The long noncoding RNA regulation at the MYC locus
Jian-Feng Xiang¹, Li Yang²,³ and Ling-Ling Chen¹,³

Aberrant expression of long noncoding RNAs (lncRNAs) has been linked to cancers. The MYC oncoprotein is a key contributor to the development of many human tumors. Recent studies have revealed that a number of lncRNAs originating from the human 8q24 locus previously known to corresponding to a ‘gene desert’ are transcribed and play important roles in MYC regulation. In this review, we highlight recent progress in how these lncRNAs participate in control of MYC levels in normal and tumor cells.

Addresses
¹ State Key Laboratory of Molecular Biology, Shanghai Key Laboratory of Molecular Andrology, Institute of Biochemistry and Cell Biology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China  
² Key Laboratory of Computational Biology, CAS-MPG Partner Institute for Computational Biology, CAS Center for Excellence in Brain Science, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China  
³ School of Life Science and Technology, ShanghaiTech University, Shanghai 200031, China

Corresponding author: Chen, Ling-Ling (linglingchen@sibcb.ac.cn)

Current Opinion in Genetics & Development 2015, 33:41–48

This review comes from a themed issue on Molecular and genetic bases of disease

Edited by Dan E Arking and Johanna M Rommens

For a complete overview see the Issue and the Editorial

Available online 7th August 2015

http://dx.doi.org/10.1016/j.gde.2015.07.001

0959-437X/© 2015 Elsevier Ltd. All rights reserved.

Introduction

Mammalian genomes encode thousands of long noncoding RNAs (lncRNAs) that impact a variety of important biological processes (reviewed in [1]). Although mechanisms are not yet completely clear, aberrant expression of lncRNAs has been observed in human cancers with distinct modes of action (reviewed in [2]). The MYC oncoprotein functions as a central hub of a cell, by acting as a sensor and effector of cellular information. MYC expression in normal cells is tightly controlled at multiple levels, but becomes dysregulated in many human cancers (reviewed in [3]). Thus, it is essential to keep MYC level under close surveillance in order to avoid deleterious oncogenic changes.

The transcription of the MYC locus is complex and is modulated at multiple levels, from the regulation of enhancers, promoters to transcription factors [4,5,6]. Also, the stability of MYC protein is highly controlled as well in cells through targeted degradation by the ubiquitin-proteasome system (reviewed in [3]). Importantly, recent studies have uncovered a number of lncRNAs transcribed from chromatin regions close to the MYC locus in different types of human cancers (Figure 1a) [7,8,9,10,11,12–14]. These newly identified lncRNAs were shown to actively participate in MYC regulation by a number of mechanisms [9,11,15], providing additional components to the complicated MYC regulation. In this review, we highlight these recent studies and our understanding of how lncRNAs are involved in the regulation of MYC levels under normal and disease conditions.

The chromatin organization at the MYC locus

The human MYC locus is located at the 8q24 region, a previously known ‘gene desert’ that exhibits a paucity of protein-coding genes (Figure 1a). This megabase-sized region of the gene desert around MYC contains a number of regulatory elements, including enhancers [16–20] and super-enhancers [6,9] (Figure 1a). Enhancers are DNA regulatory sequences that are capable of binding master transcription factors/mediators and often form long-range chromatin loops with their target genes to activate temporal-specific and tissue-specific gene expression, independent of their proximity or orientation to their target genes [21]. On the other hand, super-enhancers consist of large clusters of transcriptional enhancers and tend to be associated with genes that control and define cell identity [22,23].

The 8q24 region enhancers have been annotated in several human diseases, in particular, cancers. Such enhancers often form chromatin loops with the MYC oncogene and are located from several hundred to two thousand kilobases (kb) away (Figure 1b). While the MYC upstream enhancers are linked to prostate, breast and colorectal cancers and form loops with the MYC promoter in a tissue-specific manner in these cancers [5], the MYC downstream enhancers are linked to acute leukemia [18,20] and the development of normal facial morphogenesis [19]. In addition, the region upstream of MYC has been recently shown to contain tumor type specific super-enhancers in cancer cells, but not in their healthy counterparts [6]. Importantly, recent studies have revealed that some such enhancers can regulate MYC expression by recruiting transcriptional or epigenetic factors [18–20,24]. Clearly, a complete annotation of regulatory elements in 8q24 across different tissue types under physiological and pathological conditions is
important for the understanding of the transcriptional control of MYC.

It should be noted that the formation of these long-range chromatin loops with the MYC oncogene not only plays a role in MYC transcription regulation \([5,18,20,24**]\), but also allows the distally transcribed lncRNAs to be spatially localized to the MYC locus to exert their functions in MYC regulation \([9^*,10,11^*]\) (Figures 2 and 3).

LncRNA-mediated chromatin regulation at the MYC locus

A number of lncRNAs were recently reported from the 8q24 region in prostate cancer, colorectal cancer (CRC) patient samples and other cancers (Figure 1a). Among these, the CCAT1 locus, located 515 kb upstream of MYC, encodes at least two abundant lncRNAs, CCAT1-S (Colorectal Cancer Associated Transcript 1, CCAT1, short isoform) and CCAT1-L (CCAT1, long isoform). CCAT1-S is 2,600 nt in length, contains two exons \([8]\) and is largely located in the cytoplasm \([9^*,25]\). It is probably produced by 3’ end processing of CCAT1-L via an unknown mechanism \([9^*]\) (Figure 2). Importantly, CCAT1-S is a highly specific marker for transformation in the colon [8] and its up-regulation is evident in both pre-malignant conditions and through all disease stages in CRC \([26]\). In addition to CRC, the over-expression of both CCAT1-S/CCAT1-L has also been reported in gastric \([27]\) and gallbladder \([13]\) cancers (Figure 1a). CCAT1-S is also known as CARLo-5 (Cancer-Associated Region Long noncoding RNA) \([10]\). Knockdown of CARLo-5 decreased cell proliferation and suppressed cell transformation and tumor incidence \([10]\). These observations indicate that CCAT1-S (CARLo-5) plays a role in tumorigenesis, although the underlying mechanisms are still lacking.

CCAT1-L contains two exons that overlap with CCAT1-S at the 5’ end (Figure 2), but this lncRNA is mainly located in the nucleus. It is a polyadenylated lncRNA of 5200 nt in length \([9^*]\). CCAT1-L is also highly expressed in CRC primary tissues and is associated with increased CRC susceptibility \([9^*,10]\). While knockdown of CCAT1-L modestly reduced MYC transcription, over-expression of CCAT1-L by modulating its endogenous expression in CRC-derived cells by genome editing enhanced MYC expression and increased tumor formation in a mouse xenograft model \([9^*]\), suggesting a role of CCAT1-L in MYC regulation and tumor development.
**Figure 2**

CCAT1-L regulates chromatin looping at the MYC locus. The CCAT1 locus, which is located in the CRC-specific super-enhancer (green), produces two abundant lncRNAs in CRC, CCAT1-L and CCAT1-S. Note that the formation of chromatin loops among CCAT1-L, MYC-335, and MYC loci is mediated by the higher-order chromatin organizer CTCF. While CCAT1-S is exported to cytoplasm, CCAT1-L is exclusively localized in-cis and interacts with CTCF to achieve an enhanced super-enhancer–promoter (E-P) looping for MYC activation. See text for details.

CCAT1-L was found to act in cis to mediate chromatin looping between the MYC promoter and its enhancers. The CCAT1 locus forms multiple long-range chromatin interactions in the 8q24 region, including interactions with the MYC promoter and a well-characterized 8q24 enhancer located about 335 kb upstream of MYC (MYC-335) [9,10,16,17] (Figure 2). Interestingly, the interaction between the CCAT1 (CARLo-5) promoter and the MYC-335 enhancer also indicates that the expression of both CCAT1-S/CCAT1-L can be regulated through the long-range interaction between MYC-335 and the promoter of CCAT1 [10]. Importantly, the CCAT1 locus has been recently characterized as a super-enhancer that spans 150 kb in length in a CRC cell-specific manner [6] (Figure 2). Transcribed CCAT1-L from this super-enhancer in CRC cell lines, such as HCT116 and HT29 cells, exclusively accumulates at or near its sites of transcription [9]. The formation of chromatin loops between the CCAT1-L and MYC loci, mediated by the higher-order chromatin organizer CTCF [28], allows CCAT1-L to ‘coat’ the MYC oncogene locus and other looping sites in cis [9] (Figure 2). Further studies revealed that CCAT1-L specifically interacted with CTCF. Depletion of CCAT1-L led to reduced chromatin interaction frequencies and CTCF binding to these looping sites [9], suggesting that CCAT1-L could modulate CTCF association with chromatin at these looping regions in the MYC locus (Figure 2). The findings presented in this study are consistent with a recent report that CTCF contains an RNA binding region, distinct from its DNA-binding domain, that mediates its binding to a variety of RNAs [29]. However, how exactly CCAT1-L and CTCF work together to regulate the chromatin dynamics in the MYC locus, and to what extent CCAT1-L-regulated
looping cross talks with other aspects of MYC regulation, remain to be determined. In addition, it cannot be excluded that other protein factors may be involved in this CCAT1-L mediated MYC transcription regulation.

LncRNAs have been suggested to play a role in modulating nuclear architecture. For instance, the LncRNA NEAT1 (Nuclear Enriched Abundant Transcript 1) is required for the integrity of paraspeckles. These nuclear bodies contain specific proteins and RNA components that have been implicated in gene regulation (reviewed in [30]). Fire (Functional intergenic repeating RNA element) is involved in the topological organization of multichromosomal regions [31*]. The striking accumulation of CCAT1-L RNA at the MYC locus [9*] (Figure 2) implies that other protein components can be associated with CCAT1-L and additional multichromosomal regions might be regulated by CCAT1-L, as well. Thus, the detailed analyses of CCAT1-L RNA associated proteins and chromatin regions will provide new insights into MYC regulation.

Furthermore, the identification of CCAT1-L, transcribed from a super-enhancer [9*] suggests that some super-enhancers can be transcribed and their activities may be further regulated by RNA transcripts. Similar phenomena have been observed in gene activation mediated by enhancer RNAs (eRNAs). eRNAs are expressed from many enhancers and the resulting eRNAs are non-polyadenylated with generally very low-copy numbers [32]. Recent studies have shown that eRNAs play an enhancer-like function participating in transcriptional activation [33–36]. For example, the level of eRNA expression at neuronal enhancers positively correlated with the level of mRNA synthesis at nearby genes [32]. Depletion of some eRNAs resulted in repression of neighboring protein-coding genes in a cis regulatory manner [33]. In human breast cancer cells, 17 beta-oestradiol (E2)-induced eRNAs contributed to the E2-dependent global gene activation by stabilizing E2/ eRNAs induced enhancer–promoter looping through the interaction with the Cohesin complex [34]. As the 8q24 gene desert contains many enhancers that may act in different cancers [5,18,20], and as recent studies have uncovered pervasive transcripts produced from this region [10,11*], it will be of great interest to investigate whether other 8q24 transcripts act in a similar manner in MYC transcription regulation during tumorigenesis.

**LncRNA-mediated transcription regulation at the MYC locus**

One of the best-characterized enhancers in the MYC locus is MYC-335, the chromatin region located about 335 kb upstream of MYC at human 8q24, and this enhancer has been shown to play an important role in CRC development [16,17,24**]. MYC-335 is a highly conserved enhancer and is topologically located to the MYC oncogene promoter region [16,17,24**]. Like other enhancers that are known to control temporal-specific and tissue-specific gene expression [37], a transgenic mouse model of MYC-335-derived LacZ reporter gene revealed that MYC-335 could function as a tissue-specific enhancer in developing embryos [16]. Strikingly, mice lacking MYC-335 exhibited a modest reduction of myc transcripts and were resistant to intestinal tumors [24**], strongly suggesting that this enhancer is crucial for the development of colon tumors at least partially through the regulation of MYC. Moreover, this enhancer also contains a CRC risk G allele of rs6983267 [16,17], which shows copy number increase during CRC development [16]. Although no significant correlation between rs6983267 genotype and MYC expression was observed [16,17], this G risk allele variant has been shown to confer an enhanced binding affinity
to transcription factor 4 (TCF4), which plays a crucial role in the activation of the key CRC WNT pathway in colon cancers [16,17] (Figure 3). These findings thus support the view that both the MYC-335 enhancer and the non-coding risk variant within it can regulate CRC pathogenesis.

CCAT2 (Colorectal Cancer Associated Transcript 2) is a recently characterized lncRNA transcribed from the MYC-335 region encompassing the rs6983267 site (Figure 3). It is a 340 nt non-spliced transcript and is over-expressed in microsatellite-stable (MSS) CRC samples [11*], breast cancer [38] and lung adenocarcinoma [39] (Figure 1a). Furthermore, its expression is significantly higher in primary CRC tumors from patients with metastasis than in those without metastasis [11*]. Experimentally, over-expression of CCAT2 in CRC-derived cell lines increased subcutaneous tumor formation and promoted cell migration, while knockdown of it reduced cell invasion, suggesting that CCAT2 plays a role in promoting cancer growth and metastasis [11*].

Ling et al. [11*] also found that the expression pattern of MYC in MSS CRC samples is highly correlated with that of CCAT2. Knockdown of CCAT2 led to a reduced MYC expression and MYC target genes, including a number of miRNAs that are known to play roles in metastasis. Further analyses revealed that CCAT2 is largely localized to the nucleus in CRC-derived cell lines and in CRC patient tissues. At the molecular level, CCAT2 was suggested to bind to TCF4 and augment its transcriptional activity as shown by the activation of WNT signaling in reporter assays in CCAT2 over-expressing cells [11*] (Figure 3). Interestingly, CCAT2 expression appeared to respond to WNT, suggesting a positive feedback loop between CCAT2 and WNT signaling [11*].

The detailed mechanism of how CCAT2 regulates TCF4 is still largely unknown. It is well known that lncRNAs can regulate transcription by serving as ‘ligands’ for transcription factors (reviewed in [30]). Thus, one possibility is that the binding of CCAT2 RNA to TCF4 may allosterically affect the protein structure or modulate the association of TCF4 with its partners in the transcription complex. As a result, such modification may alter the transcription regulation of TCF4 target genes, leading to enhanced WNT and MYC activities in CRC. Furthermore, by measuring the amount of CCAT2 transcript produced from different alleles in CRC cell lines with a heterogeneous rs6983267 genotype, the G allele of rs6983267 appeared to produce more CCAT2 than the T allele [11*]. Therefore, another possibility is that the risk rs6983267 mutation may change the property of the final CCAT2 transcript, altering its structure, stability and binding capacity to TCF4 specifically in CRC. Interestingly, the expression of CARL0.5 is also correlated with this risk allele in CRC [10]. These observations have provided yet another possibility in linking this risk allele with higher CRC risk [16,17]. Finally, as CCAT2 is transcribed from the MYC-335 enhancer, it may function similarly to eRNAs or the super-enhancer transcribed CCAT1-L in activating gene expression [9*,33–36]. If this were the case, CCAT2 might promote enhancer–promoter looping at the MYC locus, resulting in an enhanced MYC transcription in CRC. Nevertheless, although the underlying mechanisms of CCAT2-mediated regulation at the MYC locus have not yet been fully defined, CCAT2 is a new 8q24 transcript involved in CRC pathogenesis [11*].

lncRNA in modulating MYC protein stability

Besides the above-mentioned MYC upstream region transcribed lncRNAs, PVT1 (Plasma-cytoma Variant Translocation 1) is transcribed from approximately 100–500 kb downstream of the MYC locus (Figure 1). In human cells, the PVT1 gene contains nine exons and produces multiple non-coding transcripts of between 2.7 and 3.3 kb in length by extensive alternative splicing [7,15]. Interestingly, recent studies have suggested that this locus also encodes circular RNAs, which are characterized by covalently closed loop structures without 5′–3′ polarity or polyadenylated tails [40*,41]. Nevertheless, while the genomic context and transcriptional activity of the PVT1 locus are conserved across species, the transcribed PVT1 RNA sequences share only very low sequence similarity (see review in [42]). Although this feature of PVT1 is consistent with the fact that many lncRNAs exhibit sequence divergence yet conserved function between species, it warrants a fuller characterization of PVT1 function.

Multiple studies have revealed that the higher expression of PVT1 was significantly associated with increased metastasis and worse prognosis in many cancers, such as hepatocellular carcinoma [43], colon [14], gastric [44], ovarian and breast cancers [7] (Figure 1a). Over-expression of PVT1 in cancers might be a consequence of the amplification and translocation of the 8q24 region, which contains MYC and its surrounding fragments ranging from several hundred kb to several Mb [7,15*]. In fact, PVT1 RNA and MYC protein expression are highly correlated, and in more than 98% of cases, there is co-increase in the copy-number of PVT1 and MYC in primary human cancer cells [15*]. These clinical observations have indicated that PVT1 might be an important player in cancer progression.

A recent study aimed at investigating whether low copy number gain of one, or more, of MYC and its adjacent genes on 8q24 could promote cancer in mouse models has uncovered a novel role of PVT1 in the regulation of MYC protein stability during tumor development [15*]. While gain of a single extra copy of the MYC gene or the region containing (PVT1, CCDC26 and GSDMC) genes was not
sufficient to induce tumors, single supernumerary gain of (MYC, PVT1, CCDC26 and GSDMC) genes promoted cancer. As CCDC26 and GSDMC transcripts could only be barely detected in mouse, these observations led to the hypothesis that PVT1 and MYC are more probably to contribute to the observed tumorigenesis. Further studies revealed that the alteration of PVT1 expression had no effect on myc mRNA level; however, the MYC protein levels did change accordingly in a positively correlated manner [15**] (Figure 4). Increase of PVT1 led to enhanced levels of MYC protein, which appeared to be a consequence of increased stability of the MYC protein. The PVT1-null HCT116 cell line exhibited significantly reduced levels of MYC protein, and retarded tumorigenic potency [15**]. Moreover, PVT1 RNA was found to interact with MYC in the nucleus and interfered with its phosphorylation at threonine 58 (Thr58) [15**], which is known to play a well-established role in promoting MYC protein degradation [48] (Figure 4). Together, these findings have suggested that PVT1 is an important regulator of tumorigenesis by controlling MYC protein stability through modulating Thr58 phosphorylation. It should be noted, however, that PVT1 might act independently of MYC as well. For example, in PVT1 over-expressed ovarian or breast cell lines, specific knockdown of PVT1 induced apoptosis that was not detected with only the depletion or knockdown of MYC [7].

Interestingly, emerging lines of evidence have shown that the interaction between an lncRNA and protein mutually affect each others function, resulting in additional layers of gene expression regulation for many cellular processes. On one hand, the interaction of lncRNA with protein can modulate the protein modification status, as seen in the case of the PVT1 mediating MYC stability. A similar observation was found in lnc-DC regulated STAT3 phosphorylation [49]. In this study, lnc-DC was shown to control human dendritic cell differentiation by activating transcription factor STAT3 through a direct binding of lnc-DC to STAT3. Such an lncRNA-protein interaction prevented STAT3 from binding to its downstream partner and promoted STAT3 phosphorylation at Tyrosine 705. On the other hand, the post-translational modification of a protein can affect the function of lncRNAs as well. For instance, methylation/demethylation of Polycomb 2 protein was found to modulate its interaction with different lncRNAs, either TUG1 or MALAT1, resulting in the coordinated gene expression program in distinct subnuclear architectural compartments in response to growth signals [50]. Apart from these striking observations, it will be of great interest to investigate how a single or a few amino acids modification on a protein might alter its binding capacity to a particular lncRNA.

**Perspectives**

Recent advances in genomic analyses and experimental systems have led to unprecedented new insights into the diverse and dynamic interplay between lncRNAs and MYC regulation. Despite such rapid progress in the functional study of 8q24 transcribed lncRNAs, some important questions remain to be addressed for MYC regulation at this locus due to the fact that MYC is a major oncogene in many cancers of distinct tissue origins. First, some 8q24 lncRNAs were shown to be regulated by MYC [51], WNT [11*], or the cancer-SNP-related enhancer [10], suggesting a positive feedback loop between aberrant tumor signals and lncRNA expression. It is thus important to further investigate whether the transcription of 8q24 lncRNAs is a consequence or a cause during tumor pathogenesis. Furthermore, as the MYC locus is involved in translocations in cancer [52], whether the expression of the co-gained lncRNAs could influence the cancer development with chromosomal rearrangements is unknown. Moreover, it has been recently reported that thousands of RNAs with new formats, such as circular RNAs [40*,41,53], are pervasively transcribed from mammalian genomes. As different types of tissues and cancers express different sets of lncRNAs from 8q24 (Figure 1a), a complete annotation and characterization of RNAs generated from this locus under physiological and pathologic conditions requires further study. Finally, it should be stressed that not all 8q24 region transcribed lncRNA may act in cis. PRNCR1 (Prostate Cancer Non-Coding RNA 1) is produced from 8q24 and is in strong association with susceptibility of prostate cancer [54]. A recent genome-wide ChIPR (Chromatin Isolation by RNA Purification) study with biotin-labelled DNA probes tiling PRNCR1 RNA revealed that PRNCR1 could

![Figure 4](https://example.com/figure4.png)

**Figure 4**

PVT1-mediated MYC protein stability regulation. Left, in normal cells, MYC degradation is promoted by phosphorylation at threonine 58 (Thr58). Right, in cancer cells, co-gain of PVT1 and MYC on 8q24 elevates the expression of MYC and PVT1. PVT1 interferes with MYC phosphorylation at Thr58, which increases MYC protein stability, leading to elevated MYC levels in cancers.
bind to the androgen receptor and was involved in the androgen-receptor-mediated gene activation in prostate cancers [55]. The findings presented in this study have broadened the mechanisms of IncRNA action at the MYC locus. Clearly, the 8q24 ‘gene desert’ contains previously underestimated information and extensive investigation will be required to fully determine the impact of this region on gene regulation during normal and disease development.

Acknowledgements

This work was supported by Grants XDA01010206 from Chinese Academy of Sciences (CAS) 91440202, 31322018 from the Natural Science Foundation of China. 31471241 from NSFC to LY.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

● of special interest
◆ of outstanding interest


This work identified several tumor-type specific super-enhancers of MYC, including a CRC specific super-enhancer containing the CCAT1 gene.


This study described the role of CCAT1-L in MYC regulation by modulating chromatin loops between distal enhancers and the MYC promoter in coordination with CTCF in CRC.


This work showed the MYC-335 enhancer derived CCAT2 in promoting MYC and WNT signal pathways via the transcription factor TCF4-mediated transcriptional regulation in CRC.


The authors demonstrated that a single copy gain of the 8q24.21 genes promoted cancer by using transgenic mice, and also illustrated the function of PVT1 IncRNA in controlling MYC protein stability.


This study showed that disruption of transcriptional coactivator binding to giant super-enhancers affected the transcription of genes with super-enhancers including MYC, suggesting the potency of components of super-enhancers in cancer therapeutics.


Mouse models have been developed in this study to demonstrate the function of the conserved MYC-335 enhancer in the regulation of MYC expression and tumor development.


This paper revealed that Firre was located across a 5-Mb on the X chromosome and five other trans-chromosomal loci. Firre interacted with hnrNPU and was required for the colocalization of these trans-chromosomal interacting loci, as revealed by the genetic deletion of the Firre locus.


This study identified a circular isoform of PVT1 which was expressed in a cell-type specific manner.


