IDENTIFICATION AND CHARACTERIZATION OF T CELL FACTOR-4 (TCF-4) ISOFORMS OF WNT SIGNALING IN HEPATOCELLULAR CARCINOMA (HCC)

BY

ORKHONTUYA TSEDENSODNOM

B.S., BOGAZICI UNIVERSITY, ISTANBUL, TURKEY, 2003

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN THE DEPARTMENT OF MOLECULAR BIOLOGY, CELL BIOLOGY, AND BIOCHEMISTRY AT BROWN UNIVERSITY

PROVIDENCE, RHODE ISLAND

MAY, 2010

This dissertation by Orkhontuya Tsedensodnom is accepted in its present form by the Department of Molecular Biology, Cell Biology, and Biochemistry as satisfying the dissertation requirement for the degree of Doctor of Philosophy

Date_____ _____
Jack R. Wands, M.D., advisor

Date_____ _____
Miran Kim, Ph.D., supervisor

Recommended to the Graduate Council

Date_____ _____
Richard N. Freiman, Ph.D., reader

Date_____ _____
Philip A. Gruppuso, M.D., reader

Date_____ _____
Gary M. Wessel, Ph.D., reader

Date_____ _____
György Baffy, M.D., Ph.D., reader

Approved by the Graduate Council

Date_____ _____
Sheila Bonde, Ph.D.
Dean of the Graduate School

Curriculum Vitae

Ms. Orkhontuya Tsedensodnom

EDUCATION

Ph.D. 2010
Molecular Biology, Cellular Biology, and Biochemistry
Brown University
185 Meeting Street, Providence, RI 02912

B.S., Honors 2003
Molecular Biology and Genetics
Bogazici University
Istanbul, Turkey

RESEARCH EXPERIENCE

Graduate Student 2005 – 2010
Dr. Jack R. Wands and Dr. Miran Kim Labs
RI Hospital and the Warren Alpert Medical School of Brown University
Liver Research Center, Providence, RI 02903
Identification and characterization of T-cell factor-4 (TCF-4) isoforms of Wnt signaling in hepatocellular carcinoma (HCC)

Rotation Student Summer 2005
Dr. Jack R. Wands Lab
RI Hospital and the Warren Alpert Medical School of Brown University
Liver Research Center, Providence, RI
Wnt signaling in HCC

Rotation Student Spring 2005
Dr. John M. Sedivy Lab
Laboratories for Molecular Medicine
Brown University, Providence, RI 02903
Upregulation of p38 MAPK and its involvement in overexpression of p16 during cellular senescence

Undergraduate Research Student 2002 – 2003
Dr. Asli Tolun Lab
Department of Molecular Biology and Genetics
Bogazici University, Istanbul, Turkey
Population genetics of consanguineous families and gene locus mapping of chromosome markers implicated in macular degeneration

PUBLICATIONS

Orkhontuya Tsedensodnom, Hironori Koga, Han Chu Lee, Kang Mo Kim, Stephen A. Rosenberg, John J. Carroll, Jack R. Wands and Miran Kim — Submitted

Hepatology

Identification of T-cell factor-4 isoforms that contribute to the malignant phenotype of hepatocellular carcinoma

Miran Kim, Han Chu Lee, **Orkhontuya Tsedensodnom**, Rochelle Hartley, Young-Suk Lim, Eunsil Yu, Philippe Merle, and Jack R. Wands. — 2008

J Hepatol. 2008 May; 48(5): 780–791.

Functional interaction between Wnt3 and Frizzled-7 leads to activation of the Wnt/β-catenin signaling pathway in hepatocellular carcinoma cells

CONFERENCE PRESENTATIONS

Orkhontuya Tsedensodnom, Hironori Koga, Jack R. Wands and Miran Kim. — Apr. 2010

The 101st Annual Meeting of the American Association for Cancer Research

Washington, DC

Characterization and function of unique T-cell factor-4 splicing isoforms that exhibit tumor repressive phenotypes in hepatocellular carcinoma

Orkhontuya Tsedensodnom, Jack R. Wands and Miran Kim. — Nov. 2009

The 60th Annual Meeting of the American Association for the Liver Diseases

Boston, MA

Identification of distinctive T-cell factor-4 splicing variants that contribute to the malignant phenotype of hepatocellular carcinoma cells

Orkhontuya Tsedensodnom, Jack R. Wands and Miran Kim. — Oct. 2009

17th Annual Research Celebration, Rhode Island Hospital, A Lifespan Partner and Care New England Affiliate

Providence, RI

Identification of distinctive T-cell factor-4 splicing variants that contribute to the malignant phenotype of hepatocellular carcinoma cells

CONFERENCE PRESENTATIONS Cont'd

Orkhontuya Tsedensodnom, Jack R. Wands and Miran Kim. Jun. 2008
14th Annual Research Forum, Department of Medicine
The Warren Alpert Medical School of Brown University, Providence, RI

Identification of unique TCF-4 isoforms that exhibit repressive functional characteristics in hepatocellular carcinoma

Orkhontuya Tsedensodnom, Jack R. Wands and Miran Kim. Jun. 2008
RI Research Alliance, Emerging Biomedical and Life Sciences Research
Rhode Island Convention Center, Providence, RI

Identification of unique TCF-4 isoforms that exhibit repressive functional characteristics in hepatocellular carcinoma

Orkhontuya Tsedensodnom, Han Chu Lee, Rochelle Hartley, Jack R. Wands, and Miran Kim. Jun. 2007
The Traditional Wnt Meeting
La Jolla, CA

Activation of the Wnt3/β-catenin pathway is mediated by Frizzled-7 receptor in human hepatoma cells

Orkhontuya Tsedensodnom, Jack R. Wands and Miran Kim. Nov. 2006
The 57th Annual Meeting of the American Association for the Liver Diseases
Boston, MA

Transcription factor TCF-4 isoforms in human hepatocellular carcinoma

Han Chu Lee, **Orkhontuya Tsedensodnom**, Jack R. Wands, and Miran Kim. Jul. 2006
NIH, NCRR 1st Biennial National IDeA Symposium of Biomedical Research Excellence (NISBRE)
Washington, D.C.

Wnt3 is a ligand for Frizzled-7 receptor and increases cell migration in human hepatoma cells

Han Chu Lee, **Orkhontuya Tsedensodnom**, Jack R. Wands, and Miran Kim. May 2006
COBRE/INBRE Symposium, Stem Cell & Cancer Therapeutics
Rhode Island Convention Center, Providence, RI

Wnt3 Is a ligand for Frizzled-7 receptor and increases cell migration in human hepatoma cells

SOCIETY MEMBERSHIP

Associate Member — 2009 – Present
The American Association for Cancer Research

TEACHING EXPERIENCE

Laboratory Mentor to Undergraduate Students — 2006 – 2010
Dr. Miran Kim Lab
RI Hospital and the Warren Alpert Medical School of Brown University
Liver Research Center
Train students and colleagues with lab techniques and data analysis

Teaching Assistant — Spring 2006
Bio20 The Foundation of Living Systems (Introductory Biology)
Brown University
Conducted laboratory sections, graded exams

Teaching Assistant — Fall 2005
Bio 47 Genetics
Brown University
Conducted laboratory sections, graded exams

LANGUAGES

Mongolian	Native language
English	Fluent
Turkish	Fluent
Russian	Intermediate
Hungarian	Intermediate

Dedication:

To my daughter Andrea Onon Tsedensodnom

Acknowledgments

My journey of pursuing a doctorate degree at Brown University and the successful completion of my thesis work wouldn't have come true without the guidance and encouragement of many wonderful people in the graduate program of MCB and in the Liver Research Center (LRC) of Rhode Island Hospital (RIH).

First and foremost, I would like to express my sincere gratitude to my thesis advisors, Jack R. Wands, MD, and Miran Kim, PhD, as I have experienced a unique opportunity of having two mentors during my graduate training. Their complementing guidance, support and encouragement tremendously helped me grow as a scientist and in accomplishing my thesis research. I am greatly thankful to Dr. Wands for accepting me into LRC and providing a stepping stone on my scientific career. I am forever grateful to Dr. Kim for confidently giving me the best project of hers and generously teaching me every biological method step by step. She is a true mentor and genuinely passionate about great education. Even at her busiest day, she would "aliquot" some time to answer my questions and show how the problem can be fixed. I often flaunt of myself for being extremely fortunate to have her as my advisor as there is nothing she would not know and she is literally a readily accessible bank of knowledge and skills.

I would like to extend my gratitude to my thesis committee members, Richard N. Freiman, PhD, Philip A. Gruppuso, MD, and Gary M. Wessel, PhD, for their scientific expertise and helpful comments throughout my research. In addition, I would like to acknowledge Dr. Gruppuso for his valuable suggestions on my writing skills. I would like to give a special thanks to György Baffy, MD, PhD, for enthusiastically accepting to read my thesis and serving in my thesis committee as the outside reader.

I would like to thank to current and former post-doctoral fellows of Dr. Kim lab, Hironori Koga, MD, Takashi Toyama, MD, and Han Chu Lee, MD for their help, mentorship, and sharing their personal tactics and tricks in solving problems. Thanks to other members of the lab, Rochelle R. Hartley, Stephen A. Rosenberg, Tai Ho Shin, and Evan J. Walker, our lab was a fun- and youth-filled vibrant place to work in. I would like to extend my appreciation to other principal investigators in LRC, Suzanne de la Monte, MD, MPH, Shuping Tong, MD, PhD, and JiSu Li, MD, PhD, for their comments and suggestions on my research throughout my graduate training. Many thanks to Rolf I. Carlson and Donna Pratt for their support and patience, and for making LRC a wonderful place for all of us. Dr. Kim and Rolf are the only non-Mongolian people who can say my name fluently and spell it by heart. I will miss the free lunches of our lab meetings and the daily cakes of our cake-hour; but I will definitely follow our www.cakehour.com blog. I would also like to extend my gratitude to my fellow graduate students in LRC for their support and friendships. The generosity of RIH to graduate student volunteers has been invaluable to our budget.

My sincere thanks to Elaine M. Butler for her instant helps in every MCB- and Brown University-related issue I had.

Finally, I would like to express my greatest tribute to my father Tsedensodnom Jamts and my mother Dungarjid Dashbazar for their constant support and unconditional love. In 2008, I was rewarded with the nature's greatest gift of giving birth to a baby girl. It was love at first sight, and the more I love her, the more I realize how much my parents must love me. Only after becoming a mother myself, I have realized the true meaning of unconditional love.

Table of Contents

List of Figures

List of Tables

Abbreviations and Alternative Names

AES	Amino-terminal enhancer of split
AFB1	Aflatoxin B1
AIH	Autoimmune hepatitis
APC	Adenomatous polyposis coli
β-catenin	*CTNNB1* gene; Armadillo in *Drosophila*
BCBD	β-catenin-binding domain
β-Trcp	β-transducin repeats-containing proteins
BMI	Body mass index
CaMK	Calcium/calmodulin-dependent protein kinase
CBP	CREB Binding Protein
CCC	Cholangiocellular carcinoma
CH	Chronic hepatitis
CK1α	Casein Kinase 1α; Gilgamesh in *Drosophila*
CNS	Central nervous system
CRC	Colorectal carcinoma
CRD	Context-dependent transcription repression domain
CtBP	C-terminal Binding Protein
CTD	C-terminal domain
DBD	DNA-binding domain
Dkk	Dickkopf
DMEM	Dulbecco's modified Eagle's medium
dn	dominant negative
Dvl or Dsh	Dishevelled
EGFP	Enhanced green fluorescent protein
ER	Endoplasmic reticulum
ERK	Extracellular signal-regulated kinase
EST	Expressed sequence tag
EV	Empty vector
FAP	Familial adenomatous polyposis
FBS	Fetal bovine serum
Fzd or Fz	Frizzled
GAPDH	Glyceraldehyde 3-phosphate dehydrogenase
GPCR	G-protein coupled receptor
GS	Glutamine synthetase
GSK3β	Glycogen synthase kinease-3β; Shaggy or Zeste-white 3 in *Drosophila*
HAT	Histone acetyltransferase
HB	Hepatoblastoma
HBV	Hepatitis B virus
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HDAC	Histone deacetylase
HEK	Human embryonic kidney
HMG	High mobility group

HMT	Histone methyltransferase
HSC	Hematopoietic stem cells
IARC	International Agency for Research on Cancer
IGF	Insulin-like growth factor
IRS	Insulin receptor substrate
JAK/STAT	Janus kinase/signal transducers and activators of transcription
JNK	c-Jun N-terminal kinase
LDL	Low-density lipoprotein
LEF	Lymphoid enhancer factor
LOH	Loss of heterozygosity
LRP	LDL receptor-related protein
MAPK	Mitogen activated protein kinase
MLL	Mixed lineage leukemia
MMP7	Matrix metalloproteinase 7
MMTV	Mouse mammary tumor virus
N	Normal liver
NAFLD	Non-alcoholic fatty liver disease
NASH	Non-alcoholic steatohepatitis
NK	Natural killer
NLK	Nemo-like kinase
NLS	Nuclear localization sequence
NTD	N-terminal domain
PBC	Primary biliary cirrhosis
PCP	Planar cell polarity
PDGF	Platelet-derived growth factor
PI3K	Phosphoinositide 3 kinase
PIC	Pre-initiation complex
PKC	Protein kinase C
PSC	Primary sclerosing cholangitis
pT	HCC matching peritumor tissue
PTEN	Phosphatase and tensin homolog
qPCR	Quantitative real time RT-PCR
Rb	Retinoblastoma
RCC	Renal cell carcinoma
RKIP	Raf kinase inhibitor protein
ROS	Reactive oxygen species
RT-PCR	Reverse transcriptase polymerase chain reaction
SD	Standard deviation
semi-qPCR	semi-quantitative RT-PCR
sFRP	Secreted Frizzled-related protein
SUMO	Small ubiquitin-like modifier
T	HCC tumor tissue
TBP	TATA box-binding protein
TBX3	T-box protein 3
TCF	T cell factor; Pangolin or dTCF in *Drosophila*
TERT	Telomerase reverse transcriptase

TGF	Transforming growth factor
TLE	Transducin-like enhancer of split proteins; Groucho in *Drosophila*
TNF-α	Tumor necrosis factor-α
VEGF	Vascular endothelial growth factor
WB	Western blot
WHO	World health organization
WIF	Wnt inhibitory factor
WRE	Wnt responsive element

CHAPTER 1

1. INTRODUCTION

PART 1

1.1. HEPATOCELLULAR CARCINOMA

1.1.1. HEPATOCELLULAR CARCINOMA: EPIDEMIOLOGY

Liver cancer consists of different types of primary hepatic neoplasms that include hepatocellular carcinoma, intrahepatic bile duct carcinoma (cholangiocellular carcinoma), hepatoblastoma, bile duct cystadenocarcinoma, hemangiosarcoma and epitheliod hemangioendothelioma [1]. The predominant type is hepatocellular carcinoma (HCC), which comprises 85% – 90% of primary liver cancers [2]. It is the fifth most common cancer worldwide and the third leading cause of cancer mortality [3,4] due to limited treatment options. Currently, the estimated worldwide incidence of new cases of HCC is about 500,000 – 1,000,000 per year, resulting in more than 600,000 deaths each year [3,5]. These incidence and mortality rates are almost equal because the general prognosis is poor with overall survival rate of 3% – 5% as it is often diagnosed at an advanced stage [6]. Moreover, HCC is resistant to radiotherapy and unresponsive to chemotherapy limiting the curative options to surgical resection and liver transplantation though recurrence and metastasis is quite common. The worldwide prevalence of HCC (Figure 1.1.1.) ranges from low rate of <5 cases per 100,000 population, which mostly dispersed in North America, Australia and Northern Europe, to medium rate of 5 – 20 cases per 100,000, typically in Southern Europe and South Central Asia, to high rate of 20 – 100 cases per 100,000, usually confined to Eastern Asia and Central Africa [2-4]. Unfortunately, the highest rate of incidence is observed in my home country of Mongolia with 99 cases per 100,000 persons [3,4], while the second highest rate is reported in South Korea (49/100,000), followed by Gambia (39/100,000), China (35/100,000), Mali (34/100,000), and Japan (29/100,000). The incidence of HCC is increasing in developed western countries as well even though they are currently in the low rate category [2,3].

Particularly, in the United States of America, the HCC incidence rate has tripled between 1975 and 2005, rising from 1.6/100,000 to 4.9/100,000 [7].

In addition to geographic variation in HCC prevalence, the rate of incidence is also affected by gender, ethnicity, and age. Cancer registries around the globe have reported that males have 2 – 4 times higher rates of both HCC incidence and mortality than females [2,4], probably due to higher body mass index, increased levels of androgenic hormones, and sex-specific exposure difference to risk factors. Moreover, various ethnic populations living in the same country show different HCC incidence profiles. For example, in the United States, the HCC rate is twice as high in Asians than in African Americans, whose rates, in turn, are two times higher than white Americans [2]. In almost all populations, the incidence of HCC increases with age, reaching its highest prevalence among people aged over 65 years, due to acquisition and prolonged exposure to major risk factors of HCC.

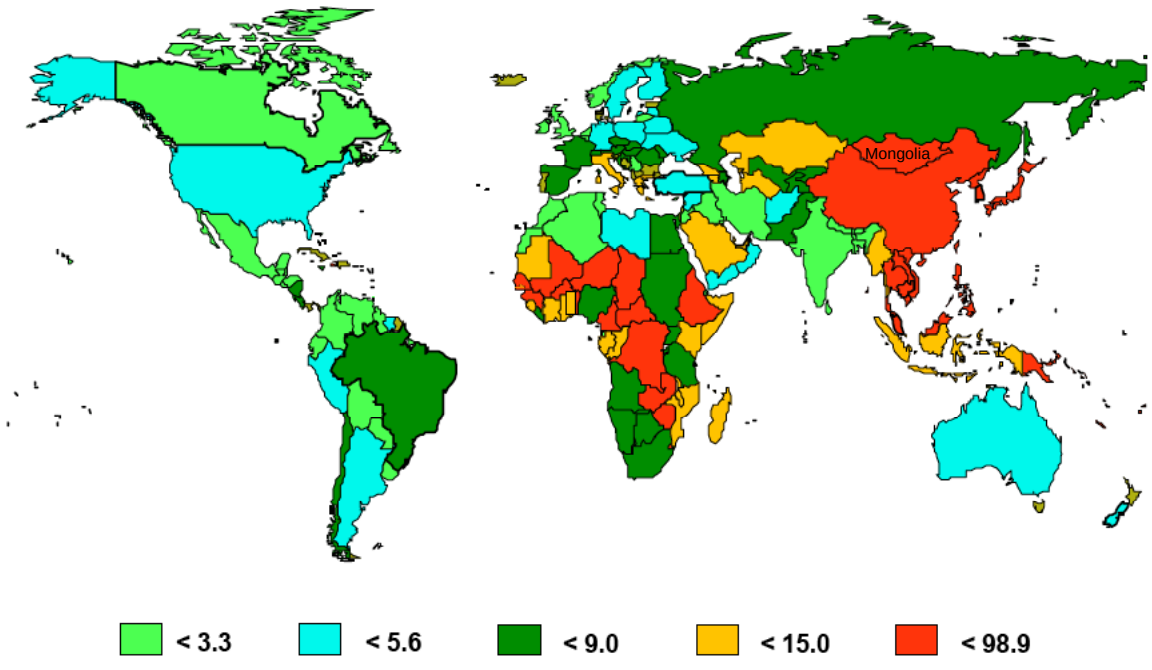


Figure 1.1.1. Worldwide incidence rate of HCC estimated in 2000 [4]

1.1.2. HEPATOCELLULAR CARCINOMA: ETIOLOGY

In the majority of cases (>80%), the main predisposing factor for development of HCC is a background liver disease of chronic hepatitis and/or cirrhosis, although HCC can arise from normal liver [5]. Chronic hepatitis is a prolonged inflammation of the liver, a process in which inflammatory cells embedded in the liver tissue determine hepatocyte death (necrosis) followed by regeneration of viable hepatocytes. Persistent chronic hepatitis progresses into cirrhosis, an irreversible end-stage of liver fibrosis with distortion of the hepatic architecture and dysplastic nodule formation. Fibrosis is the accumulation of extracellular matrix and collagen resulting in scarring of the liver tissue. The major risk factors responsible for HCC, cirrhosis and chronic hepatitis are viral (chronic hepatitis B virus and hepatitis C virus infections), toxic (alcohol and aflatoxins), metabolic (diabetes, non-alcoholic fatty liver disease, and hemochromatosis) and immune-related (primary biliary cirrhosis and autoimmune hepatitis) (Figure 1.1.2.) [3]. Combination of two or more of these etiological factors enhances the development of cirrhosis and HCC.

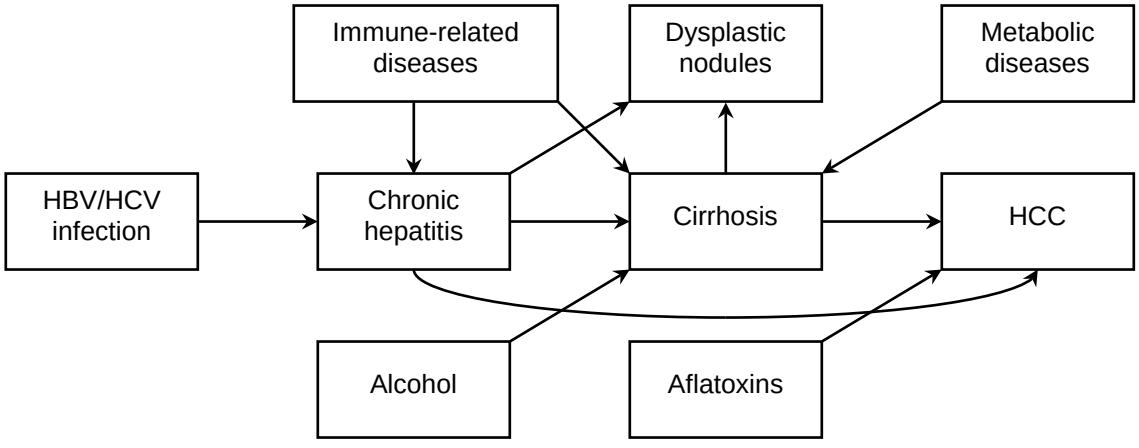


Figure 1.1.2. Major risk factors in development of HCC

Adapted from *Molecular pathogenesis of hepatocellular carcinoma* [8]

1.1.2.1. Viral risk factors

There are two main hepatitis viruses associated with the development of HCC: hepatitis B virus (HBV) and hepatitis C virus (HCV). An estimated 75% – 85% of HCC cases are due to chronic viral infections of HBV (50% – 55%) or HCV (25% – 30%) [4].

1.1.2.1.1. HBV

According to World Health Organization (WHO), HBV has infected about 2 billion people worldwide, of which 350 million are chronically infected, and results in an estimated 600,000 deaths each year [9]. Approximately 30% – 50% of HBV-related mortality is caused by HCC [1]. Fortunately, Hepatitis B vaccine has been available since 1982, and this vaccination is widely documented as the most effective technique to inhibit HBV infection, making HCC the first human cancer susceptible to prevention as the risk of HCC is substantially lower in people who are immune to HBV [2,3,9,10].

Hepatitis B virus is a hepatotropic, partially double-stranded DNA virus of the *hepadnaviridae* family [10]. Because the malignant transformation due to HBV occurs after a long period of chronic hepatitis and cirrhosis, the role of HBV in HCC is complex and involves both direct and indirect mechanisms [11]. HBV has been shown to induce hepatocarcinogenesis by interacting with the host at cellular, organelle, genome, and cell signaling levels [1]. Upon infection, the T-cells initiate an immune response that triggers hepatocyte necrosis, followed by inflammation, and consequently regeneration. In chronic HBV infections, this cycle of necrosis-inflammation-regeneration persists for long periods of time ultimately leading to carcinogenesis [11]. Mutations in HBV enable it to evade the host cell's immune response increasing the chance of chronic infection and subsequent development of HCC. Moreover, HBV has been shown to physically interact

with the host cell's endoplasmic reticulum (ER) inducing oxidative stress, which causes generation of free radicals and activation of stellate cells [1]. Once activated, stellate cells, the quiescent liver cells that store retinol, proliferate vigorously, acquire myofibroblast morphology, and robustly synthesize extracellular matrix components such as collagen, greatly contributing to liver fibrosis. Furthermore, HBV invades the host cells at the genome level by integrating into the chromosomes to produce host DNA microdeletions. Large scale analysis of HBV DNA integration in HBV-related HCCs showed that HBV recurrently targeted the cancer-relevant genes that potentially may provide cell growth advantage, such as *telomerase reverse transcriptase (TERT), platelet-derived growth factor receptor-β (PDGFRβ), PDGFβ* and *mitogen activated protein kinase 1 (MAPK1)* [12]. Finally, HBV induces hepatocarcinogenesis by altering host's cell signaling system to increase cellular proliferation and survival by either directly interacting with the components or indirectly mutating them due to long-term toxic effects of the immune response. HBx, a transcription factor encoded by HBV, activates host proto-oncogenes or growth-control genes, such as *c-Myc* and *SRC* tyrosine kinases, *Ras, Raf, MAPK, extracellular signal-regulated kinase (ERK), and c-Jun N-terminal kinase (JNK),* and inactivates tumor suppressors such as p53, APC, and p21 [1,3,5,8,11].

1.1.2.1.2. HCV

WHO estimates that about 170 million people worldwide are infected with HCV, of which 130 million are chronically infected, and 3 – 4 million people are newly infected each year [13]. It has been assessed that HCV causes 27% of cirrhosis and 25% of HCC worldwide [14]. Unfortunately, there is no vaccine against HCV, mostly due to high mutability of the HCV genome, and the antiviral therapy treatments are too costly.

Hepatitis C virus is a hepatotropic, positive-stranded RNA virus of the *flaviviridae* family [15]. Because it does not make a DNA intermediate, HCV cannot integrate into host genome and thereby it does not cause gene mutations by DNA deletion. Moreover, HCV has a higher frequency to achieve chronic infection (60% – 80%), compared to HBV, which is 10% of all HBV cases. However, similar to HBV, the exact mechanism of hepatocarcinogenesis in chronic HCV infection is unclear as it takes 25 – 30 years to develop HCC after the initial infection [2]. Studies show that HCV also induces similar biological processes as HBV, though HCV may have more tendency to activate the necrosis-inflammation-regeneration cycle by easily escaping the host cell's immune responses [1,14]. HCV interacts with the host cell's ER during its replication process, provoking ER stress and its subsequent carcinogenic consequences. Additional evidence of oxidative-stress-mediated mechanism in HCV-induced HCC development is indicated by the activation of hepatic steatosis. Steatohepatitis, or fatty liver, is a process of accumulation of fatty acids in hepatocytes and induction of reactive oxygen species that result in lipid peroxidation [1]. With its core protein that has transcriptional effects among other activities, HCV alters expression of different cellular genes including the proto-oncogene *c-Myc* to enhance cell growth; modulates cellular proteins, such as tumor necrosis factor-α (TNF-α) to inhibit apoptosis, and/or tumor suppressor p53 to increase cell survival; and interacts with components of cell signaling pathways such as ERK, MEK and Raf of the MAPK signaling pathway to increase cell proliferation [1,5,14]. Finally, the core protein and some of the non-structural proteins of HCV have been shown to have direct carcinogenic potential as they induce anchorage-independent growth when expressed in fibroblasts and tumor formation in transgenic nude mice [1,5].

1.1.2.2. Toxic risk factors

The liver is an important organ within the body that has a central role in metabolism, synthesis, storage, and redistribution of nutrients, carbohydrates, fats, and vitamins. In addition, it functions as the main detoxifying organ by removing potentially toxic compounds that are either taken up from outside of the body or made within the organism. The hepatocytes, or the parenchymal cells, which comprise ~80% of hepatic cells, are the main cell type of the liver that carries out most of these functions. The other 20% makes up the non-parenchymal cells, which include endothelial cells, Kupffer cells, lymphocytes, and stellate cells [16]. Even though the metabolism and eventual removal of toxic wastes are processed through elaborate, multi-step mechanisms, repeated exposure to these toxic compounds affects hepatocyte integrity and results in liver injury, cirrhosis, and HCC. Chronic alcohol intake and consumption of aflatoxin-contaminated foods among other drugs, medications, diet, and chemicals have been identified as major toxic, dietary risk factors in hepatocarcinogenesis.

1.1.2.2.1. Alcohol

A significant amount of epidemiological data demonstrates that alcohol is an independent and strong risk factor for HCC [17]. Although quantification of the association is not possible based on epidemiological data, it is estimated that risk of HCC appears when alcohol intake exceeds 60 grams of ethanol per day [18], which is equivalent to ~5 drinks of 12 oz (360 ml) of beer, 4 oz (120 ml) of wine, or 1.5 oz (45 ml) of spirits/liquor [19]. The risk is further increased upon development of cirrhosis even if alcohol consumption is decreased or stopped. Alcohol-induced hepatocarcinogenesis is triggered by the induction of inflammation due to increased amount of endotoxin and activation of

monocytes and Kupffer cells [1]. Kupffer cells are specialized macrophages, or monocytes, localized in the liver. When activated, Kupffer cells migrate within the liver to the site of injury or infection, initiate biochemical attack by releasing various chemokines, cytokines, reactive oxygen species (ROS) and proteolytic enzymes, and remove foreign particles or infecting organisms by phagocytosis [20]. In the context of chronic alcohol intake, hepatocytes become sensitive to repeated exposure to cytokines, chemokines, and ROS leading to recurrent cycles of necrosis-inflammation-regeneration and activation of stellate cells, which result in cirrhosis and ultimately HCC [21]. Moreover, because the metabolism of ethanol in hepatocytes produces ROS, excessive alcohol intake may trigger the development of HCC through oxidative stress mechanisms [5], which promote fibrosis and cirrhosis by activating stellate cells, alter signaling pathways involved in cell growth and proliferation, generate oncogenic mutations, increase DNA damage and chromosome instability, and accelerate telomere shortening. In addition, long-term alcohol ingestion sets the stomach, liver and intestines vulnerable for further assaults, as it depletes body's natural anti-oxidants, weakens the immune system, and permeabilizes the cells lining these organs for pathogens [1]. Finally, alcohol has been established as a strong synergistic cofactor for HBV or HCV infections on the development of HCC, as it speeds up the time of onset by 10 years, and increases the risk of HBV-mediated HCC to 3 – 4 fold and the risk of HCV-induced HCC to 2 fold [19,22].

1.1.2.2.2. Aflatoxin

Aflatoxins are mycotoxins produced by the molds *Aspergillus flavus* and *Aspergillus parasiticus*. In Asia and southern Africa, these fungi readily grow on harvest foods, such as grain, corn, nuts, and rice, when they are processed and stored improperly under hot

and humid conditions. Out of more than 13 types of aflatoxins, Aflatoxin B1 (AFB1) is the most toxic, and it has been classified as carcinogenic by the International Agency for Research on Cancer (IARC) [2]. Geographic areas with high levels of AFB1 contamination coincide with the regions with high incidence of HCC [3], but there is no apparent correlation between AFB1 exposure and development of cirrhosis, thus the hepatotoxicity or carcinogenic effect of aflatoxins may exploit a direct mechanism towards hepatocarcinogenesis [1]. Once ingested, AFB1 is metabolized into a reactive intermediate species that can bind to guanine residues of hepatocyte DNA and form a number of adducts. In an acute aflatoxin exposure, the formation of adducts leads to DNA damage, inhibition of RNA synthesis, heritable genetic mutations, acute hepatic injury, necrosis, and possible hepatic failure, whereas chronic aflatoxicity inevitably results in the development of HCC [23]. Studies show that AFB1 exerts its carcinogenic effect by functioning as a mutagen on tumor suppressor genes or oncogenes. One of the most common DNA mutations produced by AFB1 in hepatocytes is a guanine-to-thymine transversion at the third base of codon 249 of the tumor suppressor *TP53* gene [24]. This mutation, which leads to an arginine-to-serine substitution (Arg249Ser) in the p53 protein, has been observed in 50% – 70% of HCC cases in aflatoxin-endemic regions. However, the specificity of the p53 Arg249Ser mutation has been identified in <4% of HCC incidences where aflatoxin exposure is minimal or undetectable [24,25]. In addition, AFB1 exposure often coexists with HBV infection, and their synergistic effect increases the risk of developing HCC 5 – 10 fold compared to the effect of either one of the risk factors. On the other hand, such synergistic effect or any correlation between AFB1 and HCV infection has not been documented [4].

1.1.2.3. Metabolic risk factors

In countries with a relatively high amount of alcohol consumption, intermediate levels of HBV and HCV infections, and low rate of Aflatoxin exposure, these main viral and toxic etiologies account for ~85% of HCC cases leaving little space for other causes. The other causes are mostly due to metabolic risk factors, such as diabetes, non-alcoholic fatty liver diseases, and hemochromatosis.

1.1.2.3.1. Non-alcoholic fatty liver disease, obesity, and diabetes

Non-alcoholic fatty liver disease (NAFLD) is a wide spectrum of liver diseases associated with fat accumulation in the absence of alcohol that range from simple fatty liver to non-alcoholic steatohepatitis (NASH) and finally to cirrhosis [20]. NASH is the most severe form of NAFLD before developing cirrhosis characterized by accumulation of fat, inflammation, necrosis, and fibrosis of the liver. About 20% of adults suffer from NAFLD, 3% from NASH, and it progresses into cirrhosis in every fifth patient [26]. It has been estimated that the development of HCC in patients with cirrhosis caused by NASH occurs with the same frequency as in patients with chronic HCV infection. The main cause for the development of NASH is the presence of insulin resistance, and the major risk factors for insulin resistance are obesity and sedentary lifestyle, in addition to genetic predisposition. In 1997, WHO formally recognized obesity as a global epidemic, and as of 2005, WHO estimated that 10% of adults are obese, 40% are overweight [27], and 70% live in sedentary lifestyle [26]. Defined as body mass index (BMI) of more than 30 kg/m^2, obesity alone increases the risk of developing HCC 2 – 3 fold compared to individuals with normal BMI [2], as 20% – 80% of obese people suffer from NASH, and 5% of the cases progress into NASH-related cirrhosis [6]. Type 2 diabetes, one of the

consequences of insulin resistance, has been identified as an independent risk factor for chronic liver disease and HCC through development of NASH as diabetes has been associated with a 3 fold increased risk of developing HCC regardless of other risk factors [2]. In addition, as observed in other etiologies, coexistence of NAFLD, obesity, and diabetes with viral infections dramatically increases the risk of HCC and synergistically contributes to the hepatocarcinogenesis [26]. Needless to mention, obesity and diabetes mainly increase the risk of heart disease, stroke, neuropathy, and kidney failure.

1.1.2.3.2. Hemochromatosis

Excess accumulation of iron in the liver either due to hereditary hemochromatosis or dietary iron overload has been shown to induce chronic necroinflammatory hepatitis leading to cirrhosis and finally HCC [6]. Hereditary hemochromatosis is mainly caused by a defect in *HFE* gene that regulates the amount of iron absorbed from food. The two known mutations of *HFE* are Cys282Tyr and His63Asp. The Cys282Tyr missense mutation has a more pronounced effect on HFE protein function than the H63D mutation, and the homozygotes of either gene are affected with this genetic disease. Thus, 95% of patients with hereditary hemochromatosis are Cys282Tyr homozygous, 4% are Cys282Tyr/His63Asp compound heterozygous, and 1% are His63Asp homozygous [28]. Studies show that Cys282Tyr homozygotes have 2 – 3 fold increased risk of developing HCC, while there is no evidence for significant prevalence of His63Asp mutations in HCC cases [2]. In addition, alcohol intake acts synergistically with excessive iron load in hepatocarcinogenesis. Although hemochromatosis is often classified as a liver disease, the excessively absorbed iron is also stored in other organs of the body such as tissues of heart and pancreas, resulting in additional serious diseases.

1.1.2.4. Immune-related risk factors

In rare cases where the well-known etiologies are excluded, the underlying cause of such HCC incidences may be identified as non-infectious, immune-related liver diseases. Autoimmune hepatitis (AIH), primary biliary cirrhosis (PBC), and primary sclerosing cholangitis (PSC) form the triad of autoimmune liver diseases. AIH is a chronic active hepatitis with continuous hepatocellular inflammation and necrosis, which may progress to cirrhosis and liver failure if untreated [29]. In 1% – 3% of AIH cases, complication of the disease results in HCC, and the following factors in AIH patients increase the risk of hepatocarcinogenesis by 6 – 24 fold and they serve as predictive risk factors for this neoplasm: male gender, history of blood transfusions, features of portal hypertension, treatment failure, progression to cirrhosis, and immunosuppressive treatment of >3 years [6]. PBC is a chronic cholestatic liver disease characterized by inflammation and necrosis of the bile ducts, which eventually may progress to fibrosis, cirrhosis, and liver failure [30]. Because the inflammatory lesions associated with PBC are not hepatocytes, but the cells of the bile ducts, or cholangiocytes, the incidence of HCC due to complication of PBC is low – about 0.5% of PBC cases. Like AIH, age, male gender, and advanced histologic stage of PBC are found to be associated with HCC development. PSC is another chronic cholestatic liver disease that is characterized by diffuse inflammation, obstruction, and fibrosis of the bile ducts [31]. Although the most common complication of PSC is cholangiocellular carcinoma, in very rare cases the end-stage PSC progresses into liver cirrhosis and eventually HCC. In addition, patients with AIH tend to have another autoimmune liver disease, PBC or PSC, and this situation, termed as overlap syndrome, increases the risk of HCC even further.

1.1.3. CELL SIGNALING PATHWAYS INVOLVED IN HEPATOCARCINOGENESIS

Despite the well established risk factors, the cellular and molecular mechanisms that occur during the initiation, promotion, and progression of HCC either due to the viral, toxic, and metabolic factors, or etiologically non-specific, sporadic causes are less well understood. The multistep process of hepatocarcinogenesis includes chronic liver injury, inflammation, cell death, fibrosis, cirrhosis, regeneration, DNA damage, dysplasia and HCC. However, the molecular mechanism of hepatocarcinogenesis is a more complex process that involves numerous genetic and epigenetic alterations, many of which may have already been triggered during the cirrhosis level. One of these alterations includes inhibition of tumor suppressor genes and/or activation of proto-oncogenes that regulate cell cycle, migration and apoptosis thereby promoting the survival and propagation of cancerous cells, and finally formation of tumors. Extensive research over the years has revealed several signal transduction pathways that are implicated in HCC. Although the molecular mechanism of hepatic oncogenesis may differ according to the risk factors involved, the underlying cell signaling pathways are related to each other and the genetic alterations are not of independent or separate pathways. Another genetic alteration is observed as reactivation of telomerase activity. Telomerase is non-functional in normal liver, whereas it is progressively activated from cirrhotic liver, to regenerative nodules, low-grade to high-grade dysplastic nodules and HCC. Finally, chromosomal aberrations, such as allelic deletions or loss of heterozygosity (LOH), are observed as the genetic changes of the late stage hepatocarcinogenesis. Loss of chromosomal arms that contain tumor repressor genes or gain of chromosomal loci that harbor proto-oncogenes ensure the progression of HCC.

1.1.3.1. Wnt/β-catenin pathway

Wnt/β-catenin signaling pathway is highly conserved from fruit flies to humans and involved in numerous cellular processes including homeostasis, cell proliferation, differentiation, motility, and apoptosis. It is well established by its role in embryogenesis as well as in a variety of carcinogenesis. The hallmark of active Wnt signaling is the stabilization and accumulation of β-catenin in the cytoplasm and the nucleus, which eventually results in the expression of Wnt target genes such as c-Myc and cyclin D1 [32,33]. In the liver, Wnt signaling plays many important roles during hepatic development and regeneration, while its deregulation contributes to many abnormalities ranging from aberrant hepatic growth to various liver tumors [34]. This pathway is activated in HCC arising from HBV/HCV infections, aflatoxin exposure, and alcoholic liver cirrhosis [35]. In most cases of HCC, the hyper-activation of Wnt signaling is observed due to the stabilization and accumulation of β-catenin [36]. Several lines of studies have reported that the β-catenin gene, *CTNNB1*, is mutated in 25% – 35% of all HCCs [37,38]. When the etiological factors are considered, 40% of HCV-associated HCCs and 45% of aflatoxin-induced HCCs demonstrate stabilizing mutations in *CTNNB1*, but HBV-related HCCs have a much lesser rate of β-catenin mutations [34]. Likewise, alcohol-associated HCCs show less deregulation in the Wnt/β-catenin pathway, but rather it is involved Rb1 and p53 signaling cascades more often [39]. Interestingly, 60% – 70% of all HCCs with β-catenin mutations irrespective of etiology developed in the background of non-cirrhotic liver, indicating that *CTNNB1* mutations might lead to HCC in the absence of cirrhosis [40]. On the other hand, HCCs arising from cirrhosis mainly demonstrate p53 alterations [34].

Dephosphorylation of β-catenin, thus its accumulation, is frequently observed in HCC as well [41]. However, abnormal redistribution and accumulation of wild type β-catenin is detected in up to 90% of HCCs suggesting additional mechanisms that contribute to the stabilization and accumulation of β-catenin [42,43]. In addition to mutations in the β-catenin gene, other components of the Wnt signaling cascade have been affected to activate this molecule. The negative regulators of Wnt pathway, Axin 1 and Axin 2 are mutated and non-functional in 10% and 3% of HCCs, respectively [44], while inactive or phosphorylated glycogen synthase kinase-3β (p-GSK-3β), which is unable to present β-catenin for degradation, is increased in 50% of HCCs [45]. However, the tumor suppressor adenomatous polyposis coli (APC) is rarely mutated in HCC [39]. Our laboratory has revealed that Frizzled-7 (FZD7) receptor is upregulated in up to 90% of liver tumors and overexpression of FZD7 stabilizes β-catenin early in the hepatic oncogenesis [46,47]. Furthermore, we have identified Wnt3 as the ligand for FZD7 and that its expression is increased in 76% of HCCs leading to activated Wnt signaling and accumulation of β-catenin in cytoplasm and nucleus [48]. Finally, it has been shown that the inhibitor of Wnt signaling, secreted frizzled-related protein (sFRP) is epigenetically inactivated in up to 75% of HCCs [49]. Interestingly, the components of the Wnt pathway are associated with low rate of LOH suggesting that β-catenin activation may be a very important early event in hepatocarcinogenesis [24].

1.1.3.2. p53 pathway

The tumor suppressor p53 (encoded by *TP53* gene) is a stress response protein whose activation is induced by a number of stress signals, including DNA damage, oxidative stress and oncogenic activation. In addition to its direct roles in DNA repair,

recombination, genome stability, and chromatin modification, p53 is mainly employed as a transcription factor of huge number of p53-regulated genes. Activation of these genes results in three major outputs: cell cycle arrest, cellular senescence and apoptosis. Other p53-regulated gene functions include cell differentiation, motility, angiogenesis, energy metabolism, immune response, cell to cell communication, positive and negative feedback loops that enhance or attenuate the functions of the p53 protein, and activation of other signal transduction pathways to integrate these stress responses [50]. Therefore, p53 function is almost always compromised in all tumor cells, usually as a result of single point mutations, which occur in approximately half of all human cancers. In most of the remaining cancers, the functions of the p53 pathway are defective by other means [35]. Somatic *TP53* mutations are usually followed by LOH during tumor progression inactivating the remaining wild-type allele.

Mutations in *TP53* gene are found in 30% of HCC cases worldwide, while the frequency of all p53 mutations ranges from 20% in North America to 67% in Africa [51]. In addition to the frequency, the type of mutation also differs according to the geographic location and suspected etiology. A guanine-to-thymine transversion mutation at the third base position of codon 249 of the *TP53* gene leading to Arg249Ser substitution in the p53 protein has been observed in 50% – 70% of HCC cases in aflatoxin-endemic regions [24]. This mutation is identified as a “hotspot” mutation for HCC as its detection in plasma serves as a biomarker for both AFB1 exposure and risk for HCC development [52]. Another p53 mutagen is oxidative stress, and it results in the development of cirrhosis with a 200 fold risk for HCC [35]. The mutations include a guanine-to-thymine

transversion at codon 249, a cytosine-to-adenine and a cytosine-to-thymine transversions at codon 250 [35].

In HBV-associated HCC cases, the HBx protein, which is required for the transcription of the viral genome, has been detected to bind to p53 and inactivate p53-mediated apoptosis and DNA repair [52]. In addition, expression of HBx increases the frequency of Arg249Ser mutation of p53 protein in HCC cases associated with both AFB1 and HBV [51]. In HBV- and HCV-related HCCs, the frequency of p53 mutations is higher in advanced malignancies (43%) than in regenerative nodules (7%) indicating that p53 alteration may contribute to cancer progression in those HCCs [1].

1.1.3.3. Rb pathway

The tumor suppressor protein retinoblastoma (Rb) interacts with a large number of cellular proteins to regulate cell cycle progression, differentiation, apoptosis and tumor suppression. The best studied binding partner of Rb is the E2F transcription factor, which controls the expression of genes required for transition from G1 to S phase of the cell cycle [53]. In G0 and early G1 phases, hypo-phosphorylated Rb binds to E2F and inhibits its partner's transcriptional activity. Upon mitogen stimulation in the late G1 phase, Rb is inactivated through phosphorylation by the cyclin D/cyclin-dependent kinase 4/6 (cdk 4/6) or cyclin E/cdk 2 complex. Hyper-phosphorylated Rb releases E2F, allowing the expression of genes essential for S phase progression and DNA replication. In most human tumors, Rb is functionally inactivated either by direct mutation/deletion or indirectly through altered expression/activity of upstream regulators. The Rb pathway, comprised of Rb family of proteins, Rb1, Rb2, and p107, cyclins D and E, cdks 2, 4, and 6, and cdk inhibitors p16, p21, and p27, is disrupted in 69% – 100% of HCC patients

[39]. Mutations of Rb are observed in 15% of HCCs, and LOH at the *Rb1* gene locus is detected in 25% – 48% of liver tumors [24]. In 10% – 13% of HCC cases, cyclins are overexpressed. In particular, amplification of cyclin D1 gene correlates with aggressive forms of human HCC, and its overexpression is sufficient to initiate hepatic oncogenesis in transgene mouse models [51]. In addition, cyclin D1 is a well-characterized target gene of the Wnt/β-catenin pathway, whose alteration in HCC contributes to the overexpression of this protein [33]. Genetic changes in the expression of the cdk inhibitors p16, p21, and p27 leads to carcinogenesis in almost 90% of HCC cases [35]. p16 alone is inactivated in 60% of HCCs, throughout early to late stages of the disease, either due to LOH at its gene locus chromosome 9p21 (INK4A/ARF) or methylation of its promoter [51]. Another tumor suppressor gene located at INK4A/ARF locus is p14/p19, which induces G1 and G2 cell cycle arrests by stabilizing the p53 transcription factor. While p14/p19 is altered in 15% of HCCs, combined disruption of p16, p14/p19, and p53 pathway is observed in 86% of liver tumors [51]. Genetic alterations of p21 are also associated with *TP53* gene mutations in HCC and their combined effect greatly contributes to hepatocarcingenesis.

1.1.3.4. IGF pathways

The insulin/insulin-like growth factor-1/2 (IGF-1/IGF-2) signaling cascades, which play a central role in maintaining basic cellular functions such as cell proliferation, survival and metabolism through activation of numerous target genes including p27, Myc, Fos, cyclin B, and vascular endothelial growth factor (VEGF) [54], are often activated in the majority of HCCs [55]. IGF-2, which is highly expressed in fetal liver but strongly downregulated after birth, is found to be overexpressed in 16% – 40% of HCC cases, different HCC cell lines and several HCC animal models [55]. IGF-2 is partially

regulated by IGF-2 receptor (IGF-2R), which promotes degradation of IGF-2 and activation of transforming growth factor-β (TGF-β). Hence, LOH in IGF-2R locus is observed in more than 60% of HCCs [2] and expression of IGF-2R is reduced in 63% of HCC cases [55]. Insulin receptor substrate-1 (IRS-1), the principle intracellular substrate for insulin/IGF-1 signaling, is overexpressed 3 – 4 fold in 85% of human HCCs [56], and its overexpression levels are correlated with increased HCC tumor size and progression [57]. In addition, overexpression of IRS-1 leads to activation of its downstream signal transduction pathways involved in cell growth, survival, and invasion: the phosphoinositide 3 kinase (PI3K) pathway [57] and the MAPK pathway [56]. 36% of HCCs have been found to have activating mutations in the kinases of the PI3K pathway, and 5% of those mutations activated the downstream Akt (also known as protein kinase B) signaling [25]. Moreover, the most important negative regulator of the PI3K pathway, the tumor suppressor phosphatase and tensin homolog (PTEN), which also lowers the secretion of IGF and expression of VEGF in hepatocytes, is altered in about 10% of HCCs [51]. The MAPK signaling cascade is mostly activated in HBV- and HCV-associated HCCs [1]. There are three major MAPK pathways: the ERK pathway, the JNK pathway, and the p38 MAPK pathway. The protein kinases of the ERK/MAPK signaling cascade including Raf, MEK (MAPK ERK kinase), and ERK are upregulated in 50%, 64%, and 60% – 100% of HCC cases, respectively [58]. Additionally, the Raf kinase inhibitor protein (RKIP), which is the negative regulator of the ERK/MAPK pathway, is reduced in 80% – 90% of HCCs and the reduction of RKIP leads to increased expression of phospho-ERK in 86% of HCCs [59] indicating that the loss of RKIP function is an important step in hepatocarcinogenesis.

1.1.3.5. Other pathways

A number of other genetic alterations have been described in HCC over the years, many of which are implicated in different cellular pathways that regulate cell growth, proliferation, differentiation, migration, adhesion, and apoptosis. Deregulation of these pathways often lead to different types of cancer including HCC. The TGF-β signaling pathway inhibits cell growth and promotes apoptosis in hepatocytes. Many of its main components are altered due to LOH and/or mutations in 10% – 30% of HCC cases [35]. Furthermore, various chemicals trigger the Ras pathway in HCC by causing activating point mutations in Ras proteins. Ras proteins are intracellular, membrane-associated proteins that are activated by extracellular growth factors and relay their signals to many downstream pathways implicated in cell proliferation, migration, survival, and differentiation. Mutations in *Ras* genes, such as single point mutations in codon 13 of *H-Ras*, codon 12 of *N-Ras*, and codon 61 of *K-Ras* seen in HCC [35], can permanently activate these molecular switch proteins and lead to inappropriate transmissions even in the absence of extracellular stimuli. Finally, activation of Janus kinase/signal transducers and activators of transcription (JAK/STAT) signaling, which is activated by a variety of cytokines, hormones, and growth factors, is enhanced in 37% – 83% of HCC cases [60] as deregulation of this pathway leads to increased angiogenesis, cell survival and immunosuppression. Additional pathways of cellular processes, such as alcohol metabolism, stress response, intracellular transport, and ubiquitins, may play a role in promoting hepatocarcinogenesis.

1.1.4. SUMMARY

Hepatocellular carcinoma, the primary malignancy of the liver, is the fifth most common cancer worldwide and the third most frequent cause of cancer-related mortality. The incidence and mortality rates are almost equal, reaching up to half a million per year. The main reason for these unfortunate statistics is that HCC is usually diagnosed at an advanced stage and has poor prognosis. The highest rate of prevalence is observed in Eastern Asia and Central Africa. Though USA is among the countries with the lowest incidence, the rate of occurrence has tripled over the last three decades.

The major risk factors responsible for HCC and its precursors, chronic hepatitis and cirrhosis, are well established. They include chronic infection with HBV and/or HCV, dietary exposure to aflatoxin B1, prolonged abuse of alcohol, and predisposition to metabolic and immune-related diseases. Combination of two or more of these etiological factors enhances the development of cirrhosis and HCC. However, the genetic and epigenetic alterations that accumulate and lead to the process of hepatocarcinogenesis either due to above risk factors or etiologically non-specific, sporadic causes are less well understood. Nevertheless, genetic changes activating the Wnt signaling pathway and mutations in tumor suppressor proteins p53, Rb and p16 have been observed in a majority of the HCC cases. The Wnt signaling cascade alone is aberrantly activated in up to 90% of all HCCs indicating its predominant role in liver tumorigenesis. Therefore, every molecule and each step of Wnt pathway should be studied further to elucidate the molecular mechanism of hepatic oncogenesis with respect to this signaling cascade in the hopes of discovering better treatments for this devastating disease of extremely poor prognosis.

1.1.5. REFERENCES

1. Farazi PA and DePinho RA (2006) Hepatocellular carcinoma pathogenesis: from genes to environment. *Nat Rev Cancer*, 6: 674-687.
2. El-Serag HB and Rudolph KL (2007) Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology*, 132: 2557-2576.
3. Gomaa AI, Khan SA, Toledano MB, Waked I and Taylor-Robinson SD (2008) Hepatocellular carcinoma: epidemiology, risk factors and pathogenesis. *World J Gastroenterol*, 14: 4300-4308.
4. Bosch FX, Ribes J, Diaz M and Cleries R (2004) Primary liver cancer: worldwide incidence and trends. *Gastroenterology*, 127: S5-S16.
5. McKillop IH, Moran DM, Jin X and Koniaris LG (2006) Molecular pathogenesis of hepatocellular carcinoma. *J Surg Res*, 136: 125-135.
6. Schutte K, Bornschein J and Malfertheiner P (2009) Hepatocellular carcinoma--epidemiological trends and risk factors. *Dig Dis*, 27: 80-92.
7. Altekruse SF, McGlynn KA and Reichman ME (2009) Hepatocellular carcinoma incidence, mortality, and survival trends in the United States from 1975 to 2005. *J Clin Oncol*, 27: 1485-1491.
8. Wong CM and Ng IO (2008) Molecular pathogenesis of hepatocellular carcinoma. *Liver Int*, 28: 160-174.
9. World Health Organization, Hepatitis B, World Health Organization Fact Sheet # 204, http://www.who.int/mediacentre/factsheets/fs204/en/

10. Perrillo R, and Nair S (2006) Chapter 31: Hepatitis B. In Zakim and Boyer's Hepatology: A Textbook of Liver Diseases by Boyer TD, Wright TL, and Manns MP, Saunders Elsevier, Philadelphia, Fifth edition, vol. 1, pp. 635-663.

11. Cougot D, Neuveut C and Buendia MA (2005) HBV induced carcinogenesis. *J Clin Virol*, 34 Suppl 1: S75-78.

12. Murakami Y, Saigo K, Takashima H, Minami M, Okanoue T, Brechot C and Paterlini-Brechot P (2005) Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas. *Gut*, 54: 1162-1168.

13. World Health Organization, Hepatitis C, World Health Organization Fact Sheet # 164, http://www.who.int/mediacentre/factsheets/fs164/en/index.html

14. Blonski W and Reddy KR (2008) Hepatitis C virus infection and hepatocellular carcinoma. *Clin Liver Dis*, 12: 661-674, x.

15. Wright TL, and Manns MP (2006) Chapter 32: Hepatits C. In Zakim and Boyer's Hepatology: A Textbook of Liver Diseases by Boyer TD, Wright TL, and Manns MP, Saunders Elsevier, Philadelphia, Fifth edition, vol. 1, pp. 665-686.

16. McCusky RS (2006) Chapter 1: Anatomy of the Liver. In Zakim and Boyer's Hepatology: A Textbook of Liver Diseases by Boyer TD, Wright TL, and Manns MP, Saunders Elsevier, Philadelphia, Fifth edition, vol. 1, pp. 3-21.

17. La Vecchia C (2007) Alcohol and liver cancer. *Eur J Cancer Prev*, 16: 495-497.

18. Bruix J, Bru C, and Llovet JM (2006) Chapter 59: Hepatocellular Carcinoma. In Zakim and Boyer's Hepatology: A Textbook of Liver Diseases by Boyer TD, Wright TL, and Manns MP, Saunders Elsevier, Philadelphia, Fifth edition, vol. 2, pp. 1109-1131.

19. Morgan TR, Mandayam S and Jamal MM (2004) Alcohol and hepatocellular carcinoma. *Gastroenterology*, 127: S87-96.

20. Baffy G (2009) Kupffer cells in non-alcoholic fatty liver disease: the emerging view. *J Hepatol*, 51: 212-223.

21. Stewart SF, and Day CP (2006) Chapter 29: Alcoholic Liver Disease. In Zakim and Boyer's Hepatology: A Textbook of Liver Diseases by Boyer TD, Wright TL, and Manns MP, Saunders Elsevier, Philadelphia, Fifth edition, vol. 1, pp. 579-623.

22. Donato F, Tagger A, Gelatti U, Parrinello G, Boffetta P, Albertini A, Decarli A, Trevisi P, Ribero ML, Martelli C, Porru S and Nardi G (2002) Alcohol and hepatocellular carcinoma: the effect of lifetime intake and hepatitis virus infections in men and women. *Am J Epidemiol*, 155: 323-331.

23. Schiano TD, and Hunt K (2006) Chapter 28: Occupational and Environmental Hepatotoxicity. In Zakim and Boyer's Hepatology: A Textbook of Liver Diseases by Boyer TD, Wright TL, and Manns MP, Saunders Elsevier, Philadelphia, Fifth edition, vol. 1, pp. 561-577.

24. Wands JR, and Moradpour D (2006) Chapter 10: Molecular Pathogenesis of Hepatocellular Carcinoma. In Zakim and Boyer's Hepatology: A Textbook of Liver Diseases by Boyer TD, Wright TL, and Manns MP, Saunders Elsevier, Philadelphia, Fifth edition, vol. 1, pp. 165-175.

25. Laurent-Puig P and Zucman-Rossi J (2006) Genetics of hepatocellular tumors. *Oncogene*, 25: 3778-3786.

26. Neuschwander-Tetri BA (2006) Chapter 55: NASH. In Zakim and Boyer's Hepatology: A Textbook of Liver Diseases by Boyer TD, Wright TL, and Manns MP, Saunders Elsevier, Philadelphia, Fifth edition, vol. 2, pp. 1031-1063.

27. World Health Organization, Obesity and Overweight, World Health Organization Fact Sheet # 311, http://www.who.int/mediacentre/factsheets/fs311/en/index.html

28. Adams PC (2006) Chapter 67: Hemochromatosis. In Zakim and Boyer's Hepatology: A Textbook of Liver Diseases by Boyer TD, Wright TL, and Manns MP, Saunders Elsevier, Philadelphia, Fifth edition, vol. 2, pp. 1239-1256.

29. Bantel H, and Manns MP (2006) Chapter 40: Autoimmune Hepatitis and Overlap Syndromes. In Zakim and Boyer's Hepatology: A Textbook of Liver Diseases by Boyer TD, Wright TL, and Manns MP, Saunders Elsevier, Philadelphia, Fifth edition, vol. 2, pp. 795-801.

30. Talwalkar JA, and Lindor KD (2006) Chapter 41: Primary Biliary Cirrhosis. In Zakim and Boyer's Hepatology: A Textbook of Liver Diseases by Boyer TD, Wright TL, and Manns MP, Saunders Elsevier, Philadelphia, Fifth edition, vol. 2, pp. 803-820.

31. Broome U, and Bergquist A (2006) Chapter 42: Primary Sclerosing Cholangitis In Zakim and Boyer's Hepatology: A Textbook of Liver Diseases by Boyer TD, Wright TL, and Manns MP, Saunders Elsevier, Philadelphia, Fifth edition, vol. 2, pp. 821-854.

32. He TC, Sparks AB, Rago C, Hermeking H, Zawel L, da Costa LT, Morin PJ, Vogelstein B and Kinzler KW (1998) Identification of c-MYC as a target of the APC pathway. *Science*, 281: 1509-1512.

33. Tetsu O and McCormick F (1999) Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. *Nature*, 398: 422-426.

34. Monga SP (2009) Role of Wnt/beta-catenin signaling in liver metabolism and cancer. *Int J Biochem Cell Biol*.

35. Aravalli RN, Steer CJ and Cressman EN (2008) Molecular mechanisms of hepatocellular carcinoma. *Hepatology*, 48: 2047-2063.

36. Lee HC, Kim M and Wands JR (2006) Wnt/Frizzled signaling in hepatocellular carcinoma. *Front Biosci*, 11: 1901-1915.

37. Wong CM, Fan ST and Ng IO (2001) beta-Catenin mutation and overexpression in hepatocellular carcinoma: clinicopathologic and prognostic significance. *Cancer*, 92: 136-145.

38. Legoix P, Bluteau O, Bayer J, Perret C, Balabaud C, Belghiti J, Franco D, Thomas G, Laurent-Puig P and Zucman-Rossi J (1999) Beta-catenin mutations in hepatocellular carcinoma correlate with a low rate of loss of heterozygosity. *Oncogene*, 18: 4044-4046.

39. Edamoto Y, Hara A, Biernat W, Terracciano L, Cathomas G, Riehle HM, Matsuda M, Fujii H, Scoazec JY and Ohgaki H (2003) Alterations of RB1, p53 and Wnt pathways in hepatocellular carcinomas associated with hepatitis C, hepatitis B and alcoholic liver cirrhosis. *Int J Cancer*, 106: 334-341.

40. Cieply B, Zeng G, Proverbs-Singh T, Geller DA and Monga SP (2009) Unique phenotype of hepatocellular cancers with exon-3 mutations in beta-catenin gene. *Hepatology*, 49: 821-831.

41. Terris B, Pineau P, Bregeaud L, Valla D, Belghiti J, Tiollais P, Degott C and Dejean A (1999) Close correlation between beta-catenin gene alterations and nuclear accumulation of the protein in human hepatocellular carcinomas. *Oncogene*, 18: 6583-6588.

42. Nhieu JT, Renard CA, Wei Y, Cherqui D, Zafrani ES and Buendia MA (1999) Nuclear accumulation of mutated beta-catenin in hepatocellular carcinoma is associated with increased cell proliferation. *Am J Pathol*, 155: 703-710.

43. de La Coste A, Romagnolo B, Billuart P, Renard CA, Buendia MA, Soubrane O, Fabre M, Chelly J, Beldjord C, Kahn A and Perret C (1998) Somatic mutations of the beta-catenin gene are frequent in mouse and human hepatocellular carcinomas. *Proc Natl Acad Sci U S A*, 95: 8847-8851.

44. Taniguchi K, Roberts LR, Aderca IN, Dong X, Qian C, Murphy LM, Nagorney DM, Burgart LJ, Roche PC, Smith DI, Ross JA and Liu W (2002) Mutational spectrum of beta-catenin, AXIN1, and AXIN2 in hepatocellular carcinomas and hepatoblastomas. *Oncogene*, 21: 4863-4871.

45. Ban KC, Singh H, Krishnan R and Seow HF (2003) GSK-3beta phosphorylation and alteration of beta-catenin in hepatocellular carcinoma. *Cancer Lett*, 199: 201-208.

46. Merle P, de la Monte S, Kim M, Herrmann M, Tanaka S, Von Dem Bussche A, Kew MC, Trepo C and Wands JR (2004) Functional consequences of frizzled-7 receptor overexpression in human hepatocellular carcinoma. *Gastroenterology*, 127: 1110-1122.

47. Merle P, Kim M, Herrmann M, Gupte A, Lefrancois L, Califano S, Trepo C, Tanaka S, Vitvitski L, de la Monte S and Wands JR (2005) Oncogenic role of the frizzled-7/beta-catenin pathway in hepatocellular carcinoma. *J Hepatol*, 43: 854-862.

48. Kim M, Lee HC, Tsedensodnom O, Hartley R, Lim YS, Yu E, Merle P and Wands JR (2008) Functional interaction between Wnt3 and Frizzled-7 leads to activation of the Wnt/beta-catenin signaling pathway in hepatocellular carcinoma cells. *J Hepatol*, 48: 780-791.

49. Takagi H, Sasaki S, Suzuki H, Toyota M, Maruyama R, Nojima M, Yamamoto H, Omata M, Tokino T, Imai K and Shinomura Y (2008) Frequent epigenetic inactivation of SFRP genes in hepatocellular carcinoma. *J Gastroenterol*, 43: 378-389.

50. Menendez D, Inga A and Resnick MA (2009) The expanding universe of p53 targets. *Nat Rev Cancer*, 9: 724-737.

51. Tannapfel A and Wittekind C (2002) Genes involved in hepatocellular carcinoma: deregulation in cell cycling and apoptosis. *Virchows Arch*, 440: 345-352.

52. Hussain SP, Schwank J, Staib F, Wang XW and Harris CC (2007) TP53 mutations and hepatocellular carcinoma: insights into the etiology and pathogenesis of liver cancer. *Oncogene*, 26: 2166-2176.

53. Giacinti C and Giordano A (2006) RB and cell cycle progression. *Oncogene*, 25: 5220-5227.

54. Foulstone E, Prince S, Zaccheo O, Burns JL, Harper J, Jacobs C, Church D and Hassan AB (2005) Insulin-like growth factor ligands, receptors, and binding proteins in cancer. *J Pathol*, 205: 145-153.

55. Breuhahn K, Longerich T and Schirmacher P (2006) Dysregulation of growth factor signaling in human hepatocellular carcinoma. *Oncogene*, 25: 3787-3800.

56. Ito T, Sasaki Y and Wands JR (1996) Overexpression of human insulin receptor substrate 1 induces cellular transformation with activation of mitogen-activated protein kinases. *Mol Cell Biol*, 16: 943-951.

57. Tanaka S, Mohr L, Schmidt EV, Sugimachi K and Wands JR (1997) Biological effects of human insulin receptor substrate-1 overexpression in hepatocytes. *Hepatology*, 26: 598-604.

58. Hwang YH, Choi JY, Kim S, Chung ES, Kim T, Koh SS, Lee B, Bae SH, Kim J and Park YM (2004) Over-expression of c-raf-1 proto-oncogene in liver cirrhosis and hepatocellular carcinoma. *Hepatol Res*, 29: 113-121.

59. Lee HC, Tian B, Sedivy JM, Wands JR and Kim M (2006) Loss of Raf kinase inhibitor protein promotes cell proliferation and migration of human hepatoma cells. *Gastroenterology*, 131: 1208-1217.

60. Calvisi DF, Ladu S, Gorden A, Farina M, Conner EA, Lee JS, Factor VM and Thorgeirsson SS (2006) Ubiquitous activation of Ras and Jak/Stat pathways in human HCC. *Gastroenterology*, 130: 1117-1128.

PART 2

1.2. WNT SIGNALING

1.2.1. WNT SIGNALING: OVERVIEW

The Wnt signaling network is involved in almost every aspect of embryonic development as well as in homeostatic self-renewal of adult tissues [1]. As a result, genetic alterations in the Wnt pathway lead to deleterious effects: while somatic mutations are associated with numerous types of cancer and a variety of degenerative disorders, germline mutations are implicated in hereditary diseases [2,3]. The term "Wnt" was coined in the late 1980s as a combinatorial nomenclature for *Wingless* and *Int-1* [4], when the *Drosophila Wingless* gene, which controlled the segment polarity in larval development [5], was demonstrated to be the fruit fly homolog of mammalian *Int-1* proto-oncogene [6], which was initially identified in the mouse genome as a target integration site for the Mouse Mammary Tumor Virus (MMTV) [7]. Over the past two decades, the knowledge of Wnt signaling has been advanced tremendously by the discovery of its molecular pathways and their main regulators and mediators [8]. However, despite these achievements, many fundamental molecular events are still unclear and remain to be elucidated.

Many components of the Wnt signaling are conserved throughout the animal kingdom. In humans and mice, there are 19 Wnt ligands, which are a large family of small secreted cysteine-rich glycoproteins of 350 – 380 amino acids that act as short- and long-range morphogens to activate receptor-mediated signaling pathways [9]. Activation of Wnt genes is regulated in a precise temporal and spatial manner as they exhibit unique expression patterns and distinct functions during development [1]. While Wnt signaling is quiescent in adult resting liver [10], four Wnt ligands – Wnt3, Wnt5a, Wnt6, and Wnt11

– have been found to activate Wnt signaling in HCC cell lines, HepG2, Hep3B, Huh7, and FOCUS determined by RT-PCR using specific primer pairs to all 19 Wnt genes [11].

Wnt signaling is activated when the Wnt ligands bind to the receptors of the Frizzled (Fzd or Fz) family of proteins at the extracellular surface of a stimulated cell. Fzd receptors are seven transmembrane proteins that belong to the large family of G-protein coupled receptors (GPCRs) [12]. As with GPCRs, Fzd receptors are found in homomeric and heteromeric complexes with other members of the Fzd family [13]. All Fzd receptors have a highly conserved, ligand-binding, cysteine-rich domain, which is followed by a linker region, a seven-pass transmembrane domain, and a C-terminal, cytoplasmic region for the receptor signaling (Figure 1.2.1) [14]. So far, 10 Fzd receptors have been identified in humans and mice, and the Wnt glycoproteins can bind to multiple Fzd receptors and vice versa [15]. Although Fzd proteins are the primary receptors for Wnt ligands, Fzd and Wnt can interact with other partners to activate alternative signaling cascades or to inhibit the Wnt pathway [13,16,17]. Nevertheless, the Wnt signaling is initiated only when a Wnt ligand binds to a Fzd receptor, and it is categorized into two groups: canonical and non-canonical Wnt signaling pathways. In the canonical Wnt signaling, which is also called Wnt/β-catenin pathway, expression of Wnt target genes is regulated by the availability of free β-catenin molecules in the cytoplasm and nucleus. The non-canonical Wnt signaling, which is also called the β-catenin-independent pathway, is activated by at least two mechanisms, each overlapping with other signaling pathways [18,19]. First, the planar cell polarity (PCP) pathway (Wnt/PCP), which regulates tissue polarity, cell migration and cytoskeleton arrangement, activates small GTPases (Rho and Rac) and JNK pathways. Second, the calcium signaling pathway

(Wnt/Ca) inhibits the canonical Wnt signaling and promotes cell migration and cell proliferation by activating calcium/calmodulin-dependent protein kinase (CaMK) II and protein kinase C (PKC) enzymes.

The Wnt ligands are classified into two categories according to the pathways involved [20], and historically based on their ability to transform the mouse mammary epithelial cell line C57MG [21,22]. The Wnt proteins of the transforming class are Wnt1, Wnt2, Wnt2b, Wnt3, Wnt3a, Wnt7a, Wnt8a, Wnt8b, Wnt9a (Wnt14), Wnt10a and Wnt10b, and they predominantly activate canonical Wnt signaling [8,20,23]. On the other hand, the Wnt ligands of the non-transforming class, which are also called the β-catenin-independent or non-canonical Wnts, are Wnt4, Wnt5a, Wnt5b, Wnt6, Wnt7b, Wnt9b (Wnt15), Wnt11 and Wnt16, and they mainly induce cell polarity, asymmetric cell division, and morphogenetic movements during gastrulation [8,20]. However, this classification of Wnt ligands has grown difficult as some Wnt proteins, including Wnt3a and Wnt7a, regulate both canonical and non-canonical Wnt pathways [24], and activation of a certain pathway by a Wnt ligand can be dependent on the cellular context and other factors [9]. Nevertheless, several Wnt proteins, such as Wnt1, Wnt3, Wnt5a, and Wnt11, consistently activate one of the two pathways across many cellular and developmental models. Although a certain combination of Wnt and Fzd interaction and cellular context result in the activation of one of the Wnt/β-catenin, Wnt/PCP, and Wnt/Ca pathways, it has been suggested that Wnt proteins activate a complex intracellular signaling network rather than individual pathways since those pathways overlap and cross-talk at many levels of the signaling cascade [25]. Because my thesis work is solely based on the Wnt/β-catenin pathway, I will continue my thesis with the canonical Wnt signaling.

1.2.2. CANONICAL WNT SIGNALING

The main regulatory molecule of the canonical Wnt signaling is β-catenin (Armadillo in *Drosophila*), and the hallmark of activated Wnt/β-catenin pathway is the accumulation of β-catenin protein levels in the cytoplasm and nucleus [1]. In the absence of secreted Wnt ligands or when the canonical Wnt signaling is inhibited by its extracellular antagonists like sFRP, Wnt Inhibitory Factor (WIF), or Dickkopf (Dkk) [26], the unstimulated resting cells are devoid of free β-catenin. Most of the cellular β-catenin is recruited at the adherens junctions by binding to the intracellular domain of E-Cadherins and connecting them to the actin cytoskeleton via α-catenin [27,28]. At the same time, the cytoplasmic free β-catenin, which is constantly synthesized by the cell, is rapidly and continuously degraded through a group of proteins called the destruction complex [29]. The β-catenin destruction complex is composed of the scaffolding protein Axin or its homologue conductin [30,31], the tumor suppressor APC [32,33], and the Ser/Thr kinases Casein Kinase 1α (CK1α; Gilgamesh in *Drosophila*) and Glycogen Synthase Kinase 3β (GSK3β; Shaggy or Zeste-white 3 in *Drosophila*) (Figure 1.2.1A) [29,34,35]. Shortly after its synthesis in the cytosol, β-catenin is sequestered at the destruction complex via Axin and APC [36], where it is phosphorylated at its amino terminus in a C- to N-terminal direction by protein kinases CK1α and GSK3β [37,38]. CK1α phosphorylates β-catenin at Ser45 and primes it for GSK3β [39], which in turn sequentially adds phosphate groups to Thr41, Ser37, and Ser33 [40]. Once phosphorylated, β-catenin is recognized by β-Trcp (β-transducin repeats-containing proteins) subunit of the E3 ubiquitin ligase, which subsequently ubiquitylates and leads to degradation of β-catenin by a 26S proteasome [41]. This continual elimination of β-catenin prevents it from

entering to the nucleus and binding to the DNA-bound transcription factors of the T cell factor/lymphoid enhancer factor (TCF/LEF; Pangolin or dTCF in *Drosophila*) family of proteins, such that the Wnt target genes remain repressed (Figure 1.2.1A). In this context, the TCF/LEF transcription factors act as repressors by forming a complex with well-known repressors Transducin-Like Enhancer of split proteins (TLE; Groucho in *Drosophila*) [42,43] and possibly with C-terminal Binding Proteins (CtBP) [44].

The canonical Wnt signaling is activated when a Wnt ligand binds to a Fzd receptor and its co-receptor Low-Density Lipoprotein (LDL) Receptor-related Protein 6 (LRP6; Arrow in *Drosophila*), or its close relative LRP5 [45,46]. The formation of the trimeric Wnt-Fzd-LRP complex results in the phosphorylation of a downstream molecule Dishevelled (Dvl, or Dsh in *Drosophila*) by CK1ε and CK2 leading to its binding to the cytoplasmic C-terminus of the Fzd receptor [47-49]. At the same time, the cytoplasmic end of the single-pass transmembrane protein LRP is also activated and phosphorylated by membrane-anchored CK1γ and GSK3β [50,51]. The Fzd-bound Dvl, along with the cytoplasmic tail of LRP, recruits the scaffolding protein Axin to the receptors resulting in the dissociation of the destruction complex [52]. In the absence of active destruction complex, free cytosolic β-catenin escapes from phosphorylation and hence recognition by β-Trcp leading to its stabilization and accumulation in the cytoplasm (Figure 1.2.1B). The increased cytosolic β-catenin translocates into the nucleus, where it binds to the DNA-bound TCF/LEF transcription factors by displacing TLE, and promotes the expression of Wnt target genes [53]. The interaction of TCF/LEF proteins with β-catenin switches them from transcriptional repressors to transcriptional activators.

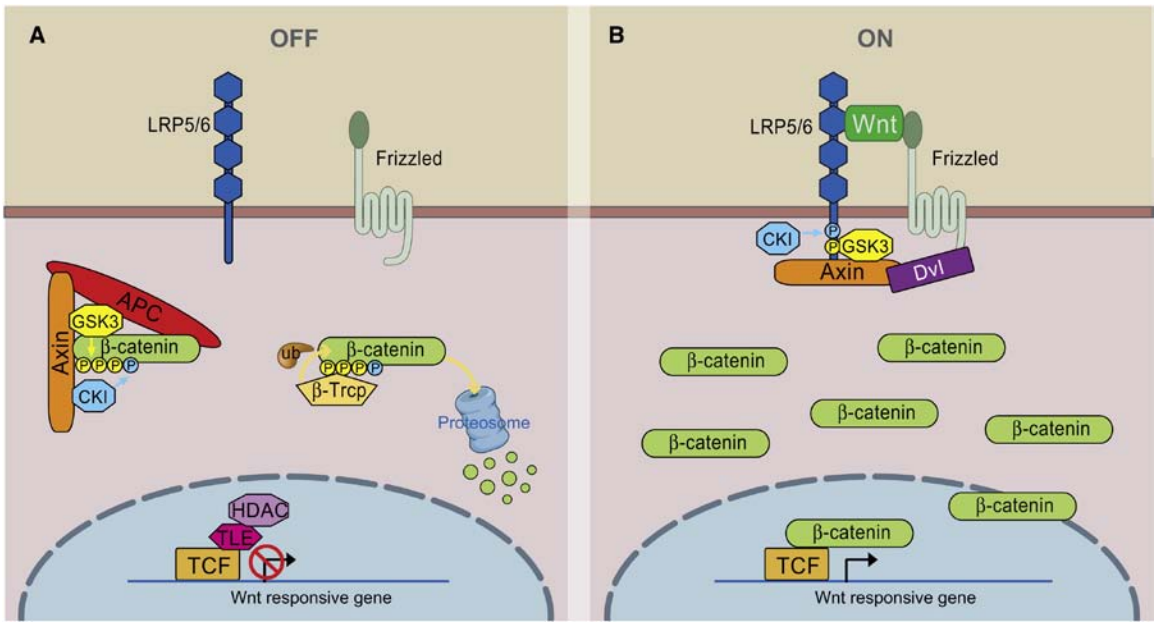


Figure 1.2.1. The canonical Wnt signaling

A. In the absence of Wnt ligand, cytoplasmic β-catenin is captured by the destruction complex composed of Axin, APC, CK1α, and GSK3β, where it is phosphorylated at its amino terminus. Phosphorylated β-catenin is recognized by the E3 ubiquitin ligase β-Trcp, which targets β-catenin for proteasomal degradation. In the nucleus, the Wnt target genes are repressed by TCF/LEFs bound to TLE that recruits other co-repressors, such as HDAC. *B*. In the presence of Wnt signals, formation of Wnt-Fzd-LRP complex recruits Dvl, which in turn along with LRP recruits Axin resulting in the dissociation of the destruction complex. Stabilized β-catenin accumulates and enters the nucleus, where it binds to TCF/LEFs to activate Wnt-responsive genes. Adapted from [46].

1.2.3. T CELL FACTOR/LYMPHOID ENHANCER FACTOR

The transcription factors of the TCF/LEF family, which is comprised of four members in mammals, are the most downstream components of the canonical Wnt signaling pathway (Figure 1.2.2) [54-56]. The founding members, TCF-1 and LEF-1, were identified in screens for lymphocyte-specific transcription factors in differentiating T and B cells. TCF-1 was identified as a T cell-specific nuclear protein that bound to the AACAAAG motif on the enhancer region of the *CD3ε* gene [57]. LEF-1 (formerly TCF-1α) was cloned independently by two different groups as a T cell- as well as a pre-B cell-specific DNA-binding protein with affinity to TTCAAAGG enhancer sequence of the *TCRα* (*T Cell antigen Receptor α*) gene [58,59]. Later, TCF-1 was also demonstrated to bind the enhancer regions of *TCRα, TCRβ,* and *TCRδ* genes recognizing a consensus sequence motif $^{A}/_{T}{}^{A}/_{T}$CAAAGG [60]. Both TCF-1 and LEF-1 contained a virtually identical promoter-binding domain called the High Mobility Group (HMG) box [61,62]. By using low stringency cloning methods, such as "guessmer" PCR with primers based on the conserved HMG box of TCF-1 and LEF-1 [63], the other two members, TCF-3 and TCF-4, were identified and characterized [64,65]. In addition, homologs of these transcription factors have been identified in other species that serve as important experimental model systems for further characterizations. The single orthologs pangolin (or dTCF) in *Drosophila melanogaster* [66,67] and POP-1 in *Caenorhabditis elegans* [68] are homologous to mammalian TCF-1, whereas *Xenopus laevis* expresses four members, named analogously with their mammalian counterparts, XTCF-3 [69], XLEF-1 [70], XTCF-4 [71], and XTCF-1 [72].

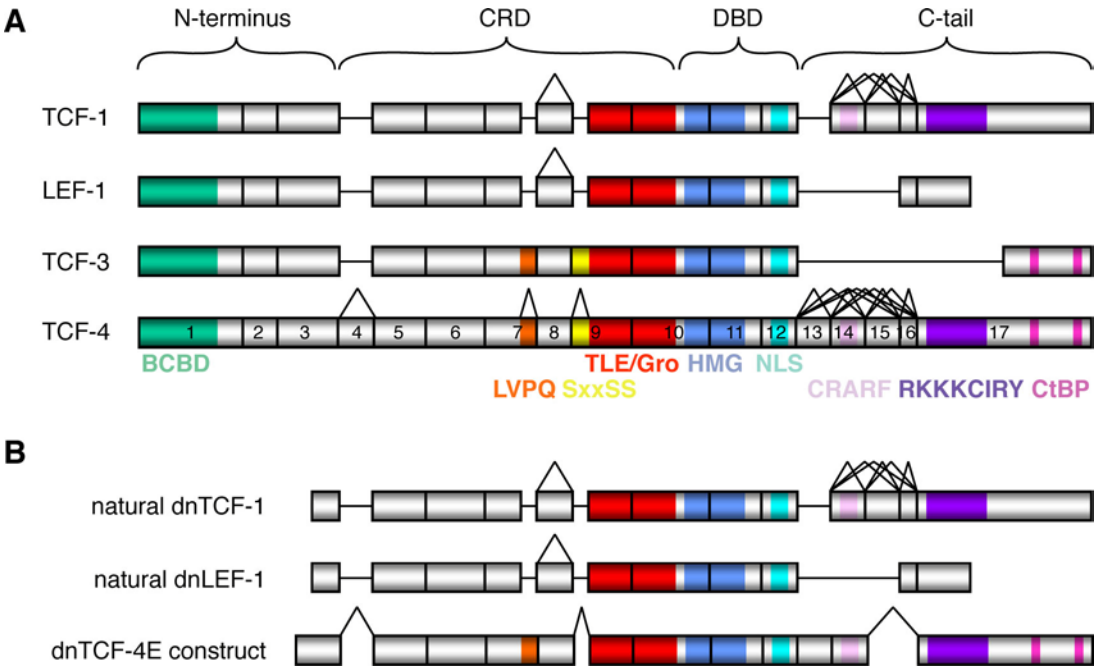


Figure 1.2.2. Structure of TCF/LEFs

A. Schematic representation of human TCF/LEF family of transcription factors. The conserved regions are marked with different colors. BCBD, β-catenin binding domain at the N-terminus; CRD, context-dependent repression domain contains the LVPQ and SxxSS motifs, and TLE/Groucho co-repressor binding site (TLE/Gro); DBD, DNA-binding domain includes the high mobility group (HMG) box and nuclear localization sequence (NLS); and the C-terminal tail harbors the C-clamp (CRARF and RKKKCIRY) and two CtBP-binding sites. Alternatively spliced sites are indicated by upside down "V" shapes. Note that hTCF-3 is not alternatively spliced, and the LVPQ and SxxSS motifs and the two CtBP repressor-binding sites are constitutively expressed in hTCF-3. On the other hand, these sites are absent in TCF-1 and LEF-1, while TCF-4 can generate any combination of these motifs by alternative splicing. In TCF-4, the numbers inside boxes indicate the exon order of the gene (*TCF7L2*). *B*. Natural dominant negative forms of TCF-1 and LEF-1 are formed due to usage of second promoter, which is located at the second intronic sequences. This results in truncated proteins without the BCBDs. No such truncated protein is found in TCF-4. Thus, dnTCF-4E is constructed by deleting the first 90 nucleotides of its gene. Because the second codon for Met is located in exon 3, the translation begins at Met107.

1.2.3.1. Expression of TCF/LEF

As with the other components of the Wnt/β-catenin signaling cascade, while TCF/LEF proteins are ubiquitously expressed in many tissues during embryonic development, their expression is downregulated in most terminally differentiated adult cells. Since *Drosophila* and *C. elegans* express single orthologs, the diversity and expression pattern of TCF/LEF proteins has been mostly obtained from developmental and genetic studies of frogs and mice.

1.2.3.1.1. Expression in Xenopus development

The primary role of the canonical Wnt signaling during embryonic development is to specify the dorsoventral axis immediately after the fertilization [73]. It is required for induction of dorsal development in the future dorsal side of the early embryo [74]. Artificial expression of Wnt signaling components in the opposite ventral side results in the induction of ectopic dorsal development and axis duplication. Extensive research in *Xenopus* development revealed that XLEF-1, XTCF-1 and XTCF-4 function, partially redundantly, on the dorsal side to activate expression of dorsal genes [75,76], while XTCF-3 represses expression of the same genes on the ventral side [77]. However, despite the redundancy, only XLEF-1, which is the only member of *Xenopus* TCF/LEF family not inherited maternally, can induce ectopic dorsal development on the ventral side [76]. In addition, overexpression of XTCF-3 leads to repression of dorsal genes in either side of the axis and hence inhibits dorsal development [78], whereas null mutations of XTCF-3 induces axis duplication, which later results in production of secondary neural tube [79], indicating that XTCF-3 functions as a repressor. Depending on their specific isoform expressed, XTCF-1 and XTCF-4 have also been observed to function as

repressors in the presence of XTCF-3 to inhibit the expression of dorsal genes ventrally and laterally [75,80,81]. Wnt signaling is further activated during gastrulation for mesoderm induction, in neural development for midbrain and hindbrain patterning, and repeatedly throughout embryonic development. In each stage, TCF/LEF transcription factors are utilized in a same pattern: LEF-1 functions to activate Wnt-responsive genes, TCF-1 and TCF-4 have dual roles as activator and repressor, and TCF-3 inhibits the expression of Wnt target genes [78,82-84].

1.2.3.1.2. Expression in mouse development

In situ hybridization studies on dissected murine embryos revealed the expression patterns of TCF/LEF mRNAs in mammalian development [85]. TCF-1 and LEF-1 are ubiquitously expressed in an overlapping pattern throughout the embryo between embryonic days 7.5 – 10.5 [86-88]. The highest expression of both genes is observed in mesoderm and a lower expression level in ectoderm, whereas no apparent expression is detected in endoderm [86]. In addition, their expression occurs along the anteroposterior gradient with the highest levels being expressed at the posterior end. At later stages, the expression pattern diverges and becomes more complex with unique expression sites for each transcript although they partially coincide on days 13.5 and 14.5 [86]. After birth and throughout life, their expression is restricted to T and B lymphocytes. While TCF-1 is expressed only in T lymphocytes, LEF-1 is found in both T and pre-B cells [57-59].

Expression of TCF-3, on the other hand, is detected only during embryonic days 6.5 – 10.5 [65]. While it is expressed strongly throughout the embryo on day 6.5, its expression level declines by day 7.5 with the highest signals remaining in the anterior part of the embryo and none in the posterior section [65,79]. The differential expression

of TCF-3 from TCF-1 and LEF-1 on day 7.5 is reminiscent of their expression pattern during dorsoventral axis formation in *Xenopus* embryo. This observation suggests that TCF-3 is probably repressing activation of posterior-specific genes in the anterior part of the embryo. In fact, TCF-3 has been consistently shown to act as a non-redundant repressor in mammalian system as well [79,89-91]. It is important to note that, however, Wnt stimulation relieves TCF-3 repression not by converting TCF-3 into an activator, but by replacing it with another member of the TCF/LEF family, or even with another transcription factor of a different family [77]. Expression of TCF-3 in mouse embryo gradually disappears and becomes undetectable beyond day 10.5 [65].

Finally, an *in situ* hybridization study demonstrated that TCF-4 is absent in the early embryo and that its expression is first observed in midbrain on embryonic day 10.5 coinciding with LEF-1 signals [65]. High-level expression is mostly restricted to the central nervous system (CNS) until embryonic day 16.5. Apart from the CNS, the intestinal epithelium is the only other embryonic tissue where TCF-4 is expressed. At day 13.5, TCF-4 is detected directly before the transition of gut endoderm into epithelium, and its expression increases during gut development until day 18.5 [85]. Between embryonic days 13.5 – 18.5, the mouse gut endoderm develops by altering the cell types from stratified to columnar epithelium and forms the villi in the lumen of the small intestine. Beyond embryonic day 16.5 and throughout organism's life, two different cellular populations occupy the lumen: the non-cycling differentiated epithelial cells on the villi, and the less differentiated, actively dividing cells in the intervillus regions, which form the future crypts, or the stem cell compartment. Further study shows that

TCF-4 remains expressed in the gut epithelium postnatally and throughout life suggesting its role in tissue homeostasis and stem cell maintenance [64,65,85].

1.2.3.1.3. Expression in knockout mice

In order to further characterize the biological functions of TCF/LEF transcription factors, various knockout mice have been generated. Mice homozygous for a mutation in the *TCF7* gene, which encodes the mammalian TCF-1 protein, are healthy, fertile, and have a normal life span [92]. However, the TCF-1-deficient mice have a smaller thymus, limited number of T cells, and many defects in thymocyte differentiation and expansion implying the importance of TCF-1 in establishment and maintenance of T lymphocytes [92,93]. Apart from the lymphoid system, no other organs are affected by the null mutation, though the mutant mice develop adenomas in the gut and mammary glands months after birth [94]. Remarkably, however, despite the defects in thymocyte development, adult *TCF7*$^{-/-}$ mice are immunocompetent and generate virus-neutralizing antibodies at normal titers [93]. This interesting observation can be explained by the possible redundancy in TCF-1 and LEF-1 function since their expression coincides at various stages of embryogenesis and continues to overlap postnatally though restricted to T lymphocytes (the latter is also expressed in adult pre-B cells). Heterozygous mice have the wild-type phenotype, including a normal T cell development [92].

On the other hand, mice with *LEF1*$^{-/-}$ gene die within two weeks after birth and show defects in many organs normally expressed during their formation in embryo [85,88]. The most striking phenotype is the absence of body hair and whiskers making the *LEF1*-null mice look nude. In addition, their body is half the size of their heterozygous siblings, which are phenotypically same as wild-type mice. Moreover, the

knockout mice lack teeth, mammary glands, and neurons in the midbrain although formation of these organs is initiated at proper embryonic developmental stage and they are undistinguishable from wild-type counterparts at the bud stage. Further examination of LEF-1-deficient embryos reveals an arrest in progression of these buds later in the development around embryonic day 15, confirming the redundant roles of TCF-1 and LEF-1 at early stages of embryogenesis until their divergence [88]. In contrast to *TCF7*$^{-/-}$ mice, no obvious defects in lymphoid cell populations are detected at birth. However, careful examination of perinatal bone marrow reveals that *LEF1* homozygous null mice have reduced number of pro-B cells due to extensive apoptosis though differentiation of B lymphocytes is not affected, demonstrating the importance of LEF-1 in cell proliferation and survival but not for cell differentiation [95]. The overall pattern of these defects suggests an important role for LEF-1 in the formation of several organs and structures [88].

Since TCF-1 and LEF-1 single knockout mutations do not lead to the phenotypes similar to the ones that arise from loss of a Wnt ligand or β-catenin, but result in more limited defects than compared to the ubiquitous expression patterns of each transcription factor during embryogenesis, double mutant mice are generated and examined [87,96]. The double knockout mutation is embryonic lethal at around day 10, and *TCF7*$^{-/-}$ *LEF1*$^{-/-}$ embryos demonstrate early developmental abnormalities. Most of the defects are detected in tissues that should have differentiated from lateral mesoderm and apical ectoderm as the mutant embryos failed to induce somites, limb buds, and tail buds. Instead, at the expense of lateral mesoderm, they develop additional neural tubes, which is an essentially identical phenotype to knockout mutations of other components of the canonical Wnt

signaling pathway [87]. In addition, thymocyte development is severely impaired and completely arrested at early stage of differentiation confirming the redundant functions of TCF-1 and LEF-1 [96].

Similar to the double knockout mutation, loss of *TCF7L1* gene, which encodes the mammalian TCF-3 protein, results in an early embryonic death at around day 11 [79]. In contrast to $TCF7^{-/-}$ $LEF1^{-/-}$ mutation, however, the TCF-3-deficient embryos display anterior truncations and develop somites along the anteroposterior axis, which derives from the lateral mesoderm, confirming the opposite function of TCF-3 from TCF-1 and LEF-1. In addition, consistent with its repressive role, developmental abnormalities of $TCF7L1^{-/-}$ embryos resembled the effects of knockout mutations of the negative regulators of Wnt/β-catenin signaling, Axin and APC. These defects arise due to anteroposterior axis duplication and include defects in neural patterning, and multiplication of nodes, notochord, neural folds and grooves [79].

Mice homozygous for a mutation in the *TCF7L2* gene, which encodes the mammalian TCF-4 protein, are born at Mendelian ratios, and do not show any decreased body weight or other anatomical abnormalities [85,97]. However, they die within 24 hours after birth despite the presence of milk in their stomachs. The only defect of the TCF-4-deficient mice is the altered epithelial organization of the small intestine when compared with that of wild-type or heterozygous littermates, of which the latter is viable and fertile [97]. The intestinal epithelium of these knockout mice lacks the stem cell population and is entirely composed of differentiated, non-dividing cells confirming the involvement of TCF-4 in tissue homeostasis and stem cell maintenance in the crypts of the intestine. In addition, the lumen is filled with decreased number of villi, and the

intervillus regions are populated with reduced amount of non-cycling epithelial cells, which result in extensive stretching and occasional tearing of the epithelial layer. Aside from the small intestine, all other organs, including CNS, do not show any differences or defects in *TCF7L2*$^{-/-}$ mice.

Finally, mice with compound homozygous mutations in the *TCF7* and *TCF7L2* genes are generated to further examine the roles of TCF-1 and TCF-4 in embryogenesis. The double knockout mutation is embryonic lethal at around day 14, and the *TCF7*$^{-/-}$ *LEF1*$^{-/-}$ embryos show major abnormalities in the posterior part of the organism [98]. They lack hind limbs, posterior body, tail, genital tubercle, and hindgut. However, the forelimbs and most internal organs including lung, heart, pancreas, liver, and even kidney, a posterior organ, appear normal, indicating that TCF-1 and TCF-4 are not required for early development and morphogenesis of these organs. However, the intestine, which derives from the posterior endoderm, is severely truncated and blindly ended, while additional stomach tissues are formed at the intestine-stomach boundary, indicating a redundant role for TCF-1 and TCF-4 in patterning of gut tube by restricting growth of stomach but promoting development of midgut and hindgut [94,97]. Similar to other embryonic lethal mutants, *TCF7*$^{-/-}$ *TCF7L2*$^{-/-}$ embryos show duplication or aberrant branching of neural tube, but unlike *TCF7*$^{-/-}$ *LEF1*$^{-/-}$, they retain lateral mesoderm. Comparison of *TCF7*$^{-/-}$ *LEF1*$^{-/-}$ and *TCF7*$^{-/-}$ *TCF7L2*$^{-/-}$ embryos suggests that TCF-1 and LEF-1 are responsible for specification of lateral mesoderm, while TCF-1 and TCF-4 are required for formation of posterior endoderm; and in the absence of these factors the respective germ layers are not specified but instead neural structures form by default [87,98].

1.2.3.1.4. Expression in stem cells

The canonical Wnt signaling cascade regulates cell fate, cell proliferation *vs*. differentiation, cell survival *vs*. apoptosis, cell behavior and migration during embryonic development. However, the pathway is strongly downregulated after birth in most terminally differentiated tissues, except at sites where continual cell growth and differentiation is provided from stem cell populations. The stem cells are unique cells characterized by the ability to self renew themselves with their proliferative potential and to differentiate into various cell types with their developmental plasticity. Wnt/β-catenin signaling has been shown to play important roles in maintaining the homeostatic relationships between stem cell self renewal and cell differentiation in many tissues including the hematopoietic system (myeloid and lymphoid lineages), the skin (hair follicle and epidermis), and the gastrointestinal tract (intestine and colon) [99]. Accordingly, postnatal expressions of TCF/LEF transcription factors have been detected in thymus (TCF-1, LEF-1) [92], bone marrow (LEF-1) [100], skin (LEF-1, TCF-3) [101], hair follicles (LEF-1) [102], intestinal mucosa (TCF-1, TCF-4) and colon crypts (TCF-4) [65]. In addition, TCF-3 has recently been identified as the transcription factor that co-occupied and inhibited the promoters of *Oct4* and *Nanog* genes, whose protein products regulated the pluripotency of embryonic stem cells [89].

The hematopoietic stem cells (HSC) located in the bone marrow are multipotent cells that perpetuate themselves as well as give rise to all types of blood cells including myeloid (neutrophils, basophils, eosinophils, macrophages, dendritic cells, erythrocytes, and platelets) and lymphoid lineages (T cells, B cells, and Natural Killer (NK) cells). During embryonic development, the HSCs are derived from mesoderm, whose formation

and patterning is regulated, in part, by the Wnt signaling cascade, suggesting an essential role for the Wnt pathway in HSC commitment, self renewal, and differentiation [103]. Indeed, activation of Wnt signaling by overexpressing Wnt3a or Wnt5a, or by virally transducing a constitutively active form of β-catenin, results in self renewal of murine HSCs *in vitro* and reconstitutes the hematopoietic system of lethally irradiated mice *in vivo* [99]. Likewise, inhibition of Wnt signaling by overexpressing Axin or the cysteine-rich, ligand-binding domain of the Fzd receptor suppresses the growth of HSCs *in vitro* and inhibits the reconstitution *in vivo* [100,103]. Finally, actual activation of Wnt signal transduction in HSCs is observed *in vivo* upon endogenous Wnt stimulation when HSCs containing TCF/LEF transcription reporter construct driving expression of green fluorescent protein (GFP) are transplanted into the irradiated mice [100]. In addition to these observations that show involvement of Wnt pathway in HSC and progenitor cell self renewal, Wnt signaling is also required for hematopoiesis and lymphopoiesis. As described above, TCF-1 and LEF-1 were first identified in the lymphoid cells, and their expression is required for the development of T cells in the thymus and the maturation of B lymphocytes in the bone marrow. Similarly, the combined presence and redundant activities of TCF-1 and LEF-1 are also essential in the development of NK cells [74]. In each stage of cell maturation, a specific isoform of TCF-1 and LEF-1 is expressed to carefully regulate the development of these lymphocytes. The role of Wnt signaling in myelopoiesis is not known.

1.2.3.1.5. Expression in cancer and other human diseases

Given the fact that Wnt/β-catenin signaling regulates important developmental processes and plays an essential role in stem cell maintenance, it is conceivable that this pathway is

often taken advantage of by cancers and other human diseases. Loss-of-function mutations in the negative regulators of the Wnt signaling and gain-of-function alterations in other components, which lead to stabilization of β-catenin, are often observed in many tumors, especially of gastrointestinal origin, such as colorectal carcinoma (CRC), HCC, hepatoblastoma, adenocarcinoma, gastric adenoma, and pancreatic carcinoma, as well as in other cancers including various hematological malignancies and skin cancers, hair follicle tumor, renal cell carcinoma (RCC), Wilms tumor, prostate cancer, and ovarian, endometrial, and thyroid carcinomas [2,8,104-106]. Moreover, many hereditary mutations in the pathway also give rise to numerous though rare genetic disorders, including loss of four limbs (Tetra-amelia), osteoporosis, focal dermal hypoplasia, vascular defects in the eye, coronary artery disease, and familial tooth agenesis [46,105]. Finally, the *TCF7L2* (TCF-4) gene has recently been implicated in type 2 diabetes through genome-wide association studies [107]. A common variant of *TCF7L2* is identified to have the strongest effect on disease risk among all genes studied, and this notion has been verified on tens of thousands of individuals in numerous populations [108]. However, because this polymorphism occurs in the non-coding intronic region of the gene, it does not change the protein sequence and alter the activity of TCF-4. Therefore, its exact molecular mechanism is unclear, although the predisposing *TCF7L2* variant is suggested to promote a decrease in insulin secretion in insulin-producing pancreatic β cells but not to cause insulin resistance [109].

The first link between tumor formation and Wnt signaling was established when APC, which was always mutated in a hereditary form of colon cancer called familial adenomatous polyposis (FAP), was discovered to be one of the binding partners of β-

catenin and hence a component of Wnt signal transduction [32,33]. In fact, not only the tumors from FAP patients, but also up to 85% of sporadic CRCs turned out to have inactivating mutations in both alleles of the *APC* gene [110]. In CRC, these mutations result in the production of truncated APC protein that is no longer able to interact with and sequester β-catenin in the destruction complex. Therefore, loss of tumor suppressor function of APC leads to stabilization of β-catenin and constitutive transcriptional activation of Wnt target genes by TCF/β-catenin complex [64]. In the rest of the CRC cases, where *APC* is not mutated, gain-of-function mutations in *CTNNB1* (β-catenin) gene involving the N-terminal phosphorylation sites as well as inactivating mutations in *Axin* genes are mostly observed [106,111].

In HCC, stabilization, accumulation, and aberrant activation of β-catenin are detected in up to 90% of cases depending on the study (Section 1.1.3.1, page 16). Because β-catenin itself is mutated in 25% – 35% of HCCs, alterations in other components of the Wnt pathway, such as Axin, may contribute to the stabilization and accumulation of β-catenin. However, APC is rarely mutated in HCC. Interestingly, the highest rate of activating mutation in *CTNNB1* gene occurs in hepatoblastomas [10], the most common primary liver tumor of childhood. Irrespective of upstream components of Wnt pathway and their loss- or gain-of-function mutations, the highly activated β-catenin would require an equally upregulated amount of TCF/LEF transcription factors in order to exert its carcinogenic effect. In this regard, only TCF-4 has been reported to be overexpressed and bound with those aberrantly accumulated β-catenin molecules in HCC [104,112-114]. No mutations have been observed in *TCF7L2* gene. However, a deletion mutation in the poly A tract of *TCF7L2* gene (see Section 2.2) has been reported in a

subset of CRC cells [115]. This mutation causes a frameshift in the coding sequence so that potential CtBP-binding sites are eliminated from translation, suggesting a mechanism where cancerous cells preferentially produce TCF-4 isoforms that lack the co-repressor-binding sites. In addition, a positive feedback loop, where TCF-1 and LEF-1 are expressed as Wnt targets, is often exploited by CRC cells such that these two members of the TCF/LEF family are overexpressed to interact with the upregulated β-catenin molecules and carry out the transcription of more target genes [74]. It is not known if this mechanism also exists in HCC.

1.2.3.2. Protein structure of TCF/LEFs

All members of the TCF/LEF family of transcription factors consist of four functional domains (Figure 1.2.2): an almost identical DNA-binding domain, a well conserved β-catenin-binding domain, a central domain, and an alternatively spliced C-terminal tail [91].

1.2.3.2.1. DNA-binding domain

The 88 amino acid-long DNA-binding domain (DBD), which harbors a 68 amino acid HMG box and a 9 amino acid Nuclear Localization Sequence (NLS) separated by a nine-residue linker sequence, is the most highly conserved feature of TCF/LEF family with 93% – 99% identity [56]. Because the TCF/LEF proteins were identified through their near identical, sequence-specific HMG box, they are now categorized as a sub-family of DNA-binding proteins of the HMG box superfamily [116]. Upon nuclear import via their NLS, TCF/LEF proteins recognize the DNA consensus sequence $^{A}/_{T}{}^{A}/_{T}$CAAAGG, now called the Wnt Responsive Element (WRE), in the promoter or enhancer regions of the Wnt target genes [57-59]. TCF/LEFs bind at WRE, located in the minor groove of the

double helix, and bend it 90° – 130° away from the protein [117-119]. This binding and bending of the promoter region alters the local chromatin structure by widening the minor groove to allow binding of other transcriptional regulators, either co-activators or co-repressors [119]. At the same time, the flexible nine-residue linker allows the NLS to reach underneath WRE into the curvy major groove and to make contacts with the phosphate backbone of the double helix to increase the DNA-binding affinity [120]. Since the DNA-binding domain is almost identical among the family members, it has been assumed that all TCF/LEF proteins can bind identically and equivalently to any WRE.

1.2.3.2.2. Beta-catenin-binding domain

The second most conserved region of TCF/LEF family is the 50 amino acid-long β-catenin-binding domain (BCBD) at the extreme N-terminus. In human TCF-1 (hTCF-1), hTCF-3, and hTCF-4, respectively, BCBD is 62%, 44%, and 62% identical compared to hLEF-1 [56]. Soon after their discovery as HMG box-containing transcription factors, TCF/LEF proteins were recognized as the β-catenin-binding partners through their N-terminus and acknowledged as essential components of the Wnt signaling by independent yeast two hybrid screens [65-67,76]. The BCBD is entirely encoded in exon 1 of every TCF/LEF gene studied so far. While TCF/LEFs bind β-catenin by amino acids 1 – 50 [121], they can also interact with a closely related cell adhesion molecule, γ-catenin (or plakoglobin), by amino acids 51 – 80 [122]. Although this interaction is weak, the functional role of γ-catenin in the canonical Wnt signaling is controversial as some claim it to activate transcription of Wnt target genes in the absence of β-catenin, while others report it to inhibit TCF/β-catenin-mediated transcription [122,123].

1.2.3.2.3. Central domain

The overall function of the central domain, or the Context-dependent transcription Repression Domain (CRD), is uniquely conserved for each TCF/LEF protein even though the amino acid sequences are not (15% – 20% identity) [56]. Towards the end of the long stretch of the CRD, which contributes to the specific properties of the individual family members, there is a repressor-binding site (Figure 1.2.2). Owing to this region, TCF/LEF proteins act as transcription repressors in the absence of β-catenin. The repressors that bind to this region are the members of the TLE/Groucho family of transcriptional co-repressors [42,43]. During the yeast two hybrid screens that established TCF/LEFs as bona fide components of the Wnt signaling, large number of clones were identified as Groucho proteins when hTCF-1 was used as a bait [67], and the binding of Groucho occurred between the BCBD and the HMG box around amino acids 176 – 359 of hTCF-1 [43].

There are two additional small motifs, referred by their amino acid sequences LVPQ and SFLSS (or SxxSS), in the central domain of TCF-3 and TCF-4 [81]. Although exact mechanism is not known, these motifs have been characterized for transcriptional repression [80]. Despite the close proximity of SxxSS motif with TLE-binding domain, neither of them influenced one another during transcriptional activity [81]. While LVPQ and SxxSS are constitutively expressed in hTCF-3, they can be alternatively spliced out from hTCF-4 proteins (Figure 1.2.2). In addition, the region flanked between these two motifs has been recognized as another repressor binding site making the CRD strong repression domain [124]. However, this same region, which can be spliced out from hLEF-1 and hTCF-1 but not from hTCF-3 and hTCF-4, is also reported to be a

transcription-activating element arguing to rename the CRD acronym from Context-dependent transcription Repression Domain to Context-dependent transcription *Regulatory* Domain [80].

1.2.3.2.4. C-terminal tail

The C-terminal tail is the most varied region among the family members. One of the fascinating features of TCF/LEF proteins is the use of alternatively spliced exons in the C-terminus and the CRD to generate hundreds of theoretical isoforms (Figure 1.2.2). Except for hLEF-1, alternative splicing at the carboxy-terminus gives rise to a variety of C-terminal tails with or without three conserved motifs: the CRARF motif, the RKKKCIRY motif, and two binding sites for CtBP transcriptional co-repressors [74]. While hTCF-4 can generate different combinations of all three motifs in one isoform, hTCF-1 produces only the CRARF and RKKKCIRY, and hTCF-3 has only the CtBP binding sites. Although not settled, the CRARF and RKKKCIRY motifs have been recognized to have transcriptional activation function [125]. The CRARF domain is proposed to be a β-catenin-dependent transcription co-activator-binding site for histone acetyltransferase (HAT) p300 [126]. The 30 amino acid-long region that harbors both CRARF and RKKKCIRY, which are together termed as C-clamp motif, are suggested to be a WRE-independent, secondary DNA-binding domain that gives specificity in target gene recognition [127]. For example, only the long isoforms of TCF-1 and TCF-4 with the C-clamp motif can bind and activate all four *TCF/LEF* genes as Wnt targets. The two CtBP-binding sites with consensus sequence PLSLxxK have been identified through yeast two hybrid screens [78]. Binding of CtBPs on TCF-3 and TCF-4 to repress transcription has been well established [78,128]. It is interesting to note that TCF-1 and

LEF-1 do not express any of the LVPQ, SxxSS, and PLSLxxK motifs, while TCF-3 harbors all of these transcription repression-specific sites as obligate sequences. The fact that TCF-3 is required for target gene repression during embryonic development and that its repression is relieved by replacing it with TCF-1 or LEF-1 could be explained by their protein structure [83]. Interestingly, the TCF-4 mRNA is theoretically able to generate any isoform with or without any of those motifs, rendering it capable of producing different splicing variants that are structurally similar to the other members of the family.

During the initial cloning of TCF-1 by the Hans Clevers laboratory, four isoforms were identified and named alphabetically as TCF-1A through TCF-1D [57,62]. The C-terminal tail of these isoforms was short and similar to that of LEF-1. Later, the same lab isolated another alternative splice variant of TCF-1, but it had a long C-terminus with CRARF and RKKKCIRY motifs [129]. To be consistent with their nomenclature order, they labeled this isoform as TCF-1E. Next, when the very first isoform of TCF-4 was cloned and turned out to harbor a long C-terminus, it was named TCF-4E (see Figure 2.1B, p. 93) to conserve the domain homology [64]. Since then, because TCF-1E and TCF-4E have been known as the only long isoforms (until my thesis research), TCF/LEFs with long C-terminus have been called to have an E-tail or Extended C-terminal domain. TCF-3 also has an extended C-terminal tail, but because it is the only isoform identified so far, *TCF7L1* gene is believed not to possess alternatively spliced exons. Likewise, TCF/LEF transcription factors with short C-terminus are recognized to have B-tails as LEF-1B, TCF-1B, and TCF-4B are the most commonly expressed short isoforms in their respective family members.

1.2.3.2.5. Dominant negative TCF/LEF

Another interesting feature of the TCF/LEF family is the use of alternative promoters to generate N-terminally truncated isoforms. The mammalian *TCF7* and *LEF1* genes have two different promoter regions for their transcription [129,130]. The first promoter, which is located upstream of exon 1 of each gene, produces mRNA that encodes full-length protein with an intact N-terminus. However, the second promoter found in the second intronic sequences results in a shorter mRNA such that an alternative downstream translation start codon is used to produce truncated TCF-1 and LEF-1 proteins without the BCBDs [129,130]. In the absence of N-terminus, they cannot recruit β-catenin and mediate Wnt signals. However, because other domains of these truncated forms are intact and functional, they bind to the WREs and constitutively assemble repressor complexes even during active Wnt signaling and presence of stabilized β-catenin. Therefore, these truncated proteins are known as natural dominant negative forms, and the expression of dnTCF-1 and dnLEF-1 in cells is balanced with their respective full-length isoforms to effectively regulate transcription of Wnt target genes [56].

Despite extensive search, alternative promoter usage and expression of natural dominant negative forms for TCF-3 and TCF-4 have not been observed. Since TCF-3 functions as a repressor [77-79,89-91], a dominant negative version of TCF-3 has never been inquired. Many laboratories attempted, but failed to find a second promoter sequence in *TCF7L2* gene. However, in my opinion, deficiency of natural dnTCF-4 could be explained by the fact that TCF-4 mRNA may produce, by alternative splicing, an isoform homologous to TCF-3, which can fulfill the repressive function of dnTCF-4. Nevertheless, scientists have constructed a dnTCF-4E isoform for research purposes

[131] since overexpression of dnTCFs, hence its inhibitory functions, have been very helpful in exploring the activity of Wnt signals [67,69]. The dnTCF-4E construct was developed by deleting the first 90 nucleotides of TCF-4E mRNA, such that the translation begins from exon 3 efficiently removing its BCBD. This deletion mutant is commercially available as a transfection plasmid and can be purchased from Millipore (formerly Upstate Biotechnology).

1.2.3.3. Transcriptional activity of TCF/LEFs

Extensive research has demonstrated that the members of the TCF/LEF family do not act as classical transcription factors as they do not contain a transcription activation domain and the DNA-binding property alone is not sufficient to initiate gene expression. Instead, co-activators or co-repressors provide transcriptional regulation, respectively, in the presence or absence of Wnt signals. In this regard, the only function of TCF/LEFs is to select and bind to the target genes until a regulatory co-factor arrives and determines the fate of these genes. Therefore, the effector role of TCF/LEF proteins, transcriptional activation *vs*. transcriptional repression, is established by the amount of co-activators *vs*. co-repressors [132].

1.2.3.3.1. TCF/LEFs as transcriptional activators

The main transcriptional activator of the canonical Wnt signaling is the β-catenin molecule, and promoter activation is only achieved when TCF/LEFs are switched from transcriptional repressors to transcriptional activators by interacting with β-catenin [85]. Encoded by *CTNNB1* (catenin, beta 1) gene, β-catenin protein was originally discovered as a component of adherens junctions, and later positioned in the Wnt signaling cascade when it was found to be the vertebrate homolog of *Drosophila* Armadillo, the segment

polarity gene product of the Wingless signaling [27]. The 92 kDa β-catenin molecule consists of 130 amino acid-long N-terminal domain (NTD), a central region that harbors 12 Armadillo repeats (R1 – R12) each of 42 amino acids, and a C-terminal domain (CTD) of 100 amino acids [133]. While the NTD includes the Ser33, Ser37, Thr41, and Ser45 phosphorylation sites, which are translated from exon 3 of the *CTNNB1* gene, the central repeats mediate binding of β-catenin to α-catenin, E-Cadherin, APC, Axin, and TCF/LEFs [132]. In particular, β-catenin associates with TCF/LEF transcription factors via R3 – R10 repeats [121], which are also part of interaction sites for E-Cadherin. Cells have two pools of β-catenin: a membrane-associated pool of stable proteins that connect E-Cadherins to actin cytoskeleton via α-catenin, and a soluble pool of highly unstable proteins that transduce the canonical Wnt signal [134]. Therefore, the highly conserved central region of β-catenin shares the partner-binding domains between different partners such that the two pools cannot be mixed. For example, β-catenin of adherens junction is not targeted for destruction complex since the Armadillo repeats that associate with the scaffolding proteins, APC and Axin, are already occupied by E-Cadherin. However, when the soluble pool of β-catenin is stabilized in the cytosol due to active Wnt signaling, it must not be recruited to the adherens junctions by α-catenin, which binds to its NTD – R1 region. Instead, a nuclear protein called Bcl-9 (Legless in *Drosophila*) that shuttles in and out of the nucleus binds to the NTD – R4 region of β-catenin, and tethers it to a constitutively nuclear protein Pygopus [135,136]. However, the preference of β-catenin between cell adhesion and Wnt signaling is reflected by competitive binding of α-catenin *vs*. Bcl-9 [134]. Because Bcl-9 and Pygopus are required for TCF/β-catenin-mediated gene activation and exclusively dedicated for the canonical Wnt signaling, these

two proteins are considered as transcriptional co-activators of Wnt signal transduction [136].

The CTD contains a potent transactivation domain that recruits additional co-activators to form a multiprotein complex necessary for transcription initiation during active Wnt signaling [46]. Although the mechanism for transcription activation is not fully elucidated, a number of co-activators have been identified.

1. The chromatin structure of the genome creates an environment that is inaccessible for transcription initiation. Therefore, the transcription factors recruit chromatin remodeling complexes that provide accessibility to the densely packed chromatin of target gene loci [137]. The HAT proteins are one of the remodeling complexes that acetylate the histone tails of nucleosomes, the basic repeating units of chromatin, and ease the chromatin conformation by inhibiting tight packaging interactions between individual nucleosomes. The well known HAT proteins p300 and CBP (CREB Binding Protein) interact with R10 – CTD region of β-catenin to activate select Wnt target genes [138,139]. The HAT activities of p300 and CBP are not required for all Wnt targets and they do not act redundantly [140]. Moreover, additional p300 proteins directly bind to CRARF motif of the long tailed TCF-1 and TCF-4 isoforms to reinforce the overall HAT activities [Hecht [126]. While p300 acetylates all four core histone types in nucleosomes, CBP can acetylate all histones as well as non-histone proteins. In addition, HAT proteins called TRAPP and TIP60 from another family also activate transcription of some other genes by binding to R11 – CTD region of β-catenin [141]. These HATs acetylate H2A and H4 of the nucleosomal histones [140]. The DNA bending

capability of TCF/LEFs facilitates in bringing distant DNA regions together, rendering them accessible to the β-catenin-bound HATs, and leading to histone acetylation of up to 30 kilobase region surrounding the WREs [142].

2. The position of nucleosomes can be rearranged by ATPase-dependent complexes to expose DNA sequences that are entangled and shielded between histones [137]. These chromatin remodeling complexes use the energy generated by ATP hydrolysis for catalyzing conformational changes in the nucleosomes or for sliding the histones along DNA to achieve the final rearrangement. In Wnt signaling, this function is mainly provided by BRG1 of SWI2/SNF2 family, which directly binds to β-catenin at R7 – R12 repeats [140]. Moreover, human ISW1, which is another family of ATPase remodeling complexes, is also found to interact with R11 – CTD region of β-catenin, indicating that both SWI/SNF and ISW1 family of ATPases are employed in Wnt target gene expression [141].
3. Histone methyltransferases (HMTs) catalyze one to three methyl groups to Lys or Arg residues of histone tails to mark for transcriptional activation or repression in a context-dependent manner. The MLL (Mixed Lineage Leukemia) complexes that contain HMT activity interact with β-catenin at R11 – CTD region, and mono-, di-, or tri-methylate H3 histone at Lys4 [141]. The H3K4 methylations are chromatin modifications that signify highly active genes: while the H3K4 tri-methylation is recognized as activation sign at transcription start sites, the H3K4 di-methylation serves as a bookmark for RNA Pol II transcription elongation.
4. Despite the efforts of β-catenin in marking transcription initiation via HMTs, and opening the chromatin structure by HATs, SWI/SNF, and ISW1 complexes, the

RNA Pol II needs to be oriented at the promoters of target genes by a pre-initiation complex (PIC). A multi-subunit complex called Mediator bridges the DNA-bound transcription factors with PIC by directly interacting with both parties [143]. The Mediator subunit 12 (MED12) binds to R12 – CTD of β-catenin and promotes the transcriptional activation [140]. The Pygopus protein that retains β-catenin in the nucleus also interacts with Mediator to enhance efficient recruitment of PIC. In addition, β-catenin also interacts with a PIC subunit, TATA box-binding protein (TBP), either directly via its R11 – CTD region or indirectly via Pontin52 [144,145]. Interestingly, Pontin52, also known as TBP-interacting protein or TIP49, not only binds to β-catenin at R2 – R5 repeats, but also serves as a component of the TRAPP/TIP60 HAT complex. Finally, β-catenin also directly binds to some of the components of the core DNA transcriptional machinery, such as DNA Topoisomerase IIα [146].

Overall, the β-catenin appears to serve as the scaffolding protein that orchestrates the recruitment of the transcriptional co-activators. However, since these cofactors are bulky multiprotein complexes, they cannot simultaneously occupy the R11 – CTD region. Instead, it has been proposed that these co-activators are employed in a sequential manner such that each complex performs its function one at a time in an orderly fashion, and the whole process can be repeatedly cycle to yield sustained Wnt target gene transcription [140].

1.2.3.3.2. TCF/LEFs as transcriptional repressors

In the absence of Wnt signaling all TCF/LEF transcription factors interact with TLE co-repressors to suppress the Wnt target genes [42,43,147]. TLE proteins are global, long-

range co-repressors that are recruited to target promoters by different families of DNA-binding transcription factors of wide range of cell signaling cascades [148]. Once recruited, TLE co-repressors oligomerize at the promoters and exert their repressive function by engaging with the chromatin remodeling complexes or interacting with the core transcriptional machinery. Therefore, TLE proteins are true antagonists of β-catenin and employ exact opposite mechanism of action. In humans, there are four full-length TLE proteins (TLE-1 – TLE-4) and one truncated protein called AES (Amino-terminal Enhancer of Split) that contains only the N-terminus of TLE [147]. TLE proteins bind to TCF/LEF transcription factors by their N-termini, which also facilitate their tetramerization at WREs [132]. The functions of HATs are counteracted with protein complexes called Histone Deacetylases (HDACs), which result in hypoacetylation of histone tails leading to a compressed local chromatin and inhibition of transcription [137]. In resting cells, TCF/LEF-bound TLE interact with HDAC enzyme called Rpd3 [132,148]. However, the truncated AES cannot bind to Rpd3 due to the absence of its C-terminal interaction domain and results in reduced repression since it is able to associate with TCF/LEF to tetramerize with other TLE members. Resultant histone deacetylations by the recruitment of Rpd3 are found to be the predominant mechanism of TLE-mediated transcription repression in Wnt signaling. TLE is also reported to interact with the components of the Mediator complex to interfere with the formation of PIC at WREs [148].

CtBP is another co-repressor that specifically binds to TCF-3 and isoforms of TCF-4 with long C-terminal tails [44,78]. The CtBP proteins dimerize and mediate short-range repression through other transcriptional inhibitors. CtBP is found to interact with

various HDACs and HMTs that specifically methylate the histone tails with repression signals, such as tri-methylation of H3K27 [128].

1.2.3.4. Posttranslational modifications of TCF/LEF

TCF/LEF transcription factors are posttranslationally modified and these covalent modifications help determine their effector roles by disrupting or enabling interactions with the transcriptional co-regulators [46,91,132,149]. They include phosphorylation, acetylation, and SUMOylation, and these modifications are specific to individual members of the TCF/LEF family.

1.2.3.4.1. Phosphorylation

Phosphorylation of TCF/LEFs is an important posttranslational modification with both positive and negative consequences depending on the kinases used and the residues targeted. TCF-3 phosphorylation by CK1ε and LEF-1 phosphorylation by CK2 enhance their binding to β-catenin and diminish LEF-1 binding to TLE [46]. On the other hand, LEF-1 phosphorylation by CK1δ at BCBD disrupts its association with β-catenin [91]. Similarly, TCF-4 phosphorylation by CK2 at Ser60 just downstream of BCBD interrupts its binding to γ-catenin. Finally, the phosphorylation of LEF-1 and TCF-4 by MAPK-related Nemo-like kinase (NLK) at the CRD (exon 6 of TCF-4) leads to disruption of their binding to β-catenin and eventually to their proteosomal degradation [46,150]. These are the phosphorylation events known to date for TCF/LEF proteins although there are a number of potential Ser, Thr, and Tyr residues that can be targeted.

1.2.3.4.2. Acetylation

The CBP HATs, which can acetylate both histone tails and non-histone proteins, are bimodal regulators of the canonical Wnt signaling. On one hand, during active Wnt

signaling, CBPs activate transcription of target genes by binding to β-catenin, acetylating histone tails in the nucleosomes, and loosening the chromatin conformation [138-140]. On the other hand, in the absence of Wnt signals, they inhibit the expression of target genes by directly interacting with TCF/LEF transcription factors [91,132,150]. These CBPs bind to TCF/LEF proteins at their HMG box and acetylate a conserved Lys residue in the BCBD, which reduces the affinity of TCF/LEFs for β-catenin and leads to disruption of their interaction.

1.2.3.4.3. SUMOylation

SUMOylation is a posttranslational modification, where Small Ubiquitin-like Modifier (SUMO**)** proteins are covalently attached to and detached from other proteins [151]. Although the SUMO proteins are structurally similar to ubiquitins and SUMOylation is directed by an enzymatic cascade analogous to that involved in ubiquitylation, SUMO is not used to tag proteins for degradation [149]. The functional consequences of SUMOylation are extremely diverse and they include altered localization, changes in activity, protein-protein interactions, and overall stability of the modified protein [151]. The ~10 kDa SUMO protein modifiers are reversibly added to Lys residues of the substrate proteins with a consensus SUMO accepter sequence of ΨKxE, where Ψ is either Leu, Ile or Val, and x is any amino acid. In TCF/LEF transcription factors, SUMOylation is observed only in TCF-4 and LEF-1 [74]. LEF-1 SUMOylation at a Lys residue in the BCBD results in sequestration of LEF-1 in nuclear bodies and repression of Wnt target genes [46,150]. In contrast, SUMOylation of TCF-4 at TLE-binding site (middle of exon 9) with consensus sequence VKQE enhances the binding of TCF-4 to β-catenin and promotes activation of target genes [150,152].

1.2.4. TARGET GENES

Because the canonical Wnt signaling cascade regulates a wide range of cellular behavior including disparate events like cell proliferation *vs*. differentiation and cell survival *vs*. apoptosis in both embryonic development and adult tissue homeostasis, the Wnt target genes are not only diverse but also cell-specific and context-dependent [2,46]. Therefore, most target genes cannot be regarded as universal targets and used for measuring the activation of Wnt/β-catenin pathway.

A plethora of Wnt target genes have been identified from various organisms by utilizing different techniques, and the full list of these genes is available online at the Wnt homepage (http://www.stanford.edu/~rnusse/wntwindow.html). Most target genes usually contain multiple WREs scattered at large distances from the transcription start sites. The *Drosophila* segment polarity gene, *engrailed*, was the first target gene to be discovered by correlating its spatial expression with the expression of Wingless [153]. Such single target strategy led to the identification of few more genes including the gap junction protein connexin-43 in *Xenopus*. As a more efficient approach, the Wnt targets were detected by comparing the transcriptomes (total RNA) from Wnt stimulated or from tumor cells containing activating mutations in the pathway with those from non-stimulated or normal cells, respectively [153]. This differential analysis of gene expression led to the discovery of numerous Wnt targets, such as c-Myc from CRC cells and glutamine synthetase (GS) from HCC [154,155]. Now, this method is developed into microarray analysis, and today it is the most commonly used method for identification of Wnt target genes. Among others, T-box protein 3 (TBX3), which is a transcriptional repressor involved in the control of cell cycle and senescence, is identified as a direct

target of the canonical Wnt signaling in HCC by microarray analysis [156]. However, the transcriptome-based methods also detect indirect target genes that are regulated on the secondary or even tertiary level of the Wnt targets. Finally, a direct BLAST screening of the genome for the optimal TCF/LEF-binding site $^{A}/_{T}{}^{A}/_{T}$CAAAGG coupled with data verification studies is also used to reveal additional Wnt target genes [153]. For example, Cyclin D1 and matrix metalloproteinase 7 (MMP7) were identified in CRC by this genome-based screening strategy [157-159]. Utilization of both transcriptome-based and genome-based screening methods further delineates the identification of target genes. For example, analysis of the transcriptome profile of hepatoma cell lines using cDNA microarray with more than 15,000 unique, liver enriched gene loci yielded 130 potential Wnt target genes, and the examination of presence of the consensus TCF/LEF-binding site narrowed down to 33 candidate genes [160].

In addition to activating genes involved in cell fate determination, cell-cell interaction, cell proliferation, survival, and migration, the canonical Wnt signaling controls the expression of components of its own pathway. Fzd, LRP6, DKK1, Axin2 (conductin), TCF-1, dnTCF-1, LEF-1, and β-Trcp are often regulated either positively or negatively by a Wnt signal, indicating that feedback control is an important feature of the Wnt/β-catenin pathway [1,8,46]. Wnt-induced expression of DKK1, Axin2, β-Trcp, and dnTCF-1, and suppression of Fzd and LRP6 constitute negative feedback loops that diminish the overall Wnt signaling [56,161]. On the other hand, Wnt induction of TCF-1 and LEF-1 reinforce Wnt signaling by forming positive feedback circuits [1,46]. However, expression of TCF-1 and LEF-1 as their own target is only achieved when the long isoforms of TCF-1 and TCF-4 with the C-clamp motif occupy their promoters [127].

1.2.5. SUMMARY

The evolutionarily well-conserved Wnt signaling network is involved in almost every aspect of embryonic development including cell fate determination, cell proliferation *vs*. differentiation, cell survival *vs*. apoptosis, and cell migration. The pathway is normally downregulated after birth in most terminally differentiated tissues, except in certain stem cells. Deregulated activation of Wnt signaling cascade and overexpression of Wnt target genes have been observed in many cancers, including hepatocellular carcinoma (HCC). The activation of Wnt signaling is characterized by the stabilization and accumulation of β-catenin in the cytoplasm, which leads to its translocation into the nucleus and transcription of Wnt target genes by binding to TCF/LEF family of transcription factors. The family consists of four members in mammals: TCF-1, LEF-1, TCF-3, and TCF-4. Upon binding to TCF/LEFs, β-catenin recruits other co-activators necessary for transcription initiation. On the other hand, in the absence of secreted Wnt ligands, the unstimulated resting cells continually eliminate the cytoplasmic free β-catenin preventing it from entering to the nucleus and binding to the TCF/LEF proteins, such that the Wnt target genes, including c-Myc and cyclin D1, remain repressed. In this context, the TCF/LEF transcription factors act as repressors by forming a complex with well-known repressors TLE, which employs other co-repressors used for chromatin compaction. Therefore, TCF/LEFs can exert functions of both repression and activation of target genes as their interaction with β-catenin during active Wnt signaling displaces the TLE and switches them from transcriptional repressors to transcriptional activators.

1.2.6. REFERENCES

1. Logan CY and Nusse R (2004) The Wnt signaling pathway in development and disease. *Annu Rev Cell Dev Biol*, 20: 781-810.
2. Clevers H (2006) Wnt/beta-catenin signaling in development and disease. *Cell*, 127: 469-480.
3. Moon RT, Kohn AD, De Ferrari GV and Kaykas A (2004) WNT and beta-catenin signalling: diseases and therapies. *Nat Rev Genet*, 5: 691-701.
4. Nusse R, Brown A, Papkoff J, Scambler P, Shackleford G, McMahon A, Moon R and Varmus H (1991) A new nomenclature for int-1 and related genes: the Wnt gene family. *Cell*, 64: 231.
5. Nusslein-Volhard C and Wieschaus E (1980) Mutations affecting segment number and polarity in Drosophila. *Nature*, 287: 795-801.
6. Rijsewijk F, Schuermann M, Wagenaar E, Parren P, Weigel D and Nusse R (1987) The Drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. *Cell*, 50: 649-657.
7. Nusse R and Varmus HE (1982) Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell*, 31: 99-109.
8. Chien AJ, Conrad WH and Moon RT (2009) A Wnt survival guide: from flies to human disease. *J Invest Dermatol*, 129: 1614-1627.
9. Mikels AJ and Nusse R (2006) Wnts as ligands: processing, secretion and reception. *Oncogene*, 25: 7461-7468.

10. Nejak-Bowen K and Monga SP (2008) Wnt/beta-catenin signaling in hepatic organogenesis. *Organogenesis*, 4: 92-99.

11. Kim M, Lee HC, Tsedensodnom O, Hartley R, Lim YS, Yu E, Merle P and Wands JR (2008) Functional interaction between Wnt3 and Frizzled-7 leads to activation of the Wnt/beta-catenin signaling pathway in hepatocellular carcinoma cells. *J Hepatol*, 48: 780-791.

12. Egger-Adam D and Katanaev VL (2008) Trimeric G protein-dependent signaling by Frizzled receptors in animal development. *Front Biosci*, 13: 4740-4755.

13. Angers S and Moon RT (2009) Proximal events in Wnt signal transduction. *Nat Rev Mol Cell Biol*, 10: 468-477.

14. Schulte G and Bryja V (2007) The Frizzled family of unconventional G-protein-coupled receptors. *Trends Pharmacol Sci*, 28: 518-525.

15. Hsieh JC (2004) Specificity of WNT-receptor interactions. *Front Biosci*, 9: 1333-1338.

16. Hendrickx M and Leyns L (2008) Non-conventional Frizzled ligands and Wnt receptors. *Dev Growth Differ*, 50: 229-243.

17. Kikuchi A, Yamamoto H and Kishida S (2007) Multiplicity of the interactions of Wnt proteins and their receptors. *Cell Signal*, 19: 659-671.

18. Veeman MT, Axelrod JD and Moon RT (2003) A second canon. Functions and mechanisms of beta-catenin-independent Wnt signaling. *Dev Cell*, 5: 367-377.

19. James RG, Conrad WH and Moon RT (2008) Beta-catenin-independent Wnt pathways: signals, core proteins, and effectors. *Methods Mol Biol*, 468: 131-144.

20. Lee HC, Kim M and Wands JR (2006) Wnt/Frizzled signaling in hepatocellular carcinoma. *Front Biosci*, 11: 1901-1915.

21. Wong GT, Gavin BJ and McMahon AP (1994) Differential transformation of mammary epithelial cells by Wnt genes. *Mol Cell Biol*, 14: 6278-6286.

22. Shimizu H, Julius MA, Giarre M, Zheng Z, Brown AM and Kitajewski J (1997) Transformation by Wnt family proteins correlates with regulation of beta-catenin. *Cell Growth Differ*, 8: 1349-1358.

23. Katoh M and Katoh M (2005) Comparative genomics on Wnt8a and Wnt8b genes. *Int J Oncol*, 26: 1129-1133.

24. Samarzija I, Sini P, Schlange T, Macdonald G and Hynes NE (2009) Wnt3a regulates proliferation and migration of HUVEC via canonical and non-canonical Wnt signaling pathways. *Biochem Biophys Res Commun*, 386: 449-454.

25. Kestler HA and Kuhl M (2008) From individual Wnt pathways towards a Wnt signalling network. *Philos Trans R Soc Lond B Biol Sci*, 363: 1333-1347.

26. Kawano Y and Kypta R (2003) Secreted antagonists of the Wnt signalling pathway. *J Cell Sci*, 116: 2627-2634.

27. McCrea PD, Turck CW and Gumbiner B (1991) A homolog of the armadillo protein in Drosophila (plakoglobin) associated with E-cadherin. *Science*, 254: 1359-1361.

28. Perez-Moreno M and Fuchs E (2006) Catenins: keeping cells from getting their signals crossed. *Dev Cell*, 11: 601-612.

29. Kimelman D and Xu W (2006) beta-catenin destruction complex: insights and questions from a structural perspective. *Oncogene*, 25: 7482-7491.

30. Hart MJ, de los Santos R, Albert IN, Rubinfeld B and Polakis P (1998) Downregulation of beta-catenin by human Axin and its association with the APC tumor suppressor, beta-catenin and GSK3 beta. *Curr Biol*, 8: 573-581.

31. Ikeda S, Kishida S, Yamamoto H, Murai H, Koyama S and Kikuchi A (1998) Axin, a negative regulator of the Wnt signaling pathway, forms a complex with GSK-3beta and beta-catenin and promotes GSK-3beta-dependent phosphorylation of beta-catenin. *Embo J*, 17: 1371-1384.

32. Rubinfeld B, Souza B, Albert I, Muller O, Chamberlain SH, Masiarz FR, Munemitsu S and Polakis P (1993) Association of the APC gene product with beta-catenin. *Science*, 262: 1731-1734.

33. Su LK, Vogelstein B and Kinzler KW (1993) Association of the APC tumor suppressor protein with catenins. *Science*, 262: 1734-1737.

34. Price MA (2006) CKI, there's more than one: casein kinase I family members in Wnt and Hedgehog signaling. *Genes Dev*, 20: 399-410.

35. Doble BW and Woodgett JR (2003) GSK-3: tricks of the trade for a multi-tasking kinase. *J Cell Sci*, 116: 1175-1186.

36. Kishida S, Yamamoto H, Ikeda S, Kishida M, Sakamoto I, Koyama S and Kikuchi A (1998) Axin, a negative regulator of the wnt signaling pathway, directly interacts with adenomatous polyposis coli and regulates the stabilization of beta-catenin. *J Biol Chem*, 273: 10823-10826.

37. Yanagawa S, Matsuda Y, Lee JS, Matsubayashi H, Sese S, Kadowaki T and Ishimoto A (2002) Casein kinase I phosphorylates the Armadillo protein and induces its degradation in Drosophila. *Embo J*, 21: 1733-1742.

38. Yost C, Torres M, Miller JR, Huang E, Kimelman D and Moon RT (1996) The axis-inducing activity, stability, and subcellular distribution of beta-catenin is regulated in Xenopus embryos by glycogen synthase kinase 3. *Genes Dev*, 10: 1443-1454.

39. Amit S, Hatzubai A, Birman Y, Andersen JS, Ben-Shushan E, Mann M, Ben-Neriah Y and Alkalay I (2002) Axin-mediated CKI phosphorylation of beta-catenin at Ser 45: a molecular switch for the Wnt pathway. *Genes Dev*, 16: 1066-1076.

40. Liu C, Li Y, Semenov M, Han C, Baeg GH, Tan Y, Zhang Z, Lin X and He X (2002) Control of beta-catenin phosphorylation/degradation by a dual-kinase mechanism. *Cell*, 108: 837-847.

41. Aberle H, Bauer A, Stappert J, Kispert A and Kemler R (1997) beta-catenin is a target for the ubiquitin-proteasome pathway. *Embo J*, 16: 3797-3804.

42. Cavallo RA, Cox RT, Moline MM, Roose J, Polevoy GA, Clevers H, Peifer M and Bejsovec A (1998) Drosophila Tcf and Groucho interact to repress Wingless signalling activity. *Nature*, 395: 604-608.

43. Roose J, Molenaar M, Peterson J, Hurenkamp J, Brantjes H, Moerer P, van de Wetering M, Destree O and Clevers H (1998) The Xenopus Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. *Nature*, 395: 608-612.

44. Cuilliere-Dartigues P, El-Bchiri J, Krimi A, Buhard O, Fontanges P, Flejou JF, Hamelin R and Duval A (2006) TCF-4 isoforms absent in TCF-4 mutated MSI-H colorectal cancer cells colocalize with nuclear CtBP and repress TCF-4-mediated transcription. *Oncogene*, 25: 4441-4448.

45. He X, Semenov M, Tamai K and Zeng X (2004) LDL receptor-related proteins 5 and 6 in Wnt/beta-catenin signaling: arrows point the way. *Development,* 131: 1663-1677.

46. MacDonald BT, Tamai K and He X (2009) Wnt/beta-catenin signaling: components, mechanisms, and diseases. *Dev Cell,* 17: 9-26.

47. Klingensmith J, Nusse R and Perrimon N (1994) The Drosophila segment polarity gene dishevelled encodes a novel protein required for response to the wingless signal. *Genes Dev,* 8: 118-130.

48. Noordermeer J, Klingensmith J, Perrimon N and Nusse R (1994) dishevelled and armadillo act in the wingless signalling pathway in Drosophila. *Nature,* 367: 80-83.

49. Willert K, Brink M, Wodarz A, Varmus H and Nusse R (1997) Casein kinase 2 associates with and phosphorylates dishevelled. *Embo J,* 16: 3089-3096.

50. Davidson G, Wu W, Shen J, Bilic J, Fenger U, Stannek P, Glinka A and Niehrs C (2005) Casein kinase 1 gamma couples Wnt receptor activation to cytoplasmic signal transduction. *Nature,* 438: 867-872.

51. Zeng X, Tamai K, Doble B, Li S, Huang H, Habas R, Okamura H, Woodgett J and He X (2005) A dual-kinase mechanism for Wnt co-receptor phosphorylation and activation. *Nature,* 438: 873-877.

52. Itoh K, Antipova A, Ratcliffe MJ and Sokol S (2000) Interaction of dishevelled and Xenopus axin-related protein is required for wnt signal transduction. *Mol Cell Biol,* 20: 2228-2238.

53. Daniels DL and Weis WI (2005) Beta-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nat Struct Mol Biol*, 12: 364-371.

54. Ravindranath A, O'Connell A, Johnston PG and El-Tanani MK (2008) The role of LEF/TCF factors in neoplastic transformation. *Curr Mol Med*, 8: 38-50.

55. van Noort M and Clevers H (2002) TCF transcription factors, mediators of Wnt-signaling in development and cancer. *Dev Biol*, 244: 1-8.

56. Waterman ML (2004) Lymphoid enhancer factor/T cell factor expression in colorectal cancer. *Cancer Metastasis Rev*, 23: 41-52.

57. van de Wetering M, Oosterwegel M, Dooijes D and Clevers H (1991) Identification and cloning of TCF-1, a T lymphocyte-specific transcription factor containing a sequence-specific HMG box. *Embo J*, 10: 123-132.

58. Travis A, Amsterdam A, Belanger C and Grosschedl R (1991) LEF-1, a gene encoding a lymphoid-specific protein with an HMG domain, regulates T-cell receptor alpha enhancer function [corrected]. *Genes Dev*, 5: 880-894.

59. Waterman ML, Fischer WH and Jones KA (1991) A thymus-specific member of the HMG protein family regulates the human T cell receptor C alpha enhancer. *Genes Dev*, 5: 656-669.

60. Oosterwegel MA, van de Wetering ML, Holstege FC, Prosser HM, Owen MJ and Clevers HC (1991) TCF-1, a T cell-specific transcription factor of the HMG box family, interacts with sequence motifs in the TCR beta and TCR delta enhancers. *Int Immunol*, 3: 1189-1192.

61. Giese K, Amsterdam A and Grosschedl R (1991) DNA-binding properties of the HMG domain of the lymphoid-specific transcriptional regulator LEF-1. *Genes Dev*, 5: 2567-2578.

62. van de Wetering M, Oosterwegel M, Holstege F, Dooyes D, Suijkerbuijk R, Geurts van Kessel A and Clevers H (1992) The human T cell transcription factor-1 gene. Structure, localization, and promoter characterization. *J Biol Chem*, 267: 8530-8536.

63. Castrop J, van Norren K and Clevers H (1992) A gene family of HMG-box transcription factors with homology to TCF-1. *Nucleic Acids Res*, 20: 611.

64. Korinek V, Barker N, Morin PJ, van Wichen D, de Weger R, Kinzler KW, Vogelstein B and Clevers H (1997) Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science*, 275: 1784-1787.

65. Korinek V, Barker N, Willert K, Molenaar M, Roose J, Wagenaar G, Markman M, Lamers W, Destree O and Clevers H (1998) Two members of the Tcf family implicated in Wnt/beta-catenin signaling during embryogenesis in the mouse. *Mol Cell Biol*, 18: 1248-1256.

66. Brunner E, Peter O, Schweizer L and Basler K (1997) pangolin encodes a Lef-1 homologue that acts downstream of Armadillo to transduce the Wingless signal in Drosophila. *Nature*, 385: 829-833.

67. van de Wetering M, Cavallo R, Dooijes D, van Beest M, van Es J, Loureiro J, Ypma A, Hursh D, Jones T, Bejsovec A, Peifer M, Mortin M and Clevers H (1997) Armadillo coactivates transcription driven by the product of the Drosophila segment polarity gene dTCF. *Cell*, 88: 789-799.

68. Lin R, Thompson S and Priess JR (1995) pop-1 encodes an HMG box protein required for the specification of a mesoderm precursor in early C. elegans embryos. *Cell*, 83: 599-609.

69. Molenaar M, van de Wetering M, Oosterwegel M, Peterson-Maduro J, Godsave S, Korinek V, Roose J, Destree O and Clevers H (1996) XTcf-3 transcription factor mediates beta-catenin-induced axis formation in Xenopus embryos. *Cell*, 86: 391-399.

70. Molenaar M, Roose J, Peterson J, Venanzi S, Clevers H and Destree O (1998) Differential expression of the HMG box transcription factors XTcf-3 and XLef-1 during early xenopus development. *Mech Dev*, 75: 151-154.

71. Konig A, Gradl D, Kuhl M and Wedlich D (2000) The HMG-box transcription factor XTcf-4 demarcates the forebrain-midbrain boundary. *Mech Dev*, 93: 211-214.

72. Roel G, van den Broek O, Spieker N, Peterson-Maduro J and Destree O (2003) Tcf-1 expression during Xenopus development. *Gene Expr Patterns*, 3: 123-126.

73. Croce JC and McClay DR (2006) The canonical Wnt pathway in embryonic axis polarity. *Semin Cell Dev Biol*, 17: 168-174.

74. Hoppler S and Kavanagh CL (2007) Wnt signalling: variety at the core. *J Cell Sci*, 120: 385-393.

75. Standley HJ, Destree O, Kofron M, Wylie C and Heasman J (2006) Maternal XTcf1 and XTcf4 have distinct roles in regulating Wnt target genes. *Dev Biol*, 289: 318-328.

76. Behrens J, von Kries JP, Kuhl M, Bruhn L, Wedlich D, Grosschedl R and Birchmeier W (1996) Functional interaction of beta-catenin with the transcription factor LEF-1. *Nature*, 382: 638-642.

77. Houston DW, Kofron M, Resnik E, Langland R, Destree O, Wylie C and Heasman J (2002) Repression of organizer genes in dorsal and ventral Xenopus cells mediated by maternal XTcf3. *Development*, 129: 4015-4025.

78. Brannon M, Brown JD, Bates R, Kimelman D and Moon RT (1999) XCtBP is a XTcf-3 co-repressor with roles throughout Xenopus development. *Development*, 126: 3159-3170.

79. Merrill BJ, Pasolli HA, Polak L, Rendl M, Garcia-Garcia MJ, Anderson KV and Fuchs E (2004) Tcf3: a transcriptional regulator of axis induction in the early embryo. *Development*, 131: 263-274.

80. Gradl D, Konig A and Wedlich D (2002) Functional diversity of Xenopus lymphoid enhancer factor/T-cell factor transcription factors relies on combinations of activating and repressing elements. *J Biol Chem*, 277: 14159-14171.

81. Pukrop T, Gradl D, Henningfeld KA, Knochel W, Wedlich D and Kuhl M (2001) Identification of two regulatory elements within the high mobility group box transcription factor XTCF-4. *J Biol Chem*, 276: 8968-8978.

82. Kunz M, Herrmann M, Wedlich D and Gradl D (2004) Autoregulation of canonical Wnt signaling controls midbrain development. *Dev Biol*, 273: 390-401.

83. Liu F, van den Broek O, Destree O and Hoppler S (2005) Distinct roles for Xenopus Tcf/Lef genes in mediating specific responses to Wnt/beta-catenin signalling in mesoderm development. *Development*, 132: 5375-5385.

84. Roel G, Hamilton FS, Gent Y, Bain AA, Destree O and Hoppler S (2002) Lef-1 and Tcf-3 transcription factors mediate tissue-specific Wnt signaling during Xenopus development. *Curr Biol*, 12: 1941-1945.

85. Roose J and Clevers H (1999) TCF transcription factors: molecular switches in carcinogenesis. *Biochim Biophys Acta*, 1424: M23-37.

86. Oosterwegel M, van de Wetering M, Timmerman J, Kruisbeek A, Destree O, Meijlink F and Clevers H (1993) Differential expression of the HMG box factors TCF-1 and LEF-1 during murine embryogenesis. *Development*, 118: 439-448.

87. Galceran J, Farinas I, Depew MJ, Clevers H and Grosschedl R (1999) Wnt3a-/--like phenotype and limb deficiency in Lef1(-/-)Tcf1(-/-) mice. *Genes Dev*, 13: 709-717.

88. van Genderen C, Okamura RM, Farinas I, Quo RG, Parslow TG, Bruhn L and Grosschedl R (1994) Development of several organs that require inductive epithelial-mesenchymal interactions is impaired in LEF-1-deficient mice. *Genes Dev*, 8: 2691-2703.

89. Cole MF, Johnstone SE, Newman JJ, Kagey MH and Young RA (2008) Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. *Genes Dev*, 22: 746-755.

90. Pereira L, Yi F and Merrill BJ (2006) Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. *Mol Cell Biol*, 26: 7479-7491.

91. Arce L, Yokoyama NN and Waterman ML (2006) Diversity of LEF/TCF action in development and disease. *Oncogene*, 25: 7492-7504.

92. Verbeek S, Izon D, Hofhuis F, Robanus-Maandag E, te Riele H, van de Wetering M, Oosterwegel M, Wilson A, MacDonald HR and Clevers H (1995) An HMG-box-containing T-cell factor required for thymocyte differentiation. *Nature*, 374: 70-74.

93. Schilham MW, Wilson A, Moerer P, Benaissa-Trouw BJ, Cumano A and Clevers HC (1998) Critical involvement of Tcf-1 in expansion of thymocytes. *J Immunol*, 161: 3984-3991.

94. Roose J, Huls G, van Beest M, Moerer P, van der Horn K, Goldschmeding R, Logtenberg T and Clevers H (1999) Synergy between tumor suppressor APC and the beta-catenin-Tcf4 target Tcf1. *Science*, 285: 1923-1926.

95. Reya T, O'Riordan M, Okamura R, Devaney E, Willert K, Nusse R and Grosschedl R (2000) Wnt signaling regulates B lymphocyte proliferation through a LEF-1 dependent mechanism. *Immunity*, 13: 15-24.

96. Okamura RM, Sigvardsson M, Galceran J, Verbeek S, Clevers H and Grosschedl R (1998) Redundant regulation of T cell differentiation and TCRalpha gene expression by the transcription factors LEF-1 and TCF-1. *Immunity*, 8: 11-20.

97. Korinek V, Barker N, Moerer P, van Donselaar E, Huls G, Peters PJ and Clevers H (1998) Depletion of epithelial stem-cell compartments in the small intestine of mice lacking Tcf-4. *Nat Genet*, 19: 379-383.

98. Gregorieff A, Grosschedl R and Clevers H (2004) Hindgut defects and transformation of the gastro-intestinal tract in Tcf4(-/-)/Tcf1(-/-) embryos. *Embo J*, 23: 1825-1833.

99. Reya T and Clevers H (2005) Wnt signalling in stem cells and cancer. *Nature*, 434: 843-850.

100. Reya T, Duncan AW, Ailles L, Domen J, Scherer DC, Willert K, Hintz L, Nusse R and Weissman IL (2003) A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature*, 423: 409-414.

101. Merrill BJ, Gat U, DasGupta R and Fuchs E (2001) Tcf3 and Lef1 regulate lineage differentiation of multipotent stem cells in skin. *Genes Dev*, 15: 1688-1705.

102. Zhou P, Byrne C, Jacobs J and Fuchs E (1995) Lymphoid enhancer factor 1 directs hair follicle patterning and epithelial cell fate. *Genes Dev*, 9: 700-713.

103. Staal FJ and Clevers HC (2005) WNT signalling and haematopoiesis: a WNT-WNT situation. *Nat Rev Immunol*, 5: 21-30.

104. Ikenoue T, Ijichi H, Kato N, Kanai F, Masaki T, Rengifo W, Okamoto M, Matsumura M, Kawabe T, Shiratori Y and Omata M (2002) Analysis of the beta-catenin/T cell factor signaling pathway in 36 gastrointestinal and liver cancer cells. *Jpn J Cancer Res*, 93: 1213-1220.

105. Luo J, Chen J, Deng ZL, Luo X, Song WX, Sharff KA, Tang N, Haydon RC, Luu HH and He TC (2007) Wnt signaling and human diseases: what are the therapeutic implications? *Lab Invest*, 87: 97-103.

106. Lustig B and Behrens J (2003) The Wnt signaling pathway and its role in tumor development. *J Cancer Res Clin Oncol*, 129: 199-221.

107. Grant SF, Thorleifsson G, Reynisdottir I, Benediktsson R, Manolescu A, Sainz J, Helgason A, Stefansson H, Emilsson V, Helgadottir A, Styrkarsdottir U, Magnusson KP, Walters GB, Palsdottir E, Jonsdottir T, Gudmundsdottir T, Gylfason A, Saemundsdottir J, Wilensky RL, Reilly MP, Rader DJ, Bagger Y, Christiansen C, Gudnason V, Sigurdsson G, Thorsteinsdottir U, Gulcher JR, Kong A and Stefansson K (2006) Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet*, 38: 320-323.

108. Weedon MN (2007) The importance of TCF7L2. *Diabet Med*, 24: 1062-1066.

109. Florez JC, Jablonski KA, Bayley N, Pollin TI, de Bakker PI, Shuldiner AR, Knowler WC, Nathan DM and Altshuler D (2006) TCF7L2 polymorphisms and progression to diabetes in the Diabetes Prevention Program. *N Engl J Med*, 355: 241-250.

110. Bienz M and Clevers H (2000) Linking colorectal cancer to Wnt signaling. *Cell*, 103: 311-320.

111. Polakis P (2007) The many ways of Wnt in cancer. *Curr Opin Genet Dev*, 17: 45-51.

112. Cui J, Zhou X, Liu Y, Tang Z and Romeih M (2003) Wnt signaling in hepatocellular carcinoma: analysis of mutation and expression of beta-catenin, T-cell factor-4 and glycogen synthase kinase 3-beta genes. *J Gastroenterol Hepatol*, 18: 280-287.

113. Jiang Y, Zhou XD, Liu YK, Wu X and Huang XW (2002) Association of hTcf-4 gene expression and mutation with clinicopathological characteristics of hepatocellular carcinoma. *World J Gastroenterol*, 8: 804-807.

114. Zhao DH, Hong JJ, Guo SY, Yang RL, Yuan J, Wen CY, Zhou KY and Li CJ (2004) Aberrant expression and function of TCF4 in the proliferation of hepatocellular carcinoma cell line BEL-7402. *Cell Res*, 14: 74-80.

115. Duval A, Gayet J, Zhou XP, Iacopetta B, Thomas G and Hamelin R (1999) Frequent frameshift mutations of the TCF-4 gene in colorectal cancers with microsatellite instability. *Cancer Res*, 59: 4213-4215.

116. Laudet V, Stehelin D and Clevers H (1993) Ancestry and diversity of the HMG box superfamily. *Nucleic Acids Res*, 21: 2493-2501.

117. van de Wetering M and Clevers H (1992) Sequence-specific interaction of the HMG box proteins TCF-1 and SRY occurs within the minor groove of a Watson-Crick double helix. *Embo J*, 11: 3039-3044.

118. Giese K, Cox J and Grosschedl R (1992) The HMG domain of lymphoid enhancer factor 1 bends DNA and facilitates assembly of functional nucleoprotein structures. *Cell*, 69: 185-195.

119. van Beest M, Dooijes D, van De Wetering M, Kjaerulff S, Bonvin A, Nielsen O and Clevers H (2000) Sequence-specific high mobility group box factors recognize 10-12-base pair minor groove motifs. *J Biol Chem*, 275: 27266-27273.

120. Love JJ, Li X, Case DA, Giese K, Grosschedl R and Wright PE (1995) Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature*, 376: 791-795.

121. Graham TA, Weaver C, Mao F, Kimelman D and Xu W (2000) Crystal structure of a beta-catenin/Tcf complex. *Cell*, 103: 885-896.

122. Miravet S, Piedra J, Miro F, Itarte E, Garcia de Herreros A and Dunach M (2002) The transcriptional factor Tcf-4 contains different binding sites for beta-catenin and plakoglobin. *J Biol Chem*, 277: 1884-1891.

123. Maeda O, Usami N, Kondo M, Takahashi M, Goto H, Shimokata K, Kusugami K and Sekido Y (2004) Plakoglobin (gamma-catenin) has TCF/LEF family-dependent transcriptional activity in beta-catenin-deficient cell line. *Oncogene*, 23: 964-972.

124. Ghogomu SM, van Venrooy S, Ritthaler M, Wedlich D and Gradl D (2006) HIC-5 is a novel repressor of lymphoid enhancer factor/T-cell factor-driven transcription. *J Biol Chem*, 281: 1755-1764.

125. Atcha FA, Munguia JE, Li TW, Hovanes K and Waterman ML (2003) A new beta-catenin-dependent activation domain in T cell factor. *J Biol Chem*, 278: 16169-16175.

126. Hecht A and Stemmler MP (2003) Identification of a promoter-specific transcriptional activation domain at the C terminus of the Wnt effector protein T-cell factor 4. *J Biol Chem*, 278: 3776-3785.

127. Atcha FA, Syed A, Wu B, Hoverter NP, Yokoyama NN, Ting JH, Munguia JE, Mangalam HJ, Marsh JL and Waterman ML (2007) A unique DNA binding domain converts T-cell factors into strong Wnt effectors. *Mol Cell Biol*, 27: 8352-8363.

128. Valenta T, Lukas J and Korinek V (2003) HMG box transcription factor TCF-4's interaction with CtBP1 controls the expression of the Wnt target Axin2/Conductin in human embryonic kidney cells. *Nucleic Acids Res*, 31: 2369-2380.

129. Van de Wetering M, Castrop J, Korinek V and Clevers H (1996) Extensive alternative splicing and dual promoter usage generate Tcf-1 protein isoforms with differential transcription control properties. *Mol Cell Biol*, 16: 745-752.

130. Hovanes K, Li TW and Waterman ML (2000) The human LEF-1 gene contains a promoter preferentially active in lymphocytes and encodes multiple isoforms derived from alternative splicing. *Nucleic Acids Res*, 28: 1994-2003.

131. van de Wetering M, Sancho E, Verweij C, de Lau W, Oving I, Hurlstone A, van der Horn K, Batlle E, Coudreuse D, Haramis AP, Tjon-Pon-Fong M, Moerer P, van den Born M, Soete G, Pals S, Eilers M, Medema R and Clevers H (2002) The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell*, 111: 241-250.

132. Brantjes H, Barker N, van Es J and Clevers H (2002) TCF: Lady Justice casting the final verdict on the outcome of Wnt signalling. *Biol Chem*, 383: 255-261.

133. Willert K and Nusse R (1998) Beta-catenin: a key mediator of Wnt signaling. *Curr Opin Genet Dev*, 8: 95-102.

134. Bienz M (2005) beta-Catenin: a pivot between cell adhesion and Wnt signalling. *Curr Biol*, 15: R64-67.

135. Kramps T, Peter O, Brunner E, Nellen D, Froesch B, Chatterjee S, Murone M, Zullig S and Basler K (2002) Wnt/wingless signaling requires BCL9/legless-

mediated recruitment of pygopus to the nuclear beta-catenin-TCF complex. *Cell*, 109: 47-60.

136. Thompson B, Townsley F, Rosin-Arbesfeld R, Musisi H and Bienz M (2002) A new nuclear component of the Wnt signalling pathway. *Nat Cell Biol*, 4: 367-373.

137. Narlikar GJ, Fan HY and Kingston RE (2002) Cooperation between complexes that regulate chromatin structure and transcription. *Cell*, 108: 475-487.

138. Hecht A, Vleminckx K, Stemmler MP, van Roy F and Kemler R (2000) The p300/CBP acetyltransferases function as transcriptional coactivators of beta-catenin in vertebrates. *Embo J*, 19: 1839-1850.

139. Takemaru KI and Moon RT (2000) The transcriptional coactivator CBP interacts with beta-catenin to activate gene expression. *J Cell Biol*, 149: 249-254.

140. Mosimann C, Hausmann G and Basler K (2009) Beta-catenin hits chromatin: regulation of Wnt target gene activation. *Nat Rev Mol Cell Biol*, 10: 276-286.

141. Sierra J, Yoshida T, Joazeiro CA and Jones KA (2006) The APC tumor suppressor counteracts beta-catenin activation and H3K4 methylation at Wnt target genes. *Genes Dev*, 20: 586-600.

142. Parker DS, Ni YY, Chang JL, Li J and Cadigan KM (2008) Wingless signaling induces widespread chromatin remodeling of target loci. *Mol Cell Biol*, 28: 1815-1828.

143. Lewis BA and Reinberg D (2003) The mediator coactivator complex: functional and physical roles in transcriptional regulation. *J Cell Sci*, 116: 3667-3675.

144. Bauer A, Huber O and Kemler R (1998) Pontin52, an interaction partner of beta-catenin, binds to the TATA box binding protein. *Proc Natl Acad Sci U S A*, 95: 14787-14792.

145. Hecht A, Litterst CM, Huber O and Kemler R (1999) Functional characterization of multiple transactivating elements in beta-catenin, some of which interact with the TATA-binding protein in vitro. *J Biol Chem*, 274: 18017-18025.

146. Huang L, Shitashige M, Satow R, Honda K, Ono M, Yun J, Tomida A, Tsuruo T, Hirohashi S and Yamada T (2007) Functional interaction of DNA topoisomerase IIalpha with the beta-catenin and T-cell factor-4 complex. *Gastroenterology*, 133: 1569-1578.

147. Brantjes H, Roose J, van De Wetering M and Clevers H (2001) All Tcf HMG box transcription factors interact with Groucho-related co-repressors. *Nucleic Acids Res*, 29: 1410-1419.

148. Jennings BH and Ish-Horowicz D (2008) The Groucho/TLE/Grg family of transcriptional co-repressors. *Genome Biol*, 9: 205.

149. Kikuchi A, Kishida S and Yamamoto H (2006) Regulation of Wnt signaling by protein-protein interaction and post-translational modifications. *Exp Mol Med*, 38: 1-10.

150. Phillips BT and Kimble J (2009) A new look at TCF and beta-catenin through the lens of a divergent C. elegans Wnt pathway. *Dev Cell*, 17: 27-34.

151. Geiss-Friedlander R and Melchior F (2007) Concepts in sumoylation: a decade on. *Nat Rev Mol Cell Biol*, 8: 947-956.

152. Yamamoto H, Ihara M, Matsuura Y and Kikuchi A (2003) Sumoylation is involved in beta-catenin-dependent activation of Tcf-4. *Embo J*, 22: 2047-2059.

153. Vlad A, Rohrs S, Klein-Hitpass L and Muller O (2008) The first five years of the Wnt targetome. *Cell Signal*, 20: 795-802.

154. He TC, Sparks AB, Rago C, Hermeking H, Zawel L, da Costa LT, Morin PJ, Vogelstein B and Kinzler KW (1998) Identification of c-MYC as a target of the APC pathway. *Science*, 281: 1509-1512.

155. Cadoret A, Ovejero C, Terris B, Souil E, Levy L, Lamers WH, Kitajewski J, Kahn A and Perret C (2002) New targets of beta-catenin signaling in the liver are involved in the glutamine metabolism. *Oncogene*, 21: 8293-8301.

156. Renard CA, Labalette C, Armengol C, Cougot D, Wei Y, Cairo S, Pineau P, Neuveut C, de Reynies A, Dejean A, Perret C and Buendia MA (2007) Tbx3 is a downstream target of the Wnt/beta-catenin pathway and a critical mediator of beta-catenin survival functions in liver cancer. *Cancer Res*, 67: 901-910.

157. Shtutman M, Zhurinsky J, Simcha I, Albanese C, D'Amico M, Pestell R and Ben-Ze'ev A (1999) The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway. *Proc Natl Acad Sci U S A*, 96: 5522-5527.

158. Tetsu O and McCormick F (1999) Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. *Nature*, 398: 422-426.

159. Brabletz T, Jung A, Dag S, Hlubek F and Kirchner T (1999) beta-catenin regulates the expression of the matrix metalloproteinase-7 in human colorectal cancer. *Am J Pathol*, 155: 1033-1038.

160. Lee HS, Park MH, Yang SJ, Park KC, Kim NS, Kim YS, Kim DI, Yoo HS, Choi EJ and Yeom YI (2007) Novel candidate targets of Wnt/beta-catenin signaling in hepatoma cells. *Life Sci*, 80: 690-698.

161. Lustig B, Jerchow B, Sachs M, Weiler S, Pietsch T, Karsten U, van de Wetering M, Clevers H, Schlag PM, Birchmeier W and Behrens J (2002) Negative feedback loop of Wnt signaling through upregulation of conductin/axin2 in colorectal and liver tumors. *Mol Cell Biol*, 22: 1184-1193.

CHAPTER 2

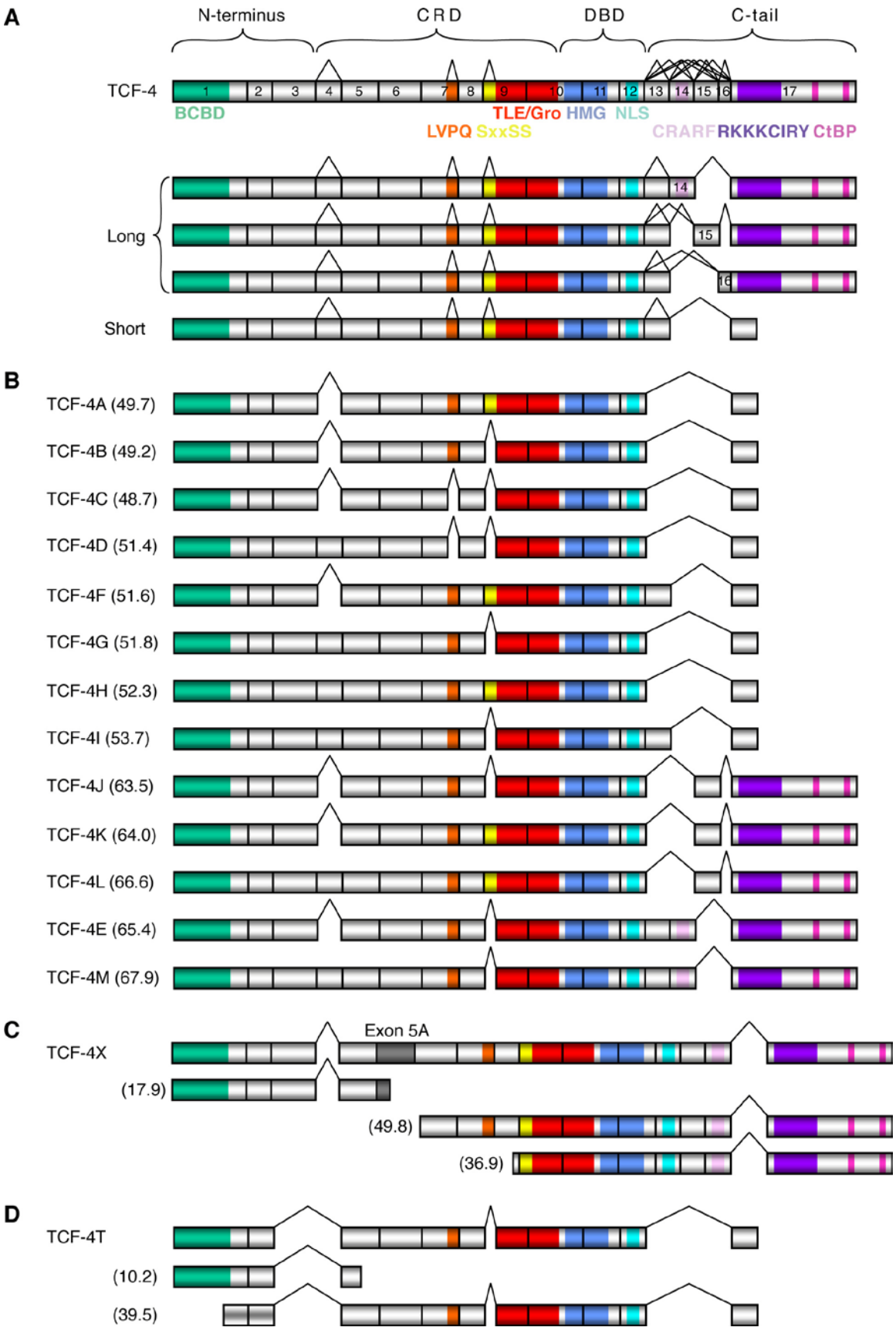
2. IDENTIFICATION OF TCF-4 ISOFORMS IN HCC

2.1. ABSTRACT

The human *TCF7L2* gene that encodes the TCF-4 transcription factor is composed of 17 exons, many of which are alternatively spliced. Because it contains 9 different splice acceptor and donor sites, it can theoretically generate more than 500 isoforms. While the well-conserved BCBD, TLE-binding domain, and DBD are constitutively expressed in all isoforms, the repression-specific motifs and the secondary DNA-binding sequences are optionally generated via alternative splicing. Therefore, I propose that *TCF7L2* expresses different isoforms in a given cell and that those isoforms may have unique functional properties depending on the presence or absence of the alternative sequences. Despite the considerable knowledge of splicing events of TCF-4 and its optionally expressed functional motifs, only two isoforms with full sequences have been described so far, and most articles regarding TCF-4 have portrayed it as if it were expressed in single form. Several reports indicated that TCF-4 was overexpressed and along with the aberrantly activated β-catenin, it upregulated the expression of Wnt target genes in HCC. However, no information is available on TCF-4 splicing variants, and functional differences between various isoforms are not known in HCC. Here, I show the analysis of TCF-4 mRNA exon by exon by RT-PCR to initially substantiate the presence of multiple isoforms in human hepatoma cells, and then describe the identification of 17 distinct TCF-4 splicing variants by sequencing cDNAs generated from complete mRNA clones derived from four HCC cell lines and human normal liver tissues.

2.2. INTRODUCTION

The human *TCF7L2* gene has been mapped to the q25.3 region of chromosome 10 [1] and its genomic structure is well characterized [2]. This 216,063 nucleotide-long gene is composed of 17 exons (1 through 17); its schematic representation and DNA sequences with exon nomenclature are shown in Figure 2.1A and Table 2.1. At the protein level, TCF-4 has several important functional domains and motifs. The N-terminus occupies the first three exons, and the BCBD resides in the first 50 amino acids of exon 1. The central CRD domain spans from exon 4 through the first half of exon 10. The TLE repressor-binding site corresponds to exons 9 and 10. The two alternatively spliced motifs, LVPQ and SxxSS, are located at the end of exon 7 (exon 7L if present or exon 7S if spliced out) and at the beginning of exon 9 (similarly exon 9L or 9S), respectively. The highly conserved DBD occupies the other half of exon 10, exon 11 and exon 12. The HMG box is covered by exons 10 and 11, while the NLS is found in the middle of exon 12. The C-terminal tail extends from exon 13 to the end of exon 17. A region in exon 14 translates into the CRARF motif, while the RKKKCIRY domain and two CtBP co-repressor-binding sites with PLSLxxK consensus sequence is expressed from a full length exon 17. The CRARF and/or RKKKCIRY sequences have been suggested to be a WRE-independent, secondary DNA-binding domain that gives specificity in target gene recognition, such as *TCF* and *LEF* genes themselves [3].


A
N-terminus
CRD
DBD
C-tail
TCF-4
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
BCBD
TLE/Gro HMG NLS
LVPQ SxxSS
CRARF RKKKCIRY CtBP
Long
Short
B
TCF-4A (49.7)
TCF-4B (49.2)
TCF-4C (48.7)
TCF-4D (51.4)
TCF-4F (51.6)
TCF-4G (51.8)
TCF-4H (52.3)
TCF-4I (53.7)
TCF-4J (63.5)
TCF-4K (64.0)
TCF-4L (66.6)
TCF-4E (65.4)
TCF-4M (67.9)
C
Exon 5A
TCF-4X
(17.9)
(49.8)
(36.9)
D
TCF-4T
(10.2)
(39.5)

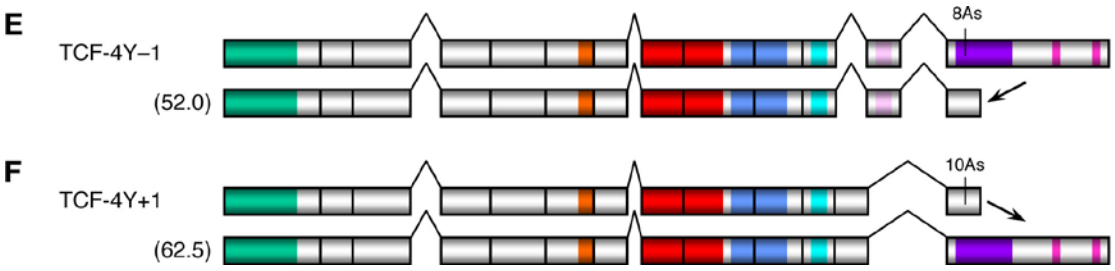


Figure 2.1. Identification of novel TCF-4 splicing variants in HCC cell lines

A. A schematic representation of human *TCF7L2* gene comprising 17 exons based on the reported cDNA sequences [2]. Organization of functional domains is same as described in Chapter 1.2. BCBD, β-catenin-binding domain; CRD, context-dependent transcription repression (or regulation) domain; TLE/Gro, TLE/Groucho co-repressor-binding site; DBD, DNA-binding domain with high mobility group (HMG) box and nuclear localization sequence (NLS); CtBP, C-terminal-binding protein co-repressor-binding site. The alternatively spliced sites, shown with a tee-pee shape, are exons 4, 7L (LVPQ), 9L (SxxSS), 13, 14, 15, 16, 16L, and 17L. *B – F*. Summary of the organizational structure of various TCF-4 isoforms identified in HCC cell lines. The calculated molecular weight of each protein product is shown in parentheses in kDa. *B*. Identification of 13 distinct TCF-4 isoforms derived from the predicted 512 possibilities. Among them were the previously known two isoforms, TCF-4B and E. To be consistent with, the remaining 11 were named alphabetically according to their increasing protein product size. *C*. Identification of TCF-4X isoform in HepG2. This unexpected isotype was termed as X because of the *unknown* and e*x*tra exon. Its mRNA included an additional exon in intron 5, which was named exon 5A to preserve the latest nomenclature order [2]. *D*. Identification of TCF-4T isoform in HepG2. It was labeled as T due to the *t*runcation and the absence of *t*hird exon. The TCF-4T mRNA had the non-alternative exon 3 spliced out from the sequence. *E*. Identification of TCF-4Y–1 isoform in Hep3B. Its mRNA had a deletion of one adenine in the poly A tract (8As), which resulted in a frameshift mutation that lead to the truncation of the C-terminal tail. *F*. Identification of TCF-4Y+1 isoform in Hep3B. Its transcript contained an additional adenine in the poly A tract (10As). The latter two isotypes were called TCF-4Y as in the question *why* (it happened).

TCF-4 mRNA is the only TCF/LEF member that can theoretically create the most splice variants due to the presence of multiple alternative exons. The exons 4, 13, 14, 15, and 16 are alternatively spliced ($2^5 = 32$), and exons 7, 9, 16, and 17 have alternative splice acceptor and donor sites ($2^4 = 16$) making this gene possible to create more than 500 ($32 \times 16 = 2^9 = 512$) isoform transcripts during pre-mRNA processing. However, because the *TCF7L2* gene has one promoter, DNA transcription of this gene produces single type of pre-mRNA. The translation begins from the first start codon in exon 1 (Table 2.1). Because the number of nucleotides that are added or removed from the coding region due to alternative splice acceptor or donor sites of exons 7L and 9L, and the base pairs of exons 4 and 13 are divisible by three (as in three nucleotides of a codon), alternative splicing of these sequences does not interrupt the reading frame until the end of exon 13. Therefore, all types of TCF-4 mRNA isoforms translate the same BCBD, the repressor TLE-binding domain, and the DBD with both HMG box and NLS (Figure 2.1). However, the presence of exons 14, 15 and 16 alters the reading frame of exon 17 and determines the length of the C-terminal tail. Acceptance of one of the exons 14, 15 or 16, but not in combination, results in the so called long isoform such that mRNA translation terminates at the last stop codon located towards the end of exon 17 owing to the utilization of reading frame 2 of exon 17 (Figure 2.1A). However, exon 16 is shown to be expressed very rarely and identified only in a single type of CRC cell line so far [2]. Long isoforms contain the CRARF motif, if exon 14 is used, the RKKKCIRY domain, and the two CtBP co-repressor-binding sites. For example, TCF-4E, one of the first two isoforms identified [4], is a long isotype harboring all of the above functional domains (Figure 1.2B).

Table 2.1. Exon sequences of *TCF7L2* gene

Complete sequences of each exon and partial intronic sequences are shown with their sizes and corresponding nucleotide positions at chromosome band 10q25.3. The coding sequences in each exon are capitalized and highlighted with dark grey color. The capitalized sequences with light grey shade in intron 5 correspond to the newly identified exon 5A. Note that exon 5A also has the ag/gt intron boundaries, which are underlined in all introns. The sequences for alternative acceptor (9L, 16L, 17L) and donor sites (7L) are shown in italic letters. The first start codon used for translation initiation in exon 1, as well as the stop codons that result in short and long C-tails in exon 17 are shown in bold letters. The poly A tract in exon 17 is presented in white font.

Exon/ Intron	Chromosome position	Length (bp)	Sequence
Exon 1	114710009–114710704	696	aataatctccgctcccagactactccgttcctccggatttcgatccccctttt tctatctgtcaatcagcgccgcctttgaactgaaaagctctcagtctaacttc aactcactcaaatccgagcggcacgagcacctcctgtatcttcggcttccccc ccccttttgctctttatatctgacttcttgttgttgttggtgtttttttttttt ttaccccccttttttatttattatttttttgcacattgatcggatccttggga acgagagaaaaaagaaacccaaactcacgcgtgcagaagatctcccccccctt ccccctcccctcctccctctttcccctccccaggagaaaaagaccccccaagca gaaaaaagttcaccttggactcgtctttttcttgcaatatttttttgggggggc aaaactttttgggggtgatttttttttggctttttcttcctccttcatttttctt ccaaaattgctgctggtgggtgaaaaaaaa**ATG**CCGCAGCTGAACGGCGGTGG AGGGGATGACCTAGGCGCCAACGACGAACTGATTTCCTTCAAAGACGAGGGCG AACAGGAGGAGAAGAGCTCCGAAAACTCCTCGGCAGAGAGGGATTTAGCTGAT GTCAAATCGTCTCTAGTCAATGAATCAGAAACGAATCAAAACAGCTCCTCCGA TTCCGAG
Intron 1	114710705–114710965	261	gtaggaaaagcccctcgggc.....ctccacccctctgggaag
Exon 2	114710966–114711032	67	GCGGAAAGACGGCCTCCGCCTCGCTCCGAAAGTTTCCGAGACAAATCCCGGGA AAGTTTGGAAGAAG
Intron 2	114711033–114711241	209	gtgagtacgccccgcgcg.....gccgccccgccatgttag
Exon 3	114711242–114711366	125	CGGCCAAGAGGCAAGATGGAGGGCTCTTTAAGGGGCCACCGTATCCCGGCTAC CCCTTCATCATGATCCCCGACCTGACGAGCCCCTACCTCCCCAACGGATCGCT CTCGCCCACCGCCCGAACC
Intron 3	114711367–114724314	12948	gtaagtgcctccgcgccc.....tttggtttctttctacag
Exon 4	114724315–114724384	69	CTCCATTTTCAGTCCGGCAGCACACATTACTCTGCGTACAAAACGATTGAACA CCAGATTGCAGTTCAG

Intron 4	114724385-114799783	75399	gtaggaaacgcaagagat.....tttgtttcctctcgccag
Exon 5	114799784-114799885	102	TATCTCCAGATGAAATGGCCACTGCTTGATGTCCAGGCAGGGAGCCTCCAGAGTAGACAAGCCCTCAAGGATGCCCGGTCCCCATCACCGGCACACATTGTC
Intron 5	114799886-114900942	101057	gtaagtaacctccca.....tattatttccaacagGCAGTGCTGTAGCTATCGATAGAAGTCTTGTCCAGGGAAACAATCCACCATTAGCTGGGAGCAAGGCCTTAGACAGAAAGGAGCAACATTCAACTAAGGAAGACCCCGGCCACGgtgatgcatgggtcc.....ttctcctccccacag
Exon 6	114900943-114901075	133	TCTAACAAAGTGCCAGTGGTGCAGCACCCTCACCATGTCCACCCCCTCACGCCTCTTATCACGTACAGCAATGAACACTTCACGCCGGGAAACCCACCTCCACACTTACCAGCCGACGTAGACCCCAAAACAG
Intron 6	114901076-114903681	2606	gtaggctgtgggctacgg.....tgctgcttcttccctcag
Exon 7	114903682-114903784	103	GAATCCCACGGCCTCCGCACCCTCCAGATATATCCCCGTATTACCCACTATCGCCTGGCACCGTAGGACAAATCCCCCATCCGCTAGGATG*GTTAGTACCACA*
Intron 7	114903785-114905769	1985	gtaaggagttccattttt.....tgtacccctttctttgtag
Exon 8	114905770-114905856	87	GCAAGGTCAACCAGTGTACCCAATCACGACAGGAGGATTCAGACACCCCTACCCCACAGCTCTGACCGTCAATGCTTCCATGTCCAG
Intron 8	114905857-114910741	4885	gtgagttccaagaaccgg.....ttctgacgatttacacag
Exon 9	114910742-114910882	141	*CTTTCTGTCTTCTAG*GTTCCCTCCCCATATGGTCCCACCACATCATACGCTACACACGACGGGCATTCCGCATCCGGCCATAGTCACACCAACAGTCAAACAGGAATCGTCCCAGAGTGATGTCGGCTCACTCCATAGTTC
Intron 9	114910883-114911483	601	gtaagtgttgctgttttt.....tgttttgtctatctccag
Exon 10	114911484-114911643	160	AAAGCATCAGGACTCCAAAAAGGAAGAAGAAAAGAAGAAGCCCCACATAAAGAAACCTCTTAATGCATTCATGTTGTATATGAAGGAAATGAGAGCAAAGGTCGTAGCTGAGTGCACGTTGAAAGAAAGCGCGGCCATCAACCAGATCCTTGGGCGGAGG
Intron 10	114911644-114912091	448	gtaggtgacgcccttctc.....accccacgtcccctccag
Exon 11	114912092-114912199	108	TGGCATGCACTGTCCAGAGAAGAGCAAGCGAAATACTACGAGCTGGCCCGGAAGGAGCGACAGCTTCATATGCAACTGTACCCCGGCTGGTCCGCGCGGGATAACTAT
Intron 11	114912200-114917779	5580	gtaggtggatcattttcg.....tttcatctcttcatctag
Exon 12	114917780-114917828	49	GGAAAGAAGAAGAAGAGGAAAAGGGACAAGCAGCCGGGAGAGACCAATG
Intron 12	114917829-114918425	597	gtaagtgacaatcatcag.....tttgttctttttttcag
Exon 13	114918424-114918476	51	AACACAGCGAATGTTTCCTAAATCCTTGCCTTTCACTTCCTCCGATTACAG
Intron 13	114918477-114919678	1202	gtgctaatgtcattttga.....tctcttgaactcattcag
Exon 14	114919679-114919751	73	ACCTGAGCGCTCCTAAGAAATGCCGAGCGCGCTTTGGCCTTGATCAACAGAATAACTGGTGCGGCCCTTGCAG
Intron 14	114919752-114920377	626	gtgtgtatagttttccag.....tttaccttatgctaacag

Exon 15	114920378-114920450	73	ATGCAAATACTCCAAAGAAGTGTCGGGCACTGTTCGGGCTTGACCGACAGACTTTATGGTGCAAACCGTGCAG
Intron 15	114920451-114921334	884	gtatattaccactgcgag.....tctttgttctgatacaag
Exon 16	114921335-114921362	28	CAGTCTTTGAATTTGGAATATTACAATG
Intron 16	114921363-114925290	3928	gtaggtatttcaacaccc.....tgtgccctctattcacag
Exon 17	114925291-114926073	783	ATAACTCTCTCCCCTGTTTCTAGGAGAAAAAAAAAGTGCGTTCGCTACATACA AGGTGAAGGCAGCTGCCTCAGCCCACCCTCTTCAGATGGAAGCTTACTAGATT CGCCTCCCCCCTCCCCGAACCTGCTAGGCTCCCCTCCCCGAGACGCCAAGTCA CAGACTGAGCAGACCCAGCCTCTGTCGCTGTCCCTGAAGCCCGACCCCCTGGC CCACCTGTCCATGATGCCTCCGCCACCCGCCCTCCTGCTCGCTGAGGCCACCC ACAAGGCCTCCGCCCTCTGTCCCAACGGGGCCCTGGACCTGCCCCCAGCCGCT TTGCAGCCTGCCGCCCCCTCCTCATCAATTGCACAGCCGTCGACTTCTTCCTT ACATTCCCACAGCTCCCTGGCCGGGACCCAGCCCCAGCCGCTGTCGCTCGTCA CCAAGTCTTTAGAATAGctttagcgtcgtgaaccccgctgctttgtttatggt tttgtttcacttttcttaatttgccccccaccccacccttgaaaggttttgtt ttgtactctcttaattttgtgccatgtggctacattagttgatgtttatcgag ttcattggtcaatatttgacccattcttatttcaatttctccttttaaatatg tagatgagagaagaacctcatgattctaccaaaatttttatcaacagctgttt aaagtctttgtagcgtttaaaaaatatatatatatacataactgttatgtagt tcggatagcttagttttaaaagactgattaaaaaacaaaaa

On the other hand, when exons 14, 15 and 16 are all spliced out of the transcript, the resultant protein product is called a short isoform because mRNA translation ends at the beginning of exon 17 just after the poly A tract due to the third reading frame of exon 17 (Figure 2.1A). The poly A tract is a sequence of nine adenines (Table 2.1) categorized as microsatellite repeats that are prone to deletion or insertion mutations, which frequently accumulate in some tumors [5]. A deletion of one nucleotide in poly A tract of *TCF7L2* has been reported as a frequent truncating (frameshift) mutation in a subset of CRC cells characterized by microsatellite instability [6]. In short isoforms, none of the functional domains of the C-terminal tail is expressed (Figure 2.1A). TCF-4B, the other isoform identified so far [4], is a very well known short isotype ubiquitously expressed in a wide range of normal and tumor cells [2,4,7-9]. Because the length of TCF-4 protein is dictated by the splicing of exons 14, 15, and 16, researchers have been focused on the C-terminus for identifying splicing variations; although they have reported various TCF-4

sequences of the C-terminus, the alternative exon combinations at the N-terminus and CRD have never been clarified [2,8,9]. Therefore, TCF-4B and E isoforms are the only splicing variants, in which the complete sequences are fully described.

Finally, when any combination of exons 14, 15, and 16 is expressed, the mRNA translation ends prematurely without reaching the exon 17, which results in truncation. Therefore, such isoform would not express any of the C-termial domains. It is important to note that a TCF-4 splicing variant containing all 17 exons is extremely rare, if at all exists, and such combination must not be thought as default or wild type isoform. Besides, such splicing variant would not express the functional domains of C-terminus.

The main characteristic of canonical Wnt signaling is stabilization of β-catenin in the cytoplasm, its subsequent translocation into the nucleus, and activation of Wnt target genes by binding to the TCF-4 transcription factors. However, although β-catenin has been described as the main effector of Wnt signaling pathway, its success on relaying the signal in the nucleus depends on the presence of TCF-4 proteins at the promoter regions of Wnt target genes. Therefore, TCF-4 is equally important in mediating the final transcriptional outcome and such activity may be regulated according to their isoform diversity and relative abundance. Unfortunately, TCF-4 has been portrayed as a docile, yet two-faced protein bound to the promoters of Wnt target genes, where its effector role (transcriptional activation *vs*. repression) is decided by the relative abundance of co-activators *vs*. co-repressors [10]. In this context, if the amount of co-activators, such as β-catenin, exceeds the amount of co-repressors, like TLE and CtBPs, TCF-4 is converted into a transcriptional activator; but when the TLE and CtBPs overwhelm the co-activators, then TCF-4 befriends with the co-repressors and inhibits the transcription of

target genes. If TCF-4, regardless of isotype difference, were such a passive protein that behaved the same way by siding with whichever co-regulatory protein was expressed the most, why would a cell waste its energy and time for generating different types of isoforms? The potential of producing more than 500 splicing variants suggests that different isoforms may have different roles in determining the final transcriptional program. Therefore, I hypothesize that TCF-4 expresses different isoforms and that those isoforms may have unique functional responsibilities: 1) activator isotypes that are specific for transcription activation and favor interaction with β-catenin; 2) target gene repression-specific isoforms, or repressor types, that make strong bonds with TLE, CtBPs and other repressors; and 3) unbiased isoforms that balance the activation and repression of Wnt responsive genes. These isoforms may also be recruited to occupy the WRE sequences temporarily until the designated type is expressed to exert its unique function.

Activation of Wnt/β-catenin pathway and overexpression of Wnt target genes have been implicated in hepatocarcinogenesis [11-13]. In addition, among the TCF/LEF family of proteins, TCF-4 has been reported as the main member expressed in HCC, and overexpression of TCF-4 transcripts has been detected in human HCC cases [14-17]. Few reports are available on the expression of LEF-1 in HCC as another member [18]. It has been reported that TCF-4/β-catenin complex formation is increased in transgenic mice that develop HCC [19]. Therefore, it is important to identify expression of various TCF-4 isoforms in the context of HCC.

In this chapter, I describe the identification of various TCF-4 isoforms from human hepatoblastoma cell line HepG2, and HCC cell lines Hep3B, Huh7, and FOCUS. These four hepatoma cells have been well characterized and their differentiation status

has been categorized as follows: undifferentiated (most malignant) < FOCUS < Huh7 < Hep3B < HepG2 < most differentiated (least malignant) [20]. The FOCUS (Friendship of China and United States) carcinoma cell line, derived from a Greek patient with HBV-associated HCC, was established in Dr. Wands' laboratory in 1984 [21]. It contained HBV sequences integrated in the short arm of chromosome 11, which resulted in a deletion of more than 13.5 kilobases (kb) of cellular sequences and hence LOH. In 1982, the human hepatoma Huh7 cell line was derived from a 57 year-old Japanese male with differentiated HCC [22]. Although the patient was negative for HCV, Huh7 cells were highly permissive for HCV infection and provided favorable environment for HCV RNA replication. The Hep3B and HepG2 cell lines were established simultaneously in 1979 at Wistar Institute of Anatomy and Biology, Philadelphia, PA [23], and they are commercially available at American Type Culture Collection. Hep3B was isolated from a liver biopsy of an 8 year-old African American boy, who later died from HBV-related, well-differentiated, and metastasized HCC. In contrast, HepG2 cell line was derived from another liver biopsy of a 15 year-old white Argentinean male with well-differentiated hepatoblastoma, an embryonal malignancy of hepatocellular origin and the most common primary liver cancer of childhood [24]. These cell lines had frequent chromosomal abnormalities with average chromosome numbers of 60 in Hep3B and 55 in HepG2. Of note, Hep3B cells had integrated HBV DNA sequences that partially deleted the *p53* gene locus leading to loss of activity of this tumor suppressor. On the other hand, the HepG2 cell line carried a large deletion in the exon 3 of the *CTNNB1* gene that contained the phosphorylation sites leading to stabilization and accumulation of β-catenin, and eventually aberrant activation of canonical Wnt signaling.

2.3. RESULTS

2.3.1. Evaluation of TCF-4 mRNA splicing variants exon by exon

Before analyzing each exon of *TCF7L2* gene, I confirmed that its protein product, TCF-4, was expressed in hepatoma cell lines HepG2, Hep3B, Huh7, and FOCUS by Western Blot (WB) analysis (Figure 2.2, middle panel). Human embryonic kidney (HEK) 293 cells were used for control. WB results showed multiple bands suggesting the presence of different TCF-4 isoforms in all HCC cell lines. HEK did not reveal any expression of TCF-4 proteins. Expression pattern of short and long isoforms transfected in HEK293 cells (HEK + Short and HEK + Long, respectively) as positive controls indicated that HCC cell lines expressed both short and long isoforms, though at lower levels in Hep3B and FOCUS cells. HEK293 cells transfected with empty vector (HEK + EV) served as negative control for transfections.

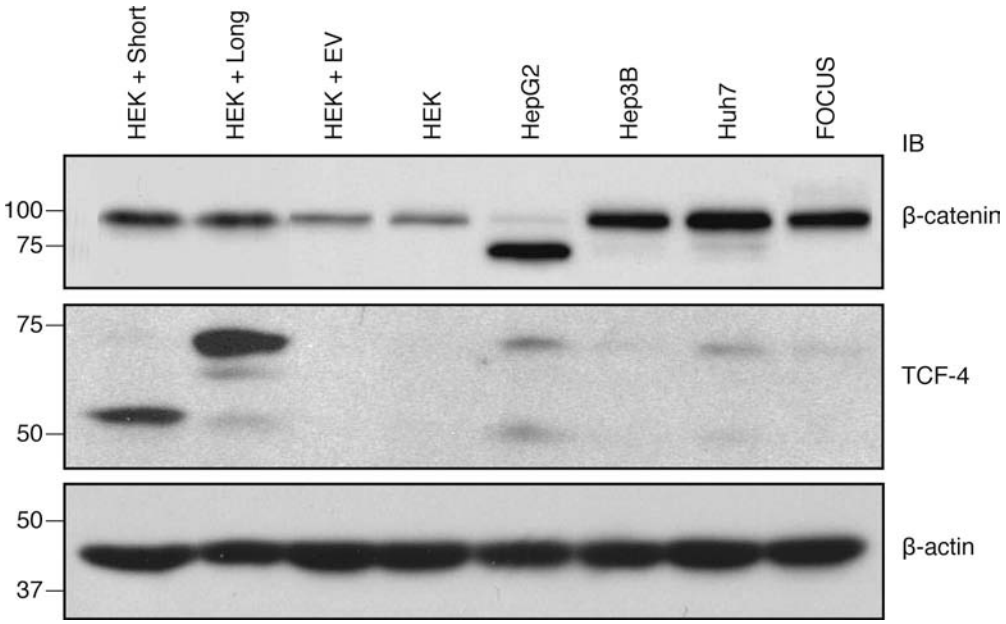


Figure 2.2. Endogenous expression of TCF-4 proteins in HCC cell lines

Expression of β-catenin and TCF-4 isoforms was detected by WB analysis from whole cell lysates. The average protein size of TCF-4 is ~51 kDa for short isoforms and ~66 kDa for long

isoforms as depicted by the expression patterns of positive controls. All four HCC cell lines expressed both short and long isotypes (middle panel). The molecular weight for β-catenin is 92 kDa. Because β-catenin is also a structural protein, it was expressed abundantly in all cell lines. Please note that HepG2 expressed a 75 kDa truncated form of β-catenin due to partial deletion of exon 3 of the *CTNNB1* gene (top panel). Expression of β-actin served as the protein loading control (bottom panel).

DNA sequences and exon characteristics of *TCF7L2* are shown in Figure 2.1A and Table 2.1. In order to evaluate TCF-4 mRNA splicing isoforms in HCC, I performed reverse transcriptase polymerase chain reaction (RT-PCR) in HepG2, Hep3B, Huh7, and FOCUS cell lines by using different primer sets designed against each exon (Figure 2.3A and Table 2.2) and analyzed on 2% agarose gels (Figure 2.3B). All PCR products migrated within the expected size range and the different band patterns on the agarose gel verified the presence of alternative exons. For example, PCR products of 1F/2R, 1F/3R, and 1F/4R had single bands, whereas starting from 1F/5R and on, there were more than one bands indicating that exon 4 was indeed alternatively spliced. In addition, the PCR product of 1F/4R gave some added information. Except for Hep3B, no visible product was detected under UV suggesting that TCF-4 isotypes with exon 4 were expressed at low abundance. In fact, the homolog of exon 4 is absent in other members of TCF/LEF family (Figure 1.2.2), and the known isoforms, TCF-4B and TCF-4E, do not contain this exon. Its existence had long been unrecognized and alternative splicing had been considered only at the C-terminal tail until exon 4 was identified and described in CRC [2]. However, the function of exon 4 is not known.

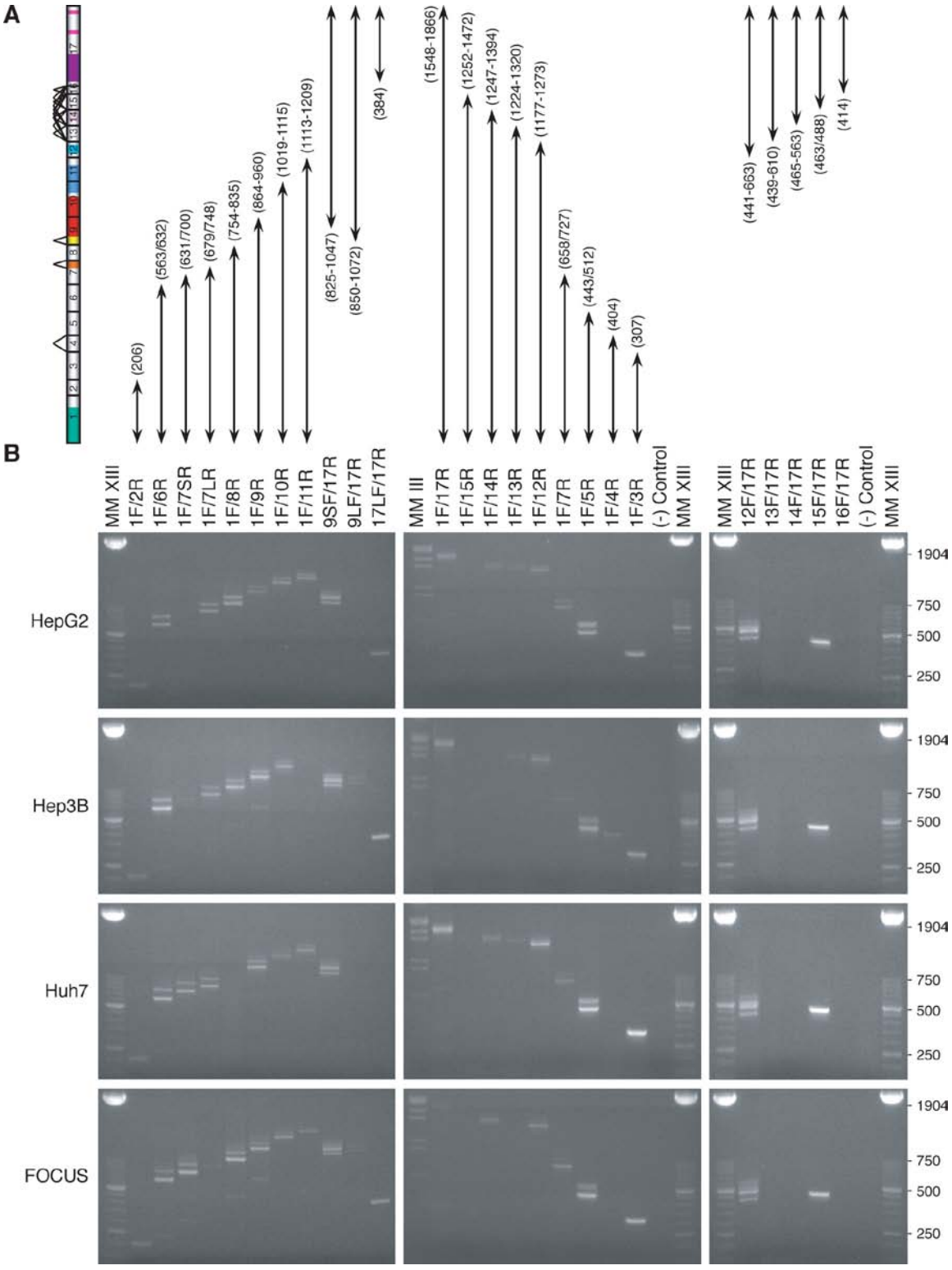

A
(206)
(563/632)
(631/700)
(679/748)
(754-835)
(864-960)
(1019-1115)
(1113-1209)
(825-1047)
(850-1072)
(384)
(1548-1866)
(1252-1472)
(1247-1394)
(1224-1320)
(1177-1273)
(658/727)
(443/512)
(404)
(307)
(441-663)
(439-610)
(465-563)
(463/488)
(414)
B
MM XIII
1F/2R
1F/6R
1F/7SR
1F/7LR
1F/8R
1F/9R
1F/10R
1F/11R
9SF/17R
9LF/17R
17LF/17R
MM III
1F/17R
1F/15R
1F/14R
1F/13R
1F/12R
1F/7R
1F/5R
1F/4R
1F/3R
(-) Control
MM XIII
MM XIII
12F/17R
13F/17R
14F/17R
15F/17R
16F/17R
(-) Control
MM XIII
HepG2
Hep3B
Huh7
FOCUS
1904
750
500
250

Figure 2.3. Evaluation of TCF-4 mRNA exons by RT-PCR

Exon-specific primer sets were used to assess various splicing possibilities. *A*. A schematic representation of the primer pairs used. The double-headed arrows represent PCR amplification between a given forward and reverse primers corresponding to each exon of the TCF-4 mRNA schematic figure. The numbers in the parenthesis indicate the size of amplicons in base pairs. If a primer set is expected to generate single product, the estimated size is given. When two different products are amplified due to presence of one alternative splicing site, both sizes are shown separated by a slash sign (e.g. 443/512 for primer set 1F/5R due to alternative exon 4). However, if more than two amplicons are anticipated due to multiple alternative splicing, then the range of sizes is represented with a hyphen between the lowest and highest estimated products (e.g. 1548 – 1866 for 1F/17R due to all alternatively spliced exons). This was the method I exploited for cloning the various TCF-4 isoforms as the primer set 1F/17R amplified all available TCF-4 isoform transcripts in each cell line. *B*. Agarose gel electrophoresis of RT-PCR products stained with ethidium bromide. Each RT-PCR product from all cell lines was run in 2% agarose gel and visualized under the UV light. The (-) control samples were obtained by RT-PCR without the template cDNA. The DNA molecular weight marker XIII (MM XIII) is a 50-base pair ladder, while the MM III (bacteriophage lambda DNA digested with *Eco* RI and *Hind* III) shows higher molecular sizes of 564, 831, 947, 1375, 1584, 1904, 2027, 3530, and etc. Most of the PCR products migrated in the agarose gels as expected. No PCR product was obtained with primer set 16F/17R, and no isoform was identified with exon 16. Despite the presence of PCR amplicon with primer pair 17LF/17R, an isoform with exon 17L was not isolated.

Another expected outcome was the PCR product of primer pair 12F/17R, which showed multiple bands confirming that the exons between 12 and 17 are alternatively spliced extensively. Finally, the primer sets specifically designed to distinguish products with exon 7S (without LVPQ) or 7L (with LVPQ) and exon 9S (without SxxSS) or 9L (with SxxSS) also confirmed the alternatively spliced acceptor sites in these exons by

showing multiple bands. As with exon 4, TCF-4 isoforms with exon 9L seemed to be expressed less abundantly than ones with 9S in HCC. Since LVPQ and SxxSS had been described to have transcription repression characteristics [25], lack of these motifs in most HCC cells was unsurprising as they might have disadvantageous effect for the propagation of cancerous cells.

Table 2.2. Primer sequences used for *TCF7L2* exon analysis

Primer	Sequence (5′ – 3′)
1F	CCGCTCGAGCGGATGCCGCAGCTGAACGGCGG
9SF	CATATGGTCCCACCACATCA
9LF	CTTTCTGTCTTCTAGGTTCCCTCC
12F	GGAAAGAAGAAGAAGAGGAAAAGG
13F	CAGCGAATGTTTCCTAAATCC
14F	ACCTCAGCGCTCCTAAGAAA
15F	GCAAATACTCCAAAGAAGTGTCG
16F	GAATTTGGAATATTACAATG
17LF	TCTCTCCCCTGTTTCTAGGAGA
2R	ACTTTCCCGGGATTTGTCTC
3R	AGGTAGGGGCTCGTCAGGT
4R	TGCAATCTGGTGTTCAATCG
5R	GACAATGTGTGCCGGTGAT
6R	TACGTCGGCTGGTAAGTGTG
7R	GGATGGGGGATTTGTCCTAC
7SR	GGCGATAGTGGGTAATACGG
7LR	TGTGGTACTAACCATCCTAGCG
8R	CTGGACACGGAAGCATTGAC
9R	AACTATGGAGTGAGCCGACA
10R	CCCAAGGATCTGGTTGATGG
11R	AGCCGGGGTACAGTTGCATA
12R	GTCTCTCCCGGCTGCTTGT
13R	GAGGAAGTGAAAGGCAAGGA
14R	GCCGCACCAGTTATTCTGTT
15R	CACGGTTTGCACCATAAAGTC
17R	CGCGGATCCGCGCTATTCTAAAGACTTGGTGACGAGCGACAGCGGCT

2.3.2. Identification of novel TCF-4 isoforms

In order to isolate different TCF-4 mRNA isotypes expressed in HCC, I performed RT-PCR using primer pair 1F/17R that spanned from the beginning of exon 1 through the end of exon 17 in HepG2, Hep3B, Huh7, and FOCUS cell lines. The PCR products, which included various combinations of the alternative exons, were cloned into pCR2.1-TOPO vector using a TA cloning kit. A total of 48 clones (12 from each cell line) were harvested, and each of the individual clones was sequenced. DNA sequencing results revealed the isolation of 17 different TCF-4 isoforms as depicted in Figure 2.1. Interestingly, not all splicing variants had the exon characteristics described previously as unexpected isotypes were uncovered. In this context, 4 out of 17 isoforms showed novel properties in *TCF7L2* exons (Figure 2.1C-F), while 13 were derived from the expected 512 possibilities (Figure 2.1B). The two previously known species, TCF-4B and TCF-4E [4], were identified among them. Except for these two, the rest of the new isoforms were named in an alphabetical order according to calculated size of their protein products (TCF-4A through M). It was worth noting that out of 48 clones sequenced, 18 of them (37.5%) were identified as TCF-4B, suggesting that this prominent isotype was expressed abundantly in HCC or that it was stabilized at mRNA level. Among the 13 splicing variants, eight of them were short isoforms while five were long ones harboring the CtBP-binding sites (Figure 2.1B). Moreover, six isotypes contained the exon 4, whereas another five included both exons 7L and 9L encoding the LVPQ and SxxSS motifs, respectively. Presence of SxxSS motif in human tumors had not been described previously (compare the well-known TCF-4B with TCF-4A). Finally, there were two isoforms incorporating the alternative exon 13, two others bearing the exon 14, and three

carrying the exon 15. However, I could not find any isoform possessing the exon 16 although it was described in a CRC cell line [2]. HCC cells might preferentially splice this short exon out of TCF-4 mRNAs since the RT-PCR with primer pair 16F/17R did not generate any PCR product in the agarose gel as well (Figure 2.3B).

One of the unexpected isotypes was named as TCF-4X (Figure 2.1C). The TCF-4X mRNA included an additional exon of 114 nucleotides embedded within intron 5, the longest intron of the *TCF7L2* gene with over 100,000 base pairs (Table 2.1). This novel exon was referred to exon 5A to keep the previously described nomenclature order [2]. DNA sequence of the *TCF7L2* gene confirmed it to be a potential exon as it contained the ag/gt dinucleotides at intron boundaries. Interestingly, the insertion of exon 5A in TCF-4X mRNA terminated the translation prematurely at the beginning of this extra exon generating a protein product of 17.9 kDa, including only the BCBD. However, two more potential translation initiation codons existed at the beginning of exon 6 and at the end of exon 8, producing protein products of 49.8 kDa and 36.9 kDa, respectively, but each without the BCBD (Figure 2.1C). If any of these three forms were expressed, they might repress transcription of Wnt target genes even in the presence of β-catenin during active Wnt signaling. In this context, the latter two might occupy the WREs but could not activate transcription since they would not be able to interact with β-catenin, while the first one could compete with other TCF-4 isoforms for binding to β-catenin and tether it away from the WREs.

Another unusual isoform identified was TCF-4T (Figure 2.1D). The TCF-4T mRNA lacked the exon 3 though it was described to be a constitutively expressed sequence of *TCF7L2*. This isoform was missing all 125 nucleotides of the exon 3 and did

not affect the sequence of exon 2 indicating that the exon 3 was yet another alternatively spliced exon. Similar to TCF-4X, however, TCF-4T was able to generate two different translation products due to the early stop codon at exon 5 and the alternative start codon at exon 1 (Figure 2.1D). In the first one, the amino acid translation was terminated prematurely generating a small, truncated protein product of 10.2 kDa, just carrying the BCBD. The second one produced a translation product of 39.5 kDa without the BCBD. However, because the start codon of the latter version was set in another reading frame, the amino acid sequences at the end of exon 1 and the entirety of exon 2 were different from the TCF-4 protein sequences. If these forms were expressed, they might as well repress transcription of Wnt target genes even in the presence of β-catenin in a similar fashion as with TCF-4X.

Finally, two distinct isoforms with potential genetic alterations were identified (Figure 2.1E and F). Termed as TCF-4Y, they had frameshift mutations at the poly A tract at the beginning of exon 17 (Table 2.1). While the TCF-4Y–1 isoform had a deletion of one adenine (8As), the TCF-4Y+1 isotype contained an additional adenine (10As). Interestingly, these frameshift mutations changed the length of C-terminal tails from long isoform to short isotype, and vice versa. In TCF-4Y–1, the deletion of a single adenine resulted in premature termination of protein translation at the beginning of exon 17 switching the expected long tail to a short tail (Figure 2.1F). On the other hand, an insertion of an extra adenine eliminated the stop codon of TCF-4Y+1 leading this isoform to have a long tail instead of the expected short one (Figure 2.1F). A deletion of a single adenine in the poly A tract of TCF-4 had been shown to be a genetic alteration in CRC [6,26]. Identification of TCF-4Y–1 isoform confirmed that this frameshift mutation also

occurred in HCC, and that hepatoma cells exploited this mutation to escape from having the CtBP repressor-binding sites in TCF-4 transcription factors. However, an insertion of single adenine in the poly A tract had never been observed. In the case for TCF-4Y+1, the addition of extra adenine resulted in the expression of the CtBP-binding sites. Although expression of repressor sites in TCF-4 was not favored for cancer cells, the insertion mutation might have other importance in different isoforms. It was interesting to note that the extra nucleotide altered the structure of TCF-4Y+1 to appear similar to that of TCF-4E, another well-known isoform. Nevertheless, the isoforms TCF-4Y–1, TCF-4Y+1 and TCF-4T were not investigated further, and they were excluded from the future experiments.

Using the same cloning method, TCF-4 isoforms expressed in normal human liver samples were identified. However, all of the positive clones (50 clones) sequenced represented the single isotype, TCF-4B, suggesting that it was the most stabilized isoform at mRNA level and/or dominantly expressed splice variant in normal liver tissues, and that other alternatively spliced TCF-4 isoforms were expressed at very low level. In addition, TCF-4B could be the canonical isoform.

2.4. DISCUSSION

TCF/LEF transcription factors are the ultimate mediators of the Wnt signaling pathway. Among the four members, TCF-4 is unique in that it can theoretically generate more than 500 isoforms, and can preferentially splice some of the functional domains in or out of the sequence. This concept suggests that not only multiple isoforms exist but also a certain isotype may have a specific function and may be required in a particular situation. For example, an isoform that contains many of the repressive motifs, such as LVPQ, SxxSS, and PLSLxxK, may be expressed specifically in the absence of Wnt signals and required for transcriptional repression of target genes. However, despite the available knowledge of splicing variations and the optional functional motifs, TCF-4 has always been regarded as a transcription factor expressed as one type, whose function is determined by the amount of co-regulators. In HCC, TCF-4 has been reported to be overexpressed along with β-catenin upregulation. However, no information is available on TCF-4 isoform specificity. It is not known if there were multiple isoforms expressed, and if all of them had the same transcriptional activity.

If my hypothesis were valid, cancerous cells would mostly express TCF-4 isoforms that are specific for transcriptional activation. Nevertheless, I have evaluated TCF-4 mRNA exon by exon, and identified 17 different TCF-4 isoforms by sequencing cDNAs generated from complete mRNA clones derived from four HCC cell lines. The full list of isoforms with cell line origin is shown in Table 2.3.

During the TCF-4 mRNA exon evaluation and the isolation of the isoforms, several features were observed:

Table 2.3. Summary of TCF-4 isoforms identified in HCC cells

Isoform	Included exons	HCC cell line origin	mRNA coding sequence (bp)	Protein size (aa)	Protein size (kDa)
TCF-4A	1-3, 5, 6, 7L, 8, 9L, 10-12, 17	HepG2, Hep3B, Huh7	1344	448	49.7
TCF-4B	1-3, 5, 6, 7L, 8, 9S, 10-12, 17	HepG2, Hep3B, Huh7, FOCUS	1329	443	49.2
TCF-4C	1-3, 5, 6, 7S, 8, 9S, 10-12, 17	Huh7	1317	439	48.7
TCF-4D	1-6, 7S, 8, 9S, 10-12, 17	Huh7	1386	462	51.4
TCF-4F	1-3, 5, 6, 7L, 8, 9L, 10-13, 17	HepG2, Hep3B	1395	465	51.6
TCF-4G	1-6, 7L, 8, 9S, 10-12, 17	HepG2, Hep3B, Huh7, FOCUS	1398	466	51.8
TCF-4H	1-6, 7L, 8, 9L, 10-12, 17	Huh7	1413	471	52.3
TCF-4I	1-6, 7L, 8, 9S, 10-13, 17	Hep3B	1449	483	53.7
TCF-4J	1-3, 5, 6, 7L, 8, 9S, 10-12, 15, 17	HepG2	1740	580	63.5
TCF-4K	1-3, 5, 6, 7L, 8, 9L, 10-12, 15, 17	Hep3B	1755	585	64.0
TCF-4L	1-6, 7L, 8, 9S, 10-12, 15, 17	HepG2	1824	608	66.6
TCF-4E	1-3, 5, 6, 7L, 8, 9S, 10-14, 17	Hep3B	1791	597	65.4
TCF-4M	1-6, 7L, 8, 9S, 10-14, 17	FOCUS	1860	620	67.9
TCF-4X	1-3, 5, 5A, 6, 7L, 8, 9L, 10-14, 17	HepG2	1920	165 336 438	17.9 36.9 49.8
TCF-4T	1, 2, 5, 6, 7L, 8, 9S, 10-12, 17	HepG2	1204	95 351	10.2 39.5
TCF-4Y-1	1-3, 5, 6, 7L, 8, 9S, 10-12, 14, 17 (8A)	Hep3B	1401	466	52.0
TCF-4+1	1-3, 5, 6, 7L, 8, 9S, 10-13, 17 (10A)	Hep3B	1719	573	62.5

1. As expected, presence of multiple bands in the agarose gels during the evaluation and the identification of 17 isoforms confirmed that HCC cells indeed expressed multiple splicing variants.

2. While the PCR products of the primer set 9SF/17R (without SxxSS) was ubiquitously observed in all cell lines, the PCR amplification from 9LF/17R (with SxxSS) was faintly detected only in Hep3B agarose gel (Figure 2.3). In addition, out of 17 isoforms identified, 11 of them expressed isoforms that did not contain the SxxSS motif corroborating with the idea that this motif is a strong repression-specific site and that cancerous cells avoided expressing isoforms with SxxSS.
3. Interestingly, the LVPQ motif was present in 15 out of 17 isoforms. During exon evaluations, the PCR products of the 1F/7LF (with LVPQ) primer pair were detected more than that of 1F/7SF (without LVPQ) in agarose gels. Preferential expression of isoforms with LVPQ motif in HCC cell lines suggested that this motif may not be specific for transcriptional repression.
4. Similar to SxxSS motif, 11 out of 17 isoforms did not contain exon 4 implying that this unique exon of TCF-4 might have an effect that was not favorable in HCC cells.
5. Identification of TCF-4X revealed presence of another exon, which was named exon 5A. Most intron sequences start with a dinucleotide gt (or gu in mRNA) and end with an ag (Table 2.1), and they are referred to as splice donor and splice acceptor sites, respectively. DNA sequence of the *TCF7L2* gene confirmed exon 5A to be a potential exon as it contained the ag/gt dinucleotides at its intron/exon/intron boundaries embedded within intron 5. Therefore, *TCF7L2* turned out to have 18^{th} exon, and since exon 5A was an alternative exon, the theoretical number of potential isoforms could go up to 1024.

6. TCF-4T was another unexpected isoform as it differed from the description of *TCF7L2* gene. Its mRNA sequence lacked the constitutively expressed exon 3, indicating that its sequence was yet of another alternatively spliced exon.
7. Identification of TCF-4Y isoforms confirmed that a frameshift mutation in the poly A tract, which was only observed in CRC this far, also took place in HCC. Therefore, HCC cells also exploited this mechanism where the potential co-repressor binding sites were eliminated by single nucleotide deletion or insertion.
8. Finally, identification of TCF-4B isoform in more than 35% of the total clones of HCC cell lines and 100% of normal liver tissues suggested that it could be the canonical isoform or that it was highly stabilized at mRNA level. In addition, TCF-4B was the only isoform that had been detected in a wide range of tissues, both normal and cancerous, in many other laboratories.

Overall, this study confirmed the expression of multiple TCF-4 isoforms in HCC. Moreover, the identification of a novel sequence of exon 5A and the potential of exon 3 being spliced out suggested that the *TCF7L2* gene had not been fully characterized. Recently, another isoform with a new exon sequence was detected by overlapping expressed sequence tags (ESTs) by The Sanger Institute (NM_001146284.1). This new exon was also embedded in the very long sequence of intron 5 (upstream of exon 5A), further implying the presence of additional rarely expressed exons. Finally, exon evaluation and splicing patterns of sequences in the 17 TCF-4 isoforms hinted that exon 4 and the SxxSS motif might be preferentially removed from the isoforms expressed in HCC. However, a comprehensive expression profile on all splicing variants needs to be performed. TCF-4T and Y isoforms were not pursued for further investigation.

2.5. METHODS

2.5.1. Cell culture

Human hepatoma cell lines HepG2, Hep3B, Huh7, and FOCUS, and human embryonic kidney cell line HEK293 were grown at 37°C in Dulbecco's modified Eagle's medium (DMEM, Sigma-Aldrich Corp., St. Louis, MO) supplemented with 10% (vol/vol) fetal bovine serum (FBS, Sigma-Aldrich), 1% (vol/vol) penicillin/streptomycin (Cellgro/Mediatech Inc., Manassas, VA) and 1% (vol/vol) L-glutamine (Cellgro).

2.5.2. Western Blot analysis

For preparing the whole cell lysates, all cell lines at 80% confluence were homogenized in RIPA lysis and extraction buffer (Pierce Biotechnology/Thermo Fisher Scientific (Pierce/Thermo), Rockford, IL) supplemented with Halt Protease and Phosphatase Inhibitor Cocktails (Pierce/Thermo), and sonicated. Protein concentration was determined with the BCA Protein Assay Kit (Pierce/Thermo) using bovine serum albumin as standard. One hundred μg of protein aliquots were resolved on sodium dodecyl sulfate-polyacrylamide gel electrophoresis and transferred onto polyvinylidene fluoride membranes (PerkinElmer Life Sciences, Inc., Waltham, MA) by electroblotting using a semi-dry transfer system. The membranes were blocked with 5% non-fat dry milk in Tris-buffered saline containing 0.1% Tween 20 and then probed with the following antibodies: mouse monoclonal anti-TCF-4 antibody diluted at 1:1,000 (Upstate/Millipore, Billerica, MA), mouse monoclonal anti-β-catenin antibody diluted at 1:1,000 (BD Biosciences, San Jose, CA), and anti-β-actin antibody diluted at 1:1,000 (Santa Cruz Biotechnology, Santa Cruz, CA). Each primary antibody was followed by incubation with a secondary horseradish-peroxidase antibody diluted 1:10,000, and the blots were

revealed with the chemiluminescence imaging Western Lightning (PerkinElmer). For controls, 1.0 μg of TCF-4I (short) and TCF-4J (long) isoforms, and empty vector (EV) plasmid were transiently transfected into 60% confluent HEK293 cells seeded in 6-well plates by using a *Trans*IT-LT1 transfection reagent (Mirus Bio LLC, Madison, WI) according to the manufacturer's instruction. Twenty-four hours post-transfection, the whole cell protein extraction was carried out as described above.

2.5.3. RT-PCR

Total cellular RNA was extracted from human HCC cell lines HepG2, Hep3B, Huh7, and FOCUS by using TRIzol reagent (Invitrogen, Carlsbad, CA). The quality of the RNA samples was determined by ethidium bromide-stained agarose gel electrophoresis verifying intact 18S and 28S rRNA bands. To obtain first-strand complementary DNA (cDNA), 0.5 μg of total RNA was treated with DNase-I, RNase inhibitor, and reverse transcribed with oligo-dT primer and AMV reverse transcriptase (Roche Diagnostics, Indianapolis, IN). The primer pairs used for evaluation of TCF-4 mRNA splicing variants exon by exon are given in Figure 2.3A and Table 2.2. Primer sets were designed by using the Primer3 website (http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi), and synthesized by Sigma-Aldrich.

2.5.4. Cloning of TCF-4 isoforms

For identifying different TCF-4 isoforms, RT-PCR was carried out with a primer pair 1F/17R that spanned from the beginning of exon 1 to the end of exon 17 regardless of splicing variations of exons in between them. The PCR products from all four HCC cell lines were cloned into the pCR2.1 vector using a TOPO TA cloning kit (Invitrogen). Twelve clones with positive inserts from each cell line (a total of 48 clones) were

obtained, and the individual clones were DNA sequenced. To verify the sequences, a program called ClustalW2, a general purpose multiple sequence alignment for DNA or proteins, (http://www.ebi.ac.uk/Tools/clustalw2/index.html) was used. The total genomic sequence of *TCF7L2* gene obtained from NCBI, and individual exon sequences described in [2] were used as the reference sequences for alignments. For transient transfections in WB analysis, the fourteen TCF-4 isoforms were sub-cloned into mammalian expression vector pcDNA3.1/myc-His (Chapter 4).

2.6. REFERENCES

1. Duval A, Busson-Leconiat M, Berger R and Hamelin R (2000) Assignment of the TCF-4 gene (TCF7L2) to human chromosome band 10q25.3. *Cytogenet Cell Genet*, 88: 264-265.
2. Duval A, Rolland S, Tubacher E, Bui H, Thomas G and Hamelin R (2000) The human T-cell transcription factor-4 gene: structure, extensive characterization of alternative splicings, and mutational analysis in colorectal cancer cell lines. *Cancer Res*, 60: 3872-3879.
3. Atcha FA, Syed A, Wu B, Hoverter NP, Yokoyama NN, Ting JH, Munguia JE, Mangalam HJ, Marsh JL and Waterman ML (2007) A unique DNA binding domain converts T-cell factors into strong Wnt effectors. *Mol Cell Biol*, 27: 8352-8363.
4. Korinek V, Barker N, Morin PJ, van Wichen D, de Weger R, Kinzler KW, Vogelstein B and Clevers H (1997) Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science*, 275: 1784-1787.
5. Ionov Y, Peinado MA, Malkhosyan S, Shibata D and Perucho M (1993) Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature*, 363: 558-561.
6. Duval A, Gayet J, Zhou XP, Iacopetta B, Thomas G and Hamelin R (1999) Frequent frameshift mutations of the TCF-4 gene in colorectal cancers with microsatellite instability. *Cancer Res*, 59: 4213-4215.
7. Korinek V, Barker N, Willert K, Molenaar M, Roose J, Wagenaar G, Markman M, Lamers W, Destree O and Clevers H (1998) Two members of the Tcf family

implicated in Wnt/beta-catenin signaling during embryogenesis in the mouse. *Mol Cell Biol*, 18: 1248-1256.

8. Howng SL, Huang FH, Hwang SL, Lieu AS, Sy WD, Wang C and Hong YR (2004) Differential expression and splicing isoform analysis of human Tcf-4 transcription factor in brain tumors. *Int J Oncol*, 25: 1685-1692.

9. Shiina H, Igawa M, Breault J, Ribeiro-Filho L, Pookot D, Urakami S, Terashima M, Deguchi M, Yamanaka M, Shirai M, Kaneuchi M, Kane CJ and Dahiya R (2003) The human T-cell factor-4 gene splicing isoforms, Wnt signal pathway, and apoptosis in renal cell carcinoma. *Clin Cancer Res*, 9: 2121-2132.

10. Brantjes H, Barker N, van Es J and Clevers H (2002) TCF: Lady Justice casting the final verdict on the outcome of Wnt signalling. *Biol Chem*, 383: 255-261.

11. Kim M, Lee HC, Tsedensodnom O, Hartley R, Lim YS, Yu E, Merle P and Wands JR (2008) Functional interaction between Wnt3 and Frizzled-7 leads to activation of the Wnt/beta-catenin signaling pathway in hepatocellular carcinoma cells. *J Hepatol*, 48: 780-791.

12. Lee HC, Kim M and Wands JR (2006) Wnt/Frizzled signaling in hepatocellular carcinoma. *Front Biosci*, 11: 1901-1915.

13. Takigawa Y and Brown AM (2008) Wnt signaling in liver cancer. *Curr Drug Targets*, 9: 1013-1024.

14. Cui J, Zhou X, Liu Y, Tang Z and Romeih M (2003) Wnt signaling in hepatocellular carcinoma: analysis of mutation and expression of beta-catenin, T-cell factor-4 and glycogen synthase kinase 3-beta genes. *J Gastroenterol Hepatol*, 18: 280-287.

15. Ikenoue T, Ijichi H, Kato N, Kanai F, Masaki T, Rengifo W, Okamoto M, Matsumura M, Kawabe T, Shiratori Y and Omata M (2002) Analysis of the beta-catenin/T cell factor signaling pathway in 36 gastrointestinal and liver cancer cells. *Jpn J Cancer Res*, 93: 1213-1220.

16. Jiang Y, Zhou XD, Liu YK, Wu X and Huang XW (2002) Association of hTcf-4 gene expression and mutation with clinicopathological characteristics of hepatocellular carcinoma. *World J Gastroenterol*, 8: 804-807.

17. Zhao DH, Hong JJ, Guo SY, Yang RL, Yuan J, Wen CY, Zhou KY and Li CJ (2004) Aberrant expression and function of TCF4 in the proliferation of hepatocellular carcinoma cell line BEL-7402. *Cell Res*, 14: 74-80.

18. Schmitt-Graeff A, Ertelt-Heitzmann V, Allgaier HP, Olschewski M, Nitschke R, Haxelmans S, Koelble K, Behrens J and Blum HE (2005) Coordinated expression of cyclin D1 and LEF-1/TCF transcription factor is restricted to a subset of hepatocellular carcinoma. *Liver Int*, 25: 839-847.

19. Yanagitani A, Yamada S, Yasui S, Shimomura T, Murai R, Murawaki Y, Hashiguchi K, Kanbe T, Saeki T, Ichiba M, Tanabe Y, Yoshida Y, Morino S, Kurimasa A, Usuda N, Yamazaki H, Kunisada T, Ito H, Murawaki Y and Shiota G (2004) Retinoic acid receptor alpha dominant negative form causes steatohepatitis and liver tumors in transgenic mice. *Hepatology*, 40: 366-375.

20. Lee HC, Tian B, Sedivy JM, Wands JR and Kim M (2006) Loss of Raf kinase inhibitor protein promotes cell proliferation and migration of human hepatoma cells. *Gastroenterology*, 131: 1208-1217.

21. He L, Isselbacher KJ, Wands JR, Goodman HM, Shih C and Quaroni A (1984) Establishment and characterization of a new human hepatocellular carcinoma cell line. *In Vitro*, 20: 493-504.

22. Nakabayashi H, Taketa K, Miyano K, Yamane T and Sato J (1982) Growth of human hepatoma cells lines with differentiated functions in chemically defined medium. *Cancer Res*, 42: 3858-3863.

23. Aden DP, Fogel A, Plotkin S, Damjanov I and Knowles BB (1979) Controlled synthesis of HBsAg in a differentiated human liver carcinoma-derived cell line. *Nature*, 282: 615-616.

24. Lopez-Terrada D, Cheung SW, Finegold MJ and Knowles BB (2009) Hep G2 is a hepatoblastoma-derived cell line. *Hum Pathol*, 40: 1512-1515.

25. Pukrop T, Gradl D, Henningfeld KA, Knochel W, Wedlich D and Kuhl M (2001) Identification of two regulatory elements within the high mobility group box transcription factor XTCF-4. *J Biol Chem*, 276: 8968-8978.

26. Cuilliere-Dartigues P, El-Bchiri J, Krimi A, Buhard O, Fontanges P, Flejou JF, Hamelin R and Duval A (2006) TCF-4 isoforms absent in TCF-4 mutated MSI-H colorectal cancer cells colocalize with nuclear CtBP and repress TCF-4-mediated transcription. *Oncogene*, 25: 4441-4448.

CHAPTER 3

3. EXPRESSION OF TCF-4 ISOFORMS IN HCC

3.1. ABSTRACT

TCF/LEF family of transcription factors of Wnt signaling pathway do not act as classical transcription factors as they can be switched between transcriptional repressors and activators depending on the absence and presence of Wnt signals, respectively. The family consists of four members in mammals: TCF-1, LEF-1, TCF-3, and TCF-4. Apart from common structural domains they share, there are several short sequences that are optionally expressed via alternative splicing. These sequences may give functional diversity between family members. For example, LVPQ, SxxSS, and CtBP motifs are known to have added repressive effect on target gene transcription. Correspondingly, TCF-3 that contains these motifs as obligate sequences is exclusively expressed to repress Wnt target genes, while TCF-1 and LEF-1, which do not have these sequences, are specifically abundant during active Wnt signaling. However, TCF-4 is expressed in both transcriptional repression and activation during development. This bi-functional property of TCF-4 has been explained by the relative amount of co-regulators that switch it between activator and repressor. In contrast, I propose the existence of TCF-4 isoforms that are specific for transcriptional activation and for transcriptional repression, and I further hypothesize that expression of each isoform depends on its functional feature.

Identification of more than 14 distinct TCF-4 isoforms in HCC revealed the presence of alternative motifs in various combinations. Since Wnt signaling is aberrantly activated in HCC, it is important to determine their expression levels. Here, I analyze the endogenous expression of these splice variants in HCC cells, as well as in HCC tumors, their matching peritumor, and normal liver tissues by RT-PCR.

3.2. INTRODUCTION

The four members of the TCF/LEF family of transcription factors, TCF-1, LEF-1, TCF-3, and TCF-4, are the most downstream components of the canonical Wnt signaling pathway [1-4]. The structural organization of TCF/LEFs is schematically shown in Figure 3.1A. Their common properties, such as binding to similar WRE and activating Wnt target genes are provided by their highly conserved HMG box and the N-terminal BCBD, respectively [5]. Conversely, transcription repression is mediated by the TLE-binding site, which is also well-conserved and constitutively expressed in all members [6,7]. Consequently, TCF/LEF transcription factors have been described as bi-functional context-dependent regulators. The option between transcription activation *vs*. repression is determined when the balance between the amount of co-activators and co-repressors is shifted toward one direction [8]. If this is how the function of TCF/LEFs is regulated, then why does TCF-3 exclusively act as transcription repressor during development? Despite the presence of BCBD, TCF-3 has hardly been reported to interact with β-catenin and activate Wnt target genes. The alternatively spliced C-terminal tail and part of the CRD are expressed differently in TCF/LEFs, which may contribute to the specific functional properties for each family member. Apart from the TLE-binding domain, the CRD contains two alternatively spliced LVPQ and SxxSS motifs that have been reported to have repressive activities [9]. In addition, there are two binding sites for co-repressor CtBP in the long C-terminal tail [10]. Interestingly, TCF-3 constitutively expresses LVPQ, SxxSS, and CtBP-binding sites, while TCF-1 and LEF-1, which are mostly employed for target gene activation, do not even contain these repression-specific domains (Figure 3.1A). On the other hand, TCF-4 can generate any combination of these

motifs by alternative splicing, and TCF-4 isoforms with structural resemblance to TCF-1, LEF-1, and TCF-3, at least with respect to these motifs, may function in a similar pattern.

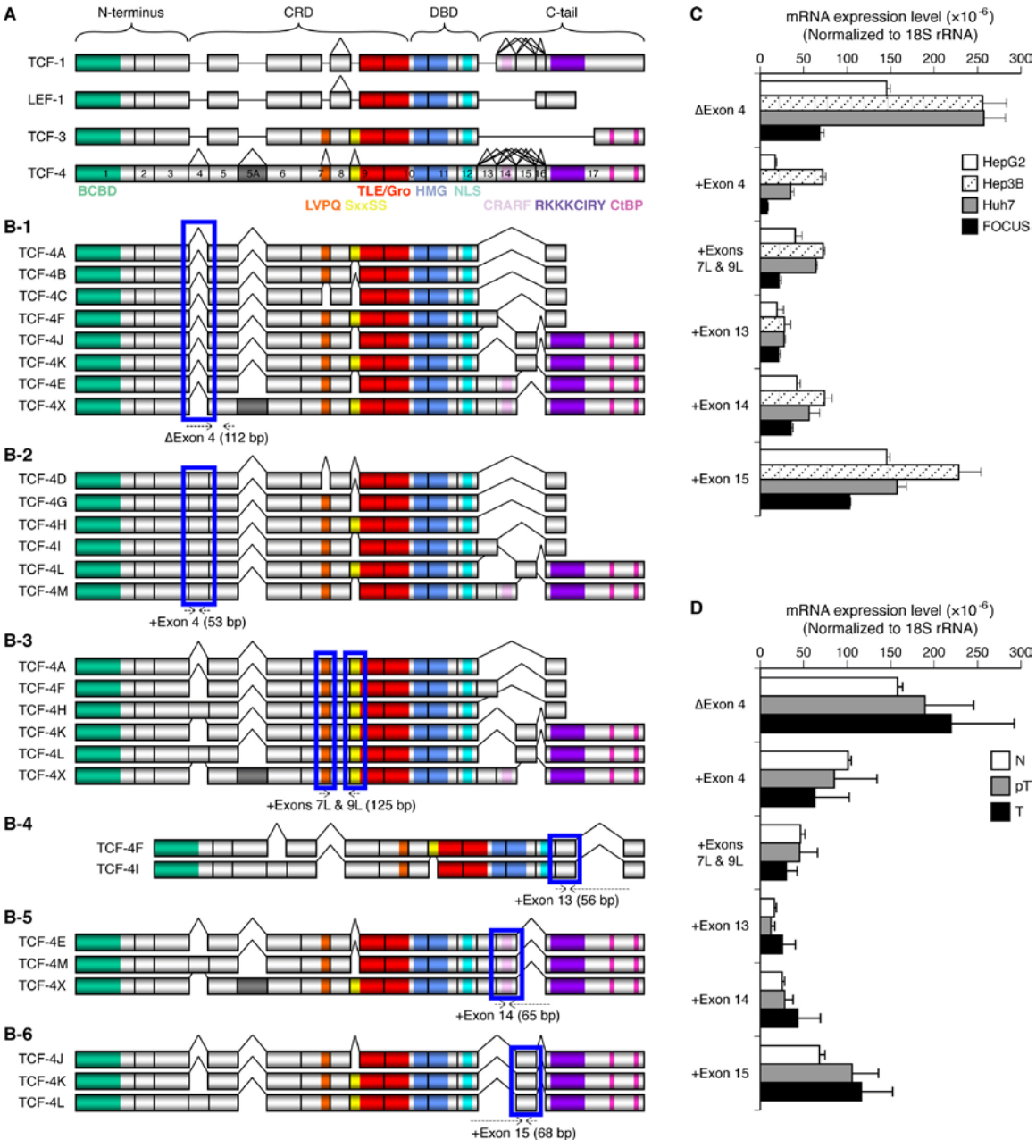


Figure 3.1. Expression of TCF-4 mRNA isoforms by qPCR

A. Schematic representation of human TCF/LEF family of transcripotion factors reflecting the newly identified exon, exon 5A. The conserved regions are marked with different colors as in

Figure 1.2.2. In TCF-4, the numbers inside boxes indicate the exon order of the gene (*TCF7L2*) including exon 5A (dark grey box). *B*. Fourteen isoforms were categorized into six different groups according to their exon characteristics (blue rectangles) and correspondingly, six sets of primers were designed. *B-1*. Isoforms that lacked exon 4 (ΔExon 4) included TCF-4A, B, C, F, J, K, E, and X. The forward primer was complementary to the DNA sequences bordering exons 3 and 5 while the reverse primer was within exon 5. The size of the qPCR product detected by these primers was 112 bps. *B-2*. Isotypes that contained exon 4 (+Exon 4) were TCF-4D, G, H, I, L, and M. The primer pair detected sequences within exon 4 and generated a 53 bp-long product. *B-3*. Splicing variants that included the LVPQ and SxxSS motifs (+Exons 7L & 9L) were TCF-4A, F, H, K, L, and X. The primer set detected each motif and produced a 125 bp-long PCR product. *B-4*. Isoforms that harbored exons 13 (+Exon 13) were TCF-4F and I. The reverse primer was designed to detect sequences along the border of exons 13 and 17. Therefore, this group excluded isoforms shown in *B-5* although they also composed of exon 13. *B-5*. Isotypes with exon 14 (+Exon 14) were TCF-4E, M, and X. The reverse primer was also designed to complement sequences along exons 14 and 17 to exclude other potential isoforms with exon 14 of different group. B-6. Splicing variants with exon 15 (+Exon 15) were TCF-4J, K, and L. Similarly, the forward primer was devised to anneal to the sequences along the border of exons 12 and 15. *C*. The mRNA expression levels of each isoform group in HepG2, Hep3B, Huh7, and FOCUS HCC cell lines measured by qPCR. The qPCR was carried out using 5 ng of cDNA samples and performed in triplicates. Results were presented as average values with SD from two independent experiments. *D*. The average mRNA expression levels of each isoform in normal liver (N), HCC tumor (T) and matching peritumor (pT) tissues assayed by qPCR. The experiments were performed in duplicates and results were illustrated as overall averages with SD across each tissue type.

During *Xenopus* development, XLEF-1 and XTCF-1 are upregulated on the dorsal side of the embryo to activate the expression of dorsal genes [2,11], while XTCF-3 is overexpressed to represses the transcription of the same genes on the ventral side [10,12]. Similar expression pattern is observed in mouse development as well [3,13,14]. As an extension to my hypothesis, I further propose that the potential activator isoforms of TCF-4, specific for Wnt target gene expression, may show structural resemblance to TCF-1 and/or LEF-1 and may be overexpressed in cancerous cells. In contrast, the TCF-4 isotypes specialized for target gene repression may appear structurally similar to TCF-3 and their expression may be downregulated in tumors. I described the identification of 17 distinct TCF-4 isoforms in Chapter 2. In this chapter, I analyze the endogenous expression of 14 of these splice variants (TCF-4A through M, and X) in the original cell lines they were isolated, in paired HCC tissues, and in normal liver tissue samples by both quantitative real time RT-PCR (qPCR) and semi-quantitative RT-PCR (semi-qPCR). In the latter, the RT-PCR products were run in agarose gels to distinguish and quantitate individual isoforms.

3.3. RESULTS

3.3.1. Expression of TCF-4 mRNA isoform groups by qPCR

As an attempt to confirm the expression of the newly identified TCF-4 isoforms in HCC by qPCR, I designed primer sets that would distinguish a group of isoforms with same exon characteristics. In this regard, six sets of primer pairs that detected isoforms without exon 4 (ΔExon 4), with exon 4 (+Exon 4), with the LVPQ and SxxSS motifs (+Exons 7L & 9L), and with exons 13 (+Exon 13), 14 (+Exon 14), and 15 (+Exon 15) were generated as shown in Figure 3.1B and Table 3.1. The mRNA expression of these groups of isoforms in hepatoma cell lines HepG2, Hep3B, Huh7, FOCUS are shown in Figure 3.1C. The isoforms that did not contain exon 4 were the most abundantly expressed group, while the ones with exon 4 were produced at much lower level indicating that HCC cell lines prefer to splice the exon 4 out of the transcript. The isotypes that harbored exon 15 was the second highest expressed group, whereas the splicing variants that contained exon 13 were the lowest. The isoforms that incorporated exon 14 and the ones that included both LVPQ and SxxSS motifs were expressed at similarly low levels. Interestingly, the most carcinogenic FOCUS HCC cell line expressed TCF-4 isoforms at low levels compared to other hepatoma cell lines.

The qPCR was also performed with the same primer sets in normal liver (N), HCC tumor (T) and matching paired peritumor (pT) tissues, and the average expression levels are shown in Figure 3.1D. The overall expression pattern was consistent with the cell line results: the isotypes that lacked exon 4 were expressed at the highest level, the group with exon 15 the second highest, ones with exon 13 the lowest, and the isoforms with LVPQ and SxxSS motifs were expressed moderately low. However, because these

isoform groups were not exclusive for one isoform, i.e. an isotype could be categorized into more than one group, this method was not very informative on individual isoforms. For example, TCF-4X was sorted into three different groups (ΔExon 4, +Exons 7L & 9L, and +Exon 14), one of which illustrated a high expression level while the other two showed lower values, rendering it impossible to deduce the expression level of this particular isoform.

Table 3.1. Primer sequences used for qPCR

Target	Primer	Sequence (5′ – 3′)
ΔExon 4	3-5F	cgcccgaacctatctccaga
	5R	gacaatgtgtgccggtgat
+Exon 4	4F	agtccggcagcacacattac
	4R	tgcaatctggtgttcaatcg
+Exons 7L & 9L	7LF	ggatggttagtaccacagca
	9LR	gggaacctagaagacagaaagc
+Exon 13	13F	cagcgaatgtttcctaaatcc
	13-17R	ttttttctcctgtaatcggagg
+Exon 14	14F	acctcagcgctcctaagaaa
	14-17R	gccgcaccagttattctgtt
+Exon 15	12-15F	gcaaatactccaaagaagtgtcg
	15R	cacggtttgcaccataaagtc
18S rRNA	F	ggacacggacaggattgaca
	R	acccacggaatcgagaaaga

3.3.2. Expression of TCF-4 mRNA isoforms in HCC cells by semi-qPCR

The qPCR results were impractical at individual isoforms and one primer set was detecting multiple isotypes. This challenge was overcome by designing eight splice site-specific primer pairs such that each would amplify the fewest number of isoforms in RT-PCR (Table 3.2 and Figure 3.2A), and by running the PCR products in agarose gels to

separate and quantitate the band intensity of each isoform by densitometry. For example, primer pair P1F/P1R detected and amplified isoforms TCF-4A, B, and C only, and they could be separated in an agarose gel at individual bands. As shown in Figure 3.2B and C, expression of 10 isoforms was revealed in four HCC cell lines by semi-qPCR. TCF-4A, H, K and X isoforms were undetectable, possibly due to their low expression levels and/or the low sensitivity of the technique. Strikingly, TCF-4A, H, and K differed by only SxxSS motif from their closely related isotypes TCF-4B, G, and J, respectively (blue rectangles in Figure 3.2A), which were expressed at high levels. This suggested that isoforms with the SxxSS motif were substantially downregulated in HCC cells.

Table 3.2. Primer sequences used for semi-qPCR

Primer	Sequence (5′ – 3′)
P1F	cccgaacctatctccagatg
P1R	ttttctccattggtctctcc
P2F	agcacacattactctgcgta
P2R	ttttctcctgtaatcgga
P3R	ttttctcctgcacggtttg
P4R	tttttctcctgcaagggc
GAPDH-F	cttagcacccctggccaag
GAPDH-R	gatgttctggagagccccg

Consistent with the qPCR results, the FOCUS cell line expressed the isoforms at lower levels than compared to other HCC cells. Moreover, TCF-4B and C isotypes, which were detected among the highest expression group, ΔExon 4, during qPCR (Figure 3.1B-1), were also expressed at high levels when measured by semi-qPCR. Interestingly, TCF-4B was found only in well differentiated HepG2 and Hep3B cells, while TCF-4C was expressed in more carcinogenic cells, Huh7 and FOCUS (Figure 3.2C). Furthermore,

TCF-4J isoform that was sorted in both highly expressed groups, ΔExon 4 and +Exon 15 (Figure 3.1B-1 and B-6), was also detected at high levels by the semi-qPCR method. Finally, TCF-4F and I isoforms, which were detected at low levels as +Exon 13 group (Figure 3.1B-4) by qPCR, were uniformly expressed at lower levels in all cell lines as well when quantitated by semi-qPCR.

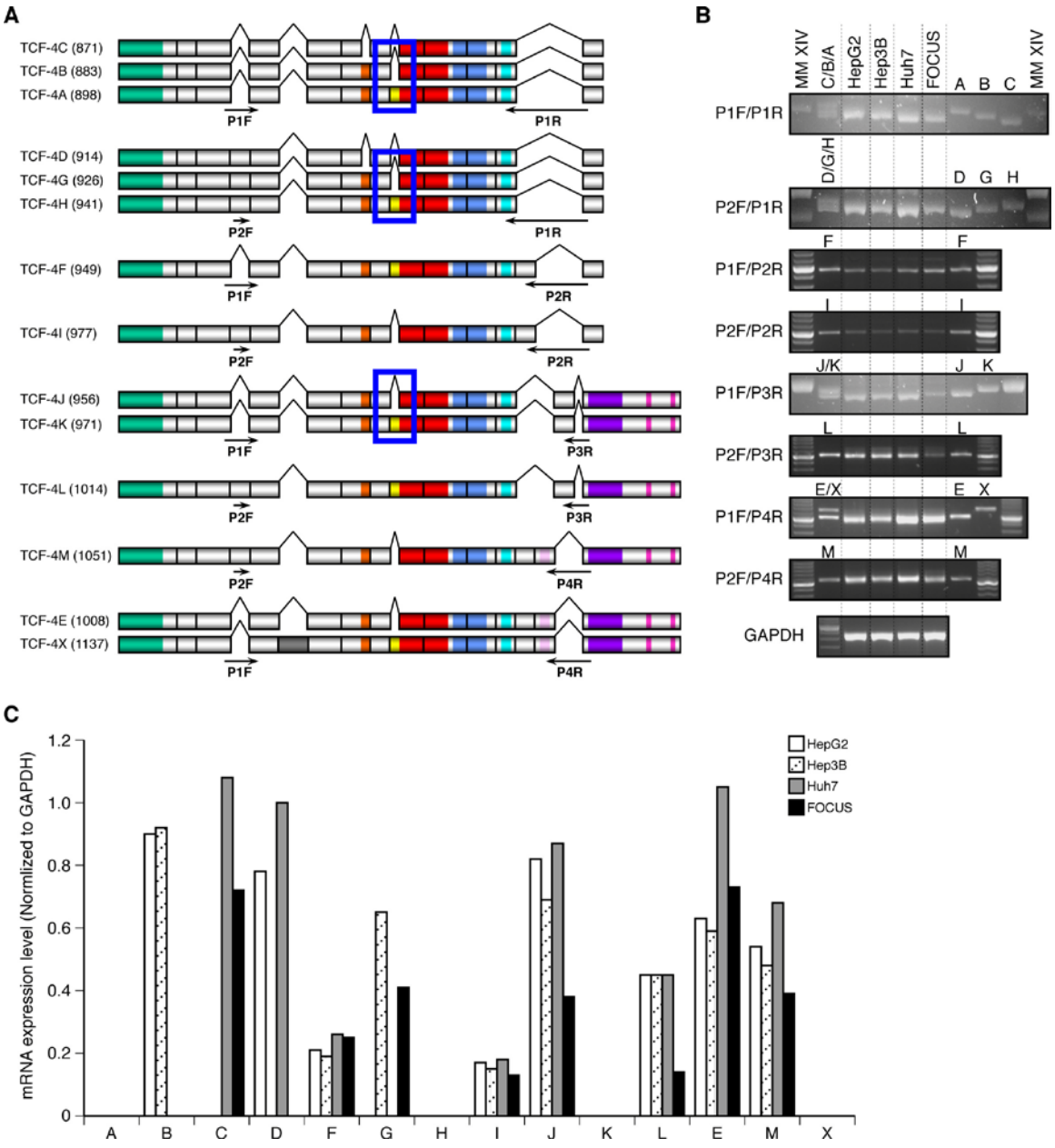


Figure 3.2. Expression of TCF-4 mRNA isoforms in HCC cells by semi-qPCR

A. Summary of 14 isoforms of TCF-4 considered for further investigation. Schematic positions of primer sets used for semi-qPCR were shown below isoforms they amplify. Numbers in parentheses represent the size of PCR products in bp. The blue rectangles were sketched to emphasize the closely related isoforms that differed by SxxSS motif only (TCF-4A *vs*. B, TCF-4H *vs*. G, and TCF-4K *vs*. J). *B*. Agarose gel electrophoresis of PCR products stained by ethidium bromide. Each TCF-4 isoform plasmid was transfected into Huh7 cells and used as a positive control for semi-qPCR. For PCRs that generated multiple products, positive controls were applied to lanes as both combined (C/B/A) and single (A, B, and C) species, and indicated above each gel. Migration of positive controls revealed the expected sizes. The PCR products detected by primer sets, P1F/P1R (A, B, and C), P2F/P1R (D, G, and H), and P1F/P3R (J and K), were run in 3.5% NuSieve 3:1 agarose gels for efficient separation of similar-sized amplicons. Although the isoforms, TCF-4E and X, were generated by a single primer set, P1F/P4R, they were separated in a 2% agarose gel as their size difference was large enough. The rest of the PCR products were analyzed in 2% agarose gels. The expression of GAPDH was used as an endogenous control. The 100-base pair molecular marker (MM XIV) served as control for band intensity between different agarose gels. *C*. The mRNA expression level of 14 distinct TCF-4 isoforms in HCC cell lines. TCF-4 expression in each sample was normalized to GAPDH.

3.3.3. Expression of TCF-4 mRNA isoforms in HCC tissues by semi-qPCR

Expression levels of all isoforms in twenty-seven paired HCC tumor and matching peritumor tissue samples (Table 3.3) were analyzed by comparing their average expression with that of normal liver tissues (dotted lines). The mRNA expression levels in these tissues are described below in detail. Representative examples of agarose gels are shown in the section A of the following figures. All data are shown in a same scale for easy comparison.

Table 3.3. Patients and histological features of liver tumors

Abbreviations: CCC, cholangiocellular carcinoma; HB, hepatoblastoma; SN, single nodule; PNE, perinodular extension; MN, multiple nodule; MD, moderately differentiated; PD, poorly differentiated, LC, liver cirrhosis; CH, chronic hepatitis

No	Age	Sex	Virus	Tumor diagnosis	Tumor size (cm)	Nodule	Grade	Vascular invasion	Lymph node metastasis	pT liver
1	46	M	HBV	HCC/CCC	6.8	SN	MD	(-)	(-)	LC
2	56	M	HBV	HCC	2.7	SN	MD	(-)	(-)	CH
3	45	M	HBV	HCC	15.0	SN-PNE	PD	(-)	(-)	LC
4	47	M	HBV	HCC	3.5	SN	MD	(-)	(-)	LC
5	57	F	HBV	HCC	1.7	SN	MD	(-)	(-)	CH
6	55	F	HBV	HCC	3.0	MN	MD	(-)	(-)	LC
7	67	M	HBV	HCC	8.5	MN	MD	(-)	(-)	CH
8	61	F	HBV	HCC	3.0	SN	MD	(-)	(-)	LC
9	37	M	HBV	HCC	9.5	MN	MD	(-)	(-)	CH
10	53	M	HBV	HCC	4.2	SN	MD	(-)	(-)	LC
11	57	M	HBV	HCC	16.0	MN	MD	(+)	(-)	CH
12	55	M	HBV	HCC	11.0	MN	PD	(-)	(-)	LC
13	43	M	HBV	HB	14.0	SN	PD	(-)	(-)	CH
14	22	M	HBV	HCC	10.0	SN-PNE	PD	(-)	(-)	CH
15	67	F	(-)	HCC	6.0	SN	PD	(-)	(-)	Normal
16	52	M	HBV	HCC	5.0	SN	PD	(+)	(+)	LC
17	52	M	HBV	HCC	5.0	SN	PD	(+)	(+)	LC
18	51	M	HBV	HCC	3.9	SN	PD	(+)	(-)	LC
19	26	F	HBV	HCC/CCC	5.0	MN	PD	(-)	(-)	CH
20	56	M	HBV	HCC	6.0	SN	PD	(-)	(-)	LC
21	49	M	HBV	HCC	3.4	SN-PNE	PD	(-)	(-)	CH
22	63	M	HBV	HCC	8.0	MN	PD	(+)	(-)	LC
23	42	M	HBV	HCC	4.7	SN	MD	(-)	(-)	LC
24	50	M	HBV	HCC	4.5	MN	MD	(-)	(-)	LC
25	41	M	HBV	HCC	4.0	SN	MD	(-)	(-)	CH
26	44	M	HBV	HCC	3.8	SN	MD	(-)	(-)	LC
27	69	F	HBV	HCC	8.0	SN	MD	(-)	(-)	LC

3.3.3.1. Expression of isoforms TCF-4A, B, and C

The RT-PCR with the primer pair P1F/P1R generated three isoforms, TCF-4A, B, and C (Figure 3.2A and Figure 3.3). Isoform TCF-4A was scarcely expressed in tumor tissues with 30% (only in 8 out of 27 (8/27) samples) and their overall expression level was lower than the average expression in normal liver tissues (Figure 3.3B, left chart) further suggesting that TCF-4A is preferentially downregulated in HCC. On the other hand, 89% (24/27) of peritumor samples expressed TCF-4A, and the expression level of 85% (23/27) was within the range of expression in normal tissues as defined by average expression in normal samples ± 3× standard deviation (SD). Correspondingly, the overall average values between normal and peritumor tissues were similar (Figure 2B, right chart), indicating that peritumoral tissues behaved similarly to normal samples.

As shown in Figure 3.3C, TCF-4B isoform was uniformly expressed throughout all types of tissues. The expression pattern of majority of peritumor (63%, 17/27) and tumor (70%, 19/27) values was within the range of normal expression (Figure 3.3C, left chart). When paired samples were compared, the TCF-4B was expressed higher in tumor than in peritumor tissues (74%, 20/27). Despite these differences in individual tissue data, the overall average expression rate was similar between normal, peritumor, and tumor tissues (Figure 3.3C, right chart), implying that TCF-4B might not be required specifically for tumorigenesis, but might be involved in regular homeostatic maintenance of the cell. Similar to cell line data, comparison between TCF-4A and B in tumor tissues revealed that the presence of SxxSS motif in TCF-4A lead to its elimination or downregulation in HCC tumors.

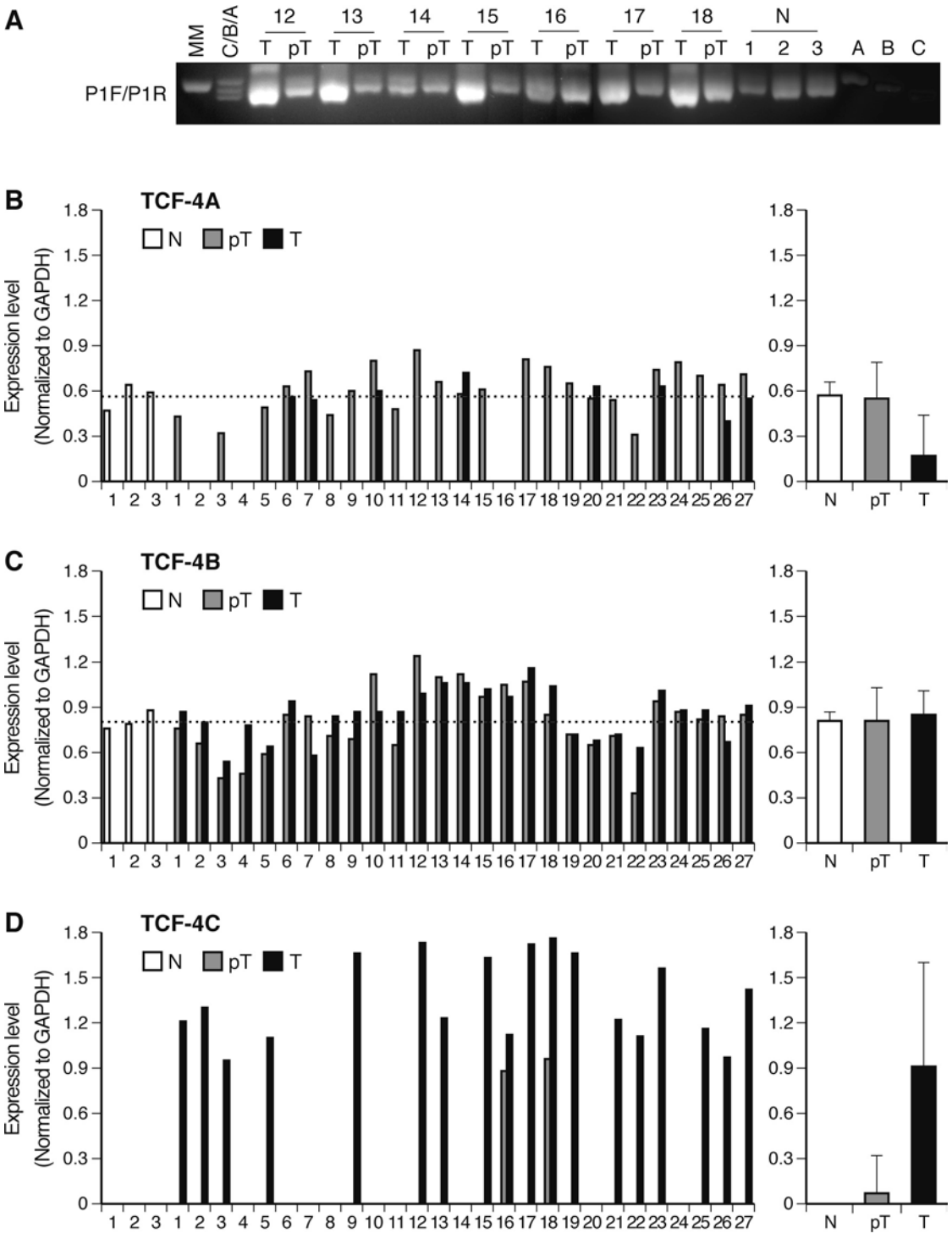


Figure 3.3. Expression of TCF-4A, B, and C isoforms in HCC tissues by semi-qPCR

RNA expression levels of TCF-4A (*B*), TCF-4B (*C*), and TCF-4C (*D*) measured by semi-qPCR with primer pair P1F/P1R in twenty-seven HCC tumor (T) and matching peritumor (pT) tissues and in three normal liver samples (N). *A*. A representative picture for separating all three isoforms

by running 3.5% NuSieve 3:1 agarose gel. MM, 1: 100 bp molecular marker; C/B/A, combined positive controls for TCF-4C, B, and A; single positive controls are indicated by their isoform names. The positive controls were obtained by transfecting corresponding TCF-4 isoform into Huh7 cells and performing RT-PCR. In *B* – *D*, the left charts illustrate the expression levels in individual samples, while the right charts show the average values in each category of samples. The white bars demonstrate the expression of isoforms in normal liver, grey bars indicate the peritumoral samples, and the black bars represent the matching tumor tissues. The dotted lines in the left charts display the average values in normal liver. Note that there is no dotted line in (*D*), because TCF-4C was not expressed in normal liver tissues.

TCF-4C, which lacked both LVPQ and SxxSS motifs, was exclusively and highly expressed in tumor tissues (Figure 3.3D). No expression was detected in normal tissues, and only in 7% (2/27) of peritumoral samples showed expression of this isoform, strongly suggesting a role for TCF-4C in progression of HCC.

3.3.3.2. Expression of isoforms TCF-4D, G, and H

The isoforms TCF-4D, G, and H were generated by RT-PCR primer pair P2F/P1R (Figure 3.2A and Figure 3.4). TCF-4D was the only isoform that expressed in the least number of tissues. Its expression was detected in 67% (2/3) of normal, 22% (6/27) of peritumor, and 41% (11/27) of tumor tissues (Figure 3.4B, left chart). However, once its transcript was detected, the expression levels were very high in both peritumor and tumor samples. The average values, calculated by including all 27 samples, together with 3× SD showed that both tumor and peritumor tissues still had similar values compared to that of normal samples (Figure 3.4B, right chart). It is interesting to note that, similar to TCF-4C, TCF-4D also lacked the LVPQ and SxxSS motifs and was expressed abundantly in HCC. The presence of exon 4 is the only difference between the two isoforms, and this

exon might be directing the differential expression pattern, suggesting that exon 4 might have a leaky repressive effect.

Similar to TCF-4B, TCF-4G was also uniformly expressed in almost all samples (Figure 3.4C, left chart). The expression pattern of majority of peritumor (81%, 22/27) and tumor (85%, 23/27) values was within the average range of normal expression (Figure 3.4C, left chart). When matching samples were compared, the number of samples with difference between tumor and peritumor tissues was similar: while the expression level of TCF-4G was higher in 52% (14/27) of tumor, 48% (13/27) was higher in peritumor samples. In addition, the overall average expression level of TCF-4G was comparable between normal, peritumor, and tumor tissues (Figure 3.4C, right chart), suggesting a similar role for TCF-4G with TCF-4B. Comparison of their expression (Figure 3.3C *vs*. Figure 3.4C), revealed near identical pattern, but TCF-4G was less abundant in tumor tissues. Structurally, the only difference between the two was exon 4 in TCF-4G, and the presence of this exon was again implicated in slightly lowered expression in tumor.

As shown in Figure 3.4D, TCF-4H was expressed in 78% (21/27) of peritumor, and 70% (19/27) of tumor tissues (Figure 3.4D, left chart). Almost half of their expression levels was lower than that of normal tissues, indicating that TCF-4G might also be downregulated in HCC. However, the overall average values between tissue types did not show any difference (Figure 3.4D, right chart). Interestingly, similar to TCF-4A and B splice variants, TCF-4H and G are isoforms that differed by SxxSS motif, and yet again presence of this motif in TCF-4H resulted in its downregulation in many tumor tissues (48%, 13/27).

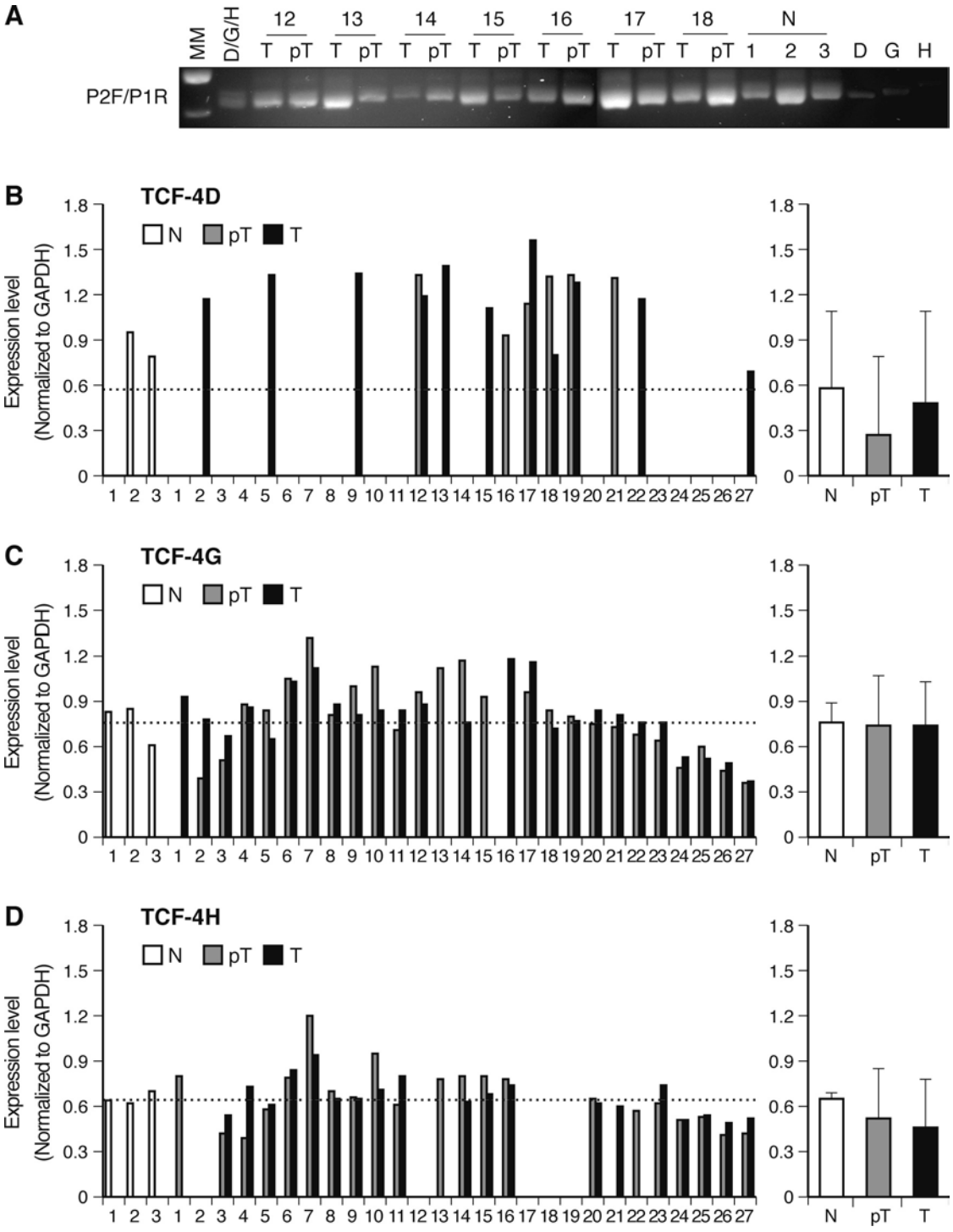


Figure 3.4 Expression of TCF-4D, G, and H isoforms in HCC tissues by semi-qPCR

RNA expression levels of TCF-4D (*B*), TCF-4G (*C*), and TCF-4H (*D*) measured by semi-qPCR with primer pair P2F/P1R in normal and HCC tissue samples. *A*. A representative picture for separating all three isoforms by running 3.5% NuSieve 3:1 agarose gel.

3.3.3.3. Expression of isoforms TCF-4J and K

Another pair of isoforms that varied from one another by SxxSS motif was TCF-4J and K, though they have long C-terminal tails (Figure 3.2A). Semi-quantitative RT-PCR was performed with the primer pair P1F/P3R to analyze their mRNA levels (Figure 3.5). TCF-4J, the one without SxxSS motif, was overexpressed in majority of tumor tissues compared to both average expression in normal (85%, 23/27) and corresponding paired peritumoral (85%, 23/27) samples (Figure 3.5B, left chart). The overall higher expression of TCF-4J in tumor tissues (Figure 3.5B, right chart) suggested that TCF-4J was involved in hepatocarcinogenesis.

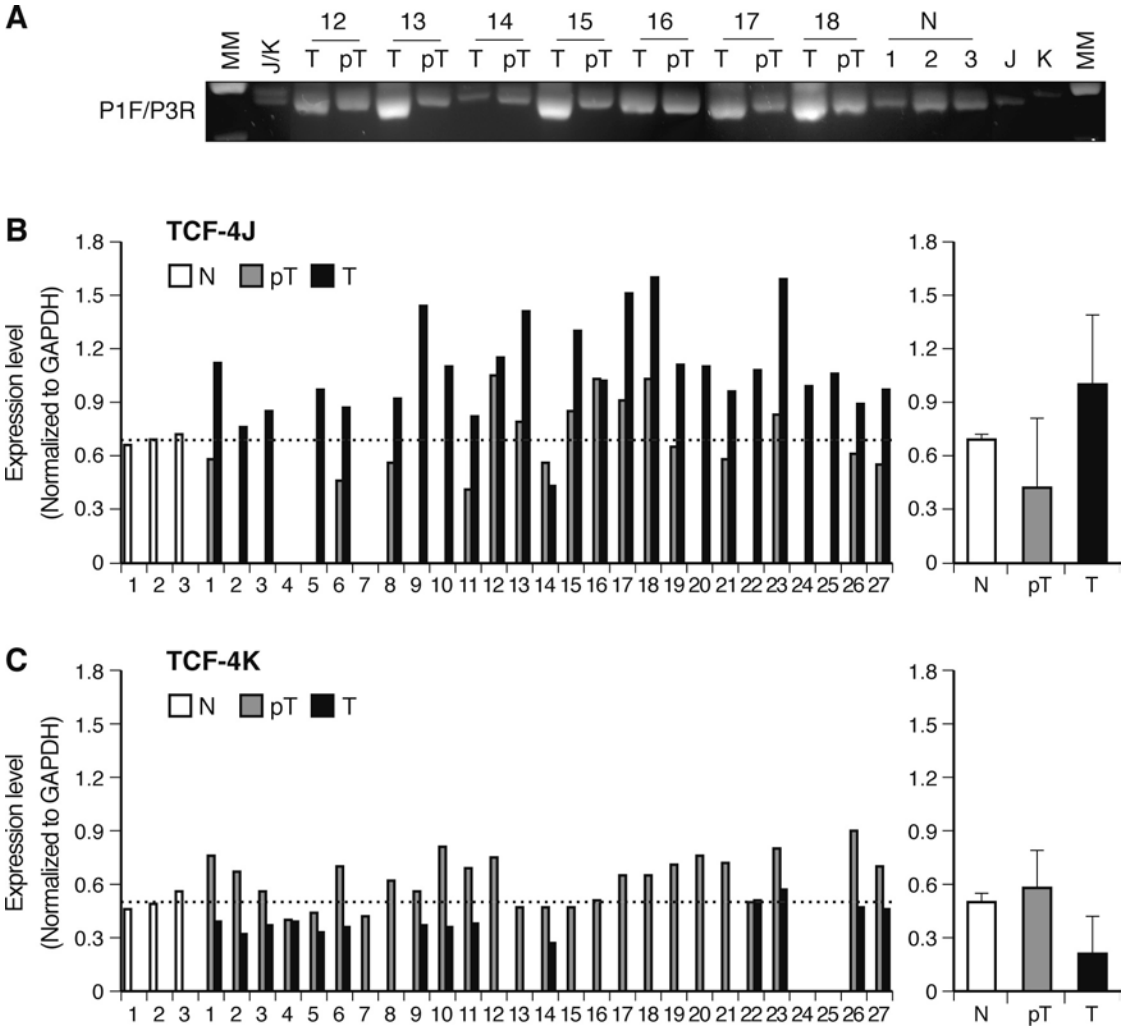


Figure 3.5 Expression of TCF-4J and K isoforms in HCC tissues by semi-qPCR

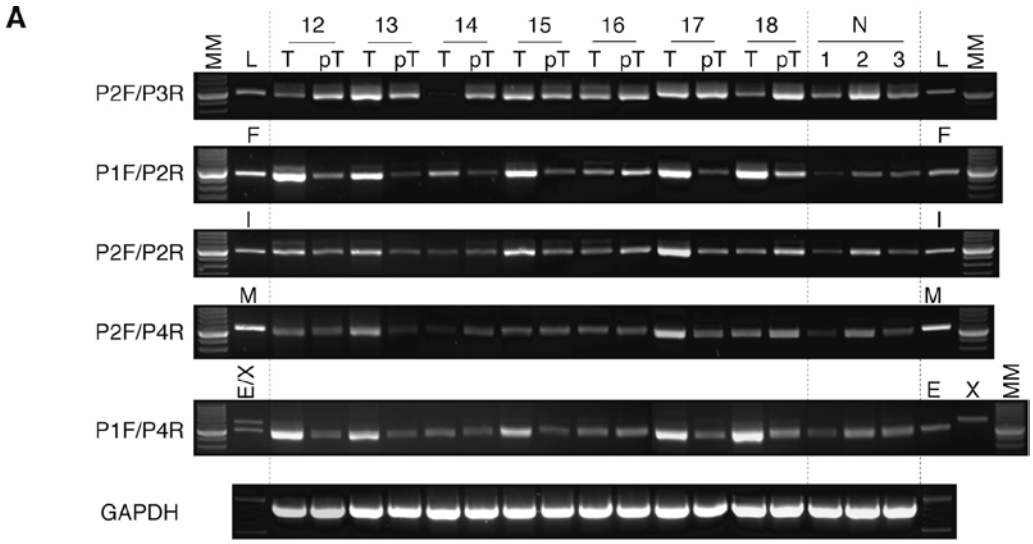
RNA expression levels of TCF-4J (*B*) and TCF-4K (*C*) measured by semi-qPCR with primer pair P1F/P3R in normal and HCC liver. *A*. A representative picture for separating both isoforms by running 3.5% NuSieve 3:1 agarose gel.

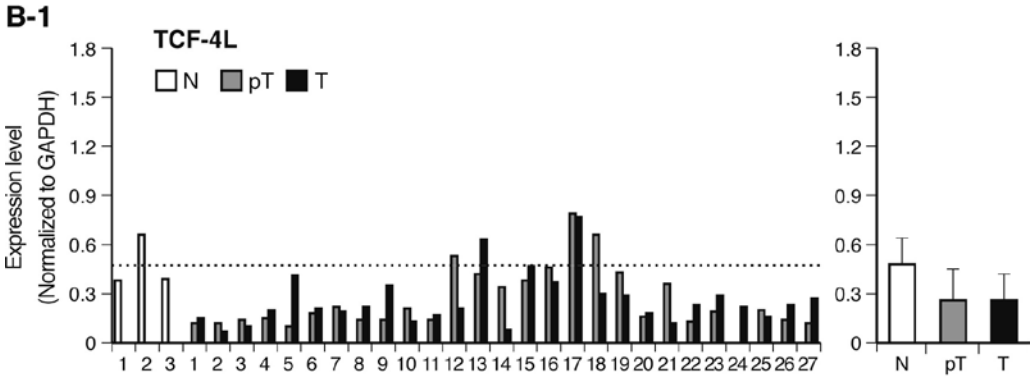
In contrast to TCF-4J, TCF-4K was hardly expressed in HCC tumor samples, and if expressed, none of them was higher than that of normal samples. In fact, the majority of them (59%, 16/27) were below the average expression of normal tissues (Figure 3.5C, left chart). However, its expression pattern in many of the peritumor tissues (48%, 13/27) fell within the average range of normal samples. Comparison of the mean values revealed efficient downregulation of TCF-4K in HCC tumors (Figure 3.5C, right chart). In addition, TCF-4K was the isoform that had the most structural resemblance to TCF-3.

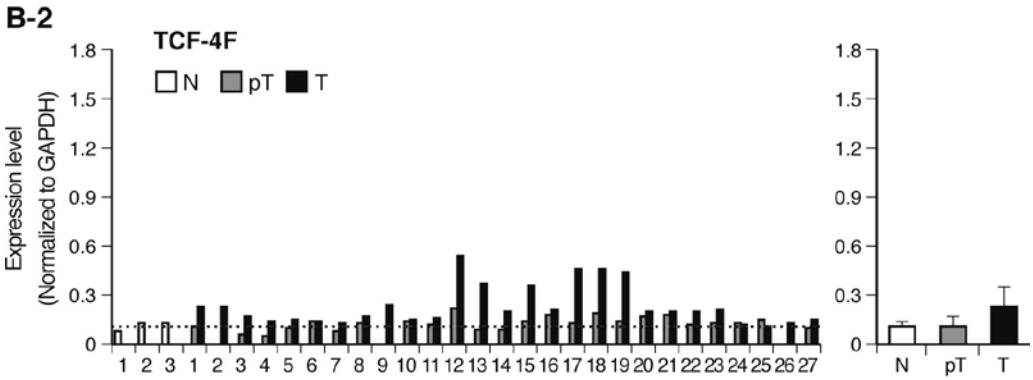
3.3.3.4. Expression of isoforms TCF-4L, F, I, M, E, and X

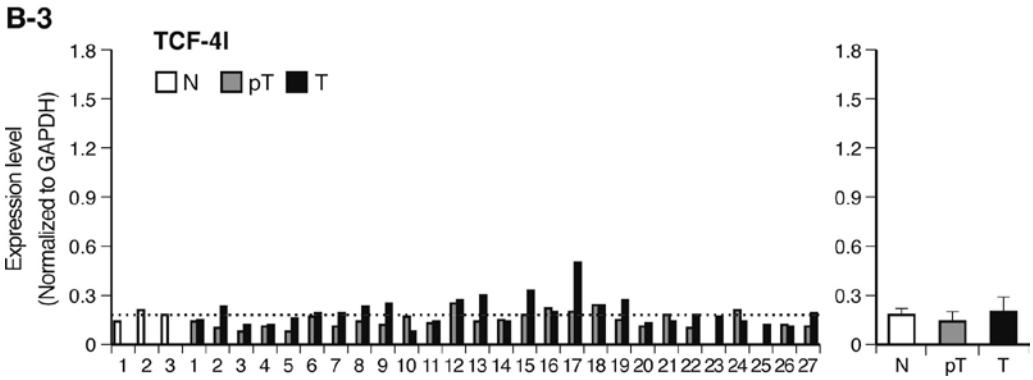
The isoform TCF-4L was generated by RT-PCR with the primer pair P2F/P3R (Figure 3.2A and Figure 3.6). It was a unique isotype in that it contained both LVPQ and SxxSS motifs, the two CtBP-biding sites, and exon 4, all of which implicated in possible transcriptional repression of Wnt target genes. As shown in Figure 3.6B, expression of TCF-4L RNA was effectively downregulated in many peritumor and tumor samples compared to normal tissues.

The primer pairs that detected and amplified the isoforms TCF-4F, I, M, E, and X are shown in Figure 3.2A. Expression of these isoforms not only was very low but also similar between normal, peritumor and tumor tissues (Figure 3.6). Interestingly, tumor samples numbered 12 – 19 were obtained from poorly differentiated HCCs (Table 3.3) and they expressed TCF-4L, F, I, M, and E at higher levels than the rest of the samples.


A
MM
L
12
13
14
15
16
17
18
N
T
pT
1
2
3
P2F/P3R
F
P1F/P2R
I
P2F/P2R
M
P2F/P4R
E/X
E
X
P1F/P4R
GAPDH


B-1
TCF-4L
N
pT
T
Expression level
(Normalized to GAPDH)


B-2
TCF-4F
N
pT
T
Expression level
(Normalized to GAPDH)


B-3
TCF-4I
N
pT
T
Expression level
(Normalized to GAPDH)

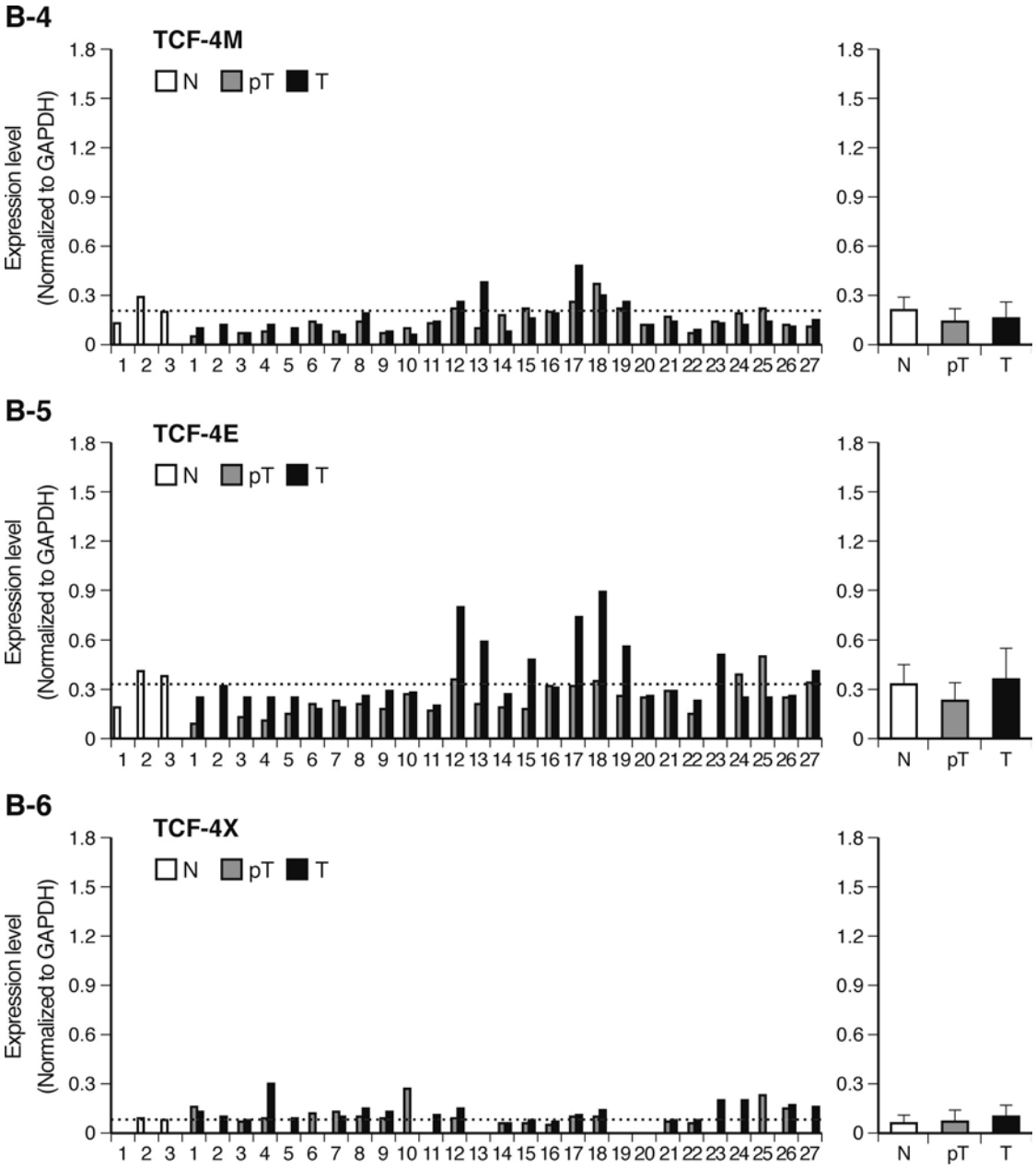


Figure 3.6 Expression of TCF-4L, F, I, M, E, and X in HCC tissues by semi-qPCR

RNA expression levels of TCF-4L (*B-1*), TCF-4F (*B-2*), TCF-4I (*B-3*), TCF-4M (*B-4*), TCF-4E (*B-5*), and TCF-4X (*B-6*) measured by semi-qPCR in twenty-seven HCC tumor and matching peritumor tissues and in three normal liver samples. *A*. Representative pictures for running above isoforms in 2.0% agarose gels. The primer sets used to amplify are indicated at the left side. The respective positive controls for each isoform are indicated above each gel picture. Note that TCF-4E and X were generated by single primer pair, P1F/P4R.

3.4. DISCUSSION

In order to confirm the expression of the newly identified TCF-4 isoforms in HCC and determine their mRNA expression levels, one would perform qPCR. However, it was impossible to make unique RT-PCR primer pair to every splice variant as each isoform was extensively spliced in both 5' and 3' ends as well as in the center. Therefore, isoforms were categorized into several groups according to their exon characteristics and the expression of the isoforms as a group was analyzed. Because the qPCR results were impractical at individual isoforms, this problem was overcome by a less sensitive semi-qPCR method. Nevertheless, both qPCR and semi-qPCR findings were quite consistent.

The most noticeable expression pattern was that of TCF-4C. Absence of this isoform in normal and peritumor tissues (Figure 3.3D) strongly suggested that TCF-4C might have an undesirable effect in normal adult liver tissues and hence it was efficiently downregulated. Conversely, overexpression of TCF-4C in more carcinogenic Huh7 and FOCUS cell lines (Figure 3.2C) as well as in the HCC tumor tissues implied a major role in progression of HCC, but not in hepatocarcinogenesis since it was not detected in less carcinogenic HepG2 and Hep3B cells and peritumoral regions, of which the latter was mostly cirrhotic or affected by chronic hepatitis (Table 3.3). Moreover, TCF-4C not only overexpressed in HCC tumors compared to their matching peritumor and normal tissues, but also was the highest expressed isotype (compare Figures 3.3 – 3.6, all graphs are at same scale), indicating it to be the key downstream effector of Wnt signaling in HCC. It is worth noting that the expression of such extreme pattern was observed in an isoform that lacked the exon 4, LVPQ, SxxSS, and CtBP-binding sequences. Therefore, TCF-4C could be a strong transcription activator isotype, or a target gene activation-specific

isoform. On the other hand, TCF-4D, which contained exon 4, otherwise structurally identical to TCF-4C, was not expressed in most of the tumor tissues suggesting a repressive role for exon 4. However, when TCF-4D was expressed in the remaining tumor tissues, its expression levels were as high as that of TCF-4C.

Another striking observation was the downregulation of isoforms with SxxSS motif in HCC cells and tumor tissues as evidenced by the expression patterns of isoform pairs, TCF-4B and A, TCF-4G and H, and TCF-4J and K. First, in the qPCR approach, the isoforms TCF-4A, H, and K were detected as the +Exons 7L & 9L group, which had a low expression level in both HCC cell lines and tumor tissues (Figure 3.1 C and D). Second, these three isoforms along with TCF-4X were expressed at such low levels the semi-qPCR could not detect them in HCC cell lines (Figure 3.2). On the other hand, their counterparts that lacked the SxxSS motif (TCF-4B, G, and J isoforms) were expressed abundantly. Third, expression of TCF-4A, H, and K was downregulated in most HCC tumor samples compared to their matched peritumor and normal tissues when analyzed by semi-qPCR (Figures 3.3 – 3.5). Similar to cell line data, the TCF-4B, G and J were highly expressed in HCC tumors. Therefore, these results strongly confirmed that the SxxSS motif was indeed implicated in transcriptional repression and thus their expression was avoided in HCC. However, the claimed repressive function of LVPQ motif could not be verified, as TCF-4B, G, and J were expressed at high levels despite its presence. Interestingly, the isoforms TCF-4B and G were both expressed highly and uniformly in all types of tissues (Figure 3.3C and 3.4C), suggesting that these splice variants served as common effectors of Wnt signaling in different conditions. TCF-4B could be the default

isoform since it was one of the first two TCF-4 isoforms identified [15], and the most abundantly expressed splice variant in a broad range of tissues.

The significance of the SxxSS motif was further elucidated in the differential expression pattern between TCF-4J and K (Figure 3.2C and 3.5). Although both isoforms contained the long C-terminal tails thus CtBP-binding sites, the SxxSS motif highly influenced their expression profile. Evidently, its presence was very disadvantageous in HCC tumors as TCF-4K was efficiently downregulated. Thus it could be the transcription repressive isoform, or target gene-inhibition-specific splice variant. In addition, TCF-4K is the isoform that structurally resembles to TCF-3 the most. Conversely, the absence of SxxSS in TCF-4J was markedly beneficial as this isoform was highly upregulated in HCC tissues, suggesting that TCF-4J could mediate the transcription of Wnt target genes along with TCF-4C as a transcription activator type. However, since TCF-4J was also moderately expressed in peritumoral tissues, it could be involved in the process of hepatocarcinogenesis as well as in HCC progression. Incidentally, it is now obvious that the qPCR result of the ΔExon 4 group showed very high expression levels as this group included the highly abundant TCF-4B, C, and J (Figure 3.1B-1, C and D). Interestingly, the same strong effect between the absence of SxxSS motif and the expression of TCF-4J was not observed with TCF-4B and G, suggesting an involvement of more intricate molecular mechanism depending on other sequences.

Finally, the expression of TCF-4F and I isoforms was uniformly low in all cell lines and across different tissues as determined by both qPCR (as +Exon 13 group in Figure 3.1) and semi-qPCR (Figures 3.2C and 3.6). Isoform TCF-4M was also expressed at low levels evenly across all tissues, indicating that these splicing variants could serve

as potential neutral isoforms, whose expression was controlled consistently at low levels in various cell conditions. The lowest expressed isotype was TCF-4X, confirming that exon 5A was rarely included in the transcript. Interestingly, the isoform TCF-4L that contained all motifs implicated in repression (exon 4, LVPQ, SxxSS, and CtBP-binding sites) was expressed higher in normal tissues compared to most tumor and peritumor samples although overall expression across all tissue types was low. Expression pattern of TCF-4E, which was one of the two original isoforms, was inconsistent between HCC cell lines and tissues as it was expressed at high levels in HCC cells but much lower in tissues.

3.5. METHODS

3.5.1. Human HCC tissues

Twenty-seven pairs of liver tumor and matched peritumor tissues were obtained from South Korean patients (Table 3.3). While the tumor samples were obtained from the source of the HCC section, the peritumor tissues were dissected from the liver area surrounding the HCC. Twenty-one patients were male, and six were female. The mean age of the patients was 50 years ranging between 22 and 69. Except one patient, all of them had HBV-related chronic liver diseases. Three normal liver tissues were acquired from individuals who had undergone hepatic resection for single metastasis of colorectal cancer to the liver and served as controls.

3.5.2. Plasmids, transfections, and RNA extraction

Fourteen of the newly identified TCF-4 isoforms (TCF-4A through M, and X) were cloned into mammalian expression vector pcDNA3.1/myc-His. As positive control for the semi-qPCR, 1.0 μg of each isoform plasmid was transiently transfected into Huh7 cells plated in 6-well plates by using a *Trans*IT-LT1 transfection reagent (Mirus) according to the manufacturer's instruction. Total cellular RNA was extracted from the transfected Huh7 cells, four HCC cell lines as wells as from normal liver and all paired HCC tissues by using TRIzol reagent (Invitrogen, Carlsbad, CA). The quality of the RNA samples was determined by ethidium bromide-stained agarose gel electrophoresis verifying intact 18S and 28S rRNA bands. To obtain first-strand cDNA, 2 μg of total RNA was treated with DNase-I, RNase inhibitor, and reverse transcribed with oligo-dT primer and AMV reverse transcriptase (Roche Diagnostics, Indianapolis, IN).

3.5.3. Quantitative real time RT-PCR (qPCR)

To confirm the expression of different isoforms of TCF-4 in HCC, all isoforms were categorized into different groups according to their exon characteristics, and six set of primers were designed as shown in Figure 3.1B and Table 3.1. The copy number of mRNAs of each group was quantified by measuring the C_t value, which was normalized by 18S rRNA value after comparing their standard curves. The standards for every group and 18S rRNA were prepared as 10 fold dilutions from each PCR product.

3.5.4. Semi-quantitative RT-PCR (semi-qPCR)

Expression of individual isoforms in both HCC cell lines and tissues was carried out by semi-qPCR with eight different sets of splicing site-specific primers (P1F/P1R, P2F/P1R, P1F/P2R, P2F/P2R, P1F/P3R, P2F/P3R, P1F/P4R, P2F/P4R) as shown in Figure 3.2A and Table 3.2. The PCR products generated by P1F/P1R, P2F/P1R, P1F/P3R primer sets, which produced multiple isoforms, were separated on 3.5% high resolution NuSieve 3:1 agarose gels (Lonza Rockland, Inc., Rockland, ME). This analytic technique is capable of resolving DNA fragments from 10 to 1500 base pairs. The rest of the isoforms were isolated on 2% agarose gels. PCR products were visualized by FluorChem 8000 gel imaging system (Alpha Innotech/Cell Biosciences, Inc., Santa Clara, CA), and the band intensity of each PCR product was analyzed and quantitated by densitometry. The expression of glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was used as an endogenous control. For quantitation, TCF-4 expression of each sample was normalized to GAPDH mRNA levels. The 100-base pair molecular marker (Marker XIV from Roche) served as control for band intensity between different agarose gels.

3.6. REFERENCES

1. Brunner E, Peter O, Schweizer L and Basler K (1997) pangolin encodes a Lef-1 homologue that acts downstream of Armadillo to transduce the Wingless signal in Drosophila. *Nature*, 385: 829-833.
2. Behrens J, von Kries JP, Kuhl M, Bruhn L, Wedlich D, Grosschedl R and Birchmeier W (1996) Functional interaction of beta-catenin with the transcription factor LEF-1. *Nature*, 382: 638-642.
3. Korinek V, Barker N, Willert K, Molenaar M, Roose J, Wagenaar G, Markman M, Lamers W, Destree O and Clevers H (1998) Two members of the Tcf family implicated in Wnt/beta-catenin signaling during embryogenesis in the mouse. *Mol Cell Biol*, 18: 1248-1256.
4. van de Wetering M, Cavallo R, Dooijes D, van Beest M, van Es J, Loureiro J, Ypma A, Hursh D, Jones T, Bejsovec A, Peifer M, Mortin M and Clevers H (1997) Armadillo coactivates transcription driven by the product of the Drosophila segment polarity gene dTCF. *Cell*, 88: 789-799.
5. Waterman ML (2004) Lymphoid enhancer factor/T cell factor expression in colorectal cancer. *Cancer Metastasis Rev*, 23: 41-52.
6. Cavallo RA, Cox RT, Moline MM, Roose J, Polevoy GA, Clevers H, Peifer M and Bejsovec A (1998) Drosophila Tcf and Groucho interact to repress Wingless signalling activity. *Nature*, 395: 604-608.
7. Roose J, Molenaar M, Peterson J, Hurenkamp J, Brantjes H, Moerer P, van de Wetering M, Destree O and Clevers H (1998) The Xenopus Wnt effector XTcf-3 interacts with Groucho-related transcriptional repressors. *Nature*, 395: 608-612.

8. Brantjes H, Barker N, van Es J and Clevers H (2002) TCF: Lady Justice casting the final verdict on the outcome of Wnt signalling. *Biol Chem*, 383: 255-261.

9. Pukrop T, Gradl D, Henningfeld KA, Knochel W, Wedlich D and Kuhl M (2001) Identification of two regulatory elements within the high mobility group box transcription factor XTCF-4. *J Biol Chem*, 276: 8968-8978.

10. Brannon M, Brown JD, Bates R, Kimelman D and Moon RT (1999) XCtBP is a XTcf-3 co-repressor with roles throughout Xenopus development. *Development*, 126: 3159-3170.

11. Standley HJ, Destree O, Kofron M, Wylie C and Heasman J (2006) Maternal XTcf1 and XTcf4 have distinct roles in regulating Wnt target genes. *Dev Biol*, 289: 318-328.

12. Houston DW, Kofron M, Resnik E, Langland R, Destree O, Wylie C and Heasman J (2002) Repression of organizer genes in dorsal and ventral Xenopus cells mediated by maternal XTcf3. *Development*, 129: 4015-4025.

13. Oosterwegel M, van de Wetering M, Timmerman J, Kruisbeek A, Destree O, Meijlink F and Clevers H (1993) Differential expression of the HMG box factors TCF-1 and LEF-1 during murine embryogenesis. *Development*, 118: 439-448.

14. Merrill BJ, Pasolli HA, Polak L, Rendl M, Garcia-Garcia MJ, Anderson KV and Fuchs E (2004) Tcf3: a transcriptional regulator of axis induction in the early embryo. *Development*, 131: 263-274.

15. Korinek V, Barker N, Morin PJ, van Wichen D, de Weger R, Kinzler KW, Vogelstein B and Clevers H (1997) Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. *Science*, 275: 1784-1787.

CHAPTER 4

4. CHARACTERIZATION OF TCF-4 ISOFORMS IN HCC CELLS

4.1. ABSTRACT

The TCF/LEF transcription factors are bipartite effectors of the Wnt signaling pathway. They repress gene expression by interacting with co-repressors TLE and CtBPs in the absence of Wnt ligands. Upon Wnt induction, the stabilized and accumulated β-catenin enters the nucleus, displaces the TLE, and converts the TCF/LEFs into transcription activators. However, this mechanism does not seem to apply for TCF-1, LEF-1, and TCF-3, leaving only TCF-4 to be such a molecular switch. TCF-3 is known with its repressive effect, and its repression is relieved by replacing it with other transcription factors. The suppressive activities of TCF-1 and LEF-1 are provided by their respective dnTCF-1 and dnLEF-1, and their repressive functions are eliminated by again replacing them with full length counterparts, indicating that a repression-specific TCF/LEF is expressed in the absence of Wnt signaling. Therefore, I further hypothesize that the potential repression- and activation-specific isoforms of TCF-4 may also be expressed specifically in the absence and presence of Wnt signals, respectively. Here, I investigate the functional characteristics of the newly identified TCF-4 isoforms by analyzing their transcriptional activity, effects on cell proliferation, migration, and anchorage-independent cell growth. The importance of the repressive effect of SxxSS motif is further validated.

4.2. INTRODUCTION

The transcription factors of the TCF/LEF family are the ultimate effectors of the Wnt signaling pathway. Developmental studies in frogs and mice have revealed that each member of the family is utilized in a particular way during embryogenesis: while TCF-1, LEF-1, and TCF-4 are used for both transcriptional activation and repression of Wnt target genes, TCF-3 is solely employed to inhibit the expression of those genes [1-4]. During active Wnt signaling, β-catenin is thought to convert TCF/LEF proteins from transcriptional repressors to activators by displacing the co-repressor TLE and recruiting other complexes necessary for transcription initiation [5,6]. However, TCF-3 has not been reported to activate transcription despite the presence of BCBD, suggesting that TCF-3 is not converted into a transcriptional activator. Instead, it is replaced by other transcription factors of the HMG superfamily, including TCF-1 and LEF-1, which are mostly required for transcriptional activation [1,2]. On the other hand, TCF-4 is equally employed as both transcription repressor owing to TLE and other co-repressors in the absence of Wnt signals and as transcription activator due to subsequent interaction with β-catenin upon Wnt stimulation.

Identification of several TCF-4 isoforms with varying combinations of alternative sequences and their differential expression pattern in HCC let me further hypothesize that the potential repression- and activation-specific isoforms of TCF-4 not only structurally resemble with TCF-3 and TCF-1/LEF-1, respectively, but may also function like TCF-3 and TCF-1/LEF-1. In other words, the repressive isoforms of TCF-4 may be upregulated in the absence of Wnt signals to efficiently suppress the expression of target genes while the activator isotypes may be expressed to oppose the function of the former when Wnt

signaling is initiated. Therefore, the activator isotypes may replace the repressive isoforms from the promoters upon Wnt stimulation, instead of a same TCF-4 molecule being converted from transcriptional repressor to activator by β-catenin.

In this chapter, I characterize the newly identified TCF-4 isoforms by *in vitro* functional assays to examine if they had different transcriptional activity and if their role affected the overall cellular behavior in the context of HCC. One of the valuable tools for studying transcriptional activity of TCF/LEFs is the optimized Wnt/β-catenin reporter plasmids called TOPFlash and FOPFlash [7]. Originally designed by the Hans Clevers lab [8], the TOPFlash plasmid contains two sets of three optimal TCF/LEF-binding sites (ATCAAAGG) upstream of a minimal promoter and a *Luciferase* open reading frame. The FOPFlash plasmid is used as negative control as its TCF/LEF-binding sites are mutated (GCCAAAGG). Upon Wnt stimulation or overexpression of β-catenin, TCF/β-catenin complexes assemble on TOPFlash, but not on FOPFlash, and the expression of the reporter gene is used to measure the activity of TCF/LEF transcription factors.

Another helpful tool is the availability of an HA-tagged dnTCF-4E construct, which was also created by the same lab [9]. Similar constructs without the BCBD have been instrumental in placing TCF/LEF factors within Wnt pathway and the discovery of natural dnTCF-1 and dnLEF-1 has been very useful in exploring their activity [10-12]. Because dnTCF-4E cannot bind to β-catenin, its transcriptional activity measured by TOPFlash is extremely low and its overexpression opposes Wnt signals by repressing the transcription of target genes, such as c-Myc and Cyclin D1. Therefore, dnTCF-4E would be useful in testing my hypothesis by comparing its functional outcomes with those of 14 isoforms of TCF-4.

4.3. RESULTS

4.3.1. Expression of TCF-4 isoforms by WB

In order to characterize the TCF-4 isoforms and study their functional properties, each isoform was subcloned into mammalian expression vectors pcDNA3.1/Myc-His and pcDNA3.1(-)IRES-EGFP. EV and 14 TCF-4 isoforms cloned in the pcDNA3.1/Myc-His plasmid were co-transfected with a β-catenin expression vector into HEK293 and Huh7 cells, and expression of each isotype was confirmed by WB analysis with anti-TCF-4 and anti-Myc tag antibodies (Figure 4.1). All isoforms were expressed at their predicted protein sizes and the expression levels were comparable (indicated by single asterisks in Figure 4.1A and B). The higher bands were probably generated as a result of TCF-4 SUMOylation, an addition of ~10 kDa SUMO protein modifiers (Section 1.2.3.5.3, p. 65). In addition, the long isoforms showed multiple lower-sized bands. These bands might represent possible cleavage products or expression of other short TCF-4 isotypes since the TCF-1 and TCF-4 transcription factors with CRARF and/or RKKKCIRY motif (long C-tail) could bind to promoters of all *TCF* and *LEF* genes (Section 1.2.3.2.4, p. 55).

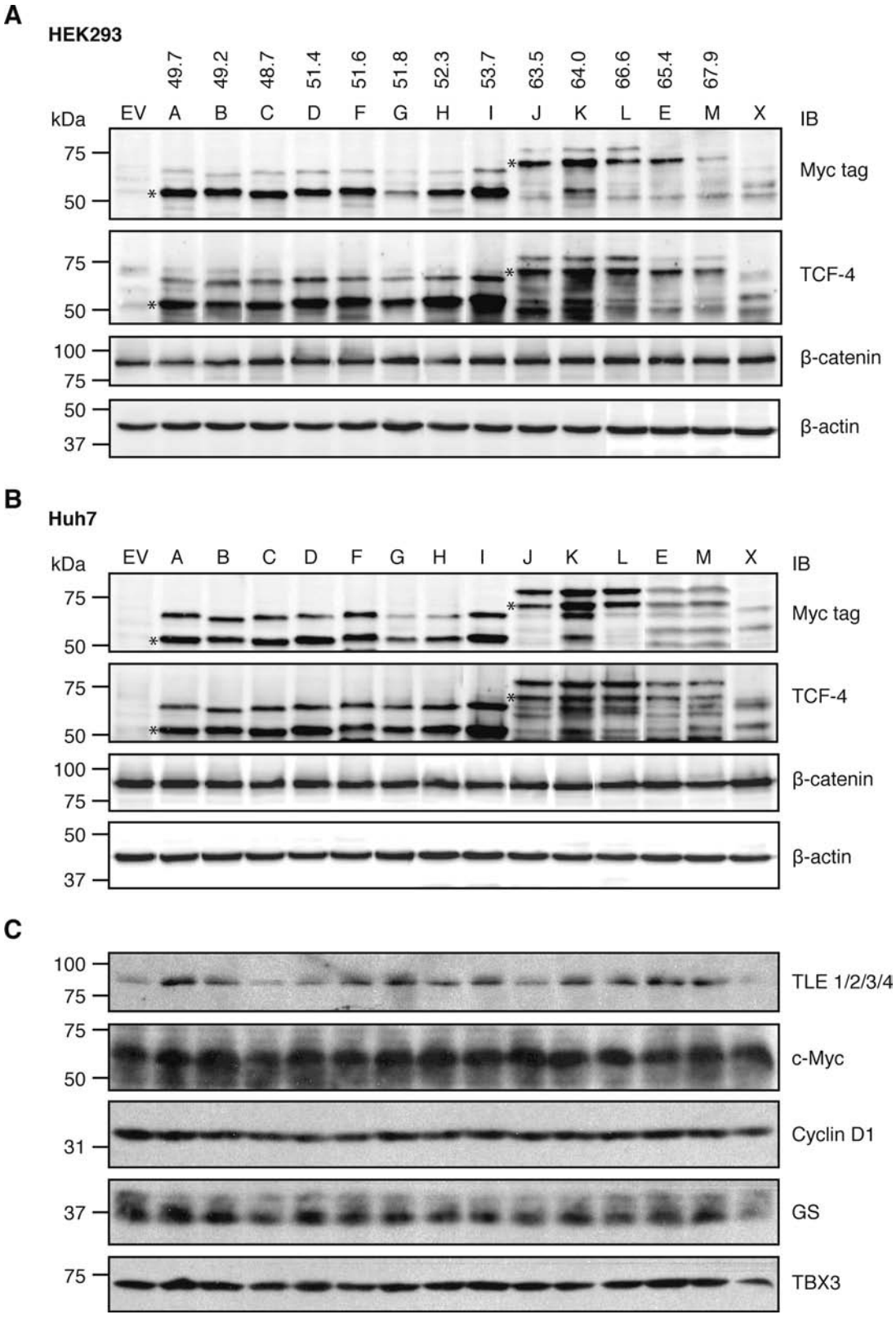

A
HEK293
49.7
49.2
48.7
51.4
51.6
51.8
52.3
53.7
63.5
64.0
66.6
65.4
67.9
kDa
EV
A
B
C
D
F
G
H
I
J
K
L
E
M
X
IB
75
50
Myc tag
75
50
TCF-4
100
75
β-catenin
50
37
β-actin
B
Huh7
kDa
EV
A
B
C
D
F
G
H
I
J
K
L
E
M
X
IB
75
50
Myc tag
75
50
TCF-4
100
75
β-catenin
50
37
β-actin
C
100
75
TLE 1/2/3/4
75
50
c-Myc
31
Cyclin D1
37
GS
75
TBX3

Figure 4.1. Ectopic expression of TCF-4 isoforms in HEK293 and Huh7 cells

WB analysis of TCF-4 isoforms, their binding partners, and known target genes of Wnt signaling. The whole cell lysates were prepared at 48 hours post-transfection of each isoform-containing vector and β-catenin plasmid into HEK293 and Huh7 cell lines. Immunoblot results from HEK293 (*A*) and Huh7 (*B*) cells were shown. TCF-4 isoforms containing the Myc tag at the C-terminus were detected with both anti-TCF-4 and anti-Myc tag antibodies (*). The predicted polypeptide sizes of TCF-4 proteins were indicated above each isoform in kDa. Expression of β-actin served as the protein loading control. *C*. WB results of TLE/Gro, c-Myc, Cyclin D1, GS, and TBX3 in Huh7 cells upon overexpression of TCF-4 isoforms and β-catenin. The protein sizes were shown in parenthesis.

The ectopic expression of β-catenin was uniform in all samples tested, including empty vector transfected cells, indicating that variation in TCF-4 isoform levels did not influence β-catenin (Figure 4.1A and B). To examine the functional effects of overexpression of each TCF-4 isoform, endogenous protein levels of TLE and several Wnt targets were assessed in Huh7 cells by WB analysis (Figure 4.1C). Expression level of co-repressor TLE, which was detected by anti-TLE 1 – 4 antibody, was overall low possibly due to overexpression of β-catenin. Interestingly, expression of TLE was even lower in TCF-4C, D, J and X isoforms as their band intensity was comparable to that of EV control, suggesting that TLE was further downregulated in the presence of these splice variants. However, there was no significant difference in expression of Wnt target genes, c-Myc, Cyclin D1, GS, and TBX3 compared to the levels in EV.

4.3.2. Transcriptional activity of each TCF-4 isoform

Once expression of TCF-4 isoforms at their expected protein sizes was confirmed, their transcriptional activity was assessed in HEK293 and Huh7 cells by using TOPFlash

(TOP) and FOPFlash (FOP) reporter system. To exclude the effect of different expression levels of TCF-4 isotypes in the pcDNA3.1(-)IRES-EGFP and β-catenin on transcriptional activity, transfection was carried out in duplicate such that both WB analysis and TCF-4 activity assay could be performed at the same time. As shown in Figure 4.2A, the immunoblots detected by anti-TCF-4 and anti-HA tag antibodies produced similar expression patterns as in Figure 4.1A and B. The dnTCF-4E isoform was also efficiently expressed in its expected protein size of 53.7 kDa (double asterisks). In addition, expression of TCF-4 isoforms was further confirmed by positive expression of EGFP directly from the transcriptional activity assay plates under the fluorescence microscope (Figure 4.2B). Finally, the transcriptional activity of TCF-4 isoforms in HEK293 and Huh7 cells was determined as shown in Figure 4.2C. As the N-terminal deletion mutant lacking the BCBD, dnTCF-4E served as a control for assessing a repressed TCF-4 transcriptional activity. On the other hand, the EV control exhibited the basal transcriptional activity because it represented the baseline activity delivered by the endogenous TCF-4 proteins. Comparing the transcriptional activities of each TCF-4 isoform with the basal activity demonstrated several interesting features:

1. Each TCF-4 isoform exhibited a consistent pattern of transcriptional activity compared to EV in both HEK293 and Huh7 cell lines.
2. TCF-4 activity was β-catenin dependent. Whether it was a dnTCF-4E, EV, or a specific isoform, the transcriptional activity always increased with β-catenin expression (black bars) as compared to samples without β-catenin transfections (white bars). This indicated that not only TCF-4 transcriptional activity was repressed in the absence of β-catenin, but also overexpression of β-catenin

efficiently simulated the activation of the canonical Wnt signaling and hence TCF-4.

3. Consistent with my hypothesis, transcriptional activity of each isoform was different from one another, and not all isotypes showed activity levels higher than that of EV control despite the presence of β-catenin. In this regard, many of the short isoforms demonstrated high transcriptional activity, while all of the long splicing variants displayed lower activity compared to EV confirming the effect of the co-repressor CtBP, which could only bind to the long isotypes. The isoforms TCF-4B, C, and G revealed the highest transcriptional activity indicating that they might have potent target gene activating properties. On the other hand, transcriptional activity of the isoforms TCF-4K, L, and M was as low as that of dnTCF-4E suggesting that these isotypes might represent the long-sought natural dominant negative forms of TCF-4 that could inhibit expression of Wnt target genes. These observations further supported my hypothesis that some TCF-4 isoforms might have a role as transcriptional repressors, and their role could even be extended in the presence of β-catenin if overexpressed.
4. Comparison of transcriptional activities and protein structure between TCF-4B and A revealed that the SxxSS motif was indeed exerting a repressive effect. The two isoforms differed from one another by only this motif, and the presence of SxxSS in TCF-4A resulted in repressed activity compared to that of TCF-4B. Therefore, these five amino acids, SFLSS, were responsible for the reduction of transcriptional activity in TCF-4A even in the presence of β-catenin expression further corroborating the fact that alternative splicing could generate repressive

TCF-4 isoforms. The same repression pattern of SxxSS could also be observed in other paired isoforms, TCF-4G *vs*. H and TCF-4J *vs*. K (also see Figure 5.2, p. 192). The transcriptional activity difference between TCF-4J and K was less obvious compared to the other paired isoforms due to overall repression caused by CtBP. Nevertheless, TCF-4J demonstrated the highest transcriptional activity among the long isotypes.

5. The effect of LVPQ motif was less clear and inconclusive as the transcriptional activity in both TCF-4B and C, which differed by this motif, was similarly high. Moreover, in another paired isoforms, expression of LVPQ in TCF-4G increased its transcriptional activity compared to TCF-4D indicating that the LVPQ motif might even have an activating function. Due to lack of more isoforms with SxxSS but without LVPQ, the functional effect of this motif could not be concluded.
6. Finally, sequences of exon 4 might provide another source of transcriptional repression as expression of this exon in TCF-4D resulted in diminished activity compared to TCF-4C. However, this observation could only be detected in short isoforms (TCF-4B *vs*. G and TCF-4A *vs*. H), but not in long isotypes (TCF-4K *vs*. L and TCF-4E *vs*. M), suggesting that sequences in this exon could contribute to a weak repression-specific effect in overall transcriptional activity.

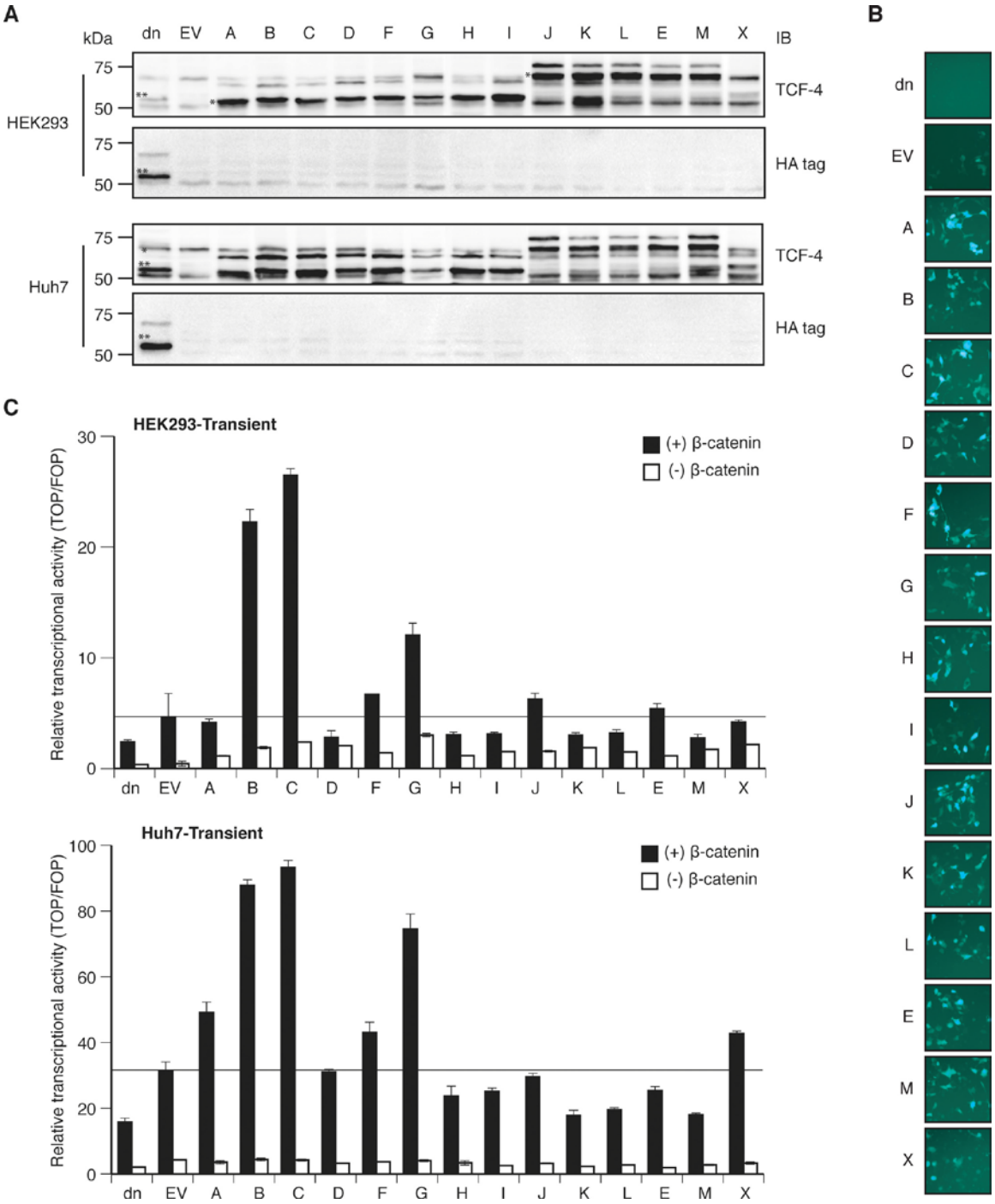


Figure 4.2. Transcriptional activity of TCF-4 isoforms

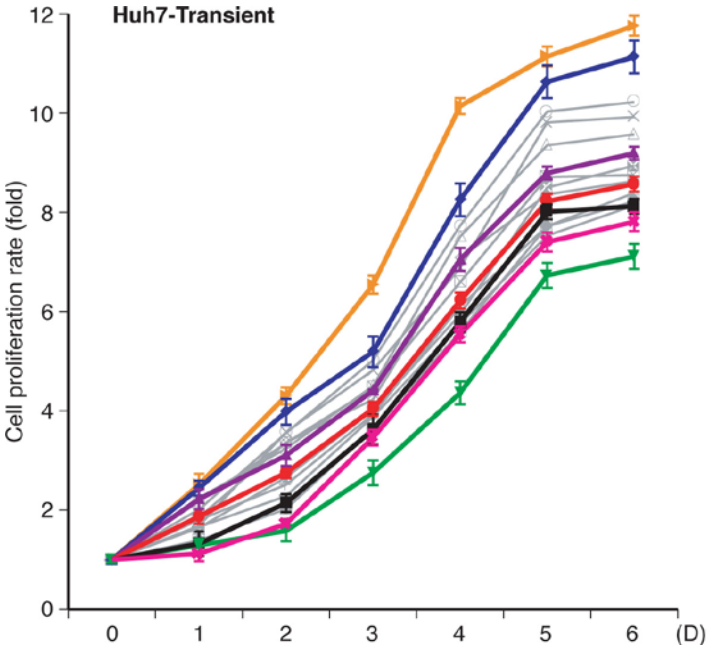
A. Confirmation of consistent expression of TCF-4 isoforms by WB analysis in the TCF activity assay setting. TCF-4 isoforms expressed from the pcDNA3.1(-)IRES-EGFP vectors were detected with anti-TCF-4 antibody (*). The dnTCF-4E that harbored the HA tag at the N-terminus was detected with both anti-TCF-4 and anti-HA tag antibodies (**). *B*. Expression of EGFP.

Efficient transfection of plasmids and expression of TCF-4 was verified from the assay plates owing to the IRES sequences in pcDNA3.1(-)IRES-EGFP plasmids before proceeding into the transcriptional assay. Note that dnTCF-4E did not express EGFP as it was constructed from a different vector that did not contain the *egfp* gene. *C*. Transcriptional activities of all TCF-4 isoforms after transient transfection into HEK293 and Huh7 cells. Activity was measured by using a luciferase reporter gene system. The relative luciferase activity was represented by the ratio of TOP value to FOP after normalizing the transfection efficiency with the β-galactosidase activity. The dnTCF-4E isoform served as control for assessing a repressed TCF-4 transcriptional activity. The horizontal line was drawn along the basal activity for comparison as EV control exhibited the endogenous activity. Results were presented as average rate ± SD from four independent experiments assayed in triplicates.

4.3.3. Effect of TCF-4 isoforms on cell proliferation

To further evaluate the functional properties of TCF-4, the effect of each isoform on HCC cell proliferation was examined. Expression plasmids containing TCF-4 splice variants were transiently co-transfected with β-catenin into Huh7 cells, and the amount of cells was measured for six days. EV and dnTCF-4E were used as controls, where dnTCF-4E would represent the rate of cell division at repressed Wnt signaling and EV would signify the normal cell proliferation of Huh7 cells. Cell proliferation profile of all isoforms is shown in Figure 4.3. In order to analyze the data, the slope of each proliferation curve was calculated. Since a horizontal line and a 45° rising line would have a slope of S = 0 and S = 1, respectively, and since a slope would illustrate the steepness of an incline, the higher the slope value, the faster the cell proliferation rate. In addition, the average cell doubling time for the first two days was also calculated to establish the time between two cell divisions for the whole population. The doubling time

was determined in hours and the lower the doubling time, the faster the cell proliferation rate. Huh7 cells transiently transfected with EV control proliferated with an average doubling time of 44.2 hours, which was consistent with previous reports [13] and with the routine maintenance schedule of growing Huh7 cells in the laboratory. As expected, the cells expressing dnTCF-4E reduced the replication rate while most of the other isoforms increased the rate of cell proliferation compared to EV control. The TCF-4C isoform that exhibited the highest transcriptional activity (Figure 3.2C) also resulted in the fastest cell proliferation with doubling time of 24.3 hours. The second highest cell proliferation rate was observed in samples with TCF-4J, which demonstrated the highest transcriptional activity among the long isoforms (though lower than EV). The cell proliferation rate of Huh7 cells transfected with TCF-4A and B was comparable with one another but higher than EV control. Interestingly, cell proliferation rate in Huh7 cells expressing TCF-4K isoform that displayed the lowest transcriptional activity was substantially reduced with doubling time of 73.4 hours, which was even below the rate of dnTCF-4E.



Symbol	Isoform	Slope	Doubling time (h)
	dn	1.34	62.4
	EV	1.43	44.2
	A	1.44	34.5
	B	1.56	32.8
	C	2.17	24.3
	D	1.82	26.7
	F	1.68	31.1
	G	1.52	34.5
	H	1.48	37.1
	I	1.51	29.3
	J	1.91	25.8
	K	1.12	73.4
	L	1.37	47.2
	E	1.74	26.5
	M	1.37	38.1
	X	1.33	41.8

Figure 4.3. Effect of TCF-4 isoforms on cell proliferation

Cell proliferation rate of Huh7 cells overexpressing TCF-4 isoforms after transient transfection. Growth rate curves were generated by normalizing each daily raw value with that of day 0, and data were presented as average rate ± SD from three independent experiments assayed in quadruplicates. The slope of each curve and the doubling time for every cell expressing various isoforms were calculated and shown next to the symbols in a chart. Only the isoforms that exhibited interesting and important characteristics were color coded for emphasis.

4.3.4. Selecting isoforms for stable expression

Since the SxxSS motif demonstrated the strongest effect in isotype expression, transcriptional activity and cell proliferation rates in HCC, the isoforms TCF-4A, B, J, and K were selected for further investigation. Huh7 and FOCUS stable cell lines overexpressing these four isotypes and the EV were generated to examine their biologic functions and long-term effects in HCC cells. Positive stable clones were confirmed by WB analysis (Figure 4.4A). The stably expressed isoforms appeared at their expected sizes. Additional short and long isotypes were revealed in the immunoblot probably due to constant expression of the stable isoforms, which might have led to activation of other splicing variants as their targets. Expression of TCF-4-binding partners (β-catenin and TLE) and Wnt target genes (c-Myc, Cyclin D1, GS, TBX3) were also examined. As previously, there was no significant difference in the expression levels of these proteins between different stable cells. Next, the transcriptional activity (Figure 4.4B) and cell proliferation (Figure 4.4C) of both Huh7 and FOCUS stable cells were analyzed. Consistent with the transient transfection results (compare Figure 4.4B with Figure 4.2C and Figure 4.4C with Figure 4.3), all isoforms exhibited similar pattern confirming a successful isolation of correct stable clones.

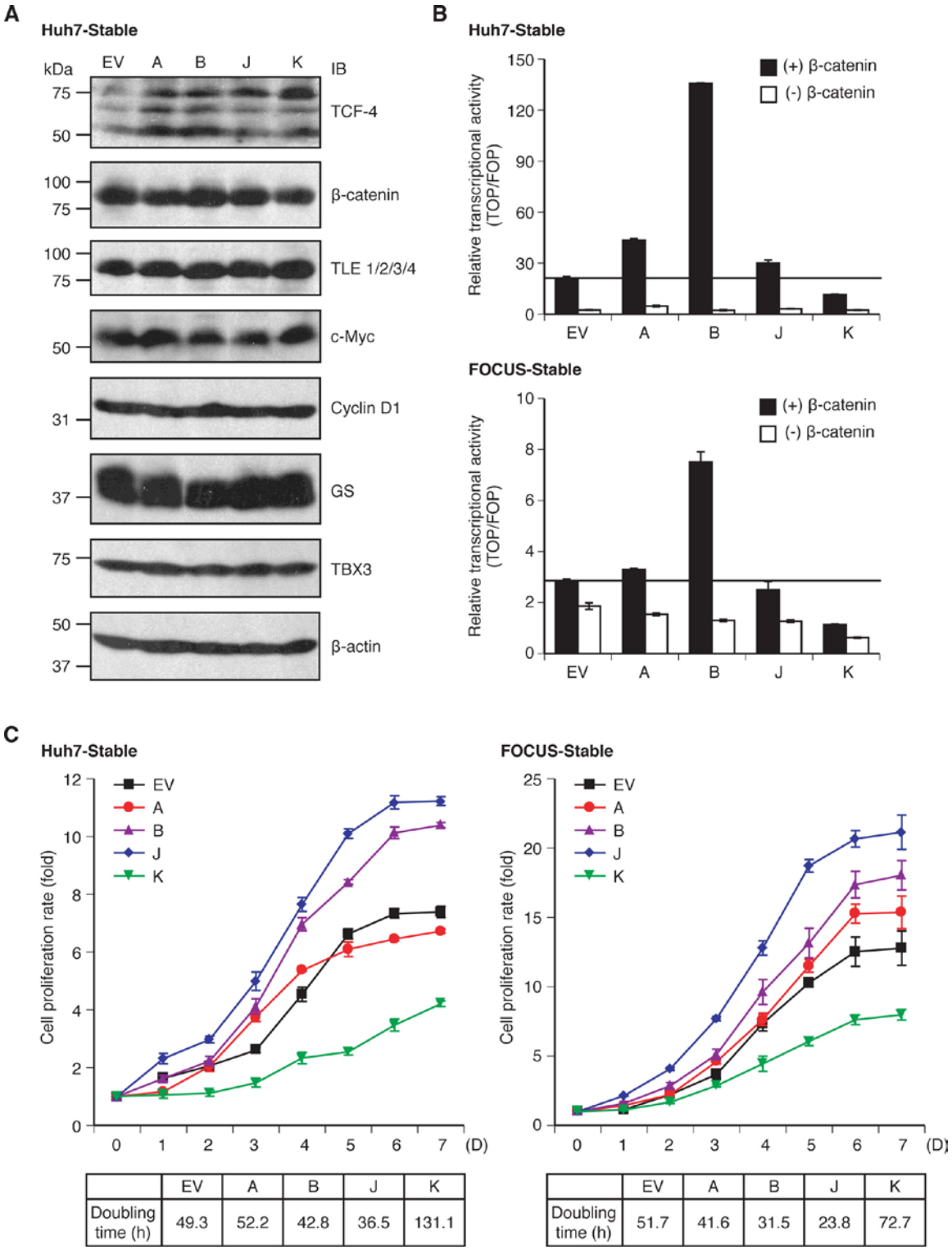


Figure 4.4. Confirmation of positive stable clones

Huh7 and FOCUS stable cell lines overexpressing EV, TCF-4A, B, J, and K were generated, and the positive clones were confirmed by WB analysis, transcriptional activity measurement, and cell

proliferation assay. *A*. Expression of these isotypes and other related proteins in Huh7 stable cells tested by immunoblotting. *B*. Transcriptional activity of isoforms stably expressed in Huh7 and FOCUS cells. Please note that the activity levels and the overall pattern are similar to their transient transfection results. *C*. Cell proliferation of Huh7 and FOCUS stable cells. The individual cell doubling times were estimated from the first 48 hours. Similarly, the growth rate pattern was consistent with their transient transfection data.

4.3.5. Effect of TCF-4 isoforms on cell migration

Once the stable clones were obtained, the effect of isoforms TCF-4A, B, J, and K on cell migration was tested since the migratory potential of tumor cells *in vitro* correlated with their invasive properties *in vivo*. Migration of stable cells was examined by two different methods: scratch wound healing and transwell/Boyden chamber cell motility assays (Figure 4.5). While wound closure was measured every 8 hours over a period of two days, cell motility from upper to lower surface of a porous PET membrane was quantitated within 24 hours.

The overall cell migration rate was much faster in FOCUS stable cell lines than in Huh7 stable cells. In wound healing assay, for example, almost all of the wounds were closed within 48 hours in FOCUS, whereas wound closure was detected 40% – 80% in Huh7 cells. Consistent with its effect on cell proliferation, stable expression of TCF-4J lead both stable cell lines to migrate the fastest. Within 48 hours, the wound was healed almost 80% in Huh7 while it was fully closed in FOCUS compared to their respective cells carrying other isoforms (Figure 4.5A). Cell growth and division should not be of concern in wound closure results as the rates were consistent at every time point. In addition, the cell doubling time for each stable cell line was calculated during the cell

proliferation assay (Figure 4.4), and comparing the rates of that time point would still give the same results. For example, at time points of 36 hours in Huh7 and 24 hours in FOCUS, the doubling times of the fastest growing cell lines, the wound closure patterns were still consistent: stable cells expressing TCF-4J lead the wound closure percentages with 40% in Huh7 and more than 73% in FOCUS cells.

Interestingly, the rate of wound closure of stable cells expressing TCF-4K was quite different. It showed the second fastest migration in Huh7 cells with 36% wound closure at 24 hours and almost 80% within 48 hours, whereas its migration was the slowest among FOCUS stable cells with 51% closure at 24 hours and 80% at 48 hours. However, the individual percentages of wound closure between the two stable cell lines were similar. Therefore, the wound healing results of TCF-4K in FOCUS stable cells, but not in Huh7, was consistent with its slow cell proliferation rate and low transcriptional activity. An exact opposite effect was observed in TCF-4A isoform. While its stable expression resulted in the lowest rate of wound healing in Huh7 cells with only 16% closure at 24 hours and 40% within 48 hours, it was the highest (along with TCF-4J) among FOCUS cells with 73% closure at 24 hours and almost 100% within the next 24 hours. Similarly, stable expression of TCF-4B revealed slow wound healing in Huh7 cells with 19% closure at 24 hours and 48% within 48 hours, and higher rate of migration in FOCUS cells with 67% closure at 24 hours and almost 100% at 48 hours. The rate of wound healing in both cell lines stably expressing TCF-4B isoform was higher than the EV control.

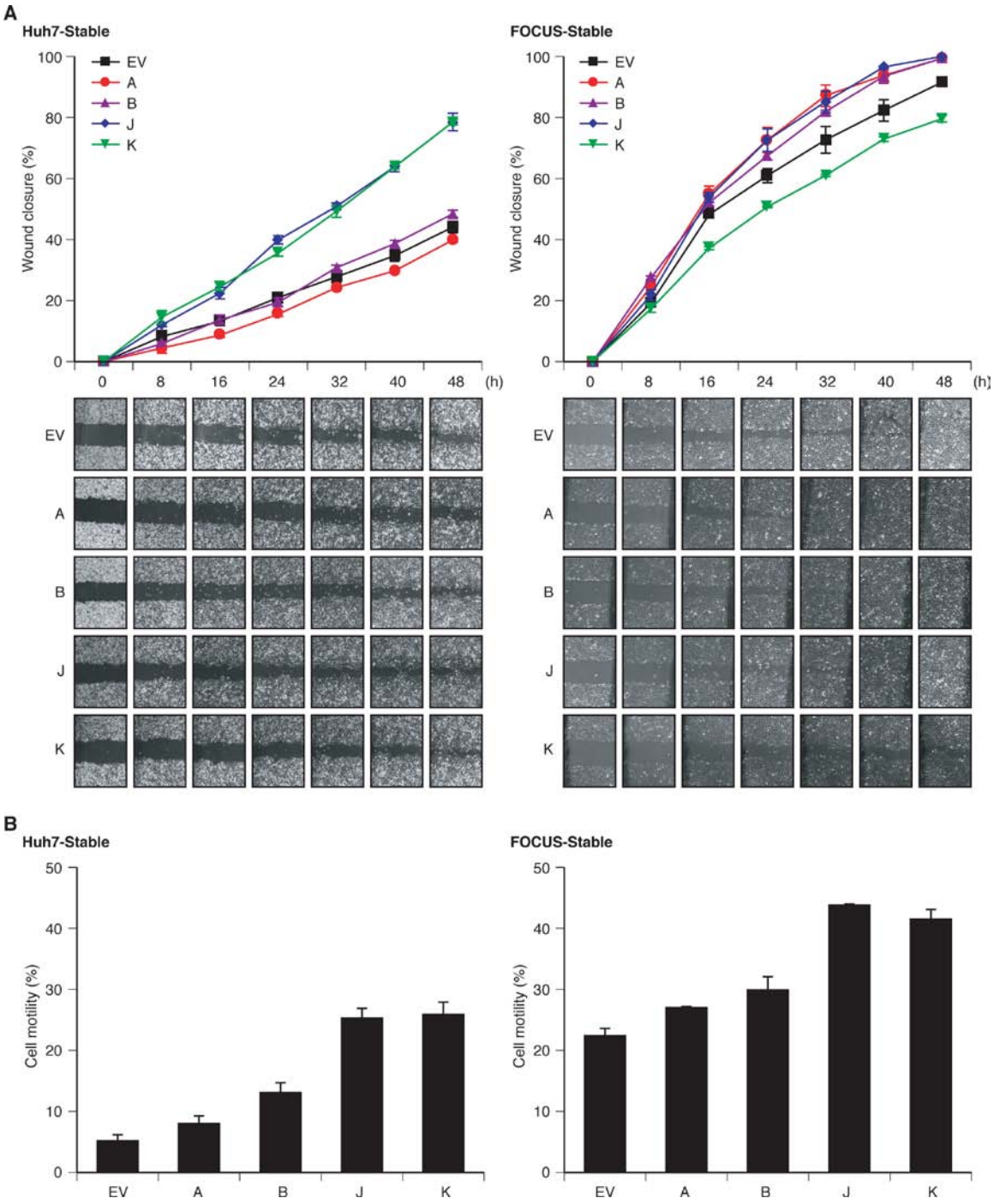


Figure 4.5. Effect of TCF-4 isoforms on cell migration

A. Scratch wound healing assay performed in Huh7 and FOCUS stable cells. Cell migration of stable cells expressing TCF-4A, B, J, and K was compared to EV control. The cells were photographed at the identical locations every eight hours as shown. The graph depicts a percent of wound closure plotted against time. The results were presented as average percentage of

wound closure ± SD from three independent experiments assayed in duplicates. *B*. Transwell/Boyden chamber cell motility assay results of both stable cell lines. Data were expressed as a percent of migrated cells to total number of cells. The assay was performed three times each in triplicates.

In transwell cell motility assay, continuous expression of TCF-4J again resulted in the highest rate of cell motility in both cell lines (Figure 4.5B). The overall result of Huh7 stable cells was consistent with their wound healing rates: TCF-4A and B caused lower cell motility while TCF-4J and K the highest. However, the cell motility results of FOCUS stable cell lines were more consistent with that of Huh7 rather than the rate of wound healing in FOCUS, suggesting that the two cell migration assays did not correlate in an already highly carcinogenic cell line. Nevertheless, the consensus results indicated that both long tailed isoforms, TCF-4J and K, might be responsible for high rate of cell migration. Except for the inconsistent results of TCF-4A and K between the two stable cell lines, it can be concluded that TCF-4J induced the highest rate of cell migration while TCF-4B moderately increased it compared to EV control.

4.3.6. Effect of TCF-4 isoforms on soft agar colony formation

Finally, to determine if TCF-4B, J, and K played a role in anchorage-independent cell growth, colony formation of the stable cells in soft agar was examined. Normally, most cell lines must be attached or anchored to a solid surface before they can divide, and they fail to grow when suspended in a viscous fluid, such as soft agar. However, when these cell lines are transformed, they become anchorage-independent and able to grow in soft agar. Therefore, soft agar colony formation assay is considered as the most stringent test for detecting malignant transformation of cells *in vitro* and positive results are viewed as

potential indications of *in vivo* carcinogenesis. As shown in Figure 4.6, the expression of the TCF-4J isoform significantly increased colony formation in both Huh7 and FOCUS stable cell lines. Notably, colonies generated from Huh7 cells expressing TCF-4J not only were increased in number, but also were larger in size than compared to EV control. In contrast, both the number and size of the colonies were reduced in Huh7 stable cells expressing TCF-4K, consistent with other opposing effects between TCF-4J and K. The number of colonies produced due to overexpression of TCF-4B was similar to that of EV. In FOCUS stable cells, TCF-4J was the only isoform that induce formation of colonies.

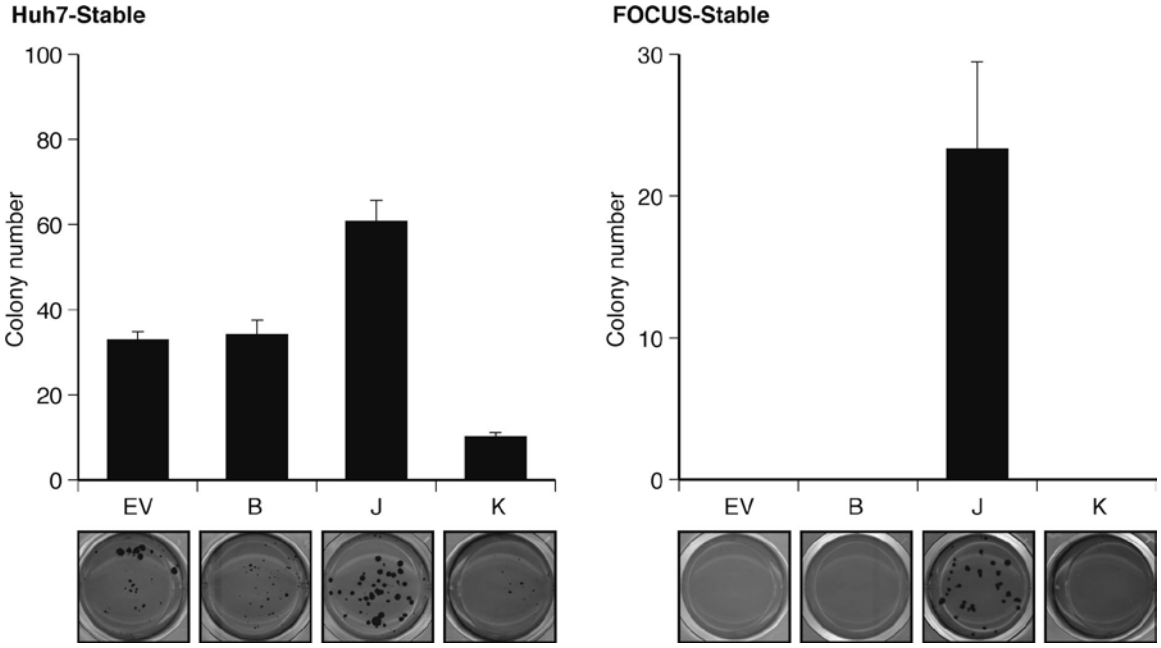


Figure 4.6. Effect of TCF-4 isoforms on anchorage-independent cell growth

Huh7 and FOCUS stable cells expressing EV, TCF-4B, J, and K were subjected to soft agar colony formation to evaluate anchorage-independent cell growth. The bar graphs show the number of colonies. The data are shown as average number of colonies ± SD from three independent experiments assayed in triplicates. Representative examples of soft agar plates illustrating transformed colonies derived from Huh7 and FOCUS stable cells are shown below the graphs.

4.4. DISCUSSION

To test my hypothesis that different TCF-4 isoforms might operate specifically either as transcription repressors or as activators instead of each isotype being converted between the two, several functional tests were performed on HEK293 and some HCC cell lines by overexpressing the newly identified TCF-4 isoforms. The activity assay measured by TOPFlash reporter system revealed the transcriptional characteristics of each isoform. If every isotype were converted from a repressor to an activator upon β-catenin binding, then each of the 14 isotypes would exhibit transcriptional activity higher than that of EV when transfected with β-catenin. However, it was not the case. Rather, it was consistent with my hypothesis because transcriptional activity of the isoforms not only was different from one another, but it ranged between as low as that of dnTCF-4E and up to six-fold of the baseline (EV). Many of the short splice variants displayed high activity levels while most of the long isoforms showed low transcriptional activity, confirming the effect of the co-repressor CtBP, which could only bind to the long C-tails. Similarly, other functional tests revealed divergent functional properties for different isoforms, and few of them significantly influenced the phenotype of HCC cells.

TCF-4C, which was highly expressed in carcinogenic HCC cells and exclusively in HCC tumor tissues, had the highest transcriptional activity in both HEK293 and Huh7 cell lines (Figure 4.2C). This observation substantiated my proposal that an isoform that lacked the exon 4, LVPQ, SxxSS, and the CtBP-binding sites, thus structurally resembled to TCF-1 and LEF-1, might be a transcription activation-specific isoform. Moreover, expression of TLE was effectively downregulated in TCF-4C-transfected Huh7 cells indicating that an interaction with such repressors was not favored (Figure 4.1C). Finally,

overexpression of TCF-4C highly accelerated cell proliferation and reduced the cell doubling time almost in half, further corroborating with my theory of TCF-4C being the major TCF-4 splice variant of Wnt signaling pathway during the course of HCC progression. However, the molecular mechanism in which TCF-4C exerted such notable characteristic could not be clarified, as the expression of well-known Wnt target genes did not change between samples with different isoform overexpressions. This can be explained by the fact that most Wnt target genes are tissue- and developmental/cancer stage-specific and that they are not always Wnt responsive [7]. Presumably, different TCF-4 isoforms might regulate diverse target genes and contribute to distinct cellular phenotypes.

In contrast, splice variants TCF-4K, L, and M showed the lowest transcriptional activities (Figure 4.2C). In fact, their activity was so low, it was comparable with that of dnTCF-4E suggesting that these isoforms might be the transcription repression-specific isotypes. In addition, their overexpression lead to the slowest cell proliferation rates (Figure 4.3). Consistent with my hypothesis, their domain structure was similar to that of TCF-3. TCF-4L contained exon 4, LVPQ, SxxSS, and the CtBP-binding sites, reflecting an exact opposite form of TCF-4C, while TCF-4K was bearing a structural resemblance to TCF-3 except for the RKKKCIRY motif, which was an obligate sequence for all long isoforms of TCF-4 (Figure 2.1A, p. 93). Speaking of other TCF/LEFs, TCF-1 and LEF-1 are able to exert repressive functions even in the context of active Wnt signaling although they are renowned with their gene activating roles. These occasional repressive activities of TCF-1 and LEF-1 are provided by their respective natural dominant negative variants, which lack the BCBD but are able to occupy the promoters of Wnt target genes [10,12].

However, such dominant negative form for TCF-4 has not been identified. Interestingly, because the repressive function of TCF-4K, L, and M was observed in the presence of β-catenin and their transcriptional activity was comparable to that of dnTCF-4E, these isotypes could represent the long-sought natural dominant negative forms of TCF-4. Evidently though, their repressive role could only be achieved if they were overexpressed to occupy all available WREs.

The importance of the SxxSS motif observed during the differential expression pattern of each isoform was further recognized by the functional assays. Three paired isoforms that differed by this motif, TCF-4B and A, TCF-G and H, and TCF-4J and K, exhibited opposite transcriptional activities. Namely, expression of the reporter gene in cells transfected with TCF-4A, H, and K was significantly lower than in samples overexpressing TCF-4B, G, and J, respectively (Figure 4.2C). Therefore, the reason why HCC cells and tumors were inclined to express isoforms that lacked SxxSS was that this motif negatively affected the transcription of Wnt target genes. Effect of SxxSS was also noticed in the cell proliferation assay. Expression of SxxSS in TCF-4A, H, and J was responsible for slower cell division than compared to their counterparts and resulted in higher doubling times (Figure 4.3). Although the isoforms TCF-4B and G revealed high transcriptional activity, their doubling times were not significantly lower than those of TCF-4A and H, respectively. Moreover, the effect of TCF-4B on cell migration and anchorage-independent cell growth was similar to that of EV (Figures 4.5 and 4.6). These data taken together with the high, uniform expression of TCF-4B and G in normal liver, HCC tumor, and peritumor suggested that they might participate in gene expression and cell homeostasis of normal hepatocytes rather than hepatocarcinogenesis. Especially for

TCF-4B, it could represent the wild type isoform of TCF-4 since its expression had been reported in a wide range of normal and carcinogenic tissues.

Surprisingly, the effect of SxxSS motif was most pronounced between the long isoforms TCF-4J and K. While presence of SxxSS in TCF-4K drastically lowered the cell proliferation rate even below that of dnTCF-4E with a doubling time of almost twice as EV, TCF-4J lead to the second highest replication rate just after TCF-4C. Interestingly, the transcriptional activity of TCF-4J was significantly lower than that of TCF-4C, which had a short C-terminal tail. This observation not only confirmed the repressive effect of CtBPs on transcriptional activity of TCF-4J, but also substantiated the possibility that distinct TCF-4 isoforms regulated diverse target genes and thus lead to different effects on HCC cell proliferation. Nevertheless, TCF-4J managed to show the highest activity among the long isoforms and demonstrated high cell proliferation rate suggesting that absence of SxxSS motif was more effective in producing aggressive phenotype than the repressive effects of CtBPs. Moreover, TCF-4J induced the highest cell migration rate in both assays (Figure 4.5), and triggered the formation of the largest and the highest number of colonies on soft agar (Figure 4.6). These consistent results together with its high expression in HCC cells, tumor and peritumor tissues suggested that TCF-4J was indeed the most carcinogenic isoform, and thus it was probably exploited for hepatic transformation and tumor progression. On the other hand, TCF-4K not only exhibited the lowest transcriptional activity and cell proliferation rate, but also effectively decreased the anchorage-independent colony formation both in number and size, implying that TCF-4K applied exactly the opposite effects of TCF-4J. However, TCF-4K illustrated an overall high cell migration rate. Although, TCF-4K lead to slowest cell migration due to

wound healing in FOCUS cells (Figure 4.5A, right), its overexpression triggered cell migration as fast as in TCF-4J in other set of experiments. This inconsistent behavior of TCF-4K could be explained as a mechanism of cell survival rather than invasiveness of cancerous cells since a wound formation resulted in an injury (wound healing) and the cells in the upper chamber competed for food (transwell cell motility). Interestingly, the repressive effect of SxxSS was detected between short isoforms TCF-4A and B. Except for wound healing in FOCUS cells, TCF-4A showed lower cell migration rate than TCF-4B, indicating that the effect of SxxSS was specific for certain conditions and that the long C-terminal tails of TCF-4J and K might selectively regulate expression of target genes involved in cell survival.

Overall, the functional tests on all 14 TCF-4 splice variants revealed the presence of target gene activation- and repression-specific isoforms. While TCF-4C showed the highest transcriptional activity and cell proliferation rate, TCF-4K exhibited the opposite results. Moreover, among the structural motifs exon 4, LVPQ, SxxSS, and CtBP-binding sites, SxxSS showed the most repressive effects. Although TCF-4J was not the obvious transcription activator type, it showed the most carcinogenic effects. TCF-4B could be the default isoform.

4.5. METHODS

4.5.1. Cloning and plasmids

The fourteen TCF-4 isoforms (TCF-4A – M, X) were sub-cloned into mammalian expression vector plasmids pcDNA3.1(-)IRES-EGFP (gift from Dr. X. Bo, University of London, UK) and pcDNA3.1/Myc-His (Invitrogen) via *Xho I* (or *Eco RI*) and *Bam HI* restriction sites using supercompetent bacterial cells (Stratagene/Agilent Technologies, La Jolla, CA). All constructs were verified by DNA sequencing. These expression vectors were useful in confirming the expression of the TCF-4 isoforms in various transfection assays. The Myc tag was utilized for confirming the expression pattern by WB when the pcDNA3.1/Myc-His plasmids were used. Simple visualization of EGFP (enhanced green fluorescent protein) expression from the assay plates under the fluorescence microscope quickly verified the expression of TCF-4 isoforms when the pcDNA3.1(-)IRES-EGFP plasmids were used. Owing to the IRES (internal ribosome entry site) sequences in these vectors, EGFP and TCF-4 isoforms were translated simultaneously. Dominant negative form of TCF-4E (dnTCF-4E), and the TCF/LEF reporter plasmids TOPFlash and FOPFlash were purchased from Upstate/Millipore The β-catenin expression plasmids were purchased from OriGene (Rockville, MD).

4.5.2. Cell culture and transfection studies

The human cell lines HepG2, Hep3B, Huh7, and FOCUS, and HEK293 were maintained as described in Chapter 2 (Section 2.5.1). For transient transfection, cells were plated in 6-well or 12-well plates, and plasmids were transfected by a *Trans*IT-LT1 transfection reagent. To establish cell lines stably expressing TCF-4 isotypes, pcDNA3.1(-)IRES-EGFP vectors carrying TCF-4A, B, C, J, and K isoforms, and an empty vector (EV) were

transfected into Huh7 and FOCUS HCC cells. The positive colonies were selected by using 400 μg/ml of geneticin (G418, Invitrogen) for Huh7 and 600 μg/ml of G418 for FOCUS. G418 is toxic to all cell types as this antibiotic blocks polypeptide synthesis by inhibiting the elongation step during mRNA translation. Resistance to G418 is conferred by the *Neomycin* gene, which is expressed from the pcDNA3.1(-)IRES-EGFP plasmids. Positive colonies were also verified by the expression of EGFP through fluorescence microscope.

4.5.3. Western blot analysis

Each pcDNA3.1/Myc-His plasmid expressing the fourteen TCF-4 isoforms was co-transfected with a β-catenin expression plasmid into HEK293 and Huh7 cells seeded in 6-well plates. Forty-eight hours after the transfection, the whole cell lysis and WB were carried out as described in Chapter 2 (Section 2.5.2). The following antibodies were used for probing: anti-TCF-4, anti-β-catenin, anti-β-actin, anti-Myc tag antibody diluted at 1:2000 (Cell Signaling, Danvers, MA), anti-HA tag antibody at 1:2000 dilution (BD Biosciences), mouse monoclonal anti-TBX3 antibody at 1:5000 dilution (Millipore), rabbit polyclonal anti-TLE 1/2/3/4 antibody at 1:500 dilution (Cell Signaling), mouse monoclonal anti-GS antibody at 1:1000 dilution (BD Biosciences), rabbit monoclonal anti-c-Myc antibody diluted at 1:2000 (Abcam, Cambridge, MA), rabbit polyclonal anti-cyclin D1 antibody at 1:500 dilution (Santa Cruz Biotechnology).

4.5.4. TCF/LEF transcriptional activity assay

To determine the transcriptional activity of the TCF-4 isoforms, a TCF/LEF reporter gene assay was employed. Each pcDNA3.1(-)IRES-EGFP plasmid expressing the fourteen TCF-4 isoforms or HA-tagged dnTCF-4E construct was co-transfected with TOPFlash or

FOPFlash, and β-galactosidase in the presence or absence of a β-catenin expression plasmid into HEK293 and Huh7 cells seeded in 12-well plates. For the stable Huh7 and FOCUS cell lines, TCF-4 isoforms were not transfected. The β-galactosidase activity was used for normalization of transfection efficiency. Transfection of β-catenin and its ectopic overexpression simulated the activation of Wnt signaling. Expression of TCF-4 isoforms was confirmed by both WB analysis and the visualization of EGFP under the Olympus IX70 fluorescence microscope (DSC Optical Services, Newton Centre, MA). Forty-eight hours after the transfection, the luciferase activity was measured using the Luciferase Assay System (Promega, Madison, WI) by a luminometer TopCount microplate reader (Packard Instrument Company, Meriden, CT). The relative transcriptional activity was determined by the ratio of TOP value to FOP basal level. FOCUS, Huh7, or HEK293 cells were at 50% confluency the day before transfection. Results were presented as the mean of triplicates and each experiment was repeated at least three times.

4.5.5. Cell proliferation assay

Cell growth rate was measured by using the CellTiter 96 AQeous Non-Radioactive Cell Proliferation Assay (Promega). Each pcDNA3.1(-)IRES-EGFP vector expressing the TCF-4 isoforms or HA-tagged dnTCF-4E construct was co-transfected with a β-catenin expression plasmid into Huh7 cells seeded in 6-well plates. For the stable Huh7 and FOCUS cell lines, only the β-catenin expression plasmid was transfected to activate Wnt signaling. Twenty-four hours after the transfection (day 0), the cells were plated in 96-well plates at a density of 2000 cells/well (Huh7) or 1000 cells/well (FOCUS). The cells were incubated at 37°C and the DMEM media containing 10% FBS were changed every 48 hours. Cell proliferation was measured every day at the same time of the day. On the

day of measurement, the cells were incubated in a combined MTS/PMS solution (Promega) for 1 hour, and the plate reading was taken at the absorbance of 490 nm by using SpectroMax M5 Multi-Mode Microplate Reader (Molecular Devices, Sunnyvale, CA). It is a colorimetric method for determining viable proliferating cells as MTS (tetrazolium salt) is converted into aqueous soluble formazan product by dehydrogenase enzymes found in metabolically active cells. The absorbance of the formazan product at 490 nm can be measured directly from 96-well assay plates, and the quantity of formazan is directly proportional to the number of living cells in the culture. Results are presented as the average absorbance of six wells in one experiment and reported as means of triplicate assays. The slope of each proliferation curve was computed by using the following equation: $S = [\sum(t - \overline{t})(q - \overline{q})]/[\sum(q - \overline{q})^2]$, where t was the time points in days and q was the rate of cell proliferation; $\overline{t}$ and $\overline{q}$ were the average values. The doubling time was calculated with the formula $Td = (t_2 - t_1)[\log 2/\log(q_2/q_1)]$, where t was the time points in hours and q was the rate of cell growth. The average cell doubling time was determined as the mean value of doubling times calculated from the cell proliferation rates of the first two days (i.e. $t_1 = 0$ h, $t_2 = 24$ h; $t_1 = 24$ h, $t_2 = 48$ h; $t_1 = 0$ h, $t_2 = 48$ h).

4.5.6. Cell migration by wound healing assay

Cell migration was assessed using a scratch wound assay. HCC cell lines Huh7 and FOCUS stably expressing isoforms TCF-4A, B, J, and K, or empty vector were transfected with β-catenin to activate the canonical Wnt signaling. Twenty-four hours after the transfection, 100% confluent monolayer cells in 6-well plate were scratched with a sterile micropipette tip to generate wound, and the cells were washed to remove floating cells and debris. The wound closure, as an index of cell migration, was observed

over 48-hour period and photographed at 0, 8, 16, 24, 32, 40, and 48 hours from the same wound area. Phase-contrast images were acquired with a MicroFire® Microscope Digital CCD Camera (Optronics, Goleta, CA) under the Olympus IX70 fluorescence microscope using Picture Frame software (Optronics). The mean wound width was calculated by dividing the wound area with the constant height using the Image J software. Results were expressed as a percent of wound closure normalized to wound width at time 0 hour. Results were derived from three independent measurements.

4.5.7. Cell migration by transwell (Boyden chamber) cell motility assay

A luminescence-based assay was used to evaluate cell migration. This assay assesses non-migrated, migrated adherent, and migrated non-adherent cells through uncoated filters in cell culture inserts partitioned with 22 mm diameter, 8 μm pore polyethylene terephthalate (PET) membranes (BD Biosciences). Huh7 or FOCUS HCC cells stably expressing control vector or any of the TCF-4 isoforms A, B, C, J, and K were transfected with β-catenin expressing plasmid. After 24 hours of starvation, 1×10^5 Huh7 or 5×10^4 FOCUS cells were resuspended in 0.5 ml of DMEM containing 0.1% FBS and seeded into cell culture inserts (upper chamber) containing translucent PET membrane within a 12-well plate (lower chamber). The lower chamber was filled with 0.75 ml of MEM containing 10% FBS as an attractant. For control in determining the total number of experimental cells, 1×10^5 Huh7 or 5×10^4 FOCUS HCC cells in 0.5 ml of DMEM containing 0.1% FBS were seeded into a lower chamber without the insert. After incubating for 18 hours at 37°C, the cells on the upper surface of the membrane (non-migrated) were harvested with a sterile cotton swab and placed into a well in 96-well plate. The cells migrated to the lower surface of the membrane (migrated adherent), and

the cells migrated and dripped to the lower chamber (migrated non-adherent) were combined and placed into another well. The cells in the control well were also harvested into a third well. The 96-well plate with non-migrated, migrated, and control cells was incubated with ATPLite substrate (PerkinElmer) for 2 minutes, and the luminescent counts per second was measured by SpectroMax M5 Multi-Mode Microplate Reader. The results were expressed as a percentage of migrated cells to total number of cells, and the experiment was performed in triplicate. For microscopic visualization, the migrated adherent cells were stained with crystal violet. The non-migrated cells from the upper surface of the upper chamber were removed with a cotton swab. The migrated cells on the lower surface and the control cells were fixed with 4% paraformaldehyde for 10 minutes in room temperature. After washing with PBS, the cells were stained with 4% crystal violet for 30 minutes in room temperature. The cells were visualized and counted at 40× magnification.

4.5.8. Soft agar colony formation

Anchorage independent cell growth was determined by soft agar colony formation assay using Cell Transformation Detection Assay kit (Chemicon/Millipore). Huh7 or FOCUS HCC cells stably expressing control vector or TCF-4B, J, and K isoforms were transfected with β-catenin expressing plasmid. Twenty-four hours post-transfection, 2.5 × 104 Huh7 or 1 × 104 FOCUS cells were resuspended in 0.4% top agar and spread over a bottom layer of 0.8% base agar in 6-well plates. The solidified soft agar was overlaid with DMEM media containing 10% FBS. The media was changed every 4-5 days. Twenty-four days later, the colonies were visualized by staining with cell stain solution, and the number and size of the colonies were analyzed by using Image J software.

4.6. REFERENCES

1. Hoppler S and Kavanagh CL (2007) Wnt signalling: variety at the core. *J Cell Sci*, 120: 385-393.
2. Arce L, Yokoyama NN and Waterman ML (2006) Diversity of LEF/TCF action in development and disease. *Oncogene*, 25: 7492-7504.
3. Brannon M, Brown JD, Bates R, Kimelman D and Moon RT (1999) XCtBP is a XTcf-3 co-repressor with roles throughout Xenopus development. *Development*, 126: 3159-3170.
4. Liu F, van den Broek O, Destree O and Hoppler S (2005) Distinct roles for Xenopus Tcf/Lef genes in mediating specific responses to Wnt/beta-catenin signalling in mesoderm development. *Development*, 132: 5375-5385.
5. Daniels DL and Weis WI (2005) Beta-catenin directly displaces Groucho/TLE repressors from Tcf/Lef in Wnt-mediated transcription activation. *Nat Struct Mol Biol*, 12: 364-371.
6. Mosimann C, Hausmann G and Basler K (2009) Beta-catenin hits chromatin: regulation of Wnt target gene activation. *Nat Rev Mol Cell Biol*, 10: 276-286.
7. Barolo S (2006) Transgenic Wnt/TCF pathway reporters: all you need is Lef? *Oncogene*, 25: 7505-7511.
8. van de Wetering M, Cavallo R, Dooijes D, van Beest M, van Es J, Loureiro J, Ypma A, Hursh D, Jones T, Bejsovec A, Peifer M, Mortin M and Clevers H (1997) Armadillo coactivates transcription driven by the product of the Drosophila segment polarity gene dTCF. *Cell*, 88: 789-799.

9. van de Wetering M, Sancho E, Verweij C, de Lau W, Oving I, Hurlstone A, van der Horn K, Batlle E, Coudreuse D, Haramis AP, Tjon-Pon-Fong M, Moerer P, van den Born M, Soete G, Pals S, Eilers M, Medema R and Clevers H (2002) The beta-catenin/TCF-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell*, 111: 241-250.

10. Hovanes K, Li TW and Waterman ML (2000) The human LEF-1 gene contains a promoter preferentially active in lymphocytes and encodes multiple isoforms derived from alternative splicing. *Nucleic Acids Res*, 28: 1994-2003.

11. Molenaar M, van de Wetering M, Oosterwegel M, Peterson-Maduro J, Godsave S, Korinek V, Roose J, Destree O and Clevers H (1996) XTcf-3 transcription factor mediates beta-catenin-induced axis formation in Xenopus embryos. *Cell*, 86: 391-399.

12. Van de Wetering M, Castrop J, Korinek V and Clevers H (1996) Extensive alternative splicing and dual promoter usage generate Tcf-1 protein isoforms with differential transcription control properties. *Mol Cell Biol*, 16: 745-752.

13. Nakabayashi H, Taketa K, Miyano K, Yamane T and Sato J (1982) Growth of human hepatoma cells lines with differentiated functions in chemically defined medium. *Cancer Res*, 42: 3858-3863.

CHAPTER 5

5. DISCUSSION

5.1. HYPOTHESIS

The TCF/LEF family of proteins mediates Wnt signals in the nucleus as bipartite factors by binding to DNA and acting either as transcriptional activators or as repressors depending on their association with different transcriptional co-regulators. Expression of target genes is achieved when TCF/LEFs interact with stabilized β-catenin during active Wnt signaling. On the other hand, in the absence of a Wnt signal and free β-catenin, TCF/LEFs associate with transcriptional co-repressors, such as TLE and CtBP, resulting in inhibition of target gene transcription. In this regard, TCF/LEFs are defined as docile transcription factors whose functional roles are dictated by the relative amount of co-activators *vs*. co-repressors (Figure 5.1).

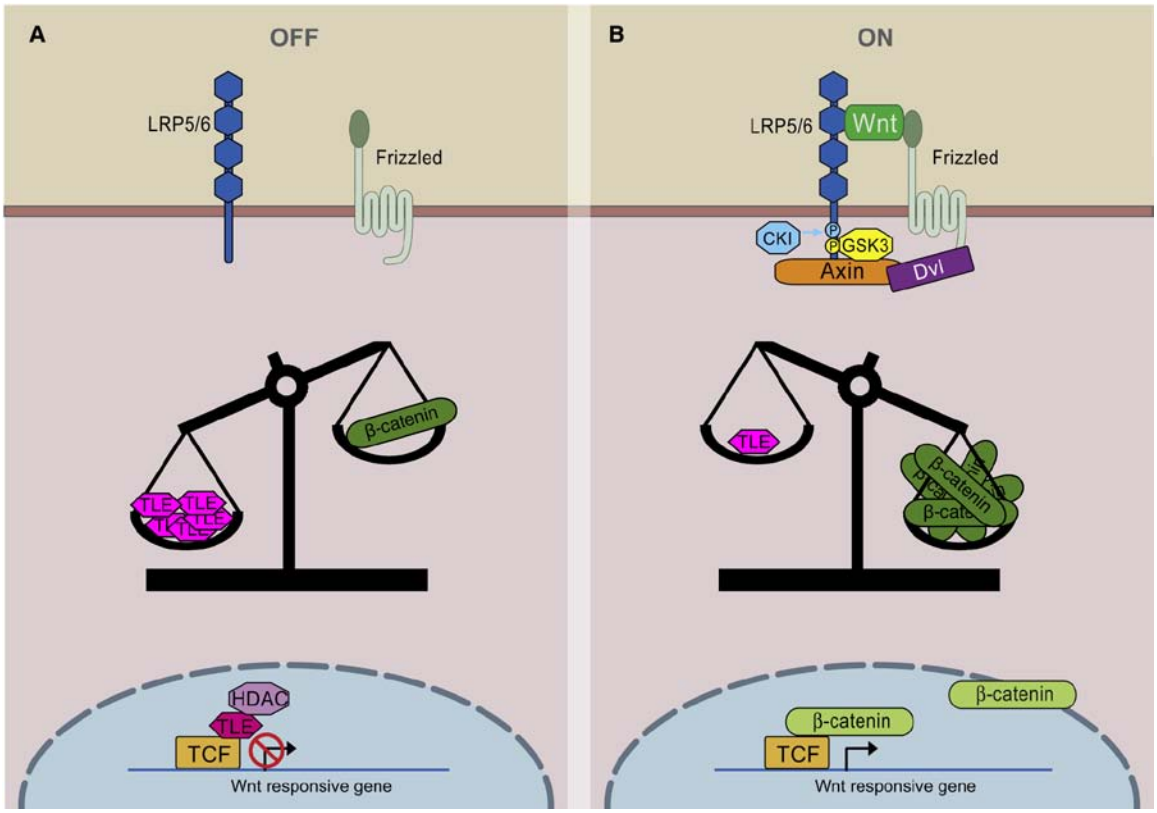


Figure 5.1. Current definition of TCF-4 transcription factors

TCF-4 is portrayed as a docile, yet two-faced protein bound to the promoters of Wnt target genes, where its effector role (transcriptional repression *vs*. activation) is decided by the relative abundance of co-repressors *vs*. co-activators. *A*. When the amount of co-repressors, like TLE, exceeds the amount of co-activators, such as β-catenin, TCF-4 "befriends" with the co-repressors and inhibits the transcription of target genes. *B*. If the amount of co-activators overwhelms the co-repressors, then TCF-4 interacts with β-catenin, converts into a transcriptional activator, and "kicks out" TLE.

TCF-4 is the highly expressed member of the family in hepatocyte transformation and HCC. Although its gene, *TCF7L2*, has been known for its ability to express more than 500 splicing variants, specificity of isotypes expressed in HCC is not available and functional differences between various isoforms have not been described. Consequently, it is not known whether all of the expressed TCF-4 isoforms fit into above description as transcriptional switches or whether diverse splice variants have different functional properties. If above description were applicable to TCF-4, then all 500+ isoforms would have the same function. In other words, regardless of isotype difference, all TCF-4 proteins on WREs would be converted from transcriptional repressors into activators upon arrival of β-catenin. Then, why would a cell need that many isoforms when one generic type could do it all by switching between on and off states? Therefore, I challenged above description, at least for TCF-4, and I aimed to identify various isoforms expressed in HCC. I hypothesized that distinctive splice variants might lead to different biologic phenotypes and that their expression might be regulated according to their specific functional role depending on particular situations. Besides above conflicting

description of TCF/LEFs, the reasoning behind my hypothesis was influenced by several other factors:

1. Although both binding domains for the primary co-activator (BCBD) and co-repressor (TLE/Gro) are conserved and constitutively expressed, the two CtBP-binding sites are only present in long isoforms. This suggests that long isotypes are more likely to show transcriptional repression than compared to short isoforms since CtBPs have been described as strong co-repressors. Therefore, splicing variation may dictate the functional difference between long and short isotypes.
2. A frameshift mutation in the poly A tract has been described as a genetic alteration in CRC. This mutation, which switches potential long isoforms into short ones, provides the cancer cells a way to avoid from expressing TCF-4 isoforms with the CtBP co-repressor-binding sites. Hence, such mechanism implies that short isoforms are favored in cancerous cells over the long counterparts, further strengthening my previous point.
3. In *TCF7L2*, the LVPQ and SxxSS sequences that emerged due to additional splice donor and acceptor sites in exons 7 and 9, respectively, have been described to enhance transcriptional repression. Therefore, isoforms that contain these functional motifs may exert stronger transcriptional inhibition than the ones that lack these motifs.
4. TCF-3, another member of the TCF/LEF family expressed as a single type without splicing variations, has been described to exclusively function as a transcriptional repressor. Interestingly, TCF-3 constitutively expresses both

LVPQ and SxxSS motifs and the two CtBP-binding sites, which may have contributed to its characteristics of transcriptional inhibition. Therefore, a TCF-4 isoform structurally homologous to TCF-3 may deliver similar functional properties.

5. Upon Wnt stimulation, TCF-3 repression is relieved not by converting TCF-3 into an activator, but by replacing it with another transcription factor, indicating that a TCF/LEF with added repressive domains is not favored for switching into an activator. This mechanism could also be employed by TCF-4 isoforms, such that repressive splice variants could be replaced with activating counterparts when a cell was stimulated by a Wnt ligand, thus these opposite isoforms would be expressed at different circumstances.
6. While TCF-3 constitutively expresses LVPQ, SxxSS, and CtBP-binding sites, TCF-1 and LEF-1, which are usually utilized for target gene activation, do not even contain these repression-specific domains. Since TCF-4 could generate any combination of these motifs by alternative splicing, TCF-4 isoforms with structural resemblance to TCF-1 and LEF-1 could represent the transcription activation-specific splice variants.
7. TCF-1 and LEF-1 have also been known to produce dominant negative forms due to the presence of second promoter sequences in their respective genes. Because these truncated proteins do not have BCBDs and since transcriptional activation is only achieved upon β-catenin binding, dnTCF-1 and dnLEF-1 constitutively inhibit the expression of target genes regardless of Wnt signals. Utilization of the second promoter is tightly controlled and the expression of these repressor forms

is balanced with their respective full-length counterparts to efficiently regulate the overall outcome of the Wnt signaling. However, such mechanism of creating an optional repressor isoform through alternative promoter usage has not been identified in TCF-4. Since TCF-3, but not TCF-1 and LEF-1, possesses the transcription repression-specific motifs, LVPQ, SxxSS, and CtBP-binding sites, expression of these motifs in TCF-4 by elective splicing mechanism might provide very strong repressive isoforms that could perform the tasks a dominant negative form.

5.2. FINDINGS AND CONCLUSION

In accordance with my proposal, I was able to identify more than a dozen different TCF-4 isoforms in HCC cell lines. Because they were obtained from complete mRNA clones, I was able to use them for further investigation by sub-cloning into mammalian expression vectors. By quantitative real-time and semi-quantitative RT-PCR, endogenous mRNA levels of these isoforms revealed a diverse expression pattern further substantiating my hypothesis that different isoforms might be required for particular circumstances. Finally, by examining them by several functional assays, their role as transcription factors was determined. The summary of all isoforms with some of their characteristics is illustrated in Figure 5.2 for comparative view to emphasize the importance of the SxxSS motif, which provided most of the exciting outcomes. Taken together, the results overall supported my theory that transcription activation- and repression-specific isoforms existed, that their upregulation differently affected the overall cellular phenotype, and thus their expression was required in precise settings according to their functional specificity. Here are some of the isoforms that significantly influenced the phenotype of HCC cells.

The repressive effect of SxxSS was first noticed in differential transcriptional activity between TCF-4A and B. Although, TCF-4B showed high activity, it was expressed uniformly in tumor, peritumor, and normal tissues, and its effect on cell proliferation, migration, and anchorage-independent cell growth was not different from those of EV control. Therefore, TCF-4B could be a neutral isoform that was expressed throughout different tissues and developmental/cancer stages as a “gatekeeper” without affecting the overall cell behavior. On the other hand, TCF-4A exhibited consistent,

though minimal, repressive function in transcriptional activity, cell proliferation, and migration. In addition, its expression was significantly reduced in HCC tumor tissues compared to normal and peritumor samples. Therefore, TCF-4A must be a transcriptional repression-specific isoform. TCF-4A and B will be useful in the future studies in exploring the mechanism of the SxxSS motif.

Expression of TCF-4C produced the most carcinogenic phenotype. Originally identified in Huh7 cells, its endogenous expression was restricted to the HCC tumor tissues and to the more carcinogenic HCC cells, Huh7 and FOCUS, immediately implying its importance in HCC. Moreover, TCF-4C had the highest transcriptional activity and the cell proliferation rate, revealing its significance in cancerous state (Figure 5.2). Therefore, TCF-4C must be a transcription activation-specific isoform, and that's why its expression was only required in tumors. In addition, TCF-4C was an isoform that did not contain the exon 4, LVPQ, SxxSS, and the CtBP-binding sites, thus structurally resembling to TCF-1 and LEF-1. During aberrant activation of Wnt signaling in HCC, TCF-4C must have been one of the main downstream isoforms to activate the expression of target genes implicated in cell proliferation and possibly other aggressive phenotypes. However, the latter could not be investigated as I was unable to propagate a stable cell line expressing TCF-4C, suggesting that overexpression of this isoform was too severe even for the cancerous cells. Therefore, TCF-4C could be better studied with an inducible expression system to see its full effects. Nevertheless, the available results of TCF-4C supported my hypothesis the most.

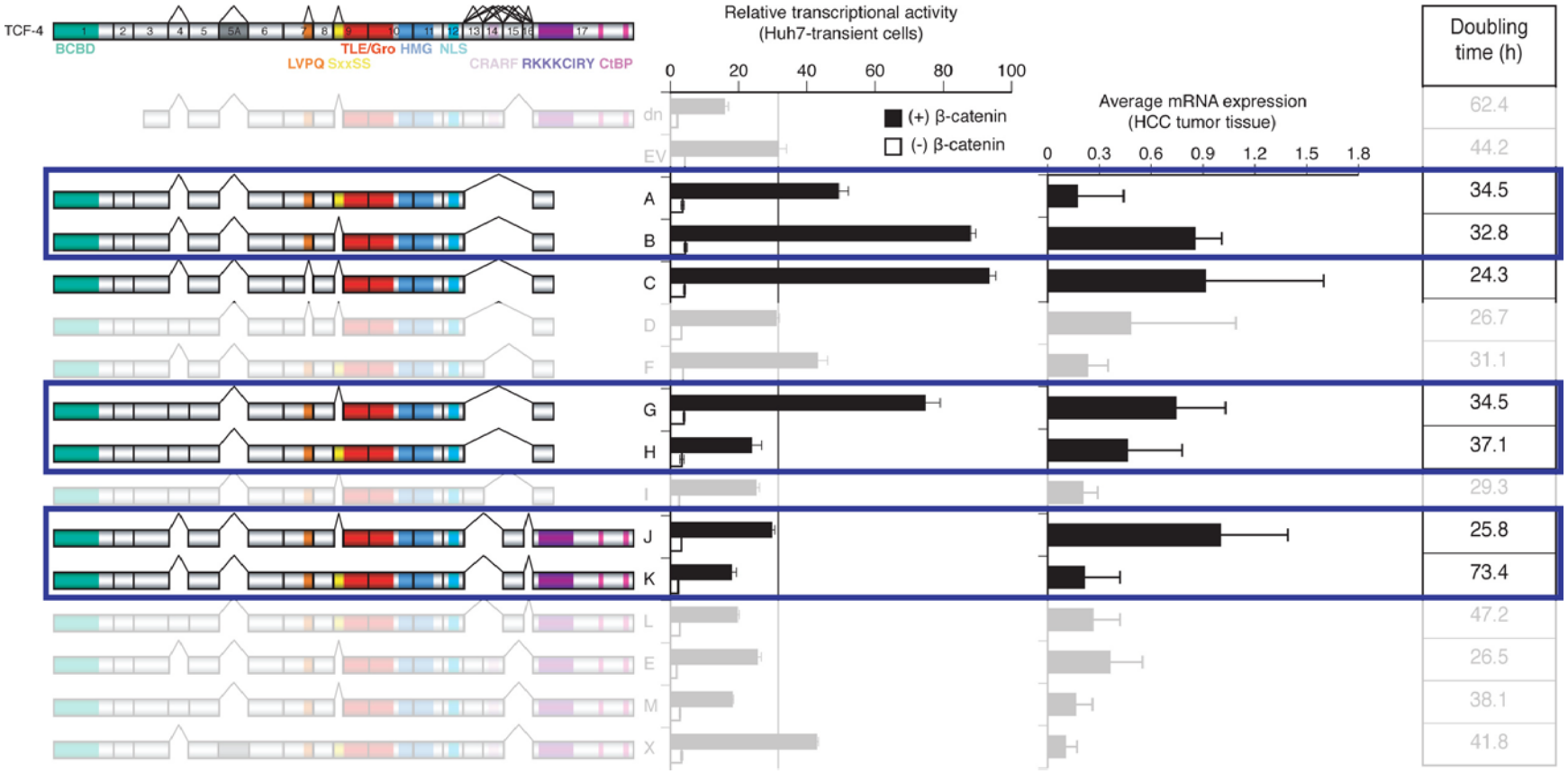


Figure 5.2. Identification and characterization of TCF-4 isoforms

The significance of repressive effect of SxxSS motif is emphasized in paired isoforms TCF-4A and B, TCF-4G and H, and TCF-4K and J as evidenced by their transcriptional activity, expression pattern in HCC tumor tissues, and influence on cell proliferation (blue rectangles). The isoform TCF-4C is also emphasized as it demonstrated the most oncogenic phenotypes.

Another splice variant that produced carcinogenic characteristics was the long isoform, TCF-4J. Although the transcriptional activity of TCF-4J was low, it was the highest among the long isoforms (Figure 5.2). Otherwise, the generous expression of this isotype in tumor tissues was comparable to that of TCF-4C. However, in contrast to TCF-4C, TCF-4J was also expressed in the peritumoral tissues suggesting that TCF-4J was involved in an early event of hepatocarcinogenesis. Moreover, TCF-4J elevated the cell proliferation rate to the second highest level just after TCF-4C, induced cell migration at the fastest rate (Figure 4.5, p. 168), and triggered the formation of numerous and large colonies on soft agar (Figure 4.6, p. 170). Interestingly, all of these oncogenic characteristics were produced because of the absence of the SxxSS motif. Although not obvious from its transcriptional activity, its expression and functional roles implied that TCF-4J could be a gene activator isoform. However, the intriguing question was why did TCF-4J, which had a long tail with CtBP co-repressor-binding sites, but not short isoforms (without the SxxSS motif), showed such tumorigenic phenotypes? The answer could be in the long C-terminal tail itself. Along with the CtBP-binding sites, the exon 17 carried another motif, the RKKKCIRY sequence (Figure 2.1, p. 93). This sequence had been proposed to be a secondary DNA-binding site that gave additional specificity in target gene recognition. Among these Wnt target genes recognized by long tails of TCF-1 and TCF-4 were *TCF* and *LEF* genes themselves, thus making a positive feedback loop (Section 1.2.4, p. 66). In CRC, this positive feedback circuit was often exploited for continuous propagation of TCF/LEFs (Section 1.2.3.1.5, p. 49). The same mechanism could be employed in HCC, where upregulation of TCF-4J promoted the expression of *TCF7L2* itself and *LEF1* as Wnt targets. Indeed, expression of LEF-1 had been

occasionally reported in HCC, while its transcription activation-specific role had been known very well. In context of *TCF7L2*, TCF-4J might lead to overexpression of isoforms that were specific for target gene activation, such as TCF-4C and TCF-4J itself for repeated expression of more such isotypes. However, because different TCF-4 isoforms were generated at the pre-mRNA level (after transcription) by other regulatory mechanisms, the job for TCF-4J must just be to recognize and activate the expression of its own gene. Therefore, the overall effects seen in overexpression of TCF-4J could be involved with its secondary and tertiary targets, such as LEF-1, TCF-4C and others. Because the transcriptional activity of TCF-4J due to CtBPs was low and because making stable cells overexpressing TCF-4C was not feasible, expression of these target genes was probably kept at low level to achieve the best results for HCC. Another observation was that the full extent of repressive effect of CtBPs was reached in the presence of SxxSS, as evidenced by TCF-4K. Therefore, each combination of alternative motifs of TCF-4 must be unique in the expression of target genes. It would be very interesting to see the effect of TCF-4J as well as TCF-4K *in vivo* models. In addition, the target genes specifically expressed by these isoforms should be identified by large-scale gene expression analysis, and difference in those targets might give an insight on molecular mechanism of these almost identical splice variants.

Finally, my theory was further strengthened by the splice variants TCF-4K, L, and M. These isoforms, not only had the domain structures opposite to TCF-4C, but also revealed exact reverse characteristics. These long isoforms were expressed in tumor tissues at very low levels, and they exhibited the lowest transcriptional activities and slowest cell proliferation rates (Figure 5.2). Therefore, these isoforms must be the

transcription repression-specific isoforms, and that's why their expression was higher in normal tissues than tumor samples. During overactivation of Wnt signaling in HCC, they must have been replaced by other isoforms, like TCF-4C and J, but not converted into activators. In addition, because their transcriptional activity was comparable with that of dnTCF-4E even in the presence of β-catenin, a possibility of these isoforms to represent endogenous dominant negative forms of TCF-4 was suggested. Indeed, further examination of TCF-4K revealed that it opposed a malignant potential by effectively reducing the anchorage-independent cell growth (Figure 4.6, p. 170). More importantly, this observation highly substantiated the repressive effect of the SxxSS motif. However, the induction of high cell migration rate implied that TCF-4K might not always monotonously behave as dominant negative when the cell survival was in question. It could be explained by the specific target genes TCF-4K was repressing, such as *TCF7L2* and *LEF1*. Because TCF-4K already had the BCBD, these target genes could be put into a good use by other molecular mechanisms in such situations. These suppositions need to be confirmed by further experimentation, though it raises the difficulty of distinguishing distinct yet near identical isoforms within each assay setup.

All in all, few of the TCF-4 isoforms structurally similar to TCF-1/LEF-1 were found to be target gene activation-specific splice variants and were preferentially expressed in HCC tumor samples, while several isoforms that included many of the motifs, such as exon 4, LVPQ, SxxSS, and CtBP-binding sites, turned out to be the repression-specific isotypes and were expressed in normal tissues more than tumor. Among these motifs, SxxSS showed the strongest repressive effect. When combined with CtBP-binding sites, the resultant isoform may even represent dominant negative forms.

5.3. PERSPECTIVES

TCF/LEF transcription factors are defined as molecular switches between repressors and activators, depending on the relative amount of co-regulators. While this could be true for some neutral isoforms, most TCF-4 isotypes may not get converted between the two. Instead, an isoform may be replaced by another one when the functional role of the former is no longer needed. Therefore, the definition of TCF/LEFs, at least for TCF-4, could be amended as follows: In the absence of Wnt signals, the target genes are not expressed as repressive isoforms of TCF-4 occupy their promoters. In the absence of β-catenin, these isoforms are upregulated and preferentially interact with co-repressors, TLE, whose expression is also increased at this time. On the other hand, upon Wnt stimulation, β-catenin is stabilized and enters the nucleus, where repressive TCF-4 isotypes are downregulated and ready to transfer the WREs to activating isoforms, which are now overexpressed. The activating isoforms interact with β-catenin to initiate the transcription of target genes (Figure 5.3). During the time when the WREs are delivered from repressive to activating isoforms, the neutral splice variants may occupy it for temporary basis until the expression of the activating isoforms reached to a certain level. Therefore, expression of Wnt target genes is not only dependent on the relative abundance of β-catenin *vs*. TLE, but also on the expression of activating *vs*. repressive isoforms. However, the exact molecular mechanism needs to be validated although it again brings up the challenge of distinguishing near identical isoforms at protein level.

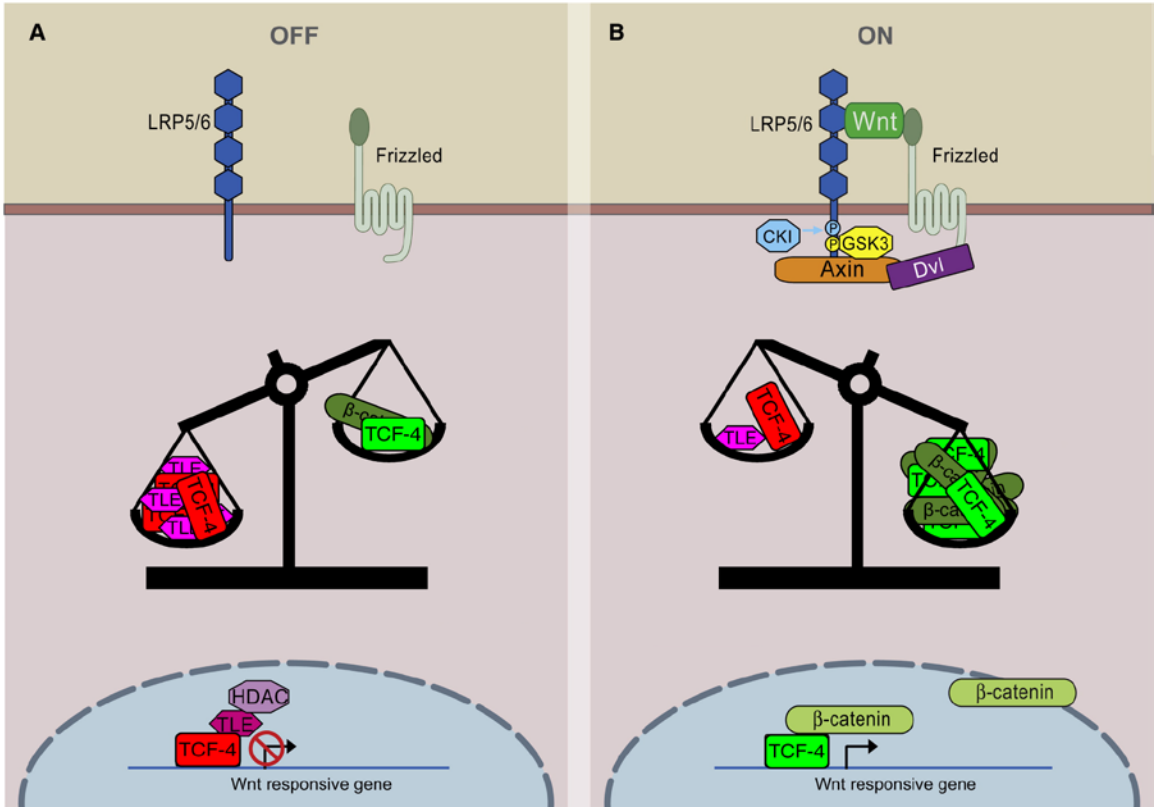


Figure 5.3. Updated definition of TCF-4 transcription factors

Instead of siding with whichever co-regulatory protein is expressed the most and getting converted between transcriptional activator and repressor, different TCF-4 isoforms replace one another according to their calling. *A*. In the absence of Wnt signals, not only the amount of TLE exceeds that of β-catenin but also the repressive TCF-4 isoforms (red) are upregulated compared to activator isotypes (bright green). These repressive splice variants and the co-repressors harmoniously interact and mutually inhibit the expression of Wnt target genes. *B*. In the presence of Wnt signaling, the opposite effect is observed. While β-catenin is stabilized and accumulated, the co-repressors are downregulated. At the same time, the amount of repressive TCF-4 isoforms decreases to free the WREs for the activator isotypes, whose expression is rapidly increasing.

5.4. SUMMARY

Background and Aims: The canonical Wnt signaling pathway is frequently activated in hepatocellular carcinoma (HCC). Although β-catenin is the principle effector protein of Wnt pathway, the success of relaying signals to the nucleus depends on the T Cell Factor (TCF) family of transcription factors that bind to promoters of Wnt-responsive target genes. The human TCF-4 gene is composed of 17 exons with multiple alternative splicing sites. In spite of theoretical possibility of constructing more than 500 isoforms, only two TCF-4 splice variants have been identified so far. Therefore, the objective of this thesis is to identify TCF-4 isoforms in HCC, and to evaluate their role in hepatocarcinogenesis.

Methods: Reverse transcription polymerase chain reaction (RT-PCR) was employed to identify distinct TCF-4 isoforms and to determine expression profiles of TCF-4 isoforms in HCC cell lines and human HCC tissues. The function of TCF-4 isoforms in HCC cell lines was examined by TCF transcriptional activity, cell proliferation, wound healing, trans-well motility, and colony formation assays.

Results: I identified 17 different TCF-4 isoforms including two previously known forms TCF-4B (short) and TCF-4E (long) from four HCC cell lines. The 15 new isotypes were named in alphabetical order according to increasing size of their protein products. Functional analysis of these isoforms revealed distinct effects on the phenotype of HCC cells. The TCF-4J isoform produced characteristic features of the malignant phenotype since it generated the highest cell proliferation rate, cell migration and colony formation. In contrast, the TCF-4K isoform displayed the lowest TCF transcriptional activity, cell proliferation rate, and colony formation. Interestingly, TCF-4J and TCF-4K differed by only five amino acids (the SxxSS motif). Moreover, TCF-4J expression was upregulated

in HCC tumors compared to peritumor and normal liver, whereas TCF-4K was downregulated in tumors.

Conclusions: Conserved splicing motifs in TCF-4 may have a major influence on the function of TCF-4 isoforms and alter the characteristics of the malignant phenotype in HCC.