CHARACTERIZATION OF NOVEL HHEX PARTNERS: SOX13 AND C-MYC

-New mechanism for the regulation of Wnt/TCF and c-Myc pathways-

Vanessa Marfil Vives

PhD supervisor Roque Bort
PhD tutor Francisco X. Real

Thesis submitted in accordance with the requirements for the degree of PhD in Biology, Universitat Pompeu Fabra.

Barcelona, May 2010

The research in this thesis has been carried out at the Institut Municipal d'Investigacions Mèdiques (IMIM) at the Parc de Recerca Biomèdica de Barcelona (PRBB), and in the Unidad de Hepatología Experimental at the Hospital de la Fe (Valencia).

A la meva iaia, la millor persona que he conegut mai.

"Life is what happens to you when you are busy making other plans" John Lennon

ACKNOWLEDGEMENTS

Va ser a primer de Batxillerat quan la meva professora de Biologia em va encomanar la seva passió per la ciència. Després al voler conèixer més aquest món vaig fer bioquímica, i aviat ja vaig saber que volia fer una tesis doctoral i seguir aquest camí. Ara bé, si m'haguessin dit el dia que vaig començar la tesis que durant el transcurs d'aquesta els meus dos directors de tesis es mudarien, que hauria de marxar de la meva ciutat i que faria dues estàncies fora del meu país, no m'ho hagués cregut. Però un no sap mai com aniran les coses, i ara, avui, no canviaria res del que m'ha tocat viure. Gràcies a això crec que m'he fet millor persona, i espero que també científica. Però el més important, he conegut a totes les persones que han fet possible aquesta tesis, i a les que vull donar les gràcies. Gràcies, Gracias, Thank you, Merci.

En primer lloc voldria donar les gràcies al meu director de tesis. Roque, gracias por enseñarme lo que es la ciencia y hacer que siga teniendo ganas de investigar. Por la paciencia que has tenido siempre conmigo, y el apoyo que me has dado durante estos años. Y sobre todo por confiar en mí. I al meu tutor, Paco, gràcies per donar-me la oportunitat de fer aquesta tesis i transmetre'm sempre l'entusiasme que sents per la ciència.

Gràcies a la meva família, pel seu suport constant. A les meves germanes Mònica, Meritxell i Mireia, que sempre estan amb mi, sigui on sigui. Mil gràcies per preocupar-vos i comptar sempre amb mi tot i no ser a casa. Als meus nebots. L'Ian, el meu coco, per transmetre'm sempre tanta alegria i fer-me tenir ganes de tornar ben aviat a casa. I al petit Pol, que mai deixa de somriure. Als meus cunyats, Vichin, Gerard i Toni, que són els millors cunyats del món! Als meus pares Antònia i Manolo, que han treballat tota la vida per a que no em faltés mai res i sempre pogués triar en aquesta vida. Als meus avis l'Antònia i el José, que tenen una gran part de culpa que avui sigui la persona que sóc.

A les meves amigues i amics del poble, els de tota la vida. Gràcies Marisol, Silvia, Mireia C i Dani per estar sempre ben a prop meu, per tenir-me sempre en compte i per ajudar-me, d'una manera o una altra amb aquesta tesis.

A Javi. Gracias sol, por todo el apoyo que me has dado siempre.

A les meves floretes, la Sara, la Tània, la Trini, l'Elsa, i la Cris. Merci per ser com sou, per fer de qualsevol dia que ens trobem un dia inoblidable. Per donar-me la oportunitat de desconnectar quan estic amb vosaltres. Quina sort que vaig de tenir de trobar-vos a l'aulari J el primer any de carrera!! Dios las cría y ellas se juntan!!

A la meva gent de Barcelona

Gracias Andreea, por ser siempre un gran apoyo para mí. Multsu mesk por escucharme, ayudarme en todo y estar siempre allí. Nunca olvidaré el primer día de los cursos de doctorado, cuando nos sentamos juntas!! Qué suerte la mía!

A la Txell, moltes gràcies per ajudar-me tant als meus principis, que mai són fàcils! Per ser sempre tan pacient, per transmetre'm el teu optimisme i donar-me tants bons consells!! I, junt amb el amico Loris, mil gràcies per animar-me i ajudar-me tant durant la meva estància als US, sou els millors!

Als meus Rabenessitos!! Sin vosotros la tesis no hubiera sido lo mismo, lo sabéis bien!! Merci beacoup Fabien, por todos los cafés que hemos compartido, y por escucharme siempre. Jordi, gràcies per les llargues xerrades, les visites a Vlc i les nits d'absenta! A Mari, Bea, Annie, Leo, Rosa, Giulia, Katy, Alberto, Momo, y todos los demás por esas noches inolvidables de Bitácora (y a veces de Magic!!) desconectando de todo. Força Rabenessen de Caramel!!!

Luisa, mi querida compañera de laboratorio! Gracias por todas esas tardes/noches (más bien noches!!) compartidas en el labo! A Andreia, mil gracias por transmitirme ese buen rollo y optimismo que siempre llevas contigo, y que tan necesario es en el laboratorio!

Al laboratori de la UBCM (o antiga UBCM) de l'IMIM, la meva familia durant tres anys. Gràcies Almu, Carles, Judit, Ilse, Elena (mil gracias por ayudarme a encontrar piso!!), Nico, Montse, Maria, Anna M, Neus, Mercè, David (que haguès fet jo sense tu a Bruseles? :D), Manolo, Alicia, Jordina, Laura, Natalia, Sandra, Rosa, Raquel, Marta, Mireia, Pilar, i tota la resta, pel meravellós dia a dia que vaig passar amb tots vosaltres.

Gràcies als meus companys de doctorat, Ivan, Xavi, Selma, Jose, Anna i d'altres, per totes les rises que vem compartir durant els primers anys de tesis!

A mi gente de Valencia

Gracias Carla, sabes que te mereces un párrafo entero!! Gracias por haberme hecho sentir en casa aún cuando no lo estaba. Por presentarme a todos tus amigos y hacerlos míos también. Por acordarte siempre de mí y hacerme reír tantas veces!! Qué grande eres!

A Marta, la pequeña motera! Gracias por escucharme siempre, porque saber escuchar es algo que pocos saben hacer. Por aconsejarme tantas y tantas veces y tan bien, aunque yo a menudo hiciera todo lo contrario… ;P! Y por ser mi compañera de laboratorio a partir de las 19h! Te va a salir una tesis impresionante!!!

Marina, muchas gracias simplemente por ser como eres. Por todo lo que me has aguantado este tiempo, por lo bien que siempre te has portado conmigo!! Porque es un placer trabajar mano a mano contigo!

Lauri, gracias por tantas horas pasadas juntas, pasando lo bueno y lo malo de la vida del laboratorio y por la confianza que siempre has depositado en mí. Gràcies Marta Benet, per transmetre'm sempre la teva energia positiva!

A los chicos del semisótano, gracias!!! A Felipe, per a que segueixis sent el fan número 1 del Barça quan jo ja no hi sigui!! A Juan Car, mi súper compañero de tenis! A Agus, Ramiro y Guillermo por esas birras de viernes tarde sin maldad que se acabaron convirtiendo en cenas de borrachera! ☺

A la gente del centro de investigación. Gracias a la mama del laboratorio, Cris, por tratarme siempre como a una más de la familia, y hacerme sentir tan bien. Esas paellas en Canet no las olvidaré nunca!!! A la boluda del laboratorio, Gaby. Gracias por tantas risas diarias, eres única! A la Lourdes, gràcies per fer-me sentir sempre una mica més a prop de casa! A Nuri, Sandra, Laura, Silvia, Ana, Loli, Laia, Álex, Amparo, Violeta, y a todos los demás por el día a día y compartir conmigo esas charlas locas durante la comida en la biblioteca!

A la millor companya de pis que es pot demanar, la Gema, mil gràcies per tot.

Gracias a Carlos, Pedro, Carmina, Salva, Nuria, Vicente, Antonio y Pepe. Por tratarme como a una más de la colla. Sabéis que tenéis a una amiga en Barcelona!!!

A mis compañeros de voley, que tantos momentos de risas me han dado. Primero, al mejor equipo de voley playa de Valencia!! Gracias Gaby, Rosana, Toni, Angels, Juan, Gonzalo, Carmen, Alberto, Fer, Rafa, Raul, Augusto, y otros, por esos miércoles por la tarde disfrutando de la playa!!! Por las cervecitas post-partido, por esas torrás de bienvenida, y de despedida, por esos partidos del Barça vividos juntos!! Y también a mi equipo de voley de Tarongers! A David (míster voley), Iñaki, Inma, Luis, Carlos, Benja, María, Jaume, César, … Por esos entrenamientos con tantas risas, permitiéndome desconectar de todo. Por esas cenas y fiestas de las que guardo un gran recuerdo!! Y claro, por esos impresionantes partidos de voley! De aquí a las olimpiadas! Os echaré de menos… Sniff…

Julio, gracias bonic. Por aparecer en mi vida justo cuando más lo necesitaba.

A la meva gent de Bruseles

Merci beacoup to all Frederic Lemaigre's Lab members for your help during my stay in Brussels. My special thanks to Christophe, Sabine and Anne Christine.

x

A la meva gent de Philadelphia

Thank you Dr. Ken Zaret. It was a great honor to have the opportunity to work in your laboratory. Thanks for transmitting me your enthusiasm for science, and for the warm hospitality during my visit.

Juanma, mil gracias! Porque hiciste que mi estancia en Philly pasara rapidísimo! Y claro, por ayudarme tanto en el labo (espero mi nombre en agradecimientos de tu próximo Science!). Lots of thanks to the rest of the Zaret's lab, Jun Sun, Cheng Ran, Dave, Andreea, Ewa, Angela, Jason, Eilleen and specially to Deborah, who helped me a lot when I was there.

Gracias a Dani, Almudena, Carmen y Juanma (otra vez!). Por esos domingos de cine, lunes de DVD, y jueves de Blue-Ox! Estoy segurísima que os irá muy bien a todos, porque los que valen, valen! Y vosotros valéis mucho. A Antonio, Mara, Salva y muchos otros que hicieron que mi estancia fuera maravilllosa!

ABSTRACT

Hhex transcription factor is expressed in multiple endoderm-derived tissues, like the liver, where it is essential for proper development. The pleiotropic effect of Hhex in the embryo and its dual role as a transcriptional repressor/activator suggest the presence of different interaction partners capable of modulating its activity and function. In the current study we identified two new Hhex protein interactors: SOX13 and c-Myc.

We show that Hhex interacts directly with SOX13. By doing so, Hhex sequesters SOX13 from the SOX13·TCF1 complex, overturning SOX13-dependent repression of the Wnt pathway. On the other hand, Hhex induces proliferation of non-tumorigenic human fibroblast through a Myc-dependent mechanism. Hhex and c-Myc interact directly upregulating Cyclin D1, a c-Myc target gene involved in cell cycle progression and proliferation. Elevation of Cyclin D1 might be the final effector of Hhex capacity to regulate cell proliferation.

ABSTRACT

El factor de transcripción Hhex se expresa en múltiples tejidos derivados del endodermo, como el hígado, siendo esencial para su correcto desarrollo. El efecto pleiotrópico de Hhex en el embrión y su posible papel como activador o represor transcripcional sugieren la existencia de interactores capaces de modular su actividad y función. En este trabajo hemos identificado dos nuevos interactores de Hhex: SOX13 y c-Myc.

La interacción entre Hhex y SOX13 bloquea el efecto represor de SOX13 sobre la vía Wnt/TCF. De este modo Hhex podría actuar modulando la vía Wnt para obtener los niveles adecuados en diferentes contextos biológicos. Por otra parte, encontramos que Hhex induce la proliferación y transformación en fibroblastos y que esta actividad es dependiente de c-Myc. Hhex y c-Myc interaccionan de forma directa resultando en la inducción de Ciclina D1, un gen diana de c-Myc implicado en la regulación del ciclo celular.

ABREVIATION AND ACRONIM INDEX

aa: Amino acids
Ade: Adenine
ADE: Anterior Definitive Endoderm
AML: Acute Myelocitic Leukemia
AP-1: Activator Protein 1
APL: Acute promyelocitic leukaemia
AVE: Anterior Visceral Endoderm
bHLH: Basic Helix Loop Helix
BMP: Bone morphogenetic protein
bp: Base Pairs
BrdU: Bromodeoxyuridine
DBD: DNA binding domain
ChIP: Chromatin Immunoprecipitation
CK1: Casein Kinase 1
CNS: Central Nervous System
Co-IP: Co-immunoprecipitation
DKK1: Dickkopf-related protein 1
Dpc: Days post coitum
DVL: Disheveled
E: Embryonic age
ECM: Extra cellular matrix
EGF: Epidermal Growth Factor
EGFP: Enhanced Green Fluoresence Protein
eIF4E: Eukaryotic translational initiator factor 4E
EMT: Epithelial to Mesenchymal Transition
ES: Embryonic stem
ESM-1: Endothelial cell-specific molecule 1
FCS: Fetal Calf Serum
FGF: Fibroblast growth factors
FGFR: Fibroblast growth factor receptor.
FZ: Frizzled Receptor
FSK: Foreskin cells
Gal: Galactose
GRG: Groucho family proteins
GST: Glutathione-S-transferase
HD: Homeodomain
HGF: Hepatocyte Growth Factor
Hhex: Hematopoietically Expressed Homeobox
His: Histidine
HMG: High Mobility Group
HNF1a: Hepatocyte Nuclear Factor 1 alpha
HNF4: Hepatocyte Nuclear Factor4
HRE: Hex response element
HSCs: Haematopoietic stem cells
ICA12: Islet cell antibody 12
ICM: Inner cell mass
IP: Immunoprecipitation
Kap7: Karyopherin/Importin 7
kDa: Kilodalton
KO: Knockout
LB: Liver Bud
LD: Liver diverticulum
LEF: Lymphoid enhancer-binding factor
Leu: Leucine

ABREVIATIONS

L-PK: L- Pyruvate Kinase
LRP: Low-density lipoprotein receptor related protein
LZ: Leucine-Zipper
Mb: Myc box
MEF: Mouse embryo fibroblast
Miz-1: Myc-interacting zinc finger1
NB: Nuclear Body
NEO: Neomicin
NF-Y: Nuclear factor Y
NIS: Sodium Iodide Symporter
NPM: Nucleophosmin
NTCP: Na-dependent bile acid cotransporter
o/n: overnight
ORF: Open Reading Frame
P/S: Penicillin/ Streptomycin
PML: Promyelocytic Leukaemia protein
Q-rich: Glutamine Rich
RA: Retinoic acid
SD: Selective Dropout
Shh: Sonic hedgehog
SNP: Single nucleotide Polymorphism
SOX: SRY-related high-mobility group box
SP1: specificity protein-1
SRF: Serum response factor
STM: Septum transversum
T1DM: Type 1 Diabetes Mellitus
T2DM: Type 2 Diabetes Mellitus
TAD: Transcriptional Activation Domain
TBP: TATA-binding protein
TCF: T-cell specific Factor
Tg: Thyroglobulin
TGF: Transforming growth factor
TLE1: Transducin-like enhacer of split 1
Trp: Triptofan
TRRAP: Transactivation/ Transformation-associated protein
UTR: Unstranlated region
VEGF: Vascular endothelial growth factor
WT: Wild type
Y2H: Yeast Two-Hybrid
YSL: Yolk syncitial layer
YY1: Yingyang-1

INTRODUCTION

I1. MOUSE DEVELOPMENT: ENDODERM AND LIVER

I1.1. Early steps in mouse development: endoderm formation

In the mouse, embryonic development is initiated when the one-cell egg is fertilised to form the zygote. The first cleavage of mouse zygote produces two identical cells, which then divides again to produce four cells. Usually the cells remain together, dividing asynchronously to produce 8 cells, 16 cells, and so on. By the third day the embryo develops to a compact ball of 16-32 cells called morula. After several more divisions, the morula cells begin to specialize and form a hollow sphere of cells called blastocyst or blastula (**see figure I1**). The blastocyst is a sphere of about 150 cells, with an outer layer (the trophoblast), a fluid-filled cavity (the blastocoel), and a cluster of cells on the interior (the inner cell mass or ICM). At 4.5 days two distinct cell types are derived from the earlier ICM: the epiblast, and the primitive endoderm which comprises the parietal and visceral endoderm (VE). Only the epiblast gives rise to the embryo proper (containing the three embryonic germ layers, ectoderm, mesoderm, and endoderm) whereas the trophoblast and the primitive endoderm give rise to extra-embryonic structures that support the intra-uterine development of the embryo and act as signalling sources to pattern the embryonic tissues prior to gastrulation.

The process of gastrulation begins at about 6.5 days post coitum (dpc) in the mouse. Epiblast cells, an apparently homogeneous epithelium, undergo an epithelial to mesenchymal transition (EMT) to form the

primitive streak, located at the posterior pole of the future embryo. During the next 12-24 hours, the streak elongates to the distal end and a specialized structure is formed in the anterior part of the streak, the node, also called organiser in other vertebrates. The node generates axial mesendoderm which gives rise to the mesoderm that will populate the midline of the embryo and the definitive endoderm. This definitive endoderm moves anteriorly, displacing the anterior visceral endoderm (AVE) and becoming the anterior definitive endoderm (ADE) [1].

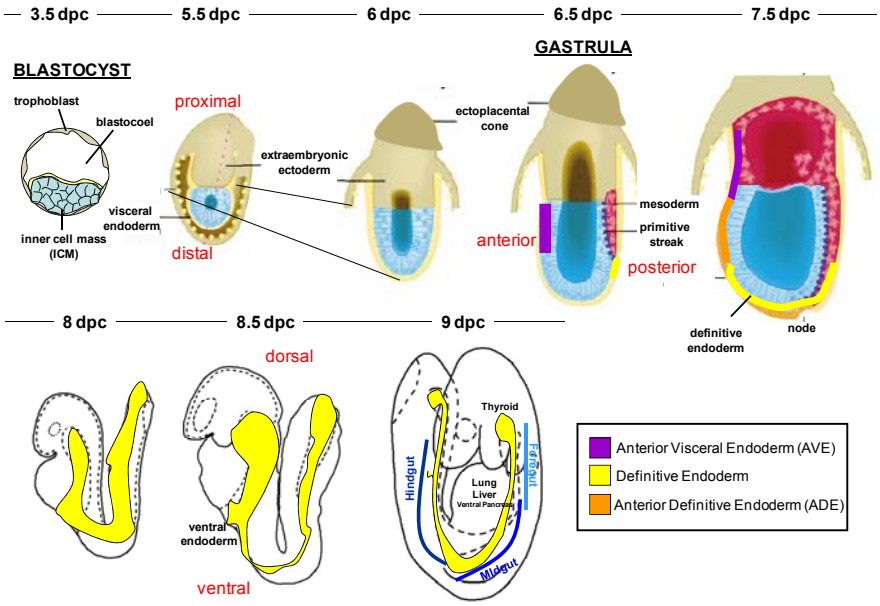


FIGURE I1. Steps on early mouse development and endoderm formation

Schematic representation of the different stages on mouse development, from 3.5 dpc to 9 dpc. It is emphasized the formation of the endoderm layer (yellow), and its patterning.

Between 7,5-8,0 dpc the flat sheet of ADE cells is transformed into the primitive gut tube as follows [2]. A crescent-shaped fold appears at the anterior end of the embryo simultaneously with the beginning of the formation of the somites, mesodermal masses which number is used to stage mouse embryos. This fold, called anterior intestinal portal, progresses posteriorly along the body axis forming the anterior part of the primitive gut tube. Shortly after, a similar fold, the caudal intestinal portal, forms at the posterior end of the embryo, and moves anteriorly to form the posterior parts of the primitive gut. Both folds meet at the yolk salk [3-4]. The gut tube is patterned in an anterior to posterior manner, and discrete regions of endoderm can be identified: foregut, midgut and hindgut sections [3, 5].

Temporal and spatial gradients of some factors secreted from the adjacent mesoderm, such as Wnt ligands, appear to be important to regulate regional identity of the primitive gut tube. Studies in chick and *Xenopus* support a model where Wnts secreted from the posterior mesoderm repress the foregut fate posteriorly and promote hindgut development [6]. Thus, Wnt signalling must be inhibited in the anterior endoderm to establish foregut identity (**figure I4**). Secreted Wnt-antagonist expressed in the foregut endoderm, such as Sfrp are predicted to keep Wnt signalling off in the foregut region [7] where the liver, gall bladder, pancreas and lungs are specified [5, 8-10].

The canonical Wnt/β–catenin pathway is an evolutionarily conserved pathway essential for normal cellular processes such as development, growth, survival, regeneration and self-renewal [11-15], and also in the progression of some kinds of cancers [16].

The Wnt/β-catenin signalling is inactive in unstimulated cells. In this steady-state condition, β-catenin, the central player in the signalling cascade, forms a complex with Axin, APC, GSK3 and CK1 (see **figure I2**). In the absence of Wnt, β-catenin is phosphorylated by casein kinase 1 (CK1) and GSK3 at serine/threonine residues located at the N-terminal domain of the protein [17]. This phosphorylation targets β-catenin for ubiquitinization and ultimate degradation by the proteasome. When Wnt proteins bind to the Frizzled (Fz) receptor on the surface of the cells, it activates the canonical Wnt pathway. The Wnt/Frizzled interaction induces association with the low-density lipoprotein receptor related protein (LRP) 5/6, and this complex then recruits Disheveled (Dvl), which is thought to inactivate GSK3 [18]. Inactivation of GSK3 leads to the absence of β-catenin phosphorylation, releasing it from the Axin/APC/GSK3 complex. β-catenin then translocates to the nucleus, where it binds to lymphoid enhancer-binding factor1/T cell specific transcription factor (LEF/TCF), displaces the transcriptional inhibitor of the Groucho family Tle, and in complex with TCF activates target genes important in proliferation and differentiation [11].

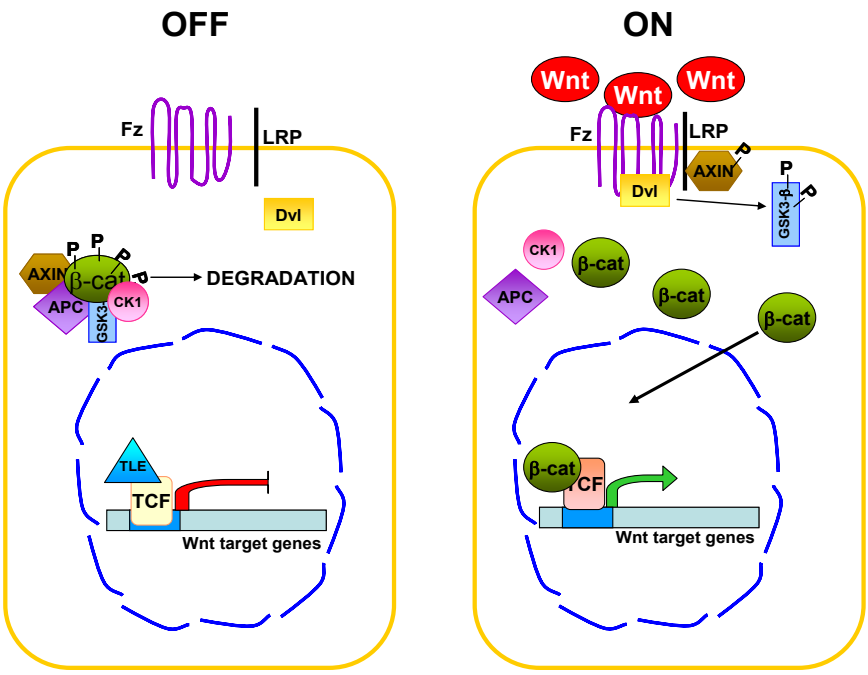


FIGURE I2. Canonical Wnt pathway

In the absence of Wnt ligands (OFF state), β-catenin is sequestered by a protein complex and it is phosphorylated and degraded by the proteasome. Upon binding of the Wnt ligand to the receptor Frizzled (Fz; ON state), it associates to the LRP protein. The Wnt/Fz/LRP complex recruits dishevelled (Dvl) protein, which in turn inactivates GSK3, and β-catenin is not phosphorylated, allowing its translocation to the nucleus where it binds to TCF proteins to activate Wnt target genes.

I1.2. Liver development

The liver is the largest internal organ providing essential metabolic, exocrine and endocrine functions. Hepatocytes are the principal cell type in the liver and these along with cholangiocytes or biliary epithelial cells (BECs) are derived from the embryonic endoderm. Hepatic development can be divided in different stages: **hepatic competence, liver induction, liver bud morphogenesis and growth and hepatic maturation (figure I3).**

The liver is specified around 8.0 dpc (embryonic age, E8.0) in the mouse [5, 19] from a region of the ventral foregut endoderm **competent** to activate the hepatic program i.e. expression of hepatic-specific genes such as Foxa2 and Gata4 [20-21]. However, the first morphological sign of the embryonic liver is the formation of the hepatic or liver diverticulum (ld), an out-pocket of thickened ventral foregut epithelium adjacent to the developing heart at E9.0. The anterior part of the hepatic diverticulum gives rise to the liver and the intrahepatic biliary tree, while the posterior portion forms the gallbladder and extrahepatic bile ducts. Between E9.0 and E9.5 the hepatic endoderm cells have to progress to form the **liver bud**. This process is divided into three morphogenetic stages: stage I, the formation of a thickened, columnar epithelium; stage II, the formation of a pseudo-stratified epithelium, and stage III, laminin breakdown and hepatoblast migration from the epithelium adjacent septum transversum mesenchyme[22-24].

Although Wnt/β-catenin pathway appears to repress liver fate during endoderm patterning stages of development, by E9.5-E10 β-catenin has the opposite effect, promoting hepatic growth in the liver bud [6, 25-29](**figure I4**). Overexpression of β-catenin in developing chick livers leads to an increase in liver size [29] and mouse embryo livers cultured in the presence of β-catenin antisense oligonucleotides show a decrease in cell proliferation rate [25].

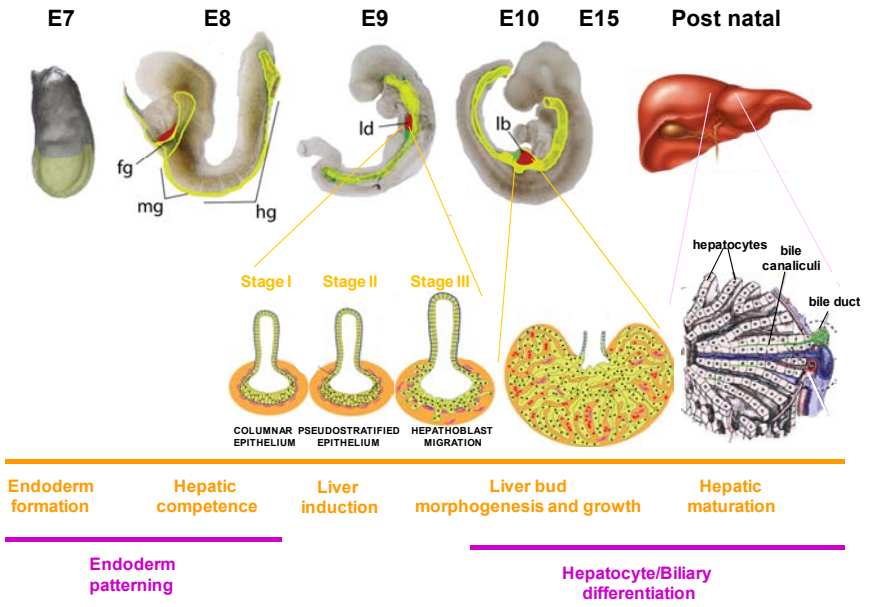


FIGURE I3. Steps on mouse liver development

The scheme shows mouse embryos at different developmental stages with the endoderm highlighted in yellow, the liver in red and the gall bladder in green. Different processes that occur on hepatic development are written in orange and purple. Fg: Foregut, ld: liver diverticulum, lb: growing liver bud, mg: midgut, hg: hindgut. *Adapted from Zorn, A. et al (2007).*

Between 10-15 dpc the liver bud undergoes a period of accelerated **growth** as it is vascularised and colonized by hematopoietic cells [23, 30].

Hepatoblast are bipotential cells that will give rise to hepatocytes and cholangiocytes. Prior to day 12 of gestation in the mouse, the hepatoblast remain in a morphological undifferentiated state, having an irregular shape, large nuclear to cytoplasmic ratio and relatively few organelles [23]. The process of **differentiation** of the hepatoblast to the hepatocyte is gradual, taking several days during embryo development from day E13. Establishment of the hepatocyte cell polarity is crucial in generating the hepatic function. The hepatic epithelium polarizes,

creating a small apical domain that lines channels between cells, called cannaliculi, that connect the bile ducts and drain into the intestine. Final **maturation** of the liver is gradual and continues into the postnatal period (see summarized scheme of liver development in **figure I3**)

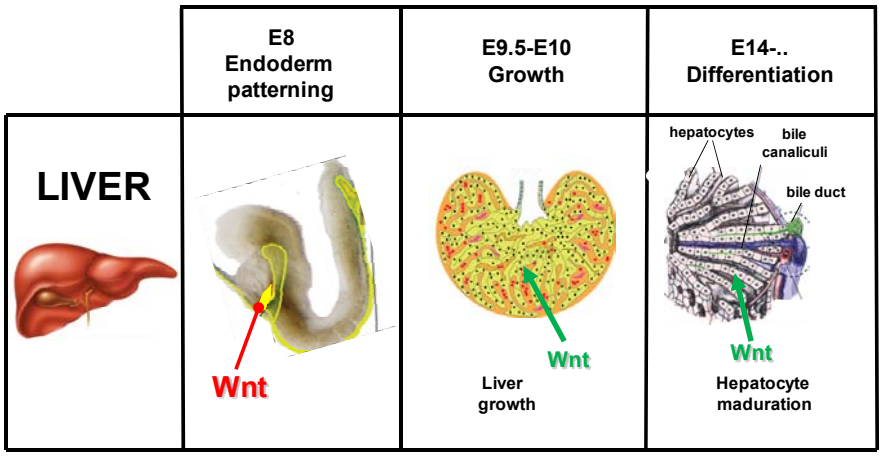


FIGURE I4. Wnt in hepatic development

Schematic table depicting the role of the Wnt pathway during different stages of liver development. Tightly regulation of the Wnt pathway is necessary for proper hepatic development. Wnt pathway should be repressed at E8 for a correct foregut formation. However liver bud growth and hepatocyte maturation require an active Wnt signalling. *Adapted from Zorn, A. et al (2007).*

I2. HHEX TRANSCRIPTION FACTOR

Hhex (Haematopoietically expressed homeobox) also known as PRH (proline-rich homeodomain) is a transcription factor that belongs to the homeobox gene family. Homeobox-containing proteins are a large family of transcription factors that share an evolutionary conserved 60 amino acids sequence encoding a DNA-binding domain called the homeodomain. This homeodomain was originally identified in the homeotic proteins of *Drosophila* and subsequently in proteins of diverse organisms such as frogs, mice and humans [31]. Homeobox genes act as master regulators of development through their ability to activate or repress downstream target genes controlling the specification and/or organogenesis of several tissues such as thyroid, heart, spleen and liver [32-35]. Hhex was first described in chicken and human by Crompton et al [36], while searching for new homeobox genes in haematopoietic cells. Later, *Hhex* was detected in several species including mouse, rat, Zebrafish, *Xenopus and C.elegans* [37-40].

I2.1 Biochemical features

I2.1.1. Genomic structure

Human HHEX gene was mapped in chromosome 10 [41], specifically in 10q24. It contains 4 exons and spans 5.7 kb of genomic DNA [42]. Mouse Hhex is found in chromosome 19 [43-44], and its structure in four exons is conserved with the human homologue (see **figure I5**).

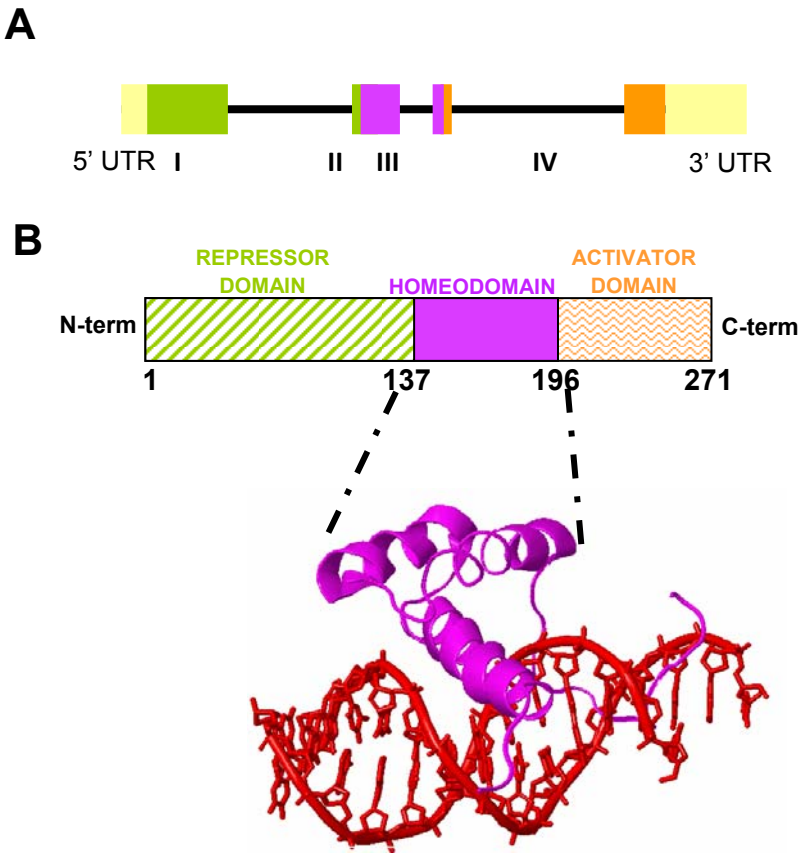


FIGURE I5. HHEX gene and protein structure

A) Diagram of human HHEX gene. Yellow boxes represent the 5' and 3' untranslated regions (UTRs), black thin lines represents the introns and coloured boxes represent the different exons (roman numbers). **B)** *Top*: A schematic view of human HHEX protein. The three functional domains are depicted in different colours that correlate with panel A, *Bottom*: Molecular model showing HHEX homeodomain interacting with DNA. Homeodomain helix III is contacting the major groove of DNA and N-terminal arm is contacting the minor groove.

I2.1.2. Functional domains

At the protein level, murine Hhex is a 271 amino acid protein with a molecular weight of approximately 37 kDa. Human and murine homologues share 95% sequence homology, emphasizing the conservation of Hhex between species [45]. As many other transcription factors it has a modular nature, containing three well-characterised domains (**figure I5**).

The N-terminal or proline-rich domain (from amino acid 1 to 137) is rich in proline, alanine and glycine residues. It has been described as the repressor domain because it is able to independently repress transcription of a GAL4-dependent reporter when fused to a GAL4 DNA-binding domain [39, 46]. Moreover, serial deletions within N-terminal domain showed that there are at least two regions capable of significantly repressing transcription, region aa28-105 and region aa107-141 [46]. On the other hand, recent findings have shown that the N-terminal domain mediates Hhex homo-oligomerisation *in vivo* and *in vitro* [47].

The homeodomain (aa138 to aa196) is composed by a flexible N-terminal arm and three alpha-helices that are separated by a short loop and a short turn (**figure I5 B, bottom panel**) [48]. This is the region responsible for DNA binding to the consensus 5'-$^{C}/_{T}{}^{A}/_{T}$**ATTAA**$^{A}/_{G}$-3' sequence (also called Hhex response elements or HRE) [36].

The C-terminal domain (from aa197 to aa271) is also known as the acid domain because it contains high levels of acidic residues. It functions as the activation domain, being a necessary region for full activation of

some genes like the sodium-dependent bile acid cotransporter gene (NTCP) or the sodium iode symporter (NIS) [49-51].

I2.1.3. Subcellular localization

Hhex is present in the nuclear and cytoplasmatic cell compartments [50, 52-56]. Overexpression experiments of Hhex-EGFP fusion proteins in HeLa cells show that Hhex is mainly nuclear and its homeodomain is essential for the nuclear localization [57]. A shift to a cytoplasmatic location is linked, in most cases, to aberrant proliferation and tumourigenesis of several tissues such as thyroid, breast and haematopoietic cells [50, 55-56].

I2.1.4. Interacting partners

In the last decade, Hhex has been identified as an interacting partner of several proteins such us PML, eIF4E, TLE1, HNF1α, CK2, GATA2 or Jun (**Table I1**).

Hhex was found to interact with RING domain of Promyelocytic leukaemia protein (PML) through its N-terminal domain. PML is found to be deregulated on acute promyelocytic leukaemia, a disease characterised by a block in myeloid cell differentiation [58]. Taking into account that Hhex has an important role in the correct differentiation of the haematopoietic system, these interaction could establish a link between growth control and haematopoiesis [59]. Another important partner for Hhex found in the literature is the eukaryotic translational initiator factor 4e (eIF4E) [54]. eIF4E promotes cellular proliferation and malignant transformation in part by promoting the selective transport of specific mRNAs, such as Cyclin D1 from the nucleus to the cytoplasm

[60]. Hhex interacts with eIF4E by its repression domain (N-terminal domain) inhibiting eIF4E mRNA transport function [54].

PARTNER NAME	HHEX INTERACTION DOMAIN	INTERACTION TECHNIQUES AND CELLULAR MODELS	BIOLOGICAL SIGNIFICANCE OF THE INTERACTION	REFS
PML	N-terminal	Y2H, co-IP, end. co-IP K562 and NB4 cells	?	[59]
JUN	Homeodomain	Y2H, *in vitro* PD	Reduction of Jun transcriptional activation	[61]
TBP	N-terminal	*In vitro* PD	Hhex transcriptional repression	[46]
eIF4E	N-terminal	PD, end co-IP K562 and U397 cells	Inhibition of transformation and growth effects of eIEF4	[54]
HC8 proteasome subunit*	N-terminal	Y2H*, PD, co-IP, end. co-IP K562 cells	?	[62]
TLE1	N-terminal	Y2H, PD, co-IP K562 cells	Increment of Hhex transcriptional repression activity	[63]
GATA	Homeodomain and C-terminal	Y2H, co-IP, end. co-IP COS7 and HUVEC	Inhibition of GATA transcriptional activity	[64]
SRF	?	Co-IP 10T1/2 cells	Induction of SSM22a expression	[65]
HNF1α	Homeodomain	*In vitro* PD, co-IP HeLa cells	Stimulation of HNF1α activity	[51]
SOX10	?	Y2H, PD 293 cell	?	[66]
CK2β*	N-term *in vitro* C-term *in vivo*	Y2H*, PD, co-IP K562 cells	Hhex phosphorylation blocking DNA binding	[67]
Kap7	?	PD and mass spectrometry HeLa cells	Transport of Hhex to nuclear compartment	[68]

TABLE I1. List of proteins known to interact with Hhex.

Y2H: Yeast Two-Hybrid; *PD: Pulldown; End: Endogenous;. PML: Promyelocytic leukaemia protein; TBP: TATA-binding protein; eIF4E: Eukaryotic translational initiator factor 4E; TLE1: Transducin-like enhacer of split 1; SRF: Serum response factor; HNF1a: Hepatocyte Nuclear Factor 1 alpha; Kap7: Karyopherin/Importin 7.** Y2H in which Hhex was used as a bait.

Hhex binding can also result in transcriptional co-repression. Hhex interacts with GATA-2 in endothelial cells resulting in reduced binding of GATA to its target genes, inducing the down-regulation of GATA-dependent gene expression [64]. Similarly, Hhex homeodomain is also reported to interact with members of the AP-1 (activator protein 1) family of transcriptions factors through direct binding to the Jun-Fos complex [61]. This interaction results in the inhibition of Jun-mediated gene activation.

I2.1.5. Target genes

Hhex can behave as a transcriptional repressor or activator. Its effect is exerted by directly binding to the promoters of regulated genes containing the consensus Hhex-binding sequence **5'-$^{C}/_{T}{}^{A}/_{T}$ATTAA$^{A}/_{G}$-3'** [36] or, as mentioned above, indirectly by interaction with other transcription factors that are ultimately responsible for gene regulation [69].

Hhex is overall considered a transcriptional repressor (**Table I2**). In fact, when a vector expressing a fusion-protein Hhex-GAL4 DNA-binding domain was co-transfected with a luciferase reporter plasmid containing several GAL4-binding sites, luciferase activities were decreased in a concentration-dependent manner [39]. As already stated, the N-terminal domain of Hhex is responsible for its repressing features [39, 46]. Goosecoid (**Gsc**), a marker of dorsal mesoderm, is a negatively regulated Hhex target gene [70]. Hhex downregulates Goosecoid expression in *Xenopus* embryos and in embryonic stem cells (ES). The Wnt antagonist Groucho-related factor Tle4 is also repressed by Hhex in *Xenopus* [71]. The repression of Gsc and Tle4 contributes to establish the anterior

identity in early developmental stages as will be discussed in **section I2.2.4**.

GENE	MODEL	METHOD	REF.
Goosecoid	*Xenopus* embryos, ES cells	Ectopic expression, Transient transfections	[70]
NTCP	E10.5 Hhex KO embryos	Transient transfection	[72]
Tg	FRLT-5 cells	Transient transfection	[53]
VEGF VEGFR1-2	HUVEC cells, ES cells, K562 cells	Adenoviral infection, microarray, ChIP, EMSA	[73-74]
L-PK	HepG2 cells HeLa cells Rat hepatocytes	Transient transfection Adenoviral infection	[51]
XTle4, Nodal	*Xenopus* embryos, ES cells	ISH, conditional expression, gene chip	[71]
ESM-1	EOMA cells HUVEC cells	Transient transfection, adenoviral infection, Microarray	[75]
NIS	HeLa cells	Transient transfection	[50]

TABLE I2. Hhex target genes

Summary of Hhex targets genes. Shown in red are repressed genes and in green activated genes. *NTCP: Na-dependent bile acid cotransporter; Tg: Thyroglobulin; VEGF: Vascular endothelial growth factor; L-PK: L-Piruvate Kinase; XTle4: Xenopus Transducin-like enhacer of split 4; ESM-1: Endothelial cell-specific molecule 1; NIS: Sodium Iodide Symporter.*

Although minor, Hhex can also act as a transcriptional activator in specific promoters (**Table I2**). This is the case, among others, of NTCP (also known as SLC10A1), a Na (+)-dependent bile acid cotransporter directly activated by Hhex [49, 72]. NTCP is only expressed in hepatocytes and encodes the major transporter of bile acids across the basolateral surface of the hepatocytes [76]. Up-regulation of NTCP by Hhex could be a mechanism to regulate the bile flow in the liver.

I2.2. Cell biology and function

I2.2.1. Expression in normal tissues and in tumorigenesis

During murine development *Hhex* mRNA is detected at 5 dpc by *in situ* hybridisation in the primitive endoderm, and later, at 5.5 dpc in the visceral endoderm (VE) [77] (see **figure I1** for a summary of early mouse development). Then, *Hhex* expression is shifted to the anterior definitive endoderm, an important early signalling centre involved in axis formation by signalling to the epiblast before and during gastrulation, at E6.0. Later, *Hhex* is expressed in the anterior definitive endoderm [77]. The anterior domain of the definitive endoderm becomes the foregut endoderm, where *Hhex* expression continues. Noteworthy, Hhex is expressed during embryogenesis in tissues derived from the foregut endoderm including liver, lung, thyroid, thymus, gallbladder and pancreas [77-79]. But the expression of Hhex is not only restricted to endoderm or derivatives, it is also expressed in mesodermal tissues that give rise to the haematopoietic and vascular progenitors and in the endocardium of the heart [77, 79]. Additionally *Hhex* is also present in regions where vessels are known to form. However, once cells have differentiated to form

committed endothelial precursors, *Hhex* expression is lost in these cells [77].

Similar patterns of *Hhex* expression also occur during early development of other vertebrates as chick and frog [38, 80]. During chick development *Hhex* is also expressed on feather skin development regulating epidermal cells proliferation [81-83].

In the adult, *Hhex* continues to be expressed in several endoderm-derived tissues such as the thyroid, pancreas, liver and lung and also in the haematopoietic compartment and the bone marrow [41, 45, 78, 84]. More recently it has been also found in epithelial breast cells [50].

There have been also several reports associating tumorigenesis to Hhex cellular location. Hhex is a transcription factor present mainly in the nucleus of normal cells [50, 52-53] However, Hhex protein has been found in the cytoplasm in some kind of epithelial tumours [50, 55]. In human thyroid tumours, *Hhex* is expressed only in the cytoplasm in contrast to normal thyroid tissue, where it is both nuclear and cytoplasmic [55]. This suggests that regulation of Hhex entry in the nucleus of thyrocytes may represent a critical step during human thyroid tumorigenesis. Similarly, Hhex is found to be mislocalizated in the cytoplasm in a subset of acute and chronic myelogenous leukaemia cells [56]. In summary, localization of Hhex in the cytoplasm of cells might be associated to tumorigenesis.

I2.2.2. Regulation of *Hhex* gene expression

Hhex expression is regulated by multiple signalling pathways and factors. **Figure I6** shows a summary of the factors that regulate *Hhex* expression,

by direct binding to its promoter or, indirectly by an unknown mechanism.

Regulatory regions, controlling *Hhex* expression in the developing embryo, were mapped by Rodriguez et al [85] using mouse transient transgenic experiments. They found a 4.2 Kb region upstream of the Hhex gene that is important for its expression in endothelial cells, liver and thyroid. They also identify a 633 bp in one of the introns that is necessary and sufficient for *Hhex* expression in mouse AVE and ADE (**figure I6**).

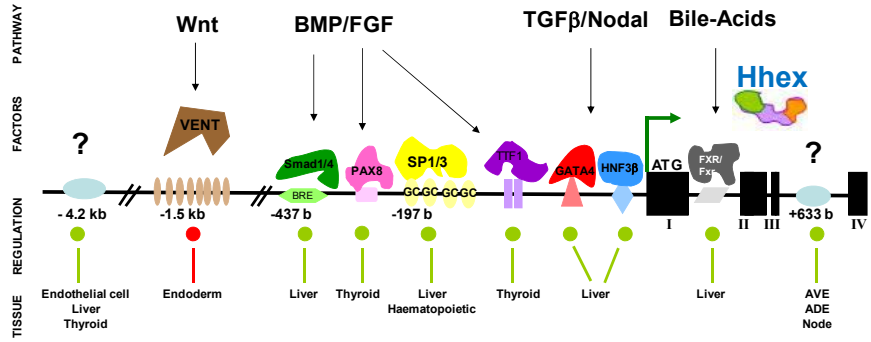


FIGURE I6. Regulation of *Hhex* expression

Summary of different pathways and factors involved in *Hhex* expression. From top to bottom you can see the signalling pathways that regulate *Hhex* expression, the factors found to bind to Hhex regulatory elements producing activation (green circle) or a repression (red circle) of its expression and the tissues where this regulation occurs.

Noteworthily, Wnt pathway plays an essential role in the developmental control of Hhex expression in *Xenopus* and Zebrafish [6, 86-87]. The homeodomain transcription factor Vent2 acts downstream of the Wnt/β–catenin pathway repressing *Hhex* expression in the posterior

endoderm of *Xenopus* [6]. This activity seems to be essential for correct endoderm antero-posterior patterning (see section **I2.2.4**). In Zebrafish, Wnt/β-catenin signalling is also controlling Hhex expression in the dorsal yolk syncitial layer [87].

I2.2.3. Animal models to study Hhex function

At least three Hhex knockout mice have been generated and characterized [88-90]. While Hhex heterozygous mouse is completely viable and phenotypically undistinguishable from the wild type, Hhex deficient embryos die in utero between E10.5 and E13.5 [88-91]. Hhex-deficient embryos show macroscopic features of "dorsalization" (**figure I7**). They present different levels of antero-ventral truncations due to defects on anterior definitive endoderm morphogenesis [88-89]. This leads to different grades of forebrain truncations, and liver, ventral pancreas and thyroid dysplasia [88, 91]. On the other hand they have also impairment on development of vascular system, heart and B-cells.

In order to investigate the embryological role of Hhex after E10-E13.5 and bypass the lethality associated to its deficiency, two lines of conditional liver-specific Hhex knockout mice were developed by crossing a $Hhex^{flox/flox}$ mouse line with Foxa3-Cre or Alfp-Cre mice [92]. Fox3-Cre;Hhex(d2,3/-) conditional mice activates Cre at the time of liver specification, resulting also in embryonic lethality and recapitulating the liver phenotype of the straight mutant. In Alfp-Cre;Hhex(d2,3/) conditional mice, Hhex is deleted in the embryonic liver at E11.5. These mice show severe defects in hepatoblast maturation and in bile duct morphogenesis that result in a small and cystic liver.

Hhex role in development has also been studied in other animals such as *Xenopus* and Zebrafish. Mutant fish line that lack Hhex locus, cycb16, presents also defects in liver, pancreas and thyroid development [93-94]. Similar effects are observed when morpholinos for Hhex are injected in Zebrafish embryos. Studies in which Hhex is inhibited in *Xenopus*, either with interfering constructs or by antisense morpholinos [70, 95], indicates that Hhex is required for correct anterior patterning, which is also in agreement with mouse models.

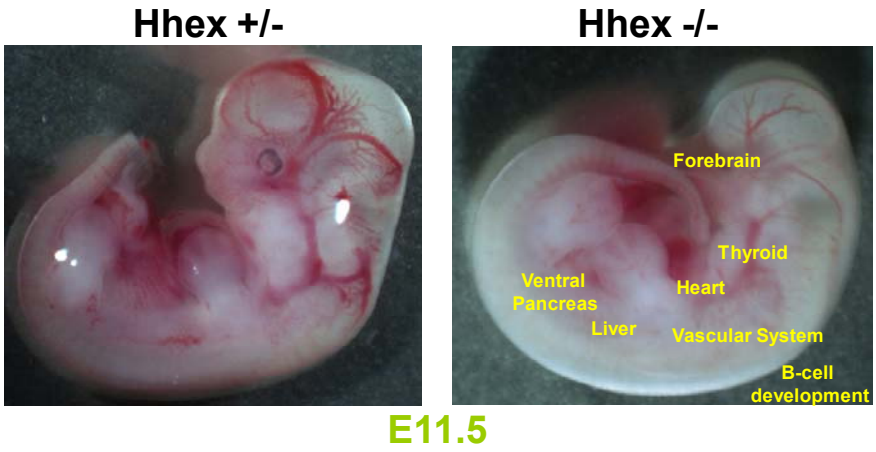


FIGURE I7. Hhex knockout mouse defects
Pictures showing E11.5 heterozygous and mutant homozygous Hhex embryos. Dorsalization of the mutant embryo is noticed by the lack of several ventral structures such as the forebrain, thyroid, liver and ventral pancreas.

Two transgenic gain-of-function mouse lines have been constructed to investigate the role of Hhex in T-cell lineage commitment. One transgenic line expresses constitutively Hhex in bone marrow, spleen and thymus, while the second transgenic line has a thymus-restricted expression through development [96]. Both types of transgenic mice have

abnormalities in T-cell development, indicating that down-regulation of Hhex is important for T-cell maturation.

I2.2.4. The biological role of Hhex

Hhex is a crucial factor for proper development and maintenance of several tissues by regulation of cell proliferation and differentiation (see **figure I9** for a summary). Here, we will discuss its essential role on endoderm patterning as well as hepatic and pancreatic development.

- Anterior identity, endoderm patterning and foregut development

In vertebrates, Hhex plays a critical role in patterning of the body axis. As mentioned in **section I2.2.1**, *Hhex* is first expressed in the primitive endoderm at 4.5 dpc and then in the VE (visceral endoderm). Real-time imaging of Hhex-expressing VE shows that these cells actively migrate anteriorly to form the AVE [97]. Hhex expression is therefore used as a marker for the earliest molecular anterioposterior asymmetry in the mouse embryo before the formation of later signalling centres such as the primitive streak [77].

The role of Hhex in early patterning of the body axis has been addressed in *Xenopus* embryos by ectopic expression of Hhex in two-cell and four-cell stage embryos. This results in embryos with enlarged heads and small trunk structures [70]. Although embryonic development in mouse and *Xenopus* are not identical, they have tissues that perform analogous functions. The anterior endoderm in *Xenopus* contains tissues analogous to the murine AVE and ADE [95]. Hhex specifies the anterior

mesendoderm (head formation) in *Xenopus* [70-71] by repressing promoters associated with mesendodermal genes such as Goosecoid [70] and the Groucho related co-repressor TLE4 [71] that would promote trunk formation. TLE4 functions as a Wnt antagonist and repression of this gene allows increased Wnt signalling in the anterior mesendoderm and formation of the anterior endoderm (**figure I8**). This mechanism is conserved between mice and *Xenopus* [71]. Hhex can also induce the expression of proteins that block the action of signalling pathways in the posterior mesoderm including the expression of Cerberus, a Nodal and Wnt antagonist [70, 86, 98].

Wnt/β-catenin is repressed in the anterior endoderm after head and axial patterning (at late gastrulation and early somite stages), in *Xenopus* embryos. Repression of Wnt signalling in the anterior endoderm allows Hhex expression since β-catenin represses Hhex expression through the homeodomain repressor Vent2 [6] (**figure I8**). Loss of Hhex through the injection of Hhex morpholino oligonucleotides resulted in the absence of liver and pancreas-specific markers. Thus Hhex expression is required for the maintenance of foregut identity and consequently for early stages of liver and pancreas development [6].

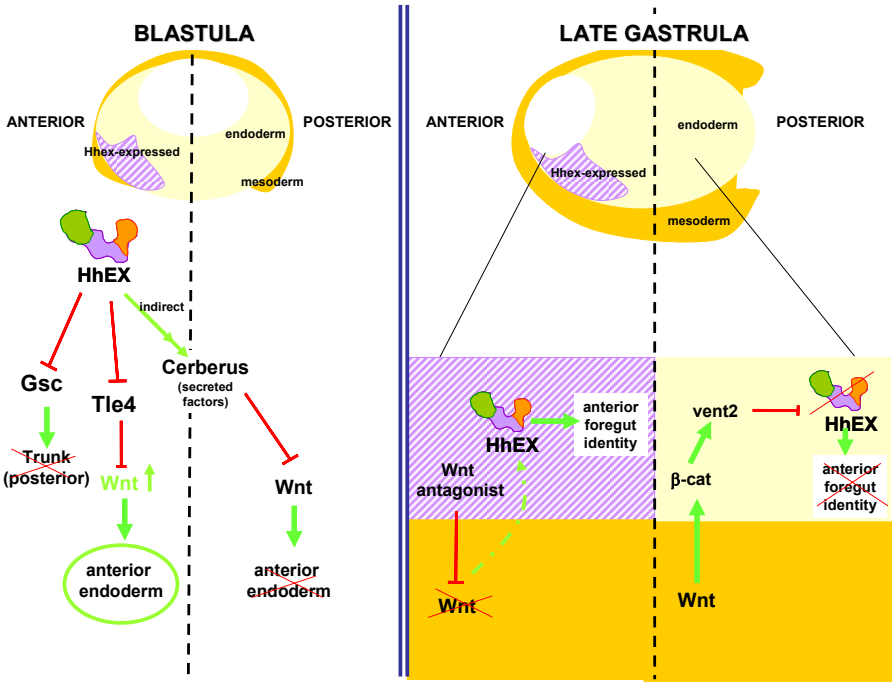


FIGURE I8. Proposed roles for Hhex promoting anterior identity in *Xenopus* embryos
Scheme of the described Hhex roles in different stages of *Xenopus* development. In blastula stage Hhex is promoting anterior endoderm by repressing mesendoderm genes such as Gsc (Goosecoid) or Tle4, promoting then Wnt activity. It also induces indirectly the expression of secreted factors such as Cerberus that antagonise the Wnt pathway in the posterior mesoderm. Later in development, Hhex promotes foregut identity in the anterior pole. In contrast, in the posterior pole Hhex expression is repressed by Vent2 homeodomain transcription factor, a downstream factor of the Wnt/β-catenin signalling. *Adapted from Brickman, J.M et al (2000) [70]; Zamparini, A.L. et al (2006) [71] and McLin, V.A et al (2007)[6].*

- Liver development

Based on the phenotypic analysis of the Hhex straight mutant and liver conditional knockout, it became clear that Hhex is essential for liver morphogenesis.

In Hhex KO mouse the hepatic diverticulum is present at E9.0, suggesting that Hhex is not required for liver specification [88]. However, the liver bud degenerates and no morphologic trace of liver is present in the E10.5 mutant embryo. Bort et al [99] demonstrated that Hhex controls the proliferation of the hepatic endoderm, and promotes the transition of this simple columnar epithelium into a pseudostratified columnar epithelium. This transition permits the hepatoblast to migrate into the stroma, invade the mesenchyme and continue differentiating. The molecular mechanism by which this morphogenetic event is possible is still unclear. The regulation of sonic hedgehog (Shh) pathway could be playing a major role in early liver morphogenesis. Shh is ectopically expressed in the nascent hepatic endoderm from Hex-/- embryos [99]. Since ectopic Shh expression is associated to a lack of pseudostratified epithelium in the neuroepithelium [100], it is feasible that ectopic Shh in the liver mutant Hhex embryo may be at the root of epithelium transition in the liver bud.

In addition to the role of Hhex during early hepatic development, it is also important in later stages of hepatobiliary morphogenesis [92]. By using a Hhex conditional mice in the liver, Hunter et al [92] showed that Hhex is necessary for bile duct morphogenesis and maturation of the hepatoblasts. In fact, Hhex conditional mutant lacks the expression of HNF4 in the nascent liver bud. HNF4 is a crucial hepatocyte maturation factor [101] suggesting that Hhex is an essential knot in the genetic network controlling hepatic differentiation.

In the adult liver, Hhex is thought to play a role in maintaining liver differentiation as Hhex is expressed in the adult liver, but not in undifferentiated hepatocytes or in undifferentiated liver cell lines [39].

- Pancreas development

In vertebrates, the pancreas originates from the dorsal and ventral domains of the gut endoderm epithelium [102]. Bort et al. [91] demonstrated that Hhex is an essential factor for proper ventral pancreas development. In Hhex KO mice there is a failure in ventral pancreas specification, however, when they isolated Hhex-null ventral endoderm prior to its interaction with cardiogenic mesoderm and cultured it *in vitro*, they observed that it is able to activate early pancreatic genes. This suggests that Hhex is not direct regulating the pancreas-specific program, but instead it has a morphogenetically indirect role. In fact, measurements of the proliferation rate of ventral endoderm cells that normally express Hhex showed that Hhex acts by controlling the proliferation and growth of a discrete domain of endoderm. Thus, Hhex-dependent proliferation of ventral endoderm cells allow them to be positioned below the cardiac mesoderm and escape the hepatic-inducing signalling (FGFs and BMPs), executing the pancreatic program.

Hhex expression is maintained in the adult pancreas, suggesting that Hhex may have additional functions in differentiated pancreatic cells [41, 78]. In fact, Hhex has been associated to the pathogenesis of type 2 diabetes in genome-wide association studies in several populations based on large-scale single polymorphism (SNP) analysis [103-105]. Type 2 diabetes mellitus (T2DM) is the most prevalent metabolic disease of the western industrialised world. T2DM is characterised by high blood glucose in the context of insulin resistance and relative insulin deficiency. It is generally agreed that T2DM is caused by environmental factors such as high-caloric diet and reduced physical activity, and a polygenic

background that confers increased susceptibility toward these environmental challenges [106]. Two of the found SNPs (rs1111875 and rs7923837) are in the 3'-flanking region of the human HHEX locus [103], establishing an association between Hhex and T2DM. This association was robustly confirmed in several studies with different populations [107-112]. As Hhex is involved in pancreatic development, the question that arises was whether these SNPs in HHex locus can affect β-cell function. What they can observe is that population with the SNP rs7923837 (G>A) but not rs1111875 (C>T) in Hhex locus presents impaired glucose-stimulated insulin secretion but not alteration in insulin sensitivity [111-113]. Thus, Hhex could be playing a role controlling insulin secretion in the adult pancreas.

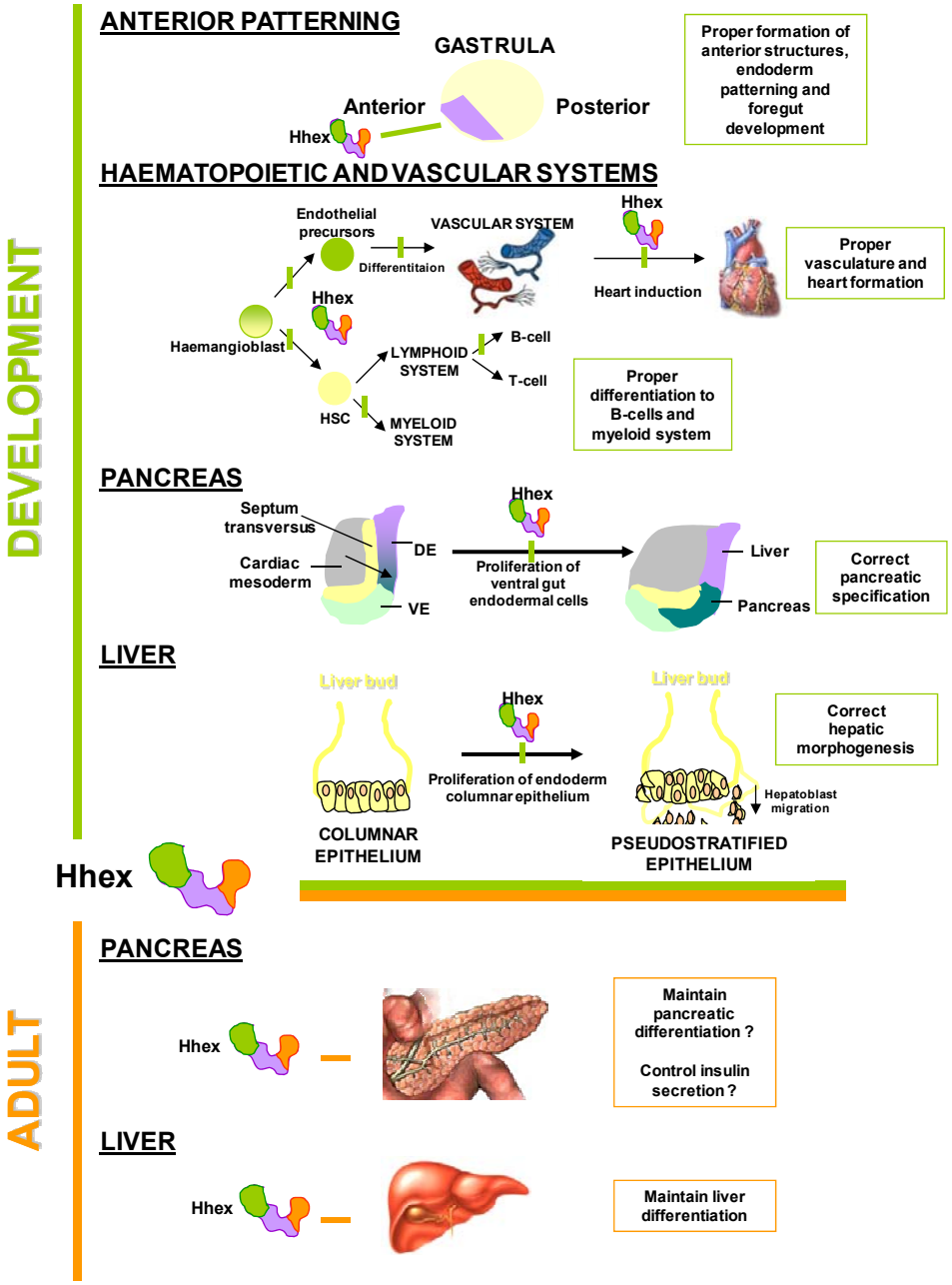


FIGURE I9. Hhex roles on developmental and adult tissues

Summarized scheme of Hhex roles controlling cell proliferation, migration and differentiation during embryo anterior patterning and development of haematopoietic, vascular, pancreas and liver systems. Described Hhex role on maintaining differentiation state of adult pancreas and liver is also shown.

I3.SOX13(ICA12) TRANSCRIPTION FACTOR

The SRY-related high-mobility-group box (SOX) gene family encodes a group of transcription factors that are characterized by a highly conserved high-mobility group (HMG) domain [114-116]. These genes are found throughout the animal kingdom and are expressed in a restricted spatial-temporal pattern playing critical roles in stem cell biology, organogenesis and animal development [116]. More recently they were also linked to human cancer, being deregulated in a wide variety of tumours [117-120].

Protein sequence comparisons showed that SOX genes fall into eight groups, A to H. SOX proteins within the same group share a high degree of similarity (generally 70-95%) inside and outside the HMG box, whereas SOX proteins from different groups share partial similarity (≥46%) in the HMG box domain and none outside this domain. SOX13 belongs to the group D, sharing a characteristic LZ (leucine zipper) and a Q-rich (glutamine rich) domains with the other two members of the group; Sox5 and Sox6. It has also been determined to be a type-1 diabetes mellitus autoantigen, also known as islet cell antibody 12 (ICA12) [121-123], although there is some controversy about its role as a type-1 diabetes mellitus marker [124-125].

I3.1. Biochemical features

I3.1.1. Genomic structure

Human SOX13 is mapped at chromosome 1q31.3–32.1. SOX13 genomic locus contains 14 exons and three different isoforms have been reported in the literature. *SOX13* [121] is an open reading frame (ORF) of 1815 bp

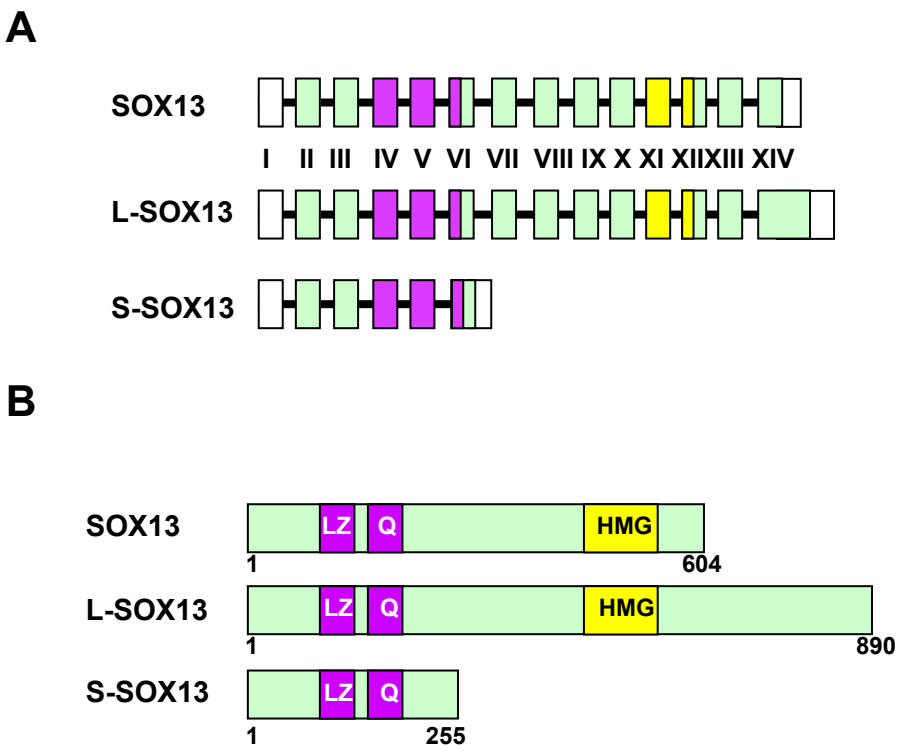


FIGURE I10. SOX13 gene variants and proteins isoforms

A) Graphic representation of SOX13 genes. Exons are depicted as boxes and black thin lines represent the introns. White boxes represent the untranslated regions. Exon size is not proportional to box size. Three different SOX13 coding sequences have been reported; SOX13 (1815 bp), L-SOX13 (2637 bp) and S-SOX13 (768 bp). **B)** SOX13 protein isoforms. SOX13 proteins have two functional domains, LZ-Q domain (purple) and the HMG domain (yellow). Functional domains are depicted in different colours that correlate with panel A.

encoding a protein of 604aa containing the HMG domain and the LZ and Q-rich domain in the N-terminus [121] (**figure I10**). *L-SOX13* (long-SOX13) [126] has an ORF of 2637 bp encoding a 890aa protein as a result of a missing G nucleotide at bp 1950 in the sequence reported by Roose et al [126] (AF083105). This deletion causes a shift of exon 14 open reading frame, leading to a longer protein. Finally, *S-SOX13* (short-SOX13) (ECRBrowser [127]) is an ORF of 768 bp encoding a 255aa protein caused by alternative splicing that results in a truncated isoform lacking the C-terminal domain, including the HMG domain (**figure I10**). Until date, it is not known which SOX13 isoforms is prevalent in human tissues.

I3.1.2. Functional domains

Human SOX13 protein has a molecular weight of approximately 67 kDa, and it has 90% homology with the murine SOX13 protein [128], suggesting conserved function of SOX13 in human and mouse.

As well as other SOX proteins, SOX13 contains a single HMG box of 79 amino acids that confers the ability to specifically bind and bend the DNA [129]. An unusual property of the HMG box is to bind the minor groove of DNA whilst most transcription factors bind the major groove [129]. Studies using oligonucleotide selection assays have determined that SRY proteins bind with high affinity to the consensus sequence **5'-$^{A}/_{T}{}^{A}/_{T}$CAA$^{A}/_{T}$G-3'** [130]. In the case of SOX13, Roose et al. [126] demonstrate by EMSA assay that it can bind to this consensus DNA sequence *in vitro*. However, in contrast with other SOX proteins such as SOX4 [131] or SOX9 [132], SOX13 is not able to transactivate transcription through a multimerized consensus binding motif [126].

The N-terminal domain of SOX D family members contains also a LZ /Q rich region. These coiled-coil domains can mediate homo and heterodimerisation between members of the family [133]. In fact, SOX13 can homodimerisate through its LZ/Q domain causing a reduction in DNA binding affinity *in vitro* [121].

I3.1.3. Interacting partner: TCF1

SOX proteins bind DNA with low sequence specificity and binding affinity when compared to other classes of transcription factors [134]. Since SOX proteins perform unique functions in different cell types, and regulate different events in the same cell type, their specificity is achieved by the differential partnership of SOX transcription factors with other transcription regulators, many of which have not been yet discovered [115-116].

Until now, there is only one well-known partner for SOX13 described in the literature, the T-cell factor 1 (**TCF1**) [135]. The interaction has been described in the context of T-lymphocytes differentiation, where SOX13 is necessary to promote $\gamma\delta$ T-cell development at the expense of $\alpha\beta$ lineage development [135].

In the thymus, where T-cell lineage migrate to continue the differentiation process, SOX13 represses Wnt the pathway producing a reduction of the proliferation and altering differentiation of the $\alpha\beta$ T-cell precursor compartment, thus promoting instead $\gamma\delta$ development [135]. At the molecular level, the repression of Wnt activity caused by SOX13 is achieved by antagonising TCF1 function, which together with β-catenin constitute the effector complex of the Wnt pathway (see **figure I11**). Specifically, SOX13 interacts with the N-terminal domain of TCF1,

sequestering TCF1 from β-cat/TCF complex, and thus, inhibiting Wnt signalling (see **figure I11**).

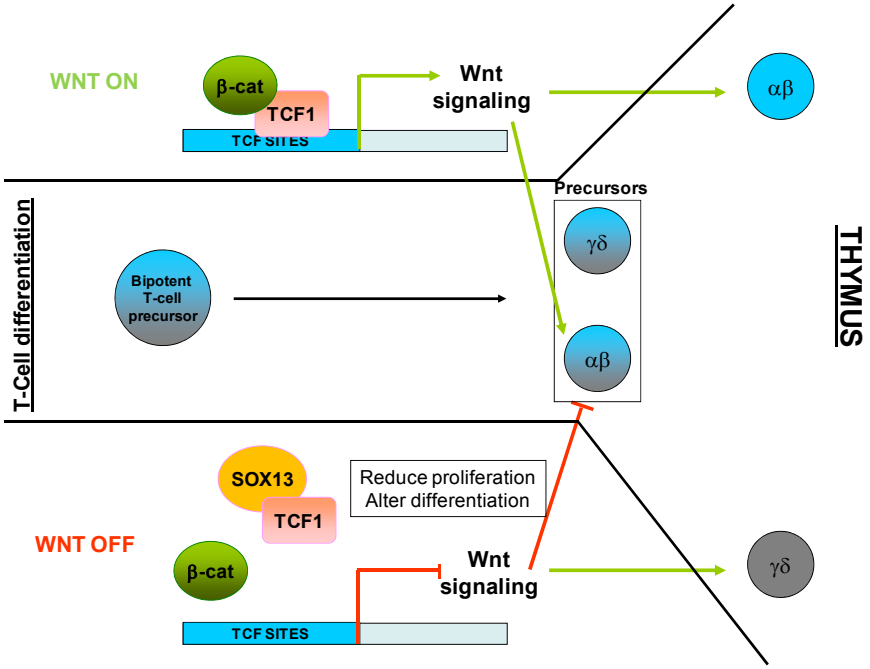


FIGURE I11. SOX13·TCF1 interaction regulates T-cell differentiation
During T-cell differentiation, a bipotent T-cell precursor present in the thymus gives rise to both αβ and γδ T-Lymphocytes. SOX13 interaction with TCF1 represses Wnt signalling reducing the proliferation and promoting γδ T-cell differentiation at the expense of αβ precursors.

I3.2. Cell biology and function

I3.2.1. Expression

During development, SOX13 is found at different developmental stages in some tissues such as thymus, arteries, cartilage of the limbs, brain, kidney, pancreas, and liver as well as in the visceral mesoderm of the extra-embryonic yolk sac and spongiotrophoblast layer of the placenta

[126, 136-137]. By Northern Blot, *SOX13* mRNA is detected in multiple adult tissues, including heart, brain, placenta, lung, liver, kidney and pancreas [121]. In the pancreas, SOX13 protein is present in the Islets of Langerhans as well as in isolated cells in the exocrine compartment [121].

I3.2.2. Animal models

Sox13 deficient mouse was generated by Melichar et al [135] by excision of exons 4 to 11 via homologous recombination. Most of Sox13 KO mice die shortly after birth, while others have significant dwarfism and severe abnormalities, and die by three weeks of age (see **figure I12**). [135]. During embryonic development, no overt developmental defects have been described [135], which could be explained by a possible functional redundancy of SOX13 with other members of the D family, i.e. SOX5 and SOX6 [138]. Double or even triple knockout animals for SOX13, SOX5 and SOX6 will be necessary to elucidate the importance of SOX13 in the different systems and organs where it is present, such as pancreas and liver.

Sox13 transgenic mice have also been generated by the group of Joosoon Kang to further study its role in T-cell development [135]. SOX13 was overexpressed under the control of Lck proximal promoter, which is active during early thymocyte development. Ectopic expression of SOX13 produces an inhibition of T-cell precursors proliferation and survival. Generation of other SOX13 transgenic animals will help to elucidate SOX13 role in other systems.

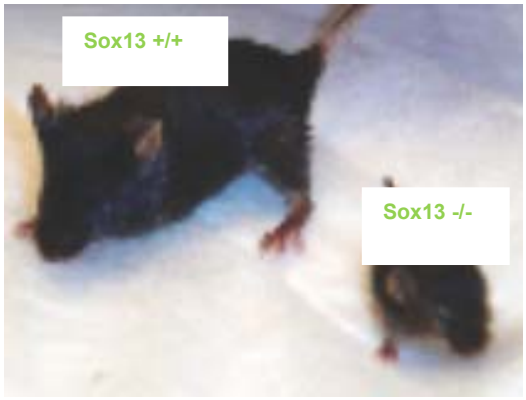


FIGURE I12. Sox13 knockout mouse phenotype
Picture of a 2-weeks old WT (left) and Sox13-/- (right) mice. Significant dwarfism is evident in the knockout animal when compared with the wild type. *Adapted from Melichar, H.J. et al (2007) [135].*

I3.2.3. The biological role of SOX13

Besides SOX13 function in T-cell differentiation mentioned above, SOX13 has been also linked to the development of the central nervous system (CNS) [139-140]. SOX13 expression during CNS development is confined to a specific subset of post-mitotic differentiating neurons while it is absent in the proliferating ventricular zone, suggesting a role for SOX13 in the specification and/or differentiation of neurons in the developing CNS [139]. On the other hand, SOX13 also seems to be important during cranial motoneurons differentiation, acting downstream of the transcription factor Phox2b. Down-regulation of *SOX13* is essential to activate the genetic program that leads to proper cranial motoneurons differentiation [140]. Further studies of the SOX13 knockout mouse in the CNS will help to elucidate the functions of SOX13 in neural development.

I4. C-MYC TRANSCRIPTION FACTOR

In 1911 Peyton Rous observed that chicken sarcoma could be transmitted through cell-free extracts from the tumours, suggesting that a virus could be the etiologic agent of these sarcomas. On the basis of the work by Bishop and co-workers, studies of a specific subgroup of avian retrovirus, which induces myeloid leukaemia, sarcomas, liver, kidney, and other tumours in chickens, led to the identification of the v-myc oncogene [141-144]. The discovery of a homologous gene, termed c-Myc, in chicken supported the view that oncogenic avian retroviruses could capture cellular growth regulatory genes and transmit the activated gene. The finding that human cancers frequently display altered expression of human c-Myc underscores the importance of this gene in the origin of human cancers. On the other hand, c-Myc was found to be also critic in physiological conditions, being a key factor in the regulation of many aspects of normal cell behaviour.

I4.1. Biochemical features

I4.1.1. Genomic structure

The c-Myc gene is located on human chromosome 8q24, consisting of three exons. Its transcription may be initiated at one of three promoters. Translation at the AUG start site in the second exon produces a major form of 439 amino acids, the 64 kDa c-Myc protein (also termed Myc2). Alternative translational initiation start sites result in both longer and shorter forms of the protein, termed p67 Myc (or Myc1) and MycS, respectively [145] (**figure I13**).

I4.1.2. Functional Domains

The c-Myc sequence contains three conserved N-terminal domains, termed Myc boxes (Mb), which are also found in closely related proteins, N-myc and L-myc.

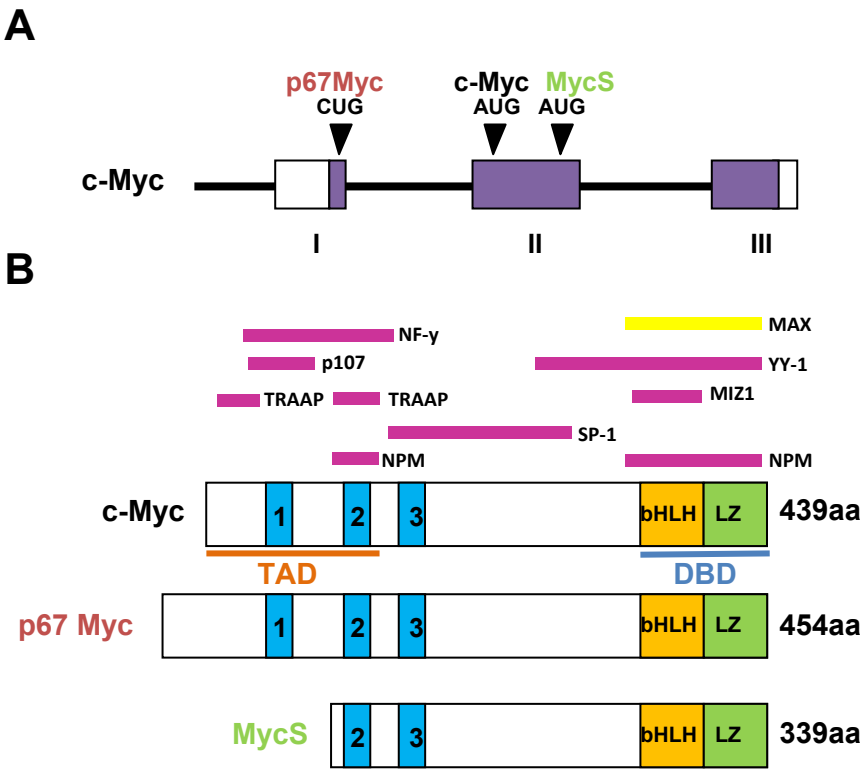


FIGURE I13. C-Myc gene, isoforms and interacting partners

A) Graphic representation of c-Myc gene. Exons are depicted as boxes (roman numbers) and black thin lines represent the introns. White boxes represent the untranslated regions and coloured boxes represent the different exons. **B)** C-Myc isoforms (c-Myc, p67 Myc and MycS) are the result of alternative translational initiation start sites. They share different domains, Myc boxes 1, 2, 3, (in blue), the bHLH (basic helix-loop-helix) domain in orange and the leucine-zipper (LZ) domain in green. Some of the c-Myc binding proteins are shown in purple boxes spaning the interaction domain in c-Myc. Max domain is shown as a yellow box. *TAD: Transcriptional Activator Domain; DBD: DNA binding domain*

The C-terminal region of c-Myc contains a dimerization motif, termed the basic helix-loop-helix leucine zipper (bHLH-LZ). The bHLH-LZ domain mediates homotypic or heterotypic dimerization with other bHLH-LZ proteins. The c-Myc dimerization domain is necessary for cellular transformation, and in 1991 the bHLH-LZ protein Max was identified as a c-Myc obligate partner protein [146-147]. The c-Myc mRNA and protein have short half lives (30 min and 20 min, respectively) as compared to those of Max (3 hrs and >24 hrs, respectively), and in most systems c-Myc appears to be the limiting, regulated component of the heterodimer. Dimerization of c-Myc and Max through the bHLH-LZ domain aligns the adjacent basic regions on each molecule to grip onto specific DNA hexanucleotide core sequences, termed **E boxes** (**5'-CA$^{C}/_{T}$GTG-3'**). Max can also bind Mad, another bHLH-LZ family protein, which mediates transcriptional repression. Mad levels, as opposed to c-Myc, increase during differentiation.

DNA-bound Myc-Max complexes activate transcription through the amino terminal 143 amino acids of c-Myc, termed the transcriptional activation domain (TAD). A small segment of this region is also required for c-Myc mediated transcriptional repression, although the mechanism is not well understood.

I4.1.3. Interacting partners

The DNA bound Myc-**Max** heterodimer interacts through the c-Myc N-terminal region with a variety of proteins involved in transcription (see **figure I13** for a summary). These include **TRRAP** (transactivation/ transformation-associated protein), which associates with histone acetylase GCN5. Acetylation of histones may then mark chromatin to

allow access of transcription factors that belong to the general transcriptional machinery, such as TFIIE or TBP, to DNA. The Mad-Max complexes, in contrast to Myc-Max complexes, recruit histone deacetylases that induce compact chromatin structures, which in turn limit access of transcription factors to DNA.

Other putative c-Myc binding proteins have been implicated in regulating c-Myc transactivating or transrepressing properties, including **p107** and **Miz-1** (Myc-interacting zinc finger-1). The binding of p107 to N-terminal domain of c-Myc facilitates the formation of a quaternary protein complex consisting of c-Myc, p107, and Cyclin A/CDK [148-149]. The recruitmentof Cyclin A/CDK to c-Myc facilitates the phosphorylation of c-Myc, hindering the ability of c-Myc to transactivate gene expression. In the case of Miz-1, the interaction occurs through the C-terminal bHLH domain of c-Myc, without disrupting Myc-Max complex. This multi-protein complex is recruited to core promoters and is thought to displace co-activators and recruit co-repressors, resulting in transcriptional repression [150-152]. Other c-Myc interactor partners that are supposed to have similar functions include specificity protein-1 (**SP1**), nuclear factor Y (**NF-Y**), and yingyang-1 (**YY1**) [153-155].

Nucleophosmin (**NPM**, also termed B23), a multifunctional protein deregulated and overexpressed in several different types of cancers, is a recent described c-Myc partner [156]. It interacts with MbII and bHLH-LZ domains of c-Myc, not altering Myc/Max binding. NPM interaction regulates expression of endogenous c-Myc targets, stimulating c-Myc induced proliferation and transformation.

I4.2.3. Target genes

To better understand c-Myc function, recent efforts have concentrated on identifying the genetic program induced by c-Myc. c-Myc target genes were identified using a variety of techniques including DNA microarray and SAGE. However, only a few c-Myc targets, identified primarily through ectopic expression approaches, have been proven to contribute to c-Myc functions.

FUNCTIONAL CLASS	TARGET GENES
Cell cycle	Cyclin D1, Cyclin D2,CDK4,cdc25a p21, p15, GADD45
Differentiation	CEBP
Cell Growth, Metabolism and Protein Synthesis	Lactate deshydrogenase, CAD, ODC, ribosomal proteins, EIF4E, EI2FA
Cell adhesion and migration	N-Cadherin, integrins
DNA breaks and chromosomal instability	MAD2, TOP1, Cyclin B1

TABLE I3. C-Myc target genes

List of some of the c-Myc target genes involved in different c-Myc dependent processes: cell cycle, differentiation, cell growth, metabolism, protein synthesis, cell adhesion and migration and chromosomal instability. c-Myc induced genes are bolded in green, whereas repressed genes are in red.

Chromatin immunoprecipitation (ChIP) approaches have allowed the identification of true direct targets of c-Myc, confirming previously identified targets and revealing new ones. A significant number of genes

are bound by c-Myc only in specific cells, and only a fraction of these are regulated by c-Myc [157-158]. These target genes are involved in several functions in which c-Myc is playing an essential role, such as proliferation, apoptosis, differentiation, growth and transformation (some of the target genes are included in **table I3**).

I4.2. Cell biology and function

I4.2.1. Expression

In normal cells, c-Myc expression is tightly regulated in response to growth signals. It is ubiquitously expressed during embryogenesis and in tissue compartments of the adult possessing high proliferative capacity. In fact, mouse embryos in which both alleles of c-Myc have been deleted die in utero around E10.5 with severe defects in vasculogenesis and angiogenesis [159]. Non-proliferating or quiescent cells generally express non-detectable levels of *c-Myc,* but the gene is rapidly induced following mitogenic stimulation and, thereafter, continues to be expressed (at very low levels) in proliferating cells.

Deregulated c-Myc expression is often associated with aggressive, poorly differentiated tumours. Activation of c-Myc gene occurs in several ways. In many tumours, c-Myc activation can be attributed to structural alterations affecting the c-Myc gene, including chromosomal translocations, retroviral promoter or enhancer insertions, and gene amplification.

I4.2.2. Regulation of c-Myc expression

Expression of the c-Myc proto-oncogene is induced as a consequence of signalling pathways downstream of a variety of mitogens such as

platelet-derived growth factor (PDGF), epidermal growth factor (EGF), basic fibroblast growth factor (bFGF), [160-162]. The components of the signal transduction pathways which drive c-Myc transcription have not been fully elucidated, yet studies have identified a few of the players such as protein kinase C (PKC), protein kinase A (PKA), src, E2F, ets, and abl [163-170].

I4.2.3. The biological role of c-Myc

FISIOLOGICAL CONDITIONS

- Proliferation: cell cycle control

Cell division consists of two consecutive processes, mainly characterized by DNA replication and segregation of replicated chromosomes into two separate cells. Originally, cell division was divided into two stages: mitosis (M), i.e. the process of nuclear division; and interphase, the interlude between two M phases. Stages of mitosis include prophase, metaphase, anaphase and telophase. Under the microscope, interphase cells simply grow in size, but different techniques revealed that the interphase includes G1, S and G2 phases [171]. Replication of DNA occurs in a specific part of the interphase called S phase. S phase is preceded by a gap called G1 during which the cell is preparing for DNA synthesis and is followed by a gap called G2 during which the cell prepares for mitosis. G1, S, G2 and M phases are the traditional subdivisions of the standard cell cycle (see **figure I14**) [172]. Cells in G1 can, before commitment to DNA replication, enter a resting state called G0. Cells in G0 account for the major part of the non-growing, non-proliferating cells in the human body.

The eukaryotic cell cycle is regulated by the coordinated action of Cyclins and Cyclin-dependent kinases (CDKs). The initial transition from G0 to G1 phase and the subsequent transition from G1 to S phase are mediated by a series of sequential regulatory events (**figure I14**). D-type Cyclins are the first group of Cyclins to be synthesized and their expression is closely coupled to the action of growth factors [173-175]. D-type Cyclins bind and activate CDK4 and CDK6 [176]. The major targets of the Cyclin D/CDK4 and Cyclin D/CDK6 complexes are the retinoblastoma protein (Rb) and the related proteins p107 and p130 [177-179]. Phosphorylation of Rb in mid-G1 leads to the release of active forms of the E2F family of transcription factors [180]. Targets of E2F identified to date include Cyclin E, Cyclin A and many S phase-specific genes [181-183]. Cyclin E forms an active complex with CDK2 and this complex, which also phosphorylates Rb, is necessary for the orderly completion of the G1 to S phase transition [184]. Cyclin A binds to CDK2 and this complex is required during S phase [185-186]. In late G2 and early M, Cyclin A complexes with CDK1 to promote entry into M phase. Mitosis is further regulated by Cyclin B1 in complex with CDK1 [187].

The activity of CDKs is regulated by a variety of mechanisms. These include Cyclin and Cyclin-dependent kinase inhibitors (CKIs) binding as well as post-translational modifications. The CKIs are currently classified into two groups [188]. The first group, known as the Cip/Kip family, consists of the p21, p27, and p57 proteins. These inhibitors require pre-formed Cyclin/CDK complexes for binding, and can inhibit all Cyclin/CDK complexes *in vitro* [189-193]. The second group of inhibitors, known as the INK family, consists of the p15, p16, p18, and p19 proteins. Unlike the Cip/Kip family, these inhibitors are only active on CDK4/6 containing

complexes. In addition, binding of the INK proteins to CDK4/6 is independent of Cyclins [194-198]. Members of both families of inhibitors have been shown to be important for executing growth arrest signals in response to a variety of signals, such as DNA damage, senescence, contact inhibition, and TGF-β treatment [188]. Phosphorylation, both inhibitory and activating, represents another major mode of CDK regulation [199-201].

C-Myc promotes cell proliferation by activating or repressing target genes involved in cell cycle progression, specifically for G0/G1 to S phase progression (see **figure I14**) [202]. CCND2 (which encodes Cyclin D2), CCND1 (Cyclin D1) and CDK4 are direct target genes activated by c-Myc [203-205]. So, c-Myc increases Cyclin D/CDK4 activity resulting in Rb hyperphosphorylation and subsequent release of E2F. E2F, in turn, induces the expression of essential genes for S-phase, including Cyclin E, permitting the progression from G0 to G1. On the other hand, c-Myc also acts by repressing genes, such as the CDK inhibitors p15 and p21 that are involved in cell cycle arrest. Later on, c-Myc induces Cyclin E/CDK2 activity early in the G1 phase of the cell cycle, which is regarded as an essential event in Myc-induced G1–S progression [206-207]. Moreover, it has also been involved in the degradation of the CDK inhibitor p27 by inducing other c-Myc target genes, CUL-1 and CKS [208-209], increasing Cyclin E/CDK2 activity [203].

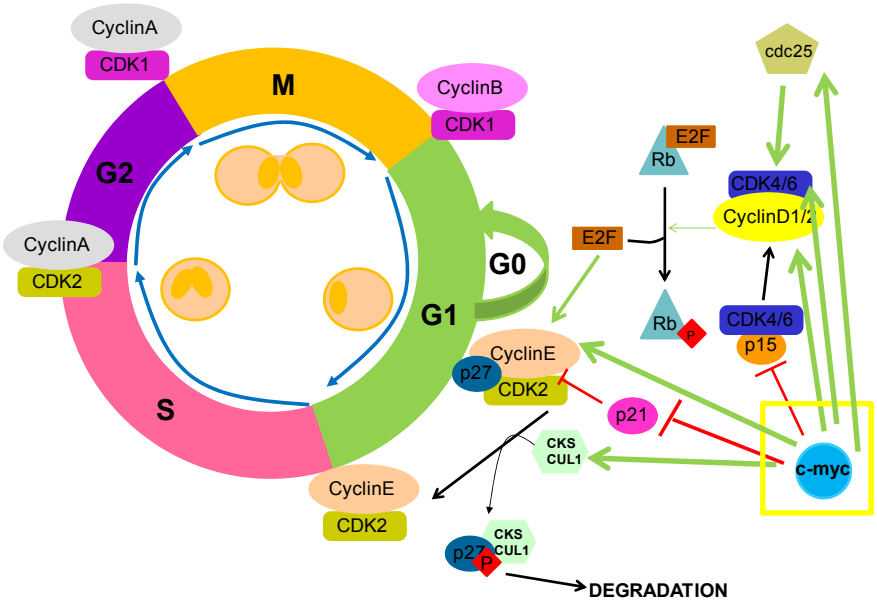


FIGURE I14. Role of c-Myc regulating G0-G1 and G1-S cell cycle transition

Summarized scheme of eukaryotic cell cycle and c-Myc's role regulating G0-G1 and G1-S phase transitions. To permit quiescent cells to enter in the cell cycle from G0, c-Myc oncoprotein induces expression of CDK4/CDK6 and Cyclins D, allowing CDK/Cyclin complex to phosphorylate Rb, leading then the release of E2F active form. E2F transcription factor, induces then the expression of S-phase genes, including Cyclin E. On the other hand, c-Myc also acts by repressing the expression of the CDK inhibitors of the INK (p15) and CIP/KIP (p21) families. Later, c-Myc induces the expression of CUL1 and CKS that participates in the degradation of other CDK inhibitor, p27, permitting the progression of the cell cycle. *Adapted from Obaya, A.J. et al (1999)[202] and Vermeulen, K. et al (2003) [172].*

- Cell growth

The ability of c-Myc to promote cell growth (causing cells to double in mass and size) was shown in normal and tumour cells, both *in vitro* and *in vivo*. c-Myc promotes cell growth by providing the cell with an abundant supply of several classes of basic building blocks as well as increasing cell metabolism and protein synthesis [210-211]. When c-Myc is activated, cellular growth is no longer rate limiting to the proliferative

process. Several c-Myc target genes are thought to have a role in this activity, including those associated with cellular metabolism, ribosomal and mitochondrial biogenesis, and protein and nucleic acid synthesis [212].

- Differentiation

Consistent with the role of c-Myc as a strong proliferative stimulus, c-Myc expression can inhibit the differentiation program in a number of cellular systems. Indeed, in a variety of cultured cells, exposure to inducers of differentiation results in the repression of c-Myc expression and a withdrawal from the cell cycle [213-217]. However, c-Myc can also stimulate differentiation of human epidermal stem cells [218]. Thus, the role of c-Myc in the inhibition of differentiation is dependent upon the cell type.

- Apoptosis

c-Myc expression not only promotes proliferation, but also induces or sensitizes cells to apoptosis. Overexpression of c-Myc under circumstances in which this gene is usually down regulated such as serum deprivation, results in apoptotic cell death [219]. Then, c-Myc deregulation results in a cellular state in which increased proliferation or apoptotic death are both equally possible depending on the cellular microenvironment.

- Cell competition

Studies in the c-Myc homolog in *Drosophila melanogaster*, dMyc, have highlighted the role of c-Myc on promoting apoptosis of neighbouring cells. Cells with higher levels of c-Myc activity act as supercompetitors to potentiate the programmed death of surrounding normal cells [220-221].

This process, known as cell competition or supercompetition, could be at the root of c-Myc depending tumourigenesis, by facilitating the expansion of high c-Myc-expressing preneoplastic clones.

PATHOLOGICAL CONDITIONS: CANCER

In humans, the array of tumour types which frequently exhibit alterations in c-Myc expression is extensive: hematopoietic malignancies such as lymphoma and leukemia, connective tissue cancers such as osteosarcoma, carcinomas derived from the epithelial layers of the breast, lung, cervix, ovary, stomach, prostate, and colon, as well as squamous cell carcinomas of head and neck [222-223] . As mentioned above, c-Myc activation can be attributed to structural alterations affecting the c-Myc gene, including chromosomal translocations, retroviral promoter or enhancer insertions, and gene amplification. Furthermore, overexpression of c-Myc in transgenic mice results in tumour development [224]. C-Myc promotes cell proliferation, genomic instability and transformation by accelerating cells through G1 and S phases of the cell cycle, abrogating cell cycle checkpoints, and increasing cell metabolism. In many settings these alterations will lead to apoptosis, or cell death. But in the background of additional mutations that activate anti-apoptotic signals, c-Myc can lead to full blown neoplastic transformation.

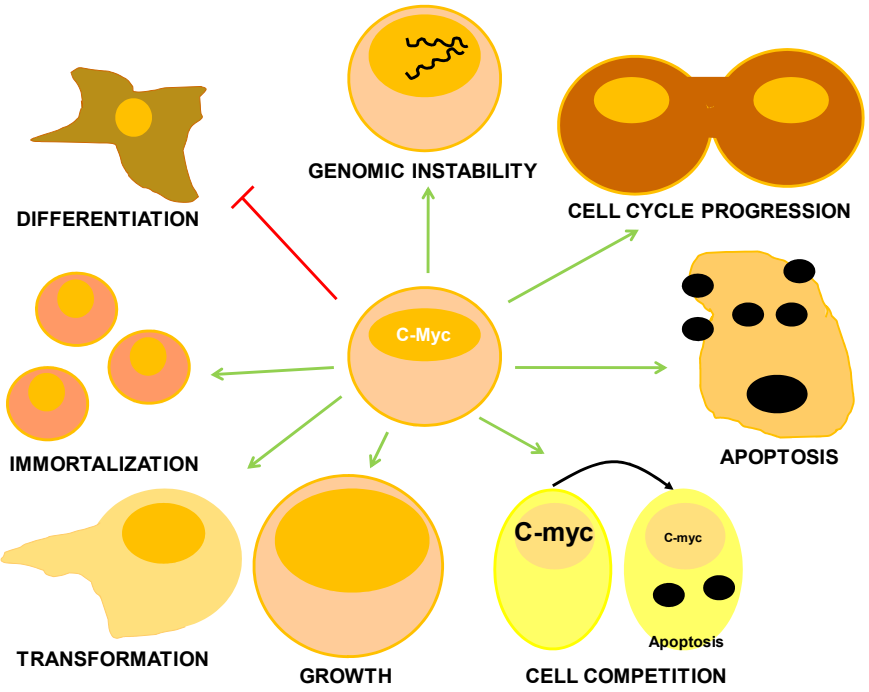


FIGURE I15. C-Myc functions

Scheme of some of the functions attributed to c-Myc protein. C-myc acts by regulating downstream target genes producing cell cycle progression, apoptosis, cell growth, genomic instability, transformation, immortalization and cell competition, among others. It also inhibits cell differentiation in some circumstances.

HYPOTHESIS

Considering Hhex pleiotropic effect during development and its dual role as a transcriptional repressor or activator, we hypothesize that Hhex activity and function should be modulated by different tissue-specific partners.

OBJECTIVES

The major goal of this project is to further characterize Hhex transcription factor by studying Hhex interaction partners, gaining insights into its molecular mechanism of action.

The specific objectives are:

1- To identify putative Hhex interactors.

2- To characterise at the biochemical level the most promising interactions.

3- To determine the functional relevance of these interactions.

EXPERIMENTAL PROCEDURES

EP1. DNA ANALYSIS TECHNIQUES

EP1.1. Plasmids construction

Gal4-Hhex fusion proteins (Yeast Two-Hybrid: bait)

The plasmids used in the yeast-two-hybrid derive from the yeast plasmid pBD-Gal4-Cam (Stratagene). It contains the TRP1 gene for selection in yeast and the chloramphenicol-resistance gene for posterior selection with chloramphenicol in E. coli. In order to construct the baits for the yeast-two hybrid screening, different Hhex fragments were amplified by polymerase chain reaction (PCR) such that an artificial EcoRI restriction site at the 5' end and a SalI restriction site at the 3' end flanked the product. Following digestion with EcoRI and SalI, the fragments were cloned into the pBD-Gal4-Cam vector, rendering 5 bait plasmids containing Hhex(1-137), Hhex(1-196), Hhex(138-271), Hhex(197-271) or full-length Hhex (1-271) in frame with the Gal4 DNA binding domain (Gal4-DBD).

GST-Hhex fusion proteins

GST-Hhex fusion proteins were produced in bacteria using plasmids derived from the pGex-6P-1 vector (GE Healthcare). EcoRI-SalI fragment from the pBD-Gal4 Cam series (above) containing the Hhex-related sequence were subcloned into the EcoRI-SalI site in pGex-6P-1.

EGFP-Hhex fusion proteins

EGFP-Hhex protein expression vectors were constructed by subcloning the BamHI-SalI fragment from pGex-6P-1 Hhex series into the BglII-SalI site in pEGFP-C1 (Clontech).

Mammalian- expression plasmids

HA-tagged Hhex

Expression plasmids containing HA-tagged full-length Hhex, Hhex(1-271) and Hhex(1-196) were obtained by EcoRI-SalI digestion of the corresponding pBD-Gal4-Cam vector and subcloning into the EcoRI-XhoI site in pIRES-hrGFP-2a (Stratagene). Hhex(138-271) was cloned by PCR to introduce an initiation codon at the 5'-end of the sequence. Hhex(138-271) was amplified using the Expand High Fidelity PCR System (Roche Applied Science), digested with EcoRI-SalI and ligated to EcoRI-XhoI fragment of pIRES-hrGFP-2a.

SOX13

Expression plasmids containing FLAG and myc tagged SOX13 isoforms (604aa and 255aa) were obtained by PCR-cloning from a human cDNA library (Ambion) such that an artificial EcoRI restriction site at the 5' end and a SalI restriction site at the 3' end flanked the product. Total RNA (1 µg) was reverse transcribed using Superscript II (Invitrogen) and oligo(dT)15 as primer. SOX13 cDNAs (1 µl) were amplified using the Expand High Fidelity PCR System (Roche Applied Science) with the addition of 6% dimethyl sulfoxide. After denaturing for 3 min at 95°C, amplification was performed by four cycles of 40 s at 94°C, 45 s at 50°C, and 2 min at 74°C plus 34 cycles of 40 s at 94°C, 45 s at 67°C and 2 min at 74°C and a final extension of 5 min at 74°C. To obtain ΔSOX13, the fragment aa1-aa100 was cloned by PCR such that an artificial EcoRI restriction site at the 5' end and a HindIII restriction site at the 3' end flanked the product. The fragment aa210-aa604 similarly but an artificial HindIII restriction site at the 5' end and a SalI restriction site at the 3' end

flanked the product. The resulting EcoRI-SalI flanked PCR products were ligated to EcoRI-SalI fragment of pFLAG-CMV2 (Sigma) and EcoRI-XhoI fragment of pcDNA3-myc-tag (a kind gift from Dr. Gabriel Gil) to obtain pFLAG-SOX13-255, pFLAG-SOX13-604, pFLAG-ΔSOX13, pcDNA3-myc-255, pcDNA3-myc-SOX13-604 (used for *in vitro* translation) and pcDNA3-myc-ΔSOX13.

C-Myc

C-Myc expression plasmid used for generating *in vitro* translated c-Myc protein was obtained by subcloning c-Myc from pECFL-Myc (a gift from Javier León) into the HindIII-KpnI site of pcDNA3. Full-length HA-tagged c-Myc expression construct as well as HA-tagged c-Myc deletion constructs c-Myc ΔMb (aa2-41 + 180-439), c-Myc ΔC1 (aa1-390), c-Myc ΔC2 (aa1-361), c-Myc ΔC3 (aa1-262) and c-Myc ΔN1 (aa148-439) were a kind gift from Luciano di Croce Laboratory.

TCF1 and β-catenin

Myc-tagged TCF1 was obtained by PCR amplification from TCF1-pcDNA3.1/Zeo vector (a kind gift from Elena Sancho) such that an artificial EcoRI restriction site at the 5' end and a SalI restriction site at the 3' end flanked the product. Following digestion with EcoRI and SalI, TCF1 was cloned in frame into the EcoRI-XhoI site in pcDNA3-myc-tag. Activated β–Catenin expression plasmid (pcDNA3 S37Y β–Catenin) was a kind gift of Antonio Garcia de Herreros.

Reporter plasmids

The reporter plasmid pGL3-NTCP was obtained by PCR-cloning from rat genomic DNA such that artificial HindIII sites were inserted at the 5' and

3' end. The PCR product was digested with HindIII and ligated to the HindIII site in pGL3-Basic (Promega). Reporter plasmid for measuring c-Myc dependent activity, pMyc TA-LUC and the control empty vector TA-LUC were from Clontech. Goosecoid Luciferase (Gsc-LUC) reporter plasmid was a gift of Joshua Brickman. Topflash and Fopflash reporter plasmids for measuring Wnt activity were kindly provided by Hans Clever Laboratory. The reporter plasmid 3xSx, to control Sox-dependent activity was a gift of Michael Wegner.

All constructed plasmids were verified by sequencing.

EP2. CELL CULTURE, TRANSFECTION AND INFECTION

EP2.1. Cell lines

HeLa cells, 3T3 cells and the rat cell lines H015.19 and TGR1 (a kind gift of Luciano di Croce Laboratory) were maintained as monolayer cultures and grown in DMEM + Glutamax (GIBCO; 31966) supplemented with 10% newborn calf serum, 50 U/ml penicillin, and 50 mg/ml streptomycin. 293 and 293T cells were grown as HeLa cells but also supplemented with 3.5 g/L of glucose. MEFs were obtained from primary culture of mouse embryos (E13.5) and maintained with complete DMEM media as HeLa cells. Human Fibroblast Foreskin cells (FSK, CRL-2429) were maintained in Iscoves Modified Dulbeccos's medium (IMDM) + Glutamax (GIBCO; 31980) supplemented with 10% newborn calf serum, 50 U/ml penicillin, and 50 mg/ml streptomycin. The hepatoma derived cell line HepG2 was maintained in F12/L15 media supplemented with 10% newborn calf serum, 50 U/ml penicillin, and 50 mg/ml streptomycin.

EP2.2. Introduction of DNA into mammalian cells

Transfection

Plasmid DNAs were purified on Qiagen Maxiprep kit columns (Qiagen) and quantified by OD_{260}. Cell lines were transfected using Fugene HD or Fugene 6 reagents (Roche) depending on the cell type following manufacturer instructions.

Infection

Adenovirus

Recombinant adenovirus encoding Hhex was prepared by homologous recombination of pJM17 and pACCMV-derived plasmid as described in McGrory et al. [225]. pAC-HA plasmid was derived from pACCMV by inserting an HA-epitope sequence in the 3' end of the multicloning site. Hhex was amplified by PCR such that an artificial EcoRI restriction site at the 5' end and a BamHI restriction site at the 3' end flanked the product. The DNA fragment was amplified using the Expand High Fidelity PCR System (Roche Applied Science), digested with EcoRI-BamHI and ligated to the EcoRI-BamHI fragment of pAC-HA to obtain pAC-Hhex-HA. pAC-HA and pAC-Hhex-HA plasmids were cotransfected with pJM17 into 293 cells (AdE1A-transformed human embryonic kidney cells) by calcium phosphate/DNA co-precipitation. The CMV driven cassette of pAC-CMV is located between the sequences representing 0–1.3 and 9.2–16mU of the adenovirus type 5, whereas pJM17 encodes a full-length adenovirus-5 genome (dl309) interrupted by the insertion of the bacterial plasmid pBRX at position 3.7 mU, thereby exceeding the packaging limit for adenovirus. Homologous recombination between adenovirus sequences in the transfer plasmid (pAC-HA) and in the pJM17 plasmid results in the

substitution of the pBRX sequences in pJM17 by the chimeric gene. This generates a genome of packable size in which most of the adenovirus early region 1 is lacking, thus rendering the recombinant virus replication defective. The resulting viruses (called Adempty and AdHhex) were plaque-purified, expanded into a high concentration stock using Vivapure Adenopack (Sartorius), and titrated by plaque forming assay (AdEmpty: $6x10^{10}$ pfu/ml; AdHhex: $2.5x10^{10}$ pfu/ml). Adenovirus toxicity was tested by infection of FSK cells with a doses range of each adenovirus. Cell mortality was controled every 24 hours by taking pictures, and levels of protein expression were confirmed by immunoblot. Recombinant adenovirus for MadMyc was acquired from Vector Biolabs with a titer of $1x10^{10}$ pfu/ml.

An adenoviral vector expressing shRNA against Hhex was obtained by homologous recombination of pJM17 and pAC-shHhex as described [226]. pAC-shHhex was obtained by inserting synthesized Hhex shRNA (targeting sequence in HHex, TGGACAGTTCCTGTGATCAGAG) into the BamHI-EcoRI site in pAC-shRNA plasmid. pAC-shRNA plasmid was obtained by replacing the CMV cassette in pAC-CMV with the CMV expression cassette from pPRIME-CMV-GFP [227]. An adenovirus expressing shRNA against luciferase was constructed similarly and used as an infection control.

Retrovirus

RetHhexGFP, RetGFP and RetMyc retrovirus were generated by transfection of the plasmids MIGR1-GFP-IRES-HHEX, MIGR1-GFP and pMX-Myc with the retrovirus packaging Vector pCL-ECO in 293T cells. The retroviral particle-containing supernatants were collected 48 hours post-

transfection, filtered through a 0.45-µm filter and stored at –80°C until use.

EP3. RT-PCRs

EP3.1. Quantification of SOX13 mRNA levels in human tissues

DNA free total RNA from human tissues was obtained from Ambion. RNA (1 µg) was reverse transcribed as described [228]. Diluted cDNA (3 µl) was amplified with a rapid thermal cycler (LightCycler Instrument; Roche Diagnostics) in 15 µl of LightCycler DNA Master SYBR Green I (Roche Applied Science), 5 mM MgCl2, and 0.3 µM of each oligonucleotide. We designed specific primer sets for three putative splicing variants of SOX13 (**table EP2**). In parallel, we always analyzed the mRNA concentration of the human housekeeping porphobilinogen deaminase (hydroxymethylbilane synthase, PBGD) as an internal control for normalization. A stable expression of the housekeeping porphobilinogen deaminase gene was validated by comparison with TATA box-binding protein expression as a second constitutive control gene (Human TBP Primer Set; Invitrogen). We found that the expression ratio of these two internal control genes was practically constant in the different tissues and cells investigated. Moreover, human porphobilinogen deaminase and TATA box-binding protein do not harbor pseudogenes and show genomic stability in cancer. PCR amplicons were confirmed to be specific by size (agarose gel electrophoresis), melting curve analysis and eventual sequencing. After denaturing for 30 s at 95 °C, amplification was performed in 40 cycles of 1 s at 94°C, 5 s at 62 °C, and 20 s at 72 °C. The real-time monitoring of the PCR reaction and the precise quantification of the products in the exponential phase of the amplification were

performed with the LightCycler quantification software according to the manufacturer's recommendations. Reproducibility of the measurements was assessed by conducting triplicate reactions.

EP3.2. Quantification of c-Myc/Wnt target genes mRNA levels in FSK cells

Total mRNA from infected FSK cells (c-Myc target genes study) or WNT activated and tranfected 293T cells (WNT target gene study) was extracted using RNeasy Plus Mini Kit (QIAGEN). Retrotranscription was performed with iScript cDNA synthesis kit (Bio-Rad) and 1/5 diluted cDNA (3 µl) was amplified with a rapid thermal cycler (LightCycler480 Instrument; Roche Diagnostics) in 15 µl of LightCycler DNA Master SYBR Green I (Roche Applied Science), and 0.3 µM of each oligonucleotide. We designed specific primer sets for human c-Myc [158] and Wnt target genes expression [229-230] (see **table EP2**). In parallel, we always analyzed the mRNA concentration of the human housekeeping PBGD PCR amplicons were confirmed to be specific by size (agarose gel electrophoresis), melting curve analysis and eventual sequencing. quantification of the products in the exponential phase of the amplification were performed with the LightCycler quantification software according to the manufacturer's recommendations. Reproducibility of the measurements was assessed by conducting triplicate reactions.

EP3.3. Quantification of c-Myc target genes mRNA in mouse embryo

Liver bud was dissected from WT and KO E9.5 embryos as described previously (see representative image in **results figure R2.6**) [91] and RNA was extracted with the RNA micro Kit (Qiagen). Retrotranscription was

performed with iScript cDNA synthesis kit (Bio-Rad) and 1/5 diluted cDNA (5 µl) was amplified with a thermal cycler (iQ5 Real-Time PCR Detection System, Bio-Rad) in 25 µl of iQ SYBR Green Supermix (Bio-Rad) and 0.3 µM of each oligonucleotide. We designed specific primer sets some c-Myc target genes as well as for mouse Hhex and c-Myc (see **table EP2**). In parallel, we always analyzed the mRNA concentration of the mouse housekeeping glyceraldehyde 3-phosphate dehydrogenase (GAPDH) as an internal control for normalization. PCR amplicons were confirmed to be specific by size (agarose gel electrophoresis) and melting curve. After denaturing for 4 min at 95°C, amplification was performed in 45 cycles of 15 s at 95°C, 15 s at 58 °C, and 30 s at 72 °C.

EP4. PROTEIN ANALYSIS

EP4.1. Immunofluoresence

HeLa cells growing in coverslips were co-transfected with Hex-EGFP and Sox13-myc plasmids. After 48 hours cells were fixed during 5 minutes in PFA 4%, and permeabilizated with PBS containing 0.2% TritonX-100 during other 5 minutes. Blocking was carried out with 1% BSA in PBS solution during 30 minutes. Primary mouse anti-myc antibody (9E10 hybridoma) was incubated using a 1/30 dilution overnight at 4ºC. Non-primary antibody incubated coverslips were used as a negative control (data not shown). Secondary anti-mouse Alexa555 conjugated antibody (1/500 dilution, Invitrogen) was incubated during 1 hour in PBS-BSA 1% solution. Cells were washed three times with PBS and the coverslips were mounted with Mowiol® Medium in microscope slides. Fluoresence was viewed through a Leica TCS-SP2 confocal microscope.

EP4.2. Protein extraction

Cell protein extracts were obtained by harvesting cells transfected/infected after 48/72 h with M-Per buffer containing Halt Protease Inhibitors Cocktail (Pierce). Cell debris was removed by centrifugation at 14000×g for 10 min at 4°C and quantified by performing a Bradford assay (Bradford Reagent, SIGMA).

EP4.3. Immunoblot

Protein extracts were mixed with 4X Laemmli buffer, heated for 5 minutes and resolved on SDS-PAGE gels of different acrylamide percentage, depending on the molecular weight of the protein, at 100V in 1x running buffer (25 mM Tris-base, 200mM glycine, 0,1% (w/v) SDS). Proteins were then transferred onto a Immobilon filters (Millipore) at 400 mA for 1 h at 4ºC in 1x transfer buffer (25 mMTris-HCl pH 8.3, 200 mM glycine, 20% (v/v) methanol) and protein transfer was checked by staining with Ponceau S (Sigma). Transferred membranes were blocked 1 h or o/n at RT or 4ºC respectively in 5% (w/v) skimmed milk in TBS-T (10 mM Tris-HCl, pH 7.5, 100 mM NaCl, and 0.1% Tween-20). Then the membrane was incubated for at least 1 hour (time is depending on the antibody) with the corresponding primary antibody diluted in 5% skimmed milk in TBS-T (see **Table EP1** for antibody working dilutions). After three washes of 10 min with TBS-T, membranes were incubated for 45 min at RT with a rabbit anti-mouse, goat anti-rabbit or rabbit anti-goat IgGs conjugated to horseradish peroxidase (1:5000; Dako) diluted in 0.5% skimmed milk in TBS-T. After three washes of 10 min with TBS-T, protein detection was by enhanced chemiluminiscence with ECL™ Western blotting detection (Amersham Life Sciences).

TABLE EP1. ANTIBODIES LIST

WESTERN BLOT ANTIBODIES				
ANTIBODY	**TARGET PROTEIN**	**HOST SPECIE**	**WORKING DILUTION**	**SOURCE**
HA	HA epitope tag	Mouse	1/500	Santa Cruz Biotechnology (sc-57592)
MYC	Myc epitope tag and mouse and human c-Myc	Mouse	1/500	Santa Cruz Biotechnology (SC-40)
Max	Max	Mouse	1/5000	Santa Cruz Biotechnology (sc-8011)
Glutathione S-transferase	GST	Goat	1/10000	Amersham Bioscience (27457701V)
FLAG	Flag epitope tag	Mouse	1/10000	SIGMA (F1804)
Gal4DBD	Gal4 protein DNA binding domain	Rabbit	1/500	Santa Cruz Biotechnology (sc-729)
Hhex	Hhex	Rabbit	1/2500	Katherine Borden´s Laboratory
Cyclin D1	Cyclin D1	Rabbit	1/1000	Cell Signalling (#2922)
Actin	Actin	Goat	1/500	Santa Cruz Biotechnology (sc-1616)
CHROMATIN IMMUNOPRECIPITATION ANTIBODIES				
c-Myc	Myc epitope and human c-Myc	Mouse	**Concentration** 1μg/ml	Millipore (4A6 clone,05-724)
HisG	HisG epitope	Mouse	**Concentration** 1'13 μg/ml	Invitrogen (46-1008)

EP4.4. GST-Hhex fusion protein expression in bacteria

GST and GST fusion proteins were isolated from the E.coli strain BL21 (DE3) after induction with 0.3 mM IPTG followed by o/n incubation at 17ºC. The cells were harvested and sonicated in cold STE buffer (10 mM Tris-HCl pH8, 150 mM NaCl, 1 mM EDTA) containing Halt Protease Inhibitors Cocktail (Pierce), 100 μg/ml lysozyme, 5mM DTT and 1,5% sarkosyl. The lysate was clarified by centrifugation and the supernatant was aliquoted and stored at -80ºC.

EP4.5. Protein-Protein interaction techniques

Yeast two-hybrid screening

We used an E9.5–10.5 dpc mouse expression library, kindly provided by Dr. Weintraub (Howard Hughes Medical Institute, USA) and Dr. Scambler (Institute of Child Health, UK). In this library, cDNAs are fused to a sequence encoding the activation domain of the herpes simplex virus protein VP16 [231]. Different Hhex fragments and the complete open reading frame were pre-selected as baits by PCR cloning into the EcoRI-SalI site in pBD-Gal4 Cam vector (Trp1 gene; Stratagene). Bait plasmids were transformed into AH109 strain (Ade–, His–, Leu–, Trp–; Stratagene) following the manufacturer instructions. Gal4-Hhex fusion proteins were detected by Immunoblotting (**figure R1.1**). pBD-Gal4-Hhex(1-137) was selected as a bait given the extremely low autoactivating capabilities and the biochemical relevance of the N-terminal domain (see **introduction section I2.1**). For the yeast two-hybrid library, 145 μg of the library (aprox. 106 transformants) was transformed into AH109 strain expressing Gal4-Hhex(1-137) and plated onto a SD/-Leu/-Trp/-His/-Ade 9 cm plate and grown at 30ºC for 7 days. The plates were daily checked for

growing colonies. Single growing colonies were transferred to SD/-Leu/-Trp/-His/-Ade α-X-Gal to reconfirm the interaction. Confirmed positive clones (blue colonies) were stocked in glycerol media at -80ºC. In parallel, all the clones appeared during the first four days of selection (a total of 522) were amplified by PCR (colony touching) using primers flanking the insert VP16-FP and VP16-RP (see **table EP2**) and the Biomek FX robot (Genomics services, UPF). After denaturing for 2 min at 94ºC, amplification was performed by 38 cycles of 30 s at 94ºC, 30 s at 50ºC and 5 min at 74ºC and a final extension of 5 min at 72ºC. Finally, the amplified products were sequenced with the BigDye Terminatorv3.1 cycle sequencing kit (Applied Biosystems) and the VP16-F primer. Translated sequences were blasted against the protein database. Selected preys were then isolated from yeast using the Wizard® Plus SV Minipreps DNA Purification System (Promega) after four rounds of vortex with glass beads to destroy the yeast cell wall. Isolated plasmids were transformed into Escherichia coli, amplified and re-transformed into AH109 yeast expressing the bait plasmids and empty pBD-Gal4 Cam and seeded onto SD/-Leu/-Trp/-His/-Ade α-X-Gal plates to reconfirm the interaction. β-Gal activity was also measured using the Beta-Glo Assay System (Promega).

GST pulldown

Volumes (15–50 µl) corresponding to approximately even amounts of individual GST fusion proteins were coupled to washed glutathione sepharose beads (40 µl suspension), washed twice with PBS and once with IP buffer (20 mM Tris-HCl pH8, 200 mM NaCl, 1 mM EDTA and 0,5% NP40) both buffers containing Halt Protease Inhibitors Cocktail (Pierce).

When *in vitro* translated protein was used, glutathione-coupled beads were incubated with 8 µl ^{35}S-methionine labelled *in vitro* translated protein (Promega's TNT® T7 Quick Coupled Transcription/Translation System) in IP-buffer containing Halt Protease Inhibitors Cocktail and then the beads were incubated o/n at 4ºC in a rocking platform. The day after the beads were washed six times with IP-buffer, resuspended in 40 µl of electrophoresis sample buffer and heated (95 ºC) for 5 min. Sample was run in a SDS polyacrylamide gel electrophoresis and signal intensities evaluated using a phosphorimager. When protein extracts were pulled down, glutathione-coupled beads were incubated with 500 µg of total protein lysate in M-Per buffer (Pierce) from cell cultures overexpressing the proteins of interest. The mixture was incubated o/n at 4ºC in a rocking platform. Then, the beads were washed six times with IP-buffer, resuspended in 40 µl of electrophoresis sample buffer and heated (95 ºC) for 5 min. Following SDS polyacrylamide gel electrophoresis, the gel was checked by immunoblotting.

Immunoprecipitation

For immunoprecipitation between Hhex and SOX13 approximately, 0.5–1.0 mg of total 293T protein extract expressing HexHA and SOX13-604-Flag was incubated with anti-HA antibody bound to agarose beads (Sigma Aldrich, E6779) overnight at 4 °C. After five washes in IP buffer, immunoprecipitates were resolved by SDS-PAGE and transferred to Immobilon filters (Millipore). Filters were then immunoblotted using HRP-conjugated anti-FLAG or anti-HA antibodies.

In the case of the immunoprecipation between Hhex and c-Myc, HeLa protein extracts expressing Hex(1-271)-HA or Hex(138-271)-HA and

endogenous c-Myc were incubated with anti-myc antibody coupled to agarose beads (Sigma Aldrich, A7470) overnight at 4°C and the immunoprecipitation protocol was followed as above.

For the competive immunoprecipitation assay, 293T cell extracts expressing SOX13-FLAG and TCF1-Myc tagged proteins were incubated with increasing amounts of GST-Hhex or GST alone (40 μg and 160 μg). Samples were immunoprecipitated as above using anti-FLAG antibodies bound to agarose beads (SIGMA, A2220). Immunoprecipitates were resolved by SDS-PAGE and transferred to Immobilon filters (Millipore). Filters were then immunoblotted using mouse anti-FLAG or anti-myc antibodies and incubated with a True Blot HRP-conjugated antimouse IgG (eBioscience). Immunoreactive proteins were visualized using the ECL Detection Reagent System (Amersham).

Immunoprecipitation of endogenous Hhex and c-Myc in human liver was performed by incubating 1mg of human liver protein extract with protein G agarose beads (control condition) or anti-myc antibody coupled to agarose beads (c-Myc condition) overnight at 4°C. The day after, the beads were washed eight times with IP buffer, immunoprecipitates were eluted by adding 40 µl of electrophoresis sample buffer and heated (95 ºC) for 5 min. Following SDS polyacrylamide gel electrophoresis, the gel was checked by anti-Hhex immunoblotting.

EP5. FUNCTIONAL ASSAYS

EP5.1. Reporter assays

Approximately $3x10^4$ of HeLa, 293T, 3T3 or HepG2 cells per well were seeded in 24 well plates. 24 hours later, when the cells are at 50-80% of

confluency, a maximum of 1µg of total DNA per well including 100 ng of the reporter plasmid of interest, 15ng of pRL-SV40 (a plasmid expressing Renilla luciferase under the SV40 immediate early enhancer/promoter) to correct variations in transfection efficiency, and different amounts of the expression plasmids, were transfected using Fugene 6 or Fugene HD (Roche Diagnostics). Each transfection condition was tested by triplicate. Luciferase activities were assayed after 48 hours using Dual-Luciferase reporter reagents (Promega) in a Microplate Luminometer (Berthold Dectection Systems).

EP5.2. Wnt stimulation and chromatin Immunoprecipitation (ChIP) assay

Wnt signaling was activated in 293T cells by treatment with LiCl [232]. Cells were grown in 10 cm plates. One fifth of the cells were used to obtain RNA and protein extracts while the rest were used for chromatin immunoprecipitation. Chromatin fixation and immunoprecipitation were performed as previously described [233] with the following modifications. After sonication, the samples were centrifuged at full speed for 10 min at 4ºC.The resultant chromatin was diluted 1/10 in ChIP dilution buffer (1% SDS, 10% Triton X-100, 1.2 mM EDTA, 50mM TrisHCl pH8, 200 mM NaCl and protease inhibitors) and precleared for 1 hour at 4ºC with Salmon Sperm DNA/Protein G Agarose Slurry (Milipore, 16-201). Then, 50 µl of diluted chromatin was incubated with 4 µg of mouse monoclonal anti-myc tag antibody (Upstate technology, clon 4A6) or monoclonal mouse anti-HisG antibody (negative control; Invitrogen) overnight at 4ºC in a rocking platform. Immunoprecipitated chromatin was eluted from beads by the addition of 250 µl of 0.1 M sodium bicarbonate, 1% SDS and incubation for 15 minutes by shaking at 42ºC.

DNA was recovered using Qiaquick PCR purification kit (Qiagen). Precipitated DNA was analysed using the LightCycler 480 Real-Time PCR system [234]. ChIP primers for Wnt target promoters are shown in **table EP2**.

EP5.3. BrdU proliferation assay

The day before transfection, 0.5 x 10^4 FSK cells per well were plated on 96 well plates. The cells were then infected with equivalent amounts of the purified AdHhex and control adenoviruses. Infected FSK cells were maintained in culture for 72 hours. Each infection condition was assayed by sixtiplicate. Cell proliferation rate was then assayed using the Cell Proliferation ELISA, BrdU (colorimetric) kit (Roche Applied Science) following manufacturer instructions.

EP5.4. Cell proliferation assay using Scan^R technology

10 cm/plates of TGR1 and H015.19 cells were infected with 3 ml of HhexGFP or GFP retroviral supernatants mixed with Polybrene (4 μg/ml) to increase infection efficiency. After 5 hours of incubation at 37ºC, 7 ml of fresh media was added to each plate. The infection procedure was repeated once 24 h later. Correct infection was confirmed by checking GFP fluorescence in a microscope. Infected TGR1 and H015.19 cells (TGR1-RetGFP; TGR1-RetHhexGFP; H015.19-RetGFP; H015.19-RetHhexGFP) were then splited in two 24 well/plates adding 10^5 cells/well and incubated at 37ºC. After 24h-48h green-fluorescence positive cells were counted using Scan^R technology (Olympus). Changes in the number of GFP cells (infected cells) were represented as a relative percentage setting as 100% the GFP cells counted at 24h. Data is the mean of two independent experiments.

EP5.5. Colony formation assay

MEFs were infected with Myc, HhexGFP or GFP retroviral supernatants as done above. Soft agar plates were prepared as follows. 7ml of Agar Media (2xMEM, 10%FCS, 1xDMEM, P/S and 0.5% bacto agar), were added to 60mm plates. Once plates solidified, they were kept at 4ºC during at least 24 hours. Then, 10^4 infected cells of each condition were diluted in 1ml of DMEM complete media and kept at 37ºC. Plates were then taken out from 4ºC and cell suspension was mixed with 2ml of Agar Media. Carefully, 1.5ml of the mixture was added to the plates and they were put in the incubator at 37ºC with 5%CO_2. New media was added every 3-4 days. Colonies were counted after 4-5 weeks. Cell viability was checked by addition of Tripan Blue. Each infection condition was assayed by triplicate. Data is the mean of two independent experiments.

EP5.6. *In vivo* electroporation of mouse embryos

All the manipulations were performed in the laboratory of Frederic Lemaigre PhD, at Univertisité Catholique de Louvain (Brussels, Belgium).

Ventral definitive endoderm was electroporated as described in Pierreux et al, [235]. Briefly, 8.5 days postcoitum (E8.5) mouse embryos were obtained from pregnant CD1 mice and placed in Hank's balanced salt solution (HBSS). Dissected embryos were transferred to a 30 µL drop of plasmid solution (concentration from 0.5 to 1.5 µg/µL) for 10 min at room temperature and then transferred with the DNA solution in the electroporation cuvette between two platinum plate electrodes placed 5 mm apart. Embryos were oriented in the cuvette so that the ectoplacental cone was tilted at an angle of approximately 45 degrees to the anode. Embryos were viable and developed correctly after three 50

ms pulses of 9 V at 1 s intervals. After electroporation, embryos were cultured for 24 h at 37°C in a roller culture system. Somites were counted and only embryos showing blood circulation and turning of the embryos to the foetal position were selected. The ventral midgut (liver bud region) was dissected and the luciferase activity was measured with the Dual-Luciferase Reporter (Promega).

EP5.7. Hhex null mouse embryo studies

All the manipulations were performed in the laboratory of Ken Zaret PhD, at the Fox Chase Cancer Center (Philadelphia, USA).

The generation of Hhex null mice in which LacZ gene and Neo resistance are in frame with the endogenous Hhex locus was described previously [88]. To generate $Hex^{LacZ/LacZ}$ null embryos (also termed $Hhex^{-/-}$), Hex^{LacZ} mice were mated and noon of the day of the appearance of a vaginal plug was considered as E0.5. To get E9.5 mouse embryos, pregnant mice were dissected after nine days, as described previously [236]. Embryos were genotyped as follows. Briefly, small piece of embryo was incubated with 250 μl of Lysys Bufer (100mM Tris/HCl pH 8.5, 5mM EDTA, 0.2% SDS, 200 mM NaCl) with fresh proteinase K (600 μg/ml) for 2-12 hours at 55ºC. After spin the samples during 10 minutes, supernatant was transferred into a new tube containing 250 μl of isopropanol. Precipitated genomic DNA was removed with a tip and resuspended with 100 μl of TE buffer. 1 μl of a 1/10 dilution of the resuspended DNA was use to perform the genotyping PCR reaction. Samples were amplified for 40 cycles (94°C for 30 seconds, 58°C for 30 seconds and 72°C for 45 seconds) with the PCR primers were Hhex-FP and Hhex-RP for the Hex

wild type allele; Neo-FP and Neo RP for the Neo gene (see **table EP2** for primers).

TABLE EP2. PRIMER LIST

YEAST-TWO-HYBRID			
	Primer	**Target**	**Sequence (5'-3')**
	VP16-FP VP16-RP	Primers flanking library insert	tcgagtttgagcagatgtttaccgtt gtaaaacgcggccagt
SOX13 isoforms			
Nr.	**Primer**	**Target**	**Sequence (5'-3')**
1	SOX13-FP*	SOX13 and S-SOX13	tgattcagcagcagcataagat
2	SOX13-RP1*	S-SOX13	taatggatgagagtggctccag
3	SOX13-RP2	SOX13	cagcggatactccactggtt
4	LSOX13-FP	L-SOX13	cgaagagaagagcgatggga
5	LSOX13-RP		gtatctgccaagtgcacagc
6	SOX13-I-FP	Intron 1 in SOX13 locus	gatcccagcaacaatcctgt
7	SOX13-I-RP		cctgcccaagctttgtttag
8	SOX13-SNP-FP	SOX13 and S-SOX13. Conflict.	aacatgcctgtgatcgtcaa
9	SOX13-SNP-RP		cgagttgggacaaggtcttc
MOUSE C-MYC TARGET GENES			
	Primer	**Target**	**Sequence (5'-3')**
	mRpl6-FP mRpl6-RP	Ribosmal protein 6	ggtgaagcttcgaaaaatgc ctggagcgtagccttctcac
	mNcl-FP mNcl-RP	Nucleophosmin	cacctgtggtcttacggttg gagcagatcgctttccagac
	mHdgf-FP mHdgf-RP	Hepatome derived Growth Factor	gtgggagatcgagaacaacc gcactgcccttcttatcacc
	mMyc-FP mMyc-RP	c-Myc	agtgctgcatgaggagacac ggtttgcctcttctccacag
	mHhex-FP mHhex-RP	Hhex	actacacgcacgccctactc acttgaccgcctttccttt
	mCcnd1-FP mCcnd1-RP	Ccnd1 (Cyclin D1)	gcgtaccctgacaccaatct ctcttcgcacttctgctcct
	mFoxa2-FP mFoxa2-RP	Foxa2	tggtcactggggacaagggaa gcaacaacagcaatagagaac
	mGapdh-FP mGapdh- RP	Glyceraldehyde 3-phosphate dehydrogenase	gagtcaacggatttggtcgt ttgattttggagggatctcg
HUMAN C-MYC TARGET GENES			
	CDK4- FP CDK4-RP	CDK4 (Cyclin-dependent kinase 4)	tactgaggcgactggaggct agttcgggatgtggcacag
	CCND2-FP CCND2-RP	CCND2 (Cyclin D2)	attgaaccatttggg atggtggtgtctgcaatgaa
	CCND1FP CCND1-RP	CCND1 (Cyclin D1)	agcgctgtttttgtt tcatcctggcaatgt

	GADD45A-FP GADD45A-RP	Growth arrest and DNA damage inducible, alpha	cggacctgcactgcgt ataagttgacttaaggcaggatcctt
	p15-FP p15-RP	CDKN2B (Cyclin-dependent kinase inhibitor 2B ,p15)	taatgaagctgagcccaggtct caccgttggccgtaaacttaac
	RCL-FP RCL-RP	RCL	atcgtgtctcggctgcg catggatgagcctgtcacc
	HHEX-FP HHEX-RP	HHEX	cggacggtgaacgactaca tttgacctgtctctc gctga
	PBGD FP PBGD RP	Porphobilinogen deaminase (hydroxymethylbilane synthase)	cggaagaaaacagcccaaaga tgaagccaggaggaagcacagt
HUMAN WNT TARGET GENES			
	DKK1-FP DKK1-RP	Dickkof Homolog 1	tggaatatgtgtgtcttctgatcaaa aagacaaggtggttcttctggaat
	SP5-FP SPR-RP	SP5 transcription factor	actttgcgcagtaccagagc acgtcttcccgtacaccttg
ChIP PRIMERS			
	ChIP-DKK1- FP ChIP-DKK1-RP	Dickkof Homolog 1	attcaacccttactgccaggc aaggctaccagcgagcgttat
	ChIP-SP5-FP ChIP-SP5-RP	SP5 transcription factor	tccagaccaacaaacacacc gcttcaggatcacctccaag
	ChIP-RPL0-FP ChIP-RPL0-RP	Ribosomal protein large P0	acccagctctggagaagtca gaggtcctccttggtgaaca
GENOTYPING			
	Hhex-FP Hhex-RP	Mouse endogenous Hhex	aggccgagtgtgaaatcagag cagaagagctgtggttaaccaa
	Neo-FP Neo-RP	Neomicin cassete	ctgtgctcgacgttgtcactgaag tattcggcaagcaggcatcgcca

RESULTS

R1. INTERACTION BETWEEN HHEX AND SOX13 MODULATES WNT/TCF ACTIVITY

R1.1. Identification of novel Hhex partners by yeast two-hybrid

To search for Hhex interactors during embryonic development, we performed a yeast two-hybrid screening using an E9.5-E10.5 mouse embryo library (see **Experimental Procedures section EP4.5**) as prey and Hhex(1–137) as bait. After the sequential transformation of the AH109 yeast strain (**figure R1.1 C**), we retrieved 1182 clones that grew in stringent media SD/-Leu/-Trp/-Ade/-His. Only 1014 grew and were α-galactosidase-positive when re-plated in SD/-Leu/-Trp/-Ade/-His/ + α-X-gal. A first round of sequencing (522 isolated clones) provided 294 in-frame coding sequences with repetitions ranging from 1 to 56 times. The sequences with higher number of repetitions are summarised in **Table R1**.

Several new putative Hhex-binding partners were identified in this screening. Paternally expressed gene 10 (Peg10), for instance, is a gene that contains two overlapping open reading frames (ORF1 and ORF2) derived from a retroelement that acquired a cellular function. It is paternally imprinted, and it is highly expressed in hepatocellular carcinoma, being interesting for further studies as a Hhex interactor [237]. Other intriguing partner is the enzyme Pyruvate Kinase isoform M2 (M2PK). It is physiological expressed as tetramer, but its interaction with certain oncoproteins switches it to a tumoural dimeric form, used as marker in several types of cancer [238-239] . On the other hand, TLE-related proteins (AES-1) and members of the Eif4 family (Eif4g1 and Eif4g3) were also present in the list, in agreement with previous reports [56, 63].

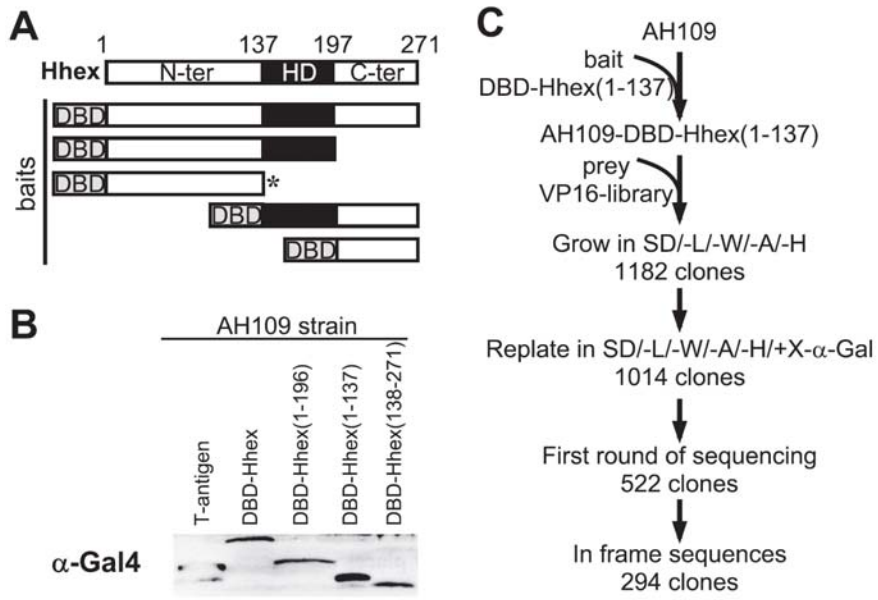


FIGURE R1.1. Hhex yeast two-hybrid screening.
A) Diagram showing Hhex fusion proteins containing the Gal4 DNA binding domain (DBD) in the N terminus. These proteins were tested as putative baits in a yeast two-hybrid screening. An asterisk indicates the protein used as the final bait in the screening. The Hhex protein showing functional domains is also shown above as a reference. N-ter, N-terminal domain; HD, homeodomain; C-ter, C-terminal domain**. B)** Immunoblotting using an anti-Gal4 antibody of extracts from the AH109 yeast transformed with the different baits. The yeast-expressing plasmid, pGBKT7-T, expressing the SV40 large antigen T protein was used as a positive control. **C)** An overview of the Hhex yeast two-hybrid screening by sequential transformation of bait and prey (VP16 library). VP16 library, E9.5-E10.5 mouse embryo library made in the pVP16; SD/-L/-W/-A/-H, synthetic dropout medium lacking leucine, tryptophan, adenine, and histidine; SD/-L/-W/-A/-H/ X-α-gal, synthetic dropout medium lacking leucine, tryptophan, adenine, and histidine, and containing 20 μg/ml of X-α-gal.

GENES	REPEATED	DOMAINS
Melanoma-associated D1 MAGED1	56	aa337-aa377
Peg10, ORF1/2	55	aa850-aa950
M2 Pyruvate Kinase (M2PK)	13	aa377-aa531
SOX13	8	aa119-aa218
Acyl-CoA binding domain containing 3	7	aa230-aa301
Eif4g3 protein	5	aa156-aa329
AES-1; TLE_N domain	5	aa29-aa152
Type IV collagen alpha-1	4	aa78-aa172
Anaphase promoting complex subunit 5	3	aa215-aa362

TABLE R1. List of the most repeated cDNAs recovered in yeast two-hybrid screening. Summary of the putative Hhex partners found in the yeast-two-hybrid screening with Hhex (1-137) as a bait and a E9.5-E10.5 mouse embryo library as a prey. Only coding and in frame cDNAs are shown. The domain corresponds to the region of the protein encoded in the prey plasmid.

R1.2. SOX13 interacts with Hhex

Among the different novel sequences, SOX13 was selected based on the high number of clones detected during the screening, expression pattern and the extensive embryological function of SOX family members. SOX proteins can activate or repress genes through interaction with different partner proteins [115]. Eight clones of SOX13 were isolated (**figure R1.2 A**), all of which contained the region aa126-aa218 comprising the leucine zipper (LZ) and the glutamine-rich (Q-rich) domain. To confirm the interaction and to exclude the auto-activation of the isolated clones, a representative prey plasmid (clone 11.23) was purified from the yeast, amplified in bacteria, and re-transformed in the AH109 yeast strain expressing each pre-selected Hhex bait (see **figure R1.1 A**). Only those transformants expressing the bait containing the N-terminal domain of Hhex (aa 1–137) were able to grow and stain positive in SD/- Leu/-Trp /

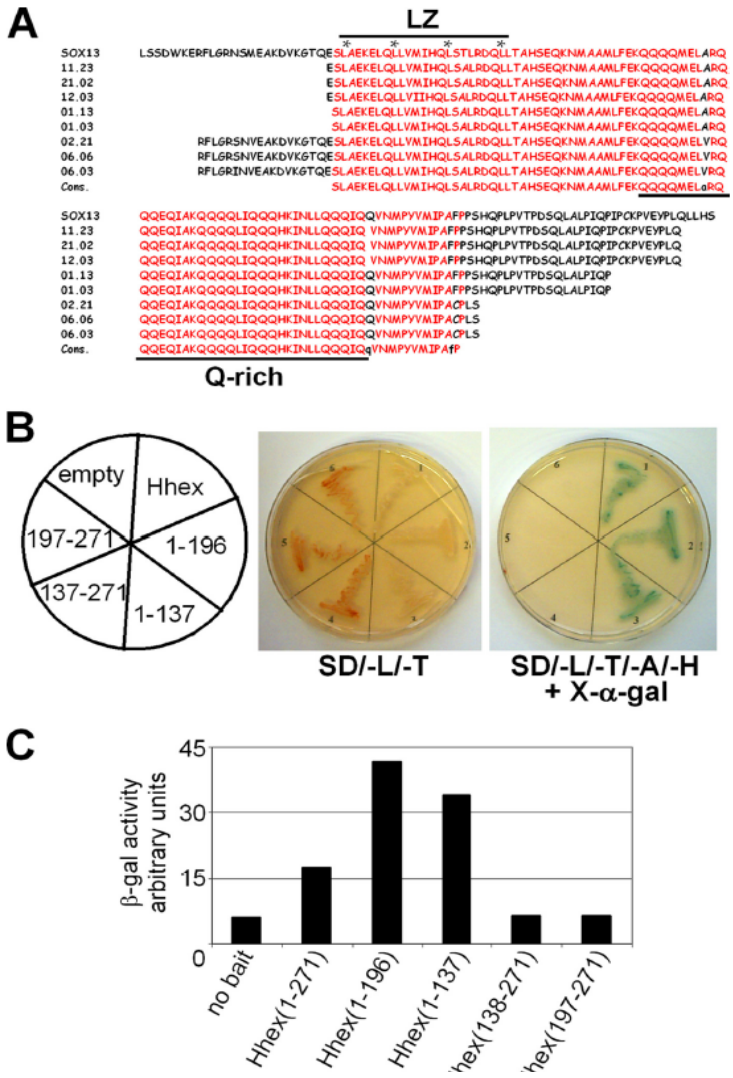


FIGURE R1.2. Interaction between Hhex and SOX13 in yeast.

A) Multiple alignments between mouse Sox13 and the prey clones isolated in the yeast two-hybrid screening. LZ, leucine zipper; Q-rich, glutamine rich. Asterisks indicate the position of leucines within the LZ domain. Consensus sequences are in red. **B)** A representative Sox13 prey plasmid (clone 11.23) was re-transformed into AH109 yeast expressing Hhex baits or empty plasmid as a negative control. The transformed yeast were plated in SD/-Leu/-Trp and SD/-Leu/-Trp/-Ade/-His/ X-α-gal. **C)** β-galactosidase activity was also measured in yeast cotransformed with each Hhex bait and Sox13 prey plasmid (clone 11.23).

-Ade/-His + α- X-gal media (**figure R1.2 B**).

Finally, we quantified the interaction by measuring β-galactosidase activity by luminescence. Once again, only the bait containing the N-terminal domain of Hhex gave β-galactosidase activity above the background level (**figure R1.2 C**). In conclusion, SOX13 (aa 126–218) interacts with Hhex in yeast.

R1.3. A 1815-bp open reading frame of SOX13 (604 aa) is prevalent in human tissues

SOX13 is a member of the group D of the SOX transcription factors that contains the characteristic high mobility group (HMG) domain together with a LZ and Q-rich domain. The SOX13 genomic locus contains 14 exons, and three different isoforms have been reported in the literature (**figure R1.3 A**). SOX13 [121] is an open reading frame of 1815 bp encoding a 604 aa protein containing the HMG domain and the LZ and Q-rich domain (**figure R1.3 B**). Long (L)-SOX13 [126] is an open reading frame of 2673 bp encoding an 890 aa protein as a result of a missing G nucleotide at base pair 1950 in the sequence reported by Roose et al. [126] (GenBank™ accession number AF083105). This putative deletion leads to a shift of exon 14 open reading frame, resulting in a longer protein. Finally, short (S)-SOX13 (ECRBrowser [127]) is an open reading frame of 768 bp encoding a 255 aa protein caused by alternative splicing that results in a truncated isoform lacking the C-terminus including the HMG domain. The three isoforms contain the putative SOX13 interaction domain detected in the yeast two-hybrid screening (LZ-Q-rich domain, aa126–218; **figure R1.3 B**).

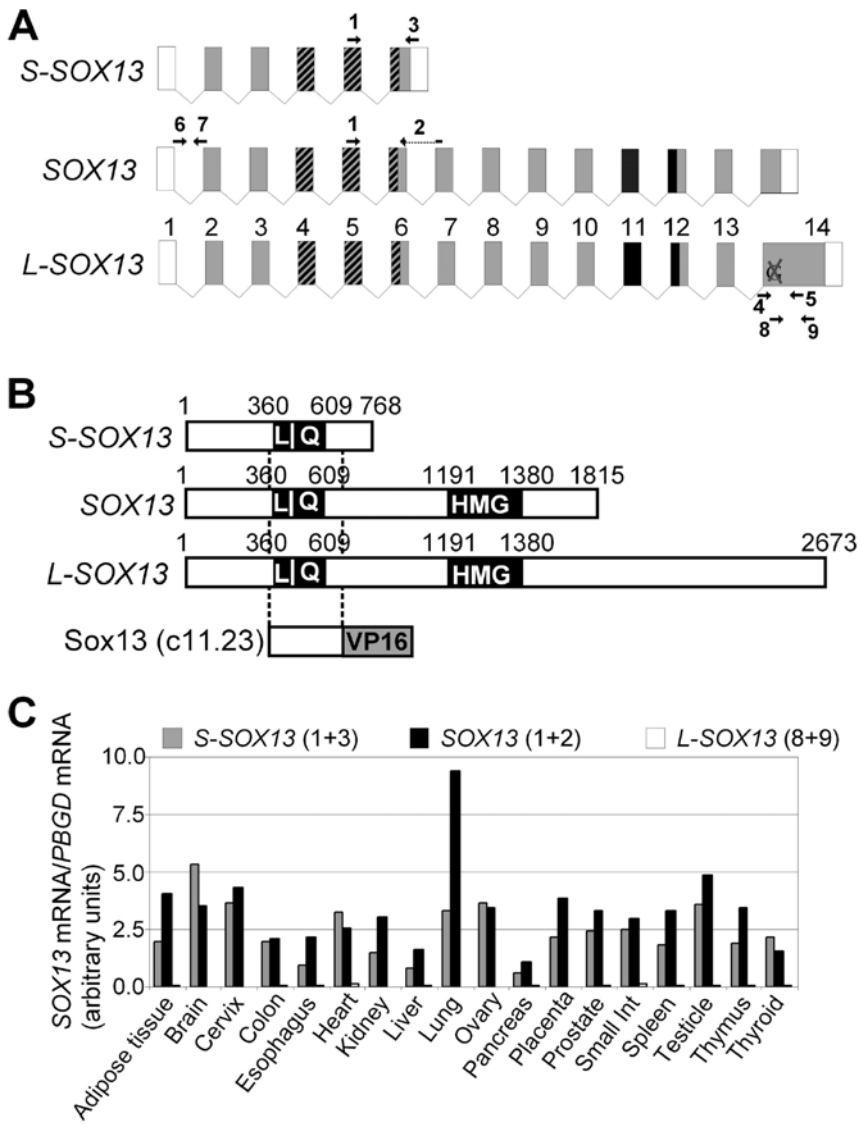


FIGURE R1.3. SOX13 isoforms present in human tissues.

A) Graphic representation of SOX13 exons as described in ECRBrowser. Exons are depicted as boxes. Exon size is not proportional to box size. Gray boxes represent translated exons. The HMG box is represented in black (exons 11–12). Dashed boxes represent the LZ-Q domain (exons 4–6). The conflictive G1950 in exon 14 is also shown. The primers used for qRT-PCR analysis are depicted with a number (1–9). Primer sequences are shown in Experimental Procedures. **B)** Alignment of the SOX13 coding sequences are compared with the prey clone 11.23 isolated from the yeast two-hybrid screening. This fragment contains the LZ-Q domain of SOX13. Numbers show base pair positions. **C)** Relative *SOX13* mRNA levels in 18 human tissues were determined by qRT-PCR. The numbers in parentheses represent the primers used for the qPCR reaction (see Experimental Procedures **table EP2**). PBGD, porphobilinogen deaminase.

In preparation for mapping the interaction domain of SOX13 with Hhex, we decided to determine which SOX13 isoform prevails in human tissues. We designed three primer pairs capable of binding specifically to each coding sequence (see **table EP2 in experimental procedures** section). These primers were validated by PCR using plasmids containing the three different cDNAs. PCR fragments representative of each PCR reaction were purified and sequenced to confirm identity. The mRNA of the SOX13 isoforms was quantified by qRT-PCR in 18 different tissues from a mixture of at least 3 different human donors (**figure R1.3 C**). Only *SOX13* and *S-SOX13* were significantly detected in all the tissues tested. *SOX13* predominated in most of the tissues, although *S-SOX13* mRNA was more abundant in the brain, heart, and thyroid. To confirm the absence of *L-SOX13*, we designed a pair of primers (**numbers 4 and 5; table EP2**) to amplify the conflicting region in exon 14. The sequencing of the amplified by region revealed no deletion at nucleotide G1950. Based on these results, *SOX13* and *S-SOX13* coding sequences were amplified by PCR from human cDNA and cloned into the pFLAG-CMV2 vector. We conclude that isoforms *SOX13* and *S-SOX13* are prevalently expressed in human tissues and that both contain the LZ and Q-rich domain.

R1.4. Mapping Hhex and SOX13 interaction domains

We investigated which Hhex domains are required for the interaction with SOX13 and vice versa by pulldown assay using GST-Hhex fusion constructs purified from bacteria (see scheme in **figure R1.4 B**). We also obtained total protein extracts from HeLa cells ectopically expressing FLAG-tagged full-length SOX13 (604 aa), S-SOX13 (255 aa),

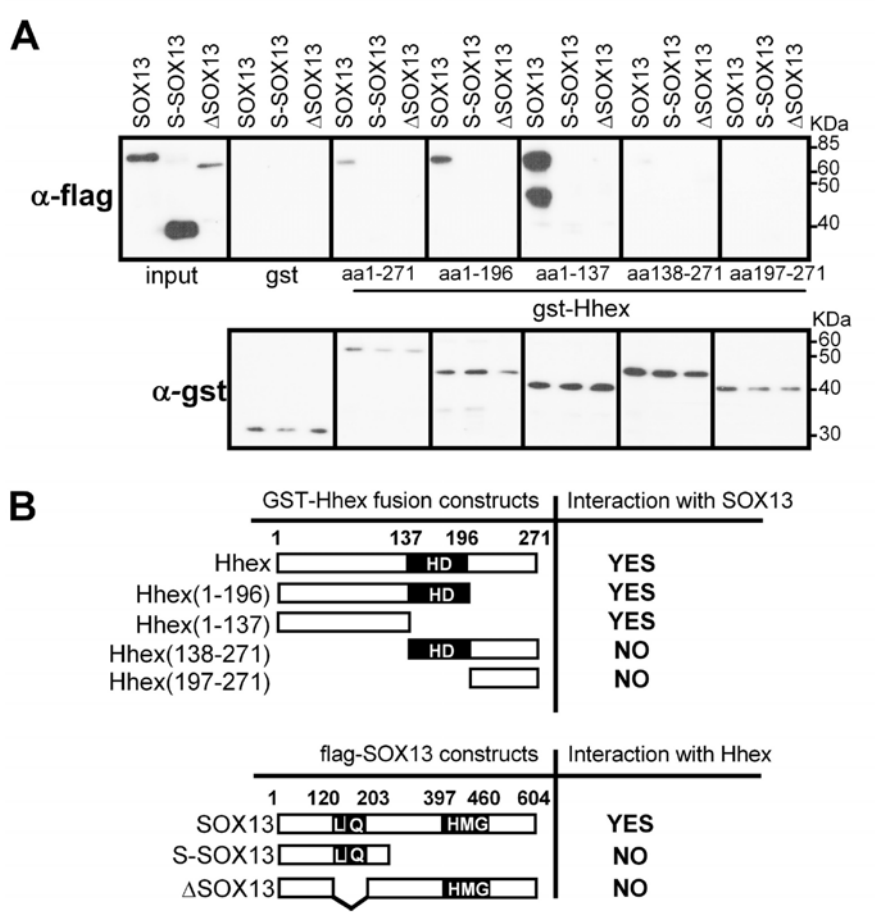


FIGURE R1.4. Hhex and SOX13 interaction domains.

A) *In vitro* binding assay between purified GST, GST-Hhex(1–271), GST-Hhex(1–196),GST-Hhex(1–137), GST-Hhex(138–271) or GST-Hhex(197–271) and HeLa cell extract expressing FLAG-tagged *SOX13*, S-*SOX13*, or Δ*SOX13*. Glutathione-sepharose was used to pulldown the GST-Hhex fusion proteins and immunoblotted using anti-FLAG antibody. One-twentieth of the extract was loaded as the input. The lower band in the SOX13 incubation was routinely observed after overnight incubation and is a C-terminal degradation product of SOX13. Lower panel is an immunoblot using anti-GST antibodies in 1% of the beads. **B)** Diagrams summarizing the results obtained in the GST pulldown assay. Numbers show amino acid positions. *HD,* homeodomain; *L*, leucine zipper; *Q*, glutamine-rich; *HMG*, high mobility group domain.

and ΔSOX13, a mutant SOX13 lacking the LZ and Q-rich domain (lacking aa 126–218, the interaction domain). As **figure R1.4** illustrates, GST-Hhex and all the fragments containing the N-terminal domain (aa 1–137) were able to co-precipitate full-length SOX13. When only the N-terminal domain of Hhex was expressed as a single fragment, it was capable of pulling down full-length SOX13 more efficiently. This result clearly establishes the N-terminus of Hhex as the binding interface with SOX13 and confirms the result obtained in the yeast two-hybrid assay (**figure R1.2**). On one hand, ΔSOX13, which specifically lacks the LZ-Q domain of SOX13, was unable to bind Hhex or any of its fragments. This result is also in agreement with the yeast two-hybrid assay. On the other hand, we found that the N-terminus of SOX13 alone (the S-SOX13 isoform), which contains the LZ-Q domain, did not interact with Hhex or any of its deletion mutants. Thus, the N-terminus, or more specifically the LZ and Q-rich domain, is essential for SOX13 interaction with Hhex, but the domain alone is not sufficient for the interaction in mammalian cells. We conclude that the N-terminus of Hhex is essential for SOX13 interaction, whereas the LZ and Q-rich domain of SOX13 is necessary, but not sufficient, for Hhex interaction.

R1.5. SOX13 directly interacts with Hhex

The results illustrated in **figure R1.4** are compatible with the existence of a direct association between SOX13 and Hhex. To test this possibility we performed a GST pulldown assay using GST-Hhex immobilized on glutathione agarose beads and *in vitro* translated ^{35}S-SOX13. The GST protein alone did not interact with SOX13 (**figure R1.5 A**). However, GST-Hhex was able to interact with full-length SOX13 and L-SOX5 (another member of the group D family of SOX; data not shown). Neither GST nor

GST-Hhex pulled down the unrelated protein Lamin C. These results show that a direct interaction takes place between SOX13 and Hhex. To further corroborate the interaction between Hhex and SOX13, we performed immunoprecipitation experiments with HeLa cells ectopically expressing both transcription factors. Full-length SOX13-FLAG co-immunoprecipitated with full-length Hhex-HA (**figure R1.5 B**). No co-precipitation was detected in either mock-transfected HeLa cells or cells transfected with a plasmid expressing SOX9-FLAG (data not shown). Endogenous Hhex was also immunoprecipitated with SOX13 after transfection of HepG2 cells with a plasmid expressing SOX13-FLAG (**figure R1.5 C**). Co-localization studies in transfected HeLa cells showed nuclear localization of the proteins and are in agreement with both co-immunoprecipitation and the pulldown analysis (**figure R1.5 D**).

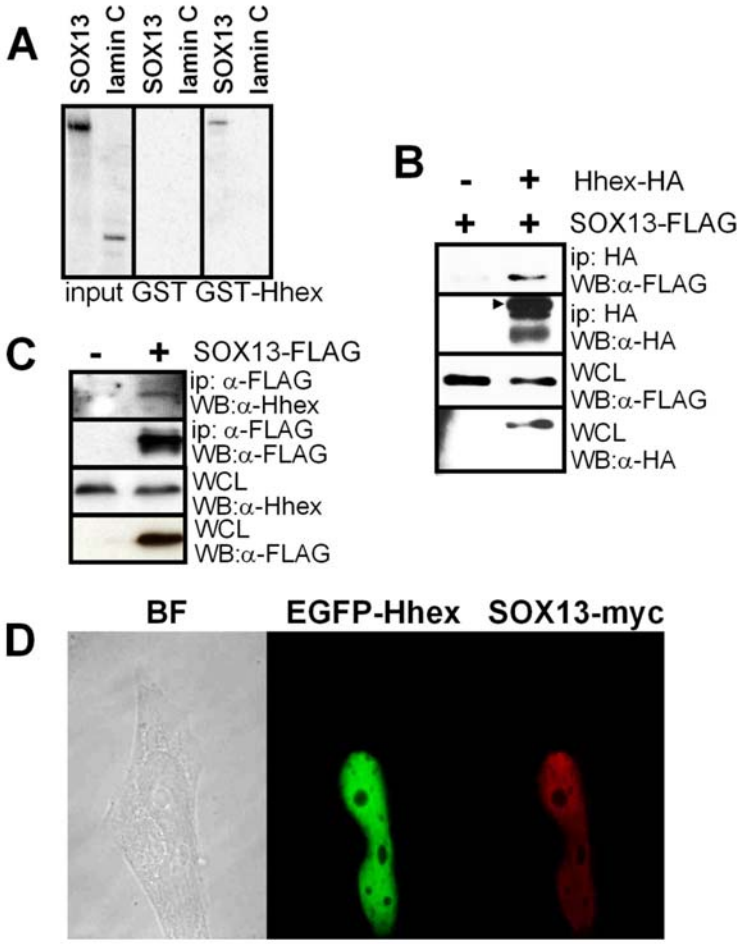


FIGURE R1.5. Interaction nature and co-localization of Hhex and SOX13.

A) *In vitro* translated ^{35}*S-SOX13* and ^{35}*S-lamin C* (negative control) were incubated with glutathione-sepharose beads coated with GST-Hhex or GST. Bound proteins were run in SDS PAGE and visualized using a Phosphorimager**. B)** Co-immunoprecipitation of SOX13 and Hhex in HeLa cell lysates transfected with SOX13-FLAG and Hhex-HA after immunoprecipitation (ip) with anti-HA antibody and immunoblotting against the FLAG tag. The whole cell lysate (WCL) was immunoblotted using anti-FLAG or anti-HA to show the expression levels of SOX13-FLAG and Hhex-HA in the assay. **C)** Immunoblot analysis of HepG2 cell lysates transfected with SOX13-FLAG detects endogenous Hhex after immunoprecipitation with anti-FLAG antibody and immunoblotting against a primary antibody against Hhex. The whole cell lysate was immunoblotted using anti-FLAG or anti-Hhex to show the expression levels of SOX13-FLAG and Hhex in the assay. **D)** Co-localization of ectopically expressed Hhex and SOX13 is shown. HeLa cells were co-transfected with plasmids expressing the fusion protein EGFP-Hhex (green) and a SOX13-myc construct and labelled with anti-Myc antibody (red). Images were acquired with a Leica TCS-SP2 confocal microscope. BF, bright field.

R1.6. Hhex modulates SOX13-dependent Wnt signalling in 293T cells

Given that Hhex can act as a transcriptional repressor or activator, we tested the possibility that SOX13 could be a co-repressor or co-activator of Hhex. To look for SOX13 co-repressing activity, we used the plasmid Gsc-LUC, a Hhex-specific reporter system based on the Goosecoid promoter [70]. SOX13 did not significantly change the transcriptional activity of Hhex (data not shown). Similarly, we tested the possible role of SOX13 as a transcriptional co-activator of SOX reporter 3xSx [240]. SOX13 did not transactivated the 3xSx luciferase reporter, although other SOX protein, SOX9, effectively activated this reporter more than 30-fold (data not shown). The addition of Hhex did not change SOX13 response (data not shown). In short, neither Hhex nor SOX13 altered the transcriptional capabilities of each other in specific luciferase reporter plasmids.

SOX13 was recently described as a key regulator of T-lymphocyte differentiation by the inhibition of Wnt signalling [135]. To test whether Hhex could modulate the function of SOX13 in Wnt signalling, we used a luciferase reporter controlled by multiple wild type (TOPflash) or mutated (FOPflash) TCF binding sites [241]. Wnt pathway was constitutively activated by co-transfecting an expression vector containing activated β-catenin-S37Y. We used 293T cells because they express TCF1 and are fully competent to activate transcription of a luciferase reporter gene controlled by TCF binding sites [241]. Full-length Hhex caused a 2-fold induction of the TCF/β-cat luciferase reporter, whereas Hhex(1–196) had no effect (**figure R1.6 A**, **compare group 1 and 2**), suggesting that the activating domain of Hhex located in the C-

terminus is necessary for this induction. In fact, when a deletion mutant of Hhex containing only the activating domain and the homeodomain was used Hhex(138–271), Wnt reporter activity increased by 4-fold. Full-length SOX13 caused a dose-dependent inhibition of the Wnt reporter in 293T cells (data not shown). A concentration of SOX13 causing an 80% reduction of TOPflash reporter activity was selected for further transfection studies (**group 3-6**). The same concentration of ΔSOX13 and S-SOX13 did not change significantly Wnt reporter activity (**figure R1.6 B**), as well as SOX9 protein that belongs to other subfamily, suggesting that the LZ-Q domain is necessary, but not sufficient, for modulation of Wnt activity. The addition of Hhex to SOX13 abrogated the repressive effect of the latter on Wnt activity in a dose-dependent manner (**figure R1.6 A**, **black bars, compare group 4–6**). We speculated that the resulting TOPflash activity may be either the simple addition of the repressing and activating features of SOX13 and Hhex, respectively, or the result of the interference caused by Hhex in the SOX13·TCF1 complex on the TOPflash reporter. To explore these hypotheses and establish the direction of this functional interaction (does Hhex rescue SOX13 inhibition or does SOX13 block Hhex induction?), we used Hhex(1–196), a fragment of Hhex that interacts with SOX13 (**figure R1.4**), but it was unable to induce Wnt reporter activity on its own. Nevertheless, Hhex(1–196) was still able to restore TOPflash activity (**figure R1.6 A**). Conversely, Hhex(138–271), a fragment of Hhex that cannot interact with SOX13 (**figure R1.4**) but is an enhancer of TOPflash reporter activity on its own, only induced Wnt reporter activity at higher doses and significantly less efficiently than Hhex.

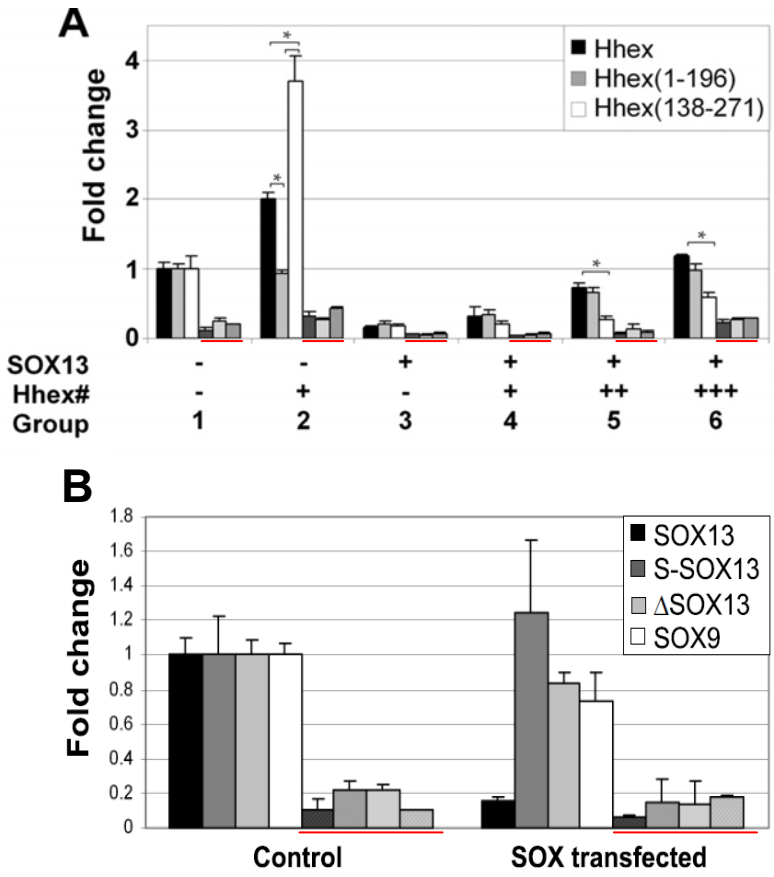


FIGURE R1.6. Effect of Hhex and SOX proteins on Wnt activity in 293T cells

HEK 293T cells were transfected with the Wnt reporter plasmid TOPflash (filled bars) and the mutated version, FOPflash (dashed bars and red line), as a negative control. An expression vector containing activated β-catenin-S37Y was cotransfected in each condition. Those plasmids expressing SOX proteins or Hhex derivatives were transfected as indicated below the graph. The means and S.E. of the relative firefly/renilla ratios and fold change over the control of at least three experiments are shown. The total firefly/renilla ratio of untreated cells transfected with activated β-catenin (S37Y) was set as 1.0. **A)** Hhex#: Hhex(1–271), Hhex(1–196), or Hhex(138 –271). **B)** Only SOX13 is able to inhibit Wnt activity. No significant repression was observed for S-SOX13, ΔSOX13 and SOX9. *, p <0.05, paired Student's t test.

Next, our aim was to definitively address whether the ability of Hhex to block the SOX13 repression of Wnt activity was directly dependent on SOX13 interaction or if it was indirect by co-repressing a factor that would ultimately alter Wnt reporter. Hhex is a transcriptional repressor that can be converted into an activator by fusing two copies of the minimal transcriptional activation domain of VP16 (Hhex-VP2 [70]; see schematic representation in **figure R1.7 A**). Retention of the entire Hhex open reading frame allows the examination of Hhex function without disrupting potential protein-protein

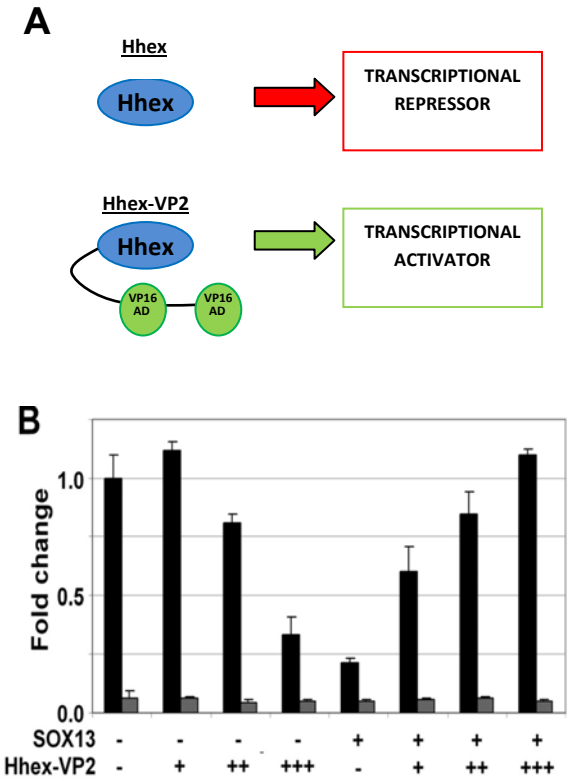


FIGURE R1.7. Effect of Hhex-VP2 on Wnt activity in 293T cells

A) Schematic representation of Hhex and Hhex-VP2 construct. The fusion of two copies of the minimal transcriptional activator domain VP16 to Hhex transforms Hhex into a transcriptional activator (*Brickman, J.M. et al 2000*) [70]. **B)** 293T cells were transfected with the Wnt reporter plasmid TOPflash (black bars) and the mutated version, FOPflash (grey bars) as a negative control. An expression vector containing activated β-catenin (S37Y) was always cotransfected in each condition. Plasmids expressing SOX13 and Hhex-VP2 proteins were transfected as indicated. Means and SEM of the relative firefly/renilla ratios and fold change over control of three experiments are shown. The total firefly/renilla ratio of untreated cells transfected with activated β-catenin (S37Y) was set as 1.0.

interactions. As previously reported [71] Hhex-VP2 alone caused a dose-dependent decrease in the activity of the TOPflash reporter, which is in agreement with the conversion of Hhex into an activator (**figure R1.7 B**). Paradoxically, when increasing amounts of Hhex-VP2 were added to SOX13, Wnt activity increased to basal levels. This indicates that abolishment of the inhibitory activity of SOX13 by Hhex does not reside in the repressing capabilities of the latter. Rather, it strongly suggests that Hhex antagonizes SOX13 activity on the Wnt signalling pathway by displacing and sequestering SOX13 and not indirectly repressing other genes ultimately involved or linked to Wnt activity.

Next, we conducted loss-of-function experiments in Hhex expressing HepG2 cells by means of an adenoviral vector expressing shRNA against Hhex. Knockdown of Hhex was validated by immunoblotting using specific Hhex primary antibodies (**figure R1.8 A**). Knockdown of Hhex resulted in inhibition of Wnt activity in agreement with the role of Hhex as a Wnt inducer (**figure R1.8 B group 1**). The addition of increasing doses of Hhex slowly increased Wnt activity in shRNA-treated HepG2 cells, although at a significantly lower rate than control HepG2 cells (**figure R1.8 B**, **group 3-5**).

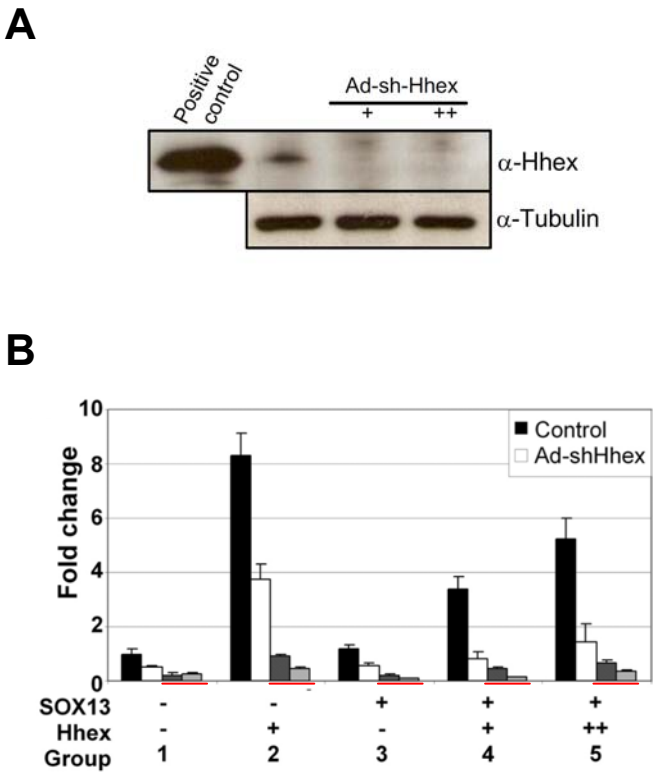


FIGURE R1.8. Endogenous Hhex in HepG2 participates in the induction of Wnt activity

A) HepG2 cells were infected with an adenoviral vector expressing a shRNA against Hhex. Total cell lysates were immunoblotted using anti-Hhex primary antibody. A clear reduction in HHEX protein content is observed. Positive control are cells transfected with a plasmid expressing Hhex. Tubulin immunoblot is shown as a loading control. **B)** HepG2 cells were transfected with the Wnt reporter plasmid TOPflash (black and hite bars) and the mutated version, FOPflash (grey bars) as a negative control. An expression vector containing activated β-catenin (S37Y) was always cotransfected for each condition. Endogenous Hhex was knocked down by adenovirus-mediated expression of shRNA (Ad-shHhex) (white bars). An adenovirus expressing shRNA against luciferase was used in HepG2 control cells. (black bars) A plasmid expressing Hhex is represented at the bottom of the panel as Hhex.

R1.7. Hhex competes with TCF1 to bind SOX13 in 293T cells

SOX13 is a potent repressor of the Wnt pathway by interacting and sequestering TCF1 from the Wnt transcriptionally active complex [135]. As shown above, the co-expression of Hhex together with SOX13 is able to restore Wnt activity to basal levels. One hypothesis to explain Hhex activity is that Hhex interacts with SOX13 to release TCF1 that is then restored into the Wnt activating complex. To test this hypothesis, we performed a competitive immunoprecipitation assay using purified GST-Hhex as a competitor. As reported, SOX13 binds full-length TCF-1 (**figure R1.9**, **upper panel, lanes 1 and 4**). The addition of Hhex was able to abolish the formation of the SOX13·TCF1 complex in a dose-dependent manner (**figure R1.9; lanes 2 and 3**). Concomitantly, the addition of Hhex restores the Wnt activity in transfected 293T cells (**figure R1.6 A**). Basically, Hhex is not only able to interact with SOX13 but is also capable of withdrawing SOX13 from the SOX13·TCF1 complex and releasing back TCF1 and restoring Wnt activity to basal levels.

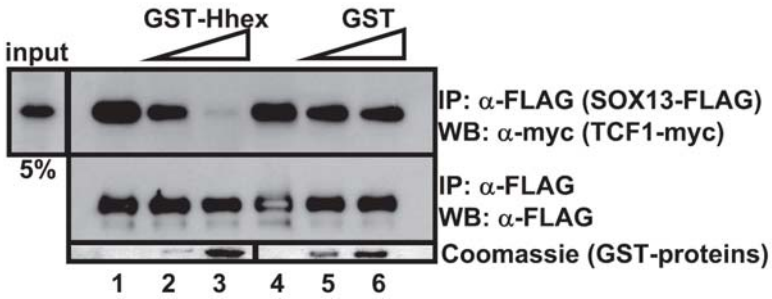


FIGURE R1.9. Hhex and SOX13 competitive immunoprecipitation assay.
A lysate from 293T cells co-transfected with SOX13-FLAG and TCF1-myc tagged-expressing plasmids was immunoprecipitated (IP) using anti-FLAG antibody in the absence of GST proteins (lanes 1 and 4) or with increasing amounts of purified GST-Hhex (lanes 2 and 3) or GST (lanes 5 and 6). The immunoprecipitated material was immunoblotted (WB) with anti-Myc antibody. The presence of increasing amounts of GST-proteins is shown with Coomassie staining of a gel run in parallel.

R1.8. Hhex influences SOX13-dependent Wnt activity in mouse embryos

To validate the results obtained in cell culture, we designed an *in vivo* reporter assay. We electroporated mouse embryos with reporter and expressing plasmids, cultured the embryos for 24 h, isolated the presumptive liver domain, and measured luciferase activity on the isolated embryonic tissue extract (**figure R1.10**). The method was developed to target the definitive ventral endoderm where the liver develops [235]. Once again, electroporation of SOX13 resulted in a severe (>80%) reduction in the Wnt reporter activity. Unlike our observation in 293T cells, Hhex electroporation also caused a profound inhibition of the Wnt pathway in mouse ventral targeted endoderm. The different behaviour of Hhex alone in this system and in cultured cells may result from tissue-specific differences in the expression of other Wnt

regulators. Co-electroporation of Hhex and SOX13 restored TOPflash reporter activity at the control levels, thus confirming our cell culture experiments. In short, *in vivo* experiments corroborated that Hhex can revert Wnt inhibition triggered by SOX13.

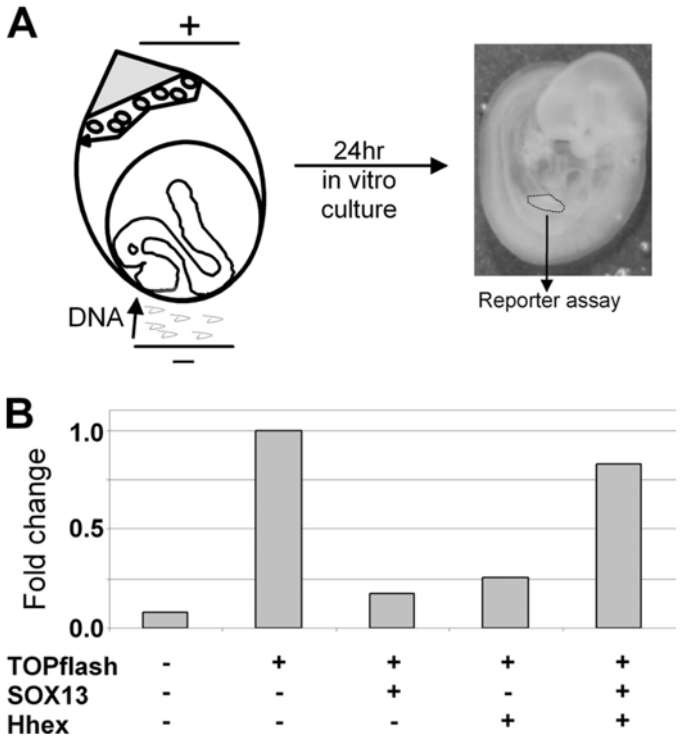


FIGURE R1.10. Wnt activity in mouse embryos electroporated with SOX13 and Hhex.

A) Diagram showing the orientation of the embryo in the electroporation cuvette. E8.5 mouse embryos were orientated with a 45° angle to specifically transfer DNA to the pre-hepatic and pre-pancreatic endoderm. After electroporation, embryos were incubated for 24 h *in vitro*. Turned-embryos, with the E9.5 developmental stage, were then dissected to isolate the liver bud region, and the reporter activity was measured. **B)** Luciferase activity is shown of isolated liver buds from non-electroporated (first chart) or electroporated embryos with TOPflash reporter, SOX13, and/or Hhex. The total firefly/renilla ratio of embryos electroporated with TOPflash alone was set at 1.0. Each condition was done at least three times, and the total number of embryos recovered and assayed was at least six.

R1.9. DKK1, a target of the Wnt/TCF pathway is regulated by Hhex·SOX13

We addressed the question of whether the modulation of the Wnt activity by the Hhex·SOX13 complex observed in reporter assay could also be detected in endogenous Wnt target. We selected DKK1, a well known Wnt target previously validated by chromatin immunoprecipitation and qRT-PCR in cultured cells [232]. mRNA levels of *DKK1* behaved in parallel to reporter activity (**compare figure R1.11B with figure R1.6 A, groups 1, 3 and 6**). After activation of Wnt signalling by LiCl, SOX13 caused a down-regulation of DKK1 mRNA. This repression was accompanied by a reduction of TCF1 bound to the DKK1 promoter (**figure R1.11 D**). Hhex coexpression increased the level of DKK1 mRNA, obliterating the inhibition caused by SOX13. Again, the derepression caused by Hhex was followed by a recovery in the amount of TCF1 bound to the DKK1 promoter. Similar results were obtained with another Wnt target, SP5 [230] (see **figure A3 in Annex**).

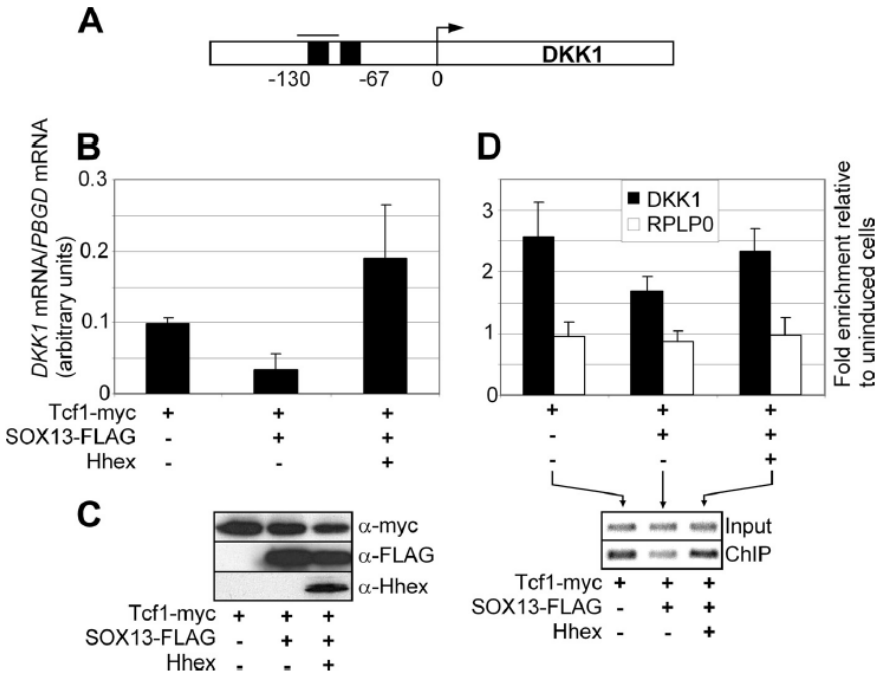


FIGURE R1.11. Regulation of DKK1, a target of the β-catenin-TCF pathway, by Hhex.
A) Schematic representation of the human DKK1 gene. TCF binding sites are depicted as *black boxes*. **B)** 293T cells were transfected with plasmids expressing TCF1-myc, SOX13, or Hhex, as shown *below the graph*. The expression of DKK1 was assessed by quantitative RT-PCR. **C**) Immunoblot analysis of the transfected 293T cells. **D)** 293T cells transfected as in B were cross-linked and lysed. The relative amounts of TCF1-myc on DKK1 promoter were analyzed by real-time qPCR of immunoprecipitated chromatin using anti-Myc antibody. Binding of TCF1to the RPLP0 gene was used as a negative control. The image of a representative qPCR stopped at the exponential phase is also shown. Data are the mean of three experiments. *PBGD*, porphobilinogen deaminase.

R2. HHEX INTERACTS WITH C-MYC ONCOGENE INDUCING C-MYC FUNCTION

R2.1. Hhex enhances c-Myc dependent proliferation and transformation

Previous results from different labs support a role for Hhex as a potent regulator of cell proliferation. Hhex controls the proliferation of endodermal progenitors of the pancreas and liver allowing correct cell positioning and permitting proper pancreatic and hepatic development [91, 99]. Ectopic expression of Hhex induces extra bud formation and enhances cell proliferation in the interbud region of the dorsal skin [83]. To investigate Hhex role in cell proliferation we first carried out proliferation assays in cultured human fibroblast cells (FSK) after Hhex expression via adenoviral infection. Human fibroblast cells have a low proliferation rate when compared with transformed cell lines, such as HeLa or 293T cells, allowing us to observe small changes in cell proliferation. As shown in **Figure R2.1 A (compare bar 1 and 2)**, FSK infected with AdHhex had a 30% increase in the proliferation rate measured by bromodeoxyuridine (BrdU) incorporation.

Several lines of evidence suggest that c-Myc plays a significant role in the ability of Hhex to promote proliferation. Hhex and c-Myc co-localize in specific subcellular location, called PML nuclear bodies and also in the vicinity of endoplasmic reticulum and Golgi derived vesicles [59, 242-243]. Mouse mutants lacking c-Myc or Hhex are embryonic lethal at about E10.5 with severe defects in vasculogenesis and angiogenesis [88-89, 159]. Moreover, in both cases there is a decrease in the cell proliferation of the tissues [91, 244]. On the other hand, when cells with unequal levels of c-Myc are apposed, those expressing lower levels are extruded and apoptose [220]. Similarly, chimeric Hhex-/-⇔Hhex+/- livers

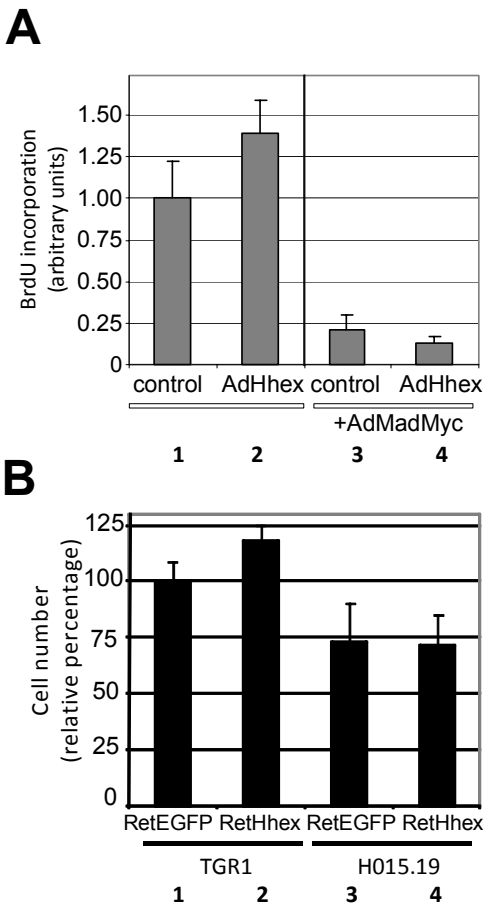


FIGURE R2.1. Hhex induction of proliferation is dependent of c-Myc

A) BrdU proliferation assay of FSK cells infected with Adcontrol or AdHhex (bars 1 and 2), or with AdMadMyc plus AdHhex (bars 3 and 4). **B)** Proliferation assay of TGR1 (c-Myc+/+) cells and H015.19 (c-Myc -/-) cells infected with a retrovirus expressing EGFP (RetEGFP) or HhexEGFP (RetHhex). Green fluorescent cells in each infection condition were counted 24 and 48hours after being seeded. Change in GFP positive cells from 24 to 48 hours was represented as a relative percentatge above total number of cells, setting as 100% the GFP cells found at 24h. Mean of two independent experiments.

result in the specific extrusion and apoptosis of Hhex-null cells into the lumen, at the time-point when Hhex is expressed in the liver bud [99].

To test whether Hhex induction of cell proliferation is dependent on c-Myc activity we used the adenovirus AdMadMyc, an adenovirus expressing a dominant negative c-Myc mutant [204]. As expected, FSK infected with AdMadMyc showed a 75% reduction of proliferation rate (**figure R2.1 A, compare bars 1 and 3**). However, the addition of AdHhex did not enhance the proliferation rate of FSK infected with the AdMadMyc (**compare bars 3 and 4**), suggesting that intact c-Myc activity is necessary for the induction of proliferation caused by Hhex.

To definitively confirm that c-Myc is necessary for Hhex induction of proliferation, we worked with the rat cell line H015.19, a c-Myc mutant rat fibroblast cell line obtained by targeted homologous recombination [245]. H015.19 has impaired proliferation and growth when compared with the parental c-Myc +/+ TGR1 cell line (**figure R2.1 B, compare bars 1 and 3**). In agreement with the experiments in FSK cells, Hhex ectopic expression in control TGR1 cell line significantly increased cell number **(figure R2.1 B, compare bars 1 and 2)**. However, no changes were observed in the c-Myc-/- cell line after Hhex expression (**figure R2.1 B, compare bars 3 and 4**) reinforcing the idea that c-Myc is an essential mediator of Hhex induced proliferation. In conclusion, Hhex is able to promote cell proliferation in fibroblasts by a mechanism dependent of the oncoprotein c-Myc.

We next evaluated whether Hhex could also affect other c-Myc

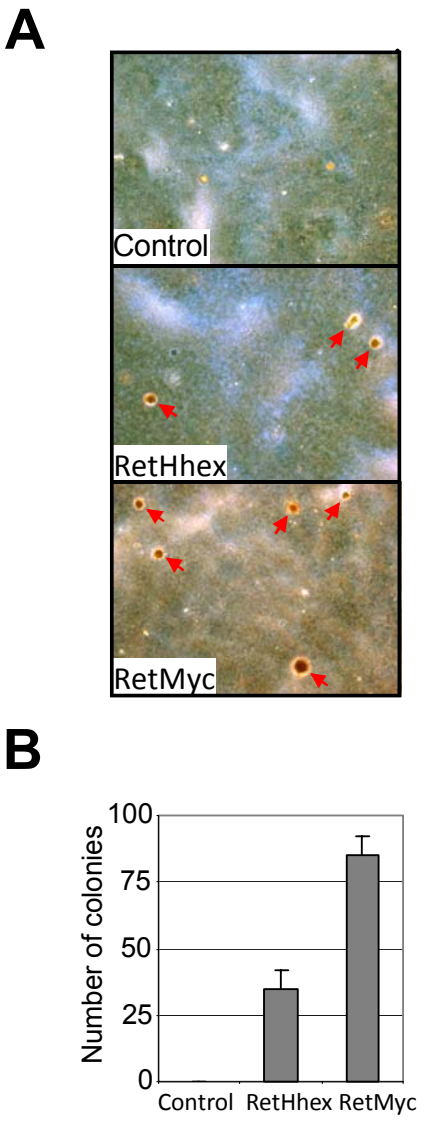


FIGURE R2.2. Hhex promotes cell transformation

A) Representative images of anchorage-independent growth assay of MEFs infected with infected with a retrovirus expressing EGFP (control, top panel) or HhexEGFP (RetHhex, middle panel) or c-Myc (RetMyc, bottom panel). **B)** Quantification of anchorage-independent growth assay of MEFs infected with control, Hhex or Myc retroviruses. Mean of two independent assays.

dependent functions, like cellular transformation. We performed anchorage-independent growth assays in soft agar, a hallmark of cell transformation, using mouse embryo fibroblast cells (MEFs) that have been previously described as suitable for this analysis [156]. While control infected MEFs did not render any colony, MEFs infected with a retrovirus expressing c-Myc dramatically induced colony formation (**figure R2.2 A-B**). Hhex expression also promoted significant colony growth in soft agar when compared with controls (**figure R2.2 A-B**). Thus, Hhex is not only able to affect c-Myc dependent proliferation, but also influences cellular transformation regulated by c-Myc.

R2.2. Hhex interacts with c-Myc

The ability of Hhex to induce c-Myc dependent functions could be the consequence of a physical interaction between both proteins. As a preliminary experiments we performed *in vitro* pulldown assay using bacterially purified GST-Hhex full-length fusion protein (aa1-271) and GST-Hhex deletion mutants (aa1-196; aa1-137; aa138-271; aa197-271) (**figure R2.3 A**) and HeLa protein extracts that endogenously express c-Myc [246]. The presence of c-Myc in the pulled down extracts was assessed by immunobloting against endogenous c-Myc. As **figure R2.3 B** illustrates, GST-Hhex full-length and all the GST-Hhex deletion mutants that contain Hhex N-terminal domain (aa1-137) were able to interact with endogenous c-Myc.

To get a deeper insight into the nature of the interaction between Hhex and c-Myc, we evaluated whether it was direct or through a bridging protein(s). To address this question, we carried out pulldown assays with ^{35}S-labeled *in vitro* translated c-Myc and different GST-Hhex fragments

(**figure R2.3 C**). While GST alone and GST-Hhex (aa138-271) were not able to pulldown c-Myc, full-length GST-Hhex and GST-Hhex (aa1-137) did. We can conclude that c-Myc and Hhex interact directly and that Hhex N-terminal domain is required for this interaction to occur.

To further corroborate the interaction in a cellular context, we set out immunoprecipitation experiments in HeLa cells overexpressing HA-tagged, Hhex and Hhex (aa138-271). Endogenous c-Myc protein co-immunoprecipited Hhex-HA, but failed to immunoprecipitate either Hhex (aa138-271) or control (empty plasmid) transfected HeLa cells (**figure R2.3 D**). Finally, we aimed to determine whether the endogenous c-Myc and Hhex proteins are also associated *in vivo*. For this purpose, human liver protein extracts that express both proteins were used and a co-immunoprecipitation assay with anti-Myc antibody coupled to agarose beads was performed. Only when immunoprecipitation against endogenous c-Myc was done HHEX protein was present (**figure R2.3 E**). In summary, Hhex and c-Myc interact directly, both *in vivo* and *in vitro* and the N-terminal domain Hhex is necessary for this interaction

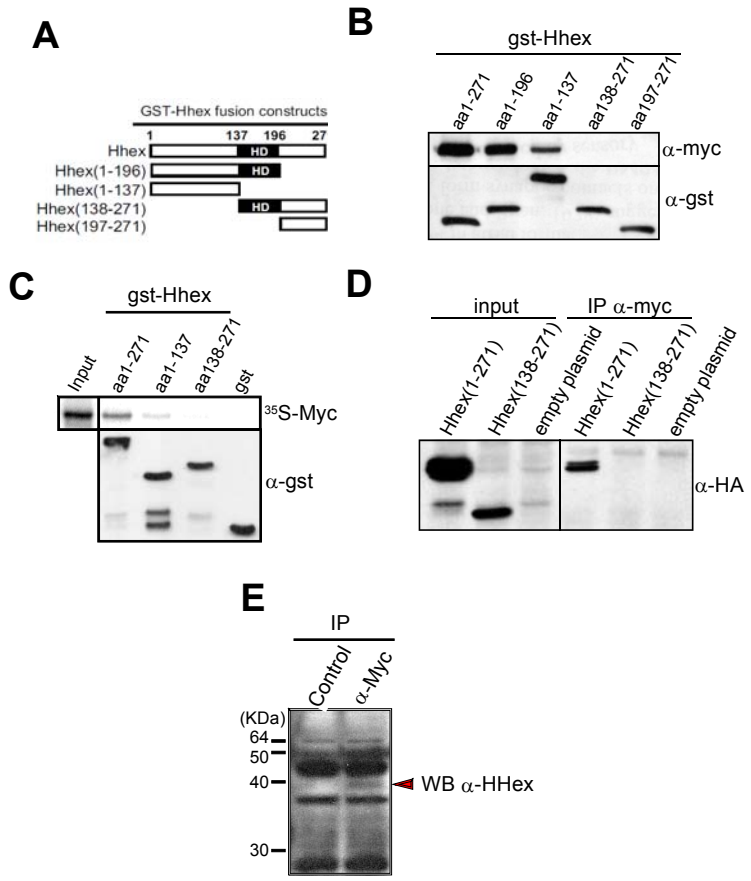


FIGURE R2.3. Hhex and c-Myc interacts *in vivo*, directly, through Hhex N-terminal domain

A) Schematic representation of GST-Hhex and different GST-Hhex deletion mutants used in pulldown assays. Numbers show amino acid positions. HD, homeodomain. **B)** *Upper panel:* Pulldown assay between GST-Hhex, GST-Hhex(1-196), GST-Hhex(1-137), GST-Hhex(138-271) or GST-Hhex(197-271) from E.Coli lysate and HeLa cell extract endogenously expressing c-Myc. HeLa total proteins were incubated with GST-Hhex proteins and gluthahtione-sepharose beads were used to pulldown the GST-HHex proteins. Anti-Myc antibody was used in immunoblot. *Lower panel:* immunoblot using anti-GST antibody of 1% of the beads. **C)** *Upper panel:* ^{35}S-c-Myc were synthesized in an *in vitro* transcription and translation system, and incubated with gluthahtione-sepharose beads coated with GST-Hhex, GST-Hhex (1-137), GST-Hhex (138-271) or GST. Bound proteins were run in an SDS/PAGE and visualized using a Phosphorimager. *Lower panel:* immunoblot using anti-GST antibody of 1% of the beads. **D)** Immunoprecipitation assay (IP) of HeLa cell lysates transfected with Hhex-HA, Hhex (138-271)-HA or empty vector. Cell extracts were immunoprecipitated with an anti-Myc antibody and an immunoblot for HA was performed. One twentieth of the extracts were loaded as the input. **E)** Immunoprecipitation assay (IP) of human liver extracts endogenously expressing HHex and c-Myc. Liver extracts were incubated with ProtG agarose beads (control) or agarose coated with c-Myc antibody (α-Myc) and an immunoblot blot anti-Hhex was performed.

R2.3. C-Myc interacts with Hhex via the C-terminal DNA binding domain (DBD) but the formation of Myc/Max heterodimer is not disrupted

To identify the regions of c-Myc that interact with Hhex, pulldown assays were carried out as described above with GST-Hhex, GST and 293T extracts expressing series of c-Myc protein fragments and deletion constructs (depicted in **figure R2.4 B**). As shown in **figure R2.4 A** and summarized in **figure R2.4 B**, c-Myc deletion mutants lacking C-terminal domain fails to interact with GST-Hhex. Indeed, deletion of only the last 50aa of the C-terminus, which contains part of the bHLH and the LZ domain, was sufficient to abolish the interaction, emphasizing the requirement of this region for the binding. In contrast, the highly conserved Myc Boxes (Mb) in the N-terminal domain seems not to be necessary for the interaction. These results clearly define the C-terminal domain of c-Myc as the binding interface for Hhex.

Taking into account that Hhex interacts with the bHLH-LZ domain of c-Myc, which is critical for the interaction with its obligate partner Max, we next examined whether the interaction affected Myc/Max heterodimer formation in HeLa cells. We run a pulldown assay with GST-Hhex, GST-Hhex (aa1-137), GST-Hhex (aa138-271) or GST alone (negative control) incubated with HeLa protein extracts expressing endogenous c-Myc and Max. As already shown above, Hhex immunoprecipitated c-Myc only when the N-terminal domain was present (**figure R2.4 C**). We also detected Max in the GST pulled down extracts, indicating that Hhex, c-Myc and Max are together in the complex. Thus, Hhex does not alter Myc/Max heterodimerization.

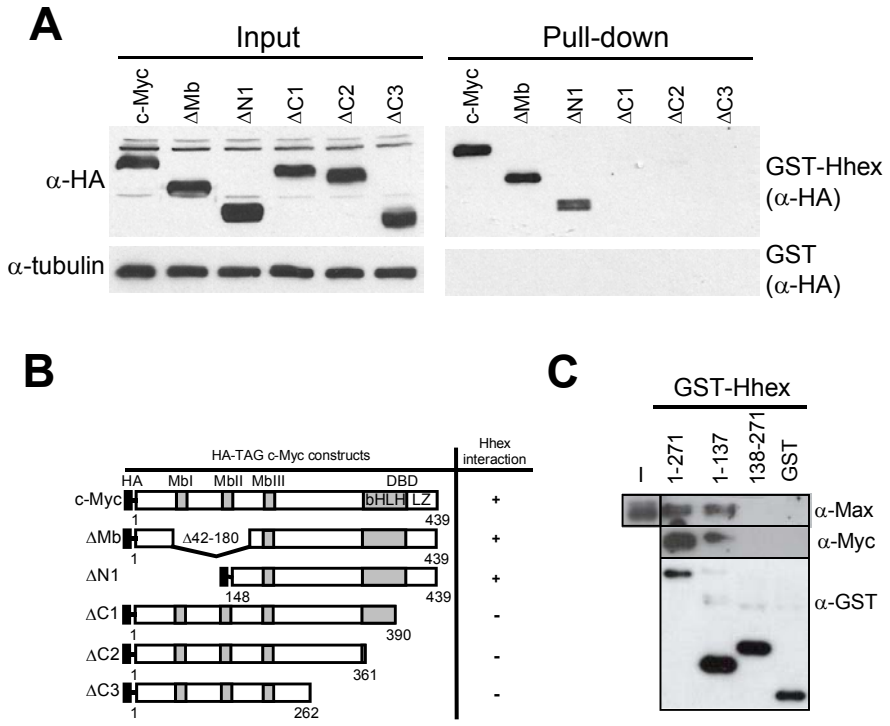


FIGURE R2.4. Hhex interacts with c-Myc C-terminal domain without disrupting Myc/Max heterodimerisation

A) Pulldown assay between GST-Hhex or GST purified proteins and 293T cell extracts expressing HA-tagged c-Myc or the mutated versions ΔMb, ΔN1, ΔC1, ΔC2 and ΔC3. Protein extracts were incubated with GST-Hhex or GST alone. Gluthahtione-sepharose beads were used to pulldown GST proteins. Anti-HA antibody was used in western blots of the pulled down samples. One twentieth of the extracts were loaded as the input. α-tubulin immunoblot was performed as loading control. **B)** Diagram summarizing the results obtained in the GST pulldown assay. Numbers show amino acid positions. *Mb,* Myc box*; DBD,* DNA Binding Domain*; bLHL,* basic helix loop helix; *LZ,* leucine zipper **C)** GST pulldown assay using GST-Hhex, GST-Hhex(aa1-137), GST-Hhex(aa138-271) or GST and HeLa cell extracts expressing endogenous c-Myc and Max. HeLa protein extracts were incubated with GST-Hhex proteins and gluthahtione-sepharose beads were used to pulldown GST proteins. Anti-Myc, anti-Max or anti-GST antibodies were used in western blots of the pulled down samples. One twentieth of the extracts were loaded as the input (I*)*

R2.4. Hhex induces c-Myc-dependent transcriptional activity

The interaction of Hhex with essential regulatory regions of c-Myc (bHLH-LZ) and the subsequent induction of c-Myc dependent functions, led us to speculate that Hhex is a modulator of c-Myc transcriptional activity. To determine whether the effects of Hhex on c-Myc biological activities, showed in **figures R2.1** and **R2.2**, could be mediated by controlling the transcriptional activity of c-Myc, we first examined the effect of Hhex overexpression on a c-Myc-dependent reporter vector. Human 3T3 cells were co-transfected with plasmids expressing c-Myc, Hhex or Hhex (aa138-271) together with the c-Myc reporter plasmid pMYC-TA-Luc. This reporter plasmid expresses luciferase controlled by a promoter containing six tandem copies of E-box consensus sequence, located upstream of the minimal TA promoter. As **figure R2.5 A** illustrates **(compare bars 1 and 2)**, c-Myc overexpression resulted in a 4.5-fold induction of the luciferase reporter. Hhex transfection caused almost 3-fold induction (**compare bars 1 and 4**), while control or Hhex (aa138-271) transfected, which lacks the interaction domain with c-Myc, did not produce any effect on the reporter activity (**compare bar 1 with bars 3 and 5**). Basal luciferase activity was observed in all the combinations when the control reporter vector TA-LUC was used (data not shown). We conclude that Hhex is able to induce c-Myc transcriptional activity by a mechanism that involves the interaction between both proteins. Similar results were obtained when hepatocellular carcinoma derived cell line, HepG2, was used.

Enhanced c-Myc transcriptional activity can be the consequence of increased c-Myc protein levels, or a synergic interaction between both transcription factors. To discern between these two possibilities we analysed c-Myc RNA and protein levels after Hhex overexpression in FSK. As shown in **figure R2.5 B** there is not a significant increase in the *c-Myc* mRNA when Hhex is expressed. Similarly, we didn't observe any changes

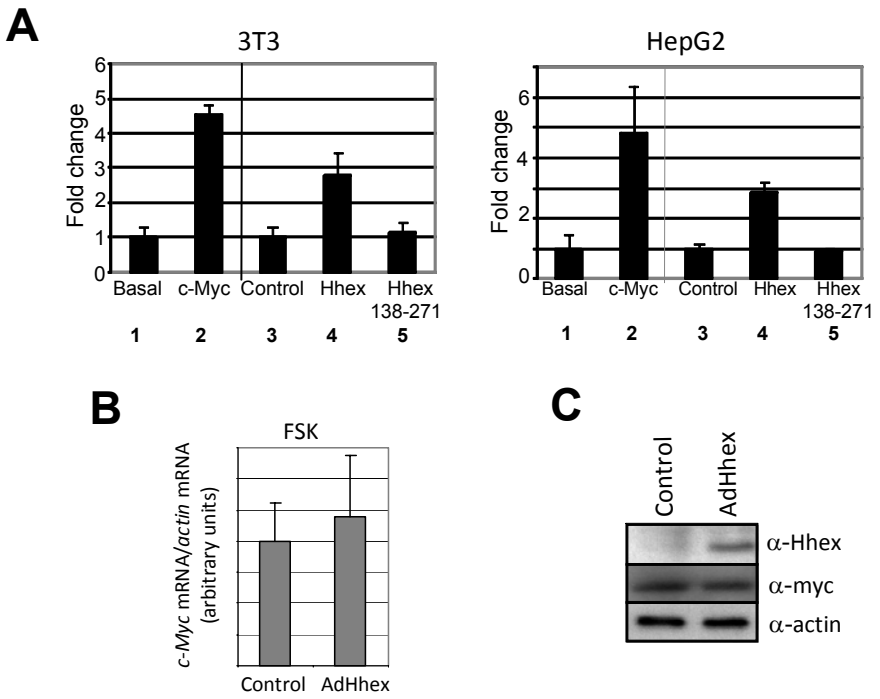


FIGURE R2.5. Hhex induces c-Myc transcriptional activity but not c-Myc levels

A) Reporter assay in 3T3 and HepG2 cells. Cells were cotransfected with empty plasmid (control) or expression plasmids for c-Myc, Hhex and Hhex (138-271), together with a c-Myc dependent reporter (pMyc TA-LUC) and a renilla luciferase plasmid (SV40-LUC) to standardize data. The means and the S.E. of the relative firefly/renilla ratios and fold change over control of at least three independent experiments are shown. The total firefly/renilla of cells transfected only with the reporter plasmid (Basal) was set as 1. **B)** Quantitative RT-PCR of cDNA from FSKs cells infected with control empty adenovirus (control) or with Ad-Hhex. **C)** Immunoblot of infected FSKs cells against Hhex, c-Myc and actin.

at the protein level (**figure R2.5 C**), suggesting that c-Myc is not target gene of Hhex. Our results suggest that Hhex interaction enhances c-Myc transcriptional activity without altering c-Myc mRNA or protein levels.

R2.5. *Cyclin D1* expression, a c-Myc target gene, is regulated by Hhex

To explore the possible synergic effect produced by Hhex on c-Myc activity, we performed loss-of-function experiments using the Hhex knockout mouse. Hhex is essential for proper development of the liver by controlling the proliferation of endodermal progenitor cells in the embryo [99], being the emerging liver bud a perfect model to study Hhex influence on c-Myc activity. We isolated nascent liver buds at 9.5 days of development in Hex-/- (knockout, KO) and Hex+/+ (wild type, WT) mouse embryos as illustrated in **figure R2.6 A**. Given the heterogeneous composition of the liver bud region, we only selected for the analysis, those liver buds with comparable and high expression of the endodermal marker FOXA2 in order to work with comparable amounts of hepatic endoderm. Analysis of *c-Myc* levels in WT and KO liver buds showed that wild type and mutant liver buds had similar levels of *c-Myc*, in agreement with our previous results (**figure R2.7 A**). In order to minimize the contribution of non-endodermal tissues to our RT-PCR analysis, we selected a list of genes that fulfil two criteria: be specifically expressed in the liver/endoderm region during development and be defined as c-Myc target genes. Using Venn diagrams we crossed different databases of embryonic liver/endoderm enriched genes [247-249] with databases of c-Myc target genes [209, 250-253] (**see figure R2.6 B**) and we obtained eight genes that achieved both criteria: Fibronectin (Fn1), Ribosome protein large 4 and 6 (Rpl6, Rpl4) , Nucleophosmin (Npm), Hepatome

derived growth factor (Hdgf), isoleucyl-tRNA synthetase(Iars), eukaryotic translation elongation factor 1 gamma (Ee1fg) and transferrin receptor (Tfrc). Quantitative RT-PCR revealed that there was no difference in the expression levels of any of the eight target genes

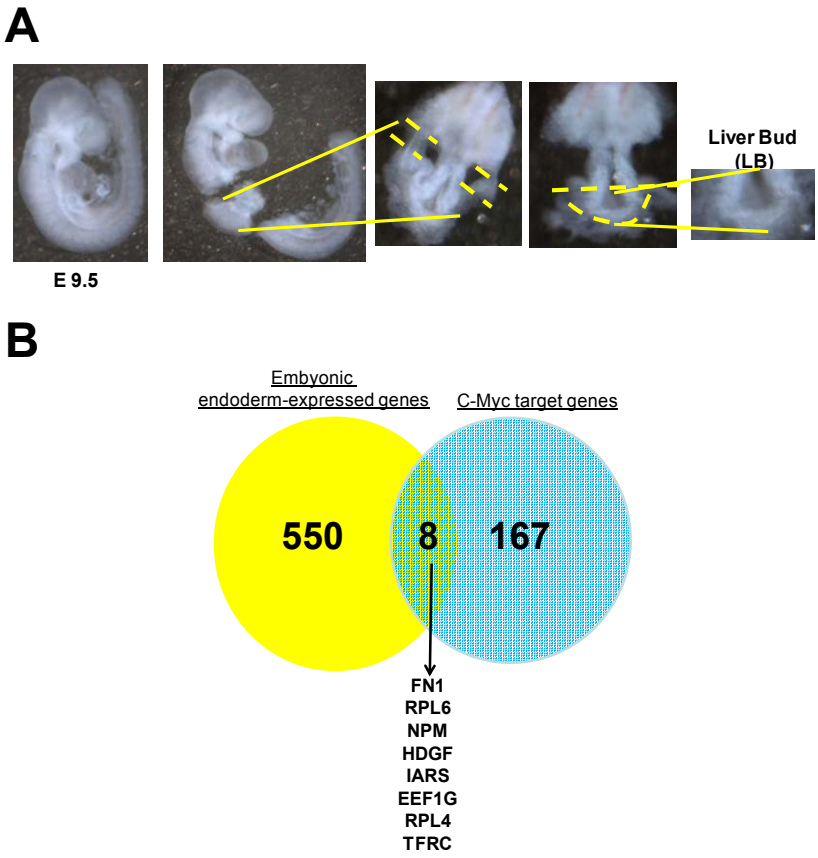


FIGURE R2.6. C-Myc target genes expression in Hhex -/- liver buds (LB)

A) Schematic representation of E9.5 mouse embryo liver bud dissection procedure. Liver buds with similar levels of hepatic endoderm were selected taking into account FOXA2 expression levels. Pool of wild type (Hhex +/+) LBs and four different KO (Hhex -/-) LBs were used for the assay. **B)** Representation of the Venn diagram analysis performed to select c-Myc target genes specifically expressed in the endoderm/liver region. Eight genes accomplish the two conditions: Fibronectin (FN1), Ribosomal protein large 6 (RPL6), Nucleophosmin (NPM), Hepathome Derived Growth Factor (HDGF), isoleucyl-tRNA synthetase(IARS), eukaryotic translation elongation factor 1 gamma (EE1FG), Ribosomal protein large 4 (RPL4) and transferrin receptor (TFRC).

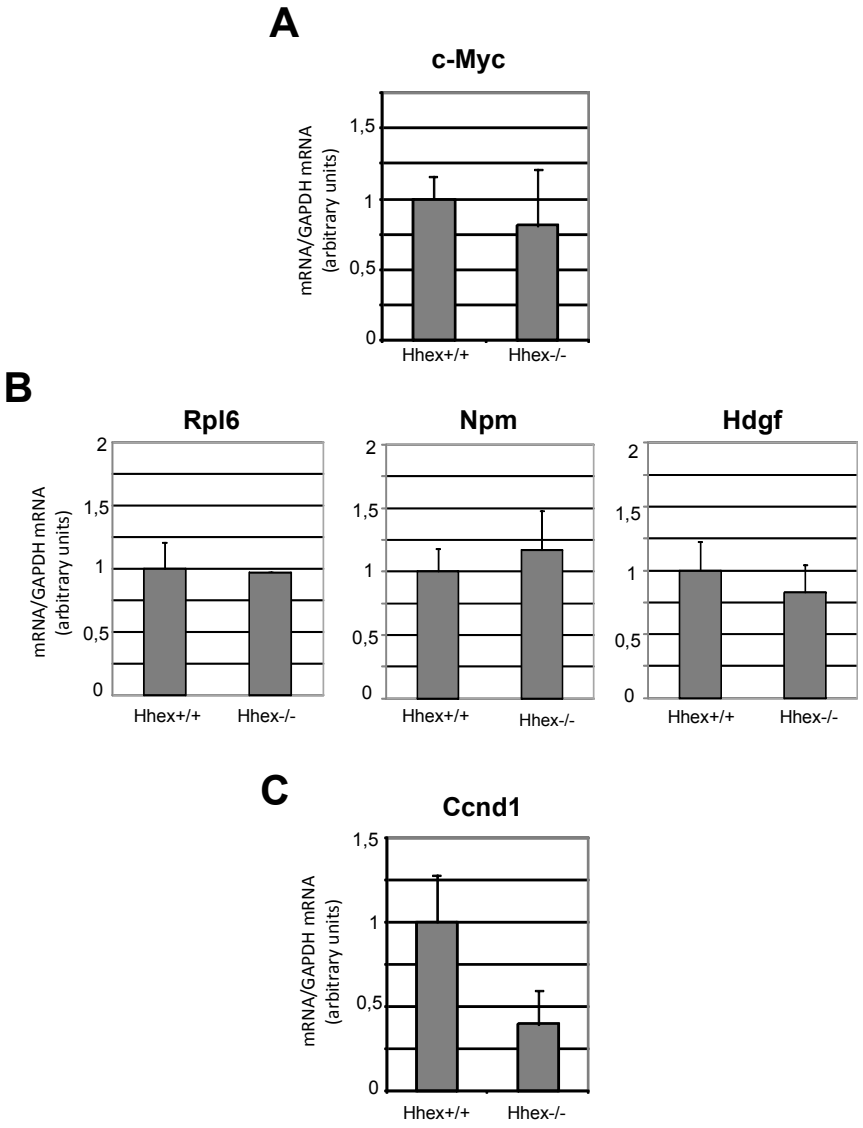


FIGURE R2.7. C-Myc target genes expression in Hhex +/+ and Hhex -/- liver buds

A) Quantitative RT PCR of c-Myc in WT (Hhex +/+) and KO (Hhex -/-) liver buds. **B)** Quantitative RT PCR of c-Myc target genes RPL6, NPM and HDGF, expressed during liver development in WT and KO LB. **C)** Quantitative RT PCR of *Cyclin D1 (CCND1)* in WT and KO liver buds buds. *RPL6: Ribosome protein large 6; NPM: Nucleophosmin; HDGF: Hepatome-derived Growth factor.*

evaluated (**figure R2.7 B** and data not shown) . We decided to expand our analysis to other c-Myc target genes not exclusively expressed in the endoderm but solidly linked to cell cycle and proliferation. We analysed Cyclin D1 (Ccnd1 gene), a c-Myc target gene, essential for cell-cycle progression and proliferation [202]. *Cyclin D1* mRNA was clearly down-regulated in the mutant liver buds when compared to wild type (**figure R2.7 A**). Thus, we conclude that Hhex is not a general modulator of c-Myc activity but rather a specific regulator of *Cyclin D1* expression.

Since Cyclin D1 has not an endoderm restricted expression, we decided to extend our study to a simpler and more general cellular model, the FSK cells. Apart from CCND1, we measured the mRNA levels of some of the most relevant c-Myc target genes involved in cell cycling: Cyclin D2 (CCND2), CDK4, GADD45 and p15 (CDKN2B), in control or AdHhex infected FSKs (**figure R2.8**). Quantitative RT-PCR analyses revealed increased expression of *CCND1*, *CCND2* and *CDK4*, with the highest increase in the case of *CCND1*, showing a 3-fold increase in the mRNA levels when compared to control. No changes were observed in other target genes with non-related cell cycle functions, such as *RCL* (**figure R2.8 A**), or in the case of *GADD45* and *p15* **(figure R2.8 B).** Noteworthily, *CCND1*, *CCND2* and *CDK4* are positively regulated c-Myc targets while *GADD45* and *p15* are transcriptional-repressed by c-Myc. Consistent with the increase in the mRNA levels, Cyclin D1 protein levels were also elevated (**figure R2.8 C**.)

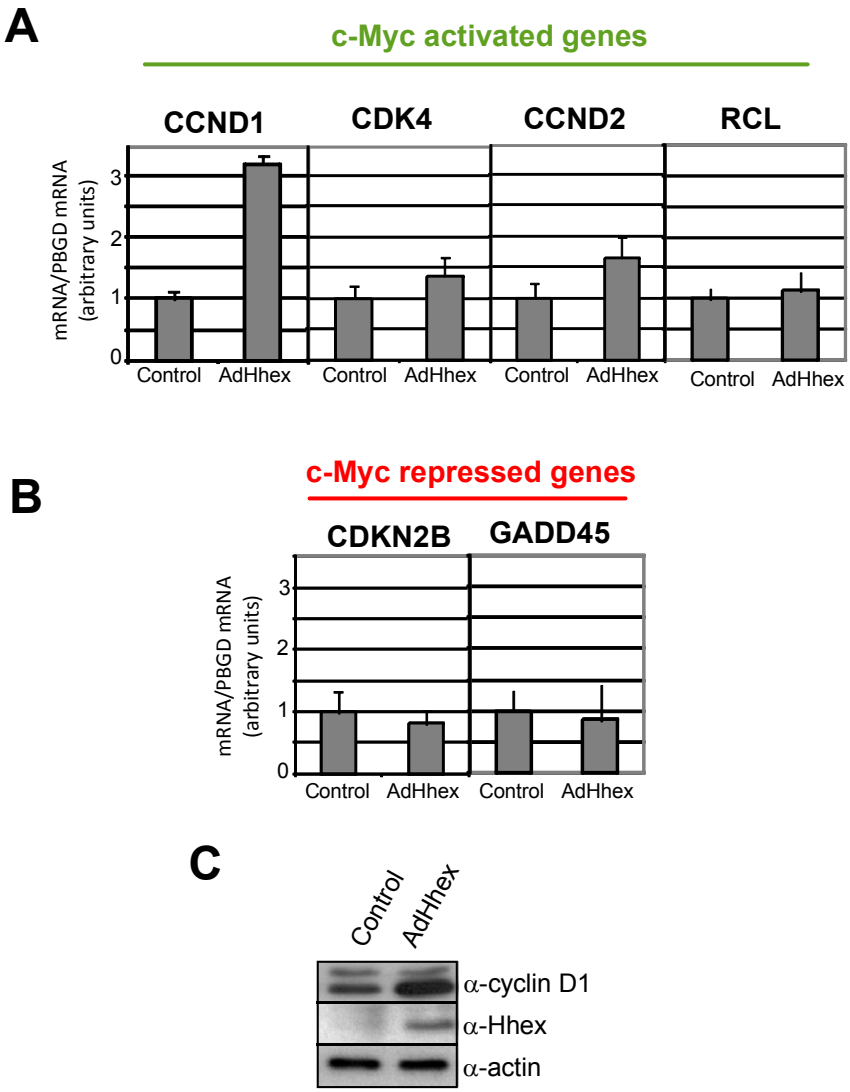


FIGURE R2.8. Hhex increases Cyclin D1 expression

A) , B) FSKs cells were infected with Adempty control adenovirus (control) or with AdHhex and quantitative RT-PCR of a set of c-Myc activated genes or c-Myc repressed genes was performed respectively. Mean of at least three independent experiments**. C)** Immunoblot of control infected or AdHhex infected FSK cells against Cyclin D1, Hhex, and actin as a loading control. *CCND1: Cyclin D1; CCND2: Cyclin D2; CDK4: Cyclin-dependent kinase 4; RCL: Deoxynucleoside 5'-monophosphate N-Glycosidase; GADD45: Growth arrest and DNA damage inducible, alpha; CDKN2B: Cyclin-dependent kinase inhibitor 2B (p15).*

To sum up, Hhex is specifically inducing positive cell cycle c-Myc targets, but not repressed genes. The most profound effect was found with *Cyclin D1* expression which correlates with an increase in its protein levels.

DISCUSSION

D1. INTERACTION BETWEEN HHEX AND SOX13 MODULATES WNT/TCF ACTIVITY

We have shown that binary interactions between Hhex, SOX13, and TCF regulate Wnt activity *in vitro* and *in vivo*. The three genes are co-expressed in several embryonic tissues, like the liver and pancreas [78, 137]. The β-catenin·TCF complex is a downstream effector of the Wnt signaling pathway **(figure D1)**. SOX13 binds the transcription factor, TCF1, possibly through the LZ-Q domain to disrupt the β-catenin·TCF complex and inhibit Wnt activity [135]. The addition of Hhex results in the formation of the Hhex·SOX13 complex **(figure R1.4)** to dislodge SOX13 from the SOX13·TCF1 complex **(figure R1.7)** and results in a restoration/elevation of the Wnt levels **(figure R1.6 A, R1.8, and R1.9).** In short, Hhex and SOX13 act as a pair of antagonists of the Wnt activity to achieve the right level in different biological contexts.

SOX13 as new Hhex partner

Through the screening for Hhex-binding proteins with an E9.5-E10.5 mouse embryo library [231], we isolated cDNA for a HMG box protein, SOX13. In a previous yeast two-hybrid screening using Hhex as bait, two interactors were isolated; the HC8 subunit of the proteasome [62] and the CK2 [67]. In our study we have used an embryo-derived library instead of a chronic myelogenous leukemia cell line (K562)-derived cDNA library, which lacks any developmental context. Other studies using different baits have identified Hhex as an interacting partner of Jun [61],

GATA2 [64], Sox10 [66], TLE1 [63], or PML [59]. Noteworthy, our screening confirmed the interaction of Hhex with TLE1. Besides SOX13, we obtained other 68 candidates, 24 of them with 2 repetitions or more (**Table R1.1**). SOX13 was isolated 8 times with a common spanning region of ~100 aa **(figure R1.2).**

Among the different SOX13 mRNAs described in the literature, we identify SOX13 (1815 bp) and S-SOX13 (768 bp) as the isoforms present in human tissues. Our identification is in conflict with Roose et al. [126] who described a longer open reading frame of 2673 bp isolated from embryonic thymic cDNA. We did not find a significant amount of this longer mRNA in any of the 18 tissues tested, including adult thymus, which is also in agreement with a later report from Harley and co-workers [121].

SOX transcription factors are involved in the differentiation of multiple cell types and developmental processes from the endoderm. Sox17 plays an essential role in definitive endoderm formation in the mouse [254]. Sox9 is expressed in the pancreas where it plays a role in the maintenance of a progenitor cell pool [255]. It is also expressed asymmetrically in the ductal plate, and it seems to control the timing of biliary tubulogenesis [256]. Sox4 is essential for the normal formation of endocrine pancreas [257]. SOX13 is expressed in the pancreas and liver, in the visceral mesoderm of the extraembryonic yolk sac, and also in the spongiotrophoblast layer of the placenta [136-137]. At the biochemical level, SOX13 contains a leucine zipper and glutamine rich domain at the C- terminus and the HMG box, an 80-amino acid sequence that mediates DNA binding [126]. Which domains are responsible for Hhex/SOX13 interaction? The N-terminus of Hhex (aa1–137) is apparently necessary

and sufficient for SOX13 interaction. When used as bait, it interacts with the LZ-Q domain of SOX13 to transactivate MEL1 and LacZ reporter genes under the control of a heterologous GAL4-responsive upstream activation sequence and promoter elements in yeast (**figure R1.2**). In a GST pulldown assay, only Hhex itself and its deletion mutants containing the N-terminal domain, i.e. Hhex(1–137) and Hhex(1–196), achieve a significant amount of SOX13 pulldown (**figure R1.4**). On the other hand, the LZ-Q domain in SOX13 is necessary, but not sufficient, for interaction in cells. S-SOX13 contains the LZ-Q domain but does not interact with any fragment of Hhex. ΔSOX13, which lacks only the LZ-Q domain, is not able to interact either. If Hhex interacts with the LZ-Q, then we may expect its effects to be mediated by a disruption of the putative SOX13 homodimer. However, SOX13 might act as a monomer, as suggested by the fact that co-transfection with S-SOX13 (containing the LZ-Q domain but not the HMG domain) did not block the inhibitory effect of SOX13 on Wnt activity **(Annex Figure A2)**. In short, the N-terminal domain of Hhex mediates the binding to SOX13 in mammalian cells, whereas the LZ-Q domain in SOX13 is necessary, but not sufficient, for the interaction. The interaction between Hhex and SOX13 seems to be specific as SOX9 did not interact with Hhex in an immunoprecipitation assay. Finally, we showed that SOX13 and Hhex interact in a cell system and can do so directly without a bridging protein (**figure R1.5**).

Hhex binding to SOX13 establish a new mechanism for the modulation of Wnt/TCF activity

The Wnt/β-catenin signaling pathway describes a complex network of proteins most well known for their roles in embryogenesis and cancer but also involved in normal physiological processes in the adult (**figure I2**). There is an emerging evidence of a role for SOX proteins as modulators of canonical Wnt/β-Catenin signaling in diverse development and disease contexts [258]. In most of the cases SOX proteins repress Wnt transcriptional responses; however, some SOX proteins appear to enhance Wnt-target gene expression. This modulation is done through a variety of mechanisms, including protein-protein interactions, DNA-binding, recruitment of cofactors and protein stability. For instance, the interaction of SOX7 with the β-catenin·TCF complex promotes β-catenin degradation resulting in repression of the pathway in colon cancer cells, suppressing the hyperactive β-catenin activity, thus producing a decrease on cell proliferation [259-260]. Conversely, SOX4, which is frequently over-expressed or genetically amplified in tumors [261-265], binds to β-catenin producing its stabilization and inducing Wnt pathway in colon cancer cells [266]. SOX13 has been described to repress Wnt-targets in T-cell development not by binding to β-catenin, but to TCF1 [135]. This interaction seems to be through TCF1 N-terminal domain, the same necessary for β-catenin interaction, suggesting that SOX13 inhibits Wnt signaling by excluding β-catenin from the complex.

Our work uncovered the independent structural requirements of Hhex and SOX13 to effectively modulate Wnt activity. Hhex has been recently described as an inducer of Wnt by repressing the expression of TLE4 [71]. However, Hhex(1–196), which retains the repressor domain but lacks the

transactivator domain, did not induce Wnt activity **(figure R1.6 A**). In parallel, Hhex(138–271), which lacks the repressor domain, induced Wnt activity by almost 4-fold, suggesting that Wnt activation depends on the C-terminus of Hhex. It is feasible that the assigned transcriptional function of specific Hhex domains might be more plastic than initially thought. Alternatively, TLE4 might be a secondary target of Hhex. Fragments of Hhex lacking the homeodomain, i.e. Hhex(197–271) and Hhex(1–137), did not alter Wnt activity (data not shown). This is possibly linked to their extranuclear location (**Annex figure A1**) and the inability to bind DNA. On the other hand, SOX13 integrity seems to be crucial for Wnt repression. A specific deletion of the LZ-Q domain in SOX13 (ΔSOX13) completely abolished Wnt repression. S-SOX13, which lacks the C-terminal domain of SOX13, was also unable to inhibit Wnt activity (**figure R1.6 B**). An inspection of S-SOX13 structure suggests that it might function as a SOX13 inhibitor. However, it did not interfere with the SOX13-dependent regulation of Wnt. Full-length Hhex is able to restore Wnt inhibition after SOX13 co-transfection in 293T cells. Hhex domains have a statistically significant difference to restore SOX13-dependent Wnt repression. Hhex(1–196) was as efficient as the full-length protein, whereas Hhex(138–271) only caused a certain degree of Wnt restoration at the highest dose, which was significantly less pronounced (**figure R1.6 A**). Given that Hhex(138–271) lacks the N- terminus involved in SOX13 interaction, this result suggests that Wnt activity restoration is caused by direct SOX13·Hhex interaction. In agreement with this hypothesis, Hhex-VP2, a modified form of Hhex containing the N-terminus that behaves as a transcriptional activator is still able to reinstate Wnt activity (**figure R1.7 B**). Thus, Hhex influences Wnt signaling in the presence of SOX13 and does not reside in Hhex transcriptional activity. Reporter assay

analysis in HepG2 cells confirmed that Hhex is able to block the inhibitory effect of SOX13 either by endogenous expression in HepG2 cells (**figure R1.8 B**) or by overexpression in 293T cells (**figureR1.6 A, group 6, black bar**). Moreover, knockdown of endogenous HHEX in HepG2 restored the inhibitory activity of SOX13 (**figure R1.8 B, group 3**).

We validated in cultured mouse embryos the results obtained in cell cultures (**figure R1.10**). It is somewhat striking that ectopic expression of Hhex cDNA in 293T cells induces Wnt activity, whereas in the embryo it inhibits Wnt activity. We can best explain this discrepancy by suggesting that Wnt induction is a cell-autonomous effect of Hhex. In the embryo, non-cell autonomous effects, such as interaction between Hhex expressing endoderm and mesoderm during patterning, may induce expression of the Wnt antagonist that ultimately accounts for a general inhibition of Wnt activity. In fact, phenotypes observed in response to overexpression of Hhex resemble those obtained after expression of Wnt antagonists [70]. Loss-of-function of Hhex in the mouse causes a dorsalization of the embryo, where most of the ventral structures are missing, i.e. forehead, thyroid, liver, or ventral pancreas [88, 91], which is compatible with extended Wnt activity [267].

Keeping in mind that our reporter assay is based on an heterologous promoter, we wanted to confirm our results using an endogenous promoter. We chose DKK1 because it is a relevant developmentally regulated gene and it has been validated extensively as a β-catenin·TCF target [232]. The mRNA levels of DKK1 reproduced the expression pattern obtained with the TOPflash reporter plasmid. Moreover, DKK1 mRNA expression profile was paralleled by TCF1 occupancy of the DKK1 promoter. In summary, the results obtained in

reporter assay were validated with endogenous TCF1-regulated promoter by qRT-PCR and a chromatin immunoprecipitation assay (**figure R1.11).**

Hhex↔SOX13↔TCF1 interactions: achieving correct Wnt intensity for proper liver development?

There seems to be an intense tissue-specific cross-regulation between Hhex and Wnt during development. Activation of Wnt/β-catenin signaling in Zebrafish [87] induces Hhex expression in the dorsal yolk syncytial layer. In agreement, β-catenin deficient mouse embryos do not express Hhex in the anteriorposterior axis on embryonic day 5.5 [268]. However, inhibition of the canonical Wnt pathway in *Xenopus* induces Hhex in the underlying endoderm of the heart field [269]. Do modulators like SOX13 determine the output of Hhex and Wnt interrelationship in different tissues? Wnt regulation is particularly essential for the specification of the liver and pancreas from the ventral foregut endoderm. Wnt down-regulation in the anterior endoderm is shown to be crucial for liver and pancreas specification in the ventral foregut endoderm. But immediately after inducing the hepatic program in the endoderm, Wnt signaling is apparently required for the endoderm to outgrow further into a liver bud [6] (**figure I4**). In Zebrafish, the expression of Wnt2b in the lateral plate mesoderm, acting through the β-catenin canonical pathway, appears essential for liver specification in the endoderm and bud induction [270]. Briefly, there must be fast and sharp control of Wnt in the early steps of liver and pancreas specification. Part of this control is exerted by the expression of the Wnt inhibitors [70]. In

this study we propose that the reciprocal interactions between the triad Hhex, SOX13, and TCF1 together with the auto regulatory loop Hhex↔Wnt contribute to achieve the correct Wnt intensity in the appropriate spatiotemporal dimension (**figure D1**).

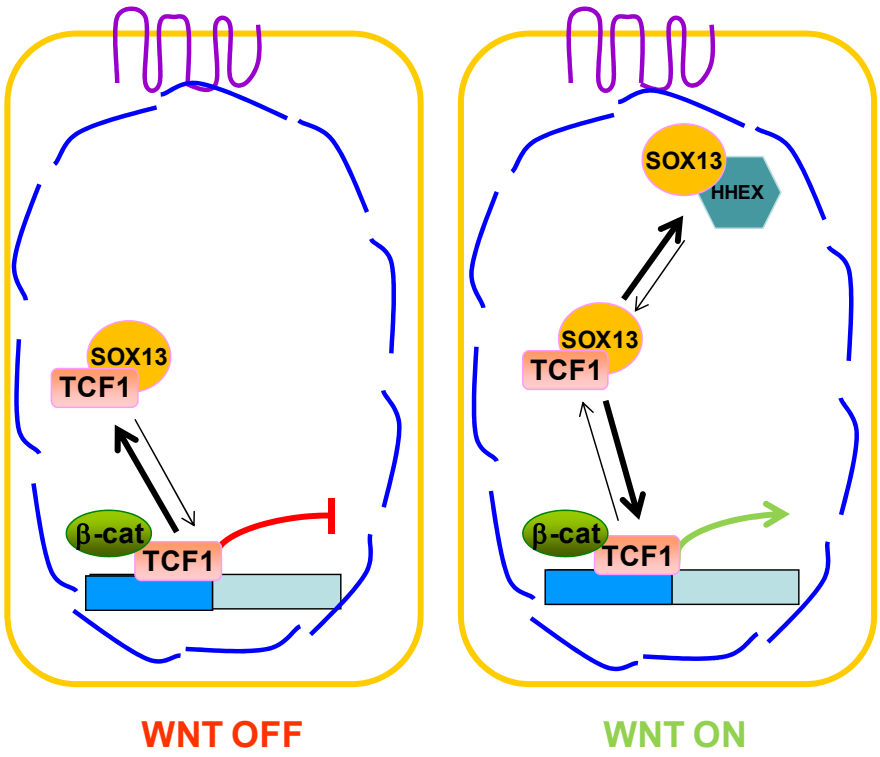


FIGURE D1. Model depicting the role of Hhex in Wnt activity modulation.
In this model Wnt-responsive genes are inhibited in the presence of SOX13 because of its interaction with the TCF1 protein, which disrupts the TCF1·β-catenin complex. The addition of Hhex results in the formation of a Hhex·SOX13 complex, displacing SOX13 from TCF1 and allowing TCF1 to form the effector complex in sensitive promoters once more.

D2. HHEX INTERACTS WITH C-MYC ONCOGENE INDUCING C-MYC FUNCTION

We have shown that the homeodomain transcription factor Hhex interacts with c-Myc oncogene inducing c-Myc dependent proliferation and transformation. We have also shown that Hhex enhances c-Myc transcriptional activity and controls the expression of Cyclin D1, Cyclin D2 and CDK4, c-Myc targets involved in the control of G0/G1 to S phase progression of the cell cycle.

Hhex as c-Myc-dependent promoter of cell proliferation

Previous studies have shown that Hhex has a complex role regulating proliferation and differentiation. During development, Hhex expression is linked to both processes, depending on the developmental stage and the tissue (see **figure I10**). Studies with the knockout mouse have highlighted the importance of Hhex in controlling the proliferation rate, and thus the positioning, of the leading edge of ventral endoderm cells that will give rise to the liver and the ventral pancreas [91, 99]. Apart from its role in cell proliferation, recent studies in embryonic stem cells and conditional knockouts suggest that Hhex also regulates hepatocyte differentiation in later developmental stages [92, 271]. In the case of T-cell development, Hhex expression is necessary for proper T-cell differentiation [96], whereas in the skin, exogenous Hhex promotes epidermal cell proliferation [83].

To explore how Hhex controls cell proliferation we used a gain-of-function model based in the expression of Hhex in human skin fibroblast cells (FSK), which are derived from normal human skin and not from tumoral tissues. We initially chose low passage FSK because they maintain the replication machinery unaltered, allowing to investigate the role of Hhex in a quasi-physiological cellular environment [272]. Hhex expression in FSK resulted in an increase in the proliferation rate measured by BrdU incorporation (**figure R2.1 A**). In a similar experiment, Hhex also caused an increase in TGR1 cell numbers (**figure 2.1 B**). We observed a lower effect of Hhex in TGR1 than FSK, probably linked to the immortalized nature of the TGR1 that alters growth and dysregulates cell proliferation control.

What is the mechanism underlying Hhex stimulus of cell proliferation? As explained in the results section, several lines of evidence led us to speculate that c-Myc participates in cell proliferation control by Hhex (see **results section R2.1**). Two different approaches confirmed this hypothesis; first, Hhex induction of cell proliferation was abolished when co-expressed with MadMyc protein, a dominant negative form of c-Myc (**figure R2.1 A**). Second, we have not observed changes in cell numbers when Hhex was ectopically expressed in the c-Myc null cell line H015.19 (**figure R2.1 B**). Cell number is the output of proliferative and death signals. In fact, cell number can also increase when apoptotic pathways are inhibited. It is still unclear the possible role of Hhex in apoptosis, but the increase in cell number we have described, is at least partially associated to an increase of cell proliferation as observed in FSK cells. Taking into account the ability of c-Myc to induce apoptosis of the neighbouring cells expressing lower levels of c-Myc (termed known as

"supercompetition", [220]), and considering that Hhex-expressing cells have an induced c-Myc activity, we cannot exclude the possibility that apoptosis triggered by Hhex could also counteract the increase in the proliferation rate and result in lower than expected cell numbers. The possible role of Hhex in apoptosis and supercompetition is currently being explored in Drosophila S2 cell line.

Cell transformation is a complex process involving alteration in the genetic material of the cell turning normal cells into tumoral cells with high capacity of migration and invasion. Both proteins, Hhex and specially c-Myc have been implicated in tumorigenesis. Hhex is expressed in thyroid and breast tumors, where it is mislocated in the cytoplasm compartment [50, 55]. Decreased Hhex expression and loss of nuclear localization of Hhex is also implicated in a number of human myeloid leukaemias [56, 273]. In addition, a chromosomal translocation resulting in a Hhex fusion protein that can activate transcription and has trans dominant-negative activity over wild type Hhex has been shown to be a causative agent in acute myeloid leukaemia (AML) [274]. On the other hand, c-Myc is capable of inducing cell transformation when ectopically expressed in normal cells. Indeed, overexpression of c-Myc in transgenic mice results in the development of mammary, haematopoietic, hepatic, pancreatic and epidermal tumours among others [275-279]. In our study, we found that Hhex promotes cell transformation in mouse embryonic fibroblast, measured by anchorage-independent growth assay **(figure R2.2).**

c-Myc oncogene, new partner of Hhex

The ability of Hhex to induce c-Myc dependent functions could be the consequence of a physical interaction between both proteins. In fact, we have found that Hhex and c-Myc interact directly, both *in vitro* and *in vivo*, in a cellular context and in the liver tissue. The N-terminus of Hhex (aa1-137) is essential for c-Myc interaction, since no interaction was observed in pull-down or immunoprecipitation assays when Hhex mutants lacking this domain were used (**figure R2.3**). C-Myc C-terminal is the interface necessary for the interaction with Hhex (**figure R2.4**). C-terminal region of c-Myc contains a basic region (aa355-367) which serves as a specific DNA binding domain (DBD) with two protein-protein interaction motifs, a basic helix-loop-helix domain (bHLH) (aa368-410) and a leucine-zipper motif (LZ) (aa411-439)[280]. Deletion of this region inhibits all of the biological activities of c-Myc, suggesting that protein-protein interaction and DNA binding are required for c-Myc function [280]. Specifically our results have shown that LZ domain and a part of the bHLH domain of c-Myc are essential for the interaction with Hhex, since no interaction was observed in the mutants that lack this domain (**figure R2.4 A, B**). C-terminal region of c-Myc is also crucial for the interaction with other proteins, such as nucleophosmin (NPM) [156]. Similar to Hhex, NPM interaction with c-Myc is involved in c-Myc dependent hyperproliferation and transformation, not affecting c-Myc·MAX heterodimer formation (**figure R2.4 C**).

Enhancement of c-Myc activity by Hhex

At this point we hypothesized that Hhex interaction with c-Myc increases the transcriptional activity of c-Myc. In agreement, reporter assays using a c-Myc-dependent reporter have shown that full-length Hhex induces 3-fold c-Myc dependent transcription (**figure R2.5 A**). Moreover, a specific deletion mutant of the Hhex (Hhex138-271), which lacks the domain necessary for the interaction with c-Myc, completely abolishes Hhex induction. We discarded the possibility of a direct effect of Hhex138-271 as a repressor of c-Myc activity per se, since this mutant lacks Hhex repressor domain (aa1-aa137). Similar results were obtained in 3T3 (mouse fibroblast) and HepG2 (human hepatoma derived) cell lines, ruling out the possibility of a cell-specific effect. In short, we have found a new interactor partner for c-Myc, Hhex, which interacts through its N-terminal domain with the C-terminal region of c-Myc, in a cell environment and can do it directly without disrupting the formation of c-Myc/MAX heterodimer and resulting in increased transcriptional activity.

The enhancement in c-Myc transcriptional activity can be associated to increased c-Myc protein levels or to a synergistic effect of Hhex in c-Myc·MAX transcriptional control. Accumulation of c-Myc protein is the result of changes in transcriptional initiation and attenuation, as well as posttranscriptional regulation at the levels of mRNA stability, translation, and protein stability [281]. Our data have demonstrated that ectopic expression of Hhex in FSK cells does not raise c-Myc mRNA or protein levels (**figure R2.5 B, C**). We have not observed changes in *c-Myc* mRNA levels in the liver bud of Hhex knockout embryos (**figure R2.7 A**), reinforcing the idea that Hhex expression does not result in the accumulation of c-Myc protein. To explore the possible synergistic effect

induced by Hhex, we have used the liver bud region of Hhex knockout mouse embryos at the time point when Hhex is normally expressed (E9.5). The liver bud region contains significant amounts of non-endodermal cells originated in the septum transversum mesenchyme and adjacent tissues [22]. Using Venn diagrams we have found a set of c-Myc target genes specifically expressed in endodermal cells (FN, RPL4, RPL6, IARS, EE1FG, NPM, HDGF and TFRC) (**figure R2.6 B**). Quantitave PCR assays in wild type and Hhex knockout liver buds revealed not significant differences in the expression levels of any of these genes (**figure R2.7 B and data not shown**). However, we found a significant decrease in *Cyclin D1* mRNA levels (**figure R2.7 C**). We may conclude that Hhex is not a wide range controller of c-Myc transcriptional control, but a selective regulator of Cyclin D1, a c-Myc target. Noteworthy, as mentioned above, absence of Hhex in the liver bud does not decrease the levels of *c-Myc* (**figure R2.7 A**), suggesting that part of c-Myc regulation of these targets might be maintained in the knockout liver bud.

Cyclin D1, an activated c-Myc target gene, is regulated by Hhex

A possible caveat in our results is that *Cyclin D1* is expressed in endoderm and non-endoderm cells of the liver bud region. In order to clarify whether Hhex is able to control *Cyclin D1* mRNA, we evaluated the expression of *Cyclin D1*, as well as other c-Myc targets involved in cell cycle control, in FSK cells overexpressing Hhex. We found a significant and profound increase in the levels of *Cyclin D1* mRNA that was

correlated with increased protein levels (**figure R2.8**). We also found a slighter, but significant, increase in *CDK4* and *Cyclin D2* expression. Other c-Myc target, the enzyme RCL, with non-related cell cycle functions, did not show any significant change in its expression in the presence of Hhex. Similarly, no changes have been observed in repressed c-Myc target genes, *GADD45* and *CDKN2B*, suggesting that Hhex may selectively regulate the transcriptional status of positive c-Myc targets related to cell cycle control. Given that Cyclin D1/CDK4 complex promotes progression through the G1-S phase of the cell cycle and increased proliferation [282-283] it is plausible that the observed decrease in S-phase transition in the Hhex-/- liver bud [99] is caused by decreased *Cyclin D1* levels in the embryonic domain. From the gain-of-function perspective, amplification or overexpression of Cyclin D1 plays pivotal roles in the development of a subset of human cancers (including parathyroid adenoma, breast cancer, colon cancer, lymphoma, melanoma, and prostate cancer) which could also contribute to the tumorigenic activity attributed to Hhex [284].

We conclude that Hhex binds to c-Myc·MAX and increases the endogenous transcriptional activity of the heterodimer resulting in upregulation of positively regulated genes but does not modify the mRNA level of negatively regulated genes. It is well known that c-Myc binds to its partner MAX and activates transcription by binding to E-box elements in the target gene promoters. Activation involves the recruitment of multiple coactivators. No monomeric Myc proteins have been found *in vivo*. However, the mechanism used by c-Myc to repress gene transcription is not fully understood. It is hypothesized that Miz1, another Myc interacting partner, binds to specific sequences in c-Myc repressed gene promoters and recruits the p300 histone

acetyltransferase to stimulate transcription. The c-Myc·MAX heterodimer blocks transactivation by Miz1, partly through binding Miz1 and disrupting the interaction between Miz1 and the gene promoter or by disrupting the interaction with p300 and by recruiting the DNA methyltransferase DNMT3a. In agreement with this model, TGFβ represses p15 transcription by releasing Miz1 from the c-Myc·MAX complex by downregulating the c-Myc protein [285]. Our hypothesis is compatible with this mode of action. Since Hhex binds to the c-Myc·MAX heterodimer and increases its transcription capabilities, but does not increase c-Myc protein levels, it is expected that Hhex expression will increase mRNA levels of positive regulated genes without interfering with negatively-regulated genes.

Another important question for debate is whether Cyclin D1 upregulation is caused exclusively by c-Myc·MAX enhanced transcriptional activity. Since the first description of the Cyclin D1 promoter emerged almost 15 years ago [286-287], many different transcription factors have been identified that directly bind to, or otherwise regulate, the Cyclin D1 promoter [288]. One of these transcriptions factors is c-Myc. Direct binding of c-Myc to the CCND1 promoter has been demonstrated by Chromatin Immunoprecipitation assay (ChIP) [158, 289-290], suggesting that there is a direct mechanism of Cyclin D1 transcriptional control by c-Myc. We have recently described that Hhex is able to positively regulate Wnt activity [57]. Cyclin D1 is a well-known target of Wnt signaling [291]. Thus, it is feasible that Wnt induction could be also responsible for Cyclin D1 upregulation. However, we show that Hhex enhanced proliferation is dependent on c-Myc (**figure R2.1**), excluding a direct Wnt-Cyclin D1 epistatic regulation to explain the observed increase in proliferation.

Nonetheless a contribution from Wnt induction to the overall increase in Cyclin D1 protein levels cannot be discarded. To dissect the specific contribution of each pathway in Cyclin D1 expression, Wnt loss-of-function mutants will be needed.

The kind of regulation exerted by c-Myc on Cyclin D1 expression is controversial. Initial studies using chimaeric inducible Myc-ER protein expression reported both transactivation and repression of *Cyclin D1* [205, 292]. Nevertheless, subsequent studies have shown that this initial observation may be compromised by the presence of a cryptic, estrogen receptor moiety of the Myc-ER fusion protein [293]. Studies in c-Myc-/- cell lines showed a decreased Cyclin D1 mRNA expression [294]. Other reports propose that the type of regulation depend on the c-Myc partner. This is the case of the tight junction protein ZO-2 (zona occludens-2) that down-regulates Cyclin D1 transcription by interacting with c-Myc in the Cyclin D1 promoter region [289]. This kind of regulation might be comparable to the one exerted by Hhex. In general, Cyclin D1 regulation by c-Myc is unclear, and it may depend on the cell type, cellular conditions and c-Myc interacting partners.

Hhex: an inductor of cell proliferation in non-tumorigenic cells?

In our hands Hhex promotes cellular proliferation in primary fibroblast cells. Loss-of-function studies in mice are in agreement with the results of this study and support that Hhex is necessary for an adequate cellular proliferation rate in definitive endoderm and liver bud [88-89, 99]. However, some reports have demonstrated that Hhex inhibits Cyclin D1

mRNA transport and subsequent transformation in lymphoma derived cell line U937 [54, 56]. In fact, Hhex overexpression causes G1 arrest in these cells. These results seem to be in disagreement with the results shown in our study, although we believe that such differences may reside in the different nature of the cells. Cultured fibroblasts are non-malignant cells obtained from healthy human tissues. U937 cells were isolated from a malignant histiocytic lymphoma and present a completely disregulated cell cycle control. In fact, when we attempted similar experiments with HeLa cells, a cell line derived from a human cervix cancer, we didn´t observed any increase in cell proliferation.

In summary, in this study we have proposed a model (**figure D2**), by which direct interaction of Hhex with c-Myc oncoprotein would induce c-Myc activity, promoting cellular proliferation and transformation. At the mechanistic level, we suggest that enhanced c-Myc activity induced by Hhex results *Cyclin D1* upregulation that ultimately will increase cell proliferation and transformation in a non-tumourigenic cell type.

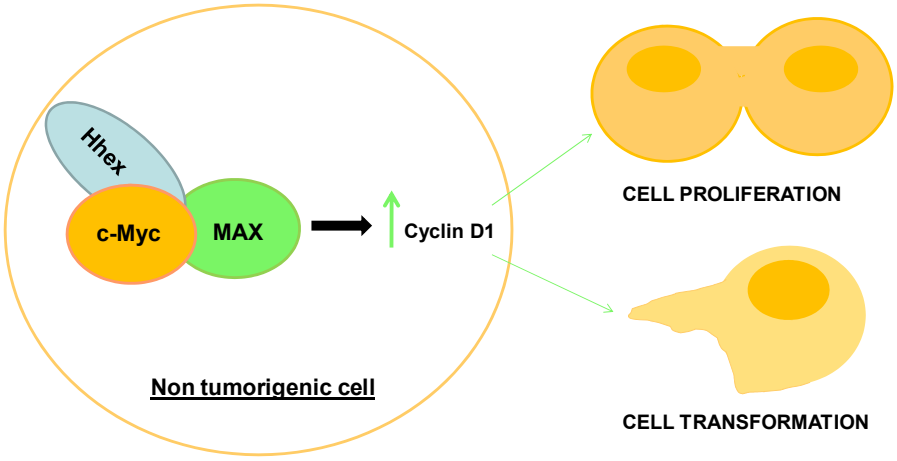


FIGURE D2. Theoretical model depicting Hhex role promoting c-Myc dependent proliferation and transformation.

Interaction between Hhex and c-Myc in non tumorigenic cells would enhance c-Myc activity inducing the expression of Cyclin D1, a positive c-Myc target gene essential for cell cycle transition from G1 to S, resulting in higher cell proliferation and transformation.

CONCLUSIONS

C1. INTERACTION BETWEEN HHEX AND SOX13 MODULATES WNT/TCF ACTIVITY

- SOX13 (fragment corresponding to amino acids 126-218) was identified as a putative Hhex-binding partner in a yeast-two-hybrid screening.
- *SOX13* (encoding a protein of 604 aa) and *S-SOX13* (encoding a protein of 255 aa) are the SOX13 mRNA isoforms prevalently expressed in human tissues.
- Hhex and SOX13 interact directly, in vi*vo* and are co-localized in the cell nuclei. The interaction is mediated by the N-terminal domain of Hhex, which is essential for this interaction to occur. The LZ and Q-rich domain of SOX13 is necessary, but not sufficient for the interaction.
- Using cell lines and cultured mouse embryos, we show that Hhex antagonizes the repressing activity of SOX13 over the Wnt/β–catenin pathway.
- Mechanistically, Hhex interaction with SOX13 withdraws SOX13 from the SOX13•TCF1 complex allowing TCF1 to form back the effector complex in Wnt sensitive promoters.
- Hhex and SOX13 are capable of modulating endogenous Wnt-target genes expression, such as Dickkopf 1 (DKK1,) by altering TCF1 binding to Wnt target promoters.

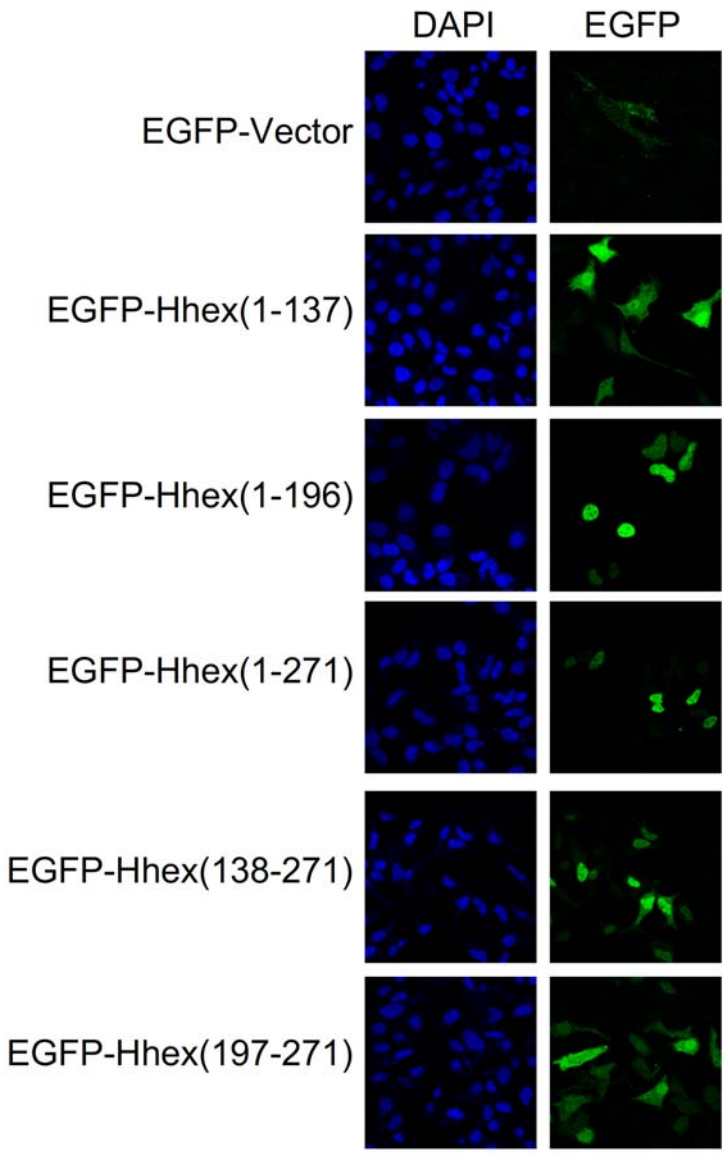
C2. HHEX INTERACTS WITH C-MYC ONCOGENE INDUCING C-MYC FUNCTION

- Hhex increases cell proliferation of human and rat fibroblast cells by a c-Myc dependent mechanism. It also promotes cellular transformation in mouse embryonic fibroblasts.
- Hhex interacts with c-Myc oncoprotein in cell lines and in human liver. Hhex N-terminal domain and c-Myc bHLH-LZ C-terminal domain are involved in the interaction. bHLH-LZ is also the interaction domain of MAX protein, however, Hhex does not interfere with the formation of c-Myc/Max heterodimer.
- Hhex expression in non-hepatic and hepatic cell lines induces c-Myc transcriptional activity on a heterologous reporter.
- We attribute the increase in c-Myc transcriptional activity to a synergistic effect caused by Hhex interaction. In fact, the levels of c-Myc mRNA and protein do not change significantly in gain-of-function (Hhex expression in human fibroblast cells) or loss-of-function (Hex-/- mouse embryonic liver bud) models.
- Hhex regulates Cyclin D1, a c-Myc target gene, at the mRNA and protein levels, in the liver bud region of E9.5 mouse embryo and in human fibroblast cells.

ANNEX

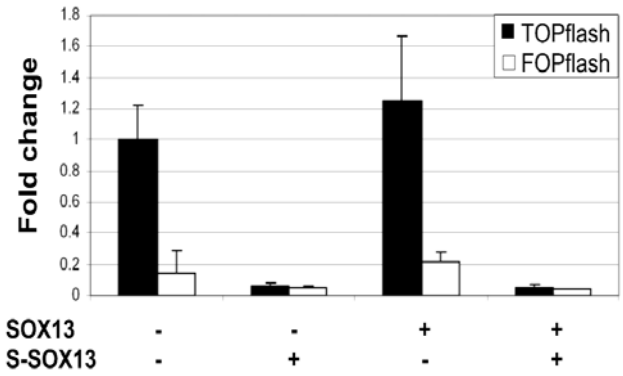
ANNEX

A1. Hhex and Hhex deletion mutants cellular localization



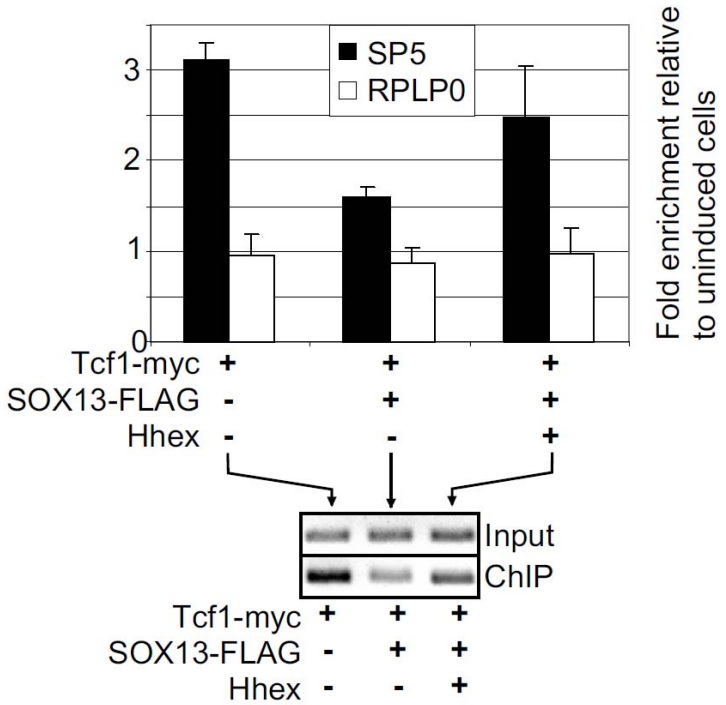
Plasmids expressing EGFPHhex fusion proteins were transfected in Hela cells. Confocal photomicrographs demonstrated localization of EGFP-Hhex and EGFP-Hhex(1-196) in the nucleus. EGFP-Hhex(137-271) was localized mainly in the nucleus but also in the cytoplasm. Images were acquired with a Leica TCS-SP2 confocal microscope.

A2. S-SOX13 does not interfere with inhibition of Wnt mediated by full length SOX13



HEK 293T cells were transfected with the Wnt reporter plasmid TOPflash (black bars) and the mutated version, FOPflash (white bars) as a negative control. An expression vector containing activated β-catenin (S37Y) was always cotransfected for each condition. Plasmids expressing SOX proteins were transfected as indicated. Means and SEM of the relative firefly/renilla ratios and fold change over control of at least three experiments are shown. The total firefly/renilla ratio of untreated cells transfected with activated β-catenin (S37Y) was set as 1.0.

A3. Regulation of SP5, a target of the β-catenin/TCF pathway by Hhex



293T cells were transfected with plasmids expressing TCF1-myc, SOX13 or Hhex, as shown below the graph. Cells were crosslinked and lysed. The relative amounts of TCF1-myc on SP5 promoter was analysed by real-time qPCR of immunoprecipitated chromatin using anti-myc antibody. Binding of TCF1 to the RPLP0 gene was used as a negative control. The image of a representative qPCR stopped at the exponential phase is also shown. Data are the mean of three experiments.

ANNEX

PUBLISHED PAPER

Marfil V, Moya M, Perreux CE, Castell JV, Lemaigre FP, Real FX, et al. Interaction between Hhex and SOX13 modulates Wnt/TCF pathway. J Biol Chem. 2010; 285(8): 5726-37.

BIBLIOGRAPHY

1. Lu, C.C., J. Brennan, and E.J. Robertson, *From fertilization to gastrulation: axis formation in the mouse embryo.* Curr Opin Genet Dev, 2001. **11**(4): p. 384-92.
2. Wells, J.M. and D.A. Melton, *Vertebrate endoderm development.* Annu Rev Cell Dev Biol, 1999. **15**: p. 393-410.
3. Lawson, K.A., J.J. Meneses, and R.A. Pedersen, *Cell fate and cell lineage in the endoderm of the presomite mouse embryo, studied with an intracellular tracer.* Dev Biol, 1986. **115**(2): p. 325-39.
4. Rosenquist, G.C., *The location of the pregut endoderm in the chick embryo at the primitive streak stage as determined by radioautographic mapping.* Dev Biol, 1971. **26**(2): p. 323-35.
5. Tremblay, K.D. and K.S. Zaret, *Distinct populations of endoderm cells converge to generate the embryonic liver bud and ventral foregut tissues.* Dev Biol, 2005. **280**(1): p. 87-99.
6. McLin, V.A., S.A. Rankin, and A.M. Zorn, *Repression of Wnt/beta-catenin signaling in the anterior endoderm is essential for liver and pancreas development.* Development, 2007. **134**(12): p. 2207-17.
7. Li, Y., et al., *Sfrp5 coordinates foregut specification and morphogenesis by antagonizing both canonical and noncanonical Wnt11 signaling.* Genes Dev, 2008. **22**(21): p. 3050-63.
8. Deutsch, G., et al., *A bipotential precursor population for pancreas and liver within the embryonic endoderm.* Development, 2001. **128**(6): p. 871-81.
9. Serls, A.E., et al., *Different thresholds of fibroblast growth factors pattern the ventral foregut into liver and lung.* Development, 2005. **132**(1): p. 35-47.
10. Chalmers, A.D. and J.M. Slack, *The Xenopus tadpole gut: fate maps and morphogenetic movements.* Development, 2000. **127**(2): p. 381-92.
11. Logan, C.Y. and R. Nusse, *The Wnt signaling pathway in development and disease.* Annu Rev Cell Dev Biol, 2004. **20**: p. 781-810.
12. Cadigan, K.M. and R. Nusse, *Wnt signaling: a common theme in animal development.* Genes Dev, 1997. **11**(24): p. 3286-305.
13. Nusse, R., *Wnt signaling in disease and in development.* Cell Res, 2005. **15**(1): p. 28-32.

14. Peifer, M. and P. Polakis, *Wnt signaling in oncogenesis and embryogenesis--a look outside the nucleus.* Science, 2000. **287**(5458): p. 1606-9.
15. Willert, K., et al., *Wnt proteins are lipid-modified and can act as stem cell growth factors.* Nature, 2003. **423**(6938): p. 448-52.
16. Polakis, P., *Wnt signaling and cancer.* Genes Dev, 2000. **14**(15): p. 1837-51.
17. Clevers, H., *Wnt/beta-catenin signaling in development and disease.* Cell, 2006. **127**(3): p. 469-80.
18. Cadigan, K.M. and Y.I. Liu, *Wnt signaling: complexity at the surface.* J Cell Sci, 2006. **119**(Pt 3): p. 395-402.
19. Gualdi, R., et al., *Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control.* Genes Dev, 1996. **10**(13): p. 1670-82.
20. Ang, S.L., et al., *The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/forkhead proteins.* Development, 1993. **119**(4): p. 1301-15.
21. Arceci, R.J., et al., *Mouse GATA-4: a retinoic acid-inducible GATA-binding transcription factor expressed in endodermally derived tissues and heart.* Mol Cell Biol, 1993. **13**(4): p. 2235-46.
22. Houssaint, E., *Differentiation of the mouse hepatic primordium. I. An analysis of tissue interactions in hepatocyte differentiation.* Cell Differ, 1980. **9**(5): p. 269-79.
23. Medlock, E.S. and J.L. Haar, *The liver hemopoietic environment: I. Developing hepatocytes and their role in fetal hemopoiesis.* Anat Rec, 1983. **207**(1): p. 31-41.
24. Douarin, N.M., *An experimental analysis of liver development.* Med Biol, 1975. **53**(6): p. 427-55.
25. Monga, S.P., et al., *Beta-catenin antisense studies in embryonic liver cultures: role in proliferation, apoptosis, and lineage specification.* Gastroenterology, 2003. **124**(1): p. 202-16.
26. Hussain, S.Z., et al., *Wnt impacts growth and differentiation in ex vivo liver development.* Exp Cell Res, 2004. **292**(1): p. 157-69.
27. Micsenyi, A., et al., *Beta-catenin is temporally regulated during normal liver development.* Gastroenterology, 2004. **126**(4): p. 1134-46.
28. Tan, X., et al., *Beta-catenin deletion in hepatoblasts disrupts hepatic morphogenesis and survival during mouse development.* Hepatology, 2008. **47**(5): p. 1667-79.
29. Suksaweang, S., et al., *Morphogenesis of chicken liver: identification of localized growth zones and the role of beta-catenin/Wnt in size regulation.* Dev Biol, 2004. **266**(1): p. 109-22.

30. Enzan, H., et al., *Development of hepatic sinusoidal structure with special reference to the Ito cells.* Microsc Res Tech, 1997. **39**(4): p. 336-49.
31. Gehring, W.J., *The homeobox in perspective.* Trends Biochem Sci, 1992. **17**(8): p. 277-80.
32. Guazzi, S., et al., *The thyroid transcription factor-1 gene is a candidate target for regulation by Hox proteins.* EMBO J, 1994. **13**(14): p. 3339-47.
33. Guazzi, S., et al., *Thyroid nuclear factor 1 (TTF-1) contains a homeodomain and displays a novel DNA binding specificity.* EMBO J, 1990. **9**(11): p. 3631-9.
34. Komuro, I. and S. Izumo, *Csx: a murine homeobox-containing gene specifically expressed in the developing heart.* Proc Natl Acad Sci U S A, 1993. **90**(17): p. 8145-9.
35. Hentsch, B., et al., *Hlx homeo box gene is essential for an inductive tissue interaction that drives expansion of embryonic liver and gut.* Genes Dev, 1996. **10**(1): p. 70-9.
36. Crompton, M.R., et al., *Identification of a novel vertebrate homeobox gene expressed in haematopoietic cells.* Nucleic Acids Res, 1992. **20**(21): p. 5661-7.
37. Ho, C.Y., et al., *A role for the extraembryonic yolk syncytial layer in patterning the zebrafish embryo suggested by properties of the hex gene.* Curr Biol, 1999. **9**(19): p. 1131-4.
38. Newman, C.S., F. Chia, and P.A. Krieg, *The XHex homeobox gene is expressed during development of the vascular endothelium: overexpression leads to an increase in vascular endothelial cell number.* Mech Dev, 1997. **66**(1-2): p. 83-93.
39. Tanaka, T., et al., *cDNA cloning and expression of rat homeobox gene, Hex, and functional characterization of the protein.* Biochem J, 1999. **339 (Pt 1)**: p. 111-7.
40. Morck, C., et al., *pha-2 encodes the C. elegans ortholog of the homeodomain protein HEX and is required for the formation of the pharyngeal isthmus.* Dev Biol, 2004. **272**(2): p. 403-18.
41. Hromas, R., J. Radich, and S. Collins, *PCR cloning of an orphan homeobox gene (PRH) preferentially expressed in myeloid and liver cells.* Biochem Biophys Res Commun, 1993. **195**(2): p. 976-83.
42. Morgutti, M., et al., *Genomic organization and chromosome mapping of the human homeobox gene HHEX.* Cytogenet Cell Genet, 2001. **94**(1-2): p. 30-2.

43. Ghosh, B., et al., *Genomic structure, cDNA mapping, and chromosomal localization of the mouse homeobox gene, Hex.* Mamm Genome, 1999. **10**(10): p. 1023-5.
44. Myint, Z., et al., *Genomic organization and promoter analysis of a mouse homeobox gene, Hex.* J Biochem, 1999. **125**(4): p. 795-802.
45. Bedford, F.K., et al., *HEX: a novel homeobox gene expressed during haematopoiesis and conserved between mouse and human.* Nucleic Acids Res, 1993. **21**(5): p. 1245-9.
46. Guiral, M., et al., *PRH represses transcription in hematopoietic cells by at least two independent mechanisms.* J Biol Chem, 2001. **276**(4): p. 2961-70.
47. Soufi, A., et al., *Oligomerisation of the developmental regulator proline rich homeodomain (PRH/Hex) is mediated by a novel proline-rich dimerisation domain.* J Mol Biol, 2006. **358**(4): p. 943-62.
48. Soufi, A. and P.S. Jayaraman, *PRH/Hex: an oligomeric transcription factor and multifunctional regulator of cell fate.* Biochem J, 2008. **412**(3): p. 399-413.
49. Kasamatsu, S., et al., *Identification of the transactivating region of the homeodomain protein, hex.* J Biochem, 2004. **135**(2): p. 217-23.
50. Puppin, C., et al., *HEX expression and localization in normal mammary gland and breast carcinoma.* BMC Cancer, 2006. **6**: p. 192.
51. Tanaka, H., et al., *Hex stimulates the hepatocyte nuclear factor 1alpha-mediated activation of transcription.* Arch Biochem Biophys, 2005. **442**(1): p. 117-24.
52. Ghosh, B., et al., *Immunocytochemical characterization of murine Hex, a homeobox-containing protein.* Pediatr Res, 2000. **48**(5): p. 634-8.
53. Pellizzari, L., et al., *Expression and function of the homeodomain-containing protein Hex in thyroid cells.* Nucleic Acids Res, 2000. **28**(13): p. 2503-11.
54. Topisirovic, I., et al., *The proline-rich homeodomain protein, PRH, is a tissue-specific inhibitor of eIF4E-dependent cyclin D1 mRNA transport and growth.* EMBO J, 2003. **22**(3): p. 689-703.
55. D'Elia, A.V., et al., *Expression and localization of the homeodomain-containing protein HEX in human thyroid tumors.* J Clin Endocrinol Metab, 2002. **87**(3): p. 1376-83.
56. Topisirovic, I., et al., *Aberrant eukaryotic translation initiation factor 4E-dependent mRNA transport impedes hematopoietic*

differentiation and contributes to leukemogenesis. Mol Cell Biol, 2003. **23**(24): p. 8992-9002.

57. Marfil, V., et al., *Interaction between Hhex and SOX13 modulates Wnt/TCF activity.* J Biol Chem, 2010. **285**(8): p. 5726-37.
58. Melnick, A. and J.D. Licht, *Deconstructing a disease: RARalpha, its fusion partners, and their roles in the pathogenesis of acute promyelocytic leukemia.* Blood, 1999. **93**(10): p. 3167-215.
59. Topcu, Z., et al., *The promyelocytic leukemia protein PML interacts with the proline-rich homeodomain protein PRH: a RING may link hematopoiesis and growth control.* Oncogene, 1999. **18**(50): p. 7091-100.
60. Rousseau, D., et al., *Translation initiation of ornithine decarboxylase and nucleocytoplasmic transport of cyclin D1 mRNA are increased in cells overexpressing eukaryotic initiation factor 4E.* Proc Natl Acad Sci U S A, 1996. **93**(3): p. 1065-70.
61. Schaefer, L.K., S. Wang, and T.S. Schaefer, *Functional interaction of Jun and homeodomain proteins.* J Biol Chem, 2001. **276**(46): p. 43074-82.
62. Bess, K.L., et al., *The transcriptional repressor protein PRH interacts with the proteasome.* Biochem J, 2003. **374**(Pt 3): p. 667-75.
63. Swingler, T.E., et al., *The proline-rich homeodomain protein recruits members of the Groucho/Transducin-like enhancer of split protein family to co-repress transcription in hematopoietic cells.* J Biol Chem, 2004. **279**(33): p. 34938-47.
64. Minami, T., et al., *Interaction between hex and GATA transcription factors in vascular endothelial cells inhibits flk-1/KDR-mediated vascular endothelial growth factor signaling.* J Biol Chem, 2004. **279**(20): p. 20626-35.
65. Oyama, Y., et al., *Homeobox protein Hex facilitates serum responsive factor-mediated activation of the SM22alpha gene transcription in embryonic fibroblasts.* Arterioscler Thromb Vasc Biol, 2004. **24**(9): p. 1602-7.
66. Wissmuller, S., et al., *The high-mobility-group domain of Sox proteins interacts with DNA-binding domains of many transcription factors.* Nucleic Acids Res, 2006. **34**(6): p. 1735-44.
67. Soufi, A., et al., *CK2 phosphorylation of the PRH/Hex homeodomain functions as a reversible switch for DNA binding.* Nucleic Acids Res, 2009.
68. Ploski, J.E., et al., *A mechanism of nucleocytoplasmic trafficking for the homeodomain protein PRH.* Mol Cell Biochem, 2009. **332**(1-2): p. 173-81.

69. Gaston, K. and P.S. Jayaraman, *Transcriptional repression in eukaryotes: repressors and repression mechanisms.* Cell Mol Life Sci, 2003. **60**(4): p. 721-41.
70. Brickman, J.M., et al., *Hex is a transcriptional repressor that contributes to anterior identity and suppresses Spemann organiser function.* Development, 2000. **127**(11): p. 2303-15.
71. Zamparini, A.L., et al., *Hex acts with beta-catenin to regulate anteroposterior patterning via a Groucho-related co-repressor and Nodal.* Development, 2006. **133**(18): p. 3709-22.
72. Denson, L.A., et al., *Divergent homeobox gene hex regulates promoter of the Na(+)-dependent bile acid cotransporter.* Am J Physiol Gastrointest Liver Physiol, 2000. **279**(2): p. G347-55.
73. Nakagawa, T., et al., *HEX acts as a negative regulator of angiogenesis by modulating the expression of angiogenesis-related gene in endothelial cells in vitro.* Arterioscler Thromb Vasc Biol, 2003. **23**(2): p. 231-7.
74. Noy, P., et al., *PRH/Hhex controls cell survival through coordinate transcriptional regulation of vascular endothelial growth factor signaling.* Mol Cell Biol, 2010. **30**(9): p. 2120-34.
75. Cong, R., et al., *Hhex is a direct repressor of endothelial cell-specific molecule 1 (ESM-1).* Biochem Biophys Res Commun, 2006. **346**(2): p. 535-45.
76. Meier, P.J., *Molecular mechanisms of hepatic bile salt transport from sinusoidal blood into bile.* Am J Physiol, 1995. **269**(6 Pt 1): p. G801-12.
77. Thomas, P.Q., A. Brown, and R.S. Beddington, *Hex: a homeobox gene revealing peri-implantation asymmetry in the mouse embryo and an early transient marker of endothelial cell precursors.* Development, 1998. **125**(1): p. 85-94.
78. Bogue, C.W., et al., *Hex expression suggests a role in the development and function of organs derived from foregut endoderm.* Dev Dyn, 2000. **219**(1): p. 84-9.
79. Keng, V.W., et al., *Expression of Hex mRNA in early murine postimplantation embryo development.* FEBS Lett, 1998. **426**(2): p. 183-6.
80. Yatskievych, T.A., S. Pascoe, and P.B. Antin, *Expression of the homebox gene Hex during early stages of chick embryo development.* Mech Dev, 1999. **80**(1): p. 107-9.
81. Obinata, A. and Y. Akimoto, *Involvement of Hex in the initiation of feather morphogenesis.* Int J Dev Biol, 2005. **49**(8): p. 953-60.
82. Obinata, A. and Y. Akimoto, *Expression of Hex during feather bud development.* Int J Dev Biol, 2005. **49**(7): p. 885-90.

83. Obinata, A., et al., *Expression of Hex homeobox gene during skin development: Increase in epidermal cell proliferation by transfecting the Hex to the dermis.* Dev Growth Differ, 2002. **44**(4): p. 281-92.
84. Manfioletti, G., et al., *Differential expression of a novel proline-rich homeobox gene (Prh) in human hematolymphopoietic cells.* Blood, 1995. **85**(5): p. 1237-45.
85. Rodriguez, T.A., et al., *Distinct enhancer elements control Hex expression during gastrulation and early organogenesis.* Dev Biol, 2001. **234**(2): p. 304-16.
86. Zorn, A.M., K. Butler, and J.B. Gurdon, *Anterior endomesoderm specification in Xenopus by Wnt/beta-catenin and TGF-beta signalling pathways.* Dev Biol, 1999. **209**(2): p. 282-97.
87. Bischof, J. and W. Driever, *Regulation of hhex expression in the yolk syncytial layer, the potential Nieuwkoop center homolog in zebrafish.* Dev Biol, 2004. **276**(2): p. 552-62.
88. Martinez Barbera, J.P., et al., *The homeobox gene Hex is required in definitive endodermal tissues for normal forebrain, liver and thyroid formation.* Development, 2000. **127**(11): p. 2433-45.
89. Hallaq, H., et al., *A null mutation of Hhex results in abnormal cardiac development, defective vasculogenesis and elevated Vegfa levels.* Development, 2004. **131**(20): p. 5197-209.
90. Keng, V.W., et al., *Homeobox gene Hex is essential for onset of mouse embryonic liver development and differentiation of the monocyte lineage.* Biochem Biophys Res Commun, 2000. **276**(3): p. 1155-61.
91. Bort, R., et al., *Hex homeobox gene-dependent tissue positioning is required for organogenesis of the ventral pancreas.* Development, 2004. **131**(4): p. 797-806.
92. Hunter, M.P., et al., *The homeobox gene Hhex is essential for proper hepatoblast differentiation and bile duct morphogenesis.* Dev Biol, 2007. **308**(2): p. 355-67.
93. Elsalini, O.A., et al., *Zebrafish hhex, nk2.1a, and pax2.1 regulate thyroid growth and differentiation downstream of Nodal-dependent transcription factors.* Dev Biol, 2003. **263**(1): p. 67-80.
94. Wallace, K.N., et al., *Zebrafish hhex regulates liver development and digestive organ chirality.* Genesis, 2001. **30**(3): p. 141-3.
95. Smithers, L.E. and C.M. Jones, *Xhex-expressing endodermal tissues are essential for anterior patterning in Xenopus.* Mech Dev, 2002. **119**(2): p. 191-200.

96. Mack, D.L., et al., *Down-regulation of the myeloid homeobox protein Hex is essential for normal T-cell development.* Immunology, 2002. **107**(4): p. 444-51.
97. Srinivas, S., et al., *Active cell migration drives the unilateral movements of the anterior visceral endoderm.* Development, 2004. **131**(5): p. 1157-64.
98. Jones, C.M., et al., *An anterior signalling centre in Xenopus revealed by the homeobox gene XHex.* Curr Biol, 1999. **9**(17): p. 946-54.
99. Bort, R., et al., *Hex homeobox gene controls the transition of the endoderm to a pseudostratified, cell emergent epithelium for liver bud development.* Dev Biol, 2006. **290**(1): p. 44-56.
100. Goodrich, L.V., et al., *Altered neural cell fates and medulloblastoma in mouse patched mutants.* Science, 1997. **277**(5329): p. 1109-13.
101. Parviz, F., et al., *Hepatocyte nuclear factor 4alpha controls the development of a hepatic epithelium and liver morphogenesis.* Nat Genet, 2003. **34**(3): p. 292-6.
102. Slack, J.M., *Developmental biology of the pancreas.* Development, 1995. **121**(6): p. 1569-80.
103. Sladek, R., et al., *A genome-wide association study identifies novel risk loci for type 2 diabetes.* Nature, 2007. **445**(7130): p. 881-5.
104. Scott, L.J., et al., *A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants.* Science, 2007. **316**(5829): p. 1341-5.
105. Zeggini, E., et al., *Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes.* Science, 2007. **316**(5829): p. 1336-41.
106. Freeman, H. and R.D. Cox, *Type-2 diabetes: a cocktail of genetic discovery.* Hum Mol Genet, 2006. **15 Spec No 2**: p. R202-9.
107. Lee, Y.H., et al., *Association between polymorphisms in SLC30A8, HHEX, CDKN2A/B, IGF2BP2, FTO, WFS1, CDKAL1, KCNQ1 and type 2 diabetes in the Korean population.* J Hum Genet, 2008. **53**(11-12): p. 991-8.
108. Wu, Y., et al., *Common variants in CDKAL1, CDKN2A/B, IGF2BP2, SLC30A8, and HHEX/IDE genes are associated with type 2 diabetes and impaired fasting glucose in a Chinese Han population.* Diabetes, 2008. **57**(10): p. 2834-42.
109. Ng, M.C., et al., *Implication of genetic variants near TCF7L2, SLC30A8, HHEX, CDKAL1, CDKN2A/B, IGF2BP2, and FTO in type 2*

diabetes and obesity in 6,719 Asians. Diabetes, 2008. **57**(8): p. 2226-33.

110. Hertel, J.K., et al., *Genetic analysis of recently identified type 2 diabetes loci in 1,638 unselected patients with type 2 diabetes and 1,858 control participants from a Norwegian population-based cohort (the HUNT study).* Diabetologia, 2008. **51**(6): p. 971-7.
111. Staiger, H., et al., *Polymorphisms within novel risk loci for type 2 diabetes determine beta-cell function.* PLoS ONE, 2007. **2**(9): p. e832.
112. Staiger, H., et al., *A candidate type 2 diabetes polymorphism near the HHEX locus affects acute glucose-stimulated insulin release in European populations: results from the EUGENE2 study.* Diabetes, 2008. **57**(2): p. 514-7.
113. Pivovarova, O., et al., *The influence of genetic variations in HHEX gene on insulin metabolism in the German MESYBEPO cohort.* Diabetes Metab Res Rev, 2009. **25**(2): p. 156-62.
114. Bowles, J., G. Schepers, and P. Koopman, *Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators.* Dev Biol, 2000. **227**(2): p. 239-55.
115. Wilson, M. and P. Koopman, *Matching SOX: partner proteins and co-factors of the SOX family of transcriptional regulators.* Curr Opin Genet Dev, 2002. **12**(4): p. 441-6.
116. Kamachi, Y., M. Uchikawa, and H. Kondoh, *Pairing SOX off: with partners in the regulation of embryonic development.* Trends Genet, 2000. **16**(4): p. 182-7.
117. Gure, A.O., et al., *Serological identification of embryonic neural proteins as highly immunogenic tumor antigens in small cell lung cancer.* Proc Natl Acad Sci U S A, 2000. **97**(8): p. 4198-203.
118. Chen, Y., et al., *The molecular mechanism governing the oncogenic potential of SOX2 in breast cancer.* J Biol Chem, 2008. **283**(26): p. 17969-78.
119. Dong, C., D. Wilhelm, and P. Koopman, *Sox genes and cancer.* Cytogenet Genome Res, 2004. **105**(2-4): p. 442-7.
120. Comtesse, N., et al., *Complex humoral immune response against a benign tumor: frequent antibody response against specific antigens as diagnostic targets.* Proc Natl Acad Sci U S A, 2005. **102**(27): p. 9601-6.
121. Kasimiotis, H., et al., *Sex-determining region Y-related protein SOX13 is a diabetes autoantigen expressed in pancreatic islets.* Diabetes, 2000. **49**(4): p. 555-61.

122. Argentaro, A., et al., *Linkage studies of SOX13, the ICA12 autoantigen gene, in families with type 1 diabetes.* Mol Genet Metab, 2001. **72**(4): p. 356-9.
123. Kasimiotis, H., et al., *Antibodies to SOX13 (ICA12) are associated with type 1 diabetes.* Autoimmunity, 2001. **33**(2): p. 95-101.
124. Steinbrenner, H., et al., *Autoantibodies to ICA12 (SOX-13) are not specific for Type I diabetes.* Diabetologia, 2000. **43**(11): p. 1381-4.
125. Lampasona, V., et al., *ICA12(SOX13) autoantibodies are unlikely to be a useful marker for pre-clinical Type I diabetes.* Diabetologia, 2001. **44**(2): p. 267.
126. Roose, J., et al., *The Sox-13 gene: structure, promoter characterization, and chromosomal localization.* Genomics, 1999. **57**(2): p. 301-5.
127. Ovcharenko, I., et al., *ECR Browser: a tool for visualizing and accessing data from comparisons of multiple vertebrate genomes.* Nucleic Acids Res, 2004. **32**(Web Server issue): p. W280-6.
128. Argentaro, A., et al., *Genomic characterisation and fine mapping of the human SOX13 gene.* Gene, 2000. **250**(1-2): p. 181-9.
129. Harley, V.R., et al., *DNA binding activity of recombinant SRY from normal males and XY females.* Science, 1992. **255**(5043): p. 453-6.
130. Harley, V.R., R. Lovell-Badge, and P.N. Goodfellow, *Definition of a consensus DNA binding site for SRY.* Nucleic Acids Res, 1994. **22**(8): p. 1500-1.
131. van de Wetering, M., et al., *Sox-4, an Sry-like HMG box protein, is a transcriptional activator in lymphocytes.* EMBO J, 1993. **12**(10): p. 3847-54.
132. Sumi, E., et al., *SRY-related HMG box 9 regulates the expression of Col4a2 through transactivating its enhancer element in mesangial cells.* Am J Pathol, 2007. **170**(6): p. 1854-64.
133. Han, Y. and V. Lefebvre, *L-Sox5 and Sox6 drive expression of the aggrecan gene in cartilage by securing binding of Sox9 to a far-upstream enhancer.* Mol Cell Biol, 2008. **28**(16): p. 4999-5013.
134. Zabel, U., R. Schreck, and P.A. Baeuerle, *DNA binding of purified transcription factor NF-kappa B. Affinity, specificity, Zn2+ dependence, and differential half-site recognition.* J Biol Chem, 1991. **266**(1): p. 252-60.
135. Melichar, H.J., et al., *Regulation of gammadelta versus alphabeta T lymphocyte differentiation by the transcription factor SOX13.* Science, 2007. **315**(5809): p. 230-3.

136. Lioubinski, O., et al., *Expression of Sox transcription factors in the developing mouse pancreas.* Dev Dyn, 2003. **227**(3): p. 402-8.
137. Wang, Y., S. Ristevski, and V.R. Harley, *SOX13 exhibits a distinct spatial and temporal expression pattern during chondrogenesis, neurogenesis, and limb development.* J Histochem Cytochem, 2006. **54**(12): p. 1327-33.
138. Smits, P., et al., *The transcription factors L-Sox5 and Sox6 are essential for cartilage formation.* Dev Cell, 2001. **1**(2): p. 277-90.
139. Wang, Y., S. Bagheri-Fam, and V.R. Harley, *SOX13 is up-regulated in the developing mouse neuroepithelium and identifies a sub-population of differentiating neurons.* Brain Res Dev Brain Res, 2005. **157**(2): p. 201-8.
140. Pla, P., et al., *Identification of Phox2b-regulated genes by expression profiling of cranial motoneuron precursors.* Neural Dev, 2008. **3**: p. 14.
141. Payne, G.S., et al., *Analysis of avian leukosis virus DNA and RNA in bursal tumours: viral gene expression is not required for maintenance of the tumor state.* Cell, 1981. **23**(2): p. 311-22.
142. Payne, G.S., J.M. Bishop, and H.E. Varmus, *Multiple arrangements of viral DNA and an activated host oncogene in bursal lymphomas.* Nature, 1982. **295**(5846): p. 209-14.
143. Neel, B.G., et al., *Avian leukosis virus-induced tumors have common proviral integration sites and synthesize discrete new RNAs: oncogenesis by promoter insertion.* Cell, 1981. **23**(2): p. 323-34.
144. Hayward, W.S., et al., *Avian lymphoid leukosis is correlated with the appearance of discrete new RNAs containing viral and cellular genetic information.* Haematol Blood Transfus, 1981. **26**: p. 439-44.
145. Hann, S.R., et al., *A non-AUG translational initiation in c-myc exon 1 generates an N-terminally distinct protein whose synthesis is disrupted in Burkitt's lymphomas.* Cell, 1988. **52**(2): p. 185-95.
146. Blackwood, E.M. and R.N. Eisenman, *Max: a helix-loop-helix zipper protein that forms a sequence-specific DNA-binding complex with Myc.* Science, 1991. **251**(4998): p. 1211-7.
147. Blackwood, E.M., B. Luscher, and R.N. Eisenman, *Myc and Max associate in vivo.* Genes Dev, 1992. **6**(1): p. 71-80.
148. Gu, W., et al., *Binding and suppression of the Myc transcriptional activation domain by p107.* Science, 1994. **264**(5156): p. 251-4.
149. Hoang, A.T., et al., *A link between increased transforming activity of lymphoma-derived MYC mutant alleles, their defective*

regulation by p107, and altered phosphorylation of the c-Myc transactivation domain. Mol Cell Biol, 1995. **15**(8): p. 4031-42.

150. Mao, D.Y., et al., *Promoter-binding and repression of PDGFRB by c-Myc are separable activities.* Nucleic Acids Res, 2004. **32**(11): p. 3462-8.
151. Kleine-Kohlbrecher, D., S. Adhikary, and M. Eilers, *Mechanisms of transcriptional repression by Myc.* Curr Top Microbiol Immunol, 2006. **302**: p. 51-62.
152. Adhikary, S., et al., *The ubiquitin ligase HectH9 regulates transcriptional activation by Myc and is essential for tumor cell proliferation.* Cell, 2005. **123**(3): p. 409-21.
153. Gartel, A.L., et al., *Myc represses the p21(WAF1/CIP1) promoter and interacts with Sp1/Sp3.* Proc Natl Acad Sci U S A, 2001. **98**(8): p. 4510-5.
154. Izumi, H., et al., *Mechanism for the transcriptional repression by c-Myc on PDGF beta-receptor.* J Cell Sci, 2001. **114**(Pt 8): p. 1533-44.
155. Shrivastava, A., et al., *Inhibition of transcriptional regulator Yin-Yang-1 by association with c-Myc.* Science, 1993. **262**(5141): p. 1889-92.
156. Li, Z., D. Boone, and S.R. Hann, *Nucleophosmin interacts directly with c-Myc and controls c-Myc-induced hyperproliferation and transformation.* Proc Natl Acad Sci U S A, 2008. **105**(48): p. 18794-9.
157. Patel, J.H., et al., *The c-MYC oncoprotein is a substrate of the acetyltransferases hGCN5/PCAF and TIP60.* Mol Cell Biol, 2004. **24**(24): p. 10826-34.
158. Fernandez, P.C., et al., *Genomic targets of the human c-Myc protein.* Genes Dev, 2003. **17**(9): p. 1115-29.
159. Baudino, T.A., et al., *c-Myc is essential for vasculogenesis and angiogenesis during development and tumor progression.* Genes Dev, 2002. **16**(19): p. 2530-43.
160. Cutry, A.F., et al., *Induction of c-fos and c-myc proto-oncogene expression by epidermal growth factor and transforming growth factor alpha is calcium-independent.* J Biol Chem, 1989. **264**(33): p. 19700-5.
161. Dean, M., et al., *Regulation of c-myc transcription and mRNA abundance by serum growth factors and cell contact.* J Biol Chem, 1986. **261**(20): p. 9161-6.
162. Kelly, K., et al., *Cell-specific regulation of the c-myc gene by lymphocyte mitogens and platelet-derived growth factor.* Cell, 1983. **35**(3 Pt 2): p. 603-10.

163. Barone, M.V. and S.A. Courtneidge, *Myc but not Fos rescue of PDGF signalling block caused by kinase-inactive Src.* Nature, 1995. **378**(6556): p. 509-12.
164. Cleveland, J.L., et al., *Tyrosine kinase oncogenes abrogate interleukin-3 dependence of murine myeloid cells through signaling pathways involving c-myc: conditional regulation of c-myc transcription by temperature-sensitive v-abl.* Mol Cell Biol, 1989. **9**(12): p. 5685-95.
165. Coughlin, S.R., et al., *c-myc gene expression is stimulated by agents that activate protein kinase C and does not account for the mitogenic effect of PDGF.* Cell, 1985. **43**(1): p. 243-51.
166. Mudryj, M., S.W. Hiebert, and J.R. Nevins, *A role for the adenovirus inducible E2F transcription factor in a proliferation dependent signal transduction pathway.* EMBO J, 1990. **9**(7): p. 2179-84.
167. Ran, W., et al., *Induction of c-fos and c-myc mRNA by epidermal growth factor or calcium ionophore is cAMP dependent.* Proc Natl Acad Sci U S A, 1986. **83**(21): p. 8216-20.
168. Roussel, M.F., et al., *Myc rescue of a mutant CSF-1 receptor impaired in mitogenic signalling.* Nature, 1991. **353**(6342): p. 361-3.
169. Roussel, M.F., et al., *Dual control of myc expression through a single DNA binding site targeted by ets family proteins and E2F-1.* Oncogene, 1994. **9**(2): p. 405-15.
170. Wong, K.K., et al., *v-Abl activates c-myc transcription through the E2F site.* Mol Cell Biol, 1995. **15**(12): p. 6535-44.
171. Norbury, C. and P. Nurse, *Animal cell cycles and their control.* Annu Rev Biochem, 1992. **61**: p. 441-70.
172. Vermeulen, K., D.R. Van Bockstaele, and Z.N. Berneman, *The cell cycle: a review of regulation, deregulation and therapeutic targets in cancer.* Cell Prolif, 2003. **36**(3): p. 131-49.
173. Matsushime, H., et al., *Colony-stimulating factor 1 regulates novel cyclins during the G1 phase of the cell cycle.* Cell, 1991. **65**(4): p. 701-13.
174. Sherr, C.J., *D-type cyclins.* Trends Biochem Sci, 1995. **20**(5): p. 187-90.
175. Won, K.A., et al., *Growth-regulated expression of D-type cyclin genes in human diploid fibroblasts.* Proc Natl Acad Sci U S A, 1992. **89**(20): p. 9910-4.
176. Bates, S., et al., *CDK6 (PLSTIRE) and CDK4 (PSK-J3) are a distinct subset of the cyclin-dependent kinases that associate with cyclin D1.* Oncogene, 1994. **9**(1): p. 71-9.

177. Beijersbergen, R.L. and R. Bernards, *Cell cycle regulation by the retinoblastoma family of growth inhibitory proteins.* Biochim Biophys Acta, 1996. **1287**(2-3): p. 103-20.
178. Beijersbergen, R.L., et al., *Regulation of the retinoblastoma protein-related p107 by G1 cyclin complexes.* Genes Dev, 1995. **9**(11): p. 1340-53.
179. Xiao, Z.X., et al., *Regulation of the retinoblastoma protein-related protein p107 by G1 cyclin-associated kinases.* Proc Natl Acad Sci U S A, 1996. **93**(10): p. 4633-7.
180. Chellappan, S.P., et al., *The E2F transcription factor is a cellular target for the RB protein.* Cell, 1991. **65**(6): p. 1053-61.
181. Botz, J., et al., *Cell cycle regulation of the murine cyclin E gene depends on an E2F binding site in the promoter.* Mol Cell Biol, 1996. **16**(7): p. 3401-9.
182. Dou, Q.P., et al., *G1/S-regulated E2F-containing protein complexes bind to the mouse thymidine kinase gene promoter.* J Biol Chem, 1994. **269**(2): p. 1306-13.
183. Geng, Y., et al., *Regulation of cyclin E transcription by E2Fs and retinoblastoma protein.* Oncogene, 1996. **12**(6): p. 1173-80.
184. Dulic, V., E. Lees, and S.I. Reed, *Association of human cyclin E with a periodic G1-S phase protein kinase.* Science, 1992. **257**(5078): p. 1958-61.
185. Girard, F., et al., *Cyclin A is required for the onset of DNA replication in mammalian fibroblasts.* Cell, 1991. **67**(6): p. 1169-79.
186. Walker, D.H. and J.L. Maller, *Role for cyclin A in the dependence of mitosis on completion of DNA replication.* Nature, 1991. **354**(6351): p. 314-7.
187. King, R.W., P.K. Jackson, and M.W. Kirschner, *Mitosis in transition.* Cell, 1994. **79**(4): p. 563-71.
188. Sherr, C.J. and J.M. Roberts, *Inhibitors of mammalian G1 cyclin-dependent kinases.* Genes Dev, 1995. **9**(10): p. 1149-63.
189. Harper, J.W., et al., *The p21 Cdk-interacting protein Cip1 is a potent inhibitor of G1 cyclin-dependent kinases.* Cell, 1993. **75**(4): p. 805-16.
190. Lee, M.H., I. Reynisdottir, and J. Massague, *Cloning of p57KIP2, a cyclin-dependent kinase inhibitor with unique domain structure and tissue distribution.* Genes Dev, 1995. **9**(6): p. 639-49.
191. Polyak, K., et al., *Cloning of p27Kip1, a cyclin-dependent kinase inhibitor and a potential mediator of extracellular antimitogenic signals.* Cell, 1994. **78**(1): p. 59-66.

192. Polyak, K., et al., *p27Kip1, a cyclin-Cdk inhibitor, links transforming growth factor-beta and contact inhibition to cell cycle arrest.* Genes Dev, 1994. **8**(1): p. 9-22.
193. Xiong, Y., et al., *p21 is a universal inhibitor of cyclin kinases.* Nature, 1993. **366**(6456): p. 701-4.
194. Chan, F.K., et al., *Identification of human and mouse p19, a novel CDK4 and CDK6 inhibitor with homology to p16ink4.* Mol Cell Biol, 1995. **15**(5): p. 2682-8.
195. Guan, K.L., et al., *Growth suppression by p18, a p16INK4/MTS1- and p14INK4B/MTS2-related CDK6 inhibitor, correlates with wild-type pRb function.* Genes Dev, 1994. **8**(24): p. 2939-52.
196. Hannon, G.J. and D. Beach, *p15INK4B is a potential effector of TGF-beta-induced cell cycle arrest.* Nature, 1994. **371**(6494): p. 257-61.
197. Hirai, H., et al., *Novel INK4 proteins, p19 and p18, are specific inhibitors of the cyclin D-dependent kinases CDK4 and CDK6.* Mol Cell Biol, 1995. **15**(5): p. 2672-81.
198. Serrano, M., G.J. Hannon, and D. Beach, *A new regulatory motif in cell-cycle control causing specific inhibition of cyclin D/CDK4.* Nature, 1993. **366**(6456): p. 704-7.
199. Gu, Y., J. Rosenblatt, and D.O. Morgan, *Cell cycle regulation of CDK2 activity by phosphorylation of Thr160 and Tyr15.* EMBO J, 1992. **11**(11): p. 3995-4005.
200. Kato, J.Y., et al., *Regulation of cyclin D-dependent kinase 4 (cdk4) by cdk4-activating kinase.* Mol Cell Biol, 1994. **14**(4): p. 2713-21.
201. Honda, R., et al., *Dephosphorylation of human p34cdc2 kinase on both Thr-14 and Tyr-15 by human cdc25B phosphatase.* FEBS Lett, 1993. **318**(3): p. 331-4.
202. Obaya, A.J., M.K. Mateyak, and J.M. Sedivy, *Mysterious liaisons: the relationship between c-Myc and the cell cycle.* Oncogene, 1999. **18**(19): p. 2934-41.
203. Bouchard, C., et al., *Direct induction of cyclin D2 by Myc contributes to cell cycle progression and sequestration of p27.* EMBO J, 1999. **18**(19): p. 5321-33.
204. Hermeking, H., et al., *Identification of CDK4 as a target of c-MYC.* Proc Natl Acad Sci U S A, 2000. **97**(5): p. 2229-34.
205. Daksis, J.I., et al., *Myc induces cyclin D1 expression in the absence of de novo protein synthesis and links mitogen-stimulated signal transduction to the cell cycle.* Oncogene, 1994. **9**(12): p. 3635-45.
206. Steiner, P., et al., *Identification of a Myc-dependent step during the formation of active G1 cyclin-cdk complexes.* EMBO J, 1995. **14**(19): p. 4814-26.

207. Berns, K., E.M. Hijmans, and R. Bernards, *Repression of c-Myc responsive genes in cycling cells causes G1 arrest through reduction of cyclin E/CDK2 kinase activity.* Oncogene, 1997. **15**(11): p. 1347-56.
208. O'Hagan, R.C., et al., *Myc-enhanced expression of Cul1 promotes ubiquitin-dependent proteolysis and cell cycle progression.* Genes Dev, 2000. **14**(17): p. 2185-91.
209. Coller, H.A., et al., *Expression analysis with oligonucleotide microarrays reveals that MYC regulates genes involved in growth, cell cycle, signaling, and adhesion.* Proc Natl Acad Sci U S A, 2000. **97**(7): p. 3260-5.
210. Schuhmacher, M., et al., *Control of cell growth by c-Myc in the absence of cell division.* Curr Biol, 1999. **9**(21): p. 1255-8.
211. Iritani, B.M. and R.N. Eisenman, *c-Myc enhances protein synthesis and cell size during B lymphocyte development.* Proc Natl Acad Sci U S A, 1999. **96**(23): p. 13180-5.
212. Dang, C.V., *c-Myc target genes involved in cell growth, apoptosis, and metabolism.* Mol Cell Biol, 1999. **19**(1): p. 1-11.
213. Bianchi Scarra, G.L., et al., *Terminal erythroid differentiation in the K-562 cell line by 1-beta-D-arabinofuranosylcytosine: accompaniment by c-myc messenger RNA decrease.* Cancer Res, 1986. **46**(12 Pt 1): p. 6327-32.
214. Gonda, T.J. and D. Metcalf, *Expression of myb, myc and fos proto-oncogenes during the differentiation of a murine myeloid leukaemia.* Nature, 1984. **310**(5974): p. 249-51.
215. Gowda, S.D., R.D. Koler, and G.C. Bagby, Jr., *Regulation of C-myc expression during growth and differentiation of normal and leukemic human myeloid progenitor cells.* J Clin Invest, 1986. **77**(1): p. 271-8.
216. Griep, A.E. and H.F. DeLuca, *Decreased c-myc expression is an early event in retinoic acid-induced differentiation of F9 teratocarcinoma cells.* Proc Natl Acad Sci U S A, 1986. **83**(15): p. 5539-43.
217. Lachman, H.M. and A.I. Skoultchi, *Expression of c-myc changes during differentiation of mouse erythroleukaemia cells.* Nature, 1984. **310**(5978): p. 592-4.
218. Gandarillas, A. and F.M. Watt, *c-Myc promotes differentiation of human epidermal stem cells.* Genes Dev, 1997. **11**(21): p. 2869-82.
219. Prendergast, G.C., *Mechanisms of apoptosis by c-Myc.* Oncogene, 1999. **18**(19): p. 2967-87.

220. Moreno, E. and K. Basler, *dMyc transforms cells into super-competitors.* Cell, 2004. **117**(1): p. 117-29.
221. de la Cova, C., et al., *Drosophila myc regulates organ size by inducing cell competition.* Cell, 2004. **117**(1): p. 107-16.
222. Field, J.K. and D.A. Spandidos, *The role of ras and myc oncogenes in human solid tumours and their relevance in diagnosis and prognosis (review).* Anticancer Res, 1990. **10**(1): p. 1-22.
223. Garte, S.J., *The c-myc oncogene in tumor progression.* Crit Rev Oncog, 1993. **4**(4): p. 435-49.
224. Pelengaris, S., M. Khan, and G. Evan, *c-MYC: more than just a matter of life and death.* Nat Rev Cancer, 2002. **2**(10): p. 764-76.
225. McGrory, W.J., D.S. Bautista, and F.L. Graham, *A simple technique for the rescue of early region I mutations into infectious human adenovirus type 5.* Virology, 1988. **163**(2): p. 614-7.
226. Gomez-Foix, A.M., et al., *Adenovirus-mediated transfer of the muscle glycogen phosphorylase gene into hepatocytes confers altered regulation of glycogen metabolism.* J Biol Chem, 1992. **267**(35): p. 25129-34.
227. Stegmeier, F., et al., *A lentiviral microRNA-based system for single-copy polymerase II-regulated RNA interference in mammalian cells.* Proc Natl Acad Sci U S A, 2005. **102**(37): p. 13212-7.
228. Jover, R., et al., *Re-expression of C/EBP alpha induces CYP2B6, CYP2C9 and CYP2D6 genes in HepG2 cells.* FEBS Lett, 1998. **431**(2): p. 227-30.
229. Thompson, B.A., et al., *CHD8 is an ATP-dependent chromatin remodeling factor that regulates beta-catenin target genes.* Mol Cell Biol, 2008. **28**(12): p. 3894-904.
230. Valenta, T., et al., *HIC1 attenuates Wnt signaling by recruitment of TCF-4 and beta-catenin to the nuclear bodies.* EMBO J, 2006. **25**(11): p. 2326-37.
231. Hollenberg, S.M., et al., *Identification of a new family of tissue-specific basic helix-loop-helix proteins with a two-hybrid system.* Mol Cell Biol, 1995. **15**(7): p. 3813-22.
232. Niida, A., et al., *DKK1, a negative regulator of Wnt signaling, is a target of the beta-catenin/TCF pathway.* Oncogene, 2004. **23**(52): p. 8520-6.
233. Weinmann, A.S., et al., *Use of chromatin immunoprecipitation to clone novel E2F target promoters.* Mol Cell Biol, 2001. **21**(20): p. 6820-32.
234. Martinez-Jimenez, C.P., et al., *Transcriptional activation of CYP2C9, CYP1A1, and CYP1A2 by hepatocyte nuclear factor*

4alpha requires coactivators peroxisomal proliferator activated receptor-gamma coactivator 1alpha and steroid receptor coactivator 1. Mol Pharmacol, 2006. **70**(5): p. 1681-92.

235. Pierreux, C.E., et al., *Gene transfer into mouse prepancreatic endoderm by whole embryo electroporation.* JOP, 2005. **6**(2): p. 128-35.
236. Rossi, J.M., et al., *Distinct mesodermal signals, including BMPs from the septum transversum mesenchyme, are required in combination for hepatogenesis from the endoderm.* Genes Dev, 2001. **15**(15): p. 1998-2009.
237. Tsou, A.P., et al., *Overexpression of a novel imprinted gene, PEG10, in human hepatocellular carcinoma and in regenerating mouse livers.* J Biomed Sci, 2003. **10**(6 Pt 1): p. 625-35.
238. Loitsch, S.M., Y. Shastri, and J. Stein, *Stool test for colorectal cancer screening--it's time to move!* Clin Lab, 2008. **54**(11-12): p. 473-84.
239. Hathurusinghe, H.R., K.S. Goonetilleke, and A.K. Siriwardena, *Current status of tumor M2 pyruvate kinase (tumor M2-PK) as a biomarker of gastrointestinal malignancy.* Ann Surg Oncol, 2007. **14**(10): p. 2714-20.
240. Schreiber, J., E. Sock, and M. Wegner, *The regulator of early gliogenesis glial cells missing is a transcription factor with a novel type of DNA-binding domain.* Proc Natl Acad Sci U S A, 1997. **94**(9): p. 4739-44.
241. Korinek, V., et al., *Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma.* Science, 1997. **275**(5307): p. 1784-7.
242. Cairo, S., et al., *PML interacts with Myc, and Myc target gene expression is altered in PML-null fibroblasts.* Oncogene, 2005. **24**(13): p. 2195-203.
243. Foster, L.J., et al., *A mammalian organelle map by protein correlation profiling.* Cell, 2006. **125**(1): p. 187-99.
244. Trumpp, A., et al., *c-Myc regulates mammalian body size by controlling cell number but not cell size.* Nature, 2001. **414**(6865): p. 768-73.
245. Mateyak, M.K., et al., *Phenotypes of c-Myc-deficient rat fibroblasts isolated by targeted homologous recombination.* Cell Growth Differ, 1997. **8**(10): p. 1039-48.
246. Durst, M., et al., *Papillomavirus sequences integrate near cellular oncogenes in some cervical carcinomas.* Proc Natl Acad Sci U S A, 1987. **84**(4): p. 1070-4.

247. Odom, D.T., et al., *Control of pancreas and liver gene expression by HNF transcription factors.* Science, 2004. **303**(5662): p. 1378-81.
248. Kelley-Loughnane, N., et al., *Independent and overlapping transcriptional activation during liver development and regeneration in mice.* Hepatology, 2002. **35**(3): p. 525-34.
249. Petkov, P.M., et al., *Gene expression pattern in hepatic stem/progenitor cells during rat fetal development using complementary DNA microarrays.* Hepatology, 2004. **39**(3): p. 617-27.
250. Menssen, A. and H. Hermeking, *Characterization of the c-MYC-regulated transcriptome by SAGE: identification and analysis of c-MYC target genes.* Proc Natl Acad Sci U S A, 2002. **99**(9): p. 6274-9.
251. Watson, J.D., et al., *Identifying genes regulated in a Myc-dependent manner.* J Biol Chem, 2002. **277**(40): p. 36921-30.
252. Schuldiner, O., S. Shor, and N. Benvenisty, *A computerized database-scan to identify c-MYC targets.* Gene, 2002. **292**(1-2): p. 91-9.
253. Zeller, K.I., et al., *An integrated database of genes responsive to the Myc oncogenic transcription factor: identification of direct genomic targets.* Genome Biol, 2003. **4**(10): p. R69.
254. Kanai-Azuma, M., et al., *Depletion of definitive gut endoderm in Sox17-null mutant mice.* Development, 2002. **129**(10): p. 2367-79.
255. Seymour, P.A., et al., *SOX9 is required for maintenance of the pancreatic progenitor cell pool.* Proc Natl Acad Sci U S A, 2007. **104**(6): p. 1865-70.
256. Antoniou, A., et al., *Intrahepatic bile ducts develop according to a new mode of tubulogenesis regulated by the transcription factor SOX9.* Gastroenterology, 2009. **136**(7): p. 2325-33.
257. Wilson, M.E., et al., *The HMG box transcription factor Sox4 contributes to the development of the endocrine pancreas.* Diabetes, 2005. **54**(12): p. 3402-9.
258. Kormish, J.D., D. Sinner, and A.M. Zorn, *Interactions between SOX factors and Wnt/beta-catenin signaling in development and disease.* Dev Dyn, 2010. **239**(1): p. 56-68.
259. Guo, L., et al., *Sox7 Is an independent checkpoint for beta-catenin function in prostate and colon epithelial cells.* Mol Cancer Res, 2008. **6**(9): p. 1421-30.

260. Zhang, Y., et al., *SOX7, down-regulated in colorectal cancer, induces apoptosis and inhibits proliferation of colorectal cancer cells.* Cancer Lett, 2009. **277**(1): p. 29-37.
261. Reichling, T., et al., *Transcriptional profiles of intestinal tumors in Apc(Min) mice are unique from those of embryonic intestine and identify novel gene targets dysregulated in human colorectal tumors.* Cancer Res, 2005. **65**(1): p. 166-76.
262. Aaboe, M., et al., *SOX4 expression in bladder carcinoma: clinical aspects and in vitro functional characterization.* Cancer Res, 2006. **66**(7): p. 3434-42.
263. de Bont, J.M., et al., *Differential expression and prognostic significance of SOX genes in pediatric medulloblastoma and ependymoma identified by microarray analysis.* Neuro Oncol, 2008. **10**(5): p. 648-60.
264. Scharer, C.D., et al., *Genome-wide promoter analysis of the SOX4 transcriptional network in prostate cancer cells.* Cancer Res, 2009. **69**(2): p. 709-17.
265. Andersen, C.L., et al., *Dysregulation of the transcription factors SOX4, CBFB and SMARCC1 correlates with outcome of colorectal cancer.* Br J Cancer, 2009. **100**(3): p. 511-23.
266. Sinner, D., et al., *Sox17 and Sox4 differentially regulate beta-catenin/T-cell factor activity and proliferation of colon carcinoma cells.* Mol Cell Biol, 2007. **27**(22): p. 7802-15.
267. Fagotto, F., K. Guger, and B.M. Gumbiner, *Induction of the primary dorsalizing center in Xenopus by the Wnt/GSK/beta-catenin signaling pathway, but not by Vg1, Activin or Noggin.* Development, 1997. **124**(2): p. 453-60.
268. Huelsken, J., et al., *Requirement for beta-catenin in anterior-posterior axis formation in mice.* J Cell Biol, 2000. **148**(3): p. 567-78.
269. Foley, A.C. and M. Mercola, *Heart induction by Wnt antagonists depends on the homeodomain transcription factor Hex.* Genes Dev, 2005. **19**(3): p. 387-96.
270. Ober, E.A., et al., *Mesodermal Wnt2b signalling positively regulates liver specification.* Nature, 2006. **442**(7103): p. 688-91.
271. Kubo, A., et al., *The homeobox gene Hex regulates hepatocyte differentiation from embryonic stem cell-derived endoderm.* Hepatology, 2010. **51**(2): p. 633-41.
272. Hovatta, O., et al., *A culture system using human foreskin fibroblasts as feeder cells allows production of human embryonic stem cells.* Hum Reprod, 2003. **18**(7): p. 1404-9.

273. George, A., H.C. Morse, 3rd, and M.J. Justice, *The homeobox gene Hex induces T-cell-derived lymphomas when overexpressed in hematopoietic precursor cells.* Oncogene, 2003. **22**(43): p. 6764-73.
274. Jankovic, D., et al., *Leukemogenic mechanisms and targets of a NUP98/HHEX fusion in acute myeloid leukemia.* Blood, 2008. **111**(12): p. 5672-82.
275. Pelengaris, S., M. Khan, and G.I. Evan, *Suppression of Myc-induced apoptosis in beta cells exposes multiple oncogenic properties of Myc and triggers carcinogenic progression.* Cell, 2002. **109**(3): p. 321-34.
276. Adams, J.M., et al., *The c-myc oncogene driven by immunoglobulin enhancers induces lymphoid malignancy in transgenic mice.* Nature, 1985. **318**(6046): p. 533-8.
277. Pelengaris, S., et al., *Reversible activation of c-Myc in skin: induction of a complex neoplastic phenotype by a single oncogenic lesion.* Mol Cell, 1999. **3**(5): p. 565-77.
278. Felsher, D.W. and J.M. Bishop, *Reversible tumorigenesis by MYC in hematopoietic lineages.* Mol Cell, 1999. **4**(2): p. 199-207.
279. Schoenenberger, C.A., et al., *Targeted c-myc gene expression in mammary glands of transgenic mice induces mammary tumours with constitutive milk protein gene transcription.* EMBO J, 1988. **7**(1): p. 169-75.
280. Kerkhoff, E. and K. Bister, *Myc protein structure: localization of DNA-binding and protein dimerization domains.* Oncogene, 1991. **6**(1): p. 93-102.
281. Sears, R.C., *The life cycle of C-myc: from synthesis to degradation.* Cell Cycle, 2004. **3**(9): p. 1133-7.
282. Albrecht, J.H. and L.K. Hansen, *Cyclin D1 promotes mitogen-independent cell cycle progression in hepatocytes.* Cell Growth Differ, 1999. **10**(6): p. 397-404.
283. Lange, C., W.B. Huttner, and F. Calegari, *Cdk4/cyclinD1 overexpression in neural stem cells shortens G1, delays neurogenesis, and promotes the generation and expansion of basal progenitors.* Cell Stem Cell, 2009. **5**(3): p. 320-31.
284. Fu, M., et al., *Minireview: Cyclin D1: normal and abnormal functions.* Endocrinology, 2004. **145**(12): p. 5439-47.
285. Adhikary, S. and M. Eilers, *Transcriptional regulation and transformation by Myc proteins.* Nat Rev Mol Cell Biol, 2005. **6**(8): p. 635-45.
286. Motokura, T. and A. Arnold, *PRAD1/cyclin D1 proto-oncogene: genomic organization, 5' DNA sequence, and sequence of a*

tumor-specific rearrangement breakpoint. Genes Chromosomes Cancer, 1993. **7**(2): p. 89-95.

287. Herber, B., et al., *Inducible regulatory elements in the human cyclin D1 promoter.* Oncogene, 1994. **9**(7): p. 2105-7.
288. Wang, C., et al., *Signal transduction mediated by cyclin D1: from mitogens to cell proliferation: a molecular target with therapeutic potential.* Cancer Treat Res, 2004. **119**: p. 217-37.
289. Huerta, M., et al., *Cyclin D1 is transcriptionally down-regulated by ZO-2 via an E box and the transcription factor c-Myc.* Mol Biol Cell, 2007. **18**(12): p. 4826-36.
290. Kidder, B.L., J. Yang, and S. Palmer, *Stat3 and c-Myc genome-wide promoter occupancy in embryonic stem cells.* PLoS ONE, 2008. **3**(12): p. e3932.
291. Shtutman, M., et al., *The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway.* Proc Natl Acad Sci U S A, 1999. **96**(10): p. 5522-7.
292. Philipp, A., et al., *Repression of cyclin D1: a novel function of MYC.* Mol Cell Biol, 1994. **14**(6): p. 4032-43.
293. Solomon, D.L., et al., *Expression of cyclin D1 mRNA is not upregulated by Myc in rat fibroblasts.* Oncogene, 1995. **11**(9): p. 1893-7.
294. Mateyak, M.K., A.J. Obaya, and J.M. Sedivy, *c-Myc regulates cyclin D-Cdk4 and -Cdk6 activity but affects cell cycle progression at multiple independent points.* Mol Cell Biol, 1999. **19**(7): p. 4672-83.