BIOLOGICAL MODELS OF COLORECTAL CANCER METASTASIS AND TUMOR SUPPRESSION PROVIDE MECHANISTIC INSIGHTS TO GUIDE PERSONALIZED CARE OF THE COLORECTAL CANCER PATIENT

By

Jesse Joshua Smith

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

In partial fulfillment of the requirements

For the degree of

DOCTOR OF PHILOSOPHY

In

Cell and Developmental Biology

May, 2010

Nashville, Tennessee

Approved:

Professor R. Daniel Beauchamp

Professor Robert J. Coffey

Professor Mark deCaestecker

Professor Ethan Lee

Professor Steven K. Hanks

To my grandparents, Gladys and A.L. Lyth and Juanda Ruth and J.E. Smith,

fully supportive and never in doubt. To my amazing and enduring parents, Rebecca Lyth and Jesse E. Smith, Jr., always there for me. . . . .my sure foundation. To Jeannine, Bill and Reagan for encouragement, patience, love, trust and a solid backing. To Granny George and Shawn for loving support and care.

And

To my beautiful wife, Kelly,

My heart, soul and great love,

Infinitely supportive, patient and graceful.

ACKNOWLEDGEMENTS

This work would not have been possible without the financial support of the Vanderbilt Medical Scientist Training Program through the Clinical and Translational Science Award (Clinical Investigator Track), the Society of University Surgeons-Ethicon Scholarship Fund and the Surgical Oncology T32 grant and the Vanderbilt Medical Center Section of Surgical Sciences and the Department of Surgical Oncology. I am especially indebted to Drs. R. Daniel Beauchamp, Chairman of the Section of Surgical Sciences, Dr. James R. Goldenring, Vice Chairman of Research of the Department of Surgery, Dr. Naji N. Abumrad, Chairman of the Department of Surgery and Dr. John L. Tarpley, Program Director and Professor of the Department of Surgery for constant support, encouragement and guidance. I would especially like to thank Mrs. Donna Putnam, Dr. Natasha G. Deane, Associate Research Professor, Department of Surgical Oncology, Dr. Nancy J. Brown, Director, Department of Clinical Pharmacology, Dr. Yu Shyr, Professor and Chief, Department of Cancer Biostatistics and Dr. Terence Dermody, Professor and Director of the Vanderbilt Medical Scientist Training Program, who have supported my career goals and who worked tirelessly to provide me with the resources, protected time and encouragement to pursue my dreams and goals. I would also like to thank two individuals who have been taken from us, Niki Carey and Dr. Neville Grant, who provided measures of heroism and inspiration during my early residency and graduate training. I am thankful to Fei Wu, M.D., Ph.D. for getting the 34-gene story off to a great start. I thank Drs. J. Chad Johnson, Carl Schmidt, Nipun B. Merchant, Steven Eschrich, Timothy J. Yeatman and Martin J. Heslin for initial efforts to organize and secure the colorectal tissues used for the microarray analysis in this work.

I am grateful to all of those with whom I have had the pleasure to work during the last 4 years in the laboratory. Each of my Dissertation Committee members has provided excellent professional

guidance and spurred me on to pursue science and research at a new level. Notably, Drs. Mark de Caestecker and Ethan Lee had tremendous roles in my development as a scientist. I am thankful for their guidance, time, patience and generosity with reagents. I would especially like to thank Dr. Bob Coffey and Dr. Steve Hanks who have exemplified a steady hand in my graduate school career and who have provided me with much wisdom and encouragement.

I would like to thank my medical school mentors, Drs. Hazim J. Safi, Charles C. Miller and Anthony L. Estrera for continued support and encouragement in my studies, my life and my career. Dr. Safi has been a stalwart and was the first to introduce me to the wonderful world of medicine and science. I am forever grateful for his inspiration and encouragement. Dr. Miller taught me how to think like a clinical scientist and kept me sane during medical school. Dr. Estrera encouraged me during my pursuit of surgical residency and taught me how to round on patients as a first year medical student. I am thankful to R. Daniel Beauchamp, M.D. and Maximilian Buja, M.D. for giving me a chance.

Nothing has been more important to me than my faith and my family as a bedrock of support during my training. My mom and dad, in-laws, my brother Nate and extended family have provided much grace and love to Kelly and myself. Most importantly, I wish to thank my dear wife Kelly and our yet unknown children for unending love, support and help during my training.

This work was supported by the following grants from the National Institutes of Health: 1) CA69457; DK52334; CA068485; CA077839 (R. Daniel Beauchamp, M.D.); 2) TL1 RR024978; CA106183 (Institutional support for J. Joshua Smith); 3) CA46413; CA95103; CA084239 (R. J. Coffey, M.D.); 4) CA112215 (Timothy J. Yeatman, M.D.); 5) CA126588; CA128323 (Natasha G. Deane, Ph.D.); *Other sources of funding*: Society of University Surgeons-Ethicon Scholarship Award (J.J.S.).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

APC = Adenomatous polyposis coli
TGFβ = Transforming Growth Factor-beta
MAPK = mitogen-activated protein kinase
PI3K = Phosphatidylinositol 3-Kinase
PTEN = phosphatase and tensin homolog
EMT = epithelial-mesenchymal transition
TCF = T cell-specific transcription factor
MET = mesenchymal-epithelial-transition
BMP = Bone Morphogenetic Protein
GDF = Growth and Differentiation Factor
MIS = Muellerian Inhibiting Substance
KCP = kielin-chordin like protein
TβRII = Transforming Growth Factor-beta receptor type II
BMP-2 = Bone Morphogenetic Protein-2
BMPR1A = Bone Morphogenetic Protein receptor IA (also known as Alk3)
FAP = familial adenomatous polyposis
JPS = juvenile polyposis syndrome
α-SMA = alpha smooth muscle actin
BMP-7 = Bone Morphogenetic Protein-7
LEF1 = lymphoid enhancer-binding factor 1
GSK3β = glycogen synthase kinase 3-beta
CK1α = casein kinase 1-alpha
PP2A = protein phosphatase 2A
β-TrCP = beta-transducin repeat containing
LRP = LDL-receptor related protein
Dsh = Dishevelled
sFRP = secreted Frizzled-related protein
DKK = Dickkopf
bHLH = basic-helix-loop-helix
SLUG = snail homolog 2 (also known as SNAI2)
siRNA = small interfering RNA
pT3 = pathologic stage T3 rectal tumor
COX-2 = cyclooxygenase-2
PGE_2 = prostaglandin E_2
EGFR = epidermal growth factor receptor
HBEGF = heparin-binding Epidermal Growth Factor-like growth factor
VEGF = Vascular endothelial growth factor
CD133 = prominin 1
CD44 = CD44 molecule (Indian blood group)
BRAF = B-Raf proto-oncogene serine/threonine-protein kinase
MSH2 = MutS homolog 2, colon cancer, nonpolyposis type 1
MSI-H = microsatellite-high

VMC = Vanderbilt Medical Center
MCC = Moffitt Cancer Center
UAB = University of Alabama-Birmingham
HR = hazard ratio
CI = confidence interval
DMEM = dulbecco's modified eagle medium
ATCC = American Type Culture Collection
BLI = bioluminescence imaging
PBSA = phosphate buffered saline A
AJCC = American Joint Commission on Cancer
FDR = False discovery rate
OS = overall survival
DSS = disease-specific survival
DFS = disease-free survival
Cox PH = Cox Proportional Hazard
CIMP = CpG island methylator phenotype
MSS = microsatellite stable
CTX = adjuvant chemotherapy
IPA = Ingenuity Pathways Analysis
TNF = tumor necrosis factor
IL-1 = interleukin-1
PCR = polymerase chain reaction
RT-PCR = reverse transcriptase polymerase chain reaction
qPCR = quantitative real time polymerase chain reaction
DRB = 5,6-Dichlorobenzimidazole, 1-β-D-ribofuranoside
GFP = green fluorescent protein
ChIP = chromatin immunoprecipitation
EDTA = ethylenediaminetetraacetic acid
CCD = charge-coupled device
KEGG = Kyoto Encyclopedia of Genes and Genomes
RMA = Robust MultiChip Averaging
ANOVA = analysis of variance
POL2RA = RNA polymerase II
SAMP = Ser-Ala-Met-Pro

CHAPTER I

INTRODUCTION TO THE PROBLEM OF COLORECTAL CANCER

Epidemiology and Pathogenesis

Colorectal adenocarcinoma is the second leading cause of cancer-related death in the United States. Currently, there are an estimated 148,000 new colorectal cancer cases and ~50,000 deaths due to this disease annually (Jemal et al. 2009). Colorectal cancer is thought to progress through a series of steps involving mutational inactivation of tumor suppressor genes, activation of oncogenes and eventual progression to metastatic lesions (Fearon and Vogelstein 1990). While localized tumor growth may cause significant organ dysfunction and even death, metastases cause the vast majority (~90%) of human cancer deaths (Hanahan and Weinberg 2000). The ability to metastasize is linked with the ability of cancer cells to invade adjacent tissues, to gain access to vascular or lymphatic channels and to survive transit through the bloodstream so that they may extravasate, then reside and colonize heterologous organs or tissues. Cancer cells acquire the capacity of invasion and metastasis through regulated processes related to tissue development and homeostasis. Many steps in this process resemble a developmental program known as an 'epithelial-mesenchymal transition' as the epithelial cells travel from a primary cancer site to a new site of distant metastasis. Identification of the key regulators of the processes of metastasis and tumor suppression may provide valuable prognostic information to patients with colorectal cancer as well as opportunities for optimal, personalized intervention.

Genomic Instability

Chromosomal instability is the most common type of genomic instability in colorectal cancer and leads to changes in chromosomal number and structure (Lengauer, Kinzler, and Vogelstein 1997; Markowitz and Bertagnolli 2009). Chromosomal instability can lead to loss of wild-type copies of tumor suppressor genes (e.g. *APC* or *Smad4*), which normally prevent malignant progression (Hahn et al. Dpc4, a candidate tumor suppressor gene at human chromosome 18q21.1 1996; Morin et al. 1997). Interestingly, colorectal cancer is not characterized by gene amplification or gene rearrangement (Leary et al. 2008); however, DNA-repair defects can occur and can even be inherited. Patients with hereditary nonpolyposis colon cancer have germ-line defects in mismatch-repair genes and this confers an 80% lifetime risk of colorectal cancer (Lynch et al. 2008). Loss of mismatch-repair function and epigenetic silencing of genes are important in a proportion of patients with colorectal cancer, but fall outside the scope of this dissertation (Ionov et al. 1993; Issa 2004) and will not be discussed in detail.

Mutational Inactivation of Tumor-Suppressor Genes

Wnt signaling pathway activation is considered the initiating event in sporadic colorectal cancer formation (Goss and Groden 2000). The most common mutation in colorectal cancer inactivates the *adenomatous polyposis coli* gene (*APC*, see Fig. 1). APC is a key component of a cellular protein complex, which normally regulates the degradation of β-catenin. β-catenin is the central mediator of Wnt signaling and has important roles both developmentally and in carcinogenesis. Inactivation of *APC* leads to inappropriate nuclear translocation of β-catenin where it activates transcription of target genes such as *cyclin D1* or *c-MYC* to promote proliferation and de-differentiation. The second key genetic step in colorectal cancer development is inactivation of the p53 pathway via mutation of *TP53*

((Markowitz and Bertagnolli 2009), also see Fig. 1). Usually both *TP53* alleles are lost by a missense mutation and loss of chromosome 17p (Baker et al. 1989). *TP53* mediates cell-cycle arrest and the cell-death checkpoint (Vazquez et al. 2008) and inactivation is thought to parallel the adenoma to invasive carcinoma sequence (Baker et al. 1990). A potential role of p53 in the 34-gene recurrence classifier will be discussed in Chapter II. A third step in the colorectal cancer progression sequence is inactivation of the Transforming Growth Factor-β (TGFβ) pathway. Approximately half of colorectal cancers have a defect at the level of the receptor (e.g., TβRII) (Grady et al. 1999; Markowitz et al. 1995) or downstream at the level of the common mediator of TGFβ signaling, Smad4, or its partner transcription factors (e.g., Smad2 or Smad3) (Takagi et al. 1996; Thiagalingam et al. 1996). Inactivation in the TGFβ pathway is thought to be coincident with the adenoma to dysplasia or carcinoma transition (Also see Fig. 1). Wnt signaling and TGFβ superfamily signaling will be discussed in extensive detail below and in Chapter III.

Oncogenic Pathway Activation

Oncogenic mutations of *RAS* or *BRAF* (also see Fig. 1) activate the mitogen-activated protein kinase (MAPK) signaling pathway in ~30% of colorectal cancers (Bos et al. 1987; Davies et al. 2002) and this can potentiate defects in the Wnt or TGFβ pathways. For example, recent observations have been made in regard to cooperative regulation between oncogenic Ras and TGFβ signaling in the process of metastatic behavior (Janda et al. 2002; Kanies et al. 2008). Also, approximately one-third of colorectal cancers have mutations in the Phosphatidylinositol 3-Kinase pathway (e.g. *PI3KA*) and a smaller percentage include activation of other downstream mediators of PI3K signaling and loss of *PTEN* which is an inhibitor of PI3K signaling (Parsons et al. 2005). Although these pathways play

important roles in the pathogenesis of colorectal cancer they are outside the scope of this work and will not be discussed in further detail.

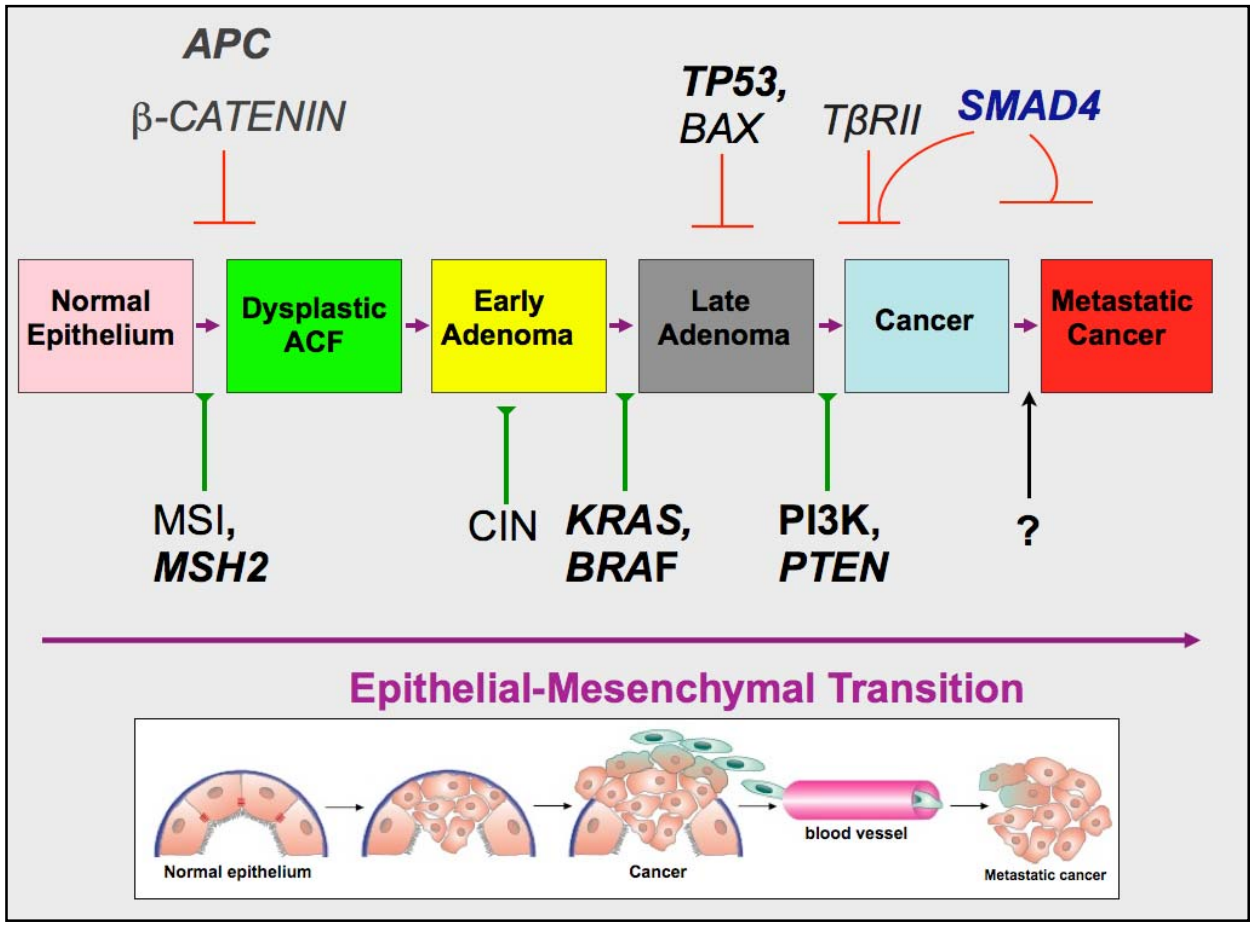


Figure 1. *Pathogenesis of colorectal cancer.* The initiating event in the development of colorectal cancer is thought to be mutation of *APC* as shown. Genetic loss of *TP53* and pathway components of TGFβ signaling are thought to follow. Other contributing pathways and processes thought to be involved in the progression of colorectal cancer are shown (e.g., MSI, *MSH2*, CIN, *KRAS*, *BRAF*, PI3K and *PTEN*). A schematic of an epithelial-mesenchymal transition (EMT) is shown along the bottom and is thought to parallel this progression in a developmentally conserved program of events that regulate the transition from an epithelial cell to a more motile metastatic phenotype. The question mark indicates yet unknown factors. Key: *APC* (adenomatous polyposis coli); MSI (microsatellite-instability); *MSH2* (MutS homolog 2, colon cancer MutS homolog 2, colon cancer, nonpolyposis type 1); ACF (aberrant crypt foci); CIN (chromosomal instability); *BRAF* (B-Raf proto-oncogene serine/threonine-protein kinase); PI3K (Phosphatidylinositol 3-Kinase); *PTEN* (phosphatase and tensin homolog). *Adapted with permission from Fearon ER and Vogelstein B, Cell. 1990 Jun 1;61(5):759-67 and Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. J Clin Invest. 2009 Jun;119(6):1420-8.*

Developmental Biology and Carcinogenesis in Intestinal Epithelial Cells

Overview of Prominent Signaling Pathways

Two major signaling pathways are critical for the development and maintenance of homeostasis in the gastrointestinal tract: the TGFβ superfamily and Wnt/β-catenin/TCF (T cell-specific transcription factor) signaling pathways. These pathways contribute to the development, differentiation and maintenance of homeostasis in intestinal epithelium, whereas mutations or aberrant regulation of these pathways contribute to tumor initiation and progression in the intestine (reviewed in (Radtke and Clevers 2005)). In this section, emphasis will be placed on interactions of the TGFβ superfamily and Wnt pathways and their potential roles in colorectal cancer cell invasiveness and metastatic potential. It is well recognized that other genetic alterations, signaling pathways and processes such as angiogenesis are also important contributors of carcinogenesis, tumor progression and metastasis, but these will only be mentioned briefly.

Metastasis and Epithelial-Mesenchymal Transition: Parallels in Developmental Biology and Cancer

Most solid human tumors are carcinomas that originate in epithelial cells. In order for these cells to invade adjacent layers they become motile and lose cell-cell adhesion then must activate a highly conserved program known as epithelial-mesenchymal transition (EMT) (Thiery 2002). The conversion of epithelial cells into mesenchymal cells was defined in the 1980s when epithelial cells from embryonic and adult lens were cultured in 3D collagen (Greenburg and Hay 1982). The reverse

process of MET can occur and is best described in nephronic epithelium of the developing kidney (Davies 1996). EMT is tightly controlled, reversible and required for embryonic development, tissue reorganization and wound healing (Yang and Weinberg 2008). During EMT, cells lose epithelial polarity and acquire a mesenchymal phenotype (see Fig. 2) with invasive characteristics (Thiery 2003). Cell-cell junctions are disrupted in EMT by mechanisms involving loss of expression of the adherens junction protein and the key gatekeeper of the epithelial state, E-Cadherin (Kang and Massague 2004). Tight junctions lose polarity and function, and the cells undergoing EMT express mesenchymal markers such as vimentin and α-smooth muscle actin (α-SMA). Also, multiple transcription factors can induce EMT and have been associated with tumor invasion and metastasis. For example, Twist1 is essential for metastasis of a mouse breast cancer line to the lung (Yang et al. 2004). In colorectal cancer, nuclear localization of β-catenin has been noted at the invasive front of tumor invasion and the cells appear to have undergone EMT as noted by loss of E-cadherin and gain of mesenchymal markers such as fibronectin (Fodde and Brabletz 2007). In many cases, cellular transformation recapitulates a molecular environment favoring EMT, which in turn disrupts normal epithelial cell polarity and results in the acquisition of invasive and metastatic potential. Thus, a growing body of evidence implicates EMT in tumor cell migration, invasiveness and metastatic behavior (Cui et al. 1996; Kalluri 2009).

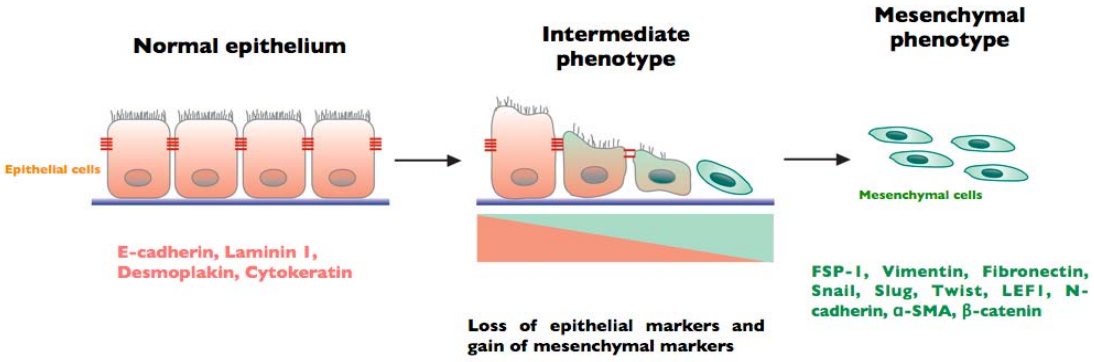


Figure 2. *Schematic of an epithelial-mesenchymal transition*. Epithelial cells marked by expression of E-cadherin undergo a highly conserved developmental program where a loss of epithelial markers is associated with gain of mesenchymal markers (e.g., N-cadherin vimentin, LEF-1). Adapted with permission from Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *J Clin Invest*. 2009 Jun;119(6):1420-8.

TGFβ Superfamily Signaling in Epithelial Biology

The TGFβ superfamily is made up of two subfamilies of cytokines: 1) the TGFβ/Activin/Nodal subfamily and 2) the Bone Morphogenetic Protein (BMP) / Growth and Differentiation Factor (GDF) / Muellerian Inhibiting Substance (MIS) subfamily (Shi and Massague 2003). For purposes of this work, the accessory receptors (e.g., Betaglycan, Cripto, or Endoglin), interactions with inhibin, GDFs, or the Mullerian subfamily (AMH/MIS) and their associations in the TGFβ superfamily will not be further discussed. Specific TGFβ and BMP-related cytokines and their shared signaling intermediates function as critical developmental regulators, cell growth inhibitors and tumor suppressors in normal tissue. Germline loss of certain components of the signaling intermediates for these cytokines often results in embryonic lethality due to their critical roles in development. On the other hand, somatic cell mutations or acquired loss of function may contribute to the development or progression of cancer (Massague 2008).

The TGFβ family of ligands (e.g., TGFβ, BMP, Activin) bind and signal through a heteromeric complex of ligand-specific type I (Alk 1-7) and type II serine/threonine kinase receptors (Shi and Massague 2003). Ligand access to the serine/threonine kinase receptors is regulated by a family of proteins known as ligand traps, proteins that selectively bind to specific TGFβ superfamily ligands, thereby blocking access to the receptors (Neilson 2005; Shi and Massague 2003). For example, the ligand trap protein, Noggin, inhibits receptor activation of the BMP cytokines 2, 4, and 7 (Fig. 3). Other proteins, like kielin-chordin like protein (see Fig. 3) have been noted to enhance BMP signaling to promote a more epithelial phenotype (Lin et al. 2007). A type II receptor is necessary for specific ligand binding. Following binding of ligand to type II receptors, the ligand-bound type II receptor

forms an oligomeric complex with the type I receptor, resulting in type I receptor phosphorylation**.** Information from the tissue microenvironment to the cell nucleus by these growth inhibitors is then transmitted by specific receptor-associated Smad proteins (R-Smads) associated with the growth factor (type II) receptors. Eight members of the vertebrate TGFβ family intracellular signaling pathways have been identified, and by consensus are now referred to as Smad1 through Smad8 (Shi and Massague 2003). As a general rule, Smads1, 5, and 8 are R-Smads

that function to transduce signals from the bone morphogenetic proteins (BMPs) and their specific type II (BMPR-II, ActR-II/B) and type I (Alk3, Alk6, Alk2) receptors (see Fig. 3).

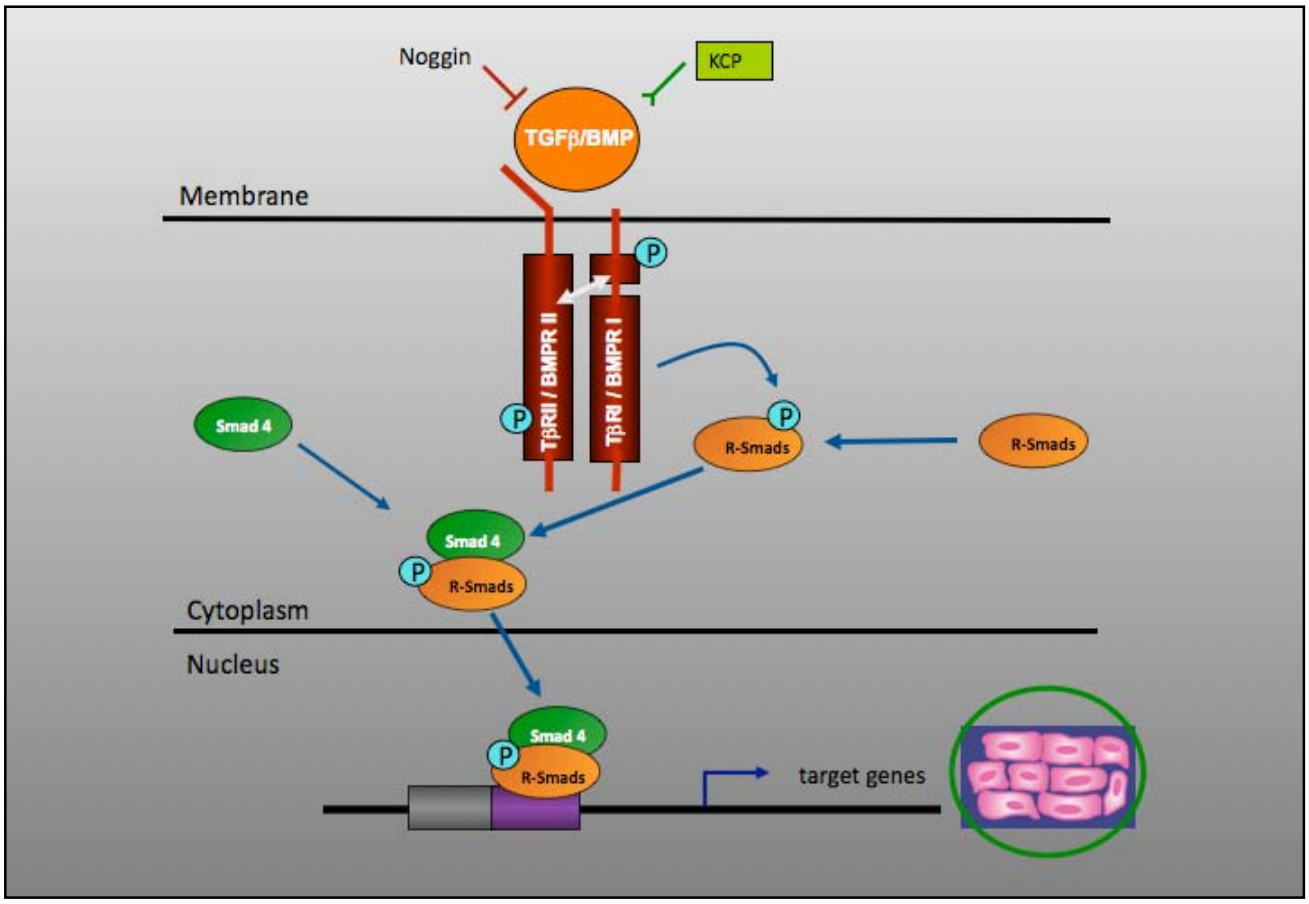


Figure 3*. Overview of TGFβ signaling*. Ligand interaction with type II receptors leads to activation of type I receptors. Type I receptors phosphorylate the receptor-associated Smads which complex with Smad4 and translocate to the nucleus to induce target gene transcripts. Ligand traps (e.g., Noggin) or agonists (e.g. KCP) can inhibit or potentiate signaling responses. Key: TGFβ = transforming growth factor beta; BMP=bone morphogenetic protein; P=phosphorylation event; RII = receptor type II; RI = receptor type I; R-Smads=receptor-associated Smads (Smads 1, 2, 3, 5 or 8); KCP: kielin-chordin like protein. *Adapted and re-drawn with permission from Ming Zhao M.D., Ph.D., Vanderbilt Medical Center.*

Smads2 and 3 are important R-Smad substrates of Alk5 that may be activated by either activin (through the type II ActR-IIB receptor) or by TGFβ selective type II receptors. Inhibitory Smads 6 and 7 (I-Smads) inhibit the signaling function of the receptor-activated Smads (not shown in Fig. 3). Smad6 preferentially inhibits BMP signaling, whereas Smad7 can inhibit both TGFβ and BMP signaling by preventing receptor-mediated phosphorylation of the receptor-activated Smad proteins (Shi and Massague 2003). When phosphorylated by the type I receptor, R-Smads associate with the common signaling intermediate, Smad4 (co-Smad4), which translocates the entire R-Smad:Co-Smad complex to the nucleus, where it associates with one or more of a number of DNA-binding partners and activates the transcription of specific target genes important for both growth inhibition and EMT. Steady-state levels of Smad proteins are regulated through the ubiquitin-proteasome degradation pathway (Massague, Seoane, and Wotton 2005).

TGFβ and Smad4 in Gastrointestinal Homeostasis and Cancer

The TGFβ family of peptides has a growth inhibitory role in intestinal epithelium. Kurokowa and colleagues (Kurokowa, Lynch, and Podolsky 1987) were the first to report that TGFβ was an inhibitor of cultured rat intestinal epithelial cells. It was subsequently determined that inhibition of cultured intestinal epithelial cell proliferation after TGFβ treatment results from mid-to-late G1 cell cycle arrest associated with down-regulation of cyclin D1 (Ko et al. 1995) and inhibition of Cdk4-associated retinoblastoma kinase activity (Ko et al. 1998). *In vivo*, there is increased expression of both TGFβ1 and type II TGFβ receptor (TβRII) in intestinal epithelial cells as they migrate from the proliferative compartment toward the lumen in both the small intestine and the colon (Barnard et al. 1989; Barnard, Warwick, and Gold 1993; Winesett, Ramsey, and Barnard 1996). This pattern of expression is inversely correlated with the mitotic activity in the gut epithelium (Zhang et al. 1997). Taken

together, these findings suggest that TGFβ plays a regulatory role in intestinal cell proliferation, and perhaps differentiation.

TGFβ regulation of epithelial cell proliferation is altered by cellular transformation. TGFβ inhibits growth of non-tumorigenic human colonic adenoma cells in culture; however, conversion of adenoma to a tumorigenic adenocarcinoma is associated with a decreased response to the inhibitory actions of TGFβ (Manning et al. 1991). Studies in human colon carcinoma cell lines have demonstrated a correlation between the differentiation state of tumors and sensitivity to the anti-proliferative and differentiation-promoting effects of TGFβ (Hoosein et al. 1989). Thus, loss of growth-inhibitory responses to TGFβ appears to be a common and important event that attends malignant transformation of epithelial cells.

One of the mechanisms by which tumor cells become resistant to the growth inhibitory actions of TGFβ may be through down-regulation or mutation of the TβRII. Several studies have suggested that a decrease in expression of TβRII is a key step for the neoplastic transformation of epithelial cells (Arteaga et al. 1988; Kimchi et al. 1988; Sun et al. 1994). Inactivation of the TβRII has been detected in a subgroup of colorectal carcinomas associated with the microsatellite instability or replication error phenotype found in approximately 13% of all colorectal cancers (Markowitz et al. 1995). Mutations of TβRII have also been identified in 15% of microsatellite stable colorectal cancers (Grady et al. 1999). Of potential importance are the observations that the subset of colorectal cancers that exhibit microsatellite instability (and TβRII mutations) tend to be proximal colon cancers and have a better prognosis (stage for stage) than the majority of sporadic colorectal cancers that do not share these genetic defects (Gryfe et al. 2000; Thibodeau, Bren, and Schaid 1993).

In addition to loss of the normal growth inhibitory processes mediated from mutations of the TβRII, mutations or epigenetic loss of expression of Smad signal transduction proteins also

contributes to tumorigenesis and tumor progression in the GI tract (Elliott and Blobe 2005). Mutations of Smad3 have not been identified in human cancers, but 37% of gastric carcinomas exhibit decreased Smad3 immunoreactivity and when Smad3 deficient gastric cancer cells are forced to express ectopic Smad3, TGFβ responsiveness is restored and xenograft growth in nude mice is suppressed in Smad3 transfected gastric cancer cells versus controls (Han et al. 2004). Smad2 mutations have been identified in a small subset (under 10%) of colorectal cancers (Eppert et al. 1996; Ohtaki et al. 2001).

The most commonly disrupted Smad mediator in cancers, including colorectal cancer, is Smad4. Mutations of Smad4 have been identified in 50% of pancreatic cancers (Hahn et al. Homozygous deletion map at 18q21.1 in pancreatic cancer 1996), 20-30% of colorectal cancer (Riggins et al. 1997; Riggins et al. 1996; Thiagalingam et al. 1996) and in 10-20% of small bowel adenocarcinomas (Blaker et al. 2004). Loss of Smad4 expression is correlated with loss of E-cadherin expression (Reinacher-Schick et al. 2004), liver metastasis and poor prognosis in colon cancer (Alazzouzi et al. 2005; Miyaki et al. 1999). Loss of Smad4 expression and deletion of chromosome 18q have recently been associated with increased incidence of lymph node metastasis in colorectal cancer (Tanaka et al. 2008). These reports indicate that *loss of Smad4* expression is an important contributing factor for tumorigenesis in colorectal cancer.

Bone Morphogenetic Proteins in the Intestinal Tract

The role of BMP subfamily ligands and receptors in the GI tract are less well known than that of the TGFβ subfamily, though recent experimental observations indicate their activity in growth, development and cancer. BMP-2 and its receptors are expressed in the mouse and human colon, predominantly in mature colonocytes at the epithelial surface. BMP-2 promotes apoptosis and growth inhibition in cultured colon cancer cells (Hardwick et al. 2004). Howe and colleagues have identified

germline mutations in the gene encoding the BMP receptor 1A in Juvenile Polyposis Syndrome (JPS) (Howe et al. 2001; Sayed et al. 2002). Of note, previous studies also implicated Smad4 mutations in a significant subset of JPS patients (Howe et al. 1998). Villin promoter-directed expression of Noggin, a natural BMP antagonist, in the intestine of mice as well as targeted disruption of the BMP signaling pathway through conditional inactivation of the *Bmpr1a* gene in mice resulted in similar histopathology as is seen in JPS (Haramis et al. 2004; He et al. 2004). Similarly, BMP-2 expression was found to be decreased in microadenomas of familial adenomatous polyposis (FAP) patients (Hardwick et al. 2004). These studies indicate that BMP signaling is critical for homeostasis in the intestinal tract, and loss of this function leads to pre-cancerous conditions such as JPS and FAP.

TGFβ Superfamily Signaling and EMT

In contrast with the tumor suppressive effects of Smad4 intrinsic to TGFβ and BMP signaling, additional data demonstrate that Smad4 may also regulate EMT, perhaps through TGFβ and/or BMP mediated signaling interactions with Ras/MAPK, PI3K or Wnt signaling pathways (Valcourt et al. 2005). Others have reported that TGFβ and BMP ligands and their respective ligand trap peptides have homeostatic opposing effects on EMT in some systems, with TGFβ promoting and BMPs inhibiting EMT (Zeisberg et al. 2003) and (reviewed in (Neilson 2005)). The balance between a differentiated epithelial phenotype and the more aggressive and invasive mesenchymal phenotype is also influenced by the abundance and activity of locally produced cytokines. Rees and colleagues recently reported evidence of EMT at the leading invasive front of esophageal adenocarcinomas that consisted of loss of membrane E-cadherin expression, increased α-SMA and vimentin, along with diffuse stromal TGF-β1 immunostaining (Rees et al. 2006). In contrast, BMP-7 immunostaining was absent at the invasive front with occasional cells showing increased immunoreactivity in the central

tumor areas. These data suggest an intricate balance in the epithelial cell microenvironment amongst members of the TGFβ superfamily to either prevent or promote EMT.

Wnt Signaling in Intestinal Growth, Differentiation and Colorectal Cancer

Wnt signaling is a critical regulator of embryonic development. The foundation for characterization of canonical Wnt signaling was laid by the discovery that the *Drosophila* segment polarity gene *Wingless* and the murine proto-oncogene *Int-1* were of a common origin (*Wg* + *Int-1* = Wnt) (Nusse and Varmus 1982; Rijsewijk et al. 1987). β-catenin-mediated canonical Wnt signaling is critical for maintenance of the intestinal epithelial stem cell compartment (Reya and Clevers 2005). In adult organisms, Wnt signaling regulates intestinal crypt progenitor cell compartments and the control of cell fate along the intestinal crypt-villus axis (Radtke and Clevers 2005). This pathway is highly conserved from invertebrates to vertebrate organisms and intracellular levels of the key signaling mediator, β-catenin, are tightly controlled. The differentiated epithelial cell goes to great lengths to restrict the levels of cytosolic β-catenin. The Wnt pathway is normally activated by Wnt ligand and receptor interactions in an embryologic or stem cell microenvironment and it is widely held that the central event in canonical Wnt signaling is cytoplasmic accumulation of β-catenin and its subsequent nuclear translocation and activity (Gordon and Nusse 2006). Although the non-canonical/Planar Cell Poloarity Wnt pathway may play an important role in epithelial pathobiology and cell movement in gastrulation (Veeman, Axelrod, and Moon 2003), it will not be discussed further in this work.

When Wnt signaling is activated, β-catenin levels transiently increase intracellularly and β-catenin translocates to the nucleus where it forms a transcriptional regulatory complex with the T cell-specific transcription factor/lymphoid enhancer-binding factor 1 (TCF/LEF) to activate transcription of *cyclin D1*, *c-MYC* and other downstream target genes (He et al. 1998; Tetsu and McCormick 1999).

In the absence of Wnt signaling, β-catenin levels in the cytoplasm are exceedingly low and the β-catenin pools are localized to the adherens junctions in a complex with E-cadherin (see Fig. 4, left panel). Normally, in the absence of a Wnt ligand, low cytoplasmic concentrations of β-catenin are maintained by the scaffold protein Axin. Axin coordinates the formation of a protein complex that consists of casein kinase 1α (CK1α), glycogen synthase kinase 3β (GSK3β), the tumor suppressor adenomatous polyposis coli (APC) and protein phosphatase 2A (PP2A) (Fig. 4, left panel, CK1α and PP2A *not shown*). Together, components of this complex act to constitutively phosphorylate β-catenin, leading to its recognition by the beta-transducin repeat containing (β-TrCP) subunit of the SCF ubiquitin ligase complex ($SCF^{\beta\text{-}TrCP}$), catalytic transfer of polyubiquitin chains to β-catenin and its rapid degradation by the 26S proteasome (Aberle et al. 1997; Ikeda et al. 1998; Kishida et al. 1998; Maniatis 1999).

Inactivating mutations in the *APC* (*adenomatous polyposis coli*) gene or activating mutations in the *β-catenin* gene (*ctnnb1*) result in failure to phosphorylate β-catenin with stabilization and accumulation of β-catenin in the cytoplasm and aberrant activation of Wnt signaling (Fig. 4, right panel). Hypophosphorylation of β-catenin leads to accumulation in the cytoplasm and translocation to the nucleus with resultant regulation of target gene expression with the TCF/LEF family of transcription factors (Gordon and Nusse 2006). The significance of the role of APC in colorectal epithelial homeostasis is revealed by the fact that over 80% of sporadic colorectal cancer cases have truncating mutations in the *APC* gene, which has often been considered the gatekeeper in the genesis of colorectal cancer because it is the earliest detectable genetic lesion in most premalignant colorectal adenomas. Aberrant stabilization of β-catenin in colorectal cancer can result either from inactivating APC mutation or stabilizing mutations in β-catenin (found in 10% of sporadic colon cancers).

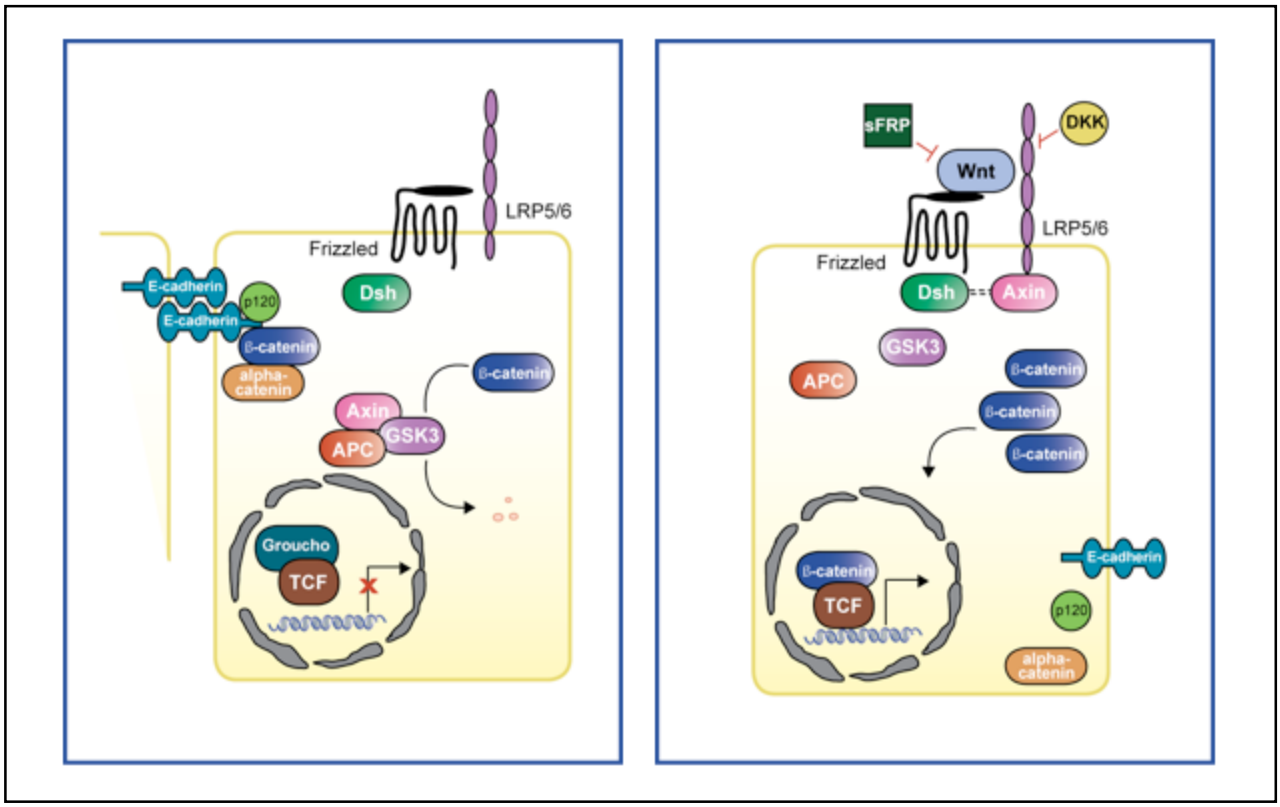


Figure 4. *Canonical Wnt signaling pathway*. In the absence of Wnt ligand, E-cadherin binds and sequesters β-catenin to the plasma cell membrane at the adherens junction (with p120 and alpha-catenin), rendering it unavailable for transcriptional activity (left panel). Wnt ligand interaction with the LRP receptor complex (Frizzled + LRP 5/6) can activate the pathway (right panel) and free cytoplasmic β-catenin levels increase. β-catenin translocates to the nucleus where it forms a transcriptional regulatory complex with TCF/LEF1 (TCF) to activate transcription of downstream targets (e.g., *c-MYC*). Key: LRP (LDL-receptor related protein), Dsh (Dishevelled), TCF/LEF1 (T-cell factor/lymphoid enhancer factor), sFRP (secreted Frizzled-related protein), DKK (Dickkopf), and GSK3 (glycogen synthase kinase 3β). *Adapted with permission from the publisher from Logan CY, Nusse R. The Wnt signaling pathway in development and disease. Annu Rev Cell Dev Biol. 2004; 20:781-810.*

Adherens Junctions, Wnt Signaling and EMT

Adherens junctions are specialized forms of cadherin-based adhesive contacts important for tissue organization in developing and adult organisms. E-cadherin is an epithelium-specific cadherin that forms protein complexes with cytoplasmic proteins (catenins) that convert the specific, homophilic-binding capacity of the extracellular domain into stable cell adhesion. Critical proteins found at the adherens junction include E-cadherin, α-catenin, β-catenin and p120 catenin (see Fig. 4, left panel). E-cadherin is a calcium-dependent adhesion protein. It is the major component of the adherens junction that facilitates cell-cell communication, and E-cadherin also functions as a tumor

suppressor (Christofori and Semb 1999) Junctional proteins may also serve important roles other than as structural components of intercellular junctions. E-cadherin binds and sequesters cytoplasmic β-catenin, rendering β-catenin unavailable for signaling in the canonical Wnt/β-catenin/TCF signaling cascade (Nelson and Nusse 2004). Therefore, E-cadherin like the APC protein is an important tumor suppressor protein. Both E-cadherin and APC have essential roles in preventing accumulation of cytoplasmic β-catenin, and thereby prevent inappropriate activation of the Wnt pathway (Christofori and Semb 1999).

Loss of E-cadherin in epithelial malignancies is associated with invasion, metastasis, and worse prognosis (Brabletz et al. 2001; Ikeguchi, Makino, and Kaibara 2001; Ikeguchi et al. 2000; Kanazawa et al. E-cadherin expression in the primary tumors and metastatic lymph nodes of poorly differentiated types of rectal cancer 2002; Kanazawa et al. Poorly differentiated adenocarcinoma and mucinous carcinoma of the colon and rectum show higher rates of loss of heterozygosity and loss of e-cadherin expression due to methylation of promoter region 2002). Loss of normal expression of E-cadherin may occur by any of several well-described mechanisms. These mechanisms may include transcriptional and post-translational regulation of E-cadherin membrane localization and stability. Reynolds and colleagues demonstrated that p120 catenin interaction with E-cadherin is essential for maintenance of E-cadherin stability and function as a tumor suppressor in epidermoid and colorectal cancer cells (Davis, Ireton, and Reynolds 2003; Ireton et al. 2002). Transcriptional repression is also a prominent regulatory mechanism by which E-cadherin expression may be suppressed in epithelial and colorectal cancer cells (Batlle et al. 2000). Two basic-helix-loop-helix (bHLH) proteins, E12/E47 (Perez-Moreno et al. 2001) and Twist (Yang et al. 2004) have been identified as interacting with the E-cadherin promoter E-box region to repress E-cadherin expression and to induce the EMT phenotype (Kang and Massague 2004). Even though it is known that E-cadherin can regulate β-

catenin levels and that E-cadherin itself is regulated via a transcriptional mechanism, the transcriptional regulation of β-catenin as it relates to colorectal cancer, Smad4 and EMT has not been described, but will be discussed in Chapter III.

Alteration of E-cadherin expression in experimental models has profound effects on the invasive and metastatic potential of human cancer cells. E-cadherin expression in human MDA MB 231 breast cancer cells suppressed the development of bone metastases in athymic (nude) mice (Mbalaviele et al. 1996). E-cadherin may interact with other Wnt target genes such as EphrinB2 to modulate tumorigenesis (Batlle et al. 2002; Cortina et al. 2007) and interestingly, we found EphrinB2 to be down-regulated in the metastatic gene expression profile found in Chapter II of this work. Suppression of the transcriptional repressor, SLUG (or SNAI2) by siRNA restores E-cadherin expression and partially reverts the transformed phenotype in Ras transformed rat intestinal epithelial cells (Schmidt et al. 2005). Recent work has shown that restoration of Smad4 in Smad4-deficient SW480 colon cancer cells induced E-cadherin expression while reducing total β-catenin protein levels, and these effects were accompanied by a marked decrease in β-catenin/TCF activity (Shiou et al. 2007). These data implicate Smad4 as a key modulator of both EMT and Wnt signaling in colon cancer cells and will be expanded upon further in Chapter III.

Tumor Biology at the Invasive Front

Increased nuclear β-catenin localization and Wnt signaling have been demonstrated at the tumor-stromal interface of the leading invasive edge of tumors and is associated with localized evidence of EMT (loss of E-cadherin at the cell membrane) and disruption of cell polarity (Brabletz et al. 2001). This group recently showed that well-to-moderately differentiated pT3 rectal adenocarcinomas express basement membrane proteins, but breakdown of the basement membrane in many of the rectal cancers occurs in discrete regions adjacent to cancer cells at the invasive front

of tumors (Spaderna et al. 2006). A provocative observation from this report was that the basement membrane was rebuilt in both lymph node and distant metastases, accompanied by the same glandular morphology observed in the majority of the primary tumor, including the predominance of the epithelial phenotype, but that the entire process was recapitulated in the invasive regions of the metastatic lesions. These observations support the notion that in most cases, EMT is a dynamic, regulated and reversible process. The notion that β-catenin may also play a prominent role at the invasive front of colorectal cancers will be introduced in Chapter III.

Other Contributing Pathways in the Pathogenesis of Colorectal Cancer

The Notch signaling cascade plays a role in intestinal homeostasis, cell fate decisions and differentiation (Radtke and Clevers 2005). This pathway contributes heavily to the development and maintenance of secretory and enteroendocrine lineages important for regulation of the intestinal epithelial milieu (Jenny et al. 2002; Jensen et al. 2000; Yang et al. 2001). Interestingly, a major target of the Notch signaling pathway, Hes1, is up-regulated in the 34-gene classifier as discussed in Chapter II. Other important growth factor pathways contribute to colorectal cancer and are worthy of mention. Activation of prostaglandins is an early step in the development of an adenoma (Markowitz 2007) and this response can be induced by inflammation or up-regulation of cyclooxygenase-2 (COX-2) that spurs the production of prostaglandin E_2 (PGE_2) and this action is strongly associated with colorectal cancer (Cha and DuBois 2007). Loss of 15-prostaglandin dehydrogenase promotes activity of PGE_2 and has been noted in the majority of colorectal adenomas and cancers (Yan et al. 2004). Interestingly, COX-2 is a Wnt target gene (Howe et al. 1999) and clinical trials have shown the

benefits of COX-2 inhibition in prevention and regression of adenomas (Bertagnolli et al. 2006). The potential role of inflammation in the metastatic process will be discussed further in Chapter II.

Epidermal growth factor has been noted to play an important role via the epidermal growth factor receptor (EGFR) in a subset of colorectal cancer patients (Saltz et al. 2004). Activation of EGFR can mediate up-regulation of prostaglandins (Coffey et al. 1997) in addition to up-regulation of MAPK and PI3K signaling pathways, all of which may play a role in lack of response to anti-EGFR therapies in advanced colorectal cancers (Wong and Cunningham 2008). Notably, prominent members of this pathway, EGFR and HBEGF were implicated in the metastasis-associated signature as described in Chapter II. Additionally, Vascular endothelial growth factor (VEGF) and its critical role in normal and transformative angiogenesis in tumor development and colorectal cancer pathogenesis have been well described (Ellis and Hicklin 2008). The role of anti-VEGF treatment appears promising even though patient selection by reliable molecular markers for this treatment remains difficult (Hurwitz et al. 2004). Finally, cancer stem cells play a prominent role in the development of colorectal cancer (O'Brien et al. 2007) and eventual metastasis (Boman and Huang 2008). Although the isolation of individual colon cancer stem cells is not currently feasible, certain cell-surface proteins hold promise as markers (e.g. CD133, CD44) (Yeung and Mortensen 2009; Zeilstra et al. 2008). The molecular mechanisms of colon cancer stem cell biology and treatment possibilities that would exist once these cells are readily identifiable hold great potential for the treatment and prevention of colorectal cancer. It is recognized that these additional pathways play important roles in the pathogenesis of colorectal cancer and may interact with molecules implicated in our 34-gene classifier and in the Wnt and TGFβ pathways; however, these topics are beyond the scope of this work and will not be addressed further.

Predictive and Prognostic Markers and Molecular Detection in Colorectal Cancer

An immense challenge to clinicians and biologists worldwide is the translation of colorectal cancer genomic data into a clinically useful and applicable prognostic test. As mentioned above, the complexities of EGFR signaling via RAS and BRAF along with anti-EGFR therapy have been applied with individual patient genomic status in mind and have led to the strong recommendation to have all patients with colorectal cancer genotyped for *KRAS* mutations (Wong and Cunningham 2008). Only a few markers can identify high-risk patients on a consistent basis and thereby be used to guide treatment decisions for colorectal cancer patients (e.g., *APC, MSH2*). Many other markers have not been tested prospectively and have unconfirmed utility as prognostic markers. It is well established that patients with microsatellite-high (MSI-H), sporadic colorectal cancer typically have favorable outcome (Samowitz et al. 2001) and recently loss of p27 has been associated with poor outcome in stage III colon cancers (Bertagnolli et al. 2009). Currently, loss of heterozygosity at chromosome 18q (Watanabe et al. 2001) is being evaluated prospectively in stage II and stage III patients for its value in selection of high-risk stage II and III colorectal cancers at resection.

Determination of high-risk colorectal cancers at the earliest stage possible is a primary goal for clinicians worldwide (e.g., at surgical resection or in a surveillance assay). Untargeted adjuvant chemotherapy in stage II patients produces a small absolute reduction in death at 5 years (Quasar Collaborative et al. 2007), but recent meta-analysis indicates that a sub-group of ‘high-risk’ stage II patients do benefit from such treatment (Figueredo, Coombes, and Mukherjee 2008). Conversely, 40-45% of stage III patients will not recur after surgical resection of their cancer (Ragnhammar et al. 2001), whereas 55-60% of stage III patients would die without adjuvant chemotherapy. Therefore, objective identification of high-risk stage II patients who would benefit from adjuvant chemotherapy

and selection of low-risk stage III patients who could be spared the cost, morbidity and potential mortality of systemic chemotherapy could prevent thousands of deaths each year in the US. To this end, use of molecular detection tools for aberrant DNA methylation in fecal DNA, plasma cell-free DNA and in resected tissues with RNA profiling are under intense investigation to uncover reliable 'high-risk' biomarkers (Imperiale et al. 2004; Jorissen et al. 2009; Li et al. 2009).

Biological Models of Colorectal Cancer

Multiple mouse models of colorectal cancer are in use and have been successful in aiding our understanding of colorectal cancer pathogenesis. For example, since it is known that inappropriate activation of the Wnt pathway is the key event in initiation of colonic polyps, the APC^{Min} mouse has been particularly informative on its own and with introduction of additional genetic lesions that potentiate the development of adenomas and/or adenocarcinomas (Moser, Pitot, and Dove 1990; Su et al. 1992; Takaku et al. 1998). Transplantation models (e.g. xenografts) are an important methodology to inform the biology of colon cancer cell tumorigenicity and metastasis and were used prior to the development of genetic mouse models (Heijstek, Kranenburg, and Borel Rinkes 2005; Kobaek-Larsen et al. 2000). There are few (if any) genetic mouse models of spontaneous tumor development that invade and then metastasize (Taketo and Edelmann 2009). The development of such a model is critical since metastasis is responsible for most of the mortality in colorectal cancer. In the meantime, xenograft models and colon cancer cell lines derived from mouse colon cancers (Corbett et al. 1975; Franks and Hemmings 1978) in addition to genetic mouse models are the workhorses for discovery in the biology and progression of colorectal cancers.

In this body of work, we demonstrate novel, *in vitro* and *in vivo* translational application to human disease in a mouse colon cancer cell line to the process of metastasis (see Chapter II) in addition to testing of human colon cancer cell lines in a xenograft model to compliment *in vitro* findings in exploring the process of tumor suppression (see Chapter III). Use of these biological models can still allow important discoveries even though a spontaneous model of colorectal tumorigenicity and metastasis is on the horizon. For example, we describe two strategies based upon biological models that may prove useful in detection of high-risk colorectal cancer patients with use of tissue available upon surgical resection to inform treatment decisions and guide novel biologic therapies in Chapters II and III of this work. First, we used a mouse model of colorectal cancer founded on the biology of metastasis to identify a gene expression signature that when applied to resected tumor tissues from colon cancer patients selected high-risk stage II patients and low-risk stage III patients. Secondly, we used an *in vitro* model of epithelial cells to uncover a novel mechanism whereby Smad4 represses the transcriptional activity of β-catenin and downstream Wnt signaling activity. This program of transcriptional repression by Smad4 led to the discovery of a gene expression signature that identified an additional subset of high-risk colorectal cancer patients.

Summary

Colorectal cancer is the second most lethal, non-cutaneous epithelial cancer in the United States. Metastasis contributes to the majority of cancer-related deaths in this disease. Metastasis closely resembles the developmental process of EMT and reliable models to recapitulate this process can be useful to shed light onto the biology and prognosis of patients with colorectal cancer. In Chapter 2, we describe an immunocompetent mouse model that led to the discovery of a gene signature of metastasis, which identified stage II and III colon cancer patients prone to recurrence and death from metastatic disease, in addition to a low-risk sub-group of stage III patients for whom adjuvant chemotherapy provided no additional survival benefit. These findings form the basis for substantive pre-clinical biomarker testing and eventual translational application to a clinical trial. This work was recently published (Smith et al.).

On the other hand, defects in tumor suppressor genes affect the majority of colorectal cancer patients. Many pathways are defective in the pathogenesis of colorectal cancer; however, more than half of the patients have defects in the Wnt/β-catenin and TGFβ signaling pathways which are critical for development and intestinal homeostasis. In Chapter 3, we use *in vitro* and *in vivo* models to describe a new role for the tumor suppressor Smad4 in the repression of β-catenin transcriptional activity in epithelial cells. This repression was associated with down-regulation of Wnt signaling and reversal of EMT. Clinical relevance of this effect was demonstrated by an epithelial cell-specific, Smad4-modulated gene expression profile associated with Wnt signaling suppression, which contributed prognostic information for colorectal cancer patients independently of pathological staging. This work will be submitted for publication immediately post-dissertation. These findings should facilitate hypothesis testing for biologically-targeted therapeutic interventions based on

TGFβ/Smad and Wnt/β-catenin pathway activity levels. Overall, these results provide insight into the biology of metastasis and tumor suppression in colorectal cancer to promote seamless translation to care of the colorectal cancer patient at the bedside.

CHAPTER II

AN EXPERIMENTALLY DERIVED METASTASIS GENE EXPRESSION PROFILE PREDICTS RECURRENCE AND DEATH IN COLON CANCER PATIENTS

Abstract

Staging inadequately predicts metastatic risk in colon cancer patients. We used a gene expression profile derived from invasive, murine colon cancer cells that were highly metastatic in an immunocompetent mouse model to identify colon cancer patients at risk for recurrence. This phase I, exploratory biomarker study used 55 colorectal cancer patients from Vanderbilt Medical Center (VMC) as the training dataset and 177 patients from the Moffitt Cancer Center as the independent dataset. The metastasis-associated gene expression profile developed from the mouse model was refined using comparative functional genomics in the VMC gene expression profiles to identify a 34-gene classifier associated with high risk of metastasis and death from colon cancer. A metastasis score derived from the biologically based classifier was tested in the Moffitt dataset. A high score was significantly associated with increased risk of metastasis and death from colon cancer across all pathological stages and specifically in stage II and stage III patients. The metastasis score was shown to independently predict risk of cancer recurrence and death in univariate and multivariate models. For example, among stage III patients, a high score translated to increased relative risk for cancer recurrence (hazard ratio = 4.7 [95% confidence interval=1.566-14.05]). Furthermore, the metastasis score identified stage III patients whose 5-year recurrence-free survival was >88% and for

whom adjuvant chemotherapy did not increase survival time. A gene expression profile identified from in an experimental model of colon cancer metastasis predicted cancer recurrence and death, independently of conventional measures, in patients with colon cancer.

Introduction

Colorectal carcinoma is the 3rd most commonly occurring non-cutaneous carcinoma and the 2nd leading cause of cancer-related death in the United States (Jemal et al. 2009). While it is well established that adjuvant chemotherapy for stage III colon cancer patients results in a survival benefit for the group, careful review of clinical trials data reveals that 40-44% of stage III patients enrolled in "surgery-only" groups did not recur in five years even without adjuvant treatment (Ragnhammar et al. 2001). Therefore, this subgroup of stage III patients, if prospectively identifiable, is unlikely to achieve benefit from and could be spared adjuvant chemotherapy. In addition, clinical trials have failed to demonstrate the benefit of adjuvant chemotherapy when applied to unselected patients with stage II colon cancer (Benson et al. 2004; Gill et al. 2004; Mamounas et al. 1999). On the other hand, some studies suggest that a subset of high-risk stage II colon cancer patients may benefit from adjuvant therapy (Figueredo et al. 2004; Figueredo, Coombes, and Mukherjee 2008; Quasar Collaborative et al. 2007). Much of these data for stage II patients come from meta-analyses and the question of whether adjuvant treatment really improves outcomes in stage II patients with "high-risk" features (e.g., T4 lesions, poorly differentiated histology or lymphovascular invasion) has not been answered in prospective clinical trials to date. Thus, an accurate and reliable method that identifies patients at greatest and least risk (e.g., "high-risk" stage II and "low-risk" stage III patients) could improve the selection of individualized therapy within these groups.

In this study, an experimental mouse model of metastasis was used to develop a gene expression profile that discriminated high versus low risk of cancer recurrence and death in colon cancer patients. This biologically based model identified distinct subsets of stage II and III colon cancer patients at greater risk of cancer recurrence and death. Using a metastasis score derived from the gene expression profile, we found that low metastasis score stage III patients did not gain significant benefit from adjuvant chemotherapy, suggesting that these patients could have been spared the potentially toxic and costly effects of these treatments. Conversely, high metastasis score stage III patients treated with adjuvant chemotherapy had markedly improved outcomes compared with high metastasis score patients who did not receive adjuvant chemotherapy. Thus, our biologically based gene expression profile provides a potential platform to facilitate selection of colon cancer patients who may benefit from adjuvant systemic therapy.

Materials and Methods

Cell Culture and Mouse Model Overview

MC-38 mouse adenocarcinoma cells were obtained from the American Type Culture Collection (ATCC, Manassas, VA) (Lafreniere and Rosenberg 1986). MC-38 cells were transfected with firefly luciferase gene (pGL3 basic, Promega, Madison, WI) and selected (G418, Invitrogen, Carlsbad, California). To enrich for invasive cells, 7.5×10^5 cells were seeded onto 6-well, 8.0 μM pore transwell polycarbonate membrane inserts (Costar, Cambridge, MA) coated with 2.5 mg/mL matrigel and incubated with serum-free DMEM in the upper chamber and complete DMEM in the bottom well. After 12 hours, cells that invaded were aseptically harvested by brief, gentle trypsinization and

transferred to new dishes (Poste, Doll, and Fidler 1981) and six serial passages through matrigel-coated Boyden chambers ensued. The selected invasive cells and the parental luciferase-expressing MC-38 cells were injected into the tail vein and development of lung metastases was assessed (Fidler and Nicolson 1976). Development of metastases was determined by bioluminescence imaging (Jenkins et al. 2003; Wu et al. 2001). The MC-38met cells were derived by culture of tumor cells from a metastatic lung tumor (Whitehead 1976). The Vanderbilt Institutional Animal Care and Use Committee approved all animal work.

Detailed Cell Culture and Mouse Methods

MC-38 mouse adenocarcinoma cells were obtained from ATCC and cultured. MC-38 cells were transfected with firefly luciferase gene in pGL3 basic (Promega, Madison, WI) and selected in 0.5 mg/mL G418 (Invitrogen, Carlsbad, California) and referred to as MC-38 parentals. To enrich for invasive MC-38 parental cells, 7.5×10^5 cells were seeded onto 6-well, 8.0 μM pore transwell polycarbonate membrane inserts (Costar, Cambridge, MA) coated with 2.5 mg/mL matrigel and incubated with serum-free DMEM in the upper chamber and complete DMEM in the bottom well. Cells collected after the 6^{th} passage were designated "MC-38inv". To quantify the enrichment of invasive cells, cells were again incubated in matrigel-coated transwell filters for 24 hours at 37°C, 5% CO_2. Filters were washed and cells on the upper surface were removed with cotton swabs. The cells that had invaded were fixed in 4% paraformaldehyde for 10 minutes and stained with 1% crystal violet. Four random fields were counted to determine the numbers of invaded cells. To determine metastatic potential *in vivo*, equivalent numbers of MC-38 parental or MC-38inv cells were injected into the tail veins of C57BL/6 mice and the mice were followed for development of pulmonary metastases. Briefly, unanesthetized C57BL/6 mice were warmed with a heat lamp to allow for venous dilation. Mice were then placed into a plastic restraining apparatus and 2.5×10^5 MC-38

parental, MC-38inv or MC-38met cells were injected via lateral tail vein (n = 10). Successful injection and growth of metastatic lung nodules was confirmed by immediate and then weekly bioluminescence imaging (BLI). MC-38 *in vivo* metastatic cells ("MC-38met") were derived from metastatic lung nodules following washing and sterilization of fresh tumor tissue in 0.04% sodium hypochlorite. The tissue was next rinsed in PBSA, minced into approximately $1mm^3$ pieces, collected and resuspended in collagenase / neutral protease solution for 90 minutes at 37°C with occasional shaking. Remaining tissue pieces were allowed to settle, and were then re-suspended in fresh collagenase / neutral protease solution and incubated at 4°C overnight. The single-cell suspension was transferred to a new tube, centrifuged at 100g for 5 minutes, washed once with growth medium and then re-suspended in growth medium and plated onto collagen-coated 25cm flasks and 24-well dishes. Cells were grown at 37°C with 5% CO_2 under G418 selection. The invasive phenotype of the MC-38met cells was confirmed both *in vitro* and *in vivo*. For the splenic assay, 2.5×10^5 cells were implanted under the splenic capsule 2-3 minutes prior to splenectomy and metastatic foci derived from resulting circulating tumor cells were followed as they formed in the liver by BLI. This work was done by Fei Wu, M.D., Ph.D. in the Beauchamp laboratory. The work that follows was performed by J. Joshua Smith, M.D.

Human Tissue and Microarray Platforms

The protocols and procedures for this study were approved by the Institutional Review Boards at the University of Alabama-Birmingham Medical Center, Vanderbilt Medical Center (VMC), the Veterans Administration Hospital (Nashville, TN) and the H. Lee Moffitt Cancer Center (MCC, Tampa, FL). Representative sections of fresh tissue specimens were flash frozen in liquid nitrogen and stored at -80ºC until RNA isolation. Quality assessment slides were obtained to verify the diagnosis. Stage was assessed by American Joint Commission on Cancer (AJCC) guidelines. RNA was purified using the RNeasy® kit (Qiagen, Valencia, CA). Mouse and human samples were hybridized to Affymetrix arrays (Mouse Genome 430 2.0 GeneChip Expression and Human Genome U133 Plus 2.0 GeneChip Expression Arrays, respectively).

Statistical Methods and Identification of VMC High-Risk Patients

Microarray data for the metastatic MC-38met cells and the parental MC-38 cells and the human data were processed using the Robust MultiChip Analysis (Irizarry et al. 2003) algorithm as implemented by Bioconductor (pmid: 12582260). Mouse probe set identifiers (IDs) were mapped to Ensembl Gene IDs based on the mapping provided by Ensembl V49 (http://www.ensembl.org). Median expression levels from multiple probe sets corresponding to the same gene were calculated. Mouse genes with one-to-one human ortholog mapping as annotated by Ensembl V49 were carried forward for differential expression analysis using the limma package in Bioconductor (Smyth, Michaud, and Scott 2005) The 300-gene metastasis-associated signature was determined with the limma package in Bioconductor (Smyth, Michaud, and Scott 2005) based upon 3 criteria: (1) Fold

change >2; (2) False discovery rate (FDR) based on the moderated t-test followed by Benjamini and Hochberg's multiple-test adjustment <0.01; and (3) Log odds ratio of differential expression (B-statistic) > 1. This analysis resulted in a subset of 300 differentially expressed genes that could be unambiguously mapped to human orthologs. Gene expression datasets were separately standardized such that each gene had a mean expression value of 0 and a standard deviation of 1 across samples in a dataset. A direction of expression was assigned for each gene in each sample based on the sign of the standardized value. Directional concordance between MC-38met cells and 19 patients from VMC with poor outcome (17 stage IV patients and 2 stage III patients who developed metastatic recurrence (n=19)) was determined (exact binomial test, P≤.10) to refine the metastasis-associated signature to the 34-gene recurrence classifier. Specifically, each gene of the murine training-set was examined in the VMC patient-derived training set in patients with stage III or stage IV disease who had experienced metastasis or cancer-related death. Nineteen patients fell into these categories and the median expression level of each of the 300 genes was examined for concordance with MC-38met median expression. The genes with concordance (+1 or -1) in at least 13 of 19 patients (Exact binomial test, P≤0.10) were selected for the putative recurrence profile. The concordance analysis resulted in a 34-gene profile that was termed the '34-gene metastasis score' or more simply the '34-gene recurrence classifier'. Integration of mouse and human microarray datasets was conducted (Lee and Thorgeirsson 2004). Clustering (Pearson's correlation coefficient) was applied to the integrated data set as follows: average linkage clustering based on Pearson's correlation coefficient was applied to the integrated data set. The ceiling was set to cover 95% of all data points (e.g., the top 5% z-scores were truncated). The corresponding z-score is 2.00 for the 34-gene classifier.

Clinical Outcomes and Testing of the 34-gene Based Metastasis Score

The association between individual gene expression level and clinical endpoint (overall survival (OS), disease-specific survival (DSS) and disease-free survival (DFS)) were first analyzed using a Cox Proportional Hazard (PH) model (Beer et al. 2002; Hedenfalk et al. 2001; Tukey 1993; Yanagisawa et al. 2003) and definitions for each are provided below (Vanderbilt Cox Model (Training Dataset)). A compound score was calculated for each patient by summation of a Cox PH weighted sum (Wald score) of log-2 gene expression for each gene in the classifier (Yanagisawa et al. 2003). The compound score was used to measure the impact of the 34-gene classifier on survival (OS and DSS) and recurrence (DFS). A Vanderbilt-derived Cox model and re-sampling Wald tests were utilized to rule out over-fitting of the model. The estimated coefficients of the Cox PH model in the VMC data (n=55 patients) were applied to the MCC dataset (n=177 patients). In addition, a censored C-index (Harrell et al. 1982) was computed to validate the predictive value of the identified expression classifier. The compound scores were used in a univariate analysis (e.g., log-rank test) to segregate patients into higher than median and lower than median compound score groups. The compound score, age, gender and tumor grade were adjusted in the multivariate Cox model for both DSS and DFS. Adjusted P-values as well as the adjusted 95% confidence intervals of the hazard ratios from the Cox model were reported. Fisher's exact test was used for the T4, histological, microsatellite instability status and the stage III metastasis score/chemotherapy analyses.

Vanderbilt Cox Model (Training Dataset)

The association between individual gene expression level and clinical endpoint, (e.g., overall survival (OS) or disease-specific survival (DSS)), was first analyzed using a Cox Proportional Hazard (PH) model. Overall-survival was defined as death from any cause. Disease-specific survival was

defined as a documented cancer-related death. A disease-free survival event was defined as incidence of recurrence (~75% distant and ~25% local) after R0 resection. Sixty Affymetrix probe sets can be found on the microarray HG U133 Plus 2.0 platform based on the thirty-four classifier genes (e.g., multiple probes for each gene were found that mapped to each of the 34 genes). Their expression data were extracted from the VMC dataset (n=55). Expression data for each Affymetrix probe set were treated as the independent variable, and the Cox proportional hazard model was used for survival analyses. A compound score was next calculated for each patient by obtaining a Cox PH weighted sum (Wald score) of $\log_2$ gene expression. The compound score for patient *i* is defined as $\Sigma_i W_i * X_{ii}$ (where W_i=Wald statistic score for gene *i* and X_{ii} = $\log_2$ gene *i* * expression level of patient *i*). Finally, the compound score was used to measure the impact of the 34-gene profile on survival and a censored c-index was computed for the validation of the predictive value of the expression profile identified. The c-index is a probability of concordance between predicted and observed survival, with c = 0.5 for random predictions, and c = 1 for a perfectly discriminating model. The univariate analysis was completed using the compound scores to segregate patients into higher than median and lower than median compound score groups by Kaplan-Meier estimates to determine differences in survival by the log-rank statistic. The compound score, age gender, and grade were adjusted in the multivariate Cox model for overall, disease-specific and disease-free survival. The adjusted P-values as well as the adjusted 95% confidence intervals of the hazard ratios from the Cox model were reported.

Moffitt Cox Model (Test Dataset)

Sixty Affymetrix probe sets can be found on the microarray HG U133 Plus 2.0 platform based on the thirty-four classifier genes. Their expression data were extracted from the MCC dataset (n=177).

Then we applied the estimated coefficients of the Cox PH model from the training set (n=55) to the MCC dataset (n=177).

Distribution of re-sampling Wald tests with the 34-gene recurrence classifier

Sixty Affymetrix probe sets can be found on the microarray HG U133 Plus 2.0 platform based on the thirty-four classifier genes. Their expression data were extracted from the MCC dataset. Expression data for each Affymetrix probe set were treated as the independent variable, and the Cox proportional hazard model was used for survival analyses. Beta and Wald statistics for each Affymetrix probe set were used along with expression data to build up a compound score for each patient. The compound score was used as the independent variable to perform overall survival analysis based on the Cox model. The Wald test P value was saved as the observed P-value. For the re-sampling test, we randomly chose 60 Affymetrix probe sets from the 54675 sets on the whole array. We repeated the above procedure and generated one re-sampling Wald test P-value from the overall Cox model survival analysis. We repeated the re-sampling and survival analysis procedure 10,000 times, generating 10,000 re-sampling Wald test P-values. We transformed both the observed and re-sampling P-values into $\log_{10}$ format, plotted a histogram of the 10,000 re-sampling $\log_{10}$ (P-values), and added the observed $\log_{10}$ (P-value).

Cox modeling of Disease-specific survival for relative risk according to percentile score

Percentiles across high-score patients were plotted related to relative risk. Hazard ratios for 50^{th}, 75^{th} and 90^{th} percentiles were plotted as compared to the 10^{th} percentile.

Histologic Analysis and Microsatellite Instability in Association with Metastasis score

Fisher's exact test for a 2 x 3 table was used to determine association between histology and metastasis score. Fisher's exact test was used for the metastasis score, T4 and histological analyses

in the stage II and III patients in the MCC dataset in addition to the analysis of the subgroup of microsatellite status patients for the combined VMC and MCC datasets.

Pathway Analysis

Data were analyzed through the use of Ingenuity Pathways Analysis (IPA) and networks were generated with the use of IPA (Ingenuity Systems, www.ingenuity.com). A dataset containing gene identifiers and corresponding expression values was uploaded into the application. Each gene identifier was mapped to its corresponding gene object in the Ingenuity Pathways Knowledge Base. The cutoffs described in the statistical sections were used to identify genes whose expression was significantly differentially regulated. These genes, called focus genes in IPA, were overlaid onto a global molecular network developed from information contained in the Ingenuity Pathways Knowledge Base. Networks of these focus genes were then algorithmically generated based on their connectivity. The Functional Analysis identified the biological functions and/or disease that were most significant in the data set. Fisher's exact test was used to calculate a P-value determining the probability that each biological function and/or disease assigned to the data set is due to chance alone.

Results

Development of an Immunocompetent Mouse Model of Colon Cancer Metastasis

Tumors are a heterogeneous mixture of cells with varied invasive and metastatic potential. Therefore, we used a conventional invasion assay to enrich for a sub-population of highly invasive MC-38 mouse colon cancer cells (Fig. 5A, MC-38inv). Following six serial passages through matrigel, MC-38inv cells were more invasive than MC-38 parental cells both *in vitro* and *in vivo* in a tail vein injection assay (Fig. 5B and Table 1, $P<.001$). Lung tumors derived from MC-38inv cells were cultured to derive a highly metastatic cell line, MC-38met. MC-38met cells were injected into the tail vein and spleen and produced extensive metastatic tumors in the lung and liver respectively (see Fig. 6 and Table 2). Please note that the results from Chapter II are published (Smith et al.).

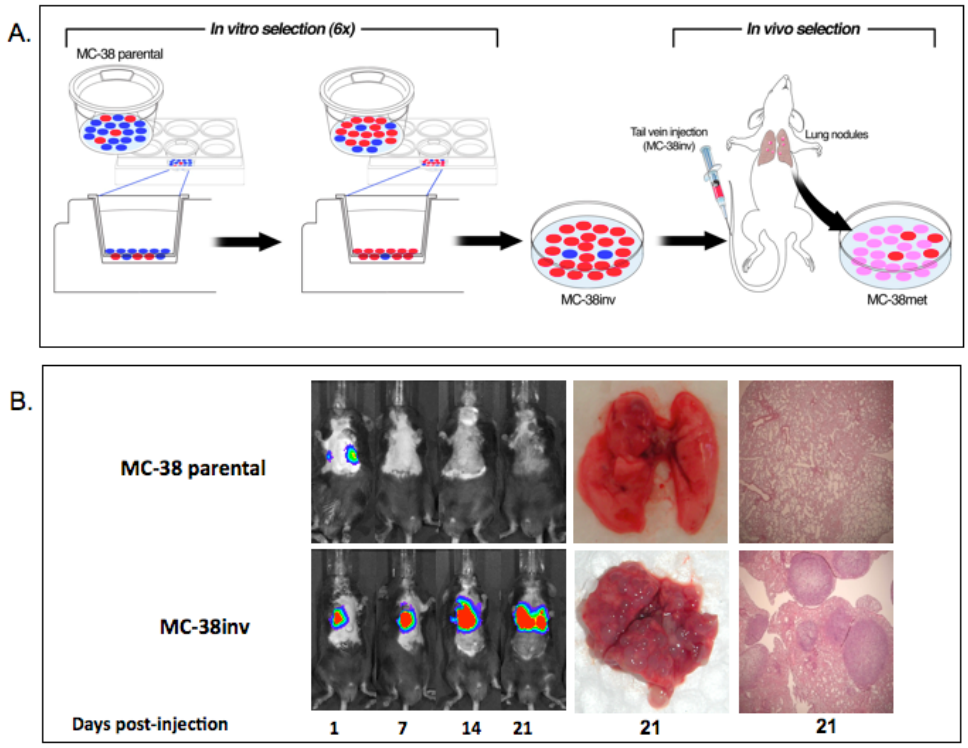


Figure 5. *Cell Culture and Mouse Model: murine model of metastasis, in vivo monitoring and ex vivo proof of metastases.* (A) MC-38 parental cells (heterogeneous: blue and red) were subjected to six sequential passages through matrigel-coated transwells (enrichment of invasive subpopulations of MC-38 cells (*red*)) called "MC-38inv". After *in vivo* passage, a stabilized cell line (*pink cells*) called "MC-38met" was established. (B) MC-38inv cells were tested alongside MC-38 parental cells for the ability to form lung metastasis in a tail vein assay. The figure shows representative tumor progression in live (C57BL/6) mice by bioluminescent imaging (days 1-21) and at the time of autopsy (day 21).

TABLE 1

QUANTIFICATION OF LUNG NODULES FROM MC-38 PARENTAL AND MC-38INV CELLS

	MC-38 parental (n=10)	**MC-38inv (n=10)**	**P-value**
Mean nodule number (median, s.d.)	1.7 (1.0, 1.89)	102.3 (102.0, 24.7)	<.001

MC-38 parental and MC-38inv cells, derived as described in Figure 4 and Materials and Methods, were injected into the tail veins of C57BL/6 mice. 21 days post-injection, mice were humanely killed, and lung nodules were counted at necropsy. Summary statistics for the data and results are displayed (Mann-Whitney U test, median, standard deviation (s.d.)).

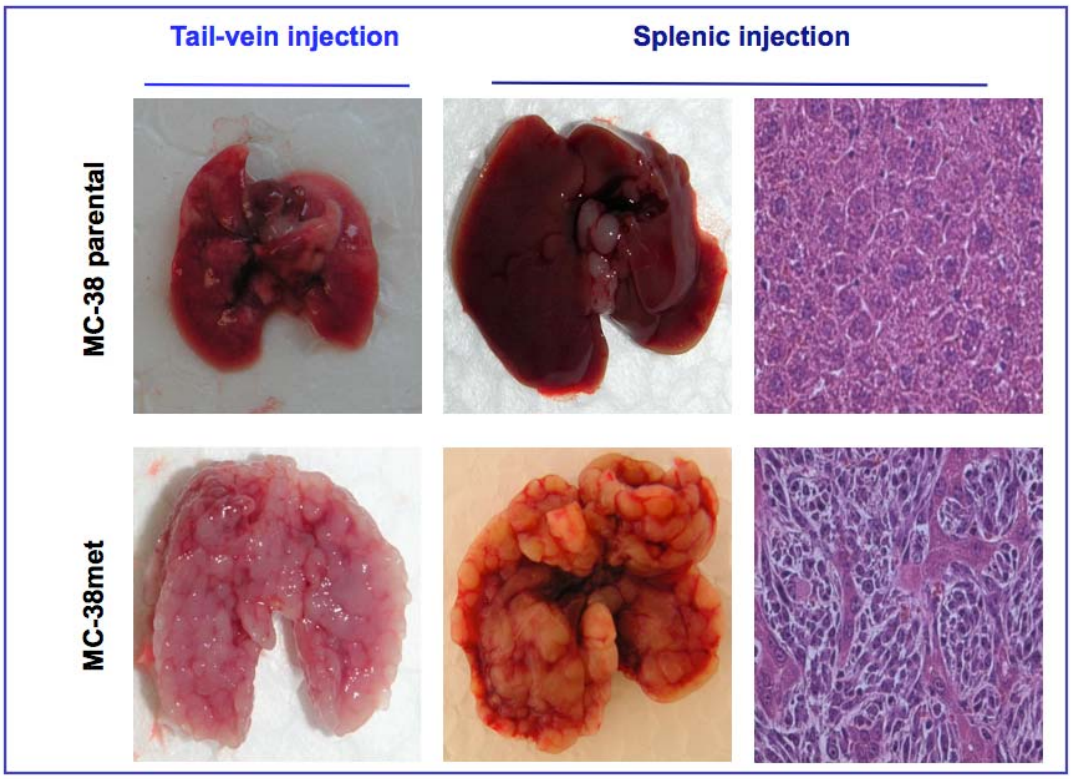


Figure 6. *MC-38 parental and MC-38met gross and histopathology in tail vein and splenic assays***.** Gross and representative histologic sections from lungs and livers of MC-38 parental (*upper panel*) and MC-38met (*lower panel*) injected mice are shown. For MC-38 parental-injected mice, minimal lung nodules (<10) and no liver metastases were noted. Prolific liver metastases (>200) were found in the MC-38met injected mice.

TABLE 2

HEPATIC AND LUNG METASTASES (SPLENIC AND TAIL VEIN MODELS): QUANTIFIED NECROPSY RESULTS

Hepatic metastases	MC-38 parental (n=8)	MC-38met (n=6)	P-value
Mean liver weight in grams (median, s.d.)	1.46 (1.45, 0.13)	5.37 (6.5, 2.27)	.002
Incidence (%)	0/8 (0%)	6/6 (100%)	<.001

Lung metastases	MC-38 parental (n=6)	MC-38met (n=5)	P-value
Mean lung weight in grams (median, s.d.)	0.32 (0.30, 0.08)	1.34 (1.5, 0.37)	.005

Quantification of the incidence of liver metastasis in the splenic assay and mean liver weights taken at necropsy with summary statistics are shown (Mann-Whitney U test (SPSS, Version 16; SPSS Inc, Chicago, IL). Median and standard deviation (s.d.) data are displayed. Of note, in regard to MC-38met liver metastasis, 2 mice died prematurely of massive liver metastasis, and analysis was done on the 6 mice that survived to the end of the 3-week experiment. Quantification of lung weights from the tail vein assay is also shown in the lower panel of the table.

Discovery of a Gene Expression Profile Associated with Metastasis: Mouse to Man

A flow diagram of the derivation of the metastatic gene expression signature and its refinement and testing is provided in Figure 7A. MC-38 parental and MC-38met cell mRNA expression profiles were examined by microarray and directly compared to evaluate the gene expression changes associated with invasion and metastasis. Gene elements from this microarray profile were mapped to 11,465 corresponding human orthologs. The initial cluster analysis and the 300 gene list are shown in Figure 8 and Table 3 (Appendix**)**. In order to refine the signature with relevance to cancer recurrence, each of these 300 genes was scored for directional concordance (see Methods) with gene expression data from 19 “high-risk” VMC colon cancer patients that either had metastatic disease or had died from cancer progression (see Methods). Thirty-four genes (Table 4) from the metastatic gene signature exhibited directional concordance in 13 of the 19 “high-risk” patients. Since the 300 genes had been selected with stringent statistical criteria, we did not further attempt to determine the minimum number of genes that could discriminate outcomes.

The resultant 34-gene expression pattern was designated the recurrence gene classifier and was next integrated with the entire VMC human microarray dataset and subjected to unsupervised cluster analysis. The recurrence gene classifier separated patients into two clusters (Fig. 7B). The cluster associated with MC-38met cells (cluster 2) contained 17 of the 19 “high-risk” patients used in the signature refinement process.

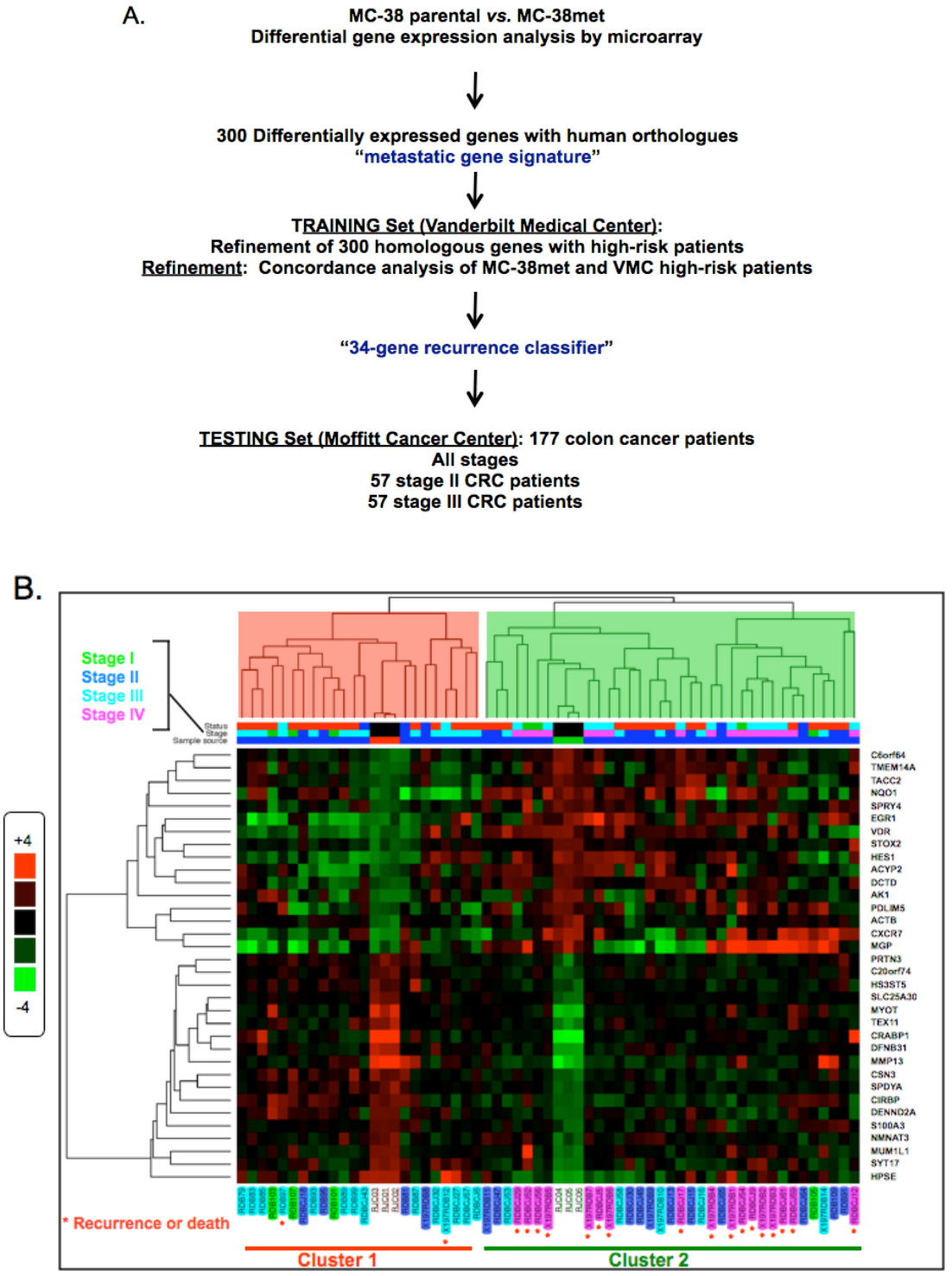


Figure 7. *A) Recurrence classifier development*. VMC two-step schematic for enrichment and establishment of the 34-gene recurrence classifier for colon cancer. Mouse genes were mapped to human orthologs and 300 differentially expressed genes (MC-38 parental vs. MC-38met) were identified. These 300 genes were next refined with 19 "high-risk" patients from the VMC training dataset for concordance. This analysis revealed 34 genes with concordant expression among the 19 high-risk patients and the MC-38met cells. The 34-gene recurrence classifier was then applied to the independent MCC database to determine whether it could be used to discriminate patients on the basis of outcomes. B) *Functional genomic cluster analysis of the 34-gene recurrence classifier.* Mean-centered gene expression data (rows) clustered with individual VMC patients (columns) results in two distinct patient groups (cluster 1, *pink*; cluster 2, *green*). The 19 VMC "high-risk" patients used in the concordance analysis are marked with a *red asterisk*.

TABLE 4

THE 34-GENE RECURRENCE CLASSIFIER

Up-regulated Genes			
Gene symbol	**Protein/Official Name**	**Fold change**	**Processes/Networks**
CXCR7	chemokine (C-X-C motif) receptor 7	3.9	Cancer, survival/growth and chemotaxis
AK1	adenylate kinase 1	2.8	Nucleotide binding
ACTB	actin, beta	2.8	Cancer, cell morphology and motility, growth, polarization and adhesion
MGP	matrix Gla protein	2.8	Cell-cell Signaling, branching, migration
HES1	hairy and enhancer of split 1, (Drosophila)	2.8	Cancer, endocrine function, cell death
TMEM14A	transmembrane protein 14A	2.6	Cell proliferation (target of CREB)
EGR1	early growth response 1	2.5	Cancer, endocrine function, cell death
VDR	vitamin D (1,25- dihydroxyvitamin D3) receptor	2.4	Cancer, endocrine function, cell death
C6orf64	chromosome 6 open reading frame 64	2.4	Membrane dynamics
NQO1	NAD(P)H dehydrogenase, quinone 1	2.4	Cancer, cell death
STOX2	storkhead box 2	2.3	Putative stem cell marker
ACYP2	acylphosphatase 2, muscle type	2.2	Mutated in aromatic rice
SPRY4	sprouty homolog 4 (Drosophila)	2.1	Cancer; cell migration, proliferation, differentiation
DCTD	dCMP deaminase	2.1	Nucleotide biosynthesis
TACC2	transforming, acidic coiled-coil containing protein 2	2.1	Cancer; biogenesis, morphology, proliferation
PDLIM5	PDZ and LIM domain 5	2.0	Cancer, actin binding, cytoskeleton organization

Down-regulated genes			
Gene symbol	**Protein/Official Name**	**Fold change**	**Processes/Networks**
CRABP1	cellular retinoic acid binding protein 1	-21.0	Cancer, CpG Island Methylation
MMP13	matrix metallopeptidase 13 (collagenase 3)	-17.3	Cell-cell signaling, immune response
MYOT	myotilin	-7.1	Cancer, actin filaments and stress fibers
DFNB31	deafness, autosomal recessive 31	-5.1	Cell-cell Signaling, actin cytoskeletal assembly
HPSE	heparanase	-4.7	Cell-cell signaling, immune response
TEX11	testis expressed 11	-3.7	Cell cycle/division
SYT17	synaptotagmin XVII	-3.1	Membrane protein
MUM1L1	melanoma associated antigen (mutated) 1-like 1	-2.8	Unknown
SLC25A30	solute carrier family 25, member 30	-2.6	Oxidative stress
CSN3	casein kappa	-2.5	Cell cycle, membrane dynamics, apoptosis
NMNAT3	nicotinamide nucleotide adenylyltransferase 3	-2.4	Nucleotide biosynthesis
DENND2A	DENN/MADD domain containing 2A	-2.4	Unknown
CIRBP	cold inducible RNA binding protein	-2.3	Cancer, nucleotide binding
SPDYA	speedy homolog A (Xenopus laevis)	-2.3	Putative cell cycle, putative biomarker in HCC
S100A3	S100 calcium binding protein A3	-2.2	Cancer
PRTN3	proteinase 3	-2.1	Cell-cell signaling, immune response
C20orf74 (AKA RALGAPA2)	Ral GTPase activating protein, alpha subunit 2 (catalytic)	-2.1	GTPase regulation
HS3ST5	heparan sulfate (glucosamine) 3-O-sulfotransferase 5	-2.0	Putative epigenetic regulation

Genes that were up-regulated or down-regulated in both MC-38met derived cells and in 19 VMC patients with poor-prognosis are shown. Gene symbols and associated fold-change in gene expression for MC-38 parental versus MC-38met as determined by microarray are given. Known relationships (connectedness) between up-regulated and down-regulated genes were determined independently and, where networks existed, functional enrichment within the networks was determined in Ingenuity Pathways Analysis (www.ingenuity.com), PubMed and on the Affymetrix website (see Materials and Methods). Ensembl protein identifiers are also noted in Table 3 (Appendix, p. 104).

The Recurrence Classifier Identifies Poor Outcome Colon Cancer Patients in an Independent Colon Cancer Dataset

An independent human colon cancer gene expression and clinical database from the H. Lee Moffitt Cancer Center (MCC) was used to test the ability of the recurrence classifier to discriminate patients at increased risk of cancer recurrence and death. The demographics for the training (VMC) and test set (MCC) are shown in Table 5. 205 patients were available for analysis in the MCC group and 195 of these patients had complete demographic, histological grade, stage and differentiation information. Of this 195 patient group, we focused on colon cancer patients (n=177) to avoid potential confounding effects of neoadjuvant and radiation therapy in rectal cancer patient tumor samples and outcomes. A method for weighted scoring of the gene expression pattern for the recurrence classifier was applied (see Methods) and metastasis scores were created. Patients were segregated into higher and lower than median metastasis score groups and survival analysis was performed between the two groups. To rule out over-fitting of the model, three separate statistical approaches were applied. First, we developed the metastasis score based on the Cox Proportional Hazard (PH) model in the VMC data (n=55 patients) and then applied the estimated coefficients of the Cox PH model from VMC data to the MCC dataset (n=177 patients). Thus, the sign and magnitude of the coefficient was completely based on the training dataset (VMC) and then tested in the MCC dataset. High metastasis score colon cancer patients across all stages in the MCC dataset had significantly worse overall and disease-specific survival compared with low metastasis score patients (Fig. 9A and B, P=.003 and P=.04 respectively). Second, multiple permutation testing was performed using the 34-gene recurrence classifier and the metastasis score was also robust in this model (see Fig. 10, Appendix). Third, a c-index was calculated to validate the predictive value of the 34-gene metastasis

score. The c-index is a probability of concordance between predicted and observed survival, with c = 0.5 for random predictions, and c = 1 for a perfectly discriminating model.

In order to determine whether the metastasis score could be used to identify stage II and III patients at high risk for recurrence, we tested it on stage II and III patients in the MCC dataset. Low metastasis score patients in each group (stage II alone, Fig. 11A-B and stage III alone, Fig. 11C-D) demonstrated significantly better outcomes than high metastasis score patients. As can be seen in the figure, this finding held true in both analyses using the endpoints of disease-specific survival (cancer-related death, panels A, C) as well as disease-free survival (recurrence, panels B, D).

Notably, there were no cancer-related deaths and only one recurrence event in low metastasis score stage II patients (see Fig. 11A, B). At five years, 31% of the high metastasis score stage II patients had died of cancer versus none of the low score patients. For the stage III patients, the five-year mortality rate was 10.7% for low metastasis score patients as compared with 37.9% for the high metastasis score patients (see Fig. 11C). The median survival time for stage III patients with a high metastasis score was 29.4 months. In sharp contrast, the low metastasis score stage III patients as a group faired so well that this group had not reached the threshold to calculate their median survival time. We noted no association with either the high or the low metastasis score on available data regarding T4 lesions, lymph nodes retrieved, histological grade or microsatellite instability status in the MCC dataset (See details, pp. 44-45). These data show that the 34-gene metastasis score can discriminate stage II and III colon cancer patients that have a low- or high-risk of cancer recurrence and death.

TABLE 5

STUDY DEMOGRAPHICS

	VMC-training n=55	*MCC-testing n=177*
Mean Age (s.d.)	62.3 (14.1)	65.5 (13.1)
Sex (%male)	30 (54.5%)	96 (54.2%)
Stage I	4 (7.3%)	24 (13.6%)
Stage II	15 (27.3%)	57 (32.2%)
Stage III	19 (34.5%)	57 (32.2%)
Stage IV	17 (30.9%)	39 (22%)
Median Follow-up Months (Min - Max)	50.2 (0.4 - 111.3)	48.1 (0.92 - 142.6)
Number of deaths	20 (36.3%)	73 (41.2%)
Caucasian (%)	50 (90.9%)	151 (85.3%)
Black (%)	4 (7.3%)	9 (5.1%)
Other (%)	1 (1.8%)	17 (9.6%)

All patients were diagnosed with colorectal adenocarcinoma (stages I-IV) according to current American Joint Commission on Cancer (AJCC) guidelines. On hundred seventy-seven of 205 MCC patients met the criteria of having AJCC stage I-IV colon cancer as well as complete demographic, histologic grade, stage and differentiation information. The Vanderbilt training set includes 14 patients from the University of Alabama-Birmingham Medical Center (tumors provided by M.J.H). *Other* in the VMC medical record implies not otherwise specified and it implies Hispanic or not otherwise specified in the MCC medical record.

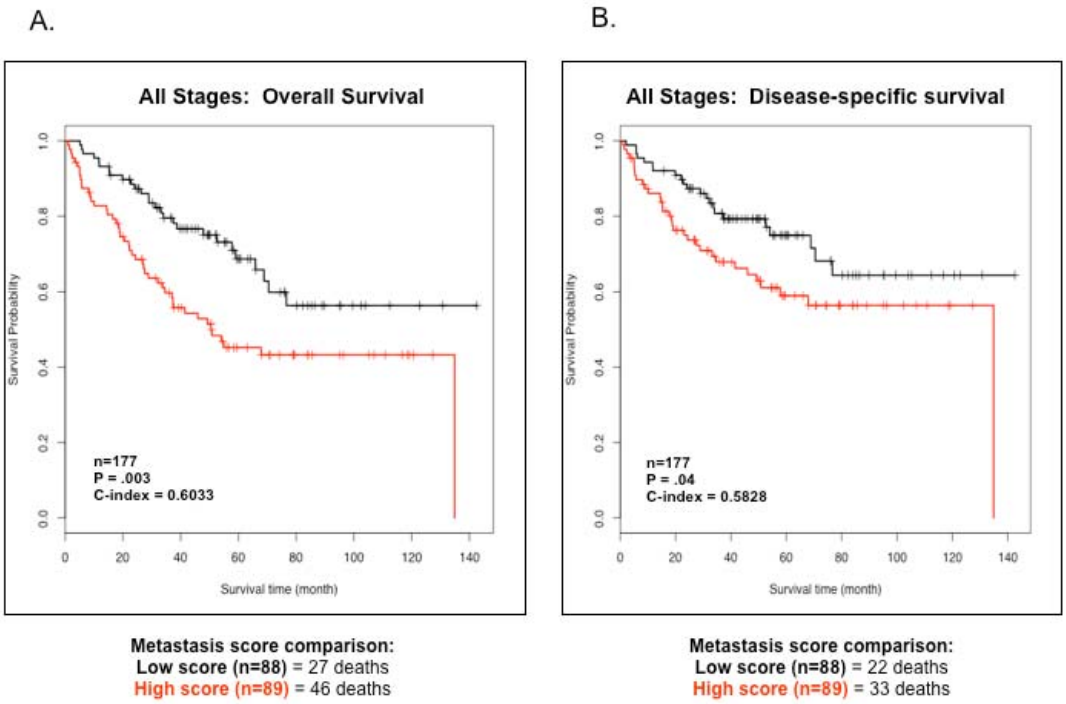


Figure 9*. The 34-gene recurrence classifier as tested in the Moffitt Cancer Center (MCC) dataset across all stages.* Kaplan-Meier estimates of overall and disease-specific survival in the MCC test set. Expression data for probes corresponding to the 34-gene recurrence classifier were used to build the Cox proportional hazard model from patient data in the Vanderbilt dataset. Plots represent survival analyses in the MCC patient data set, based upon Beta and Wald statistics (see Materials and Methods) from the Vanderbilt dataset. (A) Overall and (B) disease-specific survival analyses were performed.

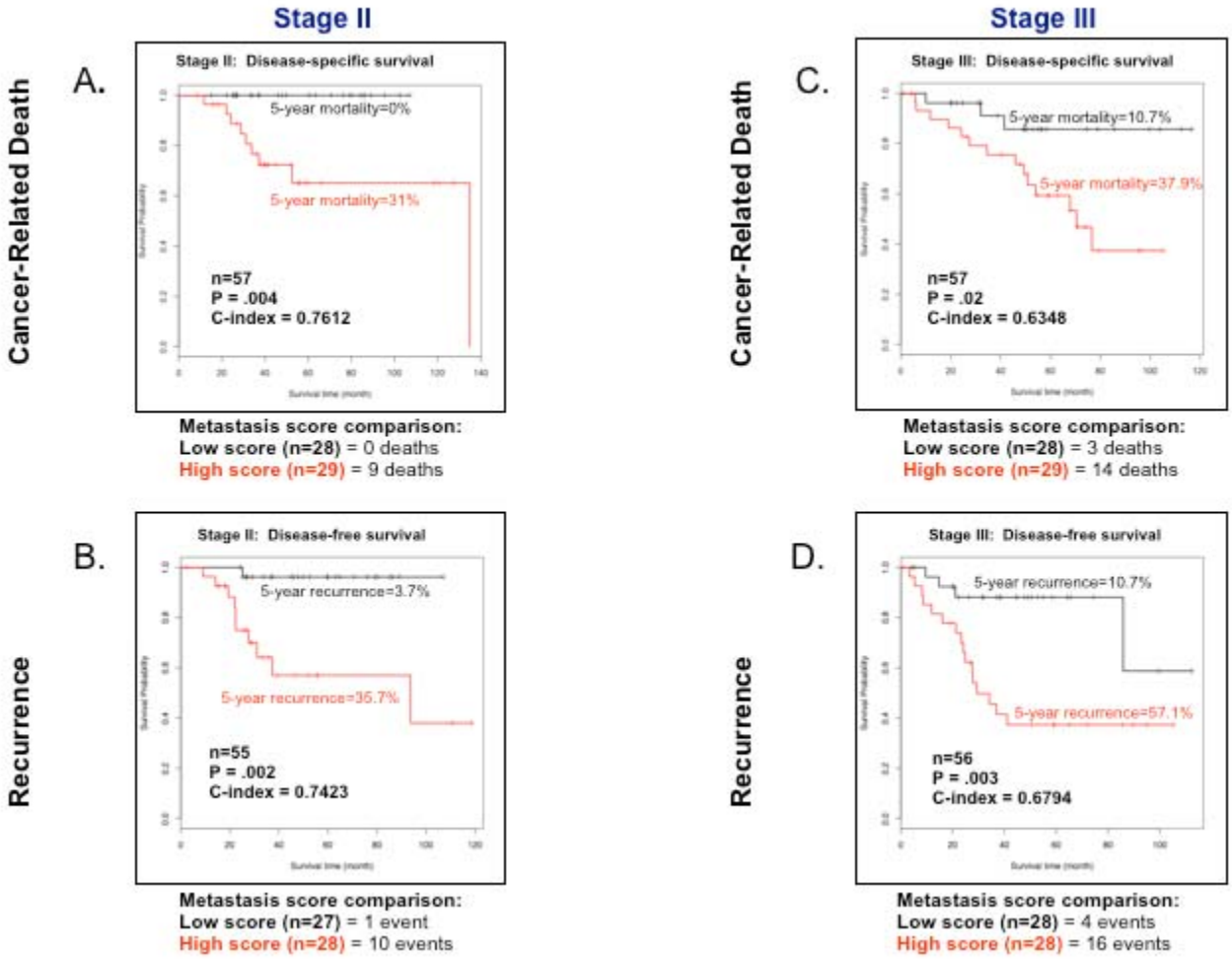


Figure 11. *Kaplan-Meier estimates from 114 colon cancer (stages II and III) patients under study at MCC analyzed with the 34-gene based metastasis score*. Lower than median metastasis score is denoted in black and higher than median metastasis score is noted in red. A low score was associated with better disease-specific and disease-free survival in stage II patients (A) Cancer-related death: n = 57 patients, high scores (9 of 9 total deaths) and (B) Disease-free survival: n = 55, high scores (10 of 11 total events). Similarly, a low score was associated with better disease-specific and disease-free survival in stage III patients (C) Cancer-related death: n = 57 patients, high scores (14 of 17 total deaths) and (D) Disease-free survival: n = 56, high scores (16 of 20 total events). 5-year mortality and recurrence rates are shown for stage II and III patients (A-D).

High Risk Characteristics and the Metastasis Score

In the stage II disease-free survival analysis, the four T4 lesions were evenly distributed between high (n=2) and low (n=2) metastasis score patients (P>.99). The number of MCC T4 tumors is small and we will need to assess these characteristics as the metastasis score is tested in additional colon cancer patients to confirm this finding in association with the 34-gene metastasis score. We found no significant differences in the distribution of well differentiated, moderately differentiated or poorly differentiated tumors in high or low score stage II patients (P=.47 and .22 respectively). Furthermore, in regard to number of lymph nodes retrieved and metastasis score in the disease-specific survival analysis, we found equivalent numbers of patients with less than 12 lymph nodes retrieved in the low and high score groups (P>.99). We also queried differentiation status and lymph nodes retrieved in the stage III patients and found no association with metastasis score and differentiation status or number of lymph nodes retrieved (data not shown). These data indicate that the metastasis score performs independently of traditional pathological markers.

To determine if previously described, yet unproved, pathologic “high-risk” features in the stage II patients were associated with either the high or the low metastasis score we reviewed the available data regarding T4 lesions, lymph node retrieval and histology in the larger MCC dataset. In the disease-specific survival analysis, four of the 28 low metastasis score patients had T4 lesions, while only one of the 29 high metastasis score patients had a T4 lesion (P=0.19). In the stage II disease-free survival analysis, the four T4 lesions were evenly distributed between high (n=2) and low (n=2) metastasis score patients (P>0.99). The numbers of T4 tumors are small in the MCC test set and we will need to assess these characteristics as the metastasis score is tested in a larger set of colon cancer patients before making any strong statements in association with the 34-gene metastasis score.

We found no significant differences in the distribution of well-differentiated, moderately differentiated or poorly differentiated tumors in high or low score stage II patients (P=.47 and .22 respectively). Furthermore, in regard to the number of lymph nodes retrieved and the metastasis score in the disease-specific survival analysis, we found that 12 of 28 low score patients had less than 12 lymph nodes retrieved while 12 of 29 high score patients had less than 12 lymph nodes retrieved (P>.99). We also queried differentiation status and lymph nodes retrieved in the stage III patients and found no association with metastasis score and differentiation status or number of lymph nodes retrieved (data not shown). These data indicate that the metastasis score performs independently of traditional pathological markers.

Microsatellite Instability and the Recurrence Classifier

Other factors known to affect prognosis (e.g., microsatellite instability status [MSI] or CpG island methylator phenotype [CIMP]) could potentially impact our signature, however, only 73 (31%) of the patients analyzed in this study (VMC + MCC) had MSI status available which prevented a robust analysis adjusting for MSI status in a multivariate model. In a preliminary analysis of the 73 patients with known MSI status in the combined VMC and MCC datasets, 13 were MSI-high and 60 were microsatellite stable (MSS). Of the MSI-high patients, 7 (54%) were in the low metastasis score group and 6 (46%) were in the high metastasis score group. Additionally, 43% of the MSS patients were in the low metastasis score group while 57% were in the high score group. No significant statistical association was noted comparing the recurrence classifier and microsatellite instability status in this study (P=.55). We are prospectively collecting MSI-status on our patients and will factor this important prognostic factor into future work in light of the 34-gene classifier. CIMP status was unavailable in the datasets used in our analysis.

A High Metastasis Score Predicts Recurrence and Survival in Univariate and Multivariate Models

The metastasis score was tested in the MCC patient dataset to determine the relative risk of recurrence and cancer-related death. Patients with a high metastasis score had increased relative risk of recurrence, as measured by hazard ratios (HR) across all stages (Table 6, HR 4.9, P<.001). High score stage II patients were also at increased relative risk of recurrence (HR 13.1, P=.01). Finally, the relative risk of recurrence in stage III patients with a high score was increased in this analysis (HR 4.7, P=.006). These data show that the metastasis score based on the 34-gene classifier is a strong predictor of recurrence and cancer-related death in colon cancer patients.

Multivariate analysis of the MCC patient data was performed to determine independent predictors of recurrence and survival. After adjusting for metastasis score, gender, tumor stage, age and tumor grade, only the metastasis score (P<.001) and tumor stage (P=.002) were significant determinants of cancer recurrence (Table 7 (see Appendix)). We observed similar results with disease-specific survival as the outcome measure in univariate and multivariate models (Tables 8 and 9 (see Appendix)). Furthermore, the magnitude of the metastasis score was significantly associated with outcome. Hazard ratios for cancer-related death demonstrated that across all stages, high metastasis score patients in the 50th, 75th and 90th percentiles are at increased relative risk of cancer-related death (HR=2.0, 3.1 and 4.6, respectively) compared with those patients with scores in the 10th percentile (Fig. 12). Therefore, the 34-gene metastasis score is an independent predictor of cancer recurrence and death in colon cancer patients even after adjustment for baseline characteristics available at resection (e.g., age, grade, stage and sex).

TABLE 6

THE 34-GENE METASTASIS SCORE ASSOCIATES WITH INCREASED RISK OF RECURRENCE

	HR (95% CI)	P-value
Disease-free survival (All stages)	4.9 (2.157-11.27)	<.001
Disease-free survival (stage II)	13.1 (1.660-103.1)	.01
Disease-free survival (stage III)	4.7 (1.566-14.050	.006

Univariate analysis was done with metastasis scores to segregate patients from the MCC data set into higher-than-median and lower-than-median score groups. Hazard ratios (HRs) were calculated for each patient group related to disease-free survival.

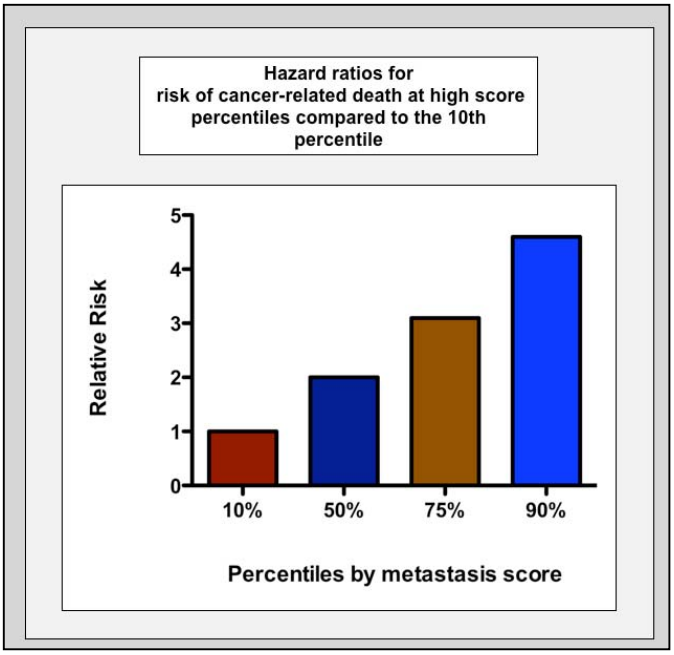


Figure 12. *Cox model hazard ratios*. Percentiles across high-score MCC patients were plotted related to relative risk. Hazard ratios for the 50th, 75th, and 90th percentiles are plotted compared with the 10th percentile and show a 2.0, 3.1, and 4.6 increased risk of cancer-related death from the 50th percentile upward.

The 34-gene Based Metastasis Score is Associated with Patient Benefit and Adjuvant Chemotherapy in Stage III Colon Cancer Patients

As described above, a significantly greater proportion of stage III patients with a high metastasis score died of cancer as compared with low metastasis score patients. Thirty percent (17 of 57 patients) of the stage III patients did not receive adjuvant chemotherapy (CTX). Therefore, we sought to determine if there was a difference in survival between high and low metastasis score patients related to CTX administration. Stage III patients with a low metastasis score had equivalent survival outcomes regardless of whether or not they received adjuvant CTX (Table 10, 10% with CTX versus 12.5% without CTX, P>.99). Among stage III patients with a high metastasis score, a significantly greater proportion of patients who did not receive CTX died from their cancer as compared with those who did receive adjuvant CTX (86% vs. 36%, P=.04). There was no statistically significant difference in follow-up interval or in the proportion of patients receiving CTX when comparing high and low metastasis score groups (P=.576 and P=.770). These data suggest that stage III patients with a low metastasis score did not gain significant benefit from adjuvant CTX, whereas stage III patients with a high metastasis score had a better outcome after adjuvant CTX.

TABLE 10

METASTASIS SCORE ASSOCIATES WITH PATIENT BENEFIT AND ADJUVANT CHEMOTHERAPY IN PATIENTS WITH STAGE III COLON CANCER FROM THE MCC DATASET

Metastasis score and survival	Alive	Cancer-related death	P-value
Low score	25 (62.5%)	3 (17.6%)	.003
High score	15 (37.5%)	14 (82.4%)	

Low score	CTX (n=20)	No CTX (n=8)	P-value
Cancer-related death	2 (10%)	1 (12.5%)	>.99
Alive	18 (90%)	7 (87.5%)	

High score	CTX (n=22)	No CTX (n=7)	P-value
Cancer-related death	8 (36.4%)	6 (85.7%)	.04
Alive	14 (63.6%)	1 (14.3%)	

Table of proportions shows associations between metastasis score, cancer-related death and exposure to adjuvant chemotherapy in stage III patients from the MCC data set. As can be seen in the low score panel, no significant association was found between a low metastasis score and cancer-related death in those stage III patients receiving chemotherapy versus those who did not receive chemotherapy (Fisher's exact test).

Pathway Analysis of the 34-gene Recurrence Classifier

In order to relate our findings to basic tumor biology, we subjected the recurrence classifier to Ingenuity Pathways Analysis (IPA, Ingenuity Systems®, www.ingenuity.com). Up-regulated genes formed a single network enriched for cancer and cell death while the down-regulated gene network was enriched for cell-to-cell signaling/interaction and immune response (see Table 4 and details below).

Ingenuity Pathways Analysis and the Recurrence Classifier

In this analysis, up-regulated genes (n=16) formed a single interacting network that is functionally enriched for disease and process including cancer (*CXCR7, EGR1, HES1, SPRY4, VDR, NQO1, ACTB, MGP*), endocrine system disorders *(HES1, VDR, EGR1)*, and cell death *(EGR1, NQO1, HES1, VDR, MGP)* (Table 4 and Fig. 13). Down-regulated genes (n=18) similarly formed a single interacting network enriched for cell-to-cell signaling/interaction *(MMP13, PRTN3, HPSE, DFBN31)* and immune response/development/function *(HPSE, PRTN3, MMP13)* (Table 4 and Fig. 14).

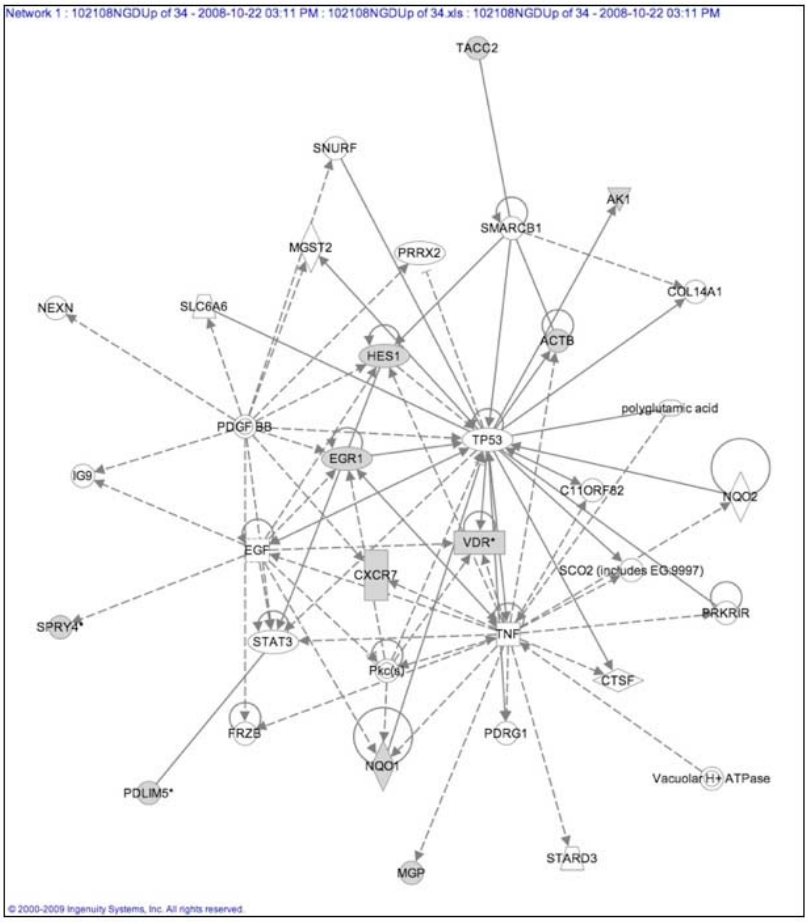


Figure 13. *Functional enrichment of the up-regulated genes in the recurrence classifier.* Genes or gene products are represented as nodes, and the biological relationship between 2 nodes is represented as an edge (*line*). All edges are supported by 1 reference from the literature, from a textbook, or from canonical information stored in the Ingenuity knowledge base. The 16 up-regulated genes in the classifier were considered for the analysis (*shaded in gray*). Predominant central regulatory nodes include p53 (*TP53*), tumor necrosis factor (TNF), and platelet-derived growth factor BB (PDGFBB).

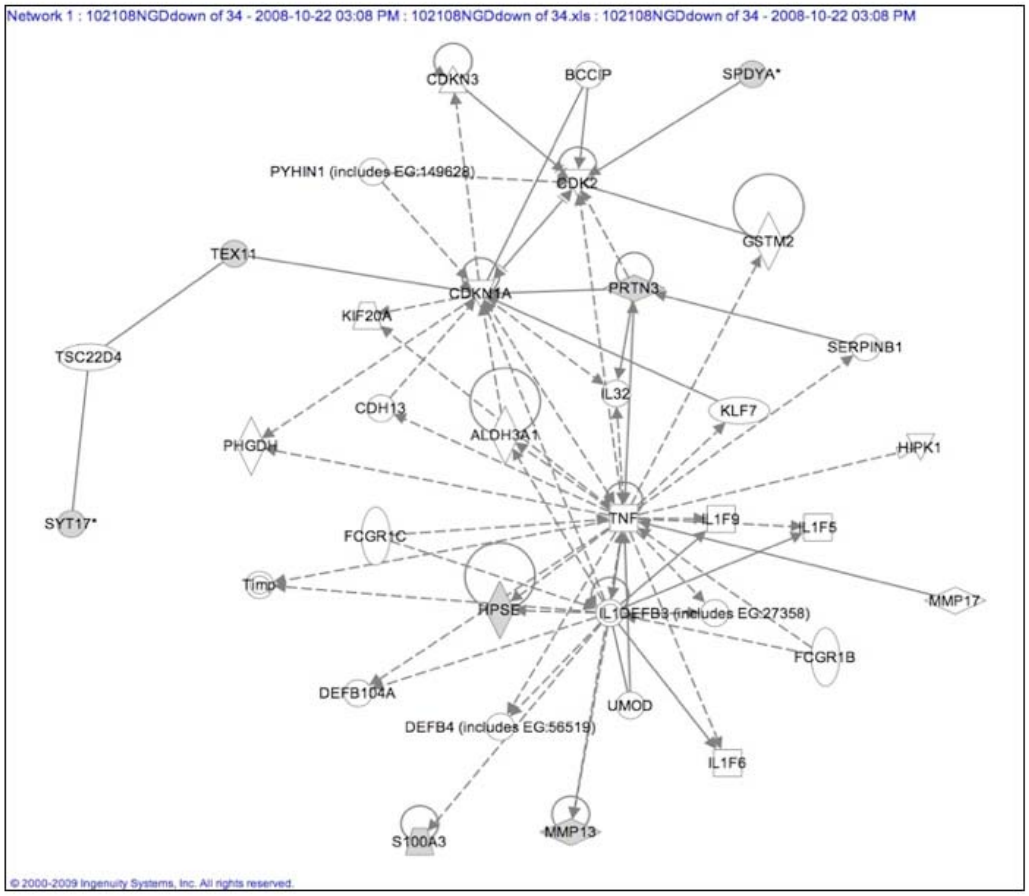


Figure 14. *Functional enrichment of the down-regulated genes in the recurrence classifier*. Genes or gene products are represented as nodes, and the biological relationship between 2 nodes is represented as an edge (*line*). All edges are supported by 1 reference from the literature, from a textbook, or from canonical information stored in the Ingenuity knowledge base. The 18 down-regulated genes from the high-risk signature were considered for the analysis (*shaded in gray*). Predominant central regulatory nodes in this network include tumor necrosis factor (*TNF*), cyclin-dependent kinase N1A (*CDKN1A or p21*), and interleukin-1 (*IL-1*).

As can be seen in Fig. 7B, the magnitude of the expression differences between up-regulated and down-regulated genes amongst patient samples from these two clusters is quite different, with differences in up-regulated genes being of higher magnitude than differences in down-regulated genes. This distinction is suggestive of two uniquely regulated gene networks. Indeed, when we performed network analysis, coordinate regulation of 11 of 16 up-regulated genes (69%) and 7 of 18 down-regulated genes (39%) is apparent. We found that the gene network represented by the up-regulated genes is predominantly enriched in genes related to carcinogenesis, while the network represented by the down-regulated genes is predominantly enriched in genes related to inflammation. Specifically, a central node of the up-regulated genes was TP53, which has a significant role upon inactivation in the progression of colorectal cancer as discussed in Chapter I. Additionally, a central node of the down-regulated genes was Tumor Necrosis Factor, which plays a major role in inflammation. Chronic inflammation is a well-recognized risk factor for cancers of the alimentary tract (reviewed in (Wang and DuBois 2008)). Experimental mouse models have been developed to demonstrate that inflammation is a potent tumor promoter in the gastrointestinal tract (Cooper et al. 2000). Indeed, the incidence of tumor formation in several genetic mouse models of intestinal neoplasia is dependent upon active immune responses induced by gut flora (Kado et al. 2001). Conversely, tumor progression is also associated with immune suppression. In the current study, we used microarray analysis to identify perturbations in gene networks related to both carcinogenesis and inflammation. Ongoing work in our laboratory aims to uncover the molecular mechanisms by which these two gene networks may interact to promote metastasis in our experimental model.

Discussion and Future Directions

In the present study, the biology of colon cancer metastasis was modeled in immunocompetent mice to develop a gene expression classifier that discriminates recurrence and survival outcomes in human colon cancer patients. Stage II and stage III patients with primary colon cancers that reflected the recurrence-associated gene expression pattern were at greater relative risk of recurrence than those who did not (hazard ratios of 13.1 and 4.7, respectively). This gene expression profile, tested with a recurrence scoring method, performed independently of conventional pathological staging.

Perhaps most importantly, this metastasis score identifies stage II patients at high risk of recurrence and death and stage III patients at low risk of recurrence and death. Our biological model has identified a subset of high-risk stage II patients that may benefit from adjuvant therapy and a subset of low-risk stage II patients who may have an excellent outcome after surgical resection without adjuvant therapy. We found that the 5-year survival rate was >95% in stage II patients with a low metastasis score, suggesting that adjuvant chemotherapy would provide minimal benefit in this group of patients. In contrast, 31% of stage II patients with a high metastasis score died of cancer. Our preliminary analyses of these data suggest that high metastasis score stage II patients should be further studied to determine whether they will benefit from adjuvant therapy.

A unique aspect of our study is the inclusion of sufficient numbers of stage III patients who did not receive adjuvant chemotherapy in the MCC database. This enabled an evaluation of whether the molecular metastasis score could predict response to adjuvant therapy. Of the high metastasis score stage III patients who were treated with adjuvant chemotherapy only 36.4% died from cancer, whereas 85.7% of the high score patients who did not receive adjuvant chemotherapy died from

cancer. Despite the small numbers in these sub-groups the differences were statistically significant. More importantly, equally low proportions of stage III patients with a low metastasis score died of cancer regardless of administration of chemotherapy. Our data suggest that there is a low-risk group of stage III patients who could be surgically cured and spared the morbidity, expense and potential mortality associated with adjuvant chemotherapy. This is consistent with prior observations from randomized clinical trials that established the benefits of adjuvant chemotherapy in stage III colon cancer where 40-44% of patients enrolled in the surgery-only groups did not recur in five years even without adjuvant treatment (Ragnhammar et al. 2001). Determination of an objective scoring method whereby the 34-gene classifier can be tested in a prospective fashion is ongoing and will be required to determine if the 34-gene based metastasis score can be used clinically to guide decisions regarding adjuvant therapy for stage III colon cancer patients.

Several investigative groups have reported gene expression classifiers with predictive power in breast, lung, liver and colorectal cancers (Barrier et al. 2006; Hoshida et al. 2008; Lin et al. 2007; Paik et al. 2004; Shedden et al. 2008; Wang et al. 2004). Like the previously described colon cancer classifiers, a weakness of our model is the retrospective analysis of prospectively collected clinical data. A 43-gene poor-prognosis signature for colorectal cancer provides a classifier for stage II and III patients as a molecular staging device (Eschrich et al. 2005). In a more recent study, a computational model was used to derive a 50-gene signature and a metastasis score for early stage colon cancer (Garman et al. 2008). We found minimal overlap between our 34-gene classifier and the previously published, computationally derived colon cancer gene signatures. We were not surprised at this finding since the prior models were computationally determined and ours is founded on the biology of metastasis.

We find it interesting that 13 of the 34 genes in our proposed classifier have previously described roles in cancer, and several others are involved in cell-cell signaling, immune response, cell proliferation, embryonic development and cell migration. Inflammation plays a potent role in gastrointestinal tract tumor promotion and tumor progression is associated with immune suppression (Cooper et al. 2000; Kado et al. 2001; Wang and DuBois 2008). In this study, microarray analysis identified perturbations in gene networks related to both carcinogenesis (e.g., TP53) and inflammation (e.g. TNF). Ongoing work in our laboratory aims to unravel the molecular mechanisms by which these two gene networks may interact to promote metastasis in our experimental model.

A number of recent examples provide additional insight into the contributions of components in the 34-gene classifier to the process of EMT. For example, TNF-α (a central node of the down-regulated genes in the classifier) has been connected to TGFβ-induced EMT via the Wnt target gene CD44 (Takahashi et al.). Additionally, hypoxia has been noted to promote EMT in breast cancer cells via up-regulation of HIF1-α to promote transcriptional up-regulation of HES1 (Notch signaling target gene) and down-regulation of E-cadherin (Chen, Imanaka, and Griffin). Also, Zeb1 (a known pro-EMT gene inversely correlated with E-cadherin (Schmalhofer, Brabletz, and Brabletz 2009)) has been found to positively correlate with levels of vitamin D receptor in colorectal tumors (Pena et al. 2009). These data open up new possibilities that may inform the process of EMT that can be further explored in the MC-38 model in immunocompetent mice.

The cross-species functional genomics approach yields insights into the molecular mechanisms of the metastatic process. Consistent with our approach, gene expression patterns identified in wound healing have been applied successfully to breast cancer outcomes (Chang et al. 2004). In addition, cell culture and mouse models have also demonstrated relevance to clinical outcomes in hepatocellular carcinoma using gene expression profiling (Kaposi-Novak et al. 2006; Lee

et al. 2004). Similarly, our approach has uncovered a gene classifier with prognostic significance in colon cancer.

Although a high score based on the 34-gene recurrence classifier worked well in this study, the number of significant genes reported was not based on the smallest number of genes that could discriminate the survival endpoint, but was based upon the combined statistical, biological and clinical evidence in the available data. There is also the possibility that some of the computationally derived genes discovered in human datasets would be missed in a mouse model; however, the biological basis of the 34-gene classifier derived from our mouse model seems to be a robust predictor. The possibility of achieving similar or better survival discrimination with different subsets of the genes certainly exists; however, we feel that the biological basis of our study provides a solid foundation for further translational application and testing of our model.

In conclusion, the 34-gene based metastasis score can identify stage II and III patients at greater risk of colon cancer recurrence and death. Our biologically based expression classifier identifies a potential method for more appropriate selection of patients for systemic therapy after curative-intent surgical resection of colon cancer. Future prospective studies are needed to confirm whether chemotherapy may be safely avoided in stage III patients with a low metastasis score and whether stage II patients with a high metastasis score can achieve a better outcome if they receive adjuvant chemotherapy. Furthermore, we have begun work to determine the strongest drivers of the 34-gene classifier with additional statistical modeling as we move forward with the determination of an objective score that can be taken into a prospective trial in stage II colon cancer patients. Finally, since the model is based on the biology of metastasis we are currently looking into the ability of the 34-gene classifier to predict outcome in other epithelial cancers such as breast or lung cancer.

CHAPTER III

SMAD4 INHIBITS WNT SIGNALING IN EPITHELIAL CELLS BY TRANSCRIPTIONAL REPRESSION OF β-CATENIN

Abstract

Transforming Growth Factor-β and Wnt signaling pathways are essential during embryonic development, and later regulate intestinal epithelial cell differentiation and proliferation. Central mediators of these pathways, Smad4 and β-catenin, have important roles in cellular signaling and colorectal cancer pathogenesis. Smad4 is a tumor suppressor frequently mutated in colon cancer; whereas β-catenin accumulation and nuclear localization contribute to both the initiation and progression of colorectal cancers. Smad4 expression in Smad4-mutant colon cancer cells results in decreased β-catenin protein levels and inhibition of β-catenin/TCF-driven transcription. Here, we determine the mechanism and significance of Smad4 inhibition of Wnt signaling in colorectal cancer. We found that Smad4 mediates a direct transcriptional repression of *β-catenin* with resultant inhibition of β-catenin-mediated gene expression in epithelial cells. Consistent with these findings, we observed that the biological consequence of Smad4 expression in colon cancer cells is reversal of epithelial-to-mesenchymal transition (EMT) and cell invasiveness. From analysis of clinical tumor samples, we found that patients whose tumors exhibited high Smad4 expression and reduced expression of Wnt target genes have significantly better prognosis than those who exhibit low Smad4 expression with elevated gene expression of Wnt targets and EMT markers.

Introduction

Mutational inactivation of *Adenomatous polyposis coli (APC)* occurs in >80% of sporadic colorectal cancers associated with the loss of a master cell regulatory mechanism for epithelial cell growth, renewal and homeostasis (Ashton-Rickardt et al. 1989; Morin et al. 1997; Rowan et al. 2000; Sansom et al. 2004). APC is part of a cellular protein complex that regulates the availability of the Wnt signaling mediator, β-catenin, which is critical for maintenance of the intestinal epithelial stem cell compartment at steady-state (Aberle et al. 1997; Ikeda et al. 1998; Kishida et al. 1998; Korinek et al. 1997; Munemitsu et al. 1995; van de Wetering et al. 2002). In the absence of functional APC, β-catenin accumulates and translocates to the nucleus along with TCF/LEF co-transcriptional mediators where it may inappropiately propagate canonical Wnt signaling outside of its normal stem cell niche (Radtke and Clevers 2005). Pathological activation of Wnt signaling participates in the initiation and progression phases of colorectal cancer and accumulating evidence suggests that it drives epithelial-mesenchymal transition (EMT), particularly where signaling is further amplified at the invasive front in colorectal cancer specimens (Brabletz et al. 2005; Hlubek et al. 2007). The EMT process has been linked to invasive phenotype and metastatic potential in cancer (Thiery 2002, 2003).

The Transforming Growth Factor-β (TGFβ) family of ligands (TGFβs, Inhibin, Activin, Nodal, Bone Morphogenetic Proteins (BMPs)) and cell signaling responses underlie fundamental epithelial cell processes of proliferation, apoptosis and differentiation (Massague 2008; Shi and Massague 2003). TGFβ inhibits epithelial cell growth and proliferation in normal cells, a property which is lost in many transformed cells (Ko et al. 1995; Manning et al. 1991) and in cancer (Hanahan and Weinberg 2000; Moses, Yang, and Pietenpol 1990; Pietenpol et al. 1990). Over 50% of colorectal cancers harbor inactivating mutations in the TGFβ/BMP signaling pathway (e.g., receptor or Smad signaling protein

(Grady et al. 1999)). TGFβ and BMP signaling occurs through interaction with heterotypic (Type I/Type II) serine/threonine kinase receptors which phosphorylate receptor-associated Smads (Smad2/3 or Smad1/5/8), and then complex with Smad4 and translocate to the nucleus to promote transcription (Shi and Massague 2003). Smad4, the common mediator of TGFβ/BMP signaling, is the most frequently disrupted Smad mediator in pancreatic (Hahn et al. Homozygous deletion map at 18q21.1 in pancreatic cancer 1996) and colorectal cancers (Thiagalingam et al. 1996). Smad4 mutations have been identified in 50% of pancreatic cancers (Hahn et al. Homozygous deletion map at 18q21.1 in pancreatic cancer 1996; Hahn et al. Dpc4, a candidate tumor suppressor gene at human chromosome 18q21.1 1996) and 10-35% of colorectal cancers (Markowitz and Bertagnolli 2009; Riggins et al. 1997; Riggins et al. 1996).

TGFβ/BMP and Wnt signaling pathway crosstalk has an important impact on embryonic development and homeostasis (Cadigan and Nusse 1997; Whitman 1998). For example, expression of BMP antagonists is associated with dedifferentiation, increased proliferation and enhanced β-catenin nuclear localization in normal human intestinal epithelium and in cancer (Kosinski et al. 2007). Transgenic Noggin expression in the mouse intestine results in intestinal polyps similar to humans with Juvenile Polyposis Syndrome (JPS) and these lesions were noted to have increased nuclear β-catenin immunostaining (Haramis et al. 2004). Smad4 loss plays a significant role in JPS as patients with reduced intestinal Smad4 expression are predisposed to the development of intestinal carcinomas (Howe et al. 1998). In *Xenopus*, Smads interact with TCF/LEF1 and activate transcription from the *Xenopus* Wnt signaling target, *Xtwn (Labbe, Letamendia, and Attisano 2000; Nishita et al. 2000)*. Interestingly, compound *Smad4/APC*-heterozygous mutant mice develop more invasive intestinal carcinomas than *APC* heterozygous mutant mice in which benign intestinal adenomas are the predominant lesion, suggesting that Smad4 expression tempers *APC*-driven tumor

formation (Takaku et al. 1998). Therefore, dual-regulation of TGFβ and Wnt pathways can modulate differential responses of growth inhibition, tumor promotion or de-differentiation in epithelial cells; however, specific mechanisms by which Smad4 interacts with components of the Wnt signaling pathway in colorectal cancer remain largely undefined.

In this study, we examined the mechanism by which Smad4 represses β-catenin expression and Wnt signaling in colorectal cancer. We found that Smad4 expression results in transcriptional repression of β-catenin expression and thereby inhibits Wnt signaling, with an associated inhibition of the downstream program of Wnt target gene expression. Finally, the clinical relevance of this effect was demonstrated by an epithelial cell-specific, Smad4-modulated gene expression profile associated with suppression of Wnt signaling in human colorectal cancer specimens. This Smad4-modulated, Wnt-associated gene expression program yields prognostic information, independent of conventional pathological staging in a large cohort of colorectal cancer patients.

Materials and Methods

Primers

Primers for all PCR reactions were obtained from either realtimeprimers.com (Elkins Park, PA) or from Integrated DNA Technologies, Inc. (Coralville, Iowa) and are listed in (Table 17, see Appendix p. 187).

Reverse Transcriptase Polymerase Chain Reaction (RT-PCR)

RNA (see collection method below) was reverse transcribed with MMLV reverse transcriptase (Promega, City, State, M170A) in each reaction as described below (Sequences obtained and then ordered as above). RNA (0.5 μg), random hexamers (2.5μL) and RNase/DNase free water (5μL)

were incubated at 70ºC for 5 minutes (11μl reaction). MMLV-RT 5X buffer (5μL), dinucleoside triphosphates (dNTPs, 10mM, 5μL), RNAse inhibitor (1μL), MMLV-RT enzyme (1μL) and water (5μL) were placed into the 11μL reaction and incubated at 37ºC for 5 minutes, 25ºC for 10 minutes, 37ºC for 60 minutes and stopped at 70ºC for 10 minutes. Red Taq® buffer (2μL), Red Taq® polymerase (Sigma-Aldrich, St. Louis, MO) enzyme (1μL), dNTPs (0.4μL), cDNA (2μL), Primer set (2μL, Forward + Reverse) and RNase/DNase free water (12.6μL) were used for each 20μL reaction. Each cDNA reaction was prepared on the same day and then run on a Gene Amp PCR System 9700, Applied Biosystems (Foster City, CA) under the following protocol (94°C--5min; 32 cycles of 94°C--20sec; 55°C-20sec; 72°C-20sec; 72°C--1min). Twenty μL of the PCR product and 4μL of 6X DNA loading dye was then loaded into 2% agarose gel and run at 100V for 1-1.5 hour and visualized on a BioRad Molecular Imager (Hercules, CA).

Quantitative real-time PCR (qPCR)

RNA Collection

SW480^{vector} and SW480^{Smad4} colon cancer cells were seeded equally into 6-well plates. At 80% confluence, cells were harvested in parallel using the Qiagen RNeasy® kit (Valencia, CA), according to the manufacturer's instructions. Samples were DNAse treated during purification and eluted with 50 μL RNAse/DNAse free water. RNA integrity was checked by visual inspection following agarose gel electrophoresis. Quantitation was conducted in triplicate by UV Spectrophotometry. RNA (300ng) was reverse transcribed using Superscript III (Invitrogen) in each reaction using gene specific primers (see Appendix, p. 187). Reactions were performed for 20 min at 50^{o}C followed by 10min at 95^{o}C to inactivate enzymes. For PCR reactions, each well contained SYBR Green Master Mix (12.5μL) from SuperArray, RNase/DNase free water (10.5μL) combined with 10μM Primer sets (1μL)

and 1μL of freshly prepared cDNA. Each cDNA reaction (25μL) was loaded in quadruplicate in 96 well plates and run on a Bio-Rad iCycler (Hercules, CA) under the following protocol (1: 95°C, 30sec; 2: 95°C, 30sec (x2); 3: 95°C, 13min-30sec; 4: 95°C, 15sec; 55°C, 30sec; 72°C, 30sec (x45); 5: 72°C--7min; 6: 55°C, 10 sec (x100)).

mRNA Stability

SW480^{vector} cells and SW480^{Smad4} cells were grown in regular medium and treated at 4-hour intervals with 5,6-Dichlorobenzimidazole, 1-β-D-ribofuranoside (DRB, Sigma, St. Louis, MO). After RNA isolation, qPCR was utilized to determine the difference in *β-catenin* steady-state mRNA levels after DRB treatment. Cells were harvested at 4-hour intervals for 24 hours. RNA was isolated from cells as above at each time point. Two doses of DRB were used. qPCR was used to analyze the amounts of *β-catenin* mRNA at each time point and analysis of covariance was used to determine if the slope if the two lines was significantly different. Cyber-green qPCR was performed using human β-catenin primer (see Appendix, p. 187 and main methods).

Data Analysis

For each individual well, fluorescence curves were log transformed and the slope of the logarithmic portion of the reaction was extracted to determine efficiency. Ct values and efficiency were utilized to calculate the relative expression levels and plotted as 1/Ct (Cronin et al. 2007; Schefe et al. 2006).

Flow Cytometry

Transient expression of Smad4 was carried out in SW480 cells and HEK-293T cells along with GFP. GFP-positive cells were separated from GFP-negative cells with a cell sorter and the populations were compared. RNA and protein were isolated from the two populations as described and the difference in *β-catenin* mRNA was determined by qPCR after Smad4 expression was

confirmed by Western analysis (Flag and Smad4 antibody, see Immunodetection section). Flow Cytometry experiments were performed in the VMC Flow Cytometry Shared Resource. Specifically, SW480 and HEK-293T cells were transiently co-transfected with pRK5 DPC4 Flag (Zhang et al. 1996) (Addgene, Cambridge, MA), pRK5 and GFP plasmid (pEGFP-C3, BD Biosciences Clontech, Mountain View, CA). After validation that the GFP+ cells represented cells transfected with pRK5 DPC4 (e.g., +Flag band and Smad4 band on western compared to control SW480 cells; data not shown) then qPCR was completed. pRK5 was used as a negative control and was found to be equivalent to GFP- cells by Western and qPCR (e.g., no Flag on Western for SW480s, no change in baseline Smad4 levels for HEK-293T cells and no difference in β-catenin by qPCR). Qiagen Effectene reagents (Qiagen, Valencia, CA) were used for the transfections and equal amounts of DNA (1.0μg) were transfected per well for each experiment.

Chromatin Immunoprecipitation (ChIP) Assay

ChIP assays were performed using the EZ-Magna ChIP™ kit (Millipore Cat. # 17-408, Billerica, MA) according to the manufacturer's instructions. Briefly, $SW480^{vector}$ and $SW480^{Smad4}$ colon cancer cells were fixed in fresh 18.5 % PFA for ten minutes at room temperature prior to cross-linking. After cell and nuclear fractionation, DNA from the nuclear fractions was sheared by sonication. For the immuno-precipitation, Protein A beads (Millipore CS200637) were blocked by coating with bovine serum albumin (BSA) (1μg/mL) overnight (50-100μL of BSA for each 20μL of Protein A beads). Cell nuclear lysates were pre-cleared by addition of 10μL of coated beads and 2μL of mouse IgG (Santa Cruz Biotechnology, Santa Cruz, CA)). Five microliters of precleared lysates were saved as input. 500μg of the precleared lysate was used in overnight incubation at 4°C with each of 5 μg Smad4 antibody (Santa Cruz Biotechnology, Santa Cruz, CA), Anti-Acetyl Histone 3 antibody (Millipore,

Billerica, MA), RNA polymerase II antibody (Millipore, Billerica, MA), or mouse IgG (Santa Cruz Biotechnology, Santa Cruz, CA). Following washes and reverse crosslinking, DNA was isolated and preparations were made for PCR reactions.

ChIP PCR

For the 20μL PCR reactions, each well contained: ChIP DNA (2μL), RNase/DNase free water (12.4μL), 10X PCR Buffer (2μL), 25 mM magnesium chloride (1.2μL), 10mM dNTPs (1.6μL), 10μM primer set (0.8μL) and Qiagen HotStart Taq (5U/μL, 0.4μL, Valencia, CA). The following protocol was used: 1: 94°C, 3 min; 2: 94°C, 45 sec; 3: 55°C, 45 sec; 4: 72°C, 45 sec (x35); 5: 72°C--2min; 6: 4°C. The PCR product and 4μL of 6X DNA loading dye were loaded into a 12% acrylamide gel and run at 100 volts for 1.5 hours and visualized on a BioRad Molecular Imager (Hercules, CA, USA).

Transcription Assays

$SW480^{vector}$ and $SW480^{Smad4}$ colon cancer cells were used as described (Shiou et al. 2007). SW480 cells from American Type Culture Collection and HEK-293T were cultured in DMEM and RPMI media (Invitrogen, Carlsbad, CA) supplemented with 10% fetal bovine serum, L-glutamine, penicillin and streptomycin. See below for collection and use of Wnt3a conditioned medium for HEK-293T cells (added post-transfection). Cells were transiently transfected with pRK5 or pRK5 DPC4 Flag (Zhang et al. 1996) (Addgene, Cambridge, MA) in addition to TOPflash or FOPflash reporter plasmids (Millipore, Billerica, MA) to determine TCF-mediated transcriptional activity (Topflash/FOPflash ratio, representative of independent experiments). Qiagen Effectene reagents (Qiagen, Valencia, CA) were used for the transfections and equal amounts of DNA were transfected per well for each experiment. SW480 and $SW480^{Smad4}$ cells were transiently transfected with wild-type β-catenin (*Xenopus* and human (*simian* CMV promoter in pCS2+), gifts from E. Lee laboratory,

Vanderbilt) and TCF-mediated transcription was determined 48 hours post-transfection via luminometer (BioTek II, Winooski, VT) and data shown is representative of at least three independent experiments. Luciferase reagents (Dual-Glo Luciferase Assay System Cat# E 2940, Promega, Madison, WI) were used per manufacturer instructions.

Microarray Experiments: Human tissues, Cell Lines and Microarray Platform

The protocols and procedures for the Vanderbilt Medical Center (VMC) cohort were approved by the Institutional Review Boards at the University of Alabama-Birmingham Medical Center, Vanderbilt Medical Center and the Veterans Administration Hospital (Nashville, TN). The Moffitt samples were also approved by their local IRB. Representative sections of fresh tissue specimens were flash frozen in liquid nitrogen and stored at -80ºC until RNA isolation. Quality assessment slides were obtained to verify the diagnosis of cancer or normal adjacent mucosa. Stage was assessed by American Joint Commission on Cancer (AJCC) guidelines for both cohorts of tumor samples. RNA for human tissue and cells was purified using the RNeasy® kit (Qiagen, Valencia, CA). Cells were harvested at 70-80% in regular medium for RNA isolation. Samples were hybridized to the Human Genome U133 Plus 2.0 GeneChip Expression Affymetrix® array. Six, independent biologic replicates were harvested for $SW480^{vector}$ cells and each of the three independent $SW480^{Smad4}$ clones. RNA was isolated as above and hybridized to the same platform.

Wnt target list

A Wnt target list was verified, updated and gene identifiers were collected. Published support for these targets was verified and the gene identifiers were converted to Affymetrix probe identifiers.

This list was then used to determine Wnt enrichment for the Smad4-modulated cell line and human data as described in the statistical methods (see Appendix, pp. 177).

Wnt3a Conditioned Medium

L Wnt-3A cells (ATCC #CRL-2647) were brought up in high glucose DMEM supplemented with 10% FBS, l-glutamine, pen-strep, and G418 at 400μg/mL. They were allowed to grow to ~60% confluence, and were then split at 1:10 into five T75 flasks in 10mL complete DMEM / flask without G418. The cells were allowed to grow for 4 days, then the conditioned medium was removed from the flasks, sterile filtered, and placed into a 50mL conical tube. The five flasks were re-fed with 10mL complete DMEM, and were cultured for an additional 4 days. The medium was removed and sterile filtered as before. The first and second harvest of conditioned medium were labeled as such and stored at 4°C. The 1st and 2nd batch of conditioned medium was mixed 1:1 and per ATCC protocol and we applied the conditioned medium at 1:1.

Immunodetection

Immunoblots

Cells were lysed in radioimmunoprecipitation assay buffer and lysates were sonicated and centrifuged and protein concentration was determined with bovine serum albumin standards. Western blotting was carried out as described (Dhawan et al. 2005) with antibodies to E-cadherin (BD Biosciences, San Jose, CA), β-catenin (BD Biosciences, San Jose, CA), Vimentin (Santa Cruz Biotechnology, sc-6260, Santa Cruz, CA), Smad4 (Santa Cruz Biotechnology, Santa Cruz, CA), anti-Flag M2 (Sigma, St. Louis, MO), PARP (nuclear control, Santa Cruz Biotechnology, Santa Cruz, CA) and β-actin (Santa Cruz Biotechnology, Santa Cruz, CA). SW480, $SW480^{vector}$ and $SW480^{Smad4}$

clones were harvested and cell lysates were subjected to western blot with use of E-cadherin and vimentin specific antibodies (β-actin was used as a loading control). Active Motif Nuclear Extract Reagents (Carlsbad, CA) were used to obtain nuclear fractions per manufacturer's instructions

Immunofluorescence

SW480^{vector} and SW480^{Smad4} cells were each stained for β-catenin, E-cadherin or p120 catenin and 4'6'-diamidino-2-phenylindole (DAPI) after methanol fixation. *Antibodies for immunofluorescence*: β-catenin (Sigma, 1:2000, anti-rabbit, Alexa-fluor 579); E-cadherin (BD Biosciences, 1:100, anti-mouse, alexa-fluor 488) and p120 catenin (courtesy of Albert B. Reynolds, Ph.D. laboratory, p120 rabbit: 1:250, p120 mouse: 1:1000 (p120-rabbit: red-594; p120-mouse: Green-488). DAPI was used as indicated in the figure legends (Vector Laboratories, Inc., Burlingame, CA). Photomicrographs of confocal images were acquired with an Olympus Fluoview microscope (Olympus FV-1000) fitted with a 40X objective. Images were merged using Olympus Fluoview software. Experiments, data analysis and presentation of figures were performed in part through the use of the VUMC Cell Imaging Shared Resource.

Immunohistochemistry

Colorectal and normal adjacent tissue samples from Vanderbilt Medical Center were stained for Smad4 and β-catenin. Slides were stained using a 1:50 dilution of mouse monoclonal anti-Smad4 antibody (100μg/mL stock, Santa Cruz Biotechnology, Santa Cruz, CA) or a 1:800 dilution of mouse monoclonal anti-β-catenin antibody (250μg/mL stock, BD Biosciences, San Jose, CA). In both cases, antigen retrieval was carried out in citrate buffer (pH = 6.0) under pressure for 15 minutes, followed by quenching with H_20_2. Antigen retrieval was carried out in EDTA at 98^oC in a decloaking chamber and quenched with H_20_2. Photomicrographs were taken with the Ariol® SL-50 system at 20X. All images were acquired with a 20X objective on a CoolSNAP-ES CCD camera. Images were captured as a

montage and selected areas in original resolution are displayed as enlarged images. Native resolution (0.322 microns/pixel) is shown.

Functional Assays

Matrigel invasion assay

Transwells (8 μm pore size, 6.5 mm in diameter) from Costar (Cambridge, MA) were coated with 40μL of 2.5 mg/mL Matrigel and incubated for 2h. Cells were trypsinized, washed with PBS, resuspended in serum-free medium with 0.2% BSA, and then seeded in Transwells (75,000 cells/well). Cells were grown on Transwells with 10% FBS containing medium in the lower chamber for 72h. Transwells were fixed in 70% ethanol for 1 hour at 4^{o}C. Cells remaining in the top chamber were removed with cotton swabs, and the cells that traversed to the reverse side of the inserts were rinsed with PBS and stained with propidium iodide for 1h to overnight at room temperature. Cells were counted under a light microscope (at 200X) and invasive cell numbers were the averages of five areas on each insert. Each invasion assay was done in triplicate and repeated thrice.

Athymic (nude) mouse tumorigenicity assay

SW480^{vector} and SW480^{Smad4} cells were grown in regular medium and 1 x 10^{6} cells were injected into sixteen, 6-week old, athymic nude (female, nu/nu ((Harlan Sprague Dawley, Indianapolis, IN)) on opposite flanks (e.g., SW480^{vector} on the left and SW480^{Smad4} colon cancer cells on the right). The mice were followed every other day and the tumors were measured at weekly intervals and the mice were sacrificed uniformly at week 4. Tumors were assumed to be spheroid and volume was calculated by the appropriate formula (volume = 4/3 (Πr^2)). The Vanderbilt Institutional Animal Care and Use Committee approved all animal work.

Gene set analysis

WebGestalt (Zhang, Kirov, and Snoddy 2005) (*http://bioinfo.vanderbilt.edu/webgestalt*) was used to analyze the top signaling pathways in the refined, Smad4-modulated, epithelial-specific target list (n=787 probesets). This tool can be used to suggest biological areas that are important to a gene set and warrant further investigation. Kyoto Encyclopedia of Genes and Genomes (KEGG, Kanehisa Laboratories, Kyoto University, Japan) pathway function was utilized and the cut-offs of P<.05 (Fisher's exact test) and minimum of 5 genes per pathway were selected.

Statistical analysis

Microarray data were normalized with the Robust MultiChip Averaging (RMA) algorithm (Irizarry et al. 2003) as implemented in the Bioconductor package *Affy*. For pairwise group comparisons, t-test in the *Limma* package (Smyth 2004) in Bioconductor was used to identify differentially expressed probe sets between the two groups under comparison (e.g., $SW480^{vector}$ versus $SW480^{Smad4}$ or normal adjacent specimens versus stage I cancers). The implementation of t-test in *Limma* uses an empirical Bayes method to moderate the standard errors of the estimated log-fold changes, this results in a more stable inference, especially for experiments with a small number of arrays. The Wilcoxon rank sum test was used to assess the inverse relationship between Smad4 and β-catenin. To test if the direction of change between Smad4 and β-catenin was significant, we tested the interaction effect between genes (*Smad4, β-catenin*) and groups (normal, cancer). An FDR of 0.0005 was used as a cut-off to evaluate the potential enrichment of Wnt targets in the $SW480^{vector}$ and $SW480^{Smad4}$ comparison. Fisher's exact test was then utilized to determine if there was a significant enrichment upon comparison of the genes significantly differentially regulated by Smad4 and an annotated and updated Wnt target list. Smad4-modulated human tumor signatures were

generated by determination of genes differentially expressed for each of the three Smad4 probes on the Affymetrix® array. We focused on available 'grade A' Affymetrix® probes for Smad4 mapped to exonic transcript regions. The genes for the probe with the most significant number of differentially expressed targets (FDR<0.05) were then overlapped with the Smad4 gene expression profile from the cells to determine epithelial-specific Smad4 target genes. We then determined if there was a significant enrichment of Wnt target genes in this epithelial-specific list. We then used the Smad4-modulated, Wnt enriched gene list to execute unsupervised hierarchical cluster analysis and clinical outcome determination. Hierarchical clustering with complete linkage and Euclidean distance was applied to generate heatmaps. The three clusters were discovered by unsupervised, hierarchical clustering and Kaplan-Meier estimates were performed. The log-rank test was used to determine if there were significant differences across three clusters for survival outcomes. To assess the independence of cluster membership and cancer stages of the patients, we used the chi-square test. Student's t-test or ANOVA was used for qPCR and TOPflash assays and P-values were designated as indicated (*P<.05, **P<.01 or ***P<.001 and non-significant (P>.05, ns)). Difference in slope between groups in Fig. 17B was found by analysis of variance (no significant interaction between groups by dose or time was noted (P=.88)). GraphPad Prism® software was used for analysis of data in Figures 15; 17; 18; 20; 22, 26 and 27 (*P<.05, **P<.01, ***P<.001 and P>.05=ns by ANOVA and Student's t-test respectively).

Results

Smad4 Expression Reduces Transcription of β-catenin mRNA

SW480 cells are microsatellite stable, $18q21^{del}$ and harbor a splice site disruption in the remaining Smad4 allele (Woodford-Richens et al. 2001). In addition, these cells have additional inactivating mutations of the tumor suppressor genes, *APC* and *p53,* plus an activating mutation of *k-Ras (Capon et al. 1983; Rodrigues et al. 1990; Rowan et al. 2000)*. We previously reported that Smad4 restoration in SW480 colon cancer cells is associated with suppression of β-catenin/TCF-dependent expression of TOPflash reporter gene activity and reduction in steady-state levels of cellular β-catenin protein (Shiou et al. 2007); however, the mechanism by which this occurs is unknown. In the present studies, we found that expression of *β-catenin* mRNA is significantly reduced in SW480 cells in which Smad4 is either stably or transiently expressed (Figure 15A-C). These findings were reproduced by overexpression of Smad4 in HEK-293T cells demonstrating that the observed effect is independent of mutations in *APC and KRAS* (Fig. 15D). Thus, Smad4 expression reduces steady-state levels of β-catenin mRNA in both colon cancer cells and in epithelial, human embryonic kidney cells.

To determine whether the decrease in *β-catenin* mRNA levels is due to decreased transcription of the β-catenin gene *(ctnnb1)*, RNA polymerase II bound to the 2nd exon (+460 to +579) was assessed by chromatin immunoprecipitation (ChIP). A clear decrease in RNA polymerase II (POL2RA) binding at the *ctnnb1* gene was observed in $SW480^{Smad4}$ cells as compared with control $SW480^{vector}$ cells (Fig. 16A). These data indicate transcriptional suppression of *ctnnb1* upon Smad4 expression in SW480 cells.

The *ctnnb1* promoter/enhancer region (-4500 to +1) contains ~15 putative Smad binding elements (Fig. 17A, Appendix). The area around which we designed the *ctnnb1* promoter/enhancer region

primers is shown in Fig. 18A (see Appendix; ~ -3400 to -3000). Four minimal Smad binding elements are noted at this location (see bold gray text and bars). To determine whether the decrease in *β-catenin* mRNA transcription is associated with Smad4 interaction with the *ctnnb1* 5' upstream region, we performed a Smad4 ChIP experiment specific for the -3393 to -3136 5' region and observed a marked increase in Smad4 binding to this region in $SW480^{Smad4}$ compared to $SW480^{vector}$ cells (Fig. 16B, lanes 5 and 6). These data indicate that direct Smad4 interaction with the *ctnnb1* 5' upstream regulatory region may contribute to suppression of *β-catenin* mRNA transcription.

It should be noted that these effects on β-catenin mRNA were reflected at the protein level. Smad4 restoration in SW480 cells is associated with significant down-regulation of β-catenin in both whole cell and nuclear lysates (Figure 17A and B). Because others have reported that *β-catenin* mRNA stability may play a role in its activity (Gherzi et al. 2006), we assessed whether Smad4 restoration alters β-catenin mRNA stability in SW480 cells. $SW480^{vector}$ and $SW480^{Smad4}$ cells were treated with the RNA polymerase II inhibitor, 5,6-Dichlorobenzimidazole, 1-β-D-ribofuranoside (DRB) and *β-catenin* mRNA levels were assessed at 4 hour intervals. No significant differences were observed in β-catenin mRNA stability in the $SW480^{vector}$ cells compared with $SW480^{Smad4}$ cells in independent experiments using two different concentrations of DRB carried out to 24 hours (P=.88, Fig. 18B, Appendix). Thus, we concluded that Smad4 restoration in SW480 colon cancer cells does not reduce *β-catenin* mRNA by decreasing β-catenin mRNA stability.

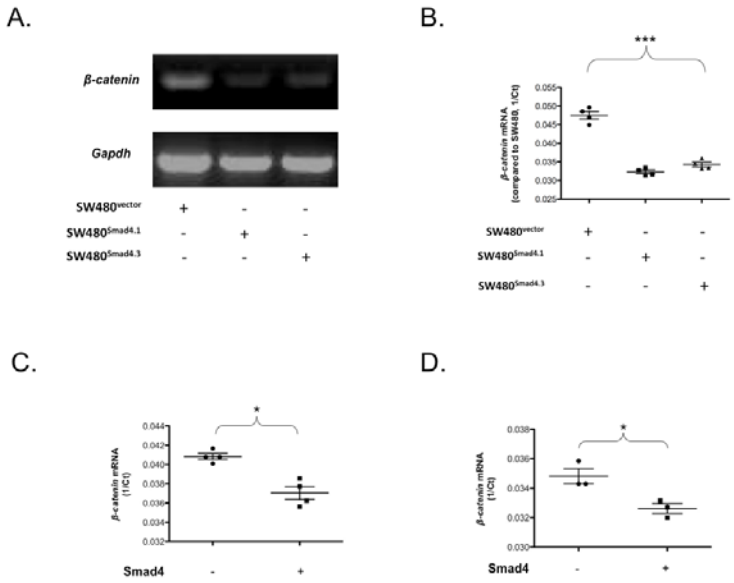


Figure 15. *Smad4 expression represses transcription of β-catenin mRNA*. (A) RT-PCR and (B) quantitative RT-PCR (qPCR) analysis of steady-state *β-catenin* mRNA levels in SW480^{Smad4} clones *vs*. SW480^{vector}. Each of two independent SW480^{Smad4} clones were compared to SW480 colon cancer cells (independent replicates are displayed; ***P<.001, ANOVA, mean +/- SEM). (C-D) Transient expression of Smad4 was carried out in SW480 cells and HEK-293T cells along with GFP. GFP positive cells were compared to GFP-negative cells and the difference in *β-catenin* mRNA was determined by qPCR (plotted as $1/C_t$, independent replicates are shown; *P<.05, student's t-test, mean +/- SEM displayed).

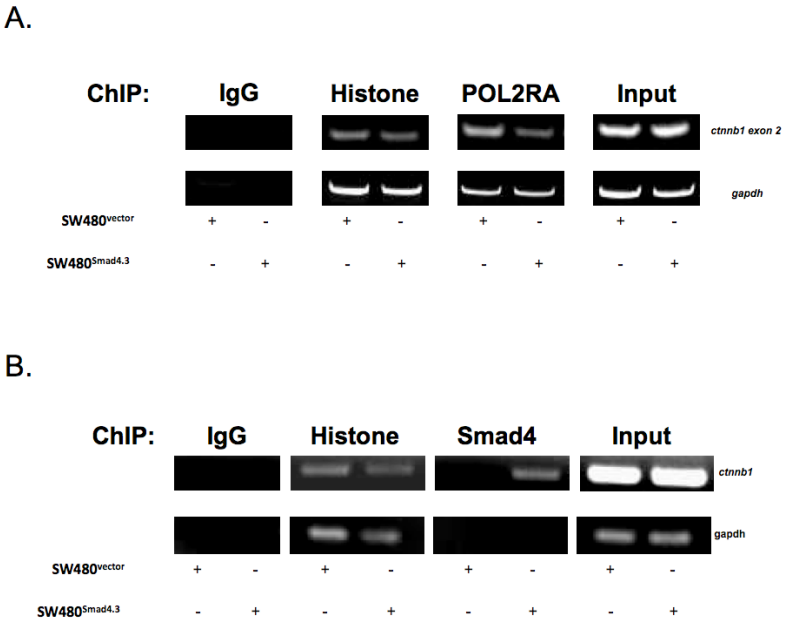


Figure 16. *Smad4 restoration in SW480 cells is associated with transcriptional down-regulation and binding to the ctnnb1 promoter/enhancer.* (A) Representative, gel resolved, 119bp PCR amplified bands from exon 2 of *ctnnb1* in a POL2RA Chromatin immunoprecipitation (ChIP) assay are displayed. Histone and IgG antibodies are used as positive and negative controls, respectively, and amplification of input DNA is shown on the right. Amplified bands for the control *gapdh* promoter (166bp) are also displayed. (B) Representative *ctnnb1* promoter-specific PCR amplified bands (225bp) from a Smad4 ChIP assay are shown on an agarose gel. IgG serves as a negative control and Histone H3 as a positive control for binding to the *ctnnb1* promoter. Amplification of input DNA using β-catenin primers is shown on the right hand side of the gel. Amplified bands for the control *gapdh* promoter (166bp) are also displayed. Lanes 1, 3, 5 and 7 represent SW480^{vector} cells while lanes 2, 4, 6 and 8 represent SW480$^{Smad4.3}$ as indicated (+ or -) for (A) and (B).

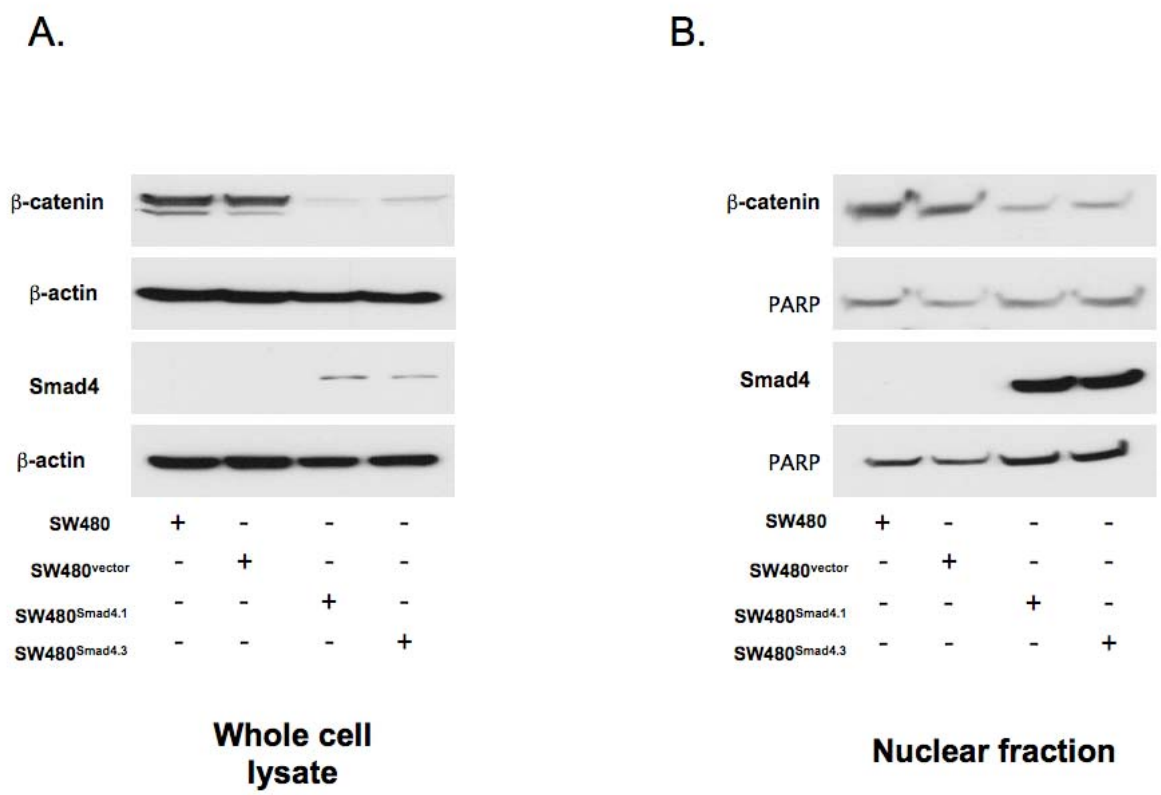


Figure 17. *Smad4 restoration in SW480 cells is associated with down-regulation of β-catenin mRNA which is reflected at the protein level.* Steady-state protein levels for β-catenin and Smad4 in cellular fractions by Western blot are displayed. (A) β-actin was the loading control for whole cell lysates. (B) PARP served as the nuclear marker. Antibodies were used as described in Methods.

Smad4 Repression of Wnt Activated Transcription is Independent of *APC* Mutation and Occurs in a Dose-Dependent Manner

SW480 cells lack functional APC and therefore exhibit relatively high Wnt transcriptional activity. We previously reported that restoration of Smad4 expression in SW480 cells causes a marked reduction in activity of the Wnt signaling reporter, TOPflash (Shiou et al. 2007). To further evaluate the inhibition of TOPflash by Smad4, we first confirmed that transient co-expression of Smad4 inhibits TOPflash activity in SW480 cells (Fig. 19A). We also determined that transient overexpression of Smad4 suppressed TOPflash activity in HEK-293T cells (wild type *APC*) in the presence or absence of Wnt3a ligand (Fig. 19B). Finally, we found that increasing doses of Smad4 in SW480 and HEK-293T cells results in increasing suppression of TOPflash activity (Fig. 19C and D). Thus, Smad4 expression suppresses TOPflash in a dose-dependent manner and it does so independently of Wnt pathway activation and in the presence of wild-type *APC*.

Smad4 Inhibition of TOPflash Activity is β-catenin Dependent

TOPflash activity is controlled by a transcriptional complex composed of several proteins, any of which might be altered by Smad4 expression to inhibit reporter activity. We next determined whether inhibition of TOPflash by Smad4 is β-catenin dependent. First, we co-expressed TOPflash along with wild-type β-catenin under the control of a constitutively active CMV promoter. We found that co-expression of wild-type β-catenin restored robust TOPflash activity in $SW480^{Smad4}$ cells to comparable levels observed in parental SW480 cells demonstrating that suppression of TOPflash activity by Smad4 is restored by expression of β-catenin with a heterologous promoter (Fig. 20A, p. 77). Next,

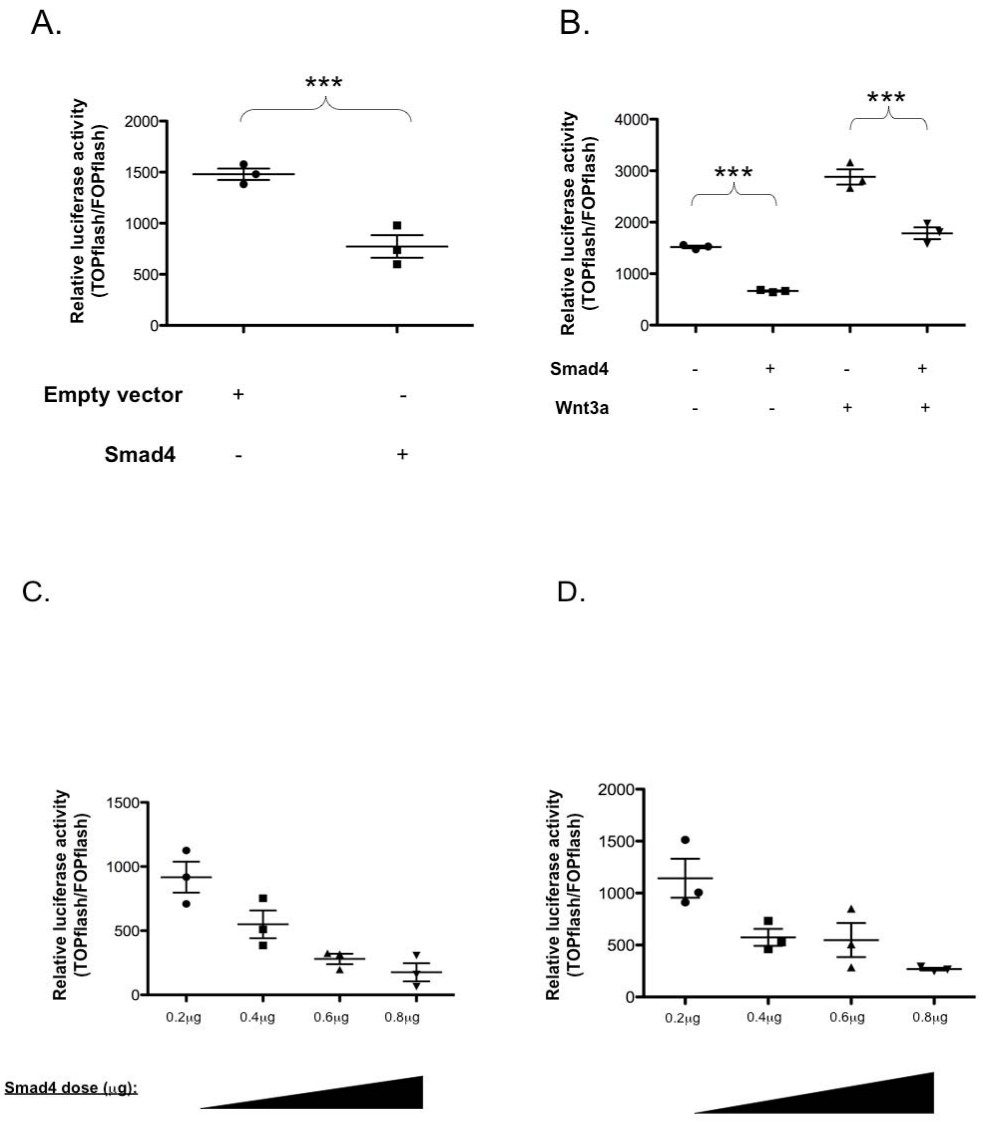


Figure 19. *Smad4 expression represses TOPflash activity in an APC independent manner.* (A) SW480 and (B) HEK-293T cells were transiently transfected with TOPflash and FOPflash reporter plasmids. The ratio of measured light units from TOP and FOPflash luciferase is graphed with either empty vector or Smad4 transfection as indicated (0.4μg used). Each graph is representative of at least 2 independent experiments (***P<.001, student's t-test, mean +/- SEM displayed). Wnt3a conditioned medium (1:1) was added to HEK-293T cells as indicated after transfection. (C) $SW480^{parental}$ and (D) HEK-293T cells were transiently transfected with TOPflash and FOPflash reporter plasmids. The ratio of measured light units (relative luciferase activity) from TOPflash luciferase and FOPflash luciferase is graphed with doses of Smad4 transfection as indicated.

we conducted a Smad4 concentration-response transfection experiment in SW480 cells with and without co-expression of wild-type β-catenin. As seen in Fig. 20B, transient Smad4 expression in SW480 cells suppressed Wnt-specific transactivation in a concentration-dependent fashion when no exogenous β-catenin was present. Consistent with the proposition that Smad4 directly suppresses transcriptional activity of the *ctnnb1* gene, Smad4 expression does not suppress TOPflash upon co-expression of exogenous β-catenin under the control of a heterologous promoter (Fig. 20C). These data further support a model in which Smad4 acts to inhibit Wnt-signaling through transcriptional repression at the β-catenin promoter.

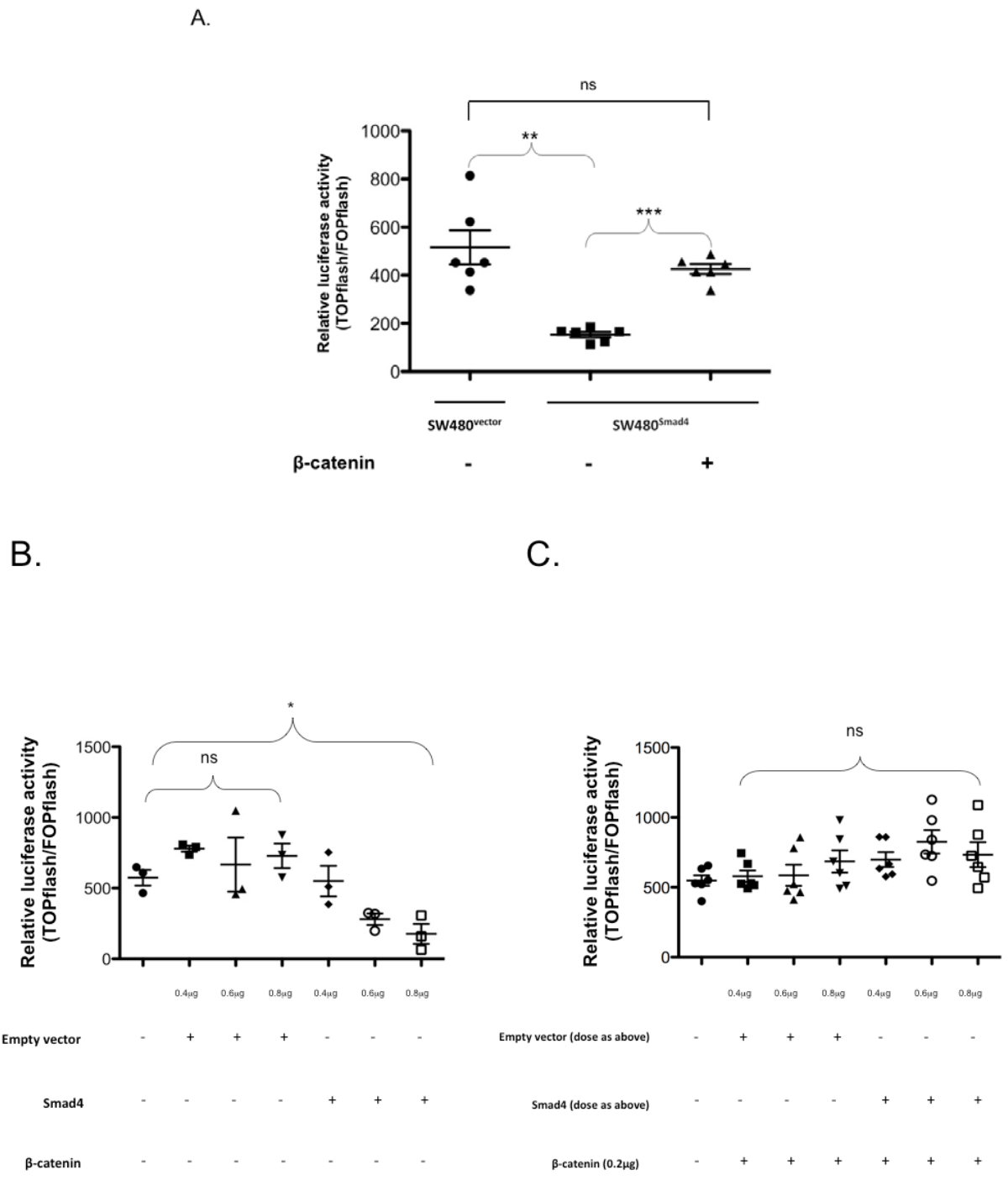


Figure 20. *Smad4 inhibition of TOPflash activity is β-catenin dependent.* (A) SW480^{vector} and SW480$^{Smad4.3}$ were transiently transfected with wild-type β-catenin (X*enopus* data shown and representative of independent replicates with human β-catenin). Luciferase activity, expressed in light units, from the TOPflash reporter is graphed for each condition (**P<.01 and ***P<.001, student's t-test, ***P<.001, ANOVA, mean +/- SEM displayed). (B) TOPflash activity was assessed in SW480 cells transiently transfected with control vector (0.4g, 0.6g and 0.8g, lanes 2-4), Smad4 (0.4g, 0.6g and 0.8g, lanes 5-7) and reporter plasmids (all lanes). Relative light units from TOP/FOPflash luciferase is displayed for one representative biological replicate (ns=P>.05, ANOVA for SW480 vs. control vector; *P<.05, ANOVA for SW480 vs. Smad4, mean +/- SEM displayed for all). (C) TOPflash activity was likewise assessed in SW480 cells transiently transfected with control vector (0.4 µg, 0.6 µg and 0.8 µg, lanes 2-4), Smad4 (0.4µg, 0.6µg, and 0.8µg, lanes 5-7) and also co-transfected with 0.2 µg wild-type human β-catenin (lanes 2-8). (two biologic replicates displayed, ns=P>.05, ANOVA, mean +/- SEM displayed).

We tested whether or not Smad4 represses TOPflash activity independently of β-catenin with use of the ΔN-Lef1-VP16 construct. This construct directs the expression of an N-terminal truncated form of Lef1 fused to a VP16 activator protein removing the β-catenin binding site on Lef1 and eliminating the requirement for β-catenin interaction with Lef1 to activate TOPflash. Thus, if Smad4 suppresses TOP-flash in the presence of ΔN-Lef1-VP16, we would conclude that Smad4-mediated inhibition of Wnt signaling is β-catenin independent, implicating other members of the Wnt-sigaling transcriptional complex. In these experiments, we found that Smad4 expression does not suppress TOPflash activity when ΔN-Lef1-VP16 is expressed (Fig. 21). These data confirm that attenuation of Wnt-specific transactivation by Smad4 restoration is β-catenin-dependent and cannot be attributed to altered activity in other members of the TCF/LEF transcriptional complex.

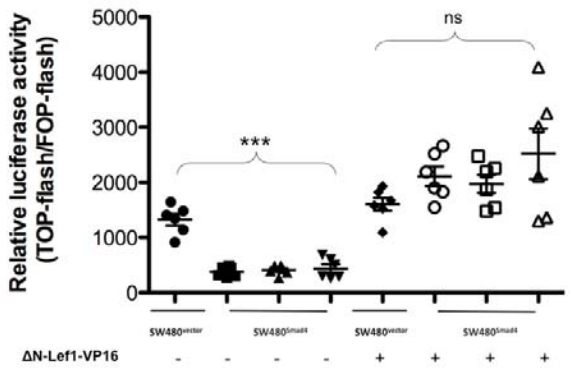


Figure 21. *Smad4 cannot suppress TOPflash independently of β-catenin.* TOPflash activity was determined as described in Materials and Methods in $SW480^{vector}$ cells (lane 1: TOP and FOP only; lane 5: TOP, FOP and the ΔN-Lef1-VP16 fusion construct) and $SW480^{Smad4}$ cells (lanes 2-4: TOP and FOP only; lanes 6-8: TOP, FOP and the ΔN-Lef1-VP16 construct); ***P<.001 and ns=P>.05, ANOVA for both, mean +/- SEM displayed.

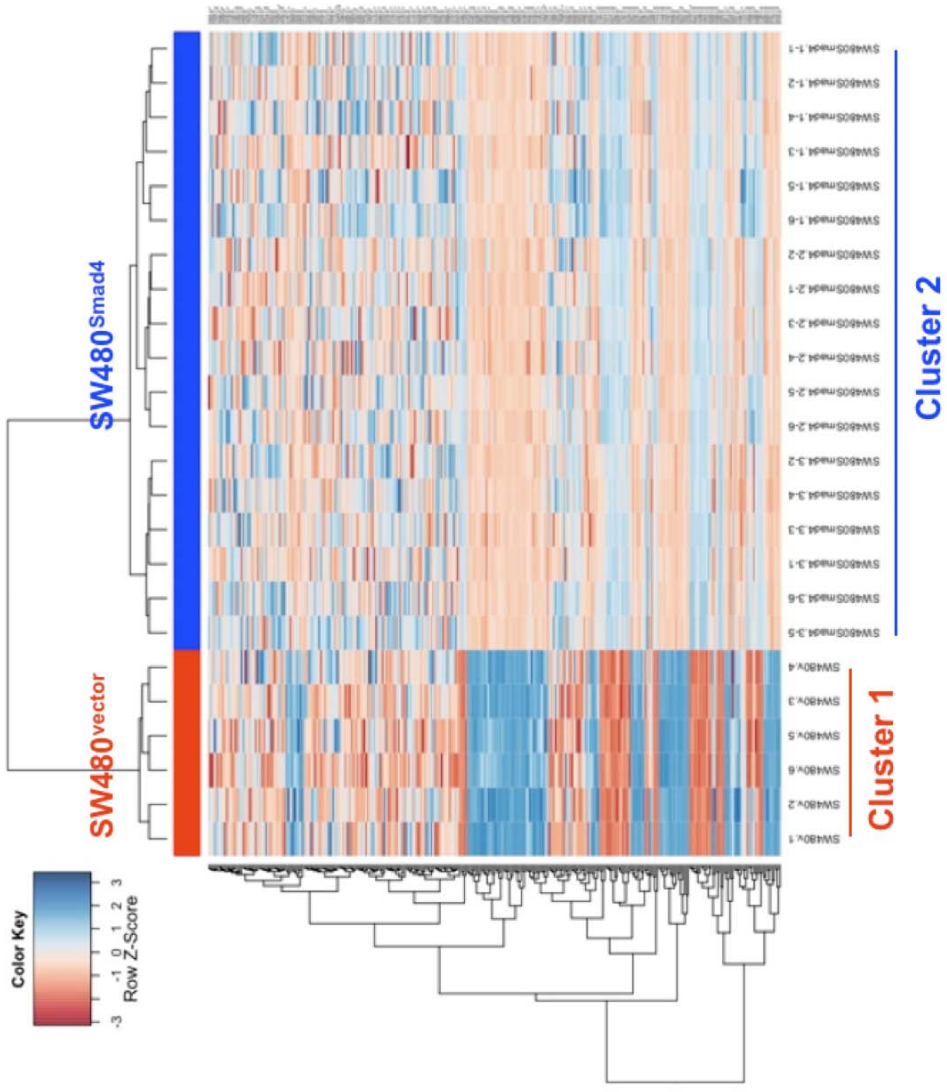
Smad4 Inhibits Wnt Target Gene Expression While Reversing EMT

Since we observed inhibition of β-catenin/TCF-dependent reporter activity in SW480 cells, we assessed whether Wnt target gene expression is also altered by Smad4 expression. Global gene expression analysis of $SW480^{vector}$ and $SW480^{Smad4}$ colon cancer cells by microarray was performed to assess changes in gene expression caused by restored Smad4 expression. We used the following stringent criteria to determine differential expression of $SW480^{vector}$ compared with $SW480^{Smad4}$ colon cancer cells: FDR<0.005 and fold-change >4 (see Materials and Methods) to obtain an epithelial-specific, Smad4 expression profile (see Table 11, Appendix).

Smad4 Inhibition of Wnt Target Gene Expression

We then queried this list of significantly altered genes for enrichment of published and annotated Wnt target genes (see Table 12 and Fig. 22 (schematic) in the Appendix). This Wnt target gene identifier list was mapped to Affymetrix® probesets and overlapped with the Smad4 expression profile (n=1668 probesets) to determine enrichment. We observed significant changes and enrichment in expression of Wnt target genes (P<.001) and EMT markers in the differentially expressed profile (see hierarchical cluster analysis in Fig. 23A) indicative of negative regulation of Wnt pathway targets. The Smad4-modulated, Wnt-enriched targets are noted in Table 13 (see Appendix). For example, expression of Wnt target genes known to be associated with negative regulation of the Wnt pathway, namely *E-cadherin (Jamora et al. 2003)* (also associated with reversal of EMT), *β-transducin repeat containing* (BTRC) and *dickkopf 1 (DKK1)* was significantly up-regulated by Smad4 restoration whereas expression of Wnt targets associated with pathway activation such as *Axin2, Lef1, Apcdd1 (Takahashi et al. 2002), CyclinD1 and Claudin1* were significantly down-regulated when Smad4 was expressed. We verified the microarray-predicted expression changes in *Axin2, Lef1, Apcdd1, CyclinD1 and E-cadherin* by RT-PCR and qPCR (Fig. 23B-C and data not shown).

A.


SW480vector
SW480Smad4
Cluster 1
Cluster 2
Color Key
Row Z-Score
-3
-1
0
1
2
3

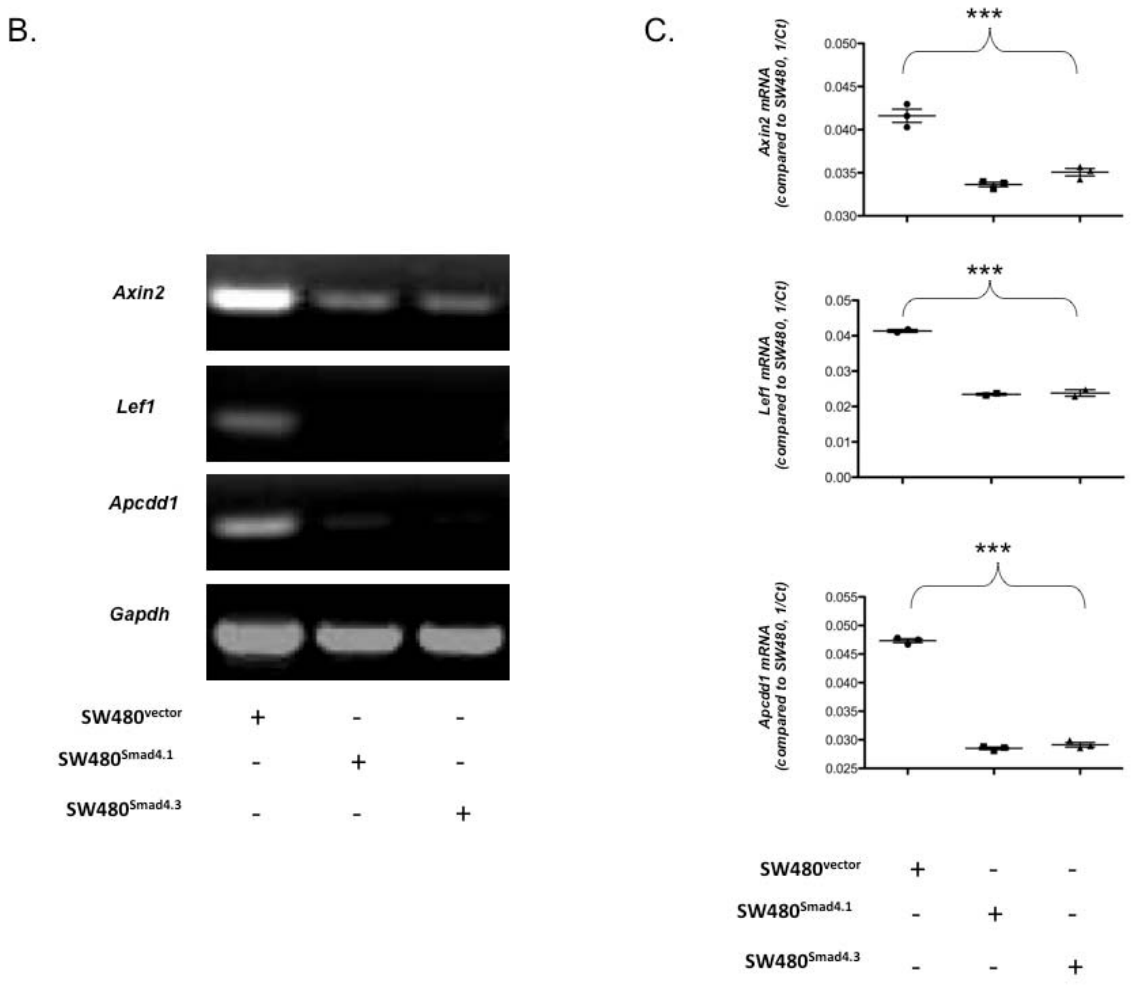


Figure 23*. A Smad4-specific gene expression profile reflects down-regulation of Wnt signaling.*
(A) Unsupervised clustering (p.80) of SW480^{vector} (indicated by red bar, n = 6) and SW480$^{Smad4.1\text{-}3}$ (indicated by blue bar, n = 6 each of 3 clones) in columns and Wnt target genes in rows are shown in the heatmap. Up-regulated genes are indicated by dark blue and down-regulated genes are indicated by dark red. (B) RT-PCR and (C) qPCR assessment of representative steady state Wnt target mRNAs (*Axin2, Lef1* and *Apcdd1*) in SW480^{Smad4} clones versus SW480^{vector} cells (three biological replicates (loading control, *Gapdh*)). For qPCR, SW480^{vector} and each of two independent SW480^{Smad4} clones were compared to the SW480 colon cancer cells. At least three biologic replicates were completed and representative biologic replicate data is shown (***P<.001, ANOVA).

Smad4 Expression is Associated with Reversal of EMT in Colon Cancer Cells

It is well established that Wnt signaling is deregulated secondary to an *adenomatous polyposis coli* (*APC*) mutation in >80% of colorectal cancers and recently, nuclear localization of β-catenin has been associated with EMT in this setting (Hlubek et al. 2007). As previously noted, the most commonly disrupted Smad mediator in cancer, including colorectal cancer, is Smad4, located on chromosome 18q21 (Hahn et al. Dpc4, a candidate tumor suppressor gene at human chromosome 18q21.1 1996; Thiagalingam et al. 1996). SW480 colon cancer cells are a useful model to investigate potential cross-talk between canonical Wnt signaling and TGFβ signaling in colorectal cancer for the following reasons: 1) They exhibit high baseline canonical Wnt activity due to a mutation in *APC*. This mutation constitutes a deletion at the carboxyl terminus at residue 1338. This missing area contains Ser-Ala-Met-Pro (SAMP) motifs otherwise known as SAMP repeats, which are important for Axin binding (Behrens et al. 1998) (the truncated protein still contains β-catenin binding sites (Yang et al. 2006)); and 2) Smad4 is mutated via loss of one allele of chromosome 18 with the second allele not expressed due to a splice site mutation (Woodford-Richens et al. 2001).

Smad4 restoration has been associated with induction of E-cadherin and P-cadherin in colon cancer cells and loss of Smad4 has been associated with loss of E-cadherin in colon cancer patients (Muller et al. 2002; Reinacher-Schick et al. 2004). These studies were some of the first to demonstrate the potential role of Smad4 in the reversal of EMT. Since β-catenin is part of the basolateral junctional complex along with E-cadherin and p120 catenin we sought to determine if Smad4 restoration in SW480 colon cancer cells is associated with a reversal of EMT and establishment of a functional cell membrane complex. We noted E-cadherin induction upon Smad4 restoration along with marked reduction in the intermediate-filament protein vimentin (Fig. 24A). Vimentin has been characterized as an EMT marker and a putative Wnt target gene (Gilles et al.

2003; Kalluri 2009). We analyzed SW480, SW480^{vector} and SW480^{Smad4} cells for localization and distribution of β-catenin, E-cadherin and p120 catenin with confocal microscopy. A marked increase in membrane-localized β-catenin in addition to a marked cuboidal phenotype was noted in SW480^{Smad4} cells compared with SW480^{vector} cells (Fig. 24B and C).

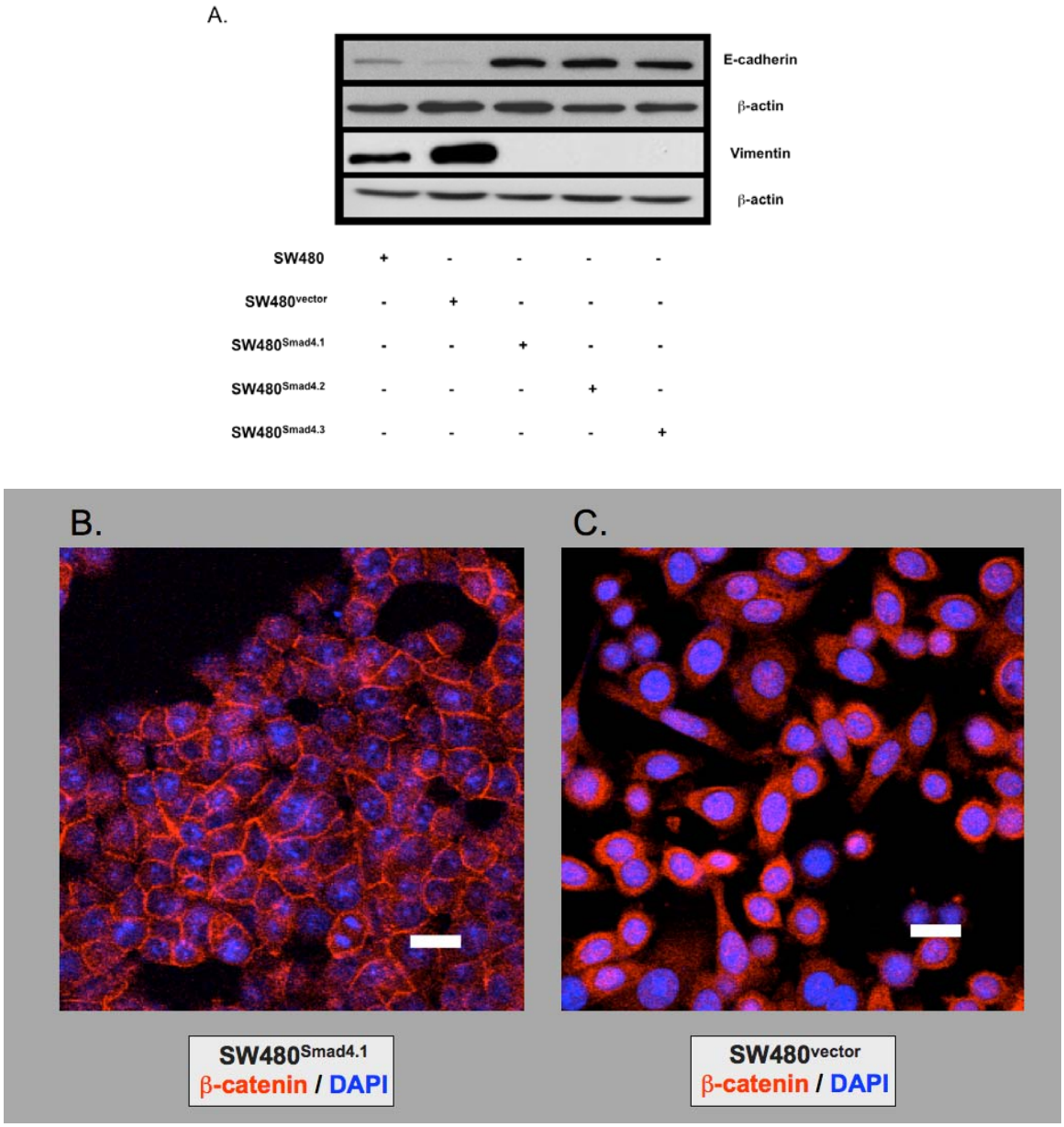


Figure 24*. Smad4 expression restores an epithelial cell phenotype in colon cancer cells.* (A) An immunoblot for SW480, SW480^{vector} and SW480^{Smad4} cells (SW480$^{Smad4.1\text{-}4.3}$) for E-cadherin and vimentin antibodies is shown (β-actin, loading control). (B and C) Confocal immunofluorescence of β-catenin (red) is shown to demonstrate junctional membrane localization in SW480^{Smad4} cells (left panel) compared with SW480^{vector} cells (right panel). The cell nuclei were stained with DAPI (blue). Images were acquired with the Olympus FV-1000 as described in Methods and scale bars are 20μM (white).

As can be seen in the vertical plane of the image (Fig. 25A, z-stack), β-catenin and E-cadherin demonstrate membranous co-localization upon Smad4 expression compared with lack of membranous co-localization of these components in the SW480^{vector} cells (Fig. 25B). These data strongly support functional reversal of EMT with re-introduction of Smad4 in SW480 colon cancer cells. Additionally, we noted markedly enhanced nuclear and cytoplasmic β-catenin in the SW480^{vector} cells along with poorly localized E-cadherin (Fig. 24C). These data support observations made by Hublek, et. al. where high nuclear β-catenin activity was noted at the invasive front of colorectal cancers (Hlubek et al. 2007). These data are supportive of the idea that a reduction in total levels of β-catenin mediated by the transcriptional repression via Smad4 is part of the Smad4-mediated tumor suppressor program in colon cancer cells.

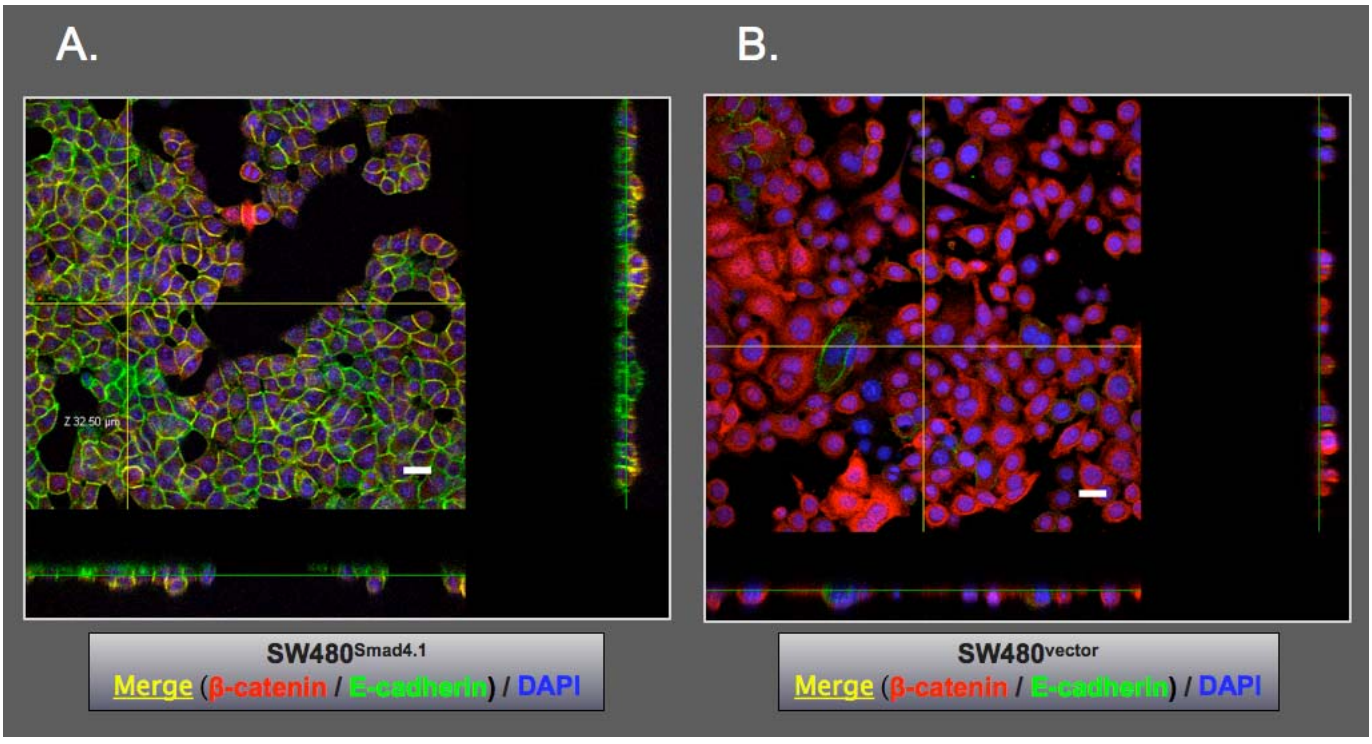


Figure 25*. Smad4 expression is associated with membrane localization of β-catenin and E-cadherin in colon cancer cells.* Confocal immunofluorescence for β-catenin (red) and E-cadherin (green) is shown for SW480^{Smad4} (A) and SW480^{vector} (B) cells. Images were acquired with the Olympus FV-1000 as described in Methods. The merged images along with DAPI (blue) are shown along with the z-stack (z height of 32.5μM is shown for both). Scale bars are 20μM (white).

Lastly, we noted that β-catenin co-localized with p120 catenin and that p120 catenin in turn co-localized with E-cadherin at the cell membrane (data not shown) in an independent $SW480^{Smad4}$ clone compared with the dense cytoplasmic and sparse membrane decoration of p120 catenin and β-catenin that we had observed in the $SW480^{vector}$ and SW480 cells (Fig. 26A and B, z-stack not shown). These data indicate that Smad4 restoration in SW480 colon cancer cells is associated with reversal of EMT.

Recent work by an independent group (Tian et al. 2009) recapitulates our previously published findings of suppression of TOPflash and down-regulation of the Wnt target gene Claudin-1 upon Smad4 re-expression (Shiou et al. 2007). This group did not demonstrate basolateral membrane component co-localization, nor do they explore further mechanistic insights into our initial observations. Our current data extend previous observations by multiple independent laboratories to demonstrate that Smad4 restoration in colon cancer cells reverses EMT and leads to an epithelial phenotype as supported by co-localization of basolateral membrane components in association with a significant down-regulation of β-catenin levels. In summary, as previously reported (Shiou et al. 2007), we observed marked up-regulation of E-cadherin in addition to the new observations of β-catenin and p120 cell membrane co-localization, reduction in nuclear β-catenin and down-regulation of the putative Wnt target and EMT marker, vimentin (Gilles et al. 2003), in response to restoration of Smad4 expression (Fig. 24-26).

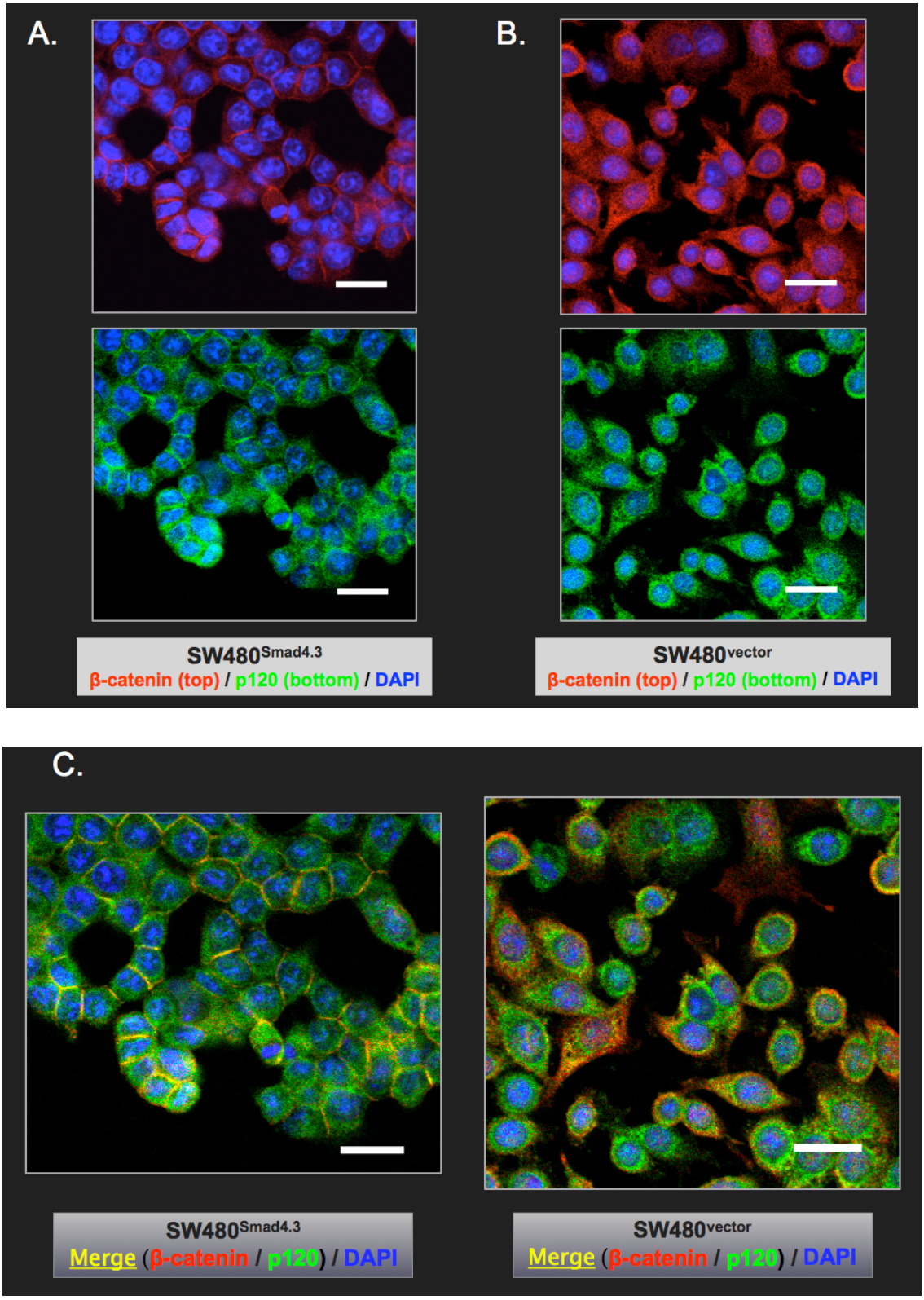


Figure 26. *Smad4 expression is associated with membrane localization of p120 catenin and β-catenin in colon cancer cells.* (A) Junctional localization of β-catenin (red) co-localized with p120 catenin (green) is show by confocal immunofluorescence in SW480^{Smad4} cells compared with (B) cytoplasmic and nuclear localization in the SW480^{vector} cells. (C) Merged images along with DAPI are shown as above. Images were acquired with the Olympus FV-1000 as described in Methods. Scale bars are 20μM for each image (white).

We further sought to determine the functional consequences of Smad4 restoration in SW480 colon cancer cells *in vitro* and *in vivo* with invasion and xenograft flank tumorigenicity assays. $SW480^{vector}$ cells are significantly more invasive in an invasion assay compared with $SW480^{Smad4}$ cells (Fig. 27, $P<.01$). Smad4 has been noted to decrease tumorigenicity in a unilateral flank model in nude mice (Schwarte-Waldhoff et al. 1999); however, this phenomenon has not been tested in a model that more closely mimics that of human tumor development and progression. We placed $SW480^{vector}$ cells on the left flank of 6-week old nude mice and $SW480^{Smad4}$ cells on the right flank. We did this so that each mouse would serve as it's on control and more closely resemble tumor cells in a human colorectal tumor where Smad4 loss and retention occur in the same organism. Tumors were monitored weekly over the course of 4 weeks and the mice were sacrificed and final *ex vivo* tumor volume was determined. $SW480^{vector}$ cells were significantly more tumorigenic than $SW480^{Smad4}$ cells in this bilateral flank xenograft model (Fig. 28, $P<.05$). These data further substantiate the tumor suppressive role for Smad4 in colon cancer and its potential role in the reversal of EMT in the context of deregulated canonical Wnt signaling in epithelial biology. These data substantiate our observations that Smad4 restoration and marked reduction in β-catenin mRNA levels and transcriptional activity mediate downstream inhibition of Wnt signaling and reverse EMT in colon cancer cells and corroborate the role of Smad4 as a tumor suppressor (Schwarte-Waldhoff et al. 1999).

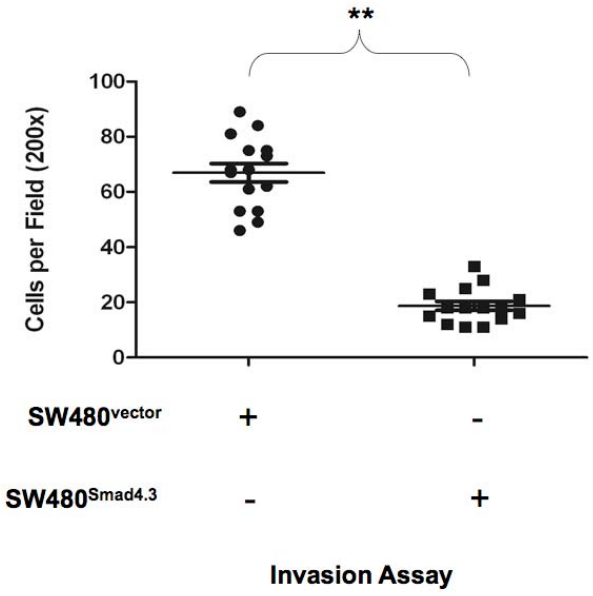


Figure 27. *Smad4 expression significantly decreases colon cancer cell invasion*. Matrigel invasion assays (representative biological replicates shown) demonstrate significant reduction of invasion ((proliferation control not shown) **P<.01, Student's t-test, mean +/- SEM displayed).

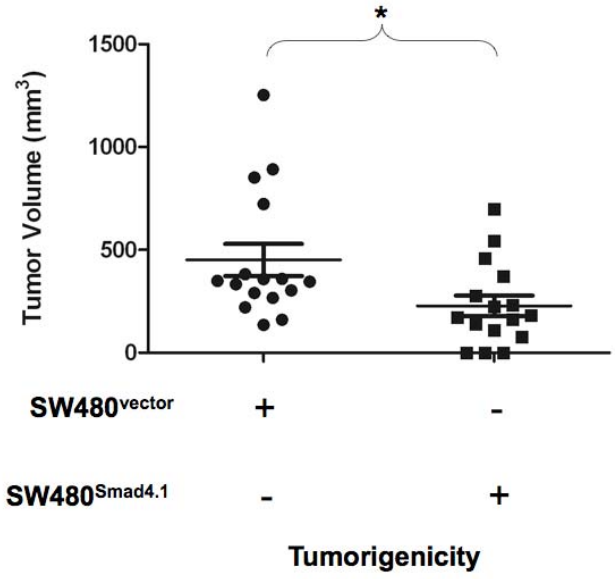


Figure 28. *Smad4 expression significantly decreases tumorigenicity in a nude mouse model.* Suppression of tumorigenicity upon Smad4 restoration in SW480 cells in a nude mouse model is displayed (tumor volumes calculated from post-sacrifice *ex vivo* measurements taken at 4 weeks post-injection of cells are shown (*P<.05, Student's t-test, mean +/- SEM displayed)).

Smad4 and β-catenin Levels Demonstrate Significant Inverse Expression Patterns in Human Colorectal Cancer Patient Specimens

In order to assess the clinical relevance of our cell culture studies, we examined the relationship between Smad4 and β-catenin expression in surgically resected colorectal cancer specimens. Smad4 and β-catenin expression were analyzed in tumor samples from 250 colorectal cancer patients as compared with ten normal adjacent colorectal tissue specimens, by microarray analysis in a combined dataset of patient samples from Vanderbilt Medical Center and the Moffitt Cancer Center (Table 14). As seen in Fig. 29A-B, we noted significant down-regulation of Smad4 and up-regulation of β-catenin in colorectal tumors, and this trend was noted for each stage of colorectal cancer as compared with normal adjacent mucosa. Furthermore, we noted highly significant inverse expression of Smad4 and β-catenin (Fig. 30, $P<.0001$). Further supportive evidence for this inverse expression pattern was noted when we stained normal adjacent colon and colon adenocarcinoma specimens for Smad4 and β-catenin. We noted positive Smad4 nuclear and cytoplasmic staining in the stroma and epithelium of normal adjacent mucosa while β-catenin demonstrated membranous staining along the crypt with nuclear and cytoplasmic staining at the base of the crypt (Fig. 31 A-B and native resolution in E-H). Interestingly, we observed negative staining for Smad4 in a colon cancer specimen except for sparse positive stromal cells. In contrast, we noted strong nuclear and cytoplasmic β-catenin staining in the serial section of the same tumor (Fig. 31 C-D and native resolution in I-J). These data support the clinical significance of the association between Smad4 down-regulation and a parallel up-regulation of β-catenin in the progression from normal colonic epithelium to colon adenocarcinoma.

TABLE 14

MICROARRAY STUDY DEMOGRAPHICS

	VMC (N=55 Colorectal adenocarcinomas)	*MCC* (N=195 Colorectal adenocarcinomas)	*MCC-NORMAL* (N=10 Normal adjacent colon specimens)
Mean Age (s.d.)	62.3 (14.1)	65.3 (12.9)	61.9 (16.3)
Sex (%male)	30 (54.5%)	106 (54.4%)	7 (70%)
Normal adjacent colon specimens	n/a	n/a	10
Stage I	4 (7.3%)	29 (14.9%)	2
Stage II	15 (27.3%)	61 (31.3%)	5
Stage III	19 (34.5%)	63 (32.3%)	3
Stage IV	17 (30.9%)	42 (21.5%)	0
Median follow-up (months, min-max)	50.2 (0.4-111.3)	44.97 (0.92-142.6)	n/a
Deaths	20 (36.3%)	64 (32.8%)	n/a
Caucasian (%)	50 (90.9%)	165 (84.6%)	9 (90%)
Black (%)	4 (7.3%)	11 (5.6%)	0 (0%)
Other (%)	1 (1.8%)	19 (9.8%)	1 (10%)

Human colorectal and normal adjacent specimen microarray dataset demographics. Colorectal cancer patients from Vanderbilt Medical Center (VMC, n=55)), Moffitt Cancer Center (MCC, n=195) and 10 MCC normal adjacent colon specimen patients used for microarray analyses are displayed. All patients were diagnosed with colorectal adenocarcinoma and staged according to American Joint Commission on Cancer (AJCC) guidelines (stages I-IV) and the 10 normal adjacent specimens were evaluated by a pathologist and determined to contain no adenocarcinoma contribution (only normal colonic mucosa). The normal specimens were normal adjacent colon mucosa specimens from patients whose colons were resected for colon cancer. VMC 55 includes 14 patients from the University of Alabama-Birmingham Medical Center. Other in the VMC medical record implies ‘not otherwise specified’ and implies ‘Hispanic, not otherwise specified’ in the MCC medical record.

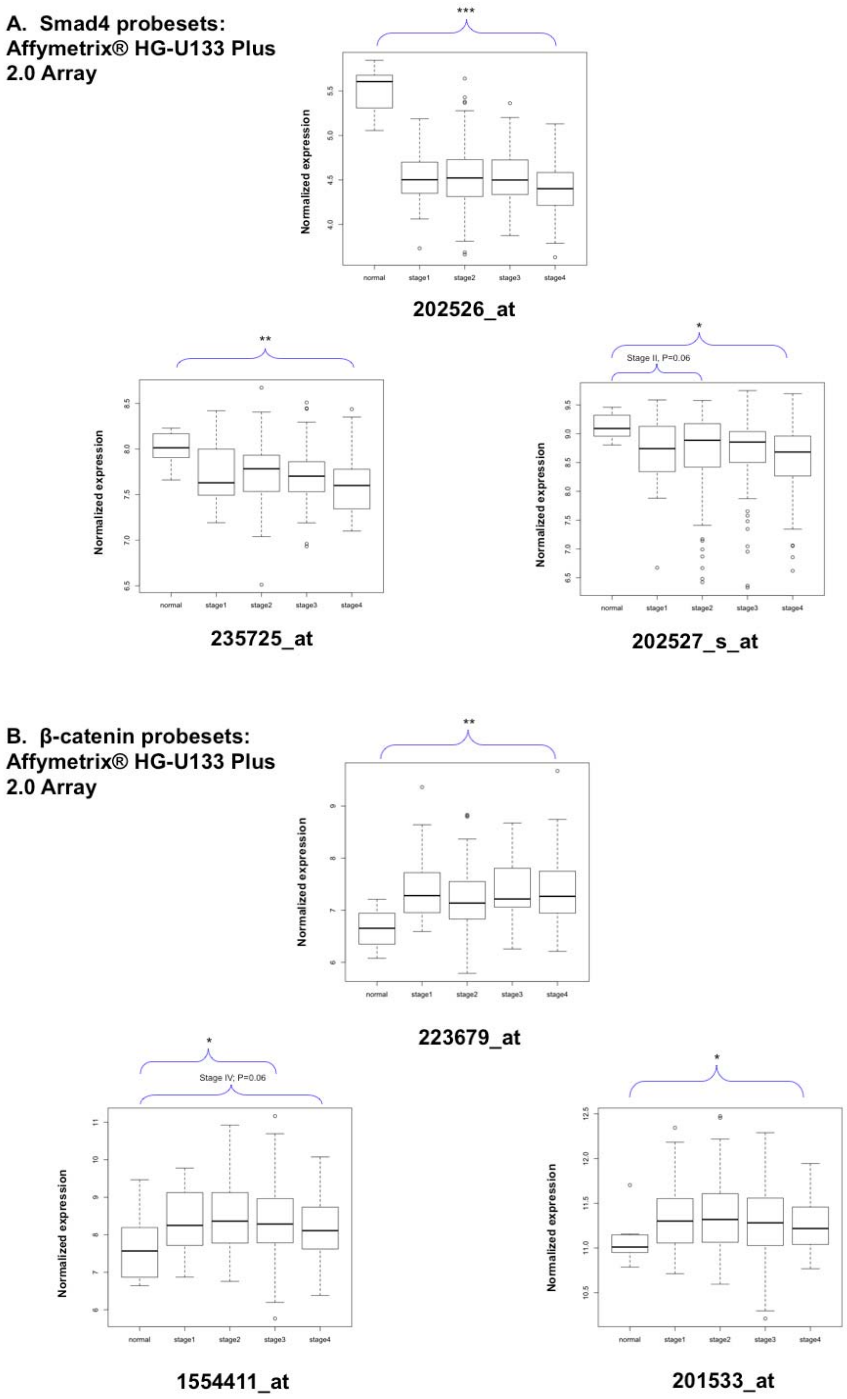


Figure 29. *(A) Smad4 expression is significantly down-regulated in colorectal cancer patient tumors compared with normal adjacent mucosal specimens.* Comparison of expression for exonic Smad4 probes on the Affymetrix® microarray platform for normal adjacent colon tissue versus colorectal cancers (stages I-IV) is shown for the combined VMC/MCC data (202526_at (**P<.001 for all comparisons); 235725_at (***P≤.01 for all comparisons); 202527_s_at (*P≤.02 for all comparisons except normal versus stage II)). *(B) β-catenin expression is significantly up-regulated in colorectal cancer patient tumors compared with normal adjacent mucosal specimens.* Comparison of expression for exonic β-catenin probes on the Affymetrix® microarray platform for normal adjacent colon tissue versus colorectal cancers (stages I-IV) is shown for the combined VMC/MCC data (223679_at (**P<.01 for all comparisons); 1554411_at (*P<.05 for all comparisons except for stage 4, P=.06); 201533_at (*P<.05 for all comparisons)).

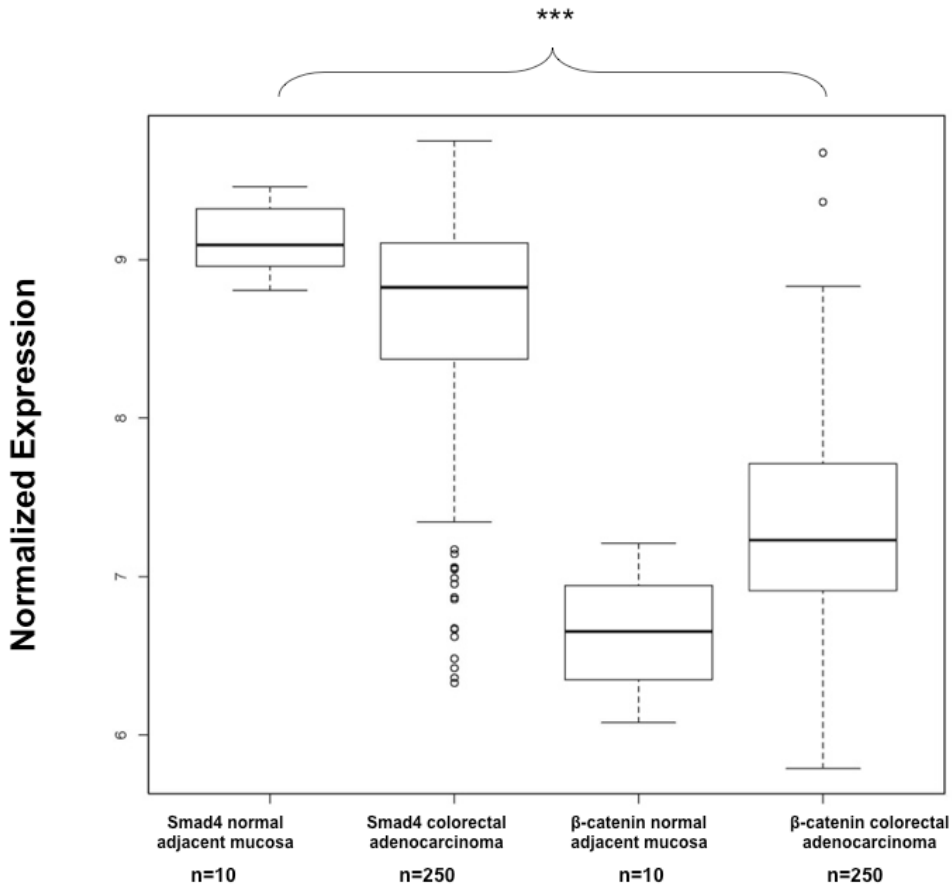
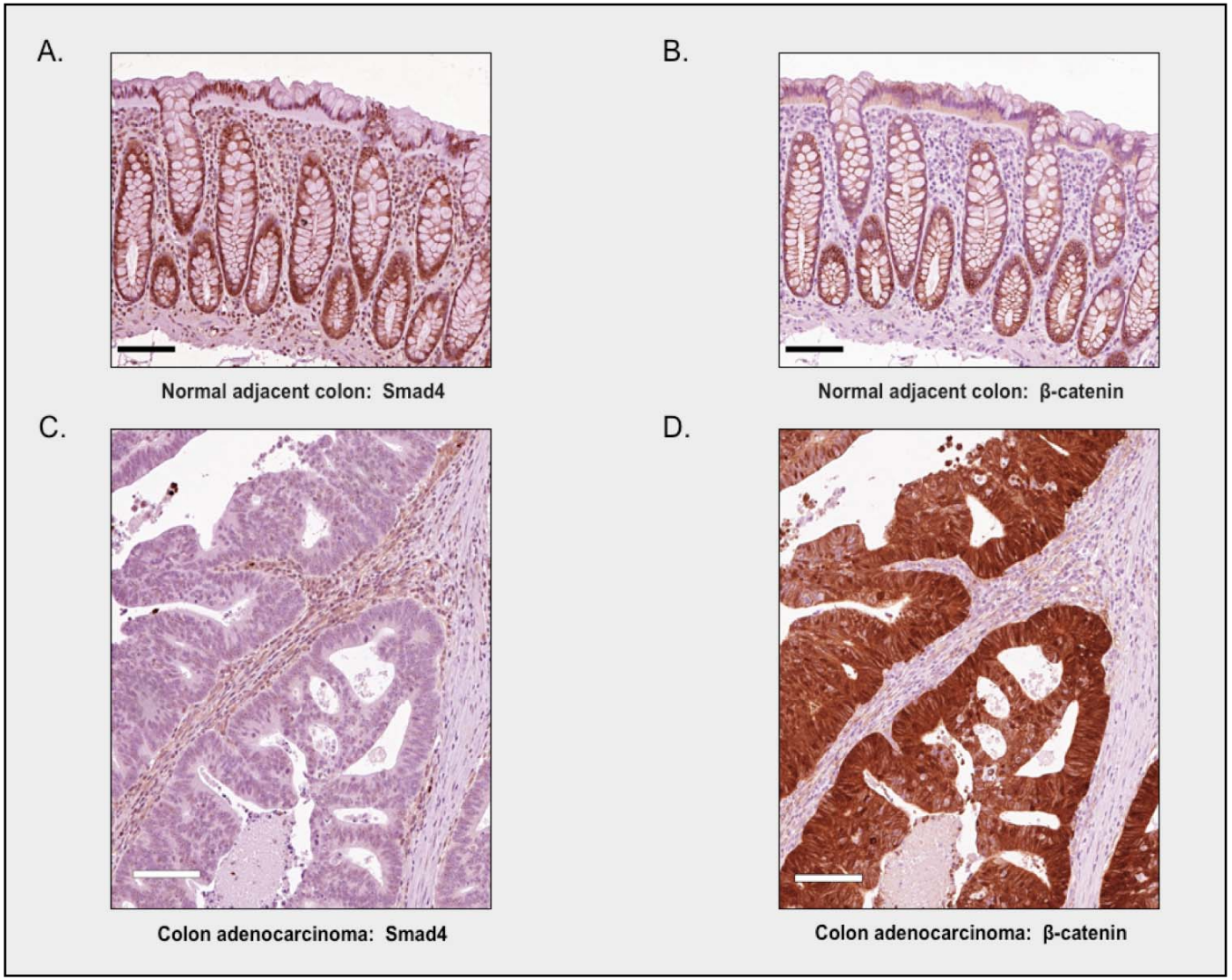
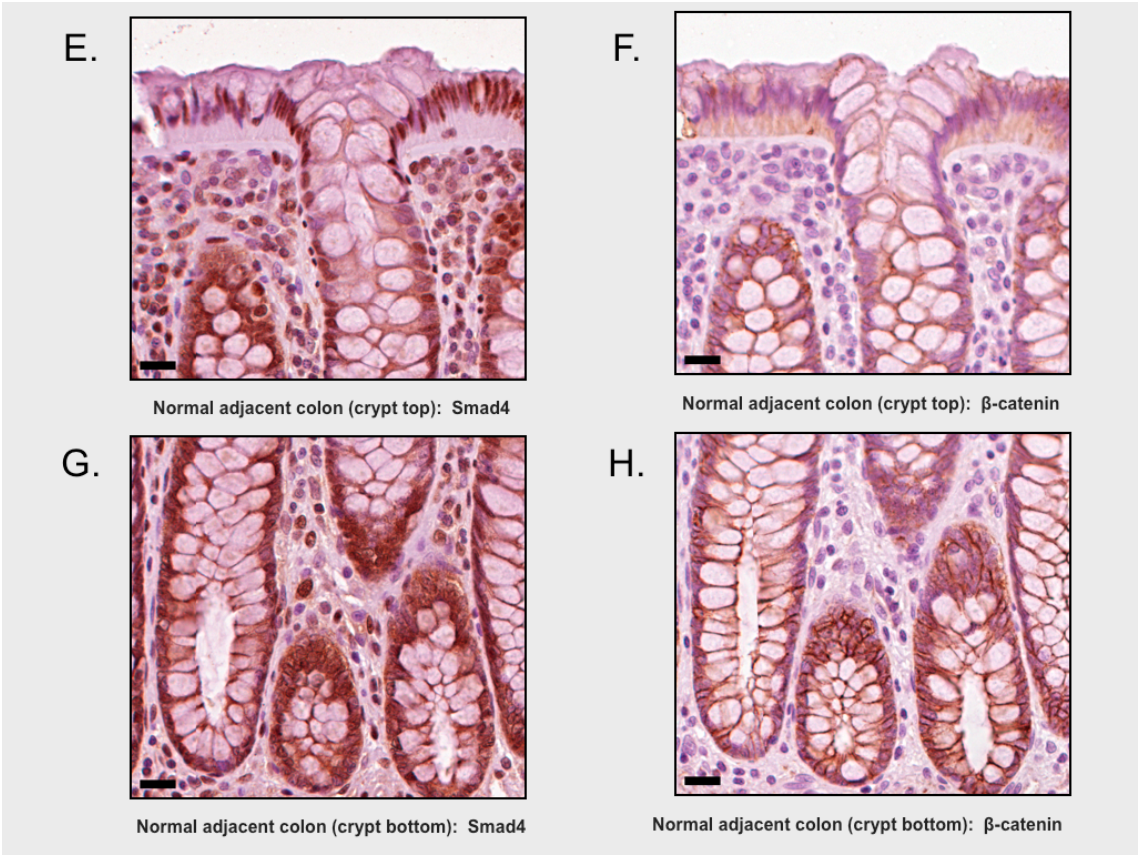


Figure 30*. Smad4 and β-catenin demonstrate significant inverse expression patterns in colorectal cancer patients.* 250 patients with colorectal cancer evaluated by microarray analysis are compared with normal adjacent colonic mucosa for Smad4 and β-catenin expression levels. Normalized expression values for normal tissue and colorectal adenocarcinomas (all stages) are displayed. The P-values of normal vs all colorectal cancers for Smad4 and for β-catenin are not shown on the graph (P=.01 for Smad4 and P<.001 for β-catenin, Wilcoxon rank sum test). The P-value for the interaction effect between Smad4 and β-catenin is displayed (P<.0001) to demonstrate the significant inverse expression patterns.


A.
Normal adjacent colon: Smad4
B.
Normal adjacent colon: β-catenin
C.
Colon adenocarcinoma: Smad4
D.
Colon adenocarcinoma: β-catenin



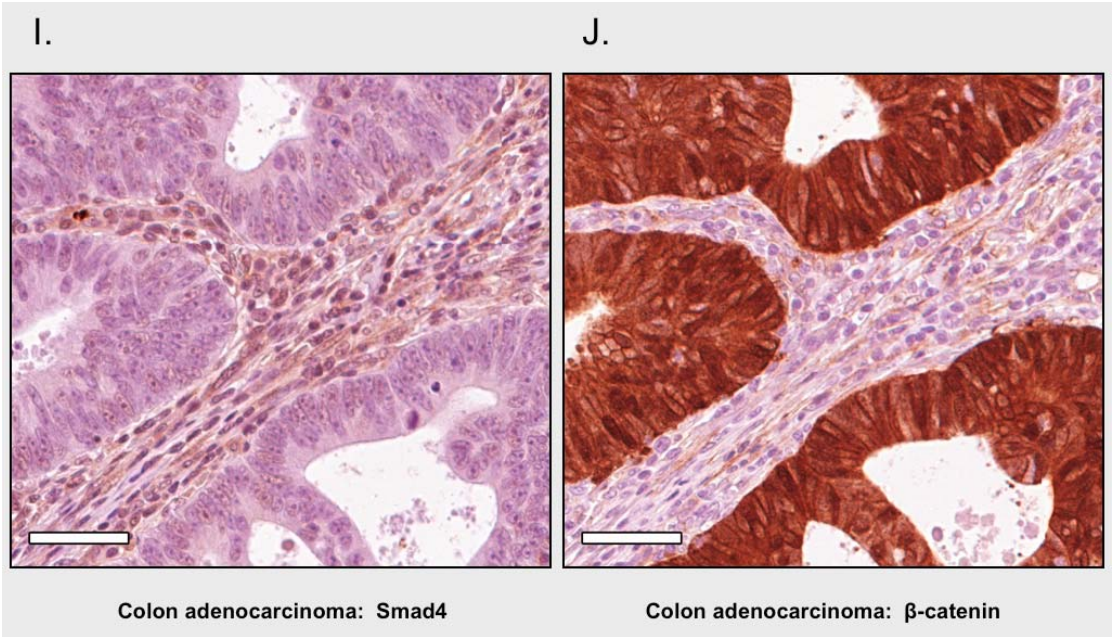


Figure 31*. Smad4 and β-catenin levels show inverse expression in colorectal cancer tumors.* Representative photomicrographs of normal adjacent colon (A-B) and a colon adenocarcinoma (C-D) stained for Smad4 and β-catenin are shown. Scale bars represent 100 microns (p. 93). Scaled images are shown in A-D (p. 93) and native resolution is shown in E-J (p. 94). Images were captured as a montage and selected areas in original resolution are displayed in E-J. All images were acquired with a 20X objective on a CoolSNAP-ES CCD camera (see Methods). Scale bars in E-H are 20 microns (normal) and in I-J serial sections are 50 microns (cancer).

A Smad4-Modulated, Wnt Target Gene Expression Pattern is Correlated with Outcome in Colorectal Cancer Patients

Given that Smad4 expression suppresses β-catenin mRNA and TOPflash activation and these changes are associated with EMT, we wanted to test whether a Smad4-dependent, Wnt transcriptional program contains predictive information for clinical outcomes in colorectal cancer patients. We reasoned that the human colorectal cancer microarray dataset would reflect a repressed Wnt transcriptional program since we observed a highly significant inverse pattern of expression between Smad4 and β-catenin. Thus, we conducted a stringent differential expression analysis between Smad4 high *vs*. Smad4 low patients in the primary tumor dataset (FDR<0.05). To generate this list, we focused on the Smad4 probe with the most significant number of differentially expressed genes. Next, in order to refine the list of differentially expressed genes for epithelial-specific targets, the differential expression profile from the colon cancer cell lines (SW480 and $SW480^{Smad4}$) was intersected with the list of differentially expressed genes in the patient tumor samples (Fig. 32A, Appendix). We queried this group of targets (451 mapped genes corresponding to 787 probes) in WebGestalt and discovered enrichment in the following signaling pathways: Mitogen-activated protein kinase (MAPK), TGFβ, Wnt, and T and B Cell Receptor signaling (Table 15, P<.05).

TABLE 15

SMAD4-MODULATED KEGG SIGNALING PATHWAYS

	Fisher's exact P-value
MAPK	<.01
TGFβ	<.01
Wnt	<.04
T Cell Receptor	<.01
B Cell Receptor	<.01

Signaling Pathways implicated in the human colorectal cancer Smad4-modulated, epithelial-specific co-expression module. The top 5 signaling pathways implicated by the Smad4-modulated, epithelial-specific global expression analysis (n=787 probesets) as found in WebGestalt with Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways are indicated. Cut-offs of P<.05 and 5 molecules per pathway were used. Fisher's exact P-value and number of implicated molecules is displayed.

We then asked if there was significant enrichment of Wnt target genes in this intersected list (see Fig. 32B, Appendix) and found that Wnt target gene representation was significant (P<.001) in this Smad4 co-regulated, epithelial-specific expression profile (32 probes, 18 distinct Wnt target genes are shown in Table 16). High Smad4 expression was associated with significant down-regulation of *Axin2* and *Claudin1* (both Wnt-induced) whereas low Smad4 expression was associated with significant up-regulation of *Lef1, Versican*, *Fibronectin and Twist* (all activated by Wnt in the absence of Smad4). *Claudin1* is associated with metastatic behavior in colorectal cancers (Dhawan et al. 2005), while *Fibronectin* and *Twist* are well-known markers of EMT (Yang and Weinberg 2008). These data compliment the cell culture array findings and provide a solid clinical parallel to our colon cancer cell line observations.

TABLE 16

SMAD4-MODULATED, EPITHELIAL-SPECIFIC
WNT TARGET GENES IN COLORECTAL CANCER PATIENTS

Gene symbol	Affymetrix ID	Fold-change
AXIN2	222696_at	0.71361550366
CLDN1	222549_at	0.74392453017
PITX2	207558_s_at	0.80555334013
BMP7	209590_at	0.83976253173
JUN	201466_s_at	0.86379993614
FGF18	211029_x_at	0.92172558033
ID2	201566_x_at	1.1043361046
FGF9	206404_at	1.1368378077
MET	211599_x_at	1.1766504952
LEF1	221558_s_at	1.2098730736
FN1	212464_s_at	1.2419832588
FGF20	220394_at	1.201414161
RUNX2	232231_at	1.2335540937
TWIST1	213943_at	1.2553732893
TNFRSF19	227812_at	1.2695599137
CYR61	210764_s_at	1.4294470874
TNFRSF11A	238846_at	1.4502180164
VCAN	211571_s_at	1.6280834871

Smad4-modulated, epithelial-specific Wnt target genes in human colorectal cancer. 32 Wnt target probesets (18 individual gene IDs displayed) in common amongst the SW480 Smad4-modulated, Wnt-enriched gene expression profile and the 202527_s_at (Smad4 probe) gene expression profile are displayed. Blue indicates down-regulation with high Smad4 levels and red indicates up-regulation upon low Smad4 levels for this display (the reverse scenario is not displayed). These Smad4 co-regulated Wnt enriched probes were then used to determine if they could separate patient groups based upon unsupervised hierarchical cluster analysis and subsequently to determine if they could predict outcome in colorectal cancer patients (see Fig. 31).

Since Smad4 is a tumor suppressor and has shown promise as a potential biomarker in colorectal cancer patients (Alazzouzi et al. 2005), we assessed whether the Wnt target gene probes co-regulated by Smad4 expression could identify distinct colorectal cancer patient groups in unsupervised cluster analysis. The Smad4-modulated Wnt target genes clustered patients into three distinct groups (Fig. 33A). Importantly, Wnt target gene expression exhibited a marked change in these groupings; for example, C*ysteine-rich angiogenic inducer, 61* (CYR61), *Versican*, and *Fibronectin* are markedly down-regulated in cluster 1 while *Lef1, Axin2* and *Twist 1* are up-regulated in clusters 2 and 3. These data are consistent with the notion that Wnt target genes previously implicated in processes of Wnt activation and EMT are down-regulated (e.g., *Fibronectin* in (ten Berge et al. 2008)) in the context of high Smad4 expression and up-regulated when Smad4 expression is reduced (e.g., *Twist1* in (Howe et al. 2003)).

In order to determine whether this Smad4 co-regulated gene list discriminates patient groups on the basis of clinical outcome, overall and disease-free survival analyses were conducted in the 250 patient colorectal cancer combined databases (VMC and MCC). We observed significant differences in outcomes for patients in each of these three clusters for overall survival (P<.001) and disease-free survival, (P=.01, Fig. 33B). These data suggest that although there are significant differences by overall survival for clusters 2 and 3 that these differences are diminished when recurrence is the primary outcome measure. One interpretation may be that loss of the Smad4-modulated transcriptional program is important for both overall survival (death from all causes including cancer-related deaths) and disease-free survival (evidence of recurrent disease) whereas the transcriptional program found in clusters 2 and 3 may become similar when the outcome measured is recurrence only. These differences in outcome could not be attributed to cancer stage as determined by analysis of stage distribution among each cluster of patients (P=.48). These data indicate that the epithelial

cell-specific transcriptional program regulated by Smad4 specifically related to suppression of Wnt signaling correlates with stage-independent prognostic information in colorectal cancer patients.

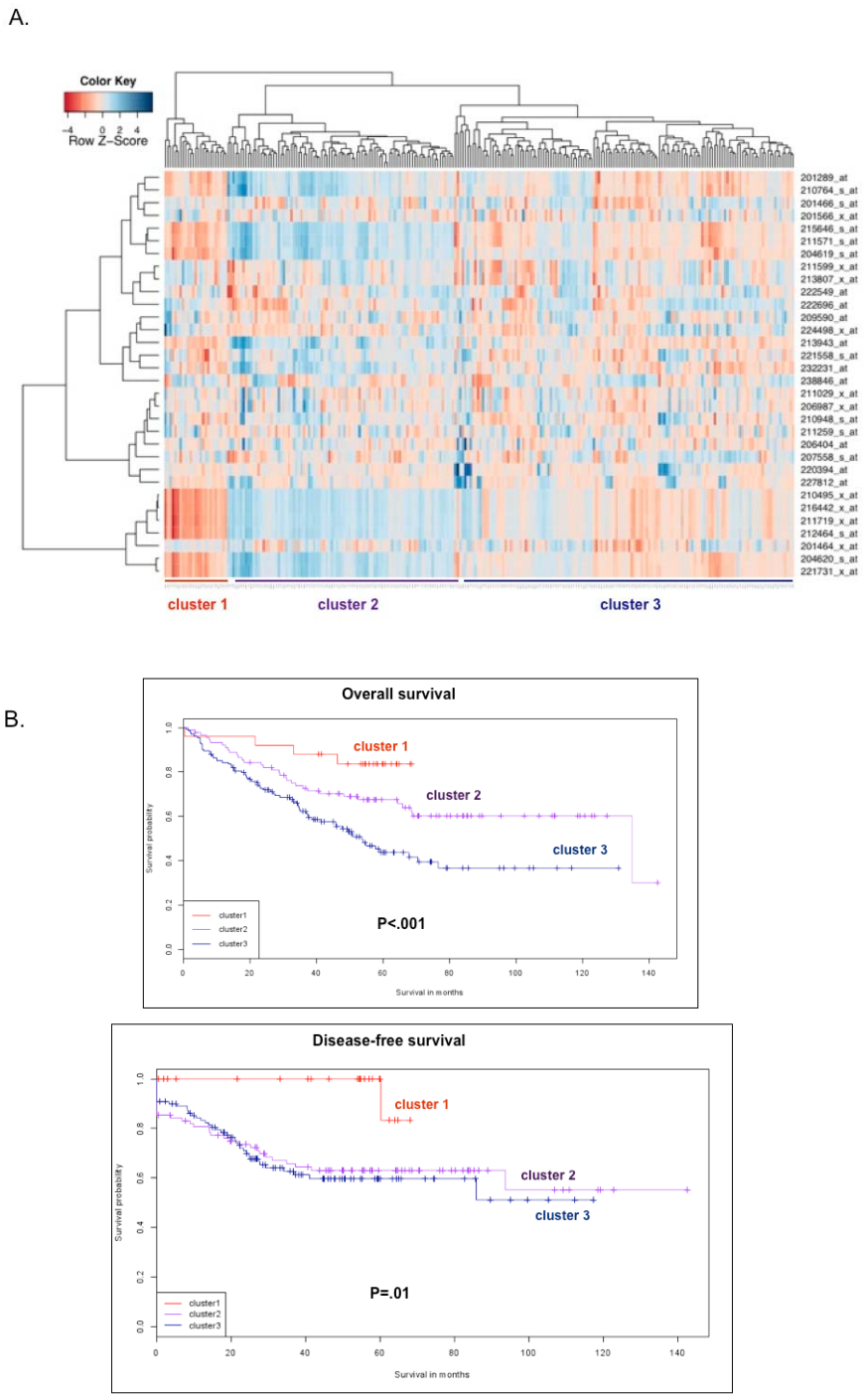


Figure 33. *Smad4-modulated Wnt target genes are significantly associated with better survival in colorectal cancer patients.* (A) Unsupervised cluster analysis of Smad4-modulated Wnt targets in 250 colorectal cancer patients. Individual clusters are denoted by red, purple and navy lines. Individual patients are represented by columns and Wnt specific Affymetrix probes are represented in rows. Up-regulated genes are indicated by dark blue and down-regulated genes are indicated by dark red. (B) Kaplan-Meier survival estimates for each cluster of patients are shown for overall and disease-free survival (OS, DFS). P-values shown compare all three clusters.

Discussion and Future Directions

In this chapter, we provide evidence that Smad4 is an important modulator of β-catenin gene expression and downstream Wnt signaling. Smad4 expression inhibits Wnt activity and associated Wnt target gene expresson patterns in colon cancer cells and in tumors resected from colon cancer patients. The clinical relevance of these findings was demonstrated by application of the Smad4-associated Wnt transcriptional program to colorectal cancer patient datasets.

Dynamic signaling interplay between Wnt and TGFβ superfamily members is evident in embryonic development and in homeostasis of the adult organism (Nishita et al. 2000; Radtke and Clevers 2005). For example, BMPs can suppress TCF/LEF transcription in hair follicles (Jamora et al. 2003) and can influence TCF4 transcription in muscle (Bonafede et al. 2006). Smads interact with TCF/LEF transcription factors in a cooperative fashion to activate target genes both developmentally (Nishita et al. 2000) and during carcinogenesis (Labbe et al. 2007). In other contexts, expression of a BMP antagonist, such as Noggin, is associated with increased nuclear and cytoplasmic β-catenin and with up-regulation of Wnt target genes (Kosinski et al. 2007). Independent studies also show that transgenic expression of Noggin *in vivo* is associated with increased expression of β-catenin (Haramis et al. 2004). Similarly, expression of other TGFβ superfamily antagonists is associated with EMT (Zeisberg et al. 2003). Notably, direct regulatory interaction between the central mediators of Wnt and TGFβ signaling, β-catenin and Smad4, has not been described, even though defects in both pathways contribute to more than half of all colorectal cancers.

Negative regulation of Wnt signaling can occur at multiple levels; however, the dominant paradigm for pathway regulation is through post-translational modification of β-catenin leading to proteasomal mediated degradation of intracellular levels to meter the Wnt signal (Clevers 2006). However, recent

work also suggests transcriptional activation of *β-catenin* as potentially important at the invasive front of colorectal cancers (Bandapalli et al. 2009; Nollet et al. 1996). Here, we provide additional evidence, in support of a new model for regulation of Wnt signaling in which Smad4 directly represses β-catenin transcription in epithelial cells.

Our data show that Smad4 associates with the 5' region upstream of *ctnnb1* suggesting that the suppression of *ctnnb1* transcription may be a direct effect of Smad4. Future work will determine precisely which *cis*-elements of the implicated area of the *ctnnb1* promoter/enhancer are necessary and sufficient for transcriptional repression induced by Smad4, and will determine which transcriptional co-regulators participate with Smad4 in this regulation (e.g., p300, Smads 1/5/8 or Smads 2/3). Additionally, determination of the ligand dependence of this effect remains to be addressed. For example, Smad4 restoration in colon cancer cells may enable autocrine TGFβ or BMP signaling ((Beck et al. 2006) and our unpublished observations) as a potential mechanism to further propagate resultant downstream tumor suppressor effects of Smad4. Our gene expression microarray observations of the significant inverse correlation of Smad4 and β-catenin expression in 250 colorectal cancer patients corroborated the observations we made in cultured colon cancer cells. In addition, analysis of the epithelial-specific, Smad4 co-regulated genes confirm that expression of Smad4 in colorectal cancers is associated with suppression of specific Wnt target genes. For example, we found that Wnt stimulated targets, *Axin2, Claudin1, Inhibitor of DNA binding 2, dominant negative helix-loop-helix protein (ID2)* and *Jun oncogene* are all significantly down-regulated when Smad4 expression levels are high in the primary colorectal tumors. Finally, we observed that the Smad4 co-regulated gene list contains stage-independent prognostic information that can be tested in future clinical datasets for potential use in prospective identification of high-risk colorectal cancer patients. These data indicate that the transcriptional profile associated with Smad4 loss may be more

informative than analysis of Smad4 loss alone in prediction of clinical outcomes in colorectal cancer patients.

In summary, our data uncover a previously unrecognized function of Smad4 in transcriptional repression of the *ctnnb1* gene. Our data also provide new insights into the modulation of the Wnt transcriptional program by Smad4 that has important implications for their cooperative roles in homeostasis in intestinal epithelium and tumorigenesis. This Smad4-modulated, epithelial-specific, Wnt-enriched gene signature has potential prognostic value beyond conventional pathological staging of colorectal cancer patients. These findings should also facilitate hypothesis testing for biologically targeted therapeutic interventions based on the level of activity of the TGFβ/Smad or Wnt signaling pathways.

CHAPTER IV

SUMMARY OF FINDINGS AND FUTURE DIRECTIONS

Brief Review

Colorectal cancer is the second leading cause of cancer-related deaths in the United States, and most of these deaths are a consequence of metastatic disease. Understanding the biology of metastasis and loss of tumor suppression is critical to identify patients at highest risk of cancer-related death. We used biological models to gain mechanistic insights into these areas to translate into optimal care of colorectal cancer patients. **Chapter II:** Staging inadequately predicts metastatic risk in colon cancer patients. We addressed this problem by developing a biological mouse model of metastasis where gene expression analysis led to the discovery of a metastasis-associated profile. Profile refinement in a colorectal cancer patient test set uncovered a 34-gene classifier that when translated to a metastasis score identified colon cancer patients in an independent test set at highest risk of death from recurrence. It also identified high-risk stage II and III patients, and importantly revealed low-risk stage III patients for whom adjuvant chemotherapy did not improve survival. This 34-gene classifier predicted poor outcome, independently of conventional measures, thereby providing insight into the biology of colon cancer metastasis. **Chapter III**: Loss of tumor suppressor genes is evident in the majority of colorectal cancer patients. TGFβ and Wnt signaling pathways are implicated in this process and are essential during development, epithelial cell differentiation and proliferation. Since we previously noted reduction in β-catenin levels and suppression of Wnt activity upon Smad4 restoration in colon cancer cells, we examined the mechanism by which Smad4

represses β-catenin expression and Wnt signaling. Smad4 expression results in transcriptional repression of β-catenin and thereby inhibits Wnt signaling, with associated inhibition of downstream Wnt target gene expression and reversal of epithelial-mesenchymal transition. The clinical relevance of this effect was demonstrated by an epithelial cell specific, Smad4-modulated gene expression profile associated with suppression of Wnt signaling, which contributed prognostic information independent of conventional pathological staging, in a large cohort of colorectal cancer patients. These findings should facilitate hypothesis testing for biologically targeted therapeutic interventions based on TGFβ/Smad or Wnt signaling pathway activity levels.

Future Directions for the 34-gene Classifier

Validation and Optimization of the 34-gene Classifier in Clinical Samples

We are currently working on three separate fronts to move the 34-gene classifier toward clinical application. First, we are developing an objective metastasis score for colorectal cancer. The validity of the proposed molecular predictor must be tested as an objective score that can be applied to individual patients prospectively. However, before this can be done we are using primary tumors resected from stage II and III patients with colorectal cancer with mature follow-up data (>3 years) to build a statistically robust model represented by data from four independent academic centers (VMC, UAB, MCC and The Ludwig Cancer Institute in Melbourne Australia). We will build the model in a combined dataset from the VMC/MCC groups (n=250 patients, see Chapter III) and test it in retrospectively collected, mature samples from UAB and Melbourne. We will combine the 34-gene classifier and several other credible colon cancer signatures (Barrier et al. 2006; Eschrich et al. 2005; Garman et al. 2008; Jorissen et al. 2009; Lin et al. 2007) proposed to predict a recurrence-prone

phenotype to systematically evaluate combinations of gene expression patterns for their association with colon cancer metastasis and recurrence according to the method proposed by Paik (Paik et al. 2004). The results of these analyses will be competed head-to-head with the biologically based 34-gene metastasis score. We will re-build the Cox Model (see Chapter II) in the combined 250 patient dataset and then test the results in an independent dataset made up of Melbourne and UAB patients for whom microarray data is readily available (n=180 patients). Our proposed sample size will allow at least 80% power to detect a hazard ratio of 1.70 between high and low score groups with an FDR = 0.003. These efforts will allow us to determine an objective metastasis score in a foundation of patients samples linked to mature clinical outcome in four independent academic settings to be carried forward in prospective analysis of the resultant metastasis score.

Secondly, we will determine the optimal platform for detection of the 34-gene metastasis score. Many technical issues stand in the way of the development of a robust clinical assay and include preservation of resected surgical specimens, amount and quality of purified RNA and reproducibility and robustness of the assay. We and our collaborators have overcome many of the collection issues and are now focusing on continual high quality RNA isolation from pathologically verified tissues that can be analyzed by the proposed assays (see below). We will compete head-to-head comparisons of expression values of the 34-genes from matched fresh frozen samples versus formalin fixed paraffin embedded (FFPE) tissues. Since routine tissue collection (e.g., community and regional hospitals) does not employ rigorous standardization methods as have been used to develop the original metastasis score (see Chapter II), a clinical test will be designed to detect the metastasis signature in FFPE tissues in order to be widely applicable. The details of this process are beyond the scope of this dissertation, but briefly we will use the following three competing platforms to determine the best method of detection: a) Applied Biosystems® qPCR; b) High Throughput Genomics' qNPA®

technology; c) GeneTitan based microarray with Affymetrix® technology. The major goals of this work will be to determine whether differential expression analysis is comparable in fresh frozen versus FFPE tumor tissue and to determine which platform for a clinical test is most reliable.

Lastly, we will test the optimized metastasis signature (determined as described above) in a blinded fashion on archival FFPE tissue collected in our multi-institutional collaboration, in order to determine the sensitivity, specificity, accuracy and clinical validity of the resultant score. This optimized signature will provide us with preliminary data that will be required to advance the work toward a prospective clinical trial to predict outcomes in stage II and III patients. The ultimate goal is to develop a deliverable assay on a robust platform that is ready to be applied to a prospective clinical trial. Future clinical trials would be based on collaboration with National Cancer Institute cooperative groups who have access to large numbers of patients across many academic centers and institutions.

Determination of the Primary Drivers of the 34-gene Classifier

We are concurrently in the process of determining which of the 34 genes in the metastasis score truly drive the predictive ability of the score. We are collaborating with our colleagues in biostatistics to model the 'driver' genes in the signature using the compound scoring method as applied to the combined VMC/MCC data and then will validate this in additional test sets. So far, our work has shown that 11 genes of the signature show equivalent predictive power compared to the 34 genes (preliminary data, not shown). We are also comparing the specificity and sensitivity of these 11 genes to the 34 genes in this dataset. Additionally, preliminary work has shown that the biologically-based metastasis score may be informative in other epithelial cancers such as breast and lung cancer (data not shown). We are in the process of finishing a manuscript to describe these findings.

Identification of Central Transcriptional Regulators in the 34-gene Classifier

In collaboration with colleagues in the biostatistics and bioinformatics departments at VMC we have begun to identify metastasis-related co-expression modules in a biologically meaningful way to reflect potential underlying regulatory mechanisms central to the transcriptional program driving the molecularly encoded, metastasis-prone phenotype. For example, the 300-gene metastatic gene signature (see Chapter II) has been used to identify 18 metastasis-related transcriptional modules that appear to be robust in prediction of colorectal cancer survival. These data are based on the idea that it is possible to identify regulatory mechanisms responsible for gene de-regulation in cancer signatures by searching for the unifying, coordinate transcription factor binding sites amongst genes in a gene expression classifier (Rhodes et al. 2005). Work to identify novel transcriptional regulators of the metastatic phenotype has begun and this work has subsequently been funded to further identify metastasis-related networks and screen for metastasis-related transcriptional modules using mouse and human colorectal cancer cell lines. These networks and transcriptional modules will then be tested and validated in our growing set of test and training datasets for clinical outcomes. For example, transcriptional factors discovered in this manner will be tested functionally via small interfering RNA and short hairpin RNA-mediated approaches to verify biological relevance in both mouse and human cell line models. Specifically, cell migration and invasion assays (see Chapters II and III) will be used to evaluate loss or gain of invasive capacity for each identified transcriptional factor that has been validated in a survival-prediction model. We will subsequently validate effects on apoptosis, proliferation and anoikis in addition to proceeding with *in vivo* experiments to confirm regulation of an invasive phenotype. Repressors of metastasis can be identified in a like manner and will be validated as above. Both metastasis effectors and metastasis repressors can then be validated in colorectal cancer samples by qPCR (and potentially immunohistochemistry) via our

continued collaboration with Dr. Kay Washington in Pathology. These approaches bring together the unique strengths of bioinformaticists, systems biologists, cancer cell biologists, biostatisticians, physician-scientists and clinical teams to address a complex clinical problem. The use of complimentary network-based models and biological models to determine functions of genes identified in metastasis-prone signatures will provide a means whereby valuable regulators of metastasis can discovered, validated and moved toward robust clinical application.

Differences and Overlap Between the 34-gene Classifier and Recently Published High-risk Gene Signatures for Colorectal Cancer Patients

As discussed in Chapter II, numerous gene expression signatures that identify high-risk colorectal cancer patients have been put forth as potentially useful in various prediction models. Unfortunately, prior prediction classifiers were determined on inconsistent microarray platforms to prevent a balanced comparison. In this section, I will focus on two of the most recent signatures, which have been developed on the same modern Affymetrix platform as described in Chapters II and III of this work, and their relevance and potential relationship with our 34-gene classifier. It should be noted that all other published gene expression signatures for colorectal cancer used computational methods to determine their classifiers, whereas we used a biological model refined with a comparative functional genomics approach to determine the 34-gene classifier.

The most recent unit of published work is from the Ludwig Colon Cancer (LCC) Initiative Laboratory in Australia (Jorissen et al. 2009). We note some interesting similarities between our proposed 34-gene recurrence classifier and their work. This group used a comparison between early stage (stage I) and late stage (stage IV) patients to determine a 128-gene signature that was

validated in an independent set of stage II and III colorectal cancer patients. Of note, some of the MCC patients used in this analysis were also used in both Chapters II and III of our work. Even though there is no specific gene overlap between our 34-gene classifier and this 128-gene classifier there is similarity in the biological processes implicated. For example, amongst the up-regulated genes in the LCC classifier, processes of extracellular matrix biology and embryonic development were implicated by Gene Ontology analysis (e.g., COL5A1 (collagen, type V, alpha 1), Integrins (ITGB1)), and similarly, in our use of Ingenuity Pathways Analysis we found that up-regulated genes were enriched for processes of cellular and embryonic development and connective tissue development / function (e.g., EGR1 (early growth response 1), HES1 (hairy and enhancer of split 1 (*Drosophila*), SPRY4 (sprouty homolog 4 (*Drosophila*)), ACTB (β-actin), PDLIM5 (PDZ and LIM domain 5)). The LCC down-regulated genes implicated processes of immune function as well as proteasomal genes (e.g. Immunoglobulins (IGHA1), chemokines (CCL28) and PSMB genes like PSMB9 (proteasome-related)). Our biologically-based signature derived from an immunocompetent mouse model showed that the down-regulated genes formed a network around Tumor Necrosis Factor (e.g., MMP13 (matrix metallopeptidase 13 (collagenase 3)) and HPSE (heparanase)) to suggest a prominent role for the immune system in the modulation of the metastasis-prone phenotype. We also found that the 300-gene metastatic gene signature had twelve genes in common with the 128-gene signature. These genes (e.g., PDGFC (platelet derived growth factor C), POSTN (periostin, osteoblast specific factor) and ITGBL1 (integrin, beta-like 1 (with EGF-like repeat domains)) are involved in processes of connective tissue development and cellular adhesion indicating that our approaches while different have implicated similar biological processes and networks that are important for the pathogenesis of a high-risk colorectal cancer phenotype. The discrepancy in overlap could be attributed to our use of a mouse model that may have missed human

genes important for the process of metastasis that were uncovered in the LCC signature. Likewise, it is possible that our approach uncovered novel regulators of metastasis with a purely biological basis that extend even beyond colorectal cancer epithelial pathobiology.

Another recent genomic approach to stratify colon cancer risk was put forth by a group from Duke University (Garman et al. 2008), wherein they developed a classifier that identified high-risk patients with computational models in early stage patients who had recurred and applied it to two independent datasets. They went on to show that this 50-gene classifier was able to identify compounds (e.g. COX-2 inhibitors, PI3K inhibitors) with biological activity against colorectal cancer cell lines in culture. Unlike our 34-gene classifier, they did not note any association with their signature and chemotherapeutic intervention in colorectal cancer patients. We found no intersection between the 50-gene classifier and our 34-gene classifier; however, upon further review it was interesting to note that the 50-gene signature implicated COX-2 (using a Connectivity Map approach (Lamb et al. 2006)), an important inflammatory mediator in colorectal cancer pathogenesis, and our signature uncovered a transcriptional module of down-regulated genes that suggest modulation of inflammation to be important for the metastatic phenotype. Prior to the publication of their 50-gene signature, I performed an independent analysis where the 34-gene signature was input into the Connectivity Map (Lamb et al. 2006), and I noted that the top compound implicated was acetylsalicylic acid (also known as aspirin, 18 Sept 2008 analysis, unpublished results, data not shown). These data compliment the Duke University group's findings. Even though we found some broad similarities, there was still minimal overlap amongst the 50-gene classifier and the 300-gene metastatic signature with only two genes overlapping, POSTN (periostin, osteoblast specific factor) and NDRG family member 2. POSTN has now been identified in three recent models and will surely be used in our development of an objective score as we more forward. Again, the minimal overlap

could be due to loss of important human genes with use of the mouse model, but it is encouraging to note that central processes amongst the analyses (e.g., inflammation and extracellular matrix / connective tissue biology) seem to remain relevant to inform factors critical for the identification of the high-risk colorectal patient. It is important to note that all three of these signatures are highly enriched for processes related to carcinogenesis such as cell-cell signaling, cellular development, cellular proliferation, cytoskeletal remodeling and modulation of inflammatory mediators. These data support our development of an objective score (as above) since it is likely that a combination of gene co-expression modules may uncover a superior gene expression classifier.

Lastly, since we used a comparative functional genomics approach originally modeled in the prediction of high-risk hepatocellular carcinoma (HCC) patients, we asked if there was any overlap amongst these pioneering studies (Lee et al. 2004; Lee et al. 2006) and our work. We found overlap amongst the genes downstream of FOS and JUN as they related to the hepatoblast subtype of HCC who had associated poor outcome (Lee et al. 2006). For example, ATF3 (activating transcription factor 3), TNC (tenascin C) and NR4A1 (nuclear receptor subfamily 4, group A, member 1) were found in both our 300-gene signature and the cluster associated with the hepatoblast sub-type of HCC. These genes are implicated in processes of apoptosis (e.g., NR4A1), extracellular matrices (e.g., TNC) and transcriptional modulation of apoptotic events and response to TGFβ signaling partners (e.g., ATF3). Indirect connections include PTGS2 (prostaglandin-endoperoxide synthase 2 (prostaglandin G/H synthase and cyclooxygenase), also known as COX-2) and MMP1 (matrix metallopeptidase 1 (interstitial collagenase)). This is interesting as an inflammatory central node that seems to become a common theme across this survey of prognostic classifiers. Also, MMP10, MMP12 and MMP13 were implicated in the 300-gene metastatic signature in our MC-38parental to MC-38met comparison as well as in the functional genomic enrichment of the rat hepatoblast

phenotype in poor prognosis patients with HCC. We noted very little overlap from the initial 2004 paper from this group other than enrichment of proteinases (MMPs) in the selection of optimal mouse models of human HCC (e.g., *Myc, E2f1* and *Myc E2f1* transgenic mice). In conclusion, gene overlap in the comparison of high-risk colorectal cancer and HCC patients reflect common processes on the pathway of carcinogenesis that were revealed in independent comparative genomic model systems (e.g., modulation of immune responses). These data support the use of these models, but suggest that specific carcinoma sub-types will result in distinctive gene expression profiles with common themes important for the metastatic process.

Interactions Amongst the 34-gene Classifier and the Smad4-expression Profile

As described in Chapter III, we discovered a Smad4-modulated, epithelial-specific, Wnt-enriched gene expression classifier that identified colorectal cancer patient sub-groups with regard to clinical outcome independently of conventional measures. One interesting notion is that this tumor suppressor-modulated gene expression profile could validate putative tumor suppressors and oncogenes in the 34-gene classifier. In regard to putative tumor suppressors, we noted that SYT17 (synaptotagmin XVII), MUM1L1 (melanoma associated antigen (mutated) 1-like 1) and HPSE (heparanse) were all up-regulated in the presence of Smad4. Very little is known about the 'syt-like' proteins. Synaptotagmins are thought to play a role in calcium sensing and neurotransmitter release and may be involved in calcium-dependent exocytosis (Koh and Bellen 2003). This might suggest that Smad4 has a role in processes of neural development and homeostasis which is supported in other neural environments as Smad4 has a transcriptional and synergistic effect with BMP2 on γ-aminobutyric acid (GABA)-modulated neuronal differentiation via Gat1 (GABA transporter sub-type I)

(Yao et al.). GABA is the chief inhibitory neurotransmitter in mammalian neurobiology. This is quite interesting as investigators from Baylor College of Medicine recently noted an association between neuronal ingrowth (e.g., perineural invasion) and tumor invasiveness / poor prognosis in colorectal cancer patients (Liebig et al. Perineural invasion in cancer: A review of the literature 2009; Liebig et al. Perineural invasion is an independent predictor of outcome in colorectal cancer 2009). This brings up the possibility of an undiscovered association with Smad4 loss and perineural invasion. One could speculate that Smad4 tempers neuronal outgrowth into epithelial structures via inhibitory action through GABAergic control of differentiated neurons. There is no published literature on MUM1L1; however, it is interesting to note that several melanoma-associated genes with generally unknown biological functions (e.g., MAGEB2 (melanoma antigen family B, 2), MAGEA3 (melanoma antigen family A, 3) are up-regulated in the presence of Smad4 in SW480 cells. These melanoma-associated genes, typically only expressed in the testes, have been implicated in a p53-dependent manner to apoptosis in colon cancer cells (Yang et al. 2007). HPSE is thought to play a role in invasiveness of renal cell carcinoma and has been found to be associated with increased tumor stage and disease-free survival (Mikami et al. 2008). These data provide some additional evidence to support the idea that down-regulation of these genes in the 34-gene classifier as putative tumor suppressors are important in the progression of the metastatic phenotype.

Regarding putative oncogenes, VDR (vitamin D (1,25- dihydroxyvitamin D3) receptor) was found in the genes down-regulated by Smad4 in SW480 colon cancer cells (see Chapter III, FDR<0.005, fold-change > 4). Interestingly, as noted in Chapter II there is now some data to support the role of VDR in EMT as related to ZEB1 expression (see page 57). As shown in Chapter III, Smad4 restoration in SW480 colon cancer cells is associated with down-regulation of β-catenin, inhibition of Wnt signaling and reversal of EMT *in vitro* and *in vivo*. The finding of VDR up-regulation in the highly invasive MC-

38met cells and as part of the 34-gene classifier indicate that the gene encoding the Vitamin D receptor may be an important target of Wnt/β-catenin signaling that plays an potential role in the acquisition of a more mesenchymal and invasive phenotype. Recent evidence suggests that VDR does interact with Wnt signaling components (e.g., TCF4, E-cadherin and Snail) in association with a less invasive phenotype (Beildeck et al. 2009; Pena et al. 2009), although it should be noted that the biological validation of the interaction in these studies was unimpressive. The potential connection between Smad4 and VDR is made more intriguing by the mounting level of evidence that implicates Vitamin D deficiency as a potential factor in colorectal cancer progression (Egan et al. ; Kure et al. 2009). Chronically low vitamin D levels could potentiate compensatory up-regulation of the vitamin D receptor in the setting of APC and Smad4 loss, which then promotes progression from dysplastic adenoma to carcinoma. Smad3-4 heteromeric complexes may bind the VDR promoter in COS cells in a synergistic manner requiring both TGFβ-responsive and Vitamin D-responsive promoter elements (Subramaniam et al. 2001). The biological effects of this transcriptional interaction are unknown, but open the door to interesting hypotheses in regard to the relationship with the central mediator of TGFβ signaling and the complex biology of VDR homeostasis.

Further Dissection of the Mechanism of Smad4 Inhibition of Wnt/β-catenin Signaling

Use of a nude mouse model as described in Chapter III could be used to provide *in vivo* evidence of the biological mechanism of Smad4 repression of β-catenin transcriptional activity. We could repeat the experiment as described in Chapter III, but this time use laser capture microdissection to isolate the invading front of the resultant tumors to prove whether Smad4 expression and the observed reduction in tumorigenicity is indeed due to repression of β-catenin mRNA at the invasive

edge of the tumors. We could also determine if there is indeed a down-regulation of β-catenin mRNA and protein in the Smad4 expressing tumors as indicated by qPCR, Western, and Immunohistochemistry/Immunofluorescence analyses. Another way to use this model would be to incorporate the use of a novel Wnt inhibitor developed by our collaborator Ethan Lee. We would be able to determine if Smad4 expression could potentiate suppression of tumorigenicity in the presence of the inhibitor or if suppression of tumorigencity occurs independently of administration of the Wnt inhibitor. Since the Wnt inhibitor prevents β-catenin-mediated Wnt transcriptional activity upstream of β-catenin transcription, we would be able to determine if Smad4 has an additional effect on β-catenin transcriptional activity downstream of this inhibitor to affect tumorigenicity and measures of invasion *in vitro*. Either way, we could determine whether Smad4 is required for suppression of Wnt activity in association with β-catenin repression *in vivo* and complimentary *in vitro* approaches.

Another *in vivo* model that we are currently developing is an inducible, epithelial-specific Smad4 knockout mouse. As we gain more numbers for each timepoint of inducible Smad4 depletion in the colony, we can determine if Smad4 loss alone is sufficient to promote up-regulation of β-catenin levels (mRNA and protein) in the areas of Smad4 loss. Another mouse model in development is the epithelial-specific Smad4 knockout mouse crossed with the $APC^{Min1638}$ mouse. This mouse could inform our hypothesis that Smad4 tempers progression toward a more invasive phenotype even in the absence of APC. It would also tell us if *APC* loss must occur in order to see the biological effects of β-catenin repression *in vivo*. Another interesting mouse model to directly inform whether this transcriptional repression occurs *in vivo* is to cross the epithelial-specific Smad4 knockout mouse with a TOPflash mouse. The development of this mouse is also underway in the laboratory and will provide an additional model whereby we can directly test our hypothesis that Smad4 does indeed

repress β-catenin transcriptional activity *in vivo* to promote a more epithelial phenotype and prevent progression to invasive adenocarcinoma via inhibition of Wnt transcriptional activation.

We are also collaborating with Kay Washington in Pathology to further examine a number of colorectal cancer specimens where the invasive front or leading edge of the tumor can be examined. We hypothesize that in the invasive, EMT-prone areas where Brabletz's group (Schmalhofer, Brabletz, and Brabletz 2009) has noted loss of E-cadherin and gain of β-catenin that we would observe loss of Smad4. Furthermore, we hypothesize that areas of colorectal cancer tumors that remain differentiated (membranous E-cadherin and β-catenin staining) also have retained Smad4 expression. We are also accumulating more colorectal cancer tumor specimens to determine if the loss of Smad4, gain of β-catenin and loss of E-cadherin can further inform colorectal cancer outcomes. It is certainly possible that the transcriptional program of Smad4 loss is more informative than simply looking at Smad4 gain or loss by immunohistochemical analysis alone.

Lastly, we have begun collaboration with local experts in transcriptional and protein regulation (Dr. Bill Tansey and Dr. Dan Liebler) to further investigate the mechanisms by which Smad4 represses Wnt/β-catenin signaling. For example, we will determine the transcriptional program regulated by Smad4 modulation of Wnt signaling in colorectal cancer cells by extending our ChIP findings with massively parallel short-read sequencing of DNA fragments obtained through ChIP-sequencing (ChIP-seq). We can conduct Smad4 and TCF4 ChIP-seq experiments where we anticipate that we will observe reduced TCF4-binding to Wnt target genes in the presence of intact BMP/Smad4 signaling. These data will provide more sensitive and quantitative support to our proposed mechanism and will allow sequence-specific detail that can inform critical co-regulators with Smad4 in this process of Wnt/β-catein suppression. Additionally, we will identify co-regulatory proteins involved in Smad4-driven repression of Wnt/β-catenin transcriptional activity in colon cancer cells. We will use

a shotgun proteomic approach (liquid chromotagraphy and tandem mass-spectrometry) to gain insight into the transcriptional regulatory protein complexes with high sensitivity. We can then use a small-interfering RNA approach to prove the necessity and sufficiency for identified co-regulatory proteins that may work in concert with Smad4 in repression of Wnt/β-catenin signaling. The multiple approaches at our disposal will provide important insights into the homeostatic, intersecting regulatory roles of the TGFβ and Wnt/β-catenin signaling pathways to illuminate our understanding of their contribution to the progression of colorectal cancer. These findings can then facilitate identification of new molecular targets to translate into therapeutic interventions that benefit colorectal cancer patients.

Summary

Colorectal cancer is the second most lethal, non-cutaneous epithelial cancer in the United States. Metastasis contributes to the majority of cancer-related deaths in this disease. Metastasis closely resembles the developmental process of epithelial-mesenchymal transition and reliable models to recapitulate this process can be useful to shed light onto the biology and prognosis of patients with colorectal cancer. Herein we describe a mouse model that led to the discovery of a gene signature of metastasis using comparative functional genomics, which identified stage II and III colon cancer patients prone to recurrence and death from metastatic disease, in addition to a low-risk sub-group of stage III patients for whom adjuvant chemotherapy provided no additional survival benefit. These findings form the basis for substantive pre-clinical biomarker testing and eventual translational application to a clinical trial. On the other hand, loss of tumor suppressor genes affects the majority of colorectal cancer patients to permit progression of disease. Many pathways are defective in the

pathogenesis of colorectal cancer; however, more than half of the patients have defects in the Wnt/β-catenin and Transforming Growth Factor-β signaling pathways, which are critical for development and intestinal homeostasis. This body of work also uses *in vitro* and *in vivo* models to describe a new role for the tumor suppressor Smad4 in the repression of β-catenin transcriptional activity in epithelial cells. This repression was associated with down-regulation of Wnt signaling and reversal of EMT. Clinical relevance of this effect was demonstrated by an epithelial cell-specific, Smad4-modulated gene expression profile associated with Wnt signaling suppression, which contributed prognostic information for colorectal cancer patients independently of pathological staging. These findings should facilitate hypothesis testing for biologically-targeted therapeutic interventions based on TGFβ/Smad and Wnt/β-catenin pathway activity levels. Overall, these results provide insight into the biology of metastasis and tumor suppression in colorectal cancer to promote seamless translation to care of the colorectal cancer patient at the bedside.

Appendix

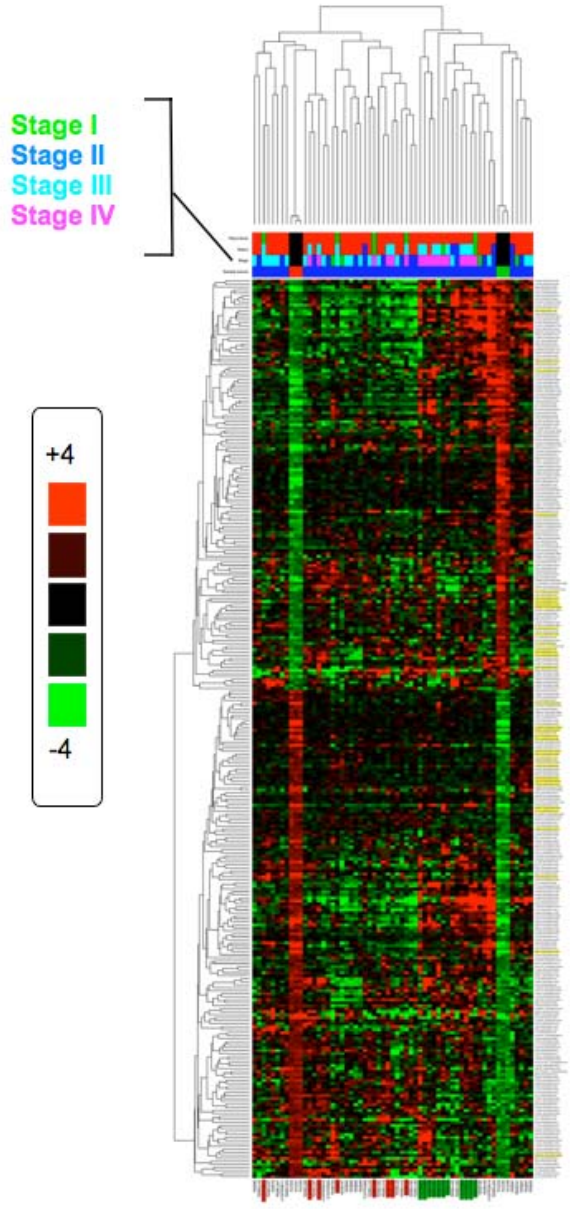


Figure 8. *Functional genomic clustering of the 300-gene metastatic signature*. Cluster analysis of mean-centered gene expression data (*rows*) with individual VMC patients (*columns*). Individual gene symbols in the metastasis-associated signature are listed at the right side of the dendrogram, and the 34 genes that make up the recurrence classifier are highlighted in *yellow*. Patient IDs are listed along the bottom. The patients who clustered with MC-38 parental cells are highlighted in *red* and those who clustered with MC-38met cells are highlighted in *green*. The heatmap key is 4-fold on the original signal intensity scale, which corresponds to 2-fold on a $\log_2$ scale. Key (1) Sample source: red, MC-38 parental; green, MC-38 invasive derivative (MC-38met); blue, patient samples; (2) Ensembl Human Gene identifiers: right side of heatmap.

TABLE 3

300 GENE METASTATIC SIGNATURE

Display ID	Gene Stable ID	Fold Change	Description (Protein name)
AQP1	ENSG00000106125	17.613	Aquaporin-1 (AQP-1)
OGN	ENSG00000106809	13.176	Mimecan precursor (Osteoglycin) (Osteoinductive factor) (OIF).
IL2RG	ENSG00000147168	13.041	Cytokine receptor common gamma chain precursor (Gamma-C) (Interleukin- 2 receptor gamma chain) (IL-2R gamma chain) (p64)
GRAMD1C	ENSG00000178075	11.665	GRAM domain-containing protein 1C.
AQP5	ENSG00000161798	10.694	Aquaporin-5 (AQP-5)
C2orf40	ENSG00000119147	8.53	Esophageal cancer related gene 4 protein precursor.
C15orf26	ENSG00000156206	6.934	Uncharacterized protein C15orf26.
EREG	ENSG00000124882	6.87	Epiregulin precursor (EPR).
OLR1	ENSG00000173391	6.555	Oxidized low-density lipoprotein receptor 1 (Ox-LDL receptor 1) (Lectin-type oxidized LDL receptor 1) (Lectin-like oxidized LDL receptor 1)
PTGS1	ENSG00000095303	6.348	Prostaglandin G/H synthase 1 precursor (EC 1.14.99.1) (Cyclooxygenase- 1) (COX-1) (Prostaglandin-endoperoxide synthase 1)
TMEM98	ENSG00000006042	6.279	Transmembrane protein 98 (Protein TADA1).
PDZRN3	ENSG00000121440	6.012	PDZ domain-containing RING finger protein 3 (Ligand of Numb-protein X 3) (Semaphorin cytoplasmic domain-associated protein 3)
C13orf33	ENSG00000102802	5.998	Uncharacterized protein C13orf33 (Activated in W/Wv mouse stomach 3 homolog) (hAWMS3).
IGFBP4	ENSG00000141753	5.842	Insulin-like growth factor-binding protein 4 precursor (IGFBP-4) (IBP- 4) (IGF-binding protein 4).
APOD	ENSG00000189058	5.723	Apolipoprotein D precursor (Apo-D) (ApoD).
DIO2	ENSG00000211448	5.179	Type II iodothyronine deiodinase (EC 1.97.1.10) (Type-II 5' deiodinase) (DIOII) (Type 2 DI) (5DII).
MYH10	ENSG00000133026	5.127	Myosin-10 (Myosin heavy chain 10) (Myosin heavy chain, nonmuscle IIb) (Nonmuscle myosin heavy chain IIb) (NMMHC II-b) (NMMHC-IIB)
MME	ENSG00000196549	5.066	Neprilysin (EC 3.4.24.11) (Neutral endopeptidase) (NEP) (Enkephalinase) (Neutral endopeptidase 24.11) (Atriopeptidase)
PRRX2	ENSG00000167157	5.022	Paired mesoderm homeobox protein 2 (PRX-2) (Paired-related homeobox protein 2).
PLA2G7	ENSG00000146070	4.896	Platelet-activating factor acetylhydrolase precursor (EC 3.1.1.47) (PAF acetylhydrolase) (PAF 2-acylhydrolase)
PID1	ENSG00000153823	4.646	PTB-containing, cubilin and LRP1-interacting protein (P-CLI1) (Phosphotyrosine interaction domain-containing protein 1)
SCARA5	ENSG00000168079	4.597	Scavenger receptor class A member 5 (Scavenger receptor hlg).
C14orf159	ENSG00000133943	4.261	UPF0317 protein C14orf159, mitochondrial precursor.
PDLIM2	ENSG00000120913	4.222	PDZ and LIM domain protein 2 (PDZ-LIM protein mystique) (PDZ-LIM protein).
NR4A2	ENSG00000153234	4.131	Orphan nuclear receptor NR4A2 (Orphan nuclear receptor NURR1) (Immediate-early response protein NOT)
AGT	ENSG00000135744	3.934	Angiotensinogen precursor (Serpin A8) [Contains: Angiotensin-1 (Angiotensin I) (Ang I); Angiotensin-2 (Angiotensin II)
SRPX2	ENSG00000102359	3.909	Sushi repeat-containing protein SRPX2 precursor.
FAM84A	ENSG00000162981	3.891	Protein FAM84A (Protein NSE1).

SYBU_HUMAN	ENSG00000147642	3.876	Syntabulin (Syntaxin-1-binding protein) (Golgi-localized syntaphilin- related protein).
CXCR7	**ENSG00000144476**	**3.854**	**C-X-C chemokine receptor type 7 (CXC-R7) (CXCR-7) (G-protein coupled receptor RDC1 homolog) (RDC-1) (Chemokine orphan receptor 1)**
ASPN	ENSG00000106819	3.79	Asporin precursor (Periodontal ligament-associated protein 1) (PLAP- 1).
C1QTNF3	ENSG00000082196	3.642	Complement C1q tumor necrosis factor-related protein 3 precursor (Secretory protein CORS26).
PCDHB15	ENSG00000113248	3.564	Protocadherin beta 15 precursor (PCDH-beta15).
SCAMP5	ENSG00000198794	3.553	Secretory carrier-associated membrane protein 5 (Secretory carrier membrane protein 5).
ODZ4	ENSG00000149256	3.486	Teneurin-4 (Ten-4) (Tenascin-M4) (Ten-m4) (Protein Odd Oz/ten-m homolog 4).
HS3ST1	ENSG00000002587	3.355	Heparan sulfate glucosamine 3-O-sulfotransferase 1 precursor (EC 2.8.2.23) (Heparan sulfate D-glucosaminyl 3-O-sulfotransferase 1)
HAS2	ENSG00000170961	3.347	Hyaluronan synthase 2 (EC 2.4.1.212) (Hyaluronate synthase 2) (Hyaluronic acid synthase 2) (HA synthase 2).
GGTL3	ENSG00000131067	3.285	Gamma-glutamyltransferase 4 precursor (EC 2.3.2.2) (Gamma- glutamyltranspeptidase 4)
ANKH	ENSG00000154122	3.285	Progressive ankylosis protein homolog (ANK).
SEMA3A	ENSG00000075213	3.193	Semaphorin-3A precursor (Semaphorin III) (Sema III).
ITGA8	ENSG00000077943	3.19	Integrin alpha-8 precursor [Contains: Integrin alpha-8 heavy chain; Integrin alpha-8 light chain].
MYBPHL	ENSG00000134222	3.145	Proline/serine-rich coiled coil protein 1.
TINAGL1	ENSG00000142910	2.994	Tubulointerstitial nephritis antigen-like precursor (Tubulointerstitial nephritis antigen-related protein) (TIN Ag-related protein)
PDE4B	ENSG00000184588	2.994	cAMP-specific 3',5'-cyclic phosphodiesterase 4B (EC 3.1.4.17) (DPDE4) (PDE32).
RRBP1	ENSG00000125844	2.991	Ribosome-binding protein 1 (Ribosome receptor protein) (180 kDa ribosome receptor homolog) (ES/130-related protein).
RENBP	ENSG00000102032	2.983	N-acylglucosamine 2-epimerase (EC 5.1.3.8) (GlcNAc 2-epimerase) (N- acetyl-D-glucosamine 2-epimerase) (AGE) (Renin-binding protein) (RnBP).
EDA2R	ENSG00000131080	2.967	Tumor necrosis factor receptor superfamily member 27 (X-linked ectodysplasin-A2 receptor) (EDA-A2 receptor).
MRGPRF	ENSG00000172935	2.908	Mas-related G-protein coupled receptor member F (Mas-related gene F protein) (G-protein coupled receptor 140)
NR4A1	ENSG00000123358	2.886	Orphan nuclear receptor NR4A1 (Orphan nuclear receptor HMR) (Early response protein NAK1) (TR3 orphan receptor)
CHRNA1	ENSG00000138435	2.873	Acetylcholine receptor subunit alpha precursor.
EVI2A	ENSG00000126860	2.869	EVI2A protein precursor (Ecotropic viral integration site 2A protein homolog) (EVI-2A).
AIG1	ENSG00000146416	2.843	Androgen-induced protein 1 (AIG-1).
AK1	**ENSG00000106992**	**2.817**	**Adenylate kinase isoenzyme 1 (EC 2.7.4.3) (ATP-AMP transphosphorylase) (AK1) (Myokinase).**
PTPLAD2	ENSG00000188921	2.807	protein tyrosine phosphatase-like A domain containing 2
ACTB	**ENSG00000075624**	**2.805**	**Actin, cytoplasmic 1 (Beta-actin).**
MGP	**ENSG00000111341**	**2.794**	**Matrix Gla-protein precursor (MGP) (Cell growth-inhibiting gene 36 protein).**
FAM43A	ENSG00000185112	2.794	Protein FAM43A.
CRYAB	ENSG00000109846	2.776	Alpha crystallin B chain (Alpha(B)-crystallin) (Rosenthal fiber component) (Heat-shock protein beta-5) (HspB5)
HES1	**ENSG00000114315**	**2.77**	**Transcription factor HES-1 (Hairy and enhancer of split 1) (Hairy- like) (HHL) (Hairy homolog).**
PCDHB16	ENSG00000196963	2.749	Protocadherin beta 16 precursor (PCDH-beta16) (Protocadherin 3X).
NIPSNAP1	ENSG00000184117	2.74	Protein NipSnap1.
EMB	ENSG000001	2.712	Embigin precursor.

	70571		
ALDH3A1	ENSG00000108602	2.684	Aldehyde dehydrogenase, dimeric NADP-preferring (EC 1.2.1.5) (ALDH class 3) (ALDHIII).
ANGPT1	ENSG00000154188	2.659	Angiopoietin-1 precursor (ANG-1).
SEMA3B	ENSG00000012171	2.629	Semaphorin-3B precursor (Semaphorin V) (Sema V) (Sema A(V)).
CAPN6	ENSG00000077274	2.611	Calpain-6 (Calpamodulin) (CalpM) (Calpain-like protease X-linked).
ITGBL1	ENSG00000198542	2.592	integrin, beta-like 1 (with EGF-like repeat domains)
TMEM14A	**ENSG00000096092**	**2.578**	**Transmembrane protein 14A.**
DUSP1	ENSG00000120129	2.571	Dual specificity protein phosphatase 1 (EC 3.1.3.48) (EC 3.1.3.16) (MAP kinase phosphatase 1) (MKP-1)
MCTP2	ENSG00000140563	2.566	multiple C2 domains, transmembrane 2
EGR1	**ENSG00000120738**	**2.477**	**Early growth response protein 1 (EGR-1) (Krox-24 protein) (Transcription factor Zif268) (Nerve growth factor-induced protein A) (NGFI-A)**
CHST2	ENSG00000175040	2.477	Carbohydrate sulfotransferase 2 (EC 2.8.2.-) (N-acetylglucosamine 6-O- sulfotransferase 1) (GlcNAc6ST-1) (Gn6ST)
F2RL1	ENSG00000164251	2.473	Proteinase-activated receptor 2 precursor (PAR-2) (Thrombin receptor- like 1) (Coagulation factor II receptor-like 1)
VDR	**ENSG00000111424**	**2.448**	**Vitamin D3 receptor (VDR) (1,25-dihydroxyvitamin D3 receptor).**
PDGFC	ENSG00000145431	2.439	platelet-derived growth factor C precursor
PCDHB18	ENSG00000146001	2.414	protocadherin beta 18 pseudogene (PCDHB18) on chromosome 5
C6orf64	**ENSG00000112167**	**2.394**	**Uncharacterized protein C6orf64.**
ELOVL6	ENSG00000170522	2.391	Elongation of very long chain fatty acids protein 6 (hELO2).
NQO1	**ENSG00000181019**	**2.385**	**NAD(P)H dehydrogenase [quinone] 1 (EC 1.6.5.2) (Quinone reductase 1) (NAD(P)H:quinone oxidoreductase 1)**
SEC16B	ENSG00000120341	2.384	leucine zipper transcription regulator 2
UBXD6	ENSG00000104691	2.351	UBX domain-containing protein 6 (Reproduction 8 protein) (Protein Rep- 8).
PDGFRB	ENSG00000113721	2.335	Beta platelet-derived growth factor receptor precursor (EC 2.7.10.1) (PDGF-R-beta) (CD140b antigen).
TMEPAI	ENSG00000124225	2.327	Transmembrane prostate androgen-induced protein (Solid tumor- associated 1 protein).
TPBG	ENSG00000146242	2.321	Trophoblast glycoprotein precursor (5T4 oncofetal trophoblast glycoprotein) (5T4 oncotrophoblast glycoprotein) (5T4 oncofetal antigen) (M6P1).
C3orf33	ENSG00000174928	2.308	Uncharacterized protein C3orf33.
CD300LB	ENSG00000178789	2.308	CD300 molecule-like family member b
PLP1	ENSG00000123560	2.299	Myelin proteolipid protein (PLP) (Lipophilin).
COL12A1	ENSG00000111799	2.295	Collagen alpha-1(XII) chain precursor.
RDH10	ENSG00000121039	2.293	retinol dehydrogenase 10
PTPN22	ENSG00000134242	2.292	Tyrosine-protein phosphatase non-receptor type 22 (EC 3.1.3.48) (Hematopoietic cell protein-tyrosine phosphatase 70Z-PEP)
SULF2	ENSG00000196562	2.284	Extracellular sulfatase Sulf-2 precursor (EC 3.1.6.-) (HSulf-2).
TLR3	ENSG00000164342	2.277	Toll-like receptor 3 precursor (CD283 antigen).
STOX2	**ENSG00000173320**	**2.275**	**storkhead box 2**
ALOX5AP	ENSG00000132965	2.255	Arachidonate 5-lipoxygenase-activating protein (FLAP) (MK-886-binding protein).
TAPBPL	ENSG00000139192	2.245	Tapasin-related protein precursor (TAPASIN-R) (Tapasin-like) (TAP- binding protein-related protein) (TAPBP-R) (TAP-binding protein-like).

ZMAT3	ENSG00000172667	2.242	p53 target zinc finger protein isoform 1
SLAMF8	ENSG00000158714	2.241	SLAM family member 8 precursor (B-lymphocyte activator macrophage expressed) (BCM-like membrane protein).
PSMB10	ENSG00000205220	2.237	Proteasome subunit beta type 10 precursor (EC 3.4.25.1) (Proteasome MECl-1) (Macropain subunit MECl-1)
PSMB9	ENSG00000204261	2.225	Proteasome subunit beta type 9 precursor (EC 3.4.25.1) (Proteasome chain 7) (Macropain chain 7)
ACYP2	**ENSG00000170634**	**2.195**	**Acylphosphatase-2 (EC 3.6.1.7) (Acylphosphate phosphohydrolase 2) (Acylphosphatase, muscle type isozyme).**
XDH	ENSG00000158125	2.186	Xanthine dehydrogenase/oxidase [Includes: Xanthine dehydrogenase (EC 1.17.1.4) (XD); Xanthine oxidase (EC 1.17.3.2) (XO)
SOD3	ENSG00000109610	2.183	Extracellular superoxide dismutase [Cu-Zn] precursor (EC 1.15.1.1) (EC-SOD).
NRP1	ENSG00000099250	2.174	Neuropilin-1 precursor (Vascular endothelial cell growth factor 165 receptor) (CD304 antigen).
COL6A2	ENSG00000142173	2.171	Collagen alpha-2(VI) chain precursor.
GLT8D4	ENSG00000172986	2.17	Glycosyltransferase 8 domain-containing protein 4 (EC 2.4.1.-).
ADH7	ENSG00000196344	2.163	Alcohol dehydrogenase class 4 mu/sigma chain (EC 1.1.1.1) (Alcohol dehydrogenase class IV mu/sigma chain) (Retinol dehydrogenase) (Gastric alcohol dehydrogenase).
TCN2	ENSG00000185339	2.158	Transcobalamin-2 precursor (Transcobalamin II) (TCII) (TC II).
PARP9	ENSG00000138496	2.154	Poly [ADP-ribose] polymerase 9 (EC 2.4.2.30) (PARP-9) (B aggressive lymphoma protein).
ATF5	ENSG00000169136	2.154	Cyclic AMP-dependent transcription factor ATF-5 (Activating transcription factor 5) (Transcription factor ATFx).
STC2	ENSG00000113739	2.15	Stanniocalcin-2 precursor (STC-2) (Stanniocalcin-related protein) (STCRP) (STC-related protein).
ADCY7	ENSG00000121281	2.133	Adenylate cyclase type 7 (EC 4.6.1.1) (Adenylate cyclase type VII) (ATP pyrophosphate-lyase 7) (Adenylyl cyclase 7).
SERPINB8	ENSG00000166401	2.131	Serpin B8 (Cytoplasmic antiproteinase 2) (CAP-2) (CAP2) (Proteinase inhibitor 8).
NAPRT1	ENSG00000147813	2.129	nicotinate phosphoribosyltransferase domain containing 1
SCRN3	ENSG00000144306	2.115	Secernin-3.
COL27A1	ENSG00000196739	2.109	collagen, type XXVII, alpha 1
EGFR	ENSG00000146648	2.1	Epidermal growth factor receptor precursor (EC 2.7.10.1) (Receptor tyrosine-protein kinase ErbB-1).
SPRY4	**ENSG00000187678**	**2.093**	**Sprouty homolog 4 (Spry-4).**
P2RX7	ENSG00000089041	2.084	P2X purinoceptor 7 (ATP receptor) (P2X7) (Purinergic receptor) (P2Z receptor).
COLEC12	ENSG00000158270	2.084	collectin sub-family member 12
GPC1	ENSG00000063660	2.075	Glypican-1 precursor.
FLCN	ENSG00000154803	2.074	Folliculin (Birt-Hogg-Dube syndrome protein) (BHD skin lesion fibrofolliculoma protein).
ALG8	ENSG00000159063	2.072	Probable dolichyl pyrophosphate Glc1Man9GlcNAc2 alpha-1,3- glucosyltransferase (EC 2.4.1.-)
RSAD2	ENSG00000134321	2.07	radical S-adenosyl methionine domain containing 2
DCTD	**ENSG00000129187**	**2.059**	**Deoxycytidylate deaminase (EC 3.5.4.12) (dCMP deaminase).**
SORBS3	ENSG00000120896	2.054	Vinexin (Sorbin and SH3 domain-containing protein 3) (SH3-containing adapter molecule 1) (SCAM-1).
TACC2	**ENSG00000138162**	**2.052**	**Transforming acidic coiled-coil-containing protein 2 (Anti Zuai-1) (AZU-1).**
EMP3	ENSG00000142227	2.05	Epithelial membrane protein 3 (EMP-3) (YMP protein) (Hematopoietic neural membrane protein) (HNMP-1).
HOXB2	ENSG00000173917	2.049	Homeobox protein Hox-B2 (Hox-2H) (Hox-2.8) (K8).
CYP51A1	ENSG000000	2.042	Cytochrome P450 51A1 (EC 1.14.13.70) (CYPLI) (P450LI) (Sterol 14-alpha demethylase) (Lanosterol 14-

	01630		alpha demethylase)
CRABP2	ENSG00000143320	2.023	Cellular retinoic acid-binding protein 2 (Cellular retinoic acid- binding protein II) (CRABP-II) (Retinoic acid-binding protein II, cellular).
NP_001009555.2	ENSG00000109686	2.021	SH3 domain protein D19
RAB3IL1	ENSG00000167994	2.021	RAB3A interacting protein (rabin3)-like 1
PTPRE	ENSG00000132334	2.016	Receptor-type tyrosine-protein phosphatase epsilon precursor (EC 3.1.3.48) (Protein-tyrosine phosphatase epsilon) (R-PTP-epsilon).]
SLC30A1	ENSG00000170385	2.016	Zinc transporter 1 (ZnT-1) (Solute carrier family 30 member 1).
GHR	ENSG00000112964	2.015	Growth hormone receptor precursor (GH receptor) (Somatotropin receptor) [Contains: Growth hormone-binding protein (GH-binding protein)
PDLIM5	**ENSG00000163110**	**2.011**	**PDZ and LIM domain protein 5 (Enigma homolog) (Enigma-like PDZ and LIM domains protein).**
NP_079178.2	ENSG00000178401	2.006	
BBS4	ENSG00000140463	-2	Bardet-Biedl syndrome 4 protein.
TMTC2	ENSG00000179104	-2.003	Transmembrane and TPR repeat-containing protein 2.
STARD8	ENSG00000130052	-2.004	StAR-related lipid transfer protein 8 (StARD8) (START domain- containing protein 8).
CORO2A	ENSG00000106789	-2.007	Coronin-2A (WD repeat-containing protein 2) (IR10).
SLC25A23	ENSG00000125648	-2.011	solute carrier family 25, member 23
MATN2	ENSG00000132561	-2.011	Matrilin-2 precursor.
HAO1	ENSG00000101323	-2.012	Hydroxyacid oxidase 1 (EC 1.1.3.15) (HAOX1) (Glycolate oxidase) (GOX).
HTATSF1	ENSG00000102241	-2.017	HIV Tat-specific factor 1 (Tat-SF1).
GPR155	ENSG00000163328	-2.018	Integral membrane protein GPR155 (G-protein coupled receptor PGR22).
ALDH18A1	ENSG00000059573	-2.025	Delta 1-pyrroline-5-carboxylate synthetase (P5CS) (Aldehyde dehydrogenase 18 family member A1)
MAOA	ENSG00000189221	-2.03	Amine oxidase [flavin-containing] A (EC 1.4.3.4) (Monoamine oxidase type A) (MAO-A).
NDRG2	ENSG00000165795	-2.035	Protein NDRG2 (Protein Syld709613).
HS3ST5	**ENSG00000175818**	**-2.039**	**Heparan sulfate glucosamine 3-O-sulfotransferase 5 (EC 2.8.2.23) (Heparan sulfate D-glucosaminyl 3-O-sulfotransferase 5)**
ASAHL	ENSG00000138744	-2.047	N-acylethanolamine-hydrolyzing acid amidase precursor (EC 3.5.1.-) (N- acylsphingosine amidohydrolase-like) (ASAH-like protein)
GAS2L3	ENSG00000139354	-2.047	GAS2-like protein 3 (Growth arrest-specific 2-like 3).
LRRC20	ENSG00000172731	-2.048	Leucine-rich repeat-containing protein 20.
PHEX	ENSG00000102174	-2.05	Phosphate-regulating neutral endopeptidase (EC 3.4.24.-) (Metalloendopeptidase homolog PEX) (X-linked hypophosphatemia protein)
FBXO32	ENSG00000156804	-2.05	F-box only protein 32 (Muscle atrophy F-box protein) (MAFbx) (Atrogin- 1).
CXCL12	ENSG00000107562	-2.051	Stromal cell-derived factor 1 precursor (SDF-1) (CXCL12) (Pre-B cell growth-stimulating factor) (PBSF) (hIRH)
MEX3D	ENSG00000181588	-2.052	RNA-binding protein MEX3D (RING finger and KH domain-containing protein 1) (RING finger protein 193). 8]
LDLRAD3	ENSG00000179241	-2.059	low density lipoprotein receptor class A domain containing 3
CA6	ENSG00000131686	-2.061	Carbonic anhydrase 6 precursor (EC 4.2.1.1) (Carbonic anhydrase VI) (Carbonate dehydratase VI) (CA-VI) (Secreted carbonic anhydrase) .
TMEM121	ENSG00000184986	-2.064	Transmembrane protein 121.
NEFL	ENSG00000104725	-2.065	Neurofilament light polypeptide (NF-L) (Neurofilament triplet L protein) (68 kDa neurofilament protein).
C7orf46	ENSG00000188732	-2.067	Uncharacterized protein C7orf46.

PDLIM7	ENSG00000196923	-2.07	PDZ and LIM domain protein 7 (LIM mineralization protein) (LMP) (Protein enigma).
NNMT	ENSG00000166741	-2.078	Nicotinamide N-methyltransferase (EC 2.1.1.1).
MMP12	ENSG00000110347	-2.081	Macrophage metalloelastase precursor (EC 3.4.24.65) (HME) (Matrix metalloproteinase-12) (MMP-12) (Macrophage elastase) (ME).
EXOC6	ENSG00000138190	-2.09	Exocyst complex component 6 (Exocyst complex component Sec15A) (Sec15- like 1).
ITPKB	ENSG00000143772	-2.1	Inositol-trisphosphate 3-kinase B (EC 2.7.1.127) (Inositol 1,4,5- trisphosphate 3-kinase B)
C6orf145	ENSG00000168994	-2.1	Uncharacterized protein C6orf145.
NP_056994.3	ENSG00000143951	-2.104	
PTPRM	ENSG00000173482	-2.104	Receptor-type tyrosine-protein phosphatase mu precursor (EC 3.1.3.48) (Protein-tyrosine phosphatase mu) (R-PTP-mu).
FAM125B	ENSG00000196814	-2.109	Protein FAM125B. [Source:Uniprot/SWISSPROT;Acc:Q9H7P6]
C20orf74	**ENSG00000188559**	**-2.118**	**250 kDa substrate of Akt (AS250). [Source:Uniprot/SWISSPROT;Acc:Q2PPJ7]**
KLHL8	ENSG00000145332	-2.123	Kelch-like protein 8. [Source:Uniprot/SWISSPROT;Acc:Q9P2G9]
C8orf32	ENSG00000156795	-2.123	Uncharacterized protein C8orf32. [Source:Uniprot/SWISSPROT;Acc:Q96HA8]
SNORA32	ENSG00000166012	-2.123	Protein JOSD3. [Source:Uniprot/SWISSPROT;Acc:Q9H5J8]
MPPED2	ENSG00000066382	-2.125	Metallophosphoesterase domain-containing protein 2 (Fetal brain protein 239) (239FB). [Source:Uniprot/SWISSPROT;Acc:Q15777]
LYPLA3	ENSG00000103066	-2.132	1-O-acylceramide synthase precursor (EC 2.3.1.-) (ACS) (Lysosomal phospholipase A2) (Lysophospholipase 3) (LPLA2)
FAM62A	ENSG00000139641	-2.138	Protein FAM62A (Membrane-bound C2 domain-containing protein).
PRTN3	**ENSG00000196415**	**-2.143**	**Myeloblastin precursor (EC 3.4.21.76) (Leukocyte proteinase 3) (PR-3) (PR3) (AGP7) (Wegener autoantigen) (P29) (C-ANCA antigen)**
NP_116012.2	ENSG00000137463	-2.15	ovary-specific acidic protein [Source:RefSeq_peptide;Acc:NP_116012]
PTPRK	ENSG00000152894	-2.152	Receptor-type tyrosine-protein phosphatase kappa precursor (EC 3.1.3.48) (Protein-tyrosine phosphatase kappa) (R-PTP-kappa).
CCNG2	ENSG00000138764	-2.159	Cyclin-G2. [Source:Uniprot/SWISSPROT;Acc:Q16589]
S100A3	**ENSG00000188015**	**-2.163**	**Protein S100-A3 (S100 calcium-binding protein A3) (Protein S-100E). [Source:Uniprot/SWISSPROT;Acc:P33764]**
C3orf28	ENSG00000114023	-2.167	E2-induced gene 5 protein. [Source:Uniprot/SWISSPROT;Acc:Q96A26]
HBEGF	ENSG00000113070	-2.192	Heparin-binding EGF-like growth factor precursor (HB-EGF) (HBEGF) (Diphtheria toxin receptor)
SLC44A2	ENSG00000129353	-2.206	Choline transporter-like protein 2 (Solute carrier family 44 member 2). [Source:Uniprot/SWISSPROT;Acc:Q8IWA5]
GFRA2	ENSG00000168546	-2.215	GDNF family receptor alpha-2 precursor (GFR-alpha-2) (Neurturin receptor alpha) (NTNR-alpha) (NRTNR-alpha)
HK1	ENSG00000156515	-2.216	Hexokinase-1 (EC 2.7.1.1) (Hexokinase type I) (HK I) (Brain form hexokinase). [Source:Uniprot/SWISSPROT;Acc:P19367]
CXorf15	ENSG00000086712	-2.219	Gamma-taxilin (Lipopolysaccharide-specific response protein 5). [Source:Uniprot/SWISSPROT;Acc:Q9NUQ3]
MARK1	ENSG00000116141	-2.221	Serine/threonine-protein kinase MARK1 (EC 2.7.11.1) (MAP/microtubule affinity-regulating kinase 1).
ID1	ENSG00000125968	-2.228	DNA-binding protein inhibitor ID-1 (Inhibitor of DNA binding 1).
B3GALNT1	ENSG00000169255	-2.231	UDP-GalNAc:beta-1,3-N-acetylgalactosaminyltransferase 1 (EC 2.4.1.79) (Beta-3-GalNAc-T1) (Beta-1,3-galactosyltransferase 3)
CLYBL	ENSG00000125246	-2.235	Citrate lyase beta subunit-like protein, mitochondrial precursor (EC 4.1.-.-) (Citrate lyase beta-like).
PRR6	ENSG00000166582	-2.249	Proline-rich protein 6 (Nuclear protein p30). [Source:Uniprot/SWISSPROT;Acc:Q7Z7K6]
PVRL2	ENSG00000130202	-2.261	Poliovirus receptor-related protein 2 precursor (Herpes virus entry mediator B) (HveB) (Nectin-2) (CD112 antigen).
STAMBPL1	ENSG000001	-2.262	AMSH-like protease (EC 3.1.2.15) (AMSH-LP) (STAM-binding protein-like 1).

	38134		[Source:Uniprot/SWISSPROT;Acc:Q96FJ0]
OSTF1	ENSG00000134996	-2.267	Osteoclast-stimulating factor 1. [Source:Uniprot/SWISSPROT;Acc:Q92882]
CLCN3	ENSG00000109572	-2.288	Chloride channel protein 3 (ClC-3). [Source:Uniprot/SWISSPROT;Acc:P51790]
C6orf141	ENSG00000197261	-2.303	Uncharacterized protein C6orf141. [Source:Uniprot/SWISSPROT;Acc:Q5SZD1]
MXI1	ENSG00000119950	-2.308	MAX-interacting protein 1 (Protein MXI1). [Source:Uniprot/SWISSPROT;Acc:P50539]
GADD45G	ENSG00000130222	-2.308	Growth arrest and DNA-damage-inducible protein GADD45 gamma (Cytokine- responsive protein CR6).
ARHGAP29	ENSG00000137962	-2.323	PTPL1-associated RhoGAP 1 [Source:RefSeq_peptide;Acc:NP_004806]
VLDLR	ENSG00000147852	-2.327	Very low-density lipoprotein receptor precursor (VLDL receptor) (VLDL- R). [Source:Uniprot/SWISSPROT;Acc:P98155]
SPDYA	**ENSG00000163806**	**-2.327**	**Speedy protein A (Speedy-1) (Spy1) (Rapid inducer of G2/M progression in oocytes A) (RINGO A) (hSpy/Ringo A).**
CIRBP	**ENSG00000099622**	**-2.336**	**Cold-inducible RNA-binding protein (Glycine-rich RNA-binding protein CIRP) (A18 hnRNP). [Source:Uniprot/SWISSPROT;Acc:Q14011]**
BCKDHB	ENSG00000083123	-2.339	2-oxoisovalerate dehydrogenase subunit beta, mitochondrial precursor (EC 1.2.4.4)
ENAH	ENSG00000154380	-2.339	Protein enabled homolog. [Source:Uniprot/SWISSPROT;Acc:Q8N8S7]
NDRG1	ENSG00000104419	-2.349	Protein NDRG1 (N-myc downstream-regulated gene 1 protein) (Differentiation-related gene 1 protein) (DRG-1)
DENND2A	**ENSG00000146966**	**-2.356**	**DENN domain-containing protein 2A. [Source:Uniprot/SWISSPROT;Acc:Q9ULE3]**
SLC2A1	ENSG00000117394	-2.36	Solute carrier family 2, facilitated glucose transporter member 1 (Glucose transporter type 1, erythrocyte/brain) (GLUT-1)
NRN1	ENSG00000124785	-2.368	Neuritin precursor. [Source:Uniprot/SWISSPROT;Acc:Q9NPD7]
EPHB2	ENSG00000133216	-2.369	Ephrin type-B receptor 2 precursor (EC 2.7.10.1) (Tyrosine-protein kinase receptor EPH-3) (DRT) (Receptor protein-tyrosine kinase HEK5) (ERK) (Tyrosine-protein kinase TYRO5)
NMNAT3	**ENSG00000163864**	**-2.369**	**Nicotinamide mononucleotide adenylyltransferase 3 (EC 2.7.7.1) (NMN adenylyltransferase 3)**
SH3BP4	ENSG00000130147	-2.396	SH3 domain-binding protein 4 (Transferrin receptor-trafficking protein) (EH-binding protein 10). [Source:Uniprot/SWISSPROT;Acc:Q9P0V3]
INSIG2	ENSG00000125629	-2.4	Insulin-induced gene 2 protein (INSIG-2). [Source:Uniprot/SWISSPROT;Acc:Q9Y5U4]
GLUL	ENSG00000135821	-2.406	Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase) (GS). [Source:Uniprot/SWISSPROT;Acc:P15104]
ETS2	ENSG00000157557	-2.407	Protein C-ets-2. [Source:Uniprot/SWISSPROT;Acc:P15036]
ENO2	ENSG00000111674	-2.424	Gamma-enolase (EC 4.2.1.11) (2-phospho-D-glycerate hydro-lyase) (Neural enolase) (Neuron-specific enolase) (NSE)
NAALAD2	ENSG00000077616	-2.427	N-acetylated-alpha-linked acidic dipeptidase 2 (EC 3.4.17.21) (N- acetylated-alpha-linked acidic dipeptidase II) (NAALADase II).
TFRC	ENSG00000072274	-2.437	Transferrin receptor protein 1 (TfR1) (TR) (TfR) (Trfr) (CD71 antigen) (T9) (p90) [Contains: Transferrin receptor protein 1, serum form (sTfR)].
CTH	ENSG00000116761	-2.458	Cystathionine gamma-lyase (EC 4.4.1.1) (Gamma-cystathionase). [Source:Uniprot/SWISSPROT;Acc:P32929]
REEP1	ENSG00000068615	-2.462	Receptor expression-enhancing protein 1. [Source:Uniprot/SWISSPROT;Acc:Q9H902]
TRIB3	ENSG00000101255	-2.474	Tribbles homolog 3 (TRB-3) (Neuronal cell death-inducible putative kinase) (p65-interacting inhibitor of NF-kappaB) (SINK).
CSN3	**ENSG00000171209**	**-2.475**	**Kappa-casein precursor. [Source:Uniprot/SWISSPROT;Acc:P07498]**
C9orf72	ENSG00000147894	-2.479	Uncharacterized protein C9orf72. [Source:Uniprot/SWISSPROT;Acc:Q96LT7]
NP_001070884.1	ENSG00000205084	-2.501	CDNA: FLJ22167 fis, clone HRC00584 (Hypothetical protein FLJ22167). [Source:Uniprot/SPTREMBL;Acc:Q9H6L2]
ZFPM2	ENSG00000169946	-2.521	Zinc finger protein ZFPM2 (Zinc finger protein multitype 2) (Friend of GATA protein 2) (FOG-2) (hFOG-2).
ARRDC4	ENSG00000140450	-2.523	Arrestin domain-containing protein 4. [Source:Uniprot/SWISSPROT;Acc:Q8NCT1]
C20orf112	ENSG00000197183	-2.53	Uncharacterized protein C20orf112. [Source:Uniprot/SWISSPROT;Acc:Q96MY1]

C9orf30	ENSG00000066697	-2.54	UPF0439 protein C9orf30 (Protein L8). [Source:Uniprot/SWISSPROT;Acc:Q96H12]
GALNT12	ENSG00000119514	-2.541	Polypeptide N-acetylgalactosaminyltransferase 12 (EC 2.4.1.41) (Protein-UDP acetylgalactosaminyltransferase 12)
GADD45A	ENSG00000116717	-2.549	Growth arrest and DNA-damage-inducible protein GADD45 alpha (DNA- damage-inducible transcript 1) (DDIT1).
RAB7L1	ENSG00000117280	-2.566	Ras-related protein Rab-7L1 (Rab-7-like protein 1). [Source:Uniprot/SWISSPROT;Acc:O14966]
NP_001071087.1	ENSG00000215114	-2.574	
LAMA4	ENSG00000112769	-2.576	Laminin subunit alpha-4 precursor. [Source:Uniprot/SWISSPROT;Acc:Q16363]
TCEAL1	ENSG00000172465	-2.576	Transcription elongation factor A protein-like 1 (TCEA-like protein 1) (Transcription elongation factor S-II protein-like 1)
SLC25A30	**ENSG00000174032**	**-2.586**	**Kidney mitochondrial carrier protein 1 (Solute carrier family 25 member 30). [Source:Uniprot/SWISSPROT;Acc:Q5SVS4]**
CTGF	ENSG00000118523	-2.599	Connective tissue growth factor precursor (Hypertrophic chondrocyte- specific protein 24). [Source:Uniprot/SWISSPROT;Acc:P29279]
ERO1L	ENSG00000197930	-2.622	ERO1-like protein alpha precursor (EC 1.8.4.-) (ERO1-Lalpha) (Oxidoreductin-1-Lalpha) (Endoplasmic oxidoreductin-1-like protein) (ERO1-L).
ABTB2	ENSG00000166016	-2.624	Ankyrin repeat and BTB/POZ domain-containing protein 2. [Source:Uniprot/SWISSPROT;Acc:Q8N961]
C9orf89	ENSG00000165233	-2.64	BINCA_HUMAN Isoform 2 of Q96LW7 - Homo sapiens (Human) [Source:Uniprot/Varsplic;Acc:Q96LW7-2]
DNAJC12	ENSG00000108176	-2.671	DnaJ homolog subfamily C member 12 (J domain-containing protein 1). [Source:Uniprot/SWISSPROT;Acc:Q9UKB3]
GNAZ	ENSG00000128266	-2.687	Guanine nucleotide-binding protein G(z) subunit alpha (G(x) alpha chain) (Gz-alpha). [Source:Uniprot/SWISSPROT;Acc:P19086]
THBS2	ENSG00000186340	-2.689	Thrombospondin-2 precursor. [Source:Uniprot/SWISSPROT;Acc:P35442]
CITED1	ENSG00000125931	-2.726	Cbp/p300-interacting transactivator 1 (Melanocyte-specific protein 1). [Source:Uniprot/SWISSPROT;Acc:Q99966]
PKIA	ENSG00000171033	-2.729	cAMP-dependent protein kinase inhibitor alpha (PKI-alpha) (cAMP- dependent protein kinase inhibitor, muscle/brain isoform).
TSC22D3	ENSG00000157514	-2.744	TSC22 domain family protein 3 (Glucocorticoid-induced leucine zipper protein) (Delta sleep-inducing peptide immunoreactor)
CLU	ENSG00000120885	-2.769	Clusterin precursor (Complement-associated protein SP-40,40) (Complement cytolysis inhibitor) (CLI) (NA1/NA2) (Apolipoprotein J) (Apo-J)
SNCA	ENSG00000145335	-2.792	Alpha-synuclein (Non-A beta component of AD amyloid) (Non-A4 component of amyloid precursor) (NACP).
ADD3	ENSG00000148700	-2.8	Gamma-adducin (Adducin-like protein 70). [Source:Uniprot/SWISSPROT;Acc:Q9UEY8]
PRELID2	ENSG00000186314	-2.802	PRELI domain containing 2 isoform c [Source:RefSeq_peptide;Acc:NP_612501]
MBOAT2	ENSG00000143797	-2.825	Membrane-bound O-acyltransferase domain-containing protein 2 (EC 2.3.-.-) (O-acyltransferase domain-containing protein 2).
PDK1	ENSG00000152256	-2.825	[Pyruvate dehydrogenase [lipoamide]] kinase isozyme 1, mitochondrial precursor (EC 2.7.11.2) (Pyruvate dehydrogenase kinase isoform 1).
MUM1L1	**ENSG00000157502**	**-2.828**	**MUM1-like protein 1 (Mutated melanoma-associated antigen 1-like protein 1). [Source:Uniprot/SWISSPROT;Acc:Q5H9M0]**
HIST1H1C	ENSG00000187837	-2.843	Histone H1.2 (Histone H1d). [Source:Uniprot/SWISSPROT;Acc:P16403]
SASH1	ENSG00000111961	-2.911	SAM and SH3 domain-containing protein 1 (Proline-glutamate repeat- containing protein). [Source:Uniprot/SWISSPROT;Acc:O94885]
IL33	ENSG00000137033	-2.916	Interleukin-33 precursor (IL-33) (Interleukin-1 family member 11) (IL- 1F11) (Nuclear factor from high endothelial venules) (NF-HEV).
MDM1	ENSG00000111554	-2.963	Mdm4, transformed 3T3 cell double minute 1, p53 binding protein isoform 1
C4orf31	ENSG00000173376	-2.968	Uncharacterized protein C4orf31. [Source:Uniprot/SPTREMBL;Acc:Q8TB73]
CDKN1C	ENSG00000129757	-2.977	Cyclin-dependent kinase inhibitor 1C (Cyclin-dependent kinase inhibitor p57) (p57KIP2). [Source:Uniprot/SWISSPROT;Acc:P49918]
P4HA2	ENSG00000072682	-2.984	Prolyl 4-hydroxylase subunit alpha-2 precursor (EC 1.14.11.2) (4-PH alpha-2) (Procollagen-proline,2-oxoglutarate-4-dioxygenase alpha-2 subunit).
ITGA6	ENSG00000091409	-3.008	Integrin alpha-6 precursor (VLA-6) (CD49f antigen) [Contains: Integrin alpha-6 heavy chain; Integrin alpha-6 light chain].
SH3YL1	ENSG000000	-3.01	SH3 domain containing, Ysc84-like 1 [Source:RefSeq_peptide;Acc:NP_056492]

	35115		
SYT17	**ENSG00000103528**	**-3.095**	**B/K protein [Source:RefSeq_peptide;Acc:NP_057608]**
KLF6	ENSG00000067082	-3.099	Krueppel-like factor 6 (Core promoter element-binding protein) (B- cell-derived protein 1) (Proto-oncogene BCD1) (Transcription factor Zf9)
EVI1	ENSG00000085276	-3.265	Ecotropic virus integration site 1 protein homolog (EVI-1). [Source:Uniprot/SWISSPROT;Acc:Q03112]
EDG3	ENSG00000213694	-3.28	Sphingosine 1-phosphate receptor Edg-3 (S1P receptor Edg-3) (Endothelial differentiation G-protein coupled receptor 3)
TMCC3	ENSG00000057704	-3.305	Transmembrane and coiled-coil domains protein 3. [Source:Uniprot/SWISSPROT;Acc:Q9ULS5]
DOCK10	ENSG00000135905	-3.311	Dedicator of cytokinesis protein 10 (Zizimin-3). [Source:Uniprot/SWISSPROT;Acc:Q96BY6]
BBS2	ENSG00000125124	-3.339	Bardet-Biedl syndrome 2 protein. [Source:Uniprot/SWISSPROT;Acc:Q9BXC9]
ENPP2	ENSG00000136960	-3.43	Ectonucleotide pyrophosphatase/phosphodiesterase family member 2 precursor (EC 3.1.4.39) (E-NPP 2)
ATF3	ENSG00000162772	-3.488	Cyclic AMP-dependent transcription factor ATF-3 (Activating transcription factor 3). [Source:Uniprot/SWISSPROT;Acc:P18847]
RUNDC3B	ENSG00000105784	-3.539	RUN domain containing 3B [Source:RefSeq_peptide;Acc:NP_612147]
RNF128	ENSG00000133135	-3.588	E3 ubiquitin-protein ligase RNF128 precursor (EC 6.3.2.-) (RING finger protein 128) (Gene related to anergy in lymphocytes protein).
TEX11	**ENSG00000120498**	**-3.655**	**testis expressed sequence 11 isoform 2 [Source:RefSeq_peptide;Acc:NP_112566]**
DEPDC7	ENSG00000121690	-3.777	novel 58.3 KDA protein isoform 1 [Source:RefSeq_peptide;Acc:NP_001070710]
DFFB	ENSG00000169598	-3.889	DNA fragmentation factor subunit beta (EC 3.-.-.-) (DNA fragmentation factor 40 kDa subunit) (DFF-40) (Caspase-activated deoxyribonuclease)
ADAM8	ENSG00000151651	-4.022	ADAM 8 precursor (EC 3.4.24.-) (A disintegrin and metalloproteinase domain 8) (Cell surface antigen MS2) (CD156a antigen) (CD156).
TMEM176B	ENSG00000106565	-4.161	Transmembrane protein 176B (Protein LR8). [Source:Uniprot/SWISSPROT;Acc:Q3YBM2]
MLF1	ENSG00000178053	-4.27	Myeloid leukemia factor 1 (Myelodysplasia-myeloid leukemia factor 1). [Source:Uniprot/SWISSPROT;Acc:P58340]
TMEM176A	ENSG00000002933	-4.5	Transmembrane protein 176A (Hepatocellular carcinoma-associated antigen 112). [Source:Uniprot/SWISSPROT;Acc:Q96HP8]
HPSE	**ENSG00000173083**	**-4.673**	**Heparanase precursor (EC 3.2.-.-) (Heparanase-1) (Hpa1) (Endo- glucoronidase) [Contains: Heparanase 8 kDa subunit; Heparanase 50 kDa**
DFNB31	**ENSG00000095397**	**-5.112**	**Whirlin (Autosomal recessive deafness type 31 protein). [Source:Uniprot/SWISSPROT;Acc:Q9P202]**
ALCAM	ENSG00000170017	-5.185	CD166 antigen precursor (Activated leukocyte-cell adhesion molecule) (ALCAM). [Source:Uniprot/SWISSPROT;Acc:Q13740]
ADM	ENSG00000148926	-5.301	ADM precursor [Contains: Adrenomedullin (AM); Proadrenomedullin N-20 terminal peptide (ProAM-N20) (ProAM N-terminal 20 peptide)
STBD1	ENSG00000118804	-5.85	Genethonin-1. [Source:Uniprot/SWISSPROT;Acc:O95210]
C9orf19	ENSG00000122694	-6.111	Golgi-associated plant pathogenesis-related protein 1 (Golgi- associated PR-1 protein) (GAPR-1) (Glioma pathogenesis-related protein 2)
POSTN	ENSG00000133110	-6.552	Periostin precursor (PN) (Osteoblast-specific factor 2) (OSF-2). [Source:Uniprot/SWISSPROT;Acc:Q15063]
TNC	ENSG00000041982	-6.868	Tenascin precursor (TN) (Tenascin-C) (TN-C) (Hexabrachion) (Cytotactin) (Neuronectin) (GMEM) (JI) (Myotendinous antigen)
MYOT	**ENSG00000120729**	**-7.135**	**Myotilin (Titin immunoglobulin domain protein) (Myofibrillar titin- like Ig domains protein) (57 kDa cytoskeletal protein).**
MMP10	ENSG00000166670	-7.886	Stromelysin-2 precursor (EC 3.4.24.22) (Matrix metalloproteinase-10) (MMP-10) (Transin-2) (SL-2). [Source:Uniprot/SWISSPROT;Acc:P09238]
SYCE2	ENSG00000161860	-8.051	Synaptonemal complex central element protein 2 (Central element synaptonemal complex protein 1). [Source:Uniprot/SWISSPROT;Acc:Q6PIF2]
ITM2A	ENSG00000078596	-9.112	Integral membrane protein 2A (E25 protein). [Source:Uniprot/SWISSPROT;Acc:O43736]
C10orf58	ENSG00000122378	-9.877	Uncharacterized protein C10orf58 precursor. [Source:Uniprot/SWISSPROT;Acc:Q9BRX8]
BLNK	ENSG00000095585	-10.908	B-cell linker protein (Cytoplasmic adapter protein) (B-cell adapter containing SH2 domain protein) (B-cell adapter containing Src homology 2 domain protein)
GAP43	ENSG00000172020	-13.043	Neuromodulin (Axonal membrane protein GAP-43) (Growth-associated protein 43) (PP46) (Neural phosphoprotein B-50).

LCN2	ENSG00000148346	-16.838	Neutrophil gelatinase-associated lipocalin precursor (NGAL) (p25) (25 kDa alpha-2-microglobulin-related subunit of MMP-9) (Lipocalin-2)
MMP13	**ENSG00000137745**	**-17.341**	**Collagenase 3 precursor (EC 3.4.24.-) (Matrix metalloproteinase-13) (MMP-13).**
CRABP1	**ENSG00000166426**	**-21.012**	**Cellular retinoic acid-binding protein 1 (Cellular retinoic acid- binding protein I) (CRABP-I) (Retinoic acid-binding protein I, cellular).**
FABP4	ENSG00000170323	-38.52	Fatty acid-binding protein, adipocyte (AFABP) (Adipocyte lipid-binding protein) (ALBP) (A-FABP). [Source:Uniprot/SWISSPROT;Acc:P15090]

MC-38 parental versus MC-38met expression changes as determined in Materials and Methods are displayed for the 300 gene metastasis-associated signature. Mouse probe set identifiers (IDs) were mapped to Ensembl Gene IDs based on the mapping provided by Ensembl V49 (http://www.ensembl.org) and are displayed in addition to fold-change and the gene name as noted in WebGestalt. The genes in the 34-gene classifier are bolded.

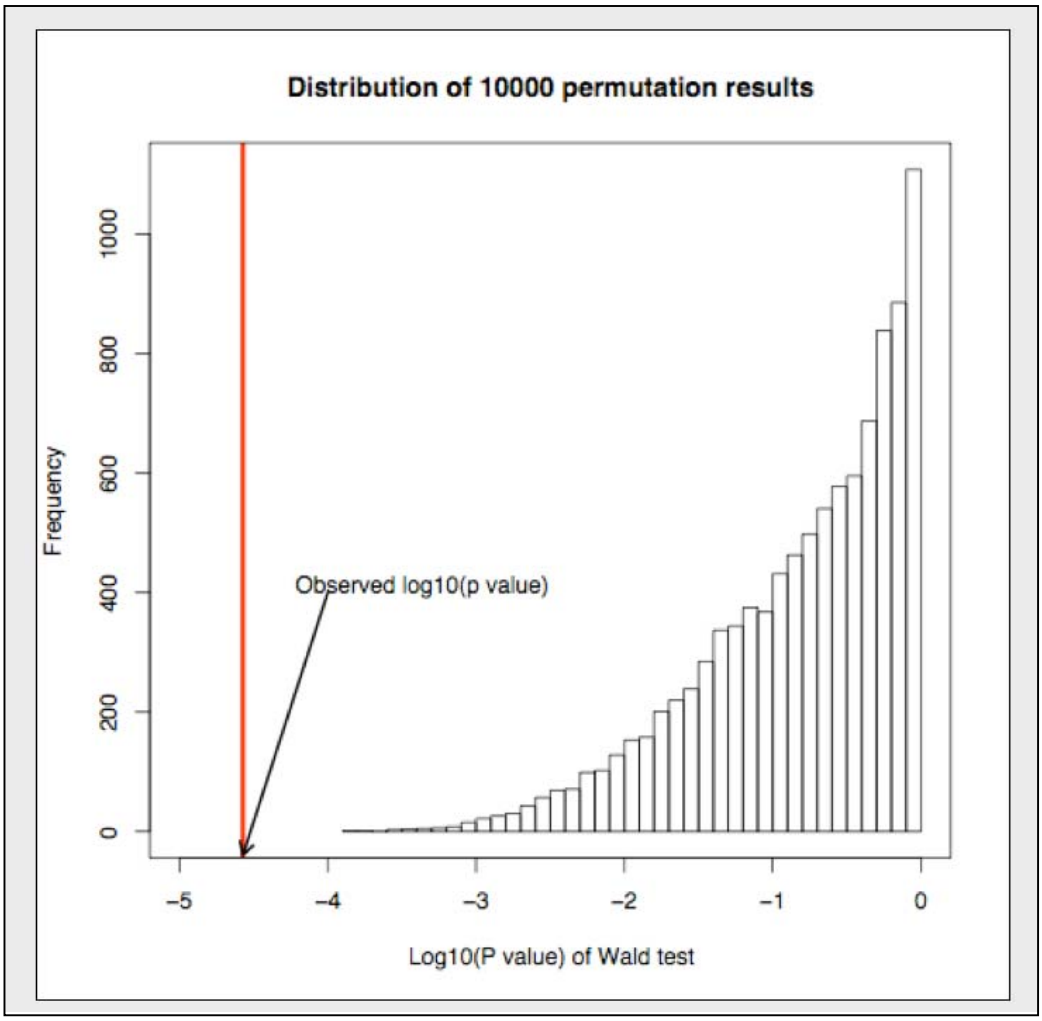


Figure 10. *Distribution of 10,000-permutation Wald tests for the 177 MCC patients with the 34-gene metastasis score*. Beta and Wald statistics for each Affymetrix probe set were used along with expression data to build up a metastasis score for each patient. The score was used as the independent variable to perform overall survival analysis based on the Cox model. The Wald test P-value was saved as the observed P-value. For the re-sampling test, we randomly chose 60 Affymetrix probe sets from the 54,675 sets on the whole array. We repeated the above procedure and generated 1 re-sampling Wald test P-value from the overall Cox model survival analysis. We repeated the re-sampling and survival analysis procedure 10,000 times, generating 10,000 re-sampling Wald test P-values. We transformed both the observed and re-sampling P-values into $\log_{10}$ format, plotted a histogram of the 10,000 re-sampling $\log_{10}$ (P values), and added the observed $\log_{10}$ (P-value).

TABLE 7

THE METASTASIS SCORE IS AN INDEPENDENT PREDICTOR OF RECURRENCE RISK

	Adjusted HR (95% CI)	P-value
Metastasis score	1.016 (1.008-1.025)	<.001
Sex	1.011 (0.481-2.124)	.98
Stage	2.119 (1.312-3.424)	.002
Age	1.001 (0.974-1.029)	.93
Grade	1.446 (0.701-2.985)	.32

A summary of a multivariate analysis, using the MCC patient variables to calculate independent risk factors, is shown. Metastasis score (P<.001) and stage (P= .002) were each found to be significant predictors of DFS in the multivariate model. Metastasis score, age, sex, and grade were adjusted in the multivariate Cox model for DFS, and the results for each risk factor and DFS are shown. Adjusted P-values as well as the adjusted 95% confidence intervals (CI) of the β-coefficient from the Cox model were reported.

TABLE 8

THE METASTASIS SCORE ASSOCIATES WITH INCREASED RISK OF CANCER-RELATED DEATH

	HR (95% CI)	P-value
Disease-specific survival (All stages)	4.4 (2.25-8.45)	<.001
Disease-specific survival (stage II)	NA (NA)	.003
Disease-specific survival (stage III)	4.1 (1.17-14.2)	.03

Univariate analysis with the use of the metastasis score to segregate patients from the MCC data set into higher-than-median and lower-than-median score groups. Hazard ratios were calculated for each patient group related to DSS. The 95% confidence intervals (CIs) and P-values for the test are given in the table. NA implies that no hazard ratio from the Cox model was calculated because no cancer-related deaths occurred in the patients with stage II colon cancer with a low score. The P-value for stage II DSS in this model was calculated according to exact log-rank test for unequal follow-up (Heinze, Gnant, and Schemper 2003).

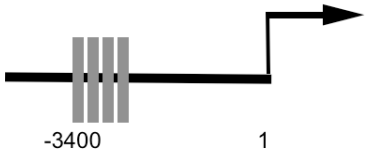
TABLE 9

THE METASTASIS SCORE IS AN INDEPENDENT PREDICTOR OF CANCER-RELATED DEATH

	Adjusted HR (95% CI)	P-value
Metastasis score	1.021 (1.014-1.028)	<.001
Sex	0.884 (0.493-1.588)	.68
Stage	4.98 (3.113-7.967)	<.001
Age	1.026 (1.002-1.051)	.04
Grade	1.61 (0.887-2.922)	.12

A summary of a multivariate analysis, using the MCC patient variables to calculate independent risk factors, is shown in the table. Metastasis score (P<.001), stage (P< .001) and age (P=.04) were each found to be significant predictors of DSS in the multivariate model. Metastasis score, age, sex, and grade were adjusted in the multivariate Cox model for both OS and DSS, and the results for each risk factor and DSS are shown. The adjusted P-values as well as the adjusted 95% confidence intervals (CIs) of the β-coefficient from the Cox model were reported.

A.




```
tgtgtgcacgtgcgtgtgttgcgggtgtgtgtctccgaaacatgatttct (-3400)
tagaaatgcagcggtacttgtctgagttcttagaaacattgagaaacata (-3350)
tatacctttatgttctttgaatatcatggacgatgactccattgcttctc (-3300)
cttcatactttgtatggtagccattaaaatgtacttccttagtctggcaa (-3250)
accaatgaaggcctcgagtacacataactccaagcatgtagggtgttcct (-3200)
gacataactgaagtgaccaggccatagcaaccaccatggccattgcctcc (-3150)
cgccaccttcaccgcccctccctcgctctgcagtatacatcagctgcaag (-3100)
ctttccacagtggctttccgccgaacttgcccagtgggttcctagattac (-3050)
ttctgtgtgcgagaatagccaaccatgtagtctgacaagaggtatttaca (-3000)
```

B.

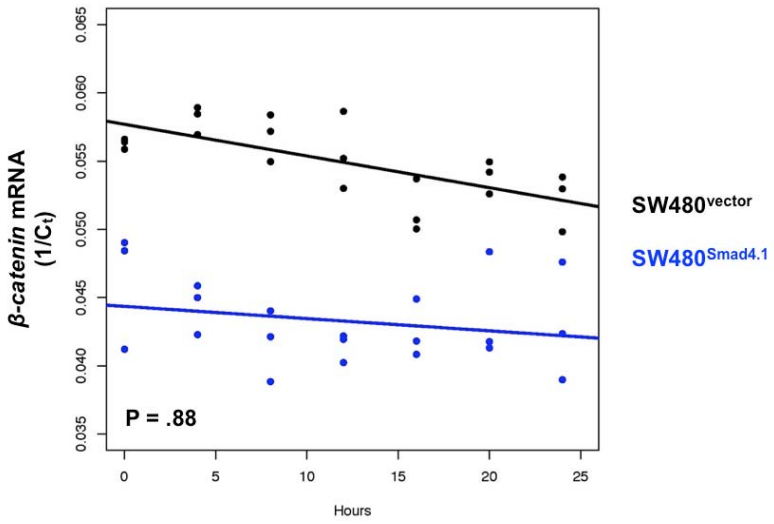


Figure 18. *(A) Ctnnb1 promoter/enhancer minimal Smad binding elements and region of amplification* schematic of *an area of the ctnnb1 promoter* (-3400 to -3000) is displayed. Smad binding elements are in gray and the 5' and 3' primer regions are underlined and bold black. *(B) Smad4 restoration in SW480 cells does not affect β-catenin mRNA stability.* SW480^{vector} and SW480^{Smad4} were treated with 5,6-Dichlorobenzimidazole 1--D-ribofuranoside (DRB, specific RNA polymerase II inhibitor) as indicated for 0, 4, 8, 12, 16, 20 and 24 hours. Relative expression for steady-state β-catenin mRNA was determined after DRB treatment and is displayed. Analysis of variance was used to determine if any significant interactions occurred between groups (P = .88). SW480^{vector} is denoted in black and SW480$^{Smad4.1}$ is denoted in blue. Two doses of DRB were used and data from the 100μM dose is shown.

TABLE 11

GLOBAL GENE EXPRESSION ANALYSIS OF SW480^{VECTOR} VERSUS SW480^{SMAD4} COLON CANCER CELLS

UP-REGULATED BY SMAD4			
Affymetrix probe	**Coef**	**FDR**	**Gene Symbol**
228640_at	-8.714577898	4.44E-25	PCDH7
227711_at	-7.958764655	1.21E-37	GTSF1
206218_at	-7.756652036	5.43E-36	MAGEB2
209942_x_at	-7.717070359	5.91E-36	MAGEA3
214612_x_at	-7.696179633	1.12E-37	MAGEA6
211124_s_at	-7.657882354	1.15E-32	KITLG
201849_at	-7.383997352	8.17E-32	BNIP3
226534_at	-7.173853692	1.02E-29	KITLG
207029_at	-7.000946956	2.42E-28	KITLG
214603_at	-6.883081796	5.43E-36	MAGEA2B
1552767_a_at	-6.784122115	2.48E-33	HS6ST2
201445_at	-6.727393674	4.08E-33	CNN3
230030_at	-6.615547924	9.59E-35	HS6ST2
205919_at	-6.535255899	4.65E-33	HBE1
202983_at	-6.531039709	2.61E-35	HLTF
218186_at	-6.442496555	2.94E-35	RAB25
1553830_s_at	-6.39486666	7.13E-35	MAGEA2
243110_x_at	-6.360154734	9.45E-34	NPW
201117_s_at	-6.217523105	3.19E-30	CPE
209301_at	-6.214799538	7.82E-29	CA2
205767_at	-6.06146787	1.31E-25	EREG
218677_at	-5.966475635	1.02E-22	S100A14
227919_at	-5.907829919	4.55E-24	UCA1
201848_s_at	-5.89585183	2.42E-28	BNIP3
206295_at	-5.885920052	5.68E-33	IL18
204971_at	-5.816499121	1.11E-28	CSTA
205466_s_at	-5.794299158	4.45E-22	HS3ST1
223551_at	-5.71991099	8.13E-25	PKIB
225381_at	-5.644466412	3.31E-34	LOC399959
231192_at	-5.553414689	3.85E-28	NA
201116_s_at	-5.534183303	6.95E-33	CPE
208837_at	-5.517704249	2.92E-29	TMED3
223592_s_at	-5.440415817	1.9E-32	RNF135

204688_at	-5.439822765	8.91E-33	SGCE
204035_at	-5.396409789	6.4E-15	SCG2
202388_at	-5.386649931	3.1E-09	RGS2
205347_s_at	-5.384823628	4.25E-19	TMSB15A
208937_s_at	-5.356431362	4.42E-29	ID1
234469_at	-5.355394233	5.42E-34	OR51B4
225945_at	-5.344472359	2.98E-30	ZNF655
226621_at	-5.343865646	9.1E-26	FGG
202527_s_at	-5.294875093	1.38E-30	SMAD4
235557_at	-5.268789269	3.68E-28	LOC150763
213816_s_at	-5.208338904	5.32E-31	MET
213258_at	-5.203966325	2.58E-34	TFPI
209199_s_at	-5.181597286	7.11E-21	MEF2C
202973_x_at	-5.152029952	4.02E-32	FAM13A
202620_s_at	-5.104496061	2.41E-25	PLOD2
229402_at	-5.094923313	9.11E-29	SAMD13
226302_at	-5.059370209	6.24E-25	ATP8B1
204083_s_at	-5.023699267	2.73E-31	TPM2
1554803_s_at	-5.017672119	5.8E-33	TRIM72
204304_s_at	-5.016762846	3.94E-21	PROM1
212094_at	-4.965591224	9.12E-17	PEG10
226909_at	-4.961649578	9.54E-32	ZNF518B
201272_at	-4.939658251	3.03E-30	AKR1B1
239913_at	-4.907918746	5.75E-31	SLC10A4
224022_x_at	-4.794553413	1.67E-24	WNT16
219412_at	-4.748236738	6.75E-31	RAB38
203708_at	-4.734864329	5.71E-24	PDE4B
201828_x_at	-4.713025367	7.59E-32	FAM127A
221690_s_at	-4.711097692	9.47E-29	NLRP2
203789_s_at	-4.610964515	3.03E-30	SEMA3C
1552766_at	-4.608769757	4.27E-25	HS6ST2
1555801_s_at	-4.588397302	4.73E-33	ZNF385B
211675_s_at	-4.519218284	3.19E-31	MDFIC
202072_at	-4.505536699	2.98E-30	HNRNPL
207826_s_at	-4.489496389	3.88E-24	ID3
201427_s_at	-4.475176247	1.05E-23	SEPP1
221731_x_at	-4.455090423	5.36E-15	VCAN
228297_at	-4.440584307	3.31E-26	CNN3
203510_at	-4.437460842	1.19E-27	MET
210664_s_at	-4.425081818	4.4E-28	TFPI
234994_at	-4.406643543	4.9E-22	TMEM200A

213664_at	-4.378061462	4.73E-26	SLC1A1
212148_at	-4.369511371	5.23E-15	PBX1
213222_at	-4.353416112	2.86E-24	PLCB1
204620_s_at	-4.349783568	2.54E-15	VCAN
213156_at	-4.3412992	3.65E-21	NA
219355_at	-4.328726428	2.12E-29	CXorf57
205729_at	-4.320124371	6.93E-28	OSMR
1555800_at	-4.317205827	4.76E-28	ZNF385B
206155_at	-4.297905664	1.8E-15	ABCC2
219932_at	-4.277046336	1.38E-27	SLC27A6
208782_at	-4.262519874	3.63E-32	FSTL1
237159_x_at	-4.255301528	1.43E-31	AP1S3
203680_at	-4.254882045	2.42E-28	PRKAR2B
225540_at	-4.209549986	2.8E-24	MAP2
213158_at	-4.203307735	1.73E-20	NA
202619_s_at	-4.190018133	3.67E-26	PLOD2
1555733_s_at	-4.188721418	6.51E-30	AP1S3
228624_at	-4.161792378	1.14E-29	TMEM144
211302_s_at	-4.146115095	4.78E-25	PDE4B
217948_at	-4.140555857	3.19E-30	FAM127B
1555731_a_at	-4.131869279	7.16E-30	AP1S3
217047_s_at	-4.1196971	8.47E-24	FAM13A
1555148_a_at	-4.111957437	1.36E-28	C2orf65
232113_at	-4.092979863	9.25E-24	NA
209676_at	-4.08367698	9.1E-26	TFPI
228063_s_at	-4.067851885	2.82E-24	NAP1L5
218546_at	-4.058686835	6.93E-28	C1orf115
220615_s_at	-4.056984743	5.66E-29	FAR2
210273_at	-4.054999962	3.74E-17	PCDH7
226789_at	-4.032101474	1.09E-27	LOC647121
209200_at	-4.023766598	7.29E-19	MEF2C
205239_at	-4.017174485	1.04E-12	AREG
232202_at	-3.994040074	3.94E-23	NA
210868_s_at	-3.993594161	6.24E-27	ELOVL6
220445_s_at	-3.968126454	4.52E-30	CSAG2
242888_at	-3.956452569	4.59E-24	NA
203989_x_at	-3.953777314	9.41E-26	F2R
210145_at	-3.936997211	9.51E-27	PLA2G4A
238689_at	-3.927653449	3.9E-10	GPR110
233364_s_at	-3.908621286	4.39E-25	NA
204256_at	-3.902968671	2.03E-31	ELOVL6

217564_s_at	-3.890419141	6.14E-25	CPS1
235148_at	-3.884689862	3.44E-24	KRTCAP3
203974_at	-3.866536097	7.0E-28	HDHD1A
227082_at	-3.864585197	2.98E-19	NA
232125_at	-3.849416984	5.29E-18	NA
225575_at	-3.845970999	4.97E-26	LIFR
201951_at	-3.841131165	1.01E-28	ALCAM
201952_at	-3.808702785	4.03E-29	ALCAM
223591_at	-3.80165121	1.22E-28	RNF135
226051_at	-3.794081845	1.37E-24	SELM
244111_at	-3.792417998	1.67E-27	KRT222P
230563_at	-3.782348339	2.02E-25	RASGEF1A
211668_s_at	-3.750162363	2.27E-28	PLAU
220513_at	-3.746918551	2.23E-22	KHDC1
204086_at	-3.7337206	1.95E-28	PRAME
205479_s_at	-3.722298056	9.49E-28	PLAU
231270_at	-3.720636164	2.0E-27	CA13
204172_at	-3.702774939	1.2E-26	CPOX
227522_at	-3.690120227	2.12E-28	CMBL
205226_at	-3.686800911	1.07E-25	PDGFRL
206125_s_at	-3.682994854	2.14E-27	KLK8
206504_at	-3.657277738	3.47E-17	CYP24A1
1554726_at	-3.613681754	2.06E-24	ZNF655
238717_at	-3.605031631	4.52E-27	NA
218976_at	-3.590523155	6.32E-15	DNAJC12
219263_at	-3.577439277	2.26E-26	RNF128
204036_at	-3.574944814	5.03E-19	LPAR1
204920_at	-3.566868522	1.03E-22	CPS1
212097_at	-3.557795675	2.44E-25	CAV1
219388_at	-3.539240511	1.94E-27	GRHL2
35201_at	-3.528824527	8.4E-32	HNRNPL
206204_at	-3.522214479	4.28E-22	GRB14
1565162_s_at	-3.517355862	3.35E-25	MGST1
218854_at	-3.503859364	8.86E-25	DSE
1555734_x_at	-3.492347773	2.65E-27	AP1S3
231766_s_at	-3.490052336	7.74E-18	COL12A1
205364_at	-3.481846545	3.14E-22	ACOX2
229019_at	-3.466191232	1.3E-23	ZNF385B
203695_s_at	-3.462992194	2.08E-23	DFNA5
225171_at	-3.443087477	6.38E-26	ARHGAP18
227121_at	-3.4281626	1.36E-20	NA

203060_s_at	-3.412554427	9.34E-27	PAPSS2
229053_at	-3.384903559	8.57E-25	SYT17
209156_s_at	-3.382179814	2.91E-26	COL6A2
219298_at	-3.37738782	5.34E-24	ECHDC3
221113_s_at	-3.374678227	3.5E-19	WNT16
203355_s_at	-3.370498665	1.16E-21	PSD3
231736_x_at	-3.368596188	7.85E-29	MGST1
213094_at	-3.367202397	1.45E-25	GPR126
202489_s_at	-3.359702254	1.62E-21	FXYD3
212092_at	-3.356624139	3.04E-13	PEG10
202525_at	-3.353412135	1.16E-26	PRSS8
207057_at	-3.349627303	8.45E-24	SLC16A7
227070_at	-3.346608462	2.39E-25	GLT8D2
212096_s_at	-3.344989331	4.44E-17	MTUS1
202022_at	-3.344215626	4.25E-23	ALDOC
203153_at	-3.33202304	4.56E-15	IFIT1
210467_x_at	-3.331867005	1.0E-24	MAGEA12
206140_at	-3.326595869	1.11E-21	LHX2
225809_at	-3.322894826	6.38E-26	DKFZP564O0823
204653_at	-3.315954782	7.79E-26	TFAP2A
228261_at	-3.314303056	7.55E-28	MIB2
230356_at	-3.30586423	3.16E-18	NA
1552849_at	-3.30525836	1.38E-26	C2orf65
204751_x_at	-3.290821488	1.8E-24	DSC2
210002_at	-3.269376566	9.4E-25	GATA6
204671_s_at	-3.26738077	5.96E-25	ANKRD6
224918_x_at	-3.248369818	7.04E-29	MGST1
243252_at	-3.246229836	3.02E-23	NA
203650_at	-3.242541388	5.97E-29	PROCR
205153_s_at	-3.242324115	2.89E-25	CD40
225173_at	-3.234731199	8.89E-27	ARHGAP18
203065_s_at	-3.225436457	1.81E-23	CAV1
226250_at	-3.222065773	1.19E-18	NA
212095_s_at	-3.210585956	2.55E-16	MTUS1
229160_at	-3.209625482	9.41E-18	MUM1L1
218692_at	-3.208583974	3.33E-20	GOLSYN
230720_at	-3.207581907	2.41E-22	RNF182
206197_at	-3.193820402	4.26E-26	NME5
205383_s_at	-3.183059099	1.9E-17	ZBTB20
219271_at	-3.177097937	2.14E-24	GALNT14

202948_at	-3.177013167	7.63E-25	IL1R1
228062_at	-3.172877406	1.61E-20	NAP1L5
205581_s_at	-3.15898932	4.48E-27	NOS3
1569878_at	-3.149466187	1.14E-24	CCNYL2
213515_x_at	-3.148007631	1.42E-27	HBG2
206153_at	-3.147855065	3.05E-18	CYP4F11
236207_at	-3.145806606	1.24E-29	SSFA2
204472_at	-3.145799656	6.29E-17	GEM
207761_s_at	-3.145321878	6.13E-24	METTL7A
213807_x_at	-3.143497429	5.0E-27	MET
236016_at	-3.143027237	7.49E-28	NA
206429_at	-3.142523286	2.42E-24	F2RL1
204955_at	-3.141921582	5.31E-25	SRPX
238423_at	-3.141814342	1.42E-09	SYTL3
219750_at	-3.132514368	2.23E-22	TMEM144
226817_at	-3.121466629	1.83E-26	DSC2
230063_at	-3.113323812	2.47E-21	ZNF264
206653_at	-3.111298887	3.8E-26	POLR3G
1564220_a_at	-3.090428336	1.61E-25	NA
215646_s_at	-3.086916374	6.75E-11	VCAN
218613_at	-3.083731234	3.94E-23	PSD3
225664_at	-3.080903286	5.73E-17	COL12A1
204602_at	-3.075203281	7.58E-13	DKK1
205382_s_at	-3.065475664	2.52E-27	CFD
224657_at	-3.063017504	1.5E-23	ERRFI1
226771_at	-3.054194491	7.15E-28	ATP8B2
219313_at	-3.049959847	2.48E-22	GRAMD1C
1555742_at	-3.035934535	6.89E-20	NA
213913_s_at	-3.035221065	2.23E-22	TBC1D30
211599_x_at	-3.014005885	5.87E-27	MET
230033_at	-3.007559616	2.74E-26	C19orf51
206654_s_at	-3.000015858	2.01E-23	POLR3G
200884_at	-2.999726868	4.81E-27	CKB
229393_at	-2.992552015	7.75E-25	L3MBTL3
227176_at	-2.985382221	3.39E-24	SLC2A13
223302_s_at	-2.976060076	5.81E-26	ZNF655
238029_s_at	-2.974147273	1.21E-25	SLC16A14
226420_at	-2.974099614	5.69E-19	EVI1
225450_at	-2.957741769	3.28E-24	AMOTL1
206354_at	-2.95763142	4.16E-21	SLCO1B3
227481_at	-2.95431255	6.92E-25	CNKSR3

218718_at	-2.947594188	1.13E-24	PDGFC
213506_at	-2.945904401	5.52E-26	F2RL1
243880_at	-2.935303789	3.17E-23	GOSR2
218858_at	-2.931609016	2.64E-20	DEPDC6
235515_at	-2.927796442	3.67E-23	C19orf46
219667_s_at	-2.922669403	6.24E-17	BANK1
213355_at	-2.922156919	2.87E-17	ST3GAL6
223674_s_at	-2.915435592	7.24E-23	CDC42SE1
230570_at	-2.912693627	1.5E-24	NA
223895_s_at	-2.901323348	3.66E-21	EPN3
231120_x_at	-2.895096213	8.18E-17	PKIB
229155_at	-2.893568353	4.66E-25	NA
212151_at	-2.878455904	3.3E-12	PBX1
221760_at	-2.878257538	3.04E-20	MAN1A1
209782_s_at	-2.872259794	2.0E-17	DBP
229842_at	-2.864040635	1.22E-21	ELF3
240304_s_at	-2.853596666	3.58E-17	TMC5
229522_at	-2.852737795	9.45E-26	SDR42E1
210508_s_at	-2.85076924	1.27E-18	KCNQ2
212503_s_at	-2.849385138	2.04E-22	DIP2C
221447_s_at	-2.843512172	6.12E-23	GLT8D2
235244_at	-2.842713496	8.1E-25	CCDC58
210445_at	-2.842248415	2.44E-12	FABP6
213309_at	-2.836487168	7.62E-25	PLCL2
222881_at	-2.833888932	4.51E-23	HPSE
219518_s_at	-2.832936119	7.63E-21	ELL3
228152_s_at	-2.831026066	2.35E-20	DDX60L
228365_at	-2.828903792	1.1E-24	CPNE8
239552_at	-2.823653202	6.72E-24	VWDE
238440_at	-2.816260391	5.65E-26	CLYBL
204750_s_at	-2.81191368	6.31E-23	DSC2
226252_at	-2.80950496	1.99E-17	NA
205418_at	-2.805670798	5.85E-27	FES
219959_at	-2.80030706	4.53E-26	MOCOS
203058_s_at	-2.79923752	1.17E-21	PAPSS2
1555461_at	-2.794082881	2.07E-23	NA
214642_x_at	-2.791302546	2.08E-21	MAGEA5
242517_at	-2.772630953	9.58E-24	KISS1R
204015_s_at	-2.7699602	3.15E-27	DUSP4
202506_at	-2.767711148	4.7E-27	SSFA2
220638_s_at	-2.758191237	4.42E-24	CBLC

208116_s_at	-2.75817569	5.15E-21	MAN1A1
207808_s_at	-2.757112623	9.63E-21	PROS1
227163_at	-2.756119378	5.86E-31	GSTO2
206698_at	-2.756093515	4.6E-21	XK
223611_s_at	-2.755286791	1.07E-14	LNX1
239108_at	-2.755162157	9.08E-22	FAR2
238439_at	-2.747656078	3.27E-15	ANKRD22
222433_at	-2.745204447	2.32E-26	ENAH
1558280_s_at	-2.742052225	3.95E-19	ARHGAP29
231769_at	-2.738768145	5.12E-29	FBXO6
224553_s_at	-2.73060438	1.66E-23	TNFRSF18
227204_at	-2.719298401	1.61E-24	PARD6G
215505_s_at	-2.714276881	1.21E-24	STRN3
228115_at	-2.713305041	1.09E-15	NA
229377_at	-2.711026806	5.36E-25	GRTP1
235057_at	-2.710892139	3.83E-22	ITCH
211571_s_at	-2.710027617	1.41E-10	VCAN
212762_s_at	-2.704396785	2.53E-26	TCF7L2
212093_s_at	-2.700261679	2.14E-14	MTUS1
220816_at	-2.700173124	3.15E-21	LPAR3
225911_at	-2.698684704	2.1E-20	NPNT
213110_s_at	-2.687661089	6.55E-22	COL4A5
216511_s_at	-2.687525169	4.16E-26	TCF7L2
242873_at	-2.683762537	3.99E-12	NA
209631_s_at	-2.669081192	9.25E-23	GPR37
205014_at	-2.668687759	9.12E-21	FGFBP1
215411_s_at	-2.659420935	1.03E-23	TRAF3IP2
202180_s_at	-2.659001081	5.12E-20	MVP
1552797_s_at	-2.640691771	1.82E-09	PROM2
231984_at	-2.638588534	1.06E-23	MTAP
225166_at	-2.632803947	1.85E-21	ARHGAP18
205917_at	-2.631145266	1.4E-19	ZNF264
206219_s_at	-2.623368947	3.79E-22	VAV1
212543_at	-2.622852136	4.62E-18	AIM1
204897_at	-2.621565623	8.05E-19	PTGER4
227785_at	-2.616760968	1.7E-25	SDCCAG8
229744_at	-2.615904284	3.11E-30	SSFA2
205139_s_at	-2.609968595	9.9E-22	UST
222662_at	-2.608194996	2.24E-19	PPP1R3B
212136_at	-2.605979124	2.49E-21	ATP2B4
203939_at	-2.595928019	4.57E-13	NT5E

216035_x_at	-2.588755147	3.12E-24	TCF7L2
238681_at	-2.58814915	3.64E-19	GDPD1
224209_s_at	-2.587297003	5.75E-20	GDA
204619_s_at	-2.586155238	1.26E-10	VCAN
233562_at	-2.584864607	1.74E-23	LOC84856
226679_at	-2.584847266	1.6E-19	SLC26A11
201564_s_at	-2.582499682	1.51E-24	FSCN1
204319_s_at	-2.58183463	4.73E-29	RGS10
244246_at	-2.574497798	5.91E-20	MIPOL1
214279_s_at	-2.574449097	5.81E-22	NDRG2
244467_at	-2.570876372	1.57E-20	C22:CTA-250D10.9
227491_at	-2.564046565	1.18E-24	ELOVL6
217820_s_at	-2.562130958	1.1E-23	ENAH
229849_at	-2.560134547	4.69E-24	NA
204014_at	-2.558299779	9.62E-21	DUSP4
201925_s_at	-2.558030298	3.93E-22	CD55
215346_at	-2.552319516	5.86E-17	CD40
200907_s_at	-2.551995251	4.77E-11	PALLD
210138_at	-2.550268024	1.5E-20	RGS20
213839_at	-2.548538976	1.89E-19	CLMN
1569025_s_at	-2.547810792	4.0E-20	FAM13A
202742_s_at	-2.546225199	0.00000109	PRKACB
228946_at	-2.542925602	6.17E-21	INTU
203130_s_at	-2.541420185	6.94E-23	KIF5C
201061_s_at	-2.54131482	4.0E-22	STOM
228551_at	-2.540600024	3.84E-20	DENND5B
206103_at	-2.540167218	2.6E-25	RAC3
219825_at	-2.539667435	3.44E-18	CYP26B1
207558_s_at	-2.53932263	5.59E-20	PITX2
227875_at	-2.536752763	3.96E-23	KLHL13
222376_at	-2.525063494	8.25E-19	NA
223125_s_at	-2.524473663	9.44E-25	C1orf21
235059_at	-2.524201803	1.54E-21	RAB12
219863_at	-2.508380659	3.83E-18	HERC5
222240_s_at	-2.506488897	6.76E-25	ISYNA1
209469_at	-2.494158538	4.26E-17	GPM6A
239468_at	-2.490602425	0.0000704	MKX
1558290_a_at	-2.489411159	6.0E-29	PVT1
239752_at	-2.485450206	7.81E-15	NA
39248_at	-2.481989948	2.6E-15	AQP3

242909_at	-2.475624144	1.98E-22	NA
208925_at	-2.47392851	4.32E-22	CLDND1
210933_s_at	-2.47288671	9.99E-25	FSCN1
216248_s_at	-2.472379771	3.55E-17	NR4A2
204419_x_at	-2.46852187	2.96E-24	HBG2
202786_at	-2.465276854	4.76E-28	STK39
241703_at	-2.462837738	1.67E-23	RUNDC3B
203595_s_at	-2.460553892	1.26E-23	IFIT5
207723_s_at	-2.457058628	1.85E-12	KLRC3
229414_at	-2.450730356	3.11E-23	PITPNC1
210827_s_at	-2.450374333	2.2E-23	ELF3
215016_x_at	-2.449757301	9.25E-19	DST
220334_at	-2.446802574	9.59E-18	RGS17
1552334_at	-2.445591258	1.26E-19	TRIOBP
204201_s_at	-2.441912187	3.58E-14	PTPN13
204622_x_at	-2.433455377	1.16E-16	NR4A2
225541_at	-2.43127256	7.31E-27	RPL22L1
228057_at	-2.424539548	1.32E-17	DDIT4L
236281_x_at	-2.423450679	8.31E-10	HTR7
236117_at	-2.42195651	8.08E-23	NA
225295_at	-2.420013443	1.52E-19	SLC39A10
219562_at	-2.41994567	3.67E-20	RAB26
216685_s_at	-2.418407723	8.44E-26	MTAP
200897_s_at	-2.413700944	4.58E-11	PALLD
204844_at	-2.411813434	4.55E-20	ENPEP
221884_at	-2.408874578	8.51E-18	EVI1
219317_at	-2.406259636	3.92E-24	POLI
226034_at	-2.405437601	2.81E-22	NA
210517_s_at	-2.403974323	2.22E-21	AKAP12
219062_s_at	-2.402414218	2.54E-23	ZCCHC2
204798_at	-2.401685117	2.55E-20	MYB
1557302_at	-2.400892598	6.56E-23	NA
238956_at	-2.39838517	2.28E-14	NA
239078_at	-2.397673309	1.64E-24	C1orf58
212254_s_at	-2.394675208	2.03E-18	DST
235419_at	-2.391357716	9.17E-20	NA
204379_s_at	-2.38948196	1.17E-16	FGFR3
213912_at	-2.387467978	3.48E-23	TBC1D30
218970_s_at	-2.384507656	1.63E-18	CUTC
218211_s_at	-2.382994542	3.37E-24	MLPH
226559_at	-2.381724319	1.45E-15	IER5L

238794_at	-2.381364868	4.81E-25	C10orf78
241726_at	-2.380834256	7.5E-21	NA
223541_at	-2.376448189	3.97E-14	HAS3
223721_s_at	-2.373262893	8.48E-11	DNAJC12
50965_at	-2.373088131	7.47E-20	RAB26
225974_at	-2.371042464	5.98E-23	TMEM64
210665_at	-2.371023084	2.27E-18	TFPI
242338_at	-2.369495102	9.63E-23	TMEM64
224837_at	-2.362078813	8.81E-21	FOXP1
223276_at	-2.360042631	2.8E-28	MST150
203324_s_at	-2.359112619	4.51E-23	CAV2
216037_x_at	-2.358301977	3.0E-24	TCF7L2
210942_s_at	-2.355113717	4.14E-15	ST3GAL6
220318_at	-2.354217024	1.68E-17	EPN3
204956_at	-2.35136885	3.89E-20	MTAP
228314_at	-2.34298296	5.4E-19	LRRC8C
205530_at	-2.341727308	3.06E-18	ETFDH
206023_at	-2.336700825	9.56E-25	NMU
213032_at	-2.336382918	4.6E-17	NFIB
241386_at	-2.335936188	1.32E-20	NA
206315_at	-2.334084738	4.31E-18	CRLF1
238649_at	-2.332228407	4.63E-21	PITPNC1
219501_at	-2.32975691	3.77E-15	ENOX1
203910_at	-2.324574039	2.74E-22	ARHGAP29
226857_at	-2.323803698	2.77E-26	ARHGEF19
1554149_at	-2.322576045	5.36E-21	CLDND1
225524_at	-2.319532144	5.52E-13	ANTXR2
231930_at	-2.316968436	1.67E-17	ELMOD1
219745_at	-2.315155002	3.42E-17	TMEM180
238846_at	-2.310058123	2.61E-18	TNFRSF11A
213715_s_at	-2.307354435	1.03E-21	KANK3
205442_at	-2.306395681	4.09E-18	MFAP3L
1554291_at	-2.299490869	7.31E-19	UHRF1BP1L
239367_at	-2.299091115	2.66E-09	BDNF
213943_at	-2.298439083	3.32E-20	TWIST1
1559072_a_at	-2.297135034	3.2E-17	ELFN2
203788_s_at	-2.295512365	1.59E-21	SEMA3C
205164_at	-2.295046411	9.87E-25	GCAT
223851_s_at	-2.290556897	2.02E-21	TNFRSF18
209290_s_at	-2.289841846	2.54E-18	NFIB

212761_at	-2.28766511	5.28E-27	TCF7L2
201860_s_at	-2.287091723	7.09E-21	PLAT
200906_s_at	-2.285367635	2.52E-09	PALLD
201510_at	-2.284003238	1.48E-21	ELF3
219403_s_at	-2.281773111	2.39E-22	HPSE
205129_at	-2.281750982	6.03E-23	NPM3
1554008_at	-2.280937971	2.33E-23	OSMR
203571_s_at	-2.274123013	6.83E-21	C10orf116
1557638_at	-2.273933699	5.87E-22	NA
235725_at	-2.273083712	5.0E-23	SMAD4
214164_x_at	-2.271581144	9.77E-20	CA12
230003_at	-2.268851852	9.36E-17	NA
211363_s_at	-2.266075096	9.21E-25	MTAP
222196_at	-2.264699986	2.27E-19	LOC286434
208146_s_at	-2.257275585	1.22E-19	CPVL
201926_s_at	-2.25406707	1.0E-22	CD55
223533_at	-2.253817136	8.11E-23	LRRC8C
211162_x_at	-2.24419709	6.8E-25	SCD
219371_s_at	-2.243542755	3.03E-16	KLF2
235046_at	-2.241750319	1.55E-09	NA
212150_at	-2.240688215	2.03E-25	EFR3A
1558111_at	-2.239388205	3.03E-23	MBNL1
206396_at	-2.238842124	1.5E-19	SLC1A1
212759_s_at	-2.234653689	8.48E-22	TCF7L2
226806_s_at	-2.232823927	1.67E-17	NA
209289_at	-2.231123874	4.08E-21	NFIB
201060_x_at	-2.228081779	2.07E-23	STOM
202887_s_at	-2.226973135	7.03E-10	DDIT4
217599_s_at	-2.226474718	1.39E-21	MDFIC
207818_s_at	-2.219821009	1.48E-08	HTR7
235429_at	-2.219775815	2.32E-20	NA
236738_at	-2.215289746	4.56E-17	LOC401097
223278_at	-2.213288887	1.0E-23	GJB2
223741_s_at	-2.212339789	1.33E-22	TTYH2
212190_at	-2.207184847	4.38E-13	SERPINE2
212192_at	-2.206906628	4.81E-12	KCTD12
238067_at	-2.206885272	9.93E-18	TBC1D8B
206108_s_at	-2.206342102	4.73E-26	SFRS6
213470_s_at	-2.203941459	1.84E-24	HNRNPH1
224391_s_at	-2.203768456	5.49E-24	SIAE
208820_at	-2.202814572	7.26E-28	PTK2

210683_at	-2.202481324	5.67E-21	NRTN
222997_s_at	-2.202198979	5.39E-25	MRPS21
228304_at	-2.20173752	2.26E-20	RBM43
210674_s_at	-2.201371666	1.94E-24	PCDHA12
205613_at	-2.199178015	9.53E-21	SYT17
212646_at	-2.197875564	6.88E-20	RFTN1
229014_at	-2.195243947	3.72E-23	FLJ42709
237301_at	-2.194549347	1.66E-12	NA
226597_at	-2.191144077	3.35E-24	REEP6
222904_s_at	-2.19096284	1.7E-14	TMC5
221002_s_at	-2.188630371	1.27E-26	TSPAN14
238935_at	-2.187654217	8.23E-20	RPS27L
232038_at	-2.186699925	5.55E-19	C6orf170
244190_at	-2.186679423	1.75E-16	PNPLA8
225387_at	-2.185394243	2.95E-24	TSPAN5
231597_x_at	-2.180740468	3.46E-18	NA
211330_s_at	-2.177374872	4.14E-21	HFE
1555950_a_at	-2.176171757	5.14E-23	CD55
214647_s_at	-2.175408957	2.17E-22	HFE
213358_at	-2.174443225	3.71E-21	KIAA0802
236489_at	-2.172194371	0.0000874	NA
225388_at	-2.168822335	2.14E-24	TSPAN5
207753_at	-2.168802352	3.88E-17	ZNF304
201724_s_at	-2.165669027	5.95E-23	GALNT1
1568612_at	-2.165363384	1.02E-16	GABRG2
201312_s_at	-2.162356546	5.56E-22	SH3BGRL
203238_s_at	-2.158115439	2.55E-21	NOTCH3
229396_at	-2.151821585	4.59E-21	OVOL1
221577_x_at	-2.151709512	2.31E-17	GDF15
227468_at	-2.149524868	6.6E-20	CPT1C
33494_at	-2.144454036	1.07E-21	ETFDH
1558605_at	-2.139598979	2.75E-13	NA
204848_x_at	-2.13747121	9.37E-22	HBG1
1554679_a_at	-2.136531549	2.14E-24	LAPTM4B
206492_at	-2.13101951	8.5E-18	FHIT
209846_s_at	-2.128564814	1.23E-21	BTN3A2
219517_at	-2.127827846	1.34E-18	ELL3
213310_at	-2.127519626	9.11E-20	EIF2C2
208029_s_at	-2.126987611	2.49E-26	LAPTM4B
203323_at	-2.123316848	2.92E-19	CAV2
242979_at	-2.123012504	1.96E-17	IRS1

218986_s_at	-2.122216262	1.44E-17	DDX60
209470_s_at	-2.119516833	5.44E-16	GPM6A
1552390_a_at	-2.117108474	2.27E-18	C8orf47
235308_at	-2.115841372	3.33E-13	ZBTB20
219358_s_at	-2.114960416	4.34E-18	ADAP2
211708_s_at	-2.114304678	9.39E-25	SCD
236045_x_at	-2.111589877	3.11E-16	NA
235004_at	-2.109642594	9.67E-16	RBM24
201311_s_at	-2.101077609	2.39E-21	SH3BGRL
219580_s_at	-2.096086728	5.07E-14	TMC5
208767_s_at	-2.095274155	1.84E-26	LAPTM4B
208818_s_at	-2.091826751	1.27E-23	COMT
210135_s_at	-2.090304766	7.66E-22	SHOX2
209615_s_at	-2.088132125	3.07E-23	PAK1
222592_s_at	-2.087742598	1.25E-09	ACSL5
230055_at	-2.087483511	5.02E-19	KHDC1
225459_at	-2.085854697	1.8E-21	AMOTL1
224838_at	-2.085138059	2.64E-20	FOXP1
202345_s_at	-2.084080705	3.87E-24	FABP5
227839_at	-2.083256509	4.26E-17	MBD5
219179_at	-2.082352139	7.19E-11	DACT1
234775_at	-2.081043614	7.7E-19	OR51B5
225959_s_at	-2.079950407	2.62E-26	ZNRF1
213603_s_at	-2.079568923	7.74E-26	RAC2
64408_s_at	-2.078774794	1.24E-18	CALML4
209890_at	-2.075707099	3.42E-26	TSPAN5
212149_at	-2.075161897	3.63E-26	EFR3A
242917_at	-2.074628664	3.4E-20	RASGEF1A
1560318_at	-2.071669901	3.42E-17	ARHGAP29
223515_s_at	-2.063004351	2.51E-25	COQ3
225666_at	-2.062031945	2.02E-25	TMTC4
218322_s_at	-2.061837645	9.4E-09	ACSL5
204621_s_at	-2.061823092	1.3E-14	NR4A2
227856_at	-2.061406395	1.45E-21	C4orf32
235259_at	-2.05960176	8.58E-19	PACRGL
228221_at	-2.057100154	1.72E-19	SLC44A3
234513_at	-2.056798045	4.76E-15	ELOVL3
227438_at	-2.052855149	9.6E-24	ALPK1
223599_at	-2.051444212	1.14E-11	TRIM6
206241_at	-2.050136178	7.15E-24	KPNA5
207821_s_at	-2.049817423	1.89E-28	PTK2

209183_s_at	-2.04766941	5.25E-19	C10orf10
202741_at	-2.046644699	0.00000455	PRKACB
239196_at	-2.045578374	1.01E-16	ANKRD22
204820_s_at	-2.044677939	5.46E-23	BTN3A3
215321_at	-2.041116616	1.47E-20	RUNDC3B
222558_at	-2.039580946	4.55E-25	RPRD1A
204560_at	-2.039232955	2.21E-21	FKBP5
220474_at	-2.034038032	3.17E-14	SLC25A21
208817_at	-2.030383448	7.1E-23	COMT
211105_s_at	-2.030065139	1.73E-20	NFATC1
224976_at	-2.02987832	1.16E-20	NFIA
223435_s_at	-2.029598288	6.64E-19	PCDHA@
224856_at	-2.029206572	5.97E-23	FKBP5
212253_x_at	-2.024077043	1.72E-16	DST
217847_s_at	-2.023683745	9.15E-23	THRAP3
229590_at	-2.023221496	5.52E-24	RPL13
222559_s_at	-2.021994511	1.91E-26	RPRD1A
215867_x_at	-2.021876556	4.06E-15	CA12
202457_s_at	-2.020728155	2.74E-20	PPP3CA
1557103_a_at	-2.019716166	6.99E-16	LMTK3
231768_at	-2.019542023	4.48E-17	USF1
224840_at	-2.017108093	2.35E-25	FKBP5
36475_at	-2.016527424	4.66E-27	GCAT
204040_at	-2.016436312	4.74E-18	RNF144A
235871_at	-2.013394239	2.51E-17	LIPH
235325_at	-2.00983031	3.38E-20	SPG7
212135_s_at	-2.008966591	1.44E-18	ATP2B4
230281_at	-2.002073034	1.96E-21	C16orf46
227126_at	-2.000756882	6.6E-16	PTPRG
228653_at	-2.00053394	2.62E-15	SAMD5

DOWN-REGULATED BY SMAD4			
Affymetrix probe	**Coef**	**FDR**	**Gene Symbol**
223948_s_at	2.001595545	1.13E-20	TMPRSS3
217585_at	2.003528658	7.13E-18	NEBL
211160_x_at	2.004615463	1.72E-24	ACTN1
1552562_at	2.007497781	1.07E-19	ZNF570
1552423_at	2.008817735	1.37E-17	ETV3
1552740_at	2.009394831	6.78E-17	C2orf15

227013_at	2.010912568	6.69E-24	LATS2
219440_at	2.011493007	2.29E-18	RAI2
224204_x_at	2.012325661	1.24E-20	ARNTL2
227808_at	2.012557765	4.40E-24	DNAJC15
1557448_a_at	2.012567813	2.06E-18	NA
228030_at	2.020441304	1.24E-19	RBM6
36920_at	2.021345788	6.65E-18	MTM1
1555852_at	2.022349974	1.28E-16	PSMB9
208712_at	2.02402084	1.88E-24	CCND1
235729_at	2.024360154	1.07E-17	ZNF514
204254_s_at	2.024505089	1.41E-20	VDR
229576_s_at	2.024874112	7.17E-23	TBX3
208682_s_at	2.025867161	8.50E-21	MAGED2
209710_at	2.026305792	2.03E-20	GATA2
241418_at	2.026645289	2.56E-14	LOC344887
230398_at	2.028443161	2.94E-19	TNS4
201826_s_at	2.030994562	6.39E-17	SCCPDH
213924_at	2.032093713	1.81E-20	MPPE1
1553713_a_at	2.032228876	5.78E-20	RHEBL1
1554576_a_at	2.033707581	1.43E-18	ETV4
1556533_at	2.033752706	2.89E-20	C17orf52
223616_at	2.034100611	3.53E-19	ZNF649
219884_at	2.03499411	8.53E-24	LHX6
243918_at	2.037096622	4.12E-17	NA
227822_at	2.037255454	8.90E-18	ZNF605
203627_at	2.037826204	1.21E-19	IGF1R
219873_at	2.038049187	1.83E-15	COLEC11
203438_at	2.039098575	7.21E-16	STC2
223408_s_at	2.039693313	7.48E-20	FOXK2
219565_at	2.04111855	4.20E-22	CYP20A1
226342_at	2.04136412	1.24E-18	SPTBN1
230352_at	2.042149583	1.42E-22	PRPS2
210869_s_at	2.043353091	2.99E-22	MCAM
202082_s_at	2.04459877	1.32E-20	SEC14L1
220921_at	2.045143936	4.89E-16	SPANXA1
243010_at	2.045331613	2.60E-23	MSI2
204683_at	2.047862438	2.17E-19	ICAM2
206286_s_at	2.047952074	2.16E-18	TDGF1
1568813_at	2.047958281	3.53E-19	NA
212549_at	2.049792864	7.68E-24	STAT5B
226092_at	2.051485079	3.83E-22	MPP5

210510_s_at	2.052339859	7.65E-18	NRP1
227897_at	2.05331385	2.78E-16	RAP2B
207401_at	2.053682625	2.22E-18	PROX1
202124_s_at	2.055352439	9.63E-19	TRAK2
227566_at	2.056882796	7.44E-13	NTM
209917_s_at	2.057188894	6.74E-20	TP53TG1
239154_at	2.057287968	4.12E-20	NA
212774_at	2.057394225	1.60E-19	ZNF238
231149_s_at	2.057565784	6.60E-19	ULK4
232088_x_at	2.059196391	2.20E-18	hCG_2039027
232140_at	2.060920271	6.20E-20	LOC100132352
209737_at	2.061799186	1.46E-19	MAGI2
229826_at	2.063509151	3.04E-23	LOC440957
209496_at	2.064566187	1.98E-18	RARRES2
1553055_a_at	2.066659568	1.30E-15	SLFN5
219826_at	2.067977547	1.21E-17	ZNF419
207468_s_at	2.069248903	6.63E-25	SFRP5
225674_at	2.069280895	2.44E-24	BCAP29
204881_s_at	2.070045873	1.20E-20	UGCG
244570_at	2.070129534	3.71E-22	LOC100130360
200953_s_at	2.071157243	2.08E-18	CCND2
240690_at	2.071931025	6.39E-18	NA
233949_s_at	2.072414505	1.60E-19	MYH7B
219389_at	2.07248453	3.26E-19	SUSD4
228293_at	2.072505967	5.32E-20	DEPDC7
1560156_at	2.073033034	5.95E-20	NA
210196_s_at	2.075470397	2.53E-21	PSG1
207413_s_at	2.076069147	8.75E-20	SCN5A
225564_at	2.076641127	1.62E-22	SPATA13
201830_s_at	2.077757208	1.53E-22	NET1
220265_at	2.078674309	3.57E-20	GPR107
226657_at	2.079603059	1.66E-21	MGC33894
243065_at	2.082862865	3.14E-19	NA
217707_x_at	2.08591528	2.16E-10	SMARCA2
200660_at	2.088782815	3.93E-26	S100A11
204831_at	2.08913978	8.25E-21	CDK8
225677_at	2.089658537	7.74E-24	BCAP29
215237_at	2.089948665	1.15E-18	DOCK9
226040_at	2.090890673	4.53E-14	NA

226267_at	2.091210936	1.40E-17	JDP2
239050_s_at	2.092050641	6.10E-17	NA
202637_s_at	2.092272057	5.21E-17	ICAM1
1555270_a_at	2.092656737	3.78E-23	WFS1
229061_s_at	2.093069871	2.45E-18	SLC25A13
219264_s_at	2.093702869	3.63E-19	PPP2R3B
205500_at	2.095203785	4.31E-19	C5
232382_s_at	2.09796432	9.34E-19	PCMTD1
223916_s_at	2.098387577	2.43E-17	BCOR
214021_x_at	2.101508248	5.05E-21	ITGB5
208637_x_at	2.103540823	8.57E-25	ACTN1
231034_s_at	2.10436072	6.78E-18	NHSL1
223974_at	2.108570947	3.77E-18	MGC11082
225844_at	2.10878755	5.15E-23	POLE4
224919_at	2.11114806	4.05E-25	MRPS6
219496_at	2.112316233	2.72E-22	ANKRD57
230002_at	2.112740538	1.19E-22	CLCC1
223874_at	2.113688945	6.77E-19	ARP11
228843_at	2.114183729	2.49E-19	NA
203279_at	2.117560595	5.97E-23	EDEM1
212298_at	2.118482295	8.49E-16	NRP1
204214_s_at	2.119549905	6.18E-24	RAB32
213393_at	2.120549932	2.40E-20	NA
220054_at	2.121974766	5.46E-20	IL23A
240151_at	2.123985943	3.56E-21	LOC404266
203002_at	2.128986284	1.11E-23	AMOTL2
238624_at	2.129004934	2.15E-21	NA
229294_at	2.12918493	7.99E-20	JPH3
230689_at	2.129656116	1.42E-20	NA
215238_s_at	2.131748535	7.00E-19	DOCK9
206337_at	2.13208729	7.74E-14	CCR7
225512_at	2.13353456	6.10E-22	ZBTB38
203628_at	2.134682951	1.29E-19	IGF1R
219321_at	2.136116032	6.08E-21	MPP5
209356_x_at	2.138117527	3.63E-17	EFEMP2
227261_at	2.138432886	3.36E-18	KLF12
201566_x_at	2.139151764	2.14E-10	ID2
209619_at	2.139220283	2.66E-17	CD74
205801_s_at	2.139690005	2.71E-23	RASGRP3
229225_at	2.140132561	2.06E-15	NRP2
200856_x_at	2.140995159	2.13E-22	LOC100133

			918
203231_s_at	2.142190856	3.23E-16	ATXN1
223405_at	2.145369713	4.13E-18	NPL
222902_s_at	2.146389349	2.01E-22	DEM1
230120_s_at	2.147590245	4.86E-20	PLGLB2
236223_s_at	2.147713742	5.30E-21	RIT1
228638_at	2.152274214	3.70E-22	FAM76A
206665_s_at	2.154857428	2.50E-21	BCL2L1
228825_at	2.15547215	1.73E-19	PTGR1
228400_at	2.157018484	4.07E-17	SHROOM3
228290_at	2.157621704	5.01E-18	NCRNA00153
1562056_at	2.157678768	6.04E-17	NA
230434_at	2.158918924	1.69E-21	KLHL23
209383_at	2.159584029	3.12E-19	DDIT3
204220_at	2.159756718	7.27E-20	GMFG
205463_s_at	2.160642566	5.79E-21	PDGFA
226453_at	2.161034619	1.01E-26	RNASEH2C
235094_at	2.163013073	2.15E-17	NA
225045_at	2.16464947	4.81E-20	CCDC88A
230954_at	2.165301557	1.95E-15	C20orf112
227410_at	2.166369676	8.09E-22	FAM43A
1555883_s_at	2.17037673	2.48E-21	SPIN3
201464_x_at	2.171310811	1.69E-21	JUN
228462_at	2.17165176	4.40E-19	IRX2
225946_at	2.172986747	5.12E-20	RASSF8
217966_s_at	2.173719784	9.05E-17	FAM129A
206632_s_at	2.174145788	6.91E-19	APOBEC3B
229512_at	2.175074274	6.07E-18	FAM120C
236571_at	2.175346953	1.38E-19	NA
1564413_at	2.175589441	5.12E-23	FLJ36116
201905_s_at	2.175816402	5.69E-25	CTDSPL
204288_s_at	2.176209802	7.95E-21	SORBS2
206463_s_at	2.181560277	1.59E-17	DHRS2
204345_at	2.183398049	6.40E-20	COL16A1
1557523_at	2.185284448	2.79E-17	LOC92270
226947_at	2.189390496	5.78E-20	GUSBL2
209396_s_at	2.18985411	1.18E-19	CHI3L1
201125_s_at	2.190253032	1.05E-25	ITGB5
205761_s_at	2.191036913	3.63E-23	DUS4L
226979_at	2.191835189	2.25E-23	MAP3K2

204416_x_at	2.192392719	8.69E-21	APOC1
227452_at	2.192587304	1.81E-11	NA
209565_at	2.192604649	3.24E-24	RNF113A
236982_at	2.193180652	5.90E-22	NA
1553992_s_at	2.194080028	4.19E-21	NBR2
229835_s_at	2.195955942	3.76E-23	SLMO2
239376_at	2.197553452	9.00E-21	NA
230752_at	2.200905719	1.35E-17	NA
225656_at	2.201408373	1.40E-19	EFHC1
229332_at	2.20261466	2.48E-23	HPDL
213229_at	2.20769639	3.82E-18	DICER1
1553587_a_at	2.208114119	3.18E-22	POLE4
223689_at	2.209418779	4.34E-23	IGF2BP1
223764_x_at	2.210973597	3.14E-21	NIPSNAP3B
209107_x_at	2.211818881	4.68E-22	NCOA1
235589_s_at	2.212601928	1.47E-19	MDM4
210051_at	2.213887075	2.04E-22	RAPGEF3
1569366_a_at	2.214491667	6.87E-22	ZNF569
232884_s_at	2.214818286	1.75E-18	ZNF853
214022_s_at	2.215183032	2.68E-12	IFITM1
218113_at	2.215609105	7.57E-14	TMEM2
200739_s_at	2.215985691	3.13E-16	SUMO3
242550_at	2.216052553	2.79E-17	EIF3B
1565347_s_at	2.216537648	9.72E-20	TFE3
243802_at	2.217170512	8.25E-21	DNAH12
205600_x_at	2.217420008	2.58E-22	HOXB5
1557128_at	2.217702135	1.34E-21	FAM111B
226283_at	2.217711829	3.93E-18	WDR51B
226137_at	2.21788027	5.12E-15	ZFHX3
244174_at	2.21822943	7.70E-19	NA
231697_s_at	2.222143822	6.69E-19	TMEM49
233825_s_at	2.225136539	1.83E-22	CD99L2
210249_s_at	2.22765983	1.64E-22	NCOA1
215923_s_at	2.228127747	1.63E-22	PSD4
219748_at	2.230102725	2.32E-19	TREML2
1557164_a_at	2.230837534	3.13E-18	NA
206777_s_at	2.232567152	1.17E-20	CRYBB2
204972_at	2.233495323	5.35E-17	OAS2
216867_s_at	2.233724387	4.64E-18	PDGFA
229128_s_at	2.240231277	1.18E-25	ANP32E
201721_s_at	2.241516614	8.31E-18	LAPTM5

218435_at	2.242415626	1.96E-26	DNAJC15
220658_s_at	2.243257146	2.04E-20	ARNTL2
219655_at	2.243716362	2.88E-18	C7orf10
1562020_s_at	2.244373268	9.51E-16	NT5DC4
236557_at	2.247268716	3.38E-21	ZBTB38
1569349_at	2.248470451	3.63E-19	C11orf30
223253_at	2.255457401	7.77E-25	EPDR1
212285_s_at	2.255846238	1.21E-23	AGRN
1553115_at	2.257716576	3.93E-22	NKD1
235728_at	2.258440694	3.37E-22	ZFP3
236313_at	2.262691633	2.80E-16	CDKN2B
238909_at	2.264083249	2.24E-19	S100A10
219833_s_at	2.266662965	2.72E-21	EFHC1
237213_at	2.267957724	1.89E-22	NA
206542_s_at	2.268495126	4.43E-11	SMARCA2
230588_s_at	2.26893631	3.24E-22	LOC285074
202166_s_at	2.26894137	4.09E-25	PPP1R2
226360_at	2.270965971	7.18E-26	ZNRF3
219594_at	2.276200112	9.30E-20	NINJ2
225619_at	2.27933342	1.70E-20	SLAIN1
243539_at	2.280464316	2.97E-18	KIAA1841
212843_at	2.280821942	7.26E-21	NCAM1
219228_at	2.281481382	1.36E-20	ZNF331
218746_at	2.283749284	2.55E-18	TAPBPL
206110_at	2.283753927	2.17E-16	HIST1H3H
242133_s_at	2.284749335	8.59E-20	NA
224917_at	2.287054718	2.25E-19	TMEM49
206482_at	2.289575164	5.99E-18	PTK6
1566129_at	2.292361106	1.51E-19	LIMS1
203382_s_at	2.297054498	9.84E-21	APOE
238530_at	2.297178011	1.08E-15	NNT
203342_at	2.297696771	5.23E-25	TIMM17B
215041_s_at	2.302335267	6.63E-19	DOCK9
214182_at	2.302450885	1.27E-23	LOC100132430
238017_at	2.30266678	3.00E-19	SDR16C5
222590_s_at	2.304525182	1.34E-20	NLK
228235_at	2.304826282	8.48E-20	MGC16121
209663_s_at	2.305761611	1.07E-17	ITGA7
228044_at	2.307392289	1.90E-20	SERP2
1556769_a_at	2.308736467	3.45E-11	NA

214079_at	2.3092358	3.59E-17	DHRS2
214257_s_at	2.317524854	4.09E-24	SEC22B
235456_at	2.31789932	8.56E-20	NA
212538_at	2.318434416	1.05E-19	DOCK9
208868_s_at	2.3191494	1.61E-18	GABARAPL1
201185_at	2.321616506	4.43E-19	HTRA1
229146_at	2.334810625	2.86E-20	C7orf31
214718_at	2.339950988	2.70E-21	GATAD1
229347_at	2.342111401	4.01E-20	NA
217967_s_at	2.342287847	1.60E-16	FAM129A
213627_at	2.343989481	6.02E-22	MAGED2
211094_s_at	2.349676385	5.48E-19	NF1
1557470_at	2.35214077	1.74E-21	SPATA13
205560_at	2.352471305	3.91E-22	PCSK5
210852_s_at	2.352904467	7.40E-20	AASS
1557129_a_at	2.355912134	5.23E-17	FAM111B
235287_at	2.356216674	3.70E-18	CDK6
213788_s_at	2.356276173	1.72E-22	BRD3
214071_at	2.356849843	5.05E-20	GNAL
225927_at	2.359979205	6.73E-23	MAP3K1
222900_at	2.362214604	1.80E-22	NA
204072_s_at	2.362433975	2.33E-15	FRY
239186_at	2.366138856	1.21E-17	MGC39372
1553539_at	2.368424487	1.64E-21	KRT74
204417_at	2.372461661	3.51E-17	GALC
218031_s_at	2.373123466	1.22E-21	FOXN3
218625_at	2.37475004	1.05E-19	NRN1
230454_at	2.375836886	3.37E-18	ICA1L
205559_s_at	2.376415248	7.13E-22	PCSK5
227396_at	2.376681062	4.62E-20	PTPRJ
230333_at	2.377484203	8.53E-18	NA
235174_s_at	2.378567624	4.40E-17	LOC100128822
224279_s_at	2.37939089	5.20E-21	CABYR
1552846_s_at	2.380488849	1.13E-16	RAB42
1555486_a_at	2.380667612	6.67E-22	FLJ14213
242005_at	2.381646968	8.49E-09	NA
203851_at	2.386195698	2.97E-16	IGFBP6
1562722_at	2.386480515	5.89E-18	PRR20
202688_at	2.386891314	3.07E-24	TNFSF10
222877_at	2.387693277	9.24E-20	NA

206154_at	2.389289678	3.16E-19	RLBP1
1568838_at	2.390377891	2.91E-19	LOC100132169
1569157_s_at	2.390953677	4.47E-20	ZNF846
223690_at	2.39209856	3.23E-20	LTBP2
212807_s_at	2.392381239	1.79E-20	SORT1
218359_at	2.393447801	1.13E-19	NRSN2
210948_s_at	2.396963985	2.17E-21	LEF1
213931_at	2.397088384	9.10E-12	ID2
203008_x_at	2.400463233	1.75E-24	TXNDC9
1569194_at	2.400501605	5.29E-23	ZNF789
236616_at	2.401208101	2.66E-20	NA
203849_s_at	2.405282122	5.29E-22	KIF1A
235033_at	2.40613372	3.05E-22	NPEPL1
1557169_x_at	2.407279654	4.04E-20	HCG11
219342_at	2.40841727	1.64E-21	CASD1
218691_s_at	2.408811819	5.68E-23	PDLIM4
219704_at	2.411197714	4.27E-18	YBX2
209185_s_at	2.411377875	9.94E-21	IRS2
1558953_s_at	2.416290339	3.41E-22	CEP164
235106_at	2.416651379	2.50E-25	MAML2
244819_x_at	2.417623502	8.57E-20	PSPH
223223_at	2.424008178	2.72E-24	ARV1
228083_at	2.427311316	1.50E-24	CACNA2D4
202778_s_at	2.428996215	3.76E-22	ZMYM2
212531_at	2.429591585	0.0025938	LCN2
232579_at	2.431224355	5.92E-22	LOC100134229
211758_x_at	2.43151623	1.34E-23	TXNDC9
224687_at	2.431689668	6.22E-23	ANKIB1
206523_at	2.432252009	4.38E-22	CYTH3
212848_s_at	2.433040461	4.37E-26	C9orf3
227444_at	2.433107577	8.87E-24	ARMCX4
231102_at	2.434022619	3.73E-20	CROT
217248_s_at	2.441686527	4.84E-19	SLC7A8
220945_x_at	2.441891651	2.20E-21	MANSC1
216242_x_at	2.442773935	3.65E-28	POLR2J2
240436_at	2.442979333	2.90E-20	LOC650794
231215_at	2.447110217	1.04E-19	NA
202165_at	2.447219345	5.31E-26	PPP1R2
243362_s_at	2.447771709	1.81E-25	LOC641518

225469_at	2.448593445	2.44E-22	LYRM5
1553112_s_at	2.448684642	2.53E-21	CDK8
230233_at	2.450919218	1.04E-20	NA
218656_s_at	2.451324483	3.32E-18	LHFP
201124_at	2.455181162	1.19E-21	ITGB5
212845_at	2.458030326	8.27E-23	SAMD4A
214830_at	2.458110752	5.50E-20	SLC38A6
229376_at	2.458315358	2.74E-18	PROX1
1553430_a_at	2.45967347	5.72E-23	EDARADD
217867_x_at	2.461823473	2.74E-26	BACE2
211029_x_at	2.462131571	4.47E-21	FGF18
211518_s_at	2.462756098	5.24E-19	BMP4
234300_s_at	2.466006077	5.91E-24	ZFP28
230493_at	2.467010452	3.12E-20	SHISA2
228222_at	2.470152644	1.45E-28	PPP1CB
238692_at	2.471620829	5.90E-20	BTBD11
225142_at	2.473334323	8.04E-20	JHDM1D
228415_at	2.481084749	4.98E-24	AP1S2
221552_at	2.482259348	4.01E-23	ABHD6
208291_s_at	2.484371863	2.13E-22	TH
45288_at	2.484689693	5.32E-23	ABHD6
225782_at	2.4885047	5.62E-17	MSRB3
242940_x_at	2.488792755	8.27E-17	DLX6
238790_at	2.49042345	1.26E-18	LOC374443
1556029_s_at	2.492654079	3.77E-21	NMNAT2
202458_at	2.493320695	8.91E-18	PRSS23
223824_at	2.497765348	5.41E-23	C10orf59
229071_at	2.5040351	5.99E-21	C17orf100
227401_at	2.505415699	2.23E-22	IL17D
228606_at	2.507506499	1.92E-24	TM4SF19
219282_s_at	2.507967101	4.90E-22	TRPV2
228846_at	2.508651133	2.25E-22	MXD1
203736_s_at	2.50875283	4.25E-18	PPFIBP1
1552566_at	2.509456886	1.30E-23	BTBD16
224938_at	2.509850856	4.06E-18	NUFIP2
210762_s_at	2.512737366	1.08E-20	DLC1
207279_s_at	2.516595637	5.77E-19	NEBL
206669_at	2.516792867	2.16E-23	GAD1
224848_at	2.523164821	2.19E-15	CDK6
227272_at	2.526906819	1.68E-20	C15orf52
210915_x_at	2.530172079	4.55E-23	IL23A

229588_at	2.534369092	1.51E-19	DNAJC10
202083_s_at	2.538327568	3.87E-22	SEC14L1
223739_at	2.538742913	6.71E-24	PADI1
210657_s_at	2.545125308	4.43E-19	4-Sep
204929_s_at	2.546628907	2.02E-23	VAMP5
220936_s_at	2.547772759	3.12E-20	H2AFJ
214708_at	2.553155318	1.45E-22	SNTB1
230149_at	2.555232646	1.52E-19	NA
206544_x_at	2.555600954	2.93E-13	SMARCA2
204588_s_at	2.558845817	7.98E-20	SLC7A7
209184_s_at	2.559764509	1.12E-20	IRS2
227062_at	2.563096141	4.82E-18	NCRNA00084
201466_s_at	2.56324932	6.42E-19	JUN
229733_s_at	2.56495409	3.65E-24	NA
209525_at	2.565347936	4.24E-22	HDGFRP3
215303_at	2.567164893	1.11E-23	DCLK1
217591_at	2.570261733	1.59E-25	NA
208869_s_at	2.573823294	1.00E-18	GABARAPL1
201109_s_at	2.573994148	4.04E-22	THBS1
205978_at	2.57408841	3.16E-16	KL
205932_s_at	2.576250783	3.08E-18	MSX1
219147_s_at	2.577576972	1.94E-24	C9orf95
242775_at	2.577963951	5.99E-21	NA
1558834_s_at	2.578003331	1.35E-20	C1orf62
229986_at	2.584916186	6.55E-20	ZNF717
222451_s_at	2.588591474	1.65E-28	ZDHHC9
210495_x_at	2.592240002	6.08E-21	FN1
214502_at	2.592304165	2.67E-15	HIST1H2BJ
229972_at	2.594771252	9.40E-20	NA
207198_s_at	2.595742123	1.49E-27	LIMS1
1555613_a_at	2.596414903	3.81E-21	ZAP70
241782_at	2.596834913	4.54E-18	NEBL
205365_at	2.598157886	3.32E-20	HOXB6
205453_at	2.601337422	2.88E-25	HOXB2
238493_at	2.602519404	2.39E-24	ZNF506
200740_s_at	2.602629908	5.91E-20	SUMO3
220663_at	2.605479691	4.84E-20	IL1RAPL1
212415_at	2.608578786	3.59E-27	SEPT6
215596_s_at	2.611561549	5.20E-25	RNF160
226546_at	2.611601139	4.78E-22	NA

1559957_a_at	2.61293232	5.46E-26	LOC642852
213496_at	2.614702672	1.68E-18	LPPR4
227976_at	2.615452007	5.14E-25	LOC644538
223827_at	2.615606039	2.54E-22	TNFRSF19
225273_at	2.616880947	2.38E-18	WWC3
231152_at	2.617052743	3.71E-19	INO80D
224847_at	2.617239964	7.22E-16	CDK6
229667_s_at	2.617988155	1.75E-13	HOXB8
217551_at	2.624708864	2.74E-19	LOC441453
235534_at	2.626553245	2.21E-23	NA
213711_at	2.6270004	1.25E-20	KRT81
212687_at	2.628671505	1.48E-25	LIMS1
218996_at	2.630093152	6.52E-25	TFPT
1570253_a_at	2.632084155	2.62E-20	RHEBL1
202638_s_at	2.632396737	5.56E-20	ICAM1
213167_s_at	2.633073058	1.29E-22	MRPS6
235371_at	2.633729214	6.39E-23	GLT8D4
1553082_at	2.638987331	5.35E-22	CRYGN
211534_x_at	2.639368024	5.39E-22	PTPRN2
229481_at	2.646663455	1.72E-23	NKD1
200857_s_at	2.647355601	6.47E-23	NCOR1
206648_at	2.647885848	7.40E-21	ZNF571
219856_at	2.651534227	4.94E-20	C1orf116
235721_at	2.651585801	3.95E-23	DTX3
213619_at	2.652091616	1.30E-26	HNRNPH1
216092_s_at	2.655620029	2.11E-22	SLC7A8
225908_at	2.661959797	1.29E-20	IAH1
211817_s_at	2.665448517	2.33E-26	KCNJ5
219569_s_at	2.666397326	6.30E-25	TMEM22
222553_x_at	2.670001961	2.75E-23	OXR1
217868_s_at	2.670411347	1.20E-26	METTL9
235433_at	2.671279443	3.16E-23	APOOL
228095_at	2.67185217	1.99E-25	PHF14
229344_x_at	2.675173631	4.78E-23	RIMKLB
214375_at	2.679820282	2.47E-20	PPFIBP1
208916_at	2.687475864	9.33E-27	SLC1A5
211458_s_at	2.687822464	3.64E-19	GABARAPL3
222589_at	2.690853033	9.39E-25	NLK
222453_at	2.691637451	1.13E-22	CYBRD1
209189_at	2.70318444	2.14E-20	FOS

219928_s_at	2.703674363	2.19E-23	CABYR
243752_s_at	2.705237358	6.85E-22	CYTH3
219520_s_at	2.705766291	6.61E-22	WWC3
226001_at	2.707334474	5.62E-18	KLHL5
227200_at	2.709102953	5.28E-27	NA
231101_at	2.709451736	2.47E-26	PPP2R5E
201009_s_at	2.711959391	1.10E-22	TXNIP
212565_at	2.714181192	9.57E-23	STK38L
219383_at	2.714389113	2.14E-21	FLJ14213
212915_at	2.715503393	5.02E-18	PDZRN3
214374_s_at	2.716634372	1.13E-19	PPFIBP1
218318_s_at	2.718494575	3.10E-22	NLK
1558733_at	2.720284614	3.29E-25	ZBTB38
204255_s_at	2.720440918	1.66E-23	VDR
202752_x_at	2.72675246	3.83E-23	SLC7A8
219561_at	2.727281932	5.46E-20	COPZ2
206987_x_at	2.733241904	8.93E-23	FGF18
210886_x_at	2.734013152	1.18E-22	TP53TG1
202086_at	2.734053393	1.34E-20	MX1
218182_s_at	2.734588961	1.43E-23	CLDN1
211538_s_at	2.736931126	1.87E-20	HSPA2
221969_at	2.739111186	9.26E-25	NA
218831_s_at	2.74041632	3.51E-23	FCGRT
204867_at	2.740742067	1.12E-22	GCHFR
226388_at	2.740900213	2.69E-19	TCEA3
222446_s_at	2.741084196	1.09E-29	BACE2
214577_at	2.744038735	7.56E-19	MAP1B
228570_at	2.744539538	8.28E-23	BTBD11
235567_at	2.746922896	1.03E-25	RORA
207156_at	2.74706823	8.98E-21	HIST1H2AG
1554711_at	2.747710605	6.88E-24	CALHM3
239253_at	2.748044295	8.25E-21	NA
226880_at	2.750824571	3.51E-29	NUCKS1
234937_x_at	2.751097385	2.70E-23	ZFP28
242218_at	2.759858224	2.16E-22	PPARD
211259_s_at	2.760066972	1.86E-25	BMP7
228642_at	2.763479147	2.92E-21	NA
235652_at	2.765957674	1.66E-21	NA
219682_s_at	2.769690004	1.30E-23	TBX3
201289_at	2.770069888	6.99E-22	CYR61
206670_s_at	2.771756305	3.63E-26	GAD1

215143_at	2.777088868	2.86E-27	DPY19L2P2
206027_at	2.777527968	2.63E-18	S100A3
223879_s_at	2.778035041	2.40E-24	OXR1
219956_at	2.778433201	6.63E-25	GALNT6
234921_at	2.779834627	1.67E-25	ZNF470
209882_at	2.780056048	8.17E-22	RIT1
221679_s_at	2.783758889	2.15E-25	ABHD6
1554588_a_at	2.787018053	5.53E-23	TTC30B
205278_at	2.787437337	8.43E-25	GAD1
204249_s_at	2.78859104	6.91E-20	LMO2
235085_at	2.788616539	1.87E-25	PRAGMIN
206529_x_at	2.790086911	3.97E-18	SLC26A4
230921_s_at	2.792987468	2.90E-23	PCBP2
210105_s_at	2.798231022	2.73E-19	FYN
223380_s_at	2.800396774	1.20E-27	LATS2
225147_at	2.806400148	1.62E-26	CYTH3
222471_s_at	2.807653539	1.75E-25	KCMF1
236646_at	2.81320894	5.58E-25	C12orf59
226558_at	2.813509149	1.95E-24	MAFIP
204890_s_at	2.815517589	3.15E-23	LCK
230621_at	2.817255849	7.34E-25	IAH1
238694_at	2.818601233	2.38E-24	NA
218747_s_at	2.820068988	7.00E-24	TAPBPL
238657_at	2.82400144	4.47E-22	UBXN10
227949_at	2.825372558	4.84E-25	PHACTR3
210281_s_at	2.828039059	8.43E-24	ZMYM2
37966_at	2.831967027	7.81E-25	PARVB
222447_at	2.840431917	3.13E-25	METTL9
227931_at	2.847320339	2.61E-21	INO80D
233555_s_at	2.847941069	1.21E-24	SULF2
1558762_a_at	2.848265911	2.08E-26	ZNF789
203962_s_at	2.850394472	2.11E-17	NEBL
201110_s_at	2.852335514	1.58E-20	THBS1
233078_at	2.8544519	2.09E-25	API5
224558_s_at	2.854875645	6.10E-22	MALAT1
228443_s_at	2.856889863	1.25E-27	SETD8
223761_at	2.85812778	1.20E-21	FGF19
236266_at	2.858944632	6.05E-23	RORA
213170_at	2.861273196	1.65E-24	GPX7
1557167_at	2.862626168	1.08E-27	HCG11
1555882_at	2.863408774	9.28E-21	SPIN3

216442_x_at	2.864553299	8.49E-22	FN1
214285_at	2.864898698	6.56E-22	FABP3
219471_at	2.875115171	1.22E-21	C13orf18
210764_s_at	2.875186002	3.15E-21	CYR61
215779_s_at	2.878489324	8.14E-21	HIST1H2BG
218704_at	2.879029597	4.53E-12	RNF43
220120_s_at	2.879997514	7.85E-21	EPB41L4A
202196_s_at	2.883172702	7.85E-21	DKK3
203780_at	2.891948589	2.12E-21	MPZL2
205896_at	2.897695863	8.50E-23	SLC22A4
219229_at	2.904763308	1.64E-22	SLCO3A1
229138_at	2.914177929	7.07E-26	PARP11
239201_at	2.92142664	4.56E-25	PFTK2
216033_s_at	2.926287099	1.24E-18	FYN
209684_at	2.926710608	1.72E-21	RIN2
218197_s_at	2.927985173	3.53E-26	OXR1
217875_s_at	2.928532323	5.16E-23	PMEPA1
220198_s_at	2.930982355	4.69E-17	EIF5A2
225999_at	2.936788241	6.42E-25	RIMKLB
207819_s_at	2.940537895	9.51E-22	ABCB4
229715_at	2.942281878	1.85E-23	NA
229374_at	2.944915381	2.32E-19	EPHA4
236224_at	2.945667905	1.34E-24	RIT1
212810_s_at	2.951099474	7.94E-21	SLC1A4
206094_x_at	2.953284642	3.30E-22	UGT1A6
203726_s_at	2.957757615	3.05E-21	LAMA3
211685_s_at	2.96149178	1.05E-23	NCALD
221210_s_at	2.96479409	3.44E-23	NPL
206018_at	2.965856536	7.22E-27	FOXG1
202520_s_at	2.966266328	1.86E-29	MLH1
235296_at	2.96673825	1.28E-17	EIF5A2
205027_s_at	2.96787861	1.54E-19	MAP3K8
221701_s_at	2.969423295	1.34E-24	STRA6
204532_x_at	2.97400928	1.90E-23	UGT1A9
212464_s_at	2.974592447	2.25E-21	FN1
218959_at	2.978045777	9.68E-24	HOXC10
208523_x_at	2.983316263	6.20E-23	HIST1H2BI
208359_s_at	2.988707177	3.11E-25	KCNJ4
209678_s_at	2.990360354	1.67E-27	PRKCI
207433_at	2.992147196	4.72E-27	IL10
209738_x_at	2.992733224	2.80E-24	PSG6

213304_at	2.994941365	1.24E-23	FAM179B
223821_s_at	3.007094886	2.34E-25	SUSD4
222772_at	3.007656919	1.50E-24	MYEF2
204682_at	3.009389542	1.25E-23	LTBP2
207457_s_at	3.010484146	2.24E-25	LY6G6D
1555935_s_at	3.011664905	1.27E-25	HUNK
230659_at	3.014414171	4.62E-24	NA
212678_at	3.018656696	1.64E-22	NF1
1557051_s_at	3.01880922	1.81E-22	NA
212486_s_at	3.021382261	8.28E-21	FYN
239302_s_at	3.022161904	2.52E-23	NA
1569003_at	3.023956205	3.13E-25	TMEM49
236893_at	3.027199566	1.57E-21	LOC404266
242053_at	3.029760682	2.21E-27	TSGA10
203973_s_at	3.033407815	1.60E-27	CEBPD
207126_x_at	3.034107725	2.01E-23	UGT1A1
204702_s_at	3.037002154	1.26E-23	NFE2L3
226985_at	3.03917817	2.59E-22	FGD5
230543_at	3.039482654	1.06E-21	USP9X
219414_at	3.042323841	3.44E-24	CLSTN2
242919_at	3.042397123	3.76E-25	ZNF253
226368_at	3.044476729	2.41E-26	CHST11
229889_at	3.049284265	2.25E-24	C17orf76
226438_at	3.05300151	6.45E-25	SNTB1
222787_s_at	3.054226348	9.72E-22	TMEM106B
226462_at	3.058577494	9.73E-22	STXBP6
212624_s_at	3.058729751	1.93E-15	CHN1
205031_at	3.066126601	5.28E-27	EFNB3
222833_at	3.067725068	3.50E-25	LPCAT2
229307_at	3.067819733	5.74E-26	ANKRD28
218559_s_at	3.068188107	7.24E-24	MAFB
224341_x_at	3.069047039	4.37E-24	TLR4
1553169_at	3.069946087	9.10E-26	LRRN4
206377_at	3.078103353	3.17E-23	FOXF2
242358_at	3.07827769	4.73E-26	NA
213317_at	3.078925255	1.28E-24	CLIC5
205569_at	3.080914454	1.49E-18	LAMP3
225227_at	3.084100368	2.93E-22	NA
210241_s_at	3.088168138	8.27E-27	TP53TG1
207173_x_at	3.088815133	2.98E-19	CDH11
201904_s_at	3.091204573	3.51E-29	CTDSPL

215821_x_at	3.091464394	4.46E-25	PSG3
209395_at	3.099208694	8.28E-24	CHI3L1
242277_at	3.101355236	3.39E-26	NA
1553708_at	3.103480987	4.81E-27	MGC16075
239503_at	3.103590838	1.88E-19	NA
209755_at	3.104083386	5.00E-27	NMNAT2
235749_at	3.104584658	1.90E-20	UGCGL2
216604_s_at	3.106659642	5.35E-27	SLC7A8
202007_at	3.110914043	9.98E-22	NID1
219044_at	3.111074496	5.26E-24	THNSL2
229026_at	3.118641983	1.29E-28	CDC42SE2
212944_at	3.119559517	7.76E-20	SLC5A3
227377_at	3.119606163	3.02E-23	IGF2BP1
239598_s_at	3.120525058	5.40E-25	LPCAT2
211596_s_at	3.124352767	5.91E-22	LRIG1
230413_s_at	3.124654725	8.64E-21	NA
210387_at	3.124719238	2.65E-25	HIST1H2BG
211719_x_at	3.128023957	1.55E-22	FN1
219221_at	3.128483641	3.35E-28	ZBTB38
203779_s_at	3.130956385	8.77E-25	MPZL2
205295_at	3.134799125	4.01E-23	CKMT2
241359_at	3.144879456	1.99E-24	NA
221795_at	3.146945213	5.05E-26	NTRK2
202687_s_at	3.147073082	3.88E-25	TNFSF10
214519_s_at	3.148534536	1.25E-23	RLN2
228425_at	3.148587345	1.70E-24	LOC654433
201906_s_at	3.151942859	1.08E-28	CTDSPL
228291_s_at	3.152102348	1.34E-28	NCRNA00153
242414_at	3.160941765	9.27E-26	QPRT
222067_x_at	3.164923606	6.51E-23	HIST1H2BD
225224_at	3.167873628	7.19E-25	C20orf112
228260_at	3.167937441	1.93E-23	ELAVL2
202150_s_at	3.169633642	1.48E-24	NEDD9
209930_s_at	3.175048161	7.98E-24	NFE2
235027_at	3.176345734	2.49E-22	NA
213916_at	3.18346116	1.53E-23	ZNF20
222760_at	3.188897031	4.01E-25	ZNF703
210195_s_at	3.189117192	4.06E-21	PSG1
225799_at	3.190146974	7.81E-24	LOC541471
228496_s_at	3.191692441	2.44E-24	CRIM1

204268_at	3.194336992	2.70E-12	S100A2
208527_x_at	3.197008809	1.13E-20	HIST1H2BE
214469_at	3.197656593	2.54E-21	HIST1H2AE
211819_s_at	3.202997316	3.66E-26	SORBS1
228834_at	3.210596069	1.84E-28	TOB1
214012_at	3.215036503	9.45E-27	ERAP1
242931_at	3.219085974	3.83E-21	LONRF3
229464_at	3.220409334	1.06E-24	MYEF2
241710_at	3.238456722	3.97E-26	hCG_1645220
209594_x_at	3.249023467	3.30E-23	PSG9
1554241_at	3.250334949	6.28E-28	COCH
205366_s_at	3.251242849	1.50E-25	HOXB6
206404_at	3.25327442	8.87E-23	FGF9
226381_at	3.255810971	3.17E-24	PS1TP4
212676_at	3.259798167	2.42E-28	NF1
203892_at	3.260253426	8.91E-26	WFDC2
227145_at	3.261117526	1.73E-26	LOXL4
230508_at	3.268772847	9.99E-22	DKK3
207210_at	3.272731269	3.25E-20	GABRA3
205258_at	3.273248638	3.05E-26	INHBB
209904_at	3.277786804	5.07E-20	TNNC1
226779_at	3.280685382	9.71E-23	NA
209815_at	3.281610512	2.66E-21	PTCH1
228915_at	3.285719453	1.55E-23	DACH1
231583_at	3.290446809	4.08E-27	KRT74
231382_at	3.291539204	4.79E-24	FGF18
206114_at	3.292026783	5.59E-18	EPHA4
222484_s_at	3.29309274	2.32E-28	CXCL14
226400_at	3.297416583	1.23E-29	CDC42
208490_x_at	3.302425158	1.54E-22	HIST1H2BF
214343_s_at	3.303704979	1.50E-23	ATXN7L1
1558103_a_at	3.304902938	1.27E-26	NA
212811_x_at	3.305983632	1.97E-23	SLC1A4
214051_at	3.312024364	7.00E-24	TMSB15B
205591_at	3.313182374	2.55E-22	OLFM1
215071_s_at	3.31669792	3.19E-20	HIST1H2AC
1562019_at	3.321844815	9.45E-26	NT5DC4
218137_s_at	3.32409875	1.09E-26	SMAP1
204724_s_at	3.328063372	2.06E-22	COL9A3
226235_at	3.328312862	5.92E-27	LOC339290

204044_at	3.331959558	3.91E-24	QPRT
236193_at	3.332513301	1.50E-19	HIST1H2BC
209806_at	3.343591279	1.58E-25	HIST1H2BK
208257_x_at	3.34607547	1.01E-24	PSG1
219961_s_at	3.354413029	2.87E-29	NCRNA00153
237186_at	3.356912568	6.01E-26	NA
205738_s_at	3.362927631	6.21E-21	FABP3
209269_s_at	3.365373432	3.08E-24	SYK
219848_s_at	3.368954135	2.70E-27	ZNF432
224833_at	3.370488127	2.52E-18	ETS1
225728_at	3.379041546	1.64E-26	SORBS2
203961_at	3.379333563	1.24E-17	NEBL
200854_at	3.379639302	1.86E-26	NCOR1
202208_s_at	3.390133072	9.55E-27	ARL4C
209911_x_at	3.390447922	9.77E-25	HIST1H2BD
227051_at	3.393997025	1.21E-24	NA
219557_s_at	3.394651404	4.89E-24	NRIP3
228256_s_at	3.395902027	8.56E-19	EPB41L4A
240277_at	3.397874068	5.77E-25	NA
229125_at	3.404690429	1.24E-24	KANK4
230518_at	3.407325711	1.56E-25	MPZL2
201147_s_at	3.408915494	4.18E-22	TIMP3
218002_s_at	3.413670852	4.90E-26	CXCL14
236180_at	3.413725375	2.04E-24	NA
209994_s_at	3.417156665	5.50E-24	ABCB1
208180_s_at	3.418227421	1.66E-19	HIST1H4H
206115_at	3.419306335	5.05E-20	EGR3
204872_at	3.426118643	1.26E-26	TLE4
203946_s_at	3.429296488	1.06E-23	ARG2
242722_at	3.43132621	1.27E-26	LMO7
221609_s_at	3.431766211	1.10E-25	WNT6
218280_x_at	3.43358477	1.91E-21	HIST2H2AA3
1553185_at	3.436712669	1.94E-26	RASEF
205234_at	3.438681084	6.08E-28	SLC16A4
229400_at	3.442130704	1.66E-24	HOXD10
214175_x_at	3.442659003	2.87E-29	PDLIM4
211084_x_at	3.447692676	7.57E-20	PRKD3
203423_at	3.451237788	1.41E-20	RBP1
204973_at	3.462579662	5.20E-26	GJB1
214639_s_at	3.464284098	3.96E-23	HOXA1

226279_at	3.465148945	4.47E-26	PRSS23
1553186_x_at	3.468866127	8.93E-28	RASEF
219634_at	3.472907883	3.71E-28	CHST11
224498_x_at	3.474353316	4.05E-24	AXIN2
209032_s_at	3.483990832	2.30E-25	CADM1
201162_at	3.486300954	3.21E-29	IGFBP7
224724_at	3.486464188	6.31E-28	SULF2
229534_at	3.487454536	1.20E-28	ACOT4
243018_at	3.49147799	2.47E-24	NA
1559584_a_at	3.493709322	6.23E-28	C16orf54
229623_at	3.499105535	2.87E-29	FLJ12993
212098_at	3.499524574	1.75E-31	LOC151162
209590_at	3.504695514	2.55E-23	BMP7
1555673_at	3.513525839	2.13E-16	KRTAP2-1
222565_s_at	3.514912198	3.25E-18	PRKD3
213131_at	3.516221723	7.26E-20	OLFM1
206022_at	3.518190011	1.95E-26	NDP
238695_s_at	3.521733381	6.95E-23	RAB39B
223949_at	3.53259321	3.39E-27	TMPRSS3
1558801_at	3.533063313	9.99E-29	NA
209610_s_at	3.533881544	2.21E-24	SLC1A4
210004_at	3.549781842	1.19E-21	OLR1
235056_at	3.551035745	3.53E-24	ETV6
227819_at	3.555562288	1.67E-24	LGR6
240633_at	3.555740535	9.45E-27	DOK7
222921_s_at	3.567858989	2.32E-26	HEY2
238183_at	3.56870017	3.75E-24	NA
212233_at	3.580045163	3.47E-21	MAP1B
1553171_x_at	3.583225164	1.05E-29	LRRN4
220122_at	3.587221935	6.69E-21	MCTP1
214290_s_at	3.593323297	2.62E-22	HIST2H2AA3
208608_s_at	3.593706296	1.16E-25	SNTB1
242844_at	3.60342123	8.37E-29	PGGT1B
1559283_a_at	3.60354118	1.55E-27	CNPY1
214091_s_at	3.604606747	3.33E-27	GPX3
228964_at	3.606325636	2.14E-22	PRDM1
226372_at	3.615380692	1.81E-27	CHST11
1557094_at	3.618353821	1.02E-24	LOC728449
222747_s_at	3.62329679	9.18E-22	SCML1
230372_at	3.624822182	3.35E-27	HAS2

213069_at	3.626213628	1.24E-25	HEG1
238969_at	3.629876991	1.19E-27	C3orf55
228948_at	3.631920862	7.17E-18	EPHA4
227850_x_at	3.637266729	1.18E-26	CDC42EP5
220994_s_at	3.639351184	3.26E-21	STXBP6
235649_at	3.64769918	4.33E-24	ADAMTS8
48106_at	3.656384257	2.87E-29	SLC48A1
225841_at	3.65684861	9.99E-29	C1orf59
235289_at	3.657854186	7.15E-22	EIF5A2
237313_at	3.662435127	1.21E-27	NA
212338_at	3.664925421	3.31E-25	MYO1D
233819_s_at	3.669702757	5.04E-27	RNF160
228865_at	3.669830923	5.52E-19	C1orf116
204830_x_at	3.677088958	5.97E-26	PSG5
238096_at	3.679980581	4.42E-26	LOC284023
218416_s_at	3.68025929	6.90E-27	SLC48A1
205646_s_at	3.680578535	6.28E-26	PAX6
230748_at	3.692286047	4.72E-12	SLC16A6
214696_at	3.699585888	7.46E-25	C17orf91
220494_s_at	3.708426655	7.21E-30	NA
206432_at	3.709321911	2.19E-25	HAS2
225872_at	3.71195561	1.41E-28	SLC35F5
228904_at	3.713381355	7.66E-30	HOXB3
236235_at	3.718283344	1.83E-28	ITCH
228345_at	3.726095245	4.01E-27	CHIC1
1555355_a_at	3.729906852	5.07E-22	ETS1
226085_at	3.738869522	1.80E-28	CBX5
221214_s_at	3.743771322	4.89E-29	NELF
232203_at	3.745645043	9.29E-32	NA
202146_at	3.750056622	1.10E-26	IFRD1
230075_at	3.755303862	7.34E-25	RAB39B
208546_x_at	3.759832814	9.32E-22	HIST1H2BH
205632_s_at	3.761877371	2.25E-28	PIP5K1B
201150_s_at	3.762011986	5.12E-23	TIMP3
205201_at	3.772931511	9.23E-27	GLI3
208579_x_at	3.780347319	4.57E-22	H2BFS
205531_s_at	3.79027671	3.22E-27	GLS2
206002_at	3.792491392	4.65E-27	GPR64
207540_s_at	3.799417725	2.05E-23	SYK
220102_at	3.805958165	3.23E-26	FOXL2
218952_at	3.814312485	3.19E-30	PCSK1N

207060_at	3.823088997	9.94E-28	EN2
220394_at	3.823936895	1.39E-27	FGF20
214660_at	3.824036967	2.61E-26	ITGA1
204790_at	3.826606855	2.00E-27	SMAD7
202147_s_at	3.828678788	1.32E-27	IFRD1
239647_at	3.838528607	2.70E-27	CHST13
219778_at	3.844163234	1.36E-26	ZFPM2
204584_at	3.849827547	8.78E-27	L1CAM
240211_at	3.851610171	3.39E-29	LOC100130468
201249_at	3.864588999	1.09E-26	SLC2A1
218793_s_at	3.869769929	3.45E-14	SCML1
243179_at	3.870345274	1.08E-27	LOC100130360
238937_at	3.88472757	3.24E-26	ZNF420
218330_s_at	3.885973773	2.99E-21	NAV2
236083_at	3.897142097	6.89E-25	BCL2L15
1553972_a_at	3.900856051	2.03E-23	CBS
232035_at	3.907804241	1.13E-20	HIST4H4
204469_at	3.91159933	5.31E-26	PTPRZ1
208596_s_at	3.913039415	1.02E-24	UGT1A3
226731_at	3.926404968	6.57E-28	PELO
209493_at	3.931165422	1.09E-26	PDZD2
236420_s_at	3.954543082	3.21E-29	ANO4
210479_s_at	3.967363693	5.75E-28	RORA
225978_at	3.96824882	2.85E-29	RIMKLB
201008_s_at	3.971573101	7.15E-31	TXNIP
202627_s_at	3.973616853	4.46E-23	SERPINE1
213620_s_at	3.975647758	1.10E-25	ICAM2
227565_at	3.975884347	2.98E-23	NA
1564706_s_at	3.980337827	6.41E-27	GLS2
222695_s_at	3.989645196	3.80E-26	AXIN2
218417_s_at	4.004491731	9.46E-29	SLC48A1
212816_s_at	4.010155616	3.38E-23	CBS
228600_x_at	4.027358429	9.54E-25	C7orf46
209201_x_at	4.028777862	7.41E-25	CXCR4
208607_s_at	4.034989312	9.99E-29	SAA2
224851_at	4.035295105	3.92E-20	CDK6
222925_at	4.03948058	8.30E-32	DCDC2
232297_at	4.039491621	4.65E-24	NA
226068_at	4.041592156	4.09E-21	SYK
213048_s_at	4.041706887	3.33E-29	NA

44790_s_at	4.042203229	1.50E-25	C13orf18
227889_at	4.046036821	7.95E-28	LPCAT2
229912_at	4.075280266	1.33E-26	SDK1
226875_at	4.078960786	9.35E-27	DOCK11
203381_s_at	4.097717865	3.23E-26	APOE
242761_s_at	4.100822533	2.58E-25	ZNF420
210140_at	4.103643002	6.80E-30	CST7
211002_s_at	4.11660951	3.08E-28	TRIM29
235740_at	4.11663046	2.81E-23	NA
242383_at	4.12562674	7.51E-25	NA
1558102_at	4.130863136	5.43E-28	NA
219476_at	4.144578882	5.94E-22	C1orf116
218236_s_at	4.150359247	2.32E-15	PRKD3
211919_s_at	4.163901543	7.34E-25	CXCR4
227647_at	4.164212433	2.89E-25	KCNE3
229613_at	4.165578886	1.21E-30	NA
204891_s_at	4.181607846	3.89E-29	LCK
236917_at	4.181801964	7.60E-30	LRRC34
210426_x_at	4.205522322	4.53E-30	RORA
216603_at	4.210926725	3.24E-28	SLC7A8
206224_at	4.213573867	1.03E-24	CST1
243495_s_at	4.216301839	3.78E-27	NA
208481_at	4.216352669	2.92E-29	ASB4
227449_at	4.217585021	2.97E-13	EPHA4
1554026_a_at	4.236308926	9.46E-25	MYO10
235795_at	4.248285776	3.31E-24	PAX6
213479_at	4.252895427	1.08E-27	NPTX2
218803_at	4.260248536	7.66E-30	CHFR
202982_s_at	4.265488808	7.10E-29	ACOT2
207038_at	4.27226843	2.34E-15	SLC16A6
214455_at	4.273043763	5.91E-21	HIST1H2BC
201250_s_at	4.282654706	2.16E-24	SLC2A1
226084_at	4.283899123	1.96E-23	MAP1B
218454_at	4.285978428	1.30E-27	PLBD1
214995_s_at	4.287614379	1.42E-30	APOBEC3F
240572_s_at	4.288868869	2.99E-27	LOC374443
228920_at	4.299633496	2.07E-29	ZNF260
227099_s_at	4.300338886	7.49E-28	LOC387763
219117_s_at	4.301041604	1.54E-29	FKBP11
236471_at	4.308036167	7.46E-24	NFE2L3
203815_at	4.310438056	7.98E-31	GSTT1

217228_s_at	4.315831944	1.11E-28	ASB4
219118_at	4.320368737	1.67E-27	FKBP11
235457_at	4.330850064	4.13E-30	MAML2
213285_at	4.334974113	4.48E-27	TMEM30B
231849_at	4.341804272	1.60E-30	KRT80
211071_s_at	4.348007242	5.99E-25	MLLT11
202783_at	4.363360503	2.22E-32	NNT
221648_s_at	4.395444773	7.37E-29	NA
204042_at	4.406209156	6.76E-25	WASF3
205003_at	4.411778434	1.81E-19	DOCK4
228266_s_at	4.418838091	2.73E-29	HDGFRP3
215125_s_at	4.423254251	1.19E-27	UGT1A@
202207_at	4.430268677	6.79E-26	ARL4C
212143_s_at	4.4325135	9.68E-31	IGFBP3
203508_at	4.45742527	8.66E-28	TNFRSF1B
212776_s_at	4.462552662	3.47E-28	OBSL1
222771_s_at	4.466520226	1.28E-29	MYEF2
206029_at	4.488030927	3.18E-28	ANKRD1
229215_at	4.495802842	1.23E-30	ASCL2
236979_at	4.506821068	6.37E-31	BCL2L15
205110_s_at	4.514206803	3.13E-27	FGF13
219109_at	4.51482857	3.96E-30	SPAG16
1558388_a_at	4.519733027	9.21E-25	NA
222449_at	4.524503386	1.64E-30	PMEPA1
203559_s_at	4.524531879	7.53E-30	ABP1
227753_at	4.532066865	1.50E-27	TMEM139
222150_s_at	4.5404784	2.07E-30	PION
221666_s_at	4.544133022	1.88E-31	PYCARD
222513_s_at	4.572953768	1.11E-28	SORBS1
226829_at	4.579320075	4.97E-29	AFAP1L2
214571_at	4.581354757	8.54E-30	FGF3
213789_at	4.584017391	3.35E-28	EBP
205858_at	4.608291805	2.45E-26	NGFR
204523_at	4.612446101	3.79E-32	ZNF140
209591_s_at	4.628920548	1.66E-25	BMP7
237720_at	4.629160973	1.69E-25	ASB4
220048_at	4.632384526	9.52E-30	EDAR
209030_s_at	4.633525554	1.96E-28	CADM1
1569342_at	4.654604987	1.19E-31	GLI3
243489_at	4.662489597	2.54E-25	NA
221024_s_at	4.68438753	4.03E-30	SLC2A10

213142_x_at	4.686111868	1.13E-30	PION
214456_x_at	4.703079815	4.23E-29	SAA2
239791_at	4.709513144	1.07E-29	LOC404266
211840_s_at	4.711666435	4.73E-29	PDE4D
239153_at	4.720703792	4.89E-29	HOTAIR
219563_at	4.726677676	1.36E-24	C14orf139
209822_s_at	4.729414324	1.10E-30	VLDLR
203799_at	4.730316936	1.86E-32	CD302
202708_s_at	4.730427145	1.75E-25	HIST2H2BE
227812_at	4.738529483	1.33E-28	TNFRSF19
227279_at	4.746540064	1.87E-31	TCEAL3
201010_s_at	4.759053982	5.51E-30	TXNIP
1553602_at	4.762446302	1.53E-21	MUCL1
209040_s_at	4.768668189	5.51E-30	PSMB8
223434_at	4.770714033	1.50E-29	GBP3
209101_at	4.777273007	9.79E-30	CTGF
203030_s_at	4.78142293	4.93E-28	PTPRN2
213164_at	4.784594136	4.26E-26	SLC5A3
221796_at	4.785325049	3.61E-31	NTRK2
211796_s_at	4.787469259	1.56E-28	IL23A
201348_at	4.802407705	1.98E-29	GPX3
229095_s_at	4.809985359	1.01E-28	LIMS3
229634_at	4.810504894	1.08E-30	TMEM139
201825_s_at	4.820029158	1.18E-29	SCCPDH
229963_at	4.828070892	2.58E-30	BEX5
202206_at	4.833388086	1.83E-28	ARL4C
1553681_a_at	4.83503507	3.05E-30	PRF1
1552626_a_at	4.843176559	6.32E-29	TMEM163
202784_s_at	4.85124007	4.02E-30	NNT
230560_at	4.856434044	3.43E-25	STXBP6
235182_at	4.866079067	2.32E-26	ISM1
211966_at	4.879895185	9.03E-31	COL4A2
232676_x_at	4.883374704	1.56E-33	MYEF2
231940_at	4.893912693	1.75E-31	ZNF529
201508_at	4.90618738	3.14E-28	IGFBP4
228262_at	4.910615305	1.86E-28	MAP7D2
209598_at	4.918544008	2.29E-31	PNMA2
211549_s_at	4.938244818	3.08E-30	HPGD
202283_at	4.943811341	6.40E-31	SERPINF1
240687_at	4.951694879	2.11E-31	PASD1
224176_s_at	4.964145784	1.71E-28	AXIN2

229830_at	4.969258134	4.21E-29	PDGFA
222450_at	4.977795355	5.32E-31	PMEPA1
227236_at	5.024594887	8.77E-30	TSPAN2
1558700_s_at	5.038946702	3.00E-31	ZNF260
216693_x_at	5.054395533	1.21E-29	HDGFRP3
204045_at	5.054625784	9.95E-33	TCEAL1
209387_s_at	5.060435864	7.03E-32	TM4SF1
231867_at	5.065919831	5.32E-31	ODZ2
214247_s_at	5.076833805	3.89E-29	DKK3
238780_s_at	5.077502968	2.44E-28	NA
204932_at	5.102161875	6.22E-26	TNFRSF11B
228523_at	5.106581599	1.84E-26	NANOS1
242013_at	5.114226909	1.55E-31	BCL2L15
218963_s_at	5.114492799	8.34E-14	KRT23
203029_s_at	5.12017628	2.12E-26	PTPRN2
202286_s_at	5.137666	3.47E-17	TACSTD2
222803_at	5.152802825	1.99E-31	PRTFDC1
209526_s_at	5.15724731	7.00E-33	HDGFRP3
222549_at	5.167696239	1.70E-33	CLDN1
240055_at	5.174996434	3.92E-31	NA
206560_s_at	5.178857464	2.57E-33	MIA
218723_s_at	5.196995571	1.44E-27	C13orf15
202628_s_at	5.213883908	2.83E-25	SERPINE1
221558_s_at	5.248422657	2.73E-29	LEF1
237721_s_at	5.275233137	3.34E-29	ASB4
231227_at	5.278884299	9.59E-35	NA
204933_s_at	5.302809105	1.66E-29	TNFRSF11B
226682_at	5.31364024	5.80E-33	RORA
201506_at	5.320158709	1.31E-30	TGFBI
211429_s_at	5.323447734	4.90E-27	SERPINA1
236918_s_at	5.339461334	2.16E-32	LRRC34
230795_at	5.340692109	8.14E-27	NA
209213_at	5.428561762	6.04E-33	CBR1
203798_s_at	5.457496327	1.43E-31	VSNL1
235940_at	5.463010338	9.49E-34	C9orf64
203304_at	5.483975709	2.16E-22	BAMBI
203797_at	5.509456097	3.95E-31	VSNL1
219836_at	5.514210182	1.11E-29	ZBED2
235845_at	5.52188532	3.92E-31	SP5
218087_s_at	5.527335081	2.82E-28	SORBS1

226226_at	5.540056241	1.72E-32	TMEM45B
203914_x_at	5.58234348	9.02E-31	HPGD
223503_at	5.605728792	2.31E-30	TMEM163
222696_at	5.615989091	1.36E-30	AXIN2
210095_s_at	5.619119696	3.65E-31	IGFBP3
209031_at	5.62842311	9.01E-32	CADM1
214617_at	5.6378784	4.20E-29	PRF1
210836_x_at	5.702551823	2.41E-32	PDE4D
202504_at	5.717674814	5.82E-32	TRIM29
212775_at	5.732467036	1.61E-34	OBSL1
209386_at	5.755728019	2.29E-31	TM4SF1
210837_s_at	5.771365504	9.54E-32	PDE4D
211548_s_at	5.797156674	1.42E-30	HPGD
202149_at	5.83374799	6.28E-31	NEDD9
203474_at	5.838363348	4.93E-30	IQGAP2
205399_at	5.885915621	9.99E-29	DCLK1
211564_s_at	5.891723978	9.54E-32	PDLIM4
221011_s_at	5.901510021	3.08E-30	LBH
202833_s_at	5.906131804	8.46E-27	SERPINA1
211003_x_at	5.906561411	2.07E-33	TGM2
229580_at	5.970583176	5.37E-34	NA
215034_s_at	6.072809343	7.03E-32	TM4SF1
211573_x_at	6.081013458	5.83E-34	TGM2
237974_at	6.093435795	4.30E-35	ABHD12B
213425_at	6.10880796	7.00E-33	WNT5A
211964_at	6.126274623	6.51E-32	COL4A2
209655_s_at	6.202016916	3.69E-33	TMEM47
217028_at	6.218488412	1.66E-25	CXCR4
209524_at	6.247652627	8.35E-34	HDGFRP3
205990_s_at	6.30708769	1.81E-34	WNT5A
204205_at	6.331443474	6.93E-30	APOBEC3G
203913_s_at	6.452092708	1.34E-34	HPGD
232231_at	6.49125603	7.28E-16	RUNX2
235619_at	6.531240266	5.60E-35	LOC285986
229800_at	6.665510164	1.81E-35	DCLK1
228962_at	6.692436894	4.08E-33	PDE4D
204115_at	6.719502089	1.61E-34	GNG11
230323_s_at	6.831382501	9.04E-35	TMEM45B
207935_s_at	6.88218859	1.96E-34	KRT13
201042_at	6.970819459	7.13E-35	TGM2
236892_s_at	7.016783909	5.91E-36	LOC404266

201426_s_at	7.036826527	5.91E-30	VIM
203413_at	7.099925614	6.04E-33	NELL2
227376_at	7.130645468	6.10E-35	GLI3
204491_at	7.24870063	8.28E-37	PDE4D
214023_x_at	7.270976799	3.97E-36	TUBB2B
201163_s_at	7.309927506	5.60E-35	IGFBP7
216236_s_at	7.387604268	9.53E-28	SLC2A14
228335_at	7.391273205	7.80E-38	CLDN11
205174_s_at	7.717056894	8.70E-37	QPCT
202497_x_at	7.75240915	3.31E-26	SLC2A3
201820_at	7.777504885	7.14E-36	KRT5
200799_at	7.801303143	1.24E-33	HSPA1A
223800_s_at	7.804629034	1.05E-32	LIMS3
202498_s_at	7.843883924	2.75E-27	SLC2A3
222088_s_at	7.847920371	6.59E-27	SLC2A14
227475_at	7.989488515	3.57E-34	FOXQ1
202499_s_at	8.332857437	3.30E-24	SLC2A3
209656_s_at	8.729730884	5.08E-35	TMEM47
225016_at	9.310729468	2.09E-36	APCDD1

Global gene expression analysis of $SW480^{vector}$ and $SW480^{Smad4}$ colon cancer cells by microarray was performed to assess changes in gene expression caused by restored Smad4 expression. We used the following stringent criteria to determine differential expression of $SW480^{vector}$ compared with $SW480^{Smad4}$ colon cancer cells: FDR<0.005 and fold-change >4 (see Materials and Methods) to obtain an epithelial specific, Smad4 expression profile. Up-regulated and down-regulated Affymetrix® probesets are displayed along with the $\log_2$ coefficient, FDR and gene symbol. A negative coefficient in this analysis indicates up-regulation in the presence of Smad4 and a positive coefficient indicates down-regulation in the presence of Smad4.

TABLE 12

WNT TARGET LIST GENE IDENTIFIERS

Gene ID	Gene	Gene Symbol	Cells or tissue validation	Reference 1: PMID	Reference 2: PMID
4609	c-myc	MYC	human colon cancer	He 1998: 9727977	
4613	n-myc	MYCN	mesenchyme limbs	Ten Berge 2008: 18776145	
595	Cyclin D	CCND1	human colon cancer	Tetsu 1999: 10201372	Shtutman 1999: 10318916
6932	Tcf-1	TCF7	human colon cancer	Roose 1999: 10489374	
51176	LEF1	LEF1	human colon cancer	Hovanes, 2001: 11326276	Filali 2002: 12052822
5467	PPARdelta	PPARD	human colon cancer	He TC,1999: 10555149	
3725	c-jun	JUN	human colon cancer	Mann B, 1999: 9990071	
8061	fra-1	FOSL1	human colon cancer	Mann B, 1999: 9990071	
5329	uPAR	PLAUR	human colon cancer	Mann B, 1999: 9990071	
4316	matrix metalloproteinase MMP-7	MMP7	human colon cancer	Brabletz 1999: 10514384	Crawford 1999: 10362259
8313	Axin-2	AXIN2	human colon cancer	Yan, 2001: 11752446	Lustig, 2002: 11809809
4897	Nr-CAM	NRCAM	human colon cancer	Conacci-Sorrell 2002: 12183361	
6925	ITF-2	TCF4	human colon cancer	Kolligs, 2002: 12086873	
2520	Gastrin	GAST	human colon cancer	Koh, 2000: 10953028	
960	CD44	CD44	human colon cancer	Wielenga 1999: 10027409	
1948	EphB/ephrin-B	EFNB2	human colon cancer	Batlle, 2002: 12408869	
1947	EphB/ephrin-B	EFNB1	human colon cancer	Batlle, 2002: 12408869	
1949	EphB/ephrin-B	EFNB3	human colon cancer	Batlle, 2002: 12408869	
652	BMP4	BMP4	human colon cancer	Kim 2002: 12019147	
9076	claudin-1	CLDN1	human colon cancer	Miwa 2002: 11939410	
332	Survivin	BIRC5	human colon cancer	Zhang, 2001: 11751382	
7422	VEGF	VEGFA	human colon cancer	Zhang, 2001: 11507052	
8817	FGF18	FGF18	human colon cancer	Shimokawa 2003: 14559787	
474	Hath1	ATOH1	human colon cancer	Leow 2004: 15342386	
4233	Met	MET	human colon cancer	Boon 2002: 12234972	
1906	endothelin-1	EDN1	human colon cancer	Kim 2005: 15558022	
26292	c-myc binding protein	MYCBP	human colon cancer	Jung 2005: 15979100	
3897	L1 neural adhesion	L1CAM	human colon cancer	Gavert 2005: 15716380	
3398	Id2	ID2	human colon cancer	Rockman 2001: 11572874	Willert 2002: 12095419 (see this paper for others)
182	Jagged	JAG1	human colon cancer	Rodilla, 2009: 19325125	
7074	Tiam1	TIAM1	Colon tumors	Malliri 2005: 16249175	
4843	Nitric Oxide Synthase 2	NOS2	Hepg2, Hct116 and dld-1 cells	Du, 2006	
22943	Dickkopf	DKK1	Various cells, tumors	Niida 2004: 15378020	Gonzalez-Sancho 2004: 15592505
2254	FGF9	FGF9	ovarian endometrioid adenocarcinoma	Hendrix, 2006: 16452189	
26281	FGF20	FGF20	Various cells, tumors	Chamorro 2004: 15592430	

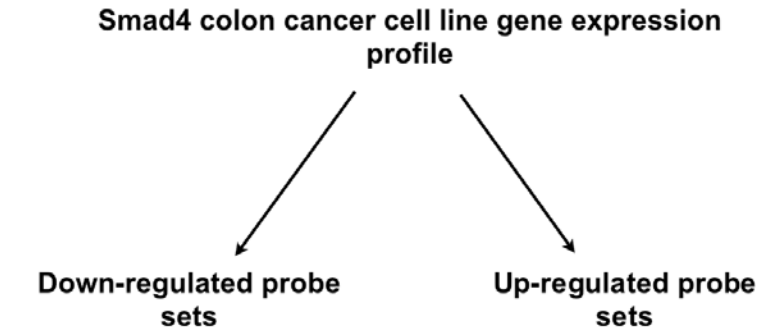
8549	LGR5/GPR49	LGR5	Intestine	Barker, 2007: 17934449	
6662	Sox9	SOX9	Intestine	Blache 2004: 15240568	
6662	Sox9	SOX9	mesenchyme	Hill, 2005: 15866163	Day 2005: 15866164
860	Runx2	RUNX2	chondrocytes	Dong 2006	Hill, 2005: 15866163
64388	Gremlin	GREM2	fibroblasts	Klapholz-Brown 2007: 17895986	
57167	SALL4	SALL4		Bohm, 2006: 16899215	
8792	RANK ligand	TNFRSF11A	Osteoblasts	Spencer 2006	
3491	CCN1/Cyr61	CYR61	Osteoblasts	Si, 2006: 16581771	
6657	Sox2	SOX2	Xenopus retina	Van Raay, 2005: 15820691	
9232	Pituitary tumor transforming gene (PTTG)	PTTG1	esophageal squamous cell carcinoma	Zhou 2004: 15514942	
28514	Delta-like 1	DLL1	somites	Galceran, 2004: 15545629	Hofmann 2004: 15545628
8456	FoxN1	FOXN1	thymus	Balciunaite 2002: 12379851	
56547	matrix metalloproteinase-26	MMP26	Human	Marchenko 2002: 11931652	
79923	nanog	NANOG	ES	Pereira, 2006: 16894029	Cole 2008: 18347094
5460	POU5F1 (aka oct4)	POU5F1	ES	Cole 2008: 18347094	
6615	snail	SNAI1	ES/EB	Ten Berge 2008: 18983966	
2335	Fibronectin	FN1	ES/EB	Ten Berge 2008: 18983966	
8324	Frizzled 7	FZD7	EC cells	Willert 2002: 12095419	
10468	Follistatin	FST	EC cells, ovary	Willert 2002: 12095419	Yao 2004: 15162500
89780	Wnt3a	WNT3A	EC cells	Zhang 2009: 19109969	
2335	Fibronectin	FN1	Mouse lung	De Langhe 2005: 15617677	
3670	Islet1	ISL1	Cardiac cells	Lin 2007: 17519333	
4313	MMP2	MMP2	T cells	Wu 2007: 17306568	
4318	MMP9	MMP9	T cells	Wu 2007: 17306568	
2335	fibronectin	FN1	Xenopus	Gradl 1999: 10409747	
652	BMP4	BMP4	Xenopus	Baker 1999: 10601040	
n/a	myogenic bHLH	MYOD1	Xenopus	Munsterberg 1995: 7498788	
2020	engrailed-2	EN2	Xenopus	McGrew 1999: 10495268	
2697	connexin43	GJA1	Xenopus, Mouse	van der Heyden 1998: 9601103	
10804	connexin 30	GJB6	Xenopus	McGrew 1999: 10495268	
5916	retinoic acid receptor gamma	RARG	Xenopus	McGrew 1999: 10495268	
4286	MITF/nacre	MITF	Zebrafish	Dorsky, 2000: 10652270	Saito 2002: 12048204
64220	Stra6	STRA6	Wnt-1 transformed mouse cells	Szeto 2001: 11358845	
58480	Wrch-1	RHOU	Wnt-1 transformed mouse cells	Tao, 2001: 11459829	
3604	TNF family 41BB ligand	TNFRSF9	Wnt-1 transformed mouse cells	Tice 2002: 11832495	
1947	ephrinB1	EFNB1	Wnt-1 transformed mouse cells	Tice 2002: 11832495	
64220	Stra6	STRA6	Wnt-1 transformed mouse cells	Tice 2002: 11832495	
5168	autotaxin	ENPP2	Wnt-1 transformed mouse cells	Tice 2002: 11832495	
3671	ISLR	ISLR	Wnt-1 transformed mouse cells	Tice 2002: 11832495	
7291	Twist	TWIST1	Wnt1 induced mammary cancer	Howe, 2003: 12702582	

4314	Stromelysin	MMP3	Wnt-1 transformed mouse cells	Prieve, 2003: 12697065	
51429	WISP	SNX9	Wnt-1 transformed mouse cells	Xu, 2000: 10716946	
2641	Proglucagon	GCG	Mouse	Ni 2003: 12421827	
632	Osteocalcin	BGLAP	Mouse	Kahler 2003: 12551949	
1044	Cdx1	CDX1	Mouse embryo	Pilon 2007: 17537796	
5743	cyclooxygenase-2	PTGS2	mouse (Wnt-1)	Howe 1999: 10197631	Haertel-Wiesmann 2000: 10884377
79191	Irx3	IRX3	Mouse brain	Braun 2003: 14522868	
6496	Six3	SIX3	Mouse brain	Braun 2003: 14522868	
4762	neurogenin 1	NEUROG1	Mouse brain	Hirabayashi 2004: 15142975	
389058	SP5	SP5	Mouse brain	Weidinger 2005: 15797017	Fujimura 2007: 17090534
4821	Nkx2.2	NKX2-2	Neural tube	Lei, 2006: 16950124	
3569	IL-6	IL6	3T3-L1 Preadipocytes	Longo, 2002: 12154096	
420322	WISP-1	WISP1	3T3-L1 Preadipocytes	Longo, 2002: 12154096	
8839	WISP-2	WISP2	3T3-L1 Preadipocytes	Longo, 2002: 12154096	
3481	IGF-II	IGF2	3T3-L1 Preadipocytes	Longo, 2002: 12154096	
10296	Emp	MAEA	3T3-L1 Preadipocytes	Longo, 2002: 12154096	
3479	IGF-I	IGF1	3T3-L1 Preadipocytes	Longo, 2002: 12154096	
7424	VEGF-C	VEGFC	3T3-L1 Preadipocytes	Longo, 2002: 12154096	
5243	MDR1	ABCB1	3T3-L1 Preadipocytes	Longo, 2002: 12154096	
10631	Â periostin	POSTN	Mouse Wnt-3	Haertel-Wiesmann 2000: 10884377	
1044	Cdx1	CDX1	Mouse Wnt-3A	Lickert 2000: 10934025	
1046	Cdx4	CDX4	Mouse Wnt-3A	Pilon, 2006: 16309666	
1046	Cdx4	CDX4	Zebrafish HSC	Davidson 2003: 13679919?	
8945	betaTRCP	BTRC	293T and HeLa cells	Spiegelman 2000: 10882123	
6423	sFRP-2	SFRP2	Mouse (Wnt-4)	Lescher 1998: 9853965	
5308	Pitx2	PITX2	pituitary	Kioussi 2002: 12464179	
1956	EGF receptor	EGFR	Liver	Tan 2005: 16012954	
1896	Eda (TNF-related)	EDA	Mouse hair follicle	Laurikkala 2002: 11973284	Durmowicz 2002: 12039047
999	E-cadherin	CDH1	Mouse hair follicle	Jamora, 2003: 12646922	
999	E-cadherin	CDH1	ES/EB	Ten Berge 2008: 18983966	
3852	Keratin	KRT5	Mouse hair follicle	Dasgupta 1999: 10498690	
182	Jagged1	JAG1	Mouse hair follicle	Estrach, 2006: 17035290	
1029	P16ink4A	CDKN2A	Melanocytes	Delmas, 2007: 18006687	
1493	CTLA-4	CTLA4	Melanomas	Shah 2008: 18563180	
688	mBTEB2	KLF5	Mouse	Ziemer 2001: 11134343	
2249	FGF4	FGF4	Mouse tooth bud	Kratochwil 2002: 12502739	
3576	Interleukin8	IL8	Endothelial cells	Masckauchan 2005: 16132617	
5979	ret	RET	rat PC12	Zheng, 1996: 8637712	
2697	connexin43	GJA1	rat cardiomyocytes	Ai 2000: 10642594	
1462	versican	VCAN	vascular smooth muscle cells	Rahmani 2005: 15668231	
55504	Tnfrsf19	TNFRSF19	Somitic mesoderm	Buttitta 2003: 12781685	
51035	Ubx	UBXN1	Drosophila	Riese 1997: 9118221	
127733	Ubx	UBXN10	Drosophila	Riese 1997: 9118221	
7993	Ubx	UBXN8	Drosophila	Riese 1997: 9118221	

7471	wingless	WNT1	Drosophila	Yu 1998: 9835654	
655	BMP7	BMP7	human colon cancer cells	Hatzis MCB 2008:	Beites CL: 19474151 (weaker indirect evidence)
652	Dpp	BMP4	Drosophila	Yang, 2000: 10934014	
2019	Engrailed 1	EN1	human epithelial cells	Bachar-Dahan, 2006: 16571670	
8325	Dfrizzled2	FZD8	Drosophila	Cadigan 1998(Dfz2 human homolog=fz8): 9630221	
3219	HOXB9 homeobox B9	HOXB9	lung cancer cells	Nguyen DX, 2009: 19576624	
9314	KLF4	KLF4	intestinal cells	Evans PM, 2009: 19901072	

A table of annotated Wnt target genes is shown. This list was used for Affymetrix® input to obtain 282 resultant Affymetrix probes (not shown). Gene identifier (ID), gene name and gene symbol are displayed in addition to validated cells/tissues as noted in the published literature. The first author and PubMed Identifier (PMID) are listed for at least two references. Source for the original list:
http://www.stanford.edu/~rnusse/pathways/targets.html

A.



B.

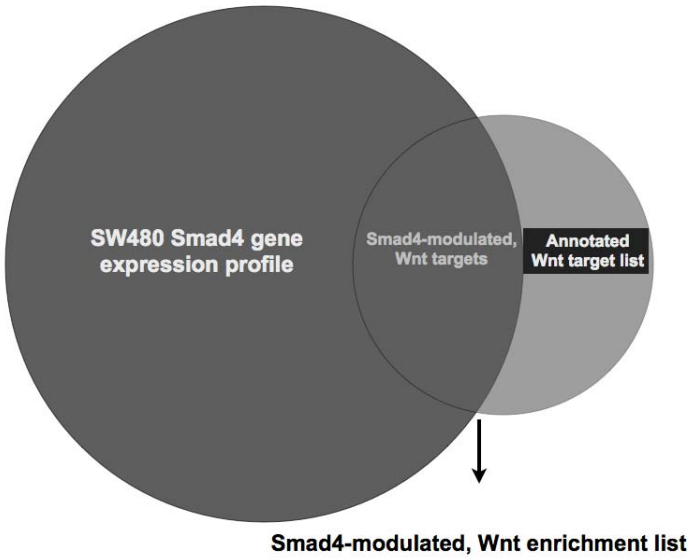


Figure. 22*: (A) Smad4 colon cancer expression profile determination: flow diagram.* The figure shows a flow-diagram of the development of a Smad4 gene expression profile in SW480 colon cancer cells (FDR<0.005 and fold-change>4). 1668 probesets were implicated which consisted of up-regulated (n=593) and down-regulated probesets (n=1075). *(B) Schema for determination of Wnt target gene enrichment in the Smad4 expression profile.* The SW480 Smad4 global expression profile (black circle, n=1668 probesets) was used to determine if there was a significant enrichment of an annotated list of published Wnt target genes (gray circle, n=282 probesets) (see Table 12). The table containing the enriched Smad4-modulated Wnt targets is Table 13 (dark gray overlap, n=150 probe sets).

TABLE 13

SMAD4-MODULATED, WNT ENRICHED TARGETS

Affymetrix probe	Coef	FDR	Gene Symbol
213816_s_at	-5.208338904	5.32E-31	MET
221731_x_at	-4.455090423	5.36E-15	VCAN
203510_at	-4.437460842	1.19E-27	MET
204620_s_at	-4.349783568	2.54E-15	VCAN
213807_x_at	-3.143497429	5.00E-27	MET
215646_s_at	-3.086916374	6.75E-11	VCAN
204602_at	-3.075203281	7.58E-13	DKK1
211599_x_at	-3.014005885	5.87E-27	MET
211571_s_at	-2.710027617	1.41E-10	VCAN
204619_s_at	-2.586155238	1.26E-10	VCAN
207558_s_at	-2.53932263	5.59E-20	PITX2
238846_at	-2.310058123	2.61E-18	TNFRSF11A
213943_at	-2.298439083	3.32E-20	TWIST1
215983_s_at	-1.755190399	3.48E-22	UBXN8
212063_at	-1.669196296	5.85E-19	CD44
204489_s_at	-1.47796131	1.43E-19	CD44
202668_at	-1.320836633	2.00E-14	EFNB2
212014_x_at	-1.218341893	6.15E-17	CD44
202669_s_at	-1.204038673	4.41E-12	EFNB2
224471_s_at	-1.198348701	6.04E-18	BTRC
210512_s_at	-1.191362159	6.80E-15	VEGFA
202095_s_at	-1.131580202	1.75E-17	BIRC5
201130_s_at	-1.119648171	3.52E-13	CDH1
209835_x_at	-1.092351031	3.57E-17	CD44
204490_s_at	-1.0777488	1.35E-16	CD44
210916_s_at	-1.073401268	2.69E-18	CD44
204188_s_at	-1.0290679	6.19E-17	RARG
210513_s_at	-1.018637972	5.13E-15	VEGFA
211527_x_at	-1.003293473	4.59E-15	VEGFA
210334_x_at	-0.983871923	2.44E-15	BIRC5
1557905_s_at	-0.921889157	1.15E-16	CD44
211421_s_at	-0.884091927	1.17E-09	RET
217523_at	-0.883786496	2.54E-07	CD44
212171_x_at	-0.868778847	5.80E-13	VEGFA
201131_s_at	-0.868599892	2.04E-11	CDH1
216091_s_at	-0.857392319	6.76E-14	BTRC

207037_at	-0.819544417	7.66E-09	TNFRSF11A
207922_s_at	-0.769870439	3.69E-16	MAEA
204189_at	-0.682444519	1.54E-11	RARG
202094_at	-0.653995507	1.08E-10	BIRC5
1565483_at	-0.653204532	7.18E-05	EGFR
204901_at	-0.621820525	1.25E-11	BTRC
205879_x_at	-0.589409335	3.89E-10	RET
203361_s_at	-0.550288811	2.98E-11	MYCBP
1565484_x_at	-0.487337782	1.40E-06	EGFR
203359_s_at	-0.480112291	6.11E-10	MYCBP
229221_at	-0.476320032	7.40E-05	CD44
219480_at	-0.462653661	2.48E-05	SNAI1
206634_at	-0.412215743	8.90E-05	SIX3
215771_x_at	-0.397667812	1.19E-05	RET
203360_s_at	-0.331871568	4.56E-08	MYCBP
221331_x_at	-0.316135616	1.34634E-04	CTLA4
220184_at	-0.298162725	2.55565E-04	NANOG
203936_s_at	-0.273585702	1.88495E-04	MMP9
207233_s_at	0.282372198	3.80852E-04	MITF
209644_x_at	0.308520031	3.50E-07	CDKN2A
223028_s_at	0.332644354	3.72E-06	SNX9
209540_at	0.333169751	2.90847E-04	IGF1
203554_x_at	0.363779621	1.07E-06	PTTG1
210636_at	0.383131464	2.09E-06	PPARD
223027_at	0.432717882	8.48E-10	SNX9
211260_at	0.49250662	5.61E-08	BMP7
204748_at	0.523747577	3.26E-05	PTGS2
224325_at	0.581595283	2.83E-07	FZD8
228038_at	0.595555086	6.04E-09	SOX2
214701_s_at	0.601019473	1.18E-07	FN1
221557_s_at	0.610681431	4.69E-07	LEF1
211577_s_at	0.622106274	3.20E-06	IGF1
221282_x_at	0.656272386	5.13E-08	RUNX2
226066_at	0.664734161	3.36E-06	MITF
204420_at	0.675402303	3.18E-07	FOSL1
227405_s_at	0.686961958	2.83E-08	FZD8
213281_at	0.700658487	2.24E-11	JUN
201871_s_at	0.732825908	1.61E-13	UBXN1
208044_s_at	0.739911957	2.13E-08	PPARD
1569334_at	0.748543991	5.49E-11	STRA6
218995_s_at	0.766521861	4.18E-10	EDN1

37152_at	0.842712279	8.64E-12	PPARD
210984_x_at	0.858900718	5.11E-12	EGFR
211607_x_at	0.863360829	9.39E-08	EGFR
209541_at	0.869308094	2.14E-11	IGF1
209542_x_at	0.952747647	1.05E-09	IGF1
1569335_a_at	0.955531689	5.00E-12	STRA6
223168_at	0.963972173	2.87E-06	RHOU
201984_s_at	0.984776338	5.53E-09	EGFR
205255_x_at	1.01704237	1.38E-14	TCF7
205254_x_at	1.026916058	5.09E-14	TCF7
201069_at	1.068712668	1.29E-11	MMP2
232109_at	1.167483103	2.81E-13	UBXN10
222802_at	1.199330545	1.62E-14	EDN1
202936_s_at	1.293576102	4.93E-10	SOX9
216417_x_at	1.298462572	1.13E-18	HOXB9
226461_at	1.349730726	1.71E-19	HOXB9
239178_at	1.439091976	1.24E-14	FGF9
201983_s_at	1.446036871	2.78E-10	EGFR
230092_at	1.459946709	1.42E-15	UBXN10
206104_at	1.51025451	2.37E-12	ISL1
201465_s_at	1.524742241	8.14E-16	JUN
202935_s_at	1.54400712	4.70E-11	SOX9
203705_s_at	1.582594184	1.73E-18	FZD7
208711_s_at	1.588606808	7.87E-24	CCND1
209993_at	1.73480427	3.59E-14	ABCB1
211485_s_at	1.74932923	2.77E-20	FGF18
202711_at	1.834688368	6.23E-19	EFNB1
216994_s_at	1.873677384	7.74E-17	RUNX2
209211_at	1.882266539	9.29E-18	KLF5
224090_s_at	1.906822818	1.12E-20	TNFRSF19
203706_s_at	1.940917236	3.28E-18	FZD7
209212_s_at	1.97330756	7.51E-23	KLF5
201565_s_at	1.982069503	6.20E-10	ID2
208712_at	2.02402084	1.88E-24	CCND1
201566_x_at	2.139151764	2.14E-10	ID2
201464_x_at	2.171310811	1.69E-21	JUN
210948_s_at	2.396963985	2.17E-21	LEF1
213931_at	2.397088384	9.10E-12	ID2
211029_x_at	2.462131571	4.47E-21	FGF18
211518_s_at	2.462756098	5.24E-19	BMP4
201466_s_at	2.56324932	6.42E-19	JUN

210495_x_at	2.592240002	6.08E-21	FN1
223827_at	2.615606039	2.54E-22	TNFRSF19
206987_x_at	2.733241904	8.93E-23	FGF18
218182_s_at	2.734588961	1.43E-23	CLDN1
242218_at	2.759858224	2.16E-22	PPARD
211259_s_at	2.760066972	1.86E-25	BMP7
201289_at	2.770069888	6.99E-22	CYR61
238657_at	2.82400144	4.47E-22	UBXN10
216442_x_at	2.864553299	8.49E-22	FN1
210764_s_at	2.875186002	3.15E-21	CYR61
221701_s_at	2.969423295	1.34E-24	STRA6
212464_s_at	2.974592447	2.25E-21	FN1
205031_at	3.066126601	5.28E-27	EFNB3
211719_x_at	3.128023957	1.55E-22	FN1
206404_at	3.25327442	8.87E-23	FGF9
231382_at	3.291539204	4.79E-24	FGF18
209994_s_at	3.417156665	5.50E-24	ABCB1
224498_x_at	3.474353316	4.05E-24	AXIN2
209590_at	3.504695514	2.55E-23	BMP7
207060_at	3.823088997	9.94E-28	EN2
220394_at	3.823936895	1.39E-27	FGF20
204584_at	3.849827547	8.78E-27	L1CAM
222695_s_at	3.989645196	3.80E-26	AXIN2
209591_s_at	4.628920548	1.66E-25	BMP7
227812_at	4.738529483	1.33E-28	TNFRSF19
224176_s_at	4.964145784	1.71E-28	AXIN2
222549_at	5.167696239	1.70E-33	CLDN1
221558_s_at	5.248422657	2.73E-29	LEF1
235845_at	5.52188532	3.92E-31	SP5
222696_at	5.615989091	1.36E-30	AXIN2
232231_at	6.49125603	7.28E-16	RUNX2
201820_at	7.777504885	7.14E-36	KRT5

Smad4-modulated, Wnt enriched target genes and their Affymetrix probe identifiers (IDs) are displayed. Fold-change on a $\log_2$-scale, False Discovery Rate (FDR) and Gene Symbol are also displayed. A negative coefficient in this analysis indicates up-regulation in the presence of Smad4 and a positive coefficient indicates down-regulation in the presence of Smad4.

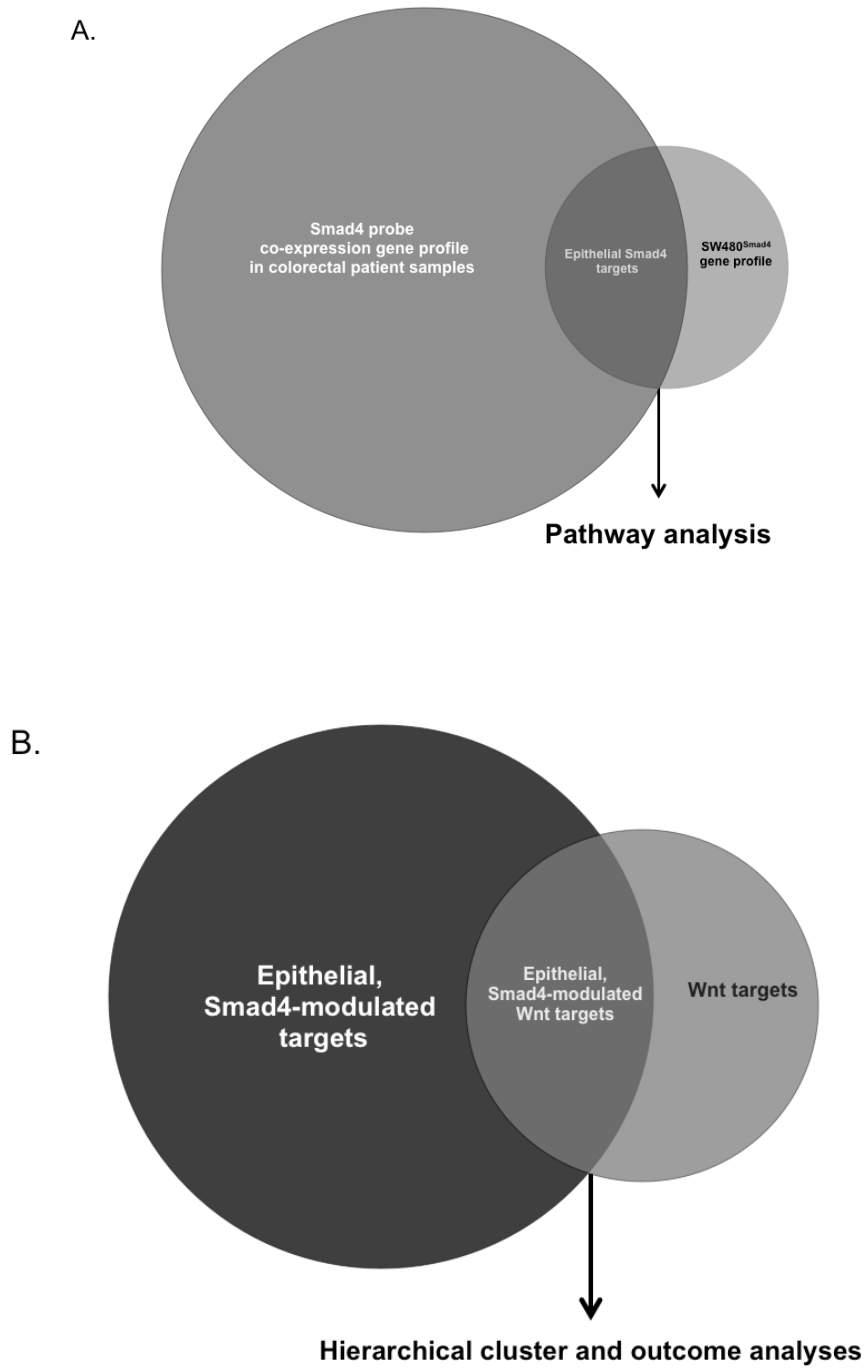


Figure 32. *Derivation of epithelial-specific Smad4 targets and determination of Wnt enrichment.* (A) Schematic depiction of human Smad4-modulated gene expression profiling and pathway analysis. The large black circle represents the human Smad4 gene expression profile (FDR<0.05, n=21889 probesets). The small light gray circle represents the SW480 cell line Smad4 gene expression profile (FDR<0.005, fold-change>4, n=1668 probesets). The darker gray overlap represents epithelial-specific Smad4 co-expression targets (n=787 probesets). (B) Schematic for determination of Smad4-modulated, Wnt-enriched epithelial-specific targets for hierarchical and outcome analyses is displayed. The dark gray circle represents epithelial-specific Smad4 co-regulated targets (n=787 probesets) and the light gray circle represents annotated Wnt targets (n=285 probesets). The overlap represents epithelial-specific Smad4-modulated Wnt target genes (n=32 probesets).

TABLE 17

PRIMERS

	Forward	Reverse
β-catenin	5' TGA GGA CAA GCC ACA AGA TTA C 3'	5' TCC ACC AGA GTG AAA AGA ACG 3'
gapdh	5' ACC ACA GTC CAT GCC ATC AC 3'	5' TCC ACC ACC CTG TTG CTG TA 3'
ctnnb1 exon2	5' GAT TTG ATG GAG TTG GAC ATG GCC 3'	5' GAA GGA GCT GTG GTA GTG GCA C 3'
gapdh promoter	5' TAC TAG CGG TTT TAC GGG CG 3'	5' TCG AAC AGG AGG AGC AGA GAG CGA 3'
ctnnb1 promoter / enhancer	5' TGC AGC GGT ACT TGT CTG AG 3'	5' GTT GCT ATG GCC TGG TCA CT 3'
Axin2	5' AGG GAG AAA TGC GTG GAT AC 3'	5' TGG AAT CAA TCT GCT GCT TC 3'
Lef1	5' AAT GAG AGC GAA TGT CGT TGC 3'	5' GCT GTC TTT CTT TCC GTG CTA 3'
Apcdd1	5' GTC TTC AAC GGG AAT GAG TG 3'	5' ACC GTT GAA CAG CAG ATA GC 3'

Primers used for quantitative real time PCR, RT-PCR and ChIP assays. Primer sets were ordered as indicated in Materials and Methods in Chapter III and forward and reverse sequences are displayed.

REFERENCES

Aberle, H., A. Bauer, J. Stappert, A. Kispert, and R. Kemler. 1997. Beta-catenin is a target for the ubiquitin-proteasome pathway. *EMBO J* 16, no. 13: 3797-804.

Alazzouzi, H., P. Alhopuro, R. Salovaara, H. Sammalkorpi, H. Jarvinen, J. P. Mecklin, A. Hemminki, S. Schwartz, Jr., L. A. Aaltonen, and D. Arango. 2005. Smad4 as a prognostic marker in colorectal cancer. *Clin Cancer Res* 11, no. 7: 2606-11.

Arteaga, C. L., A. K. Tandon, D. D. Von Hoff, and C. K. Osborne. 1988. Transforming growth factor beta: Potential autocrine growth inhibitor of estrogen receptor-negative human breast cancer cells. *Cancer Res* 48, no. 14: 3898-904.

Ashton-Rickardt, P. G., M. G. Dunlop, Y. Nakamura, R. G. Morris, C. A. Purdie, C. M. Steel, H. J. Evans, C. C. Bird, and A. H. Wyllie. 1989. High frequency of apc loss in sporadic colorectal carcinoma due to breaks clustered in 5q21-22. *Oncogene* 4, no. 10: 1169-74.

Baker, S. J., E. R. Fearon, J. M. Nigro, S. R. Hamilton, A. C. Preisinger, J. M. Jessup, P. vanTuinen, D. H. Ledbetter, D. F. Barker, Y. Nakamura, R. White, and B. Vogelstein. 1989. Chromosome 17 deletions and p53 gene mutations in colorectal carcinomas. *Science* 244, no. 4901: 217-21.

Baker, S. J., A. C. Preisinger, J. M. Jessup, C. Paraskeva, S. Markowitz, J. K. Willson, S. Hamilton, and B. Vogelstein. 1990. P53 gene mutations occur in combination with 17p allelic deletions as late events in colorectal tumorigenesis. *Cancer Res* 50, no. 23: 7717-22.

Bandapalli, O. R., S. Dihlmann, R. Helwa, S. Macher-Goeppinger, J. Weitz, P. Schirmacher, and K. Brand. 2009. Transcriptional activation of the beta-catenin gene at the invasion front of colorectal liver metastases. *J Pathol* 218, no. 3: 370-9.

Barnard, J. A., R. D. Beauchamp, R. J. Coffey, and H. L. Moses. 1989. Regulation of intestinal epithelial cell growth by transforming growth factor type beta. *Proc Natl Acad Sci U S A* 86, no. 5: 1578-82.

Barnard, J. A., G. J. Warwick, and L. I. Gold. 1993. Localization of transforming growth factor beta isoforms in the normal murine small intestine and colon. *Gastroenterology* 105, no. 1: 67-73.

Barrier, A., P. Y. Boelle, F. Roser, J. Gregg, C. Tse, D. Brault, F. Lacaine, S. Houry, M. Huguier, B. Franc, A. Flahault, A. Lemoine, and S. Dudoit. 2006. Stage ii colon cancer prognosis prediction by tumor gene expression profiling. *J Clin Oncol* 24, no. 29: 4685-91.

Batlle, E., J. T. Henderson, H. Beghtel, M. M. van den Born, E. Sancho, G. Huls, J. Meeldijk, J. Robertson, M. van de Wetering, T. Pawson, and H. Clevers. 2002. Beta-catenin and tcf mediate cell positioning in the intestinal epithelium by controlling the expression of ephb/ephrinb. *Cell* 111, no. 2: 251-63.

Batlle, E., E. Sancho, C. Franci, D. Dominguez, M. Monfar, J. Baulida, and A. Garcia De Herreros. 2000. The transcription factor snail is a repressor of e-cadherin gene expression in epithelial tumour cells. *Nat Cell Biol* 2, no. 2: 84-9.

Beck, S. E., B. H. Jung, A. Fiorino, J. Gomez, E. D. Rosario, B. L. Cabrera, S. C. Huang, J. Y. Chow, and J. M. Carethers. 2006. Bone morphogenetic protein signaling and growth suppression in colon cancer. *Am J Physiol Gastrointest Liver Physiol* 291, no. 1: G135-45.

Beer, D. G., S. L. Kardia, C. C. Huang, T. J. Giordano, A. M. Levin, D. E. Misek, L. Lin, G. Chen, T. G. Gharib, D. G. Thomas, M. L. Lizyness, R. Kuick, S. Hayasaka, J. M. Taylor, M. D. Iannettoni, M. B. Orringer, and S. Hanash. 2002. Gene-expression profiles predict survival of patients with lung adenocarcinoma. *Nat Med* 8, no. 8: 816-24.

Behrens, J., B. A. Jerchow, M. Wurtele, J. Grimm, C. Asbrand, R. Wirtz, M. Kuhl, D. Wedlich, and W. Birchmeier. 1998. Functional interaction of an axin homolog, conductin, with beta-catenin, apc, and gsk3beta. *Science* 280, no. 5363: 596-9.

Beildeck, M. E., M. Islam, S. Shah, J. Welsh, and S. W. Byers. 2009. Control of tcf-4 expression by vdr and vitamin d in the mouse mammary gland and colorectal cancer cell lines. *PLoS One* 4, no. 11: e7872.

Benson, A. B., 3rd, D. Schrag, M. R. Somerfield, A. M. Cohen, A. T. Figueredo, P. J. Flynn, M. K. Krzyzanowska, J. Maroun, P. McAllister, E. Van Cutsem, M. Brouwers, M. Charette, and D. G. Haller. 2004. American society of clinical oncology recommendations on adjuvant chemotherapy for stage ii colon cancer. *J Clin Oncol* 22, no. 16: 3408-19.

Bertagnolli, M. M., C. J. Eagle, A. G. Zauber, M. Redston, S. D. Solomon, K. Kim, J. Tang, R. B. Rosenstein, J. Wittes, D. Corle, T. M. Hess, G. M. Woloj, F. Boisserie, W. F. Anderson, J. L. Viner, D. Bagheri, J. Burn, D. C. Chung, T. Dewar, T. R. Foley, N. Hoffman, F. Macrae, R. E. Pruitt, J. R. Saltzman, B. Salzberg, T. Sylwestrowicz, G. B. Gordon, and E. T. Hawk. 2006. Celecoxib for the prevention of sporadic colorectal adenomas. *N Engl J Med* 355, no. 9: 873-84.

Bertagnolli, M. M., R. S. Warren, D. Niedzwiecki, E. Mueller, C. C. Compton, M. Redston, M. Hall, H. P. Hahn, S. D. Jewell, R. J. Mayer, R. M. Goldberg, L. B. Saltz, and M. Loda. 2009. P27kip1 in stage iii colon cancer: Implications for outcome following adjuvant chemotherapy in cancer and leukemia group b protocol 89803. *Clin Cancer Res* 15, no. 6: 2116-22.

Blaker, H., S. Aulmann, B. Helmchen, H. F. Otto, R. J. Rieker, and R. Penzel. 2004. Loss of smad4 function in small intestinal adenocarcinomas: Comparison of genetic and immunohistochemical findings. *Pathol Res Pract* 200, no. 1: 1-7.

Boman, B. M. and E. Huang. 2008. Human colon cancer stem cells: A new paradigm in gastrointestinal oncology. *J Clin Oncol* 26, no. 17: 2828-38.

Bonafede, A., T. Kohler, M. Rodriguez-Niedenfuhr, and B. Brand-Saberi. 2006. Bmps restrict the position of premuscle masses in the limb buds by influencing tcf4 expression. *Dev Biol* 299, no. 2: 330-44.

Bos, J. L., E. R. Fearon, S. R. Hamilton, M. Verlaan-de Vries, J. H. van Boom, A. J. van der Eb, and B. Vogelstein. 1987. Prevalence of ras gene mutations in human colorectal cancers. *Nature* 327, no. 6120: 293-7.

Brabletz, T., F. Hlubek, S. Spaderna, O. Schmalhofer, E. Hiendlmeyer, A. Jung, and T. Kirchner. 2005. Invasion and metastasis in colorectal cancer: Epithelial-mesenchymal transition, mesenchymal-epithelial transition, stem cells and beta-catenin. *Cells Tissues Organs* 179, no. 1-2: 56-65.

Brabletz, T., A. Jung, S. Reu, M. Porzner, F. Hlubek, L. A. Kunz-Schughart, R. Knuechel, and T. Kirchner. 2001. Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment. *Proc Natl Acad Sci U S A* 98, no. 18: 10356-61.

Cadigan, K. M. and R. Nusse. 1997. Wnt signaling: A common theme in animal development. *Genes Dev* 11, no. 24: 3286-305.

Capon, D. J., P. H. Seeburg, J. P. McGrath, J. S. Hayflick, U. Edman, A. D. Levinson, and D. V. Goeddel. 1983. Activation of ki-ras2 gene in human colon and lung carcinomas by two different point mutations. *Nature* 304, no. 5926: 507-13.

Cha, Y. I. and R. N. DuBois. 2007. Nsaids and cancer prevention: Targets downstream of cox-2. *Annu Rev Med* 58: 239-52.

Chang, H. Y., J. B. Sneddon, A. A. Alizadeh, R. Sood, R. B. West, K. Montgomery, J. T. Chi, M. van de Rijn, D. Botstein, and P. O. Brown. 2004. Gene expression signature of fibroblast serum response predicts human cancer progression: Similarities between tumors and wounds. *PLoS Biol* 2, no. 2: E7.

Chen, J., N. Imanaka, and J. D. Griffin. Hypoxia potentiates notch signaling in breast cancer leading to decreased e-cadherin expression and increased cell migration and invasion. *Br J Cancer* 102, no. 2: 351-60.

Christofori, G. and H. Semb. 1999. The role of the cell-adhesion molecule e-cadherin as a tumour-suppressor gene. *Trends Biochem Sci* 24, no. 2: 73-6.

Clevers, H. 2006. Wnt/beta-catenin signaling in development and disease. *Cell* 127, no. 3: 469-80.

Coffey, R. J., C. J. Hawkey, L. Damstrup, R. Graves-Deal, V. C. Daniel, P. J. Dempsey, R. Chinery, S. C. Kirkland, R. N. DuBois, T. L. Jetton, and J. D. Morrow. 1997. Epidermal growth factor receptor activation induces nuclear targeting of cyclooxygenase-2, basolateral release of prostaglandins, and mitogenesis in polarizing colon cancer cells. *Proc Natl Acad Sci U S A* 94, no. 2: 657-62.

Cooper, H. S., S. Murthy, K. Kido, H. Yoshitake, and A. Flanigan. 2000. Dysplasia and cancer in the dextran sulfate sodium mouse colitis model. Relevance to colitis-associated neoplasia in the human: A study of histopathology, b-catenin and p53 expression and the role of inflammation. *Carcinogenesis* 21, no. 4: 757-68.

Corbett, T. H., D. P. Griswold, Jr., B. J. Roberts, J. C. Peckham, and F. M. Schabel, Jr. 1975. Tumor induction relationships in development of transplantable cancers of the colon in mice for chemotherapy assays, with a note on carcinogen structure. *Cancer Res* 35, no. 9: 2434-9.

Cortina, C., S. Palomo-Ponce, M. Iglesias, J. L. Fernandez-Masip, A. Vivancos, G. Whissell, M. Huma, N. Peiro, L. Gallego, S. Jonkheer, A. Davy, J. Lloreta, E. Sancho, and E. Batlle. 2007. Ephb-ephrin-b interactions suppress colorectal cancer progression by compartmentalizing tumor cells. *Nat Genet* 39, no. 11: 1376-83.

Cronin, M., C. Sangli, M. L. Liu, M. Pho, D. Dutta, A. Nguyen, J. Jeong, J. Wu, K. C. Langone, and D. Watson. 2007. Analytical validation of the oncotype dx genomic diagnostic test for recurrence prognosis and therapeutic response prediction in node-negative, estrogen receptor-positive breast cancer. *Clin Chem* 53, no. 6: 1084-91.

Cui, W., D. J. Fowlis, S. Bryson, E. Duffie, H. Ireland, A. Balmain, and R. J. Akhurst. 1996. Tgfbeta1 inhibits the formation of benign skin tumors, but enhances progression to invasive spindle carcinomas in transgenic mice. *Cell* 86, no. 4: 531-42.

Davies, H., G. R. Bignell, C. Cox, P. Stephens, S. Edkins, S. Clegg, J. Teague, H. Woffendin, M. J. Garnett, W. Bottomley, N. Davis, E. Dicks, R. Ewing, Y. Floyd, K. Gray, S. Hall, R. Hawes, J. Hughes, V. Kosmidou, A. Menzies, C. Mould, A. Parker, C. Stevens, S. Watt, S. Hooper, R. Wilson, H. Jayatilake, B. A. Gusterson, C. Cooper, J. Shipley, D. Hargrave, K. Pritchard-Jones, N. Maitland, G. Chenevix-Trench, G. J. Riggins, D. D. Bigner, G. Palmieri, A. Cossu, A. Flanagan, A. Nicholson, J. W. Ho, S. Y. Leung, S. T. Yuen, B. L. Weber, H. F. Seigler, T. L. Darrow, H. Paterson, R. Marais, C. J. Marshall, R. Wooster, M. R. Stratton, and P. A. Futreal. 2002. Mutations of the braf gene in human cancer. *Nature* 417, no. 6892: 949-54.

Davies, J. A. 1996. Mesenchyme to epithelium transition during development of the mammalian kidney tubule. *Acta Anat (Basel)* 156, no. 3: 187-201.

Davis, M. A., R. C. Ireton, and A. B. Reynolds. 2003. A core function for p120-catenin in cadherin turnover. *J Cell Biol* 163, no. 3: 525-34.

Dhawan, P., A. B. Singh, N. G. Deane, Y. No, S. R. Shiou, C. Schmidt, J. Neff, M. K. Washington, and R. D. Beauchamp. 2005. Claudin-1 regulates cellular transformation and metastatic behavior in colon cancer. *J Clin Invest* 115, no. 7: 1765-76.

Egan, J. B., P. A. Thompson, E. L. Ashbeck, D. V. Conti, D. Duggan, E. Hibler, P. W. Jurutka, E. C. Leroy, M. E. Martinez, D. Mount, and E. T. Jacobs. Genetic polymorphisms in vitamin d receptor vdr/rxra influence the likelihood of colon adenoma recurrence. *Cancer Res* 70, no. 4: 1496-504.

Elliott, R. L. and G. C. Blobe. 2005. Role of transforming growth factor beta in human cancer. *J Clin Oncol* 23, no. 9: 2078-93.

Ellis, L. M. and D. J. Hicklin. 2008. Vegf-targeted therapy: Mechanisms of anti-tumour activity. *Nat Rev Cancer* 8, no. 8: 579-91.

Eppert, K., S. W. Scherer, H. Ozcelik, R. Pirone, P. Hoodless, H. Kim, L. C. Tsui, B. Bapat, S. Gallinger, I. L. Andrulis, G. H. Thomsen, J. L. Wrana, and L. Attisano. 1996. Madr2 maps to 18q21 and encodes a tgfbeta-regulated mad-related protein that is functionally mutated in colorectal carcinoma. *Cell* 86, no. 4: 543-52.

Eschrich, S., I. Yang, G. Bloom, K. Y. Kwong, D. Boulware, A. Cantor, D. Coppola, M. Kruhoffer, L. Aaltonen, T. F. Orntoft, J. Quackenbush, and T. J. Yeatman. 2005. Molecular staging for survival prediction of colorectal cancer patients. *J Clin Oncol* 23, no. 15: 3526-35.

Fearon, E. R. and B. Vogelstein. 1990. A genetic model for colorectal tumorigenesis. *Cell* 61, no. 5: 759-67.

Fidler, I. J. and G. L. Nicolson. 1976. Organ selectivity for implantation survival and growth of b16 melanoma variant tumor lines. *J Natl Cancer Inst* 57, no. 5: 1199-202.

Figueredo, A., M. L. Charette, J. Maroun, M. C. Brouwers, and L. Zuraw. 2004. Adjuvant therapy for stage ii colon cancer: A systematic review from the cancer care ontario program in evidence-based care's gastrointestinal cancer disease site group. *J Clin Oncol* 22, no. 16: 3395-407.

Figueredo, A., M. E. Coombes, and S. Mukherjee. 2008. Adjuvant therapy for completely resected stage ii colon cancer. *Cochrane Database Syst Rev*, no. 3: CD005390.

Fodde, R. and T. Brabletz. 2007. Wnt/beta-catenin signaling in cancer stemness and malignant behavior. *Curr Opin Cell Biol* 19, no. 2: 150-8.

Franks, L. M. and V. J. Hemmings. 1978. A cell line from an induced carcinoma of mouse rectum. *J Pathol* 124, no. 1: 35-8.

Garman, K. S., C. R. Acharya, E. Edelman, M. Grade, J. Gaedcke, S. Sud, W. Barry, A. M. Diehl, D. Provenzale, G. S. Ginsburg, B. M. Ghadimi, T. Ried, J. R. Nevins, S. Mukherjee, D. Hsu, and A. Potti. 2008. A genomic approach to colon cancer risk stratification yields biologic insights into therapeutic opportunities. *Proc Natl Acad Sci U S A* 105, no. 49: 19432-7.

Gherzi, R., M. Trabucchi, M. Ponassi, T. Ruggiero, G. Corte, C. Moroni, C. Y. Chen, K. S. Khabar, J. S. Andersen, and P. Briata. 2006. The rna-binding protein ksrp promotes decay of beta-catenin mrna and is inactivated by pi3k-akt signaling. *PLoS Biol* 5, no. 1: e5.

Gill, S., C. L. Loprinzi, D. J. Sargent, S. D. Thome, S. R. Alberts, D. G. Haller, J. Benedetti, G. Francini, L. E. Shepherd, J. Francois Seitz, R. Labianca, W. Chen, S. S. Cha, M. P. Heldebrant, and R. M. Goldberg. 2004. Pooled analysis of fluorouracil-based adjuvant therapy for stage ii and iii colon cancer: Who benefits and by how much? *J Clin Oncol* 22, no. 10: 1797-806.

Gilles, C., M. Polette, M. Mestdagt, B. Nawrocki-Raby, P. Ruggeri, P. Birembaut, and J. M. Foidart. 2003. Transactivation of vimentin by beta-catenin in human breast cancer cells. *Cancer Res* 63, no. 10: 2658-64.

Gordon, M. D. and R. Nusse. 2006. Wnt signaling: Multiple pathways, multiple receptors, and multiple transcription factors. *J Biol Chem* 281, no. 32: 22429-33.

Goss, K. H. and J. Groden. 2000. Biology of the adenomatous polyposis coli tumor suppressor. *J Clin Oncol* 18, no. 9: 1967-79.

Grady, W. M., L. L. Myeroff, S. E. Swinler, A. Rajput, S. Thiagalingam, J. D. Lutterbaugh, A. Neumann, M. G. Brattain, J. Chang, S. J. Kim, K. W. Kinzler, B. Vogelstein, J. K. Willson, and S. Markowitz. 1999. Mutational inactivation of transforming growth factor beta receptor type ii in microsatellite stable colon cancers. *Cancer Res* 59, no. 2: 320-4.

Greenburg, G. and E. D. Hay. 1982. Epithelia suspended in collagen gels can lose polarity and express characteristics of migrating mesenchymal cells. *J Cell Biol* 95, no. 1: 333-9.

Gryfe, R., H. Kim, E. T. Hsieh, M. D. Aronson, E. J. Holowaty, S. B. Bull, M. Redston, and S. Gallinger. 2000. Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. *N Engl J Med* 342, no. 2: 69-77.

Hahn, S. A., A. T. Hoque, C. A. Moskaluk, L. T. da Costa, M. Schutte, E. Rozenblum, A. B. Seymour, C. L. Weinstein, C. J. Yeo, R. H. Hruban, and S. E. Kern. 1996. Homozygous deletion map at 18q21.1 in pancreatic cancer. *Cancer Res* 56, no. 3: 490-4.

Hahn, S. A., M. Schutte, A. T. Hoque, C. A. Moskaluk, L. T. da Costa, E. Rozenblum, C. L. Weinstein, A. Fischer, C. J. Yeo, R. H. Hruban, and S. E. Kern. 1996. Dpc4, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science* 271, no. 5247: 350-3.

Han, S. U., H. T. Kim, D. H. Seong, Y. S. Kim, Y. S. Park, Y. J. Bang, H. K. Yang, and S. J. Kim. 2004. Loss of the smad3 expression increases susceptibility to tumorigenicity in human gastric cancer. *Oncogene* 23, no. 7: 1333-41.

Hanahan, D. and R. A. Weinberg. 2000. The hallmarks of cancer. *Cell* 100, no. 1: 57-70.

Haramis, A. P., H. Begthel, M. van den Born, J. van Es, S. Jonkheer, G. J. Offerhaus, and H. Clevers. 2004. De novo crypt formation and juvenile polyposis on bmp inhibition in mouse intestine. *Science* 303, no. 5664: 1684-6.

Hardwick, J. C., G. R. Van Den Brink, S. A. Bleuming, I. Ballester, J. M. Van Den Brande, J. J. Keller, G. J. Offerhaus, S. J. Van Deventer, and M. P. Peppelenbosch. 2004. Bone morphogenetic protein 2 is expressed by, and acts upon, mature epithelial cells in the colon. *Gastroenterology* 126, no. 1: 111-21.

Harrell, F. E., Jr., R. M. Califf, D. B. Pryor, K. L. Lee, and R. A. Rosati. 1982. Evaluating the yield of medical tests. *JAMA* 247, no. 18: 2543-6.

He, T. C., A. B. Sparks, C. Rago, H. Hermeking, L. Zawel, L. T. da Costa, P. J. Morin, B. Vogelstein, and K. W. Kinzler. 1998. Identification of c-myc as a target of the apc pathway. *Science* 281, no. 5382: 1509-12.

He, X. C., J. Zhang, W. G. Tong, O. Tawfik, J. Ross, D. H. Scoville, Q. Tian, X. Zeng, X. He, L. M. Wiedemann, Y. Mishina, and L. Li. 2004. Bmp signaling inhibits intestinal stem cell self-renewal through suppression of wnt-beta-catenin signaling. *Nat Genet* 36, no. 10: 1117-21.

Hedenfalk, I., D. Duggan, Y. Chen, M. Radmacher, M. Bittner, R. Simon, P. Meltzer, B. Gusterson, M. Esteller, O. P. Kallioniemi, B. Wilfond, A. Borg, J. Trent, M. Raffeld, Z. Yakhini, A. Ben-Dor, E. Dougherty, J. Kononen, L. Bubendorf, W. Fehrle, S. Pittaluga, S. Gruvberger, N. Loman, O. Johannsson, H. Olsson, and G. Sauter. 2001. Gene-expression profiles in hereditary breast cancer. *N Engl J Med* 344, no. 8: 539-48.

Heijstek, M. W., O. Kranenburg, and I. H. Borel Rinkes. 2005. Mouse models of colorectal cancer and liver metastases. *Dig Surg* 22, no. 1-2: 16-25.

Heinze, G., M. Gnant, and M. Schemper. 2003. Exact log-rank tests for unequal follow-up. *Biometrics* 59, no. 4: 1151-7.

Hlubek, F., T. Brabletz, J. Budczies, S. Pfeiffer, A. Jung, and T. Kirchner. 2007. Heterogeneous expression of wnt/beta-catenin target genes within colorectal cancer. *Int J Cancer* 121, no. 9: 1941-8.

Hoosein, N. M., M. K. McKnight, A. E. Levine, K. M. Mulder, K. E. Childress, D. E. Brattain, and M. G. Brattain. 1989. Differential sensitivity of subclasses of human colon carcinoma cell lines to the growth inhibitory effects of transforming growth factor-beta 1. *Exp Cell Res* 181, no. 2: 442-53.

Hoshida, Y., A. Villanueva, M. Kobayashi, J. Peix, D. Y. Chiang, A. Camargo, S. Gupta, J. Moore, M. J. Wrobel, J. Lerner, M. Reich, J. A. Chan, J. N. Glickman, K. Ikeda, M. Hashimoto, G. Watanabe, M. G. Daidone, S. Roayaie, M. Schwartz, S. Thung, H. B. Salvesen, S. Gabriel, V. Mazzaferro, J. Bruix, S. L.

Friedman, H. Kumada, J. M. Llovet, and T. R. Golub. 2008. Gene expression in fixed tissues and outcome in hepatocellular carcinoma. *N Engl J Med* 359, no. 19: 1995-2004.

Howe, J. R., J. L. Bair, M. G. Sayed, M. E. Anderson, F. A. Mitros, G. M. Petersen, V. E. Velculescu, G. Traverso, and B. Vogelstein. 2001. Germline mutations of the gene encoding bone morphogenetic protein receptor 1a in juvenile polyposis. *Nat Genet* 28, no. 2: 184-7.

Howe, J. R., S. Roth, J. C. Ringold, R. W. Summers, H. J. Jarvinen, P. Sistonen, I. P. Tomlinson, R. S. Houlston, S. Bevan, F. A. Mitros, E. M. Stone, and L. A. Aaltonen. 1998. Mutations in the smad4/dpc4 gene in juvenile polyposis. *Science* 280, no. 5366: 1086-8.

Howe, L. R., K. Subbaramaiah, W. J. Chung, A. J. Dannenberg, and A. M. Brown. 1999. Transcriptional activation of cyclooxygenase-2 in wnt-1-transformed mouse mammary epithelial cells. *Cancer Res* 59, no. 7: 1572-7.

Howe, L. R., O. Watanabe, J. Leonard, and A. M. Brown. 2003. Twist is up-regulated in response to wnt1 and inhibits mouse mammary cell differentiation. *Cancer Res* 63, no. 8: 1906-13.

Hurwitz, H., L. Fehrenbacher, W. Novotny, T. Cartwright, J. Hainsworth, W. Heim, J. Berlin, A. Baron, S. Griffing, E. Holmgren, N. Ferrara, G. Fyfe, B. Rogers, R. Ross, and F. Kabbinavar. 2004. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med* 350, no. 23: 2335-42.

Ikeda, S., S. Kishida, H. Yamamoto, H. Murai, S. Koyama, and A. Kikuchi. 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, no. 5: 1371-84.

Ikeguchi, M., M. Makino, and N. Kaibara. 2001. Clinical significance of e-cadherin-catenin complex expression in metastatic foci of colorectal carcinoma. *J Surg Oncol* 77, no. 3: 201-7.

Ikeguchi, M., T. Taniguchi, M. Makino, and N. Kaibara. 2000. Reduced e-cadherin expression and enlargement of cancer nuclei strongly correlate with hematogenic metastasis in colorectal adenocarcinoma. *Scand J Gastroenterol* 35, no. 8: 839-46.

Imperiale, T. F., D. F. Ransohoff, S. H. Itzkowitz, B. A. Turnbull, and M. E. Ross. 2004. Fecal DNA versus fecal occult blood for colorectal-cancer screening in an average-risk population. *N Engl J Med* 351, no. 26: 2704-14.

Ionov, Y., M. A. Peinado, S. Malkhosyan, D. Shibata, and M. Perucho. 1993. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* 363, no. 6429: 558-61.

Ireton, R. C., M. A. Davis, J. van Hengel, D. J. Mariner, K. Barnes, M. A. Thoreson, P. Z. Anastasiadis, L. Matrisian, L. M. Bundy, L. Sealy, B. Gilbert, F. van Roy, and A. B. Reynolds. 2002. A novel role for p120 catenin in e-cadherin function. *J Cell Biol* 159, no. 3: 465-76.

Irizarry, R. A., B. Hobbs, F. Collin, Y. D. Beazer-Barclay, K. J. Antonellis, U. Scherf, and T. P. Speed. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4, no. 2: 249-64.

Issa, J. P. 2004. Cpg island methylator phenotype in cancer. *Nat Rev Cancer* 4, no. 12: 988-93.

Jamora, C., R. DasGupta, P. Kocieniewski, and E. Fuchs. 2003. Links between signal transduction, transcription and adhesion in epithelial bud development. *Nature* 422, no. 6929: 317-22.

Janda, E., K. Lehmann, I. Killisch, M. Jechlinger, M. Herzig, J. Downward, H. Beug, and S. Grunert. 2002. Ras and tgf[beta] cooperatively regulate epithelial cell plasticity and metastasis: Dissection of ras signaling pathways. *J Cell Biol* 156, no. 2: 299-313.

Jemal, A., R. Siegel, E. Ward, Y. Hao, J. Xu, and M. J. Thun. 2009. Cancer statistics, 2009. *CA Cancer J Clin* 59, no. 4: 225-49.

Jenkins, D. E., Y. Oei, Y. S. Hornig, S. F. Yu, J. Dusich, T. Purchio, and P. R. Contag. 2003. Bioluminescent imaging (bli) to improve and refine traditional murine models of tumor growth and metastasis. *Clin Exp Metastasis* 20, no. 8: 733-44.

Jenny, M., C. Uhl, C. Roche, I. Duluc, V. Guillermin, F. Guillemot, J. Jensen, M. Kedinger, and G. Gradwohl. 2002. Neurogenin3 is differentially required for endocrine cell fate specification in the intestinal and gastric epithelium. *EMBO J* 21, no. 23: 6338-47.

Jensen, J., E. E. Pedersen, P. Galante, J. Hald, R. S. Heller, M. Ishibashi, R. Kageyama, F. Guillemot, P. Serup, and O. D. Madsen. 2000. Control of endodermal endocrine development by hes-1. *Nat Genet* 24, no. 1: 36-44.

Jorissen, R. N., P. Gibbs, M. Christie, S. Prakash, L. Lipton, J. Desai, D. Kerr, L. A. Aaltonen, D. Arango, M. Kruhoffer, T. F. Orntoft, C. L. Andersen, M. Gruidl, V. P. Kamath, S. Eschrich, T. J. Yeatman, and O. M. Sieber. 2009. Metastasis-associated gene expression changes predict poor outcomes in patients with dukes stage b and c colorectal cancer. *Clin Cancer Res* 15, no. 24: 7642-7651.

Kado, S., K. Uchida, H. Funabashi, S. Iwata, Y. Nagata, M. Ando, M. Onoue, Y. Matsuoka, M. Ohwaki, and M. Morotomi. 2001. Intestinal microflora are necessary for development of spontaneous adenocarcinoma of the large intestine in t-cell receptor beta chain and p53 double-knockout mice. *Cancer Res* 61, no. 6: 2395-8.

Kalluri, R. 2009. Emt: When epithelial cells decide to become mesenchymal-like cells. *J Clin Invest* 119, no. 6: 1417-9.

Kanazawa, N., T. Oda, N. Gunji, M. Nozue, T. Kawamoto, T. Todoroki, and K. Fukao. 2002. E-cadherin expression in the primary tumors and metastatic lymph nodes of poorly differentiated types of rectal cancer. *Surg Today* 32, no. 2: 123-8.

Kanazawa, T., T. Watanabe, S. Kazama, T. Tada, S. Koketsu, and H. Nagawa. 2002. Poorly differentiated adenocarcinoma and mucinous carcinoma of the colon and rectum show higher rates of loss of heterozygosity and loss of e-cadherin expression due to methylation of promoter region. *Int J Cancer* 102, no. 3: 225-9.

Kang, Y. and J. Massague. 2004. Epithelial-mesenchymal transitions: Twist in development and metastasis. *Cell* 118, no. 3: 277-9.

Kanies, C. L., J. J. Smith, C. Kis, C. Schmidt, S. Levy, K. S. Khabar, J. Morrow, N. Deane, D. A. Dixon, and R. D. Beauchamp. 2008. Oncogenic ras and transforming growth factor-beta synergistically regulate au-rich element-containing mrnas during epithelial to mesenchymal transition. *Mol Cancer Res* 6, no. 7: 1124-36.

Kaposi-Novak, P., J. S. Lee, L. Gomez-Quiroz, C. Coulouarn, V. M. Factor, and S. S. Thorgeirsson. 2006. Met-regulated expression signature defines a subset of human hepatocellular carcinomas with poor prognosis and aggressive phenotype. *J Clin Invest* 116, no. 6: 1582-95.

Kimchi, A., X. F. Wang, R. A. Weinberg, S. Cheifetz, and J. Massague. 1988. Absence of tgf-beta receptors and growth inhibitory responses in retinoblastoma cells. *Science* 240, no. 4849: 196-9.

Kishida, S., H. Yamamoto, S. Ikeda, M. Kishida, I. Sakamoto, S. Koyama, and A. Kikuchi. 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, no. 18: 10823-6.

Ko, T. C., H. M. Sheng, D. Reisman, E. A. Thompson, and R. D. Beauchamp. 1995. Transforming growth factor-beta 1 inhibits cyclin d1 expression in intestinal epithelial cells. *Oncogene* 10, no. 1: 177-84.

Ko, T. C., W. Yu, T. Sakai, H. Sheng, J. Shao, R. D. Beauchamp, and E. A. Thompson. 1998. Tgf-beta1 effects on proliferation of rat intestinal epithelial cells are due to inhibition of cyclin d1 expression. *Oncogene* 16, no. 26: 3445-54.

Kobaek-Larsen, M., I. Thorup, A. Diederichsen, C. Fenger, and M. R. Hoitinga. 2000. Review of colorectal cancer and its metastases in rodent models: Comparative aspects with those in humans. *Comp Med* 50, no. 1: 16-26.

Koh, T. W. and H. J. Bellen. 2003. Synaptotagmin i, a ca2+ sensor for neurotransmitter release. *Trends Neurosci* 26, no. 8: 413-22.

Korinek, V., N. Barker, P. J. Morin, D. van Wichen, R. de Weger, K. W. Kinzler, B. Vogelstein, and H. Clevers. 1997. Constitutive transcriptional activation by a beta-catenin-tcf complex in apc-/- colon carcinoma. *Science* 275, no. 5307: 1784-7.

Kosinski, C., V. S. Li, A. S. Chan, J. Zhang, C. Ho, W. Y. Tsui, T. L. Chan, R. C. Mifflin, D. W. Powell, S. T. Yuen, S. Y. Leung, and X. Chen. 2007. Gene expression patterns of human colon tops and basal crypts and bmp antagonists as intestinal stem cell niche factors. *Proc Natl Acad Sci U S A* 104, no. 39: 15418-23.

Kure, S., K. Nosho, Y. Baba, N. Irahara, K. Shima, K. Ng, J. A. Meyerhardt, E. L. Giovannucci, C. S. Fuchs, and S. Ogino. 2009. Vitamin d receptor expression is associated with pik3ca and kras mutations in colorectal cancer. *Cancer Epidemiol Biomarkers Prev* 18, no. 10: 2765-72.

Kurokowa, M., K. Lynch, and D. K. Podolsky. 1987. Effects of growth factors on an intestinal epithelial cell line: Transforming growth factor beta inhibits proliferation and stimulates differentiation. *Biochem Biophys Res Commun* 142, no. 3: 775-82.

Labbe, E., A. Letamendia, and L. Attisano. 2000. Association of smads with lymphoid enhancer binding factor 1/t cell-specific factor mediates cooperative signaling by the transforming growth factor-beta and wnt pathways. *Proc Natl Acad Sci U S A* 97, no. 15: 8358-63.

Labbe, E., L. Lock, A. Letamendia, A. E. Gorska, R. Gryfe, S. Gallinger, H. L. Moses, and L. Attisano. 2007. Transcriptional cooperation between the transforming growth factor-beta and wnt pathways in mammary and intestinal tumorigenesis. *Cancer Res* 67, no. 1: 75-84.

Lafreniere, R. and S. A. Rosenberg. 1986. A novel approach to the generation and identification of experimental hepatic metastases in a murine model. *J Natl Cancer Inst* 76, no. 2: 309-22.

Lamb, J., E. D. Crawford, D. Peck, J. W. Modell, I. C. Blat, M. J. Wrobel, J. Lerner, J. P. Brunet, A. Subramanian, K. N. Ross, M. Reich, H. Hieronymus, G. Wei, S. A. Armstrong, S. J. Haggarty, P. A. Clemons, R. Wei, S. A. Carr, E. S. Lander, and T. R. Golub. 2006. The connectivity map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science* 313, no. 5795: 1929-35.

Leary, R. J., J. C. Lin, J. Cummins, S. Boca, L. D. Wood, D. W. Parsons, S. Jones, T. Sjoblom, B. H. Park, R. Parsons, J. Willis, D. Dawson, J. K. Willson, T. Nikolskaya, Y. Nikolsky, L. Kopelovich, N. Papadopoulos, L. A. Pennacchio, T. L. Wang, S. D. Markowitz, G. Parmigiani, K. W. Kinzler, B. Vogelstein, and V. E. Velculescu. 2008. Integrated analysis of homozygous deletions, focal amplifications, and sequence alterations in breast and colorectal cancers. *Proc Natl Acad Sci U S A* 105, no. 42: 16224-9.

Lee, J. S., I. S. Chu, A. Mikaelyan, D. F. Calvisi, J. Heo, J. K. Reddy, and S. S. Thorgeirsson. 2004. Application of comparative functional genomics to identify best-fit mouse models to study human cancer. *Nat Genet* 36, no. 12: 1306-11.

Lee, J. S., J. Heo, L. Libbrecht, I. S. Chu, P. Kaposi-Novak, D. F. Calvisi, A. Mikaelyan, L. R. Roberts, A. J. Demetris, Z. Sun, F. Nevens, T. Roskams, and S. S. Thorgeirsson. 2006. A novel prognostic subtype of human hepatocellular carcinoma derived from hepatic progenitor cells. *Nat Med* 12, no. 4: 410-6.

Lee, J. S. and S. S. Thorgeirsson. 2004. Genome-scale profiling of gene expression in hepatocellular carcinoma: Classification, survival prediction, and identification of therapeutic targets. *Gastroenterology* 127, no. 5 Suppl 1: S51-5.

Lengauer, C., K. W. Kinzler, and B. Vogelstein. 1997. Genetic instability in colorectal cancers. *Nature* 386, no. 6625: 623-7.

Li, M., W. D. Chen, N. Papadopoulos, S. N. Goodman, N. C. Bjerregaard, S. Laurberg, B. Levin, H. Juhl, N. Arber, H. Moinova, K. Durkee, K. Schmidt, Y. He, F. Diehl, V. E. Velculescu, S. Zhou, L. A. Diaz, Jr., K. W. Kinzler, S. D. Markowitz, and B. Vogelstein. 2009. Sensitive digital quantification of DNA methylation in clinical samples. *Nat Biotechnol* 27, no. 9: 858-63.

Liebig, C., G. Ayala, J. A. Wilks, D. H. Berger, and D. Albo. 2009. Perineural invasion in cancer: A review of the literature. *Cancer* 115, no. 15: 3379-91.

Liebig, C., G. Ayala, J. Wilks, G. Verstovsek, H. Liu, N. Agarwal, D. H. Berger, and D. Albo. 2009. Perineural invasion is an independent predictor of outcome in colorectal cancer. *J Clin Oncol* 27, no. 31: 5131-7.

Lin, Y. H., J. Friederichs, M. A. Black, J. Mages, R. Rosenberg, P. J. Guilford, V. Phillips, M. Thompson-Fawcett, N. Kasabov, T. Toro, A. E. Merrie, A. van Rij, H. S. Yoon, J. L. McCall, J. R. Siewert, B. Holzmann, and A. E. Reeve. 2007. Multiple gene expression classifiers from different array platforms predict poor prognosis of colorectal cancer. *Clin Cancer Res* 13, no. 2 Pt 1: 498-507.

Lynch, H. T., J. F. Lynch, P. M. Lynch, and T. Attard. 2008. Hereditary colorectal cancer syndromes: Molecular genetics, genetic counseling, diagnosis and management. *Fam Cancer* 7, no. 1: 27-39.

Mamounas, E., S. Wieand, N. Wolmark, H. D. Bear, J. N. Atkins, K. Song, J. Jones, and H. Rockette. 1999. Comparative efficacy of adjuvant chemotherapy in patients with dukes' b versus dukes' c colon cancer: Results from four national surgical adjuvant breast and bowel project adjuvant studies (c-01, c-02, c-03, and c-04). *J Clin Oncol* 17, no. 5: 1349-55.

Maniatis, T. 1999. A ubiquitin ligase complex essential for the nf-kappab, wnt/wingless, and hedgehog signaling pathways. *Genes Dev* 13, no. 5: 505-10.

Manning, A. M., A. C. Williams, S. M. Game, and C. Paraskeva. 1991. Differential sensitivity of human colonic adenoma and carcinoma cells to transforming growth factor beta (tgf-beta): Conversion of an adenoma cell line to a tumorigenic phenotype is accompanied by a reduced response to the inhibitory effects of tgf-beta. *Oncogene* 6, no. 8: 1471-6.

Markowitz, S. D. 2007. Aspirin and colon cancer--targeting prevention? *N Engl J Med* 356, no. 21: 2195-8.

Markowitz, S. D. and M. M. Bertagnolli. 2009. Molecular origins of cancer: Molecular basis of colorectal cancer. *N Engl J Med* 361, no. 25: 2449-60.

Markowitz, S., J. Wang, L. Myeroff, R. Parsons, L. Sun, J. Lutterbaugh, R. S. Fan, E. Zborowska, K. W. Kinzler, B. Vogelstein, and et al. 1995. Inactivation of the type ii tgf-beta receptor in colon cancer cells with microsatellite instability. *Science* 268, no. 5215: 1336-8.

Massague, J. 2008. Tgfbeta in cancer. *Cell* 134, no. 2: 215-30.

Massague, J., J. Seoane, and D. Wotton. 2005. Smad transcription factors. *Genes Dev* 19, no. 23: 2783-810.

Mbalaviele, G., C. R. Dunstan, A. Sasaki, P. J. Williams, G. R. Mundy, and T. Yoneda. 1996. E-cadherin expression in human breast cancer cells suppresses the development of osteolytic bone metastases in an experimental metastasis model. *Cancer Res* 56, no. 17: 4063-70.

Mikami, S., M. Oya, M. Shimoda, R. Mizuno, M. Ishida, T. Kosaka, M. Mukai, M. Nakajima, and Y. Okada. 2008. Expression of heparanase in renal cell carcinomas: Implications for tumor invasion and prognosis. *Clin Cancer Res* 14, no. 19: 6055-61.

Miyaki, M., T. Iijima, M. Konishi, K. Sakai, A. Ishii, M. Yasuno, T. Hishima, M. Koike, N. Shitara, T. Iwama, J. Utsunomiya, T. Kuroki, and T. Mori. 1999. Higher frequency of smad4 gene mutation in human colorectal cancer with distant metastasis. *Oncogene* 18, no. 20: 3098-103.

Morin, P. J., A. B. Sparks, V. Korinek, N. Barker, H. Clevers, B. Vogelstein, and K. W. Kinzler. 1997. Activation of beta-catenin-tcf signaling in colon cancer by mutations in beta-catenin or apc. *Science* 275, no. 5307: 1787-90.

Moser, A. R., H. C. Pitot, and W. F. Dove. 1990. A dominant mutation that predisposes to multiple intestinal neoplasia in the mouse. *Science* 247, no. 4940: 322-4.

Moses, H. L., E. Y. Yang, and J. A. Pietenpol. 1990. Tgf-beta stimulation and inhibition of cell proliferation: New mechanistic insights. *Cell* 63, no. 2: 245-7.

Muller, N., A. Reinacher-Schick, S. Baldus, J. van Hengel, G. Berx, A. Baar, F. van Roy, W. Schmiegel, and I. Schwarte-Waldhoff. 2002. Smad4 induces the tumor suppressor e-cadherin and p-cadherin in colon carcinoma cells. *Oncogene* 21, no. 39: 6049-58.

Munemitsu, S., I. Albert, B. Souza, B. Rubinfeld, and P. Polakis. 1995. Regulation of intracellular beta-catenin levels by the adenomatous polyposis coli (apc) tumor-suppressor protein. *Proc Natl Acad Sci U S A* 92, no. 7: 3046-50.

Neilson, E. G. 2005. Setting a trap for tissue fibrosis. *Nat Med* 11, no. 4: 373-4.

Nelson, W. J. and R. Nusse. 2004. Convergence of wnt, beta-catenin, and cadherin pathways. *Science* 303, no. 5663: 1483-7.

Nishita, M., M. K. Hashimoto, S. Ogata, M. N. Laurent, N. Ueno, H. Shibuya, and K. W. Cho. 2000. Interaction between wnt and tgf-beta signalling pathways during formation of spemann's organizer. *Nature* 403, no. 6771: 781-5.

Nollet, F., G. Berx, F. Molemans, and F. van Roy. 1996. Genomic organization of the human beta-catenin gene (ctnnb1). *Genomics* 32, no. 3: 413-24.

Nusse, R. and H. E. Varmus. 1982. Many tumors induced by the mouse mammary tumor virus contain a provirus integrated in the same region of the host genome. *Cell* 31, no. 1: 99-109.

O'Brien, C. A., A. Pollett, S. Gallinger, and J. E. Dick. 2007. A human colon cancer cell capable of initiating tumour growth in immunodeficient mice. *Nature* 445, no. 7123: 106-10.

Ohtaki, N., A. Yamaguchi, T. Goi, T. Fukaya, K. Takeuchi, K. Katayama, K. Hirose, and T. Urano. 2001. Somatic alterations of the dpc4 and madr2 genes in colorectal cancers and relationship to metastasis. *Int J Oncol* 18, no. 2: 265-70.

Paik, S., S. Shak, G. Tang, C. Kim, J. Baker, M. Cronin, F. L. Baehner, M. G. Walker, D. Watson, T. Park, W. Hiller, E. R. Fisher, D. L. Wickerham, J. Bryant, and N. Wolmark. 2004. A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer. *N Engl J Med* 351, no. 27: 2817-26.

Parsons, D. W., T. L. Wang, Y. Samuels, A. Bardelli, J. M. Cummins, L. DeLong, N. Silliman, J. Ptak, S. Szabo, J. K. Willson, S. Markowitz, K. W. Kinzler, B. Vogelstein, C. Lengauer, and V. E. Velculescu. 2005. Colorectal cancer: Mutations in a signalling pathway. *Nature* 436, no. 7052: 792.

Pena, C., J. M. Garcia, M. J. Larriba, R. Barderas, I. Gomez, M. Herrera, V. Garcia, J. Silva, G. Dominguez, R. Rodriguez, J. Cuevas, A. G. de Herreros, J. I. Casal, A. Munoz, and F. Bonilla. 2009. Snai1 expression in colon cancer related with cdh1 and vdr downregulation in normal adjacent tissue. *Oncogene* 28, no. 49: 4375-85.

Perez-Moreno, M. A., A. Locascio, I. Rodrigo, G. Dhondt, F. Portillo, M. A. Nieto, and A. Cano. 2001. A new role for e12/e47 in the repression of e-cadherin expression and epithelial-mesenchymal transitions. *J Biol Chem* 276, no. 29: 27424-31.

Pietenpol, J. A., R. W. Stein, E. Moran, P. Yaciuk, R. Schlegel, R. M. Lyons, M. R. Pittelkow, K. Munger, P. M. Howley, and H. L. Moses. 1990. Tgf-beta 1 inhibition of c-myc transcription and growth in keratinocytes is abrogated by viral transforming proteins with prb binding domains. *Cell* 61, no. 5: 777-85.

Poste, G., J. Doll, and I. J. Fidler. 1981. Interactions among clonal subpopulations affect stability of the metastatic phenotype in polyclonal populations of b16 melanoma cells. *Proc Natl Acad Sci U S A* 78, no. 10: 6226-30.

Quasar Collaborative, Group, R. Gray, J. Barnwell, C. McConkey, R. K. Hills, N. S. Williams, and D. J. Kerr. 2007. Adjuvant chemotherapy versus observation in patients with colorectal cancer: A randomised study. *Lancet* 370, no. 9604: 2020-9.

Radtke, F. and H. Clevers. 2005. Self-renewal and cancer of the gut: Two sides of a coin. *Science* 307, no. 5717: 1904-9.

Ragnhammar, P., L. Hafstrom, P. Nygren, and B. Glimelius. 2001. A systematic overview of chemotherapy effects in colorectal cancer. *Acta Oncol* 40, no. 2-3: 282-308.

Rees, J. R., B. A. Onwuegbusi, V. E. Save, D. Alderson, and R. C. Fitzgerald. 2006. In vivo and in vitro evidence for transforming growth factor-beta1-mediated epithelial to mesenchymal transition in esophageal adenocarcinoma. *Cancer Res* 66, no. 19: 9583-90.

Reinacher-Schick, A., S. E. Baldus, B. Romdhana, S. Landsberg, M. Zapatka, S. P. Monig, A. H. Holscher, H. P. Dienes, W. Schmiegel, and I. Schwarte-Waldhoff. 2004. Loss of smad4 correlates with loss of the invasion suppressor e-cadherin in advanced colorectal carcinomas. *J Pathol* 202, no. 4: 412-20.

Reya, T. and H. Clevers. 2005. Wnt signalling in stem cells and cancer. *Nature* 434, no. 7035: 843-50.

Rhodes, D. R., S. Kalyana-Sundaram, V. Mahavisno, T. R. Barrette, D. Ghosh, and A. M. Chinnaiyan. 2005. Mining for regulatory programs in the cancer transcriptome. *Nat Genet* 37, no. 6: 579-83.

Riggins, G. J., K. W. Kinzler, B. Vogelstein, and S. Thiagalingam. 1997. Frequency of smad gene mutations in human cancers. *Cancer Res* 57, no. 13: 2578-80.

Riggins, G. J., S. Thiagalingam, E. Rozenblum, C. L. Weinstein, S. E. Kern, S. R. Hamilton, J. K. Willson, S. D. Markowitz, K. W. Kinzler, and B. Vogelstein. 1996. Mad-related genes in the human. *Nat Genet* 13, no. 3: 347-9.

Rijsewijk, F., M. Schuermann, E. Wagenaar, P. Parren, D. Weigel, and R. Nusse. 1987. The drosophila homolog of the mouse mammary oncogene int-1 is identical to the segment polarity gene wingless. *Cell* 50, no. 4: 649-57.

Rodrigues, N. R., A. Rowan, M. E. Smith, I. B. Kerr, W. F. Bodmer, J. V. Gannon, and D. P. Lane. 1990. P53 mutations in colorectal cancer. *Proc Natl Acad Sci U S A* 87, no. 19: 7555-9.

Rowan, A. J., H. Lamlum, M. Ilyas, J. Wheeler, J. Straub, A. Papadopoulou, D. Bicknell, W. F. Bodmer, and I. P. Tomlinson. 2000. Apc mutations in sporadic colorectal tumors: A mutational "Hotspot" And interdependence of the "Two hits". *Proc Natl Acad Sci U S A* 97, no. 7: 3352-7.

Saltz, L. B., N. J. Meropol, P. J. Loehrer, Sr., M. N. Needle, J. Kopit, and R. J. Mayer. 2004. Phase ii trial of cetuximab in patients with refractory colorectal cancer that expresses the epidermal growth factor receptor. *J Clin Oncol* 22, no. 7: 1201-8.

Samowitz, W. S., K. Curtin, K. N. Ma, D. Schaffer, L. W. Coleman, M. Leppert, and M. L. Slattery. 2001. Microsatellite instability in sporadic colon cancer is associated with an improved prognosis at the population level. *Cancer Epidemiol Biomarkers Prev* 10, no. 9: 917-23.

Sansom, O. J., K. R. Reed, A. J. Hayes, H. Ireland, H. Brinkmann, I. P. Newton, E. Batlle, P. Simon-Assmann, H. Clevers, I. S. Nathke, A. R. Clarke, and D. J. Winton. 2004. Loss of apc in vivo immediately perturbs wnt signaling, differentiation, and migration. *Genes Dev* 18, no. 12: 1385-90.

Sayed, M. G., A. F. Ahmed, J. R. Ringold, M. E. Anderson, J. L. Bair, F. A. Mitros, H. T. Lynch, S. T. Tinley, G. M. Petersen, F. M. Giardiello, B. Vogelstein, and J. R. Howe. 2002. Germline smad4 or bmpr1a mutations and phenotype of juvenile polyposis. *Ann Surg Oncol* 9, no. 9: 901-6.

Schefe, J. H., K. E. Lehmann, I. R. Buschmann, T. Unger, and H. Funke-Kaiser. 2006. Quantitative real-time rt-pcr data analysis: Current concepts and the novel "Gene expression's ct difference" Formula. *J Mol Med* 84, no. 11: 901-10.

Schmalhofer, O., S. Brabletz, and T. Brabletz. 2009. E-cadherin, beta-catenin, and zeb1 in malignant progression of cancer. *Cancer Metastasis Rev* 28, no. 1-2: 151-66.

Schmidt, C. R., Y. J. Gi, T. A. Patel, R. J. Coffey, R. D. Beauchamp, and A. S. Pearson. 2005. E-cadherin is regulated by the transcriptional repressor slug during ras-mediated transformation of intestinal epithelial cells. *Surgery* 138, no. 2: 306-12.

Schwarte-Waldhoff, I., S. Klein, S. Blass-Kampmann, A. Hintelmann, C. Eilert, S. Dreschers, H. Kalthoff, S. A. Hahn, and W. Schmiegel. 1999. Dpc4/smad4 mediated tumor suppression of colon carcinoma cells is associated with reduced urokinase expression. *Oncogene* 18, no. 20: 3152-8.

Shedden, K., J. M. Taylor, S. A. Enkemann, M. S. Tsao, T. J. Yeatman, W. L. Gerald, S. Eschrich, I. Jurisica, T. J. Giordano, D. E. Misek, A. C. Chang, C. Q. Zhu, D. Strumpf, S. Hanash, F. A. Shepherd, K. Ding, L. Seymour, K. Naoki, N. Pennell, B. Weir, R. Verhaak, C. Ladd-Acosta, T. Golub, M. Gruidl, A. Sharma, J. Szoke, M. Zakowski, V. Rusch, M. Kris, A. Viale, N. Motoi, W. Travis, B. Conley, V. E. Seshan, M. Meyerson, R. Kuick, K. K. Dobbin, T. Lively, J. W. Jacobson, and D. G. Beer. 2008. Gene expression-based survival prediction in lung adenocarcinoma: A multi-site, blinded validation study. *Nat Med* 14, no. 8: 822-7.

Shi, Y. and J. Massague. 2003. Mechanisms of tgf-beta signaling from cell membrane to the nucleus. *Cell* 113, no. 6: 685-700.

Shiou, S. R., A. B. Singh, K. Moorthy, P. K. Datta, M. K. Washington, R. D. Beauchamp, and P. Dhawan. 2007. Smad4 regulates claudin-1 expression in a transforming growth factor-beta-independent manner in colon cancer cells. *Cancer Res* 67, no. 4: 1571-9.

Smith, J. J., N. G. Deane, F. Wu, N. B. Merchant, B. Zhang, A. Jiang, P. Lu, J. C. Johnson, C. Schmidt, C. E. Bailey, S. Eschrich, C. Kis, S. Levy, M. K. Washington, M. J. Heslin, R. J. Coffey, T. J. Yeatman, Y. Shyr, and R. D. Beauchamp. Experimentally derived metastasis gene expression profile predicts recurrence and death in patients with colon cancer. *Gastroenterology* 138, no. 3: 958-68.

Smyth, G. K. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol* 3: Article3.

Smyth, G. K., J. Michaud, and H. S. Scott. 2005. Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics* 21, no. 9: 2067-75.

Spaderna, S., O. Schmalhofer, F. Hlubek, G. Berx, A. Eger, S. Merkel, A. Jung, T. Kirchner, and T. Brabletz. 2006. A transient, emt-linked loss of basement membranes indicates metastasis and poor survival in colorectal cancer. *Gastroenterology* 131, no. 3: 830-40.

Su, L. K., K. W. Kinzler, B. Vogelstein, A. C. Preisinger, A. R. Moser, C. Luongo, K. A. Gould, and W. F. Dove. 1992. Multiple intestinal neoplasia caused by a mutation in the murine homolog of the apc gene. *Science* 256, no. 5057: 668-70.

Subramaniam, N., G. M. Leong, T. A. Cock, J. L. Flanagan, C. Fong, J. A. Eisman, and A. P. Kouzmenko. 2001. Cross-talk between 1,25-dihydroxyvitamin d3 and transforming growth factor-beta signaling requires binding of vdr and smad3 proteins to their cognate DNA recognition elements. *J Biol Chem* 276, no. 19: 15741-6.

Sun, L., G. Wu, J. K. Willson, E. Zborowska, J. Yang, I. Rajkarunanayake, J. Wang, L. E. Gentry, X. F. Wang, and M. G. Brattain. 1994. Expression of transforming growth factor beta type ii receptor leads to reduced malignancy in human breast cancer mcf-7 cells. *J Biol Chem* 269, no. 42: 26449-55.

Takagi, Y., H. Kohmura, M. Futamura, H. Kida, H. Tanemura, K. Shimokawa, and S. Saji. 1996. Somatic alterations of the dpc4 gene in human colorectal cancers in vivo. *Gastroenterology* 111, no. 5: 1369-72.

Takahashi, E., O. Nagano, T. Ishimoto, T. Yae, Y. Suzuki, T. Shinoda, S. Nakamura, S. Niwa, S. Ikeda, H. Koga, H. Tanihara, and H. Saya. Tumor necrosis factor-alpha regulates transforming growth factor-beta-dependent epithelial-mesenchymal transition by promoting hyaluronan-cd44-moesin interaction. *J Biol Chem* 285, no. 6: 4060-73.

Takahashi, M., M. Fujita, Y. Furukawa, R. Hamamoto, T. Shimokawa, N. Miwa, M. Ogawa, and Y. Nakamura. 2002. Isolation of a novel human gene, apcdd1, as a direct target of the beta-catenin/t-cell factor 4 complex with probable involvement in colorectal carcinogenesis. *Cancer Res* 62, no. 20: 5651-6.

Takaku, K., M. Oshima, H. Miyoshi, M. Matsui, M. F. Seldin, and M. M. Taketo. 1998. Intestinal tumorigenesis in compound mutant mice of both dpc4 (smad4) and apc genes. *Cell* 92, no. 5: 645-56.

Taketo, M. M. and W. Edelmann. 2009. Mouse models of colon cancer. *Gastroenterology* 136, no. 3: 780-98.

Tanaka, T., T. Watanabe, Y. Kazama, J. Tanaka, T. Kanazawa, S. Kazama, and H. Nagawa. 2008. Loss of smad4 protein expression and 18qloh as molecular markers indicating lymph node metastasis in colorectal cancer--a study matched for tumor depth and pathology. *J Surg Oncol* 97, no. 1: 69-73.

ten Berge, D., W. Koole, C. Fuerer, M. Fish, E. Eroglu, and R. Nusse. 2008. Wnt signaling mediates self-organization and axis formation in embryoid bodies. *Cell Stem Cell* 3, no. 5: 508-18.

Tetsu, O. and F. McCormick. 1999. Beta-catenin regulates expression of cyclin d1 in colon carcinoma cells. *Nature* 398, no. 6726: 422-6.

Thiagalingam, S., C. Lengauer, F. S. Leach, M. Schutte, S. A. Hahn, J. Overhauser, J. K. Willson, S. Markowitz, S. R. Hamilton, S. E. Kern, K. W. Kinzler, and B. Vogelstein. 1996. Evaluation of candidate tumour suppressor genes on chromosome 18 in colorectal cancers. *Nat Genet* 13, no. 3: 343-6.

Thibodeau, S. N., G. Bren, and D. Schaid. 1993. Microsatellite instability in cancer of the proximal colon. *Science* 260, no. 5109: 816-9.

Thiery, J. P. 2002. Epithelial-mesenchymal transitions in tumour progression. *Nat Rev Cancer* 2, no. 6: 442-54.

_____. 2003. Epithelial-mesenchymal transitions in development and pathologies. *Curr Opin Cell Biol* 15, no. 6: 740-6.

Tian, X., H. Du, X. Fu, K. Li, A. Li, and Y. Zhang. 2009. Smad4 restoration leads to a suppression of wnt/beta-catenin signaling activity and migration capacity in human colon carcinoma cells. *Biochem Biophys Res Commun* 380, no. 3: 478-83.

Tukey, J. W. 1993. Tightening the clinical trial. *Control Clin Trials* 14, no. 4: 266-85.

Valcourt, U., M. Kowanetz, H. Niimi, C. H. Heldin, and A. Moustakas. 2005. Tgf-beta and the smad signaling pathway support transcriptomic reprogramming during epithelial-mesenchymal cell transition. *Mol Biol Cell* 16, no. 4: 1987-2002.

van de Wetering, M., E. Sancho, C. Verweij, W. de Lau, I. Oving, A. Hurlstone, K. van der Horn, E. Batlle, D. Coudreuse, A. P. Haramis, M. Tjon-Pon-Fong, P. Moerer, M. van den Born, G. Soete, S. Pals, M. Eilers, R. Medema, and H. Clevers. 2002. The beta-catenin/tcf-4 complex imposes a crypt progenitor phenotype on colorectal cancer cells. *Cell* 111, no. 2: 241-50.

Vazquez, A., E. E. Bond, A. J. Levine, and G. L. Bond. 2008. The genetics of the p53 pathway, apoptosis and cancer therapy. *Nat Rev Drug Discov* 7, no. 12: 979-87.

Veeman, M. T., J. D. Axelrod, and R. T. Moon. 2003. A second canon. Functions and mechanisms of beta-catenin-independent wnt signaling. *Dev Cell* 5, no. 3: 367-77.

Wang, D. and R. N. DuBois. 2008. Pro-inflammatory prostaglandins and progression of colorectal cancer. *Cancer Lett* 267, no. 2: 197-203.

Wang, Y., T. Jatkoe, Y. Zhang, M. G. Mutch, D. Talantov, J. Jiang, H. L. McLeod, and D. Atkins. 2004. Gene expression profiles and molecular markers to predict recurrence of dukes' b colon cancer. *J Clin Oncol* 22, no. 9: 1564-71.

Watanabe, T., T. T. Wu, P. J. Catalano, T. Ueki, R. Satriano, D. G. Haller, A. B. Benson, 3rd, and S. R. Hamilton. 2001. Molecular predictors of survival after adjuvant chemotherapy for colon cancer. *N Engl J Med* 344, no. 16: 1196-206.

Whitehead, R. H. 1976. The culture of tumour cells from human tumour biopsies. *Clin Oncol* 2, no. 2: 131-40.

Whitman, M. 1998. Smads and early developmental signaling by the tgfbeta superfamily. *Genes Dev* 12, no. 16: 2445-62.

Winesett, M. P., G. W. Ramsey, and J. A. Barnard. 1996. Type ii tgf(beta) receptor expression in intestinal cell lines and in the intestinal tract. *Carcinogenesis* 17, no. 5: 989-95.

Wong, R. and D. Cunningham. 2008. Using predictive biomarkers to select patients with advanced colorectal cancer for treatment with epidermal growth factor receptor antibodies. *J Clin Oncol* 26, no. 35: 5668-70.

Woodford-Richens, K. L., A. J. Rowan, P. Gorman, S. Halford, D. C. Bicknell, H. S. Wasan, R. R. Roylance, W. F. Bodmer, and I. P. Tomlinson. 2001. Smad4 mutations in colorectal cancer probably occur before chromosomal instability, but after divergence of the microsatellite instability pathway. *Proc Natl Acad Sci U S A* 98, no. 17: 9719-23.

Wu, J. C., G. Sundaresan, M. Iyer, and S. S. Gambhir. 2001. Noninvasive optical imaging of firefly luciferase reporter gene expression in skeletal muscles of living mice. *Mol Ther* 4, no. 4: 297-306.

Yan, M., R. M. Rerko, P. Platzer, D. Dawson, J. Willis, M. Tong, E. Lawrence, J. Lutterbaugh, S. Lu, J. K. Willson, G. Luo, J. Hensold, H. H. Tai, K. Wilson, and S. D. Markowitz. 2004. 15-hydroxyprostaglandin dehydrogenase, a cox-2 oncogene antagonist, is a tgf-beta-induced suppressor of human gastrointestinal cancers. *Proc Natl Acad Sci U S A* 101, no. 50: 17468-73.

Yanagisawa, K., Y. Shyr, B. J. Xu, P. P. Massion, P. H. Larsen, B. C. White, J. R. Roberts, M. Edgerton, A. Gonzalez, S. Nadaf, J. H. Moore, R. M. Caprioli, and D. P. Carbone. 2003. Proteomic patterns of tumour subsets in non-small-cell lung cancer. *Lancet* 362, no. 9382: 433-9.

Yang, B., S. M. O'Herrin, J. Wu, S. Reagan-Shaw, Y. Ma, K. M. Bhat, C. Gravekamp, V. Setaluri, N. Peters, F. M. Hoffmann, H. Peng, A. V. Ivanov, A. J. Simpson, and B. J. Longley. 2007. Mage-a, mmage-b, and mage-c proteins form complexes with kap1 and suppress p53-dependent apoptosis in mage-positive cell lines. *Cancer Res* 67, no. 20: 9954-62.

Yang, J., S. A. Mani, J. L. Donaher, S. Ramaswamy, R. A. Itzykson, C. Come, P. Savagner, I. Gitelman, A. Richardson, and R. A. Weinberg. 2004. Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* 117, no. 7: 927-39.

Yang, J. and R. A. Weinberg. 2008. Epithelial-mesenchymal transition: At the crossroads of development and tumor metastasis. *Dev Cell* 14, no. 6: 818-29.

Yang, J., W. Zhang, P. M. Evans, X. Chen, X. He, and C. Liu. 2006. Adenomatous polyposis coli (apc) differentially regulates beta-catenin phosphorylation and ubiquitination in colon cancer cells. *J Biol Chem* 281, no. 26: 17751-7.

Yang, Q., N. A. Bermingham, M. J. Finegold, and H. Y. Zoghbi. 2001. Requirement of math1 for secretory cell lineage commitment in the mouse intestine. *Science* 294, no. 5549: 2155-8.

Yao, M., G. Niu, Z. Sheng, Z. Wang, and J. Fei. Identification of a smad4/yy1-recognizedand bmp2-responsive transcriptional regulatorymodule in the promoter of mouse gaba transporter subtype i (gat1) gene. *J Neurosci* 30, no. 11: 4062-71.

Yeung, T. M. and N. J. Mortensen. 2009. Colorectal cancer stem cells. *Dis Colon Rectum* 52, no. 10: 1788-96.

Zeilstra, J., S. P. Joosten, M. Dokter, E. Verwiel, M. Spaargaren, and S. T. Pals. 2008. Deletion of the wnt target and cancer stem cell marker cd44 in apc(min/+) mice attenuates intestinal tumorigenesis. *Cancer Res* 68, no. 10: 3655-61.

Zeisberg, M., J. Hanai, H. Sugimoto, T. Mammoto, D. Charytan, F. Strutz, and R. Kalluri. 2003. Bmp-7 counteracts tgf-beta1-induced epithelial-to-mesenchymal transition and reverses chronic renal injury. *Nat Med* 9, no. 7: 964-8.

Zhang, B., S. Kirov, and J. Snoddy. 2005. Webgestalt: An integrated system for exploring gene sets in various biological contexts. *Nucleic Acids Res* 33, no. Web Server issue: W741-8.

Zhang, T., L. B. Nanney, M. O. Peeler, C. S. Williams, L. Lamps, K. J. Heppner, R. N. DuBois, and R. D. Beauchamp. 1997. Decreased transforming growth factor beta type ii receptor expression in intestinal adenomas from min/+ mice is associated with increased cyclin d1 and cyclin-dependent kinase 4 expression. *Cancer Res* 57, no. 9: 1638-43.

Zhang, Y., X. Feng, R. We, and R. Derynck. 1996. Receptor-associated mad homologues synergize as effectors of the tgf-beta response. *Nature* 383, no. 6596: 168-72.