12-2015

IN VIVO FUNCTIONAL SIGNIFICANCE OF CCAT2 LONG NON-CODING RNA IN MYELODYSPLASTIC SYNDROME

MAITRI Y. SHAH

IN VIVO FUNCTIONAL SIGNIFICANCE OF *CCAT2* LONG NON-CODING RNA IN MYELODYSPLASTIC SYNDROME

by

Maitri Yogen Shah, M.S.

APPROVED:

_____
George A. Calin, M.D., Ph.D.
Supervisory Professor

_____
Gabriel Lopez-Berestein, M.D.

_____
Xiongbin Lu, Ph.D.

_____
Anil K. Sood, M.D

_____
M. James You, M.D., Ph.D.

APPROVED:

_____

Dean, the University of Texas

Graduate School of Biomedical Sciences at Houston

IN VIVO FUNCTIONAL SIGNIFICANCE OF *CCAT2* LONG NON-CODING RNA IN MYELODYSPLASTIC SYNDROME

A

DISSERTATION

Presented to the Faculty of
The University of Texas
Health Science Center at Houston
and
The University of Texas
MD Anderson Cancer Center
Graduate School of Biomedical Sciences

in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

By

Maitri Yogen Shah, M.S.
Houston, Texas
August, 2015

Acknowledgements

I would like to extend my sincere gratitude to my mentor, Dr. George A. Calin for his persistent guidance and support. I have learned a lot from him, and I continue to learn more each day I work with him. His energy and enthusiasm for science has motivated me to work hard and add value with my work. He has always encouraged me to strive for excellence and has immensely influenced my professional development.

I would also like to thank my committee members, Drs. Anil Sood, Xiongbin Lu, M. James You and Gabriel Lopez-Berestein for their time, support and guidance. Their help and feedback has helped me become a better scientist. I really appreciate them taking out time for me in their busy schedule.

My friends and colleagues in the Calin Lab hold a special place in my heart. They have helped me immensely during this journey. I would like to especially thank Roxana Redis and Katrien Van Roosbroeck for their support and friendship, both inside and outside the lab.

I cannot express enough my love and gratitude for my parents, for their unconditional love, support and belief in me. They have made countless sacrifices for me, and always encouraged me to pursue my dreams. While thousands of miles away, they have motivated me every single day to keep moving forward.

I want to extend my deep appreciation to my closest friends, Aarthi and Shubi for putting up with me for the last five years. They have supported me through some

of the best and worst times, and they have been my pillars of strength throughout the PhD journey. They are my family away from home. Last, but not the least, I want to thank Sriharesh for being next to me on this crazy journey and helping me become a better person. This achievement would be incomplete without you.

During my PhD, I was fortunate to receive awards that afforded me unique opportunities and recognized my work. I would like to thank the Rosalie B. Hite Fellowship Committee for presenting me the Rosalie B. Hite graduate Fellowship in cancer research Fellowship. I would like to thank the Graduate School of Biomedical Sciences (GSBS) for choosing me to receive the esteemed Floyd Haar, M.D., Endowed Memorial Research Award in Memory of Freda Haar. I would also like to thank the Experimental Therapeutics Academic Program, Dr. Varsha Gandhi and Lidia Vogelsang for helping me throughout graduate school and also providing me opportunities to develop my leadership and communication skills. In addition, the faculty and staff at GSBS including: Dr. William Mattox, Joy Lademora, Bunny Perez, Dr. Andrew Bean, Carol Helton, Brenda Gaughan, and Dr. Eric Solberg, have all helped me at different stages of my time here.

In vivo functional significance of _CCAT2_ long non-coding RNA in myelodysplastic syndrome

Maitri Y. Shah, M.S.

Supervisory Professor: George A. Calin, M.D., Ph.D.

Abstract

Long non-coding RNAs form the largest part of the mammalian non-coding transcriptome and control gene expression at various levels including chromatin modification, transcriptional and post-transcriptional processing. Although the underlying molecular mechanisms are not yet entirely understood, lncRNAs are implicated in initiation and progression of several cancers. *CCAT2* is a lncRNA that spans the highly conserved 8q24 region associated with increased risk for various cancers. *CCAT2* has been shown to play an important role in inducing chromosomal instability and supporting cell proliferation and cell cycle arrest. However, a causal role of *CCAT2* in initiation of tumorigenesis and the importance of G/T SNP in *CCAT2*-induced phenotype still remains to be resolved. The purpose of this study was to elucidate the role of *CCAT2* and its specific alleles (G/T) in regulation of cellular processes that drive tumorigenesis using a genetically engineered mouse model. We generated transgenic mice for each *CCAT2* allele using random integration approach in C57Bl6/N background.

CCAT2(G/T) mice displayed spontaneous induction of widespread pancytopenias with splenomegaly and hepatomegaly. CCAT2(G/T) BM biopsies

displayed severe myeloid and erythroid hyperplasia with enhanced proliferation and excessive apoptosis, along with extramedullary hematopoiesis in spleen and liver. Percentage of HSPCs was significantly reduced in BM of these mice, with increased presence of immature erythroid blasts and granulocyte-macrophage progenitors suggesting a block in differentiation. HSPCs of CCAT2(G/T) mice showed increased frequency of cytogenetic aberrations, including breaks and chromosomal fusions. However, these mice don't develop AML, suggesting *CCAT2* is critical in initiation of MDS. Microarray expression profiling of CCAT2(G/T) HSPCs revealed enrichment of pathways associated with epigenetic regulation, chromosomal instability and cell cycle regulation. We further identified significantly higher *CCAT2* expression in the MDS patients as compared to healthy volunteers. Interestingly, patients with AML had significantly lower expression of *CCAT2* as compared to patients with only MDS.

Based on these data, we conclude that *CCAT2* plays an important role in regulation of normal hematopoiesis, and its deregulation can lead to MDS. *CCAT2* lncRNA can be developed into a diagnostic and prognostic marker, as well as a novel intervention target for MDS therapy. CCAT2(G/T) mice can serve as a robust model for studying initiation of de novo MDS and as a pre-clinical model for evaluation of new therapies for low-risk MDS.

Table of Contents

List of Figures

List of Tables

CHAPTER I: Introduction

(Parts of this section were adapted with permission in part from Shah, MY and Calin, GA, MicroRNAs as therapeutic targets in human cancers, Wiley Interdisciplinary Reviews - RNA, 5 (4), 537-548 (2014); Shah, MY and Calin, GA, Regulatory RNA, Encyclopedia of Genetics, 2013; and Shah, MY and Calin, GA, MicroRNAs and Cancer Therapeutics, Principles of Molecular Diagnostics and Personalized Cancer Medicine, 2012)

The non-coding RNA revolution

The complexity of the human genome, transcriptome and proteome organization necessitates a robust and dynamic regulatory system that helps facilitate the coordinated performance of each functional element. RNA molecules have emerged as excellent candidates for this job by the virtue of their diversity and versatility. Historically, RNA had been relegated to a secondary messenger for DNA, conveying protein-coding instructions to diverse intracellular destinations. However, accumulating evidence over the past decade has recognized RNA as a major player in some of the cell's most vital activities (Mercer et al., 2009). The large portion of the human genome, which we only now know gets transcribed but not necessarily translated and is usually considered the "dark matter" of the genome, represents a hidden layer of regulation imposed by the newly identified class of regulatory RNAs. These are a heterogeneous group of non-protein-coding RNAs (ncRNAs) that play an important role in controlling gene activity at different levels.

From genomic analyses, it is evident that over 95% of the human transcriptional output is from ncRNAs (Szymanski and Barciszewski, 2002). A part of this comprises of housekeeping RNAs, such as rRNA (ribosomal RNA) and tRNA (transfer RNA), which are constitutively expressed and are required for normal functioning of the cells. The largest fraction consists of regulatory RNAs that help coordinate and regulate the gene activity. Regulatory ncRNAs can be broadly divided into two classes based on their size, small ncRNAs (less than 200bp) and

long ncRNA (longer than 200bp). Both small and long ncRNAs can then be further categorized according to the location of their transcript origin.

Small regulatory ncRNA (sncRNA) mainly include microRNAs (miRNAs) and silencing RNAs (siRNAs). miRNAs are small (19-24 nt), endogenous, highly conserved ncRNAs which negatively regulate gene expression post-transcription. miRNAs are usually transcribed by RNA polymerase II (RNAPII) as autonomous transcription units, or as clusters from a polycistronic transcription unit which can be located in the exonic or intronic regions of non-protein coding regions, or in the intronic region of protein-coding transcription units (Ambros, 2004; Bartel, 2004). miRNA biogenesis is a multi-step process involving coordinated cropping and dicing by RNase III enzymes Drosha and Dicer to produce mature miRNA. These are incorporated into the RISC (RNA-induced silencing complex), which then cause gene suppression by binding at the 3'-UTR of mRNA. Endo-siRNAs are short, endogenous, double-stranded RNA duplexes which can cause complete gene silencing by RNA interference (RNAi) (Carthew and Sontheimer, 2009). Another important class of sncRNAs are piRNAs (Piwi-interacting RNAs), that are transcribed from transposable elements (TEs) and function in RNAi and related pathways in association with Piwi proteins of the Argonaute family (Thomson and Lin, 2009). Other functional classes of sncRNAs with limited experimental evidence include promoter-associated short RNAs, termini-associated short RNAs, 3'-UTR derived RNAs, splice-site RNAs, transcription start site associated RNAs, small nuclear RNAs and small nucleolar RNAs.

Long regulatory ncRNAs (lncRNAs) are a recently identified class of regulatory RNAs with important implications in human diseases (Mercer et al., 2009; Spizzo et al., 2009). They are defined as ncRNAs greater than 200 nucleotides in length, and encompass a board spectrum of different RNA classes, including long intergenic RNAs (formerly lincRNAs), enhancer RNAs (eRNAs), circular RNAs, pseudogenes and sense and antisense RNAs overlapping other protein-coding or non-coding transcripts (Xue and He, 2014). They make up the largest portion of the human non-coding transcriptome and act through several distinct mechanisms. The best functionally characterized are the lincRNAs (long intergenic ncRNAs), including the most studied *HOTAIR* (Hox transcript antisense RNA) gene. Other lncRNAs includes the transcribed ultraconserved regions (T-UCRs), transcripts from the regions in the genome showing extraordinary conservation patterns across several species. An example of T-UCRs is uc.73A, which acts as an oncogene in colon cancer (Calin et al., 2007). Another class are single stranded antisense RNAs which are transcribed in the opposite orientation from a protein-coding gene and are complementary to the mRNA sequence. Antisense RNAs have been implicated in gene silencing at the chromatin and transcriptional levels. An additional class of lncRNAs is long intronic ncRNAs, defined based on their site of transcription (Louro et al., 2009). These are generally cis-acting elements, which regulate the genome locally. Recently, a new type of ncRNA transcribed from the enhancer regions, eRNAs, have been identified which are also important in transcriptional regulation (Orom et al., 2010). The biological context and significance of several of these non-coding transcripts still needs to be evaluated. They might play a hitherto

unrecognized critical regulatory role in controlling gene expression in a variety of organisms.

Widespread functionality of long- and micro-non-coding RNAs

The discovery of regulatory RNAs has greatly accelerated since the availability of the complete human genome sequence. The broad functional repertoire of regulatory ncRNAs makes it difficult to classify them based on their activity. These can regulate several biological processes using a wide array of molecular mechanisms. Majority of the regulatory ncRNAs act as *cis*-regulatory elements and control the activity of the neighboring protein-coding genes (Xue and He, 2014). However, certain RNAs also act *trans*-specific, controlling the activity of genes located on different chromosome loci. This opens up the entire landscape of the human genome for RNA-mediated regulation. However, unlike the housekeeping RNAs, regulatory ncRNAs are expressed in tissue- and cell-specific manner, implying a developmentally regulated mechanism of their expression. The abundant and pervasive transcription of these ncRNAs from the human genome, but at low levels, reinforces the functional regulatory nature of these transcripts. These RNAs play a critical role in regulating each stage of the genetic process, from how the information is stored, processed and transmitted.

Regulatory ncRNAs play an important role in chromatin modification and epigenetic regulation. LncRNAs mediate epigenetic changes by recruiting histone-modifying or chromatin remodelling complexes and DNA methyltransferases to specific gene loci (Lee, 2012). One such long ncRNA, *HOTAIR*, is transcribed from

the *HOXC* cluster locus, and functions in *trans*-specific manner to silence transcription at *HOXD* locus by recruiting Polycomb chromatin remodeling complexes (Gupta et al., 2010). LncRNAs are also important in regulation of chromatin architecture and epigenetic memory. Small regulatory RNAs, like piRNAs have been associated with heterochromatin formation and transposon activity in various species (Thomson and Lin, 2009). LncRNAs also play an important role in dosage compensation by mediating X chromosome inactivation. *Xist*, a long ncRNA from the *Xist* locus (Plath et al., 2002), and its antisense transcript, *Tsix* (Sado et al., 2005), act *cis*-specific and together propagate the epigenetic silencing of an individual X chromosome in women. Another phenomenon that closely resembles dosage compensation is genetic imprinting, wherein one of the two parental alleles of a gene gets preferentially silenced to maintain parent-of-origin-specific gene expression (Koerner et al., 2009). In humans, the well-characterized imprinting lncRNAs are *H19* and *Kcnq1ot1*, which sustain silencing of the *IGF2* and *KCNQ1* genes on the maternal and paternal alleles respectively. Certain small ncRNAs are also important for regulation of dosage compensation and genetic imprinting processes (Koerner et al., 2009).

ncRNAs have been also shown to regulate gene expression at the transcriptional level through direct and indirect mechanisms. LncRNAs interact globally with the transcriptional machinery of the RNAPII, regulating the assembly of initiation complex or inhibiting transcription elongation (Rinn and Chang, 2012). Examples are *Alu* elements, heat shock transcripts that are transcribed in response to stress. These bind to RNAPII and prevent formation of active pre-initiation

complexes thus allowing for rapid gene repression in response to stress (Mariner et al., 2008). LncRNAs also afford gene-specific transcription regulation by interacting with specific transcription factors or by acting themselves as co-factors, enhancers or promoters (Rinn and Chang, 2012). For example, cell-lineage-specific-factors such as forkhead box A1 (FoxA1) have been shown to regulate the expression of androgen receptor and reprogramming of the hormonal response by causing a switch in the binding of the androgen receptor to certain enhancers (Yang et al., 2013).

LncRNA can control several steps in the post-transcriptional processing of mRNAs, including splicing, editing, transport and degradation (Wang and Chang, 2011). Post-transcriptional gene silencing is achieved mainly by small ncRNAs, mainly miRNAs and siRNAs. miRNAs bind to the 3'-UTR of the mature mRNA to achieve mRNA translational suppression (by imperfect complementarity) or mRNA degradation (by perfect complementarity), thus regulating the expression of hundreds of genes (Bartel, 2004; Carthew and Sontheimer, 2009). Similarly, siRNAs usually target homologous sequences with perfect complementarity and cause gene silencing via the RNAi pathway (Carthew and Sontheimer, 2009). miRNAs play an important role in multiple biological processes, including developmental timing, embryogenesis, cell differentiation, organogenesis, metabolism, apoptosis and various diseases, including cancers. Few miRNAs can also promote gene expression and activate translation under conditions of stress (Vasudevan et al., 2007). Additionally, specific miRNAs that carry a distinct hexanucleotide terminal motif, like *miR-29b*, were found enriched in the nucleus, suggesting extra miRNA

functions in different subcellular compartments (Hwang et al., 2007). It has been shown that miRNAs in the nucleus may act at the promoter level affecting transcription**,** as exemplified by binding of *miR-373* to the *CDH1* promoter and stimulating transcription (Place et al., 2008). Also, a novel function for miRNAs called “decoy activity” was reported. Mechanistically, miR-328 interacts with hnRNP-E2 and prevents its binding to the *CEBPA* (CCAAT/enhancer-binding protein alpha) intercistronic mRNA region; this restores C/EBPa expression which, in turn directly enhances miR-328 transcription (Eiring et al., 2010). Other classes of ncRNAs that act post-transciptionally are antisense RNAs, which are involved in alternate splicing (Mercer et al., 2009). Two cancer-associated lncRNA, *MALAT1* and *NEAT1*, were found to be important in RNA splicing, editing and transport (Chen and Carmichael, 2009; Tripathi et al., 2010). *MALAT1* has been shown to sequester phosphorylated serine/arginine-rich splicing factors (SRSFs), inducing alternative splicing of pre-mRNAs (Tripathi et al., 2010).

Few lncRNAs can directly modulate protein localization and function (Willingham et al., 2005). Similar interactions have also been reported between lncRNAs and small ncRNAs, demonstrating an important possibility of self-regulation by regulating RNAs. On certain occasions, mRNAs can also function at RNA level, by inducing translation (Yoon et al., 2012). Recently, a novel biological role was identified where mRNAs are expressed as pseudogenes, which can then regulate coding gene expression (Johnsson et al., 2013; Karreth et al., 2011). These data reveal an unexpected non-coding function for mRNA. Thus, regulatory RNAs act as

a part of a larger molecular circuitry with complex, interlaced regulatory networks that function at several levels in the human cell.

Figure 1 summarizes the various potential functions of lncRNAs that currently identified. Depending on the cellular localization of lncRNAs, nuclear lncRNAs can regulate transcription by acting as enhancer RNA (eRNA), for example regulation of p53 by p53 eRNAs (Melo et al., 2013) **(A)**, by recruiting chromatin modifying complexes, for example regulation of PRC2 complex by *HOTAIR* (Gupta et al., 2010) **(B)**, or by regulating transcription factor activity, like by association of PANDA lncRNA with transcription factor NF-YA (Hung et al., 2011) **(C)**. Moreover, they can regulate gene expression by acting on the spatial conformation of chromosomes, as illustrated by alteration of X chromosome by *Xist* RNA (Plath et al., 2002) **(D)** or by influencing pre-mRNA splicing, for example Malat1 co-localises with pre-mRNA-splicing factor SF2/ASF and CC3 antigen in the nuclear speckles (Bernard et al., 2010; Tripathi et al., 2010) **(E)**. Cytoplasmic lncRNAs can regulate mRNA expression by regulating mRNA stability, as in the example of NEAT1 that is involved in nuclear retention of mRNAs (Chen and Carmichael, 2009) **(F)**, mRNA translation as shown by neuron-specific antisense Uchl1 that specifically promotes the translation of UCHL1 under rapamycin treatment (Carrieri et al., 2012) **(G)**, or by competing for microRNA binding, like the pseudogene PTENP1 that competes with its coding counterpart, the tumor-suppressor PTEN, for a set of miRNAs (Johnsson et al., 2013; Karreth et al., 2011) **(H)**. In addition, few lncRNAs contain small open reading frames (ORFs) that can be translated in biological active small peptides, as described by Slavoff et al, 2013 (Slavoff et al., 2013) **(I)**.

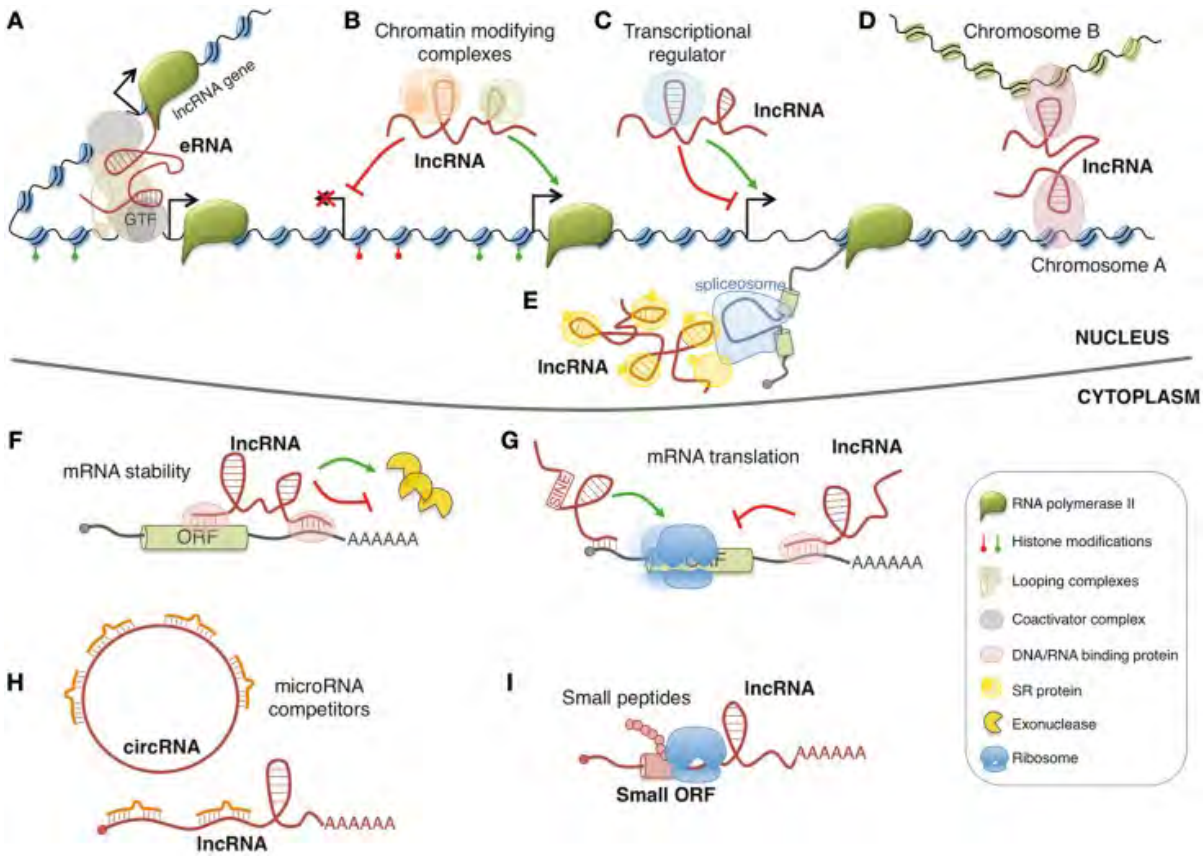


Figure 1: Models of lncRNA functions.

The figure describes the different functional mechanisms of lncRNAs

This figure is used with permission and originally published by Morlando M, Ballarino M, Fatica A. in Long Non-Coding RNAs: New Players in Hematopoiesis and Leukemia. *Frontiers in Medicine*. 2015; 2:23 (Morlando et al., 2015).

Non-coding RNAs in cancer

Cancer represents a complex multistep genetic disorder characterized by deregulation of homeostasis at the genomic, transcriptomic and proteomic levels (Vogelstein and Kinzler, 2004). The preferred choice of treatment for most human malignancies for the last century has been surgery and chemotherapy. However, the severe toxicities, adverse side effects and the poor quality of living associated with the chemotherapeutic treatment emphasize the need for new therapeutic interventions for cancer patients. The major advances in the last decade have focused on designing novel targeted therapy capable of targeting specifically the malignant cells in a more rational way. The advent of miRNAs and lncRNAs provide an additional layer of gene regulation on a broad spectrum of biological pathways by fine-tuning protein expression levels. By the virtue of their ability to target multiple protein-coding genes and their aberrant perturbations in widespread cancers, miRNAs have emerged to be promising novel therapeutic targets and intervention tools. With the advances in next-generation sequencing and genome wide association studies, the next decade will make possible the therapeutic targeting of lncRNAs. The following sections discuss the involvement of ncRNAs in cancer and their potential application in diagnosis and treatment of cancer.

Ubiquitous ncRNA alterations in cancers

Aberrant expression of miRNA and other ncRNA levels has been implicated in a broad spectrum of human diseases, including autoimmune, cardiovascular and psychiatric diseases, diabetes and cancers (Trang et al., 2008). A growing body of

evidence suggests that these ncRNAs play a vital role in cancer predisposition, initiation, maintenance, progression and metastasis (Esquela-Kerscher and Slack, 2006). miRNA expression profiling studies have identified a unique miRNA signature profile that can differentiate normal from cancer tissues and also classify different tumor types and grades (Calin and Croce, 2006). Recent expression studies have identified considerable cancer-specific alterations in lncRNAs too, in tumor-specific manner at distinct stages of cancer progression. lncRNA *MALAT1* is aberrantly overexpressed during metastasis of non-small cell lung cancer, and acts as a prognostic marker for poor survival (Xue and He, 2014). LncRNAs *HOTAIR* and *HULC* also exhibit strong expression level specifically in cancer (Gupta et al., 2010; Panzitt et al., 2007).

The widespread dysregulation of ncRNAs can be explained by their frequent location in cancer-associated hotspots of the human genome, including fragile sites, minimal regions of amplification, loss of heterozygosity sites, and common breakpoint regions (Calin et al., 2004). Other mechanisms of this dysregulation include chromosomal deletions or translocations of regions with ncRNA genes, epigenetic regulation of ncRNA expression, alterations in ncRNA promoter activity by oncogenes and tumor suppressor genes, and presence of mutated ncRNA structural variants (Croce, 2009).

ncRNAs function as oncogenes or tumor suppressor genes

ncRNAs including miRNAs and transcribed-ultraconserved regions (T-UCRs) contribute to tumorigenesis by functioning as oncogenes or tumor suppressor genes

(TSGs) (Calin et al., 2007; Croce, 2009; Esquela-Kerscher and Slack, 2006). miRNAs that target oncogenes and whose expression is lost in most human cancers are classified as tumor suppressor miRNAs (TSmiRs). Loss of TSmiR expression due to somatic alterations or germline mutations can initiate or enhance tumorigenecity. One of the important examples of TSmiRs is miR-15b/miR-16-1 cluster, which is frequently deleted in B-cell chronic lymphocytic leukemia (CLL) and targets several oncogenes such as *BCL2*, *CCND1* and *WNT3A* (Calin et al., 2008). A knockout mouse model targeting the miR-15b/miR-16-1 cluster recapitulated the CLL-associated phenotype, validating the *in vivo* functionality of these TSmiRs (Klein et al., 2010). Other important TSmiRs are the let-7 family members, miR-29, miR-34 family, miR-122 and miR-143/145 cluster (Croce, 2009). In contrast, overexpressed miRNAs that promote tumorigenecity by targeting TSGs are classified as oncomiRs (Esquela-Kerscher and Slack, 2006). One of the most prevalent oncomiR is miR-21, which is overexpressed in almost all human cancers and targets *PTEN* and *PDCD4* (Asangani et al., 2008). Spontaneous tumorigenesis in transgenic mice overexpressing miR-155 has also established miR-155 as a bona fide oncomiR (Costinean et al., 2006). Other miRNAs with oncogenic function include miR-17-92 cluster, miR-155, miR-200 family, miR-221/222 cluster and miR-372/373 cluster (Esquela-Kerscher and Slack, 2006). Another example includes the classic lncRNA *Xist* with a potent tumor suppressor effect. The deletion of *Xist* in the blood compartment leads to an aggressive myeloproliferative neoplasm and myelodysplastic syndrome in mice due to aberrant X reactivation and multiple autosomal changes (Yildirim et al., 2013).

ncRNAs in cancer metastasis

miRNAs and lncRNAs have also been shown to act as activators or suppressors of tumor metastasis. One such miRNA is miR-10b, which promotes metastasis of breast cancer to the lungs by targeting *HOXD10*, and thus increasing the expression of *RHOC*. HOTAIR lncRNA has also been implicated in breast cancer metastasis.

OncomiR addiction

The concept of addiction or dependency of tumor cells on activating mutations in certain oncogenic factors for survival and proliferation has been recently demonstrated for oncogenic miRNAs. Medina et al (Medina et al., 2010) developed a conditional transgenic mouse overexpressing miR-21 and showed that these mice developed spontaneous pre-B malignant lymphoid-like phenotype. In the absence of miR-21, malignant cells undergo apoptosis and the tumors regress. These *in vivo* experimental findings from transgenic mice strongly advocate the causative role of miRNAs in carcinogenesis, and further support their use for therapeutic intervention.

ncRNAs as biomarkers for cancer

Because of the widespread dysregulation of ncRNAs in all types of tumors, miRNA and lncRNA expression profiling in cancer patients has been a valuable signature classifier (Calin and Croce, 2006). The distinct ncRNA expression patterns between the normal and cancer tissues serve as diagnostic, predictive and

prognostic biomarkers. miRNA signatures can classify tumors of different types and grades, and also correlate with chemotherapy and incidence of drug resistance patterns. The recent detection of stable miRNAs in serum has opened up new avenues for non-invasive biomarkers in cancer prognosis (Cortez et al., 2011).

'Sponge'-based regulatory network

The recent revolutionary discovery of lncRNAs functioning as 'sponges' (Salmena et al., 2011) lend further support to the potential use of ncRNAs as therapeutic agents. The sponge hypothesis states that the extensive human transcriptome, including transcribed pseudogenes, mRNAs and lncRNAs interact or crosstalk with each other through the miRNA responsive elements (MREs), establishing a comprehensive intricate regulatory network. For example, tumor suppressor PTEN is finely regulated by its ceRNAs, PTENP1 pseudogene and ZEB2 mRNA (Karreth et al., 2011; Tay et al., 2011). In the context of miRNA therapy, miRNA modulation may have more profound manifestations in an as-yet-uncharacterized RNA-dependent aspect. miRNAs that form autoregulatory loops with other ncRNAs might undergo analogous genetic alterations in neoplasia via similar underlying mechanisms and pathways. Thus, therapeutic modulation of miRNA levels might shift the balance and set up a cellular cascade enabling a more pronounced biological effect than previously anticipated. This discovery reveals a whole new realm of therapeutic possibilities for human cancers. More research to identify other cancer-associated sponge-RNA-miRNA networks needs to be done before the full therapeutic potential of such regulatory loops can be exploited. The advent of ncRNAs has expanded the arena of novel potential targets available for

therapy. ncRNA therapy in combination with the conventional therapy might prove to be a better therapeutic option. However, insights on the importance of lncRNAs in human diseases are still developing. The new genomic and bioinformatic tools should be useful in identifying the whole spectrum of ncRNAs and characterizing their functional significance. More unbiased, focused research on the mechanistic basis of these ncRNAs is essential to understand their potential applications in medicine.

Figure 2 describes the lncRNAs and their putative functions in cancer. **Table 1** provides a list of known lncRNAs that play a critical role in cancer pathogenesis.

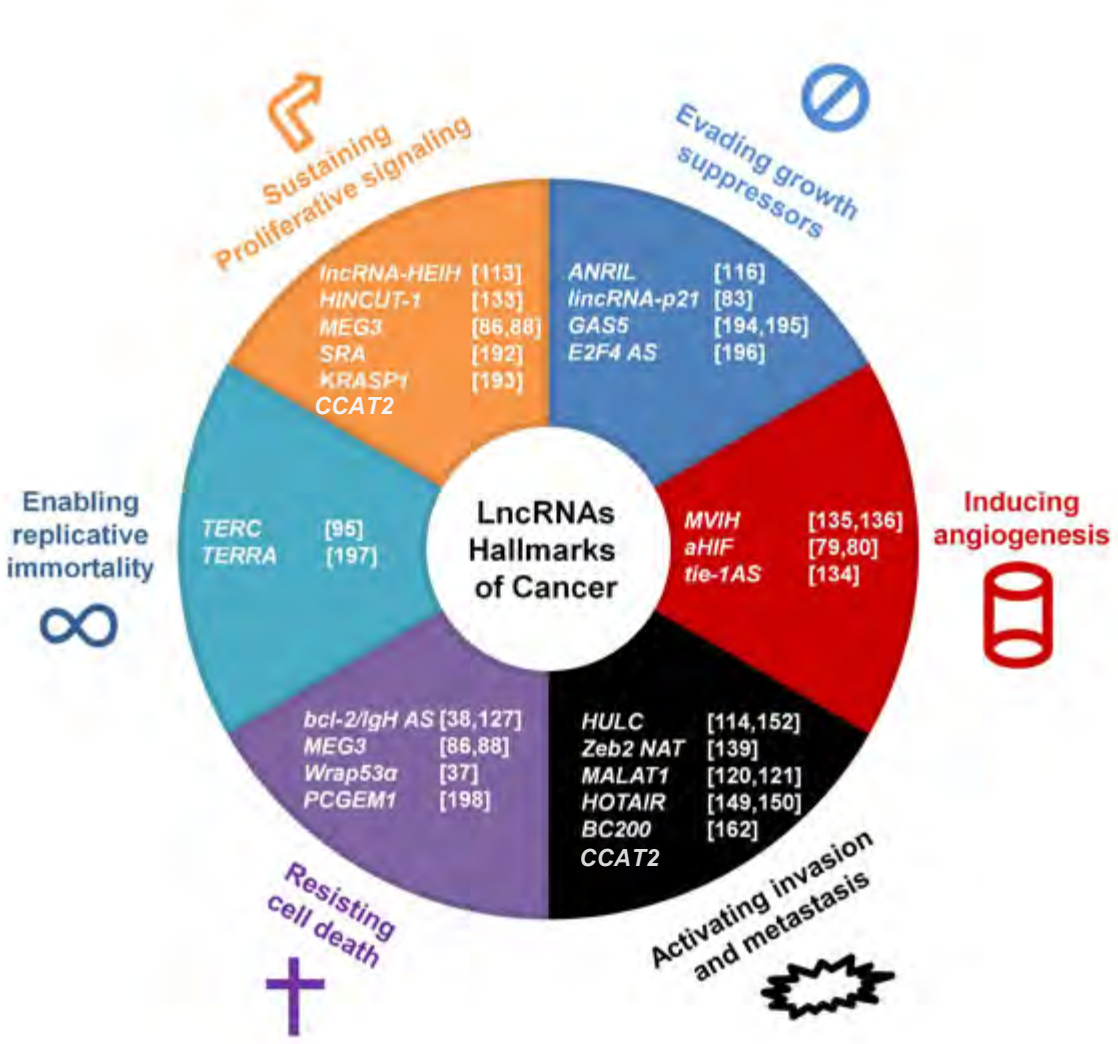


Figure 2: LncRNAs impact the hallmarks of cancer. The six hallmarks of cancer with selected associated lncRNAs that are involved in cancer onset and progression.

This figure is used with permission and originally published by Di Gesualdo F, Capaccioli S, Lulli M. (2014) in A pathophysiological view of the long non-coding RNA world. *Oncotarget* 5(22): 10976–10996 (Di Gesualdo et al., 2014).

Table 1: List of long non-coding RNAs with known function in cancers

lncRNA	Function	Involvement in cancer	Refs
AIRN	Imprinting, Chromatin-mediated repression, transcription interference	hepatocellular carcinoma	(Oliva et al., 2009)
HIF1a-AS	Messenger RNA decay	renal cancer	(Thrash-Bingham and Tartof, 1999)
ANRIL	Chromatin-mediated repression at the *INK4b–ARF–INK4a* locus	acute lymphoblastic leukemia, gastric, non-small cell lung, prostate cancers	(Iacobucci et al., 2011; Nie et al., 2015; Zhang et al., 2014)
GAS5	Repression of glucocorticoid receptor-mediated transcription	breast, kidney, prostate cancer	(Mourtada-Maarabouni et al., 2009; Pickard et al., 2013; Qiao et al., 2013)
HOTAIR	Chromatin-mediated epression at the *HOXD* locus	breast, gastric, lung, colorectal cancers	(Cheng et al., 2015; Gupta et al., 2010; Niinuma et al., 2012; Wu et al., 2014b; Zhao et al., 2014a)
HOTTIP	Chromatin-mediated activation at the *HOXA* locus	hepatocellular, pancreatic, tongue squamous cell carcinoma	(Cheng et al., 2015; Li et al., 2015; Quagliata et al., 2014; Tsang et al., 2015; Zhang et al., 2015a)
HULC	Downregulation of miRNA-mediated repression by miRNA-sequesteration	colorectal, hepatocellular carcinoma	(Du et al., 2012; Panzitt et al., 2007; Peng et al., 2014; Wang et al., 2010; Zhao et al., 2014b)
MALAT1	Scaffolding of subnuclear domains	bladder, breast, liver, lung, gastric cancers	(Ji et al., 2014; Jiang et al., 2014; Liu et al., 2014; Wu et al., 2014a; Yang et al., 2015)
MEG3	regulation of p53 transcriptional activation	cervical, non-small cell lung,	(Sun et al., 2014; Xia et al., 2015;

		gastric, ovarian cancers, pituitary adenoma	Yin et al., 2015)
Neuroblastoma *MYC* (NAT)	Inhibition of neuroblastoma *MYC* intron 1 splicing	neuroblastoma	(Krystal et al., 1990)
PTENP1 pseudogene	Upregulation of PTEN by sequestration of miRNAs	prostate cancer	(Johnsson et al., 2013; Karreth et al., 2011)
TERC	Telomere template	prostate cancer	(Bojovic and Crowe, 2011; Jones et al., 2012; Visnovsky et al., 2014; Wang et al., 2014)
XIST	X inactivation	Breast, cervical, ovarian, hematological cancers	(Kawakami et al., 2004; Laner et al., 2005; Vincent-Salomon et al., 2007; Yildirim et al., 2013)

CCAT2 - a novel oncogenic long non-coding RNA

Cancer-associated genomic regions (CAGR) are regions showing high frequency of cancer related abnormalities, such as loss of heterozygosity or amplifications (Calin et al., 2004). These have long been implicated in cancer predisposition; however the underlying molecular mechanisms responsible for the instability still remain a mystery. One such widely studied CAGR is the 8q24.21 genomic region that has been associated with several human cancers, including prostate, colorectal, breast and bladder cancer (Ghoussaini et al., 2008). The region is highly unstable, with numerous-genome wide association studies (GWAS) mapping a large number of SNPs (single nucleotide polymorphism) to within a 1.5-Mb gene-free region on this locus. One SNP of particular importance is the rs6983267, with the G allele of the SNP conferring increased risk of colorectal, prostate and ovarian cancers (Gruber et al., 2007; Haiman et al., 2007; Tomlinson et al., 2007; Zanke et al., 2007). In colorectal cancer (CRC), the G allele predisposes to colon cancer with an odds ratio of 1.27 for GT and 1.47 for GG (Haiman et al., 2007). CRC being the second most common cause of cancer-related deaths, the contribution of the SNP variant is significant, due to its widespread prevalence among Caucasians (45-50%) and Africans (85%) (Haiman et al., 2007). The SNP falls within an ultraconserved DNA enhancer region (Jia et al., 2009), regulating the expression of the *MYC* gene present 335kb downstream of the SNP (Pomerantz et al., 2009; Sotelo et al., 2010). The cancer risk allele (G) binds preferentially to the beta-catenin-TCF4 transcription factor and activates the Wnt signaling pathway (Tuupanen et al., 2009). However, in contradiction, the SNP does not affect the

efficiency of interaction with the *MYC* gene (Wright et al., 2010). Moreover, the Genome Wide Association Studies (GWAS) that reported the rs6983267 association with cancer was unable to identify any linkage between rs6983267 genotype and *POU5F1P1* or *MYC* expression, amplification, and mutations in patient samples (Tomlinson et al., 2007; Zanke et al., 2007). In spite of the extensive research of several years, no known protein-coding gene and no mutation in the known protein-coding genes near this CAGR has been identified; and thus the mechanism by which this SNP promotes CRC still remains undefined.

Our lab identified a novel 1.7-kb long non-coding transcript at this SNP locus, called Colon Cancer Associated Transcript 2 *(CCAT2*) (Ling et al., 2013). *CCAT2* spans the highly conserved 8q24 region and is transcribed into two novel transcripts, *CCAT2-G* or *CCAT2-T*, based on the G/T SNP locus.

The seminal study on CCAT2 showed this lncRNA is specifically overexpressed in microsatellite stable (MSS) CRC samples (Ling et al., 2013). Patient data analysis showed that this lncRNA is specifically overexpressed in microsatellite stable (MSS) CRC samples. *In vitro* experiments suggested that *CCAT2* increased tumor formation and metastasis, and conferred resistance to chemotherapeutics drugs. A consistent positive correlation was observed between *CCAT2* and *MYC* expression both in cell line models and CRC patient samples. We also observed a higher β-catenin expression and higher Wnt activity in *CCAT2*-overexpressing clones. Finally, exogenous *CCAT2* expression induced chromosomal instability (CIN) phenotype in HCT116 cells, a microsatellite-instable (MSI) CRC cell line with CIN-negative phenotype. The high expression of *CCAT2* in

MSS CRC, the ability to enhance tumor formation and metastasis, induce CIN and activation of Wnt pathway suggests that *CCAT2* is an important oncogene in CRC predisposition and progression.

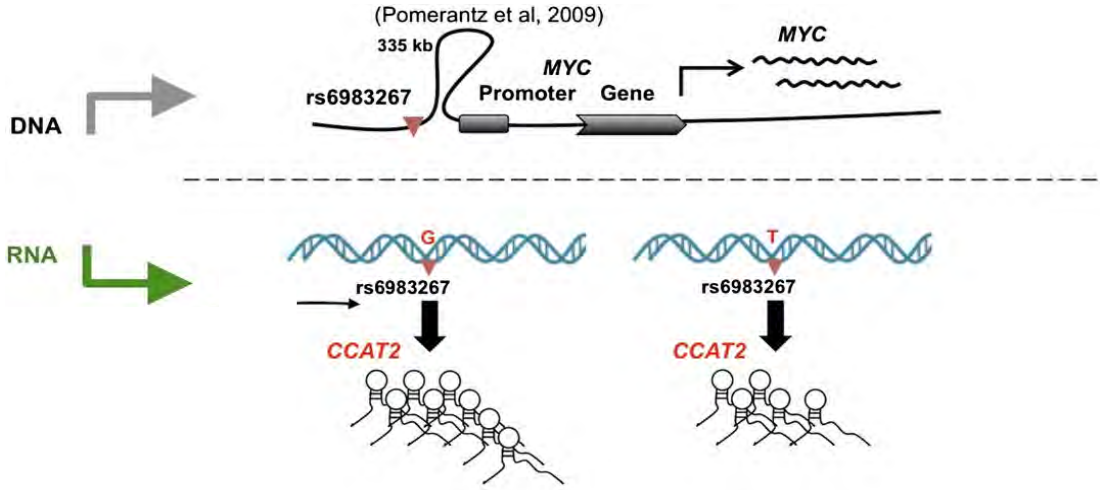


Figure 3: *CCAT2* is transcribed from the 8q24.21 genomic locus.

This figure is used with permission and originally published by Ling, H., Spizzo, R., Atlasi, Y., Nicoloso, M., Shimizu, M., Redis, R.S., Nishida, N., Gafa, R., Song, J., Guo, Z.*, et al.* (2013) in CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res *23*, 1446-1461 (Ling et al., 2013).

CCAT2 lncRNA in cancer

CCAT2 is also overexpressed in breast cancer (Redis et al., 2013), non-small cell lung cancer (Qiu et al., 2014), esophageal cancer (Wang et al., 2015b; Zhang et al., 2015b), and gastric cancer (Wang et al., 2015a). Redis et al reported that CCAT2 could serve as a predictive biomarker in lymph node positive (LNP) breast cancer. For this subgroup, high levels of *CCAT2* suggest the patients will not benefit from CMF (cyclophosphamide, methotrexate and 5-fluorouracil) adjuvant chemotherapy, with shorter metastasis-free survival and shorted overall survival. Additionally, CCAT2 up-regulated cell migration and downregulated chemosensitivity to 5'FU in breast cancer cell lines. These data suggest that CCAT2 plays an oncogenic role in breast cancer patients.

The study by Qiu et al reported that CCAT2 was significantly over-expressed in NSCLC tissues compared with paired adjacent normal tissues. Over-expression of CCAT2 was significantly associated with lung adenocarcinoma, but not squamous cell cancer. In vitro silencing of CCAT2 by siRNA led to inhibition of proliferation and invasion in NSCLC cell lines. Additionally, CCAT2 could also predict lymph metastasis in lung cancer patients when combined with carcinoembryonic antigen test. These findings additionally confirm oncogenic role of CCAT2 in lung adenocarcinoma.

The level of CCAT2 was positively correlated with TNM stages and positive lymph nodes metastasis (LNM) in esophageal squamous cell carcinoma (ESCC) (Zhang et al., 2015b). CCAT2 expression and MYC amplification were also

significantly associated with TNM stages and LNM. High CCAT2 expression and MYC amplification were significantly associated with poorer overall survival in ESCC patients. Additionally, (Wang et al., 2015b) showed that CCAT2 expression level was significantly associated with smoking status.

(Wang et al., 2015a) reported that CCAT2 expression levels were significantly higher in gastric cancer tissues than those in adjacent non-tumor tissues. It also positively correlated with higher incidence of lymph node metastasis and metastasis. Moreover, patients with high CCAT2 had shorter overall survival and progression-free survival compared with the low CCAT2 group. Finally, CCAT2 expression was identified as an independent poor prognostic factor for gastric cancer patients, valiating its oncogenic role in gastric cancer.

***CCAT2 is overexpressed in CIN positive CRC samples*.**

To detect the ncRNA gene transcribed in the rs6983267 conserved region, we designed primers spanning the SNP and adjacent 50 bp in both orientations **(Figure 4A)** and found a transcript that was expressed in human ovary, testis and brain tissues (Ling et al., 2013). RACE cloning from colon cancer cDNAs identified a ~340 nt transcript (CCAT2). Northern blot analysis confirmed the existence of *CCAT2* transcript in sense orientation (data not shown). MSS CRC tumors with CIN phenotype exhibited a significantly higher expression of *CCAT2* than MSI tumors or normal colon tissues that lacked CIN **(Figure 4B)**. These findings suggested that *CCAT2* might be a novel long ncRNA with potential oncogenic effect**.**

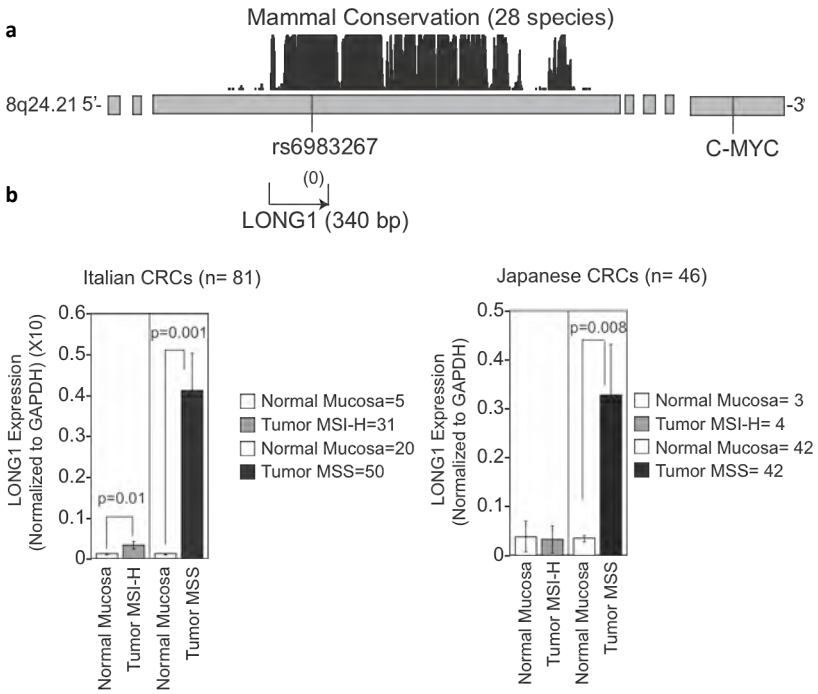


Figure 4: Transcription of *CCAT2* in CRC.

A) Genomic location of CCAT2 spanning the rs6983267 conserved region;

B) *CCAT2* expression in CRC patients.

This figure is used with permission and originally published by Ling, H., Spizzo, R., Atlasi, Y., Nicoloso, M., Shimizu, M., Redis, R.S., Nishida, N., Gafa, R., Song, J., Guo, Z.*, et al.* (2013) in CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res *23*, 1446-1461 (Ling et al., 2013).

CCAT2 increases tumor formation and metastasis.

To check if *CCAT2* had an oncogenic role, we cloned *CCAT2* cDNA in a retroviral expression vector and transfected it into the HCT116 cells, which expressed low endogenous levels of *CCAT2*. Under low adherence conditions, *CCAT2*-transduced cells showed increased proliferation and higher colony-forming ability (data not shown). Consistently, subcutaneous transplantation of *CCAT2*-overexpressing cells into Swiss nu-nu/Ncr nude mice resulted in more tumors of larger volumes **(Figure 5A).**

Next, we studied the involvement of *CCAT2* in promoting cancer metastasis. *In vitro* migration assay showed a 2-fold increase in the migration of *CCAT2*-transduced cells **(Figure 5B).** Injection of *CCAT2*-overexpressing cells into the spleen of Swiss nu-nu/Ncr nude mice resulted in increased liver metastasis rate (6 of 9 mice) as well as more metastatic tumors than that found in control group mice (1 of 5 mice) **(Figure 5C).**. Thus, *in vitro* and *in vivo* studies confirmed the tumor formation and metastatic potential of *CCAT2.*

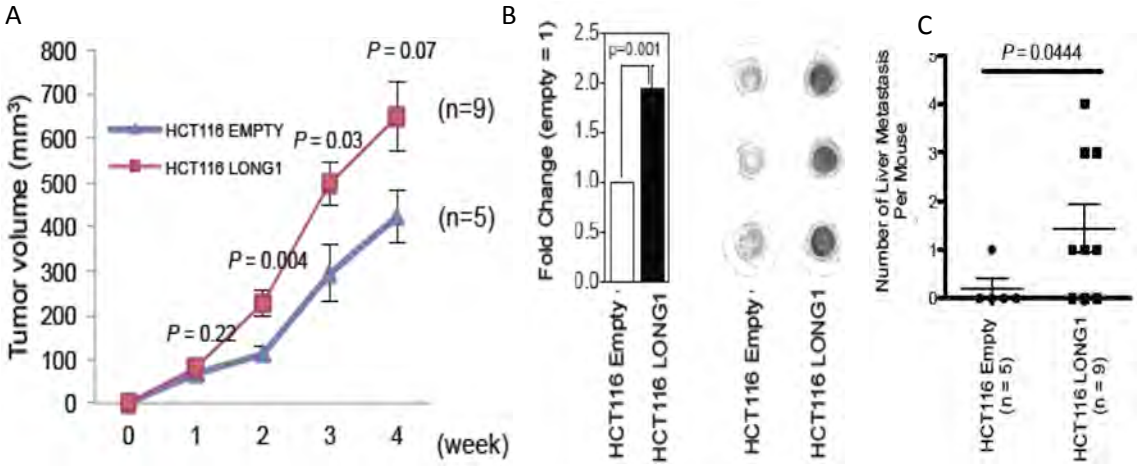


Figure 5: *CCAT2* increases tumor formation and metastasis.

A) CCAT2-overexpressing HCT116 clones increase subcutaneous tumor formation in a mouse xenograft model compared to empty HCT116 controls;

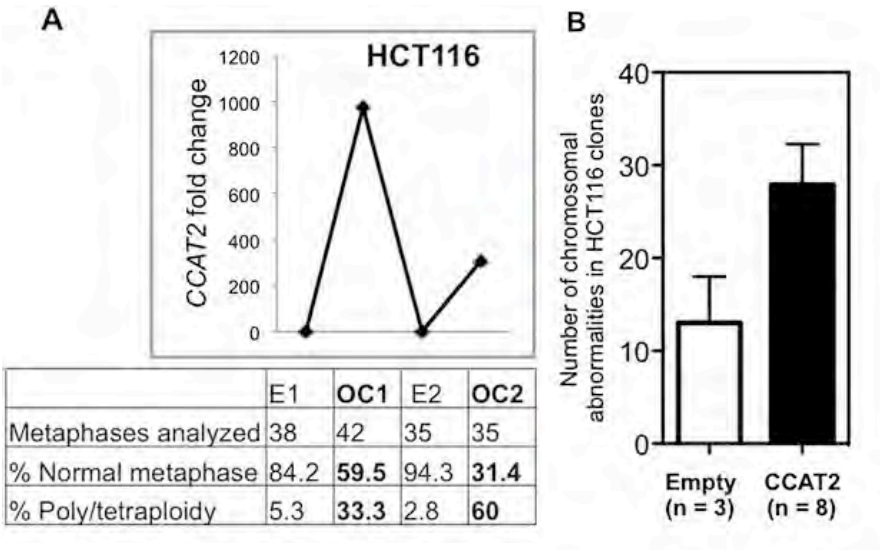
B) HCT116 cells transduced with CCAT2 showed higher migratory ability by migration chamber assay; and

C) CCAT2 promotes tumor metastasis in mice after intrasplenic injection of CCAT2-overexpressing HCT116 clones.

This figure is used with permission and originally published by Ling, H., Spizzo, R., Atlasi, Y., Nicoloso, M., Shimizu, M., Redis, R.S., Nishida, N., Gafa, R., Song, J., Guo, Z.*, et al.* (2013) in CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res *23*, 1446-1461 (Ling et al., 2013).

CCAT2 induces chromosomal instability.

MSS CRCs are characterized by their widespread CIN phenotype. Upregulation of *CCAT2* specifically in MSS CRC implied a possible involvement of *CCAT2* in CIN initiation and/or maintenance. HCT116 is an MSI CRC cell line with near-diploid karyotype (Barber 2008). We generated stable clones of HCT116 cells with higher (and #8) or basal (empty and #3) *CCAT2* expression. Karyotyping analysis revealed that the percentage of cells with normal metaphase was markedly lower in clones with high *CCAT2* expression (59.5% in #2 and 31.4% in #8, respectively) than in clones with basal *CCAT2* expression (84.2% in empty and 94.3% in #3, respectively) **(Figure 6A)**. In the same analysis, *CCAT2* overexpression dramatically increased the percentage of cells with polyploidy from 5.3% (empty) and 2.8% (#3) to 33.3% and 60.0% (#8), respectively **(Figure 6A)**. *CCAT2* also contributed to increased chromosomal exchange and greater DNA content. Consistently, we detected a near-tetraploid status in clone #8 by spectral karyotype analysis (data not shown). These findings were further validated by a second batch of eight *CCAT2* clones that exhibited similar numerical and structural chromosomal aberrations **(Figure 6B)**.



	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**

Figure 6: *CCAT2* induces chromosomal instability.

A) Increased chromosomal abnormalities in *CCAT2*-overexpressing clones, as identified by genomic instability analysis;

B) *CCAT2* induces chromosomal instability in a second set of *CCAT2*-overexpressing clones.

This figure is used with permission and originally published by Ling, H., Spizzo, R., Atlasi, Y., Nicoloso, M., Shimizu, M., Redis, R.S., Nishida, N., Gafa, R., Song, J., Guo, Z*., et al.* (2013) in CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res *23*, 1446-1461 (Ling et al., 2013).

CCAT2 regulates MYC and β-catenin expression.

Since *CCAT2* is transcribed from the region that acts as an enhancer to *MYC* gene, we analyzed the levels of *MYC* in relation to *CCAT2*. We found higher expression levels of *MYC* in cells that overexpress *CCAT2* (3.4- and 2.5-fold increase in #2 and #8 clones, respectively) **(Figure 7A)**. To further understand the regulation of *MYC* by *CCAT2*, we analyzed if other targets of *MYC* were also regulated by *CCAT2*. We observed upregulation of β-catenin (data not shown) and miRNA-17/92 cluster genes in these clones (data not shown). Conversely, down-regulation of *CCAT2* using shRNAs decreased *MYC* expression in stably transfected COLO320 cells in a *CCAT2*-dependent fashion **(Figure 7B)**. Thus, *CCAT2* regulates the transcription of *MYC* at RNA as well as protein levels. Additionally, we detected that the expression of *CCAT2* was dependent on TCF4 (transcription factor 4) and Wnt pathway. However differential G vs T allele behavior was observed, with higher TCF4-induced stimulation of *CCAT2* and hence higher transcription of *MYC* in high risk G allele (data not shown).

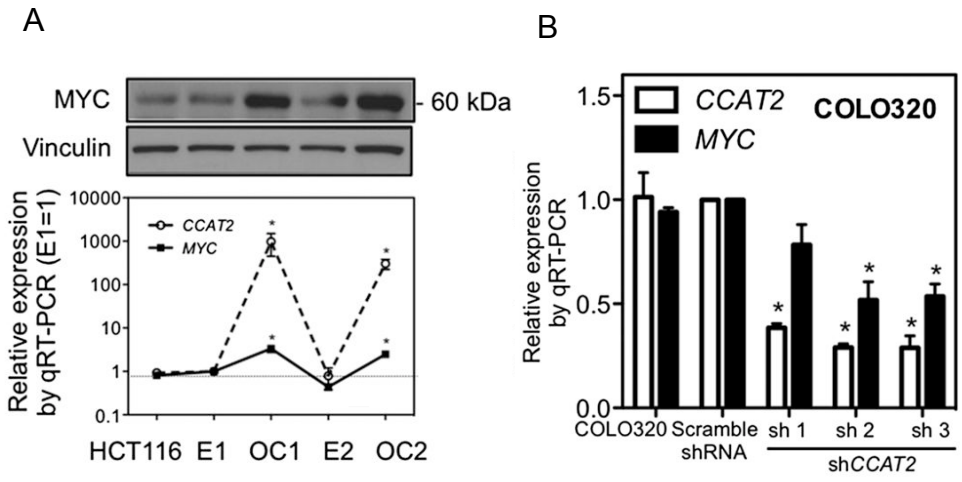


Figure 7: CCAT2 regulates MYC and β-catenin.

A, B) Correlation of *CCAT2* and *MYC* expression (A) and β-catenin (B) in CCAT2-overexpressing HCT116 clones. CCAT2 induces the expression of MYC;

C) Downregulation of *CCAT2* using shRNAs reduced *MYC* mRNA expression in COLO320 cells.

This figure is used with permission and originally published by Ling, H., Spizzo, R., Atlasi, Y., Nicoloso, M., Shimizu, M., Redis, R.S., Nishida, N., Gafa, R., Song, J., Guo, Z.*, et al.* (2013) in CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res *23*, 1446-1461 (Ling et al., 2013).

Hypothesis and Aims of the Study

These data indicates that *CCAT2* functions as an oncogene by regulating *MYC* expression and inducing chromosomal instability, thus predisposing a cell to cancer phenotype. However, several questions regarding the exact molecular mechanism of *CCAT2*-induced phenotype still remain, because it is difficult to conclusively establish the genetic events important for tumor initiation in the artificial milieu of the in vitro systems. A causal role of *CCAT2* in initiation of tumorigenesis and the importance of G/T SNP in *CCAT2*-induced phenotype still remains to be resolved. Characterizing the role in *CCAT2* in cancer initiation will help develop CCAT2 as a diagnostic and prognostic marker. The purpose of our study was to elucidate the role of *CCAT2* and its specific alleles in regulation of cellular processes that drive spontaneous tumorigenesis.

To do this, we established a genetically engineered mouse model overexpressing CCAT2 in all the body tissues at clinically relevant levels. This experimental strategy was designed to help us identify the tissue/organ most sensitive to CCAT2 expression. During the initial study, we observed that *in vivo CCAT2* overexpression induces pancytopenia, BM insufficiency and signs of de novo myelodysplastic syndrome (MDS).

We hypothesized that constitutive overexpression of CCAT2 induces chromosomal instability in HSCs and inhibits maturation of immature progenitor cells by regulating proteins important in MDS development.

To support our hypothesis, we decided to pursue the following tasks:

Aim 1: To study the oncogenic potential of *CCAT2* in initiation and progression of MDS

Aim 2: To delineate the underlying mechanism of MDS initiation by *CCAT2*

Aim 3: To detect and characterize *CCAT2* expression in BMCs of MDS patients

CHAPTER II: Methods

Cell lines and culture

HEK293 and SET2 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 Center using STR DNA fingerprinting. Both cells were maintained and cultured according to ATCC's instructions. All of the cell lines are routinely screened for Mycoplasma species (Mycoalert Mycoplasma Detection Kit, Lonza). All in vitro and in vivo experiments were conducted when cells were 70% to 80% confluent.

CCAT2 transgenic mice

CCAT2 (G/T) transgenic mice were generated using random integration approach. A 1.7kb human cDNA of CCAT2 expressing either the G or T allele was cloned into a vector backbone containing the CAG promoter, along with eGFP reporter gene followed by IRES site. Pronuclear injection of the entire 4.5kb linearized insert and generation of founder mice was performed by the MDACC Genetically Engineered Mouse Facility. The founders were mated with C57BL/6 mice. Pups were screened for presence of the transgene by both PCR and Southern blot analysis on tail-extracted DNA according to standard protocols. For PCR screening, three different primer pairs were used to detect different products in the inserted transgene. Pups showing positive detection for all three products were identified as founders. Southern blot analysis was performed to confirm insertion of CCAT2 cassette in the mice genome, and to determine to copy number of the insertion. The expression of *CCAT2* transcript was tested by Real time quantitative PCR with human specific set

of primers previously described in (Ling et al., 2013). All the protocols and experiments were conducted according to the guidelines of the MDACC Institutional animal care and use committee.

Hematological measurements and peripheral blood morphology

Blood was collected from tail vein for hematological measurements. Peripheral blood cell counts were performed on an Avida hematology analyzer. For morphological assessment, peripheral blood smears were stained with Hema III stain (Sigma-Aldrich) for 10 min followed by rinsing in dH_2O for 3 min. Images were taken using a 60× objective on a Leica microscope outfitted with a camera.

Histology and Immunohistochemistry

Mice were necropsied, and femurs, sternum, spleens, and livers were fixed in 10% buffered formalin, included in paraffin, and then cut into 4 um sections. Sections were stained with H&E according to standard protocols. For the dewaxing step, sections were heated for 1hour at 55°C, followed by rehydration steps through a graded ethanol series and distilled water, immersed in PBS, and then treated with 0.1% trypsin solution in Tris buffer for 30 min at 37°C. Endogenous peroxidase was blocked with 10% normal serum. Ki67, CD3, CD4, and CD20 were used as primary anti-mouse antibodies (BD PharMingen). Secondary antibodies and diaminobenzidine were added according to the manufacturer's instruction.

Cytochemical Staining

Cytospins of bone marrow cells were prepared following standard procedure, and

fixed in methanol for 7 min. Cytochemical assays for myeloperoxidase, (Sigma), nonspecific esterase (**α**-naphthyl butyrate, Sigma), *periodic acid-Schiff* [*PAS*], and reticulin were performed according to manufacturer's instructions.

Mouse hematopoietic progenitor assay

Methylcellulose colony formation assay (CFU-C) were performed using Methocult GF M3434 (Stem Cell Technologies). Bone marrow and spleen cells were obtained from CCAT2(G/T) and WT mice and red blood cells were lysed using Red blood cell lysis solution (Sigma). 10,000 bone marrow cells or 100,000 spleen cells were plated in duplicate. Colonies were identified and counted following 7 days culture at 37C with 5% CO_2.

RNA extraction, reverse transcription and real time qPCR

Total RNA was extracted from patient samples, mice tissues or cultured cells using Trizol and following manufacturer's protocol (Sigma). RNA concentrations were measured with spectrophotometer NanoDrop ND-1000 instrument (NanoDrop Tachnologies, Termo Scientific, Wilmington, DE, USA). Reverse transcription was performed using random hexamers with SuperScript III Reverse Polymerase according to manufacturer's protocol (Invitrogen). Quantitative RT-PCR analysis was performed with SYBR Green using specific primes (Appendix). B2M, HRPT or PGK1 were used as internal control. MicroRNA expression was tested using TaqMan microRNA assay (Applied Biosystems). The cDNA was synthesized using TaqMan Reverse Transcription Reagents kit (Applied Biosystems) and then used for quantitative RT-PCR analysis with TaqMan probes and SsoFast Supermix (Bio-rad).

Primers and probes for each miRNA were purchased from TaqMan. U6 snRNA was used as an internal control. Experiments were performed in triplicates. Relative expression levels were calculated using the $2^{-\Delta\Delta Ct}$ method.

In situ hybridization

The FFPE tissue sections were first digested with 5 µg/mL proteinase K for 20 minutes at RT, and were then loaded onto Ventana Discovery Ultra system (Ventana Medical Systems, Inc, Tucson, AZ) for in situ hybridization. The tissue slides were incubated with double-DIG labeled probe for CCAT2 or control U6 (Exiqon) for 2 hrs at 55° C. The digoxigenins were detected with a polyclonal anti-DIG antibody and alkaline phosphatase conjugated second antibody (Ventana) using NBT-BCIP as the substrate. The signal intensities of *CCAT2* and *RNU6-6P* expression were quantified by using the intensity measurement tools of the Image-Pro Plus software package (Media Cybernetics).

Genomic instability analysis

Femur bone marrow CCAT2(G/T) or WT mice was flushed and collected with RPMI medium 1640 + 20% FBS. BMCs were cultured overnight in complete RPMI medium using standard methods. Cells were then 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. Images were captured using Cytovision Imaging system (Applied Imaging) attached to a Nikon Eclipse 600 microscope. Twenty to thirty karyotypes were prepared from each sample and described using the standard chromosome nomenclature for mice.

Spectral karyotyping

Spectral karyotyping was performed according to the manufacturer's protocol using Mouse Paint probes (Applied Spectral Imaging, ASI). Images were captured using a Nikon 600 microscope equipped with Spectral karyotyping software from ASI.

Northern blot

Total RNA was isolated using using Trizol and following manufacturer's protocol (Sigma). RNA samples (30 µg each) were run on 15% acrylamide denaturing (urea) Criterion precast gels (Bio-Rad), and then transferred onto Hybond-*n* + membrane (Amersham Pharmacia Biotech). The hybridization with [α-^{32}P]ATP was performed at 42°C in 7% SDS/0.2M Na_2PO_4 (pH 7.0) overnight. Membranes were washed at 42°C, twice with 2× SSPE [standard saline phosphate/EDTA (0.18 M NaCl/10 mM phosphate, pH 7.4/1 mM EDTA)]/0.1% SDS and twice with 0.5× SSPE/0.1% SDS. CCAT2 probes used were as published previously in Hui et al. As loading control we used 5S rRNA stained with ethidium bromide.

Southern blot

Mice were screened for the presence of the transgene by Southern blot analysis on tail DNAs digested with *PshA1*. Different concentrations of the 4.5kb sequence used for pronuclear injection were used as the copy number positive controls. Blots were hybridized with the same *CCAT2* DNA fragment reported in Ling et al. Genotyping was performed on tail DNAs by PCR.

Western blot

Proteins were collected from cultured cells and lysed with RIPA buffer (SIGMA) freshly supplemented with a complete protease inhibitor cocktail (SIGMA). The Bradford assay was used to measure protein concentrations. Proteins were separated on polyacrylamide gel (Bio-rad) electrophoresis and transferred to a 0.2 μm nitrocellulose membrane (Bio-rad). Vinculin was used as loading control and quantification of protein expression was done with ImageJ. All antibodies were purchased from Cell Signaling unless mentioned otherwise. The following antibodies were used for analysis: EZH2, H3K27Me3, Histone H3, GFP, and p27.

***In vivo* detection of apoptosis via TUNEL assay**

Apoptotic cells in bone marrow tissue were detected by terminal deoxynucleotidyl transferase-mediated dUTP nick-end labeling (TUNEL) staining using an apoptotic cell detection kit following the manufacturer's directions (Promega, Madison, WI, USA). Images of the sections were captured using a Nikon 600 microscope (Nikon, Tokyo, Japan). The apoptotic index was calculated by dividing the number of TUNEL-positive cells by the total number of cells in the field. Light microscopy was

used to count the number of TUNEL-positive cells on ten randomly selected fields for each section.

Alkaline Comet Assay

Freshly isolated bone marrow cells from CCAT2(G/T) and WT mice were subjected to alkaline comet assay according to the manufacturer's instructions (catalog # 4250-050-K; Trevigen). Briefly, cells were combined with low melting agarose onto CometSlides (Trevigen). After lysis, cells were subjected to electrophoresis and stained with SYBR green. Subsequently, cells were visualized using fluorescent microscopy (Carl Zeiss, Thornwood, NY). At least 200 comet images were analyzed for each time point using Comet Score software (version 1.5; TriTek Corp.). The number of tail-positive cells with small and large nuclei was manually counted by an examiner blinded to treatment group, and expressed as a percentage of all cells evaluated. Experiments were repeated in triplicate.

Flow Cytometry analysis

Single-cell suspensions were prepared from bone marrow (from femoral and tibial bones) by passing cells through pre-separation filters (Miltenyi). Cell numbers were subsequently counted. Cells were fixed and stained according to standard protocol for FACS sorting and analysis. Cells were acquired using LSR Fortessa (BD Bioscience) and analyzed using FlowJo software (Tree Star). Dead cells were excluded by staining them with Ghost Dye (Tonbo Biosciences). Cell doublets were excluded from all analyses. Lin− cells were enriched using a Mouse hematopoietic progenitor cell enrichment kit (Stem Cell Technologies). The antibodies (all from

eBioscience, Tonbo or Biolegend) conjugated to FITC, APC, APC-Cy7, PE, PE-Cy7, PercpCy5.5, EFLUOR450, or APC-Cy7 were used for the flow cytometry analysis.

RNA immunoprecipitation

We used the Magna RIP Kit (Millipore) according to the manufacturer's instructions. Cells were prepared in RIP lysis buffer, and the RNA–protein complexes were immunoprecipitated with magnetic beads using anti-EZH2 or normal mouse IgG (control) (Appendix). Co-purified RNA was extracted using phenol:chloroform:isoamyl alcohol and subjected to reverse transcription and real-time qPCR analysis. B2M was used as a non-target internal control.

GST-MS2 pull-down assay

HEK293 cells were transfected with 2 ug of pMS2-CCAT2 or pMS2, together with 1 ug of pMS2-GST, and total lysates were prepared after 48 hrs of transfection. The protein-RNA complexes were immunoprecipitated using GSH agarose beads (GE Healthcare). Co-purified proteins were analyzed by western blot analyses. GST was used as the internal transfection control.

MicroRNA microarray

After extraction of total RNA using Trizol (Sigma), the RNA samples from bone marrow cells (n=3 each) were checked for purify and quality via an Agilent Bioanalyzer before being submitted for human miRNA array (Sanger miRBase v15, 1,087 human miRNAs) analysis. The results of the analysis were used to determine which miRNAs had significant “fold” differences in expression between WT and

CCAT2(G/T) mice. The miRNAs with the most significant differences in expression levels were then chosen for the miRNA target analysis using multiple bioinformatics prediction tools, including TargetScan 6.2 (http://www.targetscan.org/), PicTar (http://pictar.mdcberlin.de/), miRanda (http://www.microrna.org/), and miRDB (http://mirdb.org/miRDB/).

Clinical samples

CD34+ cells from bone marrow of 80 myelodysplastic syndrome patients and 33 myeloproliferative patients were obtained from MD Anderson Cancer Center tissue bank. A second set of peripheral blood from 55 myelodysplastic syndrome patients and 8 normal patient samples were obtained from Romania (University of Medicine and Pharmacy Iuliu Hatieganu). All samples were collected according to the institutional policies and obtained following patient's informed consent. Tissue samples were obtained from fresh surgical specimens frozen in liquid nitrogen and stored at −80°C. Data were de-identified prior to any analyses using standard procedures.

Statistical analysis

For the MDS patient 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. We investigated the prognostic value of the clinical and biological variables with metastasis-free survival (MFS) as the end point 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. Unless specified otherwise, all data are presented as the mean values ± the standard error of the mean from at least 3 independent experiments. Two- sided *t* tests were used to test the relationships between the means of data sets, and *P* values indicate the probability of the means compared, being equal with $*P < 0.05$, $**P < 0.01$ and $***P < 0.001$. Student's *t* tests and analysis of variance were calculated with GraphPad software. Statistical analyses were performed in R (version 3.0.1) (http:///www.r-project.org/), and *P* values less than 0.05 were considered statistically significant.

CHAPTER III: Results

Generation of CCAT2(G/T) transgenic mice

In order to study the consequence of *CCAT2* overexpression *in vivo*, we generated transgenic mice for each *CCAT2* allele (G or T) using random integration approach in C57Bl6/N background. The human *CCAT2* cDNA was expressed under the control of CAG promoter, which gets expressed ubiquitously in all mice tissues, along with eGFP fluorescent tag using IRES (internal ribosome entry site) (**Figure 8A**). All CCAT2(G/T) pups were born at expected frequencies and without any abnormalities. Pups were screened for presence of the transgene by both PCR and Southern blot analysis on tail-extracted DNA. For PCR screening, three different primer pairs were used to detect different products in the inserted transgene. Pups showing positive detection for all three products were identified as founders (**Figure 8B**). Five positive founders were identified, G26-0-0, G28-0-0, G31-0-0 (founders with G allele) and T06-0-0, T12-0-0 (founders with T allele). We also performed southern blot analysis to confirm insertion of CCAT2 cassette in the mice genome and to determine to copy number of the insertion. Comparing the band intensity with the copy number standards, founder 26-0-0 had insertion between 5-10 copies, and all the other founders between 1-5 copies **(Figure 8C)**. Sanger sequencing confirmed integration of specific G or T allele in the mouse genome **(Figure 8D)**. Thus, the genomic DNA of the G mice had only G allele, while T mice only had T allele, validating the allele-specific CCAT2 transgenic model.

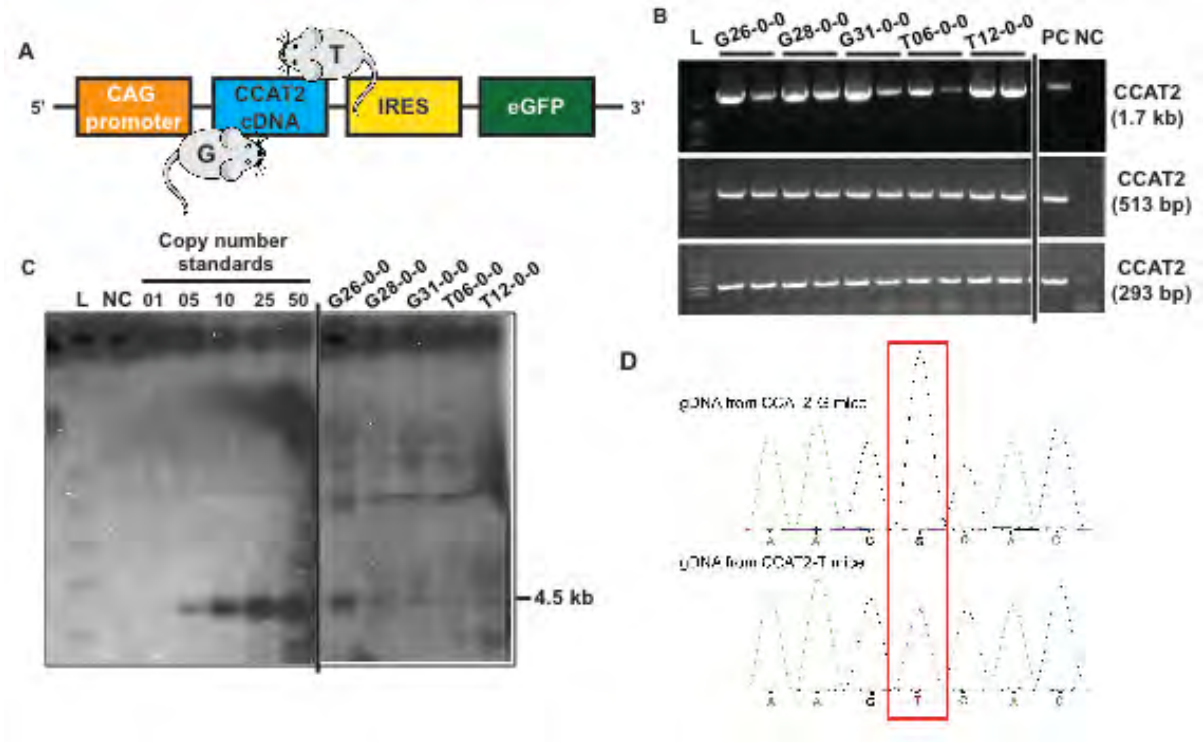


Figure 8: Generation of CCAT2(G/T) transgenic mice.

(A) The schematic of the CCAT2 vector construct that was integrated in the mice genome by random integration using pronuclear injection.

(B) PCR on tail genomic DNA from five positive founders detected using three different primer pairs.

(C) Southern blot analysis using tail genomic DNA from five positive founders to detect the integration copy number.

(D) Sanger sequencing analysis on tail genomic DNA of CCAT2-G and CCAT2-T mice.

The expression of *CCAT2* transcript was tested by quantitative PCR with human specific set of primers previously described in (Ling et al., 2013). We detected a ten- to twenty-fold increase in *CCAT2* expression in different tissues of the animal including colon compared to their non-transgenic siblings **(Figure 9A)** by qRT-PCR. Similar level of overexpression was also detected in the bone marrow cells o these mice **(Figure 9B, 9C)**. Overexpression of *CCAT2* was also confirmed by in situ hybridization in hematopoietic tissues, including bone marrow cells (BMCs) **(Figure 10A)** and spleen **(Figure 10B)**. This level of overexpression mimics what we previously reported for CCAT2 in colon cancer malignancy (Ling et al., 2013).

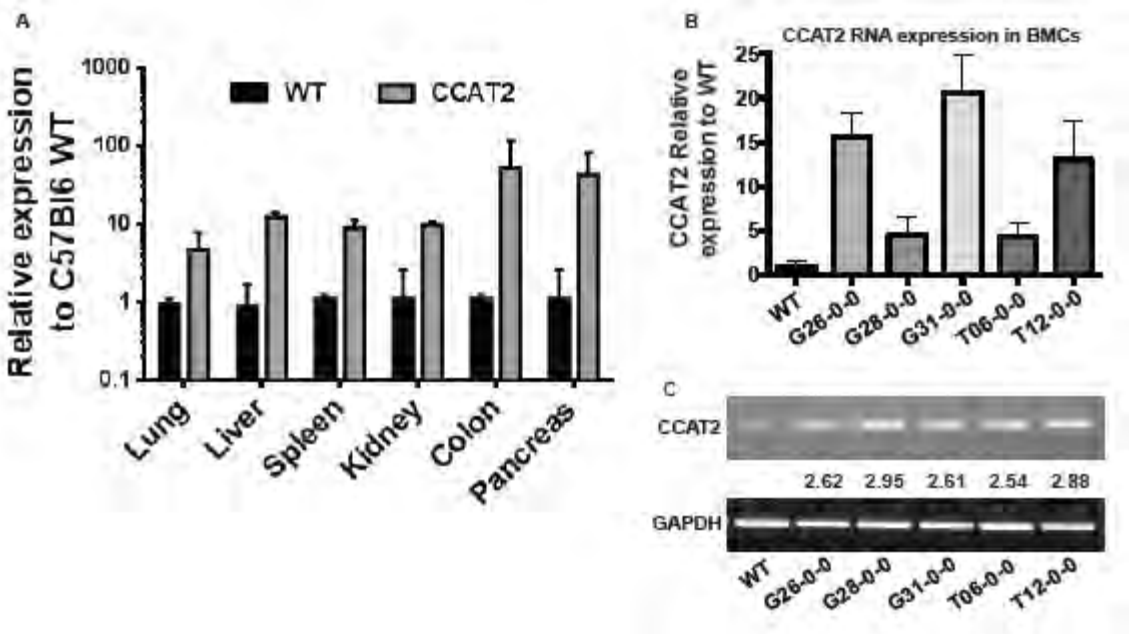


Figure 9: *CCAT2* transcript expression in *CCAT2* transgenic mice.

(A) *CCAT2* expression levels in different organs of CCAT2(G/T) mice in comparison to non-transgenic WT littermates by Real Time qPCR.

(B) *CCAT2* expression levels in bone marrow cells of CCAT2(G/T) mice in comparison to non-transgenic WT littermates by Real Time qPCR.

(C) *CCAT2* expression levels in bone marrow cells of CCAT2(G/T) in comparison to non-transgenic WT littermates by PCR on cDNA.

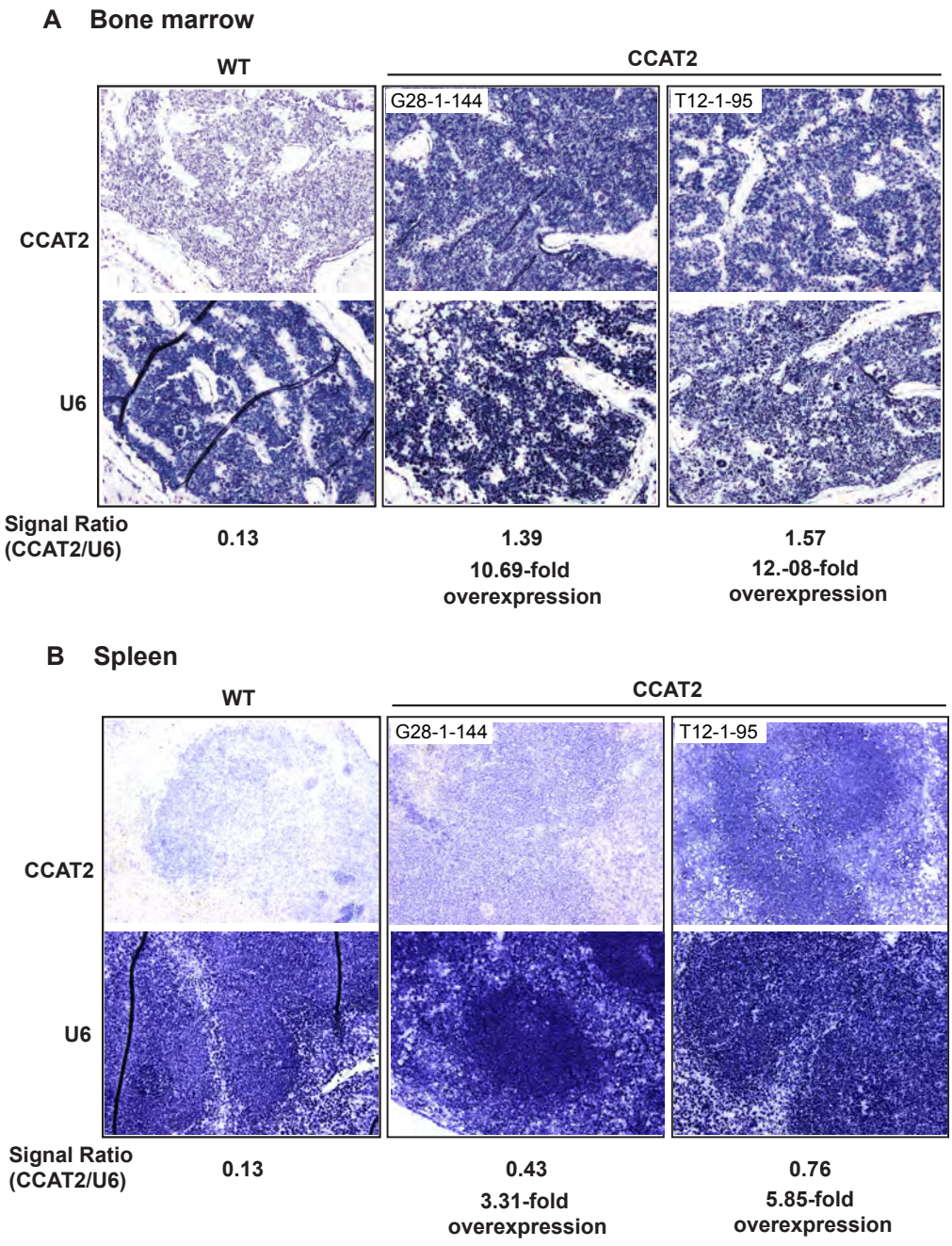


Figure 10: CCAT2 expression in bone marrow and spleen of CCAT2 transgenic mice.

(A, B) In-situ hybridization was performed using CCAT2-specific probes to detect expression of CCAT2 lncRNA in bone marrow (A) and spleen (B) of CCAT2(G/T) mice.

CCAT2(G/T) mice display bilineage cytopenias

We initially expected to see a colon phenotype by overexpressing *CCAT2*, since we previously reported that high *CCAT2* predisposes to colon cancer tumors. However, although *CCAT2* was expressed at high levels in the colon, no macroscopic (n = 54) or histological (n = 22) evidence for colon tumors or early stage polyps were observed in any adult mouse up to 12 months of age (data not shown). In contrast, within about 6-9 months of age, CCAT2(G/T) mice showed clinical signs of a hematological phenotype, predominantly widespread cytopenias. We monitored the hematological parameters of 6-9 months old WT (n = 20), CCAT2-G (n = 21) and CCAT2-T (n = 18) mice and observed that CCAT2(G/T) mice developed bilineage cytopenias. These mice exhibited significantly low WBC **(Figure 11A)** and lymphocyte counts (lymphocytopenia, **Figure 11B**); and minor anemia with lower red blood counts and hemoglobin levels (data not shown) compared to WT and non-CCAT2 expressing littermates, a phenomenon consistent with myelodysplasia and ineffective hematopoiesis. No significant difference was observed in other hematological parameters including levels of eosinophils, basophiles, monocytes, reticulocytes, hematocrit, mean platelet volume and mean corpuscular volume in these mice (data not shown).

Interestingly, about 40% of the CCAT2(G/T) mice exhibited a myeloproliferative neoplasm (MPN)-like phenotype compared to WT controls. The mixed MDS/MPN CCAT2(G/T) mice displayed significantly elevated platelet levels (**Figure 11C**) along with other myelodysplastic features (**Figure 11D, and E**). Both $CCAT2^{(G)}$ and $CCAT2^{(T)}$ mice were equally susceptible to both MDS- or MDS/MPN-

like features. These data illustrations that overexpression of *CCAT2* induces two distinct phenotypic manifestations in the BM of mice, suggesting a critical role of CAT2 in maintenance of homeostasis in hematopoietic stem cells.

We also identified a significant increase in presence of LUCs (large unstained cells) in the peripheral blood (PB) of CCAT2(G/T) mice (**Figure 11F)**. These are cells that usually cannot be distinguished according to their staining pattern, including large atypical lymphocytes and immature blasts. Increase of immature cells in circulation usually signifies compromised BM function.

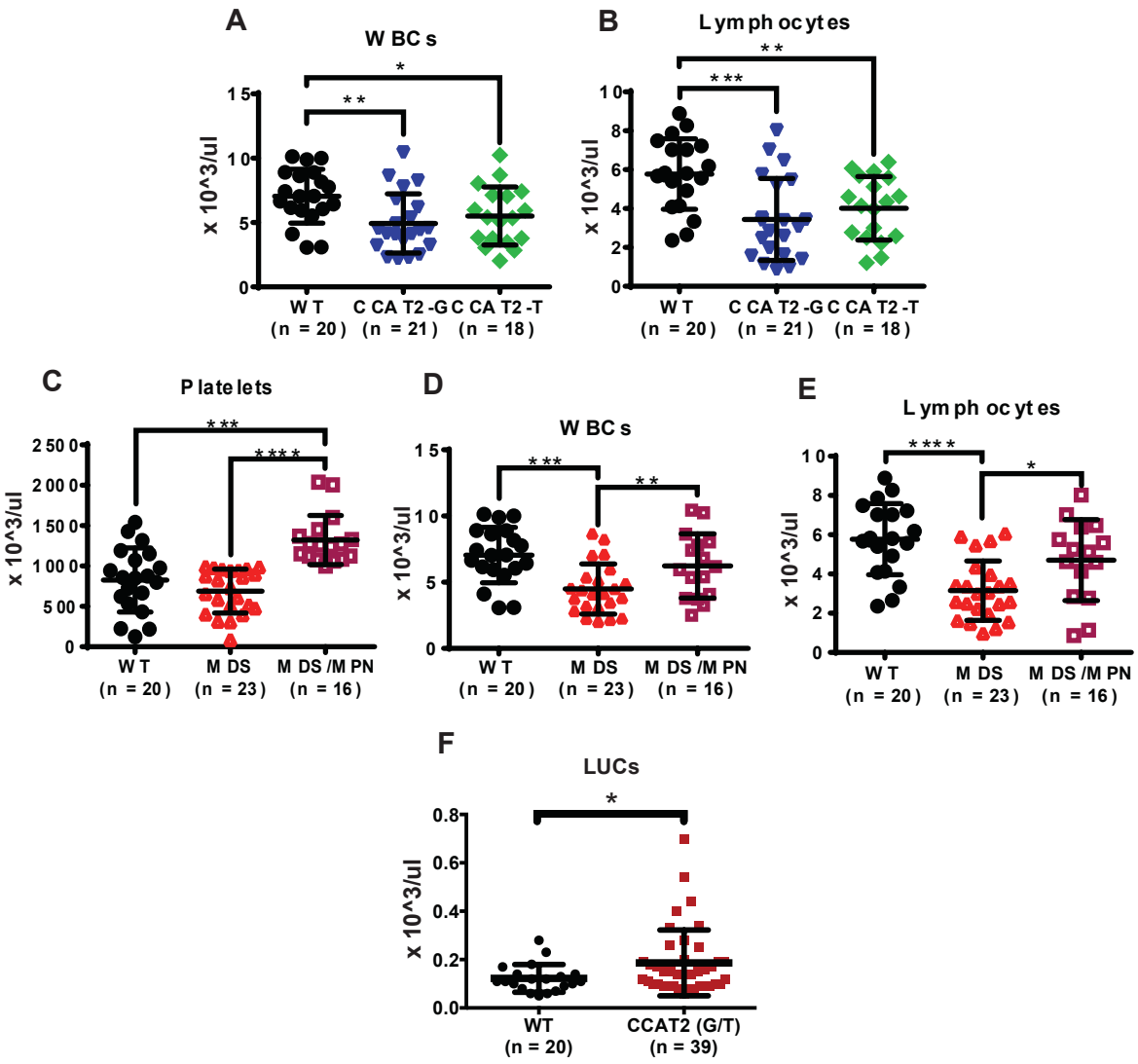


Figure 11: Peripheral complete blood counts (CBC) of CCAT2(G/T) and WT mice

(A-B) WBCs (A), and lymphocytes (B) in CCAT2-G, CCAT2-T and WT mice.

(D-G) Platelets (D), WBCs (F), and lymphocytes (E) in CCAT2(G/T) mice displaying MDS-like and MDS/MPN-like phenotype compared to non-transgenic WT littermates.

(F) Presence of large unstained cells (LUCs) in peripheral blood of CCAT2(G/T) mice.

The data are presented as ± STD. *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$.

Consistent with the complete blood counts, morphological analysis of PB smears by Hema III stain (modified May Giemsa stain) showed presence of several aberrant circulating blood cells (**Figure 12**). We observed poikilocytosis and anisocytosis (RBCs of abnormal shapes and variable size), including presence of oddly shaped RBCs such as dacrocytes, knizocytes, teardrops, schistocytes and echnocytes. An increased numbers of immature polychromatophilic RBCs, reticulocytes and Howell-Jolly bodies were also noted in circulating PB. In the myeloid lineage, we observed hypogranularity of neutrophils with abnormal lobation. These include presence of pseudo-Pelger-Huet anomaly, hypersegmented neutrophils and ringed neutrophils. In the platelet lineage, we noted megaplatelets with clustered, atypical granulation in the cytoplasm. Clumped platelets of variable sizes were also observed in PB smears. Levels of circulating platelets were significantly increased in PB of MPN-like mice.

Additionally, we also observed presence of immature blasts and unidentified mononuclear cells in PB. The presence of these immature blasts correlated with the LUCs detected in PB CBC analysis as shown in **Figure 11H**. Presence of these immature cells in circulation could potentially indicate their defective maturation in the BMCs.

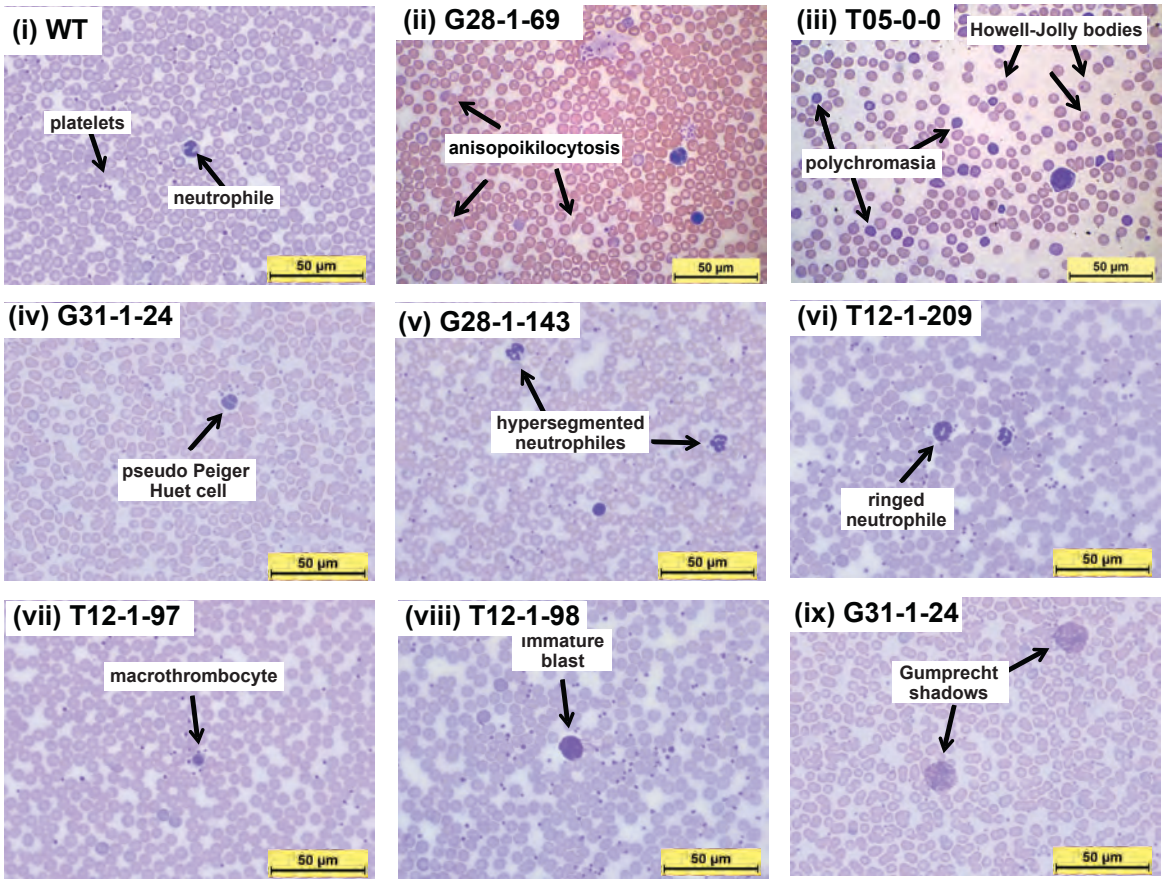


Figure 12: Peripheral blood smears of WT (i) and CCAT2(G/T) (ii-ix) mice stained with Hema III stain (modified May Giemsa stain).

Atypical cells belonging to each hematological lineage were observed in the PB of CCAT2(G/T) mice. The aberrant cells are indicated using arrows and labeled.

CCAT2(G/T) mice have hyperplasic bone marrows

To identify if the bone marrow of these animals showed signs of dysplasia, we performed histologic examinations of H&E stained BM sections of CCAT2(G/T) (n = 18) and WT (n = 9) animals at 9 months of age. In contrast with WT-mice, all CCAT2-G and CCAT2-T mice had increased cellularity of the bone marrow. However, the hematopoietic hypercellularity was higher in CCAT2-G-mice (average score 3.3) in comparison with CCAT2-T-mice (average score 2.6). The hematopoietic hypercellularity of CCAT2-G and CCAT2-T-mice was mostly due to the increase number of myeloid (granulopoietic) cells, average score of 2.8 in G-mice and 2.0 in T-mice (**Figure 13A**). This change is consistent with myeloid/granulocytic hyperplasia in these mice. Increased erythroid cells were also observed in 2/9 G-mice and 4/9 T-mice (**Figure 13B, C**). In addition to granulocytes, there was increased number of megakaryocytes (megakaryocytic hyperplasia) in 6/9 G-mice and 5/9 T-mice with increased accumulation of atypical micromegakaryocytes was also observed in these mice (**Figure 13D**). All CCAT2(G/T) mice we analyzed showed bone marrow hyperplasia and cytopenias of at-least one lineage compared to their non-transgenic WT littermates. Both MDS- and MPN-like mice showed comparable bone marrow hyperplasia and multi-lineage defects.

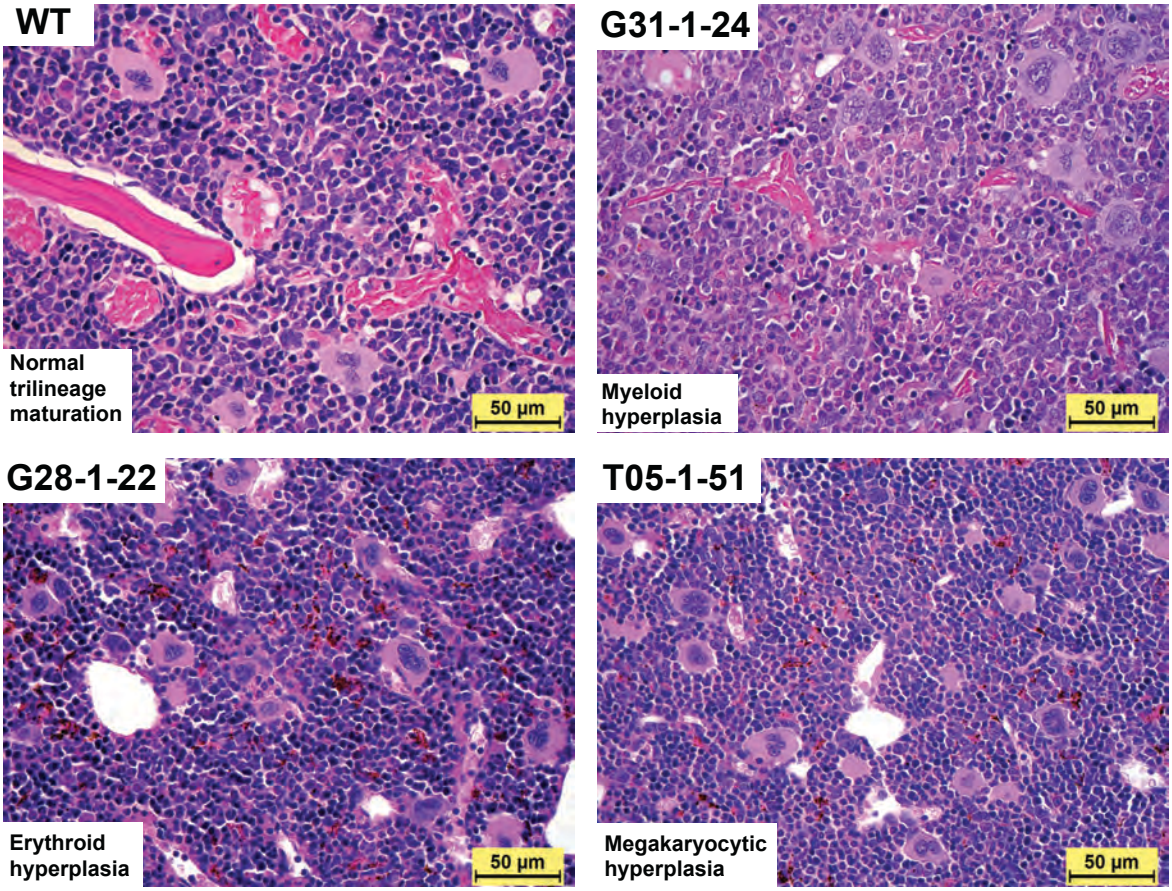


Figure 13: H&E staining of BM sections of CCAT2(G/T) and WT mice.

Femurs from WT and CCAT2(G/T) mice were fixed in 10% Formalin and paraffin embedded. H&E staining was performed and these slides were then analyzed for complete histopathology.

CCAT2(G/T) mice display splenomegaly and hepatomegaly

Along with BM hyperplasia, several alterations in the hematopoietic organs were noted in these mice, including splenomegaly and hepatomegaly. The spleen was frequently enlarged ($p < 0.01$), sometimes up to five-fold that normal weight **(Figure 14A)**. The liver was also enlarged markedly in several mice ($p < 0.01$, **Figure 14B**). Histological examination revealed noticeable extramedullary hematopoiesis (EMH) and infiltration of hematopoietic precursor cells, diminishing the white pulp and other germinal centers (**Figure 14C**). Additionally, minor EMH was also observed in organs not commonly associated with EMH, such as liver (**Figure 14D**). Extensive EMH accompanied by defects in central hematopoiesis and bone marrow insufficiency suggests that EMH might be partly compensatory. Since bone marrow insufficiency with ineffective hematopoiesis, and enlarged EMH organs are a hallmark of MDS/MPN, we concluded that CCAT2(G/T) mice displayed MDS or mixed MDS/MPN-like disease.

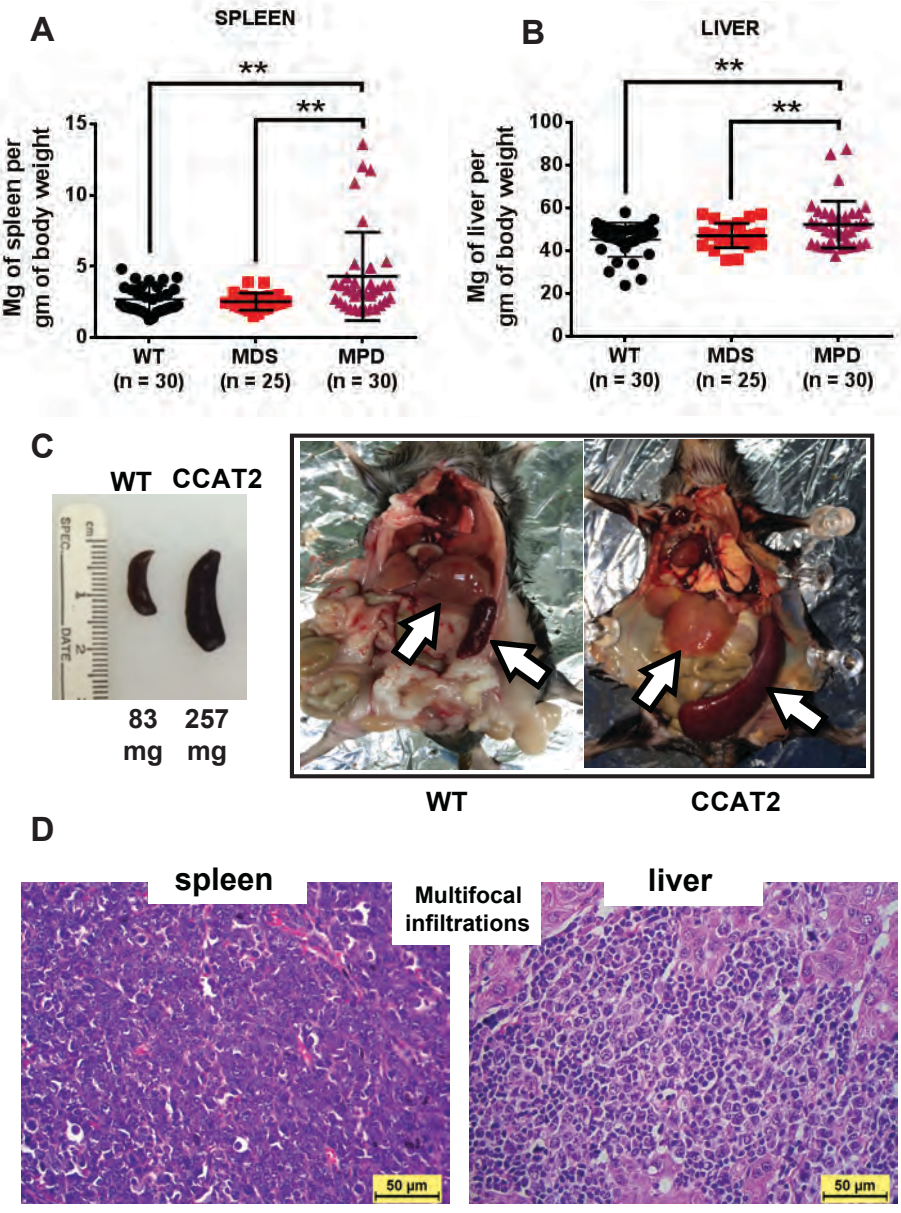


Figure 14: CCAT2(G/T) mice display splenomegaly and hepatomegaly.

(A, B) Normalized Spleen (A) and Liver (B) weights of CCAT2(G/T) mice displaying MDS-like or MDS/MPN-like phenotype compared to WT littermates.

(C) Representative images enlarged spleen and liver in CCAT2(G/T mice with mixed MDS/MPN-like phenotype.

(D) H&E staining of spleen and liver sections of CCAT2(G/T) and WT mice.

The data are presented as ± STD. *P < 0.05, **P < 0.01 and ***P < 0.001.

CCAT2 bone marrow cells display dysplastic features

We next obtained BM smears from WT and CCAT2(G/T) mice and stained them with Hema III (modified May Giemsa stain). In CCAT2(G/T) mice, bone marrow aspirates showed multilineage proliferative and dysplastic changes **(Figure 15)**. Presence of immature blast cells were observed in several CCAT2(G/T) mice. However, the proportion of these immature cells were < 5% in BM, indicating myelodysplastic changes but no significant progression to leukemias.

Findings included increased presence of hypersegmented neutrophils, and dysplastic erythroblast precursors. A distinct hyperplasia of the bone marrow, specifically myeloid hyperplasia was noted. Additionally, we observed remarkable hyperplasia and aggregates of dysplastic megakaryocytes. Micromegakaryocytes with decreased ploidy and reduced chromatin, as well as mono- and bi-lobated megakaryocytes were noticed. In spite of thrombocytopenia in some mice, megakaryocytic hyperplasia was evident in the marrow aspirates, suggesting aberrant platelet maturation.

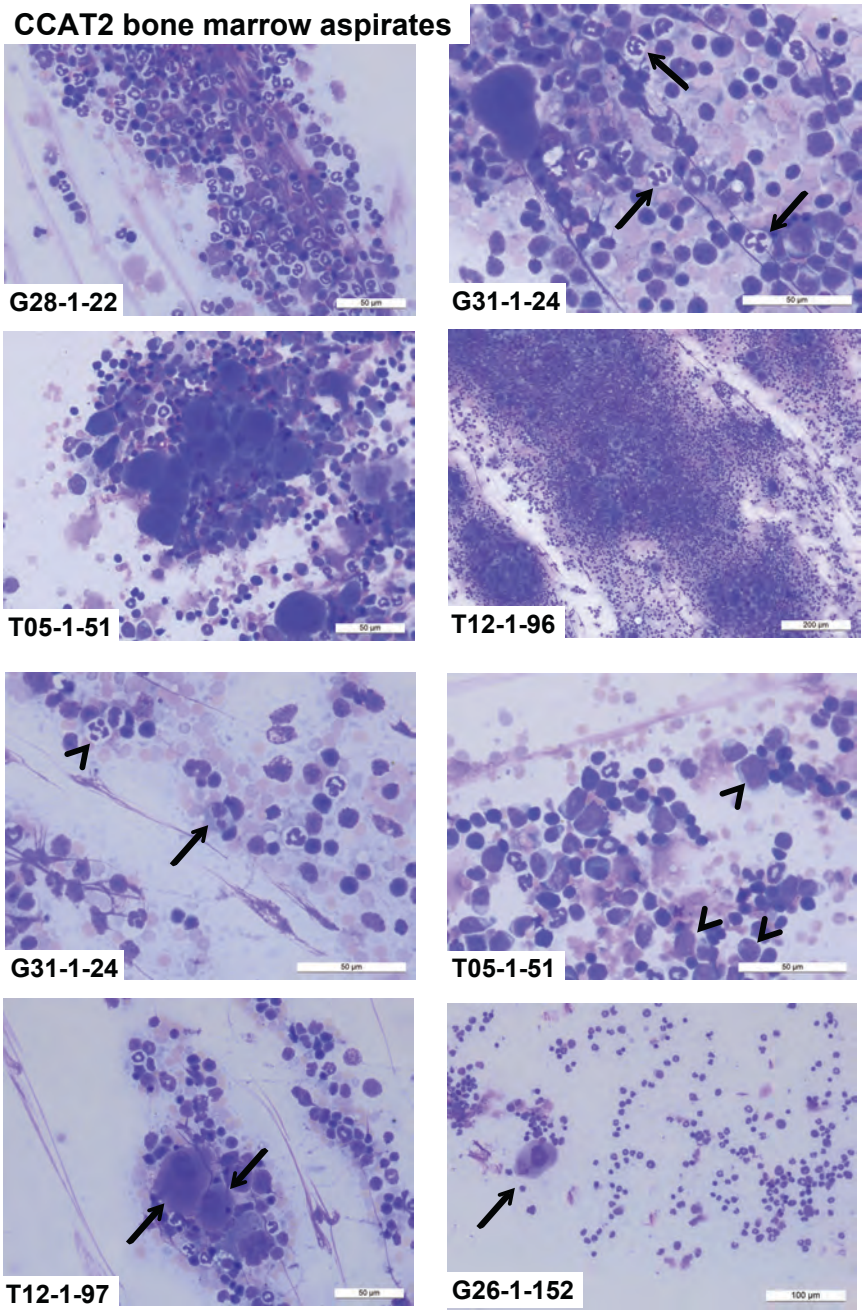


Figure 15: BM smears from WT and CCAT2(G/T) mice stained with Hema III (modified May Giemsa stain).

We then stained BM aspirates for characteristic indicators of MDS. We observed presence of ring sideroblasts in CCAT2(G/T) mice, a characteristic of MDS, while none were seen in the WT mice **(Figure 16A)**. However, the prevalence of these ring sideroblasts was <5% in the BM of CCAT2(G/T) mice. Additionally, staining with iron showed reduced iron deposition in CCAT2(G/T) mice compared to WT littermates **(Figure 16B)**, indicating minor anemia as observed from peripheral blood counts.

The blasts were PAS and myeloderoxidase negative **(Figure 16C)**. Further, reticulin and trichome staining of the bone marrow biopsies revealed no collagen deposition or myelofibrosis in these mice **(Figure 16D)**. Collectively, these data show that CCAT2(G/T) overexpression leads to a disease phenotype resembling clinical manifestation of mixed MDS/MPN in humans without myelofibrosis.

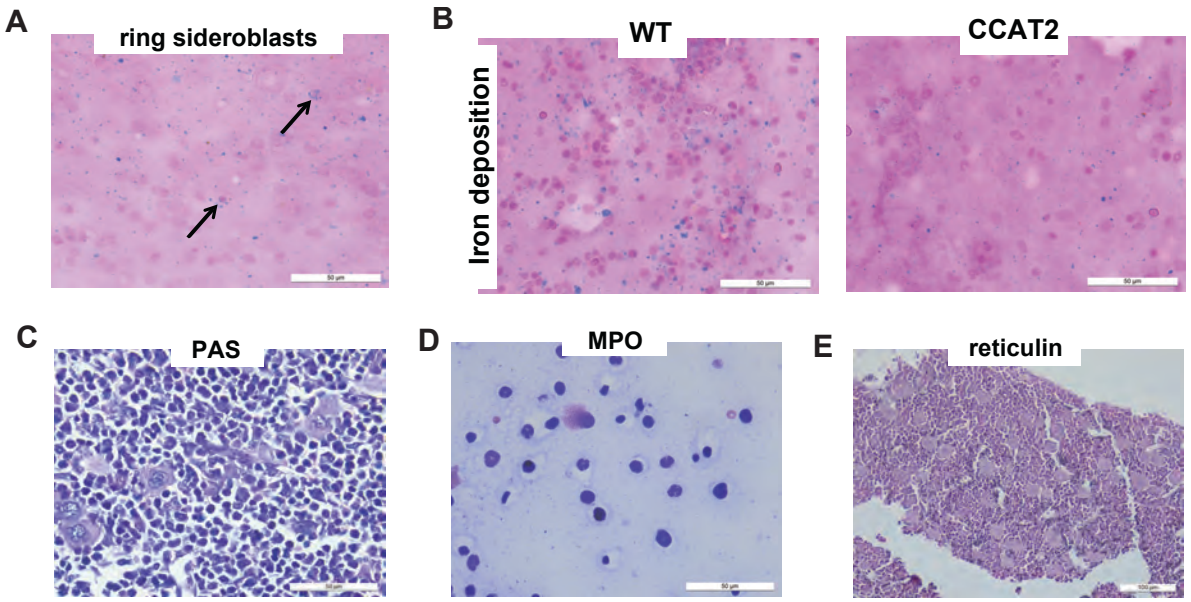


Figure 16: CCAT2(G/T) mice display characteristic features of MDS or mixed MDS/MPN phenotype.

(A, B) CCAT2(G/T) BM smears stained for iron. CCAT2(G/T) (n = 3 for each allele) mice show presence of ring sideroblasts in their BMs (A). The iron deposition is also reduced in BM of CCAT2(G/T) mice (n = 6) compared to WT littermates (n = 3).

(C, D) Negative PAS and myeloderoxidase staining in CCAT2(G/T) BM smears (n = 5 for each allele).

(E) Negative reticulin and trichome staining of CCAT2(G/T) BM smears (n = 3 for each allele).

Enhanced proliferation and excessive apoptosis of hematopoietic stem and immature progenitor cells is one of the characteristic features of MDS (REF). In order to further characterize the phenotype, we stained BM biopsies of CCAT2(G/T) (n = 6) and WT (n = 3) animals for Ki67 (proliferation marker). We observed a significant increase in Ki67-positive cells in CCAT2(G/T) mice compared to WT mice **(Figure 17A)**. Both MDS- and MDS/MPN-like mice displayed similar increase in proliferation. Next, we performed TUNEL analysis to identify apoptotic cells in BM of these mice. A significant increase in apoptosis was observed in CCAT2(G/T) mice **(Figure 17B).** These data indicate dyregulation in the production and maturation of hematopoietic cells in the BM of CCAT2(G/T) mice. **Figure 17C** shows the quantification of Ki67 and TUNEL staining in the WT and CCAT2 bone marrow biopsies. This data suggests dysregulation of hematopoietic cell maturation and maintenance, and might play an important role in bone marrow failure.

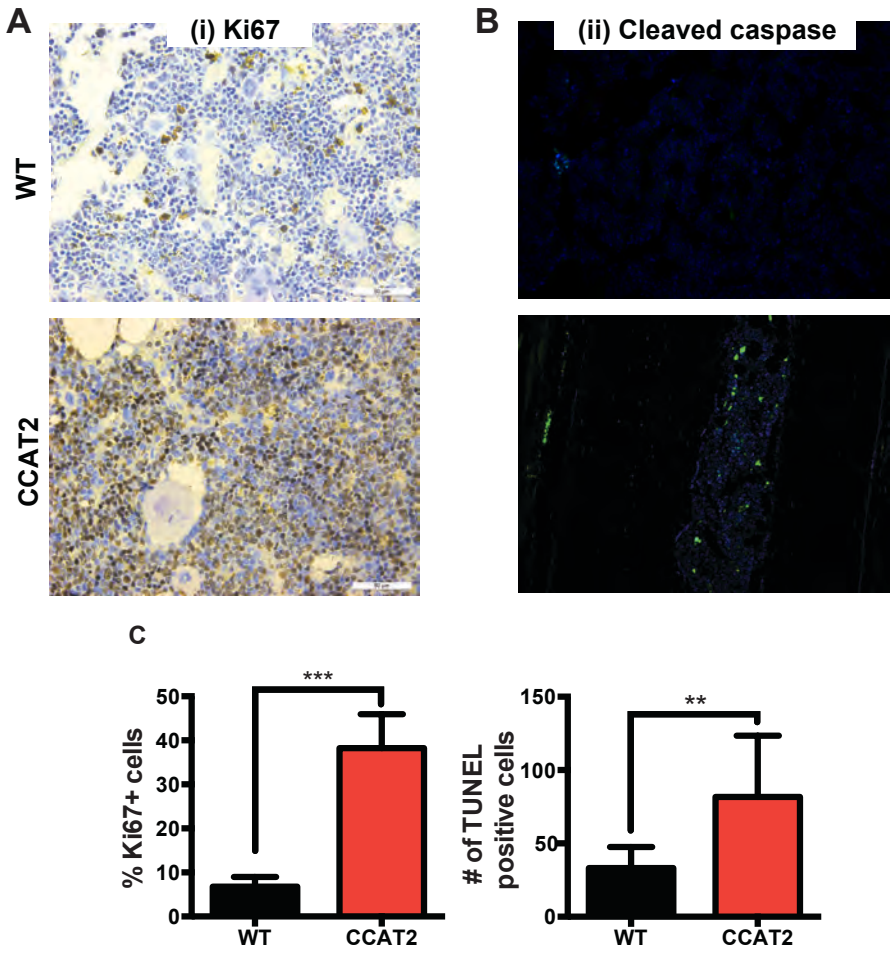


Figure 17: CCAT2(G/T) mice display excessive proliferation and enhanced apoptosis in BM.

(A) BM sections of CCAT2(G/T) and WT mice were stained with Ki67 marker by immunohistochemistry (IHC) to detect actively proliferating cells.

(B) TUNEL assay was performed on BM sections from CCAT2(G/T) and WT mice.

(C) Quantification of Ki67 and TUNEL staining in BM sections of CCAT2(G/T) and WT mice.

The data are presented as ± STD. *P < 0.05, **P < 0.01 and ***P < 0.001.

CCAT2(G/T) bone marrow cells are genomically unstable

CCAT2 overexpression is associated with genomic instability and an increase in frequency of chromosomal fusions and aneuploidy (Ling et al., 2013). Since genomic instability of the bone marrow (BM) cells is an important characteristic of MDS, it was possible that *CCAT2* overexpression could lead to increase in chromosomal abnormalities and thus contribute to MDS phenotype. To examine this possibility, metaphase spreads were prepared from bone marrow cells of WT (n = 4) and CCAT2(G/T) (n = 7) mice and were scored for abnormal mitosis. The BM cells of CCAT2(G/T) mice showed an increased frequency of cytogenetic aberrations (**Figure 18A**), and significant increase in number of cells with breaks and chromosomal fusions (**Figure 18B**). **Figure 18C** shows the different types of structural anomalies detected in CCAT2(G/T) mice, including chromosomal fusions, fragments, breaks, and ring chromosomes. Next, in order to detect if there were any specific recurrent mutations or chromosomal fusions in these BM cells, we performed karyotyping and SKY (spectral karyotyping) analysis on these mice. No significant genomic deletions or fusions were detected **(Figure 19A, 19B)**, suggesting that while the bone marrow cells show chromosomal instability, these aberrations do not follow a specific pattern of induction. Thus, induction of widespread genomic instability by *CCAT2* might be a key event in initiation of MDS.

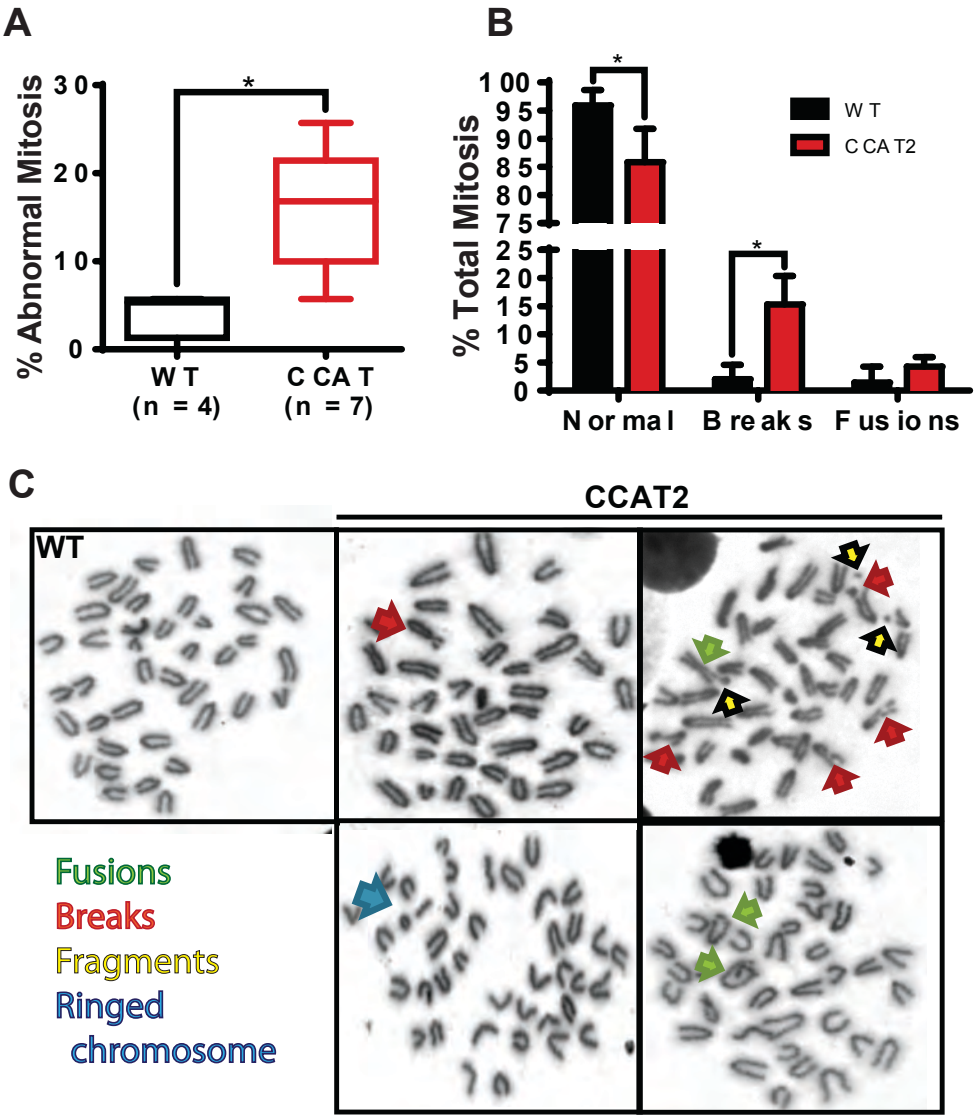


Figure 18: CCAT2(G/T) bone marrow cells are genomically unstable. Metaphase spreads were prepared from bone marrow cells of WT and CCAT2(G/T) mice, and were scored for abnormal mitosis.

(A) Quantification of abnormal mitosis in metaphase spreads of WT and CCAT2(G/T) mice.

(B) Percentage of BMCs showing fusions and breaks in WT and CCAT2(G/T) mice.

(C) Representative images of metaphase spreads from WT and CCAT2(G/T) mice showing several chromosomal aberrations present in these cells.

The data are presented as ± STD. *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$.

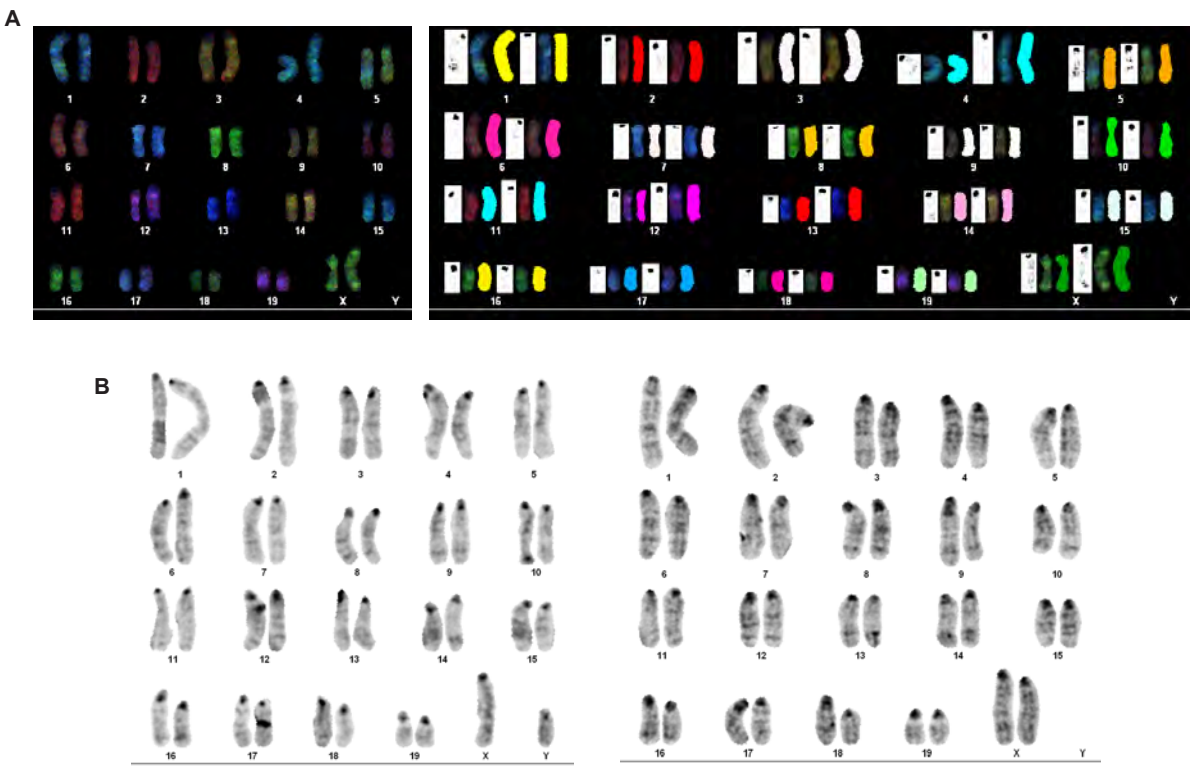


Figure 19: SKY and Karyotyping analysis on BMCs of CCAT2(G/T) and WT mice.

(A) Representative image showing SKY analysis on one CCAT2(G/T) mice.

(B) Representative image showing karyotyping analysis on two CCAT2(G/T) mice.

CCAT2(G/T) bone marrow cells do not show DNA damage

DNA damage is an important phenomenon that induces widespread genomic instability. Therefore, we sought to directly evaluate and quantify DNA damage in WT and CCAT2(G/T) BMCs. To this end, we used alkaline comet assay to measure single- and double-strand breaks in BMCs HSCs from 9-10 months old WT (n = 3) and CCAT2(G/T) (n = 6) mice. BMCs treated with peroxidase were assayed in parallel as a positive control. Analysis of Olive moment and percent tail DNA did not show significant differences between BMCs of WT and CCAT2(G/T) mice **(Figure 20A, B)**. Similar results were also obtained in CCAT2 mice displaying MDS and MPD phenotypes. These data were verified in six independent experiments.

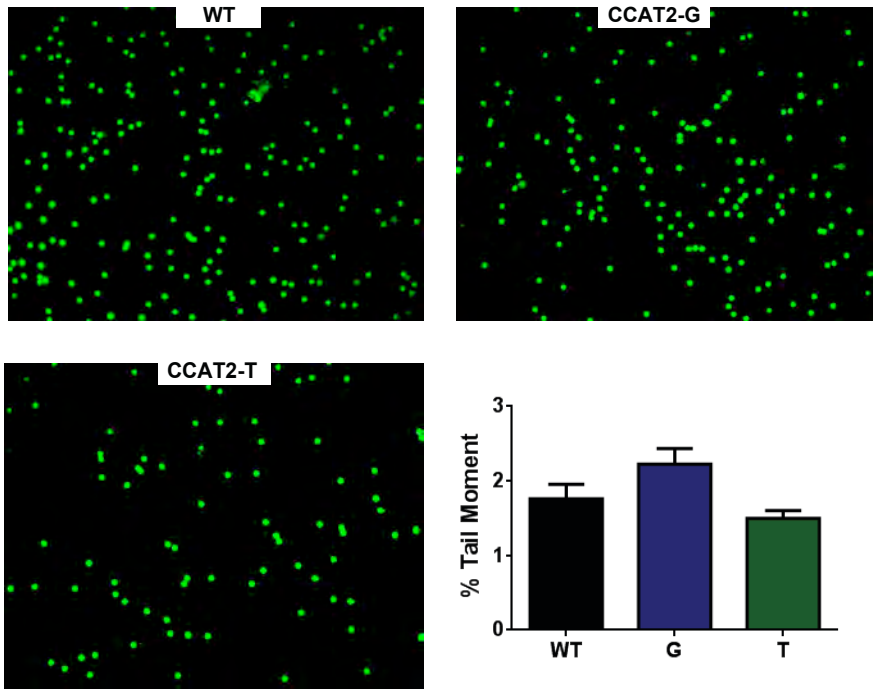


Figure 20: CCAT2(G/T) bone marrow cells do not show DNA damage. Alkaline comet assay was performed on BMCs of WT (n = 3) and CCAT2(G/T) (n = 3 for each allele) mice. Left panel shows representative images of comet. Right panel is the quantification of three independent experiments showing no significant difference between CCAT2(G/T) and WT BMCs.

CCAT2(G/T) mice have stable disease and do not progress to leukemias

We next characterized the CCAT2(G/T) and WT mice at different ages to monitor the progression of MDS or MDS/MPN in these mice. Peripheral blood analyses of young (<6 months old), 8/9-months old and older (>15 months old) detected an increased incidence of MDS/MPN phenotype in these mice with age, with almost 95% of the older mice analyzed displayed MDS/MPN (n = 8 for each allele). Early signs of mild hyperplasia were noted in the BM section of young mice **(Figure 21A)** which corroborated with mild cytopenias observed in the peripheral blood **(Figure 21B).** Interestingly, peripheral blood counts, BM sections, and BM smears of older mice showed similar widespread hyperplasia and cytopenias **(Figure 21C)** as 8/9-months of age as reported above. The data suggested an increased incidence of megakaryocytic hyperplasia in older CCAT2(G/T)-mice in comparison with 8/9-month old **(Figure 21D)** mice. Myeloid and erythroid hyperplasia were relatively constant in older CCAT2(G/T)-mice in comparison with 8/9-month old CCAT2(G/T)-mice **(Figure 21D)**. Overall, the main phenotype of myeloid and megakaryocytic hyperplasia in conjunction with various degrees of erythroid hypoplasia of older mice was still greater than 8/9-month old mice. Additionally, an increased EMH in the spleen of mice was detected in 2/4 G-mice and 1/4 T-mice. Increased number of megakaryocytes in spleen were present in 3/4 T-mice and 1/4 G-mice. Taken together, these results from cell blood count, histology and hematological/cytological examination of bone marrow smears suggest a progression to mixed MDS/MPN with age, while there was no progression to leukemia in any of the mice we analyzed.

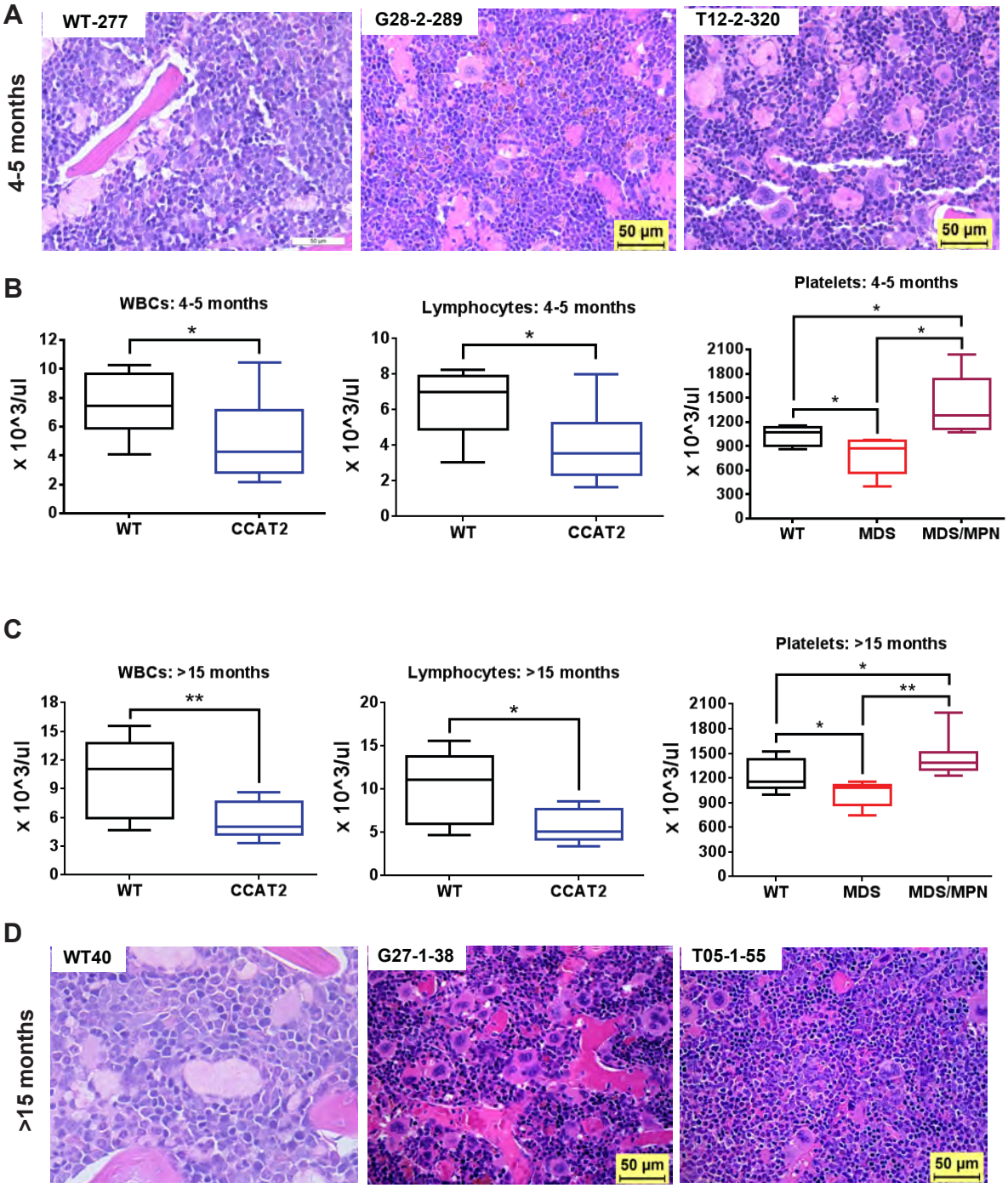

A
4-5 months
WT-277
G28-2-289
T12-2-320
50 µm
50 µm
B
WBCs: 4-5 months
x 10^3/ul
WT
CCAT2
Lymphocytes: 4-5 months
x 10^3/ul
WT
CCAT2
Platelets: 4-5 months
x 10^3/ul
WT
MDS
MDS/MPN
C
WBCs: >15 months
x 10^3/ul
WT
CCAT2
Lymphocytes: >15 months
x 10^3/ul
WT
CCAT2
Platelets: >15 months
x 10^3/ul
WT
MDS
MDS/MPN
D
>15 months
WT40
G27-1-38
T05-1-55
50 µm
50 µm

Figure 21: CCAT2(G/T) mice display stable disease with age and do not progress to leukemias.

(A) H&E staining of BM sections of 4-5-months old CCAT2(G/T) (n = 5 for each allele) and WT mice (n = 5).

(B) Total blood counts for WBCs, lymphocytes and platelets in peripheral blood of 4-5-months old CCAT2(G/T) and WT mice.

(C) Total blood counts for WBCs, lymphocytes and platelets in peripheral blood of >15-months old CCAT2(G/T) and WT mice.

(D) H&E staining of BM sections of >15-months old CCAT2(G/T) and WT mice.

The data are presented as ± STD. *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$.

CCAT2 induces exhaustion of hematopoietic stem cells

To further characterize the mechanisms underlying aberrant differentiation and maturation in BM of CCAT2(G/T) mice and given the important role of hematopoietic stem cell and progenitor cells in induction of ineffective hematopoiesis and myelodysplasia, we performed flow cytometric analyses on BMCs in CCAT2(G/T) (n = 8 for each allele) mice compared to age- and sex-matched WT control (n = 8) mice. Interestingly, the MDS and MDS/MPN-like mice showed presence of significantly different populations in their BM. The proportion of LSK ($Lin^{-}Sca1^{+}cKit^{+}$) cells was significantly decreased in MDS-CCAT2(G/T) mice, while no significant alteration was observed in MDS/MPN-CCAT2(G/T) mice compared to WT mice (**Figure 22A**).

Correspondingly, the percentage of long-term HSCs (LT-HSCs, defined by $Lin^{-}c\text{-}Kit^{+}Sca\text{-}1^{+}CD34^{lo}CD135^{lo}$ population), short-term HSCs (ST-HSCs, defined by $Lin^{-}c\text{-}Kit^{+}Sca\text{-}1^{+}CD34^{hi}CD135^{lo}$ population) and multipotent progenitor cells (MPPs, defined by ($Lin^{-}c\text{-}Kit^{+}Sca\text{-}1^{+}CD34^{hi}CD135^{hi}$) were also significantly decreased only in MDS-CCAT2(G/T) as compared to MDS/MPN-CCAT2(G/T) and WT mice (**Figure 22B**).

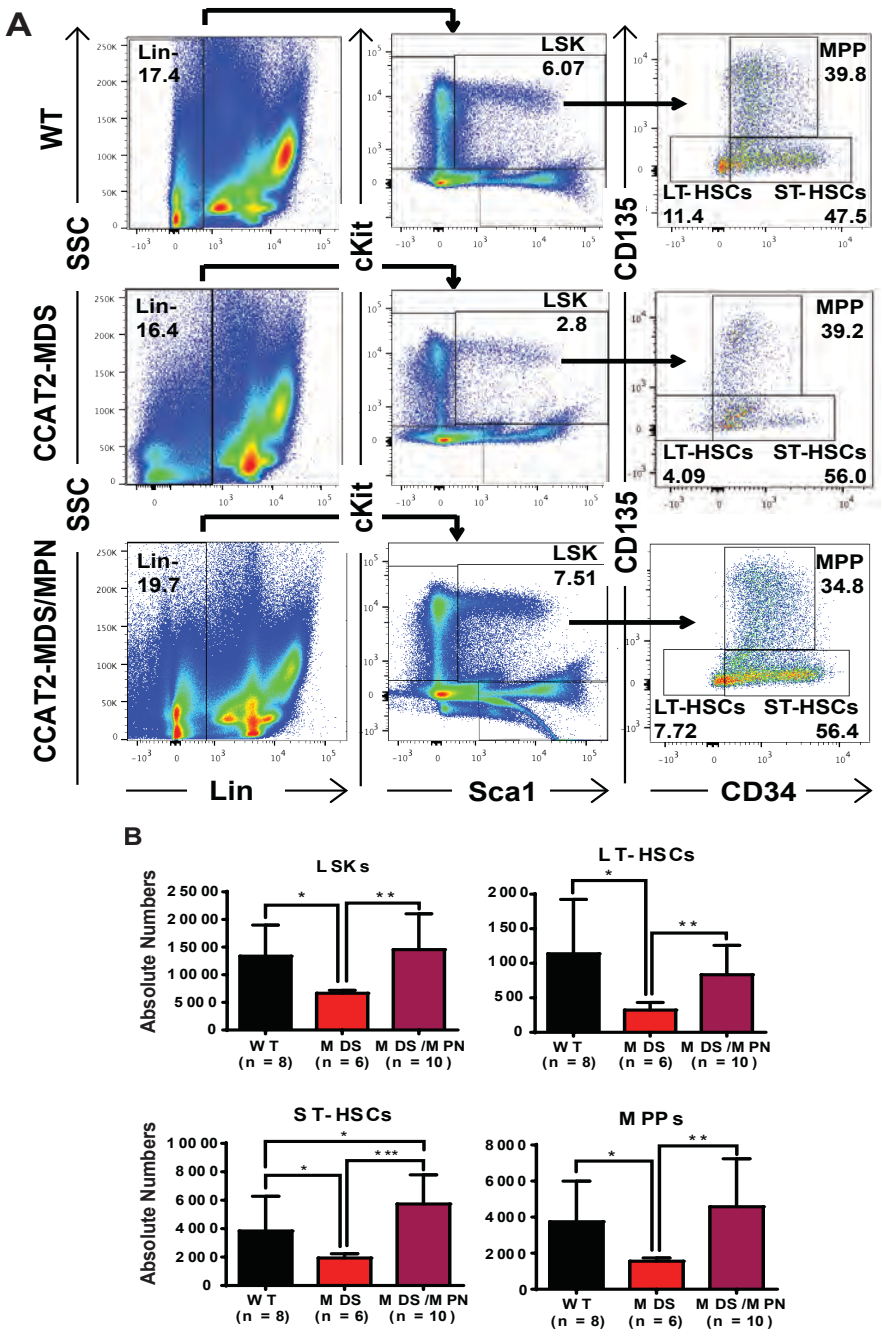


Figure 22: *CCAT2* induces exhaustion of hematopoietic stem cells

(A) Representative flow scatters of BMCs from WT and CCAT2(G/T) mice with distinct MDS-like and MDS/MPN-like profiles showing percentage of LSK HSCs, ST-HSCs, LT-HSCs and MPPs using HSC specific markers.

(B) Total number of LSK HSCs, ST-HSCs, LT-HSCs and MPPs cells in CCAT2(G/T) mice with distinct MDS-like and MDS/MPN-like profiles.

The data are presented as ± STD. **P* < 0.05, ***P* < 0.01 and ****P* < 0.001.

Among the lineage positive cells, the percentage of immature B cells (defined by $B220^{+}IgM^{-}$) was significantly reduced in CCAT2-MDS mice compared to WT controls and MDS/MPN-CCAT2(G/T) mice (**Figure 23**). Paired with lower RBC peripheral blood levels, this suggests a block in differentiation to mature erythroid in the blood. Conversely, we noted a distinct infiltration of CD4+/CD8+ T cells in the BM of MDS/MPN-CCAT2(G/T) mice (**Figure 24A**). The clonal expansion was more evident in CCAT2-T-MDS/MPN mice (**Figure 24B**), suggesting that T allele might be more important in expansion of CD8+ T cells. The flow cytometry data highlight the two distinct phenotypes induced in CCAT2(G/T) mice, suggesting an important role of CCAT2 in regulating the HSC pool, which might in turn alter the differentiation efficiency of remaining progenitor cells in mice.

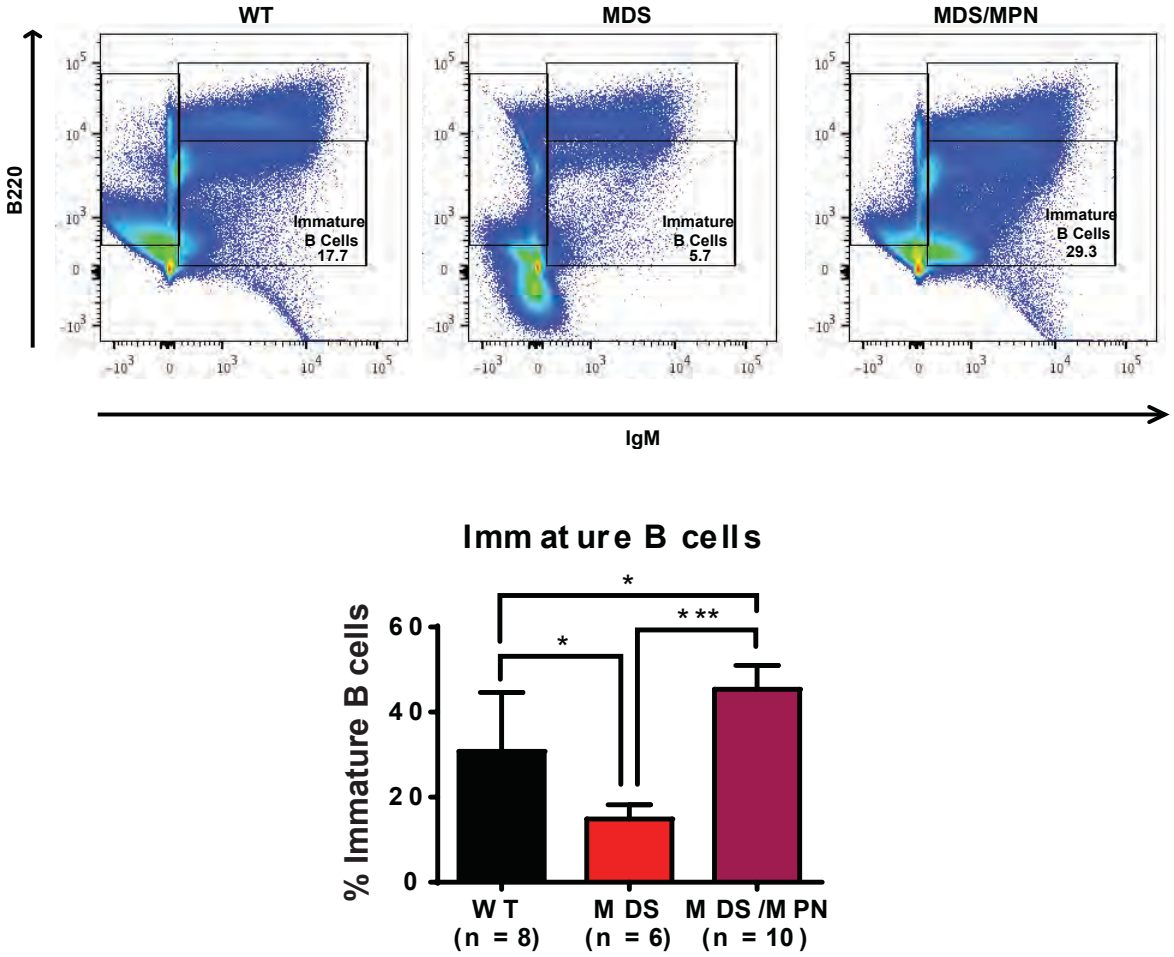


Figure 23: Immature B cells in bone marrow of *CCAT2* mice

(A) Representative flow scatters of BMCs from WT and CCAT2(G/T) mice with distinct MDS-like and MDS/MPN-like profiles showing percentage of immature B cells using specific markers.

(B) Percent Immature B cells in CCAT2(G/T) mice with distinct MDS-like and MDS/MPN-like profiles.

The data are presented as ± STD. *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$.

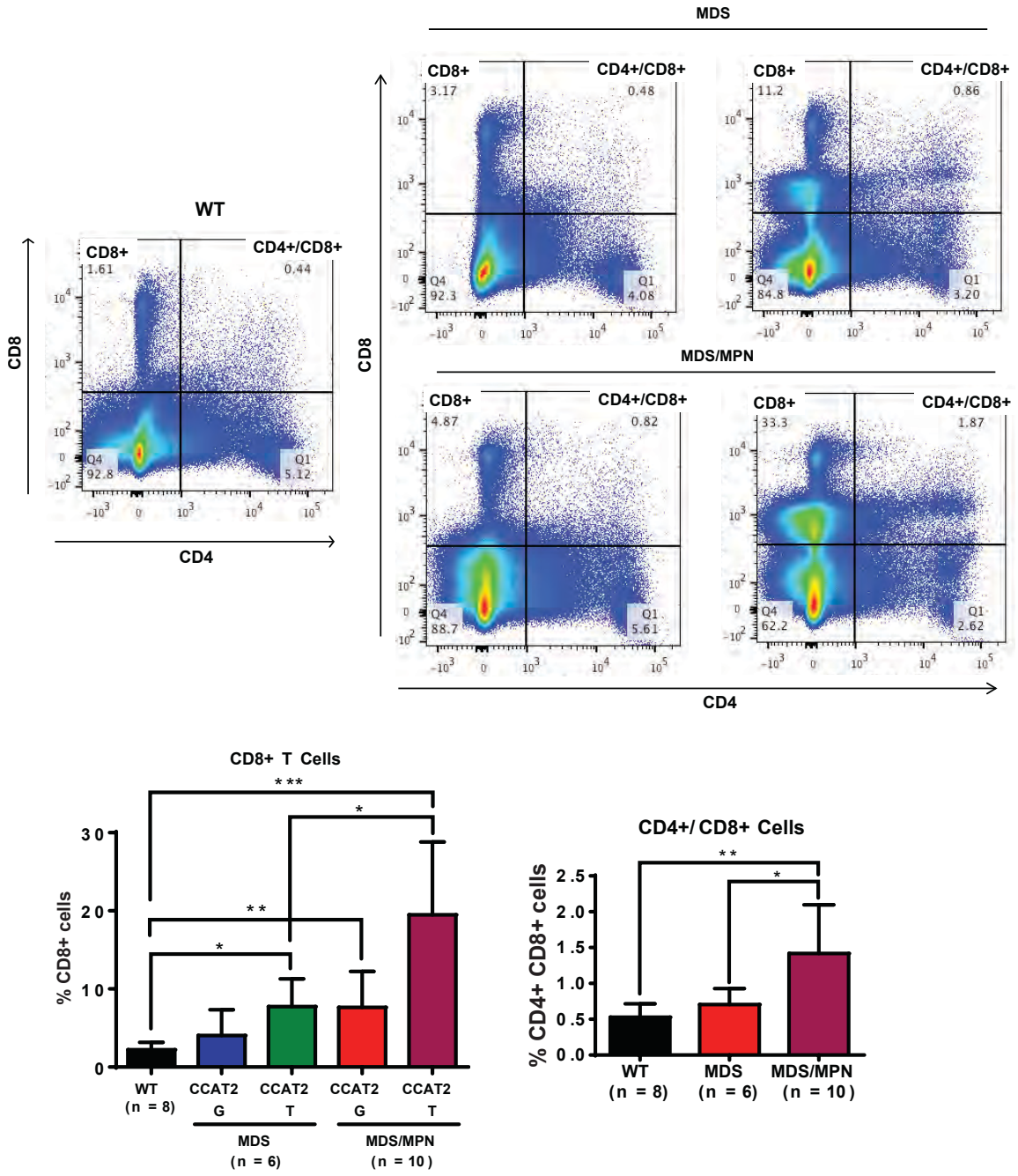


Figure 24: Infiltration of activated T cells in bone marrow of *CCAT2* mice

(A) Representative flow scatters of BMCs from WT and CCAT2(G/T) mice showing an infiltration of CD4+/CD8+ T cells in MDS/MPN-like mice and CD8+ T cells specifically in CCAT2-MDS/MPN-T mice.

(B) Percent CD4+/CD8+ (left panel) and CD8+ (right panel) T cells in CCAT2(G/T) mice with distinct MDS-like and MDS/MPN-like profiles.

The data are presented as ± STD. **P* < 0.05, ***P* < 0.01 and ****P* < 0.001.

To verify the expression of B and T cells in BM, we stained BM sections of WT and CCAT2(G/T) mice for markers of B (CD20) and T (CD3) cells. We detected a remarkable increase in CD20+ immature B-cell precursors in the BM sections of CCAT2-MDS/MPN-like mice but not CCAT2-MDS mice as compared to WT mice. Additionally, presence of CD3+ immature T cells was also specifically increased in CCAT2-MDS/MPN mice only. These data confirm our flow cytometry data and suggest defective maturation and clonal expansion of the T-lymphocytes (**Figure 25**).

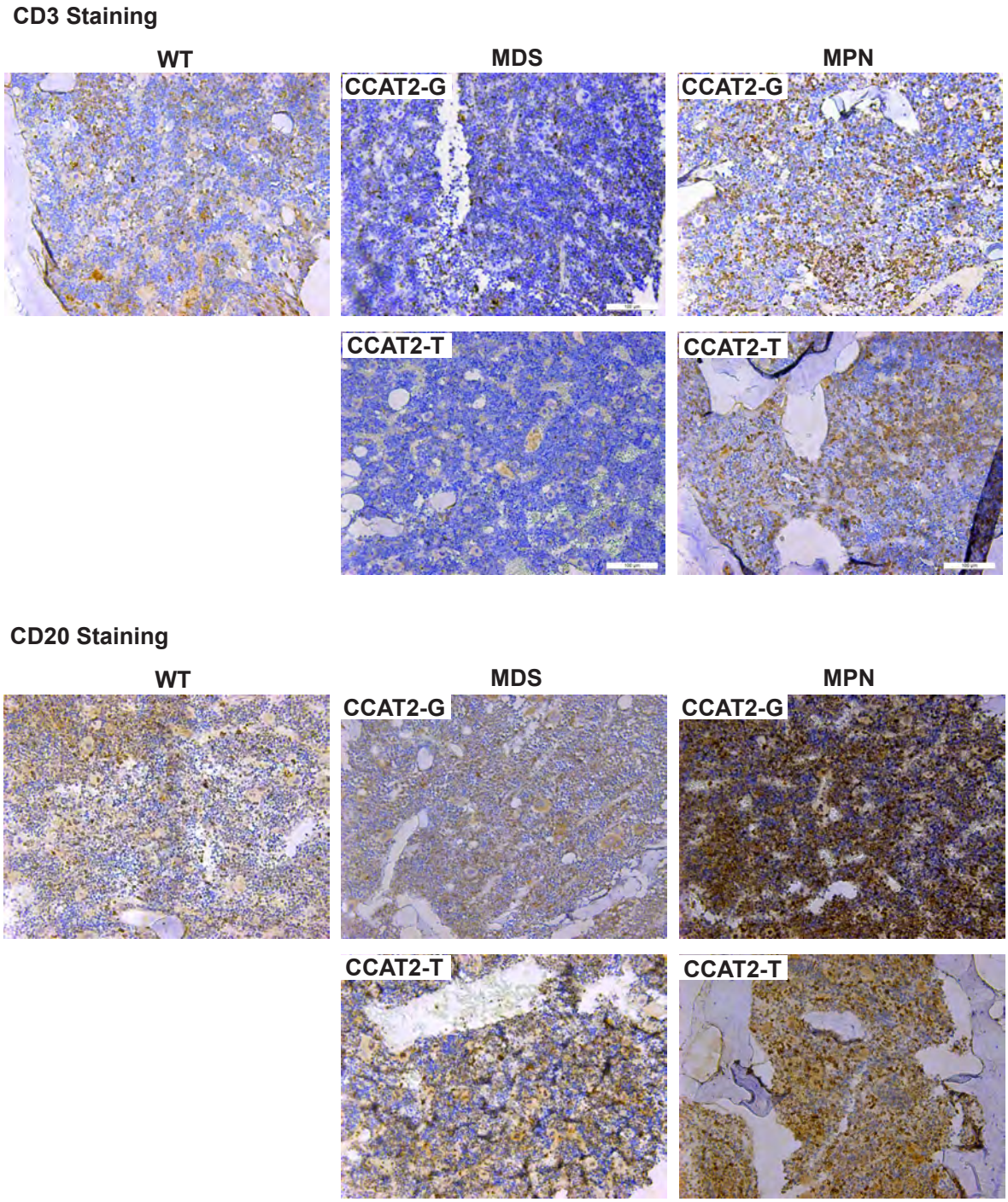


Figure 25: Immature B cells and activated T cells in bone marrow of *CCAT2* mice

Immunohistochemical stainings of CD3 (upper panel) and CD20 (lower panel) in BM sections of WT and CCAT2(G/T) mice with distinct MDS-like and MDS/MPN-like features.

In vitro hematopoietic stem cell assay demonstrated a significant decrease in colony-forming capabilities of HSCs from CCAT2 mice **(Figure 26A**). HSCs from both CCAT2-MDS and CCAT2-MDS/MPN mice showed similar reduction in colony formation, suggesting that the HSCs from CCAT2 mice have compromised differentiation. Serial re-plating assay on HSCs from CCAT2-MDS and CCAT2-MDS/MPN revealed a remarkable decrease in re-populating efficiency of CCAT2-HSCs *in vitro* (**Figure 26B).** These data suggest a significant alteration in the self-renewing and differentiation efficiency of HSCs from CCAT2(G/T) mice.

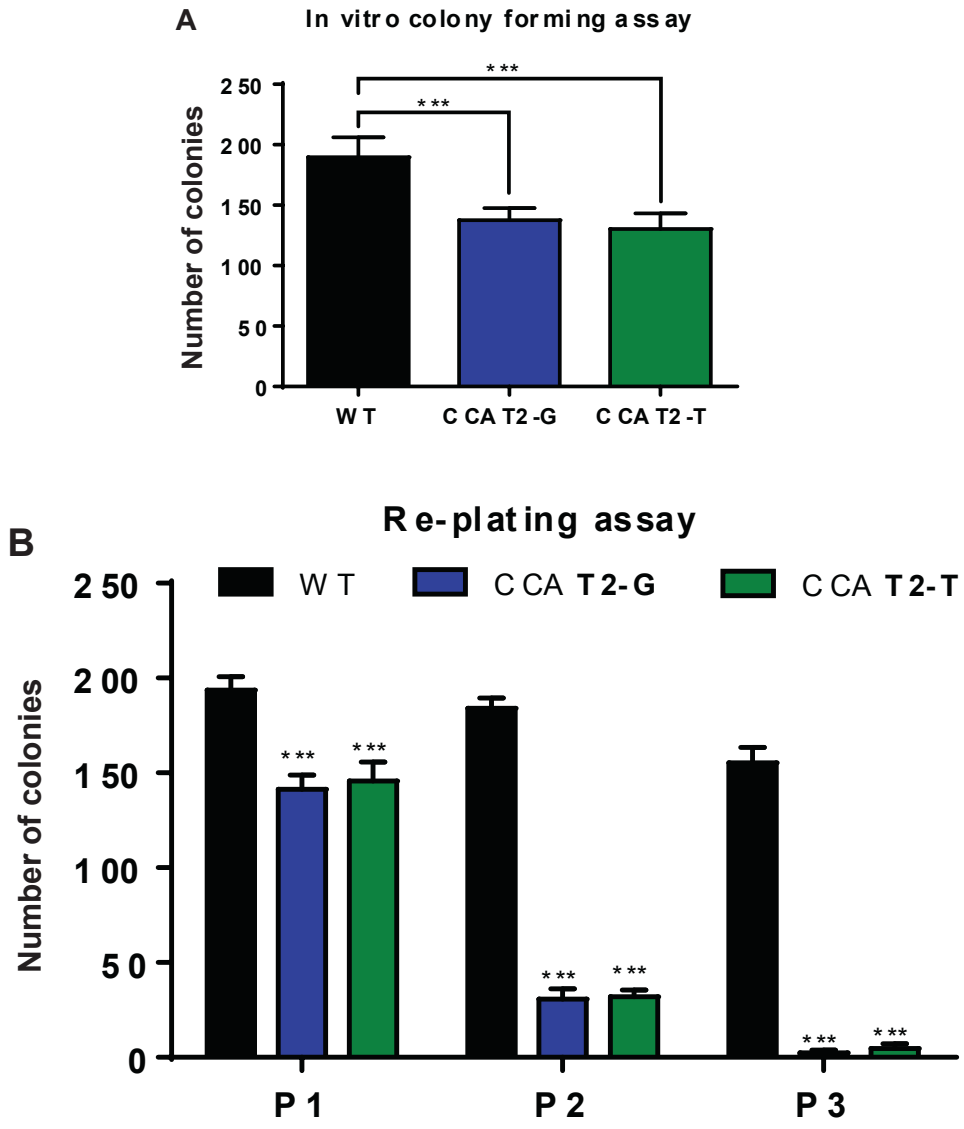


Figure 26: CCAT2 induces exhaustion of hematopoietic stem cells *in vitro*

(A) In vitro colony formation assay in methylcellulose was performed using total BMCs from WT and CCAT2(G/T) mice. The graph is the average of three independent experiments.

(B) Total BMCs from WT and CCAT2(G/T) mice were subjected to serial re-plating analysis. The graph is the average of three independent experiments.

The data are presented as ± STD. **P* < 0.05, ***P* < 0.01 and ****P* < 0.001.

CCAT2 regulates EZH2 expression in an allele specific manner

To determine the mechanism by which *CCAT2* induces genomic instability and myelodysplasia, we screened for several genes that have been previously reported to induce myelodysplasia as potential targets of CCAT2. Interestingly, EZH2 was downregulated in the BMCs of $CCAT2^{(G/T)}$ mice compared to WT littermates (**Figure 27A).** EZH2 downregulation was observed in both MDS only and MDS/MPD mice. While EZH2 is frequently overexpressed and considered to be an oncogene in cancers; nevertheless, EZH2 is considered as a candidate tumor suppressor gene in MDS/MPN (Ernst et al., 2010). 10% of MDS show loss-of-function mutations in EZH2, and these mutations are associated with poor survival (Ernst et al., 2010). Ezh2 loss has also been reported to induce development of myelodysplastic syndrome *in vivo*, but attenuates its predisposition to leukaemic transformation (Muto et al., 2013; Sashida et al., 2014). In CCAT2(G/T) mice, EZH2 and H3K27Me3 reduction was observed in hematopoietic stem and progenitor cells (HSPCs) as well as lineage positive bulk cells (**Figure 27B),** suggesting that CCAT2 might induce alteration in EZH2 levels in the HSC compartment. Next, we wanted to determine if EZH2 expression was downregulated pre- or post-transcriptionally. Using qRT-PCR analysis, we found that EZH2 transcripts were not significantly downregulated in these CCAT2(G/T) mice **(Figure 27C)**, suggesting a post-transcriptional mechanism of expression levels.

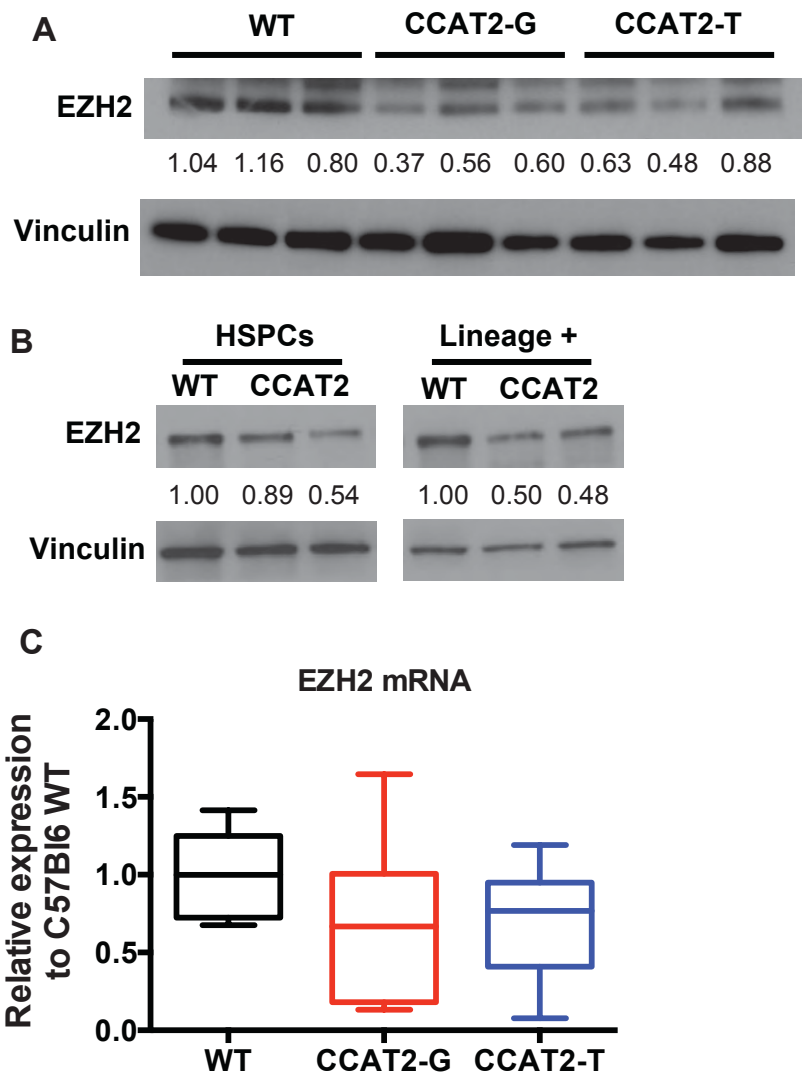


Figure 27: EZH2 is downregulated in CCAT2(G/T) mice

(A, B) Western blot analysis for EZH2 was performed on total BMCs (A) or enriched hematopoietic stem and progenitor cells (HSPCs) from WT and CCAT2(G/T) mice.

(B) Real time qPCR analysis was performed on total BMCs from WT and CCAT2(G/T) mice. The graph is the average of three independent experiments.

The data are presented as ± STD. *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$.

In order to determine if the downstream pathway of EZH2 was affected, we performed western blot analysis for H3K27Me3 (histone trimethylation at lysine 27), a primary function of EZH2. A corresponding decrease in global H3K27Me3 levels was observed in these mice **(Figure 28A)**. Since p21 is one of the most important trimethylation targets of EZH2, we checked for p21 levels in these mice. Consequently, an increase in p21 was also observed (**Figure 28B)**. Since maintenance of gene expression via histone trimethylation by EZH2 is an important event in hematopoietic stem cells, we next determined if this phenomenon was affected in CCAT2(G/T) mice. We observed a corresponding decrease in H3K27Me3 levels in the stem cells **(Figure 28C)**. This indicates a global gene expression dysregulation in the hematopoietic stem cells in CCAT2(G/T) mice. Based on these data, we concluded that EZH2 expression levels and its functional activity were downregulated in CCAT2(G/T) mice, and this was true in both MDS-like and MDS/MPN-like CCAT2(G/T) mice.

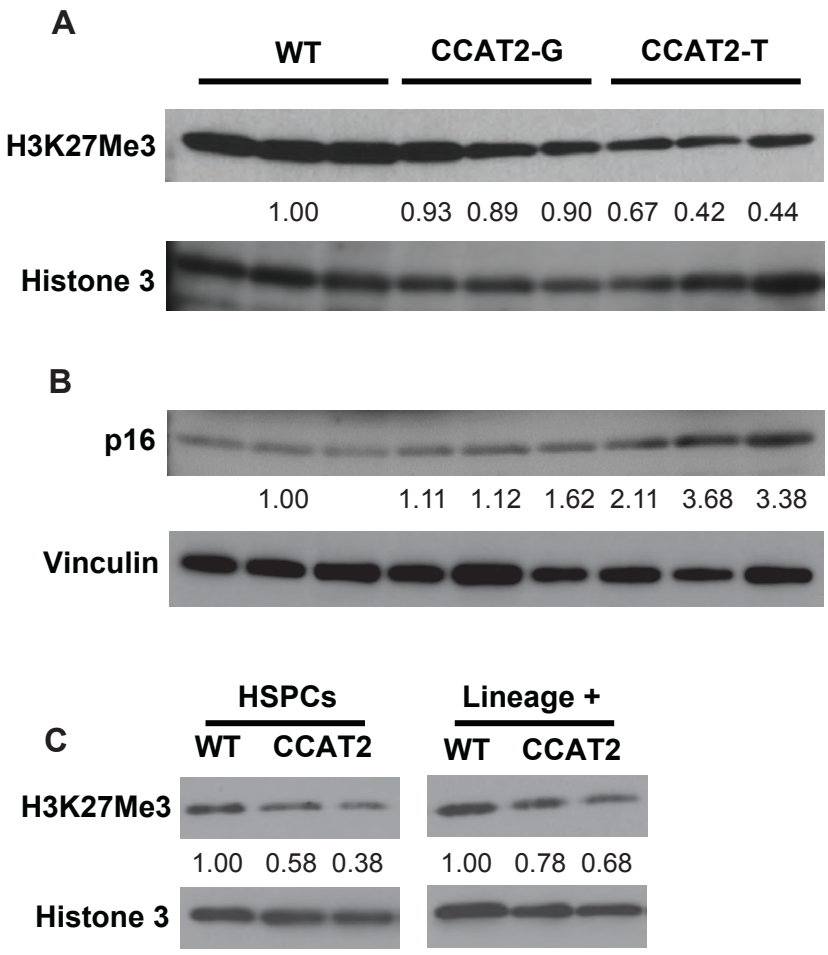


Figure 28: EZH2 functional activity is compromised in CCAT2(G/T) mice

Western blot analysis for H3K27Me3 (A) and p21 (B) levels was performed on total BMCs or enriched hematopoietic stem and progenitor cells (HSPCs, C) from WT and CCAT2(G/T) mice.

To determine the mechanism by which *CCAT2* regulates EZH2 expression, we performed microRNA microarray on the BMCs from age- and sex-matched WT (n = 3) and *CCAT2* (n = 6) BMCs. Microarray analysis identified 92 miRNAs that were dysregulated in either CCAT2-G or –T mice compared to WT control mice. Interestingly, using Ingenuity pathway analysis, one of the top significant molecular and cellular functions associated with these potential CCAT2-regulated miRNAs included a subset important in hematological malignancies (data not shown). We validated the expression of some of the potential miRNAs in BMCs of CCAT2(G/T) and WT mice, and identified miR-26a, miR-26b, miR-21, miR-150, and miR-155 to be significantly upregulated in CCAT2-G mice, while miR-130c to be significantly upregulated in both CCAT2-G and -T mice (**Figure 29A**). We next used multiple miRNA target prediction programs (RNA22, TargetScan, miRanda, microT, and PicTar) to determine whether these miRNAs could potentially target EZH2. Interestingly, we identified miR-26a, miR-26b, miR-155, and miR-150 were either already reported or predicted to target EZH2 **(Figure 29B)**. In order to further identify the role of CCAT2 in these interactions, we determined expression levels of these miRNAs in CCAT2-overexpressing SET2 cells **(Figure 29C)**. Interestingly, only CCAT2-G overexpressing cells showed a significant increase in expression levels of these miRNAs, emulating what we observed in the BMCs of WT and CCAT2-G mice. These data suggests that CCAT2-G regulates EZH2 expression primarily through regulation of target miRNAs.

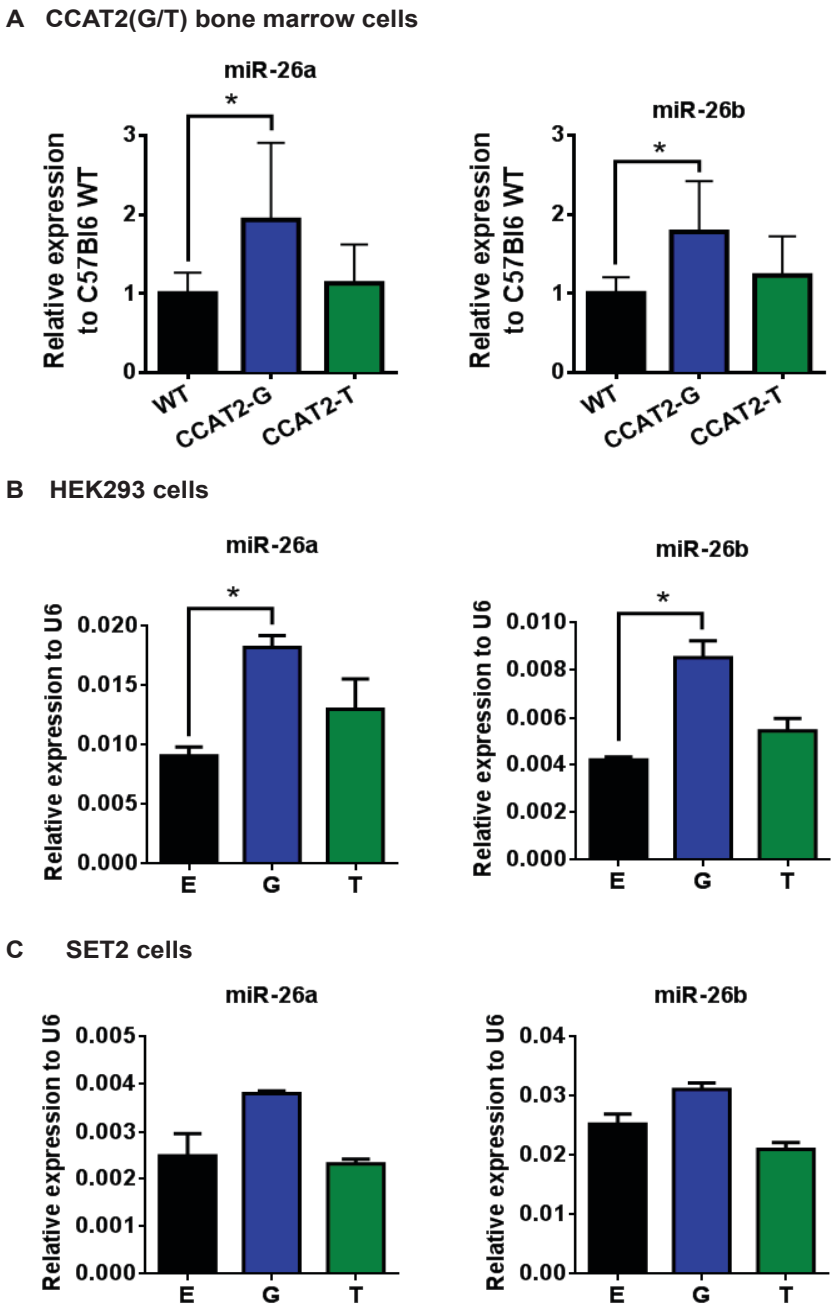


Figure 29: EZH2 targeting miRNAs are upregulated in CCAT2-G mice

Expression levels of miR-26a and miR-26b, miRNAs that target EZH2, were analyzed using Taqman real time qPCR in CCAT2(G/T) BMCs (A), CCAT2-overexpressing HEK293 cells (B) and CCAT2-overexpressing SET2 cells (C).

The graph is the average of three independent experiments. The data are presented as ± STD. *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$.

On performing CCAT2-EZH2 sequence analysis, a potential EZH2 binding motif was identified on the CCAT2 transcript. RNA-pulldown assay using MS2 vectors containing *CCAT2*- G or T allele successfully identified EZH2 to interact preferentially to CCAT2-T compared to empty MS2 or CCAT2-G transcript (**Figure 30A).** To further evaluate the binding affinity of EZH2 to a single nucleotide on CCAT2, we repeated the RNA-pulldown assay using MS2 vectors with mutated SNP into an A or C, and confirmed that EZH2 preferentially binds to only the CCAT2-T transcript (**Figure 30B).** RNA-Immunoprecipitation (RIP) assay using EZH2 antibodies successfully detected CCAT2 transcript in HEK293 stable clones overexpressing the CCAT2-T allele, but not from HEK293-CCAT2-G or the empty cells **(Figure 30C).** To confirm this interaction *in vivo*, we performed RIP on BMCs from CCAT2-G, -T and WT mice using EZH2 antibodies. qPCR amplification confirmed interaction with EZH2 in CCAT2-T mice but not in WT or CCAT2-G mice **(Figure 30D**). These data confirmed that EZH2 preferentially binds to the CCAT2 in an allele-specific manner.

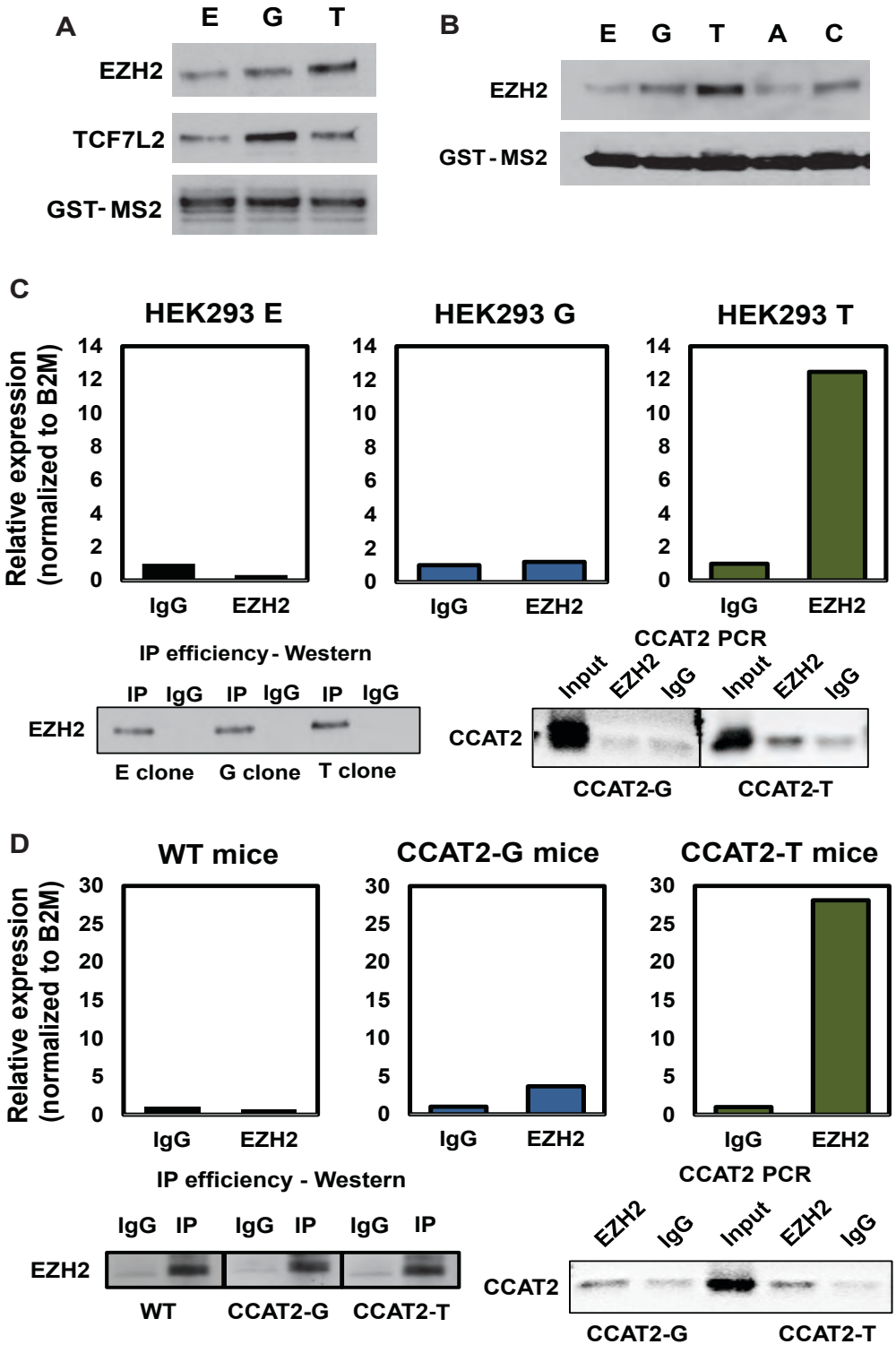

A
E G T
EZH2
TCF7L2
GST- MS2
B
E G T A C
EZH2
GST - MS2
C
HEK293 E
HEK293 G
HEK293 T
Relative expression (normalized to B2M)
14
12
10
8
6
4
2
0
IgG
EZH2
IP efficiency - Western
IP IgG IP IgG IP IgG
EZH2
E clone G clone T clone
CCAT2 PCR
Input EZH2 IgG Input EZH2 IgG
CCAT2
CCAT2-G CCAT2-T
D
WT mice
CCAT2-G mice
CCAT2-T mice
Relative expression (normalized to B2M)
30
25
20
15
10
5
0
IgG
EZH2
IP efficiency - Western
IgG IP IgG IP IgG IP
EZH2
WT CCAT2-G CCAT2-T
CCAT2 PCR
EZH2 IgG Input EZH2 IgG
CCAT2
CCAT2-G CCAT2-T

Figure 30: EZH2 specifically binds to *CCAT2-T* lncRNA

(A, B) RNA pull-down analysis was performed in vitro using GST-fusion plasmids expressing different CCAT2 alleles. Western blot analysis was performed on the pull-down lysate. TCF7L2 was used as a positive control. GST-MS2 was used as a loading control.

(C, D) RNA-Immunoprecipitation analysis was performed on CCAT2-overexpressing HEK293 cells (C) or total BMCs from CCAT2(G/T) and WT mice (D). Real time qPCR analysis was performed to identify CCAT2. B2M was used a non-target internal control. The experiment was repeated three times.

Next, to identify how *CCAT2* regulates EZH2 expression, we performed cycloheximide chase assay on CCAT2-overexpressing HEK293 stable clones. As shown in **Figure 31**, 10 hours after cycloheximide treatment, the levels of EZH2 reduced to 65% compared to its original amount only in CCAT2-T clones, while no significant alteration was noted in CCAT2-G or empty clones. This data was confirmed in three independent experiments. Together these experiments demonstrate that CCAT2 regulates the stability of EZH2 in an allele-dependent manner.

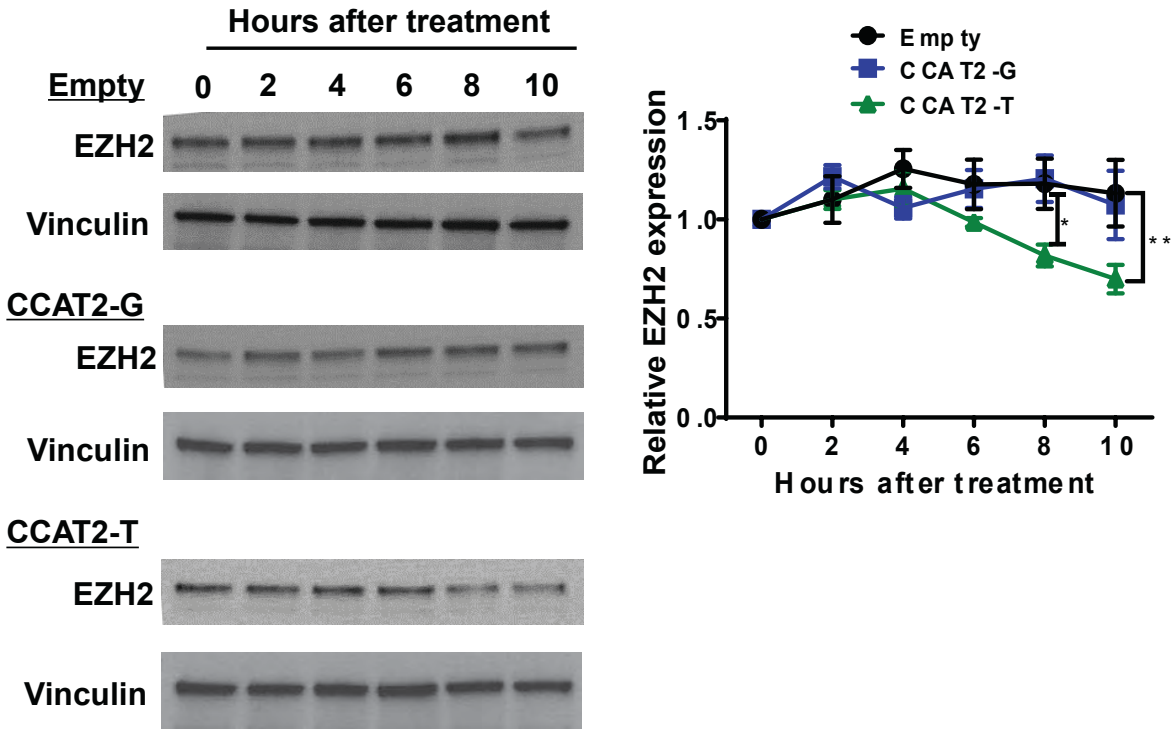


Figure 31: CCAT2-T lncRNA reduces the stability of EZH2 protein

CCAT2-overexpressing HEK293 cells were treated with cycloheximide and lysates were collected every two hours till 10 hours. Western blow analysis was performed to determine the stability of EZH2 protein in these cells. Right panel shows the average of three independent experiments.

The data are presented as ± STD. *$P < 0.05$, **$P < 0.01$ and ***$P < 0.001$.

CCAT2 is overexpressed in human MDS patients

To access the relevance of these findings to humans, we examined CCAT2 expression in CD34+ cells from the bone marrow biopsies of MDS of all subtypes including MDS that had transformed to AML. We identified significantly higher CCAT2 expression in the MDS patients (n = 70) as compared to healthy volunteers (n = 5) **(Figure 32A)**. CCAT2 overexpression was observed in both high risk and low risk patients **(Figure 32B)**, while no significant difference in CCAT2 expression was noted between groups of different risks classified according to the IPSS system **(Figure 32C)**. No significant difference in CCAT2 expression was detected between MDS patients presenting different cytogenetic aberrations (data not shown). Additionally, CCAT2 overexpression was not correlated with survival advantage in MDS patients. We next compared CCAT2 expression levels in patients with only MDS to those that progressed to AML. Interestingly, patients with AML had significantly lower expression of CCAT2 as compared to patients with MDS (**Figure 32E**). This is in agreement with our initial observation that CCAT2 levels decreased with increased risk of AML progression. Taken together, these data indicate that CCAT2 plays a critical role in initiation of MDS, while it doesn't play an important role in its progression to AML.

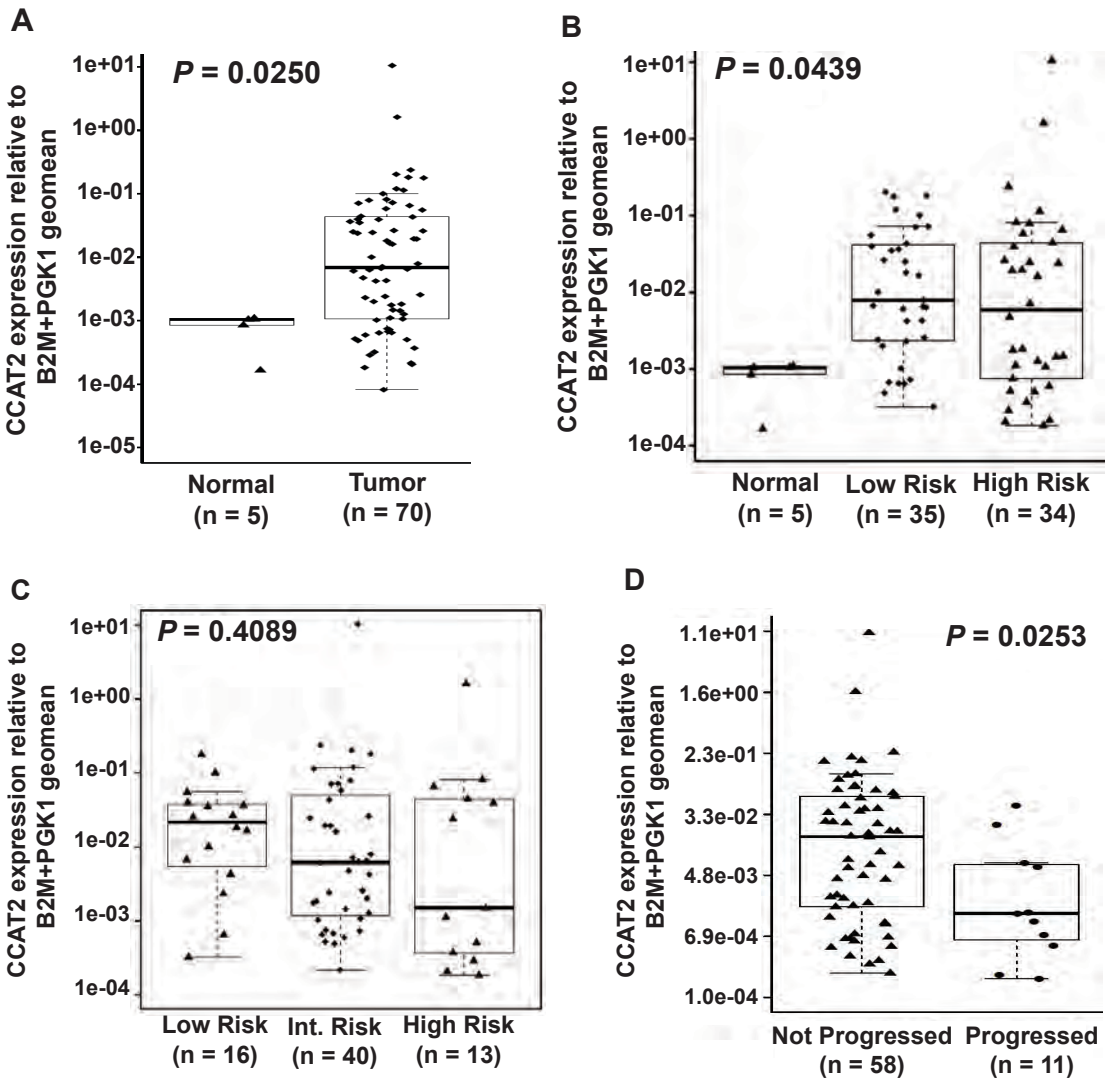


Figure 32: *CCAT2* expression in MDS patient samples

qRT-PCR analysis for CCAT2 expression in CD34+ cells from BM biopsies of MDS patients.

CHAPTER IV: Discussion

In this study, we demonstrate that overexpression of CCAT2 lncRNA alters hematopoiesis and induces spontaneous de novo MDS- or mixed MDS/MPN-like syndrome which closely emulates the human diseases. We generated a whole-body transgenic mouse in order to determine the organ, and consequently the disease phenotype most sensitive to CCAT2 overexpression. We identified that one of the most critical roles of CCAT2 is in the regulation of hematopoietic stem cells. We initially expected to see a colon phenotype by overexpressing *CCAT2*, since we previously reported that high *CCAT2* predisposes to colon cancer tumors. Interestingly we observed that CCAT2 plays an important role in maintenance of hematopoietic homeostasis, regulating the self-renewal and differentiation capacity of hematopoietic stem cells. This is the first report describing that overexpression of a lncRNA spontaneously induces myelodysplasia or myeloproliferation. Deletion of *Xist* lncRNA has been previously reported to develop an aggressive phenotype comprising mixed myelodysplastic and myeloproliferative neoplasm, leukemia, primary myelofibrosis and histiocytic sarcoma (Yildirim et al., 2013). However, our CCAT2 mice models can serve as a robust model for studying initiation of de novo MDS/MPN that does not progress to secondary AML and as a pre-clinical model for evaluation of new therapies for MDS. It has high translational potential as *CCAT2* can be developed into a novel intervention target for MDS/MPN therapy.

We report here that overexpression of CCAT2 at clinically relevant levels induces two distinct MDS- or MDS/MPN-like features in a mouse model. In this study, we generated two separate transgenic mice lines, expressing each allele of CCAT2 transcript (G or T). Interestingly, we observed that both alleles

spontaneously induced myelodysplastic features in the mice, with distinct MDS or MDS/MPN-like characteristics. However we noted subtle allelic preferences in the prevalence of MDS- or MDS/MPN-features in CCAT2 mice. Approximately 55% of CCAT2-T mice developed MDS-like features, while almost 60% of CCAT2-G mice developed MDS/MPN-like phenotype. Our data indicates that overexpression of *CCAT2* transcript, and not essentially its allelic distribution, is important in initiation of MDS or MDS/MPN-like features. Instead, CCAT2 alleles G or T play an important role in regulating the downstream targets of CCAT2, since CCAT2-G and CCAT2-T have different secondary structures. Thus, further characterization of these mice will help us understand the role of a SNP in regulating the factors important in initiation of MDS and MDS/MPN. Additionally, this is the first study providing in vivo evidence of the role of rs6983267 SNP in predisposition of MDS and mixed MDS/MPN. A large-scale prospective study to further illustrate the relevance of this SNP in MDS and MDS/MPN will enhance our understanding of the biology of these diseases.

The CCAT2 mice displaying only MDS-like and mixed MDS/MPN-like phenotype show significantly different characteristics. CCAT2-MDS-like mice display predominant leukopenia, minor anemia, multilineage BM hyperplasia, exhaustion of hematopoietic stem cells, and increased proliferation and concurrent apoptosis of maturing cells in bone marrow; characteristic features of MDS in humans. Conversely, CCAT2-MDS/MPN-like mice show distinctive thrombocytosis, splenomegaly, and hepatomegaly in addition to other characteristic BM dysplastic features similar to MDS. An important distinguishing criterion is rampant peripheral thrombocytosis in MDS/MPN-like mice. This might be a manifestation of the

widespread megakaryocytic dysplasia observed in both MDS- and MDS/MPN-like mice. Megakaryocytic dysplasia leads to thrombocytopenia in MDS- but thrombocytosis in mixed MDS/MPN, suggesting that CCAT2 play a critical role in regulation of megakaryocytic differentiation. One additional difference is that while BM of CCAT2-MDS/MPN mice are hyperplasic, they do not show exhaustion of hematopoietic stem cells, suggesting an imbalance in hematopoietic homeostasis. One potential explanation for the distinct BM presentation between these mice could be that CCAT2 induces global chromosomal instability in the HSCs, altering the fine balance between proliferation and apoptosis in BMCs.

An interesting observation was the distinct infiltration of CD4+/CD8+ T cells in the BM of CCAT2-MDS/MPN mice. There is accumulating evidence associating MDS to autoimmune disease, and that the bone marrow failure in MDS is primarily mediated via autoimmune myelosuppression from T lymphocytes (Chamuleau et al., 2009; Epperson et al., 2001; Mailloux and Epling-Burnette, 2013; Rosenfeld and List, 2000). It is speculated that this is an important component in the marrow failure that may respond to immunosuppressive treatment with antithymocyte globulin or cyclosporine (Epperson et al., 2001). In CCAT2 mice, the clonal expansion was more evident in CCAT2-T-MDS/MPN mice, suggesting that T allele might be more important in expansion of CD8+ T cells. The CCAT2 model might help address important questions about the role of autoimmunity in MDS/MPN.

We have previously reported that one of the primary mechanisms of CCAT2's oncogenic behavior in MSS colorectal cancer is through initiation of global genomic instability (Ling et al., 2013). Our study reproduces this observation *in vivo*, since

CCAT2 transgenic mice show global chromosomal instability in their BMCs. While the underlying pathogenesis of MDS and MDS/MPN is unclear, it is thought that clonal expansion of somatically mutated unstable hematopoietic stem and progenitor cells play a critical role in it (Woll et al., 2014). Cytogenetic abnormalities in MDS arise primarily from accumulation of genomic damage, failure to repair such damage, or both. Our study demonstrates that overexpression of CCAT2 is sufficient to induce de novo myelodysplasia and myeloproliferative alterations. Our data suggests that higher CCAT2 levels lead to widespread accumulation of genetic alterations, manifested by recurrent chromosomal breaks and fragmentation, which in turn drives myelodysplasia. However, we did not detect any specific translocations or deletions in the BMCs of these mice as evidenced by SKY and G-banding analyses, suggesting that CCAT2 potentially regulates ubiquitous cellular processes that maintain global genomic integrity.

In order to determine the mechanism by which CCAT2 induces myelodysplasia, we screened for several genes that have been previously reported to induce myelodysplasia as potential targets of CCAT2. We identified EZH2 to be one of the potential downstream effectors of CCAT2. EZH2 protein levels were significantly downregulated in the HSCs of both CCAT2-MDS as well as CCAT2-MDS/MPN mice, suggesting that decline in EZH2 levels might be responsible for the extensive dysregulation we observed in BMCs. Our data is strongly supported by multiple studies implicating EZH2 in both MDS and mixed MDS/MPN (Ernst et al., 2010; Muto et al., 2013; Sashida et al., 2014). While EZH2 is frequently overexpressed and considered to be an oncogene in cancers; nevertheless, EZH2 is

considered as a candidate tumor suppressor gene in MDS/MPN. 10% of MDS show loss-of-function mutations in EZH2, and these mutations are associated with poor survival (Ernst et al., 2010). Additionally, EZH2 dysregulation has also been shown to promote genomic instability by altering the global epigenetic landscape in tumor cells (Gonzalez et al., 2011). Sashida et al (Sashida et al., 2014) reported that EZH2 loss is enough to induce MDS *in vivo*; while other studies report its concomitant loss with Tet2 (Muto et al., 2013) also significantly promotes MDS initiation. Conversely, (Herrera-Merchan et al., 2012) reported that ectopic re-expression of EZH2 induced mixed MDS/MPN-like phenotype in conditional knock-in murine models. While contrasting data exist about the loss or gain of EZH2 induces MDS or MDS/MPN-like features, it is evident that epigenetic dysregulation in the HSCs compartment by EZH2 plays a critical role in regulating their homeostasis.

Interestingly, multiple studies have reported that EZH2 loss actively alleviates the leukemogenic transformation capacity of HSCs (Sashida et al., 2014; Tanaka et al., 2012). This is further supported by the information that EZH2 loss-of-function mutations are rare in AML (Ernst et al., 2010). This data is important since all CCAT2 mice display stable disease with age and do not progress to AML even after 30 months of age. This suggests that alteration in CCAT2 expression levels might be an early genetic event that is essential in the initiation of myelo-alteration, but would typically require other co-occurring alterations (genetic “hits”) to progress to AML. Thus, this mouse model could be used to study the factors that are critical in transforming MDS or MPN to AML.

Another surprising finding of the study is the allele specific regulation of EZH2 by CCAT2. We report that CCAT2-G transcript potentially targets EZH2 via post-transcriptional regulation by miR-26a and miR-26b, microRNAs previously reported to target EZH2 (Dang et al., 2012; Lu et al., 2011; Wong and Tellam, 2008). On the other hand, CCAT2-T transcript directly binds to the EZH2 protein and reduces its stability. This is the first study reporting the role of a SNP in altering the target profile of a lncRNA. We hypothesize that the differences in the secondary structures of CCAT2 at the SNP locus potentially plays a role in this. More detailed studies are required to identify the how CCAT2-G regulates microRNA expression. Additionally, identification of the mechanism by which CCAT2-T destabilizes EZH2 protein will shed light on lncRNA-protein interaction.

We report for the first time that CCAT2 lncRNA in significantly upregulated in the CD34+ BMCs of MDS patients compared to normal healthy individuals. Interestingly, CCAT2 expression levels were significantly lower in MDS patients that progressed to AML, clearly differentiating the two groups. This observation is in congruence with the CCAT2 mice, since they also do not progress to AML. Thus, CCAT2 can be potentially used as a prognostic biomarker to identify MDS patients susceptible to AML progression. Further characterization of these mice will help understand the factors that govern initiation of MDS/MPN and its further progression to AML. Interestingly, trisomy 8, presence of three copies of chromosome 8, is an important genetic aberration present in 10 – 15% MDS patients, including about 5% patients with +8 as the sole aberration (Paulsson and Johansson, 2007). Additionally, in MDS/MPN-U patients, trisomy 8 is one of the most prevalent sole

aberrations (15%) present (DiNardo et al., 2014). Since CCAT2 is transcribed from the highly unstable 8q24.21 region, genomic amplification in trisomy 8 can be one of the potential explanations for CCAT2 overexpression. Further studies are warranted to determine the mechanism of CCAT2 overexpression in MDS and MDS/MPN patients.

MDS are a group of heterogeneous clonal stem cell disorders that are primarily characterized by ineffective hematopoiesis and prevalent cytopenias (Rosenfeld and List, 2000). The main prognostic factors for diagnosis of MDS include the number and severity of cytopenias, percentage of blasts in the marrow, and incidence of cytogenetic abnormalities, allowing their classification into different subgroups according to their risk of progression to acute myeloid leukemia and predicted survival (Tefferi and Vardiman, 2009). MDS is predominantly a disease of the elderly. About 86% of patients with MDS are diagnosed after the age of 60 years, with a median age of diagnosis being 76 years (Tefferi and Vardiman, 2009). Between 2006 and 2010, there were about 15,000-20,000 new cases of MDS per year in the United States. Despite this growing prevalence, the underlying mechanisms that induce de novo MDS remains poorly understood. Mixed MDS/MPN (unclassifiable), on the other hand, are patients at diagnosis with clinical, morphologic and laboratory features which overlap both those of MDS and MPN features, and that do not satisfy criteria for chronic myelomonocytic leukemia (CMML), atypical chronic myeloid leukemia BCR-ABL1 negative (aCML), or juvenile myelomonocytic leukemia (JMML). MDS/MPN-U is a rare diagnosis, making up less than 5% of all myeloid disorders. MDS/MPN-U is formally defined as “patients with

no preceding history of MDS or MPN, no recent cytotoxic growth factor therapy, no Philadelphia chromosome, *BCR-ABL1* fusion gene, *PDGFRA*, *PDGFRB* or isolated del(5q), t(3;3)(q21;q26) or inv(3)(q21q26), and with dysplastic features in ≥1 hematopoietic cell line, <20% blasts in the blood and bone marrow, prominent myeloproliferative features (i.e., platelet count ≥450×10^9/L or white blood cell count ≥13×10^9/L, with or without splenomegaly); or *de novo* disease with mixed myeloproliferative and myelodysplastic features which cannot be assigned to any other category of MDS, MPN or of MPS/MPN" (quoted from DiNardo et al., 2014). Consequently, no standard prognostic or treatment algorithms for MDS/MPN-U exist.

Current therapeutic options for MDS or MDS/MPN are largely dependent on the risk stratification of the patients (Tefferi and Vardiman, 2009). Only three drugs have been approved in the United States specifically for MDS, and second-line therapies do not exist beyond clinical trials or supportive care. Allogeneic hematopoietic stem cell transplant (HSCT) is the only potentially curative therapeutic option. However, approximately three-quarters of the MDS patient population are ineligible for HSCT due to older age and poor overall health status (Tefferi and Vardiman, 2009). While about 30% patients that progress to AML are typically treated with hypomethylating agents or considered for allogeneic stem-cell transplantation, the remaining 70% lower-risk patients are usually offered less aggressive therapies, mainly including transfusion support, or simply supportive symptomatic care (Bejar and Steensma, 2014). Thus, improved understanding of the factors that contribute to initiation of MDS and mixed MDS/MPN and accurate identification of patients with worse prognosis and higher risk of progression to AML

could have important implications for development of novel risk-appropriate therapies.

Thus, in conclusion, we propose that constitutive overexpression of CCAT2 leads to global genetic instability that drives the spontaneous initiation of de novo MDS and MDS/MPN **(Figure 33)**. According to our model, CCAT2 induces this effect partially by regulating the EZH2 protein levels in an allele specific manner, and thus altering the epigenetic landscape of the hematopoietic stem cells. Our study improves our understanding of MDS and MDS/MPN biology and further expounds the role of lncRNAs in tumorigenesis. The CCAT2-G/T model provides a unique opportunity to study the similarities and differences in the pathogenesis of MDS and mixed MDS/MPN diseases. This mouse model may be useful as a pre-clinical tool to test potential therapies for MDS as well as mixed MDS/MPN. Further, CCAT2 can be developed as a potential therapeutic target for management of MDS and MDS/MPN. Further studies are required to clearly define the potential roles and implications of CCAT2 in myeloid malignancies.

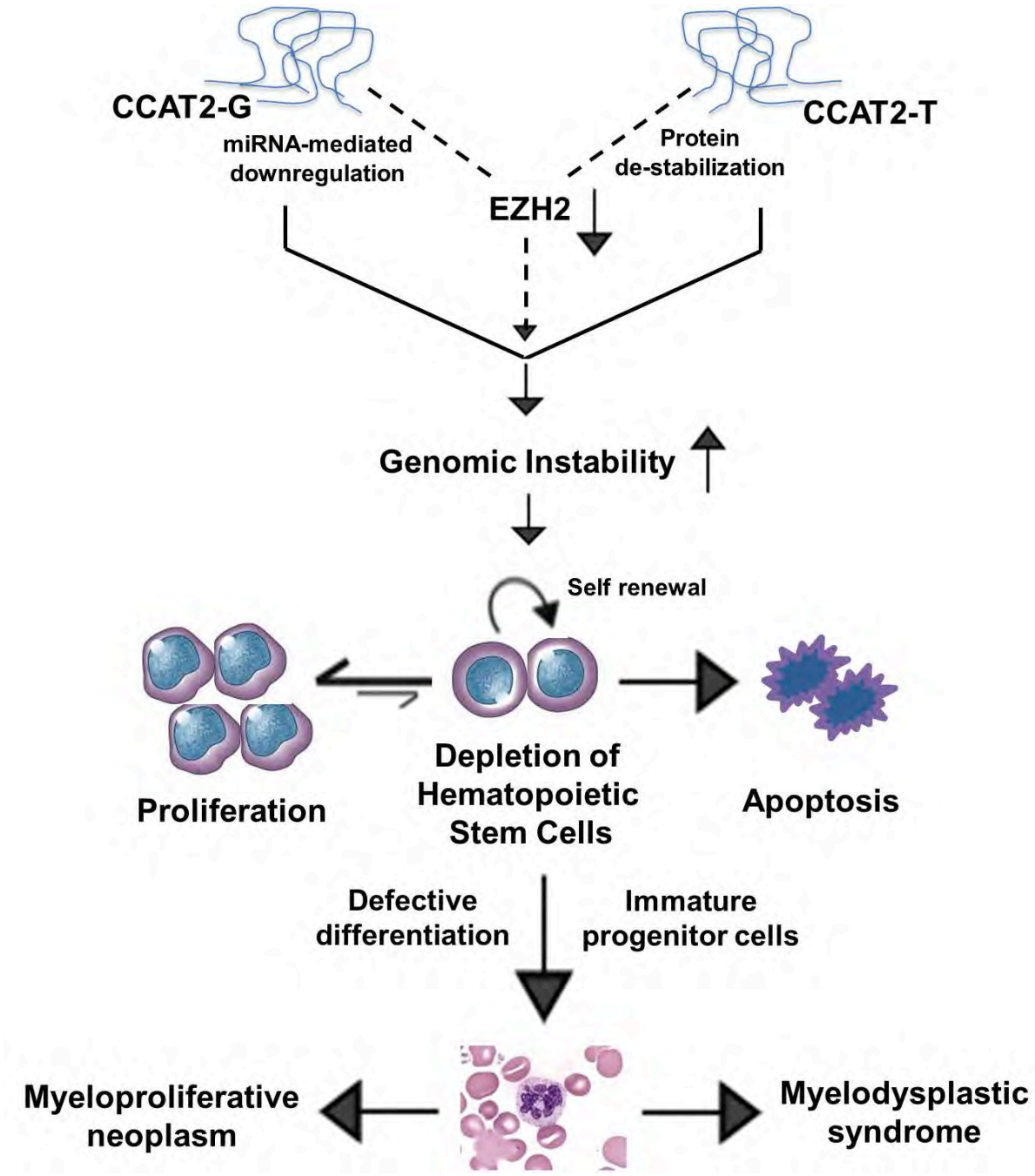


Figure 33: Graphical model of initiation of MDS and MDS/MPN by CCAT2

Conclusion

The discovery of functional non-protein-coding RNAs (ncRNAs, that do not codify for proteins) in the last decade has initiated a shift in the paradigms of cancer biology and has profoundly influenced our understanding of molecular genetics. Long non-coding RNAs (lncRNAs) form the largest part of the mammalian non-coding transcriptome and control gene expression at various levels including chromatin modification, transcriptional and post-transcriptional processing. Although the underlying molecular mechanisms are not yet entirely understood, lncRNAs are implicated in initiation and progression of several cancers. Identification of novel lncRNAs and their roles in regulation of protein-coding genes will provide vital information about the processes that drive tumorigenesis.

CCAT2 is a novel lncRNA that spans the highly conserved 8q24 region associated with increased risk for various cancers. *CCAT2* has been shown to play an important role in inducing chromosomal instability and supporting cell proliferation and cell cycle arrest. *CCAT2* lncRNA is overexpressed in colorectal, breast, and lung adenocarcinomas. However, a causal role of *CCAT2* in initiation of tumorigenesis and the importance of G/T SNP in *CCAT2*-induced phenotype still remains to be resolved. In this study, we generated transgenic mice for each *CCAT2* allele using random integration approach to elucidate the role of *CCAT2* and its specific alleles (G/T) in regulation of cellular processes that drive spontaneous tumorigenesis.

In this study, we identified that *CCAT2* plays an important role in regulation of normal hematopoiesis. Constitutive *in vivo* overexpression of each *CCAT2* transcript in the mice resulted in spontaneous induction of widespread pancytopenias with splenomegaly and hepatomegaly. CCAT2(G/T) BM biopsies displayed severe myeloid or erythroid hyperplasia, and dysplastic megakaryocytic proliferation, along with enhanced proliferation and excessive apoptosis. Interestingly, we identified two distinct phenotypes in CCAT2(G/T) mice with equal prevalence of MDS or mixed MDS/MPN. This suggests that *CCAT2* overexpression might affect regulation of hematopoietic stem cells, disturbing their self-renewal or maturation capacity, and subsequently resulting in BM failure. Percentage of HSPCs was significantly reduced in BM of MDS mice, with increased presence of immature erythroid blasts and granulocyte-macrophage progenitors suggesting a block in differentiation. HSPCs of CCAT2(G/T) mice also showed increased frequency of cytogenetic aberrations, including breaks and chromosomal fusions. However, these mice don't develop sAML, suggesting *CCAT2* is critical in initiation of MDS. Microarray expression profiling of CCAT2(G/T) HSPCs revealed enrichment of pathways associated with epigenetic regulation, chromosomal instability and cell cycle regulation. We further identified significantly higher *CCAT2* expression in the MDS patients as compared to healthy volunteers. Patients with sAML had significantly lower expression of *CCAT2* as compared to patients with only MDS.

Deciphering the role of *CCAT2* in spontaneously induced myelodysplasia and cytopenias will help us further characterize the poorly understood MDS/MPN phenotype. *CCAT2* mice can serve as a robust model for studying initiation of de

novo MDS/MPN that does not progress to secondary AML, and as a pre-clinical model for evaluation of new therapies for MDS. It has high translational potential as *CCAT2* can be developed into a diagnostic and prognostic marker, as well as a novel intervention target for MDS therapy.

Future directions

This project provided basis for a number of future studies. Detailed understanding of function and regulation of CCAT2(G/T) mice will help understand the factors involved in initiation of MDS, and thus provide better options for therapeutic intervention.

Further investigation on mechanism of CCAT2 in driving cellular process important in initiation of MDS vs MDS/MPN and its progression to AML

In order to identify the genetic regulators of MDS and mixed MDS/MPN in CCAT2 mice, an unbiased Affimetrix cDNA microarray and antisense oligomer-RNA-protein pull-down experiments can be performed on the BM cells of CCAT2(G/T) in comparison to WT mice, to obtain a list of differentially expressed genes when *CCAT2* is overexpressed. This will provide more information on the most impactful downstream genes and pathways regulated by CCAT2. Additionally, comparison between MDS-like and mixed MDS/MPN-like mice will yield useful information on the genetic differences in these two indications. This can also be useful to identify novel therapeutic targets specific for each disease. Finally, comparison between CCAT2-G and CCAT2-T mice will also provide important information on the functional significance of the SNP in predisposition to MDS and mixed MDS/MPN.

CCAT2 as a prognostic biomarker for MDS and its progression to AML

We report that CCAT2 expression levels were significantly upregulated specifically in MDS patients that do not progress to sAML. This could be very important in order to identify patients that are susceptible to AML progression. For this purpose, *CCAT2*

expression levels will need to be determined in a larger cohort of MDS and mixed MDS/MPN patients, spanning multiple institutions in several countries. Additionally, Sanger sequencing can be performed to determine the status of rs6983267 SNP at the genomic and RNA level in these patients to further understand the significance of the SNP in MDS and MDS/MPN.

CCAT2 as a potential therapeutic target in MDS and mixed MDS/MPN

Our study reported that ectopic overexpression of CCAT2 in vivo leads to spontaneous induction of MDS and mixed MDS/MPN-like diseases. Our study provides strong evidence implicating the importance of CCAT2 in these diseases, providing a basis for its development into a potential therapeutic option for treatment of MDS and/or MDS/MPN. In vivo treatment of these mice with LNA (Locked nucleic acid)-based anti-sense CCAT2 oligomers can be performed to determine if targeting CCAT2 can help in reversal of the MDS or MDS/MPN phenotype. An important therapeutic end point of this experiment would be to rescue the self-renewal capacity of HSCs and promote their normal maturation into adult hematopoietic stem cells. Additionally, CCAT2-anit-sense-LNAs can also be administered in conjunction with lenalidomide or azacitidine to see if combination therapy can afford improved therapeutic efficacy.

Bibliography

Ambros, V. (2004). The functions of animal microRNAs. Nature *431*, 350-355.

Asangani, I.A., Rasheed, S.A., Nikolova, D.A., Leupold, J.H., Colburn, N.H., Post, S., and Allgayer, H. (2008). MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdcd4 and stimulates invasion, intravasation and metastasis in colorectal cancer. Oncogene *27*, 2128-2136.

Bartel, D.P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. Cell *116*, 281-297.

Bejar, R., and Steensma, D.P. (2014). Recent developments in myelodysplastic syndromes. Blood *124*, 2793-2803.

Bernard, D., Prasanth, K.V., Tripathi, V., Colasse, S., Nakamura, T., Xuan, Z., Zhang, M.Q., Sedel, F., Jourdren, L., Coulpier, F., Triller, A., Spector, D. L., and Bessis, A. (2010). A long nuclear-retained non-coding RNA regulates synaptogenesis by modulating gene expression. The EMBO journal *29*, 3082-3093.

Bojovic, B., and Crowe, D.L. (2011). Telomere dysfunction promotes metastasis in a TERC null mouse model of head and neck cancer. Molecular cancer research : MCR *9*, 901-913.

Calin, G.A., Cimmino, A., Fabbri, M., Ferracin, M., Wojcik, S.E., Shimizu, M., Taccioli, C., Zanesi, N., Garzon, R., Aqeilan, R.I., Alder, H., Volinia, S., Rassenti, L., Liu, X., Liu, C.G., Kipps, T. J., Negrini, M., Croce, C.M. (2008). MiR-15a and miR-16-1 cluster functions in human leukemia. Proc Natl Acad Sci U S A *105*, 5166-5171.

Calin, G.A., and Croce, C.M. (2006). MicroRNA signatures in human cancers. Nat Rev Cancer *6*, 857-866.

Calin, G.A., Liu, C.G., Ferracin, M., Hyslop, T., Spizzo, R., Sevignani, C., Fabbri, M., Cimmino, A., Lee, E.J., Wojcik, S.E., Shimizu, M., Tili, E., Rossi, S., Taccioli, C., Pichiorri, F., Liu, X., Zupo, S., Herlea, V., Gramantieri, L., Lanza, G., Alder, H., Rassenti, L., Volinia, S., Schmittgen, T.D., Kipps, T.J., Negrini,

M., and Croce, C.M. (2007). Ultraconserved regions encoding ncRNAs are altered in human leukemias and carcinomas. Cancer Cell *12*, 215-229.

Calin, G.A., Sevignani, C., Dumitru, C.D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., and Croce, C.M. (2004). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A *101*, 2999-3004.

Carrieri, C., Cimatti, L., Biagioli, M., Beugnet, A., Zucchelli, S., Fedele, S., Pesce, E., Ferrer, I., Collavin, L., Santoro, C., Forrest, A.R., Carninci, P., Biffo, S., Stupka, E., and Gustincich, S. (2012). Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. Nature *491*, 454-457.

Carthew, R.W., and Sontheimer, E.J. (2009). Origins and Mechanisms of miRNAs and siRNAs. Cell *136*, 642-655.

Chamuleau, M.E., Westers, T.M., van Dreunen, L., Groenland, J., Zevenbergen, A., Eeltink, C.M., Ossenkoppele, G.J., and van de Loosdrecht, A.A. (2009). Immune mediated autologous cytotoxicity against hematopoietic precursor cells in patients with myelodysplastic syndrome. Haematologica *94*, 496-506.

Chen, L.L., and Carmichael, G.G. (2009). Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. Molecular cell *35*, 467-478.

Cheng, Y., Jutooru, I., Chadalapaka, G., Corton, J.C., and Safe, S. (2015). The long non-coding RNA HOTTIP enhances pancreatic cancer cell proliferation, survival and migration. Oncotarget *6*, 10840-10852.

Cortez, M.A., Bueso-Ramos, C., Ferdin, J., Lopez-Berestein, G., Sood, A.K., and Calin, G.A. (2011). MicroRNAs in body fluids--the mix of hormones and biomarkers. Nat Rev Clin Oncol *8*, 467-477.

Costinean, S., Zanesi, N., Pekarsky, Y., Tili, E., Volinia, S., Heerema, N., and Croce, C.M. (2006). Pre-B cell proliferation and lymphoblastic leukemia/high-grade lymphoma in E(mu)-miR155 transgenic mice. Proc Natl Acad Sci U S A *103*, 7024-7029.

Croce, C.M. (2009). Causes and consequences of microRNA dysregulation in cancer. Nat Rev Genet *10*, 704-714.

Dang, X., Ma, A., Yang, L., Hu, H., Zhu, B., Shang, D., Chen, T., and Luo, Y. (2012). MicroRNA-26a regulates tumorigenic properties of EZH2 in human lung carcinoma cells. Cancer genetics *205*, 113-123.

Di Gesualdo, F., Capaccioli, S., and Lulli, M. (2014). A pathophysiological view of the long non-coding RNA world. Oncotarget *5*, 10976-10996.

DiNardo, C.D., Daver, N., Jain, N., Pemmaraju, N., Bueso-Ramos, C., Yin, C.C., Pierce, S., Jabbour, E., Cortes, J.E., Kantarjian, H.M., Garcia-Manero, G., Verstovsek, S. (2014). Myelodysplastic/myeloproliferative neoplasms, unclassifiable (MDS/MPN, U): natural history and clinical outcome by treatment strategy. Leukemia *28*, 958-961.

Du, Y., Kong, G., You, X., Zhang, S., Zhang, T., Gao, Y., Ye, L., and Zhang, X. (2012). Elevation of highly up-regulated in liver cancer (HULC) by hepatitis B virus X protein promotes hepatoma cell proliferation via down-regulating p18. J Biol Chem *287*, 26302-26311.

Eiring, A.M., Harb, J.G., Neviani, P., Garton, C., Oaks, J.J., Spizzo, R., Liu, S., Schwind, S., Santhanam, R., Hickey, C.J., Becker, H., Chandler, J.C., Andino, R., Cortes, J., Hokland, P., Huettner, C.S., Bhatia, R., Roy, D.C., Liebhaber, S.A., Caligiuri, M.A., Marcucci, G., Garzon, R., Croce, C.M., Calin, G.A., and Perrotti, D. (2010). miR-328 functions as an RNA decoy to modulate hnRNP E2 regulation of mRNA translation in leukemic blasts. Cell *140*, 652-665.

Epperson, D.E., Nakamura, R., Sauntharajah, Y., Melenhorst, J., and Barrett, A.J. (2001). Oligoclonal T cell expansion in myelodysplastic syndrome: evidence for an autoimmune process. Leukemia research *25*, 1075-1083.

Ernst, T., Chase, A.J., Score, J., Hidalgo-Curtis, C.E., Bryant, C., Jones, A.V., Waghorn, K., Zoi, K., Ross, F.M., Reiter, A., Hochhaus, A., Drexler, H.G., Duncombe, A., Cervantes, F., Oscier, D., Boultwood, J., Grand, F.H., and Cross, N.C. (2010). Inactivating mutations of the histone methyltransferase gene EZH2 in myeloid disorders. Nat Genet *42*, 722-726.

Esquela-Kerscher, A., and Slack, F.J. (2006). Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer *6*, 259-269.

Ghoussaini, M., Song, H., Koessler, T., Al Olama, A.A., Kote-Jarai, Z., Driver, K.E., Pooley, K.A., Ramus, S.J., Kjaer, S.K., Hogdall, E., UK Genetic Prostate Cancer Study Collaborators/British Association of Urological Surgeons' Section of Oncology; UK ProtecT Study Collaborators. (2008). Multiple loci with different cancer specificities within the 8q24 gene desert. Journal of the National Cancer Institute *100*, 962-966.

Gonzalez, M.E., DuPrie, M.L., Krueger, H., Merajver, S.D., Ventura, A.C., Toy, K.A., and Kleer, C.G. (2011). Histone methyltransferase EZH2 induces Akt-dependent genomic instability and BRCA1 inhibition in breast cancer. Cancer Res *71*, 2360-2370.

Gruber, S.B., Moreno, V., Rozek, L.S., Rennerts, H.S., Lejbkowicz, F., Bonner, J.D., Greenson, J.K., Giordano, T.J., Fearson, E.R., and Rennert, G. (2007). Genetic variation in 8q24 associated with risk of colorectal cancer. Cancer biology & therapy *6*, 1143-1147.

Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L., Wang, Y., Brzoska, P., Kong, B., Li, R., West, R.B., van de Vijver, M.J., Sukumar, S., Chang, H.Y. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature *464*, 1071-1076.

Haiman, C.A., Le Marchand, L., Yamamato, J., Stram, D.O., Sheng, X., Kolonel, L.N., Wu, A.H., Reich, D., and Henderson, B.E. (2007). A common genetic risk factor for colorectal and prostate cancer. Nat Genet *39*, 954-956.

Herrera-Merchan, A., Arranz, L., Ligos, J.M., de Molina, A., Dominguez, O., and Gonzalez, S. (2012). Ectopic expression of the histone methyltransferase Ezh2 in haematopoietic stem cells causes myeloproliferative disease. Nature communications *3*, 623.

Hung, T., Wang, Y., Lin, M.F., Koegel, A.K., Kotake, Y., Grant, G.D., Horlings, H.M., Shah, N., Umbricht, C., Wang, P., Wang, Y., Kong, B., Langerød, A., Børresen-Dale, A.L., Kim, S.K., van de Vijver, M., Sukumar, S., Whitfield,

M.L., Kellis, M., Xiong, Y., Wong, D.J., and Chang, H.Y. (2011). Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. Nat Genet *43*, 621-629.

Hwang, H.W., Wentzel, E.A., and Mendell, J.T. (2007). A hexanucleotide element directs microRNA nuclear import. Science *315*, 97-100.

Iacobucci, I., Sazzini, M., Garagnani, P., Ferrari, A., Boattini, A., Lonetti, A., Papayannidis, C., Mantovani, V., Marasco, E., Ottaviani, E., Soverini, S., Girelli, D., Luiselli, D.,Vignetti, M., Baccarani, M., and Martinelli, G. (2011). A polymorphism in the chromosome 9p21 ANRIL locus is associated to Philadelphia positive acute lymphoblastic leukemia. Leukemia research *35*, 1052-1059.

Ji, Q., Zhang, L., Liu, X., Zhou, L., Wang, W., Han, Z., Sui, H., Tang, Y., Wang, Y., Liu, N., Ren, J., Hou, F., and Li, Q. (2014). Long non-coding RNA MALAT1 promotes tumour growth and metastasis in colorectal cancer through binding to SFPQ and releasing oncogene PTBP2 from SFPQ/PTBP2 complex. Br J Cancer *111*, 736-748.

Jia, L., Landan, G., Pomerantz, M., Jaschek, R., Herman, P., Reich, D., Yan, C., Khalid, O., Kantoff, P., Oh, W., Manak, J.R., Berman, B.P., Henderson, B.E., Frenkel, B., Haiman, C.A., Freedman, M., Tanay, A., and Coetzee, G.A. (2009). Functional enhancers at the gene-poor 8q24 cancer-linked locus. PLoS genetics *5*, e1000597.

Jiang, Y., Li, Y., Fang, S., Jiang, B., Qin, C., Xie, P., Zhou, G., and Li, G. (2014). The role of MALAT1 correlates with HPV in cervical cancer. Oncology letters *7*, 2135-2141.

Johnsson, P., Ackley, A., Vidarsdottir, L., Lui, W.O., Corcoran, M., Grander, D., and Morris, K.V. (2013). A pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human cells. Nature structural & molecular biology *20*, 440-446.

Jones, A.M., Beggs, A.D., Carvajal-Carmona, L., Farrington, S., Tenesa, A., Walker, M., Howarth, K., Ballereau, S., Hodgson, S.V., Zauber, A., Bertagnolli, M., Midgley, R.,Campbell, H., Kerr, D., Dunlop, M.G., and Tomlinson, I.P.

(2012). TERC polymorphisms are associated both with susceptibility to colorectal cancer and with longer telomeres. Gut *61*, 248-254.

Karreth, F.A., Tay, Y., Perna, D., Ala, U., Tan, S.M., Rust, A.G., DeNicola, G., Webster, K.A., Weiss, D., Perez-Mancera, P.A., Krauthammer, M., Halaban, R., Provero, P., Adams, D.J., Tuveson, D.A., and Pandolfi, P.P. (2011). In vivo identification of tumor- suppressive PTEN ceRNAs in an oncogenic BRAF-induced mouse model of melanoma. Cell *147*, 382-395.

Kawakami, T., Okamoto, K., Ogawa, O., and Okada, Y. (2004). XIST unmethylated DNA fragments in male-derived plasma as a tumour marker for testicular cancer. Lancet *363*, 40-42.

Klein, U., Lia, M., Crespo, M., Siegel, R., Shen, Q., Mo, T., Ambesi-Impiombato, A., Califano, A., Migliazza, A., Bhagat, G., and Dalla-Favera, R. (2010). The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. Cancer Cell *17*, 28-40.

Koerner, M.V., Pauler, F.M., Huang, R., and Barlow, D.P. (2009). The function of non-coding RNAs in genomic imprinting. Development *136*, 1771-1783.

Krystal, G.W., Armstrong, B.C., and Battey, J.F. (1990). N-myc mRNA forms an RNA-RNA duplex with endogenous antisense transcripts. Molecular and cellular biology *10*, 4180-4191.

Laner, T., Schulz, W.A., Engers, R., Muller, M., and Florl, A.R. (2005). Hypomethylation of the XIST gene promoter in prostate cancer. Oncology research *15*, 257-264.

Lee, J.T. (2012). Epigenetic regulation by long noncoding RNAs. Science *338*, 1435-1439.

Li, Z., Zhao, X., Zhou, Y., Liu, Y., Zhou, Q., Ye, H., Wang, Y., Zeng, J., Song, Y., Gao, W., Zheng, S., Zhuang, B., Chen, H., Li, W., Li, H., Li, H., Fu, Z., and Chen, R. (2015). The long non-coding RNA HOTTIP promotes progression and gemcitabine resistance by regulating HOXA13 in pancreatic cancer. J Transl Med *13*, 84.

Ling, H., Spizzo, R., Atlasi, Y., Nicoloso, M., Shimizu, M., Redis, R.S., Nishida, N., Gafa, R., Song, J., Guo, Z., Mimori, K., Mori, M., Sieuwerts, A.M., Martens,

J.W., Tomlinson, I., Negrini, M., Berindan-Neagoe, I., Foekens, J.A., Hamilton, S.R., Lanza, G., Kopetz, S., Fodde, R., and Calin, G.A. (2013). CCAT2, a novel noncoding RNA mapping to 8q24, underlies metastatic progression and chromosomal instability in colon cancer. Genome Res *23*, 1446-1461.

Liu, J.H., Chen, G., Dang, Y.W., Li, C.J., and Luo, D.Z. (2014). Expression and prognostic significance of lncRNA MALAT1 in pancreatic cancer tissues. Asian Pacific journal of cancer prevention : APJCP *15*, 2971-2977.

Louro, R., Smirnova, A.S., and Verjovski-Almeida, S. (2009). Long intronic noncoding RNA transcription: expression noise or expression choice? Genomics *93*, 291-298.

Lu, J., He, M.L., Wang, L., Chen, Y., Liu, X., Dong, Q., Chen, Y.C., Peng, Y., Yao, K.T., Kung, H.F., and Li, X.P. (2011). MiR-26a inhibits cell growth and tumorigenesis of nasopharyngeal carcinoma through repression of EZH2. Cancer Res *71*, 225-233.

Mailloux, A.W., and Epling-Burnette, P.K. (2013). Effector memory regulatory T-cell expansion marks a pivotal point of immune escape in myelodysplastic syndromes. Oncoimmunology *2*, e22654.

Mariner, P.D., Walters, R.D., Espinoza, C.A., Drullinger, L.F., Wagner, S.D., Kugel, J.F., and Goodrich, J.A. (2008). Human Alu RNA is a modular transacting repressor of mRNA transcription during heat shock. Molecular cell *29*, 499-509.

Medina, P.P., Nolde, M., and Slack, F.J. (2010). OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. Nature *467*, 86-90.

Melo, C.A., Drost, J., Wijchers, P.J., van de Werken, H., de Wit, E., Oude Vrielink, J.A., Elkon, R., Melo, S.A., Leveille, N., Kalluri, R., de Laat, W., and Agami, R. (2013). eRNAs are required for p53-dependent enhancer activity and gene transcription. Molecular cell *49*, 524-535.

Mercer, T.R., Dinger, M.E., and Mattick, J.S. (2009). Long non-coding RNAs: insights into functions. Nat Rev Genet *10*, 155-159.

Morlando, M., Ballarino, M., and Fatica, A. (2015). Long Non-Coding RNAs: New Players in Hematopoiesis and Leukemia. Frontiers in medicine *2*, 23.

Mourtada-Maarabouni, M., Pickard, M.R., Hedge, V.L., Farzaneh, F., and Williams, G.T. (2009). GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. Oncogene *28*, 195-208.

Muto, T., Sashida, G., Oshima, M., Wendt, G.R., Mochizuki-Kashio, M., Nagata, Y., Sanada, M., Miyagi, S., Saraya, A., Kamio, A., Nagae, G., Nakaseko, C., Yokote, K., Shimoda, K., Koseki, H., Suzuki, Y., Sugano, S., Aburatani, H., Ogawa, S., and Iwama, A. (2013). Concurrent loss of Ezh2 and Tet2 cooperates in the pathogenesis of myelodysplastic disorders. The Journal of experimental medicine *210*, 2627-2639.

Nie, F.Q., Sun, M., Yang, J.S., Xie, M., Xu, T.P., Xia, R., Liu, Y.W., Liu, X.H., Zhang, E.B., Lu, K.H., and Shu, Y.Q. (2015). Long noncoding RNA ANRIL promotes non-small cell lung cancer cell proliferation and inhibits apoptosis by silencing KLF2 and P21 expression. Molecular cancer therapeutics *14*, 268-277.

Niinuma, T., Suzuki, H., Nojima, M., Nosho, K., Yamamoto, H., Takamaru, H., Yamamoto, E., Maruyama, R., Nobuoka, T., Miyazaki, Y., Hosokawa, M., Hasegawa, T., Tokino, T., Hirata, K., Imai, K., Toyota, M.\, and Shinomura, Y. (2012). Upregulation of miR-196a and HOTAIR drive malignant character in gastrointestinal stromal tumors. Cancer Res *72*, 1126-1136.

Oliva, J., Bardag-Gorce, F., French, B.A., Li, J., and French, S.W. (2009). The regulation of non-coding RNA expression in the liver of mice fed DDC. Experimental and molecular pathology *87*, 12-19.

Orom, U.A., Derrien, T., Beringer, M., Gumireddy, K., Gardini, A., Bussotti, G., Lai, F., Zytnicki, M., Notredame, C., Huang, Q., Guigo, R., and Shiekhattar, R. (2010). Long noncoding RNAs with enhancer-like function in human cells. Cell *143*, 46-58.

Panzitt, K., Tschernatsch, M.M., Guelly, C., Moustafa, T., Stradner, M., Strohmaier, H.M., Buck, C.R., Denk, H., Schroeder, R., Trauner, M., *and* Zatloukal, K. (2007). Characterization of HULC, a novel gene with striking up-regulation in

hepatocellular carcinoma, as noncoding RNA. Gastroenterology *132*, 330-342.

Paulsson, K., and Johansson, B. (2007). Trisomy 8 as the sole chromosomal aberration in acute myeloid leukemia and myelodysplastic syndromes. Pathologie-biologie *55*, 37-48.

Peng, W., Gao, W., and Feng, J. (2014). Long noncoding RNA HULC is a novel biomarker of poor prognosis in patients with pancreatic cancer. Med Oncol *31*, 346.

Pickard, M.R., Mourtada-Maarabouni, M., and Williams, G.T. (2013). Long non-coding RNA GAS5 regulates apoptosis in prostate cancer cell lines. Biochim Biophys Acta *1832*, 1613-1623.

Place, R.F., Li, L.C., Pookot, D., Noonan, E.J., and Dahiya, R. (2008). MicroRNA-373 induces expression of genes with complementary promoter sequences. Proc Natl Acad Sci U S A *105*, 1608-1613.

Plath, K., Mlynarczyk-Evans, S., Nusinow, D.A., and Panning, B. (2002). Xist RNA and the mechanism of X chromosome inactivation. Annual review of genetics *36*, 233-278.

Pomerantz, M.M., Ahmadiyeh, N., Jia, L., Herman, P., Verzi, M.P., Doddapaneni, H., Beckwith, C.A., Chan, J.A., Hills, A., Davis, M., Manak, J.R., Shivdasani, R., Coetzee, G.A., and Freedman, M.L. (2009). The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. Nat Genet *41*, 882-884.

Qiao, H.P., Gao, W.S., Huo, J.X., and Yang, Z.S. (2013). Long non-coding RNA GAS5 functions as a tumor suppressor in renal cell carcinoma. Asian Pacific journal of cancer prevention : APJCP *14*, 1077-1082.

Qiu, M., Xu, Y., Yang, X., Wang, J., Hu, J., Xu, L., and Yin, R. (2014). CCAT2 is a lung adenocarcinoma-specific long non-coding RNA and promotes invasion of non-small cell lung cancer. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine *35*, 5375-5380.

Quagliata, L., Matter, M.S., Piscuoglio, S., Arabi, L., Ruiz, C., Procino, A., Kovac, M., Moretti, F., Makowska, Z., Boldanova, T., Diederichs, S., Cillo, C., and

Terracciano, L.M. (2014). Long noncoding RNA HOTTIP/HOXA13 expression is associated with disease progression and predicts outcome in hepatocellular carcinoma patients. Hepatology *59*, 911-923.

Redis, R.S., Sieuwerts, A.M., Look, M.P., Tudoran, O., Ivan, C., Spizzo, R., Zhang, X., de Weerd, V., Shimizu, M., Ling, H., Martens, J.W., Foekens, J.A., Berindan-Neagoe, I., and Calin, G.A. (2013). CCAT2, a novel long non-coding RNA in breast cancer: expression study and clinical correlations. Oncotarget *4*, 1748-1762.

Rinn, J.L., and Chang, H.Y. (2012). Genome regulation by long noncoding RNAs. Annual review of biochemistry *81*, 145-166.

Rosenfeld, C., and List, A. (2000). A hypothesis for the pathogenesis of myelodysplastic syndromes: implications for new therapies. Leukemia *14*, 2-8.

Sado, T., Hoki, Y., and Sasaki, H. (2005). Tsix silences Xist through modification of chromatin structure. Dev Cell *9*, 159-165.

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P.P. (2011). A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell *146*, 353-358.

Sashida, G., Harada, H., Matsui, H., Oshima, M., Yui, M., Harada, Y., Tanaka, S., Mochizuki-Kashio, M., Wang, C., Saraya, A., Inaba, T., Koseki, H., Huang, G., Kitamura, T., and Iwama, A. (2014). Ezh2 loss promotes development of myelodysplastic syndrome but attenuates its predisposition to leukaemic transformation. Nature communications *5*, 4177.

Shah, M. Y., Calin, G. A. (2014). Chapter 48: MicroRNA and cancer therapeutics. In: Principles of Molecular Diagnostics and Personalized Cancer Medicine, Ed(s) Tan D, Lynch H. Lippincott Williams and Wilkins: New York, NY.

Shah, M. Y., Calin, G. A. (2013). Chapter 1290: Regulatory RNA. In: Encyclopedia of genetics, 2nd ed. Ed(s) Maloy S, Hughes K. Elsevier Churchill Livingston: Philadelphia, PA.

Slavoff, S.A., Mitchell, A.J., Schwaid, A.G., Cabili, M.N., Ma, J., Levin, J.Z., Karger, A.D., Budnik, B.A., Rinn, J.L., and Saghatelian, A. (2013). Peptidomic

discovery of short open reading frame-encoded peptides in human cells. Nature chemical biology *9*, 59-64.

Sotelo, J., Esposito, D., Duhagon, M.A., Banfield, K., Mehalko, J., Liao, H., Stephens, R.M., Harris, T.J., Munroe, D.J., and Wu, X. (2010). Long-range enhancers on 8q24 regulate c-Myc. Proc Natl Acad Sci U S A *107*, 3001-3005.

Spizzo, R., Rushworth, D., Guerrero, M., and Calin, G.A. (2009). RNA inhibition, microRNAs, and new therapeutic agents for cancer treatment. Clin Lymphoma Myeloma *9 Suppl 3*, S313-318.

Sun, M., Xia, R., Jin, F., Xu, T., Liu, Z., De, W., and Liu, X. (2014). Downregulated long noncoding RNA MEG3 is associated with poor prognosis and promotes cell proliferation in gastric cancer. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine *35*, 1065-1073.

Szymanski, M., and Barciszewski, J. (2002). Beyond the proteome: non-coding regulatory RNAs. Genome biology *3*, reviews0005.

Tanaka, S., Miyagi, S., Sashida, G., Chiba, T., Yuan, J., Mochizuki-Kashio, M., Suzuki, Y., Sugano, S., Nakaseko, C., Yokote, K., Koseki, H., and Iwama, A. (2012). Ezh2 augments leukemogenicity by reinforcing differentiation blockage in acute myeloid leukemia. Blood *120*, 1107-1117.

Tay, Y., Kats, L., Salmena, L., Weiss, D., Tan, S.M., Ala, U., Karreth, F., Poliseno, L., Provero, P., Di Cunto, F., Lieberman, J., Rigoutsos, I., and Pandolfi, P.P. (2011). Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. Cell *147*, 344-357.

Tefferi, A., and Vardiman, J.W. (2009). Myelodysplastic syndromes. N Engl J Med *361*, 1872-1885.

Thomson, T., and Lin, H. (2009). The biogenesis and function of PIWI proteins and piRNAs: progress and prospect. Annual review of cell and developmental biology *25*, 355-376.

Thrash-Bingham, C.A., and Tartof, K.D. (1999). aHIF: a natural antisense transcript overexpressed in human renal cancer and during hypoxia. Journal of the National Cancer Institute *91*, 143-151.

Tomlinson, I., Webb, E., Carvajal-Carmona, L., Broderick, P., Kemp, Z., Spain, S., Penegar, S., Chandler, I., Gorman, M., Wood, W., Sieber, O., Gray, R.,Thomas, H., Peto, J., Cazier, J.B., and Houlston, R. (2007). A genome-wide association scan of tag SNPs identifies a susceptibility variant for colorectal cancer at 8q24.21. Nat Genet *39*, 984-988.

Trang, P., Weidhaas, J.B., and Slack, F.J. (2008). MicroRNAs as potential cancer therapeutics. Oncogene *27 Suppl 2*, S52-57.

Tripathi, V., Ellis, J.D., Shen, Z., Song, D.Y., Pan, Q., Watt, A.T., Freier, S.M., Bennett, C.F., Sharma, A., Bubulya, P.A., Blencowe, B.J., Prasanth, S.G., and Prasanth, K.V. (2010). The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. Molecular cell *39*, 925-938.

Tsang, F.H., Au, S.L., Wei, L., Fan, D.N., Lee, J.M., Wong, C.C., Ng, I.O., and Wong, C.M. (2015). Long non-coding RNA HOTTIP is frequently up-regulated in hepatocellular carcinoma and is targeted by tumour suppressive miR-125b. Liver international : official journal of the International Association for the Study of the Liver *35*, 1597-1606.

Tuupanen, S., Turunen, M., Lehtonen, R., Hallikas, O., Vanharanta, S., Kivioja, T., Bjorklund, M., Wei, G., Yan, J., Niittymaki, I., Karhu, A., Taipale, J., and Aaltonen, L.A. (2009). The common colorectal cancer predisposition SNP rs6983267 at chromosome 8q24 confers potential to enhanced Wnt signaling. Nat Genet *41*, 885-890.

Vasudevan, S., Tong, Y., and Steitz, J.A. (2007). Switching from repression to activation: microRNAs can up-regulate translation. Science *318*, 1931-1934.

Vincent-Salomon, A., Ganem-Elbaz, C., Manie, E., Raynal, V., Sastre-Garau, X., Stoppa-Lyonnet, D., Stern, M.H., and Heard, E. (2007). X inactive-specific transcript RNA coating and genetic instability of the X chromosome in BRCA1 breast tumors. Cancer Res *67*, 5134-5140.

Visnovsky, J., Kudela, E., Farkasova, A., Balharek, T., Krkoska, M., and Danko, J. (2014). Amplification of TERT and TERC genes in cervical intraepithelial neoplasia and cervical cancer. Neuro endocrinology letters *35*, 518-522.

Vogelstein, B., and Kinzler, K.W. (2004). Cancer genes and the pathways they control. Nat Med *10*, 789-799.

Wang, C.Y., Hua, L., Yao, K.H., Chen, J.T., Zhang, J.J., and Hu, J.H. (2015a). Long non-coding RNA CCAT2 is up-regulated in gastric cancer and associated with poor prognosis. International journal of clinical and experimental pathology *8*, 779-785.

Wang, J., Liu, X., Wu, H., Ni, P., Gu, Z., Qiao, Y., Chen, N., Sun, F., and Fan, Q. (2010). CREB up-regulates long non-coding RNA, HULC expression through interaction with microRNA-372 in liver cancer. Nucleic Acids Res *38*, 5366-5383.

Wang, J., Qiu, M., Xu, Y., Li, M., Dong, G., Mao, Q., Yin, R., and Xu, L. (2015b). Long noncoding RNA CCAT2 correlates with smoking in esophageal squamous cell carcinoma. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine.

Wang, J., Wang, H., Zhang, Y., Zhen, N., Zhang, L., Qiao, Y., Weng, W., Liu, X., Ma, L., Xiao, W., Yu, W., Chu, Q., Pan, Q., and Sun, F. (2014). Mutual inhibition between YAP and SRSF1 maintains long non-coding RNA, Malat1-induced tumourigenesis in liver cancer. Cellular signalling *26*, 1048-1059.

Wang, K.C., and Chang, H.Y. (2011). Molecular mechanisms of long noncoding RNAs. Molecular cell *43*, 904-914.

Willingham, A.T., Orth, A.P., Batalov, S., Peters, E.C., Wen, B.G., Aza-Blanc, P., Hogenesch, J.B., and Schultz, P.G. (2005). A strategy for probing the function of noncoding RNAs finds a repressor of NFAT. Science *309*, 1570-1573.

Woll, P.S., Kjallquist, U., Chowdhury, O., Doolittle, H., Wedge, D.C., Thongjuea, S., Erlandsson, R., Ngara, M., Anderson, K., Deng, Q., Hellström-Lindberg, E., Linnarsson, S., and Jacobsen, S.E. (2014). Myelodysplastic syndromes are propagated by rare and distinct human cancer stem cells in vivo. Cancer Cell *25*, 794-808.

Wong, C.F., and Tellam, R.L. (2008). MicroRNA-26a targets the histone methyltransferase Enhancer of Zeste homolog 2 during myogenesis. J Biol Chem *283*, 9836-9843.

Wright, J.B., Brown, S.J., and Cole, M.D. (2010). Upregulation of c-MYC in cis through a large chromatin loop linked to a cancer risk-associated single-nucleotide polymorphism in colorectal cancer cells. Molecular and cellular biology *30*, 1411-1420.

Wu, X.S., Wang, X.A., Wu, W.G., Hu, Y.P., Li, M.L., Ding, Q., Weng, H., Shu, Y.J., Liu, T.Y., Jiang, L.*,* Cao Y, Bao RF, Mu JS, Tan ZJ, Tao F, Liu YB. (2014a). MALAT1 promotes the proliferation and metastasis of gallbladder cancer cells by activating the ERK/MAPK pathway. Cancer biology & therapy *15*, 806-814.

Wu, Z.H., Wang, X.L., Tang, H.M., Jiang, T., Chen, J., Lu, S., Qiu, G.Q., Peng, Z.H., and Yan, D.W. (2014b). Long non-coding RNA HOTAIR is a powerful predictor of metastasis and poor prognosis and is associated with epithelial-mesenchymal transition in colon cancer. Oncology reports *32*, 395-402.

Xia, Y., He, Z., Liu, B., Wang, P., and Chen, Y. (2015). Downregulation of Meg3 enhances cisplatin resistance of lung cancer cells through activation of the WNT/beta-catenin signaling pathway. Molecular medicine reports.

Xue, B., and He, L. (2014). An expanding universe of the non-coding genome in cancer biology. Carcinogenesis *35*, 1209-1216.

Yang, L., Lin, C., Jin, C., Yang, J.C., Tanasa, B., Li, W., Merkurjev, D., Ohgi, K.A., Meng, D., Zhang, J.*,* Evans, C.P., and Rosenfeld, M.G.. (2013). lncRNA-dependent mechanisms of androgen-receptor-regulated gene activation programs. Nature *500*, 598-602.

Yang, M.H., Hu, Z.Y., Xu, C., Xie, L.Y., Wang, X.Y., Chen, S.Y., and Li, Z.G. (2015). MALAT1 promotes colorectal cancer cell proliferation/migration/invasion via PRKA kinase anchor protein 9. Biochim Biophys Acta *1852*, 166-174.

Yildirim, E., Kirby, J.E., Brown, D.E., Mercier, F.E., Sadreyev, R.I., Scadden, D.T., and Lee, J.T. (2013). Xist RNA is a potent suppressor of hematologic cancer in mice. Cell *152*, 727-742.

Yin, D.D., Liu, Z.J., Zhang, E., Kong, R., Zhang, Z.H., and Guo, R.H. (2015). Decreased expression of long noncoding RNA MEG3 affects cell proliferation and predicts a poor prognosis in patients with colorectal cancer. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine.

Yoon, J.H., Abdelmohsen, K., Srikantan, S., Yang, X., Martindale, J.L., De, S., Huarte, M., Zhan, M., Becker, K.G., and Gorospe, M. (2012). LincRNA-p21 suppresses target mRNA translation. Molecular cell *47*, 648-655.

Zanke, B.W., Greenwood, C.M., Rangrej, J., Kustra, R., Tenesa, A., Farrington, S.M., Prendergast, J., Olschwang, S., Chiang, T., Crowdy, E., Thomas G, Gallinger, S., Hudson, T.J., and Dunlop, M.G. (2007). Genome-wide association scan identifies a colorectal cancer susceptibility locus on chromosome 8q24. Nat Genet *39*, 989-994.

Zhang, E.B., Kong, R., Yin, D.D., You, L.H., Sun, M., Han, L., Xu, T.P., Xia, R., Yang, J.S., De, W., and Chen, J.F. (2014). Long noncoding RNA ANRIL indicates a poor prognosis of gastric cancer and promotes tumor growth by epigenetically silencing of miR-99a/miR-449a. Oncotarget *5*, 2276-2292.

Zhang, H., Zhao, L., Wang, Y.X., Xi, M., Liu, S.L., and Luo, L.L. (2015a). Long non-coding RNA HOTTIP is correlated with progression and prognosis in tongue squamous cell carcinoma. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine.

Zhang, X., Xu, Y., He, C., Guo, X., Zhang, J., He, C., Zhang, L., Kong, M., Chen, B., and Zhu, C. (2015b). Elevated expression of CCAT2 is associated with poor prognosis in esophageal squamous cell carcinoma. Journal of surgical oncology *111*, 834-839.

Zhao, W., An, Y., Liang, Y., and Xie, X.W. (2014a). Role of HOTAIR long noncoding RNA in metastatic progression of lung cancer. European review for medical and pharmacological sciences *18*, 1930-1936.

Zhao, Y., Guo, Q., Chen, J., Hu, J., Wang, S., and Sun, Y. (2014b). Role of long non-coding RNA HULC in cell proliferation, apoptosis and tumor metastasis of

gastric cancer: a clinical and in vitro investigation. Oncology reports *31*, 358-364.

Vita

Maitri Y. Shah was born in Mumbai, India on September 12, 1986. She received her Bachelors in Pharmacy degree from Institute of Chemical Technoloy, India in 2008. After graduation, she moved to USA for pursuing her Masters degree in Molecular biology and Biotechnology at East Carolina University, NC in 2010. She joined University of Texas Health Science Center at Houston Graduate School for Biomedical Sciences in August 2010. In May 2011, she joined Dr. George Calin's laboratory, where she completed her thesis work.