UNIVERSIDAD AUTÓNOMA DE MADRID

DEPARTAMENTO DE BIOQUÍMICA
FACULTAD DE MEDICINA

Regulation of Ephrin-A3 expression by hypoxia through a novel lncRNA-mediated mechanism.
A potential role in cancer metastasis.

LAURA GÓMEZ MALDONADO

MADRID, 2013

DEPARTAMENTO DE BIOQUÍMICA

FACULTAD DE MEDICINA

UNIVERSIDAD AUTÓNOMA DE MADRID

Regulation of Ephrin-A3 expression by hypoxia through a novel lncRNA-mediated mechanism.
A potential role in cancer metastasis.

Laura Gómez Maldonado

Licenciada en Biología, presenta esta memoria para optar al título de Doctor Internacional por la Universidad Autónoma de Madrid

Director de tésis:

Luis del Peso Ovalle, Profesor titular de la Universidad Autónoma de Madrid

Dr Luis del Peso Ovalle,

Profesor de la Universidad Autónoma de Madrid

CERTIFICA que:

Laura Gómez Maldonado, Licenciado en Biología, ha realizado bajo su dirección en el Departamento de Bioquímica, Facultad de Medicina- Instituto de Investigaciones Biomédicas "Alberto Sols" (CSIC-UAM) de Madrid, el trabajo titulado "Regulation of Ephrin-A3 expression by hypoxia through a novel lncRNA-mediated mechanism. A potential role in cancer metastasis."

El presente trabajo cumple, con los requisitos necesarios para ser presentado y defendido como Tesis Doctoral.

Madrid, 30 de Octubre de 2013

Fdo. Dr Luis del Peso Ovalle

A mis padres y Alberto

ACKNOWLEDGEMENTS

Quería dar las gracias a todas las personas que han formado parte de mi vida durante estos 4 años y que de una u otra manera han hecho que este proyecto sea posible.

En primer lugar quería dar las gracias a mi jefe Luis del Peso, por haberme dado la oportunidad de hacer la tesis bajo su tutela y enseñarme así a hacer ciencia con rigor y pasión. Gracias también por esos seminarios repletos de bollos, salidas de labo a ver magos y monólogos, sin olvidar el énfasis que pusiste en que aprendiéramos a usar Ubuntu y del que ahora soy fan incondicional. Ha habido de todo estos 4 años, momentos buenos y momentos menos buenos pero siempre has estado ahí con tu apoyo tanto a nivel profesional como a nivel personal, mil gracias.

Continuaré dando las gracias a todos los colaboradores durante mi tesis que han permitido que mi proyecto salga adelante, a Benilde por su apoyo en los experimentos de angiogenesis e invivo que me han ayudado tanto en este proyecto, así como su apoyo personal durante estos años. Olga Volpert, por darme la oportunidad de ir a su laboratorio y aprender no solo a nivel técnico sino también a nivel personal. A Adrian Harris y mi gente de Oxford, por darme la oportunidad de aprender de ellos y disfrutar de mi estancia allí.

Sin olvidar a aquellas personas que ya se fueron del labo, Amaya, Nuri, Bea y Sara, así como Jose, Valentí y Fonsi. Con vosotros empecé a dar mis primeros pasos en ciencia, gracias por estar siempre ahí para resolver mis dudas y por seguir ahí después de este tiempo, compartiendo nuestras días con barbacoas, bodorrios y mórulos de por medio. Es una alegría poder contar con gente como vosotras en mi vida.

A mis Chuchi-Cuquis del 2.5.1 y 2.5.2 que aun siguen por aquí aguantándome cada día, a Asun con sus "uff que pereza", a María por sus historias que empiezan con un "a una amiga de mi amiga Marta...", a Laura con sus "boludeces", a Olga por sus deliciosos postres y a LaPili por sus visitas al labo y su compañía en la estancia de Chicago, que a pesar de haberla conocido al final de la tesis, he cogido mucho cariño. Vosotras habéis hecho que cada mañana venga con una sonrisa a trabajar, os echaré infinitamente de menos. Se que será imposible encontrar otro trabajo en el que disfrute tanto de la compañía, ha sido una etapa maravillosa que recordaré con muchísimo cariño y del que me llevo buenas amistades.

A mi niñas Toñi, Yoly, Mar, Ana y Mamen, compañeras inseparables de la carrera, máster y de la vida. Hemos compartido infinidad de cosas durante todo este tiempo, esas cenitas en la marisquería bañadas con un buen Albariño, risas por doquier e historias miles, qué gran terapia!! Nuestra reciente incorporación, el pequeño Luca, la despedida y la boda.... Y lo que nos queda por compartir!! Nenas nos hacemos mayores pero siempre juntas, sois geniales y fundamentales para mi, os quiero un montón!!

A Gonzalo, Alvaro, Casas, Elena, Pilu, Ibi, Nacho y mi queridísimo "paisano" Kike. Ya son muchos los años compartiendo innumerables risas y aventuras... qué grandes momentos, me vienen a la cabeza miles!! Me siento muy afortunada de contar con esta pequeña gran familia, sin darnos cuenta hemos crecido juntos y habéis hecho que mis días fueran siempre dulces incluso cuando vinieron amargos, os quiero con locura!!

A Alberto, mi querido maridín, por intentar entender lo que hago en mi día a día con la efrina-A3, por ese apoyo que siempre me brindaste, en un principio desde la lejana China, India, Tailandia, mediante horas de skype, enseñándome lo importante que es hacer sacrificios por conseguir lo que uno quiere y más tarde desde la cercana convivencia con sesiones de mimos siempre que las fuerzas flaqueaban. Eres mi fuente de energía y alegrías, eres mi pilar!! Sin ti este proyecto tampoco habría sido posible.

A mi madre y hermana que siempre me motivaron para luchar por conseguir todas mis metas, e inculcaron en mí valores como el esfuerzo, la constancia y la importancia de estudiar y que me hacen ser hoy quien soy, gracias!! Y en especial a mi padre, que a pesar de no poder estar ya conmigo, siempre apoyó y se ilusionó con la idea de que su hija llegara a ser doctora, hoy puedo decir que nuestro sueño se ha hecho realidad.

ABSTRACT

Tumor hypoxia strongly correlates with metastasis and poor survival. However, the molecular mechanism by which hypoxia drives metastasis formation is not clear yet. EFNA3, a member of the ephrin type A ligands, was identified as a potential novel HIF target gene in an *in silico* search performed by our group (Ortiz-Barahona, Villar, Pescador, Amigo, & Del Peso, 2010). Although the involvement of EFNA3 in cancer has not been studied in detail, the expression of other ephrin members and their receptors are increased in human tumors and correlate with increased metastatic potential. Here, we show that hypoxia regulates Ephrin-A3 expression through a non-direct mechanism that involves the HIF-mediated transcriptional induction of a novel family of long-non-coding RNAs (lncRNAs) from the EFNA3 locus. In turn, these lncRNAs favor Ephrin-A3 accumulation acting at the postranscriptional level. As a consequence of this mechanism, EFNA3 transcript expression is significantly augmented in human renal carcinoma, where HIF is constitutively active, as compared with normal kidney tissue. Importantly, we found a strong correlation between high EFNA3 expression and shorter metastasis-free survival in breast cancer patients, suggesting a role for Ephrin-A3 in the promotion of metastatic behavior. In support of this hypothesis, Ephrin-A3 expression increased the metastatic potential of human breast cancer cells in an orthotopic xenotransplantation model. Importantly, by a combination of in vitro an in vivo experiments, we also found that Ephrin-A3 expression increases the ability of tumor cells to extravasate from the blood vessels into surrounding tissue. This function is in agreement with the role of ephrins as key mediators intercellular adhesion-repulsion and suggest a mechanisms for the promotion of metastasis. Altogether, our results suggest that hypoxia contributes to metastatic spread of cancer cells via HIF-mediated induction of EFNA3 lncRNAs and subsequent protein accumulation.

La presencia de hipoxia tumoral correlaciona fuertemente con la aparición de metástasis y baja supervivencia. Sin embargo, el mecanismo molecular por el cual la hipoxia da lugar a la formación de metástasis no está aun claro. EFNA3, miembro de las efrinas de tipo A, fue identificado como un nuevo gen potencial diana de HIF en un estudio *in silico* llevado a cabo por nuestro grupo (Ortiz-Barahona, A., 2010). Aunque la implicación de EFNA3 en cáncer no se ha estudiado en detalle, la expresión de otros miembros de esta familia así como sus receptores está aumentada en tumores humanos y correlaciona con un mayor potencial metastásico. En este trabajo, mostramos que la hipoxia regula la expresión de Efrina-A3 a través de un mecanismo indirecto que implica la inducción transcripcional, mediada por HIF, de una familia nueva de RNAs no codificantes largos (lncRNAs) desde el locus EFNA3. A su vez, estos lncRNAs favorecen la acumulación de Efrina-A3 a nivel postranscripcional. Como consecuencia de este mecanismo, la expresión de tránscritos de EFNA3 está incrementada significativamente en carcinoma renal humano, donde HIF se encuentra constitutivamente activo, en comparación con tejido renal normal. Cabe destacar, que hemos encontrado una fuerte correlación entre altos niveles de expresión de EFNA3 y una supervivencia libre de metástasis más corta en pacientes con cáncer de mama, sugiriendo que la efrina-A3 pudiera tener un papel en la promoción del comportamiento metastásico. En apoyo de esta hipótesis, la expresión de Efrina-A3 en células de cáncer de mama, aumentó su potencial metastásico en un modelo de xenotransplante ortotrópico. Además, por combinación de experimentos tanto *in vitro* como *in vivo*, también encontramos que la expresión de Efrina-A3 aumenta la capacidad de las células tumorales de extravasar de los vasos sanguíneos a los tejidos circundantes. Este función está de acuerdo con el papel de las efrinas como mediadores clave en la adhesión-repulsión intercelular y sugiere un mecanismo para la promoción de la metástasis. En conjunto, nuestros resultados sugieren que la hipoxia contribuye a la diseminación metastásica de las células cancerosas mediante la inducción mediada por HIF de los lncRNAs del locus EFNA3 y su acumulación proteica posterior.

INDEX

INDEX

ABBREVIATIONS

ARNT: aryl receptor nuclear translocator

BNIP3: bcl2/adenovirus E1B 19kDa protein-interacting protein 3

ccRCC: clear cell renal cell carcinoma

cDNA: complementary DNA

ChIP: chromatine immunoprecipitation

CTAD: carboxy-terminus transactivation domain

DAPI: 4',6-diamidino-2-phenylindole, is a fluorescent stain that binds strongly to A-T rich regions
in DNA

DMEM: dulbecco'sModified Eagle's Medium

DMOG: dimethyloxalylglycine

dsDNA: double-stranded DNA

EFNA3: ephrin A3 gen

EGLN: EGL-nine homologues

EPAS: endothelial PAS domain

EPO: erythropoietin gen

FIH: factor inhibiting HIF

GFP: green fluorescent protein

GPI: glycosylphosphatidyl-inositol

H&E: hematosin & eosin staining

HeLa: cervical cancer cells from Henrietta Lacks

HIF: hypoxia inducible factor

HRE: hypoxia response element

HUVEC: human umbilical vein endothelial cells

ID: inhibitory domain

lncRNAs: long noncoding RNAs

MDA-MB-231: human breast cancer cells

miRNAs: microRNAs

NC1: non-coding isoform 1

NC1s: non-coding isoform 1 short

NC2: non-coding isoform 2

NC2s: non-coding isoform 2 short

ncRNAs: non-protein-coding RNAs

NTAD: amino-terminus transactivation domain

OCT: ornithine transcarbamylase

ODD: oxygen dependent degradation domain

ORF: open reading frame

P4HA: prolyl 4-hydroxylase alpha subunit

PAS: (bHLH)-Per, ARNT and Sim domain

PDZ: PSD-95 post synaptic density protein, disc large, zona occludens tight junction protein

PET: paired-end diTag

PHD: prolyl-hydroxylase domain

pO2: oxygen partial pressure

Pol II: RNA polymerase II

PVDF: polyvinyl pirrolydone membranes

qPCR: quantitative PCR

RACE: rapid amplification of cDNA ends

RFP: red fluorescent protein

RT: reverse transcriptase

RTKs: receptor protein-tyrosine kinases

STT3A: subunit **o**f the oligosaccharyltransferase complex homolog A

TdT: terminal deoxynucleotidyl transferase

TFBS: transcription factor binding site

TSS: transcription start sites

VEGF: vascular endothelial growth factor

VHL: von Hippel Lindau

INTRODUCTION

1. OXYGEN HOMEOSTASIS

Oxygen is one of the most abundant molecules in our planet and its accumulation in the atmosphere 2000 million years ago was a determining factor for the evolution of life. The passage from a reducing atmosphere to the current oxidant atmosphere (21% oxygen), allowed the development of highly efficient metabolism which led to an explosion and diversification of aerobic life forms (Webster, K. A., 2007).

As complex organisms became dependent on this molecule to sustain their large energy demands, they had to develop mechanisms to ensure oxygen homeostasis by maintaining a fine balance between its supply and consumption (Semenza G. L., 2004). In simple organisms oxygen supply is ensured just by simple diffusion, but in more complex organisms its delivery to all tissues requires the development of sophisticated transport structures such as the tracheal or circulatory systems.

The evolution of these transport systems in animals imposed the need for specialized structures, such as the carotid body, pulmonary arteries and the neuroepithelial bodies in the respiratory tract, to monitor oxygen levels and ensure its systemic homeostasis. In vertebrates, a small decrease in the blood pO_2 (below 60 mmHg) is sensed, almost instantaneously, by the carotid body. Its Glomus cells, release the neurotransmitters, including dopamine, in a Ca^{2+} dependent manner, thereby stimulating respiration and cardiac beating (Lopez-Barneo, J., Ortega-Saenz, P., et al. 2009). On the other hand, muscles cells of the pulmonary arteries constrict in response to low oxygen level, resulting in blood flow distribution from poorly to well oxygenated areas, enabling optimum gas exchange between lungs and blood cells (Seta, K., et al. 2002).

In addition to the acute systemic responses described above, virtually all animal cells are able to adapt to more sustained restrictions in oxygen supply through a local chronic response that involves changes in gene expression. Specifically, this ubiquitous response is activated whenever oxygen supply does not meet oxygen consumption, a condition termed hypoxia. The induction of the hypoxia-induced transcriptional program aims to restore oxygen homeostasis by increasing oxygen supply to hypoxic regions through the induction of erythropoiesis (reviewed in Haase, V. H., 2010) and angiogenesis (Rey, S., & Semenza, G. L., 2010) and to reduce oxygen consumption by reprogramming cellular metabolism (reviewed in Wheaton, W. W., & Chandel, N. S., 2011)).

The first gene implicated in the cellular response to hypoxia was the erythropoietin gen (EPO), that encodes for a hormone responsible for the proliferation and survival of progenitor's red blood cells. Investigation on the Epo transcriptional regulation led to the identification of a *cis*-regulatory

element, known as *hypoxia response element* (HRE), in the 3'enhancer of the Epo gene (Semenza, G. L. and Wang, G. L., 1992). Subsequent investigation identified a transcription factor, *Hypoxia-Inducible Factor* (HIF), that bound led to the HRE and mediated in the response of Epo to hypoxia (Wang G. L., Jiang, B. H., Rue, E. A., & Semenza, G. L., 1995). Although the expression of EPO is restricted to liver tissue, HIF is ubiquitously expressed in all tissues, acting as the key regulator in response to hypoxia.

The transcriptional response to hypoxia, under the control of HIF, ensures oxygen homeostasis at tissue and cellular level, so only in exceptional situations such as embryonic development, tumors or inflammatory processes, an imbalance between supply and consumption of oxygen might occur.

2. HYPOXIA INDUCIBLE FACTOR (HIF)

HIF-1 has been identified in all metazoan species that have been analyzed from the simplest animal, *Trichoplax adherents* (Loenarz, C., 2011 and Rytkönen, K. T., 2011) to *Homo sapiens*, suggesting that the appearance of HIF-1 represented an adaptation that was essential to metazoan evolution (Semenza, G. L.,2004).

2.1 HIF STRUCTURE

HIF is a heterodimeric transcription factor composed by an oxygen-regulated alpha (HIFα) and an oxygen-insensitive beta subunit (HIFβ, also known as Aryl receptor nuclear translocator, ARNT). Both proteins belong to the basic helix-loop-helix (bHLH)-*Per, ARNT* and *Sim domain (*PAS) transcription factor family. The bHLH domain mediate DNA binding, whereas both, bHLH and PAS, domains are involved in the heterodimerization with other members of the family (Jiang, B. H., 1997). In contrast to HIFα, HIFβ shows a promiscuous behavior and forms dimers with other members of the bHLH-PAS superfamily in addition to HIFα.

There are three genes encoding for HIFα subunits: HIF1α, HIF2α (also known as EPAS) and HIF3α (figure 1). Functional studies have revealed that, in addition to the bHLH and PAS domains, all three HIFα proteins contain an oxygen dependent degradation domain (ODD) with critical proline residues (Pro 402/Pro 564 in HIF1α, Pro 405/Pro 531 in HIF2α and Pro 490 in HIF3α) that determine their accumulation in response to a decrease in oxygen availability (Huang L. E., et al, 1996) .

All HIFα subunits also contain transactivation domains, two in HIF1α and HIF2α (NTAD and

CTAD) and one in HIF3α (NTAD). The NTAD (amino acids 531–575 in Human HIF1α) and CTAD (amino acids 786 – 826) are separated by an inhibitory domain, ID (amino acids 576 –785) that represses CTAD under non-hypoxic conditionswith its repressive activity being prevented during hypoxia (Jiang, B. H., et al, 1997)
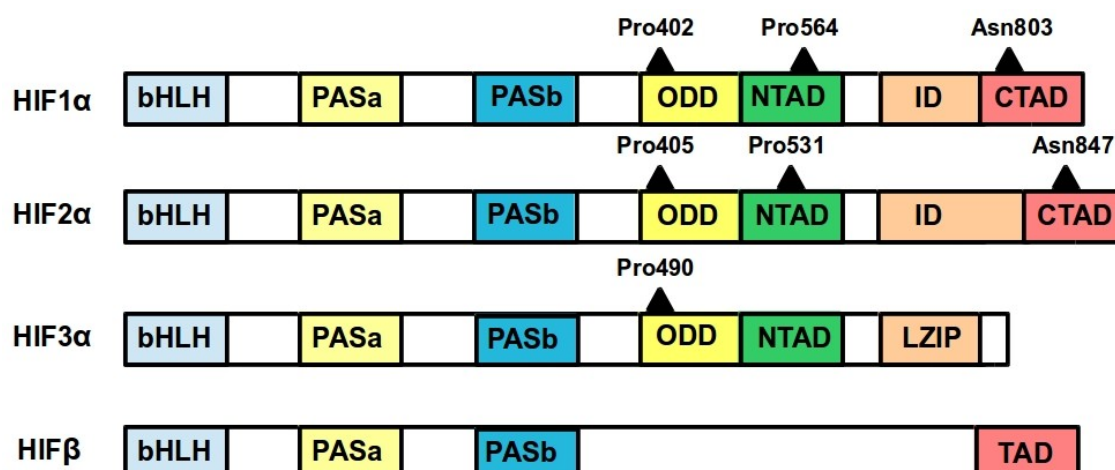
Figure 1: HIF family showing functional domains

Several mechanisms have been proposed to explain the detection pO2 cellular variations by HIF (Wang G. L., Jiang, B. H., Rue, E. A., & Semenza, G. L., 1995). However, the prevailing mode, summarized in figure 2, involves the activity of hydroxylases family, that are highly sensitive to pO2 variations. In normoxic conditions, HIFα is efficiently hydroxylated at two specific proline residues within the ODD. This reaction is catalyzed by a three-member family of prolyl-hydroxylases termed Prolyl-Hydroxylase Domain (PHD) EGL-nine homologues (EGLN), (Epstein, A. C., et al., 2001).

The hydroxylation in at least one of these prolines allows its recognition by an E3-ubiquitin ligase complex that targets HIFα for degradation via proteasome (Maxwell, P., H., et al. 1999; Ivan, M., et al. 2001; Jaakkola, P., Mole, D., et al. 2001). The product of the Von Hippel Lindau (VHL) tumor suppressor gene is the substrate recognition subunit of the E3-ubiquiting ligase complex (Lisztwan, J. 1999, Iwai K., 1999) that targets hydroxylated-HIF for proteasomal degradation.

In addition, under normoxia, another dioxygenase (Factor inhibiting HIF, FIH) catalyzes the oxygen-dependent hydroxylation of a conserved asparagine residue, located in the C-terminal transactivation domain. This postranslational modification prevents the interaction of HIFα with the

p300 coactivator, blunting its transcriptional activity (Mahon P. C., et al, 2001, Hewitson, K. S., et al. 2002; Lando, D., et al. 2002). However, the functional significance of this modification appears to be more subtle than the prolyl-hydroxylation of HIFα subunits, which is likely the main mechanism behind oxygen-dependent HIF transcription (Dayan F., 2006).

When oxygen becomes limiting, all these hydroxylation reactions are compromised and, as a consequence, HIFα subunit escapes degradation allowing its translocation to the nucleus and its interaction with HIFβ. Then, the heterodimer HIFα/HIFβ binds to the HRE (consensus sequence RCGTG), within regulatory regions of its target genes and recruits p300/CBP coactivator leading to their transcription (Min J. H., 2002, Epstein A. C., 2001).
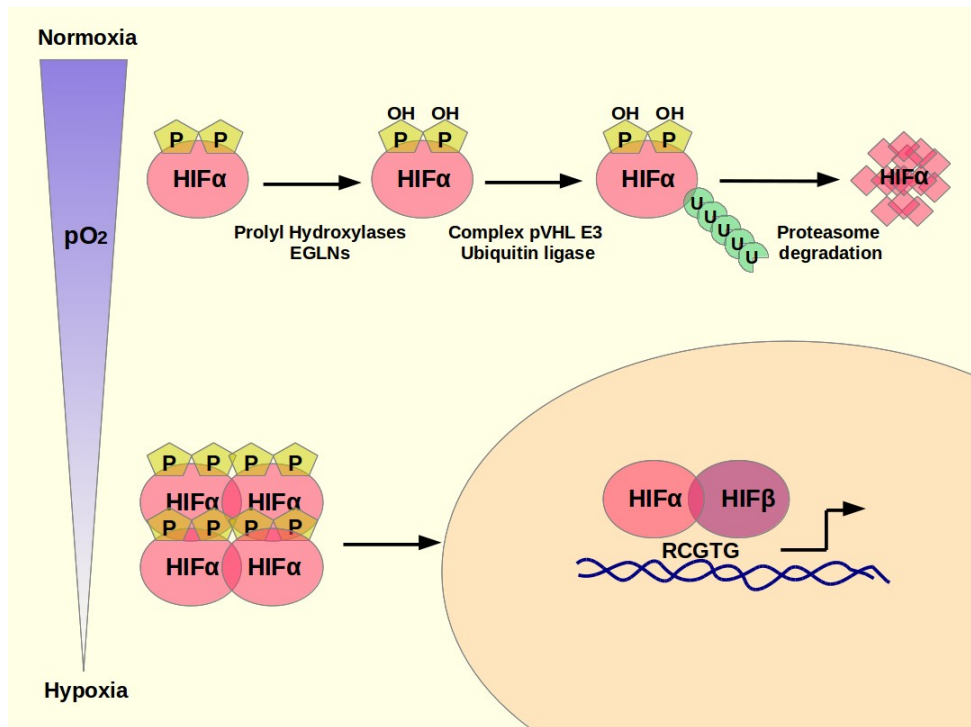
Figure 2: HIF signaling pathway

2.2 IDENTIFICATION OF HIF TARGET GENES:

Since most of the cellular adaptations to hypoxia are mediated by HIF, the identification of the complete set of genes regulated by this transcription factor is of key importance to fully understand the responses to reduced oxygen availability. This response entails the expression of multiple genes involved in different functions, as metabolic reprogramming, angiogenesis, chromatin remodeling,

proliferation and cellular differentiation (summarized in figure 3).

A large number of single gene studies have steadily increased the list of HIF-target genes (Wenger R. H., 2005). More recently, the application of high-throughput techniques such as global gene expression profiling, chromatin immunoprecipitation (Xia X., 2009 and Mole D. R., 2009) and computational prediction (Benita, Y., 2009 , Ortiz-Barahona, A., 2010), had resulted in a impressive increase in the number of known HIF target genes.
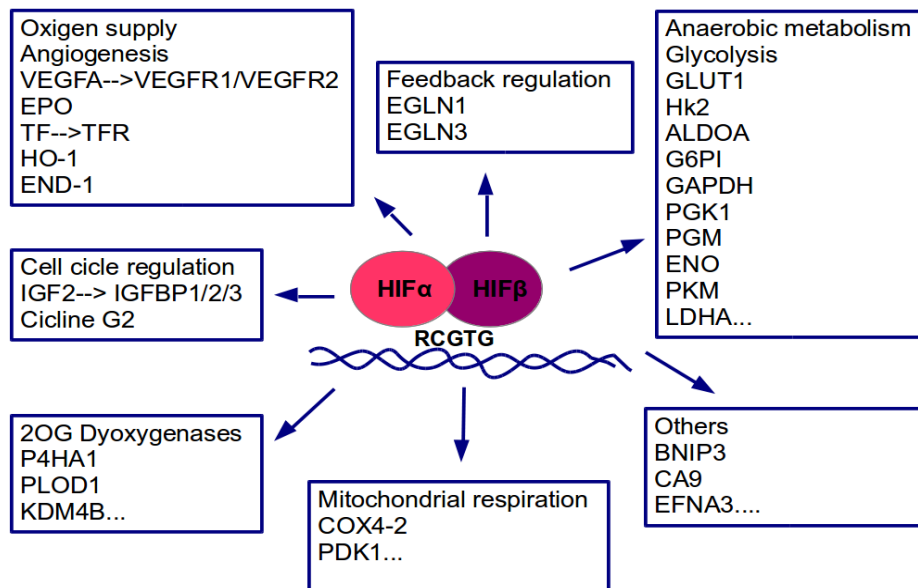
Figure 3 : The HIF transcriptome

With the aim of identifying novel HIF targets, our group developed a computational strategy based on the combination of phylogenetic footprinting and meta-analysis of publicly available gene expression profiles of cells exposed to hypoxia (Ortiz-Barahona, A., 2010). Comparison of the resulting candidates with experimentally validated HIF1α targets, indicated that this approach had a high sensitivity (78%) and specificity (97,8%).

Importantly, the implementation of this strategy led to the identification of 215 potential HIF targets, of which 152 were not previously involved in the response to hypoxia. Among the novel potential targets genes was EFNA3 gene, in which this project focuses, ranking in the top 51 position of the 215 candidates (Ortiz-Barahona, A., 2010)

2.3 THE ROLE OF HIF IN CANCER

In addition to orchestrating physiological responses to hypoxia, HIF has been involved in the ethiology and progression of multiple pathologies including cardiovascular and respiratory diseases, diabetes and cancer. In particular, the role of HIF in cancer has been extensively investigated. On one hand, hypoxia is frequently observed in solid tumors as a consequence of the rapid expansion of the cell mass compounded by aberrant and incomplete vascularization (Vaupel, P. & Mayer, A., 2007). On the other hand, the VHL tumor suppressor, whose protein product plays a key role in the control of HIFα stability (Maxwell, P., H., et al. 1999; Ivan, M., et al. 2001; Jaakkola, P., Mole, D., et al. 2001), is frequently lost in the progression of certain tumor types. Moreover, several lines of evidence point to HIFs, in particular EPAS, as key promoters of tumor progression upon VHL loss (Kaelin, W. G., 2007). Other oncogenes and tumor suppressor genes including PTEN, mTORC1, Ras, Akt and p53 also regulate HIFs activity (Semenza, G. L., 2012). Importantly, the correlation between tumor hypoxia and/or HIFα expression leads to poor prognosis and increased risk of metastasis which has been repeatedly demonstrated in diverse tumor types (Vaupel, P. & Mayer, A., 2007). Non-surprisingly, many of the HIF target genes are involved in biological processes that impact the metastatic spread of cancer cells, such as angiogenesis, epithelial-mesenchimal transition, cell motility, intra/extravasation and control of the pre-metastatic niche (Lu, X. & Kang, Y., 2010).

3. EPHRIN FAMILY

Ephrins are a large family of cell surface ligands that mediate intercellular adhesion and repulsion through interaction with the largest subfamily of receptor protein-tyrosine kinases (RTKs), the Eph receptors (Nievergall, E., Lackmann, M., & Janes, P. W., 2012; Pasquale E. B, 2010).

Ephrins are divided in two subclasses, A and B, according to their structure. Ephrin-A ligands (ephrin-A1 to -A5) are glycosylphosphatidyl-inositol (GPI)-linked membrane bound ligands, whereas ephrin-B ligands (ephrin-B1 to -B3) contain a single transmembrane domain and a short cytoplasmic tail.

Accordingly, Eph receptors are classified into two classes, A and B, based on the type of ephrin they bind to.There are nine EphA receptors (EphA1−8 and 10), and five EphB receptors (EphB1−4 and EphB6). In general, the binding between ephrin ligands and receptors is promiscuous within each A or B class although with different binding affinities (Gale N. W., et al., 1996). Still, EphB4 only

binds ephrin-B2 (Blits-Huizinga, C. T., et al., 2004). There is some inter-class cross-talk for exaple on how EphA4 binds to ephrin-B ligands (ephrin-B2–ephrin-B3) (Blits-Huizinga, C. T., et al., 2004) and EphB2 binds to ephrin-A5 (Himanen, J. P., et al., 2004).

Consistent with other types of RTKs, both A and B Eph class receptors contain a single transmembrane domain. The extracellular region of the Eph receptor is glycosylated, and consists of a ligand binding domain containing immunoglobulin-like motifs, followed by a cysteine-rich and two repeated fibronectin III-like domains. The intracellular region consists of a juxtamembrane region, a single tyrosine kinase domain, and PDZ binding motif (PSD-95 post synaptic density protein, Disc large, Zona occludens tight junction protein) within the non-catalytic region of the COOH-terminus. The kinase domain and juxtamembrane region contain tyrosine residues, and phosphorylation of these tyrosine residues creates docking sites for interactions with signaling proteins containing SH2/SH3 domains. The PDZ binding motif binds to PDZ domain-containing proteins, which are meant to serve as scaffolds for the assembly of multi-protein signaling complexes at the membrane (Cheng, N., 2002), (figure 4).
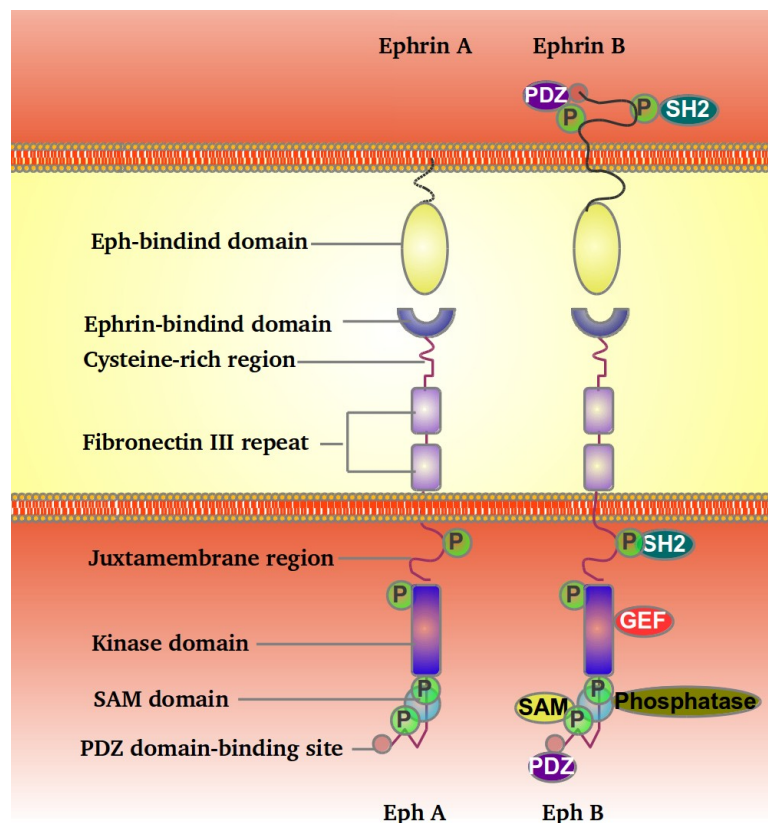
Figure 4. Structure of Eph receptors and ephrin ligands (Modified from Pasquale E., 2010).

3.1 EPHRIN CELLULAR SIGNALING

Eph/ephrin signaling is complex in several aspects. As indicated above, receptor-ligand binding is promiscuous, one ligand is able to bind multiple receptors. Another unique aspect of the ligand-receptor interaction in the Eph family is that the extent of receptor activation is dependent on the oligomerization state of the ligands (Cheng N., 2002). To facilitate the construction of macromolecular signaling complexes they are anchored to the plasma membrane at precise locations through protein-lipid interactions into lipid raft microdomains, which results in the association of downstream signaling proteins in exact positions for efficient interactions (Gauthier and Robbins, 2003; Simons and Toomre, 2000). Other complex aspect is that Eph/ephrin protein signaling results in the activation of several cytoplasmic downstream pathways (figure 6). Moreover, both A and B class ephrins complexes emanate bidirectional signaling. Forward signals that depend on Eph kinase activity propagate in the receptor-expressing cell, and reverse signals that depend on Src family kinases propagate in the ephrin-expressing cell (Pasquale E. B., 2010), but also, ephrin-dependent but kinase-independent Eph signals can occur (Gu, C. and Park, S., 2001, Miao, H., et al. 2005).
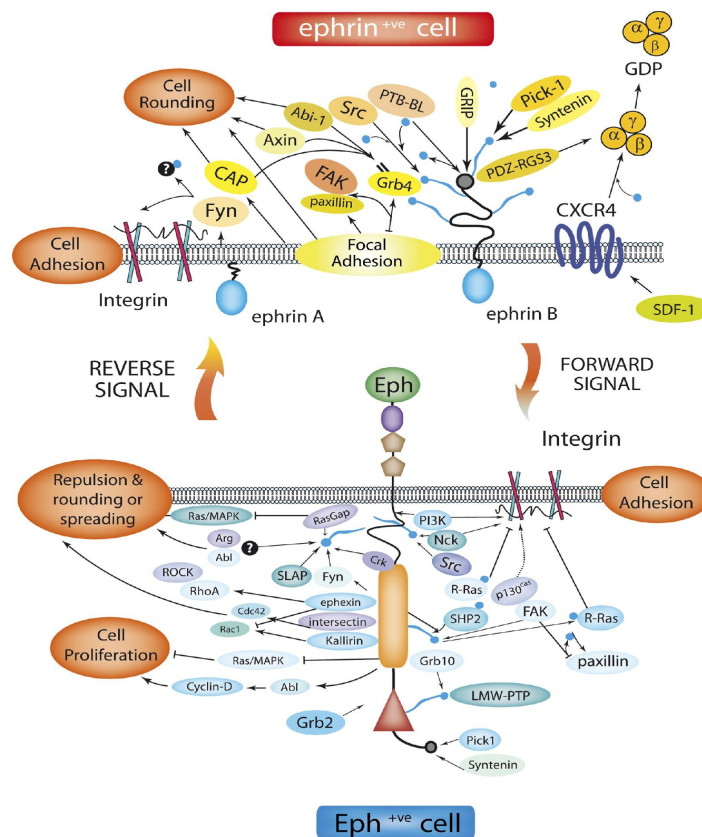

Figure 5. Eph/ephrin signaling pathways (from Coulthard, 2012).

Finally the attenuation and termination of the signaling involves proteolytic cleavage (Hattori, M., et al., 2000; Pascall, J. C., and Brown, K. D., 2004), receptor-mediated endocytosis (Zimmer, M., et al., 2003) and tyrosine phosphatase activity (Palmer, A., et al., 2002). Altogether, thiscomplex cellular signaling suggests that Eph/ephrin signaling require networks of interaction rather than simple linear pathways (Pawson, T. and Saxton, T. M., 1999).

3.2 EPHRIN FUNCTIONS

Ephrins mediate intercellular adhesion and repulsion through interaction with the Eph receptors (Nievergall E,, Lackmann M., & Janes P. W., 2012; Pasquale E. B., 2010), playing essential roles during development state where they guide migration and positioning of the cells for proper tissue patterning. Their function has been particularly well characterized in the nervous system, where Ephrins/Eph function as axon guidance molecules, and in cardiovascular system, where they control vasculogenesis (Nievergall, E,, Lackmann, M., & Janes, P. W., 2012).

In addition, recent reports suggest an emerging role of ephrins in the biology of stem cells (Sancho, E., et al., 2003) and tissue repair and maintenance (Hafner, C., et al., 2004). immune function (Wu, J. and Luo, H., 2005) and hematopoiesis (Okubo, T., et al., 2006).

3.2.1 THE ROLE OF EPH/EPHRIN IN CANCER

The protein tyrosine kinases are the largest family of oncogenes because of their central role in signal transduction pathways which control cell differentiation and proliferation (Bishop J. M., 1991; Blume-Jensen, P., and Hunter, T., 2001; Hunter, T., and Cooper, J. A., 1985). However, the Eph subfamily of RTKs are not classical oncogenes. The gene expression patterns of Eph and ephrins has been defined in a variety of benign and malignant human tumors and a number of subsequent studies report upregulation or downregulation of Eph/ephrin proteins in tumor cell lines and human cancers (Hafner, C., et al., 2004; Nakamoto, M., and Bergemann, A. D., 2002). This opposing function appears to be influenced by tissue type, oncogenic context, and ligand-independent versus ligand-dependent signaling (Pasquale E. B., 2010).

The growth of solid tumors is highly dependent on the ability to recruit blood vessels, that supply the tumor mass with growth factors, nutrients and oxygen necessary for survival and growth. It has

been proposed that Eph/ephrin expressed on the tumor and endothelial cells (Ogawa, K., 2000). regulate tumor angiogenesis, but the precise mechanism is unknown. Ephrins expressed on the tumor cells may function as contact-dependent organizing molecules to guide incoming vessels that express EphA2 receptor. Alternatively, angiogenic factors such as VEGF or TNF-α in the tumor microenvironment may induce expression and/or activation of ephrins in endothelial cells (Cheng, N., et al. 2002).

3.3 EPHRIN-A3

In spite of the intense investigation on the Eph/ephrin family, the regulation and function of some members remain relatively poorly characterized. Previous works addressed the role of Ephrin-A3, howevermost of them studied it within the context of the nervous system. Ephrin-A3 induces a repulsive interaction that regulates post synaptic morphology through the EphA4 receptor tyrosine kinase (Keith K. Murai, 2002). It has also been shown that Ephrin-A3 mediates the restriction of entorhinal axon terminals in the outer molecular layer of the dentate gyrus through the EphA5 (Stein E, 1999) and it is essential for maintaining EphA4 activation and normal spine morphology in vivo (Carmona, 2009). In addition to its role in the nervous system, Ephrin A3 has been involved in hair follicle development, where accelerates anagen development and increases the density of hair follicles (Yamanda, 2008).

Finally, in regards to its role in cancer mostly is unkown.. Ephrin-A3 mRNA expression has been found to be up-regulated 26-fold in squamous cell lung carcinoma (Hafner, 2004). Arecent study demonstrates the differential expression of Eprin-A3 between the highly metastatic and poorly metastatic colon cancer cells (Barderas et al., 2013).

4. NON-CODING RNA

Although genome transcription is pervasive, less than two percent of human genome codes for proteins and thus the majority of the genome gives rise to non-protein-coding RNAs (ncRNAs) (Djebali, S,. 2012), which are predicted to play a essential role in a variety of biological processes (Wilusz, J.E., 2009, Pauli, A., 2011). A subtype of them, defined as long non-coding RNAs (lncRNAs), with lengths of over 200 nucleotides, are transcribed by RNA polymerase II and processed as protein-coding RNAs. Many of them have small open reading frames but do not code for proteins (Yoon, J. H., 2012). Although the function of most long non-coding RNAs (lncRNAs)

is unknown, the number of characterized lncRNAs is growing and many publications suggest they play a major role in regulating gene expression during development, differentiation and human disease (Wilusz, J.E., 2009, Taft, R.J., 2010, Huarte, M. and Rinn, J. L., 2010).

lncRNAs regulate protein-coding gene expression at both, transcriptional and post-transcriptional levels (figure 8). These transcripts affect transcriptional regulation acting either in *cis*, when their effects are restricted to the chromosome/locus from which they are transcribed, or in *trans*, when they affect genes in other chromosomes/loci (Kornienko, A. E., 2013). As for the mechanism by which they regulate transcription, several possibilities have been proposed. Transcription silencing may involve epigenetic changes like promoter nucleosome repositioning (Martens, J.A., 2004), promoter histone modifications (Van Werven, F. J., 2012) and promoter DNA methylation (Tufarelli, C., 2003), but also there are transcriptional interference cases in the absence of chromatin changes at the silenced promoter (Latos, P.A., 2012). On the other hand, activation of gene expression by lncRNA may be mediated by blocking access of repressor complexes to chromatin (Cumberledge, S., 1990) or creating a permissive chromatin environment that supports enhancer function (Kornienko, A. E., 2013). This second model has not been experimentally tested yet

At postranscriptional level, lncRNAs act as competing endogenous RNAs to regulate microRNA levels as well as modulating pre-mRNA splicing, mRNA stability and translation by homologous base pairing (Kornienko, A. E., 2013, Yoon, Y. H., 2012).
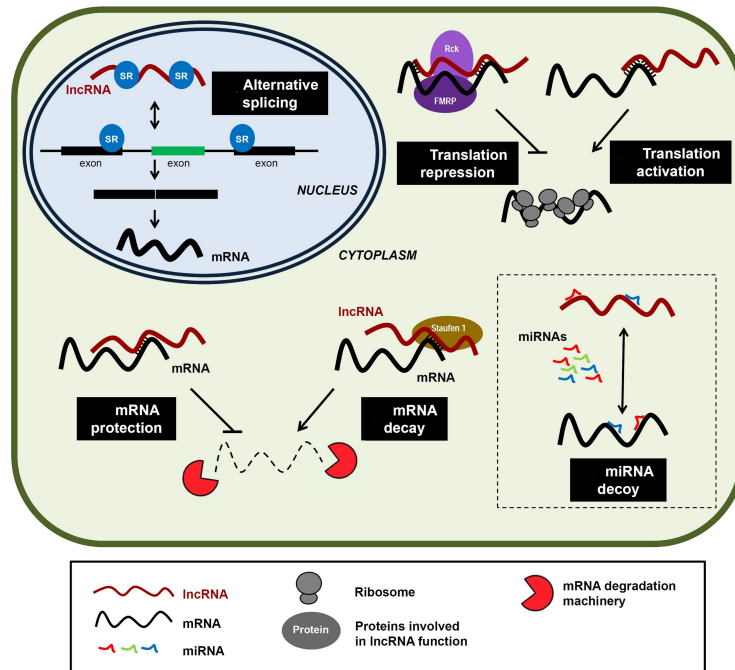
Figure 6: Levels of posttranscriptional gene regulation by lncRNAs (from Yoon, J. H., 2012)

In addition to lncRNAs, another group of non-coding RNA, microRNAs (miRNAs), play a key role in the regulation of gene expression. miRNAs constitute a family of short non-coding RNA molecules with lenghts varying from 20 to 25 nucleotides that regulate gene expression (Ambros, V., 2004; Bartel, D. P., 2004). In animals, miRNAs typically target sequences in the transcript 3'-UTRs that are only partially complementary to the miRNA, hereby causing messenger degradation and/or repression in translation of the mRNA (McCormick, R., 2010).

A pioneerstudy in 2008 showed that ephrin-A3 gene (EFNA3) was a direct target of miR-210 (Fasanaro, P., et al. 2008). The authors concluded that EFNA3 modulation by miR-210 had significant functional consequences for endothelial cell response to hypoxia, affecting cell survival, migration and differentiation. In addition, a recent report has shown that miR210 modulation of EFNA3 underlies the molecular mechanisms in preeclampsia (Zhang, Y., et al. 2012), further underscoring the role of miR210 in EFNA3 regulation.

OBJECTIVES

Is important to know the genes induced under hypoxic conditions in the tumor environment to the correct understanding of the tumorigenic response. Therefore we set the following objectives to be developed in our project:

1.   Validate EFNA as a HIF-target gene.

2.   Identify the mechanism of EFNA3 regulation by hypoxia.

3.   Study the functional role of EFNA3 regulation by hypoxia/HIF in the context of cancer.

MATERIALS AND METHODS

1. CELL CULTURE

Cell lines used in this project are detailed in Table 1, along with their growth medium and origin. All cell lines were cultured at 37ºC and 5% CO_2 in a humidified incubator.

Cell line	Origin	Culture medium
HeLa	Human cervical-carcinoma cell line	DMEM 10% FBS
HEK-293T	Human embryonic kidney epithelial cells immortalized with the T antigen of SV40	DMEM 10% FBS
MDA-MB-231	Human breast adenocarcinoma	DMEM 10% FBS
MCF7	Human breast adenocarcinoma	DMEM 10% FBS
HUVEC	Human Umbilical Vein Endothelial cells	EGM-2

Table 1. Cell lines

2. HYPOXIC CONDITIONS

For hypoxia treatments, cells were grown at 37ºC in a 1% O_2, 5% CO_2 94% N_2 gas mixture environmennt by use of a Whitley hypoxystation (don Whitley Scientific, UK). In addition, hypoxic conditions were mimicked by adding the Dimethyloxalylglycine (DMOG, Sigma), an inhibitor of HIF-prolyl-hydroxylases that prevents HIF degradation, to the cultures media at a 500 μM final concentration.

3. RNA EXTRACTION AND QUANTITATIVE RT-PCR

Total RNA was extracted and purified with the RNeasy Mini Kit (Qiagen) and treated with RNase free Dnase set (Qiagen, 79254). 1μg of total RNA of each sample was reverse-transcribed to cDNA (Transcriptor First Strand cDNA Synthesis kit, Roche) and 2μl of this reaction were diluted 1:4 as template for amplification reactions, carried out with the Power SYBR green PCR Master Mix (Applied Biosystems, 4367659) or Taqman Universal Master Mix II, no UNG (Applied Biosystems, 4440040), following the manufacturer's instructions.

PCR amplifications were carried out in a StepOne Realtime PCR System (Applied Biosystems), and data were analyzed with StepOne software. Expression levels were calculated using ΔΔCt method using β-actin as a reference.

For hnRNA quantification, total RNA was extracted, purified and treated with RNase free DNase (Qiagen). Total RNA (1μg/sample) was reverse transcribed to cDNA (Transcriptor First Strand cDNA Synthesis kit, Roche) and 1μl of cDNA was used as template in PCR amplification reactions

(see above). hnRNA was amplified by qPCR, with primers designed to amplify exon-intron boundary regions.

For chormatine immunoprecipitation (ChiP) qPCR, the threshold cycle (Ct) values for each sample were interpolated in a standard curve of input DNA dilutions to obtain % of input absolute values. The enrichment of HIF1α binding or Polymerase II binding at the target loci was calculated as the ratio of the target sequence precipitated (measured as % of input) in hypoxic over normoxic ChIP samples (% of input hypoxia / % of input normoxia).

For microRNA quantification, a specific retrotranscription protocol is required. We used Megaplex RT Primers Pools AV2.1 (Applied Biosistem, 4399966) together with the Multiscribe enzyme and 1μl of the product was used as template in PCR amplification reactions (see above).

Name	Assay	Forward sequence 5'-3'	Reverse sequence 3'-5'
1 HRE	ChiP HIF1α	CTTCTCCTTCCCCTCATTCC	CAGGCTGAATTTCCCAGAAG
2 HRE	ChiP HIF1α	CGAGCACCTCCCTAGAAGGA	CTCACCCTAGCCGGGCTTAT
3 HRE	ChiP HIF1α	TCCCCCCACCAAACCTATTC	TTGGAGCACCTCAGCGTTCT
4-8 HRE	ChiP HIF1α	GAGGGTCTCTGCCCCTTG	CCCCAGGTCTGTCAAAGGAG
9-10 HRE	ChiP HIF1α	GATGGCGAGGATTTGACAAGTT	AAGACCAGGGAGTAGGGAAAGG
11-12 HRE	ChiP HIF1α	GGAAGCTCGGAGGAAAAGTC	GAAGCGGCTCAGAGAGAAGA
EGLN3 (+ control)	ChiP HIF1α	GGTGTGCTCGGGTGTG	CGTGGAGGACTGGCTCTAAG
EGLN3 (- control)	ChiP HIF1α	ACGGGAGGCACTCGGAG	CCCTTAACGTTGACTTTCGCTC
Prom1	ChiP Pol II	CATGCGGTGTACTGGAACAG	GGATCCCCCAGGACTTCTC
Prom2	ChiP Pol II	GGAAGCTCGGAGGAAAAGTC	GAAGCGGCTCAGAGAGAAGA
P4HA (+ control)	ChiP Pol II	GAGCCCGTTAGCCCTTTTAT	GGTGTGATCGAGCTCACGTA
STT3 (- control)	ChiP Pol II	GAGCGCGGAAAGAACGTG	FGCAAGGGCCTATTTCAGCGTA

Name	Assay	Forward sequence 5'-3'	Reverse sequence 3'-5'
Exon 1+2	EFNA3 expression	AACCGGCATGCGGTGTA	ATCCAGATAGTCGTTCACGTTCA
Exon 4+5	EFNA3 expression	CACTCTCCCCCAGTTCACCAT	CGCTGATGCTCTTCTCAAGCT
Ms_Coding EFNA3 Isoform	EFNA3 expression	TGGAACAGCTCCAATCAGCA	GAGCTGTTGTAGTGCGGACA
Ms_Non-Coding EFNA3 Isoform	EFNA3 expression	CACTCTCCCCCAGTTCACCAT	CGCTGATGCTCTTCTCAAGCT
hnRNA coding EFNA3	hnRNA	ACTGGAACAGCTCCAACCAG	GCTCTGGTCTCCCAGGACTC
hnRNA lncEFNA3	hnRNA	GGCACTGATACTTCCTACCCTG	GGCACCTGAGGGTTCTCTCC
hnRNA Actin	hnRNA	CCCAGCACAATGAAGATCAA	GTGAGGACCCTGGATGTGAC
hnRNA EGLN3	hnRNA	TGGGAAGTCGACATACAACG	ATCGACAGGCTGGTCCTCTA
Exon 1	Sequencial PCR	AGCTGGGAAGCGGAGAAG	CTGGTTGGAGCTGTTCCAGT
Exon 2.1	Sequencial PCR	CTGCGGCGAGAGGGCTA	TGTAGTGCGGGCAGTAAATATCC
Exon 2.2	Sequencial PCR	AAGCGCTGGGAGTGCAAC	GTGGAACTCGTAGCCCAGAG
Intron 2	Sequencial PCR	AGCTCAGACGAGGTCGTGGGG	GTGGGCGTGGAGAAACAGCGA
Exon 4	Sequencial PCR	CATCGCACTCCGGGGAGAAG	CACGTTGATCTTCACATTGGGG
Exon 5	Sequencial PCR	GCTTGAGAAGAGCATCAGCG	CTAGGAGGCCAAGAACGTCA
3' UTR 1	Sequencial PCR	TCCTCCCATGGCTAGAAGTG	GTCCACTACAGTGCCCTACG
3' UTR 2	Sequencial PCR	GCAATAAGCACGTCCTCCTC	CAGGGGGTTAAAGAGGGAAG

Table 2. Primers

27

Taqman	Reference
Taqman 1+2 codingEFNA3	Hs00191913_m1
Taqman 4+5 lncEFNA3	Hs00900213_g1
EPAS1	Hs01026149_m1
HIF1α	Hs00936368_m1
β Actin	4326315E

Table 3. Taqman probes, all were supplied for Applied Biosistems

4. RNA INTERFERENCE

The cells were transfected with the indicated siRNAs using Lipofectamine RNAiMax (Invitrogen 13778) according to the manufacturer's instructions. Knockdown efficiency was determined by real time quantitative PCR analysis and Western blot. We used the following commercial siRNAs: HIF1α siRNA (Santa Cruz, sc-44225), EPAS-1 siRNA (Santa Cruz sc-35316), Ephrin-A3 (Santa Cruz sc-39430) and siRNA-B (Santa Cruz sc-44230) as a negative control. In all cases, siRNAs were used at 100nM concentration.

5. WESTERN BLOT

Whole-cell lysates were extracted in RIPA buffer (50mM Tris-HCl pH 8, 0,1% SDS, 150mM NaCl, 0,02% NaN3, 1% NP40, 1gr Sodium Deoxycholate) containing protease inhibitors (Complete ULTRA tablet, Roche 06538304001). Lysates were passed through a 27-gauge needle to shear genomic DNA.Cellular debries were removed by centrifugation at 4°C for 10 minutes at 14,000×g. Proteins were then resolved in 10% SDS-polyacrylamide gels and subsequently transferred to polyvinyl pirrolydone (PVDF) membranes. Membranes, were blocked with 5% BSA in TBS-T (50mM Tris, pH 7.6, 150mM NaCl, 0.1% Tween 20) for 1 hour and incubated overnight at 4°C with the requisite antibodies (Table 3). Next day, the membrane was washed 3 times with TBS-T for a total of 45 minutes and incubated with the secondary antibody in TBS-T with 5% BSA for 1 hour at room temperature. Then, membrane was washed again three times with TBS-T for 45 minutes and developed with ECL (GE Healthcare, Buckinghamshire, UK) and exposed to medical X-ray film (Agfa, ENKMV).

Antibodies	Species	Company	Dilution
Ephrin A3	Mouse	Santa cruz (sc-73954)	1:1000
HIF 1α	Mouse	BD Bioscience (610958)	1:1000
actin	Goat	Santa Cruz (sc-1616)	1:1000
anti-mouse		Promega (W402B)	1:5000
anti-goat		Santa cruz (sc-2020)	1:5000

Table 4. Antibodies

6. DEGLYCOSILATION

25μg lysates were mixed with 1μl 10X deglycosylation buffer (5%SDS, 0.4M DTT) and H_2O up to a total reaction volume of 10μl. The glycoprotein was denatured at 100ºC for 10 minutes and incubated for additional hour at 37ºC with 1μl PNGaseF (Biolabs, P0705S), in a total volume adjusted to 20μl with 2μl of 10X G7 Reaction buffer (0.5M Sodium Phosphate pH 7.5), 2μl of 10%NP40.

7. RAPID AMPLIFICATION OF of cDNA ENDS (RACE)

Rapid Amplification of cDNA Ends (RACE) is a procedure for amplification of nucleic acid sequences from a messenger RNA template between a defined internal site and unknown sequences at either the 3' or the 5' ends of the mRNA.

RACE was performed with the FirstChoice RLM-RACE Kit (Ambion) following manufacturer's instructions and using the indicated specific primers. All PCR fragments of interest were cloned into pCR2.1 vector (Invitrogen) and sequenced.

8. HIGH RESOLUTION GEN EXPRESSION PROFILING

A high resolution gene expression profiling was performed as described before (Dolken, L. et all.,2008). This assay is based on metabolic labeling of newly transcribed RNA using 4-thiouridine (4sU-tagging), a naturally occurring uridine derivative, and thus provides a means to detect newly transcribed transcripts with minimal interference in cell growth and gene expression. For these experiments we exposed HeLa cells to 400mM 4sU in normoxic or hypoxic conditions for 2 hours to allow 4sU-triphosphate incorporation into the newly transcribed RNA. Following isolation of total cellular RNA, the 4sU-labeled RNA fraction is specifically biotinylated by generating a disulfide bond between biotin and the thiol group in the 4sU-moiety, then "total cellular RNA" can

be quantitatively separated into labeled ('newly transcribed') and unlabeled ('pre-existing') RNA with high purity using streptavidin-coated magnetic beads. Finally, labeled RNA is recovered from the beads by simply adding a reducing agent cleaving the disulfide bond and releasing the newly transcribed RNA from the beads.
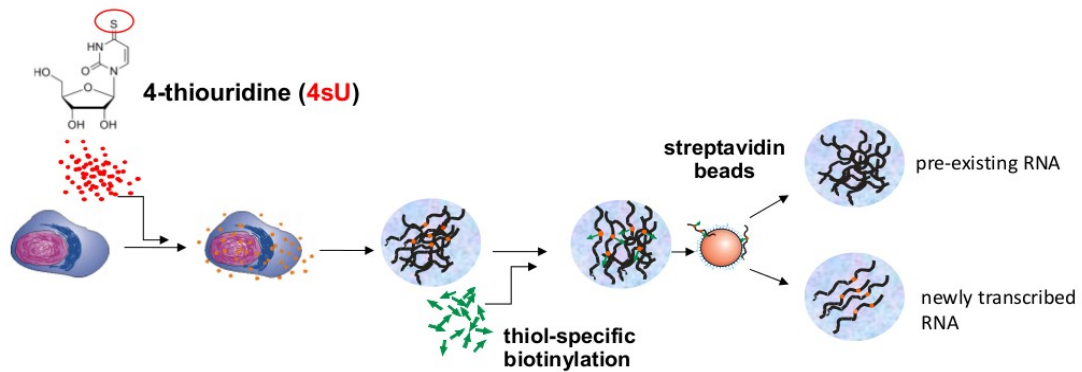
Figure 7. High resolution gene expression profiling squeme

9. HALF LIFE ASSAY

HeLa cells were seeded in a 6cm² plates ($1,12 \times 10^6$ cells/well) and 8 hours later are incubated in normoxia, hypoxia or in presence of 500 μM of DMOG for 12 hours. Then cells were treated with 5μg/ml of Actinomycin D and processed for RNA insolation at different time points.

10. ChIP (CHROMATIN IMMUNOPRECIPITATION)

For ChIP, EZ ChIP-Chromatin Inmunoprecipitation Kit (Millipore) was used following the manufacturer's instructions. Briefly, HeLa cells were grown to 85% confluence on 10 cm plates before they were exposed to hypoxia or left in normoxic conditions for 8 hours. Following treatments, cells were crosslinked with 1% formaldehyde for 12 min at 4ºC. Cell lysis was achieved by scraping in 1 ml of lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris/HCl, pH 8.1, and a protease inhibitor cocktail, Roche). Cell lysates were sonicated to shear the DNA to fragments between 200 and 1500 bp. Sonication is a critical step for successful chromatin immunoprecipitation. Therefore, quality and homogeneous size distribution of sonicated samples was assayed by DNA electrophoresis for every experiment, and only experiments that showed homogeneous results across all samples (from normoxia and hypoxia treatments) were continued. Lysates were then immunoprecipitated overnight at 4ºC with antibodies against RNA Polymerase II (Abcam 5408) or polyclonal serum for HIF1α (Abcam, ab2185) and normal Mouse IgG (Sigma M8695) or whole rabbit serum, respectively, as negative controls.

The DNA was finally purified by phenol–chloroform extraction and ethanol precipitation. The bound regions were identified by qPCR amplification of the immunoprecipitated material using the indicated primer pairs in Table 1, and represented as percentage of input material (see qPCR).

11. PLASMID CONSTRUCTION

Lentiviral shuttle vector pLOC-Ephrin-A3 (OHS5897-101186331), encoding for the canonical EFNA3 mRNA and pLOC control (OHS5832) were obtained from Open Biosystems. Alternative EFNA3 isoforms were PCR-amplified using the primers indicated in Table 5 and subcloned in pLOC into SpeI/AscI restriction sites.

For in vitro translation Ephrin-A3 and the non-coding isoforms were cloned into BamHI/XHoI restriction sites of pCDNA3 plasmid.

For reporter assays, regions containing the putative promoters were PCR-amplified (see table 5) from genomic DNA extracted from HeLa cells and cloned into pGL4 reporter vector.

Name	Assay	Forward sequence 5'-3'	Reverse sequence 3'-5'
1 HRE	Reporter Assays	CCTTGGTACCCCCGAGGGGGCAGTACGG	CCTTCTCGAGCCCCGGAGCCGCCGCCGCCG
2-8 HRE	Reporter Assays	AAGGATCCCAGAGAGATGTCGGTGTGTCA	AAGGATCCACAGTCCCCAGGTCTGTCAA
4-8 HRE	Reporter Assays	AAGGATCCCACCCTTGGAGCACCTCA	AAGGATCCGGCATATATTTGGCCATCG
4-10 HRE	Reporter Assays	AAGGATCCGGCCCCATAAATCGTGAGAG	AAGGATCCCGGGCTGCAGAGTTACAGG
11 HRE	Reporter Assays	GAGAAGCGATGGAGGGTGT	CACTCAGACTCTCGCGTTCA
Non Coding 1	Cloning isoforms	CGGACTAGTATGGTGAGCCGCAACGGCTAC	CGGGGCGCGCCGACTAAGAGCAGAGTATGAAAGTC
Non Coding 1 short	Cloning isoforms	CGGACTAGTATGGTGAGCCGCAACGGCTAC	
Non Coding 2	Cloning isoforms	CGGACTAGTGAGACCAGGGAGGAGGCGTG	CGGGGCGCGCCGACTAAGAGCAGAGTATGAAAGTC
Non Coding 2 short	Cloning isoforms	CGGACTAGTGAGACCAGGGAGGAGGCGTG	
11 HRE_mut	Mutagenesis	ACACCGTGCAGGTGAAACGACTATCTGGAT	ATCCAGATAGTCGTTTCACCTGCACGGTGT
12 HRE_mut	Mutagenesis	GCGGGGCAGAGCAGTCTGTACATGGTGAGC	GCTCACCATGTACAGACTGCTCTGCCCCGC

Table 5. Primers for plasmid construct
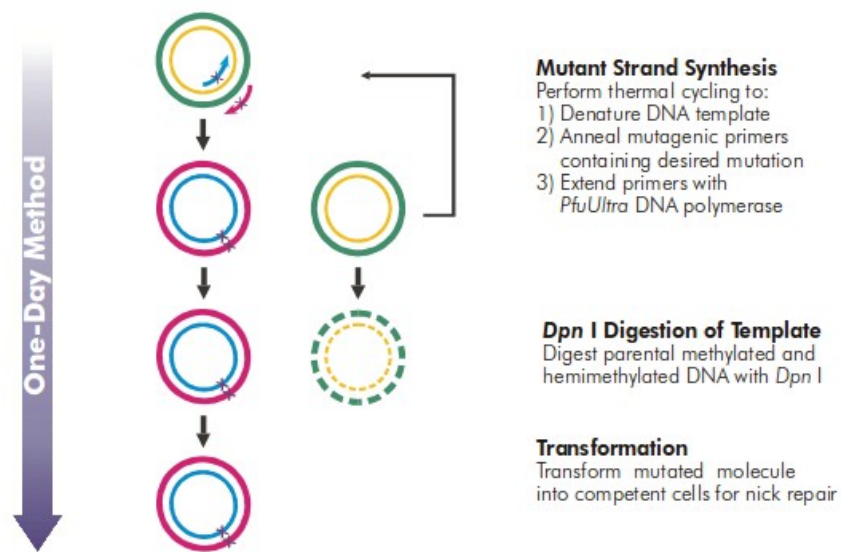
12. IN VITRO TRANSLATION

For in vitro translation we used the TNT T7 Quick Coupled Transcription/Translation system (Promega, L1170) that is optimized for expression of the most standard T7 promoter driven gene constructs. The reaction mix contained 1μg of plasmid and 2μl (20,5 mCi) of EasyTag methionine, L-(^{35}S).The reaction mix was incubated at 30ºC for 90 minutes and subsequently 12μl of this reaction were mixed with 12μl of RIPA buffer with proteases inhibitor and laemmli sample buffer. Then the mix was denatured at 72ºC for 10 minutes and resolved a 12% polyacrilamide gel.

13. REPORTER ASSAYS

HeLa cells were seeded in six-well plates ($3x10^5$ cells/well) and after 6 hours, transfected with 1μg DNA mixture (0,9μg of the indicated reporter or control vector and 0.1μg of plasmid encoding for Renilla), using 2μl of lipofectamine 2000 (Invitrogene). After 16 hours, the cells where re-plated in 24-well plates, and incubated under normoxia, hypoxia (1% oxygen) or in the presence of 500μM DMOG for an additional 16 hours. After treatments, the cells were lysed and the firefly and Renilla luciferase activities were determined using Dual-Luciferase System (Promega, Madison, WI, U.S.A.). The firefly luciferase activity was normalized to that of Renilla luciferase to correct for differences in transfection efficiency between samples.

14. MUTAGENESIS

Mutant reporter constructs were generated with the QuikChange® II XL Site-Directed Mutagenesis Kit (Agilent, 200521-5) using primers designed to delete the HRE motif (Table 5). This kit is used to replace amino acids using Pfu ultra high-fidelity (HF) DNA polymerase, the basic procedure is represented in figure 8.

Mutant Strand Synthesis
Perform thermal cycling to:
1) Denature DNA template
2) Anneal mutagenic primers
   containing desired mutation
3) Extend primers with
   PfuUltra DNA polymerase

Dpn I Digestion of Template
Digest parental methylated and
hemimethylated DNA with *Dpn* I

Transformation
Transform mutated molecule
into competent cells for nick repair

Figure 8. Mutagenesis squeme

15. EPHRIN-A3 OVEREXPRESSION

HeLa cells were seeded into 10 cm^2 plates at 90% of confluence and 16 hours after plating, transfected by adding a mixture of 24μg of the corresponding plasmids and 72μl of Lipofectamine 2000 (Invitrogene). Cells were exposed to the transfection reagents for 4 hours and then we let them recover for 24 hours in fresh media before exposing them to hypoxic or normoxic conditions for 12 additional hours prior analysis of mRNA and protein expression.

16. LENTIVIRAL PRODUCTION

16.1 Particles generation

Replication-deficient lentiviral particles with tropism for human cells were produced using constructs from Open Biosistem. Precision LentiORF have fully-sequenced human open reading frames (ORFs) derived from cDNA coding sequences that have been cloned into a lentiviral backbone, enabling gene and protein expression experiments in mammalian cell type.

Features of the Precision LentiORF backbone include a CMV promoter for robust protein expression, a GFP reporter to visually track protein expression, and blasticidin resistance for selection of stable cell lines.
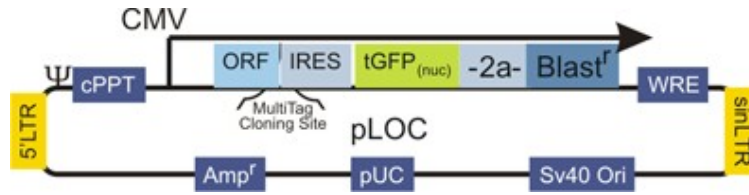
Figure 9. LentiORF backbone

The Precision LentiORF control, contains a RFP (red fluorescent protein) sequence in the Multi Tag cloning site instead of the canonical Ephrin-A3 sequence.

For Lentivirus production, $7x10^6$ HEK-293T cells were seeded per 10cm² plate. After 24 hours medium was changed for 9ml of fresh DMEM with cloroquine 30μM per plate.

Calcium phosphate transfection was performed under this conditions:

1.   Lentiviral plasmid: 14,2μg

2.   Packaging plasmids: 12,45μg of psPAX2

3.   Envelop plasmid: 4,5μg

After 8 hours of transfection, cells were washed andfresh media added to the plates. Transfection efficiency was determined, as the percentage of GFP (green fluorescent protein) positive cells under fluorescent microscopy 48 hours after transfection (typically around 90-100%).

To collect virus, culture supernatant was centrifuged, filtered (0,45μm pore size) and stored at -80ºC.

Lentivirus titration was performed in HEK-293T cell line by flow cytometry quantification of GFP positive cells after 72 hours of infection. The titer was calculated by this formula:

TU/ml=(%GFP positive cells/number of cells the day of infection)/virus volume (ml)

16.2 Cells transduction by lentiviral infection:

MDA-MB-231 cells were infected at a MOI of 4 with pLOC-EFNA3-GFP or pLOC-GFP control lentivirus in the presence of 8μg/ml polibrene (Sigma) for 8 hours. This protocol resulted in more than 95% of GFP positive cells 72 hours after infection. Cells were re-infected with Lenti-Fire Luc2-pGL4 at a MOI of 4 to allow for tracking cells in in vivo assays by means of bioluminescence imaging.
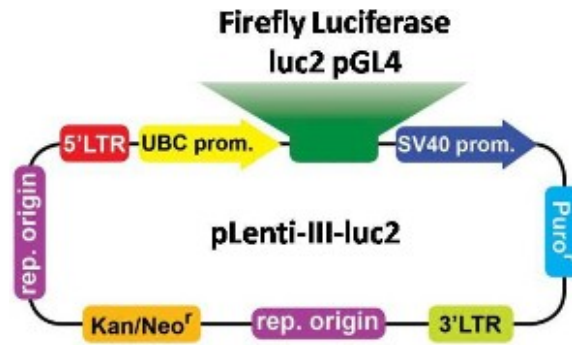
Figure 10. Lenti-Fire Luciferase Luc2 pGL4 vector

17. IN VITRO STUDIES

17.1 Transwell migration

To test the migratory cell capacity, we used a modified Boyden chamber with 6,5mm diameter and 8μm pore size polycarbonate filters (Costar) pre-coated with 0,5% gelatin in PBS. HUVECs were pre-incubated in serum-free media for 16 hours and then $1,25 \times 10^5$ cells seeded in serum-free medium on the top side of the filter and insets placed on culture media containing different combinations of NIH3T3 conditioned media (30μg total protein), 10 % FBS, 500μM DMOG or 0,2-5μg of recombinant Fc-Ephrin-A3 fusion protein (Sigma). Cells were allowed to migrate for 6 hours and then cells remaining on the top of the membrane were removed with a cotton swab and the filters fixed and stained with Diff Quik (Dade Behring). A minimum of ten 40X fields were evaluated for quantitative analysis of cell migration and expressed as the average values ± standard deviation.

17.2 Repulsion assay

For repulsion assays HUVEC cells were stained with 10μM PKH26 per million cells and seeded it in a 24 multiwell plate (1×10^5 cells/well). 24 hours later HUVEC-PKH26 had grown to form a confluent monolayer and HEK-293T (3000 cells/well), transfected cells with GFP or EFNA3-GFP, were plated on top of them. Cultures were evaluated 24 hour latter to asses the number of HEK cells attached to the plates.

18. IN VIVO STUDIES

To study the effect of Ephrin-A3 on tumor growth and metastasis formation we injected MDA-MB-231 cells luciferase alone or in combination with Ephrin-A3 into 5-7 weeks old female nude mice (Harlan). Animals were treated according to the protocols approved by Instituto de Investigaciones Biomédicas and Northwertern University Care and Use Committee in accordance with the NIH Guide for the Care and Use of Laboratory Animals.

18.1 Orthotopic tumor implantation

Cells were harvested by trypsinization, resuspended in PBS and injected into the mammary fat pad ($2x10^6$ per site, one site per mouse) in a 1:1 mix with Matrigel (BD Biosciences). The animals were subjected to weekly bioluminescence imaging in IVIS Spectrum System (Caliper, Xenogen) to monitor tumor progression and metastasis. Mice were anesthetized by isofluorane/oxygen mix and given intraperitoneal injections of D-luciferin (15 mg/ml in PBS) 5 minutes prior to imaging. Mice were examined for 5 weeks after inoculation of tumor cells and then sacrificed.Lungs, bones and livers were harvested and processed for tissue analysis.

18.2 Spontaneous metastases assay

Cells were harvested by trypsinization, resuspended in 100μl PBS and injected intravenously into nude mice (10^6 cells per mice). The animals were subjected to weekly bioluminescence imaging in IVIS Spectrum System (Caliper, Xenogen) to monitor tumor progression and metastasis. The mice were anesthetized by isofluorane/oxigen mix and given intraperitoneal injections of D-luciferin (15 mg/ml in PBS), 5 minutes prior to imaging. Lungs were harvested after 10 weeks and processed for tissue analysis.

18.3 Lung extravasation assay

Cells, were injected intravenously into nude mice ($1,5x10^6$ cells). After 2 hours or 8 days, lungs were isolated and images were acquired by confocal microscopy of the surface of fresh lungs. GFP-positive cells were counted and analysed by ImageJ software.

18.4 EFNA3 measurements in VHL-defficient mice

cDNA from liver and lung of Vhl^{fl}-UBC-Cre-ER^{T2} mice were kindly donated by Julian Aragones (Elorza. A et al., 2012). Briefly, C;129S-*Vhlh^{tm1Jae}*/J mice (Jackson Laboratories, stock no. 4081)

were used to generate the Vhl^{floxed}-UBC-Cre-ER^{T2} mice. These mice harbor two loxP sites flanking the promoter and exon 1 of the murine *Vhl* locus so that VHL-inactivation can be easily achieved by feeding them *ad libitum* with Teckland CRD TAM^{400}/CreER tamoxifen pellets (Harlan Teklad), which contain 400 mg tamoxifen citrate/kg, for 10 days(Miró-Murillo et al., 2011).

18.5 OCT samples preparation:

Isolated organs were fixed in Tissue Teck Cryomold (Dako) with OCT in cold condition with dry ice. These blocks were stoared at -80ºC and sections of 4μm were cutted by Reichert-Jung Cryocut 1900 (Leica). These sections were mounted over slides for posterior analyses.

18.6 In situ Immunofluorescence:

For immunodetection of CD31, 10μm cryosections sections were fixed in acetone, blocked with 3% BSA in PBS and incubated at 4°C overnight with anti-CD31 antibody (BD Biosciences Pharmingen, San Deigo, CA), followed by the anti-mouse secondary antibodies tagged with Alexa-546 (Invitrogen). DAPI was used to countestain nuclei. Quantification of the CD31-positive area was performed using the NIH ImageJ software. Mean values of at least 10 fields of 3 sections were calculated for each sample .

18.7 Hematosin & Eosin Stained (H&E):

The slides were stained with hematoxilina (DC Panreac) for 5 minutes at room temperature, washed 3 times with dH$_2$O for 5 minutes and then eosina stained for 1 minute at room temperature . Samples where rehydrated with ethanol increased concentration washed (50%, 70%, 95%, 100%) and cleaned with xilol for 20 minutes . Final samples were mounted with Depex (BDH Prolabo).

19. GENE EXPRESSION ANALYSIS

Gene expression data and relevant sample information were downloaded from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) and ROCK (http://www.rock.icr.ac.uk/) databases (free public access) and analyzed using custom R scripts for statistical programming (http://www.r-project.org/).

The breast cancer datasets Loi2008 (Loi S., et al., 2008), Pawitan2005 (Pawitan Y., et al., 2005) and Ur_Rheman, (Sims D., et al., 2010) and the RCC series GDS505 (Lenburg M. E., et al., 2003) and GDS2880, (Tun H. W., et al., 2010)) were generated using Affimetrix Human Genome U133A

Array platform. Thus, in these datasets EFNA3 expression was determined with the probe 210132_at, which recognizes the 3'UTR region. The breast cancer study VandeVijve (Van de Vijver M. J., et al., 2002) was performed using custom ~60 mer oligonucleotides specific for NM_004952 refseq gene. Finally, the RCC series GSE16441 utilized Agilent-014850 Whole Human Genome Microarray 4x44K G4112F.

20. STATISTICAL ANALYSES

Statistical analysis and graphic representations were performed in R, a language and environment for statistical computation and graphics (R Core Team, 2010).

RESULTS

1. EPHRIN-A3 IS INDUCED UNDER HYPOXIA CONDITIONS

We first identified EFNA3, a member of the ephrin type A ligands, as potential novel HIF target gene using an *in silico* search (Ortiz-Barahona A., 2010) . As a first step to validate this prediction, we determined the level of Ephrin-A3 protein and found that it was induced by hypoxia in several cell lines (figure 11).
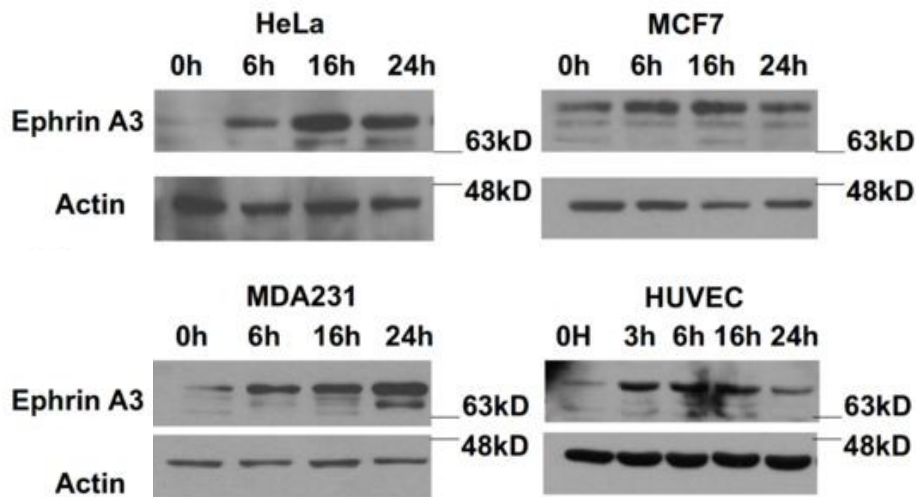
Figure 11. Hypoxia induces the expression of Ephrin-A3. EFNA3 protein was determined in HeLa, HUVEC, MCF-7 and MDA-MB23, cells were exposed to normoxia (21% oxygen) or hypoxia (1% oxygen) for the indicated periods of time.

The Ephrin-A3 species we detected had an apparent molecular weight of approximately 72kD, much higher than the 26.3kD predicted from its amino acid sequence. In orderto confirm the identity of the protein recognized by the antibody, we treated cells with a siRNA directed against EFNA3 and found that it decreased the intensity of the high molecular band specifically (figure 12A and 12B).
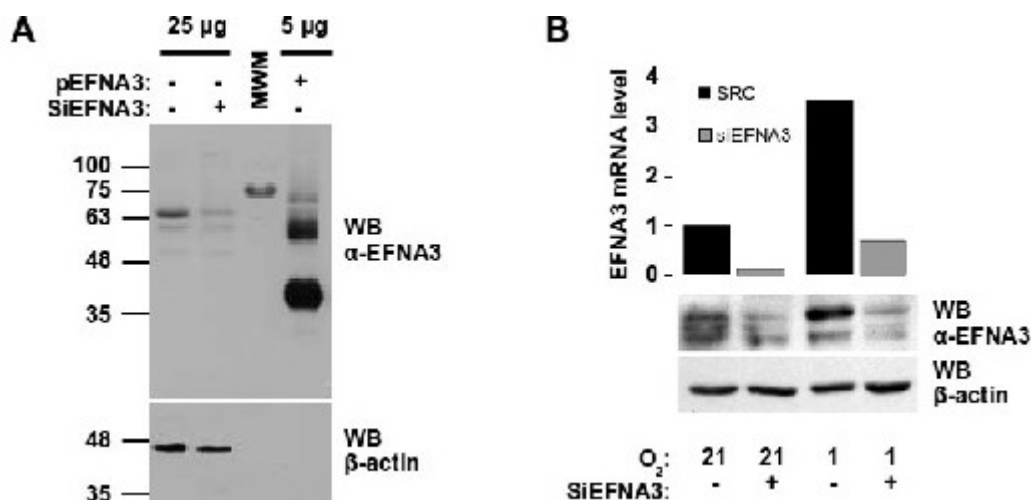
Figure 12. Ephrin-A3 migrate with a molecular weight of 72kD (A) HeLa cells were transfected with a plasmid encoding EFNA3 (pEFNA3) or siRNA directed against EFNA3 and 48h postransfection EFNA3 proteins levels were determined by immunoblot. (A) To allow comparison in the same membrane, different amounts of cell lysate were used to detect the endogenous and exogenous proteins. (B) HeLa cells were transfected with scramble siRNA (-) or siRNA directed against EFNA3 and then grown at 21% (Nx) or 1% oxygen (Hyp) for 12 hours. Duplicated samples were processed to determine EFNA3 mRNA (upper graph) or protein (lower panels) levels.

Type A Ephrins undergo several postranslational modifications, including the addition of a GPI moiety and glycans, that could account for its high apparent mass. In agreement, treatment of cell lysates with N-glycosidase, resulted in the depletion of the endogenous ~72kD bands and generation of a faster migrating band (figure 13A). Note that the intensity of the ~50kD band generated by the enzymatic treatment is much higher than that of the corresponding 72kD bands, probably due to the exposure of the antigenic determinants after the removal of the oligosaccharide chains, as has been shown for other postranslational modifications (Bütikofer, Malherbe, Boschung, & Roditi, 2001). An additional test to confirm the identity of the ~72kD bands was performed. We used a recombinant form of EphA5 (a high affinity receptor for Ephrin-A3) fused to the immunoglobulin Fc fragment to pull down endogenous Ephrin-A3 from cell lysates. As shown in panel figure 13B, the Ephrin-A3 forms that co-precipitated with the EphA5-Fc migrate as a ~72Kd bands. Altogether these results indicate that Ephrin-3A is detected as ~72Kd bands in SDS-PAGE gels due to postranslational modifications.
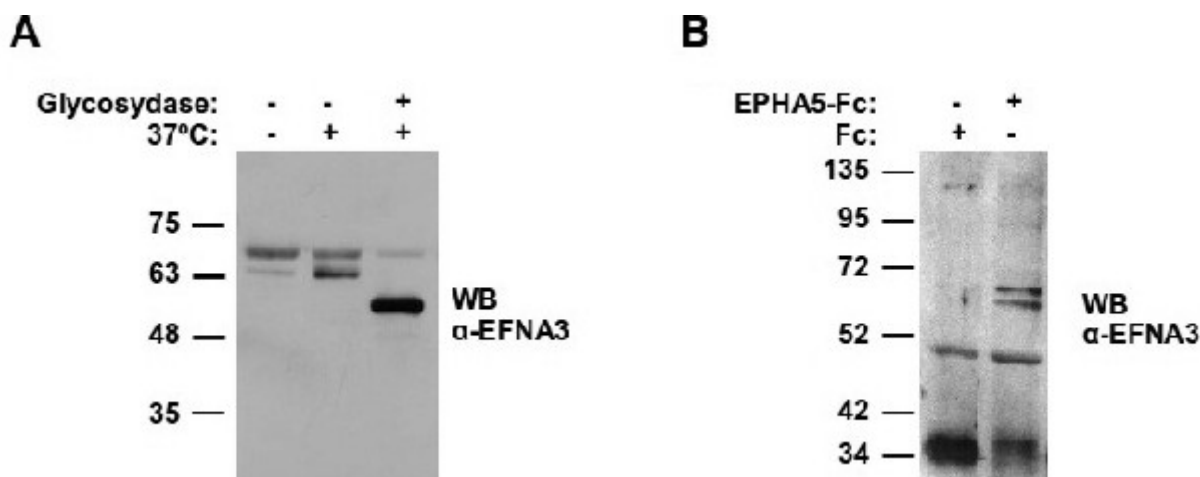
Figure 13 (A) HeLa cell lysates were incubated at 37ºC (+) or 0ºC (-) in the presence (+) or absence (-) of N-glycosidase and then resolved by SDS-PAGE and probed with anti-EFNA3 antibodies. (B) HeLa cells lysates were incubated with recombinant Fc-EphA5 receptor or Ig Fc fragments. Recombinant constructs were pulled down with protein A-sepharose and coimmunoprecipitating material was resolved by SDS-PAGE and probed with anti-EFNA3 antibodies.

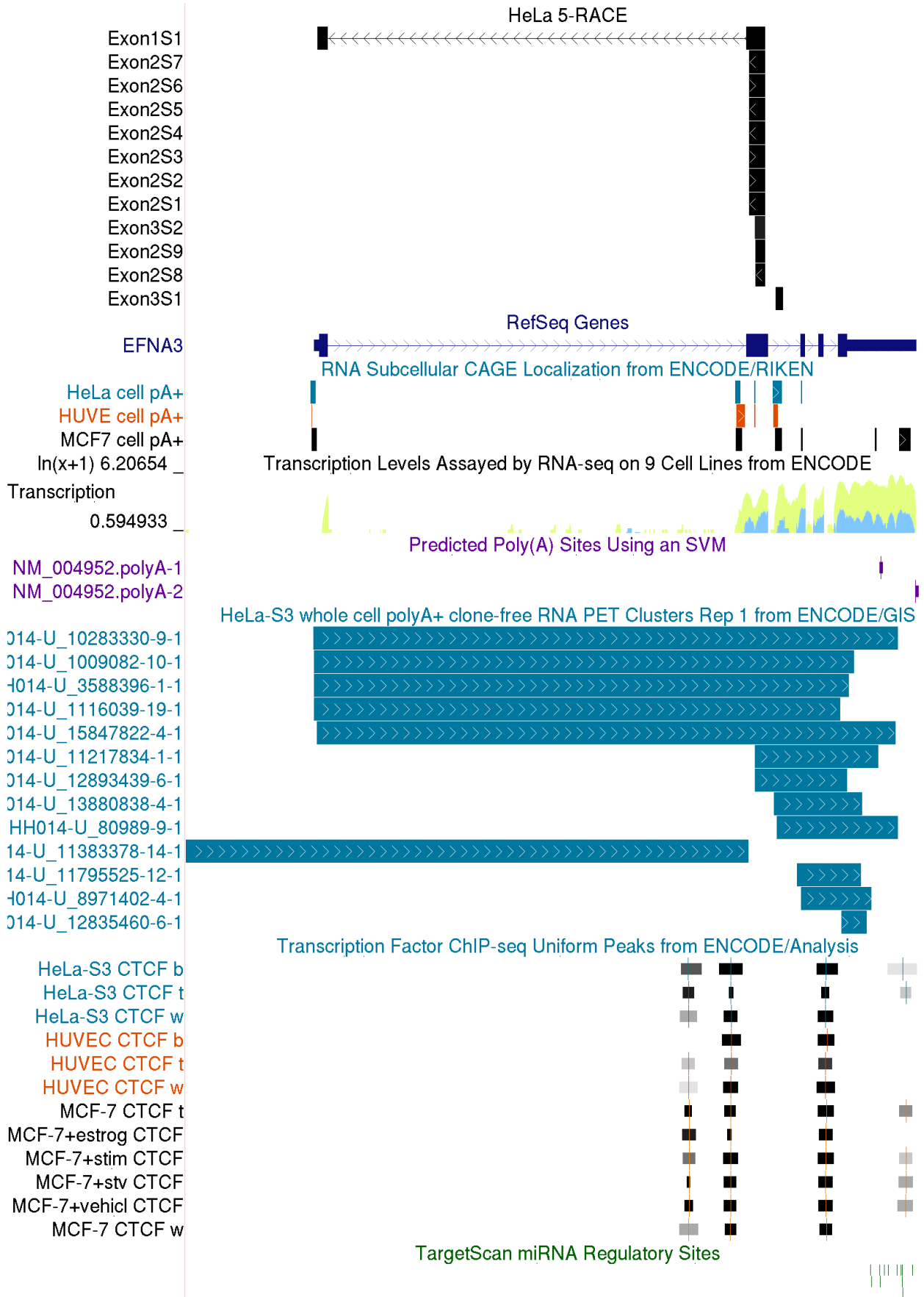## 2. EFNA3 LOCUS ENCODES LONG NON-CODING RNAs (lncRNAs) THAT ARE REGULATED BY HYPOXIA

Next we decided to study the effect of hypoxia on EFNA3 mRNA. According to curated databases (RefSeq) the EFNA3 locus encodes for a single mRNA isoform (NM_004952). However, our analysis of the open access databases, including ESTs and experimentally identified transcription start sites (TSS) (http://dbtss.hgc.jp/), suggested the existence of additional mRNAs transcribed from this locus. To identify hypothetical isoforms encoded by the EFNA3 locus, we performed 5'-RACE experiments in HeLa and LoVo cell lines ("HeLa 5-RACE") and found two transcription start sites (TSS) in addition to that of the NM_004952 mRNA (Figure 14).
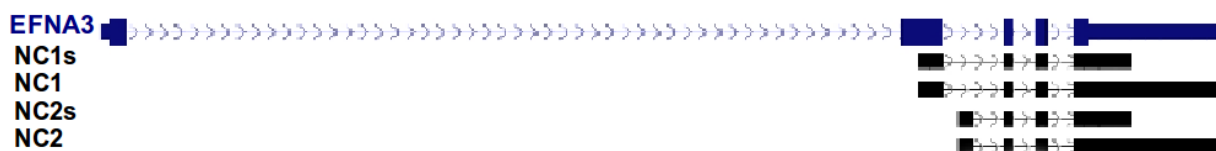
Figure 14. Diagram of the EFNA3 locus. Canonical EFNA3, NM_004952 RefSeq, gene (upper track,) and the novel isoforms (second track from the top, N1-NC2s).

The existence and location of these additional transcription start sites was supported by 5' cap analysis gene expression (CAGE) tags from multiple cell lines produced as part of the ENCODE transcriptome Project (Figure 15, "RNA Subcellular CAGE"). Interestingly, none of these novel RNAs seemed to encode for functional proteins and, accordingly, we termed them non-coding-1 (NC1) and NC2, based on their different TSS (Figure15). Specifically, NC1 isoform contains an ATG codon within its first 10 nucleotides that is in frame with the open reading frames (ORF) of NM_004952, suggesting that NC1 could encode a truncated form of Ephrin A3. However, this potential product would have an interrupted structural domain (the ephrin domain), that makes up the bulk of the protein, which would likely undermine its stability. On the other hand, the NC2 sequence does not contain any ORFs of significant length.

HeLa 5-RACE

Exon1S1
Exon2S7
Exon2S6
Exon2S5
Exon2S4
Exon2S3
Exon2S2
Exon2S1
Exon3S2
Exon2S9
Exon2S8
Exon3S1

RefSeq Genes

EFNA3

RNA Subcellular CAGE Localization from ENCODE/RIKEN

HeLa cell pA+
HUVE cell pA+
MCF7 cell pA+

ln(x+1) 6.20654 _

Transcription Levels Assayed by RNA-seq on 9 Cell Lines from ENCODE

Transcription

0.594933 _

Predicted Poly(A) Sites Using an SVM

NM_004952.polyA-1
NM_004952.polyA-2

HeLa-S3 whole cell polyA+ clone-free RNA PET Clusters Rep 1 from ENCODE/GIS

014-U_10283330-9-1
014-U_1009082-10-1
H014-U_3588396-1-1
014-U_1116039-19-1
014-U_15847822-4-1
014-U_11217834-1-1
014-U_12893439-6-1
014-U_13880838-4-1
HH014-U_80989-9-1
14-U_11383378-14-1
14-U_11795525-12-1
H014-U_8971402-4-1
014-U_12835460-6-1

Transcription Factor ChIP-seq Uniform Peaks from ENCODE/Analysis

HeLa-S3 CTCF b
HeLa-S3 CTCF t
HeLa-S3 CTCF w
HUVEC CTCF b
HUVEC CTCF t
HUVEC CTCF w
MCF-7 CTCF t
MCF-7+estrog CTCF
MCF-7+stim CTCF
MCF-7+stv CTCF
MCF-7+vehicl CTCF
MCF-7 CTCF w

TargetScan miRNA Regulatory Sites

Figure 15. EFNA3 locus encodes for several transcripts expressed from two different promoter regions.
The 5'-ends of the EFNA3 transcripts was determined by rapid amplification of cDNA ends (RACE) using the FirstChoice RLM-RACE kit (Applied Biosystems) according to manufacturer's instructions. Total RNA from HeLa and LoVo cells was used as template for 5'- RACE reactions, the products cloned into pCR2.1 (Invitrogen) and their sequence determined by Sanger sequencing. The figure shows the position of the RACE products ("HeLa 5-RACE") relative to the canonical NM_004952 RefSeq gene (EFNA3 track). The names indicate the location of EFNA3-specific primers used for the PCR amplification of RACE products: AGAAGGCGCTGTAGCGCTGGAA (Exon1 and 2) ; TTCCAGTGCAGGTTGTGAGT (Exon 3). The "RNA Subcellular CAGE" track shows 5' cap analysis gene expression (CAGE) tags from the ENCODE project and thus points to Transcription Start Sites (TSS) within this locus. Note the tight correlation between our 5'-RACE results and those reported by ENCODE. To keep image size to a minimum, in this and following tracks we only included information from cell lines used in our work (HeLa, HUVEC and MCF7). The "Transcription" track shows transcription levels for HeLa (green) and HUVEC (blue) as assayed by high-throughput sequencing of polyadenylated RNA (RNA-seq).The prediction of Poly(A) sites using a supported vector machine-based algorithm indicates the existence of two different 3-end for this locus ("predicted poly(A) sites" track). The next track shows the result of Gene Identification Signature (GIS) paired-end ditag (PET) sequencing experiments in HeLa cells. This technique provide signatures of the 5' start and the 3' end of individual mRNA transcripts (Ng et al., 2005). The "Transcription Factor ChIP-Seq" track shows the CTCF binding sites in HeLa, HUVEC and MCF7 cells. Finally, the "TargetScan miRNA" track shows the predicted binding sites (TargetScan) for several miRNAs one of them being miR210. The figure was generated by the UCSC genome browser upon loading the indicated custom tracks

In agreement, in vitro transcription-coupled translation of EFNA3 and NC1 cDNAs, produced proteins of expected sizes, whereas NC2 cDNA generated no apparent protein product (Figure 16A). To test the coding potential of NC1 and NC2 in intact cells, we transfected HeLa cells with the respective plasmids along with the canonical EFNA3 cDNA (EFNA3) (Figure 16B). Neither NC1 nor NC2 caused the expression of exogenous protein that could be recognized by the monoclonal or polyclonal antibody against EFNA3. Thus NC1 and NC2 can be considered novel long non-coding RNAs (lncRNAs).
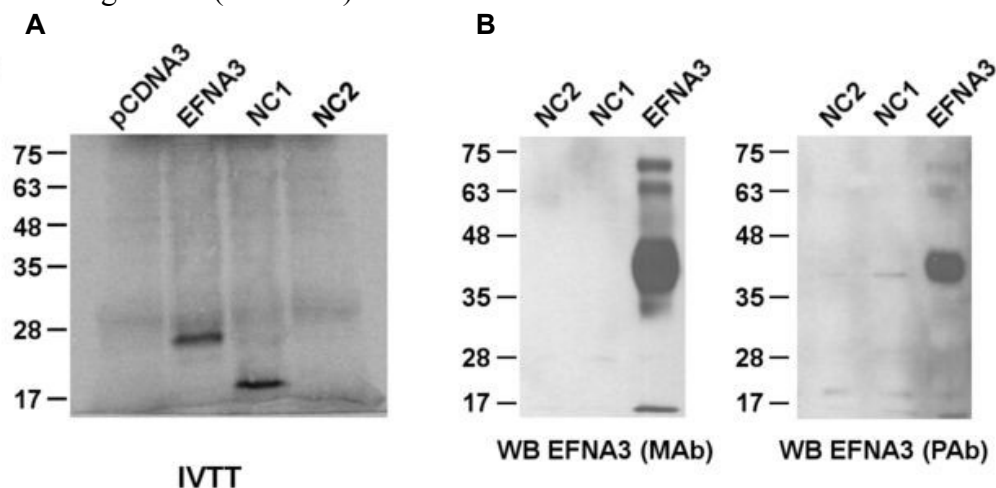
Figure 16. **EFNA3 isoforms are long-non coding RNA**. (A) cDNAs corresponding to EFNA3, NC1 or NC2 isoforms were transcribed and translated in vitro in the presence of 35S-methionine and proteins resolved by SDS-PAGE. Image shows the autoradiogram of a representative experiment. (B) HeLa cells were transfected with plasmids encoding for the indicated EFNA3 isoforms and cell lysates probed with monoclonal (MAb) or polyclonal (PAb) antibodies against the C-terminal region of human Ephrin-A3

In addition to the novel TSS, a method using support vector machine for poly(A) site prediction (Cheng Y., 2006) identifies two potential 3'-ends for this locus (figure 15 "Poly(A) Sites" track). In agreement, analysis of Paired-end diTag (PET) Sequencing data from ENCODE (Ng P., et al., 2005) indicates that all combinations of the TSSs and 3'-ends are present in cells (figure 15, "GIS-PET" track). Thus, both NC1 and NC2, can be expressed as shorter forms with a truncated 3'UTR, which we termed NC1s and NC2s respectively (figure 14).

We next investigated the regulation by hypoxia of the different transcripts encoded by the EFNA3 locus. First, we used commercially available TaqMan probes to amplify the regions of the EFNA3 gene specific to the canonical EFNA3 mRNA (Exon1+2, Figure) and a region common to all the RNA isoforms encoded by this locus (Exonn4+5, Figure). qPCR results indicate that the absolute expression levels and relative induction in response to hypoxia varied widely between RNA species (figure 17). The expression of the canonical coding isoform was low, compared to the combined expression level for all isoforms, suggesting that under normoxic conditions the transcription of the long non-coding RNAs predominates. This result was confirmed by the ENCODE genome-wide transcription analysis (figure 15, "Transcription" track). Strikingly, the canonical EFNA3 isoform was barely induced in response to hypoxia, in stark contrast with the robust upregulation of the bulk of EFNA3 RNAs (figure 17).
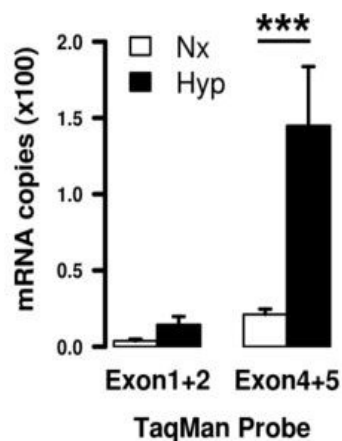
Figure 17. mRNA absolutes levels of the lncRNA and canonical isoform of the EFNA3. HeLa cells were exposed to 1% oxygen (Hyp) or left at normoxic conditions (Nx) for 12h and EFNA3 RNA levels were determined using the indicated TaqMan probes (see Table3). The graph shows the ratio of the number of EFNA3 to ACTB (beta-Actin) copy number. Bars represent the mean of three independent biological replicates and the error bars the standard deviation. The differences between groups were statistically significant (ANOVA $F_{3,8}=34.21$, $p<0.001$).The asterisks indicate mean pairs that were statistically significant (adjusted $p <0.001$) in a posteriori Tukey test.

This result suggested that the lncRNAs, but not the canonical mRNA, were regulated by hypoxia. To confirm this possibility, we designed primer pairs for all the exons in the gene to determine the response to hypoxia of every potential RNA isoform. Indeed, hypoxia strongly upregulated the novel lncRNA isoforms, while the regulation of the canonical protein-coding mRNA was only marginal (figure 18).
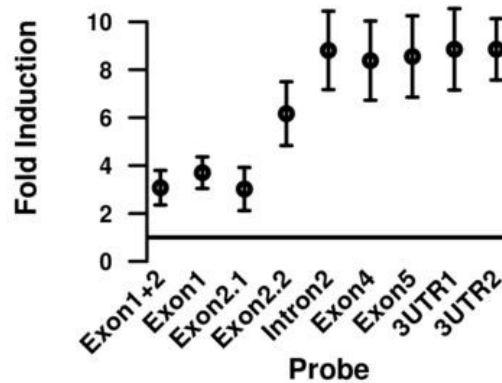
**Figure 18. Fold induction quantification of each exon in the EFNA3 locus.** HeLa cells exposed to 1% oxygen (Hyp) or at normoxic conditions (Nx) for 12 hours and EFNA3 RNA levels were determined using the indicated primer pairs (See Figure 21). The graph shows the ratio of hypoxic EFNA3 RNA to the expression in normoxia. Symbols represent the mean of three independent biological replicates and the error bars the standard deviation.

Importantly, the induction of EFNA3 transcripts in response to hypoxia was mediated by HIF as it was blocked by siRNA directed against HIF1α (figure 19).
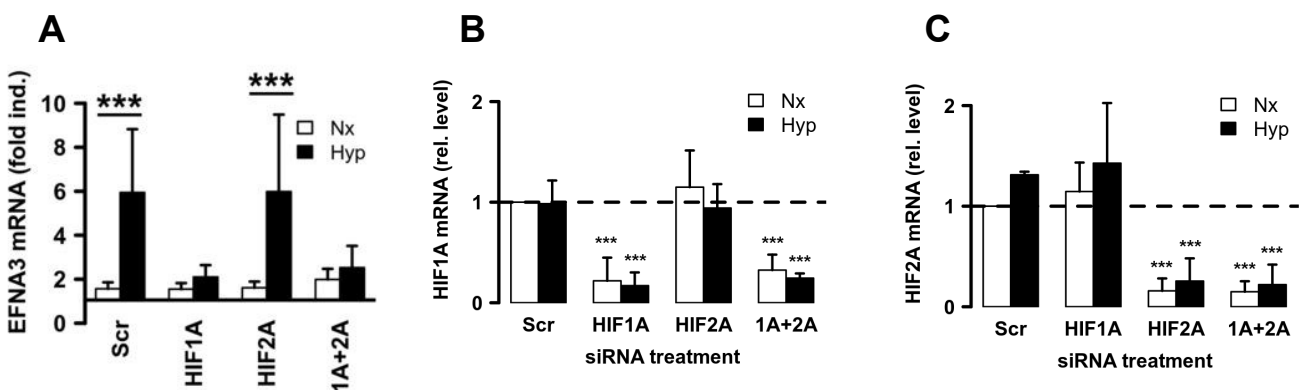
Figure 19. siRNAs treatments HeLa cells were treated with the indicated siRNAs and then grown at normoxic (21% oxygen, Nx) or hypoxic (1% oxygen, Hyp) conditions for 12 hours and the levels of EFNA3 (A), HIF1A (B) and EPAS1 (C) mRNA were determined by qPCR. The graph represents the normalized levels of mRNA as fold over control conditions (normoxic cells treated with scramble siRNA). Bars represents average values in four independent biological replicates and error bars the standard deviation. The differences between groups were statistically significant: (A) ANOVA $F_{7,48}=23.52$, $p<0.001$; (B) ANOVA $F_{7,24}=16.9$, $p<0.001$; (C) ANOVA $F_{7,20}=17.58$, $p<0.001$. The asterisks indicate mean pairs that were statistically significant (***, adjusted $p<0.001$) in a posteriori Tukey test.

To investigate the regulation of EFNA3 by HIF in vivo, we employed conditional VHL knockout mouse lines (Miró-Murillo M., et al., 2011) . VHL deletion results in constitutive HIF activity and, consistently with the in vitro results, this led to increased EFNA3 lncRNA expression without significantly altering the level of the coding EFNA3 mRNA. Importantly, the induction of EFNA3 upon VHL loss was partially prevented in animals lacking both VHL and EPAS (HIF2α) alleles (figure 20), suggesting that the in vivo HIF2α mediates the effect of VHL, at least in the liver and the lung.
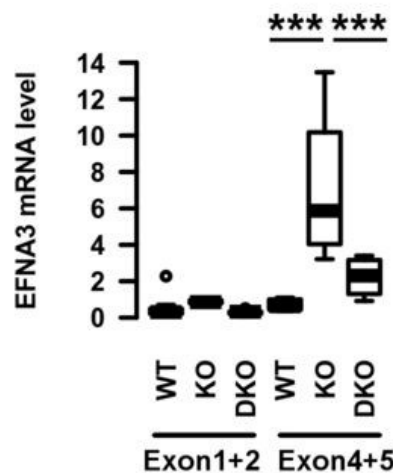

Figure 20. EFNA3 is regulated by HIF in vivo. Vhlfl/fl-UBC-Cre-ERT2 (n = 5), Vhlfl/flHIF2αfl/fl-UBC-Cre-ERT2 (n=3) and control (n=11) mice were placed on a tamoxifen diet for ten days followed by ten additional days on a normal diet. The box-and-whisker plot represents the distribution of the normalized EFNA3 mRNA expression in the liver of animals with the indicated genotypes. The box contains the values comprised between the second and third quartiles, and the horizontal black line the median. The "whiskers" extent to 1.5 times the interquartile range. The differences between groups were statistically significant (ANOVA $F_{2,18}=20.49$, $p<0.001$) and the asterisks indicate means pairs that were the statistically significant (adjusted $p <0.001$) in a posteriori Tukey test.

Regardless the specific HIF isoform involved, that could just reflect their differential tissue expression, it is clear from these set of results that EFNA3 expression is induced, both in vitro and in vivo, in response to hypoxia, in a HIF-dependent manner.

3. LncRNAs ENCODED BY THE EFNA3 LOCUS CAUSE EPHRIN-A3 ACCUMULATION

Mounting evidence indicates that lncRNAs are key regulators of gene expression that affect the mRNA transcription rate, stability and translation (Nie L., et al., 2012; Yoon J. H., 2012). Thus, in an attempt to reconcile the induction of Ephrin-A3 protein with the regulation of coding and non-coding transcripts by hypoxia, we tested whether NC1 and NC2 lncRNAs could affect EFNA3 mRNA or protein levels. As shown in figure 21 , the overexpression of the lncRNAs had no

significant effect on EFNA3 mRNA levels (figure 21A). However, exogenous expression of the different isoforms, particularly the short ones (NC1s and NC2s) in HeLa (figure 21B and 21C) and MDA-MB-231 (figure 21D), caused EFNA3 protein accumulation. These results provide an explanation for the induction of EFNA3 protein under hypoxia in spite of its modest effect on EFNA3 mRNA level.
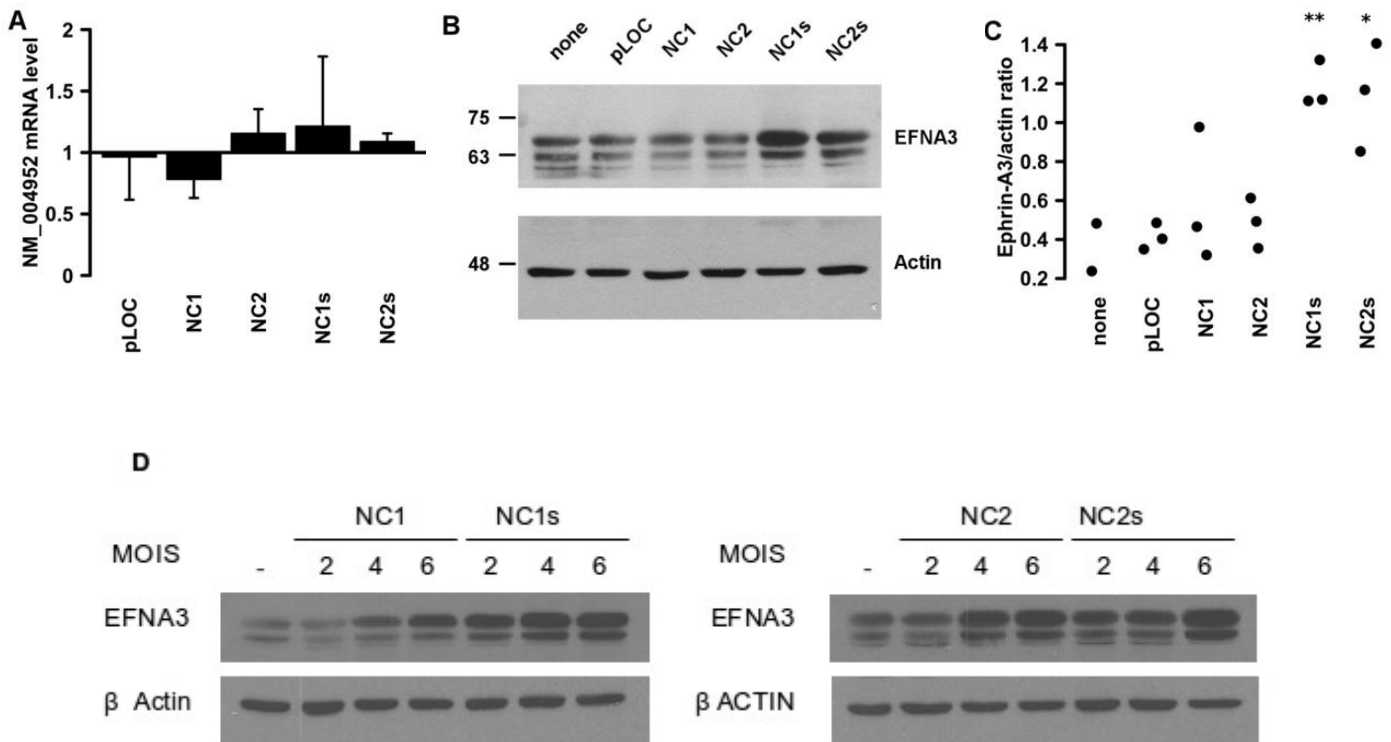

Figure 21. lncRNA expression results in increased EFNA3 protein levels. HeLa cells were transfected with constructs encoding for the indicated lncRNAs or empty plasmid (pLOC) and the level of the canonical EFNA3 mRNA (A) and protein (B and C) were determined by qPCR and immunoblot respectively. The graph in A represents the level of the NM_004952 mRNA, determined with primers Exon1+2, as a fold over the level found in untreated cells. Bars represent the mean of five independent biological replicates and the error bars the standard deviation. The differences between groups were not statistically significant (ANOVA $F_{4,20}=1.452$, $p=0.254$). The graph in C represents the level of Ephrin-A3 as the ratio of the Ephrin-A3 band intensity corrected by the Actin band intensity in three independent experiments. The differences between groups was statistically significant (ANOVA $F_{5,11}=8.635$, $p<0.01$).The asterisks indicate sample means that were significantly different from controls (pLOC samples) in a posteriori Tukey test (*, adjusted $p < 0.05$; **, adjusted $p < 0.01$). The image in B is representative of at least three independent experiments. (D) Immunoblot in MDA-MB-231 cells infected at different MOI with lentivirus encoding for the indicated lncRNA.

4. EFNA3 ISOFORMS ARISE FROM ALTERNATIVE PROMOTERS WITH DIFFERENTIAL RESPONSIVENESS TO HYPOXIA

The different transcription start sites of the EFNA3 isoforms suggested the existence of alternative promoters. In agreement, histone modifications landscape, location of the open chromatin and transcription factor binding site (TFBS) regions in the EFNA3 locus (figure 22) were consistent

with the existence of two promoter regions (figure 22, Prm1 and Prm2).
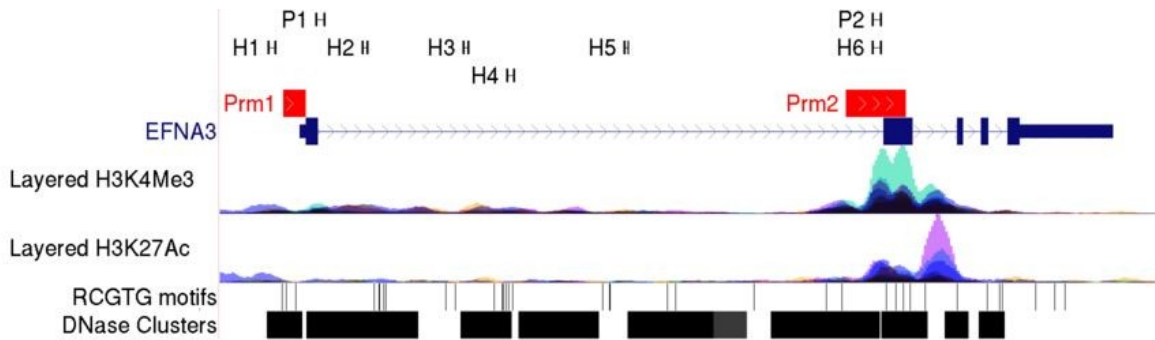
Figure 22 L**ncRNAs encoded by the EFNA locus are transcribed from an alternative hypoxia-responsive promoter**. Diagram depicting the EFNA3 locus and showing the NM_004952 RefSeq gene ("EFNA3" track) along with accessible chromatine regions ("DNase clusters track"), histone marks associated to promoters ("Layered H3K4Me3" track) and active regulatory elements ("H3K27Ac Track"). The colors in the histone tracks correspond to the signal obtained in different cell lines (see UCSC for details). The figure was generated by the UCSC genome browser upon loading custom tracks to indicate the location of primer and amplicons used in the RNA pol II (P1 and P2) and Hif1a (H1-H6) ChIP-qPCR experiments as well as the regions cloned to assay their promoter activity ("promoters" track).

To further investigate the regulation of the different EFNA3 transcripts, we analyzed the binding of RNA polymerase II (Pol II) to these putative promoter regions. As controls, we also included the promoter regions of a bona fide hypoxia inducible gene (P4HA) and a non-responsive gene STT3A. ChIP-qPCR analysis showed that Pol II bound to both regions of the EFNA3 gene locus under normoxia (figure 23).
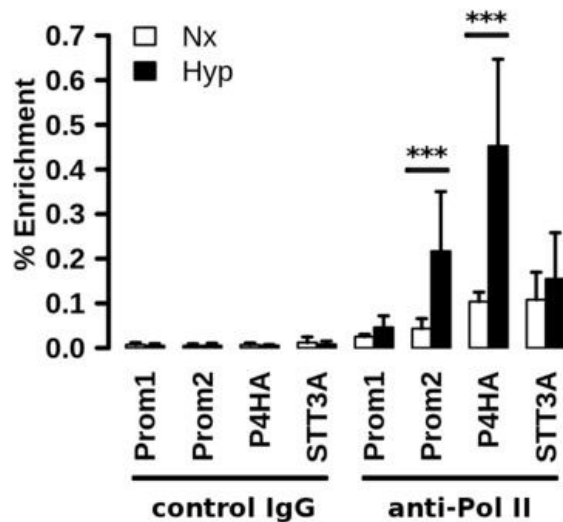
Figure 23 . Binding of RNA polymerase II by ChiP-qPCR. HeLa cells were exposed to 21% or 1% oxygen environment for 8 hours and RNA pol II binding to the indicated regions of the EFNA3 locus. Binding of P4hA and STT3S promoters were used as positive and negative controls respectively. The graph shows the amount of precipitated material as a percentage of the input (% enrichment). Bars represent the mean of three independent biological replicates and the error bars the standard deviation. The differences between groups were statistically significant (ANOVA $F_{15,69}$ =15.6, p< 0.001) and the asterisks indicate mean pairs that were statistically significant in a *posteriori* Turkey test.

50

Consistent with the relatively low EFNA3 expression in HeLa cells, the binding of Pol II to the EFNA3 promoter regions was lower than its binding to the P4HA and STT3A promoters. Importantly, Pol II binding to the P2 region was strongly induced by hypoxia (figure 23), to a level comparable to the HIF-responsive promoter of P4HA. In contrast, the Pol II binding to the Prm1 region and to the STT3A promoter remained unaffected by hypoxia (figure 23).

Next, we investigated direct binding of HIF1α to the EFNA3 locus by ChIP-qPCR. We designed six primer sets (figure 22, H1-H6) to sample most of the RCGTG motifs (figure 22 , "RCGTG motifs" track) within EFNA3 locus and found that the only fragment that showed, significant HIF binding under hypoxic conditions was H6 (figure 24). Interestingly, H6 region is close proximity to Prm2, which could explain the differential response of promoter regions 1 and 2 to hypoxia.
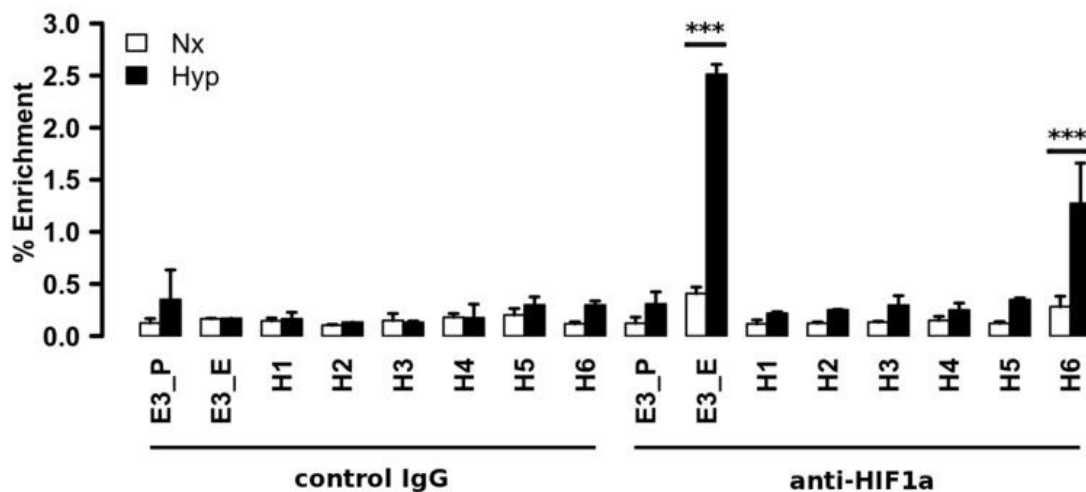
Figure 24 . **Binding of HIF1α by ChiP-qPCR.** HeLa cells were exposed to 21% or 1% of oxygen for 8 hours and HIF1α binding to the indicated regions of the EFNA3 locus. Binding of EGLN3 enhancer region (E3_E) and EGLN3 promoter region (E3_P) were used as positive and negative controls respectively. The graph shows the amount of precipitated material as a percentage of the input (% enrichment). Bars represent the mean of three independent biological replicates and the error bars the standard deviation. The differences between groups were statistically significant (ANOVA $F_{31,32}$ =39.99, p< 0.001) and the asterisks indicate mean pairs that were statistically significant in a *posteriori* Turkey test.

Finally, we cloned both EFNA3 promoter regions and studied their response to hypoxia and to the chemical inhibitor of HIF prolyl hydroxylase, DMOG, which causes acute HIF activation. Although both regions showed similar basal promoter activity (data not shown), Prm2 was strongly induced by hypoxia and DMOG, whereas Prm1 remained unaffected (figure 25). Importantly, Prm2 response to hypoxia was critically dependent on one of its RCGTG motifs whose mutation completely abrogated hypoxic induction (figure 25).
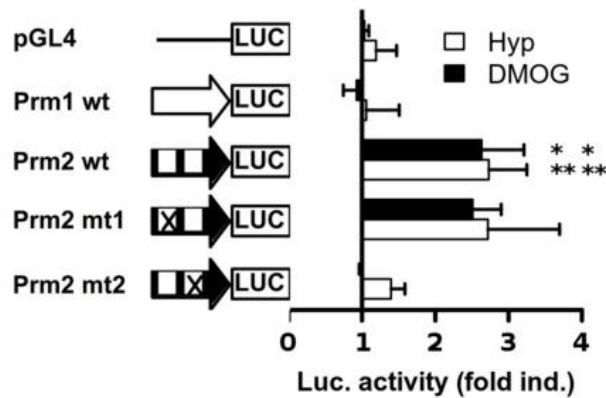
Figure 25. Reporter assay The effect of hypoxia (white bars) and DMOG (grey bars) on the transcriptional activity of the promoter region 1 and 2 was assessed by reporter assays upon transfection of the indicated constructs into HeLa cells. White boxes within the promoter 2 (Prm2) diagram represent RCGTG motifs and crossed boxes represent mutated RCGTG motifs. The graph shows the normalized luciferase activity in hypoxic (Hyp) or DMOG-treated (DMOG) samples expressed as fold over the activity obtained in normoxic conditions. Bars represent the mean of three independent biological replicates and the error bars the standard deviation. The differences between groups were statistically significant (ANOVA $F_{4,10} = 21.9$, $p < 0.01$) and the asterisks indicate mean pairs that were statistically significant (*, adjusted $p < 0.05$; **, adjusted $p < 0.01$) in a *posteriori* Turkey test.

5. HIF ACTIVATION CORRELATES WITH EFNA3 EXPRESSION IN HUMAN TUMORS

Over 80% of human clear cell renal cell carcinomas (ccRCC) samples are deficient for VHL function and, as a consequence, present constitutive HIF activity even in the presence of oxygen (Kaelin W.G., 2007) . Thus, ccRCC is highly suitable to study a putative link between HIF and EFNA3 in human tumors. In agreement, the expression of EGLN3, a well-characterized direct target of HIF, was clearly increased in the three independent ccRCC tumor series (figure 26). Similarly, EFNA3 was significantly increased in ccRCC tumor cells as compared with the normal kidney tissue in the same series, regardless of the microarray platform used to assay tumor samples (figure 26). The probes used to determine EFNA3 expression in these datasets bind to regions common to all transcripts from this locus and thus they detect the cumulative signal by the mRNA and lncRNA transcripts. However, since the expression of lncRNAs was much higher than that of canonical EFNA3 mRNA (figure 17), we assume that the microarray signal is largely generated by the lncRNA. We also examined the expression of two other EFNA family members, EFNA1 and EFNA4, whose coding genes flank the EFNA3 locus in mammalian genomes. Interestingly, EFNA1, but not EFNA4, is induced by hypoxia (YamashitaT., et al., 2008). Concordantly, the expression of EFNA1, but not that of EFNA4, is clearly increased in ccRCC (figure 26). These results rule out that the increased EFNA3 expression observed in these tumors could be caused by a gross structural alteration or transcriptional deregulation of the genomic region containing the EFNA3 gene and suggest it is due to increased HIF activity as it correlates with the changes

observed for other hypoxia regulated genes.
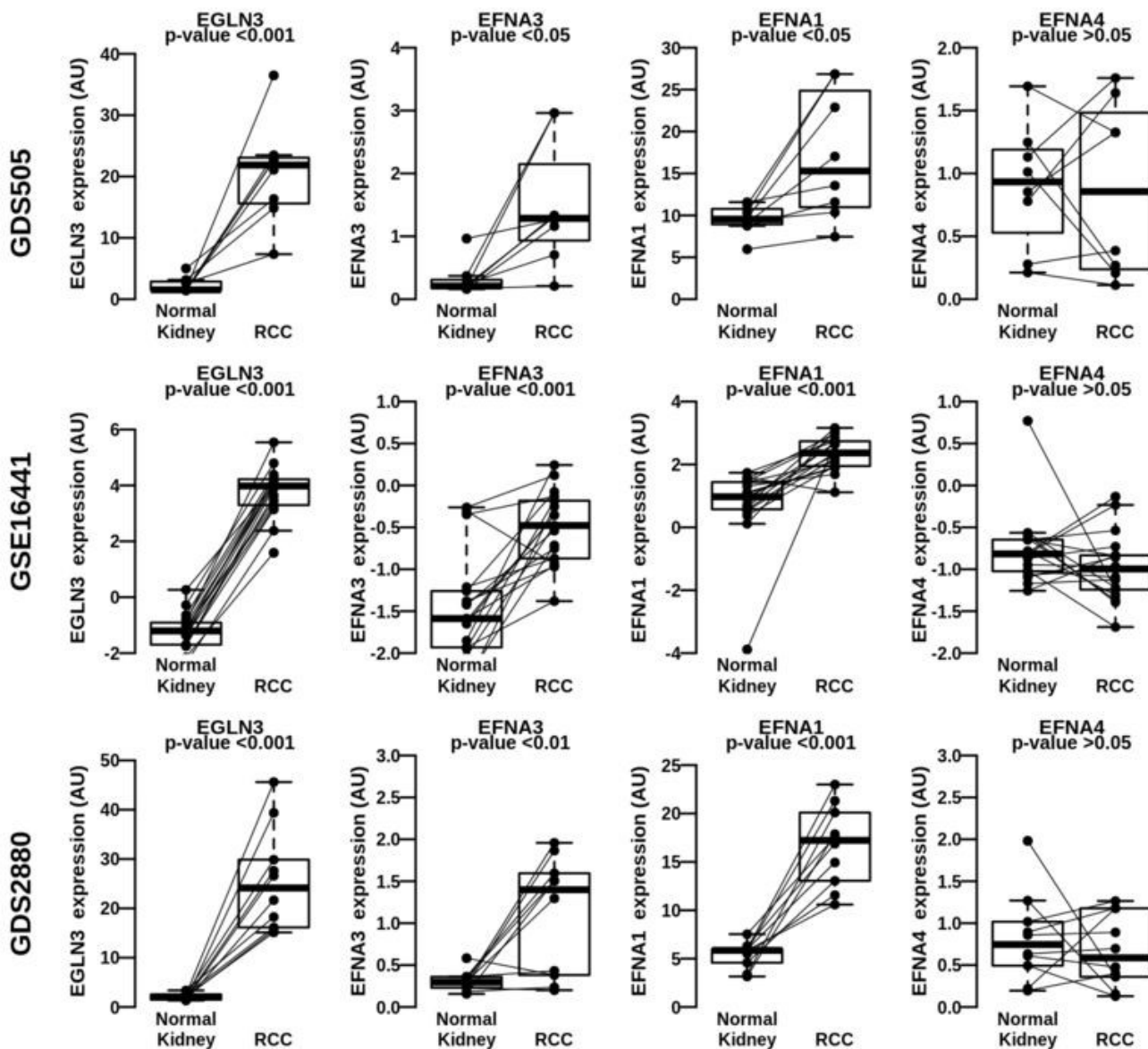
Figure 26. EFNA3 expression is induced in human renal clear cell carcinomas (RCC). The expression of EGLN3, EFNA3, EFNA1 and EFNA4 (columns) was determined in publicly available gene expression profiles of ccRCC samples from three independent studies (ID shown on the left margin of each row). Graphs represent the expression of the indicated genes in arbitrary units (normalized microarray intensity values). The individual samples are shown and pairs of tumoral and normal kidney tissue are joined by segments. . The statistical significance of mean differences in indicated on top of each graph (paired student's t test).

6. EFNA3 PROMOTES METASTATIC BEHAVIOR BY ENHANCING EXTRAVASATION OF TUMOR CELLS

Since hypoxia is a common condition in solid tumors that correlates with metastatic potential (Vaupel & Mayer, 2007) and some ephrin family members have been implicated in the promotion of metastatic behavior (Surawska et al., 2004), we next studied the metastatic potential of MDA-MB-231 cells engineered to express luciferase and *EFNA3* in an orthotopic xenotransplantation model. Figure 27A, show luminescence signal five weeks after tumor inoculation and Figure 27B the progression of luciferate activity during the course of the experiment. At five weeks after tumor inoculation, the total photon flux was significantly higher in most animals bearing EFNA3-positive tumors, compared to controls (Figure 27A and 27B). Moreover, the necropsies performed at the end of the experiment confirmed that 7 out of 8 of the animals injected with EFNA3-positive cells presented metastasis as opposed to 1 out of 7 in the control group. Metastasis were located mostly, in the peritoneal cavity, attached to the surface of internal organs (Figure 28B and 28C)
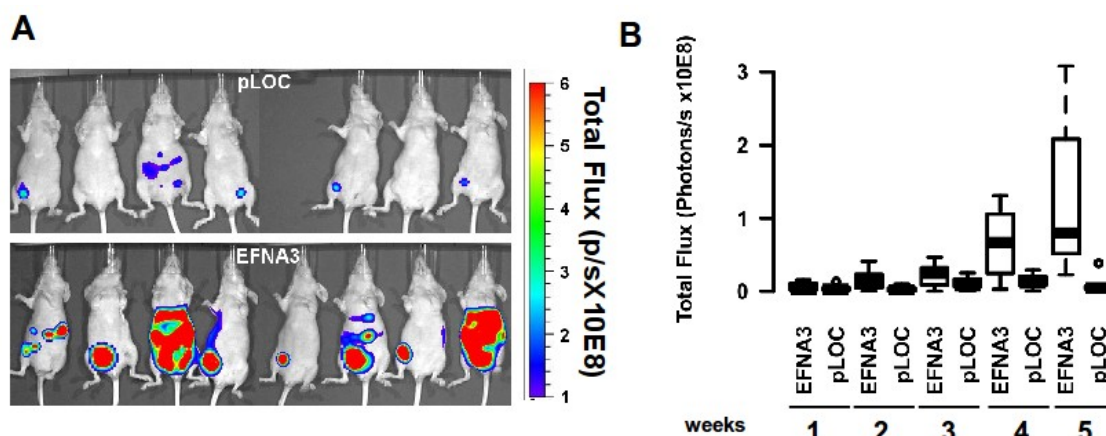
Figure 27. EFNA3 overexpression increases metastasis formation in an animal model. (A,B) MDA-MB-231 cells engineered to stably express luciferase alone (ploc-luc) or in combination with EFNA3 (pefna3-luc) were orthotopically injected into the mammary fat pad of nude mice and tumor growth followed by non-invasive whole body bioluminescence imaging. (A) Bioluminescence images of mice at end time point (5 weeks). (B) Luminescence signal during the course of the experiment. The graph represents the luciferase signal (total flux) in photons/seconds $\times 10^8$

Importantly, there were no significant differences in the volume of the primary tumor (Figure 28A) nor tumor vascularization (Figure 28D and 28E) between both animal groups, suggesting that the increased metastatic potential observed in the EFNA3 group was not due to increased tumor growth nor vascularization.
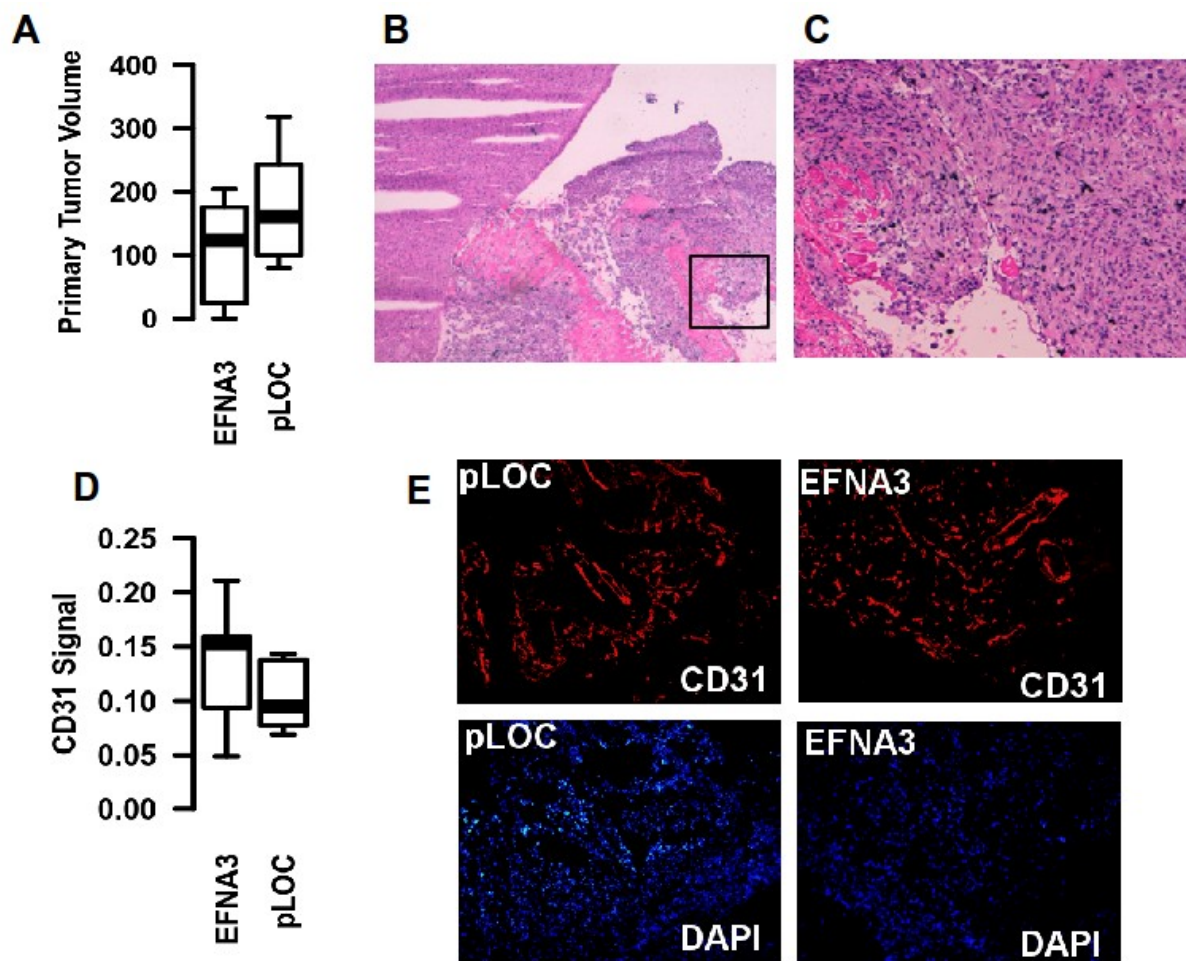
Figure 28. The increased metastatic potential observed in the EFNA3 group was not due to increased tumor growth nor vascularization (A) Tumor volume was determined after five weeks. The boxplot represents the distribution of volume in each group. The differences between controls and EFNA3-expressing tumors was not significant (independent samples t test: $t12=-1.5373$, $P=0.1499$). Animals were euthanised and tissues processed for H&E staining (B and C) or immunostaining. (E) Representative H&E staining from liver at 10X (left) and 40X (right) magnification. Blood vessels were stained with an antibody against human CD31 and cell nuclei with DAPI. (D) The CD31 signal was normalized to cellularity (DAPI staining).Te graph represents the distribution of normalized values. Differences between groups weere not significant (independent samples t-test $t10 =1.0522$, $p= 0.3182$). (E) Representative images of CD31 and DAPI staining.

Finally, the increased metastatic potential of the EFNA3 cells was confirmed in an independent experiment where cells were injected via the tail vein to generate experimental lung metastasis. As shown in figure 29, Lung weight measurement at ten weeks showed significant differences between mice control and mice overexpressing Ephrin-A3, that correlated with the differences observed in bioluminescence signal.
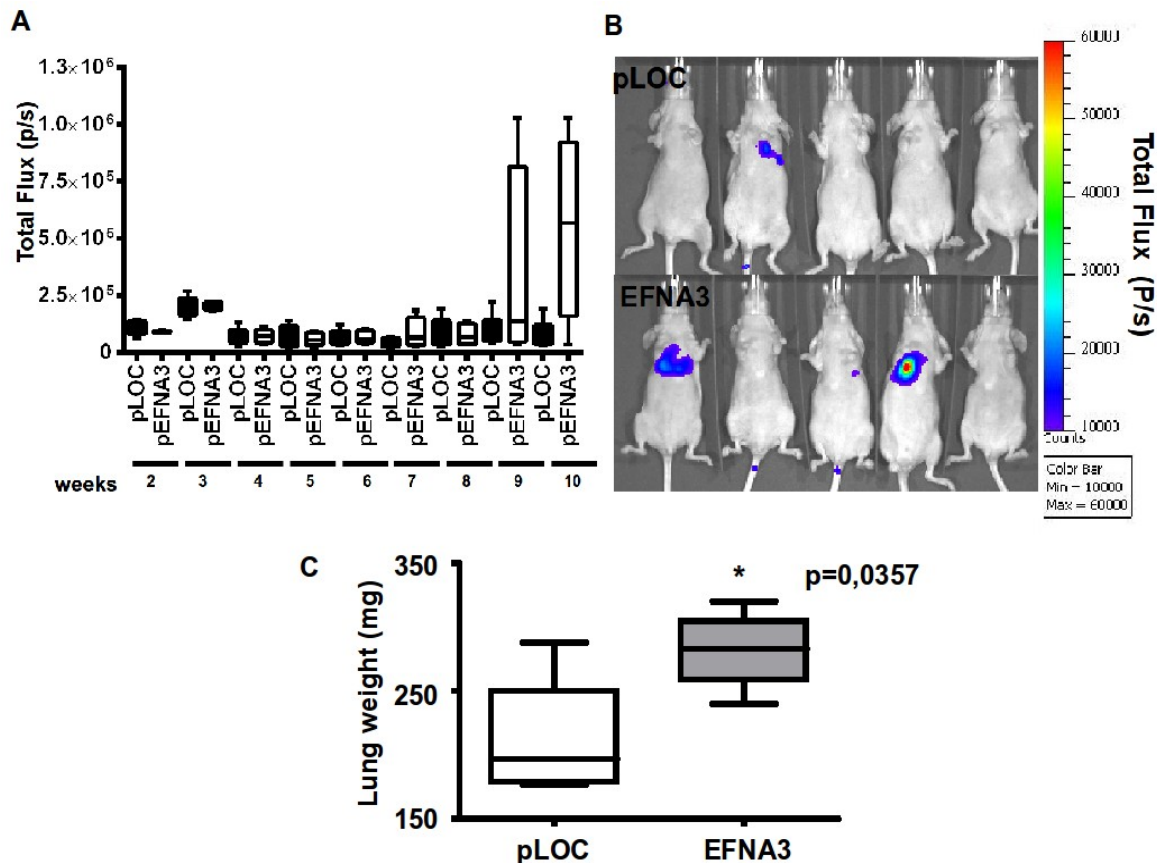
Figure 29. EFNA3 overexpression increases metastasis formation in an animal model. MDA-MB-231 cells engineered to stably express luciferase alone (pLOC) or in combination with EFNA3 (EFNA3). The cells injected via the tail vein of nude mice and tumor growth followed by non-invasive whole body bioluminescence imaging (BLI) (A) BLI image of mice at 9 weeks time point. (B) Luminescence signal during the course of the experiment. The graph represents the luciferase signal (total flux) in photons/seconds at end time point (10 weeks). (C) The weight of the lungs was determined at the end of the experiment, the boxplot represents the distribution of volume in each group. The differences between controls and EFNA3-expressing tumors were significant (independent samples t-test: t12=-2.696, p<0.05).

7. EFNA3 INDUCES ENDOTHELIAL CELL REPULSION AND PROMOTES EXTRAVASATION OF TUMOR CELLS

The experiments above demonstrate that Ephrin-A3 expression results in increased metastatic potential without affecting the microvascular density nor primary tumor growth. Ephrins, through their receptors, mediate attraction and repulsion signals between cells (Nievergall et al., 2012). This function could enable tumor cells entry and exit from blood vessels and in this way increase their metastatic potential. To test this possibility, we plated HEK293T cells transiently transfected with plasmids encoding for EFNA3 or the empty vector, on top of the HUVEC cells grown in monolayer. We measured the ability of HEK-EFNA3 cells to pass through HUVEC monolayer and attach to the plastic below. The number of GFP-positive cells attached and spread on plastic was

significantly higher when EFNA3 was expressed compared to control (figure 30). This result is consistent with EFNA3 transducing the repulsive signal to HUVECs that allowed HEK293 cells to infiltrate through the monolayer.
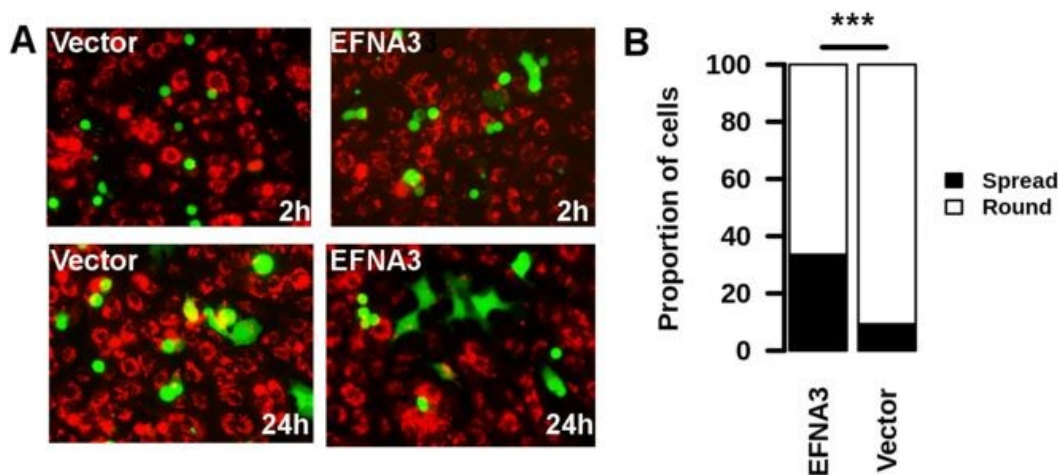
Figure 30. **EFNA3 expression results in repulsion of HUVEC cells and promotes extravasation.** EFNA3 expression results in repulsion of HUVEC cells and promotes extravasation. (A,B) HEK293 cells were transiently transfected with a plasmids encoding for GFP and EFNA3 ("EFNA3") or GFP alone ("Vector") and then plated on top of a monolayer of HUVEC cells previously labeled with a red fluorochrome. (A) Representative images of cells 2h and 24h after plating HEK293 cells on top of the HUVEC monolayer. (B) The graph represents the proportion of GFP-positive cells attached to the plastic surface and spreading at 24h. The differences between cells expressing EFNA3 and cells transfected with GFP alone was statistically significant (2x2 contingency table: χ^2_1=67.52, p<0.001).

In agreement, recombinant Ephrin-A3-Fc blocked the directional migration of HUVEC cells induced by the conditioned media from NIH 3T3 cells (Figure 31).
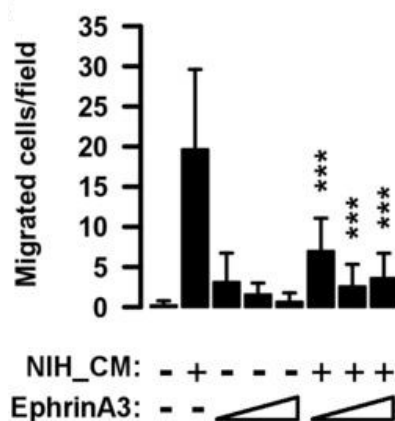
Figure 31. Recombinant Ephrin-A3-Fc block migration of HUVEC cells. HUVEC migration was determined using NIH conditioned media (NIH_CM) as chemoattractant in the absence or presence of 0.2, 1 or 5 microgr/ml of recombinant Ephrin-A3 fused to the Fc region of immunoglobulins (EphrA3). The graph represents the average number of migrated cells per field in a single experiment and error bars the standard deviation. The experiment was repeated three independent times with similar results. The differences among treatments werestatistically significant (ANOVA $F_{7,178}$=47.17, p<0.001) and the asterisks indicate means that were the statistically significant to the NIH-conditioned media treatment (***, adjusted p < 0.001) in a posteriori Tukey test.

These data suggests a mechanism for the increased metastatic potential of the tumors with EFNA3 expression whereby they can efficiently intravasate and extravasate through the vascular wall. To investigate this possibility *in vivo*, we injected MDA-MB-231 cells via the tail vein of immune deficient mice and assessed the appearance of the individual tumoral cells in the lungs in a short-term assay. Although the numbers of circulating tumor cells immediately after injection, were similar in the groups injected with EFNA3 and control cells, 8 days later the GFP signal was much higher than for the animals injected with EFNA3-expressing cells (Figures 32A and 32B). Therefore, EFNA3 facilitates metastasis, at least in part, by allowing extravasation of the tumor cells from the vasculature.
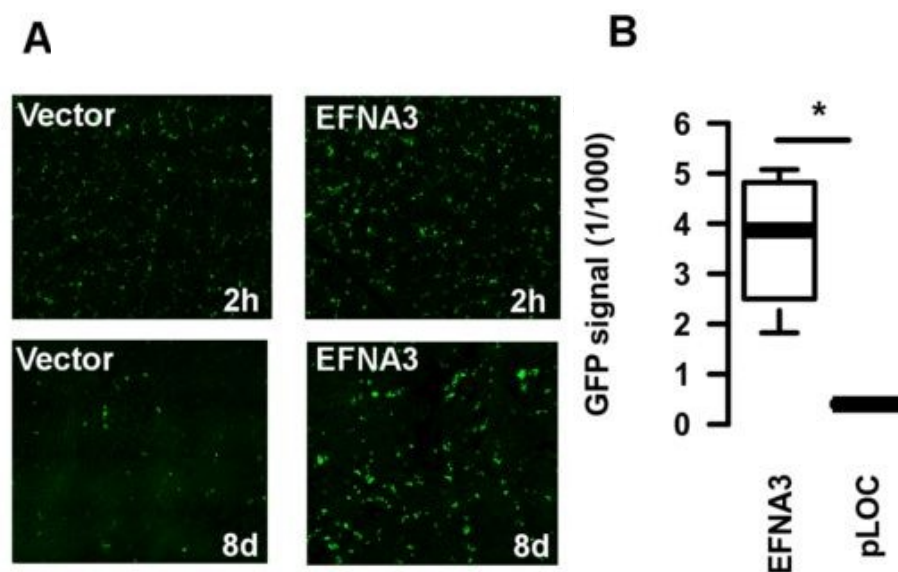
Figure 32. Lung extravasation assay. MDA-MB-231 cells stably expressing GFP alone or in combination with EFNA3 were injected in the tail of nude mice. After the indicated periods of time, mice were euthanized and lungs were examined by confocal microscopy for the presence of GFP-positive cells. The boxplot represents the distribution of total GFP signal normalized to lung area in each group of mice. Results from two independent experiments were pooled. The differences between the two groups were statistically significant (Student's t-test t(3) = -4.44, p<0.05).

8. EFNA3 EXPRESSION CORRELATES WITH METASTATIC RISK IN BREAST CANCER PATIENTS

In knowledge ofthese results, we next analyzed the correlation between EFNA3 expression and the risk of metastasis in breast cancer patients. We made use of the public access datasets of human breast cancer gene expression from the ROCK website (Online Breast Cancer Knowledgebase, http://www.rock.icr.ac.uk/). In addition to gene expression profiles, this website provides detailed clinical data allowing for correlation analyses between genes and phenotypes (Sims D., et al., 2010). We categorized patient samples according to EFNA3 expression level (high vs. low) and analyzed the incidence of metastasis over time in both groups (figure 33). In agreement with the

results of the animal studies, the risk of metastatic disease was significantly higher in patients whose primary tumors expressed higher levels of EFNA3 (figure 33). Noteworthy, the probes used to determine EFNA3 mRNA in these sets of human tumors were directed against regions common to all RNA products of the EFNA3 locus (see materials and methods). Thus they detect the cumulative signal by the mRNA and lncRNA, whereby the signal is largely generated by the lncRNA.
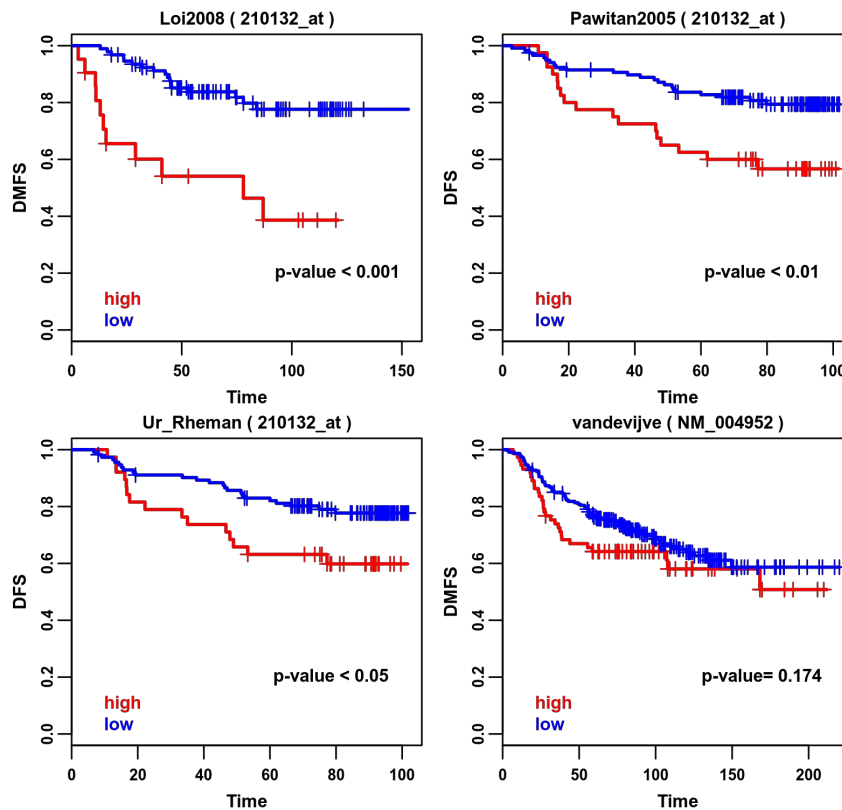
Figure 33 High levels of EFNA3 correlate with metastasis in human tumors. High levels of EFNA3 correlate with metastasis in human tumors. Gene expression profiles from the indicated series of breast cancer tumors (Loi2008, (Loi et al., 2008) ; Pawitan2005, (Pawitan et al., 2005) ; Ur-Rehman, www.rock.icr.ac.uk/; vandeVijve, (van de Vijver et al., 2002)) were downloaded from the ROCK database (www.rock.icr.ac.uk/). Samples were categorized according to EFNA3 expression into high (samples with EFNA3 expression in the top quartile of the series, labeled in red) and low expression (rest of samples in the serie, labeled in blue) and the Kaplan−Meier estimate of the distant metastasis free survival (DMFS) over time calculated (graphs). The survival of the two groups was compared using a Cox proportional hazards model, the p-values are indicated in the graphs. The probes used in the gene profiling assays are indicated in brackets in the graph title.

DISUSSION

HIF-mediated transcriptional response plays a central role in the adaptation of tissues to hypoxia. For this reason, the identification of HIF-regulated genes is essential to draw a comprehensive picture of these adaptive mechanisms. In the past few years, many HIF-target genes have been characterized and, in the majority of cases, HIF activation results in induction of transcription of the coding mRNA. In this study, we not only identified a novel HIF target gene, but also found a completely novel way of hypoxic regulation of gene expression whereby the changes in the coding mRNA are only marginal. Instead, HIF directly regulates the levels of long non-coding (lnc) RNAs, which, in turn, increase protein levels. This discovery leads to the question of how do lncRNA regulate Ephrin-A3 protein levels.

LncRNAs have been shown to regulate virtually every step of the gene expression cascade from transcription to translation (Nie, L., et al., 2012; Yoon, J. H., et al., 2012). In particular, one recent report describes lncRNA that affects Uchl1 protein levels by enhancing mRNA translation of the coding mRNA (Carrieri, C., et al, 2012), suggesting that EFNA3 lncRNA could act using a similar mechanism. Although our results show a clear effect of EFNA3 lncRNA on Ephrin-A3 expression at the postranscriptional level, further work is needed to determine whether this effect is caused by increased translation or reduced protein turnover.

Interestingly, some non-coding transcripts act in concert to regulate the expression of coding mRNAs. For example, some lncRNAs act as decoys for the miRNA that target a particular mRNA (Yoon, J. H., et al., 2012). It is known that miR-210, which is induced by hypoxia (Huang, X., et al., 2009; Kulshreshtha, R., et al., 2007), prevents the translation of several mRNA including ISCU and EFNA3 (Chan, S. Y., et al., 2009; Fasanaro, P., et al., 2008; Favaro, E., et al., 2010; Huang, X., Le, Q. B. & Giaccia, A. J., 2010). Since miR-210 binds the 3'-UTR of the EFNA3 mRNA (Fasanaro, P., et al., 2008), it is feasible that EFNA3 lncRNAs increase EFNA3 mRNA translation by depleting miR-210. However, the fact that shorter isoforms, NC1s and NC2s are more efficient in the induction of Ephrin-A3 than the long ones (NC1 and NC2), argues to the contrary because the predicted miR-210 binding sites are not represented on the shorter lncRNAs isoforms, NC1s and NC2s (figure 14). Nonetheless, it is still possible that miR-210 could bind to sites other than those predicted.

It could also be argued that lncRNA might promote Ephrin-A3 accumulation by acting at the level of EFNA3 mRNA stability rather than translation. We explored this possibility and found that

EFNA3 mRNA half-life, as determined by inhibition of transcripion by Actinomycin D treatment, was indeed increased during hypoxia compared with EFNA3 lncRNA (figure 34).
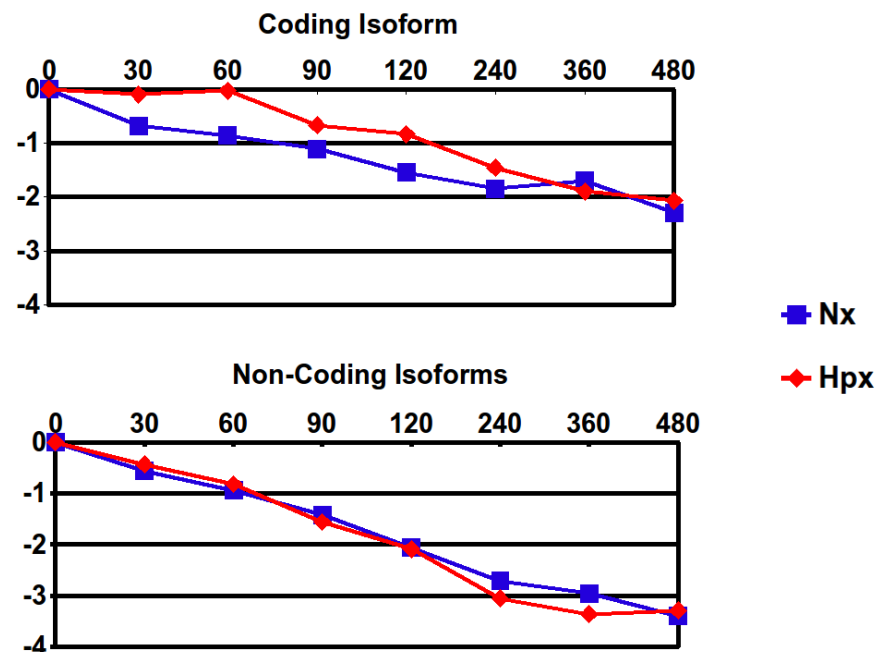
Figure 34. Hypoxia effect on the half-life of the different EFNA transcripts. HeLa cells were cultured under normoxic or hypoxic conditions for 12 h prior the inhibition of transcription with Actynomycin D (ACTD) and then kept at normoxia or hypoxia for the indicated periods of time (minutes) until processing. Non-coding isoform or coding EFNA3 mRNA levels were determined by qPCR at the indicated periods of time after the addition of ACTD.

However, when we analyzed the half life by means of 4-thio-Uridine labelling (Rädle, B., et al, 2013), we were unable to detect any increment in the half life of coding EFNA3 under hypoxic conditions (table 6). Thus, we can not draw any consistent conclusion in regards to the effect of hypoxia on EFNA3 coding mRNA stability.

	Experiment 1		Experiment 2	
	Normoxia	**Hypoxia**	**Normoxia**	**Hypoxia**
Coding EFNA3	4,49	5,37	10,52	3,13
lncRNA EFNA3	1,38	1,09	1,58	0,55
GYS1	18,93	28,65	12,29	19,17
HIF1α	19,17	50,34	14,73	21,18
EPAS	12,80	20,38	13,17	17,77

Table 6. Half life of the indicated transcripts as determined by 4-thio-Uridine labelling (Rädle, B., et al, 2013).

Although our data suggests that the major effect of lncRNA occurs at the postranscriptional level, it is possible that they act at multiple levels. Thus, another plausible mechanism by which lncRNA could contribute to Ephrin-A3 expression is that HIF-induced lncRNA transcription can act creating

a permissive chromatin environment that facilitates the expression of nearby mRNAs (Kornienko, A. E., 2013). In our case, transcription from Prom2 could facilitate transcription of the coding EFNA3 from Prom1 (figure 35). This could explain the small, but reproducible, increase in EFNA3 mRNA observed under hypoxia.
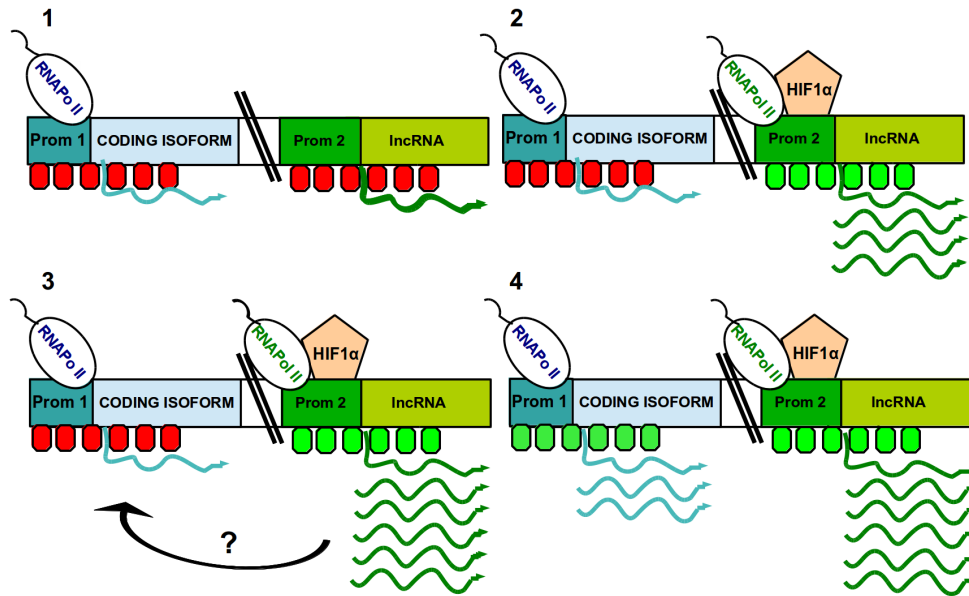
Figure 35. LncRNA transcription allows enhancer activity: 1: An inactive enhancer with closed chromatin cannot activate protein coding gene. 2 y 3 transcription of the enhancer open chromatin 4. HIF1A can bind to the enhancer and interact with coding isoform promoter allowing its expression.

In fact, this hypothesis might explain why hypoxia has a minor, but consistent, effect on the level of the coding mRNA as well. Moreover, this possibility is supported by the analysis of newly synthesized heterogeneous nuclear RNA (hnRNA) in cells grown in the presence of 4-thio-Uridine (Rädle, B., et al, 2013). As shown in figure 36, hypoxia induces newly transcribed coding EFNA3 hnRNA, but the kinetic of this induction is retarded compared with that of the hnRNA precursor of the EFNA3 lncRNA or EGLN3 hnRNA. Thus, in this model, hypoxia would result in the recruitment of HIF and RNA pol II to Prom2 (as supported by our data in figure 23) leading to increased transcription of the lncRNA. The increased transcription from Prom2, could indirectly enhance transcription from Prom1 by, for example, creating a permissive chromatin environment within the EFNA3 locus. The increased activity of Prom1 would be secondary and consequence of the transcription from Prom1 and does not involve recruitment of RNA-pol II or HIF to this promoter (figure 23 and figure 24).
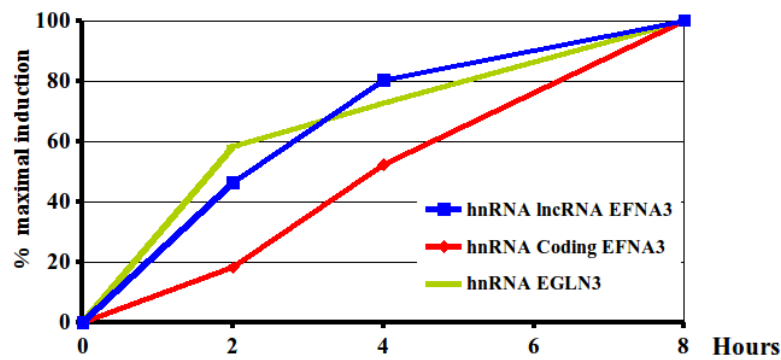
Figure 36. hnRNA levels under 4-thio-Uridine treatment. 400mM of 4-thio-Uridine was added to HeLa cells medium the last 2 hours of hypoxia treatment. hnRNA levels were quantified with the specific set of primers.

From our data, it is quite evident that hypoxia preferentially regulates the transcription of the EFNA3 lncRNA from Prom2 and yet Prom1 is not far away from the HIF binding site. Hence the question on how hypoxia affects differentially these promoters arises. Interestingly, the presence of CTCF binding sites flanking EFNA3 Prom2 region (figure 15) offers a plausible explanation for the preferential transcriptional activation of the lncRNA by HIF and the lack of the regulation of the canonical coding isoform. Other studies also demonstrate the role of insulators in determination of HIF specificity for promoters located in the close proximity (Tiana, M., et al., 2011). Of note, in agreement with the model proposed above, the presence of CTCF binding sites would block the activity of cis-regulators such as HIF, but will probably not have an effect on trans-acting regulators, such as lncRNA.

Also regarding the molecular mechanism, we found that in vitro EFNA3 induction was dependent on HIF-1α and not affected by HIF-2α knockdown whereas in animal models, EPAS1 was required for EFNA3 induction upon VHL loss. Thus, HIF isoforms responsible for EFNA3 may be context-dependent and could be dictated by availability of a specific HIF isoform, or another, HIF cell-type specific transcription factor(s), that favor interaction with a specific HIF isoform. Nonetheless, this has been observed for other hypoxia-regulated genes. For instance, the hypoxic regulation of another ephrin, Ephrin-A1, also requires HIF1α (Vihanto, M. M., et al., 2005) in cell cultures, but not in vivo, where it is dependent on HIF2α (Yamashita, T., et al., 2008).

Taken together, our results and those published elsewhere point to a complex regulation mediated

by non coding RNAs of different types and strongly suggest that Ephrin-A3 levels need to be tightly controlled in the response to hypoxia. This fact leads to the question of what is the role of Ephrin-A3 in the physiological response to hypoxia. Since the induction of angiogenesis is one of the paradigmatic responses to hypoxia and other ephrins have key roles in angiogenesis and vasculogenesis (Mosch, B., Reissenweber, B., Neuber, C. & Pietzsch, J., 2010), we decided to investigate the role of EFNA3 in angiogenesis.

To address this question we decided study the effect of forced Ephrin-A3 expression in an *in vivo* model of tumoral vascularization. Using this approach, we did not found evidence of a role for Ephrin-A3 on the control of angiogenesis. In fact, the vascular density of tumors expressing this protein was similar to that of controls (figures 28D and 28E). This is consistent with the fact that Ephrin-A3 KO mice (Pfeiffenberger, C., et al., 2005) have no vascular phenotype described so far. Earlier studies showed that Ephrin-A3 overexpression prevented the formation of capillary-like structures by HUVEC cells *in vitro* (Fasanaro, P., et al., 2008; Xiao, F., et al., 2013). We similarly found that recombinant Ephrin-A3-Fc prevents the migration of HUVECs in vitro (figure 31), however, we interpret this result as an evidence of the repulsive activity of EFNA3 than explains the enhanced ability of Ephrin-A3 expressing cells to enter/exit blood vessels.

A further biological process in which ephrins play a key role is axon pathfinding (Flanagan, J. G., 2006). Interestingly, recent works in C. elegans (Pocock, R. & Hobert, O., 2008) and zebrafish (Stevenson, T. J., et al., 2012) have shown that hypoxia disrupts normal CNS pathfinding through the induction of Eph/ephrins in a HIF-dependent manner. Moreover, Ephrin-A3 protein is induced at very high levels in post-ischemic mouse hippocampus in vivo (Pulkkinen, K., Malm, T., Turunen, M., Koistinaho, J. & Ylä-Herttuala, S., 2008). Thus, it is tempting to speculate that the physiological role of Ephrin-A3 induction by hypoxia could be related to axon guidance in the CNS.

Regardless their physiological roles, Eph/ephrins are frequently co-opted by transformed cells to modulate cell-cell and cell-matrix attachment and survival during invasion, angiogenesis and metastasis (Chen, J., 2012; Nievergall, E., et al., 2012). Accordingly, we found that the expression of Ephrin-A3 in MDA-MB231 breast cancer cells resulted in increased metastatic potential (figure 27 and 29) suggesting that the activation of HIF as a consequence of intratumoral hypoxia or

oncogenic alterations could promote metastasis through the induction of Ephrin-A3. Supporting this possibility we found a strong correlation between high EFNA3 level and development of metastasis (figure 26 and 33).

Also in agreement with our findings, Ephrin-A3 was recently identified as one of the proteins differentially expressed in metastatic colon cancer cells (Barderas, R., et al., 2013) and in squamous cell carcinoma in lung associated with poor prognosis (Li, J. J., 2012). As for the mechanism by which Ephrin-A3 promotes metastasis, our results suggests that tumor cells expressing Ephrin-A3 might have enhanced ability to intra and extravasate and this ability be related to the repulsion of endothelial cells by tumor cells expressing Ephrin-A3 (figure 30). Interestingly, this mechanism is similar to the one employed by angiopoietin-like 4 (ANGPTL4), which is also HIF-dependent hypoxia-inducible gene, to promote metastasis (Zhang , H., et al., 2011).

In summary, we demonstrate a novel mechanism of hypoxic regulation, where EFNA3 expression is regulated by previously non-annotated intragenic lncRNAs whose transcription is driven by HIF in response to hypoxia. We propose that in hypoxic regions of solid tumors or in tumors harboring mutations leading to HIF stabilization, this mechanism allows Ephrin-A3 accumulation and the promotion of metastasis by augmenting the ability of tumoral cell to pass through the endothelium of blood vessels (figure 37).
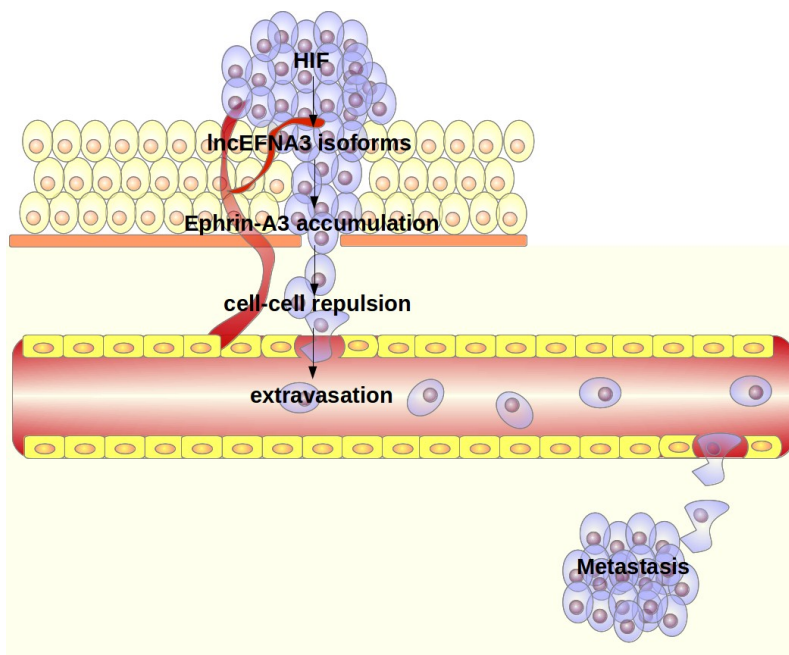
Figure 37: EphrinA3 metastasis induction scheme

CONCLUSIONS

➢ **Ephrin-A3 protein and EFNA3 transcripts are induced by hypoxia.**

➢ **EFNA3 locus encodes for a coding mRNA and several lncRNAs that are transcribed from alternative promoters.**

➢ **Hypoxia induces the transcription of the lncRNA through the binding of HIF to an HRE element located within the second exon of the gene.**

➢ **Hypoxia promotes Ephrin-A3 protein accumulation by a mechanism that is likely to be controlled by EFNA3 lncRNAs acting at multiple levels including the postrancriptional one.**

➢ **EFNA3 expression correlates with HIF activity in human renal carcinomas**

➢ **The overexpression of Ephrin-A3 results in increased metastatic potential without altering tumor growth nor vascularization.**

➢ **The increased metastatic potential associated to Ephrin-A3 expression is probably due to enhanced extravasation of tumor cells as a consequence of their repulsive effect on endothelial cells.**

➢ **EFNA3 expression correlates with higher metastatic potential and poor prognosis in human breast cancer samples**

➢ **La proteína Efrina-A3 y los tránscritos de EFNA3 son inducido por hipoxia**

➢ **El locus EFNA3 codifica para un mRNA codificante y varios RNAs no codificantes largos (lncRNAs) desde el locus EFNA3 transcritos desde un promotor alternativo**

➢ **La hipoxia induce la transcripción de RNAs no codificantes largos (lncRNAs) desde el locus EFNA3 mediante la unión de HIF a un HRE localizado en el segundo exón del gen**

➢ **La hipoxia promueve la acumulción de Efrina-A3 por un mecanismo que parece controlado por RNAs no codificantes largos (lncRNAs) desde el locus EFNA3, actuando a múltiples niveles de regulación.**

➢ **La expresión de EFNA3 correlaciona con actividad de HIF en carcinomas renales humanos**

➢ **La sobreexpresión de Efrina-A3 se traduce en mayor potencial metastásico sin alterar el crecimiento del tumor ni su vascularización**

➢ **El mayor potencial metastásico asociado a la expresión de Efrina-A3, es probablemente debido al aumento de la extravasación de las células tumorales como consecuencia de su efecto de repulsión sobre las células endoteliales.**

➢ **La expresión de EFNA3 correlaciona con un mayor potencial metastásico y peor prognosis en muestras de cáncer de mama humano**

BIBLIOGRAPHY

Ambros, V.(2004). **The functions of animal microRNAs**. Nature. 2004 Sep 16;431(7006):350-5. Review.

Barderas, R., Mendes, M., Torres, S., Bartolomé, R. A., López-Lucendo, M., Villar-Vázquez, R., Peláez-García, A., et al. (2013). **In-depth Characterization of the Secretome of Colorectal Cancer Metastatic Cells Identifies Key Proteins in Cell Adhesion, Migration, and Invasion**. Molecular & cellular proteomics: MCP, 12(6),1602–20.

Bartel, D. P. (2004). **MicroRNAs: genomics, biogenesis, mechanism, and function.** Cell. 2004 Jan 23;116(2):281-97. Review.

Benita Y, Kikuchi H, Smith AD, Zhang MQ, Chung DC, Xavier RJ. (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(14):4587-602.

Bishop J. M. (1991). **Molecular themes in oncogenesis.** Cell, Vol. 64, 235-246, January 25, 1991.

Blits-Huizinga, C. T., Nelersa, C. M., Malhotra, A., Liebl, D. J. (2004). **Ephrins and their receptors: binding versus biology.** IUBMB Life; 56(5):257-65.

Blume-Jensen P., Hunter T. (2001). **Oncogenic kinase signaling**. Nature. 2001 May 17;411(6835):355-65.

Bütikofer P., Malherebe T., Bosching M., Roditi I. (2001). **GPI-anchored proteins: now you see 'em, now you don't.** FASEB J. 2001 Feb;15(2):545-8.

Carmona M. A., Murai K. K., Wang L., Roberts A. J., Pasquale E. B. (2009).**Glial ephrin-A3 regulates hippocampal dendritic spine morphology and glutamate transport.** Proc Natl Acad Sci USA. 2009 Jul 28;106(30):12524-9.

Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., Pesce, E., et al. (2012). **Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat**. Nature, 491(7424), 454–7.

Chan, S. Y., Zhang, Y.-Y., Hemann, C., Mahoney, C. E., Zweier, J. L., & Loscalzo, J. (2009). **MicroRNA-210 controls mitochondrial metabolism during hypoxia by repressing the iron-sulfur cluster assembly proteins ISCU1/2**. Cell metabolism,10(4),273–84.

Chen, J. (2012). **Regulation of tumor initiation and metastatic progression by Eph receptor tyrosine kinases**. Advances in cancer research (1st ed., Vol. 114, pp. 1– 20). Elsevier Inc.

Cheng, N., Brantley, D. M., Chen, J. (2002).**The ephrins and Eph receptors in angiogenesis**. Cytokine Growth Factor Rev.; 13(1):75-85. Review.

Cumberledge, S., Zaratzian, A., Sakonju, S. (1990).**Characterization of two RNAs transcribed from the cis-regulatory region of the abd-A domain within the Drosophila bithorax complex.** Proc Natl Acad Sci USA. 1990 May;87(9):3259-63.

Dayan, F., Roux, D., Brahimi-Horn, M. C., Pouyssegur, J., Mazure, N. M. (2006). **The oxygen sensor factor-inhibiting hypoxia-inducible factor-1 controls expression of distinct genes through the bifunctional transcriptional character of hypoxia-inducible factor-1alpha.** Cancer Res. 1;66(7):3688-98.

Djebali, S., Carrie, A. D., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A. M., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., Xue, C., Marinov, G. K., Khatun, J., Williams, B. A., Zaleski, C., Rozowsky, J., Röder, 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 F., 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., 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., Guigó, R., Gingeras, T. R. (2013). **Landscape of transcription in human cells.** Nature. 2012 September 6; 489(7414): 101–108.

Elorza A., Soro-Arnáiz I., Meléndez-Rodríguez F., Rodríguez-Vaello V., Marsboom G., de Cárcer G., Acosta-Iborra B., Albacete-Albacete L., Ordóñez A., Serrano-Oviedo L., Giménez-Bachs J. M., Vara-Vega A., Salinas A., Sánchez-Prieto R., Martín del Río R., Sánche-Madrid F., Malumbres M., Landázurri M. O., Aragonés J. **HIF2α acts as an mTORC1 activator through the amino acid carrier SLC7A5.** Mol Cell. 2012 Dec 14;48(5):681-91.

Epstein, A. C., Gleadle, J. M., McNeill, L. A., Hewitson, K. S., O'Rourke, J., Mole, D. R., Mukherji, M., et al. (2001). **C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation.** Cell, 107(1),43-54. Retrieved from http://www.sciencedirect.com/science/article/pii/S0092867401005074

Fasanaro, P., D'Alessandra, Y., Di Stefano, V., Melchionna, R., Romani, S., Pompilio, G., Capogrossi, M. C., et al. (2008). **MicroRNA-210 modulates endothelial cell response to**

hypoxia and inhibits the receptor tyrosine kinase ligand Ephrin-A3. The Journal of biological chemistry, 283(23), 15878–83.

Favaro, E., Ramachandran, A., McCormick, R., Gee, H., Blancher, C., Crosby, M., Devlin, C., et al. (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(4), e10345.

Firth JD, Ebert BL, Pugh CW, Ratcliffe PJ. (1994) **Oxygen-regulated control elements in the phosphoglycerate kinase 1 and lactate dehydrogenase A genes: similarities with the erythropoietin 3' enhancer**. Proc Natl Acad Sci U S A. 1994 Jul 5;91(14):6496-500.

Flanagan, J. G. (2006). **Neural map specification by gradients**. Current opinion in neurobiology, 16(1), 59–66.

Gale, N. W., Holland, S. J., Valenzuela, D. M., Flenniken, A., Pan, L., Ryan, T. E., Henkemeyer, M., Strebhardt, K., Hirai, H., Wilkinson, D. G., Pawson, T., Davis, S., Yancopoulos, G. D. (1996). **Eph receptors and ligands comprise two major specificity subclasses and are reciprocally compartmentalized during embryogenesis.** Neuron.; 17(1):9-19.

Gauthier, L. R., Robbins, S. M. (2003)**Ephrin signaling: One raft to rule them all? One raft to sort them? One raft to spread their call and in signaling bind them?** Life Sci. 2003 Dec 5;74(2-3):207-16. Review.

Gu, C., Park, S. (2001). **The EphA8 receptor regulates integrin activity through p110gamma phosphatidylinositol-3 kinase in a tyrosine kinase activity- independent manner**. Mol Cell Biol.; 21(14):4579-97.

Haase, V. H. (2010). **Hypoxic regulation of erythropoiesis and iron metabolism**. American journal of physiology. Renal physiology, 299(1), F1–13.

Hafner, C., Schmitz, G., Meyer, S., Bataile, F., Hau, P., Langmann, T., Dietmaier, W., Landthaler, M., Vogt, T. (2004). **Differential gene expression of Eph receptors and ephrins in benign human tissues and cancers**. Clin Chem.;50(3):490-9.

Hattori, M., Osterfield, M., Flanagan, J. G. (2000). **Regulated cleavage of a contact-mediated axon repellent**. Science 25;289(5483):1360-5.

Hewitson, K. S., McNeill, L. a, Riordan, M. V, Tian, Y.-M., Bullock, A. N., Welford, R. W., Elkins, J. M., et al. (2002). **Hypoxia-inducible factor (HIF) asparagine hydroxylase is identical to factor inhibiting HIF (FIH) and is related to the cupin structural famil**y.

The Journal of biological chemistry, 277(29), 26351–5.

Himanen JP, Chumley MJ, Lackmann M, Li C, Barton WA, Jeffrey PD, Vearing C, Geleick D., Feldheim DA, Boyd AW, Henkemeyer M, Nikolov DB. (2004). **Repelling class discrimination: ephrin-A5 binds to and activates EphB2 receptor signaling.** Nat Neurosci. 2004 May;7(5):501-9. Epub 2004 Apr 25.

Huang L. E., Z. Arany, Z., Livingston, D. M., Bunn H. F. (1996). **Activation of hypoxia-inducible transcription factor depends primarily upon redox-sensitive stabilization of its alpha subunit**. J. Biol Chem 271 (50):32253-32259.

Huang, X., Ding, L., Bennewith, K. L., Tong, R. T., Welford, S. M., Ang, K. K., Story, M., et al. (2009). **Hypoxia-inducible mir-210 regulates normoxic gene expression involved in tumor initiation**. Molecular cell, 35(6), 856–67.

Huang, X., Le, Q.-T., & Giaccia, A. J. (2010). **MiR-210--micromanager of the hypoxia pathway**. Trends in molecular medicine, 16(5), 230–7.

Huarte, M., Rinn, J. L. (2010). **Large non-coding RNAs: missing links in cancer?**. Hum Mol Genet. 2010 Oct 15;19(R2):R152-61.

Hunter T., Cooper J. A. (1985). **Protein-tyrosine kinases.** Annu Rev Biochem. 1985;54:897-930.

Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., Salic, a, et al. (2001). **HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing**. Science (New York, N.Y.), 292(5516), 464–8.

Iwai K, Yamanaka K, Kamura T, Minato N, Conaway RC, Conaway JW, Klausner RD, Pause A. (1999). **Identification of the von Hippel-lindau tumor-suppressor protein as part of an active E3 ubiquitin ligase complex.** Proc Natl Acad Sci USA. 26;96(22):12436-41.

Jaakkola, P., Mole, D. R., Tian, Y. M., Wilson, M. I., Gielbert, J., Gaskell, S. J., Kriegsheim A., et al. (2001). **Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. Scienc**e (New York, N.Y.), 292(5516), 468–72.

Jiang, B. H., Zheng, J. Z., Leung, S. W., Roe, R., Semenza, G. L.(1997). **Transactivation and inhibitory domains of hypoxia-inducible factor 1alpha**. **Modulation of transcriptional activity by oxygen tension.**The Journal of biological chemistry, 272(31), 19253–60. Retrieved from http://www.jbc.org/content/272/31/19253.short

Kaelin, W. G. (2007). **Von Hippel-Lindau disease**. Annual review of pathology, 2, 145–73.

Kornienko, A. E., Guenzl, P. M., Barlow, D. P., Pauler, F. M. (2013) **Gene regulation by the act of**

long non-coding RNA transcription. BMC Biol. **2013** May 30;11:59.

Kulshreshtha, R., Ferracin, M., Wojcik, S. E., Garzon, R., Alder, H., Agosto-Perez, F. J., Davuluri, R., et al. (2007). **A microRNA signature of hypoxia**. Molecular and cellular biology, 27(5), 1859–67.

Latos, P. A., Pauler, F. M., Koerner, M. V., Şenergin, H. B., Hudson, Q. J., Stocsits, R. R., Allhoff, W., Stricker, S. H., Klement, R. M., Warczok, K. E., Aumayr, K., Pasierbek, P., Barlow, D. P. (2012). **Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing.** Science. 2012 Dec 14;338(6113):1469-72.

Lando, D., Peet, D. J., Gorman, J. J., Whelan, D. a, Whitelaw, M. L., & Bruick, R. K. (2002). **FIH-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor**. Genes & development, 16(12),1466–71.

Lenburg, M. E., Liou, L. S., Gerry, N. P., Frampton, G. M., Cohen, H. T., & Christman, M. F. (2003). **Previously unidentified changes in renal cell carcinoma gene expression identified by parametric analysis of microarray dat**a. BMC cancer, 3, 31.

Li, J. J., Xie, D. (2012). **The roles and therapeutic potentials of Ephs and ephrins in lung cancer.** Exp Cell Res. 2013 Jan 15;319(2):152-9.

Lisztwan, J., Imbert, G., Wirbelauer, C., Gstaiger, M., Krek, W. (1999). **The von Hippel-Lindau tumor suppressor protein is a component of an E3 ubiquitin-protein ligase activity.** Genes Dev. 1999 Jul 15;13(14):1822-33.

Loenarz, C., Coleman, M. L., Bleininger, A., Schierwater, B., Holand, P. W., Ratcliffe, P. J., Schofield, C. J. (2011).**The hypoxia-inducible transcription factor pathway regulates oxygen sensing in the simplest animal, Trichoplax adhaerens.** 2011 Jan;12(1):63-70.

Loi, S., Haibe-Kains, B., Desmedt, C., Wirapati, P., Lallemand, F., Tutt, A. M., Gillet, C., et al. (2008). **Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen**. BMC genomics,9,239.

Lopez-Barneo, J., Ortega-Saenz, P., et al. (2009). **Oxygen sensing in the carotid body**. Ann N Y Acad Sci 1177:119-131.

Lu, X., & Kang, Y. (2010). **Hypoxia and hypoxia-inducible factors: master regulators of metastasis**. Clinical cancer research 16(24), 5928–35.

Mahon PC, Hirota K, Semenza GL. (2001). **FIH-1: a novel protein that interacts with HIF-1alpha and VHL to mediate repression of HIF-1 transcriptional activity.** Genes Dev. 15;15(20):2675-86.

Martens, J. A., Laprade, L., Winston, F. (2004) **Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene.** Nature. 2004 Jun 3;429(6991):571-4.

Maxwell, P. H., Wiesener, M. S., Chang, G. W., Clifford, S. C., Vaux, E. C., Cockman, M. E., Wykoff, C. C., Pugh, C. W., Maher, E. R., & Peter J. Ratcliffe. (1999). **The tumor suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis.** Nature, 399(6733), 271–5.

McCormick, R., Buffa, F. M., Ragoussis, J., Harris, A.L. (2010). T**he role of hypoxia regulated microRNAs in cancer.** Curr Top Microbiol Immunol. 2010;345:47-70.

Miao, H., Strebhardr, K., Pasquale, E. B., Shen, T. L., Guan, J. L., Wang, B. (2005). Inhibition of integrin-mediated cell adhesion but not directional cell migration requires catalytic activity of EphB3 receptor tyrosine kinase. Role of Rho family small GTPases. J Biol Chem. 14;280(2):923-32.

Min J. H., Yang, H., Ivan, M. Gertler,, F., Kaelin, W. G Jr., Pavletich, N. P. (2002). **Structure of an HIF-1alpha -pVHL complex: hydroxyproline recognition in signaling**. Science 2002 Jun 7;296(5574):1886-9.

Miró-Murillo, M., Elorza, A., Soro-Arnáiz, I., Albacete-Albacete, L., Ordoñez, A., Balsa, E., Vara-Vega, A., et al. (2011). **Acute Vhl Gene Inactivation Induces Cardiac HIF-Dependent Erythropoietin Gene Expression**. (M. Rojas, Ed.)PLoS ONE, 6(7), e22589.

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.**

Mosch, B., Reissenweber, B., Neuber, C., & Pietzsch, J. (2010). **Eph receptors and ephrin ligands: important players in angiogenesis and tumor angiogenesis**. Journal of oncology, 2010, 135285.

Murai, K. K., Pasquale, E. B. (2004). **Eph receptors, ephrins, and synaptic function.** Neuroscientist.;10(4):304-14. Review.

Nakamoto, M., Bergemann, A. D. (2002). **Diverse roles for the Eph family of receptor tyrosine kinases in carcinogenesis.** Microsc Res Tech. 2002 Oct 1;59(1):58-67.

Ng, P., Wei, C., Sung, W., Chiu, K. P., Lipovich, L., Ang, C. C., Gupta, S., et al. (2005). **Gene identification signature (GIS) analysis for transcriptome characterization and genome annotation**. Nature methods, 2(2), 105–11.

Nie, L., Wu, H., Hsu, J.-M., Chang, S., Labaff, A. M., Li, C., Wang, Y., et al. (2012). **Long non-coding RNAs: versatile master regulators of gene expression and crucial players in cancer.** American journal of translational research, 4(2), 127–50. Retrieved fromhttp://www.pubmedcentral.nih.gov/articlerender.fgi?artid=3353529&tool=pmcentrez&rendertype=abstract

Nievergall, E., Lackmann, M., & Janes, P. W. (2012). **Eph-dependent cell-cell adhesion and segregation in development and cancer.** Cellular and molecular life sciences CMLS, 69(11), 1813–42.

Ogawa, K., Pasqualini, R., Lindberg, R. a, Kain, R., Freeman, a L., & Pasquale, E. B. (2000). **The ephrin-A1 ligand and its receptor, EphA2, are expressed during tumor neovascularization**. Oncogene, 19(52), 6043–52.

Okubo, T., Yanai, N., Obinata, M. (2006). **Stromal cells modulate ephrinB2 expression and transmigration of hematopoietic cells**. Exp Hematol.34(3):330-8.

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 research, 38(7), 2332–45.

Palmer, A., Zimmer, M., Erdmann, K. S., Eulenburg, V., Porthin, A., Heumann, R., Deutsch, U., Klein, R. (2002). **EphrinB phosphorylation and reverse signaling: regulation by Src kinases and PTP-BL phosphatase**. Mol Cell;9(4):725-37.

Pascall JC, Brown KD. (2004). **Intramembrane cleavage of ephrinB3 by the human rhomboid family protease, RHBDL2.** Biochem Biophys Res Commun. 23;317(1):244-52.

Pasquale, E. B. (2010). **Eph receptors and ephrins in cancer: bidirectional signaling and beyond**. Nature reviews. Cancer, 10(3), 165–80.

Pauli, A., Valen, E., Lin, M. F., Garber, M., Vastenhouw, N. L., Levin, J. Z., Fan, L., Sandelin, A., Rinn, J. L., Regev, A., Schier, A. F. (2011)**Systematic identification of long noncoding RNAs expressed during zebrafish embryogenesis.** Genome Res. 2012 Mar;22(3):577-91.

Pawson T, Saxton TM. (1999). **Signaling networks--do all roads lead to the same genes?** Cell.;97(6):675-8. Review.

Pawitan, Y., Bjöhle, J., Amler, L., Borg, A.-L., Egyhazi, S., Hall, P., Han, X., et al. (2005). **Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts**. Breast cancer

research : BCR, 7(6), R953–64.

Pfeiffenberger, C., Cutforth, T., Woods, G., Yamada, J., Rene, C., Copenhagen, D. R., Flanagan, J. G., et al. (2005). **Ephrin-As and neural activity are required for eye-specific patterning during retinogeniculate mapping.** NIH Public Access, 8(8), 1022–1027.

Pocock, R., & Hobert, O. (2008). **Oxygen levels affect axon guidance and neuronal migration in Caenorhabditis elegans**. Nature neuroscience, 11(8), 894–900.

Pulkkinen, K., Malm, T., Turunen, M., Koistinaho, J., & Ylä-Herttuala, S. (2008). **Hypoxia induces microRNA miR-210 in vitro and in vivo ephrin-A3 and neuronal pentraxin 1 are potentially regulated by miR-210**. FEBS letters, 582(16), 2397– 401.

Rädle, B., Rutkowski, A. J., Ruzsics, Z., Friedel, C. C., Koszinowski, U. H., Dölken, L. (2013). **Metabolic labeling of newly transcribed RNA for high resolution gene expression profiling of RNA synthesis, processing and decay in cell culture.** J Vis Exp. 2013 Aug 8; (78).

Rey, S., & Semenza, G. L. (2010). **Hypoxia-inducible factor-1-dependent mechanisms of vascularization and vascular remodelling**. Cardiovascular research, 86(2),   236–42.

Rytkönene, K. T., Williams, T. A., Renshaw, G. M., Primmer, C. R., Nikinmaa, M. (2011). **Molecular evolution of the metazoan PHD-HIF oxygen-sensing system.** Mol Biol Evol. 2011 Jun;28(6):1913-26.

Sancho, E., Batle, E., Clevers, H. (2003). **Live and let die in the intestinal epithelium**. Curr Opin Cell Biol.; 15(6):763-70.

Semenza, G. L (2004). **Hydroxylation of HIF-1 for cancer therapy.** Nat Rev Cancer 3 (10):721-732

Semenza, G. L. (2012). **Hypoxia-inducible factors: mediators of cancer progression and targets for cancer therapy.** Trends in pharmacological sciences, 33(4), 207–14.

Semenza, G. L. and Wang, G. L. (1992). **A nuclear factor induced by hypoxia via de novo protein synthesis binds to the human erythropoietin gene enhancer at a site requiered for transcriptional activation**. Mol Cell Biol 12(12): 5447-5454

Seta K, Kim HW, Ferguson T, Kim R., Pathrose P., Yuan Y., Lu G., Spicer Z., Millhorn D. E. (2002). **Genomic and physiological analysis of oxygen sensitivity and hypoxia tolerance in PC12 cells**. Ann N Y Acad Sci. 2002 Oct;971:379-88. Review.

Simons, K., Toomre, D. (2000). **Lipid raft and signal transduction**. Nat Rev Mol Cell Biol. 2000 Oct;1(1):31-9. Review. Erratum in: Nat Rev Mol Cell Biol 2001 Mar;2(3):216

Sims, D., Bursteinas, B., Gao, Q., Jain, E., MacKay, A., Mitsopoulos, C., & Zvelebil, M. (2010). **ROCK: a breast cancer functional genomics resource.** Breast cancer   research and treatment, 124(2), 567–72.

Stein, E., Savaskan, N. E., Ninnemann, O., Nitsch, R., Zhou, R., Skutella, T. (1999). **A role for the Eph ligand ephrin-A3 in entorhino-hippocampal axon targeting.** J Neurosci. 1999 Oct 15;19(20):8885-93.

Stevenson, T. J., Trinh, T., Kogelschatz, C., Fujimoto, E., Lush, M. E., Piotrowski, T., Brimley, C. J., et al. (2012). **Hypoxia disruption of vertebrate CNS pathfinding through ephrinB2 Is rescued by magnesium**. PLoS genetics, 8(4), e1002638.

Taft, R. J., Pang, K. C., Mercer, T. R., Dinger, M., Mattick, J. S. (2010). **Non-coding RNAs: regulators of disease.** J Pathol. 2010 Jan;220(2):126-39.

Tiana, M., Villar, D., Pérez-Guijarro, E., Gómez-Maldonado, L., Moltó, E., Fernández- Miñán, A., Gómez-Skarmeta, J. L., et al. (2011). **A role for insulator elements in the regulation of gene expression response to hypoxia**. Nucleic acids research, 1–12.

Tufarelli, C., Stanley, J. A., Garrick, D., Sharpe, J. A., Ayyub, H., Wood, W. G., Higgs, D. R. (2003). **Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease.** Nat Genet. 2003 Jun;34(2):157-65.

Tun, H. W., Marlow, L. a, Von Roemeling, C. a, Cooper, S. J., Kreinest, P., Wu, K., Luxon, B. a, et al. (2010). **Pathway signature and cellular differentiation in clear cell renal   cell carcinoma**. PloS one, 5(5), e10696.

Van de Vijver, M. J., He, Y. D., van't Veer, L. J., Dai, H., Hart, A. A. M., Voskuil, D. W., Schreiber, G. J., et al. (2002). **A gene-expression signature as a predictor of survival in breast cancer**. The New England journal of medicine, 347(25), 1999–2009.

Van Werven, F. J., Neuert, G., Hendrick, N., Lardenois, A., Buratowski, S., Van Oudenaarden, A., Primig, M., Amon, A. (2012). **Transcription of two long noncoding RNAs mediates mating-type control of gametogenesis in budding yeast.** Cell. 2012 Sep 14;150(6):1170-81.

Vaupel, P., & Mayer, A. (2007). **Hypoxia in cancer: significance and impact on clinical outcome.** Cancer metastasis reviews, 26(2), 225–39.

Vihanto, M. M., Plock, J., Erni, D., Frey, B. M., Frey, F. J., & Huynh-Do, U. (2005). **Hypoxia up-regulates expression of Eph receptors and ephrins in mouse skin.** FASEB journal : official publication of the Federation of American Societies for Experimental Biology, 19(12), 1689–91.

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**. Proceedings of the National Academy of Sciences of the United States of America, 92(12), 510–4. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=41725&tool=pmcentrez&rendertype=abstract

Webster, K. A. (2007). **Hypoxia: life on the edge**. Antioxid Redox Signal 9 (9):1303-1307.

Wenger, R. H., Stiehl, D. P., Camenisch, G., (2005). **Integration of oxygen signaling at the consensus HRE**. Sci STKE. 18;2005(306):re12. Review.

Wheaton, W. W., & Chandel, N. S. (2011). Hypoxia. 2. **Hypoxia regulates cellular metabolism**. American journal of physiology. Cell physiology, 300(3), C385–93.

Xiao, F., Qiu, H., Zhou, L., Shen, X., Yang, L., & Ding, K. (2013). WSS25 inhibits **Dicer, downregulating microRNA-210, which targets Ephrin-A3, to suppress human microvascular endothelial cell (HMEC-1) tube formation**. Glycobiology, 23(5),524–35.

Wilusz, J. E., Sunwoo, H., Spector, D. L. (2009). **Long noncoding RNAs: functional surprises from the RNA world.** Genes Dev. 2009 Jul 1;23(13):1494-504.

Wu, J., Luo, H. (2005). **Recent advances on T-cell regulation by receptor tyrosine kinases.** Curr Opin Hematol.;12(4):292-7. Review.

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 USA. 2009 Mar 17;106(11):4260-5.

Xiao, F., Qiu, H., Zhou, L., Shen, X., Yang, L., Ding, K. (2013). **WSS25 inhibits Dicer, downregulating microRNA-210, which targets Ephrin-A3, to suppress human microvascular endothelial cell (HMEC-1) tube formation.** Glycobiology. 2013 May;23(5):524-35.

Yamanda, Y., Midorikawa, T., Oura, H., Yoshino, T., Ohdera, M., Kubo, Y., Arase, S. (2008). **Ephrin-A3 not only increases the density of hair follicles but also accelerates anagen development in neonatal mice**. J Dermatol Sci. 2008 Dec;52(3):178-85.

Yamashita, T., Ohneda, K., Nagano, M., Miyoshi, C., Kaneko, N., Miwa, Y., Yamamoto, M., et al. (2008). **Hypoxia-inducible transcription factor-2alpha in endothelial cells regulates**

tumor neovascularization through activation of ephrin A1. The Journal of biological chemistry, 283(27), 18926–36.

Yoon, J. H., Abdelmohsen, K., & Gorospe, M. (2012). **Posttranscriptional Gene Regulation by Long Noncoding RNA.** Journal of molecular biology, 2.

Zhang, H., Wong, C. C. L., 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., Sulumar, S., Semenza, G. L.,(2011). **HIF-1-dependent expression of angiopoietin-like 4 and L1CAM mediates vascular metastasis of hypoxic breast cancer cells to the lungs**. Oncogene, 1–14.

Zhang, Y., Fei, M., Xue, G., Zhou, Q., Jia, Y., Li, L., Xin, H., Sun, S. (2012). **Elevated levels of hypoxia-inducible microRNA-210 in pre-eclampsia: new insights into molecular mechanisms for the disease.** J Cell Mol Med. 2012 Feb;16(2):249-59.

Zimmer, M., Palmer, A., Köhler, J., Klein, R. (2003). **EphB-ephrinB bi-directional endocytosis terminates adhesion allowing contact mediated repulsion**. Nat Cell Biol.;5(10):869-78. Epub 2003 Sep 14.

PUBLICATIONS

Hypoxia regulates pro-metastatic Ephrin-A3 by the induction of EFNA3 long-non coding RNAs

Laura Gómez-Maldonado[1,8], María Tiana[1,8], Olga Roche[1,6],Alfonso Prado-Cabrero[1,7], Asunción Fernandez-Barral[1], Irene Guijarro-Muñoz[2], Elena Favaro[3], Gema Moreno-Bueno[1], Laura Sanz[2], Julian Aragones[4], Adrian Harris[3], Olga Volpert[5], Benilde Jimenez[1] and Luis del Peso[1,6,9].

1 Department of Biochemistry, Universidad Autónoma de Madrid and Instituto de Investigaciones Biomédicas Alberto Sols, CSIC-UAM. Madrid 28029. Spain

2 Molecular Immunology Unit, Hospital Universitario Puerta de Hierro Majadahonda, Joaquin Rodrigo 2, 28222, Madrid, Spain

3 Molecular Oncology Laboratories, Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DS, UK.

4 Research Unit, Hospital Universitario Santa Cristina, Research Institute Princesa, Autonomous University of Madrid, 28009 Madrid, Spain.

5 Urology Department, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA.

6 IdiPaz, Instituto de Investigación Sanitaria del Hospital Universitario La Paz, Madrid 28029, Spain.

7 present address King Abdullah University of Science and Technology, Thuwal, Saudi Arabia.

8 equal contribution

9 Correspondence to: Luis del Peso, Department of Biochemistry-Instituto de Investigaciones Biomédicas Alberto Sols. Arturo Duperier, 4. Madrid 28029. Spain. Telf: +34 (91) 5854440. e-amil: luis.peso@uam.es; lpeso@iib.uam.es

Running title: Regulation of Ephrin-A3 by hypoxia. Role in cancer

Conflict of interest: The authors declare no competing financial interests

1

ABSTRACT

The presence of hypoxic regions in solid tumors is an adverse prognostic factor for patient outcome. Here, we show that hypoxia induces the expression of Ephrin-A3 through a novel Hypoxia Inducible Factor (HIF)-mediated mechanism. In response to hypoxia, the coding *EFNA3* mRNA levels remained relatively stable, but HIFs drove the expression of previously unknown long non-coding (lnc) RNAs from *EFNA3* locus and these lncRNA caused Ephrin-A3 protein accumulation. Ephrins are cell surface proteins that regulate diverse biological processes by modulating cellular adhesion and repulsion. Mounting evidence implicates deregulated ephrin function in multiple aspects of tumor biology. We demonstrate that Ephrin-A3 expression increased the metastatic potential of human breast cancer cells in an xenotransplantation model, likely by increasing the ability of tumor cells to extravasate from the blood vessels into surrounding tissue. In agreement, we found a strong correlation between high *EFNA3* expression and shorter metastasis-free survival in breast cancer patients. Together, our results suggest that hypoxia could contribute to metastatic spread of breast cancer via HIF-mediated induction of *EFNA3* lncRNAs and subsequent Ephrin-A3 protein accumulation.

Keywords: Ephrin, HIF, hypoxia, lncRNA, metastases,

2

INTRODUCTION

In a variety of pathological situations, oxygen demand exceeds it supply to affected tissues leading to a condition known as hypoxia. In particular, hypoxia is a frequently observed in solid tumors and, importantly, it has been suggested to be an adverse prognostic factor for patient outcome (Bertout, Patel, & Simon, 2008; Vaupel & Mayer, 2007) .

At the cellular level, oxygen homeostasis is largely dependent on the induction of a specific gene expression program under the control of the Hypoxia Inducible Factors (HIFs). HIFs are heterodimers composed of alpha and beta subunits that belong to the basic helix-loop-helix (bHLH) Per-ARNT-Sim (PAS) family (Wang, Jiang, Rue, & Semenza, 1995) . In mammals, three genes encode HIFα subunits, *HIF1A*, *EPAS1/HIF2A* and *HIF3A*, of which HIF-1α and *EPAS1* are the best characterized (Kaelin & Ratcliffe, 2008) . Oxygen regulates the stability (Epstein et al., 2001; Salceda & Caro, 1997) and transcriptional activity (Hewitson et al., 2002; Jiang, Zheng, Leung, Roe, & Semenza, 1997; Lando, Peet, Whelan, Gorman, & Whitelaw, 2002) of HIFα without affecting HIFβ function. The induction of the HIF transcriptional program results in cellular adaptation to hypoxia, a response that aims to restore oxygen supply to hypoxic regions through the induction of erythropoiesis and angiogenesis and to reduce oxygen consumption by reprogramming cellular metabolism.

HIF's role in cancer has been extensively investigated. On one hand, as hypoxia develops in tumors as a consequence of the rapid expansion of the transformed cells, HIF is activated contributing to critical aspects of tumor progression (Bertout et al., 2008; Semenza, 2003) . On the other hand, VHL tumor suppressor, which plays a key role in the control of HIFα stability (Ivan et al., 2001; Jaakkola et al., 2001; Maxwell et al., 1999) , is frequently lost in clear cell renal cell carcinomas (ccRCC). Moreover, several lines of evidence point to HIFs, in particular *EPAS1*, as a ethiological factor in this kind of cancer (Kaelin, 2007) . Other oncogenes and tumor suppressor genes including PTEN, mTORC1, Ras, Akt and p53 also regulate of HIF activity (Semenza, 2012) . Importantly, the correlation between

tumor hypoxia and/or HIFα expression with poor prognosis and increased risk of metastasis has been repeatedly demonstrated in diverse tumor types (Vaupel & Mayer, 2007) . Non-surprisingly, many of the HIF target genes are involved in biological processes that impact the metastatic spread of cancer cells, such as angiogenesis, epithelial-mesenchymal transition, cell motility, intra/extravasation and control of the pre-metastatic niche (Lu & Kang, 2010) .

Ephrins are a large family of cell surface ligands that mediate intercellular adhesion and repulsion through interaction with a large group of receptor tyrosine kinases, the Eph receptors (Nievergall, Lackmann, & Janes, 2012; Pasquale, 2010) . Ephrins play essential roles during development where they guide migration and positioning of the cells for proper tissue patterning. Their function has been particularly well characterized in the nervous system, where Ephrins/Eph function as axon guidance molecules, and in cardiovascular system, where they control vasculogenesis and angiogenesis (Nievergall et al., 2012) . Ephrins are separated in two families according to their structure. Type A Ephrins (ephrin-A1 to -A5) are glycosylphosphatidyl-inositol (GPI)-linked membrane bound ligands, whereas type B ephrins (ephrin-B1 to -B3) contain a single transmembrane domain and a short cytoplasmic tail (Pasquale, 2008) . Interestingly, many of the biological functions of ephrins and Ephs are co-opted by transformed cells and contribute to tumor progression (Chen, 2012a) . Accordingly, many ephrin family members are altered in human cancers and their expression often correlates with a more aggressive phenotype, invasion, metastasis and poor prognosis (Surawska, Ma, & Salgia, 2004) . In addition to their role in adhesion and repulsion, Eph/Ephrins may contribute tumor progression and metastasis by altering angiogenesis (Surawska et al., 2004) . Intriguingly, HIF-dependent regulation of Ephrin-A1, Ephrin-B2, EphA2 and Eph4 has been described in a mouse model of skin hypoxia (Vihanto et al., 2005) , suggesting a potential link between intratumoral hypoxia, ephrin/Eph expression and tumor progression. In agreement, Ephrin-A1 is upregulated at transcription level, via *EPAS1*, in hypoxic tumors and this upregulation contributes to increased tumor vascularization (Ogawa

4

et al., 2000; Yamashita et al., 2008) . Altogether, these works suggest a link between tumoral hypoxia and progression to aggressive phenotype via HIF-mediated regulation of Ephrin expression.

In this study we show that hypoxia leads to Ephrin-A3 protein accumulation via a previously unknown mechanism that involves HIF-mediated transcriptional upregulation of a novel group of lncRNAs encoded by the *EFNA3* locus. Using animal models and *in vitro* assays, we demonstrate that Ephrin-A3 expression leads to metastatic spread. In contrast with previous studies we observed no effect on vascularization, but a strong repulsive action, which leads to increased intra- and extravasation that underlies the promotion of metastatic spread by *EFNA3*. Finally, the analysis of public gene expression profiling datasets revealed that *EFNA3* is induced by HIF in human tumors and this induction is predictive of poor prognosis and increased risk of metastasis.

RESULTS

***EFNA3* locus encodes long non-coding RNAs (lncRNAs) that are regulated by hypoxia**

We first identified *EFNA3*, a member of the ephrin type A ligands, as potential novel HIF target gene using an *in silico* search (Ortiz-Barahona, Villar, Pescador, Amigo, & del Peso, 2010) . As a first step to validate this prediction, we determined the level of Ephrin-A3 protein and found that it was induced by hypoxia in several cell lines (Figure 1A). The Ephrin-A3 species we detected had an apparent molecular weight of approximately 72kD, much higher than the 26.3kD molecular weight predicted from its amino acid sequence, due to postranslational modifications (supplementary Figure 1). Next we decided to study the effect of hypoxia on *EFNA3* mRNA. According to curated databases (RefSeq) the *EFNA3* locus encodes for a single mRNA isoform (NM_004952). However, inspection of the publicly available information, including ESTs and experimentally identified transcription start sites (http://dbtss.hgc.jp/), suggested the existence of additional mRNAs transcribed from this locus. To identify the hypothetical RNAs isoforms encoded by the *EFNA3* locus, we performed 5'-RACE experiments in HeLa and LoVo cell lines (Supplementary Figure 2) and found two transcription start sites (TSS) in addition to that of the NM_004952 mRNA (Figure 1B). The existence and location of these additional transcription start sites was supported by 5' cap analysis gene expression (CAGE) tags from multiple cell lines produced as part of the ENCODE transcriptome Project (Supplementary figure 2). Interestingly, none of these novel RNAs seemed to encode for functional proteins and, accordingly, we termed them non-coding-1 (NC1) and NC2, based on their different TSS (Figure 1B). Specifically, NC1 isoform contains an ATG codon within its first 10 nucleotides that is in frame with the open reading frames (ORF) of NM_004952, suggesting that NC1 could encode a truncated form of Ephrin A3. However, this potential product would have an interrupted structural domain (the ephrin domain), that makes up the bulk of the protein, which would likely undermine its stability. On the other hand, the

6

NC2 sequence does not contain any ORFs of significant length. In agreement, *in vitro* transcription-coupled translation of *EFNA3* and NC1 cDNAs, produced the proteins of expected sizes, whereas NC2 cDNA generated no apparent protein product (Figure 1C). To test the coding potential of NC1 and NC2 in intact cells, we transfected HeLa cells with the respective plasmids along with the canonical *EFNA3* cDNA (*EFNA3*) (Figure 1D). Neither NC1 nor NC2 caused the expression of exogenous protein that could be recognized by the monoclonal or polyclonal antibody against *EFNA3*. Thus NC1 and NC2 can be considered novel long non-coding RNAs (lncRNAs). In addition to the novel TSS, a method using support vector machine for poly(A) site prediction (Cheng, Miura, & Tian, 2006) identifies two potential 3'-ends for this locus (Supplementary Figure 2, "Poly(A) Sites" track). In agreement, analysis of Paired-end diTag (PET) Sequencing data from ENCODE (Ng et al., 2005) indicates that all combinations of the TSSs and 3'-ends are present in cells (Supplementary figure 2, "GIS-PET" track). Thus, both NC1 and NC2, can be expressed as shorter forms with a truncated 3'UTR, which we termed NC1s and NC2s respectively (Figure 1B).

We next investigated the regulation by hypoxia of the different transcripts encoded by the *EFNA3* locus. First, we used commercially available TaqMan probes to amplify the regions of the *EFNA3* gene specific to the canonical *EFNA3* mRNA (TaqMan1+2, Figure 1E) and a region common to all the RNA isoforms encoded by this locus (TaqMan4+5, Figure 1E). qPCR results indicate that the absolute expression levels and relative induction in response to hypoxia varied widely between RNA species (Figure 1E). The expression of the canonical coding isoform was low, compared to the combined expression level for all isoforms, suggesting that under normoxic conditions the transcription of the long non-coding RNAs predominates. This result was confirmed by the ENCODE genome-wide transcription analysis (Supplementary Figure 2, "Transcription" track). Strikingly, the canonical *EFNA3* isoform was barely induced in response to hypoxia, in stark contrast with the robust upregulation of the bulk of *EFNA3* RNAs (Figure 1E). This result suggested that the lncRNAs, but not

7

the canonical mRNA, were regulated by hypoxia. To confirm this possibility, we designed primer pairs for all the exons in the gene to determine the response to hypoxia of every potential RNA isoform. Indeed, hypoxia strongly upregulated the novel lncRNA isoforms, while the regulation of the canonical protein-coding mRNA was only marginal (figure 1F). Importantly, the induction of *EFNA3* transcripts in response to hypoxia was mediated by HIF as it was blocked by siRNA directed against HIF1α (Figure 1G, Supplemntary Figure 3). To investigate the regulation of *EFNA3* by HIF *in vivo*, we employed conditional VHL knockout mouse lines (Miró-Murillo et al., 2011) . VHL deletion results in constitutive HIF activity and, consistently with the in vitro results, this led to increased *EFNA3* lncRNA expression without significantly altering the level of the coding *EFNA3* mRNA (Figure 1H). Importantly, the induction of *EFNA3* upon VHL loss was partially prevented in animals lacking both VHL and EPAS (HIF2α) alleles (Figure 1H), suggesting that the *in vivo* HIF2α mediates the effect of VHL, at least in the liver and the lung. Regardless the specific HIF isoform involved, that could just reflect their differential tissue expression, it is clear from these set of results that *EFNA3* expression is induced, both *in vitro* and *in vivo*, in response to hypoxia, in a HIF-dependent manner.

Mounting evidence indicates that lncRNAs are key regulators of gene expression that affect the mRNA transcription rate, stability and translation (Nie et al., 2012; Yoon, Abdelmohsen, & Gorospe, 2012) . Thus, in an attempt to reconcile the induction of Ephrin-A3 protein with the regulation of coding and non-coding transcripts by hypoxia, we tested whether NC1 and NC2 lncRNAs could affect *EFNA3* mRNA or protein levels. As shown in Figure 2A, the overexpression of the lncRNAs had no significant effect on *EFNA3* mRNA levels. However, exogenous expression of NC1/2, particularly the short isoforms (NC1s and NC2s), caused *EFNA3* protein accumulation (Figure 2B, 2C and data not shown). These results provide an explanation for the induction of *EFNA3* protein under hypoxia in spite of its modest effect on *EFNA3* mRNA level.

HIF binding to an intragenic regulatory region induces the transcription of *EFNA3* lncRNAs and, indirectly, regulates Ephrin-A3 expression postranscriptionally

The different transcription start sites of the *EFNA3* isoforms suggested the existence of alternative promoters. In agreement, histone modifications landscape, location of the open chromatin and transcription factor binding site (TFBS) regions in the *EFNA3* locus (Figure 3A) were consistent with the existence of two promoter regions (Figure 3A, Prm1and Prm2). To further investigate the regulation of the different *EFNA3* transcripts, we analyzed the binding of RNA polymerase II (Pol II) to these putative promoter regions. As controls, we also included the promoter regions of a bona fide hypoxia inducible gene (P4HA) and a non-responsive gene STT3A. ChIP-qPCR analysis showed that Pol II bound to both regions of the *EFNA3* gene locus under normoxia (Figure 2B). Consistent with the relatively low *EFNA3* expression in HeLa cells, the binding of Pol II to the *EFNA3* promoter regions was lower than its binding to the P4HA and STT3A promoters. Importantly, Pol II binding to the Prm2 region was strongly induced by hypoxia (Figure 3B), to a level comparable to the HIF-responsive promoter of P4HA. In contrast, the Pol II binding to the Prm1 region and to the STT3A promoter remained unaffected by hypoxia (Figure 3B).

Next, we investigated direct binding of HIF1α to the *EFNA3* locus by ChIP-qPCR. We designed six primer sets (Figure 3A, H1-H6) to sample most of the RCGTG motifs (Figure 3A, "RCGTG motifs" track) within *EFNA3* locus and found that the only fragment that showed, significant HIF binding under hypoxic conditions was H6 (Figure 3C). Interestingly, H6 region is close proximity to Prm2, which could explain the differential response of promoter regions 1 and 2 to hypoxia.

Finally, we cloned both *EFNA3* promoter regions and studied their response to hypoxia and to the chemical inhibitor of HIF prolyl hydroxylase, DMOG, which causes acute HIF activation. Although both regions showed similar basal promoter activity (data not shown), Prm2 was strongly induced by hypoxia and DMOG, whereas Prm1 remained unaffected (Figure 3D). Importantly, Prm2 response to

9

hypoxia was critically dependent on one of its RCGTG motifs whose mutation completely abrogated hypoxic induction (figure 3D).

HIF activation correlates with *EFNA3* expression in human tumors

Over 80% of human clear cell renal cell carcinomas (ccRCC) samples are deficient for VHL function and, as a consequence, present constitutive HIF activity even in the presence of oxygen (Kaelin, 2007) . Thus, ccRCC is highly suitable to study a putative link between HIF and *EFNA3* in human tumors. In agreement, the expression of EGLN3, a well-characterized direct target of HIF, was clearly increased in the three independent ccRCC tumor series (Figure 4A). Similarly, *EFNA3* was significantly increased in ccRCC tumor cells as compared with the normal kidney tissue in the same series, regardless of the microarray platform used to assay tumor samples (Figure 4B). The probes used to determine *EFNA3* expression in these datasets bind to regions common to all transcripts from this locus (see materials and methods) and thus they detect the cumulative signal by the mRNA and lncRNA transcripts. However, since the expression of lncRNAs was much higher than that of *EFNA3* mRNA in all cell types and tissues tested so far (Figure1, Supplementary Figure 2 and data not shown), we assume that the microarray signal is largely generated by the lncRNA. We also examined the expression of two other EFNA family members, EFNA1 and EFNA4, whose coding genes flank the *EFNA3* locus in mammalian genomes. Interestingly, EFNA1, but not EFNA4, is induced by hypoxia ((Yamashita et al., 2008) , and data not shown). Concordantly, the expression of EFNA1, but not that of EFNA4, is clearly increased in ccRCC (Figure 4). These results rule out that the increased *EFNA3* expression observed in these tumors could be caused by a gross structural alteration or transcriptional deregulation of the genomic region containing the *EFNA3* gene and suggest it is due to increased HIF activity as it correlates with the changes observed for other hypoxia regulated genes.

***EFNA3* promotes metastatic behavior by enhancing extravasation of tumor cells**

Since hypoxia is a common condition in solid tumors that correlates with metastatic potential (Vaupel & Mayer, 2007) and some ephrin family members have been implicated in the promotion of metastatic behavior (Surawska et al., 2004), we next studied the metastatic potential of MDA-MB-231 cells engineered to express luciferase and *EFNA3* in an orthotopic xenotransplantation model. At five weeks after tumor inoculation, the total photon flux was significantly higher in most animals bearing *EFNA3*-positive tumors, compared to controls (Figure 5A and 5B). Importantly, the luminiscence signal in *EFNA3* animals was not restricted to the original injection site but diffused throughout the abdominal area, suggesting metastatic spread rather than increased primary tumor growth. In agreement, the volumes of the primary tumors did not differ significantly between groups at the end of the experiment (Figure 5C) and necropsies confirmed that 7 out of 8 of the animals injected with *EFNA3*-positive cells presented with metastases as opposed to 1 out of 7 in control group (Figure 5D). Most metastases were located in the peritoneal cavity, attached to the surface of internal organs (Figure 5E). The increased metastatic potential of the expressing cells was confirmed in another model where the same cells were injected via the tail vein to generate experimental lung metastases (Supplementary figure 4).

Importantly, the microvascular density, as determined by analysis of CD31 staining, was similar between the *EFNA3*-positive and negative tumors (Figure 5F). Thus, the increased metastatic potential of *EFNA3*-expressing cells was not due to increased vascular density. Ephrins, through their receptors, mediate attraction and repulsion signals between cells (Nievergall et al., 2012). This function could enable tumor cells entry and exit from blood vessels and in this way increase their metastatic potential. To test this possibility, we plated HEK293T cells, transiently transfected with plasmids encoding *EFNA3*, on top of the HUVEC cells grown as a cell monolayer and measured the ability of HEK-*EFNA3* cells to pass through HUVEC monolayer and attach to the plastic surface below. As shown in

11

Figures 6A and 6B, the number of GFP-positive cells attached and spread on the plate was significantly higher when *EFNA3* was expressed. This result is consistent with *EFNA3* transducing a repulsive signal to HUVECs that allowed HEK293 cells to transmigrate through the monolayer. In agreement, recombinant Ephrin-A3-Fc blocked the directional migration of HUVEC cells induced by the conditioned media from NIH 3T3 cells (Figure 6C).

These data suggest a mechanism for the increased metastatic potential of the tumors cells expressing *EFNA3* whereby they can efficiently intravasate and extravasate through the vascular wall. To investigate this possibility *in vivo*, we injected MDA-MB-231 cells in the tail vein of immune defficient mice and assessed the appearance of the individual tumoral cells in the lungs in a short-term assay. Although the numbers of circulating tumor cells two hours after injection, were similar in the control and experimental groups, seven days later the GFP signal was much higher in the animals injected with *EFNA3*-expressing cells (Figures 6D and 6E). Therefore, *EFNA3* facilitates metastasis, at least in part, by allowing the extravasation of the tumor cells from the vasculature.

In view of these results, we next analyzed the correlation between *EFNA3* expression and the risk of metastases in breast cancer patients. To this end, we made use of the public access datasets of human breast cancer gene expression and associated clinical data from the ROCK website (Online Breast Cancer Knowledgebase, http://www.rock.icr.ac.uk/) (Sims et al., 2010) . We categorized patient samples according to *EFNA3* expression level (high vs. low) and analyzed the incidence of metastasis over time in both groups (Figure 7). In agreement with the results of the animal studies, the risk of metastatic disease was significantly higher in patients whose primary tumors expressed higher levels of *EFNA3* (Figure 7).

12

DISCUSSION

The majority of known HIF targets are regulated by direct induction of coding mRNA. In this study, using Ephrin A3 as an example, we demonstrate a completely novel way of hypoxic regulation of gene expression whereby the changes in the coding mRNA are only marginal. Instead, HIF directly regulates the levels of long non-coding (lnc) RNAs, which, in turn, increase protein levels. Further work is needed to determine whether this effect is caused by increased translation or reduced protein turnover, however one recent report describes lncRNA that affects Uchl1 protein levels by enhancing mRNA translation of the coding mRNA (Carrieri et al., 2012) . In fact, lncRNAs have been shown to regulate virtually every step of the gene expression cascade from transcription to translation (Nie et al., 2012; Yoon et al., 2012) . Interestingly, some non-coding transcripts act in concert to regulate the expression of coding mRNAs. For example some lncRNAs act as decoys for the miRNA that target a particular mRNA (Yoon et al., 2012) . It is known that miR-210, which is induced by hypoxia (Huang et al., 2009; Kulshreshtha et al., 2007) , prevents the translation of several mRNA including ISCU and *EFNA3* (Chan et al., 2009; Fasanaro et al., 2008; Favaro et al., 2010; Huang, Le, & Giaccia, 2010). Since miR-210 binds the 3'-UTR of the *EFNA3* mRNA (Fasanaro et al., 2008) , it is feasible that *EFNA3* lncRNAs increase *EFNA3* mRNA translation by depleting miR-210. However, the fact that shorter isoforms, NC1s and NC2s are more efficient in the induction of Ephrin-A3 than the long ones (NC1 and NC2), argues to the contrary because the miR-210 binding sites are not represented on the shorter lncRNAs isoforms, NC1s and NC2s Supplementary Figure 2). In agreement, we found that expression of truncated forms of NC1 and NC2, lacking all the sequences corresponding to the 3'-UTR, were also able to increase Ephrin-A3 protein levels (data not shown). Our study and reports by other groups (Fasanaro et al., 2008; Larsen, Muz, Khong, Feldmann, & Paleolog, 2012) suggest that *EFNA3* response to hypoxia is tightly controlled.

Regarding the mechanism responsible for the selective transcription of specific *EFNA3* isoforms, we

13

found that it is mediated by discriminatory recruitment of HIF and RNA pol II to the Prm2 promoter region. Interestingly, the presence of CTCF binding sites flanking *EFNA3* Prm2 region (Supplementary Figure 2) offers a plausible explanation for the preferential transcriptional activation of the lncRNA by HIF and the lack of the regulation of the canonical coding isoform. Other studies also demonstrate the role of insulators in determination of HIF specificity for promoters located in the close proximity (Tiana et al., 2011) . Also regarding the molecular mechanism, we found that *in vitro EFNA3* induction was dependent on HIF-1α and not affected by HIF-2α knockdown whereas in animal models, *EPAS1* was required for *EFNA3*A induction upon VHL loss. Thus, HIF isoforms responsible for *EFNA3* may be context-dependent and could be dictated by availability of a specific HIF isoform, or another, HIF cell-type specific transcription factor(s), that favor interaction with a specific HIF isoform. Nonetheless, this has been observed for other hypoxia-regulated genes. For instance, the hypoxic regulation of another ephrin, Ephrin-A1, also requires HIF1 α (Vihanto et al., 2005) in cell cultures, but not *in vivo*, where it is dependent on HIF-2α (Yamashita et al., 2008) .

An important question arising from our results relates to the role of Ephrin-A3 in the physiological response to hypoxia. The induction of angiogenesis is one of the classical responses to hypoxia and other ephrins play key roles in angio- and vasculogenesis. However, we did not find a significant effect of Ephrin-A3 expression on tumor vascular density as determined by CD31 staining. Although we cannot rule out effects on other vascular parameters like maturation or normalization of tumor vasculature, our results demonstrate a direct role of Ephrin-A3 on tumor cell intra-/extravasation that could underlie the pro-metastatic action of this ephrin. Earlier studies showed that Ephrin-A3 overexpression prevents the *in vitro* formation of capillary-like structures by HUVE cells (Fasanaro et al., 2008; Xiao et al., 2013) . We similarly found that recombinant Ephrin-A3-Fc prevents the migration of HUVECs *in vitro*, however these *in vitro* effects did not manifest in the *in vivo* models.

Regardless of their physiological role, ephrins are frequently co-opted by the transformed cells

14

whereby they contribute to the changes in cell-cell and cell-matrix attachment and survival during invasion, angiogenesis and metastasis (Chen, 2012b; Nievergall et al., 2012) . In agreement, we found that the expression of Ephrin-A3 in MDA-MB-231 breast cancer cells increased their metastatic potential. The significant correlation that we found between high *EFNA3* levels and metastasis in clinical samples strongly supports this conclusion. Also in agreement with our findings, a recent study demonstrate the differential expression of Eprin-A3 between the highly metastatic and poorly metastatic colon cancer cells (Barderas et al., 2013) . Although we have not formally demonstrated it, it is tempting to speculate that the correlation of hypoxia with metastasis risk observed in human tumors is, at least in part, explained by the induction of Ephrin-A3 downstream of HIF activation due to the tumor hypoxia or oncogenic alterations. As for the mechanism, our results suggest that Ephrin-A3 promotes metastasis by repelling vascular endothelium and therefore creating points of intra- and extravasation for the tumor cells. Similar events are caused by angiopoietin-like 4 (ANGPTL4), another hypoxia-inducible gene, to promote metastasis (Zhang et al., 2011) .

In summary, we demonstrate a novel mechanism of hypoxic regulation, where *EFNA3* expression is driven by previously non-annotated intragenic lncRNAs whose transcription is controlled by HIF. We propose that in hypoxic regions of solid tumors or in tumors harboring genetic alterations leading to constitutively active HIF, this novel mechanism leads to Ephrin-A3 accumulation and the promotion of metastasis.

MATERIALS AND METHODS

Cell culture

Cells were grown in standard culture conditions. For hypoxia treatments, cells were placed in a 1% O2, 5% CO2, 94% N2 gas mixture in a Whitley hypoxystation (don Whitley Scientific, UK). Dimethyloxalylglycine (DMOG, Frontier Scientific, CA, USA) was added where indicated at a 500 μM final concentration. MDA-MB231-EphrinA3-Luc and MDA-MB231-Luc cells were generated by infection with pLOC-*EFNA3* or pLOC lentivirus respectively and then reinfected with Lenti-Fire Luciferase Luc2 (CellCyto) lentivirus.

ChIP (chromatin immunoprecipitation)

ChIP was performed as previously described (Villar et al., 2012)　 with minor modifications. Briefly, EZ Chromatin Inmunoprecipitation Kit (Millipore) was used following the manufacturer's instructions. HeLa cells were grown to 85% confluence on 10 cm plates before they were exposed to hypoxia or left in normoxic conditions. Chromatin was immunoprecipitated with antibodies against RNA Polymerase II (Abcam 5408) or Hif1a (Abcam, ab2185). Mouse IgG (Sigma M8695) or whole rabbit serum were used as negative controls. DNA was purified by phenol–chloroform extraction and ethanol precipitation and bound regions were identified by qPCR with indicated primer pairs (supplementary Table I).

Repulsion assay

HUVE cells were stained with 10μmoles/10^6 cells of a membrane labelling reagent PKH26 (Sigma), seeded into a 24 multi-well plate (1X10　 cells/well) and allowed spread to form a cell monolayer overnight. Then HEK-293T cells, expressing GFP alone or in combination with Eprhrin-A3, were plated on top of the HUVEC (3000 cells/well) and mixed cell culture followed over time.

16

Orthotopic tumor implantation and extravasation assays.

Tumor studies were performed on 5–7 week-old female nude mice (Harlan) according to the protocols in accordance with the NIH Guide for the Care and Use of Laboratory Animals. MDA-MB-231-Luc and MDA-MB-231-EphrinA3-Luc cells were harvested by trypsinization, resuspended in PBS and injected into the mammary fat pad (10 per site, one site per mouse) of immune deficient mice in a 1:1 mix with Matrigel (BD Biosciences). The animals were subjected to weekly bioluminescence imaging in an IVIS Spectrum System (Caliper, Xenogen) to monitor tumor progression and metastasis. The mice were given intraperitoneal injections of D-luciferin (15 mg/ml in PBS), 5 min prior to imaging. Tumors, lungs, bones and livers were harvested after 9 weeks and processed for tissue analysis. Primary tumor volumes were calculated as length x width^2 x 0.52. For the extravasation assays cells were injected into the tail vein. Extravasation of cells was monitored by confocal images of whole fresh lungs.

Gene expression profile analysis

Gene expression data and relevant sample information were downloaded from the Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) and ROCK (http://www.rock.icr.ac.uk/) databases (free public access) and analyzed using custom scripts written in R language.

Statistical Analysis

Statistical analysis and graphic representations were performed in R, a language and environment for statistical computation and graphics (R Core Team, 2010) .

17

CONFLICT OF INTERESTS

Authors declare no conflict of interests.

REFERENCES

Barderas, R., Mendes, M., Torres, S., Bartolomé, R. A., López-Lucendo, M., Villar-Vázquez, R., … Casal, J. I. (2013). In-depth Characterization of the Secretome of Colorectal Cancer Metastatic Cells Identifies Key Proteins in Cell Adhesion, Migration, and Invasion. *Molecular & cellular proteomics : MCP*, *12*(6), 1602–20. doi:10.1074/mcp.M112.022848

Bertout, J. a, Patel, S. a, & Simon, M. C. (2008). The impact of O2 availability on human cancer. *Nature reviews. Cancer*, *8*(12), 967–75. doi:10.1038/nrc2540

Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., … Gustincich, S. (2012). Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature*, *491*(7424), 454–457. doi:10.1038/nature11508

Chan, S. Y., Zhang, Y.-Y., Hemann, C., Mahoney, C. E., Zweier, J. L., & Loscalzo, J. (2009). MicroRNA-210 controls mitochondrial metabolism during hypoxia by repressing the iron-sulfur cluster assembly proteins ISCU1/2. *Cell metabolism*, *10*(4), 273–84. doi:10.1016/j.cmet.2009.08.015

Chen, J. (2012a). *Regulation of tumor initiation and metastatic progression by Eph receptor tyrosine kinases. Advances in cancer research* (1st ed., Vol. 114, pp. 1–20). Elsevier Inc. doi:10.1016/B978-0-12-386503-8.00001-6

Chen, J. (2012b). *Regulation of tumor initiation and metastatic progression by Eph receptor tyrosine kinases. Advances in cancer research* (1st ed., Vol. 114, pp. 1–20). Elsevier Inc. doi:10.1016/B978-0-12-386503-8.00001-6

Cheng, Y., Miura, R. M., & Tian, B. (2006). Prediction of mRNA polyadenylation sites by support vector machine. *Bioinformatics (Oxford, England)*, *22*(19), 2320–5. doi:10.1093/bioinformatics/btl394

Epstein, A. C., Gleadle, J. M., McNeill, L. A., Hewitson, K. S., O'Rourke, J., Mole, D. R., … Ratcliffe, P. J. (2001). C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. *Cell*, *107*(1), 43–54.

Fasanaro, P., D'Alessandra, Y., Di Stefano, V., Melchionna, R., Romani, S., Pompilio, G., … Martelli, F. (2008). MicroRNA-210 modulates endothelial cell response to hypoxia and inhibits the receptor tyrosine kinase ligand Ephrin-A3. *The Journal of biological chemistry*, *283*(23), 15878–83. doi:10.1074/jbc.M800731200

Favaro, E., Ramachandran, A., McCormick, R., Gee, H., Blancher, C., Crosby, M., … 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*(4), e10345. doi:10.1371/journal.pone.0010345

Hewitson, K. S., McNeill, L. a, Riordan, M. V, Tian, Y.-M., Bullock, A. N., Welford, R. W., … Schofield, C. J. (2002). Hypoxia-inducible factor (HIF) asparagine hydroxylase is identical to factor inhibiting HIF (FIH) and is related to the cupin structural family. *The Journal of biological chemistry*, *277*(29), 26351–5. doi:10.1074/jbc.C200273200

Huang, X., Ding, L., Bennewith, K. L., Tong, R. T., Welford, S. M., Ang, K. K., … Giaccia, A. J. (2009). Hypoxia-inducible mir-210 regulates normoxic gene expression involved in tumor initiation. *Molecular cell*, *35*(6), 856–67. doi:10.1016/j.molcel.2009.09.006

19

Huang, X., Le, Q.-T., & Giaccia, A. J. (2010). MiR-210--micromanager of the hypoxia pathway. *Trends in molecular medicine*, *16*(5), 230–7. doi:10.1016/j.molmed.2010.03.004

Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., Ohh, M., … Kaelin, W. G. (2001). HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing. *Science (New York, N.Y.)*, *292*(5516), 464–8. doi:10.1126/science.1059817

Jaakkola, P., Mole, D. R., Tian, Y. M., Wilson, M. I., Gielbert, J., Gaskell, S. J., … Ratcliffe, P. J. (2001). Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation. *Science (New York, N.Y.)*, *292*(5516), 468–72. doi:10.1126/science.1059796

Jiang, B. H., Zheng, J. Z., Leung, S. W., Roe, R., & Semenza, G. L. (1997). Transactivation and inhibitory domains of hypoxia-inducible factor 1alpha. Modulation of transcriptional activity by oxygen tension. *The Journal of biological chemistry*, *272*(31), 19253–60.

Kaelin, W. G. (2007). Von Hippel-Lindau disease. *Annual review of pathology*, *2*, 145–73. doi:10.1146/annurev.pathol.2.010506.092049

Kaelin, W. G., & Ratcliffe, P. J. (2008). Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. *Molecular cell*, *30*(4), 393–402. doi:10.1016/j.molcel.2008.04.009

Kulshreshtha, R., Ferracin, M., Wojcik, S. E., Garzon, R., Alder, H., Agosto-Perez, F. J., … Ivan, M. (2007). A microRNA signature of hypoxia. *Molecular and cellular biology*, *27*(5), 1859–1867. doi:10.1128/MCB.01395-06

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 (New York, N.Y.)*, *295*(5556), 858–61. doi:10.1126/science.1068592

Larsen, H., Muz, B., Khong, T. L., Feldmann, M., & Paleolog, E. M. (2012). Differential effects of Th1 versus Th2 cytokines in combination with hypoxia on HIFs and angiogenesis in RA. *Arthritis research & therapy*, *14*(4), R180. doi:10.1186/ar3934

Loi, S., Haibe-Kains, B., Desmedt, C., Wirapati, P., Lallemand, F., Tutt, A. M., … Sotiriou, C. (2008). Predicting prognosis using molecular profiling in estrogen receptor-positive breast cancer treated with tamoxifen. *BMC genomics*, *9*, 239. doi:10.1186/1471-2164-9-239

Lu, X., & Kang, Y. (2010). Hypoxia and hypoxia-inducible factors: master regulators of metastasis. *Clinical cancer research : an official journal of the American Association for Cancer Research*, *16*(24), 5928–35. doi:10.1158/1078-0432.CCR-10-1360

Maxwell, P. H., Wiesener, M. S., Chang, G. W., Clifford, S. C., Vaux, E. C., Cockman, M. E., … Ratcliffe, P. J. (1999). The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature*, *399*(6733), 271–5. doi:10.1038/20459

Miró-Murillo, M., Elorza, A., Soro-Arnáiz, I., Albacete-Albacete, L., Ordoñez, A., Balsa, E., … Aragonés, J. (2011). Acute Vhl Gene Inactivation Induces Cardiac HIF-Dependent Erythropoietin Gene Expression. (M. Rojas, Ed.)*PLoS ONE*, *6*(7), e22589. doi:10.1371/journal.pone.0022589

Ng, P., Wei, C., Sung, W., Chiu, K. P., Lipovich, L., Ang, C. C., … Ruan, Y. (2005). Gene identification signature (GIS) analysis for transcriptome characterization and genome annotation. *Nature methods*, *2*(2), 105–11. doi:10.1038/nmeth733

Nie, L., Wu, H., Hsu, J.-M., Chang, S., Labaff, A. M., Li, C., … Hung, M. (2012). Long non-coding RNAs: versatile master regulators of gene expression and crucial players in cancer. *American*

journal of translational research, *4*(2), 127–50.

Nievergall, E., Lackmann, M., & Janes, P. W. (2012). Eph-dependent cell-cell adhesion and segregation in development and cancer. *Cellular and molecular life sciences : CMLS*, *69*(11), 1813–42. doi:10.1007/s00018-011-0900-6

Ogawa, K., Pasqualini, R., Lindberg, R. a, Kain, R., Freeman, a L., & Pasquale, E. B. (2000). The ephrin-A1 ligand and its receptor, EphA2, are expressed during tumor neovascularization. *Oncogene*, *19*(52), 6043–52. doi:10.1038/sj.onc.1204004

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 research*, *38*(7), 2332–45. doi:10.1093/nar/gkp1205

Pasquale, E. B. (2008). Eph-ephrin bidirectional signaling in physiology and disease. *Cell*, *133*(1), 38–52. doi:10.1016/j.cell.2008.03.011

Pasquale, E. B. (2010). Eph receptors and ephrins in cancer: bidirectional signalling and beyond. *Nature reviews. Cancer*, *10*(3), 165–80. doi:10.1038/nrc2806

Pawitan, Y., Bjöhle, J., Amler, L., Borg, A.-L., Egyhazi, S., Hall, P., … Bergh, J. (2005). Gene expression profiling spares early breast cancer patients from adjuvant therapy: derived and validated in two population-based cohorts. *Breast cancer research : BCR*, *7*(6), R953–64. doi:10.1186/bcr1325

Pescador, N., Cuevas, Y., Naranjo, S., Alcaide, M., Villar, D., Landázuri, 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. *The Biochemical journal*, *390*(Pt 1), 189–97. doi:10.1042/BJ20042121

R Core Team. (2010). R : A Language and Environment for Statistical Computing. Vienna.

Salceda, S., & Caro, J. (1997). Hypoxia-inducible factor 1alpha (HIF-1alpha) protein is rapidly degraded by the ubiquitin-proteasome system under normoxic conditions. Its stabilization by hypoxia depends on redox-induced changes. *The Journal of biological chemistry*, *272*(36), 22642–7.

Semenza, G. L. (2003). Targeting HIF-1 for cancer therapy. *Nature reviews. Cancer*, *3*(10), 721–32. doi:10.1038/nrc1187

Semenza, G. L. (2012). Hypoxia-inducible factors: mediators of cancer progression and targets for cancer therapy. *Trends in pharmacological sciences*, *33*(4), 207–14. doi:10.1016/j.tips.2012.01.005

Sims, D., Bursteinas, B., Gao, Q., Jain, E., MacKay, A., Mitsopoulos, C., & Zvelebil, M. (2010). ROCK: a breast cancer functional genomics resource. *Breast cancer research and treatment*, *124*(2), 567–72. doi:10.1007/s10549-010-0945-5

Surawska, H., Ma, P. C., & Salgia, R. (2004). The role of ephrins and Eph receptors in cancer. *Cytokine & growth factor reviews*, *15*(6), 419–33. doi:10.1016/j.cytogfr.2004.09.002

Tiana, M., Villar, D., Pérez-Guijarro, E., Gómez-Maldonado, L., Moltó, E., Fernández-Miñán, A., … Del Peso, L. (2011). A role for insulator elements in the regulation of gene expression response to hypoxia. *Nucleic acids research*, *40*(5), 1916–1927. doi:10.1093/nar/gkr842

Van de Vijver, M. J., He, Y. D., van't Veer, L. J., Dai, H., Hart, A. A. M., Voskuil, D. W., … Bernards, R. (2002). A gene-expression signature as a predictor of survival in breast cancer. *The New England journal of medicine*, *347*(25), 1999–2009. doi:10.1056/NEJMoa021967

Vaupel, P., & Mayer, A. (2007). Hypoxia in cancer: significance and impact on clinical outcome. *Cancer metastasis reviews*, *26*(2), 225–39. doi:10.1007/s10555-007-9055-1

Vihanto, M. M., Plock, J., Erni, D., Frey, B. M., Frey, F. J., & Huynh-Do, U. (2005). Hypoxia up-regulates expression of Eph receptors and ephrins in mouse skin. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, *19*(12), 1689–91. doi:10.1096/fj.04-3647fje

Villar, D., Ortiz-Barahona, A., Gómez-Maldonado, L., Pescador, N., Sánchez-Cabo, F., Hackl, H., … Del Peso, L. (2012). Cooperativity of Stress-Responsive Transcription Factors in Core Hypoxia-Inducible Factor Binding Regions. *PloS one*, *7*(9), e45708. doi:10.1371/journal.pone.0045708

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. *Proceedings of the National Academy of Sciences of the United States of America*, *92*(12), 5510–4.

Xiao, F., Qiu, H., Zhou, L., Shen, X., Yang, L., & Ding, K. (2013). WSS25 inhibits Dicer, downregulating microRNA-210, which targets Ephrin-A3, to suppress human microvascular endothelial cell (HMEC-1) tube formation. *Glycobiology*, *23*(5), 524–35. doi:10.1093/glycob/cwt004

Yamashita, T., Ohneda, K., Nagano, M., Miyoshi, C., Kaneko, N., Miwa, Y., … Fujii-Kuriyama, Y. (2008). Hypoxia-inducible transcription factor-2alpha in endothelial cells regulates tumor neovascularization through activation of ephrin A1. *The Journal of biological chemistry*, *283*(27), 18926–36. doi:10.1074/jbc.M709133200

Yoon, J.-H., Abdelmohsen, K., & Gorospe, M. (2012). Posttranscriptional Gene Regulation by Long Noncoding RNA. *Journal of molecular biology*, *2*. doi:10.1016/j.jmb.2012.11.024

Zhang, H., Wong, C. C. L., Wei, H., Gilkes, D. M., Korangath, P., Chaturvedi, P., … Semenza, G. L. (2011). HIF-1-dependent expression of angiopoietin-like 4 and L1CAM mediates vascular metastasis of hypoxic breast cancer cells to the lungs. *Oncogene*, 1–14. doi:10.1038/onc.2011.365

Figure 1. Hypoxia regulates *EFNA3* expression. (A) Ephrin-A3 protein level was determined in HeLa, HUVEC, MCF-7 and MDA-MB231 cells were exposed to normoxia (21% oxygen) or hypoxia (1% oxygen) for the indicated periods of time. (B) Diagram of the *EFNA3* locus showing the canonical, NM_004952 RefSeq, gene (upper track, *EFNA3*) and the novel isoforms (second track from the top, N1-NC2s). The bottom tracks show the position of the qPCR oligonucleotides employed in figures 2D-2G (black boxes) and target regions (thin lines). The figure was generated by the UCSC genome browser upon loading the indicated custom tracks. (C) cDNAs corresponding to *EFNA3*, NC1 or NC2 isoforms were transcribed and translated *in vitro* in the presence of 35S-methionine and proteins resolved by SDS-PAGE. Image shows the autoradiogram of a representative experiment. (D) HeLa cells were transfected with plasmids encoding for the indicated *EFNA3* isforms and cell lysates probed with monoclonal (MAb) or polyclonal (PAb) antibodies against the C-terminal region of human Ephrin-A3. (E) HeLa cells were exposed to 1% oxygen (Hyp) or left at normoxic conditions (Nx) for 12h and *EFNA3* RNA levels were determined using the indicated TaqMan probes (see panel B). The graph shows the ratio of *EFNA3* to ACTB (beta-Actin) copy number. Bars represent the mean of three independent biological replicates and the error bars the standard deviation. The differences between groups were statistically significant (ANOVA $F_{3,8}=34.21$, $p<0.001$) and the asterisks indicate means pairs that were the statistically significant (adjusted $p <0.001$) in a *posteriori* Tukey test. (F) Cells were treated as in E and *EFNA3* RNA levels were determined using the indicated primer pairs (see panel B). The graph shows the ratio of hypoxic *EFNA3* RNA to the expression in normoxia. Symbols represent the mean of three independent biological replicates and the error bars the standard deviation. (G) HeLa cells were treated with the indicated siRNAs and then grown at normoxic (21% oxygen, Nx) or hypoxic (1% oxygen, Hyp) conditions for 12 hours. The graph represent the normalized levels of *EFNA3* mRNA as fold over control conditions (normoxic cells treated with scramble siRNA). Bars

represent average values in four independent biological replicates and error bars the standard deviation. The differences between groups were statistically significant (ANOVA $F_{7,48}$=23.52, p<0.001) and the asterisks indicate means pairs that were the statistically significant (***, adjusted p <0.001) in a *posteriori* Tukey test. (H) Vhlfl/fl-UBC-Cre-ERT2 (n = 5), Vhlfl/flHIF2αfl/fl-UBC-Cre-ERT2 (n=3) and control (n=11) mice were placed on a tamoxifen diet for ten days followed by ten additional days on a normal diet. The box-and-whisker plot represents the distribution of the normalized *EFNA3* mRNA expression in the liver of animals with the indicated genotypes. The box contains the values comprised between the second and third quartiles, and the horizontal black line the median. The "whiskers" extent to 1.5 times the interquartile range. The differences between groups were statistically significant (ANOVA $F_{2,18}$=20.49, p<0.001) and the asterisks indicate means pairs that were the statistically significant (adjusted p <0.001) in a *posteriori* Tukey test.

Figure 2. lncRNA expression results in increased *EFNA3* protein levels. HeLa cells were transfected with constructs encoding for the indicated lncRNAs or empty plasmid (pLOC) and the level of the canonical *EFNA3* mRNA (A) and protein (B and C) were determined by qPCR and immunoblot respectively. The graph in A represents the level of the NM_004952 mRNA, determined with primers Exon1+2 (Figure 3A), as a fold over the level found in untreated cells. Bars represent the mean of five independent biological replicates and the error bars the standard deviation. The differences between groups were not statistically significant (ANOVA $F_{4,20}$=1.452, p=0.254). The graph in C represents the level of Ephrin-A3 as the ratio of the Ephrin-A3 band intensity corrected by the Actin band intensity in three independent experiments. The differences between groups was statistically significant (ANOVA $F_{5,11}$=8.635, p<0.01) and the asterisks indicate sample means that were was significantly different from controls (pLOC samples) in a *posteriori* Tukey test (*, adjusted p < 0.05; **, adjusted p < 0.01). The image in B is representative of at least three independent experiments.

24

Figure 3. lncRNAs encoded by the EFNA locus are transcribed from an alternative hypoxia-responsive promoter. (A) diagram depicting the *EFNA3* locus and showing the NM_004952 RefSeq gene ("*EFNA3*" track) along with accessible chromatine regions ("DNase clusters track"), histone marks associated to promoters ("Layered H3K4Me3" track) and active regulatory elements ("H3K27Ac Track"). The colors in the histone tracks correspond to the signal obtained in different cell lines (see UCSC for details). The figure was generated by the UCSC genome browser upon loading custom tracks to indicate the location of primer and amplicons used in the RNA pol II (P1 and P2) and Hif1a (H1-H6) ChIP-qPCR experiments as well as the regions cloned to assay their promoter activity ("promoters" track). (B, C) HeLa cells were exposed to 21% or 1% oxygen for 8 hours and RNApol-II (B) or Hif1a (C) binding to the indicated regions of the *EFNA3* locus was determined by ChIP-qPCR. Binding to the P4HA and STT3S promoters were used in the RNApol II ChIP experiment as positive and negative controls respectively. In the case of Hif1a ChIP, the EGLN3 enhancer region (E3_E) and EGLN3 promoter region (E3_P) were used as positive and negative controls (Pescador et al., 2005) . The graphs show the amount of precipitated material as a percentage of the input (%Enrichment). Bars represent the mean of three (B) or just one (C) independent biological replicates and the error bars the standard deviation. The differences between groups were statistically significant (ANOVA $F_{15,69}$=15.6, p<0.001; ANOVA $F_{31,32}$=39.99, p<0.001) and the asterisks indicate means pairs that were the statistically significant (adjusted p <0.001) in a *posteriori* Tukey test. (D) The effect of hypoxia (white bars) and DMOG (black bars) on the transcriptional activity of the promoter regions 1 and 2 was assessed by reporter assays upon transfection of the indicated constructs into HeLa cells. White boxes within the promoter 2 (Prom2) diagram represent RCGTG motifs and crossed boxes represent deleted RCGTG motifs. The graph shows the normalized luciferase activity in hypoxic (Hyp) or DMOG-treated (DMOG) samples expressed as fold over the activity obtained in normoxic conditions. Bars

25

represent the mean of three (B) independent biological replicates and the error bars the standard deviation. The differences between groups were statistically significant (ANOVA $F_{4,10}$=21.9, p<0.001) and the asterisks indicate means pairs that were the statistically significant (*, adjusted p < 0.05; **, adjusted p < 0.01) in a *posteriori* Tukey test.

Figure 4. *EFNA3* **expression is induced in human renal clear cell carcinomas (RCC).** The expression of EGLN3, EFNA3, EFNA1 and EFNA4 (columns) was determined in publicly available gene expression profiles of ccRCC samples from three independent studies (ID shown on the left margin of each row). Graphs represent the expression of the indicated genes in arbitrary units (normalized microarray intensity values). The individual samples are shown and pairs of tumoral and normal kidney tissue are joined by segments. Graphs in each row represent the data from an independent study. The statistical significance of mean differences in indicated on top of each graph (paired student's t test).

Figure 5. *EFNA3* **overexpression increases metastasis formation in a spontaneous metastasis assay.** MDA-MB-231 cells were engineered to stably express luciferase alone (pLOC) or in combination with *EFNA3* (*EFNA3*). Then cells were orthotopically injected into the mammary fat pad of nude mice and tumor growth followed by non-invasive whole body bioluminescence imaging (BLI). (A) BLI image of mice at end time point (5 weeks). (B) Luminiscence signal during the course of the experiment. The graph represents the luciferase signal (total flux) in photons/second x10^8. (C) Tumor volume was determined at five weeks. The boxplot represents the distribution of volume in each group. The differences between controls and *EFNA3*-expressing tumors was not significant (independent samples t test: t_{12}=-1.5373, P=0.1499). Animals were euthanized and tissues processed for H&E staining (D,E) or immunostaining (F,G). (D) The box-and-whisker plot represents the distribution of the number

26

of macroscopic metastasis per mice identified by visual inspection during necropsy. The difference between groups was statistically significant (independent samples t test: t_7=2.5923, P<0.05). (E) Representative H&E staining images at 10x (left) and 40x (right) magnification (F) Blood vessels were stained with an antibody against human CD31 and cell nuclei with DAPI. The CD31 signal was normalized to cellularity (DAPI staining). The graph represents the distribution of normalized values. Differences between groups was not significant (independent samples t test: t_{10}=1.0522, P=0.3182) (G) Representative images of CD31 and DAPI staining.

Figure 6. *EFNA3* expression results in repulsion of HUVEC cells and promotes extravasation. (A,B) HEK293 cells were transiently transfected with a plasmids encoding for GFP and *EFNA3* ("*EFNA3*") or GFP alone ("Vector") and then plated on top of a monolayer of HUVEC cells previously labeled with a red fluorochrome. (A) Representative images of cells 2h and 24h after plating HEK293 cells on top of the HUVEC monolayer. (B) The graph represents the proportion of GFP-positive cells attached to the plastic surface and spreading at 24h. The differences between cells expressing *EFNA3* and cells transfected with GFP alone was statistically significant (2x2 contingency table: χ^2_1=67.52, p<0.001). (C) HUVEC migration was determined using NIH conditioned media (NIH_CM) as chemoattractant in the absence or presence of 0.2, 1 or 5 microgr/ml of recombinant Ephrin-A3 fused to the Fc region of immunoglobulins (EphrA3). The graph represents the average number of migrated cells per field in a single experiment and error bars the standard deviation. The experiment was repeated three independent times with similar results. The differences among treatments was statistically significant (ANOVA $F_{7,178}$=47.17, p<0.001) and the asterisks indicate means that were the statistically significant to the NIH-conditioned media treatment (***, adjusted p < 0.001) in a *posteriori* Tukey test. (D,E) MDA-MB-231 cells stably expressing GFP alone or in combination with *EFNA3* were injected in the tail of nude mice. After the indicated periods of time, mice were

27

euthanized and lungs were examined by confocal microscopy for the presence of GFP-positive cells. The boxplot represents the distribution of total GFP signal normalized to lung area in each group of mice. Results from two independent experiments were pooled. The differences between the two groups was statistically significant (Student's t-test $t(3) = -4.44$, $p<0.05$).

Figure 7. High levels of *EFNA3* correlate with metastasis in human tumors. Gene expression profiles from the indicated series of breast cancer tumors (Loi2008, (Loi et al., 2008) ; Pawitan2005, (Pawitan et al., 2005) ; Ur-Rehman, www.rock.icr.ac.uk/; vandeVijve, (van de Vijver et al., 2002)) were downloaded from the ROCK database (www.rock.icr.ac.uk/). Samples were categorized according to *EFNA3* expression into high (samples with *EFNA3* expression in the top quartile of the series, labeled in red) and low expression (rest of samples in the serie, labeled in blue) and the Kaplan−Meier estimate of the distant metastasis free survival (DMFS) over time calculated (graphs). The survival of the two groups was compared using a Cox proportional hazards model, the p-values are indicated in the graphs. The probes used in the gene profiling assays are indicated in brackets in the graph title.
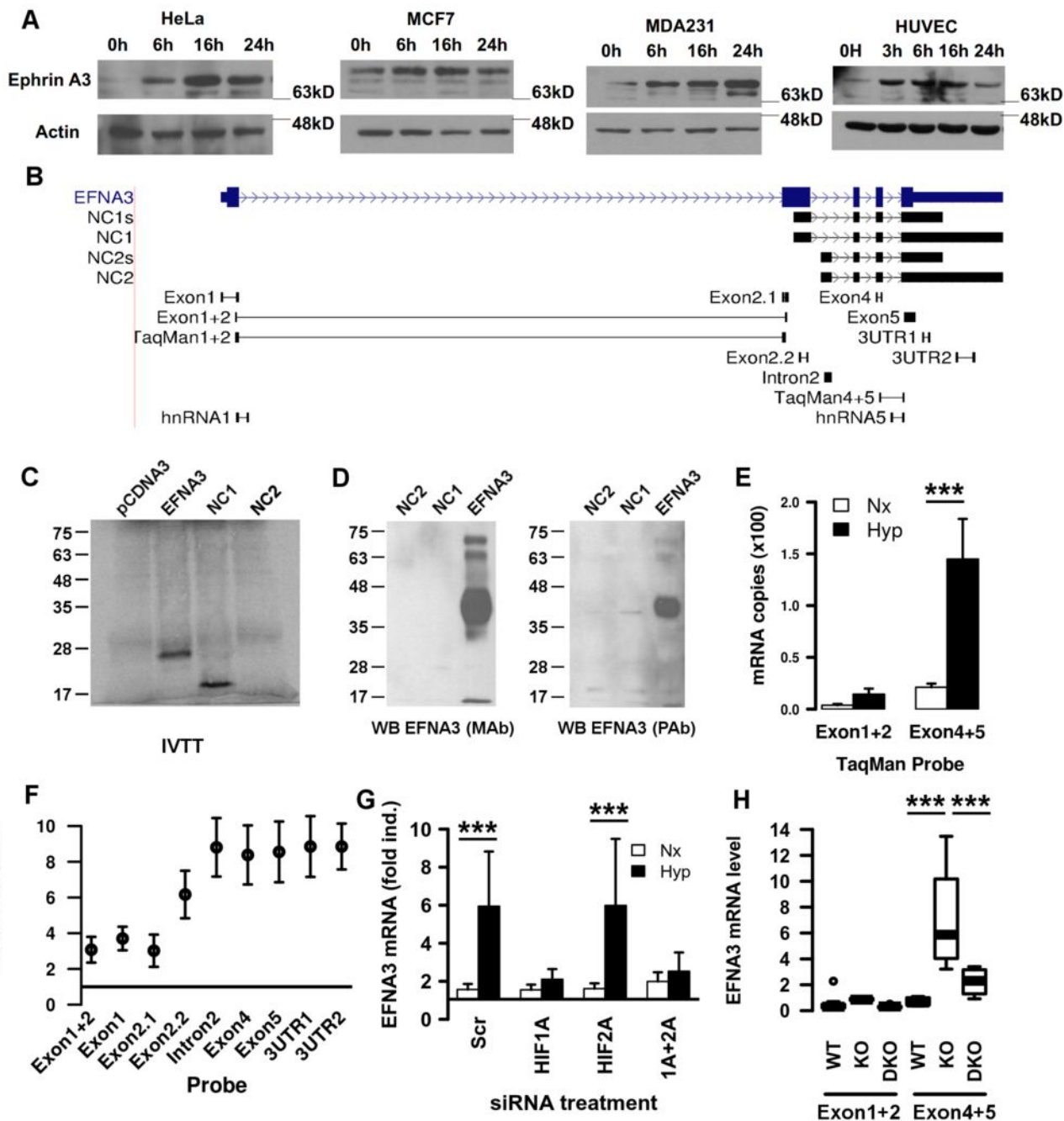
FIGURE 1

FIGURE 2

A

NM_004952 mRNA level

pLOC NC1 NC2 NC1s NC2s

B

none pLOC NC1 NC2 NC1s NC2s

75 —
63 —   EFNA3

48 —   Actin

C

Ephrin-A3/actin ratio

none pLOC NC1 NC2 NC1s NC2s

FIGURE 3

FIGURE 4

FIGURE 5

FIGURE 6

A

Vector · 2h   EFNA3 · 2h

Vector · 24h   EFNA3 · 24h

B

Proportion of cells

***

EFNA3   Vector

■ Spread □ Round

C

Migrated cells/field

NIH_CM: - - + - - - - + + +
EphrinA3: - - - ◿ - ◿ - ◿

D

Vector · 2h   EFNA3 · 2h

Vector · 8d   EFNA3 · 8d

E

GFP signal (1/1000)

*

EFNA3   pLOC
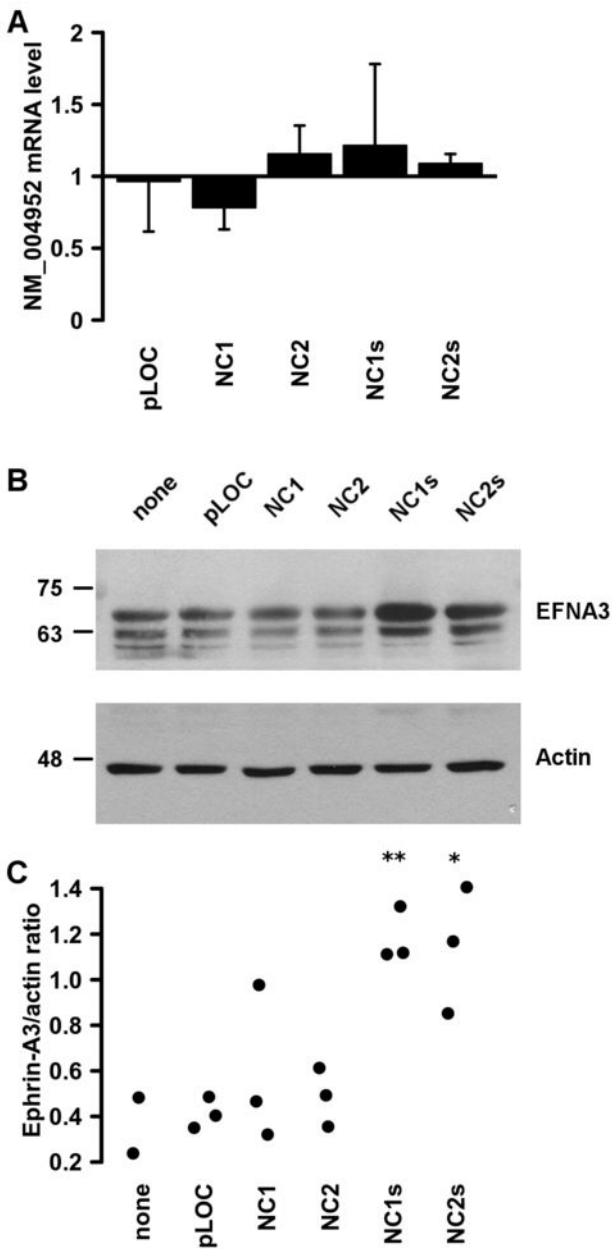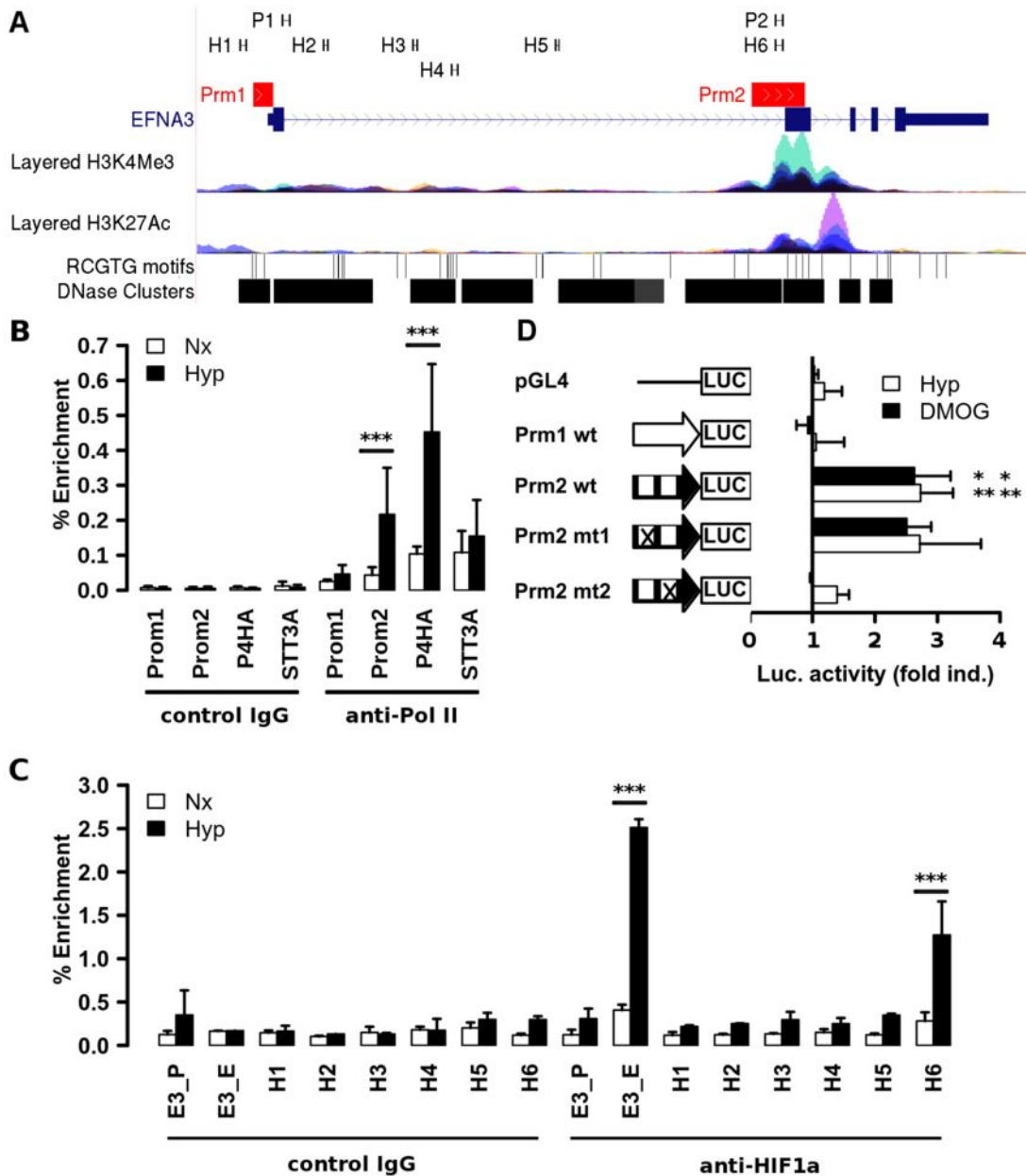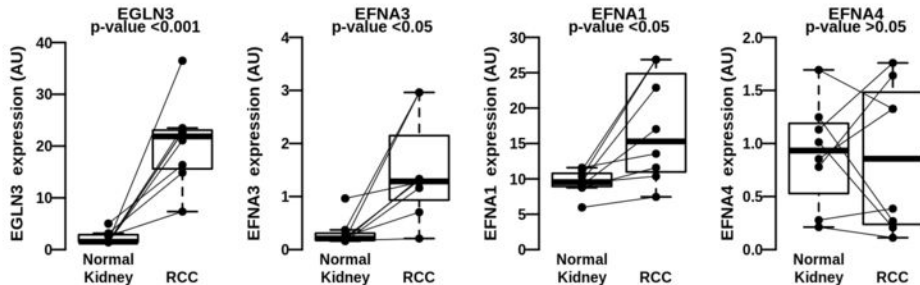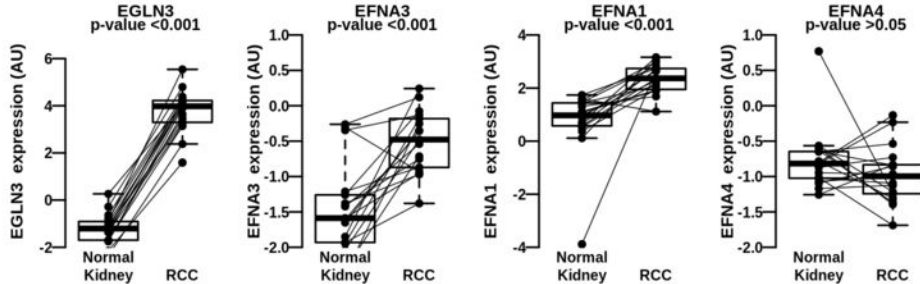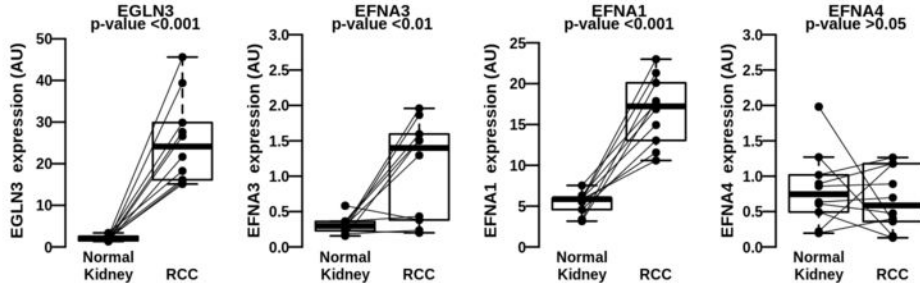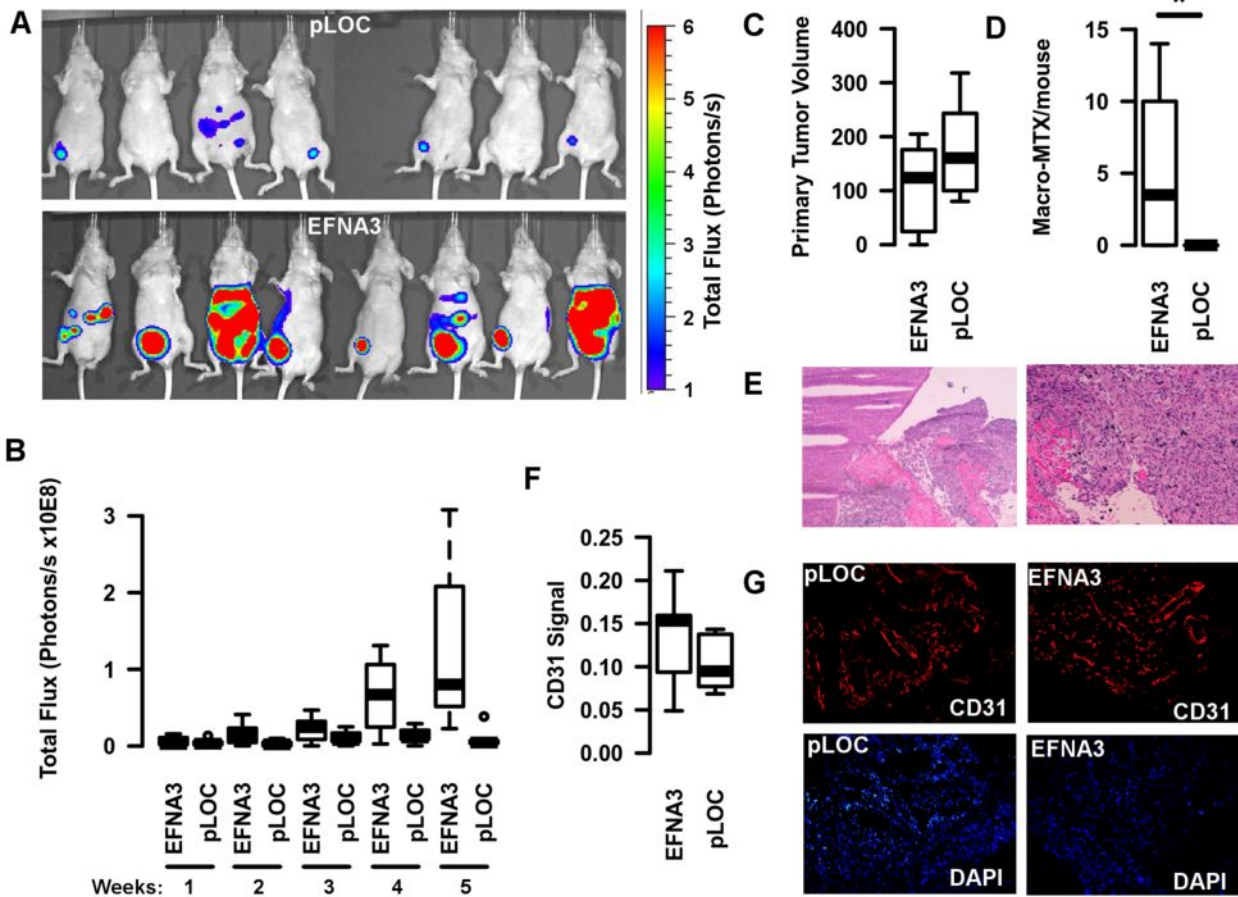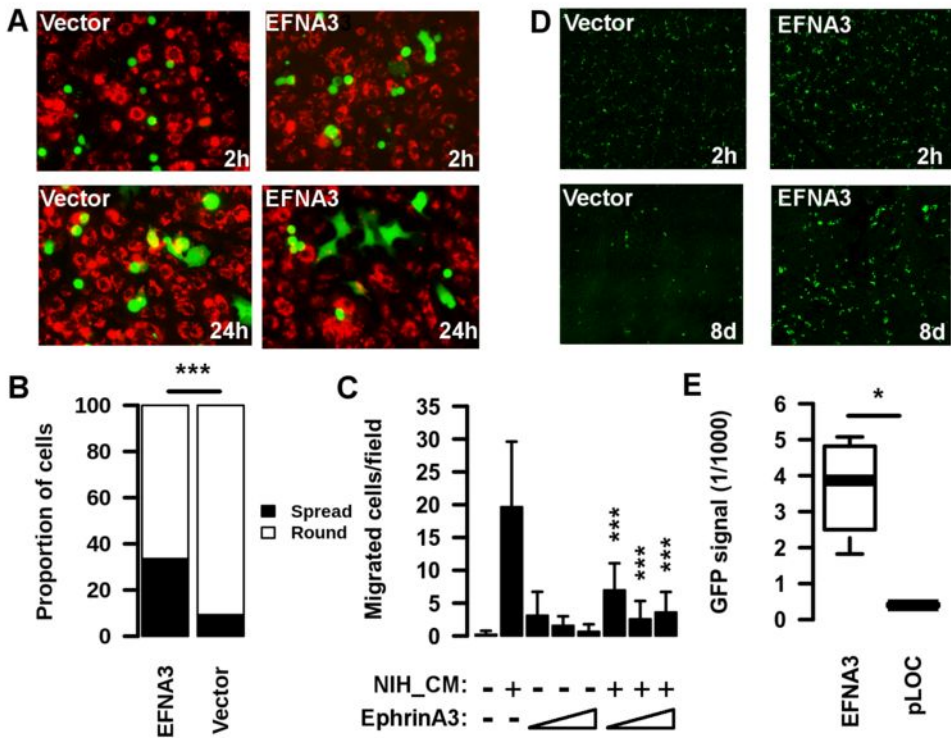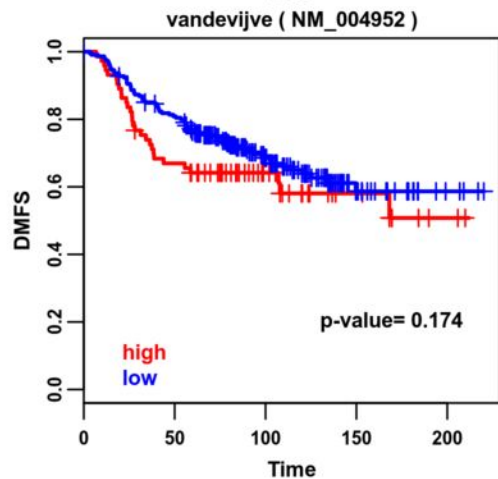
FIGURE 7

Supplementary Figure 1
EphrinA3 migrates as a ~72kD band in SDS-PAGE due to postranslational modifications.
EFNA3 mRNA levels were low in all cell types tested and, accordingly, the immunoblot detection of Ephrin-3A proved difficult. However, Ephrin-A3 migrates with an apparent molecular weight of about 72kD (figure 1C and supplementary figure 1A), much higher than the hypothetical size of 26.3kD predicted from its amino acid sequence. Thus, to confirm the identity of the protein recognized by the antibody, we treated cells with a siRNA directed against EFNA3 and found that it decreased tha intensity of the high molecular band specifically (A and B). Type A Ephrins undergo several postranslational modifications, including the addition of a GPI moiety and glycans, that could account for its high apparent mass. In agreement, overexpression of EFNA3 cDNA also resulted in high molecular weight bands (A). Moreover, treatment of cell lysates 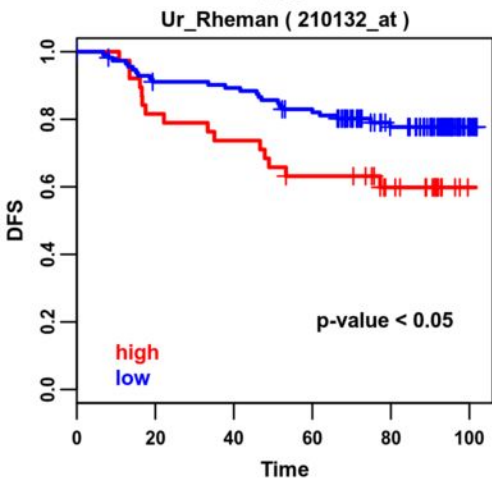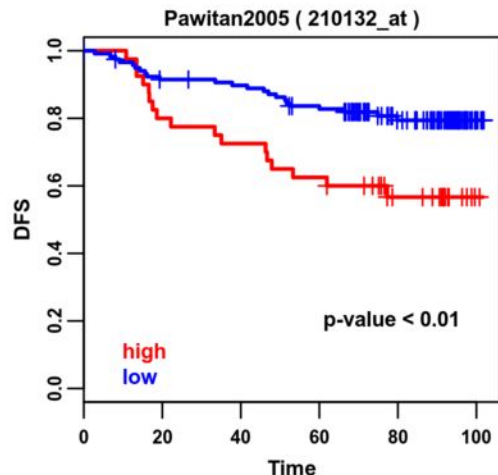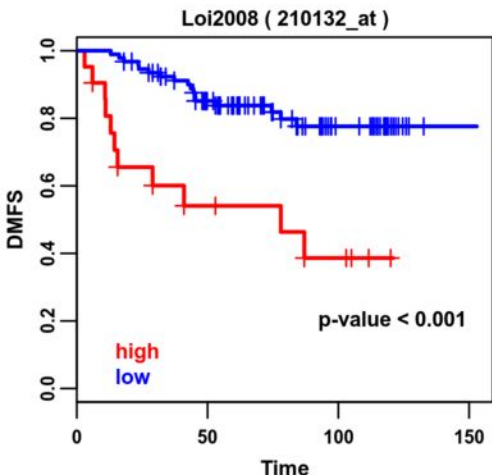with N-glycosidase, resulted in the depletion of the endogenous ~72kD bands and generation of a faster migrating band. Note that the intensity of the ~50kD band generated by the enzymatic treatment is much higher than that of the corresponding 72kD bands, probably due the exposure of the antigenic determinants after the removal of the oligosaccharide chains, as has been shown for other postranslational modifications (Bütikofer, Malherbe, Boschung, & Roditi, 2001) (C). Finally, we used a recombinant form of EphA5 (a high affinity receptor for Ephrin-A3) fused to the immunoglobulin Fc fragment to pull down endogenous Ephrin-A3 from cell lysates. As shown in panel D, the Ephrin-A3 forms that co-precipitated with the EphA5-Fc migrate as a ~72Kd bands. Altogether these results indicate that Ephrin-3A is detected as ~72Kd bands in SDS-PAGE gels due to postranslational modifications.
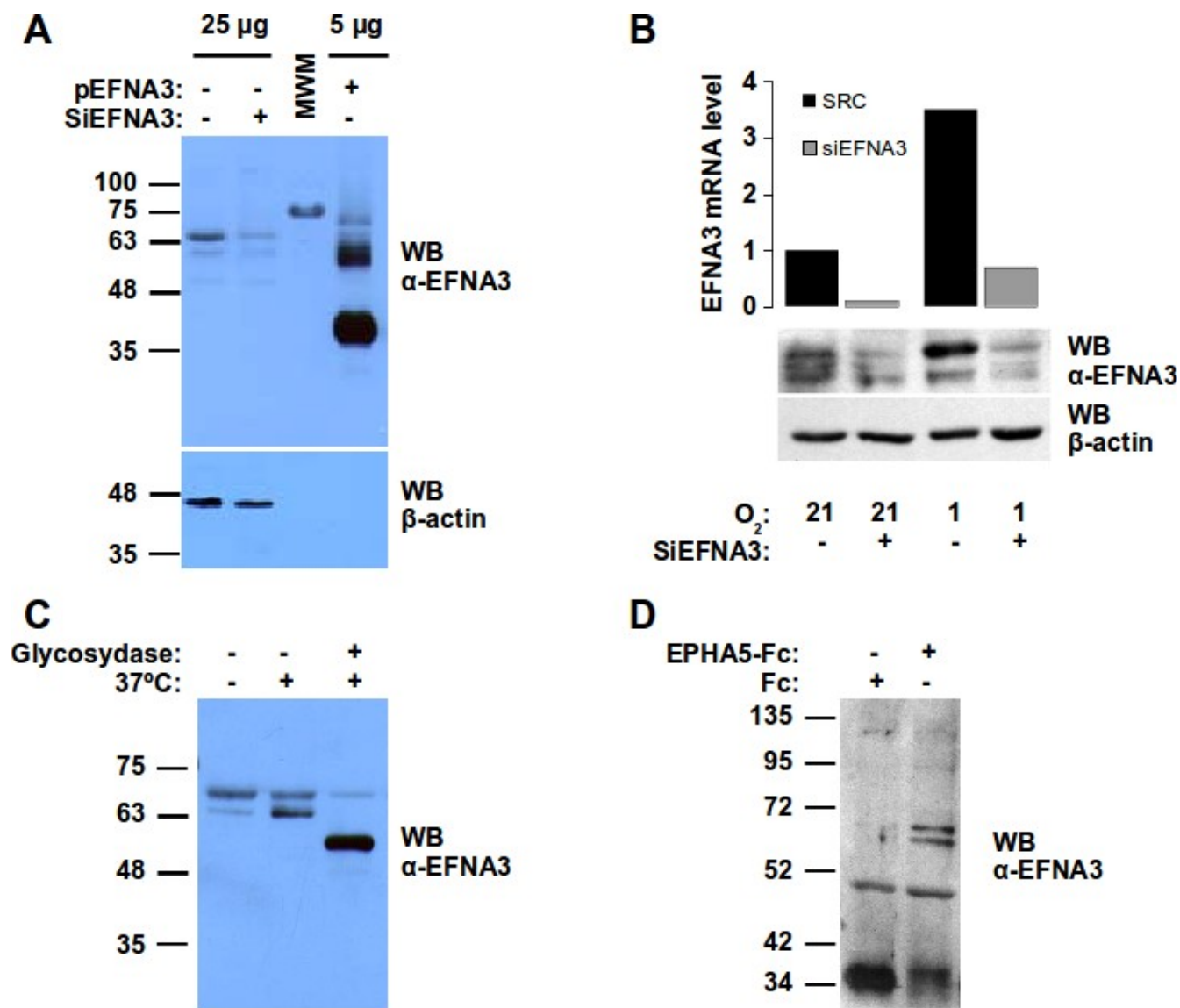
Figure legend

(A) HeLa cells were transfected with a plasmid encoding EFNA3 (pEFNA3) or siRNA directed against EFNA3 and 48h postransfection EFNA3 proteins levels were determined by inmmunoblot. To allow comparison in the same membrane, different amounts of cell lysate were used to detect the endogenous and exogenous proteins. MWM, molecular weight markers, the cross-reactivity band correspond to the 75kD marker. (B) HeLa cells were transfected with scramble siRNA (-) or siRNA directed against EFNA3 and then grown at 21% (Nx) or 1% oxygen (Hyp) for 12 hours. Duplicated samples were processed to determine EFNA3 mRNA (upper graph) or protein (lower panels) levels. (C) HeLa cell lysates were incubated at 37°C (+) or 0°C (-) in the presence (+) or absence (-) of N-glycosidase and then resolved by SDS-PAGE and probed with anti-EFNA3 antibodies. (D) HeLa cells lysates were incubated with recombinant Fc-EphA5 receptor or Ig Fc fragments. Recombinant contructs were pulled down with protein A-sepharose and coinmunoprecipitating material was resolved by SDS-PAGE and probed with anti-EFNA3 antibodies.
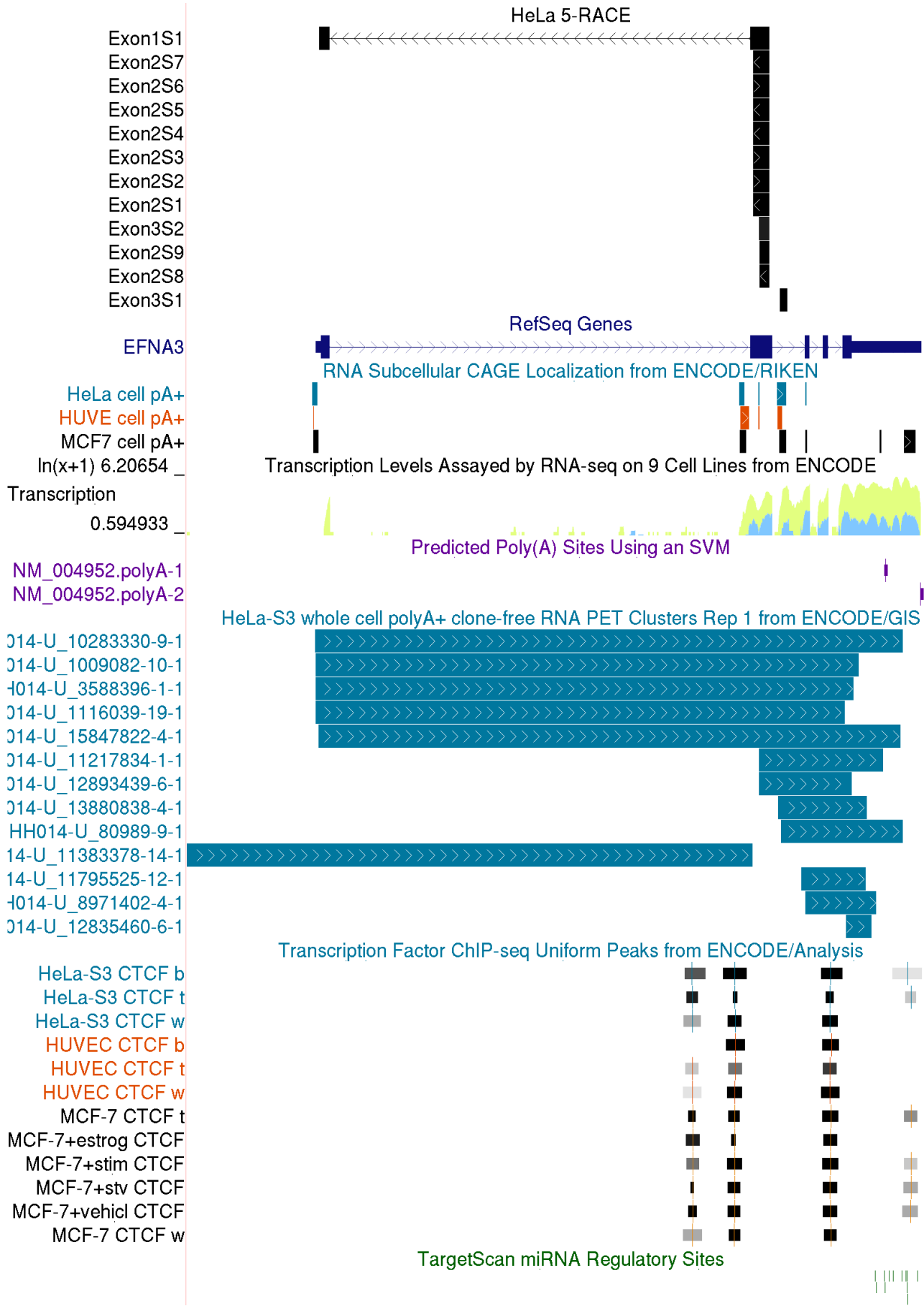
Supplementary Figure 2
EFNA3 locus encodes for several transcripts expressed from two different promoter regions.
The 5'-ends of the EFNA3 transcripts was determined by rapid amplification of cDNA ends (RACE) using the FirstChoice RLM-RACE kit (Applied Biosystems) according to manufacturer's instructions. Total RNA from HeLa and LoVo cells was used as template for 5'- RACE reactions, the products cloned into pCR2.1 (Invitrogen) and their sequence determined by Sanger sequencing. The figure shows the position of the RACE products ("HeLa 5-RACE") relative to the canonical NM_004952 RefSeq gene (EFNA3 track). The names indicate the location of EFNA3-specific primers used for the PCR amplification of RACE products: AGAAGGCGCTGTAGCGCTGGAA (Exon1 and 2) ; TTCCAGTGCAGGTTGTGAGT (Exon 3).
Similar results were obtained when RNA form LoVo cells was used as template (data not shown).
In addition to our RACE results, the figure includes data from other sources as follows (tracks from top to bottom). The "RNA Subcellular CAGE" track shows 5' cap analysis gene expression (CAGE) tags from the ENCODE project and thus points to Transcription Start Sites (TSS) within this locus. Note the tight correlation between our 5'-RACE results and those reported by ENCODE. To keep image size to a minimum, in this and following tracks we only included information from cell lines used in our work (HeLa, HUVEC and MCF7). The "Transcription" track shows transcription levels for HeLa (green) and HUVEC (blue) as assayed by high-throughput sequencing of polyadenylated RNA (RNA-seq). Note that, in close agreement with our data (figure 2), the expression level of the first exon is below the rest. This pattern is particularly pronounced in HUVEC were there is no detectable expression of the first exon. The prediction of Poly(A) sites using a supported vector machine-based algorithm insicates the existence of two different 3-end for this locus ("predicted poly(A) sites" track). The next track shows the result of Gene Identification Signature (GIS) paired-end ditag (PET) sequencing experiments in HeLa cells. This technique provide signatures of the 5' start and the 3' end of individual mRNA transcripts (Ng et al., 2005). The "Transcription Factor ChIP-Seq" track shows the CTCF binding sites in HeLa, HUVEC and MCF7 cells. Note that the HIF binding site and prom2 region (figure 2) identified in our work are flanked by strong CTCF-binding signal and thus probably isolated from other regions of this locs and in particular de prom1 region. Finally, the "TargetScan miRBNA" track shows the predicted binding sites (TargetScan) for several miRNAs one of them being miR210.
The figure was generated by the UCSC genome browser upon loading the indicated custom tracks.

HeLa 5-RACE

Exon1S1
Exon2S7
Exon2S6
Exon2S5
Exon2S4
Exon2S3
Exon2S2
Exon2S1
Exon3S2
Exon2S9
Exon2S8
Exon3S1

RefSeq Genes

EFNA3

RNA Subcellular CAGE Localization from ENCODE/RIKEN

HeLa cell pA+
HUVE cell pA+
MCF7 cell pA+
ln(x+1) 6.20654 _

Transcription Levels Assayed by RNA-seq on 9 Cell Lines from ENCODE

Transcription

0.594933 _

Predicted Poly(A) Sites Using an SVM

NM_004952.polyA-1
NM_004952.polyA-2

HeLa-S3 whole cell polyA+ clone-free RNA PET Clusters Rep 1 from ENCODE/GIS

014-U_10283330-9-1
014-U_1009082-10-1
H014-U_3588396-1-1
014-U_1116039-19-1
014-U_15847822-4-1
014-U_11217834-1-1
014-U_12893439-6-1
014-U_13880838-4-1
HH014-U_80989-9-1
14-U_11383378-14-1
14-U_11795525-12-1
H014-U_8971402-4-1
014-U_12835460-6-1

Transcription Factor ChIP-seq Uniform Peaks from ENCODE/Analysis

HeLa-S3 CTCF b
HeLa-S3 CTCF t
HeLa-S3 CTCF w
HUVEC CTCF b
HUVEC CTCF t
HUVEC CTCF w
MCF-7 CTCF t
MCF-7+estrog CTCF
MCF-7+stim CTCF
MCF-7+stv CTCF
MCF-7+vehicl CTCF
MCF-7 CTCF w

TargetScan miRNA Regulatory Sites

Supplementary Figure 3
Efficiency of HIF1A and EPAS1 knock-down in HeLa cells
HeLa cells were treated with the indicated siRNAs and then grown at normoxic (21% oxygen, Nx) or hypoxic (1% oxygen, Hyp) conditions for 12 hours and the levels of HIF1A (A) and EPAS1 (B) mRNA were determined by qPCR. The graph represent the normalized levels of mRNA as fold over control conditions (normoxic cells treated with scramble siRNA). Bars represent average values in four independent biological replicates and error bars the standard deviation. The differences between groups were statistically significant: (A) ANOVA $F7,24=1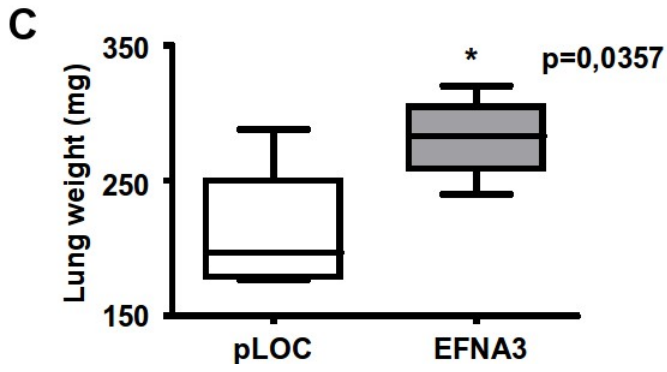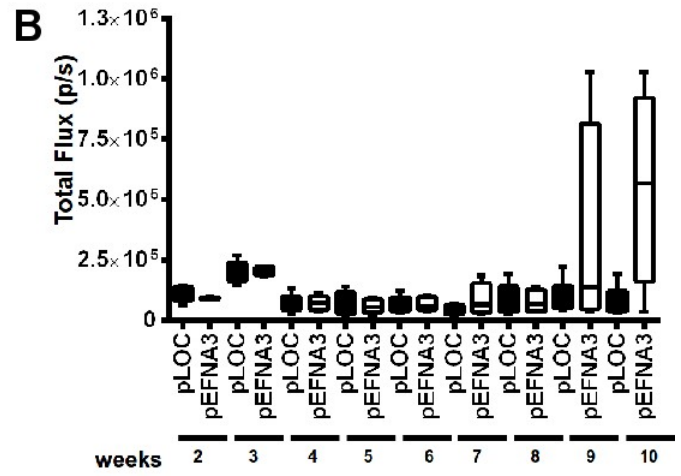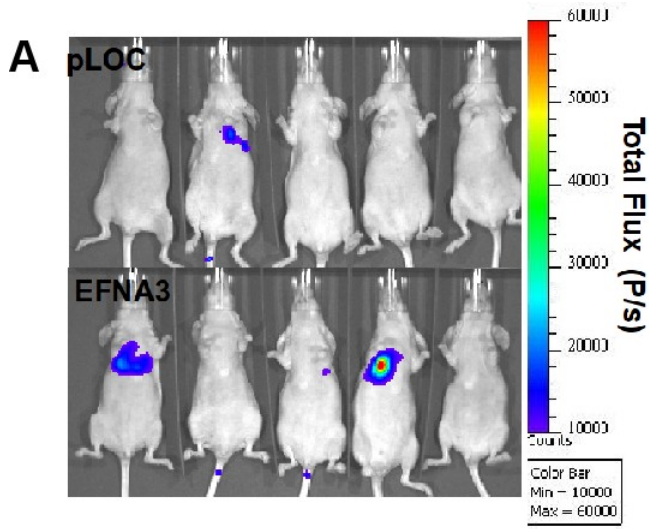6.9$, $p<0.001$; (B) ANOVA $F7,20=17.58$, $p<0.001$. The asterisks indicate means pairs that were the statistically significant (***, adjusted p <0.001) in a posteriori Tukey test.
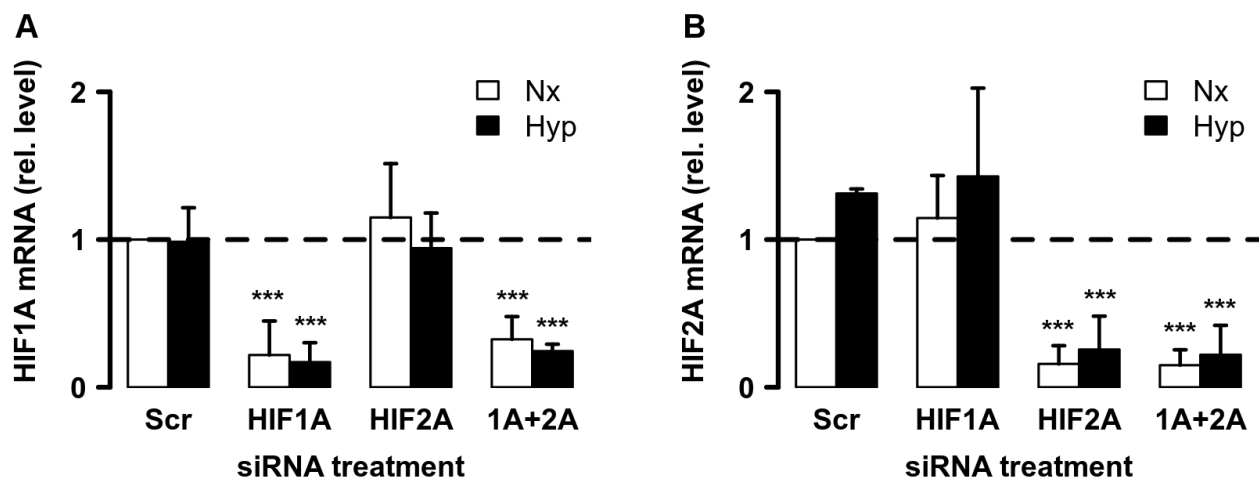
Supplementary Figure 4
EFNA3 overexpression increases metastasis formation in an experimental metastasis assay.
MDA-MB-231 cells were engineered to stably express luciferase alone (pLOC) or in combination with EFNA3 (EFNA3). Then cells were injected into the tail vein of nude mice and tumor growth followed by non-invasive whole body bioluminescence imaging (BLI). (A) BLI image of mice at 9 weeks time point. (B) Luminiscence signal during the course of the experiment. The graph represents the luciferase signal (total flux) in photons/second. (C) The weight of the lungs was determined at the end of the experiment, the boxplot represents the distribution of volume in each group. The differences between controls and EFNA3-expressing tumors was significant (independent samples t test: $t_{12}=-2.696$, $p<0.05$).

A

pLOC

EFNA3

Total Flux (P/s)

Counts

Color Bar
Min = 10000
Max = 60000

B

Total Flux (p/s)

weeks   2   3   4   5   6   7   8   9   10

pLOC / pEFNA3

C

Lung weight (mg)

* p=0,0357

pLOC   EFNA3

Methods
Deglycosylation
25µg lysates were mixed with 1µl 10X deglycosylation buffer (5%SDS, 0.4M DTT) and H20 to a 10µl total reaction volume. The glycoprotein was denature at 100ºC for 10 minutes and incubated for additional hour at 37ºC with 1µl PNGaseF (Biolabs, P0705S), in a total volume adjusted to 20µl with 2µl of 10X G7 Reaction buffer (0.5M Sodium Phosphate pH 7.5), 2µl of 10%NP40.   The deglycosylation was analyzed by westernblotting.

Rapid Amplification of cDNA Ends (RACE)
RACE was performed with the FirstChoice RLM-RACE Kit (Ambion) following manufacturer's instructions and using the indicated specific primers. All PCR fragments of interest were cloned into pCR2.1 vector (Invitrogen) and sequenced.

References

Supplementary Table I
TaqMan probes (Applied Biosystems) used in this study

Probe Name	Reference
TaqMan 1+2	Hs00191913_m1
TaqMan4+5	Hs00900213_g1
EPAS1	Hs01026149_m1
HIF1α	Hs00936368_m1
Actin	4326315E

Supplementary Table II
Oligonucleotides used in this study

Primer Name	Assay	Forward sequence 5'-3'	Reverse sequence 3'-5'
H1	ChiP HIF1α	CTTCTCCTTCCCCTCATTCC	CAGGCTGAATTTCCCAGAAG
H2	ChiP HIF1α	CGAGCACCTCCCTAGAAGGA	CTCACCCTAGCCGGGCTTAT
H3	ChiP HIF1α	TCCCCCCACCAAACCTATTC	TTGGAGCACCTCAGCGTTCT
H4	ChiP HIF1α	GAGGGTCTCTGCCCCTTG	CCCCAGGTCTGTCAAAGGAG
H5	ChiP HIF1α	GATGGCGAGGATTTGACAAGTT	AAGACCAGGGAGTAGGGAAAGG
H6	ChiP HIF1α	GGAAGCTCGGAGGAAAAGTC	GAAGCGGCTCAGAGAGAAGA

EGLN3_E	ChiP HIF1α	GGTGTGCTCGGGTGTG	CGTGGAGGACTGGCTCTAAG
EGLN3_P	ChiP HIF1α	ACGGGAGGCACTCGGAG	CCCTTAACGTTGACTTTCGCTC
Prm1	ChiP Pol II	CATGCGGTGTACTGGAACAG	GGATCCCCCAGGACTTCTC
Prm2	ChiP Pol II	GGAAGCTCGGAGGAAAAGTC	GAAGCGGCTCAGAGAGAAGA
P4HA	ChiP Pol II	GAGCCCGTTAGCCCTTTTAT	GGTGTGATCGAGCTCACGTA
STT3	ChiP Pol II	GAGCGCGGAAAGAACGTG	FGCAAGGGCCTATTTCAGCGTA
Exon 1+2	qPCR	CACTCTCCCCCAGTTCACCAT	CGCTGATGCTCTTCTCAAGCT
Exon 4+5	qPCR	AACCGGCATGCGGTGTA	ATCCAGATAGTCGTTCACGTTCA
Coding EFNA3 (mouse)	qPCR	TGGAACAGCTCCAATCAGCA	GAGCTGTTGTAGTGCGGACA
Non-coding (mouse)	qPCR	AACCGGCATGCGGTGTA	ATCCAGATAGTCGTTCACGTTCA
hnRNA1	qPCR	ACTGGAACAGCTCCAACCAG	GCTCTGGTCTCCCAGGACTC
hnRNA5	qPCR	GGCACTGATACTTCCTACCCTG	GGCACCTGAGGGTTCTCTCC
bACTIN hnRNA	qPCR	CCCAGCACAATGAAGATCAA	GTGAGGACCCTGGATGTGAC
Exon 1	qPCR	AGCTGGGAAGCGGAGAAG	CTGGTTGGAGCTGTTCCAGT
Exon 2.1	qPCR	CTGCGGCGAGAGGGCTA	TGTAGTGCGGGCAGTAAATATCC
Exon 2.2	qPCR	AAGCGCTGGGAGTGCAAC	GTGGAACTCGTAGCCCAGAG
Exon 4	qPCR	CATCGCACTCCGGGGAGAAG	CACGTTGATCTTCACATTGGGG
Exon 5	qPCR	GCTTGAGAAGAGCATCAGCG	CTAGGAGGCCAAGAACGTCA
3UTR1	qPCR	TCCTCCCATGGCTAGAAGTG	GTCCACTACAGTGCCCTACG
3UTR2	qPCR	GCAATAAGCACGTCCTCCTC	CAGGGGGTTAAAGAGGGAAG
Prm1	Cloning	CCTTGGTACCCCGAGGGGGCAGTACGG	CCTTCTCGAGCCCCGGAGCCGCC

			GCCGCCG
Prm2	Cloning	GAGAAGCGATGGAGGGTGT	CACTCAGACTCTCGCGTTCA
NC 1	Cloning	CGGACTAGTATGGTGAGCCGCAACGGCTAC	CGGGGCGCGCCGACTAAGAGCAGAGTATGAAAGTC
NC 2	Cloning	CGGACTAGTGAGACCAGGGAGGAGGCGTG	CGGGGCGCGCCGACTAAGAGCAGAGTATGAAAGTC
11 HRE_mut	Mutagenesis	ACACCGTGCAGGTGAAACGACTATCTGGAT	ATCCAGATAGTCGTTTCACCTGCACGGTGT
12 HRE_mut	Mutagenesis	GCGGGGCAGAGCAGTCTGTACATGGTGAGC	GCTCACCATGTACAGACTGCTCTGCCCCGC

A role for insulator elements in the regulation of gene expression response to hypoxia

Maria Tiana[1,2], Diego Villar[1], Eva Pérez-Guijarro[1], Laura Gómez-Maldonado[1], Eduardo Moltó[2,3], Ana Fernández-Miñán[4], Jose Luis Gómez-Skarmeta[4], Lluís Montoliu[2,3] and Luis del Peso[1,5,*]

[1]Departamento de Bioquímica, Universidad Autónoma de Madrid and Instituto de Investigaciones Biomedicas Alberto Sols, CSIC-UAM, 28029 Madrid [2]Department of Molecular and Cellular Biology, Centro Nacional de Biotecnología (CNB-CSIC), Campus de Cantoblanco, Darwin 3, 28049 Madrid [3]Centro de Investigación Biomédica en Red de Enfermedades Raras (CIBERER), ISCIII, 28049 Madrid [4]Centro Andaluz de Biología del Desarrollo (CABD) CSIC-UPO, 41013 Sevilla and [5]Department of Systemic Pathology, Hospital La Paz/Autónoma University School of Medicine, IdiPAZ, 28029 Madrid, Spain

ABSTRACT

Hypoxia inducible factor (HIF) up-regulates the transcription of a few hundred genes required for the adaptation to hypoxia. This restricted set of targets is in sharp contrast with the widespread distribution of the HIF binding motif throughout the genome. Here, we investigated the transcriptional response of GYS1 and RUVBL2 genes to hypoxia to understand the mechanisms that restrict HIF activity toward specific genes. GYS1 and RUVBL2 genes are encoded by opposite DNA strands and separated by a short intergenic region (∼1 kb) that contains a functional hypoxia response element equidistant to both genes. However, hypoxia induced the expression of GYS1 gene only. Analysis of the transcriptional response of chimeric constructs derived from the intergenic region revealed an inhibitory sequence whose deletion allowed RUVBL2 induction by HIF. Enhancer blocking assays, performed in cell culture and transgenic zebrafish, confirmed the existence of an insulator element within this inhibitory region that could explain the differential regulation of GYS1 and RUVBL2 by hypoxia. Hence, in this model, the selective response to HIF is achieved with the aid of insulator elements. This is the first report suggesting a role for insulators in the regulation of differential gene expression in response to environmental signals.

INTRODUCTION

A large number of biochemical reactions require oxygen as a substrate and metazoa metabolism is largely dependent on oxidative phosphorylation. At the cellular level, the unbalance between oxygen demand and supply (hypoxia) results in the activation of a specific gene expression program aimed at increasing oxygen delivery and reducing its consumption through metabolic reprogramming. This transcriptional response is mostly mediated by an evolutionarily conserved family of transcription factors termed hypoxia inducible factors (HIFs), that belong to the basic helix-loop-helix superfamily (1). HIFs are heterodimers of a constitutive beta subunit (HIFβ also known as ARNT), that partners with several factors and an alpha subunit (HIFα), whose stability (2) and transcriptional activity (3) is regulated by oxygen. Under hypoxia, HIFα subunits avoid degradation, bind to the constitutively expressed beta subunits and the heterodimers translocate to the nucleus where they bind to the RCGTG motif within the regulatory regions of target genes to promote their transcription (4–6). Several works have identified individual HIF targets that, taken together, account for the metabolic adaptation and induction of angiogenesis observed under hypoxia (7). To gain insight into the full range of cellular adaptations to hypoxia, several groups recently attempted the global identification of HIF-targets (5,6,8–11). Interestingly, all these works coincide in that only a few hundred, out of all the genes containing RCGTG motifs, are regulated by hypoxia. Thus, as it is the case for other transcription factors (12), HIF binds only to a small proportion of the

potential binding sites (5,6,8–11). The basis for this selectivity is incompletely understood, but several mechanisms have been proposed. Among them, the cooperation with other transcription factors, have been well characterized in some instances (13,14). In the case of HIF, requirement of functional HNF-4 (15), AP-1 (16), GATA-2 (16) or ETS (17,18) sites for proper hypoxic induction of selected targets have been described. In agreement with these single locus studies, global analysis of HIF binding sites by means of experimental (5) and computational methods (11) showed the existence of overrepresented transcription factor binding sites (TFBS) in close proximity to the hypoxia response element (HRE) that might account for factors cooperating with HIF. However, the experimental characterization of the role of these TFBS in the regulation of HIF targets by hypoxia is yet to be determined. Thus, the cooperation between HIF and other factors could contribute to the target selectivity, but it is yet unclear to what extent this mechanism explains the observed pattern of targets.

A further mechanism that could dictate the target selectivity is the accessibility to the TFBS. Histone modifications alter the structure of chromatin and hence the availability of the underlying nucleotide sequence for the binding of transcription factors (19). In addition, DNA methylation can preclude the binding of specific transcription factors (20–22). In this regard, a recent study addressed the cell-type specificity in response to hypoxia and concluded that only those loci that were transcriptionally active under basal (normoxic) conditions were permissive to HIF-regulation (8). However, although these results explain most of the intercellular variation in the hypoxic transcriptome, it is clear that an additional layer of regulation is required, as only a small fraction of all the active genes under basal conditions were induced by hypoxia in any of the cell lines studied.

Finally, insulators are included among the regulatory mechanisms employed by eukaryotes to ensure specific patterns of gene expression and as such, they could be involved in the selection of genes to be activated by HIF in response to hypoxia. Insulators are defined as DNA elements that partition chromatin into independent transcriptional domains, thereby contributing, in combination with additional epigenetic mechanisms, to the tight control of gene expression (23) and to the nuclear structure and dynamic organization (24). These elements have been functionally described, according to their ability to block the spread of heterochromatin (barrier function) into adjacent loci and to prevent the promiscuous interaction of distal enhancers with proximal promoters, when placed in between (enhancer blocking function) (25). The role of insulators in the determination of linage-specific patterns of gene expression is well characterized, as illustrated in the chicken β-globin locus (26). However, the participation of these elements in the target discrimination by acutely activated transcription factors, in response to environmental factors, has not been reported.

Here, we investigate the mechanism that restricts HIF activity toward specific genes by the study of a locus that we consider paradigmatic, the *GYS1/RUVBL2* genomic region from the human genome. The *GYS1* gene was recently described as a novel hypoxia-inducible gene and a functional HRE was identified upstream its promoter (27). Interestingly, very close to *GYS1*, but encoded by the opposite DNA strand, is located the *RUVBL2* gene. In spite of the location of the HRE between both genes, we found that only *GYS1*, but not *RUVBL2*, was induced in response to hypoxia. The lack of RUVBL2 response to hypoxia was not due to epigenetic silencing of its promoter as it showed a substantial transcriptional activity and the level of RUVBL2 mRNA was comparable with that of GYS1. Instead, the analysis of different reporter constructs derived from the intergenic GYS1/RUVBL2 sequence, revealed an inhibitory region, located between the HRE and the RUVBL2 minimal promoter, that prevented the induction of RUVBL2 by HIF. Removal of this region allowed the up-regulation of the RUVBL2 promoter upon HRE activation. This result hinted the existence of an enhancer blocking element within the inhibitory region that prevented the interaction between the HRE and the RUVBL2 promoter. By means of specific enhancer blocking assays (EBA), performed in cell culture and using heterologous constructs in transgenic zebrafish, we confirmed the existence of an insulator element within this locus that could explain the differential regulation of GYS1 and RUVBL2 by hypoxia. Altogether, our results suggest that HIF selectivity is achieved, at least in this locus, by an insulator element that prevents the activity of the HRE/HIF complex on the *RUVBL2* promoter.

MATERIALS AND METHODS

Cell culture and reagents

Human cervical-carcinoma cells (HeLa) and Human embryonic kidney 293 (HEK293) cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% (v/v) fetal bovine serum and 100 U/ml penicillin and 100 mg/ml streptomycin and cultured at 37°C and 5% CO_2 in a humidified incubator.

For hypoxia treatments, cells were grown at 37°C in sealed chambers (Billups–Rothenberg) flushed with a 1% O_2, 5% CO_2, 94% N_2 gas mixture or in a Whitley hypoxystation (don Whitley Scientific, UK) set at 1% oxygen concentration.

Dimethyloxalylglycine (DMOG, Frontier Scientific, CA, USA) was added to the indicated cultures at a 500 μM final concentration. For the analysis of RUVBL2 and GYS1 expression, cDNA obtained from the following cell lines were also used: HepG2, HepaC1, NIH/3T3; N2a, HEK293 and A549.

Analysis of gene profiling datasets

The expression profiles corresponding to the indicated datasets and series were downloaded from Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) (28) database. In all the cases, untreated normoxic cells were used as reference. For each data set we calculated the mean of probe values in the biological replicates. All probes mapping to the locus of interest, except those with null values, were included in the analysis. Then, for each probe, the effect of hypoxia was

calculated as the logarithm of the ratio of the means of treated and control samples. Finally, individual log-ratio values were normalized by subtraction of the mean of all the log-ratios across the data set and division by their standard deviation. In the case multiple probes mapping to a given gene locus, the average of the log ratio was calculated. Information regarding GEO and probes ID can be found in the Supplementary Table S1.

RNA extraction and quantitative PCR

Total RNA was extracted and purified with the RNeasy Mini Kit (Qiagen). One microgram of RNA from each sample was reverse-transcribed to cDNA (Improm-II reverse transcriptase; Promega) and 1μl of cDNA samples were used as template for amplification reactions carried out with the LC Fast Start DNA master SYBR Green I kit (Roche Applied Science), following the manufacturer's instructions. PCR amplifications were carried out in a Light Cycler System (Roche Applied Science), and data were analyzed with LightCycler software 3 version 3.5.28 (Idaho Technology Inc.). For each sample, duplicate determinations were made, and the gene expression determined by the $\Delta\Delta$Ct method using β-actin as reference gene. The primers used in this study are shown in the Supplementary Table S2.

Plasmid construction

Human genomic DNA extracted from HeLa cells was used as template for PCR amplification of *GYS1* and *RUVBL2* promoter regions. For reporter assays, PCR products were first cloned into pCR2.1-TOPO (invitrogen) and subsequently subcloned into the KpnI/XhoI restriction sites of pGL3-Basic.

For the *in vitro* EBA, the putative insulators were cloned in the plasmid pELuc (29) and assayed their activity by transfection of the resulting constructs into HEK293 cells. This vector carries a CMV enhancer and a minimal promoter controlling the firefly luciferase report gene cassette with the polyadenylation site from SV40. All inserts were cloned into the XhoI site, between the enhancer and the CMV promoter (IN-position), and into SmaI, upstream of the enhancer (OUT-position). The insertion in SmaI site is a control for the assay that tests the potential silencer/repressor effects of our insert. For the *in vivo* EBA the putative insulator elements were cloned in the vector 48RCar (30) and injected resulting constructs into zebrafish embryos to assess their effect on transcription. In this plasmid the GFP gene is carried in the vector under the control of an actin promoter, that directs expression in heart and muscle. The insulator is cloned in the KpnI site, between the actin promoter and the Enhancer 48, which targets the expression to the central nervous system (31).

The identity of all constructs was verified by sequencing. All primer sequences are available in the Supplementary Table S1.

Reporter assays

Reporter assays were performed using the HeLa. Cells were seeded on six-well plates (3×10^5 cells/well) 6 h

prior to transfection. A 9 μg DNA mixture containing 3 μg of the indicated reporter construct or empty plasmid and 0.6 μg of a plasmid encoding for renilla (sea pansy) luciferase under the control of a null promoter (Promega, Madison, WI, USA.) was used for transfection using the calcium phosphate method. On completing 16 h after transfection, cells where washed, replated in 24-well plates and incubated in normoxia, in the presence of DMOG or under hypoxia for additional 16 h. After treatments, the cells were lysed and the firefly and renilla luciferase activities of the lysate were determined using a dual-luciferase system (Promega, Madison, WI, USA). The firefly luciferase activity was normalized to that of renilla luciferase.

In vitro EBA

Cells were transiently transfected using Lipofectamine 2000 (invitrogen) and OPTI-MEM® medium (Invitrogen) according to the supplier's instructions. Briefly, 1.8×10^5 HEK293 cells were seeded the day before transfection in 24-well plates. For each well, 0.66 μg of the linearized reporter vector was transfected together with 0.14μg of pCMV-lacZ control plasmid (for normalization purposes). The pELuc-derived constructs were linearized prior to transfection, to avoid bidirectional enhancer activity, using the restiction enzyme Asp-718. This site is located 3′ downstream from the polyadenylation signal. The pCMV-lacZ plasmid was also linearized using ScaI. The cells were incubated 24 h with the transfection mixture and were thereafter lysed with Reporter Lysis Buffer (Promega).

The luciferase activity was determinad using the Luciferase Assay Reagent (Promega) according to the manufacturer's specifications in a microplate luminometer (Orion, Berthold Detection Systems). The sample luciferase activity was corrected by the β-galactosidase activity in and the number of molecules (picomol) of the transfected plasmid construct (according to each plasmid size). Finally, activities were normalized as a fraction of the mean luciferase values obtained for the empty (pEluc) plasmid.

Transgenesis and *in vivo* EBA

Transgenesis and *in vivo* EBA were performed as described (30,31). For zebrafish transgenesis, the Tol2 transposon/transposase method of transgenesis (32) was used with minor modifications. In total, 1 nl was injected in the cell of one-cell stage embryos containing 50 ng/μl of transposase mRNA, 40 ng/μl of phenol/chloroform purified DNA and 0.05% phenol red.

Statistical analysis of data

Statistical analysis of the experimental data was performed with the R software package [(33), http://www.R-project.org/]. The statistical tests applied to each data set are indicated in the figure legends. We adopted the following code to indicate the magnitude of *P*-values throughout the manuscript figures as: ***P = [0, 0.001]; **P = [0.001, 0.01] and*P = [0.01, 0.05].

RESULTS

Differential regulation of *GYS1* and *RUVBL2* by hypoxia

The muscle glycogen synthase gene, *GYS1*, is regulated by HIF as part of the cellular metabolic reprogramming required for the adaptation to hypoxia (27). The regulation of human *GYS1* by hypoxia is mediated by a functional RCGTG element located 255 bp upstream its transcription start site (TSS) (27). Very close to *GYS1*, but encoded by the opposite DNA strand, is located the *RUVBL2* gene (Figure 1A). The relative position, intergenic distance and orientation of these two genes are conserved across mammals (data not shown). The TSS of *RUVBL2* is located at 288 bp of the HRE driving *GYS1* expression in response to hypoxia (location of the HRE is shown by a black box in the 'blat' track, Figure 1A), raising the possibility of a coordinated regulation of these two genes by HIF. In fact, the intergenic region between *GYS1* and *RUVBL2* can be considered a bidirectional promoter (Figure 1A, 'Elnitski bidirectional promoters' prediction track). To study this possibility, we analyzed publicly available gene expression profiles of cells exposed to hypoxia and found that, whereas *GYS1* mRNA levels were induced by hypoxia in most of the profiles, the expression of *RUVBL2* remained constant or was even repressed under low oxygen tension (Figure 1B). To confirm these results we exposed myotubes to hypoxia and determined its effect on *GYS1* and *RUVBL2* mRNA levels. As shown in Figure 1C, *GYS1* mRNA level significantly increased in response to hypoxia, in agreement with published results (27), but the treatment did not induce *RUVBL2* mRNA. In order to rule out a cell-type specific effect, we determined the effect of hypoxia on the expression of *GYS1* and *RUVBL2* in a variety of cell types (Figure 1D). These analyses confirmed that, whereas *GYS1* expression was increased by hypoxia in virtually all cell lines studied, *RUVBL2* remained largely unaffected. In agreement with this conclusion, a meta-analysis of gene expression profile experiments suggested that, unlike *GYS1*, *RUVBL2* is not significantly modulated by hypoxia (10). These results indicated that hypoxia specifically affects *GYS1*, but not *RUVBL2* transcription. To confirm this possibility, we investigated the effect of hypoxia on a set of reporter constructs derived from this locus. As shown in Figure 1E, a reporter construct containing the region flanking *GYS1* gene and including the HRE (region cG spanning residues +84 to −429 relative to *GYS1* TSS, Figure 1A), was robustly induced by hypoxia and the hypoxia mimetic DMOG in HeLa cells, as expected from the presence of the evolutionarily conserved block containing the HRE within this cG region. In contrast, the analogous region upstream of *RUVBL2* (cR, spanning −396 to +12 relative to *RUVBL2* TSS) did not respond to HIF activation in spite of including the same HRE-containing region (Figure 1E). Importantly, this same result was obtained when the whole intergenic region, maintaining the original genomic structure, was used to drive luciferase expression from the GYS1 (cRcG) or the RUVBL2 (cGcR) promoters (Figure 1E).

Collectively, these results indicate that the HRE located between *GYS1* and *RUVBL2* genes selectively drives the transcription of the former in response to hypoxia.

The lack of *RUVBL2* induction by hypoxia is not due to gene silencing

The accessibility of promoter regions is one that the mechanisms that restricts the activity of a TF toward specific genes. This can be achieved by methylation of promoter regions and/or by altering chromatin compactness through histone modification. In fact, it has been recently found that basal promoter activity determines cell type-specific HIF transcription (8). Thus, we investigated whether a lack of *RUVBL2* transcriptional activity could explain the observed selectivity of hypoxia within this locus. As shown in Figure 2A, the normoxic levels of *RUVBL2* and *GYS1* mRNAs were of similar magnitude, at least for the set of cell lines included in our study. Moreover, the differential effect of hypoxia (Figure 1D) was observed even for cell lines, such as HeLa or A549, in which the relative basal level of *RUVBL2* mRNA was much higher than that of *GYS1*. On the other hand, the genomic region adjacent to the *RUVBL2* gene (cR) showed a strong basal promoter activity when assayed in HeLa cells (Figure 2B). In fact, the promoter activity of the cR region was significantly higher than that of the cG region (Figure 2B). In agreement with our results, published data of genome-wide RNA polymerase II binding shows a similar signal in the *GYS1* and *RUVBL2* promoter regions in a wide range of cell types (Supplementary Figure S1).

Altogether, these results indicate that *RUVBL2* and *GYS1* are transcribed under normoxia to a similar extent and thus, we discarded *RUVBL2* promoter accessibility as a potential explanation for the differential regulation observed under hypoxia.

An inhibitory region prevents *RUVBL2* induction by the HRE

To gain insight into the molecular mechanism responsible for the differential regulation of *RUVBL2* and *GYS1* by hypoxia, we generated a set of reporter constructs containing different deletions and rearrangements of the genomic region between *GYS1* and *RUVBL2* genes (Figure 3A and diagrams on the left of Figure 3B). Based on the evolutionary conservation (Phas Cons elements, see Figure 1A), we differentiated five blocks within this region (Figures 1A and 3): proximal *GYS1* (pG), upstream *GYS1* (uG), HRE-containing block (HRE), upstream *RUVBL2* (uR) and proximal *RUVBL2* (pR). As shown in Figure 3, analysis of the transcriptional activity of these reporter constructs in HeLa cells showed that neither pG nor pR proximal regions responded to HIF activation. However, combination of these proximal regions with the HRE block, regardless of the orientation of the latter, resulted in constructs (HREF_pG, HRER_pG, HREF_pR and HRER_pR) that were robustly induced by the hypoxia mimetic DMOG. The induction of these constructs was of similar magnitude to that observed for the complete *GYS1* construct (cG), suggesting that upstream regions
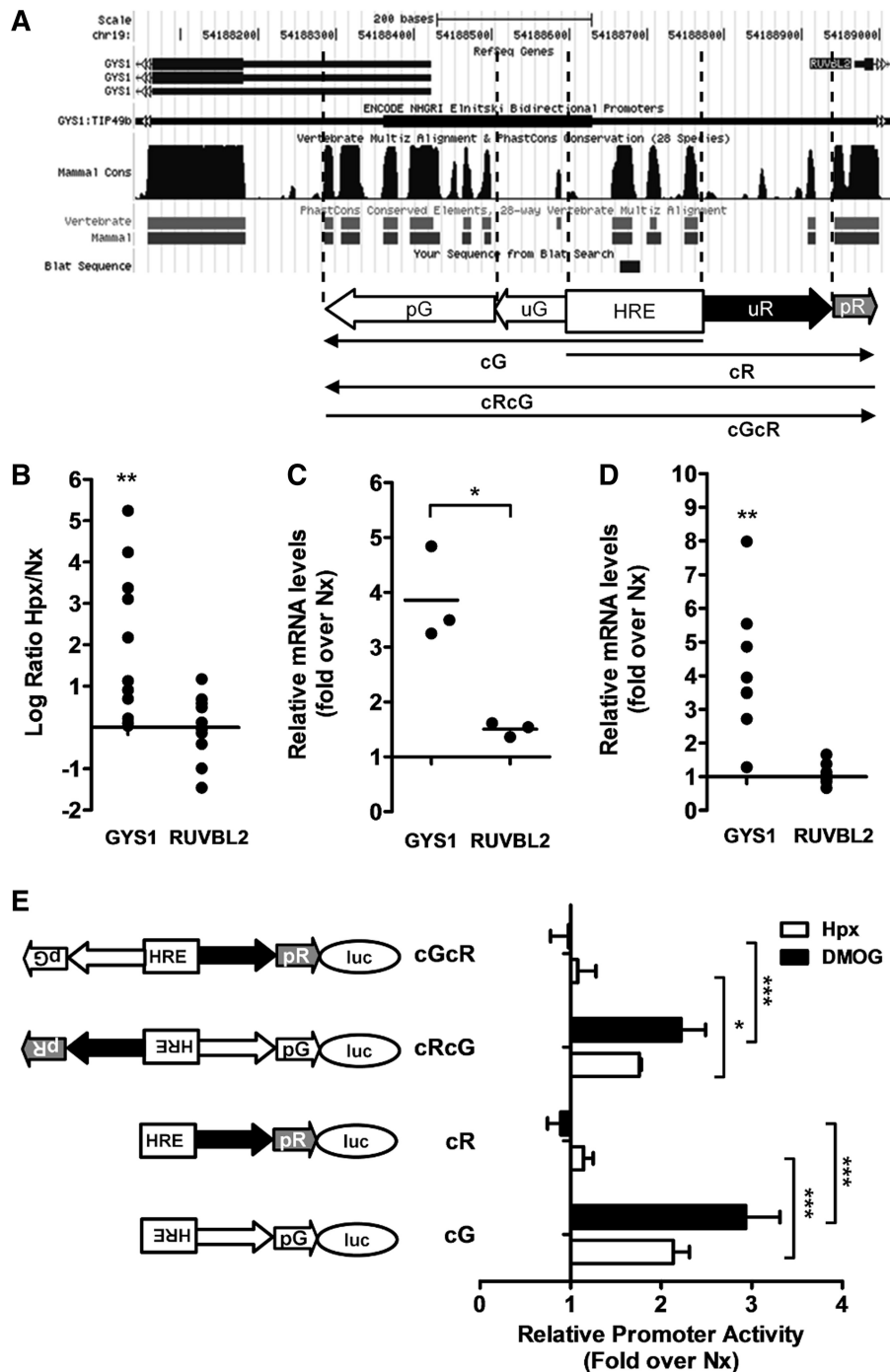
Figure 1. Differential regulation of *GYS1* and *RUVBL2* by hypoxia. (**A**) Schematic diagram depicting the human (hg18 assembly) genomic region containing the intergenic region between *GYS1* and *RUVBL2* and its sequence conservation across mammals [adapted from UCSC genomic browser, http://genome.ucsc.edu/(53)]. The boxes below the diagram represent the different blocks identified within this region according to their evolutionarily conservation and solid lines indicate the regions cloned to generate reporter constructs, cG, cR and cG+cR. (**B**) Effect of hypoxia or the hypoxia-mimetic deferoxamine (GSE5579) on *GYS1* and *RUVBL2* expression extracted from gene expression profiles of human breast carcinoma cell line MCF7 (GSE3188), mouse embryo fibroblast (GSE3196), human B lymphocyte P493-6 cells (GSE4086), human monocyte-derived macrophages (GDS2036), human lymphatic endothelial cells (GSE5579), human aortic smooth muscle cells (GSE4725), human colon adenocarcinoma cell line HT29 cells (GSE9234), mouse hepatocytes (GDS1648), human embryonic kidney cell line HEK293 (GSE2020), human astrocytes (GSE3045) and human epithelial cervical cancer cell line HeLa (GSE3051) exposed to hypoxia. Asterisks indicate that the set of data values was significantly different (one sample *t*-test, $t = 3.6988$, df = 11, $P = 0.003509$) from the value of zero (no induction). (**C**) c2c12 myoblast were exposed to normoxia or hypoxia for 12–24 h and *GYS1* and *RUVBL2* expression was determined by quantitative PCR from total RNA samples. Data were calculated relative to β-actin and expressed as fold change relative to normoxia. Data shown are the results of three independent experiments and their mean (bar). The relative induction of both mRNAs was significantly different (*t*-test, $t = 4.9995$, df = 2, $P = 0.03776$). (**D**) A variety of cell lines (HepG2 and HepaC1, mouse hepatocarcinoma cell lines; primary mouse hepatocytes; NIH3T3, mouse fibroblast cell line; HeLa; N2a, mouse neuroblastoma cell line; HEK293; A549, human lung adenocarcinoma cell line) were exposed to normoxia or hypoxia and the levels of *GYS1* and *RUVBL2* mRNA determined and represented as indicated in C. Asterisks indicate that the set of data values was significantly different (one sample *t*-test, $t = 4.4522$,
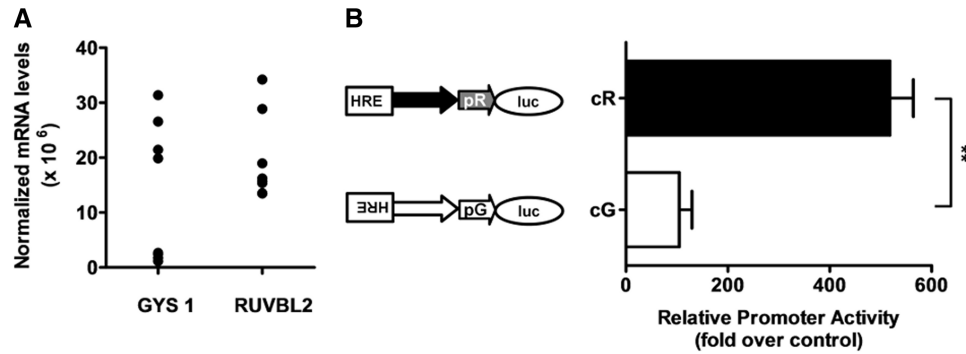
Figure 2. *RUVBL2* is transcribed under normoxia. (**A**) The basal (normoxic) expression of *GYS1* and *RUVBL2* was determined by quantitative PCR for the cell lines indicated in Figure 1D. The graph represents the normalized mRNA levels for each gene across the different cell lines. The difference between both groups was not statistically significant (paired *t*-test, $t = -0.9365$, df = 7, $P = 0.3802$) (**B**) HeLa cells were transfected with a reporter construct containing the indicated region upstream a firefly luciferase gene or the corresponding empty plasmid lacking insert (pGL3basic). The graph represents the corrected luciferase activity values obtained for each plasmid as fold over the activity contained in cells transfected with empty vector. Bars represent the average of values obtained in three independent experiments and errors bars their standard deviation. The difference between the means of both groups was statistically significant (paired *t*-test, $t = -10.2002$, df = 2, $P = 0.009475$).

(uG and uR) are not required for the activity of the HRE-containing block. These data ruled out a collaboration between HIF and other transcription factors binding to pG or uG as a potential explanation for the differential induction of *GYS1* and *RUVBL2* by hypoxia. On the other hand, comparison of response of the cR and HREF_pR constructs, hinted the existence of an inhibitory region within the uR region as removal of this region significantly increased the response to HIF activation. In agreement, the up-regulation of the HREF_pG construct was blunted by the inclusion of the upstream region of *RUVBL2* between the HRE and proximal *GYS1* block (Figure 3, HREF_pG and HREF_uR_pG).

In summary, these results suggest that the genomic region from −202 to −30, relative to the *RUVBL2* TSS, (uR) prevents the induction of *RUVBL2* by HIF.

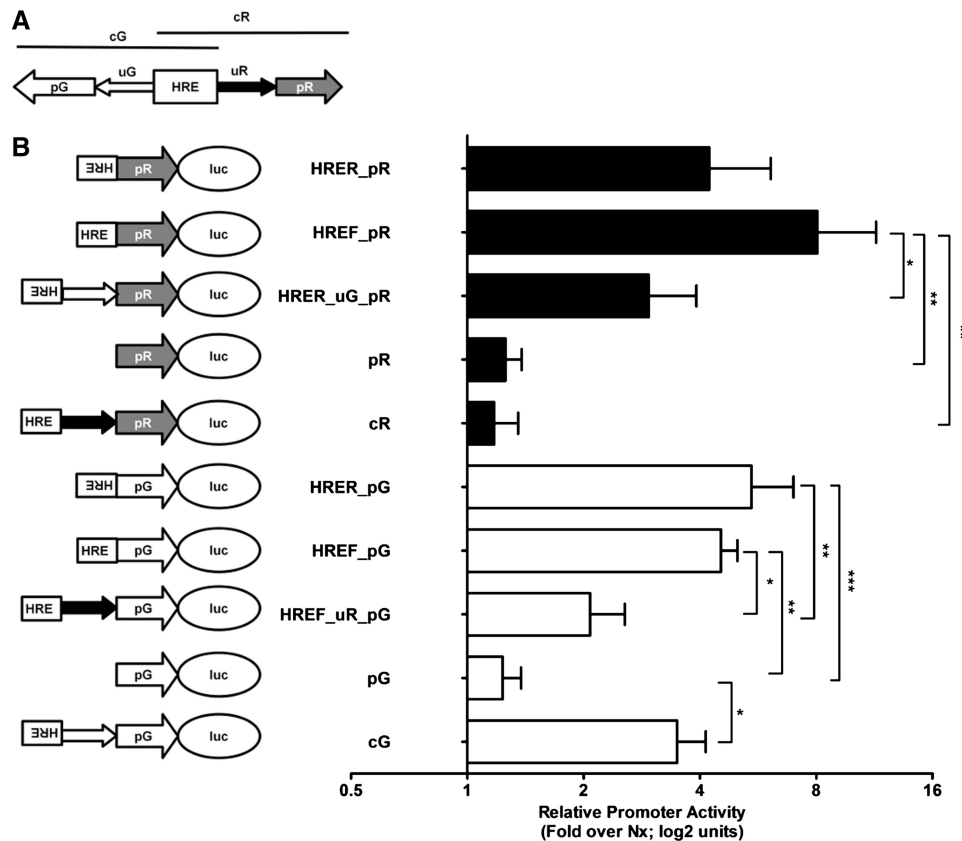### The *GYS1/RUVBL2* intergenic region contains an enhancer blocking element

Insulators are DNA elements that can prevent the promiscuous effect of enhancers or silencers, restricting their interactions with promoters. Our results suggested that the uR region prevents the interaction of the HRE over the *RUVBL2* promoter, a function that is compatible with the enhancer blocking type activity commonly associated to insulators. To test this possibility, we studied the ability of different sequences, derived from the *RUVBL2/GYS1* intergenic region, to interfere with the activity of a heterologous enhancer/promoter pair in a standard EBA, in HEK293 cells (29). Each of the genomic fragments under study was cloned either, between the enhancer and promoter (IN position) or

upstream and the enhancer (OUT position) (Figure 4A). As a positive control, we also included the boxes II/III (E II/III) derived from the classical chicken 5′HS4 β-globin insulator element (34), known to bind the nuclear factor CTCF and responsible for the enhancer blocking effect of the 5′HS4 element (35). The enhancer blocking activity of these sequences, represented as fold reduction of the activity of a vector lacking insert, is represented in Figure 4B. In agreement with previous reports (35), the E II/III boxes, but not a mutant form, interfered with the activity of the CMV enhancer when inserted into the XhoI site (Figure 4B, white bars). Enhancer blockers work only in the 'IN' configuration as demonstrated by the lack of activity of the EII/III constructs in the 'OUT' configuration (Figure 4B, black bars). Significantly, all the constructs containing the uR in the 'IN' configuration, except 1R, showed a significant reduction in the enhancer activity (Figure 4B, white bars corresponding to constructs 1F, 2F, 2R, 3F and 3R). Importantly, these same regions had no significant inhibitory effect when cloned in the 'OUT' configuration (Figure 4B, black bars corresponding to constructs 1F, 2F, 2R, 3F and 3R). In contrast to these results, neither the HRE nor uG regions were able to suppress the activity of the CMV enhancer (Figure 4B, constructs 4F, 4R, 5F, 5R, 6F, 6R and Supplementary Figure S2).

The EBA results strongly support the existence of an enhancer blocking element within the genomic region from −202 to −30 relative to the *RUVBL2* TSS (Figure 4B, constructs 3F and 3R). Thus, we decided to functionally validate the presence of an insulator in this region *in vivo* by means of an independent assay using

Figure 1. Continued

df = 7, $P = 0.002964$) from the value of one (no induction). (**E**) HeLa cells were transiently transfected with reporter constructs containing the indicated region (cG, cR, cGcR or cRcG, see Figure 1A) upstream a firefly luciferase gene and exposed to normoxia, hypoxia or the hypoxia mimetic DMOG for 12 h. The graphs represent the corrected luciferase activity values of each construct in cells exposed to hypoxia/DMOG and represented as fold change over the activity in normoxic cells. Bars represent the average of values obtained in three independent experiments and error bars their standard deviation. Statistically significant differences between pairs of constructs are indicated by asterisks (one-way ANOVA, $F_{9,20} = 36.704$, $P = 1.6 \times 10^{-10}$, followed by Tukey's multiple comparison test).

Figure 3. Region uR prevents the increase in transcription mediated by the HRE. (**A**) Schematic diagram of the *GYS1/RUVBL2* intergenic showing the different elements included in the reporter constructs. (**B**) HeLa cells were transfected with the indicated constructs and treated with the hypoxia mimetic DMOG or left untreated for 12 h. The graph represents the corrected luciferase activity values of each construct in treated cells represented as fold change over the activity obtained in control (normoxic) cells. Bars represent the average of values obtained in three independent experiments and error bars, their standard deviation. The diagrams on the left are a schematic representation of the genomic elements included in the reporter construct and their orientation relative to the luciferase gene. Statistically significant differences between pairs of constructs within the groups containing pG or pR are indicated by asterisks (one-way ANOVA followed by Tukey's multiple comparison test. pG-containing constructs $F_{4,10} = 13.608$, $P = 0.0004705$; pR-containing constructs $F_{4,10} = 7.518$, $P = 0.004598$).

transgenic zebrafish. To this end, we used a reporter construct in which EGFP expression is under the control of the cardiac actin promoter from *Xenopus laevis* and the Z48 enhancer, which drives transgene expression in the midbrain (Figure 5A) (31). Transgenic zebrafish embryos injected with this construct showed EGFP expression in the heart and in the developing somites (due to the cardiac actin promoter), as well as in the CNS (Figure 5B and C, control), due to the Z48 enhancer, as reported earlier (30,31). Micro-injection of a construct containing the uR region cloned between the enhancer and the promoter, regardless of its orientation, resulted in a strong attenuation of the CNS signal, whereas retaining the EGFP signal in the developing somites (Figure 5C, uR_F and uR_R). Analysis of the somites/CNS EGFP-mediated fluorescence signal ratio in 62 independent transgenic zebrafish lines demonstrated an enhancer blocking activity associated with the uR sequence (Figure 5B, uR_F and uR_R). Importantly, in agreement with the *in vitro* EBA assays (Figure 4), the effect was specific for the uR region as cloning of the uG sequence between the Z48 enhancer and the actin promoter did not interfere with the EGFP expression in the CNS (Figure 5B and C, uG_F and uG_R).

Altogether, these results demonstrate the existence of a powerful enhancer blocking element in the uR region that could explain the differential regulation of *GYS1/RUVBL2* by HIF. The activity of this insulator element is observed when assayed in a distant heterologous system (zebrafish) suggesting that its function and potential *trans*-acting factors have been evolutionary conserved in vertebrates.

DISCUSSION

The recognition of short DNA motifs by transcription factors is a key step in the regulation of transcription. However, the low information content of most of the TFBS predicts a promiscuous binding that is in contrast with the observed specificity. This apparent paradox raises the question of how transcription factor selectivity is achieved. Although several mechanisms including chromatin accessibility and TF cooperation, have been shown to restrict the target space for a given TF, a complete explanation is still lacking in most of the cases. In this scenario, HIF transcription factors are not an exception. HIF heterodimers bind to the RCGTG motif
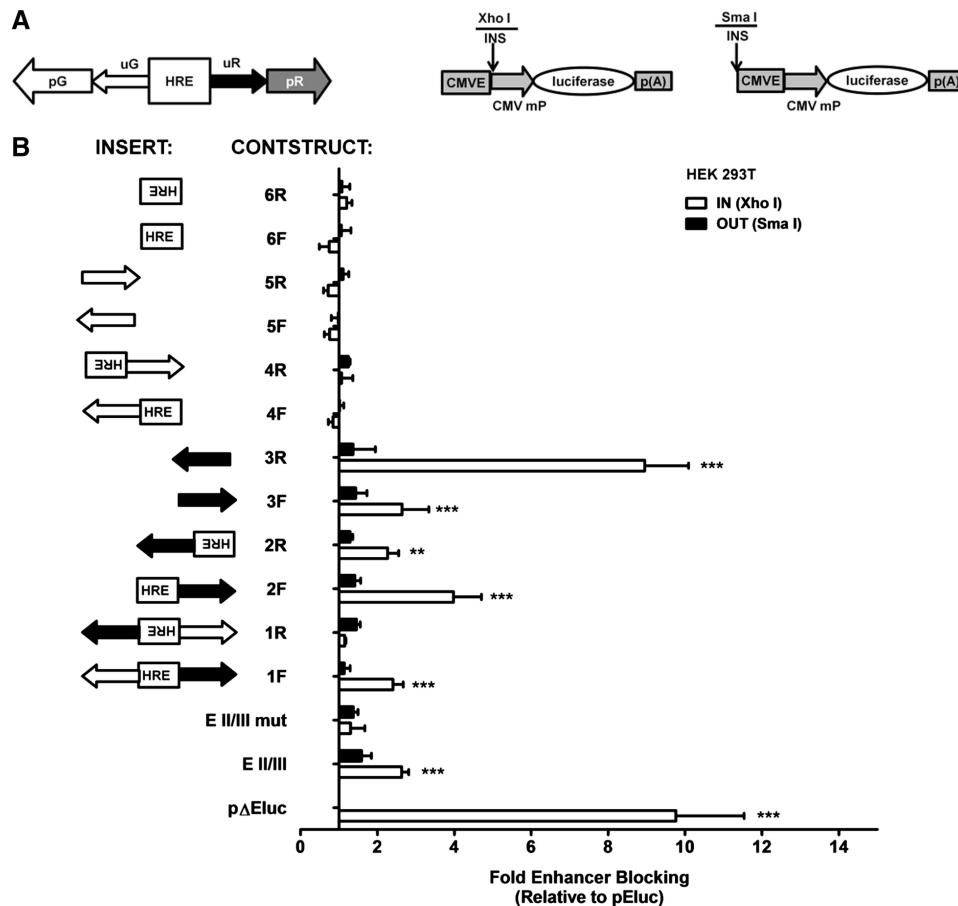
Figure 4. The uR region contains an enhancer blocking element. (**A**) Schematic diagrams of the different elements within the *GYS1/RUVBL2* intergenic region (left diagram) and the EBA vector showing the 'IN' (middle diagram) and 'OUT' (right diagram) cloning sites. (**B**) Each of the indicated elements (diagrams on the left under 'INSERT') derived from the *GYS1/RUVBL2* intergenic region, or from the β-globin gene (E II/III and a mutated version, E II/III mut), were cloned into the XhoI ('IN' position, white bars) or SmaI ('OUT' position, black bars) sites of the pELuc vector to generate the indicated constructs ('CONST.'). The HEK293 cells were transiently transfected with each one of these constructs, an empty EBA vector lacking insert (pEluc) or an empty EBA vector lacking the enhancer (pΔEluc). On completing 24 h after transfection, the cells were processed to determine the transcriptional activity of the constructs. The figure represents the corrected luciferase activity in each sample and is expressed as fold reduction of the activity observed in cells transfected with the empty vector, pEluc (29). Bars represent the average of values obtained in three independent experiments and errors bars, their standard deviation. Asterisks indicate constructs whose activity was significantly different from that observed for the control, pEluc (one-way ANOVA, $F_{29,237} = 131.10$, $P = 2.2 \times 10^{-16}$, followed by Tukey's multiple comparison test).

(4,6,9), which is present in almost every gene in the human genome, yet hypoxia results in the regulation of a few hundred genes only and, accordingly, HIF only binds to a subset of the potential binding sites (5,6,9,10). It has been recently demonstrated that HIF binds preferentially to RCGTG motifs present in the promoters of genes actively transcribed under normoxic conditions (8). Although this restriction results in a large reduction in the number of potential targets, it does not fully explain HIF selectivity as many genes transcribed under normoxia and containing RCGTG motifs are not induced by hypoxia. With the aim of gaining insight into the mechanisms of HIF target selectivity, we investigated the *GYS1/RUVBL2* genes as a particularly striking example of differential regulation. Given their close proximity and the location of the functional HRE between both genes, we had expected their coordinated regulation by HIF. However, our results showed that while *GYS1* was

induced by hypoxia, *RUVBL2* levels remained unchanged (Figure 1). The basal level of mRNA, RNA polymerase II binding and promoter activity (Figure 2 and Supplementary Figure S1) ruled out the accessibility of chromatin as an explanation for the lack of induction of *RUVBL2* by hypoxia. Detailed analysis of the promoter activity of different fragments derived from the *GYS1/RUVBL2* intergenic region revealed the existence of an inhibitory region between the HRE and the *RUVBL2* gene (Figure 3). Finally, specific EBAs performed *in vitro* (Figure 4), as well as *in vivo* (Figure 5) demonstrated the existence of an insulator element that could explain the lack of effect of the HRE over the *RUVBL2* promoter and thus, the differential regulation of *GYS1* and *RUVBL2* genes by hypoxia.

Insulators have been shown to play a key role in the differential patterns of gene expression during development and cell-lineage specification (23). However, to our
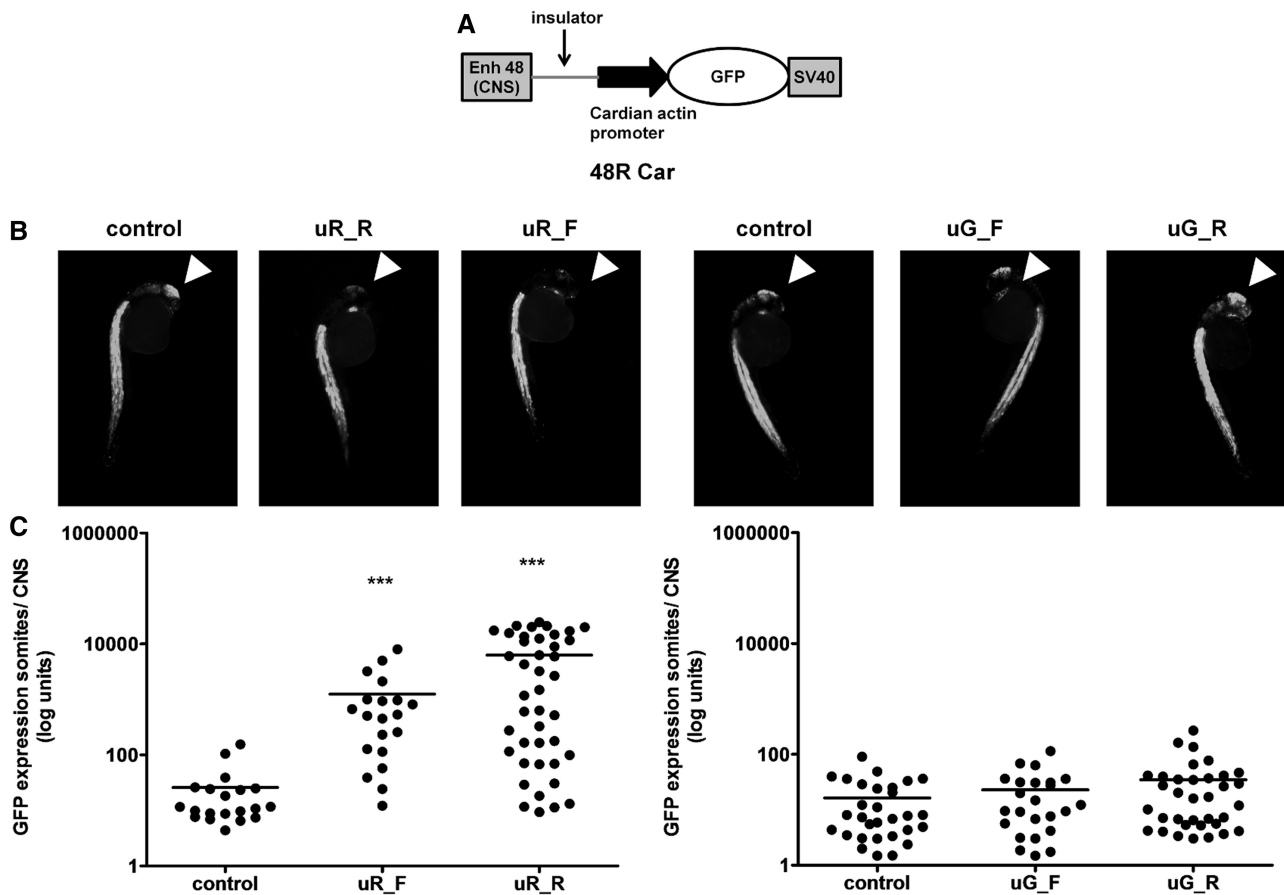
Figure 5. The insulator element derived from *RUVBL2* genomic region blocks enhancer activity *in vivo*. (**A**) Schematic representation of the ZED construct. (**B** and **C**) zebrafish embryos were micro-injected with an empty ZED vector (control), lacking insert between the enhancer and promoter, or constructs containing the *RUVBL2*-derived enhancer blocking element (uR) or the corresponding region upstream the GYS1 gene (uG). Both elements were cloned in the forward (uR_F, uG_F) or reverse (uR_R, uG_R) orientation. After 36 h after micro-injection, the GFP signal in CNS and somites was determined by fluorescence microscopy and quantified with LaserPix (Bio-Rad) image analysis software, as reported before (30). (**B**) A representative image of each group of animals is shown. Arrowheads indicate the location of the CNS. (**C**) The graph represents the ratio between the GFP signals in somites and CNS for each of the transgenic fish. Horizontal line represents the median of the ratios for each group of animals. Asterisks indicate set of values that were significantly different from those obtained for the empty construct Kruskal–Wallis, chi-squared = 89.2081, df = 5, $P < 2.2e-16$; *P*-values for comparisons were calculated by Wilcoxon rank and corrected for multiple comparison using Bonferoni's method.

knowledge, this is the first example of an enhancer blocking element contributing to the selectivity of a transcription factor acutely induced by environmental factors, such as HIF. Although further work is required to determine whether this is a general mechanism contributing to HIF specificity, our working hypothesis is that this mechanism could be particularly relevant in cases of bidirectional promoters, such as the one described herein, where chromatin accessibility and DNA methylation is likely to be similar across the intergenic region. In support to this possibility, we have found at least one further example of bidirectional promoters differentially regulated by hypoxia, *BCKDHA* and *EXOSC5*. Preliminary results indicate that, in spite of similar basal transcription, *BCKDHA*, but not *EXOSC5*, is induced by hypoxia (Supplementary Figure S3).

Another issue raised by our results relates to the molecular identity of the insulator element located upstream of the *RUVBL2* gene. In vertebrates, several regulatory elements including CTCF binding motifs (36–38),

repetitive elements, [such as ALUs (39), SINE B2 (29) and SINE B1 (30)] and scaffold/matrix-attachment regions [S/MARs; (40,41)], have been shown to function as insulators (25,42). Among them, the most widespread and well characterized are CTCF-binding elements (43). We have not found obvious CTCF binding motif within the −202 to −30 region upstream of *RUVBL2*, and published ChIP data shows CTCF binding to the GYS1 TSS region, but not to the region between the HRE and *RUVBL2* gene (Supplementary Figure S4A, CTCF ChIP signal track). On the other hand, there are no repetitive elements within this genomic region (Supplementary Figure S4A, repeats tracks) arguing against the existence of a SINE element that could explain the observed enhancer blocking activity. Finally, although we found no locally high proportion of A/T nucleotides, typically associated with S/MARs elements (41) in the region upstream of *RUVBL2* (Supplementary Figure S4A, GC percent track), computer prediction of S/MAR sites, found a significant signal in the upstream region of

RUVBL2 (Supplementary Figure S4B). Thus, the insulator activity described herein could be mediated by S/MAR elements. However, as this evidence has been obtained through *in silico* approaches, further work is required to identify the minimal region required for the enhancer blocking function, including its molecular identity, underlying mechanism and the associated potential *trans*-acting factors. This is yet another perfect example demonstrating the diversity of mechanisms, most of them not known to date, that cells are using to organize functional insulator elements (25).

The mechanism by which enhancer blockers prevent the activity of upstream enhancers is unclear, although our current understanding is that insulators will probably not be using unique mechanisms but, rather, adaptations of pre-existing ones already in place for the normal regulation of gene expression (23). Several of the proposed models invoke the generation of chromatin loops that segregate enhancer-sensitive and resistant promoters in distinct domains (44,45). In the case described herein, the model is further complicated by the short distance existing between the *cis*-elements involved. A piece of information that could shed light into the mechanism is the intriguing observation that the inclusion of the uG region seems to abolish the enhancer blocking effect of uR in EBA (Figure 4B, compare constructs 1R/1F with 3R/3F). Although we cannot currently explain this behavior, it would suggest that the EBA of uR can be modulated by elements located in its proximity. In this regard, it has been previously found that the EBA of the 'gypsy' element is affected by the number of copies of this element. When two copies, instead of one, are located between the enhancer and promoter, its blocking effect is abolished (46). However, the enhancer blocking effect is restored by the insertion of a third copy (47), depending on the order and distance of the insulator elements, indicating that complex protein–protein interactions are responsible for these unexpected effects and underlying a major role of insulators in whole nuclear genome organization (24,48). Thus, it is possible that the impairment of the EBA of uR by uG can be reverted by other *cis*-elements, present in the native genomic context, but not included in this set of constructs. A further possibility is that the HRE-containing block could enhance the transcription from the minimal CMV promoter. If this is the case, the enhancer blocking element within uR would be located upstream of the HRE element in the construct 1R and would be thus, unable to prevent its action upon the minimal CMV promoter. Regardless of the specific mechanism by which uR exerts its effect, our data clearly demonstrates that it contains an EBA.

Reprogramming of cellular metabolism, in particular glucose metabolism, is central in the cellular adaptation to hypoxia. The hypoxic induction of *GYS1*, encoding for an isoform of glycogen synthase, is part of this reprogramming (27). On the other hand, the existence of an enhancer blocking element between the HRE and the *RUVBL2* promoter raises the question of why *RUVBL2* expression has to be shielded from the HIF-mediated induction. *RUVBL2* gene encodes for Reptin, an AAA+ ATPase that forms part of chromatin remodeling complexes. Interestingly, it has been recently shown that hypoxia leads to reptin methylation and that, upon this post-translational modification, it is able to repress HIF-mediated transcription (49). Thus, it is plausible that *RUVBL2* expression might not be induced by hypoxia so as not to upset the balance between methylated/unmethylated reptin that could lead to premature termination of HIF-mediated transcription. However, this hypothesis does not provide an explanation for the close proximity of these genes. The conservation of the *GYS1/RUVBL2* genomic arrangement from opossums to humans suggests a selective pressure to maintain both genes in close proximity. Previous studies have shown that a substantial proportion of mammalian genes is arranged in a divergent head-to-head structure and controlled by bidirectional promoters (50), so that the pair of genes tend to be co-expressed (50,51). The need for co-regulation of the pair of genes under the control of a bidirectional promoter could explain the selective pressure that keeps them in close proximity. However, in the case of *GYS1/RUVBL2*, co-regulation of both genes is an unlikely reason for their close proximity as the existence of the insulator element would prevent the action of flanking *cis*-elements on the opposite promoter.

In summary, we have identified an insulator, acting as a functional enhancer blocking element, that explains the differential response of *GYS1* and *RUVBL2* genes to hypoxia. To our knowledge, this is the first report describing a role for this type of genetic elements in dictating the specificity of acutely induced transcription factors in response to environmental, as opposed to developmental, signals. Importantly, the generalization of this model adds to the arsenal of strategies, including chromatin accessibility and combinatorial assembly of TFs, that are employed by eukaryotes to ensure a highly specific gene expression based on an otherwise promiscuous set of *cis*-regulatory elements.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Tables 1 and 2 and Supplementary Figures 1–4, Supplementary Reference [54].

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Wang,G.L., Jiang,B.H., Rue,E.A. and 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. USA*, **92**, 5510–5514.

2. Salceda,S. and Caro,J. (1997) Hypoxia-inducible factor 1alpha (HIF-1alpha) protein is rapidly degraded by the ubiquitin-proteasome system under normoxic conditions. Its stabilization by hypoxia depends on redox-induced changes. *J. Biol. Chem.*, **272**, 22642–22647.

3. Jiang,B.H., Zheng,J.Z., Leung,S.W., Roe,R. and Semenza,G.L. (1997) Transactivation and inhibitory domains of hypoxia-inducible factor 1alpha. Modulation of transcriptional activity by oxygen tension. *J. Biol. Chem.*, **272**, 19253–19260.

4. Wang,G.L. and Semenza,G.L. (1993) Characterization of hypoxia-inducible factor 1 and regulation of DNA binding activity by hypoxia. *J. Biol. Chem.*, **268**, 21513–21518.

5. Mole,D.R., Blancher,C., Copley,R.R., Pollard,P.J., Gleadle,J.M., Ragoussis,J. and 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**, 16767–16775.

6. Xia,X., Lemieux,M.E., Li,W., Carroll,J.S., Brown,M., Liu,X.S. and Kung,A.L. (2009) Integrative analysis of HIF binding and transactivation reveals its role in maintaining histone methylation homeostasis. *Proc. Natl Acad. Sci. USA*, **106**, 4260–4265.

7. Wenger,R.H., Stiehl,D.P. and Camenisch,G. (2005) Integration of oxygen signaling at the consensus HRE. *Sci. STKE*, **2005**, re12.

8. Xia,X. and Kung,A.L. (2009) Preferential binding of HIF-1 to transcriptionally active loci determines cell-type specific response to hypoxia. *Genome Biol.*, **10**, R113.

9. Schodel,J., Oikonomopoulos,S., Ragoussis,J., Pugh,C.W., Ratcliffe,P.J. and Mole,D.R. (2011) High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. *Blood*, **117**, e207–e217.

10. Ortiz-Barahona,A., Villar,D., Pescador,N., Amigo,J. and 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**, 2332–2345.

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

12. Pan,Y., Tsai,C.J., Ma,B. and Nussinov,R. (2010) Mechanisms of transcription factor selectivity. *Trends Genet.*, **26**, 75–83.

13. Ogata,K., Sato,K. and Tahirov,T.H. (2003) Eukaryotic transcriptional regulatory complexes: cooperativity from near and afar. *Curr. Opin. Struct. Biol.*, **13**, 40–48.

14. Panne,D., Maniatis,T. and Harrison,S.C. (2007) An atomic model of the interferon-beta enhanceosome. *Cell*, **129**, 1111–1123.

15. Zhang,W., Tsuchiya,T. and Yasukochi,Y. (1999) Transitional change in interaction between HIF-1 and HNF-4 in response to hypoxia. *J. Hum. Genet.*, **44**, 293–299.

16. Yamashita,K., Discher,D.J., Hu,J., Bishopric,N.H. and Webster,K.A. (2001) Molecular regulation of the endothelin-1 gene by hypoxia. Contributions of hypoxia-inducible factor-1, activator protein-1, GATA-2, AND p300/CBP. *J. Biol. Chem.*, **276**, 12645–12653.

17. Aprelikova,O., Wood,M., Tackett,S., Chandramouli,G.V. and Barrett,J.C. (2006) Role of ETS transcription factors in the hypoxia-inducible factor-2 target gene selection. *Cancer Res.*, **66**, 5641–5647.

18. Hu,C.J., Sataur,A., Wang,L., Chen,H. and Simon,M.C. (2007) The N-terminal transactivation domain confers target gene specificity of hypoxia-inducible factors HIF-1alpha and HIF-2alpha. *Mol. Biol. Cell*, **18**, 4528–4542.

19. Lupien,M., Eeckhoute,J., Meyer,C.A., Wang,Q., Zhang,Y., Li,W., Carroll,J.S., Liu,X.S. and Brown,M. (2008) FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. *Cell*, **132**, 958–970.

20. Iguchi-Ariga,S.M. and Schaffner,W. (1989) CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation. *Genes Dev.*, **3**, 612–619.

21. Perini,G., Diolaiti,D., Porro,A. and Della Valle,G. (2005) In vivo transcriptional regulation of N-Myc target genes is controlled by E-box methylation. *Proc. Natl Acad. Sci. USA*, **102**, 12117–12122.

22. Wenger,R.H., Kvietikova,I., Rolfs,A., Camenisch,G. and Gassmann,M. (1998) Oxygen-regulated erythropoietin gene expression is dependent on a CpG methylation-free hypoxia-inducible factor-1 DNA-binding site. *Eur. J. Biochem.*, **253**, 771–777.

23. Gaszner,M. and Felsenfeld,G. (2006) Insulators: exploiting transcriptional and epigenetic mechanisms. *Nat. Rev. Genet.*, **7**, 703–713.

24. Capelson,M. and Corces,V.G. (2004) Boundary elements and nuclear organization. *Biol. Cell*, **96**, 617–629.

25. Molto,E., Fernandez,A. and Montoliu,L. (2009) Boundaries in vertebrate genomes: different solutions to adequately insulate gene expression domains. *Brief. Funct. Genomic Proteomic.*, **8**, 283–296.

26. Burgess-Beusse,B., Farrell,C., Gaszner,M., Litt,M., Mutskov,V., Recillas-Targa,F., Simpson,M., West,A. and Felsenfeld,G. (2002) The insulation of genes from external enhancers and silencing chromatin. *Proc. Natl Acad. Sci. USA*, **99**, 16433–16437.

27. Pescador,N., Villar,D., Cifuentes,D., Garcia-Rocha,M., Ortiz-Barahona,A., Vazquez,S., Ordonez,A., Cuevas,Y., Saez-Morales,D., Garcia-Bermejo,M.L. *et al.* (2010) Hypoxia promotes glycogen accumulation through hypoxia inducible factor (HIF)-mediated induction of glycogen synthase 1. *PLoS One*, **5**, e9644.

28. Barrett,T., Troup,D.B., Wilhite,S.E., Ledoux,P., Rudnev,D., Evangelista,C., Kim,I.F., Soboleva,A., Tomashevsky,M., Marshall,K.A. *et al.* (2009) NCBI GEO: archive for high-throughput functional genomic data. *Nucleic Acids Res.*, **37**, D885–D890.

29. Lunyak,V.V., Prefontaine,G.G., Nunez,E., Cramer,T., Ju,B.G., Ohgi,K.A., Hutt,K., Roy,R., Garcia-Diaz,A., Zhu,X. *et al.* (2007) Developmentally regulated activation of a SINE B2 repeat as a domain boundary in organogenesis. *Science*, **317**, 248–251.

30. Roman,A.C., Gonzalez-Rico,F.J., Molto,E., Hernando,H., Neto,A., Vicente-Garcia,C., Ballestar,E., Gomez-Skarmeta,J.L., Vavrova-Anderson,J., White,R.J. *et al.* (2011) Dioxin receptor and SLUG transcription factors regulate the insulator activity of B1 SINE retrotransposons via an RNA polymerase switch. *Genome Res.*, **21**, 422–432.

31. Bessa,J., Tena,J.J., de la Calle-Mustienes,E., Fernandez-Minan,A., Naranjo,S., Fernandez,A., Montoliu,L., Akalin,A., Lenhard,B., Casares,F. *et al.* (2009) Zebrafish enhancer detection (ZED) vector: a new tool to facilitate transgenesis and the functional analysis of cis-regulatory regions in zebrafish. *Dev. Dyn.*, **238**, 2409–2417.

32. Kawakami,K. (2004) Transgenesis and gene trap methods in zebrafish by using the Tol2 transposable element. *Methods Cell Biol.*, **77**, 201–222.

33. R Development Core Team. (2008) *R: A Language and Environment for Statistical Computing*. Vienna, Austria.

34. Recillas-Targa,F., Bell,A.C. and Felsenfeld,G. (1999) Positional enhancer-blocking activity of the chicken beta-globin insulator in transiently transfected cells. *Proc. Natl Acad. Sci. USA*, **96**, 14354–14359.

35. Recillas-Targa,F., Pikaart,M.J., Burgess-Beusse,B., Bell,A.C., Litt,M.D., West,A.G., Gaszner,M. and Felsenfeld,G. (2002) Position-effect protection and enhancer blocking by the chicken beta-globin insulator are separable activities. *Proc. Natl Acad. Sci. USA*, **99**, 6883–6888.

36. Bell,A.C., West,A.G. and Felsenfeld,G. (1999) The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell*, **98**, 387–396.

37. Furlan-Magaril,M., Rebollar,E., Guerrero,G., Fernandez,A., Molto,E., Gonzalez-Buendia,E., Cantero,M., Montoliu,L. and Recillas-Targa,F. (2011) An insulator embedded in the chicken alpha-globin locus regulates chromatin domain configuration and differential gene expression. *Nucleic Acids Res.*, **39**, 89–103.

38. Giraldo,P., Martinez,A., Regales,L., Lavado,A., Garcia-Diaz,A., Alonso,A., Busturia,A. and Montoliu,L. (2003) Functional dissection of the mouse tyrosinase locus control region identifies a new putative boundary activity. *Nucleic Acids Res.*, **31**, 6290–6305.

39. Willoughby,D.A., Vilalta,A. and Oshima,R.G. (2000) An Alu element from the K18 gene confers position-independent expression in transgenic mice. *J. Biol. Chem.*, **275**, 759–768.

40. McKnight,R.A., Shamay,A., Sankaran,L., Wall,R.J. and Hennighausen,L. (1992) Matrix-attachment regions can impart position-independent regulation of a tissue-specific gene in transgenic mice. *Proc. Natl Acad. Sci. USA*, **89**, 6943–6947.

41. Millot,B., Montoliu,L., Fontaine,M.L., Mata,T. and Devinoy,E. (2003) Hormone-induced modifications of the chromatin structure surrounding upstream regulatory regions conserved between the mouse and rabbit whey acidic protein genes. *Biochem. J.*, **372**, 41–52.

42. Lunyak,V.V. (2008) Boundaries. Boundaries...Boundaries???. *Curr. Opin. Cell Biol.*, **20**, 281–287.

43. West,A.G., Gaszner,M. and Felsenfeld,G. (2002) Insulators: many functions, many mechanisms. *Genes Dev.*, **16**, 271–288.

44. Farrell,C.M., West,A.G. and Felsenfeld,G. (2002) Conserved CTCF insulator elements flank the mouse and human beta-globin loci. *Mol. Cell. Biol.*, **22**, 3820–3831.

45. Byrd,K. and Corces,V.G. (2003) Visualization of chromatin domains created by the gypsy insulator of Drosophila. *J. Cell Biol.*, **162**, 565–574.

46. Muravyova,E., Golovnin,A., Gracheva,E., Parshikov,A., Belenkaya,T., Pirrotta,V. and Georgiev,P. (2001) Loss of insulator activity by paired Su(Hw) chromatin insulators. *Science*, **291**, 495–498.

47. Savitskaya,E., Melnikova,L., Kostuchenko,M., Kravchenko,E., Pomerantseva,E., Boikova,T., Chetverina,D., Parshikov,A., Zobacheva,P., Gracheva,E. *et al*. (2006) Study of long-distance functional interactions between Su(Hw) insulators that can regulate enhancer-promoter communication in Drosophila melanogaster. *Mol. Cell. Biol.*, **26**, 754–761.

48. Phillips,J.E. and Corces,V.G. (2009) CTCF: master weaver of the genome. *Cell*, **137**, 1194–1211.

49. 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. *et al*. (2010) Negative regulation of hypoxic responses via induced Reptin methylation. *Mol. Cell*, **39**, 71–85.

50. Trinklein,N.D., Aldred,S.F., Hartman,S.J., Schroeder,D.I., Otillar,R.P. and Myers,R.M. (2004) An abundance of bidirectional promoters in the human genome. *Genome Res.*, **14**, 62–66.

51. Li,Y.Y., Yu,H., Guo,Z.M., Guo,T.Q., Tu,K. and Li,Y.X. (2006) Systematic analysis of head-to-head gene organization: evolutionary conservation and potential biological relevance. *PLoS Comput. Biol.*, **2**, e74.

52. Birney,E., Stamatoyannopoulos,J.A., Dutta,A., Guigo,R., Gingeras,T.R., Margulies,E.H., Weng,Z., Snyder,M., Dermitzakis,E.T., Thurman,R.E. *et al*. (2007) Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature*, **447**, 799–816.

53. Kent,W.J., Sugnet,C.W., Furey,T.S., Roskin,K.M., Pringle,T.H., Zahler,A.M. and Haussler,D. (2002) The human genome browser at UCSC. *Genome Res.*, **12**, 996–1006.

54. Rosenbloom,K.R., Dreszer,T.R., Pheasant,M., Barber,G.P., Meyer,L.R., Pohl,A., Raney,B.J., Wang,T., Hinrichs,A.S., Zweig,A.S. *et al*. (2010) ENCODE whole-genome data in the UCSC Genome Browser. *Nucleic Acids Res.*, **38**, D620–D625.

PLOS ONE

Cooperativity of Stress-Responsive Transcription Factors in Core Hypoxia-Inducible Factor Binding Regions

Diego Villar[1]*[¤], Amaya Ortiz-Barahona[1], Laura Gómez-Maldonado[1], Nuria Pescador[1], Fátima Sánchez-Cabo[2], Hubert Hackl[3], Benjamin A. T. Rodriguez[4], Zlatko Trajanoski[3], Ana Dopazo[2], Tim H. M. Huang[4], Pearlly S. Yan[4], Luis del Peso[1]

1 Department of Biochemistry, Universidad Autónoma de Madrid and Instituto de Investigaciones Biomedicas Alberto Sols, Madrid, Spain, 2 Genomics Unit, Centro Nacional de Investigaciones Cardiovasculares (CNIC), Madrid, Spain, 3 Biocenter, Division of Bioinformatics, Innsbruck Medical University, Innsbruck, Austria, 4 Human Cancer Genetics Program, Department of Molecular Virology, Immunology, and Medical Genetics, The Ohio State University, Columbus, Ohio, United States of America

Abstract

The transcriptional response driven by Hypoxia-inducible factor (HIF) is central to the adaptation to oxygen restriction. Despite recent characterization of genome-wide HIF DNA binding locations and hypoxia-regulated transcripts in different cell types, the molecular bases of HIF target selection remain unresolved. Herein, we combined multi-level experimental data and computational predictions to identify sequence motifs that may contribute to HIF target selectivity. We obtained a core set of *bona fide* HIF binding regions by integrating multiple HIF1 DNA binding and hypoxia expression profiling datasets. This core set exhibits evolutionarily conserved binding regions and is enriched in functional responses to hypoxia. Computational prediction of enriched transcription factor binding sites identified sequence motifs corresponding to several stress-responsive transcription factors, such as activator protein 1 (AP1), cAMP response element-binding (CREB), or CCAAT-enhancer binding protein (CEBP). Experimental validations on HIF-regulated promoters suggest a functional role of the identified motifs in modulating HIF-mediated transcription. Accordingly, transcriptional targets of these factors are over-represented in a sorted list of hypoxia-regulated genes. Altogether, our results implicate cooperativity among stress-responsive transcription factors in fine-tuning the HIF transcriptional response.

Introduction

Oxygen is essential for the survival of all eukaryotic cells, and metazoans are heavily dependent on this element to meet their large metabolic demands. At the cellular level, 90% of oxygen is consumed in oxidative phosphorylation. Consistent with a central role of oxygen in aerobic metabolism, all metazoan cells respond to an imbalance between demand and supply of oxygen (hypoxia) by activating a gene expression program aimed at restoring oxygen supply and reducing its consumption. The cellular response to hypoxia is mainly controlled by the evolutionarily conserved Hypoxia Inducible Factor (HIF) family of basic helix-loop-helix transcription factors. HIFs are heterodimers of a beta subunit (HIFβ, also known as ARNT), and an alpha subunit (HIFα) [1]. While ARNT levels are not sensitive to oxygen, both HIFα stability [2] and its transcriptional activity [3] are regulated by oxygen-dependent hydroxylation [4–6]. Under oxygen restriction, HIFα subunits escape proteasomal degradation, heterodimerize with HIFβ subunits and translocate to the cell nucleus, where they bind the RCGTG consensus sequence (termed Hypoxia Response Element, HRE) within regulatory regions of target genes, leading

to their transcriptional activation in hypoxia [7]. Mammals present three isoforms of HIFα (HIF1α, HIF2α and HIF3α) that differ in their tissue distribution, HIF1α being the more ubiquitous and best characterized [8].

A large number of studies focusing on single genes have identified individual HIF targets that, collectively, account for the functional responses to hypoxia, mainly metabolic adaptation and induction of angiogenesis [7]. More recently, works employing HIF1α and HIF2α chromatin immunoprecipitation coupled to genomic microarrays (ChIP-chip) or high-throughput sequencing (ChIP-Seq) have addressed the genome-wide identification of HIF binding locations [9–12], thereby improving the existing knowledge on the HIF-modulated transcriptome and largely confirming the RCGTG HIF binding consensus. Additionally, these studies have provided important insights into the global properties of HIF1 binding and transactivation. First, these works reported a significant association between the presence of a HIF binding site (HBS) and hypoxic induction of the neighboring genes. The same trend was not found for genes repressed by hypoxia, suggesting that hypoxia-mediated repression is largely indirect or HIF-

independent [9,12,13]. Furthermore, they have clearly shown that only a small subset of about a hundred of all RCGTG-containing genes is robustly regulated by hypoxia. Hence, and in agreement with work on other transcription factors [14], HIFs bind a small proportion of potential binding sites, albeit the basis of their binding and target selectivity are incompletely understood.

Understanding the mechanisms that explain HIFs transactivation selectivity is of paramount importance to expand our knowledge on transcriptional regulation and to improve the sensitivity and specificity of genome-wide efforts to characterize the HIF transcriptional response. DNA accessibility of transcription factor binding sites (TFBSs) can clearly contribute to binding selectivity [15]. For HIFs, recent evidence includes enhanced HIF1 and HIF2 binding to normoxic DNAse hypersensitivity sites [12] and enrichment of HIF1 binding in the proximity of genes with a "permissive" transcriptional state in normoxia, as evidenced by significant basal expression [11]. Additionally, DNA methylation has been also shown to modulate HIF1 binding, as originally demonstrated for the 3′ enhancer of the erythropoietin gene [16,17]. A further mechanism that can impact target selectivity is direct or indirect cooperativity between transcription factors (TFs). Models of direct cooperativity have been mainly derived from developmental enhancers, and include the strict enhanceosome model [18], where cooperative occupancy occurs through extensive protein-protein interactions between TFs or common cofactors, and the more flexible billboard model [19], which suggests that enhancers contain submodules that interact independently or redundantly with promoters. Conversely, indirect cooperativity is based on the equilibrium competition between nucleosomes and DNA-binding proteins, thereby not requiring protein-protein interactions [20]. In the case of HIF-mediated transcription, the binding of cooperating transcription factors has been demonstrated for several target genes. In particular, HIF-mediated expression of the erythropoietin gene requires an adjacent HNF4 binding site [21], both GATA2 and AP1 binding sites are necessary for epithelial induction of *ET-1* under hypoxia [22], and *PAI-1* induction by hypoxia has been linked to cooperative promoter activation by CEBPα, HIF1α and EGR-1 [23]. Other examples include cooperation with Smads [24], Sp1 [25] or CREB [26]. Additionally, USFs have been shown to complement HIF binding either at neighbouring (*LDHA* promoter) or identical sites (*BNIP3*) [27], while collaboration with ETS transcription factors has been proposed to play a role in HIF2α target selectivity [28,29]. Recent genome-wide approaches relying on experimental [9–11] and computational [13,30] identification of HIF binding sites have reported overrepresented transcription factor binding sites in the flanking sequences that might be indicative of transcriptional cooperativity. However, significant differences exist in the overrepresented TFBSs predicted in each study, and the functional significance of these enriched motifs remains unclear.

Gene expression profiling indicates that the expression of thousands of genes changes with hypoxia, with vast cell-type differences in the specific genes being regulated [31–38]. HIF1α ChIP-chip binding locations have been reported in cell lines of diverse tissue origin, namely HepG2 hepatocarcinoma cells [9], MCF-7 breast cancer cells [10] and U87 glioma cells [11], showing differences in the binding sites identified in each experiment. In previous studies we integrated microarray expression profiling experiments and HIF binding site predictions in a core set of tissue-independent HIF target genes [13]. To further investigate the selectivity of HIF1 binding, in this work we conducted HIF1α ChIP-chip in cervical carcinoma HeLa cells and observed largely non-overlapping binding locations with previous studies. To explore the role of cooperativity in HIF target selection, we integrated HIF1 alpha ChIP-chip binding locations across cell-types with a meta-analysis of gene expression profiles of cells exposed to hypoxia [13]. Computational prediction of enriched transcription factor binding sites in this integrated set suggested several stress-responsive transcription factors as potential HIF1 collaborators. Experimental validation of these predictions in cell-based reporter assays indicates that binding sites for stress-responsive transcription factors other than HIFs, such as CEBPs, contribute to cooperative hypoxic activation of individual targets.

Materials and Methods

Gene-expression Profiling Analysis

Gene profiling experiments of hypoxic cell cultures were downloaded from the Gene Expression Omnibus database (GEO, http://www.ncbi.nlm.nih.gov/geo/) [39]. The average raw signal from biological replicates was used in the analysis. When fifty percent of the replicates had null values the average signal was set to null. All probes mapping to a particular locus were considered independently. A gene (identified by a particular probe) was recorded as having no basal expression when the signal for the probe under normoxic conditions had a null value. A gene (probe) was considered to be induced by hypoxia when the log-ratio of the hypoxic over normoxic signal values exceeded by 2.6 standard deviations the average log-ratio of all the probes in the array. Genes (probes) with a null normoxic value and not-null hypoxic values were also considered as induced by hypoxia. The presence and absence of conserved RCGTG motifs at each locus was determined as described previously [13].

Cell Culture

HeLa cells were maintained in Dulbecco's modified Eagle medium supplemented with 100 units/ml penicillin, 100 μg/ml streptomycin and 5% (v/v) fetal bovine serum. Cells were grown at 37°C and 5% CO_2. Hypoxic treatments were carried out in sealed chambers flushed with a $1\%O_2/5\%CO_2/94\%N_2$ gas mixture (Billups-Rothenberg, Inc.; CA, USA).

Chromatin Immunoprecipitation

Chromatin immunoprecipitation was performed as previously described [40]. Briefly, 10^7 HeLa cells were subjected to hypoxia (1% oxygen) for six hours or left untreated (normoxic conditions, 21% oxygen). Following treatments, cells were crosslinked with 1% formaldehyde for 12 min at 4°C. Cross-linking was stopped by the addition of glycine (0.125 M final concentration). Cell lysis was achieved by scraping in 1 ml of lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris/HCl, pH 8.1, and a protease inhibitor cocktail, Roche). Cell lysates were incubated on ice for 10 min and then sonicated to shear DNA to fragments between 200 and 1500 bp. Only experiments that showed homogeneous sonication across all samples (from normoxia and hypoxia treatments) were continued. 50 μl of each sample was stored (input), while 100 μl were diluted in 1 ml of immunoprecipitation buffer (1% Triton X-100, 2 mM EDTA, 150 mM NaCl and 20 mM Tris/HCl, pH 8.1). Lysates were precleared with 200 μg of a Salmon Sperm DNA/Protein A agarose 50% slurry (Upstate Biotechnology, Lake Placid, NY, U.S.A.) for 1 h at 4°C; and then immunoprecipitated twice, initially with whole rabbit serum for 6 h (IgG control) and then overnight at 4°C with a polyclonal anti-HIF1 alpha antiserum (Abcam, ab2185). Immunocomplexes were recovered by addition of 400 μg of Salmon Sperm DNA/Protein A agarose 50% slurry, and then sequentially washed in Low Salt Wash Buffer (0.1%SDS, 1%Triton X-100, 2 mM EDTA,

20 mM Tris/HCl, pH 8.1, and 150 mM NaCl), High Salt Wash Buffer (0.1% SDS, 1% Triton X-100,2 mM EDTA, 20 mM Tris/HCl, pH 8.1, and 500 mM NaCl), LiCl buffer (0.25 M LiCl,1% Nonidet P40, 1% deoxycholate, 1 mM EDTA and, 10 mM Tris/HCl, pH 8.1), and twice in TE buffer (10 mM Tris, pH 8.0, and 1 mM EDTA). Elution of protein-bound DNA was performed twice with 1% SDS 0.1 M NaHCO3. Eluates were pooled, and crosslinking was reversed by the addition of 200 mM NaCl (final concentration) and overnight incubation at 65°C. Protein and RNA were removed by the addition of proteinase K (30 µg/sample) and RNAse (200 mg/µl) for 2 hours at 42°C, and immunoprecipitated DNA was purified by phenol-chloroform extraction and ethanol precipitation. Successful ChIP was assayed by standard PCR using two primer pairs, targeting the functional EGLN3 HRE and a control region in the same *locus* [40].

Chromatin Immunoprecipitation on Microarray

The ChIP-chip method was carried out as previously described [41]. First, purified DNA from chromatin immunoprecipitation was amplified by ligation-mediated PCR. DNA ends were extended by incubation with T4 DNA polymerase (New England Biolabs), and blunted DNA was ligated to pre-annealed oligonucleotide linkers (JW102 gcggtgacccgggagatctgaattc and JW103 gaattcagatc) using T4 DNA ligase (New England Biolabs), and subsequently amplified by two rounds of PCR using JW102 as primer, to yield 2–5 µg of amplified DNA. An aliquot of this material was run in a microfluidics platform (Agilent 2100 Electrophoresis Bioanalyzer) to accurately measure size distribution of amplified material and discard gross amplification bias. Additionally, quantitative PCR against both a Hypoxia-Response Element (HRE) in the EGLN3 locus [40] and a control negative region in the same locus was routinely performed to assess loss of enrichment during amplification.

Amplified DNA from normoxic and hypoxic chromatin immunoprecipitation samples was labelled with Cy3 and Cy5 fluorescent dyes, respectively, and hybridized to microarrays following guidelines from the microarray manufacturer (Agilent Mammalian ChIP-on-chip Protocol v.10). Hybridized microarray slides were scanned in an Agilent DNA microarray scanner (Agilent Technologies, CA, USA) at 5 µm resolution, and acquired microarray images were quantified with GenePix 6.0 software (Molecular Devices, CA, USA). A total of six hybridizations were conducted, corresponding to four biological replicates. The two technical replicates were dye-swap experiments, where normoxic samples were labelled with Cy5 and hypoxic samples with Cy3.

Analysis of ChIP-chip Data

A custom alternative promoter microarray was used for ChIP-chip hybridizations [42]. Probes in the array cover 34000 known or putative promoters representing about 7000 human genes, and tile a region from −200 to +200 of known or predicted transcription start sites, with an average probe spacing of 80 bp. Genomic coordinates of the probes in the array (hg17, May 2004) were updated to the hg19 assembly using the alignment tool Exonerate [43] with 97% sequence identity. Probes having non-unique matches to this version of the Human Genome were excluded from ChIP-chip analysis.

The R/Bioconductor packages "Ringo" and "limma" were used to analyze ChIP-chip readouts [44,45]. Limma analysis was performed after normalization of ChIP-chip data with the variance-stabilizing method. A separate linear model was fitted to each biological replicate, which comprised a single readout or two in the case of dye-swap experiments, and these models were averaged to obtain a single linear model that includes estimation of moderated t-statistic p values. The Benjamini-Hochberg correction for multiple comparisons (False Discovery Rate) was applied to these p values. Ringo analysis was performed essentially as described [46], using the parameters indicated below. Raw data were again normalized with the variance-stabilizing method. First, for the calculation of the average smoothed signal across replicates, we used: winHalfSize = 100 (based on probe density and spacing in the array) and quant = 0.75. To obtain a threshold intensity value for bound probes, a 0.99 quantile was used as upper bound for the null distribution. For the identification of ChIP-enriched regions on the smoothed signal, distCutOff = 200 (maximum probe spacing within a single ChIP-enriched region) and minProbesInRow = 4 (minimum number of probes per region) were used. Minor modifications to Ringo functions "cherByThreshold" and "findChersOnSmoothed" were made to take into account probe-wise p values (as previously calculated by limma) for ranking of ChIP-enriched regions. Specifically, ChIP-enriched regions found by Ringo were required to harbour one or more probes with an FDR-adjusted p value lower than 0.02 (2% false discovery rate). Finally, ChIP-enriched regions in poorly covered regions of the array (having less than 8 total probes, 4 inside the region and 4 surrounding it) and those mapping to repetitive regions were identified with in-house Perl scripts and taken out of the analysis.

The microarray experiments described above have been deposited in ArrayExpress under accession number E-MEXP-3499.

Quantitative PCR

Purified DNA from chromatin immunoprecipitation samples was used in quantitative PCR with SYBR green-based detection (Applied Biosystems) following the manufacturer's recommendations. Primers targeting candidate ChIP-enriched regions were designed with Primer Express software v2.0 (Applied Biosystems) and Primer3 (http://frodo.wi.mit.edu/primer3/). All measurements were carried out in triplicate. Threshold cycle (Ct) values for each sample were interpolated in a standard curve of input DNA dilutions to obtain % of input absolute values. Enrichment of HIF1 alpha binding to target loci was calculated as the ratio of the amounts of target sequence (measured as % of input) detected in hypoxic vs noxmoxic ChIP samples (% of input hypoxia/% of input normoxia). For validation of ChIP-chip candidates, three negative control regions (in the EGLN3, IRS4 and HIPEV1 loci, [13]) were used to estimate an average background enrichment, and a 90% confidence interval was applied on these values to set a threshold for successful validation of candidate regions.

Obtaining a High-confidence Set of Core HIF Binding Regions and a Background Set of Control Regions

Custom scripts written in Perl were used to indentify evolutionarily conserved HIF binding sites (HBSs) within ChIP-chip regions and to select HBSs that showed evolutionary conservation. Conservation of the HIF binding consensus RCGTG in four mammalian species, including mouse, was required for the evolutionary conservation filter. HBSs were further selected to map to genes robustly induced by hypoxia, as indicated by the results of a meta-analysis of gene expression profiling experiments in hypoxic cell cultures [13], using a p value threshold of 0.02 (FDR). Finally, HBS coordinates were extended into HIF binding regions (HBRs) that spanned surrounding non-coding conserved sequences, as determined by >50% presence of phastCons elements [47], and the corresponding multiple sequence alignment of each HBR was retrieved. Multiple

sequence alignments were downloaded from the UCSC Genome Browser's Table Browser [48].

To obtain a set of background (control) genomic regions, custom perl scripts were used to exploit the microarray meta-analysis results for the identification of genes harbouring conserved RCGTG motifs but that are unlikely to be modulated by hypoxia. To this end, gene *loci* that contained conserved RCGTG motifs in their non-coding sequences were first selected. For these genes, each of their probes was examined, and only genes for which all of their associated probes exhibited a mean fold value within 0.25 standard deviations of the global mean in each of the 19 datasets employed in the meta-analysis were considered as not induced by hypoxia. The selected coordinates of conserved RCGTG motifs mapping to these *loci* were extended as previously described for the set of core HIF binding sites. Genomic regions from this collection were further selected to match the frequency of genomic locations (relative to the TSS) found in the core HBR set. Briefly, Perl scripts were used to annotate core HBRs as promoter, 5'UTR, intronic or 3'UTR genomic locations and to choose, from the whole collection, a random sample according to the proportions of genomic locations found in the core HBR set. Similarly as with the set of core HIF binding regions, multiple sequence alignments corresponding to the selected control regions were retrieved.

In silico TFBS Prediction

In silico transcription factor binding site predictions were carried out employing custom scripts written in Perl. Position-weight matrices (PWMs) collections were downloaded from the public databases JASPAR (2010 release) and Transfac (7.0 version) [49,50]. Raw frequency matrices were transformed into log-odd matrices to take into account the background nucleotide frequencies found in the whole collection of HIF binding regions (core and background together), and the information content of each matrix position was calculated as proposed by Stormo [51]. The formulae used for these calculations are detailed below.

PWM conversion:

The weight of base b in position i is calculated as:

$$W_{b,i} = \log_2 \frac{p(b,i)}{p(b)},$$

where p(b,i) is the corrected probability of base b in position i and p(b) the background probability of base b.

The corrected probability is obtained from the raw matrix by adding a pseudocount:

$$p(b,i) = \frac{f(b,i) + \frac{1}{N}}{N},$$

where f(b,i) are the counts of base b in position i and N the number of sites used to construct the matrix.

Information content calculation:

$$D_i = \sum_b p(b,i) \cdot \log_2 \frac{p(b,i)}{p(b)}$$

Subsequently, log-odd matrices were used to screen core HIF binding regions and genomic-matched background regions for the presence of putative TFBSs or other sequence motifs. Perl scripts were employed to split each sequence into overlapping fragments of length equal to that of the PWM under analysis. For each fragment, a score value was calculated by summing up the log-odd

frequencies obtained by substitution of nucleotides found in the fragment in the corresponding position of the PWM. The contribution of each base to the score was weighted by the information content of its position in the matrix, as detailed below.

Score calculation:

$$Sc = \sum_i D_i \cdot W_{b,i},$$

where D_i is the information content of position i and $W_{b,i}$ the log-odd weight of base b in position i.

Finally, the resulting score was normalized by subtraction of the minimum score and division by the score range, and compared with a threshold value. Fragments showing a score above the threshold were considered as putative TFBSs, and the evolutionary conservation of nucleotides in each motif was evaluated for matrix positions with information content over 60%. Putative TFBSs showing evolutionary conservation in four mammalian species (including mouse) and whose score was over the threshold value were recorded as present (respresented as 1). Otherwise, they were considered absent from the analyzed sequence (represented as 0).

For each PWM, the three strategies proposed in MATCH [52] were used for the calculation of threshold values. The minFN strategy aims at minimizing false negative predictions (low stringency), and was obtained by setting the threshold value that detects 90% of cases in a randomly generated sample of sequences in which the probability of nucleotides at each position is dictated by the matrix frequencies. Conversely, the minFP threshold focuses on minimizing false positives (high stringency), and its calculation is based on the assumption that coding sequences in the Genome are impoverished in functional TFBSs. We used the threshold value that results in a single hit (on average) per 10000 bp when the matrix is used on all human coding sequences. Finally, the minSum threshold was obtained by estimating the false positive and false negative rate for all threshold values between minFN and minFP, and then choosing the value that minimizes the sum of both (medium stringency).

Statistical Analysis of Enriched TFBSs

Fisher's exact test was used to identify PWMs showing significant enrichment in the set of core HIF binding regions *versus* the background collection. In particular, we considered significant PWMs with a p value lower than 0.05. No correction for multiple comparisons was applied to these p values.

Additionally, the Weka machine learning workbench [53] (3.6 version) was used to identify the most informative PWMs, that were better able to distinguish core HIF binding regions from the background set. To this end, the correlation-based feature selection variable selection procedure was used. The algorithm was run 10 times using stratified 10-fold cross-validation in each iteration. Finally, the number of times that each variable had been selected was recorded. This number ranges from zero (never chosen) to a hundred (corresponding to every cross-validation fold and every iteration).

Plasmid Construction

Human genomic DNA extracted from HeLa cells was used as template for PCR amplification of the *CA9* promoter region (hg19 coordinates chr9:35673508–35673956), which was subcloned into the pGL3-Basic plasmid (Invitrogen). The Human *GYS1* reporter construct (hg19 chr19:49496421–49496978) has been previously described [54]. The mouse *LDHA* promoter construct (mm9 chr7:54101027–54101258) and the derived HRE and CREB

binding site mutations were a kind gift from Peter Ratcliffe [26]. The remaining HRE, predicted AP1 or CEBPB binding sites and control mutations were generated by site-directed mutagenesis, employing QuikChange Site-directed mutagenesis kit (Stratagene). The introduced mutations are detailed in Table S1. The identity of all constructs was verified by sequencing.

Reporter Assays

Reporter assays were performed using the human cervical-carcinoma cell line HeLa. Cells were seeded on six-well plates (2.5·105 cells/well) 6 h prior to transfection. Per well, a 4.5 µg DNA mixture containing 1.5 µg of the indicated reporter construct or empty plasmid and 0.25 µg of a plasmid encoding for Renilla (sea pansy) luciferase under the control of a null promoter (Promega, Madison, WI, U.S.A.) was used for transfection using the calcium phosphate method. 16 h after transfection, cells where washed, replated in 24-well plates, and incubated in normoxia, in the presence of DMOG (dimethyloxaloylglycine, 500 µM) or in hypoxia for an additional 16 hours. After treatments, cells were lysed and the firefly and Renilla luciferase activities of the lysate were determined using a dual-luciferase system (Promega, Madison, WI, U.S.A.). The firefly luciferase activity was normalized to that of Renilla luciferase. Each experimental condition was assayed in duplicate. Hypoxia or DMOG fold induction values for each experiment were analyzed by repeated measures ANOVA with a Dunnet posthoc correction, comparing values of the wild-type promoter construct to each of the introduced mutations.

Gene-set Enrichment Analysis

Gene-set enrichment analysis (GSEA) was carried out as previously described ([55] and http://www.broadinstitute.org/gsea/index.jsp). We employed a ranked list of core hypoxia-regulated genes derived from a meta-analysis of 16 hypoxia gene expression experiments [13], where genes are sorted by their mean hypoxic induction across cell lines/tissues represented in the meta-analysis. We studied the distribution of transcription factor targets in this list employing 3000 gene-sets from the GSEA molecular signatures database, which includes experimentally derived lists of targets for specific transcription factors. GSEA analysis provides and enrichment score (ES) for each gene-set across the ranked list of hipoxia-responsive genes. In order to compare several gene-sets, enrichment scores are normalized to produce NES values (normalized enrichment scores). Comparison with NES values obtained from random gene-sets allows estimation of statistical significance. We used an FDR-adjusted p-value of 0.05.

Results

1. Basal Gene Expression is Necessary but not Sufficient for HIF Target Selection

Previous studies have proposed that chromatin accessibility and basal gene expression mediate HIF target selectivity [11,12]. In order to independently assess the contribution of this mechanism to HIF target selection, we exploited publicly available genome profiling experiments of hypoxic cell cultures [13] to look at the association of basal expression and hypoxic induction. We analyzed the basal expression of a list of well-characterized HIF targets [7] and correlated it with their response to hypoxia. In agreement with previous reports [11,12], we found that the response to hypoxia, scored as the percentage of HIF-target genes induced by the treatment, was significantly higher among genes that were already expressed in the basal (normoxic) condition (Figure 1A, $p < 0.01$ Wilcoxon matched test). Moreover, when the

HIF target genes across all datasets were pooled and categorized according to their basal expression and response to hypoxia, the distribution was significantly different to that expected by chance (Table S2, $p < 0.001$ Chi-square test). These results further suggest that chromatin accessibility contributes to HIF target selectivity by modulating the availability of RCGTG motifs. However, given the large number of genes with basal (normoxic) expression and the widespread distribution of the RCGTG motifs, it is expected that many RCGTG motifs would lie within open chromatin regions. To look at the contribution of chromatin accessibility in more detail, we next studied the response to hypoxia of all the genes with detectable normoxic values represented in each array. To this end, within the group of genes with basal gene expression, we identified those harbouring conserved HIF-binding motifs in their non-coding sequences (Figure 1B, HBS) and categorized them according to their induction by hypoxia. For each microarray dataset, the large majority of genes harbouring a conserved RCGTG motif were not induced by hypoxia (Figure 1B, yellow bar segments) in spite of proximity of the motif to genes with significant normoxic expression. This observation strongly suggests that although basal gene expression correlates with hypoxia inducibility of a gene, additional mechanisms are needed to specify HIF target selection.

2. Comparative Analysis of HIF1 Alpha Binding Locations in Cell Lines of Diverse Tissue Origin

The binding of additional transcription factors in proximity of HIFs constitutes a plausible mechanism that could contribute to HIF target selection. In this regard, previous works have addressed the identification of sequence motifs overrepresented in collections of HIF binding regions obtained from ChIP-chip datasets or combinations of computational predictions and gene expression profiling experiments [9,10,30]. A recent work failed to identify clearly overrepresented sequences [9], while the predictions reported in two other studies showed very small overlap [10,30]. On the other hand, the wealth of HIF1 alpha binding and hypoxic gene expression data obtained in different cell types provides a unique opportunity to construct integrated sets of HIF1 binding sites that may overcome the limitations of datasets based on a single experiment. In order to study the role of cooperativity in HIF target selectivity, we determined the genome-wide pattern of HIF1 alpha binding sites in cervical carcinoma HeLa cells and compared our results to previously published HIF1 ChIP-chip experiments in hepatocellular carcinoma HepG2 cells [9], breast cancer MCF-7 cells [10] and U87 glioma cells [11], as detailed below.

For the determination of HIF binding sites in HeLa cells, we performed HIF1α chromatin immunoprecipitation in HeLa cells exposed to normoxia or hypoxia (1% oxygen) for six hours. Amplified samples from normoxic and hypoxic cells were competitively hybridized to a proximal promoter microarray that tiles a subset of human 7000 genes [42]. ChIP-chip data was analyzed with the R/Bioconductor packages Ringo [44] and limma [45] to identify hypoxic HIF1-bound genomic regions. Stringent statistical thresholds (2% FDR) were applied to normalized signals from four biological replicates (Figure 2A, all). ChIP-enriched regions were required to harbour four or more probes above background average signal (Figure 2A, blue horizontal line) and one or more probes robustly induced by hypoxia in a linear model of the four replicates (Figure 2B, red dots, 2% FDR). This analysis produced a ranked list of 57 HIF1 binding regions (Table S3), spanning the coordinates of previously characterized HIF binding sites [7] and including many potentially novel HIF1 binding sites and HIF1 targets. Quantitative PCR
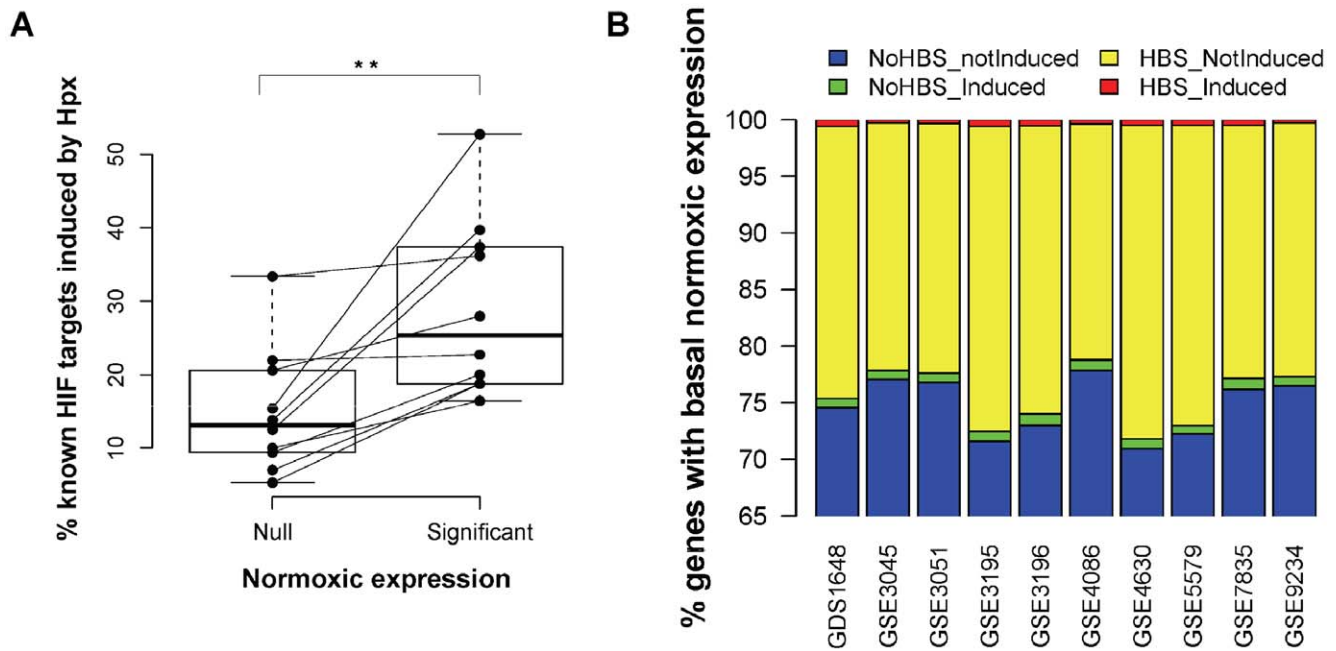
Figure 1. Basal gene expression is not sufficient for HIF transactivation. (A) A list of well-characterized HIF target genes (from ref. 7) present in individual gene expression profiling (microarray) datasets (see B for GEO IDs) were categorized according to their basal (normoxic) expression level into two groups: Null, no detectable basal expression; Significant, detectable basal expression. In addition, for each microarray experiment, HIF-target genes were further classified into those whose expression was significantly induced by hypoxia (ratio hypoxia/normoxia greater than 2.6Sd above the mean) and non-responsive genes. The graph represents the percentage of HIF target genes in each category that were induced by hypoxia. Each pair of joined dots represents the data from a single microarray experiment. Box and whisker plots represent the distribution of values in each category. **: p<0.01 (Wilcoxon matched test) **(B)** For each of the indicated microarray datasets (GEO identifiers in horizontal axis), we identified all the genes showing a significant basal (normoxic) expression. Then, we classified them according to their response to hypoxia ("Induced" and "notInduced", see A) and the presence of conserved RCGTG motifs in their regulatory regions ("HBS" and "NoHBS", respectively). The graph depicts cumulative percentages (vertical axis) of genes in each of the four combinations of the two categories: no conserved HIF binding motifs and no hypoxic induction (blue, NoHBS_notInduced), no conserved HIF binding motifs but hypoxic induction (green, NoHBS_Induced), conserved HIF binding motifs but no hypoxic induction (yellow, HBS_notInduced) and conserved HIF binding sites and hypoxic induction (red, HBS_Induced).
doi:10.1371/journal.pone.0045708.g001

validation of the ChIP-chip results in independent HIF1 alpha chromatin immunoprecipiations confirmed hypoxic enrichment of the majority of tested candidates (Figure 2C).

Next, we compared HIF1α ChIP-chip predictions in the four cell lines by analyzing the overlap of reported binding locations (Figure 2D). The majority (36 sequences) of HeLa HIF1α binding locations did not overlap with ChIP-chip results obtained in other cell types, although many were also found in at least one of the previous reports. A similar trend was observed taking any of the other studies as reference, suggesting that most HIF1 binding is cell-type specific. To test the role of cooperativity in dictating HIF1 target selection, we focused on HIF1α ChIP-chip binding locations common to two or more studies as a bona-fide set of core HIF1 binding regions. Analysis of evolutionary conservation in these sequences, defined as RCGTG motifs within PhastCons elements [47] and conserved in at least four mammalian species including human and mouse, showed a strong enrichment of conserved sequences in the core set of common HIF1 binding sites, versus those found uniquely in a single ChIP-chip study (Table S4). Since HBSs identified in more than one study are more likely to correspond to functional sites, this analysis suggest that evolutionary conservation of HIF binding motifs can be predictive of functionality as has been shown for other TFBSs [56,57].

3. Binding Sites for Diverse Stress-responsive Transcription Factors are Enriched in bona fide HIF Binding Regions

We employed the previous set of core, bona fide HBSs to computationally identify enriched TFBSs that could be indicative of transcription factor cooperation. To focus on binding locations for which there is evidence of transcriptional modulation of nearby genes in hypoxia, we sought to combine the core set of HIF1 binding locations with HIF transactivation data. To this end, we employed our previous microarray meta-analysis study [13] of 16 gene expression profiling experiments comparing normoxic and hypoxic cell cultures. This integrated gene-expression dataset was used to select, from the binding dataset, HIF1 binding locations that mapped close to genes showing robust hypoxic induction (p<0.02, false discovery rate) (Figure 3A, right). Lastly, and in order to reduce the number of spurious predictions in in silico sequence analyses, we focused on HIF binding sites whose sequence showed evolutionary conservation in mammalian species (Figure 3A). These selection criteria produced an integrated set of core HIF binding sites (Table S5). A gene annotation enrichment analysis of the sites in this integrated set revealed enriched annotation terms clearly associated with functional responses to hypoxia, such as glycolysis, 2-oxoglutarate dioxygenase activity and glycogen metabolism [7,58] (Table S6), strongly suggesting that it faithfully represents bona fide HIF binding locations.

We next sought to identify putative TFBS enriched in the vicinity of the selected core HIF binding sites (Figure 3B). For this purpose,
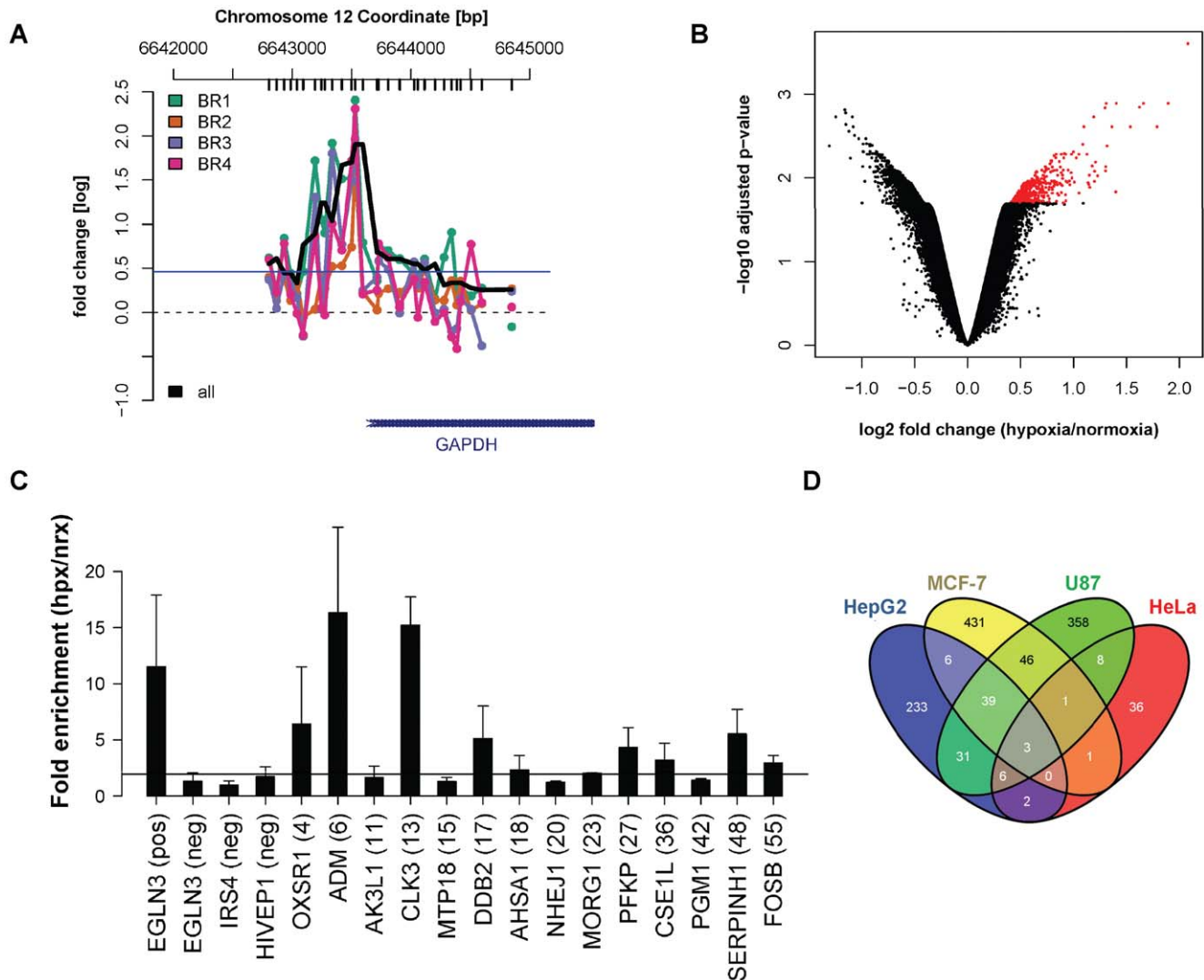
Figure 2. Comparative analysis of HIF1A ChIP-chip data in cell lines of different tissue origin. (A) Visualization of HIF1A ChIP-chip readouts for the GAPDH promoter region in HeLa cells. The plot represents normalized intensities (log fold change hypoxia/normoxia, vertical axis) along the hg19 genomic coordinate (horizontal axis). Vertical black bars (top of the graph) mark the center position of array probes. The signal of four independent biological replicates is indicated in different colors (BR1 to 4). The smoothed black line corresponds to the averaged signal across replicates. The horizontal blue line indicates the intensity threshold for bound probes. (B) Volcano plot of HeLa ChIP-chip data (linear model across the four biological replicates). Spots in the plot correspond to individual probes in the array. Probes significantly enriched in hypoxic samples (p<0.02, FDR) are highlighted as red spots. (C) Quantitative PCR validation of HeLa ChIP-chip candidates. The official Gene Symbol corresponding to each region is indicated in the horizontal axis. Four control regions (one positive and three negative) were used as reference to estimate the successful validation of the indicated ChIP-chip candidates (candidate ranks in parenthesis). Bars represent the average fold enrichment in hypoxic versus normoxic ChIP samples (vertical axis), as obtained in three independent experiments. Error bars represent the standard deviation. The horizontal black line indicates the threshold for successful validation (90% confidence interval). (D) 4-way Venn diagram indicating the overlap of HIF1A bound regions as reported by ChIP-chip studies in hepatoma (HepG2 cells, ref. 9), mammary gland (MCF-7 cells, ref. 10), glioma (U87 cells, ref. 11) and cervix (HeLa cells, this study) origin.
doi:10.1371/journal.pone.0045708.g002

we obtained the sequences flanking each HBS (Table S5). The length of flanking non-coding sequences was based on evolutionary conservation, as indicated by genomic annotation of PhastCons elements [47]. Statistical assessment of sequence motif enrichment in this set of sequences requires comparison with a background set, the election of this set greatly influencing the results of the analysis [59]. We constructed a set of sequences resembling those in the set of core HIF binding regions by screening the non-coding Human Genome for evolutionarily conserved HIF binding consensus sequences, and extended these motifs to span the flanking conserved sequence (Figure 3B and Figure S1). From this set, we selected

regions that are unlikely to be transcriptionally modulated by hypoxia, as judged by no differential expression in any of the 16 hypoxia experiments included in our previously reported genome profiling meta-analysis [13]. Finally, a subset from these sequences was chosen that matched the genomic locations and base composition found in the core set (Figure S1). We thereby obtained a custom set of circa 3500 background sequences containing a RCGTG HIF binding consensus.

The sequences in the core HIF binding regions and background sets were screened for TFBSs employing the mammalian position-weight matrixes from the public Transfac 7.0 and Jaspar (2010
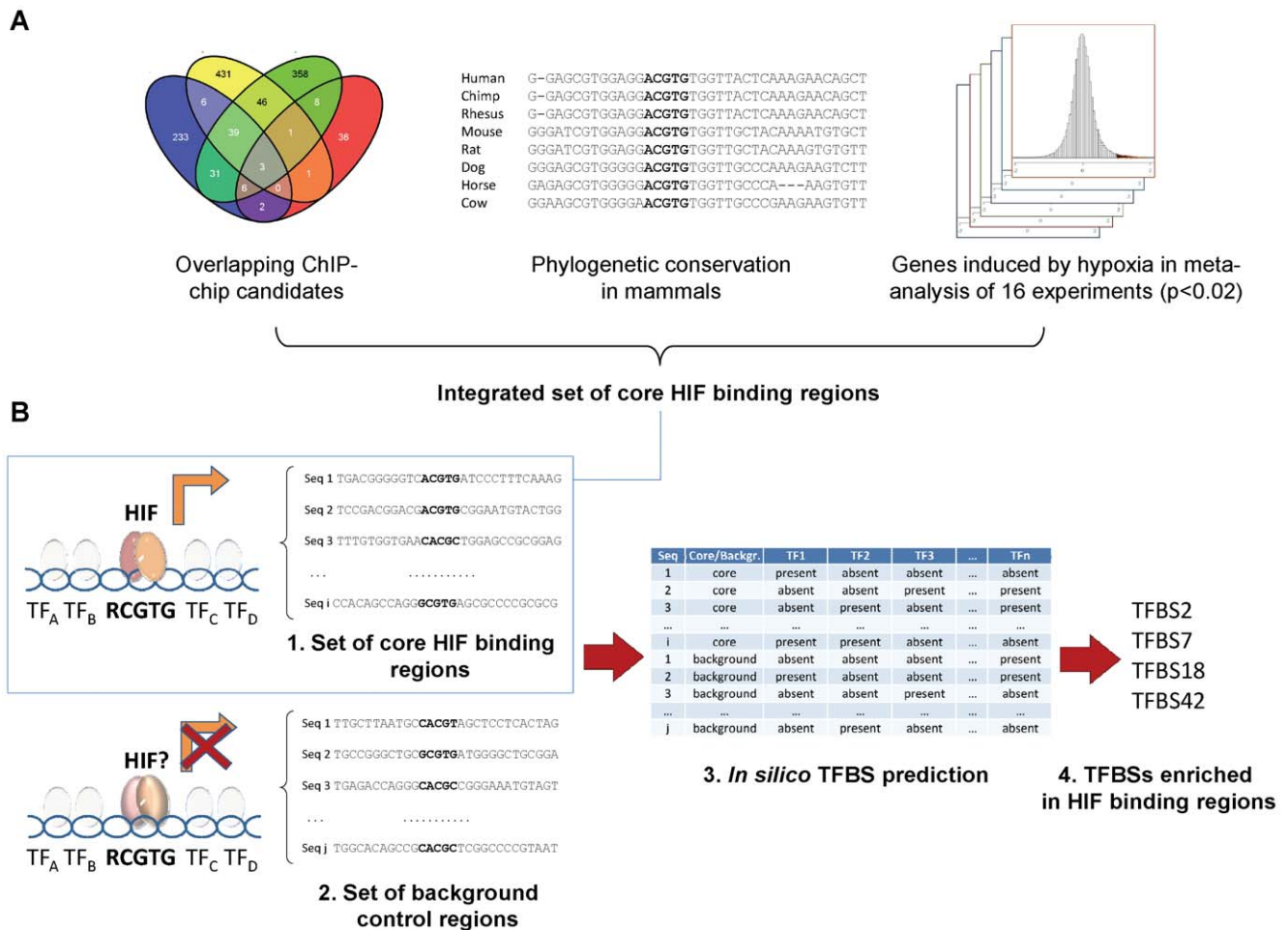
Figure 3. Integrative strategy for prediction of cooperativity in HIF binding regions. (**A**) HIF1 binding locations common to at least two out of four different ChIP-chip studies in HeLa, HepG2, MCF-7 and U87 cells (left), mammalian sequence conservation of the HIF binding regions (center) and regions close to genes robustly induced in hypoxia in a meta-analysis of 16 gene expression experiments (right) were integrated into a set of *bona-fide* core HIF binding regions (**B**) Stepwise diagram for prediction and validation of TFBSs enriched in core HIF binding regions: collection of a set of core HIF binding regions and a background set of control sequences (left), *in silico* prediction of transcription factor binding sites present or absent in the sequences of core and background sets (center), statistical analyses of enriched TFBSs in sequences from the core set (right, top) and experimental validation of these predictions (right, bottom).
doi:10.1371/journal.pone.0045708.g003

release) databases and custom scripts based on the MATCH algorithm [52], recording the presence or absence of a total of 605 sequences motifs in each sequence by using three different stringencies (Table 1). In order to reduce spurious hits, we only considered as positive hits those motifs that were conserved in mammalian species. Fisher's exact test was applied to these datasets to identify motif predictions enriched in the set of core HIF binding regions. Enriched motifs were consistently found across different stringencies and database sets. In addition to HIF PWMs, we found a significant enrichment for PWMs associated to CREB1, FOS/AP1 and NFY (Table 1).

As an independent assessment of enriched motifs that is less dependent on the composition of the core set, we compared the results of the previous analysis with a variable selection approach implemented in the Weka machine learning software [53]. Correlation-based feature selection was applied to the complete set of high-stringency predictions to detect non-redundant variables (PWMs) able to distinguish between the core and background sets. As expected, a number of the top-ranked PWMs, such as those for HIF1, AP1/ATF3 or NFY were coincident with the Fisher's exact

test predictions (Table 2). However, additional enriched motifs were found (such as CEBPB or NFAT), probably reflecting an increased predictive power after stratified cross-validation.

We next asked whether the TFs associated to the enriched TFBSs may share any common characteristics. Gene annotation enrichment analysis (Table S7) of these enriched transcription factors pointed at stimulus-responsive transcription factors as significantly enriched in core HIF binding regions, and indeed most of the identified DNA-binding proteins have been reported to function as transcription factors of stress responses [60], including hypoxia-responsive TFs [61]. On the whole, our results suggest that binding sequences of several additional TFs other than HIFs, and in particular diverse stress-responsive TFs, are enriched in *bona fide* HIF binding regions.

4. Functional Impact of Transcription Factor Binding Sites Proximal to Hypoxia Response Elements

In order to address the functional relevance of the enriched TFBS identified *in silico*, we next set out to validate some of these predictions experimentally. To this end, we selected hits for

Table 1. Enriched TFBSs in core HIF binding regions (Fisher's exact test).

PWM Collection	Stringency	Overrepresented PWM	Hits	Transcription factor	P value
JASPAR CORE 2010	minFN (low)	MA0033.1_FOXL1	53	FOXL1	0,001
		MA0259.1_HIF1A::ARNT	54	**HIF1**	0,0076
	minFP (high)	MA0018.2_CREB1	7	**CREB1**	0,0203
		MA0060.1_NFYA	9	**NFYA**	0,0234
		MA0259.1_HIF1A::ARNT	44	**HIF1**	6E−15
	minSum (intermediate)	MA0032.1_FOXC1	36	FOXC1	0,0065
		MA0060.1_NFYA	14	**NFYA**	0,0218
		MA0099.1_Fos	20	**FOS**	0,0305
		MA0259.1_HIF1A::ARNT	52	**HIF1**	3E−06
JASPAR PHYLOFACTS 2010	minFN (low)	PF0014_TGACGTCA	4	**FOS/AP1**	0,0377
		PF0032_TGASTMAGC	3	NF-E2	0,0268
	minFP (high)	PF0014_TGACGTCA	4	**FOS/AP1**	0,0445
		PF0032_TGASTMAGC	3	NF-E2	0,0308
	minSum (intermediate)	PF0014_TGACGTCA	4	**FOS/AP1**	0,0383
		PF0032_TGASTMAGC	3	NF-E2	0,0272
TRANSFAC 7.0	minFN (low)	M00055_V$NMYC_01	34	NMYC	0,0174
		M00244_V$NGFIC_01	5	NGFIC	0,0425
		M00246_V$EGR2_01	5	EGR2	0,0499
		M00251_V$XBP1_01	19	XBP1	0,02
	minFP (high)	M00185_V$NFY_Q6	6	**NFY**	0,0284
		M00188_V$AP1_Q4	7	**AP1**	0,0096
	minSum (intermediate)	M00040_V$CREBP1_01	6	**CREBP1**	0,0465
		M00185_V$NFY_Q6	14	**NFY**	0,0244
		M00244_V$NGFIC_01	5	NGFIC	0,0437
		M00287_V$NFY_01	14	**NFY**	0,0248
		M00394_V$MSX1_01	16	MSX1	0,0229

Enriched sequence motifs in core HIF binding regions, as indicated by statistical analysis (Fisher's exact test, $p < 0.05$ with no correction for multiple comparisons). For each overrepresented sequence motif/PWM, the table indicates the following: the database collection (PWM collection), the stringency used in *in silico* TFBS identification (Stringency), the number of hits obtained in the set of core HBRs (Hits), the transcription factor (Tr. Factor) associated to the PWM and the p value of the enrichment (p value). Robust predictions across different stringencies and PWM datasets are highlighted in bold.
doi:10.1371/journal.pone.0045708.t001

enriched TFBSs focusing on: 1) HREs located close to the TSSs of genes, to be able to study these promoter regions in *in cellulo* reporter assays, and 2) TFBS predictions located close to the Hypoxia Response Element (HRE). According to these criteria, we evaluated a CREB binding site prediction in the *LDHA* promoter (Figure 4A), a CEBPB binding site identified *in silico* in the *GYS1* promoter (Figure 4B), and a predicted AP1 site in the *CA9* promoter (Figure 4C). The selected promoters were cloned upstream of a firefly luciferase gene, either in their wild-type version or harbouring mutations in the predicted TFBSs. We then compared the effect of these mutations with that of the hypoxia response element (HRE). Finally, and in order to evaluate non-specific effects of the introduced changes, we also generated control mutations in these promoters by altering randomly-selected DNA motifs in the vicinity of the HRE (Figure 4). These control mutations lay in non-conserved (LDHA and GYS1 CONTROL 2) as well as conserved (CA9 and GYS1 CONTROL 1) genomic regions.

We next measured the luciferase activity of these constructs in normoxia, hypoxia (1% oxygen) and upon treatment with the prolyl hydroxylase inhibitor DMOG (500 μM). As expected, mutation of the HRE in all the studied promoters completely

abrogated induction by either hypoxia or DMOG treatment (Figure 4). Importantly, mutation of the predicted CREB site in the proximity of the HRE within the *LDHA* promoter led to a partial reduction in the inducibility of the construct, while introduction of a random mutation had a negligible effect in the response of the promoter to either hypoxia or DMOG. Similarly, mutation of the CEBPB binding site proximal to the *GYS1* HRE led to a partial abrogation of the hypoxic induction when compared to mutation of the HRE core (Figure 4B). This reduction was not consistently recapitulated when two distinct control mutations were introduced in the promoter (Figure 4B, Control 1 and Control 2), strongly suggesting that it is a specific effect. Importantly, similar results were obtained upon DMOG treatment (Figure 4B). Finally, in contrast to the two previous cases, directed mutagenesis of the AP1 site proximal to the *CA9* HRE led to slightly increased inducibility of the construct by either hypoxia or DMOG (Figure 4C), reaching statistical significance only for the latter. This effect was distinct from that of a control mutation or the expected abrogation of the induction produced by mutation of the HRE.

Collectively, these results indicate that at least some of the TFBSs computationally predicted as enriched in a core set of *bona*

Table 2. Enriched TFBSs in core HIF binding regions (variable selection).

Overrepresented PWM	Number chosen	Transcription factor
MA0259.1	100	**HIF1**
PF0146	86	unknown (RRCCGTTA)
PF0032	82	NFE-2
M00222	80	Hand1:E47
M00188	77	**AP1**
PF0096	67	unknown (YGCANTGCR)
PF0014	58	**ATF3**
M00109	57	CEBPB
M00185	45	**NFY**
M00246	36	EGR2
M00244	20	NGFIC
MA0154.1	15	EBF1
PF0009	13	**ATF3**
M00302	11	NFAT
M00002	10	E47

The Table indicates sequence motifs/PWMs identified as discriminative of core HBRs employing correlation-based feature selection in 10 iterations of 10-fold stratified cross-validation. The results are ranked according to the total number of folds (up to a hundred) in which the variable was chosen by the algorithm (Number chosen). The associated transcription factor, were known, is indicated along with the PWM (Tr. factor). Predictions coincident with Fisher's exact test (Table 1) are highlighted in bold.
doi:10.1371/journal.pone.0045708.t002

fide HIF binding regions play a functional role in the transactivation by hypoxia or DMOG treatment of HIF-responsive promoters. Moreover, our data also suggests that diverse stress-responsive transcription factors, probably through modulation of basal transcription or recruitment of common cofactors, contribute to the specification of HIF target selectivity.

5. Transcriptional Targets of Stress-responsive Transcription Factors are Enriched Among HIF Target Genes

The results in the previous section were restricted to a limited set of validated promoters. However, if the involved transcription factors are of general relevance to HIF mediated transcription, some of their targets would be expected to be common with HIFs. To judge the potential generality of the involvement of CEBPs, CREB and AP1 in modulation induction of HIF transcriptional targets, we employed a gene-set enrichment analysis (GSEA) [55] as an unbiased way to explore the distribution of other transcription factor targets among hypoxia inducible genes. For this analysis, we employed a list of over 11000 genes sorted according to their response to hypoxia, and derived from our previous meta-analysis of gene expression profiles [13]. This sorted list was then queried against the curated collection (C2) of the GSEA molecular signatures database (http://www.broadinstitute.org/gsea/index.jsp). This collection comprises over 3000 gene sets from various sources including experimentally derived lists of targets for specific transcription factors. Thus, this analysis identifies sets of functionally related genes, such as those co-regulated in response to specific genetic and chemical perturbations, that are significantly enriched in the top positions of a list of genes induced by hypoxia.

As expected, GSEA analysis revealed a statistically significant enrichment of well-characterized HIF targets in this sorted list (Figure 5A, black circles). Moreover, enrichment of CEBPA/B targets was also significant for three different gene-sets (Figure 5A, purple circles, and Figure 5B). These gene sets derive from independent expression profiling experiments performed in cells overexpressing different members of the CEBP family [62–64]. In addition, the analysis also revealed enrichment for two gene sets containing targets regulated by the ATF/CREB family (Figure 5A, orange circles), albeit the FDR-adjusted p-values did not reach statistical significance (0.106 and 0.229 respectively). Finally, enrichment of AP1 targets was not statistically significant (Figure 5A, green circle). Altogether, these results suggest, at least for the case of CEBPs (Figure 5B), that transcription factor collaboration can be a general mechanism contributing to HIF target selectivity.

Discussion

The complete elucidation of the molecular principles governing the translation of genomic information to gene regulation remains a central question in biology. In particular, understanding the mechanisms dictating target selection by HIF transcription factors is of fundamental importance to truly dissect the genes directly modulated by HIFs, and therefore to completely characterize the transcriptional response to hypoxia that these factors orchestrate, and its interactions with other transcriptional pathways. Several mechanisms have been proposed to contribute to selective DNA binding and gene regulation by transcription factors with largely generic DNA binding domains [65], among them the co-binding of several transcription factor molecules [14,66,67]. In order to dissect these mechanisms, high-quality collections of binding sites are an obvious pre-requisite. The recent development of high-throughput chromatin immunoprecipitation experiments [41,68] has spurred knowledge on the genome-wide DNA binding locations of transcription factors, and these techniques hence constitute an essential tool to explore mechanisms of transcriptional regulation on a global scale [69–71]. In this work, we employed an integrative approach to identify additional transcription factors that could contribute to HIFs binding and target selectivity. This strategy was based on computational prediction of enriched sequence motifs in a set of core HIF binding regions constructed through selection of HIF1 alpha binding locations derived from genome-wide chromatin immunoprecipitation experiments in HeLa (this study), HepG2 [9], MCF-7 [10] and U87 cells [11]. During preparation of this manuscript, a fourth study employing ChIP-Seq in MCF-7 cells was published [12], providing high-resolution data on genome-wide HIF binding locations independently of gene architecture.

Chromatin accessibility has been shown to play an important role in dictating transcription factor binding [72–74]. In this regard, integration of HIF1 alpha binding locations in U87 and HepG2 cells with gene expression data in the same cell types revealed a preference for HIF1 binding to map to transcriptionally active genes in normoxia [11], therefore suggesting that chromatin accessibility, as indirectly evidenced by basal transcriptional activity, determines HIF1 binding. As an independent approach to test this hypothesis, we looked at the correlation of normoxic gene expression and induction of known HIF targets in publicly available microarray datasets of hypoxic cell cultures [13]. In agreement, we found a statistically significant association between basal expression and hypoxia inducibility of known targets (Figure 1A and Table S2). Furthermore, comparison of HIF1α and HIF2α binding locations in MCF-7 cells with DNAse
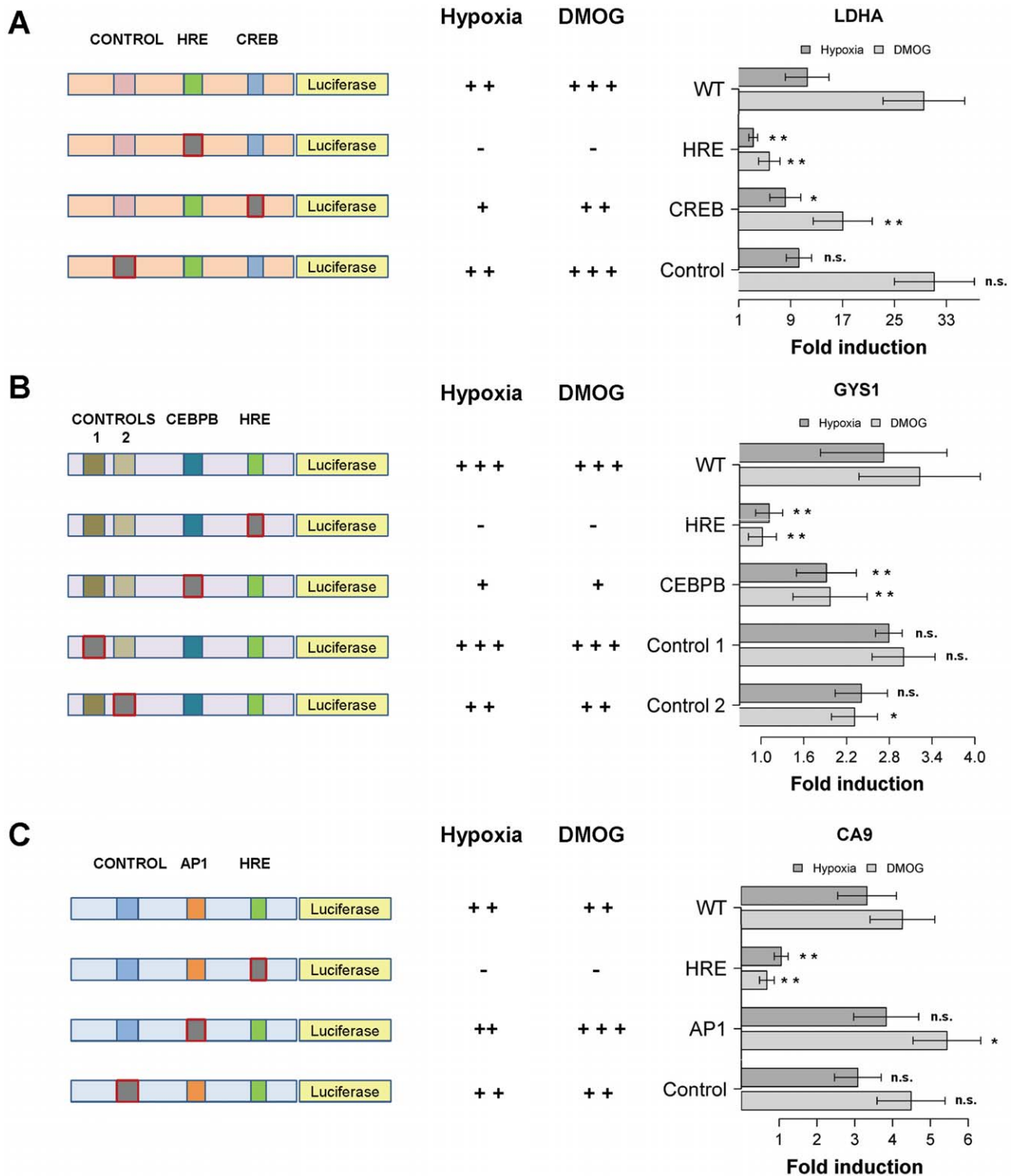
Figure 4. Effect of stress-responsive transcription factor binding sites in the proximity of hypoxia-response elements on hypoxic induction of HIF-responsive promoters. HeLa cells were transfected with reporter plasmids containing promoter regions of mouse LDHA (**A**), human GYS1 (**B**) and human CA9 (**C**) in their wild-type form or harbouring the indicated mutations. Diagrams to the left of each graph indicate the location of the different mutations in the employed promoter constructs (grey blocks, highlighted with red border). Effects on reporter induction by hypoxia and the hypoxia mimetic DMOG are summarized in the central columns: -, no hypoxic/DMOG induction; +, ++, +++: increasing strength of hypoxic/DMOG induction. Graphs represent the fold induction over normoxia of the wild-type promoter (WT) upon hypoxia or DMOG treatment, compared to that of promoter versions harbouring mutations in the hypoxia response element (HRE), in CREB (**A**), CEBPB (**B**), or AP1 (**C**) binding sites proximal to the HRE, or in control genomic regions (CONTROL). Bars represent average values in four to six independent experiments, and error bars the standard deviation. Statistical significance of observed activity compared to the wild-type promoters are indicated: n.s.: not significant, *: p<0.05, **: p<0.01 (repeated measures ANOVA with Dunnet post-hoc correction).
doi:10.1371/journal.pone.0045708.g004
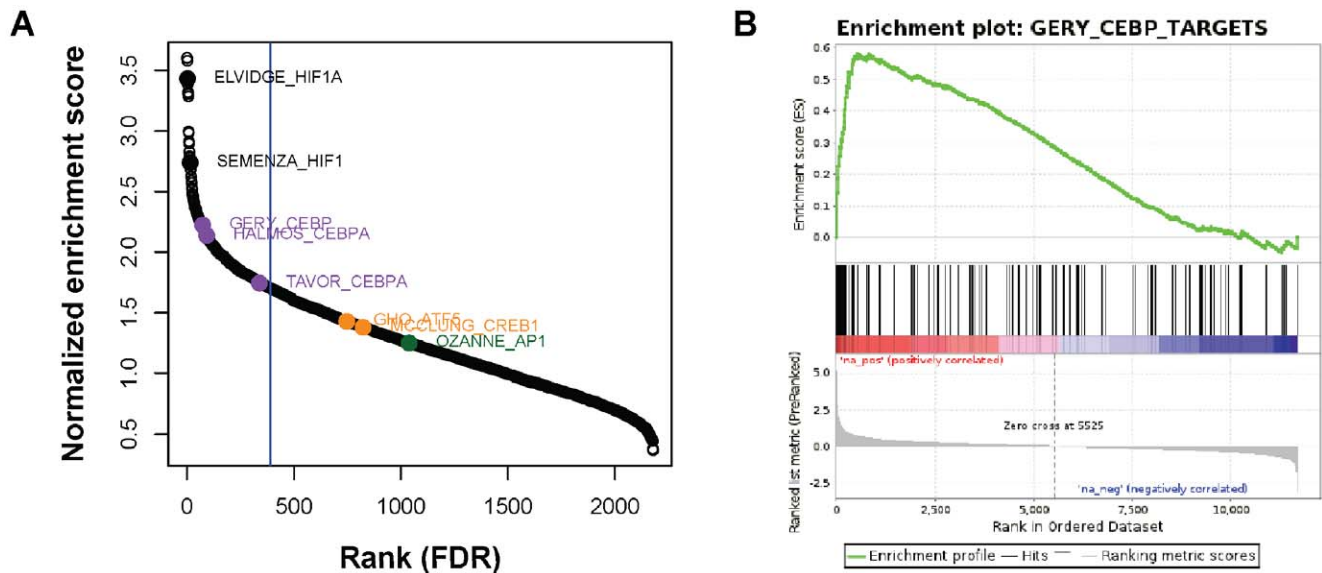
Figure 5. Transcriptional targets of stress-responsive transcription factors are enriched among core hypoxia-responsive genes. (**A**) Gene-set enrichment analysis on a set of 11673 genes sorted by their response to hypoxia according to a meta-analysis of hypoxia gene expression experiments (ref. 13). The graph depicts the normalized enrichment score of 3174 gene sets from the curated collection (C2) of the GSEA molecular signatures database v3.0, that includes sets of transcription factor target genes. Solid circles highlight gene-sets derived from studies on HIF1 (black, ELVIDGE_HYPOXIA_UP and SEMENZA_HIF1_TARGETS), CEBPA/B (purple, GERY_CEBP_TARGETS, HALMOS_CEBPA_TARGETS_UP and TAVOR_CEBPA_- TARGETS_UP), CREB1/ATF5 (orange, GHO_ATF5_TARGETS_DN and MCCLUNG_CREB1_TARGETS_UP) and AP1 (green, OZANNE_AP1_TARGETS_UP) transcriptional targets. The vertical blue line corresponds to an FDR-adjusted p-value of 0.05. (**B**) GSEA analysis of hypoxia-responsive genes (see A) against the GERY_CEBPA_TARGETS (M12338, derived from the GEO dataset GSE2188) gene-set. Hypoxic response is rank-ordered in the horizontal axis (Rank in ordered dataset). Black bars indicate the position of individual targets in the CEBPA gene-set. The graph on top (green curve) represents enrichment scores of CEBPA targets across hypoxia responsive genes, indicating positive correlation between the two. The gradient color bar indicates positive (red) and negative (blue) correlation boundaries.
doi:10.1371/journal.pone.0045708.g005

hypersentitivity data in the same cell type [12] also revealed a significant association of HIF binding with normoxic DNAse hypersensitive sites, again pointing at an important role of open chromatin regions in dicating HIF binding. However, when conserved RCGTG HIF binding consensus motifs are identified in non-coding regions of genes showing basal expression, a majority of these are not induced by hypoxia (Figure 3B). Therefore, although chromatin accessibility clearly favors HIF1 binding, additional mechanisms are likely needed to fully specify HIF target selectivity.

DNA methylation of a HIF binding site was originally shown to block HIF1α binding to the 3′ erythropoietin enhancer [16], and indeed erythropoietin expression appears to be restricted to cell types in which the hypoxia response element is unmethylated. Altered HIF binding due to methylation changes in HREs has been further validated in additional target genes, such as *BNIP3* [75] or *HIF1A* [76], and is often associated with cancer progression. However, a global view on the effects of DNA methylation in HIF binding selectivity is lacking, and may be challenging to analyze in view of recent evidence arguing for dynamic DNA methylation in hypoxia [77].

Additional transcription factors binding in the proximity of a HIF1 binding site could impact either HIF1 binding or transcriptional modulation of the target gene. In agreement with this possibility, a recent study addressing the functional validation of common genetic variants at a renal cancer susceptibility locus [78] found HIF2 binding to be dependent on a polymorphism falling outside the RCGTG HIF binding consensus, strongly suggesting that sequences outside the HIF binding site can be functionally important in determining HIF binding. We tested this hypothesis by computational prediction of transcription factor

binding sites enriched in a core set of *bona fide* HIF binding regions (Figure 3B). These were obtained through integration of HIF1α ChIP-chip data with a gene expression meta-analysis of hypoxic cell cultures [13] (Figure 3A), thereby combining multiple HIF DNA binding and hypoxic gene expression datasets. Our approach has the advantage of using an integrated set of sequences that could overcome the limitations of analyses based on a single dataset [10,30], where a proportion of binding sites could potentially correspond to false positives or non-functional sites. In addition to HIF matrices, we observed additional sequence motifs that were enriched in core HIF binding regions (Tables 1 and 2) and that could potentially impact HIF binding and transactivation selectivity. Of note, the transcriptional activity of several of these proteins, such as AP-1, CREB, EGR-2 or CEBPB is known to be induced by hypoxia [61]. Nevertheless, and in agreement with previous predictions of enriched TFBSs in the vicinity of experimentally [10] or computationally [30] identified HIF binding sites, the statistical significance of these predictions is relatively low and, even on an integrated dataset, no single collaborating TF stands out. In fact, HIF PWMs are the most enriched in core HIF binding regions. Since sequences in the background set used for comparison also contain RCGTG motifs (Figure S1), this enrichment likely arises from the well known preference for A versus G in the first position of the HIF binding consensus. These results collectively suggest that several additional transcription factors could influence HIF transcriptional activity.

Importantly, we noted that most of the enriched TFBSs corresponded to stress-responsive transcription factors. Varied stress-responsive TFs have been shown to coordinately regulate the same genes [23,79], and indeed several transcription factors are activated by the same stresses in mammalian cells [61,80,81].

However, it is unclear whether this cooperation among stress-responsive pathways translates at the genomic level. In order to evaluate the functional significance of the TFBSs enriched in core HIF binding regions, we carried out an experimental validation by disrupting selected sequences in *bona fide* HIF-responsive promoters (Figure 4). Importantly, no experimental confirmation had been attempted on previously reported predictions [10,30], and therefore the biological significance of those findings remained unclear. In spite of being limited to three selected promoters, our results clearly indicate that, compared to control mutations, alteration of binding sequences of transcription factors enriched in HIF binding regions, and different from HIFs themselves, have a specific effect on the transcriptional activation of HIF-responsive promoters. In particular, we found negative effects on hypoxic induction of *LDHA* and *GYS1* promoters upon disruption of CREB and CEBPB binding sites proximal to the HRE, whereas mutation of an AP1 site proximal to the *CA9* HRE led to a slightly augmented hypoxic induction of the promoter. In agreement with our results, mutation of the same CREB binding site was been previously shown to alter *LDHA* hypoxic induction [26]. Interestingly, USF binding to a palindrome CACGTG HRE in the *LDHA* promoter was suggested to complement HIF binding [27]. However, our results do not allow us to corroborate these findings, as mutation of this HRE was not evaluated in our experiments (Materials and Methods and Table S1). Furthermore, hypoxic *CA9* expression has been linked to cooperation between AP1 family member ATF4 and HIF1α [82]. In this study, ATF4 over-expression led to an augmented *CA9* induction in hypoxia, with reduced hypoxic expression of *CA9* being observed upon ATF4 knock-down. Chromatin immunoprecipitation experiments mapped ATF4 binding to the $-1400/-1000$ region of the *CA9* promoter, which falls outside of the promoter region employed in our experiments. Nevertheless, the apparent paradox with our results argues for careful interpretation of the role of AP1 in the HIF transcriptional response. In fact, both positive and negative effects of AP1 have been reported on hypoxic gene expression [81] and, given the number of AP1 family members, these probably arise from compositional differences in AP1 complexes.

Importantly, the effects observed upon mutation of CREB, CEBPB or AP1 binding sites (Figure 4) were always moderate when compared to mutation of the HIF binding consensus RCGTG, suggesting that rather than being an absolute requirement for hypoxic induction, the integrity of these neighboring TFBSs fine-tunes the HIF-mediated transcriptional response. Thus, it is possible that multiple independent factors contribute, in an additive fashion, to HIF-mediated transcription. This model could also explain why we found a relatively large number of enriched TFBSs in HIF binding regions, but all of them sharing a modest statistical significance. On the whole, these observations indicate that several of the enriched TFBSs identified in our approach are of functional relevance for HIF-mediated transcription. Nevertheless, it should be noted that other TFs for which collaboration with HIFs has been previously suggested [61] are not recovered as enriched in our approach. An inherent assumption in our method is that such TFBSs will be enriched in bona fide HIF binding regions (Figure 3), which may not hold true if cooperation is specific to a small number of target genes. Furthermore, the employed HIF binding data in this study is for the HIF1α subunit only, whereas transcription factor cooperativity may well apply to other HIF subunits. In fact, several reports have implicated the ETS family of transcription factors in target selection by HIF2α [28,29].

We observed very similar tendencies when transcriptional activation of reporter constructs was elicited by DMOG or hypoxia treatment (Figure 4), additionally suggesting that, at least in our experimental conditions, the contribution of these factors could occur mainly in basal conditions, as it is unlikely that hypoxia and DMOG treatment induce completely overlapping cellular responses. Several recent reports [11,12] have suggested that chromatin accessibility determines HIF1 binding, although this mechanism may not fully explain HIFs binding and target selectivity (Figure 1). Our results indicate that an additional layer of specificity comes from proximal co-binding of other transcription factors and HIFs to open chromatin regions, thereby facilitating or restricting HIF-mediated transcription. Elucidation of the underlying molecular mechanisms falls outside the scope of our work, although it is tempting to speculate that transcription factors binding in proximity of HIFs may be involved in recruitment of co-activator or co-repressor proteins. Of note, a recent mammalian two-hybrid survey of protein-protein interactions for human and mouse TFs [83] reported a physical association between HIF1A and AP-1 family member JUN, as well as the previously known interaction between CEBPB and p300 [84]. p300/CBP is a master co-activator of HIF-mediated transcription whose recruitment can also be mediated by CREB [85]. In this regard, evidence from a synthetic transactivation screen on the EGLN1 promoter [86] pointed to ETV4 as an additional p300-dependent coactivator of HIF-mediated transcription. Moreover, HIF1 is known to interact with Jab1/CSN5 [87,88], a protein originally identified as a transcriptional co-activator for AP1 [89]. Future investigations on protein composition of HIF-bound enhancers should be pivotal in supporting this model.

The associations between HIFs and AP1, CREB and CEBPs analyzed in our reporter results could be general across many HIF targets or be restricted to individual targets. To judge the generality of these results, we conducted a gene-set enrichment analysis (GSEA) of transcription factor targets in a sorted list of genes regulated by hypoxia [13]. The results of this analysis showed a significant enrichment of CEBP targets among hypoxia-inducible genes (Figure 5), suggesting that at least for this family of transcription factors, the functional association with HIFs could be relatively general. Of note, recent works have reported a direct protein-protein interaction between HIF1α and CEBPα [90,91], and have implicated CEBPα activity in regulation of the HIF target genes galectin-1 [92] and PAI-1 [23]. Hypoxic induction of both *galectin-1* and *PAI-1* was found to be synergistically dependent on both HIF1α and CEBPα activity and their co-binding to the promoter region. Our results further suggest that this functional association may be general across a wider collection of HIF targets.

In conclusion, the data presented herein demonstrates that integration of high-throughput chromatin immunoprecipitation and gene expression data is a successful approach to select high-quality core HIF binding regions, and provides experimental proof of principle for the biological relevance of enriched transcription factor binding sites other than the HIF binding consensus in HIF-mediated transcription. Specifically, our results suggest that diverse stress-responsive transcription factors, in particular CEBPs, contribute to fine-tuning of the HIF-mediated transcriptional response.

Supporting Information

Figure S1 Construction of a background set of control sequences resembling core HIF binding regions.
(PDF)

Table S1 HRE, AP1, CREB, CEBPB binding sites and control mutations introduced in LDHA, CA9 and GYS1 promoter constructs.
(XLSX)

Table S2 Correlation between normoxic basal expression and hypoxia inducibility (known HIF targets).
(XLSX)

Table S3 Candidate HIF1alpha-binding regions identified by ChIP-chip analysis in HeLa cells.
(XLSX)

Table S4 Distribution of conserved HBSs in ChIP-chip results across different cell types.
(XLSX)

Table S5 Integrated set of core HIF binding regions.
(XLSX)

Table S6 Enriched gene annotation clusters in core HIF binding regions.
(XLSX)

Table S7 Enriched gene annotation terms in transcription factors associated to sequence motifs overrepresented in core HIF binding regions.
(XLSX)

Author Contributions

Conceived and designed the experiments: DV FSC HH PY LdP. Performed the experiments: DV AOB LGM NP. Analyzed the data: DV AOB LGM LdP. Contributed reagents/materials/analysis tools: HH BR ZT AD TH PY LdP. Wrote the paper: DV LdP.

References

1. Wang GL, Jiang BH, Rue EA, Semenza GL (1995) Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. Proc Natl Acad Sci U S A 92: 5510–5514.
2. Salceda S, Caro J (1997) Hypoxia-inducible factor 1alpha (HIF-1alpha) protein is rapidly degraded by the ubiquitin-proteasome system under normoxic conditions. Its stabilization by hypoxia depends on redox-induced changes. J Biol Chem 272: 22642–22647.
3. Mahon PC, Hirota K, Semenza GL (2001) FIH-1: a novel protein that interacts with HIF-1alpha and VHL to mediate repression of HIF-1 transcriptional activity. Genes Dev 15: 2675–2686.
4. Epstein AC, Gleadle JM, McNeill LA, Hewitson KS, O'Rourke J, et al. (2001) C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. Cell 107: 43–54.
5. Bruick RK, McKnight SL (2001) A conserved family of prolyl-4-hydroxylases that modify HIF. Science 294: 1337–1340.
6. Lando D, Peet DJ, Whelan DA, Gorman JJ, Whitelaw ML (2002) Asparagine hydroxylation of the HIF transactivation domain a hypoxic switch. Science 295: 858–861.
7. Wenger RH, Stiehl DP, Camenisch G (2005) Integration of oxygen signaling at the consensus HRE. Sci STKE 2005: re12.
8. Kaelin WG, Jr., Ratcliffe PJ (2008) Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway. Mol Cell 30: 393–402.
9. Xia X, Lemieux ME, Li W, Carroll JS, Brown M, et al. (2009) Integrative analysis of HIF binding and transactivation reveals its role in maintaining histone methylation homeostasis. Proc Natl Acad Sci U S A 106: 4260–4265.
10. Mole DR, Blancher C, Copley RR, Pollard PJ, Gleadle JM, et al. (2009) Genome-wide association of hypoxia-inducible factor (HIF)-1{alpha} and HIF-2{alpha} DNA binding with expression profiling of hypoxia-inducible transcripts. J Biol Chem 284: 16767–16775.
11. Xia X, Kung AL (2009) Preferential binding of HIF-1 to transcriptionally active loci determines cell-type specific response to hypoxia. Genome Biol 10: R113.
12. Schodel J, Oikonomopoulos S, Ragoussis J, Pugh CW, Ratcliffe PJ, et al. (2011) High-resolution genome-wide mapping of HIF-binding sites by ChIP-seq. Blood 117: e207–217.
13. 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: 2332–2345.
14. Pan Y, Tsai CJ, Ma B, Nussinov R (2010) Mechanisms of transcription factor selectivity. Trends Genet 26: 75–83.
15. Bell O, Tiwari VK, Thoma NH, Schubeler D (2011) Determinants and dynamics of genome accessibility. Nat Rev Genet 12: 554–564.
16. Wenger RH, Kvietikova I, Rolfs A, Camenisch G, Gassmann M (1998) Oxygen-regulated erythropoietin gene expression is dependent on a CpG methylation-free hypoxia-inducible factor-1 DNA-binding site. Eur J Biochem 253: 771–777.
17. Rossler J, Stolze I, Frede S, Freitag P, Schweigerer L, et al. (2004) Hypoxia-induced erythropoietin expression in human neuroblastoma requires a methyl-ation free HIF-1 binding site. J Cell Biochem 93: 153–161.
18. Panne D (2008) The enhanceosome. Curr Opin Struct Biol 18: 236–242.
19. Kulkarni MM, Arnosti DN (2003) Information display by transcriptional enhancers. Development 130: 6569–6575.
20. Segal E, Widom J (2009) From DNA sequence to transcriptional behaviour: a quantitative approach. Nat Rev Genet 10: 443–456.
21. Zhang W, Tsuchiya T, Yasukochi Y (1999) Transitional change in interaction between HIF-1 and HNF-4 in response to hypoxia. J Hum Genet 44: 293–299.
22. Yamashita K, Discher DJ, Hu J, Bishopric NH, Webster KA (2001) Molecular regulation of the endothelin-1 gene by hypoxia. Contributions of hypoxia-inducible factor-1, activator protein-1, GATA-2, AND p300/CBP. J Biol Chem 276: 12645–12653.
23. Liao H, Hyman MC, Lawrence DA, Pinsky DJ (2007) Molecular regulation of the PAI-1 gene by hypoxia: contributions of Egr-1, HIF-1alpha, and C/EBPalpha. Faseb J 21: 935–949.
24. Sanchez-Elsner T, Ramirez JR, Sanz-Rodriguez F, Varela E, Bernabeu C, et al. (2004) A cross-talk between hypoxia and TGF-beta orchestrates erythropoietin gene regulation through SP1 and Smads. J Mol Biol 336: 9–24.
25. Miki N, Ikuta M, Matsui T (2004) Hypoxia-induced activation of the retinoic acid receptor-related orphan receptor alpha4 gene by an interaction between hypoxia-inducible factor-1 and Sp1. J Biol Chem 279: 15025–15031.
26. Firth JD, Ebert BL, Ratcliffe PJ (1995) Hypoxic regulation of lactate dehydrogenase A. Interaction between hypoxia-inducible factor 1 and cAMP response elements. J Biol Chem 270: 21021–21027.
27. Hu J, Stiehl DP, Setzer C, Wichmann D, Shinde DA, et al. (2011) Interaction of HIF and USF signaling pathways in human genes flanked by hypoxia-response elements and E-box palindromes. Mol Cancer Res 9: 1520–1536.
28. Aprelikova O, Wood M, Tackett S, Chandramouli GV, Barrett JC (2006) Role of ETS transcription factors in the hypoxia-inducible factor-2 target gene selection. Cancer Res 66: 5641–5647.
29. Hu CJ, Sataur A, Wang L, Chen H, Simon MC (2007) The N-terminal transactivation domain confers target gene specificity of hypoxia-inducible factors HIF-1alpha and HIF-2alpha. Mol Biol Cell 18: 4528–4542.
30. Benita Y, Kikuchi H, Smith AD, Zhang MQ, Chung DC, et al. (2009) An integrative genomics approach identifies Hypoxia Inducible Factor-1 (HIF-1)-target genes that form the core response to hypoxia. Nucleic Acids Res 37: 4587–4602.
31. Hu CJ, Wang LY, Chodosh LA, Keith B, Simon MC (2003) Differential roles of hypoxia-inducible factor 1alpha (HIF-1alpha) and HIF-2alpha in hypoxic gene regulation. Mol Cell Biol 23: 9361–9374.
32. Vengellur A, Woods BG, Ryan HE, Johnson RS, LaPres JJ (2003) Gene expression profiling of the hypoxia signaling pathway in hypoxia-inducible factor 1alpha null mouse embryonic fibroblasts. Gene Expr 11: 181–197.
33. Greijer AE, van der Groep P, Kemming D, Shvarts A, Semenza GL, et al. (2005) Up-regulation of gene expression by hypoxia is mediated predominantly by hypoxia-inducible factor 1 (HIF-1). J Pathol 206: 291–304.
34. Chi JT, Wang Z, Nuyten DS, Rodriguez EH, Schaner ME, et al. (2006) Gene expression programs in response to hypoxia: cell type specificity and prognostic significance in human cancers. PLoS Med 3: e47.
35. Vengellur A, Phillips JM, Hogenesch JB, LaPres JJ (2005) Gene expression profiling of hypoxia signaling in human hepatocellular carcinoma cells. Physiol Genomics 22: 308–318.
36. Elvidge GP, Glenny L, Appelhoff RJ, Ratcliffe PJ, Ragoussis J, et al. (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: 15215–15226.

37. Katada K, Naito Y, Mizushima K, Takagi T, Handa O, et al. (2006) Gene expression profiles on hypoxia and reoxygenation in rat gastric epithelial cells: a high-density DNA microarray analysis. Digestion 73: 89–100.

38. Sung FL, Hui EP, Tao Q, Li H, Tsui NB, et al. (2007) Genome-wide expression analysis using microarray identified complex signaling pathways modulated by hypoxia in nasopharyngeal carcinoma. Cancer Lett 253: 74–88.

39. Edgar R, Domrachev M, Lash AE (2002) Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 30: 207–210.

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

41. Ren B, Robert F, Wyrick JJ, Aparicio O, Jennings EG, et al. (2000) Genome-wide location and function of DNA binding proteins. Science 290: 2306–2309.

42. Singer GA, Wu J, Yan P, Plass C, Huang TH, et al. (2008) Genome-wide analysis of alternative promoters of human genes using a custom promoter tiling array. BMC Genomics 9: 349.

43. Slater GS, Birney E (2005) Automated generation of heuristics for biological sequence comparison. BMC Bioinformatics 6: 31.

44. Toedling J, Skylar O, Krueger T, Fischer JJ, Sperling S, et al. (2007) Ringo–an R/Bioconductor package for analyzing ChIP-chip readouts. BMC Bioinformatics 8: 221.

45. Smyth GK (2005) Limma: linear models for microarray data. In: R. Gentleman V, Carey V, S R Irizarry, W Huber, editors. Bioinformatics and Computational Biology Solutions using R and Bioconductor. New York: Springer. 397–420.

46. Toedling J, Huber W (2008) Analyzing ChIP-chip data using bioconductor. PLoS Comput Biol 4: e1000227.

47. Siepel A, Bejerano G, Pedersen JS, Hinrichs AS, Hou M, et al. (2005) Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res 15: 1034–1050.

48. Rhead B, Karolchik D, Kuhn RM, Hinrichs AS, Zweig AS, et al. (2010) The UCSC Genome Browser database: update 2010. Nucleic Acids Res 38: D613–619.

49. Portales-Casamar E, Thongjuea S, Kwon AT, Arenillas D, Zhao X, et al. (2010) JASPAR 2010: the greatly expanded open-access database of transcription factor binding profiles. Nucleic Acids Res 38: D105–110.

50. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, et al. (2006) TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res 34: D108–110.

51. Stormo GD (2000) DNA binding sites: representation and discovery. Bioinformatics 16: 16–23.

52. Kel AE, Gossling E, Reuter I, Cheremushkin E, Kel-Margoulis OV, et al. (2003) MATCH: A tool for searching transcription factor binding sites in DNA sequences. Nucleic Acids Res 31: 3576–3579.

53. Frank E, Hall M, Trigg L, Holmes G, Witten IH (2004) Data mining in bioinformatics using Weka. Bioinformatics 20: 2479–2481.

54. Pescador N, Villar D, Cifuentes D, Garcia-Rocha M, Ortiz-Barahona A, et al. (2010) Hypoxia promotes glycogen accumulation through hypoxia inducible factor (HIF)-mediated induction of glycogen synthase 1. PLoS One 5: e9644.

55. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, et al. (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102: 15545–15550.

56. Hemberg M, Kreiman G (2011) Conservation of transcription factor binding events predicts gene expression across species. Nucleic Acids Res 39: 7092–7102.

57. He Q, Bardet AF, Patton B, Purvis J, Johnston J, et al. (2011) High conservation of transcription factor binding and evidence for combinatorial regulation across six Drosophila species. Nat Genet 43: 414–420.

58. Brahimi-Horn MC, Bellot G, Pouyssegur J (2011) Hypoxia and energetic tumour metabolism. Curr Opin Genet Dev 21: 67–72.

59. Ji H, Vokes SA, Wong WH (2006) A comparative analysis of genome-wide chromatin immunoprecipitation data for mammalian transcription factors. Nucleic Acids Res 34: e146.

60. Tacchini L, Fusar-Poli D, Bernelli-Zazzera A (2002) Activation of transcription factors by drugs inducing oxidative stress in rat liver. Biochem Pharmacol 63: 139–148.

61. Cummins EP, Taylor CT (2005) Hypoxia-responsive transcription factors. Pflugers Arch 450: 363–371.

62. Tavor S, Park DJ, Gery S, Vuong PT, Gombart AF, et al. (2003) Restoration of C/EBPalpha expression in a BCR-ABL+ cell line induces terminal granulocytic differentiation. J Biol Chem 278: 52651–52659.

63. Halmos B, Basseres DS, Monti S, D'Alo F, Dayaram T, et al. (2004) A transcriptional profiling study of CCAAT/enhancer binding protein targets identifies hepatocyte nuclear factor 3 beta as a novel tumor suppressor in lung cancer. Cancer Res 64: 4137–4147.

64. Gery S, Gombart AF, Yi WS, Koeffler C, Hofmann WK, et al. (2005) Transcription profiling of C/EBP targets identifies Per2 as a gene implicated in myeloid leukemia. Blood 106: 2827–2836.

65. Georges AB, Benayoun BA, Caburet S, Veitia RA (2010) Generic binding sites, generic DNA-binding domains: where does specific promoter recognition come from? Faseb J 24: 346–356.

66. Remenyi A, Scholer HR, Wilmanns M (2004) Combinatorial control of gene expression. Nat Struct Mol Biol 11: 812–815.

67. Tomancak P, Ohler U (2010) Mapping the complexity of transcription control in higher eukaryotes. Genome Biol 11: 115.

68. Robertson G, Hirst M, Bainbridge M, Bilenky M, Zhao Y, et al. (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. Nat Methods 4: 651–657.

69. Lefterova MI, Zhang Y, Steger DJ, Schupp M, Schug J, et al. (2008) PPARgamma and C/EBP factors orchestrate adipocyte biology via adjacent binding on a genome-wide scale. Genes Dev 22: 2941–2952.

70. Farnham PJ (2009) Insights from genomic profiling of transcription factors. Nat Rev Genet 10: 605–616.

71. Lupien M, Eeckhoute J, Meyer CA, Wang Q, Zhang Y, et al. (2008) FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription. Cell 132: 958–970.

72. Kaplan T, Biggin MD (2012) Quantitative models of the mechanisms that control genome-wide patterns of animal transcription factor binding. Methods Cell Biol 110: 263–283.

73. Pique-Regi R, Degner JF, Pai AA, Gaffney DJ, Gilad Y, et al. (2011) Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. Genome Res 21: 447–455.

74. Degner JF, Pai AA, Pique-Regi R, Veyrieras JB, Gaffney DJ, et al. (2012) DNase I sensitivity QTLs are a major determinant of human expression variation. Nature 482: 390–394.

75. Bacon AL, Fox S, Turley H, Harris AL (2007) Selective silencing of the hypoxia-inducible factor 1 target gene BNIP3 by histone deacetylation and methylation in colorectal cancer. Oncogene 26: 132–141.

76. Koslowski M, Luxemburger U, Tureci O, Sahin U (2011) Tumor-associated CpG demethylation augments hypoxia-induced effects by positive autoregulation of HIF-1alpha. Oncogene 30: 876–882.

77. Liu Q, Liu L, Zhao Y, Zhang J, Wang D, et al. (2011) Hypoxia induces genomic DNA demethylation through the activation of HIF-1alpha and transcriptional upregulation of MAT2A in hepatoma cells. Mol Cancer Ther 10: 1113–1123.

78. Schodel J, Bardella C, Sciesielski LK, Brown JM, Pugh CW, et al. (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: 420–425, S421–422.

79. Alam J, Cook JL (2007) How many transcription factors does it take to turn on the heme oxygenase-1 gene? Am J Respir Cell Mol Biol 36: 166–174.

80. Malhi H, Kaufman RJ (2011) Endoplasmic reticulum stress in liver disease. J Hepatol 54: 795–809.

81. Laderoute KR (2005) The interaction between HIF-1 and AP-1 transcription factors in response to low oxygen. Semin Cell Dev Biol 16: 502–513.

82. van den Beucken T, Koritzinsky M, Niessen H, Dubois L, Savelkouls K, et al. (2009) Hypoxia-induced expression of carbonic anhydrase 9 is dependent on the unfolded protein response. J Biol Chem 284: 24204–24212.

83. Ravasi T, Suzuki H, Cannistraci CV, Katayama S, Bajic VB, et al. (2010) An atlas of combinatorial transcriptional regulation in mouse and man. Cell 140: 744–752.

84. Kovacs KA, Steinmann M, Magistretti PJ, Halfon O, Cardinaux JR (2003) CCAAT/enhancer-binding protein family members recruit the coactivator CREB-binding protein and trigger its phosphorylation. J Biol Chem 278: 36959–36965.

85. Chrivia JC, Kwok RP, Lamb N, Hagiwara M, Montminy MR, et al. (1993) Phosphorylated CREB binds specifically to the nuclear protein CBP. Nature 365: 855–859.

86. Wollenick K, Hu J, Kristiansen G, Schraml P, Rehrauer H, et al. (2012) Synthetic transactivation screening reveals ETV4 as broad coactivator of hypoxia-inducible factor signaling. Nucleic Acids Res 40: 1928–1943.

87. Bae MK, Ahn MY, Jeong JW, Bae MH, Lee YM, et al. (2002) Jab1 interacts directly with HIF-1alpha and regulates its stability. J Biol Chem 277: 9–12.

88. Bemis L, Chan DA, Finkielstein CV, Qi L, Sutphin PD, et al. (2004) Distinct aerobic and hypoxic mechanisms of HIF-alpha regulation by CSN5. Genes Dev 18: 739–744.

89. Claret FX, Hibi M, Dhut S, Toda T, Karin M (1996) A new group of conserved coactivators that increase the specificity of AP-1 transcription factors. Nature 383: 453–457.

90. Jiang Y, Xue ZH, Shen WZ, Du KM, Yan H, et al. (2005) Desferrioxamine induces leukemic cell differentiation potentially by hypoxia-inducible factor-1 alpha that augments transcriptional activity of CCAAT/enhancer-binding protein-alpha. Leukemia 19: 1239–1247.

91. Yang L, Jiang Y, Wu SF, Zhou MY, Wu YL, et al. (2008) CCAAT/enhancer-binding protein alpha antagonizes transcriptional activity of hypoxia-inducible factor 1 alpha with direct protein-protein interaction. Carcinogenesis 29: 291–298.

92. Zhao XY, Zhao KW, Jiang Y, Zhao M, Chen GQ (2011) Synergistic induction of galectin-1 by CCAAT/enhancer binding protein alpha and hypoxia-inducible factor 1alpha and its role in differentiation of acute myeloid leukemic cells. J Biol Chem 286: 36808–36819.