Wnt Signaling in Stem Cells and Cancer

Yaser Atlasi

Wnt Signaling in Stem Cells and Cancer

Wnt Signalering in Stamcellen en Kanker

Thesis

to obtain the degree of Doctor from the
Erasmus University Rotterdam
by command of the
rector magnificus

Prof.dr. H.A.P. Pols

and in accordance with the decision of the Doctorate Board

The public defense shall be held on
Wednesday 27 November 2013 at 9:30 hours

by

Yaser Atlasi
born in Baghdad, Iraq

Doctoral committee:

Prof.dr. R. Fodde

Promoter

Other members

Prof.dr. J. Gribnau
Prof.dr. L.H.J. Looijenga
Prof.dr. C.P. Verrijzer

ISBN: 978-94-6191-973-1

Cover: Yaser Atlasi
Layout: Legatron Electronic Publishing
Printing: Ipskamp Drukkers

Cover illustrates the puzzle of embryonic stem cells and cancer. Two mouse embryonic stem cell colonies are depicted in the picture.

The work presented in this thesis was performed in the Department of Pathology, Josephine Nefkens Institute, Erasmus MC-University Medical Center, Rotterdam, The Netherlands.

The printing of this thesis was financially supported by:
Erasmus University Rotterdam
Josephine Nefkens Institute
BD Bioscience
Greiner Bio One
Promega

"The more you know,
the more you realize
that you know nothing"
(Socrates)

To my parents: Farid and Farah

CONTENTS

GENERAL INTRODUCTION AND OUTLINE OF THE THESIS

Mammalian development starts from a fertilized egg that initially generates few pluripotent cells which eventually give rise to the embryo proper. Different 'flavors' of pluripotency have been captured *in vitro* which led to the establishment of different pluripotent cell lines. Mouse embryonic stem cells (mESCs) are derived from the pre-implantation embryo and have three defining properties: self-renewal, pluripotency, and contribution to chimera formation. By applying specific culture conditions or ectopic expression of the pluripotency factors, similar pluripotent cells can be derived from germ cells or differentiated cells referred to as embryonic germ (EG) and induced pluripotent cells (iPSCs), respectively. When established from post-implantation embryo, the cultured cells are termed epi stem cells (EpiSCs). EpiSCs have limited potential for chimerism and germ line transmission and require different culture conditions when compared to ESCs. Hence, mouse ESCs and EpiSCs represent two different phases of pluripotency usually referred to as the naïve and primed states. By employing genetic manipulation or specific culture conditions, the different pluripotent cells can be interconverted which leads to several intermediate states. Unlike their murine counterparts, human ESCs closely resemble the rodent primed EpiSCs and respond to similar signaling pathways. Tumorigenic transformation of primordial germ cells (PGC) and gonocytes can also give rise to pluripotent cells known as embryonal carcinoma cells (hECs), thought to represent the malignant counterpart of hESCs. Among different signaling pathways, Wnt signaling plays a central role in self-renewal and differentiation of pluripotent cells.

In **Chapter 1**, we outline the current literature on Wnt signaling, as a major stemness pathway, and establish a parallel between its role in embryonic and cancer stem cells. We also discuss germ cell tumors (GCTs) as a model to study the role of stem cells in neoplasia. In ***Chapter 2 and 3***, we employ a series of *Apc*-mutant ESCs encoding for different Wnt signaling levels, to study the mechanisms through which this pathway regulates lineage differentiation in mouse ESCs.

In **Chapter 2,** we investigate Tcf3 downregulation downstream of Wnt signaling and examine how this effect is implicated in lineage commitment of mESCs.

In **Chapter 3** we report that miR-302-367, one of the highly specific miRNAs in mouse ES cells, is negatively regulated by Wnt signaling. We further investigate the impact of this epistatic interaction on ES self-renewal and differentiation.

In **Chapter 4**, we explore the functional role of Wnt signaling in human embryonal carcinoma cells and provide evidence that WNT induction induces a range of diverse differentiation responses among the different employed hEC lines, thus underscoring the context-dependent role of Wnt signaling in mouse and human pluripotent stem cells.

It is estimated that human genome encodes 20000-25000 genes which constitute approx. 2% of the DNA. However, more than 90% of the human genome is transcribed into different RNA species. microRNAs are the smallest non-coding RNAs and are involved in the regulation of gene expression. Highly conserved genomic regions (ultra-conserved regions) are also frequently transcribed as long non coding-RNAs (>200 bp) in both normal and tumor cells. In **Chapter 5** we describe the identification and functional characterization of CCAT2, a novel WNT-regulated long noncoding RNA located on

chromosome 8q24. We show that this non-coding RNA is expressed in microsatellite-stable colorectal cancers where it promotes tumor growth, metastasis, and chromosomal instability.

The interaction between tumor cells and their niche plays a pivotal role in tumor initiation, progression and metastasis. In **Chapter 6,** we studied the role of *S100a4*, a Wnt target gene that is expressed from within the tumor microenvironment in intestinal neoplasia. By using different transgenic and knockout mouse models, we examined the *in vivo* role of *S100a4* in intestinal tumor initiation and progression. Moreover, we identified a novel role for *S100a4* in the etiology of desmoid disease. Desmoids are mesenchymal tumors associated with constitutive activation of Wnt signaling due to *APC* or *CTNNB1* mutations and are believed to arise in sites of wound healing. In ***Chapter 6*** we show that specific depletion of S100a4 in $Apc^{1638N/+}$ mice significantly reduces desmoids formation.

chapter ONE

CANCER STEM CELLS, PLURIPOTENCY, AND CELLULAR HETEROGENEITY: A WNTer PERSPECTIVE

Yaser Atlasi, Leendert Looijenga, and Riccardo Fodde*

Dept. of Pathology, Erasmus MC, Rotterdam, The Netherlands

Current Topics in Developmental Biology: in press

chapter TWO

WNT SIGNALING REGULATES THE LINEAGE DIFFERENTIATION POTENTIAL OF MOUSE EMBRYONIC STEM CELLS THROUGH TCF3 DOWN-REGULATION

Yaser Atlasi[1], Rubina Noori[1], Claudia Gaspar[1], Patrick Franken[1], Andrea Sacchetti[1], Haleh Rafati[2], Tokameh Mahmoudi[2], Charles Decraene[3,4], George A. Calin[5], Bradley J. Merrill[6] and Riccardo Fodde[1]

[1]Dept. of Pathology, Josephine Nefkens Institute, Erasmus MC, Rotterdam, The Netherlands; [2]Dept. of Biochemistry, Erasmus MC, Rotterdam, The Netherlands; [3]Translational Research Dept, Institut Curie, Centre de Recherche, Paris, France; [4]CNRS, UMR144, Paris, France; [5]Department of Experimental Therapeutics and Center for RNA Interference and non-coding RNAs, MD Anderson Cancer Center, Houston, TX, USA; [6]Dept. of Biochemistry and Molecular Genetics, University of Illinois, Chicago, IL, USA.

PLoS Genet. 2013 May;9(5):e1003424.

ABSTRACT

Canonical Wnt signaling plays a rate-limiting role in regulating self-renewal and differentiation in mouse embryonic stem cells (ESCs). We have previously shown that mutation in the Apc (adenomatous polyposis coli) tumor suppressor gene constitutively activates Wnt signaling in ESCs and inhibits their capacity to differentiate towards ecto-, meso- and endodermal lineages. However, the underlying molecular and cellular mechanisms through which Wnt regulates lineage differentiation in mouse ESCs remains to date largely unknown. To this aim, we have derived and studied the gene expression profiles of several Apc-mutant ESC lines encoding for different levels of Wnt signaling activation. We found that down-regulation of *Tcf3*, a member of the Tcf/Lef family and a key player in the control of self-renewal and pluripotency, represents a specific and primary response to Wnt activation in ESCs. Accordingly, rescuing *Tcf3* expression partially restored the neural defects observed in Apc-mutant ESCs, suggesting that *Tcf3* down-regulation is a necessary step towards Wnt-mediated suppression of neural differentiation. We found that *Tcf3* down-regulation in the context of constitutively active Wnt signaling does not result from promoter DNA methylation but is likely to be caused by a plethora of mechanisms both at the RNA and protein level as shown by the observed decrease in activating histone marks (H3K4me3 and H3-acetylation) and the upregulation of miR-211, a novel Wnt-regulated microRNA that targets *Tcf3* and attenuates early neural differentiation in mouse ESCs.

Our data show for the first time that Wnt signaling down-regulates *Tcf3* expression, possibly at both the transcriptional and post-transcriptional levels, and thus highlight a novel mechanism through which Wnt signaling inhibits neuro-ectodermal lineage differentiation in mouse embryonic stem cells.

AUTHOR SUMMARY

The future successes of regenerative medicine largely rely on our knowledge of, and our capacity to manipulate the cellular and molecular mechanisms governing stem cell differentiation. Growing body of evidence suggests that in mouse embryonic stem cells, canonical Wnt/β-catenin signaling not only enhances self-renewal but also directs the cell fate decision towards non-neuroectodermal lineages. However, little is known about the mechanisms underlying the differentiation defects caused by constitutive active Wnt signaling. Using a set of *Apc*-mutant ESCs harbouring different levels of Wnt signaling, we found that, among others, down-regulation of *Tcf3*, a key member of the pluripotency circuit, as well as induction of a novel Wnt-regulated microRNA, miR-211, represent two important downstream effects through which Wnt signaling inhibits neural differentiation in mouse ESCs. We also provide a more detailed picture on how Wnt signaling counteracts Tcf3 function in stem cells by showing that Tcf3 repression, in the context of active Wnt signaling, involves histone modifications at the *Tcf3* promoter and the activation of miR-211 which post-transcriptionally stabilizes Tcf3 downregulation. Understanding the downstream effects of Wnt signaling in ESCs is both of fundamental and translational relevance as it may be exploited to manipulate ESC differentiation towards specific cell lineages.

INTRODUCTION

Embryonic stem cells (ESCs) are *in vitro* cultured cells derived from the preimplantation-stage embryo, which possess unconfined capacity for self-renewal and multi-lineage differentiation towards different embryonic germ layers. Pluripotency and self-renewal are two essential features of ESCs, which make them not only a very robust and suitable model for stem cell research, but also a promising source for regenerative medicine. Also, with the emergence of induced pluripotent stem cells (iPS) technology, understanding the basic mechanisms governing the embryonic stem state becomes of great interest for safe clinical applications in regenerative medicine and stem cell programming.

Among different signaling pathways, Wnt/β-catenin signaling has been shown to play a major role in maintaining self-renewal as well as in regulating ESCs differentiation [1,2,3,4,5,6]. The canonical Wnt/β-catenin signaling pathway is controlled by post-translational modifications of β-catenin leading to its differential protein stability and sub-cellular localization. In the absence of active Wnt signaling, β-catenin is negatively regulated by the so-called "destruction complex", consisting of the Apc and Axin scaffolding proteins and the glycogen synthase and casein kinases (GSK and CK1), resulting in proteolytic degradation and low levels of cytoplasmic β-catenin. Ligand-mediated Wnt signaling activation leads to nuclear translocation of β-catenin where it binds to members of the Tcf/Lef family of transcriptional factors thus modulating the expression of a broad spectrum of downstream target genes [7,8,9].

In vertebrates, the Tcf/Lef family encompasses four functionally specialized members including Tcf1 (also known as Tcf7), Tcf3 (also known as Tcf7l1), Tcf4 (also known as Tcf7l2) and Lef1 [10]. Whereas Tcf1, Tcf4 and Lef1 are known to activate different Wnt target genes in the context of active Wnt signaling, Tcf3 primarily functions as a transcriptional repressor [5,11,12,13,14,15,16]. Tcf3 is the most abundant Tcf/Lef member in mouse ES cells [14] and is an integral component of the core pluripotency circuit, co-occupying Oct4, Nanog and Sox2 DNA binding sites [17,18,19,20]. Loss of function experiments have shown that Tcf3 down-regulation enhances self-renewal and confers differentiation resistance in mouse ESCs [14,17,19,20,21,22]. In fact, both the zebrafish *headless/tcf3* mutant and the *Xenopus* embryo depleted of TCF3 reveal anterior head defects resembling the Wnt-gain of function phenotype [11,15,16]. Similarly, Tcf3 ablation in mice resulted in expanded axial mesoderm and loss of anterior neural tissues [21]. *Tcf3* is ubiquitously expressed through the mouse embryo at embryonic day 6.5 (E6.5) and is gradually localized in the anterior part of the embryo at E7.5 and the anterior neuroectoderm at E8.5 [23,24].

Although several studies have demonstrated the key role played by Wnt signaling in regulating self-renewal and differentiation of both mouse and human ESCs, the downstream effects through which Wnt exerts these functions have been a matter of controversy. To date, three models have been suggested in this regard: a. Tcf-independent, β-catenin/Oct4 signaling [25]; b. Tcf3 antagonism by nuclear β-catenin which relieves Tcf3 repression and enhances self-renewal. A minimal role for the canonical Tcf/β-catenin signaling has been suggested in this model [6]; and c. synergistic action of Tcf3 antagonism and the canonical β-catenin/Tcf1 signaling [5]. Although these studies have shed some light on the underlying mechanisms through which Wnt signaling controls self-renewal, none of the above-mentioned models

explains how this signaling pathway regulates the lineage differentiation potential of ESCs.

In order to elucidate the downstream effects of Wnt signaling on lineage commitment and differentiation in embryonic stem cells, we examined several *Apc*-mutant ESCs harboring different levels of Wnt signaling and compared their gene expression profiles with wild type ESCs. We show that activation of Wnt signaling down-regulates Tcf3 expression in mouse ESCs. We provide evidence that *Tcf3* down-regulation represents a main downstream effect through which Wnt signaling directs the differentiation of pluripotent ESCs towards non-neuroectodermal lineages. Moreover, we show that Wnt-mediated repression of *Tcf3* involves epigenetic regulation associated with histone modifications and Wnt-mediated induction of miR-211. Our data demonstrate that Wnt signaling counteracts Tcf3 function at multiple levels, which ultimately ensures the delicate balance between self-renewal and differentiation in mouse ESCs.

RESULTS

LINEAGE DIFFERENTIATION IN APC-MUTANT ESCS CORRELATES WITH THE LEVEL OF WNT SIGNALING

To attempt the elucidation of the mechanisms underlying lineage differentiation in the context of Wnt activation, we have derived several ES clones from pre-implantation blastocysts carrying different hypomorphic *Apc* alleles: $Apc^{1638T/1638T}$ (*Apc*TT), $Apc^{1638N/1638T}$ (*Apc*NT), $Apc^{1638N/1638N}$ (*Apc*NN) [26,27], together with $Apc^{+/+}$ as wild type controls. As previously reported, *Apc*TT, *Apc*NT, and *Apc*NN encode for a gradient of different Wnt signaling dosages [1,26], as also confirmed by TOP-Flash reporter assay [28] with *Apc*NN showing the highest Wnt activity (*Apc*NN>>*Apc*NT>*Apc*TT> $Apc^{+/+}$) (Figure 1A). The potential of the *Apc*-mutant ES cells to differentiate into ecto-, meso- and endodermal lineages was also evaluated and confirmed by the teratoma formation assay followed by immunohistochemistry (IHC) analysis, matching our previous results obtained with ES clones obtained by two rounds of gene targeting by homologous recombination [1]. As expected, no expression of neuroectodermal markers (GFAP, SV2, and neurofilaments) was observed in teratomas derived from *Apc*NN ES cells (Figure 1B).

ES cells can be cultured in serum-free medium supplemented with LIF, GSK inhibitor (CHIRON) and Mek inhibitor (PD), the so-called 2i medium [29]. Using the serum-free culture supplemented with a single inhibitor, we found that *Apc*NN cells have the highest colony-forming capacity when cultured in LIF + Mek inhibitor, suggesting that their constitutive Wnt signaling activity replaces the need for additional pathway activation by the GSK inhibitor (Figure 1C). Of note, culturing *Apc*NN ESCs in medium supplemented with CHIRON reduced the colony formation capacity of these cells suggesting that a very high dosage of Wnt signaling can compromise the growth of *Apc*NN cells. We also observed that *Apc*TT and *Apc*NT cells formed similar number of colonies in different culture conditions independently of CHIRON supplementation, possibly pointing to the Wnt-independent effects of *Apc* mutations in these cells (Figure 1C).

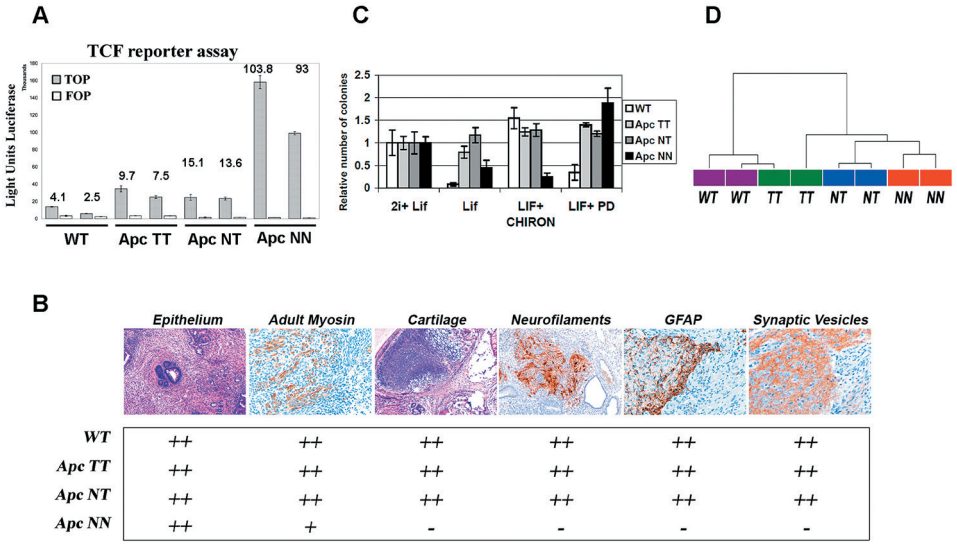


Figure 1. Wnt signaling regulates the differentiation potential of mouse ESCs in a dosage dependent manner.
(**A**) b-catenin/TCF reporter assay in wild type and *Apc*-mutant ESCs. Measurements are reported as the average luciferase units performed in triplicate for the TOP (filled bars) and FOP (empty bars) reporter constructs (data reported is mean ± SD). Numbers in the histogram represent the calculated TOP/FOP ratios; (**B**) Table summarizing the results obtained by teratoma differentiation assay from different *Apc*-mutant ESCs and their wild type controls. Tissue sections were stained with hematoxylin and eosin (H&E) or used in immunohistochemical analysis using specific antibodies against the neural markers: GFAP, neurofilaments and synaptic vesicles. Adult myosin was used as a mesodermal marker to stain the striated muscle differentiation. Cartilage differentiation was assessed either by H&E or theonin staining. Two independent clones were used for each genotype and differentiation was scored as: (-) not present, (+) weakly present, and (++) present; (**C**) Histogram showing the percent of colonies formed after plating 500 FACS-sorted cells in N2B27 medium supplemented with different combinations of LIF, Mek inhibitor (PD) and GSK-inhibitor (CHIRON). Bars represent mean ± SD, n=3; (**D**) Dendrogram derived from unsupervised hierarchical clustering of global gene expression in wild type, *Apc*TT, *Apc*NT and *Apc*NN ES cells. Pearson's correlation coefficient and Ward's method were used after MAS 5.0 normalization of all probe sets.

WNT SIGNALING DOWN-REGULATES *TCF3* EXPRESSION IN MOUSE ESCS

To elucidate the molecular mechanisms underlying the altered cell fate decision in *Apc*-mutant ES cells, genome-wide transcriptional analysis was performed on the newly derived cells. Unsupervised hierarchical clustering analysis showed that global gene expression in *Apc*NN ESCs is already influenced before differentiation is induced, resolving *Apc*NN from WT expression profiles in different branches of the dendogram (Figure 1D). Among the genes differentially expressed between *Apc*NN ES cells and their wild type counterparts (Table S1), we found that, unlike other pluripotency markers (e.g. *Oct4*, *Nanog*, and *Sox2*), *Tcf3* was specifically down-regulated in *Apc*NN ES cells, an observation which was further confirmed by qRT-PCR and western blot analysis (Figure 2A and B; and Figure S1). Further

qRT-PCR analysis revealed that the observed downregulation is specific for *Tcf3* but not for other members of the Tcf/Lef family (Figure S2A). Whereas Tcf3 was down-regulated in both *Apc*NN and *Apc*$^{Min/Min}$ ESCs, the latter encode for the most severely truncated Apc mutant allele and therefore for a very high level of Wnt signaling, other members of the Tcf/Lef family were exclusively up-regulated in *Apc*$^{Min/Min}$ ESCs.

Accordingly, Wnt activation achieved in wild type cells either by Wnt3a conditioned medium or by a GSK3-small molecule inhibitor (SB-216763), confirmed that *Tcf3* down-regulation is a specific response to canonical Wnt signaling in mouse ESCs (Figure 2C, D, and E, and Figure S2B and C). Moreover, using a gradient of the GSK inhibitor SB-216763, we observed that unlike the canonical Wnt targets *Axin2* and *Cdx1*, downregulation of *Tcf3* required a higher Wnt signaling level, possibly explaining why *Tcf3* downregulation is observed in *Apc*NN cells but not *Apc*TT or *Apc*NT ESCs (Figure 2F and Figure S1).

RESCUING *TCF3* EXPRESSION IN *APC*NN ESCS PARTIALLY RESTORES NEURAL DIFFERENTIATION

It has been previously shown that *Tcf3* not only functions as a controller of self-renewal in wild type ESCs, but it is also required for proper neurogenesis in zebrafish, xenopus and mice [11,16,21]. We therefore hypothesized that *Tcf3* down-regulation in *Apc*NN ESCs might mediate the neural differentiation defects observed in these cells. To test this hypothesis, we rescued *Tcf3* expression by stably over-expressing its full-length cDNA in *Apc*NN ES cells (Figure 3A, B). *Tcf3* over-expression decreased TOP-Flash reporter activity (Figure 3C) and, accordingly, reduced the transcript levels of *Cdx1* and *Brachyury* (*T*), two well-known Wnt downstream targets. Gene expression profiling of *Tcf3*-expressing *Apc*NN cells confirmed that *Tcf3* effectively reverses the expression pattern of several genes differentially expressed in *Apc*NN when compared to wild type ESCs (Figure S3).

Since it has been previously reported that Tcf3 over-expression in wild type ESCs induces differentiation under self-renewing conditions [5], we first assessed whether over expressing Tcf3 in *Apc*NN ESCs induces similar effects in these cells. As reported above, *Apc*NN cells can grow in 1i medium (i.e. in LIF + Mek inhibitor) in the absence of GSK inhibitor (Figure 1C). To investigate whether Tcf3 can restore their dependency on the GSK inhibitor in serum-free culture, *Tcf3*-over expressing *Apc*NN cells were seeded at clonal density under different conditions and subsequently stained for alkaline phosphatase (AP) to evaluate the percentage of undifferentiated colonies. We found that, similar to the parental *Apc*NN cells, *Tcf3*-rescued clones show the highest colony forming capacity in the presence of LIF and Mek inhibitor (Figure 3D). Moreover, by applying the short term differentiation assay in N2B27 medium, we found that both *Apc*NN and their *Tcf3*-rescued counterparts retain expression of the pluripotency markers *Nanog* while fail to express the early differentiation markers *Fgf5* (Figure 3E). Hence, constitutive Wnt signaling prevents differentiation in a short-term assay despite the ectopic *Tcf3* expression.

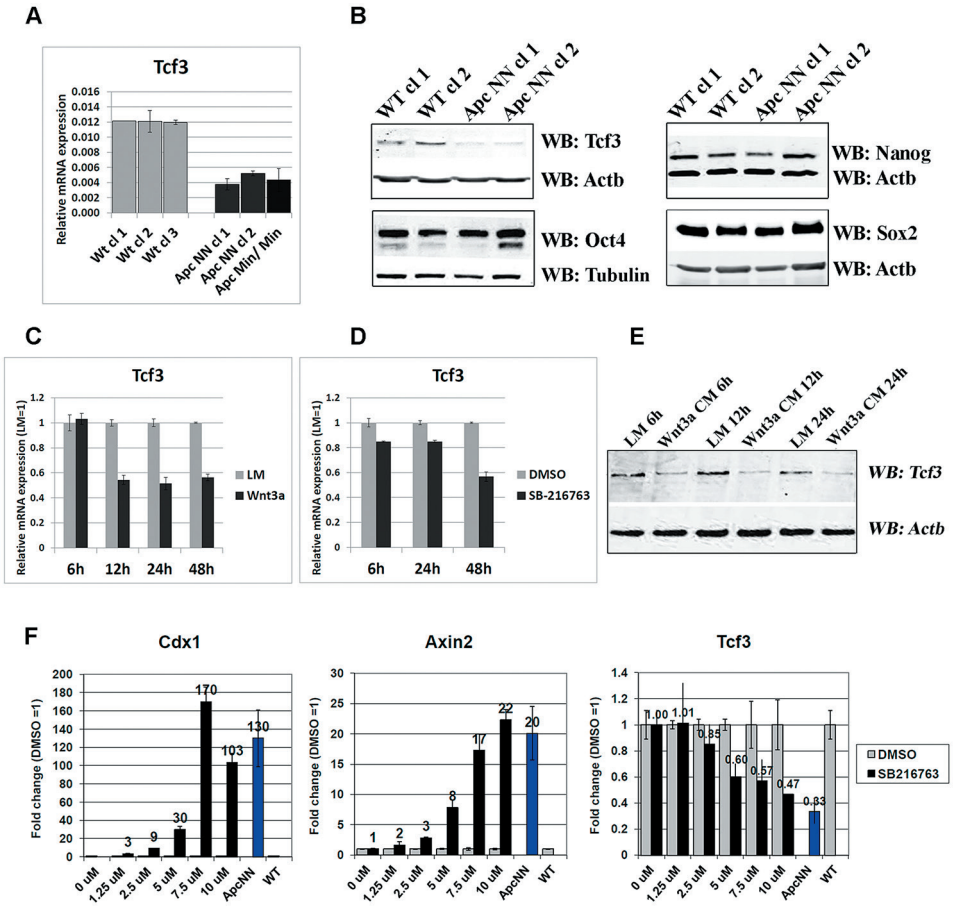


Figure 2. Wnt signaling downregulates Tcf3 expression in mouse ESCs.
(**A**) qRT-PCR analysis of *Tcf3* in wild type, *Apc*NN and *Apc*$^{Min/Min}$ ESCs. *Actb* was used as an internal control; bars represent n=2 ± SD; (**B.** Western blot analysis of the core pluripotency markers Oct4, Nanog, Sox2 and Tcf3 on protein lysates isolated from two independent *Apc*NN clones and wild type control ESCs. Actb and Tubulin were used as an internal control; (**C-D**) qRT-PCR analysis of *Tcf3* in wild type ESCs treated for different time intervals with Wnt3a conditioned medium (**C**), or with the GSK-inhibitor SB-216763 (**D**). L-medium and DMSO were employed as controls, respectively. *Actb* was used as an internal control; bars represent n=2 ± SD; (**E**) Time course western blot analysis of Tcf3 expression in wild type ESCs treated with Wnt3a conditioned medium (Wnt3a CM) or control L-medium (LM). *Actb* was used as an internal control; (**F**) qRT-PCR analysis of *Tcf3* and Wnt target genes Axin2 and Cdx1 in wild type ESC treated for 48h with different concentrations of GSK-inhibitor SB-216763 or DMSO as control. *Actb* was used as an internal control; bars represent n=2 ± SD.

We then asked whether rescuing Tcf3 expression in *Apc*NN cells could affect the neural differentiation potential of these cells. To this aim, we applied the *in vitro* neural differentiation assay previously described by Bibel et al. [30]. We found that, whereas wild

type ESCs readily gave rise to Tuj1-positive cells, no staining could be detected in *Apc*NN cells, while only few dispersed Tuj1-expressing cells were observed in the *Tcf3* rescued clones (Figure 3F). In contrast, a clear increase in *Nestin* expression was observed in *Tcf3* over-expressing cells (Figure 3G and Figure S4). This suggests that, although *Tcf3* could not restore the formation of fully mature Tuj1-proficient neurons, it does affect neural differentiation *in vitro* in a more subtle fashion towards neural progenitor-like cells. Next, we examined the differentiation potential of *Tcf3*-rescued ES cells *in vivo* by teratoma assay. We injected the newly generated clones into recipient isogenic mice to generate teratomas and analyzed them for the expression of different neuroectodermal markers by IHC. Interestingly, in contrast to the control *Apc*NN teratomas which did not express any neuroectodermal marker (0/20 analyzed teratomas), approximately 50% of all teratomas generated from different *Tcf3* over-expressing ES clones were positive for the same set of markers (6/10, 6/10, and 4/10 teratomas originated from clones 1, 2 and 3, respectively) (Figure 4). However, the extent of neural differentiation was lower compared to teratomas originated from wild type ESCs. Unlike neuroectodermal lineages, Tcf3 did not rescue the mesodermal cartilage-differentiation defect.

The observed difference in the results obtained by *in vivo* and *in vitro* differentiation assay might reflect the presence of different microenvironmental factors and the longer period of differentiation *in vivo*, which result in a larger extent of neural differentiation in teratomas.

Overall, these results indicate that *Tcf3* expression in *Apc*NN cells can partially rescue the neural differentiation defect characteristic of these cells. Next, we then asked whether Tcf3 down-regulation in wild type embryonic stem cells is sufficient to induce neural differentiation defects, characteristic of Wnt-high ESCs. To this aim, teratomas were obtained by subcutaneous transplantation of *Tcf3*$^{-/-}$ ESCs [14] followed by IHC and qRT-PCR analysis of different neural markers. We observed reduced neural differentiation in *Tcf3*$^{-/-}$ teratomas when compared to wild type controls (Figure 5). However, high expression of the pluripotency markers *Oct4* and *Nanog* was also observed in *Tcf3*$^{-/-}$ teratomas (Figure 5). IHC analysis of Oct4 also showed that *Tcf3*$^{-/-}$ teratomas are largely composed of undifferentiated, embryonic carcinoma (EC) -like cells, confirming their undifferentiated nature. This is in contrast with *Apc*NN teratomas where pluripotency markers were down-regulated. These results suggest that *Tcf3* down-regulation in wild type ES cells is necessary but insufficient to fully inhibit neural differentiation, and that canonical Wnt signaling is still required for redirecting the differentiation towards non-neuroectodermal lineages.

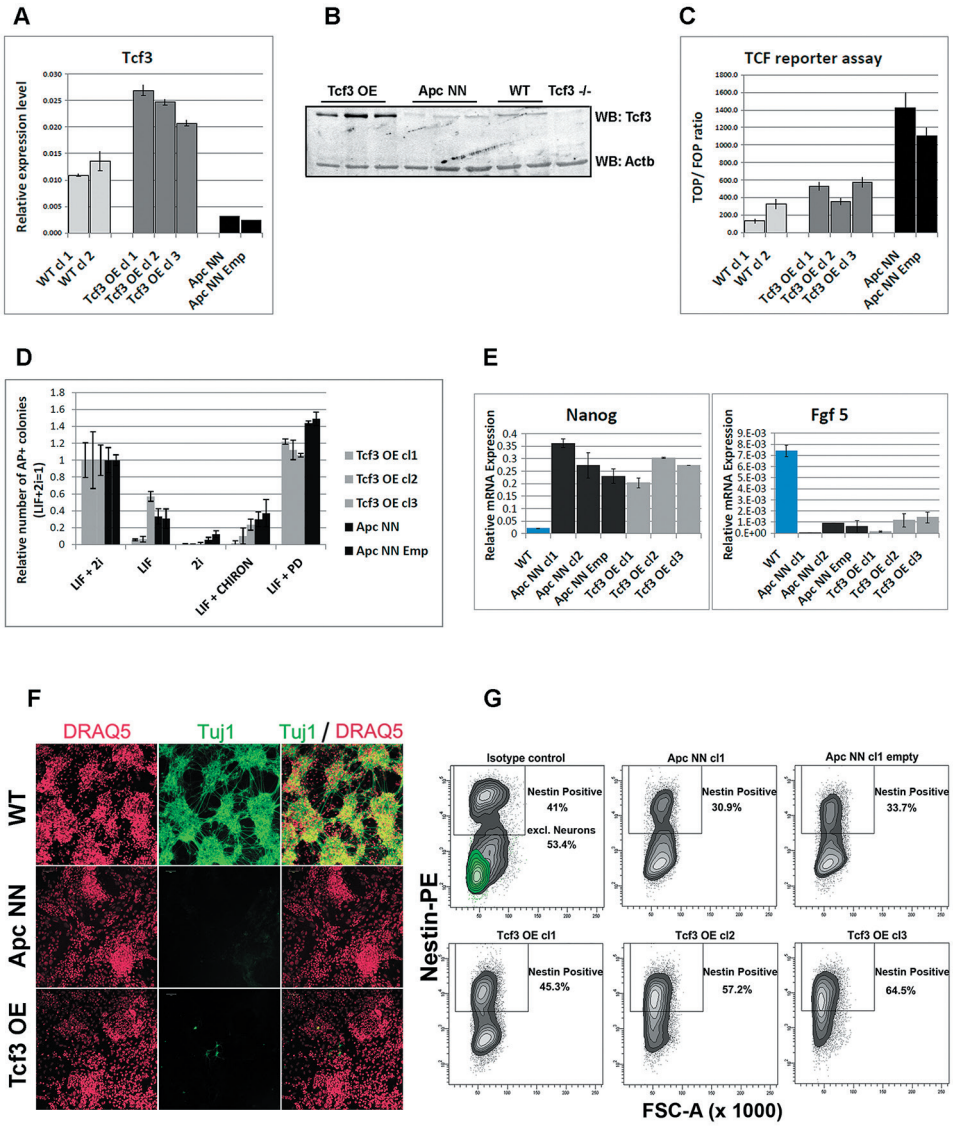


Figure 3. Characterization of Tcf3 over expressing ESCs

(**A-B**) qRT-PCR (A) and western blot (B) analysis of Tcf3 in *Apc*NN ESCs stably expressing Tcf3. Wild type and *Tcf3*$^{-/-}$ ESCs were used for comparison. Actb was used as an internal control; (**C**) Histogram showing reduction of β-catenin/Tcf reporter activity in *Apc*NN cells stably expressing Tcf3 (Tcf3 OE) compared to parental *Apc*NN cells and cells expressing the corresponding empty vector. Luciferase signal from TOP or FOP reporter constructs were measured and TOP/FOP ratios are shown in the graph. Bars represent n=3 ± SD; (**D**) Histogram showing the percent of alkaline phosphatase (AP) positive colonies formed by plating 500 FACS-sorted cells in N2B27 medium after 7 days. N2B27 medium was supplemented with different combinations of LIF, PD and CHIRON. Two independent *Apc*NN ESC clones (parental clone and transfected with empty vector) and three independent *Apc*NN ESC clones expressing *Tcf3* (Tcf3 OE) were used.

Figure 3. *Continued*
Bars represent n=3 ± SD; (**E**) Histograms showing relative expression of the pluripotency markers *Nanog* and the early differentiation markers *Fgf5* in different ESCs cultured for 48h in N2B27 medium; (**F**) Confocal analysis of ES cells stained with Tuj-1-Alexa 488 and counterstained with the far-red nuclear stain DRAQ5. Wild type, *Apc*NN and *Apc*NN expressing *Tcf3* (Tcf3 OE) ESCs were used in -4/+4 neural differentiation assay and analyzed by immunofluorescence after 13 days of culture; (**G**) Flow cytometric analysis showing expression of the neural progenitor marker Nestin in *Apc*NN ESCs stably expressing *Tcf3* (Tcf3 OE) and their control cells (parental *Apc*NN clone and *Apc*NN transfected with the corresponding empty vector) or wild type ESCs. Cells were analyzed by the -4/+4 neural differentiation assay and stained with specific antibody against Nestin and Tuj1 after 13 days of culture. Wild type (WT) ESCs are shown as control to indicate the Tuj1 positive population which is absent in other genotypes (0.1% in average in Tcf3 OE clones). Numbers in the graph represent the percent of Nestin-positive cells. For wild type ESCs the Nestin-positive populations before and after excluding the mature neurons are shown. See also Figure S4 for defining different FACS gates.

TCF3 DOWN-REGULATION IN *APC*NN ESCS IS ASSOCIATED WITH HISTONE MODIFICATIONS

To elucidate the mechanisms underlying Wnt-driven repression of *Tcf3* expression, we first analyzed its promoter activity in *Apc*NN and wild type ESCs to localize the responsible regulatory elements. We employed luciferase reporter constructs containing a 6.7 kb genomic fragment upstream of the mouse *Tcf3* ATG translation start site as well as a series of different deletion constructs spanning 4.5kb, 3.5kb, 2.2kb, 1.2kb and 176 bp fragments of the same region (Figure S5A) [31]. The 4.5 kb fragment was previously shown to resemble endogenous *Tcf3* expression in mouse embryo as well as embryonic derived neural stem cells [31]. To test whether Wnt signaling affects Tcf3 promoter activity, we transfected the different *Tcf3* promoter constructs in *Apc*NN and wild type ESCs. Likewise, transfected wild type ESCs were also treated with Wnt3a conditioned medium or L-control medium to examine *Tcf3* promoter activity. Using both approaches, we found that the Wnt-mediated repression of Tcf3 is not regulated by elements located within the 6.7 kb promoter region (Figure S5A). However, we cannot exclude the possibility that long-range enhancer elements located outside the 6.7 kb promoter region might still contribute to the observed *Tcf3* repression in Wnt context.

The mouse *Tcf3* promoter contains a large CpG island extending over exon 1, 2 and 3. This indicates that DNA methylation may play a role in the regulation of *Tcf3* expression [32]. To test whether the observed *Tcf3* down-regulation in *Apc*NN ESCs results from DNA methylation, we employed the bisulfite-conversion method followed by sequencing and methylation-specific PCR to analyze the *Tcf3* promoter in *Apc*NN cells and compare its methylation pattern to wild type ESCs. As depicted in Figure S5B, we found that similar to wild type ESCs, the *Tcf3* promoter is unmethylated in *Apc*NN cells thus suggesting that DNA methylation is unlikely to represent the mechanism underlying Wnt-driven *Tcf3* down-regulation in mouse ESCs.

Active and repressed promoters are thought to be associated with histone marks, which reflect the gene expression status of the corresponding genes. To test whether *Tcf3* down-regulation in *Apc*NN cells is regulated via chromatin modifications, we performed chromatin immunoprecipitation (ChIP) to analyze post-translational histone modifications associated with active and repressed promoters. We studied the active-chromatin marks H3K4me3 and H3-acetylation as well as the repressed-

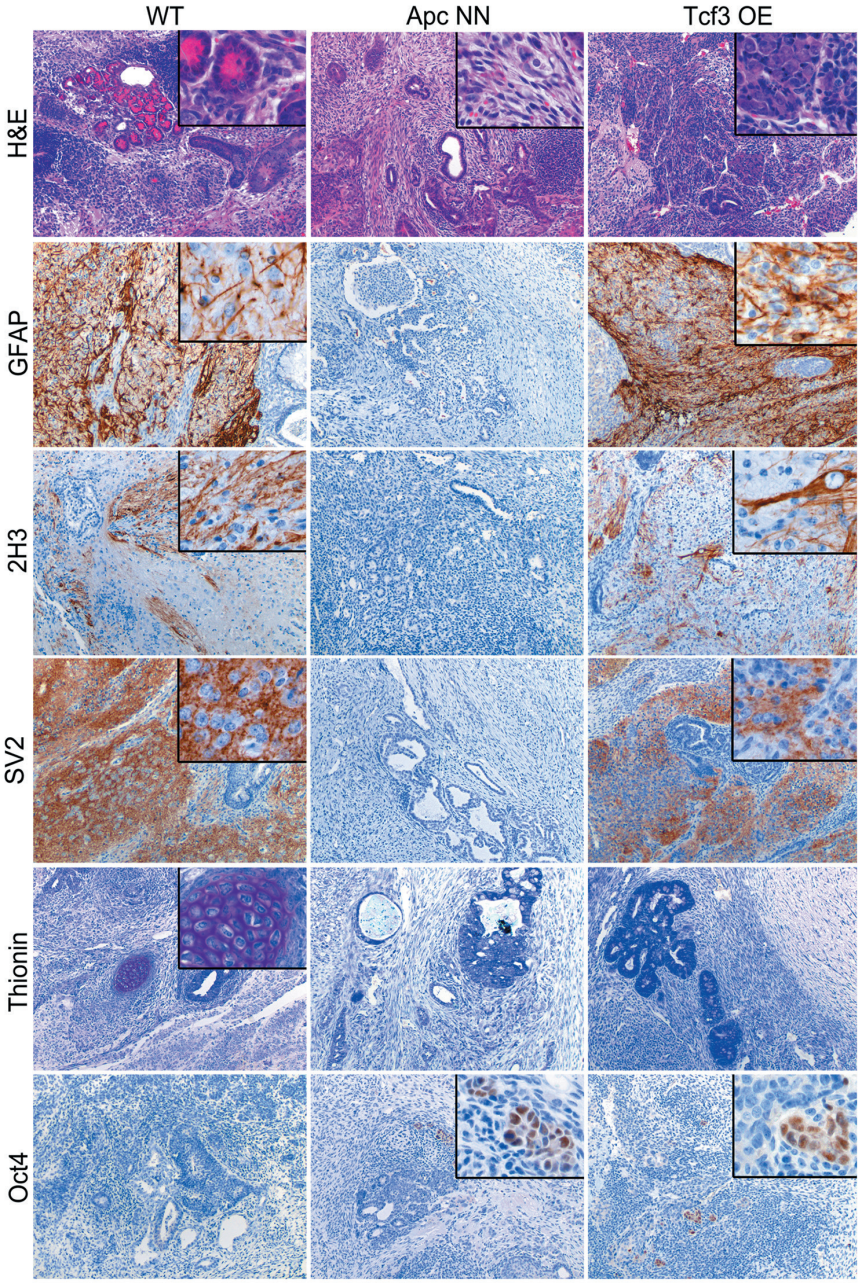


Figure 4. Rescue of *Tcf3* expression in *Apc*NN ESCs partially restores *in vivo* neural differentiation. Teratoma samples were obtained from wild type, *Apc*NN and *Apc*NN stably expressing *Tcf3* (Tcf3 OE) ESCs. Tissue sections were stained by H&E, thionin (marker of cartilage differentiation), and by IHC with specific antibodies against the neural differentiation markers GFAP, 2H3 (neurofilaments) and SV2 (synaptic vesicles). Oct4 IHC analysis was used to asses the presence of undifferentiated EC-like cells in the teratomas.

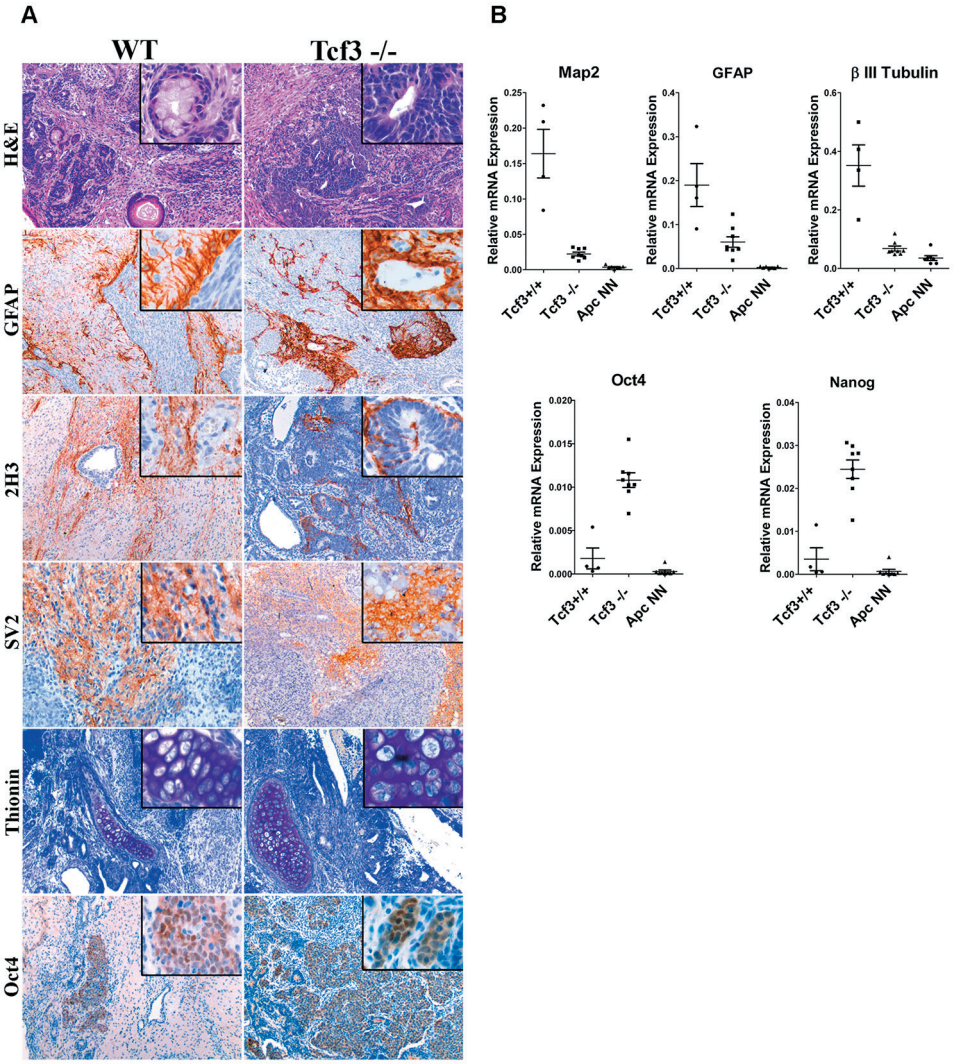


Figure 5. Tcf3 downregulation in wild type ES cells impairs but does not fully inhibit neural differentiation.

(**A**) Immunohistochemistry analysis was used to evaluate the neural differentiation in teratoma samples derived from *Tcf3*$^{-/-}$ or their wild type control (GS1) ESCs. Immunostaining with specific antibodies revealed retention of the pluripotency marker Oct4 and expression of the neural markers GFAP, neurofilaments (2H3) and synaptic vesicles (SV2) in *Tcf3*$^{-/-}$ teratomas. Thionin staining was used to evaluate cartilage differentiation; (**B**) RNAs were isolated from different teratoma samples and analyzed by qRT-PCR for differentiation markers. Dot plots show normalized qRT-PCR values for the neural markers *Map2*, β-*III-Tubulin* and *GFAP* and for the pluripotency markers *Oct4* and *Nanog* among the different teratoma samples. Each dot represents one sample.

chromatin marks H3K27me3 and H3K9me3 in the *Tcf3* promoter and compared the histone modification patterns between *Apc*NN and wild type ESCs. The immunoprecipitated chromatin was then assessed by qPCR analysis with a panel of specific primers covering a region encompassed between -2kb to +2kb from the transcription start site (TSS), as well as 20 kb of the gene body within the *Tcf3* locus. In accordance with the observed *Tcf3* downregulation in *Apc*NN cells, we found a decrease in the activating marks H3K4me3 and H3Ac and, to a lesser extent, a slight increase in the repressive marks H3K27me3 and H3K9me3 (Figure 6). Similarly, 12h treatment of wild type ESCs with Wnt3a conditioned medium significantly reduced the H3Ac and H3K4me3 activating marks but had no effect on the H3K27me3 and H3K9me3 repressing marks (Figure S6). These data demonstrate a correlation between *Tcf3* expression and histone modifications in its promoter suggesting that Wnt signaling might regulate Tcf3 expression through epigenetic mechanisms. However, the mediator of this regulation still remains elusive.

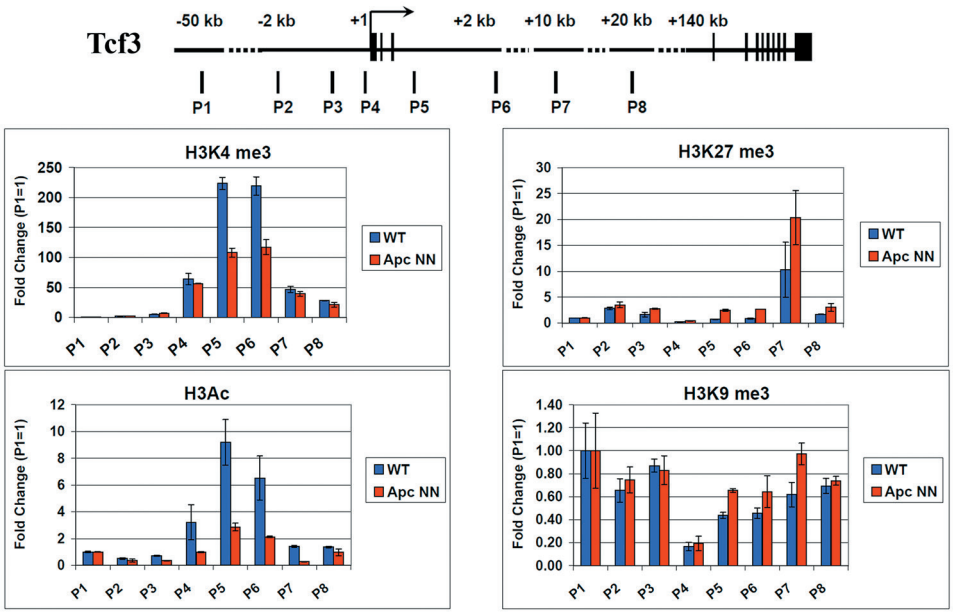


Figure 6. Regulation of Tcf3 in *Apc*NN ESCs is associated with histone modifications
Schematic representation of mouse Tcf3 locus and the different amplicons (P1-P8) analyzed by QPCR in chromatin immunoprecipitation experiment. Chromatin was isolated from *Apc*NN and wild type ESCs and was immunoprecipitated with specific antibodies against the activating histone marks (H3K4me3 and H3Ac) and the repression histone marks (H3k27me3 and H3K9me3). The input DNA (chromatin before immunoprecipitation) and immunoprecipitated DNA was quantified by QPCR and using specific primers as described in materials and methods. Values from each amplicon were normalized to input chromatin and fold change was calculated relative to the corresponding negative region (P1). Bars represent n=2 ± SD.

MIR-211, A NOVEL WNT-REGULATED MICRORNA, TARGETS *TCF3* AND ATTENUATES EARLY NEURAL DIFFERENTIATION IN WILD TYPE ESCS

It has been previously shown that members of the core pluripotency circuit are fine-tuned via microRNA-mediated regulation in embryonic stem cells [33,34,35,36,37]. Therefore we tested the idea whether Wnt-driven repression of *Tcf3* expression might also be mediated, post-transcriptionally, by Wnt-induced miRNAs. To this aim, we profiled the different *Apc*-mutant ESCs for microRNA expression by using a miRNA array encompassing specific probes for all known mouse miRNAs [38] (data not shown). Of the different candidate miRNAs induced upon Wnt activation, mmu-miR-211 showed a Wnt dosage-dependent up-regulation among the different *Apc*-mutant ESCs (Figure 7A). Accordingly, activation of Wnt signaling in wild type ESCs either by Wnt3a conditioned medium (CM) or by GSK3 inhibition, confirmed that miR-211 is a novel Wnt-regulated microRNA in mouse embryonic stem cells (Figure 7B and C).

In silico analysis with three software packages, namely Miranda [39], Targetscan [40] and PicTar [41], pointed to several potential miR-211 target genes predicted by all three programs. To narrow down the list of potential targets, qRT-PCR analysis was performed on wild type ESCs compared with *Apc*NN (Figure S7A) as well as on wild type ESCs treated with Wnt3a CM (Figure S7B). We excluded those predicted targets that showed up-regulation upon Wnt signaling. Based on these results *Sox11*, *Sf3b1* and *Tcf3* were selected for further analysis. Several stable ESC clones were generated which ectopically over-express miR-211 in an otherwise wild type background (Figure S7C). Western blot analysis showed that, unlike Sox11 and Sf3b1 (Figure S7D), Tcf3 protein level was decreased upon miR-211 ectopic expression (Figure 7D). To confirm that miR-211 directly targets Tcf3, we cloned the 3' untranslated region (3'UTR) of the mouse *Tcf3* gene in the pmirGLO reporter plasmid (Figure 7E) and performed a luciferase-based reporter assay. Transfection of HEK293 cells with the *Tcf3*-3'UTR reporter plasmid confirmed that *Tcf3* is a direct target of miR-211 (Figure 7F). The inhibitory effects of miR-211 were not observed when mutant forms of the 3'UTR, i.e. lacking 7 or 4 nucleotides of the miRNA seed sequence target (MTR1 and MTR2 respectively) were used (Figure 7F).

We next assessed the differentiation potential of miR-211 over-expressing clones using *in vitro* neural differentiation assay as well as *in vivo* teratoma formation. FACS analysis for Tuj1, a marker for mature neurons and Nestin, revealed that both miR-211 over-expressing ES cells and their wild type controls give rise to similar number of neurons and neural progenitor cells after 13 days of *in vitro* differentiation, thus suggesting that miR-211 does not affect terminal neural differentiation. As expected, *Apc*NN cells show a dramatic reduction in mature, Tuj1-proficient neurons (Figure 7G). Teratoma formation assay also confirmed that miR-211 does not suffice to inhibit neural differentiation (data not shown).

To evaluate the role of miR-211 at earlier stages of differentiation, we derived embryoid bodies (EBs) from miR-211 over-expressing cells and their wild type controls and analyzed lineage differentiation at different time points. EBs derived from wild type ES cells encompass differentiated lineages from the three germ layers, thus providing an *in vitro* assay recapitulating the early steps of embryonic development. qRT-PCR analysis for different lineage-specific markers indicated that, unlike mesodermal, endodermal and pluripotency markers (data not shown), early neuroectodermal differentiation was specifically attenuated by miR-211. We found

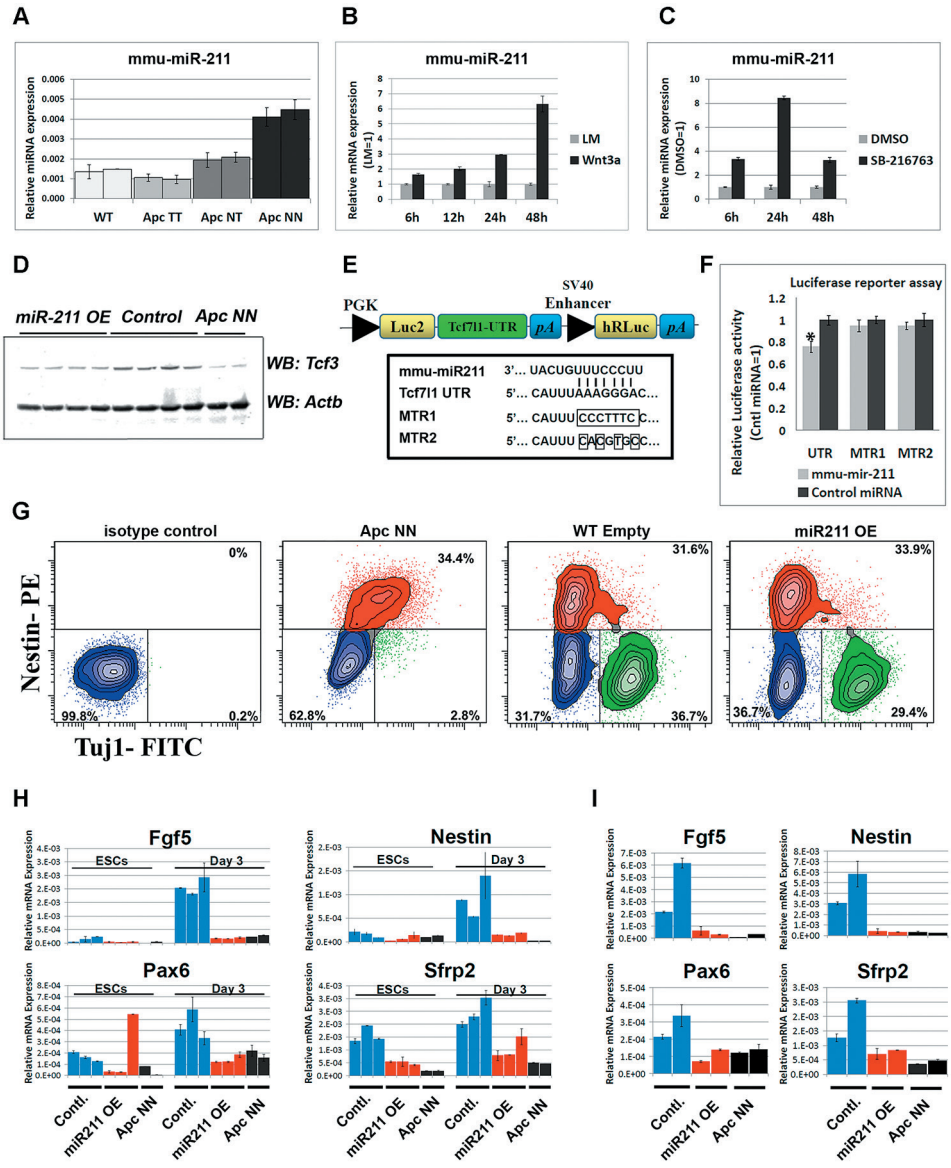


Figure 7. The Wnt-regulated miR-211, targets Tcf3 in mouse ESCs.
(**A**) qRT-PCR analysis showing a dosage-dependent up-regulation of miR-211 in different *Apc*-mutant ESCs. SnoRNA-234 was used as an internal control; bars represent n=2 ± SD; (**B-C**) Time course analysis of wild type ESCs treated with Wnt3a conditioned medium (B) or with the GSK-inhibitor SB-216763 (C). L-medium and DMSO were used as controls, respectively. RNAs were isolated at different time points and were subjected to qRT-PCR analysis of miR-211 or snoRNA-234 as an internal control. Bars represent n=2 ± SD; (**D**) Western blot analysis of Tcf3 expression in protein lysates isolated from independent clones of wild type ESCs stably expressing miR-211 (miR-211 OE) or the corresponding empty vector (control).

Figure 7. *Continued*
Two independent *Apc*NN clones were included for comparison; (**E**) Schematic representation of the Tcf3-3'-UTR luciferase vector derived from the pmirGLO construct (Promega). Sequence alignment between miR-211 and its target site on Tcf3-3'-UTR. Site directed mutagenesis was used to introduce 7-bp or 4-bp mutations in Tcf3-3'-UTR; (**F**) HEK-293 cells were co-transfected with the Tcf3-3'-UTR luciferase vector, and either with miR-211 or a non-targeting miRNA. Luciferase activity was measured 24h post-transfection and normalized to Renilla luciferase signal. The same experiment was repeated with the mutant luciferase vectors, MTR1 and MTR2. Asterix represent *P*-value <0.01 and bars represent n=3 ± SEM; (**G**) Flow cytometric analysis of Tuj1 and Nestin in miR-211 over expressing ESCs (miR-211 OE) and their controls (Emp) after 13 days of neural differentiation. Two independent clones were used for each genotype and representative example of each genotype is shown. Numbers in the graph represent the percent of cells in neural (green), progenitor (red) or negative (blue) populations; (**H**) Histogram showing the relative expression of early neural markers *Fgf5*, *Nestin*, *Pax6* and *Sfrp2* in embryoid bodies derived from independent wild type ESCs clones stably expressing miR-211 or the corresponding empty vector. RNAs were isolated at different time points and were analyzed by qRT-PCR for different lineage markers. Bars represent n=2 ± SD; (**I**) qRT-PCR analysis of *Fgf5*, *Nestin*, *Pax6* and *Sfrp2* in wild type ESCs stably expressing miR-211 or the corresponding empty vector, cultured for 24h in N2B27 medium. Bars represent n=2 ± SD.

that expression of the primitive ectoderm marker *Fgf5* and of the neural progenitor markers *Nestin* and *Pax6* as well as the early neural differentiation marker *Sfrp2* were repressed at day 3 of EB formation. Notably, these effects could not be detected at later time points (day 6, 9 or 12; data not shown). Similar results were obtained at early time points (i.e. after 24h) in N2B27 culture medium, previously described to induce neural differentiation in mESCs [42] (Figure 7H and I). These results suggest that miR-211 functions at early stages of neural differentiation and its ectopic expression in wild type ES cells is not sufficient to inhibit further neural commitment as differentiation proceeds.

Altogether, our results indicate that miR-211, a novel Wnt-regulated miRNA, can fine-tune *Tcf3* expression and attenuate early neural differentiation in wild type ESCs.

DISCUSSION

The role of Wnt/b-catenin signaling in controlling self-renewal and lineage differentiation in pluripotent embryonic stem cells has been a matter of controversy. Although both GSK3 inhibitors and Wnt ligands are essential to support ESCs self-renewal, it is yet unclear whether this occurs through b-catenin- and TCF-dependent mechanisms [43]. Among the members of the Tcf/Lef family of transcription factors, Tcf3 and Tcf1 are the most abundant in ES cells. This is of relevance as, while Tcf1 appears to function as a canonical transcriptional activator upon association with b-catenin, Tcf3 acts as a b-catenin-independent transcriptional repressor of self-renewal, suppressing genes such as *Nanog*, *Oct4* and other members of the core pluripotency circuitry [17,19]. In this scenario, it is yet unclear how canonical Wnt signaling controls the balance between Tcf1- and Tcf3-mediated gene activation and repression in the regulation of self-renewal and differentiation in ESCs.

During the last few months, several studies have been published on the specific roles of b-catenin and Tcf3 in these processes [4,5,6,44]. In the classical Wnt model, Tcf factors bind DNA and repress gene expression in the absence of active Wnt signaling. Activating the signaling pathway leads to the binding of b-catenin to Tcf proteins thus converting them from transcriptional repressors to transcriptional activators. Among the four members of Tcf/Lef family, Tcf3 seems to be different as its repressor function is not directly affected by Wnt signaling. In this perspective, two modes of action have been described for the relief of Tcf3 repression by Wnt signaling: 1) Tcf3 phosphorylation by homeodomain interacting protein kinase 2 (HIPK2) which is mediated by b-catenin and results in displacement of Tcf3 from its target sites[45]; and 2) direct physical interaction between β-catenin and Tcf3 which displaces Tcf3 and inhibits its repressive role in the context of active Wnt signaling [6,46]. Recently, using a knock-in mouse model lacking the β-catenin-interaction domain of Tcf3, Wu et al have demonstrated that counteracting Tcf3 function is not mediated by the physical interaction between β-catenin and Tcf3 during the first stages of embryonic development [47]. In view of these models, our data suggest that transcriptional and post-transcriptional down-regulation of Tcf3 expression might be yet another mechanism by which Wnt signaling inhibits Tcf3 function. It is worthwhile mentioning, however, that Wnt signaling does not seem to fully suppress Tcf3 expression and that residual levels of Tcf3 are retained even in the most severely truncated *Apc* mutant alleles (i.e. $Apc^{Min/Min}$ ESCs; Figure 2A) which encode for extremely high Wnt signaling dosages. Altogether these observations suggest that Wnt/β-catenin signaling regulates Tcf3 at several levels and by a combination of multiple mechanisms during different stages of embryonic development.

Although over-expression of a dominant negative form of Tcf1 or Tcf4 reduced the canonical Wnt reporter activity (TOP-Flash), it failed to rescue the neural differentiation in GSK-null ESCs [25]. Inhibition of β-catenin in GSK3b-null ESCs, however, was sufficient to rescue the neural differentiation defect thus confirming the central role of β-catenin-dependent mechanisms in this process [25]. The partial rescue of neural differentiation by *Tcf3* expression in *Apc*NN cells, as shown here, highlights the distinct role of Tcf3 from other members of the Tcf/Lef family and suggests that a plethora of Tcf3-dependent and -independent mechanisms underlie the Wnt-regulated lineage differentiation in embryonic stem cells.

As for self-renewal maintenance in ES cells, Wray et al have shown that a mutant form of β-catenin where the trans-activating domain was deleted, can still maintain self-renewal in mESCs cultured in 2i medium [6]. This suggests that maintenance of self-renewal is mediated by Tcf3 displacement rather than β-catenin signaling in 2i culture. Based on this, one can hypothesize that forced overexpression of *Tcf3* in Wnt context could restore the dependency on CHIRON in serum-free culture. Our data show that *Tcf3* overexpression in *Apc*NN cells does not induce differentiation in 2i culture, highlighting the dominant role of Wnt signaling in this process. This is in line with the report by Yi et al. which showed that over expressing Tcf3 in the context of Wnt signaling activation has minimal effect on self-renewal suggestive of a synergistic action of Tcf3 antagonism and β-catenin/Tcf1 signaling [5].

In an attempt to elucidate the mechanisms underlying Tcf3 downregulation in the context of active Wnt signaling, we found that Tcf3 down-regulation does not require DNA methylation but is associated with alterations in histone marks at the core Tcf3 promoter

region which are likely to regulate Tcf3 expression. Notably, these modifications occur shortly after Wnt stimulation and it is plausible to think that the chromatin modifications within the *Tcf3* locus can trigger the downregulation process of *Tcf3* expression which can be stabilized further on via miR-211 function. Epigenetic regulation through histone modification or DNA methylation was also shown previously for other antagonists of Wnt signaling such as *DACT3*, *sFRPs*, *WIF1 and DKK-1* in different cancer cells [48,49,50,51]. Further experiments are required to clarify whether this mode of gene repression is a general mechanism for Wnt-induced gene silencing in embryonic stem cells and tumor cells. Although the mediator of the observed chromatin modifications downstream of Wnt signaling remains elusive, we found that the putative *cis*-acting element, if any, is not located in the 6.7 kb promoter region which was previously described to regulate Tcf3 expression in different cell types [31]. Further work is needed to identify and study these *cis*-acting elements which might be of potential interest for providing further insight into the transcriptional repression downstream of Wnt signaling.

As an additional regulatory mechanism, we also described a novel Wnt-induced micro RNA, miR-211, and demonstrated that it targets Tcf3 in *Apc*NN ESCs. However, miR-211 over-expression in wild type ESCs does not reduce Tcf3 levels to the same degree as observed in *Apc*NN ES cells thus suggesting that multiple Wnt-mediated mechanisms are likely to exist. On the other hand, microRNAs usually exert their function by targeting multiple genes and it is plausible that miR-211 inhibits early neural differentiation in mESCs by repressing target genes other than *Tcf3*. Further experiments are required to characterize the loss of miR-211 function phenotype in mouse ESCs in order to evaluate the long-term effects on neural differentiation. The observation that Wnt signaling induces miR-211 expression might also be of interest for other disciplines of research and in particular cancer. In line with our observation, a tumor promoting function has recently been described for miR-211 in colorectal cancer cells [52]. Accordingly, miR-211 has also been shown to play a key role in melanoma tumor formation and metastasis, as well as mesenchymal to epithelial transition (MET) [53,54,55].

Taken together, we have revealed two downstream effects of Wnt signaling which contribute to the differentiation defects observed upon constitutive activation of canonical Wnt signaling, namely downregulation of Tcf3 expression and induction of miR-211. These cooperatively contribute to the inhibition of neural differentiation previously observed in *Apc*-mutant mouse ESCs [1]. We suggest that Wnt signaling represses *Tcf3* expression possibly by altering the histone marks at the Tcf3 promoter and by activating miR-211 expression, thus extending our understanding of Tcf3 regulation in stem cells. In the future, additional studies are required to elucidate how these mechanisms contribute to the regulation of *Tcf3* expression and, more in general, how Wnt signaling regulates stemness in embryonic and adult stem cells.

MATERIALS AND METHODS

ETHICS STATEMENT

This study was carried out in strict accordance with the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health. The protocol was approved by the Committee on the Ethics of Animal Experiments of the Erasmus Medical Center (DEC permit numbers EMC 2351). All efforts were made to minimize suffering.

ES CELL CULTURES AND EXPRESSION VECTORS.

Apc$^{1638N/+}$ and *Apc*$^{1638T/+}$ animals, kept on an inbred C57Bl6/J background, were bred to derive ES cells from pre-implantation blastocyts according to previously described protocols[56]. Cells were cultured on MEFs inactivated by Mitomycin-C (Sigma) in Dulbecco's Modified Eagle's Medium (DMEM, Gibco) supplemented with 10% fetal calf serum (FCS, Gibco), L-glutamine (2nM, Gibco), Na-Pyruvate (1mM, Gibco), non essential amino acids (0.1 mM each, Gibco), 2-mercaptoethanol (55mM, Gibco) and LIF (1000U/ml, Milipore). Bruce 4 ESCs were purchased from American Type Culture Collection (ATCC) and *Tcf3*$^{-/-}$ and their wild type control GS1 ESCs were obtained as previously decribed [14]. To stimulate Wnt signaling in wild type ESCs, cells were cultured on gelatin coated dishes and treated with Wnt3a-conditioned medium (collected from L-cells expressing Wnt3a plasmid) or L-control medium (collected from parental L-cells). Conditioned media were diluted 1:1 with ES medium and added to wild type ESCs for different time points. The Gsk-inhibitor SB-216763 was purchased from Sigma, dissolved in DMSO and used at 10 μM final concentration. DMSO was used as control in all the experiments.

Stable clones over expressing mmu-miR-211 were generated by transfecting Bruce4 wild type ESCs with miR-211 expressing plasmid pEZX-MR01 (Genecopoeia), or the corresponding empty vector. Several G418 resistant clones (200 μg/ml) were selected and validated for miRNA expression. In order to generate Tcf3 over expressing ESCs, *Apc*NN ESCs were co-transfected with pCAG-HA-*Tcf3*-IRES-EGFP (gift of Dr. Bing Lim, National University of Singapore, Singapore,) and Hygromycin resistance plasmid. Transfected ES cells were selected for Hygromycin (150 μg/ml). GFP expression in resistant clones was employed for validation purposes. Several independent clones were isolated and, upon validation by qPCR and western blot analysis for Tcf3 expression, employed for subsequent experiments.

The *Tcf3*-3′-UTR plasmid was obtained by PCR amplification from mouse genomic DNA of a 565 bp fragment encompassing the *Tcf3*-3′-UTR inclusive of the miR-211 target site (forward primer 5′-AAATTGAGCTCTCCCCTTGCGCTGTGGTG-3′; reverse primer 5′-AAAAACTCGAGGGTGGGGGAAGGGGCAGA-3′). PCR products were digested with *Sac*I and *Xho*I and ligated into *Sac*I and *Xho*I-cut pmirGlo plasmid (Promega). All constructs were sequenced to verify their authenticity.

MICROARRAY ANALYSIS

RNA was isolated using the RNeasy Mini Kit (QIAGEN) from cells lysed directly on the plate; a DNase step on the column was performed according to manufacturer's instructions. RNA quality was controlled by RNA 6000 Nano LabChip kit (Agilent Technologies). RNA was labeled using the GeneChip One-Cycle Target Labeling kit, hybridized to MOE430 2.0 arrays (Affymetrix) according to manufacturer's instructions. For data analysis, CEL

files were uploaded and normalized using MAS 5.0 algorithm in Expression Console software (Affymetrix, Inc). Expression analysis was performed using Partek Genomics Suite 6.5 ((Partek Inc., St. Louis, MO) and Excel 2010 (Microsoft). A robust empirical method coupled with a validation step using qRT-PCR was used to confirm the modulation of gene expressions between different genotypes. A modulation of gene expression was validated when the observed fold-change is ≥1.5 and corresponding to none overlapping individual values, not present in the background. The unsupervised hierarchical clustering was performed after MAS 5.0 normalization, using Pearson's dissimilarity as distance measure and Ward's method for linkage analysis.

DNA METHYLATION ANALYSIS

ESCs were cultured on 0.1% gelatin-coated dishes without MEFs for 2 passages and genomic DNA was isolated using DNeasy Blood & Tissue Kit (Qiagen). 1µg of genomic DNA was used in bisulfite conversion reaction using EZ DNA Methylation Kit (Zymo Research) according to the manufacturer's instructions. Converted DNA was amplified by PCR using specific primers (Table S2) designed with Methyl Primer Express Software 1.0 (Applied Biosystems) or MethPrimer software [57].

The PCR amplification was carried out using KAPA2G Robust HotStart Taq DNA polymerase (Kapa biosystems) and PCR conditions were: 95°C for 3 min and 39 cycles of 95°C for 15 sec, 57 or 53°C for 15 sec and 72°C for 15 sec, followed by 10 min at 72°C. PCR products from Region A, B and D were employed in direct sequencing using ABI BigDye Terminator and ABI 3130xl genetic analyzer (Applied Biosystems).

EMBRYOID BODY FORMATION

ESCs were trypsinized, re-suspended in ES medium and plated on gelatin-coated culture dishes for 30 min to remove MEFs. Non-attached ESCs were resuspended in EB medium (ESCs medium without LIF) at a cell density of $2x10^5$ cells/ml and plated on a non-adherent bacterial dish to initiate EB formation. EBs were collected by centrifugation (800 rpm for 5 min) every three days and re-suspended in fresh medium. 1/5 volume of EBs suspension was used for RNA extraction while the remaining EBs were kept in culture until day 12.

N2B27 SHORT TERM DIFFERENTIATION

Differentiation assays were performed as previously described [42]. Shortly, cells were trypsinized and plated on gelatine-coated dishes in N2B27 medium consisting of DMEM/F12:Neurobasal medium (1:1, Gibco) supplemented with N2 and B27 (Gibco). Cells were harvested after 24h or 48h of differentiation for further analysis.

-4/+4 NEURAL DIFFERENTIATION ASSAY

Neuronal differentiation of ESCs was induced as previously described[30]. Briefly, ESCs were trypsinized and incubated in ES medium on gelatine-coated dishes for 30 min. to allow attachment of MEFs. Non attached cells were collected and $3x10^6$ cells were cultured in 10 cm. non-adherent bacterial dishes (Greiner Bio-One) in EB medium for 8 days. Medium was refreshed every 2 days and 5 µM all-trans retinoic acid (Sigma) was added at day 4 and 6. On day 8 cells were trypsinized and plated on poly-L-ornithine/laminin-coated dishes at a density of $2x10^5$ cells/cm^2 in N2 medium. Poly-L-ornithine (Sigma) and laminin (Roche) were used at final concentrations of 0.1 mg/ml and 20 µg/ml, respectively.

N2 medium was refreshed after 2 and 24 hrs. from cell plating to remove dead cells. The N2 medium consisted of: DMEM/F12 (Gibco) supplemented with L-glutamine (Gibco), Nonessential amino acids (GIBCO), Insulin (25ug/ml, Sigma), Progesterone (20nM, Sigma), Putrescine (100 nM, Sigma), Transferrin (50 µg/ml, Sigma), Bovine serum albumin (50 µg/ml, Sigma), Sodium selenite (30 nM, Sigma) and Penicillin-Sterptomycin (Gibco).

After 48h from cell plating, medium was changed to N2B27 and refreshed every 2 days. Cells were collected after 5 days of plating for further analysis.

COLONY FORMING ASSAY AND ALKALINE PHOSPHATASE STAINING

Cells were trypsinized and plated on 0.1% gelatin-coated dishes for 30 min to remove MEFs. 500 FACS sorted cells were plated on each well of a gelatinized 24-well plate in N2B27 medium supplemented with different combinations of CHIR99021 (3 µM, Stemgent), PD0325901 (1 µM, Stemgent) and LIF (1000 U/ml, Milipore). Total number of colonies were counted after 5 days from plating upon staining with alkaline phosphatise (Milipore) according to manufacture's instructions.

TERATOMA FORMATION

Teratomas were obtained upon subcutaneous injection of $5x10^6$ cells (in PBS) into C57Bl6/J, for *Apc*-mutant ESCs (and their wild type controls), and NOD/SCID, for *Tcf3*$^{-/-}$ ESCs (and their wild type controls), recipient mice. Teratomas were collected after 2-3 weeks and used for further experiments.

RNA ISOLATION, CDNA SYNTHESIS AND QRT-PCR

RNA was isolated using Trizol (Invitrogen) or RNeasy Mini Kit (QIAGEN) and treated with DNase (Ambion) to remove contaminating genomic DNA. For gene expression analysis cDNA was synthesized using 1µg RNA and the RevertAid™ H Minus First Strand cDNA Synthesis Kit (Thermo). microRNA expression analysis was performed using 40 ng of total RNA isolated by Trizol and employed in cDNA synthesis reaction using TaqMan™ MicroRNA Reverse Transcription kit (ABI). Real-time RT-PCR was performed using Applied Biosystems inventoried assays or TaqMan™ MicroRNA Assays on a 7900HT ABI real-time PCR system (Applied Biosystems). The Delta-Ct method was used to quantify the mRNA or miRNA relative gene expressions. *Actb* or *snoRNA234* were used for normalization, respectively. qPCR analysis of the selected genes were performed using Fast SYBR® Green Master Mix (ABI) and the primers listed in table S2 .

IMMUNOHISTOCHEMISTRY

Isolated teratomas were fixed in PFA (4%) and embedded in paraffin. Five µm sections were mounted on slides stained by H&E for routine histology. Antibodies employed for IHC analysis included: rabbit anti-GFAP (1:5000, Z0334, DAKO,); mouse 2H3 against Neurofilaments (1:50, Developmental Studies Hybridoma Bank); mouse SV2 against Synaptic vesicles (1:50, Developmental Studies Hybridoma Bank); mouse A4.1025 against Adult myosin (1:50, Developmental Studies Hybridoma Bank); goat anti-Oct3/4 (1:100, sc-8629, Santa Cruz). Signal detection was performed using HRP-conjugated Goat anti mouse (1:250, Jackson ImmunoResearch), rabbit-anti-Goat-HRP (Dako) or Rabbit Envision™ kit (Dako).

IMMUNOFLUORESCENCE AND CONFOCAL MICROSCOPY

Cells were harvested and fixed in 2% PFA for 20 min, washed with PBS, and permeabilized with 0.1% triton X-100 in PBS for 15 minutes. Cells were then incubated in Blocking solution (PBS, 4% FCS) for 30 min., stained overnight at 4°C with the primary antibody, washed and finally incubated with the secondary antibody for 2h. Confocal analysis was performed with a Zeiss LSM510 confocal microscope. Tuj-1-Alexa488 was detected using a 488 nm laser and BP 500-550 emission filter. DRAQ5 was detected using a 633 nm laser and LP650 nm emission filter. Alexa 488-conjugated monoclonal anti-Tuj-1 was from Covance (A488-435L) and was used at 1:4000 dilution.

FLOW CYTOMETRY ANALYSIS

Flow cytometric analysis was performed with a BD FACSAria III, using a yellow-green laser at 561nm and a BP582/15 emission filter to detect anti-Nestin-PE antibodies, and 488nm laser and LP502 and BP530/30 emission filters for Tuj-Alexa-488 antibodies. A Live-Dead-Fixable red staining (Invitrogen) was performed before fixation, to exclude dead cells and was detected using a 633nm laser and BP660/20 emission filter.

Alexa 488-conjugated anti-Tuj-1 antibody was used at a 1:4000 dilution and the mouse anti-nestin antibody was from BD (556309) and was used at a 1:500 dilution together with a 2nd Rat-anti-mouse PE-conjugated antibody (BD; 1:1000). DRAQ5 was from Biostatus and was used as recommended by the manufacturer.

LUCIFERASE REPORTER ASSAYS

For the b-catenin/TCF reporter assay, 5x10^5 ES cells were plated on 24-well plates seeded with MEFs and subsequently transfected by Fugene HD (Roche) with 250ng of the TOP-Flash or FOP-Flash reporter constructs [28] together with 25ng of the Renilla luciferase vector for normalization purposes. Luciferase activity was measured by Dual-Luciferase Reporter Assay System (Promega). Tcf3 promoter activity was evaluated in *Apc*NN and wild type ESCs similar to b-catenin/TCF reporter assay, as mentioned above and by using Tcf3-promoter constructs (kindly provided by Nina Solberg, SCI-CAST Innovation Center, Norway) and pGL3 empty vector a control. To examine the effect of Wnt3a treatment on Tcf3 promoter activity, cells were transfected with luciferase constructs and treated with Wnt3a condition medium or L-control medium for 48h and luciferase activity was measured using Dual-Luciferase® Reporter Assay System (Promega).

For the 3'UTR-Luciferase reporter assay, HEK293 cells were plated in 24-well plates at a density of 0.5x10^5 cells per well. Cells were co-transfected with 250 ng of UTR (or MTR) reporter plasmid and either mmu-miR-211 mimic or non-targetting oligos (40 nM, Dharmacon) using lipofectamin 2000 (Invitrogen). Twenty-four hrs. after transfection, firefly-luciferase activity was measured by Dual-Luciferase® Reporter Assay System (Promega) and normalized to the co-expressed Renilla luciferase signal.

WESTERN BLOT ANALYSIS

ES cells were lysed using Cell Lysis Buffer (9803, Cell Signaling) and a cocktail of protease inhibitors (11836170001, Roche). Subsequently, NuPage LDS Sample Buffer (NP0008, Invitrogen) and DTT (1 mM) were added. Primary antibodies employed in western blot analysis included: Tcf3 (sc-8635, Santa Cruz); Sox2 (AF2018, R&D Systems); Sox11 (sc-20096, Santa Cruz); Oct4 (sc-5279, Santa Cruz); Dyrk1A G-19 (G2905, Santa Cruz);

Sap155/Sf3b1 (D138-3, MBL); Nanog (AB5731, Milipore); β-actin (A5441, Sigma); β-tubulin (ab6046, Abcam). Lysates were loaded on 10% SDS-PAGE (BIO-RAD System), and transferred onto Immobilon-FL PVDF membrane (IPFL00010, Millipore). Blocking was performed at room temperature using LI-COR Blocking buffer (Part#927-40000) diluted 1:1 with PBS. Incubation with the first antibody was performed overnight at 4°C. Blots were subsequently incubated with fluorescent-labeled secondary antibodies for 30 min. at room temperature. Goat anti-mouse IgG – IRDye 680 (1:5000, LI-COR Biosiences), Goat anti-rabbit IgG – IRDye 800CW (1:5000, LI-COR Biosiences) and Donkey anti-goat-IRDye 800CW (1:5000, LI-COR Biosiences) were used as secondary antibodies. Fluorescent signal was detected using LI-COR scanner (LI-COR Biosiences).

SITE DIRECTED MUTAGENESIS

Two mutant forms of the Tcf3-3'UTR-luc plasmid were generated using QuikChange Lightning Site-Directed Mutagenesis Kit (Agilent, 210518). We introduced either 7 bp substitutions in the miRNA binding site (AAAGGGA into CCCTTTC) to generate the MTR1-Luc plasmid, or 4 bp (AAAGGGA into cAcGtGc) to generate the MTR2-Luc plasmid. The following mutagenesis primers were employed in the reaction:

For MTR1, sense primer is 5'-tctgaaatggtccccccccctgcatttccctttcctcaag-gtgcctaccactgccttc-3' and antisense primer is 5'-gaaggcagtggtaggcaccttgaggaaagg-gaaatgcagggggggggaccatttcaga-3'. For MTR2 plasmid the sense primer is 5'-gtccccccccctgcatttcacgtgcctcaaggtgcctacc-3' and the antisense primer is 5'-ggtaggcaccttgaggcacgtgaaatgcagggggggggac-3'.

The mutagenesis reaction was performed according to manufacture's instruction. Briefly, mutant strands were synthesized using the described primers followed by DpnI digestion of the amplification products to remove the parental methylated strands. Digestion reactions were transformed in XL10-Gold Ultracompetent cells and bacterial clones with correct nucleotide substitutions were used for further plasmid extraction.

CHROMATIN IMMUNOPRECIPITATION (CHIP)

ChIP was performed on wild type and *Apc*NN ESCs or on wild type ESCs treated for 12h with Wnt3a conditioned medium or L-control medium (1:1 dilluted with ES medium). Briefly, cells were fixed in 1% PFA for 30 minutes at room temperature and PFA was quenched afterwards with 125 mM glycine. Cells were washed with buffer B (0.25% Triton-X 100, 1 mM EDTA, 0.5 mM EGTA, 20 mM Hepes, pH 7.6), buffer C (150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 20 mM Hepes, pH 7.6). Cells were then sonicated in ChIP incubation buffer (0.3% SDS, 1% Triton-X 100, 0.15 M NaCl, 1 mM EDTA, 0.5 mM EGTA, 20 mM Hepes, pH 7.6) using a BioRuptor sonicator (Cosmo Bio Co., Ltd) to obtain DNA fragments 200-700 base pairs. Chromatin was diluted in ChIP dilution buffer (with 0.15 % SDS) and incubated with BSA-blocked protein-A/G Sepharose beads (Amersham) and 5 µg antibody overnight at 4°C. Antibodies used in this study include: H3K4me3 (Abcam, Ab8580-50), H3K27me3 (Upstate, 07-449), H3K9me3 (Abcam, Ab8898-100), H3Ac (Millipore #06-599)

Beads were washed with buffer 1 (0.1% SDS, 0.1% deoxycholate, 1% Triton-X 100, 150 mM NaCl, 1 mM EDTA, 0.5 mM EGTA, 20 mM Hepes pH 7.6), buffer 2 (0.1% SDS, 0.1% deoxycholate, 1% Triton-X 100, 0.5 M NaCl, 1 mM EDTA, 0.5 mM EGTA, 20 mM Hepes pH 7.6), buffer 3 (250 mM LiCl, 0.5% deoxycholate, 0.5%

NP-40, 1 mM EDTA, 0.5 mM EGTA, 20 mM Hepes, pH 7.6), and buffer 4 (1 mM EDTA, 0.5 mM EGTA, 20 mM Hepes, pH 7.6). Chromatin was eluted for 30 min. at room temperature in elution buffer (1% SDS, 0.1 M $NaHCO_3$) and together with input chromatin, decrosslinked overnight at 65°C in the presence of 200 mM NaCl. DNA was extracted using QIAquick PCR Purification Kit and was used in QPCR analysis using Fast SYBR Green Master Mix (ABI) and primers indicated in table S2.

ACKNOWLEDGMENTS:

We thank Wilfred F. van Ijcken and Wilfrid Richer for performing gene array analysis; Simona Rossi for performing miRNA array analysis; Wolter Oosterhuis for assistance with the teratoma analysis, Nina Solberg for kindly providing the Tcf3-promoter-constructs, Juanjiangmeng Du and Rosalie Joosten for technical assistance and Frank van der Panne for assistance with figure preparation.

REFERENCES

1. Kielman MF, Rindapaa M, Gaspar C, van Poppel N, Breukel C, et al. (2002) Apc modulates embryonic stem-cell differentiation by controlling the dosage of beta-catenin signaling. Nat Genet 32: 594-605.
2. Ogawa K, Nishinakamura R, Iwamatsu Y, Shimosato D, Niwa H (2006) Synergistic action of Wnt and LIF in maintaining pluripotency of mouse ES cells. Biochem Biophys Res Commun 343: 159-166.
3. Sato N, Meijer L, Skaltsounis L, Greengard P, Brivanlou AH (2004) Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. Nat Med 10: 55-63.
4. Berge DT, Kurek D, Blauwkamp T, Koole W, Maas A, et al. (2011) Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. Nat Cell Biol 13: 1070-1075.
5. Yi F, Pereira L, Hoffman JA, Shy BR, Yuen CM, et al. (2011) Opposing effects of Tcf3 and Tcf1 control Wnt stimulation of embryonic stem cell self-renewal. Nat Cell Biol 13: 762-770.
6. Wray J, Kalkan T, Gomez-Lopez S, Eckardt D, Cook A, et al. (2011) Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nat Cell Biol 13: 838-845.
7. Rubinfeld B, Albert I, Porfiri E, Fiol C, Munemitsu S, et al. (1996) Binding of GSK3beta to the APC-beta-catenin complex and regulation of complex assembly. Science 272: 1023-1026.
8. Aberle H, Bauer A, Stappert J, Kispert A, Kemler R (1997) beta-catenin is a target for the ubiquitin-proteasome pathway. EMBO J 16: 3797-3804.
9. Hart MJ, de los Santos R, Albert IN, Rubinfeld B, Polakis P (1998) Downregulation of beta-catenin by human Axin and its association with the APC tumor suppressor, beta-catenin and GSK3 beta. Curr Biol 8: 573-581.
10. Hoppler S, Kavanagh CL (2007) Wnt signaling: variety at the core. J Cell Sci 120: 385-393.
11. Kim CH, Oda T, Itoh M, Jiang D, Artinger KB, et al. (2000) Repressor activity of Headless/Tcf3 is essential for vertebrate head formation. Nature 407: 913-916.
12. Merrill BJ, Gat U, DasGupta R, Fuchs E (2001) Tcf3 and Lef1 regulate lineage differentiation of multipotent stem cells in skin. Genes Dev 15: 1688-1705.
13. Nguyen H, Rendl M, Fuchs E (2006) Tcf3 governs stem cell features and represses cell fate determination in skin. Cell 127: 171-183.
14. Pereira L, Yi F, Merrill BJ (2006) Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. Mol Cell Biol 26: 7479-7491.
15. Dorsky RI, Itoh M, Moon RT, Chitnis A (2003) Two tcf3 genes cooperate to pattern the zebrafish brain. Development 130: 1937-1947.
16. Houston DW, Kofron M, Resnik E, Langland R, Destree O, et al. (2002) Repression of organizer genes in dorsal and ventral Xenopus cells mediated by maternal XTcf3. Development 129: 4015-4025.
17. Tam WL, Lim CY, Han J, Zhang J, Ang YS, et al. (2008) T-cell factor 3 regulates embryonic stem cell pluripotency and self-renewal by the transcriptional control of multiple lineage pathways. Stem Cells 26: 2019-2031.
18. Marson A, Levine SS, Cole MF, Frampton GM, Brambrink T, et al. (2008) Connecting microRNA genes to the core transcriptional regulatory circuitry of embryonic stem cells. Cell 134: 521-533.
19. Cole MF, Johnstone SE, Newman JJ, Kagey MH, Young RA (2008) Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes Dev 22: 746-755.
20. Yi F, Pereira L, Merrill BJ (2008) Tcf3 functions as a steady-state limiter of transcriptional programs of mouse embryonic stem cell self-renewal. Stem Cells 26: 1951-1960.
21. Merrill BJ, Pasolli HA, Polak L, Rendl M, Garcia-Garcia MJ, et al. (2004) Tcf3: a transcriptional regulator of axis induction in the early embryo. Development 131: 263-274.
22. Guo G, Huang Y, Humphreys P, Wang X, Smith A (2011) A PiggyBac-based recessive screening method to identify pluripotency regulators. PLoS One 6: e18189.
23. Korinek V, Barker N, Willert K, Molenaar M, Roose J, et al. (1998) Two members of the Tcf family implicated in Wnt/beta-catenin signaling during embryogenesis in the mouse. Mol Cell Biol 18: 1248-1256.

24. Galceran J, Farinas I, Depew MJ, Clevers H, Grosschedl R (1999) Wnt3a-/--like phenotype and limb deficiency in Lef1(-/-)Tcf1(-/-) mice. Genes Dev 13: 709-717.
25. Kelly KF, Ng DY, Jayakumaran G, Wood GA, Koide H, et al. (2011) beta-catenin enhances Oct-4 activity and reinforces pluripotency through a TCF-independent mechanism. Cell Stem Cell 8: 214-227.
26. Smits R, Kielman MF, Breukel C, Zurcher C, Neufeld K, et al. (1999) Apc1638T: a mouse model delineating critical domains of the adenomatous polyposis coli protein involved in tumorigenesis and development. Genes Dev 13: 1309-1321.
27. Fodde R, Edelmann W, Yang K, van Leeuwen C, Carlson C, et al. (1994) A targeted chain-termination mutation in the mouse Apc gene results in multiple intestinal tumors. Proc Natl Acad Sci U S A 91: 8969-8973.
28. Korinek V, Barker N, Morin PJ, van Wichen D, de Weger R, et al. (1997) Constitutive transcriptional activation by a beta-catenin-Tcf complex in APC-/- colon carcinoma. Science 275: 1784-1787.
29. Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, et al. (2008) The ground state of embryonic stem cell self-renewal. Nature 453: 519-523.
30. Bibel M, Richter J, Lacroix E, Barde YA (2007) Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. Nat Protoc 2: 1034-1043.
31. Solberg N, Machon O, Krauss S (2012) Characterization and functional analysis of the 5′-flanking promoter region of the mouse Tcf3 gene. Mol Cell Biochem 360: 289-299.
32. Spieker N, Peterson J, Reneman S, Destree O (2004) Analysis of the Tcf-3 promoter during early development of Xenopus. Dev Dyn 231: 510-517.
33. Tay YM, Tam WL, Ang YS, Gaughwin PM, Yang H, et al. (2008) MicroRNA-134 modulates the differentiation of mouse embryonic stem cells, where it causes post-transcriptional attenuation of Nanog and LRH1. Stem Cells 26: 17-29.
34. Tay Y, Zhang J, Thomson AM, Lim B, Rigoutsos I (2008) MicroRNAs to Nanog, Oct4 and Sox2 coding regions modulate embryonic stem cell differentiation. Nature 455: 1124-1128.
35. Xu N, Papagiannakopoulos T, Pan G, Thomson JA, Kosik KS (2009) MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. Cell 137: 647-658.
36. Zhong X, Li N, Liang S, Huang Q, Coukos G, et al. (2010) Identification of microRNAs regulating reprogramming factor LIN28 in embryonic stem cells and cancer cells. J Biol Chem 285: 41961-41971.
37. Wu L, Belasco JG (2005) Micro-RNA regulation of the mammalian lin-28 gene during neuronal differentiation of embryonal carcinoma cells. Mol Cell Biol 25: 9198-9208.
38. Liu CG, Calin GA, Volinia S, Croce CM (2008) MicroRNA expression profiling using microarrays. Nat Protoc 3: 563-578.
39. Betel D, Wilson M, Gabow A, Marks DS, Sander C (2008) The microRNA.org resource: targets and expression. Nucleic Acids Res 36: D149-153.
40. Friedman RC, Farh KK, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19: 92-105.
41. Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, et al. (2005) Combinatorial microRNA target predictions. Nat Genet 37: 495-500.
42. Ying QL, Stavridis M, Griffiths D, Li M, Smith A (2003) Conversion of embryonic stem cells into neuroectodermal precursors in adherent monoculture. Nat Biotechnol 21: 183-186.
43. Watanabe K, Dai X (2011) A WNTer Revisit: New Faces of {beta}-Catenin and TCFs in Pluripotency. Sci Signal 4: pe41.
44. Lyashenko N, Winter M, Migliorini D, Biechele T, Moon RT, et al. (2011) Differential requirement for the dual functions of beta-catenin in embryonic stem cell self-renewal and germ layer formation. Nat Cell Biol 13: 753-761.
45. Hikasa H, Ezan J, Itoh K, Li X, Klymkowsky MW, et al. (2010) Regulation of TCF3 by Wnt-dependent phosphorylation during vertebrate axis specification. Dev Cell 19: 521-532.
46. Solberg N, Machon O, Machonova O, Krauss S (2012) Mouse Tcf3 represses canonical Wnt signaling by either competing for beta-catenin binding or through occupation of DNA-binding sites. Mol Cell Biochem 365: 53-63.

47. Wu CI, Hoffman JA, Shy BR, Ford EM, Fuchs E, et al. (2012) Function of Wnt/beta-catenin in counteracting Tcf3 repression through the Tcf3-beta-catenin interaction. Development 139: 2118-2129.

48. Suzuki H, Watkins DN, Jair KW, Schuebel KE, Markowitz SD, et al. (2004) Epigenetic inactivation of SFRP genes allows constitutive WNT signaling in colorectal cancer. Nat Genet 36: 417-422.

49. Jiang X, Tan J, Li J, Kivimae S, Yang X, et al. (2008) DACT3 is an epigenetic regulator of Wnt/beta-catenin signaling in colorectal cancer and is a therapeutic target of histone modifications. Cancer Cell 13: 529-541.

50. Aguilera O, Fraga MF, Ballestar E, Paz MF, Herranz M, et al. (2006) Epigenetic inactivation of the Wnt antagonist DICKKOPF-1 (DKK-1) gene in human colorectal cancer. Oncogene 25: 4116-4121.

51. He B, Reguart N, You L, Mazieres J, Xu Z, et al. (2005) Blockade of Wnt-1 signaling induces apoptosis in human colorectal cancer cells containing downstream mutations. Oncogene 24: 3054-3058.

52. Cai C, Ashktorab H, Pang X, Zhao Y, Sha W, et al. (2012) MicroRNA-211 expression promotes colorectal cancer cell growth in vitro and in vivo by targeting tumor suppressor CHD5. PLoS One 7: e29750.

53. Levy C, Khaled M, Iliopoulos D, Janas MM, Schubert S, et al. (2010) Intronic miR-211 assumes the tumor suppressive function of its host gene in melanoma. Mol Cell 40: 841-849.

54. Mazar J, DeYoung K, Khaitan D, Meister E, Almodovar A, et al. (2010) The regulation of miRNA-211 expression and its role in melanoma cell invasiveness. PLoS One 5: e13779.

55. Wang FE, Zhang C, Maminishkis A, Dong L, Zhi C, et al. (2010) MicroRNA-204/211 alters epithelial physiology. FASEB J 24: 1552-1571.

56. Joyner AL, Sedivy JM (2000) Gene targeting : a practical approach. Oxford ; New York: Oxford University Press. xviii, 293 p. p.

57. Li LC, Dahiya R (2002) MethPrimer: designing primers for methylation PCRs. Bioinformatics 18: 1427-1431.

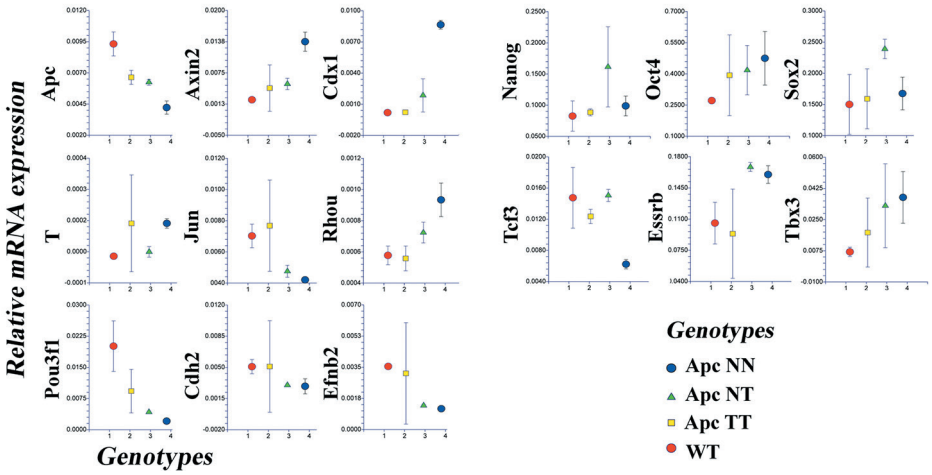


Figure S1. qRT-PCR validation of microarray results. Selected differentially expressed genes include Wnt and pluripotency-related genes. Measurements were performed in duplicates and using two independent cell lines per genotype. *Actb* was used for normalization. Plots represent average ± SD of normalized qRT-PCR values for two independent clones of each genotype.

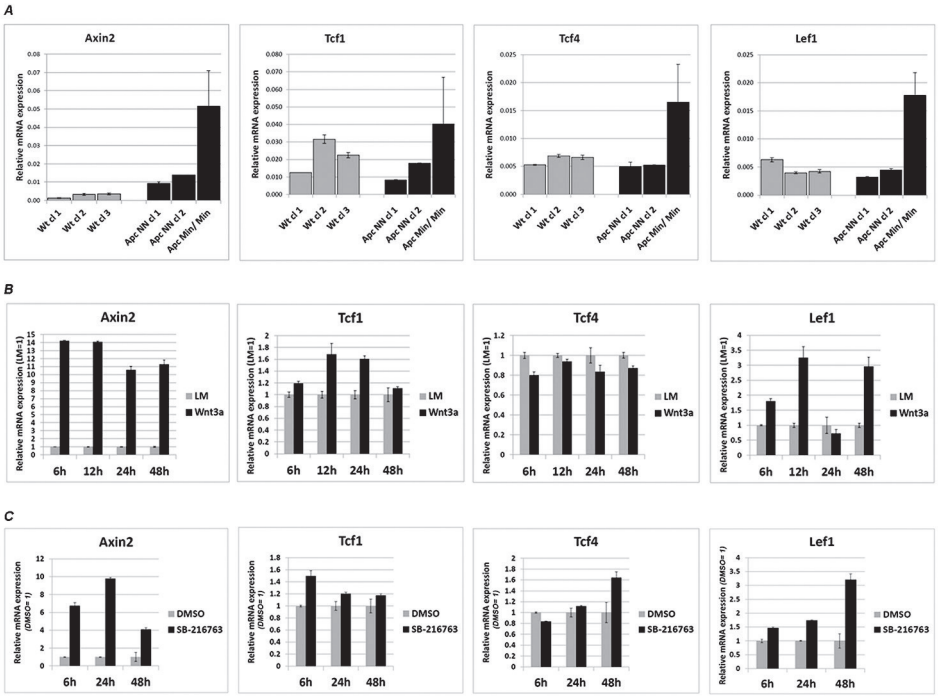


Figure S2.
(**A**) Histogram showing relative expression of *Axin2* and of members of the Tcf/Lef family in wild type, *Apc*NN and *Apc*$^{Min/Min}$ ESCs. *Actb* was used for normalization. Bars represent n=2 ± SD; (**B-C**) qRT-PCR analysis of *Axin2* and of members of the Tcf/ Lef family in wild type ESCs treated for different time intervals with Wnt3a conditioned medium (B) and with the GSK inhibitor SB-216763 (C). L-medium and DMSO were used as control media. *Actb* was used for normalization. Bars represent n=2 ± SD.

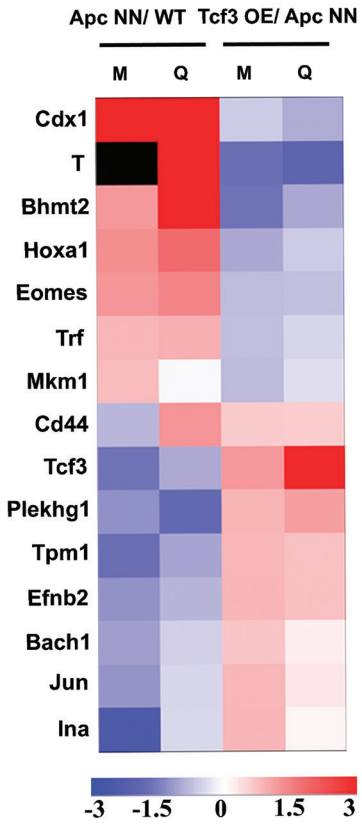


Figure S3. Heat map showing the results of the qRT-PCR validation of microarray data relative to selected genes. Genes differentially expressed between *Apc*NN and wild type ESCs were compared to the list of genes differentially expressed between ApcNN and Tcf3OE cells (table S3). Among several genes overlapping between the two microarray studies, 15 were selected for QPCR validation. The heat map shows the fold change values obtained from the microarray (M) and qRT-PCR (Q) data. *Apc*NN/WT values represent the average fold change of 2 *Apc*NN vs. 2 WT ES clones for each gene. Tcf3 OE/ *Apc*NN values represent the average fold change of three Tcf3 OE vs. three *Apc*NN clones (parental cells as well as empty vector transfected cells) for each gene. Scale represents $\log_2$ values.

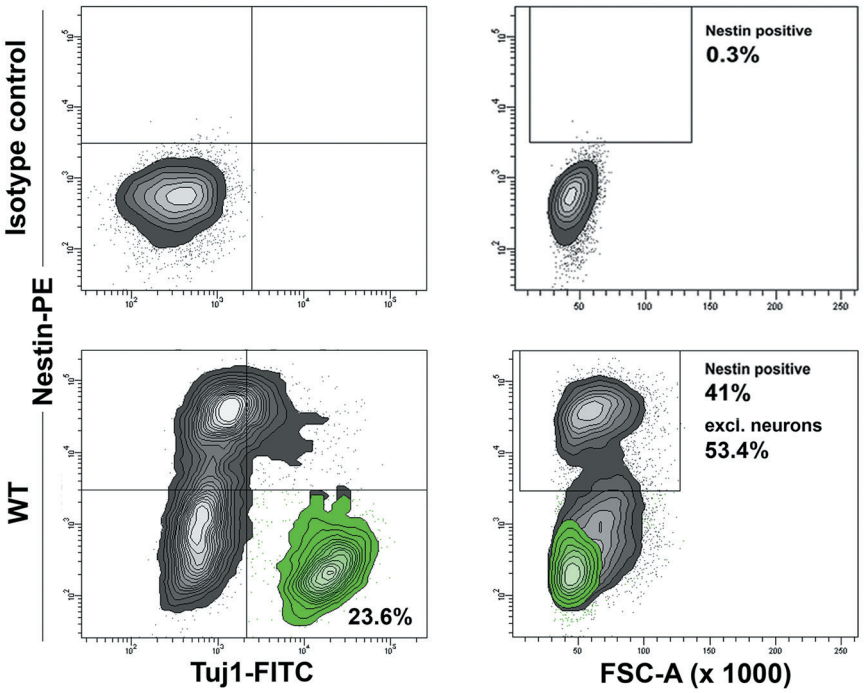


Figure S4. Supporting data to Figure 3. Flow cytometric analysis of wild type (WT) ESCs cells stained with isotype controls or Nestin and Tuj1 specific antibodies. The double staining allows the identification of Nestin positive neural progenitors and Nestin-negative/Tuj-positive mature neurons (highlighted in green). The left panel shows the Nestin-PE vs. Tuj1 and the right panel indicate the Nestin-PE against forward scattering (FSC) of the same sample. Since Tcf3 over expressing clones gave rise to 0.1% mature neurons in average, the Nestin-PE vs. FSC has been used in Figure 3.

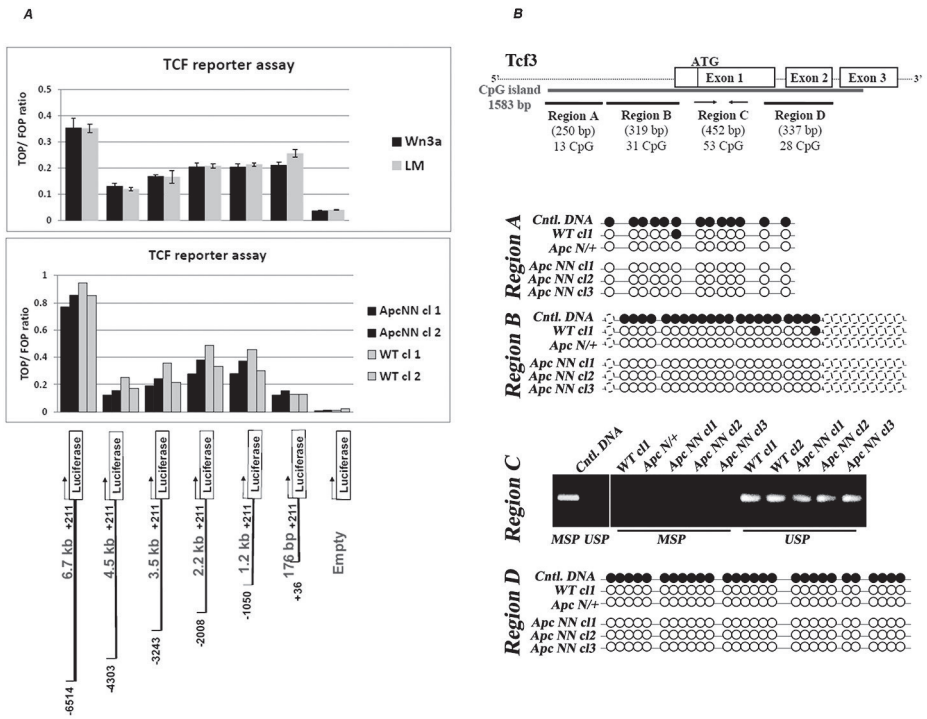


Figure S5.

(**A**) Tcf3 promoter activity in Wnt high and Wnt low ESCs. Luciferase constructs containing different Tcf3 promoter fragments were co-transfected with Renilla luciferase and the relative promoter activity is shown after normalization to Renilla-luciferase values. To monitor the effect of Wnt signaling on Tcf3 promoter activity, luciferase constructs were transfected in wild type ESCs, followed by 24h treatment with Wnt3a-condition medium (Wnt3a) or L-control medium (LM). Similarly, luciferase constructs were transfected in 2 independent clones of *Apc*NN or wild type ESCs and promoter activity was measured after 48h of transfection. The genomic location of different Tcf3 promoter fragments is depicted in the scheme. Bars represent n=2 ± SD; (**B**) DNA methylation analysis of Tcf3 promoter. Schematic representation of the mouse *Tcf3* promoter defined by the 5'UTR and ~2 kb large CpG island extending into exons 1-3. For the purpose of methylation analysis, the CpG island was subdivided into regions A, B, C and D. Genomic DNA was first bisulfite-converted and the individual regions either employed in bisulfite-specific PCR followed by DNA sequencing (region A, B and D), or used in methylation-specific PCR assays (region C). Arrows represent methylated (and un-methylated) specific primers for region C. PCR products from regions A, B and D were obtained from *Apc*NN, wild type and *Apc*N/+ (*Apc*$^{1638N/+}$) ESCs and directly sequenced. *Apc*N/+ ESCs were employed as controls since they express similar *Tcf3* levels compared to wild type ESCs (data not shown). The sequencing results are depicted as open and solid circles for unmethylated and methylated CpG dinucleotides, respectively. Dashed circles represent CpGs which were not included in the PCR products. Because of its extremely GC-rich sequence, PCR amplification of region C was carried out using methylated- and unmethylated-specific primers (MSP and USP) covering 4 different CpG dinucleotides. Control DNA (mouse genomic DNA where all CpGs sites are enzymatically methylated by means of CpG Methylase) was used in all bisulfite reactions as a positive control.

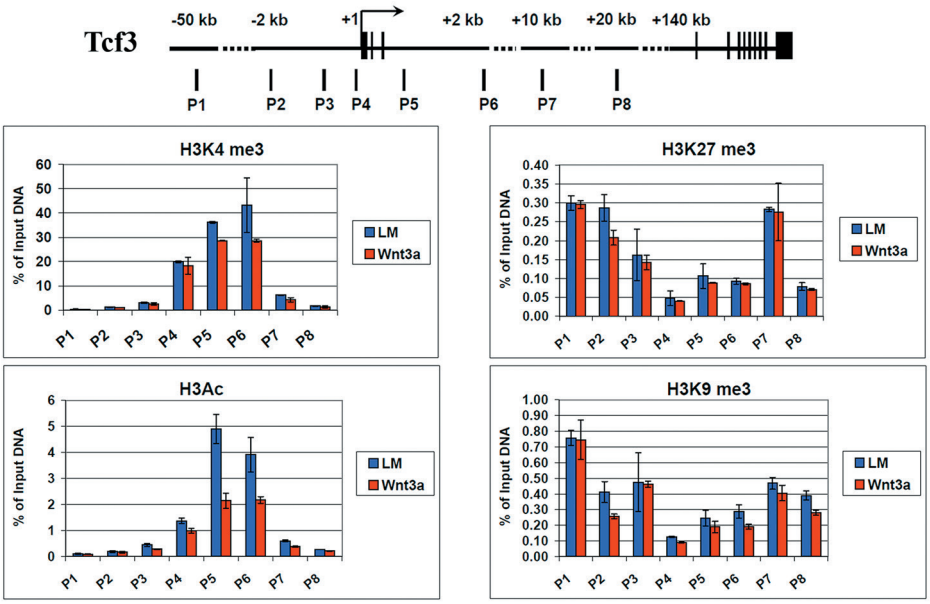


Figure S6. Transient activation of Wnt signaling in wild type ESCs reduces H3Ac and H3K4me3 activating marks in Tcf3 promoter. Bruce 4 wild type ESCs were cultured on gelatin-coated dishes and treated with Wnt3a condition medium or L-control medium for 12h. Cells were used for ChIP-QPCR as described before. Values from each amplicon were normalized to input chromatin. Since no amplification was detected at the negative region (P1) from some of the immunoprecipitated chromatin, values are shown as percent of input DNA. Bars represent n=2 ± SD.

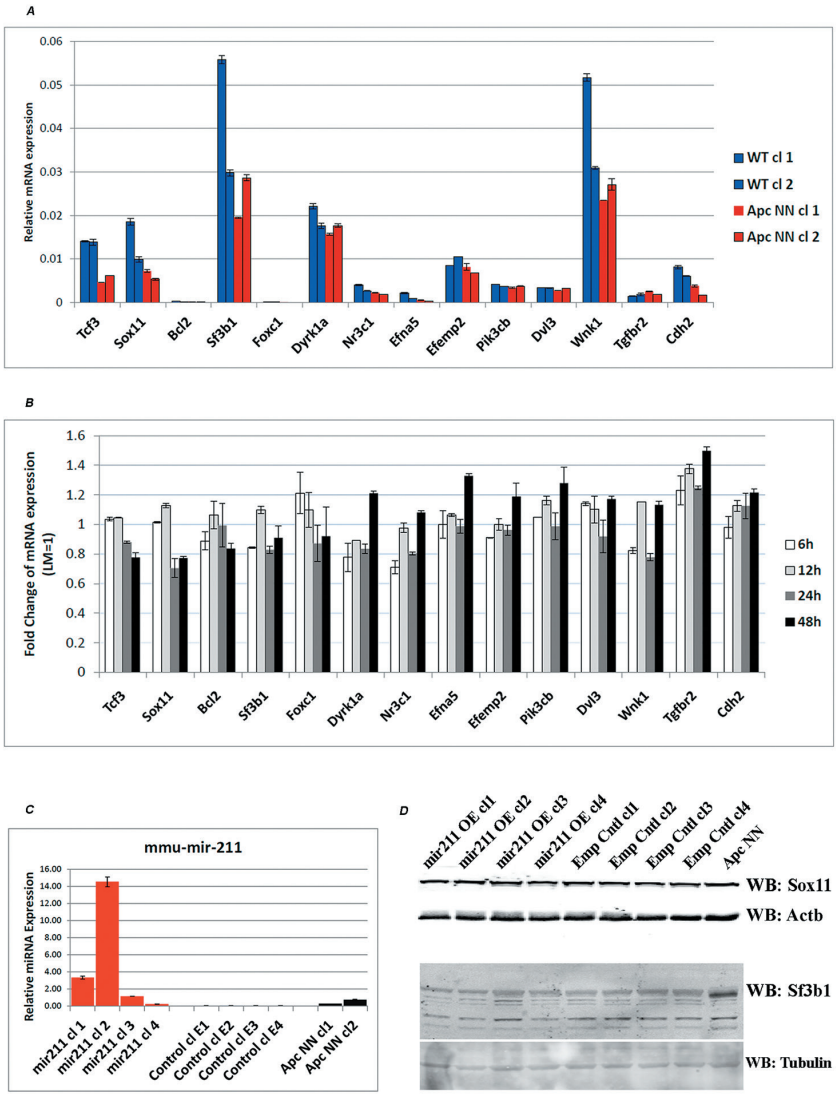


Figure S7.

(**A**) Histogram showing relative expression of selected miR-211 predicted targets in *Apc*NN and wild type ESCs. Two independent clones were used for each genotype. *Actb* was used for normalization. Bars represent n=2 ± SD; (**B**) Histogram showing relative expression of selected miR-211 predicted targets in wild type ESCs treated with Wnt3a condition medium or L-medium for different time intervals. The ratios of Wnt3a CM/ L-medium are shown in the graphs. Bars represent n=2 ± SD; (**C**) qRT-PCR analysis of miR-211 expression in wild type ESCs stably expressing miR-211 or the corresponding empty vector. Two independent *Apc*NN ESC clones were included for comparison. *snoRNA-234* was used for normalization.Bars represent n=2 ± SD; (**D**) Western blot analysis of the miR-211 predicted targets Sox11 and Sf3b1 in miR-211 over expressing cells and their wild type controls.

Table S1. Differentially expressed genes between WT, *Apc*TT, *Apc*NT and *Apc*NN ESCs.
http://www.plosgenetics.org/article/info%3Adoi%2F10.1371%2Fjournal.pgen.1003424#s5

Table S2. Primer sequences used in QPCR and DNA methylation analysis.
http://www.plosgenetics.org/article/info%3Adoi%2F10.1371%2Fjournal.pgen.1003424#s5

Table S3. Differentially expressed genes between *Apc*NN and *Apc*NN-over expressing Tcf3 (Tcf3 OE) ESCs.
http://www.plosgenetics.org/article/info%3Adoi%2F10.1371%2Fjournal.pgen.1003424#s5

chapter THREE

THE EMBRYONIC STEM CELL-SPECIFIC MIR-302-367 FAMILY IS NEGATIVELY REGULATED BY CANONICAL WNT SIGNALING IN MOUSE EMBRYONIC STEM CELLS

Yaser Atlasi[1], Rosalie Joosten[1], Andrea Sacchetti[1], Rubina Noori[1], Caterina Purini[1], Rebecca T. van Dorsten[1], Martin Rijlaarsdam[1], Simona Rossi[2], George A. Calin[2], Leendert Looijenga[1], and Riccardo Fodde[1]

[1]Dept. of Pathology, Josephine Nefkens Institute, Erasmus MC, Rotterdam, The Netherlands; [2]Dept. of Experimental Therapeutics and Center for RNA Interference and non-coding RNAs, MD Anderson Cancer Center, Houston, TX, USA

Manuscript in preparation

chapter FOUR

ECTOPIC ACTIVATION OF WNT SIGNALING IN HUMAN EMBRYONAL CARCINOMA CELLS DIFFERENTIALLY AFFECT THEIR PLURIPOTENCY IN SHORT- AND LONG-TERM IN VITRO CULTURE

Yaser Atlasi*, Rebecca T. van Dorsten*, Andrea Sacchetti, Hans Stoop, Martin A. Rijlaarsdam, J. Wolter Oosterhuis, Leendert H.J. Looijenga and Riccardo Fodde[1]

Dept. of Pathology, Josephine Nefkens Institute, Erasmus MC, Rotterdam, The Netherlands;

*These authors are contributed equally to this work

Submitted

chapter FIVE

CCAT2, A NOVEL NON-CODING RNA MAPPING TO 8q24, UNDERLIES METASTATIC PROGRESSION AND CHROMOSOMAL INSTABILITY IN COLON CANCER

Hui Ling[1]*, Riccardo Spizzo[1,]*, **Yaser Atlasi**[2], et al.

[1]Dept. of Experimental Therapeutics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA; [2]Dept. of Pathology, Josephine Nefkens Institute, Erasmus Medical Center, Rotterdam, The Netherlands.

*These authors contributed equally to this work.

Genome Res. 2013 Sep;23(9):1446-61.

ABSTRACT

The functional roles of SNPs within the 8q24 gene desert in the cancer phenotype are not yet well understood. Here, we report that *CCAT2*, a novel long non-coding RNA transcript (lncRNA) encompassing the rs6983267 SNP, is highly overexpressed in microsatellite-stable colorectal cancer and promotes tumor growth, metastasis and chromosomal instability. We demonstrate that *MYC*, *miR-17-5p*, and *miR-20a* are up-regulated by *CCAT2* through TCF7L2-mediated transcriptional regulation. We further identify the physical interaction between *CCAT2* and TCF7L2 resulting in an enhancement of WNT signaling activity. We show that *CCAT2* is itself a WNT downstream target, which suggests the existence of a feedback loop. Finally, we demonstrated that the SNP status affects *CCAT2* expression and the risk allele G produces more *CCAT2* transcript. Our results support a new mechanism of *MYC* and WNT regulation by the novel lncRNA *CCAT2* in colorectal cancer pathogenesis, and provide an alternative explanation on the SNP-conferred cancer risk.

INTRODUCTION

Notwithstanding the considerable advancements in our understanding of the molecular genetic basis of cancer, in the majority of cancer-associated genomic regions the responsible protein-coding genes have not been identified yet. The discovery of short (19-22 nt), non-coding RNAs (ncRNAs) - called microRNAs (miRNAs) [1] – not only revealed a novel mechanism of gene regulation but also led to the identification of miRNAs directly involved in cancer development [2]. It is therefore plausible that as-yet unidentified members of the broader category of ncRNA mapping to cancer-associated genomic regions play rate-limiting roles in tumor initiation and/or progression [3]. For instance, we previously reported that highly conserved genomic regions (ultra-conserved regions or UCRs) [4] are frequently transcribed as long (>200 bp) ncRNAs (lncRNAs) in both normal and tumor tissues [5]. Furthermore, germline mutations, as well as single nucleotide polymorphisms (SNPs) in ultra-conserved ncRNAs were found to occur more frequently in patients with colon cancer and chronic leukemia than in the general population [6].

The rs6983267 SNP, mapping to the 8q24.21 chromosomal region, has been consistently associated with an increased risk of colorectal cancer (CRC) [7]: the G allele was associated with greater predisposition to CRC than the T allele (odds ratios of 1.27 and 1.47 for heterozygotes and homozygotes, respectively, P=1.27x10^{-14}) [8]. The increased cancer risk from this SNP variant was also observed in other cancer types, including prostate, ovarian, and inflammatory breast cancer [9,10]. Despite the consistent association between rs6983267 and cancer risk, the underlying molecular and cellular mechanisms remain largely unknown. The genomic region spanning rs6983267 was found to contain DNA enhancer elements [11,12], and the allelic variants were shown to confer different binding affinity to TCF7L2 (transcription factor 7-like 2 [T-cell specific, HMG-box]), a transcription factor that, together with CTNNB1, plays a central role in the transcriptional activation of WNT target genes. In view of the well-established role of constitutive WNT signaling activation in CRC, these findings suggest that rs6983267 itself resides in a functional element that directly participates in colon cancer pathogenesis.

On the base of the localization of rs6983267 within a highly conserved region of the genome (http://genome.ucsc.edu) [13] and our previous report of transcription in UCRs [5], we hypothesized that an as-yet unidentified, highly conserved ncRNA could be transcribed from this genomic region. The association between rs6983267 and CRC risk further suggested that this ncRNA is likely to be aberrantly expressed in colon cancer and to play a role in CRC carcinogenesis through WNT signaling.

RESULTS

CCAT2, A NOVEL LNCRNA TRANSCRIPT MAPS TO THE HIGHLY CONSERVED 8Q24.21 REGION ENCOMPASSING RS6983267 AND IS OVEREXPRESSED IN MICROSATELLITE-STABLE CRC SAMPLES

The 8q24.21 genomic region encompasses a highly conserved segment with homologies found in 28 species-conserved tracks (University of California, Santa Cruz Genome Browser; http://genome.ucsc.edu/; Figure 1A). We designed primers spanning the rs6983267 SNP and found that this region was transcribed at high levels in human pancreas, lung, prostate, and testis tissues (data not shown). Subsequently, we cloned by RACE (rapid

amplification of cDNA ends) a 340-bp non-spliced transcript from bone marrow cDNA (Figure 1A). To exclude the possibility that this transcript was part of long intergenic non-coding RNAs (lincRNAs) mapped in this region (RP11-382A18.2 and RP11-382A18.1), we performed polymerase chain reaction (PCR) amplification using lincRNA-specific primers combined with primers from within the 340-bp transcript. We did not observe any amplified product in the COLO320 colon cancer cell line that expresses the highest levels of the newly identified transcript, suggesting that it is distinct from the two lincRNAs. We further cloned an intronless 630-bp transcript from mouse testis cDNA. Notably, the sequence overlapping between the transcripts cloned from the human and mouse tissues shared 93% homology and encompassed the rs6983267 SNP.

Because a previous study using a tiling assay did not find any transcriptional activity in this region [14], we sought other independent data to support the existence of this newly identified RNA transcript. By searching the RNA-seq database from ENCODE/CSHL (http://genome.ucsc.edu/), we found that the rs6983267 region is actively transcribed in the genomic sense orientation in GM12878 lymphoblastoid cells (Figure S1). We next determined whether this transcript has protein coding potential. The transcript's non-coding nature was suggested by the negative results of the *in vitro* translation assay, along with the lack of consistent open reading frames (ORF) by ORF Finder [http://www.ncbi.nlm.nih.gov/gorf/gorf.html], negative score by phyloCSF (-271.5366, meaning that *CCAT2* is 10^(27.15366) times more likely to be a non-coding sequence than a coding one), a comparative genomic method differentiating coding and non-coding RNA [15], lack of coding potential determined by coding potential assessment tool (CPAT) (http://lilab.research.bcm.edu/cpat/) [16], and negative data from small peptide database (Pepbank; http://pepbank.mgh.harvard.edu/search/blast) (Figure S2A). We further confirmed the existence of this novel RNA by using Northern blot analysis on a panel of colon cancer cell lines probed by sense-specific oligonucleotides spanning rs6983267: a transcript of approximately 400 bp in genomic sense orientation (centromere to telomere) was detected with the highest expression in COLO320 cells (Figure S2B). Consistent with this observation, Southern blot analysis revealed that among analyzed samples, COLO320 cells had the highest degree of genomic amplification of the 8q24 region (data not shown). By taking advantage of a cohort of nine paired colon specimens, we determined by quantitative real-time PCR (qRT-PCR) analysis that the newly transcript was expressed at much higher levels in tumor tissue than in the adjacent normal mucosa, and confirmed that the transcription occurs in the genomic sense orientation (Figure S2C). We named this novel gene transcript **C**olon **C**ancer **A**ssociated **T**ranscript **2** (*CCAT2*).

We further measured *CCAT2* expression by qRT-PCR in 215 CRC and 94 paired non-neoplastic mucosal specimens obtained from patients from four different geographical regions (i.e., Italy, Japan, Australia, and North America). We identified significantly higher *CCAT2* expression in the CRC tissue as compared with the adjacent mucosae (Figure 1B). Moreover, microsatellite-stable (MSS) cancers had higher *CCAT2* expression (~10 times) than microsatellite-unstable (MSI-H) tumors and adjacent normal colon mucosae (Figures 1B and S2C). To visualize *CCAT2* expression, we performed *in situ* hybridization by using a locked-nucleic acid probe designed against *CCAT2* RNA. In agreement with the qRT-PCR findings (Figure 1B), the epithelial expression pattern observed had stronger staining in MSS than in MSI-H cancers and normal mucosae (Figure 1C). We did not observe a significant association between *CCAT2* expression and genomic DNA amplification in an additional set of CRC samples from Japan (n=64, P=0.46, Figure S2D), suggesting that

CCAT2 expression is, at least partially, independent of the genomic DNA copy number alterations in this chromosomal region.

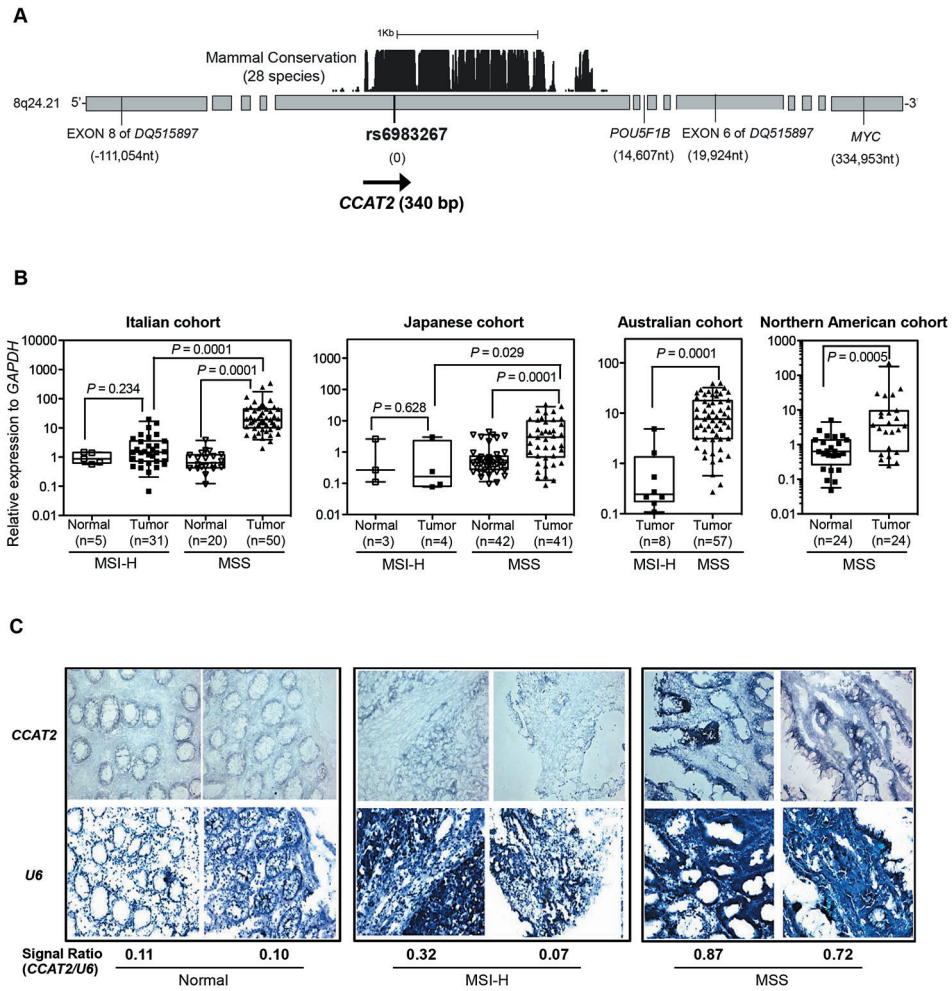


Figure 1. ***CCAT2*, a novel lncRNA spanning rs6983267, is overexpressed in colorectal cancer samples.** (**A**) Genomic features of the 8q24.21 region spanning SNP rs6983267 and the genomic location of *CCAT2*. Numbers in parentheses show the genomic distance relative to rs6983267 (image not to scale). (**B**) *CCAT2* expression in CRC patient samples. *CCAT2* expression was quantified using qRT-PCR, and data are presented as box-whisker plots showing the five statistics (lower whisker is 5% minimum, lower box part is the 25th percentile, solid line in box represents the median, upper box part is 75th percentile, and upper whisker is 95% maximum). (**C**) In situ hybridization of *CCAT2* and *U6* (as normalizer) in CRC patient samples.

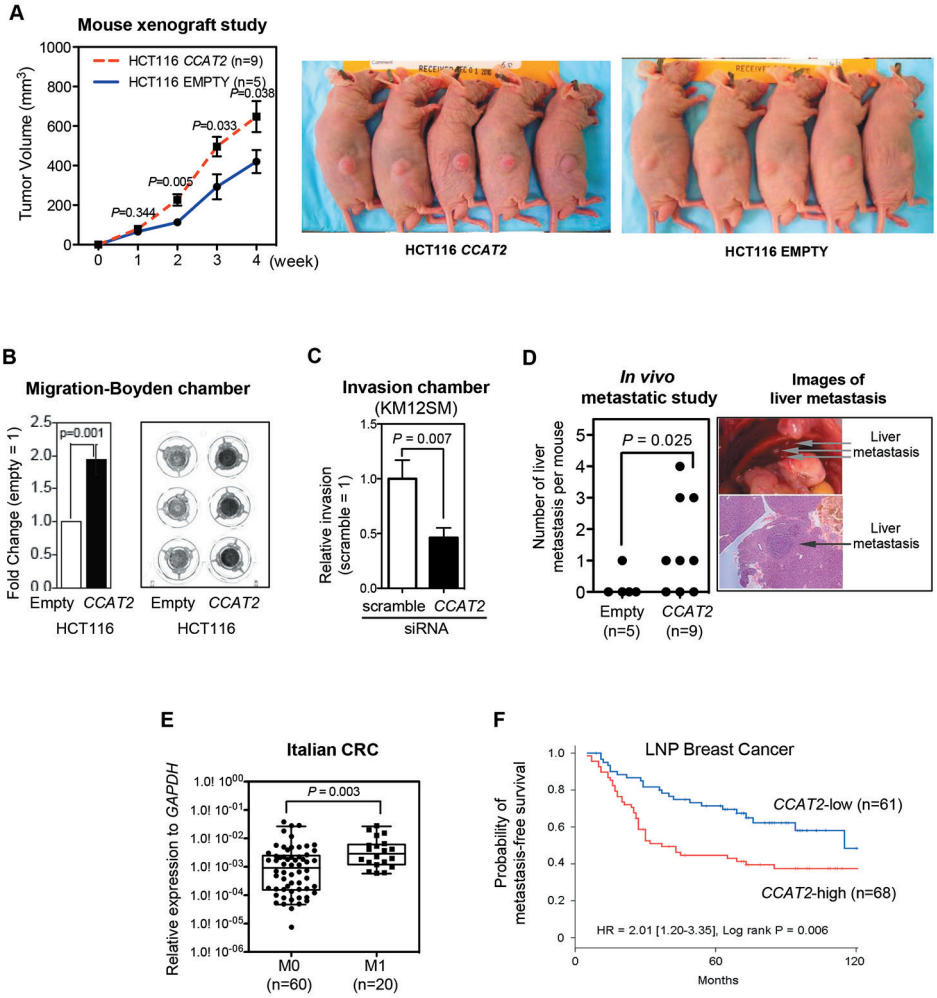


Figure 2. *CCAT2* promotes tumor growth and metastasis.
(**A**) *CCAT2* increased subcutaneous tumor formation in a mouse xenograft model. Comparison was made between the empty vector and *CCAT2* groups at the indicated weekly time points using the *t* test. (**B**) HCT116 cells transduced with *CCAT2* showed significantly higher migration ability, as measured by using a migration chamber. Data are presented as mean ± SEM from three independent experiments. (**C**) Knockdown of *CCAT2* reduces invasion ability of KM12SM colon cancer cells. After treatment with *CCAT2* siRNA (50 nM) for 24 h, cells were seeded onto invasion assay chamber for 48 h, and invaded cells were stained and counted. (**D**) *CCAT2* enhances liver metastasis in mice that were incubated by intrasplenic injection with HCT116 cells. A representative image of liver metastasis from the *CCAT2* group is shown. (**E**) *CCAT2* expression levels in Italian primary CRC samples from patients with metastasis (M1) were higher than those from without metastasis (M0). Comparison was made using Mann-Whitney test. Data are presented as box-whisker plots. (**F**) Breast cancer patients with high *CCAT2* had shorter metastasis-free survival. Kaplan-Meier analysis as a function of *CCAT2* levels in 129 lymph node-positive breast cancer patients who received adjuvant combination chemotherapy with cyclophosphamide, methotrexate, and 5-fluorouracil. *CCAT2* high level to normalizer, >0.0045; *CCAT2* low level, ≤0.0045.

CCAT2 PROMOTES CANCER GROWTH AND METASTASIS

To test the hypothesis that *CCAT2* plays an oncogenic role in CRC, we cloned *CCAT2* into a retroviral expression vector and transduced it into HCT116 colon cancer cells. This is an MSI-H cell line (*MSH6* mutant) with a near-diploid karyotype [17] that expresses low levels of *CCAT2* (when compared to COLO320; Figure S2B). Retrovirus transfection of the *CCAT2* construct caused a 630-times increase in *CCAT2* expression compared with the cells transfected with empty vector (Figure S3A). For comparison the expression difference between high expresser COLO320 cells and low-expresser HCT116 cells was ~1500 times (Figure S3A**)**. The *CCAT2*-transduced HCT116 cells did not affect cell proliferation in 2-D culture conditions (Figure S3B), but showed a growth advantage under low adherence conditions and increased colony-forming potential (Figures S3C and S3D). Subcutaneous transplantation of *CCAT2*-overexpressing HCT116 cells resulted in larger xenograft tumors in Swiss nu-nu/Ncr nude mice when compared to empty vector-transduced cells (Figure 2A). This observed *in vivo* tumor promotion effect of *CCAT2* was independently confirmed by a xenograft study using pcDNA construct-established *CCAT2*-overexpressing HCT116 cells (data not shown).

We next studied whether *CCAT2* is involved in promoting the metastatic phenotype in colon cancer cells. *In vitro* migration assay showed a 2-fold increase in the migration of *CCAT2*-transduced HCT116 cells (*P*=0.001) (Figure 2B). To validate the role of *CCAT2* in the metastatic phenotype, we used KM12SM cells, a CRC cell line established from the spontaneous liver metastasis of KM12C cells [18]. Knockdown of *CCAT2* by siRNA, which reduced the *CCAT2* expression by 70%, significantly reduced the invasive ability of KM12SM cells (that express high levels of *CCAT2*), suggesting the involvement of *CCAT2* in metastasis **(**Figures 2C and S4**)**. More importantly, spleen injection of *CCAT2*-overexpressing HCT116 cells into Swiss nu-nu/Ncr nude mice resulted in a higher incidence of liver metastasis (6 of 9 mice vs. 1 of 5 mice) as well as a greater number of metastatic tumors (*P*=0.025) than in the control group (Figure 2D). These studies suggest that modulation of *CCAT2* expression levels changes the metastatic capacity of cancer cells. In addition, in the Italian cohort of CRC samples (which was the larger available set), we found significantly higher *CCAT2* expression levels in primary CRC tumors from patients with metastasis (TNM category M1) than in those without metastasis (M0) (*P*=0.003) (Figure 2E).

To further expand on this issue, and in view of the previous report on the correlation between rs6983267 and breast cancer aggressiveness [10], we measured *CCAT2* RNA expression levels in 129 lymph node-positive breast cancer patients who received adjuvant cyclophosphamide, methotrexate, and 5-fluorouracil combination regimen. High *CCAT2* RNA levels were associated with shorter metastasis-free survival when analyzed as both a continuous variable (Table S1) and as a dichotomized variable using the median level as a cutoff point (Figure 2F).

CCAT2 INDUCES CHROMOSOMAL INSTABILITY

As chromosomal instability (CIN) represents a main feature of MSS CRCs, and in view of the high *CCAT2* expression levels among MSS colon cancers (Figures 1B and 1C), we hypothesized that *CCAT2* may underlie CIN development. We generated stable HCT116 clones **o**verexpressing ***C****CAT2* (OC1 and OC2) or with basal *CCAT2* expression (E1 and E2) by using the pcDNA expression vector (Figure 3A). Genomic instability analysis showed that the percentage of cells with normal metaphases was markedly lower in OC (59.5%

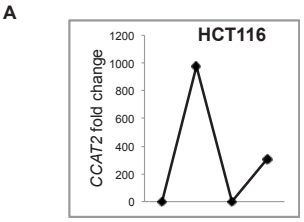
in OC1 and 31.4% in OC2) than in E (84.2% in E1 and 94.3% in E2) clones (Figure 3A). In the same analysis, *CCAT2* dramatically increased the percentage of polyploid cells from 5.3% (E1) and 2.8% (E2) to 33.3% (OC1) and 60.0% (OC2) (Figure 3A). We detected near-tetraploid status in clone OC2 by spectral karyotype analysis (Figure 3B) and observed a doubling of nuclear size and DNA content by DAPI staining and flow cytometric analysis (Figures S5A and S5B). The CIN-inducing ability by *CCAT2* was further validated by a second batch of eight overexpressing clones that exhibited numerical and structural chromosomal changes (Figures 3C, 3D and S5C). CIN was also observed in HCT116 cells transduced with *CCAT2*-containing retrovirus (Figure S5D). Stable knockdown of *CCAT2* expression in COLO320 cells partially reversed the abnormal metaphase observed in COLO320 cells (Figure S5E). To further identify the mechanism by which *CCAT2* overexpression may underlie CIN, we evaluated spindle and CREST/kinetochore structure by immunocytochemistry analysis of the *CCAT2* clones (E1, E2, OC1 and OC2) with antibodies directed against beta-tubulin and the centromere. Normal numbers of centrosomes were observed in both E1 and E2 with spindles separating chromosomes from two centromeres in the nucleus (Figure 3E). However, we detected aberrant numbers of centrosomes pulling chromosomes from three or more different points in the *CCAT2*-overexpressing clones (Figure 3E). The observed defects in centrosomes are likely to result in abnormal chromosomal breakage and fusion, and eventually aneuploidy.

As CIN scores were already available from 218 breast cancer patients with lymph node negative disease [19], we assessed the *CCAT2* expression levels in these tumor samples. We identified a significant positive correlation between *CCAT2* level and CIN score (Spearman correlation: Rs=0.17, *P*=0.012).

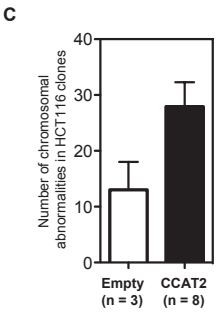
CCAT2 REGULATES *MYC* TRANSCRIPTION

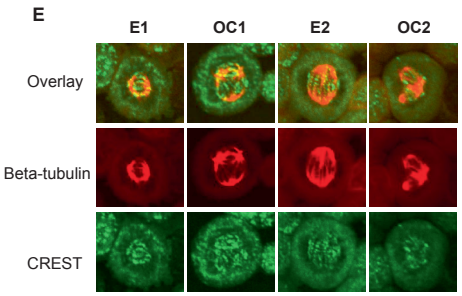
To explore the molecular mechanism underlying *CCAT2*'s oncogenic potential, we characterized the expression of other transcripts located in the same region: *MYC* and exons 6 and 8 of *DQ515897* (but not *POU5F1B*) showed expression patterns similar to that of *CCAT2*, although with smaller differences among MSI-H, and MSS cancers and normal colorectal mucosa (Figure 4A and S6A). In view of these associative data and of previous reports describing a long-range interaction between rs6983267 and *MYC* [13,14], we asked whether *CCAT2* plays a role in regulating *MYC* expression. Among the HCT116 clones, we found higher *MYC* expression at both the RNA and protein levels in cells overexpressing *CCAT2* (3.4- and 2.5-fold increases in OC1 and OC2, respectively; *P*=0.030 and *P*=0.038, respectively) (Figure 4B). Higher *MYC* expression was also found in three additional independent HCT116 clones overexpressing *CCAT2* (Figure S6B). Expression of downstream MYC protein-coding gene targets such as *BAX* [20], *CDC25A* [21], *CDKN2A* [22] and the microRNA targets such as the *MIR17HG* [23], and *miR-146a* [24] were also changed accordingly (Figures 4C, S6C and S6D). Additionally, the *MIR17HG* was also up-regulated in retrovirus-transfected cells (Figure S6E), and was significantly reduced by transient *MYC* down-regulation (Figure S6F), suggesting that these microRNAs are regulated by MYC in our experimental systems. Conversely, transient knockdown of *CCAT2* reduced *MYC* expression (both RNA and protein) in *CCAT2*-overexpressing clones (Figure 4D), suggesting that *CCAT2* RNA, and not merely its genomic locus, is required for the maintenance of high *MYC* expression levels. In addition, sustained downregulation of *CCAT2* by shRNA retroviruses decreased *MYC* expression and *MIR17HG* expression in COLO320 cells in a *CCAT2*-dependent fashion (Figures 4E and S6G). This effect was

further validated in Tet-on inducible COLO320 clones where *CCAT2* expression was reduced by specific shRNAs in response to doxycycline (Figure 4F).



	E1	OC1	E2	OC2
Metaphases analyzed	38	42	35	35
% Normal metaphase	84.2	**59.5**	94.3	**31.4**
% Poly/tetraploidy	5.3	**33.3**	2.8	**60**





D

HCT116 clones	SKY results
CCAT2 S2	39~45,X,?3[8],del(3?p)[2],?4[2],?5q[8],?6[2],del(8q)[2],der(10q)t(10;16),-11[3],?12[3], der(16?p)t(8;16),der(18p)t(17;18)[cp11] 91,idemx2[1]
CCAT2 S7	44~45,X,?5q[4],der(10q)t(10;16),der(16?p)t(8;16),der(18p)t(17;18)[cp10]
CCAT2 S13	35~45,X,del(2q)[3],?5q[4],?6[3],der(10q)t(10;16),der(16?p)t(8;16),der(18p)t(17;18)[8],-19[3][cp10] / 62,X<3n>,+1,-2,-4,-4,-5,7p/q+,-9,-9,-10,-11,+16,der(16?p)t(8;16),der(16?p)t(8;16),+18, der(18p)t(17;18),der(18p)t(17;18)[1]
CCAT2 S14	45,X,der(10q)t(10;16),der(16?p)t(8;16),der(18p)t(17;18)[9]/ 56,X,<2n>,+1,+2,?2,+3,+4,-5,+6,+6,+8,+8,-9, der(10q)t(10;16),+der(10q)t(10;16),-11,+12,+13,+14,+14,-15,-16, +17,der(18p)t(17;18),+18,+19[1]
CCAT2 S1	42~47,X,der(9q)t(9;22),der(10q)t(10;16),?der(11q)t(5;11)[2],der(16?p)t(8;16), der(18p)t(17;18)[cp12]
CCAT2 S17	42~46,X,del(2?p)[9],der(q)t(2;9),der(10q)t(10;16),der(16?p)t(8;16),der(18p)t(17;18)[cp10]
CCAT2 S19	41~56,X,del(3q)[2],?3p[2],der(8p)t(7;8)[10],der(10q)t(10;16)[10],der(16?p)t(8;16)[10], der(18p)t(17;18)[10][cp[11]
CCAT2 S23	43~46,X,der(10q)t(10;16),der(16?p)t(8;16),der(18p)t(17;18),der(22)t(21;22)[3][cp11]/ 45,der(Xq)t(X;11),r(4),der(8q)t(8;21),der(10q)t(10;16),der(11q)t(5;11),der(16?p)t(8;16),fr(17),der(18p)t(17;18),der(21)t(8;21)[1]
E10	45,X,der(10q)t(10;16),der(16?p)t(8;16)[13],der(18p)t(17;18)[cp14]
E13	42~46,X,der(10q)t(10;16)[6],der(16?p)t(8;16),der(18p)t(17;18)[6][cp7]
E15	39~45,X,?6[2],der(10q)t(10;16),der(16?p)t(8;16),der(18p)t(17;18)[cp8]

Figure 3. *CCAT2* induces chromosomal instability in HCT116 cells.
(**A**) Increased chromosomal abnormalities in *CCAT2*-overexpressing clones, as identified by genomic instability analysis. (**B**) Example of spectral karyotypes of HCT116 clones revealed a change from the original diploid (E1) to the tetraploid karyotype in the OC2 clone. (**C**) and (**D**) *CCAT2* induces chromosomal instability in an independent set of *CCAT2*-overexpressing clones. The number of chromosomal abnormalities in each clone was counted and data are expressed as mean ± SD. (**E**) Abnormal centrosomes in *CCAT2*-overexpressing clones OC1 and OC2 by immunostaining. Beta-tubulin in red and CREST/kinetochore in green.

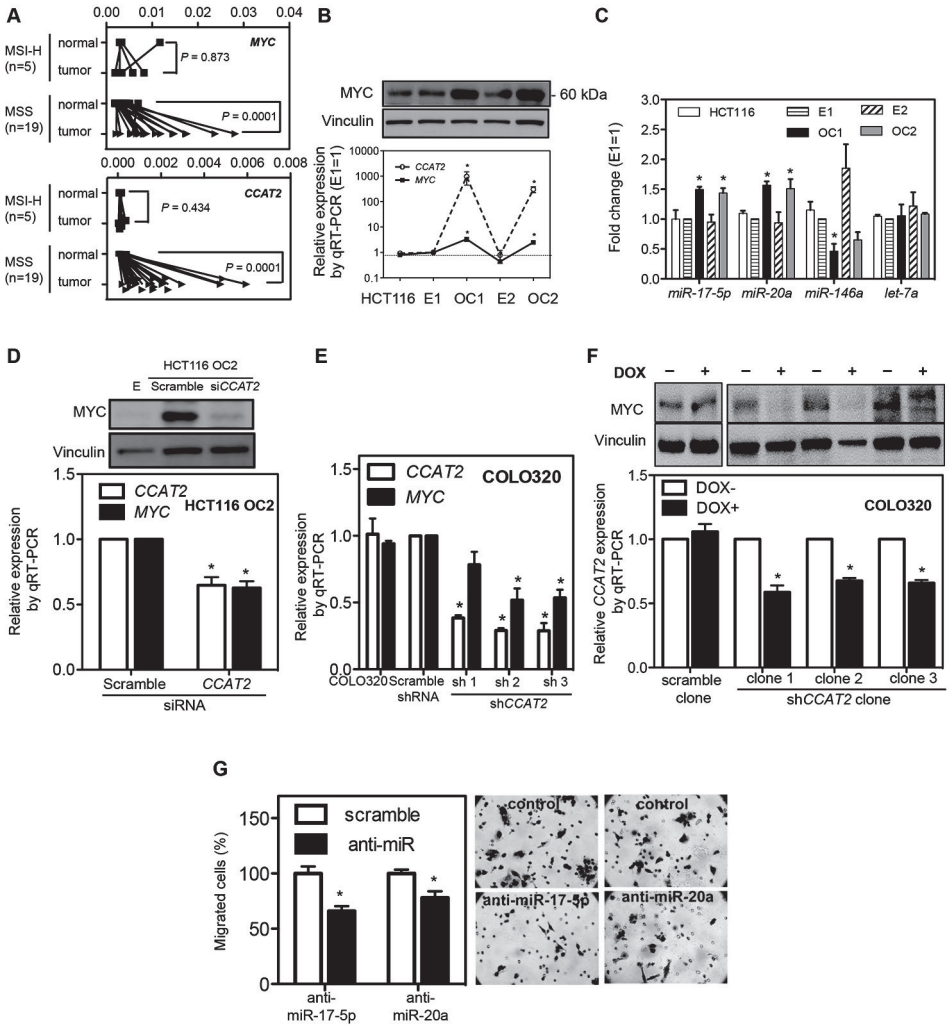


Figure 4. Regulation of *MYC* expression by *CCAT2*.
(**A**) Increased *CCAT2* and *MYC* expression in MSS CRC samples. Expression profile of *MYC* and *CCAT2* in paired CRC/control mucosa samples from Italy using qRT-PCR. (**B**) Correlation of *MYC* protein and *CCAT2* RNA levels in HCT116 clones with high and basal *CCAT2* levels. (**C**) *CCAT2* increased miRNA expression of the *MIR17HG*, decreased *miR-146a* expression, and exerted no consistent effect on *let-7a* (used as negative control). (**D**) Downregulation of *MYC* (upper panel: protein; lower panel: RNA) by *CCAT2* siRNA in HCT116 OC2 clone. (**E**) Downregulation of *CCAT2* reduced *MYC* expression in COLO320 shRNA stable clones. (**F**) Reduction of *MYC* in doxycycline-inducible sh*CCAT2* clones of COLO320. (**G**) Reduction of migration in *CCAT2*-transduced HCT116 cells by anti-miR-17-5p and anti-miR-20a. 24 h after anti-miR treatment, cells were seeded onto migration assay chamber for another 24 h, and migrated cells were stained and counted. **P*<0.05.

Next, we moved on to determine whether *MYC* and/or its targets are mediators of the above observed *CCAT2*-driven enhanced cell migration. In our experimental system, reducing *MYC* expression did not significantly decrease the migration of *CCAT2*-transduced cells (Figure S6H). However, consistent with the consensus that the *MIR17HG* increases metastasis [25,26], knockdown of *miR-17-5p* or *miR-20a* in *CCAT2*-overexpressing cells led to a significant decrease in cell migration (Figures 4G and S6I). These findings suggest a not-yet-identified complex network underlying *CCAT2*-elicited metastatic phenotype.

MECHANISMS OF *CCAT2*'S REGULATORY EFFECT

As *CCAT2*'s subcellular localization may provide additional clues about its functional role, we measured *CCAT2* expression in nuclear and cytosolic fractions from *CCAT2*-expressing HCT116 subclones and the COLO320 cell line by qRT-PCR. The differential enrichment of *GAPDH* and *U6* RNA were employed as fractionation indicators (Figure S7). In these samples, we found a considerable increase in *CCAT2* expression in the nucleus versus the cytosol, thus indicating that *CCAT2* is mainly localized in the nucleus (Figure 5A). This was further confirmed by the *in situ* hybridization findings in CRC samples (Figure 5B). Consistent with our findings, RNA-seq data showed that the transcript encompassing rs6983267 is exclusively expressed in the nucleus of GM12878 lymphoblastoid cells (Figure S1). We also investigated a potential effect of *CCAT2* on *MYC* messenger RNA stability (Figure S8), as well as an RNA enhancer effect [27] (Figure S9 and Supplemental Results), but could not find any significant variations related to *CCAT2* levels.

Aberrant WNT activation occurs in the majority of CRCs [28], and risk allele of SNP rs6983267 has been shown to confer enhanced response to WNT signaling through TCF7L2 binding [12]. Since *MYC* is an established target of TCF7L2 as confirmed by the presence of TCF7L2 binding sites in its promoter as well as by functional studies [29], it is possible that *CCAT2* affects *MYC* expression through the mediation of TCF7L2. Consistent with previous findings, we identified TCF7L2 binding to the *MYC* promoter region, and observed enhanced binding in the OC1 clone when compared with E1 clone (Figure 5C). Next, we sought to determine whether this was due to an increase in nuclear CTNNB1 or TCF7L2 accumulation. Using Western blot analysis, we observed a comparable level of total CTNNB1 expression levels with no increase in nuclear CTNNB1 between *CCAT2* overexpressing clones and their control cells (Figure S10). Similarly, TCF7L2 expression levels were comparable in all four clones (Figure S10). Notably, by using a TOP-Flash luciferase assay, we observed a 2-fold increase in WNT reporter activity in OC1 and OC2 clones with increased *CCAT2* expression levels (Figure 5D). Increased luciferase activity was also observed in an independently generated set of high-*CCAT2*-expressing HCT116 clones (Figure S11). This result was further confirmed in HEK293 WNT reporter cells transiently transfected with *CCAT2* expression vector (Figure 5E). Consistent with these observations, WNT reporter activity in COLO320 cells was effectively reduced by *CCAT2* siRNA and also as expected by *TCF7L2* siRNA (Figure S12). These findings suggest that *CCAT2* augments TCF7L2 transcriptional activity without increasing the quantity of TCF7L2.

We employed RNA immunoprecipitation analysis to examine the physical interaction between *CCAT2* and TCF7L2 by pulling down RNAs co-localized with TCF7L2 protein. We found an enrichment of *CCAT2* RNA, but not *GAPDH* and the *HOTAIR* ncRNA by TCF7L2 antibody compared with the IgG control in the OC1 clone (Figures 5F and S13). To test whether endogenous *CCAT2* binds to TCF7L2 protein, we analyzed the RNA-protein

association in COLO320 cells that express high *CCAT2* levels. Similar to the findings in the OC1 HCT116 clone, a significant enrichment of *CCAT2* was observed in TCF7L2 pull-down RNA samples from COLO320 cells (Figures 5G and S13).

Overall, these results suggest that the main mechanism through which *CCAT2* increases *MYC* expression relies on activation of WNT signaling by enhancing TCF7L2 transcriptional activity. Accordingly, the expression of specific WNT target genes such as *CD44* [30] and *VIM* [31] was increased in *CCAT2*-overexpressing clones (Figure 5H and 5I). Conversely, knockdown of *CCAT2* expression caused a reduction in the expression level of *VIM* in both KM12SM (with high endogenous *CCAT2* expression) and COLO320 (with very high endogenous *CCAT2* expression) cells (Figure S14).

CCAT2 RESPONDS TO WNT SIGNALING

Previously, a TCF7L2 binding element was found to be located within the rs6983267 genomic span [12,14], which suggested that *CCAT2* expression levels might be regulated by this transcription factor. Therefore, we treated three colon cancer cell lines (COLO320, HCT116, and RKO) with lithium chloride (LiCl) in order to further activate canonical WNT signaling through inhibition of glycogen synthesis kinase 3 beta (GSK3B) [12]. The effect of LiCl was monitored by TOP-Flash luciferase activity and by *AXIN2* expression [32]. In all these cell lines, we observed an increase in *CCAT2* expression after LiCl treatment (Figure 6A). *TCF7L2* knockdown using two different siRNAs impaired the LiCl-induced *CCAT2* expression increase in HCT116 cells (Figures 6B and S15), suggesting that the effect on WNT signaling was TCF7L2-dependent. In the *CCAT2*-overexpressing clones, *TCF7L2* siRNA consistently reduced *CCAT2* expression, either with or without LiCl stimulation, indicating that TCF7L2 is necessary for the maintenance of high *CCAT2* expression levels (Figure 6C). The endogenous expression of *CCAT2* was also significantly down-regulated by *TCF7L2* siRNA in KM12SM and COLO320 cells (Figures 6D and 6E). These findings suggest the existence of a complex feedback loop of *CCAT2* and WNT signaling.

RS6983267 ALLELES AFFECT *CCAT2* EXPRESSION AND FUNCTION

The discovery of *CCAT2* and its role in CRC suggests that it may underlie the mechanism of the rs6983267 SNP-conferred cancer risk. We first test if rs6983267 status affects *CCAT2* expression in patient samples. We genotyped the rs6983267 SNP in previously profiled MSS samples (Italian, Japan and Australian cohorts, Figure 1B), and separated them into GG, GT and TT groups. We did not observe significant difference (P=0.450) in *CCAT2* expression levels between the GG (n=52) and TT (n=45) groups (Figure S16). However, further separation of the groups according to the individual cohort revealed that the difference (P=0.009) is significant in the Italian cohort, which has the largest sample size (Figure 6F). Inconsistently, there was no significant difference in *CCAT2* expression between the GG and TT alleles in 93 CRC patients from the United Kingdom (data not shown). These data suggest that the effect of rs6983267 allele on *CCAT2* expression levels may be very subtle and for detection better-controlled experimental design are needed.

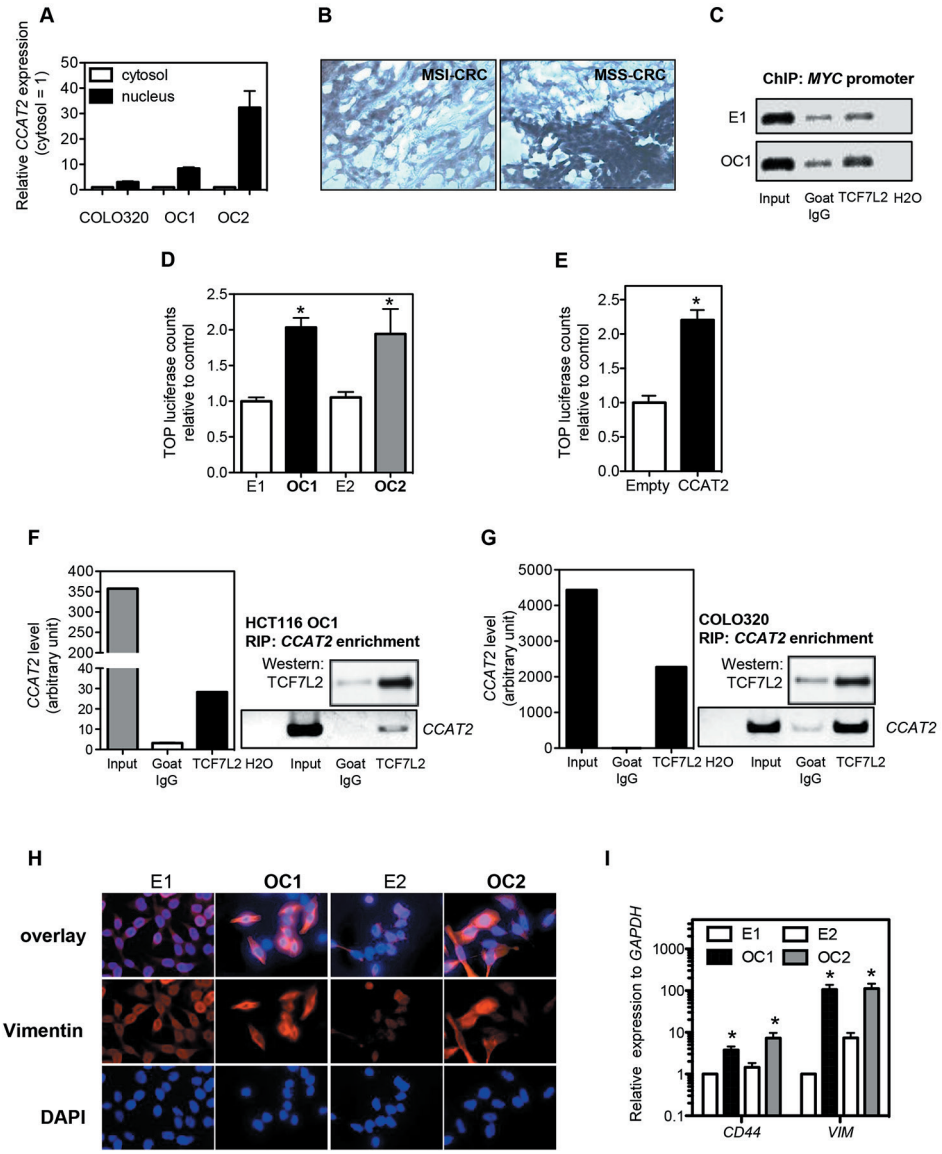


Figure 5. Regulation of TCF7L2 activity by *CCAT2*.
(**A)** *CCAT2* nuclear localization, as identified using qRT-PCR in fractionated COLO320 cells and HCT116 *CCAT2* clones. (**B)** *CCAT2* nuclear localization, as identified in CRC samples using *in situ* hybridization. (**C)** *CCAT2* increased TCF7L2 binding to the *MYC* promoter. ChIP analysis showed higher binding in *CCAT2*-overexpressing clone (OC1) than cells with basal *CCAT2* expression (E1). Goat IgG was used as a control antibody. (**D**) Higher WNT activity in *CCAT2* stable clones, as identified by TOP-Flash reporter assay. (**E)** Transient expression of *CCAT2* induced a more than 2-fold increase of TOP-Flash luciferase activity in HEK293 cells. Data in C and D are presented as mean ± SEM (n=3). * $P<0.05$ **F)** Co-localization of *CCAT2* RNA with TCF7L2 protein by RNA immunoprecipitation using anti-TCF7L2 antibodies in OC1 with

Figure 5. *Continued*
ectopic *CCAT2* overexpression. **(G)** Co-localization of *CCAT2* RNA with TCF7L2 protein by RNA immunoprecipitation using anti-TCF7L2 in COLO320 cells, which have endogenous high *CCAT2* expression. A quantitative analysis using qRT-PCR is shown along with the PCR gel images. **(H)** Immunostaining of HCT116 clones with vimentin antibody or DAPI. Stronger vimentin signal as seen in epithelial-mesenchymal transition was observed in *CCAT2*-overexpressing clones than in cells with basal *CCAT2* expression. **(I)** *CCAT2* increases the mRNA expression levels of *CD44* and *VIM* in HCT116 clones. Data are presented as mean ± SEM from at least three independent experiments. *$P<0.05$.

To address this question, we took advantage of CRC cell lines (DLD1 and KM12SM cells) with heterogeneous rs6983267 genotype, to measure the amount of *CCAT2* transcript produced from different alleles. We amplified the region spanning the SNP using genomic DNA or RNA as template, and sequenced the PCR products by pyrosequencing, a method that allows for accurate and quantitative analysis of DNA sequence [33]. In both DLD1 and KM12SM cells, the percentage of G allele is slightly (about 6%), but significantly higher in RNA samples than the corresponding genomic DNA samples (Figure 6G). These data prompted us to determine the allele ratio in genomic DNA and RNA of CRC samples with GT genotype. As shown in Figure 6H, three out of four samples showed consistent findings – that is, significantly higher percentage (7-15%) of G allele in RNA than in DNA samples. These data suggest that G allele of rs6983267 produces more *CCAT2* transcripts than the T allele.

We next determined whether *CCAT2* transcripts containing G or T have different biological functions. We mutated the SNP site of the retrovirus *CCAT2* construct from G to T, and transfected HCT116 cells with *CCAT2*-G or *CCAT2*-T retroviruses. Although we observed higher *MYC* expression (both RNA and protein) in *CCAT2*-G than *CCAT2*-T transfected cells (Figures S17A and S17B), the difference becomes not significant after normalizing to *CCAT2* expression levels (Figure S17C). To test whether the rs6983267 allele affects *CCAT2*'s regulation of *MYC* in CRC patients, we analyzed *CCAT2* and *MYC* expression correlation in previously genotyped MSS CRC samples. A significant positive correlation between *CCAT2* and *MYC* expression was observed in GG samples (n=52, coefficient=0.3; P=0.001), but not in TT samples (n=45, coefficient=0.26; P=0.078) of CRCs (Figure S18). Furthermore, by using the Vienna RNA Secondary Structure Package (http://www.tbi.univie.ac.at/~ivo/RNA/), we identified that the *CCAT2* G and T alleles have distinct conformations (Figure S19). Although further studies are needed to verify the different biological functions of the rs6983267 G and T alleles, our data indicate that the specific alleles may also differentially affect *MYC* expression.

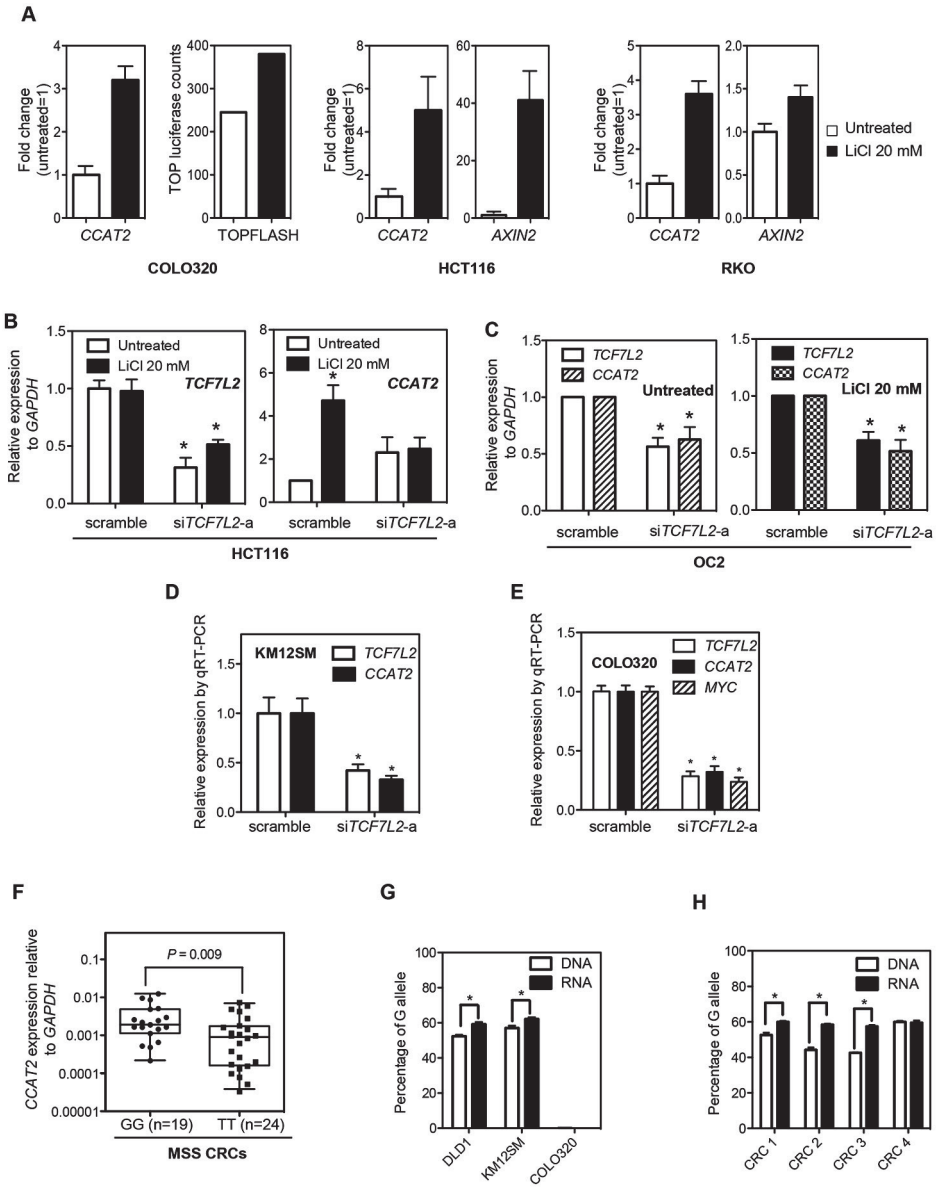


Figure 6. Regulation of *CCAT2* expression by WNT signaling and SNP variance on *CCAT2* expression.

(A) *CCAT2* expression is induced by lithium chloride (LiCl) in three colon cancer cell lines. *AXIN2* or TOP-Flash luciferase activity served as a positive control for the activation of WNT signaling. Data are presented as mean ± SD (n=3). **(B)** TCF7L2 is indispensible for LiCl-induced *CCAT2* expression. HCT116 cells were treated with *TCF7L2* siRNA (siGENOME SMARTpool TCF7L2, Dharmacon) for 24 h, and then stimulated with LiCl for another 24 h. *CCAT2* expression levels were measured by qRT-PCR. **(C)** *TCF7L2* siRNA downregulates *CCAT2* expression in overexpressing OC2 clone, either with or without LiCl stimulation. **(D, E)** *TCF7L2* siRNA (sc43525, Santa Cruz) downregulates *CCAT2* expression in KM12SM cells and

Figure 6. *Continued*
COLO320 cells. Data are presented as mean ± SEM (n = 3); * $P<0.05$ when compared with the respective control. **(F)** Comparison of *CCAT2* expression in cohort of CRC samples from Italy with GG and TT genotype showed higher *CCAT2* expression associated with the G allele. *P*-value was calculated using Mann-Whitney test. Data are presented as box-whisker plots. **(G, H)** Pyrosequencing data showed higher percentage of G allele in the *CCAT2* transcripts than its genomic DNA counterpart in heterogeneous rs6983267 cell lines (DLD1 and KM12SM) and CRC patient samples with GT genotype. COLO320, which is TT genotype, served here as a negative control.

DISCUSSION

We report the identification of *CCAT2*, a novel lncRNA encompassing the cancer-related rs6983267 SNP and located in a highly conserved genomic region enriched for markers of regulatory elements, such as H3K4me1, p300, and H3K27ac [11]. This lncRNA is expressed at high levels in MSS CRC tumors tissues and influences *MYC* levels, but has very low-level expression in normal colon tissues. This could explain how this new gene has escaped identification, until now, after genome-wide association studies identified the association between the SNP rs6983267 and cancer risk, as well as why tilling arrays did not identify it. A recent study showed that removal of the region encompassing the rs6983267 SNP confers resistance to intestinal tumors induced by the *APCmin* mutation in mice [34]. As we also cloned a novel transcript encompassing this SNP region from mouse tissue, it can be envisioned that deletion of the region also leads to removal of the *CCAT2* transcript. Thus, this study lends further support to the oncogenic function of *CCAT2* in CRC pathogenesis. In addition, the findings of decreased *MYC* expression and reduced TCF7L2 binding upstream of *MYC* in knockout mice from the same study are also consistent with our findings that *CCAT2* regulates *MYC* transcriptional activity.

CCAT2-enhanced invasion and metastasis appear to occur through the MYC-regulated miRNAs *miR-17-5p* and *miR-20a*. The role of MYC in metastasis has been controversial and a recent publication shows that, in breast cancer, MYC suppresses metastasis by direct transcriptional silencing of integrins [35]. Of interest, *miR-20a* was identified as being overexpressed in the transition from colon mucosa to early adenoma, and was predicted to target multiple WNT pathway genes [26]. The observation according to which *CCAT2*-enhanced invasion is mediated by MYC-regulated miRNAs and also possibly by other pro-metastatic targets such as *CDC25A* [36], rather than by *MYC* itself suggests a complex regulatory network that needs to be further explored in the future. This is also in agreement with the recent observation by the The Cancer Genome Atlas (TCGA) consortium that *MYC*-driven transcriptional activation and repression play important roles in colon cancer (The Cancer Genome Atlas Network, [37]). It is also possible that the effect of *CCAT2* on metastasis is mediated through its general activation of WNT signaling [38,39] or other unknown mechanism.

The TOP-Flash luciferase assay suggests increased WNT activity in *CCAT2*-over-expressing cells. Therefore, we postulated that the increased *MYC* transcription could result from increased WNT activity, as *MYC* itself represents a well-established WNT target [29]. Recent studies have identified RNA-protein interaction as one of the main regulatory mechanisms of lncRNAs [3]. Co-immunoprecipitation of *CCAT2* (RNA) and TCF7L2 (protein) suggests that *CCAT2* may broadly regulate WNT target genes by binding to

TCF7L2 and modulating its transcriptional activity. Various mechanisms can be envisaged to explain how the *CCAT2*/TCF7L2 association can enhance WNT activity. We speculate that *CCAT2* may affect TCF7L2-driven gene transcription through modulating association of TCF7L2 with its partners, which include the negative regulators Groucho/TLE and CTBP1 and the activator CTNNB1 [28], or changing TCF7L2 protein function by bridging TCF7L2 with other proteins in the transcription complexes. Alternatively, the association of *CCAT2* with TCF7L2 may modify the protein structure and enhancer its function. Regarding the mechanism of *MYC* regulation, it is also possible that *CCAT2* participates in the loop formation between the genomic loci of rs6983267 with *MYC* promoter [14] and works together with the enhancer element to activate *MYC* transcription. Notably, we also show that WNT signaling regulates *CCAT2* expression, suggesting a feedback loop between *CCAT2* and WNT signaling as previously shown for other targets such as *CDH1* [40] and *AXIN2* [41].

We demonstrated *CCAT2*'s ability to underlie chromosomal instability, manifested by the loss or gain of large portions or whole chromosomes, eventually resulting in aneuploidy, a characteristic trait of MSS CRCs. These data support a scenario in which, together with mutations in coding genes such as *APC* [42,43] and *BUB1* [44], *CCAT2* contributes to CIN in colon cancer. This is likely to occur through *CCAT2's* effects on downstream mediators such as *MYC* and/or other WNT target genes. The direct involvement of *MYC* in CIN has been previously demonstrated by showing that transient excess of *MYC* activity could elicit genomic instability and carcinogenesis [45]. Inversely, *MYC* deletion rescues APC deficiency in the mouse intestine [46]. Thus, it is plausible that *MYC* acts as a mediator of *CCAT2*-induced CIN phenotype. Furthermore, the *MIR17HG* miRNAs, known to be MYC regulated, can also affect CIN as evidenced by diffuse large B-cell lymphoma that has high levels of these miRNAs and non-random chromosomal abnormalities [47]. Alternatively, the CIN-inducing effect can result from the general activation of WNT signaling by *CCAT2*. This hypothesis is supported by previous findings that CTNNB1/TCF7L2-mediated transcription [48] or ectopic expression of the WNT target gene conductin (AXIN2) [49] directly induce CIN. However, it should be noted that although CIN can be the underlying mechanism for C*CAT2*-promoted tumor growth and metastasis, mechanistic connections of these phenotypes are not established in this study.

Finally, our pyrosequencing data on heterogeneous cell line models and CRC samples suggest that rs6983267 status affects the *CCAT2* expression, providing an additional mechanism linking the risk allele of rs6983267 with higher CRC risk. Although we demonstrated that WNT signaling regulates *CCAT2* expression, it remains to be determined if this accounts for the SNP-caused differential expression of *CCAT2*. In addition, in CRC patients with GG (but not TT) genotype, *CCAT2* shows significant association with *MYC* expression levels, suggesting the SNP status may also affect *CCAT2* function as a RNA transcript. Although most of the previous studies failed to identify a correlation between this SNP and *MYC* expression, a recent research found higher *MYC* expression in CRC samples containing GG alleles, probably due to the improved purity of the cancer tissues owing to the use of laser microdissection [50]. Our findings on SNP-induced differential *CCAT2* expression and function, together with the DNA enhancer element [12,14], provide an explanation on the cancer risk conferred by the rs6983267 risk allele (Figure 7).

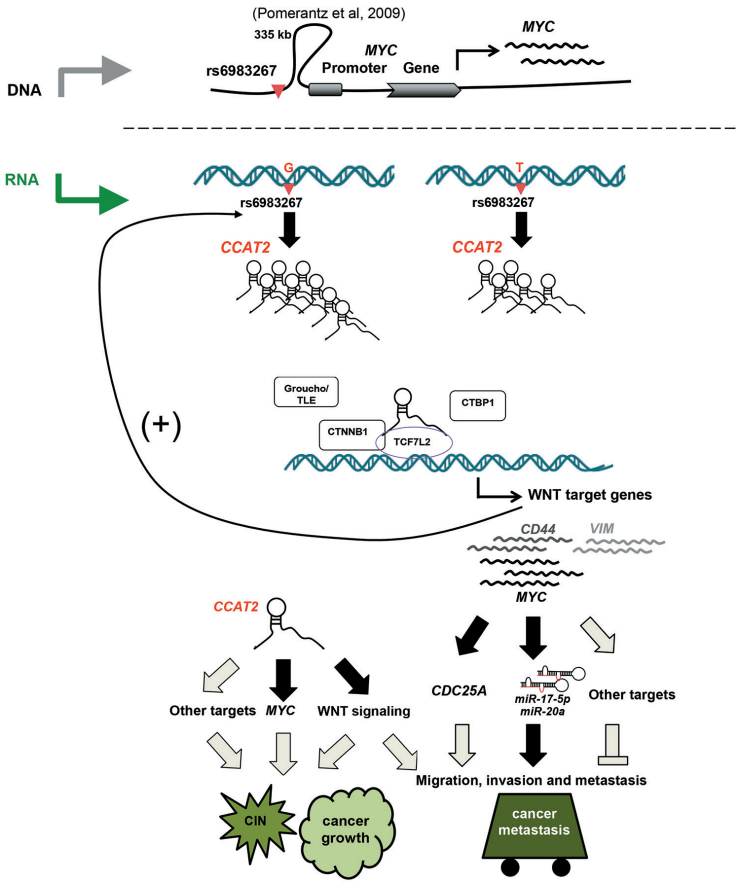


Figure 7 A model of *CCAT2* locus involvement in CRC. Upper panel. An approximately 335-kb DNA loop brings the rs6983267 genomic region close to the *MYC* locus, and this physical association may contribute to the enhancer function of the SNP-containing region on *MYC* transcription [14]. Lower panel. The enhancer region is transcribed into a long non-coding RNA (*CCAT2*), and the SNP status affects *CCAT2* expression by an as-yet-unknown mechanism. The *CCAT2* transcript upregulates WNT activity, and increases expression levels of WNT target genes (including *MYC*). This regulation by *CCAT2,* possibly through its physical interaction with TCF7L2, may lead to genomic instability and promote cell growth. We demonstrated that MYC-regulated *miR-17-5p* and *miR-20a* participate in the *CCAT2*-enhanced cell invasion, and speculate that other mechanisms, such as MYC-related mechanisms (*CDC25A*) or enhanced WNT signaling (*VIM* and *CD44*), may exist to coordinate the metastatic phenotype elicited by *CCAT2*. Finally, we demonstrated that *CCAT2* expression is regulated by transcriptional factors TCF7L2, indicating a positive feedback loop between *CCAT2* and WNT signaling. Our findings provide an additional explanation on the SNP-conferred CRC risk. Black: demonstrated; Grey: hypothesized interactions.

The discovery of *CCAT2* introduces a novel mechanism for the involvement of ncRNAs in WNT signaling and colon cancer pathogenesis. The recent findings showing that the CIN phenotype is a predictive marker for survival and may be used to select high-

risk patients with stages II and III CRC [51], combined with the specific involvement of *CCAT2* in MSS CRCs indicate that *CCAT2* has the potential to become a useful diagnostic and/or prognostic marker, although further validation in larger cohorts and by multiple independent studies is necessary.

METHODS

PATIENT SAMPLES

One hundred and ninety-one colorectal cancer samples and seventy non-neoplastic mucosa samples were used in this study. The tumor and control samples were obtained from three different sources: an Italian cohort from the University of Ferrara (CRC=81 and normal colon mucosa=25), a Japanese cohort from the University of Kyushu (CRC=45 and paired normal colon mucosa=45), an Australian cohort from Princess Alexandra Hospital (CRC=65), and a C-CFR cohort from Northern American colon cancer patients with MSS, CIMP-negative colon cancer at three participating centers of the Colorectal Cancer Family Registry (Mayo Clinic, Rochester, MN, USA; Mount Sinai Hospital, Ontario, Canada; and Cleveland Clinic, Cleveland, Ohio, USA) (CRC=24 and paired normal colon mucosa=24). Tissue samples were obtained from fresh surgical specimens frozen in liquid nitrogen and stored at -80°C. All the samples were obtained with the patients' informed consent and were histologically confirmed.

Breast tumors were selected from the tumour bank at the Erasmus Medical Center (Rotterdam, The Netherlands). The frozen tumor samples were collected from female patients with breast cancer who entered the clinic from 1979 to 1998 and for whom detailed clinical follow-up data and tumor RNA for *CCAT2* analysis were available [19,52]. The protocol for studying biological markers associated with disease outcome was approved by the medical ethics committee of the Erasmus Medical Center (MEC 02.953).

The British CRC samples were derived from a set of CRC patients recruited as part of the CORGI study who had known genotypes at the 8q24 tagSNP [8].

RNA ISOLATION, CDNA SYNTHESIS, AND QUANTIFICATION OF SPECIFIC RNA SPECIES

RNA was extracted using the TRIzol protocol (Invitrogen) and DNase digested (Ambion); the quality of the RNA was assessed using agarose gel electrophoresis. cDNA was synthesized using the Superscript III cDNA kit (Invitrogen), and diluted cDNA was used for RT-PCR analysis using iQ SYBR Green Supermix (Bio-Rad) with the appropriate primers (Table S2). The $2^{-\text{delta Ct}}$ method was used to calculate the relative abundance of RNA genes compared with *GAPDH* expression. qRT-PCR analysis for miRNAs was performed with the TaqMan MicroRNA Assay kit (Applied Biosystems). Primers and probes for *miR-17*, *miR-20a*, *miR-146a*, *let-7a*, and *U6* snRNA were purchased from Applied Biosystems. The $2^{-\text{delta Ct}}$ method was used to calculate the relative abundance of miRNA genes compared with *U6* expression.

CELL LINES

COLO320DM, HCT116, RKO, and HEK293 cells were obtained from the American Type Culture Collection and validated by the Characterized Cell Line Core at The University of Texas MD Anderson Cancer Centre using STR DNA fingerprinting.

STABLE CELL LINE GENERATION

Stable HCT116 cells for *CCAT2* overexpression experiments were produced by pMX retrovirus or pcDNA3.1 expression vector transfection. For the pcDNA3.1 single clones, we screened *CCAT2* expression in the HCT116 cells transfected with *CCAT2* construct and randomly chose two clones with high *CCAT2* expression (OC1 and OC2), and one clone with basal level of *CCAT2* expression (E2). E1 was originated from the HCT116 cells transfected with the empty pcDNA3.1 vector.

COLO320 stable cells for *CCAT2* knockdown experiments were generated using pRS-Retro-GFP/neo vector retrovirus (OligoEngine) or the Knockout™ Single Vector Inducible RNAi System (Clontech, CA).

VIRUS PRODUCTION

The *CCAT2*-containing genomic region was amplified with a primer combination (*CCAT2* F2 and R2) from HCT116 genomic DNA with Pfu polymerase (Invitrogen). After sequencing of the PCR product, we cloned it into the pMX plasmid. Then, we produced *CCAT2*-containing retrovirus in 293 GP2 cell lines and used virus-containing supernatant to infect HCT116 cells. After infection, HCT116 cells were grown in complete media containing puromycin (1 mg/mL). The shRNAs against *CCAT2* (Table S2) were inserted into the pRS-Retro-GFP/neo vector according to the manufacturer's protocol (OligoEngine).

PLASMID PRODUCTION

The sequence used for retrovirus production was also cloned into a pcDNA3.1+ vector. The pcDNA *CCAT2* plasmid was linearized with BglII and transfected into HCT116 cells using lipofectamine. Selection with G418 (0.5 mg/mL) was carried out until single colonies were generated. The tet-on inducible sh*CCAT2* clones (Table S2) were generated using the Knockout™ Single Vector Inducible RNAi System according to the manufacturer's instructions.

CCAT2 IN SITU HYBRIDIZATION

The frozen tissue slides of two MSS CRCs, two MSI-H CRCs, and two normal mucosae were incubated with a double-DIG-labeled *CCAT2* probe or control *U6* snRNA probe (Exiqon) and detected with a polyclonal anti-DIG antibody and alkaline phosphatase-conjugated secondary antibody (Ventana) using NBT-BCIP as the substrate. The signal intensities of *CCAT2* and *U6* expression were quantified by using the intensity measurement tools of the Image-Pro Plus software package (Media Cybernetics).

IN VIVO TUMORIGENIC ASSAY

Swiss nu-nu/Ncr nude mice (spontaneous mutant T-cell-deficient mice) were obtained from the Animal Production Area of the Department of Experimental Radiation Oncology at MD Anderson Cancer Center. The mice were maintained under specific pathogen-free conditions and used in accordance with institutional guidelines. HCT116 cells (1×10^6) were injected subcutaneously into the right flanks of the mice, and tumor size was measured weekly for 4 wks.

IN VIVO METASTASIS ASSAY

A midline incision was made on the left side of the flank, and the spleen was exteriorized. HCT116 cells (1×10^6 cells) were injected into the spleen, after which the wound was

closed with surgical metal clips. The mice were sacrificed after 8 wks, and their spleens and livers were removed and examined for tumor metastases on the surface of the liver. The liver specimens were formalin fixed and paraffin embedded for histological analysis.

IN VITRO MIGRATION ASSAYS

We used two assays: in the first, HCT116 cells (200,000 cells in 200 mL serum-free media) were seeded into a Transwell insert in a 24-well format (8 mm, BD Biosciences). To the bottom part of the well, 800 mL of medium with serum was added. Twenty-four hours after seeding, cells were fixed with paraformaldehyde and stained with crystal violet/methanol solution. Cells at the top of the Transwell were removed using a cotton swab, and photos of the Transwells were taken using Geldoc (Bio-Rad). The intensity of the crystal violet staining in each Transwell was evaluated using Photoshop software (Adobe Systems, Inc.) and compared among the different treatments.

The second type of assay, a cell migration assay, was performed according to the protocol previously described [53]. Cells were suspended in serum-free media (65,000 cells/insert), and seeded onto the gelatin-coated inserts. The cells that migrated to the bottom of the wells were fixed, stained, and counted using a microscope. For each well, 10 different fields were counted and the average number of cells was determined. Both types of assays were performed in triplicate experiments.

SIRNA TREATMENT

Cells were seeded onto 24-well or 6-well plate. After reaching 50% confluence, cells were transfected with 50 nM of siRNA by lipofectamin 2000 (Life Technologies) for 48 h before collecting RNA or protein for analysis.

GENOMIC INSTABILITY ANALYSIS

Cells were exposed to Colcemid (0.04 μg/mL) for 25 min at 37°C and to hypotonic treatment (0.075 M KCl) for 20 min at room temperature and then were fixed in a methanol and acetic acid mixture (3:1 by volume) for 15 min and washed three times in the fixative. The slides were air-dried, stained in 4% Giemsa, and coded for the blind analysis. Later, the slides were decoded for the evaluation of results. Slides were analyzed for several parameters, including chromosome aberrations (as evidenced by both chromosome- and chromatid-type breaks), fragments, tetraploidy, fusions, and formation of tri-radials.

SPECTRAL KARYOTYPING

Spectral karyotyping was performed according to the manufacturer's protocol using Human Paint probes (Applied Spectral Imaging [ASI], Inc.). Images were captured using a Nikon 80i microscope equipped with Spectral karyotyping software from ASI.

CREST STAINING

After fixation, cells were incubated for 1 h at room temperature with human serum from scleroderma patients (kindly provided by Dr. Maureen Mayes from The University of Texas – Houston Medical School) that produced, during the course of their disease, anti-centromere antibodies used for staining. Staining was visualized using FITC-conjugated donkey anti-human antibody. Beta-tubulin was detected using Cy3-conjugated antibody (Sigma).

SUBCELLULAR FRACTIONATION

The separation of nuclear and cytosolic fractions was performed using the PARIS™ Kit (Life Technologies) according to the manufacturer's instructions.

RNA IMMUNOPRECIPITATION

We used the Magna RIP™ Kit (Millipore, Billerica, MA) according to the manufacturer's instructions. Cells were prepared in RIP lysis buffer, and the RNA-protein complexes were immunoprecipitated using anti-TCF7L2 or normal goat IgG (control). RNA was purified using phenol:chloroform:isoamyl alcohol and subjected to reverse transcription-PCR or real-time PCR analysis (Table S2). Control amplification was carried out on input RNA before immunoprecipitation.

TOP-FLASH WNT REPORTER

Cells were transfected with 250 ng of the TOP-FLASH or FOP-FLASH reporter constructs together with 25 ng of the *Renilla* luciferase vector. Luciferase activity was measured by the Dual-Luciferase® Reporter Assay System (Promega) 48 h after transfection and TOP or FOP values were normalized to *Renilla* values. The TOP/FOP ratios were calculated and used as indicators of the endogenous level of WNT signaling.

IMMUNOFLUORESCENT ASSAY

Cells were seeded on a 96-well plate. After 48 h, the cells were permeabilized and fixed in 4% paraformaldehyde in PBS buffer at room temperature for 15 min. The cells were then incubated with anti-vimentin (V9, Novus) overnight, and finally incubated with secondary antibody and DAPI for 1 h. All matched samples were photographed using an immunofluorescence microscope and identical exposure times. Each experiment was performed in triplicate and repeated three times.

ALLELE QUANTIFICATION BY PYROSEQUENCING

We used the Pyrosequencing technology (Qiagen) to measure the frequency of the G and T alleles of the *CCAT2* transcript. The principles of the method and analysis have been previously reported [54]. In our study, nested PCR reactions were done for each sample. The first PCR done using the primers 1-F (TTTAGCAGCTGCATCGCTCCATAG) and 2-R (CTCCCTCCCCCACATAAAAT). The second PCR combined a universal biotinylated primer for subsequent single-strand DNA capture for the pyrosequencing reaction. One microliter of the first PCR was used for the nested amplification, and a reaction was done to capture the forward strand using the primers SNP/F1 (GGGACGAATAAACTCTCCTCCTAC), SNP/RU1 (GGGACACCGCTGATCGTTTACTCCCTCCCCCACATAAAAT) and the universal biotinylated primers UBP (Bio-GGGACACCGCTGATCGTTTA). An independent PCR reaction was done to capture the reverse strand using the primers SNP/FU2 (GGGACACCGCTGATCGTTTAGAGGTGTAGCCAGAGTTAATACCC), SNP/R2 (CTCCCTCCCCCACATAAAAT) and UBP (Bio-GGGACACCGCTGATCGTTTA). After clean-up in a vacuum prep station, the single-strand DNA was combined with sequencing primers to read the forward (SNP/S1, AGCTCAGCAGATGAAA and the sequence to analyze GG/TCACTG) and reverse (SNP/S2, TCTTTGTACTTTTCTCAGTG and the sequence to analyze A/CCTTTCATC) strands of the target SNP. Each measurement was done in duplicate and the values from the forward and reverse strands was combined to give the frequency of the G and T alleles.

STATISTICAL ANALYSIS

For the colorectal cancer samples, the statistical analyses were performed in R (version 2.14.2). The Shapiro-Wilk test was applied to determine whether data followed a normal distribution. Accordingly, the t-test or the nonparametric Mann-Whitney-Wilcoxon test was applied to assess the relationship between *CCAT2* expression levels and clinical parameters. Fisher's exact test was used to assess the association between *CCAT2* and the number of metastasis in mice. For the breast cancer cohorts, data were analyzed using the STATA statistical package, release 9 (STATA Corp.), and SPSS 15.0 (IBM Corporation). We investigated the prognostic value of the clinical and biological variables with metastasis-free survival (MFS) as the endpoint using Cox univariate and multivariate regression analyses. *CCAT2* RNA levels were analyzed as a log-transformed continuous variable and also, for visualization in Kaplan-Meier survival analysis, as a dichotomized variable based on the median level of *CCAT2* RNA, with the log-rank test used to evaluate differences. Spearman's rank-order correlation test was applied to measure the strength of association between the expressions of *CCAT2* and *MYC* in colorectal cancer samples and between *CCAT2* RNA and CIN score in breast cancer samples. All tests were 2-sided, and $P<0.05$ was considered statistically significant.

DATA ACCESS

The *CCAT2* RNA sequence reported in this paper has been deposited in GenBank (http://www.ncbi.nlm.nih.gov/nuccore/GQ911591.1).

ACKNOWLEDGMENTS

Dr. Calin is the Alan M. Gewirtz Leukemia & Lymphoma Society Scholar. He is also supported as a Fellow at The University of Texas MD Anderson Research Trust, as a University of Texas System Regents Research Scholar, and by the CLL Global Research Foundation. Work in Dr. Calin's laboratory is supported in part by the NIH/NCI (CA135444), a Department of Defense Breast Cancer Idea Award, Developmental Research Awards in Breast Cancer, Ovarian Cancer, Brain Cancer, Prostate Cancer, Multiple Myeloma, and Leukemia SPOREs, an MDACC Sister Institution Network Fund grant in colorectal cancer, the Laura and John Arnold Foundation, the RGK Foundation, and the Estate of C.G. Johnson, JR. Dr. Ling is an Odyssey Fellow, and his work is supported in part by the Odyssey Program and The Estate of C. G. Johnson, JR at The University of Texas MD Anderson Cancer Center. Drs. Fodde and Atlasi were supported by grants from the Dutch Cancer Society (EMCR 2001-2482), the Netherlands Organisation for Scientific Research (NWO/Vici 016.036.636), the BSIK (Kennisinfrastructuur) program of the Dutch Government grant 03038 (www.stemcells.nl), the Netherlands Institute for Regenerative Medicine (NIRM; www.nirm.nl), and the EU FP6 and FP7 consortia Migrating Cancer Stem Cells program (MCSCs; www.mcscs.eu) and TuMIC (integrated concept of tumor metastasis (http://itgmv1.fzk.de/www/tumic/tumic_main.htm). Drs. Mimori and Mori were supported by the Funding Program for Next Generation World-Leading Researchers (LS094) and Core Research for Evolutional Science and Technology, respectively. Dr. Hamilton is the recipient of the Frederick F. Becker Distinguished University Chair in Cancer Research from The University

of Texas. Genomic instability analysis was done in the MD Anderson Cancer Center Core laboratory supported by NCI grant #CA016672. Short tandem repeat DNA fingerprinting was done by the Cancer Center Support grant-funded Characterized Cell Line Core, NCI #CA16672. Allele quantification was done in the MD Anderson Cancer Center DNA Methylation Analysis Core. The C-CFR Cohort study was supported by the National Cancer Institute, National Institutes of Health under RFA # CA-95-011; and through cooperative agreements with the University of Hawaii Colorectal Cancer Family Registry (U01 CA074806), Mayo Clinic Cooperative Family Registry for Colon Cancer Studies (U01 CA074800), Ontario Registry for Studies of Familial Colorectal Cancer (U01 CA074783), the University of Southern California Familial Colorectal Neoplasia Collaborative Group (U01 CA074799), and members of the Colon Cancer Family Registry and P.I.s. In addition, this work was partly supported by grant 5U24 CA074806 and an award by Affymetrix Collaborations in Cancer Research Program. We also thank the participants of the Colon Family Registry who have contributed to a better understanding of the genetic contributions to colon cancer. The *CCAT2* expression analysis for C-CFR samples was done at the Genomics Shared Resource of the University of Hawaii Cancer Center, which is supported by 2P30CA071789-13. We thank Dawn Chalaire from the Department of Scientific Publications at The University of Texas MD Anderson Cancer Center for her help with the editing of this manuscript. We acknowledge Dr. Bryant G. Darnay, who kindly provided the pMX vector and offered assistance with the retrovirus production and Dr. Maureen Mayes for kindly providing serum from scleroderma patients.

AUTHOR CONTRIBUTIONS

G.A.C. and R.F. designed this study. H.L., R.S., R.F., S.R.H. and G.A.C. wrote the manuscript. R.S., H.L., Y.A., M.N., M. Shimizu, R.R., J.S., N.N., E.B., I.D.V., M.C., J.F.G., B.H.B., S.S., M. Shariati, F.H., V.B., A.S.M., J.W., M.A.S., W.Z., and A.M.S. performed most of the wet-lab experiments. R.G., G.A.P., D.C.G., K.Y., I.T., I.B-N., K.M., M.M., A.M.S., J.W., M.M., I.T., S.G., G.C., S.N.T., L.L.M., M.T, J.A.F., S.R.H, and G.L. provided the patient specimens. X.Z. designed probes and performed the ISH experiments, and S.R.H. assisted with interpretation. H.L., R.S., Y.A., M.F., C.I., X.Z., R.V.D.,A.M.S., J.A.F., M.N., S.K., R.F. and G.A.C. analyzed the data. All authors read and approved the manuscript's content.

DISCLOSURE DECLARATION

The authors declare no competing financial interests.

REFERENCES

1. Ambros V (2001) microRNAs: tiny regulators with great potential. Cell 107: 823-826.
2. Spizzo R, Nicoloso MS, Croce CM, Calin GA (2009) SnapShot: MicroRNAs in Cancer. Cell 137: 586-586 e581.
3. Rinn JL, Chang HY (2012) Genome Regulation by Long Noncoding RNAs. Annual review of biochemistry 81: 145-166.
4. Bejerano G, Pheasant M, Makunin I, Stephen S, Kent WJ, et al. (2004) Ultraconserved elements in the human genome. Science 304: 1321-1325.
5. Calin GA, Liu CG, Ferracin M, Hyslop T, Spizzo R, et al. (2007) Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. Cancer Cell 12: 215-229.
6. Wojcik SE, Rossi S, Shimizu M, Nicoloso MS, Cimmino A, et al. (2010) Non-codingRNA sequence variations in human chronic lymphocytic leukemia and colorectal cancer. Carcinogenesis 31: 208-215.
7. Haiman CA, Le Marchand L, Yamamato J, Stram DO, Sheng X, et al. (2007) A common genetic risk factor for colorectal and prostate cancer. Nat Genet 39: 954-956.
8. Tomlinson I, Webb E, Carvajal-Carmona L, Broderick P, Kemp Z, et al. (2007) A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat Genet 39: 984-988.
9. Ghoussaini M, Song H, Koessler T, Al Olama AA, Kote-Jarai Z, et al. (2008) Multiple loci with different cancer specificities within the 8q24 gene desert. Journal of the National Cancer Institute 100: 962-966.
10. Bertucci F, Lagarde A, Ferrari A, Finetti P, Charafe-Jauffret E, et al. (2012) 8q24 cancer risk allele associated with major metastatic risk in inflammatory breast cancer. PloS one 7: e37943.
11. Jia L, Landan G, Pomerantz M, Jaschek R, Herman P, et al. (2009) Functional enhancers at the gene-poor 8q24 cancer-linked locus. PLoS Genet 5: e1000597.
12. Tuupanen S, Turunen M, Lehtonen R, Hallikas O, Vanharanta S, et al. (2009) The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. Nat Genet 41: 885-890.
13. Sotelo J, Esposito D, Duhagon MA, Banfield K, Mehalko J, et al. (2010) Long-range enhancers on 8q24 regulate c-Myc. Proceedings of the National Academy of Sciences of the United States of America 107: 3001-3005.
14. Pomerantz MM, Ahmadiyeh N, Jia L, Herman P, Verzi MP, et al. (2009) The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet 41: 882-884.
15. Lin MF, Jungreis I, Kellis M (2011) PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics 27: i275-282.
16. Wang L, Park HJ, Dasari S, Wang S, Kocher JP, et al. (2013) CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. Nucleic Acids Res.
17. Lengauer C, Kinzler KW, Vogelstein B (1997) Genetic instability in colorectal cancers. Nature 386: 623-627.
18. Camps J, Morales C, Prat E, Ribas M, Capella G, et al. (2004) Genetic evolution in colon cancer KM12 cells and metastatic derivates. Int J Cancer 110: 869-874.
19. Smid M, Hoes M, Sieuwerts AM, Sleijfer S, Zhang Y, et al. (2011) Patterns and incidence of chromosomal instability and their prognostic relevance in breast cancer subtypes. Breast cancer research and treatment 128: 23-30.
20. Mitchell KO, Ricci MS, Miyashita T, Dicker DT, Jin Z, et al. (2000) Bax is a transcriptional target and mediator of c-myc-induced apoptosis. Cancer Res 60: 6318-6325.
21. Galaktionov K, Chen X, Beach D (1996) Cdc25 cell-cycle phosphatase as a target of c-myc. Nature 382: 511-517.
22. Zindy F, Eischen CM, Randle DH, Kamijo T, Cleveland JL, et al. (1998) Myc signaling via the ARF tumor suppressor regulates p53-dependent apoptosis and immortalization. Genes Dev 12: 2424-2433.
23. O'Donnell KA, Wentzel EA, Zeller KI, Dang CV, Mendell JT (2005) c-Myc-regulated microRNAs modulate E2F1 expression. Nature 435: 839-843.

24. Chang TC, Yu D, Lee YS, Wentzel EA, Arking DE, et al. (2008) Widespread microRNA repression by Myc contributes to tumorigenesis. Nature genetics 40: 43-50.
25. Zhang J, Xiao Z, Lai D, Sun J, He C, et al. (2012) miR-21, miR-17 and miR-19a induced by phosphatase of regenerating liver-3 promote the proliferation and metastasis of colon cancer. British journal of cancer 107: 352-359.
26. Bartley AN, Yao H, Barkoh BA, Ivan C, Mishra BM, et al. (2011) Complex patterns of altered MicroRNA expression during the adenoma-adenocarcinoma sequence for microsatellite-stable colorectal cancer. Clinical cancer research : an official journal of the American Association for Cancer Research 17: 7283-7293.
27. Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, et al. (2010) Long noncoding RNAs with enhancer-like function in human cells. Cell 143: 46-58.
28. Clevers H (2006) Wnt/beta-catenin signaling in development and disease. Cell 127: 469-480.
29. He TC, Sparks AB, Rago C, Hermeking H, Zawel L, et al. (1998) Identification of c-MYC as a target of the APC pathway. Science 281: 1509-1512.
30. Wielenga VJ, Smits R, Korinek V, Smit L, Kielman M, et al. (1999) Expression of CD44 in Apc and Tcf mutant mice implies regulation by the WNT pathway. The American journal of pathology 154: 515-523.
31. Gilles C, Polette M, Mestdagt M, Nawrocki-Raby B, Ruggeri P, et al. (2003) Transactivation of vimentin by beta-catenin in human breast cancer cells. Cancer research 63: 2658-2664.
32. Jho EH, Zhang T, Domon C, Joo CK, Freund JN, et al. (2002) Wnt/beta-catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. Molecular and cellular biology 22: 1172-1183.
33. Fakhrai-Rad H, Pourmand N, Ronaghi M (2002) Pyrosequencing: an accurate detection platform for single nucleotide polymorphisms. Hum Mutat 19: 479-485.
34. Sur IK, Hallikas O, Vaharautio A, Yan J, Turunen M, et al. (2012) Mice Lacking a Myc Enhancer That Includes Human SNP rs6983267 Are Resistant to Intestinal Tumors. Science.
35. Liu H, Radisky DC, Yang D, Xu R, Radisky ES, et al. (2012) MYC suppresses cancer metastasis by direct transcriptional silencing of alpha(v) and beta(3) integrin subunits. Nature cell biology 14: 567-574.
36. Feng X, Wu Z, Wu Y, Hankey W, Prior TW, et al. (2011) Cdc25A regulates matrix metalloprotease 1 through Foxo1 and mediates metastasis of breast cancer cells. Mol Cell Biol 31: 3457-3471.
37. (2012) Comprehensive molecular characterization of human colon and rectal cancer. Nature 487: 330-337.
38. Nguyen DX, Chiang AC, Zhang XH, Kim JY, Kris MG, et al. (2009) WNT/TCF signaling through LEF1 and HOXB9 mediates lung adenocarcinoma metastasis. Cell 138: 51-62.
39. Yu M, Ting DT, Stott SL, Wittner BS, Ozsolak F, et al. (2012) RNA sequencing of pancreatic circulating tumour cells implicates WNT signalling in metastasis. Nature.
40. Heuberger J, Birchmeier W (2010) Interplay of cadherin-mediated cell adhesion and canonical Wnt signaling. Cold Spring Harb Perspect Biol 2: a002915.
41. Lustig B, Jerchow B, Sachs M, Weiler S, Pietsch T, et al. (2002) Negative feedback loop of Wnt signaling through upregulation of conductin/axin2 in colorectal and liver tumors. Mol Cell Biol 22: 1184-1193.
42. Fodde R, Kuipers J, Rosenberg C, Smits R, Kielman M, et al. (2001) Mutations in the APC tumour suppressor gene cause chromosomal instability. Nat Cell Biol 3: 433-438.
43. Kaplan KB, Burds AA, Swedlow JR, Bekir SS, Sorger PK, et al. (2001) A role for the Adenomatous Polyposis Coli protein in chromosome segregation. Nat Cell Biol 3: 429-432.
44. Cahill DP, Lengauer C, Yu J, Riggins GJ, Willson JK, et al. (1998) Mutations of mitotic checkpoint genes in human cancers. Nature 392: 300-303.
45. Felsher DW, Bishop JM (1999) Transient excess of MYC activity can elicit genomic instability and tumorigenesis. Proc Natl Acad Sci U S A 96: 3940-3944.
46. Sansom OJ, Meniel VS, Muncan V, Phesse TJ, Wilkins JA, et al. (2007) Myc deletion rescues Apc deficiency in the small intestine. Nature 446: 676-679.

47. Li C, Kim SW, Rai D, Bolla AR, Adhvaryu S, et al. (2009) Copy number abnormalities, MYC activity, and the genetic fingerprint of normal B cells mechanistically define the microRNA profile of diffuse large B-cell lymphoma. Blood 113: 6681-6690.

48. Aoki K, Aoki M, Sugai M, Harada N, Miyoshi H, et al. (2007) Chromosomal instability by beta-catenin/TCF transcription in APC or beta-catenin mutant cells. Oncogene 26: 3511-3520.

49. Hadjihannas MV, Bruckner M, Jerchow B, Birchmeier W, Dietmaier W, et al. (2006) Aberrant Wnt/beta-catenin signaling can induce chromosomal instability in colon cancer. Proceedings of the National Academy of Sciences of the United States of America 103: 10747-10752.

50. Takatsuno Y, Mimori K, Yamamoto K, Sato T, Niida A, et al. (2012) The rs6983267 SNP Is Associated with MYC Transcription Efficiency, Which Promotes Progression and Worsens Prognosis of Colorectal Cancer. Ann Surg Oncol.

51. Watanabe T, Kobunai T, Yamamoto Y, Matsuda K, Ishihara S, et al. (2012) Chromosomal Instability (CIN) Phenotype, CIN High or CIN Low, Predicts Survival for Colorectal Cancer. Journal of clinical oncology : official journal of the American Society of Clinical Oncology 30: 2256-2264.

52. Sieuwerts AM, Usher PA, Meijer-van Gelder ME, Timmermans M, Martens JW, et al. (2007) Concentrations of TIMP1 mRNA splice variants and TIMP-1 protein are differentially associated with prognosis in primary breast cancer. Clinical chemistry 53: 1280-1288.

53. Spannuth WA, Nick AM, Jennings NB, Armaiz-Pena GN, Mangala LS, et al. (2009) Functional significance of VEGFR-2 on ovarian cancer cells. International journal of cancer Journal international du cancer 124: 1045-1053.

54. Wang H, Elbein SC (2007) Detection of allelic imbalance in gene expression using pyrosequencing. Methods Mol Biol 373: 157-176.

55. Zhao X, Li C, Paez JG, Chin K, Janne PA, et al. (2004) An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. Cancer Res 64: 3060-3071.

SUPPLEMENTARY RESULTS

CCAT2 DID NOT AFFECT *MYC* MESSENGER RNA STABILITY AND HAD NO RNA ENHANCER EFFECT

To decipher the mechanism by which *CCAT2* regulates *MYC*, we first examined whether *CCAT2* increases *MYC* mRNA stability. To this aim, we treated cells with actinomycin to inhibit RNA synthesis and monitored time-course changes in *MYC* mRNA expression in the four clones with different *CCAT2* levels. We observed a trend of degradation of messenger RNA independent of *CCAT2* level of expression (Figure S8). Thus, the increased *MYC* levels observed with higher *CCAT2* expression may result from increased *MYC* transcription rather than from augmented RNA stability.

The nuclear localization of *CCAT2* suggests its potential role as a transcriptional enhancer for specific target genes. To test this possibility, we constructed vectors with inserts containing the *CCAT2* genomic region or a previously reported enhancer sequence encompassing rs6983267 cloned downstream of the luciferase gene driven by a thymidine kinase promoter (pGL3 TK) [14,27]. In all of these constructs, we identified greater enhancer activity than in the empty pGL3 TK vector (Figure S9A). We also observed a difference between the G and T plasmids, with G plasmids displaying greater enhancer activity (Figure S9A). The small 0.5-kb sequence containing *CCAT2* exhibited the strongest enhancer activity among all these plasmids, while the 1.7-kb sequence had the weakest enhancer function (Figure S9A). These findings may be explained by the existence of the transcriptional repressor CCCTC-binding factor (CTCF)-binding sites at the 1.7-kb sequence (Figure S9B), which may compromise the enhancer activity brought about by TCF7L2. In all the luciferase constructs, knockdown of the *CCAT2* transcript by siRNA did not cause a decrease in luciferase activity (data not shown). These data do not support the possibility that *CCAT2* has an enhancer-like function and refine the previously described enhancer activity of this DNA genomic locus [14].

Figure S1. RNA-seq data from ENCODE/CSHL showed that rs6983267 region is actively transcribed in the genomic sense orientation in GM12878 lymphoblastoid cells, and the transcript is exclusively located in the nucleus. pA: poly(A) signal; strand was indicated as + (sense strand) or – (antisense strand); cyt: cytoplasm; nlus: nucleus.

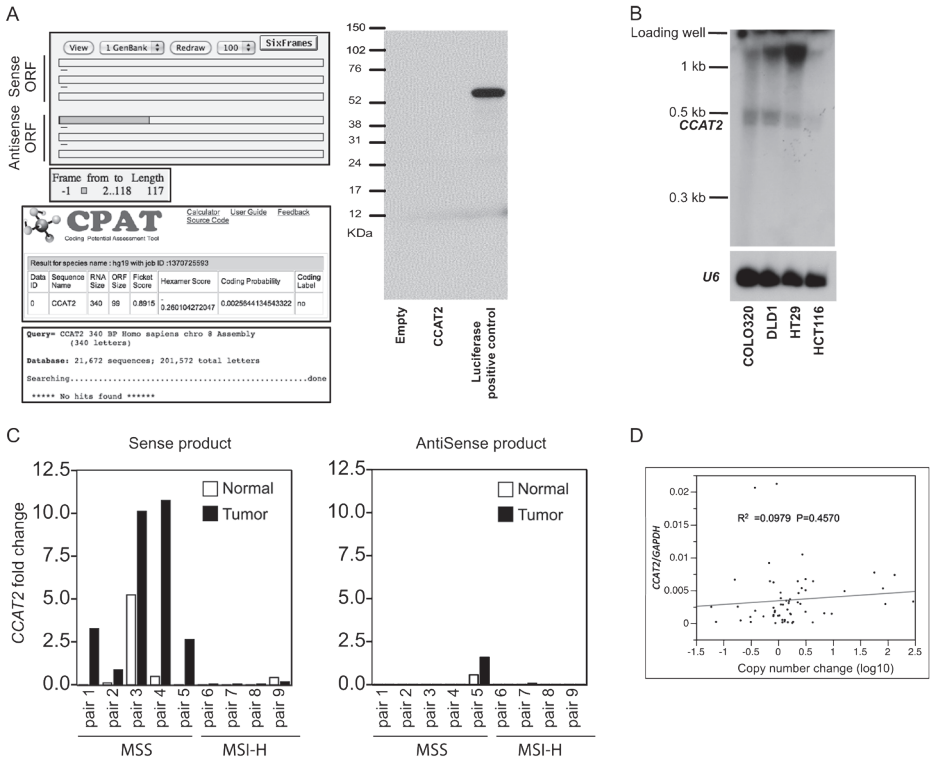


Figure S2. *CCAT2* identification and its specific expression in MSS CRCs. (**A**) *CCAT2* is a long non-coding RNA. Lack of open reading frame in the 340-bp *CCAT2* sense sequence cloned from RACE, as verified by ORF Finder (left upper) and CPAT (left middle, http://lilab.research.bcm.edu/cpat/), Pepbank (left lower, http://pepbank.mgh.harvard.edu/search/blast), and *in vitro* translation assay (right). In the translation assay, a 1752-bp genomic region containing *CCAT2* (amplified using primer *CCAT2* F2 and R2) was cloned into a pcDNA vector and expressed using the TnT Quick Coupled Transcription/Translation System (Promega). The absence of a specific band indicated that *CCAT2* is a transcript with no protein-coding capacity. Luciferase *in vitro* translation served as positive control. (**B**) Northern blot analysis of *CCAT2* expression in CRC cell lines. RNA was isolated with Trizol (Invitrogen) and run on a MOPS/formaldehyde denaturing agarose gel. RNA was then blotted on a nitrocellulose blot and hybridized with a P32-labelled DNA oligo probe. The band lower than 0.5 kb is the *CCAT2* signal, and the band higher than 1 kb is non-specific signal since it does not correlate with *CCAT2* expression levels as determined by qRT-PCR, where COLO320 cells have highest expression. (**C**) Sense-specific measurement of *CCAT2* expression relative to *GAPDH* in paired CRC patient samples using qRT-PCR identified a sense transcript that was overexpressed in all MSS tumor samples. RNA was reverse transcribed into strand-specific cDNA using primers that recognize either the sense or antisense strand and was subsequently subjected to qPCR amplification using primer *CCAT2* F1 and R1. Fold change to *GAPDH* normalizer is shown. (**D**) Lack of correlation between *CCAT2* expression levels and copy number change in CRC samples (n=64). We used genomic real-time PCR to detect genome amplification of the *CCAT2* region, with Line-1, a copy number indicator, as the normalizer. Copy number change was determined by using the formula [*CCAT2* (tumor genomic DNA)/*LINE*-1 (tumor genomic DNA)] / [*CCAT2* (normal genomic DNA)/*LINE*-1 (normal genomic DNA)] [55]. The expression of *CCAT2* RNA and amplification of its genomic locus at 8q24, as analyzed by Spearman's rank-order correlation test, are not significantly correlated (R^2=0.0979, *P*=0.457).

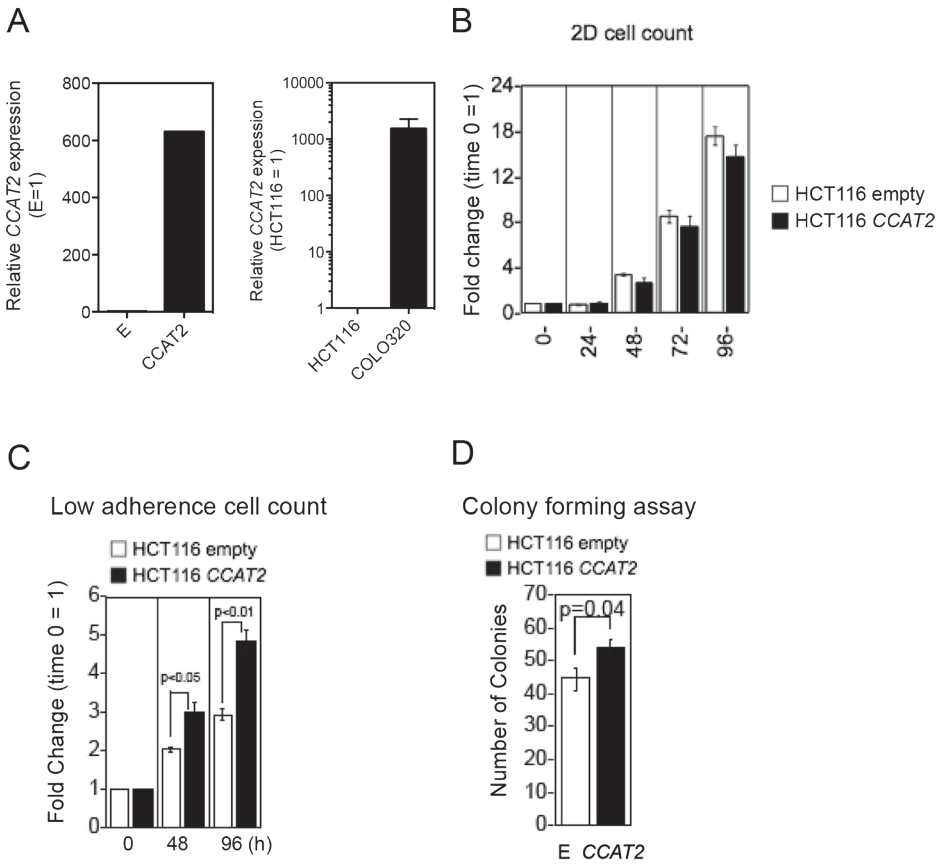


Figure S3. *CCAT2* promotes cancer cell growth *in vitro* under specific conditions. (**A**) Increased *CCAT2* expression in retrovirus-transfected HCT116 cells (left panel), which is comparable with the fold difference between HCT116 and COLO320 cells (right panel). (**B**) Lack of effect on 2D cell growth by *CCAT2*. (**C**) HCT116 cells transduced with *CCAT2* showed greater growth ability in a low-adherence condition than cells transduced with empty vector. Low adherence assay was performed in 6 well ultra low attachment plates. 48 and 96 hours after seeding, cells were collected and counted by cell counter (Vi-CELL Cell Viability Analyzer, Beckman Coulter, Irving, TX). This assay was performed two independent times, and each time three replicates were used for empty and *CCAT2* cells. (**D**) *CCAT2* enhanced the clonogenic ability of HCT116 cells. Cells (1000 cells in 100 mL of complete medium) were plated in 10cm dishes and incubated for 2 weeks. After the incubation, cells were stained by crystal violet/methanol solution, and three researchers counted colonies. The experiment was repeated three independent times, and each time three replicates were used for empty and *CCAT2* cells.

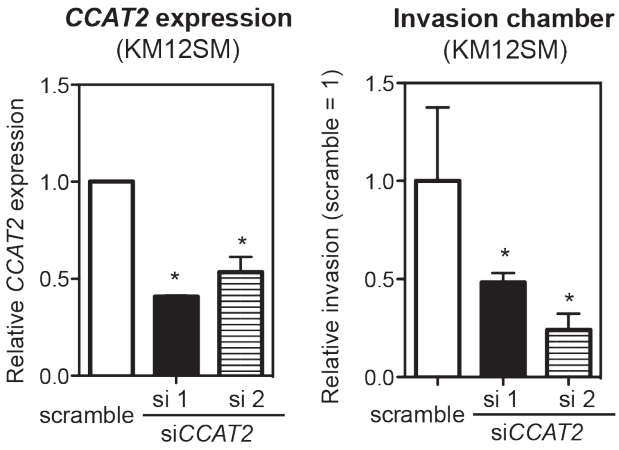
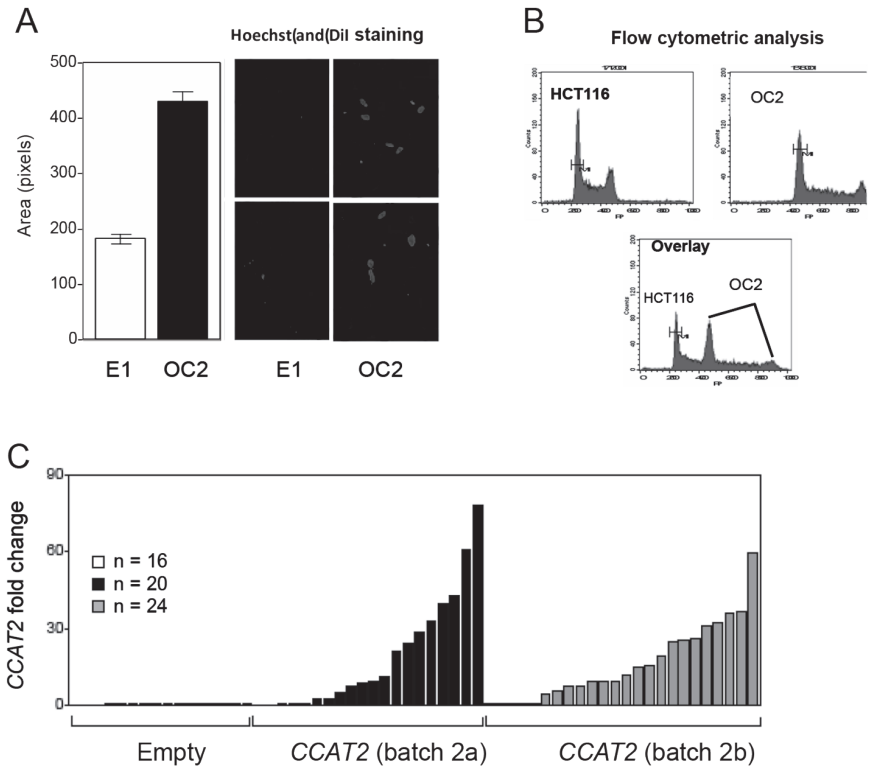


Figure S4. The presence of mitomycin C in invasion assay exludes the possiblity of proliferation interference on the invasion of KM12SM cells. Cells were treated with two independent *CCAT2* siRNA for 24 h, and then were seeded onto invasion assay chamber for 48 h with the treatment of mitomycin C (to block cell proliferation), and invaded cells were stained and counted. The experiment was repeated three independent times, and each time three replicates were used for empty and *CCAT2* cells.



D

	No. of meta. analy	% normal meta	% aberr. meta	% cells with breaks
HCT 116 pMX empty	33	90.9	6.06	6.06
HCT 116 pMX *CCAT2*	31	83.9	16.1	16.1

E

	colo 320 sh (-)	colo 320 sh *CCAT2*
Metaphases analyzed	38	34
% normal metaphase	63.2	73.5
% Poly/ tetraploidy	18.4	20.6
% cells with breaks	13.2	0

Figure S5. Chromosomal aberrances caused by *CCAT2*. (**A**) Nuclear size was increased in *CCAT2*-overexpressing tumor cells (OC2 clone). Cells were stained with Hoechst and DiI (Invitrogen), and images were captured using a fluorescence microscope. The areas of the nuclei (36 nuclei from E1 and 27 nuclei from OC2 clone) were quantified using Image J software; data are presented in a bar graph as means ± SD. (**B**) *CCAT2* caused an increase in the genomic content of HCT116 cells. E1 control, clone OC2, and a combination of these two clones were analyzed for their DNA content using flow cytometry. (**C**) *CCAT2* expression level in an independent batch of pcDNA clones in Figure 3C (the average *CCAT2* expression in 16 empty clones was set as 1 in the clones generated by using qRT-PCR). (**D**) Abnormal chromosomal changes were also observed in retrovirus-transfected HCT116 cells. (**E**) Partial reverse of abnormal metaphase in COLO320 sh*CCAT2* clone.

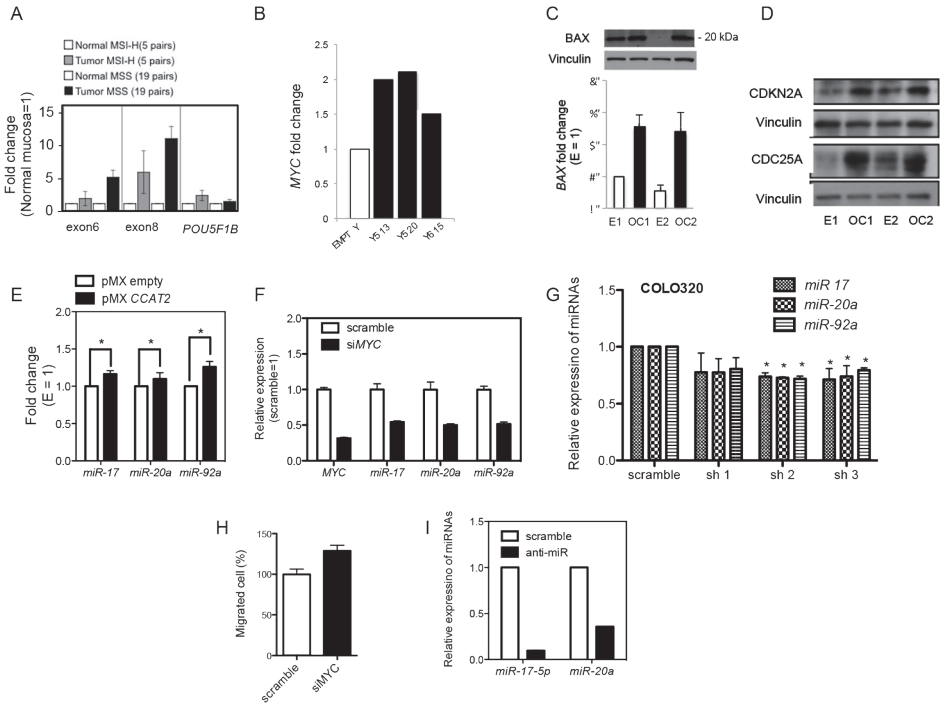


Figure S6. Effect of *CCAT2* on MYC and MYC regulated genes. (**A**) Expression levels of *CCAT2* nearby genes (exons 6 and 8 of *DQ515897* and *POU5F1B*) in CRC samples. (**B**) Higher *MYC* expression levels in *CCAT2*-overexpression clones (generated in the Department of Pathology, Josephine Nefkens Institute, Erasmus Medical Center, Rotterdam, The Netherlands) as checked by qRT-PCR. Y513, Y520, and Y615 are *CCAT2* stable clones. Data show mean values of two measurements. (**C**, **D**) *CCAT2* increased the expression level of Bax (upper panel: western blot; lower panel: qRT-PCR), CDC25A and CDKN2A in HCT116 cell. (**E**, **F**) Expression of *miR-17, -20a and -92a* is higher in retrovirus-transfected *CCAT2* overexpression cells (E), and their expression was significantly reduced by *MYC* siRNA (F). (**G**) Reduced expression of *miR-17, -20a and -92a* in stable sh*CCAT2* clones of COLO320. (**H**) Down-regulation of *MYC* does not reduce *CCAT2*'s migration activity in retrovirus-transfected *CCAT2* overexpression cells. (**I**) Reduced expression of *miR-17* and *miR-20a* in retrovirus transfected *CCAT2* cells after anti-miR treatment.

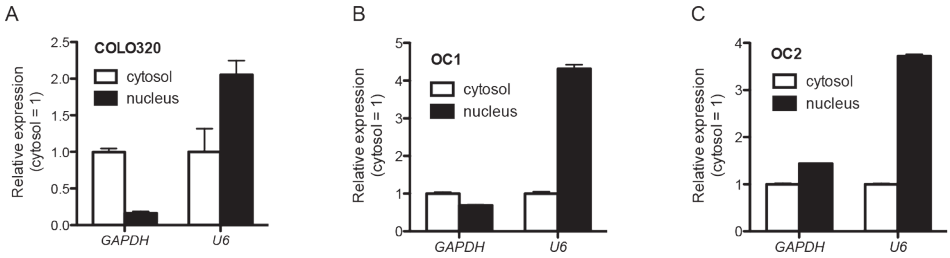


Figure S7. Nuclear and cytosolic separation in COLO320 and HCT116 *CCAT2* clones. After separation of cells using the PARIS kit (Life Technologies), RNA expression levels in (**A**) COLO320, (**B**) OC1, and (**C**) OC2 cells were measured by qRT-PCR. *GAPDH* was used as a cytosol marker and *U6* was used as a nucleus marker. Data show mean values of two measurements.

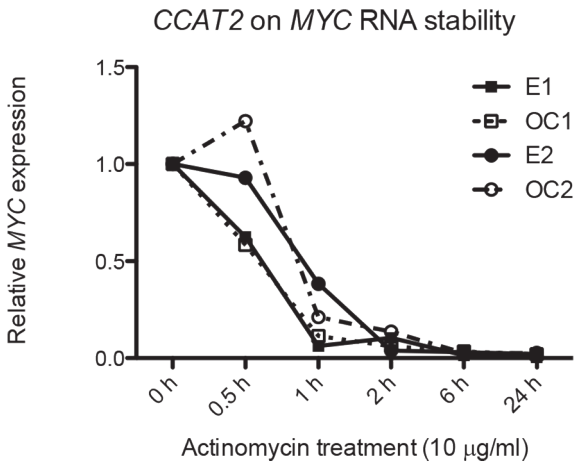


Figure S8. Effect of *CCAT2* on *MYC* mRNA stability. *MYC* mRNA stability was not affected by *CCAT2* expression levels. The four HCT116 clones were treated with Actinomycin and after the indicated incubation periods total RNA was extracted using Trizol method. *MYC* expression levels were normalized to *GAPDH* expression, and the *MYC* expression at the 0 h was set at 1.

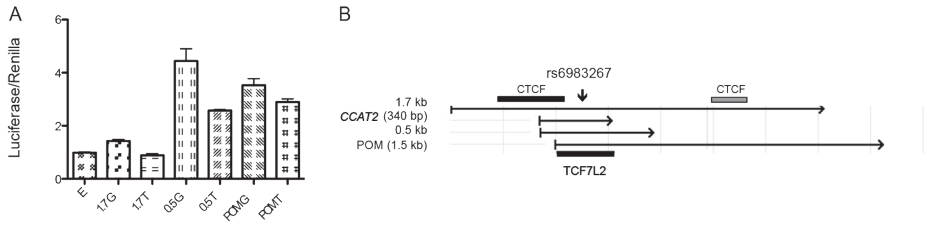


Figure S9. Enhancer activity of rs6983267-containing sequences. (**A**) Sequence-specific and allele-specific enhancer activity of the pGL3 luciferase constructs. The inserts containing rs6983267 were cloned downstream of firefly luciferase gene into pGL3 TK vector. COLO320 cells were co-transfected with the constructed plasmid and pRL-TK renilla construct. Subsequently, after 48 h, the luciferase activity was measured using a dual luciferase kit (Promega) and a Veritas luminometer (Turner Biosystems, Sunnyvale, CA, USA). Results are expressed as firefly luciferase activity normalized to *Renilla*, and data are presented as mean ± SEM from at least three independent experiments, each with duplicate or triplicate experiments. (**B**) The meaning of the constructs and the alignment of the sequences cloned into pGL3 vector and the *CCAT2* sequence. The image was constructed based on "Transcription Factor ChIP-seq from ENCODE" from the UCSC genome browser. The 1.7-kb transcript has two CTCF binding sites (black suggests strongest signal intensity while gray means weaker signal intensity). All these sequences contain a TCF7L2 binding site.

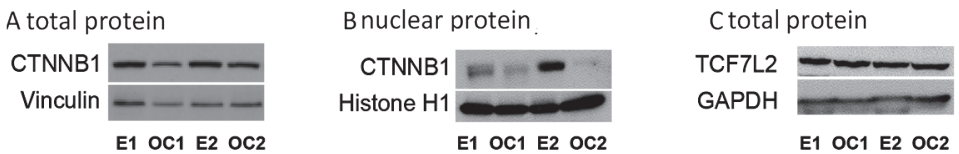


Figure S10. *CCAT2* does not increase expression of CTNNB1 and TCF7L2. Protein expression levels of CTNNB1 (A: total protein; (B: nuclear protein) and TCF7L2 C: total protein) were measured in HCT116 clones with different *CCAT2* expression levels (E = low; OC = high).

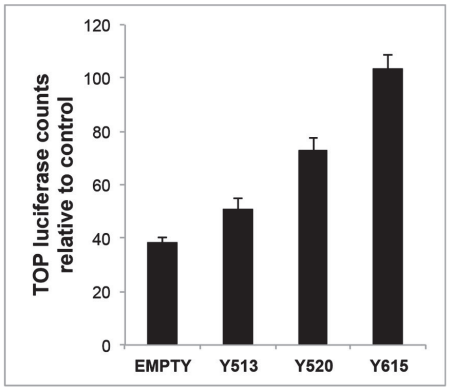


Figure S11. Higher WNT activity was also observed in independently generated *CCAT2* stable clones in the Rotterdam laboratory. These data confirm the results obtained by using the OC1 and OC2 clones prepared at MD Anderson Cancer Center.

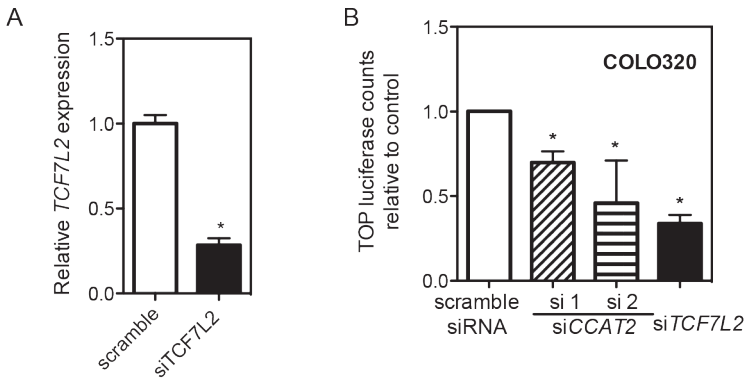


Figure S12. WNT activity in COLO320 cells is effectively reduced by *TCF7L2* siRNA, and also significantly reduced by *CCAT2* siRNA. COLO320 cells were transfected with indicated siRNA (50 nM) by lipofectamine 2000 for 24 h. Equal number of cells were then seeded onto 24-well plate and transfected with TOP-flash WNT reporter. Luciferase activity was measured by the Dual-Luciferase® Reporter Assay System (Promega) 48 h after transfection and TOP or FOP values were normalized to Renilla values. The TOP/FOP ratios were calculated and used as indicators of the endogenous level of WNT signaling. Data are presented as mean ± SEM from two independent experiments repeated in duplicate.

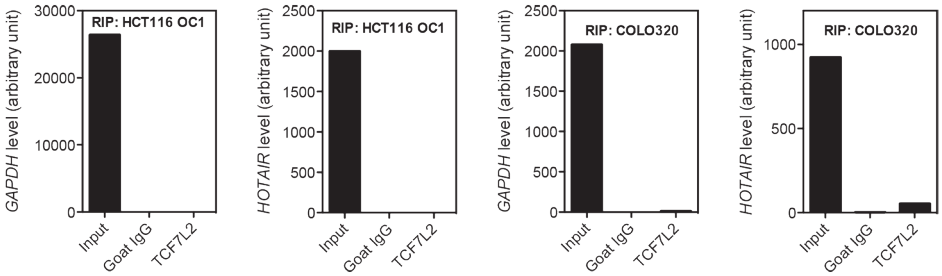


Figure S13. No enrichment of *GAPDH* and *HOTAIR* RNAs in the TCF7L2-pulled samples where TCF7L2-*CCAT2* association was detected. The RNA level in input and TCF7L2-pulled samples were normalized to the level in IgG control set as one.

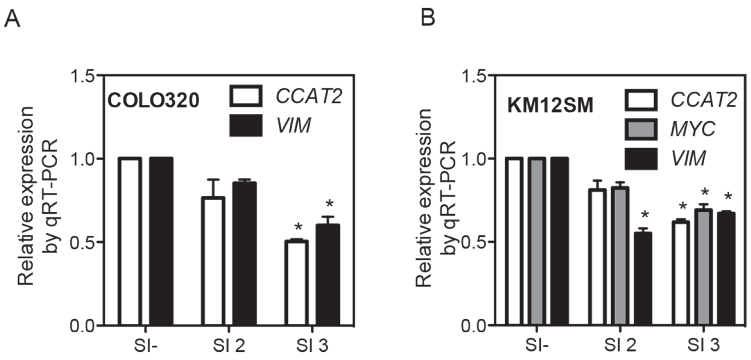


Figure S14. *VIM* and/or *MYC* expression in COLO320 and KM12SM cells was regulated by *CCAT2*. Cells were transfected with indicated siRNA (50 nM) by lipofectamine 2000, and RNA was collected 48 h after treatment for qRT-PCR analysis. Data were from 2-3 independent experiments.

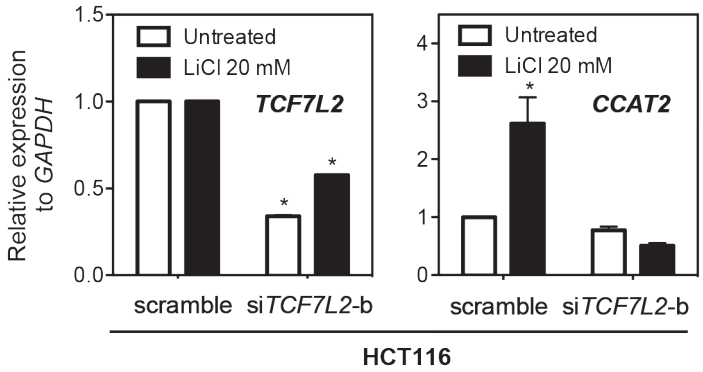


Figure S15. Demonstration of the essential role of TCF7L2 in LiCl-induced *CCAT2* expression by a second *TCF7L2* siRNA. Cells were treated with the same condition as described in Figure 6B using a different *TCF7L2* siRNA (sc-43525, Santa Cruz). Similar with the findings in Figure 6B, TCF7L2 depletion abrogated LiCl-stimulated *CCAT2* expression.

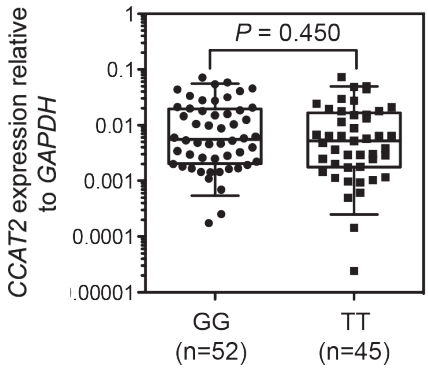


Figure S16. *CCAT2* expression in MSS CRC samples (combining Italian, Japanese and Australian cohort samples used in Figure 1B) with GG or TT rs6983267 genotype**.** Data are presented as box-whisker plots.

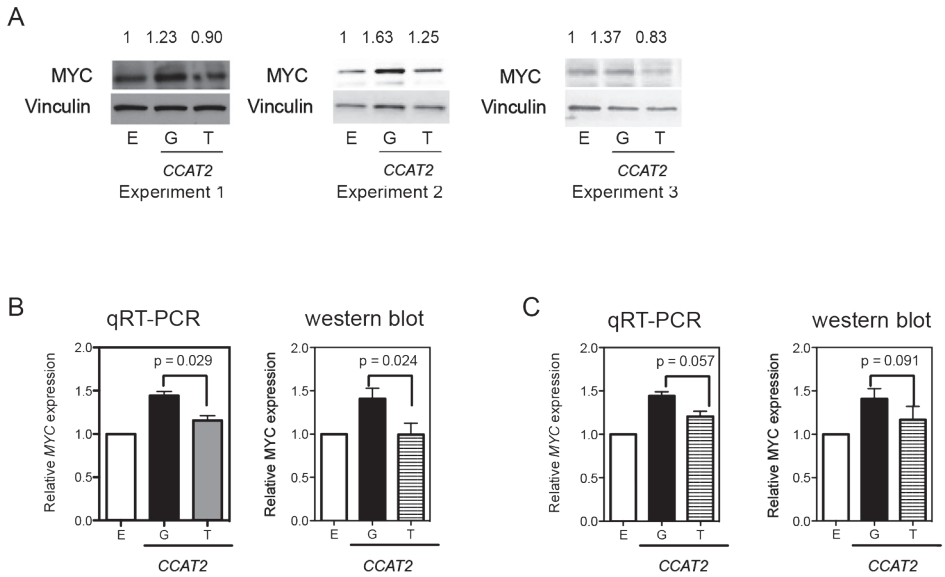


Figure S17. MYC expression levels in retrovirus (*CCAT2* G or *CCAT2* T) -transduced *CCAT2*-overexpressing HCT116 cells. (**A**) Western blog images of three independent experiments showing higher MYC expression in *CCAT2* G than *CCAT2* T-transduced cells. (**B**) Higher *MYC* mRNA (left panel) and MYC protein (right panel) expression in *CCAT2* G than *CCAT2* T-transduced cells. The band intensity of western blot images in A was quantified using image J and presented as mean ± SEM, (n=3). (**C**) After normalizing *MYC* or MYC expression to *CCAT2* expression levels in retrovirus-transduced cells, the difference between *CCAT2* G and *CCAT2* T-transduced cells is not significant.

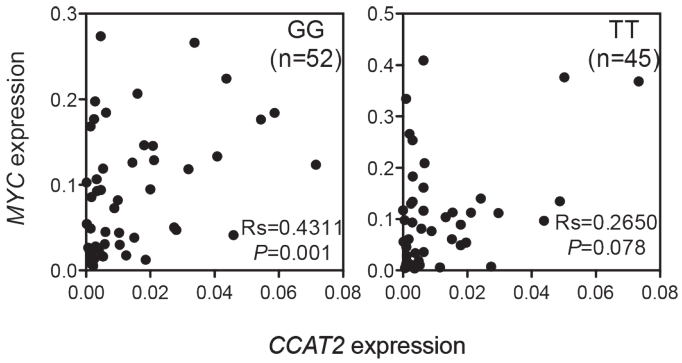


Figure S18. Significant association between *CCAT2* and *MYC* expression level was observed in GG type, but not in TT type of MSS CRCs. *CCAT2* and *MYC* expression in CRC samples (combining Italian, Japanese and Australian cohort samples used in Figure 1B) were quantified using qRT-PCR analysis, and the strength of association between the expressions of *CCAT2* and *MYC* was evaluated using Spearman's rank-order correlation test.

Figure S19. Predicted secondary structure of *CCAT2* G versus *CCAT2* T using Vienna RNA Secondary Structure Package.

Table S1. Cox univariate and multivariate analysis for metastasis-free survival as a function of *CCAT2* levels.[a]

Factor	No. of patients	Univariate analysis HR	95% CI	*P*	Multivariate analysis HR	95% CI	*P*
					Base model		
Age at start of therapy, years							
≤40	35	1			1		
41–50	91	0.64	0.37-1.10		0.61	0.35-1.07	
51–70	3	0.26	0.03-1.99	0.14	0.17	0.02-1.56	0.09
Menopausal status at start of therapy							
premenopausal	116	1			1		
postmenopausal	13	0.88	0.40-1.96	0.76	1.29	0.49-3.39	0.60
Tumor size							
pT1, ≤2 cm	32	1			1		
pT2, >2- ≤5 cm	79	2.65	1.19-5.91		2.38	1.05-5.40	
pT3, >5 cm, + pT4	18	4.93	1.93-12.55	**0.002**	5.77	2.16-15.42	**0.001**
Lymph nodes involved							
1–3	86	1			1		
>3	43	1.72	1.03-2.87	**0.041**	1.44	0.84-2.47	0.18
Grade							
poor	64	1			1		
unknown	47	0.62	0.36-1.09		0.46	0.24-0.86	
moderate	18	0.43	0.17-1.11	0.08	0.35	0.13-0.93	**0.013**
ER status, mRNA level							
negative, <0.2	25	1			1		
positive, ≥0.2	104	0.64	0.34-1.19	0.18	0.88	0.34-2.23	0.78
PgR status, mRNA level							
negative, <0.1	45	1			1		
positive, ≥0.1	84	0.79	0.47-1.32	0.37	0.70	0.32-1.52	0.37
					Additions to the base model[b]		
***CCAT2* level**							
log-transformed continuous	129	1.17	1.05-1.31	**0.005**	1.16	1.05-1.28	**0.002**
***CCAT2* level**[c]							
low	61	1			1		
high	68	2.01	1.20-3.35	**0.008**	2.08	1.19-3.62	**0.010**

[a] Primary breast tumors from 129 lymph node-positive breast cancer patients who received adjuvant CMF at Erasmus Medical Center were analyzed.

[b] CCAT2 was separately introduced to the base multivariate model that included the following factors: age, menopausal status, nodal status, pathologic tumor size, grade, and ER and PgR status.

[c] Dichotomized at the median level to low (≤0.0045) and high (>0.0045).

Table S2. Primers, oligoprobes, siRNA, shRNAs, and antibodies used in this study.

PCR Primers	Sequence
CCAT2 F1	CCCTGGTCAAATTGCTTAACCT
CCAT2 R1	TTATTCGTCCCTCTGTTTTATGGAT
CCAT2 F2	CCGAGGTGATCAGGTGGACTTTC
CCAT2 R2	GTCTTCTGGGCTGATGTTGC
MYC_F	CAGCTGCTTAGACGCTGGATT
MYC_R	GTAGAAATACGGCTGCACCGA
TCF7L2_F	CCGAAAGTTTCCGAGACAAA
TCF7L2_R	AGTGGCCATTTCATCTGGAG
GAPDH_F	ACCCAGAAGACTGTGGATGG
GAPDH_R	TCTAGACGGCAGGTCAGGTC
BAX_F	CTGGACCCGGTGCCTCAGGA
BAX_R	CGGCGGCAATCATCCTCTGCA
C17_F	GGAGTAGGGTCTCAACAAGGGTCTC
C17_R	ACTCAGGTGGTCCCAGGAAGTGTGG
Exon6_F	GATGAACATGGTCACCTTGAAA
Exon6_R	AGGGTTCCTGAGGAGCAAGT
Exon8_F	AGCCTATGGAGAGGATGCAA
Exon8_R	AGGGACAGCTTGTGTCTGCT
CD44_F	TGCCATGTAGATCCTGTTTGAC
CD44_R	AAGTTCTAGGACCAATCGGACA
VIM_F	GAGAACTTTGCCGTTGAAGC
VIM_R	GCTTCCTGTAGGTGGCAATC
U6_F	CTCGCTTCGGCAGCACA
U6_R	AACGCTTCACGAATTTGCGT
Northern blot	**Oligo probe**
CCAT2 probe	CTTTGTACTTTTCTCAGTGCCTTTCATCTGC
siRNA	**Target Sequence**
CCAT2 siRNA 1	TTAACCTCTTCCTATCTCA
CCAT2 siRNA 2	AGGTGTAGCCAGAGTTAAT
CCAT2 siRNA 3	CGAATAAACTCTCCTCCTA
shRNA	**Sequences cloned into pRS vector**
CCAT2 shRNA1	GATCCCCAATTGCTTAACCTCTTCCTTTCAAGAGAAGGAAGAGGTTAAGCAATT TTTTTA
CCAT2 shRNA2	GATCCCCGATGAAAGGCACTGAGAAATTCAAGAGATTTCTCAGTGCCTTTCATC TTTTTA
CCAT2 shRNA3	GATCCCCTTAACCTCTTCCTATCTCATTCAAGAGATGAGATAGGAAGAGGTTAA TTTTTA

PCR Primers	Sequence
Tet-on shRNA	**Sequences cloned into pSINGLE vector**
sense	TCGAGGTTAACCTCTTCCTATCTCATTCAAGAGATGAGATAGGAAGAGGTTAATT TTTTACGCGTA
antisense	AGCTTACGCGTAAAAAAATTAACCTCTTCCTATCTCATCTCTTGAATGAGATAGGA AGAGGTTAACC
RIP	**Primer sequence**
RT-PCR primer_F	TTTAGCAGCTGCATCGCTCCATAG
RT-PCR primer_R	TTTGTACTTTTCTCAGTGCCTTTCA
qRT-PCR primer_F	CCCTGGTCAAATTGCTTAACCT
qRT-PCR primer_R	TTATTCGTCCCTCTGTTTTATGGAT
Antibodies	**Product information**
MYC	Millipore (06-340)
vinculin	Santa Cruz (sc-7649)
BAX	Cell Signaling (2772)
CTNNB1	Santa Cruz (sc47752)
TCF7L2	Abcam (AB76151)
GAPDH	Calbiochem (D33622)

SUPPLEMENTARY REFERENCES

Orom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q et al. 2010. Long noncoding RNAs with enhancer-like function in human cells. Cell 143(1): 46-58.

Pomerantz MM, Ahmadiyeh N, Jia L, Herman P, Verzi MP, Doddapaneni H, Beckwith CA, Chan JA, Hills A, Davis M et al. 2009. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet 41(8): 882-884.

Zhao X, Li C, Paez JG, Chin K, Janne PA, Chen TH, Girard L, Minna J, Christiani D, Leo C et al. 2004. An integrated view of copy number and allelic alterations in the cancer genome using single nucleotide polymorphism arrays. Cancer Res 64(9): 3060-3071.

chapter SIX

THE ROLE OF *S100A4* (*MTS1*) IN *APC*- AND SMAD4-DRIVEN TUMOR ONSET AND PROGRESSION

Yaser Atlasi[1], Rubina Noori[1], Ivana Marolin[1], Patrick Franken[1], Joana Brandao[1], Katharina Biermann[1] Paola Collini[2], Mariam Grigorian[3,4], Eugene Lukanidin[3,4], Noona Ambartsumian[3,4] and Riccardo Fodde[1]

[1]Dept. of Pathology, Josephine Nefkens Institute, Erasmus MC, Rotterdam, The Netherlands; [2]Dept. of Pathology, Istituto Nazionale Tumori, Milan, Italy; [3]Dept'. of Tumor Microenvironment and Metastasis, Danish Cancer Society, Copenhagen Ø, Denmark; [4]Dept. of Neuroscience and Pharmacology, Faculty of Health Sciences, University of Copenhagen, Copenhagen N, Denmark

Submitted

chapter SEVEN

DISCUSSION

In higher multicellular organisms life starts from the fertilized egg, a totipotent cell that is capable of giving rise to every single cell in the body. This developmental program is well-controlled, reproducible, and has highly predictable outcomes. The primary cells of the developing embryo display a unique potential for multi-lineage differentiation and hence are referred to as pluripotent. Different pluripotent stem cells have been captured and propagated *in vitro* and their culture requirements were investigated accordingly. These cells include, among others, the embryonic stem cells (ESC) derived from the inner cell mass (ICM) of the pre-implantation embryo, the epi stem cells (EpiSC) derived from the post-implantation embryo, and the embryonic germ cells derived from the primordial germ cells. The inter-conversion of these stem cell lineages obtained through the manipulation of the culture conditions also led to the identification of different intermediate states with various pluripotency potentials [1,2]. Importantly, the source of pluripotent stem cells has been extended to differentiated cells with the discovery of the reversibility of the differentiation process, a process referred to as reprogramming [3]. Currently, ESCs are employed both in regenerative medicine and as a basic model for stem cell research. Moreover, they have proven an excellent model to explore the genetic and epigenetic alterations implicated in cancer. Hence, ESCs provide a unique *in vitro* model to elucidate the cellular and molecular mechanisms which underlie the capacity of stem cells to adapt to environmental changes (i.e. to specific culture conditions) [4].

Among the different signaling pathways implicated in ESCs, Wnt signaling has revolutionized the derivation and propagation of ESCs. Manipulation of this signaling pathway forms an integral component of many differentiation protocols. In this section, I will discuss how the findings reported in **chapter 2,3** and **4** of this thesis contributed to our improved understanding of the role of Wnt signaling in pluripotent stem cells. Moreover, Wnt signaling is one of the most frequently deregulated signaling pathways in cancer. In **chapter 5** and **6** we studied two Wnt targets, namely the non-coding RNA *CCAT2* and the stromal-associated factor *S100a4* in colon cancer and desmoid disease. Here, I will discuss how these findings also improved our understanding of the role of Wnt signaling in cancer.

I WNT SIGNALING SUPPORTS SELF-RENEWAL IN NAÏVE ESCS

Whether Wnt signaling plays any functional role in mouse ESCs has been a matter of controversy. Although β-catenin-null ESCs retain their undifferentiated phenotype under self-renewing condition, these cells exit the self-renewal state faster when challenged for differentiation [5,6]. On the other hand, gain of function approaches revealed a robust self-renewal promoting role for β-catenin in mESCs. Genetic downregulation of Apc or Gsk confer differentiation resistance and enhance mESCs self-renewal [7,8]. Accordingly, activation of Wnt signaling, in combination with MEK/ERK inhibition has become the golden-standard culture method to maintain ESCs in a self-renewing state. Moreover, Wnt signaling significantly enhances the derivation and *in vitro* establishment of ES cells from pre-implantation embryos [9]. Overall, these data suggest that Wnt signaling promotes self-renewal in naïve ESCs.

However, the downstream effects of Wnt signaling include both self-renewal and mesoderm differentiation. Gene expression profiles of Gsk-inhibited or *Apc*-mutant

(*Apc*NN) ES cells revealed induction of several meso- and endoderm commitment genes (including *Cdx1*, *Cdx2* and *Brachyury*) in addition to the self-renewal program [10,11]. This observation was also confirmed by analyzing the chromatin distribution pattern of β-catenin in ESCs treated with the Gsk inhibitor CHIRON [10]. Hence, activation of Wnt signaling simultaneously induces both self-renewal and lineage commitment programs which results in a metastable state in mESCs. Importantly, the combined inhibition of MEK/ERK signaling and activation of Wnt signaling (2i medium) suppresses the lineage differentiation module and synergistically promotes the self-renewal program of mESCs.

In contrast to mESCs, it is still not clear whether Wnt signaling plays any *in vivo* role in the ICM and how the *in vitro* results relative to this signaling pathway reflect its alleged function in the pre-implantation embryo. As mentioned before, activation of Wnt signaling represents one of the main components of the 2i medium which supports ESC self-renewal and maintenance. Accordingly, the 2i-cultured cells resemble the naïve cells found in the pre-implantation embryo. However, depleting the maternal and zygotic β-catenin does not impair blastocyst formation [12-14]. Similarly, Wnt signaling inhibition through small molecule inhibitors or the Wnt antagonist DKK1 did not block blastocyst formation [15]. Moreover, most of the Wnt-reporter constructs failed to detect Wnt signaling activity in ICM cells and β-catenin localization is restricted to cell membrane at this stage [16-18]. These findings suggest that canonical Wnt signaling might be dispensable in the pre-implantation-stage mouse embryo. Further research is required to investigate the parallel role of Wnt signaling *in vitro* and in the ICM.

II A LARGE PART OF WNT SIGNALING EFFECTS IS MEDIATED THROUGH TCF3 ANTAGONISM

The complex and diverse cellular responses to Wnt ligands or *Apc*/Gsk inhibition pose a major challenge towards the dissection of the molecular mechanisms downstream of β-catenin signaling. A β-catenin isoform that lacks the transactivation domain (C-terminal deletion) is still able to mediate the self-renewal response to Gsk inhibitors [6]. These data suggest that the β-catenin transcriptional activity plays only a limited role in the activation of downstream Wnt targets. In **chapter 2**, we analyzed the gene expression profiles of mESCs that harbor different *Apc* hypomorphic alleles and found specific *Tcf3* downregulation in Wnt-high ESCs. Recent studies suggested that β-catenin stabilization promotes the self-renewal downstream of the Gsk inhibitor via Tcf3 antagonism [6,19,20]. This model differs from the classical scenario where β-catenin binding converts Tcf factors from transcriptional repressors into activators. In fact, it implies that Tcf3 antagonism is sufficient to mediate the downstream effects of Wnt signaling (Figure 1, adapted from [21]). Indeed, comparison of the transcriptional profiles of Wnt3a-stimulated and *Tcf3*-depleted ESCs revealed a high degree of similarity centered around the self-renewal genes [19]. Accordingly, a substantial portion of genes induced by β-catenin stabilization and *Tcf3* depletion is also under the control of Oct4 and Nanog regulation [6,19,20]. Overall, these studies indicate that a large part of the self-renewal promoting effects of Wnt/β-catenin signaling can be narrowed down to *Tcf3* antagonism.

But how does β-catenin counteract Tcf3 function? The physical interaction between β-catenin and Tcf3 or the β-catenin mediated phosphorylation of Tcf3, both of which lead to Tcf3 displacement, have been suggested [6,22]. In **chapter 2**, we showed that

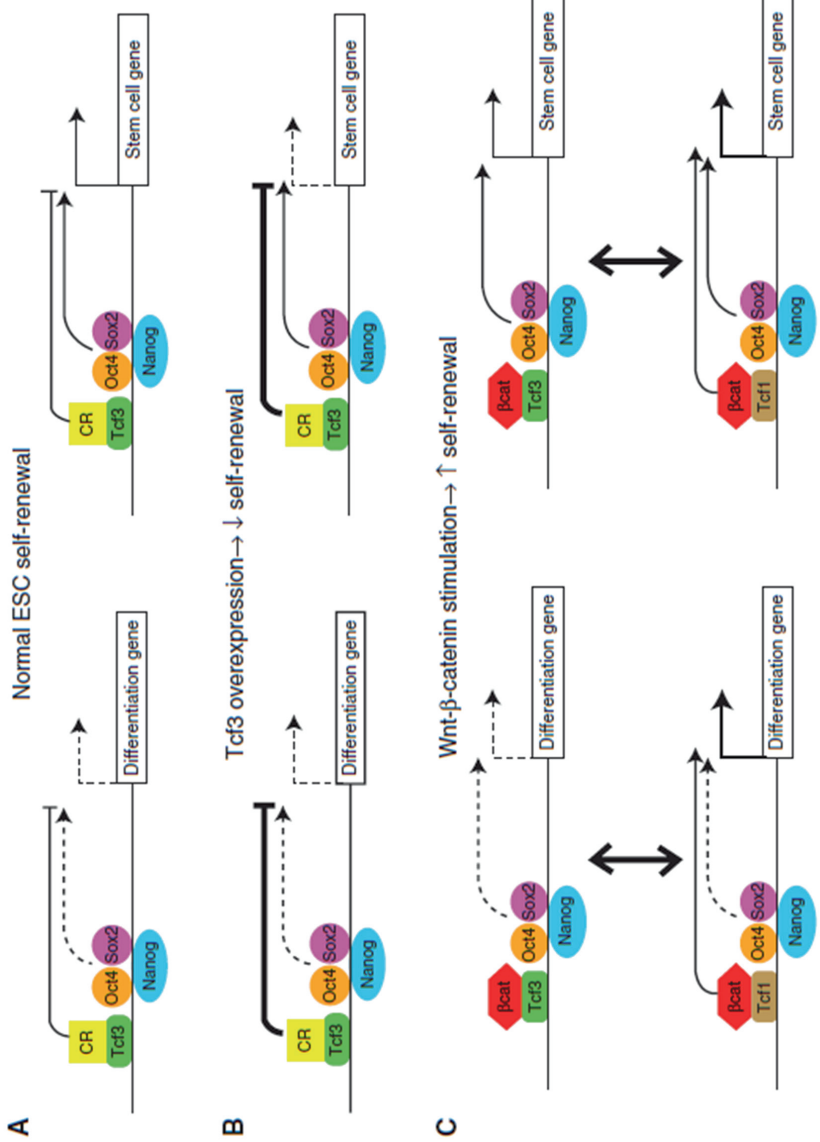


FigurE 1. Model illustrating the control of pluripotency and differentiation through Wnt/ β-catenin and Tcf proteins in mESCs (adapted from [21]). CR = Tcf3–corepressor.

Wnt signaling also downregulates *Tcf3* expression at RNA and protein levels generating a feed-forward loop to enhance Wnt effects. Moreover, Wnt signaling induces *miR-211* which in turn targets *Tcf3* thus maintaining its repression. Collectively, these studies indicate that Wnt signaling antagonizes Tcf3 function at different levels in mouse ESCs.

Further studies demonstrated that, in addition to Tcf3 antagonism, β-catenin-Tcf1 signaling provides additional positive effects on self-renewal, thus suggesting that a combination of *Tcf1*- and *Tcf3* -mediated effects contribute to enhanced self-renewal downstream of β-catenin signaling [19]. These observations indicate that the effect of stromal (i.e. micro-environmental) signals, and in this specific case of Wnt signaling ligands, is directly translated to the core pluripotency circuitry thus highlighting a functional link between microenvironment and the pluripotency module in ESCs.

A question that still remains unanswered is why *Tcf3* is expressed at its highest level in self-renewing ESCs provided that its main function is to repress the self-renewal circuitry? One explanation might be that this is yet another example of an incoherent feed-forward loop where both positive and negative regulators are simultaneously expressed to ensure a rapid differentiation response upon external stimuli.

III THE LINEAGE DIFFERENTIATION EFFECTS OF WNT SIGNALING DEPEND ON THE CELLULAR STATE

In addition to promoting self-renewal, constitutive Wnt signaling activation also affects cell fate decisions and multi-lineage differentiation. Different studies suggest that the differentiation effect of Wnt signaling is dependent on the cellular context and in particular on the developmental stage of the target cell. Previously, we employed a set of *Apc*-mutant ESCs encoding for different Wnt signaling dosages and showed that high Wnt signaling activation induces a broad spectrum of differentiation defects towards different lineages and in particular the neuroectodermal lineages [7]. A similar observation was also reported for *Gsk*-null ESCs [7,8,23]. In **chapter 2**, we showed that *Tcf3*-downregulation plays an important, role in the modulation of the Wnt downstream effects which suppress the neural differentiation in mESCs [24].

On the other hand, Wnt signaling is also required for lineage differentiation. In fact, expression of Wnt ligands and the activity of *Tcf/Lef* Wnt reporter constructs was shown to increase during ESC differentiation [25]. Similarly, Wnt signaling inhibition (e.g by using DKK1) suppresses mesoderm and endoderm lineage differentiation [25]. Also β-catenin–null ESCs do not form teratomas when injected in recipient mice [26] and, when subjected to embryoid body formation, exhibit pronounced defects in mesoderm, endoderm and neuroectoderm differentiation [5,26]. Taken together, these data indicate that the outcome of Wnt signaling is largely dependent on the developmental stage and cellular context of the target cell. While in undifferentiated ESCs Wnt signaling supresses differentiation and enhances self-renewal, it is also capable of playing a promoting role during differentiation where the signaling pathway becomes necessary for proper lineage commitment.

IV WNT SIGNALING AND TRANSCRIPTIONAL REPRESSION

As described in **chapter 2**, we observed downregulation of many genes, including some of the Wnt targets (such as *c-Myc*) in *Apc*NN ESCs [7,24]. According to the current model, β-catenin functions as a transcriptional activator displacing the Tcf-bound repressors and recruiting a plethora of transcriptional activators. However, β-catenin mediated gene repression was also reported for some of the Wnt targets including *stripe* in the *Drosophila* epidermis [27], *Decapentaplegic* (*Dpp*) in the fly leg imaginal disc [28] and *CDH1* (E-cadherin) in mouse keratinocytes [29]. A study by Cadigan and colleagues identified new TCF binding sites which mediate the transcriptional repression of several Wnt target genes in *Drosophila*. These DNA elements were distinct from the classical Wnt response elements (WREs) and mediated repression of *Ugt36Bc*, *Pxn*, *Tig*, and *Ugt58Fa* genes when Wnt signaling was activated [30]. Whether similar Tcf binding sites exist in mammalian genomes and whether these sites mediate the observed gene repression in *Apc*NN ESCs requires further investigation. Moreover, RNA pol II pausing was recently proposed as a mechanism of gene repression in 2i-cultured ESCs [31]. In 2i-cultured-cells, no increase in H3K27m3 or DNA methylation was observed in the downregulated loci. However, the transcriptional elongation was attenuated in these genes. Whether a similar mechanism mediates gene repression downstream of Wnt signaling remains elusive. In line with these findings, *c-Myc* is implicated in RNA Pol II pause release [32] and both *Apc*NN and 2i-cultured ESCs display *c-Myc* downregulation when compared to wild type ESCs cultured in serum, further highlighting a potential role of RNA pol II pausing downstream of Wnt signaling.

V WNT SIGNALING AND MIR-302 REPRESSION

In **chapter 3** we showed that constitutive activation of Wnt signaling suppresses miR-302 expression in mESCs. When compared to naïve ESCs, miR-302 is upregulated in EpiSCs. Recently Wnt signaling was found to suppress the transition from naïve ESCs to primed EpiSCs [33]. On the other hand, cells cultured in serum were shown to display transcriptional and epigenetic programs closely resembling the post-implantation embryo whereas cells cultured in 2i medium resembled the pre-implantation naïve state [31]. Despite similar expression of many pluripotency markers, the two culture systems display fundamental differences in their transcriptional and epigenetic programs. These results indicate that the serum cultured cells might represent the earliest differentiation step in mESCs. In this view, we observed induction of the naïve ES-markers (such as *Klf4*, *Tbx3* and *Essrb*) and downregulation of some of the primed ES-associated genes (such as *Dnmt3b*, *Oct6* and *Otx2*) in *Apc*NN cells when compared to wild type controls. We therefore hypothesize that ApcNN cells cultured in serum closely resemble the naïve state which is associated with miR-302 down-regulation. In line with these observations, when wild-type ESCs cultured in serum are switched to 2i medium, a significant downregulation of miR-302 is observed (**chapter 3**).

VI DOES WNT SIGNALING PROMOTE SELF-RENEWAL OR DIFFERENTIATION IN PLURIPOTENT CELLS?

In **chapter 4**, we discussed the role of Wnt signaling in a different pluripotent cell type, namely human embryonic carcinoma (hEC) cells. We showed that short-term activation of Wnt signaling induces loss of pluripotency in hEC cells. The observed role of Wnt in hEC lines adds yet another piece to the puzzle of Wnt signaling in different pluripotent stem cells. hECs share many similarities with hESCs. However, the role of WNT signaling in hESCs is somewhat controversial. Several reports demonstrated that Wnt3a or the Gsk-inhibitor BIO support short term self-renewal in hESCs but fail to maintain it in the long term [34-37]. A recent report by Blauwkamp and colleagues demonstrated that Wnt3a supports the self-renewal of hESCs in standard ES-medium [38]. On the other hand, Davidson and colleagues provided evidence that *OCT4* represses β-catenin signaling and that targeted knock down of *OCT4*, leads to WNT activation. In this study, WNT signaling was shown to induce hESCs differentiation [39]. Similarly, Wnt signaling induces mesodermal differentiation in mouse epiblast cells which share similar features to hESCs [14,40,41]. These results suggest that Wnt signaling cannot maintain the long-term self-renewal of hESCs and mouse EpiSCs.

The intriguing and yet unanswered question is how the response to Wnt signaling is interpreted differently in mouse and human pluripotent cells. While Wnt signaling promotes self-renewal in mouse ESCs, its activation induces differentiation in hES, hEC and mouse EpiS cells. One plausible explanation might be that the developmental stages of mouse and human ESCs are different and that Wnt signaling induces opposite responses depending on the cellular state of the target cell. Of note, in a recent study Kim et al provide another explanation for the seemingly different effects of Wnt signaling in pluripotent cells [42]. This study suggests that the nuclear and cytoplasmic β-catenin have different effects on self-renewal. While nuclear β-catenin induces differentiation, the cytoplasmic β-catenin supports self-renewal. The idea that β-catenin promotes self-renewal or differentiation depending on its sub-cellular localization as well as the cellular stage of the target cell, might also shed light on the role of Wnt signaling in cancer. In colorectal tumors, despite Wnt activation in all tumor cells (due to *APC* or *CTNNB1* mutations), only a subset of tumor cells, mainly those located at the invasive front of the tumor, display nuclear β-catenin whereas more cytoplasmic or membranous subcellular localizations are observed in other tumor cells. Interestingly, tumor cells with nuclear β-catenin, often located in direct proximity to the stroma, usually display features reminiscent of epithelial-to-mesenchymal transition (EMT) [43]. On the other hand, in both hESCs and murine EpiSCs, mesodermal differentiation represents the major outcome when high levels of Wnt signaling are induced. These observations suggest that nuclear β-catenin might induce mesenchymal phenotype and EMT in both hES and CRC tumor cells analogous to the normal function of Wnt signaling during mesodermal differentiation in early developing embryo. Interestingly, desmoids tumors, as discussed in **chapter 6**, are among the few tumor types where almost every single cell within the tumor mass displays nuclear β-catenin staining. Of note, desmoids are mesenchymal tumors entirely composed by fibroblast-like cells. As reported in **chapter 2**, the same is true for teratomas generated from *Apc*NN ESCs where a high percent of cells undergoing mesenchymal differentiation also display nuclear β-catenin staining [7,24]. Taken together, these observations might indicate that high levels of Wnt signaling, manifested as nuclear

β-catenin, might induce commitment towards a mesenchymal phenotype and that this role of β-catenin signaling might be conserved in different cellular contexts ranging from developing embryo, ESCs to tumor cells.

VII WNT SIGNALING AND GERM CELL TUMORS

In **chapter 4**, we showed that in the majority of examined hEC cells WNT signaling is active at very basal levels. Moreover, and similar to their hES counterparts and the mouse EpiSCs, we showed that short-term activation of WNT signaling suppresses pluripotency and induces a range of differentiation responses in EC cells. The majority of the employed EC lines adapted to WNT stimulation and activated expression of pluripotency markers overtime.

These data might also shed light on the genesis and differentiation of embryonal carcinoma (ECs) *in vivo*. Carcinoma *in situ* (CIS) have been suggested as the cell of origin for both seminomas (SEs) and ECs. However, the molecular mechanisms controlling the cell fate decision of CIS towards SEs or ECs still remain unknown. Although both SEs and ECs express the majority of pluripotency genes, only ECs display extensive pluripotency and can give rise to differentiated GCTs such as yolk sac tumors, choriocarcinoma and teratomas. Moreover, SEs are thought to represent the default tumor type generated from CIS whereas activation of pluripotency in CIS (or in some cases in SEs), leads to EC formation. Among the main differences observed between ECs and SEs, the switch from SOX17 to SOX2 expression and the activation of WNT signaling in ECs are of interest for the present discussion [44-46]. Hence it is plausible to think that activation of WNT signaling might underlie the conversion of CIS into ECs. Whether activation of WNT signaling in EC cells is the cause or the consequence of CIS/SE to EC conversion remains unknown. The remarkable switch from SOX17 to SOX2 might also correlate with WNT activation during CIS/SE to EC transition. In line with these observations, several studies have reported Sox17 as an antagonist of WNT signaling in xenopus, mouse and human cells[47-50]. The high level of SOX17 in SE/CIS cells might therefore maintain low levels of WNT signaling in the cells and the switch from SOX17 to SOX2 expression might activate WNT signaling in ECs. In addition to the basal level of WNT signaling in ECs, our results also demonstrate that high level of WNT signaling induces loss of pluripotency. This observation might reflect the *in vivo* behavior of ECs where high levels of WNT signaling are correlated with the differentiation of EC cells to yolk sac and teratomas.

VIII WNT REGULATED NON-CODING RNAS IN STEM CELLS AND CANCER

In **chapter 5**, we described the WNT-regulated non-coding RNA, CCAT2, in human colorectal cancer. CCAT2 transcript is highly conserved across different species and mice deficient for a 1740-bp DNA element, encompassing CCAT2, are resistant to intestinal tumorigenesis induced by the *Apc*Min mutation [51]. These results suggest a conserved role for CCAT2 in human and mouse intestinal tumorigenesis. In **chapter 5** we showed that CCAT2 is positively correlated with chromosomal instability (CIN), Wnt signaling and C-MYC expression in colorectal tumors. However, we did not observe CCAT2 up-

regulation in *Apc*NN ESCs which display both high Wnt signaling and CIN phenotype (**chapter 2 and 3**). Moreover *c-Myc* is downregulated in *Apc*NN ESCs compared to wild type controls and similarly in 2i- vs. serum-cultured ESCs. We also observed *C-MYC* downregulation in CHIRON-treated hEC cells (chapter 4). Taken together, these findings indicate a differential regulation of c-Myc in pluripotent stem and intestinal tumor cells and suggest that the epistatic interactions between WNT, C-MYC and CCAT2 non-coding RNA might be dependent on the cellular context. Further research is required to explore the role of CCAT2 non-coding RNA in different pluripotent cells.

IX *S100A4* IN TUMOR ONSET AND PROGRESSION

The causal role of *S100A4* in colorectal cancer remains poorly understood. In **chapter 6** we showed that *S100a4* is dispensable for intestinal tumor initiation as both *Apc*- or *Smad4*–mutant mice developed similar number of tumors when *S100a4* was depleted. On the other hand, a large body of evidence suggest a strong role for S100a4 in tumor metastasis. We could not address this issue with the employed mouse models which do not form metastases due to the discomfort caused by primary tumors. Therefore, novel models for colon cancer and the resulting liver metastasis will be necessary in the future to examine the *in vivo* role of S100a4 in this process. In contrast, we showed that specific depletion of S100a4 in $Apc^{1638N/+}$ (male) mice significantly reduces desmoids formation (**chapter 6**). However, desmoids are benign tumors and, in $Apc^{1638N/+}$ mice, are relatively quiescent. Therefore, the development of mouse models for aggressive fibromatosis (desmoid disease) as observed in man is of relevance to further study the role of *S100A4* in the onset and progression of desmoid tumors and in particular in desmoid-associated MSCs.

REFERENCES

1. Nichols J, Smith A (2012) Pluripotency in the embryo and in culture. Cold Spring Harb Perspect Biol 4: a008128.
2. Buecker C, Geijsen N (2010) Different flavors of pluripotency, molecular mechanisms, and practical implications. Cell Stem Cell 7: 559-564.
3. Takahashi K, Yamanaka S (2006) Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell 126: 663-676.
4. Baker DE, Harrison NJ, Maltby E, Smith K, Moore HD, et al. (2007) Adaptation to culture of human embryonic stem cells and oncogenesis in vivo. Nat Biotechnol 25: 207-215.
5. Lyashenko N, Winter M, Migliorini D, Biechele T, Moon RT, et al. (2011) Differential requirement for the dual functions of beta-catenin in embryonic stem cell self-renewal and germ layer formation. Nat Cell Biol 13: 753-761.
6. Wray J, Kalkan T, Gomez-Lopez S, Eckardt D, Cook A, et al. (2011) Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. Nat Cell Biol 13: 838-845.
7. Kielman MF, Rindapaa M, Gaspar C, van Poppel N, Breukel C, et al. (2002) Apc modulates embryonic stem-cell differentiation by controlling the dosage of beta-catenin signaling. Nat Genet 32: 594-605.
8. Doble BW, Patel S, Wood GA, Kockeritz LK, Woodgett JR (2007) Functional redundancy of GSK-3alpha and GSK-3beta in Wnt/beta-catenin signaling shown by using an allelic series of embryonic stem cell lines. Dev Cell 12: 957-971.
9. Ying QL, Wray J, Nichols J, Batlle-Morera L, Doble B, et al. (2008) The ground state of embryonic stem cell self-renewal. Nature 453: 519-523.
10. Zhang X, Peterson KA, Liu XS, McMahon AP, Ohba S (2013) Gene Regulatory Networks Mediating Canonical Wnt Signal Directed Control of Pluripotency and Differentiation in Embryo Stem Cells. Stem Cells.
11. Price FD, Yin H, Jones A, van Ijcken W, Grosveld F, et al. (2012) Canonical Wnt signaling induces a primitive endoderm metastable state in mouse embryonic stem cells. Stem Cells 31: 752-764.
12. De Vries WN, Evsikov AV, Haac BE, Fancher KS, Holbrook AE, et al. (2004) Maternal beta-catenin and E-cadherin in mouse development. Development 131: 4435-4445.
13. Haegel H, Larue L, Ohsugi M, Fedorov L, Herrenknecht K, et al. (1995) Lack of beta-catenin affects mouse development at gastrulation. Development 121: 3529-3537.
14. Huelsken J, Vogel R, Brinkmann V, Erdmann B, Birchmeier C, et al. (2000) Requirement for beta-catenin in anterior-posterior axis formation in mice. J Cell Biol 148: 567-578.
15. Xie H, Tranguch S, Jia X, Zhang H, Das SK, et al. (2008) Inactivation of nuclear Wnt-beta-catenin signaling limits blastocyst competency for implantation. Development 135: 717-727.
16. Kemp C, Willems E, Abdo S, Lambiv L, Leyns L (2005) Expression of all Wnt genes and their secreted antagonists during mouse blastocyst and postimplantation development. Dev Dyn 233: 1064-1075.
17. Kemler R, Hierholzer A, Kanzler B, Kuppig S, Hansen K, et al. (2004) Stabilization of beta-catenin in the mouse zygote leads to premature epithelial-mesenchymal transition in the epiblast. Development 131: 5817-5824.
18. Ferrer-Vaquer A, Piliszek A, Tian G, Aho RJ, Dufort D, et al. (2010) A sensitive and bright single-cell resolution live imaging reporter of Wnt/ss-catenin signaling in the mouse. BMC Dev Biol 10: 121.
19. Yi F, Pereira L, Hoffman JA, Shy BR, Yuen CM, et al. (2011) Opposing effects of Tcf3 and Tcf1 control Wnt stimulation of embryonic stem cell self-renewal. Nat Cell Biol 13: 762-770.
20. Cole MF, Johnstone SE, Newman JJ, Kagey MH, Young RA (2008) Tcf3 is an integral component of the core regulatory circuitry of embryonic stem cells. Genes Dev 22: 746-755.
21. Merrill BJ (2012) Wnt pathway regulation of embryonic stem cell self-renewal. Cold Spring Harb Perspect Biol 4: a007971.
22. Hikasa H, Ezan J, Itoh K, Li X, Klymkowsky MW, et al. (2010) Regulation of TCF3 by Wnt-dependent phosphorylation during vertebrate axis specification. Dev Cell 19: 521-532.

23. Kelly KF, Ng DY, Jayakumaran G, Wood GA, Koide H, et al. (2011) beta-catenin enhances Oct-4 activity and reinforces pluripotency through a TCF-independent mechanism. Cell Stem Cell 8: 214-227.

24. Atlasi Y, Noori R, Gaspar C, Franken P, Sacchetti A, et al. (2013) Wnt Signaling Regulates the Lineage Differentiation Potential of Mouse Embryonic Stem Cells through Tcf3 Down-Regulation. PLoS Genet 9: e1003424.

25. Lindsley RC, Gill JG, Kyba M, Murphy TL, Murphy KM (2006) Canonical Wnt signaling is required for development of embryonic stem cell-derived mesoderm. Development 133: 3787-3796.

26. Wagner RT, Xu X, Yi F, Merrill BJ, Cooney AJ (2010) Canonical Wnt/beta-catenin regulation of liver receptor homolog-1 mediates pluripotency gene expression. Stem Cells 28: 1794-1804.

27. Piepenburg O, Vorbruggen G, Jackle H (2000) Drosophila segment borders result from unilateral repression of hedgehog activity by wingless signaling. Mol Cell 6: 203-209.

28. Theisen H, Syed A, Nguyen BT, Lukacsovich T, Purcell J, et al. (2007) Wingless directly represses DPP morphogen expression via an armadillo/TCF/Brinker complex. PLoS One 2: e142.

29. Jamora C, DasGupta R, Kocieniewski P, Fuchs E (2003) Links between signal transduction, transcription and adhesion in epithelial bud development. Nature 422: 317-322.

30. Blauwkamp TA, Chang MV, Cadigan KM (2008) Novel TCF-binding sites specify transcriptional repression by Wnt signalling. EMBO J 27: 1436-1446.

31. Marks H, Kalkan T, Menafra R, Denissov S, Jones K, et al. (2012) The transcriptional and epigenomic foundations of ground state pluripotency. Cell 149: 590-604.

32. Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, et al. (2010) c-Myc regulates transcriptional pause release. Cell 141: 432-445.

33. ten Berge D, Kurek D, Blauwkamp T, Koole W, Maas A, et al. (2011) Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. Nat Cell Biol 13: 1070-1075.

34. Dravid G, Ye Z, Hammond H, Chen G, Pyle A, et al. (2005) Defining the role of Wnt/beta-catenin signaling in the survival, proliferation, and self-renewal of human embryonic stem cells. Stem Cells 23: 1489-1501.

35. Cai L, Ye Z, Zhou BY, Mali P, Zhou C, et al. (2007) Promoting human embryonic stem cell renewal or differentiation by modulating Wnt signal and culture conditions. Cell Res 17: 62-72.

36. Villa-Diaz LG, Pacut C, Slawny NA, Ding J, O'Shea KS, et al. (2009) Analysis of the factors that limit the ability of feeder cells to maintain the undifferentiated state of human embryonic stem cells. Stem Cells Dev 18: 641-651.

37. Ullmann U, Gilles C, De Rycke M, Van de Velde H, Sermon K, et al. (2008) GSK-3-specific inhibitor-supplemented hESC medium prevents the epithelial-mesenchymal transition process and the up-regulation of matrix metalloproteinases in hESCs cultured in feeder-free conditions. Mol Hum Reprod 14: 169-179.

38. Blauwkamp TA, Nigam S, Ardehali R, Weissman IL, Nusse R (2012) Endogenous Wnt signalling in human embryonic stem cells generates an equilibrium of distinct lineage-specified progenitors. Nat Commun 3: 1070.

39. Davidson KC, Adams AM, Goodson JM, McDonald CE, Potter JC, et al. (2012) Wnt/beta-catenin signaling promotes differentiation, not self-renewal, of human embryonic stem cells and is repressed by Oct4. Proc Natl Acad Sci U S A 109: 4485-4490.

40. Liu P, Wakamiya M, Shea MJ, Albrecht U, Behringer RR, et al. (1999) Requirement for Wnt3 in vertebrate axis formation. Nat Genet 22: 361-365.

41. Sumi T, Oki S, Kitajima K, Meno C (2013) Epiblast ground state is controlled by canonical Wnt/beta-catenin signaling in the postimplantation mouse embryo and epiblast stem cells. PLoS One 8: e63378.

42. Kim H, Wu J, Ye S, Tai CI, Zhou X, et al. (2013) Modulation of beta-catenin function maintains mouse epiblast stem cell and human embryonic stem cell self-renewal. Nat Commun 4: 2403.

43. Fodde R, Brabletz T (2007) Wnt/beta-catenin signaling in cancer stemness and malignant behavior. Curr Opin Cell Biol 19: 150-158.

44. Korkola JE, Houldsworth J, Chadalavada RS, Olshen AB, Dobrzynski D, et al. (2006) Down-regulation of stem cell genes, including those in a 200-kb gene cluster at 12p13.31, is associated with in vivo differentiation of human male germ cell tumors. Cancer Res 66: 820-827.

45. Sperger JM, Chen X, Draper JS, Antosiewicz JE, Chon CH, et al. (2003) Gene expression patterns in human embryonic stem cells and human pluripotent germ cell tumors. Proc Natl Acad Sci U S A 100: 13350-13355.

46. de Jong J, Stoop H, Gillis AJ, van Gurp RJ, van de Geijn GJ, et al. (2008) Differential expression of SOX17 and SOX2 in germ cells and stem cells has biological and clinical implications. J Pathol 215: 21-30.

47. Jia Y, Yang Y, Liu S, Herman JG, Lu F, et al. (2010) SOX17 antagonizes WNT/beta-catenin signaling pathway in hepatocellular carcinoma. Epigenetics 5: 743-749.

48. Zorn AM, Barish GD, Williams BO, Lavender P, Klymkowsky MW, et al. (1999) Regulation of Wnt signaling by Sox proteins: XSox17 alpha/beta and XSox3 physically interact with beta-catenin. Mol Cell 4: 487-498.

49. Zhang W, Glockner SC, Guo M, Machida EO, Wang DH, et al. (2008) Epigenetic inactivation of the canonical Wnt antagonist SRY-box containing gene 17 in colorectal cancer. Cancer Res 68: 2764-2772.

50. Sinner D, Kordich JJ, Spence JR, Opoka R, Rankin S, et al. (2007) Sox17 and Sox4 differentially regulate beta-catenin/T-cell factor activity and proliferation of colon carcinoma cells. Mol Cell Biol 27: 7802-7815.

51. Sur IK, Hallikas O, Vaharautio A, Yan J, Turunen M, et al. (2012) Mice lacking a Myc enhancer that includes human SNP rs6983267 are resistant to intestinal tumors. Science 338: 1360-1363.

SUMMARY

Mammalian development starts from a fertilized egg that initially generates few pluripotent cells which eventually give rise to the embryo proper. This developmental program is well-controlled, reproducible, and has highly predictable outcomes. The primary cells of the developing embryo display a unique potential for multi-lineage differentiation and hence are referred to as pluripotent. Different pluripotent stem cells have been captured and propagated *in vitro* and their culture requirements were investigated accordingly. These cells include, among others, the embryonic stem cells (ESC) derived from the inner cell mass (ICM). Tumorigenic transformation of primordial germ cells (PGC) or gonocytes can also give rise to pluripotent cells known as embryonal carcinoma cells (hECs) which are thought to represent the malignant counterpart of hESCs. Among different signaling pathways, Wnt signaling plays a central role in self-renewal and differentiation of pluripotent cells.

In **Chapter 1**, we outline the current literature on Wnt signaling, as a major stemness pathway, and establish a parallel between its role in embryonic and cancer stem cells. We also discuss germ cell tumors (GCTs) as a model to study the role of stem cells in neoplasia.

In **Chapter 2 and 3**, we employ a series of *Apc*-mutant ESCs encoding for different Wnt signaling levels, to study the mechanisms through which this pathway regulates lineage differentiation in mouse ESCs.

In **chapter 2**, we analyzed the gene expression profiles of mESCs that harbor different *Apc* hypomorphic alleles and found specific *Tcf3* downregulation as well as induction of a novel Wnt-regulated micro-RNA, miR-211, in Wnt-high ESCs. We provide further evidence that Wnt signaling suppresses neuroectodermal differentiation in mouse embryonic stem cells, partly through Tcf3 downregulation and induction of miR-211.

In **Chapter 3** we report that miR-302-367, one of the highly specific miRNAs in mouse ES cells, is negatively regulated by Wnt signaling. We further investigate the impact of this epistatic interaction on ES self-renewal and differentiation.

In **Chapter 4**, we explore the functional role of Wnt signaling in human embryonal carcinoma cells and provide evidence that WNT induction induces a range of diverse differentiation responses among the different employed hEC lines, thus underscoring the context-dependent role of Wnt signaling in mouse and human pluripotent stem cells. The observed role of Wnt in hEC lines adds yet another piece to the puzzle of Wnt signaling in different pluripotent stem cells.

Highly conserved genomic regions (ultra-conserved regions) are frequently transcribed as long non coding-RNAs (>200 bp) in both normal and tumor cells. In **Chapter 5** we describe the identification and functional characterization of CCAT2, a novel WNT-regulated long noncoding RNA located on chromosome 8q24. We show that this non-coding RNA is expressed in microsatellite-stable colorectal cancers where it promotes tumor growth, metastasis, and chromosomal instability.

The interaction between tumor cells and their niche plays a pivotal role in tumor initiation, progression and metastasis. In **Chapter 6**, we studied the role of *S100a4*, a Wnt target gene that is expressed from within the tumor microenvironment in intestinal

neoplasia. By using different transgenic and knockout mouse models, we examined the *in vivo* role of *S100a4* in intestinal tumor initiation and progression. We show that *S100a4* is dispensable for intestinal tumor initiation in both *Apc*- or *Smad4*–mutant mice.

Moreover, desmoids are mesenchymal tumors associated with constitutive activation of Wnt signaling due to *APC* or *CTNNB1* mutations. We show that specific depletion of S100a4 in $Apc^{1638N/+}$ mice significantly reduces desmoids formation.

SAMENVATTING

De ontwikkeling van zoogdieren start met een bevruchte eicel, welke zich allereerst tot een paar pluripotente cellen ontwikkelt om uiteindelijk een compleet embryo te worden. De lijn waarlangs deze ontwikkeling plaatsvindt is strak gecontroleerd, reproduceerbaar en de uitkomst is zeer voorspelbaar. De primaire cellen van het zich vormende embryo bezitten de unieke eigenschap om zich tot verschillende soorten cellen te ontwikkelen en worden daarom ook pluripotente cellen genoemd. Verschillende van deze pluripotente cellen zijn verzameld en *in vitro* gekweekt, waarbij hun verschillende eisen voor de te gebruiken cultuur is bestudeerd. Onder deze cellen bevonden zich, naast anderen, de embryonale stamcellen (ESC) welke afkomstig zijn uit de binnenste celmassa (ICM).

Maligne ontwikkeling van primaire zaadcellen (PGC) of gonocyten kunnen ook pluripotente cellen opleveren: de zogenaamde embryonale kankercellen (hECs). Deze laatste cellen worden gezien als de kwaadaardige tegenhanger van hESCs.. Van de verschillende signaling pathways speelt de Wnt signaling een centrale rol in de zelfvernieuwing en de differentiatie van de pluripotente cellen.

In **Hoofdstuk 1** geven we een overzicht van de bestaande literatuur over Wnt signaling als een belangrijke stemness pathway, en stellen een belangrijke parallel vast in haar rol bij embryonale en bij kankerstamcellen. We bespreken ook zaadceltumoren (GCTs) als een model om de rol van stamcellen in neoplasie te bestuderen.

In de **Hoofdstukken 2 en 3** gebruiken we een serie *Apc*-mutante ESCs welke coderen voor verschillende niveaus van Wnt signaling. Hiermee bestuderen we de mechanismen die deze pathway gebruikt om de lineage differentiation in embryonale stamcellen van de muis (mESCs) te reguleren.

In **Hoofdstuk 2** beschrijven we de analyse van de gen expressie profielen van mECS met verschillende *Apc* hypomorphic alleles. Hierbij vonden we specifieke *Tcf3* downregulation en de inductie van een nieuw Wnt gereguleerd micro-RNA, miR-211, in Wnt-high ESCs. We leveren meer bewijs voor de onderdrukking van neuroectodermale differentiatie door Wnt signaling in mESCs. Deze onderdrukking gebeurt gedeeltelijk door Tcf3 downregulation en de inductie van miR-211.

In **Hoofdstuk 3** beschrijven we dat miR-302-367, een van de zeer specifieke niRNAs in mESCs, negatief wordt gereguleerd door Wnt signaling. Verder onderzoeken we de impact van deze epistatische interactie op zelfvernieuwing en differentiatie van embryonale stamcellen.

In **Hoofdstuk 4** onderzoeken we de functionele rol van Wnt signaling in humane embryonale kankercellen (hECs). We leveren bewijs voor het feit dat WNT inductie een reeks aan verschillende differentiatie reacties geeft onder de verschillende hEC lijnen, waarmee we de omgevingsafhankelijke rol van Wnt signaling in muis en humane pluripotente stamcellen onderstrepen. De gesignaleerde rol van Wnt in hEC lijnen levert een nieuw stukje van de puzzel van Wnt signaling in verschillende pluripotente stamcellen.

Hoog geconserveerde genomische regio's (ultra-conserved regions) worden vaak tot transcriptie gebracht als lange non coding-RNAs (>200 bp) in zowel normale als in tumorcellen. **In Hoofdstuk 5** beschrijven we de identificatie en de functionele van CCAT2, een nieuw WNT-gereguleerd lange non coding RNA welke is gesitueerd op chromosoom

8q24. We tonen de expressie aan van dit non coding RNA in microsatellite-stabiele darmkankers waar het de tumorgroei, metastasering en de chromosomale instabiliteit bevordert.

De interactie tussen tumorcellen en hun niche speelt een centrale rol in tumorontwikkeling, -progressie en metastasering. In **Hoofdstuk 6** bestuderen we de rol van *S100a4*, een Wnt target gen dat tot expressie komt in de tumor micro-omgeving in neoplasia in de darm. Voor de bestudering van de *in vivo* rol van *S100a4* bij het ontstaan en de ontwikkeling van darmkanker hebben we verschillende transgene en knockout muismodellen gebruikt. We tonen aan dat *S100a4* overbodig is voor het ontstaan van darmtumoren in zowel *Apc*- als *Smad4*–mutante muizen. Bovendien zijn desmoids, mesenchymal tumoren, welke geassocieerd worden met constitutieve activering van Wnt signaling veroorzaakt door *APC* of *CTNNB1* mutaties. We tonen aan dat specifieke uitputting van S100a4 in $Apc^{1638N/+}$ muizen de vorming van desmoids significant vermindert.

ACKNOWLEDGMENTS

The research described in this thesis would not have been possible without the help and support of many to whom I am grateful.

First and foremost, I would like to thank you dear **Riccrardo** for giving me the chance of working in your lab, for your trust and for your patience during all these years. I never forget the day when you wrote to me: "Yaser, if you cannot come to The Netherlands, I will travel to UK". Ironically, I had my PhD interview at Gatwick airport and that was the first of a series of events where I realized that besides being a great scientist, you are a wonderful person. I learnt from you not only the science but also the whole package of thinking and behaving like a scientist. I am truly grateful for your support and supervision through all these years.

I would like to thank all members of the PhD committee for joining the PhD Viva and for the great comments and discussions.

My special thanks to our collaborators: Dear **Dr. George Calin**, I feel very lucky that I was part of your research team and that I had the chance of working on the exciting non-coding RNA project. Your enthusiasm and support taught me how "to work hard and smile" as you always said. Dear **Dr. Leendert Looijenga**, I am very happy that we had one of the best experts in germ cell tumors in our institute, thank you for all your scientific support and input through all these years. Dear **Dr. Noona Ambartsumian** and **Dr. Eugene Lukanidin**, thank you for your endless support with the S100a4 project. Without your great collaboration, this work would not have been possible. My special thanks also to Riccardo Spizzo and Ling Hui for the great collaboration.

I would like to thank **Dr. Peter Andrews** for giving me the 1st chance to experience a world-class research environment. Dear Peter, thank you for introducing me to the most wonderful cell types, the plurioptent stem cells. I have to admit that you have imprinted the love for these cells in my entire scientific carrier. I would also like to thank my former MSc supervisor, **Dr. Seyed Javad Mowla.** Dear Dr. Mowla, thank you for teaching me the basic principles of scientific research. You taught me that science is not complete and that there are still so many scientific gaps out there to be filled. Your advices will remain invaluable in my scientific carrier.

My special thanks to my Paranimfs, Mehrnaz and Patrick. Dear **Mehrnaz**, you are a true friend and a dear colleague. Together with **Saeed**, you warmly welcomed me to The Netherlands and made me feel at home while I was thousands of kilometers away from my family. I must say that I even felt sometimes that I am becoming a members of your family. Your kind help and support will always be on my mind, many thanks to both of you. Dear **Patrick**, thank you for standing by me on this special day and for being in the lab whenever I had a question or needed an advice. Some people call you the "everyone's-buddy" and I think it reflects what a kind and nice person you are, thank you for being a true friend.

I would also like to thank all past and present members of Fodde lab.

Dear **Joana** (**Monteiro**), I have shared some of my best years of PhD with you. I truly enjoyed our friendship and I always appreciated your wisdom whenever I needed a second opinion on different subjects. I am very glad that I have a friend like you. Dear **Sabrina**, I have to admit that you are one of most determined persons I have seen. You have been always a friend that I could count on and I am very grateful for that. With you

and Joana I shared some of the best time of my PhD. Dear **Marieke**, you have been always the one behind the scenes who played a great role in my PhD. I am truly grateful for all your countless efforts, help and dedication, without which I could not be in The Netherlands in the first place. Dear **Claudia**, whenever I see you, I still have a question to ask! This reflects the great help you provided during the first year of my PhD and I am glad that I have met a good friend like you. Dear **Rubina**, I cannot say how much I am grateful for all the assistance that you provided with the different projects that I had in the lab. In short, you were one of the few people who I felt relieved whenever you helped me with an experiment, because I knew that everything is going to be performed perfectly! Dear **Andrea**, your help with all the FACS experiments was invaluable. My research without your great expertise and support was not possible and I am very thankful for that. Dear **Rosalie**, I feel very lucky that I shared the office with you. You were truly a nice person, a great colleague and a good friend. Thank you for all the Dutch lessons that you taught me and for all the good advices that you provided me whenever I needed.

I would also like to thank my former master students Juliana, Caterina and Rebecca. Dear **Juliana**, everyone in the lab admits what a lovely and sweet girl you are. I am very glad that you were my first MSc student, thank you for the enthusiasm and dedication that you showed during your MSc project. Dear **Caterina**, we had our ups and downs but at the end we managed to perform our experiments and I would like to thank you for all your efforts. Dear **Rebecca**, you are one of the brightest girls I have seen. Thank you for the nice time that we had and for the great contribution you made in my thesis. I am very proud that I had a student like you.

I would also like to thank you dear **Ron** for all the brainstorming that we had. I always appreciated your critical thinking, your comments and your patience; you always represented a beautiful mind to me. Dear **Katia**, we worked for a short time together in the lab but it was enough to find a good friend like you. I really enjoyed our time and I still remember the great hospitality that you showed in the beautiful Siena.

I would like also to thank all the present PhD students in Fodde lab namely **Joel**, **Matthias** and **Medine**. Thank you for all the nice discussions that we had during and after the lab meetings. I truly enjoyed your company, your comments and your alternative points of view on different subjects.

I would like also to thank all the former members of Fodde lab and in particular Marten van der zee, Joana Brandao, Ivana, Kim, Ingrid, Marti, Youngi, Brechtje, Martino, Maarten van Keulen, Jos, Petra, Lia, Britt, Davide and Ulrike for the support that you provided, for the nice time that we had and for making the lab such a lovely place to work.

Furthermore, I would like to thank all members of Looijenga lab and in particular **Dr. Wolter Oosterhuis**, **Lambert, Hans, Martin, Ronak and Ad** for the great support you provided. I feel very lucky that I could ask your help with all the different projects that we had on germ cell tumors, your support is very much appreciated.

Dear **Ali** and **Sevda**, I feel very lucky that I have good friends like you. Your help and support during the difficult moments and the joy that we shared during the happy times are invaluable, many thanks for being my true friends. Dear **Alireza,** thank you for all the nice time that we had in Rotterdam. The image of you playing the piano, will always remind me of the good old days in Rotterdam.

I would like also to thank the Iranian community in Rotterdam and in particular**: Farhad, Azadeh, Parham, Rahele, Farzin, Sima, Sahar, Farshid, Setare, Mohsen** and **Abbas** for the wonderful time that I had in Rotterdam.

My special thanks to my friends **Dana**, **Claudia**, **Nadia**, **Dellila**, **Akhgar**, **Bianca**, **Gianluca**, **Alessandro**, **Luigi** and **Tahsin-Stefan,** thank you for your friendship and for all the fun that we had. Every one of you is in different part of the world now, but I hope we will always stay in touch.

Finally, few words to my family: My deepest gratitude and thanks to my parents Farid and Farah, my sister Rasha and my brothers Nader, Baher and Hani. Dear dad and mom, you are the most important thing in my life. Your endless love and support have always been the sources of my strength and progress and I am deeply grateful for your hard work and for setting a high moral standard. Dear Rasha, Nader, Baher and Hani, you are my love. Being there with our parents, you have always been providing me the most important thing that I needed and that is the peace of mind, thank you for everything.

PHD PORTFOLIO

Personal Data

Full name	Yaser Atlasi
Date of Birth	21-11-1980
Nationality	Iranian
PhD period	March 2008- September 2013
Department	Experimental pathology
Promoter	Prof.dr. Riccardo Fodde
Research Schools	Erasmus Molecular Medicine School (MolMed) Medisch-Genetisch Centrum Zuid-West Nederland (MGC)

Academc Degrees

BSc. Degree	Cell and Molecular Biology, Faculty of Science, Tehran University, Iran
MSc. Degree	Molecular Genetics, Faculty of Science, Tarbiat Modares University, Iran

Courses

2012	Epigenetic Regulation in health and disease, leiden UMC
2011	From development to disease, ErasmusMC
2011	6th workshop on Innovative mouse models, Leiden UMC
2010	FEBS Advanced Course on Translational Cancer Research
2009	5th workshop on Innovative mouse models, Leiden UMC
2009	Transgenesis, gene targeting and in vivo imaging, Leiden UMC
2009	Photoshop and illustrator CS5 workshop for PhD students and researchers
2008	Laboratory animal course (Artikel 9), ErasmusMC
2008	Workshop on Basic data analysis on gene expression arrays

Oral presentation

2011	21th MGC-symposium, Leiden
2011	5th Dutch stem cell meeting of the DSSCR, Sanquin, Amsterdam
2010	TUMIC consortium (integrated concept of tumour metastasis, FP7) Annual meeting, Paris
2012	TUMIC consortium (integrated concept of tumour metastasis, FP7) Annual meeting, Ghent
2010	FEBS Advanced Course on Translational Cancer Research Annual JNI scientific meeting presentation

Poster presentation

2013	ISSCR Boston 2013
2011	Cell symposia: Stem cell programming and reprogramming, Lisbon
2012	Molmed day

Conferences

2013	ISSCR Annual meeting, Boston
2012	5th Dutch stem cell meeting of the DSSCR, Sanquin, Amsterdam
2011	Cell symposia: Stem cell programming and reprogramming, Lisbon
2010	European Wnt meeting, Stockholm
2009	International Josephine Nefkens Symposium "The Stem of Cancer", Rotterdam
2009	Stem Cells, Development and regulation, Amsterdam
2008	Stem Cells, Development and regulation, Amsterdam

Teaching activities

2012-2013	MSc thesis in Molecular Medicine program: Rebecca van Dorsten, "Evaluating the role of *Wnt*-signaling in genrm cell tumors"
2012-2013	MSc thesis in Molecular Medicine program: Caterina Purini, "Wnt-regulated miRNAs in mouse embryonic stem cells"
2010-2011	MSc thesis in Molecular Medicine program: Juliana Du, "Role of Tcf3 in colorectal carcinogenesis"

ABOUT THE AUTHOR

After finishing high school and having the options to choose between medicine and molecular biology, Yaser entered the cell and Molecular Biology program at Tehran University (Tehran, Iran). He graduated as the first rank honor student and obtained the silver medal in the national Olympiad of Biology. Yaser then performed his MSc program in molecular genetics at Tarbiat Modares University (Tehran, Iran) where he studied the role of OCT4 (the master regulator of pluripotency) in bladder cancer, under the supervision of Dr. Seyed Javad Mowla. After obtaining his MSc degree as the first rank honor student, Yaser visited Sheffield to perform a short research project at Center for stem cell biology, University of Sheffield (Sheffield, UK) under the supervision of Dr. Peter W. Andrews. In Sheffield, Yaser studied the different spliced variants of OCT4 and identified the novel spliced variant, termed OCT4B1. In 2008, Yaser started the PhD program at Erasmus Medical Center (Rotterdam, The Netherlands) where he studied the role of Wnt signaling in mouse ESCs and intestinal neoplasia under the supervision of Dr. Riccardo Fodde.

LIST OF PUBLICATIONS

1. **Atlasi Y***, van Dorsten RT*, Sacchetti A, Stoop H, Rijlaarsdam MA, Oosterhuis JW, Looijenga L, Fodde R
   Ectopic activation of WNT signaling in human embryonal carcinoma cells differentially affect their pluripotency and differentiation capacity in short- and long-term in vitro culture
   Submitted

2. **Atlasi Y**, Noori R, Marolin I, Franken P, Brandao J, Alman BA, Lukanidin E, Ambartsumian N, Fodde R
   The stromal factor S100a4 (fibsroblast specific protein 1) regulates desmoid formation but has limited effect on gastro-intestinal tumorigenesis in Apc-deriven cancer
   Submitted

3. **Atlasi Y**, Looijenga L, Fodde R
   Cancer stem cells, pluripotency, and cellular heterogeneity: a WNTer perspec
   Current Topics in Developmental Biology (in press)

4. Spizzo R*, Ling H*, **Atlasi Y**, et al
   CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer
   Genome Research, 2013 Sep;23(9)

5. Ghazvini M; Sonneveld P; Kremer A; Franken P; Sacchetti A; **Atlasi Y**; Roth S; Joosten R; Smits R; Fodde R
   Cancer stemness in Apc- vs. Apc/KRAS-driven intestinal tumorigenesis
   PLOS one, 2013 Sep 17;8(9):e73872

6. **Atlasi Y**, Noori R, Gaspar C, Franken P, Sacchetti A, Rafati H, Mahmoudi T, Decraene C, Calin GA, Merrill BJ and Fodde R
   Wnt Signaling Regulates the Lineage Differentiation of Mouse Embryonic Stem Cells through Tcf3 Down-regulation
   PLOS Genetics, 2013 May;9(5):e1003424

7. Asadi MH, Mowla SJ, Fathi F, Aleyasin SA, Asadzadeh J, **Atlasi Y**
   OCT4B1, a novel spliced variant of OCT4, is highly expressed in gastric cancer and acts as an anti-apoptotic factor
   Int J Cancer. 2011 Jun 1;128(11)

8. Malakootian M, Mowla SJ, Saberi H, Asadi MH, **Atlasi Y**, Shafaroudi AM
   Differential expression of nucleostemin, a stem cell marker, and its variants in different types of brain tumors
   Mol Carcinog. 2010;49(9):818-25

9. Mowla SJ, Jafarnejad SM, **Atlasi Y**
   Detection of OCT-4 in Bladder Cancer: Role of Cancer Stem Cell
   Methods of Cancer Diagnosis, Therapy and Prognosis, 2010, Volume 6, Part 4, 211-226,Springer

10. Shafaroudi AM, Mowla SJ, Ziaee SA, Bahrami AR, **Atlasi Y**, Malakootian M
    Overexpression of BMI1, a polycomb group repressor protein, in bladder tumors: a preliminary report
    Urol J. 2008; 5(2):99-105

11. **Atlasi Y**, Mowla SJ, Ziaee AM
    Expression of survivin and its splice variants survivin-2B and survivin-DEx3 in Bladder cancer
    Cancer Detect Prev. 2009;32(4):308-13

12. **Atlasi Y**, Mowla SJ, Ziaee AM, Gokhale P, Andrews PW
    OCT4 spliced variants are differentially expressed in human pluripotent and nonpluripotent cells
    Stem Cells. 2008;26(12):3068-74

13. **Atlasi Y**, Mowla SJ, Ziaee AM, Bahrami AR
    Oct3/4, an embryonic stem cell marker, is highly expressed in bladder cancer
    Int J Cancer, 2007; 120:1598-602