 2008; Kulshreshtha et al., 2007). Besides, some miRNAs were found to target *HIF1A* and therefore regulate its expression; miR-20b (Cascio et al., 2010; Lei et al., 2009), miR-199a (Rane et al., 2009; Rane et al., 2010), cluster miR-17-92 (Taguchi et al., 2008; Yan et al., 2009), miR-519c (Cha et al., 2010), and miR-155 (Bruning et al., 2011) (reviewed in (Shen et al., 2013)). Therefore, identification of upstream *HIF1A* regulators and its downstream targets are essential for understanding complete HIF1A regulatory network. Several studies identified HIF1 targets based on response to hypoxia and the presence of conserved HIF1 binding sites within target genes proximal promoters, named hypoxia response elements (HRE). Benita et al. (2009) performed pathway enrichment analysis using known and predicted HIF1A targets and discovered that HIF1A modulates metabolic pathways, bioenergetics and processes relevant to cancer onset and progression. Some TFs, such as *TWIST1* (Yang et al., 2008) and *GATA1* (Zhang et al., 2012) have been shown to be regulated by HIF1A, whereas also *HIF1A* mRNA expression can be regulated by other TFs, like SP1 (Minet et al., 1999), NFKB1/RELA (Belaiba et al., 2007) and EGR1 (Sperandio et al., 2009). For HIF1A it is known to carry an important character in diseases and phenotype outcomes (Majmundar et al., 2010; Semenza, 2003). Current knowledge about *HIF1A* is fragmented among numerous publications; therefore to construct an Atlas of *HIF1A* gene regulatory network it is important to summarize the current knowledge and to facilitate novel discoveries about its regulation. Moreover, the understanding how the expression of TF *HIF1A* is controlled is essential for reconstructing gene regulatory networks.

Hypoxia is known as one of the classic features of the neoplastic microenvironment and is causally associated with resistance to conventional anticancer therapy (Harris, 2002). As mentioned 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 $O_2$ concentration have shown that in addition to protein-encoding genes, the noncoding transcriptome is also affected by hypoxia. Some studies previously reported that hypoxia regulates specific miRNAs, which in turn regulate signals of relevance for tumor biology (Chan et al., 2009; Devlin et al., 2011; Favaro et al., 2010; Kulshreshtha et al., 2007), including angiogenesis, cell cycle, DNA repair, and energy metabolism (Devlin et al., 2011; Huang et al., 2010).

During the past few years, it has become clear that long ncRNAs (>200 base pairs (bp)), also participate in complex genetic events in eukaryotes (reviewed in (Mattick, 2009)). Currently the knowledge about miRNAs is in great advantage compared to other groups of ncRNA. Significant progress has also been made towards understanding the mechanisms of gene regulation by miRNAs, but much less is known about other ncRNA mechanism. Beside that their role is not explored and there is still not known which environmental factors influence on their expression.

One intriguing family of long ncRNAs (Calin et al., 2007) is transcribed from regions, which are absolutely conserved (100% identity with no insertions or deletions) between orthologous regions of the human, rat, and mouse genomes (Bejerano et al., 2004a; Bejerano et al., 2004b). Therefore, these 481 segments of variable length (from 200 to 779 bp) were termed “transcribed-ultraconserved regions” (T-UCRs) or ultraconserved genes (UCGs) (Calin et al., 2007). 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, alterations in T-UCR expression have been described in adult chronic lymphocytic leukemia (CLL), colorectal cancer (CRC) (Calin et al., 2007), hepatocellular carcinomas (Braconi et al., 2011; Calin et al., 2007), and neuroblastomas (Mestdagh et al., 2010; Scaruffi et al., 2010). 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 (Bejerano et al., 2006; Bejerano et al., 2004b). While ultraconserved sequences appear to act as tissue-specific regulators of gene expression during development (Bejerano et al., 2004b; Mattick, 2009; Pennacchio et al., 2006), very little is currently known 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 approach to previously published work on hypoxia-regulated miRNAs (Kulshreshtha et al., 2007).

## 1.2 PURPOSE OF THE STUDY

- To develop a central Atlas of *HIF1A* gene regulatory network consisting of transcriptional, post-transcriptional regulators, and its downstream targets.

- To discover possible functional association between hypoxia and transcribed-ultraconserved regions of the genome. We wanted to analyze the role of T-UCRs in cell adaptation under hypoxic environment and their function in overall survival of a cancer cell.

## 1.3 RESEARCH HYPOTHESES

Our hypotheses were:

- Hypoxia alters expression of transcribed ultraconserved regions (UCR).

- Transcription factor HIF directly regulates expression of UCR.

- Hypoxia affects UCR expression through HIF as a mediator and effects cancer cell proliferation.

- Silencing of overexpressed UCR in hypoxia represses cancer cell line proliferation.

## 1.4 UVOD IN OPREDELITEV ZNANSTVENEGA PROBLEMA

Molekularni kisik ($O_2$) je nujna komponenta aerobnega metabolizma, saj se z njim vzdržuje znotrajcelična bioenergetika in kot sprejemnik elektronov služi pri številnih organskih ter anorganskih reakcijah. Hipoksija, definirana kot zmanjšana vsebnost $O_2$, je prisotna pri številnih patofizioloških pogojih vključno s kapjo, tkivno ishemijo, vnetjem in rastjo čvrstih tumorjev (Bertout in sod., 2008). Še do pred nekaj leti so se pri raziskavah genoma človeka osredotočali večinoma na protein-kodirajoče regije, ki predstavljajo 1-2% celotnega genoma. V zadnjem času večjo pozornost namenjajo tudi nekodirajočim regijam genoma, ki ne nosijo zapisa za proteine in so bile v preteklosti obravnavane kot odvečna DNA genoma. Te regije se imenujejo nekodirajoče molekule RNA (nkRNA, angl. *noncoding RNA*; ncRNA), kamor uvrščamo tudi do sedaj najbolj poznane in preučene mikro RNA (angl. *microRNA*; miRNA).

V celicah sesalcev je primaren odziv na nizko vsebnost $O_2$ hiter, saj se morajo le-te čimhitreje prilagoditi danim pogojem. V hipoksičnih celicah se zato aktivirajo signali potrebni za zagotovitev metabolnih, bioenergetskih in redoks potreb celice (pregled v Majmundar in sod., 2010). Primarni odzivi se sprožijo tudi s transkripcijskimi dejavniki HIF (angl. *hypoxia-inducible factor*), ki šele v okolju z znižano vsebnostjo $O_2$ postanejo aktivni. Proteini HIF so znani osrednji regulatorji hipoksije in imajo poleg vpliva na odziv ter prilagoditev celice okolju tudi vlogo pri raznih fizioloških in patoloških procesih celice (Majmundar in sod., 2010; Semenza, 2003). Protein HIF je heterodimer sestavljeni iz na $O_2$-občutljive (alfa, A) in -neobčutljive enote (beta, B) (Wang in Semenza, 1993). V celicah sesalcev so znane tri različne izooblike proteina HIF-A, med katerimi je HIF1A izražen konstantno, medtem ko je prisotnost HIF2A in HIF3A proteinov odvisna od vrste tkivnih celic (Bertout in sod., 2008). Do danes je bilo že veliko preučenega o dinamiki uravnavanja proteinov HIF (Wang in sod., 1995). Splošno znano je, da so v okolju z normalno vsebnostjo $O_2$ (normoksija) HIF-A podenote proteina na ohranjenih prolinskih regijah (*p.Pro402* in *p.Pro564*) (Ivan in sod., 2001) hidroksilirane s prolil hidroksilazami, encimi odvisni od $O_2$ (EGLN1, 2, 3 ali PHD1, 2, 3) (Chan in sod., 2005; Kaelin in Ratcliffe, 2008; Wang in sod., 1995). Tako označene podenote HIF-A prepozna E3 ubikvitin ligaza, ki je del proteinskega kompleksa von Hippel-Lindau (pVHL) nujnega za razgradnjo proteinov (Majmundar in sod., 2010; Wang in sod., 1995). Protein VHL prepozna N-terminalno domeno (N-TAD regija; angl. *N-terminal activation domain*) znotraj na $O_2$ občutljive domene, ki uravnava obstoj proteina HIF-A (ODD domena; angl. *oxygen-dependent degradation domain*). Domena ODD obsega 200 amino kislin in v normoksiji uravnava razgradnjo HIF1A preko ubikvitin-proteasomske poti (Huang in sod., 1998). Odstanitev domene ODD omogoča stabilnost HIF1A tudi v prisotnosti $O_2$ in posledično nemoteno heterodimerizacijo HIF1A s HIF1B, sledi vezava dimera na specifična DNA vezavna mesta in prepisovanje genov neodvisno od okolja hipoksije (Huang in sod., 1998).

V hipoksiji HIF1A uravnava prepisovanje velikega števila genov, ki omogočajo nadzor zaloge $O_2$ v celici in vzdržujejo njeno viabilnost v času nizke vsebnosti $O_2$ (Keith in sod., 2012; Vilela in sod., 2008; Wang in sod., 1995; Wenger in sod., 2005). Do sedaj je bila vloga HIF1A že omenjena pri uravnavanju metabolnih in redoks procesov (razgradnja glukoze, vloga pri metabolizmu vključno z metabolizmom lipidov), pri odzivu ožilja v hipoksiji (angiogeneza zaradi pomanjkanja dotoka krvi (ishemija)), pri raku (tumorigeneza, metastaziranje, angiogeneza tumorjev, zarodne celice raka, regulacija metabolizma raka), pri vnetnih procesih (regulacija vnetnih celic, funkcija mieloidnih celic in s tumorji povezanimi makrofagi) in poleg tega pri sistemskem odzivu na hipoksijo (pregled Majmundar in sod., 2010). Za celoten pregled v katere vse molekularne poti je vpleten HIF1, je potrebno zbrati vse njegove eksperimentalno dokazane tarče. Protein HIF1 je že dokazan regulator več sto genov, vključno z nekaj miRNA.

Povezava med okoljem hipoksije in miRNA (HRM; angl. *hypoxia-regulated miRNA*) je bila prvič predstavljena v raziskavi Kulshreshtha in sod. (2007), ki so tudi predpostavili, da imajo nekatere med njimi verjetno ključno vlogo pri preživetju celic v hipoksičnem okolju (Crosby in sod., 2009; Devlin in sod., 2011; Kulshreshtha in sod., 2008; Kulshreshtha in sod., 2007). Kljub temu, da je HIF1A dokazan regulator nekaterih miRNA, je bilo tudi za določene miRNA ugotovljeno, da kot svojo tarčo prepoznajo gen *HIF1A*. Poznanih je pet miRNA, ki z vezavo v območje 3'-UTR gena *HIF1A*, povzročijo njegovo destabilizacijo in razgradnjo: miR-20b (Cascio in sod., 2010; Lei in sod., 2009), miR-199a (Rane in sod., 2009; Rane in sod., 2010), gruča miR-17-92 (Taguchi in sod., 2008; Yan in sod., 2009), miR-519c (Cha in sod., 2010) in miR-155 (Bruning in sod., 2011). Zato je za čimboljši pregled regulatorne mreže gena *HIF1A* potrebno poleg njegovih tarč določiti tudi vse njegove regulatorje.

V številnih študijah so nove HIF1A tarče določili na podlagi izražanja genov v hipoksiji in iskanju ohranjenih vezavnih mest za HIF1A (HRE; angl. *hypoxia-response element*) znotraj promotorjev teh genov. Benita in sod. (2009) so z bioinformacijsko analizo preverili vpletenost znanih in predvidenih tarč HIF1A v biološke boti in prišli do zaključka, da HIF1A vpliva na številne metabolne poti, bioenergetiko celice in procese potrebne za razvoj in napredek raka pri človeku. Dokazano je bilo, da je izražanje nekaterih transkripcijskih dejavnikov, kot sta na primer *TWIST1* (Yang in sod., 2008) in *GATA1* (Zhang in sod., 2012), uravnavano s HIF1A, vendar je tudi *HIF1A* lahko tarča drugih transkripcijskih dejavnikov, kar je dokazano za SP1 (Minet in sod., 1999), NFKB1/RELA (Belaiba in sod., 2007) in EGR1 (Sperandio in sod., 2009). Veliko je znanega o vlogi HIF1A pri nastanku bolezni in številnih fenotipov (Majmundar in sod., 2010; Semenza, 2003), vendar je celostno znanje o HIF1A precej razpršeno med številnimi publikacijami. Za izdelavo Atlasa regulatorne mreže gena *HIF1A* je nujno zbrati vse pomembne informacije na enem mestu, jih urediti ter iz njih potegniti nove ugotovitve ključne za uravnavanje njegovega izražanja.

Hipoksija je pogosta karakteristika tumorjev in je bila do sedaj že velikokrat razlog za neuspešno zdravljenje raka (Harris, 2002). Kot že omenjeno, se celice hipoksiji prilagodijo z aktivacijo genov, potrebnih za angiogenezo, invazijo celic, metabolizem in signalne poti nujne za preživetje celic. Izsledki raziskav so pokazali, da se poleg protein-kodirajočih genov v hipoksičnem okolju celice spremeni tudi raven izražanja nekodirajoči prepisov. V dosedanjih študijah je bilo ugotovljeno tudi, da hipoksija uravnava izražanje miRNA, ki nato posredno vplivajo na signale pomembne za biologijo tumorskih celic (Chan in sod., 2009; Devlin in sod., 2011; Favaro in sod., 2010; Kulshreshtha in sod., 2007), vključno z vplivom na angiogenezo, cikel celice, popravljanje DNA in energijski metabolizem (Devlin in sod., 2011; Huang in sod., 2010).

V preteklih letih je postalo vse bolj jasno, da h kompleksnim genetskim procesom evkariotov poleg miRNA prispevajo tudi daljše (več kot 200 bp) nkRNA (pregled Mattick, 2009). Medtem ko je bil pri raziskovanju vloge in mehanizma uravnavanja izražanja z miRNA narejen velik napredek, je poznavanje vloge in mehanizma vpliva drugih nkRNA še precej omejeno. Poleg tega, da ni poznana natančna vloga drugih nkRNA v procesih celice, je potrebno dognati tudi, kateri okoljski dejavniki imajo vpliv na njihovo izražanje.

Zanimiva skupina dolgih nkRNA molekul so regije, ki so 100% identične med ortolognimi regijami genoma človeka, podgane in miši (Bejerano in sod., 2004a; Bejerano in sod., 2004b, Calin in sod., 2007). Gre za 481 zelo ohranjenih elementov genoma variabilnih dolžin (od 200 do 779 nukleotidov), po katerih se imenujejo prepisi zelo ohranjenih regij (T-UCR; angl. *transcribed-ultraconserved regions*) ali zelo ohranjeni geni (UCGs; angl. *ultraconserved genes*) (Calin in sod., 2007). Njihova presenetljivo visoka raven ohranjenosti nakazuje na pomembno biološko vlogo pri različnih odzivih celice. Spremenjeno izražanje T-UCR je bilo do sedaj že povezano s specifičnim fenotipom tumorjev predvsem vpletenost v razvoj raka. Vloga spremenjenega izražanja T-UCR je bila opisana pri kronični limfocitni levkemiji odraslih (CLL; angl. *chronic lymphocytic leukemia*), raku debelega črevesja in danke (CRC; angl. *colorectal cancer*) (Calin in sod., 2007), pri hepatocelularnem karcinomu (Braconi in sod., 2011; Calin in sod., 2007) in neuroblastomu (Mestdagh in sod., 2010; Scaruffi in sod., 2010). V genomu človeka se lokacije T-UCR pogosto prekrivajo z eksoni genov, vpletenih v izrezovanje intronov RNA, ali pa so prisotne v intronih gostiteljskih genov oziroma v bližini genov, ki so vpleteni v uravnavanje prepisovanja genov in razvoj organizma (Bejerano in sod., 2006; Bejerano in sod., 2004b). Kljub očitni tkivno-specifični vlogi UCR pri uravnavanju izražanja genov potrebnih za razvoj organizma (Bejerano in sod., 2004b; Mattick, 2009; Pennacchio in sod., 2006), je njihov mehanizem, ki privede do spremenjenega izražanja T-UCR, še nerazjasnjen. Da bi vsaj deloma razjasnili ta mehanizem, smo predpostavili, da ima zmanjšana vsebnost $O_2$ v

rakavi celici vpliv na specifično izražanje T-UCR. Preučevanje tega mehanizma smo se lotili na način, predhodno opisan pri preučevanju miRNA v hipoksiji (Kulshreshtha in sod., 2007).

## 1.5 NAMEN DELA

- Izdelati Atlas regulatorne mreže transkripcijskega dejavnika *HIF1A*, ki bo vseboval njegove regulatorje izražanja na ravni prepisovanja in po njem, in bo vključeval do sedaj znane tarče tega transkripcijskega dejavnika.

- Ugotoviti funkcijsko povezavo med okoljem hipoksije in zelo ohranjenimi regijami genoma. Želeli smo ugotoviti vlogo T-UCR pri prilagoditvi celic na hipoksično okolje in način kako le-ta omogoča delovanje in preživetje rakave celice.

## 1.6 RAZISKOVALNE HIPOTEZE

V okviru raziskave bomo preizkusili naslednje hipoteze:

- Hipoksija vpliva na spremenjeno izražanje zelo ohranjenih regij (UCR).

- Transkripcijski dejavnik HIF neposredno uravnava izražanje UCR-jev.

- Hipoksija preko transkripcijskega dejavnika HIF vpliva na spremenjeno izražanje UCR-jev, kar se odraža v proliferaciji celičnih linij raka.

- Utišanje prekomernega izražanja UCR v hipoksičnem okolju zavira proliferacijo celičnih linij raka.

# 2 LITERATURE REVIEW

## 2.1 HYPOXIA AND CANCER

Cancer is a generic term for a large group of diseases that can affect any part of the body. Other terms used are malignant tumors and neoplasms. There are over 100 different types of cancer, and each is classified by the type of cell that is initially affected. One defining feature of cancer is the rapid creation of abnormal cells that grow beyond their usual boundaries, and which can then invade adjoining parts of the body and spread to other organs. Tumors can grow and interfere with the digestive, nervous, and circulatory systems and they can release hormones that alter body function.

Tumors that stay in one spot and demonstrate limited growth are generally considered to be benign. More dangerous, or malignant, tumors form when two things occur: 1) a cancerous cell manages to move throughout the body using the blood or lymph systems, destroying healthy tissue in a process called invasion; 2) that cell manages to divide and grow, making new blood vessels to feed itself in a process called angiogenesis. When a tumor successfully spreads to other parts of the body and grows, invading and destroying other healthy tissues, it is said to have metastasized. This process itself is called metastasis, and the result is a serious condition that is very difficult to treat. Cancer arises from one single cell, but the transformation from a normal cell into a tumor cell is a multistage process, typically a progression from a pre-cancerous lesion to malignant tumors. These changes are the result of the interaction between a person's genetic factors and external agents/environment.

The paradigm that cancer is a cellular disease that is defined only by events within the genomes of cancer cells has given way in recent years to one in which cancer is viewed as an ecological disease involving a dynamic interplay between malignant and non-malignant cells. This shift redirects attention to the tumor microenvironment, which encompasses signals, proteins and cells (immune cells, fibroblasts) present in the tumor mass that are necessary for tumor growth and progression (Barcellos-Hoff et al., 2013). Therefore, mutations in cancer cells alone are not able to induce tumor growth, so active changes in the surrounding tissue cells must be essential for the clinical development of cancer (Barcellos-Hoff et al., 2013). Consistent with this idea, the initiation of tumorigenesis by the expression of an oncogene or deletion of tumor suppressor genes in mouse tissue is remarkably inefficient at generating cancers (Barcellos-Hoff et al., 2013). Oncogenic changes in tumor cells can moderate angiogenesis and so influence the maturation of the cancer niche into the tumor microenvironment. Just the ability of the tumor to generate pre-metastatic niche is likely to be a major factor that determines metastatic spread (Psaila and Lyden, 2009). It is speculated that niche-forming interactions affect which transformed cells survive, expand and progress to clinical

disease, and thus determine cancer incidence (Barcellos-Hoff et al., 2013). The multiple cellular candidates that might establish a cancer niche early in cancerogenesis all serve the same purpose: to provide a suitable environment for survival of the cancer cells (Barcellos-Hoff et al., 2013).

Hypoxia as an environmental factor is usually defined as lower than 2% $O_2$ level, while severe hypoxia (or anoxia) is defined as lower than 0.02% $O_2$ level. However in most mammalian tissue level of $O_2$ ranges from 2-9% (40 mm Hg), meanwhile ambient air contains 21% of $O_2$ (150 mm Hg). The beginnings of hypoxia research in tumor biology can be tracked back to observations made in the early 20$^{th}$ century by Otto Warburg, who demonstrated that, unlike normal cells, tumor cells favor glycolysis independent of cellular oxygenation levels (Warburg, 1956). There is number of other molecular mechanisms promoting "aerobic glycolysis" already proposed, including mutations and epigenetic mechanisms in genes encoding tumor suppressors, activation of oncogenes, and hypoxic adaptations (Denko, 2008). It is still unclear how this "aerobic glycolysis" offers tumor cells a growth advantage. New data suggests that this metabolic switch may provide benefit to the tumor by decreasing mitochondrial activity and not by increasing glycolysis (Denko, 2008).

Hypoxia has been already known as an essential feature of the neoplastic microenvironment (Harris, 2002). When tumors' diameters are longer than 2–3 mm, the center of the tumor's microenvironment will become hypoxic (Folkman et al., 1971). Tumor vascularization supplies nutrition and $O_2$ to proliferating cells, that is why cellular adaptation to hypoxia strongly correlates with the risk of cancer cell invasion and metastasis (Dang and Semenza, 1999). Consequentially cells in hypoxic conditions continue to proliferate in decreased $O_2$ level, but the environment forces them to alter transcription and translation of genes involved in angiogenesis, cell invasion, cell metabolism, and cell survival. Tumors with an extensive low $O_2$ environment are also more invasive and resistant to conventional therapy (Harris, 2002). Uncontrolled proliferation of cancer cells often results in hypoxia in cancer cell masses, and indeed up to 50–60% of solid cancers contain hypoxic tissue areas (Rankin and Giaccia, 2008). As a result, HIF1A is over-expressed in a majority of human cancers and for some of them an indicator of unfavorable prognosis (Birner et al., 2000; Osada et al., 2007; Park et al., 2009; Schindl et al., 2002; Shibaji et al., 2003; Talks et al., 2000).

A near-universal property of primary and metastatic cancers is upregulation of glycolysis, resulting in increased glucose consumption, which can be observed with clinical tumor imaging. Upregulation of glycolysis leads to microenvironmental acidosis requiring evolution of phenotypes resistant to acid-induced cell toxicity. Subsequent cell populations with upregulated glycolysis and acid resistance have a powerful growth advantage, which promotes unconstrained proliferation and invasion (Gatenby and Gillies, 2004). Hypoxia research has in the past century provided

significant advancement in our understanding of the molecular pathways responsible to give tumor cells growth advantage. Better understanding and further knowledge of these mechanisms will contribute to development of new strategies in systemic cancer therapies. Hypoxia induced factor 1 (HIF1) has been as transcription factor (TF) implicated in regulating many of genes that are responsible for metabolic difference. Tumors derived from cells lacking HIF1A or HIF1B show significantly reduced vascularization and growth rates compared to parental cells (Kung et al., 2000; Ryan et al., 2000).

## 2.2 HIF

Hypoxia-inducible factors (HIFs) belong to a family of environmental sensors known as bHLH–PAS (basic helix-loop-helix–Per-Arnt-Sim) transcription factors (Gu et al., 2000), which regulate diverse biological processes. The HIF1 heterodimer consists of an alpha-subunit (HIF1A) and a beta-subunit (HIF1B; also known as the aryl hydrocarbon receptor nuclear translocator (ARNT)). Both HIF1A and ARNT are bHLH–PAS transcription factors that contain two PAS domains of 100–120 amino acids, designated PAS-A and PAS-B, which are necessary for heterodimerization and DNA binding (Simon and Keith, 2008) (Figure 1) and two transactivation domains (N-terminal transactivation domain or N-TAD and C-terminal transactivation domain or C-TAD) (Kaelin and Ratcliffe, 2008). A bHLH domain is a conserved DNA binding domain found in a number of transcription factors, while PAS is a protein–protein dimerization domain that is related to the conserved signal sensing protein motifs in the *Drosophila melanogaster* period (PER), the mammalian aryl hydrocarbon receptor nuclear translocator (ARNT) and *D. melanogaster* single-minded (SIM) proteins (Keith et al., 2012). PAS domains play important roles as sensory modules for $O_2$ tension, redox potential or light intensity. In response to stimuli, the domain either mediates protein-protein interactions or binds co-factors within their hydrophobic cores to regulate protein-protein interactions. In response to hypoxic conditions, the PAS-B domain of HIF-A heterodimerizes with PAS-B of ARNT that is involved in transcriptional activation (Figure 2).


Figure legend: bHLH, basic helix-loop-helix; PAS, PER-ARNT-SIM; TAD, transactivation domain; ODD, $O_2$-dependent domain.

Figure 1: Subunits of HIF proteins (Simon and Keith, 2008)

Slika 1: Podenote proteinov HIF (Simon in Keith, 2008)

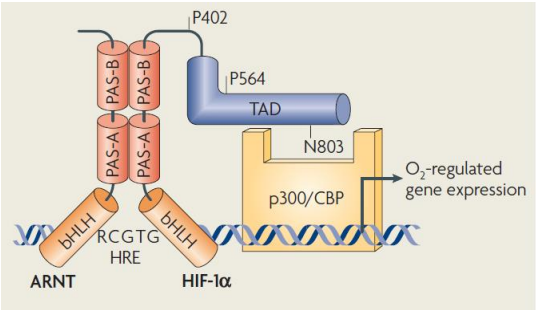



Figure legend: HRE, hypoxia response element; bHLH, basic helix-loop-helix; PAS, Per-Arnt-Sim;TAD, transactivation domain.

Figure 2: HIF1A transcriptional activation of its target genes (Simon and Keith, 2008)

Slika 2: Aktivacija prepisovanja tarčnih genov proteina HIF1A (Simon in Keith, 2008)

Semenza and colleagues first described HIF1A in 1995, when it was shown to have a central role as a mediator of $O_2$-dependent transcriptional responses in hypoxia (Wang et al., 1995). In 1997, the identification by independent groups of HIF2A − which was initially called endothelial PAS protein 1 (EPAS1) (Tian et al., 1997), HIF-related factor (HRF) (Flamme et al., 1997), HIF1A-like factor (HLF) (Ema et al., 1997) or member of PAS family 2 (MOP2) (Hogenesch et al., 1997) − indicated that HIF regulation was more complex. In mammals, three genes have been shown to encode HIF-A subunits that appeared to be similarly regulated by $O_2$ availability; however HIF1A and HIF2A appeared to be most extensively characterized. HIF1A is expressed ubiquitously, whereas HIF2A and HIF3A are found in a subset of tissues (Simon and Keith, 2008). Both, HIF1A and HIF2A, binds DNA at specific locations termed HREs containing the core sequence 5'-G/ACGTG-3', and upregulates gene expression of their target genes. By contrast, HIF3A acts as a dominant negative regulator of HIF1A and HIF2A-mediated transcription.

Unlike the HIF-A proteins, HIF-B is constitutively expressed and insensitive to changes in $O_2$ levels (Bertout et al., 2008). However, HIF-A genes are transcribed and translated at a high rate, but HIF-A proteins are rapidly degraded in the presence of sufficient $O_2$ levels. Under normoxic conditions, HIF-A subunit stability is largely regulated through the region of 200 amino acids (aa), characterized as ODD domain (403–602 aa, Figure 1) (Huang et al., 1998). In cells with normal $O_2$ level (normoxia) HIF1A stability is regulated through prolyl- and asparaginyl-hydroxylases. Within ODD domain HIF-A is hydroxylated on two conserved proline residues, *p.Pro402* and *p.Pro564* (Ivan et al., 2001; Jaakkola et al., 2001; Masson et al., 2001; Yu et al., 2001), by a family of three HIF-specific prolyl hydroxylases: PHD1, PHD2 and PHD3 (EGLN1, 2, 3) (Figure 3) (Bruick and McKnight, 2001; Ivan et al., 2002; Percy et al., 2003; Semenza, 2007b). Hydroxylated HIF-A proteins are recognized by the pVHL tumor suppressor gene product (a component of a multisubunit ubiquitin-ligase complex), then covalently

tagged with polyubiquitin and subsequently degraded by the 26S proteasome (Kallio et al., 1999; Schofield and Ratcliffe, 2004). Asparaginyl-hydroxylases (HIF1AN or FIH1) hydroxylate HIF1A at asparagine residue (*p.Asn803*) in the C-TAD and thus prevents interaction of HIF1 with transcriptional co-activators, which inhibits HIF1 transcriptional activity. In cells with low $O_2$ level (hypoxia) all these processes are inhibited and HIF1A subunit is stabilized, co-factor recruitments is enabled and HIF1A downstream target genes are transcribed (Bertout et al., 2008).

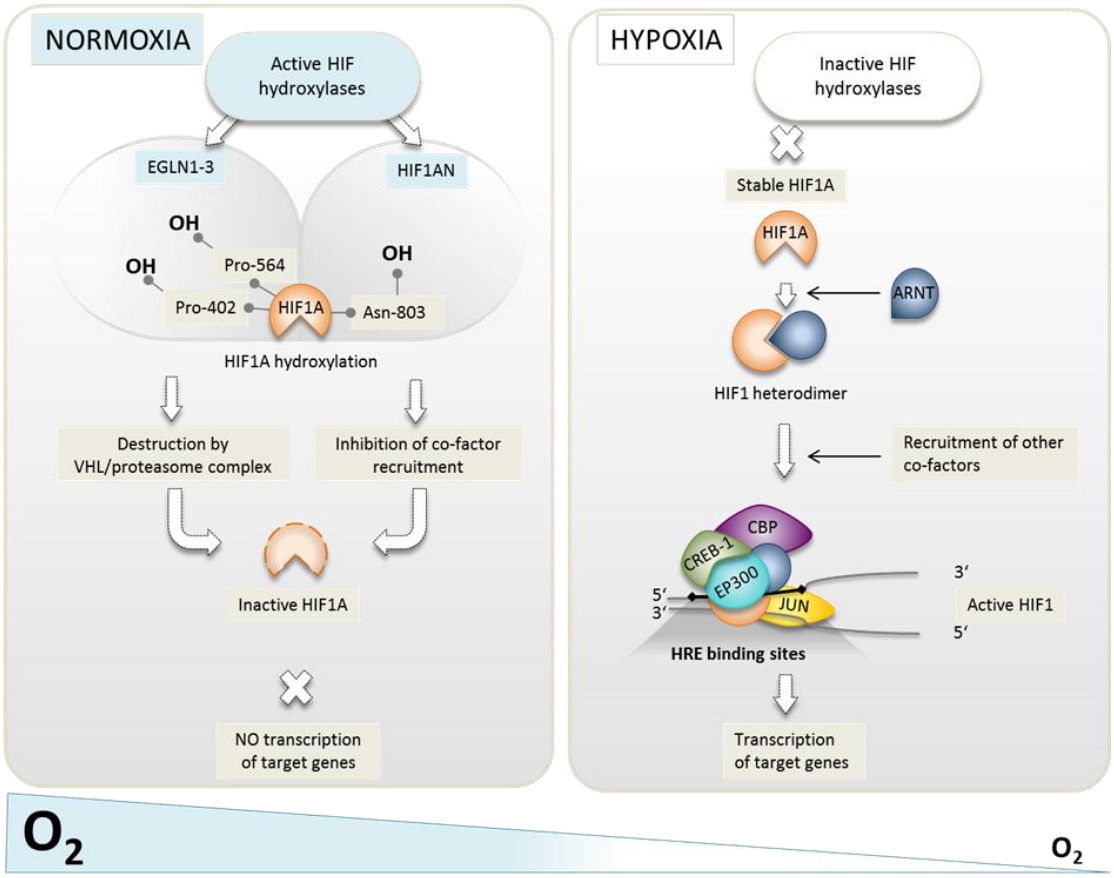



Figure legend: EGLN1-3, egl nine homolog 1-3; HIF1AN, HIF1A subunit inhibitor; ARNT, aryl hydrocarbon receptor nuclear translocator; EP300, E1A binding protein p300; JUN, jun proto-oncogene; CREB1, cAMP responsive element binding protein 1; CBP, CREB binding protein; HRE, hypoxia response element.

Figure 3: HIF1A post-translational regulation depending on the level of $O_2$. Figure was modified according to (Bertout et al., 2008)

Slika 3: Potranslacijsko uravnavanje proteina HIF1A v odvisnosti od prisotnosti $O_2$ (prirejeno po (Bertout in sod., 2008))

Under normoxic conditions, the half-life of HIF1 is less than one minute (Yu et al., 1998), but when subjected to low $O_2$ tension (< 30 mm Hg), HIF1 protein levels increase exponentially (Jiang et al., 1996). Under hypoxic conditions (lower than 2%

$O_2$), hydroxylation of the ODD of HIF-A and the interaction with pVHL is inhibited. The subunits of HIF-A therefore accumulate in the cytoplasm of $O_2$-starved cells and translocate to the nucleus, where they dimerized with HIF-B through their PAS domains (Figure 2). This interaction is thought to be required for full HIF activity, but is also regulated by normoxic hydroxylation reactions; in this case at *p.Asn803* (Lando et al., 2002a). Asparagine hydroxylation is carried out by factor inhibiting HIF (FIH), and FIH activity is inhibited under hypoxic conditions in a manner that is reminiscent of the prolyl hydroxylases (Lando et al., 2002b; Mahon et al., 2001) (Figure 3).

However, HIF1A-HIF1B dimers bind to HREs, which contain the core recognition sequence 5′-RCGTG‐3′ (where 'R' denotes a purine residue)) that are located within the promoters, introns and 3′ enhancers of a large number of $O_2$-regulated target genes. During this process, the C-TAD also interacts with co-activators such as p300/CBP (cyclic-AMP responsive element binding (CREB)-binding protein) (Ema et al., 1999).

## 2.3 HIF AND CANCER

On the basis of the human cancer biopsy samples analysis and experimental animal models, it has become clear that HIF has a crucial role in cancer progression. Overexpression of HIF1A was reported in broad range of human malignancies, because of genetic or environmental reasons associated with hypoxia (Maxwell et al., 2001). The accumulation of HIF1A has been associated with poor patient survival in early-stage cervical cancer, breast cancer, oligodendroglioma, ovarian cancer, oropharyngeal squamous cell carcinoma, CRC, and renal cell carcinoma (RCC) (Audenet et al., 2012; Bertout et al., 2008; Cao et al., 2009; Kim et al., 2012). Interestingly, HIF1A overexpression was associated with decreased patient mortality for head and neck cancer (Beasley et al., 2002) and non-small-cell lung cancer (Volm and Koomagi, 2000), for type of cancers where increased patient mortality was reported to be associated with HIF2A overexpression (Giatromanolaki et al., 2001; Koukourakis et al., 2002). These studies suggest that these two A-subunits can have opposing effect on disease progression depending on tumor type. Consecutively this has spurred a new interest in understanding the molecular differences between the A-subunits and has also demonstrated that the relative contributions of the A-subunits to tumor growth *versus* suppression are likely to be tissue specific (Bertout et al., 2008).

Studies focusing on HIF1A deficiency resulted in delayed tumor onset, reduced tumor growth and rather fewer tumor blood vessels. However, even in the number of metastases to the lung was significantly decreased. These results indicate that HIF1A has an unexpected role potential in metastasis that was additionally validated by *in vivo* experiments (Hiraga et al., 2007). Increased HIF1A expression can also be caused by a variety of genetic alterations that activate oncogenes (ERBB2 in SRC) or inactivate

tumor suppressor genes (VHL, PTEN, CDKN2A) (Semenza, 2003). One of the first links between tumor hypoxia and a specific cancer genetic program was that with TP53. HIFs also affect the MYC pathway, where HIF1A opposes MYC (Koshiji et al., 2004), as a potent regulator of proliferation and anabolic metabolism, meanwhile HIF2A promotes MYC activity (Gordan et al., 2007). Because MYC is commonly deregulated in cancer, this is also an important interaction to consider when developing anticancer therapies (Bertout et al., 2008).

### 2.3.1 Effect of *HIF1A* variations on cancerous and non-cancerous phenotypes

Several single nucleotide polymorphisms (SNPs) have been identified in *HIF1A*. However, the density of SNP appears particularly high in exon 12, which includes synonymous and non-synonymous SNPs (Shieh et al., 2010). These non-synonymous SNP result, respectively, in amino acid changes from proline 582 to serine (*p.Pro582Ser;* rs11549465) and from alanine 588 to threonine (*p.Ala588Thr;* rs11549467). Exon 12 translates a portion of HIF1A encompassing the N-TAD domain and the partial ODD domain, and has been reported to increase HIF1A expression, even under normoxia (Fu et al., 2005; Tanimoto et al., 2003). The association between rs11549465 or rs11549467 and the risk or progression of various cancers like carcinoma of the cervix, endometrium, colorectum, esophagus, bladder, and kidney has been reported (Fransen et al., 2006; Konac et al., 2007; Ling et al., 2005). Both HIF1A variations have been demonstrated to increase HIF transcription activator function and effect on normal physiological function of HIF1A including increase of HIF1A stability (Fu et al., 2005; Hebert et al., 2006; Tanimoto et al., 2003). Factors inhibiting the HIF1 (FIH1 or HIF1AN) protein interacts with amino acid 531–826 of HIF1A, which includes the rs11549465 and rs11549467 polymorphisms (Mahon et al., 2001). Both polymorphisms might decrease the binding affinity between FIH1 and HIF1A and increase the transcription activator function of HIF1 protein (Shieh et al., 2010).

To date many *HIF1A* SNP association studies were performed and various genotype-phenotype associations were identified. Previous studies performed on genetically engineered *Hif1a* knockout (KO) mouse already indicated the importance of HIF1 pathway in the metabolic control of muscle function (Mason et al., 2004), heart function (reductions in contractility, vascularization, high-energy phosphate content, and lactate production) (Huang et al., 2004), the essential role of *Hif1a* in controlling both embryonic and tumorigenic responses to variations in micro-environmental oxygenation (Ryan et al., 1998), and effect on immunological system, where chimeric mice model showed abnormal B lymphocyte development and autoimmunity (Kojima et al., 2002; Kojima et al., 2004). Moreover, Kline et al (Kline et al., 2002) disclosed that partial *Hif1a* deficiency has a dramatic effect on carotid body neural activity and ventilatory adaptation to chronic hypoxia. Chronic hypoxia in laboratory animals (mouse or rat)

was shown to induce multiple systemic responses similar to physiological responses in human, like polycythemia, pulmonary hypertension, right ventricular hypertrophy, and weight loss (DiCarlo et al., 1995; Rabinovitch et al., 1979; Yu et al., 1999).

Studies showed that the occurrence of local hypoxia in the muscle causes a drop in $O_2$ pressure within the monocyte during exercise (Richardson et al., 2001; Richardson et al., 1995). This causes an induction of the HIF1A-mediated signaling pathway in human skeletal muscle, thereby provoking a shift toward increased use of oxidative pathways for energy production (Ameln et al., 2005; Hoppeler and Fluck, 2003). Transcriptional events largely contribute to increases in mitochondrial mass in human skeletal muscle with endurance training. Expression of mitochondrial proteins from the nuclear and mitochondrial genomes is coordinated and involves the nuclear-encoded transcription factors NRF-1 and TFAM. Transcription of genes encoding the mitochondrial proteins involved in beta oxidation can be regulated separately from the genes of the Krebs cycle and the respiratory chain. Transcription factors AP-1 and PPAR alpha/gamma and the protein kinase AMPK are signaling molecules that transduce the metabolic and mechanical factors sensed during endurance training into the complex transcriptional adaptations of mitochondrial proteins (Hoppeler and Fluck, 2003). Furthermore, a tissue-specific loss of function of HIF1A in mice led to a decrease of exercise-induced changes in gene expression of glycolytic enzymes in skeletal muscle (Mason et al., 2004) and caused reductions in contractility and vascularization of the heart (Huang et al., 2004). The muscle damage in *Hif1a* KO mouse is similar to that detected in humans in diseases caused by deficiencies in skeletal muscle glycogenolysis and glycolysis. Thus, these results demonstrate an important role for the HIF1-pathway in the metabolic control of muscle function (Mason et al., 2004).

## 2.4 HYPOXIA AND NONCODING RNAs

### 2.4.1 Regulatory ncRNAs

Controlling protein-coding gene expression can no longer be attributed purely to proteins involved in transcription, RNA processing, and translation. The role that ncRNAs play as potent and specific regulators of gene expression is now widely recognized in almost all species studied to date (Brosnan and Voinnet, 2009). Noncoding RNAs are defined as RNAs that do not encode proteins (Eddy, 2001; Mattick and Makunin, 2006) and until a decade ago, ncRNA genes were invisible even by complete genome sequencing. A central dogma of molecular biology is that RNA plays the role of intermediate messenger between the gene and the final protein encoded by this gene (Eddy, 2001; Kelley and Kuroda, 2000; Mattick and Makunin, 2006). Scientists discovered prevalent roles for ncRNA that were based on specific nucleic acid recognition without any catalytic function, like ribosomes (Eddy, 2001).

Genes of ncRNA produce transcripts that function as structural, catalytic, or regulatory RNAs (Brosnan and Voinnet, 2009; Eddy, 2001; Pang et al., 2007). Some groups of ncRNAs have been known for a long time and have well-known functions in cell life. Some of them are involved in RNA translation mechanisms (ribosomal RNA (rRNA) and transfer RNA (tRNA)) or in splicing (small nuclear RNA (snRNA)), and some are known to have the ability to modify other groups of ncRNA, such as small nucleolar RNAs (snoRNA) that modify rRNA (Mattick and Makunin, 2006). Many ncRNAs have to be alternatively spliced or/and processed into smaller RNA molecules like miRNA to become functional (Mattick and Makunin, 2006). In general, RNA molecules have wide biological functions on all levels of gene regulation, mainly influencing transcription initiation, stability of messenger RNA (mRNA), RNA editing, splicing, and even control of chromosome dynamics (*i.e.,* maintenance and segregation). RNA molecules might have influence on cell cycle regulation and were found to function in response to stress (hypoxia); furthermore, they contribute to cell death or apoptosis (Allen et al., 2004; Espinoza et al., 2004; Haaland et al., 2005; Kulshreshtha et al., 2007; Mattick and Makunin, 2006). The regulatory function of ncRNA mainly relies on base-pairing with complementary sequences of RNA from other RNA classes or of DNA from promoter regions, so they can form RNA::RNA or RNA::DNA complexes (Mattick and Makunin, 2006).

Recently, tiling array experiments (*i.e.,* arrays with probes designed to cover the entire genome) suggested that the majority of the genomes of mammals and other complex organisms are actually transcribed into ncRNA and not into mRNAs. In last few years, the number of identified ncRNAs has increased very quickly and is still rising suggesting that our understanding of the genome structure and components of functional networks is far from complete (Costa, 2007; Guttman et al., 2009; Kozomara and Griffiths-Jones, 2011; Mattick and Makunin, 2006).

In general ncRNA are transcribed in long and short RNA transcripts (Table 1). Long ncRNAs were found to upregulate and downregulate gene expression in both eukaryotes and prokaryotes and are essential in processes such as dosage compensation, genomic imprinting, developmental patterning and differentiation, and stress response. Meanwhile, small ncRNAs also play essential roles in diverse organisms, although are limited to eukaryotes (Brosnan and Voinnet, 2009). Small RNAs control a wide range of developmental and physiological pathways in animals, including hematopoietic differentiation, adipocyte differentiation and insulin secretion in mammals, and have been shown to be perturbed in cancer and other diseases (Mattick and Makunin, 2005).

Table 1: Classes of noncoding RNAs (modified from (Taft et al., 2010; Wang and Chang, 2011))

Tabela 1: Razredi nekodirajočih RNA (prirejeno po (Taft in sod., 2010; Wang in Chang, 2011))

| **sncRNAs** Small ncRNAs (< 200 nt) | | | **lncRNAs** Long ncRNAs (> 200 nt) |
|---|---|---|---|
| **siRNAs** Small interfering RNAs | **miRNAs** microRNAs | **snoRNAs** small nucleolar RNAs | Signal lncRNAs |
| **piRNAs** PIWI-interacting RNAs | **PARs** Promoter-associated RNAs | **xiRNAs** X-inactivation RNAs | Decoy lncRNAs |
| **sdRNAs** Sno-derived RNAs | **moRNAs** microRNA-offset | **tel-sRNAs** Thelomere small RNAs | Guide lncRNAs |
| **MSY-RNAs** MSY2-associated RNAs | **crasiRNAs** Centrosome-associated RNAs | / | Scaffold lncRNAs |

As ncRNAs represent a very heterogeneous group, they can be sub grouped in different classes based on their length and finally their function: (i) miRNAs; (ii) other small ncRNAs (such as snoRNA, snRNA, piwi-interacting RNA (piRNA)); and (iii) long ncRNAs (lncRNAs). The most thoroughly studied class of ncRNAs is by far the group of miRNAs (Ambros, 2008). However, large numbers of lncRNAs with little or no protein-coding potential are being identified in eukaryotes. Increasing data describing the expression profiles, molecular features and functions of individual lncRNAs in a variety of systems are accumulating. For systematic compilation and updating information about lncRNA, a database (lncRNAdb) containing a comprehensive list of lncRNAs was developed (Amaral et al., 2011). Besides, long noncoding transcripts, such as T-UCRs, were also cloned, but their functions in normal cells and abnormalities in cancer cells have not been fully explored (Calin et al., 2007).

#### 2.4.1.1 Short ncRNAs

The most widely studied and very fast evolving group among short ncRNA are miRNAs. They represent a group of endogenous ncRNA molecules that modulate post-transcriptional gene regulation. Their size range from 18 nucleotides to approximately 25 nucleotides and according to many reports miRNAs control expression of at least one-third of human protein-coding genes (Lewis et al., 2005; Xie et al., 2005). They exert their regulatory role on protein-coding gene expression by binding to either full or partial complementary sequences primarily in the 3′-untranslated region (UTR), but also inside the coding DNA sequence (CDS) and the 5′-UTR, of the corresponding mRNAs (reviewed in (Kunej et al., 2012)). Eventually, this affects mRNA stability and translation (Filipowicz et al., 2008; Miranda et al., 2006; Tay et al., 2008).

Biogenesis of miRNA begins in the cell nucleus, where it is derived from a long primary miRNA molecule (pri-miRNAs) and which is further processed in cytoplasm

from hairpin structured precursor miRNAs (pre-miRNA) to mature miRNAs (Mattick and Makunin, 2005) (Figure 4). Post-transcriptional regulation due to miRNA can occur either by non-perfect pairing, which causes suppression of transcription, or by perfect-match pairing, which provokes degradation of target mRNAs (Du and Zamore, 2005). Single-stranded mature miRNAs are recognized by the multi-protein infrastructure known as RNA-induced silencing complex (RISC) and guided to target regions on mRNA (Mattick and Makunin, 2006). MicroRNAs roles of gene expression regulations includes cell-cycle regulation, proliferation, cell differentiation, apoptosis, stress response, and fat metabolism (Ambros, 2003; Lagos-Quintana et al., 2001; Mattick and Makunin, 2006; Sevignani et al., 2006).

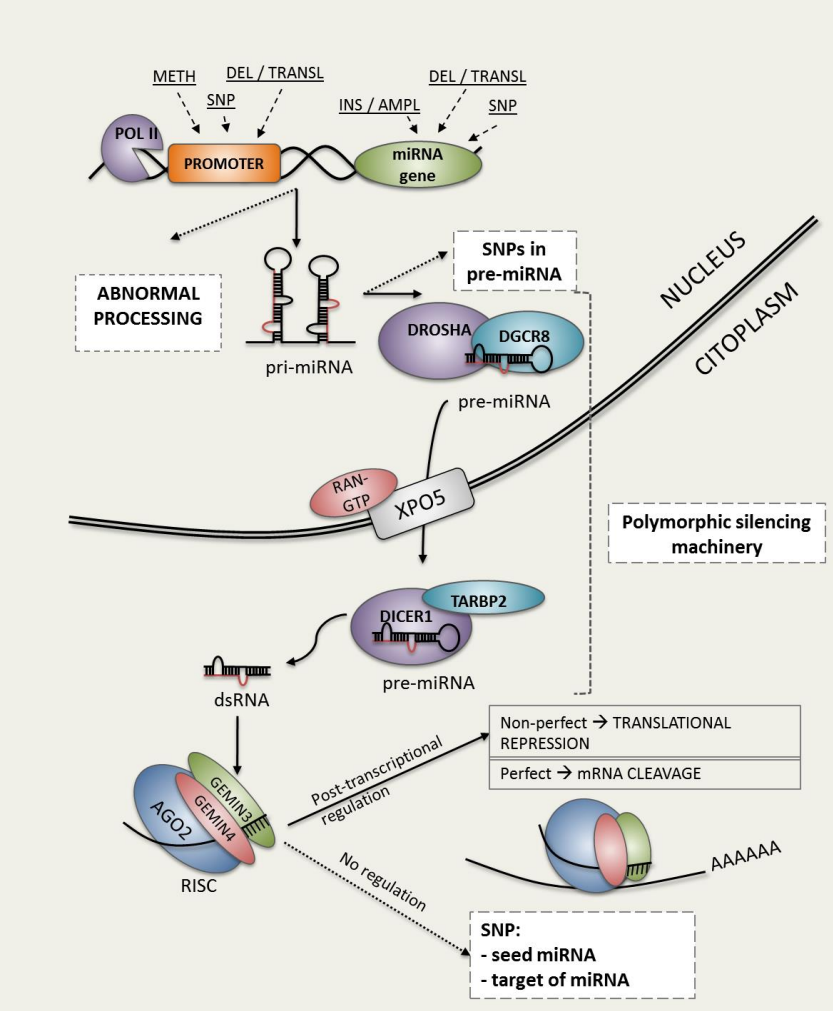



Figure legend: METH, methylation; SNP, single nucleotide polymorphism; DEL, deletion; TRANSL, translocation; INS, insertion; AMPL, amplification; MUT, mutation; POL II, RNA polymerase II; AAAAAA, polyA tail; RISC, RNA-induced silencing complex.

Figure 4: MicroRNA biogenesis and critical points that can cause miRNA deregulation and cancer predisposition when mutated (Ferdin et al., 2011)

Slika 4: Proces zorenja miRNA in kritične točke, ki lahko povzročijo spremembo uravnavanja miRNA in dovzetnost za nastanek raka v primerih mutacij (Ferdin in sod., 2011)

However, since the discovery of miRNA and the explosion in miRNA research, this field has quickly moved from basic molecular research to clinical applications. In clinical applications miRNAs are used as markers to classify cancers, as prognostic indicators, or even as predictors of the efficacy of cancer therapies (Calin et al., 2005; Ferracin et al., 2010). They can act either as oncogenes or tumor-suppressor genes, while some of them have both functions depending on the cancer type and cellular context (Figure 5). Moreover, miRNAs were believed to function as hormones or body fluids circulation biomarkers for cancers (reviewed in (Cortez et al., 2011).

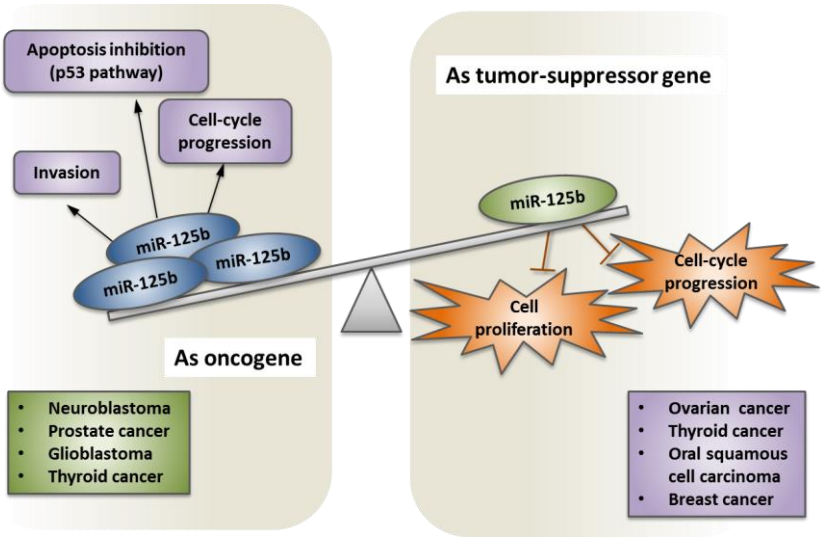



Figure 5: MicroRNA with a role of oncogene in one cancer type and tumor-suppressor gene in different cancer type (Cortez et al., 2011)

Slika 5: Mikro RNA z vlogo onkogena pri enem tipu raka ter z vlogo tumor-supresorja pri drugem tipu raka (Cortez in sod., 2011)

Studies of miRNA expression revealed tissue- or cell-specific expression patterns, and differences between normal and cancer tissue of same cellular origin (Barbarotto et al., 2008; Calin and Croce, 2006; Nicoloso et al., 2009; Spizzo et al., 2009). Changes in miRNA expression profiles, in their specific sequences, and in binding sites in mRNA targets have been identified in a variety of human genetic diseases, including cancer, heart disease, diabetes, and neural disorders (Costa, 2007; Ikeda et al., 2007; Krichevsky et al., 2003; Makeyev et al., 2007; Mattick and Makunin, 2006; Moreau et al., 2013; Schratt et al., 2006). The causes of the altered miRNAs expression in malignant compared to normal cells can be explained by the location of miRNA genes within cancer-associated genomic regions (CAGRs), by epigenetic mechanisms and by alterations in the miRNA processing machinery (Calin and Croce, 2006) (Figure 4 and Figure 6).

When researchers combined miRNA locations with previously reported genetic alterations they concluded that > 50% of the studied miRNAs located in CAGRs or/and

in chromosomal fragile sites (FRAs). Chromosomal changes such as substitutions, insertions, and deletions that were associated with altered miRNA expression were found to be principally located in the promoter regions of individual miRNA genes or their gene clusters (Sevignani et al., 2007). Most recently, SNPs in the genes encoding miRNAs, the miRNA machinery, and miRNA targets were shown to interrupt normal cellular activity, and these SNPs can also contribute to an increased susceptibility for cancer development, poor patient prognosis, and reduced treatment efficacy (Chin et al., 2008; Frazer et al., 2007; Saunders et al., 2007; Tay et al., 2008) (presented on Figure 4 and Figure 6). As an example, Nicoloso et al. (2010) study proposed that transcribed SNPs, which were known to be associated with breast cancer risk, altered miRNA binding sites recognition and consequently miRNA gene regulation. Moreover, miRNA genes might also be silenced in human tumors by aberrant hypermethylation of CpG (-C-phosphate-G- or CG site) islands that encompass or lie adjacent to miRNA (reviewed in (Kunej et al., 2011)).

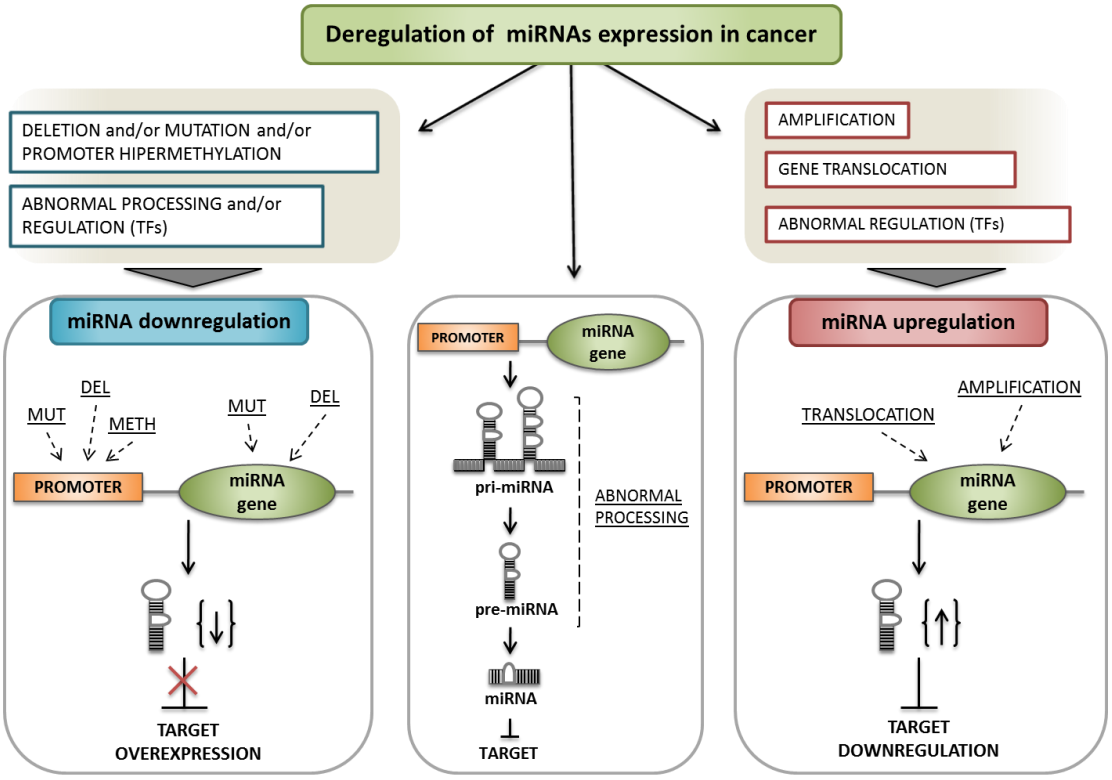



Figure legend: TF, transcription factor; MUT, mutation; DEL, deletion; METH, methylation.

Figure 6: Mechanisms responsible for deregulation of miRNAs expression in cancer (Ferdin et al., 2010)

Slika 6: Mehanizmi odgovorni za spremenjeno izražanje miRNA pri raku (Ferdin in sod., 2010)

Most of research on miRNA in cancer are focusing on creating a diagnostic tool for cancer classification and a possible prognostic tool to define disease stage and progression (Calin et al., 2005; Lee and Dutta, 2009; Lu et al., 2005). To date there are

many miRNA profiling methods available to detect different miRNA signatures in a variety of cancers in which miRNA expression patterns change because of tissue-specific signaling pathways or developmental effects (reviewed in (Ferdin et al., 2010).

While significant progress has been made towards understanding the mechanisms of gene regulation by miRNAs, much less is known about factors affecting the expression of these and others noncoding transcripts. Description for other small RNAs (snoRNA, snRNA, and piRNA) also called non-miRNA small RNAs are described in the following references (Amaral and Mattick, 2008; Castel and Martienssen, 2013; Mattick and Makunin, 2005; Mattick and Makunin, 2006).

#### 2.4.1.2 Long ncRNAs

During the past few years, it has become clear that long ncRNAs (> 200 bp in length) also participate in complex genetic events in eukaryotes (reviewed in (Mattick, 2009)). Initially, lncRNAs were considered to be background noise associated with RNA polymerase activity (Mercer et al., 2009; Struhl, 2007). Further detailed investigations, however, have proven lncRNA involvement in fundamental cellular processes such as RNA processing, gene regulation, chromatin modification, gene transcription (controlling the RNA polymerase II activity taking part as co-factors in RNA-protein complexes), post-transcriptional gene regulation (splicing, editing, transport, translation, and degradation) based on RNA sequence complementarity interactions, and other functions yet to be discovered (Bejerano et al., 2004b; Feng et al., 2006; Lareau et al., 2007; Mercer et al., 2009; Pennacchio et al., 2006; Visel et al., 2008). One intriguing family of lncRNAs (Calin et al., 2007) is transcribed from regions that exhibit extremely high conservation between the orthologous regions of human, rat, and mouse genomes (Bejerano et al., 2004a; Bejerano et al., 2004b), hence the term "transcribed-ultraconserved regions" (T-UCRs) or ultraconserved genes (UCG) (Calin et al., 2007).

Researchers became aware of UCRs when alignment of human, rat, and mouse genomes showed that, despite 300 million years of negative selection, there were genomic regions that remained highly conserved (100% identity) (Bejerano et al., 2004b; Mercer et al., 2009). In 2004, Bejerano presented 481 of those UCRs, defined as genomic sequences longer than 200 bp and perfectly conserved (Bejerano et al., 2004b). Although several other sets of conserved elements were identified, such as conserved non-genic sequences (CNSs) (Mouse Genome Sequencing et al., 2002), long conserved noncoding sequences (LCNS) (Sakuraba et al., 2008), multiple species conserved sequences (MCSs) (Thomas et al., 2003) and highly conserved regions (HCRs) (Duret et al., 1993), but UCRs are regarded as a class of genomic elements with the strongest level of conservation. In fact, UCRs are even more conserved than coding genes and are

now believed to have fundamental functions in vertebrate evolution, including that of mammals (Ban et al., 2005; Bejerano et al., 2004b; Elgar and Vavouri, 2008).

Although originally only 225 (47%) of the UCRs were classified as transcribed (exonic) or possibly transcribed (possibly exonic) by overlapping with coding exons (Bejerano et al., 2004b), an experimental analysis evidenced the transcription of 325 (68%) UCRs and called them transcribed UCRs (Calin et al., 2007). Transcribed UCRs also showed tissue-specific expression and their alteration was already associated with cancer, suggesting that UCRs may represent yet another class of functional ncRNAs (Calin et al., 2007). Recent studies suggest that UCRs play important roles in genome regulatory mechanisms such as acting as distal enhancers (Paparidis et al., 2007; Pennacchio et al., 2006), regulating alternative splicing (Lareau et al., 2007; Ni et al., 2007) and serving as transcriptional co-activators (Feng et al., 2006). Non-transcribed UCRs are generally associated with transcriptional cis-regulation; for example, they may behave as long-range enhancers in gene deserts (Calin et al., 2007; Nobrega et al., 2003; Pennacchio et al., 2006) and as tissue-specific enhancers (Pennacchio et al., 2006). Transcribed or exonic UCRs are believed to play roles in post-transcriptional regulation such as alternative splicing and mRNA processing ((Bejerano et al., 2004b; Lareau et al., 2007; Ni et al., 2007; Sathirapongsasuti et al., 2011). Besides, UCRs are often found to overlap with exons of genes involved in RNA splicing or locating within host gene introns or even in close proximity to genes, which are involved in transcription and development regulation (Bejerano et al., 2006; Bejerano et al., 2004b).

Recent studies identified alterations in T-UCR-expression patterns that are associated with specific tumor phenotypes and are pointing towards a mechanistic involvement of T-UCRs in cancer development. In particular, alterations in T-UCR expression have been described for adult CLL, CRC (Calin et al., 2007), hepatocellular carcinomas (Braconi et al., 2011; Calin et al., 2007), and neuroblastomas (Mestdagh et al., 2010; Scaruffi et al., 2010).

While UCRs appear to act as tissue-specific regulators of gene expression during development (Bejerano et al., 2004b; Mattick, 2009; Pennacchio et al., 2006), very little if anything is currently understood about the mechanisms underlying the cancer-associated profile alterations of T-UCRs. Similar to miRNAs, T-UCRs can act as oncogenes or tumor-suppressor genes, and besides their expression can be controlled by miRNAs (Calin et al., 2007). Moreover, UCRs are frequently observed in fragile genomic regions that are usually associated with cancers and therefore deregulated T-UCR signatures were identified as cancer-specific with a potential for prognostic implications in the future (Calin et al., 2007).

### 2.4.2 Cancer-associated chromosomal regions correlated with ncRNA genes

In human cancer research, bioinformatics and statistical tools have been used to demonstrate that ncRNAs (miRNAs, T-UCRs) are located in fragile genomic regions termed CAGRs (Calin et al., 2004), in tumor susceptibility loci (Sevignani et al., 2007) and in regions involved in cancers, such as common integrative sites (CISs, which are clustered inserts from CAGRs) (Makunin et al., 2007).

Human genome contains more than 100 fragile sites (Albertson, 2006; Glover, 2006) and some of them have already been associated with cancer (Pichiorri et al., 2008). In several studies where they compared tumor cells to normal common fragile sites were shown as sites of a frequent chromosome breakage and rearrangements (Arlt et al., 2003; Arlt et al., 2006). They are known to be highly recombinogenic, serving as sites where chromosome translocations, deletions and amplifications occur (Pichiorri et al., 2008). One of the area in the human genome where CAGRs could be present are minimal numbers of LOH (loss of heterozygosity), that occurs when a remaining functional allele becomes inactivated as a consequence of a somatic mutation and are predisposed to be associated with tumor-suppressor genes (Gryfe et al., 1997). Other CAGRs are genomic regions with a minimal number of amplifications, which are suggestive of regions harboring oncogenes (Albertson, 2006). Beside those two important groups of regions associated with cancer are also common regions of the genome where breakpoints occur and are often found near potential tumor-suppressor genes or oncogenes (Calin et al., 2004). In addition, many CAGRs with deletions, amplifications, or epigenetic alterations have been found to contain miRNA genes that may have a role in malignant transformation (Figure 4 and Figure 6) (Calin et al., 2002; Calin et al., 2005). Inherited or somatic mutations in the human genome are known to induce instability in a genomic region containing miRNA genes. These types of mutations can cause alterations in miRNA biogenesis or affect their post-transcriptional regulatory function. The genomic changes of miRNAs may not only cause cancer but may also increase the risk in individuals to develop it (Figure 6) (Sevignani et al., 2006).

### 2.4.3 Methods used for ncRNA profiling

Some methods used in the ncRNA expression analysis studies are used more frequently than the others. Their selection is mainly based on their sensitivity, specificity, timesaving, technical simplicity, low expense, and in correlation to the amount of useful information that can be gained from a single assay. Each method has its own advantages and disadvantages, which must be weighed against their intended use.

Profiling of ncRNA is used to document miRNA and other classes of ncRNA expression variability, and was found to be more accurate for cancer classification than

using whole sets of known protein-coding genes (Calin and Croce, 2006). Besides for disease prognosis the profiling of ncRNA may be used to define disease markers and, in some cases, even to predict cancer therapy outcome (Nicoloso and Calin, 2008; Schetter et al., 2008).

The most frequently used techniques for gaining information about ncRNA signatures are presented in Figure 7 and Table 2. Northern blot analysis was considered as the gold standard for gene expression detection and quantification before the development of microarray analysis. The first high-throughput technology was miRNA microarray analysis, which was initially developed in Carlo Croce's laboratory (Liu et al., 2008a). Other than microarray, frequently used strategies for ncRNA profiling are quantitative real-time polymerase chain reaction (RT-qPCR), bead-based hybridization, cloning, and deep sequencing (Barbarotto et al., 2008; Kong et al., 2009). Each technique is based on different principles, and all of them have some advantages and disadvantages (Table 2). Evolution of high-throughput analysis for ncRNA happened fast and methods for their analysis can be classified into two general categories, which distinguish methods according to the possibility of discovering new miRNAs (Figure 7). Most of the methods are based on hybridization (microarray, bead-based flow cytometry, and *in situ* hybridization (ISH)); some are based on amplification (RT-qPCR and sequencing), and the best combine both principles (cloning) (Creighton et al., 2009; Kong et al., 2009; Liu et al., 2008a).

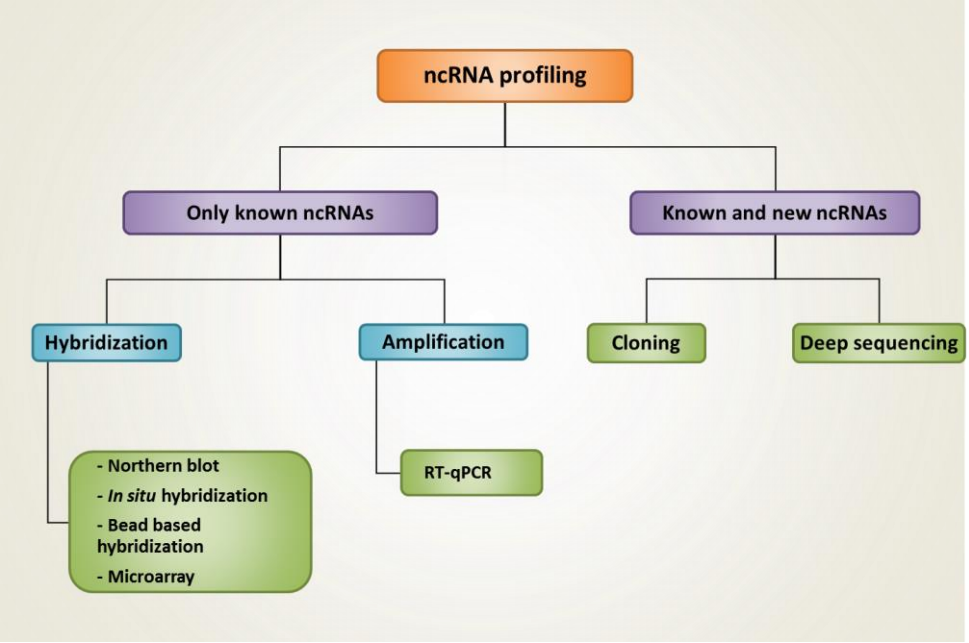



Figure legend: RT-qPCR, reverse transcription quantitative PCR.

Figure 7**:** Methods used for ncRNA profiling (Ferdin et al., 2010)

Slika 7: Metode, ki se uporabljajo za preučevanje izražanja nkRNA (Ferdin in sod., 2010)

Table 2: Methods used for ncRNA profiling (Ferdin et al., 2010)

Tabela 2: Metode, ki se uporabljajo za preučevanje izražanja nkRNA (Ferdin in sod., 2010)

| Method | Principle | Advantages | Disadvantages | Ref. |
|---|---|---|---|---|
| Microarray | • Sample preparation (RNA → cDNA)<br>• cDNA labeling<br>• Hybridization with solid-phase probes →signal<br>• Washing<br>• Detection of signal<br>• Data analysis | • High throughput method (simultaneous testing of numerous samples)<br>• More information gained in a single run<br>• Simple method to standardize<br>• Probes are spotted on solid surface by precise automated machine (thousands of probes on 1 $cm^2$)<br>• Genome-wide expression analysis | • Hard to differentiate expression of highly similar miRNA<br>• Expression data presented in relative format<br>• Detection of only known miRNAs<br>• Background noise and cross-hybridization problems | (Lu et al., 2005) (Liu et al., 2008a) |
| Bead-based array | • Small RNA isolation<br>• Adaptor ligation<br>• Adaptor specific reverse transcription<br>• PCR amplification<br>• PCR product labeling<br>• Signal detection with flow cytometry (hybridization of PCR amplicons with bead-fixed probes) | • High specificity (differentiation of closely related miRNA), sensitivity, and efficiency<br>• High speed, low cost<br>• Quantitative results<br>• Classification of miRNA based on fluorescent dye mixtures<br>• High-throughput analysis, but not the best one | • Long procedure (adaptor ligation, reverse transcription, target amplification)<br>• Limited (but not low) number of miRNAs can be analyzed in one assay<br>• Competition of hybridization for target PCR amplicon between its own complementary strand and specific probe<br>• Enrichment by fractionation is necessary | (Chen et al., 2008) |
| *In situ* hybridization (LNA probes) | • Tissue or cell preparation<br>• Digestion with protease (pepsin)<br>• Probe denaturation<br>• Hybridization → washing<br>• Conjugate incubation<br>• Chromogen incubation<br>• Reaction monitoring under microscope<br>• Immunohistochemistry | • RNA preservation (no cell lysis is performed)<br>• Actual localization of miRNA inside tissue or cell system<br>• Easy and not expensive in comparison to others | • Low-throughput method<br>• Poor detection of miRNAs present in low copy numbers<br>• Background signal problems in samples with a high number of miRNAs<br>Background signal problems in samples with a high number of miRNAs | (Nuovo et al., 2009) |
| RT-qPCR | • RNA extraction from samples<br>• RNA to cDNA reverse transcription (gene specific or total cDNA)<br>• Target sequence amplification | • Low numbers of ncRNA detection<br>• Qualitative and quantitative technique<br>• Rapid, highly sensitive, specific, and affordable | • Low throughput method<br>• Extraction of RNA and reverse transcription are steps where RNA can be degraded<br>• Quality control checks are necessary for | (Nolan et al., 2006; Schmittgen et al., 2004) |

*(Continued)*

*(Continued)*

| Method | Principle | Advantages | Disadvantages | Ref. |
|---|---|---|---|---|
| | | • Used as a confirmation method of microarray data<br>• No post-amplification detection is necessary | reliable quantitative results | |
| Cloning | • Small RNA isolation<br>• Adaptor ligation on both ncRNA ends<br>• RNA → cDNA reverse transcription<br>• Digestion of cDNA (sticky ends)<br>• Ligation with empty vectors<br>• Transformation and amplification in competent bacterial cells<br>• Sequencing for identification | • Quantification method<br>• Possibility of discovering new miRNA<br>• Ability to evaluate miRNA expression in a single cell | • Hard to detect low-expressed, instable, and tissue-specific miRNAs<br>• Low throughput method<br>• Enrichment by fractionation is necessary<br>• Technically difficult to perform<br>• Decision for right and best genomic region | (Cummins et al., 2006) |
| Northern blot | • RNA extraction from sample<br>• Size separation by electrophoresis<br>• UV/heat fixation<br>• Membrane hybridization by specific probe<br>• X-ray RNA visualization | • Quantitative method<br>• Separates RNA molecules by size and class<br>• Quantity comparison among loaded samples<br>• Simple, specific, cheap<br>• Detection of amount, size, and quality of transcripts<br>• Same blot can be re-probed several years later | • Low throughput method<br>• RNA degradation problem during electrophoresis<br>• Unhealthy chemicals needed for performing assay (radioactive material, formaldehyde, DEPEC, ethidium bromide, UV light)<br>• Difficult to perform using more than one probe<br>• Low sensitivity | (Streit et al., 2009) |
| Deep sequencing | • Library generation<br>• Genome/DNA of interest<br>• Fragmentation<br>• Adaptor ligation (DNAse, sonification)<br>• PCR amplification (*in situ* polonies, emulsion PCR, bridge PCR)<br>• Parallel sequencing | • Potential to discover novel ncRNA<br>• High throughput method (simultaneous sequencing of many different RNA molecules)<br>• Wide range of applications (polymorphism, mutation, and methylation detection, detection of structure variation, gene expression, and miRNA profiles)<br>• Absolute quantification of miRNA profiles | • Expensive (reagents, maintenance, bioinformatics software)<br>• Loss of miRNA during necessary purification steps | (Mardis, 2008; Shendure and Ji, 2008) |

## 2.5 CELLULAR POST-TRANSLATIONAL MODIFICATIONS

Within the last few decades, scientists have discovered that the human proteome is vastly more complex than the human genome (International Human Genome Sequencing, 2004; Jensen, 2004). This demonstrates that single genes encode multiple proteins. Genomic recombination, transcription initiation at alternative promoters, differential transcription termination, and alternative splicing of the transcript are mechanisms that generate different mRNA transcripts from a single gene (Ayoubi and Van De Ven, 1996).

The increase in complexity from the level of the genome to the proteome is facilitated by protein post-translational modifications. They are chemical modifications that play a key role in functional proteomics, because they regulate activity, localization and interaction with other cellular molecules such as proteins, nucleic acids, lipids, and co-factors. They occur at distinct amino acid side chains or peptide linkages and are most often mediated by enzymatic activity. It is estimated that 5% of the proteome comprises enzymes that perform more than 200 types of post-translational modifications (Walsh and Jefferis, 2006). These enzymes include kinases, phosphatases, transferases and ligases, which add or remove functional groups, proteins, lipids or sugars to or from amino acid side chains, and proteases, which cleave peptide bonds to remove specific sequences or regulatory subunits.

Many proteins are modified shortly after translation is completed to mediate proper protein folding, stability or its localization in cell (Figure 8). Other modifications occur after folding and localization are completed to activate or inactivate catalytic activity or to otherwise influence the biological activity of the protein. Proteins are also covalently linked to tags that target a protein for degradation. However, protein post-translational modification increases the functional diversity of the proteome by the covalent addition of functional groups or proteins, proteolytic cleavage of regulatory subunits or degradation of entire proteins. These modifications include phosphorylation, glycosylation, ubiquitination, nitrosylation, methylation, acetylation, lipidation and proteolysis and influence almost all aspects of normal cell biology and pathogenesis (Table 3) (Jensen, 2006).

The analysis of proteins and their post-translational modifications is particularly important for the study of heart disease, cancer, neurodegenerative diseases and diabetes. The identifying and understanding post-translational modifications is overall critical in the study of cell biology and disease treatment and prevention (Walsh and Jefferis, 2006).

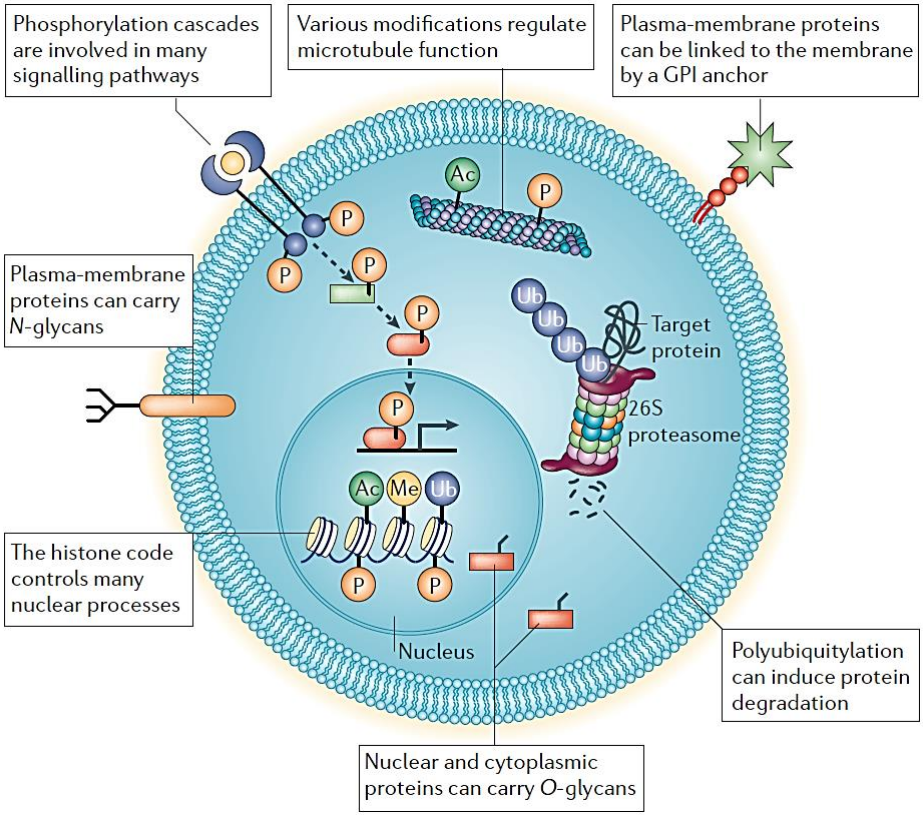



Figure 8: Post-translation modifications present in a cell (Jensen, 2006)

Slika 8: Potranslacijske spremembe v celici (Jensen, 2006)

Table 3: A selection of the most common post-translational modifications and their biological roles (modified from (Jensen, 2006))

Tabela 3: Izbor najpogostejših potranslacijskih sprememb in njihova biološka vloga (prirejeno po (Jensen, 2006))

| **POST-TRANSLATIONAL MODIFICATION*** • **modified residues** | **BIOLOGICAL FUNCTIONS** |
|---|---|
| **PHOSPHORYLATION** • Serine, Threonine, Tyrosine | Signal transduction, regulation of enzyme activity, involved in protein-protein and protein-ligand interactions. |
| **GLYCOSYLATION** • N-linked (Asparagine) • O-linked (Serine, Threonine) (single sugar or a chain sugar) • glycosylphosphatidylinositol (GPI) anchor | Protein stability, solubility, secretion signal, regulator of interactions, extracellular recognition and interactions. |
| **ACYLATION** • Palmitoylation • Farnesylation • Myristoylation | Protein localization and activity, involved in protein–protein and protein-membrane interactions. |
| **SULPHATION** • Tyrosine | Signaling and protein localization, involved in protein-protein interactions. |

*(Continued)*

*(Continued)*

| **POST-TRANSLATIONAL MODIFICATION*** <br>• **modified residues** | **BIOLOGICAL FUNCTIONS** |
|---|---|
| **UBIQUITYLATION** <br>• Lysine | Protein degradation signal, involved in protein-protein interactions. |
| **METHYLATION** <br>• Lysine (mono-, di- and trimethylation) <br>• Arginine (mono- and dimethylation) | Regulates protein activity, protein-protein and protein-nucleic acid interactions, chromatin dynamics and gene activity (histone modification). |
| **ACETYLATION** <br>• N-terminal residue, Lysine | Protein stability and activity, regulates protein-protein and protein-ligand interactions. |
| **DISULPHIDE-BOND FORMATION** <br>• Cysteine | Stabilizes protein structure and activity, involved in redox processes. |
| **OXIDATION** <br>• Methionine <br>• Tryptophan | Might regulate protein activity. |
| **DEAMIDATION** <br>• Asparagine <br>• Glutamine | Associated with ageing, might regulate protein activity and interactions; often a chemical artifact. |
| **HYDROXYLATION** <br>• Proline | Structural stability (collagens). |

### 2.5.1 Glycosylation

Protein glycosylation is acknowledged as one of the major post-translational modifications, with significant effects on protein folding, conformation, distribution, stability and activity. Glycosylation encompasses a diverse selection of sugar-moiety additions to proteins that ranges from simple monosaccharide modifications of nuclear transcription factors to highly complex branched polysaccharide changes of cell surface receptors. Carbohydrates in the form of aspargine-linked (N-linked) or serine/threonine-linked (O-linked) oligosaccharides are major structural components of many cell surface and secreted proteins (Hang and Bertozzi, 2005).

**O-GlcNAcylation** is the covalent attachment of β-D-N-acetylglucosamine (GlcNAc) sugars to serine or threonine residues of nuclear and cytoplasmic proteins, and it is involved in extensive crosstalk with other post-translational modifications, such as phosphorylation (Slawson and Hart, 2011). O-GlcNAcylation is becoming realized to have important roles in cancer-relevant processes, such as: cell signaling, transcription, cell division, metabolism, and cytoskeletal regulation (Slawson and Hart, 2011). Still little is known about the specific roles of aberrant O-GlcNAcylation in cancer.

O-GlcNAc-modified proteins have diverse cellular roles in signaling, metabolism, transcriptional regulation, cell cycle control, protein trafficking, and the regulation of cellular structure (Hart et al., 2011; Slawson et al., 2010). Because many of these processes are perturbed during tumorigenesis, alterations in O-GlcNAcylation are likely to have role in cancer.

### 2.5.2 O-GlcNAc and cancer

O-GlcNAcylation is catalyzed by a single enzyme, termed O-linked N-acetylglucosamine transferase (OGT). Role of OGT is to transfer N-acetylglucosamine from uridine diphosphate N-acetylglucosamine to protein substrates (Haltiwanger et al., 1990), meanwhile N-acetyl-β-glucosaminidase (OGA or O-GlcNAcase; official name is MGEA5) rapidly removes the O-GlcNAc modification from proteins (Dong and Hart, 1994). These enzymes enable alteration in proteins post-translational state and function in response to cellular signals (Slawson and Hart, 2011).

Cellular levels of O-GlcNAc, OGT and OGA have a tight relationship with multiple cellular activities. Loss of OGA activities induce changes of genes involved in life-span and aging, stress response, carbohydrate metabolism, membrane transport, and chromatin remodeling. Meanwhile, loss of OGT activity reflected in alterations of genes involved in amino acid metabolism, mitochondrial function, reproduction, stress response, and aging (Love et al., 2010).

Therefore, the O-GlcNAc modification is a crucial mechanism for cells to respond to various stimuli and allows cells to link nutrient availability and cellular metabolism to the regulation of vital cellular processes, such as cell cycle regulation, stress response and gene expression (Slawson and Hart, 2011). Many of these processes can become corrupted in cancer and despite conflicting data in primary tumor samples increased O-GlcNAcylation does seem to be a general characteristic of cancer cells. Although no human cancer were yet directly linked to mutations in OGT or OGA (Slawson and Hart, 2011).

However, changes in O-GlcNAcylation may disrupt the flow of information from other signaling system, such as phosphorylation, ubiquitilation and acetylation (Slawson and Hart, 2011). Transcription factors have predominance to be affected by O-GlcNAc modification (Ozcan et al., 2010) and beside that they can be widely regulated by other post-transcriptional modification like methylation, phosphorylation, and acetylation (Slawson and Hart, 2011). Alterations in transcription factor activity either by mutation or by gene expression are a well-known mechanism in the progress of cells from normal state to cancerous state (Hanahan and Weinberg, 2011).

### 2.5.3 OGT

Gene coding for OGT is highly conserved among eukaryotes, including mouse, rat, and humans. In humans its genomic location is on the long arm of the chromosome X (43 kb of genomic DNA in Xq13.1). In humans this region, Xq13, was associated with neurologic disease (Shafi et al., 2000). Transcripts of *OGT* are highly conserved in its 3' terminal part; meanwhile 5' terminal shows variability (Nolte and Muller, 2002). In

humans different sizes of OGT transcripts have been reported: 9.3-5 kb, 7.9-8.0 kb, 6.3-4 kb, 4.4 kb and 4.2 kb (Kreppel et al., 1997; Lubas et al., 1997; Nolte and Muller, 2002). Studies clearly identified that the largest transcripts (9.5 kb, 8.0 kb and 6.4 kb) are concern of alternative splicing. Nolte and Muller (2002) revealed that the alternative splicing of exon 2 and intron 4 occurs (later named retained intron), independently to each other. To mention retained intron itself does not contain any open reading frame (ORF), but does include a "second" start codon (ATG) (Lubas et al., 1997). This second start codon is also the only physiological solution for a translation of these two long transcripts (Nolte and Muller, 2002). The "fist" start codon is located at exon 1 of *OGT* appears to be used for translation of 6.4 kb long transcript (1036-1046 aa), where intron 4 is normally spliced out. Smaller transcripts like 4.4.kb and 4.2 kb according to Nolte and Miller most likely originate from transcript with retained intron (Nolte and Muller, 2002).

A complete *OGT* gene is composed of 22 exons and has two transcription stat codons (at exon 1 and retained intron) (Nolte and Muller, 2002). Beside, this protein is reported to be a heterodimer of two catalytic 110 kDa subunits and one 78 kDa subunit, which is supposed to originate from larger subunits as a result of alternative splicing (Kreppel et al., 1997).

# 3 MATERIAL AND METHODS

## 3.1 DATABASES

Locations of miRNA genes were obtained from miRBase (release16.0; www.mirbase.org/). Information for TF ChIP data were obtained from UCSC Genome Browser (University of California, Santa Cruz; http://genome.ucsc.edu/index.html). Experimentally validated miRNA::*HIF1A* direct interactions were collected using databases: miRWalk (http://www.umm.uni-heidelberg.de/apps/zmf/mirwalk/), miRecords (http://mirecords.biolead.org/), and mirTarBase (http://mirtarbase.mbc.nctu.edu.tw/). Gene locations and genetic variability were obtained using UCSC and Ensembl (http://www.ensembl.org/index.html) Genome Browsers. The nomenclature of literature-collected genes was uniformed according to the HUGO Gene Nomenclature Committee (HGNC, http://www.genenames.org/). The nomenclature for SNPs was unified to reference SNP (rs) according to the Human Genome Variation Society (HGVS) guidelines (http://www.hgvs.org/mutnomen/). Version of dbSNP build 137 from NCBI was used to obtain information about variations.

### 3.1.1 Publication databases

The literature reviews were performed using publication databases: PubMed (http://www.ncbi.nlm.nih.gov/pubmed) and Web of Science (WoS; http://apps.webofknowledge.com/WOS). Timeframes of the literature reviews depended on the time of the first and the lates publication on which these reviews were performed.

The literature review of *HIF1A* polymorphisms and associated diseases and phenotypes was performed on publications published between 2003 and 3/2013. Relevant publications were collected using following keywords: HIF1A, polymorphisms, SNP, human, and association studies. The literature review of studies describing profiling of miRNA in cancer tissue or cell lines was performed on publications published between 8/2004 and 10/2009. Relevant publications were collected using following keywords: microRNA, expression, profile, cancer, and expression profiles. The literature review of epigenetically regulated miRNAs was performed on publications published between 2006 and 7/2010. Relevant publications were collected using following keywords: noncoding RNA, microRNA, oncomiR, epigenetic gene regulation, DNA methylation, CpG island, histone modifications, cancer, and oncogene. Literature review of HIF1A downstream targets was performed on publications from 1992 to 3/2013 to obtain the list of experimentally confirmed HIF1A downstream target genes. Relevant publications were collected using following keywords: HIF1A, transcription regulation, direct targets, ChIP, luciferase reporter assay, and HRE binding site.

## 3.2 BIOINFORMATICS TOOLS

### 3.2.1 Transcription factor binding sites analysis

The information related to experimentally confirmed TF binding sites (TFBS) was obtained using ChIP–seq data from ENCODE project at the UCSC database (http://genome.ucsc.edu/index.html) and published literature. Prediction of TFBS was performed using TRANSFAC Match tool (accessed through Biobase, MDAnderson licence), which uses positional weight matrices (PWMs) from the TRANSFAC matrix library to search DNA sequences for potential TFBS. Predictions were performed for polymorphisms located within 5'-flanking *HIF1A* region and a set of two allele-specific sequences centered to SNP (51 bp) were used for the analysis. The analysis was limited on vertebrate non redundant matrices with minimal false positive results. Locations of HIF1A promoter region and transcription start site (TSS) were used as defined by Suzuki et al. (2002) (http://genomics.senescence.info).

### 3.2.2 MicroRNA target site predictions

MicroRNA target prediction was performed using RNA22 (https://cm.jefferson.edu/rna22v1.0-homo_sapiens/) and miRanda algorithms (http://www.microrna.org/microrna/home.do) and applied on *HIF1A* regions flanking SNPs, which were included in association analysis and located within exons, upstream and downstream gene regions. A set of two allele-specific sequences centered to SNP (51 bp) was used for the analysis. Human miRNAs were downloaded from miRBase (http://www.mirbase.org/index.shtml) (Kozomara and Griffiths-Jones, 2011). Algorithm RNA22 is based on a search for statistically significant miRNA motifs created from known mature miRNAs and allows to identify sites targeted by yet undiscovered miRNAs, but is not using cross-species conservational filter (Miranda et al., 2006). MiRanda algorithm in the analysis takes into consideration the cross-species conservation of miRNAs and allows one wobble pairing within the seed region (John et al., 2004). For miRNA predictions following parameter settings were used: 1.) the minimum number of base pairs between miRNA and mRNA was 14; 2.) the minimum number of seed paired bases was six; 3) the maximum folding energy was -14 kcal/mol.

### 3.2.3 Pathway enrichment analysis of HIF1A downstream targets

Enrichment pathway analysis with known HIF1A downstream targets was employed using DAVID Bioinformatics Resources 6.7 (Huang da et al., 2009) (http://david.abcc.ncifcrf.gov/) as described previously Cannistraci et al. (2013). The result of the enrichment analysis was obtained using Bonferroni multiple test correction and a *P*-value significant threshold of 0.01.

### 3.2.4 HIF1A-binding site predictions

The putative promoter sequence (−5 kb upstream and +10 kb downstream of the 5'end of the annotated T-UCR) was scanned for predicted HIF1A-binding sites generated using the MATCH program and the V$HIF1_Q3 (gnnkACGTGcggnn), V$HIF1_Q5 (ngtACGTGcngb), and V$HIF1_Q6 (nRCGTGngn) position weight matrices (PWMs) from the TRANSFAC database, version 9.1 (Matys et al., 2003) (http://www.gene-regulation.com). The position weight matrices describe the position preferences of different nucleotides in the HIF-binding site (Kulshreshtha et al., 2007).

### 3.2.5 ORF detection

To find The *OGT* isoforms 1 and 2, and the retained intron sequence (intron 4) of the *OGT* gene were tested for open reading frame (ORF) using the NCBI ORF finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html).

## 3.3 MOLECULAR GENETICS METHODS

### 3.3.1 Cell cultures and growth conditions

We used different cancer cell lines for screening and functional studies to determine the biological roles of the T-UCRs. Cell lines were selected based on previously made associations between HIF1A accumulation and poor patient survival in following cancer types. Colon cancer (HCT-116, HT-29 and COLO-205), breast cancer (MCF-7), bladder cancer (UC-2), and glioblastoma (U-87) cell line were obtained from ATCC (Manassas, VA, USA). Renal cell carcinoma (RCC) cell lines: 786-O VHL-deficient (VHL −) and 786-O VHL -restored (VHL +) (Kibel et al., 1995) were obtained from Dr. Maria Czyzyk-Krzeska (University of Cincinnati, Cincinnati, OH). The 786-O cell lines we used differ in ability to form a functional VHL protein (Kibel et al., 1995), and were used since the alteration of the VHL/HIF/VEGF pathway is well characterized in carcinogenesis and is the target of the current therapies of the metastatic RCC. Cell lines 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 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 VHL-deficient (VHL −) and 786-O VHL-restored (VHL +) cell lines (Kibel et al., 1995) were cultured in DMEM:Ham's F12 supplemented with G418 (200 μg/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 AmpFℓSTR Identifier kit, according to the manufacturer's instructions (Applied

Biosystems, cat # 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/) (Romano et al., 2009), and the University of Texas MD Anderson Cancer Center fingerprint database. The STR profiles of all the cell lines matched known DNA fingerprints.

### 3.3.2 Clinical samples

Our starting point in this study was previously published UCR microarray expression data for CRC patients and healthy controls. Seventy-eight primary CRC and 21 normal colonic mucosa samples that were used for microarray analysis had been collected at the University of Bologna (Italy) and were previously reported (Calin et al., 2007). For all patients written informed consent was obtained and the patient samples were de-identified for the molecular study.

### 3.3.3 Hypoxia experiments

Hypoxia experiments were performed to observe the expression levels of T-UCR and to determine the role of T-UCR in environment limited with $O_2$. In hypoxia experiments, we used different cancer cell lines to represent different cancer types: colon cancer (HCT-116), breast cancer (MCF-7), bladder cancer (UC-2), glioblastoma (U-87), and renal cell carcinoma (786-O). We plated $6*10^5$ cells/well under normoxic conditions and $8*10^5$ cells/well under hypoxic conditions in complete McCoy's growth medium (HCT-116), DMEM (MCF-7), MEM (UC-2) or DMEM/F12 (U-87) growth medium. Cell concentrations under normal conditions were lower to avoid cell overgrowth and consecutive cell death. Concentrations were measured using the Vi-Cell™ cell viability analyzer (Beckman Coulter, Inc, Miami, FL, USA). Experiments were performed in six-well culture plates (CELLSTAR; Greiner Bio-one, Capital Drive, Monroe, NC, USA). Each cell line was plated in duplicate for RNA extraction and protein harvesting. Cultivation under normoxic conditions was set on 37 °C at 21% $O_2$ for 24 and 48 h. A hypoxic environment was generated in an *InVivo*$_2$ 200 hypoxia workstation (Ruskinn, Inc., Cincinnati, OH, USA), where $O_2$ concentration was maintained at 1% or 0.2% (in 5% $CO_2$ and 95% $N_2$) for 24, 48, 72 and 96 h. RNA was isolated and reverse transcribed with random hexamers (Applied Biosystems, Foster City, CA, USA) or gene-specific primers to gain total complementary (cDNA) or sense-specific cDNA, respectively. Complementary DNA was used in the RT-qPCR analysis to determine gene expression. Proteins were collected to confirm the presence of hypoxia-specific proteins under hypoxic conditions, especially HIF1A protein and HIF1A-induced proteins such as carbonic anhydrase IX (CA9). The 786-O and MCF-7 cell lines were plated in 35-mm plastic dishes for 24 h and incubated under normoxic or hypoxic conditions (0.2% $O_2$) for 24 or 48 h. After treatment, cells were washed with phosphate-buffered saline (PBS), and RNA was extracted using the miRNeasy kit (Qiagen). Chemical induction of HIF1A expression under normoxic conditions was performed using

dimethyloxalylglycine (DMOG; Frontier Scientific, Logan, UT, USA). A stock solution of 100 mM DMOG was prepared in dimethyl sulfoxide (DMSO). The DMOG was added to the growth medium to a final concentration of 500 μM; meanwhile, a control set of cells was incubated in the presence of 0.5% DMSO.

### 3.3.4 T-UCR validation using RT-qPCR

#### 3.3.4.1 RNA extractions and reverse transcription (RT)

Extraction of RNA and reverse transcription were performed to prepare samples for T-UCR expression analysis measured by RT-qPCR method. Total RNA was extracted from cultured cells using TRIzol by following manufacturer's protocol (Invitrogen). RNA concentrations were measured with spectrophotometer NanoDrop ND-1000 instrument (NanoDrop Tachnologies, Termo Scientific, Wilmington, DE, USA) using 1 μl per sample. Reverse transcription of RNA to total cDNA was performed on 800 ng of RNA. Final reaction of 20 μl volume contained 1 μl of random hexamers (50 ng/μl), 1 μl of 10 mM dNTPs, 8.5 μl RNAse free water, 4 μl of 5 x First Strand Buffer, 1 μl of dithiothreitol (DTT; 0.1 M ), and finally 0.5 μl of SuperScript®III Reverse Polymerase (200 U/μl). All steps were precisely followed by manufacturer's instructions (Invitrogen). Reverse transcription by gene-specific primers was also made in total of 20 μl, where only 100 ng of RNA was transcribed to strand specific cDNA. Reaction mix for gene-specific reverse transcription was composed of 1 μl RNA (100 ng/μl), 1 μl of a mix with strand-specific primers (2 μM), 1 μl dNTPs (10 mM), 4 μl of 5 x First Strand Buffer, 1 μl of DTT (0.1 M), 1 μl of RNAse OUT (Invitrogen) and finally 0.5 μl of SuperScript®III Reverse Polymerase (200 U/μl), and then RNAse free $H_20$ was added to final volume of 20 μl. Two reverse transcription reaction mix were made for strand specific cDNA synthesis: one to detect sense oriented transcripts (using mix of reverse primers also reference gene) and one for antisense oriented transcripts (using reverse primer of a reference gene together with forward primers of other genes). Reference gene was transcribed in sense orientation in both reaction mixes. MicroRNA genes were reverse transcribed using specific primers (Applied Biosystems) and following manufacturer's instructions (Applied Biosystems). Reverse transcription mixture was further incubated in Dyad® Pettier Thermal Cycler (DNA Engine, Bio-Rad, Hercules, CA, USA). Set of temperatures for total cDNA transcription was as followed: 5 min / 65 °C, 2 min / 4 °C, 12 min / 25 °C, 50 min / 42 °C, and 15 min / 70 °C (Invitrogen). Incubation for gene-specific reverse transcription was executed as follows: 5 min / 65 °C, 60 min / 55 °C, and 15 min / 70 °C to terminate the reaction incubation. Gene-specific RT and RT-qPCR were performed only for selected UCGs. Specific primers were designed by using Primer3Plus (http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi) and were further tested with OligoAnalyzer for self- or hetero-dimers (IDT, Science tools; http://www.idtdna.com/analyzer/Applications/OligoAnalyzer/). Before use total cDNA

and gene-specific cDNA were diluted with nuclease-free water to final volume of 200 µl (10-times) and 100 µl (5-times), respectively. All examined samples including clinical samples that were screened with RT-qPCR assay were simultaneously reverse transcribed without a presence of an enzyme (named RT (-) samples). This was necessary to exclude possibility of genomic DNA (gDNA) contamination and only samples, which were reported as RT (-) negative, were further incorporated in RT-qPCR analysis.

#### 3.3.4.2 Reverse Transcription Quantitative Polymerase chain reaction (RT-qPCR)

Quantitative RT-PCR (RT-qPCR) was the first method chosen to confirm microarray data in ANNEX A from reference (Calin et al., 2007). This data gave us 61 altered expressed UCGs in CRC clinical samples (two down and 59 up) and have been normalized with normal colonic mucosa tissue expression data (Calin et al., 2007). Reactions for RT-qPCR were performed differently to measure expression of UCGs and differently to detect miRNAs. UCGs were detected by iQ™ SYBR®Green Super Mix (Bio-Rad, Hercules, CA, USA). Final volume for reaction mix was 10 µl and contained 5 µl of 2 x iQ™ SYBR®Green Super Mix, 2.6 µl of DNAse free water, 2 µl of cDNA (cca. 40 ng / µl, dilution 1:10), and 0.2 µl of an individual primer (10 µM). MicroRNA gene-specific Taq-Man based RT-qPCR for miRNA was performed using miRNA-specific and labeled hydrolysis probes (20 x, Applied Biosystems). Final volume of PCR reaction mix was 10 µl and consisted of 5 µl 2 x TaqMan Universal PCR Master Mix (Applied Biosystems), 2.6 µl of DNAse free water, 2 µl of diluted gene-specific cDNA (1:5), and finally 0.4 µl of labeled miRNA-specific hydrolysis probe (Applied Biosystems). Each cDNA sample was tested in duplicates (two wells per sample) and final volume 10 µl. All gene specific primer pairs are presented in ANNEX B, details about miRNA-specific products can be found at manufacturer's homepage (https://products.appliedbiosystems.com). Analysis was executed on 384-well white well plates using C1000™ TermalCycler with CFX384™ Real-Time System (Bio-Rad). Amplification has started with 10 min heating at 95 °C and then followed by 40 amplification cycles consisted of two steps: 95 °C for 15 sec and 60 °C for 30 sec (previously described by Schmittgen et al. (2004)). After completed amplification PCR products were validated by melting curve designing step, which started by increment of 0.5 °C / 5 sec beginning at 65 °C and finished at 95 °C. Melting curve was performed to validate specificity of transcripts amplified and detected by Sybr Green, since it non-specifically incorporates into everything that consists of a double-stranded DNA (dsDNA) structure. Amplification condition in miRNA gene-specific Taq-Man based RT-qPCR also started with 10 min heating at 95 °C and then followed by 40 amplification cycles consisted of: 95 °C for 15 sec and 60 °C for 1 min.

#### 3.3.4.3 Real time PCR controls, reference genes

By performing RT-qPCR assay we had to use appropriate reference genes, which do not change expression under examined experimental conditions. As the positive control we used a gene known to be induced in hypoxia; for longer transcripts we checked HIF1A induced expression of *CA9* (Wykoff et al., 2000), while in miRNA gene-specific RT-qPCR (TaqMan) expression of hsa-miR-210 (Huang et al., 2010; Kulshreshtha et al., 2007) was checked. Several different reference genes were tested in hypoxia and normoxia conditions, to identify which of them is not influenced by conditions we are studying. In the choice of the reference gene even a size of a transcript that we were amplifying was taken into account. MicroRNA expression analysis and T-UCR transcripts expression analysis were normalized with U6 snRNA (*RNU6-1*; RNA, U6 small nuclear 1) and *PPIA* (peptidylprolyl isomerase A (cyclophilin A)), *GAPDH* (glyceraldehyde-3-phosphate dehydrogenase) or *RPLP0* (ribosomal protein, large, P0), respectively. To avoid contamination we included the negative control, which was without any DNA template (no-template control, NTC), to check for potential intra-assay contaminations.

#### 3.3.4.4 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 quantification cycle (Cq) values for each sample of duplicates. Relative expression, as the fold change or difference (Safdar et al., 2009), was calculated using the $2^{(-\Delta\Delta Cq)}$ formula, in which $\Delta$ Cq represents the difference in Cq values between the target gene and the mean Cq of the reference gene (Livak and Schmittgen, 2001).

### 3.3.5 Western blot

This method was used to observe protein expression levels in colorectal and breast cancer cell lines under induced conditions. Detection of HIF1A protein in environment with lower $O_2$ concentrations was performed using western blot assay after incubation in hypoxia/normoxia for several experiments. Whole cell proteins were harvested from cells after washing with PBS. Proteins were collected by lysation with pre-prepared lysation mix (depend on cell density) composed of 1 x Lysis buffer (0.5% NP40 (nonyl phenoxypolyethoxylethanol), 250 mM NaCl, 50 mM Hepes, 5 mM EDTA (ethylenedinitrilo-tetraacetic acid), 0.5 mM EGTA (ethylene glycol tetraacetic acid)), 1 x Proteinase Inhibitor Cocktail 2 (100 x; SIGMA), 1 x Protease Inhibitor Cocktail (25 x; Roche, Mannheim, Germany), and 1 mM DTT (1 M; Invitrogen). Protein samples were further homogenized by vortex-ing, kept on ice for 30 min, vortexed again, and then centrifuged for 10 min at 4 °C. Supernatant was transferred into fresh clean tube and

stored at -80 °C until use. Protein concentrations were measured with Smart Spec™ Plus spectrophotometer (Bio-Rad). Standard curve was made with BSA (1 µg/µl bovine serum albumin; SIGMA). One µl of proteins was diluted in 1 ml 1 x Bio-Rad Protein assay dye reagent (5 x, Bio-Rad) prior measurement. If possible 40 µg, but not less than 20 µg of proteins per sample (amount of proteins on the same precast gel was the same) were separated on 4-20% Tris-HCl Criterion™ Precast Gel (Bio-Rad) and transferred onto 0.2 µm nitrocellulose membrane (Bio-Rad). The membrane was hybridized with specific antibodies detecting proteins of interest. Primary antibodies were detected by species specific secondary antibodies linked with Horseradish peroxidase (HRP) (ECL™, GE Healthcare, UK) and then detected using enhanced Chemiluminescence detection kit (GE Healthcare UK Limited, Little Chalfont Buckinghamshire, UK). For reference proteins in hypoxic/normoxic environment we use Vinculin (Santa Cruz Biotechnology Inc., cat # sc-7649), Beta-actin (SIGMA,cat #A5316) or Lamin B1 (Abcam, cat # 16048), which expressions are not effected by hypoxic conditions. HIF1A was a protein to confirm hypoxic conditions (BD Biosciences, San Jose, cat # 610958 or Nouvus Biologicals, cat # NB100-134) and due to its fast decomposition under short time of normoxic condition, CA9 was used as a control (Novus Biologicals, cat # NB100-417) whereas it is directly induced by HIF1A. Detecting the presence of HIF proteins in VHL-deficient cells we used HIF2A antibody (R&D Systems, cat # AF2886). Target protein to associate the presence/absence of apoptosis was of cleaved/uncleaved PARP protein (Cell Signaling, cat #9542). Antibody to detect OGT was purchased from SIGMA (cat # O 6264) meanwhile primary and secondary mouse specific antibodies to detect addition of O-GlcNAc on proteins were previously reported by Caldwell et al. (2010).

### 3.3.6 RNA interference with siRNA transfection

To determine the biological significance of UCR we performed knockdown experiments for selected T-UCRs using custom-designed siRNAs. Transfection with three different siRNA against selected gene targets was performed in 24-well plate using transfection reagent Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA). Transfection for each siRNA was made in six replicates. Two wells were used for cell counting (to detect proliferation effect), one well for RNA expression analysis to confirm efficient siRNA knockdown, and the rest for protein collection (western blot). Cell lines COLO-205 and HT-29 were used in siRNA interference experiment based on their highest expression of target T-UCR (COLO-205: uc.63; HT-29: uc.475) among a panel of five cell lines (Figure 24). Only siRNA with at least 60% target repression were used for further functional analysis. Sequences of used siRNAs are presented in ANNEX C. Cells ($0.3*10^5$/well) were plated together with transfection mix and were considered as time zero hour concentration. Transfection mix (500 µl) was composed of 1µl Lipofectamine 2000 (Invitrogen), 1 µl of siRNA (50 µM; SIGMA), 98 µl OPTIMEM® (GIBCO 35050, Invitrogen™, Grand Island, NY, USA), and 400 µl of an

appropriate cell growth media without antibiotics. Transfection was performed following manufacturer's protocol. Cells transfected with siRNA were maintained under hypoxic (1% $O_2$) and normoxic (21% $O_2$) condition. After 4-6 h media was substituted with growth media containing antibiotics. Cells were incubated for 24/48/72/96/120 h post-transfection to determine whether T-UCR downregulation exhibited an effect on cell proliferation. Cell counting at each time point was performed with Vi-Cell™ cell viability analyzer (Beckman Coulter), while RNA samples were collected to confirm knockdown efficiency and proteins extraction to quantify the gene product of interest by western blot analysis.

### 3.3.7 Fluorescence-activated cell sorting (FACS)

Fluorescence-activated cell sorting was used to determine percentage of cells in cell cycle phases after silencing with siRNAs. Cells were plated in duplicates in a 6-well plate and a final 100 nM siRNA concentration was used for siRNA transfection (SIGMA). Cells were plated in concentration $3*10^5$ cells/well. After 48/72/96 h post-transfection cells were collected by trypsinization and centrifuged so the media could be properly removed. Cell pellet was washed with 700 µl PBS and then fixed by 1400 µl of freezing cold ethanol (100%) added drop wise during gentle vortex-ing. Following cells were stored for minimum 1 h at -20 °C before adding 10 ml of PBS (without $Ca^{2+}$/$Mg^{2+}$) and mixed by tube inverting. After maximum speed (3500 rpm) centrifugation for 10 min supernatant was separated from the cell pallet. One ml of staining solution was added on top of a cell pallet, mixed, and left in the dark for 30 min at 37 °C. Used cell staining solution was composed of 50 µl propidium iodide (1 mg/ml; PI, SIGMA), 50 µl RNase A (2 mg/ml; Roche Applied Science), and 900 µl of PBS. After incubation cells were stored at +4 °C till the time when flow cytometric measurement on FACS Calibur machine (Beckman Coulter, Irving, TX; www.coulterflow.com) in program for cell cycle analysis were implemented. Cell cycle analysis was made based on fluorescence-activated cell sorting (FACS) described above. Cells were sorted by cycle stage and were presented in percentages.

### 3.3.8 Chromatin immunoprecipitation (ChIP)

Chromatin immunoprecipitation (ChIP) was used to validate predicted HIF1A-binding sites. The ChIP assay was performed using the Millipore chromatin-immunoprecipitation kit (Millipore, Billerica, MA; cat # 17-295) following the recommendations from manufacturer, with slight modification. In brief, MCF-7 cells were exposed to DMOG (1 mM) for 24 h. Cells were fixed in the presence of 1% formaldehyde for 10 min at room temperature. For DNA shearing, $1*10^6$ cells were sonicated (6-times 10-second pulses) to achieve DNA fragmentation of 200-1000 base pairs length. Immunoprecipitation was carried out using rabbit polyclonal antibody against HIF1A (Novus Biologicals, cat # NB100-134) overnight at +4 °C. For DNA extraction, the immunoprecipitates were incubated with 10 µl of 0.5 M EDTA, 20 µl of

1 M Tris-HCl, pH 6.5 and 2 µl of 10 mg/mL Proteinase K (Roche) for one hour at +45 °C. The DNA was purified by QIAquick PCR Purification Kit (Qiagen, cat # 28104) and eluted in 50 µl of water. Five µL of this DNA was used as a template for conventional PCR. All primers were designed with Primer3 software available online (http://frodo.wi.mit.edu/primer3/). PCR amplifications were carried out in Eppendorf Mastercycler Gradient with HotStartTaq Master Mix Kit (Qiagen). PCR cycling conditions were as follows: 95 °C for 15 min, 32 × (30 sec at 94 °C, 45 sec (56-60 °C) then 60 sec at 72 °C), followed by a final 10-minutes extension at 72 °C.

### 3.3.9 Promoter reporter assay

This assay using promoter fragments was used to confirm the transcriptional activation of T-UCR by hypoxia and HIF1A. The uc.475 promoter fragment containing predicted HIF1A motif (ANNEX D) was amplified and cloned into basic pGL4 vector (pGL4.10(luc2)) (Promega) with Forward 5'-GAAGATCTGCTCTCAGAGATCCAACTTGTAGCCTCC-3' and Reverse 5'-GAAGATCTAGCTTTTGATCGTCCTCGGGTGATG-3' (marked as wild type, WT, promoter variants). Plasmid construct without HIF1A-motifs were amplified on WT promoter variant later named as truncated promoter 1 (PT1) and truncated promoter 2 (PT2). Primer pair used to amplified 5'-promoter region before HIF motifs (PT1 region) was Forward 5'- GAAGATCTGCTCTCAGAGATCCAACTTGTAG -3' and Reverse 5'- GAAGATCTTACTTAGCAGGTGGTGTCGTTT -3', meanwhile primer pair used to amplified 3'-promoter region after HIF1A-motifs (PT2 region) was Forward 5'-GAAGATCTTTAGAATGTCGTGGAATGGACT-3' and Reverse 5'-GAAGATCTTTAGAATGTCGTGGAATGGACT-3'. Amplified products were directly cloned into basic pGL4-empty vector with *Bgl*II restriction sites, upstream of luciferase reporter, and later all three of them were used for Dual-Luciferase assay to test inserted promoter region for transcription activation by HIF1A transcription factor. Each construct plasmid, as showed in Figure 22 (pGL4-empty, pGL4-uc.475-promoter, pcDNA 3.1-empty, pcDNA 3.1-HIF1A (P/A or 3M), were introduced into HCT-116 and HT-29 cells separately using Lipofectamine 2000 (Invitrogen). The pRL-TK Vector (Promega) was used as an internal control reporter vector in each of transfection mix. Luciferase assays were performed 48 h post-transfection following the manufacturer's protocol (Promega). Luciferase activity (measured as Relative Light Units, RLU) was scored using a Veritas microplate Luminometer (Turner Biosystems, Sunnyvale, CA, USA). As controls we produced two truncated inserts lacking both interaction sites, one containing only the 5'-part and the second only the 3'-part of the WT construct (Figure 23**,** bottom).

### 3.3.10 Detection of global O-GlcNAc cellular-protein modifications

To detect O-GlcNAc-modified proteins, we chemically (PUGNAc as a 1,5-hydroximolactone is an inhibitor of various N-acetylhexosaminidases) prevented the

removal of O-GlcNAc modifications by the glycosidase MGEA5 (meningioma-expressed antigen 5 (hyaluronidase)) (Gao et al., 2001). Treatment with PUGNAc was initiated 24 h after siRNA transfection. Fresh complete media containing PUGNAc was added to final concentrations of 0, 40, and 80 μM, respectively, and in duplicate wells. The stock PUGNAc solution (SIGMA; cat # 132489-69-1) was prepared following manufacturer's instructions and stored at -20 °C. After 72 h under normoxic or hypoxic conditions, cells were washed with PBS and later harvested for RNA and protein extraction 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. (2010).

### 3.3.11 Northern blot

Northern blot using gene specific probes was performed to identify transcript present in normoxia and altered in hypoxia. Total RNA isolation was performed using the TRIzol reagent (Invitrogen) according to the manufacturer's instructions. The RNA samples (10 μg each) were size-fractionated on pre-made 1% formaldehyde agarose gel and then transferred onto a Hybond-N+ membrane (Amersham Pharmacia Biotech, cat # RPN303B). Probes were radiolabeled with dCTP [$\alpha^{-32}$P] (stock 10 mCi/ml; PerkinElmer, cat # 102783-51-7) by using Prime-it II Random Primer labeling kit (Stratagene-Agilent Technologies, cat # 300385). The hybridization with [$\alpha^{-32}$P] was performed at 42 °C in Modified Church's buffer (7% sodium dodecyl sulfate (SDS)–0.2 M $Na_2PO_4$ (pH 7.0)) overnight. Membranes were washed at room temperature, twice with 2 x saline-sodium citrate (SSC)–0.1% SDS and once with 0.1 x SSC–0.1% SDS. Blots were stripped by boiling in 0.1% aqueous SDS–0.1 x SSC for 10 min and later re-probed several times. As a loading control we used probing for *PPIA*. Before labeling DNA template for northern blot, probes were amplified using following primers for *OGT* (5'-TTAGCTGAGTTGGCACATCG-3' (Forward), 5'-TTGTACTGCCCCTTCCATGT-3' (Reverse)), for uc.475 (5'-TTACAGACTTGAGACTGTCCTCATT-3' (Forward), 5'-AAGCTGTTTCTGAAGCCACAA-3' (Reverse)) and reference gene *PPIA* (5'-CAGAC AAGGTCCCAAAGACAG-3' (Forward), 5'-TTGCCATCCAACCACTCAGTC-3' (Reverse)).

### 3.3.12 Cell fractionation

In order to identify whether the location of T-UCR transcript is nuclear or cytoplasmic cell fractionation was performed. Cells (HCT-116, $2*10^6$) 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).

### 3.3.13 Oxygen and glucose deprivation (OGD) assay

This method was used to identify glucose- and $O_2$-deprivation effect on the mRNA expression level of T-UCR and its host gene. Cells (HT-29, $0.3*10^6$) were seeded with DMEM media containing 4.5 g/L D-glucose concentrations supplemented with 10% fetal bovine serum (FBS) without antibiotics in 6-well tissue culture plate. Cells were plated one-day prior experiment to reach 50% confluence. On the day of the experiment cells were washed three times with PBS to remove the excess of glucose from the culture. One set of cells was cultured in normoxic while the other was cultured in parallel under hypoxic conditions for 12, 24, and 48 h. For each time point and conditions (hypoxia or normoxia) parallel plates with D-glucose (4.5 g/L) or without glucose (DMEM, Cellgro Mediatech, cat # 17-207-CV) were set up. Protein and RNA were collected to be further analyzed for coding and noncoding gene expression regions of the genome.

### 3.3.14 Enhancer activity of DNA region

To test T-UCR-containing region for enhancer activity, at first the pGL3 basic vector had to be transformed by the thymidine kinase (TK) promoter insertion as described in (Orom et al., 2010). Using primer walking, we confirmed the cDNA transcript (Figure 35) for the retained intron (2694 bp) containing uc.475 (365 bp) (pGL3-TK-2.7 kb); their genomic regions were directly cloned into empty *Bam*HI-restricted pGL3-TK promoter vector. The primer pair used to amplify the retained intron of *OGT* gene (2694 bp) was 5'-AGCGGATCCGGTCAGAAGGACAAAGCTGGT-3' (forward) and 5'-AGCGGATCCAGCCAAATTTCCCCTTGTG-3' (reverse). The insert was cloned downstream of Firefly luciferase, driven by the TK-promoter. A dual-luciferase assay was used to test and enhancer-like function and a vector containing Renilla luciferase (pRL-TK) was used as a control for transfection efficiency. Two independent biological repetitions under normoxic and hypoxic conditions were performed. All samples were treated in quadruplicates and dual-luciferase assay measurements were run in two wells for each sample.

### 3.3.15 Noncoding RNA microarray data analyses

For T-UCR expression analysis microarray chips were developed with a total of 481 human UCR sequences as in http://www.soe.ucsc.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 (Liu et al., 2004) and the detailed microarray description is in (Calin et al., 2007). 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 software (Agilent Technologies).

Hierarchical cluster analysis was done using average linkage and Pearson correlation as measures of similarity. Statistical comparisons of tumors and normal tissues were performed by filtering on fold change and then using the the Welch *t*-test with Benjamini-Hochberg correction for reduction of false positives and as previously published (Calin et al., 2005; Cimmino et al., 2005; Iorio et al., 2005). All data were submitted to ArrayExpress database (accession number E-TABM-184).

### 3.3.16 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 *t*-tests were used to address differences between groups.

# 4 RESULTS

We developed an Atlas of *HIF1A* gene regulatory network, integrating genomic information related to HIF1A obtained from publications, databases, and bioinformatics predictions. We also confirmed a new functional link between hypoxia and the expression of long ncRNA, T-UCRs. One member of T-UCRs, uc.475, and its host gene *OGT*, were validated as novel HIF1A targets.

## 4.1 ATLAS OF HIF1A GENE REGULATORY NETWORK

We developed an Atlas of *HIF1A* gene regulatory network integrating: 1.) functional *HIF1A* SNPs from genotype-phenotype association studies, 2.) experimentally validated and predicted TF and miRNA binding sites, and 3.) HIF1A downstream targets and their involvement in biological pathways. An overview of an Atlas of *HIF1A* gene regulatory network consisting of transcriptional and post-transcriptional regulation and downstream targets is presented in Figure 9.

On the Atlas of HIF1A gene regulatory network following HIF1A domains are presented: at the 5'-terminus of HIF1A is basic helix-loop-helix domain (bHLH; 17-71 aa), essential for DNA binding. In the central region of the protein are Per-ARNT-Sim domain (PAS; 85-298 aa), involved in HRE binding and heterodimerization with ARNT (HIF1B), and an $O_2$-dependent degradation domain (ODD, 401-603 aa), with proline residues (*p.Pro402* and *p.Pro564*) hydroxylated for VHL/proteasome degradation (Ivan et al., 2001). At the 3'-terminus of the protein are two transacting domains (N-TAD and C-TAD), that mediate transcriptional activation and interact with co-activators, and a nuclear localization signal (NLS, 718-721 aa), which labels a protein for import into the cell nucleus by nuclear transport (Heino et al., 2008; Pugh and Ratcliffe, 2003; Shieh et al., 2010; Tanimoto et al., 2003). Variation presented in CDS regions were collected from numerous genotype-phenotype association studies performed on *HIF1A* (Figure 10). Protein domains are evidently very important for interactions with other proteins (TFs (Figure 11) and miRNAs (Figure 12, Figure 13 and Figure 14)) and HRE binding sites within or nearby target genes promoters (experimentally validated HIF1A target genes; Figure 15). When the HIF1A protein stability is achieved then its role is observed in various cell processes.

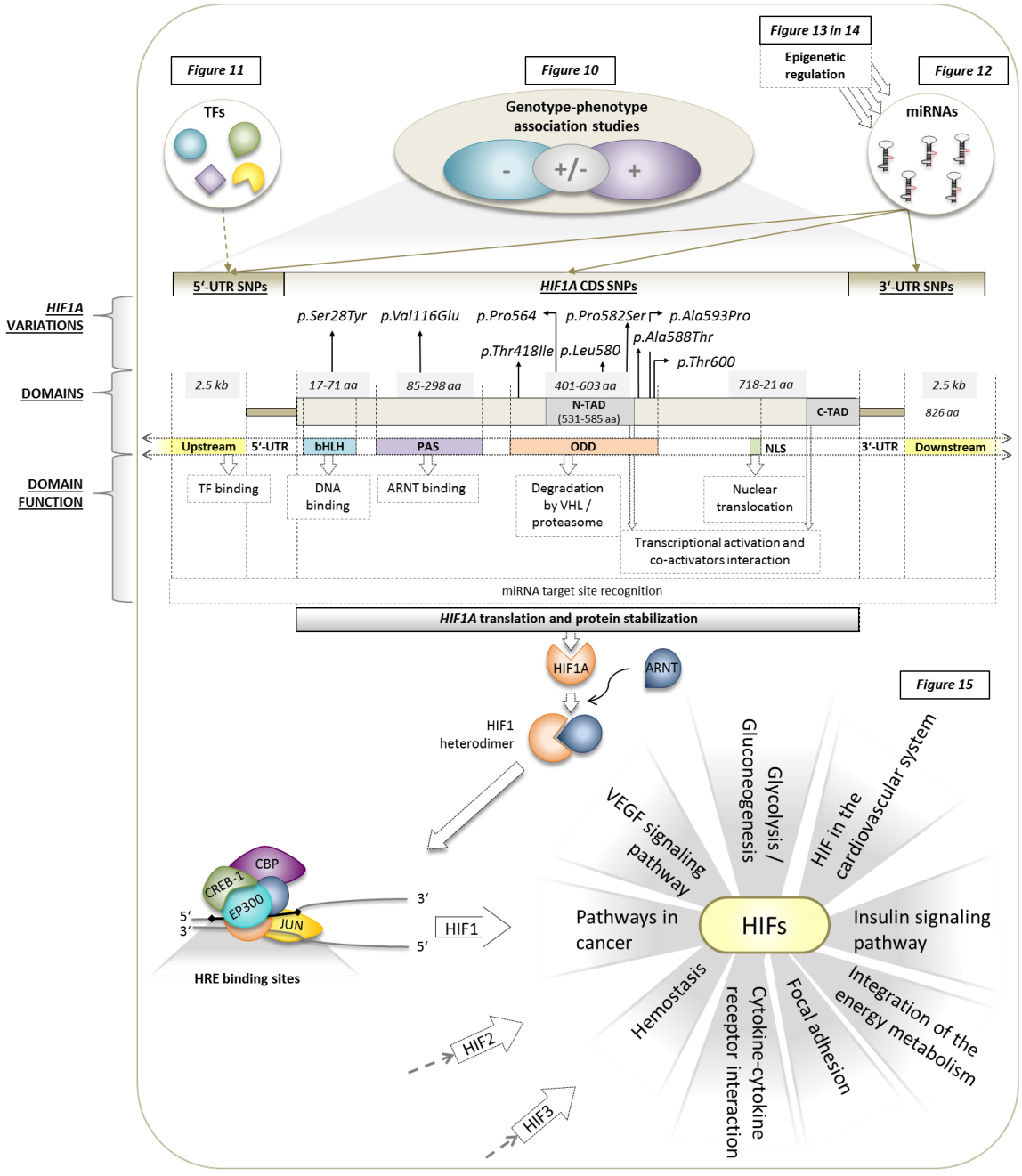



Figure legend: HIF1, HIF2, HIF3: hypoxia induced factors 1, 2, 3; HIF1 transcription co-factors: EP300, E1A binding protein p300; JUN, jun proto-oncogene; CREB1, cAMP responsive element binding protein 1; CBP, CREB binding protein. HRE, hypoxia response element; SNP, single nucleotide polymorphism; TF, transcription factor, aa, amino acid; UTR, untranslated region; CDS, coding sequence; miRNA, microRNA.

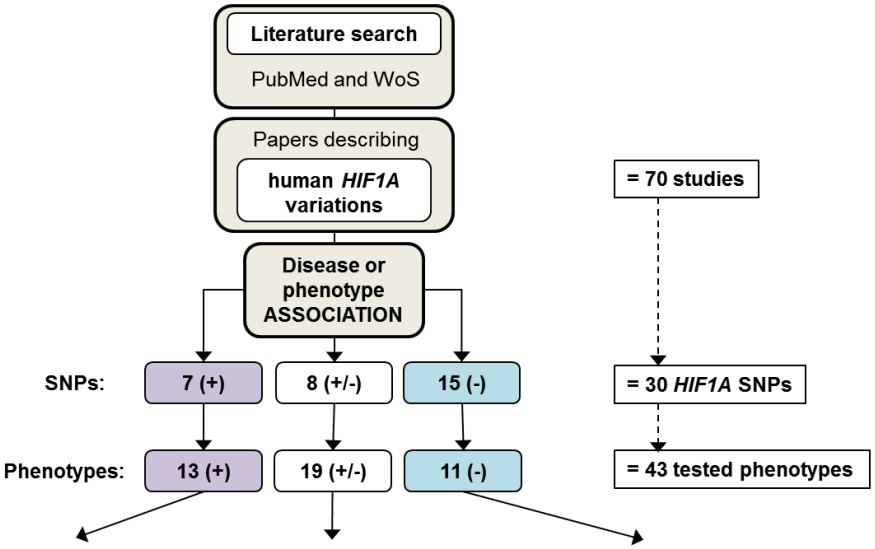
Figure 9: Atlas of *HIF1A* gene regulatory network

Slika 9: Atlas regulatorne mreže gena *HIF1A*

### 4.1.1 Literature review of HIF1A association studies

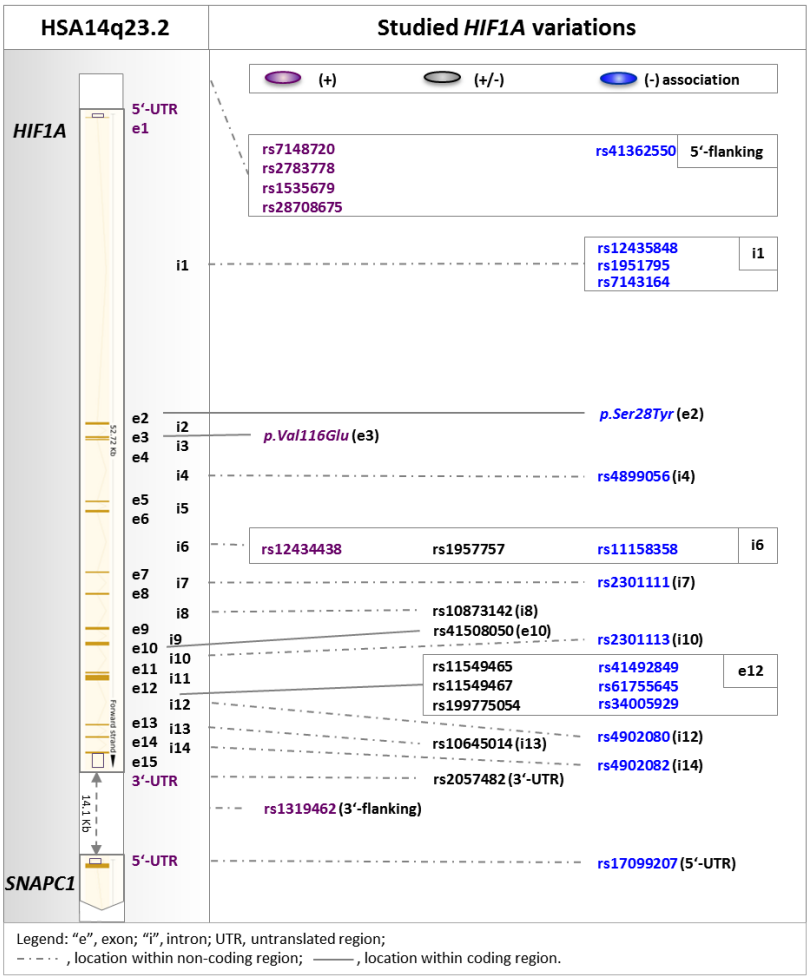
Literature review of 70 association studies in human revealed that to date 30 *HIF1A* SNPs were tested for an association to 43 phenotypes and among them 15 SNPs showed associations with 32 phenotypes including 16 cancer types (Figure 10A, ANNEX F, and ANNEX G). Ensembl nucleotide sequence of *HIF1A* is presented in ANNEX E and includes all known SNPs including 30 SNPs from association studies, which are highlighted. Studied *HIF1A* SNPs reside within all genomic regions: 5'-flanking, introns, exons, 3'-UTR, and 3'-flanking region (Figure 10B). Tested *HIF1A* genotype-phenotype associations were visualized in Figure 10C and D. Among 30 studied variations, eight SNPs showed variable associations with phenotypes (observed association, no association or opposing results) (ANNEX F), seven SNPs showed only associations with tested phenotypes, while for 15 SNPs no association with tested phenotype has been reported (ANNEX G). Among all SNPs, non-synonymous SNPs (nsSNPs) rs11549465 and rs11549467 within ODD domain were most frequently studied and for which in publications a various nomenclature was used (alternative SNPs names are presented in ANNEX F and ANNEX G). Moreover, several studies that tested associations of these two SNPs with phenotypes, showed conflicting results for the same cancer type (ANNEX F). For example, opposing results were reported for rs11549465 in prostate cancer susceptibility studies (with association (Chau et al., 2005; Foley et al., 2009; Fu et al., 2005; Jacobs et al., 2008; Orr-Urtreger et al., 2007) and no association (Li et al., 2007; Li et al., 2012)) and for rs11549467 in breast cancer risk studies (association (Zhao et al., 2009) and no association (Apaydin et al., 2008; Kim et al., 2008; Naidu et al., 2009)).

**(A)**

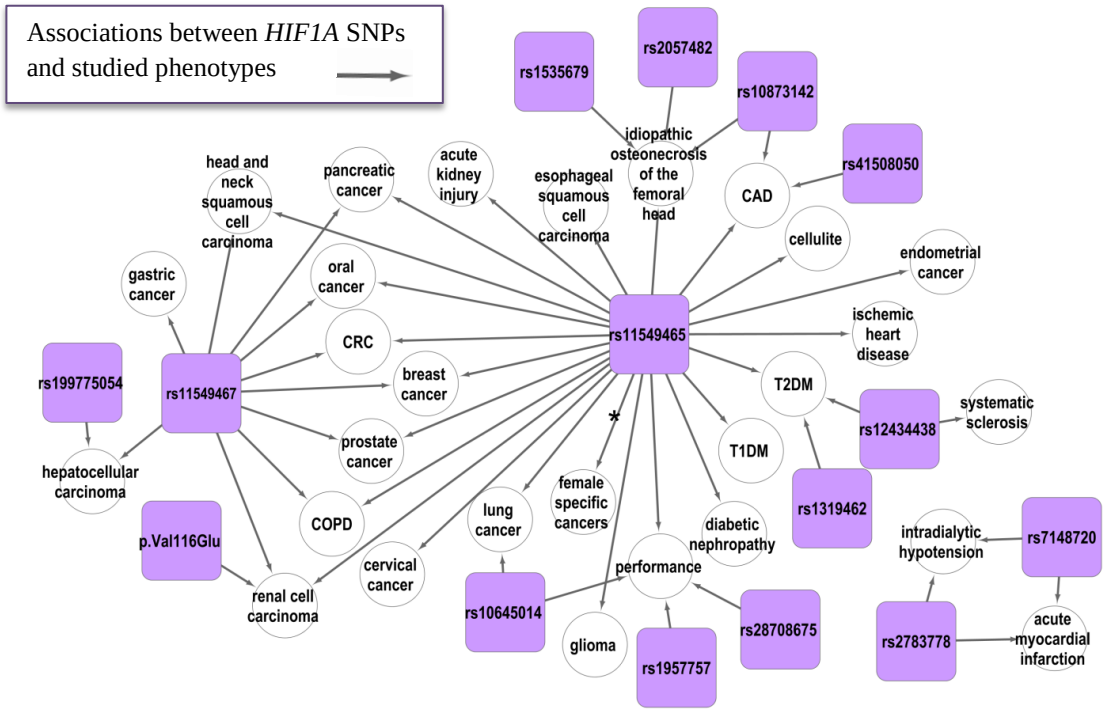



| | Association | | (+/-) | | No association |
|---|---|---|---|---|---|
| **1.** | acute kidney injury | **1.** | acute myocardial infarction | **1.** | acute leukemia |
| **2.** | adaptation to living at high altitude | **2.** | breast cancer | **2.** | acute mountain sickness symptoms |
| **3.** | cellulite | **3.** | cervical cancer | **3.** | bladder cancer |
| **4.** | chronic obstructive pulmonary disease | **4.** | CRC | **4.** | colon cancer |
| **5.** | diabetic nephropathy | **5.** | coronary artery disease with stable exertional angina | **5.** | erythrocytosis |
| **6.** | esophageal squamous cell carcinoma | **6.** | endometrial cancer | **6.** | giant cell arteritis |
| **7.** | female specific cancers * breast, endometrial, ovarian, cervical cancer | **7.** | endurance status | **7.** | ovarian cancer |
| **8.** | glioma | **8.** | frequent intradialytic hypotension | **8.** | peripheral artery disease |
| **9.** | head and neck squamous cell carcinoma | **9.** | gastric cancer | **9.** | pre-eclampsia |
| **10.** | idiopathic osteonecrosis of the femoral head | **10.** | hepatocellular carcinoma | **10.** | rectal cancer |
| **11.** | pancreatic cancer | **11.** | ischemic heart disease | **11.** | response to acute hypoxia |
| **12.** | symptom-limited exercise test duration over time | **12.** | lung cancer (NSCLC) | | |
| **13.** | T1DM | **13.** | maximal oxygen consumption | | |
| | | **14.** | oral squamous cell carcinoma & oral cancer | | |
| | | **15.** | power-oriented athletes | | |
| | | **16.** | prostate cancer | | |
| | | **17.** | renal cell carcinoma | | |
| * Meta-analysis | | **18.** | systematic sclerosis | | |
| | | **19.** | T2DM | | |

**(B)**


HSA14q23.2
Studied HIF1A variations
(+)
(+/-)
(-) association
HIF1A
5'-UTR
e1
rs7148720
rs2783778
rs1535679
rs28708675
rs41362550
5'-flanking
i1
rs12435848
rs1951795
rs7143164
e2
e3
e4
e5
e6
e7
e8
e9
e10
e11
e12
e13
e14
e15
3'-UTR
i2
i3
i4
i5
i6
i7
i8
i9
i10
i11
i12
i13
i14
p.Ser28Tyr (e2)
p.Val116Glu (e3)
rs4899056 (i4)
rs12434438
rs1957757
rs11158358
rs2301111 (i7)
rs10873142 (i8)
rs41508050 (e10)
rs2301113 (i10)
rs11549465
rs11549467
rs199775054
rs41492849
rs61755645
rs34005929
e12
rs4902080 (i12)
rs10645014 (i13)
rs4902082 (i14)
rs2057482 (3'-UTR)
rs1319462 (3'-flanking)
SNAPC1
5'-UTR
rs17099207 (5'-UTR)
Legend: "e", exon; "i", intron; UTR, untranslated region;
, location within non-coding region; , location within coding region.

**(C)**


Associations between HIF1A SNPs and studied phenotypes
rs1535679
rs2057482
rs10873142
rs41508050
idiopathic osteonecrosis of the femoral head
CAD
cellulite
endometrial cancer
ischemic heart disease
T2DM
rs12434438
systematic sclerosis
T1DM
rs1319462
intradialytic hypotension
rs7148720
acute myocardial infarction
rs2783778
rs28708675
rs1957757
glioma
rs10645014
performance
diabetic nephropathy
female specific cancers
lung cancer
cervical cancer
COPD
renal cell carcinoma
p.Val116Glu
rs199775054
hepatocellular carcinoma
rs11549467
rs11549465
head and neck squamous cell carcinoma
pancreatic cancer
acute kidney injury
esophageal squamous cell carcinoma
gastric cancer
oral cancer
CRC
breast cancer
prostate cancer
*

**(D)**

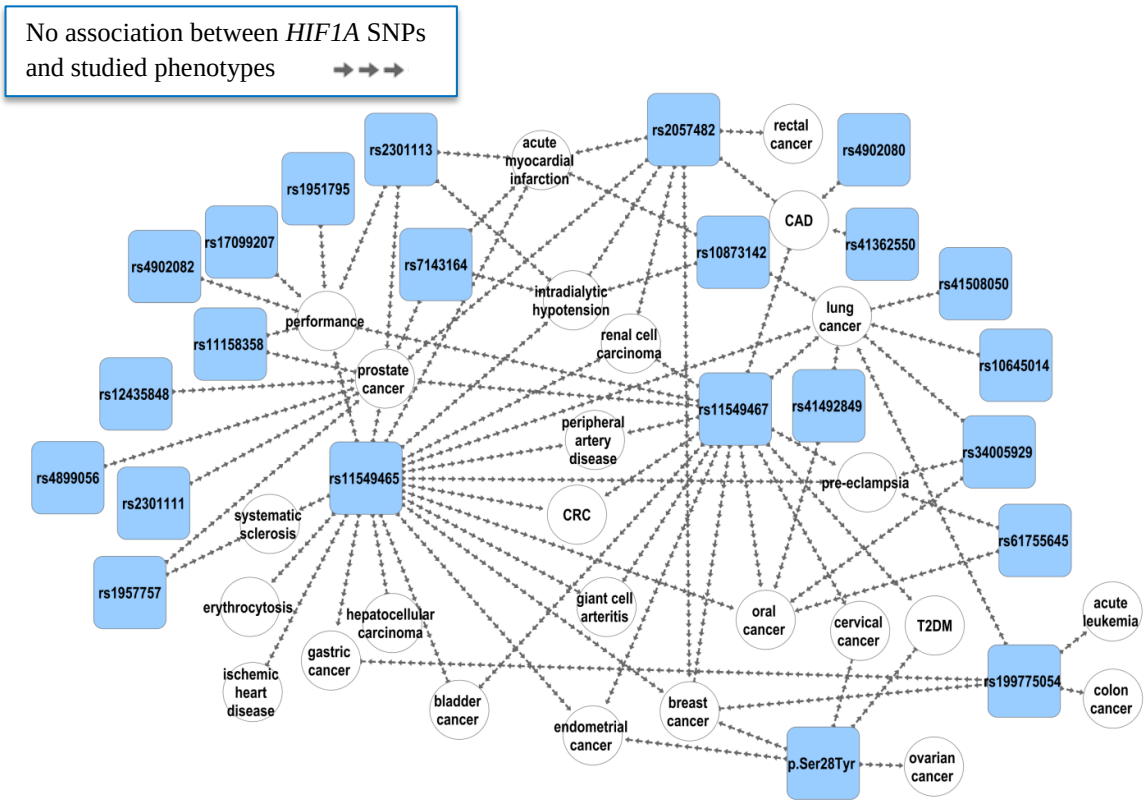



Figure legend: CAD for coronary artery disease with stable exertional angina; CRC, colorectal cancer; T1DM, diabetes mellitus type 1; T2DM, diabetes mellitus type 2; COPD, chronic obstructive pulmonary disease. Performance is a parental term for: adaptation to living at high altitude, endurance status, maximal oxygen consumption, power-oriented athletes, symptom-limited exercise test duration over time, and response to acute hypoxia. * Meta-analysis is including female specific cancers: breast, endometrial, ovarian and cervical.

Figure 10: HIF1A genotype-phenotype association studies in human

Slika 10: Analize povezav genotipa *HIF1A* s fenotipom pri človeku

Among 15 *HIF1A* SNPs associated with 32 different phenotypes five SNPs were associated with 16 cancer types: rs10654014 (Koukourakis et al., 2006), rs11549465 (Chau et al., 2005; Chen et al., 2009; Foley et al., 2009; Fransen et al., 2006; Fu et al., 2005; Hong et al., 2007; Jacobs et al., 2008; Kang et al., 2011; Kim et al., 2008; Konac et al., 2007; Koukourakis et al., 2006; Kuwai et al., 2004; Lee et al., 2008; Ling et al., 2005; Naidu et al., 2009; Ollerenshaw et al., 2004; Orr-Urtreger et al., 2007; Ruiz-Tovar et al., 2012; Tanimoto et al., 2003; Xu et al., 2011; Zhao et al., 2009), rs11549467 (Chen et al., 2009; Fransen et al., 2006; Hsiao et al., 2010; Li et al., 2009; Li et al., 2012; Munoz-Guerra et al., 2009; Ollerenshaw et al., 2004; Ruiz-Tovar et al., 2012; Tanimoto et al., 2003; Zhao et al., 2009), *p.Val116Glu* (Morris et al., 2009) and rs199775054 (Park et al., 2009) (ANNEX F and ANNEX G). Those five SNPs were most frequently associated with breast, lung, CRC, gastric, prostate, oral cancer, and

RCC. Genetic variability of *HIF1A* was also found to be associated with cardiovascular system diseases like: ischemic heart disease, coronary artery disease (CAD) with stable exertional angina, acute myocardial infarction and frequent intradialytic hypotension (ANNEX F and ANNEX G). Polymorphisms associated with cardiovascular system were: rs10873142 (Hlatky et al., 2007), rs11549465 (Hlatky et al., 2007; Resar et al., 2005), rs2783778 (Zheng et al., 2009), rs41508050 (Hlatky et al., 2007), and rs7148720 (Gaudio et al., 2012). Since cardiovascular system subsequently influences human body performance, variations associated with one or more of the following performance related phenotypes: maximal oxygen consumption, adaptation to living at high altitude, idiopathic osteonecrosis of the femoral head, power-oriented athlete performance and muscle activity, endurance status, and symptom-limited exercise test duration over time; were: rs10645014 (Suzuki et al., 2003), rs10873142 (Hong et al., 2007), rs11549465 (Ahmetov et al., 2008; Cieszczyk et al., 2011; Doring et al., 2010; Gabbasov et al., 2012; Hong et al., 2007; McPhee et al., 2011; Prior et al., 2003), rs1535679 (Hong et al., 2007), rs1957757 (Sarzynski et al., 2010), rs2057482 (Hong et al., 2007), and rs28708675 (Prior et al., 2003) (ANNEX F and ANNEX G). Some polymorphisms were also associated with other non-cancerous phenotypes such as systemic sclerosis (rs12434438 (Wipff et al., 2009)), acute kidney injury (rs11549465 (Kolyada et al., 2009)), cellulite (rs11549465 (Emanuele et al., 2010)), and chronic obstructive pulmonary disease (COPD) (rs11549465 (Putra et al., 2011a; Putra et al., 2011b)). Moreover, *HIF1A* SNPs were shown to be involved in metabolic disorders such as: diabetic nephropathy (rs11549465 (Gu et al., 2012)), which is a result of longstanding diabetes mellitus type 1 (T1DM) (rs11549465 (Nagy et al., 2009)), and type 2 (T2DM) (rs11549465 (Nagy et al., 2009; Yamada et al., 2005), rs12434438 (Yamada et al., 2005), and rs1319462 (Yamada et al., 2005)). About half of *HIF1A* SNPs genotyped to date are intronic (13/30), however according to the Ensembl none of them is defined as a splice region variant. Our predictions also showed that coding genotyped *HIF1A* SNPs (9/30) may affect the HIF1A stability or its function as TF.

#### 4.1.1.1 Re-opposing results observed in *HIF1A* association studies

In the group of SNPs for which variable associations were observed, rs1957757 showed an association with limited exercise duration in adults (Sarzynski et al., 2010), but no involvement with prostate cancer or systematic sclerosis (Jacobs et al., 2008; Wipff et al., 2009). Likewise, rs41508050 was associated with CAD with stable exertional angina (Hlatky et al., 2007), but not with lung cancer (Konac et al., 2009), and rs199775054 involvement was confirmed in hepatocellular cancer occurrence, while no such association was seen in other cancers like: CRC, gastric, breast, lung cancer or acute leukemia (Park et al., 2009).

Meta-analysis nsSNP within ODD domain concluded that rs11549465 has its role in appearance of female specific cancers (Zhao et al., 2009). Most often rs11549465 was

tested to determine potential functional role in cancers like: prostate, oral, breast cancer, and CRC. Moreover, many studies, where they were testing rs11549465, presented opposing results of association with prostate (+ (Chau et al., 2005; Foley et al., 2009; Fu et al., 2005; Jacobs et al., 2008; Orr-Urtreger et al., 2007) *vs.* − (Li et al., 2007)), oral (+ (Chen et al., 2009) *vs.* − (Munoz-Guerra et al., 2009; Shieh et al., 2010)), endometrial (+ (Konac et al., 2007) *vs.* − (Horree et al., 2008)), breast (+ (Kim et al., 2008; Lee et al., 2008; Naidu et al., 2009) *vs.* − (Apaydin et al., 2008; Vleugel et al., 2005; Zagouri et al., 2012)), lung (+ (Koukourakis et al., 2006) *vs.* − (Konac et al., 2009; Kuo et al., 2012)), renal (+ (Ollerenshaw et al., 2004) *vs.* − (Morris et al., 2009; Qin et al., 2011)), and CRC (+ (Fransen et al., 2006; Kang et al., 2011; Kuwai et al., 2004) *vs.* − (Knechtel et al., 2010; Szkandera et al., 2010)). Associations were confirmed for glioma cancer (Xu et al., 2011), pancreatic cancer (Ruiz-Tovar et al., 2012), esophageal (Ling et al., 2005) and head & neck squamous cell carcinoma (Tanimoto et al., 2003). Other non-cancerous phenotypes were also tested with rs11549465 effect on athlete performance capability (grater gains in endurance training, muscle activity effect, $O_2$-consumption), diabetes (T1DM and T2DM), CAD with stable exertional angina, erythrocytosis, pre-eclampsia, and also an effect on cellulite appearance etc. Similarly, opposing associations were observed with the second nsSNP within ODD domain, rs11549467 (G>A), and breast cancer (+ (Zhao et al., 2009) *vs.* − (Apaydin et al., 2008; Kim et al., 2008; Naidu et al., 2009)), CRC (+ (Fransen et al., 2006) *vs.* − (Knechtel et al., 2010; Szkandera et al., 2010), oral squamous cell carcinoma (OSCC) (+ (Chen et al., 2009; Munoz-Guerra et al., 2009) *vs.* − (Shieh et al., 2010)) and renal cell carcinoma (RCC) (+ (Ollerenshaw et al., 2004) *vs.* − (Morris et al., 2009; Qin et al., 2011)). Lung, prostate, ovarian and bladder cancer tested showed no relationship with rs11549467, while head & neck squamous cell carcinoma, gastric, pancreatic, hepatocellular cancer and susceptibility to COPD were apparently related with analyzed rs11549467. Other SNPs were associated as following: rs2057482 within 3'-UTR was associated with idiopathic osteonecrosis of the femoral head in men (Hong et al., 2007), intronic SNP rs10873142 (intron 8) was linked to CAD (Hlatky et al., 2007) and idiopathic osteonecrosis of the femoral head (Hong et al., 2007), while rs10645014 (intron 13) was associated with non-small cell lung carcinomas (Koukourakis et al., 2006) and possible adaptation to living at high altitude (Suzuki et al., 2003).

### 4.1.2 Transcription factor and miRNA binding site analysis

We investigated if *HIF1A* SNPs either create new or delete existing TF and miRNA binding sites, through which the stability or transcription of mRNA could be affected. Therefore, in addition to previously published TF and miRNA interactions with *HIF1A*, we performed predictions if TF and miRNA binding sites overlap with *HIF1A* SNPs tested in human genotype-phenotype association studies.

#### 4.1.2.1 Transcription factor binding site analysis

To elucidate *HIF1A* regulation by TF, we integrated an existing knowledge of TFBS upstream or nearby *HIF1A* TSS and also performed bioinformatics predictions if 5'-flanking *HIF1A* SNPs affect gain/loss of TFBSs. Literature search related to TFs regulating *HIF1A* mRNA (SP1 (Minet et al., 1999), NFKB1/RELA (Belaiba et al., 2007) and EGR1 (Sperandio et al., 2009)) and extraction of validated TFBS from the ENCODE project ChIP-seq data showed that 62 TFBSs locate nearby *HIF1A* TSS (-654 bp and +1073 bp) (ANNEX H). Four out of five SNPs located within 5' upstream *HIF1A* region were previously associated with phenotypes tested in association studies. However, polymorphism rs41362550, whose functional effect is not known, was the only one overlapping with TFBSs (42/62). Additionally, bioinformatics prediction showed that four of five SNPs were predicted to be accountable for gain/loss of six TFBSs: rs2783778 (for TAL1 and TCF3), rs1535679 (for FOXC1), rs25708675 (for FOXA1 and FOXI1), and rs41362550 (for HNF4A) (Figure 11 and ANNEX I**)**.

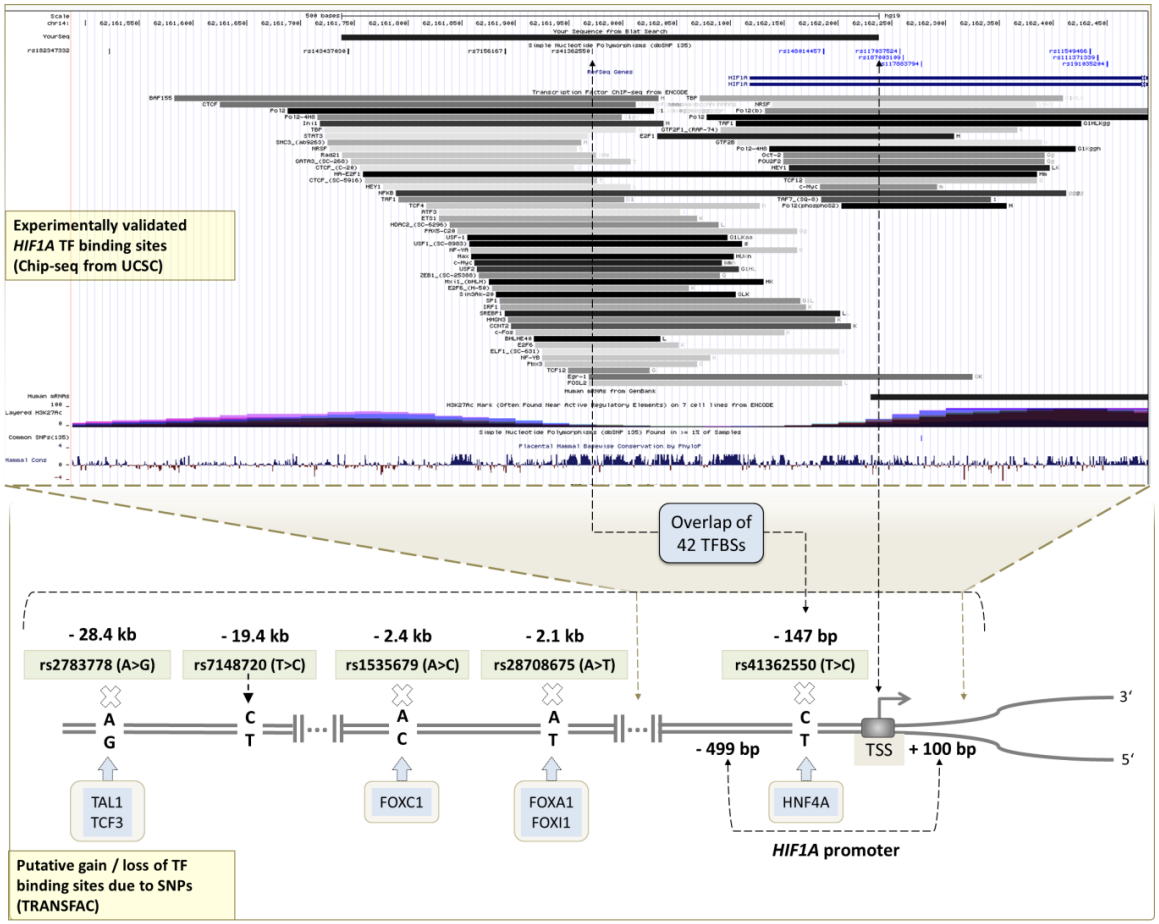



Figure legend: FOXA1, forkhead box A1; FOXC1, forkhead box C1; FOXI1 (forkhead box I1); HNF4A, hepatocyte nuclear factor 4, alpha; TCF3, transcription factor 3 (E2A immunoglobulin enhancer binding factors E12/E47); TAL1, T-cell acute lymphocytic leukemia 1; kb, kilobase; bp, base pair; TSS, transcription start site.

Figure 11: *HIF1A* SNPs within 5'-region affecting gain/loss of transcription factor binding sites. Overlap between experimentally confirmed (UCSC) and *in silico* predicted TFBS (TRANSFAC)

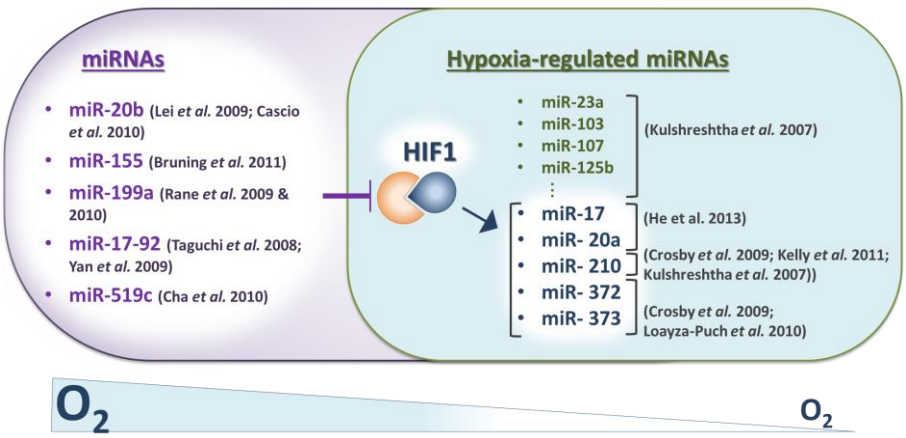
Slika 11: Vpliv *HIF1A* SNP-jev v 5'-regiji gena na pridobitev/izgubo vezavnih mest za transkripcijske dejavnike. Prekrivanje eksperimentalno potrjenih (UCSC) in računalniško predvidenih (TRANSFAC) vezavnih mest za transkripcijske dejavnike

### 4.1.2.2 MicroRNA target site analysis

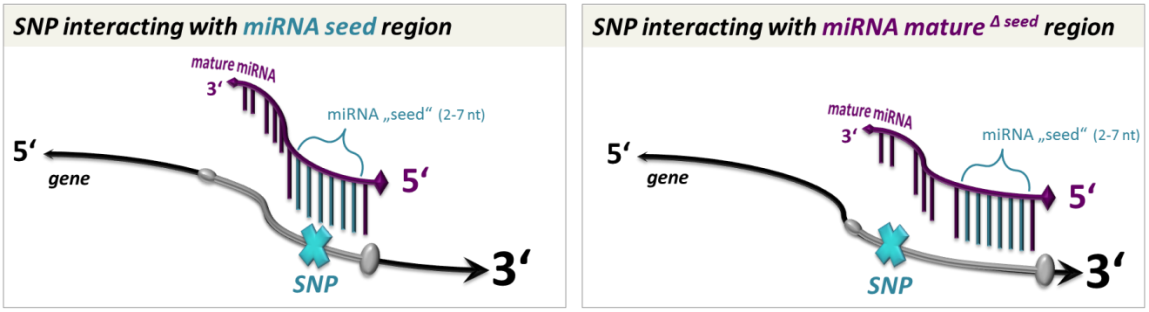
Our literature review revealed that *HIF1A* expression was previously validated to be regulated by miRNAs and by now miR-20b (Cascio et al., 2010; Lei et al., 2009), miR-199a (Rane et al., 2009; Rane et al., 2010), miR-519c (Cha et al., 2010), miR-155 (Bruning et al., 2011), and cluster miR-17-92 (Taguchi et al., 2008; Yan et al., 2009) were experimentally confirmed to target *HIF1A* and regulate its expression using various assays including luciferase reporter assay, western blot, ELISA, ChIP, microarray, RT-qPCR, or northern blot, but however, some miRNAs were also validated to be regulated by HIF1A; miR-17 (He et al., 2013), miR-20a (He et al.,

2013), miR-210 (Crosby et al., 2009b; Kelly et al., 2011; Kulshreshtha et al., 2007), miR-372 (Loayza-Puch et al., 2010) and miR-373 (Crosby et al., 2009b; Loayza-Puch et al., 2010)) (Figure 12A). In previous studies, 3'-UTR of *HIF1A* was tested for miRNA interactions and none of to date validated miRNAs target sites overlapped with tested *HIF1A* SNPs within 3'-UTR. However, the standard ''dogma'' that miRNAs bind to a site in the 3'-UTR of a gene has been extended with observations that in addition to 3'-UTR, miRNAs also bind to promoter region, 5'-UTR, and coding regions (reviewed in (Kunej et al., 2012)). Thus, we examined if putative miRNA binding sites overlap with 17 non-intronic *HIF1A* SNPs genotyped in human association studies: five SNPs within 5'-flanking region, nine exonic SNPs, and three SNPs in 3'-flanking region. The prediction analysis resulted in gain/loss of miRNA-target sites, which belonged to 150 unique human mature miRNAs. In the Figure 12B SNP location complementary to miRNA seed region (left) and complementary to mature miRNA region outside of seed (miRNA mature Δ seed) (right) is presented. Effects of 17 *HIF1A* SNPs on gain/loss (miRNA names in red/blue) of putative miRNAs target sites are presented in (Figure 12C). Each graph is presenting alteration of putative miRNA binding sites due to miRNA overlapping SNP with seed region (white box) or mature region excluding seed (grey box). Allele variant of a SNP within seed region is predicted to effect binding of miRNAs written under each graph. Due to SNPs located within *HIF1A* target region complementary to miRNA seed region (first 2-7 nt of mature miRNA) 28/35 miRNA-target sites were gained/lost (Figure 12C). Our prediction analysis suggested that *HIF1A* SNPs are involved in gain/loss of miRNA-target sites located within exonic, 3'-, and 5'-UTR regions (Figure 12C and Figure 12D). In the Figure 12D a detailed presentation of three examples from the Figure 12C is presented. Among miRNAs with gained/lost binding sites miR-23a and -23b were previously described as HRMs (Figure 12C) (Kulshreshtha et al., 2007). Our prediction analysis showed that these two miRNAs were predicted to lose binding site due to SNP *p.Val116Glu* within ARNT (PAS) binding domain in exon 3. This *HIF1A* SNP is of interest for further investigations, since HIF1A is functional as transcriptional regulator only when dimerized with ARNT.

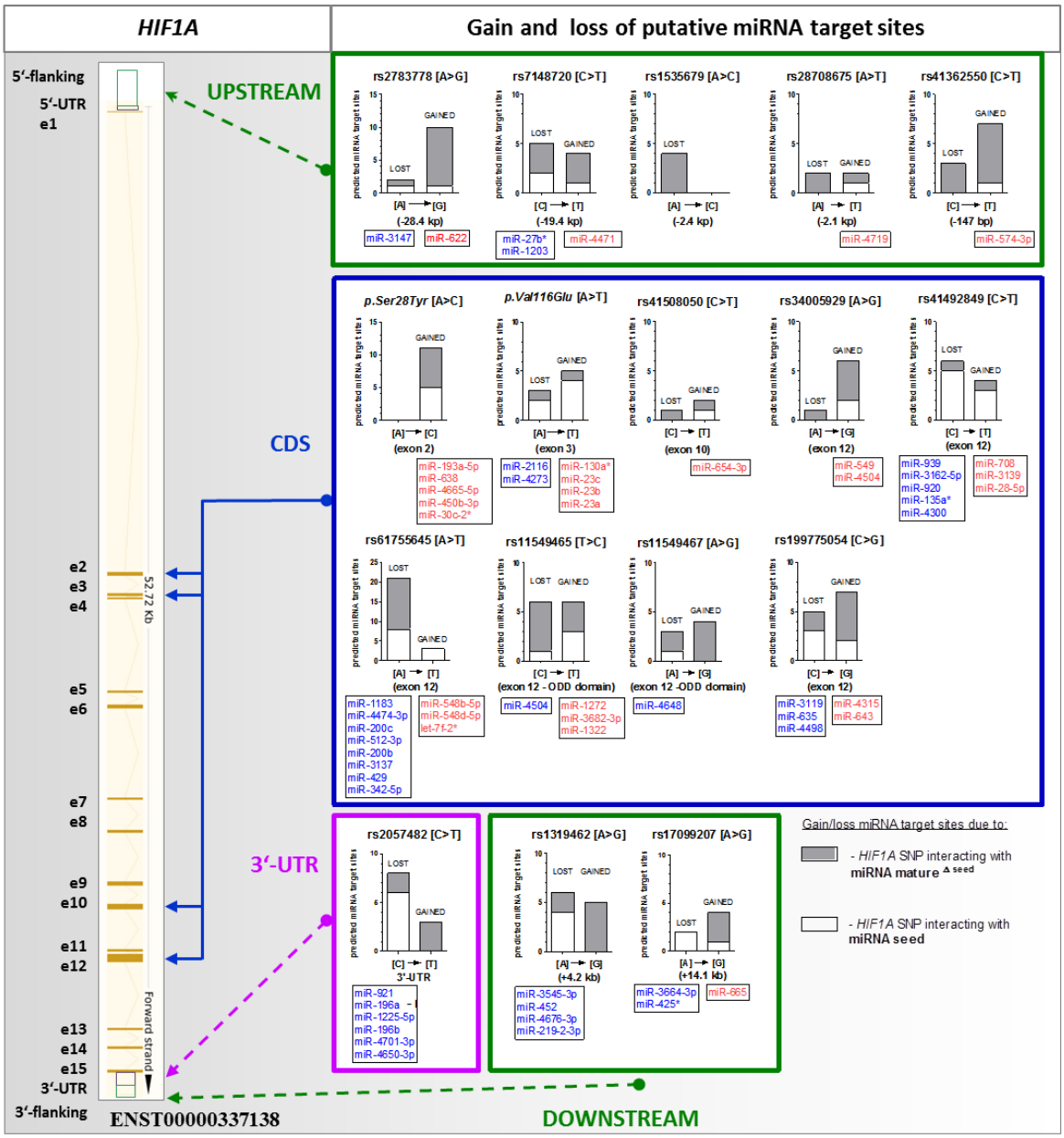
**(A)**



**(B)**


SNP interacting with miRNA seed region
mature miRNA
3'
miRNA „seed" (2-7 nt)
5'
gene
5'
3'
SNP
SNP interacting with miRNA mature Δ seed region
mature miRNA
3'
miRNA „seed" (2-7 nt)
5'
gene
5'
3'
SNP

**(C)**


HIF1A
Gain and loss of putative miRNA target sites
5'-flanking
5'-UTR
e1
UPSTREAM
CDS
e2
e3
e4
e5
e6
e7
e8
e9
e10
e11
e12
e13
e14
e15
3'-UTR
3'-flanking
52.72 Kb
Forward strand
ENST00000337138
3'-UTR
DOWNSTREAM
rs2783778 [A>G]
rs7148720 [C>T]
rs1535679 [A>C]
rs28708675 [A>T]
rs41362550 [C>T]
p.Ser28Tyr [A>C]
p.Val116Glu [A>T]
rs41508050 [C>T]
rs34005929 [A>G]
rs41492849 [C>T]
rs61755645 [A>T]
rs11549465 [T>C]
rs11549467 [A>G]
rs199775054 [C>G]
rs2057482 [C>T]
rs1319462 [A>G]
rs17099207 [A>G]
Gain/loss miRNA target sites due to:
- HIF1A SNP interacting with miRNA mature Δ seed
- HIF1A SNP interacting with miRNA seed

**(D)**

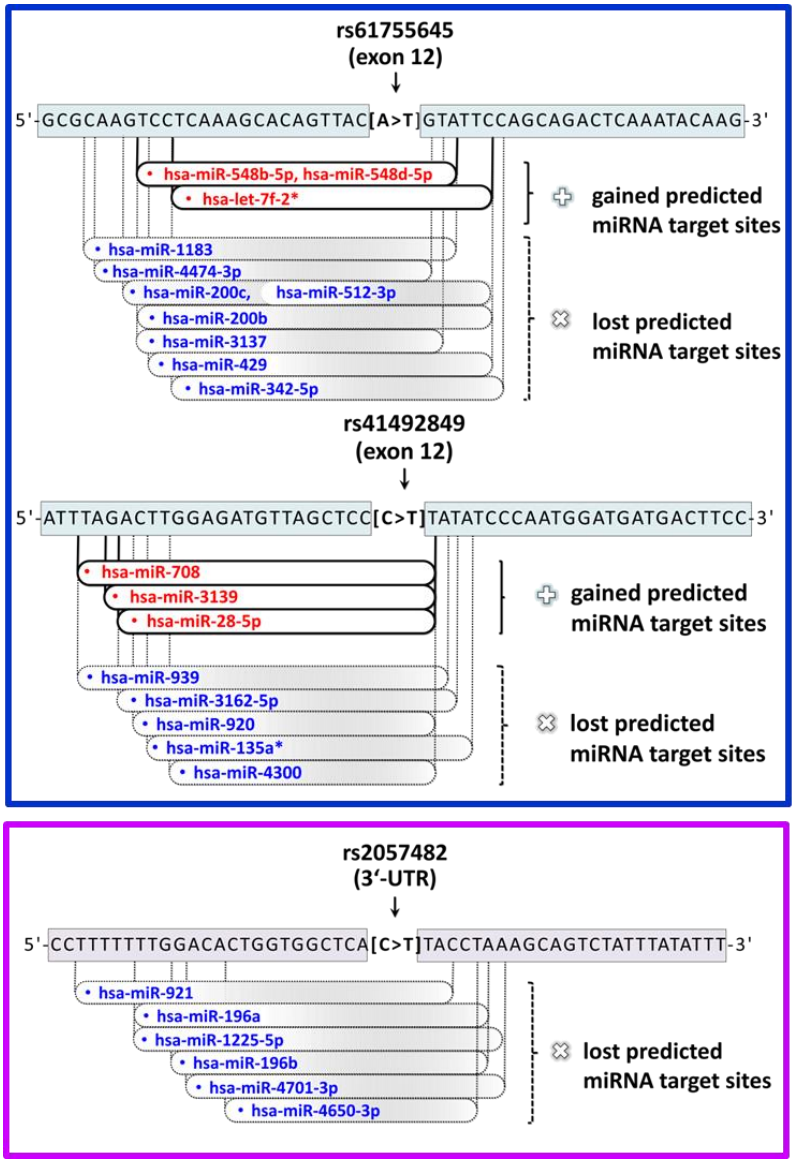



Figure legend: "e", exon; UTR, untranslated region; CDS, coding DNA sequence; ODD domain, $O_2$-dependent degradation domain; broken line, location of SNPs within noncoding region; continual line, location within of SNPs within coding region; miRNA target gain/loss due to SNP within coding-sequence (blue frame) and 3'-UTR (violet frame) of a gene.

Figure 12: Known and putative miRNAs::*HIF1A* interactions involved in HIF1A regulatory network

Slike 12: Znane in predvidene interakcije miRNA::*HIF1A* vpletene v regulatorno mrežo HIF1A

#### 4.1.2.3 Prioritization of miRNAs predicted to target *HIF1A* for further studies

MicroRNAs were proved to be involved in cancer development and progression (Calin and Croce, 2006; Esquela-Kerscher and Slack, 2006). Their profiling in various cancer tissues is used to characterize diversified solid tumors. A plethora of published studies have reported differences in miRNA profiles of cancer cells. In the review, Ferdin et al. (2010), we discussed a general mechanism of miRNA function and how their function is affected by cytogenetic alterations. We identified genomic regions containing deregulated or epigenetically regulated miRNAs in cancer, since regions frequently deleted or amplified in cancer often code for miRNA. A PubMed database search for expression studies linking miRNA genes with cancer development (ANNEX J) revealed clusters of as few as three miRNAs that had been previously associated with cancer in several expression studies (ANNEX M with a list of studies in ANNEX N) (Ferdin et al., 2011). The genomic locations most frequently associated with cancer were distributed among all chromosomes but more often among seven separate chromosomes presented in ANNEX M (some with two locations per chromosome: 4q25, 7q32.2, 9q22.32, 13q31.3, 14q32.2 and 14q32.13, 19p13.13 and 19q13.42, Xq26.2 and Xq27.3). Besides we identified that many epigenetically regulated miRNAs genes were located on the long arms of chromosomes 1, 7, 11, 14, and 19 (Kunej et al., 2011).

In order to identify which miRNAs can potentially be upstream regulators of *HIF1A*, in addition to validated ones (Figure 13), we prioritized the list of 150 miRNA obtained from the prediction analysis of miRNA::*HIF1A* interactions according to ours and others publication overviews about miRNA deregulated in cancer and hypoxia environment. Therefore, we used the list of miRNAs that were previously reported as up- or downregulated in different types of cancer tissue or cells (ANNEX J, ANNEX K and ANNEX L) (Ferdin et al., 2010), the list of miRNAs proved to be epigenetically regulated in cancer (Kunej et al., 2011), and the list of miRNAs that are currently known as hypoxia regulated (Kulshreshtha et al., 2007; Nallamshetty et al., 2013).

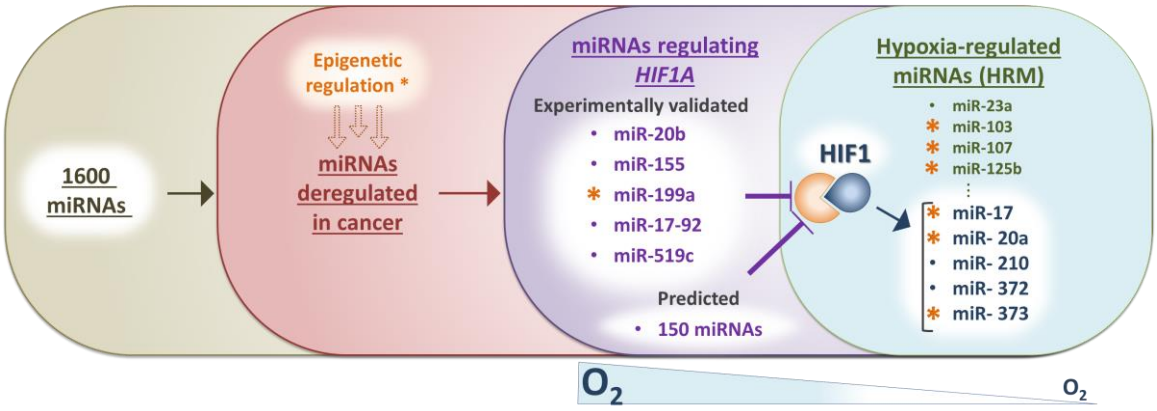



Figure 13: Involvement of miRNAs in HIF1A regulatory network in cancer

Slika 13: Vpletenost miRNA v regulatorno mrežo HIF1A pri raku

The prioritization of 150 miRNAs predicted to target *HIF1A* using information obtained from the literature reviews revealed four miRNAs (miR-23b, miR-130a, miR-135a, and miR-200b), which were previously already identified as cancer deregulated, epigenetically regulated and deregulated in hypoxia (Figure 14).

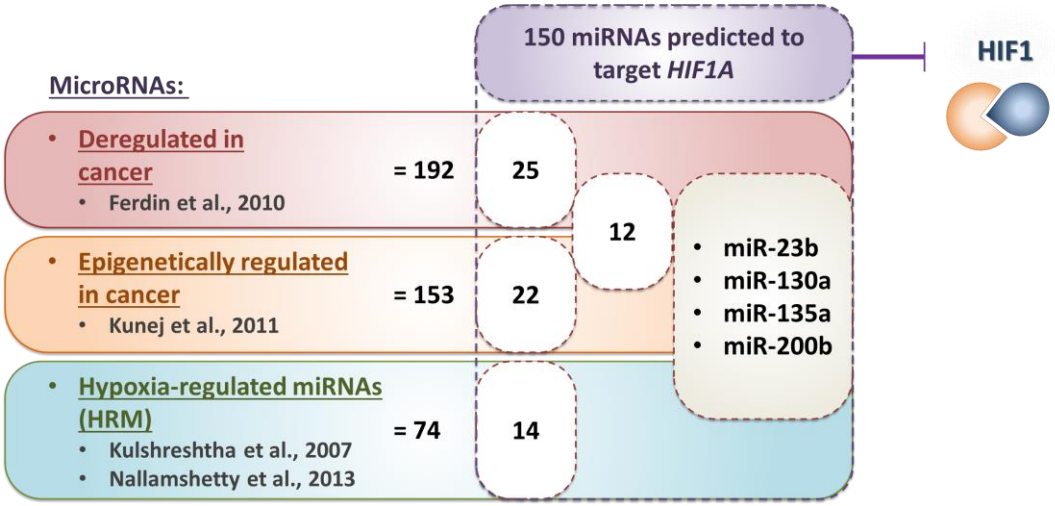


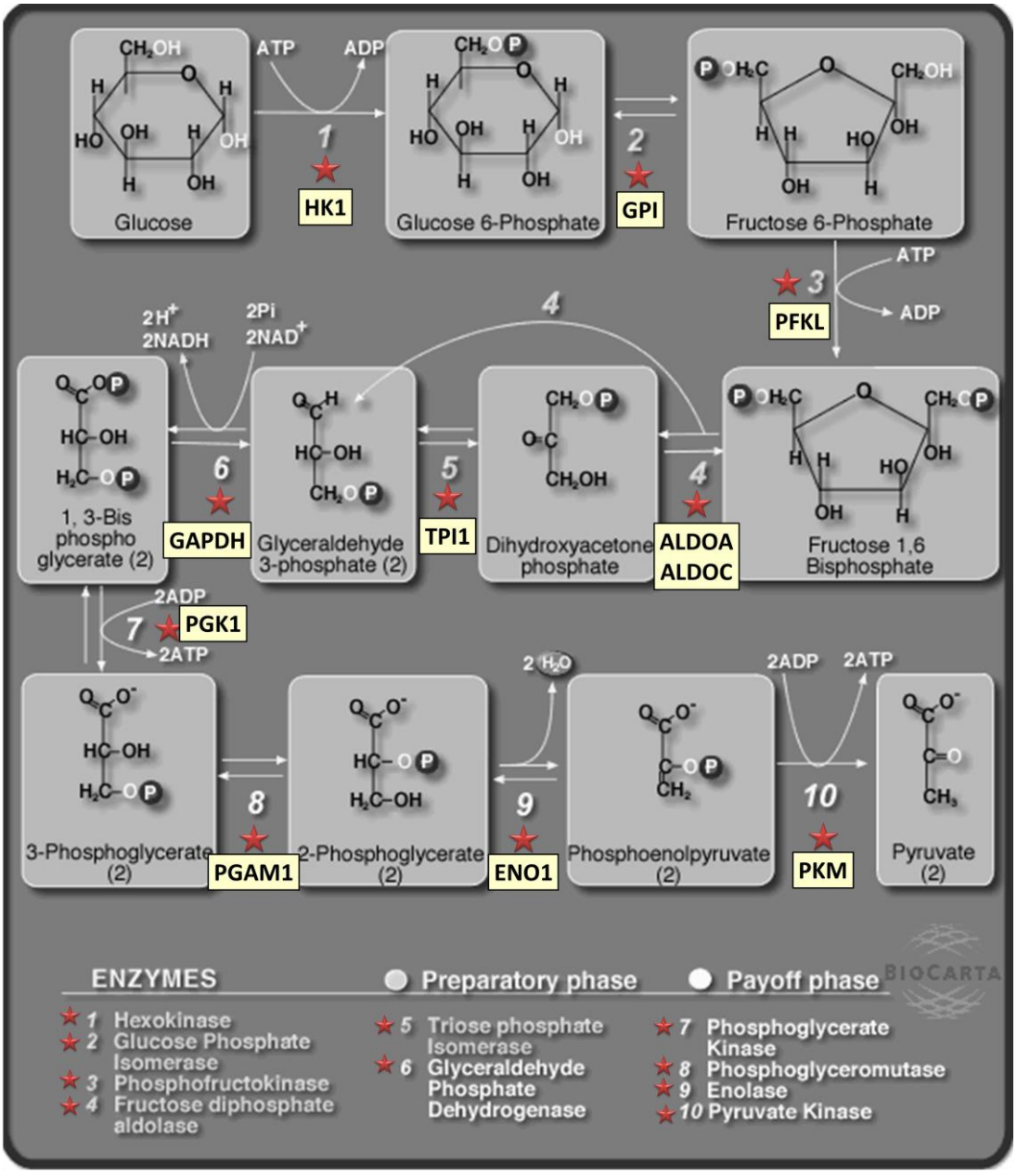

Figure 14: MicroRNAs predicted to be *HIF1A*-upstream regulators

Slika 14: Mikro RNA, za katere je predvideno, da uravnavajo izražanje gena *HIF1A*

### 4.1.3 Pathway enrichment analysis of HIF1A downstream targets

Literature mining revealed 274 experimentally validated direct HIF1A targets including 268 protein-coding genes, five miRNA genes (miR-17 (He et al., 2013), miR-20a (He et al., 2013), miR-210 (Crosby et al., 2009b; Kelly et al., 2011; Kulshreshtha et al., 2007), miR-372 (Loayza-Puch et al., 2010), miR-373 (Crosby et al., 2009b; Loayza-Puch et al., 2010)) and one T-UCR (uc.475 (Ferdin et al., 2013)). Pathway enrichment analysis, conducted by applying very stringent criteria (Bonferroni multiple test correction and *P*-value significant threshold of 0.01), yielded the presence of 21 significant pathways associated with HIF1A targets (ANNEX O). The highest number of HIF1A targets was found in: *diabetes pathways* (N=27), *metabolism of carbohydrates* (N=25), and *integration of energy metabolism* (N=19). Pathway diagrams of *glycolysis* and of *renal cell carcinoma* with marked HIF1A targets are visualized in Figure 15**.** In Figure 15A modified BioCarta diagram of *glycolysis* in human is showing that all 10 pathway components are HIF1A targets, while in Figure 15B modified KEGG diagram of *renal cell carcinoma* in human is showing that out of 51 pathway components 13 are HIF1A targets. The rest of the genes from those pathways not yet shown to be HIF1A targets, present novel candidates to be regulated by HIF1 within promoter region or in protein-protein interactions.

**(A)**




★ - HIF1A targets

**(B)**

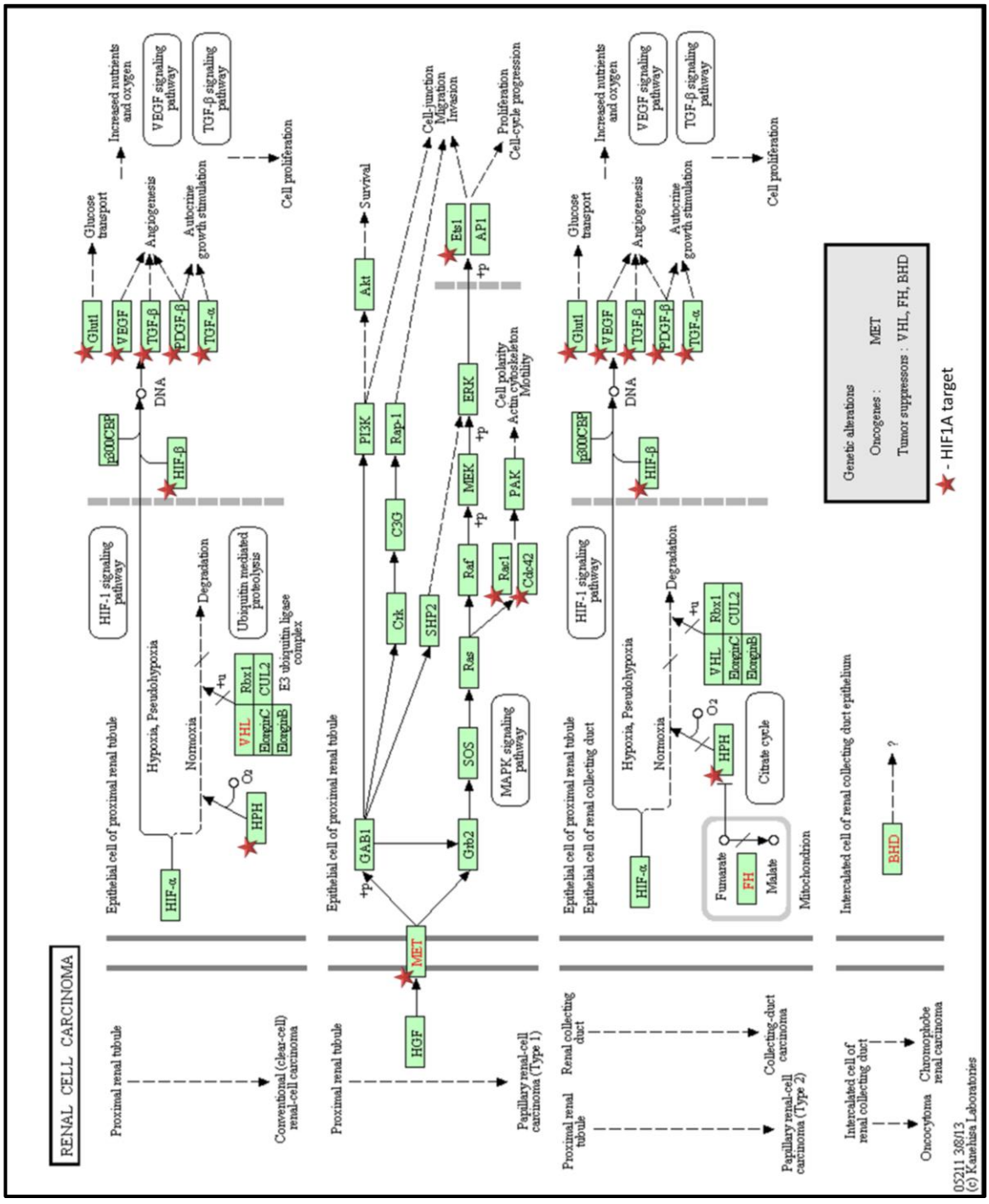



Figure 15: Two pathway diagrams containing known HIF1A targets

Figure 15: Shema dveh bioloških poti, ki vsebujeta znane HIF1A tarče

## 4.2 VALIDATION OF A NOVEL HIF1A TARGET

In cancer environment level of $O_2$ gets decreased by the depth and distance of blood flow in vessels. Previously studies showed that miRNA are among ncRNA already known to be regulated by low $O_2$ levels. Kulshreshtha et al. (2007) performed study on cancer cell lines, breast and colon, and confirmed that a group of miRNAs (HRM, hypoxia regulated miRNAs) can be influenced by variety of hypoxic factors and were proved to decrease pro-apoptotic signaling in a hypoxic environment, suggesting an impact of these transcripts even on tumor formation (Kulshreshtha et al., 2007). Nevertheless, it has been suggested that HIF1A gives a big contribution to cancer cell line adaptation in hypoxia. Microarray assays on clinical CRC samples revealed altered expression not only for miRNAs but also for T-UCRs (Calin et al., 2007; Ferdin et al., 2013). Consecutively our idea was to analyze if hypoxia and their factors influence on expression of T-UCRs.

Whereas ultraconserved sequences appear to act as tissue-specific regulators of gene expression during development (Bejerano et al., 2004b; Mattick, 2009; Pennacchio et al., 2006), 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 the original work of Kulshreshtha et al. on hypoxia-regulated miRNAs (Kulshreshtha et al., 2007). They initiated the search by exposing several cancer cell lines to low $O_2$ followed by quantification of changes in miRNA expressions *versus* normoxic controls using a genome-wide microarray profiling, while we initiated our study from the genome wide expression profiles of T-UCRs obtained from CRC samples. Subsequently, we interrogated the contribution of 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.

### 4.2.1 T-UCR-expression changes under hypoxic conditions in malignant cells

We re-interrogated previous microarray data (Array Expression accession: E-TABM-184) obtained from CRC samples (78 CRCs compared with 21 normal colonic mucosas) (Calin et al., 2007) to identify overexpressed T-UCRs in CRC that could be potentially induced by hypoxia (for a detailed workflow, see Figure 16). The same type of microarrays was used for the discovery of hypoxia-induced miRNAs (Kulshreshtha et al., 2007). 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 (ANNEX A).

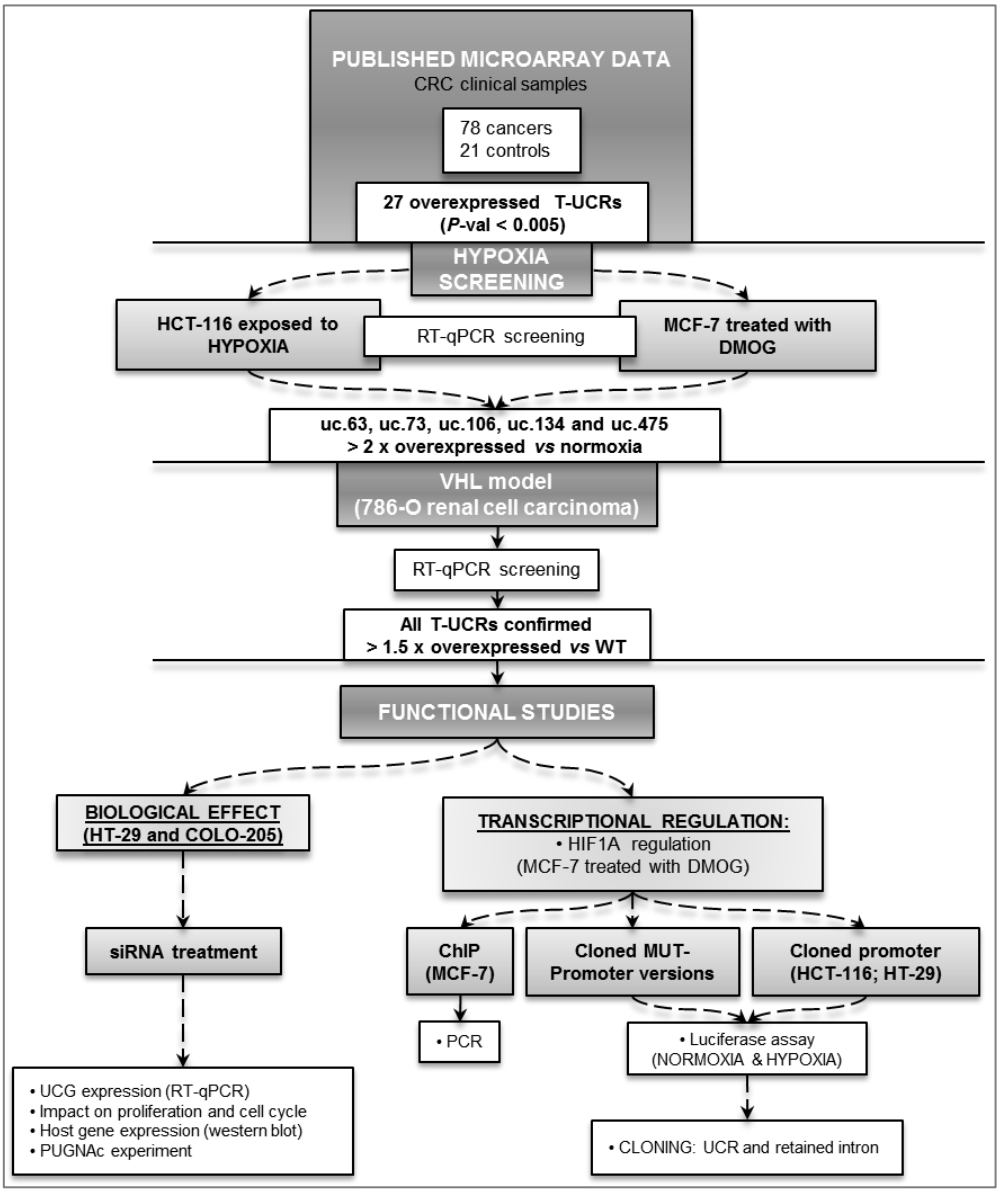



Figure 16: Workflow of the molecular genetic part of the study

Slika 16: Potek molekularno-genetskega dela raziskave

In order to validate the microarray data, we exposed two cancer cell lines (HCT-116 CRC and MCF-7 breast cancer cells) to 1% $O_2$ or 500 μM DMOG - a widely used hypoxia mimetic - for 24 and 48 h and measured T-UCR expressions using RT-qPCR. Comparisons were performed with normoxic controls (21% $O_2$) or 0.5% DMSO, respectively. 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 (Table 4 and Figure 17A). The RT-qPCR was used for confirmation of the T-UCR-expression induction in HCT-116 colon cancer cells and hypoxia-simulation conditions by DMOG in the MCF-7 breast cancer cell line. The T-UCR expression was normalized *versus* normal conditions, which was normoxia in HCT-116 and DMSO in MCF-7.

Table 4: Genomic information for the five HINCUTs identified in the present study and their differential expression in two cancer cell lines

Tabela 4: Genomske informacije petih prepisov HINCUT identificiranih v raziskavi in njihovo spremenjeno izražanje v dveh celičnih linijah raka

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

* Host gene orientation is reported versus HINCUT orientation considered as reference. 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 names for the T-UCRs as for the microarray probes found in E-TABM-184.

** Fold change of T-UCR expressions in hypoxia/hypoxia mimics compared to normoxia/normal conditions.

All of these transcripts were predominantly present in the nucleus, as identified by nuclear and cytoplasmic fractionation experiments performed on HCT-116 cell lines (Figure 18). Robust activation of hypoxia signaling was confirmed by HIF1A (or HIF2A) protein overexpression with HIF1A (or HIF2A) specific antibodies (Figure 17A and B) and *CA9* (data not shown) or hsa-miR-210 (Figure 19) gene expression levels, which are all well-established coding and noncoding HIF targets, respectively (Kulshreshtha et al., 2007). Henceforth, these T-UCRs will be referred to as HINCUTs; their functional dissection under low $O_2$ conditions is described below.

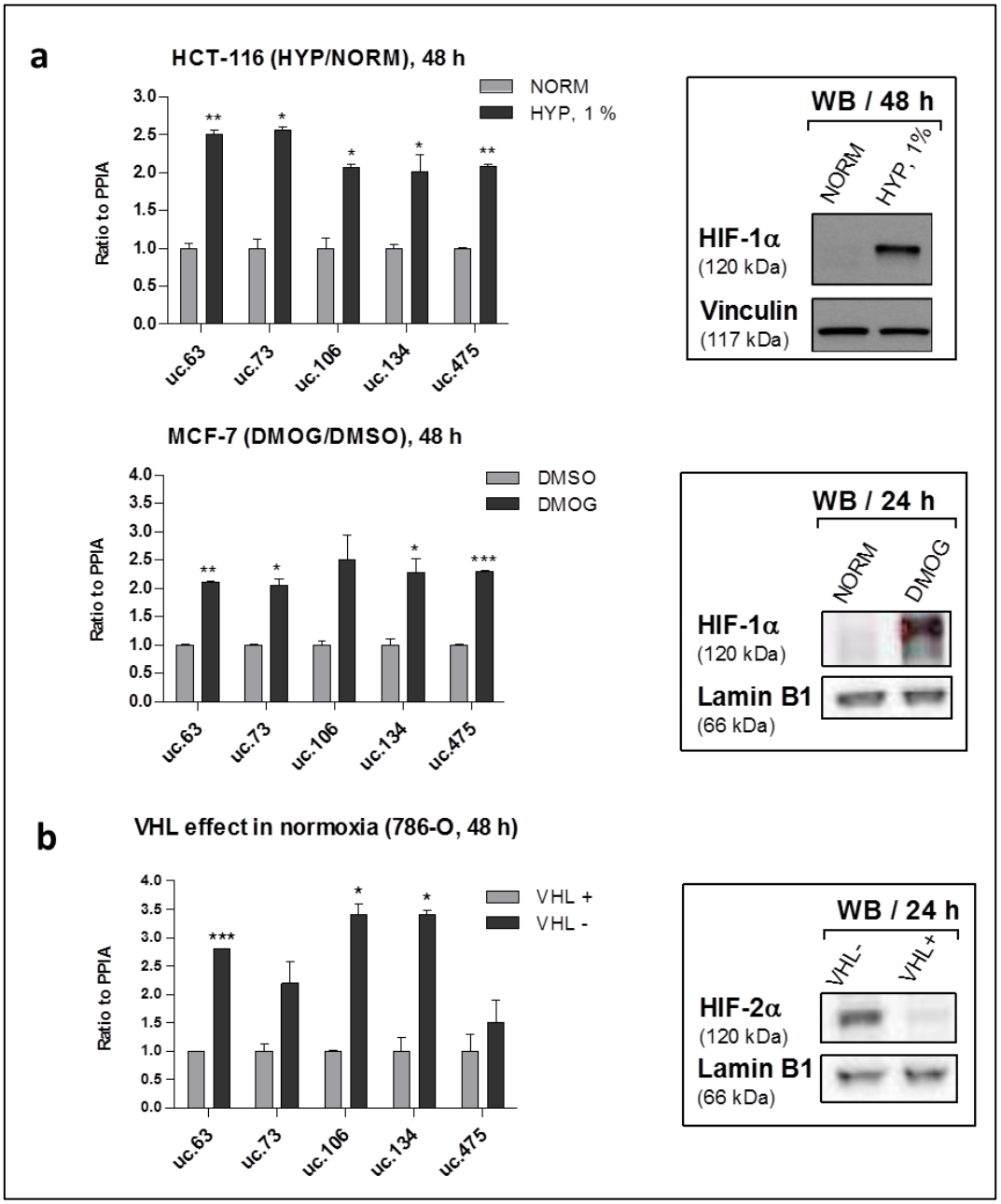



Figure legends: Error bars on the graphs indicate the standard deviation from two sample duplicates. * *P*-value < 0.05, ** *P*-value < 0.01, *** *P*-value < 0.001.

Figure 17: T-UCR expression induction by hypoxia

Slika 17: S hipoksijo inducirano izražanje T-UCR

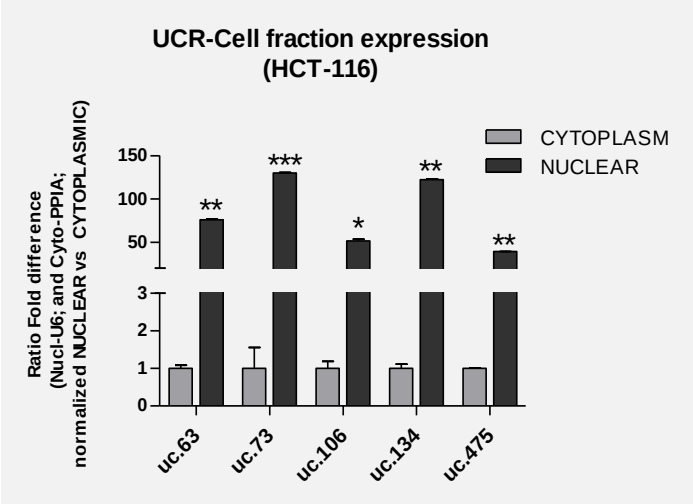



Figure legend: mean + SD (standard deviation). * *P*-value < 0.05, ** *P*-value < 0.01, *** *P*-value < 0.001. T-UCR expression in nuclear compared to cytoplasmic fractions was normalized with U6 snRNA (nuclear) or PPIA (cytoplasmic) for expression in nuclear or cytoplasmic fractions, respectively.

Figure 18: Comparison of T-UCR expression in the nucleus and in the cytoplasm

Slika 18: Primerjava izražanja T-UCR v jedru in v citoplazmi

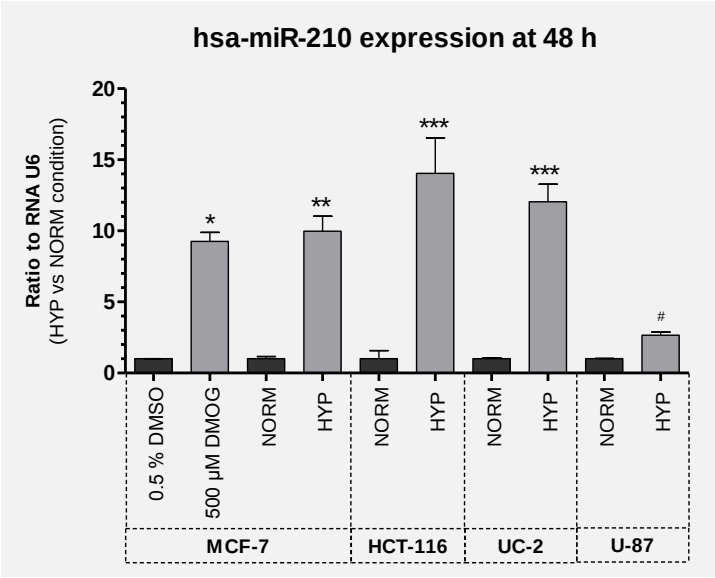



Figure legend: miR, microRNA; U6 snRNA as a reference gene.*** *P*-value < 0.001; ** *P*-value < 0.01;* *P*-value < 0.05; # *P*-value = 0.05. Expression was normalized to baseline conditions (normoxia or DMSO)

Figure 19: Expression of hsa-miR-210 in cancer cell lines after 48 h of exposure to hypoxia or hypoxia mimics

Slika 19: Izražanje hsa-miR-210 v različnih celičnih linijah raka po 48-urni izpostavljenosti hipoksiji ali posnemovalcu hipoksije, DMOG

### 4.2.2 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 RCC cell line 786-O, which exhibits homozygous inactivation of the tumor suppressor gene *VHL*, the well-established E3 ubiquitin ligase for all HIF family members. Cells 786-O exhibit constitutive activation of the HIF pathway (Xu et al., 2012) including normoxic overexpression of classic hypoxia-inducible genes, such as *CA9* or hsa-miR-210. We hypothesized that HIF2A-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* HIF2A targets, all HINCUTs were upregulated at least 1.5-fold in the VHL-deficient cell line (VHL −) compared with VHL-restored (VHL +) cells (Figure 17B). Furthermore, time course experiments in hypoxic models revealed that uc.475 was the most consistently induced at two different time points (Figure 20).

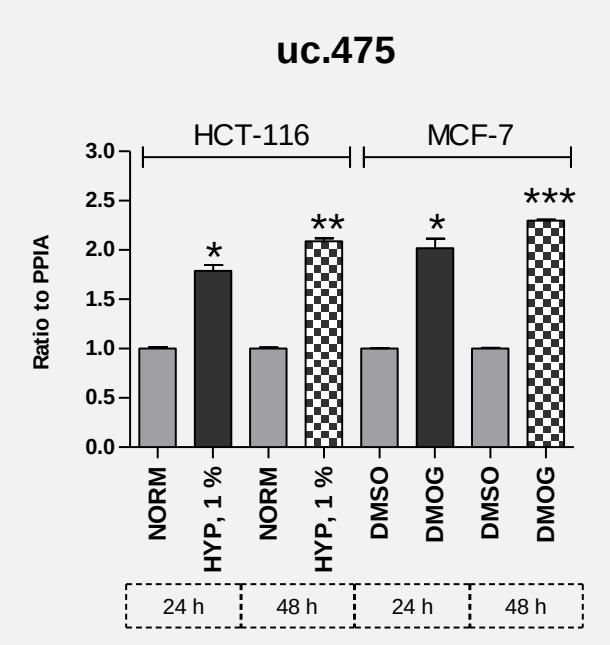



Figure legend: * *P*-value < 0.05, ** *P*-value < 0.01, *** *P*-value < 0.001.

Figure 20: Expression of uc.475 after exposure to hypoxia or hypoxia mimics, DMOG, compare to exposure to normoxia or DMSO

Slika 20: Izražanje uc.475 po izpostavitvi hipoksiji ali posnemovalcu hipoksije, DMOG, v primerjavi z izpostavitvijo normoksiji ali DMSO

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

(ANNEX D). However, only T-UCRs with yet unknown functional role and more than one binding sites identified were further included in these studies. Thus, in the uc.63 promoter region, we identified three candidate HIF-binding sites: Q3_uc.63A 5'-GGCAACGTGCATG-3' (63A.1), Q3_uc.63A 5'-GAGTGCGTGCTCCT-3' (63A.2), and Q6_uc.63A 5'-CGCGTGTGC-3' (63A.3); whereas for uc.475 the following three candidate sites were found: Q3_uc.475 5'-GCTCGCATGCGCGG-3' (475.1) and two in close proximity to each other, Q5_uc.475 5'-CACACGTGCGCG-3' and Q6_uc.475 5'-CACGTGCGC-3' (475.2) (Figure 21).

To further substantiate the role of HIF in the regulation of uc.475 and uc.63, we performed ChIP of HIF1A in MCF-7 cells under normoxic conditions or following treatment with the hypoxia mimetic, DMOG. We amplified genomic regions with HIF-interacting sites (HRE) presented on the genome graphic, using specific primers for pointed locations. We confirmed direct recruitment of endogenous HIF1A at the promoters of uc.63 (data not shown) and uc.475, which was more evident during hypoxic conditions (Figure 21; genomic regions from UCSC Genome Browser; 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. The ChIP figures are inverted for a better view of expression induction.

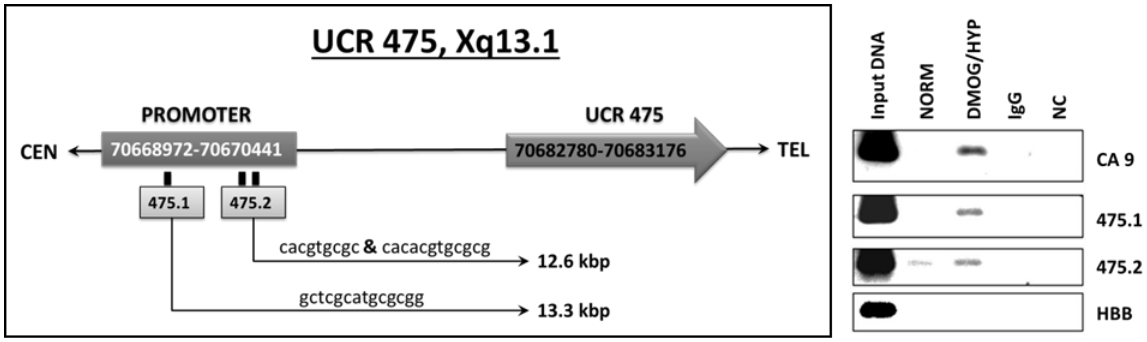


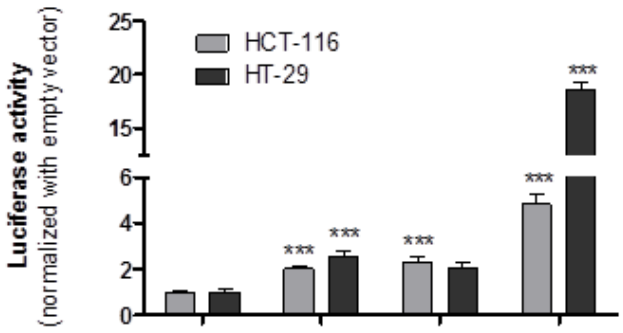

Figure legend: NC, negative control; *HBB*, hemoglobin beta, which is supposed to be transcriptional inactive; and carbonic anhydrase IX (*CA9*), which was used as a positive control in the ChIP experiment.

Figure 21: Confirmation of a direct association between HIF and T-UCR using ChIP assay

Slika 21: Potrditev neposredne povezave med proteinom HIF in T-UCR z uporabo metode ChIP

We further confirmed the transcriptional activation of uc.475 by hypoxia and HIF1A using promoter fragments subcloned into the pGL4-10 (luc2) basic expression vector. These constructs were co-expressed with pRL-TK (Renilla) control vector and constitutively active HIF1A constructs (pcDNA 3.1-HIF1A P/A (proline/alanine mutant) and pcDNA 3.1-HIF1A 3M (proline/alanine/asparagine mutant) (Ivan et al., 2001)) in two CRC cell lines, HCT-116 and HT-29 cells (Figure 22). We found that

hypoxia and the active HIF1As induced an increase (up to five-fold) in luciferase activity (Figure 22), which was correlated with the uc.475 expression in these cells (Figure 24**)**. In contrast, luciferase activity of the mutant clones that lack the HIF1A-binding sites (PT1 and PT2) was significantly reduced (Figure 23) (*P*-value < 0.005), suggesting that the two predicted binding sites are functional HIF-recognition sites that mediate hypoxia effects on uc.475 transcription.




| $O_2$ level | NORM | | | HYP |
|---|---|---|---|---|
| pcDNA 3.1-empty | (+) | (-) | (-) | (+) |
| pcDNA 3.1-HIF-1α P/A | (-) | (+) | (-) | (-) |
| pcDNA 3.1-HIF-1α 3M | (-) | (-) | (+) | (-) |
| pGL4-uc.475 (WT) | (+) | (+) | (+) | (+) |
| pGL4-empty | normalizer | | | |

Figure legend: The error bars represent the standard deviation from two independent experiments; each performed in triplicate and measured in duplicate with a luciferase assay. WT, wild type; HYP, hypoxia; *** *P*-value < 0.001 compared with luciferase activity of an empty vector.

Figure 22: The direct effect of HIF on the uc.475 promoter reporter construct

Slika 22: Neposreden vpliv HIF na reporterski konstrukt s promotorjem uc.475

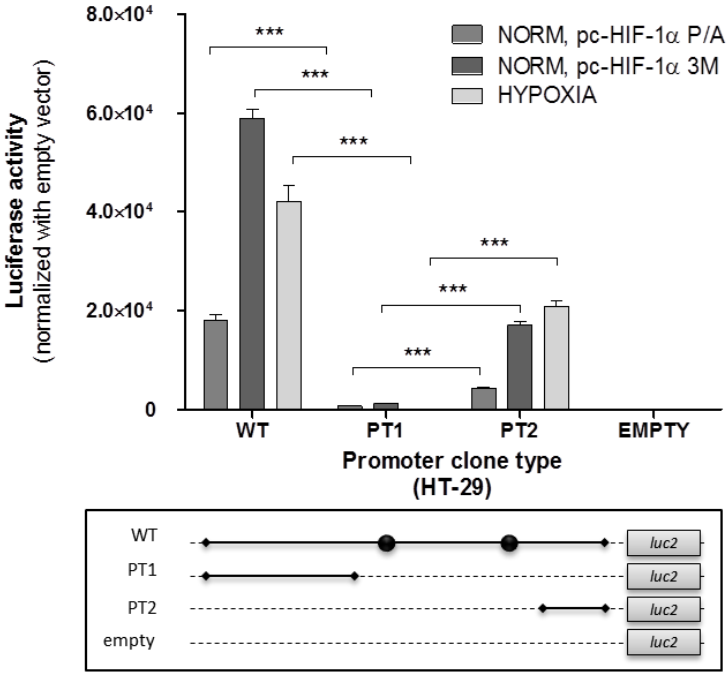



Figure legend: WT, wild type; PT1, truncated promoter 1; PT2, truncated promoter 2. Filled circles show the position of the HIF-A-predicted binding site. *** *P*-value < 0.001 compared with luciferase activity of an empty vector.

Figure 23: Luciferase activity in HCT-116 and HT-29 cells measured after transfection with truncated promoter reporter constructs (PT1 and PT2). These constructs had a complete deletion of the HIF1A-binding sites that were present in the originally designed uc.475 promoter reporter construct (WT)

Slika 23: Luciferazna aktivnost v HCT-116 in HT-29 celicah izmerjena po transfekciji s spremenjenima promoter-reporterskima konstruktoma (PT1 in PT 2). Konstrukti so imeli popolno delecijo vezavnih mest za HIF1A, ki so bili prisotni v prvotno izdelanemu uc.475 promotor-reporteskemu konstruktu (WT)

### 4.2.3 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 (Calin et al., 2007). Small interfering RNA (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 (Figure 24 and Figure 25). The uc.63 was identified to be expressed in genomic antisense orientation (the antisense transcript being expressed at 26.46 to 27.97 Cq, while the sense transcript was practically absent, Cq values of 33.08 to 36.15) in both 786-O cells and MCF-7 cells. The uc.475 was 16-times more expressed in the genomic sense orientation (Cq of 24.64 to 27.09 cycles) than antisense orientation (Cq of 30.92 to 33.08) in 786-O and MCF-7 cells. Therefore, the siRNAs we designed targeted the predominant form of each T-UCR.

The cell lines with the highest expressions of selected T-UCR were used for gene silencing experiments with siRNA; COLO-205 for uc.63 and HT-29 for uc.475. No major biological effect on the dynamic of COLO-205 culture was identified for uc.63 (Figure 26B), where cell line proliferation transfected with siRNA-uc.63A (2) was compared with siRNA-NEG as negative control or untransfected cell line. Henceforth, we focused our efforts on uc.475, which primary transcript has the same orientation as its host gene, O-linked N-acetylglucosamine (GlcNAc) transferase (*OGT*) (Figure 25).

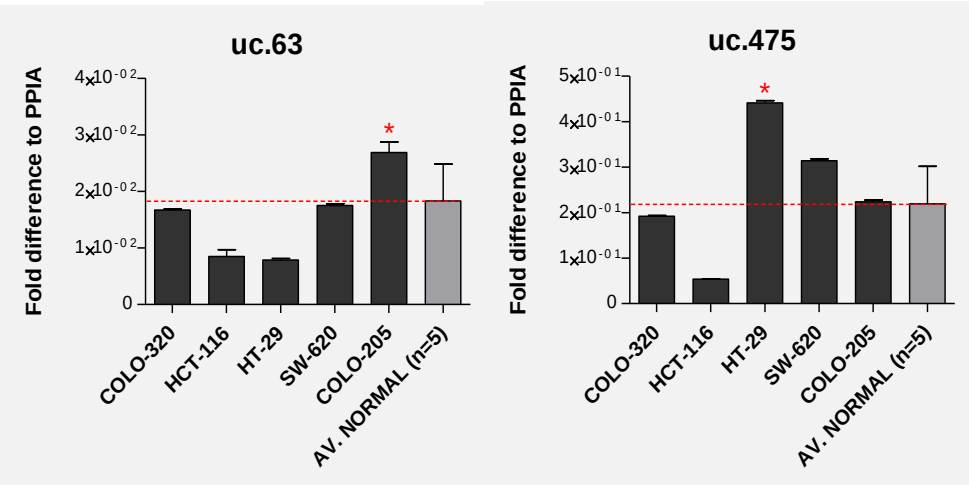



Figure legend: * *P*-value < 0.05.

Figure 24: Uc.475 and uc.63 expression in colon cancer cell lines

Slika 24: Izražanje uc.475 in uc.63 v celičnih linijah raka debelega črevesja

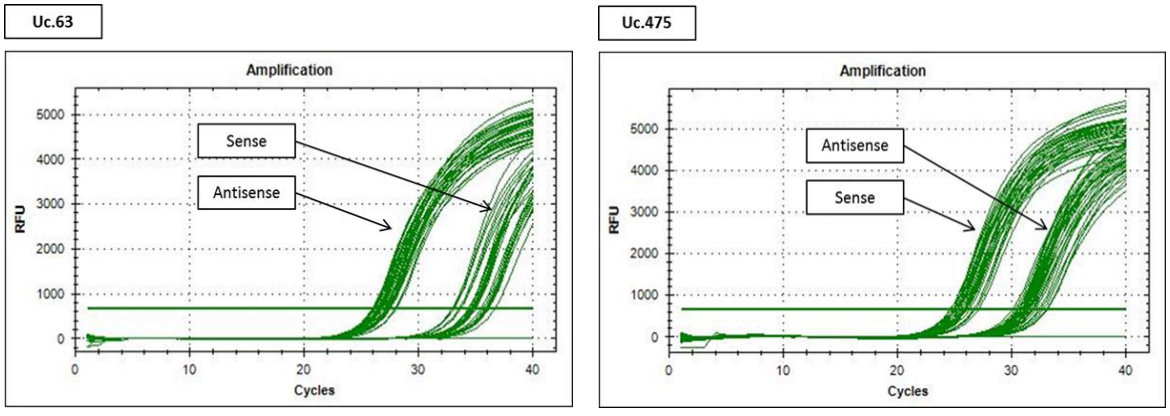



Figure legend: RFU, relative fluorescence units.

Figure 25: Strand-specific RT-qPCR expression profile of uc.63 and uc.475 transcripts

Slika 25: Verižno specifičen RT-qPCR profil izražanja prepisov uc.63 in uc.475

The effect of uc.475 downregulation on cell survival was tested in HT-29 cells under hypoxic and normoxic conditions. Again these cells were chosen based on their high endogenous uc.475 expression level, where the average of five normal colonic mucosa samples was used as a reference of normal T-UCR expressions (Figure 24). We employed a custom-designed pool of siRNAs to knockdown uc.475; each of the three siRNAs tested to be efficient to reduce a basal expression of uc.475 by more than 65% as confirmed by RT-qPCR (Figure 26A, Figure 27 and ANNEX C). Gene expressions were normalized against siRNA-NEG samples. T-UCR (uc.475) specific location to human genome (UCSC; GRCh/hg19) is presented in Figure 27 and is including visualization of siRNAs targeting host gene and T-UCR, respectively.

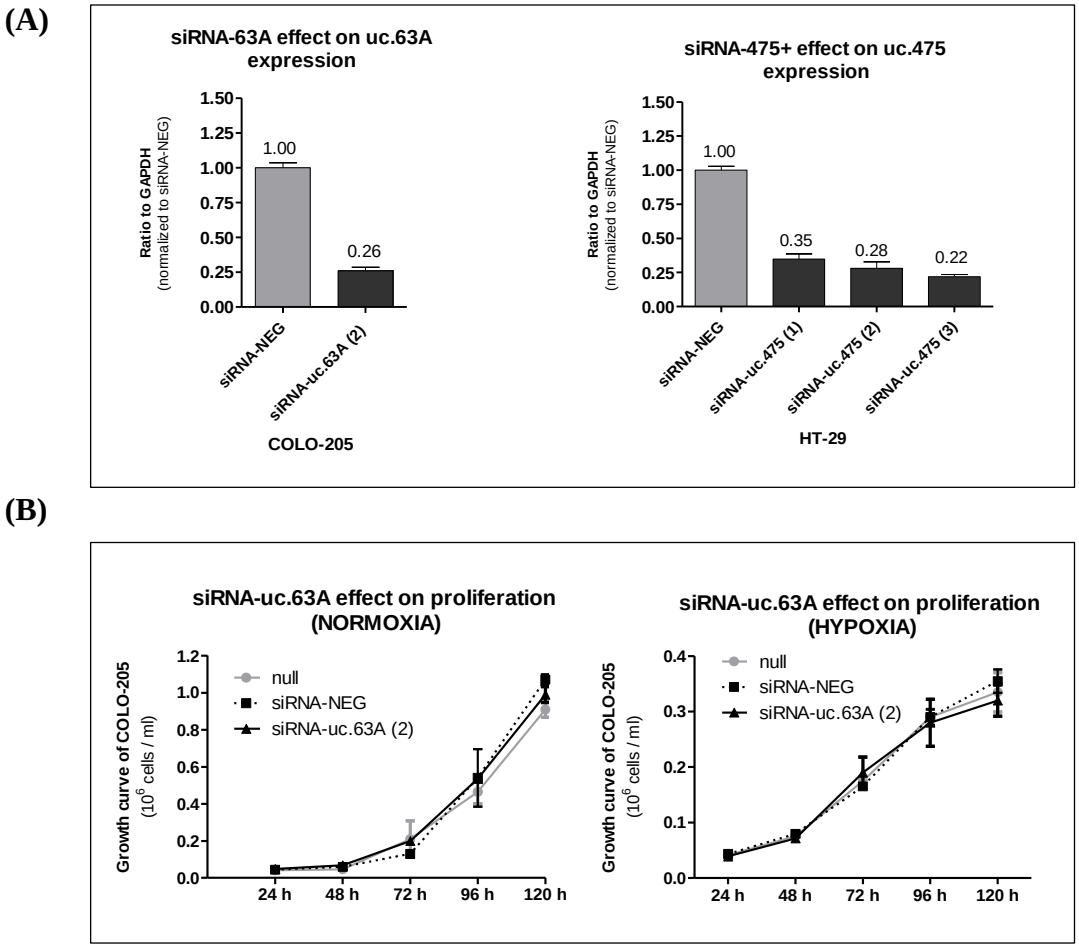
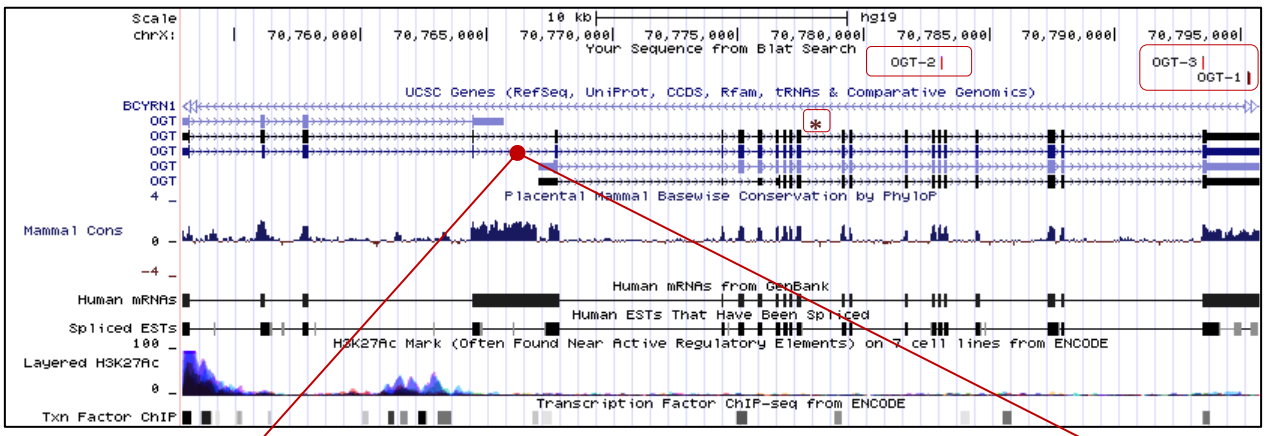



Figure legend: GAPDH, reference gene; COLO-205 and HT-29, colorectal cancer cell lines.

Figure 26: SiRNA-uc.475 and siRNA-uc.63A testing for silencing effect (Figure 26A) and cell proliferation effects of siRNA-uc.63A (Figure 26B)

Figure 26: Testiranje siRNA-uc.475 in siRNA-uc.63A na učinkovitost utišanja (Slika 26A) in vpliv siRNA-63A na proliferacijo celic (Slika 26B)

**siRNA-OGT**




**siRNA-uc.475**

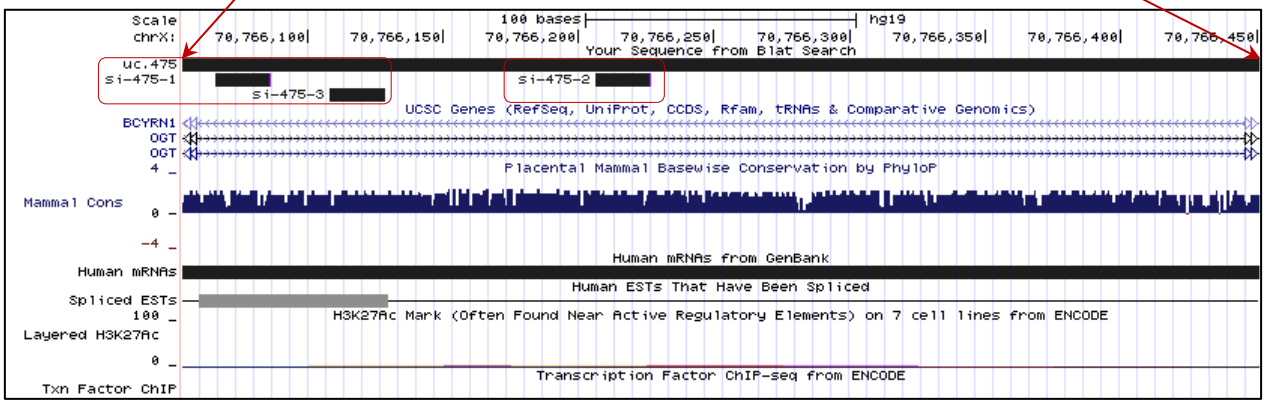



Figure legend: *- control siRNA for *OGT* silencing experiments was designed on intron 12.

Figure 27: SiRNA-uc.475 and siRNA-OGT overlaying human *OGT*

Slika 27: Prileganje siRNA-uc.475 in siRNA-OGT regijam gena *OGT* pri človeku

Under normoxic conditions, downregulation of uc.475 by using a pool of three siRNAs against uc.475 (siRNA-uc.475) had no significant impact on the total cell number *versus* control (siRNA-NEG) cells for up to 120 h (Figure 28A). In contrast, under hypoxic conditions (1% $O_2$), uc.475 knockdown significantly decreased cell proliferation (Figure 28B). Three independent experiments were performed and as control scrambled siRNA sequence, siRNA-NEG, was used. This effect was not the result of increased apoptosis, as poly ADP-ribose polymerase (PARP), a critical component of programmed cell death, remained uncleaved (data not shown). Instead, through the analysis of the cell cycle, by fluorescence-activated cell sorting (FACS), we observed that uc.475 knockdown under hypoxic condition enriched the G2/M fraction greater 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 (Figure 29**)**.

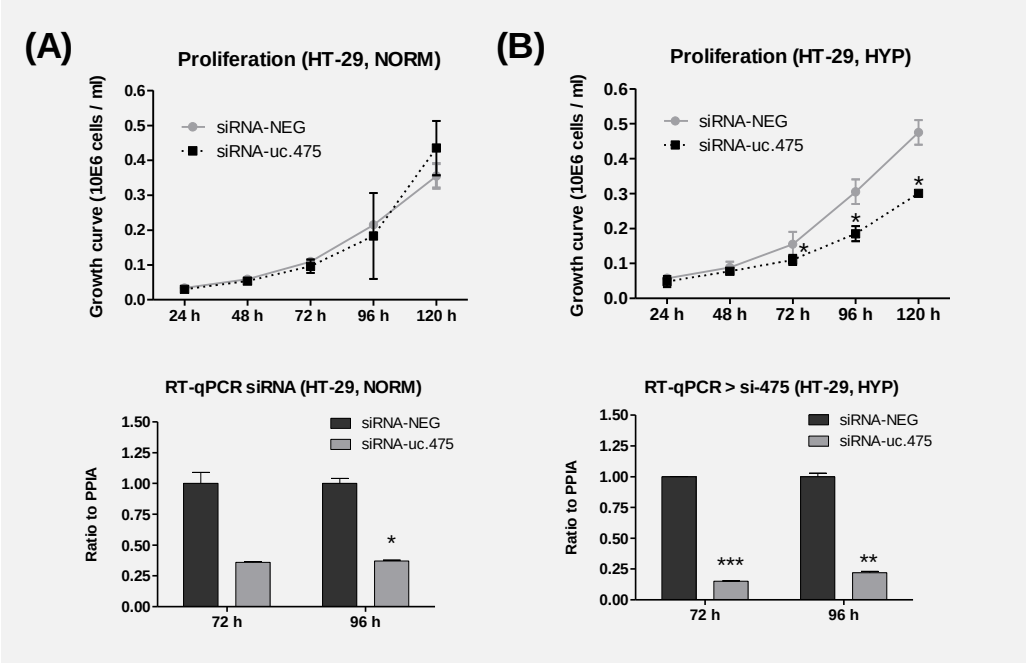



Figure legend: * *P*-value < 0.05, ** *P*-value < 0.01, *** *P*-value < 0.001.

Figure 28: Uc.475 biological significance under normoxia and hypoxia

Slike 28: Biološki pomen uc.475 v normoksiji in hipoksiji

On the contrary, no similar consistent variations in the percentage of cells in the various phases of the cell cycle under normoxic conditions were identified (Figure 29). Percentage on graphs presents average of two independent biological experiments. Effects of siRNA-uc.475 were compared to cell line transfected with siRNA-NEG as the negative control. The OGT protein expression in HT-29 cell line was measured after transfection with either siRNA-NEG or siRNA-uc.475 and 72 and 96 h incubation in normoxia and hypoxia. Protein specific bands were normalized to beta-actin or vinculin. The induction of the hypoxic marker, CA9, was evaluated as hypoxia specific marker (Figure 30).

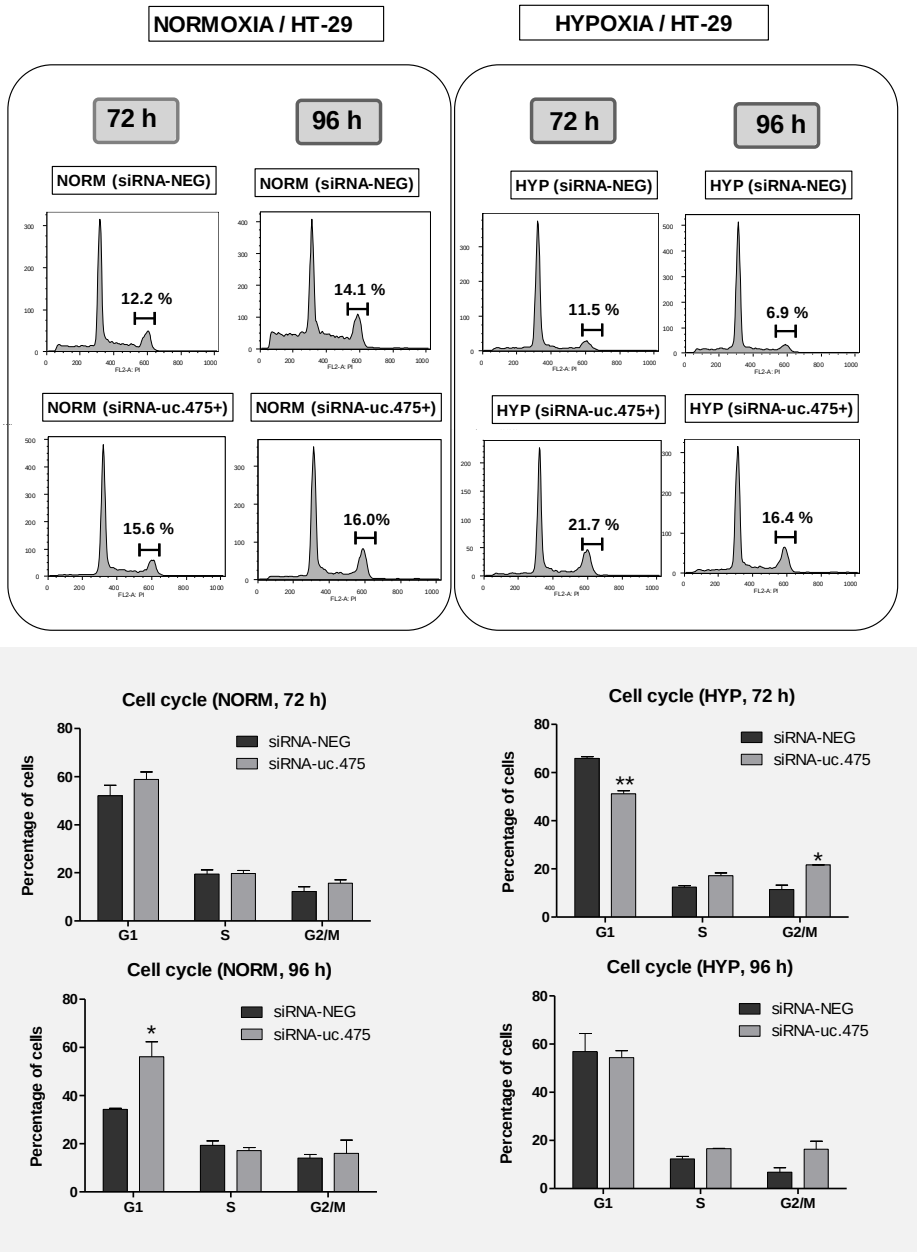



Figure legend: Error bars on graphs present the percentage of cells in different stages of the cell cycle. * *P*-value < 0.05, ** *P*-value < 0.01.

Figure 29: Cell cycle analysis by flow cytometry (FACS) of cells after uc.475 silencing in normoxia (left) and hypoxia (right)

Slika 29: Analiza celičnega cikla s pretočno citometrijo (FACS) po utišanju regije uc.475 v okolju normoksije (na sliki levo) in hipoksije (na sliki desno)

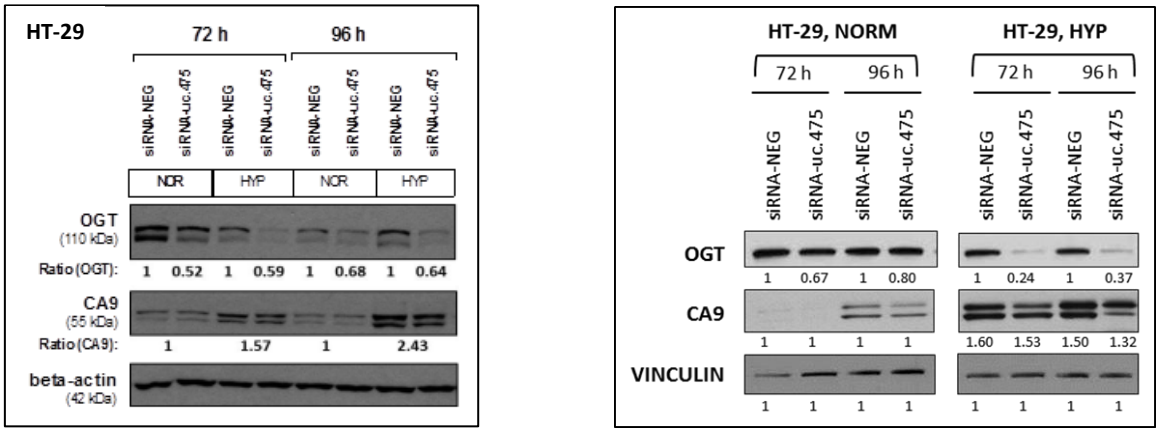



Figure 30: Expression analysis of OGT and CA9 proteins after silencing with siRNAs-uc.475

Slika 30: Analiza izražanja proteinov OGT in CA9 po utišanju s siRNAs-uc.475

Cell cycle distribution was also checked after treatment with siRNA against *OGT* (ANNEX C and Figure 27) under normoxia and hypoxia, and only minimal effects on cell cycle progression were observed (Figure 31). Small interfering RNA (siRNA) silencing *OGT* transcripts as a host gene of uc.475 were designed on the opposite site of uc.475 location within *OGT* (Figure 27).

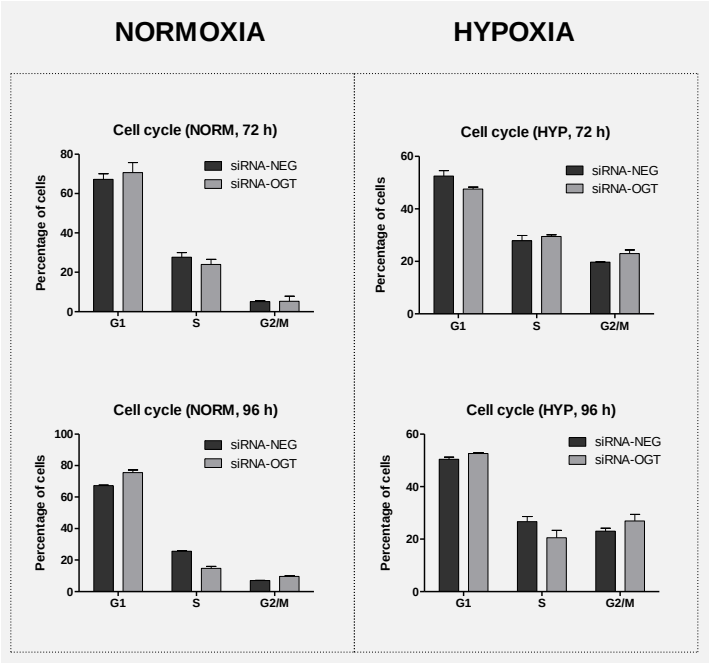



Figure legend: Error bars on graphs present the percentage of cells in different stages of the cell cycle.

Figure 31: Cell cycle analysis by flow cytometry (FACS) of cells after *OGT* silencing in normoxia (left) and hypoxia (right)

Figure 31: Analiza celičnega cikla s pretočno citometrijo (FACS) po utišanju gena *OGT* v normoksiji (na sliki levo) in hipoksiji (na sliki desno)

To exclude the possibility that siRNA against uc.475 downregulates pre-mRNA of *OGT* and leads to a phenotypic effects, 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 32), and no proliferation inhibition or altered cell cycle distribution was observed after knockdown of *OGT*-intron 12 (Figure 33).

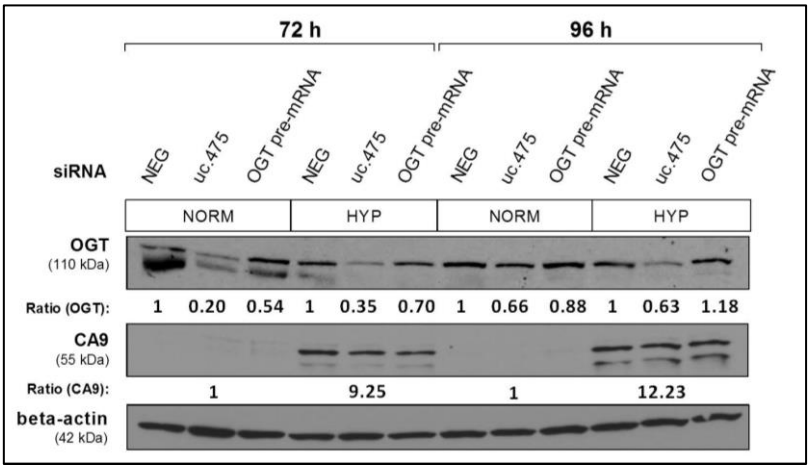



Figure 32: Differences in OGT protein expression after treatment with different siRNAs

Slika 32: Razlike v izražanju proteina OGT po tretiranju z različnimi siRNA

RT-qPCR confirmations of a successful *OGT* pre-mRNA silencing (> 50%) in hypoxia by a single custom designed siRNA and the effect of uc.475 silencing on *OGT* pre-mRNA expression is presented in Figure 33A. In Figure 33 (B and C) the effect of an *OGT*-intron 12 siRNA silencing in hypoxia and cell cycle analysis by flow cytometry (FACS) of cells after *OGT* pre-mRNA silencing in hypoxia is presented. 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.

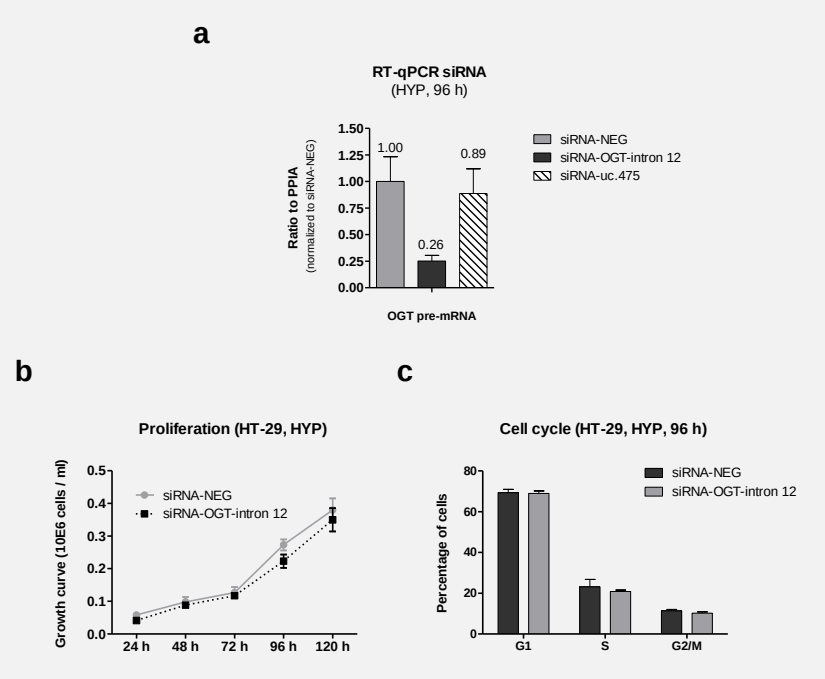



Figure 33: Biological effect of *OGT* silencing with siRNA-OGT-intron 12

Slika 33: Biološki vpliv utišanja *OGT* s siRNA-OGT-intron 12

### 4.2.4 OGT locus: a mix of OGT and HINCUT-1 transcripts

As mentioned previously, and in line with earlier reports (Kreppel et al., 1997; Lubas et al., 1997; Nolte and Muller, 2002) the uc.475 sequence (396 bp in length) is found inside the fourth intron of the *OGT* gene (UCSC Genome browser; GRCh/hg19, Figure 34). The *OGT* has two known isoforms (*OGT-1* of 5497 bp and *OGT-2* of 5467 bp) because of a splice site variant, which is known to encode proteins of 1046 or 1036 aa (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 (Nolte and Muller, 2002). Recently, this modification was proposed to be a new link between glucose metabolism and cancer growth and invasion in breast (Caldwell et al., 2010), colon and lung cancer types (Mi et al., 2011). In addition, *OGT* was found to be highly expressed in prostate cancer compared to the normal epithelium (Lynch et al., 2012). 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.

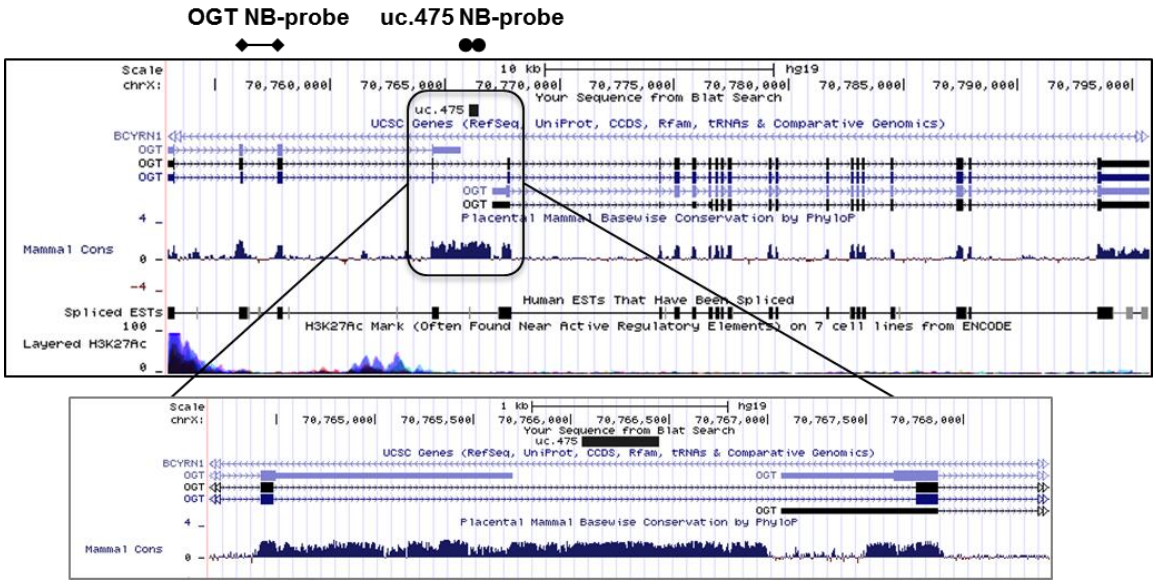



Figure legend: The rounded rectangle represents region containing uc.475. Northern blot probes locations are indicated. The intron 4 region where uc.475 is located is zoomed in for better view of region of conservation in mammals.

Figure 34: The genomic region of *OGT* from the UCSC Genome Browser with the known transcripts of *OGT* isoforms 1 and 2

Slika 34: Genomska regija gena *OGT* iz genomskega brskalnika UCSC z znanimi prepisi *OGT* izooblik 1 in 2

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 (Figure 35). In Figure 35A a scheme of cloning using the multiple primers amplification approach to define the length of T-UCRs is presented. Each column on a gel electrophoresis figure (Figure 35B and C) presents an approximate sum of primer "walking" distances subtracted from originally reported UCR length. Transcripts containing tested T-UCRs are indicated by arrows including with PCR products sizes. Complementary DNA on which we performed primer "walking"/cloning was transcribed from RNA of MCF-7 cell line exposed to 0.2% hypoxic conditions for 24 h. The same cDNA was used in RT-qPCR analysis.

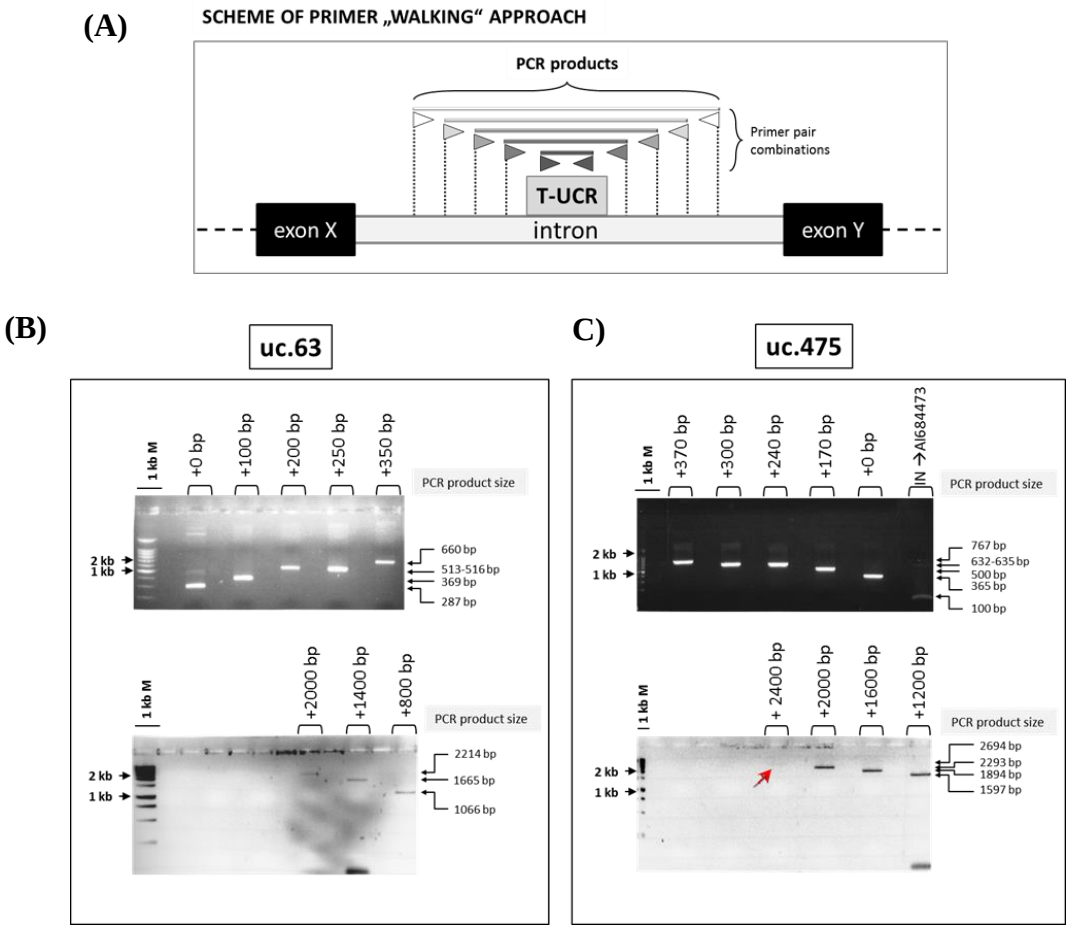



Figure legends: Primers for selected T-UCRs were designed on 3' and 5' side of each T-UCR at specific distances. "0 bp" represents the amplification of the originally reported UCR length: uc.63 = 278 bp (B) and uc.475 = 397 bp (C) (Bejerano et al., 2004b). The red arrow shows the presence of a weak band seen with UV-camera (Bio-Rad).

Figure 35: Complementary DNA (cDNA) "walking" to identify uc.63 and uc.475 transcript lengths

Slika 35: Pristop pomikanja začetnih oligonukleotidov po cDNA za določanje dolžine prepisov uc.63 in uc.475

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 36). However, RT-qPCR data are showing that the uc.475 noncoding transcript and the coding *OGT* transcript are expressed at a similar level, whereas the "hybrid" amplicons between the intron and the two adjacent exons are barely expressed. Therefore, 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).

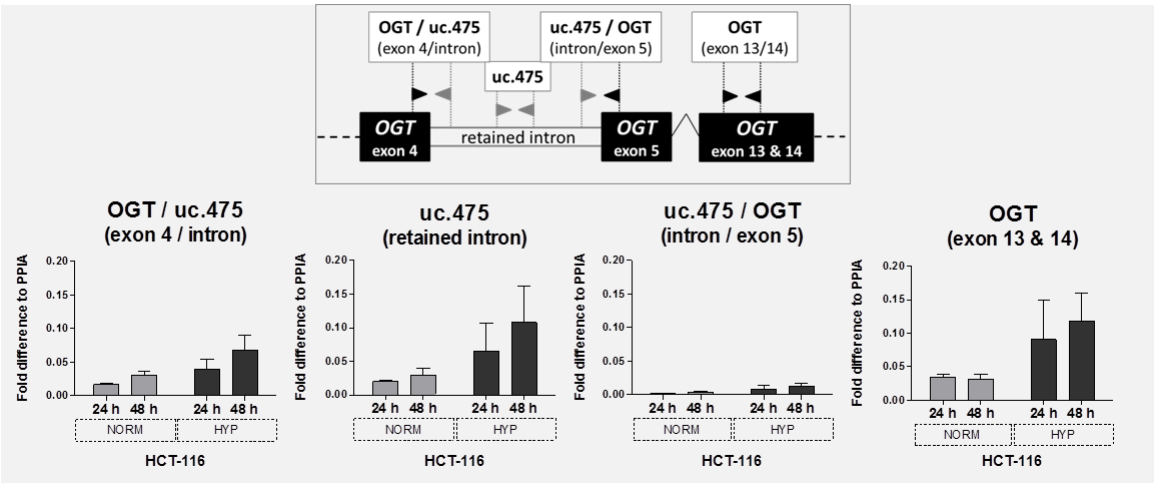



Figure 36: Transcripts produced at uc.475 genomic locus in an environment deprived of $O_2$ and/or glucose

Slika 36: Nastali prepisi genomskega lokusa uc.475 v okolju z znižano vsebnostjo $O_2$ in/ali glukoze

To further substantiate the tight link between *OGT* and uc.475, we performed experiments under stress conditions for which OGT function is relevant (Taylor et al., 2008). As OGT-mediated O-GlcNAcylation is proportional to substrate (glucose) availability (Robinson et al., 1995), we performed experiments under glucose- as well as oxygen-deprivation (OGD experiments) − two tightly associated characteristics of the tumor microenvironment. In HT-29 cells, the expression 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 in hypoxia in the presence or absence of glucose (Figure 37). These data show that the responses of uc.475 and *OGT* mRNA expression to $O_2$ and glucose deprivation are similar.

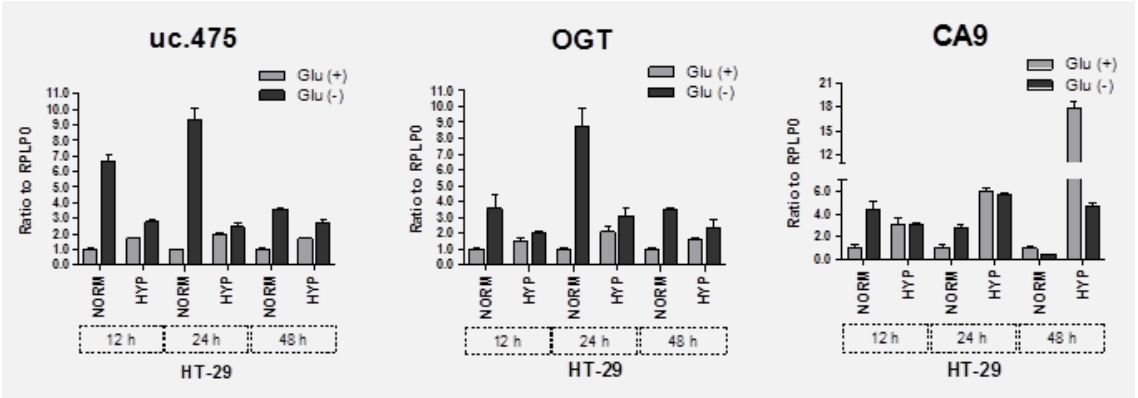



Figure 37: Glucose- and $O_2$-deprivation effect on mRNA expression level of uc.475 and *OGT*

Slika 37: Učinek pomanjkanja glukoze in $O_2$ na raven izražanja uc.475 in *OGT*

In parallel, we attempted to visualize the sizes of the transcripts from this locus by using the northern blot analysis. This would allow us to determine whether uc.475 exists as an independent transcript, as 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 38, 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 38, right). Transcript variants signals identified by northern blot assay in MCF-7 cells under normoxia and hypoxia are marked by arrows and visualized by a drawing (Figure 38, middle). The panel *PPIA* was used as a loading control. Therefore, we believe that uc.475 is the part of a larger *OGT* primary transcript, which retains intron four and which is expressed in a hypoxia-dependent manner.

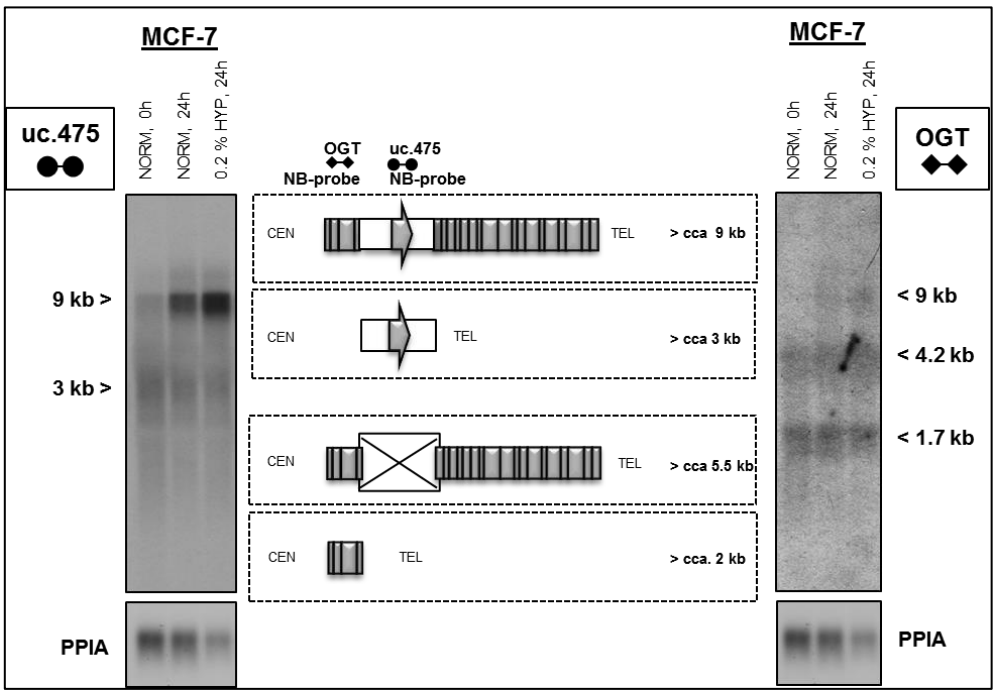



Figure 38: Uc.475-transcripts induction by hypoxia. Northern blot analysis of MCF-7 RNA identified transcripts of different sizes detected with uc.475 (left) or *OGT* probes (right)

Slika 38: Indukcija prepisov uc.475 v hipoksiji. Analiza prenosa northern RNA MCF-7 je pokazala prepise različne velikosti določene s sondo za uc.475 (na sliki levo) ali sondo za *OGT* (na sliki desno)

When we screened for potential open reading frames (ORFs) for *OGT* isoform 1 (NM_181672.2) and isoform 2 (NM_181673.3), both isoforms are showing one single ORF in the length of 1046 and 1036 aa, respectively (Figure 39A). When we screened for potential ORFs in the "retained intron", we detected no ORF > 100 bp (Figure 39B).

In addition, the 9-kb transcript was not predicted to encode for either of the two *OGT* isoforms, as the predicted ORF corresponds to 665 aa, potentially representing the smaller 78-kDa version of OGT reported recently (Butkinaree et al., 2010; Kreppel et al., 1997; Nolte and Muller, 2002) (Figure 39C). Moreover, we have searched for miRNAs and other putative ncRNAs in uc.475 region using 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.

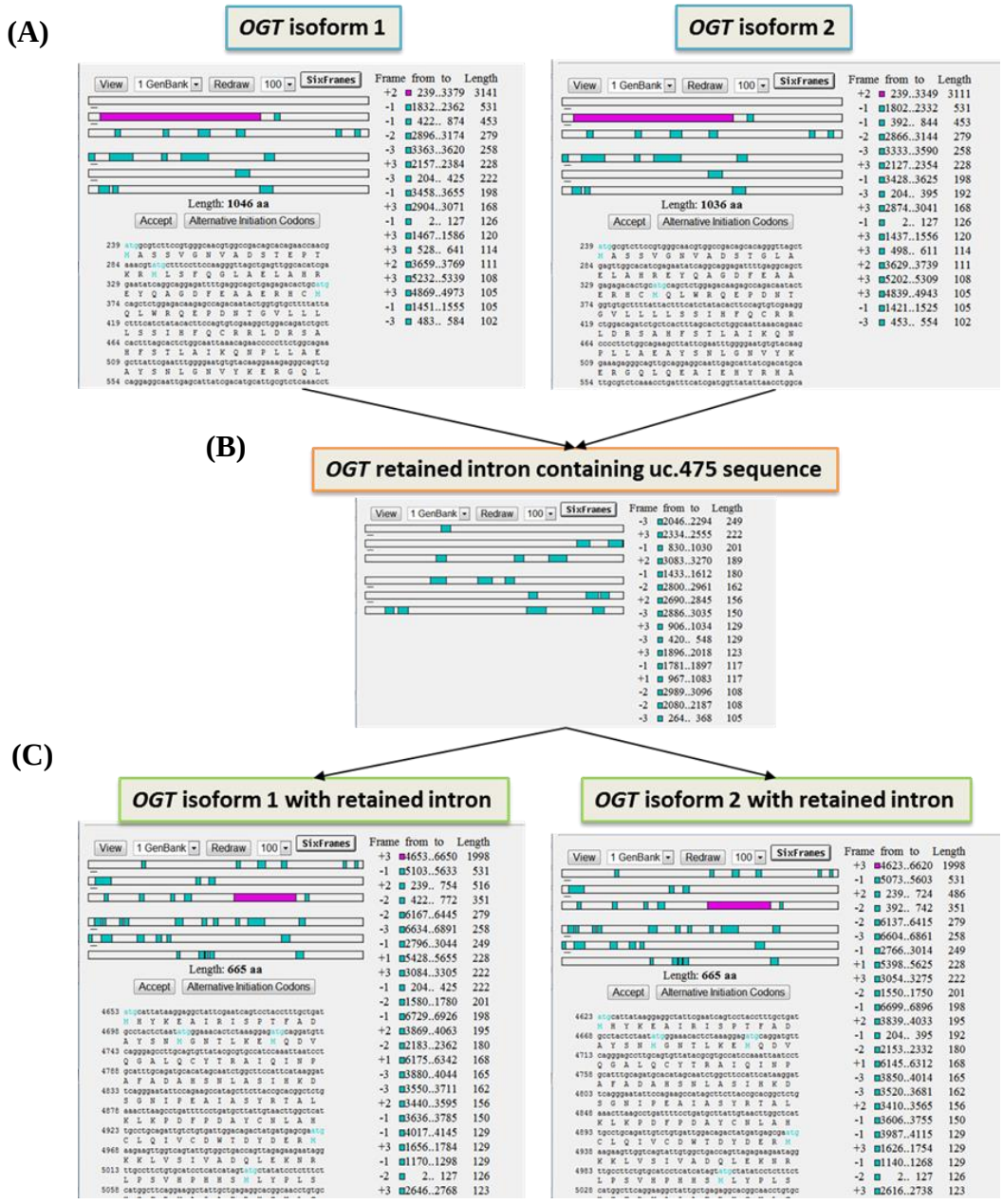



Figure 39: Search for open reading frames (ORFs) within different *OGT* transcripts

Slika 39: Iskanje odprtih bralnih okvirjev (ORF) v različnih prepisih *OGT*

### 4.2.5 Uc.475 inhibition has an impact on OGT expression and function

We first investigated whether OGT expression at RNA and protein level was affected by uc.475 knockdown in HT-29 cells. Expression of *OGT* mRNA levels was decreased by knocking down uc.475 in normoxia and hypoxia **(**Figure 40A), as well as of OGT protein levels, which was more evident in hypoxia (Figure 30). Silencing *OGT* with siRNAs (siRNA-OGT) resulted in decreased levels of uc.475 and *OGT* mRNA (Figure 40B), and OGT protein expression (Figure 40C).

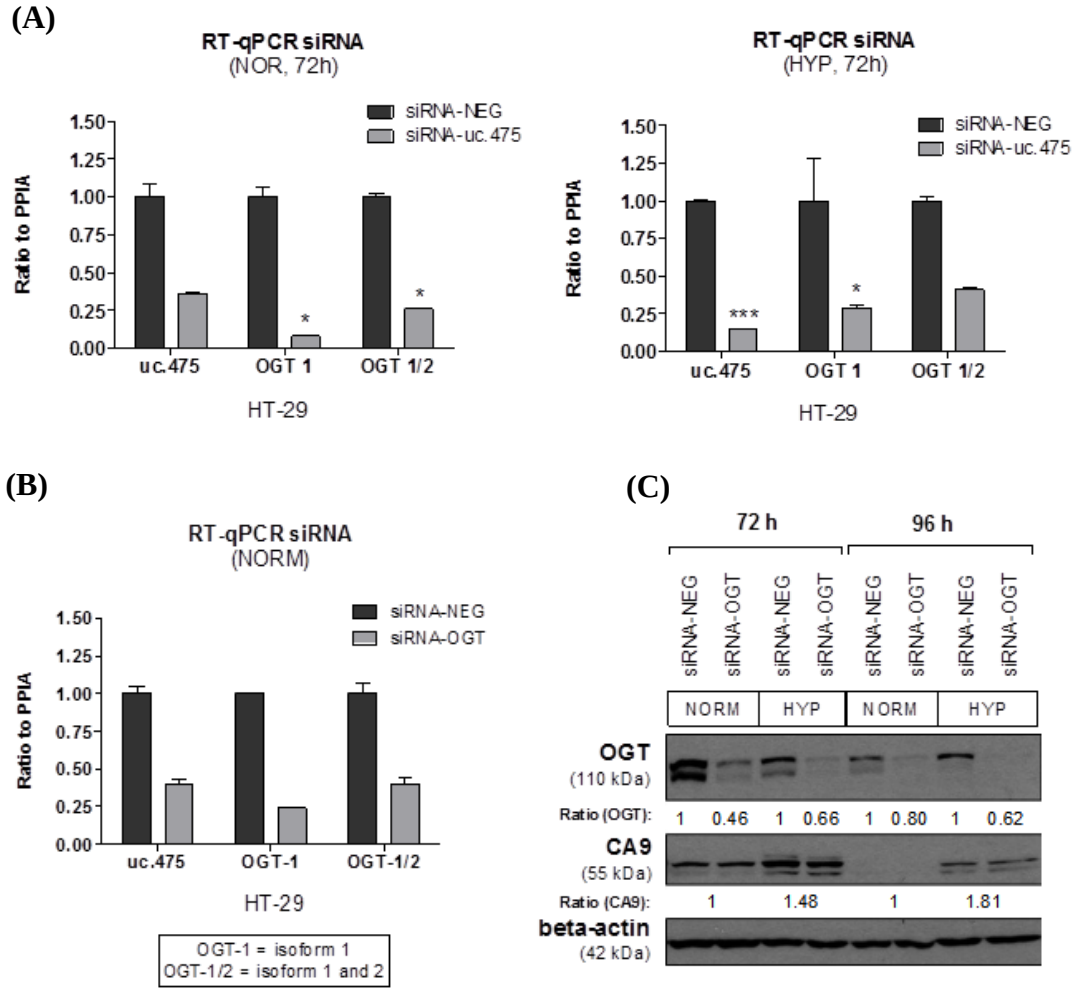



Figure legend: Primers were designed to detect both *OGT-1* and *OGT-2* isoforms. A bar indicates standard deviation. UCR or uc, ultraconserved gene; OGT, O-linked N-acetylglucosamine (GlcNAc) transferase; siRNA, small interfering RNA; HT-29, colon cancer cell line; NOR, normoxia; HYP, hypoxia; PPIA, reference gene cyclophilin A; * *P*-value < 0.05, *** *P*-value < 0.001. Hypoxic conditions were confirmed by CA9 induction. Protein specific bands were normalized to beta-actin as a reference protein.

Figure 40: Impacts of siRNA-uc.475 or siRNA-OGT on uc.475 or *OGT* mRNA level and OGT protein expression

Slika 40: Vpliv siRNA-uc.475 ali siRNA-OGT na raven izražanja uc.475 ali *OGT* mRNA in proteina OGT

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 after 48 h exposure to hypoxia (1%) and normoxia (Figure 41**)**. Moreover, OGT even decreases during prolonged (72 - 96 h) 1% $O_2$ treatment (data not shown).

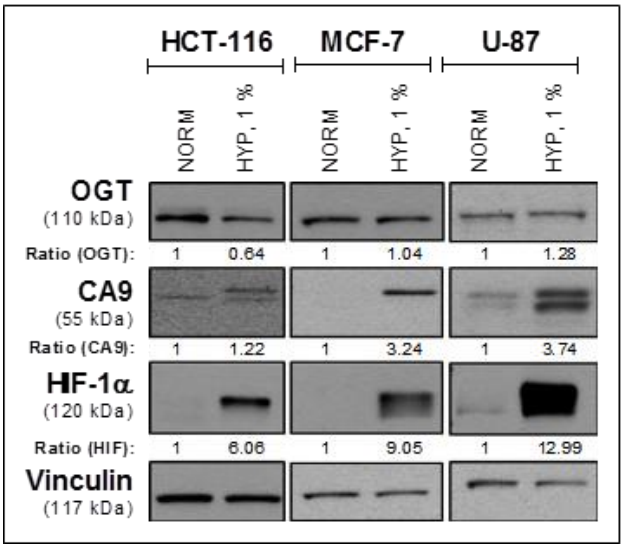



Figure legend: CA9 and HIF1A were used as confirmation of hypoxic condition. Protein loading and expression was compared to Vinculin.

Figure 41: OGT protein expression level is similar in normoxia and hypoxia

Slika 41: Raven izražanja proteina OGT v normoksiji in hipoksiji je podobna

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. This was performed in HT-29 cells treated with different concentration of PUGNAc, an inhibitor of the N-acetylhexosaminidases (OGA or MGEA5) added in cell media (40 and 80 μM) in order to prevent the removal of O-GlcNAc from proteins (Dong and Hart, 1994). The O-GlcNAc protein levels were detected in samples exposed to normoxic or hypoxic conditions for 72 h or simultaneously transfected with siRNA targeting, either uc.475, *OGT* pre-mRNA or *OGT*. Downregulation of uc.475 led to a significant decrease in the global O-GlcNAcylation rate, an effect similar to treatment with siRNA-OGT (Figure 42), 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 (Figure 43). All levels were compared with sample transfected with siRNA-NEG.

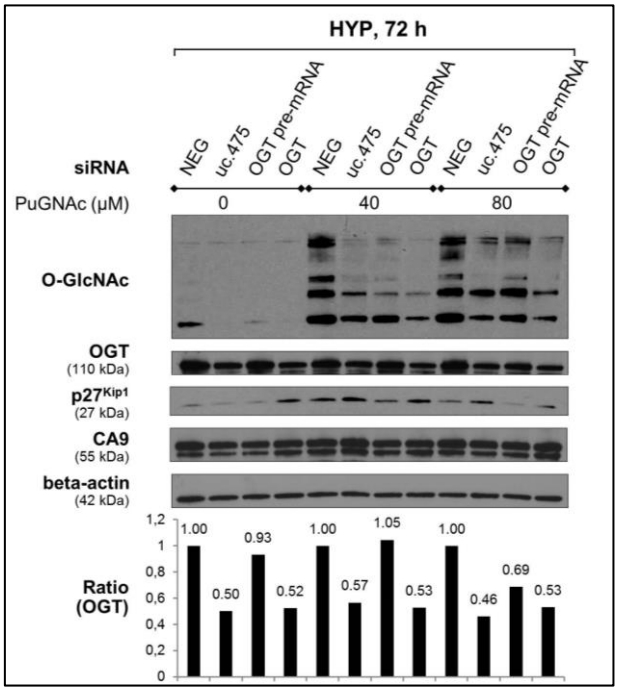



Figure legend: siRNA against *OGT* pre-mRNA was designed to target *OGT*-intron 12.

Figure 42: O-GlcNAcylation change after knocking down uc.475 or *OGT* gene

Slika 42: Sprememba O-GlcNAcilacije po utišanju uc.475 ali gena *OGT*

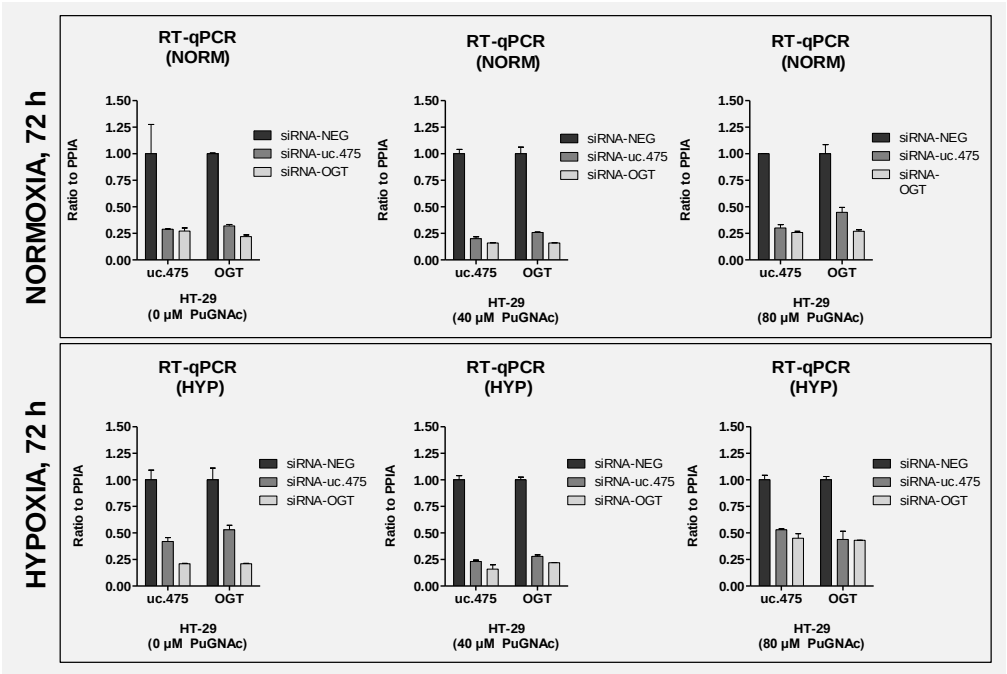



Figure 43: Expression analysis by RT-qPCR after treatment of HT-29 cancer cell line with an inhibitor of N-acetylhexosaminidases, PUGNAc

Slika 43: Analiza izražanja z RT-qPCR po tretiranju celične linije HT-29 z inhibitorjem N-acetilheksominidaz, PUGNAc

### 4.2.6 The enhancer-like function of the DNA region containing uc.475 ncRNA

It has been previously 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 (Visel et al., 2008). For doing this, we used sequence-transcription stimulation by an adjacent heterologous promoter region as described in (Orom et al., 2010). The 2.7-kb long region (−1 kb upstream and +1.3 kb downstream of uc.475) (Figure 35) was subcloned into the pGL3 basic vector, primarily transformed by the thymidine kinase (TK) promoter insertion (pGL3-TK empty) (Figure 44A). This cloned insert revealed a 2-3-fold luciferase activity enhancement under normoxia and hypoxia, in U-87 cells (Figure 44B). These findings demonstrate that the uc.475 DNA region has also an enhancer activity both under hypoxic and normoxic conditions.

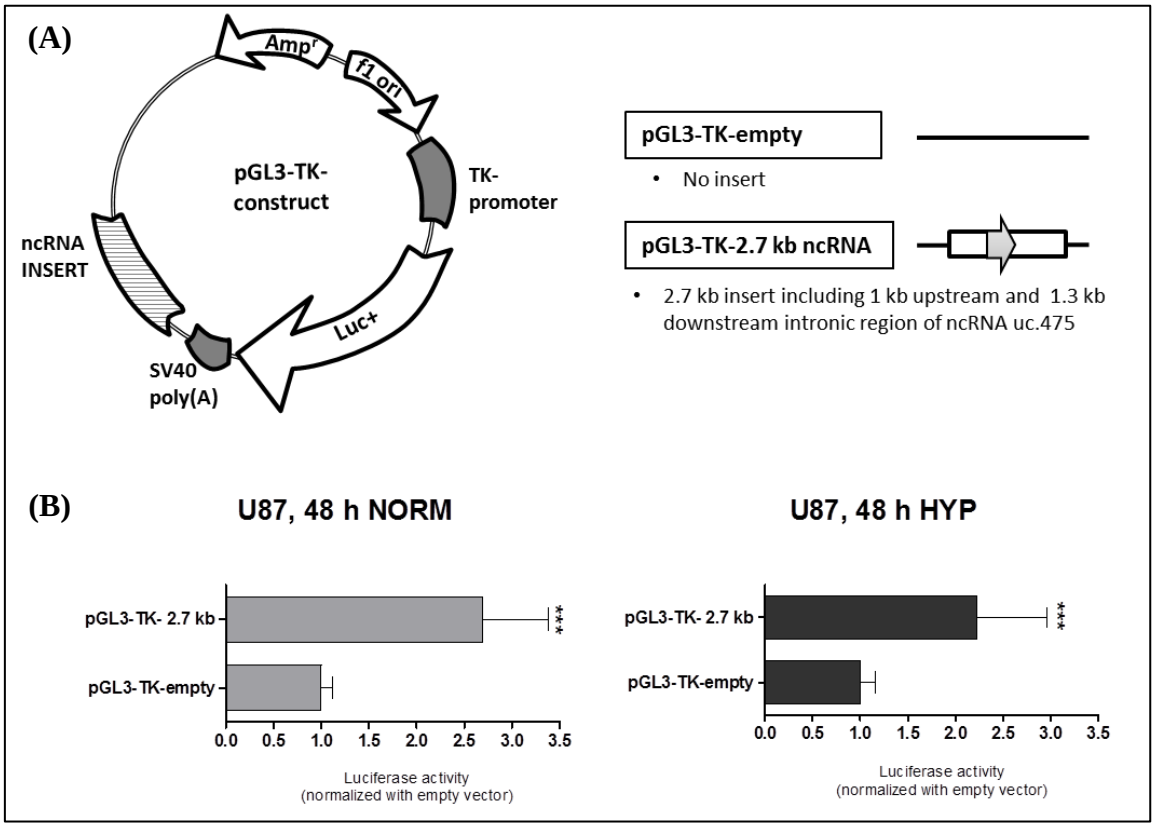



Figure legend: Axis *x* represents the relative Firefly to Renilla luciferase activity, whereas *y* axis presents the plasmids. All measurements generated were normalized to the pGL3-TK empty vector. All data shown are the mean of two biological repetitions (at least eight luciferase measurements per experiment). *** *P*-value < 0.001.

Figure 44: Testing cloned uc.475-containing region for enhancer activity

Slika 44: Testiranje klonirane regije, ki vsebuje uc.475, za aktivnost ojačevalca

# 5 DISCUSSION

## 5.1 ATLAS OF HIF1A GENE REGULATORY NETWORK

Complete identification of *HIF1A* gene regulatory network is essential for understanding its regulation and its downstream target effects. Therefore, we developed an Atlas of *HIF1A* gene regulatory network to expand the understanding of cellular and systemic response to hypoxia. Previously analyzed *HIF1A* SNPs were associated with various phenotypes in human. Further, *HIF1A* SNPs were predicted to be involved in gain/loss of TF and miRNA binding sites, which could be a possible explanation for functional effect of analyzed HIF1A polymorphisms on HIF1A expression. Additionally, various biological pathways were found to be enriched with HIF1A target genes and potentially essential for cells to adapt and survive in hypoxia. In this study we developed an Atlas of HIF1A gene regulatory network, and additionally presented a general scheme to design an Atlas of regulatory network for any TF of interest (Figure 45). Schematic presentation is showing bioinformatics tools and databases that can be used to obtain validated and predicted information about TF regulatory network.

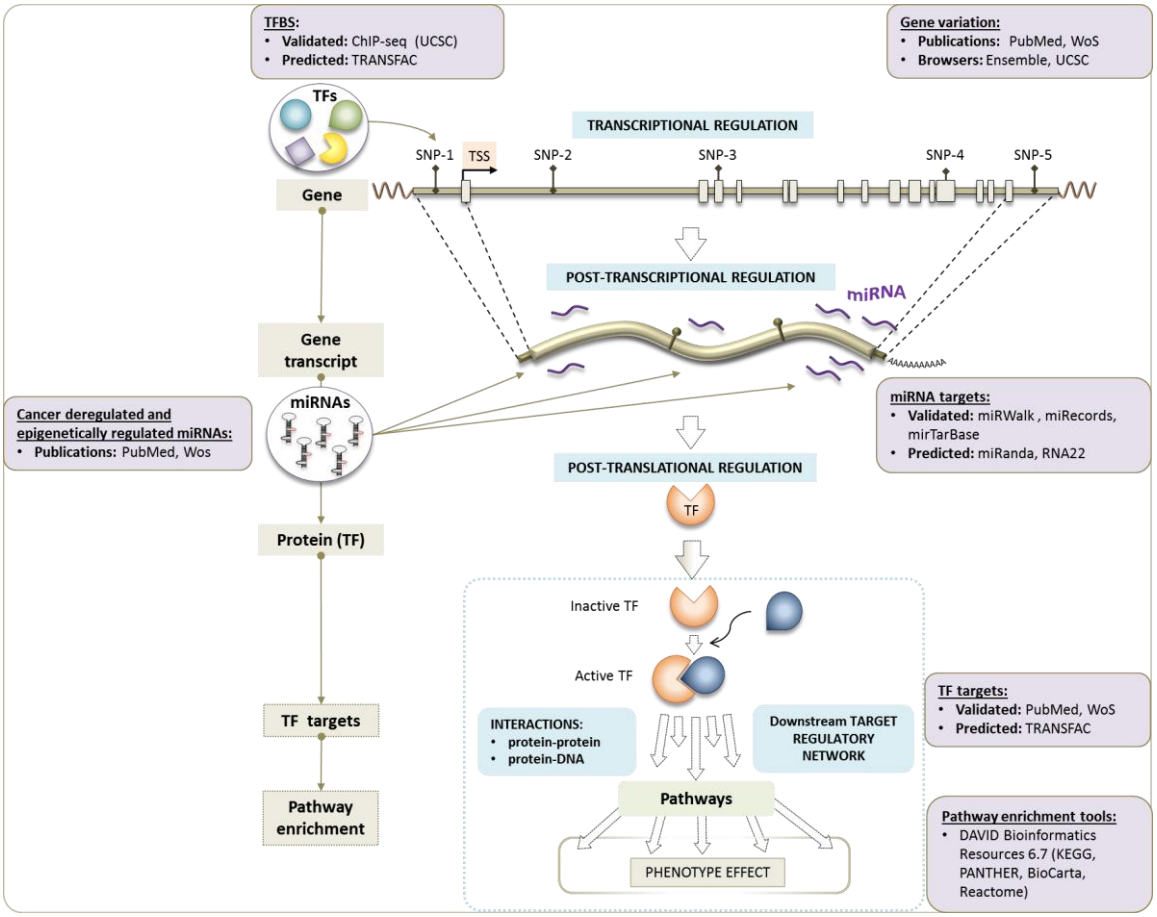



Figure legend: TF, transcription factor; TSS, transcription start site, SNP, single nucleotide polymorphism; miRNA, microRNA; WoS, Web of Science; KEGG, Kyoto Encyclopedia of Genes and Genomes; PANTHER, Protein ANalysis THrough Evolutionary Relationships; TRANSFAC, TRANScription FACtor database; ChIP-seq, chromatin immunoprecipitation (ChIP) with massively parallel DNA sequencing.

Figure 45: Atlas of TF regulatory network

Slika 45: Atlas regulatorne mreže za TF

### 5.1.1 HIF1A association with cancerous and non-cancerous phenotypes

*HIF1A* was previously associated with various phenotypes and our literature review revealed that *HIF1A* variations were often associated with cancer, cardiovascular diseases, and performance. The review revealed that SNPs tested in association studies were in half of the cases associated with cancer (16/32 phenotypes). It is known that solid tumors frequently have low levels of $O_2$. This can be a result of high cancer cell proliferation capability, which cause they outgrow their vascular network and consequentially limiting $O_2$ diffusion within the tumor. Hypoxic stress might also be caused by perfusion defect as a result of abnormal tumor blood vessel structure and function (Majmundar et al., 2010). Consequentially, these events cause HIF1A to increase in solid tumors (Bertout et al., 2008), but its levels can additionally be

increased by HIF-independent pathways (Majmundar et al., 2010). HIF1A has a capability to directly reprogram the metabolic state in cells, which is important in hypoxic settings such as vascular disease and cancer (Majmundar et al., 2010). Thus, HIF1 in hypoxic cells regulates the transcription of many genes involved in key aspects of cancer biology, including immortalization, maintenance of stem cell pools, cellular differentiation, genetic instability, vascularization, metabolic reprogramming, autocrine growth factor signaling, invasion/metastasis, and treatment failure (Semenza, 2007a). In cancer HIF1A also often associates with poor clinical prognosis in many cancer types (Semenza, 2007a).

An association of HIF1A with cardiovascular diseases is additional confirmation of HIF1A systemic effect. Cardiovascular diseases like atherosclerosis of tissues, such as heart, brain, and limb muscle, are known to be susceptible to ischemic injury (Beckman et al., 2002; Kett-White et al., 2002). Besides, myocardial ischemia is the most common cause of cardiac hypoxia in clinical medicine and it occurs when $O_2$ delivery cannot meet myocardial metabolic requirements in the heart (Shohet and Garcia, 2007). The HIF1A expression is essential and sufficient to promote reperfusion in ischemic skeletal muscle (Majmundar et al., 2010). Moreover, reduced HIF1A activation was also found in hypoxic skin wounds of aged diabetic mice (Liu et al., 2008b), which emphasize the role of age in ischemic response. This is important since peripheral arterial disease is associated with age (Beckman et al., 2002). Pro-angiogenic roles of HIF1A were also associated with hypertrophic cardiomyopathy, myocardial infarction, skin wound healing, and retinal neovascularization (reviewed in (Majmundar et al., 2010)).

Many studies suggested that the occurrence of local hypoxia in the muscle causes a drop in $O_2$ pressure within the myocyte during exercise. Consequentially this cause an induction of HIF1A-mediated signaling pathways in human skeletal and thereby provoking a shift towards increased use of oxidative pathways for energy production (Doring et al., 2010), which effects human performance. Besides, cardiovascular system subsequently influences human body performance, these genotype-phenotype associations are just additional confirmation of an extensive HIF1A effect.

The review of genotype-phenotype studies revealed that HIF1A association studies were most often focused on two nsSNPs located within ODD domain: rs11549465 (*p.Pro582Ser*), and rs11549467 (*p.Ala588Thr*). For these two SNPs also opposing associations with phenotype were found, most probably due to being an object of numerous studies. In addition, both rs11549465 and rs11549467 are germline SNPs within the ODD/pVHL interaction domain (Clifford et al., 2001). However, rs11549465 is also known to have an ability to enhance transactivation (Huang et al., 1998; Tanimoto et al., 2003), but was not identified as a site for *HIF1A* hydroxylation and is not mediating VHL binding (Yamada et al., 2005). Moreover, these results are in

agreement with previous study of Percy et al. (2003), who showed that the substitution *p.Pro582Ser in vitro* does not prevent VHL binding to a fragment of *HIF1A* after hydroxylation at *p.Pro564*.

As in most association studies, also *HIF1A* association studies focused on nsSNPs and much less attention was given to synonymous (sSNP) and other noncoding SNPs. This could reflect a long-term assumption that sSNPs are inconsequential, since the primary amino-acid sequence of the protein stays unchanged. Even though, *HIF1A* association studies mostly focused on nsSNPs within ODD domain and $O_2$-dependent proteolysis is the primary means of regulating HIF family of transcription factors, our anticipation was that genetic variability within different *HIF1A* domains, beside ODD region, have a potential to affect HIF1A activity. Here we present few possible scenarios that may disrupt the role of HIF1A functional domains and therefore influence on HIF1A stability and role as a main TF in hypoxia (Figure 9): i) variations within bHLH domain may prevent binding of HIF1A to HRE-recognition sites within promoter region of target genes, and likewise variations within HRE may influence on creation or destruction of HRE-binding sites for HIF1A to influence on downstream targets; ii) variations within PAS domain could theoretically affect dimerization with ARNT (HIF1B), which would result in HIF1A inability to function as transcriptional regulator; iii) variations within ODD domain could influence on stability of a protein in normoxia, since conserved proline residues (*p.Pro402* and *p.Pro564*) are usually targeted for VHL/proteasome degradation; iv) variations within NLS may influence on translocation of HIF1A into the cell nucleus by nuclear transport; and v) variations within N-TAD and C-TAD could influence on transcriptional activation of HIF1A and interaction with its co-activators. Results of our literature review showed that *HIF1A* variations outside of ODD domain also have functional effect and were associated with diseases and phenotypes. For the following HIF1A regions genotype-phenotype associations were found: TF binding (four SNPs within 5-flanking region; rs7148720, rs2783778, rs1535679, rs28708675), ARNT binding (*p.Val116Glu* within PAS domain), ODD region (four SNPs: rs41508050, rs199775054, rs11549467, and rs11549465, which is located in N-TAD within ODD domain), 3'-UTR (rs2057482) or downstream (rs1319462), and four within introns (rs12434438, rs1957757, rs10873142, and rs10645014).

Moreover, as mentioned *HIF1A* variations could potentially affect binding with other co-activators (EP300/CBP) or degradation complexes (nuclear transporter, VHL/proteasome, PHD1-3). However, in association studies seven SNPs within ODD domain, recognized by VHL/proteasome, were tested. This region is also overlapping with N-TAD domain; important for co-factor binding, where 3/7 SNPs were tested for phenotype association and only for rs15549465 association with phenotypes were confirmed. Although many SNPs (13/30) from *HIF1A* association studies were intronic,

none of them was previously classified as a splice region variant. According to Ensembl five *HIF1A* variations are classified as splice-region variants (either located within 1-3 bases of the exon or 3-8 bases of the intron): TMP_ESP_14_62188452, rs200037213, COSM193393, rs200439820, TMP_ESP_14_6221364, however none of these variations was previously tested in association studies. These SNPs are of interest for future experiments since they could affect alternative splicing. Additionally, other *HIF1A* gene regions like upstream and downstream are important for binding other TFs and functional noncoding RNA molecules like miRNAs.

### 5.1.2 Analysis of TF and miRNAs binding sites

Existence of *HIF1A* gene polymorphisms within promoter, 5'- and 3'-UTR regions were recognized to alter gene transcription or stability of a transcript, and potentially influence on expression of downstream targets (Hong et al., 2007). In addition to previously published experimental data related to TF and miRNA interactions with *HIF1A*, our predictions suggested that *HIF1A* SNPs either create new or delete existing TF and miRNA binding sites, through which the stability or transcription of mRNA could be affected. Integration of an existing knowledge related to TFBS revealed that SP1 (Minet et al., 1999), NFKB1/RELA (Belaiba et al., 2007) and EGR1 (Sperandio et al., 2009) were already published to regulate transcription of HIF1A and 62 TFBSs from ChIP-seq were fund to be present nearby *HIF1A* TSS. Additionally, TFBS predictions showed that upstream SNPs could affect alteration of binding sites for six TFs (Figure 11**)**. Analysis showed that three different forkhead/winged helix transcription factors (FOXs) binding sites (FOXA1, FOXC1 and FOXI1) were affected due to two *HIF1A* SNPs: rs1535679 and rs28708675. The FOX family of transcription factors is known to regulate differentiation, metabolism, and development. Moreover, FOXA1 plays a suppressive role in the early phase of adipogenesis, acting under the control of CEBPB (Fujimori and Amano, 2011). The putative transcription factor binding motifs for TAL1 and TCF3 were altered due to a SNP, rs2783778, which is the most distal *HIF1A* SNP tested in association studies. Transcription factor TAL1 plays a role in anti-apoptosis, cell proliferation and megakaryocyte differentiation (Palii et al., 2011; Reynaud et al., 2005). Meanwhile, TCF3 regulates B- and T-cell development (Bain et al., 1994; Bradney et al., 2003), apoptosis, cell cycle (Chiaro et al., 2012; Zhao et al., 2001), and somatic recombination of immunoglobulin gene segments (Lazorchak et al., 2006; Sakamoto et al., 2012). Additionally, sequence variation rs41362550 altered the predicted transcription factor binding sites of HNF4, which has a role in hepatocyte differentiation (Khurana et al., 2010), blood coagulation (Inoue et al., 2006), and its expression is upregulated in hepatocellular cancer to block carcinogenesis and metastasis (Ning et al., 2010). The appearance of type 2 diabetes was also associated with mutation of *HNF4* (Lu et al., 2008; Yamagata et al., 1996). Interestingly, polymorphism rs41362550, whose functional effect is not known and is most proximal

to TSS, was the only one overlapping with TFBSs (42/62) from ENCODE project and is also predicted to be involved in gain/loss of HNF4A TFBS. This SNP is therefore of interest for future functional studies.

Additionally, studies have shown that HIF1 regulates a panel of miRNAs, whereas some of miRNAs have also been shown to target HIF1A (Bruning et al., 2011; Cascio et al., 2010; Cha et al., 2010; Lei et al., 2009; Rane et al., 2009; Rane et al., 2010; Taguchi et al., 2008) (Figure 12A). Therefore, when low $O_2$ level induces expression of HRMs, genes targeted by these miRNAs stabilize *HIF1A* by forming positive-feedback loops (Ghosh et al., 2010; Kelly et al., 2011; Liu et al., 2010; Shen et al., 2013). However, some HRMs destabilize *HIF1A* by direct targeting of 3'-UTR region and these miRNAs are: miR-20b (Cascio et al., 2010; Lei et al., 2009), miR-199a (Rane et al., 2009; Rane et al., 2010), cluster miR-17-92 (Taguchi et al., 2008; Yan et al., 2009), miR-519c (Cha et al., 2010), and miR-155 (Bruning et al., 2011). Recently, some reports revealed that miRNAs are associated with several key signaling pathways that respond to hypoxia and played important roles in hypoxic adaptation (Chan et al., 2009; Chen et al., 2010; Favaro et al., 2010; Puissegur et al., 2011). To investigate if miRNA binding sites are affected by *HIF1A* SNPs we performed prediction analysis for miRNA binding sites on sequences flanking previously tested *HIF1A* SNPs. Our predictions revealed that *HIF1A* SNPs could be involved in gain/loss of miRNA target sites located within exons, 3'- and 5'-UTR. Therefore, predicted miRNA binding sites are distributed through whole gene regions and not only within 3'-UTR. Further experiments can now be performed to analyze if *HIF1A* SNPs are associated with gain/loss of TF binding or miRNA target sites, which are possible explanations for functional effect of *HIF1A* polymorphisms. Variations within *HIF1A* ODD domain (exon 12) and its promoter region are interesting for future analysis, since these might be involved in HIF1A stability or could distract binding of its transcriptional activators within *HIF1A* promoter, respectively. Since HIF1A together with ARNT (HIF1B) functions as master transcription regulator of various genes involved in hypoxia pathway, this would be a great contribution to a known regulatory network, where TFs and miRNA can reciprocally regulate one another. While HIF1A might be responsible for transcription of others TFs and miRNA host genes, in contrast specific TF and miRNA profiles might be responsible for *HIF1A* transcription activity, degradation or mRNA stability. However, it remains an open question if miRNAs that potentially target *HIF1A* promoter region affect the transcriptional regulation with TF and *vice versa*.

### 5.1.3 Cancer deregulated and epigenetically controlled miRNAs in HIF1A regulatory network

MicroRNAs have been proven to be involved in cancer development and progression (Calin and Croce, 2006; Esquela-Kerscher and Slack, 2006). Therefore, miRNA profiles can be very useful in tumor classification, prognosis, and prediction of the therapeutic outcome. Examples of miRNA used as diagnostic tools, as prognostic tools or as predictors of response to therapy were reviewed by Ferdin et al. (2010). Besides, some miRNA expressions were proved to be epigenetically regulated in different types of cancers and cell lines. Since hypoxia is a common characteristic of solid tumors we were prone to identify if among miRNAs predicted to gain/loss their target site due to *HIF1A* polymorphisms, have been previously associated with cancer and potentially regulated by epigenetic mechanisms. We discovered that out of 150 miRNAs predicted to target *HIF1A* 25 were previously associated with cancer, out of which 12 were found to be epigenetically regulated and among them four associated to hypoxia response.

Besides, we identified common regions of cancer deregulated and epigenetically regulated miRNAs compared to location of a gene that code for HIF1A, which is a master regulator of hundreds of genes, including miRNAs. In conclusion, some miRNAs previously associated with cancer are located near chromosome/gene alterations, which clearly suggest that some miRNA gene clusters with as few as three miRNAs may have a greater influence on cancer development, progression, and overall survival than previously thought. MicroRNAs most likely work together, where each miRNA, either alone or as a group, has a specific role in response to stress, development and progression of complex diseases such as cancer.

### 5.1.4 Pathways associated with HIF1A target genes

Pathway enrichment analysis using known HIF1A targets revealed that pathways *metabolism of carbohydrates*, *diabetes pathways*, and *integration of energy metabolism* were most significantly associated to HIF1A targets. Pathway enrichment analysis is important for further investigation to focus on HIF1A targets with significant roles for cells to survive in environment with low $O_2$ level. Therefore genes involved within these enriched pathways are more likely to be suitable as a therapeutic targets for instance in a cancer research or performance endurance improvement studies. To date, functional HREs within promoters of HIF1A targets have been identified as far as 12 kb downstream (Pescador et al., 2005) or 5 kb upstream of a target TSS (Coulet et al., 2003).

Results of the study by Benita et al. (2009) using 21 validated and 60 predicted HIF1A targets and KEGG pathway analysis tool revealed eleven associated pathways. Twelve

of these 60 predicted HIF1A targets were already experimentally validated by now and therefore also included in our pathway analysis. We performed pathway analysis of 274 validated HIF1A targets using DAVID bioinformatics tool, which enables a discovery of enriched functional-related gene groups from four different pathway databases (KEEG, PANTHER, Reactome, and BioCarta) with very stringent criteria (Bonferroni and *P*-value < 0.01). Our results confirmed six enriched pathways previously identified by Benita et al (2009) (marked with stars in ANNEX O) and revealed 15 additional associated enriched pathways.

When $O_2$ levels are insufficient, cell's metabolism shifts from mitochondrial oxidative phosphorylation to glycolysis (Zhang et al., 2007). Increased glycolysis and decreased respiration, even under aerobic conditions is characterized in many cancer cells. However, HIF1 may be involved in induction of genes necessary for this adaptation (Zhang et al., 2007). Therefore, cell proliferation and survival depend on nutrient availability and a tight link between nutrient and hormone signaling pathways (Marshall, 2006). Nutrient signaling pathways are interconnected coupled to insulin pathways and linked to release of metabolic hormones from adipose tissue (reviewed in (Marshall, 2006)). Moreover, hypoxia actually stimulates lipids storage and inhibits lipid catabolism through β-oxidation (Bostrom et al., 2006).

Pathway based systems biology approach may prove to be significant for discovery of novel HIF1A targets within these enriched pathways, including potentially novel HIF1A targets with HRE-binding sites already identified within its promoter region. It is likely that in the future novel HIF1A target genes will be identified since several HRE-binding sites have been predicted within HIF1A targets and remain to be validated (Benita et al., 2009). These potentially novel HIF1A targets could now be tested with additional functional analysis.

## 5.2 VALIDATION OF A NOVEL HIF1A TARGET

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 (Kulshreshtha et al., 2007). 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 (Kulshreshtha et al., 2007) 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 HIF1As on luciferase reporters containing fragments from the candidate promoters of HINCUTs. We identified HIF1A-binding sites upstream of uc.475. This region in fact overlaps *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 HIF1A in close proximity to *OGT* intron 4, where uc.475 is located. Recent reports have demonstrated that many intronic miRNAs possess their own promoter, which is independent to the host gene transcription (Ozsolak et al., 2008). Further investigation is required to reveal whether there are additional TF-binding sites 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 (are viewed in (Butkinaree et al., 2010; Love et al., 2010)). The *OGT* was also recently reported to be overexpressed in breast cancer and primary and metastatic prostate cancer (Caldwell et al., 2010; Lynch et al., 2012), further supporting *OGT* as a therapeutic target in solid malignancies.

This 9-kb hybrid *OGT*-uc.475 transcript is mostly nuclear and leads to the generation of a significant percentage, if not most, of the mature *OGT* mRNA and functional OGT protein(s) that perform most of the detectable O-GlcNAcylation in the cells. Additionally, this transcript contains a significant proportion, if not the majority, of the uc.475-containing RNA pool in the cell, since *OGT*-specific siRNA dramatically decreased the amount of uc.475. Consistently, a siRNA specifically directed against uc.475 sequences, can efficiently inactivated both the ultraconserved transcript and *OGT*. Based on the northern blot pattern of bands, uc.475 may also exist as a separate transcript; however this may be generated from the primary mRNA and it is not feasible to be selectively inactivated independently from *OGT* using siRNA based approaches.

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 *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) (Lubas et al., 1997).

Inhibition of uc.475 by siRNA lead to an increase in the number of cells in the G2/M fraction specifically in hypoxia, which may in fact be related to the inability of cells to maintain adequate O-GlcNAcylation during hypoxic stress. We suggest that at least some of the biochemical effects of uc.475-inhibition are in fact due to the interference with *OGT*. Based on the northern blotting results, the amount of uc.475 that exists outside the hybrid nuclear transcript is relatively modest. We also noticed a contrast between the hypoxia-inducible nature of OGT/uc.475 at RNA and at protein level. Indeed, measurements in multiple cell lines indicate that OGT is either unchanged (Figure 41**)** or even decreased during $O_2$-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 44) suggest that on one hand this sequence may serve as an enhancer, which could cooperate with HIF for elevated transcription. The predominant 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 the mRNA stability. However, some studies showed that intron retention has little or no effects on mRNA half-life (Chang et al., 2006; Nott et al., 2003; Ryu and Mertz, 1989), whereas others confirmed that it is important for mRNA stability and had an effect on prolonged mRNA half-life (Gencheva et al., 2010; Wang et al., 2007; Zhao and Hamilton, 2007). In addition, cells in 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 (Gardner, 2008). Moreover, many studies found that constitutively spliced introns are required for optimal gene expression (Buchman and Berg, 1988; Le Hir et al., 2003; Nott et al., 2003), and one possible way through which introns can affect transcription is by acting as repositories for transcriptional regulatory elements such as enhancers and repressors (Le Hir et al., 2003). 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.

## 5.3 FUTURE PERSPECTIVES

- *HIF1A* polymorphisms associated with phenotypes were found within different regions and domains: including: 5'-upstream region (TFs and miRNAs binding), PAS domain (ARNT binding), N-TAD region within ODD domain (transactivation and $O_2$ dependent regulation of a gene), 3'-UTR region (miRNA binding; in addition to binding sites within exons and promoter), and introns (alternative splicing). All these SNPs are suitable for future functional analysis to determine if they disrupt the role of HIF1A functional domains and therefore influence HIF1A stability and its role as master TF in hypoxia.

- Additional functional studies are needed at least for four prioritized miRNAs (miR-23b, miR-130a, miR-135a, and miR-200b), whose expression was previously shown to change by hypoxia environment and were found deregulated and epigenetically regulated in cancer. These studies should verify whether the presence of SNPs within miRNA binding sites influence their regulation of *HIF1A.*

- Further functional studies are also needed to verify if four SNPs within upstream *HIF1A* region, predicted to alter binding sites for six TFs, influence the transcription activity of *HIF1A*. Polymorphism rs41362550 is particularly interesting, since it is the most proximal to the TSS of a gene, it is overlapping with 42 previously validated TF binding sites from ChIP-seq data, and since its functional role is yet unknown.

- Pathways based systems biology approach using validated HIF1A targets may prove to be significant for discovery of novel HIF1A targets within enriched pathways.

- In addition to uc.475 and its host gene, the role of other HINCUTs should be identified. In our opinion uc.63 and its host gene *exportin 1* (*XPO1*) are very promising for the future analysis.

# 6 CONCLUSIONS (SKLEPI)

## 6.1 CONCLUSIONS

- We constructed a central Atlas of *HIF1A* gene regulatory network, consisting of: 1.) functional *HIF1A* polymorphisms from 70 association studies; 2.) validated and predicted HIF1A interactions with TFs and miRNAs; 3.) HIF1A downstream targets and pathways. The developed Atlas presents a central location for genomic information related to HIF1A, which will accelerate the research of this field.

- Our systematic review of *HIF1A* association studies revealed that 15/30 tested *HIF1A* variations were associated with 32 phenotypes including 16 cancer types.

- Five miRNAs were previously identified as upstream regulators of *HIF1A* and additional 150 miRNAs were predicted to target polymorphic regions residing within *HIF1A* exons, 5'- and 3'-UTR regions. Out of 150 miRNAs predicted to target regions of *HIF1A,* 25 were previously found to be deregulated in cancer and 12 of those epigenetically regulated in cancer, including four miRNAs whose expression is already known to be changed by hypoxia environment. Further experiments can now be performed to analyze if *HIF1A* SNPs are associated with gain/loss of miRNA target sites. This would also partly explain the functional effect of *HIF1A* polymorphisms.

- Out of 62 TFBSs identified within upstream *HIF1A* region from ChIP-seq data (UCSC, ENCODE project), 42 overlapped with polymorphism rs41362550. Six additional TFBS were predicted to be gained/lost due to four SNPs. Further experiments can now be performed to analyze if *HIF1A* SNPs are associated with gain/loss of TF binding sites, which is another possible explanation for functional effect of *HIF1A* polymorphisms.

- Pathway enriched analysis revealed that HIF1A targets are involved in significant pathways including: *diabetes pathways*, *metabolism of carbohydrates*, and *integration of energy metabolism*. The proposed pathways have potential for development of new generation of pathway based hypoxia and cancer biomarkers.

- We confirmed the hypothesis that in hypoxia the expression of transcribed ultraconserved regions (T-UCR) alters, since in multiple cell types hypoxia-induced noncoding ultraconserved transcripts, which we termed HINCUTs, were identified. *This conclusion refers to hypothesis 1.*

- HIF1A was confirmed to play a critical role for hypoxic regulation of HINCUTs based on induction with: 1.) chemical inactivators of the HIF prolyl hydroxylases, 2.) effect of the VHL tumor-suppressor inactivation and 3.) effect of exogenous HIF1As on luciferase reporters containing fragments from the candidate promoters of HINCUTs. *This conclusion refers to hypothesis 2.*

- We showed that studied T-UCRs were predominantly nuclear and that HIF1A is partly responsible for the induction of several members of this group. The transcript uc.475 or HINCUT-1, analyzed in this study, is part of a retained intron of the host protein-coding gene, *OGT* (O-linked N-acetylglucosamine transferase), which is overexpressed in epithelial cancer types. *This conclusion refers to hypothesis 3.*

- Based on the northern blot, we suggest uc.475 may also exist as a separate transcript; however this may be generated from the primary mRNA transcript that also includes the coding sequences of its host gene *OGT*. Therefore uc.475 transcript is not feasible to be selectively inactivated independently from *OGT* using siRNA based approaches.

- Inhibition of uc.475 by siRNA lead to an increase in the G2/M fraction specifically in hypoxia, which may in fact be related to the inability of cells to maintain adequate O-GlcNAcylation during hypoxic stress. With that we also confirmed that silencing of overexpressed T-UCR in hypoxia represses cancer cell line proliferation. *This conclusion refers to hypothesis 3 and 4.*

- The luciferase experiments suggested that tested sequence containing uc.475 may serve as an enhancer, which could cooperate with HIF for elevated transcription of *OGT*.

## 6.2 SKLEPI

- Izdelali smo centralni Atlas regulatorne mreže gena *HIF1A*, ki obsega: 1.) funkcijske polimorfizme gena *HIF1A* zbranih iz 70-ih analiz povezav genotipa s fenotipom; 2.) znane in predvidene interakcije HIF1A z drugimi transkripcijskimi dejavnki (TF) in miRNA; 3.) tarčne gene HIF1A in njihovo vpletenost v biološke poti. Izdelani Atlas predstavlja centralno mesto genomskih informacij v povezavi s HIF1A, kar bo omogočilo pospešitev raziskav na tem področju.

- Naš sistematski pregled analiz povezav genotipa s fenotipom je pokazal, da je bilo 15 od 30 testiranih *HIF1A* SNP-jev povezanih z 32 fenotipi in boleznimi, vključno s 16 tipi raka.

- Za pet miRNA je bilo že dokazano, da uravnavajo izražanje *HIF1A*. Bioinformacijska analiza je pokazala, da se na mestih preučevanih polimorfizmov znotraj eksonov, 5'- in 3'-neprevedenih regijah gena *HIF1A* nahajajo predpostavljena vezavna mesta za dodatnih 150 miRNA. Med temi 150 miRNA, ki predvideno prepoznavajo območja *HIF1A*, je bilo za 25 že dokazano spremenjeno izražanje pri raku, za 12 med njimi, da so epigenetsko uravnavane pri raku, in za štiri izmed teh je bilo dokazano spremenjeno izražanje v okolju hipoksije. Potrebne so nadaljnje raziskave za potrditev vpletenosti *HIF1A* SNP-jev v pridobitev/izgubo tarčnih mest za miRNA. S tem bi deloma razložili vlogo polimorfizmov gena *HIF1A*.

- Ugotovili smo, da se izmed 62 identificiranih vezavnih mest za TF-je iz ChIP-seq (UCSC, projekt ENCODE) v območju navzgor od gena *HIF1A*, 42 prekriva s polimorfizmom rs41362550. Dodatnih šest TF-jev je bilo predvidenih, da pridobijo/izgubijo svoje vezavno mesto zaradi prisotnosti štirih SNP-jev gena *HIF1A*. Potrebne so nadaljne raziskave za preverjanje ali so *HIF1A* SNP-ji znotraj odgovorni za pridobitev/izgubo vezavnih mest TF-jev. To bi bila dodatna obrazložitev vloge polimorfizmov gena *HIF1A* na njegovo delovanje.

- Analiza bioloških poti, v katere so uvrščene tarče HIF1A, je pokazala, da je večina teh vključena v poti: *diabetes pathways, metabolism of carbohydrates,* in *integration of energy metabolism*. Predlagane biološke poti imajo potencial za razvoj nove generacije biooznačevalcev za hipoksijo in raka.

- Potrdili smo hipotezo, da se s hipoksijo spremeni tudi izražanje prepisov T-UTR, saj smo v številnih tipih celic določili povišano izražanje prepisov teh nekodirajočih regij, ki smo jih poimenovali HINCUTs (angl. *hypoxia induced noncoding ultraconserved transcripts*). *Ta zaključek se nanaša na hipotezo 1.*

- HIF1A je bil v hipoksiji potrjen kot regulator izražanja HINCUT-ov, na podlagi njegove indukcije: 1.) s kemikalijami, ki inaktivirajo HIF prolil hidroksilaze, 2.) z inaktivacijo tumor-supresorskega gena VHL ter 3.) z vnosom ekspresijskih vektorjev z zapisom za HIF1A in njihov vpliv na poročevalce luciferaze, ki so vsebovali promotor pručevanega HINCUT-a. *Ta zaključek se nanaša na hipotezo 2.*

- Dokazali smo, da se preučevani T-UCR-ji nahajajo pretežno v jedru celice in da je HIF vsaj deloma odgovoren za indukcijo izražanja nekaterih izmed njih. Prepis uc.475 ali HINCUT-1 preučevan v tej študiji, je del neizrezanega introna gostiteljskega gena *OGT* (*O-linked N-acetylglucosamine (GlcNAc) transferase*), ki je čezmerno izražen pri raku epitelijskih celic. *Ta zaključek se nanaša na hipotezo 3.*

- Na podlagi rezultatov prenosa northern sklepamo, da lahko regija uc.475 obstaja tudi kot samostojen prepis ali pa predstavlja del primarnega transkripta, ki vsebuje tudi kodirajoče regije gostiteljskega gena *OGT*. Zato torej s siRNA ni mogoče samostojno utišati regije uc.475, da s tem ne bi vplivali tudi na raven gena *OGT*.

- Inhibicija regije uc.475 s siRNA je predvsem v hipoksiji povzročila povišanje deleža celic v G2/M fazi celičnega cikla. To bi lahko bila tudi razlaga, da je celica po inhibiciji uc.475 v hipoksiji neuspešna vzdrževati ustrezno raven O-GlcNAcilacije. To pomeni tudi, da ima utišanje prekomernega izražanja T-UCR v hipoksiji vpliv na zmanjšano pomnoževanje celic rakavih linij. *Ta zaključek se nanaša na hipotezo 3 in 4.*

- Na podlagi rezultatov luciferaznih eksperimentov predlagamo, da ima preučevana regija uc.475 vlogo ojačevalca, ki lahko skupaj s HIF vpliva na povišano raven prepisovanja gena *OGT*.

# 7 SUMMARY (POVZETEK)

## 7.1 SUMMARY

In cancer environment level of $O_2$ gets decreased by the depth and distance of blood flow in vessels. In mammal cells primary response to low $O_2$ conditions (hypoxia) is made rapidly through adaptation to conditions they are being captured. In hypoxic cells important responses are activated for metabolic, bioenergetics, and redox demands and primary transcriptional response to hypoxia is mediated by the hypoxia-inducible factors (HIFs).

HIF heterodimer belongs to a family of transcription factors (TF) responsive to a low $O_2$ availability, which is often a characteristic of solid tumors. Alpha subunit of HIF heterodimer is $O_2$-sensitive, and once stabilized in hypoxia, it functions as master regulator of various genes involved in hypoxia pathway. Identification of upstream *HIF1A* regulators and its downstream targets are essential for understanding the cellular and systemic response to hypoxia. Despite its importance as transcriptional regulator, HIF1A regulatory network is still not yet completely understood and data is scattered among publications and databases. Therefore, we created a central Atlas of *HIF1A* gene regulatory network integrating both the existing knowledge and results of bioinformatics analyses.

Previously studies showed that microRNA (miRNA) are among ncRNA already known to be regulated by low $O_2$ levels. Study on cancer cell lines, breast and colon, confirmed that a group of hypoxia regulated miRNAs can be influenced by variety of hypoxic factors and were proved to decrease pro-apoptotic signaling in a hypoxic environment, suggesting an impact of these transcripts even on tumor formation. However, it has been suggested that HIF gives a big contribution to cancer cell line adaptation in hypoxia. Even though that recent data linked hypoxia, a classic feature of the tumor microenvironment, to the function of specific miRNAs, very little is known how hypoxia affects other types of noncoding transcripts. Moreover, microarray assays on clinical colorectal cancer (CRC) samples revealed altered expression not only for miRNAs but also for “transcribed ultraconserved regions” (T-UCRs). Therefore we analyze if hypoxia and their factors influence on expression of T-UCRs, since to date the association between T-UCRs and HIF1A has not been confirmed yet. By now altered T-UCRs expression was described in colorectal cancer and hepatocellular carcinomas, hematologic malignancies, and neuroblastoma.

To create an Atlas of *HIF1A* gene regulatory network we used PubMed and WoS to collect functional *HIF1A* polymorphisms from HIF1A association studies, to collect

cancer deregulated and epigenetically regulated miRNAs, and all downstream targets of HIF1A. Experimentally confirmed *HIF1A* TF binding sites (TFBS) were extracted from publications and ENCODE ChIP-Seq data and TFBS predictions were performed using TRANFAC Match tool. Validated *HIF1A*-miRNA interactions were extracted from miRWalk, miRecords, and mirTarBase databases and miRNA binding was predicted using RNA22 and miRanda algorithms. Pathway enrichment analysis was applied using DAVID Bioinformatics Resources 6.7.

To find a functional link between hypoxia, HIF1A and T-UCR we re-interrogated previously published microarray data of CRC samples. On the selected T-UCR a hypoxia screening (1% $O_2$ level and exposure to DMOG) was performed and their expression was also tested in a cell model with proficient/deficient VHL protein. T-UCR expression was measured using RT-qPCR method. Protein levels altered by hypoxia and gene silencing experiments were detected using western blot analysis. T-UCR promoters with HIF1A-binding sites were predicted using TFRANSFAC tool. Chromatin immunoprecipitation (ChIP) was used to validate predicted HIF1A-binding sites. Promoter fragments with or without HIF1A-binding sites were cloned upstream of luciferase reporter, while T-UCR sequence that was tested for enhancer activity was cloned downstream of luciferase reporter gene. Dual-Luciferase reporter assays were performed to measure effect of cloned sequences into reporter vectors. Knock down experiments using specific siRNA sequences were used to measure biological significance of tested T-UCRs; cell counting for effect on proliferation and FACS analysis to determine cell cycle affects. Northern blot using gene specific probes was performed to identify transcript present in normoxia and altered in hypoxia. Open-reading frame (ORF) finder was used for detection of ORF within tested T-UCR sequences.

An Atlas of *HIF1A* gene regulatory network that we constructed comprises of: 1.) functional *HIF1A* polymorphisms from 70 association studies; 2.) validated and predicted TF and miRNA binding sites; 3.) HIF1A downstream targets and pathways. To date, 30 *HIF1A* single nucleotide polymorphisms (SNPs) were genotyped and 15 of them were found to be associated with 32 phenotypes including 16 cancer types. Out of 62 TFBSs identified within upstream HIF1A region, 42 overlapped polymorphism rs41362550 and six were predicted to be gained/lost due to four SNPs. Five miRNAs were previously identified to regulate *HIF1A* and 150 miRNAs were predicted to target polymorphic regions residing within exons, 5'- and 3'-UTR regions. Out of 150 miRNAs targeting regions of *HIF1A,* 25 were previously found deregulated and 12 of those epigenetically regulated in cancer, including four miRNAs whose expression was already reported to be deregulated in hypoxia environment. Downstream HIF1A targets were suggested to be involved in 21 enriched pathways.

To validate a novel target of HIF1A, we started with a genome-wide expression profiling and we demonstrated for the first time a functional link between $O_2$ deprivation and long noncoding transcripts of ultraconserved regions, termed transcribed-ultraconserved regions (T-UCRs). Interestingly, several hypoxia-upregulated T-UCRs, 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 is 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, OGT (O-linked N-acetylglucosamine transferase), which is overexpressed in epithelial cancers and important for post-translational modifications of many proteins in the cell.

Our results suggest that *HIF1A* variations might be involved in gain/loss of TF and miRNA binding sites, which could be a possible explanation for functional effect of *HIF1A* polymorphisms. The developed genetic regulatory Atlas of HIF1A gene regulatory network presents a starting point for further research related to HIF1A and an example of an Atlas of regulatory network for any TF of interest.

In addition, we validated a novel target of HIF1A and 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 $O_2$ tension is limiting. Our data gives a first glimpse of a novel functional hypoxic networks composing of protein-coding transcripts and noncoding RNAs from the T-UCRs category.

## 7.2 POVZETEK

V celicah sesalcev je primaren odziv na nizko vsebnost $O_2$ hiter, saj se morajo le-te čimhitreje prilagoditi danim pogojem. V hipoksičnih celicah so zato aktivirani signali, potrebni za zagotovitev metabolnih, bioenergetskih in redoks potreb celice (pregled v Majmundar in sod., 2010). Primarni odzivi se sprožijo tudi s transkripcijskimi dejavniki HIF, ki šele v okolju z znižano vsebnostjo $O_2$ postanejo aktivni. Proteini HIF so znani osrednji regulatorji hipoksije in imajo poleg vpliva na odziv ter prilagoditev celice okolju tudi vlogo pri različnih fizioloških in patoloških procesih celice (Majmundar in sod., 2010; Semenza, 2003). Protein HIF je heterodimer sestavljen iz na $O_2$-občutljive (alfa, A) in -neobčutljive enote (beta, B) (Wang in Semenza, 1993). V celicah sesalcev so znane tri izooblike proteina HIF-A, med katerimi je HIF1A izražen konstantno, medtem ko je prisotnost proteinov HIF2A in HIF3A odvisna od vrste tkivnih celic (Bertout in sod., 2008). V zadnjih nekaj letih je bilo na področju uravnavanja in delovanja proteinov HIF narejenih že veliko raziskav (Wang in sod., 1995), saj le-ta v hipoksiji uravnava prepisovanje velikega števila genov, ki omogočajo nadzor zaloge $O_2$ v celici in njeno preživetje v okolju z nizko vsebnostjo $O_2$ (Keith in sod., 2012; Vilela in sod., 2008; Wang in sod., 1995; Wenger in sod., 2005). Prav tako je veliko znanega tudi o vlogi HIF1A pri nastanku bolezni in številnih fenotipov (Majmundar in sod., 2010; Semenza, 2003), vendar je znanje o HIF1A precej razpršeno med publikacijami. Atlas regulatorne mreže HIF1A smo izdelali z namenom, da bi zbrali vse genomske informacije o HIF1A na enem mestu, jih uredili ter na tej podlagi razkrili nove smernice za raziskave uravnavanja izražanja HIF1A in njegovega vpliva.

Za izdelavo Atlasa regulatorne mreže *HIF1A* smo iz literarnih zbirk PubMed in WoS zbrali publikacije, povezane z znanimi regulatorji izražanja gena *HIF1A*. Določitev regulatorjev izražanja HIF1A in njegovih tarčnih genov je ključno za razumevanje celičnega in sistemskega odziva na hipoksijo. Kljub temu da je bila hipoksija kot običajna lastnost mikrookolja tumorjev že povezana s funkcijo miRNA, je trenutno malo znanega, kako le-ta vpliva na druge skupine ne-kodirajočih prepisov (Ferdin in sod, 2013). Atlas regulatorne mreže gena HIF1A tako sestavljajo: 1) funkcionalni polimorfizmi HIF1A iz 70-ih analiz povezav genotipa s fenotipom; 2) potrjena in predvidena vezavna mesta za transkripcijske dejavnike (TF) in miRNA; 3) HIF1A tarče in biološke poti, v katere so uvrščene. Eksperimentalno potrjene transkripcijske dejavnike, ki se vežejo v promotorsko območje gena HIF1A, smo pridobili iz podatkovne zbirke UCSC (ChiP-sequencing projekt ENCODE). Predvidena vezavna mesta za TF-je v promotorju gena *HIF1A* smo določili z uporabo bioinformacijskega orodja TRANFAC Match. Že znane interakcije miRNA-*HIF1A* so bile zbrane iz podatkovnih zbirk (miRWalk, miRecords, in mirTarBase), analiza predvidenih vezavnih mest za miRNA pa je bila izvedena z uporabo algoritmov RNA22 in miRanda. Iz literature smo zbrali tudi miRNA s spremenjenim izražanjem pri raku, vključno s

tistimi, ki so epigenetsko uravnavane, in miRNA, katerih spremenjeno izražanje je bilo dokazano v hipoksiji. Protein HIF1 je dokazan regulator več sto tarčnih genov, med katerimi so prisotne tako protein-kodirajoči kot tudi ne-kodirajoči zapisi (miRNA). Da bi dobili čimboljši vpogled v regulatorno mrežo gena *HIF1A*, je potrebno poleg tarč določiti tudi regulatorje, ki lahko vplivajo na njegovo izražanje.

Do danes je bilo genotipiziranih že 30 polimorfizmov posameznega nukleotida (SNP) gena *HIF1A* in 15 izmed njih povezanih z 32 fenotipi in boleznimi. Pregled literature je pokazal, da so bili polimorfizmi gena *HIF1A* pogosto povezani z rakom, srčno-žilnimi boleznimi in aktivnostjo. Preučevani SNP-ji so bili v analizah povezav genotipa s fenotipom kar v polovici primerov povezani z rakom (16/32 fenotipov). Do sedaj je bilo sicer že veliko tumorjev povezanih z nizko vsebnostjo $O_2$, kar je lahko posledica prekomernega pomnoževanja rakavih celic, saj le-te povzročijo prerast žilnega sistema tumorja in posledično omejijo difuzijo $O_2$ v njegovo notranjost. Hipoksični stres v celicah je lahko tudi posledica motnje pretoka krvi zaradi spremenjene strukture in delovanja ožilja tumorjev (Majmundar in sod., 2010). Posledično to privede do povišane ravni HIF1A v celicah (Bertout in sod., 2008), katerega raven je lahko dodatno zvišana tudi neodvisno od poti gena *HIF* (Majmundar in sod., 2010). Protein HIF1A ima sposobnost reprogramirati metabolični stadij celic, kar je nujno za hipoksično okolje tudi v primeru bolezni ožilja in pri raku (Majmundar in sod., 2010). Na ta način HIF1 v hipoksičnih celicah uravnava prepisovanje mnogih genov, ki so vpleteni v ključne vidike biologije raka kot so: imortalizacija celic, vzdrževanje zarodnih celic, diferenciacija celic, genetska nestabilnost, vaskularizacija, metabolno reprogramiranje celic, avtokrina signalizacija rastnega dejavnika, invazija in metastaziranje celic ter odpornost na zdravljenje (Semenza, 2007a). Poleg tega je HIF1A pogosto dejavnik slabe klinične napovedi za številne tipe raka (Semenza, 2007a).

Povezava HIF1A s srčno-žilnimi boleznimi je le dodatni pokazatelj sistemskega vpliva HIF1A. Srčno-žilne bolezni, kot so arterioskleroza tkiv srca, možgan in mišic okončin, so pogosto dovzetne za poškodbe, ki so posledica pomanjkanja pretoka krvi (ishemija) (Beckman in sod., 2002; Kett-White in sod., 2002). Poleg tega je ishemija srčne mišice najpogostejši vzrok za nastanek srčne hipoksije oz. kapi v kliniki in je posledica okolja, kadar količina dovedenega $O_2$ ne zagotovi metabolnih potreb srčnih celic (Shohet in Garcia, 2007). Izražanje HIF1A je ključno, da se ponovno vzpostavi pretok krvi v ishemične predele skeletnih mišic (Majmundar in sod., 2010). Zmanjšana aktivacija HIF1A je bila ugotovljena tudi v hipoksičnih ranah na koži ostarelih mišk z diabetesom (Liu in sod., 2008b), kar izpostavi tudi pomen starosti pri odzivu na ishemijo. Pri človeku so periferna arterijska obolenja prav tako povezali s starostjo (Beckman in sod., 2002). Vloga HIF1A pri angiogenezi je tako znana tudi s hipertrofno kardiomiopatijo, odpovedjo srčne mišice, celjenjem kožnih ran in neovaskularizacijo retine (pregled v Majmundar in sod., 2010).

Veliko študij je predvidevalo, da do pojava lokalne hipoksije v mišicah med telovadbo pride zaradi padca pritiska $O_2$ v mišičnih celicah (miocite). To naj bi v mišicah povzročilo indukcijo signalne poti, povezane s HIF1A, in tako pospešilo uporabo oksidacijskih poti za proizvajanje energije (Doring in sod., 2010), kar se posledično odraža v aktivnosti človeka. Srčno-žilni sistem je ključen za telesno aktivnost pri človeku, zato so analize povezav genotipa s fenotipom le dodatna potrditev sistemskega vpliva HIF1A.

Pregled objavljenih analiz povezav genotipa s fenotipom je pokazal, da se je večina teh osredotočila na dva nesinonimna *HIF1A* SNP-ja (nsSNP-ja) znotraj domene ODD: rs11549465 (*p.Pro582Ser*) in rs11549467 (*p.Ala588Thr*). Preučevanje teh dveh SNP-jev je pokazalo kar nekaj nasprotujočih povezav, kar je lahko odraz tega, da sta bila predmet proučevanja v številnih študijah. Oba SNP-ja, rs11549465 in rs11549467, imata lokacijo znotraj domene interakcij med ODD/pVHL (Clifford in sod., 2001). Polimorfizem rs11549465 je bil povezan tudi s pospešitvijo transaktivacije tarčnih genov proteina HIF1A (Huang in sod., 1998; Tanimoto in sod., 2003), kljub temu da ni bil identificiran za mesto na *HIF1A*, ki je hidroksilirano in vključeno v interakcijo z VHL (Yamada in sod., 2005). Ti rezultati so v skladu s predhodnjimi raziskavami Percy in sod. (2003), ki so dokazali, da zamenjava *p.Pro582Ser in vitro* ne prepreči vezave VHL na HIF1A po hidroksilaciji na mestu *p.Pro564*.

Tako kot večina analiz povezav genotipa s fenotipom, so se tudi študije, ki so preučevale vpliv polimorfizmov gena *HIF1A*, osredotočile predvsem na nsSNP-je in veliko manj pozornosti namenile sinonimnim (sSNP-ji) ter drugim tipom ne-kodirajočih SNP-jev. To bi bila lahko posledica dolgoletnega prepričanja, da sSNP-ji nimajo velikega vpliva na fenotip glede na to, da ne vplivajo na spremembo zaporedja amino-kislin. Četudi se je večina analiz povezav genotipa s fenotipom osredotočila na preučevanje nsSNP-jev domene ODD gena *HIF1A* in ima od $O_2$-odvisna razgradnja primaren pomen pri uravnavanju transkripcijskih dejavnikov HIF, je možno, da imajo poleg domene ODD tudi polimorfizmi znotraj različnih domen *HIF* vpliv na aktivnost proteina HIF1A. Predvidevamo, da na vlogo domen *HIF1A* in posledično na njegovo stabilnost in delovanje kot glavnega transkripcijskega dejavnika hipoksije, lahko vplivajo: i) polimorfizmi znotraj domene bHLH, ki lahko preprečijo vezavo HIF1A na vezavna mesta HRE znotraj promotorja tarčnih genov. Poleg tega lahko tudi polimorfizmi znotraj HRE povzročijo nastanek ali izgubo vezavnega mesta za HIF1A in posledično vpliv na njegove tarče; ii) polimorfizmi znotraj domene PAS bi lahko ovirali dimerizacijo z ARNT (HIF1B), kar bi povzročilo nesposobnost proteina HIF1A v vlogi regulatorja transkripcije; iii) polimorfizmi v domeni ODD lahko vplivajo na obstoj proteina v normoksiji, saj so ostanki prolina (*p.Pro402* in *p.Pro564*) običajno označeni za razgradnjo s kompleksom VHL/proteasom; iv) polimorfizmi znotraj regije NLS

(angl. *nuclear location signal*), ki je odgovorna za oznako proteina in prenos v jedro celice, lahko vplivajo na prenos proteina v jedro; in v) polimorfizmi znotraj domen N-TAD in C-TAD imajo možnost vpliva na sposobnost transkripcijske aktivacije proteina HIF1A in interakcijo s svojimi koaktivatorji. Naš sistematičen pregled analiz povezav genotipa s fenotipom je pokazal, da imajo tudi polimorfizmi *HIF1A* izven domene ODD funkcionalen vpliv, saj so bili le-ti povezani z različnimi boleznimi in fenotipi. Povezave genotipa s fenotipom so bile dokazane za sledeča območja *HIF1A* in SNP-je: območje vezave TF-jev (štirje SNP-ji znotraj 5' območja navzgor od gena *HIF*: rs7148720, rs2783778, rs1535679 in rs28708675), območje vezave z ARNT (*p.Val116Glu* znotraj domene PAS), domena ODD (štirje SNP-ji: rs41508050, rs199775054, rs11549467 in rs11549465, katerega lokacija je znotraj N-TAD domene ODD), 3'-UTR (rs2057482), območje navzdol od gena (rs1319462), in znotraj intronov (štirje SNP-ji: rs12434438, rs1957757, rs10873142 in rs10645014).

Kot že omenjeno, lahko polimorfizmi *HIF1A* predvideno vplivajo na vezavo z drugimi kofaktorji (EP300/CBP) ali kompleksi (jedrni prenašalci, VHL/proteasom, PHD1-3). V analizah povezav genotipa s fenotipom je bilo preučevanih sedem SNP-jev znotraj domene ODD, ki jo za razgradnjo prepozna VHL/proteasom. Ta regija se prekriva z domeno N-TAD, odgovorno za vezavo s kofaktorji, znotraj katere so bili 3/7 SNP-jev analizirani za povezavo s fenotipom in le za rs15549465 je bila ta povezava potrjena. Kljub temu da je veliko preučevanih SNP-jev gena *HIF1A* tudi intronskih (13/30), do danes za še nobenega ni bil dokazan vpliv na proces izrezovanja intronov. V podatkovni zbirki Ensembl je navedenih pet polimorfizmov *HIF1A*, ki se nahajajo na območju prvih treh nukleotidov eksonov ali 3-8 nukleotidov intronov: TMP_ESP_14_62188452, rs200037213, COSM193393, rs200439820, TMP_ESP_14_6221364, vendar nobeden izmed njih še ni bil preučevan v analizah povezav genotipa *HIF1A* s fenotipom.

Območja navzgor in navzdol od gena *HIF1A* so pomembna za vezavo transkripcijskih dejavnikov in funkcionalnih ne-kodirajočih RNA, kot so miRNA. Tako polimorfizmi v promotorju, 5'- in 3'-UTR regijah gena *HIF1A* predstavljajo mesta, ki lahko spremenijo njegovo prepisovanje ali stabilnost prepisov ter posledično vplivajo na izražanje njegovih tarč (Hong in sod., 2007). Predhodno objavljene študije so že eksperimentalno potrdile interakcije med *HIF1A* in TF ter miRNA, naša analiza zaporedij *HIF1A* za TF in miRNA vezavna mesta pa je pokazala, da tudi polimorfizmi v *HIF1A* lahko vplivajo na nastnek/izgubo vezavnih mest za TF-je in miRNA.

Povezava med okoljem hipoksije in miRNA (HRM; angl. *hypoxia-regulated miRNA*) je bila do danes opisana že večkrat, za nekatere so celo ugotovili, da imajo pomembno vlogo pri preživetju celic v hipoksičnem okolju (Crosby in sod., 2009; Devlin in sod., 2011; Kulshreshtha in sod., 2008; Kulshreshtha in sod., 2007). Kljub temu da je HIF1A dokazan regulator nekaterih miRNA, je bilo tudi za določene miRNA ugotovljeno, da

kot svojo tarčo prepoznajo gen *HIF1A*. Poznanih je pet miRNA, ki z vezavo v območje 3'-UTR gena *HIF1A*, povzročijo njegovo destabilizacijo in razgradnjo: miR-20b (Cascio in sod., 2010; Lei in sod., 2009), miR-199a (Rane in sod., 2009; Rane in sod., 2010), gruča miR-17-92 (Taguchi in sod., 2008; Yan in sod., 2009), miR-519c (Cha in sod., 2010) in miR-155 (Bruning in sod., 2011). Nedavne študije so celo ugotovile, da so miRNA povezane s ključnimi signalinimi potmi, ki se aktivirajo v hipoksiji in imajo posledično pomembno vlogo pri prilagajanju celic na takšno okolje (Chan in sod., 2009; Chen in sod., 2010; Favaro in sod., 2010; Puissegur in sod., 2011). V naši raziskavi smo želeli preveriti, ali se s prisotnostjo *HIF1A* SNP-jev lahko spremenijo tudi vezavna mesta za miRNA. Zato smo analizirali, če se v območjih polimorfizmov *HIF1A* z vplivom na fenotip nahajajo predvidena vezavna mesta za miRNA. Rezultati so pokazali pridobitev/izgubo tarčnih mest za 150 miRNA, katerih vezavna mesta so bila razporejena preko celotnega gena (v eksonih, 3'- in 5'-UTR območjih) in ne samo v območju 3'-UTR. Ker je bilo za miRNA že dokazano, da so vpletene v nastanek in razvoj raka (Calin in Croce, 2006; Esquela-Kerscher in Slack, 2006) in da so vzorci njihovega izražanja lahko uporabljeni pri klasifikaciji tumorjev, prognozi in uspešnosti zdravljenja tumorjev, smo za teh 150 predvidenih miRNA želeli ugotoviti, koliko med njimi jih je že bilo spremenjeno izraženih pri raku, vključno z epigenetsko uravnavanimi, ter za katere med njimi je bilo ugotovljeno spremenjeno izražanje tudi v okolju hipoksije. Rezultati povezovanja treh različnih seznamov miRNA so pokazali, da je bilo med 150 predvidenih miRNA, 25 miRNA že povezanih z rakom, od teh 12 miRNA tudi epigenetsko uravnavanih. Med temi miRNA je bilo za štiri (miR-23b, miR-130a, miR-135a in miR-200b) dokazano, da so spremenjeno izražene tudi v hipoksiji in zato predstavljajo nove potencialne regulatorje gena *HIF1A*. Običajno se zapisi za miRNA, ki so že bile povezane z rakom, nahajajo blizu kromosomskih/genskih sprememb, zato se predvideva, da imajo lahko skupine več miRNA obsežnejši vpliv na razvoj, napredovanje raka in preživetje. Vsaka miRNA, posamično ali v skupini, ima specifično vlogo pri odzivu na stres, razvoj in napredovanje kompleksnih bolezni.

Poleg genov za miRNA protein HIF1A dokazano uravnava izražanje tudi nekaterih transkripcijskih dejavnikov (TF), kot sta TWIST1 (Yang in sod., 2008) in GATA1 (Zhang in sod., 2012). Poznani so tudi regulatorji, ki uravnavajo prepisovanje *HIF1A*, in to so SP1 (Minet in sod., 1999), NFKB1/RELA (Belaiba in sod., 2007) in EGR1 (Sperandio in sod., 2009). V območju navzgor od gena *HIF1A* (od TSS; angl. *transcription start site*) je bilo v okviru projekta ENCODE (UCSC) z metodo Chip-sequencing določenih 62 vezavnih mest za različne TF-je. Rezultati naše analize vezavnih mest za TF-je na zaporedjih petih SNP-jev navzgor od gena *HIF1A* iz analiz povezav so pokazali, da se zaradi štirih SNP-jev spremenijo vezavna mesta za šest TF-jev. Vezavna mesta za tri predstavnike družine FOX (angl. *forkhead/winged helix transcription factors*; FOXA1, FOXC1 in FOXI1) so bila spremenjena s rs1535679 in rs28708675. Predvideno je bilo tudi, da se zaradi SNP-ja, rs2783778 spremenita

vezavni mesti za transkripcijska dejavnika TAL1 in TCF3. Za polimorfizem rs41362550 je bilo predvideno, da naj bi v *HIF1A* zaporedju spremenil vezavno mesto za transkripcijski dejavnik HNF4A. Zanimivo je tudi, da se je edino polimorfizem najbližje TSS gena *HIF1A*, rs41362550, katerega vpliv še ni poznan, prekrival z znanimi vezavnimi mesti TF-jev (42/62) določenih z metodo ChIP-sequencing projekta ENCODE (UCSC).

Polimorfizmi *HIF1A* znotraj domene ODD (exon 12) in njegovega promotorja so še posebno zanimivi za nadaljnje raziskave, s katerimi bi lahko preverili, ali preučevani polimorfizmi vplivajo na prepisovanje/stabilnost *HIF1A* oziroma preprečijo vezavo njegovih regulatorjev prepisovanja v območju promotorja. Te raziskave bi pojasnile, ali so SNP-ji res vzrok za pridobitev/izgubo vezavnih mest za TF-je in miRNA, to pa bi bila tudi dodatna razlaga vpliva polimorfizmov na delovanje HIF1A. Kljub temu ostaja odprto vprašanje, ali miRNA, ki prepoznavajo promotor *HIF1A*, lahko tekmujejo za vezavna mesta TF-jev in obratno.

Po pregledu vpliva polimorfizmov na vezavna mesta miRNA in TF-jev smo iz literature zbrali vse do sedaj eksperimentalno dokazane tarče HIF1A, od tega 268 protein kodirajočih in šest ne-kodirajočih (pet miRNA in en UCR). Za analizo obogatenosti bioloških poti s HIF1A tarčami smo uporabili bioinformacijsko orodje DAVID, ki omogoča odkritje funkcionalno povezanih skupin genov iz štirih različnih baz bioloških poti (KEGG, PANTHER, Reactome, in BioCarta). Rezultati analize so pokazali, da so tarče proteina HIF1A uvrščene v 21 različnih bioloških poti, največ theh pa v sledeče biološke poti: *metabolism of carbohydrates, diabetes pathways*, in *integration of energy metabolism*. V številnih študijah so nove tarče HIF1A določili na podlagi izražanja genov v hipoksiji in iskanju ohranjenih vezavnih mest za HIF1A (HRE; angl. *hypoxia-response element*) znotraj promotorjev teh genov. Eksperimentalno dokazana vezavna mesta za HIF1A so bila od začetka prepisovanja tarč oddaljena tudi do 12 kb navzdol (Pescador in sod., 2005) ali 5 kb navzgor od začetka tarčnega gena (Coulet in sod., 2003). Rezultati študije Benita in sod. (2009), kjer so uporabili 21 preverjenih in 60 predvidenih tarč HIF1A ter KEGG-ovo orodje za analizo poti, so pokazali povezavo tarč z enajstimi biološkimi potmi. Dvanajst od 60 predvidenih tarč HIF1A je bilo do danes že eksperimentalno potrjenih in vključenih tudi v našo analizo obogatenosti bioloških poti. Z upoštevanjem potrjenih HIF1A tarč iz študije Benita in sod (2009) smo z našo analizo potrdili šest bioloških poti njihove študije ter določili dodatnih 15 poti, kamor so bile uvrščene ostale tarče HIF1A zbrane za izdelavo Atlasa regulatorne mreže HIF1A. Sistemska biologija se je izkazala kot pomemben pristop določanja bioloških poti, v katere so uvrščene znane tarče transkripcijskega dejavnika HIF1A. Med potencialne nove tarče HIF1A spadajo tudi geni, ki imajo v svojih promotorjih že identificirana vezavna mesta za HIF1A (Benita in sod., 2009), vendar so za to potrebne dodatne raziskave. Pristop sistemske biologije je pomemben tudi pri identifikaciji nove

generacije biooznačevalcev hipoksije in z njo povezanimi boleznimi kot je rak pri človeku.

Metabolizem celice se ob nezadostni ravni $O_2$ preklopi iz mitohondrijsko oksidativne fosforilacije na proces glikolize (Zhang in sod., 2007). Povišana raven glikolize in znižano dihanje je bilo tudi v aerobnih pogojih opaženo pri številnih celicah raka in domneva se, da je ravno HIF1 vpleten v indukcijo genov, nujnih za prilagoditev celice (Zhang in sod., 2007). Razmnoževanje celic in preživetje je odvisno tudi od dostopnosti hranil, obstaja pa celo povezava med hranili in signalizacijo hormonskih poti (Marshall, 2006). Signalne poti hranil so povezane tudi s potjo inzulina in s sproščanjem metabolnih hormonov iz maščobnega tkiva (Marshall, 2006). Hipoksija dejansko pospešuje shranjevanje lipidov in inhibira razgradnjo le-teh z β-oksidacijo (Bostrom in sod., 2006).

V preteklih letih je postalo vse bolj jasno, da h kompleksnim genetskim procesom evkariotov poleg miRNA prispevajo tudi daljše (več kot 200 bp) nekodirajoče molekule RNA (nkRNA; angl. *noncoding RNA, ncRNA*) (pregled Mattick, 2009). Medtem ko je bil pri raziskovanju vloge in mehanizma uravnavanja izražanja z miRNA narejen velik napredek, je poznavanje vloge in mehanizma vpliva drugih nkRNA še precej omejeno. Do sedaj še ni poznana natančna vloga drugih nkRNA v procesih celice, potrebno pa je tudi dognati, kateri okoljski dejavniki imajo vpliv na njihovo izražanje.

Zanimiva skupina daljših nkRNA molekul so regije, ki so 100% identične med ortolognimi regijami genoma človeka, podgane in miši (Bejerano in sod., 2004a; Bejerano in sod., 2004b, Calin in sod., 2007). Gre za 481 zelo ohranjenih elementov genoma variabilnih dolžin (od 200 do 779 nukleotidov), po katerih se imenujejo prepisi zelo ohranjenih regij (T-UCR; angl. *transcribed-ultraconserved regions*) ali zelo ohranjeni geni (UCGs; angl. *ultraconserved genes*) (Calin in sod., 2007). Njihova presenetljivo visoka raven ohranjenosti nakazuje na pomembno biološko vlogo pri različnih odzivih celice. Spremenjeno izražanje T-UCR je bilo do sedaj že povezano s specifičnimi fenotipi tumorjev predvsem vpletenostjo v razvoj raka. Vloga spremenjenega izražanja T-UCR je bila do sedaj opisana pri kronični limfocitni levkemiji, raku debelega črevesja in danke (CRC; angl. *colorectal cancer*) (Calin in sod., 2007), pri hepatocelularnem karcinomu (raku jeter) (Braconi in sod., 2011; Calin in sod., 2007) in neuroblastomu (Mestdagh in sod., 2010; Scaruffi in sod., 2010). V genomu človeka se lokacije T-UCR pogosto prekrivajo z eksoni genov, vpletenih v izrezovanje intronov RNA, ali pa so prisotne v intronih gostiteljskih genov oziroma v bližini genov, ki so vpleteni v uravnavanje prepisovanja genov in razvoj organizma (Bejerano in sod., 2006; Bejerano in sod., 2004b). Kljub očitni tkivno-specifični vlogi T-UCR pri uravnavanju izražanja genov, potrebnih za razvoj organizma (Bejerano in sod., 2004b; Mattick, 2009; Pennacchio in sod., 2006), je njihov mehanizem, ki privede

do spremenjenega izražanja T-UCR, še nerazjasnjen. Da bi vsaj deloma razjasnili ta mehanizem, smo predpostavili, da ima zmanjšana vsebnost $O_2$ v rakavi celici vpliv na specifično izražanje T-UCR (Ferdin in sod., 2013). Preučevanja tega mehanizma smo se lotili na način, ki je bil predhodno že opisan za preučevanje miRNA v hipoksiji (Kulshreshtha in sod., 2007).

V naši raziskavi smo uporabili različne humane celične linije raka, ki smo jih izpostavili hipoksiji (1% vsebnost $O_2$) ali jih tretirali z DMOG (angl. *N-(Methoxyoxoacetyl)-glycine methyl ester*), posnemovalcem hipoksije. Te celične linije so bile: HT-29, HCT-116, rak debelega črevesja; MCF-7, rak dojk; U-87, glioblastom; in UC-2, rak mehurja. Pri preučevanju vloge HIF1A smo uporabili celično linijo karcinoma ledvic (786-O), kjer smo izražanje preiskovanih T-UCR preverili v dveh genetsko različnih linijah; z zapisom za VHL (VHL +) in brez zapisa za VHL (VHL −). Celice linij so se med seboj razlikovale v aktivnosti tumor-supresorja VHL, ki v normoksiji sodeluje pri razgradnji HIF1A (Kibel in sod., 1995). Izhodišče našega raziskovalnega dela so predstavljali predhodno objavljeni rezultati primerjave izražanja UCR raka debelega črevesja in izražanja UCR zdrave črevesne sluznice z uporabo DNA-mikromrež (Calin in sod., 2007). Z metodo RT-qPCR smo dokazali spremenjeno izražanje T-UCR v vzorcih rakavih celičnih linij, ki so bile izpostavljene hipoksičnemu ter normoksičnemu okolju. Metodo prenosa western smo uporabili za dokaz prisotnosti proteinov hipoksije (HIF1A in CA9) in za dokaz ravni izražanja gostiteljskega gena T-UCR. Naš namen je bil primarno določiti prisotnost transkripcijskega dejavnika HIF1A, vendar smo zaradi njegove izredne občutljivosti in razgradnje v prisotnosti $O_2$ za alternativo izbrali še CA9, za katerega je znano, da je neposredno induciran s HIF1A in je za razliko od njega stabilen v normoksiji.

Napoved vezavnih mest za HIF1A znotraj promotorskih regij 5 kb navzgor in 10 kb navzdol od gena z bioinformacijskim orodjem TRANSFAC Match (http://www.gene-regulation.com/pub/programs; Matys in sod., 2003) nam je omogočila, da smo se s funkcionalnimi analizami lahko osredotočili le na izbrane T-UCR-je, ki bi potencialno lahko bili uravnavani s transkripcijskim dejavnikom HIF1A. Le-te T-UCR-je smo kasneje usmerjeno utišali z vnosom tarčno specifične male interferenčne RNA (angl. *small interfering RNA*; siRNA) in opazovali njen vpliv na potek prepisovanja (transkripcije) po tretiranju s posnemovalcem hipoksije, DMOG, ter po transfekciji plazmidnega konstrukta s HIF1A. S kromatinsko imunoprecipitacijo (ChIP, angl. *chromatin immunoprecipitation*) smo ugotavljali neposredne interakcije med proteini in DNA v celici. V našem primeru smo metodo ChIP uporabili za dokaz HIF-vezavnih mest v bližini promotorskih regij gostiteljskih genov (angl. *host genes*), kjer so locirani T-UCR-ji. S tem smo lahko ločili tiste T-UCR-je, za katere je HIF1A kot transkripcijski dejavnik pomemben regulator. Po določitvi neposredne interakcije med HIF1A in promotorjem gostiteljskega gena smo v ekspresijski vektor navzgor od reporterskega

luciferaznega gena vstavili promotorski odsek z in brez vezavnih mest, ki so bila predvidena za interakcijo s HIF1A. Potrjevanje vpliva HIF1A in HIF2A konstruktov, ki omogočajo njuno nemoteno izražanje v normoksičnem okolju, in njun vpliv na klonirano promotorsko regijo v ekspresijskem vektorju, smo določevali z uporabo dvojnega luciferaznega testa (angl. *Dual-Luciferase assay*). S transfekcijo celic s specifičnimi siRNA smo ugotavljali vpliv RNA-interference na proliferacijo celičnih linij raka z neposrednim usmerjenim delovanjem siRNA na T-UCR-je v normoksičnih ter hipoksičnih pogojih. Sprva smo določili učinkovitost delovanja siRNA z določanjem kopij mRNA z RT-qPCR, nato smo v funkcionalne analize vključili le siRNA, katerih efektivnost presega 60% utišanje tarčnega gena. Te tarčno specifične siRNA smo vnesli v izbrane celične linije raka, preverili utišanje izražanja tarčnega gena na mRNA in proteinski ravni, vpliv utišanja gena na proliferacijo celic in njihov celični cikel. Analizo celičnega cikla smo izvedli po metodi ločevanje celic na osnovi fluorescence (FACS, angl. *fluorescence activated cell sorting*) in tako določili, v kateri fazi celičnega cikla je utišanje tarčnega gena s siRNA imelo vpliv na rast celic. Območje preučevanega T-UCR-ja smo nato vstavili v ekspresijski vektor navzdol od reporterskega luciferaznega gena, da bi preučili, ali ima klonirano območje vlogo ojačevalca prepisovanja. Potrjevanje vpliva T-UCR regije kot potencialnega ojačevalca smo prav tako določevali z uporabo dvojnega luciferaznega testa.

Namen naše študije je bil preiskati, ali hipoksija vpliva na povišano izražanje zelo ohranjenih ne-kodirajočih prepisov na podoben način kot pri protein-kodirajočih genih, ki so uravnavani s hipoksijo, oziroma že predhodno preučenvanih miRNA. V ta namen smo uporabili strategijo, ki vključuje začeten pregled izražanja genov iz podatkov kliničnih vzorcev tumorjev, pridobljenih z DNA-mikromrežami (Kulshreshtha in sod., 2007). V številnih tipih celic je bilo s hipoksijo povišano izražanje več zelo ohranjenih ne-kodirajočih prepisov (T-UCR); uc.63, uc.73, uc.106, uc.134, uc.475, ki smo jih imenovali HINCUTs (angl. *hypoxia induced noncoding ultraconserved transcripts*). S tem smo tudi potrdili našo hipotezo, da se v hipoksiji spremeni tudi izražanje nekaterih prepisov T-UTR (Ferdin in sod., 2013). Podobno kot je bilo že objavljeno za miRNA (Kulshreshtha in sod., 2007) smo dokazali, da ima HIF tudi ključno vlogo pri uravnavanju T-UCR v hipoksiji; i) z indukcijo zaradi kemikalij, ki preprečijo delovanje HIF prolil hidroksilaz, ii) zaradi inaktivacije delovanja tumor-supresorja VHL in iii) z učinkom delovanja vnesenih stabilnih HIF na izražanje ekspresijskih vektorjev z vključkom promotorja preučevanega HINCUT-a.

Naši rezultati so predvideli, da je preučevan HINCUT uc.475 del primarnega prepisa (po možnosti kot neizrezan intron; angl. *retained intron*), ki vključuje tudi kodirajoča zaporedja njegovega gostiteljskega gena *OGT*. Vloga OGT je kataliza O-GlcNAc (sladkorjev) na proteine, ki so vpleteni v prepisovanje, celični cikel, celične procese in signalne poti (pregled v Butkinaree in sod., 2010; Love in sod., 2010). Nedavno je bilo

tudi objavljeno, da je izražanje OGT povišano pri raku dojk ter pri primarnemu in metastaziranem raku prostate (Caldwell in sod., 2010; Lynch in sod., 2012), kar podpira idejo uporabe OGT kot tarče za zdravljenje tumorjev. V raziskavi dokazan prepis hibrida kodirajočih (*OGT*) in ne-kodirajočih (uc.475) zaporedij dolžine 9 kb je bil prisoten predvsem v jedru celice in iz katerega verjetno izvira velik ali celo precejšen delež zrele mRNA *OGT* delujočega/-ih proteina/-ov, odgovornih za velik del O-GlcNAcilacije v celici. Poleg tega ta prepis vsebuje pomemben ali celo večinski delež prepisov z uc.475 RNA v celici, saj se z utišanjem s siRNA-OGT močno zniža tudi količina prepisov z uc.475. Obratno tudi siRNA specifična za uc.475 dosledno in učinkovito inaktivira prepise uc.475 ter *OGT*. Iz prenosa northern je razvidno, da lahko uc.475 obstaja tudi kot samostojen prepis, vendar je le-ta lahko tudi ostanek primarnega prepisa *OGT* in zato verjetno območja uc.475 ni mogoče utišati neodvisno od njegovega gostiteljskega gena *OGT*.

Naši rezultati bodo pripomogli k boljšemu razumevanju kompleksnosti genomskega lokusa gena *OGT*, saj le-ta v središču hibridnega prepisa v jedru vsebuje ne-kodirajočo zelo ohranjeno sekvenco (UCR) kot del neizrezanega introna. Dokazali smo, da se iz lokusa gena *OGT* prepiše več kodirajočih in ne-kodirajočih prepisov. Neizrezan intron, ki vsebuje regijo uc.475, ne vsebuje kodirajočih odprtih bralnih okvirjev (ORF), ampak vključuje drugi začetni kodon (ATG). Neizrezan intron z uc.475 je lahko tudi del dolgega 5'-neprevedenega območja (3,27 kb) enega izmed prepisov gena *OGT* (Lubas in sod., 1997).

Inhibicija uc.475 s siRNA je predvsem v hipoksiji povzročila zvišanje števila celic v G2/M delu celičnega cikla, kar pa je lahko povezano tudi s tem, da celica med hipoksičnim stresom ni sposobna vzdrževati ustrezne ravni O-GlcNAcilacije. Menimo, da je vsaj del biokemijskih efektov utišanja uc.475 posledica vmešavanja gena *OGT*. Na podlagi rezultatov prenosa northern lahko trdimo, da je količina uc.475 izven hibridnega prepisa majhna. Poleg tega smo opazili tudi razliko med s hipoksijo induciranim OGT/uc.475 na ravni RNA in proteina. Pravzaprav so meritve v številnih celičnih linijah pokazale, da ostane protein OGT nespremenjen ali je med zmanjšano vsebnostjo $O_2$ celo znižan. To je lahko tudi posledica znižanega prepisovanja genov, ki je opaženo tudi v hipoksiji. Višja raven prepisov uc.475 lahko potencialno služi za nadomestitev zmanjšane količine proteina OGT. Zato je razumljivo, da v odsotnosti indukcije mRNA OGT, lahko raven O-GlcNAcilacije pade do mere, ki bi potencialno že lahko ogrožala preživetje celic.

Kaj je torej vloga ne-kodirajoče regije z uc.475 v kontekstu regulacije OGT v hipoksiji? Luciferazni eksperimenti predlagajo, da bi lahko ta sekvenca imela vlogo ojačevalca gena in bi lahko skupaj s HIF-om sodelovala pri pospeševanju transkripcije. Pretežno jedrna lokacija hibridnega prepisa induciranega s hipoksijo, bi poleg pospeševanja

transkripcije gena *OGT* lahko imela tudi vlogo regije, ki vzdržuje stabilnosti mRNA gostiteljskega gena. Do danes so bili neizrezani introni v številnih genih že dokazani, da pospešijo skoraj vsak korak izražanja gena in poleg tega tudi vplivajo na stabilnost mRNA. Medtem ko so nekatere študije pokazale, da ima neizrezan intron zelo malo ali nič vpliva na razpolovno dobo mRNA (Chang in sod., 2006; Nott in sod., 2003; Ryu in Mertz, 1989), so druge potrdile pomembnost neizrezanega introna za stabilnost mRNA in pri podaljševanju njene razpolovne dobe (Gencheva in sod., 2010; Wang in sod., 2007; Zhao in Hamilton, 2007). V hipoksičnih celicah je bilo tudi opaženo, da neizrezani introni zavirajo mehanizme razpada RNA (NMD; angl. *nonsense-mediated RNA decay)*, da bi se s tem omogočilo izražanje genov, ki so kritični za hitro prilagoditev celice na stres (Gardner, 2008). Poleg tega številne študije poročajo, da so različno izrezani introni ključni za optimalno izražanje genov (Buchman in Berg, 1988; Le Hir in sod., 2003; Nott in sod., 2003). Ena možnost, kako introni lahko vplivajo na prepisovanje, je, da imajo vlogo skladišč za regulatorne elemente prepisovanja, kot so na primer ojačevalci in represorji (Le Hir in sod., 2003). Ne samo, da so potrebne nadaljnje raziskave mehanizma indukcije gena *OGT*, ampak je nujno tudi podrobnejše razumevanje vloge številnih prepisov tega gena. To je sicer tudi prva študija, ki je usmerjeno preučevala vlogo T-UCR-jev kot del odziva na hipoksijo in tako deloma prispeva k boljšemu razumevanju medsebojne interakcije med kodirajočimi in nekodirajočimi prepisi v stresnih okoliščinah. S tem pa smo dokazali vsaj eno, če ne celo dve, novi tarči gena *HIF*, ki jo/ju lahko vključimo v njegovo regulatorno mrežo.

Izdelani Atlas regulatorne mreže gena *HIF1A* predstavlja centralno mesto, ki bo omogočilo pospešitev raziskav za preučevani TF. Če povzamemo, so bili analizirani SNP-ji gena *HIF1A* do sedaj povezani že s številnimi fenotipi in boleznimi pri človeku. V naši študiji smo predvideli, da imajo *HIF1A* SNP-ji vpliv na pridobitev/izgubo vezavnih mest za TF in miRNA, kar bi lahko bila tudi ena izmed razlag funkcije polimorfizmov na izražanje *HIF1A*. Z metodo obogatitve bioloških poti, kamor so uvrščene tarče HIF1A, smo določili tiste, ki bi lahko bile ključne za prilagoditev celice in njeno preživetje v hipoksiji. Predstavljen način izdelave Atlasa regulatorne mreže gena *HIF1A* v tej študiji je možno uporabiti za izdelavo Atlasa regulatorne mreže TF-ja po izbiri. V tej raziskavi je bila prvič dokazana tudi povezava med znižano vsebnostjo $O_2$ in dolgimi prepisi zelo ohranjenih regij (T-UCR), ki smo jih poimenovali HINCUT. Te regije se izražajo pretežno v jedru celice in je HIF vsaj deloma odgovoren za indukcijo prepisovanja nekaterih izmed njih. Prepis uc.475, ki smo ga podrobneje preučevali, je del neizrezanega introna gostiteljskega gena *OGT*, ki je ključen za potranslacijsko modifikacijo proteinov celice. Naši rezultati tako predstavljajo prvi vpogled v nove funkcije omrežja hipoksije, ki ga sestavljenjo protein-kodirajoči in nekodirajoči prepisi RNA.

# 8 REFERENCES

Ahmetov I., Hakimullina A.M., Lyubaeva E.V., Vinogradova O.L., Rogozkin V.A. 2008. Effect of HIF1A gene polymorphism on human muscle performance. Bull. Exp. Biol. Med., 146: 351-353

Albertson D.G. 2006. Gene amplification in cancer. Trends. Genet., 22: 447-455

Alidoosti M., Ghaedi M., Soleimani A., Bakhtiyari S., Rezvanfard M., Golkhu S., Mohammadtaghvaei N. 2011. Study on the role of environmental parameters and HIF-1A gene polymorphism in coronary collateral formation among patients with ischemic heart disease. Clin. Biochem., 44: 1421-1424

Allen T.A., Von Kaenel S., Goodrich J.A., Kugel J.F. 2004. The SINE-encoded mouse B2 RNA represses mRNA transcription in response to heat shock. Nat. Struct. Mol. Biol., 11: 816-821

Amaral P.P., Clark M.B., Gascoigne D.K., Dinger M.E., Mattick J.S. 2011. lncRNAdb: a reference database for long noncoding RNAs. Nucleic Acids Res., 39: D146-151

Amaral P.P., Mattick J.S. 2008. Noncoding RNA in development. Mamm. Genome, 19: 454-492

Ambros V. 2003. MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. Cell, 113: 673-676

Ambros V. 2008. The evolution of our thinking about microRNAs. Nat. Med., 14: 1036-1040

Ameln H., Gustafsson T., Sundberg C.J., Okamoto K., Jansson E., Poellinger L., Makino Y. 2005. Physiological activation of hypoxia inducible factor-1 in human skeletal muscle. FASEB J., 19: 1009-1011

Apaydin I., Konac E., Onen H.I., Akbaba M., Tekin E., Ekmekci A. 2008. Single nucleotide polymorphisms in the hypoxia-inducible factor-1alpha (HIF-1alpha) gene in human sporadic breast cancer. Arch. Med. Res., 39: 338-345

Arlt M.F., Casper A.M., Glover T.W. 2003. Common fragile sites. Cytogenet. Genome Res., 100: 92-100

Arlt M.F., Durkin S.G., Ragland R.L., Glover T.W. 2006. Common fragile sites as targets for chromosome rearrangements. DNA Repair. (Amst.), 5: 1126-1135

Audenet F., Yates D.R., Cancel-Tassin G., Cussenot O., Roupret M. 2012. Genetic pathways involved in carcinogenesis of clear cell renal cell carcinoma: genomics towards personalized medicine. BJU Int., 109: 1864-1870

Ayoubi T.A., Van De Ven W.J. 1996. Regulation of gene expression by alternative promoters. FASEB J., 10: 453-460

Bahadori B., Uitz E., Mayer A., Harauer J., Dam K., Truschnig-Wilders M., Pilger E., Renner W. 2010. Polymorphisms of the hypoxia-inducible factor 1 gene and peripheral artery disease. Vasc. Med., 15: 371-374

Bain G., Maandag E.C., Izon D.J., Amsen D., Kruisbeek A.M., Weintraub B.C., Krop I., Schlissel M.S., Feeney A.J., van Roon M., et al. 1994. E2A proteins are required for proper B cell development and initiation of immunoglobulin gene rearrangements. Cell, 79: 885-892

Ban M., Maranian M., Yeo T.W., Gray J., Compston A., Sawcer S. 2005. Ultraconserved regions in multiple sclerosis. Eur. J. Hum. Genet., 13: 998-999

Barbarotto E., Schmittgen T.D., Calin G.A. 2008. MicroRNAs and cancer: profile, profile, profile. Int. J. Cancer, 122: 969-977

Barcellos-Hoff M.H., Lyden D., Wang T.C. 2013. The evolution of the cancer niche during multistage carcinogenesis. Nat. Rev. Cancer, 13: 511-518

Beasley N.J., Leek R., Alam M., Turley H., Cox G.J., Gatter K., Millard P., Fuggle S., Harris A.L. 2002. Hypoxia-inducible factors HIF-1alpha and HIF-2alpha in head and neck cancer: relationship to tumor biology and treatment outcome in surgically resected patients. Cancer Res., 62: 2493-2497

Beckman J.A., Creager M.A., Libby P. 2002. Diabetes and atherosclerosis: epidemiology, pathophysiology, and management. JAMA, 287: 2570-2581

Bejerano G., Haussler D., Blanchette M. 2004a. Into the heart of darkness: large-scale clustering of human non-coding DNA. Bioinformatics, 20, Suppl 1: i40-48

Bejerano G., Lowe C.B., Ahituv N., King B., Siepel A., Salama S.R., Rubin E.M., Kent W.J., Haussler D. 2006. A distal enhancer and an ultraconserved exon are derived from a novel retroposon. Nature, 441: 87-90

Bejerano G., Pheasant M., Makunin I., Stephen S., Kent W.J., Mattick J.S., Haussler D. 2004b. Ultraconserved elements in the human genome. Science, 304: 1321-1325

Belaiba R.S., Bonello S., Zahringer C., Schmidt S., Hess J., Kietzmann T., Gorlach A. 2007. Hypoxia up-regulates hypoxia-inducible factor-1alpha transcription by involving phosphatidylinositol 3-kinase and nuclear factor kappaB in pulmonary artery smooth muscle cells. Mol. Biol. Cell, 18: 4691-4697

Benita Y., Kikuchi H., Smith A.D., Zhang M.Q., Chung D.C., Xavier R.J. 2009. An integrative genomics approach identifies Hypoxia Inducible Factor-1 (HIF-1)-target genes that form the core response to hypoxia. Nucleic Acids. Res., 37: 4587-4602

Bertout J.A., Patel S.A., Simon M.C. 2008. The impact of O2 availability on human cancer. Nat. Rev. Cancer, 8: 967-975

Birner P., Schindl M., Obermair A., Plank C., Breitenecker G., Oberhuber G. 2000. Overexpression of hypoxia-inducible factor 1alpha is a marker for an unfavorable prognosis in early-stage invasive cervical cancer. Cancer Res., 60: 4693-4696

Bostrom P., Magnusson B., Svensson P.A., Wiklund O., Boren J., Carlsson L.M., Stahlman M., Olofsson S.O., Hulten L.M. 2006. Hypoxia converts human macrophages into triglyceride-loaded foam cells. Arterioscler. Thromb. Vasc. Biol., 26: 1871-1876

Braconi C., Valeri N., Kogure T., Gasparini P., Huang N., Nuovo G.J., Terracciano L., Croce C.M., Patel T. 2011. Expression and functional role of a transcribed noncoding RNA with an ultraconserved element in hepatocellular carcinoma. Proc. Natl. Acad. Sci. U S A, 108: 786-791

Bradney C., Hjelmeland M., Komatsu Y., Yoshida M., Yao T.P., Zhuang Y. 2003. Regulation of E2A activities by histone acetyltransferases in B lymphocyte development. J. Biol. Chem., 278: 2370-2376

Brosnan C.A., Voinnet O. 2009. The long and the short of noncoding RNAs. Curr. Opin. Cell Biol., 21: 416-425

Bruick R.K., McKnight S.L. 2001. A conserved family of prolyl-4-hydroxylases that modify HIF. Science, 294: 1337-1340

Bruning U., Cerone L., Neufeld Z., Fitzpatrick S.F., Cheong A., Scholz C.C., Simpson D.A., Leonard M.O., Tambuwala M.M., Cummins E.P., Taylor C.T. 2011. MicroRNA-155 promotes resolution of hypoxia-inducible factor 1alpha activity during prolonged hypoxia. Mol. Cell. Biol., 31: 4087-4096

Buchman A.R., Berg P. 1988. Comparison of intron-dependent and intron-independent gene expression. Mol. Cell Biol., 8: 4395-4405

Butkinaree C., Park K., Hart G.W. 2010. 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, 1800: 96-106

Caldwell S.A., Jackson S.R., Shahriari K.S., Lynch T.P., Sethi G., Walker S., Vosseller K., Reginato M.J. 2010. Nutrient sensor O-GlcNAc transferase regulates breast cancer tumorigenesis through targeting of the oncogenic transcription factor FoxM1. Oncogene, 29: 2831-2842

Calin G.A., Croce C.M. 2006. MicroRNA signatures in human cancers. Nat. Rev. Cancer, 6: 857-866

Calin G.A., Dumitru C.D., Shimizu M., Bichi R., Zupo S., Noch E., Aldler H., Rattan S., Keating M., Rai K., Rassenti L., Kipps T., Negrini M., Bullrich F., Croce C.M. 2002. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc. Natl. Acad. Sci. U S A, 99: 15524-15529

Calin G.A., Ferracin M., Cimmino A., Di Leva G., Shimizu M., Wojcik S.E., Iorio M.V., Visone R., Sever N.I., Fabbri M., Iuliano R., Palumbo T., Pichiorri F., Roldo C., Garzon R., Sevignani C., Rassenti L., Alder H., Volinia S., Liu C.G., Kipps T.J., Negrini M., Croce C.M. 2005. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N. Engl. J. Med., 353: 1793-1801

Calin G.A., Liu C.G., Ferracin M., Hyslop T., Spizzo R., Sevignani C., Fabbri M., Cimmino A., Lee E.J., Wojcik S.E., Shimizu M., Tili E., Rossi S., Taccioli C., Pichiorri F., Liu X., Zupo S., Herlea V., Gramantieri L., Lanza G., Alder H., Rassenti L., Volinia S., Schmittgen T.D., Kipps T.J., Negrini M., Croce C.M. 2007. Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. Cancer Cell, 12: 215-229

Calin G.A., Sevignani C., Dumitru C.D., Hyslop T., Noch E., Yendamuri S., Shimizu M., Rattan S., Bullrich F., Negrini M., Croce C.M. 2004. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc. Natl. Acad. Sci. U S A, 101: 2999-3004

Cannistraci C.V., Ogorevc J., Zorc M., Ravasi T., Dovc P., Kunej T. 2013. Pivotal role of the muscle-contraction pathway in cryptorchidism and evidence for genomic connections with cardiomyopathy pathways in RASopathies. BMC Med. Genomics, 6: 5

Cao D., Hou M., Guan Y.S., Jiang M., Yang Y., Gou H.F. 2009. Expression of HIF-1alpha and VEGF in colorectal cancer: association with clinical outcomes and prognostic implications. BMC Cancer, 9: 432

Cascio S., D'Andrea A., Ferla R., Surmacz E., Gulotta E., Amodeo V., Bazan V., Gebbia N., Russo A. 2010. miR-20b modulates VEGF expression by targeting HIF-1 alpha and STAT3 in MCF-7 breast cancer cells. J. Cell Physiol., 224: 242-249

Castel S.E., Martienssen R.A. 2013. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nat. Rev. Genet., 14: 100-112

Cha S.T., Chen P.S., Johansson G., Chu C.Y., Wang M.Y., Jeng Y.M., Yu S.L., Chen J.S., Chang K.J., Jee S.H., Tan C.T., Lin M.T., Kuo M.L. 2010. MicroRNA-519c suppresses hypoxia-inducible factor-1alpha expression and tumor angiogenesis. Cancer Res., 70: 2675-2685

Chan D.A., Sutphin P.D., Yen S.E., Giaccia A.J. 2005. Coordinate regulation of the oxygen-dependent degradation domains of hypoxia-inducible factor 1 alpha. Mol. Cell Biol., 25: 6415-6426

Chan S.Y., Zhang Y.Y., Hemann C., Mahoney C.E., Zweier J.L., Loscalzo J. 2009. MicroRNA-210 controls mitochondrial metabolism during hypoxia by repressing the iron-sulfur cluster assembly proteins ISCU1/2. Cell Metab., 10: 273-284

Chang C.Y., Hong W.W., Chong P., Wu S.C. 2006. 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, 24: 1132-1141

Chau C.H., Permenter M.G., Steinberg S.M., Retter A.S., Dahut W.L., Price D.K., Figg W.D. 2005. Polymorphism in the hypoxia-inducible factor 1alpha gene may confer susceptibility to androgen-independent prostate cancer. Cancer Biol. Ther., 4: 1222-1225

Chen J., Lozach J., Garcia E.W., Barnes B., Luo S., Mikoulitch I., Zhou L., Schroth G., Fan J.B. 2008. Highly sensitive and specific microRNA expression profiling using BeadArray technology. Nucleic Acids Res., 36: e87

Chen M.K., Chiou H.L., Su S.C., Chung T.T., Tseng H.C., Tsai H.T., Yang S.F. 2009. The association between hypoxia inducible factor-1alpha gene polymorphisms and increased susceptibility to oral cancer. Oral Oncol., 45: e222-226

Chen Z., Li Y., Zhang H., Huang P., Luthra R. 2010. Hypoxia-regulated microRNA-210 modulates mitochondrial function and decreases ISCU and COX10 expression. Oncogene, 29: 4362-4368

Chiaro C., Lazarova D.L., Bordonaro M. 2012. Tcf3 and cell cycle factors contribute to butyrate resistance in colorectal cancer cells. Biochem. Biophys. Res. Commun., 428: 121-126

Chin L.J., Ratner E., Leng S., Zhai R., Nallur S., Babar I., Muller R.U., Straka E., Su L., Burki E.A., Crowell R.E., Patel R., Kulkarni T., Homer R., Zelterman D., Kidd K.K., Zhu Y., Christiani D.C., Belinsky S.A., Slack F.J., Weidhaas J.B. 2008. A SNP in a let-7 microRNA complementary site in the KRAS 3' untranslated region increases non-small cell lung cancer risk. Cancer Res., 68: 8535-8540

Cieszczyk P., Eider J., Arczewska A., Ostanek M., Leonska-Duniec A., Sawczyn S., Ficek K., Jascaniene N., Kotarska K., Sygit K. 2011. The Hif1a Gene Pro582ser Polymorphism in Polish Power-Orientated Athletes. Biology of Sport, 28: 111-114

Cimmino A., Calin G.A., Fabbri M., Iorio M.V., Ferracin M., Shimizu M., Wojcik S.E., Aqeilan R.I., Zupo S., Dono M., Rassenti L., Alder H., Volinia S., Liu C.G., Kipps T.J., Negrini M., Croce C.M. 2005. miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc. Natl. Acad. Sci. U S A, 102: 13944-13949

Clifford S.C., Astuti D., Hooper L., Maxwell P.H., Ratcliffe P.J., Maher E.R. 2001. The pVHL-associated SCF ubiquitin ligase complex: molecular genetic analysis of elongin B and C, Rbx1 and HIF-1alpha in renal cell carcinoma. Oncogene, 20: 5067-5074

Cortez M.A., Bueso-Ramos C., Ferdin J., Lopez-Berestein G., Sood A.K., Calin G.A. 2011. MicroRNAs in body fluids--the mix of hormones and biomarkers. Nat. Rev. Clin. Oncol., 8: 467-477

Costa F.F. 2007. Non-coding RNAs: lost in translation? Gene, 386: 1-10

Coulet F., Nadaud S., Agrapart M., Soubrier F. 2003. Identification of hypoxia-response element in the human endothelial nitric-oxide synthase gene promoter. J. Biol. Chem., 278: 46230-46240

Creighton C.J., Reid J.G., Gunaratne P.H. 2009. Expression profiling of microRNAs by deep sequencing. Brief Bioinform., 10: 490-497

Crosby M.E., Devlin C.M., Glazer P.M., Calin G.A., Ivan M. 2009a. Emerging roles of microRNAs in the molecular responses to hypoxia. Curr. Pharm. Des., 15: 3861-3866

Crosby M.E., Kulshreshtha R., Ivan M., Glazer P.M. 2009b. MicroRNA regulation of DNA repair gene expression in hypoxic stress. Cancer Res., 69: 1221-1229

Cummins J.M., He Y., Leary R.J., Pagliarini R., Diaz L.A., Jr., Sjoblom T., Barad O., Bentwich Z., Szafranska A.E., Labourier E., Raymond C.K., Roberts B.S., Juhl H., Kinzler K.W., Vogelstein B., Velculescu V.E. 2006. The colorectal microRNAome. Proc. Natl. Acad. Sci. U S A, 103: 3687-3692

Dang C.V., Semenza G.L. 1999. Oncogenic alterations of metabolism. Trends Biochem. Sci., 24: 68-72

Denko N.C. 2008. Hypoxia, HIF1 and glucose metabolism in the solid tumour. Nat. Rev. Cancer, 8: 705-713

Devlin C., Greco S., Martelli F., Ivan M. 2011. miR-210: More than a silent player in hypoxia. IUBMB Life, 63: 94-100

DiCarlo V.S., Chen S.J., Meng Q.C., Durand J., Yano M., Chen Y.F., Oparil S. 1995. ETA-receptor antagonist prevents and reverses chronic hypoxia-induced pulmonary hypertension in rat. Am. J. Physiol., 269: L690-697

Dong D.L., Hart G.W. 1994. Purification and characterization of an O-GlcNAc selective N-acetyl-beta-D-glucosaminidase from rat spleen cytosol. J. Biol. Chem., 269: 19321-19330

Doring F., Onur S., Fischer A., Boulay M.R., Perusse L., Rankinen T., Rauramaa R., Wolfarth B., Bouchard C. 2010. A common haplotype and the Pro582Ser polymorphism of the hypoxia-inducible factor-1alpha (HIF1A) gene in elite endurance athletes. J. Appl. Physiol., 108: 1497-1500

Droma Y., Ota M., Hanaoka M., Katsuyama Y., Basnyat B., Neupane P., Arjyal A., Pandit A., Sharma D., Ito M., Kubo K. 2008. Two Hypoxia Sensor Genes and Their Association with Symptoms of Acute Mountain Sickness in Sherpas. Aviation Space and Environmental Medicine, 79: 1056-1060

Du T., Zamore P.D. 2005. microPrimer: the biogenesis and function of microRNA. Development, 132: 4645-4652

Duret L., Dorkeld F., Gautier C. 1993. Strong conservation of non-coding sequences during vertebrates evolution: potential involvement in transcriptional regulation of gene expression. Nucleic Acids Res., 21: 2315-2322

Eddy S.R. 2001. Non-coding RNA genes and the modern RNA world. Nat. Rev. Genet., 2: 919-929

Elgar G., Vavouri T. 2008. Tuning in to the signals: noncoding sequence conservation in vertebrate genomes. Trends Genet., 24: 344-352

Ema M., Hirota K., Mimura J., Abe H., Yodoi J., Sogawa K., Poellinger L., Fujii-Kuriyama Y. 1999. Molecular mechanisms of transcription activation by HLF and HIF1alpha in response to hypoxia: their stabilization and redox signal-induced interaction with CBP/p300. EMBO J., 18: 1905-1914

Ema M., Taya S., Yokotani N., Sogawa K., Matsuda Y., Fujii-Kuriyama Y. 1997. A novel bHLH-PAS factor with close sequence similarity to hypoxia-inducible factor 1alpha regulates the VEGF expression and is potentially involved in lung and vascular development. Proc. Natl. Acad. Sci. U S A, 94: 4273-4278

Emanuele E., Bertona M., Geroldi D. 2010. A multilocus candidate approach identifies ACE and HIF1A as susceptibility genes for cellulite. J. Eur. Acad. Dermatol. Venereol., 24: 930-935

Espinoza C.A., Allen T.A., Hieb A.R., Kugel J.F., Goodrich J.A. 2004. B2 RNA binds directly to RNA polymerase II to repress transcript synthesis. Nat. Struct. Mol.. Biol., 11: 822-829

Esquela-Kerscher A., Slack F.J. 2006. Oncomirs - microRNAs with a role in cancer. Nat. Rev. Cancer, 6: 259-269

Eynon N., Alves A.J., Meckel Y., Yamin C., Ayalon M., Sagiv M., Sagiv M. 2010. Is the interaction between HIF1A P582S and ACTN3 R577X determinant for power/sprint performance? Metabolism, 59: 861-865

Favaro E., Ramachandran A., McCormick R., Gee H., Blancher C., Crosby M., Devlin C., Blick C., Buffa F., Li J.L., Vojnovic B., Pires das Neves R., Glazer P., Iborra F., Ivan M., Ragoussis J., Harris A.L. 2010. MicroRNA-210 regulates mitochondrial free radical response to hypoxia and krebs cycle in cancer cells by targeting iron sulfur cluster protein ISCU. PLoS One, 5: e10345

Feng J., Bi C., Clark B.S., Mady R., Shah P., Kohtz J.D. 2006. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. Genes Dev., 20: 1470-1484

Ferdin J., Kunej T., Calin G.A. 2010. Non-coding RNAs: identification of cancer-associated microRNAs by gene profiling. Technol. Cancer. Res. Treat., 9: 123-138

Ferdin J., Kunej T., Calin G.A. 2011. MicroRNA: Genomic Association with Cancer Predisposition. J. Assoc. Genet. Technol., 37: 11-19

Ferdin J., Nishida N., Wu X., Nicoloso M.S., Shah M.Y., Devlin C., Ling H., Shimizu M., Kumar K., Cortez M.A., Ferracin M., Bi Y., Yang D., Czerniak B., Zhang W., Schmittgen T.D., Voorhoeve M.P., Reginato M.J., Negrini M., Davuluri R.V., Kunej T., Ivan M., Calin G.A. 2013. HINCUTs in cancer: hypoxia-induced noncoding ultraconserved transcripts. Cell Death Differ. doi: 10.1038/cdd.2013.119 (in press)

Ferracin M., Veronese A., Negrini M. 2010. Micromarkers: miRNAs in cancer diagnosis and prognosis. Expert Rev. Mol. Diagn., 10: 297-308

Filipowicz W., Bhattacharyya S.N., Sonenberg N. 2008. Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat. Rev. Genet., 9: 102-114

Flamme I., Frohlich T., von Reutern M., Kappel A., Damert A., Risau W. 1997. HRF, a putative basic helix-loop-helix-PAS-domain transcription factor is closely related to hypoxia-inducible factor-1 alpha and developmentally expressed in blood vessels. Mech. Dev., 63: 51-60

Foley R., Marignol L., Thomas A.Z., Cullen I.M., Perry A.S., Tewari P., O'Grady A., Kay E., Dunne B., Loftus B., Watson W.R., Fitzpatrick J.M., Woodson K., Lehman T., Hollywood D., Lynch T.H., Lawler M. 2009. The HIF-1alpha C1772T polymorphism may be associated with susceptibility to clinically localised prostate cancer but not with elevated expression of hypoxic biomarkers. Cancer Biol. Ther., 8: 118-124

Folkman J., Merler E., Abernathy C., Williams G. 1971. Isolation of a tumor factor responsible for angiogenesis. J. Exp. Med., 133: 275-288

Frank B., Hoffmeister M., Klopp N., Illig T., Chang-Claude J., Brenner H. 2010. Single nucleotide polymorphisms in Wnt signaling and cell death pathway genes and susceptibility to colorectal cancer. Carcinogenesis, 31: 1381-1386

Fransen K., Fenech M., Fredrikson M., Dabrosin C., Soderkvist P. 2006. Association between ulcerative growth and hypoxia inducible factor-1alpha polymorphisms in colorectal cancer patients. Mol. Carcinog., 45: 833-840

Frazer K.A., Ballinger D.G., Cox D.R., Hinds D.A., Stuve L.L., Gibbs R.A., Belmont J.W., Boudreau A., Hardenbol P., Leal S.M., Pasternak S., Wheeler D.A., Willis T.D., Yu F., Yang H., Zeng C., Gao Y., Hu H., Hu W., Li C., Lin W., Liu S., Pan H., Tang X., Wang J., Wang W., Yu J., Zhang B., Zhang Q., Zhao H., Zhou J., Gabriel S.B., Barry R., Blumenstiel B., Camargo A., Defelice M., Faggart M., Goyette M., Gupta S., Moore J., Nguyen H., Onofrio R.C., Parkin M., Roy J., Stahl E., Winchester E., Ziaugra L., Altshuler D., Shen Y., Yao Z., Huang W., Chu X., He Y., Jin L., Liu Y., Sun W., Wang H., Wang Y., Xiong X., Xu L., Waye M.M., Tsui S.K., Xue H., Wong J.T., Galver L.M., Fan J.B., Gunderson K., Murray S.S., Oliphant A.R., Chee M.S., Montpetit A., Chagnon F., Ferretti V., Leboeuf M., Olivier J.F., Phillips M.S., Roumy S., Sallee C., Verner A., Hudson T.J., Kwok P.Y., Cai D., Koboldt D.C., Miller R.D., Pawlikowska L., Taillon-Miller P., Xiao M., Tsui L.C., Mak W., Song Y.Q., Tam P.K., Nakamura Y., Kawaguchi T., Kitamoto T., Morizono T., Nagashima A., Ohnishi Y., Sekine A., Tanaka T., Tsunoda T., Deloukas P., Bird C.P., Delgado M., Dermitzakis E.T., Gwilliam R., Hunt S., Morrison J., Powell D., Stranger B.E., Whittaker P., Bentley D.R., Daly M.J., de Bakker P.I., Barrett J., Chretien Y.R., Maller J., McCarroll S., Patterson N., Pe'er I., Price A., Purcell S., Richter D.J., Sabeti P., Saxena R., Schaffner S.F., Sham P.C., Varilly P., Stein L.D., Krishnan L., Smith A.V., Tello-Ruiz M.K., Thorisson G.A., Chakravarti A., Chen P.E., Cutler D.J., Kashuk C.S., Lin S., Abecasis G.R., Guan W., Li Y., Munro H.M., Qin Z.S., Thomas D.J., McVean G., Auton A., Bottolo L., Cardin N., Eyheramendy S., Freeman C., Marchini J., Myers S., Spencer C., Stephens M., Donnelly P., Cardon L.R., Clarke G., Evans D.M., Morris A.P., Weir B.S., Mullikin J.C., Sherry S.T., Feolo M., Skol A., Zhang H., Matsuda I., Fukushima Y., Macer D.R., Suda E., Rotimi C.N., Adebamowo C.A., Ajayi I., Aniagwu T., Marshall P.A., Nkwodimmah C., Royal C.D., Leppert M.F., Dixon M., Peiffer A., Qiu R., Kent A., Kato K., Niikawa N., Adewole I.F., Knoppers B.M., Foster M.W., Clayton E.W., Watkin J., Muzny D., Nazareth L., Sodergren E., Weinstock G.M., Yakub I., Birren B.W., Wilson R.K., Fulton L.L., Rogers J., Burton J., Carter N.P., Clee C.M., Griffiths M., Jones M.C., McLay K., Plumb R.W., Ross M.T., Sims S.K., Willey D.L., Chen Z., Han H., Kang L., Godbout M.,

Wallenburg J.C., L'Archeveque P., Bellemare G., Saeki K., An D., Fu H., Li Q., Wang Z., Wang R., Holden A.L., Brooks L.D., McEwen J.E., Guyer M.S., Wang V.O., Peterson J.L., Shi M., Spiegel J., Sung L.M., Zacharia L.F., Collins F.S., Kennedy K., Jamieson R., Stewart J. 2007. A second generation human haplotype map of over 3.1 million SNPs. Nature, 449: 851-861

Fu X.S., Choi E., Bubley G.J., Balk S.P. 2005. Identification of hypoxia-inducible factor-1alpha (HIF-1alpha) polymorphism as a mutation in prostate cancer that prevents normoxia-induced degradation. Prostate, 63: 215-221

Fujimori K., Amano F. 2011. Forkhead transcription factor Foxa1 is a novel target gene of C/EBPbeta and suppresses the early phase of adipogenesis. Gene, 473: 150-156

Gabbasov R.T., Arkhipova A.A., Borisova A.V., Hakimullina A.M., Kuznetsova A.V., Williams A.G., Day S.H., Ahmetov, II. 2012. The HIF1A gene Pro582Ser polymorphism in Russian strength athletes. J. Strength. Cond. Res. http://journals.lww.com/nsca-jscr/pages/articleviewer.aspx?year=9000&issue=00000&article=97961&type=abstract

Gao Y., Wells L., Comer F.I., Parker G.J., Hart G.W. 2001. Dynamic O-glycosylation of nuclear and cytosolic proteins: cloning and characterization of a neutral, cytosolic beta-N-acetylglucosaminidase from human brain. J. Biol. Chem., 276: 9838-9845

Gardner L.B. 2008. Hypoxic inhibition of nonsense-mediated RNA decay regulates gene expression and the integrated stress response. Mol. Cell Biol., 28: 3729-3741

Gatenby R.A., Gillies R.J. 2004. Why do cancers have high aerobic glycolysis? Nat. Rev. Cancer, 4: 891-899

Gaudio E., Spizzo R., Paduano F., Luo Z., Efanov A., Palamarchuk A., Leber A.S., Kaou M., Zanesi N., Bottoni A., Costinean S., Rassenti L.Z., Nakamura T., Kipps T.J., Aqeilan R.I., Pekarsky Y., Trapasso F., Croce C.M. 2012. Tcl1 interacts with Atm and enhances NF-kappaB activation in hematologic malignancies. Blood, 119: 180-187

Gencheva M., Lin T.Y., Wu X., Yang L., Richard C., Jones M., Lin S.B., Lin R.J. 2010. Nuclear retention of unspliced pre-mRNAs by mutant DHX16/hPRP2, a spliceosomal DEAH-box protein. J. Biol. Chem., 285: 35624-35632

Ghosh G., Subramanian I.V., Adhikari N., Zhang X., Joshi H.P., Basi D., Chandrashekhar Y.S., Hall J.L., Roy S., Zeng Y., Ramakrishnan S. 2010. Hypoxia-induced microRNA-424 expression in human endothelial cells

regulates HIF-alpha isoforms and promotes angiogenesis. J. Clin. Invest., 120: 4141-4154

Giatromanolaki A., Koukourakis M.I., Sivridis E., Turley H., Talks K., Pezzella F., Gatter K.C., Harris A.L. 2001. Relation of hypoxia inducible factor 1 alpha and 2 alpha in operable non-small cell lung cancer to angiogenic/molecular profile of tumours and survival. Br. J. Cancer, 85: 881-890

Glover T.W. 2006. Common fragile sites. Cancer Lett., 232: 4-12

Gordan J.D., Bertout J.A., Hu C.J., Diehl J.A., Simon M.C. 2007. HIF-2alpha promotes hypoxic cell proliferation by enhancing c-myc transcriptional activity. Cancer Cell, 11: 335-347

Gryfe R., Swallow C., Bapat B., Redston M., Gallinger S., Couture J. 1997. Molecular biology of colorectal cancer. Curr. Probl. Cancer, 21: 233-300

Gu H.F., Zheng X., Abu Seman N., Gu T., Botusan I.R., Sunkari V.G., Lokman E.F., Brismar K., Catrina S.B. 2012. Impact of the Hypoxia-Inducible Factor-1 alpha (HIF-1alpha) Pro582Ser Polymorphism on Diabetes Nephropathy. Diabetes Care, 36:415-21

Gu Y.Z., Hogenesch J.B., Bradfield C.A. 2000. The PAS superfamily: sensors of environmental and developmental signals. Annu. Rev. Pharmacol. Toxicol., 40: 519-561

Guttman M., Amit I., Garber M., French C., Lin M.F., Feldser D., Huarte M., Zuk O., Carey B.W., Cassady J.P., Cabili M.N., Jaenisch R., Mikkelsen T.S., Jacks T., Hacohen N., Bernstein B.E., Kellis M., Regev A., Rinn J.L., Lander E.S. 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature, 458: 223-227

Haaland R.E., Herrmann C.H., Rice A.P. 2005. siRNA depletion of 7SK snRNA induces apoptosis but does not affect expression of the HIV-1 LTR or P-TEFb-dependent cellular genes. J. Cell. Physiol., 205: 463-470

Haltiwanger R.S., Holt G.D., Hart G.W. 1990. Enzymatic addition of O-GlcNAc to nuclear and cytoplasmic proteins. Identification of a uridine diphospho-N-acetylglucosamine:peptide beta-N-acetylglucosaminyltransferase. J. Biol. Chem., 265: 2563-2568

Hanahan D., Weinberg R.A. 2011. Hallmarks of cancer: the next generation. Cell, 144: 646-674

Hang H.C., Bertozzi C.R. 2005. The chemistry and biology of mucin-type O-linked glycosylation. Bioorg. Med. Chem., 13: 5021-5034

Harris A.L. 2002. Hypoxia--a key regulatory factor in tumour growth. Nat. Rev. Cancer, 2: 38-47

Hart G.W., Slawson C., Ramirez-Correa G., Lagerlof O. 2011. Cross talk between O-GlcNAcylation and phosphorylation: roles in signaling, transcription, and chronic disease. Annu. Rev. Biochem., 80: 825-858

He M., Wang Q.Y., Yin Q.Q., Tang J., Lu Y., Zhou C.X., Duan C.W., Hong D.L., Tanaka T., Chen G.Q., Zhao Q. 2013. HIF-1alpha downregulates miR-17/20a directly targeting p21 and STAT3: a role in myeloid leukemic cell differentiation. Cell Death Differ., 20: 408-418

Hebert C., Norris K., Parashar P., Ord R.A., Nikitakis N.G., Sauk J.J. 2006. Hypoxia-inducible factor-1alpha polymorphisms and TSC1/2 mutations are complementary in head and neck cancers. Mol. Cancer, 5: 3

Heino S., Kaare M., Andersson S., Laivuori H. 2008b. Non-synonymous sequence variants within the oxygen-dependent degradation domain of the HIF1A gene are not associated with pre-eclampsia in the Finnish population. BMC Med. Genet., 9: 96

Hennis P.J., Bussell C., Darlison M.G. 2010. The lack of associations between alleles at the hypoxia-inducible factor 1A C1772T loci and responses to acute hypoxia. Wilderness Environ. Med., 21: 219-228

Hiraga T., Kizaka-Kondoh S., Hirota K., Hiraoka M., Yoneda T. 2007. Hypoxia and hypoxia-inducible factor-1 expression enhance osteolytic bone metastases of breast cancer. Cancer Res., 67: 4157-4163

Hlatky M.A., Quertermous T., Boothroyd D.B., Priest J.R., Glassford A.J., Myers R.M., Fortmann S.P., Iribarren C., Tabor H.K., Assimes T.L., Tibshirani R.J., Go A.S. 2007. Polymorphisms in hypoxia inducible factor 1 and the initial clinical presentation of coronary disease. Am. Heart J., 154: 1035-1042

Hogenesch J.B., Chan W.K., Jackiw V.H., Brown R.C., Gu Y.Z., Pray-Grant M., Perdew G.H., Bradfield C.A. 1997. Characterization of a subset of the basic-helix-loop-helix-PAS superfamily that interacts with components of the dioxin signaling pathway. J. Biol. Chem., 272: 8581-8593

Hong J.M., Kim T.H., Chae S.C., Koo K.H., Lee Y.J., Park E.K., Choi J.Y., Ryoo H.M., Kim S.Y. 2007. Association study of hypoxia inducible factor 1alpha (HIF1alpha) with osteonecrosis of femoral head in a Korean population. Osteoarthritis Cartilage, 15: 688-694

Hoppeler H., Fluck M. 2003. Plasticity of skeletal muscle mitochondria: structure and function. Med. Sci. Sports Exerc., 35: 95-104

Horree N., Groot A.J., van Hattem W.A., Heintz A.P., Vooijs M., van Diest P.J. 2008. HIF-1A gene mutations associated with higher microvessel density in endometrial carcinomas. Histopathology, 52: 637-639

Hsiao P.C., Chen M.K., Su S.C., Ueng K.C., Chen Y.C., Hsieh Y.H., Liu Y.F., Tsai H.T., Yang S.F. 2010. Hypoxia inducible factor-1alpha gene polymorphism G1790A and its interaction with tobacco and alcohol consumptions increase susceptibility to hepatocellular carcinoma. J. Surg. Oncol., 102: 163-169

Huang da W., Sherman B.T., Lempicki R.A. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc., 4: 44-57

Huang L.E., Gu J., Schau M., Bunn H.F. 1998. Regulation of hypoxia-inducible factor 1alpha is mediated by an O2-dependent degradation domain via the ubiquitin-proteasome pathway. Proc. Natl. Acad. Sci. U S A, 95: 7987-7992

Huang X., Le Q.T., Giaccia A.J. 2010. MiR-210--micromanager of the hypoxia pathway. Trends Mol. Med., 16: 230-237

Huang Y., Hickey R.P., Yeh J.L., Liu D., Dadak A., Young L.H., Johnson R.S., Giordano F.J. 2004. Cardiac myocyte-specific HIF-1alpha deletion alters vascularization, energy availability, calcium flux, and contractility in the normoxic heart. FASEB J., 18: 1138-1140

Ikeda S., Kong S.W., Lu J., Bisping E., Zhang H., Allen P.D., Golub T.R., Pieske B., Pu W.T. 2007. Altered microRNA expression in human heart disease. Physiol. Genomics, 31: 367-373

Inoue Y., Peters L.L., Yim S.H., Inoue J., Gonzalez F.J. 2006. Role of hepatocyte nuclear factor 4alpha in control of blood coagulation factor gene expression. J. Mol. Med. (Berl), 84: 334-344

International Human Genome Sequencing C. 2004. Finishing the euchromatic sequence of the human genome. Nature, 431: 931-945

Iorio M.V., Ferracin M., Liu C.G., Veronese A., Spizzo R., Sabbioni S., Magri E., Pedriali M., Fabbri M., Campiglio M., Menard S., Palazzo J.P., Rosenberg A., Musiani P., Volinia S., Nenci I., Calin G.A., Querzoli P., Negrini M., Croce C.M. 2005. MicroRNA gene expression deregulation in human breast cancer. Cancer Res., 65: 7065-7070

Ivan M., Haberberger T., Gervasi D.C., Michelson K.S., Gunzler V., Kondo K., Yang H., Sorokina I., Conaway R.C., Conaway J.W., Kaelin W.G., Jr. 2002. Biochemical purification and pharmacological inhibition of a mammalian prolyl hydroxylase acting on hypoxia-inducible factor. Proc. Natl. Acad. Sci. U S A, 99: 13459-13464

Ivan M., Kondo K., Yang H., Kim W., Valiando J., Ohh M., Salic A., Asara J.M., Lane W.S., Kaelin W.G., Jr. 2001. HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. Science, 292: 464-468

Jaakkola P., Mole D.R., Tian Y.M., Wilson M.I., Gielbert J., Gaskell S.J., von Kriegsheim A., Hebestreit H.F., Mukherji M., Schofield C.J., Maxwell P.H., Pugh C.W., Ratcliffe P.J. 2001. Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. Science, 292: 468-472

Jacobs E.J., Hsing A.W., Bain E.B., Stevens V.L., Wang Y., Chen J., Chanock S.J., Zheng S.L., Xu J., Thun M.J., Calle E.E., Rodriguez C. 2008. Polymorphisms in angiogenesis-related genes and prostate cancer. Cancer Epidemiol. Biomarkers Prev., 17: 972-977

Jensen O.N. 2004. Modification-specific proteomics: characterization of post-translational modifications by mass spectrometry. Curr. Opin. Chem. Biol., 8: 33-41

Jensen O.N. 2006. Interpreting the protein language using proteomics. Nat. Rev. Mol. Cell Biol., 7: 391-403

Jiang B.H., Semenza G.L., Bauer C., Marti H.H. 1996. Hypoxia-inducible factor 1 levels vary exponentially over a physiologically relevant range of O2 tension. Am. J. Physiol., 271: C1172-1180

John B., Enright A.J., Aravin A., Tuschl T., Sander C., Marks D.S. 2004. Human MicroRNA targets. PLoS Biol., 2: e363

Kaelin W.G., Jr., Ratcliffe P.J. 2008. Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. Mol. Cell, 30: 393-402

Kallio P.J., Wilson W.J., O'Brien S., Makino Y., Poellinger L. 1999. Regulation of the hypoxia-inducible transcription factor 1alpha by the ubiquitin-proteasome pathway. J. Biol. Chem., 274: 6519-6525

Kang M.J., Jung S.A., Jung J.M., Kim S.E., Jung H.K., Kim T.H., Shim K.N., Yi S.Y., Yoo K., Moon I.H. 2011. Associations between single nucleotide polymorphisms of MMP2, VEGF, and HIF1A genes and the risk of developing colorectal cancer. Anticancer Res., 31: 575-584

Keith B., Johnson R.S., Simon M.C. 2012. HIF1alpha and HIF2alpha: sibling rivalry in hypoxic tumour growth and progression. Nat. Rev. Cancer, 12: 9-22

Kelley R.L., Kuroda M.I. 2000. Noncoding RNA genes in dosage compensation and imprinting. Cell, 103: 9-12

Kelly T.J., Souza A.L., Clish C.B., Puigserver P. 2011. A hypoxia-induced positive feedback loop promotes hypoxia-inducible factor 1alpha stability through miR-210 suppression of glycerol-3-phosphate dehydrogenase 1-like. Mol. Cell Biol., 31: 2696-2706

Kett-White R., Hutchinson P.J., Al-Rawi P.G., Czosnyka M., Gupta A.K., Pickard J.D., Kirkpatrick P.J. 2002. Cerebral oxygen and microdialysis monitoring during aneurysm surgery: effects of blood pressure, cerebrospinal fluid drainage, and temporary clipping on infarction. J. Neurosurg., 96: 1013-1019

Khurana S., Jaiswal A.K., Mukhopadhyay A. 2010. Hepatocyte nuclear factor-4alpha induces transdifferentiation of hematopoietic cells into hepatocytes. J. Biol. Chem., 285: 4725-4731

Kibel A., Iliopoulos O., DeCaprio J.A., Kaelin W.G., Jr. 1995. Binding of the von Hippel-Lindau tumor suppressor protein to Elongin B and C. Science, 269: 1444-1446

Kim H.O., Jo Y.H., Lee J., Lee S.S., Yoon K.S. 2008. The C1772T genetic polymorphism in human HIF-1alpha gene associates with expression of HIF-1alpha protein in breast cancer. Oncol. Rep., 20: 1181-1187

Kim S.J., Hwang S.H., Kim I.J., Lee M.K., Lee C.H., Lee S.Y., Lee E.Y. 2010. The association of 18F-deoxyglucose (FDG) uptake of PET with polymorphisms in the glucose transporter gene (SLC2A1) and hypoxia-related genes (HIF1A, VEGFA, APEX1) in non-small cell lung cancer. SLC2A1 polymorphisms and FDG-PET in NSCLC patients. J. Exp. Clin. Cancer Res., 29: 69

Kim J.H., Park M.Y., Kim C.N., Kim K.H., Kang H.B., Kim K.D., Kim J.W. 2012. Expression of endothelial cell-specific molecule-1 regulated by hypoxia inducible factor-1alpha in human colon carcinoma: impact of ESM-1 on prognosis and its correlation with clinicopathological features. Oncol. Rep., 28: 1701-1708

Kline D.D., Peng Y.J., Manalo D.J., Semenza G.L., Prabhakar N.R. 2002. Defective carotid body function and impaired ventilatory responses to chronic hypoxia in mice partially deficient for hypoxia-inducible factor 1 alpha. Proc. Natl. Acad. Sci. U S A, 99: 821-826

Knechtel G., Szkandera J., Stotz M., Hofmann G., Langsenlehner U., Krippl P., Samonigg H., Renner W., Langner C., Dehchamani D., Gerger A. 2010. Single nucleotide polymorphisms in the hypoxia-inducible factor-1 gene and colorectal cancer risk. Mol. Carcinog., 49: 805-809

Kojima H., Gu H., Nomura S., Caldwell C.C., Kobata T., Carmeliet P., Semenza G.L., Sitkovsky M.V. 2002. Abnormal B lymphocyte development and autoimmunity in hypoxia-inducible factor 1alpha -deficient chimeric mice. Proc. Natl. Acad. Sci. U S A, 99: 2170-2174

Kojima H., Jones B.T., Chen J., Cascalho M., Sitkovsky M.V. 2004. Hypoxia-inducible factor 1alpha-deficient chimeric mice as a model to study abnormal B lymphocyte development and autoimmunity. Methods Enzymol., 381: 218-229

Kolyada A.Y., Tighiouart H., Perianayagam M.C., Liangos O., Madias N.E., Jaber B.L. 2009. A genetic variant of hypoxia-inducible factor-1alpha is associated with adverse outcomes in acute kidney injury. Kidney Int., 75: 1322-1329

Konac E., Dogan I., Onen H.I., Yurdakul A.S., Ozturk C., Varol A., Ekmecki A. 2009. Genetic variations in the hypoxia-inducible factor-1alpha gene and lung cancer. Exp. Biol. Med. (Maywood), 234: 1109-1116

Konac E., Onen H.I., Metindir J., Alp E., Biri A.A., Ekmekci A. 2007. An investigation of relationships between hypoxia-inducible factor-1 alpha gene polymorphisms and ovarian, cervical and endometrial cancers. Cancer. Detect. Prev., 31: 102-109

Kong W., Zhao J.J., He L., Cheng J.Q. 2009. Strategies for profiling microRNA expression. J. Cell Physiol., 218: 22-25

Koshiji M., Kageyama Y., Pete E.A., Horikawa I., Barrett J.C., Huang L.E. 2004. HIF-1alpha induces cell cycle arrest by functionally counteracting Myc. EMBO J., 23: 1949-1956

Koukourakis M.I., Giatromanolaki A., Sivridis E., Simopoulos C., Turley H., Talks K., Gatter K.C., Harris A.L. 2002. Hypoxia-inducible factor (HIF1A and HIF2A), angiogenesis, and chemoradiotherapy outcome of squamous cell head-and-neck cancer. Int. J. Radiat. Oncol. Biol. Phys., 53: 1192-1202

Koukourakis M.I., Papazoglou D., Giatromanolaki A., Panagopoulos I., Maltezos E., Harris A.L., Gatter K.C., Sivridis E. 2006. C2028T polymorphism in exon 12 and dinucleotide repeat polymorphism in intron 13 of the HIF-1alpha gene define HIF-1alpha protein expression in non-small cell lung cancer. Lung Cancer, 53: 257-262

Kozomara A., Griffiths-Jones S. 2011. miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res., 39: D152-157

Kreppel L.K., Blomberg M.A., Hart G.W. 1997. Dynamic glycosylation of nuclear and cytosolic proteins. Cloning and characterization of a unique O-GlcNAc transferase with multiple tetratricopeptide repeats. J. Biol. Chem., 272: 9308-9315

Krichevsky A.M., King K.S., Donahue C.P., Khrapko K., Kosik K.S. 2003. A microRNA array reveals extensive regulation of microRNAs during brain development. RNA, 9: 1274-1281

Kulshreshtha R., Davuluri R.V., Calin G.A., Ivan M. 2008. A microRNA component of the hypoxic response. Cell Death Differ., 15: 667-671

Kulshreshtha R., Ferracin M., Wojcik S.E., Garzon R., Alder H., Agosto-Perez F.J., Davuluri R., Liu C.G., Croce C.M., Negrini M., Calin G.A., Ivan M. 2007. A microRNA signature of hypoxia. Mol. Cell. Biol., 27: 1859-1867

Kunej T., Godnic I., Ferdin J., Horvat S., Dovc P., Calin G.A. 2011. Epigenetic regulation of microRNAs in cancer: an integrated review of literature. Mutat. Res., 717: 77-84

Kunej T., Godnic I., Horvat S., Zorc M., Calin G.A. 2012. Cross talk between microRNA and coding cancer genes. Cancer J., 18: 223-231

Kung A.L., Wang S., Klco J.M., Kaelin W.G., Livingston D.M. 2000. Suppression of tumor growth through disruption of hypoxia-inducible transcription. Nat. Med., 6: 1335-1340

Kuo W.H., Shih C.M., Lin C.W., Cheng W.E., Chen S.C., Chen W., Lee Y.L. 2012. Association of hypoxia inducible factor-1alpha polymorphisms with susceptibility to non-small-cell lung cancer. Transl. Res., 159: 42-50

Kuwai T., Kitadai Y., Tanaka S., Kuroda T., Ochiumi T., Matsumura S., Oue N., Yasui W., Kaneyasu M., Tanimoto K., Nishiyama M., Chayama K. 2004. Single nucleotide polymorphism in the hypoxia-inducible factor-1alpha gene in colorectal carcinoma. Oncol. Rep., 12: 1033-1037

Lagos-Quintana M., Rauhut R., Lendeckel W., Tuschl T. 2001. Identification of novel genes coding for small expressed RNAs. Science, 294: 853-858

Lando D., Peet D.J., Gorman J.J., Whelan D.A., Whitelaw M.L., Bruick R.K. 2002a. FIH-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor. Genes. Dev., 16: 1466-1471

Lando D., Peet D.J., Whelan D.A., Gorman J.J., Whitelaw M.L. 2002b. Asparagine hydroxylation of the HIF transactivation domain a hypoxic switch. Science, 295: 858-861

Lareau L.F., Inada M., Green R.E., Wengrod J.C., Brenner S.E. 2007. Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. Nature, 446: 926-929

Lazorchak A.S., Schlissel M.S., Zhuang Y. 2006. E2A and IRF-4/Pip promote chromatin modification and transcription of the immunoglobulin kappa locus in pre-B cells. Mol. Cell Biol., 26: 810-821

Le Hir H., Nott A., Moore M.J. 2003. How introns influence and enhance eukaryotic gene expression. Trends Biochem. Sci., 28: 215-220

Lee J.Y., Choi J.Y., Lee K.M., Park S.K., Han S.H., Noh D.Y., Ahn S.H., Kim D.H., Hong Y.C., Ha E., Yoo K.Y., Ambrosone C.B., Kang D. 2008. Rare variant of

hypoxia-inducible factor-1alpha (HIF-1A) and breast cancer risk in Korean women. Clin. Chim. Acta, 389: 167-170

Lee Y.S., Dutta A. 2009. MicroRNAs in cancer. Annu. Rev. Pathol., 4: 199-227

Lei Z., Li B., Yang Z., Fang H., Zhang G.M., Feng Z.H., Huang B. 2009. Regulation of HIF-1alpha and VEGF by miR-20b tunes tumor cells to adapt to the alteration of oxygen concentration. PLoS One, 4: e7629

Lewis B.P., Burge C.B., Bartel D.P. 2005. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell, 120: 15-20

Li H., Bubley G.J., Balk S.P., Gaziano J.M., Pollak M., Stampfer M.J., Ma J. 2007. Hypoxia-inducible factor-1alpha (HIF-1alpha) gene polymorphisms, circulating insulin-like growth factor binding protein (IGFBP)-3 levels and prostate cancer. Prostate, 67: 1354-1361

Li K., Zhang Y., Dan Z., Wang Y., Ren Z.C. 2009. Association of the hypoxia inducible factor-1alpha gene polymorphisms with gastric cancer in Tibetans. Biochem. Genet., 47: 625-634

Li P., Cao Q., Shao P.F., Cai H.Z., Zhou H., Chen J.W., Qin C., Zhang Z.D., Ju X.B., Yin C.J. 2012. Genetic polymorphisms in HIF1A are associated with prostate cancer risk in a Chinese population. Asian. J. Androl., 14: 864-869

Ling T.S., Shi R.H., Zhang G.X., Zhu H., Yu L.Z., Ding X.F. 2005. Common single nucleotide polymorphism of hypoxia-inducible factor-1alpha and its impact on the clinicopathological features of esophageal squamous cell carcinoma. Chin. J. Dig. Dis., 6: 155-158

Liu C.G., Calin G.A., Meloon B., Gamliel N., Sevignani C., Ferracin M., Dumitru C.D., Shimizu M., Zupo S., Dono M., Alder H., Bullrich F., Negrini M., Croce C.M. 2004. An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. Proc. Natl. Acad. Sci. U S A, 101: 9740-9744

Liu C.G., Spizzo R., Calin G.A., Croce C.M. 2008a. Expression profiling of microRNA using oligo DNA arrays. Methods, 44: 22-30

Liu C.J., Tsai M.M., Hung P.S., Kao S.Y., Liu T.Y., Wu K.J., Chiou S.H., Lin S.C., Chang K.W. 2010. miR-31 ablates expression of the HIF regulatory factor FIH to activate the HIF pathway in head and neck carcinoma. Cancer Res., 70: 1635-1644

Liu L., Marti G.P., Wei X., Zhang X., Zhang H., Liu Y.V., Nastai M., Semenza G.L., Harmon J.W. 2008b. Age-dependent impairment of HIF-1alpha expression in diabetic mice: Correction with electroporation-facilitated gene therapy increases

wound healing, angiogenesis, and circulating angiogenic cells. J. Cell. Physiol., 217: 319-327

Livak K.J., Schmittgen T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods, 25: 402-408

Loayza-Puch F., Yoshida Y., Matsuzaki T., Takahashi C., Kitayama H., Noda M. 2010. Hypoxia and RAS-signaling pathways converge on, and cooperatively downregulate, the RECK tumor-suppressor protein through microRNAs. Oncogene, 29: 2638-2648

Love D.C., Ghosh S., Mondoux M.A., Fukushige T., Wang P., Wilson M.A., Iser W.B., Wolkow C.A., Krause M.W., Hanover J.A. 2010. Dynamic O-GlcNAc cycling at promoters of Caenorhabditis elegans genes regulating longevity, stress, and immunity. Proc. Natl. Acad. Sci U S A, 107: 7413-7418

Lu J., Getz G., Miska E.A., Alvarez-Saavedra E., Lamb J., Peck D., Sweet-Cordero A., Ebert B.L., Mak R.H., Ferrando A.A., Downing J.R., Jacks T., Horvitz H.R., Golub T.R. 2005. MicroRNA expression profiles classify human cancers. Nature, 435: 834-838

Lu P., Rha G.B., Melikishvili M., Wu G., Adkins B.C., Fried M.G., Chi Y.I. 2008. Structural basis of natural promoter recognition by a unique nuclear receptor, HNF4alpha. Diabetes gene product. J. Biol. Chem., 283: 33685-33697

Lubas W.A., Frank D.W., Krause M., Hanover J.A. 1997. O-Linked GlcNAc transferase is a conserved nucleocytoplasmic protein containing tetratricopeptide repeats. J. Biol. Chem., 272: 9316-9324

Lynch T.P., Ferrer C.M., Jackson S.R., Shahriari K.S., Vosseller K., Reginato M.J. 2012. Critical role of O-GlcNAc transferase in prostate cancer invasion, angiogenesis and metastasis. J. Biol. Chem., 287: 11070-11081

Mahon P.C., Hirota K., Semenza G.L. 2001. FIH-1: a novel protein that interacts with HIF-1alpha and VHL to mediate repression of HIF-1 transcriptional activity. Genes Dev., 15: 2675-2686

Majmundar A.J., Wong W.J., Simon M.C. 2010. Hypoxia-inducible factors and the response to hypoxic stress. Mol. Cell, 40: 294-309

Makeyev E.V., Zhang J., Carrasco M.A., Maniatis T. 2007. The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. Mol. Cell, 27: 435-448

Makunin I.V., Pheasant M., Simons C., Mattick J.S. 2007. Orthologous microRNA genes are located in cancer-associated genomic regions in human and mouse. PLoS One, 2: e1133

Mardis E.R. 2008. The impact of next-generation sequencing technology on genetics. Trends Genet., 24: 133-141

Marshall S. 2006. Role of insulin, adipocyte hormones, and nutrient-sensing pathways in regulating fuel metabolism and energy homeostasis: a nutritional perspective of diabetes, obesity, and cancer. Sci. STKE, 2006: re7

Mason S.D., Howlett R.A., Kim M.J., Olfert I.M., Hogan M.C., McNulty W., Hickey R.P., Wagner P.D., Kahn C.R., Giordano F.J., Johnson R.S. 2004. Loss of skeletal muscle HIF-1alpha results in altered exercise endurance. PLoS Biol., 2: e288

Masson N., Willam C., Maxwell P.H., Pugh C.W., Ratcliffe P.J. 2001. Independent function of two destruction domains in hypoxia-inducible factor-alpha chains activated by prolyl hydroxylation. EMBO J., 20: 5197-5206

Mattick J.S. 2009. The genetic signatures of noncoding RNAs. PLoS Genet., 5: e1000459

Mattick J.S., Makunin I.V. 2005. Small regulatory RNAs in mammals. Hum. Mol. Genet., 14 Spec No 1: R121-132

Mattick J.S., Makunin I.V. 2006. Non-coding RNA. Hum. Mol. Genet., 15 Spec No 1: R17-29

Matys V., Fricke E., Geffers R., Gossling E., Haubrock M., Hehl R., Hornischer K., Karas D., Kel A.E., Kel-Margoulis O.V., Kloos D.U., Land S., Lewicki-Potapov B., Michael H., Munch R., Reuter I., Rotert S., Saxel H., Scheer M., Thiele S., Wingender E. 2003. TRANSFAC: transcriptional regulation, from patterns to profiles. Nucleic Acids Res., 31: 374-378

Maxwell P.H., Pugh C.W., Ratcliffe P.J. 2001. Activation of the HIF pathway in cancer. Curr. Opin. Genet. Dev., 11: 293-299

McPhee J.S., Perez-Schindler J., Degens H., Tomlinson D., Hennis P., Baar K., Williams A.G. 2011. HIF1A P582S gene association with endurance training responses in young women. Eur. J. Appl. Physiol., 111: 2339-2347

Mercer T.R., Dinger M.E., Mattick J.S. 2009. Long non-coding RNAs: insights into functions. Nat. Rev. Genet., 10: 155-159

Mestdagh P., Fredlund E., Pattyn F., Rihani A., Van Maerken T., Vermeulen J., Kumps C., Menten B., De Preter K., Schramm A., Schulte J., Noguera R., Schleiermacher G., Janoueix-Lerosey I., Laureys G., Powel R., Nittner D., Marine J.C., Ringner M., Speleman F., Vandesompele J. 2010. An integrative genomics screen uncovers ncRNA T-UCR functions in neuroblastoma tumours. Oncogene, 29: 3583-3592

Mi W., Gu Y., Han C., Liu H., Fan Q., Zhang X., Cong Q., Yu W. 2011. O-GlcNAcylation is a novel regulator of lung and colon cancer malignancy. Biochim. Biophys. Acta, 1812: 514-519

Minet E., Ernest I., Michel G., Roland I., Remacle J., Raes M., Michiels C. 1999. HIF1A gene transcription is dependent on a core promoter sequence encompassing activating and inhibiting sequences located upstream from the transcription initiation site and cis elements located within the 5'UTR. Biochem. Biophys. Res. Commun., 261: 534-540

Miranda K.C., Huynh T., Tay Y., Ang Y.S., Tam W.L., Thomson A.M., Lim B., Rigoutsos I. 2006. A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell, 126: 1203-1217

Moreau M.P., Bruse S.E., Jornsten R., Liu Y., Brzustowicz L.M. 2013. Chronological Changes in MicroRNA Expression in the Developing Human Brain. PLoS One, 8: e60480

Morris M.R., Hughes D.J., Tian Y.M., Ricketts C.J., Lau K.W., Gentle D., Shuib S., Serrano-Fernandez P., Lubinski J., Wiesener M.S., Pugh C.W., Latif F., Ratcliffe P.J., Maher E.R. 2009. Mutation analysis of hypoxia-inducible factors HIF1A and HIF2A in renal cell carcinoma. Anticancer Res., 29: 4337-4343

Mouse Genome Sequencing C., Waterston R.H., Lindblad-Toh K., Birney E., Rogers J., Abril J.F., Agarwal P., Agarwala R., Ainscough R., Alexandersson M., An P., Antonarakis S.E., Attwood J., Baertsch R., Bailey J., Barlow K., Beck S., Berry E., Birren B., Bloom T., Bork P., Botcherby M., Bray N., Brent M.R., Brown D.G., Brown S.D., Bult C., Burton J., Butler J., Campbell R.D., Carninci P., Cawley S., Chiaromonte F., Chinwalla A.T., Church D.M., Clamp M., Clee C., Collins F.S., Cook L.L., Copley R.R., Coulson A., Couronne O., Cuff J., Curwen V., Cutts T., Daly M., David R., Davies J., Delehaunty K.D., Deri J., Dermitzakis E.T., Dewey C., Dickens N.J., Diekhans M., Dodge S., Dubchak I., Dunn D.M., Eddy S.R., Elnitski L., Emes R.D., Eswara P., Eyras E., Felsenfeld A., Fewell G.A., Flicek P., Foley K., Frankel W.N., Fulton L.A., Fulton R.S., Furey T.S., Gage D., Gibbs R.A., Glusman G., Gnerre S., Goldman N., Goodstadt L., Grafham D., Graves T.A., Green E.D., Gregory S., Guigo R., Guyer M., Hardison R.C., Haussler D., Hayashizaki Y., Hillier L.W., Hinrichs A., Hlavina W., Holzer T., Hsu F., Hua A., Hubbard T., Hunt A., Jackson I., Jaffe D.B., Johnson L.S., Jones M., Jones T.A., Joy A., Kamal M., Karlsson E.K., Karolchik D., Kasprzyk A., Kawai J., Keibler E., Kells C., Kent W.J., Kirby A., Kolbe D.L., Korf I., Kucherlapati R.S., Kulbokas E.J., Kulp D., Landers T., Leger J.P., Leonard S., Letunic I., Levine R., Li J., Li M., Lloyd C., Lucas S., Ma B., Maglott D.R., Mardis E.R., Matthews L., Mauceli E., Mayer J.H., McCarthy M., McCombie W.R., McLaren S., McLay K., McPherson J.D.,

Meldrim J., Meredith B., Mesirov J.P., Miller W., Miner T.L., Mongin E., Montgomery K.T., Morgan M., Mott R., Mullikin J.C., Muzny D.M., Nash W.E., Nelson J.O., Nhan M.N., Nicol R., Ning Z., Nusbaum C., O'Connor M.J., Okazaki Y., Oliver K., Overton-Larty E., Pachter L., Parra G., Pepin K.H., Peterson J., Pevzner P., Plumb R., Pohl C.S., Poliakov A., Ponce T.C., Ponting C.P., Potter S., Quail M., Reymond A., Roe B.A., Roskin K.M., Rubin E.M., Rust A.G., Santos R., Sapojnikov V., Schultz B., Schultz J., Schwartz M.S., Schwartz S., Scott C., Seaman S., Searle S., Sharpe T., Sheridan A., Shownkeen R., Sims S., Singer J.B., Slater G., Smit A., Smith D.R., Spencer B., Stabenau A., Stange-Thomann N., Sugnet C., Suyama M., Tesler G., Thompson J., Torrents D., Trevaskis E., Tromp J., Ucla C., Ureta-Vidal A., Vinson J.P., Von Niederhausern A.C., Wade C.M., Wall M., Weber R.J., Weiss R.B., Wendl M.C., West A.P., Wetterstrand K., Wheeler R., Whelan S., Wierzbowski J., Willey D., Williams S., Wilson R.K., Winter E., Worley K.C., Wyman D., Yang S., Yang S.P., Zdobnov E.M., Zody M.C., Lander E.S. 2002. Initial sequencing and comparative analysis of the mouse genome. Nature, 420: 520-562

Munoz-Guerra M.F., Fernandez-Contreras M.E., Moreno A.L., Martin I.D., Herraez B., Gamallo C. 2009. Polymorphisms in the hypoxia inducible factor 1-alpha and the impact on the prognosis of early stages of oral cancer. Ann. Surg. Oncol., 16: 2351-2358

Nadaoka J., Horikawa Y., Saito M., Kumazawa T., Inoue T., Narita S., Yuasa T., Satoh S., Nishiyama H., Ogawa O., Tsuchiya N., Habuchi T. 2008. Prognostic significance of HIF-1 alpha polymorphisms in transitional cell carcinoma of the bladder. Int. J. Cancer, 122: 1297-1302

Nagy G., Kovacs-Nagy R., Kereszturi E., Somogyi A., Szekely A., Nemeth N., Hosszufalusi N., Panczel P., Ronai Z., Sasvari-Szekely M. 2009. Association of hypoxia inducible factor-1 alpha gene polymorphism with both type 1 and type 2 diabetes in a Caucasian (Hungarian) sample. BMC Med. Genet., 10: 79

Naidu R., Har Y.C., Taib N.A. 2009. Associations between hypoxia-inducible factor-1alpha (HIF-1alpha) gene polymorphisms and risk of developing breast cancer. Neoplasma, 56: 441-447

Nallamshetty S., Chan S.Y., Loscalzo J. 2013. Hypoxia: A Master Regulator of MicroRNA Biogenesis and Activity: Hypoxic Regulation of MicroRNAs. Free. Radic. Biol. Med., 64: 20-30

Nava-Salazar S., Sanchez-Rodriguez E.N., Mendoza-Rodriguez C.A., Moran C., Romero-Arauz J.F., Cerbon M.A. 2011. Polymorphisms in the hypoxia-inducible factor 1 alpha gene in Mexican patients with preeclampsia: A case-control study. BMC Res. Notes, 4: 68

Ni J.Z., Grate L., Donohue J.P., Preston C., Nobida N., O'Brien G., Shiue L., Clark T.A., Blume J.E., Ares M., Jr. 2007. Ultraconserved elements are associated with homeostatic control of splicing regulators by alternative splicing and nonsense-mediated decay. Genes Dev., 21: 708-718

Nicoloso M.S., Calin G.A. 2008. MicroRNA involvement in brain tumors: from bench to bedside. Brain. Pathol., 18: 122-129

Nicoloso M.S., Spizzo R., Shimizu M., Rossi S., Calin G.A. 2009. MicroRNAs--the micro steering wheel of tumour metastases. Nat. Rev. Cancer, 9: 293-302

Nicoloso M.S., Sun H., Spizzo R., Kim H., Wickramasinghe P., Shimizu M., Wojcik S.E., Ferdin J., Kunej T., Xiao L., Manoukian S., Secreto G., Ravagnani F., Wang X., Radice P., Croce C.M., Davuluri R.V., Calin G.A. 2010. Single-nucleotide polymorphisms inside microRNA target sites influence tumor susceptibility. Cancer Res., 70: 2789-2798

Ning B.F., Ding J., Yin C., Zhong W., Wu K., Zeng X., Yang W., Chen Y.X., Zhang J.P., Zhang X., Wang H.Y., Xie W.F. 2010. Hepatocyte nuclear factor 4 alpha suppresses the development of hepatocellular carcinoma. Cancer Res., 70: 7640-7651

Nobrega M.A., Ovcharenko I., Afzal V., Rubin E.M. 2003. Scanning human gene deserts for long-range enhancers. Science, 302: 413

Nolan T., Hands R.E., Bustin S.A. 2006. Quantification of mRNA using real-time RT-PCR. Nat. Protoc., 1: 1559-1582

Nolte D., Muller U. 2002. Human O-GlcNAc transferase (OGT): genomic structure, analysis of splice variants, fine mapping in Xq13.1. Mamm. Genome, 13: 62-64

Nott A., Meislin S.H., Moore M.J. 2003. A quantitative analysis of intron effects on mammalian gene expression. RNA, 9: 607-617

Nuovo G.J., Elton T.S., Nana-Sinkam P., Volinia S., Croce C.M., Schmittgen T.D. 2009. A methodology for the combined in situ analyses of the precursor and mature forms of microRNAs and correlation with their putative targets. Nat. Protoc., 4: 107-115

Ollerenshaw M., Page T., Hammonds J., Demaine A. 2004. Polymorphisms in the hypoxia inducible factor-1alpha gene (HIF1A) are associated with the renal cell carcinoma phenotype. Cancer Genet. Cytogenet., 153: 122-126

Orom U.A., Derrien T., Beringer M., Gumireddy K., Gardini A., Bussotti G., Lai F., Zytnicki M., Notredame C., Huang Q., Guigo R., Shiekhattar R. 2010. Long noncoding RNAs with enhancer-like function in human cells. Cell, 143: 46-58

Orr-Urtreger A., Bar-Shira A., Matzkin H., Mabjeesh N.J. 2007. The homozygous P582S mutation in the oxygen-dependent degradation domain of HIF-1 alpha is associated with increased risk for prostate cancer. Prostate, 67: 8-13

Osada R., Horiuchi A., Kikuchi N., Yoshida J., Hayashi A., Ota M., Katsuyama Y., Melillo G., Konishi I. 2007. Expression of hypoxia-inducible factor 1alpha, hypoxia-inducible factor 2alpha, and von Hippel-Lindau protein in epithelial ovarian neoplasms and allelic loss of von Hippel-Lindau gene: nuclear expression of hypoxia-inducible factor 1alpha is an independent prognostic factor in ovarian carcinoma. Hum. Pathol., 38: 1310-1320

Ozcan S., Andrali S.S., Cantrell J.E. 2010. Modulation of transcription factor function by O-GlcNAc modification. Biochim. Biophys. Acta, 1799: 353-364

Ozsolak F., Poling L.L., Wang Z., Liu H., Liu X.S., Roeder R.G., Zhang X., Song J.S., Fisher D.E. 2008. Chromatin structure analyses identify miRNA promoters. Genes Dev., 22: 3172-3183

Palii C.G., Perez-Iratxeta C., Yao Z., Cao Y., Dai F., Davison J., Atkins H., Allan D., Dilworth F.J., Gentleman R., Tapscott S.J., Brand M. 2011. Differential genomic targeting of the transcription factor TAL1 in alternate haematopoietic lineages. EMBO J., 30: 494-509

Pang K.C., Stephen S., Dinger M.E., Engstrom P.G., Lenhard B., Mattick J.S. 2007. RNAdb 2.0--an expanded database of mammalian non-coding RNAs. Nucleic Acids Res., 35: D178-182

Paparidis Z., Abbasi A.A., Malik S., Goode D.K., Callaway H., Elgar G., deGraaff E., Lopez-Rios J., Zeller R., Grzeschik K.H. 2007. Ultraconserved non-coding sequence element controls a subset of spatiotemporal GLI3 expression. Dev. Growth Differ., 49: 543-553

Park S.W., Chung N.G., Hur S.Y., Kim H.S., Yoo N.J., Lee S.H. 2009. Mutational analysis of hypoxia-related genes HIF1alpha and CUL2 in common human cancers. APMIS, 117: 880-885

Pennacchio L.A., Ahituv N., Moses A.M., Prabhakar S., Nobrega M.A., Shoukry M., Minovitsky S., Dubchak I., Holt A., Lewis K.D., Plajzer-Frick I., Akiyama J., De Val S., Afzal V., Black B.L., Couronne O., Eisen M.B., Visel A., Rubin E.M. 2006. In vivo enhancer analysis of human conserved non-coding sequences. Nature, 444: 499-502

Percy M.J., Mooney S.M., McMullin M.F., Flores A., Lappin T.R., Lee F.S. 2003. A common polymorphism in the oxygen-dependent degradation (ODD) domain of hypoxia inducible factor-1alpha (HIF-1alpha) does not impair Pro-564 hydroxylation. Mol. Cancer, 2: 31

Pescador N., Cuevas Y., Naranjo S., Alcaide M., Villar D., Landazuri M.O., Del Peso L. 2005. Identification of a functional hypoxia-responsive element that regulates the expression of the egl nine homologue 3 (egln3/phd3) gene. Biochem J., 390: 189-197

Pichiorri F., Ishii H., Okumura H., Trapasso F., Wang Y., Huebner K. 2008. Molecular parameters of genome instability: roles of fragile genes at common fragile sites. J Cell Biochem., 104: 1525-1533

Prior S.J., Hagberg J.M., Phares D.A., Brown M.D., Fairfull L., Ferrell R.E., Roth S.M. 2003. Sequence variation in hypoxia-inducible factor 1alpha (HIF1A): association with maximal oxygen consumption. Physiol. Genomics, 15: 20-26

Psaila B., Lyden D. 2009. The metastatic niche: adapting the foreign soil. Nat. Rev. Cancer, 9: 285-293

Pugh C.W., Ratcliffe P.J. 2003. Regulation of angiogenesis by hypoxia: role of the HIF system. Nat. Med., 9: 677-684

Puissegur M.P., Mazure N.M., Bertero T., Pradelli L., Grosso S., Robbe-Sermesant K., Maurin T., Lebrigand K., Cardinaud B., Hofman V., Fourre S., Magnone V., Ricci J.E., Pouyssegur J., Gounon P., Hofman P., Barbry P., Mari B. 2011. miR-210 is overexpressed in late stages of lung cancer and mediates mitochondrial alterations associated with modulation of HIF-1 activity. Cell Death Differ., 18: 465-478

Putra A.C., Tanimoto K., Arifin M., Antariksa B., Hiyama K. 2011a. Genetic variations in detoxification enzymes and HIF-1alpha in Japanese patients with COPD. Clin. Respir. J., 7:7-15

Putra A.C., Tanimoto K., Arifin M., Hiyama K. 2011b. Hypoxia-inducible factor-1alpha polymorphisms are associated with genetic aberrations in lung cancer. Respirology, 16: 796-802

Qin C., Cao Q., Ju X., Wang M., Meng X., Zhu J., Yan F., Li P., Ding Q., Chen J., Gu M., Zhang W., Yin C., Zhang Z. 2011. The polymorphisms in the VHL and HIF1A genes are associated with the prognosis but not the development of renal cell carcinoma. Ann. Oncol., 23:981-989

Rabinovitch M., Gamble W., Nadas A.S., Miettinen O.S., Reid L. 1979. Rat pulmonary circulation after chronic hypoxia: hemodynamic and structural features. Am. J. Physiol., 236: H818-827

Rane S., He M., Sayed D., Vashistha H., Malhotra A., Sadoshima J., Vatner D.E., Vatner S.F., Abdellatif M. 2009. Downregulation of miR-199a derepresses hypoxia-inducible factor-1alpha and Sirtuin 1 and recapitulates hypoxia preconditioning in cardiac myocytes. Circ. Res., 104: 879-886

Rane S., He M., Sayed D., Yan L., Vatner D., Abdellatif M. 2010. An antagonism between the AKT and beta-adrenergic signaling pathways mediated through their reciprocal effects on miR-199a-5p. Cell Signal., 22: 1054-1062

Rankin E.B., Giaccia A.J. 2008. The role of hypoxia-inducible factors in tumorigenesis. Cell Death Differ., 15: 678-685

Resar J.R., Roguin A., Voner J., Nasir K., Hennebry T.A., Miller J.M., Ingersoll R., Kasch L.M., Semenza G.L. 2005. Hypoxia-inducible factor 1alpha polymorphism and coronary collaterals in patients with ischemic heart disease. Chest, 128: 787-791

Reynaud D., Ravet E., Titeux M., Mazurier F., Renia L., Dubart-Kupperschmitt A., Romeo P.H., Pflumio F. 2005. SCL/TAL1 expression level regulates human hematopoietic stem cell self-renewal and engraftment. Blood, 106: 2318-2328

Richardson R.S., Newcomer S.C., Noyszewski E.A. 2001. Skeletal muscle intracellular PO(2) assessed by myoglobin desaturation: response to graded exercise. J. Appl. Physiol., 91: 2679-2685

Richardson R.S., Noyszewski E.A., Kendrick K.F., Leigh J.S., Wagner P.D. 1995. Myoglobin O2 desaturation during exercise. Evidence of limited O2 transport. J. Clin. Invest., 96: 1916-1926

Robinson K.A., Weinstein M.L., Lindenmayer G.E., Buse M.G. 1995. Effects of diabetes and hyperglycemia on the hexosamine synthesis pathway in rat muscle and liver. Diabetes, 44: 1438-1446

Romano P., Manniello A., Aresu O., Armento M., Cesaro M., Parodi B. 2009. Cell Line Data Base: structure and recent improvements towards molecular authentication of human cell lines. Nucleic Acids Res., 37: D925-932

Ruiz-Tovar J., Fernandez-Contreras M.E., Martin-Perez E., Gamallo C. 2012. Association of thymidylate synthase and hypoxia inducible factor-1alpha DNA polymorphisms with pancreatic cancer. Tumori, 98: 364-369

Ryan H.E., Lo J., Johnson R.S. 1998. HIF-1 alpha is required for solid tumor formation and embryonic vascularization. EMBO J., 17: 3005-3015

Ryan H.E., Poloni M., McNulty W., Elson D., Gassmann M., Arbeit J.M., Johnson R.S. 2000. Hypoxia-inducible factor-1alpha is a positive factor in solid tumor growth. Cancer Res., 60: 4010-4015

Ryu W.S., Mertz J.E. 1989. Simian virus 40 late transcripts lacking excisable intervening sequences are defective in both stability in the nucleus and transport to the cytoplasm. J. Virol., 63: 4386-4394

Safdar A., Abadi A., Akhtar M., Hettinga B.P., Tarnopolsky M.A. 2009. miRNA in the regulation of skeletal muscle adaptation to acute endurance exercise in C57Bl/6J male mice. PLoS ONE, 4: e5610

Sakamoto S., Wakae K., Anzai Y., Murai K., Tamaki N., Miyazaki M., Miyazaki K., Romanow W.J., Ikawa T., Kitamura D., Yanagihara I., Minato N., Murre C., Agata Y. 2012. E2A and CBP/p300 act in synergy to promote chromatin accessibility of the immunoglobulin kappa locus. J. Immunol., 188: 5547-5560

Sakuraba Y., Kimura T., Masuya H., Noguchi H., Sezutsu H., Takahasi K.R., Toyoda A., Fukumura R., Murata T., Sakaki Y., Yamamura M., Wakana S., Noda T., Shiroishi T., Gondo Y. 2008. Identification and characterization of new long conserved noncoding sequences in vertebrates. Mamm. Genome, 19: 703-712

Sarzynski M.A., Rankinen T., Sternfeld B., Grove M.L., Fornage M., Jacobs D.R., Jr., Sidney S., Bouchard C. 2010a. Association of single-nucleotide polymorphisms from 17 candidate genes with baseline symptom-limited exercise test duration and decrease in duration over 20 years: the Coronary Artery Risk Development in Young Adults (CARDIA) fitness study. Circ. Cardiovasc. Genet., 3: 531-538

Sathirapongsasuti J.F., Sathira N., Suzuki Y., Huttenhower C., Sugano S. 2011. Ultraconserved cDNA segments in the human transcriptome exhibit resistance to folding and implicate function in translation and alternative splicing. Nucleic Acids Res., 39: 1967-1979

Saunders M.A., Liang H., Li W.H. 2007. Human polymorphism at microRNAs and microRNA target sites. Proc. Natl. Acad. Sci. U S A, 104: 3300-3305

Scaruffi P., Stigliani S., Coco S., Valdora F., De Vecchi C., Bonassi S., Tonini G.P. 2010. Transcribed-ultra conserved region expression profiling from low-input total RNA. BMC Genomics, 11: 149

Schetter A.J., Leung S.Y., Sohn J.J., Zanetti K.A., Bowman E.D., Yanaihara N., Yuen S.T., Chan T.L., Kwong D.L., Au G.K., Liu C.G., Calin G.A., Croce C.M., Harris C.C. 2008. MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. JAMA, 299: 425-436

Schindl M., Schoppmann S.F., Samonigg H., Hausmaninger H., Kwasny W., Gnant M., Jakesz R., Kubista E., Birner P., Oberhuber G., Austrian B., Colorectal Cancer Study G. 2002. Overexpression of hypoxia-inducible factor 1alpha is associated with an unfavorable prognosis in lymph node-positive breast cancer. Clin. Cancer Res., 8: 1831-1837

Schmittgen T.D., Jiang J., Liu Q., Yang L. 2004. A high-throughput method to monitor the expression of microRNA precursors. Nucleic Acids Res., 32: e43

Schofield C.J., Ratcliffe P.J. 2004. Oxygen sensing by HIF hydroxylases. Nat. Rev. Mol. Cell Biol., 5: 343-354

Schratt G.M., Tuebing F., Nigh E.A., Kane C.G., Sabatini M.E., Kiebler M., Greenberg M.E. 2006. A brain-specific microRNA regulates dendritic spine development. Nature, 439: 283-289

Semenza G.L. 2003. Targeting HIF-1 for cancer therapy. Nat Rev Cancer, 3: 721-732

Semenza G.L. 2007a. Evaluation of HIF-1 inhibitors as anticancer agents. Drug Discov. Today, 12: 853-859

Semenza G.L. 2007b. Hypoxia-inducible factor 1 (HIF-1) pathway. Sci. STKE, 2007: cm8

Sevignani C., Calin G.A., Nnadi S.C., Shimizu M., Davuluri R.V., Hyslop T., Demant P., Croce C.M., Siracusa L.D. 2007. MicroRNA genes are frequently located near mouse cancer susceptibility loci. Proc. Natl. Acad. Sci U S A, 104: 8017-8022

Sevignani C., Calin G.A., Siracusa L.D., Croce C.M. 2006. Mammalian microRNAs: a small world for fine-tuning gene expression. Mamm. Genome, 17: 189-202

Shafi R., Iyer S.P., Ellies L.G., O'Donnell N., Marek K.W., Chui D., Hart G.W., Marth J.D. 2000. The O-GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. Proc. Natl. Acad. Sci. U S A, 97: 5735-5739

Shen G., Li X., Jia Y.F., Piazza G.A., Xi Y. 2013. Hypoxia-regulated microRNAs in human cancer. Acta. Pharmacol. Sin., 34: 336-341

Shendure J., Ji H. 2008. Next-generation DNA sequencing. Nat. Biotechnol., 26: 1135-1145

Shibaji T., Nagao M., Ikeda N., Kanehiro H., Hisanaga M., Ko S., Fukumoto A., Nakajima Y. 2003. Prognostic significance of HIF-1 alpha overexpression in human pancreatic cancer. Anticancer Res., 23: 4721-4727

Shieh T.M., Chang K.W., Tu H.F., Shih Y.H., Ko S.Y., Chen Y.C., Liu C.J. 2010. Association between the polymorphisms in exon 12 of hypoxia-inducible factor-1alpha and the clinicopathological features of oral squamous cell carcinoma. Oral Oncol., 46: e47-53

Shohet R.V., Garcia J.A. 2007. Keeping the engine primed: HIF factors as key regulators of cardiac metabolism and angiogenesis during ischemia. J. Mol. Med. (Berl), 85: 1309-1315

Simon M.C., Keith B. 2008. The role of oxygen availability in embryonic development and stem cell function. Nat. Rev. Mol. Cell Biol., 9: 285-296

Slawson C., Copeland R.J., Hart G.W. 2010. O-GlcNAc signaling: a metabolic link between diabetes and cancer? Trends Biochem. Sci., 35: 547-555

Slawson C., Hart G.W. 2011. O-GlcNAc signalling: implications for cancer cell biology. Nat. Rev. Cancer, 11: 678-684

Sperandio S., Fortin J., Sasik R., Robitaille L., Corbeil J., de Belle I. 2009. The transcription factor Egr1 regulates the HIF-1alpha gene during hypoxia. Mol. Carcinog., 48: 38-44

Spizzo R., Nicoloso M.S., Croce C.M., Calin G.A. 2009. SnapShot: MicroRNAs in Cancer. Cell, 137: 586-586 e581

Streit S., Michalski C.W., Erkan M., Kleeff J., Friess H. 2009. Northern blot analysis for detection and quantification of RNA in pancreatic cancer cells and tissues. Nat. Protoc., 4: 37-43

Struhl K. 2007. Transcriptional noise and the fidelity of initiation by RNA polymerase II. Nat. Struct. Mol. Biol., 14: 103-105

Suzuki K., Kizaki T., Hitomi Y., Nukita M., Kimoto K., Miyazawa N., Kobayashi K., Ohnuki Y., Ohno H. 2003. Genetic variation in hypoxia-inducible factor 1alpha and its possible association with high altitude adaptation in Sherpas. Med. Hypotheses, 61: 385-389

Suzuki Y., Yamashita R., Nakai K., Sugano S. 2002. DBTSS: DataBase of human Transcriptional Start Sites and full-length cDNAs. Nucleic Acids Res., 30: 328-331

Szkandera J., Knechtel G., Stotz M., Hofmann G., Langsenlehner U., Krippl P., Langsenlehner T., Dehchamani D., Samonigg H., Renner W., Gerger A. 2010. Association of hypoxia-inducible factor 1-alpha gene polymorphisms and colorectal cancer prognosis. Anticancer Res., 30: 2393-2397

Taft R.J., Pang K.C., Mercer T.R., Dinger M., Mattick J.S. 2010. Non-coding RNAs: regulators of disease. J. Pathol., 220: 126-139

Taguchi A., Yanagisawa K., Tanaka M., Cao K., Matsuyama Y., Goto H., Takahashi T. 2008. Identification of hypoxia-inducible factor-1 alpha as a novel target for miR-17-92 microRNA cluster. Cancer Res., 68: 5540-5545

Talks K.L., Turley H., Gatter K.C., Maxwell P.H., Pugh C.W., Ratcliffe P.J., Harris A.L. 2000. The expression and distribution of the hypoxia-inducible factors HIF-1alpha and HIF-2alpha in normal human tissues, cancers, and tumor-associated macrophages. Am. J. Pathol., 157: 411-421

Tanimoto K., Yoshiga K., Eguchi H., Kaneyasu M., Ukon K., Kumazaki T., Oue N., Yasui W., Imai K., Nakachi K., Poellinger L., Nishiyama M. 2003. Hypoxia-inducible factor-1alpha polymorphisms associated with enhanced transactivation capacity, implying clinical significance. Carcinogenesis, 24: 1779-1783

Tay Y., Zhang J., Thomson A.M., Lim B., Rigoutsos I. 2008. MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. Nature, 455: 1124-1128

Taylor R.P., Parker G.J., Hazel M.W., Soesanto Y., Fuller W., Yazzie M.J., McClain D.A. 2008. Glucose deprivation stimulates O-GlcNAc modification of proteins through up-regulation of O-linked N-acetylglucosaminyltransferase. J. Biol. Chem., 283: 6050-6057

Thomas J.W., Touchman J.W., Blakesley R.W., Bouffard G.G., Beckstrom-Sternberg S.M., Margulies E.H., Blanchette M., Siepel A.C., Thomas P.J., McDowell J.C., Maskeri B., Hansen N.F., Schwartz M.S., Weber R.J., Kent W.J., Karolchik D., Bruen T.C., Bevan R., Cutler D.J., Schwartz S., Elnitski L., Idol J.R., Prasad A.B., Lee-Lin S.Q., Maduro V.V., Summers T.J., Portnoy M.E., Dietrich N.L., Akhter N., Ayele K., Benjamin B., Cariaga K., Brinkley C.P., Brooks S.Y., Granite S., Guan X., Gupta J., Haghighi P., Ho S.L., Huang M.C., Karlins E., Laric P.L., Legaspi R., Lim M.J., Maduro Q.L., Masiello C.A., Mastrian S.D., McCloskey J.C., Pearson R., Stantripop S., Tiongson E.E., Tran J.T., Tsurgeon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.S., Young A.C., Zhang L.H., Osoegawa K., Zhu B., Zhao B., Shu C.L., De Jong P.J., Lawrence C.E., Smit A.F., Chakravarti A., Haussler D., Green P., Miller W., Green E.D. 2003. Comparative analyses of multi-species sequences from targeted genomic regions. Nature, 424: 788-793

Tian H., McKnight S.L., Russell D.W. 1997. Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells. Genes Dev., 11: 72-82

Torres O., Palomino-Morales R., Vazquez-Rodriguez T.R., Gamallo C., Morado I.C., Miranda-Filloy J.A., Amigo-Diaz E., Callejas-Rubio J.L., Fernadez-Gutierrez B., Castaneda S., Martin J., Gonzalez-Gay M.A. 2010. Lack of association between hypoxia inducible factor-1 alpha gene polymorphisms and biopsy-proven giant cell arteritis. Clin. Exp. Rheumatol., 28: 40-45

Vilela B.J., Carvalho L.C., Amancio S. 2008. Integration of stress produced reactive oxygen species in the stomatal regulation of micropropagated Vitis vinifera L. plantlets impaired in ABA signaling. Plant Signal. Behav., 3: 558-559

Visel A., Prabhakar S., Akiyama J.A., Shoukry M., Lewis K.D., Holt A., Plajzer-Frick I., Afzal V., Rubin E.M., Pennacchio L.A. 2008. Ultraconservation identifies a

small subset of extremely constrained developmental enhancers. Nat. Genet., 40: 158-160

Vleugel M.M., Greijer A.E., van der Wall E., van Diest P.J. 2005. Mutation analysis of the HIF-1alpha oxygen-dependent degradation domain in invasive breast cancer. Cancer Genet. Cytogenet., 163: 168-172

Volm M., Koomagi R. 2000. Hypoxia-inducible factor (HIF-1) and its relationship to apoptosis and proliferation in lung cancer. Anticancer Res., 20: 1527-1533

Walsh G., Jefferis R. 2006. Post-translational modifications in the context of therapeutic proteins. Nat. Biotechnol., 24: 1241-1252

Wang G.L., Jiang B.H., Rue E.A., Semenza G.L. 1995. Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. Proc. Natl. Acad. Sci. U S A, 92: 5510-5514

Wang G.L., Semenza G.L. 1993. General involvement of hypoxia-inducible factor 1 in transcriptional response to hypoxia. Proc. Natl. Acad. Sci. U S A, 90: 4304-4308

Wang H.F., Feng L., Niu D.K. 2007. Relationship between mRNA stability and intron presence. Biochem. Biophys. Res. Commun., 354: 203-208

Wang K.C., Chang H.Y. 2011. Molecular mechanisms of long noncoding RNAs. Mol. Cell, 43: 904-914

Warburg O. 1956. On the origin of cancer cells. Science, 123: 309-314

Wenger R.H., Stiehl D.P., Camenisch G. 2005. Integration of oxygen signaling at the consensus HRE. Sci. STKE, 2005: re12

Wipff J., Dieude P., Avouac J., Tiev K., Hachulla E., Granel B., Diot E., Sibilia J., Mouthon L., Meyer O., Kahan A., Boileau C., Allanore Y. 2009. Association of hypoxia-inducible factor 1A (HIF1A) gene polymorphisms with systemic sclerosis in a French European Caucasian population. Scand. J. Rheumatol., 38: 291-294

Wykoff C.C., Beasley N.J., Watson P.H., Turner K.J., Pastorek J., Sibtain A., Wilson G.D., Turley H., Talks K.L., Maxwell P.H., Pugh C.W., Ratcliffe P.J., Harris A.L. 2000. Hypoxia-inducible expression of tumor-associated carbonic anhydrases. Cancer Res., 60: 7075-7083

Xie X., Lu J., Kulbokas E.J., Golub T.R., Mootha V., Lindblad-Toh K., Lander E.S., Kellis M. 2005. Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. Nature, 434: 338-345

Xu G., Wang M., Xie W., Bai X. 2011. Hypoxia-inducible factor-1 alpha C1772T gene polymorphism and glioma risk: a hospital-based case-control study from China. Genet. Test. Mol. Biomarkers, 15: 461-464

Xu J., Wang B., Xu Y., Sun L., Tian W., Shukla D., Barod R., Grillari J., Grillari-Voglauer R., Maxwell P.H., Esteban M.A. 2012. Epigenetic regulation of HIF-1alpha in renal cancer cells involves HIF-1alpha/2alpha binding to a reverse hypoxia-response element. Oncogene, 31: 1065-1072

Yamada N., Horikawa Y., Oda N., Iizuka K., Shihara N., Kishi S., Takeda J. 2005. Genetic variation in the hypoxia-inducible factor-1alpha gene is associated with type 2 diabetes in Japanese. J. Clin. Endocrinol. Metab., 90: 5841-5847

Yamagata K., Furuta H., Oda N., Kaisaki P.J., Menzel S., Cox N.J., Fajans S.S., Signorini S., Stoffel M., Bell G.I. 1996. Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). Nature, 384: 458-460

Yan H.L., Xue G., Mei Q., Wang Y.Z., Ding F.X., Liu M.F., Lu M.H., Tang Y., Yu H.Y., Sun S.H. 2009. Repression of the miR-17-92 cluster by p53 has an important function in hypoxia-induced apoptosis. EMBO J., 28: 2719-2732

Yang M.H., Wu M.Z., Chiou S.H., Chen P.M., Chang S.Y., Liu C.J., Teng S.C., Wu K.J. 2008. Direct regulation of TWIST by HIF-1alpha promotes metastasis. Nat. Cell Biol., 10: 295-305

Yu A.Y., Frid M.G., Shimoda L.A., Wiener C.M., Stenmark K., Semenza G.L. 1998. Temporal, spatial, and oxygen-regulated expression of hypoxia-inducible factor-1 in the lung. Am. J. Physiol., 275: L818-826

Yu A.Y., Shimoda L.A., Iyer N.V., Huso D.L., Sun X., McWilliams R., Beaty T., Sham J.S., Wiener C.M., Sylvester J.T., Semenza G.L. 1999. Impaired physiological responses to chronic hypoxia in mice partially deficient for hypoxia-inducible factor 1alpha. J. Clin. Invest., 103: 691-696

Yu F., White S.B., Zhao Q., Lee F.S. 2001. HIF-1alpha binding to VHL is regulated by stimulus-sensitive proline hydroxylation. Proc. Natl. Acad. Sci. U S A, 98: 9630-9635

Zagouri F., Sergentanis T.N., Gazouli M., Tsigginou A., Dimitrakakis C., Papaspyrou I., Eleutherakis-Papaiakovou E., Chrysikos D., Theodoropoulos G., Zografos G.C., Antsaklis A., Dimopoulos A.M., Papadimitriou C.A. 2012. HSP90, HSPA8, HIF-1 alpha and HSP70-2 polymorphisms in breast cancer: a case-control study. Mol. Biol. Rep., 39: 10873-10879

Zhang F.L., Shen G.M., Liu X.L., Wang F., Zhao Y.Z., Zhang J.W. 2012. Hypoxia-inducible factor 1-mediated human GATA1 induction promotes erythroid differentiation under hypoxic conditions. J. Cell. Mol. Med., 16: 1889-1899

Zhang H., Gao P., Fukuda R., Kumar G., Krishnamachary B., Zeller K.I., Dang C.V., Semenza G.L. 2007. HIF-1 inhibits mitochondrial biogenesis and cellular

respiration in VHL-deficient renal cell carcinoma by repression of C-MYC activity. Cancer Cell, 11: 407-420

Zhao C., Hamilton T. 2007. Introns regulate the rate of unstable mRNA decay. J. Biol. Chem., 282: 20230-20237

Zhao F., Vilardi A., Neely R.J., Choi J.K. 2001. Promotion of cell cycle progression by basic helix-loop-helix E2A. Mol. Cell. Biol., 21: 6346-6357

Zhao T., Lv J., Zhao J., Nzekebaloudou M. 2009. Hypoxia-inducible factor-1alpha gene polymorphisms and cancer risk: a meta-analysis. J. Exp. Clin. Cancer Res., 28: 159

Zheng Z.L., Hwang Y.H., Kim S.K., Kim S., Son M.J., Ro H., Sung S.A., Lee H.H., Chung W.K., Joo K.W., Yang J. 2009. Genetic polymorphisms of hypoxia-inducible factor-1 alpha and cardiovascular disease in hemodialysis patients. Nephron. Clin. Pract., 113: c104-111

## ACKNOWLEDGEMENTS (ZAHVALA)

This thesis is a collaboration of several great people, which I had a pleasure to work with in Slovenia and in USA. This work was supported by the Slovenian Research Agency (ARRS) and George A. Calin who covered all the expenses for the material and reagents used for experimental part of this study.

I would like to express my gratitude to my supervisor prof. dr. Tanja Kunej for her guidance and help with this dissertation as well as for her encouragement during my study. She gave me the opportunity to experience/be involved in an interesting research at home and abroad. I would like to give my special thanks to my co-supervisor, prof. dr. George A. Calin, who gave me an excellent opportunity to work with the best researcher in optimal research environment, which also enabled me to grow as a person and researcher. Besides I'm really thankful to him for all his guidance through challenges faced when dealing with noncoding RNAs and moody HIF, and for help and motivation, which he always knew when to give it. Both of my supervisors are especially deserved for the professional assistance in all publications included in this thesis.

I would also like to thank my other two committee members' prof. dr. Simon Horvat and prof. dr. Maja Čemažar for their useful comments and evaluation of this work. Additionally, to prof. Simon Horvat for his comments and ideas used in several manuscripts that we prepared togehter.

To dr. Ivan for his experts in hypoxia field that was used in each step of a manuscript preparation about hypoxia, HIF, and T-UCRs. To Mei Koh from dr. Powis laboratory for her assistance with some of the hypoxia experiments.

Thanks to all my co-workers in Domžale for their support, offered moments of relaxation, and good wishes during my work, especially to Jernej Ogorevc, Luka Bolha, Ivanka Cizelj, Irena Oven, Anja Palandačič, and Minja Zorc.

I would also like to thank my co-workers, current and former members in dr. Calin's laboratory, for taking over my non-research and research work. Special thanks to my Italian friends Milena S. Nicoloso, Riccardo Spizzo, and Simona Rossi, which showed me what it takes to be a good scientist.

Finally, I wish to express my deepest gratitude for the consistent support, understanding and love that I received from my family and friends. Thanks to my mom and dad for keeping me up, when I was down and to my friends Katarina Kovač, Polona Kovač, Urša Bizič, and Jerneja Stanišič for all the hours of support to get physical and psychical strength for the hardest obstacles on my life path.

## ANNEXES

### ANNEX A:

The most significantly overexpressed T-UCRs in colorectal cancer and their expression in hypoxia *versus* normoxia after 24 and 48 h

| Probe for T-UCR | CRC/colon mucosa | *P*-value | HCT-116 HYP/NORM (24 h) | HCT-116 HYP/NORM (48 h) | MCF-7 DMOG/DMSO (24 h) | MCF-7 DMOG/DMSO (48 h) |
|---|---|---|---|---|---|---|
| **uc.339** | 2.17 | 2.33E-09 | 1.53 | NA | 0.93 | 1.81 |
| **uc.145A** | 1.79 | 0.000253 | NA | NA | 2.09 | NA |
| **uc.112** | 1.71 | 1.88E-14 | 0.71 | 0.73 | 2.71 | 1.04 |
| **uc.398** | 1.54 | 0.000777 | 2.50 | 0.43 | 0.57 | 2.08 |
| **uc.134A** | 1.53 | 2.92E-08 | 1.06 | 2.01 | 2.69 | 2.28 |
| **uc.18** | 1.50 | 0.000536 | 2.56 | 1.85 | 2.71 | 1.04 |
| **uc.230** | 1.40 | 0.00000316 | 0.25 | 0.07 | NA | NA |
| **uc.300A** | 1.37 | 0.000208 | 1.57 | 0.14 | NA | NA |
| **uc.388** | 1.37 | 0.00000305 | 1.22 | 1.22 | 1.2 | 3.19 |
| **uc.63A** | 1.36 | 0.00305 | 0.97 | 2.51 | 2.04 | 2.11 |
| **uc.266A** | 1.36 | 0.00081 | 0.27 | 0.06 | 0.82 | NA |
| **uc.412** | 1.31 | 0.000105 | 1.30 | 1.91 | 1.33 | 1.76 |
| **uc.73A** | 1.31 | 0.000972 | 1.10 | 2.56 | 1.28 | 2.05 |
| **uc.383** | 1.31 | 0.000308 | NA | NA | NA | NA |
| **uc.362** | 1.29 | 0.000603 | NA | NA | NA | NA |
| **uc.470A** | 1.29 | 0.000253 | NA | NA | NA | NA |
| **uc.246** | 1.28 | 0.000511 | 1.03 | 1.76 | 1.9 | 2.77 |
| **uc.106** | 1.28 | 0.00137 | 1.02 | 2.07 | 1.44 | 2.51 |
| **uc.420A** | 1.27 | 0.000013 | 1.86 | 2.36 | 0.69 | 1.25 |
| **uc.341** | 1.24 | 0.000045 | 0.61 | 0.56 | 0.65 | 1.06 |
| **uc.200** | 1.24 | 0.000763 | 1.04 | 2.75 | 2.97 | 1.28 |
| **uc.475** | 1.24 | 0.00455 | 1.79 | 2.09 | 2.02 | 2.30 |
| **uc.468A** | 1.23 | 0.000198 | 0.72 | NA | 0.84 | 1.38 |
| **uc.63** | 1.22 | 0.000511 | 0.97 | 2.51 | 2.04 | 2.11 |
| **uc.182** | 1.22 | 0.000355 | 0.95 | 0.21 | 1.01 | 1.99 |
| **uc.111** | 1.21 | 0.0000404 | 1.26 | 2.45 | 1.44 | 1.85 |
| **uc.326** | 1.21 | 0.000355 | 1.22 | 0.59 | 1.27 | 2.77 |

NA, data not available; DMSO, dimethyl sulfoxide; DMOG, dimethyloxalylglycine; HYP, hypoxia; NORM, normoxia; CRC, colorectal cancer; T-UCR or uc, transcribed-ultraconserved region.

# ANNEX B:

Oligonucleotide pairs used for gene expression analysis

| Oligonucleotide ID | Sequence | Direction | Primer length (nt) | Product size (nt) | Tm (°C) |
|---|---|---|---|---|---|
| uc.18 | TTTAAGTGATGTCGGGGAGAA | Forward | 21 | 151 | 60 |
| | TTTGATGGGCTGGGATAAAG | Reverse | 20 | 151 | 60 |
| uc.63 | AGCCAAATGTCATAGTGCAT | Forward | 20 | 104 | 52 |
| | GAGATGATTAGAGGAACAGAAAG | Reverse | 23 | 104 | 51 |
| uc.73 | TACTCCATGCAGGACCCAAC | Forward | 20 | 155 | 61 |
| | GAGCATTGGTAAAGCCTTCG | Reverse | 20 | 155 | 60 |
| uc.106 | TGCAGCCTAATTGTGGACAT | Forward | 20 | 132 | 55 |
| | GATTTCCACATCCCGGAGTA | Reverse | 20 | 132 | 56 |
| uc.111 | GAAGCCCCCTACCTTTCAA | Forward | 19 | 126 | 55 |
| | CCCTGGTACTGCCTGGAAT | Reverse | 19 | 126 | 56 |
| uc.112 | TGACAGCGAAACACACCTGTA | Forward | 21 | 134 | 56 |
| | CAACACTCTGTGCCCCATTA | Reverse | 20 | 134 | 56 |
| uc.134 | GAGCACAAAGCATTAGTTTGGA | Forward | 22 | 131 | 56 |
| | CCTTCCCCATTTTAGTGTTCA | Reverse | 21 | 131 | 55 |
| uc.145 | GCGTGGATGGGAATTTCTA | Forward | 19 | 102 | 55 |
| | TCTCGGTTATTTCACAGGCA | Reverse | 20 | 102 | 55 |
| uc.182 | ACAGCTGGCACCTGCTACT | Forward | 19 | 85 | 55 |
| | CCCTTTAAGCCATTCTTGAGT | Reverse | 21 | 85 | 54 |
| uc.200 | CTGCCTCGCCATACTAATGA | Forward | 20 | 105 | 55 |
| | CCTAGAAAGGACCTGTCCTGA | Reverse | 21 | 105 | 55 |
| uc.230 | CTCCTGGCAAATATGGGACT | Forward | 20 | 77 | 55 |
| | TGAGAAGCAGGTGCTGGAGT | Reverse | 20 | 77 | 57 |
| uc.246 | AACAAAAGGGCGAGACAAGA | Forward | 20 | 165 | 60 |
| | TCTTTACAGGCGGGAAGATG | Reverse | 20 | 165 | 60 |
| uc.266 | TGGAACGCTCATCTGGCT | Forward | 18 | 107 | 56 |
| | CTTCTGCCAGCCTCGTTT | Reverse | 18 | 107 | 55 |
| uc.300 | TCTGGAAAGAATGGAAGGGA | Forward | 20 | 106 | 56 |
| | CTTTCACCTTGCCAAGAGCT | Reverse | 20 | 106 | 56 |
| uc.326 | TTGTCCTGACACCGTTTGAA | Forward | 20 | 100 | 56 |
| | TACTGGTCGGCAGAGACTTTC | Reverse | 21 | 100 | 56 |
| uc.339 | CGACGCCATTACAGGAATG | Forward | 19 | 90 | 56 |
| | AAGATGGATTGGCCGCAT | Reverse | 18 | 90 | 57 |
| uc.341 | GGAAATTGGCTGACAGCAA | Forward | 19 | 79 | 56 |
| | CTTTCTCCAATTCCAGCGTC | Reverse | 20 | 79 | 56 |
| uc.362 | GTGTGCTGCCTTTGAGCA | Forward | 18 | 111 | 55 |
| | CCTGGTTGCCATAGCATCTA | Reverse | 20 | 111 | 55 |
| uc.383 | TCTCTCTGCGGCCGTTTA | Forward | 18 | 99 | 57 |
| | TGCCCTAAACAATGTGCATC | Reverse | 20 | 99 | 56 |
| uc.388 | ATGAAGTCATCAGGCATGGA | Forward | 20 | 108 | 55 |
| | TCACAGTCTATCGGGCACAA | Reverse | 20 | 108 | 56 |
| uc.398 | GCTGCACTGCAGAGACACAT | Forward | 20 | 83 | 56 |

*(Continued)*

*(Continued)*

| Oligonucleotide ID | Sequence | Direction | Primer length (nt) | Product size (nt) | Tm (°C) |
|---|---|---|---|---|---|
| | ACTCCGCTCCTCACATCCT | Reverse | 19 | 83 | 56 |
| uc.412 | GAAGACAGTGGAATCGGTGTT | Forward | 21 | 106 | 55 |
| | AGCAAATCCTTTGCTTTGCT | Reverse | 20 | 106 | 56 |
| uc.420 | CCACATCTGTAAGGTGTTGCA | Forward | 21 | 66 | 56 |
| | GCTGTTGTGGTGTGCAAAACT | Reverse | 21 | 66 | 57 |
| uc.468 | TAGCAGCGGCTCTGCATTA | Forward | 19 | 110 | 57 |
| | GCTTTGCTCAACAACACCATT | Reverse | 21 | 110 | 57 |
| uc.470 | TGGAGACTTCGCTGAAGTCA | Forward | 20 | 66 | 56 |
| | TCATATCCCACGCTTCTCTG | Reverse | 20 | 66 | 55 |
| uc.475 | CAGGTGGGTGAAACCTCCTA | Forward | 20 | 100 | 56 |
| | TGGAAGAGGAGGAGGAGGTT | Reverse | 20 | 100 | 56 |
| OGT 1 | TTAGCTGAGTTGGCACATCG | Forward | 20 | 363 | 64 |
| OGT 1 | TTGTACTGCCCCTTCCATGT | Reverse | 20 | 363 | 64.3 |
| OGT 1/2 | TGAGTTCCGACTTTGGGAAT | Forward | 20 | 285 | 63.4 |
| OGT 1/2 | GGAAGTTTGTGCCATCGTCT | Reverse | 20 | 285 | 64 |
| OGT-intron 12 | GCACTTCATAAGACGGAATCC | Forward | 21 | 87 | 58.7 |
| OGT-intron 12 | ATTATGGCTTTCAACTTAAAG | Reverse | 21 | 87 | 51 |
| **Oligonucleotides for northern blot** | | | | | |
| uc.475 probe | TTACAGACTTGAGACTGTCCTCATT | Forward | 25 | 365 | 58.6 |
| | AAGCTGTTTCTGAAGCCACAA | Reverse | 21 | 365 | 60 |
| OGT probe | TTAGCTGAGTTGGCACATCG | Forward | 20 | 363 | 60 |
| | TTGTACTGCCCCTTCCATGT | Reverse | 20 | 363 | 60.4 |
| PPIA probe | CAGACAAGGTCCCAAAGACAG | Forward | 21 | 298 | 59.7 |
| | TTGCCATCCAACCACTCAGTC | Reverse | 21 | 298 | 63.4 |
| **Oligonucleotides for reference genes** | | | | | |
| GAPDH | ACCCAGAAGACTGTGGATGG | Forward | 20 | 200 | 60.0 |
| | TCTAGACGGCAGGTCAGGTC | Reverse | 20 | 200 | 60.4 |
| PPIA | CAGACAAGGTCCCAAAGACAG | Forward | 21 | 298 | 59.7 |
| | TTGCCATCCAACCACTCAGTC | Reverse | 21 | 298 | 63.4 |
| RPLP0 | CCCAATTGTCCCCTTACCTT | Forward | 20 | 225 | 60 |
| | ACCCAGCTCTGGAGAAGTCA | Reverse | 20 | 225 | 60 |

Nt, nucleotide.

# ANNEX C:

Small interfering RNA (siRNA) sequences used for silencing assays

| siRNA | Target name | Orientation of oligoNT | OligoNT sequence | Start position within target | Target sequence |
|---|---|---|---|---|---|
| 1 | uc.475+ | S | GACUGUCCUCAUUCCCAUAdTdT | 13 | GACTGTCCTCATTCCCATA |
| | uc.475+ | AS | UAUGGGAAUGAGGACAGUCdTdT | 13 | TATGGGAATGAGGACAGTC |
| 2 | uc.475+ | S | GCUUCAGAAACAGCUUAAAdTdT | 350 | GCTTCAGAAACAGCTTAAA |
| | uc.475+ | AS | UUUAAGCUGUUUCUGAAGCdTdT | 350 | TTTAAGCTGTTTCTGAAGC |
| 3 | uc.475+ | S | GGUGAAACCUCCUAUUCCAdTdT | 153 | GGTGAAACCTCCTATTCCA |
| | uc.475+ | AS | UGGAAUAGGAGGUUUCACCdTdT | 153 | TGGAATAGGAGGTTTCACC |
| 1 | uc.63+A | S | GCACUAUGACAUUUGGCUUdTdT | 206 | GCACTATGACATTTGGCTT |
| | uc.63+A | AS | AAGCCAAAUGUCAUAGUGCdTdT | 206 | AAGCCAAATGTCATAGTGC |
| 1 | OGT | S | GGAAUAUCCGUGAAGGGAUdTdT | 4727 | GGAATATCCGTGAAGGGAT |
| | OGT | AS | AUCCCUUCACGGAUAUUCCdTdT | 4727 | ATCCCTTCACGGATATTCC |
| 2 | OGT | S | CAAAGCAUUUCUUGAUAGUdTdT | 2158 | CAAAGCATTTCTTGATAGT |
| | OGT | AS | ACUAUCAAGAAAUGCUUUGdTdT | 2158 | ACTATCAAGAAATGCTTTG |
| 3 | OGT | S | GAAGAAAGUUCGUGGCAAAdTdT | 2905 | GAAGAAAGTTCGTGGCAAA |
| | OGT | AS | UUUGCCACGAACUUUCUUCdTdT | 2905 | TTTGCCACGAACTTTCTTC |
| 4 | OGT pre-mRNA | S | GCACUUCAUAAGACGGAAUdTdT | intron 12 | GCACTTCATAAGACGGAA |
| | OGT pre-mRNA | AS | AUUCCGUCUUAUGAAGUGCdTdT | intron 12 | ATTCCGTCTTATGAAGTGC |

OligoNT, oligonucleotide; S, sense, AS, antisense.

# ANNEX D:

Predicted HIF1A-binding sites for selected T-UCRs

| **Region of prediction** | **Genome location** | | | |
|---|---|---|---|---|
| Type of weight matrices | Weight | Genomic location * | Position sequence | Orientation |
| uc.63_Promoter.txt | chr2:61616997-61623173 | | | |
| Q3 | 7.05 | chr2:61617762-61617775 | ggccacgtgatgca | + |
| Q3 | 5.73 | chr2:61618994-61619007 | ctgggcctccgccc | + |
| Q3 | 6.60 | chr2:61622957-61622970 | ctgggcgtgatggt | + |
| Q3 | 5.94 | chr2:61620076-61620089 | ggcaacgtgcatgt | - |
| Q3 | 8.65 | chr2:61617333-61617346 | gagtgcgtgctcct | - |
| Q5 | 8.50 | chr2:61618122-61618135 | ccggcgtggggc | - |
| Q6 | 8.15 | chr2:61617951-61617964 | cgcgtgtgc | - |
| uc.106_Promoter.txt | chr2:174743290-174747867 | | | |
| Q6 | 7.02 | chr2:174746352-174746360 | ggcgtgagg | + |
| uc.73_Promoter.txt | chr6:99378079-99392321 | | | |
| Q3 | 5.37 | chr6:99379409-99379422 | cggtggctgcgcgg | + |
| Q3 | 5.47 | chr6:99383930-99383943 | gtgtgcgtgtgcgc | + |
| Q3 | 6.41 | chr6:99383932-99383945 | gtgcgtgtgcgcgc | + |
| Q3 | 7.04 | chr6:99386614-99386627 | gtacacgtgcagcc | + |
| Q3 | 5.58 | chr6:99387154-99387167 | ctgagtctgcgcgc | + |
| Q3 | 9.79 | chr6:99387252-99387265 | ccgcgcttgcgctg | + |
| Q3 | 8.03 | chr6:99387308-99387321 | gcgcgggtgcgcgc | + |
| Q3 | 7.36 | chr6:99387316-99387329 | gcgcgcgtggtggg | + |
| Q3 | 5.67 | chr6:99390524-99390537 | ttccgcttgcgccc | - |
| Q3 | 6.54 | chr6:99390220-99390233 | gcgcgcctgggtgg | - |
| Q3 | 6.35 | chr6:99389124-99389137 | acgcgcctgggccc | - |
| Q3 | 5.51 | chr6:99387763-99387776 | agcggcgtgaccgc | - |
| Q3 | 5.55 | chr6:99387252-99387265 | cagcgcaagcgcgg | - |
| Q3 | 7.06 | chr6:99387046-99387059 | aggcgcgtgacttt | - |
| Q3 | 5.33 | chr6:99385946-99385959 | gttgacgtggagag | - |
| Q3 | 5.32 | chr6:99383946-99383959 | gcgcgcgtgtgtgt | - |
| Q5 | 8.25 | chr6:99380305-99380316 | tcggcgtgggcg | + |
| Q5 | 9.33 | chr6:99382048-99382059 | gccgcgtgcggg | + |
| Q5 | 7.10 | chr6:99386615-99386626 | tacacgtgcagc | + |
| Q5 | 8.58 | chr6:99390336-99390347 | cggacgtggggc | + |
| Q5 | 7.57 | chr6:99388240-99388253 | cctgcgtgggct | - |
| Q5 | 9.94 | chr6:99388130-99388143 | cctgcgtgcgcg | - |
| Q5 | 7.87 | chr6:99387762-99387775 | gcggcgtgaccg | - |
| Q5 | 7.25 | chr6:99387045-99387058 | ggcgcgtgactt | - |
| Q5 | 9.27 | chr6:99382700-99382713 | cccgcgtggggg | - |
| Q5 | 8.60 | chr6:99379726-99379739 | caggcgtgcgtg | - |
| Q5 | 9.61 | chr6:99379722-99379735 | cgtgcgtgagcg | - |

*(Continued)*

*(Continued)*

| Region of prediction | Genome location | | | |
|---|---|---|---|---|
| Q6 | 8.22 | chr6:99382050-99382058 | cgcgtgcgg | + |
| Q6 | 7.16 | chr6:99383933-99383941 | tgcgtgtgc | + |
| Q6 | 7.62 | chr6:99389133-99389141 | cgcgtgccg | + |
| Q6 | 6.43 | chr6:99390251-99390264 | ggcgtgtcc | - |
| Q6 | 8.79 | chr6:99388128-99388141 | tgcgtgcgc | - |
| Q6 | 7.98 | chr6:99387760-99387773 | ggcgtgacc | - |
| Q6 | 7.62 | chr6:99387407-99387420 | cacgtgtcc | - |
| Q6 | 8.71 | chr6:99379720-99379733 | tgcgtgagc | - |
| uc.475_Promoter.txt | chrX:70668972-70670441 | | | |
| Q3 | 5.90 | chrX:70669486-70669499 | gctcgcatgcgcgg | + |
| Q3 | 5.76 | chrX:70669486-70669499 | ccgcgcatgcgagc | - |
| Q5 | 9.35 | chrX:70670194-70670205 | cacacgtgcgcg | + |
| Q6 | 9.85 | chrX:70670196-70670204 | cacgtgcgc | + |
| uc.134_Promoter.txt | chr3:159306900-159312000 | | | |
| Q3 | 10.73 | chr3:159310839-159310852 | cggggcgtgcgcag | + |
| Q3 | 8.21 | chr3:159310860-159310873 | gtgcgcatgcgtcc | + |
| Q3 | 6.85 | chr3:159310860-159310873 | ggacgcatgcgcac | - |
| Q3 | 6.08 | chr3:159310662-159310675 | gttgacgtgcaagc | - |
| Q5 | 7.72 | chr3:159310999-159311012 | ggaacgtggggc | - |
| Q6 | 8.66 | chr3:159310842-159310850 | ggcgtgcgc | + |

*Location data are according to the UCSC genome browser (assembly NCBI36/hg18). Chr, chromosome; Position weight matrices for HIF1: Q3, V$HIF1_Q3 (gnnkACGTGcggnn); Q5, V$HIF1_Q5 (ngtACGTGcngb); Q6, V$HIF1_Q6 (nRCGTGngn); +, sense; -, antisense.

# ANNEX E:

*HIF1A* sequence with highlighted variations obtained from Ensembl Genome Browser, including 30 SNPs from human association studies

| # | SNP ID [substitution] | Genomic location | SNP type |
|---|---|---|---|
| 1 | rs2783778 [C>T] | 5'-flanking | |
| 2 | rs7148720 [T>C] | | |
| 3 | rs1535679 [A>C] | | |
| 4 | rs28708675 [A>T] | | |
| 5 | rs41362550 [T>C] | | |
| 6 | rs7143164 [G>C] | intron 1 | |
| 7 | rs1951795 [A>C] | | |
| 8 | rs12435848 [A>G | | |
| 9 | c.111C>A *p.Ser28Tyr* | exon 2 | non-synonymous |
| 10 | c.347T>A *p.Val116Glu* | exon 3 | non-synonymous |
| 11 | rs4899056 [T>C] | intron 4 | |
| 12 | rs1957757 [T>C] | intron 6 | |
| 13 | rs12434438 [G>A] | | |
| 14 | rs11158358 [G>C] | | |
| 15 | rs2301111 [G>C] | intron 7 | |
| 16 | rs10873142 [C>T] | intron 8 | |
| 17 | rs41508050 [C>T] *p.Thr418Ile* | exon 10 | non-synonymous |
| 18 | rs2301113 [C>A] | intron 10 | |
| 19 | rs41492849 [C>T] *p.Pro564=* | exon 12 | synonymous |
| 20 | rs34005929 [G>A] *p.Leu580=* | | synonymous |
| 21 | rs11549465 [C>T] *p.Pro582Ser* | | non-synonymous |
| 22 | rs11549467 [G>A] *p.Ala588Thr* | | non-synonymous |
| 23 | rs199775054 [G>C] *p.Ala593Pro* | | non-synonymous |
| 24 | rs61755645 [A>T] *p.Thr600=* | | synonymous |
| 25 | rs4902080 [T>C] | intron 12 | |
| 26 | rs10645014 [insGT] | intron 13 | |
| 27 | rs4902082 [C>T] | intron 14 | |
| 28 | rs2057482 [T>C] | 3'-UTR | |
| 29 | rs1319462 [G>A] | 3'-flanking | |
| 30 | rs17099207 [G>A] | 3'-flanking of *HIF1A*;<br>5'-UTR of *SNAPC1* | |

## Transcript: HIF1A-001 ENST00000337138

| | |
|---|---|
| **Description** | hypoxia inducible factor 1, alpha subunit (basic helix-loop-helix transcription factor) [S |
| **Location** | Chromosome 14: 62,162,258-62,214,976 forward strand. |
| **Gene** | This transcript is a product of gene ENSG00000100644 |

This gene has 13 transcripts (splice variants) Hide transcript table

Show All entries Show/hide columns Filter

| Name | Transcript ID | Length (bp) | Protein ID | Length (aa) | Biotype | CCDS |
|---|---|---|---|---|---|---|
| HIF1A-001 | ENST00000337138 | 3919 | ENSP00000338018 | 826 | Protein coding | CCDS9753 |
| HIF1A-002 | ENST00000323441 | 3555 | ENSP00000323326 | 735 | Protein coding | CCDS9754 |
| HIF1A-201 | ENST00000539097 | 3468 | ENSP00000437955 | 850 | Protein coding | CCDS58324 |
| HIF1A-003 | ENST00000394997 | 3922 | ENSP00000378446 | 827 | Protein coding | - |
| HIF1A-004 | ENST00000557538 | 3468 | ENSP00000451696 | 767 | Protein coding | - |
| HIF1A-012 | ENST00000557206 | 767 | No protein product | - | Processed transcript | - |
| HIF1A-011 | ENST00000555014 | 739 | No protein product | - | Processed transcript | - |
| HIF1A-005 | ENST00000553999 | 616 | No protein product | - | Processed transcript | - |
| HIF1A-007 | ENST00000556237 | 595 | No protein product | - | Processed transcript | - |
| HIF1A-006 | ENST00000554177 | 561 | No protein product | - | Processed transcript | - |
| HIF1A-008 | ENST00000557446 | 559 | No protein product | - | Processed transcript | - |
| HIF1A-009 | ENST00000556827 | 752 | No protein product | - | Retained intron | - |
| HIF1A-010 | ENST00000547430 | 711 | No protein product | - | Retained intron | - |

Variations

| | | |
|---|---|---|
| 3 prime UTR | 5 prime UTR | Coding sequence |
| Downstream | Frameshift | Intronic |
| Missense | Splice donor | Splice region |
| Stop gained | Synonymous | Upstream |

| **HIF1A-001**UCSC >hg19_dna range=chr14:62133556-62133956 5'pad=200 3'pad=200 | ***HIF1A* SNPs** |
|---|---|
| acaccattgcactccagcctgggtgacaagagcaaaactccgtctcaaaaaaaaaaagaaaaagaaaaagaaaaaaaatgcaaattaaaacca cactccagcctgggcaacagagtgagactccatctcaaaaaacagaaaaaagaaataaaacactgtttctgcacttaaggacttctcagactggta ctgggcacagtcgggtaaggaggcagttgtcctccagggaagtaagtgccatggtagaaactccagctatctatgtgttgataacgctcagcctat aactttgtctcacatactcaggagctttctgggtacctccacttggatgttcctcagaaaatcagaaatcaccctgaacatttaaaagagaaagtgta atgctgcacagggaagac..... | 1.) rs2783778 |
| UCSC >hg19_dna range=chr14:62142589-62142989 5'pad=200 3'pad=200 | |
| gcctgagtgccaaacacattctttactgcccctactgtggagtagtagccctacattcttctgacgatcaggaacaattacttttttgccaggatagtga ttccttttcttgcttgctagtctcttgtaaaatagagcccaaagtgaccaggtaacagctggagctcaaaatgcaatacttttttttttaaattttattaag ttccgggatacatgtgcaagtttgttacataggtaaacatatgccatggtggtttgatgcacctatcaacccatcacctaggtattaagccctgcatg cattagctatttatcctgatgctctacctcaccctgccactacccctcaccccaacaagccccagcatgtcttgtgcccctccctgtgtccatgtgttctc at..... | 2.) rs7148720 |
| Ensembl 5'upstream sequence (2500 bp) | |
| .....tgaatattcaaaaatttttttgctgatcaaattatcattctgctttaaacttttgaaatgatccaaaaaaatttttaaatgacttagat ttactgttacaaaatgcttgtcttttgatgtcacaaacattatatactataatcactggccagagataattgctattaagtataatgaaaagggaaatg atggaagatctctgcagctatcctcataaatgaagggtgggaacacgatgggcagttccaaagttgaaaatagagaatatatgttgtggatttatattaa cataattggtattcttggatagttaaaaatggctaaactgtaggagaagcccgagtaattactgttaacagaggaataaatttgagggcaataataa tgatgataggccaggcactgtggctcatgcctgtaatcccagcactttgggaacccgaggcgagcggaccacctgaggtcaggagttcgagagcag cctggccaacatggtgaaacctcgtctctactaaaaatagaaaaattatccgagtgtggtggtgcgtgcctgtaatcccagctacttgggaggctga ggcaggagaatcacttgtacctgggaggcggagttgcagtgagccgaaatcgcgccactgcgctccagcctgtgggccagagcgagactccgcctc agaataataataatgataataataataacgccaccaacaatactaagagctaacatttactgagtgcttactatgcaccagatattgttctaagtata catttattatctcatttaaccatccataatactgtggtatagacacttttatatccattttataaataagtaaactgagttatggagagattaaacgact tgccagtaagattcaaagcctgtgtacaagctcacgcttgattctggagccagtgttcttaacacagtatcttgagaatgttaaactaaaaagttttta atttacagtattctttccacaattaaaaaagaaattatgagtaattatttttagttctttcttctcttcaggcatttcccatggttctttttcaagacataat acatatcatttagtgttgtagatctgaaaaaacaaaagtagcgtgaagatcaaaaattttctaaagagacggagtctcgctacgttccctaggctgg | 3.) rs1535679<br>4.) rs28708675 |

*(Continued)*

*(Continued)*

| | |
|---|---|
| aacacccaggcttctccagcctcacacctctgagtagctggaaccaccctgtccgctaaggtcaatgtttaatcgtatctttgtaggtctactgaccag<br>ttaaaaagaggtgctgtatacattggttgttgtcttgtcagagtttgatgcttctatatagaccattgtttttacatgctaatacaattgaaagccactac<br>agatatttatatttacaacccaaagctaggttttaacaagaaactcataaggcaaaggtgagaagtaaaataatttagcgccaagtggagatatat<br>gtgcaatgctactttgttgggctcaaaacatatttttctttttagaagactgacaggcttgaagtttatgcctccaaagacaaaagtgattatgttttgttt<br>agtagcttgcaaagttgccaaaggccatttttttctactctttccctgaaattggtttatatgcttattaaagtcatttatacctatttgcaaatgcttaaca<br>tagtttcagattttaagatttccctgcaactttatttcccttgaagtttacagcaacaggagttcatttttatttttaattgcatttattcagtaagtaaact<br>ccgccacagaaaaacttagtagacaaggtgagttcccctgtgctccgtggcaaagagtgcggtgggtgacattgacccatggttaggtaatctggta<br>aggaaagaccccgttgtaacacatctgagcaacgagaccaaaggaagggcttgctgccacgaggcgaagtctgctttttgaacagagagcccag<br>cagagttgggcggcaatcgtgcccagcactgaggccgaggagaaagagagcaggagcattacattactgcaccaagagtaggaaaatatgatgc<br>atgtttgggaccaggcaaccgaaatcccttctcagcagcgcctcccaaagccgggcaccgccttccttcggagaaggcgcagagtccccagactcg<br>ggctgagccgcacccccatctccttttctctttcctccgccgctaaacacagacgagcacgtgagcgtccgcagcccgtcccagctgtgcctcagctga<br>ccgcctcctgattggctgagagcggcgtgggctggggtggggacttgccgcctgcgtcgctcgccattggatctcgaggaacccgcctccacctcag<br>gtgaggcgggcttgcgggagcgcgcgccggcctgggcaggcgagcgggcgcgctcccgccccctctccctccccgcgcgcccgagcgcgcctccg<br>cccttgcccgccccctgacgctgcctcagctcctcagtgcacagtgctgcctcgtctgaggg | 5.) rs41362550 |
| ENSE00001520256 62,162,258-62,162,557 (300 bp) | |
| GACAGGAGGATCACCCTCTTCGTCGCTTCGGCCAGTGTGTCGGGCTGGGCCCTGACAAGCCACCTGAGGAGAG<br>GCTCGGAGCCGGGCCCGGACCCCGGCGATTGCCGCCCGCTTCTCTCTAGTCTCACGAGGGGTTTCCCGCCTCGCA<br>CCCCCACCTCTGGACTTGCCTTTCCTTCTCTTCTCCGCGTGTGGAGGGAGCCAGCGCTTAGGCCGGAGCGAGCCT<br>GGGGGCCGCCCGCCGTGAAGACATCGCGGGGACCGATTCACCATGGAGGGCGCCGGCGGCGCGAACGACAAG<br>AAAAA | |
| Intron1-2 62,162,558-62,187,099 (24,542 bp) | |
| gtaagcccattccctcggcccgccgccttctgccccggcgaccccgcccgcctgcccgccctgggctcctgggccggcctcggcgttaatgggattggg<br>ggggggcagccttttgtttctgctgctgctcccctccctctcttcccccaacctcgccggccgggctcccccgctgtccacgtcgccatcttgtcgtgg<br>ggggtggagacgcctcgaaagtgcttctcaggggccggggtctgagccctgcttgccctcccgccggccgtgggggcctcgcgccgcccacctacc<br>cgcctcaaaaacccagcctgctctgtggccccatccggaggggacttttacccagcctgaaaaccccgggaagagaaatgagctgcagctcggtag<br>ccgcggttgcacccggagcttccgctccttcccgcccccatcctctccagttccattgaaaactcggccctggggcggaccctgcacgctggtcctgg<br>ctttgcagtggacttggggccttgagttcccgactgagggactcgcgtggtcggatgcgatgttgtcctgtagttgtccagccgtcgcgggtgtctttgc<br>ctttgtgcattagggatttgccgcgatggccttaagatgcgaacttttagtttgcacgtgcaggttttgttttcgttttaatcgccttgaaaaacttgcct<br>agactgagagtcagagtaatgggaatttagggaaatggcaacattttaaagagaacttcagaattggatacttgagttcatatcacctgtcacgag<br>aacgcagatattataaatgaatatatgcctcattcattcttcaaataatgaaaatgtaggggctggttaaatttaggcagttttaatgatactgaaaa<br>aagtatatgatgagtgaatgaaatgcggcactaaaatgttgcaaaaattttcgaactctgtctcattttcctgaaattgaagtatattaaaggaaaac<br>cgtcaacatatatctaaagtaagtaatcactcggttagaacttaatgcaagttttataaatcaccttgaagtttgagtctaaggggtacattagagatt<br>aagaattgtgagttggaccagtggtgttaagagcggactccccccatccccccaacacacacacaattttgcccacttttggcattttaacttttaaggaa<br>atcacttaaggaattgaagatttagagtaagagtttttggttagtagactggctttgctgttaaatccttccactcttctggcagagagattaatttccct<br>aatcagtatcagcagaagataaacttgtttatattcctgctgttttgtagatcccttctcctggtccttcttcaatagaatattaaattcttagtttgtata<br>cagcagagaaggtcacttataaaattcaaaaagtgagcaaacaggtctagattaattccaagagttaccaggaattaattgcagtttattttgcgga<br>ggtgattacagtgcttttgatgaaatgataaagctgctatattgtaaacctaaggcagattacctctgtgtagtgccagttttctatccttattatatatt<br>gaatcatacttaatacaatgcattaaattatgtaccacttttttatatacagtatcgaactcattgttttgccattcatccgttcagaatatcagaagca<br>gttttgaaagaattaataaattagctactgttcatcagccccaattctaaataagctcttagattttcctcagcccatctgttactttcaaaatttctc<br>atttgaaaacttggcaaccttggattggatggattcatatttcttagtatagaagttcttgatataactgaaaaattaagttaaacacttaataagtgg<br>tggttactcagcacttttagatgctgtttataatagatgaccttttctaactaatttacagtttttgaaagataactgagaggttgagggacggagatt<br>ttcttcaagcaattttttttttcattttaaatgagctcccaatgtcggagtttggaaaacaaatttgtcgttttaaaagaaggtctaggaaactcaaaac<br>ctgaagaattggaagaaatcagaatagaaaatgggtatggttatgatactgtagatttaacgcaggacatttcatgttgttcctagttataggggctg<br>aacttatttaatagcacgtgcattttgatttttagatttttaagggaatgtcaagagagtaatgattctgtttcaggcttcaggccagactccttcagag<br>ttttccaaaacaaataattactgaatcattaaagtaaaatttctgagaatagatattccttaatttccttcattaactttggccattaaaagtcaagaa<br>gctctctcatttattagcaaacttttctccttatgattctattttgattgttcctttgtttgaggaagcagcatatggtggttaagagcataggatctagag<br>gcagatacctctgagttaagggtcccagcccttcacttgtgagcttgagcaagttactgaatgcctctgagcctctttcctcctttgaaatgatgataa<br>gaatagcagccatctgagcagttattgtaaaggttaaatgagataatgcttgtgaagcacttagcccattgcaggagtcttgatgacactgtgtactt<br>gaaaatagatgttacctgttaaaattcttgtttaaacttccacaactcttaaaactctttttgctagtccttccagctgtttccttagtttcttttctgtgt<br>cttcatgcatcttttctatctcctgaaagtgaaaagactaacattggatccagagcttgaaaagcgttttttttcctgttacaatgggcaaaagagtaca<br>tccttgggttatattggcacctagtatcagttatttttcttgagcatctgatctgctctctactctagtggaggcctcctgcttcacaattgctcacccctg<br>tgttttctccccaaatagaatactgagtttactctggactctagagtcaaacatacacagtattctagtcttactgttcatttaagcagagatatgtgcaa | |

*(Continued)*

*(Continued)*

| | |
|---|---|
| gacactgcattcttagtactggcagtaagttaaaacatttttcgtcttgatgccaaagtttagacaatttataaaaattaacctttgtaaaagataat<br>gagttgataaaatattctcagtaaagcagctaggtggtagaaaaactgtcctttgcttatgagtttctccagagttaagaccattgggttccatctgaa<br>ggcaagacttcaagcttgtcttactggtctgttttgtggctcaatttgtatgaagtctatgcactcttccacacgtgtgtatttactgaactatcgagtta<br>ttttagactgagaaagtattggagttcattcctacggtccactgcagagcaccttgtgcagtttggagaatgtcaacttttctacctgttaacttccatt<br>gtctttacttttaacgccattgtctgtgactctaatggtgtcacggctcagggtttagattttgtggttacattctattcttgtatgtcaagagtggtgtat<br>agaaagctgaggggattatttagtctcttgactgatttttttttttttctgaagaactcagtttattatgtttggtggtgaaataaaaattgatgtgcat<br>ggatgttaaagatttgggttaaattgtgtgttcatagatgccttctcttagtatataatttttaaatttagatacttaaaatactgtatccctttatctaa<br>gattaacataagtctgtttcttaaccaggataaaaaaatctaaatttaaatgtgatgttggatgagtttccaatcaagaaattgatttttaaacttttgt<br>gactagttatccagtgggtggattttacccagtgtgtgtatgtttttctgcttaactctggaaggttagaaagagaatttgaaactaagacaagcca<br>agcttcttgttgctcagtatttttggtaaaaatatggtcagattgtttaaattaactataggctttggaattttaaaaataattatatctcttggtctcttg<br>acacatcaagaattaactgttttgtatatgcgttgagtattaatgttcatgttttctgcagtagaaatttataaacccttatttatttgccagacatgatc<br>cctttagaggaaatctagtatctaaaacctgaatttttaaaacaaaatttaaaatttttgtttcataaaaacaaaaatgtgattacctcatggctttttc<br>ttatagcttttgattgttttttaaaatcgtagttcaaaaacattaacctaaaattttaccatcttaaccatttctaagtactgttcagtagtgttaagtatat<br>tcacattgtgccactaacttccagaacttttttcatcttgcaaagctgaaatcttacccattaaacaactcccaatttccccctctcctcagcctctggca<br>accaccatttttactttctgtttctgcaaatttaactactctagatgcctcatataaatagaattatagggttttaatatttttgtgatgggcttatttcactt<br>tgtgtaatgtcctcaaggttcatccatgttgtagcatgtgtcagaatttccttccttttttgagtctgagcaatattccattatatgttgcatattttgtttctc<br>cattcatccagcaacggacacttgggttgcttccacatcttggttattgtgctgctctgaacatgagtctgcaaatctctctttgaagctttcacttttttt<br>ggatacatatccagaagagggatgctggatcatatggtaacccttttaattttcaaggaaccaccatattgttttctatagcagttgcaccagtttac<br>attcccaccaacagtgcacaagggttcctatttctccacatccttctaaacacttgttttcttttcttccttccttctcttctctttcttcttagccatctaat<br>gtggcaaagtggtagccatctaatatgttgaagtgattgtttttaagggcttgtttgtggataattaaccagctgaaagctaactacagtttgccagtg<br>gaagctttaactgaaaggagagtaagtacctctaaaaggagaattcaatttttctagtgacttagatttgttatgccagtacttttttcacagaaacact<br>ttttgggtaaaatagtgttacacctgttctattgttgataaagcccaatttaattaggaaatttgttctctaagatttaaaacaataattgaaataatgta<br>tttttattaaaaactgttcccaagatgttagcttttagctgttctggtgatctcaactgttatttatgagtgtttctttatttttaaaatttcaccttaacggt<br>tacagttttaaccataaagattatttcaacatatgattttgaaaattttattatcttgtaaatgggaaaatgtagtgatggaacatagtttactgtatgta<br>gttcttcacttgtttgaaaagtcacaatatatttaggcaaattaatttaaaagtgtctagtatttaatattgcaattttcactcattaaggacaggtcccc<br>cgtgtttccccctttttttttttccaagtagtttgggaggatttgtttttccagctgaaaaatactatggttaaaaataaggtttaaaggcgaaagttgaa<br>gtctttgagggttgggatacgtttctgttcttaagagtcttgtaaattcagatgctaagcaaatttctttaaaatgatttctaccctccccctttccattat<br>aaaactggatatgtttcagtggaccaaatgccaagtaggctgaatttgaaatttgtgggctgggcgcggtggctcatgcttgtaatcccagtactttg<br>ggatgccgaggtgggtggatcacctgaggtcaggagttcgagaccagcctggccaacatggtgaaaccccatctctactaaaaatacaaaattag<br>ccaggcgtggtggcgggtgcctgtaatcccagctacttaggaggctgaggcaggagaattgcttgaacctgggaggcggaggttgcggtgagcca<br>agatcgccccattgcactccagcctgggtgacagagcaagactgtgtttcaaaaaaattaaaaaagaaatctgtggtgtgaatactggtacgtggt<br>gtacacagtgagctgttaataagtatttgaattaacaaatgagacaatgattgaataattggatgaacaaagagaatgcaggtttttaaaaggtttc<br>tttagaaatattgtcggcccggcacggtggctcctgcctgtaatcccaccattttgggaggccggggcaggtgaatcacctgaggtcaggagttcaa<br>gacaagcctgaccaacttggagaaaccccgtctctactaaaaatacaaaaaaaaaaaaaaaaaaatagcaggatgtggtggcacatgcctgtaa<br>tcccagctactcggaggctgaggcaggagaatcgcttgaacctgggaagcagaggttgcagtgagccaagatcgcgccactgcactccagcctgat<br>gacagtgtgagatgctgtctccaaaaaaaaaaaaaaaaaaaattaaaaagaatgttttaattctttagttccctgtctgagattcactgattggtaaga<br>agaaagttaaagaatctcctttgacttttttgatatagatatttaaattctattactttatagtaaggttggggtttatttttctttgctttataatagaag<br>agcattgattattctctttgcttttataatagaataccatttaaataggagttccctgagtgtgtttacaatcatttgatctggctaaactattttaatgtta<br>atgaaattttaaaatttggaggaaaaaatttaaaaactacacaggtgcacaaagaaataaaaatcacctgctttttcactatgtagagaccattgt<br>ctactatttctcaattctgtgttacatctgtatgttaataactgtaggattagggactgagtactgttttaacctgcttttaaaaaatttacatctacatt<br>ttttcccatctaaatagtgaggaagagtatcagaattttgtaggcttgtggtgatggttaaattagataatattaatgttgggtacttaacataatatat<br>ggctcttaatactctccagatttcagatatagtctgttttaccattactgccttttttatcaaacctattctcaaaaaagtgagaaaagtgctgagattac<br>aggcgtgagccaccatgccggcctcatggttctttcttaataataaattagaagaagtagaattacagggtcaaaaagtatccattttaaagctttc<br>aatgtaattgcctgtttatcttctagaaagtttgacctagttgtattttagagtgtcattttcttgaacttttatcatcattaaagttttaaatttggaacact<br>ggcaatttgataagtatattaggattcttcttattgcaagtagcaaaatacaactcaatctagtttaagaggggaaaatgtagtcattggctaacaca<br>atctaattttggtttaagagacaaatctagagtctcaaatgatctcagagtgtaataatccctgacttttgtcttgatattacttggcttgtatacctttg<br>ctctatttgcatgctggccttactctgccactgacaggctgtctgtatggtgtggaagaggacggctagcatccccatacctgcatccatacagtttgt<br>aatataaaaaaaaaaaaagtaaaaaaaactccctctctcttctagtgtctatatatcagtttcctagaagaaaacgttttgccctacttggccatgtg<br>aatggagttccctgattacatgagtcaaatatgtcttattgtagcatatttgatggtcttcttgtagaatattatcttactatacacagaactcttgacca<br>gtaattaatgggccatgagtttttgttgcaagtcatttgaattcatattctatagttttctaccaagtgtagtcattctgcaagctgttcttgtcatgactt<br>ttgggaagttgagtatttcttctatgggttagggttttcatctcaagaaaaagatgatccttttctctactaaatatgtgttaagatcacacatttttcta<br>gatcgtttagctctactgtgtgatgcttacacaaattgctttattgggatgataagaataattgccttataggattgttatgagaatgaaatgatacatca | 6.) rs7143164 |

*(Continued)*

*(Continued)*

| | |
|---|---|
| actcatatgaaacactcagaacagctcttggcacaaagtaagggcttaattaagtagaaactatccatatattcataatattatagtattggttaagt<br>tgttttcaacattgtttagaatcggctcaagccttctttgtgataatctgacgaaggctattcaccaccagtgagtaaataatagtggcagaatagttac<br>tgatgcttttccttttacttagttttttttccataaacatctggcctttgcagactaaatactggtttatgtatagacatgttattctaaaataatttttccata<br>gtggtaatactaaaggaagaaaaatgttctcaaagctatttatttgggatgttaaaggaggggggaaattaagaaagcctacatttccatgtccttttgt<br>gtccagaatctcattaaatgtcttttaacttgttagcagaggaaagttggatattgcctgcctttgtagctaacatagttaaaatatttaaatggttata<br>gtgtcaaaccagtagtcaaaagccttcactgtgaatggatgaagggatattttcttgaataatttaagttgacttatttcagtggttcaaaaaaatttctt | 7.) rs1951795 |
| cacgcttaaccatgactcaggcacctaactattatactatgtcctgtaacagattgttgtgcattcatttattcaacaggtatttgtgcagctaatttat<br>tgagtacagcattgaatcgttgatggcttaggccacagttgaacattccattttttatgttcattcattcattcatagcatattccatttttaaatttttcag<br>ttcattgcactttaaagtttgaggttcttgcgaagtacagacttttgggtttaagttttgttatttaatgtcaaccaccacaggcgcattggccagtctgc<br>ttttagaattttcagacatacatacacaaaacattctcacaagacaatctacttatttttcttttttattcctgtgtttcttaacacaggattaatgttcaga<br>tctcttttggagcaaaataatcctctgaatttttgagatgtacccagtgacctcagtctgagtatgtatactgcattaaaaaatgtaaccttgttccttttt<br>agtggtcatttggtaacagtttgatcataaacaaatgcagcctcaaacacagaaggcttgaggcaagtatacagaactatggagagatcatttaga<br>tgatgtagaatatgccttttctttttttacaatgccaccaaaatgaaaacacggttttaaaaattctcatagagtgtaacttcaacactgcttttaactct<br>attaaacaaagcactgccatgttgtaattcctatttattactctctggagttgtataaattaccaaatccgccttttgtttgatatccttttcaaatatctg<br>agggtagctatcatgtttcttccttctattcttaaaaaatagtcccaaatttcttgaatctttaatttaaaaattatatattgagcatctgatttgtggaa<br>aggcataggccatattaaaaatggggcttcatattaaaatggggaaaagggtggagattctcaggtggaatctgagatctgccacacactaatagt<br>gttacctaacccttttaaagacaaagaaacaggatcagaaggtcactttggaaaatttatttggtaatattggataggatggattagtatagttgga<br>aaacagagactcttgctttaggagagctgctcctttgtcatttccagaatcttaatcatggtcaaggtttagagctaaatatttaatagaagaagtctt<br>tagggtatgctttctattgtacacccttatttcaatacatgtgtttttttcctgttatgtaagtactttattattatttatgcatcttctattaaagttaagcaa<br>ataattatttcaaggacacattgttctacatacacacaaagtttagggtcactgaccttcttaggttctagtcttagatctgttaccatctaagagcata<br>taaataagggaaacagaaagaaaaggatttacaagctgagaaggaagcaatgcagagaaagaagagtgatagagtaggtaatttgggggaaagt<br>cagtgatacacagctcttaaccatgaacagtgattcttcactcttgaatgtttgtgacattcatgaaggtattaaaagctgacttttaaaaaattgtttc<br>agagaactggaaaaaaattcagttgccacattcttccttaggtcattctttgaactctactcatgcacttacgtgtttaaggcaaagttttactaaacgc<br>acacttgttcttgctggcttattgacttttactgctagcttcttattcttagcaattatactcacattacatagtattgtgaaactcactatattcagtgtt<br>ttgcctgacaaacatggtatgttataggatgtgtattcagttatagctaaaaataaattattctcgtttttcaaaatttgctggcctacctgttaagcttt<br>tgctttaagacctgctaatgtttctcaaacttctgtggttaaatcacctgagtgtctagttgctctatggattcccagggacccattcgccagagattct<br>gatttggtaattttgggatggaactcagggatctgtaaattttacaagcactcagaaatgaaacatagactttaaacagctaagagtgctcatcagg<br>attatgttgatattatttttaaacagatgtgccaagcctttaatttgaatttccagggttgggatttggccttctatatttgggggaaaaaagttctatt<br>gatgattgtggatatataccacaggtcaaccattgaatagtctagtcagtgtagttagtgtattttataattactaagttctaagtatgtggtgtattaa<br>tgtcttaggaggtggatatatttcctgtatttgtaaagcattttgggtaggttttttaaagagaaaagtatgtaacaaactagttttgagcgttgctctttt<br>acttctttgggcatttttgaagaacacgttaagtatcttcttagagcagaggggctcagagtggtccccagattatcatcattggtaacacctagttgg<br>tgcattactaacttgttagaaatgcacattctcaggcgccattcagacttcataaatcagaaactctggaagtaaggctcagcattctgtgttttttttt<br>tctttattatactttaagttttagggtacatgtgcacaacgtgcaggttagttacatatgtatacatgtgccatgttggtgtgctgcacccagtaactcgat<br>catttaacattaggtatatctcctaatgctatccctccccgctccccccacccgacaacaggccccggcgtgtgatgttcccctttcctgtgtccatgtgtt<br>ctcattgttcagttcctacctatgagtgagaacacgcggtgtttggttttttgtccttgcggatagtttgctgagaatgatggtttccggcttcatccatgtc<br>cctacaaaggacatgaactcatcctttttttatggctgaatagtattccatggtgtatatgtgccacattttcttaatccagtctatcattattggacattt<br>gggttggttccaagtctttgctattgtgaatagtgccacaataaacatacgtgtgcatgtgtctttatagcagcatgatttataatccttgggtatata<br>cccagtaatgggatggctgggtcaaatggtatttctagttctagatccctgaggaatcgccacactgacttccacaatggttgaactagtttacagtcc<br>cactaacagtgtaaaagtgttcctgtttctccacatcctctccagcacctgttgtttcctgacttttaatgatcgccattctaactggtgtgagatggta<br>tctcattgtggttttgatttgcatttctctgatggccagtgatgatgagcatttttttcatgtgtcttttggcagcataaatgtcgttcttttgagaagtgtctg<br>ttcatatcgtttgcccacttttttgatggggttgtttttttttcttgtaaatttgtttgagttcattgtagattctggatactagccctttgtcagatgagtagatt<br>gcaaaaattttctcccattctgtaggttgcctgttcactctgatggtagtttctttttgctgtgcagaagctctttagtttaattagatcctatttgtcaatttt<br>tggcttctgttgccatggcttttggtgttttaaacatgaagtccttgcccatgcctatgtcctgaatggtattgcctaggttttattctacggtttttatggt<br>tttaggtctaacatttaagtctttaatccatcttgaattaatttttagcataaggtgtaaggaagggatccagtttcagctttctgcatatggctagccag<br>ttttcccagcaccatttattaaatagggaatccttttccccatttcttgtttttgtcaggtttgtcaaagatcagatggttgtagataagcggcattatttctt<br>gagggctctgttctgttccattggtctatatctctgttttggtaccagtaccatgctgttttggttactgcatccttgtagtatagtttgaagtcaggtagtt<br>gtgatgcctccagctttgttcttttggcttaggattgacttggcaagcattctgtgttttgagaattcttccaggggactgtgatgaaaactgacgtttga<br>gaacctttcatcttagagtaaaaactttacatacacatttttgttgttttatttatctagcacaatacttctttttttttgaaatggagttttgctcttgttgccc<br>aggctggagtgcaatggtgtaatctcagctcaccacaacctccatctcccaggttcagttgattctcctgcctcagcctcccgagtagctggggttaca<br>ggcacgtggcaacatgcctagctaattttgtatttttagtagagacggggtttctccatgttggtcaggctggtctcgaactcccgacctcaggtgatc<br>cgcccacctcagcctcccaaagtgctgggattacaggcgtgagccactgcacctggcacaatacttatatataatcagggctcaaagatttgttga<br>gaggctcaacaccaattctggaccaggaaagattttatttatatcactagtcaggaataatctaaaaacaaaaagcacattcttcttacaagtaata | |

*(Continued)*

*(Continued)*

| | |
|---|---|
| tttcaatacacattaatgtaaacacatggaaaagtattagctacttaataaattaacatgtaaatgaaaaatttacacattatggctatttcagatgtg<br>atatagatttcattttcagaaggaaccctccaatgtaaaacagtgattcttttccccggtttattttactggattagaaaatcacatttaaagtaagcatt<br>ttggtgaggtttggaaggtgaataaatccatcttttctttaattatggatatttaagagagatgttgttgtgccgtttagataataatgatctaaaccaa<br>gaaatttagttgctttcaaaaataaaataagtgtatgcattctgaacatttttctttagaaacaaaccatttcatctgttttttttgaatttcaaattaatta<br>tacagaattttcaaaatttgaaaattaggttagcatgagaaactgaagatactgaattatattgcctgttcagtctatacttttctttaggatatacagt<br>aggaaagaaatatgatagttcaagttagattactacttctttcagagttttttgacaaatgcaggtacagtgatagtgtcagttcatggtgaatttttgt<br>taaaataaattacaaaaaatttgtgatcctggtatcttgaaactagttaatatttgtaaactttgctaacactgtatatcactgtattctggttttatctg<br>tgcatctatgagttatatgtgtgtatagctacatatgtttatatttatacacatacattacacacaggagtggaatcatactcaatttttttttgtatagcc<br>tgctctgttcatataatactatattgtagcatctagtaataagcaaagattaatttttgtagactttgcttttatcctgaaattttgtggtagctggtttaatt<br>ggaaagacaatttctgtgacggtgttttgtcagttagggattgacgctggtaaaatattgctggataacaacaagcaatgtaaaaatacatttgttcca<br>taagataacctccgtgaaggtagagacttggtctgttttgtttattgcaccgtgtcctgttctgggaagagtgttagactcatagaagatgatcaagaa<br>atatttttttgaatacatcaataacattctctaacatgtgggtatcctaaaggtttatttttaaagtttattgattagaattcagaagatattttcccagat<br>aaaataatagattgctagctgtcttgaaaatgtaatttatatttaatttgaaatgtcaggtttttgctatttttttccattaagtagagataggggttttaa<br>aaattacatgtgatgttttaagtattctggttttgcaacaattactagatagaaaatgtaacaacagatcctattaataatacttccaataatacatat<br>aaaatacttgtctaaaagtaaccctccttaaaaaaacaaagctggccaggcgcggtggctcacgcctgtaatcccagcactttgggaggctgaggc<br>aggcggatcaagaggtcaggagttcaagacgcagcctggccaacatagtgaaaccccatctctagtaaaaatacaaaaaattagccgggtgtggtg<br>gcaggcgcctgtaaccccagctactcaggaaaatcgcttgaacctgggaggcggaggttgcagtgagcggagatcgcaccactgtacttcagcctt<br>gggcaacagtgcgagactctgtctcaaaaaaaaaaaaaaaaaaaggcaataggattaggtatcaacttaatgaaaacttcgtgacagcactttctt<br>gaaaaagactgtggaaaccaaagttagtaaactcctgtttctgcctgggttcggaaaacataaagatgataaagatgtttaagtattccttttttttttt<br>ttttttttttttgagacagtgtcttgctctgtctggagtgcagtggcacaatcacagctcactgcagccttgaactcctgggctcaaataatcctcctgc<br>ctcagcctcctgagtatctggaactacaggagtgcaccattacactcggctagtaatttgattggttaagaacattaactataactcacacattttcct<br>gaccacatttgcttaggacaaaacagtaaaagacatgagtgtagatgaaagcgataagggaactaatcttaaacactgaacctcttttcagcaaat<br>tggctttctagtttctcagctctctctttacacctctaaatctctttcctggcaagatcatttatttgccttggtttatggtgatactcttcattgttatactg<br>gtgggtgattgttttaattgatagctgttttttctacttcaggaagatgacactgctggctctgctggctctgatgtttaccttgtggctaatgcctgtgt<br>ttgcctgtgttcacatttattccacgattcatttgttaacatttactaagctgcttttctgtgccaggaacttggctagataaataaatggttgtttttgta<br>cacagaattagctgtcataatcagttactgtagcatttattcttgcaaaaatatatatttatacttcaactagtgatcgaatctcaacttattaattcata<br>cattcagccagcacataattgaatacttcttatgtgtcagaaactgttctaggtgcttgggatgttcattgaacaaaatagacaaaagtctccgcctct<br>atggaacttactttccagtgaaggtgtggattggtgggatagaaaataaaataatcaagtaagatatgtacttaggctttcataaaaatacagcagg<br>gcaagaggaccaagatggaggcagtgatcagggaatctcaatgagggtgagactgcgacaaagacttgaaaaaggtggagaagcaagccttgt<br>gggtatttagggtagcagtagtccaggcaaggggaacaactagtgcaaaggctctaggaggcaatgtgtttgaagtgttttaagaacagtaaggag<br>gctagtatggttagaacagaatgagcaaaggggcaaagtggtagaaggtgagatcaaagaggtaatgaggccattgtggaggcccatatggac<br>tattggaagggctttggcttttactctaaatgaggcaaaaaccatttaagcagagaggagtgatatgacttgatttcttgttaaaaggattattctag<br>ttgctgttacagaaaaagattacaggggtgcaaagaaacaggaggagacaaaagaatataagattttcactgtaacttatatctagtatgcttgcttat<br>acttgaaaatgcatatccagataattgtagtaaattcaaatattatgtttatttaatagtactaacattgatatgctggttaattatgattaggagcact<br>aataaagcacaaatcagggattcccaaaagaatgttgaaagggcagtcagcttttcctgtgccagaaatcaaagtcatagcagatttggggcaa<br>atatgtcaaagtcaaacttacgcacatcactactgagaagacaaagatgaatgtgtgacagtttcctgcccccaagaatctttaagcattgtgaagg<br>aagattaatatagccaaataactagagtgatcagttctaccagagaggaccagttttggaagccagaggaaaaaaaaaaaaaacagaaacaaa<br>atgatgtttgaattaaatctttaaaagtttctcttataaatttaccaagccacatattgggaatggtaccccaggcagaaggagtagagtaagcaagc<br>cagaaaggaaatactatggtgcttttgagtaactgcagtgtggctgaagaatgtggaaaatgatgaggataaagaggtggacagggaactaggta<br>agggagggcttcctttttaaataattagaccttgtcctgtgtacatttaatgggattttaatcaggccataatgccaaatttctttacttcggaaggatct<br>ttatggtgatggtttcagaaagaaattaaagcagagtaacagtggttagcaataatgatcagctagtggttcccaaacttacgtatcatatgcatctt<br>ggaagtttttaaaaactcagattttgggatcctgacttagatctactgaatcagaatttacagattcaaattcccagtgaggcctaggaatttgaaat<br>gttgaatgtccttcacgatgcagctagacaagcatttgggaataaagcattaggtgactatttcagtagactaaggagtgggaggccatttaagctc<br>aaaggctattctacttctcactatatttctagtacctagcacagtgcatggtacttgatagatgcatcctttctcccatacctcgcccttacacatctcttc<br>atgtgtatccttattaatatcctctattataaactggtaaacatgtttccctgagttctgtgagctgctccagcaaagatgggtttgtgagaatcccaac<br>ttttgaagcctgtcagtcagaagttcctgaggccagacttgcaactcctgttgaggggggcagtcttggggactgagccctcaacctgacactgtctcc<br>aggtagatagtgttagaattgaattgaaggacacccagttggtgtccgctgcagaactgattgctcacctggtggtggagagaaccccctcctctccc<br>gatagggttgcagaagttgtcttctgtgttgttgattgctgtggtgtgggagcagagggggggaaaaaagctgttggagagtttttttccaaaacaatag<br>gagattatttagatttataaaaatagaatcaaagtagattaactgagcacattgtgaaatatagagtagagctgtgtgtaaggagtatatcttaatgt<br>caagctgacaccaaattgaatgtttgctggaacgttcaaaaatctaagcttcccaaatctgtgaaaacactcaggttagtaaacagtctttatgcaaa<br>cagcaagacaatgctcaaagccatttaaggaaaaagaacagtaactgaattctcttatggaaatgtgagatgttgttttagtaagtactgatggtgtt<br>atacttttgtttattcgtttgctggtatttcagttcctaaaattccttcaaatatgctgcaaaatacaaaccaagaacttggtggattttccatttgttttc | 8.) rs12435848 |

*(Continued)*

*(Continued)*

*(Continued)*

*(Continued)*

| | |
|---|---|
| aggaagacagcttttgcttggtgtgacaatgaaaatctttctgatttccttcttggaagagttccctgaagatatgtcattgtattgacaccttattttt<br>gctaacctatccctctaaattctggatattgtgtgtgccacagcttttttcttccatattcctgcatttatttggcacctgttgtgccagtaatagataag<br>gggctgctaagggaggaggcaacctgcactggcttatagctgctaatgtcagttcctatagcttatcgtcagtgttattcatgtggtaaaagggtgag<br>aaagtactggagtctaaagaaacaagtagaaatcagtttgtagctattaccgttctacctgctaacaactcctgttttcaagttattatgtacaacttt<br>aggtagtttctctagccttaatcgtggtttctctgtattgagactacttttgaattctatgaagtacagccttagatgtacaggctactttaaatttttgcc<br>taaaataaaaacattctctccaattacatatgctggggaggaaacacctgcttccgacaggtttaaagcttggttttggacttttttgtgagagttcctt<br>atgtgtgcagtaatccaaaatttgtatagttgcccttttataaagtacattaatctagtagacaaatctccatgtaacttaattacatggcatcttctaa<br>tccttctgtgataagcagaaatgtaaagttttattcaagttaaggcaaactaacttgtatacactttccatctcgtgtttttcttgttgttgttaagtag | |
| ENSE00000867524 62,187,100-62,187,290 (191 bp) | |
| GATAAGTTCTGAACGTCGAAAAGAAAAGTCTCGAGATGCAGCCAGATCTCGGCGAAGTAAAGAATCTGAAGTT<br>TTTTATGAGCTTGCTCATCAGTTGCCACTTCCACATAATGTGAGTTCGCATCTTGATAAGGCCTCTGTGATGAGGC<br>TTACCATCAGCTATTTGCGTGTGAGGAAACTTCTGGATGCTG | 9.) *p.Ser28Tyr*<br>(c.111C>A) |
| Intron2-3 62,187,29162,188,226 (936 bp) | |
| gtgagttatttgacaagggtataaataggcctgaaaattagaagttagaagtaaatagaaattatttttagaaggtggtcgcaatgttttgattttgta<br>tacctctttatattgtgatatgtacacgtttaaaaatttttctgtaattctcactatttttatcaagcttcatttttttctcatcagttattctttgaaataatc<br>attctttatgcacataatttgttttgctttattctcttaaacatactctcaattctttttctaatataacatccttttttattacctgcttttaaagctttagtcag<br>gaataagatactggctttttcccctccccctttttctcctgttgcatctaccttttcttcctttaaaaaacatgactcaggcgggcgcggtggctcacgcc<br>tgtaatcccagaacttgggaggctgaggcgggtggatcatgaggtcaggagttcaagaccagcctggccaagatggtgaaaccccatatatacca<br>aaaatataaaaaattagatgggcatggctggtaggtgcctgtaatctcagctactagggaggctgaggcaggagaattgcttaaactcagagggcg<br>gagcttgcagtaagccgagatcaagccactgcactccagcctgggcggcagagtgagactccatctcaaaaataataaaataaataaataaataa<br>aaaacattactcttctttcttcttctatggtttgctttgctgcattactttaatcatgaaaagcagctggcacatctaattatagtttttctagcttctggcc<br>tgcacttttctgtgttgaaatggctgtatatattaaataaagtgtctgcgagaaaactttgtaaaaacatctaaatattatatcatttaagtacaacttt<br>ttaactaattattttcctcttcttgtgccctttttag | |
| ENSE00000658281 62,188,227-62,188,372(146) | |
| GTGATTTGGATATTGAAGATGACATGAAAGCACAGATGAATTGCTTTTATTTGAAAGCCTTGGATGGTTTTGGTTA<br>TGGTTCTCACAGATGATGGTGACATGATTTACATTTCTGATAATGTTGAACAAATACATGGGATTAACTCAG | 10.) *p.Val116Glu*<br>(c.347T>A) |
| Intron3-4 62,188,37362,188,456(84) | |
| gtaaaatgcacacatattaagagcttcttctatatgtttttatgattttatgatctagccctaatttttaaaaatgtgtttacag | |
| ENSE00000658282 62,188,457-62,188,541(85) | |
| TTTGAACTAACTGGACACAGTGTGTTTGATTTTACTCATCCATGTGACCATGAGGAAATGAGAGAAATGCTTACA<br>CACAGAAATG | |
| Intron4-5 62,188,542-62,193,423 (4,882 bp) | |
| gtaagaaaagtctgttgtttgatttaatgtgacaggtggttttacataataagatactattgctaattattaaactttgctattgtacttacccaaggca<br>aaatgttatttcatgtttaataaaatgtctattcttttgttaaaactattattttagtttttaggaatttcattttgaaagcccacctaattgcataaataatt<br>gtgtgggtgtgagaaataaaatggaaaagtaaaatcatgaccaagagagttacaaataactttttttttttttttttaagatggggtctcgctctttgc<br>ccatgctggagtgcagtggcacaatcagctgactgcagccttgaccgctgggactcaagcgatcctcccacctcagtctcccaagttagctgggacc<br>acagacgcgtgctaccatgcccagctaaatttttaaaaattatttgtagagacaaagtctcactatgctgctcaggctggtcttgaactactgggctta<br>agccatcctctcacctcggcctctcaaagtgttgggattacaggcatgagccaccacgcccaggctacctttttttccttttcttttaaattgtgatag<br>gggttcttgctgtattgcccaggctggtcttaaactcctggactcaagtgatcctcctggctcagcctcccaaagtgctaggattataggcatgcgcca<br>ccacacctggtggagttaaaaattaaaatacaccattaaggcaaggagaaattataatacaaatggcagataataggactttagacagtcattaaa<br>gttgaggtgccagtttgagtctaaggcccaataaaaaaagttcaccagaattttaagacaaacaactgcttatttgacttctttggatgttctcaataa<br>ttcgagaccgtgtagttagattataaagtattacattgtggatgcccacatattaacaaaaatagagagtaagacctctaattcttaggaattaatgtt<br>gttaaaaataatcaagtgttccaagatttttggaaactacctcttgaattaaaaaattaaagtctttctacattttttatcttgttaaacagtgtatactg<br>atcataattatttaaaaaatcatgtgttctaagatttttggaaagtacctcttgaattacaaaaacaagaaagtctttccacatttgtgccttcttaagc<br>agtgtatactgatcataattgaacttttcttcatgatggaaagttaccacaaggaaaatttcttatgttctgctgttctttgttgctctccaatttaagtgc<br>atacgtttgtttgcttctatattataaaacctcaaatttacttttttgtataatttttgaggttttcttttttcatctcatttattataataatagctaacctccat<br>tgagagaatgctgtgtgccaggacactgttcttcctattttatatgcttttaactcctttattcctcacaacaacaaccctgtgaagttaactgttagacaatt<br>tctattttactaggaaactgaggtacagagttactaagtaacttcccaacattatttggttagtaaatggcagagcttgggctgaacttcagtagactt<br>ggcttcagagtccacgctcattagtcctttggagcgcttttcatattcttgaattctcacattctgtcttttttcactctgtcagcaggacctgactcctgtt<br>tttaaatttcatattgtgtttttactgttaatttggaaaacaaatgcatacttttagaattctgtataaaggaggagtaaatatgctgtgaacaaggac<br>ctaagtggttgtcaatgagtttaatatatgagttctaatgtgcagagttgaggtttatattgactgctcagtgcttccctggggctagactataaatgg | 11.) rs4899056 |

*(Continued)*

*(Continued)*

| | |
|---|---|
| atggatattaggaagtcttgttctgatttggtaatgatgttaatgcattattctaaatcagatagtcttaatatagtttaaatgtatgtttcgaaccaaat<br>gttctttttttaaagcacacaaacattttgaaatcattactaatgtggttaatgaattattgatgttccattgggaaactaaaatgcagattttttctctttt<br>agaaatcagggactattgcaaagcatcacatttttagtgatacactgagagccagtggtgtgtttatacaaatagtcctattttccaaataaattctag<br>aaaaatgctttagaatttataaattatacaaaatatgacttatttttagagagtttaaaatttaggttttttttaatggtttgttttttgtttgtttttttgt<br>ttttttttttcctcattaggaaaacactagtacttttcagttaccttgatttttaaattaatctgcaggtccccattcaaaggccttgggttcctttcaaagg<br>tcagtataattcaagcttagttatgaaggactgaacatacccaaaggattttgcatgtggatctttactgccactaccacaaccatcaacacctaca<br>cacacacgacacacacacattctctctctctctctctctctctctctctccccctcctcccgcactccttcgcttcccccctcctttgctctcatggcatc<br>ttttaaaaatatactcttaaatccttccagggagggcaaattcacttcttaatctaagtaaacccaaatggcatgcatcagcaccaggactgcccatc<br>tttcctagttccattattcatagagtataggctggaattcatcttgttcctcaagagtccagcatttctagttaaccatgcctacatttaaacttactctc<br>atttcttttctactttacagtgtttttttcaatatactagcattacagttccagatttgatttctctcctgtcttatttccatcagttttcaagtctattaagat<br>tctacctcttcatttgtcttttgccaccattcttttccctcatactctactggctcagccctctcattacagtcacctaattctaacatatatattgctgcta<br>agttaattttccttaagttactgattgtgctttttttaaagccccttgttgaatatttaggcaggactccatgtggacatccacagccctccgtggtacagc<br>cctaaccttcccttctagctttgccttactactcttctacgtgtactctacattgtggacaaactactatatgctgtttttcaaacatgtcctatttttccta<br>cctctgtgcttttcattctcttacttctccttggaataccccttctaacccatctctacttactgacattctaatgtctctttttctaagcaagacttcttgattt<br>cccttgactagaaattatcttctaagctctccctatccttctttaaagcatttttataagtctcaagtaccaactctacattgtgttttttgttgaccttacta<br>tatctactacatttttaacttcttcaggaaaggtggcgtatcttactcatcttttgtattgcctacaatatctagtccaggttctgaataataaatatttttta<br>tatgtgttctgaagcacactgaccaatgaagataagaaatcaagaggctagttccttatttttttttaatttttttttttgagacagtgtctcactttgtcac<br>ccaggctggagtgcagtggcacaatctcagttcactacaacctctgcctcccgggttcaagtgattctcacgcctcaacctcccaagtagctgggatt<br>ataggcatgtgccaccacacctagctgatatttatatttttagtagagatggggttttgccatgatggccagcatggtctcaaacttctgtcctcaagtg<br>atcttcctgcctcagcctcccaaagtgctgggattacaggcatgaggcataagccactgcgcccagcaagatgctcttttctcagtcacctaaatata<br>atctcatttttagttatagaaggtttgaaattggagtgaatagactttacttaattctgactttatttctgtagctttttttttttgagatggattctcgctct<br>atatcccaggttggagtgcagtggcacagtctcagctcactgcaacctctgcctcccacgttcgagtgattcccctgcctcagtctcccaagtagctgg<br>gattacaggcacccactatcacacccagctaatttttgtatttttagtagagacagggtttcaccatgttggccaggccggtttcgaactcctgacctc<br>aagtgatcctcttgcctcagcctcccaaagtgctgggattacaggcatgagccaccgtgccctgcctatttctgtaacttttgataagtcatttgatctg<br>ttgttgttgttttctcatagtaacaaagtagaagtaattttctgcctgctttactagataaattaagggggaaaaaaataagatacgtaaaaatgttattt<br>gttattaaaaagaaagttgttatttttaaaggttctataaagacatagagtgcttattagaaattgagctaacacattcaggaaaggataggaagagt<br>ttgctgaagttctttctttagggattcttgtgtacgatagcacagttaaagagcaaactcataccatttttatatttctgtgtatttgactaagcttactg<br>gcttcaatgattaactgttatcccaaatatggattatctttcagccaactcagggaatcacagctactgagtagtgtgtgtcagatctcttgggtgtgct<br>ggagtgagtaaaaggggaatgaattactgtgttcatgctgagacttaattgaacgggtattcagttgatctaggtgatgggcacttgttgttacttttatt<br>gtaacaaatttgtatatttagttgctttaaaactttatttcatgctttcattag | |
| ENSE00000658283 62,193,424-62,193,536 (113 bp) | |
| GCCTTGTGAAAAAGGGTAAAGAACAAAACACACAGCGAAGCTTTTTTCTCAGAATGAAGTGTACCCTAACTAGC<br>CGAGGAAGAACTATGAACATAAAGTCTGCAACATGGAAG | |
| Intron5-6 62,193,537-62,194,170 (634 bp) | |
| gtaagtgaaaattatttgtgattgattatacacttttatttatacatagacattgtagtattaagataactttagaattgtgagggaaggtttacagttcc<br>atggtgtttggttatgtaacatttatatcttcaactcatttgcatgtgatctccaaaatgcagaaccgtgtagtaatttggccaatttgaggcacaaactt<br>aaattacgtgaattgtggcactggtgttccaggcttaatcagttggctttgccagccacacaatatttgaatcctgataggggcttaaattttctattaatc<br>atggttttatatctttgttcaatgttgaaacatagtcatcagtgcaagaaataactatcaaacagccatgatgatgagatgaatgaaaaagcagcct<br>agactttatacgaggggaatttttttaaagagtaatgtataggccctgggcaggaagtaggtcataggtggtatcataggaaaaatgttcattgatttt<br>caaaaacgtgattaatccactagtgacagtaaattttatcaaagcttactggccatgtcagactcaactacttatctctgctttttttttccctagcattg<br>taaatatttttttttaactgctttgttgttcatacacag | |
| ENSE00000658284 62,194,171-62,194,373 (203 bp) | |
| GTATTGCACTGCACAGGCCACATTCACGTATATGATACCAACAGTAACCAACCTCAGTGTGGGTATAAGAAACCA<br>CCTATGACCTGCTTGGTGCTGATTTGTGAACCCATTCCTCACCCATCAAATATTGAAATTCCTTTAGATAGCAAGA<br>CTTTCCTCAGTCGACACAGCCTGGATATGAAATTTTCTTATTGTGATTGAAAG | |
| Intron6-7 62,194,374-62,199,135 (4,762 bp) | |
| gtaaattagatctaaaatgtgaatttgaaatttttaattagtctacagcattactgaatattcaccatagcaaagattcagcgctggccatgcatggtg<br>gctcacacctgtaatcccagcactttggaaggctgaggcaagcgggggtggatcatctgaggtcaggagattgagaccagcctggccaatgtggt<br>gaaaccccatctctactaaaaaatacaaaaattagtgggacgtggtggcaggcactactcaggaggctgaggcaggagaatcgcttgaacctggg<br>aggtggatgttgtggtgagctgagctcacaccaccacactgcaagcctggatgacagagcaagactcccatttcaaaaaaaaaaaaaaattac<br>tcaatgttaaactatactttccactaaattgaacagaatgatacatcctataatattagattaactttgtaaattaattcagccacatttattgaacatt | |

*(Continued)*

*(Continued)*

| | |
|---|---|
| tactctgtactatgaacacttactttactaggtgctatccagaagttaagatgagtctttttttccccaataggggctctacttacttagagaatttcaa<br>agatatgcagtgtgtattttgagcaaagatagattaccttaggttggggactagaaagccaagtgtttgtacatctcttcaatcctacatattttccctga<br>gaagcttcaaccttgcccatggtttctattactatttcccacatttcttcctgtaactaattctatttaattgccaacttaatatttctatctggatattcttc<br>tgtattgtaaactaagtattactgtaacaactgtactactactgccccccaaacaacatcatcatcaaaaactgcctttcttgctataatgcttattgtgg<br>tttaatacaccaccatacacacatgactccagcaaaactttggaagtcatctgtaacttttctttttacattcattggctacatacagttggtgtctaaat<br>cttacagatttactatctacatatatctcttgatccatttcctcctttccatccttgcactcctgccattgaattcattagctcattattactcttgacttgag<br>ttgttggcatagctgccttttttgccaacagatttgtaccccttataatctttcatctaagttgccagaaagtgggtgtcctaatgtgaaaatcagatcatg<br>tcattctgttgttgaaaatgcctcaaatgcttccctccatctttgcacacaaaaatattttgtttataaaaatactagatgagggaagtaaatttttcatt<br>tatcaaaagaagatgtgtatttttagaagactgaaaaaaaatagacctacacaatacaatctaaacttagcatggcaaacaaagatatttatgctct<br>ggccctaactctgtctttggaatcagatgttagattcactcatggcttgcagctctgatacttacaaatgtggccttggccttggtacttaactgttgtaaa<br>attcacattccttatctataaaataagaatcatggctgggtggggtggctcatgcctataatcctagcactgtgggaggccgaggtgggtggatcacc<br>tgaggtcaggagtttgaaaccagcctggccaacatggtaaaaccccatctctactaaaaatacaaaaattagctgggtatgggggcacatgtccgt<br>aatcccagctacttgggaggctgaggtaggagaattgcttgaatccaggaggcggaggttgcagtgaaccaagcttgcaccactgcactccggcct<br>gggagacggagtgagactccatctcaaaaaacaaaaacaaaacaaaaaaagacctcagaaggatgttgtcaggattaaaggagtccattgagt<br>gcctagtacagatagtgaatgcttcactactggtgtcaactttaagaaaatgaatatagaaaagctaagaattattttaaggtgtttactactagcatt<br>gtaaatgtatgatgggacagagatttccatcctattttgaggaattatttttttatttttttgaaaacttaaggtaacaaagtagagaggaggccaggg<br>agaaaggaaggtagtggagcaaaaatgagaaagggagtgacattcccctctagttatagcagaaaattagcaaaatgatcatgacaggaggtaa<br>cagtaaagacagccagctcatatatcaaccaagacagttttgagtttgaccagcagactgttatttttctggtttagagctctttccaggaacttcttgc<br>atctataacccctgagaaccaagctatggaaaaaattttgctcaattttaagaaaatctaacatatcaagctcctcaactccaaaatattccacaaat<br>agctgctatttactatactgagtaataatcatttaaaattattcaacactttatttgagcatctactatgttcatggcactaaagtagaaatgaagatg<br>aacagttcctgcctcaaaataaatgagtagtatactgctttagatcatgggtttcctagtccattaaaaacacttttggtcatattttctggacacccc<br>gacccttttggtatagaatataacctatgtaattctctaaagttaaattaacctcacttttcttgctctaatatgtgtaaaactgaccttctaggaaagc<br>atatacagtttatatttttgacttcttggtatctttttagtgatagacatacctcagattgagaagcactgattgacattagattaaatcagagcttcctat<br>gacaatataaacaatacctttcattaatctgatccccctacctacttcttcagcatcatctcatatctgtctccactaatcatattatagaatctttgttac<br>ctgcaccatgttaagcatttttaaaaatctttgtttataccatacctttttcctgaaagcggttttgcctttccttttgtctctagtcataagtctcctata<br>agaggctgttcctcattctaccattcctttgcatggataggattccatggaatagattctcatcactgcatttatcacattatttcctaagtagtacagta<br>catctactggaagattagccacgtattgagttttgtctttgcattttcatgcctagaataatgccgggcacacataggcatattaagatttgaatagtg<br>aaaaagtttttaattccatggggattttatttaaacagaaaaatataagaccaattagaattatttttaaagcataatttcaagaaatatgactgattt<br>tgtttaaaaacatgttttccttttataatgctgccacctggtgttgctgtgtttagagatgtccctttgtaaagaattgagggtttgagttgagtttggtttg<br>gtttttggcaaatcagcttttcgtttgtatatttattttgtaataaactatggaagatcttgcctttaagtgtgagaacacaagcaatgttacttttatacc<br>tttatagaatatcttgcctatgtccttcctgtagttaggtagggttttttttttgacacacagcatgttatataaggtttgcttgcacctcggtaggaaagt<br>cctctgaaatctaaaggctgagaatctaaaagcttaactcatgttttgctcctagaaagacttgagaagagagtatttctgttcagcatggtactaag<br>aagacagctttctcttcctcatgtcatggttgccatttcatactgcttacagagaataagatctagtctctgtcttaaataaaggtctactctctgccag<br>cgagctagatagggtaattggattgttttccaatctattttcatttgaaatattgttttatgtgaaattactcccataatttcatgtaatgccaaaaacta<br>aactaagtacaagagcatcttcaaaaaccaacataattcctttagttcccatttagtgtagatgctctttggttgatgatattagaattgtgtaatggct<br>attgatctctcaaagtgaggtgttgcctaggggcttaaaagttactacataaagaatttggctttatgaagaaatgttacagattttatctatattttaa<br>aataagtgtaagtgactacctttataacttttaccatgtagtttagtagtatttcttatctgtttattaataccctgccttgttaccaaaagtatgtataat<br>gagatgtaataagaataggtaacaagtaggctgggcacgttggctcatgcctgtaatcccagtactttgggaggccaaggcgggtgaattacctga<br>gttcaggagttcaagacgagcctgaccaacatggagaaaccccatccctactaaaaatacaaaattagctgggcatggtggcacatgcctgtaatc<br>ccagctacttgggaggctgaggcaggggaatcgcttgaacctaggaggtggaggttgcggtgagccaagatcacacctcattgtgctctccagcct<br>gagcaacacgagggaaactcttgtctcaaaaaaaaagaccaggtaacaagtttgggtgaacaggattaaagagttaaataacaggaggaatcta<br>gaggacttaaagaaatgtgtggtgttggatttaataactgtagttgccaaaggtgaggtgtaaatttattctaagcaaaggaggatgctcatttttg<br>aaaattcacttgtccataagattaatgcctatcagttaacttgggaggagaaaaatttttctttatcagtgtctccctttttttttcttaaatcttgtatttttt<br>tactaacag | 12.) rs1957757<br>13.) rs12434438<br>14.) rs11158358 |
| ENSE00000658285 62,199,136-62,199,242 (107 bp) | |
| AATTACCGAATTGATGGGATATGAGCCAGAAGAACTTTTAGGCCGCTCAATTTATGAATATTATCATGCTTTGGA<br>CTCTGATCATCTGACCAAAACTCATCATGATA | |
| Intron7-8 62,199,243-62,200,855 (1,613 bp) | |
| gtaagtacaatggaagaactcagagatattctaattacttaactgttgcaacctctgtacagtttggctacccatctaattctctggttaaaagttcta<br>gactaaatgtgttaacaggcctattcagtagagatcttgaccattttgtgttttgtatgtgttgcaacaaatatcagtaaaaatagaatcatttaatcat<br>agaaaaaacttcctggcattttaaatacaaagacttttgaaaatgcaaatattatagagtattgaatagcataatttttcagaattcacataaatactc | |

*(Continued)*

*(Continued)*

| | |
|---|---|
| agaacagtggttggtatgtaaaaggcactcagaaagtatttgtacaatcaatgaatgtgaaggtggtgaacatcacctttggtaataagtaccatttt<br>aaaaaatgcttataagtgcatagttaggtatttatatttatgggttcatgaaatattttgatataggcatgcagtgcataaggataaatggagtaccta<br>tcacctcaagcattatcttgtgtgacaaacaatccagttatactctttggttatttttattttattttattttttctttgagacaggatctcactct<br>cgcccaggctggagtgcagtggagcaatctcagctcactgcaacccccgcctaccgggttcaagagattctcctgcctcatcctcccaagtagctgg<br>gattataagcatgtaccaccatgcctggctaatttttgtatttttagtatagacagggttttgccatgttggccaggctggtctcgaactcctgacctca<br>ggttatccacctgccttggcacccggcctcttttagtttctttaaaatgtacaattaaattatttttttactatagtcacccaaaacaagtacctttgaca<br>taagatttgattctgaattttactcaaatgaatgttaagatccccaagataagttaaactttggactatctcacctgtttaatctgtacctatgcatgact<br>tcccactgtgcttgaggatacctgaatatcactgagtttgtgtgactgatcagccttgaactcaagagtaaatccaagtctgcagtcaggacacccca<br>atcctcaaaataataccatcattagcatttatttagtactttctcccaaatcagtatttaatttaaattgccaaaagacttacaatgtggtatcaatttat<br>atttaaatatgctacatatagctttttaaagcatctttggttctctggaaaccatagtcagaatttaaggaagttattgtggcaccatttttcttgaaaaa<br>ggctattgattattctctaatctgacaccaacctaagtcattaaaggaatttgagttactgaagattgtatattcatgaactcttcacttagctcactgg<br>cagcaaaggagttttatttaggggggtttgaaaaaggaaatgggtacattttcagctattctgggacgcactgtcagaatgtaagcagttacaactga<br>ttccactaaataaacatttgttttccaaaacaatgatgaacattcagcatctgttcatttaattgaaaattcaaagttaaaatattttctctgcatgattc<br>ttttttctttccccccctag | 15.) rs2301111 |
| ENSE00000658286 62,200,856-62,201,003 (148 bp) | |
| TGTTTACTAAAGGACAAGTCACCACAGGACAGTACAGGATGCTTGCCAAAAGAGGTGGATATGTCTGGGTTGAA<br>ACTCAAGCAACTGTCATATATAACACCAAGAATTCTCAACCACAGTGCATTGTATGTGTGAATTACGTTGTGAG | |
| Intron8-9 62,201,004-62,203,606 (2,603 bp) | |
| gtaagtaagtttgagaaataaacatttttggggaacaaatagtaattcttttggatactctgttcatttataggaagataagataataaatattaact<br>aaattttaattcttttacatcgctaccaaattattattttctatactctgacctaggtttccagtccagctattccacagtgatgctgctaaacactgtca<br>gtagttgtctatccccataccttcactcctatttttaaaaagaccatgaaaaaaataccagatccattgattggtttggtctaattatacagatatcggc<br>atatactatctcaagacagctgtgttcttttgtaggaagaatcctggcctagatttgtatcatagctctaccactcattagctccctgaccttggggaa<br>gtctcttcatttttctgaatttcatctatgtagataatccttcagaaggttataatgaaaattaaatgaaattctatgagattagggaggggggaggga<br>tagcattaggagatatacttaatgtgaatgatgagttaatggatgtagcacaccaacatggcacatgtatacatatgtaacaaacctgcacgttgtg<br>cacatgcacccctagaacttaaagtataatttaaaaaagaaaagaaattctatgagattaataagctatatgatgtaatacatggctcttgtatattca<br>tgaactcttcacttagctctttggcttgtgaatattatgtacatcaaaatttaatttttcatttgatctattttactagactcctgccccatctagtctacct<br>gtccacattattaccacattctagtccatcttgcccattactaccaggctaagctttctagtgtggatatgtcatcatcttattttccttagaattttagcg<br>atctttttatcatttccaagataaacacttgcctaggtgtacagcatccttgtttaccatcatactcacgcattagagatttagccttccctttaaaatct<br>agggtcactcctctgtaggaagacttttgggcagttttttatttttgctacttctgacaccatcctttaatgttttaatattagtgccacagagttcttttgtga<br>ctttaccattatgtaagaatcttccacttggaatgtgtttctcttcctcacaccccagtctgcctagcaaatgccacttgatcccaagtatcagcttgtta<br>gcttctcagtgaagcaagccttctctattttagcagttatcacagtgtattttaattgtttacatatctactttcacaatgggttataaatttcttaaggtc<br>aaggagttggctattttaatctttgcattatcagttcatttcagatagtgaacatttaatacgttaattaaaggaataatttacatttaagccaaacgtga<br>agataaactattgctcatcatcccctcttcagccgtatcctgtaggtggtatcaccttatattcttaccaccaaagaaaatatggcccctctcttagaaa<br>gatcttaatcatttatctgtgtatgtttaggagtatccttagatcatgcctcacatatigatgccaaagagttcttttgtgccaatttcataatgtgtgtca<br>gcacaacaattctgaagatttgttggtgtctttcatgtacttgactacaaattgccttgccattactactcttctcaaaggatatctgaaattctttttttc<br>ttttttttttgaagatggagtctcactgtcacccaggctggagtgcagtggcgtgatcttggctcactccatttcccgagctcaaggtgattctcatgcct<br>cagcctcccaagtagctgggactacaggtgtgcaccaccacaccgggctaatttttgtatttttagtagagacagggttttgccatgttggccaggct<br>cttgaactcccaggctcaagcgatccacccgcctcagcctcccaaagtcctgggattacaggcatgagccaccacgcccagcctggatatctgaaat<br>tcttaagtgaaattagtcaaattatcttgtactggggatttttttttaatttcaacttttattttgattcaggggatacatgcataggtttgttacatggg<br>tatatcatgtgatgctgaggtttggggtacaattgatcctgtcacccaggtagtgagcataatacccaacagttgttcaaccccttgcccctctccccta<br>gtagtcctcagtgtctattgatgccatctttatgtccacaagtaaccagtgtttagctcccacttacaagtgagaacatggagcatttggttttctgttc<br>ctgggttatctcacttaggataatggtctctggatgcatccatgttgctgcaaaggacattatttcattctttttatggttgcatactgtggattttattg<br>ggtctttattttgtattagcattttaaaaccctaaatgtgacacagtacggcatgagtgatcatgcatctcaagaaatcttgaaatgttcctgtccataa<br>agcagaatttttttaagagaccatttcacagtctcccttcccctcactgtatcaagtgctcatttgtgaattaccaatttctcttgttttgacag | 16.) rs10873142 |
| ENSE00000658287 62,203,607-62,203,827 (221 bp) | |
| TGGTATTATTCAGCACGGACTTGATTTTCTCCCTTCAACAAACAGAATGTGTCCTTAAACCGGTTGAATCTTCAGAT<br>ATGAAAATGGACTCAGCTATTCACCAAAGTTGAATCAGAAGATACAAGTAGCCTCTTTGACAAACTTAAGAAGGA<br>ACCTGATGCTTTAACTTTGCTGGCCCCAGCCGCTGGAGACACAATCATATCTTTAGATTTTGGCAGCAACG | |
| Intron9-10 62,203,828-62,204,804 (977 bp) | |
| gtgagtagttatttttgttaatcccctaaattgtgtctgttgctacaagccccatttcaactaaacattactttacggtttttgttggtaatcatttggaca<br>ttacaagctaatatatgtttatagttttcttaaatgtatttgcttaaatatttttgcccccgtaatttcttaccattcttgctttttatactgttggaaattgt | |

*(Continued)*

*(Continued)*

| | |
|---|---|
| gcttcaaagtgtccttaaggtatttcttcttcccacataaatttttcctggctactctatttctgtatcctgctgtcagattttctccacagtttagcagagt<br>tatatggaagtaggcattgttgcattaaaggataaaaaagtagtcatactatataacatcaagcattgaagatgaaaactgcaattttaaagtagaga<br>acattttaatgtataaaaaggttggtattgccttttgtcttttatgccatagagattaagacggcggtatcaatagtggattgtaaaggtaactcagactt<br>atggttatactatactattgtatgtaaactttctgatgaaggaaaatttggtgacattttgttgtttgatgaattagacaaacctttgtgaaaaagaac<br>ataaatttttttatatgtgaaaatccttgtggccgggcgcagtggctcacgcctgtaatcccagcactttgggaggccgaggcgggtggatcacttgag<br>gttaggagttcgagaccagcctggccaccatggtgaaaccccgtctctaccaaaaatacaaaagttagctgggcgtggtggtgtgcgcctgtaatcc<br>cagctacttgggaggctgaggcaggggaattgcttgaacctgggaggcagaggttgcagtgagccaagattgcgccattgcactccagcctgggca<br>acagagcaagactctgtcttgggtaaaaaaaaaaaaaaatccttctatactttagattgactcatattttttccccacag | |
| ENSE00000658289 62,204,805-62,205,091 (287 bp) | |
| ACACCAGAAACTGATGACCAGCAACTTGAGGAAGTACCATTATATAATGATGTAATGCTCCCCTCACCCAACGAA<br>AAATTACAGAATATAAATTTGGCAATGTCTCCATTACCCACCGCTGAAACGCCAAAGCCACTTCGAAGTAGTGCT<br>GACCCTGCACTCAATCAAGAAGTTGCATTAAAATTAGAACCAAATCCAGAGTCACTGGAACTTTCTTTTACCATGC<br>CCCAGATTCAGGATCAGACACCTAGTCCTTCCGATGGAAGCACTAGACAAAGTTCACCTGAG | 17.) rs41508050 |
| Intron10-11 62,205,092-62,207,222 (2,131 bp) | |
| gtaggtgtcatgatataatcagaaagggacaactttcagattttaacattcaagaatgtatttataagtttgattcaaacactatttgaaccacaaat<br>tacatttgtgtgtgtgtttgaattttagcactttaaaattattgcaagagctactgcctaacctagacctgagcacatgttttaggctcaaagatagtca<br>ggaacatgggaagaaactagcttaatataaaccaaaaggtgaaacgtacattgtttctctattatttatatcagtaggacaaaaacatcttgaatttg<br>gacatttaaagagaatagtactaagtgtgctcaaggtagctacagcctatacctgttacccctttagtttgttttattgtgttttgttttgttttgagaaa<br>gagtctcactatcacccaggctggagtgcagtggtgcaatcacagcctcaacctcccaggctcaaatgattctcccacctcagcctcccaagtagct<br>gggactacaggcctgcatcaccatgcctggctaatttttttaagctttttttgtgtgtgtgtgtggagttcgggttctcactatgttgctcaggctggtttta<br>aactcctgggctcaagcgatcctcctgccttggcctcccaaagtactaggattacaggcgtgagctaccatgcctggcccattacccctttgagttgg<br>agaactgtctggtagcaatagacttacgagggtttaaatgggaaaggaccttataaattctttgcccaatttagtctaatttccatcactattttgaaa<br>ttttgggtaagtataatatgaaaataacaagtgttacataaaataaatacttagtaactggtctttttttattctggatctgtcttgatattaattgtccta<br>tgaacacaaaaataatctttaaaggctaggctggccaagacttagagatatcacacaggctctatttctaaatctagaatgattccatttttagggct<br>tcctacatctaaaaatatgctcaggagtagggcaacttagatctgaacattataacttgataaatgaggcataaataagctttaataagtggtaaat<br>aattctacattaggtatttgttgaataaaactgacaagctaagagtaggggatttgacatctcacagccttgtgttgaatgaatatatatcctatgctc<br>tggttgcttaatttacccagaaaaaaaaatgtttgattcatcttggtttttatctaacaaaagtaaatctaacaaaaacgttagaatgaggaaagcaa<br>aatttcttgtttagaatacacagctatagtttttgttaaacttcttgcccagaactcttaaaatagtaataatgtacattcgttcaggtatatgcaggta<br>aaataacttaggtttctactcccaccccgcgacagtaacagtgagatttttaggtagctcagtcaccacgaggagtgtgccttctcagttcaaaggtaa<br>attccagtgaatgtagcatctagttaattggtcaattaggtaccattgtgggatgtgaattaccaaataggttttattctttagaataaggtgtttctttt<br>catctcaattttgtaaatgatgttatattacatagtcagaaatatatatattggcaaaattagttaccagtataagcttcaaaatgtcactattttcaca<br>aatttttttttttttttttttttttgacatggagtctcactctgtcgccaggctggagtgcagtggcatgatcttggctcactgcaacctctgcctcccaggt<br>tcaagtgattctcctgcctcagcctcctgagtagctgggattacaggcgtgccaccatgcctagctaatttttgtatttttagtagagacgaggtttc<br>accatgttggccaggatggtctcgatctcttgacctcattatccctccaccttggcttcccaaagtgctgggattacaggcgtgagccactgagcccgg<br>cctagttaaataaaatttgatataaacacgatggacttggttgtgtgttttctggtttttctgagatctagttgtgaaaattctgacaactagcaaagtatat<br>ggaagcttcttcaggaaatagtaaacatatttcttttttacag | 18.) rs2301113 |
| ENSE00000658291 62,207,223-62,207,345 (123 bp) | |
| CCTAATAGTCCCAGTGAATATTGTTTTTATGTGGATAGTGATATGGTCAATGAATTCAAGTTGGAATTGGTAGAA<br>AAACTTTTTGCTGAAGACACAGAAGCAAAGAACCCATTTTCTACTCAG | |
| Intron11-12 62,207,346-62,207,472 (127bp) | |
| gtatatgaacttatttgttttatattaaatttcattaatttttagtctgaagtgactttgagtttcacttgtttttttatttataaggtgtggccattgtaaaaa<br>ctcatgtatttgctgttttaaag | |
| ENSE00000658293 62,207,473-62,207,906 (434 bp) | |
| GACACAGATTTAGACTTGGAGATGTTAGCTCCCCTATATCCCAATGGATGATGACTTCCAGTTACGTTCCTTCGAT<br>CAGTTGGTCACCCATTAGAAAGCAGTTCCGGCAAGCCCTGAAAGCGGCAAGTCCTCAAAGCACAGTTACAGTATTC<br>CAGCAGACTCAAATACAAGAACCTACTGCTAATGCCACCACTACCACTGCCACCACTGATGAATTAAAAACAGTG<br>ACAAAAGACCGTATTGGAAGACATTAAAATATTGATTGCATCTCCATCTCCTACCCACATACATAAAGAAACTACT<br>AGTGCCACATCATCACCATATAGAGATACTCAAAGTCGGACAGCCTCACCAAACAGAGCAGGAAAAGGAGTCAT<br>AGAACAGACAGAAAAATCTCATCCAAGAAGCCCTAACGTGTTATCTGTCGCTTTGAGTCAAAG | 19.) rs41492849<br>20.) rs34005929<br>21.) rs11549465<br>22.) rs11549467<br>23.) rs199775054<br>24.) rs61755645 |
| Intron12-13 62,207,907-62,211,422 (3,516 bp) | |

*(Continued)*

*(Continued)*

| | |
|---|---|
| gtatttatatgtaacattcaagttatagttctttttattattttgagataaatgtatgtgatagtacatgatttttaaacttatagcaaactttctgatata<br>tatgccctaacgcaaattcttgagaactcaaaaaactttctaaattaacctcatatatttttttctttttctttctttttttttttttgagacagagtctcgct<br>ttgtcgcccaggctggagtgcaatggcatggcaccatctcagctcacggcaacctctgcctcctgggtgcaagagattctcctgcctcagcctcccga<br>gtagctgggattacaggcatgcaccaccacgcccggctgatttttttggtatttttcatagagacagggtttctccacgttggtcaggctggtctcaaa<br>ctcccgacttcaggtgatccgcctgcctcagcctccgaaagagctgggattacaggtgtgagccaccatgcccgctcctatttttttctaaaataattat<br>aaattctaaaattacctatctaaatggaggagggtcttctgacacctttaaaataaaatccagctcagtactgtaaatgtgtttacagaacttgtttaa<br>agttcttacagttgtttaaatcagactagttaactaccctcactacttagatgcttccatttcttagagctctttttaagcttatctgaagaaaagcccctt<br>ccaatttaagggttatttccaattgcacattccaaattgagccttccatcttcagcattcaatatagatatttacaggcccctcttttaaaattttattata<br>gttaacttgtattaaagttgcttttatttttcattacgtatttgtagaacattagctatatatatattgcaggctacataggttttcaaactgtacaacag<br>gaatctaagcatgaattgttacttctatggagctagttgaaacaaacatatggacatgacccaatttttaagttatactttctgtatataatttgtaagg<br>ggatttcacatattttaagtttgaggctatagctagaagaaattaagttttatctaataagtgtgtggaaaagggaaatgattccttctctactatgtct<br>agactaagccagatatcaatagcaataggaaagaaccactgtcgtagccagaacacatagctttttccctgcctaacattccccaccttgacctaga<br>gtgctgggagaggtcttttccctaagcttggaaaagacattggggctttagatgaactcagaagtactttacattacttttatttactgtgtcacttactc<br>acttttgactctgagctccacggagggcaatcacagtgtcttgggcattttagtgatactaatacttagctcatgacctaatgtgtagtacttcctcaata<br>aatggttgttgaggcagggcgcagtggctcatcactgtaatctcagcactttgggaggctgaagcgggtggatcacctgaggccaagagtttcagac<br>cagcctggccaacatggtgaaacccggtctactaaaaatgcaaaaattagctgggcgtggtggcacgtgcctgtaatcccagctactttgggaggtt<br>gaggcaggagaattgcttgaaccctggaggtggaggctgcagtgagccgagatcgtgccattgcactccagcctgggcgagaagagtgaaactcg<br>gtttcaaaaaaaaaaaaaaaaaaaaaagttgttggactgacagatgcatgaatacagtagtaaaaatgacaatcacttataagttacagtttacta<br>tcagctacagaggatgggatatccagttttctgaacaactgttctcttgtacttgtcaaagccaaagtgtaacaacacatcaagtcactttagcaattt<br>atttttgagacggagttttgctcttgttgcccaggctggagtgcaatggcgtgatctcggatcttggctcaccgcaactgccgcctcccctggttcaagc<br>aaatctcgtgcctcagcctcccgagtagctgggattacaggcatgcaccaccacacccagctaattttgtatttttagtagagacgagtgtttctccagg<br>ttgatcaggctggtctcaaactcccgacctcaggtgatccacctgcctcagcctcccaaagtgctgggatgacagttgtgagccactgtgcccagcta<br>gcaactgtttttaaacattagttccaatgtagtgtacactgaaaacttttatgaaaggaatttcaaaaattaagataaaccattaaaaacgtaattac<br>taagtactactactactacaatgatatttacataatagactgagttacatttcataaagacaatatatctgtataagaatttttaaacttccctgtctat<br>ataatagaagttttagagaaatttttttaaaaaccaaagaaaactgcaaaataagatcacttacctattggcattctcaactgtctggaacagcaag<br>gagccattatgattatgcatttggtttgtggggtgtcttgaaaagtcaaaataatgtaacaaagctgatgtactttactcattagaacaattcttcaca<br>atttaatattaattttagatatacatagttcatgtttgataaccagatcaatactgagtgaaaaatagcatagtgggaagagcaggggaggggaggt<br>agggatctggagacctagagtgtacttccatattgcaactagtgagcagtaggactttgagaaagttacccaataggcctcagggttctaatttataa<br>aatgggtatgatatgcctgccttatctgtcttgggaacttaagtaaggttaaaatgaactaatgaacttgaaatgttttataaactgaaaatgctatac<br>gaatgtgagattgatcttgtatttcaatagtcccaacaatatcactgcattgttatattaggtggaataaaaggacaatatttaactgttttgactctac<br>aatagtgtcaatttagttgtgttcagctctattttataaaaatagggatacgcatactgtagaaaatttcctgttaaattaagctttgacggccaggtgct<br>cacgcctgtaatcccagcactttgggaggccaaggtgggcagatcacttgcgctcaggagtttgagaccagcctgagcaacatagtgaaatcctgtc<br>tctacaaaaatatgtatatataaattagtcataatcccagctacttgagaggctgaggtgggaggatcacttgattccagaggcagggcttggttgc<br>agtaagcagagatcacgtctctgcactccagcctggctgacagagtaagaccgtgtttcaccaaaaaaaaaaaaaaaaaaattaagcttttacttttta<br>agatgataaactttagtgatcaggaaagttatcttatgtatattatattccttaatattggagaactaaagaattatgtatttttcttttaaaagcgctcac<br>tggatatttttttttaaaaaacgctatattttcatttagaatttttttctttttcag | 25.) rs4902080 |
| ENSE00000658295 62,211,42362,211,531 (109 bp) | |
| AACTACAGTTCCTGAGGAAGAACTAAATCCAAAGATACTAGCTTTGCAGAATGCTCAGAGAAAGCGAAAAATGG<br>AACATGATGGTTCACTTTTTCAAGCAGTAGGAATT | |
| Intron13-14 62,211,53262,212,408 (877 bp) | |
| gtaagtatgagtagtaggttttgcttttctagctaatgtgctatttcgtgtgtgtgtgtgtgtgtgtgtgtgtttccacgtttcttccaaatag<br>taaagttatattttcagaagttatacattgggttttttttactctgtatgcactggtttttaaaaatacaaatgtttaatacatacattcttggtataaaaat<br>tccaaacaattccagtgtatttgagttaaaaagtgaagttctccccttactccaccctgaatatcaccaccaatctcattctcttccctttaagttactt<br>tgccttattaaaagaactgctattggccaggcacagtgcctcacgcctgtaatccgagcactttgggaggccaagatgaggatcacttgaggtcagg<br>agttcgagaccagcctggccaacttggtgaaaccctgtctctactaaaaatacaaaaattagccaggcgtgttggtgcacaactataatcccagcca<br>ctctggaggctgaggcaggagaatagcttgaacccgggaggtggaggttgcgatgagctgagatcaggccactgcactccagcctgggtaagaga<br>gtgagagtccatctcatatttaaaaaagaactgctatgttttggggtaagtcaatggtggtataatacattctgatattttcaaactaaattaactgga<br>aagtatttatagacagaatggtcataatggatgacaaataacttaagaaagaaattcaaaataatttagggtagtatttaagaaactgcctataatgt<br>tattaaatttacaccaatttcaaggtttttggttgtttaaaaaaaaaattcaacaaactaaacttgaaataactttactgtttatag | 26.) rs10645014 |
| ENSE00000658297 62,212,40962,212,535(127) | |
| GGAACATTATTACAGCAGCCAGACGATCATGCAGCTACTACATCACTTTCTTGGAAACGTGTAAAAGGATGCAAA | |

*(Continued)*

*(Continued)*

| | |
|---|---|
| TCTAGTGAACAGAATGGAATGGAGCAAAAGACAATTATTTTAATACCCTCTG | |
| Intron14-1562,212,53662,213,651(1,116) | |
| gttagtttattcttttgaccttgaacatcacaaagacaaaatacatgaaacatttttatttaggagctttaatctaagtgagaatgactttggttcctta gcaagattaaaaagtaaagttgtggctgggcgcggtggctcacacctgtaatcccagcactttgggaggccgaggcagccagatcatctgaggtc aggagttggagaccagcctggccaccatggtgaaaccccgtctctactaaaaatacaaaaattagctgggcgtggtggcgggcgcctgtaatccca gctacttgggaggctgaggcatgagaattgcttgaacccggaaggcagaggttgcagtgagccaagatggcaccactgcactccagcctgggcga caagggtgagactctgcctcaaaaaaaaaaaaaaaaaaaagtacagttgtatttcatgtgatggtcttaatacagagattaacatttcaaggtgga gcttttcatttttagtaatttttctttgatttctctatgtccatgtgctgtcaatattgatagaagctgaaatttgtgaacttttatgacttctttttttttttt ttttttgagacagggtctcgctctgttgcccaggcctggagtgcagtggcatgatcatagctcactgcagtctcaaactcctgtgctcaagctcaagc aatcatcctacctcagcctcctgagtagctcgcactacagacatgcctcaccacacccggttgctttttgtagagatggggtctcactatgttgcctag gctggtttcaaactcctggcctcaagtgatcctcctgcctcagcctgtgctaggattacaggcatcagctttgatgccaccatatttatgccttttcca aattgttatttctttgtgcctttattgtatcctgtaaacatttctgacacagcaacagtatcactggattatacttacttttaacatagttgtggttttgcc aggtaaactaaaaaccctccagaatttttgctttatttttctatgatacctaacacattgtgggtgtttaataaatattgattgactagatgaatgtatac ttaggtatctcttttgttttttcag | 27.) rs4902082 |
| ENSE00001660793 62,213,65262,214,976(1,325) | |
| ATTTAGCATGTAGACTGCTGGGGCAATCAATGGATGAAAGTGGATTACCACAGCTGACCAGTTATGATTGTGAA GTTAATGCTCCTATACAAGGCAGCAGAAACCTACTGCAGGGTGAAGAATTACTCAGAGCTTTGGATCAAGTTAA CTGAGCTTTTTCTTAATTTCATTCCTTTTTTTGGACACTGGTGGCTCATTTACCTAAAGCAGTCTATTTATATTTTCT ACATCTAATTTTAGAAGCCTGGCTACAATACTGCACAAACTTGGTTAGTTCAATTTTGATCCCCTTTCTACTTAATT TACATTAATGCTCTTTTTTAGTATGTTCTTTAATGCTGGATCACAGACAGCTCATTTTCTCAGTTTTTTGGTATTTAA ACCATTGCATTGCAGTAGCATCATTTTAAAAAATGCACCTTTTTATTTATTTATTTTTGGCTAGGGAGTTTATCCCT TTTTCGAATTATTTTTAAGAAGATGCCAATATAATTTTTGTAAGAAGGCAGTAACCTTTCATCATGATCATAGGCA GTTGAAAAATTTTTACACCTTTTTTTTCACATTTTACATAAATAATAATGCTTTGCCAGCAGTACGTGGTAGCCACA ATTGCACAATATATTTTCTTAAAAAATACCAGCAGTTACTCATGGAATATATTCTGCGTTTATAAAACTAGTTTTTA AGAAGAAATTTTTTTTGGCCTATGAAATTGTTAAACCTGGAACATGACATTGTTAATCATATAATAATGATTCTTA AATGCTGTATGGTTTATTATTTAAATGGGTAAAGCCATTTACATAATATAGAAAGATATGCATATATCTAGAAGG TATGTGGCATTTATTTGGATAAAATTCTCAATTCAGAGAAATCATCTGATGTTTCTATAGTCACTTTGCCAGCTCA AAAGAAAACAATACCCTATGTAGTTGTGGAAGTTTATGCTAATATTGTGTAACTGATATTAAACCTAAATGTTCTG CCTACCCTGTTGGTATAAAGATATTTTGAGCAGACTGTAAACAAGAAAAAAAAAATCATGCATTCTTAGCAAAAT TGCCTAGTATGTTAATTTGCTCAAAATACAATGTTTGATTTTATGCACTTTGTCGCTATTAACATCCTTTTTTCATG TAGATTTCAATAATTGAGTAATTTTAGAAGCATTATTTTAGGAATATATAGTTGTCACAGTAAATATCTTGTTTTTT CTATGTACATTGTACAAATTTTTCATTCCTTTTGCTCTTTGTGGTTGGATCTAACACTAACTGTATTGTTTTGTTACA TCAAATAAACATCTTCTGTGGACCAGG | 28.) rs2057482 |
| UCSC>hg19_dnarange=chr14:62219025-622194255'pad=2003'pad=200strand=+repeatMasking=none | |
| .....atatttatcacaaagggtcatgctttttacaaagtagagaaacactgtgtaaagtcagacaagtagattaggattctgtgaagtaaaa tctaattttaaagatctttgaaatattaacattaaaaaccctgttttatgtagagatcctttccaaatatgactggataatattatttcctctgaaatgga ttttttaaaaacggttctacatacaactgctatgacaaaatctttcataggcgtacagaaacaagctggacagaagtcataaactgttgcctgactcc ctaatctgctggtagaatattaaagcaaaaacctaaagaaagtgagctacaaatacggaaaattagaaaggagccaatacagcctgggaggtact ttgcatcagcctttgcctct..... | 30.) rs17099207 |
| UCSC>hg19_dnarange=chr14:62228913-622291135'pad=2003'pad=200strand=+repeatMasking=none | |
| .....ccaaagaagggtggatctcagcactctgggtgaaagcaatctcagccagtaaaacatcagaacgctttcacgcacgcgcagtaagc acacgaccggtggcaagagcccggataaaggcggagctaaaggtgccggataatgcatcacggagcgcgtgtgcgcacgccgtagggggcgggc cgttcgggcttggtttcctgggcgaccaccgctggctagtccgttagaggcgtgcgggcttcggaggcgtgcgggcttcgggtgccatggggactcc tcccggcctgcagaccgactgcgaggcgctgctcagccgcttccaggagacggacagtgtacgcttcgaggacttcacggagctctggagaaacat gaagttcgggactatcttctggtgggtg..... | 30.) rs17099207 |

# ANNEX F:

The list of *HIF1A* SNPs for which variable associations with phenotype were reported: observed association (+), no association (-) or opposing results (underlined)

| # | SNP ID/substitution | Genomic location | Association to diseases and phenotypes |
|---|---|---|---|
| **1** | rs1957757 [T>C] | intron 6 | |
| | | (+) | **NON-CANCEROUS disease:**<br>• symptom-limited exercise test duration (Sarzynski et al., 2010) |
| | | (-) | **NON-CANCEROUS disease:**<br>• systematic sclerosis (Wipff et al., 2009)<br>**CANCER:**<br>• prostate cancer risk (Jacobs et al., 2008) |
| **2** | rs10873142 [C>T] | intron 8 | |
| | | (+) | **NON-CANCEROUS diseases:**<br>• coronary artery disease with stable exertional angina (Hlatky et al., 2007)<br>• idiopathic osteonecrosis of the femoral head in men (Hong et al., 2007) |
| | | (-) | **NON-CANCEROUS diseases:**<br>• acute myocardial infarction (Zheng et al., 2009)<br>• frequent intradialytic hypotension (Zheng et al., 2009)<br>**CANCER:**<br>• lung cancer (Konac et al., 2009) |
| **3** | rs41508050 [C>T]<br>*p.Thr418Ile* | exon 10 | |
| | | (+) | **NON-CANCEROUS diseases:**<br>• coronary artery disease with stable exertional angina (Hlatky et al., 2007) |
| | | (-) | **CANCER:**<br>lung cancer (Konac et al., 2009) |
| **4** | rs11549465 [C>T]<br>*p.Pro582Ser* | exon 12 | |
| | | (+) | **CANCERS:**<br>• glioma (Xu et al., 2011)<br>• oral cancer risk (together with rs11549467) (Chen et al., 2009)<br>• esophageal squamous cell carcinoma (↑size and ↑lymph node metastasis (Ling et al., 2005))<br>• head and neck squamous cell carcinoma (Tanimoto et al., 2003)<br>• renal cell carcinoma (Ollerenshaw et al., 2004)<br>• lung cancer (NSCLC) (Koukourakis et al., 2006)<br>• breast cancer risk (poor prognosis) (Kim et al., 2008; Lee et al., 2008; Naidu et al., 2009)<br>• CRC risk and metastasis (Kang et al., 2011; Kuwai et al., 2004), ulcerative CRC (Fransen et al., 2006)<br>• prostate cancer risk/susceptibility (Foley et al., 2009; Jacobs et al., 2008; Orr-Urtreger et al., 2007); androgen-independent prostate cancer (increased transcription activity) (Chau et al., 2005) (Fu et al., 2005)<br>• pancreatic cancer (Ruiz-Tovar et al., 2012)<br>• cervical and endometrial cancers (Konac et al., 2007)<br>• increased female specific cancer risk (meta-analysis) (Zhao et al., 2009) |
| | | | **NON-CANCEROUS diseases:**<br>• T2DM (T allele with protective effect / decreased T2DM risk) (Nagy et al., 2009; Yamada et al., 2005) |

*(Continued)*

*(Continued)*

| # | SNP ID/substitution | Genomic location | Association to diseases and phenotypes |
|---|---|---|---|
| | | | • T1DM (Nagy et al., 2009)<br>• diabetic nephropathy (Gu et al., 2012)<br>• CAD with stable exertional angina (Hlatky, Quertermous et al. 2007)<br>• ischemic heart disease (collaterals formation) (Resar et al., 2005);<br>• chronic obstructive pulmonary disease susceptibility (with rs11549467) (Putra et al., 2011a)<br>• risk of cellulite (Emanuele et al., 2010)<br>• acute kidney injury (Kolyada et al., 2009)<br>• idiopathic osteonecrosis of the femoral head in men (Hong et al., 2007)<br>**PERFORMANCE:**<br>• endurance status (Doring et al., 2010; McPhee et al., 2011)<br>• maximal oxygen consumption (Prior et al., 2003)<br>• power-oriented athletes and muscle activity (Ahmetov et al., 2008; Cieszczyk et al., 2011; Gabbasov et al., 2012) |
| | | (-) | **CANCERS:**<br>• CRC risk, progression and metastasis (Knechtel et al., 2010; Szkandera et al., 2010)<br>• gastric cancer (Li et al., 2009)<br>• breast cancer risk (Vleugel et al., 2005; Zagouri et al., 2012); sporadic breast cancer (Apaydin et al., 2008)<br>• ovarian cancer (Konac et al., 2007)<br>• prostate cancer (Li et al., 2007; Li et al., 2012)<br>• bladder cancer occurrence (stat. sig. prognosis affect after radical cystectomy of invasive cancers) (Nadaoka et al., 2008)<br>• oral squamous cell carcinoma (Munoz-Guerra et al., 2009; Shieh et al., 2010)<br>• lung cancer (Konac et al., 2009; Kuo et al., 2012) (associated with lung cancer with *TP53* LOH) (Putra et al., 2011b), (NSCLC) (Kim et al., 2010)<br>• renal cell carcinoma (Morris et al., 2009) (Qin et al., 2011)<br>• hepatocellular carcinoma risk (Hsiao et al., 2010)<br>• endometrial cancer risk (potential association with tumorigenesis and increased tumor vasculature) (Horree et al., 2008) |
| | | | **NON-CANCEROUS diseases:**<br>• systematic sclerosis (Wipff et al., 2009)<br>• acute myocardial infarction and frequent intradialytic hypotension (Zheng et al., 2009)<br>• pre-eclampsia (Heino et al., 2008; Nava-Salazar et al., 2011)<br>• erythrocytosis (Percy et al., 2003)<br>• giant cell arteritis (Torres et al., 2010)<br>• peripheral artery disease (Bahadori et al., 2010)<br>• ischemic heart disease (coronary artery collaterals) (Alidoosti et al., 2011)<br>**PERFORMANCE:**<br>• power-oriented athletes (pos. association in interaction with *ACTN3* R577X polymorphism) (Eynon et al., 2010)<br>• response to acute hypoxia (Hennis et al., 2010)<br>• acute mountain sickness (Droma et al., 2008) |

*(Continued)*

*(Continued)*

| # | SNP ID/substitution | Genomic location | Association to diseases and phenotypes |
|---|---|---|---|
| **5** | rs11549467 [G>A] *p.Ala588Thr* | exon 12 | |
| | | (+) | **CANCERS:**<br>• oral squamous cell carcinoma (marker of disfavorable prognosis at early stages) (Chen et al., 2009; Munoz-Guerra et al., 2009)<br>• head and neck squamous cell carcinoma (Tanimoto et al., 2003)<br>• decreased breast cancer risk (Zhao et al., 2009)<br>• gastric cancer (Li et al., 2009)<br>• ulcerative CRC (together with rs11549467) (Fransen et al., 2006)<br>• hepatocellular carcinoma risk (Hsiao et al., 2010)<br>• renal cell carcinoma (Ollerenshaw et al., 2004)<br>• pancreatic cancer (Ruiz-Tovar et al., 2012)<br>• prostate cancer (Li et al., 2012) |
| | | | **NON-CANCEROUS disease:**<br>• chronic obstructive pulmonary disease susceptibility (with rs11549465) (Putra et al., 2011a) |
| | | (-) | **CANCERS:**<br>• oral squamous cell carcinoma (Shieh et al., 2010)<br>• breast cancer risk (Apaydin et al., 2008; Kim et al., 2008; Naidu et al., 2009)<br>• lung cancer (Konac et al., 2009), (associated with adenocarcinoma with 1p34 LOH) (Putra et al., 2011b)) and lung carcinoma (NSCLC) (Kim et al., 2010; Koukourakis et al., 2006; Kuo et al., 2012)<br>• CRC risk, progression and metastasis (association with tumor location and size (Knechtel et al., 2010)) (Szkandera et al., 2010)<br>• prostate cancer (Li et al., 2007) (Orr-Urtreger et al., 2007) (Chau et al., 2005)<br>• ovarian, cervical and endometrial cancers (Konac et al., 2007)<br>• renal cell carcinoma (Qin et al., 2011) (Morris et al., 2009)<br>• bladder cancer (stat. sig. prognosis affect after radical cystectomy of invasive cancers) (Nadaoka et al., 2008) |
| | | | **NON-CANCEROUS diseases:**<br>• T2DM (Yamada et al., 2005)<br>• pre-eclampsia (Heino et al., 2008; Nava-Salazar et al., 2011)<br>• peripheral artery disease (Bahadori et al., 2010)<br>• giant cell arteritis (Torres et al., 2010)<br>• coronary artery disease with stable exertional angina (Hlatky et al., 2007)<br>**PERFORMANCE:**<br>• endurance status (Doring et al., 2010) |
| **6** | rs199775054 [G>C] *p.Ala593Pro* | exon 12 | |
| | | (+) | **CANCER:**<br>• hepatocellular carcinomas (Park et al., 2009) |
| | | (-) | **CANCERS:**<br>• colon, gastric, breast, lung, and acute leukemia (Park et al., 2009) |
| **7** | rs10645014 [insGT] | intron 13 | |
| | | (+) | **CANCER:**<br>• lung carcinoma (NSCLC) (Koukourakis et al., 2006) |
| | | possible (+) | **PERFORMANCE:**<br>• adaptation to living at high altitude (Suzuki et al., 2003) |
| | | (-) | **CANCER:**<br>• lung cancer (Konac et al., 2009) |

*(Continued)*

*(Continued)*

| # | SNP ID/substitution | Genomic location | Association to diseases and phenotypes |
|---|---|---|---|
| **8** | rs2057482 [T>C] | 3'-UTR | |
| | | (+) | **NON-CANCEROUS disease:**<br>• idiopathic osteonecrosis of the femoral head in men (Hong et al., 2007) |
| | | (-) | **CANCERS:**<br>• breast cancer risk (Lee et al., 2008)<br>• rectal cancer risk (Frank et al., 2010)<br>• renal cell carcinoma (Qin et al., 2011)<br>• prostate cancer (Li et al., 2012)<br>**NON-CANCEROUS diseases:**<br>• acute myocardial infarction (Zheng et al., 2009)<br>• frequent intradialytic hypotension (Zheng et al., 2009)<br>• coronary artery disease with stable exertional angina (Hlatky et al., 2007) |

# ANNEX G:

*HIF1A* SNPs with observed phenotype associations and no association with investigated phenotypes

| # | SNP ID/substitution | Genomic location | Association to diseases and phenotypes |
|---|---|---|---|
| • **ASSOCIATION** | | | |
| 1 | rs2783778 [C>T] | 5'-flanking | • acute myocardial infarction and frequent intradialytic hypotension (Hemodialysis patients) (Zheng et al., 2009) |
| 2 | rs7148720 [T>C] | 5'-flanking | • acute myocardial infarction and frequent intradialytic hypotension (Hemodialysis patients) (Zheng et al., 2009) |
| 3 | rs1535679 [A>C] | 5'-flanking | • idiopathic osteonecrosis of the femoral head in men (Hong et al., 2007) |
| 4 | rs28708675 [A>T] | 5'-flanking | • maximal oxygen consumption (Prior et al., 2003) |
| 5 | c.347T>A *p.Val116Glu* | exon 3 | • renal cell carcinoma (influence on transcriptional activity of HIF1A) (Morris et al., 2009) |
| 6 | rs12434438 [G>A] | intron 6 | • systemic sclerosis (Wipff et al., 2009)<br>• T2DM (together with rs1319462) (Yamada et al., 2005) |
| 7 | rs1319462 [G>A] | 3'-flanking | • T2DM (together with rs12434438) (Yamada et al., 2005) |
| • **NO ASSOCIATION** | | | |
| 1 | rs41362550 [T>C] | 5'-flanking | • CAD with stable exertional angina (Hlatky et al., 2007) |
| 2 | rs7143164 [G>C] | intron 1 | • acute myocardial infarction and frequent intradialytic hypotension (Hemodialysis patients) (Zheng et al., 2009)<br>• prostate cancer (Jacobs et al., 2008) |
| 3 | rs1951795 [A>C] | intron 1 | • elite endurance status (elite endurance athletes (EEA)) (Doring et al., 2010) |
| 4 | rs12435848 [A>G] | intron 1 | • prostate cancer risk (Jacobs et al., 2008) |
| 5 | c.111C>A *p.Ser28Tyr* | exon 2 | • T2DM (Yamada et al., 2005)<br>• ovarian, cervical and endometrial cancers (Konac et al., 2007)<br>• sporadic breast cancer (Apaydin et al., 2008)<br>• breast cancer (Naidu et al., 2009) |
| 6 | rs4899056 [T>C] | intron 4 | • prostate cancer (Jacobs et al., 2008) |
| 7 | rs11158358 [G>C] | intron 6 | • prostate cancer risk (Jacobs et al., 2008)<br>• elite endurance status (elite endurance athletes (EEA)) (Doring et al., 2010) |
| 8 | rs2301111 [G>C] | intron 7 | • prostate cancer risk (Jacobs et al., 2008) |
| 9 | rs2301113 [C>A] | intron 10 | • acute myocardial infarction and frequent intradialytic hypotension (Hemodialysis patients) (Zheng et al., 2009)<br>• prostate cancer risk (Jacobs et al., 2008)<br>• elite endurance status (elite endurance athletes (EEA)) (Doring et al., 2010) |
| 10 | rs41492849 [C>T] *p.Pro564=* | exon 12 | • OSCC (Shieh et al., 2010)<br>• lung cancer (Konac et al., 2009) |

*(Continued)*

*(Continued)*

| # | SNP ID/substitution | Genomic location | Association to diseases and phenotypes |
|---|---|---|---|
| 11 | rs34005929 [G>A] *p.Leu580=* | exon 12 | • pre-eclampsia (Heino et al., 2008)<br>• OSCC (Shieh et al., 2010)<br>• lung cancer (Konac et al., 2009) |
| 12 | rs61755645 [A>T] *p.Thr600=* | exon 12 | • pre-eclampsia (Heino et al., 2008)<br>• OSCC (Shieh et al., 2010) |
| 13 | rs4902080 [T>C] | intron 12 | • CAD with stable exertional angina (Hlatky et al., 2007) |
| 14 | rs4902082 [C>T] | intron 14 | • maximal oxygen consumption (Prior et al., 2003) |
| 15 | rs17099207 [G>A] | 3'-flanking *HIF1A* and 5'-UTR *SNAPC1* | • elite endurance status (elite endurance athletes (EEA)) (Doring et al., 2010) |

OSCC, oral squamous cell carcinoma; CAD, coronary artery disease; RCC, renal cell carcinoma; CRC, colorectal cancer; T2DM, diabetes mellitus (type 2); T1DM, diabetes mellitus (type 1); NSCLC, non-small-cell lung cancer;

# ANNEX H:

Validated TFBS within promoter region of *HIF1A*

| # | TFBS | HGNC name | Location | SNP overlap |
|---|---|---|---|---|
| 1 | SMARCC1 | SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily c, member 1 | 3p21.31 | rs41362550 |
| 2 | CTCF | CCCTC-binding factor (zinc finger protein) | 16q21-q22.3 | rs41362550 |
| 3 | POLR2A | polymerase (RNA) II (DNA directed) polypeptide A, 220kDa | 17p13.1 | rs41362550 |
| 4 | POLR2A | polymerase (RNA) II (DNA directed) polypeptide A, 220kDa | 17p13.1 | rs41362550 |
| 5 | SMARCB1 | SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily b, member 1 | 22q11.23 | rs41362550 |
| 6 | TBP | TATA box binding protein | 6q27 | rs41362550 |
| 7 | STAT3 | signal transducer and activator of transcription 3 (acute-phase response factor) | 17q21 | / |
| 8 | SMC3 | structural maintenance of chromosomes 3 | 10q25 | / |
| 9 | REST | RE1-silencing transcription factor | 4q12 | / |
| 10 | RAD21 | RAD21 homolog (S. pombe) | 8q24.11 | rs41362550 |
| 11 | GATA3 | GATA binding protein 3 | 10p15 | rs41362550 |
| 12 | CTCF | CCCTC-binding factor (zinc finger protein) | 16q21-q22.3 | / |
| 13 | E2F1 | E2F transcription factor 1 | 20q11 | rs41362550 |
| 14 | CTCF | CCCTC-binding factor (zinc finger protein) | 16q21-q22.3 | rs41362550 |
| 15 | HEY1 | hairy/enhancer-of-split related with YRPW motif 1 | 8q21 | rs41362550 |
| 16 | NFKB1/RELA | nuclear factor of kappa light polypeptide gene enhancer in B-cells 1 / v-rel reticuloendotheliosis viral oncogene homolog A (avian) | 4q24 / 11q13 | rs41362550 |
| 17 | TAF1 | TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250kDa | Xq13.1 | rs41362550 |
| 18 | TCF7L2 | transcription factor 7-like 2 (T-cell specific, HMG-box) | 10q25.3 | rs41362550 |
| 19 | ATF3 | activating transcription factor 3 | 1q32.3 | rs41362550 |
| 20 | ETS1 | v-ets erythroblastosis virus E26 oncogene homolog 1 (avian) | 11q23.3 | rs41362550 |
| 21 | HDAC2 | histone deacetylase 2 | 6q21 | rs41362550 |
| 22 | PAX5 | paired box 5 | 9p13.2 | rs41362550 |
| 23 | USF1 | upstream transcription factor 1 | 1q22-q23 | rs41362550 |
| 24 | USF1 | upstream transcription factor 1 | 1q22-q23 | rs41362550 |
| 25 | NFYA | nuclear transcription factor Y, alpha | 6p21.3 | rs41362550 |
| 26 | MAX | MYC associated factor X | 14q23 | rs41362550 |
| 27 | MYC | v-myc myelocytomatosis viral oncogene homolog (avian) | 8q24 | rs41362550 |

*(Continued)*

*(Continued)*

| | | | | |
|---|---|---|---|---|
| 28 | USF2 | upstream transcription factor 2, c-fos interacting | 19q13 | rs41362550 |
| 29 | ZEB1 | zinc finger E-box binding homeobox 1 | 10p11.22 | rs41362550 |
| 30 | MXI1 | MAX interactor 1, dimerization protein | 10q24-q25 | rs41362550 |
| 31 | E2F6 | E2F transcription factor 6 | 2p25.1 | rs41362550 |
| 32 | SIN3A | SIN3 transcription regulator homolog A (yeast) | 15q22.33 | rs41362550 |
| 33 | SP1 | Sp1 transcription factor | 12q13.1 | rs41362550 |
| 34 | IRF1 | interferon regulatory factor 1 | 5q23-q31 | rs41362550 |
| 35 | SREBF1 | sterol regulatory element binding transcription factor 1 | 17p11.2 | rs41362550 |
| 36 | HMGN3 | high mobility group nucleosomal binding domain 3 | 6q14.3 | rs41362550 |
| 37 | CCNT2 | cyclin T2 | 2q21.3 | rs41362550 |
| 38 | FOS | FBJ murine osteosarcoma viral oncogene homolog | 14q24.3 | rs41362550 |
| 39 | BHLHE40 | basic helix-loop-helix family, member e40 | 3p26 | rs41362550 |
| 40 | E2F6 | E2F transcription factor 6 | 2p25.1 | rs41362550 |
| 41 | ELF1 | E74-like factor 1 (ets domain transcription factor) | 13q13 | rs41362550 |
| 42 | NFYB | nuclear transcription factor Y, beta | 12q22-q23 | rs41362550 |
| 43 | PBX3 | pre-B-cell leukemia homeobox 3 | 9q33.3 | rs41362550 |
| 44 | TCF12 | transcription factor 12 | 15q21 | rs41362550 |
| 45 | EGR1 | early growth response 1 | 5q23-q31 | rs41362550 |
| 46 | FOSL2 | FOS-like antigen 2 | 2p23.3 | rs41362550 |
| 47 | E2F1 | E2F transcription factor 1 | 20q11 | / |
| 48 | TBP | TATA box binding protein | 6q27 | / |
| 49 | POLR2A | polymerase (RNA) II (DNA directed) polypeptide A, 220kDa | 17p13.1 | / |
| 50 | GTF2F1 | general transcription factor IIF, polypeptide 1, 74kDa | 19p13.3 | / |
| 51 | TAF1 | TAF1 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 250kDa | Xq13.1 | / |
| 52 | GTF2B | general transcription factor IIB | 1p22-p21 | / |
| 53 | POLR2A | polymerase (RNA) II (DNA directed) polypeptide A, 220kDa | 17p13.1 | / |
| 54 | POLR2A | polymerase (RNA) II (DNA directed) polypeptide A, 220kDa | 17p13.1 | / |
| 55 | REST | RE1-silencing transcription factor | 4q12 | / |
| 56 | POU2F2 | POU class 2 homeobox 2 | 19q13.2 | / |
| 57 | POU2F2 | POU class 2 homeobox 2 | 19q13.2 | / |
| 58 | HEY1 | hairy/enhancer-of-split related with YRPW motif 1 | 8q21 | / |
| 59 | TCF12 | transcription factor 12 | 15q21 | / |
| 60 | MYC | v-myc myelocytomatosis viral oncogene homolog (avian) | 8q24 | / |
| 61 | TAF7 | TAF7 RNA polymerase II, TATA box binding protein (TBP)-associated factor, 55kDa | 5q31 | / |
| 62 | POLR2A | polymerase (RNA) II (DNA directed) polypeptide A, 220kDa | 17p13.1 | / |

# ANNEX I:

Predicted TFBS overlapping with sequences centered to HIF1A SNP (50 bp)

| # of TFBS /SNP | TF name (HGNC) | SNP (allele 1) | TFBS name (TRANSFAC) | SNP (allele 2) |
|---|---|---|---|---|
| 1 | CUX1 | | V$CDP_02 | |
| 2 | CUX1 | | V$CDPCR3HD_01 | |
| 3 | CUX1 | | V$CLOX_01 | |
| 4 | FOXA2 | | V$HNF3B_01 | |
| 5 | GATA3 | | V$GATA3_03 | |
| 6 | POU2F1 | | V$OCT1_02 | |
| 7 | POU2F1 | | V$OCT1_04 | |
| 8 | POU3F2 | | V$BRN2_01 | |
| 9 | HSF1 | | V$HSF1_01 | |
| 10 | HSF2 | | V$HSF2_01 | |
| 11 | EP300/FOXC1/FOXD1/FOXD3/FOXF1/FOXF2/FOXH1/FOXI1/FOXJ1/FOXJ2/FOXL1/FOXO1/FOXO3/FOXO4/MEIS2/RRN3/SMAD2/TGIF1 | | V$XFD1_01 | |
| 12 | EP300/FOXC1/FOXD1/FOXD3/FOXF1/FOXF2/FOXH1/FOXI1/FOXJ1/FOXJ2/FOXL1/FOXO1/FOXO3/FOXO4/MEIS2/RRN3/SMAD2/TGIF1 | | V$XFD3_01 | |
| 13 | FOXC1 | rs1535679_C | V$FREAC3_01 | rs1535679_A |
| 14 | PAX4 | | V$PAX4_04 | |
| 15 | MEF2A/MEF2D | | V$AMEF2_Q6 | |
| 16 | ALX1 | | V$CART1_01 | |
| 17 | PDX1 | | V$IPF1_Q4 | |
| 18 | NKX6-2 | | V$NKX62_Q2 | |
| 19 | PDX1 | | V$IPF1_Q6 | |
| 20 | DBP | | V$DBP_Q6 | |
| 21 | FOXA1 | | V$HNF3ALPHA_Q6 | |
| 22 | POU1F1 | | V$POU1F1_Q6 | |
| 23 | FOXA1/FOXA2/FOXA3 | | V$HNF3_Q6 | |
| 24 | PLXNA2 | | V$OCT_Q6 | |
| 25 | PDX1 | | V$IPF1_Q4_01 | |
| 26 | / | | V$CMAF_01 | |
| 27 | PDX1 | | V$IPF1_01 | |
| 28 | PDX1 | | V$IPF1_02 | |
| 1 | MYCN | | V$NMYC_01 | |
| 2 | FOXO1 | | V$FOXO1_01 | |
| 3 | SOX10/11/12/15/18/2/4/5/6/9/SRY | | V$SOX_Q6 | |
| 4 | TAL1/TCF3/TCF4 | rs41362550_T | V$TAL1BETAE47_01 | rs41362550_C |
| 5 | TAL1/TCF3/TCF4 | | V$TAL1ALPHAE47_01 | |
| 6 | TAL1/TCF3/TCF4 | | V$TAL1BETAITF2_01 | |
| 7 | TFCP2 | | V$CP2_01 | |

*(Continued)*

*(Continued)*

| # of TFBS /SNP | TF name (HGNC) | SNP (allele 1) | TFBS name (TRANSFAC) | SNP (allele 2) |
|---|---|---|---|---|
| 8 | TFCP2 | | V$CP2_01 | |
| 9 | MAX/MYC/MYCN | | V$MYCMAX_01 | |
| 10 | MAX | | V$MAX_01 | |
| 11 | USF1/USF2 | | V$USF_01 | |
| 12 | USF1/USF2 | | V$USF_02 | |
| 13 | MAX/MYC/MYCN | | V$MYCMAX_02 | |
| 14 | USF1/USF2 | | V$USF_C | |
| 15 | REPIN1 | | V$AP4_Q5 | |
| 16 | USF2 | | V$USF2_Q6 | |
| 17 | REPIN1 | | V$AP4_Q6 | |
| 18 | USF1/USF2 | | V$USF_Q6 | |
| 19 | ARNT | | V$ARNT_01 | |
| 20 | HIF1A | | V$HIF1_Q5 | |
| 21 | TCF12 | | V$HTF_01 | |
| 22 | ARNT | | V$ARNT_02 | |
| 23 | MAX/MYC/MYCN | | V$MYCMAX_03 | |
| 24 | HNF4A | | V$HNF4ALPHA_Q6 | |
| 25 | SP3 | | V$SP3_Q3 | |
| 26 | USF2 | | V$USF2_Q6 | |
| 27 | TAL1 | | V$TAL1_Q6 | |
| 28 | USF1/USF2 | | V$USF_Q6_01 | |
| 29 | HIF1A | | V$HIF1_Q3 | |
| 30 | MYC | | V$MYC_Q2 | |
| 31 | BHLHE40/BHLHE41 | | V$DEC_Q1 | |
| 32 | MYC | | V$CMYC_01 | |
| 33 | MYC | | V$CMYC_02 | |
| 34 | EPAS1 | | V$HIF2A_01 | |
| 1 | MEF2A/MEF2C/MEF2D | | V$MEF2_02 | |
| 2 | ATF1/ATF2/ATF3/ATF4/ATF6/ATF7/CREB1/CREM/JUN | | V$CREBP1_01 | |
| 3 | MEF2A/MEF2C/MEF2D | | V$MEF2_03 | |
| 4 | ATF1/ATF2/ATF3/ATF4/ATF6/ATF7/CREB1/CREM/JUN | | V$CREBP1_01 | |
| 5 | / | | V$AP3_Q6 | |
| 6 | CDX1/CDX2 | rs28708675_T | V$CDXA_01 | rs28708675_A |
| 7 | CDX1/CDX2 | | V$CDXA_02 | |
| 8 | HNF4A | | V$HNF4ALPHA_Q6 | |
| 9 | PLXNA2 | | V$OCT_C | |
| 10 | MEF2A/MEF2C/MEF2D | | V$MEF2_04 | |
| 11 | POU2F1 | | V$OCT1_07 | |
| 12 | TBP | | V$TATA_01 | |

*(Continued)*

*(Continued)*

| # of TFBS /SNP | TF name (HGNC) | SNP (allele 1) | TFBS name (TRANSFAC) | SNP (allele 2) |
|---|---|---|---|---|
| 13 | EP300/FOXC1/FOXD1/FOXD3/FOXF1/FOXF2/FOXH1/FOXI1/FOXJ1/FOXJ2/FOXL1/FOXO1/FOXO3/FOXO4/MEIS2/RRN3/SMAD2/TGIF1 | | V$XFD2_01 | |
| 14 | EP300/FOXC1/FOXD1/FOXD3/FOXF1/FOXF2/FOXH1/FOXI1/FOXJ1/FOXJ2/FOXL1/FOXO1/FOXO3/FOXO4/MEIS2/RRN3/SMAD2/TGIF1 | | V$XFD3_01 | |
| 15 | FOXC1/FOXD1/FOXD3/FOXF1/FOXF2/FOXH1/FOXI1/FOXJ1/FOXJ2/FOXL1/FOXO1/FOXO3/FOXO4 | | V$HFH3_01 | |
| 16 | FOXC1 | | V$FREAC3_01 | |
| 17 | ZNF333 | | V$ZNF333_01 | |
| 18 | / | | V$MTATA_B | |
| 19 | POU3F2 | | V$POU3F2_02 | |
| 20 | TEF | | V$TEF_Q6 | |
| 21 | FOXA1 | | V$HNF3ALPHA_Q6 | |
| 22 | FOXA1/FOXA2/FOXA3 | | V$HNF3_Q6 | |
| 23 | PLXNA2 | | V$OCT_Q6 | |
| 24 | TBP | | V$TBP_Q6 | |
| 1 | POU2F1 | rs7148720_T | V$OCT1_06 | rs7148720_C |
| 2 | EP300/FOXC1/FOXD1/FOXD3/FOXF1/FOXF2/FOXH1/FOXI1/FOXJ1/FOXJ2/FOXL1/FOXO1/FOXO3/FOXO4/MEIS2/RRN3/SMAD2/TGIF1 | | V$XFD1_01 | |
| 3 | MITF/TFE3/TFEB/TFEC | | V$TFE_Q6 | |
| 4 | EP300/FOXC1/FOXD1/FOXD3/FOXF1/FOXF2/FOXH1/FOXI1/FOXJ1/FOXJ2/FOXL1/FOXO1/FOXO3/FOXO4/MEIS2/RRN3/SMAD2/TGIF1 | | V$XFD2_01 | |
| 5 | FOXC1/FOXD1/FOXD3/FOXF1/FOXF2/FOXH1/FOXI1/FOXJ1/FOXJ2/FOXL1/FOXO1/FOXO3/FOXO4 | | V$HFH3_01 | |
| 6 | FOXC1 | | V$FREAC3_01 | |
| 7 | / | | V$MTATA_B | |
| 8 | FOXO3 | | V$FOXO3_01 | |
| 9 | CEBPG | | V$CEBPGAMMA_Q6 | |
| 10 | CDX2 | | V$CDX2_Q5 | |
| 11 | FOXA1/FOXA2/FOXA3 | | V$HNF3_Q6 | |
| 12 | FOXC1/FOXD1/FOXD3/FOXF1/FOXF2/FOXH1/FOXI1/FOXJ1/FOXJ2/FOXL1/FOXO1/FOXO3/FOXO4 | | V$FOX_Q2 | |
| 13 | HMGA1 | | V$HMGIY_Q3 | |
| 14 | ZBTB16 | | V$PLZF_02 | |
| 15 | FOXO3 | | V$FOXO3A_Q1 | |
| 16 | PDX1 | | V$IPF1_01 | |
| 1 | TCF3 | rs2783778_A | V$E47_01 | rs2783778_G |
| 2 | MYB | | V$CMYB_01 | |
| 3 | TAL1/TCF3/TCF4 | | V$TAL1BETAE47_01 | |
| 4 | TAL1/TCF3/TCF4 | | V$TAL1ALPHAE47_01 | |

*(Continued)*

*(Continued)*

| # of TFBS /SNP | TF name (HGNC) | SNP (allele 1) | TFBS name (TRANSFAC) | SNP (allele 2) |
|---|---|---|---|---|
| 5 | TAL1/TCF3/TCF4 | | V$TAL1BETAITF2_01 | |
| 6 | MYB | | V$MYB_Q6 | |
| 7 | MYOD1 | | V$MYOD_Q6 | |
| 8 | NR3C1 | | V$GR_Q6 | |
| 9 | MYB | | V$MYB_Q3 | |
| 10 | TCF3 | | V$E2A_Q2 | |
| 11 | MYB | | V$MYB_Q5_01 | |
| 12 | TAL1 | | V$TAL1_Q6 | |
| 13 | AIRE | | V$AIRE_01 | |

* TFBS recognized when allele 1, but lost when allele 2 is present. Grey shaded rows present lost TFBS due to SNP.
/ - presents mammalian TFs, not found in the HGNC database.

# ANNEX J:

The list of cancer deregulated miRNAs confirmed in at least five research studies (Ferdin et al., 2010)

| MicroRNA name | Genome location | Exp (↑/↓) | Cancer association | Clinical correlations | References |
|---|---|---|---|---|---|
| **hsa-let-7b** | 22q13.31 | ↓ | leukemia (ALL, CLL), ovary, prostate, liver, brain ca | poor prognosis | [88], [93], [95], [97], [102], [110], [116] |
| | | ↑ | leukemia (ALL) | | |
| **hsa-let-7c** | 21q21.1 | ↓ | prostate, ovary, liver ca | poor prognosis | [88], [97], [102], [106], [107] |
| **hsa-let-7d** | 9q22.32 | ↓ | ovary, prostate, breast, colon, liver, brain ca | (↓) bad prognosis | [71], [86], [95], [97], [102], [107] |
| | | ↑ | pancreas ca | | |
| **hsa-let-7e** | 19q13.33 | ↓ | liver, brain ca | | [72], [87], [95], [97], [99], [120] |
| | | ↑ | leukemia (AML), retinoblastoma, colon ca | | |
| **hsa-let-7i** | 12q14.1 | ↓ | breast, brain ca | (↓) bad prognosis | [71], [85-87], [95] |
| | | ↑ | head and neck, pancreas, breast, prostate, lung ca | | |
| **hsa-miR-9** | 15q26.1 (hsa-miR-9-3) | ↓ | lung, ovary, brain ca | | [73], [87], [94-95], [107] |
| | | ↑ | liver ca | | |
| **hsa-miR-15b** | 3q26.1 | ↓ | leukemia (AML), thyroid ca | | [70], [86], [88], [99], [103], [109] |
| | | ↑ | leukemia (CLL), pancreas, ovary, colorectal ca | | |
| **hsa-miR-16** | 13q14.3 (hsa-miR-16-1) 3q26.1 (hsa-miR-16-2) | ↓ | leukemia (CLL), prostate ca | good prognosis (mutation in familiar CLL) | [28], [86], [88], [93], [100], [103], [117] |
| | | ↑ | ovary, prostate, lung, pancreas ca | | |
| **hsa-miR-17-5p** | 13q31.3 | ↑ | prostate, tongue, lung, bladder, ovary, breast, colon, pancreas, brain ca | | [75], [88], [90], [95], [104] |
| **hsa-miR-19a** | 13q31.3 | ↑ | leukemia (CLL, ALL), B-cell lymphomas, colorectal, lung, gastric, brain ca | | [70], [76-77], [95], [110] |

*(Continued)*

*(Continued)*

| MicroRNA name | Genome location | Exp (↑/↓) | Cancer association | Clinical correlations | References |
|---|---|---|---|---|---|
| **hsa-miR-20a** | 13q31.3 | ↓ | leukemia (CLL), breast, colon ca | | [42], [70], [75], [77], [87-88], [95], [101] |
| | | ↑ | colon, pancreas, ovary, prostate, lung, gastric, brain ca | | |
| **hsa-miR-21** | 17q23.1 | ↓ | breast, brain ca | poor survival, lymph node metastasis, poor prognosis | [27], [42], [70-96] |
| | | ↑ | leukemia (CLL), liver, breast, colon, lung, pancreas, prostate, stomach, colorectal, brain, ovary, tongue, thyroid, uterine, head and neck ca | | |
| **hsa-miR-23a** | 19p13.12 | ↓ | leukemia (CLL), brain, prostate ca | | [70], [72], [83], [102], [104] |
| | | ↑ | brain, breast, ovary, prostate, lung, bladder, pancreas ca | | |
| **hsa-miR-23b** | 9q22.32 | ↓ | brain, prostate ca | | [72], [84], [101-102], [104] |
| | | ↑ | brain, ovary, colon, pancreas, bladder, uterine ca | | |
| **hsa-miR-24-1** | 9q22.32 | ↓ | leukemia (CLL), brain ca | cancer hystotype | [70], [72], [75], [86], [96] |
| | | ↑ | leukemia (CLL), colon, pancreas, stomach, thyroid, brain ca | | |
| **hsa-miR-24-2** | 19p13.12 | ↓ | leukemia (CLL), brain ca | | [70], [72], [75], [86], [94], [96] |
| | | ↑ | brain, colon, pancreas, stomach, lung, thyroid ca | | |
| **hsa-miR-25** | 7q22.1 | ↓ | leukemia (CLL), colon ca | | [70], [73], [75], [87], [89], [95], [110], [113] |
| | | ↑ | leukemia (CLL, ALL), prostate, stomach, pancreas, liver, colorectal, brain ca | | |
| **hsa-miR-26a** | 3p22.2 (hsa-miR-026a-1) 12q14.1 (hsa-miR-026a-2) | ↓ | leukemia (AML), thyroid, prostate, brain ca | | [72], [95], [99], [102], [111] |
| | | ↑ | brain ca | | |

*(Continued)*

*(Continued)*

| MicroRNA name | Genome location | Exp (↑/↓) | Cancer association | Clinical correlations | References |
|---|---|---|---|---|---|
| **hsa-miR-26b** | 2q35 | ↓ | brain, tongue ca | (↑) good survival, cancer hystotype | [72], [74], [83], [90], [95], [104], [110] |
| | | ↑ | leukemia (CLL, ALL), brain, liver, colon, pancreas, bladder ca | | |
| **hsa-miR-29b** | 7q32.3 (hsa-miR-029b-1) 1q32.2 (hsa-miR-029b-2) | ↓ | leukemia (CLL, ALL), brain ca | | [79], [87], [95-96], [108], [110] |
| | | ↑ | leukemia (CLL), breast, bladder, thyroid ca | | |
| **hsa-miR-29c** | 1q32.2 | ↓ | ovary, brain ca | (↑) good survival | [70], [79], [85], [88], [95-96], [107], [110] |
| | | ↑ | leukemia (CLL, ALL), breast, pancreas, thyroid, liver, head and neck ca | | |
| **hsa-miR-30d** | 8q24.22 | ↓ | thyroid, brain ca | | [72], [83], [87-88], [99], [111] |
| | | ↑ | leukemia (AML), serous, pancreas, brain ca | | |
| **hsa-miR-31** | 9p21.3 | ↓ | prostate, gastric, breast, brain ca | (↑) good survival | [74], [76-79], [90], [95], [100] |
| | | ↑ | colorectal, tongue, liver ca | | |
| **hsa-miR-92** | 13q31.3 (hsa-miR-092a-1) Xq26.2 (hsa-miR-092a-2) 1q22 (hsa-miR-092b) | ↓ | leukemia (CLL), ovary, prostate ca | | [80], [88], [102], [110], [113] |
| | | ↑ | leukemia (CLL, ALL), colorectal, ovary ca | | |
| **hsa-miR-96** | 7q32.2 | ↑ | leukemia (CLL, ALL), liver, prostate, colorectal ca, | | [70], [73], [76], [91], [100], [110] |
| **hsa-miR-99a** | 21q21.1 | ↓ | leukemia (CLL), tongue, prostate, breast, ovary ca | | [74], [87-88], [90], [93], [102], [107], [110] |
| | | ↑ | leukemia (ALL), liver ca | | |
| **hsa-miR-100** | 11q24.1 | ↓ | leukemia (CLL), ovary, prostate, tongue ca | (↑) good survival | [74], [86], [88], [90], [93], [101], [107], [110] |
| | | ↑ | leukemia (ALL), liver, ovary, stomach, pancreas ca | | |
| **hsa-miR-103** | 20p13 (hsa-miR-103-2) 5q35.1 (hsa-miR-103-2) | ↓ | leukemia (AML), prostate, brain ca | | [72], [83], [88], [95], [99], [102] |
| | | ↑ | pancreas, ovary, brain ca | | |

*(Continued)*

*(Continued)*

| MicroRNA name | Genome location | Exp (↑/↓) | Cancer association | Clinical correlations | References |
|---|---|---|---|---|---|
| **hsa-miR-106a** | Xq26.2 | ↓ | colon ca | (↓) bad prognosis | [42], [75], [77], [87-88], [94-95], [110] |
| | | ↑ | leukemia (CLL, ALL), colon, gastric, lung, ovary, pancreas, prostate, brain ca | | |
| **hsa-miR-106b** | 7q22.1 | ↑ | prostate, gastric, colon, ovary, brain ca | | [77], [87-88], [95], [121] |
| **hsa-miR-107** | 10q23.31 | ↓ | leukemia (CLL), tongue, liver, brain ca | | [70], [72], [75], [83], [86], [88], [90], [91], [95] |
| | | ↑ | colon, pancreas, stomach, ovary, brain, breast ca | | |
| **hsa-miR-10b** | 2q31.1 | ↓ | breast, ovary ca | (↓) metastatic potential (↑) cancer progression | [70-71], [75], [87], [89], [91], [118] |
| | | ↑ | leukemia (CLL), brain, breast, bladder, liver ca | | |
| **hsa-miR-125a** | 19q13.33 | ↓ | breast, ovary, prostate, liver, lung, brain ca | | [71], [75], [93-95], [102], [103], [108], [112] |
| | | ↑ | thyroid ca | | |
| **hsa-miR-125b** | 11q24.1 (hsa-miR-125b-1) | ↓ | leukemia (CLL), prostate, breast, thyroid, ovary, tongue, brain ca | cancer progression | [71], [86-88], [90], [95], [100], [102], [106], [110], [111] |
| | | ↑ | leukemia (ALL), pancreas, colon, liver ca | (↑) good survival | |
| **hsa-miR-125b-1** | 11q24.1 | ↓ | breast, ovary, prostate, tongue, liver, brain ca | | [71-72], [86], [89], [103], [107] |
| | | ↑ | pancreas, thyroid, stomach, brain ca | | |
| **hsa-miR-126** | 9q34.3 | ↓ | leukemia (CLL), breast, ovary, liver, lung ca | (↓) poor metastasis-free survival | [87-88], [94], [105], [107], [110], [119] |
| | | ↑ | leukemia (ALL), bladder ca | | |
| **hsa-miR-128a** | 2q21.3 (hsa-miR-128-1) | ↓ | breast, prostate, colon, pituitary, brain ca | brain enriched | [72], [89-90], [95], [116] |
| | | ↑ | leukemia (ALL), tongue ca | | |

*(Continued)*

*(Continued)*

| MicroRNA name | Genome location | Exp (↑/↓) | Cancer association | Clinical correlations | References |
|---|---|---|---|---|---|
| **hsa-miR-128b** | 3p22.3 (hsa-miR-128-2) | ↓ | leukemia (CLL), breast, brain ca | | [70-71], [75], [95], [110], [116] |
| | | ↑ | leukemia (CLL, ALL), colon, lung, pancreas, breast ca | | |
| **hsa-miR-132** | 17p13.3 | ↓ | leukemia (ALL), liver, brain ca | | [70], [72], [87], [90], [95], [97], [110] |
| | | ↑ | leukemia (CLL), colon, tongue ca | | |
| **hsa-miR-137** | 1p21.3 | ↓ | ovary, brain ca | | [70], [72-73], [90], [95], [107], [114-115] |
| | | ↑ | leukemia (CLL), colon, brain, tongue, liver ca | | |
| **hsa-miR-139** | 11q13.4 | ↓ | leukemia (ALL), tongue, liver, pancreas ca | (↑) good survival | [73-74], [86], [90], [110] |
| | | ↑ | liver ca | | |
| **hsa-miR-141** | 12p13.31 | ↓ | brain, liver ca | (↑) poor prognosis | [70], [72], [88], [93], [97], [101], [107] |
| | | ↑ | prostate, leukemia (CLL), ovary ca | | |
| **hsa-miR-143** | 5q33.1 | ↓ | breast, prostate, colon, liver, lung, ovary, brain ca | cancer metastases | [71], [78], [87-88], [90], [94-95], [97], [101-102], [107] |
| | | ↑ | tongue ca | | |
| **hsa-miR-145** | 5q33.1 | ↓ | leukemia (CLL), breast, colorectal, lung, prostate, bladder, ovary, liver, brain ca | metastasis bad prognosis | [70-71], [73], [75-76], [78], [87-88], [93-95], [97], [100-102], [106-107] |
| **hsa-miR-146** | 5q33.3 (hsa-miR-146a) 10q24.32 (hsa-miR-146b) | ↓ | leukemia (CLL), liver ca | | [70], [75], [85], [94], [96-97], [110] |
| | | ↑ | leukemia (CLL, ALL), breast, pancreas, prostate, head and neck, thyroid, lung ca | | |
| **hsa-miR-149** | 2q37.3 | ↓ | breast, tongue, prostate, brain ca | | [71-72], [90], [95], [100] |
| | | ↑ | brain ca | | |
| **hsa-miR-150** | 19q13.33 | ↓ | leukemia (ALL), brain ca | (↑) good survival | [74], [80], [87], [94], [97], [110] |
| | | ↑ | leukemia (CLL), lung, bladder, breast, liver ca | | |

*(Continued)*

*(Continued)*

| MicroRNA name | Genome location | Exp (↑/↓) | Cancer association | Clinical correlations | References |
|---|---|---|---|---|---|
| **hsa-miR-155** | 21q21.3 | ↓ | pancreas, breast, liver, brain ca | (↑) poor survival, bad prognosis | [70-71], [73], [75], [79], [80], [82-83], [85-87], [90], [94-99] |
| | | ↑ | leukemia (CLL, AML), liver, breast, pancreas, lung, head and neck, thyroid, tongue ca | | |
| **hsa-miR-181a** | 9q33.3 (hsa-miR-181a-2) | ↓ | brain ca | brain enriched | [82], [86], [89], [96], [99], [103] |
| | | ↑ | pancreas, thyroid ca | | |
| **hsa-miR-181b** | 1q31.3 (hsa-miR-181b-1) | ↓ | leukemia (CLL), prostate, brain ca | brain enriched | [42], [79], [82], [88-89], [95], [99-100], [103], [108-110] |
| | | ↑ | leukemia (CLL, AML, ALL), colorectal, breast, pancreas, thyroid, ovary ca | | |
| **hsa-miR-181c** | 19p13.12 | ↓ | leukemia (CLL), liver, lung, brain ca | | [86], [89-90], [95-97], [99], [103], [110] |
| | | ↑ | leukemia (ALL, AML), tongue, thyroid, pancreas ca | | |
| **hsa-miR-182** | 7q32.2 | ↓ | leukemia (AML) | | [73], [88], [99-100], [105], [110] |
| | | ↑ | leukemia (CLL, ALL), prostate, liver, ovary ca | | |
| **hsa-miR-191** | 3p21.31 | ↓ | leukemia (CLL), breast, brain ca | | [70-71], [75], [94-95], [109-110] |
| | | ↑ | leukemia (ALL), colon, lung, pancreas, prostate, stomach, breast ca | | |
| **hsa-miR-195** | 17p13.1 | ↓ | gastric, prostate, tongue, ovary, liver ca | | [77], [88], [90], [97], [99], [102], [110], [112] |
| | | ↑ | leukemia (CLL, ALL, AML) | | |
| **hsa-miR-197** | 1p13.3 | ↓ | leukemia (CLL) | | [70], [72], [90], [94], [122] |
| | | ↑ | thyroid, tongue, lung, brain ca | | |
| **hsa-miR-199a** | 19p13.2 (hsa-miR-199a-1) 1q24.3 (hsa-miR-199a-2) | ↓ | leukemia (AML), ovary, liver, prostate ca | (↓) poor prognosis | [74], [93], [99], [102], [107], [112] |
| | | ↑ | liver ca | | |
| **hsa-miR-199b** | 9q34.11 | ↓ | leukemia (CLL), ovary, liver, lung ca | | [74], [95], [97], [107], [110] |
| | | ↑ | leukemia (ALL), brain, liver ca | | |

*(Continued)*

*(Continued)*

| MicroRNA name | Genome location | Exp (↑/↓) | Cancer association | Clinical correlations | References |
|---|---|---|---|---|---|
| **hsa-miR-200a** | 1p36.33 | ↓ | liver ca | | [70], [72], [88], [103], [107], [112] |
| | | ↑ | leukemia (CLL), thyroid, ovary, brain ca | | |
| **hsa-miR-200c** | 12p13.31 | ↓ | liver ca | (↑) good survival, poor survival (mutation in p53) | [74], [88], [91], [107], [109] |
| | | ↑ | ovary, colorectal ca | | |
| **hsa-miR-203** | 14q32.33 | ↓ | breast, brain, liver ca | | [42], [71-72], [88], [91], [94-95], [104] |
| | | ↑ | colon, breast, lung, bladder, ovary, brain ca | | |
| **hsa-miR-205** | 1q32.2 | ↓ | breast, prostate ca | (↓) cancer progression | [87-88], [94], [100], [104] |
| | | ↑ | lung, bladder, ovary ca | | |
| **hsa-miR-210** | 11p15.5 | ↓ | leukemia (CLL), breast ca | | [70-71], [75], [94], [102], [110] |
| | | ↑ | leukemia (CLL, ALL), breast, lung, prostate, liver ca | | |
| **hsa-miR-212** | 17p13.3 | ↓ | leukemia (CLL), stomach, colon, pancreas, brain ca | | [70] [72], [86], [94-95], |
| | | ↑ | pancreas, lung ca | | |
| **hsa-miR-213** | 1q31.3-q32.1 | ↓ | leukemia (CLL), breast, brain ca | | [70-72], [75], [90], [96], [103] |
| | | ↑ | breast, lung, thyroid, tongue ca | | |
| **hsa-miR-214** | 1q24.3 | ↓ | ovary, liver, liver ca | | [73-75], [88], [94], [97] |
| | | ↑ | pancreas, prostate, stomach, lung, liver ca | | |
| **hsa-miR-221** | Xp11.3 | ↓ | leukemia (CLL), prostate, brain ca | (↑) good survival | [70], [73-75], [82], [86], [88-[89], [95-97], [99], [33], [101-105] |
| | | ↑ | leukemia (AML), brain, pancreas, liver, thyroid, ovary, bladder, colon, stomach ca | | |
| **hsa-miR-222** | Xp11.3 | ↓ | leukemia (CLL), prostate, ovary ca | | [73], [80], [82], [88], [91], [96], [99-103], [105], [107] |
| | | ↑ | leukemia (AML), pancreas, thyroid, ovary, stomach, colon, liver, bladder ca | | |

*(Continued)*

*(Continued)*

| MicroRNA name | Genome location | Exp (↑/↓) | Cancer association | Clinical correlations | References |
|---|---|---|---|---|---|
| **hsa-miR-223** | Xq12 | ↓ | leukemia (CLL, AML), liver, brain ca | | [70], [72], [74-75], [97], [104-105], [116] |
| | | ↑ | colon, pancreas, prostate, stomach, liver ca | | |
| **hsa-miR-224** | Xq28 | ↓ | ovary, lung, brain ca | | [72-73], [88], [91], [94], [103], [107], [112] |
| | | ↑ | liver, pancreas, thyroid, ovary ca | | |
| **hsa-miR-320** | 8p21.3 (hsa-miR-320a) 1p13.1 (hsa-miR-320b-1) 1q42.11 (hsa-miR-320b-2) 18q11.2 (hsa-miR-320c-1) 18q11.2 (hsa-miR-320c-2) 13q14.11 (hsa-miR-320d-1) Xq27.1 (hsa-miR-320d-2) | ↓ | leukemia (CLL), brain, breast, bladder ca | | [79], [87], [95], [102], [110], [120] |

ALL-acute lymphoblastic leukemia; AML-acute myeloid leukemia; CLL-chronic lymphocytic leukemia; (↑) high expression; (↓) low expression; ca-cancer.

Reference numbers correspond to publications listed in ANNEX L.

# ANNEX K:

The list of cancer deregulated miRNAs confirmed in at least two research studies (Ferdin et al., 2010)

| microRNA NAME | EXP (↑/↓) | CANCER ASSOCIATION | REFERENCES |
|---|---|---|---|
| **hsa-let-7** | ↓ | breast ca | [71], [75], [84] |
| | ↑ | uterine ca | |
| **hsa-let-7a** | ↓ | liver, prostate, brain ca | [95], [97], [102] |
| **hsa-let-7a-1** | ↓ | ovary, prostate, liver, brain ca | [72], [97] |
| **hsa-let-7a-2** | ↓ | leukemia (CLL), breast, liver, prostate, lung ca | [70-71], [97] |
| **hsa-let-7a-3** | ↓ | prostate, liver, breast ca | [71], [97] |
| **hsa-let-7b** | ↓ | leukemia (ALL, CLL), ovary, prostate, liver, brain ca | [88], [93], [95], [97], [102], [110], [116] |
| | ↑ | leukemia (ALL) | |
| **hsa-let-7c** | ↓ | prostate, ovary, liver ca | [88], [97], [102], [106], [107] |
| **hsa-let-7d** | ↓ | ovary, prostate, breast, colon, liver, brain ca | [71], [86], [95], [97], [102], [107] |
| | ↑ | pancreas ca | |
| **hsa-let-7e** | ↓ | liver, brain ca | [72], [87], [95], [97], [99], [120] |
| | ↑ | leukemia (AML), retinoblastoma, colon ca | |
| **hsa-let-7f** | ↓ | breast, liver, thyroid, brain ca | [79], [95] |
| | ↑ | breast ca | |
| **hsa-let-7f-1** | ↓ | leukemia (CLL), thyroid, brain ca | [70], [72], [86], [103] |
| | ↑ | pancreas ca | |
| **hsa-let-7f-2** | ↓ | breast, liver ca | [71], [97], [104] |
| | ↑ | renal ca | |
| **hsa-let-7g** | ↓ | prostate, liver, brain ca | [88], [95], [97], [102] |
| | ↑ | ovary ca | |
| **hsa-let-7i** | ↓ | breast, brain ca | [71], [85-87], [95] |
| | ↑ | head and neck, pancreas, breast, prostate, lung ca | |
| **hsa-miR-9** | ↓ | lung, ovary, brain ca | [73], [87], [94-95], [107] |
| | ↑ | liver ca | |
| **hsa-miR-9*** | ↓ | leukemia (AML), ovary ca | [73], [99], [107] |
| | ↑ | liver ca | |
| **hsa-miR-9-1** | ↓ | breast ca | [71], [123] |
| **hsa-miR-9-2** | ↑ | leukemia (CLL), brain ca | [70], [89] |
| **hsa-miR-9-3** | ↓ | ovary, lung, thyroid, brain ca | [72], [96] |
| **hsa-miR-15a** | ↓ | leukemia (CLL), prostate ca | [28], [88], [110], [117] |
| | ↑ | leukemia (CLL, ALL), ovary ca | |
| **hsa-miR-15b** | ↓ | leukemia (AML), thyroid ca | [70], [86], [88], [99], [103], [109] |
| | ↑ | leukemia (CLL), pancreas, ovary, colorectal ca | |
| **hsa-miR-16** | ↓ | prostate ca, leukemia (CLL) | [28], [86], [88], [93], [100], [103], [117] |
| | ↑ | ovary, prostate, lung, pancreas ca | |

*(Continued)*

*(Continued)*

| microRNA NAME | EXP (↑/↓) | CANCER ASSOCIATION | REFERENCES |
|---|---|---|---|
| **hsa-miR-17** | ↑ | gastric ca | [70], [77] |
| | ↓ | leukemia (CLL) | |
| **hsa-miR-17-3p** | ↑ | leukemia (CLL, ALL), lung ca | [94], [110] |
| **hsa-miR-17-5p** | ↑ | prostate, tongue, lung, bladder, ovary, breast, colon, pancreas, brain ca | [75], [88], [90], [95], [104] |
| **hsa-miR-18** | ↑ | liver, head and neck ca | [74], [85], [112] |
| **hsa-miR-18a** | ↑ | gastric, ovary, brain ca | [77], [93], [95] |
| **hsa-miR-19a** | ↑ | leukemia (CLL, ALL), colorectal, lung, gastric, brain ca | [70], [76-77], [95], [110] |
| **hsa-miR-20** | ↑ | leukemia (CLL, ALL), colorectal, lung, liver ca | [76], [110] |
| **hsa-miR-20a** | ↓ | leukemia (CLL), breast, colon ca | [42], [70], [75], [77], [87-88], [95], [101] |
| | ↑ | colon, pancreas, ovary, prostate, lung, gastric, brain ca | |
| **hsa-miR-20b** | ↓ | colon ca | [77], [87-88] |
| | ↑ | gastric, ovary ca | |
| **hsa-miR-21** | ↓ | breast, brain ca | [27], [42], [70-96] |
| | ↑ | leukemia (CLL), liver, breast, colon, lung, pancreas, prostate, stomach, colorectal, tongue, thyroid, uterine, head and neck, ovary, brain ca | |
| **hsa-miR-22** | ↓ | brain ca | [70], [95] |
| | ↑ | leukemia (CLL) | |
| **hsa-miR-23a** | ↓ | leukemia (CLL), brain, prostate ca | [70], [72], [83], [102], [104] |
| | ↑ | brain, breast, ovary, prostate, lung, bladder, pancreas ca | |
| **hsa-miR-23b** | ↓ | brain, prostate ca | [72], [84], [101-102], [104] |
| | ↑ | brain, ovary, colon, pancreas, bladder, uterine ca | |
| **hsa-miR-24-1** | ↓ | leukemia (CLL), brain ca | [70], [72], [75], [86], [96] |
| | ↑ | leukemia (CLL), colon, pancreas, stomach, thyroid, brain ca | |
| **hsa-miR-24-2** | ↓ | leukemia (CLL), brain ca | [70], [72], [75], [86], [94], [96] |
| | ↑ | brain, colon, pancreas, stomach, lung, thyroid ca | |
| **hsa-miR-25** | ↓ | leukemia (CLL), colon ca | [70], [73], [75], [87], [89], [95], [110], [113] |
| | ↑ | leukemia (CLL, ALL), prostate, stomach, pancreas, liver, colorectal, brain ca | |
| **hsa-miR-26a** | ↓ | leukemia (AML), thyroid, prostate, brain ca | [72], [95], [99], [102], [111] |
| | ↑ | brain ca | |
| **hsa-miR-26a-1** | ↓ | thyroid, ovary, prostate, lung ca | [83], [94], [96] |
| | ↑ | colon, pancreas ca | |
| **hsa-miR-26b** | ↓ | tongue, brain ca | [72], [74], [83], [90], [95], [104], [110] |
| | ↑ | leukemia (CLL, ALL), brain, liver, colon, pancreas, bladder ca | |
| **hsa-miR-27a** | ↑ | leukemia (CLL, ALL, AML) | [99], [110] |
| **hsa-miR-27b** | ↓ | leukemia (CLL), lung ca | [70], [94] |
| **hsa-miR-28** | ↑ | leukemia (CLL, ALL), renal ca | [104], [110] |

*(Continued)*

*(Continued)*

| microRNA NAME | EXP (↑/↓) | CANCER ASSOCIATION | REFERENCES |
|---|---|---|---|
| **hsa-miR-29a** | ↓ | ovary, brain ca | [87-88], [95], [110] |
| | ↑ | leukemia (CLL, ALL), bladder, ovary, breast, prostate, pancreas, thyroid ca | |
| **hsa-miR-29a-2** | ↑ | leukemia (CLL), thyroid ca | [70], [96] |
| **hsa-miR-29b** | ↓ | leukemia (CLL, ALL), brain ca | [79], [87], [95-96], [108], [110] |
| | ↑ | leukemia (CLL), breast, bladder, thyroid ca | |
| **hsa-miR-29b-2** | ↓ | lung ca | [75], [94] |
| | ↑ | breast, colon, pancreas, prostate, thyroid ca | |
| **hsa-miR-29c** | ↓ | ovary, brain ca | [70], [79], [85], [88], [95-96], [107], [110] |
| | ↑ | leukemia (CLL, ALL), breast, pancreas, thyroid, liver, head and neck ca | |
| **hsa-miR-30a** | ↓ | leukemia (CLL),brain ca | [70], [72], [87] |
| | ↑ | brain ca | |
| **hsa-miR-30a-3p** | ↓ | leukemia (AML), breast ca | [79], [90], [99], [110] |
| | ↑ | leukemia (ALL), tongue ca | |
| **hsa-miR-30a-5p** | ↓ | thyroid, lung ca | [88], [94], [111] |
| | ↑ | ovary ca | |
| **hsa-miR-30b** | ↓ | brain ca | [72], [87-88], [95] |
| | ↑ | breast, ovary, brain ca | |
| **hsa-miR-30c** | ↓ | brain ca | [72], [75], [88], [95] |
| | ↑ | brain, colon, pancreas, prostate, ovary ca | |
| **hsa-miR-30d** | ↓ | thyroid, brain ca | [72], [83], [87-88], [99], [111] |
| | ↑ | leukemia (AML), serous, pancreas, brain ca | |
| **hsa-miR-31** | ↓ | prostate, gastric, breast, brain ca | [74], [76-79], [90], [95], [100] |
| | ↑ | colorectal, tongue, liver ca | |
| **hsa-miR-32** | ↓ | lung ca | [75], [94], [121] |
| | ↑ | prostate, colon, pancreas ca | |
| **hsa-miR-33** | ↓ | lung ca | [70], [74], [94] |
| | ↑ | leukemia (CLL), liver ca | |
| **hsa-miR-34** | ↓ | breast ca | [70-71], [103] |
| | ↑ | leukemia (CLL), breast, tongue, liver, thyroid ca | |
| **hsa-miR-34a** | ↓ | ovary ca | [88], [96], [110] |
| | ↑ | leukemia (CLL, ALL), ovary, thyroid ca | |
| **hsa-miR-34b** | ↓ | ovary ca | [88], [90] |
| | ↑ | ovary, tongue ca | |
| **hsa-miR-34c** | ↓ | leukemia (AML), ovary ca | [88], [90], [99] |
| | ↑ | tongue ca | |
| **hsa-miR-92** | ↓ | leukemia (CLL), ovary, prostate ca | [80], [88], [102], [110], [113] |
| | ↑ | leukemia (CLL, ALL), colorectal, ovary ca | |
| **hsa-miR-92-1** | ↑ | leukemia (CLL), pancreas ca | [70], [86] |
| **hsa-miR-92-2** | ↓ | leukemia (CLL) | [70], [75] |
| | ↑ | pancreas, prostate, stomach ca | |
| **hsa-miR-96** | ↑ | leukemia (CLL, ALL), liver, prostate, colorectal ca | [70], [73], [76], [91], [100], [110] |

*(Continued)*

*(Continued)*

| microRNA NAME | EXP (↑/↓) | CANCER ASSOCIATION | REFERENCES |
|---|---|---|---|
| **hsa-miR-98** | ↓ | brain ca | [79], [95], [110] |
| | ↑ | leukemia (CLL, ALL), breast ca | |
| **hsa-miR-99a** | ↓ | leukemia (CLL), tongue, prostate, breast, ovary ca | [74], [87-88], [90], [93], [102], [107], [110] |
| | ↑ | leukemia (ALL), liver ca | |
| **hsa-miR-99b** | ↓ | breast, brain ca | [72], [74], [87] |
| | ↑ | colon, liver, brain ca | |
| **hsa-miR-100** | ↓ | leukemia (CLL), ovary, prostate, tongue ca | [74], [86], [88], [90], [93], [101], [107], [110] |
| | ↑ | leukemia (ALL), liver, ovary, stomach, pancreas ca | |
| **hsa-miR-101** | ↓ | breast, ovary, prostate, colon, liver, lung ca | [70], [74], [107] |
| | ↑ | leukemia (CLL), liver ca | |
| **hsa-miR-101-1** | ↓ | breast, lung ca | [70-71], [94] |
| | ↑ | leukemia (CLL) | |
| **hsa-miR-103** | ↓ | leukemia (AML), prostate, brain ca | [72], [83], [88], [95], [99], [102] |
| | ↑ | pancreas, ovary, brain ca | |
| **hsa-miR-103-1** | ↓ | leukemia (CLL) | [70], [72], [83], [104] |
| | ↑ | brain, pancreas, breast, stomach, colon, bladder ca | |
| **hsa-miR-103-2** | ↑ | brain, pancreas, prostate, colon ca | [72], [83] |
| **hsa-miR-106a** | ↓ | colon ca | [42], [75], [77], [87-88], [94-95], [110] |
| | ↑ | leukemia (CLL, ALL), colon, gastric, lung, ovary, pancreas, prostate, brain ca | |
| **hsa-miR-106b** | ↑ | prostate, gastric, colon, ovary, brain ca | [77], [87-88], [95], [121] |
| **hsa-miR-107** | ↓ | leukemia (CLL), tongue, liver, brain ca | [70], [72], [75], [83], [86], [88], [90], [91], [95] |
| | ↑ | colon, pancreas, stomach, ovary, brain, breast ca | |
| **hsa-miR-10b** | ↓ | breast, ovary ca | [70-71], [75], [87], [89], [91], [118] |
| | ↑ | leukemia (CLL), brain, breast, bladder, liver ca | |
| **hsa-miR-122** | ↓ | liver, breast ca | [71], [91], [97], [105] |
| **hsa-miR-123** | ↓ | leukemia (CLL) | [70], [72], [89] |
| | ↑ | brain ca | |
| **hsa-miR-124a** | ↓ | colorectal, lung, ovary, brain ca | [76], [90], [95], [108] |
| | ↑ | tongue ca | |
| **hsa-miR-124a-2** | ↓ | liver ca | [70], [97] |
| | ↑ | leukemia (CLL) | |
| **hsa-miR-125a** | ↓ | breast, ovary, prostate, liver, lung, brain ca | [71], [75], [93-95], [102], [103], [108], [112] |
| | ↑ | thyroid ca | |
| **hsa-miR-125b** | ↓ | leukemia (CLL), prostate, breast, thyroid, ovary, tongue, brain ca | [71], [86-88], [90], [95], [100], [102], [106], [110-111] |
| | ↑ | leukemia (ALL), pancreas, colon, liver ca | |
| **hsa-miR-125b-1** | ↓ | breast, ovary, prostate, tongue, liver, brain ca | [71-72], [86], [89], [103], [107] |
| | ↑ | pancreas, thyroid, stomach, brain ca | |
| **hsa-miR-125b-2** | ↓ | breast, ovary, prostate, tongue, brain ca | [72], [89], [103] |
| | ↑ | brain, stomach, pancreas, thyroid ca | |

*(Continued)*

*(Continued)*

| microRNA NAME | EXP (↑/↓) | CANCER ASSOCIATION | REFERENCES |
|---|---|---|---|
| **hsa-miR-126** | ↓ | leukemia (CLL), breast, ovary, liver, lung ca | [87-88], [94], [105], [107], [110], [119] |
| | ↑ | leukemia (ALL), bladder ca | |
| **hsa-miR-126*** | ↓ | lung, liver, breast ca | [91], [94] |
| **hsa-miR-127** | ↓ | breast, brain ca | [72], [79], [95] |
| | ↑ | brain ca | |
| **hsa-miR-128a** | ↓ | breast, prostate, colon, pituitary, brain ca | [72], [89-90], [95], [116] |
| | ↑ | leukemia (ALL vs AML), tongue ca | |
| **hsa-miR-128b** | ↓ | leukemia (CLL), breast, brain ca | [70-71], [75], [95], [110], [116] |
| | ↑ | leukemia (CLL, ALL), colon, lung, pancreas, breast ca | |
| **hsa-miR-129** | ↓ | colorectal, brain ca | [76], [95] |
| **hsa-miR-130a** | ↓ | leukemia (CLL), breast, liver, lung ca | [72], [89], [97], [110] |
| | ↑ | leukemia (ALL), brain ca | |
| **hsa-miR-130b** | ↓ | breast ca | [74], [87], [110] |
| | ↑ | leukemia (CLL, ALL), liver ca | |
| **hsa-miR-132** | ↓ | leukemia (ALL), liver, brain ca | [70], [72], [87], [90], [95], [97], [110] |
| | ↑ | leukemia (CLL), colon, tongue ca | |
| **hsa-miR-133a** | ↓ | ovary, tongue ca | [90], [107] |
| **hsa-miR-133b** | ↓ | gastric, colorectal brain, tongue ca | [76-77], [90], [95] |
| **hsa-miR-134** | ↓ | brain ca | [70], [72], [95] |
| | ↑ | leukemia (CLL), brain ca | |
| **hsa-miR-135a** | ↓ | brain ca | [74], [95] |
| | ↑ | liver ca | |
| **hsa-miR-135b** | ↓ | leukemia (ALL, CLL) | [76], [110] |
| | ↑ | colorectal ca | |
| **hsa-miR-136** | ↓ | breast, brain, stomach, colon, liver ca | [70-71], [72] |
| | ↑ | leukemia (CLL) | |
| **hsa-miR-137** | ↓ | ovary, brain ca | [70], [72-73], [90], [95], [107], [114-115] |
| | ↑ | leukemia (CLL), colon, brain, tongue, liver ca | |
| **hsa-miR-138** | ↓ | tongue, brain ca | [87], [90], [95] |
| | ↑ | colon ca | |
| **hsa-miR-138-1** | ↓ | lung, thyroid ca | [70], [87], [96] |
| | ↑ | leukemia (CLL) | |
| **hsa-miR-138-2** | ↓ | brain, tongue, stomach, colon, pancreas, thyroid ca | [72], [96] |
| **hsa-miR-139** | ↓ | leukemia (ALL), tongue, liver, pancreas ca | [73-74], [86], [90], [110] |
| | ↑ | liver ca | |
| **hsa-miR-140** | ↓ | ovary, lung, thyroid ca | [70], [94], [107] |
| | ↑ | leukemia (CLL) | |
| **hsa-miR-141** | ↓ | brain, liver ca | [70], [72], [88], [93], [97], [101], [107] |
| | ↑ | prostate, leukemia (CLL), ovary ca | |
| **hsa-miR-142** | ↓ | leukemia (CLL), liver, thyroid ca | [70], [97] |
| **hsa-miR-142-5p** | ↓ | leukemia (CLL) | [87], [99], [110] |
| | ↑ | leukemia (AML), bladder, lung ca | |

*(Continued)*

*(Continued)*

| microRNA NAME | EXP (↑/↓) | CANCER ASSOCIATION | REFERENCES |
|---|---|---|---|
| **hsa-miR-143** | ↓ | breast, prostate, colon, liver, lung, ovary, brain ca | [71], [78], [87-88], [90], [94-95], [97], [101-102], [107] |
| | ↑ | tongue ca | |
| **hsa-miR-144** | ↓ | brain, ovary ca | [72], [107] |
| **hsa-miR-145** | ↓ | leukemia (CLL), breast, colorectal, lung, prostate, ovary, bladder, liver, brain ca | [70-71], [73], [75-76], [78], [87-88], [93-95], [97], [100-102], [106-107] |
| **hsa-miR-146** | ↓ | leukemia (CLL), liver ca | [70], [75], [85], [94], [96-97], [110] |
| | ↑ | leukemia (CLL, ALL), breast, pancreas, prostate, head and neck, thyroid, lung ca | |
| **hsa-miR-146b** | ↑ | thyroid, head and neck, ovary ca | [85], [88], [96] |
| **hsa-miR-147** | ↓ | leukemia (AML), ovary ca | [90], [99], [107] |
| | ↑ | tongue ca | |
| **hsa-miR-148** | ↓ | brain, thyroid ca | [70], [72], [103] |
| | ↑ | leukemia (CLL) | |
| **hsa-miR-148a** | ↑ | leukemia (CLL, ALL), liver, brain, breast, lung ca | [74], [87], [95], [110] |
| **hsa-miR-149** | ↓ | breast, tongue, prostate, brain ca | [71-72], [90], [95], [100] |
| | ↑ | brain ca | |
| **hsa-miR-150** | ↓ | leukemia (ALL), brain ca | [74], [80], [87], [94], [97], [110] |
| | ↑ | leukemia (CLL), lung, bladder, breast, liver ca | |
| **hsa-miR-151** | ↓ | leukemia (AML) | [73], [99], [103] |
| | ↑ | liver ca | |
| **hsa-miR-152** | ↓ | ovary, colon, pancreas ca | [70], [87-88], [110] |
| | ↑ | leukemia (CLL), lung ca | |
| **hsa-miR-153** | ↓ | brain ca | [70], [95] |
| | ↑ | leukemia (CLL) | |
| **hsa-miR-154** | ↓ | brain ca | [70], [72], [90], [95] |
| | ↑ | leukemia (CLL), brain, tongue ca | |
| **hsa-miR-155** | ↓ | pancreas, breast, liver, brain ca | [70-71], [73], [75], [79], [80], [82-83], [85-87], [90], [94-99] |
| | ↑ | leukemia (CLL, AML), liver, breast, pancreas, lung, head and neck, thyroid, tongue ca | |
| **hsa-miR-181a** | ↓ | brain ca | [82], [86], [89], [96], [99], [103] |
| | ↑ | pancreas, thyroid ca | |
| **hsa-miR-181b** | ↓ | leukemia (CLL), prostate, brain ca | [42], [79], [82], [88-89], [95], [99-100], [103], [108-110] |
| | ↑ | leukemia (CLL, AML, ALL), colorectal, breast, pancreas, thyroid, ovary ca | |
| **hsa-miR-181c** | ↓ | leukemia (CLL), liver, lung, brain ca | [86], [89-90], [95-97], [99], [103], [110] |
| | ↑ | leukemia (ALL, AML), tongue, thyroid, pancreas ca | |
| **hsa-miR-181d** | ↑ | breast, pancreas ca | [79], [82] |
| **hsa-miR-182** | ↓ | leukemia (AML) | [73], [88], [99-100], [105], [110] |
| | ↑ | leukemia (CLL, ALL), prostate, liver, ovary ca | |
| **hsa-miR-182*** | ↑ | liver, prostate ca | [73], [100] |
| **hsa-miR-183** | ↑ | prostate, colorectal, liver ca | [73], [76], [100] |

*(Continued)*

*(Continued)*

| microRNA NAME | EXP (↑/↓) | CANCER ASSOCIATION | REFERENCES |
|---|---|---|---|
| **hsa-miR-184** | ↓ | leukemia (AML), prostate ca | [70], [90], [99-100] |
| | ↑ | leukemia (CLL), tongue ca | |
| **hsa-miR-190** | ↓ | brain ca | [70], [95] |
| | ↑ | leukemia (CLL) | |
| **hsa-miR-191** | ↓ | leukemia (CLL), breast, brain ca | [70-71], [75], [94-95], [109-110] |
| | ↑ | leukemia (ALL), colon, lung, pancreas, prostate, stomach, breast ca | |
| **hsa-miR-192** | ↓ | leukemia (CLL), brain ca | [70], [83], [94-95] |
| | ↑ | pancreas, stomach, liver, lung ca | |
| **hsa-miR-194** | ↓ | brain, tongue ca | [70], [90], [95], [110] |
| | ↑ | leukemia (CLL, ALL) | |
| **hsa-miR-195** | ↓ | gastric, prostate, tongue, ovary, liver ca | [77], [88], [90], [97], [99], [102], [110], [112] |
| | ↑ | leukemia (CLL, ALL, AML) | |
| **hsa-miR-196-2** | ↓ | breast ca | [70-71], [103] |
| | ↑ | leukemia (CLL), breast, thyroid ca | |
| **hsa-miR-197** | ↓ | leukemia (CLL) | [70], [72], [90], [94], [122] |
| | ↑ | thyroid, tongue, lung, brain ca | |
| **hsa-miR-198** | ↓ | lung ca | [90], [94], [120] |
| | ↑ | retinoblastoma, tongue ca | |
| **hsa-miR-199a** | ↓ | leukemia (AML), ovary, liver, prostate ca | [74], [93], [99], [102], [107], [112] |
| | ↑ | liver ca | |
| **hsa-miR-199a*** | ↓ | prostate, liver ca | [74], [102], [112] |
| | ↑ | liver ca | |
| **hsa-miR-199a-1** | ↓ | leukemia (ALL, CLL), thyroid ca | [75], [103], [110] |
| | ↑ | lung, pancreas, prostate ca | |
| **hsa-miR-199b** | ↓ | leukemia (CLL), ovary, liver, lung ca | [74], [95], [97], [107], [110] |
| | ↑ | leukemia (ALL), brain, liver ca | |
| **hsa-miR-200a** | ↓ | liver ca | [70], [72], [88], [103], [107], [112] |
| | ↑ | leukemia (CLL), ovary, thyroid, brain ca | |
| **hsa-miR-200b** | ↓ | ovary, liver ca | [74], [88], [97], [107] |
| | ↑ | ovary, liver ca | |
| **hsa-miR-200c** | ↓ | liver ca | [74], [88], [91], [107], [109] |
| | ↑ | ovary, colorectal ca | |
| **hsa-miR-202** | ↓ | breast ca | [71], [102-103] |
| | ↑ | breast, prostate, thyroid ca | |
| **hsa-miR-203** | ↓ | breast, brain, liver ca | [42], [71-72], [88], [91], [94-95], [104] |
| | ↑ | colon, breast, lung, bladder, ovary, brain ca | |
| **hsa-miR-204** | ↓ | breast, colon, lung ca | [71], [83], [87], [110] |
| | ↑ | leukemia (ALL), pancreas ca | |
| **hsa-miR-205** | ↓ | breast, prostate ca | [87-88], [94], [100], [104] |
| | ↑ | lung, bladder, ovary ca | |
| **hsa-miR-206** | ↓ | leukemia (CLL), breast ca | [70-71], [75], [119] |
| | ↑ | breast, prostate ca | |

*(Continued)*

*(Continued)*

| microRNA NAME | EXP (↑/↓) | CANCER ASSOCIATION | REFERENCES |
|---|---|---|---|
| **hsa-miR-210** | ↓ | leukemia (CLL), breast ca | [70-71], [75], [94], [102], [110] |
| | ↑ | leukemia (CLL, ALL), breast, lung, prostate, liver ca | |
| **hsa-miR-211** | ↓ | ovary ca | [83], [107] |
| | ↑ | pancreas ca | |
| **hsa-miR-212** | ↓ | leukemia (CLL), stomach, colon, pancreas, brain ca | [70] [72], [86], [94-95], |
| | ↑ | pancreas, lung ca | |
| **hsa-miR-213** | ↓ | leukemia (CLL), breast, brain ca | [70-72], [75], [90], [96], [103] |
| | ↑ | breast, lung, thyroid, tongue ca | |
| **hsa-miR-214** | ↓ | ovary, liver ca | [73-75], [88], [94], [97] |
| | ↑ | pancreas, prostate, stomach, lung, liver ca | |
| **hsa-miR-216** | ↓ | ovary ca | [73], [107] |
| | ↑ | liver ca | |
| **hsa-miR-217** | ↓ | pancreas ca | [70], [124] |
| | ↑ | leukemia (CLL) | |
| **hsa-miR-218-2** | ↓ | colon, stomach, prostate, pancreas, lung ca | [75], [94] |
| | ↑ | lung, breast ca | |
| **hsa-miR-219** | ↓ | brain, tongue ca | [90], [95] |
| **hsa-miR-219-1** | ↓ | tongue, lung, thyroid ca | [94], [96] |
| **hsa-miR-220** | ↓ | leukemia (CLL), lung ca | [70], [94], [96], [103] |
| | ↑ | thyroid ca | |
| **hsa-miR-221** | ↓ | leukemia (CLL), prostate, brain ca | [70], [73-75], [82], [86], [88-[89], [95-97], [99], [33], [101-105] |
| | ↑ | leukemia (AML), brain, pancreas, liver, thyroid, ovary, bladder, colon, stomach ca | |
| **hsa-miR-222** | ↓ | leukemia (CLL), prostate, ovary ca | [73], [80], [82], [88], [91], [96], [99-103], [105], [107] |
| | ↑ | leukemia (AML), pancreas, thyroid, ovary, stomach, colon, liver, bladder ca | |
| **hsa-miR-223** | ↓ | leukemia (CLL, AML), liver, brain ca | [70], [72], [74-75], [97], [104-105], [116] |
| | ↑ | colon, pancreas, prostate, stomach, liver ca | |
| **hsa-miR-224** | ↓ | ovary, lung, brain ca | [72-73], [88], [91], [94], [103], [107], [112] |
| | ↑ | liver, pancreas, thyroid, ovary | |
| **hsa-miR-296** | ↑ | leukemia (CLL), prostate ca | [102], [110] |
| **hsa-miR-301** | ↑ | leukemia (CLL, ALL), pancreas, liver ca | [73-74], [86], [110] |
| **hsa-miR-320** | ↓ | leukemia (CLL), brain, breast, bladder ca | [79], [87], [95], [102], [110], [120] |
| | ↑ | leukemia (ALL), retinoblastoma, prostate ca | |
| **hsa-miR-324-5p** | ↓ | brain ca | [73], [95], [99], [110] |
| | ↑ | leukemia (CLL, ALL, AML), liver ca | |
| **hsa-miR-325** | ↑ | tongue ca | [90], [99] |
| | ↓ | leukemia (AML) | |
| **hsa-miR-326** | ↓ | brain, pancreas ca | [83], [95], [99] |
| | ↑ | leukemia (AML) | |
| **hsa-miR-328** | ↓ | colorectal, brain ca | [76], [95], [99] |
| | ↑ | leukemia (AML) | |

*(Continued)*

*(Continued)*

| microRNA NAME | EXP (↑/↓) | CANCER ASSOCIATION | REFERENCES |
|---|---|---|---|
| **hsa-miR-331** | ↓ | brain ca | [95], [99], [110] |
| | ↑ | leukemia (CLL, ALL, AML) | |
| **hsa-miR-335** | ↓ | breast ca | [110], [119] |
| | ↑ | leukemia (CLL, ALL) | |
| **hsa-miR-340** | ↓ | leukemia (CLL, ALL) | [87], [99], [110] |
| | ↑ | leukemia (AML, ALL) | |
| **hsa-miR-342** | ↓ | ovary ca | [83], [88] |
| | ↑ | pancreas ca | |
| **hsa-miR-345** | ↓ | pancreas, thyroid ca | [74], [86], [96] |
| **hsa-miR-346** | ↓ | brain ca | [95], [122] |
| | ↑ | thyroid ca | |
| **hsa-miR-370** | ↓ | brain ca | [95], [102] |
| | ↑ | prostate ca | |
| **hsa-miR-372** | ↓ | leukemia (AML), brain ca | [90], [95], [99] |
| | ↑ | tongue ca | |
| **hsa-miR-373*** | ↑ | retinoblastoma, prostate ca | [102], [120] |
| **hsa-miR-374** | ↑ | leukemia (AML), liver ca | [73], [99] |
| **hsa-miR-375** | ↓ | pancreas, liver ca | [86], [91], [100] |
| | ↑ | prostate ca | |
| **hsa-miR-377** | ↑ | ovary, liver ca | [74], [88] |
| **hsa-miR-422b** | ↓ | liver ca | [88], [91] |
| | ↑ | ovary ca | |
| **hsa-miR-494** | ↓ | head and neck ca | [85], [120] |
| | ↑ | retinoblastoma | |
| **hsa-miR-497** | ↓ | gastric, breast, prostate ca | [77], [79], [102] |
| **hsa-miR-498** | ↑ | retinoblastoma, prostate ca | [102], [120] |

ALL-acute lymphoblastic leukemia; AML-acute myeloid leukemia; CLL-chronic lymphocytic leukemia; (↑) high expression; (↓) low expression; ca-cancer.

Reference numbers correspond to publications listed in ANNEX L.

# ANNEX L:

The list of studies reviewed for cancer deregulated miRNAs (Ferdin et al., 2010)


| | |
|---|---|
| 27. | Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R., Golub, T. R., MicroRNA expression profiles classify human cancers. *Nature 435* (7043), 834-838 (2005). |
| 28. | Calin, G. A., Ferracin, M., Cimmino, A., Di Leva, G., Shimizu, M., Wojcik, S. E., Iorio, M. V., Visone, R., Sever, N. I., Fabbri, M., Iuliano, R., Palumbo, T., Pichiorri, F., Roldo, C., Garzon, R., Sevignani, C., Rassenti, L., Alder, H., Volinia, S., Liu, C. G., Kipps, T. J., Negrini, M., Croce, C. M. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. *N Engl J Med 353* (17), 1793-1801 (2005). |
| 42. | Schetter, A. J., Leung, S. Y., Sohn, J. J., Zanetti, K. A., Bowman, E. D., Yanaihara, N., Yuen, S. T., Chan, T. L., Kwong, D. L., Au, G. K., Liu, C. G., Calin, G. A., Croce, C. M., Harris, C. C. MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *JAMA 299* (4), 425-436 (2008). |
| 70. | Calin, G. A., Liu, C. G., Sevignani, C., Ferracin, M., Felli, N., Dumitru, C. D., Shimizu, M., Cimmino, A., Zupo, S., Dono, M., Dell'Aquila, M. L., Alder, H., Rassenti, L., Kipps, T. J., Bullrich, F., Negrini, M., Croce, C. M. MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc Natl Acad Sci U S A 101* (32), 11755-11760 (2004). |
| 71. | Iorio, M. V., Ferracin, M., Liu, C. G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M., Menard, S., Palazzo, J. P., Rosenberg, A., Musiani, P., Volinia, S., Nenci, I., Calin, G. A., Querzoli, P., Negrini, M., Croce, C. M. MicroRNA gene expression deregulation in human breast cancer. *Cancer Res 65* (16), 7065-7070 (2005). |
| 72. | Bottoni, A., Zatelli, M. C., Ferracin, M., Tagliati, F., Piccin, D., Vignali, C., Calin, G. A., Negrini, M., Croce, C. M., Degli Uberti, E. C. Identification of differentially expressed microRNAs by microarray: a possible role for microRNA genes in pituitary adenomas. *J Cell Physiol 210* (2), 370-377 (2007). |
| 73. | Wang, Y., Lee, A. T., Ma, J. Z., Wang, J., Ren, J., Yang, Y., Tantoso, E., Li, K. B., Ooi, L. L., Tan, P., Lee, C. G. Profiling microRNA expression in hepatocellular carcinoma reveals microRNA-224 up-regulation and apoptosis inhibitor-5 as a microRNA-224-specific target. *J Biol Chem 283* (19), 13205-13215 (2008). |
| 74. | Jiang, J., Gusev, Y., Aderca, I., Mettler, T. A., Nagorney, D. M., Brackett, D. J., Roberts, L. R., Schmittgen, T. D. Association of MicroRNA expression in hepatocellular carcinomas with hepatitis infection, cirrhosis, and patient survival. *Clin Cancer Res 14* (2), 419-427 (2008). |
| 75. | Volinia, S., Calin, G. A., Liu, C. G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., Prueitt, R. L., Yanaihara, N., Lanza, G., Scarpa, A., Vecchione, A., Negrini, M., Harris, C. C., Croce, C. M. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci U S A 103* (7), 2257-2261 (2006). |
| 76. | Bandres, E., Cubedo, E., Agirre, X., Malumbres, R., Zarate, R., Ramirez, N., Abajo, A., Navarro, A., Moreno, I., Monzo, M., Garcia-Foncillas, J. Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues. *Mol Cancer 5*, 29 (2006). |


*(Continued)*

*(Continued)*


| | |
|---|---|
| 77. | Guo, J., Miao, Y., Xiao, B., Huan, R., Jiang, Z., Meng, D., Wang, Y. Differential expression of microRNA species in human gastric cancer versus non-tumorous tissues. *J Gastroenterol Hepatol 24* (4), 652-657 (2009). |
| 78. | Slaby, O. , Svoboda, M., Fabian, P., Smerdova, T., Knoflickova, D., Bednarikova, M., Nenutil, R., Vyzula, R. Altered expression of miR-21, miR-31, miR-143 and miR-145 is related to clinicopathologic features of colorectal cancer. *Oncology 72* (5-6), 397-402 (2007). |
| 79. | Yan, L. X., Huang, X. F., Shao, Q., Huang, M. Y., Deng, L., Wu, Q. L., Zeng, Y. X., Shao, J. Y. MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. *RNA 14* (11), 2348-2360 (2008). |
| 80. | Fulci, V., Chiaretti, S., Goldoni, M., Azzalin, G., Carucci, N., Tavolaro, S., Castellano, L., Magrelli, A., Citarella, F., Messina, M., Maggio, R., Peragine, N., Santangelo, S., Mauro, F. R., Landgraf, P., Tuschl, T., Weir, D. B., Chien, M., Russo, J. J., Ju, J., Sheridan, R., Sander, C., Zavolan, M., Guarini, A., Foa, R., Macino, G. Quantitative technologies establish a novel microRNA profile of chronic lymphocytic leukemia. *Blood 109* (11), 4944-4951 (2007). |
| 81. | Si, M. L., Zhu, S., Wu, H., Lu, Z., Wu, F., Mo, Y. Y. miR-21-mediated tumor growth. *Oncogene 26* (19), 2799-2803 (2007). |
| 82. | Bloomston, M., Frankel, W. L., Petrocca, F., Volinia, S., Alder, H., Hagan, J. P., Liu, C. G., Bhatt, D., Taccioli, C., Croce, C. M. MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis. *JAMA 297* (17), 1901-1908 (2007). |
| 83. | Roldo, C., Missiaglia, E., Hagan, J. P., Falconi, M., Capelli, P., Bersani, S., Calin, G. A., Volinia, S., Liu, C. G., Scarpa, A., Croce, C. M. MicroRNA expression abnormalities in pancreatic endocrine and acinar tumors are associated with distinctive pathologic features and clinical behavior. *J Clin Oncol 24* (29), 4677-4684 (2006). |
| 84. | Wang, T., Zhang, X., Obijuru, L., Laser, J., Aris, V., Lee, P., Mittal, K., Soteropoulos, P., Wei, J. J. A micro-RNA signature associated with race, tumor size, and target gene activity in human uterine leiomyomas. *Genes Chromosomes Cancer 46* (4), 336-347 (2007). |
| 85. | Chang, S. S., Jiang, W. W., Smith, I., Poeta, L. M., Begum, S., Glazer, C., Shan, S., Westra, W., Sidransky, D., Califano, J. A. MicroRNA alterations in head and neck squamous cell carcinoma. *Int J Cancer 123* (12), 2791-2797 (2008). |
| 86. | Lee, E. J., Gusev, Y., Jiang, J., Nuovo, G. J., Lerner, M. R., Frankel, W. L., Morgan, D. L., Postier, R. G., Brackett, D. J., Schmittgen, T. D. Expression profiling identifies microRNA signature in pancreatic cancer. *Int J Cancer 120* (5), 1046-1054 (2007). |
| 87. | Baffa, R., Fassan, M., Volinia, S., O'Hara, B., Liu, C. G., Palazzo, J. P., Gardiman, M., Rugge, M., Gomella, L. G., Croce, C. M., Rosenberg, A. MicroRNA expression profiling of human metastatic cancers identifies cancer gene targets. *J Pathol 219* (2), 214-221 (2009). |
| 88. | Lee, C. H., Subramanian, S., Beck, A. H., Espinosa, I., Senz, J., Zhu, S. X., Huntsman, D., van de Rijn, M., Gilks, C. B. MicroRNA profiling of BRCA1/2 mutation-carrying and non-mutation-carrying high-grade serous carcinomas of ovary. *PLoS One 4* (10), e7314 (2009). |
| 89. | Ciafre, S. A., Galardi, S., Mangiola, A., Ferracin, M., Liu, C. G., Sabatino, G., Negrini, M., Maira, G., Croce, C. M., Farace, M. G. Extensive modulation of a set of microRNAs in primary glioblastoma. *Biochem Biophys Res Commun 334* (4), 1351-1358 (2005). |
| 90. | Wong, T. S., Liu, X. B., Wong, B. Y., Ng, R. W., Yuen, A. P., Wei, W. I. Mature miR-184 as Potential Oncogenic microRNA of Squamous Cell Carcinoma of Tongue. *Clin Cancer Res 14* (9), 2588-2592 (2008). |


*(Continued)*

*(Continued)*


| | |
|---|---|
| 91. | Ladeiro, Y., Couchy, G., Balabaud, C., Bioulac-Sage, P., Pelletier, L., Rebouissou, S., Zucman-Rossi, J. MicroRNA profiling in hepatocellular tumors is associated with clinical features and oncogene/tumor suppressor gene mutations. *Hepatology 47* (6), 1955-1963 (2008). |
| 92. | Meng, F., Henson, R., Wehbe-Janek, H., Ghoshal, K., Jacob, S. T., Patel, T. MicroRNA-21 regulates expression of the PTEN tumor suppressor gene in human hepatocellular cancer. *Gastroenterology 133* (2), 647-658 (2007). |
| 93. | Nam, E. J., Yoon, H., Kim, S. W., Kim, H., Kim, Y. T., Kim, J. H., Kim, J. W., Kim, S. MicroRNA expression profiles in serous ovarian carcinoma. *Clin Cancer Res 14* (9), 2690-2695 (2008). |
| 94. | Yanaihara, N., Caplen, N., Bowman, E., Seike, M., Kumamoto, K., Yi, M., Stephens, R. M., Okamoto, A., Yokota, J., Tanaka, T., Calin, G. A., Liu, C. G., Croce, C. M., Harris, C. C. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell 9* (3), 189-198 (2006). |
| 95. | Ferretti, E., De Smaele, E., Po, A., Di Marcotullio, L., Tosi, E., Espinola, M. S., Di Rocco, C., Riccardi, R., Giangaspero, F., Farcomeni, A., Nofroni, I., Laneve, P., Gioia, U., Caffarelli, E., Bozzoni, I., Screpanti, I., Gulino, A. MicroRNA profiling in human medulloblastoma. *Int J Cancer 124* (3), 568-577 (2009). |
| 96. | He, H., Jazdzewski, K., Li, W., Liyanarachchi, S., Nagy, R., Volinia, S., Calin, G. A., Liu, C. G., Franssila, K., Suster, S., Kloos, R. T., Croce, C. M., de la Chapelle, A. The role of microRNA genes in papillary thyroid carcinoma. *Proc Natl Acad Sci U S A 102* (52), 19075-19080 (2005). |
| 97. | Gramantieri, L., Ferracin, M., Fornari, F., Veronese, A., Sabbioni, S., Liu, C. G., Calin, G. A., Giovannini, C., Ferrazzi, E., Grazi, G. L., Croce, C. M., Bolondi, L., Negrini, M. Cyclin G1 is a target of miR-122a, a microRNA frequently down-regulated in human hepatocellular carcinoma. *Cancer Res 67* (13), 6092-6099 (2007). |
| 98. | Wang, B., Majumder, S., Nuovo, G., Kutay, H., Volinia, S., Patel, T., Schmittgen, T. D., Croce, C., Ghoshal, K., Jacob, S. T. Role of microRNA-155 at early stages of hepatocarcinogenesis induced by choline-deficient and amino acid-defined diet in C57BL/6 mice. *Hepatology 50* (4), 1152-1161 (2009). |
| 99. | Dixon-McIver, A., East, P., Mein, C. A., Cazier, J. B., Molloy, G., Chaplin, T., Andrew Lister, T., Young, B. D., Debernardi, S. Distinctive patterns of microRNA expression associated with karyotype in acute myeloid leukaemia. *PLoS One 3* (5), e2141 (2008). |
| 100. | Schaefer, A., Jung, M., Mollenkopf, H. J., Wagner, I., Stephan, C., Jentzmik, F., Miller, K., Lein, M., Kristiansen, G., Jung, K. Diagnostic and prognostic implications of microRNA profiling in prostate carcinoma. *Int J Cancer* 000, 000-000 (2009). |
| 101. | Tong, A. W., Fulgham, P., Jay, C., Chen, P., Khalil, I., Liu, S., Senzer, N., Eklund, A. C., Han, J., Nemunaitis, J. MicroRNA profile analysis of human prostate cancers. *Cancer Gene Ther 16* (3), 206-216 (2009). |
| 102. | Porkka, K. P., Pfeiffer, M. J., Waltering, K. K., Vessella, R. L., Tammela, T. L., Visakorpi, T. MicroRNA expression profiling in prostate cancer. *Cancer Res 67* (13), 6130-6135 (2007). |
| 103. | Pallante, P., Visone, R., Ferracin, M., Ferraro, A., Berlingieri, M. T., Troncone, G., Chiappetta, G., Liu, C. G., Santoro, M., Negrini, M., Croce, C. M., Fusco, A. MicroRNA deregulation in human thyroid papillary carcinomas. *Endocr Relat Cancer 13* (2), 497-508 (2006). |
| 104. | Gottardo, F., Liu, C. G., Ferracin, M., Calin, G. A., Fassan, M., Bassi, P., Sevignani, C., Byrne, D., Negrini, M., Pagano, F., Gomella, L. G., Croce, C. M., Baffa, R. Micro-RNA profiling in kidney and bladder cancers. *Urol Oncol 25* (5), 387-392 (2007). |


*(Continued)*

*(Continued)*


| | |
|---|---|
| 105. | Wong, Q. W., Lung, R. W., Law, P. T., Lai, P. B., Chan, K. Y., To, K. F., Wong, N. MicroRNA-223 is commonly repressed in hepatocellular carcinoma and potentiates expression of Stathmin1. *Gastroenterology 135* (1), 257-269 (2008). |
| 106. | Ozen, M., Creighton, C. J., Ozdemir, M., Ittmann, M. Widespread deregulation of microRNA expression in human prostate cancer. *Oncogene 27* (12), 1788-1793 (2008). |
| 107. | Iorio, M. V., Visone, R., Di Leva, G., Donati, V., Petrocca, F., Casalini, P., Taccioli, C., Volinia, S., Liu, C. G., Alder, H., Calin, G. A., Menard, S., Croce, C. M. MicroRNA signatures in human ovarian cancer. *Cancer Res 67* (18), 8699-8707 (2007). |
| 108. | Pekarsky, Y., Santanam, U., Cimmino, A., Palamarchuk, A., Efanov, A., Maximov, V., Volinia, S., Alder, H., Liu, C. G., Rassenti, L., Calin, G. A., Hagan, J. P., Kipps, T., Croce, C. M. Tcl1 expression in chronic lymphocytic leukemia is regulated by miR-29 and miR-181. *Cancer Res 66* (24), 11590-11593 (2006). |
| 109. | Xi, Y., Formentini, A., Chien, M., Weir, D. B., Russo, J. J., Ju, J., Kornmann, M. Prognostic Values of microRNAs in Colorectal Cancer. *Biomark Insights 2*, 113-121 (2006). |
| 110. | Zanette, D. L., Rivadavia, F., Molfetta, G. A., Barbuzano, F. G., Proto-Siqueira, R., Silva-Jr, W. A., Falcao, R. P., Zago, M. A. miRNA expression profiles in chronic lymphocytic and acute lymphocytic leukemia. *Braz J Med Biol Res 40* (11), 1435-1440 (2007). |
| 111. | Visone, R., Pallante, P., Vecchione, A., Cirombella, R., Ferracin, M., Ferraro, A., Volinia, S., Coluzzi, S., Leone, V., Borbone, E., Liu, C. G., Petrocca, F., Troncone, G., Calin, G. A., Scarpa, A., Colato, C., Tallini, G., Santoro, M., Croce, C. M., Fusco, A. Specific microRNAs are downregulated in human thyroid anaplastic carcinomas. *Oncogene 26* (54), 7590-7595 (2007). |
| 112. | Murakami, Y., Yasuda, T., Saigo, K., Urashima, T., Toyoda, H., Okanoue, T., Shimotohno, K. Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and nontumorous tissues. *Oncogene 25* (17), 2537-2545 (2006). |
| 113. | Lanza, G., Ferracin, M., Gafa, R., Veronese, A., Spizzo, R., Pichiorri, F., Liu, C. G., Calin, G. A., Croce, C. M., Negrini, M. mRNA/microRNA gene expression profile in microsatellite unstable colorectal cancer. *Mol Cancer 6*, 54 (2007). |
| 114. | Silber, J., Lim, D. A., Petritsch, C., Persson, A. I., Maunakea, A. K., Yu, M., Vandenberg, S. R., Ginzinger, D. G., James, C. D., Costello, J. F., Bergers, G., Weiss, W. A., Alvarez-Buylla, A., Hodgson, J. G. miR-124 and miR-137 inhibit proliferation of glioblastoma multiforme cells and induce differentiation of brain tumor stem cells. *BMC Med 6*, 14 (2008). |
| 115. | Huang, Z. M., Yang, J., Shen, X. Y., Zhang, X. Y., Meng, F. S., Xu, J. T., Zhang, B. F., Gao, H. J. MicroRNA expression profile in non-cancerous colonic tissue associated with lymph node metastasis of colon cancer. *J Dig Dis 10* (3), 188-194 (2009). |
| 116. | Mi, S., Lu, J., Sun, M., Li, Z., Zhang, H., Neilly, M. B., Wang, Y., Qian, Z., Jin, J., Zhang, Y., Bohlander, S. K., Le Beau, M. M., Larson, R. A., Golub, T. R., Rowley, J. D., Chen, J. MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia. *Proc Natl Acad Sci U S A 104* (50), 19971-19976 (2007). |
| 117. | Bonci, D., Coppola, V., Musumeci, M., Addario, A., Giuffrida, R., Memeo, L., D'Urso, L., Pagliuca, A., Biffoni, M., Labbaye, C., Bartucci, M., Muto, G., Peschle, C., De Maria, R. The miR-15a-miR-16-1 cluster controls prostate cancer by targeting multiple oncogenic activities. *Nat Med 14* (11), 1271-1277 (2008). |
| 118. | Ma, L., Teruya-Feldstein, J., Weinberg, R. A. Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. *Nature 449* (7163), 682-688 (2007). |


*(Continued)*

*(Continued)*


| | |
|---|---|
| 119. | Tavazoie, S. F., Alarcon, C., Oskarsson, T., Padua, D., Wang, Q., Bos, P. D., Gerald, W. L., Massague, J. Endogenous human microRNAs that suppress breast cancer metastasis. *Nature 451* (7175), 147-152 (2008). |
| 120. | Zhao, J. J., Yang, J., Lin, J., Yao, N., Zhu, Y., Zheng, J., Xu, J., Cheng, J. Q., Lin, J. Y., Ma, X. Identification of miRNAs associated with tumorigenesis of retinoblastoma by miRNA microarray analysis. *Childs Nerv Syst 25* (1), 13-20 (2009). |
| 121. | Ambs, S., Prueitt, R. L., Yi, M., Hudson, R. S., Howe, T. M., Petrocca, F., Wallace, T. A., Liu, C. G., Volinia, S., Calin, G. A., Yfantis, H. G., Stephens, R. M., Croce, C. M. Genomic profiling of microRNA and messenger RNA reveals deregulated microRNA expression in prostate cancer. *Cancer Res 68* (15), 6162-6170 (2008). |
| 122. | Weber, F., Teresi, R. E., Broelsch, C. E., Frilling, A., Eng, C. A limited set of human MicroRNA is deregulated in follicular thyroid carcinoma. *J Clin Endocrinol Metab 91* (9), 3584-3591 (2006). |
| 123. | Lehmann, U., Hasemeier, B., Christgen, M., Muller, M., Romermann, D., Langer, F., Kreipe, H. Epigenetic inactivation of microRNA gene hsa-mir-9-1 in human breast cancer. *J Pathol 214* (1), 17-24 (2008). |
| 124. | Szafranska, A. E., Davison, T. S., John, J., Cannon, T., Sipos, B., Maghnouj, A., Labourier, E., Hahn, S. A. MicroRNA expression alterations are linked to tumorigenesis and non-neoplastic processes in pancreatic ductal adenocarcinoma. *Oncogene 26* (30), 4442-4452 (2007). |


The number infront of each publication corresponds to the references listed in ANNEX J and K.

# ANNEX M:

Top ten chromosome locations associated with miRNAs and human cancers (Ferdin et al., 2011)

| Genomic location | microRNA | CANCER type | Ref.erences |
|---|---|---|---|
| **4q25** | hsa-miR-302a, hsa-miR-302c | ovary | [1] |
| | hsa-miR-302b | leukemia (AML), ovary | [1] [2] |
| | hsa-miR-302d, hsa-miR-367 | leukemia (AML) | [2] |
| **7q32.2** | hsa-miR-96 | leukemia (CLL, ALL), colorectal, liver, prostate | [3] [4] [5] [6] [7] [8] |
| | hsa-miR-182 | leukemia (CLL, ALL, AML), liver, blood, prostate | [4] [5] [2] [9] [6] [10] [8] |
| | hsa-miR-183 | leukemia (CLL), colorectal, prostate, liver | [3] [8] [6] [5] |
| | hsa-miR-335 | leukemia (CLL, ALL), breast | [11] [4] |
| **9q22.32** | hsa-let-7a-1 | liver, brain | [12] [13] |
| | hsa-let-7b | leukemia (CLL, ALL), blood, liver, ovary, prostate, brain | [12] [10] [14] [4] [15] [16] [17] |
| | hsa-let-7d | leukemia (CLL), ovary, pancreas, brain, breast, liver, prostate | [1] [18] [16] [12] [5] [19] [17] |
| | hsa-let-7f-1 | leukemia (CLL), brain, thyroid, pancreas | [5] [13] [20] [19] |
| | hsa-miR-23b | brain, prostate, bladder, uterus | [13] [16] [21] [22] [23] |
| | hsa-miR-24-1 | leukemia (CLL), brain, colon, pancreas, stomach, thyroid | [5] [13] [24] [19] [25] |
| | hsa-miR-27b | leukemia (CLL), lung | [5] [26] |
| **13q31.3** | hsa-miR-17 | leukemia (CLL, ALL), stomach, breast, colon, lung, pancreas, prostate, blood, tongue, bladder, brain | [27] [4] [5] [24] [26] [10] [28] [22] [29] [17] |
| | hsa-miR-18a | stomach, ovary | [27] [15] |
| | hsa-miR-19a | leukemia (CLL, ALL), colorectal, lymphatic gland, lung, stomach, brain | [3] [5] [4] [27] [17] |
| | hsa-miR-20a | leukemia (CLL), breast, colon, pancreas, prostate, stomach, blood, brain | [24] [5] [29] [30] [27] [10] [21] [17] |
| | hsa-miR-92a-1 | leukemia (CLL), pancreas | [5] [19] |
| **14q32.2** | hsa-miR-127 | breast, brain, colon | [31] [29] [13] [17] |
| | hsa-miR-136 | leukemia (CLL), brain, breast, liver | [12] [18] [13] [5] |
| | hsa-miR-342 | blood, pancreas, bladder, colon | [10] [32] [29] |
| | hsa-miR-345 | pancreas, thyroid | [19] [25] |
| | hsa-miR-370 | brain, prostate | [17] [16] |
| **14q32.31** | hsa-miR-134, hsa-miR-323, hsa-miR-381, hsa-miR-382 | brain | [13] [17] |
| | hsa-miR-154 | leukemia (CLL), brain, tongue | [17] [13] [5] [28] |
| | hsa-miR-369 | lung | [29] |
| | hsa-miR-376a | pancreas | [19] |
| | hsa-miR-377 | blood | [10] |
| | hsa-miR-409 | colon | [29] |
| | hsa-miR-494 | head and neck, eye | [33] [34] |
| **19p13.13** | hsa-miR-23a | leukemia (CLL), prostate, pancreas, bladder, brain | [13] [5] [16] [32] [22] |
| | hsa-miR-24-2 | leukemia (CLL), brain, colon, stomach, lung, pancreas, thyroid | [5] [13] [24] [26] [19] [25] |
| | hsa-miR-27a | leukemia (CLL, ALL, AML) | [2] [4] |

*(Continued)*

*(Continued)*

| Genomic location | microRNA | CANCER type | Ref.erences |
|---|---|---|---|
| | hsa-miR-181c | leukemia (CLL, ALL, AML), lung, brain, pancreas, liver, thyroid, tongue | [12] [35] [17] [26] [4] [2] [5] [19] [25] [20] [28] |
| | hsa-miR-181d | breast, pancreas | [31] [36] |
| **19q13.42** | hsa-miR-372 | leukemia (AML), tongue, brain | [28] [2] [17] |
| | hsa-miR-373 | lung, eye, prostate | [29] [34] [16] |
| | hsa-miR-498 | prostate, eye | [16] [34] |
| | hsa-miR-518c | eye | [34] [29] |
| **Xq26.2** | hsa-miR-18b | stomach | [27] |
| | hsa-miR-20b | colon, stomach, blood | [29] [27] [10] |
| | hsa-miR-92a-2 | leukemia (CLL), pancreas, prostate, stomach | [5] [24] |
| | hsa-miR-106a | leukemia (CLL, ALL), stomach, blood, lung, colon, pancreas, prostate, brain | [29] [30] [4] [27] [10] [26] [24] [17] |
| | hsa-miR-363 | blood | [10] |
| | hsa-miR-424 | pancreas | [19] |
| **Xq27.3** | hsa-miR-508, hsa-miR-509, hsa-miR-514 | blood | [10] |
| | hsa-miR-513a | eye | [34] |

CLL: chronic lymphocytic leukemia; ALL: acute lymphocytic leukemia;

Reference numbers correspond to publications listed in ANNEX N.

# ANNEX N:

The list of studies reviewed to define genomic locations of miRNAs altered in human cancers (Ferdin et al., 2011)

| | |
|---|---|
| 1. | Iorio, M., et al., MicroRNA signatures in human ovarian cancer. Cancer Res, 2007. 67(18): p. 8699-707. |
| 2. | Dixon-McIver, A., et al., Distinctive patterns of microRNA expression associated with karyotype in acute myeloid leukaemia. PLoS One, 2008. 3(5): p. e2141. |
| 3. | Bandres, E., et al., Identification by Real-time PCR of 13 mature microRNAs differentially expressed in colorectal cancer and non-tumoral tissues. Mol Cancer, 2006. 5: p. 29. |
| 4. | Zanette, D., et al., miRNA expression profiles in chronic lymphocytic and acute lymphocytic leukemia. Braz J Med Biol Res, 2007. 40(11): p. 1435-40. |
| 5. | Calin, G., et al., MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. Proc Natl Acad Sci U S A, 2004. 101(32): p. 11755-60. |
| 6. | Wang, Y., et al., Profiling microRNA expression in hepatocellular carcinoma reveals microRNA-224 up-regulation and apoptosis inhibitor-5 as a microRNA-224-specific target. J Biol Chem, 2008. 283(19): p. 13205-15. |
| 7. | Ladeiro, Y., et al., MicroRNA profiling in hepatocellular tumors is associated with clinical features and oncogene/tumor suppressor gene mutations. Hepatology, 2008. 47(6): p. 1955-63. |
| 8. | Schaefer, A., et al., Diagnostic and prognostic implications of microRNA profiling in prostate carcinoma. Int J Cancer, 2009. |
| 9. | Wong, Q., et al., MicroRNA-223 is commonly repressed in hepatocellular carcinoma and potentiates expression of Stathmin1. Gastroenterology, 2008. 135(1): p. 257-69. |
| 10. | Lee, C.H., et al., MicroRNA profiling of BRCA1/2 mutation-carrying and non-mutation-carrying high-grade serous carcinomas of ovary. PLoS One, 2009. 4(10): p. e7314. |
| 11. | Tavazoie, S., et al., Endogenous human microRNAs that suppress breast cancer metastasis. Nature, 2008. 451(7175): p. 147-52. |
| 12. | Gramantieri, L., et al., Cyclin G1 is a target of miR-122a, a microRNA frequently down-regulated in human hepatocellular carcinoma. Cancer Res, 2007. 67(13): p. 6092-9. |
| 13. | Bottoni, A., et al., Identification of differentially expressed microRNAs by microarray: a possible role for microRNA genes in pituitary adenomas. J Cell Physiol, 2007. 210(2): p. 370-7. |
| 14. | Mi, S., et al., MicroRNA expression signatures accurately discriminate acute lymphoblastic leukemia from acute myeloid leukemia. Proc Natl Acad Sci U S A, 2007. 104(50): p. 19971-6. |
| 15. | Nam, E., et al., MicroRNA expression profiles in serous ovarian carcinoma. Clin Cancer Res, 2008. 14(9): p. 2690-5. |
| 16. | Porkka, K., et al., MicroRNA expression profiling in prostate cancer. Cancer Res, 2007. 67(13): p. 6130-5. |
| 17. | Ferretti, E., et al., MicroRNA profiling in human medulloblastoma. Int J Cancer, 2009. 124(3): p. 568-77. |
| 18. | Iorio, M.V., et al., MicroRNA gene expression deregulation in human breast cancer. Cancer Res, 2005. 65(16): p. 7065-70. |
| 19. | Lee, E.J., et al., Expression profiling identifies microRNA signature in pancreatic cancer. Int J Cancer, 2007. 120(5): p. 1046-54. |

*(Continued)*

*(Continued)*


| | |
|---|---|
| 20. | Pallante, P., et al., MicroRNA deregulation in human thyroid papillary carcinomas. Endocr Relat Cancer, 2006. 13(2): p. 497-508. |
| 21. | Tong, A., et al., MicroRNA profile analysis of human prostate cancers. Cancer Gene Ther, 2009. 16(3): p. 206-16. |
| 22. | Gottardo, F., et al., Micro-RNA profiling in kidney and bladder cancers. Urol Oncol, 2007. 25(5): p. 387-92. |
| 23. | Wang, T., et al., A micro-RNA signature associated with race, tumor size, and target gene activity in human uterine leiomyomas. Genes Chromosomes Cancer, 2007. 46(4): p. 336-47. |
| 24. | Volinia, S., et al., A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A, 2006. 103(7): p. 2257-61. |
| 25. | He, H., et al., The role of microRNA genes in papillary thyroid carcinoma. Proc Natl Acad Sci U S A, 2005. 102(52): p. 19075-80. |
| 26. | Yanaihara, N., et al., Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell, 2006. 9(3): p. 189-98. |
| 27. | Guo, J., et al., Differential expression of microRNA species in human gastric cancer versus non-tumorous tissues. J Gastroenterol Hepatol, 2009. 24(4): p. 652-7. |
| 28. | Wong, T.S., et al., Mature miR-184 as Potential Oncogenic microRNA of Squamous Cell Carcinoma of Tongue. Clin Cancer Res, 2008. 14(9): p. 2588-92. |
| 29. | Baffa, R., et al., MicroRNA expression profiling of human metastatic cancers identifies cancer gene targets. J Pathol, 2009. 219(2): p. 214-21. |
| 30. | Schetter, A., et al., MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. JAMA, 2008. 299(4): p. 425-36. |
| 31. | Yan, L.X., et al., MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. RNA, 2008. 14(11): p. 2348-60. |
| 32. | Roldo, C., et al., MicroRNA expression abnormalities in pancreatic endocrine and acinar tumors are associated with distinctive pathologic features and clinical behavior. J Clin Oncol, 2006. 24(29): p. 4677-84. |
| 33. | Chang, S.S., et al., MicroRNA alterations in head and neck squamous cell carcinoma. Int J Cancer, 2008. 123(12): p. 2791-7. |
| 34. | Zhao, J., et al., Identification of miRNAs associated with tumorigenesis of retinoblastoma by miRNA microarray analysis. Childs Nerv Syst, 2009. 25(1): p. 13-20. |
| 35. | Ciafre, S.A., et al., Extensive modulation of a set of microRNAs in primary glioblastoma. Biochem Biophys Res Commun, 2005. 334(4): p. 1351-8. |
| 36. | Bloomston, M., et al., MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis. JAMA, 2007. 297(17): p. 1901-8. |


The number infront of each publication corresponds to the references listed in ANNEX N.

# ANNEX O:

Pathways associated with the literature-collected HIF1A target genes

| # | Database | *Pathway* | *P*-value | Bonferroni | HIF1A targets involved in the pathway | no. of targets |
|---|---|---|---|---|---|---|
| **1** | **Reactome** | Metabolism of carbohydrates | 7.6E-18 | 2.9E-16 | *GCK*, *BSG*, *PGK1*, *PCK1*, *SLC2A3*, *PFKFB4*, *PFKL*, *SLC2A1*, *TPI1*, *PFKFB3*, *PFKFB2*, *GAPDH*, *PFKFB1*, *GYS1*, *HK2*, *HK1*, *GBE1*, *PGAM1*, *ALDOA*, *ALDOC*, *LDHA*, *ENO1*, *PKM*, *PFKP*, *GPI*, | 25 |
| **2** | **PANTHER** | Glycolysis | 8.6E-11 | 4.6E-9 | *HK2*, *HK1*, *GCK*, *PGAM1*, *PGK1*, *ALDOA*, *ALDOC*, *ENO1*, *PKM*, *PFKP*, *GPI*, *PFKL* | 12 |
| **3** | **KEGG** | Glycolysis / Gluconeogenesis * | 5.7E-11 | 6.4E-9 | *HK2*, *HK1*, *GCK*, *PGAM1*, *PGK1*, *ALDOA*, *ALDOC*, *PCK1*, *PFKL*, *LDHA*, *TPI1*, *GAPDH*, *ENO1*, *PKM*, *PFKP*, *GPI* | 16 |
| **4** | **Reactome** | Diabetes pathways | 2.5E-8 | 9.6E-7 | *GCK*, *PGK1*, *KCNB1*, *PFKFB4*, *PFKL*, *IGF2*, *TPI1*, *PFKFB3*, *IGFBP1*, *PFKFB2*, *GAPDH*, *IGFBP2*, *IGFBP3*, *DDIT3*, *PFKFB1*, *MMP2*, *PGAM1*, *MMP1*, *ALDOA*, *ALDOC*, *IGFBP5*, *ERO1L*, *KLK3*, *ENO1*, *PKM*, *PFKP*, *GPI* | 27 |
| **5** | **KEGG** | Fructose and mannose metabolism * | 1.9E-8 | 2.1E-6 | *HK2*, *HK1*, *TPI1*, *PFKFB3*, *PFKFB2*, *ALDOA*, *PFKFB1*, *ALDOC*, *PFKP*, *PFKFB4*, *PFKL* | 11 |
| **6** | **KEGG** | Renal cell carcinoma * | 4.7E-7 | 5.2E-5 | *PDGFB*, *TGFA*, *EGLN1*, *SLC2A1*, *VEGFA*, *ARNT*, *VEGFB*, *ETS1*, *MET*, *RAC1*, *EGLN3*, *CDC42*, *TGFB3* | 13 |
| **7** | **BioCarta** | Glycolysis pathway | 5.0E-7 | 6.4E-5 | *HK1*, *PGAM1*, *TPI1*, *PGK1*, *GAPDH*, *ENO1*, *GPI*, *PFKL* | 8 |
| **8** | **Reactome** | Integration of energy metabolism | 2.4E-5 | 9.0E-4 | *GCK*, *PGAM1*, *PGK1*, *PDK3*, *ALDOA*, *ALDOC*, *PDK1*, *PFKFB4*, *PFKL*, *IGF2*, *TPI1*, *PFKFB3*, *PFKFB2*, *GAPDH*, *PFKFB1*, *ENO1*, *PKM*, *PFKP*, *GPI* | 19 |
| **9** | **KEGG** | Pathways in cancer | 1.4E-5 | 1.6E-3 | *EGFR*, *FN1*, *CDKN1B*, *CDKN1A*, *SLC2A1*, *KITLG*, *VEGFA*, *VEGFB*, *HSP90B1*, *MET*, *TGFB3*, *MMP2*, *PDGFA*, *MMP1*, *PDGFB*, *TGFA*, *BCL2L1*, *NOS2*, *EGLN1*, *KLK3*, *ARNT*, *ETS1*, *RAC1*, *EGLN3*, *CDC42P2*, | 25 |
| **10** | **PANTHER** | Fructose galactose metabolism | 3.8E-4 | 2.0E-2 | *HK2*, *HK1*, *GCK*, *ALDOA*, *ALDOC* | 5 |
| **11** | **Reactome** | Hemostasis | 6.5E-4 | 2.4E-2 | *MMP1*, *PDGFA*, *PDGFB*, *BSG*, *ALDOA*, *FN1*, *CAV1*, *ANGPT2*, *SERPINE1*, *VEGFA*, *ITGB2*, *VEGFB*, *TF*, *TTN*, *RAC1*, *PLAUR*, *TGFB3* | 17 |

*(Continued)*

*(Continued)*

| # | Database | *Pathway* | *P*-value | Bonferroni | HIF1A targets involved in the pathway | no. of targets |
|---|---|---|---|---|---|---|
| **12** | **Reactome** | Signaling by VEGF | 3.0E-3 | 1.1E-1 | *VEGFA, VEGFB, FLT1, KDR* | 4 |
| **13** | **KEGG** | Cytokine-cytokine receptor interaction * | 1.1E-3 | 1.2E-1 | *PDGFA, EPO, PDGFB, EGFR, CTF1, FLT1, KDR, GHR, INHBB, CXCL12, CX3CR1, KITLG, VEGFA, LEP, VEGFB, CXCR4, MET, TGFB3* | 18 |
| **14** | **BioCarta** | Hypoxia-inducible factor in the cardiovascular System | 1.1E-3 | 1.4E-1 | *NOS3, EPO, ARNT, VEGFA, LDHA, EDN1* | 6 |
| **15** | **KEGG** | Prostate cancer | 3.6E-3 | 3.3E-1 | *IGF2, PDGFA, KLK3, PDGFB, TGFA, EGFR, CDKN1B, CDKN1A, HSP90B1* | 9 |
| **16** | **KEGG** | Pancreatic cancer * | 4.1E-3 | 3.7E-1 | *VEGFA, TGFA, BCL2L1, EGFR, VEGFB, RAC1, CDC42, TGFB3* | 8 |
| **17** | **KEGG** | Focal adhesion | 4.5E-3 | 4.0E-1 | *PDGFA, PDGFB, EGFR, FLT1, FN1, KDR, VASP, CAV1, VEGFA, VEGFB, COL5A1, MET, RAC1, CDC42* | 14 |
| **18** | **KEGG** | Pentose phosphate pathway | 5.0E-3 | 4.3E-1 | *ALDOA, ALDOC, PFKP, GPI, PFKL* | 5 |
| **19** | **KEGG** | Galactose metabolism | 5.8E-3 | 4.8E-1 | *HK2, HK1, GCK, PFKP, PFKL* | 5 |
| **20** | **KEGG** | Bladder cancer | 6.4E-3 | 5.1E-1 | *MMP2, MMP1, VEGFA, EGFR, VEGFB, CDKN1A* | 6 |
| **21** | **KEGG** | Starch and sucrose metabolism * | 6.4E-3 | 5.1E-1 | *HK2, HK1, GCK, GBE1, GYS1, GPI* | 6 |

Grey shaded columns present pathways with statistically significant *P*-value or Bonferroni levels (< 0.001); * - enriched pathways also identified in the study by Benita et al (2009).