Genome-wide analysis of the hypoxic breast cancer transcriptome using next generation sequencing

Hani Choudhry

Wellcome Trust Centre for Human Genetics

Nuffield Department of Medicine

University of Oxford

A thesis submitted for the degree of Doctor of Philosophy

Lincare College

Hilary Term 2014

Hani Choudhry
Nuffield Department of Medicine, Lincare College
University of Oxford
Thesis submitted for the degree of Doctor of Philosophy
Hilary Term 2014

Genome-wide analysis of the hypoxic breast cancer transcriptome using next generation sequencing

Abstract

Hypoxia pathways are associated with the pathogenesis of both ischaemic and neoplastic diseases. In response to hypoxia the transcription factor hypoxia-inducible factor (HIF) induces the expression of hundreds of genes with diverse functions. These enable cells to adapt to low oxygen availability. To date, pan-genomic analyses of these transcriptional responses have focussed on protein-coding genes and microRNAs. However, the role of other classes of non-coding RNAs, in particular lncRNAs, in the hypoxia response is largely uncharacterised. My thesis aimed at improving understanding of the transcriptional regulation of the non-coding transcriptome in hypoxia. I performed an integrated genomic analysis of both non-coding and coding transcripts by massively parallel sequencing. This was interfaced with pan-genomic analyses of DNAse hypersensitivity and HIF, H3k4me3 and RNApol2 binding in hypoxic cells. These analyses have revealed that hypoxia profoundly regulated all RNA classes. snRNAs and tRNAs are globally downregulated in hypoxia, whilst miRNAs, mRNAs and lncRNAs are both up- and downregulated with an overall trend towards slight upregulation. In addition, a significant number of previously non-annotated (and largely hypoxia upregulated) transcripts were identified, including novel intergenic transcripts and natural antisense transcripts. HIF bound close to genes for mRNAs, miRNAs and lncRNAs that were upregulated by hypoxia, but was excluded from binding at genes for RNA classes that showed global downregulation. This suggests that HIF acts as a transcriptional activator (but not repressor), of lncRNAs as well as mRNAs and miRNAs. Consistent with direct regulation by HIF, many of these hypoxia-inducible, HIF-binding lncRNAs were downregulated following HIF knockdown. Analysis of RNApol2 binding and DNAse HSS signals at HIF transcriptional target genes indicated that HIF-dependent transcriptional activation occurs through release of RNApol2 that is pre-bound to open promoters of lncRNAs as well as mRNAs. In these datasets, NEAT1 was the most hypoxia-upregulated, HIF-targeted lncRNA in MCF-7 cells and, despite binding of both HIF-1 and HIF-2 isoforms at its promoter, was selectively regulated by HIF-2 alone. Furthermore, NEAT1 was induced by hypoxia in a wide range of breast cancer cell lines and in hypoxic xenograft models. Functionally, NEAT1 is required for the assembly of nuclear paraspeckle structures. Increased nuclear paraspeckle formation was observed in hypoxia and was dependent on both NEAT1 and HIF-2. Knockdown of hypoxia-induced NEAT1 significantly reduced cell proliferation and survival and induced apoptosis. Finally, high expression of NEAT1 correlated with poor clinical outcome in a large cohort of breast cancer patients. These findings extend the role of the hypoxic transcriptional response in cancer into the spectrum of non-coding transcripts and provide new insights into molecular roles of hypoxia-regulated lncRNAs, which may provide the basis for novel therapeutic targets in the future.

Declaration

Unless otherwise indicated, all the work described in this thesis was performed by the author at the Wellcome Trust Centre for Human Genetics, University of Oxford. The work reported in this thesis has not been submitted for any other degree in this or any other university or institute of learning.

Acknowledgments

My humble gratitude goes to my supervisors; Prof. Ioannis Ragoussis, Prof. Adrian Harris and Dr. David Mole for their direct supervision and help to successfully complete this research. They inspired me and enriched my growth as a researcher and as a scientist.

Thanks to:

Prof. Ioannis Ragoussis (Head of Genomics at the Wellcome Trust Centre for Human Genetics, Head of Genomics, Head of Genome Sciences, McGill University), for taking me as DPhil student and giving me the opportunity to conduct all my DPhil work in his laboratory at the Wellcome Trust Centre for Human Genetics. I have learned a lot from him about technology development and Next Generation Sequencing. I would like to thank him for his valuable advice in scientific discussion, which has helped me to pursue my future career

Prof. Adrian Harris (Head of Molecular Oncology Laboratories), for directing my project in hypoxia and breast cancer. His depth of knowledge and experience in these fields has benefited me greatly. He provided me with unflinching encouragement and an ideal environment to perform translational cancer research, which is what I always wanted.

Dr. David Mole (Clinical Senior Lecturer & Honorary Consultant) for taking me under his supervision to complete my DPhil work. His help and support were very valuable during the term of my project. I really appreciate his critical input and help in writing my paper and thesis.

Also I am grateful to Prof. Sir Peter Ratcliffe (Head of Nuffield Department of Medicine) for valuable scientific discussions about hypoxia and critical evaluation and guidance for my publication. His philosophical discussions about science have been most interesting and wise.

I like to thanks many lab members with whom I had the pleasure working and collaborating during my DPhil (Dr. Carme Camps, Dr. Spyros Oikonomopoulos, Dr. Francesca Buffa, Ashwag Albukhari, Dr. Syed Haider).

Last but never the least, I would thank and dedicate this work to my family (my Dad Mohammed Zubair, my Mom Naheed Nissar, Bilal, Ayisha, Daniah, Hana and my life partner Fatima). Words fail me when trying to show appreciation for my parents. They deserve special mention for their unconditional support, prayers, love, persistence and encouragement to achieve the best. I owe them gratitude for their support and encouragement in allowing me to pursue my ambitions. Your pride in me has given me a sense of worth.

Publications arising from this thesis

CHOUDHRY, H., SCHODEL, J., OIKONOMOPOULOS, S., CAMPS, C., GRAMPP, S., HARRIS, A. L., RATCLIFFE, P. J., RAGOUSSIS, J. & MOLE, D. R. 2014. Extensive regulation of the non-coding transcriptome by hypoxia: role of HIF in releasing paused RNApol2. *EMBO Rep*, 15, 70-6.

CAMPS, C., SAINI, H. K., MOLE, D. R., **CHOUDHRY, H.,** RECZKO, M., GUERRA-ASSUNCAO, J. A., TIAN, Y. M., BUFFA, F. M., HARRIS, A. L., HATZIGEORGIOU, A. G., ENRIGHT, A. J. & RAGOUSSIS, J. 2014. Integrated analysis of microRNA and mRNA expression and association with HIF binding reveals the complexity of microRNA expression regulation under hypoxia. *Mol Cancer*, 13, 28.

CHOUDHRY, H., ALBUKHARI, A., MOROTTI, M., HIYDER, S., MORALLI, D., GREEN, G., CUMPS, C., BUFFA, F.,CAMPS, C., RATCLIFFE, P. J., RAGOUSSIS, J., HARRIS, A. L & MOLE, D. R. 2014. Tumour hypoxia induces nuclear paraspeckle formation through HIF-2α dependent transcriptional activation of NEAT1 leading to enhanced tumourigenicity. *Oncogene, Submitted.*

Book Editor

Next Generation Sequencing in Cancer Research
Volume 1: Decoding the Cancer Genome
Editors: Wu Wei & **Hani Choudhry**
Springer, 398 pages. 59 illus. in colour, 2013
ISBN 978-1-4614-7644-3

Selected awards

- (2014) **EACR-23 Susan G. Komen for the Cure Award**, 23rd Biennial European Association for Cancer Research Congress, Germany.
- (2014) **Best Poster in Medical and Health Sciences,** 7th Saudi Scientific International Conferences-University of Edinburgh, UK.
- (2013) **the ASHG/Charles J. Epstein Trainee Award for Excellence in Human Genetics Research (SF)**, 63nd Annual Meeting of the American Society of Human Genetics, USA.
- (2013) **The** N**ational Cancer Research Institute (NCRI) Prize Award**, the NCRI Annual Cancer Conference, UK.
- (2012) **Distinguish Student Award,** from HRH Prince Mohammed bin Nawaf Al Saud, the Ambassador of Saudi Arabia in UK. The Royal Embassy of Saudi Arabia, UK.

Abbreviations

2-OG	2-oxoglutarate
aCELSR2	Antisense cadherin egf lag seven-pass g-type receptor 2
ADM	Adrenomedullin
AGO3	Argonaute risc catalytic component 3
ASO	Antisense oligonucleotides
aSPAG4	Antisense sperm associated antigen 4
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
BCKDK	Branched chain ketoacid dehydrogenase kinase
BCOR	Bcl6 corepressor
BDP1	B double prime 1, subunit of RNA polymerase iii transcription initiation
BNIP3	Adrenomedullin 3
bp	Base pairs
BRF1	RNA polymerase iii transcription initiation factor 90
CA9	Carbonic anhydrase ix
CAGE	Cap analysis of gene expression
cDNA	Complementary DNA
CELSR2	Cadherin egf lag seven-pass g-type receptor 2
ChIP	Chromatin immunoprecipitation
ChIP-seq	Chromatin immunoprecipitation sequencing
CPC	Coding potential calculator
DHX30	Deah (asp-glu-ala-his) box helicase 30

DMEM	Dulbecco's modified eagle's medium
DMOG	Dimethyloxallyl glycine
DMSO	Dimethyl sulphoxide
DNA	Deoxyribonucleic acid
DNAse1	Deoxyribonuclease 1
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
dTTP	Thymidine triphosphate
dUTP	Deoxyuridine triphosphate
EDTA	Ethylenediaminetetraacetic acid
ENCODE	Encyclopedia of DNA Elements
EPO	Erythropoietin
ER	Oestrogens receptor
FACS	Fluorescence activated cell sorting
FBS	Fetal bovine serum
FOS	Fbj murine osteosarcoma viral oncogene homolog
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GAS5	Growth arrest specific 5
GSEA	Gene set enrichment analysis
H19	Imprinted maternally expressed transcript
H3K27me3	Histone 3 lysine 27 trimethylation
H3K36me3	Histone 3 lysine 36 trimethylation
H3K4me3	Histone 3 lysine 4 trimethylation

H3K9me2	Histone 3 lysine 9 dimethylation
HATs	Histone acetyltransferases
HBS	HIF binding site
HCC	Hepatocellular carcinoma
HDACs	Histone deacetylases
HG19	Human genome release 19
HIF	Hypoxia Inducible Factor
HIF-1α	Hypoxia-Inducible Factor 1-alpha
HIF-1β	Hypoxia-Inducible Factor-1β
HIF-2α	Hypoxia-Inducible Factor 2-alpha
HINCUTs	Hypoxia-Induced Noncoding Ultraconserved Transcripts
hist1h1d	Histone cluster 1, h1d
hist1h3c	Histone cluster 1, h3c
hist1h3i	Histone cluster 1, h3i
hist1h3j	Histone cluster 1, h3j
HMTs	Histone methyltransferases
HNCC	Head and neck squamous cell carcinoma
HOTAIR	Hox transcript antisense RNA
HR	Hazard ratio
HRE	Hypoxia response element
HUVEC	Human umbilical vein endothelial cells
IPA	Ingenuity pathways analysis
ISCU	Iron-sulfur cluster scaffold homolog

kb	Kilo base
LCL	LncRNA-containing locus
lincRNA-LET	LncRNA-low expressed in tumour
LNA	Locked nucleic acid
lncRNA	Long non-coding RNA
log_2FC	Log_2 fold change
MALAT1	Metastasis associated lung adenocarcinoma transcript 1
MCF-7	Michigan Cancer Foundation-7
miRNA	MicroRNA
mRNA	Messenger RNA
MYC	V-myc avian myelocytomatosis viral oncogene homolog
NATS	Natural antisense transcripts
NCBI	National Center for Biotechnology Information
NDRG1	N-myc downstream regulated 1
NEAT1	Nuclear Enriched Abundant Transcript 1
NIT	Novel intergenic transcripts
NONO	Non-pou domain containing octamer binding
NSCLC	Non-small-cell lung carcinoma
Nt	Nucleotide
OGT	O-linked n-acetylglucosamine transferase
PBS	Phosphate buffered saline
PIWI	P-Element Induced Wimpy Testis
piwiRNA	Piwi-interacting RNA

PNK	Polynucleotide kinase
POLR	Polymerase (RNA) ii (DNA directed) polypeptide
poly A RNA	Poly-adenylated RNA
PSPC1	Paraspeckle component 1
PVRL4	Poliovirus receptor-related 4
qPCR	Quantitative real-time polymerase chain reaction
RAB11FIP5	Rab11 family interacting protein
RACE	Rapid amplification of cDNA ends
RBAK	Rb-associated krab zinc finger
RCC	Renal cell carcinoma
RCN3	Reticulocalbin 3, ef-hand calcium binding domain
RNA	Ribonucleic acid
RNA FISH	RNA fluorescent in situ hybridization
RNApol2	RNA polymerase 2
RNApol3	RNA polymerase iii
RNase P	Ribonuclease P
RNA-seq	Next-generation RNA sequencing
ROS	Reactive oxygen species
rpm	Revolutions per minute
RPMI-1640	Roswell park memorial institute-1640
rRNAs	Ribosomal RNA
RT	Reverse transcription
SAGE	Serial analysis of gene expression

SCC	Squamous cell carcinoma
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SFPQ	Splicing factor proline/gluatamine-rich
siRNA	Small interfering RNA
snoRNPs	Small nucleolar ribonucleoproteins complex
SNP	Single-nucleotide polymorphism
snRNA	Small nuclear RNA
SP1	Sp1 transcription factor
SPAG4	Sperm associated antigen 4
STON1	Stonin 1
TBE	Tris/Borate/EDTA
TCF25	Transcription factor 25
TCGA	The cancer genome atlas
TFIIIB	Transcription factor-iiib
TFIIIC	Transcription factor-iiic
TGFB1	Transforming growth factor, beta 1
TH	Tyrosine hydroxylase
TI	Transcriptional interference
TMSB10	Thymosin beta 10
TNBC	Triple-negative breast cancer
Total- RNA-seq	Ribosomally depleted directional RNA-seq
TRIS	Tris(hydroxymethyl)aminomethane

tRNALeu	tRNA leucine
tRNAs	Transfer RNA
tRNATyr	tRNA tyrosine
TSS	Transcription start site
TSSa-RNA	Transcription start site-associated RNA
UCSC	University of California, Santa Cruz
VEGF	Vascular endothelial growth factor
VHL	Von hippel-lindau
VTCN1	V-set domain containing t cell activation inhibitor 1
XIST	X (inactive)-specific transcript
ZNF320	Zinc finger protein 320

Table of Contents

Table of Figures

Chapter 1: Introduction

1 Introduction

1.1 Breast cancer

Breast cancer is the most commonly diagnosed cancer in women worldwide. It is the second most common cause of cancer deaths in women, after lung cancer. It is reported that one in eight women carry an average lifetime risk of developing breast cancer and the risk of dying from it is 3.4 % (Libson and Lippman, 2014).

In 2012, 1,676,000 cases of breast cancer were diagnosed worldwide and more than 500,000 women died. In Europe, more than 464,000 cases of breast cancer were diagnosed and around 131,000 died (Cancer Research UK, 2014). In UK for the year 2011, approximately 350 men and 50,000 women were diagnosed with breast cancer, an average of 130 women per day. Moreover, breast cancer killed an average of 32 women every day in the UK, adding up to 11,700 women for the year 2011 (Cancer Research UK, 2014).

The UK has the 6th highest breast cancer incidence rate in Europe (Cancer Research UK, 2014). Since 1970, breast cancer incidence rates have increased by 72 %, and by 7 % in the last ten years. Nevertheless, despite increasing incidence rates, survival rates have been effectively improving. Three-quarters of the women diagnosed are now likely to survive for at least 10 years, twice the rate compared to 40 years ago **(Figure 1.1)**. In 1970s, 5 out of 10 women would have survived breast cancer, whereas, today more than 8 out of 10 women are likely to survive (Cancer Research UK, 2014).

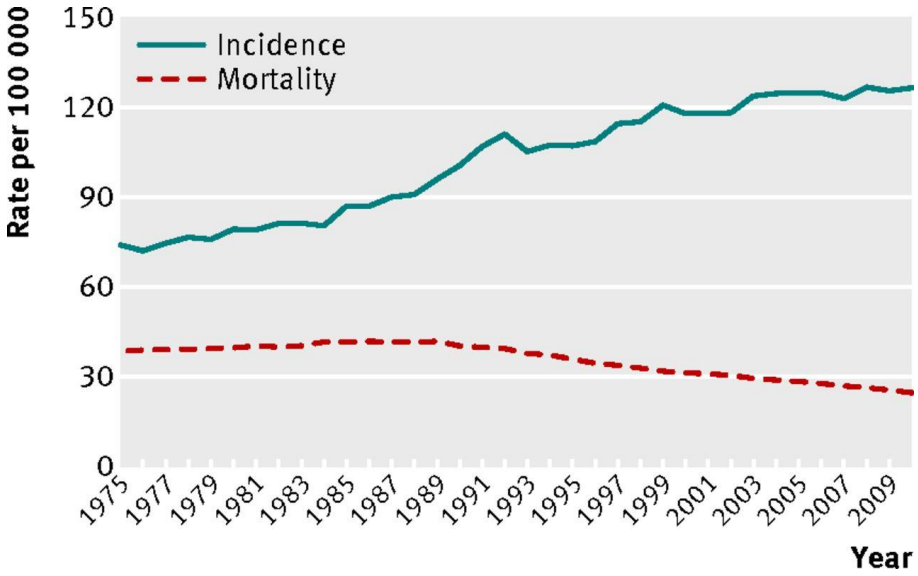


Figure 1.1 Breast cancer incidences and mortality rates for woman in the United Kingdom for last 30 years. The image is adapted from (Yeo et al., 2014).

There are number of risk factors associated with breast cancer development, including age, geographical variation, family history, breast associated diseases, exposure to radiation, lifestyle, oral contraceptives and hormone replacement therapy **(Table 1.1)**. The incidence of breast cancer increases with age. For instance, breast cancer incidence is exceptionally low for ages 35 and below (3 % of all cases), whereas breast cancer accounts for 22 % of women diagnosed below the age of 50. The 50-59 age group has the highest reported incidence of breast cancer, comprising 32 % of all the cases (Murawa et al., 2014, Ferlay et al., 2010).

Moreover, women that have a family history of a mother, sister or daughter being diagnosed with breast cancer carry almost double the risk of developing breast cancer themselves (Cancer Research UK, 2014). It is reported that mutations in the breast cancer 1 early onset (BRCA1) and breast cancer 2 early onset (BRCA2) genes are highly associated with breast cancer (Melchor and Benitez, 2013, Singletary, 2003). These two genes are located on the long arms of chromosomes 17 and 13, respectively and encode tumour suppressor proteins. These proteins have an important role in DNA damage repair and therefore essential for the stability of DNA in the cell (Maxwell and Domchek, 2012). The prevalence of BRCA1 and BRCA2 mutation in

inherited breast cancer is 20 - 25 % and 5 – 10 % in all breast cancer cases (Campeau et al., 2008). Lifestyle risk factors associated with breast cancer development include alcohol consumption, diet, weight and smoking (Cancer Research UK, 2014).

It is clear that breast cancer is a multifactorial disease and is a result of interaction between various risk factors such as genetics, environment, lifestyle, and hormonal factors.

Table 1.1 Risk factors associated with breast cancer:

Increases Risk (Convincing evidence)	**May Increase Risk** (Probable evidence)	**Decreases Risk** (Convincing evidence)	**May Decrease Risk** (Probable evidence)
• Alcohol consumption • Oestrogen-progestogen contraceptives • Oestrogen-progestogen menopausal therapy • Radiation • Obesity	• Hormone replacement therapy • Circadian disruption • Tobacco consumption • Higher weight at birth • Greater abdominal fat • Weight gain • Higher consumption of dietary fat	• Breastfeeding	• Low body fat • Physical activities

1.2 Clinical management of breast cancer

Breast cancer is a complex and heterogeneous disease encompassing diverse subtypes with different clinical, morphological, and biological features. The commonly used phenotypical classification of breast cancer subdivides tumours based on the expression of oestrogen receptor (ER), progesterone receptor (PR), and the human epidermal growth factor receptor 2 (HER2) (Libson and Lippman, 2014, Cadoo et al., 2013).

In recent years, there have been significant advances in molecular profiling of tumours based on gene expression (Cancer Genome Atlas, 2012). These advances have revealed five distinct molecular subtypes of breast cancer based on gene expression: Luminal A (LA), Luminal B (LB), Basal-like (BL), HER2-enriched, and Normal-like (NL) **(Figure1.2)** (Perou et al., 2000, Sorlie et al., 2001). These molecular subtypes reflect all stages of breast cancer from in situ carcinoma to metastatic tumours and display clinical phenotypes based on hormonal receptor status of ER and HER2 (van 't Veer et al., 2002). For example, Luminal A tumours are characterised as less proliferative, with poor chemosensitivity. They are, however, highly sensitive to hormonal therapy. On the other hand, luminal B, Basal-like, and HER2-enriched patients are commonly resistant to hormonal therapy but sensitive to chemotherapy (Cadoo et al., 2013). Despite the advances in molecular profiling and our understanding of underlying genetics of breast cancer, a molecular classification system is not widely used in the clinical practice (Sotiriou and Piccart, 2007).

Currently the clinical decisions made are based on various clinicopathological factors such as histological grade, patients' age, tumour size, lymph node status, and expression of hormonal receptors including ER, PR, and HER2. These factors estimate the prognosis and/or predictive outcome hence guiding for adjuvant therapies (chemotherapy, hormonal therapy, and receptor inhibitors).

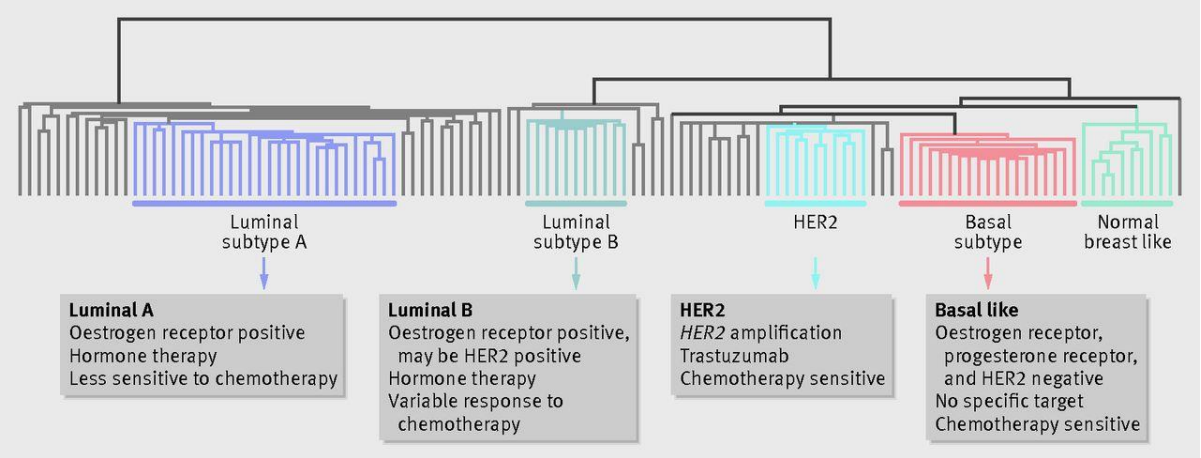


Figure 1.2 Major molecular subtypes of breast cancer. Luminal A, luminal B, *HER2* amplified, basal-like and normal like. Luminal-type breast cancers are ER+ and can be targeted with hormonal therapies. HER2 breast cancers have overexpressed HER2 with amplification of the *HER2* gene and can be targeted with the monoclonal antibody trastuzumab. Basal-like breast cancers commonly do not express oestrogen or progesterone receptors or overexpress *HER2*, a cancer phenotype often referred to as 'triple negative'. The image is adapted from (Yeo et al., 2014).

Hormone receptor status is used for prognostic marker and to predict response to hormonal therapies. The majority of breast cancers, about 80%, are reported to be oestrogen receptor (ER) positive (Nadji et al., 2005, Yeo et al., 2014). ER positive tumours are generally characterised by smaller size, lower grade and negative lymph nodes, as compared to ER negative tumours (Cadoo et al., 2013). Anti-oestrogen therapies use two approaches, either by modulating the growth of cancer cells by antagonising oestrogen receptors (tamoxifen), or via inhibiting the production of oestrogen (aromatase inhibitors). Tamoxifen is commonly used for the treatment of ER breast cancer patients. Five years of adjuvant tamoxifen reduces death by 31 % and relapse by 41 % from breast cancer, and it remains the standard of care for premenopausal women. While, aromatase inhibitors have shown improved clinical outcomes for postmenopausal women (Yeo et al., 2014).

The progesterone receptor (PR) gene is regulated by oestrogen; therefore, its expression is used as an intact marker for activation of ER pathway. Although it is estimated 40 % of ER positive

patients are PR negative(Nadji et al., 2005). As per clinical outcomes, ER positive/PR negative patients are commonly less sensitive to tamoxifen treatment compared to ER positive/PR positive patients(Rakha et al., 2010).

The oncogene HER2 is amplified/overexpressed in approximately 20 % of breast cancers (Yeo et al., 2014). HER2 overexpressed tumours are characterised by an increase in cell proliferation, angiogenesis and invasiveness of tumours and with a decreased rate of apoptosis (Yeo et al., 2014). HER2 positivity is reported in 13–20 % of invasive breast cancers, and commonly linked with younger age, high grade, involvement of lymph node, and non-expression of hormone receptors (Dandachi et al., 2002). In addition, HER2 positive tumours have a worse prognosis than other breast cancer types (Cadoo et al., 2013).

Trastuzumab is a humanised monoclonal antibody against extracellular domain of HER2 receptor. Trastuzumab has shown to improve disease free survival and overall survival of HER2 positive breast cancer (Hudis, 2007). Recently other drugs targeting HER2 such as lapatinib, pertuzumab, and ado-trastuzumab-emtansine have been developed, further improving the outcome of HER2 positive breast cancer (Yeo et al., 2014).

The triple negative breast cancers (TNBCs) are a heterogeneous group of tumours that lack expression of ER, PR, and HER2. They represent about 15 % of breast cancer cases overall, with relatively high incidence among younger females (Yeo et al., 2014). TNBCs are characterised with an aggressive phenotype and poor overall outcome. To date there are not any validated targeted therapies for TNBCs, hence chemotherapy remains the standard clinical practice for care. According to molecular classification TNBCs overlap with the Basal-like tumours, however, these classifications do not always translate to each other, since not all Basal-like tumours have the TNBCs feature and not all TNBCs are molecularly Basal-like subtypes (Cadoo et al., 2013, Valentin et al., 2012).

There are a number of physiological and genetic alterations that may aggravate breast cancer development, progression, and metastasis. These include loss or overexpression of hormonal

receptors, genomic instability, epigenomic aberration, epithelial to mesenchymal transition (EMT), response to stress, regulation of non-coding RNAs and changes in microenvironments including hypoxia. There have been significant attempts to understand and target these alterations and processes to find novel pathways for drug development and improve clinical management of cancer.

1.3 The hypoxic microenvironment

1.3.1 Cellular adaptation to hypoxia

Eukaryotic cells require a constant supply of oxygen to generate energy-rich adenosine triphosphate (ATP) molecules. The average ambient oxygen concentration is 21 % O_2 (150 mm Hg) in air; however, in most mammalian tissues it is around 2 %–9 % O_2 (on average 40 mm Hg) (Bertout et al., 2008). The normal level of oxygen in tissues (physoxia) differs widely among different organs (Hammond et al., 2014). Reduced oxygen delivery leads to alterations in cellular oxygen tension (Wilson and Hay, 2011). Hypoxia is defined as a reduction of oxygen tension below a normal level in tissues (generally when oxygen concentration is ≤ 2 %, whereas severe hypoxia (anoxia) is generally defined as ≤0.01 % O_2) (Favaro et al., 2011).

Cellular oxygen levels fall in a number of physiological conditions including growth, exercise or living at high altitude. These cells adapt to function in these microenvironments by increasing the oxygen supply from the lungs to the other tissues and/or by reducing oxygen consumption within cells (Kaelin and Ratcliffe, 2008). For example, a change from aerobic to anaerobic metabolism occurs during the cellular response to hypoxia microenvironments. However, these hypoxic cells undergo a broad spectrum of adaptations involving a wide variety of biological processes in response to low oxygen tension through coordinated expression of a large number of functional genes involved in processes such as differentiation, proliferation, apoptosis and

embryonic development. In addition to physiological roles, hypoxia is associated with many common pathological conditions such as wound healing, stroke, ischemia, inflammation and cancer and hypoxic responses play an important role in the pathogenesis of these diseases (Harris, 2002, Kaelin and Ratcliffe, 2008).

1.3.2 Hypoxia and Hypoxia-inducible factor (HIF)

The majority of hypoxic transcriptional responses are mediated by a key transcription factor known as Hypoxia-Inducible Factor (HIF). Hypoxia regulates both HIF protein stability and its transactivating activity such that in low oxygen levels HIF is activated and induces expression of its target genes. HIF activation is critical for the initiation of many HIF-dependent signalling pathways that are crucial for cellular adaptation and survival in hypoxia involving angiogenesis, metabolism regulation, erythropoiesis, autophagy, metastasis, cell survival and proliferation (Huang, 2013, Ratcliffe, 2013)

HIF is a heterodimeric complex consisting of α and a β subunits that bind hypoxia response elements (HRE) of target genes in hypoxia (Wang et al., 1995, Webb et al., 2009). There are three isoforms of HIF; HIF-1, HIF-2 and HIF-3. All contain HIF-1β and one of the three different HIF-α subunits HIF-1α, HIF-2α or HIF-3α, respectively. All subunits are members of the basic helix-loop-helix (bHLH), Per-ARNT-Sim (PAS) protein family. The N terminal of both HIF-1α/β contains the bHLH DNA-binding domain followed by a PAS domain, which is involved in heterodimerisation (Kaelin and Ratcliffe, 2008). HIF-1α is virtually ubiquitous and expressed in a wide range of tissues, whereas expression of HIF-2α is limited to specific cell types; for example, endothelial cells, cardiomyocytes, mammary cell and hepatocytes (Holmquist-Mengelbier et al., 2006, Hu et al., 2003).

The role and function of HIF-3α is not well understood and is complicated by the existence of multiple splice variants. Studies have described contrasting roles for HIF-3α, which may

synergise with other subunits assisting HIF-1α induction in hypoxic alveolar epithelial cells (Zhang et al., 2014, Li et al., 2006), or may act as a negative regulator of hypoxic response by altering formation of functional heterodimers with other HIF isoforms(Makino et al., 2001).

The molecular mechanisms by which oxygen regulates HIF stability and activity are well characterised (Semenza, 2014, Ratcliffe, 2013). In the presence of oxygen (normoxia), HIF is modified by post-translational hydroxylation of specific proline residues by a family of enzymes termed HIF prolyl 4-hydroxylases. The hydroxylation is catalysed by three prolyl hydroxylase domain-containing (PHD) proteins of the iron (II) and 2-oxoglutarate (2OG)-dependent dioxygenase family known as PHD1, PHD2, and PHD3 (Epstein et al., 2001, Bruick and McKnight, 2001). This hydroxylation of HIF occurs at two prolyl residues (Proline 402 and 564 in HIF-1a; Proline 405 and 531 in HIF-2a), which mediate interactions with the von Hippel-Lindau (pVHL) E3 ubiquitin ligase complex that targets HIF-α for proteasomal degradation. The interaction between hydroxylated residues and VHL E3 triggers rapid proteolysis of HIF-α in normoxia. Thus, in the presence of oxygen HIF-α is both inactivated and degraded and the expression of HIF target genes is therefore no longer triggered **(Figure 1.3)** (Ratcliffe, 2013)**.**

In hypoxia, active HIF-α subunits accumulate and dimerize with HIF-1β. The HIF heterodimer then translocates to the nucleus, where it binds to its consensus-binding motif at the hypoxia response element (HRE) (core-binding motif is RCGTG (R = A or G), and subsequently recruits transcriptional cofactors such as P300 and CBP (Huang, 2013, Ratcliffe, 2013). This results in the transcriptional activation of hundreds of HIF target genes that are involved in angiogenesis, metabolism, PH regulation, erythropoiesis, cell survival and proliferation **(Figure 1.3)** (Bertout et al., 2008, Semenza, 2014)**.**

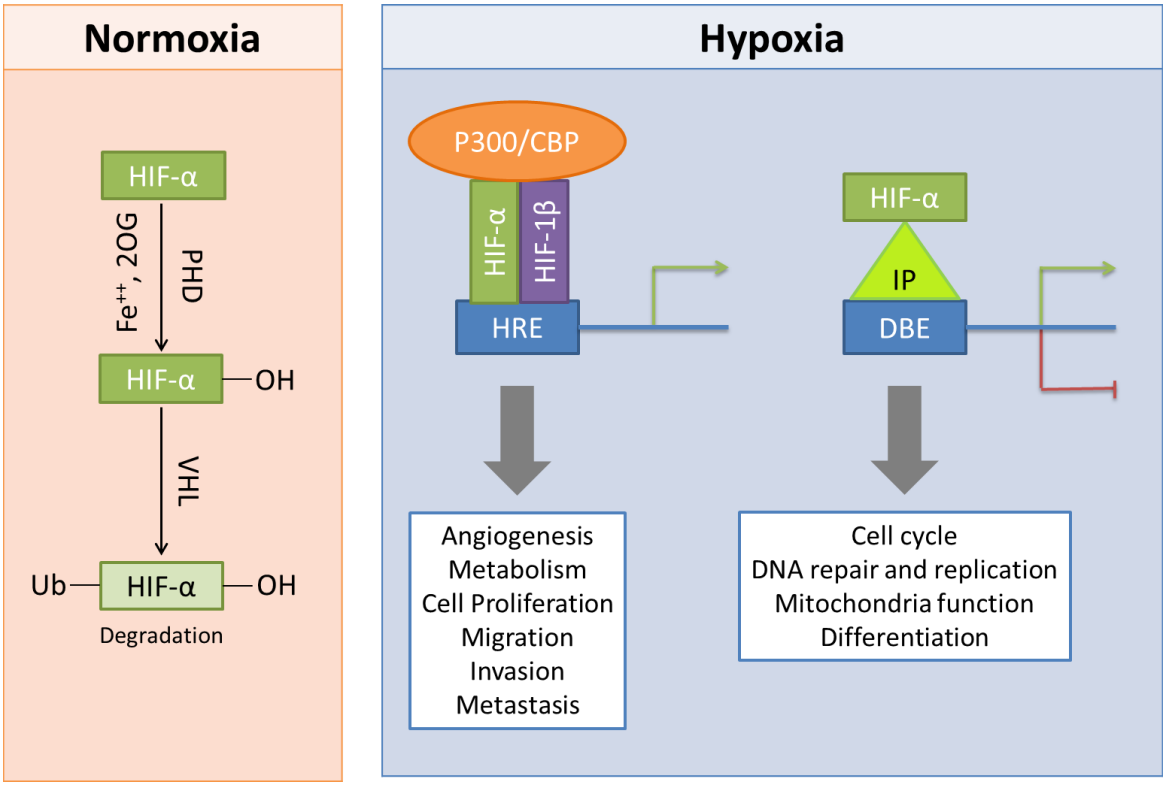


Figure 1.3 Mechanism of action and regulation of hypoxia inducible factor. In present of oxygen (normoxia), HIF-α undergo proly hydroxylation in the presence of ferrous ion (FE^{++}) and 2-oxoglutarate (2-OG). The hydroxylated HIF-α is recognised by ubiquitin ligase VHL for polyubiquitination (-Ub) and proteasome degradation. In low levels of oxygen (hypoxia), HIF-α is stabilised and dimerises with HIF-1β and binds with hypoxia responsive element (HRE), which consequently recruits p300/CBP in target genes. There are other mechanisms that modulate HIF-α activity in hypoxia; for example, HIF-α binds with a number of interacting proteins (IP) that occupy DNA binding elements (DBE) leading to the up/downregulation of a number of genes in hypoxia (Tiburcio et al., 2014).

There are several other mechanisms for modulating HIF activity in cells, including factor inhibiting HIF-1 (FIH1) regulation, oncogenes, tumour suppressor genes, acetylation, reactive oxygen species (ROS), nitric oxide, SUMOylation and non-coding RNAs. These modulatory mechanisms add to the complexity of HIF regulation in hypoxia (Majmundar et al., 2010).For instance, similar to PHDs the factor inhibiting HIF-1 (FIH1) is an oxygen-dependent asparaginyl hydroxylase, belonging to a family of 2-oxoglutarate- and iron(II)-dependent dioxygenases. FIH1 modifies Asn803 and block the interaction of the C-terminal transactivation domain (C-

TAD) of HIF-α with the transcriptional co-activator CBP/p300, therefore, controlling HIF activity in cells (Lando et al., 2002, Mahon et al., 2001).

Mitochondria also participate in oxygen sensing and regulating PHDs. It has been reported that under reduced levels of oxygen, mitochondria increases the production of reactive oxygen species (ROS), which then inhibit PHD activity and increase the HIF stability (Kaelin, 2005). Hypoxia induced oxidants are present in the inner-membrane space of mitochondria and in the cytosol. These ROS could potentially regulate PHD activity and thereby HIF in hypoxia (Waypa et al., 2010).

Another mechanism of modulating HIF expression is by Sirtuins. These are a stress-responsive family of nicotinamide adenine dinucleotide (NAD+)-dependent histone deacetylases. Sirtuin expression has effects on transcription, metabolism and DNA repair (Haigis and Sinclair, 2010). In hypoxia, Sirt1 binds with HIF-2α and forms a complex that deacetylates lysine residues in the HIF-2α protein, leading to its stabilisation. In addition, Sirt1 binds the *Epo* promoter to promote HIF-2α transcriptional activity in hypoxia (Dioum et al., 2009). In addition, activation of several oncogenic signalling pathways such as phosphotidylinositol 3-kinase (PI3K)/AKT/mammalian target of rapamycin (mTOR) and mitogen-activated protein kinase (MAPK) pathways stimulate HIF-1α and HIF-2α synthesis (Semenza, 2010, Harris, 2002).

HRE sequences where HIF binds may be located within either promoter or enhancer regions of target genes and may lie close to the TSS, within the gene body (in introns or exons), within the 5'-UTR or 3'-UTR or in intergenic regions in great distances from their transcriptional target. However, epigenetic changes such as methylation of the cytosine residue within the HRE or structural changes including genetic variants within the HRE may influence HIF binding and transcriptional responses (Zagorska and Dulak, 2004, Schodel et al., 2013).

Despite the presence of HIF-1α and HIF-2α in mammalian cells and the fact that they have similar biochemical properties, regulatory mechanisms, and number of target genes, these isoforms have distinct functional roles in mediating the hypoxic transcriptional response (Keith et al., 2012, Kaelin and Ratcliffe, 2008). For instance, HIF-1α mediates rapid stabilisation and activation of target genes such as glycolytic enzymes, whereas HIF-2α is active in moderate levels of hypoxia and its induction is more prolonged (Keith et al., 2012). It is also noted in some instances that HIF-1α and HIF-2α utilise opposing mechanisms to regulate cell proliferation, differentiation, DNA replication and mitochondrial function (Huang, 2008, Lendahl et al., 2009). These non-canonical mechanisms of HIF-1α and HIF-2α regulation require crosstalk with other interacting proteins (IPs) that occupy DNA-binding elements (DBE), thereby regulating gene expression **(Figure 1.3)**. For example, HIF-1α and HIF-2α control cell-cycle genes in an opposite manner; HIF-1α supresses cell-cycle progression by inhibiting c-Myc activity, while HIF-2α has a reverse effect (Koshiji et al., 2004, Gordan et al., 2007). Similarly, some DNA repair genes are supressed by HIF-1α repression but induced by HIF-2α (Gordan et al., 2007, To et al., 2006, Gordan et al., 2008). HIF-1α is involved in blocking cell differentiation via the Notch signalling pathway, but at the same time is involved in regulating embryonic and adult cell proliferation by the β-catenin signalling pathway (Gustafsson et al., 2005, Mazumdar et al., 2010, Kaidi et al., 2007).

The role of HIF-1α in metabolic adaptation by hypoxia is well established. HIF-1α increases the transcription of many genes involved in glucose transport and glycolysis, such as genes encoding aldolase A (ALDOA), phosphoglycerate kinase 1 (PGK1) and lactate dehydrogenase A (LDHA) (Majmundar et al., 2010). In normoxic cells, glucose is metabolised to pyruvate, which is then converted by pyruvate dehydrogenase to acetyl-CoA for oxidative phosphorylation in the tricarboxylic acid (TCA) cycle. Under low level of oxygen, cells decrease oxidative phosphorylation in the mitochondria to adopt anaerobic glycolysis by converting pyruvate to lactate (Yang et al., 2014a). HIF-1α plays a key role in this process. It decreases oxygen consumption in mitochondria by upregulating both pyruvate dehydrogenase kinases 1 and 3

(PDK1 and PDK3) (Papandreou et al., 2006, Lu et al., 2008), which then leads to suppression of pyruvate dehydrogenase and consequently inhibits pyruvate conversion and flux of acetyl-CoA into the TCA cycle. The shift into aerobic glycolysis in hypoxia is not only for bioenergetics but also more importantly for biosynthesis of macromolecules that are required for cell adaptation and proliferation (Vander Heiden et al., 2009, Cairns et al., 2011).

Laboratory experiments of hypoxia in mammalian cell cultures have been vital tools to investigate hypoxia biology and have significantly improved our understanding about hypoxia/ HIF response in cells, both at cellular and molecular levels. Two main approaches have been used to induce hypoxia in cell cultures models: use of chemical agents that mimic hypoxia in normoxic conditions, such as Cobalt chloride ($CoCl_2$), Deferoxamine mesylate (DFO) and 2-oxoglutarate analogue dimethyloxalylglycine (DMOG); the second approach uses hypoxic chambers that allow the growth of cells in conditions that accurately maintain and control the temperature, oxygen and carbon dioxide. This method includes normal cell culture incubators with reduced oxygen levels or hypoxia work stations (connected to a gas control system in a chamber with a glove box)(Wu and Yotnda, 2011, Richter et al., 1972).

The first approach has the advantages of being inexpensive and quick. It allows opening the culture plate and flask many times for experimental manipulation in normoxia without affecting the "hypoxic condition". These agents induce and or stabilise HIF in normoxic conditions. For instance, $CoCl_2$, DFO, DMOG stabilizes HIF-1α by inhibiting prolyl hydroxylase enzymes (Piret et al., 2002, Theriault et al., 2012, Jaakkola et al., 2001). The disadvantage of the approach is that cells are grown independently of the oxygen level and potentially trigger expression of other genes that are drug-specific.

The second approach has the advantage of not using drugs that mimic hypoxia independently of the oxygen tension. The normal hypoxic incubator allows cells to grow in controlled oxygen tension. However, not all experiments are feasible, as oxygen re-enters the incubator at each opening and thus alters the hypoxic level. Hypoxia workstations allow the experimentation to

take place in a continuous hypoxic environment, including changing media and manipulation of cells (Atkuri et al., 2007, Wu and Yotnda, 2011).

The use of different approaches to induce hypoxia and HIF *in vitro* has profound effects on gene expression. Ortiz-Barahona and colleagues performed "meta-analysis" of 19 gene expression profiling datasets that were available in the GEO database that were derived from different cells exposed to hypoxia or hypoxia mimicking agents (Ortiz-Barahona et al., 2010). They found that the number of hypoxia-regulated genes that were common to all 19 datasets was very small, with about 24 core hypoxia regulated genes identified in all cell types studied (Ortiz-Barahona et al., 2010). These differences may result from use of different approaches to induce hypoxia, incubation period and use of different cell-types that activate or repress different genes in response to hypoxia.

It is also to be noted, in addition to HIF, there are a number of other transcription factors such as AP1, NFκB, and p53 that have all been reported to be regulated in hypoxia(Renton et al., 2003, Rupec and Baeuerle, 1995, Kunz and Ibrahim, 2003).

1.4 Hypoxia, HIF and cancer

Most solid tumours have regions of oxygen deficiency. This is due to a number of factors inducing uncontrolled and rapid growth of cancer cells, highly defective architecture of tumour blood vessels and interrupted blood rheology leading to alteration in tumour blood flow (Hammond et al., 2014).

Tumour cells are usually characterised by their uncontrolled proliferation and aberrant vessel formation. Moreover, their vasculature is often chaotic, spatially disorganised and has poor blood flow, leading to inefficient oxygen delivery and significantly lower oxygen tension in tumours compared to healthy tissues (Bertout et al., 2008). In hypoxic tumours, oxygen tension

averages 10mm Hg that corresponds to 1.5 % oxygen, but can be variable depending on tissues (Greijer et al., 2005). Generally, hypoxia is a source of stress for both cancer and normal cells; however cancer cells may adapt to these changes by virtue of their genetic and epigenetic instability, allowing them to survive and proliferate in hypoxic conditions. These hypoxia adaptations can contribute to the malignant phenotype and aggressive tumour behaviour (Harris, 2002, Kaelin and Ratcliffe, 2008). It is believed that regional tumour hypoxia occurs and develops in early stages of carcinogenesis, even before metastases, and is often detected in non-invasive tumours (Harris, 2002).

Several studies have reported that hypoxic tumours are more resistant to anticancer treatments, including chemotherapy, radiotherapy and surgery (Brown and Giaccia, 1998, Nordsmark et al., 2005). The resistance of hypoxic tumour cell to chemotherapy are the result of many factors, including distance from functional vessels, low proliferation rates in hypoxic regions of tumours and loss of apoptosis (Hammond et al., 2014). Hypoxia also increases cellular radioresistance by triggering changes in both the DNA repair pathway and cell death-signalling pathways. There is strong evidence that tumour hypoxia is associated with poor clinical outcome, even after removing the tumour surgically. This potentially may be explained by the role of hypoxia in inducing invasion and metastasis (Brown and Wilson, 2004).

Among the early studies to show tumour hypoxia, Thomlison and Gray observed that human tumour cells undergo greater necrosis when further away from blood vessels and that necrotic areas lack capillaries in histological sections (Thomlinson and Gray, 1955). They postulated that the rapid growth rate of tumours caused them to outgrow their blood supply and therefore limit oxygen diffusion; this is defined as ‘chronic hypoxia’. Conversely, aberrant blood vessels may constrict and/or occlude leading to ‘acute hypoxia’. In addition to low oxygen, increased levels of free radicals and activation of a number of stress related genes may also be seen and may further contribute to tissue damage (Brown and Giaccia, 1998, Wilson and Hay, 2011).

Peter Vaupel and colleagues, who measured tumour oxygen levels using oxygen electrodes, have showed further evidence for hypoxia in tumours. They found that low oxygen tension in tumours was associated with increased metastasis and poor survival in patients suffering from squamous tumours of the head and neck, cervical or breast cancers (Hockel et al., 1999, Hockel and Vaupel, 2001). Several other studies have also reported that tumour hypoxia is strongly associated with poorer prognosis in a large number of cancers (Jubb et al., 2010, Shaida et al., 2011, Fyles et al., 2002).

High levels of HIF have been reported in many solid tumours (Semenza, 2009). Moreover, both hypoxia and increased levels of HIF are commonly linked to bad prognosis in cancer patients, suggesting that HIF increases aggressiveness of cancer cells *in vivo* (Semenza, 2009). For example, immunohistochemical analyses showed overexpressed HIF-1a in many human cancers including breast, cervix, brain, oropharynx, ovary, prostate, gastric and lung (Semenza, 2003, Kizaka-Kondoh et al., 2009, Ben Lassoued et al., 2013).

HIF regulate many key aspects of cancer biology inducing cell proliferation and survival, angiogenesis, stem cell maintenance, metabolism, epithelial-mesenchymal transition, invasion and metastasis (Semenza, 2012b). In laboratory experiments, HIF-α found to enhance cell proliferation in different cancer cell types. In mouse cancer models, depletion of HIF-1α, or HIF-1β significantly impairs tumour growth *in vivo* (Semenza, 2012b).

HIF induces cell proliferation and survival through regulation of many growth/survival factors in many cancer for example, transforming growth factor-A (TGFA), insulin-like growth factor-2 (IGF2), vascular endothelial growth factor (VEGF), erythropoietin (EPO), cyclin-dependent kinase inhibitor 1A (CDKN1A) and B-cell lymphoma 2 (Bcl2) (Grimshaw, 2007, Feldser et al., 1999, Szenajch et al., 2010).

HIF also regulates epithelial–mesenchymal transition by the transcriptional activation of genes (ID2, SNAI1, SNAI2, TCF3, ZEB1, ZEB2) that have essential functions in regulating the

expression of E-cadherin, the flexibility of cytoskeleton and cell-cell adhesion (Esteban et al., 2006, Krishnamachary et al., 2006). HIF-1α promotes expression of genes (TGFA, VIM) that regulate the mesenchymal phenotype(Gunaratnam et al., 2003).

During metabolic reprogramming, many genes are regulated by HIF in hypoxia, such as Glucose transporter1 (GLUT1), GLUT3, Lactate dehydrogenase A (LDHA) and Pyruvate dehydrogenase kinase 1 (PDK1) and Carbonic anhydrase 9 (CA9) (Semenza, 2012b).

During the invasion and metastasis, HIF-1α promotes expression of genes encoding: proteases that degrade (CTSC, MMP2, MMP9, MMP14, PLAUR) or remodel (LOX, LOXL2, LOXL4) the tissue (Erler and Giaccia, 2006, Wong et al., 2011), motility factors (AMF, MET) and it also upregulates (ANGPTL4) proteins that enhance extravasation of cancer cells into metastatic sites (Zhang et al., 2012b). These and many other genes are regulated by HIF in hypoxia and play crucial roles in cancer progression.

Adding to hypoxia, there are a number of cancer-associated mutations that trigger and enhance HIFa accumulation. For instance, mutations that inactive PTEN or TSC2, or activate PI3Ka or AKT1 lead to activation of mTOR, thereby increasing HIF-1α transcription and translation (Brugarolas and Kaelin, 2004). Activating mutations in Ras promotes production of ROS resulting in HIF1a hydroxylation inhibition (Gerald et al., 2004). On the other hand, inactivating mutations in metabolic enzymes such as succinate dehydrogenase or fumarate hydratase inhibits HIF hydroxylation (Kaelin and Ratcliffe, 2008, Baldewijns et al., 2010). In addition, mutations in VHL alter the ability of pVHL to polyubiquitylate HIF, increasing HIF stability. Mutations in VHL are associated with the development of hemangioblastoma and clear cell renal carcinoma (Kaelin and Ratcliffe, 2008, Baldewijns et al., 2010). The role of HIF-α in tumourigenesis has been well known through studies of pVHL-defective tumours. For example, the risk of renal carcinoma development with germline VHL mutations correlates with the level to which these mutations activate HIF-α. This is also true for VHL mutations identified in sporadic clear cell renal carcinomas (Baldewijns et al., 2010, Kaelin, 2007). In preclinical

models, VHL$^{-/-}$ clear cell renal carcinomas overexpress both HIF-1α and/or HIF-2α. In addition, HIF2-α knockdown in VHL$^{-/-}$ renal carcinoma cells is sufficient to inhibit tumour growth (Baldewijns et al., 2010, Kaelin, 2007). These findings suggest that drugs that block HIF or downstream HIF targets will be beneficial for treating VHL deficient cancers, for example, VHL$^{-/-}$ clear cell carcinomas. *In vitro*, molecules such as ascorbic acid and membrane-permeable 2-oxoglutarate antagonists have shown to supress HIF-α in tumour cells (Knowles et al., 2003, Page et al., 2008).

Hypoxia is one of the key physiological changes differentiating cancer and normal cells and therefore constitutes a therapeutic target. There are several drugs that supress HIF-α either indirectly, such as mTOR inhibitors, HSP90 inhibitors, and HDAC inhibitors (Kaelin, 2007) or directly such as polyamides that block HIF binding at the VEGF promoter (Olenyuk et al., 2004, Viger and Dervan, 2006). A list of anticancer agent targeting HIF-1α activity is presented in **Table 1.2**

There is increasing evidence that hypoxia supresses many key proteins involved in DNA repair pathways (Bristow and Hill, 2008). Ataxia telangiectasia and rad3-related (ATR)-mediated signalling and ataxia telangiectasia mutated (ATM)-mediated signalling pathways are activated in hypoxia and have important functions during hypoxia-induced replication stress and re-oxygenation (Bencokova et al., 2009, Olcina et al., 2013). Both ATR and ATM are considered as potential therapeutic targets in hypoxia. VE-821 and VE-822 are specific inhibitors of cellular ATR activity that have recently shown to be sensitive to hypoxia in a range of cancer cells (Reaper et al., 2011, Fokas et al., 2012). Notably, critical proteins in homologous recombination are repressed in hypoxia, such as Rad51 and BRCA1. These could potentially be targeted by poly ADP ribose polymerase (PARP) inhibitors. These inhibitors are highly sensitive to BRCA1 and BRCA2 deficient cancer cells, since hypoxia suppresses DNA repair and PARP inhibitors have shown increased sensitivity to hypoxic cancer cells (Chan et al., 2010).

Many HIF-target genes are translated into proteins that play key roles in cancer development and progression, including VEGF, TGFb, SDF-1, CXCR4 and MMP1 (Kaelin, 2007, Kaelin and Ratcliffe, 2008). These proteins have been critical for drug development and for clinical use. For example, inhibitors for VEGF or its receptor, known as kinase insert domain receptor (KDR) have been use in clinic for advanced kidney cancer patients. Two such agents are sorafenib and sunitinib (Kaelin, 2007).

Angiogenesis is a hallmark of cancer and is a compelling therapeutic target for hypoxic tumours. Induction of HIF-1α promotes synthesis of the angiogenic factor VEGF, causing vascular development that is required for tumour growth, however oxygen tension may suffer due vessel malformation. In this case, VEGF inhibition leads to vessel normalisation and increased oxygen levels, permitting access of chemotherapeutic drugs and improving sensitivity to radiation (Ward et al., 2013). Bevacizumab is a monoclonal antibody targeting VEGF. It increases the sensitivity to chemotherapy but does not increase overall survival (Ward et al., 2013, Rakha and Chan, 2011). On the other hand, it is reported that long-term Bevacizumab treatment induces tumour hypoxia, which may eventually lead to resistance. There have been several attempts to use combined therapies with VEGF inhibitors, for example combined blockade of CA9 with Bevacizumab to reduce tumour growth (McIntyre et al., 2012) and the use of anti-angiogenic agents with Notch inhibitors to reduce breast cancer stem cells (Ward et al., 2013, Grudzien et al., 2010).

There are number of novel compounds designed for targeting HIF-1α. Bortezomib is a proteasome inhibitor approved for hematological malignancies that has been found to downregulate CA9 in a metastatic colorectal cancer phase II trial (Mackay et al., 2005). PX-478 is novel compound that blocks HIF-1α transcription and HIF-1α protein expression in a p53-independent and pVHL-independent manner (Koh et al., 2008). In addition, YC-1, is a synthetic compound that blocks HIF through factor inhibiting HIF (FIH)-dependent inactivation of the carboxy-terminal transactivation domain (CAD) of HIF-1α (Li et al., 2008).

Table 1.2 List of anticancer agents that modulate HIF-1α activity:

Targets	Agents
HSP90 inhibitor	Geldanamycin, 17-AAG (Geldanamycin analogue), Radiciol, KF58333 (Radicicol analogue)
Topoisomerase inhibitor	Topotecan, GL331, Anthracycline
Microtubule modifier	Taxane (Paclitaxel, Docetaxel), Vinca alkaloid (Vincristine, Vinoblastine), 2-methoxyoestradiol (2ME2), Epothilone B, Colchicine
sGC stimulator	YC-1
Trx-1 inhibitor	Pleurotin, PX-12/1-methylpropyl 2-imidazolyl-disulphide
Histone deacetylase inhibitor	FK228
P300 CH1 inhibitor	Chetomin
Proteasome inhibitor	Bortezomib
PIK3 inhibitor	Wortmannin, LY294002
mTOR inhibitor	Rapamycin, CCI-779, rad-001
MEK inhibitor	PD98059, BAY43-9006 (Sorafenib)
ErbB2 receptor tyrosine kinase inhibitor	Trastuzumab (Herceptin)
Tyrosine kinase inhibitor	Imatinib (Glivec)
EGFR tyrosine kinase inhibitor	ZD-1839 (Iressa), Erlotinib (Tarceva)
COX2 inhibitor	Celecoxib
Tyrosine kinase inhibitor	Genistein

Abbreviations*:* HSP90, heat-shock protein 90; sGC, soluble guanylate cyclase; Trx, thioredoxin-1; cGMP, cycline guanosine monophosphate; PIK3, phosphatidylinositol-3-kinase; mTOR, mammalian target of rapamycin; MEK, MAP/ERK Kinase; ErbB2, epidermal growth factor receptor 2; EGFR, epidermal growth factor receptor; COX2, cyclooxigenase-2. The table is redesigned and updated from (Milani and Harris, 2008).

A number of methods have been developed to detect hypoxia in solid cancer (Vaupel and Mayer, 2007). It is crucial to reliably detect oxygen levels in tumours for prognosis and for evaluating resistance to specific treatments. There are direct and indirect methods for measuring tumour hypoxia. The Eppendorf polarographic oxygen needle electrode is the most common and direct method to measure oxygen levels, however it has many limitations, including its invasiveness, non representativity of the whole tumour and the oxygen consumption by the electrodes can result in false positive quantification (Favaro et al., 2011, Peters et al., 2010).

In clinical practise, the measurement of hypoxia utilises endogenous and exogenous markers. For example, immunohistochemistry is commonly used for patient biopsies and detects both endogenous and exogenous markers of hypoxia. Endogenous markers such as HIF-1α and its target genes, including GLUT1, CA9 and VEGF are widely used. Limitation of these endogenous

markers is their association with by non-hypoxia-related factors (for example, pH or the concentrations of glucose and glutamine)(Favaro et al., 2011).

Exogenous markers of hypoxia utilise nitroimidazole compounds, such as pimonidazole, 2-(2-nitro-1H-imidazol-1-yl)-N-(2,2,3,3,3-pentafluoropropyl)-acetamide (EF5)(Varia et al., 1998). These markers are administered to patients and interact with proteins in hypoxic conditions to form stable adducts which later can be detected by using specific antibodies on tumours (Airley et al., 2003). The major limitations of these endogenous and exogenous markers include their invasiveness (they require tumour biopsies), the fact that biopsies are not necessarily representative of the whole tumour, and the inability to perform multiple evaluations for follow up (Favaro et al., 2011, Tatum et al., 2006).

Recently, new techniques have been developed for imaging hypoxic tumours in clinic (Shao et al., 2013, Hammond et al., 2014), for example, the use of nitroimidazole derivatives including 18F-fluoromisonidazole (18F-MISO) in combination with positron emission tomography (PET)(Horsman et al., 2012). While the 18F-MISO-PET technique has several advantages (non-invasive and permits serial imaging of hypoxia), the accumulation of 18F-MISO in hypoxic tumours is moderately low. The Blood oxygen level-dependent magnetic resonance imaging (BOLD MRI) has also been used for tumour hypoxia imaging (Hammond et al., 2014). A limitation of this method is the fact that it does not quantify oxygen directly and could be affected by factors such as blood flow and tumour perfusion (Favaro et al., 2011).

It is becoming clear in order to enhance our ability to target hypoxic tumour cells and to assess their response to therapies, more comprehensive approaches for hypoxic tumour detection are required.

1.5 Hypoxia and breast cancer

Although tumour cells can tolerate hypoxic stress and develop multiple ways of adapting and surviving limited oxygen levels, severe and prolonged hypoxic conditions can eventually lead to necrosis and then to cell death. Indeed, necrotic regions are frequently found in many cancer types (Zhong et al., 1999, Tomes et al., 2003). Hypoxic and necrotic regions are found in breast cancer (Lundgren et al., 2007).

Several studies have described hypoxic regions occurring in all stages of breast cancer, from *in situ carcinomas* to both ductal and lobular high-grade invasive carcinomas (Hohenberger et al., 1998, Cornfield et al., 2004). Generally, breast cancer patients with tumour characteristics such as ER negative , PR negative and node-positive status have higher proportions of anoxic and hypoxic areas (Vaupel et al., 2002, Vaupel et al., 2005). Hypoxia is also observed in tumour metastasis. Van den Eynden et al. studied hypoxia markers in 60 breast cancer patients and found that the presence of hypoxia in the primary tumours correlates with hypoxia markers in lymph node metastases (Van den Eynden et al., 2005). These data indicate that the hypoxia signature is maintained after tumour cells migrate from their primary organ and move to a new environment.

In addition, many studies have shown that HIF isoforms (HIF-1α or HIF-2α) are adverse prognostic factors in breast cancer patients with or without lymph node status (Dales et al., 2005, Bos et al., 2004, Bachtiary et al., 2003, Helczynska et al., 2008). Using HIF-1α as a marker for hypoxia, it is estimated around 25 – 40 % of all invasive breast cancer samples are hypoxic (Kronblad et al., 2006, Gruber et al., 2004, Dales et al., 2005, Bos et al., 2003, Wang et al., 2014). HIF-1α positivity is commonly associated with increased tumour size, high grade, high proliferation, lack of lymph-node metastases and hormone-receptor negativity (Lundgren et al., 2007, Wang et al., 2014). Moreover, high levels of HIF-1α have been shown to be a predictor for breast cancer therapy responses, such as chemotherapy (Generali et al., 2006). On the other hand, high levels of HIF-2α have been found in infiltrating ductal carcinomas and were

associated with highly metastatic tumours (Giatromanolaki et al., 2006). In other studies, HIF-2α expression in patients with invasive breast cancer correlates with histological grade and Ki67 expression in tumours (Xiang et al., 2012) and is associated with poor outcome and relapse (Helczynska et al., 2008). Furthermore, interventional studies using xenograft models of breast cancer cells, in which HIF-1α has been deleted, showed a significant reduction in tumour progression and metastases in these tumours (Hiraga et al., 2007, Liao et al., 2007).

In hypoxic tumours, HIF drives the expression of many target genes some of which promote tumour growth and progression (Semenza, 2012b). Similarly, HIF regulated non-coding RNAs can also predict hypoxic responses in breast cancer, for example, microRNA-210 is a HIF-1α-regulated small non-coding RNA that is highly induced in hypoxia and is an indicator of poor prognosis in breast cancer (Volinia et al., 2012, Hong et al., 2012).

In breast cancer, evidence suggests that high HIF-1α expression and its targets have been associated with worse prognosis (Starmans et al., 2012, Favaro et al., 2011, Wang et al., 2014). One important example is carbonic anhydrase (CA9), which is a HIF-1α target gene and present on the plasma membrane and functions by hydrating CO_2 to form H^+ and HCO_3^- extracellularly (Wykoff et al., 2000). The upregulation of CA9 has been associated with aggressive features and linked to the reduction of relapse-free survival in breast cancer patients with grade I and II tumours (Beketic-Oreskovic et al., 2011, Brennan et al., 2006) and is associated with poor clinical outcome in invasive breast carcinomas (Chia et al., 2001, Hussain et al., 2007).

It is also suggested that high expression of CA9 is a predictive marker for breast cancer chemotherapy responses and adjuvant treatments (Span et al., 2007, Betof et al., 2012). In addition, high levels of GLUT1 and VEGF (HIF-1α target genes) have also been linked with increased risk of recurrence, higher tumour grade and poor prognosis (Linderholm et al., 2000, Kang et al., 2002), Chen et al. reported overexpression of HIF-1α and MET were associated together in breast cancer, indicating that HIF-1α promotes aggressive breast cancer by inducing MET (Chen et al., 2007). Another study showed a strong association between lymphatic

microvessel density and expression of HIF-1α and VEGFC in breast cancer. This study proposes that HIF-1α could have an effect on tumour-associated lymphangiogenesis, which is a critical requirement for breast cancer metastases (Schoppmann et al., 2006).

Activation of HIF-1α in breast cancer has also been associated with resistance to anticancer therapies (Unruh et al., 2003). In a clinical study conducted on 187 breast cancer patients treated with either neoadjuvant epirubicin alone or with combined epirubicin and tamoxifen it was reported that HIF-1α and CA9 were associated with treatment resistance (Generali et al., 2006). Another study included 114 breast cancer patients who received aromatase inhibitor. It showed that the level of HIF-1α was an independent factor that was significantly linked with treatment resistance (Generali et al., 2009). High levels of serum CA9 in metastatic breast cancer patients treated with trastuzumab have shown to be associated with reduced progression-free survival (Favaro et al., 2011). These studies are in agreement with previous evidence that tumours with low CA9 expression are more sensitive to adjuvant endocrine or chemotherapy treatment (Favaro et al., 2011, Span et al., 2003). In addition, high levels of *VEGF* have been linked with resistance to both hormonal and chemotherapies for breast cancer (Foekens et al., 2001). These reports highlight the significance of assessing tumour hypoxia, not only to allow for selection of patients who would more likely profit most from treatments, but also to avoid the use of treatments in cases where they might be unfavourable.

There have been attempts to utilise hypoxia gene expression signatures (expression of multiple genes responsive to hypoxia) as markers for detecting hypoxic tumours. These gene signatures can potentially improve our understanding of hypoxic responses to a specific therapy and mechanisms of resistance in terms of gene regulation (Favaro et al., 2011). Buffa et al. described a hypoxia signature based on a meta-analysis derived from head and neck squamous cell carcinomas (HNSCCs) and breast cancers. They selected genes that were consistently co-expressed with the hypoxia from *in vitro* and *in vivo* experiments of different cancers that might reflect tumour response to hypoxia and suitable for clinical use. They reported that these

metagenes are highly clinically relevant to hypoxia in cancers, in particular breast cancer (Buffa et al., 2010).

Since hypoxia is central to the pathology of breast cancer, it is important to understand and characterise in-depth mechanisms of tumour hypoxia in breast cancer that may lead to determine patient prognosis, optimise treatment and provide better understanding on how to target hypoxia and to discover novel therapeutic targets.

1.6 The cancer genome

Comprehensive analysis of cancer genomes has revealed widespread genomic alterations the complexity of which has been highlighted through recent advances in genome-wide analysis. These alterations include point mutations, altered methylation, chromosomal rearrangements, aberrant genomic structure and changes in gene expression (Cowin et al., 2010). Cancers may accumulate a number of such alterations ultimately leading to changes in cancer related genes such oncogenes (which drive cancer) and/or tumour suppressor genes (which suppress cancer). Profiling the cancer genome has had a profound impact on our understanding of many new fundamental genetic mechanisms that derive cancer initiation, progression and maintenance and on our management of specific cancers such as melanoma and breast cancer (Alexandrov et al., 2013, (TCGA), 2012, Yu et al., 2012, Ellis et al., 2012). The scope of these studies has included activation and suppression of cancer signalling pathways, the effects of epigenetic regulation, alteration of the splicing machinery, identification of novel therapeutic targets, understanding molecular mechanisms of resistance to targeted therapies, translating genome-scale variation into the clinic and discovering a variety of novel molecular diagnostic biomarkers (Wang and Wheeler, 2014). Detecting both interpatient heterogeneity and intratumoral heterogeneity can facilitate treatment decisions for personalised patient therapy leading to improved outcomes (Cowin et al., 2010, Boehm and Hahn, 2011).

1.7 The transcriptome

The transcriptome is the output of the genome that is transcribed into RNA and includes the entire set of transcripts in a cell. This consists mainly of ribosomal RNA 'rRNA' (80–90 %) with smaller proportions of 2–4 % messenger RNA (mRNA), 15 % transfer RNA (tRNA) and a small fraction of non-coding RNAs (1 %) including small interfering RNA (siRNA), microRNA (miRNA), small nuclear/nucleolar RNAs (snoRNAs), piwiRNA (piRNA), natural antisense transcripts (NATs), transcription start site-associated RNA (TSSa-RNA) and long non coding RNA (lncRNA) (Lindberg and Lundeberg, 2010) **(Figure 1.4)**. Transcriptome expression varies according to a specific cell type, developmental stage or physiological stimulus (Licatalosi and Darnell, 2010). Studying the transcriptome is essential for understanding the functional elements of the genome, identifying the cellular constituents and understanding how transcriptome is involved in development and disease. The aims of transcriptomic analysis are to catalogue all types of transcripts in cells, determine the transcriptional boundaries of genes (start sites, 5' and 3' ends), identify splicing variants and post-transcriptional modifications and to quantify the expression levels of each transcript under different conditions (McGettigan, 2013, Mutz et al., 2013).

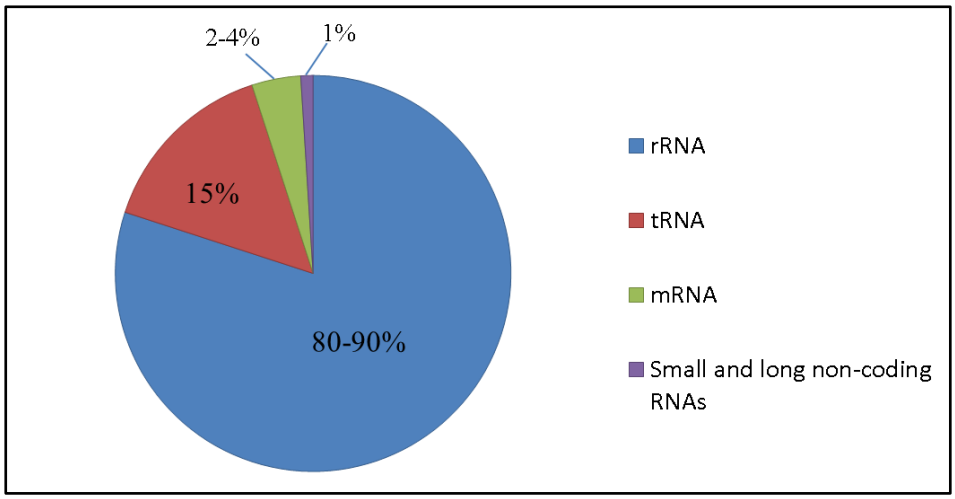


Figure 1.4 Percentage for different RNA classes in human transcriptome.

Recent transcriptomic analysis by large project such as the Encyclopaedia of DNA Elements (ENCODE) project (ENCODE, 2012) has revealed that less than 2 % of the transcriptional output of the human genome encodes proteins and around 98 % encodes different classes of non-coding RNAs, as listed above with potential transcriptional and epigenetic regulatory functions (Raz et al., 2011, Kapranov et al., 2010, Djebali et al., 2012). Many of these non-coding transcripts are reported to be deregulated in cancers, including breast, colon and prostate cancers (Maruyama et al., 2012, Martens-Uzunova et al., 2013, Hauptman and Glavac, 2013). However, the function and regulation of many of these non-coding transcripts class in hypoxia has not been studied.

1.7.1 Small non-coding RNAs

Small non-coding RNAs are defined as transcripts that are less than 200 base pairs (bp) in length that are not translated into proteins. Small non-coding RNAs include several classes; microRNA (miRNA), small nuclear/nucleolar RNAs (snoRNAs), transfer RNA (tRNA) and piwi-interacting RNA (piRNA). These varied classes of RNAs have an important impact in many biological processes (Esteller, 2011, Maute et al., 2014).

The most widely studied class of small non-coding RNAs are miRNAs, which are around 22 nucleotides in length and are highly conserved across species. The function of miRNA is to mediate post-transcriptional gene silencing by inhibiting the translation of mRNA into proteins (He and Hannon, 2004). miRNAs regulate the expression of a number of genes that are involved in key processes, including proliferation, differentiation, apoptosis and development (Mendell, 2005).

Other classes of small non-coding RNAs are less well characterised. For instance, piRNAs are 24–30 bp in length and bind the PIWI subfamily of Argonaute family proteins. Their function is mainly linked to the maintenance of genome stability in germline cells (Aravin et al., 2007,

Brennecke et al., 2007). snoRNAs are considered as intermediate-size non-coding RNAs with lengths ranging from 60 to 300 nt. snoRNAs are essential for the formation of small nucleolar ribonucleoproteins complex (snoRNPs), which are responsible for sequence-specific 2'-O-methylation and pseudouridylation of ribosomal RNA (Kiss-Laszlo et al., 1996, King et al., 2003).

1.7.1.1 Hypoxia and small non-coding RNAs

Several small non-coding RNAs of different types, including miRNAs, snRNAs and piRNAs have been reported to be altered in cancer (Stahlhut and Slack, 2013, Nallar and Kalvakolanu, 2013, Mei et al., 2013). However, the role of hypoxia in their dysregulation in these tumours is less well studied. Indeed, although hypoxic regulation of miRNAs is widely reported (Gee et al., 2013, Camps et al., 2014), the effect of hypoxia on tRNAs, snRNAs, and piwiRNAs have not been studied before.

Amongst those studies examining the hypoxic regulation of miRNAs, miR-210, -155, -372/373 and -10b are most frequently reported to be upregulated (Huang et al., 2009, Bruning et al., 2011, Crosby et al., 2009), while, miR-20b and miR-200b are commonly reported to be downregulated in hypoxia (Chan et al., 2011, Lei et al., 2009). Some of the hypoxia-induced miRNAs such as miR-210, -155, and -373 have been shown to contain HIF binding sites indicating direct transcriptional regulation of these miRNAs by HIF (Huang et al., 2009, Bruning et al., 2011, Huang et al., 2010).

1.7.2 Long non-coding RNAs (lncRNAs)

lncRNA are defined as non-protein coding transcripts of longer than 200 bp. The human genome encodes thousands of lncRNA transcripts (Geisler and Coller, 2013) (The FANTOM consortium, 2005). Recently, the GENCODE project has published the largest manually curated catalogue of lncRNAs in man. The GENCODE-identified 9,640 lncRNAs loci which represent 15,512 transcripts(Derrien et al., 2012). These lncRNAs showed lack of homology to any protein, absence of open reading frames and no high conservation through the majority of their exons. 16.8 % of lncRNAs are polyadenylated and 13.9 % of lncRNAs have chromatin marks, in particular histone 3 lysine 4 trimethylation "H3K4me3" at the promoter and histone 3 lysine 36 trimethylation "H3K36me3" across the gene body (Mattick, 2009, Derrien et al., 2012).

ENCODE has classified all the lncRNAs based on their location with respect to protein-coding genes to five classes; **antisense RNAs,** are those which overlap with any exon of protein-coding gene on the opposite strand, **lincRNAs,** are transcribed in an intergenic non-coding loci, **sense overlapping RNAs,** are transcribed within an intron of coding gene on the same strand, **sense intronic RNAs,** are resides within an intron of a coding gene, but does not overlap any exons on the same strand, **processed transcripts,** have no open reading frame (ORF) and cannot be classified in any of the other categories because of complexity in their structure **(Figure 1.5A)**. Based on this classification, ENCODE lncRNAs data consists of 5,058 lincRNAs, 3,214 antisense RNAs, 378 sense intronic RNAs and 930 processed transcripts **(Figure 1.5B)** (Derrien et al., 2012).

Interestingly, ENCODE analyses suggested that lncRNAs are transcribed through pathways similar to protein-coding genes including, histone-modification, splicing, and exon/intron structure. However, lncRNAs mainly localised in chromatin and nucleus (Derrien et al., 2012, Djebali et al., 2012).

The majority of lncRNAs are found spliced, accounting for 98 % and bias toward only two exons. Around 42 % of lncRNAs have only two exons compared with 6 % of mRNAs (Derrien et al., 2012). It is also noted that lncRNAs have slightly longer exons and introns compared to protein-coding genes (median length of exons for lncRNAs 149 bp and for mRNAs 132 bp; median length of introns for lncRNAs 2280 bp and for mRNAs 1602 bp). However, because lncRNAs have less exons, the overall length of lncRNAs is shorter than mRNAs (median length for lncRNAs 592 bp compared with 2453bp for mRNAs) (Derrien et al., 2012, Djebali et al., 2012)

Several transcription factor-binding sites are present around their promoters (Ulitsky and Bartel, 2013, Guttman et al., 2009, Mercer et al., 2009). The function of the majority of lncRNAs remains unclear and their biological role has been emerging (Ulitsky and Bartel, 2013, Flintoft, 2013). Various functions have been proposed; for example, they may act as enhancer marks (De Santa et al., 2010), in genome reprogramming (Loewer et al., 2010) and could act in cis to activate transcription or in trans to repress transcription (Mercer et al., 2009, Rinn et al., 2007, Huarte et al., 2010).

There is emerging evidence to suggest that lncRNAs play an important role in tumour biology (Du et al., 2013, Qiu et al., 2013). Aberrant expression of a number of lncRNAs in cancer has been associated with disease progression and reported as an independent predictor for disease outcomes. Such dysregulated lncRNAs include the HOX Transcript Antisense RNA (HOTAIR) in breast cancer (Gupta et al., 2010) and the Metastasis Associated Lung Adenocarcinoma Transcript 1 (MALAT-1) in lung cancer (Ji et al., 2003).

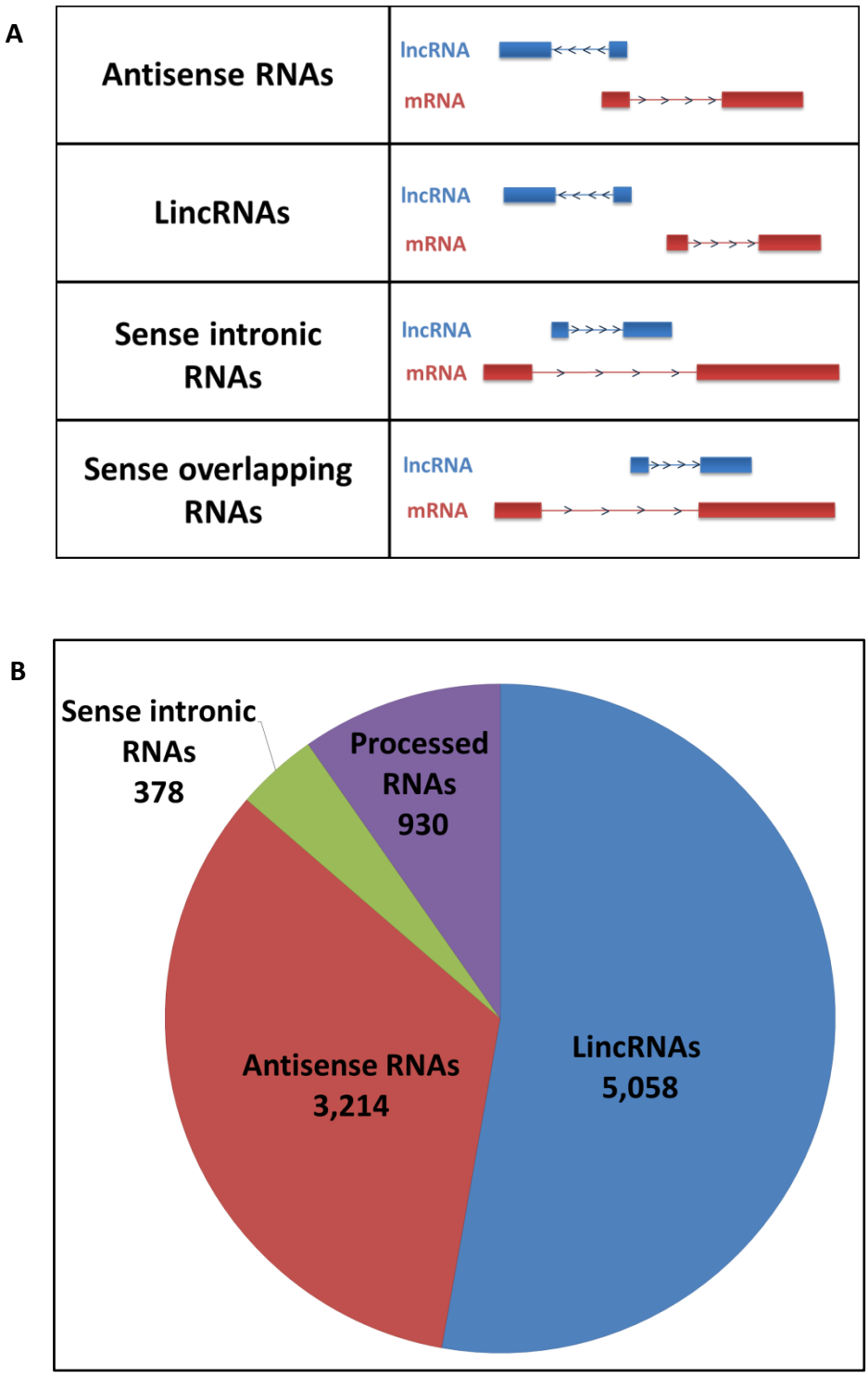


Figure 1.5 Annotation of lncRNAs in the human genome. A) lncRNAs are classified based on their locations with respect to protein-coding genes. **B)** Number of lncRNAs based on their classification.

1.7.2.1 Hypoxia and lncRNAs

Although the regulation and role of lncRNAs in hypoxia has not been well characterised, during the current project a number of studies have reported the function and regulation of specific and individual lncRNAs in hypoxia. Yang et al., (2014) used a bioinformatic analysis to identity six lncRNA (LincRNA-p21, lnc-ROR, MALAT1, NEAT1, HOTAIR and A7) that have potential HREs within their promoter regions (Yang et al., 2014b). They focused their analysis on LincRNA-p21, since it showed the highest induction in Hela cells when exposed to hypoxia (Yang et al., 2014b). They reported that lincRNA-p21 is induced by HIF-1α and binds to both HIF-1α and VHL, disrupting the VHL-HIF-1α interaction and generating a positive feedback loop that further enhances HIF-1α accumulation in hypoxia (Yang et al., 2014b). In addition, they found that forced over-expression of lincRNA-p21 promotes tumour growth in xenograft models (Yang et al., 2014b).

Fredin et al. exposed two cancer cell lines (HCT-116 and MCF-7) to 1 % O_2 or DMOG for 24 and 48 hours and found five long non-coding transcripts that were overexpressed and transcribed from ultraconserved regions named 'hypoxia-induced noncoding ultraconserved transcripts' (HINCUTs) (Ferdin et al., 2013). These transcripts were over expressed in colorectal tumours. Expression of HINCUT-1 (c.475) lncRNA, which is transcribed as part of a retained intron of the host protein-coding gene, O-linked N-acetylglucosamine transferase (OGT), was selected for validation (Ferdin et al., 2013). Knockdown of HINCUT-1 resulted in reduced cell proliferation, with G2/M blockade and reduced expression of OGT (host gene) in hypoxia. They also demonstrated that HINCUT-1 potentially has enhancer-like functions in normoxia and hypoxia (Ferdin et al., 2013).

In another study, Young and colleagues demonstrated that lncRNA-Low Expressed in Tumour (lincRNA-LET) is expressed at low levels in hepatocellular carcinoma (HCC), colorectal cancers and squamous-cell lung cancer tumours compared to their normal tissues (Yang et al., 2013a). In addition, they found that hypoxia suppresses lncRNA-LET expression though activation of

histone deacetylase 3 (HDAC3), which decreases histone H3 and H4 acetylation levels around the lncRNA-LET promoter region. Moreover, suppression of lncRNA-LET in hypoxia is important for the accumulation and stabilisation of nuclear factor 90 (NF90) and HIF-1α, which induce invasion (Yang et al., 2013a). Furthermore, lncRNA-LET expression was found inversely correlated with CA9 expression in primary HCC tumours, suggesting that lncRNA-LET is involved in hypoxia responses in cancer (Yang et al., 2013a).

Among early studies, Matouk and colleagues showed that imprinted maternally expressed transcript (H19) is also a HIF-1α-dependent lncRNA and that its expression increases in HCC cell lines and bladder carcinoma cell lines in response to hypoxia (Matouk et al., 2007). They reported that suppression of H19 in Hep3B xenograft models resulted in a very significant reduction of tumour volumes. In a later study, Matouk et al. showed elevation of H19 expression in hypoxia is associated with the status of the p53 tumour suppressor, as cells with p53wt show no induction in H19 expression upon hypoxia, whereas the p53null cells exhibited a high degree of H19 induced by hypoxia (Matouk et al., 2010). These data indicate lncRNA expression may also be dependent on DNA mutation status.

Although the studies discussed above have been useful in understanding role and regulation of specific lncRNA in hypoxia, they have a number of limitations. The use of bioinformatic approaches and/or microarray based technologies limits the discovery of potentially hypoxia regulated lncRNAs for several reasons: many microarray based technologies frequently lack updated lncRNAs probes; many lncRNA are expressed at low abundance and may be difficult to detect by microarray; microarray technology has a relatively low dynamic range and often underestimates fold-changes; and bioinformatic and microarray analysis don't allow the identification of non-annotated hypoxia-specific lncRNAs. Therefore, profiling lncRNAs at the genome wide level in hypoxic tumours using technologies such as RNA sequencing will shed light on another level of regulation of the hypoxic transcriptome and clinical translation of these

hypoxia-associated lncRNAs may assist development of new markers and therapeutic targets in cancer.

1.7.3 Natural antisense transcripts

At a number of genomic loci, overlapping RNA molecules may be transcribed from opposing strands of the DNA and are termed natural antisense transcripts (NATs), also known as opposite strand transcripts (Faghihi and Wahlestedt, 2009, Werner, 2013). NATs were first discovered in bacteria and found to control plasmid numbers (Lacatena and Cesareni, 1981). In the mammalian genome both sense and antisense transcripts can be non-coding or coding. The commonest pattern in the human genome, however is a non-coding antisense transcript to overlap a protein coding sense transcript (Wight and Werner, 2013). Antisense transcripts are not evenly distributed across the gene and most commonly overlap one or other end of the sense genes.

The function of antisense transcripts is not well understood. Recent studies indicate a regulatory role for natural antisense transcripts at both transcriptional and post-transcriptional levels, including epigenetic silencing, regulating expression of sense and neighbouring transcripts (Faghihi and Wahlestedt, 2009, Rosok and Sioud, 2004, Wahlestedt, 2006). It is proposed that natural antisense transcripts can regulate by either inducing or suppressing expression of the sense transcript. The terms co-regulation and anti-regulation are used to describe inductive or suppressive effects of NATs, respectively **(Figure 1.6).**

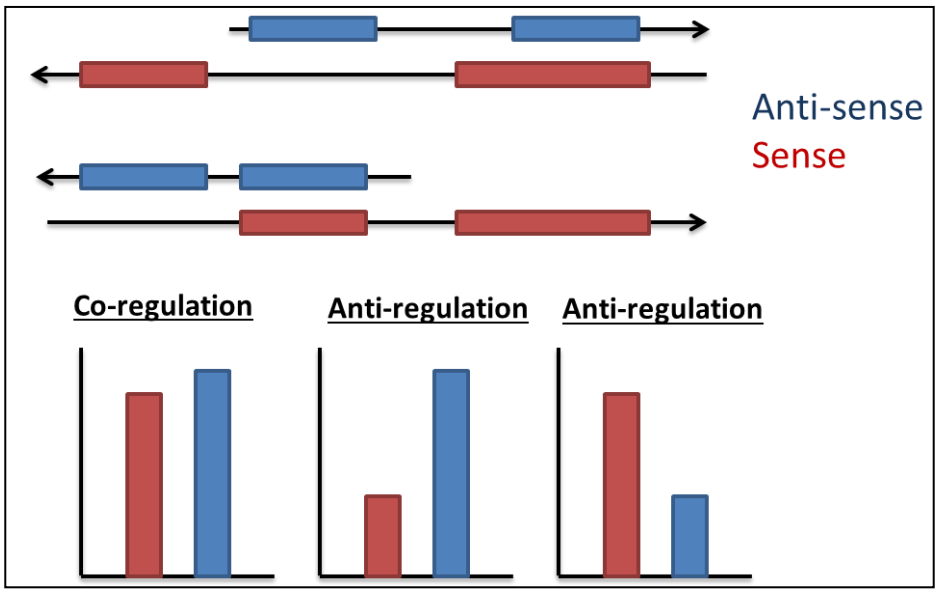


Figure 1.6 Genomic configuration of natural antisense transcripts. Top panels: sense transcript is indicated in red, antisense transcript in blue. Bottom panel: NATs can affect the expression of the overlapping sense gene in a co-regulation (NAT, sense transcript) or anti-regulation (NAT, sense transcript) manner.

1.7.3.1 Hypoxia and natural antisense transcripts

The expression of natural antisense transcripts has been found to be altered in several cancers, including breast, renal and colon cancer (Kohno et al., 2010, Grigoriadis et al., 2009). In hypoxia, a natural antisense transcript of hypoxia-inducible factor 1 (aHIF-1α) has been described originating from the 3'-untranslated region and body of the HIF-1α gene in hypoxia (Rossignol et al., 2004). It is proposed that aHIF-1α regulates expression of HIF-1a by destabilising a splice variant of HIF-1α mRNA (Uchida et al., 2004, Cayre et al., 2003). This destabilisation exposes the AU-rich elements of HIF-1α mRNA where antisense RNA can bind (Uchida et al., 2004, Rossignol et al., 2004). In addition, expression of aHIF-1α is associated with poor disease-free survival and patient outcome with breast and renal cancer (Cayre et al., 2003, Rossignol et al., 2002). However, transcriptome sequencing analyses have shown that many genes overlap with antisense transcripts raising the possibility that hypoxic regulation of antisense transcripts is a

common phenomenon. Consequently, characterising the hypoxic regulation of these natural antisense transcripts at the genome-wide level and their role in regulating sense transcripts will potentially shed a light on networks of transcriptional regulation in hypoxia.

1.7.4 Non polyA transcripts

As well as short RNAs, a number of functional long transcripts (>200 bp in length) are known to lack polyA tails, including rRNAs, some lncRNAs and replication dependent histone mRNAs (Marzluff et al., 2008, Livyatan et al., 2013).

In eukaryotic cells, nascent pre-RNAs are subjected to a number of transcriptional and post-transcriptional modification processes during their maturation. These processes are highly regulated(Zhang et al., 2013b). The 3'end maturation of nascent pre-RNAs is essential for RNA release from the transcription template and for its function after maturation. However, some classes of RNAs, such as replication-dependent histone mRNAs, do not undergo the mutation. The 3'end processing of RNApol2 transcribed RNAs involve cleavage and polyadenylation of nascent pre-RNAs. The machinery of cleavage/polyadenylation recognises the AAUAAA hexanucleotide, along with G/U rich sequences present in nascent pre-mRNAs, leading to endonucleolytic cleavage of the pre-RNA by cleavage and polyadenylation specificity factor-37 (CPSF-73) (Mandel et al., 2006). Then, poly(A) polymerase adds a polyA tail (up to 200 to 250 adenosines) at the 3'-end of the transcript (Richard and Manley, 2009).

However, there is an increasing evidence, that around 70 % of all human genes are subjected to alternative cleavage and polyadenylation (Hoque et al., 2013). These alternative cleavage and polyadenylation sites result in formation of a range of RNA isoforms with different coding potentials and/or 3' untranslated regions and additionally affect their stability and subcellular localisation (Tian and Manley, 2013). Thus, it is become clear that the proper 3' end processing is key to the functionality of a mature RNA (Zhang et al., 2013b).

Interestingly, although it is generally believed that replication-dependent histone mRNAs are the only cellular mRNAs without polyA tails, a number of recent transcriptome profiling studies reported many mature mRNAs with short polyA tails (> 30 bp), non polyA lncRNAs and non polyA enhancer RNAs (eRNAs) in human cells (Marzluff et al., 2008, De Santa et al., 2010, Memczak et al., 2013). The molecular mechanism and function of the short poly A RNAs still remains unclear. There are a number of reported mechanisms by with RNAs are stabilised without polyadenylation of 3′ ends, including RNase P cleavage to produce a mature 3′ -end (Brown et al., 2012, Wilusz et al., 2012), snoRNP complexes capping at both ends of the transcript (Yin et al., 2012), or by creating circular structures to protect them from degradation (Memczak et al., 2013, Hansen et al., 2013).

Recently, the ENCODE project reported that most annotated elements (exons, splice sites, transcripts and genes) are similar for both polyadenylated and non-polyadenylated RNAs. In addition, a small proportion of ENCODE elements were exclusively found in the non-polyadenylated RNA fractions (0.4 % of exons, 2.8 % of splice sites, 3.3 % of transcripts and 4.7 % of genes). Comparing gene expression levels, polyadenylated RNAs span six orders of magnitude for (from 10^{-2} to 10^{4} reads per kilobase per million reads (RPKM)), while non-polyadenylated RNAs five orders of magnitude (from 10^{-2} to 10^{3} RPKM) (Djebali et al., 2012).

In contrast to mRNA expression, relatively little is known about non-polyA transcripts in cancer. One possible function of non-polyA transcripts is regulation of the cell cycle through effects on replication-dependent histone expression (Yang et al., 2011a). Previous studies have shown that a significant proportion of eukaryotic gene transcripts may be non-polyA or bimorphic (containing both polyA+ and polyA- transcripts) (Wu et al., 2008, Yang et al., 2011a). In a recent study by Zhao et al. transcriptome sequencing identified both polyA+ and polyA- RNA in a panel of cancer cell lines including HMEC, MCF-7, NHLF, and A549 (Zhao et al., 2013). Differential expression of both polyA+ and polyA- RNAs were found in cancer with many altered non poly A mRNAs involved in various cancer related pathways such as cell cycle, apoptosis, and cell death

(Zhao et al., 2013), suggesting that many non-poly A mRNAs have key functions and play important roles in cancer. However, the link between hypoxia and non-polyA transcripts has not been studied previously. Understanding the role of non-polyA transcripts in hypoxic tumours could enhance our understanding of new aspects of hypoxia regulation.

1.8 Next-generation sequencing technologies

In the past decade, rapid advancements in high-throughput sequencing technologies have revolutionized our understanding about structural and functional genomics (Koboldt et al., 2013). These technologies are broadly divided into two main classes, namely next-generation sequencing (NGS) also known as second generation sequencing, which includes both conventional and desk-top instruments. The second class, known as third generation sequencing includes single DNA molecule sequencing technologies (Mardis, 2013). The conventional NGS technologies include platforms such as Roche 454 GS FLX, Illumina GA/HiSeq and Life Technologies SOLiD/5500 sequencing instruments. The advantages of conventional NGS technologies are that they are cost-effective and compatible with large scale sequencing studies, for example, whole-genome sequencing (WGS), whole transcriptome sequencing and whole-exome sequencing (WES) (Ku et al., 2013).

The NGS technologies use different chemistries for sequencing. Illumina (GA/ Hiseq/ Miseq) exploits reversible terminator chemistry, whereas, Life technologies platforms (SOLiD excluding Ion torrent) employ DNA ligase enzyme to mediate sequencing by ligation to synthesise the complementary strand of the fragment (Mardis, 2013). The NGS technologies are also known as 'massive parallel' sequencing because of their high throughput, for example, Illumina's HiSeq2000 and Life Technologies SOLiD/5500 instruments have the ability to generate more than hundred gigabases of data per single run.

The Illumina sequencing platforms are widely used and are based on bridge amplification on a solid surface known as a flowcell. Single stranded DNA fragments are first attached to a flowcell using specific adaptors. These adaptors have complementary sequence to probes on the flowcell. The unattached end of the fragment binds and hybridises to other probes on the flowcell forming bridge-like structures. Then the templates are subjected to amplification in order to produce clusters for sequencing. The sequencing is performed using reverse terminator chemistry where the four nucleotides are labelled by different fluorescent colours in the sequencing reaction. The sequencing reagents and fluorescent labelled nucleotides are added onto the flowcell. At any one time, single fluorescently labelled nucleotides are incorporated in the synthesis of the DNA strand per cycle followed by a wash step to remove excess reagents. The fluorescent signals are captured by imaging across the whole flowcell. These steps are then repeated to complete the sequencing reaction. A single flowcell has more than hundred million clusters, which allow for the spontaneous sequencing of vast numbers of DNA templates (Mardis, 2013, Metzker, 2010).

Although NGS sequencing technologies have their advantages in sequencing throughput, there are a number of limitations. They require longer run times of several days and amplification steps to generate templates for sequencing. For Roche and Life Technology sequencing platforms (Roche 454 GS FLX, GS Junior, Life Technologies SOLiD and Ion Torrent) the amplification method is based on emulsion PCR solution, while for Illumina platform technologies (Illumina GA, Hiseq and Miseq) it is based on bridge amplification on a solid surface. It is clear that the amplification steps can result in errors that can cause amplification bias. In addition, these technologies require an amplification step during the clustering process, which is prone to introduce errors and may result in uneven amplification in the sequencing read. Under-sequencing the template can result in incorrect detection and calculation, while over-sequencing the template is not cost-effective. Another limitation of NGS technologies is base-calling error rates, for instance Life Technologies platforms have the highest accuracy with

less than 0.1 % of error rate, whereas Illumina platforms have error rate of 0.5 % to 1.5 % (Mardis, 2013, Metzker, 2010).

Although NGS technologies have a number of limitations, with current advances and development in these technologies these limitations are consistently improving. In addition, some limitations are overcome by the introduction of single DNA molecule sequencing which are performed in real time and omit the requirement of amplification, thereby avoiding or reducing the errors resulted by clonal amplification.

1.8.1 Bench-top sequencing technologies

Recent studies have evaluated the application of NGS technologies as diagnostic tools for Mendelian diseases (Koboldt et al., 2013, Ku et al., 2013). The need of sequencing small panels of genes in clinical samples for routine clinical utility has led to the development of several bench-top NGS platforms such as Roche 454 GS Junior, Ion Proton, Ion Torrent, and Illumina Miseq. Since conventional NGS technologies might not be suitable for the targeted gene sequencing studies, bench-top platforms have a niche market for clinical diagnosis, where targeted gene sequencing is required.

The throughput of bench-top NGS platforms (Ion Torrent, and Illumina Miseq) range from 10 Mb to more than 1 Gb and with read lengths ranging from 100 to 200 bp. The Illumina Miseq is based on a reversible terminator sequencing technology (as explained above), in contrast the Ion Torrent and Ion Proton sequencing are based on post-light sequencing technology.

The Ion Torrent and Ion Proton sequencing platforms contain a sequencer and a range of semiconductor sequencing chips. In these two platforms, the DNA polymerase incorporates a nucleotide for synthesis of the complementary strand of the template, resulting in a release of a hydrogen ion and a subsequent voltage change. In the case that the nucleotide is not

complementary then no change in the voltage will be recorded. Since these technologies (Ion Torrent and Ion Proton sequencing) do not require light scan, and cameras as other technologies to detect the incorporated nucleotide, these technologies significantly simplify overall sequencing process and reduce sequencing time.

Owing to their throughput and quick turnover time, bench-top platforms have a promising future for targeted gene sequencing, for instance as diagnostic tools in clinics.

1.8.2 Single molecule sequencing technologies

Notwithstanding advancements in next generation technologies that have improved our understanding of genomics, single molecule sequencing (the third generation sequencing) is emerging with improved aspects of sequencing.

Compared to NGS technologies, single molecule sequencing has several advantages, including quick turnover time, higher throughput, ability to sequence longer read lengths, improved consensus accuracy and use of smaller amounts of template. These improvements may be utilised in many aspects of genomic studies, such as *de novo* assembly, detection of haplotypes, detection of rare variants and use of smaller amounts for clinical settings (Pushkarev et al., 2009). There are a handful of commercial single molecule sequencing platforms, for instance, Helicos Biosciences also known as true Single Molecule Sequencing (tSMS), Pacific Biosciences known as Single Molecule Real Time (SMRT) sequencing, and Oxford Nanopore (Mardis, 2013, Schadt et al., 2010).

The first commercially available single molecule-sequencing instrument is Helicos Biosciences. Briefly, the technology is based on extending the individual DNA molecules by attaching it on a planar surface using specific primer, fluorescent nucleotide (Virtul Terminator nucleotides), and modified DNA polymerase followed by imaging of the DNA molecules (Harris et al., 2008). The tSMS technology has been also applied for sequencing direct RNA molecules without the need of

converting RNA to cDNA or incorporating ligation and amplification steps (Pushkarev et al., 2009).

Pacific Biosciences has developed the first real time sequencing technology that allows the direct detection of a single molecule of DNA polymerase as it synthesis the complementary DNA strand (Korlach et al., 2010). This Pacific Biosciences sequencing technology has many advantages compared to other sequencing platforms. It requires less reagent and sample preparation, requires no amplification, less time consumption (scanning and washing steps are absent), and longer read length (maximum read length of 10 Kb) (Carneiro et al., 2012, Shin et al., 2013).

Despite the fact that Helicos tSMS and Pacific Biosciences SMRT have a number of potential advantages, there are a number of limitations. Both technologies have a high read error rate of more than 5 %, dominated by insertions and deletions (Schadt et al., 2010).

Nanopore-sequencing technologies are based on threading single DNA molecules through a nanopore or by positioning DNA molecules in the vicinity of nanopore. The individual nucleotides are read as they are run through the nanopore (Ying et al., 2013). These technologies rely on passing the DNA molecule or its components through a nano hole and the nucleotides are detected by the emission of their electric currents or optical signals. Since nanopore sequencing technologies use single unmodified DNA molecule in real time, they have the ability to run quickly using very small amounts of DNA material (Ying et al., 2013, Yang et al., 2013b). While nanopore technologies have a favourable potential in omic sciences they are still under development.

1.9 Transcriptome technologies

Several technologies have been developed to study transcriptomes. These are typically hybridization or sequence-based methods. Hybridization-based approaches use fluorescently labelled complementary DNA (cDNA) conjugated to high-density microarrays; for example, a microarray is a high-resolution method that allows mapping of transcribed regions in the genome (Bertone et al., 2004). Although hybridization-based approaches are relatively inexpensive and could be used for high throughput analysis to interrogate large genomes, these methods have a number of limitations. These include a limited dynamic range and a high level of background due to cross-hybridization (Shendure, 2008, Russo et al., 2003). In addition, complicated normalization methods are required to compare expression analyses from different experiments and platforms (Russo et al., 2003).

In contrast, sequence-based methods use direct cDNA for sequencing. Sanger sequencing was initially used to determine cDNA sequences, but this method is expensive and has low throughput (Boguski et al., 1994). To overcome some of these limitations, alternative sequencing methods were developed using high throughput tag-based approaches to quantify gene expression, such as serial analysis of gene expression (SAGE) (Harbers and Carninci, 2005) and cap analysis of gene expression (CAGE)(Kodzius et al., 2006). However, these methods are generally based on the expensive Sanger sequencing technology and have limitations in distinguishing different isoforms and in mapping the short tags uniquely to the reference genome (Wang et al., 2009). These disadvantages limit an accurate expression quantification and annotation of transcriptome structure.

The development of next generation sequencing methods that allow hundreds of millions of DNA fragments to be sequenced simultaneously have revolutionised the way genomes are analysed and our understanding of genetics (Wang et al., 2009, Ozsolak et al., 2009). The recent application of these massively parallel sequencing methods to the analysis of RNA, termed

"RNA-Seq" (RNA sequencing), has produced similar changes in the way the transcriptome can be studied.

1.9.1 RNA-seq

Next generation RNA sequencing "RNA-seq" has significant advantages to existing approaches that characterise and quantify the transcriptome. In particular, RNA-seq allows the accurate detection and quantification of transcript expression levels, splice-junctions, gene fusions, RNA editing, transcript isoforms and allele-specific expression (mutation and single-nucleotide polymorphism 'SNP' within transcripts) (Wang et al., 2009, Costa et al., 2010, Sanchez-Flores and Abreu-Goodger, 2013). This vital information enhances our understanding of transcriptome differences in health and disease states. Furthermore, unlike microarray technologies, RNA-seq does not require prior knowledge of transcript sequences, so it allows the identification, characterisation and quantification of novel transcripts (Costa et al., 2010, Roberts et al., 2011). Another important feature of RNA-seq is that it has a larger dynamic range of expression levels compared to microarrays (Marguerat and Bahler, 2010, Wilhelm et al., 2010) and is highly reproducible for both technical and biological replicates (Marioni et al., 2008, Liu et al., 2014).

A number of studies have used RNA-seq to determine transcriptome-wide differences in polyA RNA and/or miRNAs between normal tissues and different types of breast cancers, both in cell lines and in primary tumours (Sinicropi et al., 2012, Horvath et al., 2013). These studies found a long list of transcripts that are altered in breast cancer and which may contribute to the metabolic and growth aberrations associated with cellular transformation (Sun et al., 2011, Hah et al., 2011, Inaki et al., 2011, Craig et al., 2013, Horvath et al., 2013). Other studies have used RNA-seq for molecular classification of breast cancer or to study treatment responses in order to find transcripts that are associated with clinical outcome (Sun et al., 2011, Wu et al., 2010). However, a comprehensive analysis of hypoxic changes in the tumour transcriptome by RNA-seq has yet to be performed.

1.9.2 Challenges and advances in RNA-seq

Recently, advances in RNA-seq methodologies have allowed comprehensive cataloging of the transcriptome, including the uniform detection of transcript coverage, distinction of sense and antisense transcripts, sequencing of low amount of RNA and improved detection of splicing events (Ozsolak and Milos, 2011, Linnarsson, 2010).

When the current project started, commercially available RNA-seq methods from Illumina were designed either for polyA RNA or for sequencing small RNAs, which together represent only a small fraction of the entire transcriptome. These methods are poor at analysing some regulatory classes of transcript such as non-polyA RNA, tRNA and lncRNA (which lack poly A). Analysis of the whole transcriptome, in a manner that is not limited to polyA-selected RNAs, by coupling RNA-seq to the analysis of RNA depleted of abundant rRNA molecules (5S, 5.8S, 18S and 28S) has been reported (Cui et al., 2010, Raz et al., 2011, Kapranov et al., 2010, O'Neil et al., 2013). These studies used commercially available kits to selectively deplete large rRNA molecules from total RNA. They reported enriched detection of tRNA, non polyA mRNA, pre-processed RNA, lncRNA and small regulatory RNA such as miRNA and small RNA transcripts with undefined functions (Kapranov et al., 2010, van Dijk et al., 2011, Cui et al., 2010, O'Neil et al., 2013, Jenjaroenpun et al., 2013). These findings suggested that "ribo-depleted" RNA-seq allows a much more comprehensive identification and quantification of the transcriptome across all classes of RNA when compared to polyA RNA-seq.

Until very recently, another limitation of the commercially available Illumina RNA-seq methods has been the inability to determine orientation of the RNA transcript i.e. the DNA strand that had been transcribed. Distinguishing strand-specificity for an RNA molecule substantially enhances transcriptome analysis, by allowing for accurate identification of potential antisense transcripts with a regulatory role, annotates exact boundaries of overlapping transcribed genes and accurately quantifies coding or non-coding sense and antisense transcripts (Ozsolak and Milos, 2011, Parkhomchuk et al., 2009). Recently, several methods have been developed for

strand-specific RNA-seq. These methods are generally based on two approaches. One approach relies on using distinct adaptors that attach only to either the 5' or 3' end of the RNA transcript (Levin et al., 2010). These methods convert adapter ligated RNA to a cDNA library, maintaining the 5' end and the 3' end of the original mRNA molecule **(Figure 1.7a).**

A second approach to strand specific RNA-seq is based on chemically modifying either the RNA or DNA; examples of RNA modifications include attaching an RNA linker (Lister et al., 2008) or bisulfite treating the RNA molecules (He et al., 2008), whereas for DNA modifications these include incorporation of deoxyuridine triphosphate "dUTP" instead of thymidine triphosphate "dTTP" during second-strand cDNA synthesis followed by digestion with uracil-N-glycosylase of the unmarked strand (Parkhomchuk et al., 2009) **(Figure 1.7b/c).** Other approaches, such as reverse transcription on flow cells and computational methods have also been implemented to determine the directionality of RNA (Guttman et al., 2010, Mamanova et al., 2010).

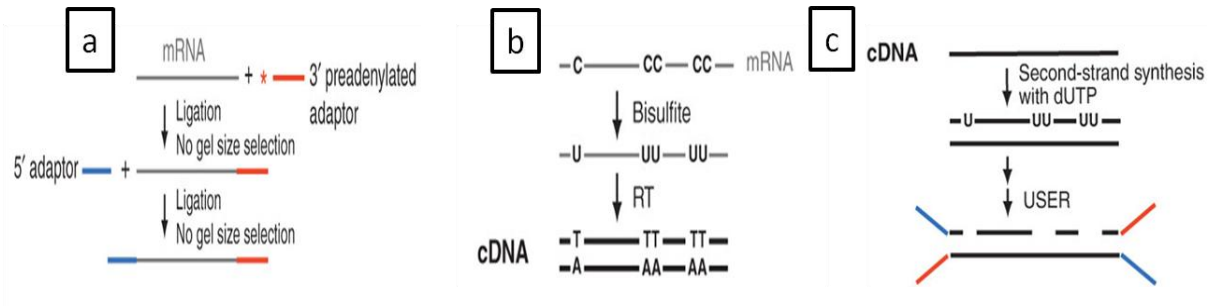


Figure 1.7 Strand specific RNA-seq approaches. A) Specific 3' and 5' RNA adaptors ligation method **B)** Chemical modification of RNA by bisulfite treatment method **C)** Chemical modification during dUTP during second-strand cDNA synthesis (adapted from (Levin et al., 2010)).

A comprehensive study was conducted by Levin and colleagues to evaluate seven published strand-specific RNA-seq library protocols using polyA-selected RNA from *Saccharomyces cerevisiae* (Levin et al., 2010). They found significant differences in expression profiling, strand specificity, sequence coverage and library complexity between the different protocols. The

authors concluded that the dUTP modification in second-strand cDNA method and adaptor ligation to RNA (the Illumina pre-released method) were the preferred protocols having high reproducibility and accuracy (Levin et al., 2010).

Overall, there is a need to standardise methods that allow comprehensive transcriptome sequencing. These should allow for identification of the original orientation of transcripts and are not limited to specific classes of transcripts. Such a method will provide a true global view of transcriptome changes and accurate annotation transcripts.

1.10 Hypoxia Genomics

1.10.1 Genome-wide analysis of the hypoxia-regulated transcriptome

Tumour cells utilise a variety of mechanisms to adapt to hypoxic stress. Hypoxia induces the expression of a large number of genes that regulate many biological processes. This includes cell proliferation, angiogenesis, metabolism, apoptosis, immortalization and migration (Harris, 2002). Several studies have reported hypoxia-inducible genes in different types of cancer (Favaro et al., 2011). These studies have defined a range of hypoxic gene signatures that are linked to poor cancer outcome and to a variety of pathological conditions, such as resistance to treatment (Wilson and Hay, 2011, Teicher, 1994, Generali et al., 2009).

Microarray expression analyses have defined hundreds of hypoxia-induced genes that are regulated by the HIF pathways. These include genes with important functions in processes such as proliferation, angiogenesis and apoptosis (Favaro et al., 2011, Buffa et al., 2010). In addition, the expression of many other transcription factors is also dependent on HIF, indicating that HIF regulates a network of responses in hypoxia. Both HIF-1α and HIF-2α are closely related and activate expression of HRE-dependent genes (Gordan et al., 2007, Greijer et al., 2005, Zagorska and Dulak, 2004). However, the extent to which HIF has direct or indirect control over gene

expression in hypoxia, as well as with other factors that are induced by HIF-regulated genes remains unclear.

In a comprehensive analysis to identify hypoxia regulated genes, Ortiz-Barahona and colleagues, performed "meta-analysis" of 19 gene expression profiling datasets that were available in the GEO database which were derived from different cells exposed to hypoxia or hypoxia mimetics (Ortiz-Barahona et al., 2010). They found that the number of non-redundant hypoxia-regulated genes increased significantly with the number of experiments. Across all 19 datasets, the total number of non-redundant hypoxia upregulated genes was 2864 with a further 2929 showing hypoxic downregulation. Further analysis revealed the number of hypoxia-regulated genes that were common to all 19 datasets is very small, with about 24 core hypoxia regulated genes identified in all cell types studied (Ortiz-Barahona et al., 2010). Whilst some of the differences observed between the datasets may result from random statistical noise, these differences strongly suggest that different cell-types activate or repress different genes in response to hypoxia.

Several other studies have reported the importance of the HIF system as the central mediator of transcriptional responses to hypoxia. Elvidge et al. reported microarray expression profiles in MCF-7 (breast cancer) cells treated with normoxia, hypoxia (1 % O_2) or DMOG (dimethyloxalyglycine). These expression analyses revealed that 246 transcripts were significantly upregulated and 190 transcripts were significantly downregulated by hypoxia with the vast majority also regulated by DMOG in the same manner (Elvidge et al., 2006). Combined HIF-1α and HIF-2α siRNA treatments significantly reduced the expression of 141 out of 246 of the hypoxia-induced transcripts (Elvidge et al., 2006). Treatment of the hypoxic MCF-7 cells with siRNA directed against the HIF-α subunit alone resulted in significant down regulation of 127 from 246 hypoxia-upregulated genes. However, in marked contrast, HIF-2α siRNA resulted in statistically significant downregulation of only 5 of 246 genes in hypoxia. Taken together

these results indicate that a large proportion of hypoxia-induced genes are regulated by HIF and that in MCF-7 cells, this regulation appears to be mediated predominantly through HIF-1.

In addition, many studies have profiled the genome-wide hypoxic regulation of microRNAs (Kulshreshtha et al., 2007, Hua et al., 2006, Camps et al., 2008, Radovich and Ragoussis, 2013). Although these studies have reported a number of hypoxia-regulated miRNAs, there has been little overlap in those identified by each study (McCormick et al., 2010). Indeed, only miR-210 has been robustly induced in hypoxia in all the reported profiling studies.

Whilst the studies mentioned above have been very useful in identifying hypoxic gene signatures and in understanding the transcriptional response to hypoxia at the genome-wide level, the microarray-based approaches employed had focused on protein coding genes or miRNAs and these represent only part of the transcriptome. The role and regulation of other RNA classes, in particular non-coding RNAs and regulatory RNAs such as lncRNA, antisense RNA and non-poly A RNA have not been studied previously at the genome-wide level in hypoxia. Importantly, these transcripts may have a vital function in hypoxia or play a critical role in its regulation.

1.10.2 Genome-wide analysis of HIF binding sites

The central role of HIF in the hypoxic regulation of gene expression raises important questions as to how many of these responses result from direct transcriptional activation by HIF and how many are indirect consequences of HIF activation. Several studies have attempted to identify direct transcriptional targets of HIF using *in silico* methods to identify HIF binding sites (Benita et al., 2009, Xie et al., 2005). These studies have identified many previously documented HIF-binding sites and HIF-dependent genes (Benita et al., 2009). Generally, these *in silico* methods use two approaches; *de novo* prediction of motifs (Gupta and Liu, 2005) or redefined position weight matrices (PMW) (Xie et al., 2005). However, both approaches have limitations in

defining HIF-binding sites genome-wide. For example, the binding sequence of HIF is very short, occurs very frequently and is ubiquitously spread across the human genome resulting in unreliable *in silico* genome-wide prediction.

Recent advances in the analysis of protein-DNA interactions methods and DNA sequences have allowed identification of transcription factor binding sites across the whole genome. Five studies have profiled HIF binding across the genome of different cell lines using chromatin immunoprecipitation (ChIP) of HIF-1α, HIF-2α, and HIF-1β coupled to either tiled arrays (ChIP-on-chip) or to high-throughput sequencing (ChIP-seq)(Xia et al., 2009, Tanimoto et al., 2010, Schodel et al., 2011, Schodel et al., 2012, Mimura et al., 2012).

In summary, the majority of these studies have reported around 500 high-stringency HIF-binding sites across the genome in each cell line studied, with the exception of Mimura and colleagues who used less stringent thresholds when compared with other studies and identified 575 HIF-1α binding sites in normoxia (when HIF-1α levels are very low) and 2060 HIF-1α binding sites in hypoxia (Mimura et al., 2012).

Schödel and colleagues, mapped HIF-binding sites in a breast cancer cell line (MCF-7) using ChIP-seq. Using highly stringent criteria, they identified 400 HIF-1 binding sites (HIF-1α and HIF-1β signal) at 356 gene loci and 425 HIF-2 binding sites (HIF-2α and HIF-1β signal) at 357 gene loci (Schodel et al., 2011). Moreover, they examined relationships between identified positions of HIF binding sites and published hypoxia MCF-7 gene expression datasets. HIF-1 and HIF-2 binding sites were strongly clustered among the most highly upregulated genes in hypoxia. In contrast, they observed no enrichment of HIF-binding sites among hypoxia-downregulated genes. They proposed that these HIF binding loci may exert a direct control over gene expression in response to hypoxia (Schodel et al., 2011).

In this study, genome-wide analysis revealed that HIF can bind to HREs remote from any documented promoter (promoter-distant sites) and can act over long genomic distances to

promote transcription of their target genes. Furthermore, these transcriptional targets of HIF may not be the closest gene to the binding site. A microarray-based approach analysis may not definitively identify a strongly regulated gene in the vicinity of a HIF-binding site. This raises the possibility that other nearby transcripts not represented on the microarray, could in fact be targets for some of these “orphan” binding sites.

1.11 Aim of this study

The aim of this project is to comprehensively profile the hypoxic transcriptome using next generation RNA sequencing (RNA-seq), taking the ER+ MCF-7, breast cancer cells as a model. To date there is no published data for hypoxic transcriptome analysis of human cells using RNA-seq. Since existing expression data in hypoxia has focused on the expression of mRNAs or miRNAs, I have extended the analysis to include other classes of RNA that may be regulated by hypoxia and/or HIF. MCF-7 cells were chosen because of the presence of existing microarray expression data and genome-wide data of HIF binding sites in hypoxia, which will allow comparing and validating the findings of the current project.

Specifically the work performed in this thesis will allow:

1. Measurement of the regulation of different RNA classes including piwiRNA, tRNA, miRNA, snRNA, mRNA and lncRNA in hypoxia
2. Identification of hypoxia-regulated natural antisense transcripts and hypoxia non-annotated transcripts at the genome wide level
3. Correlation of HIF-binding sites with different RNA classes
4. Identification of the HIF-dependent transcriptome, in particular lncRNAs
5. Investigation of RNApol2 and histone markers with the hypoxic transcriptome
6. Understanding new mechanisms of the mode of action of RNApol2 in hypoxia
7. Characterisation of hypoxia-inducible lncRNA and its role in hypoxia tumour biology
8. Utility of hypoxia inducible lncRNA as a clinicopathological and prognosis marker in breast cancer.

Outline of thesis

Chapter 2 of the thesis describes this transcriptome analysis and the development of a bespoke "total RNA-seq" method that allows the detection and quantitation in a single run of both small and long RNA transcripts in a strand-specific manner irrespective of whether these are poly-adenylated.

Chapter 3 profiles the HIF-dependence of these pan-genomic changes in hypoxia by correlating RNA-seq data with previously published HIF-α ChIP-seq data in MCF-7 cells from our group. The analysis takes into account all classes of transcripts (including previously unannotated transcripts) and provides the most comprehensive analysis of HIF transcriptional responses to date. Moreover, in an attempt to identify the HIF dependent transcriptome, HIF-1α, HIF-2α and combined HIF-1α and HIF-2α siRNAs were transfected in hypoxic MCF-7 cells that were then subjected to RNA-seq.

In Chapter 4, RNA pol2 and histone modification (H3K4me3: marker of active promoter) ChIP-seq data were integrated with hypoxic transcriptome data (Total RNA-seq), to gain further insight into the transcriptional regulation of the coding and non-coding transcriptome, with particular emphasis on lncRNAs and non-annotated transcripts.

Finally, in Chapter 5, I selected one of the most hypoxia-induced HIF-binding and HIF-regulated lncRNAs, NEAT1, for further validation and functional analysis. NEAT1 is hypoxia-regulated in a wide array of breast cancer cell lines and *in vivo* models. The hypoxic induction of NEAT1 had effects on tumorigenicity and on the formation of nuclear paraspeckle bodies. Finally, tumour NEAT1 expression correlated with clinical survival in a large cohort of breast cancer patients.

Chapter 2: Transcriptome analysis of MCF-7 cells in hypoxia and normoxia using next generation RNA sequencing

2.1 Introduction

Adaptation to altered levels of oxygen is a vital physiological process within cells. As cancers grow, they outstrip their blood supply, generating a hypoxic microenvironment. This is a common feature of most solid tumours and is associated with an aggressive phenotype (Semenza, 2012a). Hypoxia activates a cascade of key cellular pathways by regulating a number of factors including transcriptional and post-transcriptional changes in gene expression, which allows cells to survive under hypoxic conditions (Kaelin and Ratcliffe, 2008, Harris, 2002). It is estimated that up to 1.5 % of the human genome is transcriptionally activated in response to hypoxia (Denko et al., 2003).

To date, studies that have attempted to characterise the hypoxia transcriptional output have used microarray-based technologies. Major findings of these studies were:

1) The use of different cell types *in vitro* revealed heterogeneity of the induced genes. This may be explained by the specialized role and functions of different cells, which therefore show different gene regulation patterns in hypoxia (Denko et al., 2003, Denko et al., 2000, Koong et al., 2000).
2) Gene regulation differs between early and prolonged hypoxic incubation (Chi et al., 2006, Seigneuric et al., 2007).
3) There is set of core protein coding genes that consistently induced by hypoxia (Favaro et al., 2011).
4) A set of core hypoxia-regulated genes (hypoxia gene signature) could be used to predict clinical outcome for several cancer types, including breast cancer, prostate cancer and HNCC (Koop et al., 2009, Winter et al., 2007, Kong et al., 2007).

These reports have been valuable for our understanding of the transcriptional response to hypoxia, but they have focused largely on protein coding genes and some microRNAs, which represent only a small proportion of the total transcriptional output of the cell.

Recently, large genome studies and projects such as ENCODE have transformed our perspective on regulatory transcriptional networks (Djebali et al., 2012, Kapranov et al., 2007). It is now recognised that mRNA forms only part of a much broader transcriptional output that includes even greater numbers of RNAs that are not translated to protein (Esteller, 2011). These "non-coding" RNAs have important regulatory roles that further shape the transcriptional regulation of the cell. They include short non-coding RNA (<200 nucleotides) such, miRNA, snRNA, piwiRNA, and tRNA as well as long non-coding RNAs (lncRNA) (>200 nucleotides). Little is known about the hypoxic regulation of this non-coding transcriptome, beyond microRNAs.

In this part of work, I have performed a comprehensive profiling of hypoxic transcriptome using next generation RNA sequencing technologies, to investigate both the coding and the non-coding transcriptional landscape in hypoxia.

2.2 Exprimental desigen

Total RNA was isolated from MCF-7 cells grown in normoxia or hypoxic condition (1 % O_2 for 24 hours). It was then subjected to total RNA-seq or polyA selected RNA-seq. The total RNA-seq libraries were prepared with two rounds of ribosomal depletion. To assess the percentage of ribosomal fractions, libraries were cloned in the Zero Blunt® TOPO® vector and sequenced using Sanger sequencing. The optimised total RNA-seq method allows sequencing of both small (< 200 bp) and long (≥ 200 bp) RNAs and utilizes the strand specific adapters that preserve the information from which DNA strand the transcript is produced. The polyA selected RNA-seq libraries were prepared based on the Illumina TruSeq protocol. The total RNA-seq and polyA selected RNA-seq libraries of MCF-7 cells were sequenced on Illumina GXII or HiSeq 2000. The raw data were normalised to the total number of mapped reads and analysed by an in house-developed analysis pipeline. The pipeline allows detection and calculates the differential expression of the different RNA classes in hypoxia. The expression of five different transcripts from each snRNA, mRNA and lncRNA class, with median regulation in hypoxia, was measured by TaqMan expression analysis in MCF-7 cells grown in hypoxia (1 % O_2 for 24 hours) in order to confirm the findings of the pipeline. Additionally, the regulation of these transcripts was also confirmed with DMOG (500 nm, 24 hours) and HIF-1/2α siRNA treatments in MCF-7 cells. The suppression of HIF-1α and HIF-2α subunits was confirmed with western blots in hypoxic MCF-7 cells (1 % O_2 for 24 hours). Induction of several newly reported hypoxia regulated genes in this study was confirmed by qPCR in hypoxic MCF-7 cells (1 % O_2 for 24 hours). Moreover, expression of selected hypoxia regulated miRNAs and non-polyA mRNAs was confirmed by qPCR in a panel of breast cancer cell lines grown in normoxia and hypoxia (1 % O_2 for 24 hours). Regulation of ten most upregulated lncRNAs, non-annotated intergenic and antisense transcripts in hypoxia were validated by qPCR in MCF-7 cells. The coding optional calculator algorithm (http://cpc.cbi.pku.edu.cn/programs/run_cpc.jsp) was used to evaluate protein-coding potential for all identified non-annotated intergenic transcripts in hypoxia. The Ingenuity

Pathway Analysis (IPA) software tool (Version IPA 8.8-3204, http://www.ingenuity.com/products/ipa) was used for gene ontology, gene network and pathway analysis in hypoxia. All the experiments were performed using three independent biological replicates.

2.3 Results

2.3.1 Comprehensive analysis of hypoxia transcriptome

In this study, the oestrogen receptor positive (ER+) breast cancer cell line MCF-7 (the acronym of Michigan Cancer Foundation-7 (Soule et al., 1973)) was used as a model to profile the transcriptome landscape in response to hypoxia (1 % O_2) and normoxia (21 % O_2). This cell line expresses both the HIF-1α and HIF-2α isoforms and was chosen so that results could be cross-referenced to previous microarray and ChIP-seq studies. Total RNA preparations were depleted of ribosomal species and analysed by high-throughput next-generation sequencing (RNA-seq). This methodology also identified the strands from which the RNA-species were transcribed and allowed a complete and unbiased characterisation of the hypoxic transcriptome. High-throughput sequencing of more conventional libraries generated from poly-adenylated (poly-A) RNA were also analysed for comparison **(Figure 2.1A)**.

2.3.2 Development of in house ribosome-depleted strand specific RNA-seq method (Total RNA-seq)

To maximise the transcriptome resolution of the cell by capturing both small and large polyA and non A transcripts in a in strand specific manner, I developed a method for isolating total RNA by depleting the cellular ribosomal RNA and subjected them for total RNA sequencing (Detailed method development is discussed in methods section, flowchart in **(Figure 2.1B)**).

Briefly, isolated total RNA was subjected to DNAse treatment to remove DNA contamination, since the presence of DNA could bias sequencing signals later. Ribosomal fractions are highly abundant and account for 90-95 % of the total RNA (Lindberg and Lundeberg, 2010). Because rRNAs are not targets for gene expression, the rRNA fraction was depleted by hybridizing the total RNA with Locked Nucleic Acid (LNA) probes designed specifically for abundant rRNA

molecules (RiboMinus™,Invitrogen). Depletion of abundant rRNA fractions was confirmed by Agilent Bioanalyzer 2100 **(Figure 2.1C)**. Ribosome-depleted RNAs were then chemically fragmented (see methods) and isolated using a method optimised to include the small RNA fraction. The RNA was then subjected to 5' phosphorylation to facilitate ligation of strand-specific adaptors. These adaptors were designed such that the 5' adapters bound specifically at phosphorylated 5' ends and 3' adaptor bound at non-phosphorylated 3' ends, thereby conserving information about the original orientation of transcripts. 5' and 3' ligated RNA was then reverse transcribed for cDNA synthesis and amplified. The final library was purified using optimised AMPure beads ratio (1: 10µl) to obtain a cDNA library ranging in size from 75 to 200 bp (insert size from 20 to180 bp).

To investigate the effectiveness of rRNA depletion, libraries were prepared after either one or two rounds of rRNA depletion and cloned into the Zero Blunt TOPO vector. Eighty plasmid colonies from each library containing inserts were sequenced using standard Dye-terminator sequencing (Sanger et al., 1977) and showed a comparable reduction in the rRNA fraction with one round of rRNA depletion compared to two rounds **(Figure 2.1D).** However, in the interests of caution two rounds of rRNA depletion were used for this study.

Hypoxia and normoxia libraries were then forwarded to the Genomic Core services for sequencing using Illumina GAIIX (single end, 51 bp length reads).

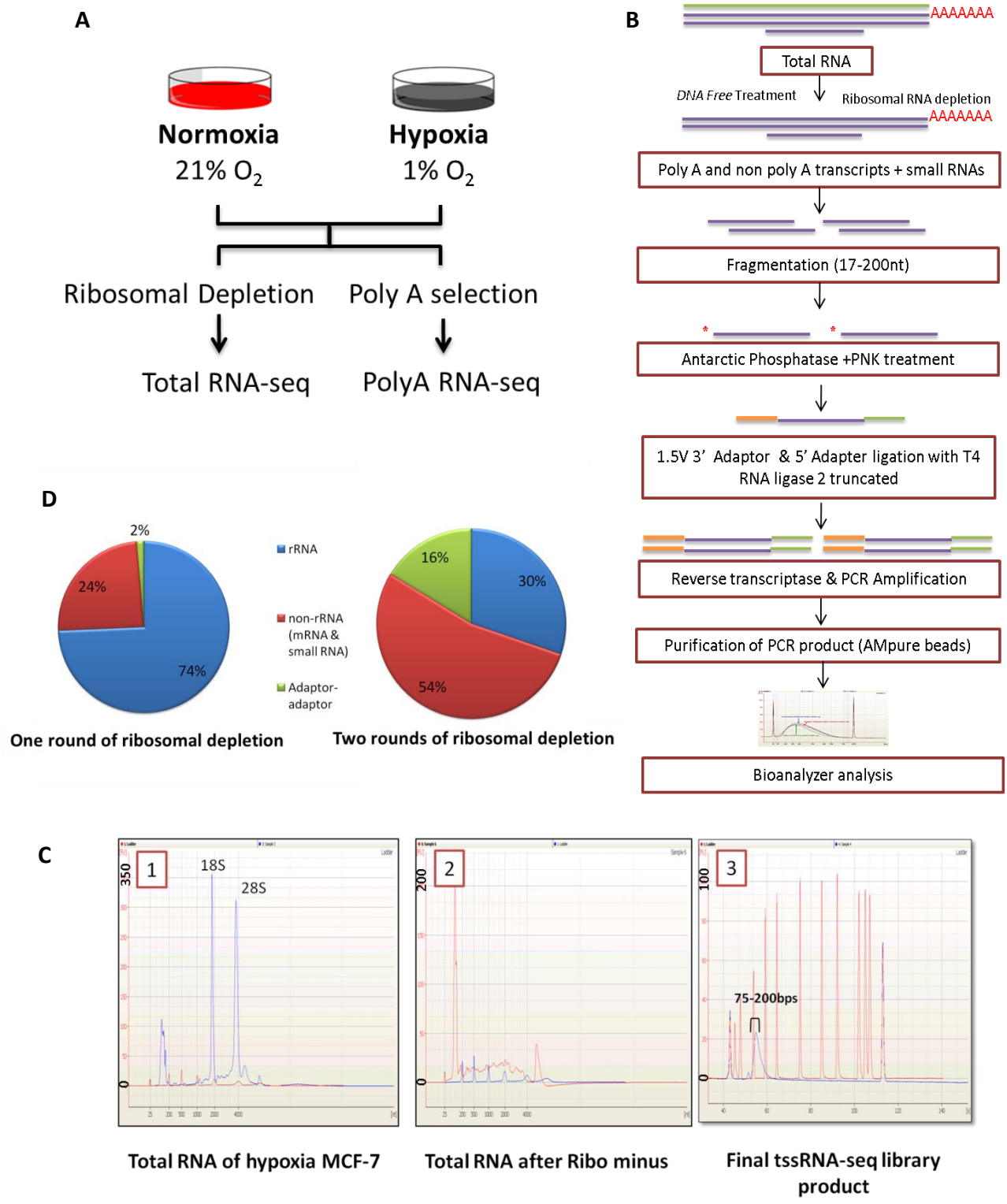


Figure 2.1 Development of total RNA-seq method. A) Total RNA was isolated from MCF-7 cells after 24 hours of culture in 21 % O_2 (normoxia) or 1 % O_2 (hypoxia). The total RNA was subjected to either ribosome-depletion or polyA selection and subjected to high-throughput sequencing. **B)** Flowchart of the total RNA-seq procedure. **C)** Agilent bioanalyzer 2100 electropherogram. (1) Total RNA from hypoxic MCF-7 cells before and (2) after total RNA after ribosomal depletion (3) the final purified total RNA-seq library with product size (75-200 bp). **D)** Percentages of different inserts in total RNA-seq library with one and two rounds of ribosomal depletion. 80 colonies were sequenced using Sanger sequencing per library.

2.3.3 Development of bioinformatic analysis pipeline for total RNA-seq

Currently RNA-seq analysis tools and packages are designed specifically for either small RNA or for polyA RNA analyses. There are no general methods proposed for total RNA-seq analysis that unambiguously detect and calculate expression of small and long RNAs. In the absence of such methods a number of approaches have been evaluated for optimal RNA detection and expression analysis.

Initially, the raw sequencing data were subjected to adapter sequence trimming using the FASTX-tool kit (http://hannonlab.cshl.edu/fastx_toolkit/). After trimming, the data were aligned to the human genome assembly (HG19) using TopHat (v.1.2.0) (Trapnell et al., 2009). TopHat uses Bowtie (v.1.0.0) (Langmead et al., 2009) as an alignment engine and additionally preserves unaligned reads which are called segments. Both aligned and unaligned (segments) reads are processed independently. TopHat attempts to align the segments ranging from 100 bp to 100 kbp apart from one another and when several segments are aligned to the genome it infers and estimates the junction splice sites. This allows TopHat to build an index of splice sites in the transcriptome without *a priori* gene or splice site annotations.

Reads were mapped with the '–bowtie-n' option and the rest of the parameters were left as default. Each library was mapped twice by accepting reads that map either uniquely (unique mapping) or in at least 20 positions (multiple mapping).

After mapping normoxia and hypoxia sequencing data, the Cufflinks package (Trapnell et al., 2010) was used to assemble aligned reads into the human reference genome model (RefSeq) for quantifying the expression levels of genes. Cufflinks measures abundances in Fragments Per Kb of exon per Million fragments mapped (FPKM). Then, the Cuffdiff tool was used to calculate the differential expression between the normoxia and the hypoxia samples. The Cuffdiff uses a statistical model that evaluates the expression changes and assumes that the number of reads mapped to each transcript is proportional to their abundance.

Although the above analysis provided a list of hypoxia regulated genes, it has a number of limitations including a bias towards longer transcripts (including mRNA and lncRNAs) and the restricted number of genes in the RefSeq database. The use of cufflinks assembly was computationally demanding for the total RNA-seq data, especially because of the low read depth of the total RNA-seq. The Cufflinks assembly and Cuffdiff rely on RefSeq databases for annotating mapped reads into genes for expression analysis. This analysis is restricted to transcripts that are present in the RefSeq database. However, there are a number of regulatory RNA classes such piwiRNA and tRNAs that are not annotated in RefSeq. Moreover, the RefSeq database lacks the updated annotation of some regulatory RNAs, such as lncRNAs, sn/snoRNAs and antisense transcripts, which are continually emerging with large scale genomic projects such as the ENCODE project. Since Cufflinks measures the transcript abundances with FPKM which is normalised by the length of the transcript, this measurement is not ideal for expression analysis of total RNA-seq data which contain both small and long RNAs.

In an attempt to analyse the global transcriptome changes in hypoxia without being limited to the RefSeq annotations, I established an in house total RNA-seq mapping and differential expression analysis pipeline that allows analysis of all RNA types together.

The pipeline allowed sequential mapping of the raw sequencing data to specific RNA databases using Bowtie tool (v.1.0.0) (Langmead et al., 2009) and calculated the expression level by counting the number of mapped reads to a particular transcript (normalised to the total number of mapped reads in each sample). The analysis pipeline was initiated by trimming the adapter sequences from the raw dataset using the FASTX-tool kit. After trimming, the data were mapped successively across different RNA databases, starting with an rRNA database, to filter any ribosomal fractions, and then mapping successively to databases of RNA classes of progressively increasing length. The flowchart of the analysis pipeline is explained in **Figure 2.2.** The first round of mapping used stringent criteria (one mismatch in seed length). Reads unmapped after

the first round were then re-mapped using less stringent criteria (two mismatches in seed length).

The total reads mapping to each transcript were then counted in each condition and used to calculate the differential expression. Finally, any unmapped reads (after the two rounds) were used to construct a novel transcript database by comparing and overlapping clusters of tags with and without genome annotations using the Cuffcompare tool from the Cufflinks package (Trapnell et al., 2010). Using default settings of Cufflinks, unmapped reads were *de novo* assembled in a set of transcripts. These transcripts were filtered with abundance of less than 13 % of the major transcript in the locus. Then, the Cuffcompare utility was used to compare the *de novo* assembly with the RefSeq database and to merge the transcripts into a combined set of annotations. The 5' and 3' exons of transcripts were allowed to vary by up to 200 bp during the comparison process. These analyses identified sets of novel transcripts, which were then classified as Novel Intergenic Transcripts (NITs) when there was no overlap with any annotated transcript and as Novel Antisense Transcripts (NATs) when they overlapped with a known transcript but were transcribed from the opposite DNA strand.

The databases of different RNA classes used in the pipeline were the human protein coding mRNA database (University of California, Santa Cruz "UCSC" genome browser), rRNA database (from Linda Hughes at Wellcome Trust bioinformatics group), piwiRNA database (National Center for Biotechnology Information "NCBI"), microRNA database (mirBase), sn/snoRNA database (UCSC genome browser), tRNA database (UCSC genome browser) and long non coding RNAs (ENCOD). All the databases were based on the 2009 release (HG19) of the annotated human genome and contained chromosomal locations and sequences.

Two technical replicates of the hypoxic and normoxic conditions were sequenced and were combined for analysis to increase depth and coverage of the transcriptome. Excluding reads mapping to ribosomal RNAs, 22,842,080 reads mapped to 42,989 transcripts in normoxia and 21,171,641 reads mapped to 43,169 transcripts in hypoxia. Reads mapped to all categories of

RNA in both normoxic and hypoxic MCF-7 cells. In addition, approximately 100 novel (intergenic and antisense) transcripts were identified.

The pipeline was tested with different alignment variables and read thresholds for optimised mapping and reliable expression analysis, including unique mapping, alignment with different number of mismatches in seed regions, initial mapping with different RNA sizes (longer to smaller RNAs and vice versa), mapping with/without rRNA filtrations, different normalisation methods (total raw sequencing normalisation, total mapped normalisation, specific class read normalisation) and different cut-off of mapped reads in expression analysis. All the steps were rigorously tested with statistical and computational analyses. Manual inspection of transcript regulation was performed by visualising the data on the UCSC genome custom tracks.

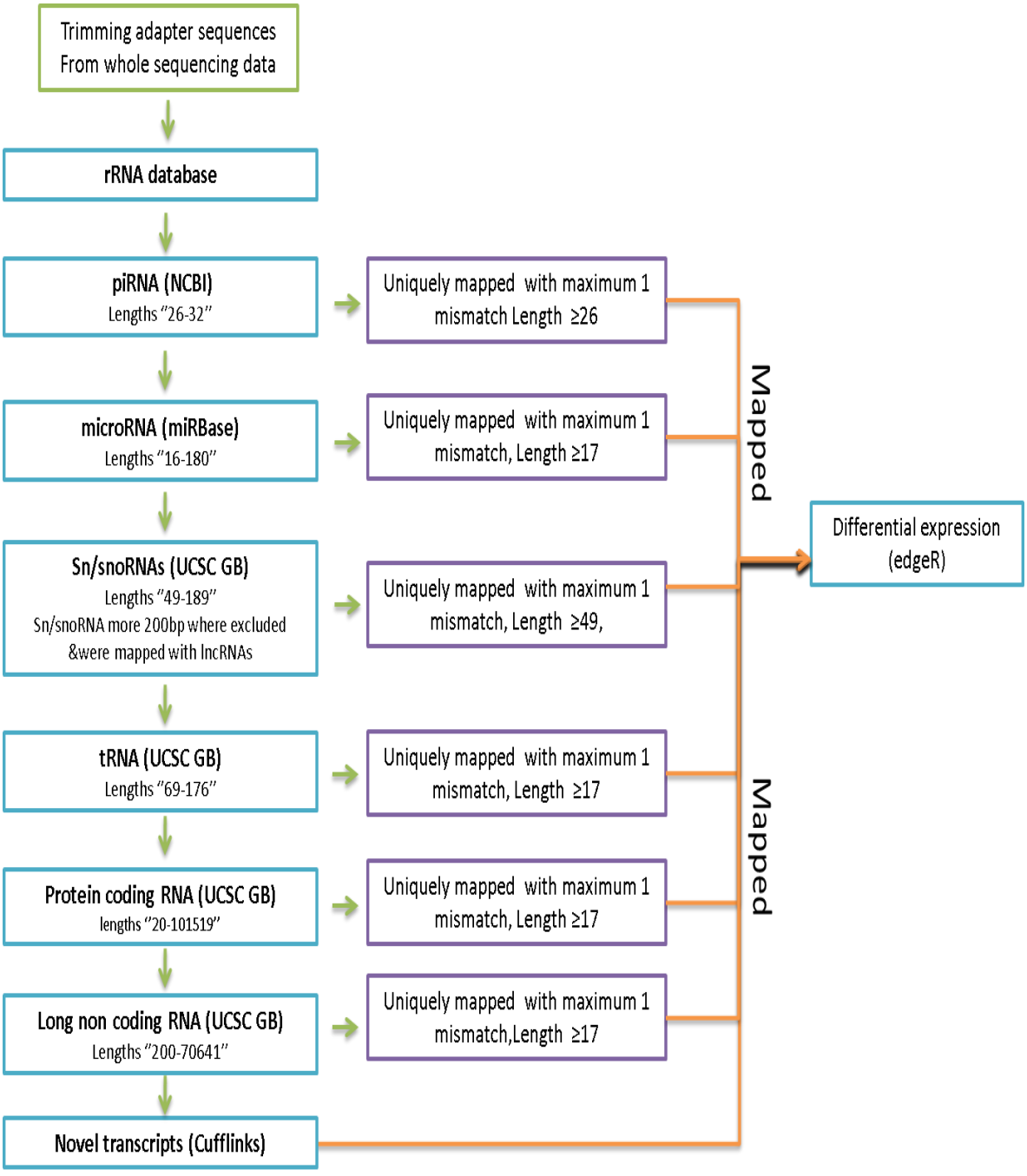


Figure 2.2 Comprehensive profiling of hypoxia transcriptome landscape. Flowchart of the pipeline used for ribosome-depleted RNA-seq analysis. The adaptor sequences were "trimmed" from the raw sequencing data and the resulting reads were mapped successively to databases of different RNA classes (using high-stringency criteria, followed by a second round of mapping using slightly lower-stringency criteria). After normalisation, the mapped data was used to calculate the differential expression for each RNA species.

2.3.4 Correlation between polyA RNA-seq and ribosome-depleted RNA-seq

The main advantage of ribosome-depleted RNA-seq is that it allows both polyA and non poly-A transcripts to be analysed simultaneously. However, many of these non poly-A transcripts, such as histone mRNAs and lncRNAs, are very abundant and account for a high proportion of sequencing reads in the total RNA-seq. The sequencing depth across the polyA RNAs may be less in total RNA-seq compare to polyA selected RNA-seq. Therefore, to compare the two methods directly, polyA selected cDNA libraries (normoxia and hypoxia) were sequenced and the transcript abundance and fold-regulation was compared for those RNA-species present in both datasets.

The analysis of the polyA selected normoxia and hypoxia datasets was performed similarly to the total RNA-seq by using the same *in-house* pipeline. For transcripts present in both the polyA selected and the ribosome-depleted RNA-seq datasets, a high degree of correlation was observed between expression levels in normoxia (R^2= 0.701) and hypoxia (R^2=0.715) in the two methods. In addition, these methods showed some degree of correlation when filtered for transcripts with at least 1.5 $\log_2$ fold regulation **(Figure 2.3).** These correlations validate the ribosome-depleted RNA-seq methodology (at least with respect to polyA RNA species) and provide a basis for extrapolation to non poly-A species of RNA.

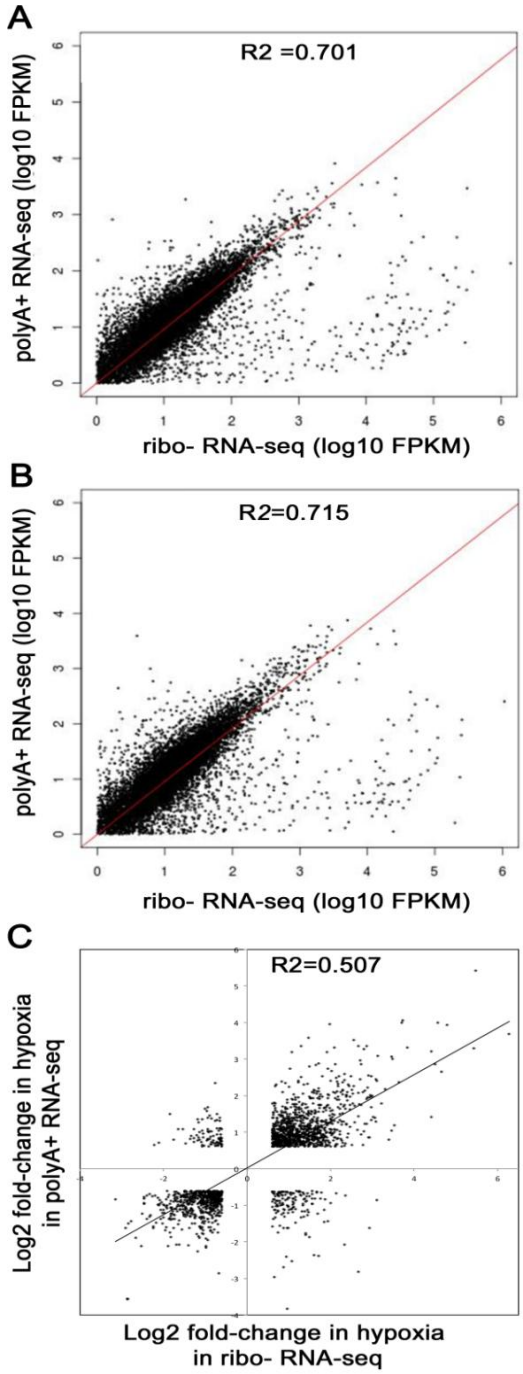


Figure 2.3 Correlation between polyA RNA-seq and total-RNA-seq. Using transcripts present in both PolyA RNA-seq and total-RNA-seq datasets, the expression level assessed by polyA RNA-seq was plotted against the transcript level in the total-RNA-seq dataset for **A)** normoxic and **B)** hypoxic samples. Linear regression shows a high degree of correlation between the two types of analysis. (FPKM is fragments per kilobase of exon per million fragments mapped) **C)** Transcripts were then filtered for those showing at least 1.5 fold regulation by hypoxia in both the polyA and total- RNA-seq datasets and the two values plotted against each other. Log_2 fold-change in hypoxia again correlated well between the two types of analysis.

2.3.5 Extensive regulation of the coding and non-coding transcriptome by hypoxia

To determine the hypoxic regulation of each RNA across all classes, the fold-change in expression after hypoxic exposure was calculated as the normalised ratio of reads mapping to the relevant locus in hypoxia, to reads mapping to the sample locus in normoxia. For each expressed RNA, the log_2-fold-change (log_2 FC) for was plotted against the log10-expression level and the data is presented by class of RNA **(Figure 2.4A)**. At low expression levels the scatter associated with log_2-fold-change increased indicating that the estimation of hypoxia regulation was poor and potentially unreliable at low transcript abundance. These low-abundance species were therefore filtered by applying thresholds to the read number for each transcript (10 reads per transcript for small RNA classes - piwiRNA, miRNA, sn/snoRNA and tRNA, and 100 reads per transcript for mRNAs and lncRNAs). Using data from these filtered transcripts the median fold-change in hypoxia, together with the interquartile and full ranges, were calculated for each class of RNA **(Figure 2.4B)**.

This analysis demonstrates that all classes of RNA are regulated by hypoxia. The median fold-change in hypoxia for each class of transcript differed significantly between the different classes of RNA ($p<0.001$, Kruskal-Wallis), with pairwise analysis showing statistically significant differences between all groups except between mRNAs and miRNAs ($p<0.001$, Mann-Whitney U test). Both piwiRNAs and tRNAs showed slight downregulation (median log_2 FC = -0.72 and -0.90, respectively), whilst snRNAs exhibited more profound downregulation in hypoxia (median log_2 FC = -2.05). miRNAs and pcRNAs showed modest upregulation (median log_2 FC = 0.48 and 0.47 respectively). lncRNAs exhibited a slightly greater fold upregulation (median log_2 FC = 0.81) **(Figure 2.4C).**

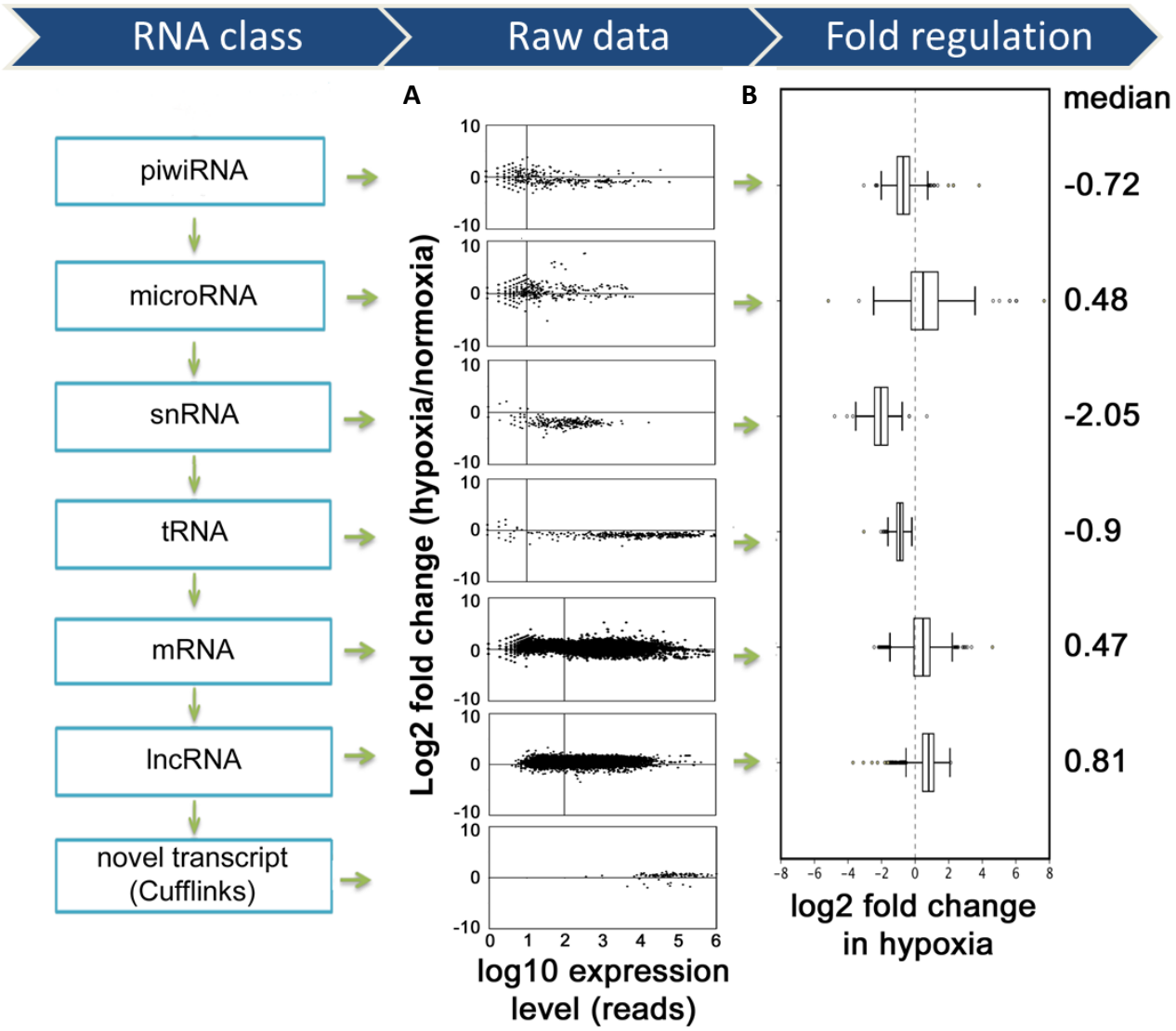


Figure 2.4 Extensive regulation of transcript abundance by hypoxia. A) Raw data for each RNA class plotted as $\log_2$ fold-change by hypoxia on the vertical axis versus $\log_{10}$ expression level on the horizontal axis. Dotted vertical lines denote thresholds for the selection of transcripts for further analysis. For small RNA (piwiRNA, miRNA, snRNA and tRNA) a threshold of 10 reads and for mRNA and lncRNA a threshold of 100 reads in at least one condition were used. **B)** Box-and-whisker plots of $\log_2$ fold-change by hypoxia for filtered transcripts in each class of RNA. The vertical dotted line denotes no fold-change. Median regulation for each class in hypoxia is shown. Two technical replicates of the hypoxic and normoxic conditions were sequenced and were combined for analysis to increase depth and coverage of the transcriptome.

To confirm these findings, representative samples of snRNA, pcRNA and lncRNA transcripts (n= 5) exhibiting regulation close to the median values were tested by quantitative real time polymerase chain reaction (qPCR). This analysis recapitulated similar differences between these RNA classes, consistent with the finding of overall class-specific regulation **(Figure 2.5A)**.

To determine whether this hypoxic regulation was conferred by a 2-oxoglutarate (2-OG)-dependent dioxygenase enzyme such as those responsible for regulating the transcription factor HIF, MCF-7 cells were treated with the inhibitory 2-OG analogue DMOG for 24 hours and expression levels of the same representative snRNA, mRNA and lncRNA transcripts were again analysed by qPCR. CA9 expression significantly increased with DMOG and was used as a control to confirm effective DMOG treatment in mimicking hypoxic conditions. This analysis showed a comparable pattern of regulation to that observed in hypoxia for all three classes of RNA, with upregulation of mRNA and lncRNA transcripts and a reduction of snRNA transcripts **(Figure 2.5B).**

Since the HIF transcription factor is the archetypal example of an oxygen and 2-OG dependent pathway, I next determined the dependence of the hypoxia-dependent regulation of mRNAs, lncRNAs and snRNAs on the presence of HIF. MCF-7 cells were incubated in hypoxia, following treatment with either control siRNA or after combined suppression of HIF-1α and HIF-2α using siRNAs to both HIF-α species. Suppression of HIF-1α and HIF-2α was confirmed by western blotting **(Figure 2.5C)**. Hypoxic mRNA and lncRNA expression was reduced while expression of snRNAs was slightly increased **(Figure 2.5D)** indicating a degree of HIF-dependence**.**

In addition to global changes in hypoxia, a number of transcripts in each category showed strong upregulation or downregulation, ranging from 15-fold up to 8-fold down (piwiRNA), >100-fold up to 30-fold down (miRNA) and 4-fold up to 12-fold down (lncRNA). The top five most regulated RNAs from each class are listed in **Supplementary table 1**.

Many differentially expressed mRNAs identified in the total RNA-seq have been previously reported to be hypoxia regulated in MCF-7 cells (Elvidge et al., 2006). For example, protein coding genes N-myc downstream regulated 1 (NDRG1), adrenomedullin 3 (BNIP3), vascular endothelial growth factor (VEGF) and carbonic anhydrase IX (CA9) were significantly upregulated in hypoxia with 4.2, 2.3, 2.1, and 2 (log_2) folds, respectively in this study and were also reported in previous hypoxia expression studies (Koop et al., 2009, Liu et al., 1995, Chi et al., 2006, Turner et al., 2002, Said et al., 2008). This high level of concordance provides further validation of the current dataset.

Microarray technology tends to underestimate fold-change in gene expression when compared to RNA-seq and qPCR. It is therefore not surprising that the current analysis has identified additional protein coding genes that were significantly upregulated but have not been previously reported. These include Thymosin Beta 10, TMSB10 (3.38), poliovirus receptor-related 4, PVRL4 (3.29), RAB11 family interacting protein, RAB11FIP5 (3.24), Stonin 1, STON1 (3.24), branched chain ketoacid dehydrogenase kinase, BCKDK (2.8), reticulocalbin 3, EF-hand calcium binding domain, RCN3 (2.75), BCL6 corepressor BCOR (2.4) and V-set domain containing T cell activation inhibitor 1 VTCN1 (2.37) (Fold change). Induction of these novel hypoxically regulated protein genes was confirmed in MCF-7 cells by qPCR **(Figure 2.5E).**

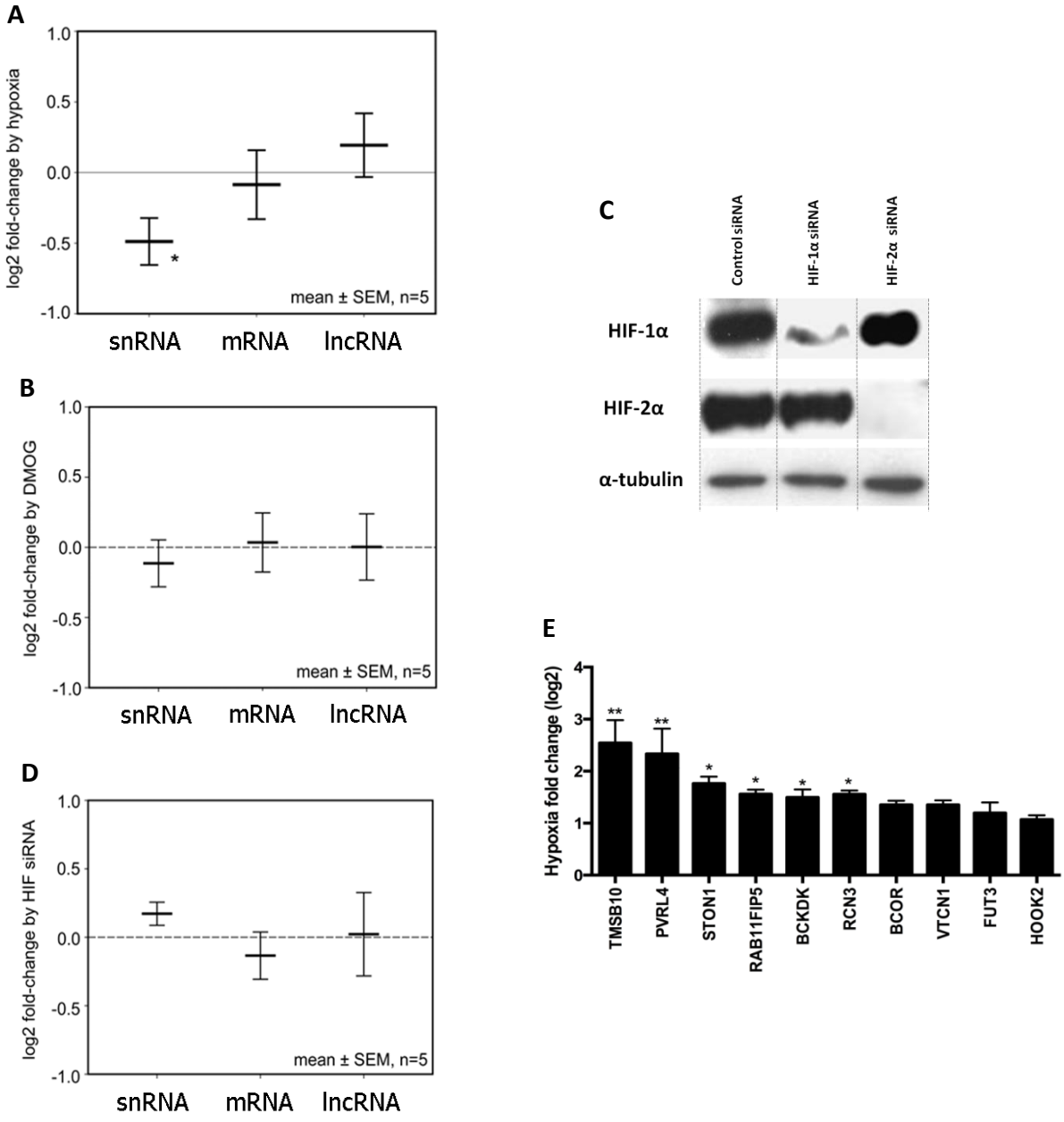


Figure 2.5 qPCR validation of global changes in transcript abundance in hypoxia. For each of the RNA classes shown, five transcripts exhibiting hypoxic regulation close to the median value were assayed by TaqMan gene expression assays. **A)** In hypoxia, five snRNAs showed significant downregulation, whilst the mean changes of lncRNAs was increased and mRNAs was unchanged, which is similar the pipeline analysis. **B)**DMOG treatment, cells were treated with (500 uM) DMOG for 24 hours in normoxia, snRNAs were marginally downregulated, whilst mRNA and lncRNAs did not show significant change. **C)** Immunoblot of protein extract obtained from hypoxic MCF7 cells (1 % O_2, 24 hours) treated with scrambled control siRNA or with siRNA targeting HIF-1α or HIF-2α. Substantial suppression of each HIF-α protein was seen with each HIF-specific siRNA treatment. **D)** HIF knockdown, cells were transfected with both HIF-1α and HIF-2α siRNAs in hypoxic condition, mean changes of snRNA was increased, mRNA expression was reduced and lncRNA was not affected. All the experiments were performed with three biological replicates for each of the five (n=5) transcripts in each class. snRNAs: SNO38, SNO117, SNO60, SNO21, SNO66. mRNAs: TCF25, C15orf42, DHX30, ZNF320, RBAK. LncRNAs: TCONS_l2_00027397, TCONS_00019713, TCONS_l2_00016774, TCONS_00020202, TCONS_00000020. **E)** Validation of previously unreported hypoxia-regulated protein coding genes in MCF-7 cells by qPCR (* $p<0.05$, **$P\leq0.005$, Student's t-test, n=3).

2.3.6 Small RNA (<200 bp) regulation in hypoxia

Importantly, several small RNAs classes (snRNAs, tRNAs and piwiRNAs) demonstrated overall downregulation, while miRNA showed general upregulation in response to hypoxia. Regulation of snRNAs, tRNAs and piwiRNAs in hypoxia has not been studied before.

2.3.6.1 Hypoxic regulation of snRNAs

snRNAs have a key role in modification and processing of rRNA (Matera et al., 2007). rRNAs undergo several modifications during their synthesis. These include 2'-O-ribose methylation and pseudouridylation (Jady and Kiss, 2000). Studies have reported that snRNA play an essential role in assembling ribonucleoprotein partials that are essential elements of the spliceosome in cells. These ribonucleoprotein partials contain modifying enzymes, such as methyltransferase and pseudouridine synthase, which direct the appropriate rRNA sequence modifications (Jady and Kiss, 2001, Filipowicz and Pogacic, 2002). The majority of snRNAs are of two types: C/D snRNA and H/ACA snRNAs. In most cases, these two types are not independently transcribed but processed differently from the pre-mRNA (Morrissey and Tollervey, 1995).

Several proteins are essential for snRNA processing and maturation. Fibrillarin (methyltransferase), NOP56 ribonucleoprotein and NOP58 ribonucleoprotein form a protein complex for maturation of C/D snRNAs, while proteins including dyskerin (DKC1), GAR1 ribonucleoprotein, NHP2 ribonucleoprotein, and NOP10 ribonucleoprotein are required for H/ACA snRNA maturation. Moreover, an RNA-binding protein NAF1 (nuclear assembly factor 1) is required to ensure a proper accumulation of H/ACA box snRNAs (Lafontaine and Tollervey, 1998, Terns and Terns, 2002).

Because snRNAs are downregulated in hypoxia, I postulated that generic pathways leading to snRNA biogenesis and maturation could be affected or altered in hypoxia. Indeed, I found that genes encoding proteins, which are required for C/D snRNA and H/ACA snRNA maturation, are suppressed by hypoxia **(Figure 2.6A),** indicating that downregulation of snRNAs may indeed result from altered maturation in hypoxia.

Other studies have shown that aberrant expression of snRNAs in cancers can lead to changes in cell growth and viability and may be used as potential biomarkers (Mannoor et al., 2012, Chen and Zhu, 2013, Dong et al., 2009). Gee et al. reported that low expression of the snRNA RNU44 is correlated with aggressive tumour pathology and poor prognosis in breast cancer and in head and neck squamous cell carcinoma (Gee et al., 2011). I found that RNU44 expression was significantly downregulated -1.8-fold in response to hypoxia. RNU44 is an intronic gene in the growth arrest specific 5 (GAS5) transcript, which is normally upregulated to arrest cell growth under stress (Renganathan et al., 2014, Pickard and Williams, 2014). Its downregulation in hypoxic MCF-7 breast cancer cells may therefore contribute directly to the aggressive phenotype seen in hypoxic breast tumours.

2.3.6.2 Hypoxic regulation of tRNAs

All tRNAs are downregulated in hypoxia with median expression $\log_2$=-0.90. tRNAs are transcribed by RNA polymerase III and their maturation and synthesis involves several processing enzymes, such as ribonuclease P (RNase P) and tRNA nucleotidyl transferase. The promoter regions of tRNAs consist of two separated elements, known as Box A and Box B, located downstream of the TSS (Kassavetis et al., 2005). The General Transcription Factor-IIIC (TFIIIC) binds to Box B; this binding allows correct positioning of General Transcription Factor-IIIB (TFIIIB), also known as Pre-Initiation Complex, which then recruits RNA Pol-III. The DNA, TFIIIB and Pol III form a very stable complex, which undergoes many rounds of tRNA

transcription. Transcription and elongation start immediately after assembly of the Initiation Complex (Andrau and Werner, 2001, Schroder et al., 2003).

Ernens et al. studied the expression of $tRNA^{Leu}$ and $tRNA^{Tyr}$ in hypoxia and saw a significant decrease in levels (Ernens et al., 2006). They also reported a 40 % reduction in levels of Pol III transcripts derived from the middle repetitive gene family B2 in hypoxia (Ernens et al., 2006). They showed that the inhibition of RNA pol III transcription is due to altered interaction between TFIIIB and its regulators c-Myc in hypoxia, which then limits RNA polymerase III recruitment. Both $tRNA^{Leu}$ and $tRNA^{Tyr}$ were also significantly downregulated in the current study.

Since RNApol III is suppressed in hypoxia, I hypothesized that the overall downregulation of tRNAs observed in hypoxia is due to deregulation of the tRNA synthesis and processing machinery. To test this hypothesis, I examined the expression of genes that are involved in tRNA processing and synthesis. These genes include subunits of RNA polymerase III (POLR3A-K), RNA polymerase III transcription initiation factors (BDP1, BRF1, BRF2), tRNA splicing endonucleases (TSEN2,34,54) and general transcription factor IIIC (GTF3C1-5), which are required for RNA polymerase III-mediated transcription and tRNA processing enzymes. A large proportion of these tRNA synthesis and processing genes were downregulated in hypoxia. Further, the hypoxic expression of many of these genes was restored following HIF-1α and/or HIF-2α knockdown by siRNA **(Figure 2.6B).** Thus, I concluded that downregulation of tRNA in hypoxia is due (at least in part) to inhibition of its synthetic and processing pathways by HIF-dependent mechanisms.

2.3.6.3 Hypoxic regulation of microRNAs

MiRNAs generally showed slightly increased expression in hypoxia compared to normoxia, although there was wide variability (median expression log_2=0.48). A number of the most deregulated miRNAs in our dataset have been previously identified as hypoxia regulated, including the two most upregulated, mir-184 and mir-210. Indeed, consistent with my analysis, a recent comprehensive study in which small RNAs were profiled in hypoxic MCF-7 (1 % O_2) at different time points (16, 48, 32 hours) using small RNA-seq showed mir-184 and mir-210 were also the most upregulated miRNAs after 48H of hypoxia (Camps et al., 2014).

Among the most commonly described hypoxia-regulated miRNAs is mir-210, which is reported to be a direct transcriptional target of HIF in several cancers (Camps et al., 2008, Kulshreshtha et al., 2007). Mir-210 was highly upregulated (3.4 fold) in the current dataset. Several studies reported that mir-210 acts by targeting and regulating expression of key genes that are involved in pathways such as DNA repair, apoptosis, metabolism and cell cycle and contributes to an aggressive tumour phenotype (Favaro et al., 2010, Chan and Loscalzo, 2010, Zhang et al., 2009). A recent study has shown that mir-210 decreases Krebs cycle activity and mitochondrial function by targeting and supressing iron- sulphur cluster scaffold homolog (ISCU), leading to an increase in free radicals and cell survival in hypoxia (Favaro et al., 2010). Consistent with this, I found that expression of ISCU was 1.5-fold downregulated in the total RNA-seq analysis.

In addition, I have undertaken a comparison of miRNA regulation in this study with the findings from five previous hypoxia microarray based studies (Kulshreshtha et al., 2007, Hua et al., 2006, Hebert et al., 2007, Donker et al., 2007, Guimbellot et al., 2009) **(Figure 2.7)**. Despite significant heterogeneity between the published reports, the degree of overlap between the current analysis and microarray-based approaches is comparable to that between these previous studies. In particular both mir-210 (described in 5/ 5 previous studies) and mir-27 (described in 2/5 previous studies) were upregulated in the current study **(Figure 2.7).**

Although a number of previously reported hypoxia regulated miRNAs have been described, several new hypoxia regulated miRNAs were also identified in this study. The expression of the most hypoxic regulated miRNAs in the current study (miR-184, miR-1, miR-612, miR-100 and miR-210) was validated in MCF-7cells and in a panel of breast cancer cell lines (MDA-MB-231, T47D, SKBR3 and BT474) grown in hypoxia (1 % O_2, 24 hours). All miRNAs tested were upregulated in MCF-7 cells. Each miRNA was also upregulated in between 2 and 4 of the four additional cell lines **(Figure 2.6C).**

These newly identified hypoxia-regulated miRNAs are suggested to have roles in cancer. For example, overexpression of miR-184 is reported to play an oncogenic role in several cancers including squamous cell carcinoma (SCC), neuroblastoma and head and neck cancer (Liu et al., 2011, Tivnan et al., 2010, Wong et al., 2009). It is reported that miR-184 acts as an antiapoptotic molecule and promotes proliferative processes by regulating c-Myc expression (Foley et al., 2010, Wong et al., 2008), while miR-1 and miR-100 have both oncogenic and tumour suppressing abilities in multiple cancer types and function in a context-dependent manner (Singh et al., 2013, Tominaga et al., 2013, Leite et al., 2013, Wang et al., 2013). Regulation of these miRNAs in hypoxia could provide new insights into post-transcriptional regulatory mechanisms.

2.3.6.4 Hypoxic regulation of piwiRNAs

The role and function piwiRNAs is not well characterised, however, studies have suggested a potential role in silencing DNA elements and the maintenance of the DNA integrity (Luteijn and Ketting, 2013, Ishizu et al., 2012). These RNAs require argonaute proteins for their function and maturation, in particular Argonaute RISC Catalytic Component 3 (AGO3), which is essential for their biogenesis (Nagao et al., 2010, Thomson and Lin, 2009). Like other class of RNAs, piwiRNAs are also aberrantly expressed in human cancers (Cheng et al., 2012, Cheng et al., 2011). I found AGO2, which is important for small RNA function, and AGO3 were -1.2 and -1.1 (log_2) fold downregulated in hypoxia, respectively. This may contribute to the downregulation of piwiRNAs in hypoxia.

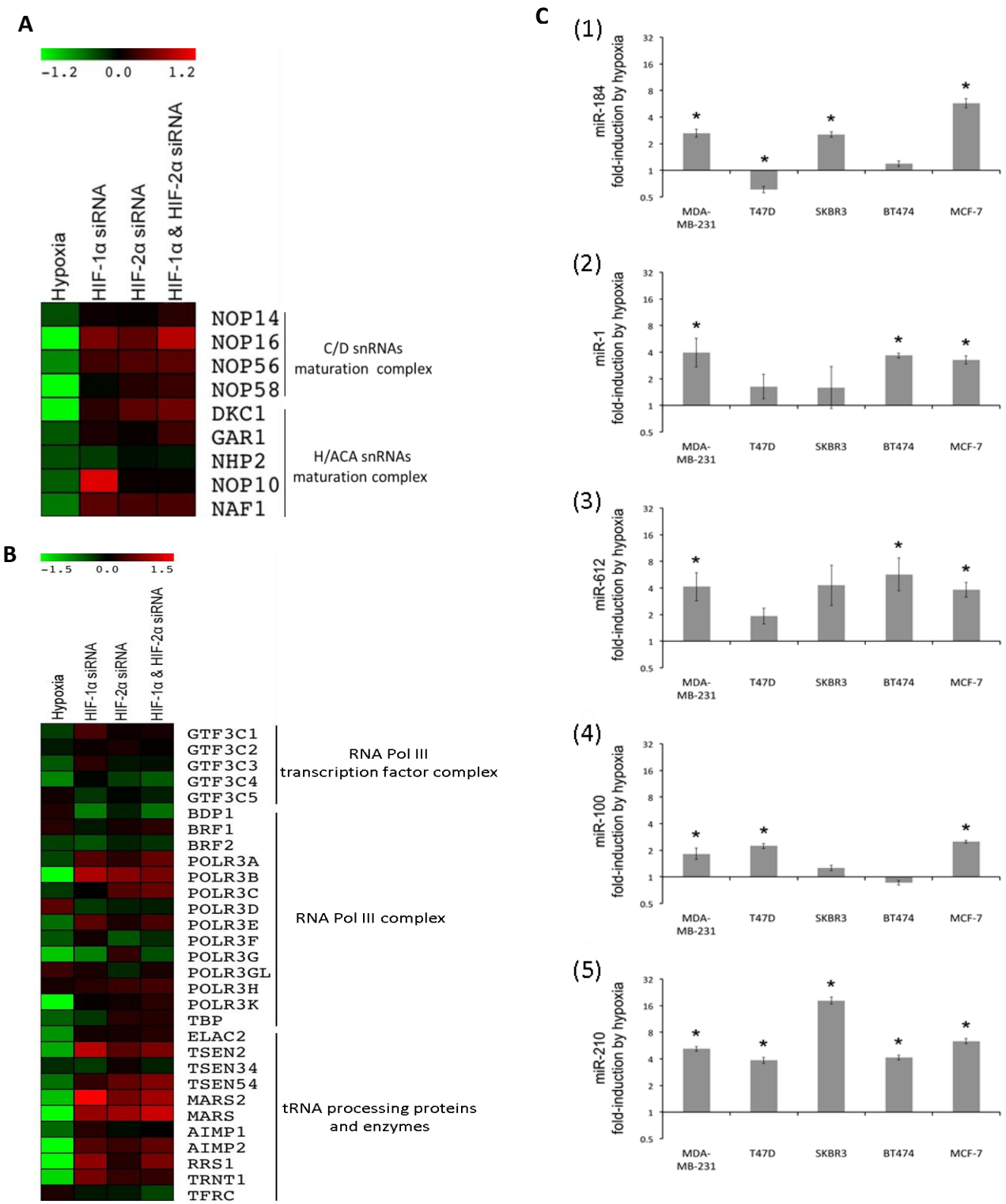


Figure 2.6 Regulation of small RNAs in hypoxia. A) Heat map of mRNA expression which encode C/D snRNA and H/ACA snRNA maturation proteins in $\log_2$ fold-changes in response to hypoxia and HIF-α siRNAs. **B)** Heat map of mRNA expression which encode tRNA processing proteins and RNA Pol III complex in $\log_2$ fold-changes in response to hypoxia and HIF-α siRNA. **C)** qPCR analysis of most upregulated hypoxic regulated miRNAs (1) miR-184, (2) miR-1, (3) miR-612, (4) miR-100 and (5) miR-210 in breast cancer cell lines (MCF-7, MDA-MB-231, T47D, SKBR3 and BT474). Bars represent fold-change after 24 hours in 1 % O_2 hypoxia ± SEM (* $p < 0.02$, Student's t-test, n=3).

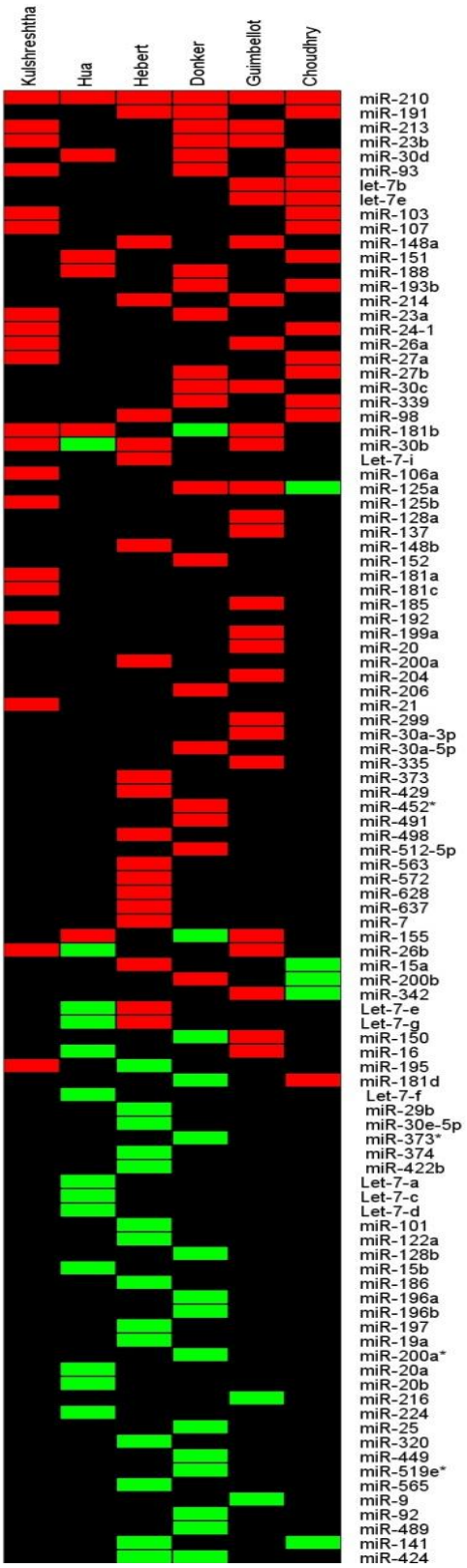


Figure 2.7 Comparison of hypoxia regulated miRNAs with previous reports. Heat map indicating miRNA regulation between studies (Upregulated in red, downregulated in green, black is non-regulated or not expressed). Studies: **Kulshreshtha et al. (Kulshreshtha et al., 2007)** - colon and breast cancer cells, 0.2 % O_2, 8–48 h; **Hua et al. (Hua et al., 2006)** - nasopharyngeal carcinoma cells, DFOM treatment, 20 h; **Hebert et al. (Hebert et al., 2007)** - head and neck squamous carcinoma cells, 1 % O_2, 1 h or 5 % O_2, 8 h; **Donker et al. (Donker et al., 2007)** - primary human cytotrophoblasts, 1 % O_2, 48 h; **Guimbellot et al. (Guimbellot et al., 2009)** - colon cells, liquid–liquid interface.

2.3.7 Long RNA (>200 bp) regulation in hypoxia

2.3.7.1 Regulation of non-polyA transcripts in hypoxia

The standard polyA-selected RNA-seq library preparation begins with positive selection of polyA RNAs, whereas the total RNA-seq method isolates both polyA and non-polyA transcripts. In this study, a number of replication-dependent histone non-polyA mRNAs including HIST1H1D, HIST1H3J, HIST1H3C and HIST1H3I were detected and found to be downregulated in hypoxia. These transcripts were only detected in the total RNA-seq analysis, and not in the polyA-selected RNA-seq dataset **(Figure 2.8 A-B).** Expression of specific non-polyA histone mRNAs (HIST1H1D, HIST1H3J and HIST1H3C) was validated by qPCR in MCF-7 cells and in a panel of breast cancer cell lines. This analysis revealed that the expression of all three non-polyA mRNAs was suppressed by hypoxia in the majority of breast cancer cell lines including ZR-571, T47D, MCF-7 and BT474 **(Figure 2.8C).** Conversely, SKBR3 cells (ER-, HER++) showed significant upregulation of all three non-polyA histone mRNAs. In conclusion regulation of non-polyA histone mRNAs by hypoxia is a frequent occurrence and is likely to have an important role in epigenetic and transcriptional regulation.

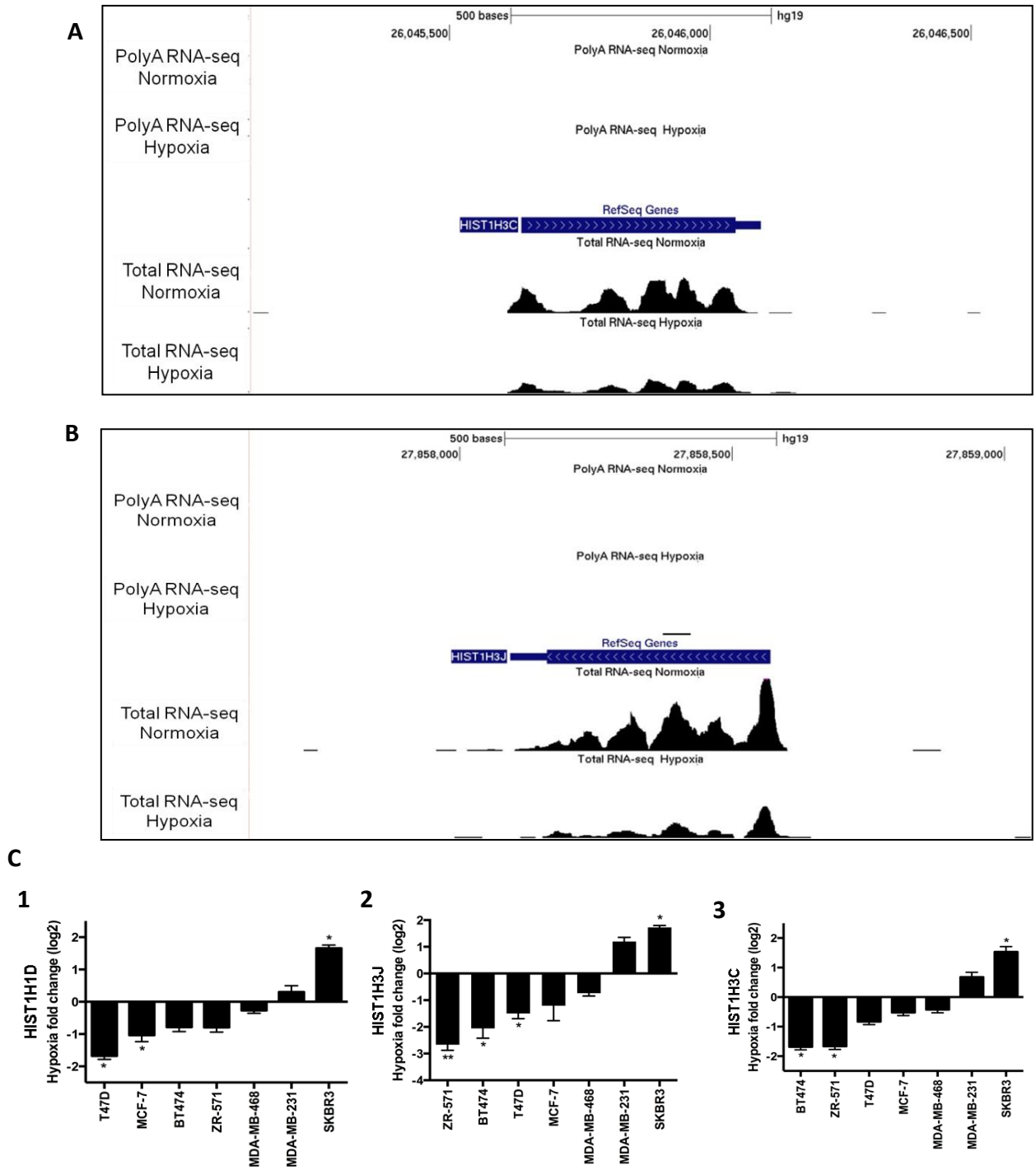


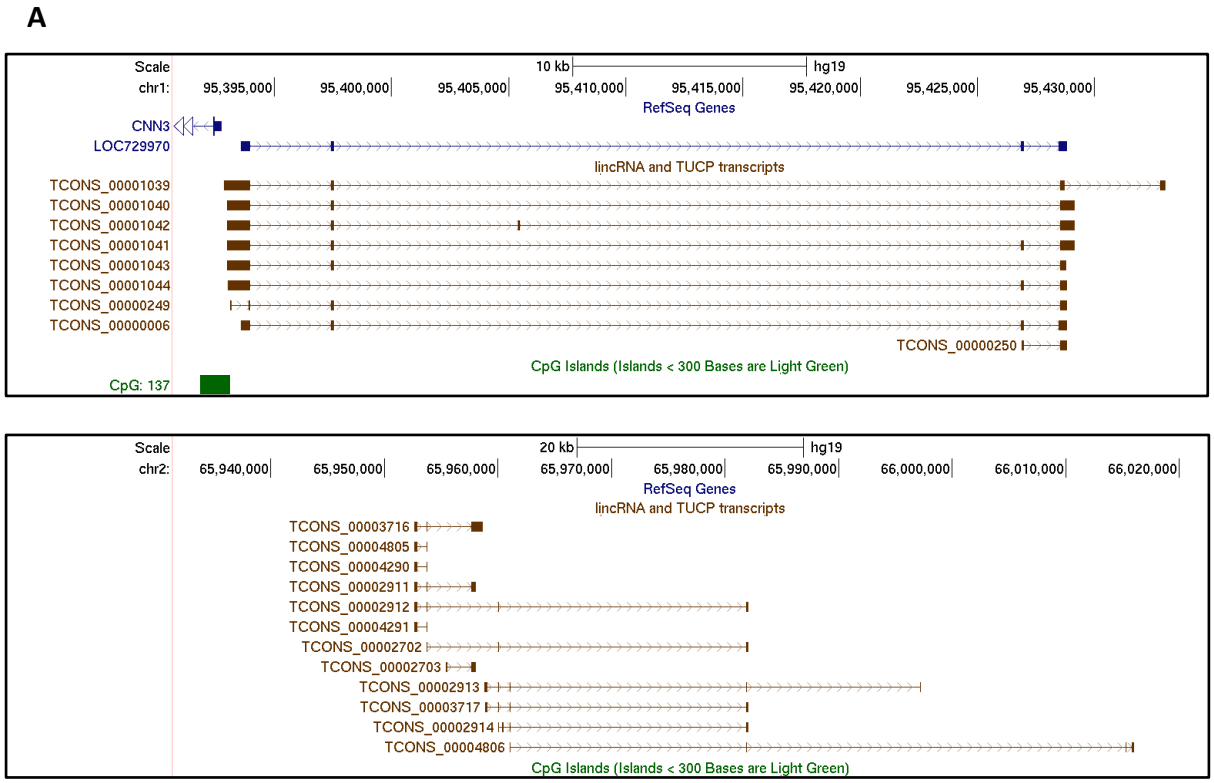
Figure 2.8 Hypoxic regulation of non-polyA mRNAs. Snap-shot of the UCSC genome browser (http://genome.ucsc.edu/) at non polyA histone mRNA loci, **A)** HIST1H3C and **B)** HIST1H3J in polyA RNA-seq (top tracks) and total RNA-seq (bottom tracks) in MCF-7 cells. Histone non polyA mRNAs were only detected in total RNA-seq dataset and were downregulated in hypoxia. **C)** Expression non-polyA histone mRNAs (1) HIST1H1D, (2) HIST1H3J and (3) HIST1H3C were validated in a panel of breast cancer cell lines grown in hypoxia. Non-polyA histone mRNAs were commonly downregulated in hypoxic breast cancer, with the exception of SKBR3 and MDA-MB-231. (* $p<0.05$, Student's t-test, n=3).

2.3.7.2 Regulation of lncRNAs by hypoxia

The recently released ENCODE (v.17) lncRNA database, containing 21,630 lncRNA transcripts was used to map the sequencing datasets using an *in house* pipeline to identify hypoxia-regulated lncRNAs. However, the majority of these transcripts are not well characterised (a large numbers of transcripts are transcribed from overlapping loci **(Figure 2.9A)** and lack a well-defined TSS). This could potentially bias the mapping and expression analysis. Thus, I created an lncRNA-containing locus (LCL) database by combining the exons of lncRNA transcripts that are transcribed from the same overlapping region and in the same orientation, creating a meta lncRNA transcript database. This identified 6,437 distinct lncRNA containing loci with putative TSSs that were then used for the mapping and analysis. Of these loci, 6,235 expressed a lncRNA with a minimum of 100 reads in at least one condition (normoxia or hypoxia or both). Differential analysis demonstrated that lncRNAs are upregulated in hypoxia compared to other RNA classes (Median $\log_2$ fold-change = 0.81). Expression of the 10 most upregulated lncRNAs was validated by qPCR in MCF-7 cells and confirmed their regulation by hypoxia **(Figure 2.9B).**

Widespread hypoxic regulation of lncRNAs has not been described previously. The majority of lncRNA transcripts are functionally uncharacterised. However, the most upregulated lncRNA is NEAT1 (4.1-fold), which is an architectural component of interchromatin paraspeckles that regulate the expression of Adenosine to Inosine (A-to-I) edited RNAs through nuclear retention (Clemson et al., 2009, Chen and Carmichael, 2009). In addition, the second most upregulated lncRNA, metastasis associated lung adenocarcinoma transcript 1 (MALAT1) was induced by 3.6-fold in hypoxic cells. MALAT1 is frequently reported to be upregulated in solid tumours and is associated with metastasis and tumour recurrence (Gutschner et al., 2013, Xu et al., 2011). Several studies have shown that increased expression of MALAT1 enhances cellular proliferation *in vitro* and tumour growth *in vivo* (Han et al., 2013, Tripathi et al., 2013). It is suggested that MALAT1 may regulate pre-mRNA splicing factors including serine arginine dipeptide-containing SR family splicing factors (Tripathi et al., 2010, Polymenidou et al., 2011)

and may also regulate some key transcription factors, such as E2F1, in cancer (Yang et al., 2011b). Another lncRNA known as imprinted maternally expressed transcript (H19) was also upregulated (2.8-fold) in hypoxia. The H19 transcript has been reported to be hypoxia-regulated and has been shown to have oncogenic properties both *in vitro* and *in vivo* (Matouk et al., 2007, Yang et al., 2012). These workers reported that hypoxic induction of H19 only occurs in the absence of a functional p53, or in the presence of mutated p53 (Matouk et al., 2010). The hypoxic regulation of lncRNAs in cancer likely reflects a previously unknown aspect of hypoxia biology in cancer.



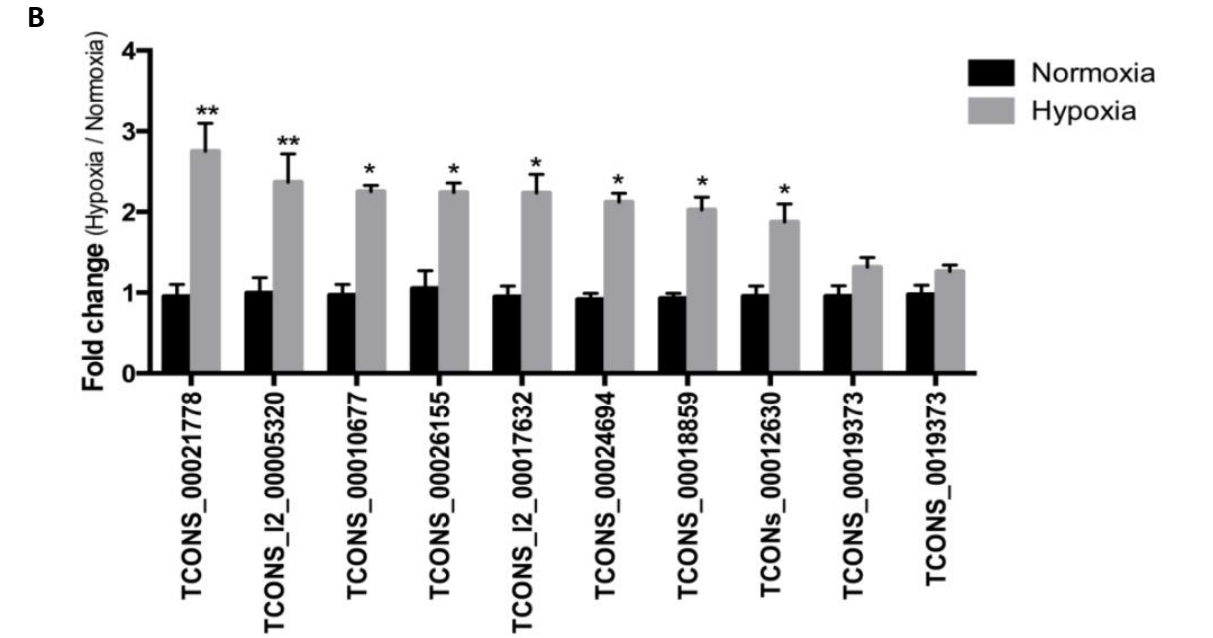


Figure 2.9 Hypoxia regulation of lncRNAs. A) Two examples of lncRNAs locus having multipliable transcript annotations with same transcriptional orientation. For analysis, exons of such kind of transcripts were combined to a single lncRNA to create lncRNA containing locus (LCL) database **B)** qPCR validation of top 10 hypoxia upregulated lncRNAs in MCF-7 cells in hypoxia. All the experiments were performed with three biological replicates (* $p<0.05$, ** $p<0.005$ Student's t-test, n=3).

2.3.8 Correlation between lncRNA and protein coding RNA transcription

lncRNAs can transcriptionally regulate the expression of protein coding genes either *in cis* (these effects are limited to the chromosome from which lncRNAs are transcribed) or *in trans* (these affect genes on other chromosomes) (Sotillo and Thomas-Tikhonenko, 2011, Ulitsky and Bartel, 2013). The process by which the transcription of lncRNAs mediates silencing or repression of other genes is known as 'transcriptional interference' (TI) (Hartzog and Martens, 2009, Shearwin et al., 2005). An important example of this process is silencing of the whole X-chromosome in *cis* by the Xist lncRNA (Herzing et al., 1997). Similarly, HOTTIP is another *cis*-acting lncRNA, which is expressed in the HOXA cluster and activates transcription of flanking genes (Wang et al., 2011).

To evaluate whether hypoxia-regulated lncRNAs might be acting *in cis* to regulate flanking protein coding genes in MCF-7 cells, the $\log_2$ fold-change in expression of each lncRNA was correlated with the $\log_2$ fold-change in expression of the closest expressed protein coding gene **(Figure 2.10A)**. Data were also censured for lncRNA-mRNA pairs that were within 1 kb of each other **(Figure 2.10B).** Most upregulated lncRNAs were associated with upregulated mRNAs, which likely reflects the observation that both classes of RNA showed slight upregulation in the analyses. Overall there was no discernable correlation between hypoxic regulation of a lncRNA and its nearest mRNA neighbour, nor was there any discernable change in scatter to suggest a pattern of both up and downregulation of neighbouring mRNAs.

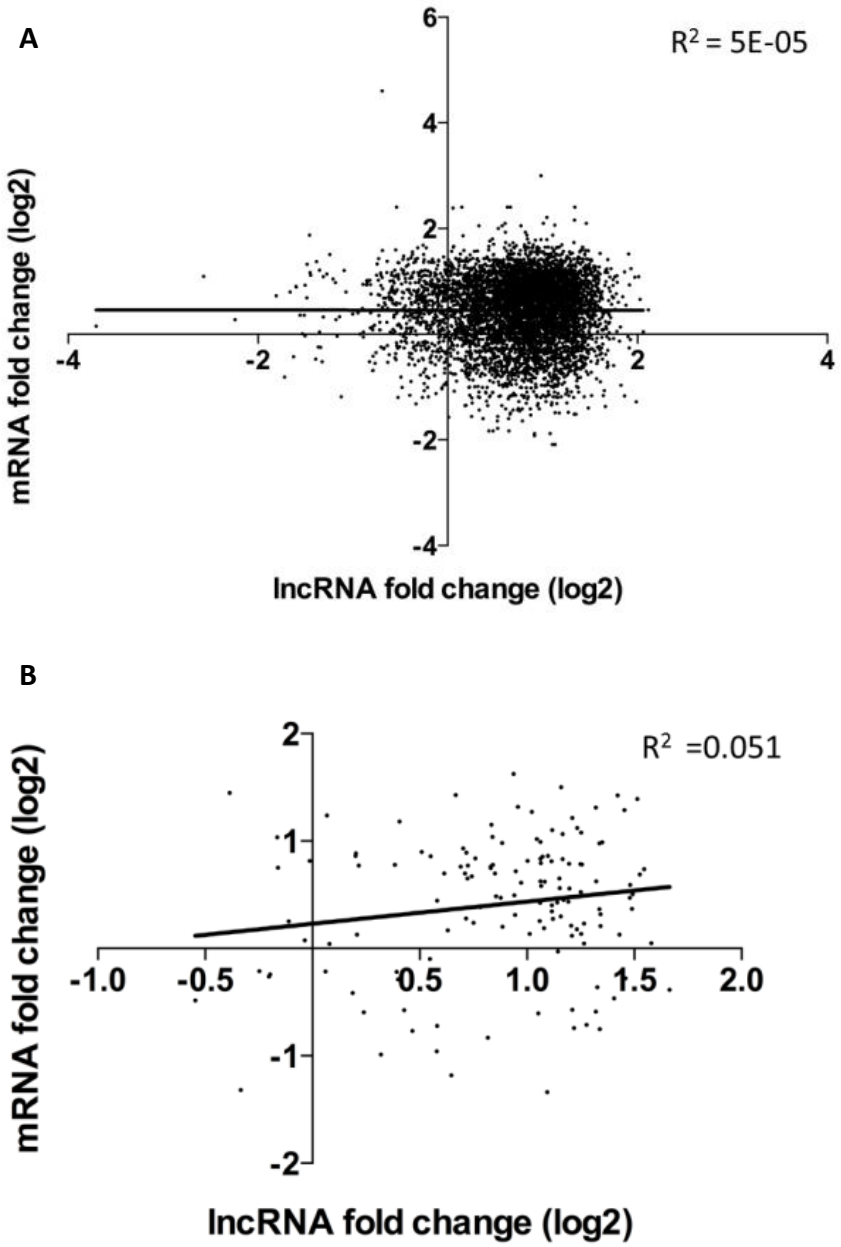


Figure 2.10 Correlation between lncRNA and protein coding RNA transcription. Correlation plot between log_2 expression of lncRNAs against **A)** the closest expressed protein coding genes and **B)** against protein coding genes within 1 kb of lncRNAs. Linear regression was represented in both plots and R-square was calculated.

2.3.9 Hypoxic induction of non-annotated intergenic and antisense transcripts

A proportion of sequence reads did not map to an annotated transcript of any class. To define whether these reads represented previously non-annotated transcripts they were first examined for contiguity using the Cufflinks transcriptome assembly algorithm (http://cufflinks.cbcb.umd.edu). Use of uniquely mapped data with a threshold minimum coverage of 100 reads in at least one condition identified 102 novel RNA transcripts. To confirm these were transcripts that had not been detected previously, I compared the chromosomal location of these novel transcripts with other transcriptome assemblies, such as Ensembl, UCSC, and ENCODE datasets. I found that 11 of these transcripts were detected in one or more of the additional databases but not in the RefSeq database.

The remaining 91 novel RNA transcripts had a median length of 9,988 bp (range 269 to 382,549 bp) and many were upregulated in hypoxia **(Supplementary tables 2 & 3).** In keeping with their assignment as *bona fide* non-annotated transcripts, reference to UCSC data (http://genome.ucsc.edu/cgi-bin/hgTrackUi?hgsid=336849981&g= cpgIslandExt) revealed that 50 (55 %) have a CpG Island within 1000 bp of the putative promoter. Furthermore, all 91 transcripts were also detected in the independently derived polyA-selected RNA-seq dataset, indicating that these transcripts are polyadenylated. Of the 91 previously non-annotated transcripts, 37 did not overlap with any annotated RefSeq gene, and were therefore classified as non-annotated intergenic transcripts **(Supplementary tables 2).** Examples of non-annotated intergenic transcripts are shown in **Figure 2.11A.**

In silico analysis was conducted to evaluate the protein coding potential of these transcripts using a Coding Potential Calculator (CPC) algorithm, which has been previously used to distinguish coding from non-coding transcripts with high accuracy (Kong et al., 2007, Flockhart et al., 2012). The algorithm provides a numerical score for a given transcript, which indicates its protein coding potential (a positive score indicates a likely coding transcript and a negative or low score indicates a likely low-coding or non-coding transcript). As controls for the

methodology, five known protein-coding transcripts were evaluated and all gave highly positive CPC scores with this algorithm. In addition, five well-annotated non-coding RNAs yielded low or negative scores **(Table 2.1).** I then evaluated the CPC score of the previously unannotated intergenic transcripts. All had low or negative CPC scores comparable to the well-annotated non-coding RNAs, indicating that these transcripts are likely lncRNAs **(Supplementary table 2).** qPCR of the most strongly hypoxia upregulated non-annotated intergenic transcripts confirmed both expression and hypoxic induction in MCF-7 cells **(Figure 2.11C).**

The remaining 54 previously unannotated transcripts all overlapped with a previously documented RNA, but were expressed from the opposite DNA strand and were therefore defined as non-annotated antisense transcripts (NATS) **(Supplementary table 3)**. Examples of these non-annotated antisense transcripts are shown in **Figure 2.11B.** These transcripts mainly overlap with either the 3' or the 5' end of the sense genes. However, a few (n=10) transcripts spanned the entire sense gene. Three of these NATs have been previously reported in other cell types, including antisense transcripts at the *homeobox C11 (HOXC11) locus, the homeobox C13 (HOXC13)* locus and at the *HIF-1a* locus itself (Uchida et al., 2004, Rinn et al., 2007, Mainguy et al., 2007). Expression of the top 10 most upregulated NATs was confirmed by qPCR and showed a high concordant with the RNA-seq expression **(Figure 2.11D).**

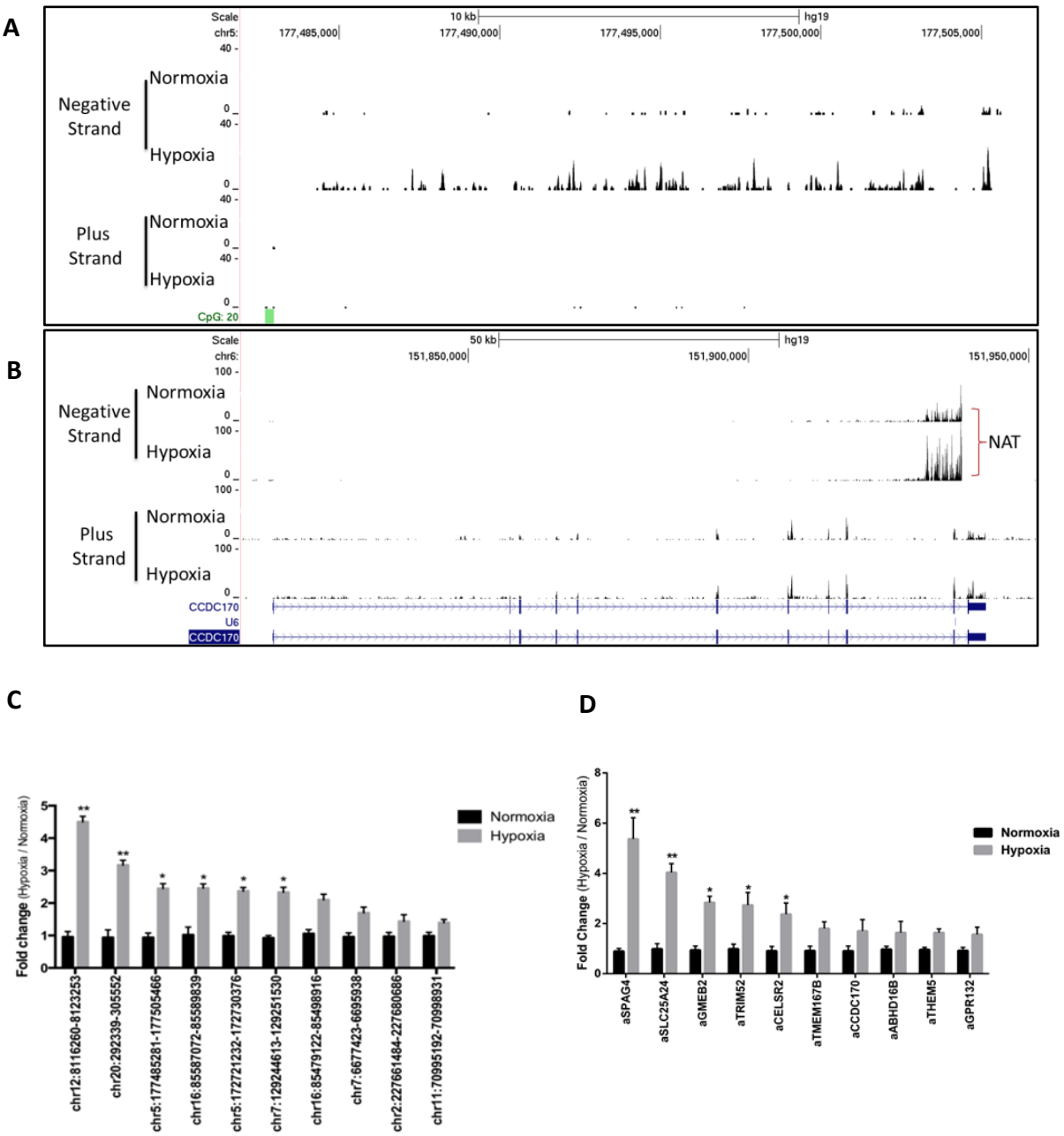


Figure 2.11 Hypoxia regulation of non-annotated intergenic and antisense transcripts in MCF-7 cells. A) Snapshot of the UCSC genome browser showing an example of a non-annotated intergenic transcript, that is transcribed from the negative DNA strand and is induced in hypoxic MCF-7 cells. **B)** Snapshot of the UCSC genome browser showing an example of a non-annotated antisense transcript which is transcribed from the 3' end of the coiled-coil domain containing 170 (CCDC170) gene locus in MCF-7 cells. **C)** qPCR validation of the 10 most hypoxia-upregulated non-annotated intergenic transcripts in MCF-7 cells. **D)** qPCR validation of the 10 most hypoxia-upregulated non-annotated antisense transcripts in MCF-7 cells. "a" before gene symbol indicates antisense transcript of the given gene. All qPCR experiments were performed with three biological replicates(* $p < 0.05$, **$p<0.005$ Student's t-test, n=3).

Table 2.1 CPC scores for the well-known protein coding RNAs and lncRNAs:

Gene name	Gene ID	Class	CPC score
ALDOC	NM_005165	Coding	8.12
NDRG1	NM_001258433	Coding	6.57
CA9	NM_001216	Coding	5.14
FOSB	NM_001114171	Coding	3.49
VEGFA	NM_001171623	Coding	3.03
MALAT1	NR_002819	Non-Coding	0.35
XIST	NR_001564	Non-Coding	-0.95
GAS5	NR_002578	Non-Coding	-1.02
PCA3	NR_015342	Non-Coding	-1.12
HOTAIR	NR_003716	Non-Coding	-1.19

2.3.10 Effect of hypoxic antisense transcripts in cis

As for other lncRNAs, naturally occurring antisense transcripts may act *in cis* to regulate the expression of the sense transcript at the same gene locus (Morris, 2009, Faghihi and Wahlestedt, 2009). Total RNA-seq analysis identified 54 antisense transcripts with minimum 100 reads in either condition in MCF-7cells. These transcripts overlap either 3' or 5' end of sense genes (25 transcripts overlap 5' end of sense gene, 22 transcripts overlap 3'end of sense genes and 7 transcripts overlap meddle part of sense genes). Interestingly, the majority of the detected non-annotated antisense transcripts were upregulated by hypoxia, hypoxic induction being commonly associated either with counter-regulation (n=16) for example, HIF-1α, TBX2, CELSR2 or with co-regulation of the overlapping sense transcript (n=23) for example SPAG4 **(Figure 2.12).** Thus, hypoxia-induced antisense transcripts are allied to both induction and repression *in cis,* further adding to hypoxic regulation of the transcriptome.

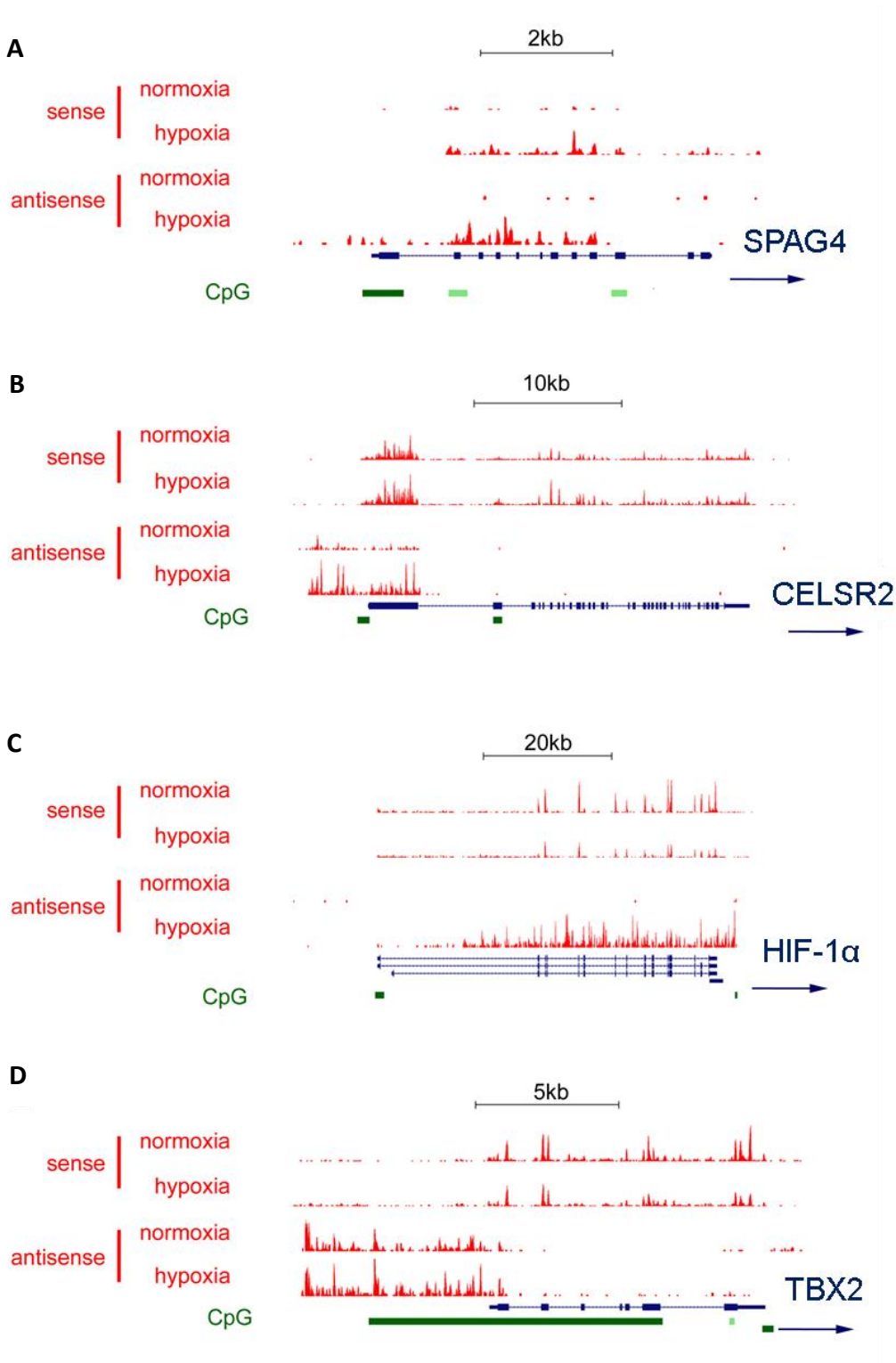


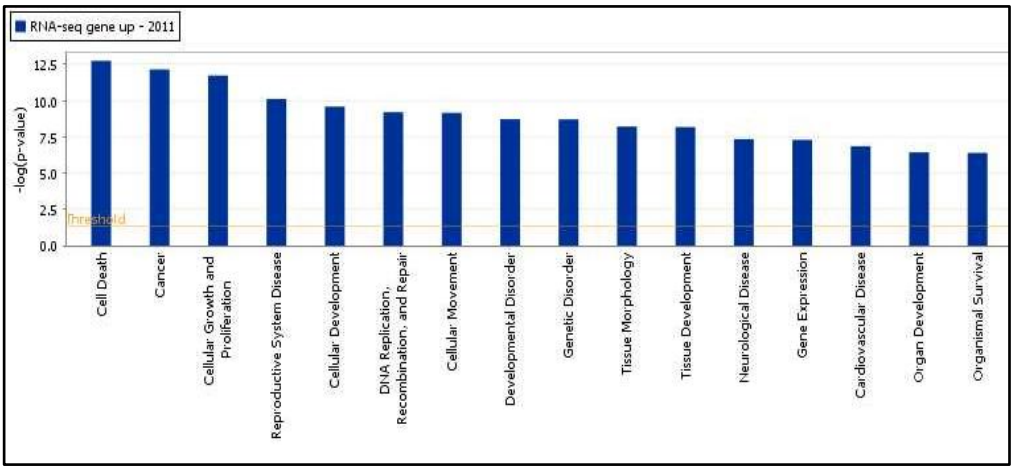
Figure 2.12 Co- and anti-regulation of overlapping sense and antisense transcripts. Ribosomal depleted directional RNA-seq tracks illustrating co-regulation of sense and antisense transcripts at **A)** sperm associated antigen 4 (SPAG4), and counter-regulation at **B)** cadherin, EGF LAG seven-pass G-type receptor 2 (CELSR2) **C)** HIF-1α and **D)** T-box 2 (TBX2). The arrow under each gene symbol represents the strand from which each RNA is transcribed.

2.3.11 Pathway analysis and gene ontology

For pathway and gene ontology analyses I focused on hypoxia-regulated mRNAs, since many small and long non-coding RNAs are poorly documented in the databases upon which such analyses rely.

The functions of the 200 most hypoxia-upregulated mRNAs were annotated using the Ingenuity Pathways Analysis software tool (Version IPA 8.8-3204) http://www.ingenuity.com/products/ipa. As expected from previous analyses of hypoxia pathways, large numbers of hypoxia-upregulated genes are involved in cancer, cellular growth, proliferation and cell death functions (P <0.05) **(Figure 2.13A).** Canonical pathways of hypoxia-upregulated genes were then examined. Key hypoxia related pathways that were statistically over-represented included HIF-1α signalling, Notch signalling, p53 signalling, Estrogen-dependent breast cancer signalling, arginine and proline metabolism and the endothelin-1 signalling pathway (P value <0.05) **(Figure 2.13B).** Network analysis identified 4 major networks, the largest containing 63 focus molecules with major nodes centered on HIF-1α, ERK1/2, TGFB1, MYC, SP1 and FOS indicating interactions between hypoxic regulation and these pathways.

A



B

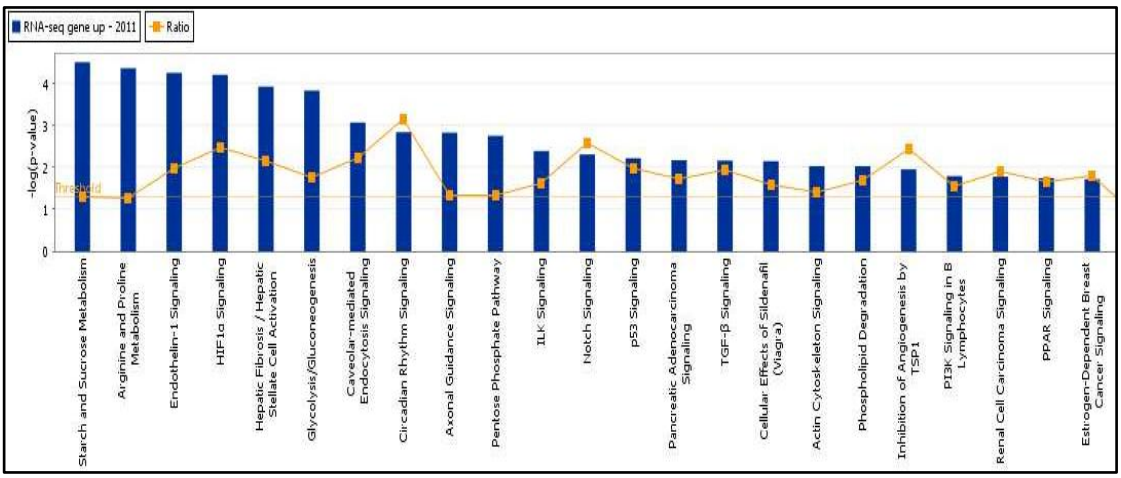


Figure 2.13 Functional annotation of hypoxia regulated genes. **A)** Top 16 functional categories that contain the largest number of hypoxia upregulated mRNA. **B)** Hypoxia upregulated mRNAs were examined for functional pathway analysis. The top 23 pathways containing most number of upregulated genes ($P<0.05$) are shown. The analysis was done by using the Ingenuity Pathways Analysis (IPA) software.

2.4 Discussion

In this chapter of the thesis I applied next generation RNA-seq technologies including polyA selected RNA-seq and strand specific ribosome-depleted RNA-seq (total RNA-seq) on the Illumina platform for comprehensive hypoxia transcriptome analysis. The numbers of uniquely mapable reads using *in-house* pipeline are similar to other total RNA-seq studies reported in the literature (O'Neil et al., 2013, Chen and Duan, 2011, Costa et al., 2011, Cui et al., 2010). There was a high degree of concordance in transcript detection when the standard polyA RNA-seq was compared with the total RNA-seq. However, total RNA-seq has benefits over polyA-selected RNA-seq in identifying different RNA classes such as small RNAs, and non-polyA mRNAs and lncRNAs as well as polyA RNAs and provides the most extensive survey of genes induced by hypoxia to date.

Although the total RNA-seq method has significant potential advantages, some limitations were encountered. For example, a proportion of rRNA sequences were detected even after two rounds of ribosomal depletion. Since these are not analysed in the expression studies they dilute the sequencing depth across the classes of RNA of interest. Low sequencing depth is particularly important when trying to determine the abundance and regulation of low expressed transcripts such as antisense transcripts and lncRNAs, as well as when studying post-transcriptional processing of RNAs, such as splicing or A-to-I editing. Increasing the total number of sequencing reads can compensate for this effect to an extent. Indeed, in the current study each sample was sequenced in two lanes of an Illumina GXII sequencer and the reads were combined from both sequencing runs to increase the depth. However, this approach does incur increased cost.

Another limitation of total RNA-seq is that there are as yet no established analysis tools that allow the unambiguous identification of different types of RNA of all classes (short and long RNAs) and measure their expression independently. Consequently, I established a computational pipeline, which allows reads to be mapped sequentially to databases of different

classes of RNA, together with subsequent quantification of the expression levels of each type of RNA. This pipeline maps sequentially to classes of RNA of increasing size. However, where short transcripts are embedded in longer "host genes" (e.g. miRNAs) it can be difficult to determine unambiguously whether a given read derived from the short "mature" RNA or from a longer "immature" form. For this reason, separate experiments to isolate short or long RNAs, which can then be sequenced separately, may provide better distinction between processed and unprocessed forms of RNA.

Using the *in-house* pipeline analysis on hypoxic and normoxic MCF-7 sequencing data a range of RNA classes were detected including piwiRNA, sn/snoRNA, tRNA, miRNA, lncRNA and protein coding RNA and showed class-specific hypoxic regulation. Expression of small RNAs such as piwiRNA, sn/snoRNA and tRNA showed overall downregulation, while miRNA, lncRNA and mRNA demonstrated overall upregulation in hypoxia. These results suggest that hypoxia induces widespread changes in transcriptome expression, beyond those previously described for mRNA and miRNA (Favaro et al., 2011, Kulshreshtha et al., 2007).

Many hypoxia-regulated mRNAs identified in this study, such as CA9, VEGF, ENO2 and NDGR1 have been reported in several microarray based studies (Chi et al., 2006, Seigneuric et al., 2007, Winter et al., 2007, Buffa et al., 2010, Conesa-Zamora et al., 2013). In addition, the current study has detected a number of novel hypoxia induced protein genes, which were not reported to be hypoxia regulated, for example TMSB10, PVRL4, RAB11FIP5, STON1, BCKDK, and RCN3. This increased sensitivity may reflect the generally greater dynamic range and fold-induction seen in RNA-seq studies compared to microarray analyses.

Functional pan-genomic pathway analysis of hypoxia-regulated mRNA transcripts has revealed functions linked to cancer, cellular growth, proliferation and cell death with significant numbers of these genes affecting key tumourigenesis pathways such as notch, p53 and oestrogen-dependent breast cancer signalling pathways. For example, among the newly identified hypoxia-regulated mRNAs, thymosin β10 (TMSB10) is an intracellular G-actin-sequestering protein that

regulates actin dynamics and cell motility (Sanders et al., 1992). Elevated TMSB10 expression is correlated with cell proliferation, angiogenesis and metastasis in various tumour types such as pancreatic, colon, breast and gastric cancer (Santelli et al., 1999, Sribenja et al., 2009). Poliovirus receptor-related 4 (PVRL4) is a member of the nectin family and is also a cell adhesion molecule. High PVRL4 expression is associated with poor outcome in many types of cancer, including breast cancer (Fabre-Lafay et al., 2007), NSCLC (Takano et al., 2009) and ovarian cancer (Derycke et al., 2010). The function and role of these and other newly identified hypoxia-regulated protein-coding genes may provide better understanding of the relationship between hypoxia biology and tumourigenesis and aid the identification of novel therapeutic targets.

The current study has shown that snRNAs, tRNAs and piwiRNAs are downregulated in hypoxia. The transcriptional regulation of these RNA classes has not been previously studied in hypoxia. I hypothesised that the downregulation of these small RNAs could due to alteration in processing and machinery of these RNA under hypoxic conditions. Indeed, I found that tRNA synthesis machinery, which is controlled by RNA Pol III, and tRNA transcription factors, was significantly downregulated in hypoxia. Previous studies have shown that RNA Pol III is repressed in hypoxia . However, the current study reports genome-wide expression of tRNA transcripts in hypoxia that is only possible through sequencing based technology. Alteration of tRNA expression could potentially have a profound effect on codon usage and protein synthesis in hypoxia.

In addition, the majority of piwiRNAs and snRNAs were also downregulated in hypoxia. This may be explained by the observed downregulation of their maturation and biogenesis proteins in hypoxia, including argonaute proteins (AGO3) for piwiRNAs and ribonucleoprotein complexes for snRNAs.

A number of microarray-based pan-genomic analyses have reported regulation of the miRNA in response to hypoxia. These studies have used a range of different cell types, hypoxic conditions and profiling platforms and have exhibited a high degree of heterogeneity in the response. For instance, in a systematic review of the hypoxic regulation of miRNAs (McCormick et al., 2010), only 23/103 (22 %) were regulated in more than one study with 9 of these (9 %) regulated in opposing rather than congruent directions. Nevertheless, I have analysed the data in the light of previous pan-genomic analyses of the miRNA response to hypoxia (Kulshreshtha et al., 2007, Hua et al., 2006, Hebert et al., 2007, Donker et al., 2007, Guimbellot et al., 2009), to identify commonly hypoxia regulated miRNAs. The data shows similar levels of overlap with each of these previously published studies as was observed between the studies **(Table 2.1).** In addition, the current study has identified a number of new hypoxia-regulated miRNAs, such miR-184, miR-1 and miR-100, that have been reported to have tumourigenic properties in cancer (Tivnan et al., 2010, Wong et al., 2009, Singh et al., 2013, Koberle et al., 2013, Tominaga et al., 2013).

One of the advantages of total RNA-seq is the detection of hypoxia-regulated non-polyA RNAs. Expression of replication dependent histone (H1D, H3C and H3J) non-polyA mRNAs were evaluated and showed significant downregulation in hypoxic MCF-7 cells and a number of breast cancer cell lines, suggesting that the non-polyA class of mRNAs may respond differently in hypoxia potentially adding another layer of transcriptional complexity to the hypoxic response. Recently, it was reported that H1D, H3C and H3J were preferentially expressed in undifferentiated stem cells and in reprogrammed cells, however, their expression significantly decreased upon differentiation (Yang et al., 2011a). The changes in histone mRNAs expression suggests that chromatin status and structure may be altered in hypoxia.

lncRNAs comprise a class of RNA that was previously not known to be widely regulated by hypoxia. Recent ENCODE data has identified a large number of these transcripts that are transcribed in multiple cell types and regulated by a variety of conditions and stresses (Derrien et al., 2012). However, the function of the majority of these transcripts remains unclear and their biological role has been widely debated in the literature (Ulitsky and Bartel, 2013, Batista and Chang, 2013). A number of studies have proposed potential functions; for example, they can act as enhancer marks (De Santa et al., 2010), can trigger genome reprogramming (Loewer et al., 2010) and can act *in cis* to activate transcription, or *in trans* to repress transcription (Rinn et al., 2007, Huarte et al., 2010). Although ENCODE provides the most updated and comprehensive dataset for lncRNAs, this database has a number of limitations. For example, multiple transcripts are annotated at the same genomic locus with different exonic structures (potential isoforms) and poorly defined TSSs. Thus, I created a "lncRNA containing locus" (LCL) database by combining transcript exons, which are transcribed from the same genomic locus to build a unique lncRNA meta-database with defined TSSs. I found that a large number of lncRNAs (502) were upregulated by more than 1.5-fold and 33 lncRNAs were downregulated by more than 1.3-fold in hypoxia.

Several well-described lncRNAs such as NEAT1, MALAT1 and H19 were among the most hypoxia-upregulated lncRNAs. To date there have been a limited number studies, which have reported regulation of lncRNAs in hypoxia, although none of these studies examined the pan-genomic regulation of lncRNAs. Matouk and colleagues showed that H19 lncRNA is abundantly expressed in hepatocellular carcinoma and in a panel of human tumour cell lines. Furthermore, they found that the hypoxic regulation of H19 expression in HCC and bladder carcinoma cell lines was dependent on HIF-1α (Matouk et al., 2007). They reported that knockdown of H19 in Hep3B xenografts in CD-1 nude mice resulted in a significant (82 %) reduction of tumour volumes (Matouk et al., 2007). In another study, Matouk et al. showed that elevation of H19 expression in hypoxia is associated with the status of the p53 tumour suppressor, as cells with $p53^{wt}$ show no induction in H19 expression in hypoxia. In contrast, $p53^{null}$ cells exhibited a high

degree of H19 induction by hypoxia (Matouk et al., 2010). These studies together with the current study, which identifies a number of new hypoxia regulated lncRNAs, provide a potential novel molecular mechanism that integrates lncRNAs and the hypoxic stress response in cancer. The role and function of NEAT1 and MALAT1 in cancer is discussed in more detail in a later chapter.

lncRNAs can transcriptionally regulate expression of protein coding genes *in cis* (Zhang et al., 2012a, Wutz, 2011). However, analysis of hypoxia regulated lncRNAs show no significant correlation between lncRNA level and expression of the closest expressed protein-coding genes even when censured for protein-coding genes within 1kb. However, there are few exceptions where upregulated lncRNAs are anti-regulated with protein coding genes (downregulated), suggesting that few hypoxia regulated lncRNAs may act *in cis* to regulate the expression. Nevertheless, confirmation of these hypotheses would require extensive interventional studies, each as knockdown and/or overexpression hypoxia-regulated lncRNAs to determine their effect on neighbouring protein-coding genes directly and is this proposed as a future project to identify *cis* acting lncRNAs in hypoxia.

Finally, a proportion of reads did not map to databases of any RNA class. These reads were mapped uniquely and in contiguity to the genome with minimum threshold coverage of 100 reads in at least one condition, to detect non-annotated transcripts with high confidence. This analysis identified 91 novel RNA transcripts, which have not been reported in any databases. The presence of CpG Islands close to the putative promoters of 50 % of these transcripts and their observation in multiple datasets, suggests that these are *bone fide* transcripts.

In silico analysis confirmed that all previously unannotated transcripts had a low or negative CPC score, suggesting that they belong to the lncRNAs class. These transcripts were divided into two types depending on their chromosomal location and orientation; (1) non-annotated intergenic transcripts, (2) non-annotated antisense transcripts (NATS). Among the list of antisense transcripts, I found antisense HIF-1α transcript (aHIF-1α) that was upregulated (2.1-

fold) in hypoxia, whilst the sense transcript was downregulated. Identification and regulation of aHIF-1α has been previously reported. The antisense transcript is transcribed from the 3'-untranslated region of the sense HIF-1α mRNA and is highly expressed in many cancers (Bertozzi et al., 2011, Thrash-Bingham and Tartof, 1999, Cayre et al., 2003). It is suggested that the aHIF-1α transcript acts in *cis* to regulate expression of the sense HIF-1α mRNA (Uchida et al., 2004). Several reports support the notion that naturally occurring antisense transcripts may act in *cis* to regulate expression of the sense transcript at the given gene locus in this way (Morris, 2009). Thus, hypoxia-induced antisense transcripts are allied to both induction and repression *in cis,* adding further complexity to hypoxic transcriptional regulation. The role of antisense in hypoxia can be further investigated by supressing its expression in hypoxia or by overexpressing its expression in normoxia and evaluating the expression of sense gene both at the RNA and protein level. Suppression of antisense can be performed using short interference RNA (siRNA) if the antisense is cytoplasmic or suppression by antisense oligonucleotides if the antisense is localise in nucleus. Effect of antisense transcript suppression can be evaluated on oncogenesis by a number of assays such as on cell proliferation, apoptosis, cell survival, tumour growth in cancer xenograft models.

Overall this section provides the most comprehensive cataloguing of the hypoxia-regulated transcriptome to date, reporting class-specific patterns of regulation in hypoxia that were not previously recognised, such as downregulation of many piwiRNAs, snRNAs and tRNAs, and upregulation of many lncRNA and extends the range of hypoxia-regulated mRNAs. This work also identifies a number of previously non-annotated transcripts that are regulated by hypoxia.

Chapter 3: HIF-dependent transcriptional regulation of the non-coding transcriptome in hypoxia

3.1 Introduction

Reduction in the level of oxygen (hypoxia) induces profound changes in the transcriptional activity of the cell, resulting in major modifications that alter its metabolic state and reduce oxygen consumption, as well as attempts to restore oxygen delivery. Hypoxia-Inducible Factors (HIFs) play a major role in mediating this transcriptional response to hypoxia by controlling the expression of hundreds of genes involved in oxygen-sensitive pathways. In addition to these physiological roles, altered expression of HIF is also linked to the pathogenesis of a wide range of common diseases, including ischemic conditions such as cardiovascular and cerebrovascular disease, inflammatory conditions, anaemia and cancer (Semenza, 2012a).

There are two main isoforms of the HIF-α subunit, HIF-1α and HIF-2α. The HIF-1β subunit is constitutively expressed, while the HIF-1α and HIF-2α subunits are highly regulated by oxygen. In the presence of oxygen the α-subunits are hydroxylated, facilitating binding of von Hippel-Lindau protein (pVHL), which is part of an E3-ubiquitin ligase complex that ubiquitinates HIF-1α and HIF-2α leading to their rapid degradation. Hypoxia inhibits this hydroxylation, protecting HIF-1α and HIF-2α from pVHL-dependent degradation by the ubiquitin-proteasome pathway and allowing them to dimerize with HIF-1β. This dimerized HIF complex then binds Hypoxia Response Elements (HREs) on genomic DNA, leading to transactivation of HIF target genes (Kaelin and Ratcliffe, 2008).

To date, a number of approaches have been used to identify the pan-genomic transcriptional response to HIF, most frequently by genome-wide expression profiling using microarrays (Hu et al., 2003, Elvidge et al., 2006, Warnecke et al., 2008, Dekervel et al., 2014, Irigoyen et al., 2007, Mense et al., 2006, Tang et al., 2012, Lee et al., 2010, Lu et al., 2010, Ortiz-Barahona et al., 2010). These studies have identified a wide range of HIF regulated genes, which are important for hypoxia survival (e.g. VEGF, CA9, and ALDO). The majority of these genes are regulated by HIF-1α. However, HIF-2α levels remain elevated longer than HIF-1α and it is thought that HIF-2

plays a more substantial role later in hypoxic exposure (Koh and Powis, 2012, Ratcliffe, 2007). Nevertheless, in these studies a high proportion of the genes that are transcriptionally regulated by hypoxia, are regulated by direct interventions on the HIF pathway, indicating that HIF is the major transcriptional regulator of the response to hypoxia. However, these expression studies are not able to distinguish between direct and indirect transcriptional targets of HIF.

One approach to distinguish directly from indirectly regulated genes is to use bioinformatic analysis of their promoter regions to identify HREs from the known consensus motif (Pugh et al., 1991, Pescador et al., 2005). Binding of HIF at individual sites can then be determined by Chromatin Immunoprecipitation (ChIP) and combined with reporter gene assays to test the function of potential HIF binding sites. However, these approaches can examine only one potential HIF-binding site at a time and it is now known that actual HIF-binding sites frequently lie at considerable distance from the transcriptional start site (TSS), with only a minority lying within the promoter. Therefore, more recent attempts to determine direct HIF transcriptional targets have employed more systematic approaches to identify HIF-binding sites using Chromatin Immunoprecipitation coupled with hybridization to ‘ tiled ’ arrays (ChIP on chip) or with massively parallel DNA sequencing (ChIP-seq)(Park, 2009, Mole et al., 2009, Schodel et al., 2011, Mimura et al., 2012, Tanimoto et al., 2010).

To date, there have been relatively few pan-genomic analyses of HIF binding sites using ChIP-seq (Schodel et al., 2012, Schodel et al., 2011, Tanimoto et al., 2010, Mimura et al., 2012). Most of these studies have identified approximately 500 high-stringent HIF binding sites across the genome in each cell line studied. However, Mimura and colleagues performed HIF ChIP-seq on human umbilical vein endothelial cells (HUVEC) cells in normoxia and hypoxia. They used less stringent thresholds compared to other studies and defined 575 binding sites in normoxia and 2060 binding sites in hypoxia for HIF-1(Mimura et al., 2012). Conversely, Xia et al. used highly stringent thresholds to define HIF binding sites based on the presence of an HRE in the sequence data (Xia and Kung, 2009, Xia et al., 2009), whereas, Schödel et al. defined the HIF

binding sites based on previously published HIF ChIP-on-chip data set and on the overlap between HIF-α and HIF-β binding (Xia et al., 2009, Mole et al., 2009, Schodel et al., 2011). It is estimated that these high stringency criteria may exclude approximately 50 % of the binding sites, suggesting that less stringent criteria may identify as many as 1,000 HIF binding sites in hypoxic MCF-7 cells (Schodel et al., 2013). These numbers are comparable with the binding sites identified for other transcription factors, such as p53 in HCT116 colorectal cancer and FOXP3 in T-cells (Zheng et al., 2007, Wei et al., 2006).

Generally, a HIF-regulated gene has been classified as a direct transcriptional target if it is adjacent to a HIF-binding site. However, this likely underestimates the number of direct transcriptional targets, as HIF does not necessarily regulate the closest or even a nearby gene. In addition, ChIP-seq and ChIP-chip analyses performed to date have only been correlated with expression analyses using microarrays, which are inherently focused on protein-coding genes.

In this chapter, I have correlated my total RNA-seq analysis, describing the hypoxic regulation of all classes of RNA, with a previously published ChIP-seq dataset (Schodel et al., 2011). In addition, I performed polyA RNA-seq of hypoxic cells treated with siRNA to each of the HIF-α isoforms to determine HIF-dependent regulation of polyA mRNA and lncRNA transcripts.

3.2 Exprimental desigen

The raw ChIP-seq data of HIF-1α, HIF-2α and HIF-1β were downloaded from the GEO database and were mapped against the recent human genome assembly (HG19) using Bowtie V 1.0.0(Langmead et al., 2009). The CisGenome software (Ji et al., 2008) was used for peak calling of HIFs in MCF-7 cells. The association between HIF-binding and regulation by hypoxia across different classes of RNA was performed using Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005). In addition, HIF-1α, HIF-2α and combined HIF-1/2α knockdown was performed in hypoxic MCF-7 cells (1 % O_2 for 24 hours). Total RNA was isolated from individual knockdowns and subjected to polyA selected RNA-seq using lllumina TruSeq protocol. The libraries were sequenced on Illumina HiSeq 2000. The raw sequencing data were mapped to human genome assembly (HG19) and analysed using in house pipeline as described in the previous chapter. The Ingenuity Pathway Analysis (IPA) software tool (Version IPA 8.8-3204, http://www.ingenuity.com/products/ipa) was used for gene ontology, gene network and pathway analysis for transcript downregulated with combined HIF-1/2α siRNA. HIF dependent lncRNA MALAT1 regulation was validated in 11 breast cancer cell lines in hypoxia using qPCR. In addition, expression of MALAT1 was evaluated in MCF-7 cells at different hypoxia time points (1 % O_2 for 16, 24, 32, 48 hours), after DMOG treatment (500 nM, 24 hours) and HIF siRNAs by qPCR. Furthermore, two HIF dependent antisense transcripts (aSPAG4 and aCELACR2) were selected for validation. The expression of antisense and its overlapping sense transcript (for the aSPAG4, SPAG4 and aCELACR2, CELACR2 transcripts) were measured by qPCR in a panel of breast cancer cell lines grown in hypoxic condition (1 % O_2 for 24 hours). In addition, using qPCR their expression was evaluated at different hypoxia time points (1 % O_2 for 16, 24, 32, 48 hours) and with HIF siRNAs in MCF-7 cells. The 3' Rapid Amplification of cDNA Ends (RACE) PCR was performed for both aSPAG4 and aCELACR2 to identify the transcript length and the exonic structure of the transcripts. Moreover, to identify the cellular localisation of aSPAG4 and aCELACR2, cytoplasmic and nuclear fractions of MCF-7 cells grown in hypoxia (1 % O_2 for 24

hours) and normoxia were fractionated and subjected to protein and total RNA isolation. Western blot for GAPDH (cytoplasmic marker) and histone H3 (nuclear marker) was performed in both fractions to ensure effective isolation of each fraction. Expression of aSPAG4, SPAG4, aCELACR2 and CELACR2 were measured by qPCR in cytoplasmic and nuclear fractions of MCF-7 cells in normoxia and hypoxia. Finally, RNA fluorescence *in situ hybridization* (RNA FISH) was performed against aSPAG4, aCELACR2 and RPL11 (control probe) in MCF-7 cells in normoxia and hypoxia (1 % O_2 for 24 hours). All the experiments were performed in three independent biological replicates.

3.3 Results

3.3.1 Defining HIF binding in hypoxic MCF-7 cells

To define HIF binding sites in MCF-7 cells across the genome, I used previously published ChIP-seq data for HIF-1α, HIF-2α and HIF-1β that has been well validated in MCF-7 cells (Schodel et al., 2011). Since this data was originally mapped to the NCBI build 36 of the human genome (HG 18), I re-analysed the raw sequencing data for each ChIP and remapped them (using Bowtie Version 1.0.0) (Langmead et al., 2009) to the NCBI build 37 of the human genome (HG19), which had been used for the RNA-seq analysis. Subsequent analysis including peak finding was performed using the CisGenome software (Ji et al., 2008) using a similar pipeline as Schödel et al. Since HIF-1α and HIF-2α subunits bind hypoxia response elements (HREs) as a heterodimer with HIF-1β subunits, the binding sites were defined by an estimated false discovery rate (FDR<0.05) and were filtered for the presence of both HIF-α and HIF-1β to define highly stringent HIF-1 and HIF-2 binding sites.

I then analysed which classes of RNA were bound by HIF. Each HIF binding site was annotated according to the closest active promoter (defined as the promoter of a gene with an expressed transcript). The binding sites were then classified according to the class of transcript to which this gene belonged. The number of HIF-binding transcripts of each class broadly mirrored the total number of transcripts belonging to that class **(Figure 3.1A).** Although the numbers in many classes were too small to permit statistical analysis, some patterns were observed. In particular, no HIF binding was observed in the vicinity of classes of RNA that showed global downregulation (piwiRNAs, snRNAs and tRNAs), suggesting that they are unlikely to be directly regulated by HIF. In contrast, HIF-binding was observed adjacent to both mRNAs and lncRNAs **(Figure 3.1A)** with HIF-2 binding a slightly higher proportion of lncRNAs than HIF-1. HIF binding was also annotated at a number of miRNA gene loci. However, this may be an

underestimate, since miRNAs are frequently “embedded” within host transcripts, the promoter of which may lie closer to the HIF-binding site than the miRNA.

Previous analysis had indicated that approximately 60 % of HIF-1 binding sites and 80 % of HIF-2 binding sites lie more than 2.5-kb from the nearest promoter (Schodel et al., 2011). However, this work used only the promoters of protein-coding genes and did not take into account the possibility that the binding site was close to the promoter of non-coding transcripts. I therefore repeated the analysis using the promoter (TSS) of all transcribed genes (coding, non-coding and non-annotated) identified by total RNA-seq. Despite the inclusion of these additional genes, the majority of HIF-1 and HIF-2 binding sites (50 % and 70 % respectively) were still found to lie in excess of 2.5-kb from the promoter of the nearest expressed transcript **(Figure 3.1B-C),** confirming previous findings that HIF frequently binds promoter-distant enhancers.

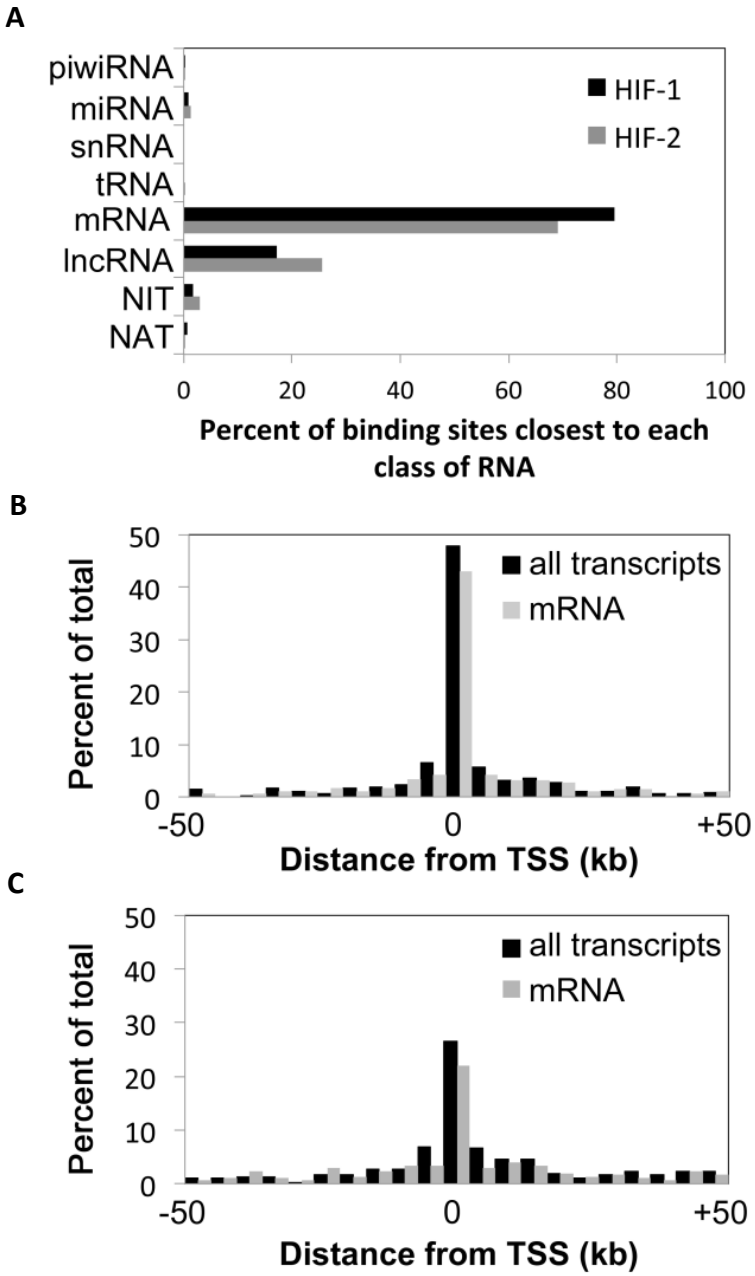


Figure 3.1 HIF-binding upregulates both the coding and non-coding transcriptome. A) The proportion of HIF-1α and HIF-2α binding sites that are closest to transcribed loci of each RNA class. Raw number of HIF-1α binding sites in different RNA classes (piwiRNA 1/150, miRNA 3/176, mRNA 282/16846, lncRNA 61/6235, NIT 6/54, NAT 2/44) and taw number of HIF-2α binding sites for each RNA claess (piwiRNA 1/150, miRNA 4/176, tRNA 1/552, mRNA 208/16846, lncRNA 77/6235, NIT 9/54, NAT 1/44).The distribution of **B)** HIF-1α and **C)** HIF-2α binding sites around the transcriptional start site of the nearest expressed gene irrespective of class (black bars). The grey bars in **B-C** show the distribution when only active mRNA genes are used (analogous to previous analyses (Schodel et al., 2011)).

3.3.2 Correlation of HIF binding sites with gene expression

To assess the functional role of HIF in the regulation of different classes of RNA by hypoxia, I looked for an association between HIF-binding and regulation by hypoxia across different classes of RNA using Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005). All expressed transcripts were pooled and ranked according to their hypoxic induction. HIF-binding transcripts were defined as the closest expressed transcript to a HIF-binding site using their promoters as above. HIF-binding transcripts (HIF-1 or HIF-2) were strongly enriched amongst hypoxia-upregulated transcripts but not amongst hypoxia-downregulated transcripts **(Figure 3.2A).** This is consistent with previous analyses (limited to protein-coding genes) and indicates that HIF is functioning solely as a transcriptional enhancer and not as a repressor.

Not surprisingly, given their global downregulation in hypoxia and their lack of HIF binding, snRNAs, tRNAs and piwiRNAs were not present in the "core-enrichment" group of HIF-binding hypoxia-inducible transcripts, which contained only transcripts belonging to mRNA, lncRNA, and miRNA classes. When HIF-binding mRNAs, lncRNAs and miRNAs were analysed separately, both HIF-binding mRNAs and lncRNAs showed enrichment amongst the hypoxia-upregulated genes **(Figure 3.2B-C)**, although the number of HIF-binding miRNAs was too small to permit meaningful analysis **(Figure 3.2D)**. Taken together, these analyses provide a functional link between HIF-binding and hypoxic gene regulation at both mRNA and lncRNA gene loci, pointing to a major new role for HIF in the transcriptional regulation of lncRNAs.

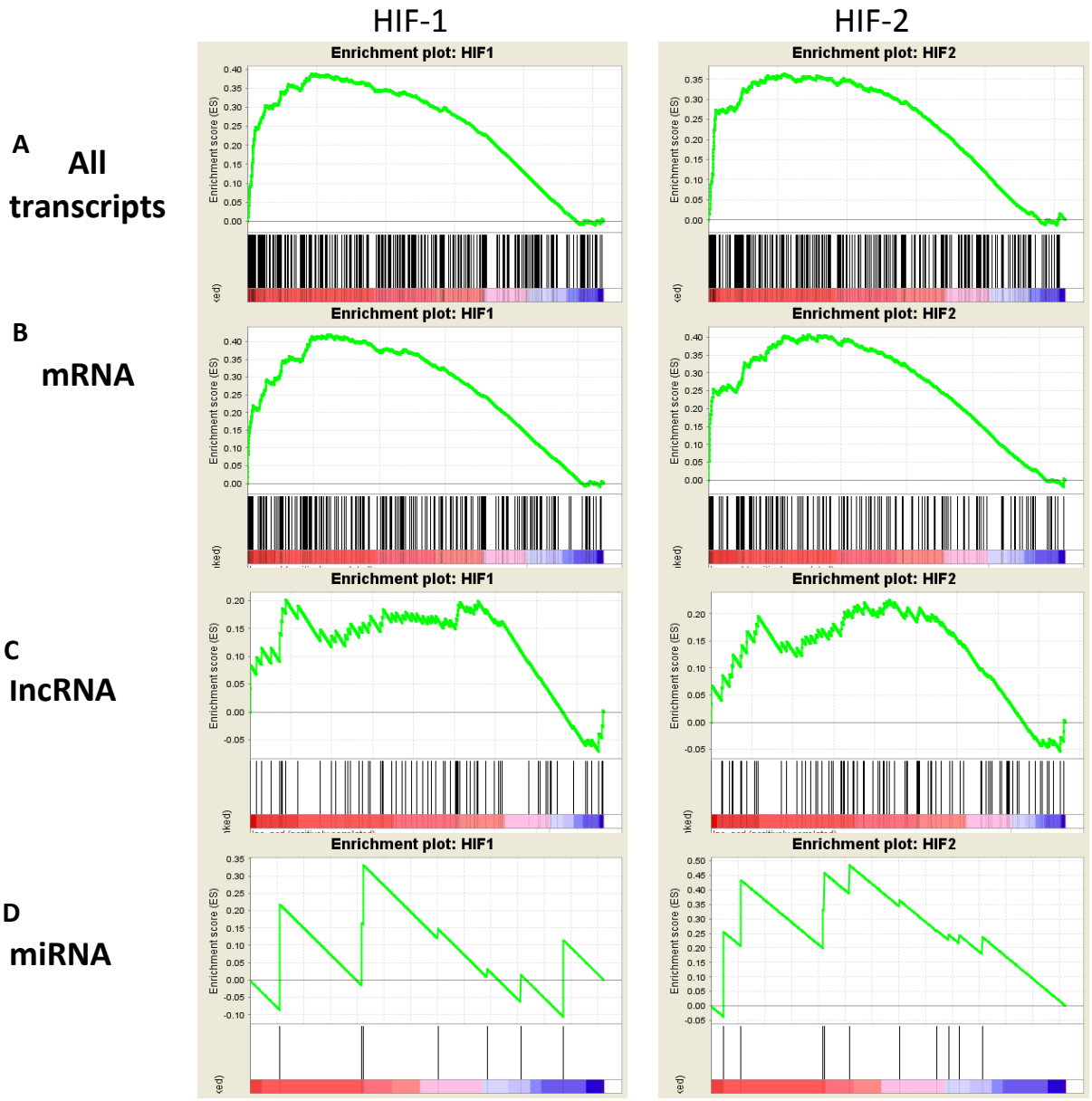


Figure 3.2 Gene Set Enrichment Analysis for HIF binding sites against fold-regulation by hypoxia of different RNA class. The GSEA of HIF-1α and HIF-2α against **A)** all transcripts **B)** mRNA, **C)** lncRNA **D)** miRNA**.** All expressed transcripts were ranked from most upregulated on the left (red) to most downregulated on the right (blue) according to the $\log_2$ fold-change in transcript abundance in hypoxia. A vertical black bar denotes each HIF-binding gene. The green line denotes a running metric with a positive score for each HIF-binding gene and a negative score for each non-binding transcript, weighted by the proportion of HIF-binding genes and by $\log_2$ fold-change. The positive deflection towards the left of each graph denotes enrichment of HIF-binding genes amongst the most upregulated transcripts. The negative gradient across the remainder of the graph denotes an excess of non-binding genes (and hence a lack HIF binding) amongst the non-regulated and downregulated genes. Similar patterns were observed for both HIF-1 and HIF-2 binding genes.

3.3.3 HIF-dependent regulation of the coding and long non-coding transcriptome

Given that HIF plays a key role in transcriptional regulation in hypoxia and that HIF-bound lncRNA loci are preferentially upregulated in hypoxia, I next examined the extent to which these transcripts were regulated by HIF itself by suppressing HIF expression. Rather than repeating the RNA-seq analysis of all classes of RNA, I focussed on mRNAs and a subset of lncRNAs using RNA-seq analysis of polyA-selected RNA, which allowed greater sequencing depth of these transcripts for the same number of total reads. HIF-1α, HIF-2α, or combined HIF-1/2α suppression in hypoxic (1 % hypoxia) MCF-7 cells was achieved using siRNAs and compared with a scrambled control siRNA. These siRNAs have been previously validated and tested for off target effects (Elvidge et al., 2006).

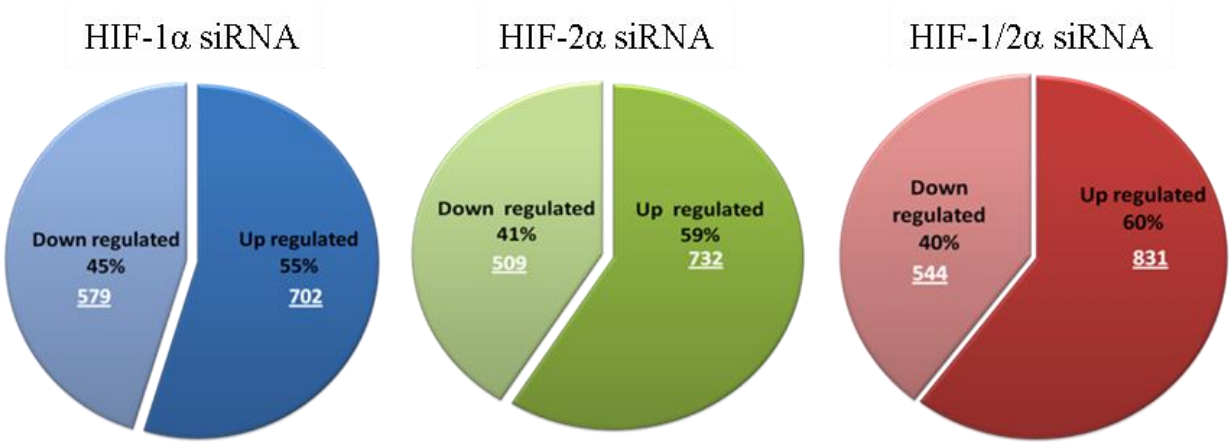
Poly-adenylated RNA was isolated from total RNA and the resultant cDNA libraries were subjected to RNA-sequencing (paired-end, length: 51/100 bp) using the Illumina HiSeq 2000 platform**.** The data were mapped using an in-house pipeline (as described in Chapter 2) and the differential expression of each expressed transcript was calculated for each siRNA treatment compared to scrambled controls. Differential expression analysis was performed using the edgeR package (Robinson et al., 2010). Applying a threshold FDR < 0.05 and fold-change ≥ 1.5 identified 702 genes that were upregulated and 579 genes that were downregulated by HIF-1α siRNA, 732 genes that were upregulated and 509 genes that were downregulated by HIF-2α siRNA and 831 genes that were upregulated and 544 genes that were downregulated by combined HIF-1α and HIF-2α siRNA. The range of changes induced by HIF knockdown varied from 735-fold upregulation to 94-fold downregulation **(Figure 3.3A).**

Since this is the first analysis, by RNA-seq, of both the coding and the non-coding transcriptional response to HIF, I compared my results to the expression microarray findings of Elvidge et al. who used the same MCF-7 cells and siRNAs to study mRNAs alone (Elvidge et al., 2006). In contrast to my findings, they reported that only 127 mRNAs were suppressed by HIF-1α siRNA and that just 5 mRNAs were suppressed by HIF-2α siRNA. However, all differentially-expressed

genes identified by Elvidge et al. were regulated in the RNA-seq analysis, but with a greater estimated fold-change. Taken together, this suggests that RNA-seq is more sensitive in identifying differentially regulated genes, possibly as a result of a greater dynamic range. It also indicates a greater functional role for HIF-2α in MCF-7 cells than that described by Elvidge *et al.* 2006.

Given the greater number of differentially-expressed mRNAs identified by the RNA-seq approach, I performed gene function and canonical pathway analysis based on the genes that responded to combined HIF-1α and HIF-2α siRNAs to determine whether novel pathways or functions could be identified. However, mRNA transcripts downregulated by HIF siRNA (i.e. upregulated by HIF) had functions that had mostly been described previously, in particular genes related to cancer, cell death, cellular growth and proliferation, with enrichment in key cancer pathways, including EGF, ERK/MAP, renal cell carcinoma, TGF-beta and VEGF signalling pathways ($P<0.05$) **(Figure 3.3B)**. Overall, this analysis provides the most comprehensive description of the response of the HIF- dependent transcriptome to date.

A



B

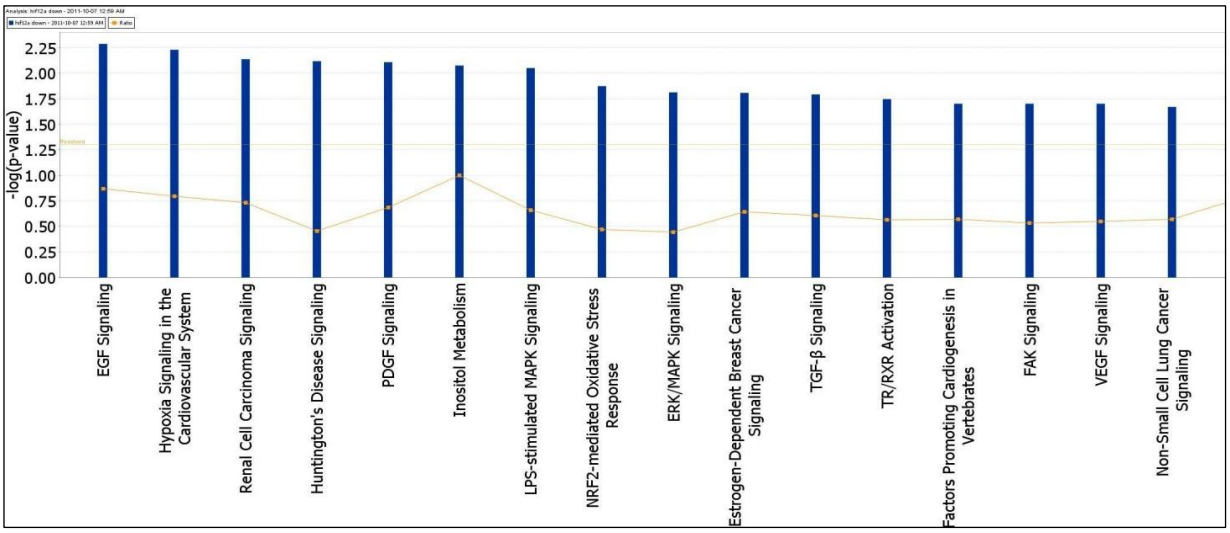


Figure 3.3 Differentially-expressed transcripts in hypoxic MCF-7 following HIF-1α, HIF-2α and HIF-1/2α siRNA. A) The number of up- and downregulated transcripts following the indicated HIF siRNA transfection in hypoxic MCF-7 cells. **B)** Pathway analysis for the top 100 downregulated genes after combined HIF-1/2α siRNA. The 16 most significantly enriched pathways (P<0.05) are shown. Analysed using Ingenuity Pathways Analysis software tool (Version IPA 8.8-3204) http://www.ingenuity.com/products/ipa.

I next used the polyA RNA-seq to examine the HIF-dependent regulation of the enriched group of HIF-binding, hypoxia-regulated lncRNAs**.** Approximately 25 % of all lncRNA transcripts that were detected in the ribosome-depleted total RNA-seq were also detected in the polyA RNA-seq analysis. Of these, 46 were HIF-binding lncRNAs. **Figure 3.4A** portrays the results of HIF-α siRNA interventions on these 46 HIF-binding lncRNA transcripts detected in this dataset. As described above, the heatmap confirms upregulation of the majority of these lncRNAs by hypoxia. In addition, many show subsequent downregulation by HIF-1α siRNA, HIF-2α siRNA or both, confirming the expected HIF dependence of the hypoxic induction of these genes. Furthermore, GSEA of HIF-binding lncRNAs shows significant enrichment amongst lncRNAs downregulated by HIF-1α, HIF-2α or combined siRNAs **(Figure 3.4 B-D).** Notably, the two most hypoxia-upregulated HIF-binding and HIF-regulated lncRNAs were NEAT1 and MALAT1. Analysis of RNA-seq and ChIP-seq tracks demonstrating the hypoxia and HIF-dependent regulation, together with the HIF-binding sites at these loci are illustrated in **Figure 3.4E-F.** Taken together, these datasets identify a set of lncRNAs with proximate HIF-binding together with hypoxia and HIF-dependent regulation, which define lncRNAs as a new group of HIF target genes. Furthermore, since the polyA RNA-seq analysis identifies only 25 % of the lncRNA transcriptome, the likely number of lncRNAs directly regulated by HIF is likely to be considerably higher than those described.

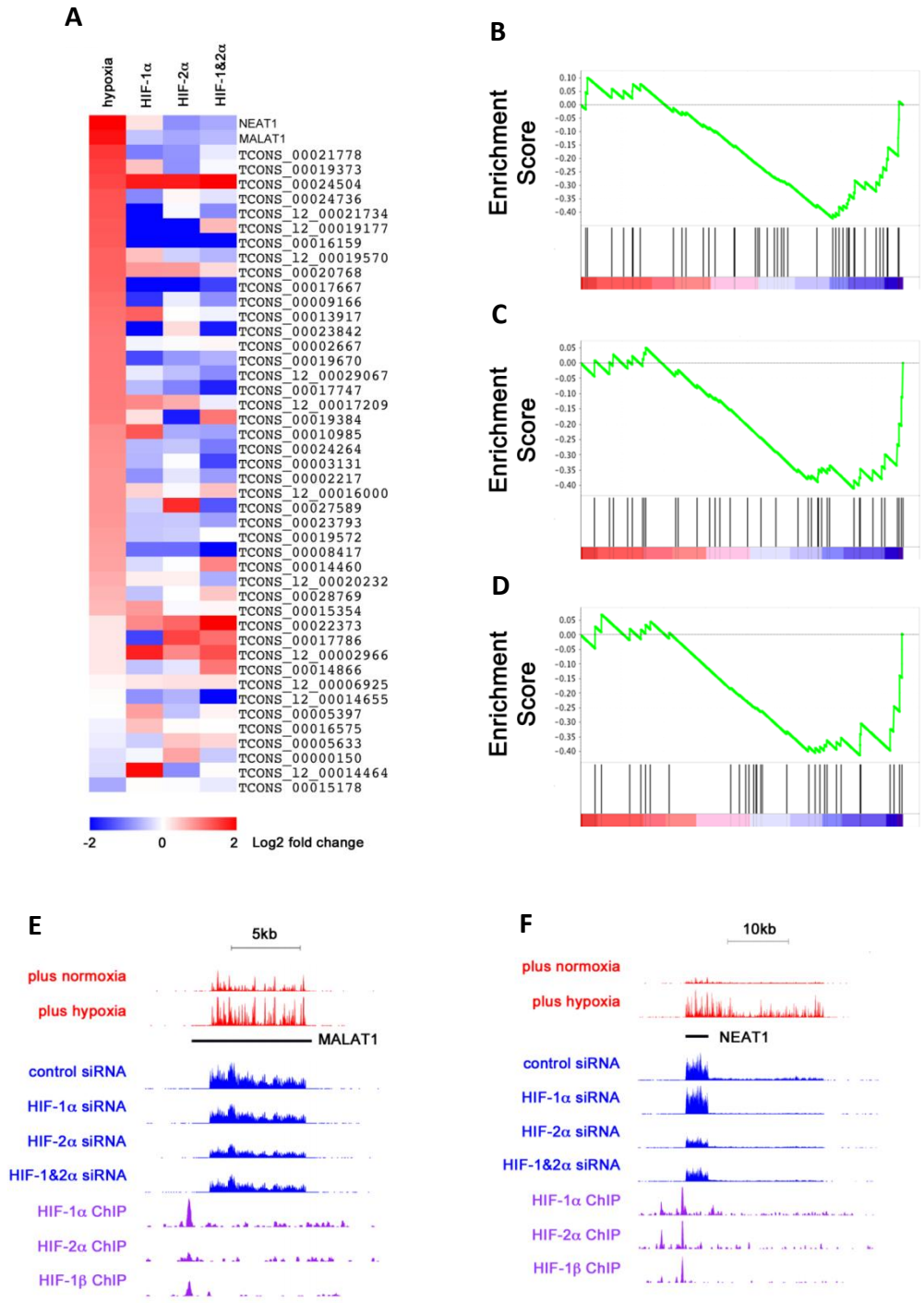


Figure 3.4 Direct transcriptional regulation of lncRNAs by HIF. A) Heat map showing fold regulation by hypoxia and subsequent fold-change following suppression of the indicated HIF subunit by siRNA for the subset of lncRNAs that are both adjacent to a HIF binding site and are detected in the polyA RNA-seq analyses. **B)** Gene Set Enrichment Analysis showing enrichment of lncRNAs adjacent to HIF-binding sites amongst those downregulated by combined HIF-1α and HIF-2α siRNAs in hypoxia. The same analysis was repeated for **C)** HIF-1α siRNA and **D)** HIF-2α siRNA alone. RNA-seq and HIF ChIP-seq genome browser tracks for the two most hypoxia-upregulated HIF-binding lncRNAs **E)** MALAT1 and **F)** NEAT1.

3.3.4 MALAT1 is a HIF regulated lncRNA

The previous analysis identified NEAT1 and MALAT1 as the two most highly regulated lncRNAs that are directly induced by HIF in hypoxia. In Chapter 5 I will focus on a detailed functional analysis of the HIF-dependent regulation of NEAT1. However, here I confirm the pan-genomic findings at the MALAT1 locus using gene-specific techniques. Elevated expression of MALAT1 has been reported in number of cancers (Xu et al., 2011, Gutschner et al., 2013). However, its regulation in hypoxia has not been investigated. I therefore screened a panel of breast cancer cells representing different breast cancer sub-types (ER+, ER- and triple receptor negative) for hypoxic regulation of MALAT1. Statistically significant hypoxic upregulation of MALAT1 was seen in the majority (8/11) of the cell lines studied **(Figure 3.5A).** The highest induction was observed in MCF-7 and ZR-75-1 cells, with mean fold-increases in hypoxia of 2.4 and 2.3, respectively. In addition, analysis in MCF-7 cells following different durations of 1 % hypoxia (16, 24, 32 and 48H), showed significant increases in MALAT1 expression at 24, 32 and 48H when compared to normoxic levels **(Figure 3.5B).** Statistical analysis showed no significant further increase in MALAT1 induction after 24 hours (ANOVA P-value= 0.80).

HIF is downregulated by a family of 2-oxoglutarate-dependent dioxygenases and I therefore determined whether inhibition of these enzymes would induce MALAT1 expression. MCF-7 cells were treated with dimethyloxalylglycine (DMOG) (500 nM, 24 hours) in normoxia. Expression of the known HIF-regulated gene, *carbonic anhydrase IX (CA9)*, was measured as a positive control to confirm effectiveness of the inhibitor. In addition, MALAT1 expression was significantly induced in MCF-7 cells (2 fold, P=0.001), suggesting that MALAT1 is regulated by a 2-oxoglutarate-dependent process **(Figure 3.5C).** PolyA RNA-seq analysis at the MALAT1 locus showed downregulation of MALAT1 RNA following either HIF-1α knockdown or HIF-2α knockdown (**Figure 3.4D).** These findings were confirmed by qPCR, and showed significant MALAT1 downregulation after HIF-1α siRNA (>50 %, p=0.001), while HIF-2α siRNA induced a more minor reduction of MALAT1 level that did not reach statistical significance **(Figure 3.5E).**

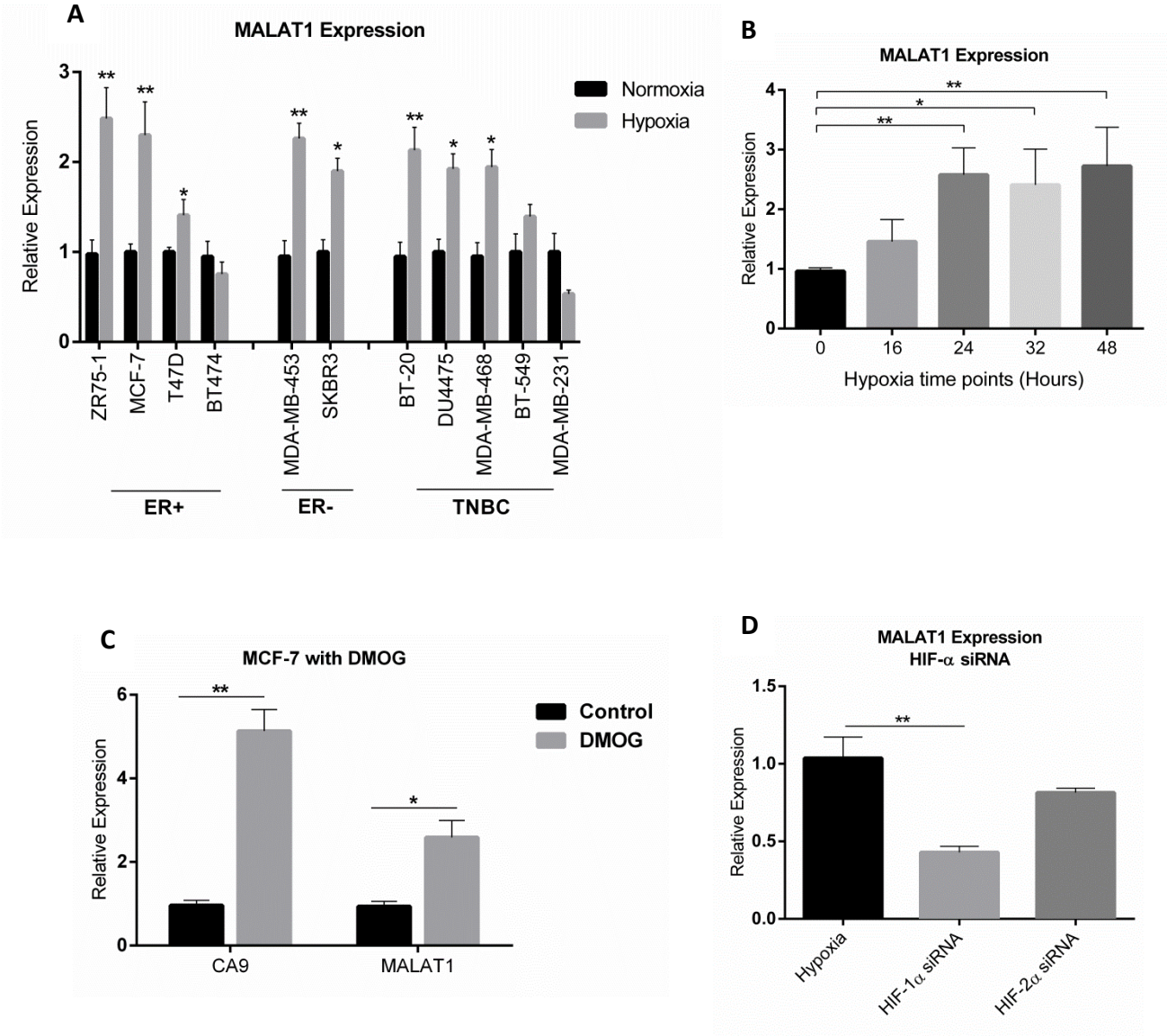


Figure 3.5 Regulation of MALAT1 in hypoxia. QPCR analyses **A)** MALAT1 expression in a panel of breast cancer cell lines grown in normoxia and hypoxia (1 % O_2 for 24 hours). The majority of cell lines showed upregulation of MALAT1 when exposed to hypoxia. **B)** Expression of MALAT1 at different hypoxia time points (16, 24, 32 and 48 hours) in MCF-7. **C)** CA9 and MALAT1 are significantly upregulated in MCF-7 after DMOG treatment **D)** MALAT1 expression was significantly reduced upon HIF-1α siRNA and showed slight decrease with HIF-2α knockdown. (*p≤0.04, **p≤0.009, Student's t-test, n=3).

3.3.5 Transcriptional regulation of non-annotated transcripts by HIF

The ribosome-depleted total RNA-seq analysis of hypoxic and normoxic MCF-7 cells identified 37 non-annotated intergenic transcripts and 54 non-annotated antisense transcripts (*In Chapter 2*). All of these transcripts were also detected in the polyA RNA-seq experiments, indicating that they were all poly-adenylated. I next examined the regulation of these transcripts following treatment with HIF-α siRNA. Heat maps of non-annotated intergenic transcripts and antisense transcripts show that, while induction of some of the non-annotated antisense transcripts required HIF, the large majority of the hypoxia-induced non-annotated intergenic transcripts were downregulated by either HIF siRNA **(Figure 3.6A-B).** These data suggests that many of these non-annotated transcripts are directly or indirectly regulated by HIF and thus many of these genes have not been previously identified in databases derived from normoxic cells.

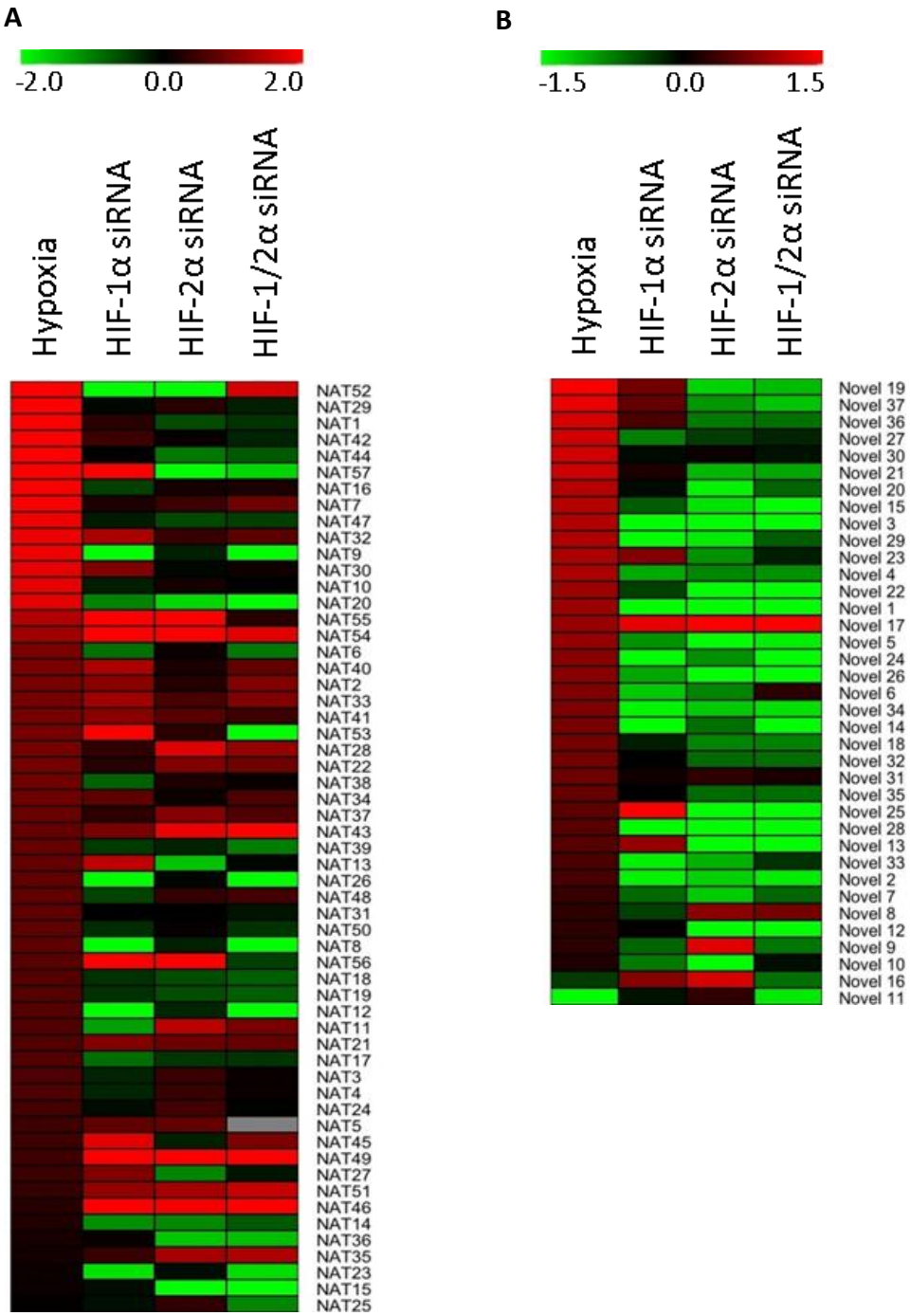


Figure 3.6 Transcriptional regulation of non-annotated transcripts by HIF. Heat maps showing fold regulation (log_2) by hypoxia and by the indicated HIF siRNA for **A)** Non-annotated antisense transcripts **B)** Non-annotated intergenic transcripts. Large numbers of hypoxia-induced transcripts were downregulated following HIF siRNA.

3.3.6 aSPAG4 and aCELACR2 are hypoxia antisense transcripts

Two of the most highly induced non-annotated antisense transcripts in hypoxia that had neighbouring HIF-binding sites and were regulated by HIF were aSPAG4 and aCELSR2. These were therefore selected for further characterisation**.** RNA-seq analysis showed that both aSPAG4 were co-regulated with the sense gene SPAG4 (both are hypoxia-upregulated), while induction of aCELSR2 was counter-regulated with the sense CELSR2 gene in hypoxia.

Both antisense transcripts were detected in the ribosome-depleted, total RNA-seq and the polyA RNA-seq datasets, confirming they are polyadenylated transcripts. 3' Rapid Amplification of cDNA Ends (RACE) PCR coupled with Sanger sequencing was used to examine the exonic structure and length of the transcripts. A series of primers were designed across the bodies of the two transcripts that were predicted by the RNA-seq analyses. This revealed that both aSPAG4 and aCELSR2 contain a single large exon which for aSPAG4 is 11,395 bp in length and for aCELSR2 is 7,543 bp long.

I next examined the hypoxic regulation of both the antisense and sense transcripts for each gene in a panel of breast cancer cell lines. Statistically significant hypoxic induction of both *SPAG4* and *aSPAG4* was observed in all breast cancer cell lines studied **(Figure 3.7A)**. Conversely, aCELSR2 was detected in only 4 cell lines, whereas, CELSR2 was expressed in all the examined cell lines. The four cell lines with detectable aCELSR2 all showed hypoxic induction of the aCELSR2 with concomitant downregulation of the sense transcript, whilst the cell lines without aCELSR2 did not show downregulation of the sense transcript **(Figure 3.7B).**

In addition, hypoxic induction of aSPAG4, aCELSR2 was measured in MCF-7 cells following differing durations of hypoxia. Both antisense transcripts were significantly induced after 16 hours of hypoxia and essentially maximal by 24 hours with sustained levels thereafter **(Figure 3.7C-D).**

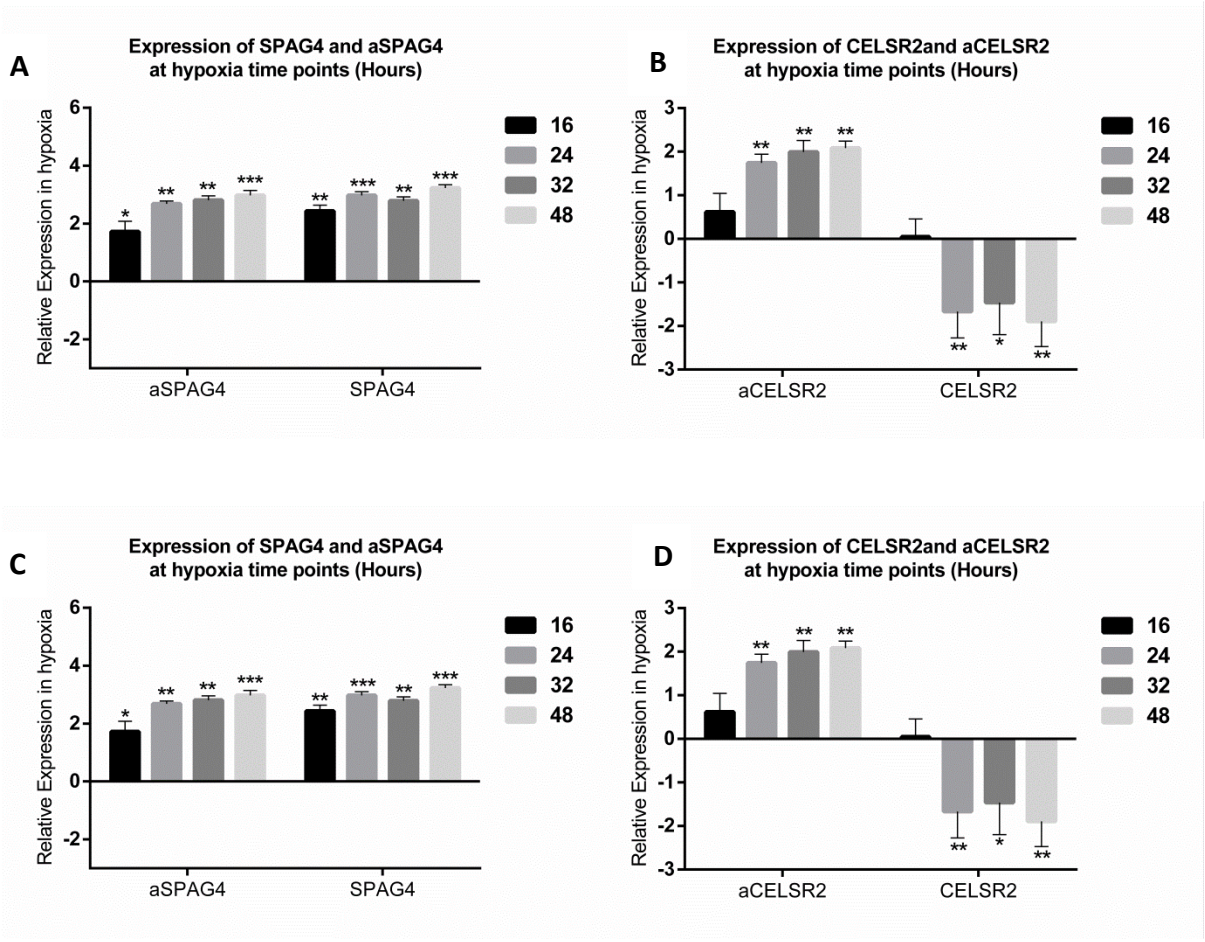


Figure 3.7 Regulation of aSPAG4 and aCELACR2 by hypoxia. Expression of antisense and sense transcripts in a panel of breast cancer cell lines grown in normoxia and hypoxia was measured by qPCR for **A)** SPAG4 and aSPAG4 and **B)** CELSR2 and aCELSR2. **C)** SPAG4 and aSPAG4 and **D)** CELSR2 and aCELSR2 expression in MCF-7 cells following differing durations of hypoxia. (*p≤0.04, **p≤0.006 ***p≤0.0005, Student's t-test, n=3).

3.3.7 aSPAG4 and aCELACR2 are HIF-1 dependent antisense transcripts

Inspection of the HIF ChIP-seq tracks at the SPAG4 and CELSR2 loci revealed that both transcripts have HIF binding sites **(Figure 3.8A-B)**. Furthermore, inspection of the polyA RNA-seq analyses of hypoxic MCF-7 cells treated with HIF siRNAs also showed that the hypoxic induction of the two antisense transcripts was dependent upon HIF. This was confirmed by qPCR analysis of aSPAG4 and aCELSR2 levels in hypoxic MFC-7 cells treated with siRNAs to suppress the individual HIF subunits. Both aSPAG4 and SPAG4 were significantly downregulated (4-fold and 3.8-fold respectively) in HIF-1α siRNA transfected cells, whereas, HIF-2α knockdown showed no effects on their expression **(Figure 3.8C)**. Similarly, expression of aCELSR2 was significantly downregulated 3-fold by HIF-1α knockdown, whilst CELSR2 expression was slightly increased (1.8-fold) by HIF-1α knockdown **(Figure 3.8D)**. HIF-2α knockdown again induced no significant change in aCELSR2 or CELSR2 expression. The above findings indicate that aSPAG4, SPAG4, aCELSR2 and CELSR2 are all HIF-1 regulated transcripts. However, from the current analyses it is not possible to determine whether all are direct transcriptional targets of HIF-1 or whether HIF-transcriptional activation of the antisense transcript regulates the sense transcript *in cis* at each locus or *vice versa*.

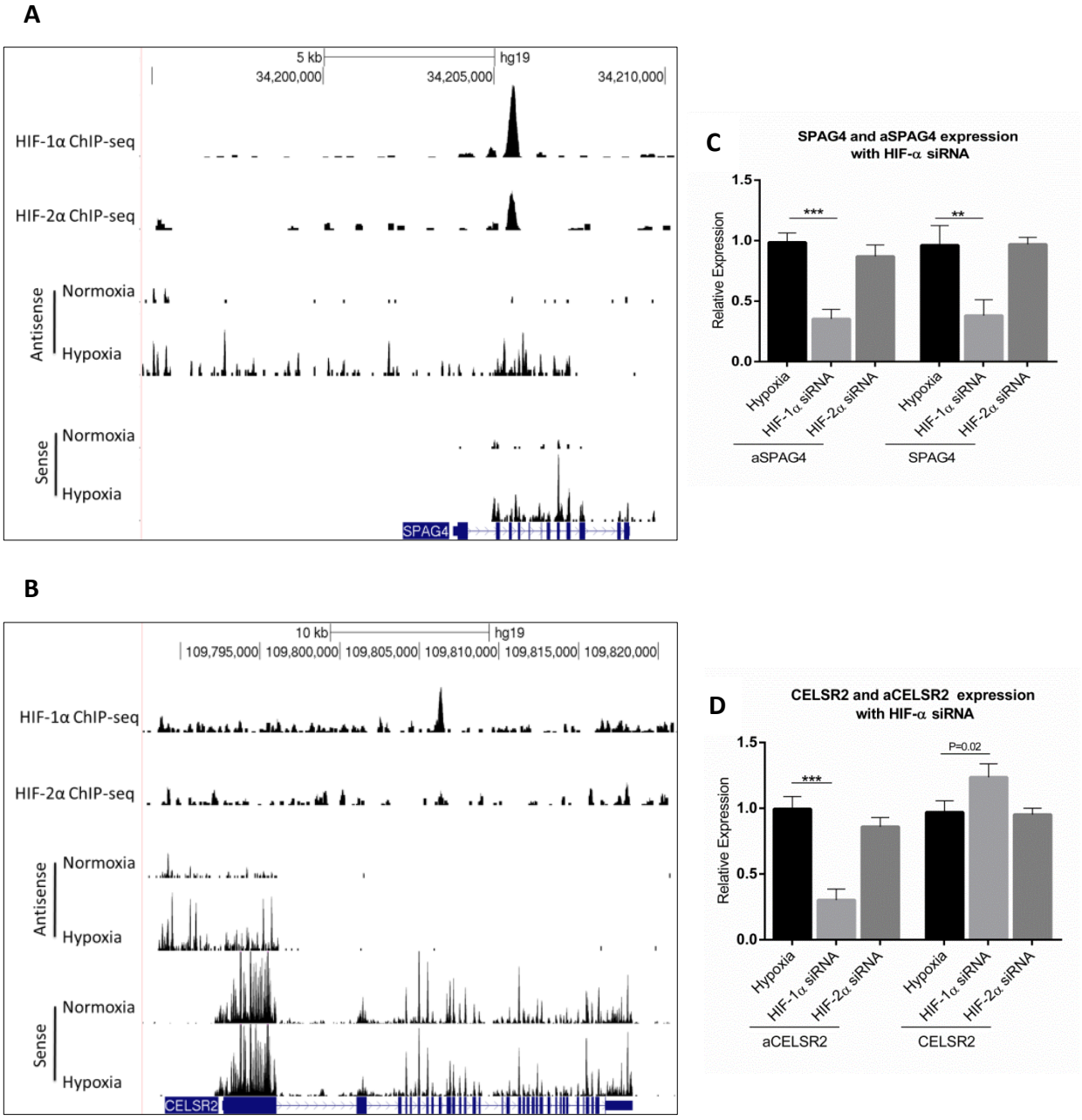


Figure 3.8 aSPAG4 and aCELACR2 are HIF dependent antisense transcripts. Snapshot of ribosome-depleted strand-specific RNA-seq tracks using the UCSC genome browser at the **A)** aSPAG4 and **B)** aCELSR loci, showing increased expression of both antisense transcripts in hypoxic MCF-7. qPCR was performed following knockdown of the indicated HIF isoform by siRNA in hypoxic MCF-7 cells **C)** SPAG4 and aSPAG4 expression were both significantly downregulated by HIF-1α knockdown but not by HIF-2α suppression, suggesting that both SPAG4 and aSPAG4 are HIF-1α dependent, **D)** aCELACR2 was downregulated and CELSR2 expression was upregulated following HIF-1α knockdown, but were unaffected by HIF-2α suppression, again suggesting that both aCELSR2 and CELSR2 are also HIF-1α dependent (*p=0.02, **p=0.008 ***p=0.0006, Student's t-test, n=3).

3.3.8 Cellular localisation and coding potential of aSPAG4 and aCELACR2

While coding mRNAs are predominantly cytoplasmic in location, non-coding RNAs are frequently nuclear in their localisation, consistent with a regulatory role in transcription. I therefore examined the cellular localisation of aSPAG4 and aCELSR2. Cytoplasmic and nuclear fractions were isolated from MCF-7 cells after incubating in normoxia and in hypoxia (1 % O_2 for 24 hours). To ensure effective isolation, immunoblotting was performed on each fraction for the cytoplasmic marker Glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and the nuclear marker histone H3 **(Figure 3.9A)**. qPCR analysis of cytoplasmic and nuclear fraction revealed that both aSPAG4 and aCELSR2 are exclusively nuclear transcripts **(Figure 3.9B-C)**. In addition, RNA fluorescent *in situ* hybridization (FISH) was performed on MCF-7 cells using probes designed against aSPAG4 and aCELSR2. Analysis confirmed hypoxic induction of both transcripts in the cell nucleus **(Figure 3.10).**

Both antisense transcripts are localised in the nucleus, suggesting that they are not translated into proteins. It is thus highly probable that they belong to the long non-coding RNA class of transcripts. *In silico* analysis using protein translation tools such as ExPASy was used to evaluate the protein coding potential of aSPAG4 and aCELSR2. All possible translation frames for each transcript contain multiple stop codons inconsistent with translation of a meaningful peptide. Furthermore, negative coding potential calculator (CPC) scores were obtained for both transcripts (aSPAG4 = -0.98 and aCELSR2 =-1.25) and are summarised in **Table 3.1.** The cellular localization and different *in silico* analyses strongly suggest that aSPAG4 and aCELSR2 are long non-coding, antisense transcripts, which are regulated by HIF in hypoxia.

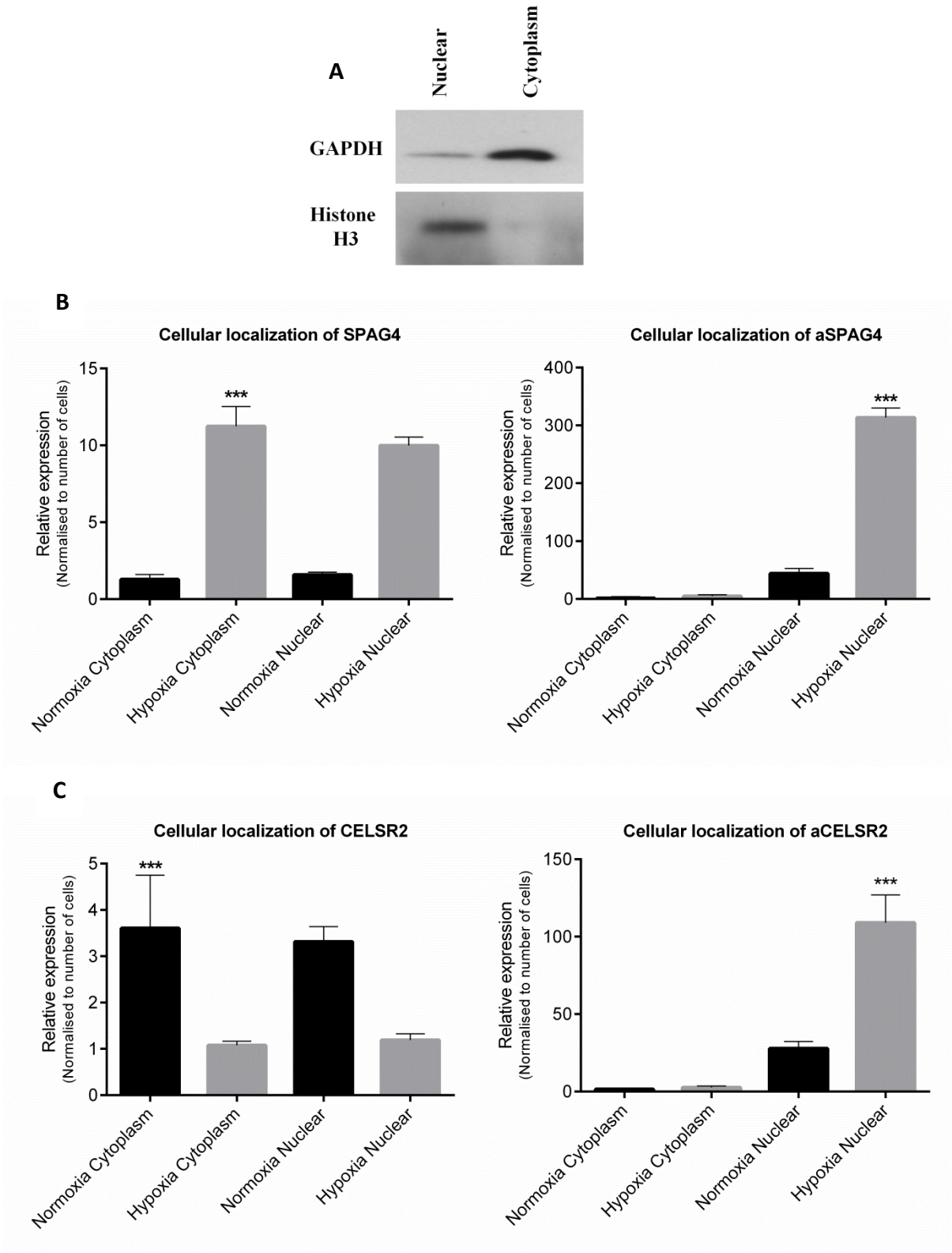


Figure 3.9 Cellular localisation of aSPAG4 and aCELACR2 in hypoxia. A) Immunoblot of nuclear and cytoplasm extracts from hypoxic MCF7 cells. To ensure affective isolation, immunoblotting was performed on isolated fractions with GAPDH (cytoplasmic marker) and histone H3 (nuclear marker) antibodies to confirm effective isolation of fractions. qPCR was performed on nuclear and cytoplasmic fractions from MCF-7 cells grown in hypoxia (1 % O_2 for 24 hours) and in normoxia **B)** Expression of SPAG4 gene was detected in both fractions and enriched in the cytoplasm, whereas, aSPAG4 was only detected in nuclear fractions in hypoxia. Similarly **C)** CELSR2 gene was detected in both fraction and was down regulated in hypoxia, whereas, aCELSR2 was induced in hypoxia and only detected in nuclear fractions (***p<0.0005, Student's t-test, n=3).

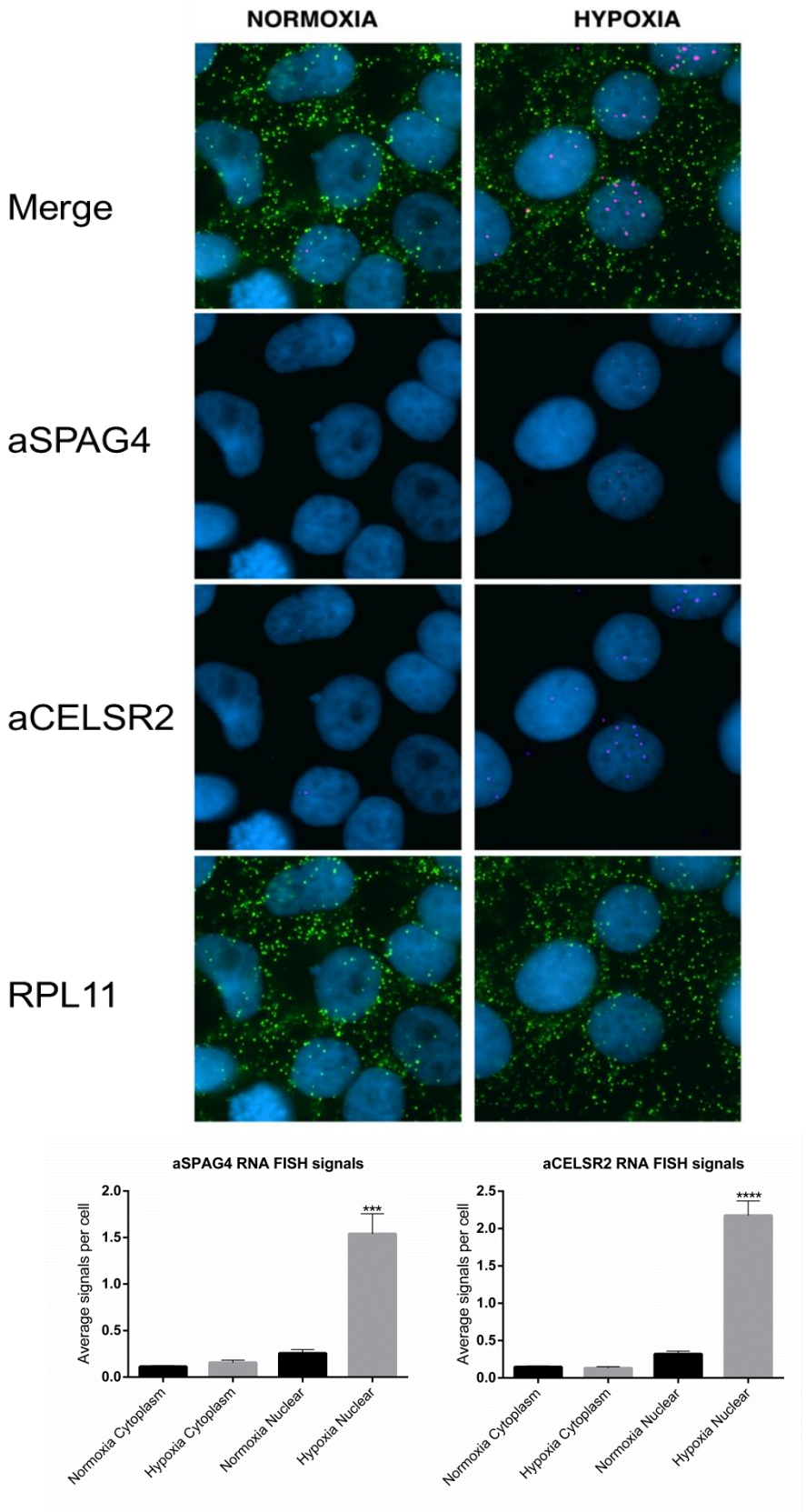


Figure 3.10 RNA FISH for aSPAG4 and aCELSR2 in MCF-7 cells. RNA FISH probes were designed against aSPAG4 and aCELSR2 transcripts and hybridized to MCF-7 cells in normoxia and hypoxia (1 % O_2 for 24 hours). RPL11 probe was used as control and unchanged in normoxia and hypoxia. The transcripts were only detected in hypoxia and confirm their nuclear localisation. aSPAG4 probe (red signal), aCELSR2 probe (pink signal) and RPL11 control RNA probe (green signal). Below bar charts showing average signals of aSPAG4 and aCELSR2 per cell in hypoxia and normoxia (***p<0.0005,****p<0.00005 Student's t-test, n=3).

Table 3.1 Summary of *in silico* analysis of two selected antisense transcripts detected in hypoxic MCF-7 cells:

Features	aSPAG4	aCELSR2
Chromosomal location	chr20:34,196,650-34,207,240	chr1:109788620;109796162
Strand	- strand	- strand
Transcript size	11395 bp	7543 bp
Number of exons	One large exon	One large exon
Coding potential (ExPASy tool)	NO coding frames	NO coding frames
Coding potential (CPC score)	-0.98	-1.25
3' Tail	Poly A	Poly A
Transcript localisation	Nucleus	Nucleus
Expression of original gene in hypoxia	Upregulated	Downregulated
Position to original gene	5' of original gene	5' of original gene
HIF binding sites	HIF-1α and HIF-2α	HIF-1α and HIF-2α
siRNA HIFs (Downregulation)	HIF-1a and HIF-1/2α siRNA	HIF-1α siRNA

3.4 Discussion

This chapter focused on the correlation between HIF-binding sites across the genome (derived from previous ChIP-seq date) and expression analysis of the coding and non-coding transcriptome by ribosome-depleted strand-specific total RNA-seq and polyA RNA-seq. This included RNA-seq data from MCF-7 cells incubated in normoxia or hypoxia, or in hypoxia following suppression of HIF-1α, HIF-2α or combined HIF-1/2α by siRNA. It has allowed, for the first time, a pan-genomic analysis of the role of the HIF transcription factor in the regulation of all classes of RNA.

The HIF-1α, HIF-2α and HIF-1β ChIP-seq datasets have been previously published (Schodel et al., 2011) and correlated with microarray analyses of mRNA regulation in hypoxia. In the current study, the HIF ChIP-seq raw data were remapped and analysed based on human genome HG19. This was the reference genome used for the RNA-seq analysis, since it contains a more comprehensive annotation of the genome than HG18. Since HIF-α binds chromatin as a HIF-α/β heterodimer, HIF-1α and HIF-2α binding sites were filtered for the presence of HIF-1β signal to define a set of high-stringency sites as per Schoedel et al. 2010. Using these criteria, 500 HIF-1 or HIF-2 binding sites were annotated across the genome.

Previous studies reported that HIF-α binding sites were observed hundreds of kilobases distant from any documented mRNA promoter region. However, it was possible that these "promoter-distant" binding sites lay close to a non-coding or previously non-annotated gene. Indeed, several studies have reported HIF binding at specific noncoding genes, such as microRNAs (Blick et al., 2013, McCormick et al., 2013). However, to date there is no reported genome wide association of HIF-binding sites with different non-coding RNA classes. I therefore reanalysed the relationship between HIF-binding sites and promoters using all active transcriptional start sites identified in my ribosome-depleted strand-specific total RNA-seq datasets. Whilst some of the HIF-binding sites that were distant from an mRNA promoter were bound close to a non-

coding or non-annotated transcripts, many were not. The majority of HIF-1 and HIF-2 binding sites were still found to lie in excess of 2.5-kb from the nearest active promoter irrespective of the class of gene. In particular, HIF-2 binding sites were generally more distant from any active promoter than were HIF-1 binding sites. This analysis further supports the hypothesis that HIF (and HIF-2 in particular) has an important role in long-range transcriptional regulation (Schodel et al., 2013).

Previous work has shown an association between the gene closest to a HIF-binding site and its functional regulation by hypoxia and by HIF, irrespective of its distance from the binding site (Schoedel et al., 2010). I therefore postulated that if HIF were involved in the transcriptional regulation of specific classes of RNA, then at least some of these would be the closest gene to a HIF-binding site. Importantly, non-coding RNAs of the piwiRNA, snRNA and tRNA classes were virtually never the closest transcript to a HIF-binding site. Furthermore, these classes of RNA were globally downregulated in hypoxia. This suggests that they are unlikely to be direct transcriptional targets of HIF. In contrast, HIF-binding was enriched in the vicinity of both mRNAs and lncRNAs. Furthermore, mRNAs and lncRNAs that were closest to HIF-binding sites were likely to be upregulated suggesting that HIF directly enhances (but does not repress) the expression of both these classes of RNA. In addition, HIF binding was also observed close to miRNAs and non-annotated transcripts. Whilst some of these were again upregulated by hypoxia, the numbers were too small to statistically correlate with gene expression. Together, these data indicates that mRNAs, lncRNAs, miRNAs and some of the previously non-annotated transcripts can all be directly activated by the HIF transcription factor, providing new insights into its role in the response to hypoxia.

Previous work has shown that, although HIF-1 and HIF-2 binding sites overlap greatly, there are unique patterns of binding. These manifest as clear functional differences between the genes closest to binding sites of each isoform as defined by gene ontology and pathway analyses. For instance, HIF-1 predominantly binds and regulates genes involved in carbohydrate metabolism,

while HIF-2 predominantly binds genes involved in stem cell pluripotency (Covello et al., 2006, Schodel et al., 2011). In my analyses, HIF-2 bound close to lncRNAs more frequently than HIF-1 binding sites, indicating a potentially new isoform-specific function in regulation of the non-coding transcriptome.

To determine the extent to which genes are regulated by HIF-1, HIF-2 or both isoforms, polyA RNA-seq was performed on hypoxic MCF-7 cells, after HIF-1α, HIF-2α and combined HIF-1/2α knockdown. This approach detected many more HIF-regulated mRNA genes and in particular, HIF-2 regulated genes, than a previous microarray analysis including all of the previously described genes (Elvidge et al., 2006). In addition, it detected about 25 % of the lncRNA transcriptome. Polyadenylated lncRNAs whose gene loci bound HIF were regulated by both HIF-1 and HIF-2 isoforms, although the numbers were too small to permit a statistical analysis of any differences. Nevertheless, it does confirm that the large majority of hypoxia-regulated, HIF-binding lncRNAs that were examined are also regulated by HIF.

These analyses revealed that MALAT1 is one of the most highly induced lncRNAs in hypoxia and is a HIF target lncRNA (HIF binding and HIF regulation). High expression of MALAT1 has been reported in number of cancers including non-small cell lung, colorectal and prostate cancers (Weber et al., 2013, Ji et al., 2013, Ren et al., 2013) and is associated with metastasis and recurrence (Gutschner et al., 2013, Xu et al., 2011). It has also been reported that MALAT1 regulates pre-mRNA splicing factors, including serine arginine dipeptide-containing SR family splicing factors (Tripathi et al., 2010, Polymenidou et al., 2011). Increased expression of MALAT1 enhances tumour proliferation and growth both *in vitro* and *in vivo* (Han et al., 2013, Tripathi et al., 2013). In the current study, expression analysis in a panel of breast cancer cell lines showed widespread induction of MALAT1 in hypoxia. Moreover, treatment with DMOG and knockdown of HIF isoforms confirmed that MALAT1 is a predominantly HIF-1 regulated lncRNA. Overall, these finding suggest that MALAT1 induction is regulated by HIF-1 and is

common when breast cancer cells are exposed to hypoxia. Further investigation is required to understand the functional role of MALAT1 in hypoxia in these cells.

In addition to lncRNAs, a number of previously non-annotated transcripts, including antisense RNAs, also bound HIF and were regulated by both hypoxia and by HIF. Antisense transcripts have been proposed to have potential roles in regulating gene expression *in cis*. Two HIF-binding and HIF-regulated antisense transcripts that demonstrated concurrent regulation of the overlapping sense transcript (aSPAG4 and aCELSR2) were selected for further validation. SPAG4 is a SUN domain-containing protein (Tzur et al., 2006, Razafsky and Hodzic, 2009) and was initially identified as a testis-specific protein, localised to microtubule-containing spermatid structures (Shao et al., 1999, Tarnasky et al., 1998). Recent studies reported that suppression of SPAG4 reduces the cellular proliferation of cancer cells and is a potential prognostic marker in renal cell cancer (RCC) (Shoji et al., 2013). Moreover, SPAG4 knockdown reduces the invasion and overexpression leads to increase in migration capability of RCC cells *in vitro* (Knaup et al., 2013). A number of studies have reported hypoxic induction of SPAG4 in multiple cell lines (Knaup et al., 2013, Shoji et al., 2013, Knowles et al., 2010). In addition to confirming this finding, the current study also identified a previously non-annotated antisense transcript (aSPAG4) transcribed from the negative strand and overlapping the 5' end of the SPAG4 gene. Both aSPAG4 and SPAG4 were upregulated by hypoxia in all the breast cancer cell lines studied. Furthermore, the expression of both the sense and antisense transcript in hypoxia is dependent on HIF-1α. *In silico* analysis together with the cellular localisation of aSPAG4 indicates that aSPAG4 is likely a nuclear long non-coding RNA. However, whether SPAG4 and aSPAG4 are both transcriptionally regulated by HIF directly or whether one transcript regulates the other in *cis* remains to be determined.

In contrast to the situation for SPAG4 and aSPAG4, the antisense transcript, aCELSR2, is induced in hypoxia whilst the sense CELSR2 transcript is reduced. Low or absent expression of CELSR2 has been reported previously in breast cancer (Huang et al., 2005). In my experiments, aCELSR2

transcript is expressed in limited breast cell lines. However, those cell lines in which aCELSR2 was detected all showed increased expression in hypoxia with concomitantly reduced expression of CELSR2. The suppression of individual HIF-α isoforms in hypoxia showed that the aCELSR2 transcript is HIF-1α-dependent. Interestingly, suppression of aCELSR2 by HIF-1α knockdown is associated with increased expression of CELSR2, suggesting that aCELSR2 may suppress CELSR2 expression in *cis*. Whether this contributes to the suppression of CELSR2 observed in breast cancer requires further evaluation.

Overall, this section provides the most comprehensive characterisation of the HIF-dependent transcriptional response to hypoxia to date by integrating HIF-binding sites (ChIP-seq), ribosome-depleted, strand-specific, total RNA-seq analysis of the hypoxia-regulated transcriptome analysis and polyA RNA-seq analysis of the HIF siRNA-dependent transcriptome. These findings indicate a major role for HIF in the direct transcriptional regulation of mRNA and non-coding RNAs, including lncRNA, antisense RNA and miRNA in response to hypoxia.

Chapter 4: Correlation of epigenetic markers and RNApol2 with hypoxia transcriptome

4.1 Introduction

Genome-wide analyses of histone modifications and RNA polymerase activity have been made possible by pan-genomic projects such as ENCODE and by advances in next generation sequencing technologies. These have revealed new insights into the epigenetic features of regulatory elements of active and inactive genes in different cell types under different conditions. Changes in epigenetic features do not alter DNA sequences, but rather result in modifications that control and regulate gene expression in cells by several mechanisms, including nucleosome modelling, histone modification, DNA methylation and regulation by non-coding RNA (Dawson and Kouzarides, 2012). Histone modifications include acetylation, methylation, phosphorylation and ubiquitination (Li et al., 2007), which occur on specific amino acids in the tails of various histones and are associated with altered gene expression. For instance, acetylation is generally a marker of gene activation, whereas deacetylation is a marker of transcription repression. Alternately, methylation of specific amino acids of histone tails can lead to different effects on transcription. For example, histone 3 lysine 4 methylation (H3K4me, me2 and me3) is a marker of gene activation, whereas histone 3 lysine 9 methylation (H3K9me2 and me3) and H3K27me3 are markers of transcription repression (Li et al., 2007, Tsai and Wu, 2013).

It has been reported that a number of chromatin reprogramming events such histone methylation, acetylation or deacetylation occur in hypoxia and contribute to transcriptional regulation (Johnson and Barton, 2007). These modifications may lead to the recruitment of transcription factors or histone re-modellers, such as histone acetyltransferases (HATs), histone deacetylases (HDACs) and histone methyltransferases (HMTs), which are crucial for transcriptional regulation (Johnson and Barton, 2007). In addition, some of these modifying enzymes are regulated by HIF itself, including the Jumonji-domain-containing proteins (JMJD1A, JMJD2B and JMJD2C), which decrease demethylation of histone tails, histone methyltransferases (MLL5) and histone acetyltransferases (SAP30) (Beyer et al., 2008, Krieg et al., 2010, Pollard et

al., 2008, Xia et al., 2009). Several studies have described bulk changes in histone marks such as histone 3 lysine 4 trimethylation H3K4me3, H3K9me2, H3K36me3 and H3K27me3 around promoters of hypoxia-regulated genes (Johnson et al., 2008, Zhou et al., 2010, Xia et al., 2009, Tsai and Wu, 2013). For example, Johnson et al. reported changes in chromatin modifications including an increase in H3K4me3 levels and a decrease in H3K27me3 levels around promoters of hypoxia-regulated genes (Johnson et al., 2008).

On the other hand, little is known about the molecular mechanisms of RNA polymerase (RNApol2) activity in hypoxia. A recent study by Galbraith and colleagues reported that promoters of hypoxia-inducible genes carry proximally-paused RNApol2 in normoxia and that HIF-1α stabilisation in hypoxia induces binding of CDK8 to its target genes and interaction with CDK8 mediator, which in turn stimulates release of the paused RNApol2 (Galbraith et al., 2013).

In this chapter I have correlated RNApol2 and histone modification (H3K4me3) ChIP-seq data in hypoxia and normoxia (raw data provided by Dr David Mole) with my hypoxic transcriptome data from the total RNA-seq experiment. I have identified patterns of RNApol2 and H3K4me3 signals around the promoters of transcripts and correlated these with their regulation by hypoxia and/or by HIF. In particular, I have identified the distribution of RNApol2 across the genome in hypoxia and extended the analysis of these markers both to the lncRNAs class of RNA and to previously non-annotated transcripts, since all the previous reports have focused on mRNAs. These analyses provide further insights into our understanding of the transcriptional regulation of non-coding RNAs through the control of RNApol2 and histone modification by HIF.

4.2 Exprimental desigen

The regulation of different classes of RNA by hypoxia obtained from the total RNAs-seq was correlated with ChIP-seq of RNApol2, H3K4me3 and DNAse I hypersensitive sites derived from MCF-7 cells in hypoxia (1 % O_2 for 24 hours) and normoxia. Dr. David Mole and his lab provided the raw ChIP-seq data for RNApol2 and H3K4me3. DNAse I hypersensitive sites data were obtained from ENCODE v.7. All the raw data were mapped in the human genome assembly (HG19) using Bowtie v.1.0.0 (Langmead et al., 2009). First, RNApol2 and H3K4me3 signals were correlated at transcription start sites (TSS) of 100 silent, intermediate and highly expressed transcripts in hypoxia. Then, RNApol2, H3K4me3 markers assessed around putative TSS of all non–annotated transcripts and similar analyses were performed after dividing non-annotated transcripts into non-annotated intergenic transcripts and antisense transcripts. Furthermore, the distribution of RNApol2, H3K4me3 and DNAse I hypersensitive sites was assessed at the TSS of the 100 most upregulated, downregulated and non-regulated transcripts by hypoxia. Similar analysis was performed for HIF dependent and non-HIF dependent upregulated transcripts in hypoxia. The traveling ratio of RNApol2 was calculated in normoxia and hypoxia. Finally, RNApol2 and H3K4me3 signals were assessed at the TSS of the individual RNA classes including mRNA, lncRNAs and miRNAs.

4.3 Results

4.3.1 Correlation of RNApol2 and histone ChIP-seq signals with gene expression

To gain mechanistic insights into the transcriptional regulation of the genome, in particular the non-coding transcriptome, I correlated my ribosome-depleted, strand-specific total RNA-seq data with pre-existing RNApol2 and H3K4me3 ChIP-seq (MTAB-1994) and DNAse hypersensitivity (ENCODE) datasets from MCF-7 cells incubated in normoxia and hypoxia. The raw data was mapped to NCBI build 37 (HG19) of the human genome as described in the method section.

As a preliminary check on data integrity, I selected the transcriptional start sites of 100 silent genes, 100 genes of intermediate expression level and the 100 most highly expressed genes. This selection was irrespective of gene class, although the vast majority were mRNAs and lncRNAs. Plotting the average distribution of signal for each set of 100 genes for RNApol2 and H3K4me3 **(Figure 4.1A)** showed similar spatial distributions and correlations with gene expression to those previously described (Barski et al., 2007) **(Figure 4.1B)**. Similar results were seen for both the normoxic and hypoxic datasets.

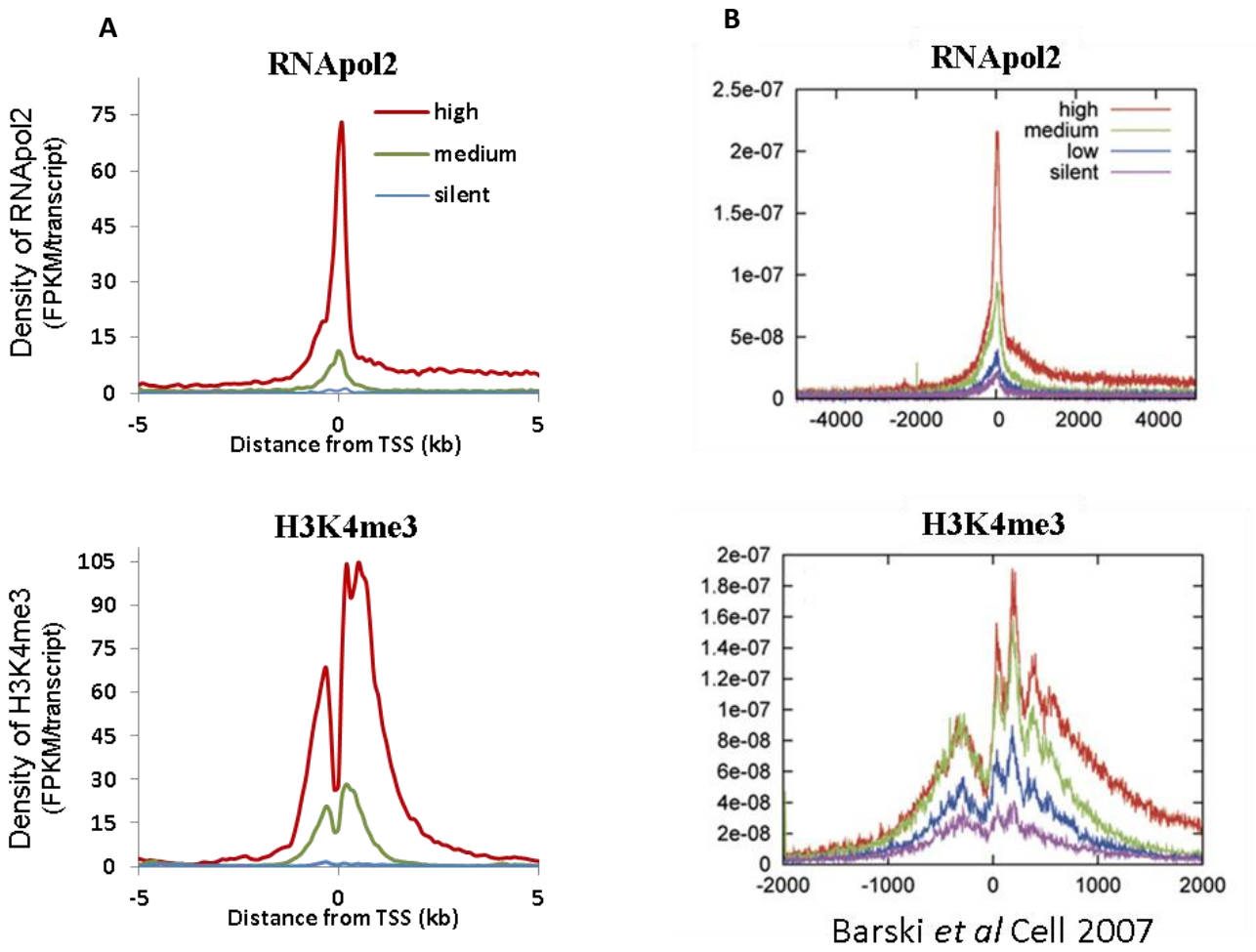


Figure 4.1 Distribution of RNApol2 and H3K4me3 signals around TSS of genes separated based on their expression levels. A) Marker distribution around TSS +/- 5 kb of 100 genes with high expression (red) or medium expression (green) or silent genes (blue) in normoxia. Signal patterns were similar to **B)** Distribution of same markers correlated with absolute gene expression in lymphocytes (Adapted from Barski et al., 2007).

4.3.2 RNApol2 and histone modifications of non-annotated transcripts

Total RNA-seq identified a number of previously non-annotated transcripts in MCF-7 cells. As an independent check, I first interrogated the RNApol2 and H3K4me3 ChIP-seq signals at the transcriptional start site of each of these putative RNAs. As anticipated, each of these promoter-associated marks was increased at the TSS when averaged across all of the newly identified transcripts **(Figure 4.2A-B).** Similar patterns were observed when these transcripts were subdivided into intergenic and antisense transcripts, providing independent corroborative evidence for the newly described transcripts **(Figure 4.2C)**. Examples of tracks at a non-annotated intergenic transcript **(Figure 4.3)** and at an antisense transcript of HIF-1α (aHIF-1α) are shown **(Figure 4.4)**. RNApol2 and H3K4me3 peaks were observed at both ends of HIF-1α transcript. Peak at the 5' end of HIF-1α correspond to the original transcript, while peak at 3' end of the transcript correspond to the antisense transcript **(Figure 4.4)**.

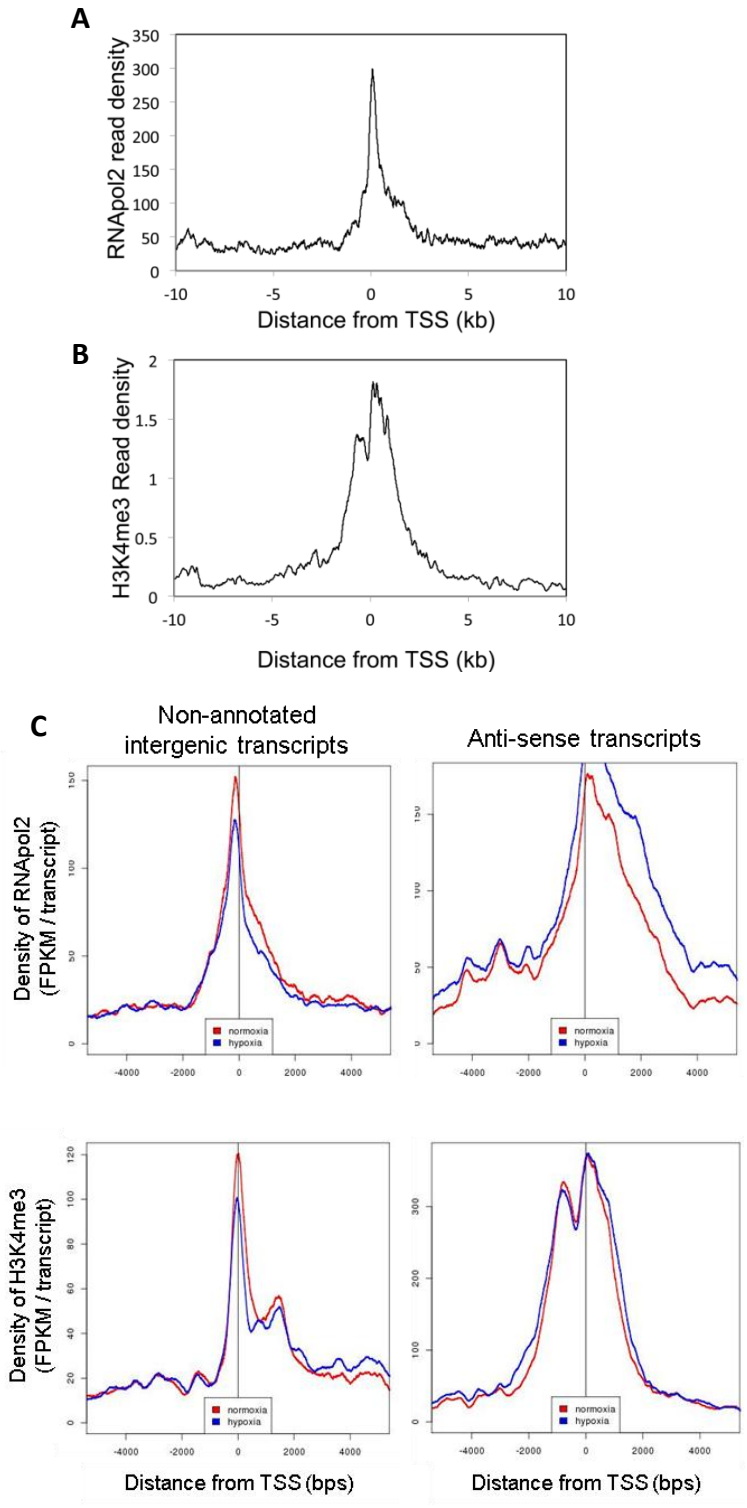


Figure 4.2 RNApol2 and H3K4me3 signals averaged over the putative TSS regions of previously non-annotated transcripts identified in the total RNA-seq. ChIP-seq signals for the promoter associated marks, **A)** RNApol2 and **B)** H3K4me3 at the putative TSS of non-annotated transcripts**. C)** Distribution of RNApol2 and H3K4me3 signals at TSS of all detected non-annotated intergenic transcripts and antisense transcripts in hypoxia and normoxia. Both class of transcripts showed slight increase in markers were obscured in hypoxia (blue) and normoxic (red).

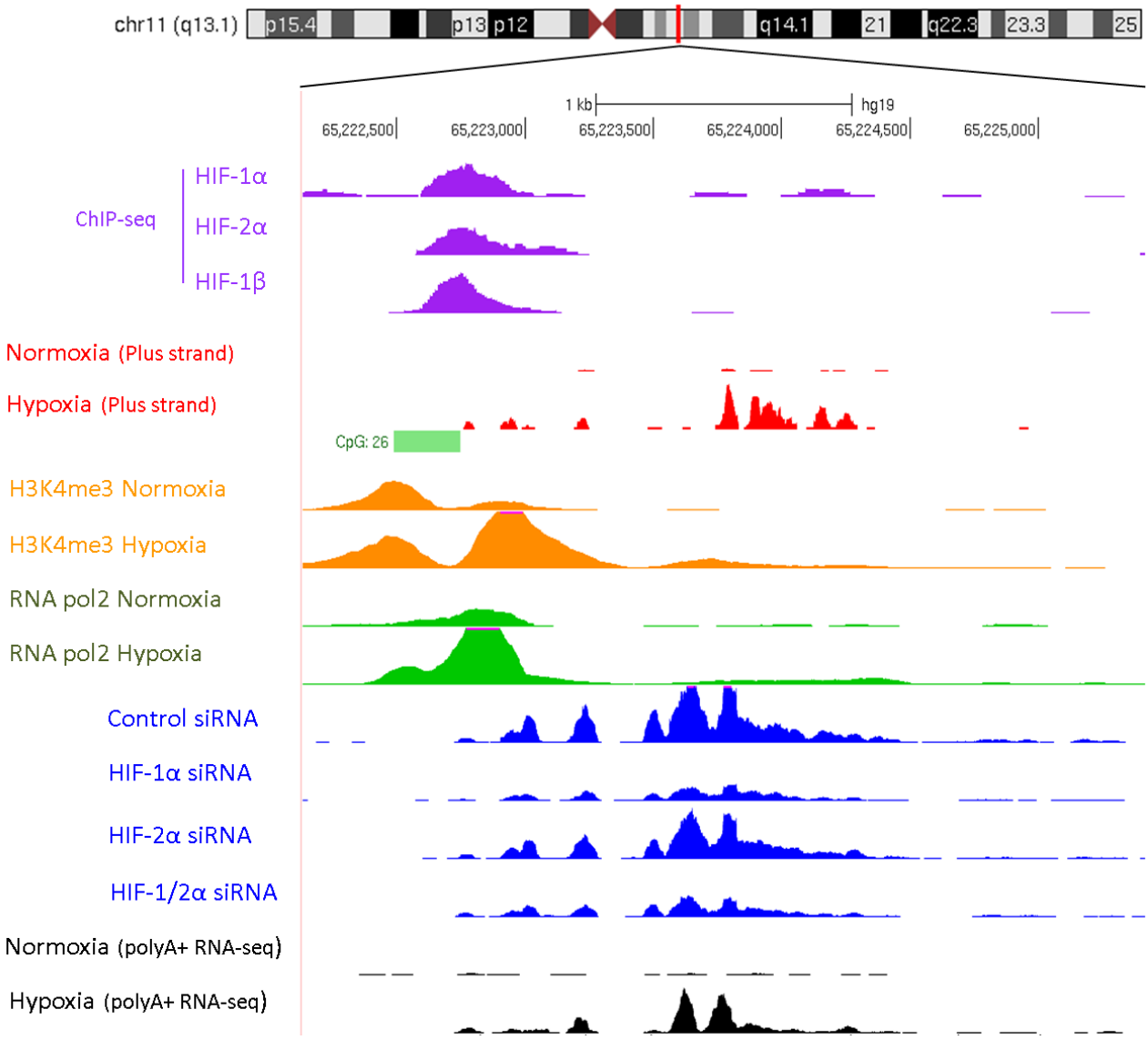


Figure 4.3 Integration of HIF, RNA Pol2, H3K4me3 ChIP-seq and polyA and total-RNA-seq data at non-annotated intergenic transcript locus. All the data were uploaded at the UCSC genome browser. Snapshot of HIF regulated non-annotated intergenic transcript showed increase in H3K4me3 level and run-through of RNApol2 in hypoxia. The transcripts was detected in different polyA RNA-seq datasets and was downregulated with HIF-1α siRNA, indicating that non-annotated intergenic transcript is HIF-1α-dependent.

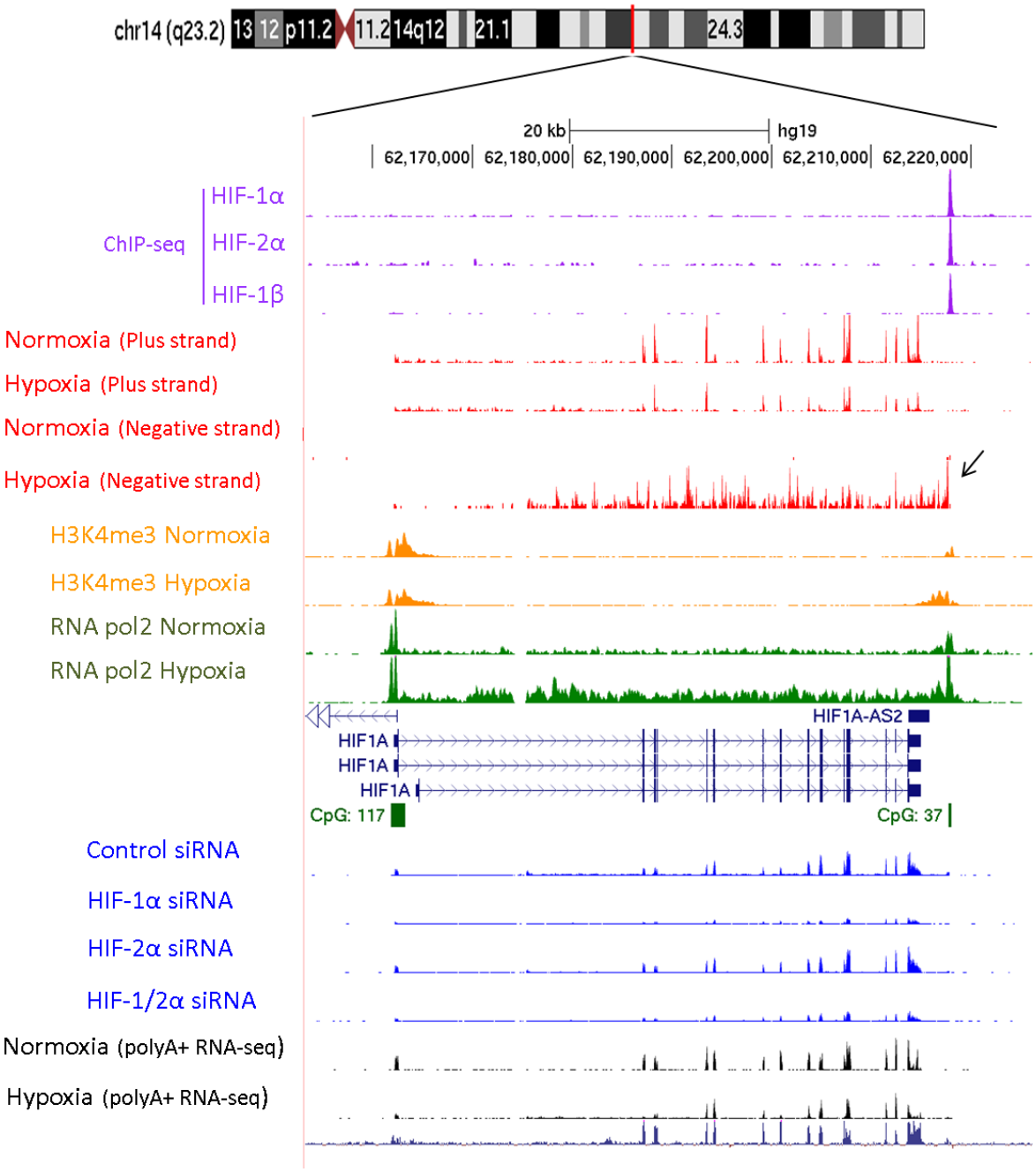


Figure 4.4 Integration of HIF, RNA Pol2, H3K4me3 ChIP-seq and polyA and total-RNA-seq data at antisense transcript of HIF-1α. All the data were uploaded at the UCSC genome browser. Snapshot of **a**ntisense transcripts of HIF-1α showed increase in H3K4me3 level and run-through of RNApol2 in hypoxia. The transcript was detected in different polyA RNA-seq datasets and was downregulated with HIF-1α siRNA, suggesting that antisense transcript is HIF-1α-dependent.

4.3.3 RNApol2 and H3K4me3 changes in hypoxia

From the total RNA-seq data I next identified the 100 most hypoxia-upregulated and downregulated transcripts irrespective of class. Plotting the average signal for the 100 most upregulated genes revealed distinct patterns for each RNApol2, H3K4me3 and DNAse hypersensitivity mark. Specifically, a strong DNAse hypersensitivity signal was observed at the TSS of these hypoxia-upregulated genes even in normoxia, with no significant further increase in hypoxia **(Figure 4.5A)**. Similarly, RNApol2 was already bound at the promoter in normoxic cells, again with no significant increase in RNApol2 signal at the TSS in hypoxia. The peak of RNApol2 signal was observed 40-45 bp downstream of the TSS, consistent with promoter-proximal pausing of RNApol. In contradistinction, RNApol2 signals downstream of the TSS (i.e. across the body of the gene) were increased in hypoxia at hypoxia-upregulated genes. This is consistent with the release of paused promoter-proximal RNApol2 to effect increased transcription **(Figure 4.5B)**. Actively transcribing RNApol2 can recruit histone methyl transferase activity that trimethylates H3K4 either directly (Ng et al., 2003) or through p300 (Tang et al., 2013). The distribution of H3K4me3 was very similar to that of RNApol2, with little change in signals at the promoter in hypoxia, but an increase downstream of the transcriptional start site, consistent with the above hypothesis **(Figure 4.5C)**.

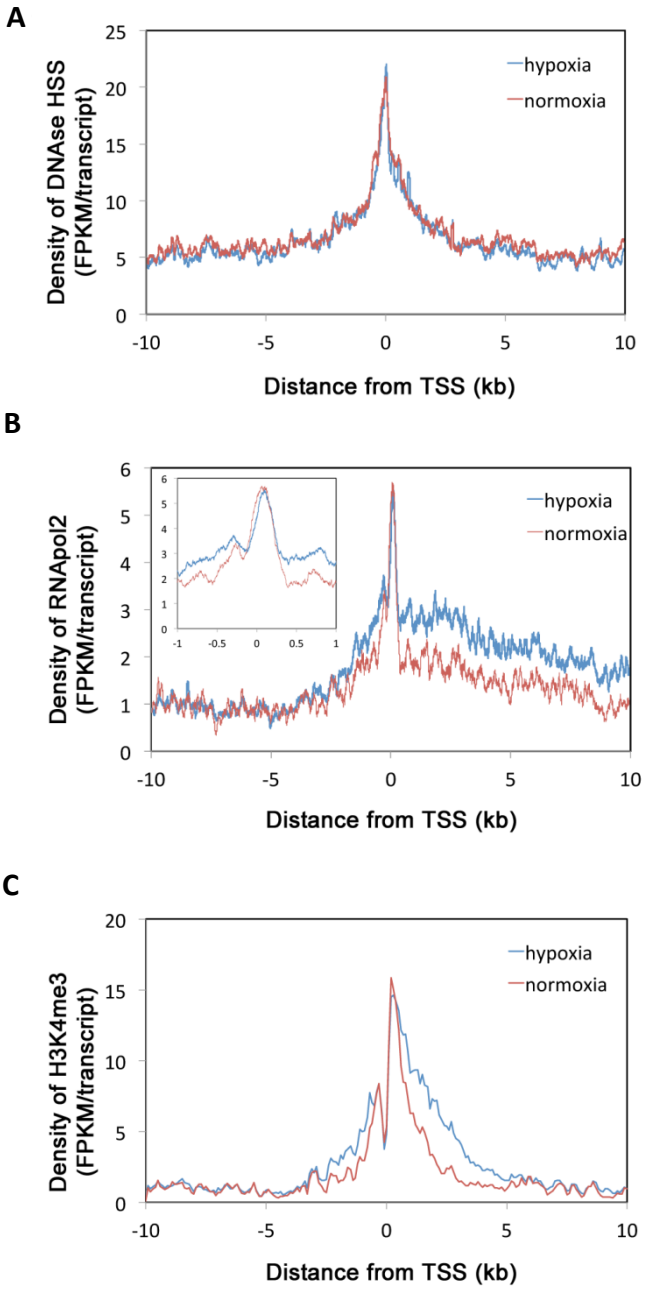


Figure 4.5 Distribution of RNApol2, H3K4me3 and DNAse hypersensitivity around the TSS of the 100 most hypoxia-upregulated transcripts of all classes. Mean distribution of normoxic (red) and hypoxic (blue) **A)** DNAse1 hypersensitivity and **B)** RNApol2 binding at the 100 transcripts most upregulated by hypoxia. The inset shows expanded view at the TSS. The same plots for **C)** H3K4me3 (FPKM = fragments per kilobase per million reads).

No significant changes were seen in any of these signals averaged over 100 non-regulated genes **(Figure 4.6A)**. Interestingly, no significant changes observed either, when signals were averaged over the 100 most hypoxia-downregulated genes **(Figure 4.6B)**.

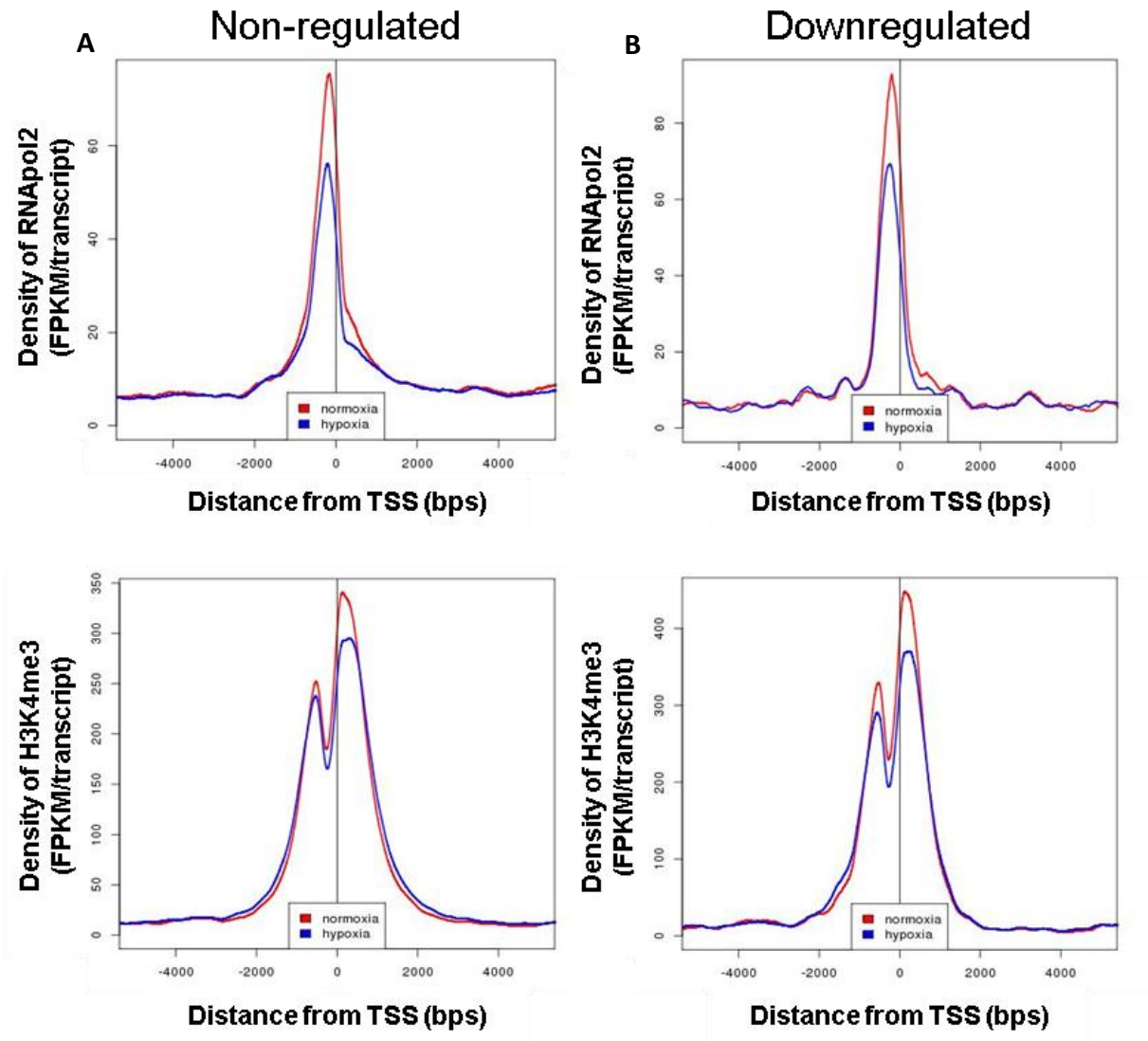


Figure 4.6 Distribution of RNApol2and H3K4me3signals at TSS of top 100 non-regulated and downregulated genes in hypoxia. There were no changes in markers in either **A)** Non-regulated or **B)** Downregulated genes in hypoxic MCF-7 cells. Hypoxia (blue); normoxia (red).

4.3.4 Promoter-paused and de novo recruited RNApol2 in hypoxia

In the previous analysis, changes in RNApol2 and H3K4me3 signal were limited to hypoxia-upregulated genes and not observed at non-regulated or downregulated genes. Similarly, GSEA analysis has shown that HIF-binding is also restricted to hypoxia-induced genes. I therefore examined whether the changes in RNApol2 and H3K4me3 binding were a feature of HIF-binding genes or simply of hypoxia-upregulated genes. I divided hypoxia-inducible genes into those that bound HIF (a high stringency peak within 100 kb of the promoter) or genes that had no evidence of HIF-binding (no peak within 1 MB). The groups were then matched for fold-induction in gene expression by hypoxia. Changes in RNApol2 and H3K4me3 signals were observed for genes that bound HIF **(Figure 4.7 A-C)**, but importantly no changes were observed at hypoxia-upregulated genes that did not bind HIF **(Figure 4.7 B-D).** This strongly suggests that HIF-binding effects transcriptional activation by releasing paused, promoter-proximal RNApol2, with concomitant methylation of H3K4 downstream from the TSS. An example of RNApol2 releasing in hypoxia from a promoter of HIF-dependent lncRNA, is the NEAT1 locus **(Figure 4.7 E).**

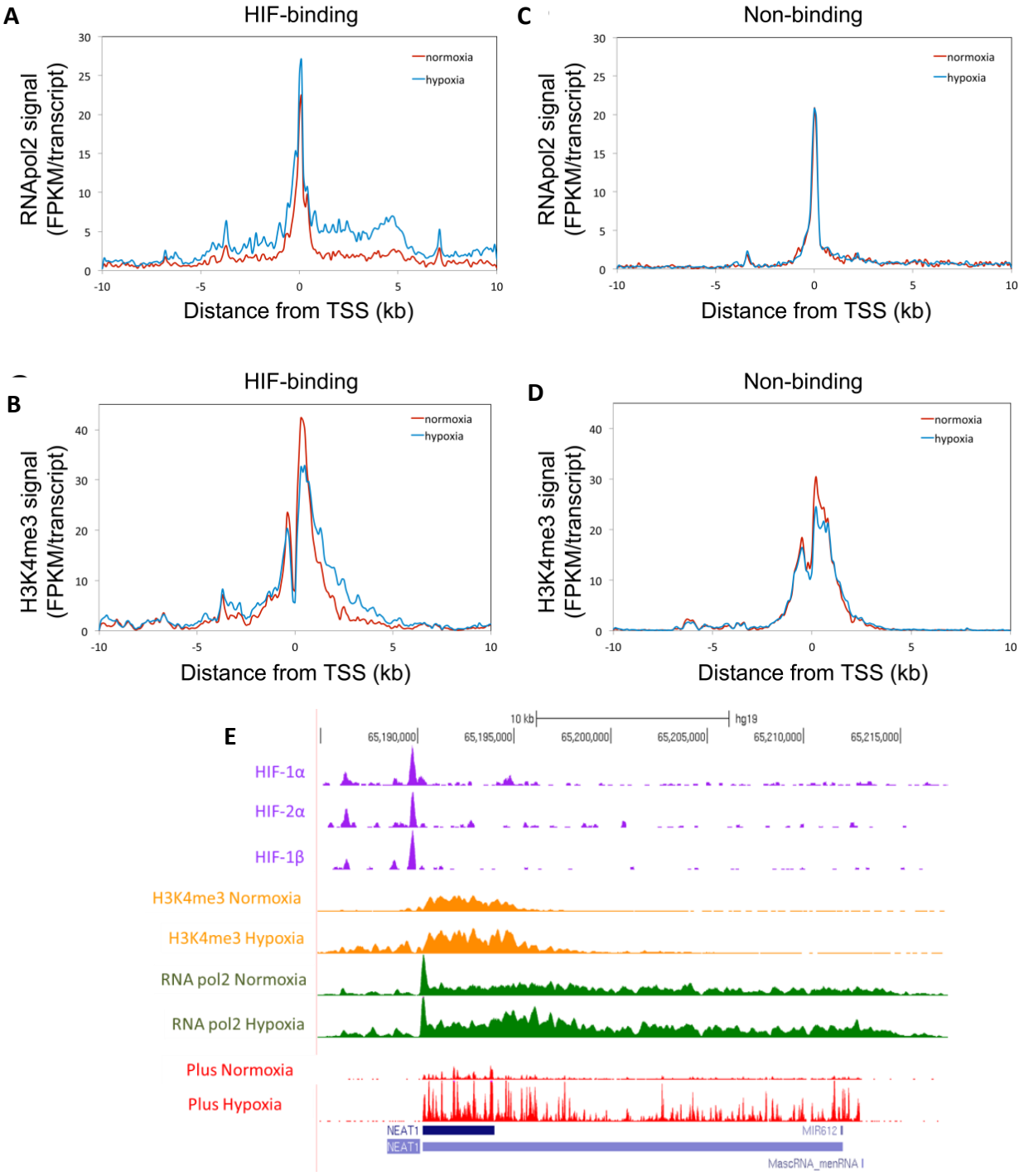


Figure 4.7 Hypoxic regulation of RNApol2 and H3K4me3 at HIF-binding and non-binding gene loci upregulated by hypoxia. mRNAs upregulated by hypoxia were classified according to HIF-binding status. Transcripts closest to and within 100 kb of a high-stringency HIF binding site were defined as HIF-binding. Transcripts with no discernable HIF ChIP-seq signal within 1 Mb of the promoter were defined as non-binding. Matched pairs of upregulated transcripts from each group were selected to have equivalent fold-regulation. Shown are normoxic (red) and hypoxic (blue) signal for **A)** RNApol2 and **B)** H3K4me3 averaged over HIF-binding transcripts and **C)** and **D)** the same plots for non-binding loci. **E)** RNA-seq and ChIP-seq tracks illustrating increase releasing of RNApol2 in hypoxia at the HIF dependent lncRNA. The present data is adapted from (Choudhry et al., 2014).

In addition, close inspection of the travelling ratio of RNApol2 was calculated across HIF genes in normoxia and hypoxia **(Figure 4.8 A-B)**. Analysis identified a small subset of HIF target genes that showed *de novo* recruitment of promoter RNApol2 in hypoxia, suggesting transcriptional regulation of HIF dependent genes by recruitment of *de novo* RNApol2. An example of *de novo* RNApol2 recruitment of HIF genes is the Adrenomedullin (ADM) gene **(Figure 4.8 C).**

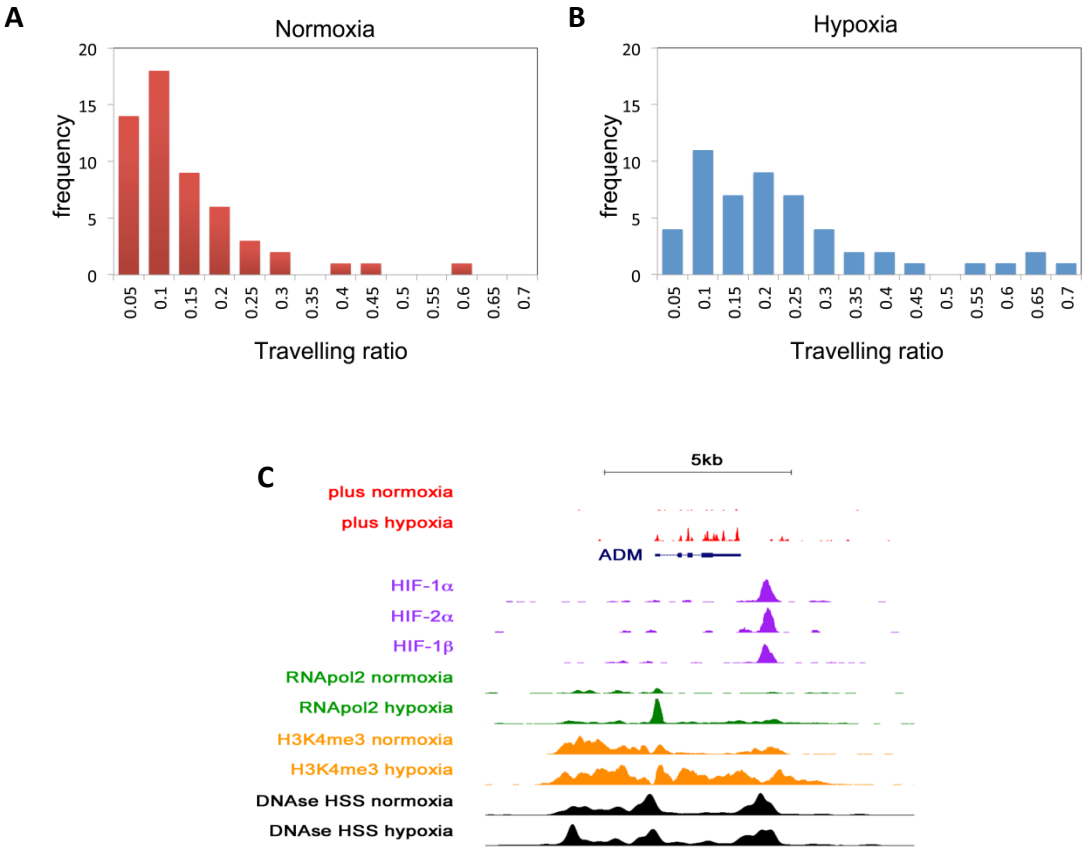


Figure 4.8 De novo recruitment of promoter RNApol2 in hypoxia. Travelling ratio of RNApol2 (number of reads across the body of the gene / number of reads at promoter (-150 bp to +250 bp)) for the most upregulated HIF-binding transcripts in **A)** Normoxia and **B)** Hypoxia. FPKM = fragments per kb per million reads. C**)** RNA-seq and ChIP-seq tracks illustrating *de novo* recruitment of RNApol2, despite constitutive DNAse1 hypersensitivity.

4.3.5 RNApol2 and histone changes at lncRNAs

The above findings are similar to those reported at protein-coding gene loci alone (Schoedel – personal communication and Galbraith). This is not surprising, since a large proportion of the genes included in each dataset were themselves protein-coding genes. To determine whether similar changes were observed at non-coding gene loci as well as protein-coding loci, I first categorised each gene according to RNA class. I then examined RNApol2 and H3K4me3 ChIP-seq signals at the 100 most hypoxia-upregulated and downregulated mRNA and lncRNA gene loci.

Broadly comparable changes were seen for both RNApol2 and H3K4me3 signals at hypoxia-induced mRNA and lncRNA loci **Figure 4.9**, although changes were less at lncRNAs, possibly reflecting a slightly lower hypoxic induction of RNA expression in this class. Again, no significant changes were seen at hypoxia-downregulated loci of either class. These data indicate that HIF regulates transcriptional activation through similar mechanisms at both protein-coding and long non-coding gene loci and, by extension to the work of Galbraith et al. that this occurs predominantly by release of pre-bound, paused, promoter-proximal RNApol2.

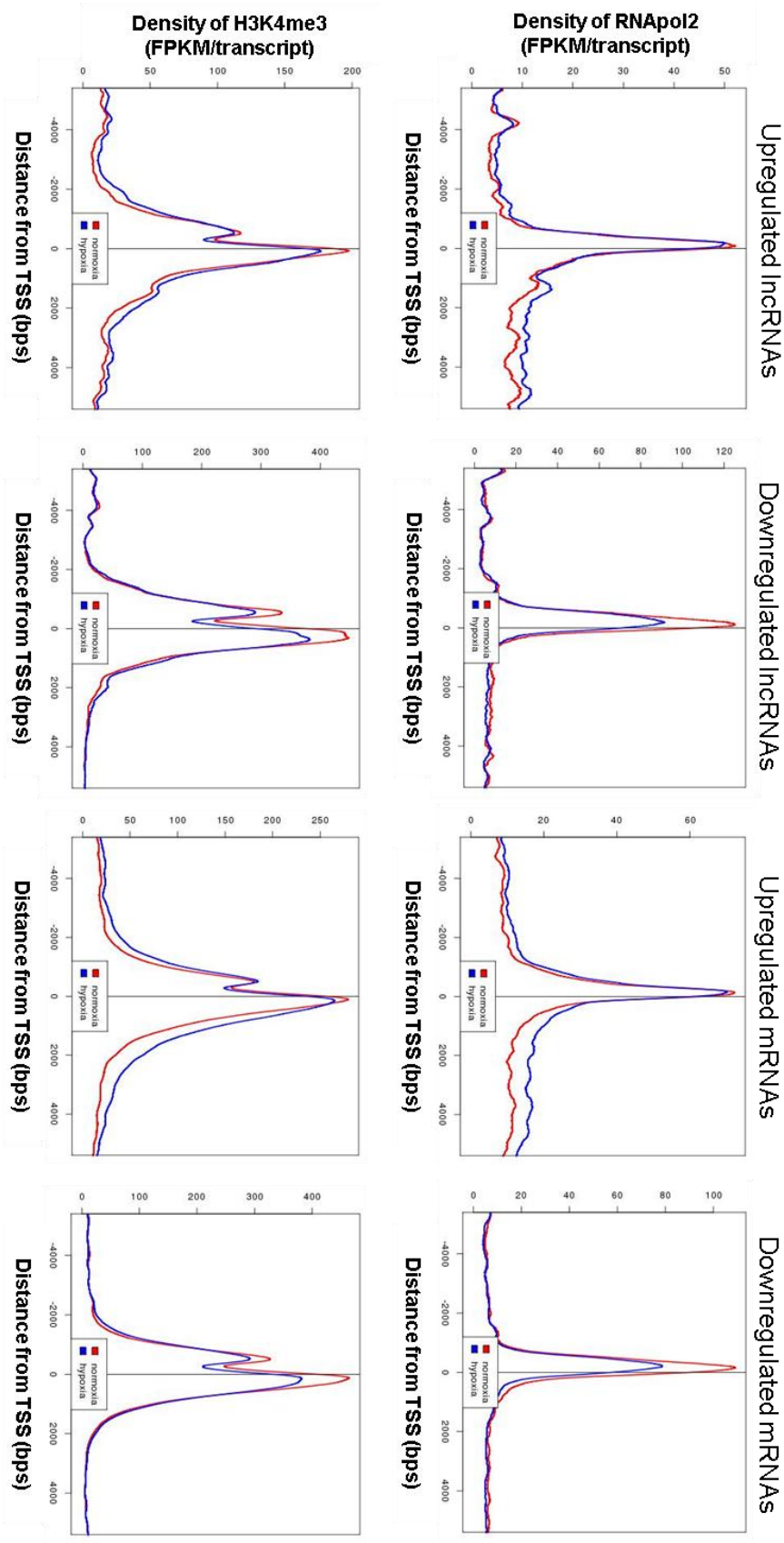


Figure 4.9 Distribution of RNApol2 and H3K4me3 around TSS of the 100 most hypoxia-upregulated lncRNAs and mRNAs. Mean distribution of normoxic (red) and hypoxic (blue) RNApol2 and H3K4me3 reads of the 100 transcripts most up/downregulated lncRNAs and mRNAs in hypoxia.

4.3.6 RNApol2 and histone changes at other non-coding RNAs

Marker analyses of other classes of non-coding RNAs including piwiRNAs, tRNA, snRNA and a small number of lncRNAs did not permit meaningful interruption as they are transcribed via RNA polymerase III (RNApol3) (White, 2008, Dieci et al., 2007). On the other hand, RNApol2 and H3K4me3 signals were analysed for hypoxia-regulated miRNAs, since the majority of miRNAs are embedded and transcribed from their host genes. Thus, marker analyses were performed around host genes that contain hypoxia regulated miRNAs. Signals were highly noisy for miRNA host genes, due to low number of these genes and the possibility that host genes may regulate differently from miRNAs in hypoxia **(Figure 4.10 A-C).**

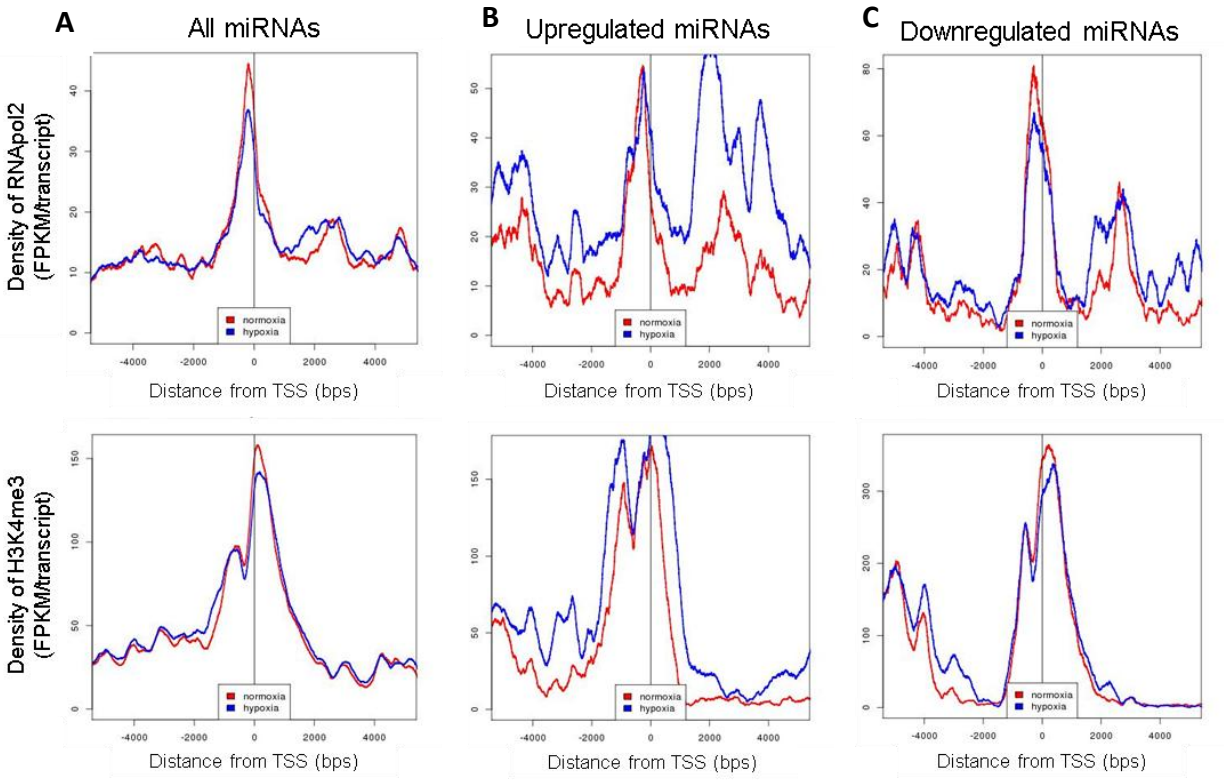


Figure 4.10 Distribution of RNApol2 and H3K4me3 around TSS of hypoxia regulated miRNAs. Mean distribution of normoxic (red) and hypoxic (blue) RNApol2 and H3K4me3 reads around host genes which contain hypoxia regulated miRNAs. **A)** The first panel shows marker signals around all the host genes that contain hypoxia regulated miRNAs. **B)** The second panel of makers for host genes that contain upregulated micRNAs. Although the data is noisy, RNApol2 and H3K4me3 signals are slightly higher in hypoxia compared to normoxia. **C)** The third panel for host genes that contain downregulated miRNAs showed no differences.

4.3.7 RNApol2 and histone modifications of transcripts altered with HIF siRNAs

Finally, I examined the RNApol2 and H3K4me3 changes of up/downregulated transcripts with HIF siRNA sequencing. Since combined HIF-1/2α siRNAs in hypoxia have general suppressive effects on both HIF isoforms, the top hundred up and downregulated genes with combined HIF-1/2α siRNAs knockdown were used for the marker analysis.

As expected, genes downregulated with HIF siRNA showed increased levels of RNApol2 and H3K4me3 signals through the gene body, indicating that these genes were transcriptionally active and upregulated in hypoxia **(Figure 4.11A)**. Genes that were upregulated by HIF siRNA showed no changes in RNApol2 and H3K4me3 levels in hypoxia. They showed similar marker patterns to downregulated genes in hypoxia **(Figure 4.11B)**. Overall, these analyses show the expected reciprocal changes of markers when genes were altered by HIF siRNA compared to genes regulated by hypoxia.

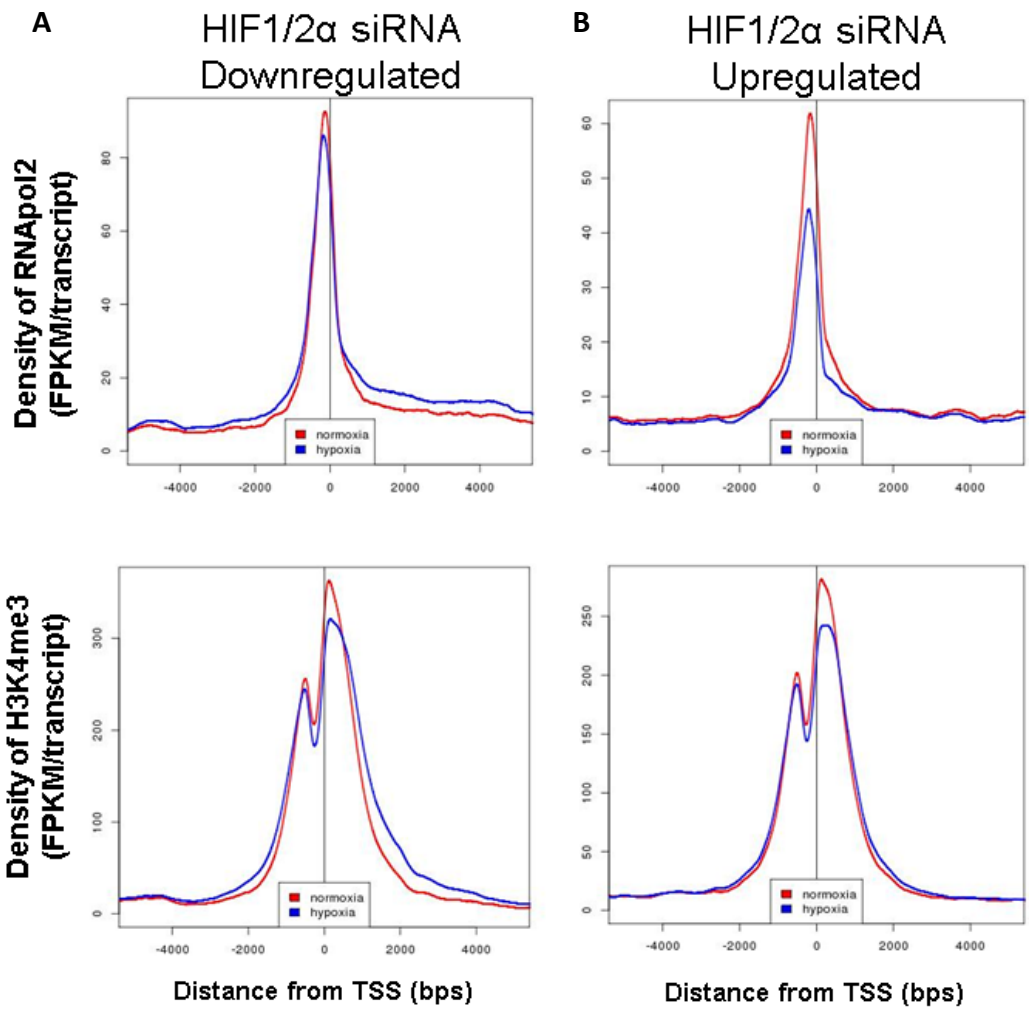


Figure 4.11 Distribution of RNApol2 and H3K4me3 signal around TSS of the 100 transcripts most up/downregulated by HIF siRNA respective of class of RNA. 100 mot **A)** Downregulated **B)** Upregulated transcripts with HIF-1/2α siRNA**.** Mean distribution of normoxic (red) and hypoxic (blue) RNApol2 and H3K4me3 reads of the 100 transcripts most up/downregulated with HIF-1/2α siRNAs in hypoxia.

4.4 Discussion

In this chapter I correlated ChIP-seq signals of RNApol2 and epigenetic markers of active promoters (H3K4me3 and DNAse hypersensitivity) with hypoxic regulation of the coding and non-coding transcriptome determined by total RNA-seq in MCF-7 cells. This analysis revealed that the majority of hypoxia-regulated genes have RNApol2 and H3K4me3 already bound close to the promoter in normoxic cells, in correlation with basal levels of gene expression. This finding is consistent with previous work by Barski et al. 2007, which Freported a correlation between gene expression and RNApol2 and H3K4me3 signals in lymphocytes (Barski et al., 2007).

These findings are also consistence with a number of previous studies that examined the effect of hypoxia on RNApol2 and histone marks. For instance, Tanimoto et al. investigated RNApol2 and the number histone markers (H3K4me3, H3AC and H3K27me3) in hypoxic DLD-1 and TIG-3 cells (Tanimoto et al., 2010), while Xia et al. examined the levels of RNApol2 and H3K4me3 in hypoxic HepG2 and U87 cells (Xia and Kung, 2009). These studies and others have concluded that the presence of RNApol2 and H3K4me3 at gene promoters in normoxia correlates with hypoxic induction of transcripts and that HIF-regulated genes recruit RNApol2 to promoters with open chromatin configuration (Tanimoto et al., 2010, Xia and Kung, 2009, Semenza et al., 1991). However, although no change was observed in DNAse hypersensitivity signals, this technique cannot distinguish individual nucleosome positions, so it is possible that individual nucleosome positioning may change in hypoxia (Zentner and Henikoff, 2012).

Interestingly, examination of RNApol2 and H3K4me3 signals at the most hypoxia-upregulated transcripts of any class showed no increase in RNApol2 at the promoter in hypoxia, while an increase was observed in RNApol2 signals across the gene body. Previous reports indicate that active RNApol2 can recruit histone methyltransferase activity to genes either directly (Ng et al., 2003) or through p300 and leads to increased trimethylation H3K4 within the body of the gene

(Tang et al., 2013). My analyses of H3K4me3 signals at the most hypoxia-upregulated genes again showed no increase of bound H3Kme3 at the TSS but were consistent with the above studies. Furthermore, the changes in H3K4me3 levels in hypoxia observed in the current study were also in agreement with previous work by Tchou-Wong et al. in A549 cells treated with nickel to induce a pseudo-hypoxic response (Tchou-Wong et al., 2011). They reported the most upregulated genes with nickel treatment were highly concordant with HIF-upregulated genes in hypoxia. In addition, they reported increased broadening of H3K4me3 markers around the genes increased by nickel treatment (Tchou-Wong et al., 2011), which is consistent with findings of the current study. Importantly, hypoxia-downregulated genes did not show the reverse changes in RNApol2 and H3k4me3 levels, suggesting different mechanisms in regulating expression of downregulated genes in hypoxia.

A significant hypoxic increase in RNApol2 signals along the gene body was observed at direct HIF transcriptional targets; however, this was not observed for genes upregulated by hypoxia without evidence of HIF binding. Although the exception rather than the rule, close inspection of the travelling ratio of RNApol2 across HIF genes identified a small subset of HIF target genes that showed *de novo* recruitment of promoter RNApol2 in hypoxia, hinting at gene-specific functions of HIF in recruiting RNApol2 *de novo*. This has been described for other transcription factors such as IRF3 and NFκB, both of which can release paused RNApol2 and recruit *de novo* RNApol2 (Freaney et al., 2013).

After dissecting the lncRNAs class from the transcriptome, analyses of markers showed similar patterns as entire transcriptome regulation in hypoxia (increased run through in gene body of RNApol2 and broadening of H3K4me3 in hypoxia), suggesting that lncRNAs have similar mechanisms of transcription activation to mRNAs in hypoxia. Strong peaks of RNApol2 and H3K4me3 signals were also detected at the putative transcriptional start sites of both non-annotated intergenic transcripts and antisense transcripts in normoxia, confirming the presence of these transcripts and their common mechanisms of transcription.

In summary, the analysis of markers and their integration with the hypoxia transcriptome in this section supports a model of HIF-dependent transcription. In normoxia, the majority of direct HIF target genes display an open chromatin conformation with the presence of H3K4me3 and paused RNApol2 (Galbraith et al., 2013, Mole et al., 2009, Schodel et al., 2011), while in hypoxia, HIF-α isoforms dimerize with HIF-1β and bind to the HREs of their target genes, which then recruits several co-activators. These include acetyl-transferases, p300 and Cyclin-dependent kinase 8 "CDK8" mediator, as well as the Super Elongation Complex (SEC), which contains AFF4 and positive transcription elongation factor b "P-TEFb" (Bedford et al., 2010, Ruas et al., 2010, Galbraith et al., 2013). Formation of this complex results in an increase of histone acetylation and bromodomain protein "BRD4" binding, which then triggers CDK8-Mediator along with the SEC to release paused RNApol2 and its elongation (Galbraith et al., 2013, Dengler et al., 2013). However, some direct HIF target genes showed exceptions to the proposed model, where RNApol2 was not present in normoxia and was only recruited *de novo* in hypoxia, providing a new mechanism of transcription in hypoxia.

Chapter 5: The expression of NEAT1, an hypoxia-regulated lncRNA, is associated with the clinical outcome of breast cancer

5.1 Introduction

Recent genomic analyses suggest that approximately 22,000 protein-coding genes are encoded by about 2 % of the genome. The majority of the remaining transcribed genome (accounting for approximately 34,000 transcripts) consists of non-coding RNAs (Carninci et al., 2005, Djebali et al., 2012). These non-coding RNAs may be broadly separated, based on their length, into small non-coding RNAs (<200 bp) and long non-coding RNA (>200 bp). There is accumulating evidence that long non-coding RNAs (lncRNA) have functional, and regulatory roles in health and disease (Mercer et al., 2009, Batista and Chang, 2013, Du et al., 2013) and in particular, during tumourgenesis (Eddy, 2001, Esteller, 2011, Stefani and Slack, 2008).

To date, studies of the non-coding RNA response to hypoxia have largely focused on microRNAs, typically composed of very short non-coding sequences of around 22 bp (Kulshreshtha et al., 2008, Camps et al., 2008), whilst (with a few exceptions) the role and regulation of lncRNAs in hypoxia remains largely unstudied. Matouk et al. have shown that imprinted maternally expressed transcript (H19), an lncRNA, is upregulated in Hepatocellular carcinoma (HCC) and bladder carcinoma cell lines in response to hypoxia and that this hypoxic regulation was dependent on HIF-1α (Matouk et al., 2007). Conversely, the lncRNA Low Expression in Tumour (lncRNA-LET) is commonly downregulated in a number of cancers including hepatocellular, colorectal and squamous-cell lung carcinomas (Yang et al., 2013a). This hypoxic suppression of lncRNA-LET enhances the accumulation and stability of HIF-1α mRNA under hypoxic conditions. Functionally, downregulation of lncRNA-LET promotes HCC metastasis (Yang et al., 2013a).

In previous Chapters, I characterised the hypoxic regulation of a large number of lncRNAs through transcriptional profiling of MCF-7 breast cancer cells cultured in hypoxic (24 hours, 1 % O_2) or in normoxic conditions (21 % O_2) using next generation RNA-sequencing. This identified a large number of lncRNAs that are induced in hypoxia and whose induction is

dependent on HIF. The most hypoxia-upregulated lncRNA in this dataset was Nuclear Enriched Abundant Transcript 1 (NEAT1). NEAT1 is a non-coding RNA that is expressed in the nucleus. It has two isoforms that are transcribed from a single promoter, a polyadenylated short transcript, which is 3.7 kb in length, and a longer non-polyadenylated transcript (Guru et al., 1997, Sasaki et al., 2009). NEAT1 is an essential architectural component that is required for the formation and maintenance of "paraspeckles". These are nuclear-specific structures that lie in close proximity to nuclear "speckles" (Sasaki et al., 2009, Chen and Carmichael, 2009, Clemson et al., 2009). Paraspeckles bodies are amongst the most recently identified nuclear bodies and were first described in 2002 (Fox et al., 2002). These bodies contain three main Drosophila Behaviour and Human Splicing (DBHS) proteins, including Paraspeckle Component 1 (PSPC1), Non-POU Domain Containing Octamer Binding (NONO or p54nrb) and Splicing Factor Proline/Gluatamine-Rich (SFPQ or PSF) (Fox et al., 2002, Nakagawa and Hirose, 2012). They are approximately 360 nm in diameter, and occur at between 2 and 20 foci in each cell (Fox et al., 2002). They are restricted to mammalian nuclei and found in both cultured cell lines and primary cultures (Bond and Fox, 2009) with except for embryonic stem cells (Chen and Carmichael, 2009).Recently, increased NEAT1 expression and paraspeckles body formation have been reported during myotube differentiation and in a variety of pathological conditions including intrauterine growth restriction disorders and viral infection (Gremlich et al., 2013, Zhang et al., 2013a, Sunwoo et al., 2009). However, the role and regulation of NEAT1 in cancer and during hypoxia is not been determined.

The current chapter focuses on analysis of the hypoxic regulation of NEAT1 in breast cancer cells subjected to culture in normoxia or hypoxia. I then evaluated the effect of NEAT1 knockdown on cell proliferation, survival and apoptosis. Finally, the clinical relevance of NEAT1 was examined in a large cohort of breast cancer patients (Curtis et al., 2012). This chapter provides novel insights into the role of lncRNAs in the hypoxic microenvironment of solid tumours in general, and breast cancer in particular.

5.2 Exprimetnal desigen

In this chapter, NEAT1 lncRNA regulation was measured by qPCR in 11 breast cancer cell lines grown in normoxia and hypoxia (1 % O_2 for 24 hours). Expression of NEAT1 was also evaluated by qPCR in MCF-7 and ZR-75-1 cells at different hypoxia points (1 % O_2 for 16, 24, 32, 48 hours) and with DMOG treatment (500 nM, 24 hours). In addition, NEAT1 expression in cytoplasmic and nuclear fractions of MCF-7 cells grown in hypoxia (1 % O_2 for 24 hours) and normoxia was evaluated by qPCR. RNA fluorescence *in situ hybridization* (RNA FISH) was performed against NEAT1 and RPL11 (control probe) in MCF-7 and ZR-75-1 cells in normoxia and hypoxia in order to define NEAT1 cellular localisation. HIF-1α and HIF-2α siRNAs were transfected in MCF-7 and ZR-75-1 cells in hypoxia (1 % O_2 for 24 hours) and NEAT1 expression was measured by qPCR and RNA FISH. Next, NEAT1 levels were evaluated in MCF-7, MDA-MB-231 and MDA-MB-468 xenograft models treated with Bevacizumab or with PBS in control groups. The xenograft tumours were obtained from Prof. Adrian Harris lab at the WIMM. Total RNA was isolated from treated and control groups and was used to measure CA9 (control hypoxia gene) and NEAT1 expression by qPCR. Furthermore, total protein expression of paraspeckles proteins (PSPC1 and NONO) was measured in MCF-7 by western blot in normoxia and hypoxia (1 % O_2 for 24 hours). Immunostaining for PSPC1, NONO and SFPQ proteins was performed in normoxia and hypoxia in MCF-7 and ZR-75-1 cells. Two different antisense oligonucleotides (ASO) targeting NEAT1 and a control scrambled ASO were designed to supress NEAT1 expression in MCF-7 and ZR-75-1 cells. After ASOs transfection with Lipofectamine® RNAiMAX in the cells, the NEAT1 expression was measured by qPCR to confirm effective knockdown. Following individual NEAT1, HIF-1α and HIF-2α knockdowns in hypoxic MCF-7 cells (1 % O_2 for 24 hours), the paraspeckles proteins were immunostained to evaluate the dependence of paraspeckles body formation on the NEAT1 transcript and the HIF regulation. To study the role NEAT1 on tumourigenesis, NEAT1 knockdown was performed on MCF-7 and ZR-75-1 cells in normoxia and hypoxia (1 % O_2 for 24 hours) and subjected to cell proliferation and clonogenic assays.

Furthermore, effect of NEAT1 knockdown on apoptosis was measured by the Caspase 3/7 activity and by theFACS analysis for annexin V stained cells in MCF-7 and ZR-75-1 cells. The Cancer Genome Atlas (TCGA) data (http://www.cbioportal.org/public-portal/) was used to screen somatic alteration harbouring NEAT1 and paraspeckle protein coding genes in a range of cancer types. Finally, NEAT1 and paraspeckle protein coding gene expression were evaluated in previously published expression data of a large cohort of breast cancer (n=200) (Curtis et al., 2012) for their association with clinical outcome and with clinicopathological features of breast cancer.

5.3 Results

5.3.1 Genome features of NEAT1 loci in hypoxic MCF-7 cells

To examine the transcriptional features across the NEAT1 locus in normoxia and in hypoxia, I used the "pile-up" tracks derived from the RNA-seq and ChIP-seq datasets (HIF-binding, RNA polymerase II and histone marker H3K4me3) in MCF-7 cells. Reads were mapped to Human (HG19) reference genome using Bowtie and viewed using the USCS genome browser.

A significant increase in the expression (RNA-seq signals) of NEAT1 (both short and long transcripts) was observed in hypoxia compare to normoxia **(Figure 5.1A).** Consistent with the findings at other HIF-regulated genes, no change in RNA Pol II binding was observed at the transcription start site (TSS) of NEAT1. However, a clear increase in RNA Pol II binding across the NEAT1 transcript body was observed in hypoxia commensurate with the increased expression level. In addition, the observed H3K4me3 peak (a marker of promoter activity) increased in breadth at the NEAT1 TSS in hypoxia compared to normoxia. The increased run-through of RNA Pol II across the NEAT1 gene body and the broadening of the H3K4me3 peak at the TSS both suggest increased transcription of NEAT1 during hypoxia. Furthermore, analysis of the HIF-1α, HIF-2α and HIF-1β ChIP-seq tracks identified binding peaks for all three isoforms within the promoter region of NEAT1 **(Figure 5.1A).**

5.3.2 NEAT1 expression in a panel of hypoxic breast cancer cell lines

To evaluate whether NEAT1 regulation in response to hypoxia is specific to MCF-7 cells or is a general feature of breast cancer cells, an extended panel of additional breast cancer cell lines was studied. These exhibited different breast cancer genotypes (ER+ and HER2- cell lines "ZR-75-1, MCF-7, T47D, BT474", ER- and HER+ cell lines "MDA-MB-453, SKBR3" and triple negative breast cancer cell lines " MDA-MB-468, MDA-MB-231, DU4475, BT549, BT20"). All the cell lines were grown in normoxia and hypoxia 1 % O_2 for 24 hours and experiments were performed with three biological replicates. NEAT1 expression was upregulated in the majority of breast cancer cell lines when grown in hypoxia **(Figure 5.1B).** ZR-75-1 and MCF-7 cells showed the most profound upregulation with 3.5- and 3.1-fold upregulation, respectively (P value<0.005). These two cell lines were selected for subsequent experiments.

Using both MCF-7 and ZR-75-1 cells, I next evaluated NEAT1 expression at different time points (16, 24, 32, 48 hours) after hypoxic incubation at 1 % O_2. The expression of NEAT1 was induced at the earliest time point (16 hours) and was maximal (and sustained) by 24 hours, with no further induction thereafter **(Figure 5.1C).**

5.3.3 NEAT1 expression is induced by the HIF hydroxylase inhibitor, dimethyloxalylglycine (DMOG)

As outlined in the introduction, HIF is regulated primarily through stability of its alpha subunits and through their transactivating activity. This in turn is governed by oxygen-dependent hydroxylation of specific amino-acids in HIF by HIF hydroxylase enzymes. In addition to oxygen these require 2-oxoglutarate (2-OG) and can be inhibited by small molecule mimetics of 2-OG such as dimethyloxalylglycine (DMOG) leading to activation of the HIF-transcriptional response in normoxia. I therefore determined whether DMOG could induce the expression of NEAT1 in normoxia.

MCF-7 and ZR-75-1 cells were treated with DMOG (500 nM, 24 hours) in normoxia. The expression of NEAT1 was measured after treatment with DMOG or a sterile nucleic acid free water for control. The expression of a known HIF-target gene, carbonic anhydrase IX (CA9), was measured as a positive control. CA9 expression was significantly induced by DMOG in both MCF-7 cells (5.4 fold) and ZR-75-1 cells (6.1 fold) confirming the effectiveness of the DMOG treatment **(Figure 5.1D).** NEAT1 was also significantly upregulated in MCF-7 cells (2.8-fold, $p = 0.006$) and in ZR-75-1 cells (3.2 fold $p = 0.003$) following treatment with DMOG, indicating that NEAT1 induction is dependent upon both oxygen and 2-oxogulutarate **(Figure 5.1D).**

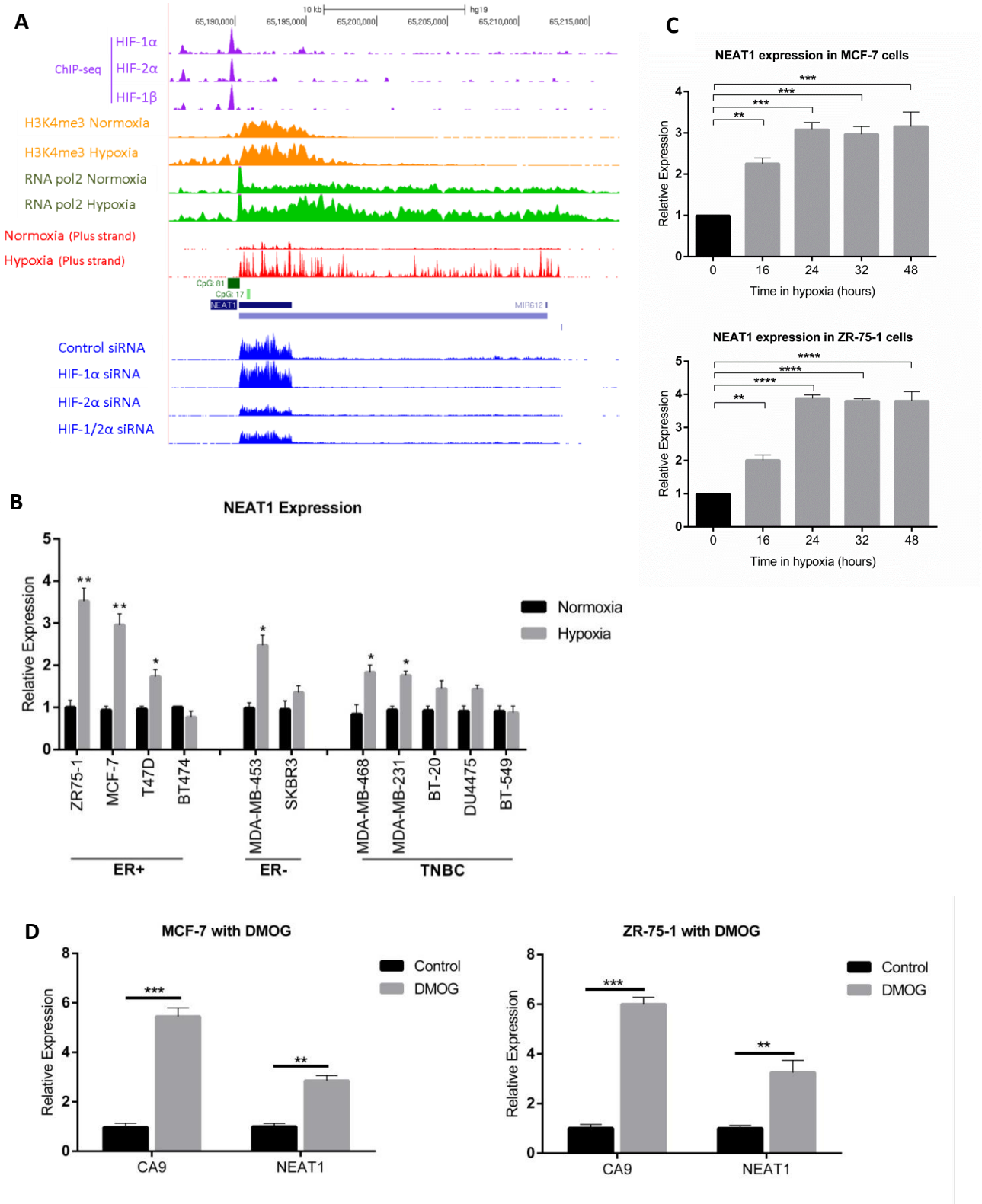


Figure 5.1 NEAT1 lncRNA is regulated by hypoxia and DMOG. A) UCSC Genome Browser tracks at the NEAT1 locus for RNA-seq and ChIP-seq (HIF-1α, HIF-2α, HIF-1β, RNApol2, H3k4me4) data in normoxia and hypoxia. **B)** NEAT1 expression in a panel of breast cancer cell lines grown in normoxia and hypoxia (TNBC: Triple Negative Breast Cancer). The majority of cell lines showed upregulation of NEAT1 when exposed to hypoxia. **C)** NEAT1 expression at different hypoxia time points (16, 24, 32, 48 hours) in MCF-7 and ZR-75-1 cells. **D)** Relative expression of CA9 and NEAT1 in MCF-7 and ZR-75-1 cells after DMOG treatment for 24 hours. Each experiment was performed with three biological replicates. (*p<0.05 , **p≤0.009, ***p≤0.0009, Student's t-test, n=3).

5.3.4 Hypoxia-induced NEAT1 is retained in the nucleus

Previous reports have shown that NEAT1 is usually exclusively located within the cellular nucleus. I determined whether expression of hypoxia-induced NEAT1 was also restricted to the nucleus or whether hypoxic stress might alter its cellular localisation. I fractionated cytoplasmic and nuclear fraction from MCF-7 cells grown in hypoxia and normoxia. qPCR analysis of each cellular fraction showed that NEAT1 is highly enriched in the nuclear fraction and that this fraction alone is induced in hypoxia **(Figure 5.2A).** In addition, I undertook RNA *In Situ Hybridization* (RNA FISH) of NEAT1 in MCF-7 and ZR-75-1 cells incubated in normoxia or in hypoxia. This confirmed the qPCR findings showing that hypoxia-induced NEAT1 is localised to the nucleus in both cell lines **(Figure 5.2B-C and Figure 5.3A-C).** These finding indicate that hypoxia does not redistribute NEAT1 to the cytoplasm and suggests that it may perform the same function in hypoxia as it does when present in normoxia.

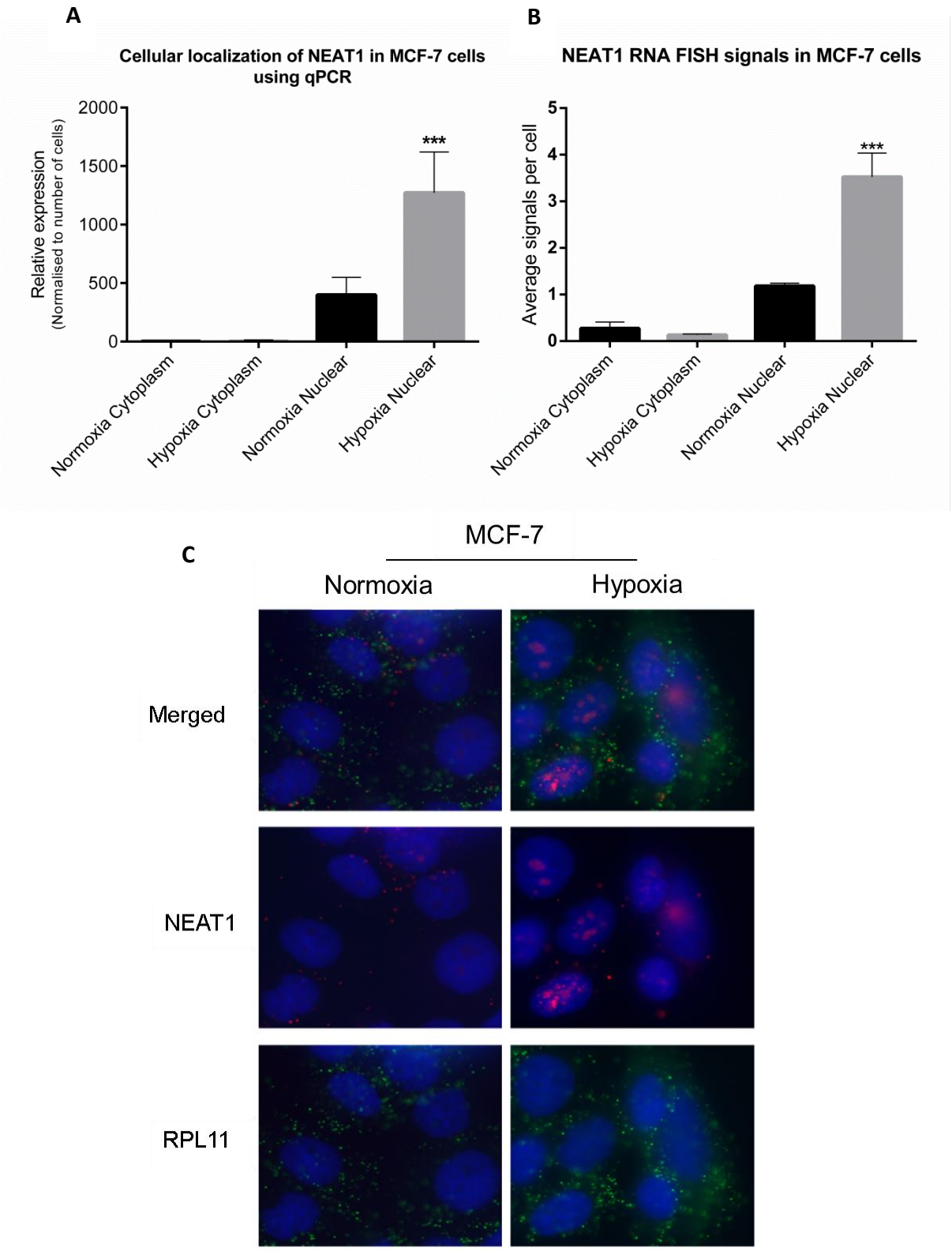


Figure 5.2 Hypoxia-induced NEAT1 is retained in MCF-7 nucleus. **A)** qPCR expression analysis of cytoplasmic and nuclear fractions normalized to equal numbers of MCF-7 cells in hypoxia or normoxia showing induction of NEAT1 in the nucleus in hypoxic conditions. **B)** Average NEAT1 RNA FISH signals in cytoplasmic and nuclear fractions of MCF-7 cells grown in normoxia and hypoxia 1 % O_2 for 24 hours. **C)** RNA FISH of NEAT1 (red) and RPL11 (green) in normoxia and hypoxia in MCF-7 cells. RPL11 was used as cytoplasmic control and showed no changes in both conditions. Experiments were performed with three biological replicates, (***$p<0.0001$, Student's t-test, n=3).

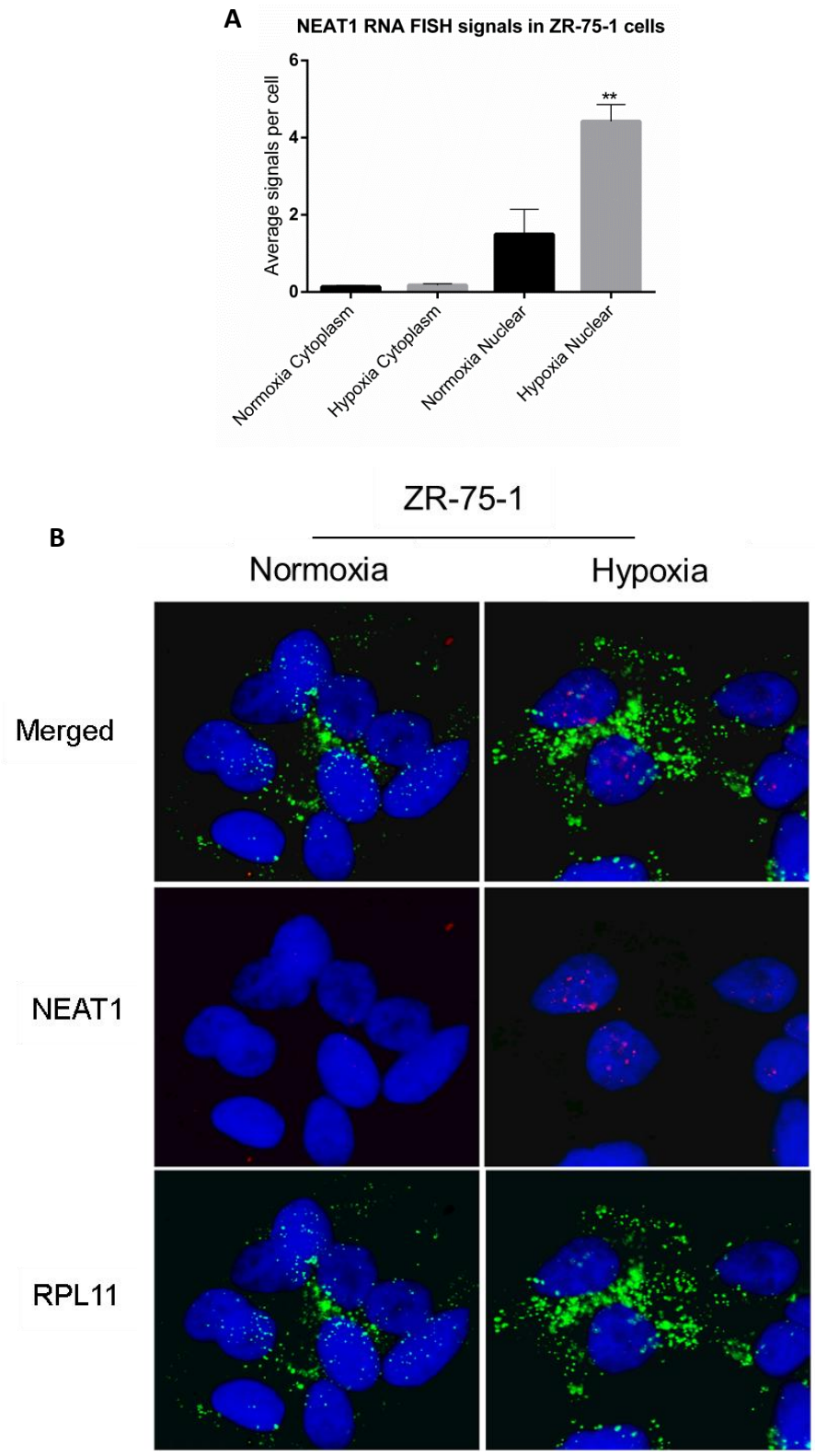


Figure 5.3 Hypoxia-induced NEAT1 is retained in ZR-75-1 nucleus. **A)** Average NEAT1 RNA FISH signals in cytoplasmic and nuclear fractions of ZR-75-1 cells grown in normoxia and hypoxia 1 % O_2 for 24 hours, shows NEAT1 is induced and enriched in nucleus **B)** RNA FISH of NEAT1 (red) and RPL11 (green) in normoxia and hypoxia in ZR-75-1 cells, showing a significant increase in nuclear NEAT1 signals in hypoxia. RPL11 was used as cytoplasmic control and showed no changes. Experiments were performed with three biological replicates, (**p<0.005, Student's t-test, n=3).

5.3.5 NEAT1 expression is HIF dependent in hypoxia

The presence of a HIF binding site within the promoter region of NEAT1, together with its regulation by both hypoxia and DMOG, strongly suggests that NEAT1 is a direct transcriptional target of the HIF transcription factor. siRNA directed against HIF-1α and HIF-2α that have previously been shown to produce suppression of these HIF subunits (Elvidge et al., 2006) were transfected into MCF-7 and ZR-75-1 cells cultured in hypoxic conditions. CA9 is a HIF-1α dependent mRNA. Expression of CA9 was significantly reduced with HIF-1α siRNA in both cell lines, confirming HIF-1α knockdown **(Figure 5.4A)**. Expression of NEAT1 was significantly downregulated by siRNAs directed against HIF-2α (around 95 % in MCF-7 and 65 % in ZR-75-1 reduction, $P<0.001$), whilst siRNAs directed against HIF-1α did not affect NEAT1 expression **(Figure 5.4B).** In addition, I analysed the effect of HIF-1α and HIF-2α suppression on hypoxic NEAT1 expression in MCF-7 cells using RNA FISH. Again, hypoxic NEAT1 expression was inhibited by HIF-2α suppression, but not by HIF-1α suppression **(Figure 5.4C).** Taken together, these findings indicate that although both HIF isoforms bind to the NEAT1 promoter region, hypoxic-induction of NEAT1 expression depended only on the HIF-2α isoform in both cell lines.

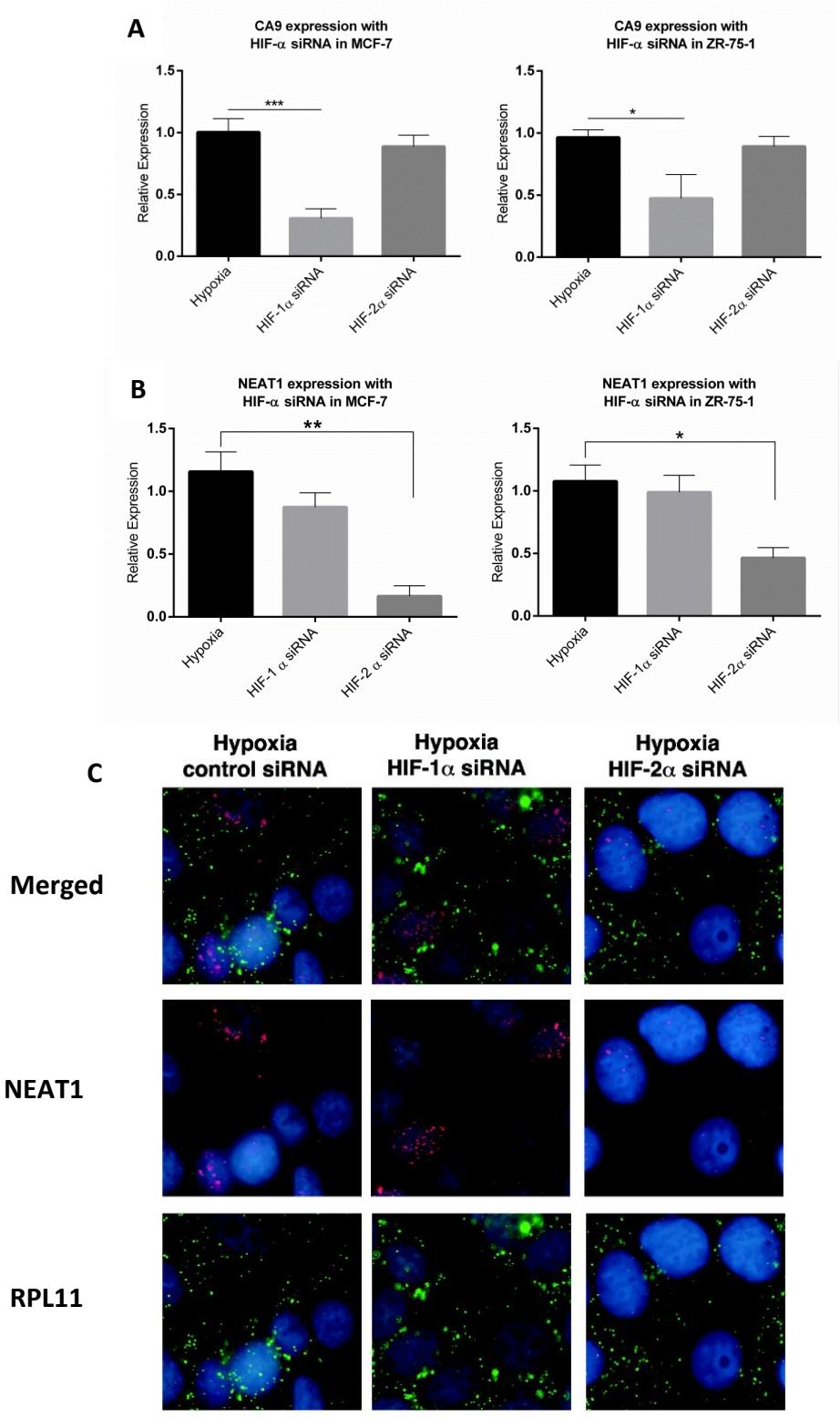


Figure 5.4 Hypoxia-induced NEAT1 expression is dependent on HIF-2α, but not HIF-1α. A) CA9 expression was significantly reduced by HIF-1α siRNA but not following HIF-2α knockdown in MCF-7 and ZR-75-1 cells in hypoxia. **B)** NEAT1 expression after HIF-1α and HIF-2α knockdown in MCF-7 and ZR-75-1 cells in hypoxia. NEAT1 expression was significantly reduced by HIF-2α siRNA but not following HIF-1α knockdown. **C)** RNA FISH of NEAT1 (red) and RPL11 (green) in MCF-7 cells grown in hypoxia and treated with either control (scrambled) siRNA, HIF-1α siRNA or HIF-2α siRNA. HIF-2α siRNA, but not HIF-1α siRNA, significantly reduced hypoxic NEAT1 expression. RPL11 expression was unaffected by either siRNA. Each experiment was performed with three biological replicates, (*$p<0.05$, **$p<0.008$, Student's t-test, n=3).

5.3.6 Antiangiogenic therapy with Bevacizumab upregulates NEAT1 expression in tumour xenografts

Tumour angiogenesis stimulated by the upregulation of HIF pathways in response to tumour hypoxia and oncogenic signals is an important component of the pathogenesis of many types of cancer, including breast cancer. When angiogenesis is impeded by antiangiogenic therapies such as the anti-VEGF antibody, Bevacizumab, tumours become increasingly hypoxic, further upregulating HIF. I therefore evaluated NEAT1 expression in xenograft models of three breast cancer cell lines MCF-7, MDA-MB-231 and MDA-MB-468 that were rendered more hypoxic by treatment with the VEGF inhibitor Bevacizumab (Favaro et al., 2012, Blouw et al., 2003, McIntyre et al., 2012). Xenograft tumours were obtained from Prof. Adrian Harris lab at WIMM. Total RNA was isolated from tumours treated with Bevacizumab and from a control group treated with PBS (n=5 for each breast cancer cell line). Expression of the HIF-responsive gene CA9 was measured by qPCR as a control to show that bevacizumab treatment induced hypoxia *in vivo*. CA9 expression was significantly upregulated in Bevacizumab treated tumours when compared to control tumours in all three models **(Figures 5.5A)**, indicating that Bevacizumab treatment induces sufficient hypoxia to activate HIF transcriptional pathways. I then measured the expression of NEAT1 in these tumour samples, again by qPCR analysis. This revealed a significant increase in NEAT1 expression when treated with Bevacizumab in all three xenograft models. NEAT1 showed median induction of 2.7-fold in MCF-7 xenografts, 3.6-fold in MDA-MB-231 xenografts and 2.3-fold in MDA-MB-468 xenografts with Bevacizumab treatment **(Figures 5.5B).** These findings indicate that the levels of hypoxia encountered within solid tumours *in vivo* are sufficient to induce NEAT1 expression.

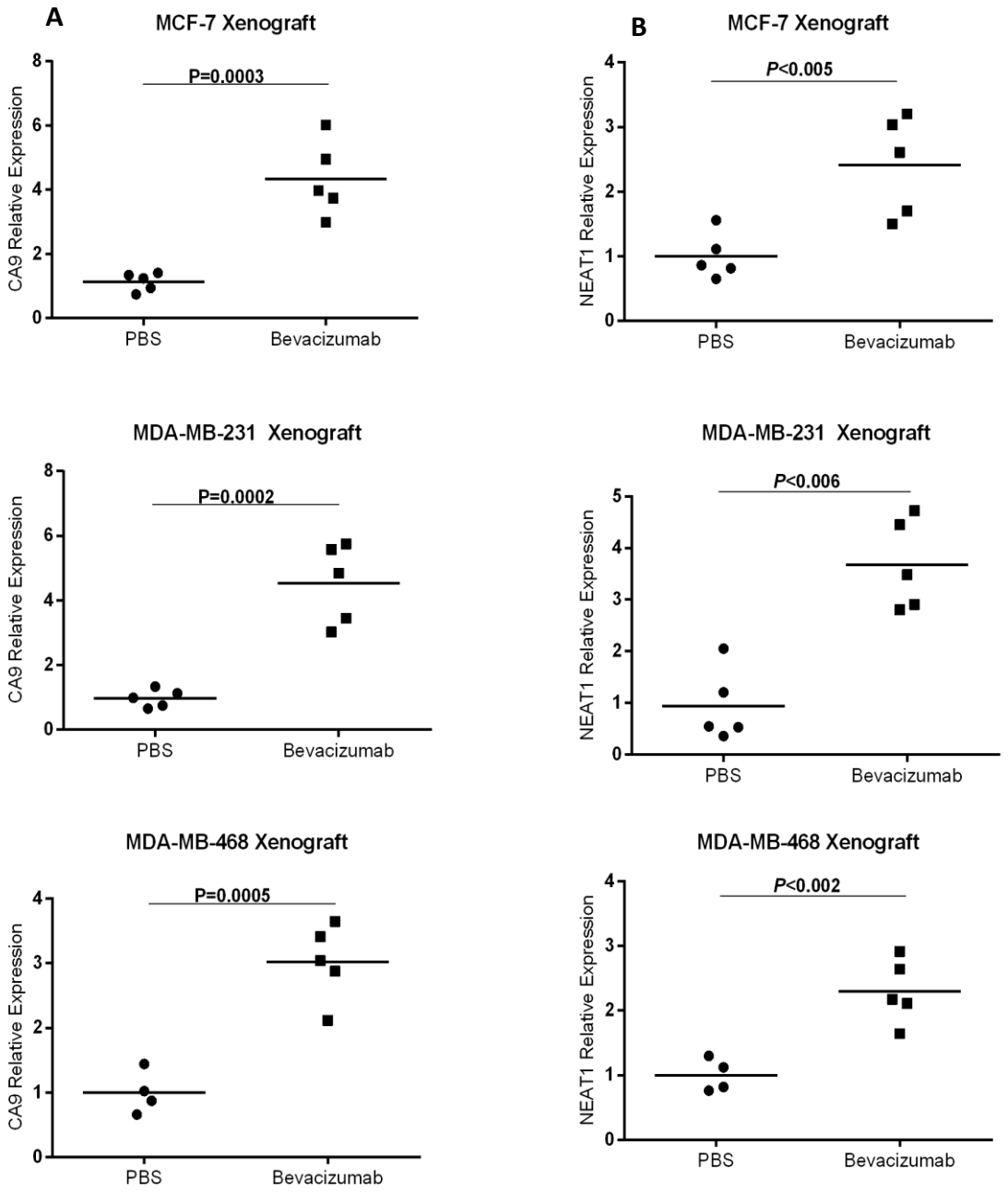


Figure 5.5 CA9 and NEAT1 expression are induced by hypoxia in solid tumour models. A) CA9 **B)** NEAT1 expression in MCF-7, MDA-MB-231 and MDA-MB-468 xenografts models treated with the VEGF inhibitor (Bevacizumab) to induce hypoxia or with PBS control alone. Both CA9 and NEAT1 expression were significantly upregulated with Bevacizumab treatment in all three models.

5.3.7 Hypoxia induces the formation of paraspeckles

Since NEAT1 is essential for paraspeckle formation (Sasaki et al., 2009, Chen and Carmichael, 2009, Clemson et al., 2009), I next investigated whether hypoxic stress can induce the formation of paraspeckles. First, I measured the expression levels of the main protein components of the paraspeckle complex, PSPC1 and NONO, in hypoxia (1 % O_2, 24 Hours) and normoxia by Western blot. There was no significant change in total PSPC1 and NONO protein expression in response to hypoxia **(Figure 5.6),** while SFPQ antibodies didn't work in immunoblots despite using several commercial available antibodies.

Since the hypoxic induction of NEAT1 might induce aggregation of pre-formed paraspeckle proteins, rather than their *de novo* synthesis, I next determined whether hypoxia altered the distribution of paraspeckle proteins within the cell by immunostaining for the paraspeckle components, PSPC1, NONO and SFPQ. MCF-7 cells were subjected to hypoxia (1 % O_2, 24 Hours) and normoxia as before. Importantly, all three component proteins were diffusely distributed in normoxic conditions with clear aggregation into paraspeckle foci in hypoxia **(Figure 5.7),** indicating that hypoxia doesn't alter expression, but rather the distribution of paraspeckle proteins. Similar findings were observed in ZR-75-1 cells **(Figure 5.8).**

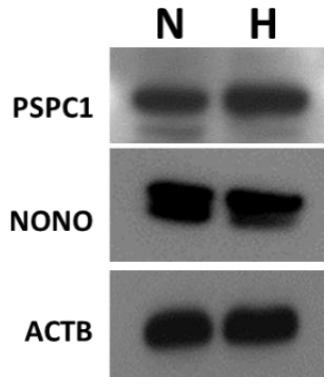


Figure 5.6 Protein expression of PSPC1 and NONO in hypoxic MCF-7 cells. Total expression of PSPC1 and NONO protein was unchanged in normoxic (N) and hypoxic (H) conditions in MCF-7 cells.

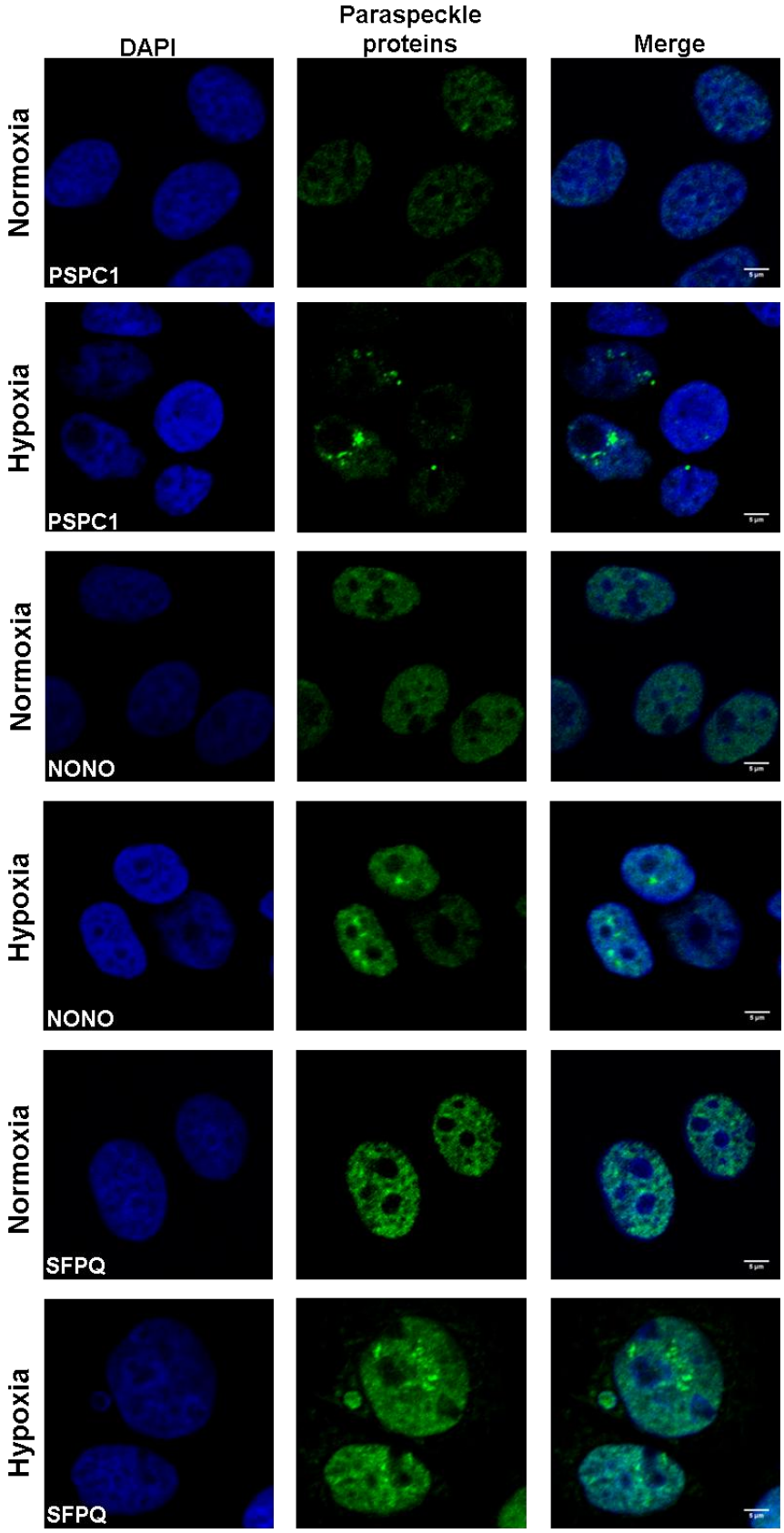


Figure 5.7 Hypoxia induces the formation of paraspeckle aggregates in MCF-7 cells. Immunostaining for PSPC1, NONO and SFPQ proteins showed diffuse expression in normoxic condition, but in hypoxia there was a clear increase in the number of paraspeckle foci (Scale 5um).

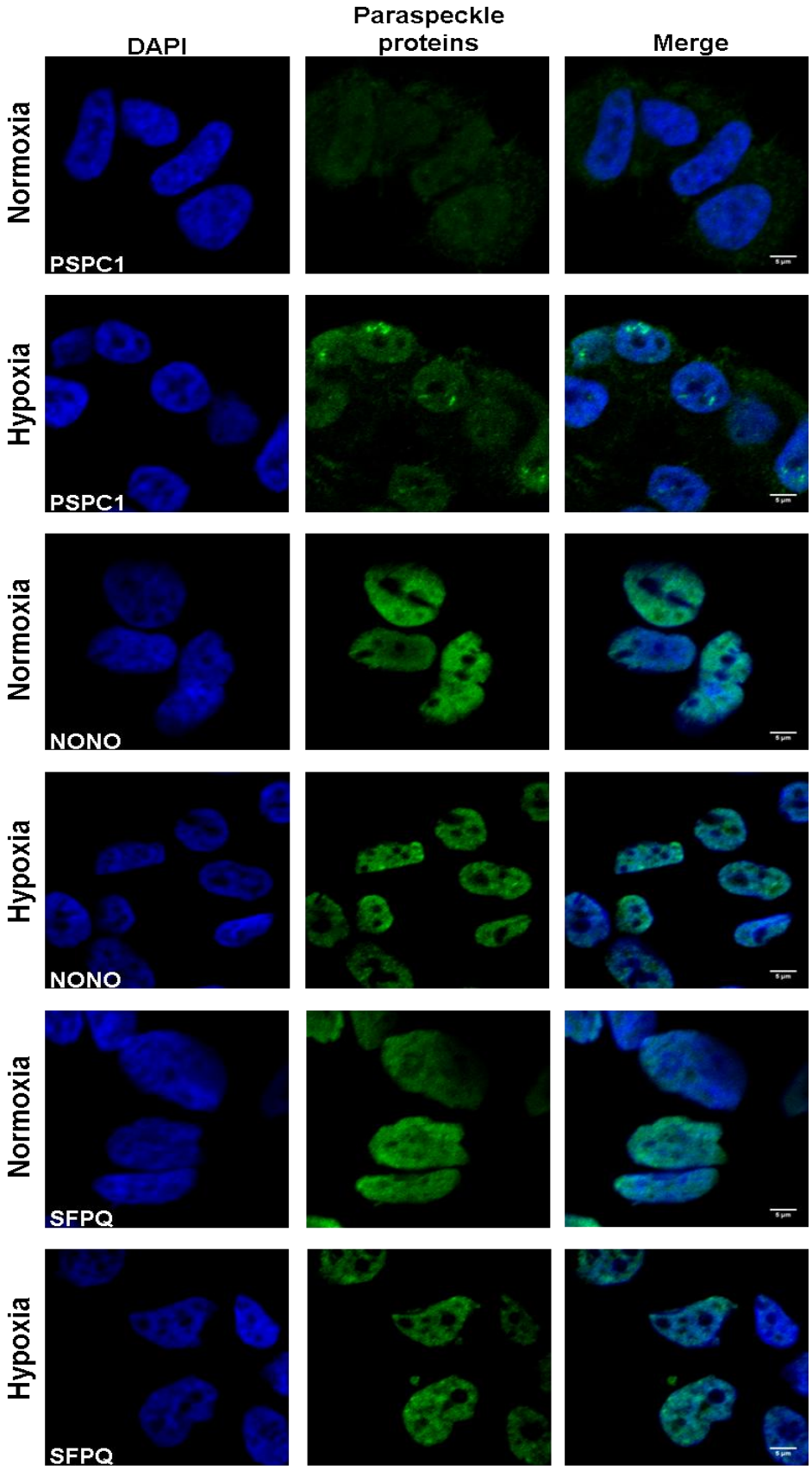


Figure 5.8 Hypoxia induces the formation of paraspeckle aggregates in ZR-75-1 cells. Immunostaining for PSPC1, NONO and SFPQ proteins in ZR-75-1 cells in normoxia and hypoxia. Proteins were diffused in normoxic condition, but in hypoxia there was a clear increase in the number of paraspeckle foci (Scale 5um).

5.3.8 NEAT1 is essential for paraspeckle formation in hypoxia

Since NEAT1 is reported to be important for formation of paraspeckle, I hypothesised that the formation of paraspeckles in hypoxia is due to induction of NEAT1 in hypoxic conditions. To test this hypothesis, I used antisense oligonucleotides (ASO) to knockdown NEAT1 transcript. As a consequence of their intracellular localisation, nuclear transcripts have proven relatively resistant to knockdown by siRNA and shRNA technologies. However, antisense oligonucleotides, utilising the RNA H activity of cell, have previously been used to knockdown nuclear transcripts with high specificity and minimum off target affects (Ideue et al., 2009, Phillips and Zhang, 2000). Indeed, several studies have successfully used ASO to knockdown NEAT1 expression in different models (Nishimoto et al., 2013, Zhang et al., 2013a). In the current study, I used two different ASO sequences (ASO1 and ASO2) both targeting the NEAT1 transcript. These two ASOs have previously been shown to induce good knockdown of NEAT1 (Sasaki et al., 2009, Nishimoto et al., 2013). To validate the degree of knockdown, MCF-7 and ZR-75-1 cells were treated with ASO1, ASO2 or scrambled control ASO and cells were then exposed to hypoxia or normoxia. qPCR analysis showed >85 % knockdown of NEAT1 expression in hypoxia in both cell lines following treatment with either ASO1 or ASO2 **(Figure 5.9)**.

After NEAT1 suppression using ASO1, MCF-7 cells were exposed to hypoxia for 24 hours. Cells were then immunostained for PSPC1 and NONO. The number of paraspeckle foci was significantly decreased in hypoxic cells treated with NEAT1 ASO compared to hypoxic cells transfected with control scrambled ASO **(Figure 5.10)**. Thus NEAT1 is necessary for the formation of paraspeckle aggregates in hypoxia.

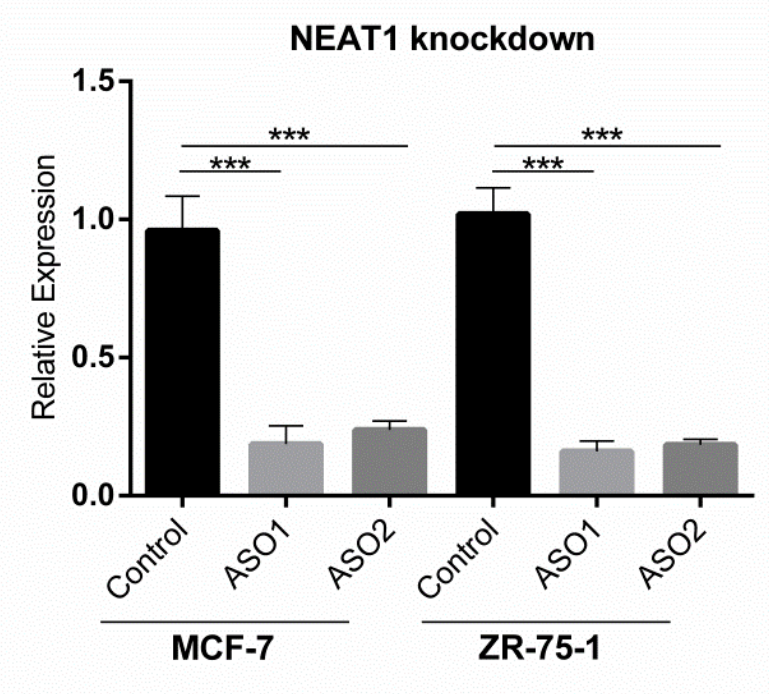


Figure 5.9 NEAT1 knockdown using antisense oligonucleotides (ASO) in hypoxia. QPCR analysis demonstrates that NEAT1 expression is significantly suppressed (>85 %) using two different ASOs in MCF-7 and ZR-75-1 cells grown in hypoxic conditions (1 % O_2 for 24 hours) compared to control scramble ASO. Each experiment was performed with three biological replicates (***p≤0.0005, Student's t-test, n=3).

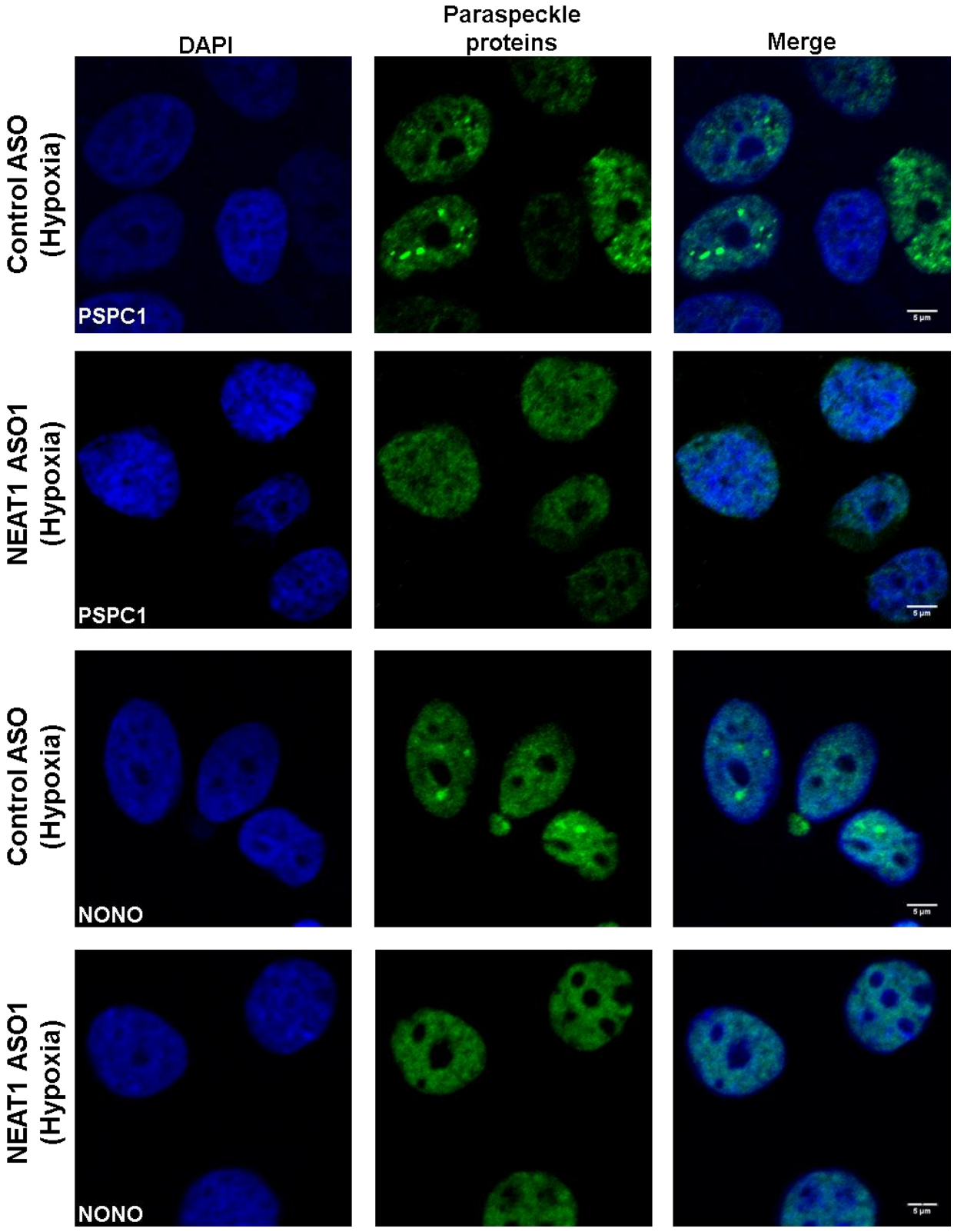


Figure 5.10 Hypoxia induces the formation of paraspeckle foci via NEAT1. Immunostaining for PSPC1 and NONO proteins in MCF-7 cells transfected with scramble control ASO and NEAT1 ASO1 in hypoxia. NEAT1 knockdown in hypoxia significantly decrease numbers of paraspeckle foci (Scale 5um).

5.3.9 HIF is required for the formation of paraspeckle aggregates in hypoxia

Since the hypoxic formation of paraspeckle aggregates requires NEAT1 and NEAT1 is induced in hypoxia by HIF-2α, I hypothesised that the hypoxic formation of paraspeckles foci would also require HIF-2α. I therefore treated MCF-7 cells with HIF-1α or HIF-2α siRNAs and exposed the cells to hypoxia as before. Hypoxic cells treated with scrambled control siRNA showed a clear increase in paraspeckle foci. Cells treated with HIF-1α siRNA showed a mild reduction in the number of paraspeckle foci in hypoxia **(Figure 5.11).** However, cells treated with HIF-2α siRNA exhibited almost complete absence of paraspeckle formation in hypoxia **(Figure 5.11).** Taken together, these results indicate that the increased formation of paraspeckles in hypoxia is dependent upon HIF-2α mediated induction of NEAT1.

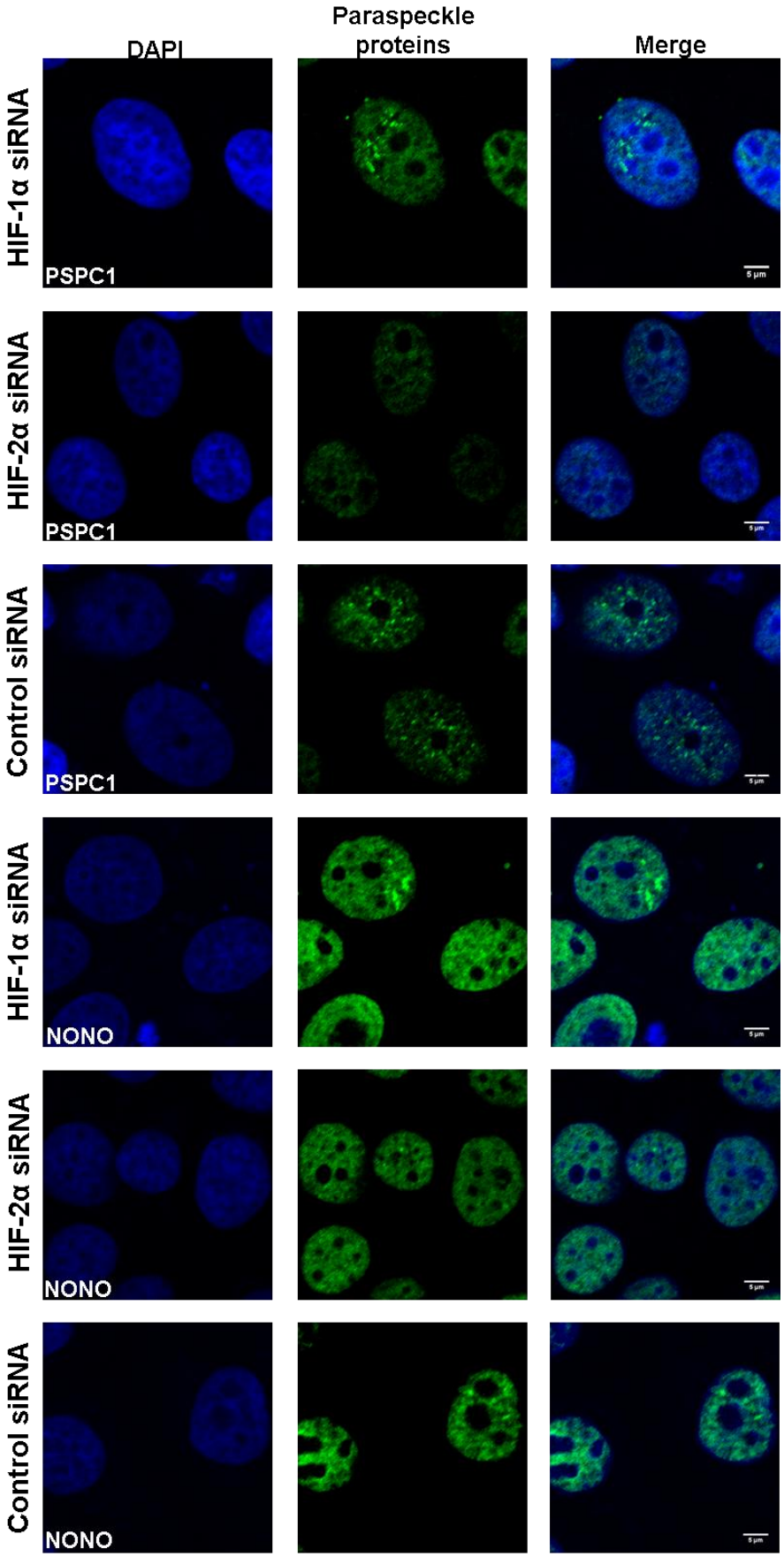


Figure 5.11 HIF-2α is required for the formation of paraspeckles in hypoxia. Immunostaining for PSPC1 and NONO proteins in MCF-7 cells transfected with scrambled siRNA, HIF-1α siRNA and HIF-2α siRNA in hypoxia. Numbers of paraspeckle foci were reduced with HIF-2α knockdown in hypoxia (Scale 5um).

5.3.10 NEAT1 knockdown in hypoxic MCF-7 and ZR-75-1 cells reduces cell proliferation and cell survival

Activation of hypoxia pathways and HIF in particular is associated with an aggressive tumour phenotype and poor clinical outcome in many types of cancer, including breast cancer. Whilst HIF-dependent activation of many coding genes is thought to contribute to these properties, the extent to which HIF-dependent activation of long non-coding transcripts contributes to the aggressive tumour phenotype remains unknown. I therefore examined the contribution of hypoxic NEAT1 expression to the behaviour of breast cancer both *in vitro* and *in vivo*.

To analyse the consequences of NEAT1 depletion in hypoxia, I used antisense oligonucleotides (ASO) as previously described in 5.2.8 to knockdown NEAT1 transcript. I then assessed the effect of NEAT1 knockdown on cell proliferation by measuring the cellular DNA content via fluorescent dye binding (CyQUANT® NF). Following transfection with NEAT1 ASO1 or scrambled control ASO, cells were incubated in hypoxia (1 % O_2 for 24 hours) and normoxia. Cellular proliferation was significantly reduced following NEAT1 knockdown in both MCF-7 and ZR-75-1 cells. Inhibition of cellular proliferation was observed in both hypoxia and normoxia. However, the suppression of cellular proliferation by NEAT1 ASO was more profound in hypoxia (>50 % reduction, $p<0.0001$) when NEAT1 levels would normally be higher **(Figure 5.12A).** I then examined whether NEAT1 knockdown would have an effect on cell survival using a clonogenic assay **(Figure 5.12B)**. NEAT1 suppression using ASO1 led to a significant decrease in clonogenic survival of both MCF-7 and ZR-571 cells in hypoxia when compared to a scrambled control ASO, while NEAT1 depletion caused similar or milder effects in normoxic cells **(Figure 5.12C-D).** The assay was repeated using the alternative ASO (ASO2) and showed similar effects in MCF-7 cells in both conditions, confirming that effects on cell survival were due to specific NEAT1 suppression rather than "off target" effects **(Figure 5.12B-D)**. Taken together, these findings indicate that NEAT1 promotes both cellular proliferation and clonogenic survival of cancer cells and that this phenotype is induced in hypoxia.

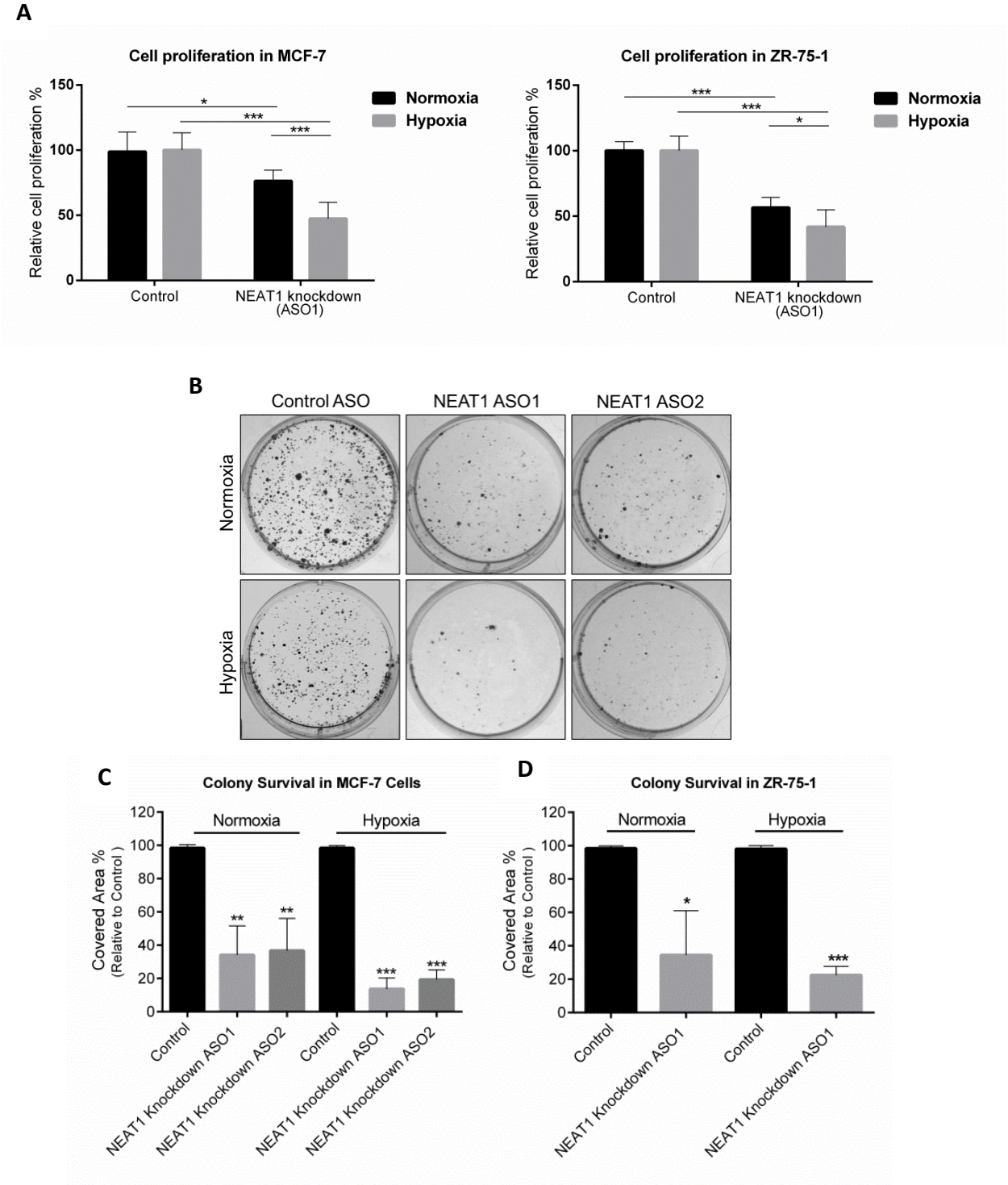


Figure 5.12 NEAT1 knockdown affects cell proliferation and cell survival. A) Cell proliferation rates relative to scrambled control ASO after NEAT1 knockdown in normoxic and hypoxic MCF-7 and ZR-75-1 cells (1 % O_2 for 24 hours). **B)** The influence of NEAT1 knockdown on formation of MCF-7 cells colonies in hypoxia 1 % O_2 for 24 hours and normoxia. Two different ASOs were used for knockdown and showed similar effects on cell survival. Calculation of colony survival relative to scrambled control ASO following NEAT1 suppression in **C)** MCF-7 cells **D)** ZR-75-1 cells in normoxia and hypoxia 1 % O_2 for 24 hours. Each experiment was performed with three biological replicates (*P<0.05, **p≤0.009, ***p≤0.0009, Student's t-test, n=3).

5.3.11 NEAT1 knockdown in hypoxic MCF-7 and ZR-571 cells increases apoptosis

Since NEAT1 depletion inhibits cellular proliferation and survival, I next determined whether NEAT1 suppression affects apoptosis by measuring caspase-3 & 7 activity, a key effector of apoptosis in mammalian cells. As previously, NEAT1 knockdown was achieved using ASO and ZR-75-1 cells were subjected to hypoxia (1 % O_2) or normoxia. NEAT1 knockdown resulted in the induction of caspase-3/7 activity by 1.61-fold in normoxia and 2.1-fold in hypoxia when compared to scrambled control ASO **(Figure 5.13A).** The caspease-3/7 activity was not measurable in MCF-7 cells (data not shown), since MCF-7 do not express caspase-3 (hiroseJanicke et al., 1998, Janicke, 2009). In addition, I also confirmed apoptotic activity by counting annexin V-positive cells using fluorescence-activated cell sorting (FACS) following NEAT1 suppression by ASO. Depletion of NEAT1 increased the number of annexin-V positive cells from 6.5 to 18.47 % ($P < 0.03$) in hypoxic MCF-7 cells **(Figure 5.13B)** Suppression of NEAT1 in normoxia did not show increase annexin-V positive cells, suggesting that NEAT1 induction in hypoxia is critical for cell survival. Taken together, these data indicate that NEAT1 suppression increases apoptosis significantly in hypoxic cells and to a lesser extent in normoxic cells (in ZR-75-1 cells), suggesting that NEAT1 has anti-apoptotic properties in hypoxia.

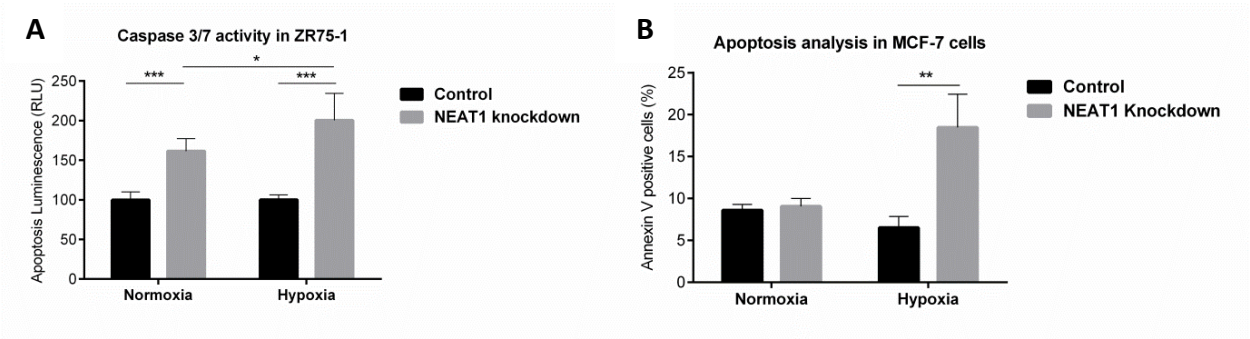


Figure 5.13 NEAT1 knockdown induces apoptosis. A) Induction of caspase3/7 activity after NEAT1 knockdown in normoxia and hypoxia (1 % O_2 for 24 hours) of ZR-75-1 cells. **B)** FACS analysis for annexin-V stain cells with and without NEAT1 knockdown in hypoxia (1 % O_2 for 24 hours) and nromoxia in MCF-7 cells (n=3, *p=0.03, **p=0.003, ***p<0.0001, Student's t-test).

5.3.12 Genetic alteration of NEAT1 and paraspeckle components in cancers

Genetic alteration of NEAT1 and paraspeckle components (PSPC1, NONO and SFPQ) has not been described before in cancer. I therefore examined the frequency of somatic variation of NEAT1, PSPC1, NONO and SFPQ in different tumour samples. Analysis of data from The Cancer Genome Atlas (TCGA - http://www.cbioportal.org/public-portal/), and other cancer sequencing projects demonstrates NEAT1 copy number gains in a number of cancer types, although with low frequency **(Figure 5.14A).** On the other hand, PSPC1, NONO and SFPQ showed a large number of somatic alterations, including mutations, amplifications and a few deletions in a variety of cancer types, including bladder, prostate and ovarian serous cancers **(Figure 5.14B-D).** These data suggest that NEAT1 and paraspeckle components harbour genetic changes in different cancer types. Mutation in cancer genes causes activation or inhibition of pathways that are likely to interfere with gene function. NEAT1 and paraspeckle components are commonly found amplified and mutated in cancers, indicating the potential role for this pathway in cancer.

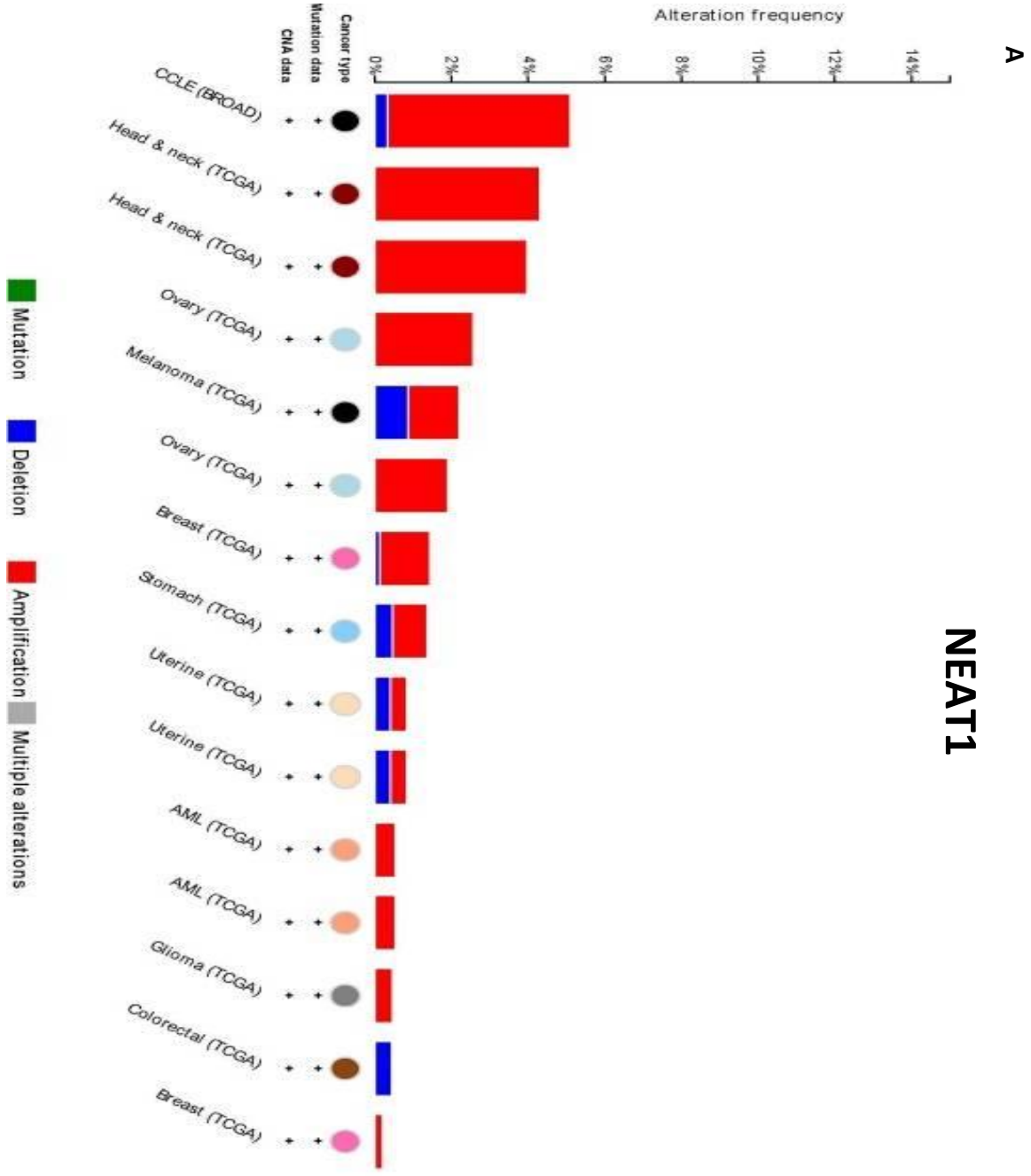


Figure 5.14 Somatic alterations of NEAT1 in different cancers. A) Percentage of each listed tumour type harbouring amplification (red), somatic mutation (green), deletion (blue) and multiple alterations (grey) in NEAT1. Plots were generated from the cBioPortal for Cancer Genomic (http://www.cbioportal.org/public-portal/).

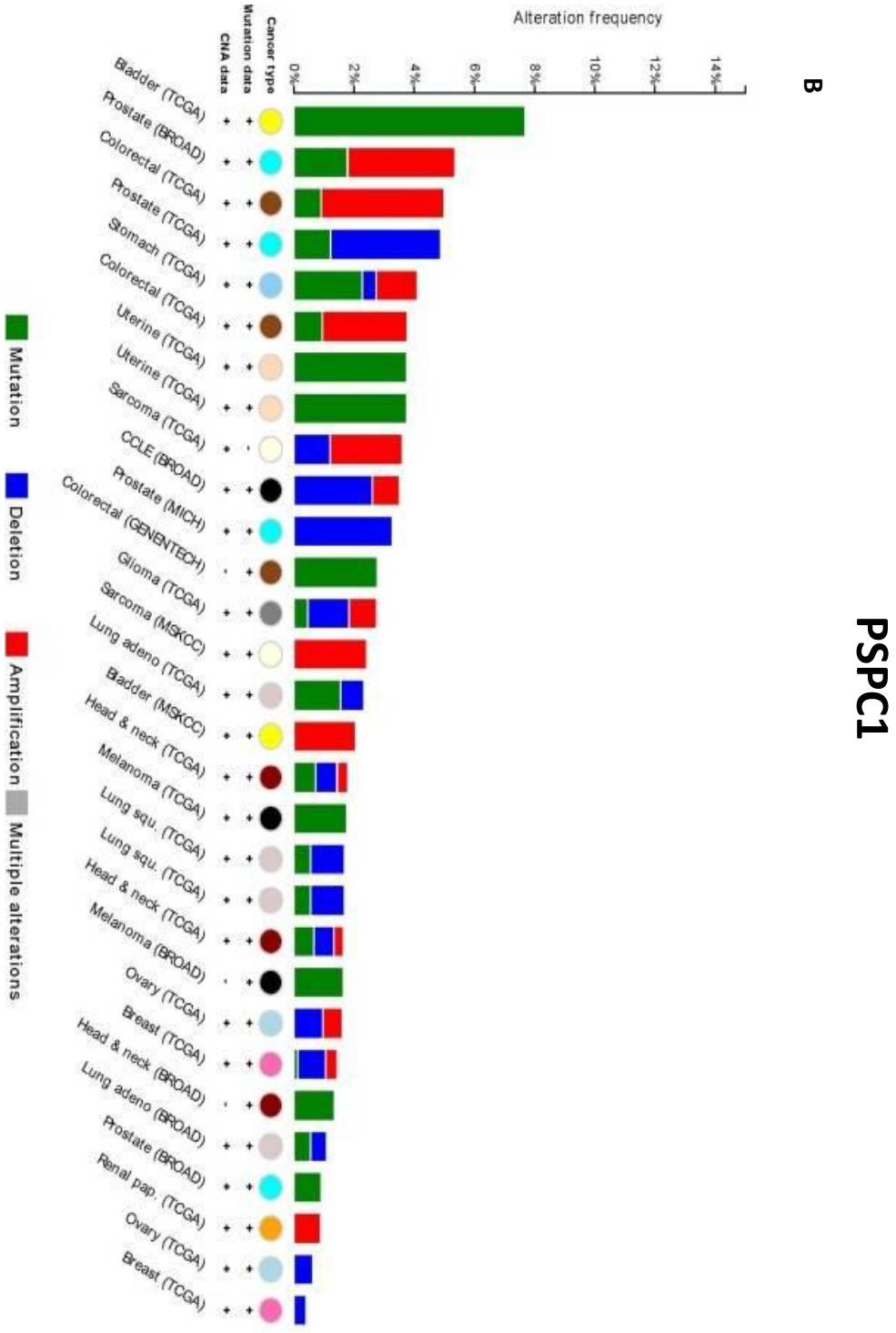


Figure 5.14 Somatic alterations of PSPC1 in different cancers. B) Percentage of each listed tumour type harbouring amplification (red), somatic mutation (green), deletion (blue) and multiple alterations (grey) in PSPC1. Plots were generated from the cBioPortal for Cancer Genomic (http://www.cbioportal.org/public-portal/).

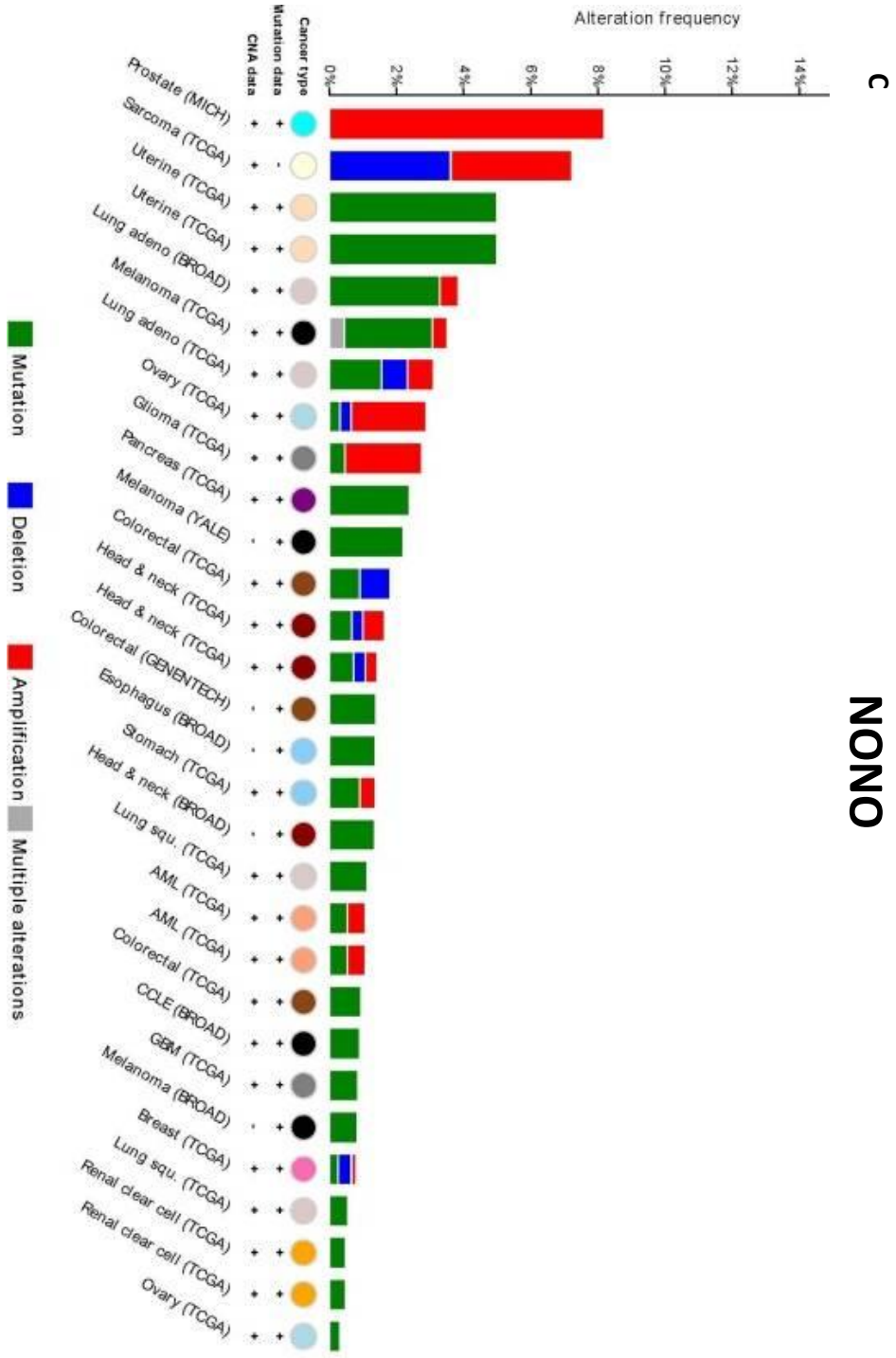


Figure 5.14 Somatic alterations of NONO in different cancers. C) Percentage of each listed tumour type harbouring amplification (red), somatic mutation (green), deletion (blue) and multiple alterations (grey) in NONO. Plots were generated from the cBioPortal for Cancer Genomic (http://www.cbioportal.org/public-portal/).

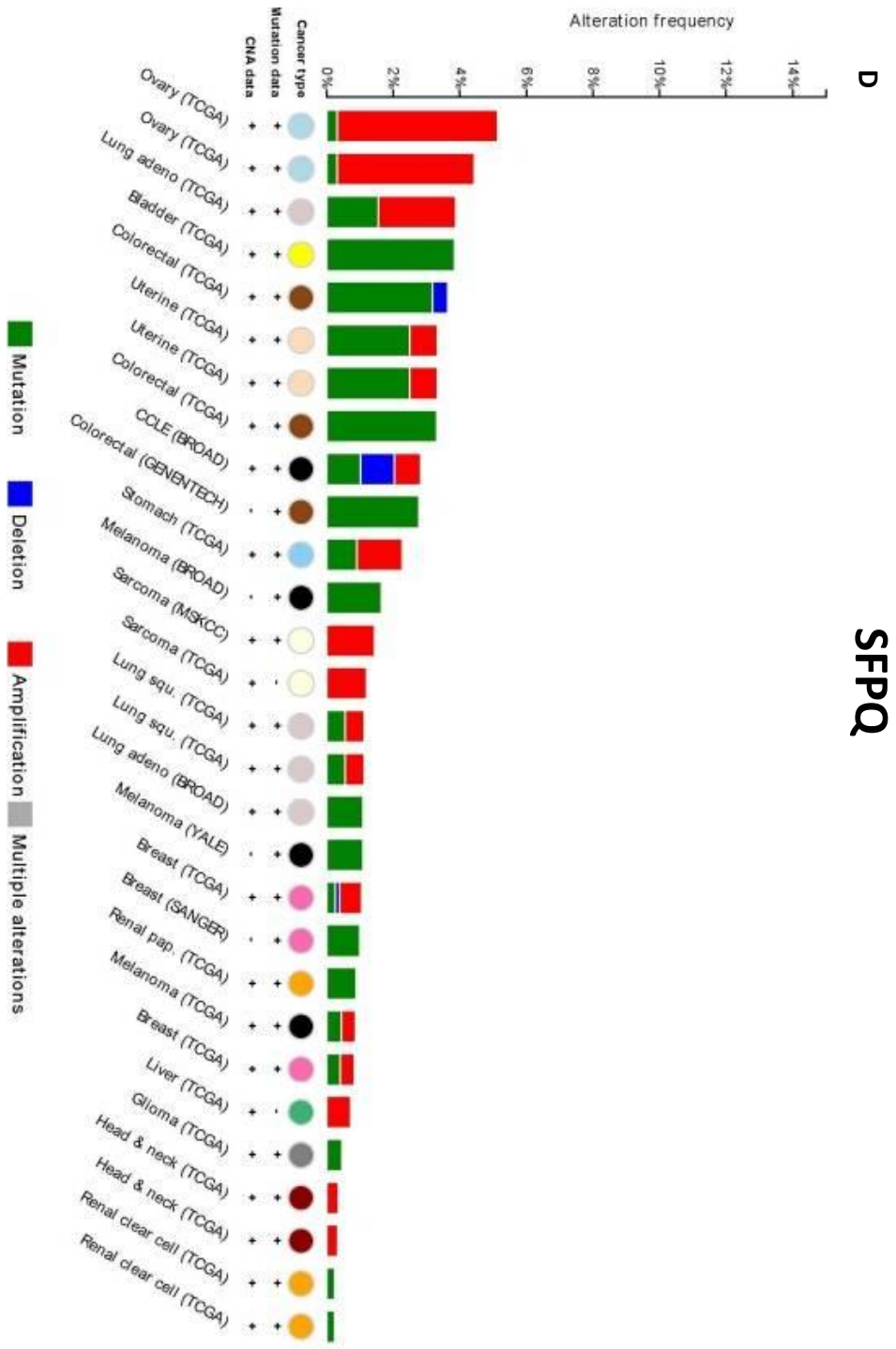


Figure 5.14 Somatic alterations of SFPQ in different cancers. D) Percentage of each listed tumour type harbouring amplification (red), somatic mutation (green), deletion (blue) and multiple alterations (grey) in SFPQ. Plots were generated from the cBioPortal for Cancer Genomic (http://www.cbioportal.org/public-portal/).

5.3.13 High NEAT1 tumour expression is associated with poor survival in a cohort of 2000 breast cancer patients

The expression levels of a number of lncRNAs are associated with clinical outcome and pathological features in a variety of cancers (Du et al., 2013, Song et al., 2013). To evaluate the clinical relevance of NEAT1 in breast cancer, I utilized published expression data derived from a large (n=2000) cohort of breast cancers for which clinical outcome data were also available (Curtis et al., 2012). The median expression of tumour NEAT1 was calculated across the cohort. Patients were then divided into two categories; patients with high tumour NEAT1 expression (>median expression), and patients with low tumour NEAT1 expression (<median expression). Kaplan-Meier analysis of overall mortality revealed that patients with high tumour NEAT1 expression have a significantly poorer clinical outcome when compared to patients with low tumour NEAT1 expression (P=0.005, Hazard Ratio "HR" 1.22, 95 % CI= 1.06-1.41) **(Figure 5.15A).** Multivariate analyses of NEAT1 expression, adjusted for clinicopathological features known to correlate with outcome (age, tumour size, grade, stage and node status), showed that NEAT1 remains significantly associated with poor outcome after adjustment for age (p= 0.013), tumour size (p= 0.008), grade (p= 0.017) and node status (p= 0.007) **(Table 5.1).** This indicates that tumour NEAT1 expression is an independent prognostic marker in breast cancer patients at least with respect to these clinical variables.

Table 5.1 High tumour NEAT1 expression is associated with poor patient survival in breast cancer independently of clinicopathological features:

Variable	*P* value	Hazard Ratio (HR)	95 % CI	
			Lower	Upper
Age	0.013	1.2	1.04	1.38
Size	0.008	1.21	1.058	1.4
Grade	0.017	1.19	1.38	1.38
Node Status	0.007	1.21	1.05	1.4
Stage	0.186	1.12	0.95	1.32

I then examined whether high tumour NEAT1 expression is associated with clinical outcome in different subtypes of breast cancer. The cohort of 2000 breast cancers were sub-classified according to estrogen receptor (ER) status, human epidermal growth factor receptor 2 (HER2) status and progesterone receptor (PR) status. I found that high tumour NEAT1 expression was associated with an adverse outcome in ER+ tumours (p=0.026) and in HER2- tumours (p=0.006). Patients with HER2-/ER+ tumours did not show a significant (P=0.068) association between tumour NEAT1 expression and overall survival. However, when interpreting these results it should be noted that the number of patients with HER2-/ER+ tumours was considerably fewer (n=121) than those who had ER+ or HER2- tumours **(Table 5.2, Figure 5.16).**

Next I performed qPCR analysis of an independent cohort of 150 breast cancers and found that high tumour NEAT1 expression correlates with tumour size (P=0.03, Rho= 0.17). Despite this correlation, high tumour NEAT1 levels had predicted patient survival independently of tumour size in the previous analysis of 2000 patients with breast cancer. This suggests that tumour NEAT1 levels have a strong influence on patient prognosis.

In addition, I correlated tumour NEAT1 expression levels with a previously described hypoxia gene expression signature (Buffa et al., 2010)**.** This signature was derived from selecting genes that were consistently co-expressed with hypoxia-regulated genes (Seeds) in multiple cancers including head and neck, breast and lung cancers. The signature is highly enriched for hypoxia-regulated pathways and reflects tumour response to hypoxia *in vivo* (Buffa et al., 2010). Applying this signature to the cohort of 150 breast cancers, I found that tumour NEAT1 expression is positively correlated with the signature of hypoxia upregulated genes (p=0.01 Rho correlation 0.2) and negatively correlated with the hypoxia downregulated signature (p=0.0004, Rho correlation -0.28) in 150 cases. However, due limited number of samples, the difference in overall survival was not significant.

5.2.14 Paraspeckle component gene expression in breast cancer cohort

Although the expression of paraspeckle components did not change in hypoxia, the presence of high levels of paraspeckle component mRNAs may indicate activation and increase formation of paraspeckle in cancer and its progression. I next examined for associations between the tumour expression levels (mRNA) of the main paraspeckle components and overall survival in the cohort of 2000 breast cancer patients. High tumour expression of NONO mRNA (P<0.00001, HR=1.31, 95 % CI 1.14 - 1.51) and SFPQ mRNA (P=0.002, HR=1.25, 95 % CI 1.09-1.44) were significantly associated with adverse prognosis in breast cancer patients **(Figure 5.15C-D).** Analysis of NONO and SFPQ expression in different breast cancer sub-types (classified according to receptor status as before) revealed significant associations between mRNA levels and overall patient survival in the same subgroups as NEAT1. Specifically, both genes were significantly associated with outcome in patients with ER+ tumours (NONO, p=0.0009 & SFPQ, p=0.002) and HER2- tumours (NONO, p=0.005 & SFPQ, p= 0.001) **(Table 5.2, Figures 5.17 & Figure 5.18)**. On the other hand, PSPC1 mRNA levels showed no association with overall survival in the total cohort of 2000 breast cancer patients **(Figure 5.15B),** or when patients

were subdivided according to the receptor subtype of the tumour **(Table 5.2, Figures 5.19).** Taken together, these analyses indicate that high tumour expression of NEAT1 and the paraspeckle components, NONO and SFPQ, are associated with adverse clinical outcomes independently of other clinicopathological features of breast cancer. Thus, they may act as a potential prognostic biomarkers for breast cancer.

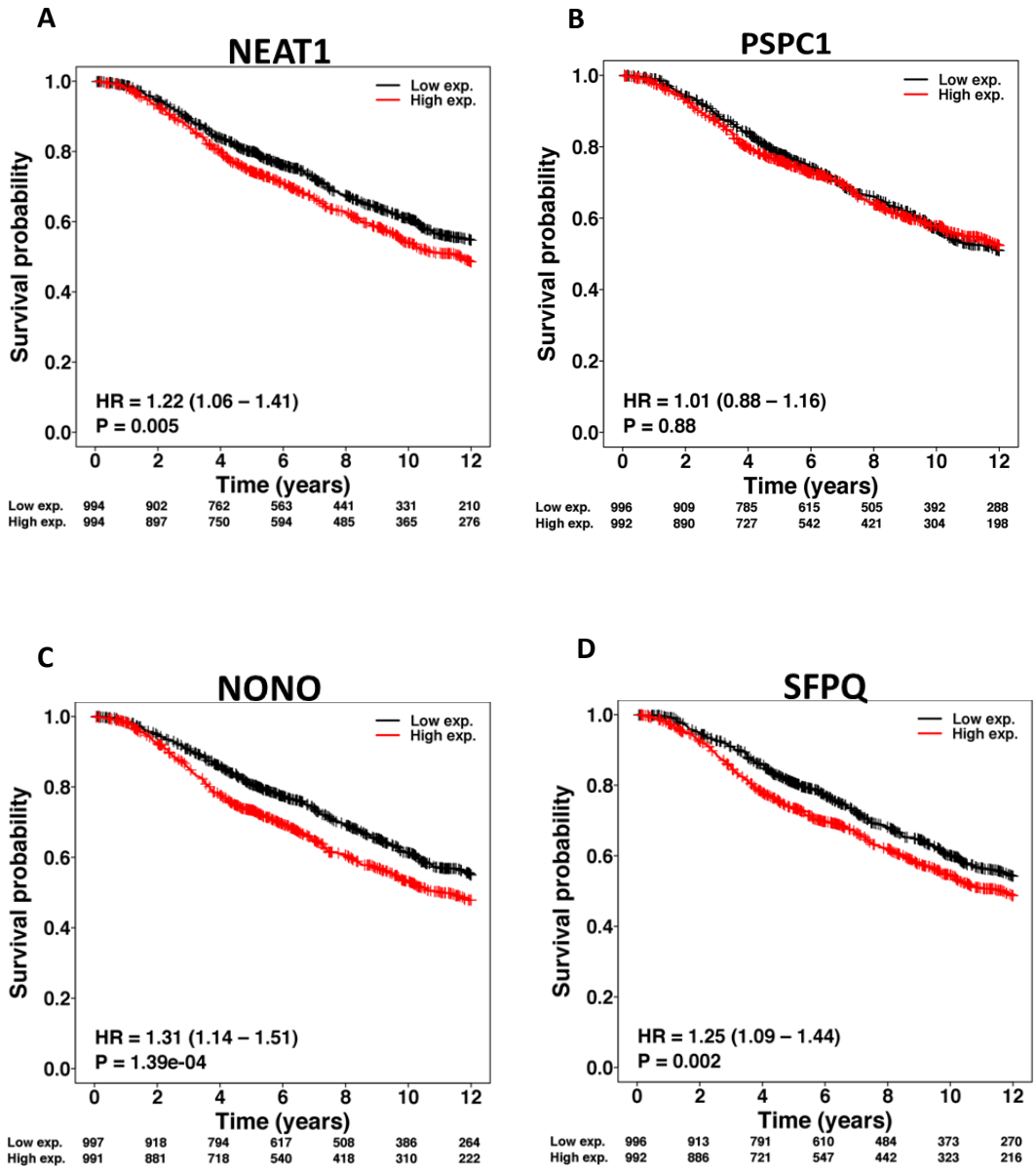


Figure 5.15 Tumour NEAT1 and paraspeckle component expression levels are associated with adverse clinical outcome in breast cancer. Kaplan-Meier curves showing significant association between tumour expression level of A) NEAT1, B) PSPC1 C)NONO and D) SFPQ expression and poor overall survival in a cohort of 2000 breast cancer patients.

Table 5.2 The association between tumour NEAT1 and paraspeckle component expression is restricted to specific receptor sub-types of breast cancer:

Type	P value	Hazard Ratio (HR)	95 % Cl		Number of samples
NEAT1			**Lower**	**Upper**	
ER+	0.02	1.21	1.02	1.42	1508
ER-	0.44	1.12	0.85	1.47	437
HER2+	0.62	0.92	0.65	1.30	247
HER2-	0.006	1.24	1.06	1.45	1741
HER2+ ER+	0.74	0.92	0.56	1.51	118
HER2+ ER-	1.00	1.00	0.61	1.64	122
Triple -	0.40	1.16	0.82	1.66	276
PSPC1					
ER+	0.79	0.98	0.83	1.15	1508
ER-	0.85	0.97	0.74	1.28	437
HER2+	0.83	0.96	0.68	1.37	247
HER2-	0.88	0.99	0.85	1.15	1741
HER2+ ER+	0.34	0.78	0.47	1.29	118
HER2+ ER-	0.45	1.21	0.74	1.98	122
Triple -	0.97	0.99	0.70	1.41	276
NONO					
ER+	0.0009	1.32	1.12	1.56	1508
ER-	0.32	1.15	0.87	1.52	437
HER2+	0.51	1.12	0.79	1.59	247
HER2-	0.005	1.24	1.07	1.45	1741
HER2+ ER+	0.14	1.46	0.89	2.42	118
HER2+ ER-	0.33	0.78	0.47	1.28	122
Triple -	0.06	1.41	0.99	2.01	276
SFPQ					
ER+	0.002	1.30	1.10	1.53	1508
ER-	0.32	1.15	0.87	1.52	437
HER2+	0.57	1.11	0.78	1.57	247
HER2-	0.001	1.29	1.10	1.50	1741
HER2+ ER+	0.49	1.19	0.72	1.96	118
HER2+ ER-	0.77	0.93	0.57	1.53	122
Triple -	0.20	1.26	0.88	1.79	276

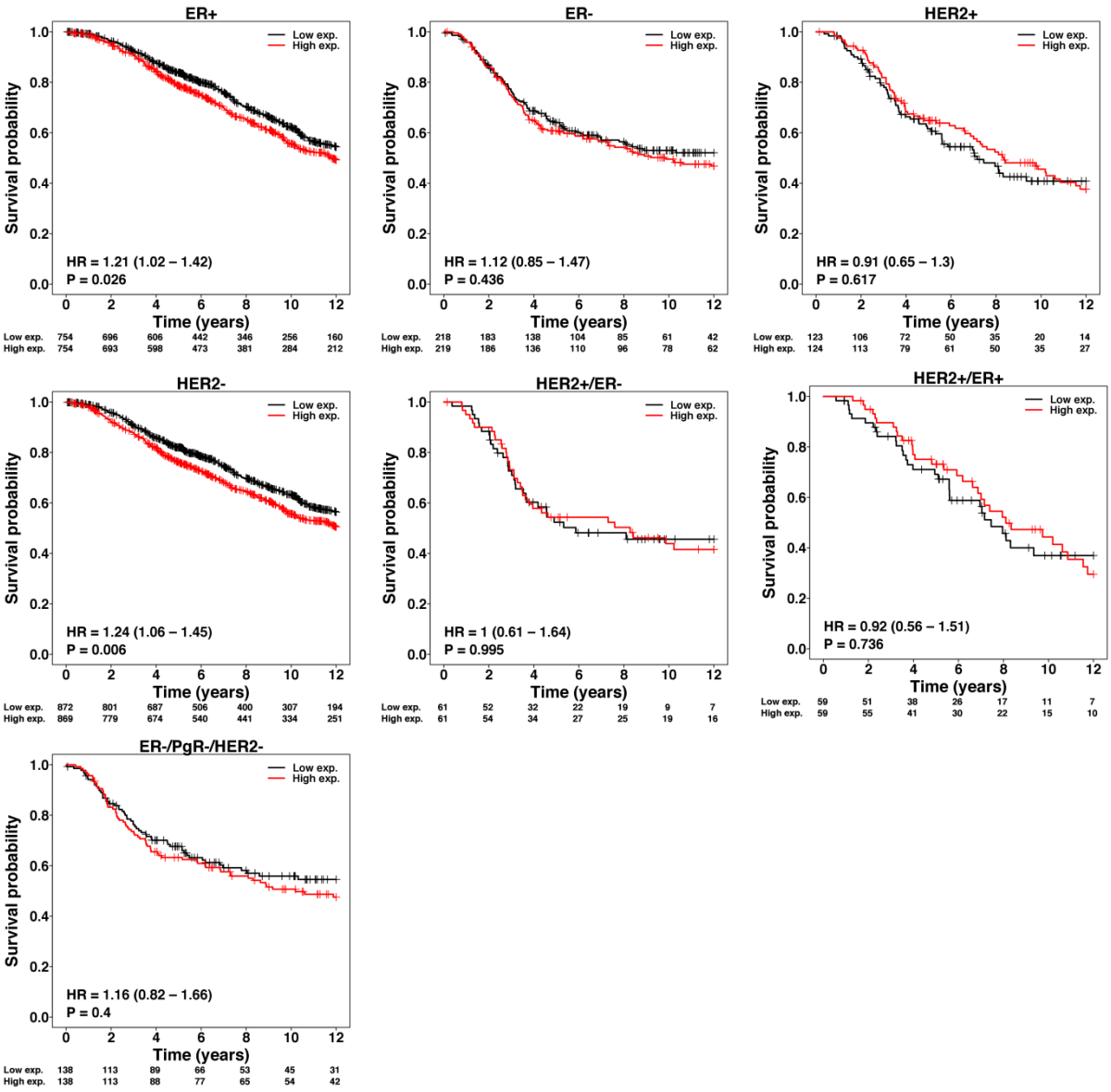


Figure 5.16 Clinical association of NEAT1 in different type of breast cancer. Kaplan-Meier curves for NEAT1 showed significant association of high NEAT1 expression with ER+ and HER- types of breast cancer.

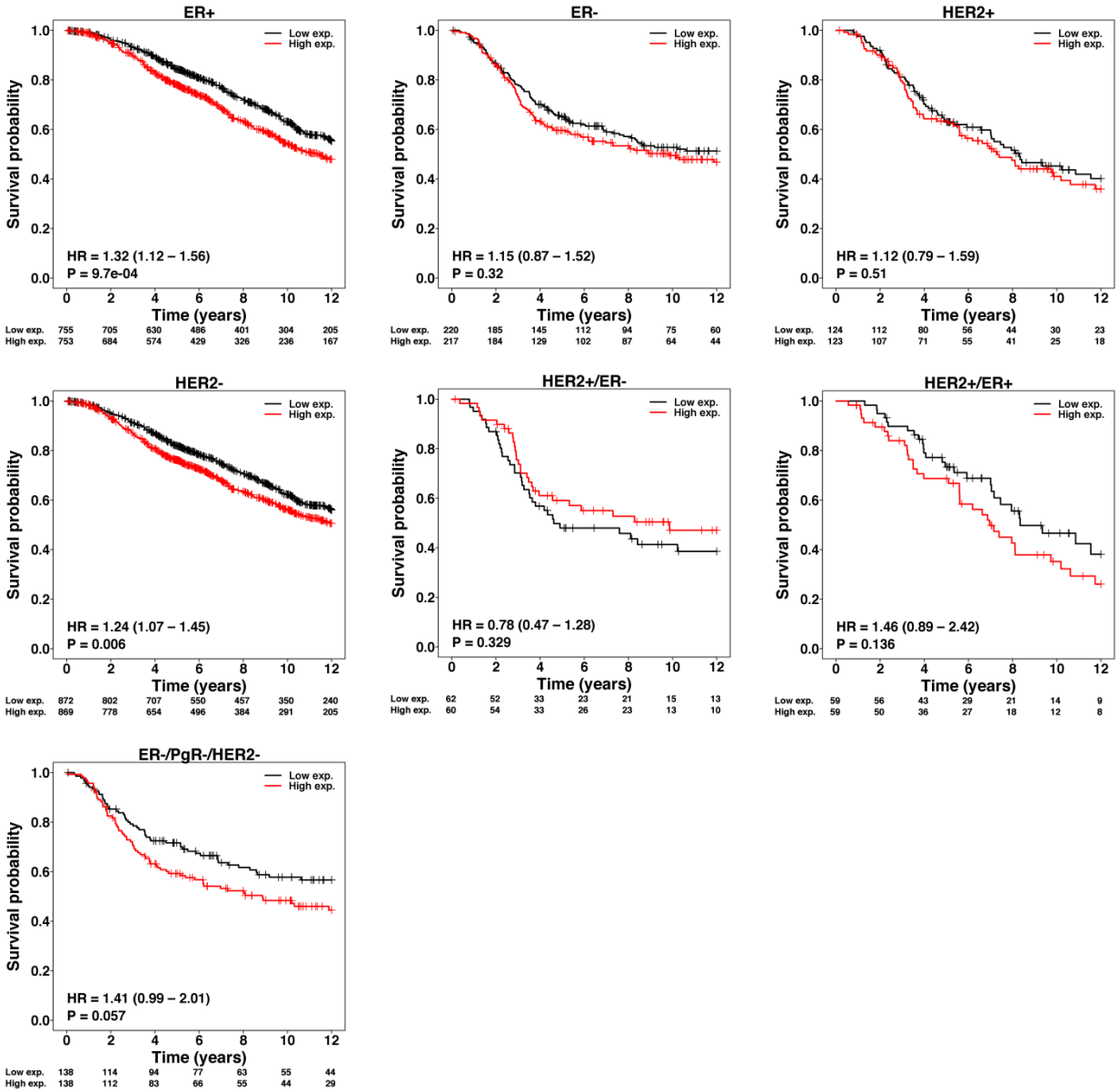


Figure 5.17 Clinical association of NONO in different type of breast cancer. Kaplan-Meier curves for NONO and SFPQ showed significant association of high NONO expression in ER+ and HER- types of breast cancer.

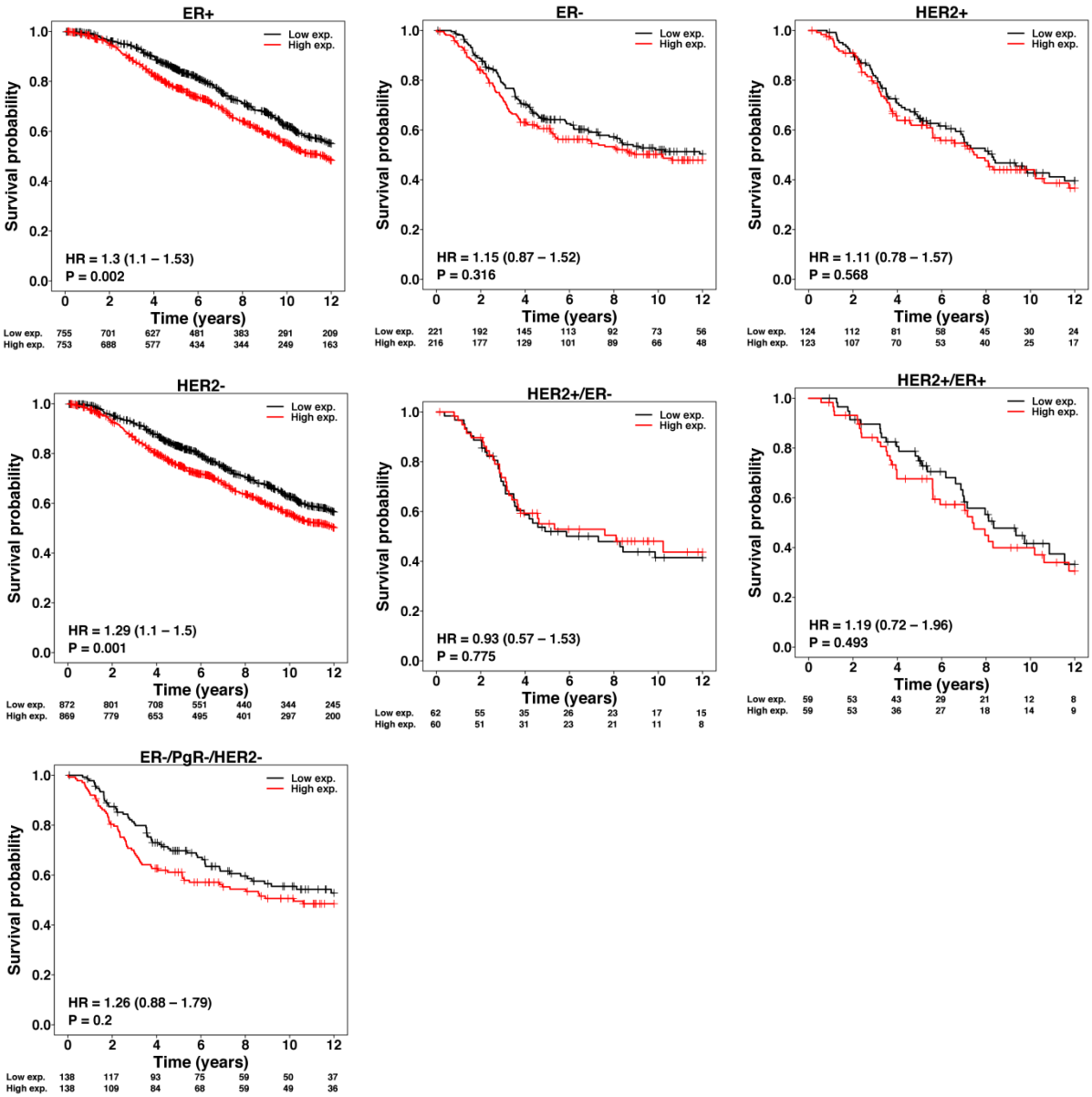


Figure 5.18 Clinical association of SFPQ in different type of breast cancer. Kaplan-Meier curves for NONO and SFPQ showed significant association of high SFPQ expression in ER+ and HER- types of breast cancer.

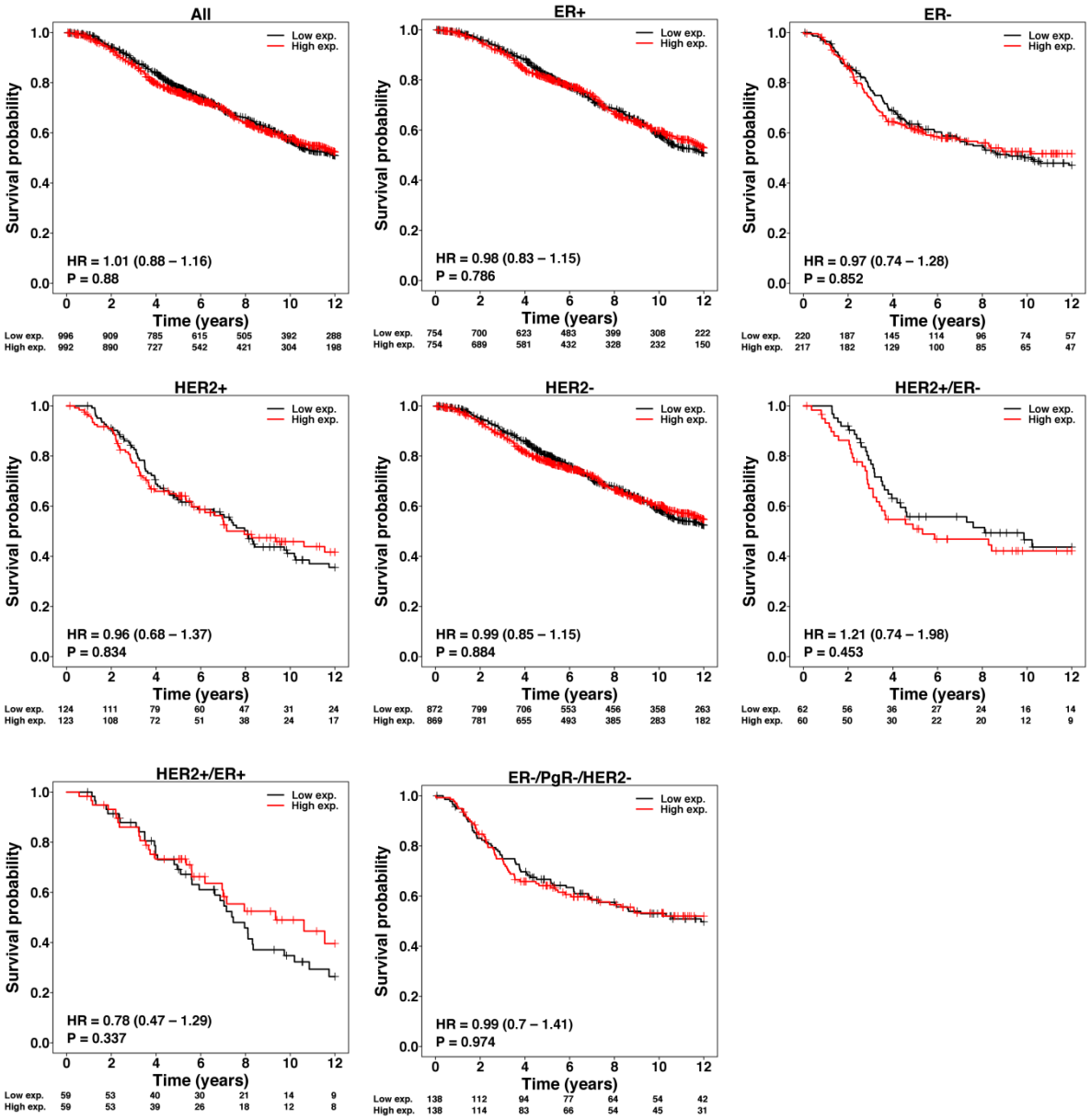


Figure 5.19 Kaplan-Meier analysis for PSPC1 in different type of breast cancer. There was no clinical association between PSPC1 expression and overall outcome and different type of breast cancer.

5.4 Discussion

Chapter five of this thesis identified lncRNAs as frequent transcriptional targets of hypoxic transcriptional pathways mediated by HIF. Amongst these apparent direct transcriptional targets of HIF, NEAT1 was the most hypoxically upregulated lncRNA. Analysis of ChIP-seq and RNA-seq signals across the datasets at this locus revealed the "typical" features observed at other direct transcriptional targets of HIF (Figure 5.1A). Specifically, RNA Poll is already pre-bound at TSS in normoxia and is released in hypoxia with increased RNApol2 ChIP-seq signal throughout the NEAT1 gene body. Concurrently, the H3K4me3 ChIP-seq signal broadens in hypoxia. Both these changes are consistent with increased transcriptional activity at the NEAT1 locus in hypoxia.

Moreover, hypoxia increases NEAT1 RNA expression not only in MCF-7 cells, but also in a range of breast cancer cell lines derived from tumours with differing receptor status. The two exceptions that do not show hypoxic induction of NEAT1 are BT-474 cells, which showed slight downregulation (not significant) of NEAT1 RNA and BT-549 cells, which show no change in expression. MCF-7 and ZR-75-1 are both ER+ positive cells and these cell lines had the highest induction of NEAT1 (3.7- and 3.1-fold respectively) amongst the screened cell lines and were therefore used for subsequent analyses **(Figure 5.1B).** Both cell lines also showed significant NEAT1 induction with DMOG treatment, suggesting that NEAT1 expression is 2-oxoglutarate-dependent as well as oxygen-dependent. Finally, consistent with its classification as a non-coding RNA, hypoxia-induced NEAT1 remained localised in the cell nucleus in hypoxia.

HIF is a major regulator of transcription in hypoxia (Ratcliffe et al., 1998, Semenza, 1998, Ratcliffe, 2013) and the regulation of protein coding RNAs and microRNAs by HIF is well described. However, the role of HIF as a widespread regulator of lncRNAs is a newly described feature of the current work. In this pan-genomic analysis, NEAT1 was the most hypoxia-upregulated lncRNA that had a HIF-binding site close to the promoter. Furthermore,

suppression of HIF-2α, but not HIF-1α, by siRNA demonstrated that the hypoxic induction of NEAT1 was dependent upon HIF-2α alone when analysed by RNA-seq, by qPCR and by RNA FISH. Taken together this provides strong evidence that NEAT1 is directly and specifically transcriptionally activated by HIF-2. This is despite binding of both HIF-1 and HIF-2 isoforms to the NEAT1 promoter region and indicates the operation of post-binding mechanisms of transcriptional selectivity, similar to those seen at protein-coding genes such as cyclin D1 (CCND1), transforming growth factor alpha (TGFβ), vascular endothelial growth factor (VEGF), BNIP3 and carbonic anhydrase 9 (CA9)(Raval et al., 2005).

Antiangiogenic therapies induce tumour hypoxia (Ebos et al., 2009, Rapisarda and Melillo, 2009). Several studies have used xenografts treated with the antiangiogenic drug (Bevacizumab) as models to study hypoxic microenvironments *in vivo* (Favaro et al., 2012, Blouw et al., 2003, McIntyre et al., 2012). In this chapter, MCF-7, MDA-MB-231 and MDA-MB-468 xenografts were treated with Bevacizumab to evaluate NEAT1 expression *in vivo*. NEAT1 expression was upregulated in the Bevacizumab treated group, indicating that NEAT1 is increased by the hypoxic microenvironment *in vivo* as well as *in vitro*.

NEAT1 has a scaffolding function essential to the formation of paraspeckles (Sasaki et al., 2009, Chen and Carmichael, 2009, Clemson et al., 2009). These paraspeckle bodies contain several RNA binding proteins, including PSPC1 and NONO, which may act as RNA depots retaining RNA transcripts in the nucleus (Nakagawa and Hirose, 2012). NEAT1 expressed in hypoxia shows a punctate nuclear distribution. Furthermore, whilst the total expression of PSPC1 and NONO proteins were not altered in hypoxia, immunostaining for these proteins and for SPQF showed a diffuse pattern of expression in normoxia with clear aggregation of protein into a punctate distribution in hypoxia. Furthermore, this hypoxic aggregation of paraspeckle-associated proteins into a punctate distribution not only required NEAT1, but was also dependent on HIF-2α. Taken together these findings indicate that in hypoxia, HIF-2α dependent transcriptional

activation of NEAT1 leads to aggregation of nuclear paraspeckles in a manner similar to that described for other stress conditions (Nakagawa and Hirose, 2012, Zhang et al., 2013a).

The role of NEAT1 in cancer and oncogenesis has not been previously studied. I therefore evaluated the impact of NEAT1 knockdown on the behaviour of MCF-7 and ZR-75-1 cells in various models of tumourigenicity, in hypoxia and normoxia. Both cellular proliferation and colony survival were reduced following NEAT1 suppression. Effects were seen in both normoxia and hypoxia, but were more profound in hypoxia in which NEAT1 levels would normally be higher. Moreover, apoptosis was increased upon NEAT1 knockdown in hypoxia measured by capases 3 & 7 activity and annexin V positivity in ZR-75-1 and MCF-7 cells, respectively. These findings suggest that NEAT1 may have an oncogenic role in hypoxic tumours.

Cancers arise as a result of somatic changes at gene loci that lead to disordered growth of tumour cells. Recently, a number of large genome-wide studies have provided powerful tools to understand the genetic makeup of many cancers and in the discovery of novel cancer-related genes. I utilized genomic data from the TCGA and other sequencing projects deposited at *cbioportal* to investigate somatic alterations in NEAT1, PSPC1, NONO and SFQP in different cancers. Copy number gains at the NEAT1 locus were reported, albeit at low frequency, in a number of cancer types including head & neck cancers, ovarian cancer and melanoma. In addition, PSPC1, NONO and SFPQ showed a spectrum of genetic alteration including mutations, amplifications and a few reported deletions in a range of cancer types. These data may underestimate somatic changes in NEAT1 locus such as mutation, since to date most large genomic mutational analyses have focussed on protein-coding RNAs. Overall, these findings indicate that NEAT1 and paraspeckle components are the subject of somatic genetic changes in assorted types of cancer. Further research will be needed to evaluate the impact of these changes on NEAT1 and paraspeckle function.

Although expression levels of some paraspeckle proteins are associated with cancer (Clark et al., 1997, Schiffner et al., 2011, Duhoux et al., 2011, Jiang et al., 2013), the clinical relevance of

NEAT1, PSPC1, NONO and SFPQ in breast cancer has not been evaluated previously. Using expression data derived from a cohort of 2000 patients with breast cancer (Curtis et al., 2012), I found that high levels of NEAT1, NONO and SFPQ were significantly associated with poor prognosis. In particular, tumour NEAT1 expression was significantly associated with overall mortality, in a manner independent of other clinicopathological features associated with adverse prognosis in breast cancer.

This is the first report to investigate the role of NEAT1 lncRNA in the hypoxia responsive signalling pathway and in the tumourigensis of breast cancer. Overall, these data suggest that NEAT1 is likely to have an important function in the hypoxic microenvironment of breast cancer and may have clinical value. Paraspeckles are reported to be involved in regulating gene expression through nuclear retention of A-to-I-edited mRNAs (Chen and Carmichael, 2009, Zhang et al., 2013a, Hirose et al., 2014), although their functions remain incompletely defined. Their role in control of RNA editing and its retention in hypoxia will be a key question for further research.

Chapter 6: Discussion

6. Discussion

In humans, aberrant oxygen homeostasis is associated with many common pathological conditions, while reduced oxygen tension (hypoxic microenvironment) is a hallmark of most solid tumours. An hypoxic microenvironment contributes to the failure of many anti-cancer therapies in clinical oncology including radiotherapy, chemotherapies and many metabolic and epigenetic therapies (Wilson and Hay, 2011). Failure or resistance to these therapies in a hypoxic microenvironment is due to many mechanisms, including reduced drug concentrations as a consequence of poor blood flow, cellular acidification, resistance to apoptosis, suppression of DNA repair, lack of oxidation of DNA free radicals, cell cycle arrest and genomic instability (Minchinton and Tannock, 2006). In hypoxia, cancer cells shift their energy production from oxidative phosphorylation (tricarboxylic acid (TCA) cycle) to glycolysis, a phenomenon termed the 'Warburg effect' (Vander Heiden et al., 2009, Gatenby and Gillies, 2004, Schulze and Harris, 2012).

Hypoxia triggers cancer cells to alter the expression of hundreds of genes (including carbohydrate and lipid metabolism genes) that allow them to survive the hypoxic stress and promote a malignant and metastatic phenotype. It is therefore associated with poor clinical outcome. In consequence, hypoxic tumour cells represent an important target for new therapeutic interventions, so examining changes in gene expression in hypoxia will provide a better understanding of transcriptional gene regulation in hypoxia responses.

Although previous studies of metabolism, proteins and mRNA transcription in hypoxia have provided considerable insight into the molecular physiology of cancer cells, our understanding of the global transcriptional response to hypoxia is still evolving. This is due to rapid advancement of next-generation sequencing technologies and better understanding of new regulatory RNAs and functional elements in the genome. The central dogma of molecular biology as originally proposed stated that encoded genetic information flows in a unidirectional

manner from DNA to RNA to protein, and that proteins are ultimately responsible for the cellular phenotype (Crick, 1970). However, this dogma is being challenged by the discovery of non-coding RNAs which constitute as much as 90 % of the transcribed genome (Consortium et al., 2012, Djebali et al., 2012). These non-coding RNAs include the well-characterised rRNAs and tRNAs, as well as miRNAs, which have been studied and documented as useful biomarkers with potential therapeutic targets in many cancers. lncRNAs, piwiRNAs and sn/snoRNAs are less characterised classes of non-coding RNAs with emerging and unexpected roles in cancer initiation, progression and metastasis.

Although many studies have characterised the protein coding genes and miRNAs regulated by hypoxia (Camps et al., 2014, Schodel et al., 2013), the molecular mechanisms orchestrating these complex phenomena still remain unclear. For example, many regulated genes (especially downregulated genes) do not have a HIF-binding site in their vicinity, suggesting indirect mechanisms of regulation by HIF. This could partially be explained by changes of other non-coding RNAs, such as lncRNAs and other epigenetic factors such as histone modifications. This drives the hypothesis that non-coding RNAs and epigenetic factors can play an important role in adaption to hypoxic stress. These non-coding RNA responses can reshape our understanding of the classical hypoxia transcriptional landscape and ultimately may identify novel therapeutic targets.

6.1 Complexity of the hypoxia transcriptome

In the current work, I have comprehensively profiled the transcriptome of hypoxic cells using next generation RNA sequencing technologies to enhance the understanding of transcriptional regulation in response to hypoxic stress. Using MCF-7 cells, which are a well-characterised and established *in vitro* breast cancer model, coupled with a bespoke strand-specific total-RNA-seq method, analyses, I have identified distinct patterns of regulation of the different RNA classes in hypoxia. In addition to identifying a number of well-known hypoxia regulated genes that have been reported previously, this study has extended knowledge of hypoxic regulation of the

transcriptome to other RNA classes such as piwiRNA, tRNAs, snRNAs, non poly A RNAs, lncRNAs and previously non-annotated transcripts comprising both non-annotated intragenic transcripts and antisense transcripts **(Figure 6.1)**. Since each class of RNA can affect transcriptional and post-transcriptional regulation through different mechanisms this suggests a complex hierarchy of gene regulation in response to hypoxic stress.

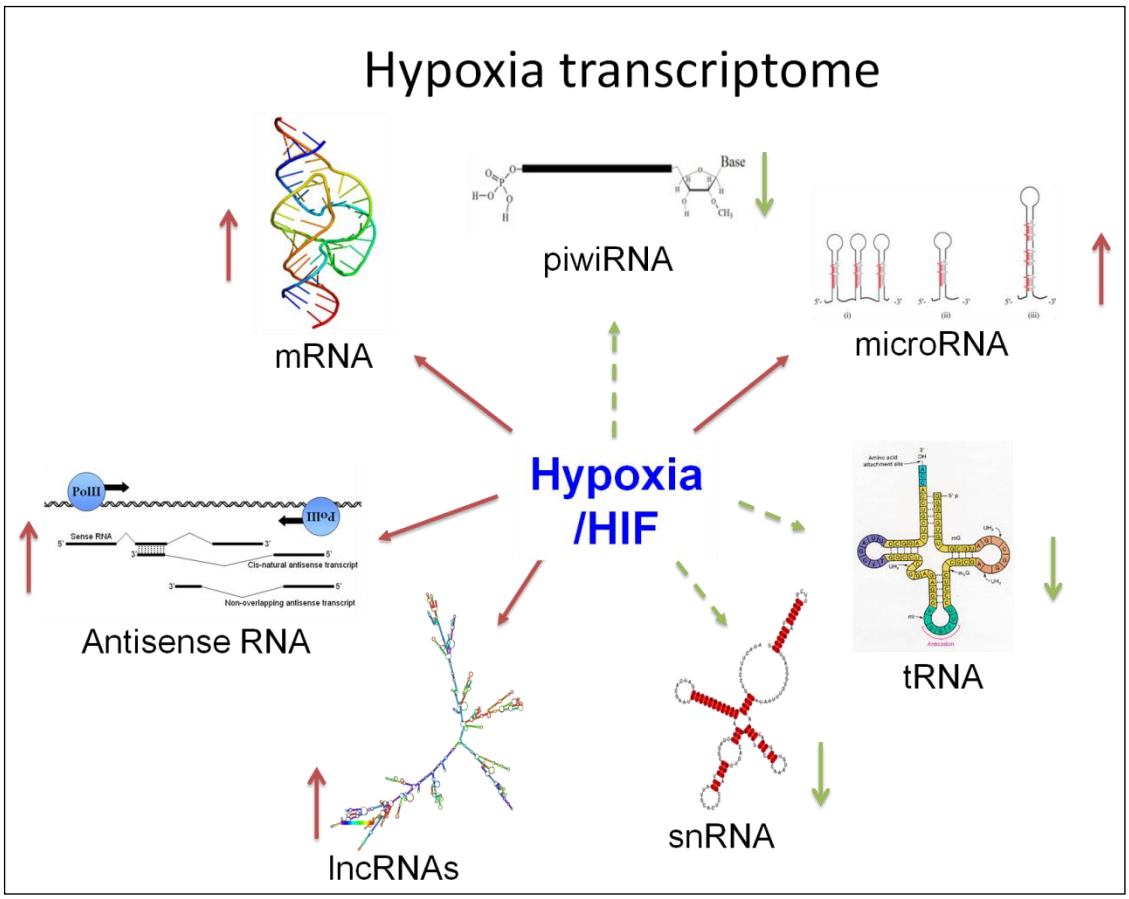


Figure 6.1 Influence of hypoxia/HIF on the cell transcriptome. All the known classes of RNA are regulated by hypoxia. Green arrow indicates downregulation (↓) and red arrow indicates upregulation (↑).

Antisense transcripts

There is increasing evidence to indicate that antisense transcripts can interact with the sense transcripts at the same locus to either suppress or induce expression of each other. Indeed, it is estimated that as many as one-third of protein coding genes are regulated by an antisense transcript (Wahlestedt, 2013). This regulation may occur by one of two mechanisms. First, antisense transcripts can bind complementary sequences on the sense transcript to form an RNA hybrid that promotes its own degradation in a manner analogous to mechanisms of siRNA degradation. Second, antisense transcripts may promote epigenetic modulations that regulate transcription of the sense transcript (e.g. chromatin structure is altered through recruitment of DNA methyltransferases to the locus by the antisense transcript which promotes DNA methylation, or by recruiting histone-modifying enzymes)(Wahlestedt, 2013, Faghihi and Wahlestedt, 2009). By using strand-specific RNA-seq methods, the current study identified a large number of hypoxia regulated antisense transcripts leading to both co- and counter-regulation and providing a new mechanism for modulating gene expression in hypoxia, which will require further locus-specific experimental and functional analysis to understand fully.

Long non-coding RNAs

In addition, there is growing evidence that lncRNAs also have functional regulatory roles. This class of RNA is dysregulated in cancer but was not previously known to be widely regulated by hypoxia. However, in the current work, large numbers of lncRNAs were upregulated in response to hypoxia and more specifically to HIF. It has been reported that the majority of lncRNAs act by physical interaction with epigenetic modifiers or other regulatory protein complexes to control expression at specific loci either *in cis* or *in trans* (Tsai et al., 2010). By doing so, lncRNAs can dynamically regulate large-scale transcriptional programs required for adaption and survival through environmental stress. In addition to regulating transcription, lncRNAs can control post-transcriptional RNA processing, for example by regulating alternative splicing (Anko and Neugebauer, 2010).

One example of an lncRNA playing a role in regulating RNA splicing from the current project is MALAT1. This is a nuclear-retained lncRNA that has been directly linked to cancer metastasis (Han et al., 2013, Schmidt et al., 2011). In addition, MALAT1 is also reported to contribute to cancer metastasis through several mechanisms, for example, by controlling transcription of pro-migratory genes (Tano et al., 2010) and epithelial-mesenchymal transition (EMT) mediating genes (Ying et al., 2012). I found that MALAT1 is induced when breast cancer cells are exposed to hypoxia and that this is dependent on HIF. MALAT1 is present in nuclear speckles which are involved in phosphorylation of the pre-mRNA splicing factors thereby controlling their cellular localization and activity (Tripathi et al., 2010). Since MALAT1 is induced in hypoxia and regulates RNA splicing and tumour metastasis, which are hallmarks of cancer, this provides new insights into how lncRNAs may influence gene expression in hypoxia.

The importance of specific lncRNAs in promoting cancer progression in hypoxia has been described very recently. lncRNA low expression in tumour (lncRNA-LET) physically interacts with NF90 and promotes its proteasomal degradation (Yang et al., 2013a). NF90 is an RNA-binding protein that stabilizes HIF1-α transcripts. LncRNA-LET reduces NF90 levels destabilizing HIF1-α mRNA, thus acting to decrease tumour metastasis under hypoxic conditions (Vumbaca et al., 2008, Yang et al., 2013a).

Another lncRNA, known as "Linc-RoR," is an extracellular RNA that is released by malignant liver cancer cells (Takahashi et al., 2014). Takahashi et al. demonstrated that extracellular levels of Linc-RoR, are induced by hypoxia and reported that Linc-RoR knockdown decreases expression of HIF-1α and PDK1 in hypoxia. In addition, they observed expression of miR-145 (target of HIF-1α) was paradoxically increased upon Linc-RoR knockdown (Takahashi et al., 2014). They proposed that linc-RoR may act as a miRNA sponge to limit miRNA such as miR-145 that subsequent regulate expression of key hypoxia response mRNAs such as HIF-1α expression (Takahashi et al., 2014). Although I was unable to comment on the hypoxic

regulation of Linc-RoR, as it is not expressed in MCF-7 cells, their findings suggest that lncRNAs may behave differently in different cancers in response to hypoxia.

lincRNA-p21 has also been also reported as a hypoxia-responsive and HIF-1α dependent lncRNA (Yang et al., 2014b). Yang et al. demonstrated that lincRNA-p21 is essential for modulating glycolysis in hypoxia. They found that lincRNA-p21 binds both HIF-1α and VHL and that thus disrupts the VHL-HIF-1α interaction (Yang et al., 2014b). Disruption of HIF-VHL binding blocks VHL-mediated HIF-1α ubiquitination and causes HIF-1α accumulation. These authors proposed that there is a positive feedback loop between HIF-1α and lincRNA-p21 that increases glycolysis under hypoxia (Yang et al., 2014b).

The studies discussed above are very recent examples of how lncRNAs can influence and modulate hypoxia responses (e.g. promote tumour metastasis, regulate expression of HIF-1α and promote the Warburg effect). My thesis has analysed lncRNA regulation by hypoxia at the genome- wide level and has determined that hypoxic regulation of lncRNAs by HIF is widespread, including the regulation of many lncRNAs with as yet undetermined function. To what extent and by what mechanisms each of these hypoxia-regulated lncRNAs contributes to hypoxic regulation of the transcriptome and the role that these play in cancer biology remains to be determined.

snRNAs

snRNAs are generally considered as "housekeeping" genes in different cell types and conditions, and have been used as control genes for normalising miRNA expression in cancer (Chang et al., 2010). However, I found that the majority of snRNAs are altered and downregulated in hypoxia. This finding is consistent, with those of Gee et al. who found that snRNA expression is as variable as miRNA expression in tumours. In addition, they reported that SNORD48 and SNORD44 expression were significantly decreased in breast cancer tumours, and that their downregulation was associated with a better prognosis in both breast and non-small cell lung

cancer tumours (Gee et al., 2011). Similarly, Esteller and co-workers identified a number of snoRNAs which were "silenced" in a panel of cancer cell lines and tumour samples due to DNA methylation (Ferreira et al., 2012). Functionally, many studies have reported that snoRNAs are implicated in the control of splicing of many genes including genes driving metastasis and those regulating oxidative stress (Chu et al., 2012, Kishore et al., 2010). Since epigenetic silencing of tumour suppressor genes and aberrant splicing is a common feature of cancer, these may indicate that at least some snoRNAs may play an important function in tumourigenesis and hypoxia stress.

tRNAs

tRNAs are essential molecules, which decode mRNA codons into amino acids during the synthesis of proteins. Significant downregulation of tRNAs in hypoxia is consistent with previous findings that hypoxia significantly reduces protein synthesis as a potential mechanism of reducing cellular energy consumption (Hochachka and Lutz, 2001). In severe hypoxia (O_2 <0.2), it is estimated that ATP demand for protein synthesis is reduced by about 7 % compared to normoxic cells (Hochachka et al., 1996). The extent to which this global reduction in protein synthesis is driven by changes in tRNA expression remains to be determined.

PiwiRNAs

The current study reports hypoxic regulation of the class of small RNAs, known piwiRNAs, for the first time. PiwiRNAs are involved in epigenetic control and disruption of piwiRNAs can lead to uncontrolled transposition of mobile genetic elements and consequent genomic instability. Recent studies have shown that some piwiRNAs are deregulated in cancer cells (Mei et al., 2013) and play a role in tumour metastasis (Mei et al., 2013). It is also reported that mobile genetic elements such retrotransposons are affected in hypoxia (Sehgal et al., 2007). This adds another layer of complexity to the regulation of the transcriptome by hypoxia, in which mobile genetic elements may be regulated by hypoxia to alter genomic stability.

miRNAs

miRNAs are a class of small RNA that regulate gene expression at the post-transcriptional and translational levels. This is the best-characterised class of small RNAs. They are involved in a wide range of cellular processes in both normal and pathological conditions, such as cell death, development, metabolism, and tumorigenesis (Yates et al., 2013, Ventura and Jacks, 2009). They are also the best-studied class of non-coding RNAs in hypoxia and previous microarray analyses have identified a number of hypoxia-regulated miRNAs. Amongst these studies there is a surprising lack of consistently hypoxia-regulated miRNAs, with the exception of miR-210, which has been reported to be upregulated in all the studies to date (Huang et al., 2010, Shen et al., 2013). These differences between studies could be due to technical variability, including the array platform used, the severity and duration of hypoxia used, or the cellular context. The current study identified hypoxia-regulated miRNAs that were both common to previous studies (e.g. miR-210) and were novel, including some with potential roles in cancer progression including miR-184, miR-1 and miR-100 (Liu et al., 2011, Tivnan et al., 2010, Koberle et al., 2013).

6.2 HIF, RNApol2, H3K4me3 and the hypoxia transcriptome

The integration of ChIP-seq with RNA-seq is a powerful tool to investigate and examine gene expression and to correlate it with genomic features. Here, I have integrated my RNA-seq analyses with pre-existing datasets of HIF, RNA Pol2, H3K4me3 ChIP-seq and DNAse hypersensitivity. HIF is the major transcriptional activator of gene expression in hypoxia. Five hundred HIF-binding sites were annotated genome-wide using high-stringency criteria. The distribution of HIF-binding sites with respect to the different RNA classes was strongly associated with the frequency of each class of transcript in the genome. However, superimposed upon this there was no enrichment of HIF-binding sites at piwiRNA, snRNA or tRNA gene loci, all classes of transcript that showed globally downregulation in hypoxia. Conversely, HIF-binding was observed in the vicinity of both mRNAs and lncRNAs, and specifically amongst the most

upregulated members of each class. These finding suggests that HIF acts as a transcriptional activator (but not repressor) of lncRNA as well as mRNA gene expression.

Previous studies have reported that a large proportion of HIF binding sites are distant from the promoters of any annotated protein coding genes. However, these studies did not take in account different RNA classes. Since a number of HIF binding sites that were previously thought to be intergenic were actually close to the promoter of a non-coding transcript in the current study, the data were reanalysed in the light of my total RNA-seq. Despite the inclusion of non-coding transcripts of all classes as well as previously non-annotated transcripts expressed in MCF-7 cells, a high proportion of HIF-1 and HIF-2 binding sites are still distant from any active promoter, confirming that HIF can act as a long-range enhancer.

After mRNAs, lncRNAs are next largest class of HIF-binding gene loci. Interestingly, lncRNA gene loci demonstrated a bias towards binding HIF-2 over HIF-1 when compared to mRNA gene loci, suggesting that HIF-2 plays a greater role in the hypoxic regulation of lncRNAs. The majority of HIF-binding lncRNAs were up regulated in hypoxia and were down regulated following suppression of HIF-α expression, providing strong evidence for widespread regulation of lncRNAs by HIF.

Histones are a family of proteins that are responsible for organizing DNA structure in nucleus. These proteins are subject to a range of biochemical modifications on their N-terminal tails including methylation, acetylation, and ubiquitination. These modifications dynamically influence the role of these histones in determining chromatin structure and function and may be associated with either activation or repression of gene expression. For example, trimethylation of histone 3 at lysine 4 (H3K4me3) is marker of active chromatin and is found close to the promoters of actively transcribed genes (Black et al., 2012). Furthermore, hypoxia may induce bulk changes in the manifestation of these modifications, either through transcriptional regulation of histone modifying enzymes or through regulation of their enzyme activity (Xia et al., 2009, Zhou et al., 2010, Wellmann et al., 2008).

Correlation of existing RNA Pol2, H3K4me3 and DNAse HSS datasets with my total RNA-seq has shown that hypoxia-upregulated, HIF-binding gene loci have RNApol2 prebound to accessible promoters that manifest H3K4me3 in normoxia. In hypoxia promoter DNA accessibility and RNApol2 binding does not change, but RNApol2 is released to run across the gene body. Concomitantly, histone H3 within the gene body is trimethylated on the K4 residue. These changes occur irrespective of whether the transcript is of the mRNA or lncRNA class. This release of prebound promoter-paused RNApol2 following HIF binding potentially facilitates a more rapid transcriptional response to hypoxia.

6.3 The role of NEAT1, an HIF-2α dependent lncRNA, in hypoxia

The integration of RNA-seq and ChIP-seq data has identified a role for HIF in the transcriptional regulation of lncRNAs as well as mRNAs. To illustrate role of lncRNAs in hypoxia, NEAT1 lncRNA was selected for further analysis of its functional and regulatory role in response to hypoxic stress. NEAT1 is the most upregulated HIF-binding lncRNA in hypoxia, is highly abundant, is predominantly upregulated by HIF-2 and has not been studied before in hypoxia. Hypoxic upregulation of NEAT1 was observed across a wide range of breast cancer cell lines as well as several xenograft models treated with Bevacizumab, suggesting that induction of NEAT1 is a common response to hypoxia. The current study demonstrated, for the first time, that NEAT1 knockdown reduces cell proliferation and survival and increases apoptosis of hypoxic cancer cells, indicating that NEAT1 has oncogenic properties in hypoxia. These findings are consistent with unreported work by Christopher Halford from Keele University, who has demonstrated that NEAT1 knockdown reduces cell viability and increases levels of apoptosis in a Burkitt's lymphoma cell line (Preliminary results). Moreover, Hirose et al. treated Neat1−/− or Neat1+/+ fibroblasts with proteasome inhibitors (MG132 and bortezomib) and showed significantly increased levels of apoptosis in NEAT1−/− fibroblasts compared to Neat1+/+ cells (Hirose et al., 2014), suggesting that NEAT1 attenuates the cell death pathway.

In addition, the current study also demonstrates that hypoxia induces the formation of nuclear paraspeckle bodies in hypoxia. NEAT1 is an important architectural component necessary for the formation of these bodies and paraspeckle formation in hypoxia was dependent on both NEAT1 and HIF-2. Previous studies have reported that paraspeckle bodies control the expression of adenosine-to-inosine (A-to-I) hyperedited RNAs by retaining them in the nucleus. Thus, the formation of paraspeckles in hypoxia, provides a potential new mechanism by which HIF-2 might control the nuclear/cytoplasmic localisation of A-to-I edited transcripts in hypoxia and potentially release them rapidly into the cytoplasm on re-oxygenation **(Figure 6.1).**

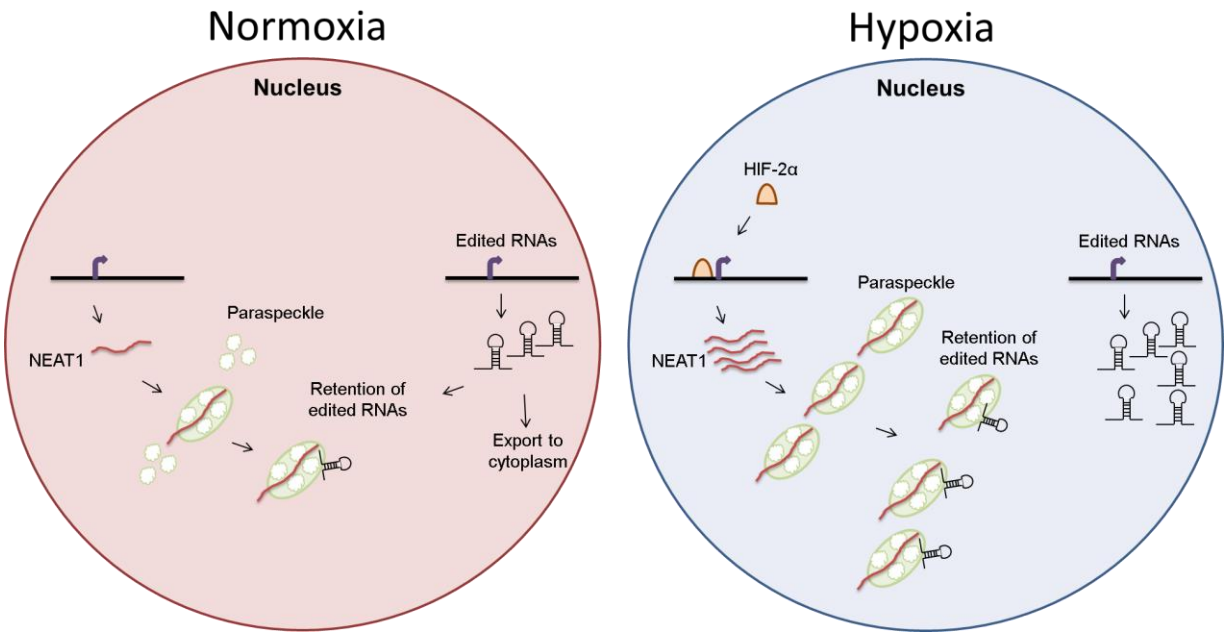


Figure 6.2 A proposed model for NEAT1 function regulating expression of hyperedited genes following hypoxic stress. In normoxia, a fraction of NEAT1 is transcribed. Paraspeckle proteins are diffuse in normoxic condition and forms limited paraspeckle bodies. Few edited RNAs are bound to these bodies while, the rest of edited RNAs are exported to the cytoplasm for translation. In hypoxia, HIF-2α is induced and binds promoter region of NEAT1 which significantly increase it expression and consequently increase formation of paraspeckle bodies. In addition upon hypoxia the editing of RNA is also increased. The majority of the hyperedited molecules binds paraspeckle bodies and is retained in the nucleus in large amounts. This model provides a new mechanism of gene regulation in hypoxia through HIF dependent lncRNA.

Editing of RNA molecules is among several biochemical modifications occurs following transcription. This "editing" can change the RNA sequence by modifying nucleoside from C-to-U as well as by A-to-I deamination (Brennicke et al., 1999). These modifications can post-transcriptionally alter RNA sequence and consequently changes the amino acid sequence of the protein that would be predicted from the DNA sequence, and could have profound effect on protein function (Chen et al., 2013). These substitutions can be studied globally using RNA-seq technologies to determine the effects of hypoxia and whether any other lncRNAs are involved in their regulation.

Studies of individual transcripts have reported differential A-to-I RNA editing of several RNAs in hypoxia, including mediator complex subunit 13 (MED13), signal transducer and activator of transcription 3 (STAT3), and F11 receptor (F11R) (Nevo-Caspi et al., 2011, Ben-Zvi et al., 2013). Specifically, Ben-Zvi et al. reported that both the expression and A-to-I editing of the 3' UTR of F11R increased in hypoxia (Ben-Zvi et al., 2013). F11R is a cell adhesion molecule that plays an important role in cell migration and cell proliferation (Goetsch et al., 2013, Brennan et al., 2013). They demonstrated that hyper-edited mature F11R RNA is retained in the nucleus and is associated with NONO (p54nrb) in hypoxic HeLa cells (Ben-Zvi et al., 2013). Since NONO (p54nrb) is one of the main components of paraspeckles, I hypothesised that NEAT1 may be involved in regulating the localisation of A-to-I edited F11R in hypoxia and this is the basis of ongoing work.

In addition, to study the effects of hypoxia on the nuclear/cytoplasmic localisation of RNAs more globally and of edited transcripts in particular, I have undertaken further RNA-seq experiments on nuclear and cytoplasmic fractions from MCF-7 cells incubated in normoxia and hypoxia. Analysis of this data is ongoing and is therefore not included in thesis. Since inosine mimics guanine, both biologically and in sequencing reactions, it is also possible to identify A-to-I edited transcripts as well as total transcript levels by examining for A-to-G substitutions in the transcriptome that are absent in the genomic DNA. I then aim to determine whether these

behave differently from non-edited transcripts. A-to-I edited transcripts that alter their localisation in hypoxia can then be examined for their dependence upon NEAT1 and on HIF-2.

Several classes of hypoxia-regulated non-coding RNAs (snRNA, piwiRNA and lncRNA) have overlapping functions in controlling and regulating splicing of mRNAs. Alternative splicing serves as a regulatory mechanism. It is estimated that around 95 % of human genes are alternatively spliced (Pan et al., 2008, Wang et al., 2008), which dramatically increases genomic complexity. There is evidence that hypoxia regulates pre-mRNA splicing, however these studies have been largely based on microarray technologies (Moller-Levet et al., 2009, Hang et al., 2009). Further analyses of RNA-seq datasets generated in this thesis are also ongoing in an attempt to identify alternative splicing events in hypoxia. In particular, sequencing of cytoplasmic fractions in hypoxia and normoxia allows the assessment of mature cytoplasmic RNAs; this will provide an ideal substrate for detecting differential splicing events in hypoxia.

6.4 Network of hypoxia transcriptome

Despite being a major cause of resistance to therapies and poor clinical outcome in cancer, hypoxia remains poorly understood at the transcriptional and genomic level. To date pan-genomic analyses of hypoxic responses have largely focussed on mRNAs or miRNAs. This thesis has shown that other classes of RNA are also regulated by hypoxia, some by HIF-transcriptional regulation. These classes of RNA can have diverse effects on both the transcriptional and post-transcriptional regulation of gene expression and/or function and may underlie widespread molecular networks of gene regulation. In support of this hypothesis, there is increasing evidence that different classes of non-coding RNAs are functionally interconnected, and that they coordinates a range of interactions that regulate cellular changes in both physiological and pathological conditions (Jalali et al., 2013). Moreover, genome-wide association studies (GWAS) have reported that the majority of cancer risk loci lie outside of protein-coding genes, potentially affecting non-coding RNAs. Currently, there around 301 single-nucleotide polymorphisms (SNPs) are associated with different types of cancer, only 12 (3.3 %) of SNPs

changes amino-acid sequence, while, 40 % of them located in the introns of protein-coding genes or 44 % in intergenic regions of genome (Gomes et al., 2013). This raises important questions as to the function of these non-coding loci, whether these loci contain non-coding RNAs or regulatory regions and if they are implicated in cancer development.

The deregulation of RNA networks in hypoxia and the observations that non-coding RNAs link with cancer progression, raise the hypothesis that protein coding genes cooperate with non-coding RNAs to control gene expression at different levels of regulation in response to hypoxia **(Figure 6.2).** The rapid development of next-generation sequencing technologies in detecting and quantifying transcripts at high resolution will result in rapid cataloging of functional non-coding RNAs. Understanding the role and regulation of non-coding RNA interactome in hypoxia will significantly enhance our understanding of molecular mechanisms of hypoxia response and will allow dissecting the molecular interactions that can be utilized as therapeutic targets.

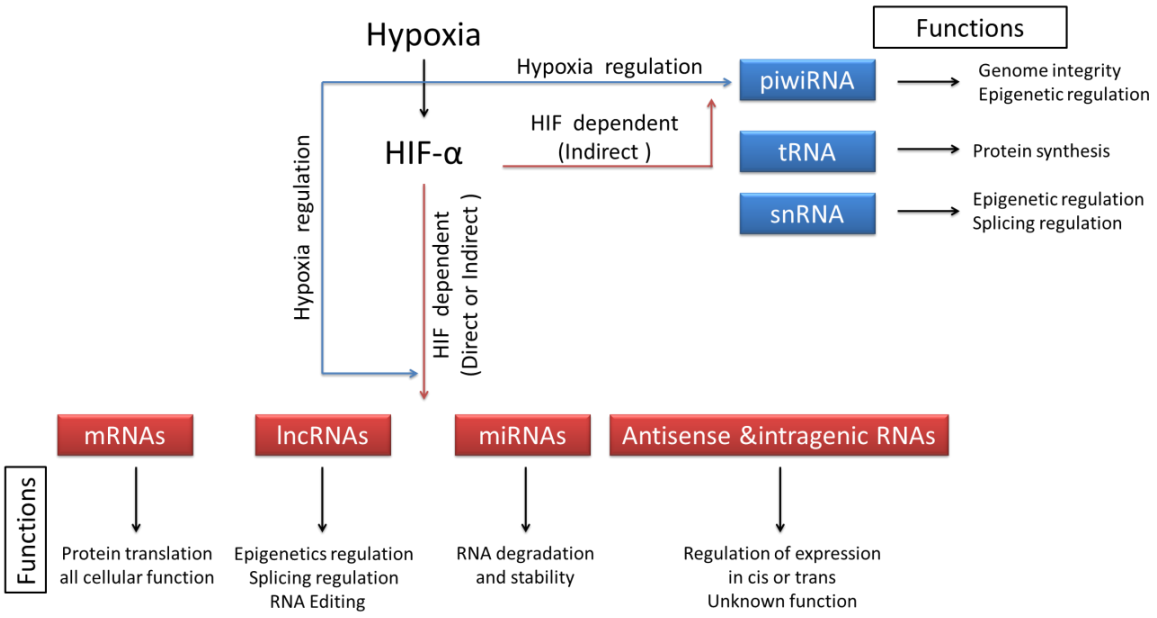


Figure 6.3 Influence of hypoxia/HIF on cellular RNA networks. mRNAs, lncRNAs, miRNAs and antisense transcripts (commonly upregulated) have HIF binding sites, thus they are potentially directly regulated by HIF, in addition they could be regulated by indirect HIF regulation or by other factors in hypoxia. piwiRNA, tRNA and snRNA (globally downregulated) lack HIF binding sites, therefore they are indirectly regulated by HIF or by hypoxia. All these RNA classes have range of functional and regulatory roles in cells.

Overall, this study provides first comprehensive genome-wide catalogue of the hypoxic transcriptome, the HIF dependent transcriptome, HIF binding sites and RNA Pol2 and H3K4me3 distribution. It should help us understand hypoxic transcriptional and gene regulation. In addition, it provides first evidence that lncRNAs are highly regulated in hypoxia via HIF and some of these lncRNAs have potential role in tumour biology and breast cancer development.

Chapter 7: Material and Methods

7.1 Cancer cell lines

Human breast cancer cell lines used in this project included MCF-7, ZR75-1, T47D, BT474, SKBR3, MDA-MB-453, MDA-MB-468, MDA-M-231, BT-20, DU4475, BT-549, these were obtained from the American Type Culture Collection 'ATCC' (Manassas, VA). All the cell lines were received frozen and were cultured upon arrival.

7.2 Cell culture

For optimal cell growth, all the cell lines were cultured in a specific media as suggested by the ATCC (list of media used for each cell line below in the table). All the media was supplemented with 2 mM L-glutamine, 10 % Fetal Bovine Serum (FBS) and 1 % penicillin-streptomycin solution (Sigma-Aldrich). For normoxic conditions: cells were cultured in a standard cell culture incubator at 37 °C. For hypoxic conditions: cells were inducted at 1 % O_2 in an *In vivo2 Hypoxia Work Station* (Ruskin Technologies, Kent, UK).

For sub-culturing and passages, cells were detached by adding 0.25 % trypsin-EDTA solution (PAA Laboratories) followed by incubating for 5 minutes at 37 °C. Once the cells were detached, fresh complete media was added and reseeded at 1:10 in fresh media. For cell storage, cells were trypsinized and resuspended in complete media, then spun down at 1,000 xg for 5 minutes. The pellet was resuspended in freezing medium containing 90 % DMEM and 10 % dimethyl sulphoxide "DMSO". Cell aliquots of 1.8 ml was added in Cryovials and transferred to the NALGENE cryo freezing container for storage at -80 °C for 24 hours and then in liquid nitrogen for long term storage.

MCF-7 is a near triploid cell line with high rearrangements. Translocations involve all chromosomes with the exception of chromosomes 4. Deletions were detected on chromosomes 3 and 13. Several markers are very complex due to multiple rearrangements. The DNA fingerprinting of the MCF-7 cells was recently performed as part of the National Cancer

Institute's NCI- cell line panel (Lorenzi et al., 2009). All the cell lines used in this project were in early passages (less than 5).

Table 7.1 List of breast cancer cell lines and medium used for culture:

breast cancer type	Cell line	Medium
ER+	MCF-7	DMEM
	ZR75-1	RPMI-1640
	T47D	RPMI-1640
	BT474	RPMI-1640
ER-	MDA-MB-453	Leibovitz's L-15
	SKBR3	DMEM
TNBC	MDA-MB-468	DMEM
	MDA-MB-231	DMEM
	BT-20	DMEM
	DU4475	RPMI-1640
	BT-549	RPMI-1640

DMEM: Dulbecco's modified Eagle's Medium; RPMI-1640: Roswell Park Memorial Institute-1640; ER: Estrogen receptors; TNBC: Triple-negative breast cancer. All the mediums were purchased from Sigma-Aldrich.

7.3 Mycoplasma testing

All the cell lines were frequently screened for the mycoplasma contamination using the MycoAlert™ mycoplasma detection kit (lonza, US) as manufacturer's instructions. Briefly,

1- All the reagents were brought to room temperature before use.
2- 2 ml of cell culture was transferred into a centrifuge tube and cells were pelleted at 200 x g for 5 minutes.
3- 100 µl of cleared supernatant was added into a luminometer plate.
4- 100 µl of MycoAlert™ reagent was then added to each sample and incubated at room temperature for 5 minutes.
5- The plate the placed in luminometer and initiated for reading A
6- Then 100 µl of MycoAlert™ substrate was added to each sample and incubated at room temperature for 10 minutes.

7- The plate is placed again in luminometer and initiated for reading B

8- The Calculate ratio = Reading B/Reading A.

7.4 Nucleic acid isolation

7.4.1 Genomic DNA isolation

The DNeasy Blood & Tissue Kit (Qiagen) was used to isolate genomic DNA from the cell lines.

1- Cells were grown in 100 mm dishes to about 80-90 % confluency.

2- After trypsinsation, cells were centrifuged for 5 minutes at 300 xg.

3- Cells were resuspend in 200 µl PBS, 20 µl proteinase K was then added.

4- 200 µl Buffer AL was mixed thoroughly by vortexing.

5- Next 200 µl ethanol (100 %) was added on the mixture and mixed thoroughly by vortexing.

6- The mixture was then transferred into a DNeasy Mini spin column placed in a 2 ml collection tube. The column was centrifuged at ≥ 6,000 xg for 1 minute. The flow-through was discarded.

7- 500 µl of Buffer AW1 was added on column and centrifuged for 1 minute at ≥ 6,000 xg. The flow-through was discarded.

8- The spin column was placed in a new 2 ml collection tube. 500 µl of Buffer AW2 was added on column and centrifuged for 3 minutes at 20,000 xg. The flow-through was discarded.

9- The spin column was transferred to a fresh 2 ml microcentrifuge.

10- 200 µl Buffer AE was added to the center of the spin column membrane. Incubated for 1 minute at room temperature (25 °C) then centrifuged for 1 minute at ≥ 6,000 xg to elute the DNA.

11- Samples were kept at -20 °C.

7.4.2 Total RNA isolation

The mirVana miRNA Isolation Kit was used to isolate total RNA from the cell lines, according to the manufacturer's instructions (Ambion, Inc.). Briefly,

1- Cells were grown in 100 mm dishes to about 80-90 % confluency.
2- Cells were washed with cold 1X phosphate buffered saline (PBS).
3- PBS was removed and kept on ice.
4- 600 µl of Lysis/Binding Solution was added on the cells.
5- Cells were scrubbed and collected in an eppendorf tube then vortexed vigorously to lyse the cells and to obtain a homogenous lysate.
6- 1/10 volume of miRNA Homogenate Additive was added to the cell lysate and mixed well by vortexing and then incubated on ice for 10 minutes.
7- Equal volume of Acid-Phenol: Chloroform was added and mixed for 60 seconds and then centrifuged for 5 minutes at 10,000 xg at room temperature.
8- The upper aqueous phase was carefully removed without disturbing the lower phase, and transferred into a fresh tube.
9- 1.25 volumes of 100 % ethanol was added to the aqueous phase, the lysate/ethanol mixture was transferred onto the filter cartridge provided in the kit, followed by centrifuging it for ~15 seconds at 10,000 xg to pass the mixture through the filter.
10- The flow-through was discarded and 700 µl of miRNA Wash Solution 1 was added on the filter cartridge followed by a centrifuge for ~10 seconds.
11- The flow-through was discarded and the filter cartridge was washed twice with 500 µl Wash Solution 2/3.
12- After discarding the flow-through from the last wash the filter cartridge was spun for 1 minute to remove residual fluid from the filter.

13- The filter cartridge was then transferred into a fresh collection tube (provided with the kit). 100 µl of pre-heated (at 95 °C) nuclease-free water was added to the center of the filter.

14- The tube was centrifuged for 30 seconds at maximum speed to recover the RNA.

15- Collected elute (which contains the RNA) was stored at –80 °C.

7.4.3 Quantification of nucleic acid concentration

Concentrations of DNA and RNA samples were determined using a NanoDrop® ND-1000 spectrophotometer (Thermo Scientific, U.K.). 1 µl of nucleic acid was loaded on the NanoDrop to measure the absorbance. DNA and RNA with absorbance ratios 260/280 > 1.8 were considered highly purified.

7.4.4 Nucleic acid quality control

The integrity of isolated DNA and RNA were verified on an Agilent Bioanalyzer 2100 (Agilent Technologies, Palo Alto, CA). For RNA, the Agilent RNA 6000 Pico kit and for DNA, Agilent DNA 1000 kit was used (Agilent). High quality total RNAs were used in the study with a RNA Integrity Number (RIN) over 9.

7.4.4.1 Agilent DNA 1000 Kit

Preparing the Gel-Dye Mix:

1- Both DNA dye concentrate and DNA gel matrix were first equilibrated to room temperature for 30 minutes.

2- The DNA dye concentrate was vortexed, 25 µl of the dye was added to a DNA gel matrix vial. The solution was mixed well and spun down.

3- The solution was then transferred to a spin filter and centrifuged at 2,240 g ± 20 % for 15 minutes and stored at 4 °C.

Loading the Gel-Dye Mix:

1- The gel-dye mix was first equilibrated to room temperature for 30 minutes before use.
2- A new DNA chip was placed on the chip priming station.
3- 9 µl of gel-dye mix was loaded in the well-marked (G).
4- Plunger in chip priming station was pressed until it is held by the clip.
5- Exactly after 60 seconds, the plunger was released.
6- Slowly the plunger was pulled back to 1 ml position.
7- 5 µl of marker was loaded in all the 12 sample wells and ladder well.

Loading the Ladder and the Sample:

1- 1 µl of DNA ladder was loaded in the well-marked ‡.
2- In each of the 12 sample wells, 1 µl of sample or 1 µl of de-ionized water (unused wells) was loaded.
3- The chip was horizontally placed in the adapter and vortexed for 1 minute at 2400 xg.
4- The chip was run in the Agilent 2100 bioanalyzer within 5 minutes.

7.4.4.2 Agilent RNA 6000 Pico

Preparing the Gel:

1- First, 550 µl of RNA 6000 Pico gel matrix was loaded into a spin filter and centrifuged at 1,500 g ± 20 % for 10 minutes at room temperature.
2- Aliquots of 65 µl filtered gel split into 0.5 ml RNase-free microfuge tubes.

Preparing the Gel-Dye Mix:

1- The RNA 6000 Pico dye concentrate was equilibrated to room temperature for 30 minutes, vortexed for 10 seconds and spun down.
2- 1 µl of RNA 6000 Pico dye concentrate was added into a 65 µl aliquot of filtered gel. The solution was mixed well and centrifuged at 13,000 g for 10 minutes at room temperature.

Loading the Gel-Dye Mix:

1- A new RNA 6000 Pico chip was placed on the chip priming station.
2- 9 µl of gel-dye mix was then loaded in the well marked G.
3- The chip was closed in the priming station and the plunger pressed until held by the clip.
4- Exactly after 30 seconds the plunger released the clip.
5- Then, 9 µl of gel-dye mix was loaded in the wells marked G.

Loading the RNA 6000 Pico Conditioning Solution and Marker:

1- 9.0 µl of the RNA 6000 Pico Conditioning Solution was loaded in the well marked CS.
2- 5 µl of RNA 6000 Pico marker was loaded in all the 11 sample wells and in the well-marked ‡.

Loading the Diluted Ladder and Samples:

1- 1 µl of the heat denatured and aliquoted ladder loaded in the well-marked ‡.
2- 1 µl of sample was loaded in each of the 11 sample wells, and 1 µl of RNA 6000 Pico Marker in each unused sample well.
3- The chip was placed in the adapter and vortexed for 1 minute at 2400 xg.
4- The chip was then run in the Agilent 2100 bioanalyzer for 5 minutes.

7.5 Cellular fractionation

Nuclear, cytoplasmic RNA and protein were isolated from hypoxic MCF-7 cells using PARIS™ Protein and RNA Isolation kit (Ambion, Inc.), according to the manufacturer's instructions. Briefly:

1- First, the solutions provided in the kit were prepared; 415 µl 2-mercaptoethanol was added into the 2X Lysis/Binding Solution and mixed well, 64 ml 100 % ethanol was added into Wash Solution 2/3 and mixed well.
2- The culture medium was aspirated and discarded, and then rinsed with PBS.
3- The culture plate was then placed on ice.
4- 500 µl of ice-cold Cell Disruption Buffer was added to the cells and vortexed vigorously to completely lyse the cells and to obtain a homogenous lysate.
5- Lysate incubated on ice for 10 minutes.
6- Sample is centrifuged for 5 minutes at 4 °C and 500 xg.
7- The cytoplasmic fraction was carefully aspirated away from the nuclear pellet.
8- The nuclear pellet was washed again on ice-cold Cell Fractionation Buffer and repelleted the nuclei for 1 minute at 4 °C and 500 xg .The supernatant was carefully discarded without disturbing the pellet.
9- 500 µl of ice-cold Cell Disruption Buffer was added to the nuclear pellet, mixed vigorously to lyse the nuclei, and continued disruption until the lysate was homogenous. The sample was kept on ice throughout.
10- The lysates were split in two portions: one for RNA isolation and the other one for protein analysis.

For proteins

Nuclear fraction: the nuclear lysate was incubated on ice for 10 minutes to ensure complete cell disruption.

Cytoplasmic fraction: the cytoplasmic fraction protein samples did not require incubation on ice.

For RNA isolation from both nuclear and cytoplasmic fractions, these steps were followed separately for both the fractions:

1- Equal volume of 2X Lysis/Binding Solution was added on the lysate and mixed well.
2- Equal volume of the intimal lysate of 100 % ethanol was added into the mixture.
3- Sample mixture was applied to a filter cartridge assembled in collection tube (supplied),then centrifuged for 1 minute at 15,000 xg. The flow-through was discarded.
4- 700 µl of Wash Solution 1 was applied to the filter cartridge. Centrifuged for 1 minute at 15,000 xg. The flow-through was discarded.
5- 500 µl of Wash Solution 2/3 was applied to the filter cartridge. Centrifuged for 1 minute at 15,000 xg. The flow-through was discarded.
6- Last wash was repeated twice and centrifuged for 30 seconds to remove the last traces of Wash Solution 2/3.
7- 100 µl of pre-heated nuclease-free water at 95 °C was applied to the center of the filter.
8- RNA was recovered by centrifugation for 30 seconds at 15,000 xg.
9- RNA samples were kept at –80 °C.

7.6 DNAse treatment

Isolated total RNA was treated with DNase I digestion (TURBO DNA-free™, Ambion).

1- Following reagents were added to a microcentrifuge tube having a total final volume of 50 µl:

Reagent	**Amount per reaction**
Buffer	5 µl
rDNAse I	3 µl
RNA	10 µg
Water	Up to 50 µl
Total volume	**50 µl**

1) Sample was incubated at 37 ˚C for 30 minutes.

2) Then 10 µl of DNAse Inhibitor reagent was added in the sample.
3) Sample was incubated at room temperature for 5 minutes and mixed occasionally.
4) After centrifuging at 10,000 xg for 1.5 minutes the supernatant was transferred (which contains the RNA) to a fresh tube.

7.7Ribosomal RNAs depletion

Ribosomal RNAs were depleted using RiboMinus™ Locked Nucleic Acid probes (Invitrogen). This step allows depletion of abundant ribosomal RNA fraction for total RNA-seq. The kit contains four main steps for rRNA depletion including probe hybridization, beads preparation, rRNA removal and concentrating Ribo depleted RNA.

7.7.1 Probe hybridization

1) Maximum volume of 20 µl of DNAse was treated, total RNA was added into 1.5 ml microcentrifuge tube.
2) Following reagents were added in the tube containing RNA:
    - 10 µl of RiboMinus Probe
    - 300 µl of Hybridization buffer
3) Incubated at 75 ˚C for 5 minutes to denature RNA.
4) The tube was immediately transferred into a water bath at 37 ˚C in order to allow the sample to slowly cool to that temperature over a period of 30 minutes. This is necessary to promote sequence-specific hybridization. During these 30 minutes steps 5 to 12 were performed.

7.7.2 Beads preparation

5) Ribominus magnetic beads were resuspend with thorough vortexing.

6) 750 µl of the bead suspension was transferred on a magnetic rack for 1 minute. The beads settled to the tube side that faces the magnet. Gently the supernatant was aspirated and discarded.
7) 750 µl of sterile RNAse free water was added to the beads and beads were resuspended slowly by vortexing.
8) The tube was placed on a magnetic rack for 1 minute. The supernatant was aspirated and discarded.
9) Steps 7 and 8 were repeated once more.
10) The beads were resuspended in 750 µl hybridization buffer and 250 µl of beads were transferred to a new tube for use at a later step.
11) The tube with 500 µl beads was placed on a magnetic rack for 1 minute. The supernatant was aspirated and discarded.
12) Beads were resuspended in 200 µl hybridization buffer and kept at 37 ˚C until use.

7.7.3 Removing rRNA

13) After 30 minutes at 37 ˚C, the sample from step 4 was briefly centrifuged to collect the sample to the bottom of the tube.
14) The sample (~330 µl) was transferred to the prepared Ribominus magnetic beads from step 12. Mixed well by pipetting up and down or low speed vortexing.
15) The tube was then incubated at 37 ˚C for 15 minutes. During incubation, the tube was gently mixed. The tube was briefly centrifuged to collect the sample to the bottom.
16) The tube was placed on a magnetic rack for 1 minute to pellet the rRNA-probe complex. The supernatant was transferred to a new tube without disturbing the beads and kept for step 18. Note: Do not discard the supernatant since it contains RiboMinus RNA (depleted on rRNA).
17) Tube containing 250 µl beads from step 10 was placed on a magnetic rack for 1 minute. The supernatant was aspirated and discarded.

18) 530 µl supernatant containing RiboMinus RNA from step 16 was added into the tube from step 17. Mixed well by pipetting up and down or low speed vortexing.
19) The tube was incubated at 37 ˚C for 15 minutes. During incubation, the tube was gently mixed occasionally. The tube then was briefly centrifuged to collect the sample to the bottom.
20) The supernatant (~530 µl) containing RiboMinus RNA was transferred into a new tube.
21) The RiboMinus RNA was then concentrated using ethanol precipitation to ensure the recovery of smaller RNA (<200 nt). Other methods of purification can be used if small RNA recovery is not required.

7.7.4 Concentrating Ribo depleted RNA

22) The RiboMinus RNA sample was transferred into a 2 ml microcentrifuge tube.
23) Following reagents were added to RiboMinus RNA:

Reagent	**Amount per reaction**
Glycogen (20 µg/µl)	1 µl
3 M sodium acetate	53 µl (1/10th of sample volume)
100 % Ethanol	1325 µl (2.5X sample volumes)
TOTAL	**1379 µl**

24) The solution was mixed well and incubated at -80 ˚C overnight.
25) The tube was centrifuged for 15 minutes at > 12,000 xg at 4 ˚C. The supernatant carefully discarded without disturbing the pellets.
26) 500 µl of 70 % cold ethanol was then added to the tube.
27) Centrifuging it again for another 5 minutes at > 12,000 xg at 4 ˚C. The supernatant carefully discarded without disturbing the pellets.
28) Steps 26 and 27 were repeated once more.
29) The pellet was air dried for ~5 minutes and resuspended in 20 µl of RNAse free water.
30) The sample was stored at -80 ˚C.

7.8 Selection of poly adenylated RNA

To isolate polyA RNA from total RNA fraction, Dynabeads® mRNA Purification Kit (Ambion) was used. The kit contains two steps: Bead preparation and Isolation of poly adenylated RNA.

1) 1 µg of DNAse I treated RNA was completed to final volume of 100 µl by using nuclease-free water.
2) The sample was the heated at 65 ˚C for 5 minutes in order to disrupt secondary structures and then immediately placed on ice.

7.8.1 Beads preparation

3) Dynabeads magnetic beads were resuspended with thorough vortexing.
4) 100 µl of well resuspended Dynabeads were transferred to a microcentrifuge tube. The tube was placed on the magnet until all Dynabeads were migrated to the tube wall. The supernatant was discarded.
5) The tube was removed from the magnet and 100 µl of Binding Buffer was added to calibrate the beads. The tube was then placed back on the magnet until all Dynabeads had migrated to the tube wall. The supernatant was discarded and the tube was removed from the magnet.
6) 100 µl Binding Buffer was added to the Dynabeads and mixed gently.

7.8.2 Isolation of polyadenylated RNA

7) The DNAse I treated RNA from step 1 (100 µl of volume) was added to the Dynabeads/Binding Buffer suspension from step 6. Solution was mixed thoroughly and rotated for 5 minutes at room temperature to allow mRNA to anneal to the oligo (dT) probes on the beads.

8) The tube was placed on the magnet until the solution became clear. The supernatant was removed without disturbing the beads.
9) The tube was removed from the magnet and 200 µl of Washing Buffer B was added.
10) The tube was placed on the magnet until the solution became clear. The supernatant was removed without disturbing the beads.
11) Steps 9 and 10 were repeated once more.
12) To elute the polyadenylated RNA from the beads, 50 µl of RNAse free water was added to the beads and mixed thoroughly. Then heated at 80 ˚C for 2 minutes and placed immediately on the magnet.
13) 50 µl of Binding Buffer was placed into a new tube. Once the bead solution in the magnet was clear, the supernatant (which contains the polyadenylated RNA) was transferred into this tube containing 50 µl of Binding Buffer. Then 200 µl of Washing Buffer B was added to the beads.
14) The sample containing the polyadenylated RNA was heated at 65 ˚C for 5 minutes and placed immediately on ice.
15) The beads from step 13 ware placed onto the magnet and the supernatant was removed once the solution became clear. The beads were washed again using 200 µl of Washing Buffer B and the supernatant was removed once the solution was clear.
16) The beads 100 µl of the sample containing polyadenylated RNA from step 14 was added. The tube was rotated at room temperature for 5 minutes. The tube was then placed on the magnet and the supernatant was removed once the solution became clear.
17) Washed the beads with 200 µl of Washing Buffer. The tube was placed on the magnet and the supernatant was removed when the solution became clear. This step was repeated one more time.
18) 20 µl of RNAse free water was added to the beads and heated at 80 ˚C for 2 minutes to elute the polyadenylated RNA from the beads.

19) The tube was immediately placed on the magnet stand and then the supernatant (which contains the polyadenylated RNA) was transferred to a new tube.
20) The sample was stored at -80 ˚C.

7.9 Next generation RNA sequencing

7.9.1 Ribosomally depleted directional RNA-seq (Total- RNA-seq)

Total RNA was isolated from MCF-7 cells grown in normoxia (21 % O_2) or in hypoxia (1 % O_2) for 24 hours. After DNAse treatment (Ambion, Inc), the rRNA fraction was depleted by hybridization with Locked Nucleic Acid (LNA) probes (RiboMinus™, Invitrogen). rRNA depletion was assessed using the Agilent RNA 6000 PicoKit (Agilent Technologies). The cDNA library was created using in house modified Illumina directional RNA-seq protocol as following:

7.9.1.1 Fragmentation reaction

1- 100 ng of RiboMinus RNA was placed in a microcentrifuging tube, the volume was topped up to 16 µl with RNAse free water.
2- 4 µl of fragmentation buffer was added.
3- The tube was incubated in PCR thermal cycler at 94 °C for exactly 8 minutes then immediately transferred onto ice and 2 µl of 0.5 M EDTA solution was added.

7.9.1.2 Purification of fragmented ribominus RNA using Qiagen RNeasy kit (Qiagen)

1- The sample was transferred into a new 1.5 ml microcentrifuge tube and the volume was adjusted to 100 µl using RNAse free water. 350 µl of RLT buffer was added and mixed well.
2- 675 µl of 100 % ethanol was added and mixed well by inversion.

3- 700 µl of the sample was transferred to an RNeasy MinElute spin column and centrifuged at 11,000 xg for 1 minute. The flow-through was discarded.

4- The column was washed with 500 µl of RPE buffer and then with 750 µl of 80 % ethanol.

5- After washing, the column was placed in a fresh 2 ml collection tube with the lid open and spun at 11,000 xg for 5 minutes.

6- The column was placed into a new 1.5 ml microcentrifuge tube. 18 µl of RNAse free water was added into the centre of the filter (without touching it) and spun at 11,000 xg for 1 minute to elute. The recovered volume was approximately 17 µl.

7- 1 µl of elute (containing fragmented RiboMinus /polyadenylated RNA) was kept in a separate tube for bioanalyser (stored at -80 ˚C) and the remaining amount was transferred to a 200 µl PCR tube.

7.9.1.3 Phosphatase treatment of fragmented ribominus RNA

1- The following reagents were added in the 200 µl PCR tube containing 16 µl of Ribominus RNA:

Reagent	**Amount for 1 reaction**
10X phosphatise buffer	2 µl
Antarctic phosphatase	1 µl
RNAse OUT	1 µl
TOTAL	**4 µl**

The final volume of the reaction should be 20 µl.

2- The sample was incubated in a thermal cycler with the following protocol:

37 ˚C for 30 minutes

65 ˚C for 5 minutes

4 ˚C indefinite hold

7.9.1.4 T4 Polynucleotide Kinase (PNK) treatment of fragmented ribominus RNA

1- The Following components were added to the 200 µl PCR tube from the previous step:

Reagent	**Amount for 1 reaction**
RNAse free water	17 µl
10X phosphatise buffer	5 µl
ATP (10mM)	5 µl
RNAseOUT	1 µl
PNK enzyme	2 µl
TOTAL	**30 µl**

The final volume of the reaction should be 50 µl.

2- The sample was incubated in a thermal cycler at 37 ˚C for 60 minutes then held at 4 ˚C.

7.9.1.5 Purification of PNK treated RNA using Qiagen RNeasy kit (Qiagen)

1- The sample obtained from the previous step was transferred into a new 1.5 ml microcentrifuge tube. The volume adjusted to 100 µl using RNAse free water.

2- 350 µl of RLT buffer added and mixed well.

3- 675 µl of 100 % ethanol was added and mixed well by inversion.

4- 700 µl of sample was loaded in an RNeasy MinElute spin column and centrifuged at 11,000 xg for 1 minute. The flow-through was discarded.

5- The column was washed with 500 µl of RPE buffer and then with 750 µl of 80 % ethanol.

6- The column was then placed in a 2 ml collection tube with the lid open and spun at 11,000 xg for 5 minutes.

7- The column was transferred into a new 1.5 ml microcentrifuge tube. 15 µl of RNAse free water was added into the center of the column filter (without touching it) and spun at 13,000 xg for 1 minute to elute. The recovered volume should be around 14 µl.

7.9.1.6 Ligation of 3' and 5' adapters

Illumina SRA 5' adaptor and the 10Xv1.5 sRNA 3' adaptor were used. The second adaptor was diluted by mixing 1 µl with 9 µl of RNAse free water

1- The following were mixed in a 200 µl PCR tube:

   RNA sample obtained in previous section (14 µl)

   Diluted v1.5 sRNA 3' adaptor (2 µl)

   Incubated at 70 ˚C for 2 minutes and placed immediately on ice.

2- Then the following reagents were added:

Reagent	**Amount for 1 reaction**
10X T4 RNL2 truncated reaction buffer	2.5 µl
100mM MgCl2	2.0 µl
RNAse OUT	1.0 µl
T4 RNA Ligase2, truncated	3.0 µl
TOTAL	**8.5 µl**

3- Mixed well and spun down.

4- Incubated at 22 ˚C for 1 hour in a thermal cycler.

5- With 5 minutes remaining in the above incubation, the required amount of the 5' adaptor from the stock (2 µl per reaction) was determined and heated at 70 ˚C for 2 minutes then transferred onto ice.

6- The following reagents were added to the ligation mixture from the previous step and mixed well:

Reagent	**Amount for 1 reaction**
10mM ATP	3 µl
Heat denatured SRA 5'adaptor	2 µl
TOTAL	**5 µl**

7- Mixed well and spun down prior to adding enzyme: 2 µl of T4 RNA ligase.

8- Incubated at 20 ˚C for 1 hour in a thermal cycler.

7.9.1.7 Reverse transcription and amplification of the adapter-ligated RNA

Illumina supplied SRA RT primer was used for the reverse transcription reaction and the GX1 and GX2 primers for the PCR reaction. The SRA RT primer needs to be diluted by mixing 1 µl of primer with 4 µl of RNAse free water

1- The following reagent were combined in a 200 µl PCR tube:

   5' and 3' ligated RNA from previous section (4 µl)

   Diluted SRA RT primer (1 µl)

2- The mixture heated at 70 ˚C for 2 minutes in a thermal cycler.

3- The tube was placed on ice and spun in order to collect all the contents.

4- The following components were added to the 200 µl PCR tube from the previous step:

Reagent	**Amount for 1 reaction**
5X first strand buffer	2.0 µl
12.5mM dNTPs	0.5 µl
100mM DTT	1.0 µl
RNAse OUT	0.5 µl
Superscript II Reverse Transcriptase	1.0 µl
TOTAL	**5 µl**

5- Heated at 48 ˚C for 3 minutes in a thermal cycler and then at 44 ˚C for 1 hour.

6- For PCR reaction, the following components were added in a 200 µl PCR tube:

Reagent	**Amount for 1 reaction**
5X Phusion HF buffer	10 µl
Primer GX1	1 µl
Primer GX2	1 µl
25mM dNTPs	0.5 µl
Phusion DNA polymerase	0.5 µl
RNAse free water	27 µl
TOTAL	**40 µl**

7- 10 µl of the product from reverse transcription reaction was added in the PCR reaction.

8- For the PCR amplification in the thermal cycler the following protocol was used:

Temperature	**Time**	**Number of cycles**
98 °C	30 seconds	1 cycle
98 °C	10 seconds	15 cycles
60 °C	30 seconds	
72 °C	15 seconds	
72 °C	10 minutes	1 cycle
4 °C		hold

7.9.1.8 Purification of library using AMPure beads

Agencourt AMPure beads (Beckman Coulter, USA) were used for library purification. First the beads were equilibrated to room temperature for at least 30 minutes prior their use. They also need to be well resuspended by vortexing vigorously for few minutes.

1- The following were added to a new 1.5 ml special low binding microcentrifuge tube:

   PCR reaction product (50 µl)

   AMPure bead solution (110 µl)

2- The tube was vortexed well to ensure thorough mixing. Incubated at room temperature for 5 minutes.

3- Then placed on the magnetic rack for 5 minutes, which allowed the complete capture of the beads. After the suspension was completely clear, the supernatant was carefully removed and discarded without disturbing the bead pellets.

4- With the tube on the magnetic rack, the bead pellet with 500 µl of 80 % ethanol was carefully rinsed. The magnetic rack was inverted with the tubes several times (the tubes were not removed from the rack). Next, the supernatant was carefully removed and discarded without disturbing the bead pellet.

5- With the tube on the magnetic rack, second wash was performed with 500 µl of 80 % ethanol as explained in the previous step. The supernatant was carefully removed and discarded without disturbing the bead pellet.

6- The remaining ethanol was carefully spun down using a mini-centrifuge. The tube was placed in the magnetic rack to capture the beads. Remaining ethanol was removed with a pipette, without disturbing the bead pellet.

7- The tube lid was opened and transferred into a 37 ˚C heat block for 3 minutes or until the beads were completely dry. Small cracks can be observed in the dried bead pellet surface.

8- 30 µl of RNAse free water was added directly to the pellet. Vortexed thoroughly and the beads left for soaking in water for 2 minutes, then vortexed thoroughly again to ensure that the beads were completely rehydrated and resuspended.

9- The tube was placed on the magnetic rack for at least 2 minutes to allow complete capturing of the beads. Once the suspension was cleared, the library-containing supernatant was transferred to a new tube (avoid taking beads).

10- The total-RNA-seq library was stored at -20 ˚C.

7.9.2 PolyA strand specific RNA-seq

First polyA RNAs were isolated from total-RNA using Dynabeads® mRNA Purification Kit (Ambion). For strand specific sequencing, 10 ng isolated polyA RNAs were subjected to the ScriptSeq v2 RNA-Seq kit (Epicentre, USA).

1- 10 ng of polyA RNAs were completed to final volume of 9.5 µl.

2- The following reagents were added in the tube:

Reagent	**Amount for 1 reaction**
PolyA RNA	90 µl
RNA Fragmentation Solution	1 µl
cDNA Synthesis Primer	2 µl
TOTAL	**12 µl**

3- The tube was incubated at 85 °C for 5 minutes in a thermocycler.

4- The fragmentation reaction was stopped by placing the tube on ice.

5- On ice, the cDNA synthesis master mix was prepared by combining the following:

Reagent	**Amount for 1 reaction**
cDNA Synthesis PreMix	3 µl
100mM DDT	0.5 µl
StarScript Reverse Transcriptase	0.5 µl
TOTAL	**4 µl**

6- The cDNA synthesis master mix was mixed gently but thoroughly by pipetting.

7- 4 µl of the cDNA synthesis master mix was added to the reaction containing fragmented RNA on ice. Mixed well.

8- Incubated at 25 ˚C for 5 minutes following by 42 ˚C for 60 minutes, then the reactions were cooled to 37 °C.

9- 1 µl of Finishing Solution was added to the reaction and mixed gently.

10- The reaction was placed on thermocycler and incubated as following

Temperature	**Time**
37 °C	15 minutes
95 °C	3 minutes
4 °C	hold

11- On ice, the Terminal Tagging Master Mix was prepared by combining the following:

Reagent	**Amount for 1 reaction**
Terminal Tagging Premix	7.5 µl
DNA Polymerase	0.5 µl
TOTAL	**8 µl**

12- The reaction from step 10 was removed and 8 µl of the Terminal Tagging Master Mix was add to the reaction and gently mixed by pipetting.

13- Reaction incubated as following:

Temperature	Time
25 °C	10 minutes
95 °C	3 minutes
25 °C	hold

14- The cDNA was next purified prior to PCR amplification using the MinElute PCR Purification Kit (Qiagen).

15- The purified cDNA was eluted in 22.5 µl of Nuclease-Free Water.

16- The purified cDNA was then amplified by combining the following reagents in a tube containing 22.5 µl of cDNA.

Reagent	Amount for 1 reaction
FailSafe PCR PreMix E	25 µl
Forward PCR Primer	1 µl
Reverse PCR Primer (different barcode Reverse PCR Primer for multiplexing)	1 µl
FailSafe PCR Enzyme (1.25 U)	0.5 µl
TOTAL	**50 µl**

17- For the PCR amplification in the thermal cycler the following protocol was used:

Temperature	Time	Number of cycles
95 °C	60 seconds	1 cycle
95 °C	30 seconds	15 cycles
55 °C	30 seconds	
68 °C	3 minutes	
68 °C	7 minutes	1 cycle

18- After library amplification, the library was purified using the MinElute PCR Purification system (Qiagen) and eluted by adding 20 µl of Nuclease-Free Water.

19- The library was stored at -20 °C

7.9.3 High throughput sequencing

All prepared libraries were sequenced using the GAII or HiSeq platforms (Illumina) at the Core Genomics Facilities as part of serveries at the Wellcome Trust Centre for Human Genetics, Oxford.

7.10 Sanger sequencing

Sanger sequencing of vectors were performed by Source Bioscience, U.K. as commercial sequencing servers.

7.11Cloning

7.11.1 Ligation of the RNA-seq libraries to the blunt vector

Conventional cloning was performed in order to confirm successful libraries preparation.

1- First the following 10 µl ligation reaction was setup:

Reagent	Amount per reaction
RNA library (PCR product)	2 µl
pCR®-Blunt (25 ng)	1 µl
10X Ligation Buffer (with ATP)	1 µl
Sterile water	5 µl
T4 DNA Ligase (4 U/µl)	1 µl
Total Volume	10 µl

2- The ligation reaction was incubated at 16 °C for 1 hour, then stored at -20 °C or proceeded with transformation.

7.11.2 Transformation of competent cells

1- 50 µl vial of One Shot® TOP10 cells were thawed on ice for each transformation.

2- 2 µl of each ligation reaction was added directly into the vial of competent cells and mixed by stirring gently with a pipette tip.

3- The vial was incubated on ice for 30 minutes. The remaining ligation mixtures were stored at -20 °C.

4- The cells were heat shocked for 45 seconds at 42 °C without shaking and immediately placed on ice for 2 minutes.

5- 250 µl of room temperature S.O.C. medium was added to each vial.

6- The vial was left for shaking horizontally at 37 °C for 1 hour at 225 xg in a shaking incubator.

7- 60 µl and 80 µl from each transformation vial was spread on pre-warmed LB plates containing 50 µg/ml kanamycin.

8- The plates were incubated overnight at 37 °C.

7.11.3 Transformation analysis

1- At least 80 transformants were picked for analysis. The colony was picked with a tip and put in a tube containing 15 µl of water. The tip was kept inside of the tube.

2- PCR was performed in order to verify if the colony had the vector+insert, using M13F and M13R primers (both priming the blunt vector).

3- The PCR reaction was prepared by mixing the following :

Reagent	**Amount for 1 reaction**
10X AmpliTaq buffer	2.5 µl
MgCl2 (25 mM)	0.5 µl
Ford. Primer (10 uM)	1 µl
Rev. primer (10 uM)	1 µl
dNTP (25 mM)	0.5 µl
AmpliTaq polymerase	0.5 µl
Sterile water	17 µl
Total volume	**23 µl**

4- For each colony picked, 2 µl of the water where the colony had been disposed was added in the PCR mix.

5- PCR cycling conditions:

95 °C	5 minutes	Hold
95 °C	30 seconds	30 cycles
56 °C	30 seconds	
72 °C	45 seconds	
15 °C		hold

6- In order to visualise the PCR result, PCR product was run in a 2 % agarose gel with at least one well with 2 µl of Hypperladder V for resolving the size of the bands.

7- If the colony contains an empty vector (with no library product ligated), the PCR product size should be ~238 bp. If the colony has incorporated a vector containing a library product, then the PCR product size should be ~360 bp for directional RNA libraries.

8- The remaining PCR product was used for sequencing.

7.11.4 Sequencing of PCR products of positive colonies containing vector and Insert

Simi-skirted plates (Greiner Bio-one 96 well half skirted plate -652290) were used for these reactions as these are compatible with ABI 3730 sequencers.

1- 1:1 of Exonuclease I and Shrimp alkaline phosphatase mix was prepared.

2- In a 96 well plate, 1 µl of Exonuclease I/ Shrimp alkaline phosphatise mixture disposed in each well and 1 µl of PCR product. The plate was centrifuged briefly.

3- Initially incubated at 37 ˚C for 30 minutes and then incubated at 80 ˚C for 15 minutes, to deactivate the enzymes. The plate was centrifuged briefly.

4- The following sequencing master mix was prepared:

Reagent	**Amount for 1 reaction**
Big Dye version 3.1	0.25 µl
Big dye buffer	1.875 µl
Primer	2 µl (at 2.5 pmol/µl)
water	3.875 µl
Total volume	**8 µl**

5- 8 µl of sequencing mix was added to each well of the plate containing the treated PCR products and briefly centrifuged.

6- Then DNA was amplified with the following cycling conditions:

96 °C	10 seconds	25 cycles
50 °C	5 seconds	
60 °C	4 minutes	
4 °C		hold

7.11.5 Cleaning up the amplified product

1) 2 µl of 125 mM EDTA was added to each well, 2 µl of 3M sodium acetate was added to each well.
2) 50 µl of 100 % ethanol was added to each well.
3) The plate was covered with aluminium tape and inverted 4 times.
4) Left at room temperature for 15 minutes.
5) Centrifuged at 3,000 xg at 4 ˚C for 30 minutes.
6) Aluminium tape was removed gently. The plate was inverted and spun at 185 xg for 15 seconds.
7) Then 70 µl of 70 % ethanol added to each well.
8) Centrifuged at 1,650 xg at 4 ˚C for 15 minutes.
9) The plate was inverted and spun at 185 xg for 1 minute.

10) The plate was left to air dry the pellets at room temperature for 15 minutes. The plate protected from the light. Once dried stored at -20 ℃.

11) The plate was sent for the Sanger sequencing.

7.12 Small RNA analyses

7.12.1 Reverse transcription

The TaqMan Micro-RNA Reverse Transcription Kit and 5X miRNA-specific RT primers (Applied Biosystems) were used for miRNA reverse transcription.

1- 5 ng of total RNA was used for reverse transcription (RT).

2- The RT master mix was prepared on ice by combining the following:

Reagent	**Amount for 1 reaction**
100 mM dNTPs (with dTTP)	0.15 µl
MultiScribe™ Reverse Transcriptase, 50 U/µl	1 µl
10✕ Reverse Transcription Buffer	1.5 µl
RNase Inhibitor, 20 U/µl	0.19 µl
Nuclease-free water	4.16 µl
TOTAL	**7 µl**

3- The master mix was mixed gently and centrifuged to bring the solution to the bottom of the tube.

4- 5 µl of total RNA was added into 7 µl RT master mix and gently mixed.

5- 3 µl of 5✕miRNA-specific RT primer was added the corresponding RT reaction tube.

6- Mixed thoroughly and incubated on ice for 5 minutes.

7- For RT, the following protocol was followed:

16 °C	30 minutes
42 °C	30 minutes
85 °C	5 minutes
4 °C	hold

8- The RT products were stored at – 20 °C until further analysis.

7.12.2 Small RNA TaqMan® Assays

Expression of microRNA and small nucleolar RNA were assessed by 6-carboxyfluorescein(FAM)-based qPCR with TaqMan Universal PCR Master Mix, individual miRNA primers and hydrolysis probes (all from Applied Biosystems). TaqMan assay IDs are in **supplementary table 5.**

1- In a sterile 1.5 ml microcentrifuge tube for each sample, the following reagents were combined for qPCR:

Reagent	**Amount for 1 reaction (20 µl reaction)**
TaqMan® Small RNA Assay (20✕)	1 µl
Product from RT reaction	1.33 µl
TaqMan® Universal PCR Master Mix II (2✕), no UNG	10 µl
Nuclease-free water	7.67 µl
TOTAL	**20 µl**

2- The tube was inverted several times to mi and centrifuged briefly.

3- 20 µl of the complete qPCR reaction mix was transferred to Bio-red 96 white plate (Bio-Rad).

4- qPCR was performed on the CFX96 Real-Time System C1000 Thermal Cycler (Bio-Rad) as follows:

50 °C	2 minutes	
95 °C	10 minutes	
95 °C	15 seconds	
60 °C	60 seconds	40 cycles

5- Each sample was run in triplicates. At the end of each reaction, the Ct value was determined with instrument default threshold settings.

7.13 mRNA and lncRNA analysis

7.13.1 First-Strand cDNA Synthesis

The SuperScript® III First-Strand Synthesis System (Invitrogen) was used for reverse transcription.

1- 500 ng DNase treatment total RNA was used for cRNA synthesis by combining the following:

Reagent	**Amount for 1 reaction**
50 ng/µl random hexamers	1 µl
10 mM dNTP mix	1 µl
DNase treatment total RNA (500 ng)	X
Nuclease-free water	Up to 10 µl

2- The tube was incubated at 65 °C for 5 minutes, then placed on ice for at least 1 minute.

3- The following cDNA Synthesis Mix was prepared by adding each component in the indicated order:

Reagent	**Amount for 1 reaction**
10X RT buffer	2 µl
25 mM MgCl2	4 µl
0.1 M DTT	2 µ
RNaseOUT™ (40 U/µl)	1 µl
SuperScript® III RT (200 U/µl)	1 µl
TOTAL	**10 µl**

4- 10 µl of cDNA Synthesis Mix was added to each RNA/primer mixture, mixed gently and centrifugation briefly.

5- The RT was carried out in a Thermal Cycler (GeneAmp, PCR System 9700, Applied Biosystems) as following:

25 °C	10 minutes
50 °C	50 minutes
85 °C	5 minutes
Placed on ice	5 minutes

6- The reaction was briefly centrifuged and 1 µl of RNase H was added to each tube and incubated for 20 minutes at 37 °C to complete cDNA synthesis.

7- The cDNA was stored at – 20 °C until further analysis.

7.13.2 Quantitative real-time Polymerase Chain Reaction (qPCR)

To validate the relative expression levels of transcripts in different cell lines, cDNA was synthesised and mixed with Bio-Rad IQ SYBR Green Mix (Bio-Rad) and specific primers for each of the target transcript. 60S ribosomal protein L11, RPL11, was tested for the stability of the expression under hypoxia and was used to normalise the expression levels of different transcripts. Primer sequences are given in **Table 3.1**. QPCR reactions (20 µl) were carried out in a 96-well plate containing 10 µl 2X IQ SYBR Green Mix (Bio-Rad), 1 µl of forward and reverse primers (10 mmol/L), 7 µl nuclease-free H2O, and 2 µl cDNA product. Real time qPCR was performed on the CFX96 Real-Time System C1000 Thermal Cycler (Bio-Rad) as follows:

95 °C	3 minutes	
95 °C	30 seconds	40 cycles
60 °C	30 seconds	
72 ^{0}C	30 seconds	

To obtain the melting curves (thermal profile) of the amplicons, the signal readings were taken at every 0.5 ^{0}C change (81 cycles) with decrease in temperature from 95 ^{0}C to 55 ^{0}C.

7.13.3 TaqMan® gene expression analysis

TaqMan® gene expression assays (Applied Biosystems) were designed to evaluate expression of transcripts. 500 ng of total RNA was subjected for cDNA synthesis using SuperScript III Reverse Transcriptase (Invitrogen). The TaqMan Universal PCR Master Mix and TaqMan® Gene Expression Assay (Applied Biosystems) designed for specific transcript (TaqMan assay IDs are in **supplementary table 5)** were used to amplify cDNA. The following reagents were combined for the reaction:

Reagent	**Amount for 1 reaction (final volume 20 µl)**
TaqMan® Universal Master Mix II (2✕)	10 µl
TaqMan® Gene Expression Assay (20✕)	1 µl
cDNA template	2 µl
Nuclease-free water	7 µl
TOTAL	**20 µl**

20 µl of the complete qPCR reaction mix was transferred to Bio-red 96 white plate (Bio-Rad). The amplification was performed on the CFX96 Real-Time System C1000 Thermal Cycler (Bio-Rad) by using the following protocol:

50 °C	2 minutes	
95 °C	10 minutes	
95 °C	15 seconds	
60 °C	60 seconds	40 cycles

7.14 Rapid amplification of cDNA ends (RACE) PCR

3' Rapid amplification of cDNA ends was performed using the 3´ RACE System kit (Invitrogen™). The cDNA was first synthesised using SuperScript ® III First-Strand Synthesis System (Invitrogen). The RACE PCR was performed using supplied 3´ RACE primers (Reverse Primer), and appropriate forward gene-specific primers. RACE-PCR products were purified using QIAquick PCR Purification Kit (Qiagen). The purified products were cloned into pcr4-TOPO vectors (Invitrogen), and subjected to conventional Sanger sequencing using M13 forward and reverse.

7.15 Primer design

Primers were designed using the publically available primer3 web tool (http://frodo.wi.mit.edu/). To insure specific primer design and avoided amplification of non-target products, each designed primer was assessed using the *silico* PCR tool provided by the UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgPcr?command=start). Primer sequences are in **supplementary table 4.**

7.16 Gel electrophoresis

Melting agarose was added in the TRIS/TBE buffer to prepare the gel. The solution was heated in a microwave to dissolve the agarose. Upon cooling, 5 µl of 0.5µg/mL ethidium bromide was added in a 100 ml gel solution. Then the gel was casted and left to solidify. 6x loading buffer (bromophenol blue and xylene cyanol) (Thermo Scientific) was added in the sample and loaded on the gel. The electrophoresis was performed using a Biorad gel electrophoresis chamber (BioRad, U.K.). Different DNA markers such as HyperLadder I, II, IV, V (Bioline) were used to identify band size of sample. The bands were visualised using ultraviolet light in the AutoChemi Bioimaging System from UVP (U.K.).

7.17 Western blotting

Cells were grown on 10 cm plates, allowed to adhere and were harvested for protein isolation in different conditions. First, cells were washed with ice-cold PBS, scrapped off with RIPA (Radio-Immunoprecipitation Assay) Buffer (150 mM NaCl, 1.0 % IGEPAL® CA-630, 0.5 % sodium deoxycholate, 0.1 % SDS, and 50 mM Tris, pH 8.0, Sigma-Aldrich) supplemented with protease inhibitor cocktail (Roche). The lysates were then centrifuged at 4 °C for 20 minutes at 10,000 xg to remove any cell debris. The supernatants were transferred into new tubes. Protein concentration was measured using the Bradford protein assay (Bio-Rad) following manufacturer's instructions and performed in triplicates.

Protein concentration was normalised between sample and 1X NuPAGE® LDS sample buffer (Invitrogen) supplemented with 2.5 % β-mercaptoethanol was added to the sample. The mixture was boiled for 10 minutes on the heat block at 95 °C and run on readymade 4 % SDS/PAGE gels (Invitrogen) at 100 V for 2 hours in the cold room.

For immunoblotting, the polyvinylidene fluoride (PVDF) membrane (Millipore) was first soaked in 100 % methanol for 1 minute and then washed in 1X transfer buffer for 10 minutes. Proteins were semi-dried and transferred onto membrane (Millipore) using the Mini-Transblot apparatus (BioRad) for 2 hours at 12 V. After that membranes were blocked with 5 % bovine serum albumin (BSA) (Sigma-Aldrich) in Tris buffered saline (pH 7.5) with 0.2 % Tween-20 (TBS-T) for 2 hours at room temperature and then incubated overnight at 4 °C with the primary antibodies in 5 % BSA TBS-T. The membrane was then washed three times for 10 minutes with 1 % non-fat dry milk in TBS-T, followed by incubation with secondary goat horseradish peroxidase-conjugated anti-rabbit IgG (Invitrogen) diluted 1:10000 in 5 % non-fat dry milk in TBS-T for 45 minutes at room temperature. After that the membrane was washed three times with 5 % non-fat dry milk in TBS-T. Immunodetection was visualised using the chemiluminescence reagents from the Amersham ECL Detection Reagents (GE Healthcare). The X-OMAT LS films (Kodak, U.K.) were exposed to the membranes protein-antibody complexes.

7.18 Immunofluorescence and confocal microscopy

Immunofluorescence assay was performed for paraspeckles proteins on MCF-7 and ZR75-1 cells in hypoxia and normoxia. 20×10^3 cells were seeded onto cover slips in 24 wells plate. Cells were left overnight to attach and settle down. Next day, the medium was removed and cells were rinsed twice with PBS. The cells were fixed with 4 % paraformaldehyde (Pierce) in PBS for 10 minutes at room temperature and were permeabilized with 0.2 % Triton X-100 (T8532, Sigma-Aldrich) in PBS for 10 minutes at room temperature. After permeabilizion, cells were washed twice with PBS. After that, cells were blocked with 1 % BSA (Sigma-Aldrich) in PBS for 1

hour, followed by incubation with either mouse antibodies; PSPC1 (1:500, sc-374181, Santa Cruz Biotechnology), NONO (1:500, 611279, BD Transduction Laboratories™), or SFPQ (1:700, WH0006421M2-100UG, Sigma-Aldrich) and kept overnight at 4 °C. The coverslips were washed with PBST three times before conjugating with fluorescent 488- secondary antibody (Life Technologies) diluted 1/200 in 1 % BSA-PBST for 30 minutes at room temperature. The coverslips were rinsed twice with PBS and distilled water. The coverslips were then mounted into a microscope slide using the DAPI Fluoromount G (Southern Biotech). The slides were visualized via the Confocal microscopy and all images were taken using LSM 510 META Ziess confocal microscopy.

7.19 Apoptosis analysis

7.19.1 Caspase 3 and 7 activities

The Caspase 3/7 activity was measured using Caspase-Glo® 3/7 Assay (Promega) according to the manufacturer protocol. Around 1,000 cells were seeded per well in 96-well plate and allowed to adhere overnight. Cells were then treated with NEAT1 and control ASO , then the cells were cultured in hypoxia and normoxia for 24 hours. 100 µl of Caspase-Glo® 3/7 reagent was added to each well and incubated for 60 minutes. Caspase 3/7 activity was measured using fluorescence microplate reader (SpectraMax M2, USA). Three independent experiments were performed with three technical replicates.

7.19.2 Fluorescence Activated Cell Sorting (FACS) analysis

In 6-well tissue culture plates, about 150,000 cells were seeded per well and allowed to adhere overnight. Cells were transfected with NEAT1 ASO and control ASO and incubated in normoxic and hypoxic conditions for 24 hours. After incubation, culture medium was collected in 5 ml polypropylene tubes (BD Falcon). The cells were washed twice with cold 0.5 % Bovine Serum Albumin diluted in Phosphate Buffered Saline (PBSA) and detached using 0.25 % trypsin-EDTA

solution at 37 °C. Detached cells were washed and resuspended with their complete culture medium-PBSA mixture. Next, cells were pelleted by centrifugation at 1,000 xg for 5 minutes and washed with PBSA twice. The cell pellet was then resuspended in 350 µl of Annexin V (AV) and with Propidium iodide (PI) solution. Then cells were incubated in a dark at room temperature for 15 minutes. The stained cells were analysed by CyAn ADP analyser (Beckman Coulter). AV-Alexa Flour 647 was measured by excitation at 642 nm and emission detection at 665 nm was measured by excitation at 405 nm and emission detection at 675 nm.

7.20 Clonogenic assay

Clonogenic assay was performed as described in Nicolass AP et al., 2006. Cells were seeded in 6-well plates and allowed to adhere overnight. NEAT1 and control antisense oligonucleotides were transfected in cells and cultured for 24 hours. Next day, the medium were changed and cells were cultured in hypoxia and normoxia for another 24 hours. Cells were washed with PBS and stained with 0.2 % methylene blue in ethanol for 60 minutes. The image was taken and the cover area was quantified using ImageJ software. Data were analysed by the average of three independent experiments.

7.21 Cell proliferation assay

Effect of knockdowns on cell proliferation was measured using the CyQUANT® NF Cell Proliferation Assay Kit (Invitrogen) according to the manufacturer's instructions. The assay is based on measuring the cellular DNA content of cell. Since the amount of cellular DNA content is closely proportional with number of cells, the extent of proliferation can be determined by comparing cell counts with/without knockdowns. First, the solutions were prepared as follows:

1- 2.2 ml of 5X HBSS buffer was added in 8.8 ml of deionized water to obtain 1X dye binding solution.

2- 22 µl of CyQUANT® NF dye reagent was added into 11 ml of 1X HBSS buffer.
3- Around 1,000 cells were seeded per well in a 96-well plate and allowed to adhere overnight. Cells were then treated with NEAT1 and control ASO, then the cells were cultured in hypoxia and normoxia for 24 hours.
4- After incubation, 100 µl of 1X dye binding solution was added in each well.
5- The microplate was covered and incubated at 37 °C for 60 minutes.
6- The fluorescence intensity of each sample was measured by a fluorescence microplate reader (SpectraMax M2, USA) with excitation at ~485 nm and emission detection at ~530 nm. The three independent experiments were performed with three technical replicates.

7.22 RNA fluorescent *in situ* hybridization (RNA FISH)

Approximately 10^4 of MCF7 cells or 10^3 of ZR75-1 cells were seeded onto 13 mm circular coverslips in 24 well dishes and incubated in normoxic and hypoxic conditions or subjected for different knockdowns. The QuantiGene ViewRNA slide-based assay kit (Affymetrix, QVC0001) was used according to the manufacturer's recommendations. RNA FISH experiment was optimised and performed by Daniela Moralli (Cytogenetics and Microscopy Core, WTCHG, University of Oxford). The kit allows to semi-quantify transcript expression and track the cellular localization of transcript in hypoxia. A viewRNA probe set designed against RPL11, (Affymetrix VA4-11811) and was used as a positive control. ViewRNA probe sets were designed against NEAT1 (and custom probes were designed against aSPAG4 and aCELSR2). The RNA FISH experiments were analysed using an Olympus BX60 microscope for epifluorescence equipped with a Sensys CCD camera (Photometrics, USA). Images were collected using Genus Cytovision 7.1 software (Leica). Three independent experiments were performed with a minimum of 100 cells which were analysed in each replicated experiment.

7.23 Transfection of small interfering RNA (siRNA)

Previously validated siRNAs targeting HIF-1α, HIF-2α and control scramble were obtained from Dharmacon, UK. HIF siRNAs sequences are in **supplementary table 6.** The siRNAs were transfected into cells using Lipofectamine® RNAiMAX reagent (Invitrogen) according to the manufacturer's instructions.

1. Cells were plated into 6-well plates and cultured until a confluency of 60 % was reached.
2. Per well, 0.3 ml of reduced serum Optimem medium (Invitrogen) was mixed with 3 µl of Lipofectamine® RNAiMAX in a 1.5 ml tube.
3. In a new tube, 1.2 ml of reduced serum Optimem medium was added with 1.5 µl of siRNA (40 µM) for final concentration of 40 nM in a final volume of 1.5 ml. Tubes were briefly mixed and kept at room temperature for 5 minutes.
4. The Lipofectamine® RNAiMAX mix was transferred into the siRNA mix. The tubes were briefly mixed and kept at room temperature for 15 minutes.
5. Cells were washed with PBS, then 1.5 ml of transfection mix was then added onto the cells and incubated for 4 hours at 37 °C.
6. After incubation, 0.5 ml of reduced Serum Optimem medium was added into the cells.
7. The transfection was repeated after 24 hours.
8. After the second transfection, the medium was replaced with a fresh complete medium, the cells were then incubated in hypoxia.

7.24 Transfection of antisense oligonucleotides (ASO)

Antisense oligonucleotides targeting NEAT1 and control scramble were designed and obtained by Integrated DNA Technologies. ASOs were transfected into cells using Lipofectamine® RNAiMAX reagent (Invitrogen) according to the manufacturer's instructions. Control and

NEAT1 ASOs sequences are in **supplementary table. 7.** For effective knockdown, two transfections were carried out 24 hours apart. Transfection was performed in 6 well plates.

1- Per well, 250 µl of reduced serum Optimem medium (Invitrogen) was mixed with 6 µl of Lipofectamine® RNAiMAX in a 1.5 ml tube.
2- In a new tube, 250 µl of reduced serum Optimem medium was added with 3 µl of siRNA (100 µM) for final concentration of 50 nM in a final volume of 3 ml. Tubes were briefly mixed and kept at room temperature for 5 minutes.
3- The Lipofectamine® RNAiMAX mix was transferred into the siRNA mix. The tubes were briefly mixed and kept at room temperature for 15 minutes.
9. 2.5 ml of transfection mix was added in an empty well. 25,000 cells were added in each well containing the transfection mix and then incubated for 10 minutes at 37 °C.
4- After incubation, 0.5 ml of reduced Serum Optimem medium was added into the cells.
5- The transfection was repeated after 24 hours.
6- After the second transfection the medium was replaced with a fresh complete medium, the cells were then incubated in hypoxia.

7.25 Statistical analyses

All the experiments were performed in duplicate or triplicate. Statistical analyses using two-tailed *t* tests were performed in SPSS (IBM Corp, New York, NY, USA) and R (http://www.R-project.org). The graphs were produced using GraphPad Prism® v5.0 software (GraphPad).

7.26 Bioinformatic analysis

7.26.1 Mapping pipeline

RNA-seq data were mapped using *in-house* analysis pipelines. These allowed sequential mapping of the raw data to specific RNA databases and for calculation of the expression level of the mapped transcripts. The analysis pipeline was initiated by trimming the adapter sequences from the raw sequence data using the FASTX-tool kit (http://hannonlab.cshl.edu/fastx_toolkit/). The sequences were mapped sequentially to databases of different RNA classes, according to Figure 1B, using Bowtie (version 1.0.0). The RNA databases used were: human protein coding database (UCSC), rRNA database (Wellcome Trust Centre for Human Genetics bioinformatics group), piwiRNA database (NCBI), microRNA database (mirBase), sn/snoRNA database (UCSC), tRNA database (UCSC), and long coding RNAs (ENCODE). All databases were based on the 2009 human genome release, hg19. The first round of mapping used stringent criteria (1 mismatch, threshold of mismatches ≤ 70) and was followed by a second less stringent round (2 mismatches).

7.26.2 Analysis of non-annotated transcripts

We next used the Cufflinks package (version 2.1.1), with default settings, to *de novo* assemble transcripts from contiguous reads that had not mapped to the annotated databases above. We excluded transcripts with an abundance of less than 13 % of neighbouring annotated transcripts or that overlapped an annotated transcript from the same orientation. Transcripts were also filtered for the presence both RNAPol2 and H3K4me3 at their putative TSS. The remaining filtered transcripts were defined as intergenic (no overlap with annotated transcripts) or antisense (overlapping with but in the opposite orientation to an annotated transcript). We further required that antisense transcripts overlapped the “sense” transcript by at least 100 reads. To assess the coding potential of transcripts ‘‘CPC calculator ’’ tool was used.

7.26.3 Quantification of gene expression

To determine transcript regulation, mapped sequence data was normalised to the total number of mapped reads (excluding the ribosomal reads). Minimum thresholds were then applied to transcript abundance (10 reads per transcript piwiRNA, miRNA, snRNA and tRNA and 100 reads per transcript for pcRNAs and lncRNAs), since at low abundance the estimation of hypoxic regulation was poor . The ratio of normalised reads between two samples was calculated as fold change.

7.26.4 ChIP-seq and peak calling analysis

All ChIP-seq data sets were mapped using Bowtie (version 1.0.0) to the human genome (build GRCh37/hg19). Data were normalised to the total number of mapped reads in each dataset. Peak finding was performed using CisGenome as previously described.

7.26.5 Data depository

Raw and mapped data for RNA-seq and ChIP-seq is available at EMBL-EBI Array Express (E-MTAB-1994, E-MTAB-1995).

Chapter 8: Supplementary

Supplementary table 1. Transcripts that manifest greatest fold-regulation in hypoxia for each class

PiwiRNA					
Genes	**Chr**	**start**	**end**	**strand**	**Fold Change**
Upregulated					
piRNApiR-52416	chr1	229567886	229567917	-	14
piRNApiR-57604	chr6	74228271	74228299	-	4.9
piRNApiR-47161	chr1	228823233	228823260	+	4.9
piRNApiR-52681	chr17	17146601	17146629	+	4
piRNApiR-35058	chr14	90341397	90341425	-	2.5
Downregulated					
piRNApiR-32376	chr11	10530773	10530804	-	0.21
piRNApiR-36248	chr17	37023897	37023926	+	0.2
piRNApiR-31612	chr2	207026644	207026674	+	0.2
piRNApiR-49039	chr5	180633907	180633933	+	0.2
piRNApiR-36247	chr17	37023897	37023925	+	0.12

miRNA					
Genes	**Chr**	**start**	**end**	**strand**	**Fold Change**
Upregulated					
mir-184	chr15	79502130	79502213	+	205
mir-9-3	chr15	89911248	89911337	+	65
mir-9-1	chr1	156390133	156390221	-	64
mir-1-1	chr20	61151513	61151583	+	49
mir-1-2	chr18	19408965	19409049	-	49
Downregulated					
mir-539	chr14	101513658	101513735	+	0.35
mir-3614	chr17	54968631	54968716	-	0.3
mir-3658	chr1	165877158	165877213	+	0.18
mir-329-1	chr14	101493122	101493201	+	0.1
mir-620	chr12	116586365	116586459	-	0.03

snRNA/snoRNA					
Genes	**Chr**	**start**	**end**	**strand**	**Fold Change**
Downregulated					
ACA67B	chr2	10586839	10586974	-	0.1
ACA24	chr4	119200345	119200475	+	0.09
U31	chr11	62620796	62620866	-	0.08
U18C	chr15	66793589	66793655	-	0.06
SNORD124	chr17	38183794	38183897	-	0.04

tRNA					
Genes	**Chr**	**start**	**end**	**strand**	**Fold Change**
Downregulated					
tRNA114- Gln CTG	chr1	146476760	146476831	-	0.3
tRNA18- Tyr	chr14	21125623	21125716	-	0.29
tRNA3- Asp GTC	chr3	184366095	184366165	-	0.29
tRNA2- Leu TAA	chr4	156384978	156385052	-	0.28
tRNA12- Tyr	chr8	66609532	66609619	-	0.12

mRNA					
Genes	**Chr**	**start**	**end**	**strand**	**Fold Change**
Upregulated					
NDRG1	chr8	134249413	134309547	-	24
ALDOC	chr17	26900132	26903951	-	10
NR4A1	chr12	52445185	52453291	+	8.6
FOS	chr14	75745480	75748937	+	8
AHNAK2	chr14	105403590	105444694	-	7.6
Downregulated					
PCNA	chr20	5095598	5100647	-	0.23
HNRNPAB	chr5	177631507	177638184	+	0.23
PKIB	chr6	122973873	123047518	+	0.23
LTV1	chr6	144164507	144184943	+	0.22
CHAC2	chr2	53994928	54002287	+	0.18

lncRNA					
Genes	**Chr**	**start**	**end**	**strand**	**Fold Change**
Upregulated					
NEAT1	chr11	65190245	65192231	+	4.3
TCONS_00013573	chr7	123560905	123564184	+	4.2
TCONS_l2_00017632	chr22	22012077	22012712	+	4.2
TCONS_00015697	chr9	78194585	78203464	+	4.1
TCONS_00028106	chr20	10844845	10846396	+	4
Downregulated					
TCONS_00015820	chr9	46661444	46662837	-	0.29
TCONS_00007500	chr4	29119930	29204392	+	0.21
TCONS_l2_00023536	chr5	129096155	129100756	+	0.17
TCONS_00011521	chr6	79314184	79315753	-	0.12
TCONS_00011596	chr6	153115631	153153048	-	0.08

Supplementary table 2. Hypoxia regulated non-annotated intergenic transcripts

ID	Chr	Start	End	Strand	Fold change in hypoxia	CPC
Novel 1	chr1	202053848	202082554	+	1.5	-0.63
Novel 2	chr1	202053658	202057903	+	1.2	-0.6
Novel 3	chr1	202002600	202088108	+	1.6	-0.56
Novel 4	chr1	201983168	202077686	+	1.6	0.84
Novel 5	chr1	181069443	181074240	-	1.5	-0.95
Novel 6	chr1	148241355	148241861	+	1.4	-1.32
Novel 7	chr1	107292908	107293420	+	1.2	-1.04
Novel 8	chr1	107073060	107424139	+	1.1	0.73
Novel 9	chr1	106734807	106876406	+	0	-0.24
Novel 10	chr1	106735450	106762265	+	1.1	-0.77
Novel 11	chr1	45196573	45196842	-	0.3	-1.36
Novel 12	chr10	8911651	8932114	+	1.1	-0.99
Novel 13	chr10	8909994	8931788	+	1.3	0.26
Novel 14	chr10	5524638	5535702	-	1.4	-0.52
Novel 15	chr11	65222490	65244737	+	1.7	-0.84
Novel 16	chr12	25956136	25959121	+	0.8	0.41
Novel 17	chr12	8116260	8123253	-	1.5	-0.81
Novel 18	chr15	51328348	51690847	+	1.4	-0.78
Novel 19	chr16	85587072	85589839	-	2	-0.88
Novel 20	chr16	85495495	85603598	+	1.7	0.005
Novel 21	chr16	85479122	85498916	-	1.8	-0.77
Novel 22	chr16	14395080	14494860	+	1.5	-0.57
Novel 23	chr16	85496110	85524891	+	1.6	-0.46
Novel 24	chr2	132300855	132350688	+	1.5	-0.58
Novel 25	chr2	75136093	75145898	+	1.3	-0.86
Novel 26	chr2	1571357	1629598	-	1.5	-0.34
Novel 27	chr20	292339	305552	-	1.8	-0.87
Novel 28	chr3	64022004	64104438	+	1.3	-0.79
Novel 29	chr5	177485281	177505466	-	1.6	-0.73
Novel 30	chr5	172721232	172730376	+	1.8	0.43
Novel 31	chr5	102088692	102198533	+	1.4	-0.86
Novel 32	chr5	66509857	66520819	+	1.4	-0.87
Novel 33	chr5	66509437	66891986	+	1.2	0.75
Novel 34	chr7	157283010	157312766	+	1.4	-0.27
Novel 35	chrX	95228382	95229237	-	1.3	0.02
Novel 36	chr11	70995192	70998931	+	1.8	-0.52
Novel 37	chr7	129244613	129251530	-	1.9	-0.99

Supplementary table 3. Hypoxia non-annotated antisense transcripts (NATs)

Name	Chr	Start	End	Strand	Fold regulation	Overlapping region of original gene	Regulation of sense gene
aARMC5	chr16	31461767	31470706	-	1.6	5'	Upregulated
aATF7IP2	chr16	10527741	10712970	-	1.5	3'	Upregulated
aMAMDC4/ PHPT1	chr9	139738901	139759957	-	1.3	Cover	Upregulated
aTSEN54	chr17	73517363	73521618	-	1.3	3'	Downregulated
aTPBG	chr6	83072075	83075226	-	1.3	5'	Upregulated
aTBX2	chr17	59470718	59477783	-	1.5	5'	Downregulated
aSTXBP2	chr19	7697325	7702009	-	1.6	5'	Upregulated
aSPAG4	chr20	34196650	34207240	-	1.4	5'	Upregulated
aCHRAC1	chr8	141516112	141521724	-	1.5	5'	Downregulated
aGRIN1	chr9	140056711	140062733	-	1.5	3'	Not expressed
aHIF1A	chr14	62175449	62214842	-	1.3	3'	Downregulated
aLENG8	chr19	54959291	54960259	-	1.3	5'	Upregulated
aLOC338758	chr12	90056440	90103653	-	1.4	5'	Upregulated
aSPTSSB	chr3	161089733	161146159	+	1.2	5'	Upregulated
aBRAT1	chr7	2577658	2587164	+	1.1	3'	Downregulated
aZIC4	chr3	147111164	147128972	+	1.6	5'	Non regulated
aTUT1	chr11	62341769	62346553	+	1.3	Middle	Non regulated
aTRIM17	chr1	228594940	228602400	+	1.3	3'	Downregulated
aTHEM5	chr1	151812700	151825963	+	1.3	3'	Not expressed
aTBC1D10B	chr16	30366581	30370364	+	1.5	3'	Upregulated
aSRMS	chr20	62170196	62175250	+	1.3	3'	Upregulated
aSLC39A4	chr8	145638718	145644681	+	1.5	5'	Non regulated
aRASEF	chr9	85677150	85684988	+	1.1	5'	Upregulated
aPTPRN	chr2	220163641	220168902	+	1.3	3'	Not expressed
aPRICKLE2	chr3	64021527	64102071	+	1.1	5'	Upregulated
aCPLX1	chr4	776511	783395	+	1.4	3'	Non regulated
aGMEB2	chr20	62258636	62259915	+	1.3	5'	Upregulated
aGPR132	chr14	105499226	105522613	+	1.5	3'	Upregulated
aHDLBP	chr2	242157867	242184464	+	1.8	3'	Not expressed
aKPNA1	chr3	122135105	122146064	+	1.5	3'	Downregulated
aLCAT	chr16	67969845	67982503	+	1.4	Cover	Non regulated
aBCAS3	chr17	59432883	59480160	-	1.5	3'	Downregulated
aCELSR2	chr1	109788560	109796206	-	1.5	5'	Downregulated
aBCAN	chr1	156607605	156616958	-	1.4	5'	Upregulated
aADRM1	chr20	60880596	60884025	-	1.1	Middle	Downregulated
aAPC2	chr19	1457662	1461617	-	1.1	Middle	not expressed
aPCIF1	chr20	44562017	44563625	-	1.4	5'	Upregulated

aABHD16B	chr20	62492845	62496432	-	1.4	Cover	Not expressed
aTPD52L2	chr20	62492903	62497100	-	1.4	5'	Downregulated
aTMEM167B	chr1	109631814	109633438	-	1.5	5'	Upregulated
aTM6SF1	chr15	83781780	83784537	-	1.5	Middle	Not expressed
aTRIM52	chr5	180673525	180684320	+	1.6	3'	Upregulated
aPRICKLE2	chr3	64015517	64092155	+	1.4	3'	Upregulated
aACOT1	chr14	74034390	74036131	-	1.6	5'	Downregulated
aCCDC170	chr6	151931492	151937995	-	1.3	3'	Upregulated
aZFYVE16	chr5	79773878	79783866	-	1.2	3'	Downregulated
aHOXC13	chr12	54329157	54332757	-	1.6	5'	Non regulated
aC12orf61	chr12	62996543	63010159	+	1.4	5'	Upregulated
aPUF60	chr8	144911368	144914231	+	1.3	5'	Downregulated
aSLC25A24	chr1	108742224	108744182	+	1.4	5'	Downregulated
aMIPEP	chr13	24289332	24309508	+	1.2	3'	Downregulated
aGPR132	chr14	105498949	105522080	+	1.2	3'	Not expressed
aIRS1	chr2	227661484	227680686	+	1.2	5'	Upregulated
aTPO	chr2	1476408	1626668	-	1.2	3'	Not expressed

Supplementary table 4. Oligonucleotides for qPCR used in this project**.**

Transcript name	Direction	Sequence 5' to 3'
TMSB10	Forward	GACCATTGAGCAGGAGAAGC
TMSB10	Reverse	GTCCCTTCAGAGACCCACAG
PVRL4	Forward	CAAAATCTGTGGCACATTGG
PVRL4	Reverse	GCTGACATGGCAGACGTAGA
STON1	Reverse	TTGCCTCTGCTTTGTGAATG
STON1	Forward	TGAAACGCATCAGCTCAAAC
RAB11FIP5	Forward	GCTCGAAAAGCAAAAAGGTG
RAB11FIP5	Reverse	GCTCGAAAAGCAAAAAGGTG
BCKDK	Forward	CAGAACTTGGAGCAGGGAAG
BCKDK	Reverse	ACGAACCAGTTCAACCCAAG
RCN3	Forward	GACTTCCGGGATCTGAACAA
RCN3	Reverse	CCATAGTTGGTGGCCTGACT
BCOR	Forward	GCAAAGAGAATCGCCAACTC
BCOR	Reverse	TCTCGGAGTCTTTGGTTGCT
VTCN1	Forward	CCCAATCCGAAGTGTCAACT
VTCN1	Reverse	TATCCTGGTGCCCGATAGAG
FUT3	Forward	AGGTGTACCCACAGGCAGAC
FUT3	Reverse	GCGGTAGGACATGGTGAGAT
HOOK3	Forward	GTGAACCTTATTGGGGAGCA
HOOK3	Reverse	CAGCTGACGATCAAGGTCAA
CA9	Forward	ACAATGTGCATATCGTGGCA
CA9	Reverse	CAGAGGGGCTCTGATAGGTTT
VEGF	Forward	CCTTGCTGCTCTACCTCC
VEGF	Reverse	AGCTGCGCTGATAGACAT
RPL11	Reverse	TCCAAGATTTCTTCTGCCTTG
RPL11	Forward	CTTTGGCATCCGGAGAAAT
XIST	Forward	TGCTGATCATTTGGTGGTGT
XIST	Reverse	TGACTTCCTCTGCCTGACCT
HOTAIR	Forward	GGTAGAAAAAGCAACCACGAAGC
HOTAIR	Reverse	ACATAAACCTCTGTCTGTGAGTGCC
H3J	Forward	GAGATCCGCCGTTATCAGAA
H3J	Reverse	CTTGGCGTGAATAGCACAGA
H1D	Forward	CTGCTCCACTTGCTCCTACC
H1D	Reverse	CCAAGCTTGATACGGCTGTT
H3C	Forward	GTGGCGTGAAGAAACCTCAT
H3C	Reverse	TTAGCGTGAATAGCGCACAG
aSPAG4	Forward	TTAAGAAGGAGCAGGGCATC
aSPAG4	Reverse	TGTGGGTATGACGATTGGTG
SPAG4	Forward	AAGGAGGTGTCCACTGTTCG
SPAG4	Reverse	GCGTAATCGTGGGATGTCTT
aSLC25A24	Forward	GTGTCTGCCAGCTGCACTTA
aSLC25A24	Reverse	GACTACCCGGAGACGCTATG
aGMEB2	Forward	GACGCTGCTGTGTGTGAATG

aGMEB2	Reverse	GCCCTCCTCCTACAAGTGTG
aTMEM167B	Reverse	AGAAGCCGTAGTACGCAGGA
aTMEM167B	Forward	AGTTACCACGGCCAAGTACG
aTRIM52	Forward	CTAAAGCTGAAGGGCAGGTG
aTRIM52	Reverse	TTTCAAAACTGTGGGTGTGC
aCCDC170	Forward	TAGTCTTGCGCTTCCTGACA
aCCDC170	Reverse	CAGCTGTTTACCGTGTAGGTCA
aABHD16B	Forward	CGGGGACTAGGGACATGTAG
aABHD16B	Reverse	CCTAGCTCTGTCCCCTCCTT
aGPR132	Forward	GACCTGGGAGAACAGTTCCA
aGPR132	Reverse	TCCAGCCTCTAGCCACAGTT
aTHEM5	Forward	GTTGCCAAGCATGTGACTGT
aTHEM5	Reverse	TTGGATCCTCCACAGACTCC
aCELSR2	Forward	TTCCACCCAAAGGCAGATAG
aCELSR2	Reverse	CCCAAGTCCCGTCTCCTCTA
CELSR2	Forward	GAGAAGCCCTACTGCCAGGT
CELSR2	Reverse	TCATTGAAACGCCCATTGTA
MALAT1	Forward	TGTGTGCCAATGTTTCGTTT
MALAT1	Reverse	AGGAGAAAGTGCCATGGTTG
NEAT1	Forward	ATGCTGATCTGCTGCGTATG
NEAT1	Reverse	CCAGTTTTCCGAGAACCAAA
chr16:85587072-85589839	Forward	GGTCGGAGAGAGAGCCTTTT
chr16:85587072-85589839	Reverse	TTCCATAGTGGCCAGAGAGG
chr11:70995192-70998931	Forward	TGCTGTGCACTTGTGTGTTC
chr11:70995192-70998931	Reverse	CCCCTGCAGTCACCTTTTAT
chr7:129244613-129251530	Forward	TGAAGGAGTAGGAGGGCAAA
chr7:129244613-129251530	Reverse	AAACACGTGGCTTTCCTGTT
chr20:292339-305552	Forward	GACCAGCATCTCAACTGCAA
chr20:292339-305552	Reverse	TCTTCCCCTCCTCTTCCAAT
chr5:172721232-172730376	Forward	CGCAACTGTAGCACATGGAT
chr5:172721232-172730376	Reverse	TGCCTTCGCAAGAGTCTTCT
chr5:177485281-177505466	Forward	GTAGGGCCTGTGCCTTTTCT
chr5:177485281-177505466	Reverse	TGGTTGAGACGTCTGTGAGC
chr12:8116260-8123253	Forward	AAAAAGGAGAGCGCGTACTG
chr12:8116260-8123253	Reverse	GCCCGAACTACAATTCTTGG
chr16:85479122-85498916	Forward	ACCACCTAAGCCGACATCTG
chr16:85479122-85498916	Reverse	CTCTGTGTGCACCTCTCTGC
TCONS_00021778	Forward	GCCTAGGTTCCACCCTTTTC
TCONS_00021778	Reverse	CGGGAACAGAAAGAAAGCAG
TCONS_00010677	Forward	GGCTGGAGTACAGTGGCATAG
TCONS_00010677	Reverse	GCGAAACACGGTGTCTACAA
TCONS_00026155	Forward	GCAGCAGAACTGATGGTGAA
TCONS_00026155	Reverse	AGTGCCACAACCCCTGATAG
TCONS_l2_00017632	Forward	ACTCTTGTTGCAGGCGAAGT
TCONS_l2_00017632	Reverse	ACGGGGAGAGGTTGATCC

TCONS_00024694	Forward	GCTAGAGGGCAGCTACAGGA
TCONS_00024695	Reverse	ATTTGGTAGGAGGGCAGCTT
TCONS_00018859	Forward	ATGGGCAAGAACAGGATGTC
TCONS_00018859	Reverse	GGACCAGGGTTGACTAATGC
TCONS_00012630	Forward	TGTTCAGGGTGTCTCTGCTG
TCONS_00012630	Reverse	CACTGCCTCTGGTGATCTGA
TCON_00019373	Forward	GCTTCACGACGTGTTGCTAA
TCON_00019373	Reverse	TCAGACTGGGCTTTGTAGGG
TCONS_00019713	Forward	AAGGCTCCTGACTGTGGAGA
TCONS_00019713	Reverse	GAGGAGTCTCACCTGCAAGC
TCONS_l2_00016774	Forward	TGTCAAGCCCTGCTCTTTCT
TCONS_l2_00016774	Reverse	GTTCCTTCTGCTGCAACTCC
TCONS_00020202	Forward	CAGCCTATGAGCTCCTCCAG
TCONS_00020202	Reverse	GAGACTGCTCTCCACGAACC
TCONS_00000020	Forward	GCCGTTTTACATTGGCTGTT
TCONS_00000020	Reverse	ATTTGGAAGGTCGTCCACTG
TCONS_l2_00027397	Forward	AGGAAGCAACTTTCCAGCAA
TCONS_l2_00027397	Reverse	CTCTGGGTCAGGAATTTGGA
TMSB10	Forward	GACCATTGAGCAGGAGAAGC
TMSB10	Reverse	GTCCCTTCAGAGACCCACAG
PVRL4	Forward	CAAAATCTGTGGCACATTGG
PVRL4	Reverse	GCTGACATGGCAGACGTAGA
STON1	Reverse	TTGCCTCTGCTTTGTGAATG
STON1	Forward	TGAAACGCATCAGCTCAAAC
RAB11FIP5	Forward	GCTCGAAAAGCAAAAAGGTG
RAB11FIP5	Reverse	GCTCGAAAAGCAAAAAGGTG
BCKDK	Forward	CAGAACTTGGAGCAGGGAAG
BCKDK	Reverse	ACGAACCAGTTCAACCCAAG
RCN3	Forward	GACTTCCGGGATCTGAACAA
RCN3	Reverse	CCATAGTTGGTGGCCTGACT
BCOR	Forward	GCAAAGAGAATCGCCAACTC
BCOR	Reverse	TCTCGGAGTCTTTGGTTGCT
VTCN1	Forward	CCCAATCCGAAGTGTCAACT
VTCN1	Reverse	TATCCTGGTGCCCGATAGAG
FUT3	Forward	AGGTGTACCCACAGGCAGAC
FUT3	Reverse	GCGGTAGGACATGGTGAGAT
HOOK3	Forward	GTGAACCTTATTGGGGAGCA
HOOK3	Reverse	CAGCTGACGATCAAGGTCAA
CA9	Forward	ACAATGTGCATATCGTGGCA
CA9	Reverse	CAGAGGGGCTCTGATAGGTTT
VEGF	Forward	CCTTGCTGCTCTACCTCC
VEGF	Reverse	AGCTGCGCTGATAGACAT
RPL11	Reverse	TCCAAGATTTCTTCTGCCTTG
RPL11	Forward	CTTTGGCATCCGGAGAAAT
XIST	Forward	TGCTGATCATTTGGTGGTGT

XIST	Reverse	TGACTTCCTCTGCCTGACCT
HOTAIR	Forward	GGTAGAAAAAGCAACCACGAAGC
HOTAIR	Reverse	ACATAAACCTCTGTCTGTGAGTGCC
H3J	Forward	GAGATCCGCCGTTATCAGAA
H3J	Reverse	CTTGGCGTGAATAGCACAGA
H1D	Forward	CTGCTCCACTTGCTCCTACC
H1D	Reverse	CCAAGCTTGATACGGCTGTT
H3C	Forward	GTGGCGTGAAGAAACCTCAT
H3C	Reverse	TTAGCGTGAATAGCGCACAG
aSPAG4	Forward	TTAAGAAGGAGCAGGGCATC
aSPAG4	Reverse	TGTGGGTATGACGATTGGTG
SPAG4	Forward	AAGGAGGTGTCCACTGTTCG
SPAG4	Reverse	GCGTAATCGTGGGATGTCTT
aSLC25A24	Forward	GTGTCTGCCAGCTGCACTTA
aSLC25A24	Reverse	GACTACCCGGAGACGCTATG
aGMEB2	Forward	GACGCTGCTGTGTGTGAATG
aGMEB2	Reverse	GCCCTCCTCCTACAAGTGTG
aTMEM167B	Reverse	AGAAGCCGTAGTACGCAGGA
aTMEM167B	Forward	AGTTACCACGGCCAAGTACG
aTRIM52	Forward	CTAAAGCTGAAGGGCAGGTG
aTRIM52	Reverse	TTTCAAAACTGTGGGTGTGC
aCCDC170	Forward	TAGTCTTGCGCTTCCTGACA
aCCDC170	Reverse	CAGCTGTTTACCGTGTAGGTCA
aABHD16B	Forward	CGGGGACTAGGGACATGTAG
aABHD16B	Reverse	CCTAGCTCTGTCCCCTCCTT
aGPR132	Forward	GACCTGGGAGAACAGTTCCA
aGPR132	Reverse	TCCAGCCTCTAGCCACAGTT
aTHEM5	Forward	GTTGCCAAGCATGTGACTGT
aTHEM5	Reverse	TTGGATCCTCCACAGACTCC
aCELSR2	Forward	TTCCACCCAAAGGCAGATAG
aCELSR2	Reverse	CCCAAGTCCCGTCTCCTCTA
CELSR2	Forward	GAGAAGCCCTACTGCCAGGT
CELSR2	Reverse	TCATTGAAACGCCCATTGTA
MALAT1	Forward	TGTGTGCCAATGTTTCGTTT
MALAT1	Reverse	AGGAGAAAGTGCCATGGTTG
NEAT1	Forward	ATGCTGATCTGCTGCGTATG
NEAT1	Reverse	CCAGTTTTCCGAGAACCAAA
chr16:85587072-85589839	Forward	GGTCGGAGAGAGAGCCTTTT
chr16:85587072-85589839	Reverse	TTCCATAGTGGCCAGAGAGG
chr11:70995192-70998931	Forward	TGCTGTGCACTTGTGTGTTC
chr11:70995192-70998931	Reverse	CCCCTGCAGTCACCTTTTAT
chr7:129244613-129251530	Forward	TGAAGGAGTAGGAGGGCAAA
chr7:129244613-129251530	Reverse	AAACACGTGGCTTTCCTGTT
chr20:292339-305552	Forward	GACCAGCATCTCAACTGCAA
chr20:292339-305552	Reverse	TCTTCCCCTCCTCTTCCAAT

chr5:172721232-172730376	Forward	CGCAACTGTAGCACATGGAT
chr5:172721232-172730376	Reverse	TGCCTTCGCAAGAGTCTTCT
chr5:177485281-177505466	Forward	GTAGGGCCTGTGCCTTTTCT
chr5:177485281-177505466	Reverse	TGGTTGAGACGTCTGTGAGC
chr12:8116260-8123253	Forward	AAAAAGGAGAGCGCGTACTG
chr12:8116260-8123253	Reverse	GCCCGAACTACAATTCTTGG
chr16:85479122-85498916	Forward	ACCACCTAAGCCGACATCTG
chr16:85479122-85498916	Reverse	CTCTGTGTGCACCTCTCTGC
TCONS_00021778	Forward	GCCTAGGTTCCACCCTTTTC
TCONS_00021778	Reverse	CGGGAACAGAAAGAAAGCAG
TCONS_00010677	Forward	GGCTGGAGTACAGTGGCATAG
TCONS_00010677	Reverse	GCGAAACACGGTGTCTACAA
TCONS_00026155	Forward	GCAGCAGAACTGATGGTGAA
TCONS_00026155	Reverse	AGTGCCACAACCCCTGATAG
TCONS_l2_00017632	Forward	ACTCTTGTTGCAGGCGAAGT
TCONS_l2_00017632	Reverse	ACGGGGAGAGGTTGATCC
TCONS_00024694	Forward	GCTAGAGGGCAGCTACAGGA
TCONS_00024695	Reverse	ATTTGGTAGGAGGGCAGCTT
TCONS_00018859	Forward	ATGGGCAAGAACAGGATGTC
TCONS_00018859	Reverse	GGACCAGGGTTGACTAATGC
TCONS_00012630	Forward	TGTTCAGGGTGTCTCTGCTG
TCONS_00012630	Reverse	CACTGCCTCTGGTGATCTGA
TCON_00019373	Forward	GCTTCACGACGTGTTGCTAA
TCON_00019373	Reverse	TCAGACTGGGCTTTGTAGGG
TCONS_00019713	Forward	AAGGCTCCTGACTGTGGAGA
TCONS_00019713	Reverse	GAGGAGTCTCACCTGCAAGC
TCONS_l2_00016774	Forward	TGTCAAGCCCTGCTCTTTCT
TCONS_l2_00016774	Reverse	GTTCCTTCTGCTGCAACTCC
TCONS_00020202	Forward	CAGCCTATGAGCTCCTCCAG
TCONS_00020202	Reverse	GAGACTGCTCTCCACGAACC
TCONS_00000020	Forward	GCCGTTTTACATTGGCTGTT
TCONS_00000020	Reverse	ATTTGGAAGGTCGTCCACTG
TCONS_l2_00027397	Forward	AGGAAGCAACTTTCCAGCAA
TCONS_l2_00027397	Reverse	CTCTGGGTCAGGAATTTGGA

Supplementary table 5. Sequences of HIF siRNAs

siRNA	Sense	Antisense
HIF-1	CUGAUGACCAGCAACUUGAtt	UCAAGUUGCUGGUCAUCAGtt
HIF-2	CAGCAUCUUUGAUAGCAGUtt	ACUGCUAUCAAAGAUGCUGtt

Supplementary table 6. List of TaqMan® assays.

TaqMan assay	Assay ID
has-miR-612	1579
has-miR-184	485
has-miR-210	512
has-miR-100	437
has-miR-1	2222
RNU43	1095
SNO38	Hs03200042_s0
SNO117	Hs03299042_s1
SNO60	Hs03309311_s1
SNO21	Hs04274615_s1
SNO66	Hs03297862_s1
TCF25	Hs00248217_m1
C15orf42	Hs00738302_m1
DHX30	Hs00209180_m1
ZNF320	Hs02598857_g1
RBAK	Hs00372722_m1

Chapter 9: Bibliography

9. Bibliography

(TCGA), C. G. A. 2012. Comprehensive molecular portraits of human breast tumours. *Nature,* 490**,** 61-70.

AIRLEY, R. E., LONCASTER, J., RALEIGH, J. A., HARRIS, A. L., DAVIDSON, S. E., HUNTER, R. D., WEST, C. M. & STRATFORD, I. J. 2003. GLUT-1 and CAIX as intrinsic markers of hypoxia in carcinoma of the cervix: relationship to pimonidazole binding. *Int J Cancer,* 104**,** 85-91.

ALEXANDROV, L. B., NIK-ZAINAL, S., WEDGE, D. C., APARICIO, S. A., BEHJATI, S., BIANKIN, A. V., BIGNELL, G. R., BOLLI, N., BORG, A., BORRESEN-DALE, A. L., BOYAULT, S., BURKHARDT, B., BUTLER, A. P., CALDAS, C., DAVIES, H. R., DESMEDT, C., EILS, R., EYFJORD, J. E., FOEKENS, J. A., GREAVES, M., HOSODA, F., HUTTER, B., ILICIC, T., IMBEAUD, S., IMIELINSKI, M., JAGER, N., JONES, D. T., JONES, D., KNAPPSKOG, S., KOOL, M., LAKHANI, S. R., LOPEZ-OTIN, C., MARTIN, S., MUNSHI, N. C., NAKAMURA, H., NORTHCOTT, P. A., PAJIC, M., PAPAEMMANUIL, E., PARADISO, A., PEARSON, J. V., PUENTE, X. S., RAINE, K., RAMAKRISHNA, M., RICHARDSON, A. L., RICHTER, J., ROSENSTIEL, P., SCHLESNER, M., SCHUMACHER, T. N., SPAN, P. N., TEAGUE, J. W., TOTOKI, Y., TUTT, A. N., VALDES-MAS, R., VAN BUUREN, M. M., VAN 'T VEER, L., VINCENT-SALOMON, A., WADDELL, N., YATES, L. R., AUSTRALIAN PANCREATIC CANCER GENOME, I., CONSORTIUM, I. B. C., CONSORTIUM, I. M.-S., PEDBRAIN, I., ZUCMAN-ROSSI, J., FUTREAL, P. A., MCDERMOTT, U., LICHTER, P., MEYERSON, M., GRIMMOND, S. M., SIEBERT, R., CAMPO, E., SHIBATA, T., PFISTER, S. M., CAMPBELL, P. J. & STRATTON, M. R. 2013. Signatures of mutational processes in human cancer. *Nature,* 500**,** 415-21.

ANDRAU, J. C. & WERNER, M. 2001. B ''-associated factor(s) involved in RNA polymerase III preinitiation complex formation and start-site selection. *European Journal of Biochemistry,* 268**,** 5167-5175.

ANKO, M. L. & NEUGEBAUER, K. M. 2010. Long noncoding RNAs add another layer to pre-mRNA splicing regulation. *Mol Cell,* 39**,** 833-4.

ARAVIN, A. A., SACHIDANANDAM, R., GIRARD, A., FEJES-TOTH, K. & HANNON, G. J. 2007. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science,* 316**,** 744-7.

ATKURI, K. R., HERZENBERG, L. A., NIEMI, A. K., COWAN, T. & HERZENBERG, L. A. 2007. Importance of culturing primary lymphocytes at physiological oxygen levels. *Proc Natl Acad Sci U S A,* 104**,** 4547-52.

BACHTIARY, B., SCHINDL, M., POTTER, R., DREIER, B., KNOCKE, T. H., HAINFELLNER, J. A., HORVAT, R. & BIRNER, P. 2003. Overexpression of hypoxia-inducible factor 1alpha indicates diminished response to radiotherapy and unfavorable prognosis in patients receiving radical radiotherapy for cervical cancer. *Clin Cancer Res,* 9**,** 2234-40.

BALDEWIJNS, M. M., VAN VLODROP, I. J., VERMEULEN, P. B., SOETEKOUW, P. M., VAN ENGELAND, M. & DE BRUINE, A. P. 2010. VHL and HIF signalling in renal cell carcinogenesis. *J Pathol,* 221**,** 125-38.

BARSKI, A., CUDDAPAH, S., CUI, K., ROH, T. Y., SCHONES, D. E., WANG, Z., WEI, G., CHEPELEV, I. & ZHAO, K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell,* 129**,** 823-37.

BATISTA, P. J. & CHANG, H. Y. 2013. Long noncoding RNAs: cellular address codes in development and disease. *Cell,* 152**,** 1298-307.

BEDFORD, D. C., KASPER, L. H., FUKUYAMA, T. & BRINDLE, P. K. 2010. Target gene context influences the transcriptional requirement for the KAT3 family of CBP and p300 histone acetyltransferases. *Epigenetics,* 5**,** 9-15.

BEKETIC-ORESKOVIC, L., OZRETIC, P., RABBANI, Z. N., JACKSON, I. L., SARCEVIC, B., LEVANAT, S., MARIC, P., BABIC, I. & VUJASKOVIC, Z. 2011. Prognostic significance of carbonic anhydrase IX (CA-IX), endoglin (CD105) and 8-hydroxy-2'-deoxyguanosine (8-OHdG) in breast cancer patients. *Pathol Oncol Res,* 17**,** 593-603.

BEN-ZVI, M., AMARIGLIO, N., PARET, G. & NEVO-CASPI, Y. 2013. F11R expression upon hypoxia is regulated by RNA editing. *PLoS One,* 8**,** e77702.

BEN LASSOUED, A., BEAUFILS, N., DALES, J. P. & GABERT, J. 2013. Hypoxia-inducible factor-1alpha as prognostic marker. *Expert Opin Med Diagn,* 7**,** 53-70.

BENCOKOVA, Z., KAUFMANN, M. R., PIRES, I. M., LECANE, P. S., GIACCIA, A. J. & HAMMOND, E. M. 2009. ATM activation and signaling under hypoxic conditions. *Mol Cell Biol,* 29**,** 526-37.

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**,** 4587-602.

BERTONE, P., STOLC, V., ROYCE, T. E., ROZOWSKY, J. S., URBAN, A. E., ZHU, X., RINN, J. L., TONGPRASIT, W., SAMANTA, M., WEISSMAN, S., GERSTEIN, M. & SNYDER, M. 2004. Global identification of human transcribed sequences with genome tiling arrays. *Science,* 306**,** 2242-6.

BERTOUT, J. A., PATEL, S. A. & SIMON, M. C. 2008. The impact of O2 availability on human cancer. *Nat Rev Cancer,* 8**,** 967-75.

BERTOZZI, D., IURLARO, R., SORDET, O., MARINELLO, J., ZAFFARONI, N. & CAPRANICO, G. 2011. Characterization of novel antisense HIF-1alpha transcripts in human cancers. *Cell Cycle,* 10**,** 3189-97.

BETOF, A. S., RABBANI, Z. N., HARDEE, M. E., KIM, S. J., BROADWATER, G., BENTLEY, R. C., SNYDER, S. A., VUJASKOVIC, Z., OOSTERWIJK, E., HARRIS, L. N., HORTON, J. K., DEWHIRST, M. W. & BLACKWELL, K. L. 2012. Carbonic anhydrase IX is a predictive marker of doxorubicin resistance in early-stage breast cancer independent of HER2 and TOP2A amplification. *Br J Cancer,* 106**,** 916-22.

BEYER, S., KRISTENSEN, M. M., JENSEN, K. S., JOHANSEN, J. V. & STALLER, P. 2008. The histone demethylases JMJD1A and JMJD2B are transcriptional targets of hypoxia-inducible factor HIF. *J Biol Chem,* 283**,** 36542-52.

BLICK, C., RAMACHANDRAN, A., WIGFIELD, S., MCCORMICK, R., JUBB, A., BUFFA, F. M., TURLEY, H., KNOWLES, M. A., CRANSTON, D., CATTO, J. & HARRIS, A. L. 2013. Hypoxia regulates FGFR3 expression via HIF-1alpha and miR-100 and contributes to cell survival in non-muscle invasive bladder cancer. *Br J Cancer,* 109**,** 50-9.

BLOUW, B., SONG, H., TIHAN, T., BOSZE, J., FERRARA, N., GERBER, H. P., JOHNSON, R. S. & BERGERS, G. 2003. The hypoxic response of tumors is dependent on their microenvironment. *Cancer Cell,* 4**,** 133-46.

BOEHM, J. S. & HAHN, W. C. 2011. Towards systematic functional characterization of cancer genomes. *Nat Rev Genet,* 12**,** 487-98.

BOGUSKI, M. S., TOLSTOSHEV, C. M. & BASSETT, D. E., JR. 1994. Gene discovery in dbEST. *Science,* 265**,** 1993-4.

BOND, C. S. & FOX, A. H. 2009. Paraspeckles: nuclear bodies built on long noncoding RNA. *J Cell Biol,* 186**,** 637-44.

BOS, R., VAN DER GROEP, P., GREIJER, A. E., SHVARTS, A., MEIJER, S., PINEDO, H. M., SEMENZA, G. L., VAN DIEST, P. J. & VAN DER WALL, E. 2003. Levels of hypoxia-inducible factor-1alpha independently predict prognosis in patients with lymph node negative breast carcinoma. *Cancer,* 97**,** 1573-81.

BOS, R., VAN DIEST, P. J., VAN DER GROEP, P., SHVARTS, A., GREIJER, A. E. & VAN DER WALL, E. 2004. Expression of hypoxia-inducible factor-1alpha and cell cycle proteins in invasive breast cancer are estrogen receptor related. *Breast Cancer Res,* 6**,** R450-9.

BRENNAN, D. J., JIRSTROM, K., KRONBLAD, A., MILLIKAN, R. C., LANDBERG, G., DUFFY, M. J., RYDEN, L., GALLAGHER, W. M. & O'BRIEN, S. L. 2006. CA IX is an independent prognostic marker in premenopausal breast cancer patients with one to three positive lymph nodes and a putative marker of radiation resistance. *Clin Cancer Res,* 12**,** 6421-31.

BRENNAN, K., MCSHERRY, E. A., HUDSON, L., KAY, E. W., HILL, A. D., YOUNG, L. S. & HOPKINS, A. M. 2013. Junctional adhesion molecule-A is co-expressed with HER2 in breast tumors and acts as a novel regulator of HER2 protein degradation and signaling. *Oncogene,* 32**,** 2799-804.

BRENNECKE, J., ARAVIN, A. A., STARK, A., DUS, M., KELLIS, M., SACHIDANANDAM, R. & HANNON, G. J. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila. *Cell,* 128**,** 1089-103.

BRENNICKE, A., MARCHFELDER, A. & BINDER, S. 1999. RNA editing. *FEMS Microbiol Rev,* 23**,** 297-316.

BRISTOW, R. G. & HILL, R. P. 2008. Hypoxia and metabolism. Hypoxia, DNA repair and genetic instability. *Nat Rev Cancer,* 8**,** 180-92.

BROWN, J. A., VALENSTEIN, M. L., YARIO, T. A., TYCOWSKI, K. T. & STEITZ, J. A. 2012. Formation of triple-helical structures by the 3'-end sequences of MALAT1 and MENbeta noncoding RNAs. *Proc Natl Acad Sci U S A,* 109**,** 19202-7.

BROWN, J. M. & GIACCIA, A. J. 1998. The unique physiology of solid tumors: opportunities (and problems) for cancer therapy. *Cancer Res,* 58**,** 1408-16.

BROWN, J. M. & WILSON, W. R. 2004. Exploiting tumour hypoxia in cancer treatment. *Nat Rev Cancer,* 4**,** 437-47.

BRUGAROLAS, J. & KAELIN, W. G., JR. 2004. Dysregulation of HIF and VEGF is a unifying feature of the familial hamartoma syndromes. *Cancer Cell,* 6**,** 7-10.

BRUICK, R. K. & MCKNIGHT, S. L. 2001. A conserved family of prolyl-4-hydroxylases that modify HIF. *Science,* 294**,** 1337-40.

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**,** 4087-96.

BUFFA, F. M., HARRIS, A. L., WEST, C. M. & MILLER, C. J. 2010. Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *Br J Cancer,* 102**,** 428-35.

CADOO, K. A., FORNIER, M. N. & MORRIS, P. G. 2013. Biological subtypes of breast cancer: current concepts and implications for recurrence patterns. *Q J Nucl Med Mol Imaging,* 57**,** 312-21.

CAIRNS, R. A., HARRIS, I. S. & MAK, T. W. 2011. Regulation of cancer cell metabolism. *Nat Rev Cancer,* 11**,** 85-95.

CAMPEAU, P. M., FOULKES, W. D. & TISCHKOWITZ, M. D. 2008. Hereditary breast cancer: new genetic developments, new therapeutic avenues. *Hum Genet,* 124**,** 31-42.

CAMPS, C., BUFFA, F. M., COLELLA, S., MOORE, J., SOTIRIOU, C., SHELDON, H., HARRIS, A. L., GLEADLE, J. M. & RAGOUSSIS, J. 2008. hsa-miR-210 Is induced by hypoxia and is an independent prognostic factor in breast cancer. *Clin Cancer Res,* 14**,** 1340-8.

CAMPS, C., SAINI, H. K., MOLE, D. R., CHOUDHRY, H., RECZKO, M., GUERRA-ASSUNCAO, J. A., TIAN, Y. M., BUFFA, F. M., HARRIS, A. L., HATZIGEORGIOU, A. G., ENRIGHT, A. J. & RAGOUSSIS, J. 2014. Integrated analysis of microRNA and mRNA expression and association with HIF binding reveals the complexity of microRNA expression regulation under hypoxia. *Mol Cancer,* 13**,** 28.

CANCER GENOME ATLAS, N. 2012. Comprehensive molecular portraits of human breast tumours. *Nature,* 490**,** 61-70.

CARNEIRO, M. O., RUSS, C., ROSS, M. G., GABRIEL, S. B., NUSBAUM, C. & DEPRISTO, M. A. 2012. Pacific biosciences sequencing technology for genotyping and variation discovery in human data. *BMC Genomics,* 13**,** 375.

CARNINCI, P., KASUKAWA, T., KATAYAMA, S., GOUGH, J., FRITH, M. C., MAEDA, N., OYAMA, R., RAVASI, T., LENHARD, B., WELLS, C., KODZIUS, R., SHIMOKAWA, K., BAJIC, V. B., BRENNER, S. E., BATALOV, S., FORREST, A. R., ZAVOLAN, M., DAVIS, M. J., WILMING, L. G., AIDINIS, V., ALLEN, J. E., AMBESI-IMPIOMBATO, A., APWEILER, R., ATURALIYA, R. N., BAILEY, T. L., BANSAL, M., BAXTER, L., BEISEL, K. W., BERSANO, T., BONO, H., CHALK, A. M., CHIU, K. P.,

CHOUDHARY, V., CHRISTOFFELS, A., CLUTTERBUCK, D. R., CROWE, M. L., DALLA, E., DALRYMPLE, B. P., DE BONO, B., DELLA GATTA, G., DI BERNARDO, D., DOWN, T., ENGSTROM, P., FAGIOLINI, M., FAULKNER, G., FLETCHER, C. F., FUKUSHIMA, T., FURUNO, M., FUTAKI, S., GARIBOLDI, M., GEORGII-HEMMING, P., GINGERAS, T. R., GOJOBORI, T., GREEN, R. E., GUSTINCICH, S., HARBERS, M., HAYASHI, Y., HENSCH, T. K., HIROKAWA, N., HILL, D., HUMINIECKI, L., IACONO, M., IKEO, K., IWAMA, A., ISHIKAWA, T., JAKT, M., KANAPIN, A., KATOH, M., KAWASAWA, Y., KELSO, J., KITAMURA, H., KITANO, H., KOLLIAS, G., KRISHNAN, S. P., KRUGER, A., KUMMERFELD, S. K., KUROCHKIN, I. V., LAREAU, L. F., LAZAREVIC, D., LIPOVICH, L., LIU, J., LIUNI, S., MCWILLIAM, S., MADAN BABU, M., MADERA, M., MARCHIONNI, L., MATSUDA, H., MATSUZAWA, S., MIKI, H., MIGNONE, F., MIYAKE, S., MORRIS, K., MOTTAGUI-TABAR, S., MULDER, N., NAKANO, N., NAKAUCHI, H., NG, P., NILSSON, R., NISHIGUCHI, S., NISHIKAWA, S., et al. 2005. The transcriptional landscape of the mammalian genome. *Science,* 309**,** 1559-63.

CAYRE, A., ROSSIGNOL, F., CLOTTES, E. & PENAULT-LLORCA, F. 2003. aHIF but not HIF-1alpha transcript is a poor prognostic marker in human breast cancer. *Breast Cancer Res,* 5**,** R223-30.

CHAN, N., PIRES, I. M., BENCOKOVA, Z., COACKLEY, C., LUOTO, K. R., BHOGAL, N., LAKSHMAN, M., GOTTIPATI, P., OLIVER, F. J., HELLEDAY, T., HAMMOND, E. M. & BRISTOW, R. G. 2010. Contextual synthetic lethality of cancer cell kill based on the tumor microenvironment. *Cancer Res,* 70**,** 8045-54.

CHAN, S. Y. & LOSCALZO, J. 2010. MicroRNA-210: a unique and pleiotropic hypoxamir. *Cell Cycle,* 9**,** 1072-83.

CHAN, Y. C., KHANNA, S., ROY, S. & SEN, C. K. 2011. miR-200b targets Ets-1 and is down-regulated by hypoxia to induce angiogenic response of endothelial cells. *J Biol Chem,* 286**,** 2047-56.

CHANG, K. H., MESTDAGH, P., VANDESOMPELE, J., KERIN, M. J. & MILLER, N. 2010. MicroRNA expression profiling to identify and validate reference genes for relative quantification in colorectal cancer. *BMC Cancer,* 10**,** 173.

CHEN, H. H., SU, W. C., LIN, P. W., GUO, H. R. & LEE, W. Y. 2007. Hypoxia-inducible factor-1alpha correlates with MET and metastasis in node-negative breast cancer. *Breast Cancer Res Treat,* 103**,** 167-75.

CHEN, L., LI, Y., LIN, C. H., CHAN, T. H., CHOW, R. K., SONG, Y., LIU, M., YUAN, Y. F., FU, L., KONG, K. L., QI, L., LI, Y., ZHANG, N., TONG, A. H., KWONG, D. L., MAN, K., LO, C. M., LOK, S., TENEN, D. G. & GUAN, X. Y. 2013. Recoding RNA editing of AZIN1 predisposes to hepatocellular carcinoma. *Nat Med,* 19**,** 209-16.

CHEN, L. L. & CARMICHAEL, G. G. 2009. Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Mol Cell,* 35**,** 467-78.

CHEN, W. D. & ZHU, X. F. 2013. Small nucleolar RNAs (snoRNAs) as potential non-invasive biomarkers for early cancer detection. *Chin J Cancer,* 32**,** 99-101.

CHEN, Z. & DUAN, X. 2011. Ribosomal RNA depletion for massively parallel bacterial RNA-sequencing applications. *Methods Mol Biol,* 733**,** 93-103.

CHENG, J., DENG, H., XIAO, B., ZHOU, H., ZHOU, F., SHEN, Z. & GUO, J. 2012. piR-823, a novel non-coding small RNA, demonstrates in vitro and in vivo tumor suppressive activity in human gastric cancer cells. *Cancer Lett,* 315**,** 12-7.

CHENG, J., GUO, J. M., XIAO, B. X., MIAO, Y., JIANG, Z., ZHOU, H. & LI, Q. N. 2011. piRNA, the new non-coding RNA, is aberrantly expressed in human cancer cells. *Clin Chim Acta,* 412**,** 1621-5.

CHI, J. T., WANG, Z., NUYTEN, D. S., RODRIGUEZ, E. H., SCHANER, M. E., SALIM, A., WANG, Y., KRISTENSEN, G. B., HELLAND, A., BORRESEN-DALE, A. L., GIACCIA, A., LONGAKER, M. T., HASTIE, T., YANG, G. P., VAN DE VIJVER, M. J. & BROWN, P. O. 2006. Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. *PLoS Med,* 3**,** e47.

CHIA, S. K., WYKOFF, C. C., WATSON, P. H., HAN, C., LEEK, R. D., PASTOREK, J., GATTER, K. C., RATCLIFFE, P. & HARRIS, A. L. 2001. Prognostic significance of a novel hypoxia-regulated marker, carbonic anhydrase IX, in invasive breast carcinoma. *J Clin Oncol,* 19**,** 3660-8.

CHOUDHRY, H., SCHODEL, J., OIKONOMOPOULOS, S., CAMPS, C., GRAMPP, S., HARRIS, A. L., RATCLIFFE, P. J., RAGOUSSIS, J. & MOLE, D. R. 2014. Extensive regulation of the non-coding transcriptome by hypoxia: role of HIF in releasing paused RNApol2. *EMBO Rep,* 15**,** 70-6.

CHU, L., SU, M. Y., MAGGI, L. B., JR., LU, L., MULLINS, C., CROSBY, S., HUANG, G., CHNG, W. J., VIJ, R. & TOMASSON, M. H. 2012. Multiple myeloma-associated chromosomal translocation activates orphan snoRNA ACA11 to suppress oxidative stress. *J Clin Invest,* 122**,** 2793-806.

CLARK, J., LU, Y. J., SIDHAR, S. K., PARKER, C., GILL, S., SMEDLEY, D., HAMOUDI, R., LINEHAN, W. M., SHIPLEY, J. & COOPER, C. S. 1997. Fusion of splicing factor genes PSF and NonO (p54nrb) to the TFE3 gene in papillary renal cell carcinoma. *Oncogene,* 15**,** 2233-9.

CLEMSON, C. M., HUTCHINSON, J. N., SARA, S. A., ENSMINGER, A. W., FOX, A. H., CHESS, A. & LAWRENCE, J. B. 2009. An architectural role for a nuclear noncoding RNA: NEAT1 RNA is essential for the structure of paraspeckles. *Mol Cell,* 33**,** 717-26.

CONESA-ZAMORA, P., GARCIA-SOLANO, J., GARCIA-GARCIA, F., TURPIN MDEL, C., TRUJILLO-SANTOS, J., TORRES-MORENO, D., OVIEDO-RAMIREZ, I., CARBONELL-MUNOZ, R., MUNOZ-DELGADO, E., RODRIGUEZ-BRAUN, E., CONESA, A. & PEREZ-GUILLERMO, M. 2013. Expression profiling shows differential molecular pathways and provides potential new diagnostic biomarkers for colorectal serrated adenocarcinoma. *Int J Cancer,* 132**,** 297-307.

CONSORTIUM, E. P., BERNSTEIN, B. E., BIRNEY, E., DUNHAM, I., GREEN, E. D., GUNTER, C. & SNYDER, M. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature,* 489**,** 57-74.

CORNFIELD, D. B., PALAZZO, J. P., SCHWARTZ, G. F., GOONEWARDENE, S. A., KOVATICH, A. J., CHERVONEVA, I., HYSLOP, T. & SCHWARTING, R. 2004. The prognostic significance of multiple morphologic features and biologic markers in ductal carcinoma in situ of the breast: a study of a large cohort of patients treated with surgery alone. *Cancer,* 100**,** 2317-27.

COSTA, V., ANGELINI, C., D'APICE, L., MUTARELLI, M., CASAMASSIMI, A., SOMMESE, L., GALLO, M. A., APRILE, M., ESPOSITO, R., LEONE, L., DONIZETTI, A., CRISPI, S., RIENZO, M., SARUBBI, B., CALABRO, R., PICARDI, M., SALVATORE, P., INFANTE, T., DE BERARDINIS, P., NAPOLI, C. & CICCODICOLA, A. 2011. Massive-scale RNA-Seq analysis of non ribosomal transcriptome in human trisomy 21. *PLoS One,* 6**,** e18493.

COSTA, V., ANGELINI, C., DE FEIS, I. & CICCODICOLA, A. 2010. Uncovering the complexity of transcriptomes with RNA-Seq. *J Biomed Biotechnol,* 2010**,** 853916.

COVELLO, K. L., KEHLER, J., YU, H., GORDAN, J. D., ARSHAM, A. M., HU, C. J., LABOSKY, P. A., SIMON, M. C. & KEITH, B. 2006. HIF-2alpha regulates Oct-4: effects of hypoxia on stem cell function, embryonic development, and tumor growth. *Genes Dev,* 20**,** 557-70.

COWIN, P. A., ANGLESIO, M., ETEMADMOGHADAM, D. & BOWTELL, D. D. 2010. Profiling the cancer genome. *Annu Rev Genomics Hum Genet,* 11**,** 133-59.

CRAIG, D. W., O'SHAUGHNESSY, J. A., KIEFER, J. A., ALDRICH, J., SINARI, S., MOSES, T. M., WONG, S., DINH, J., CHRISTOFORIDES, A., BLUM, J. L., AITELLI, C. L., OSBORNE, C. R., IZATT, T., KURDOGLU, A., BAKER, A., KOEMAN, J., BARBACIORU, C., SAKARYA, O., DE LA VEGA, F. M., SIDDIQUI, A., HOANG, L., BILLINGS, P. R., SALHIA, B., TOLCHER, A. W., TRENT, J. M., MOUSSES, S., VON HOFF, D. & CARPTEN, J. D. 2013. Genome and transcriptome sequencing in prospective metastatic triple-negative breast cancer uncovers therapeutic vulnerabilities. *Mol Cancer Ther,* 12**,** 104-16.

CRICK, F. 1970. Central dogma of molecular biology. *Nature,* 227**,** 561-3.

CROSBY, M. E., KULSHRESHTHA, R., IVAN, M. & GLAZER, P. M. 2009. MicroRNA regulation of DNA repair gene expression in hypoxic stress. *Cancer Res,* 69**,** 1221-9.

CUI, P., LIN, Q., DING, F., XIN, C., GONG, W., ZHANG, L., GENG, J., ZHANG, B., YU, X., YANG, J., HU, S. & YU, J. 2010. A comparison between ribo-minus RNA-sequencing and polyA-selected RNA-sequencing. *Genomics,* 96**,** 259-65.

CURTIS, C., SHAH, S. P., CHIN, S. F., TURASHVILI, G., RUEDA, O. M., DUNNING, M. J., SPEED, D., LYNCH, A. G., SAMARAJIWA, S., YUAN, Y., GRAF, S., HA, G., HAFFARI, G., BASHASHATI, A., RUSSELL, R., MCKINNEY, S., GROUP, M., LANGEROD, A., GREEN, A., PROVENZANO, E., WISHART, G., PINDER, S., WATSON, P., MARKOWETZ, F., MURPHY, L., ELLIS, I., PURUSHOTHAM, A., BORRESEN-DALE, A. L., BRENTON, J. D., TAVARE, S., CALDAS, C. & APARICIO, S. 2012. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature,* 486**,** 346-52.

DALES, J. P., GARCIA, S., MEUNIER-CARPENTIER, S., ANDRAC-MEYER, L., HADDAD, O., LAVAUT, M. N., ALLASIA, C., BONNIER, P. & CHARPIN, C. 2005. Overexpression of hypoxia-inducible factor HIF-1alpha predicts early relapse in breast cancer: retrospective study in a series of 745 patients. *Int J Cancer,* 116**,** 734-9.

DANDACHI, N., DIETZE, O. & HAUSER-KRONBERGER, C. 2002. Chromogenic in situ hybridization: a novel approach to a practical and sensitive method for the detection of HER2 oncogene in archival human breast carcinoma. *Lab Invest,* 82**,** 1007-14.

DAWSON, M. A. & KOUZARIDES, T. 2012. Cancer epigenetics: from mechanism to therapy. *Cell,* 150**,** 12-27.

DE SANTA, F., BAROZZI, I., MIETTON, F., GHISLETTI, S., POLLETTI, S., TUSI, B. K., MULLER, H., RAGOUSSIS, J., WEI, C. L. & NATOLI, G. 2010. A large fraction of extragenic RNA pol II transcription sites overlap enhancers. *PLoS Biol,* 8**,** e1000384.

DEKERVEL, J., HOMPES, D., VAN MALENSTEIN, H., POPOVIC, D., SAGAERT, X., DE MOOR, B., VAN CUTSEM, E., D'HOORE, A., VERSLYPE, C. & VAN PELT, J. 2014. Hypoxia driven gene expression is an independent prognostic factor in stage II and III colon cancer patients. *Clin Cancer Res*.

DENGLER, V. L., GALBRAITH, M. D. & ESPINOSA, J. M. 2013. Transcriptional regulation by hypoxia inducible factors. *Crit Rev Biochem Mol Biol*.

DENKO, N., SCHINDLER, C., KOONG, A., LADEROUTE, K., GREEN, C. & GIACCIA, A. 2000. Epigenetic regulation of gene expression in cervical cancer cells by the tumor microenvironment. *Clin Cancer Res,* 6**,** 480-7.

DENKO, N. C., FONTANA, L. A., HUDSON, K. M., SUTPHIN, P. D., RAYCHAUDHURI, S., ALTMAN, R. & GIACCIA, A. J. 2003. Investigating hypoxic tumor physiology through gene expression patterns. *Oncogene,* 22**,** 5907-14.

DERRIEN, T., JOHNSON, R., BUSSOTTI, G., TANZER, A., DJEBALI, S., TILGNER, H., GUERNEC, G., MARTIN, D., MERKEL, A., KNOWLES, D. G., LAGARDE, J., VEERAVALLI, L., RUAN, X., RUAN, Y., LASSMANN, T., CARNINCI, P., BROWN, J. B., LIPOVICH, L., GONZALEZ, J. M., THOMAS, M., DAVIS, C. A., SHIEKHATTAR, R., GINGERAS, T. R., HUBBARD, T. J., NOTREDAME, C., HARROW, J. & GUIGO, R. 2012. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res,* 22**,** 1775-89.

DERYCKE, M. S., PAMBUCCIAN, S. E., GILKS, C. B., KALLOGER, S. E., GHIDOUCHE, A., LOPEZ, M., BLISS, R. L., GELLER, M. A., ARGENTA, P. A., HARRINGTON, K. M. & SKUBITZ, A. P. 2010. Nectin 4 overexpression in ovarian cancer tissues and serum: potential role as a serum biomarker. *Am J Clin Pathol,* 134**,** 835-45.

DIECI, G., FIORINO, G., CASTELNUOVO, M., TEICHMANN, M. & PAGANO, A. 2007. The expanding RNA polymerase III transcriptome. *Trends Genet,* 23**,** 614-22.

DIOUM, E. M., CHEN, R., ALEXANDER, M. S., ZHANG, Q., HOGG, R. T., GERARD, R. D. & GARCIA, J. A. 2009. Regulation of hypoxia-inducible factor 2alpha signaling by the stress-responsive deacetylase sirtuin 1. *Science,* 324**,** 1289-93.

DJEBALI, S., DAVIS, C. A., MERKEL, A., DOBIN, A., LASSMANN, T., MORTAZAVI, A., TANZER, A., LAGARDE, J., LIN, W., SCHLESINGER, F., XUE, C., MARINOV, G. K., KHATUN, J., WILLIAMS, B. A., ZALESKI, C., ROZOWSKY, J., RODER, M., KOKOCINSKI, F., ABDELHAMID, R. F., ALIOTO, T.,

ANTOSHECHKIN, I., BAER, M. T., BAR, N. S., BATUT, P., BELL, K., BELL, I., CHAKRABORTTY, S., CHEN, X., CHRAST, J., CURADO, J., DERRIEN, T., DRENKOW, J., DUMAIS, E., DUMAIS, J., DUTTAGUPTA, R., FALCONNET, E., FASTUCA, M., FEJES-TOTH, K., FERREIRA, P., FOISSAC, S., FULLWOOD, M. J., GAO, H., GONZALEZ, D., GORDON, A., GUNAWARDENA, H., HOWALD, C., JHA, S., JOHNSON, R., KAPRANOV, P., KING, B., KINGSWOOD, C., LUO, O. J., PARK, E., PERSAUD, K., PREALL, J. B., RIBECA, P., RISK, B., ROBYR, D., SAMMETH, M., SCHAFFER, L., SEE, L. H., SHAHAB, A., SKANCKE, J., SUZUKI, A. M., TAKAHASHI, H., TILGNER, H., TROUT, D., WALTERS, N., WANG, H., WROBEL, J., YU, Y., RUAN, X., HAYASHIZAKI, Y., HARROW, J., GERSTEIN, M., HUBBARD, T., REYMOND, A., ANTONARAKIS, S. E., HANNON, G., GIDDINGS, M. C., RUAN, Y., WOLD, B., CARNINCI, P., GUIGO, R. & GINGERAS, T. R. 2012. Landscape of transcription in human cells. *Nature,* 489**,** 101-8.

DONG, X. Y., GUO, P., BOYD, J., SUN, X. D., LI, Q. N., ZHOU, W. & DONG, J. T. 2009. Implication of snoRNA U50 in human breast cancer. *Journal of Genetics and Genomics,* 36**,** 447-454.

DONKER, R. B., MOUILLET, J. F., NELSON, D. M. & SADOVSKY, Y. 2007. The expression of Argonaute2 and related microRNA biogenesis proteins in normal and hypoxic trophoblasts. *Mol Hum Reprod,* 13**,** 273-9.

DU, Z., FEI, T., VERHAAK, R. G., SU, Z., ZHANG, Y., BROWN, M., CHEN, Y. & LIU, X. S. 2013. Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. *Nat Struct Mol Biol,* 20**,** 908-13.

DUHOUX, F. P., AUGER, N., DE WILDE, S., WITTNEBEL, S., AMEYE, G., BAHLOULA, K., VAN DEN BERG, C., LIBOUTON, J. M., SAUSSOY, P., GRAND, F. H., DEMOULIN, J. B. & POIREL, H. A. 2011. The t(1;9)(p34;q34) fusing ABL1 with SFPQ, a pre-mRNA processing gene, is recurrent in acute lymphoblastic leukemias. *Leukemia Research,* 35**,** E114-E117.

EBOS, J. M., LEE, C. R. & KERBEL, R. S. 2009. Tumor and host-mediated pathways of resistance and disease progression in response to antiangiogenic therapy. *Clin Cancer Res,* 15**,** 5020-5.

EDDY, S. R. 2001. Non-coding RNA genes and the modern RNA world. *Nat Rev Genet,* 2**,** 919-29.

ELLIS, M. J., DING, L., SHEN, D., LUO, J., SUMAN, V. J., WALLIS, J. W., VAN TINE, B. A., HOOG, J., GOIFFON, R. J., GOLDSTEIN, T. C., NG, S., LIN, L., CROWDER, R., SNIDER, J., BALLMAN, K., WEBER, J., CHEN, K., KOBOLDT, D. C., KANDOTH, C., SCHIERDING, W. S., MCMICHAEL, J. F., MILLER, C. A., LU, C., HARRIS, C. C., MCLELLAN, M. D., WENDL, M. C., DESCHRYVER, K., ALLRED, D. C., ESSERMAN, L., UNZEITIG, G., MARGENTHALER, J., BABIERA, G. V., MARCOM, P. K., GUENTHER, J. M., LEITCH, M., HUNT, K., OLSON, J., TAO, Y., MAHER, C. A., FULTON, L. L., FULTON, R. S., HARRISON, M., OBERKFELL, B., DU, F., DEMETER, R., VICKERY, T. L., ELHAMMALI, A., PIWNICA-WORMS, H., MCDONALD, S., WATSON, M., DOOLING, D. J., OTA, D., CHANG, L. W., BOSE, R., LEY, T. J., PIWNICA-WORMS, D., STUART, J. M., WILSON, R. K. & MARDIS, E. R. 2012. Whole-genome analysis informs breast cancer response to aromatase inhibition. *Nature,* 486**,** 353-60.

ELVIDGE, G. P., GLENNY, L., APPELHOFF, R. J., RATCLIFFE, P. J., RAGOUSSIS, J. & GLEADLE, J. M. 2006. Concordant regulation of gene expression by hypoxia and 2-oxoglutarate-dependent dioxygenase inhibition: the role of HIF-1alpha, HIF-2alpha, and other pathways. *J Biol Chem,* 281**,** 15215-26.

EPSTEIN, A. C., GLEADLE, J. M., MCNEILL, L. A., HEWITSON, K. S., O'ROURKE, J., MOLE, D. R., MUKHERJI, M., METZEN, E., WILSON, M. I., DHANDA, A., TIAN, Y. M., MASSON, N., HAMILTON, D. L., JAAKKOLA, P., BARSTEAD, R., HODGKIN, J., MAXWELL, P. H., PUGH, C. W., SCHOFIELD, C. J. & RATCLIFFE, P. J. 2001. C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell,* 107**,** 43-54.

ERLER, J. T. & GIACCIA, A. J. 2006. Lysyl oxidase mediates hypoxic control of metastasis. *Cancer Res,* 66**,** 10238-41.

ERNENS, I., GOODFELLOW, S. J., INNES, F., KENNETH, N. S., DERBLAY, L. E., WHITE, R. J. & SCOTT, P. H. 2006. Hypoxic stress suppresses RNA polymerase III recruitment and tRNA gene transcription in cardiomyocytes. *Nucleic Acids Research,* 34**,** 286-294.

ESTEBAN, M. A., TRAN, M. G., HARTEN, S. K., HILL, P., CASTELLANOS, M. C., CHANDRA, A., RAVAL, R., O'BRIEN T, S. & MAXWELL, P. H. 2006. Regulation of E-cadherin expression by VHL and hypoxia-inducible factor. *Cancer Res,* 66**,** 3567-75.

ESTELLER, M. 2011. Non-coding RNAs in human disease. *Nat Rev Genet,* 12**,** 861-74.

FABRE-LAFAY, S., MONVILLE, F., GARRIDO-URBANI, S., BERRUYER-POUYET, C., GINESTIER, C., REYMOND, N., FINETTI, P., SAUVAN, R., ADELAIDE, J., GENEIX, J., LECOCQ, E., POPOVICI, C., DUBREUIL, P., VIENS, P., GONCALVES, A., CHARAFE-JAUFFRET, E., JACQUEMIER, J., BIRNBAUM, D. & LOPEZ, M. 2007. Nectin-4 is a new histological and serological tumor associated marker for breast cancer. *BMC Cancer,* 7**,** 73.

FAGHIHI, M. A. & WAHLESTEDT, C. 2009. Regulatory roles of natural antisense transcripts. *Nat Rev Mol Cell Biol,* 10**,** 637-43.

FAVARO, E., BENSAAD, K., CHONG, M. G., TENNANT, D. A., FERGUSON, D. J., SNELL, C., STEERS, G., TURLEY, H., LI, J. L., GUNTHER, U. L., BUFFA, F. M., MCINTYRE, A. & HARRIS, A. L. 2012. Glucose utilization via glycogen phosphorylase sustains proliferation and prevents premature senescence in cancer cells. *Cell Metab,* 16**,** 751-64.

FAVARO, E., LORD, S., HARRIS, A. L. & BUFFA, F. M. 2011. Gene expression and hypoxia in breast cancer. *Genome Med,* 3**,** 55.

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.

FELDSER, D., AGANI, F., IYER, N. V., PAK, B., FERREIRA, G. & SEMENZA, G. L. 1999. Reciprocal positive regulation of hypoxia-inducible factor 1alpha and insulin-like growth factor 2. *Cancer Res,* 59**,** 3915-8.

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,* 20**,** 1675-87.

FERLAY, J., SHIN, H. R., BRAY, F., FORMAN, D., MATHERS, C. & PARKIN, D. M. 2010. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer,* 127**,** 2893-917.

FERREIRA, H. J., HEYN, H., MOUTINHO, C. & ESTELLER, M. 2012. CpG island hypermethylation-associated silencing of small nucleolar RNAs in human cancer. *RNA Biol,* 9**,** 881-90.

FILIPOWICZ, W. & POGACIC, V. 2002. Biogenesis of small nucleolar ribonucleoproteins. *Curr Opin Cell Biol,* 14**,** 319-27.

FLINTOFT, L. 2013. Non-coding RNA: Structure and function for lncRNAs. *Nat Rev Genet,* 14**,** 598.

FLOCKHART, R. J., WEBSTER, D. E., QU, K., MASCARENHAS, N., KOVALSKI, J., KRETZ, M. & KHAVARI, P. A. 2012. BRAFV600E remodels the melanocyte transcriptome and induces BANCR to regulate melanoma cell migration. *Genome Res,* 22**,** 1006-14.

FOEKENS, J. A., PETERS, H. A., GREBENCHTCHIKOV, N., LOOK, M. P., MEIJER-VAN GELDER, M. E., GEURTS-MOESPOT, A., VAN DER KWAST, T. H., SWEEP, C. G. & KLIJN, J. G. 2001. High tumor levels of vascular endothelial growth factor predict poor response to systemic therapy in advanced breast cancer. *Cancer Res,* 61**,** 5407-14.

FOKAS, E., PREVO, R., POLLARD, J. R., REAPER, P. M., CHARLTON, P. A., CORNELISSEN, B., VALLIS, K. A., HAMMOND, E. M., OLCINA, M. M., GILLIES MCKENNA, W., MUSCHEL, R. J. & BRUNNER, T. B. 2012. Targeting ATR in vivo using the novel inhibitor VE-822 results in selective sensitization of pancreatic tumors to radiation. *Cell Death Dis,* 3**,** e441.

FOLEY, N. H., BRAY, I. M., TIVNAN, A., BRYAN, K., MURPHY, D. M., BUCKLEY, P. G., RYAN, J., O'MEARA, A., O'SULLIVAN, M. & STALLINGS, R. L. 2010. MicroRNA-184 inhibits

neuroblastoma cell survival through targeting the serine/threonine kinase AKT2. *Mol Cancer,* 9**,** 83.

FOX, A. H., LAM, Y. W., LEUNG, A. K., LYON, C. E., ANDERSEN, J., MANN, M. & LAMOND, A. I. 2002. Paraspeckles: a novel nuclear domain. *Curr Biol,* 12**,** 13-25.

FREANEY, J. E., KIM, R., MANDHANA, R. & HORVATH, C. M. 2013. Extensive cooperation of immune master regulators IRF3 and NFkappaB in RNA Pol II recruitment and pause release in human innate antiviral transcription. *Cell Rep,* 4**,** 959-73.

FYLES, A., MILOSEVIC, M., HEDLEY, D., PINTILIE, M., LEVIN, W., MANCHUL, L. & HILL, R. P. 2002. Tumor hypoxia has independent predictor impact only in patients with node-negative cervix cancer. *J Clin Oncol,* 20**,** 680-7.

GALBRAITH, M. D., ALLEN, M. A., BENSARD, C. L., WANG, X., SCHWINN, M. K., QIN, B., LONG, H. W., DANIELS, D. L., HAHN, W. C., DOWELL, R. D. & ESPINOSA, J. M. 2013. HIF1A employs CDK8-mediator to stimulate RNAPII elongation in response to hypoxia. *Cell,* 153**,** 1327-39.

GATENBY, R. A. & GILLIES, R. J. 2004. Why do cancers have high aerobic glycolysis? *Nat Rev Cancer,* 4**,** 891-9.

GEE, H. E., BUFFA, F. M., CAMPS, C., RAMACHANDRAN, A., LEEK, R., TAYLOR, M., PATIL, M., SHELDON, H., BETTS, G., HOMER, J., WEST, C., RAGOUSSIS, J. & HARRIS, A. L. 2011. The small-nucleolar RNAs commonly used for microRNA normalisation correlate with tumour pathology and prognosis. *Br J Cancer,* 104**,** 1168-77.

GEE, H. E., IVAN, C., CALIN, G. A. & IVAN, M. 2013. HypoxamiRs and Cancer: From Biology to Targeted Therapy. *Antioxid Redox Signal*.

GEISLER, S. & COLLER, J. 2013. RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. *Nat Rev Mol Cell Biol,* 14**,** 699-712.

GENERALI, D., BERRUTI, A., BRIZZI, M. P., CAMPO, L., BONARDI, S., WIGFIELD, S., BERSIGA, A., ALLEVI, G., MILANI, M., AGUGGINI, S., GANDOLFI, V., DOGLIOTTI, L., BOTTINI, A., HARRIS, A. L. & FOX, S. B. 2006. Hypoxia-inducible factor-1alpha expression predicts a poor response to primary chemoendocrine therapy and disease-free survival in primary human breast cancer. *Clin Cancer Res,* 12**,** 4562-8.

GENERALI, D., BUFFA, F. M., BERRUTI, A., BRIZZI, M. P., CAMPO, L., BONARDI, S., BERSIGA, A., ALLEVI, G., MILANI, M., AGUGGINI, S., PAPOTTI, M., DOGLIOTTI, L., BOTTINI, A., HARRIS, A. L. & FOX, S. B. 2009. Phosphorylated ERalpha, HIF-1alpha, and MAPK signaling as predictors of primary endocrine treatment response and resistance in patients with breast cancer. *J Clin Oncol,* 27**,** 227-34.

GERALD, D., BERRA, E., FRAPART, Y. M., CHAN, D. A., GIACCIA, A. J., MANSUY, D., POUYSSEGUR, J., YANIV, M. & MECHTA-GRIGORIOU, F. 2004. JunD reduces tumor angiogenesis by protecting cells from oxidative stress. *Cell,* 118**,** 781-94.

GIATROMANOLAKI, A., SIVRIDIS, E., FISKA, A. & KOUKOURAKIS, M. I. 2006. Hypoxia-inducible factor-2 alpha (HIF-2 alpha) induces angiogenesis in breast carcinomas. *Appl Immunohistochem Mol Morphol,* 14**,** 78-82.

GOETSCH, L., HAEUW, J. F., BEAU-LARVOR, C., GONZALEZ, A., ZANNA, L., MALISSARD, M., LEPECQUET, A. M., ROBERT, A., BAILLY, C., BROUSSAS, M. & CORVAIA, N. 2013. A novel role for junctional adhesion molecule-A in tumor proliferation: modulation by an anti-JAM-A monoclonal antibody. *Int J Cancer,* 132**,** 1463-74.

GOMES, A. Q., NOLASCO, S. & SOARES, H. 2013. Non-coding RNAs: multi-tasking molecules in the cell. *Int J Mol Sci,* 14**,** 16010-39.

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**,** 335-47.

GORDAN, J. D., LAL, P., DONDETI, V. R., LETRERO, R., PAREKH, K. N., OQUENDO, C. E., GREENBERG, R. A., FLAHERTY, K. T., RATHMELL, W. K., KEITH, B., SIMON, M. C. & NATHANSON, K. L. 2008. HIF-alpha effects on c-Myc distinguish two subtypes of sporadic VHL-deficient clear cell renal carcinoma. *Cancer Cell,* 14**,** 435-46.

GREIJER, A. E., VAN DER GROEP, P., KEMMING, D., SHVARTS, A., SEMENZA, G. L., MEIJER, G. A., VAN DE WIEL, M. A., BELIEN, J. A., VAN DIEST, P. J. & VAN DER WALL, E. 2005. Up-regulation of gene expression by hypoxia is mediated predominantly by hypoxia-inducible factor 1 (HIF-1). *J Pathol,* 206**,** 291-304.

GREMLICH, S., DAMNON, F., REYMONDIN, D., BRAISSANT, O., SCHITTNY, J. C., BAUD, D., GERBER, S. & ROTH-KLEINER, M. 2013. The long non-coding RNA NEAT1 is increased in IUGR placentas, leading to potential new hypotheses of IUGR origin/development. *Placenta*.

GRIGORIADIS, A., OLIVER, G. R., TANNEY, A., KENDRICK, H., SMALLEY, M. J., JAT, P. & NEVILLE, A. M. 2009. Identification of differentially expressed sense and antisense transcript pairs in breast epithelial tissues. *BMC Genomics,* 10**,** 324.

GRIMSHAW, M. J. 2007. Endothelins and hypoxia-inducible factor in cancer. *Endocr Relat Cancer,* 14**,** 233-44.

GRUBER, G., GREINER, R. H., HLUSHCHUK, R., AEBERSOLD, D. M., ALTERMATT, H. J., BERCLAZ, G. & DJONOV, V. 2004. Hypoxia-inducible factor 1 alpha in high-risk breast cancer: an independent prognostic parameter? *Breast Cancer Res,* 6**,** R191-8.

GRUDZIEN, P., LO, S., ALBAIN, K. S., ROBINSON, P., RAJAN, P., STRACK, P. R., GOLDE, T. E., MIELE, L. & FOREMAN, K. E. 2010. Inhibition of Notch signaling reduces the stem-like population of breast cancer cells and prevents mammosphere formation. *Anticancer Res,* 30**,** 3853-67.

GUIMBELLOT, J. S., ERICKSON, S. W., MEHTA, T., WEN, H., PAGE, G. P., SORSCHER, E. J. & HONG, J. S. 2009. Correlation of microRNA levels during hypoxia with predicted target mRNAs through genome-wide microarray analysis. *BMC Med Genomics,* 2**,** 15.

GUNARATNAM, L., MORLEY, M., FRANOVIC, A., DE PAULSEN, N., MEKHAIL, K., PAROLIN, D. A., NAKAMURA, E., LORIMER, I. A. & LEE, S. 2003. Hypoxia inducible factor activates the transforming growth factor-alpha/epidermal growth factor receptor growth stimulatory pathway in VHL(-/-) renal cell carcinoma cells. *J Biol Chem,* 278**,** 44966-74.

GUPTA, M. & LIU, J. S. 2005. De novo cis-regulatory module elicitation for eukaryotic genomes. *Proc Natl Acad Sci U S A,* 102**,** 7079-84.

GUPTA, R. A., SHAH, N., WANG, K. C., KIM, J., HORLINGS, H. M., WONG, D. J., TSAI, M. C., HUNG, T., ARGANI, P., RINN, J. L., WANG, Y., BRZOSKA, P., KONG, B., LI, R., WEST, R. B., VAN DE VIJVER, M. J., SUKUMAR, S. & CHANG, H. Y. 2010. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature,* 464**,** 1071-6.

GURU, S. C., AGARWAL, S. K., MANICKAM, P., OLUFEMI, S. E., CRABTREE, J. S., WEISEMANN, J. M., KESTER, M. B., KIM, Y. S., WANG, Y., EMMERT-BUCK, M. R., LIOTTA, L. A., SPIEGEL, A. M., BOGUSKI, M. S., ROE, B. A., COLLINS, F. S., MARX, S. J., BURNS, L. & CHANDRASEKHARAPPA, S. C. 1997. A transcript map for the 2.8-Mb region containing the multiple endocrine neoplasia type 1 locus. *Genome Res,* 7**,** 725-35.

GUSTAFSSON, M. V., ZHENG, X., PEREIRA, T., GRADIN, K., JIN, S., LUNDKVIST, J., RUAS, J. L., POELLINGER, L., LENDAHL, U. & BONDESSON, M. 2005. Hypoxia requires notch signaling to maintain the undifferentiated cell state. *Dev Cell,* 9**,** 617-28.

GUTSCHNER, T., HAMMERLE, M., EISSMANN, M., HSU, J., KIM, Y., HUNG, G., REVENKO, A., ARUN, G., STENTRUP, M., GROSS, M., ZORNIG, M., MACLEOD, A. R., SPECTOR, D. L. & DIEDERICHS, S. 2013. The noncoding RNA MALAT1 is a critical regulator of the metastasis phenotype of lung cancer cells. *Cancer Res,* 73**,** 1180-9.

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**,** 223-7.

GUTTMAN, M., GARBER, M., LEVIN, J. Z., DONAGHEY, J., ROBINSON, J., ADICONIS, X., FAN, L., KOZIOL, M. J., GNIRKE, A., NUSBAUM, C., RINN, J. L., LANDER, E. S. & REGEV, A. 2010. Ab

initio reconstruction of cell type-specific transcriptomes in mouse reveals the conserved multi-exonic structure of lincRNAs. *Nat Biotechnol,* 28**,** 503-10.

HAH, N., DANKO, C. G., CORE, L., WATERFALL, J. J., SIEPEL, A., LIS, J. T. & KRAUS, W. L. 2011. A rapid, extensive, and transient transcriptional response to estrogen signaling in breast cancer cells. *Cell,* 145**,** 622-34.

HAIGIS, M. C. & SINCLAIR, D. A. 2010. Mammalian sirtuins: biological insights and disease relevance. *Annu Rev Pathol,* 5**,** 253-95.

HAMMOND, E. M., ASSELIN, M. C., FORSTER, D., O'CONNOR, J. P., SENRA, J. M. & WILLIAMS, K. J. 2014. The meaning, measurement and modification of hypoxia in the laboratory and the clinic. *Clin Oncol (R Coll Radiol),* 26**,** 277-88.

HAN, Y., LIU, Y., NIE, L., GUI, Y. & CAI, Z. 2013. Inducing cell proliferation inhibition, apoptosis, and motility reduction by silencing long noncoding ribonucleic acid metastasis-associated lung adenocarcinoma transcript 1 in urothelial carcinoma of the bladder. *Urology,* 81**,** 209 e1-7.

HANG, X., LI, P., LI, Z., QU, W., YU, Y., LI, H., SHEN, Z., ZHENG, H., GAO, Y., WU, Y., DENG, M., SUN, Z. & ZHANG, C. 2009. Transcription and splicing regulation in human umbilical vein endothelial cells under hypoxic stress conditions by exon array. *BMC Genomics,* 10**,** 126.

HANSEN, T. B., JENSEN, T. I., CLAUSEN, B. H., BRAMSEN, J. B., FINSEN, B., DAMGAARD, C. K. & KJEMS, J. 2013. Natural RNA circles function as efficient microRNA sponges. *Nature,* 495**,** 384-8.

HARBERS, M. & CARNINCI, P. 2005. Tag-based approaches for transcriptome research and genome annotation. *Nat Methods,* 2**,** 495-502.

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

HARRIS, T. D., BUZBY, P. R., BABCOCK, H., BEER, E., BOWERS, J., BRASLAVSKY, I., CAUSEY, M., COLONELL, J., DIMEO, J., EFCAVITCH, J. W., GILADI, E., GILL, J., HEALY, J., JAROSZ, M., LAPEN, D., MOULTON, K., QUAKE, S. R., STEINMANN, K., THAYER, E., TYURINA, A., WARD, R., WEISS, H. & XIE, Z. 2008. Single-molecule DNA sequencing of a viral genome. *Science,* 320**,** 106-9.

HARTZOG, G. A. & MARTENS, J. A. 2009. ncRNA transcription makes its mark. *EMBO J,* 28**,** 1679-80.

HAUPTMAN, N. & GLAVAC, D. 2013. MicroRNAs and long non-coding RNAs: prospects in diagnostics and therapy of cancer. *Radiol Oncol,* 47**,** 311-8.

HE, L. & HANNON, G. J. 2004. MicroRNAs: small RNAs with a big role in gene regulation. *Nat Rev Genet,* 5**,** 522-31.

HE, Y., VOGELSTEIN, B., VELCULESCU, V. E., PAPADOPOULOS, N. & KINZLER, K. W. 2008. The antisense transcriptomes of human cells. *Science,* 322**,** 1855-7.

HEBERT, C., NORRIS, K., SCHEPER, M. A., NIKITAKIS, N. & SAUK, J. J. 2007. High mobility group A2 is a target for miRNA-98 in head and neck squamous cell carcinoma. *Mol Cancer,* 6**,** 5.

HELCZYNSKA, K., LARSSON, A. M., HOLMQUIST MENGELBIER, L., BRIDGES, E., FREDLUND, E., BORGQUIST, S., LANDBERG, G., PAHLMAN, S. & JIRSTROM, K. 2008. Hypoxia-inducible factor-2alpha correlates to distant recurrence and poor outcome in invasive breast cancer. *Cancer Res,* 68**,** 9212-20.

HERZING, L. B., ROMER, J. T., HORN, J. M. & ASHWORTH, A. 1997. Xist has properties of the X-chromosome inactivation centre. *Nature,* 386**,** 272-5.

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**,** 4157-63.

HIROSE, T., VIRNICCHI, G., TANIGAWA, A., NAGANUMA, T., LI, R., KIMURA, H., YOKOI, T., NAKAGAWA, S., BENARD, M., FOX, A. H. & PIERRON, G. 2014. NEAT1 long noncoding RNA regulates transcription via protein sequestration within subnuclear bodies. *Mol Biol Cell,* 25**,** 169-83.

HOCHACHKA, P. W., BUCK, L. T., DOLL, C. J. & LAND, S. C. 1996. Unifying theory of hypoxia tolerance: molecular/metabolic defense and rescue mechanisms for surviving oxygen lack. *Proc Natl Acad Sci U S A,* 93**,** 9493-8.

HOCHACHKA, P. W. & LUTZ, P. L. 2001. Mechanism, origin, and evolution of anoxia tolerance in animals. *Comp Biochem Physiol B Biochem Mol Biol,* 130**,** 435-59.

HOCKEL, M., SCHLENGER, K., HOCKEL, S. & VAUPEL, P. 1999. Hypoxic cervical cancers with low apoptotic index are highly aggressive. *Cancer Res,* 59**,** 4525-8.

HOCKEL, M. & VAUPEL, P. 2001. Tumor hypoxia: definitions and current clinical, biologic, and molecular aspects. *J Natl Cancer Inst,* 93**,** 266-76.

HOHENBERGER, P., FELGNER, C., HAENSCH, W. & SCHLAG, P. M. 1998. Tumor oxygenation correlates with molecular growth determinants in breast cancer. *Breast Cancer Res Treat,* 48**,** 97-106.

HOLMQUIST-MENGELBIER, L., FREDLUND, E., LOFSTEDT, T., NOGUERA, R., NAVARRO, S., NILSSON, H., PIETRAS, A., VALLON-CHRISTERSSON, J., BORG, A., GRADIN, K., POELLINGER, L. & PAHLMAN, S. 2006. Recruitment of HIF-1alpha and HIF-2alpha to common target genes is differentially regulated in neuroblastoma: HIF-2alpha promotes an aggressive phenotype. *Cancer Cell,* 10**,** 413-23.

HONG, L., YANG, J., HAN, Y., LU, Q., CAO, J. & SYED, L. 2012. High expression of miR-210 predicts poor survival in patients with breast cancer: a meta-analysis. *Gene,* 507**,** 135-8.

HOQUE, M., JI, Z., ZHENG, D., LUO, W., LI, W., YOU, B., PARK, J. Y., YEHIA, G. & TIAN, B. 2013. Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. *Nat Methods,* 10**,** 133-9.

HORSMAN, M. R., MORTENSEN, L. S., PETERSEN, J. B., BUSK, M. & OVERGAARD, J. 2012. Imaging hypoxia to improve radiotherapy outcome. *Nat Rev Clin Oncol,* 9**,** 674-87.

HORVATH, A., PAKALA, S. B., MUDVARI, P., REDDY, S. D., OHSHIRO, K., CASIMIRO, S., PIRES, R., FUQUA, S. A., TOI, M., COSTA, L., NAIR, S. S., SUKUMAR, S. & KUMAR, R. 2013. Novel insights into breast cancer genetic variance through RNA sequencing. *Sci Rep,* 3**,** 2256.

HU, C. J., WANG, L. Y., CHODOSH, L. A., KEITH, B. & SIMON, M. C. 2003. Differential roles of hypoxia-inducible factor 1alpha (HIF-1alpha) and HIF-2alpha in hypoxic gene regulation. *Mol Cell Biol,* 23**,** 9361-74.

HUA, Z., LV, Q., YE, W., WONG, C. K., CAI, G., GU, D., JI, Y., ZHAO, C., WANG, J., YANG, B. B. & ZHANG, Y. 2006. MiRNA-directed regulation of VEGF and other angiogenic factors under hypoxia. *PLoS One,* 1**,** e116.

HUANG, H., GROTH, J., SOSSEY-ALAOUI, K., HAWTHORN, L., BEALL, S. & GERADTS, J. 2005. Aberrant expression of novel and previously described cell membrane markers in human breast cancer cell lines and tumors. *Clin Cancer Res,* 11**,** 4357-64.

HUANG, L. E. 2008. Carrot and stick: HIF-alpha engages c-Myc in hypoxic adaptation. *Cell Death Differ,* 15**,** 672-7.

HUANG, L. E. 2013. Biochemistry. How HIF-1alpha handles stress. *Science,* 339**,** 1285-6.

HUANG, X., DING, L., BENNEWITH, K. L., TONG, R. T., WELFORD, S. M., ANG, K. K., STORY, M., LE, Q. T. & GIACCIA, A. J. 2009. Hypoxia-inducible mir-210 regulates normoxic gene expression involved in tumor initiation. *Mol Cell,* 35**,** 856-67.

HUANG, X., LE, Q. T. & GIACCIA, A. J. 2010. MiR-210--micromanager of the hypoxia pathway. *Trends Mol Med,* 16**,** 230-7.

HUARTE, M., GUTTMAN, M., FELDSER, D., GARBER, M., KOZIOL, M. J., KENZELMANN-BROZ, D., KHALIL, A. M., ZUK, O., AMIT, I., RABANI, M., ATTARDI, L. D., REGEV, A., LANDER, E. S., JACKS, T. & RINN, J. L. 2010. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell,* 142**,** 409-19.

HUDIS, C. A. 2007. Trastuzumab--mechanism of action and use in clinical practice. *N Engl J Med,* 357**,** 39-51.

HUSSAIN, S. A., GANESAN, R., REYNOLDS, G., GROSS, L., STEVENS, A., PASTOREK, J., MURRAY, P. G., PERUNOVIC, B., ANWAR, M. S., BILLINGHAM, L., JAMES, N. D., SPOONER, D., POOLE, C. J., REA, D. W. & PALMER, D. H. 2007. Hypoxia-regulated carbonic anhydrase IX expression is associated with poor survival in patients with invasive breast cancer. *Br J Cancer,* 96**,** 104-9.

IDEUE, T., HINO, K., KITAO, S., YOKOI, T. & HIROSE, T. 2009. Efficient oligonucleotide-mediated degradation of nuclear noncoding RNAs in mammalian cultured cells. *RNA,* 15**,** 1578-87.

INAKI, K., HILLMER, A. M., UKIL, L., YAO, F., WOO, X. Y., VARDY, L. A., ZAWACK, K. F., LEE, C. W., ARIYARATNE, P. N., CHAN, Y. S., DESAI, K. V., BERGH, J., HALL, P., PUTTI, T. C., ONG, W. L., SHAHAB, A., CACHEUX-RATABOUL, V., KARUTURI, R. K., SUNG, W. K., RUAN, X., BOURQUE, G., RUAN, Y. & LIU, E. T. 2011. Transcriptional consequences of genomic structural aberrations in breast cancer. *Genome Res,* 21**,** 676-87.

IRIGOYEN, M., ANSO, E., MARTINEZ, E., GARAYOA, M., MARTINEZ-IRUJO, J. J. & ROUZAUT, A. 2007. Hypoxia alters the adhesive properties of lymphatic endothelial cells. A transcriptional and functional study. *Biochim Biophys Acta,* 1773**,** 880-90.

ISHIZU, H., SIOMI, H. & SIOMI, M. C. 2012. Biology of PIWI-interacting RNAs: new insights into biogenesis and function inside and outside of germlines. *Genes Dev,* 26**,** 2361-73.

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**,** 468-72.

JADY, B. E. & KISS, T. 2000. Characterisation of the U83 and U84 small nucleolar RNAs: two novel 2'-O-ribose methylation guide RNAs that lack complementarities to ribosomal RNAs. *Nucleic Acids Res,* 28**,** 1348-54.

JADY, B. E. & KISS, T. 2001. A small nucleolar guide RNA functions both in 2'-O-ribose methylation and pseudouridylation of the U5 spliceosomal RNA. *EMBO J,* 20**,** 541-51.

JALALI, S., BHARTIYA, D., LALWANI, M. K., SIVASUBBU, S. & SCARIA, V. 2013. Systematic transcriptome wide analysis of lncRNA-miRNA interactions. *PLoS One,* 8**,** e53823.

JANICKE, R. U. 2009. MCF-7 breast carcinoma cells do not express caspase-3. *Breast Cancer Res Treat,* 117**,** 219-21.

JANICKE, R. U., SPRENGART, M. L., WATI, M. R. & PORTER, A. G. 1998. Caspase-3 is required for DNA fragmentation and morphological changes associated with apoptosis. *J Biol Chem,* 273**,** 9357-60.

JENJAROENPUN, P., KREMENSKA, Y., NAIR, V. M., KREMENSKOY, M., JOSEPH, B. & KUROCHKIN, I. V. 2013. Characterization of RNA in exosomes secreted by human breast cancer cell lines using next-generation sequencing. *PeerJ,* 1**,** e201.

JI, H., JIANG, H., MA, W., JOHNSON, D. S., MYERS, R. M. & WONG, W. H. 2008. An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nat Biotechnol,* 26**,** 1293-300.

JI, P., DIEDERICHS, S., WANG, W., BOING, S., METZGER, R., SCHNEIDER, P. M., TIDOW, N., BRANDT, B., BUERGER, H., BULK, E., THOMAS, M., BERDEL, W. E., SERVE, H. & MULLER-TIDOW, C. 2003. MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene,* 22**,** 8031-41.

JI, Q., LIU, X., FU, X., ZHANG, L., SUI, H., ZHOU, L., SUN, J., CAI, J., QIN, J., REN, J. & LI, Q. 2013. Resveratrol Inhibits Invasion and Metastasis of Colorectal Cancer Cells via MALAT1 Mediated Wnt/beta-Catenin Signal Pathway. *PLoS One,* 8**,** e78700.

JIANG, F. N., HE, H. C., ZHANG, Y. Q., YANG, D. L., HUANG, J. H., ZHU, Y. X., MO, R. J., CHEN, G., YANG, S. B., CHEN, Y. R., ZHONG, W. D. & ZHOU, W. L. 2013. An Integrative Proteomics and Interaction Network-Based Classifier for Prostate Cancer Diagnosis. *Plos One,* 8.

JOHNSON, A. B. & BARTON, M. C. 2007. Hypoxia-induced and stress-specific changes in chromatin structure and function. *Mutat Res,* 618**,** 149-62.

JOHNSON, A. B., DENKO, N. & BARTON, M. C. 2008. Hypoxia induces a novel signature of chromatin modifications and global repression of transcription. *Mutat Res,* 640**,** 174-9.

JUBB, A. M., BUFFA, F. M. & HARRIS, A. L. 2010. Assessment of tumour hypoxia for prediction of response to therapy and cancer prognosis. *J Cell Mol Med,* 14**,** 18-29.

KAELIN, W. G. 2007. Von Hippel-Lindau disease. *Annu Rev Pathol,* 2**,** 145-73.

KAELIN, W. G., JR. 2005. ROS: really involved in oxygen sensing. *Cell Metab,* 1**,** 357-8.

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.

KAIDI, A., WILLIAMS, A. C. & PARASKEVA, C. 2007. Interaction between beta-catenin and HIF-1 promotes cellular adaptation to hypoxia. *Nat Cell Biol,* 9**,** 210-7.

KANG, S. S., CHUN, Y. K., HUR, M. H., LEE, H. K., KIM, Y. J., HONG, S. R., LEE, J. H., LEE, S. G. & PARK, Y. K. 2002. Clinical significance of glucose transporter 1 (GLUT1) expression in human breast carcinoma. *Jpn J Cancer Res,* 93**,** 1123-8.

KAPRANOV, P., CHENG, J., DIKE, S., NIX, D. A., DUTTAGUPTA, R., WILLINGHAM, A. T., STADLER, P. F., HERTEL, J., HACKERMULLER, J., HOFACKER, I. L., BELL, I., CHEUNG, E., DRENKOW, J., DUMAIS, E., PATEL, S., HELT, G., GANESH, M., GHOSH, S., PICCOLBONI, A., SEMENTCHENKO, V., TAMMANA, H. & GINGERAS, T. R. 2007. RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science,* 316**,** 1484-8.

KAPRANOV, P., ST LAURENT, G., RAZ, T., OZSOLAK, F., REYNOLDS, C. P., SORENSEN, P. H., REAMAN, G., MILOS, P., ARCECI, R. J., THOMPSON, J. F. & TRICHE, T. J. 2010. The majority of total nuclear-encoded non-ribosomal RNA in a human cell is 'dark matter' un-annotated RNA. *BMC Biol,* 8**,** 149.

KASSAVETIS, G. A., SORAGNI, E., DRISCOLL, R. & GEIDUSCHEK, E. P. 2005. Reconfiguring the connectivity of a multiprotein complex: Fusions of yeast TATA-binding protein with Brf1, and the function of transcription factor IIIB. *Proceedings of the National Academy of Sciences of the United States of America,* 102**,** 15406-15411.

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.

KING, T. H., LIU, B., MCCULLY, R. R. & FOURNIER, M. J. 2003. Ribosome structure and activity are altered in cells lacking snoRNPs that form pseudouridines in the peptidyl transferase center. *Mol Cell,* 11**,** 425-35.

KISHORE, S., KHANNA, A., ZHANG, Z., HUI, J., BALWIERZ, P. J., STEFAN, M., BEACH, C., NICHOLLS, R. D., ZAVOLAN, M. & STAMM, S. 2010. The snoRNA MBII-52 (SNORD 115) is processed into smaller RNAs and regulates alternative splicing. *Hum Mol Genet,* 19**,** 1153-64.

KISS-LASZLO, Z., HENRY, Y., BACHELLERIE, J. P., CAIZERGUES-FERRER, M. & KISS, T. 1996. Site-specific ribose methylation of preribosomal RNA: a novel function for small nucleolar RNAs. *Cell,* 85**,** 1077-88.

KIZAKA-KONDOH, S., TANAKA, S., HARADA, H. & HIRAOKA, M. 2009. The HIF-1-active microenvironment: an environmental target for cancer therapy. *Adv Drug Deliv Rev,* 61**,** 623-32.

KNAUP, K. X., MONTI, J., HACKENBECK, T., JOBST-SCHWAN, T., KLANKE, B., SCHIETKE, R. E., WACKER, I., BEHRENS, J., AMANN, K., ECKARDT, K. U., WARNECKE, C. & WIESENER, M. S. 2013. Hypoxia regulates the sperm associated antigen 4 (SPAG4) via HIF, which is expressed in renal clear cell carcinoma and promotes migration and invasion in vitro. *Mol Carcinog*.

KNOWLES, H. J., CLETON-JANSEN, A. M., KORSCHING, E. & ATHANASOU, N. A. 2010. Hypoxia-inducible factor regulates osteoclast-mediated bone resorption: role of angiopoietin-like 4. *FASEB J,* 24**,** 4648-59.

KNOWLES, H. J., RAVAL, R. R., HARRIS, A. L. & RATCLIFFE, P. J. 2003. Effect of ascorbate on the activity of hypoxia-inducible factor in cancer cells. *Cancer Res,* 63**,** 1764-8.

KOBERLE, V., KRONENBERGER, B., PLELI, T., TROJAN, J., IMELMANN, E., PEVELING-OBERHAG, J., WELKER, M. W., ELHENDAWY, M., ZEUZEM, S., PIIPER, A. & WAIDMANN, O. 2013. Serum microRNA-1 and microRNA-122 are prognostic markers in patients with hepatocellular carcinoma. *Eur J Cancer,* 49**,** 3442-9.

KOBOLDT, D. C., STEINBERG, K. M., LARSON, D. E., WILSON, R. K. & MARDIS, E. R. 2013. The next-generation sequencing revolution and its impact on genomics. *Cell,* 155**,** 27-38.

KODZIUS, R., KOJIMA, M., NISHIYORI, H., NAKAMURA, M., FUKUDA, S., TAGAMI, M., SASAKI, D., IMAMURA, K., KAI, C., HARBERS, M., HAYASHIZAKI, Y. & CARNINCI, P. 2006. CAGE: cap analysis of gene expression. *Nat Methods,* 3**,** 211-22.

KOH, M. Y. & POWIS, G. 2012. Passing the baton: the HIF switch. *Trends Biochem Sci,* 37**,** 364-72.

KOH, M. Y., SPIVAK-KROIZMAN, T., VENTURINI, S., WELSH, S., WILLIAMS, R. R., KIRKPATRICK, D. L. & POWIS, G. 2008. Molecular mechanisms for the activity of PX-478, an antitumor inhibitor of the hypoxia-inducible factor-1alpha. *Mol Cancer Ther,* 7**,** 90-100.

KOHNO, K., CHIBA, M., MURATA, S., PAK, S., NAGAI, K., YAMAMOTO, M., YANAGISAWA, K., KOBAYASHI, A., YASUE, H. & OHKOHCHI, N. 2010. Identification of natural antisense transcripts involved in human colorectal cancer development. *Int J Oncol,* 37**,** 1425-32.

KONG, L., ZHANG, Y., YE, Z. Q., LIU, X. Q., ZHAO, S. Q., WEI, L. & GAO, G. 2007. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res,* 35**,** W345-9.

KOONG, A. C., DENKO, N. C., HUDSON, K. M., SCHINDLER, C., SWIERSZ, L., KOCH, C., EVANS, S., IBRAHIM, H., LE, Q. T., TERRIS, D. J. & GIACCIA, A. J. 2000. Candidate genes for the hypoxic tumor phenotype. *Cancer Res,* 60**,** 883-7.

KOOP, E. A., VAN LAAR, T., VAN WICHEN, D. F., DE WEGER, R. A., WALL, E. & VAN DIEST, P. J. 2009. Expression of BNIP3 in invasive breast cancer: correlations with the hypoxic response and clinicopathological features. *BMC Cancer,* 9**,** 175.

KORLACH, J., BJORNSON, K. P., CHAUDHURI, B. P., CICERO, R. L., FLUSBERG, B. A., GRAY, J. J., HOLDEN, D., SAXENA, R., WEGENER, J. & TURNER, S. W. 2010. Real-time DNA sequencing from single polymerase molecules. *Methods Enzymol,* 472**,** 431-55.

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**,** 1949-56.

KRIEG, A. J., RANKIN, E. B., CHAN, D., RAZORENOVA, O., FERNANDEZ, S. & GIACCIA, A. J. 2010. Regulation of the histone demethylase JMJD1A by hypoxia-inducible factor 1 alpha enhances hypoxic gene expression and tumor growth. *Mol Cell Biol,* 30**,** 344-53.

KRISHNAMACHARY, B., ZAGZAG, D., NAGASAWA, H., RAINEY, K., OKUYAMA, H., BAEK, J. H. & SEMENZA, G. L. 2006. Hypoxia-inducible factor-1-dependent repression of E-cadherin in von Hippel-Lindau tumor suppressor-null renal cell carcinoma mediated by TCF3, ZFHX1A, and ZFHX1B. *Cancer Res,* 66**,** 2725-31.

KRONBLAD, A., JIRSTROM, K., RYDEN, L., NORDENSKJOLD, B. & LANDBERG, G. 2006. Hypoxia inducible factor-1alpha is a prognostic marker in premenopausal patients with intermediate to highly differentiated breast cancer but not a predictive marker for tamoxifen response. *Int J Cancer,* 118**,** 2609-16.

KU, C. S., COOPER, D. N. & ROUKOS, D. H. 2013. Clinical relevance of cancer genome sequencing. *World J Gastroenterol,* 19**,** 2011-8.

KULSHRESHTHA, R., DAVULURI, R. V., CALIN, G. A. & IVAN, M. 2008. A microRNA component of the hypoxic response. *Cell Death Differ,* 15**,** 667-71.

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**,** 1859-67.

KUNZ, M. & IBRAHIM, S. M. 2003. Molecular responses to hypoxia in tumor cells. *Mol Cancer,* 2**,** 23.

LACATENA, R. M. & CESARENI, G. 1981. Base pairing of RNA I with its complementary sequence in the primer precursor inhibits ColE1 replication. *Nature,* 294**,** 623-6.

LAFONTAINE, D. L. & TOLLERVEY, D. 1998. Birth of the snoRNPs: the evolution of the modification-guide snoRNAs. *Trends Biochem Sci,* 23**,** 383-8.

LANDO, D., PEET, D. J., WHELAN, D. A., GORMAN, J. J. & WHITELAW, M. L. 2002. Asparagine hydroxylation of the HIF transactivation domain a hypoxic switch. *Science,* 295**,** 858-61.

LANGMEAD, B., TRAPNELL, C., POP, M. & SALZBERG, S. L. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol,* 10**,** R25.

LEE, J. S., KIM, Y., KIM, I. S., KIM, B., CHOI, H. J., LEE, J. M., SHIN, H. J., KIM, J. H., KIM, J. Y., SEO, S. B., LEE, H., BINDA, O., GOZANI, O., SEMENZA, G. L., KIM, M., KIM, K. I., HWANG, D. & BAEK, S. H. 2010. Negative regulation of hypoxic responses via induced Reptin methylation. *Mol Cell,* 39**,** 71-85.

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.

LEITE, K. R., MORAIS, D. R., REIS, S. T., VIANA, N., MOURA, C., FLOREZ, M. G., SILVA, I. A., DIP, N. & SROUGI, M. 2013. MicroRNA 100: a context dependent miRNA in prostate cancer. *Clinics (Sao Paulo),* 68**,** 797-802.

LENDAHL, U., LEE, K. L., YANG, H. & POELLINGER, L. 2009. Generating specificity and diversity in the transcriptional response to hypoxia. *Nat Rev Genet,* 10**,** 821-32.

LEVIN, J. Z., YASSOUR, M., ADICONIS, X., NUSBAUM, C., THOMPSON, D. A., FRIEDMAN, N., GNIRKE, A. & REGEV, A. 2010. Comprehensive comparative analysis of strand-specific RNA sequencing methods. *Nat Methods,* 7**,** 709-15.

LI, B., CAREY, M. & WORKMAN, J. L. 2007. The role of chromatin during transcription. *Cell,* 128**,** 707-19.

LI, Q. F., WANG, X. R., YANG, Y. W. & LIN, H. 2006. Hypoxia upregulates hypoxia inducible factor (HIF)-3alpha expression in lung epithelial cells: characterization and comparison with HIF-1alpha. *Cell Res,* 16**,** 548-58.

LI, S. H., SHIN, D. H., CHUN, Y. S., LEE, M. K., KIM, M. S. & PARK, J. W. 2008. A novel mode of action of YC-1 in HIF inhibition: stimulation of FIH-dependent p300 dissociation from HIF-1{alpha}. *Mol Cancer Ther,* 7**,** 3729-38.

LIAO, D., CORLE, C., SEAGROVES, T. N. & JOHNSON, R. S. 2007. Hypoxia-inducible factor-1alpha is a key regulator of metastasis in a transgenic model of cancer initiation and progression. *Cancer Res,* 67**,** 563-72.

LIBSON, S. & LIPPMAN, M. 2014. A review of clinical aspects of breast cancer. *Int Rev Psychiatry,* 26**,** 4-15.

LICATALOSI, D. D. & DARNELL, R. B. 2010. RNA processing and its regulation: global insights into biological networks. *Nat Rev Genet,* 11**,** 75-87.

LINDBERG, J. & LUNDEBERG, J. 2010. The plasticity of the mammalian transcriptome. *Genomics,* 95**,** 1-6.

LINDERHOLM, B., GRANKVIST, K., WILKING, N., JOHANSSON, M., TAVELIN, B. & HENRIKSSON, R. 2000. Correlation of vascular endothelial growth factor content with recurrences, survival, and first relapse site in primary node-positive breast carcinoma after adjuvant treatment. *J Clin Oncol,* 18**,** 1423-31.

LINNARSSON, S. 2010. Recent advances in DNA sequencing methods - general principles of sample preparation. *Exp Cell Res,* 316**,** 1339-43.

LISTER, R., O'MALLEY, R. C., TONTI-FILIPPINI, J., GREGORY, B. D., BERRY, C. C., MILLAR, A. H. & ECKER, J. R. 2008. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell,* 133**,** 523-36.

LIU, Y., COX, S. R., MORITA, T. & KOUREMBANAS, S. 1995. Hypoxia regulates vascular endothelial growth factor gene expression in endothelial cells. Identification of a 5' enhancer. *Circ Res,* 77**,** 638-43.

LIU, Y., ZHOU, J. & WHITE, K. P. 2014. RNA-seq differential expression studies: more sequence or more replication? *Bioinformatics,* 30**,** 301-4.

LIU, Z., WEI, S., MA, H., ZHAO, M., MYERS, J. N., WEBER, R. S., STURGIS, E. M. & WEI, Q. 2011. A functional variant at the miR-184 binding site in TNFAIP2 and risk of squamous cell carcinoma of the head and neck. *Carcinogenesis,* 32**,** 1668-74.

LIVYATAN, I., HARIKUMAR, A., NISSIM-RAFINIA, M., DUTTAGUPTA, R., GINGERAS, T. R. & MESHORER, E. 2013. Non-polyadenylated transcription in embryonic stem cells reveals novel non-coding RNA related to pluripotency and differentiation. *Nucleic Acids Res,* 41**,** 6300-15.

LOEWER, S., CABILI, M. N., GUTTMAN, M., LOH, Y. H., THOMAS, K., PARK, I. H., GARBER, M., CURRAN, M., ONDER, T., AGARWAL, S., MANOS, P. D., DATTA, S., LANDER, E. S., SCHLAEGER, T. M., DALEY, G. Q. & RINN, J. L. 2010. Large intergenic non-coding RNA-RoR modulates reprogramming of human induced pluripotent stem cells. *Nat Genet,* 42**,** 1113-7.

LORENZI, P. L., REINHOLD, W. C., VARMA, S., HUTCHINSON, A. A., POMMIER, Y., CHANOCK, S. J. & WEINSTEIN, J. N. 2009. DNA fingerprinting of the NCI-60 cell line panel. *Mol Cancer Ther,* 8**,** 713-24.

LU, C. W., LIN, S. C., CHEN, K. F., LAI, Y. Y. & TSAI, S. J. 2008. Induction of pyruvate dehydrogenase kinase-3 by hypoxia-inducible factor-1 promotes metabolic switch and drug resistance. *J Biol Chem,* 283**,** 28106-14.

LU, X., YAN, C. H., YUAN, M., WEI, Y., HU, G. & KANG, Y. 2010. In vivo dynamics and distinct functions of hypoxia in primary tumor growth and organotropic metastasis of breast cancer. *Cancer Res,* 70**,** 3905-14.

LUNDGREN, K., HOLM, C. & LANDBERG, G. 2007. Hypoxia and breast cancer: prognostic and therapeutic implications. *Cell Mol Life Sci,* 64**,** 3233-47.

LUTEIJN, M. J. & KETTING, R. F. 2013. PIWI-interacting RNAs: from generation to transgenerational epigenetics. *Nat Rev Genet,* 14**,** 523-34.

MACKAY, H., HEDLEY, D., MAJOR, P., TOWNSLEY, C., MACKENZIE, M., VINCENT, M., DEGENDORFER, P., TSAO, M. S., NICKLEE, T., BIRLE, D., WRIGHT, J., SIU, L., MOORE, M. & OZA, A. 2005. A phase II trial with pharmacodynamic endpoints of the proteasome inhibitor bortezomib in patients with metastatic colorectal cancer. *Clin Cancer Res,* 11**,** 5526-33.

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**,** 2675-86.

MAINGUY, G., KOSTER, J., WOLTERING, J., JANSEN, H. & DURSTON, A. 2007. Extensive polycistronism and antisense transcription in the mammalian Hox clusters. *PLoS One,* 2**,** e356.

MAJMUNDAR, A. J., WONG, W. J. & SIMON, M. C. 2010. Hypoxia-inducible factors and the response to hypoxic stress. *Mol Cell,* 40**,** 294-309.

MAKINO, Y., CAO, R., SVENSSON, K., BERTILSSON, G., ASMAN, M., TANAKA, H., CAO, Y., BERKENSTAM, A. & POELLINGER, L. 2001. Inhibitory PAS domain protein is a negative regulator of hypoxia-inducible gene expression. *Nature,* 414**,** 550-4.

MAMANOVA, L., ANDREWS, R. M., JAMES, K. D., SHERIDAN, E. M., ELLIS, P. D., LANGFORD, C. F., OST, T. W., COLLINS, J. E. & TURNER, D. J. 2010. FRT-seq: amplification-free, strand-specific transcriptome sequencing. *Nat Methods,* 7**,** 130-2.

MANDEL, C. R., KANEKO, S., ZHANG, H., GEBAUER, D., VETHANTHAM, V., MANLEY, J. L. & TONG, L. 2006. Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature,* 444**,** 953-6.

MANNOOR, K., LIAO, J. & JIANG, F. 2012. Small nucleolar RNAs in cancer. *Biochim Biophys Acta,* 1826**,** 121-8.

MARDIS, E. R. 2013. Next-generation sequencing platforms. *Annu Rev Anal Chem (Palo Alto Calif),* 6**,** 287-303.

MARGUERAT, S. & BAHLER, J. 2010. RNA-seq: from technology to biology. *Cell Mol Life Sci,* 67**,** 569-79.

MARIONI, J. C., MASON, C. E., MANE, S. M., STEPHENS, M. & GILAD, Y. 2008. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res,* 18**,** 1509-17.

MARTENS-UZUNOVA, E. S., BOTTCHER, R., CROCE, C. M., JENSTER, G., VISAKORPI, T. & CALIN, G. A. 2013. Long Noncoding RNA in Prostate, Bladder, and Kidney Cancer. *Eur Urol*.

MARUYAMA, R., SHIPITSIN, M., CHOUDHURY, S., WU, Z., PROTOPOPOV, A., YAO, J., LO, P. K., BESSARABOVA, M., ISHKIN, A., NIKOLSKY, Y., LIU, X. S., SUKUMAR, S. & POLYAK, K. 2012. Altered antisense-to-sense transcript ratios in breast cancer. *Proc Natl Acad Sci U S A,* 109**,** 2820-4.

MARZLUFF, W. F., WAGNER, E. J. & DURONIO, R. J. 2008. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet,* 9**,** 843-54.

MATERA, A. G., TERNS, R. M. & TERNS, M. P. 2007. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. *Nat Rev Mol Cell Biol,* 8**,** 209-20.

MATOUK, I. J., DEGROOT, N., MEZAN, S., AYESH, S., ABU-LAIL, R., HOCHBERG, A. & GALUN, E. 2007. The H19 non-coding RNA is essential for human tumor growth. *PLoS One,* 2**,** e845.

MATOUK, I. J., MEZAN, S., MIZRAHI, A., OHANA, P., ABU-LAIL, R., FELLIG, Y., DEGROOT, N., GALUN, E. & HOCHBERG, A. 2010. The oncofetal H19 RNA connection: hypoxia, p53 and cancer. *Biochim Biophys Acta,* 1803**,** 443-51.

MATTICK, J. S. 2009. The genetic signatures of noncoding RNAs. *PLoS Genet,* 5**,** e1000459.

MAUTE, R. L., DALLA-FAVERA, R. & BASSO, K. 2014. RNAs with multiple personalities. *Wiley Interdiscip Rev RNA,* 5**,** 1-13.

MAXWELL, K. N. & DOMCHEK, S. M. 2012. Cancer treatment according to BRCA1 and BRCA2 mutations. *Nat Rev Clin Oncol,* 9**,** 520-8.

MAZUMDAR, J., O'BRIEN, W. T., JOHNSON, R. S., LAMANNA, J. C., CHAVEZ, J. C., KLEIN, P. S. & SIMON, M. C. 2010. O2 regulates stem cells through Wnt/beta-catenin signalling. *Nat Cell Biol,* 12**,** 1007-13.

MCCORMICK, R., BUFFA, F. M., RAGOUSSIS, J. & HARRIS, A. L. 2010. The role of hypoxia regulated microRNAs in cancer. *Curr Top Microbiol Immunol,* 345**,** 47-70.

MCCORMICK, R. I., BLICK, C., RAGOUSSIS, J., SCHOEDEL, J., MOLE, D. R., YOUNG, A. C., SELBY, P. J., BANKS, R. E. & HARRIS, A. L. 2013. miR-210 is a target of hypoxia-inducible factors 1 and 2 in renal cancer, regulates ISCU and correlates with good prognosis. *Br J Cancer,* 108**,** 1133-42.

MCGETTIGAN, P. A. 2013. Transcriptomics in the RNA-seq era. *Curr Opin Chem Biol,* 17**,** 4-11.

MCINTYRE, A., PATIAR, S., WIGFIELD, S., LI, J. L., LEDAKI, I., TURLEY, H., LEEK, R., SNELL, C., GATTER, K., SLY, W. S., VAUGHAN-JONES, R. D., SWIETACH, P. & HARRIS, A. L. 2012. Carbonic anhydrase IX promotes tumor growth and necrosis in vivo and inhibition enhances anti-VEGF therapy. *Clin Cancer Res,* 18**,** 3100-11.

MEI, Y., CLARK, D. & MAO, L. 2013. Novel dimensions of piRNAs in cancer. *Cancer Lett,* 336**,** 46-52.

MELCHOR, L. & BENITEZ, J. 2013. The complex genetic landscape of familial breast cancer. *Hum Genet,* 132**,** 845-63.

MEMCZAK, S., JENS, M., ELEFSINIOTI, A., TORTI, F., KRUEGER, J., RYBAK, A., MAIER, L., MACKOWIAK, S. D., GREGERSEN, L. H., MUNSCHAUER, M., LOEWER, A., ZIEBOLD, U., LANDTHALER, M., KOCKS, C., LE NOBLE, F. & RAJEWSKY, N. 2013. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature,* 495**,** 333-8.

MENDELL, J. T. 2005. MicroRNAs: critical regulators of development, cellular physiology and malignancy. *Cell Cycle,* 4**,** 1179-84.

MENSE, S. M., SENGUPTA, A., ZHOU, M., LAN, C., BENTSMAN, G., VOLSKY, D. J. & ZHANG, L. 2006. Gene expression profiling reveals the profound upregulation of hypoxia-responsive genes in primary human astrocytes. *Physiol Genomics,* 25**,** 435-49.

MERCER, T. R., DINGER, M. E. & MATTICK, J. S. 2009. Long non-coding RNAs: insights into functions. *Nat Rev Genet,* 10**,** 155-9.

METZKER, M. L. 2010. Sequencing technologies - the next generation. *Nat Rev Genet,* 11**,** 31-46.

MILANI, M. & HARRIS, A. L. 2008. Targeting tumour hypoxia in breast cancer. *Eur J Cancer,* 44**,** 2766-73.

MIMURA, I., NANGAKU, M., KANKI, Y., TSUTSUMI, S., INOUE, T., KOHRO, T., YAMAMOTO, S., FUJITA, T., SHIMAMURA, T., SUEHIRO, J., TAGUCHI, A., KOBAYASHI, M., TANIMURA, K., INAGAKI, T., TANAKA, T., HAMAKUBO, T., SAKAI, J., ABURATANI, H., KODAMA, T. & WADA, Y. 2012.

Dynamic change of chromatin conformation in response to hypoxia enhances the expression of GLUT3 (SLC2A3) by cooperative interaction of hypoxia-inducible factor 1 and KDM3A. *Mol Cell Biol,* 32**,** 3018-32.

MINCHINTON, A. I. & TANNOCK, I. F. 2006. Drug penetration in solid tumours. *Nat Rev Cancer,* 6**,** 583-92.

MOLE, D. R., BLANCHER, C., COPLEY, R. R., POLLARD, P. J., GLEADLE, J. M., RAGOUSSIS, J. & RATCLIFFE, P. J. 2009. Genome-wide association of hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha DNA binding with expression profiling of hypoxia-inducible transcripts. *J Biol Chem,* 284**,** 16767-75.

MOLLER-LEVET, C. S., BETTS, G. N., HARRIS, A. L., HOMER, J. J., WEST, C. M. & MILLER, C. J. 2009. Exon array analysis of head and neck cancers identifies a hypoxia related splice variant of LAMA3 associated with a poor prognosis. *PLoS Comput Biol,* 5**,** e1000571.

MORRIS, K. V. 2009. Long antisense non-coding RNAs function to direct epigenetic complexes that regulate transcription in human cells. *Epigenetics,* 4**,** 296-301.

MORRISSEY, J. P. & TOLLERVEY, D. 1995. Birth of the snoRNPs: the evolution of RNase MRP and the eukaryotic pre-rRNA-processing system. *Trends Biochem Sci,* 20**,** 78-82.

MURAWA, P., MURAWA, D., ADAMCZYK, B. & POLOM, K. 2014. Breast cancer: Actual methods of treatment and future trends. *Rep Pract Oncol Radiother,* 19**,** 165-72.

MUTZ, K. O., HEILKENBRINKER, A., LONNE, M., WALTER, J. G. & STAHL, F. 2013. Transcriptome analysis using next-generation sequencing. *Curr Opin Biotechnol,* 24**,** 22-30.

NADJI, M., GOMEZ-FERNANDEZ, C., GANJEI-AZAR, P. & MORALES, A. R. 2005. Immunohistochemistry of estrogen and progesterone receptors reconsidered: experience with 5,993 breast cancers. *Am J Clin Pathol,* 123**,** 21-7.

NAGAO, A., MITUYAMA, T., HUANG, H., CHEN, D., SIOMI, M. C. & SIOMI, H. 2010. Biogenesis pathways of piRNAs loaded onto AGO3 in the Drosophila testis. *RNA,* 16**,** 2503-15.

NAKAGAWA, S. & HIROSE, T. 2012. Paraspeckle nuclear bodies--useful uselessness? *Cell Mol Life Sci,* 69**,** 3027-36.

NALLAR, S. C. & KALVAKOLANU, D. V. 2013. Regulation of snoRNAs in cancer: close encounters with interferon. *J Interferon Cytokine Res,* 33**,** 189-98.

NEVO-CASPI, Y., AMARIGLIO, N., RECHAVI, G. & PARET, G. 2011. A-to-I RNA editing is induced upon hypoxia. *Shock,* 35**,** 585-9.

NG, H. H., ROBERT, F., YOUNG, R. A. & STRUHL, K. 2003. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell,* 11**,** 709-19.

NISHIMOTO, Y., NAKAGAWA, S., HIROSE, T., OKANO, H. J., TAKAO, M., SHIBATA, S., SUYAMA, S., KUWAKO, K., IMAI, T., MURAYAMA, S., SUZUKI, N. & OKANO, H. 2013. The long non-coding RNA nuclear-enriched abundant transcript 1_2 induces paraspeckle formation in the motor neuron during the early phase of amyotrophic lateral sclerosis. *Mol Brain,* 6**,** 31.

NORDSMARK, M., BENTZEN, S. M., RUDAT, V., BRIZEL, D., LARTIGAU, E., STADLER, P., BECKER, A., ADAM, M., MOLLS, M., DUNST, J., TERRIS, D. J. & OVERGAARD, J. 2005. Prognostic value of tumor oxygenation in 397 head and neck tumors after primary radiation therapy. An international multi-center study. *Radiother Oncol,* 77**,** 18-24.

O'NEIL, D., GLOWATZ, H. & SCHLUMPBERGER, M. 2013. Ribosomal RNA depletion for efficient use of RNA-seq capacity. *Curr Protoc Mol Biol,* Chapter 4**,** Unit 4 19.

OLCINA, M. M., FOSKOLOU, I. P., ANBALAGAN, S., SENRA, J. M., PIRES, I. M., JIANG, Y., RYAN, A. J. & HAMMOND, E. M. 2013. Replication stress and chromatin context link ATM activation to a role in DNA replication. *Mol Cell,* 52**,** 758-66.

OLENYUK, B. Z., ZHANG, G. J., KLCO, J. M., NICKOLS, N. G., KAELIN, W. G., JR. & DERVAN, P. B. 2004. Inhibition of vascular endothelial growth factor with a sequence-specific hypoxia response element antagonist. *Proc Natl Acad Sci U S A,* 101**,** 16768-73.

ORTIZ-BARAHONA, A., VILLAR, D., PESCADOR, N., AMIGO, J. & DEL PESO, L. 2010. Genome-wide identification of hypoxia-inducible factor binding sites and target genes by a probabilistic model integrating transcription-profiling data and in silico binding site prediction. *Nucleic Acids Res,* 38**,** 2332-45.

OZSOLAK, F. & MILOS, P. M. 2011. RNA sequencing: advances, challenges and opportunities. *Nat Rev Genet,* 12**,** 87-98.

OZSOLAK, F., PLATT, A. R., JONES, D. R., REIFENBERGER, J. G., SASS, L. E., MCINERNEY, P., THOMPSON, J. F., BOWERS, J., JAROSZ, M. & MILOS, P. M. 2009. Direct RNA sequencing. *Nature,* 461**,** 814-8.

PAGE, E. L., CHAN, D. A., GIACCIA, A. J., LEVINE, M. & RICHARD, D. E. 2008. Hypoxia-inducible factor-1alpha stabilization in nonhypoxic conditions: role of oxidation and intracellular ascorbate depletion. *Mol Biol Cell,* 19**,** 86-94.

PAN, Q., SHAI, O., LEE, L. J., FREY, B. J. & BLENCOWE, B. J. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet,* 40**,** 1413-5.

PAPANDREOU, I., CAIRNS, R. A., FONTANA, L., LIM, A. L. & DENKO, N. C. 2006. HIF-1 mediates adaptation to hypoxia by actively downregulating mitochondrial oxygen consumption. *Cell Metab,* 3**,** 187-97.

PARK, P. J. 2009. ChIP-seq: advantages and challenges of a maturing technology. *Nat Rev Genet,* 10**,** 669-80.

PARKHOMCHUK, D., BORODINA, T., AMSTISLAVSKIY, V., BANARU, M., HALLEN, L., KROBITSCH, S., LEHRACH, H. & SOLDATOV, A. 2009. Transcriptome analysis by strand-specific sequencing of complementary DNA. *Nucleic Acids Res,* 37**,** e123.

PEROU, C. M., SORLIE, T., EISEN, M. B., VAN DE RIJN, M., JEFFREY, S. S., REES, C. A., POLLACK, J. R., ROSS, D. T., JOHNSEN, H., AKSLEN, L. A., FLUGE, O., PERGAMENSCHIKOV, A., WILLIAMS, C., ZHU, S. X., LONNING, P. E., BORRESEN-DALE, A. L., BROWN, P. O. & BOTSTEIN, D. 2000. Molecular portraits of human breast tumours. *Nature,* 406**,** 747-52.

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**,** 189-97.

PETERS, L. J., O'SULLIVAN, B., GIRALT, J., FITZGERALD, T. J., TROTTI, A., BERNIER, J., BOURHIS, J., YUEN, K., FISHER, R. & RISCHIN, D. 2010. Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: results from TROG 02.02. *J Clin Oncol,* 28**,** 2996-3001.

PHILLIPS, M. I. & ZHANG, Y. C. 2000. Basic principles of using antisense oligonucleotides in vivo. *Methods Enzymol,* 313**,** 46-56.

PICKARD, M. R. & WILLIAMS, G. T. 2014. Regulation of apoptosis by long non-coding RNA GAS5 in breast cancer cells: implications for chemotherapy. *Breast Cancer Res Treat,* 145**,** 359-70.

PIRET, J. P., MOTTET, D., RAES, M. & MICHIELS, C. 2002. CoCl2, a chemical inducer of hypoxia-inducible factor-1, and hypoxia reduce apoptotic cell death in hepatoma cell line HepG2. *Ann N Y Acad Sci,* 973**,** 443-7.

POLLARD, P. J., LOENARZ, C., MOLE, D. R., MCDONOUGH, M. A., GLEADLE, J. M., SCHOFIELD, C. J. & RATCLIFFE, P. J. 2008. Regulation of Jumonji-domain-containing histone demethylases by hypoxia-inducible factor (HIF)-1alpha. *Biochem J,* 416**,** 387-94.

POLYMENIDOU, M., LAGIER-TOURENNE, C., HUTT, K. R., HUELGA, S. C., MORAN, J., LIANG, T. Y., LING, S. C., SUN, E., WANCEWICZ, E., MAZUR, C., KORDASIEWICZ, H., SEDAGHAT, Y., DONOHUE, J. P., SHIUE, L., BENNETT, C. F., YEO, G. W. & CLEVELAND, D. W. 2011. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. *Nat Neurosci,* 14**,** 459-68.

PUGH, C. W., TAN, C. C., JONES, R. W. & RATCLIFFE, P. J. 1991. Functional analysis of an oxygen-regulated transcriptional enhancer lying 3' to the mouse erythropoietin gene. *Proc Natl Acad Sci U S A,* 88**,** 10553-7.

PUSHKAREV, D., NEFF, N. F. & QUAKE, S. R. 2009. Single-molecule sequencing of an individual human genome. *Nat Biotechnol,* 27**,** 847-50.

QIU, M. T., HU, J. W., YIN, R. & XU, L. 2013. Long noncoding RNA: an emerging paradigm of cancer research. *Tumour Biol,* 34**,** 613-20.

RADOVICH, M. & RAGOUSSIS, J. 2013. Methods of quantifying microRNAs for hypoxia research: classic and next-generation. *Antioxid Redox Signal*.

RAKHA, E. A. & CHAN, S. 2011. Metastatic triple-negative breast cancer. *Clin Oncol (R Coll Radiol),* 23**,** 587-600.

RAKHA, E. A., REIS-FILHO, J. S. & ELLIS, I. O. 2010. Combinatorial biomarker expression in breast cancer. *Breast Cancer Res Treat,* 120**,** 293-308.

RAPISARDA, A. & MELILLO, G. 2009. Role of the hypoxic tumor microenvironment in the resistance to anti-angiogenic therapies. *Drug Resist Updat,* 12**,** 74-80.

RATCLIFFE, P. J. 2007. HIF-1 and HIF-2: working alone or together in hypoxia? *J Clin Invest,* 117**,** 862-5.

RATCLIFFE, P. J. 2013. Oxygen sensing and hypoxia signalling pathways in animals: the implications of physiology for cancer. *J Physiol,* 591**,** 2027-42.

RATCLIFFE, P. J., O'ROURKE, J. F., MAXWELL, P. H. & PUGH, C. W. 1998. Oxygen sensing, hypoxia-inducible factor-1 and the regulation of mammalian gene expression. *J Exp Biol,* 201**,** 1153-62.

RAVAL, R. R., LAU, K. W., TRAN, M. G., SOWTER, H. M., MANDRIOTA, S. J., LI, J. L., PUGH, C. W., MAXWELL, P. H., HARRIS, A. L. & RATCLIFFE, P. J. 2005. Contrasting properties of hypoxia-inducible factor 1 (HIF-1) and HIF-2 in von Hippel-Lindau-associated renal cell carcinoma. *Mol Cell Biol,* 25**,** 5675-86.

RAZ, T., KAPRANOV, P., LIPSON, D., LETOVSKY, S., MILOS, P. M. & THOMPSON, J. F. 2011. Protocol dependence of sequencing-based gene expression measurements. *PLoS One,* 6**,** e19287.

RAZAFSKY, D. & HODZIC, D. 2009. Bringing KASH under the SUN: the many faces of nucleo-cytoskeletal connections. *J Cell Biol,* 186**,** 461-72.

REAPER, P. M., GRIFFITHS, M. R., LONG, J. M., CHARRIER, J. D., MACCORMICK, S., CHARLTON, P. A., GOLEC, J. M. & POLLARD, J. R. 2011. Selective killing of ATM- or p53-deficient cancer cells through inhibition of ATR. *Nat Chem Biol,* 7**,** 428-30.

REN, S., LIU, Y., XU, W., SUN, Y., LU, J., WANG, F., WEI, M., SHEN, J., HOU, J., GAO, X., XU, C., HUANG, J., ZHAO, Y. & SUN, Y. 2013. Long noncoding RNA MALAT-1 is a new potential therapeutic target for castration resistant prostate cancer. *J Urol,* 190**,** 2278-87.

RENGANATHAN, A., KRESOJA-RAKIC, J., ECHEVERRY, N., ZILTENER, G., VRUGT, B., OPITZ, I., STAHEL, R. A. & FELLEY-BOSCO, E. 2014. GAS5 long non-coding RNA in malignant pleural mesothelioma. *Mol Cancer,* 13**,** 119.

RENTON, A., LLANOS, S. & LU, X. 2003. Hypoxia induces p53 through a pathway distinct from most DNA-damaging and stress-inducing agents. *Carcinogenesis,* 24**,** 1177-82.

RICHARD, P. & MANLEY, J. L. 2009. Transcription termination by nuclear RNA polymerases. *Genes Dev,* 23**,** 1247-69.

RICHTER, A., SANFORD, K. K. & EVANS, V. J. 1972. Influence of oxygen and culture media on plating efficiency of some mammalian tissue cells. *J Natl Cancer Inst,* 49**,** 1705-12.

RINN, J. L., KERTESZ, M., WANG, J. K., SQUAZZO, S. L., XU, X., BRUGMANN, S. A., GOODNOUGH, L. H., HELMS, J. A., FARNHAM, P. J., SEGAL, E. & CHANG, H. Y. 2007. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell,* 129**,** 1311-23.

ROBERTS, A., PIMENTEL, H., TRAPNELL, C. & PACHTER, L. 2011. Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics,* 27**,** 2325-9.

ROBINSON, M. D., MCCARTHY, D. J. & SMYTH, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics,* 26**,** 139-40.

ROSOK, O. & SIOUD, M. 2004. Systematic identification of sense-antisense transcripts in mammalian cells. *Nat Biotechnol,* 22**,** 104-8.

ROSSIGNOL, F., DE LAPLANCHE, E., MOUNIER, R., BONNEFONT, J., CAYRE, A., GODINOT, C., SIMONNET, H. & CLOTTES, E. 2004. Natural antisense transcripts of HIF-1alpha are conserved in rodents. *Gene,* 339**,** 121-30.

ROSSIGNOL, F., VACHE, C. & CLOTTES, E. 2002. Natural antisense transcripts of hypoxia-inducible factor 1alpha are detected in different normal and tumour human tissues. *Gene,* 299**,** 135-40.

RUAS, J. L., BERCHNER-PFANNSCHMIDT, U., MALIK, S., GRADIN, K., FANDREY, J., ROEDER, R. G., PEREIRA, T. & POELLINGER, L. 2010. Complex regulation of the transactivation function of hypoxia-inducible factor-1 alpha by direct interaction with two distinct domains of the CREB-binding protein/p300. *J Biol Chem,* 285**,** 2601-9.

RUPEC, R. A. & BAEUERLE, P. A. 1995. The genomic response of tumor cells to hypoxia and reoxygenation. Differential activation of transcription factors AP-1 and NF-kappa B. *Eur J Biochem,* 234**,** 632-40.

RUSSO, G., ZEGAR, C. & GIORDANO, A. 2003. Advantages and limitations of microarray technology in human cancer. *Oncogene,* 22**,** 6497-507.

SAID, H. M., POLAT, B., STAAB, A., HAGEMANN, C., STEIN, S., FLENTJE, M., THEOBALD, M., KATZER, A. & VORDERMARK, D. 2008. Rapid detection of the hypoxia-regulated CA-IX and NDRG1 gene expression in different glioblastoma cells in vitro. *Oncol Rep,* 20**,** 413-9.

SANCHEZ-FLORES, A. & ABREU-GOODGER, C. 2013. A Practical Guide to Sequencing Genomes and Transcriptomes. *Curr Top Med Chem*.

SANDERS, M. C., GOLDSTEIN, A. L. & WANG, Y. L. 1992. Thymosin beta 4 (Fx peptide) is a potent regulator of actin polymerization in living cells. *Proc Natl Acad Sci U S A,* 89**,** 4678-82.

SANGER, F., NICKLEN, S. & COULSON, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A,* 74**,** 5463-7.

SANTELLI, G., CALIFANO, D., CHIAPPETTA, G., VENTO, M. T., BARTOLI, P. C., ZULLO, F., TRAPASSO, F., VIGLIETTO, G. & FUSCO, A. 1999. Thymosin beta-10 gene overexpression is a general event in human carcinogenesis. *Am J Pathol,* 155**,** 799-804.

SASAKI, Y. T., IDEUE, T., SANO, M., MITUYAMA, T. & HIROSE, T. 2009. MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. *Proc Natl Acad Sci U S A,* 106**,** 2525-30.

SCHADT, E. E., TURNER, S. & KASARSKIS, A. 2010. A window into third-generation sequencing. *Hum Mol Genet,* 19**,** R227-40.

SCHIFFNER, S., ZIMARA, N., SCHMID, R. & BOSSERHOFF, A. K. 2011. p54nrb is a new regulator of progression of malignant melanoma. *Carcinogenesis,* 32**,** 1176-82.

SCHMIDT, L. H., SPIEKER, T., KOSCHMIEDER, S., SCHAFFERS, S., HUMBERG, J., JUNGEN, D., BULK, E., HASCHER, A., WITTMER, D., MARRA, A., HILLEJAN, L., WIEBE, K., BERDEL, W. E., WIEWRODT, R. & MULLER-TIDOW, C. 2011. The long noncoding MALAT-1 RNA indicates a poor prognosis in non-small cell lung cancer and induces migration and tumor growth. *J Thorac Oncol,* 6**,** 1984-92.

SCHODEL, J., BARDELLA, C., SCIESIELSKI, L. K., BROWN, J. M., PUGH, C. W., BUCKLE, V., TOMLINSON, I. P., RATCLIFFE, P. J. & MOLE, D. R. 2012. Common genetic variants at the 11q13.3 renal cancer susceptibility locus influence binding of HIF to an enhancer of cyclin D1 expression. *Nat Genet,* 44**,** 420-5, S1-2.

SCHODEL, J., MOLE, D. R. & RATCLIFFE, P. J. 2013. Pan-genomic binding of hypoxia-inducible transcription factors. *Biol Chem,* 394**,** 507-17.

SCHODEL, J., OIKONOMOPOULOS, S., RAGOUSSIS, J., PUGH, C. W., RATCLIFFE, P. J. & MOLE, D. R. 2011. High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. *Blood,* 117**,** e207-17.

SCHOPPMANN, S. F., FENZL, A., SCHINDL, M., BACHLEITNER-HOFMANN, T., NAGY, K., GNANT, M., HORVAT, R., JAKESZ, R. & BIRNER, P. 2006. Hypoxia inducible factor-1alpha correlates with VEGF-C expression and lymphangiogenesis in breast cancer. *Breast Cancer Res Treat,* 99**,** 135-41.

SCHRODER, O., BRYANT, G. O., GEIDUSCHEK, E. P., BERK, A. J. & KASSAVETIS, G. A. 2003. A common site on TBP for transcription by RNA polymerases II and III. *Embo Journal,* 22**,** 5115-5124.

SCHULZE, A. & HARRIS, A. L. 2012. How cancer metabolism is tuned for proliferation and vulnerable to disruption. *Nature,* 491**,** 364-73.

SEHGAL, A., LEE, C. Y. & ESPENSHADE, P. J. 2007. SREBP controls oxygen-dependent mobilization of retrotransposons in fission yeast. *PLoS Genet,* 3**,** e131.

SEIGNEURIC, R., STARMANS, M. H., FUNG, G., KRISHNAPURAM, B., NUYTEN, D. S., VAN ERK, A., MAGAGNIN, M. G., ROUSCHOP, K. M., KRISHNAN, S., RAO, R. B., EVELO, C. T., BEGG, A. C., WOUTERS, B. G. & LAMBIN, P. 2007. Impact of supervised gene signatures of early hypoxia on patient survival. *Radiother Oncol,* 83**,** 374-82.

SEMENZA, G. L. 1998. Hypoxia-inducible factor 1: master regulator of O2 homeostasis. *Curr Opin Genet Dev,* 8**,** 588-94.

SEMENZA, G. L. 2003. Targeting HIF-1 for cancer therapy. *Nat Rev Cancer,* 3**,** 721-32.

SEMENZA, G. L. 2009. Regulation of cancer cell metabolism by hypoxia-inducible factor 1. *Semin Cancer Biol,* 19**,** 12-6.

SEMENZA, G. L. 2010. HIF-1: upstream and downstream of cancer metabolism. *Curr Opin Genet Dev,* 20**,** 51-6.

SEMENZA, G. L. 2012a. Hypoxia-inducible factors in physiology and medicine. *Cell,* 148**,** 399-408.

SEMENZA, G. L. 2012b. Hypoxia-inducible factors: mediators of cancer progression and targets for cancer therapy. *Trends Pharmacol Sci,* 33**,** 207-14.

SEMENZA, G. L. 2014. Oxygen sensing, hypoxia-inducible factors, and disease pathophysiology. *Annu Rev Pathol,* 9**,** 47-71.

SEMENZA, G. L., NEJFELT, M. K., CHI, S. M. & ANTONARAKIS, S. E. 1991. Hypoxia-inducible nuclear factors bind to an enhancer element located 3' to the human erythropoietin gene. *Proc Natl Acad Sci U S A,* 88**,** 5680-4.

SHAIDA, N., CHAN, P., TURLEY, H., JONES, C. M., KANGA, S., RITCHIE, R. W., MALONE, P. R., HARRIS, A. L. & FOX, S. B. 2011. Nuclear localization of factor inhibitor hypoxia-inducible factor in prostate cancer is associated with poor prognosis. *J Urol,* 185**,** 1513-8.

SHAO, Q., MORGOUNOVA, E., JIANG, C., CHOI, J., BISCHOF, J. & ASHKENAZI, S. 2013. In vivo photoacoustic lifetime imaging of tumor hypoxia in small animals. *J Biomed Opt,* 18**,** 076019.

SHAO, X., TARNASKY, H. A., LEE, J. P., OKO, R. & VAN DER HOORN, F. A. 1999. Spag4, a novel sperm protein, binds outer dense-fiber protein Odf1 and localizes to microtubules of manchette and axoneme. *Dev Biol,* 211**,** 109-23.

SHEARWIN, K. E., CALLEN, B. P. & EGAN, J. B. 2005. Transcriptional interference--a crash course. *Trends Genet,* 21**,** 339-45.

SHEN, G., LI, X., JIA, Y. F., PIAZZA, G. A. & XI, Y. 2013. Hypoxia-regulated microRNAs in human cancer. *Acta Pharmacol Sin,* 34**,** 336-41.

SHENDURE, J. 2008. The beginning of the end for microarrays? *Nat Methods,* 5**,** 585-7.

SHIN, S. C., AHN DO, H., KIM, S. J., LEE, H., OH, T. J., LEE, J. E. & PARK, H. 2013. Advantages of Single-Molecule Real-Time Sequencing in High-GC Content Genomes. *PLoS One,* 8**,** e68824.

SHOJI, K., MURAYAMA, T., MIMURA, I., WADA, T., KUME, H., GOTO, A., OHSE, T., TANAKA, T., INAGI, R., VAN DER HOORN, F. A., MANABE, I., HOMMA, Y., FUKAYAMA, M., SAKURAI, T., HASEGAWA, T., ABURATANI, H., KODAMA, T. & NANGAKU, M. 2013. Sperm-associated

antigen 4, a novel hypoxia-inducible factor 1 target, regulates cytokinesis, and its expression correlates with the prognosis of renal cell carcinoma. *Am J Pathol,* 182**,** 2191-203.

SINGH, A., HAPPEL, C., MANNA, S. K., ACQUAAH-MENSAH, G., CARRERERO, J., KUMAR, S., NASIPURI, P., KRAUSZ, K. W., WAKABAYASHI, N., DEWI, R., BOROS, L. G., GONZALEZ, F. J., GABRIELSON, E., WONG, K. K., GIRNUN, G. & BISWAL, S. 2013. Transcription factor NRF2 regulates miR-1 and miR-206 to drive tumorigenesis. *J Clin Invest,* 123**,** 2921-34.

SINGLETARY, S. E. 2003. Rating the risk factors for breast cancer. *Ann Surg,* 237**,** 474-82.

SINICROPI, D., QU, K., COLLIN, F., CRAGER, M., LIU, M. L., PELHAM, R. J., PHO, M., DEI ROSSI, A., JEONG, J., SCOTT, A., AMBANNAVAR, R., ZHENG, C., MENA, R., ESTEBAN, J., STEPHANS, J., MORLAN, J. & BAKER, J. 2012. Whole transcriptome RNA-Seq analysis of breast cancer recurrence risk using formalin-fixed paraffin-embedded tumor tissue. *PLoS One,* 7**,** e40092.

SONG, H., SUN, W., YE, G., DING, X., LIU, Z., ZHANG, S., XIA, T., XIAO, B., XI, Y. & GUO, J. 2013. Long non-coding RNA expression profile in human gastric cancer and its clinical significances. *J Transl Med,* 11**,** 225.

SORLIE, T., PEROU, C. M., TIBSHIRANI, R., AAS, T., GEISLER, S., JOHNSEN, H., HASTIE, T., EISEN, M. B., VAN DE RIJN, M., JEFFREY, S. S., THORSEN, T., QUIST, H., MATESE, J. C., BROWN, P. O., BOTSTEIN, D., LONNING, P. E. & BORRESEN-DALE, A. L. 2001. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A,* 98**,** 10869-74.

SOTILLO, E. & THOMAS-TIKHONENKO, A. 2011. The long reach of noncoding RNAs. *Nat Genet,* 43**,** 616-7.

SOTIRIOU, C. & PICCART, M. J. 2007. Taking gene-expression profiling to the clinic: when will molecular signatures become relevant to patient care? *Nat Rev Cancer,* 7**,** 545-53.

SOULE, H. D., VAZGUEZ, J., LONG, A., ALBERT, S. & BRENNAN, M. 1973. A human cell line from a pleural effusion derived from a breast carcinoma. *J Natl Cancer Inst,* 51**,** 1409-16.

SPAN, P. N., BUSSINK, J., DE MULDER, P. H. & SWEEP, F. C. 2007. Carbonic anhydrase IX expression is more predictive than prognostic in breast cancer. *Br J Cancer,* 96**,** 1309; author reply 1310.

SPAN, P. N., BUSSINK, J., MANDERS, P., BEEX, L. V. & SWEEP, C. G. 2003. Carbonic anhydrase-9 expression levels and prognosis in human breast cancer: association with treatment outcome. *Br J Cancer,* 89**,** 271-6.

SRIBENJA, S., LI, M., WONGKHAM, S., WONGKHAM, C., YAO, Q. & CHEN, C. 2009. Advances in thymosin beta10 research: differential expression, molecular mechanisms, and clinical implications in cancer and other conditions. *Cancer Invest,* 27**,** 1016-22.

STAHLHUT, C. & SLACK, F. J. 2013. MicroRNAs and the cancer phenotype: profiling, signatures and clinical implications. *Genome Med,* 5**,** 111.

STARMANS, M. H., CHU, K. C., HAIDER, S., NGUYEN, F., SEIGNEURIC, R., MAGAGNIN, M. G., KORITZINSKY, M., KASPRZYK, A., BOUTROS, P. C., WOUTERS, B. G. & LAMBIN, P. 2012. The prognostic value of temporal in vitro and in vivo derived hypoxia gene-expression signatures in breast cancer. *Radiother Oncol,* 102**,** 436-43.

STEFANI, G. & SLACK, F. J. 2008. Small non-coding RNAs in animal development. *Nat Rev Mol Cell Biol,* 9**,** 219-30.

SUBRAMANIAN, A., TAMAYO, P., MOOTHA, V. K., MUKHERJEE, S., EBERT, B. L., GILLETTE, M. A., PAULOVICH, A., POMEROY, S. L., GOLUB, T. R., LANDER, E. S. & MESIROV, J. P. 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A,* 102**,** 15545-50.

SUN, Z., ASMANN, Y. W., KALARI, K. R., BOT, B., ECKEL-PASSOW, J. E., BAKER, T. R., CARR, J. M., KHREBTUKOVA, I., LUO, S., ZHANG, L., SCHROTH, G. P., PEREZ, E. A. & THOMPSON, E. A. 2011. Integrated analysis of gene expression, CpG island methylation, and gene copy number in breast cancer cells by deep sequencing. *PLoS One,* 6**,** e17490.

SUNWOO, H., DINGER, M. E., WILUSZ, J. E., AMARAL, P. P., MATTICK, J. S. & SPECTOR, D. L. 2009. MEN epsilon/beta nuclear-retained non-coding RNAs are up-regulated upon muscle differentiation and are essential components of paraspeckles. *Genome Res,* 19**,** 347-59.

SZENAJCH, J., WCISLO, G., JEONG, J. Y., SZCZYLIK, C. & FELDMAN, L. 2010. The role of erythropoietin and its receptor in growth, survival and therapeutic response of human tumor cells From clinic to bench - a critical review. *Biochim Biophys Acta,* 1806**,** 82-95.

TAKAHASHI, K., YAN, I. K., HAGA, H. & PATEL, T. 2014. Modulation of hypoxia-signaling pathways by extracellular long non-coding RNA regulator of reprogramming. *J Cell Sci*.

TAKANO, A., ISHIKAWA, N., NISHINO, R., MASUDA, K., YASUI, W., INAI, K., NISHIMURA, H., ITO, H., NAKAYAMA, H., MIYAGI, Y., TSUCHIYA, E., KOHNO, N., NAKAMURA, Y. & DAIGO, Y. 2009. Identification of nectin-4 oncoprotein as a diagnostic and therapeutic target for lung cancer. *Cancer Res,* 69**,** 6694-703.

TANG, X., LUCAS, J. E., CHEN, J. L., LAMONTE, G., WU, J., WANG, M. C., KOUMENIS, C. & CHI, J. T. 2012. Functional interaction between responses to lactic acidosis and hypoxia regulates genomic transcriptional outputs. *Cancer Res,* 72**,** 491-502.

TANG, Z., CHEN, W. Y., SHIMADA, M., NGUYEN, U. T., KIM, J., SUN, X. J., SENGOKU, T., MCGINTY, R. K., FERNANDEZ, J. P., MUIR, T. W. & ROEDER, R. G. 2013. SET1 and p300 act synergistically, through coupled histone modifications, in transcriptional activation by p53. *Cell,* 154**,** 297-310.

TANIMOTO, K., TSUCHIHARA, K., KANAI, A., ARAUCHI, T., ESUMI, H., SUZUKI, Y. & SUGANO, S. 2010. Genome-wide identification and annotation of HIF-1alpha binding sites in two cell lines using massively parallel sequencing. *Hugo J,* 4**,** 35-48.

TANO, K., MIZUNO, R., OKADA, T., RAKWAL, R., SHIBATO, J., MASUO, Y., IJIRI, K. & AKIMITSU, N. 2010. MALAT-1 enhances cell motility of lung adenocarcinoma cells by influencing the expression of motility-related genes. *FEBS Lett,* 584**,** 4575-80.

TARNASKY, H., GILL, D., MURTHY, S., SHAO, X., DEMETRICK, D. J. & VAN DER HOORN, F. A. 1998. A novel testis-specific gene, SPAG4, whose product interacts specifically with outer dense fiber protein ODF27, maps to human chromosome 20q11.2. *Cytogenet Cell Genet,* 81**,** 65-7.

TATUM, J. L., KELLOFF, G. J., GILLIES, R. J., ARBEIT, J. M., BROWN, J. M., CHAO, K. S., CHAPMAN, J. D., ECKELMAN, W. C., FYLES, A. W., GIACCIA, A. J., HILL, R. P., KOCH, C. J., KRISHNA, M. C., KROHN, K. A., LEWIS, J. S., MASON, R. P., MELILLO, G., PADHANI, A. R., POWIS, G., RAJENDRAN, J. G., REBA, R., ROBINSON, S. P., SEMENZA, G. L., SWARTZ, H. M., VAUPEL, P., YANG, D., CROFT, B., HOFFMAN, J., LIU, G., STONE, H. & SULLIVAN, D. 2006. Hypoxia: importance in tumor biology, noninvasive measurement by imaging, and value of its measurement in the management of cancer therapy. *Int J Radiat Biol,* 82**,** 699-757.

TCHOU-WONG, K. M., KIOK, K., TANG, Z., KLUZ, T., ARITA, A., SMITH, P. R., BROWN, S. & COSTA, M. 2011. Effects of nickel treatment on H3K4 trimethylation and gene expression. *PLoS One,* 6**,** e17728.

TEICHER, B. A. 1994. Hypoxia and drug resistance. *Cancer Metastasis Rev,* 13**,** 139-68.

TERNS, M. P. & TERNS, R. M. 2002. Small nucleolar RNAs: versatile trans-acting molecules of ancient evolutionary origin. *Gene Expr,* 10**,** 17-39.

THERIAULT, J. R., FELTS, A. S., BATES, B. S., PEREZ, J. R., PALMER, M., GILBERT, S. R., DAWSON, E. S., ENGERS, J. L., LINDSLEY, C. W. & EMMITTE, K. A. 2012. Discovery of a new molecular probe ML228: an activator of the hypoxia inducible factor (HIF) pathway. *Bioorg Med Chem Lett,* 22**,** 76-81.

THOMLINSON, R. H. & GRAY, L. H. 1955. The histological structure of some human lung cancers and the possible implications for radiotherapy. *Br J Cancer,* 9**,** 539-49.

THOMSON, T. & LIN, H. 2009. The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. *Annu Rev Cell Dev Biol,* 25**,** 355-76.

THRASH-BINGHAM, C. A. & TARTOF, K. D. 1999. aHIF: a natural antisense transcript overexpressed in human renal cancer and during hypoxia. *J Natl Cancer Inst,* 91**,** 143-51.

TIAN, B. & MANLEY, J. L. 2013. Alternative cleavage and polyadenylation: the long and short of it. *Trends Biochem Sci,* 38**,** 312-20.

TIVNAN, A., FOLEY, N. H., TRACEY, L., DAVIDOFF, A. M. & STALLINGS, R. L. 2010. MicroRNA-184-mediated inhibition of tumour growth in an orthotopic murine model of neuroblastoma. *Anticancer Res,* 30**,** 4391-5.

TO, K. K., SEDELNIKOVA, O. A., SAMONS, M., BONNER, W. M. & HUANG, L. E. 2006. The phosphorylation status of PAS-B distinguishes HIF-1alpha from HIF-2alpha in NBS1 repression. *EMBO J,* 25**,** 4784-94.

TOMES, L., EMBERLEY, E., NIU, Y., TROUP, S., PASTOREK, J., STRANGE, K., HARRIS, A. & WATSON, P. H. 2003. Necrosis and hypoxia in invasive breast carcinoma. *Breast Cancer Res Treat,* 81**,** 61-9.

TOMINAGA, E., YUASA, K., SHIMAZAKI, S. & HIJIKATA, T. 2013. MicroRNA-1 targets Slug and endows lung cancer A549 cells with epithelial and anti-tumorigenic properties. *Exp Cell Res,* 319**,** 77-88.

TRAPNELL, C., PACHTER, L. & SALZBERG, S. L. 2009. TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics,* 25**,** 1105-11.

TRAPNELL, C., WILLIAMS, B. A., PERTEA, G., MORTAZAVI, A., KWAN, G., VAN BAREN, M. J., SALZBERG, S. L., WOLD, B. J. & PACHTER, L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol,* 28**,** 511-5.

TRIPATHI, V., ELLIS, J. D., SHEN, Z., SONG, D. Y., PAN, Q., WATT, A. T., FREIER, S. M., BENNETT, C. F., SHARMA, A., BUBULYA, P. A., BLENCOWE, B. J., PRASANTH, S. G. & PRASANTH, K. V. 2010. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell,* 39**,** 925-38.

TRIPATHI, V., SHEN, Z., CHAKRABORTY, A., GIRI, S., FREIER, S. M., WU, X., ZHANG, Y., GOROSPE, M., PRASANTH, S. G., LAL, A. & PRASANTH, K. V. 2013. Long noncoding RNA MALAT1 controls cell cycle progression by regulating the expression of oncogenic transcription factor B-MYB. *PLoS Genet,* 9**,** e1003368.

TSAI, M. C., MANOR, O., WAN, Y., MOSAMMAPARAST, N., WANG, J. K., LAN, F., SHI, Y., SEGAL, E. & CHANG, H. Y. 2010. Long noncoding RNA as modular scaffold of histone modification complexes. *Science,* 329**,** 689-93.

TSAI, Y. P. & WU, K. J. 2013. Epigenetic regulation of hypoxia-responsive gene expression: Focusing on chromatin and DNA modifications. *Int J Cancer*.

TURNER, K. J., CREW, J. P., WYKOFF, C. C., WATSON, P. H., POULSOM, R., PASTOREK, J., RATCLIFFE, P. J., CRANSTON, D. & HARRIS, A. L. 2002. The hypoxia-inducible genes VEGF and CA9 are differentially regulated in superficial vs invasive bladder cancer. *Br J Cancer,* 86**,** 1276-82.

TZUR, Y. B., WILSON, K. L. & GRUENBAUM, Y. 2006. SUN-domain proteins: 'Velcro' that links the nucleoskeleton to the cytoskeleton. *Nat Rev Mol Cell Biol,* 7**,** 782-8.

UCHIDA, T., ROSSIGNOL, F., MATTHAY, M. A., MOUNIER, R., COUETTE, S., CLOTTES, E. & CLERICI, C. 2004. Prolonged hypoxia differentially regulates hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha expression in lung epithelial cells: implication of natural antisense HIF-1alpha. *J Biol Chem,* 279**,** 14871-8.

ULITSKY, I. & BARTEL, D. P. 2013. lincRNAs: genomics, evolution, and mechanisms. *Cell,* 154**,** 26-46.

UNRUH, A., RESSEL, A., MOHAMED, H. G., JOHNSON, R. S., NADROWITZ, R., RICHTER, E., KATSCHINSKI, D. M. & WENGER, R. H. 2003. The hypoxia-inducible factor-1 alpha is a negative factor for tumor therapy. *Oncogene,* 22**,** 3213-20.

VALENTIN, M. D., DA SILVA, S. D., PRIVAT, M., ALAOUI-JAMALI, M. & BIGNON, Y. J. 2012. Molecular insights on basal-like breast cancer. *Breast Cancer Res Treat,* 134**,** 21-30.

VAN 'T VEER, L. J., DAI, H., VAN DE VIJVER, M. J., HE, Y. D., HART, A. A., MAO, M., PETERSE, H. L., VAN DER KOOY, K., MARTON, M. J., WITTEVEEN, A. T., SCHREIBER, G. J., KERKHOVEN, R. M.,

ROBERTS, C., LINSLEY, P. S., BERNARDS, R. & FRIEND, S. H. 2002. Gene expression profiling predicts clinical outcome of breast cancer. *Nature,* 415**,** 530-6.

VAN DEN EYNDEN, G. G., VAN DER AUWERA, I., VAN LAERE, S. J., COLPAERT, C. G., TURLEY, H., HARRIS, A. L., VAN DAM, P., DIRIX, L. Y., VERMEULEN, P. B. & VAN MARCK, E. A. 2005. Angiogenesis and hypoxia in lymph node metastases is predicted by the angiogenesis and hypoxia in the primary tumour in patients with breast cancer. *Br J Cancer,* 93**,** 1128-36.

VAN DIJK, E. L., CHEN, C. L., D'AUBENTON-CARAFA, Y., GOURVENNEC, S., KWAPISZ, M., ROCHE, V., BERTRAND, C., SILVAIN, M., LEGOIX-NE, P., LOEILLET, S., NICOLAS, A., THERMES, C. & MORILLON, A. 2011. XUTs are a class of Xrn1-sensitive antisense regulatory non-coding RNA in yeast. *Nature,* 475**,** 114-7.

VANDER HEIDEN, M. G., CANTLEY, L. C. & THOMPSON, C. B. 2009. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science,* 324**,** 1029-33.

VARIA, M. A., CALKINS-ADAMS, D. P., RINKER, L. H., KENNEDY, A. S., NOVOTNY, D. B., FOWLER, W. C., JR. & RALEIGH, J. A. 1998. Pimonidazole: a novel hypoxia marker for complementary study of tumor hypoxia and cell proliferation in cervical carcinoma. *Gynecol Oncol,* 71**,** 270-7.

VAUPEL, P., BRIEST, S. & HOCKEL, M. 2002. Hypoxia in breast cancer: pathogenesis, characterization and biological/therapeutic implications. *Wien Med Wochenschr,* 152**,** 334-42.

VAUPEL, P. & MAYER, A. 2007. Hypoxia in cancer: significance and impact on clinical outcome. *Cancer Metastasis Rev,* 26**,** 225-39.

VAUPEL, P., MAYER, A., BRIEST, S. & HOCKEL, M. 2005. Hypoxia in breast cancer: role of blood flow, oxygen diffusion distances, and anemia in the development of oxygen depletion. *Adv Exp Med Biol,* 566**,** 333-42.

VENTURA, A. & JACKS, T. 2009. MicroRNAs and cancer: short RNAs go a long way. *Cell,* 136**,** 586-91.

VIGER, A. & DERVAN, P. B. 2006. Exploring the limits of benzimidazole DNA-binding oligomers for the hypoxia inducible factor (HIF) site. *Bioorg Med Chem,* 14**,** 8539-49.

VOLINIA, S., GALASSO, M., SANA, M. E., WISE, T. F., PALATINI, J., HUEBNER, K. & CROCE, C. M. 2012. Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA. *Proc Natl Acad Sci U S A,* 109**,** 3024-9.

VUMBACA, F., PHOENIX, K. N., RODRIGUEZ-PINTO, D., HAN, D. K. & CLAFFEY, K. P. 2008. Double-stranded RNA-binding protein regulates vascular endothelial growth factor mRNA stability, translation, and breast cancer angiogenesis. *Mol Cell Biol,* 28**,** 772-83.

WAHLESTEDT, C. 2006. Natural antisense and noncoding RNA transcripts as potential drug targets. *Drug Discov Today,* 11**,** 503-8.

WAHLESTEDT, C. 2013. Targeting long non-coding RNA to therapeutically upregulate gene expression. *Nat Rev Drug Discov,* 12**,** 433-46.

WANG, E. T., SANDBERG, R., LUO, S., KHREBTUKOVA, I., ZHANG, L., MAYR, C., KINGSMORE, S. F., SCHROTH, G. P. & BURGE, C. B. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature,* 456**,** 470-6.

WANG, G., CHEN, L., MENG, J., CHEN, M., ZHUANG, L. & ZHANG, L. 2013. Overexpression of microRNA-100 predicts an unfavorable prognosis in renal cell carcinoma. *Int Urol Nephrol,* 45**,** 373-9.

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**,** 5510-4.

WANG, K. C., YANG, Y. W., LIU, B., SANYAL, A., CORCES-ZIMMERMAN, R., CHEN, Y., LAJOIE, B. R., PROTACIO, A., FLYNN, R. A., GUPTA, R. A., WYSOCKA, J., LEI, M., DEKKER, J., HELMS, J. A. & CHANG, H. Y. 2011. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature,* 472**,** 120-4.

WANG, L. & WHEELER, D. A. 2014. Genomic sequencing for cancer diagnosis and therapy. *Annu Rev Med,* 65**,** 33-48.

WANG, W., HE, Y. F., SUN, Q. K., WANG, Y., HAN, X. H., PENG, D. F., YAO, Y. W., JI, C. S. & HU, B. 2014. Hypoxia-inducible factor 1alpha in breast cancer prognosis. *Clin Chim Acta,* 428**,** 32-7.

WANG, Z., GERSTEIN, M. & SNYDER, M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet,* 10**,** 57-63.

WARD, C., LANGDON, S. P., MULLEN, P., HARRIS, A. L., HARRISON, D. J., SUPURAN, C. T. & KUNKLER, I. H. 2013. New strategies for targeting the hypoxic tumour microenvironment in breast cancer. *Cancer Treat Rev,* 39**,** 171-9.

WARNECKE, C., WEIDEMANN, A., VOLKE, M., SCHIETKE, R., WU, X., KNAUP, K. X., HACKENBECK, T., BERNHARDT, W., WILLAM, C., ECKARDT, K. U. & WIESENER, M. S. 2008. The specific contribution of hypoxia-inducible factor-2alpha to hypoxic gene expression in vitro is limited and modulated by cell type-specific and exogenous factors. *Exp Cell Res,* 314**,** 2016-27.

WAYPA, G. B., MARKS, J. D., GUZY, R., MUNGAI, P. T., SCHRIEWER, J., DOKIC, D. & SCHUMACKER, P. T. 2010. Hypoxia triggers subcellular compartmental redox signaling in vascular smooth muscle cells. *Circ Res,* 106**,** 526-35.

WEBB, J. D., COLEMAN, M. L. & PUGH, C. W. 2009. Hypoxia, hypoxia-inducible factors (HIF), HIF hydroxylases and oxygen sensing. *Cell Mol Life Sci,* 66**,** 3539-54.

WEBER, D. G., JOHNEN, G., CASJENS, S., BRYK, O., PESCH, B., JOCKEL, K. H., KOLLMEIER, J. & BRUNING, T. 2013. Evaluation of long noncoding RNA MALAT1 as a candidate blood-based biomarker for the diagnosis of non-small cell lung cancer. *BMC Res Notes,* 6**,** 518.

WEI, C. L., WU, Q., VEGA, V. B., CHIU, K. P., NG, P., ZHANG, T., SHAHAB, A., YONG, H. C., FU, Y., WENG, Z., LIU, J., ZHAO, X. D., CHEW, J. L., LEE, Y. L., KUZNETSOV, V. A., SUNG, W. K., MILLER, L. D., LIM, B., LIU, E. T., YU, Q., NG, H. H. & RUAN, Y. 2006. A global map of p53 transcription-factor binding sites in the human genome. *Cell,* 124**,** 207-19.

WELLMANN, S., BETTKOBER, M., ZELMER, A., SEEGER, K., FAIGLE, M., ELTZSCHIG, H. K. & BUHRER, C. 2008. Hypoxia upregulates the histone demethylase JMJD1A via HIF-1. *Biochem Biophys Res Commun,* 372**,** 892-7.

WERNER, A. 2013. Biological functions of natural antisense transcripts. *BMC Biol,* 11**,** 31.

WHITE, R. J. 2008. RNA polymerases I and III, non-coding RNAs and cancer. *Trends Genet,* 24**,** 622-9.

WIGHT, M. & WERNER, A. 2013. The functions of natural antisense transcripts. *Essays Biochem,* 54**,** 91-101.

WILHELM, B. T., MARGUERAT, S., GOODHEAD, I. & BAHLER, J. 2010. Defining transcribed regions using RNA-seq. *Nat Protoc,* 5**,** 255-66.

WILSON, W. R. & HAY, M. P. 2011. Targeting hypoxia in cancer therapy. *Nat Rev Cancer,* 11**,** 393-410.

WILUSZ, J. E., JNBAPTISTE, C. K., LU, L. Y., KUHN, C. D., JOSHUA-TOR, L. & SHARP, P. A. 2012. A triple helix stabilizes the 3' ends of long noncoding RNAs that lack poly(A) tails. *Genes Dev,* 26**,** 2392-407.

WINTER, S. C., BUFFA, F. M., SILVA, P., MILLER, C., VALENTINE, H. R., TURLEY, H., SHAH, K. A., COX, G. J., CORBRIDGE, R. J., HOMER, J. J., MUSGROVE, B., SLEVIN, N., SLOAN, P., PRICE, P., WEST, C. M. & HARRIS, A. L. 2007. Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. *Cancer Res,* 67**,** 3441-9.

WONG, C. C., GILKES, D. M., ZHANG, H., CHEN, J., WEI, H., CHATURVEDI, P., FRALEY, S. I., WONG, C. M., KHOO, U. S., NG, I. O., WIRTZ, D. & SEMENZA, G. L. 2011. Hypoxia-inducible factor 1 is a master regulator of breast cancer metastatic niche formation. *Proc Natl Acad Sci U S A,* 108**,** 16369-74.

WONG, T. S., HO, W. K., CHAN, J. Y., NG, R. W. & WEI, W. I. 2009. Mature miR-184 and squamous cell carcinoma of the tongue. *ScientificWorldJournal,* 9**,** 130-2.

WONG, T. S., LIU, X. B., WONG, B. Y., NG, R. W., YUEN, A. P. & WEI, W. I. 2008. Mature miR-184 as Potential Oncogenic microRNA of Squamous Cell Carcinoma of Tongue. *Clin Cancer Res,* 14**,** 2588-92.

WU, D. & YOTNDA, P. 2011. Induction and testing of hypoxia in cell culture. *J Vis Exp*.

WU, Q., KIM, Y. C., LU, J., XUAN, Z., CHEN, J., ZHENG, Y., ZHOU, T., ZHANG, M. Q., WU, C. I. & WANG, S. M. 2008. Poly A- transcripts expressed in HeLa cells. *PLoS One,* 3**,** e2803.

WU, Z. J., MEYER, C. A., CHOUDHURY, S., SHIPITSIN, M., MARUYAMA, R., BESSARABOVA, M., NIKOLSKAYA, T., SUKUMAR, S., SCHWARTZMAN, A., LIU, J. S., POLYAK, K. & LIU, X. S. 2010. Gene expression profiling of human breast tissue samples using SAGE-Seq. *Genome Res,* 20**,** 1730-9.

WUTZ, A. 2011. Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nat Rev Genet,* 12**,** 542-53.

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**,** 7075-83.

XIA, X. & KUNG, A. L. 2009. Preferential binding of HIF-1 to transcriptionally active loci determines cell-type specific response to hypoxia. *Genome Biol,* 10**,** R113.

XIA, X., LEMIEUX, M. E., LI, W., CARROLL, J. S., BROWN, M., LIU, X. S. & KUNG, A. L. 2009. Integrative analysis of HIF binding and transactivation reveals its role in maintaining histone methylation homeostasis. *Proc Natl Acad Sci U S A,* 106**,** 4260-5.

XIANG, L., LIU, Z. H., HUAN, Q., SU, P., DU, G. J., WANG, Y., GAO, P. & ZHOU, G. Y. 2012. Hypoxia-inducible factor-2a is associated with ABCG2 expression, histology-grade and Ki67 expression in breast invasive ductal carcinoma. *Diagn Pathol,* 7**,** 32.

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**,** 338-45.

XU, C., YANG, M., TIAN, J., WANG, X. & LI, Z. 2011. MALAT-1: a long non-coding RNA and its important 3' end functional motif in colorectal cancer metastasis. *Int J Oncol,* 39**,** 169-75.

YANG, C., JIANG, L., ZHANG, H., SHIMODA, L. A., DEBERARDINIS, R. J. & SEMENZA, G. L. 2014a. Analysis of hypoxia-induced metabolic reprogramming. *Methods Enzymol,* 542**,** 425-55.

YANG, F., BI, J., XUE, X., ZHENG, L., ZHI, K., HUA, J. & FANG, G. 2012. Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells. *FEBS J,* 279**,** 3159-65.

YANG, F., HUO, X. S., YUAN, S. X., ZHANG, L., ZHOU, W. P., WANG, F. & SUN, S. H. 2013a. Repression of the long noncoding RNA-LET by histone deacetylase 3 contributes to hypoxia-mediated metastasis. *Mol Cell,* 49**,** 1083-96.

YANG, F., ZHANG, H., MEI, Y. & WU, M. 2014b. Reciprocal regulation of HIF-1alpha and lincRNA-p21 modulates the Warburg effect. *Mol Cell,* 53**,** 88-100.

YANG, L., DUFF, M. O., GRAVELEY, B. R., CARMICHAEL, G. G. & CHEN, L. L. 2011a. Genomewide characterization of non-polyadenylated RNAs. *Genome Biol,* 12**,** R16.

YANG, L., LIN, C., LIU, W., ZHANG, J., OHGI, K. A., GRINSTEIN, J. D., DORRESTEIN, P. C. & ROSENFELD, M. G. 2011b. ncRNA- and Pc2 methylation-dependent gene relocation between nuclear structures mediates gene activation programs. *Cell,* 147**,** 773-88.

YANG, Y., LIU, R., XIE, H., HUI, Y., JIAO, R., GONG, Y. & ZHANG, Y. 2013b. Advances in nanopore sequencing technology. *J Nanosci Nanotechnol,* 13**,** 4521-38.

YATES, L. A., NORBURY, C. J. & GILBERT, R. J. 2013. The long and short of microRNA. *Cell,* 153**,** 516-9.

YEO, B., TURNER, N. C. & JONES, A. 2014. An update on the medical management of breast cancer. *BMJ,* 348**,** g3608.

YIN, Q. F., YANG, L., ZHANG, Y., XIANG, J. F., WU, Y. W., CARMICHAEL, G. G. & CHEN, L. L. 2012. Long noncoding RNAs with snoRNA ends. *Mol Cell,* 48**,** 219-30.

YING, L., CHEN, Q., WANG, Y., ZHOU, Z., HUANG, Y. & QIU, F. 2012. Upregulated MALAT-1 contributes to bladder cancer cell migration by inducing epithelial-to-mesenchymal transition. *Mol Biosyst,* 8**,** 2289-94.

YING, Y. L., ZHANG, J., GAO, R. & LONG, Y. T. 2013. Nanopore-based sequencing and detection of nucleic acids. *Angew Chem Int Ed Engl,* 52**,** 13154-61.

YU, M., TING, D. T., STOTT, S. L., WITTNER, B. S., OZSOLAK, F., PAUL, S., CICILIANO, J. C., SMAS, M. E., WINOKUR, D., GILMAN, A. J., ULMAN, M. J., XEGA, K., CONTINO, G., ALAGESAN, B., BRANNIGAN, B. W., MILOS, P. M., RYAN, D. P., SEQUIST, L. V., BARDEESY, N., RAMASWAMY, S., TONER, M., MAHESWARAN, S. & HABER, D. A. 2012. RNA sequencing of pancreatic circulating tumour cells implicates WNT signalling in metastasis. *Nature,* 487**,** 510-3.

ZAGORSKA, A. & DULAK, J. 2004. HIF-1: the knowns and unknowns of hypoxia sensing. *Acta Biochim Pol,* 51**,** 563-85.

ZENTNER, G. E. & HENIKOFF, S. 2012. Surveying the epigenomic landscape, one base at a time. *Genome Biol,* 13**,** 250.

ZHANG, B., ARUN, G., MAO, Y. S., LAZAR, Z., HUNG, G., BHATTACHARJEE, G., XIAO, X., BOOTH, C. J., WU, J., ZHANG, C. & SPECTOR, D. L. 2012a. The lncRNA Malat1 is dispensable for mouse development but its transcription plays a cis-regulatory role in the adult. *Cell Rep,* 2**,** 111-23.

ZHANG, H., WONG, C. C., WEI, H., GILKES, D. M., KORANGATH, P., CHATURVEDI, P., SCHITO, L., CHEN, J., KRISHNAMACHARY, B., WINNARD, P. T., JR., RAMAN, V., ZHEN, L., MITZNER, W. A., SUKUMAR, S. & SEMENZA, G. L. 2012b. HIF-1-dependent expression of angiopoietin-like 4 and L1CAM mediates vascular metastasis of hypoxic breast cancer cells to the lungs. *Oncogene,* 31**,** 1757-70.

ZHANG, P., YAO, Q., LU, L., LI, Y., CHEN, P. J. & DUAN, C. 2014. Hypoxia-inducible factor 3 is an oxygen-dependent transcription activator and regulates a distinct transcriptional response to hypoxia. *Cell Rep,* 6**,** 1110-21.

ZHANG, Q., CHEN, C. Y., YEDAVALLI, V. S. & JEANG, K. T. 2013a. NEAT1 long noncoding RNA and paraspeckle bodies modulate HIV-1 posttranscriptional expression. *MBio,* 4**,** e00596-12.

ZHANG, Y., YANG, L. & CHEN, L. L. 2013b. Life without A tail: New formats of long noncoding RNAs. *Int J Biochem Cell Biol*.

ZHANG, Z., SUN, H., DAI, H., WALSH, R. M., IMAKURA, M., SCHELTER, J., BURCHARD, J., DAI, X., CHANG, A. N., DIAZ, R. L., MARSZALEK, J. R., BARTZ, S. R., CARLETON, M., CLEARY, M. A., LINSLEY, P. S. & GRANDORI, C. 2009. MicroRNA miR-210 modulates cellular response to hypoxia through the MYC antagonist MNT. *Cell Cycle,* 8**,** 2756-68.

ZHAO, G., JIAO, F., LIAO, Q., LUO, H., LI, H., SUN, L., BU, D., YU, K., ZHAO, Y. & CHEN, R. 2013. Genome-wide identification of cancer-related polyadenylated and non-polyadenylated RNAs in human breast and lung cell lines. *Sci China Life Sci,* 56**,** 503-12.

ZHENG, Y., JOSEFOWICZ, S. Z., KAS, A., CHU, T. T., GAVIN, M. A. & RUDENSKY, A. Y. 2007. Genome-wide analysis of Foxp3 target genes in developing and mature regulatory T cells. *Nature,* 445**,** 936-40.

ZHONG, H., DE MARZO, A. M., LAUGHNER, E., LIM, M., HILTON, D. A., ZAGZAG, D., BUECHLER, P., ISAACS, W. B., SEMENZA, G. L. & SIMONS, J. W. 1999. Overexpression of hypoxia-inducible factor 1alpha in common human cancers and their metastases. *Cancer Res,* 59**,** 5830-5.

ZHOU, X., SUN, H., CHEN, H., ZAVADIL, J., KLUZ, T., ARITA, A. & COSTA, M. 2010. Hypoxia induces trimethylated H3 lysine 4 by inhibition of JARID1A demethylase. *Cancer Res,* 70**,** 4214-21.