The missing piece

Winkle, Melanie

CHAPTER 5

KTN1-AS1, an important downstream effector of Myc, controls B-cell lymphoma proliferation

Melanie Winkle†, Mina Tayari†, Annika Seitz†, Debora de Jong†, Jasper Koerts†
Lydia Visser†, Arjan Diepstra†, Joost Kluiver† and Anke van den Berg†

†Department of Pathology and Medical Biology, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands.

Manuscript in preparation

Abstract

Myc is an important transcription factor with oncogenic properties in many types of cancer, including Burkitt lymphoma (BL). As Myc itself is considered undruggable, downstream targets need to be identified to allow development of novel therapeutic approaches. Recently, long noncoding (lnc)RNAs were described to be important downstream genes of Myc in several types of cancer. To identify high-confidence Myc-regulated lncRNAs in B-cell lymphoma we generated and overlapped diverse data sets including transcript level changes upon *MYC* knockdown in BL, transcripts with an early response to Myc activation in P493-6 B cells and transcripts with a Myc binding site in close vicinity to its transcription start site. This resulted in the identification of 2 Myc-induced and 4 Myc-repressed lncRNAs. In-depth characterization of the Myc-induced KTN1-AS1 lncRNA revealed severely impaired BL cell growth without increasing the rate of apoptosis upon shRNA-mediated knockdown. KTN1-AS1 was enriched in the chromatin fraction, but showed no *cis*-regulatory effect on the nearby protein-coding gene KTN1. Genome wide expression analysis upon KTN1-AS1 knockdown indicated potent effects on the transcription levels of 295 genes. Intriguingly, gene set enrichment analysis showed enrichment of Myc-regulated gene sets and gene sets related to metabolism. Moreover, the KTN1-AS1 regulated genes were also enriched for Myc binding sites. A possible direct interaction between Myc and KTN1-AS1 was supported based on reduced Myc transcript and protein levels upon KTN1-AS1 knockdown. In conclusion, we identified MYC-induced KTN1-AS1 as an important oncogenic lncRNA involved in the pathogenesis of BL by reinforcing high Myc levels via a positive feedback loop. Our data depict targeting of KTN1-AS1 as a potential novel therapeutic approach.

1 Introduction

Long noncoding (lnc)RNAs have gained much attention in the past decade as novel modulators of normal and cancer cell biology. As lncRNAs are expressed in a cell type specific manner[1, 2] and expression can be cancer cell specific[3], studying their function can potentially lead to an expansion of druggable targets for cancer treatment. Diverse transcriptional as well as post-transcriptional regulatory roles have been described for lncRNAs, including gene regulation in *cis* and *trans*, as well as guiding localization and functionality of proteins, and stability and processing of mRNAs or miRNAs (reviewed in[4]). Many lncRNAs are under tight control of the well-known oncogenic transcription factor Myc[5-7]. These lncRNAs are likely to have a role in the mediation of growth and proliferation supportive effects in malignant cells.

Next to many other types of cancer, Myc also plays important roles in several subtypes of B-cell lymphoma. The most prominent example is Burkitt lymphoma (BL), characterized by a translocation involving the *MYC* gene locus and the immunoglobulin heavy or one of the light chain gene loci. This translocation results in high levels of Myc[8]. Diffuse large B-cell lymphoma (DLBCL) harbors *MYC* rearrangements or amplifications in 5 to 14% of cases[9]. DLBCL *MYC* rearrangements may coincide with *BCL2* or *BCL6* rearrangements in 'double hit' lymphomas, which is thought to confer additional aggressiveness[10]. Other subtypes of B-cell lymphoma more rarely contain *MYC* translocations. In line with an important role for Myc in lymphoma, Myc signatures have been identified in several subtypes and these have been linked to survival.

Myc is thought to bind 10-15% of all human gene loci and accordingly affects the expression of thousands of genes. These genes are involved in cell cycle progression, growth, metabolism and apoptosis[11]. In contrast, Myc was also suggested to act as a general amplifier targeting most or all genes already expressed in the cell[12, 13]. This might possibly cause its broad effects on metabolism, biosynthesis and growth, arguing for a rather general role in transcriptional activation. However, Myc also induces diverse cell type-specific effects[14, 15], indicating an essential role for the cell type in which Myc acts on the functional outcome. In line with this, Myc has a large repertoire of co-factors. Heterodimerization with its primary interactor Max is essential for Myc binding to DNA, which occurs primarily at E-box elements (CANNTG) and often within CpG islands[16, 17]. Other Myc co-factors may cause either transcriptional activation (e.g. TRRAP, YY1, p300) or repression (e.g. Sp1, Nf-y). Furthermore, other Max-interacting factors such as Miz1 or Mxi1 may compete with the binding of Myc to Max and result, via additional recruitment of transcriptional repressors and chromatin modifiers, in repression of Myc target genes (reviewed in[15, 18]).

Multiple lncRNAs have been described to act within the Myc pathway, including lncRNAs regulating Myc levels at the transcriptional (CCAT1, PCGEM1, CUDR)[19-23] or post-

transcriptional level (GAS5, GHET1, PCAT-1, RoR, MIF)[24-28]. In addition, Myc regulates expression of several lncRNAs, including Myc-induced (CCAT1, H19, HOTAIR, MINCR, MIF, MYU, BCYRN1, MYCLo-1, CCAT6)[7, 19, 21, 28-33] as well as Myc-repressed (MYCLo-4, LPP-AS2)[34] lncRNAs.

In this study we used multiple *in vitro* models in combination with a Myc binding profile analysis to identify lncRNAs directly regulated by Myc. This resulted in the identification of two Myc-induced and four Myc-repressed lncRNAs. In-depth analysis of the top MYC-induced candidate lncRNA KTN1-AS1 revealed that this lncRNA is a critical factor supporting BL cell growth, likely by regulating Myc target genes via a positive-feedback regulatory loop between Myc and KTN1-AS1.

2 Materials and Methods

2.1 Cell culture and cell line panel

BL cell lines were purchased from ATCC (ST486) and DSMZ (CA46 and DG75), while P493-6 B cells were a kind gift of Prof. D. Eick (Helmholtz center Munich). These cell lines were cultured at 37°C under an atmosphere containing 5% CO2 in RPMI-1640 medium supplemented with 2mM ultra-glutamine, 100U/mL penicillin, 0.1mg/mL streptomycin and 20% (ST486) or 10% (P493-6, CA46 and DG75) fetal calf serum (all reagents from Cambrex Biosciences, Walkersville, MD, USA). At regular intervals, cell lines were tested for mycoplasma contamination by PCR[35] and subjected to short tandem repeat (STR) genotyping using the PowerPlex 16HS system (Promega, Madison, WI, USA) to confirm cell line identity. The cell lines included in the panel were cultured according to their proposed culture conditions and included: PMBL - Primary mediastinal B-cell lymphoma (K1106P, MedB1); HL - Hodgkin lymphoma (L428, L540, L1236, KM-H2, SUP-HD1, DEV), DLBCL - Diffuse large B-Cell lymphoma (SUDHL-2, -4 to -6 and -10, SC-1, OCILy3, U2932, DOHH2); BL - Burkitt lymphoma (ST486, CA46, DG75, RAMOS, BL65, Namalwa, Raij, Jijoye).

Q-VD-Oph (Cat# S7311; Selleckchem, Munich, Germany) stock of 5mM dissolved in 100% DMSO was used to inhibit caspase acivity. Medium was supplemented with Q-VD (10uM) or DMSO starting 24h after lentiviral infection and on every following culture day.

2.2 Patient samples

Frozen tissue sections of 13 BL (MYC translocation, EBV-, CD20+, CD10+ and BCL2-) and 9 lymph node derived CLL (CD20+, CD5+, cyclin D1- and variable ZAP-70 expression)

primary cases used for qRT-PCR analysis were described previously[5]. Each individual diagnosis was reviewed by an experienced hematopathologist according to the World Health Organization classification[36]. For Nanostring analysis, FFPE tissue sections were obtained from 7 BL (including 2 overlapping with the FFPE selection), 12 CLL (no overlap with FFPE) cases and 180 DLBCL cases with an ABC (n = 81) or GCB (n = 99) cell of origin based on the Lymph2X cell of origin classifier[37]. The unclassifiable (n=25) cases were not included. The DLBCL cases were retrieved from the HOVON 46 and 84 cohorts based on availability of sufficient tumor tissue in the FFPE blocks. The procedures were performed according to the guidelines of the medical ethics board of the UMCG.

2.3 Subcellular fractionations

Cytoplasmic, nuclear and chromatin fractions of BL cell lines were separated as described earlier[38] with some modifications. In brief, cell pellets were lysed on ice in 1x lysis buffer (10mM Tris-HCL pH8, 300mM sucrose, 10mM NaCl, 2mM MgAc2, 3mM CaCl2, 0.1% Nonidet P-40, 0.5mM DTT) and centrifuged at 1,000xg for 5min at 4°C to yield the cytoplasmic fraction. The nuclei fraction was washed once in glycerol buffer (50mM Tris-HCL pH8.0, 25% glycerol, 5mM MgAc2, 0.1mM EDTA, 5mM DTT). To separate nucleoplasmic and chromatin fractions, nuclei were lysed in Urea Buffer (20mM HEPES pH7.5, 7.5mM MgCl2, 0.1mM EGTA, 0.3M NaCl, 1M Urea, 1% Nonidet P-40, 1mM DTT) and centrifuged at 14,000xg for 10min at 4°C. All buffers were supplemented with 20u/mL RNAseOUT and 1x Complete™ EDTA-free protease inhibitor cocktail (Roche Penzberg, Germany) prior to use. Qiazol lysis reagent (Qiagen, Hilden, Germany) was added to all fractions for RNA isolation.

2.4 shRNA-mediated knockdown

ShRNAs were designed using the Invivogen siRNA wizard and sense (S) and antisense (AS) oligo's (see **TABLE S1**) were ordered from Integrated DNA Technologies (IDT, Coralville, Iowa, USA) and cloned into the pGreenpuro lentiviral vector (SBI, MountainView, CA, USA). Generation of viral particles and infection of cells was performed as described previously[5]. Flow cytometry was used to assess the percentage of infected cells based on GFP expression. Cells were harvested directly if GFP+ cells were >90% or after sorting GFP+ cells at day 4 and day 6 (KTN1 and KTN1-AS1 knockdown) or day 8 (MYC knockdown) after infection.

2.5 Western blotting

Cells were washed with cold PBS and lysed either in Lysis Buffer (#9803, Cell Signaling Technology, Danvers, MA, United States) (BL cell lines, MYC knockdown samples) or in RIPA buffer (50mM Tris, 150mM NaCl, 2,5mM Na2EDTA, 1% Triton X-100, 0.5%mM sodium deoxycholate, 0.1% SDS) (all other samples) with 1mM phenylmethanesulphonyl fluoride for 45 minutes on ice. Subcellular fractionation samples (cytoplasm,

nucleoplasm) were used in their respective lysis buffers. Lysates were cleared by centrifugation (14,000xg, 10min, 4°C) and protein concentrations were determined using Pierce BCA Protein Assay (Thermo scientific) according to the manufacturer's instructions. 10% polyacrylamide gels were used for separation followed by transfer onto nitrocellulose membranes using standard procedures. Membranes were blocked in 5% milk supplemented with Tris-buffered saline and 0.1% Tween-20 followed by incubation overnight at 4°C with the following antibodies: monoclonal anti-human Myc antibody (Y69) (Cat# 1472-1, Epitomics, Burlingame, USA); rabbit monoclonal anti-human Myc antibody (Y69) (Cat# ab32072; Abcam, Cambridge, UK); polyclonal rabbit anti-human (cleaved) PARP (Cat# 9542; Cell Signaling, Danvers, MA, USA); Rabbit polyclonal anti-human histone H3 (FL-136) (Cat# sc-10809; Santa Cruz Biotechnology, Dallas, TX, USA); rabbit polyclonal anti-human α-tubulin (H-300) (Cat# sc-5546; Santa Cruz); and rabbit polyclonal anti-human Rb (C-15) (Cat# sc-50; Santa Cruz). Polyclonal horseradish peroxidase–conjugated goat anti-rabbit Ig (1000x) and rabbit anti-mouse Ig (1000x; both from Dako, Glostrup, Denmark) were used as secondary antibodies. Membranes were incubated with Super Signal West Pico Chemiluminescent Substrate (Thermo Scientific, Rockford, IL, USA) according to the manufacturer's instructions and signals were visualized on the ChemiDoc MP scanner (Biorad, Veenendaal, The Netherlands). Image Lab 4.0.1 Software (Bio-Rad) was used for quantification of protein bands.

2.6 GFP competition assays

GFP measurements were performed tri-weekly for a period of three weeks starting at day four after infection. The relative amount of GFP+ cells is calculated by normalization to the first measurement (range 20-60% GFP+ cells). All GFP competition assays were performed in triplicate per shRNA and per cell line. Non-targeting control shRNAs were included in all experiments.

2.7 Microarray analysis

Two array designs were used in this study. MYC knockdown samples were hybridized to a custom design lncRNA / mRNA array ('Custom array', AMADID: 039731). Array design and procedures were described previously[5]. KTN1-AS1 knockdown samples were hybridized to a commercially available microarray (AMADID: 072363; Agilent Technologies, Santa Clara, CA, USA) following the same procedures. The Custom and Agilent arrays contained 26,426 and 33,680 probes against coding and 31,054 and 19,049 probes against noncoding transcripts, respectively. 29,779 probes are shared between both arrays, while 5,520 noncoding probes match >50% [30nt] resulting in a total overlap of 35,299 probes. Raw data were extracted with Agilent Feature Extraction software v12 and analyzed with Genespring GX 13.1.1 software (Agilent Technologies). Probes flagged as present by the feature extracting software and consistently expressed in the 10th

to 100th percentile in at least 1 out of 2 conditions (i.e. knockdown or control) were used for further analysis. Using these settings, 7,192 lncRNA probes and 13,741 mRNA probes (5,040 and 10,670 loci) were expressed above background in ST486 cells with Myc knockdown, and 3,012 lncRNA and 14,760 mRNA probes (2,425 and 11,368 loci) in ST486 cells with KTN1-AS1 knockdown. LncRNA probes that may also detect coding genes with high affinity as well as probes detecting pseudogenes were excluded from the final lists of differentially expressed lncRNA transcripts. Heatmaps were generated with Genesis software v1.8.1[39] (Institute for Genomics and Bioinformatics Graz, Graz, Austria) using unsupervised hierarchical clustering with Pearson correlation as the distance metric.

2.8 Enrichment Analyses

Gene Set Enrichment Analysis (GSEA) was performed using the Hallmarks genesets of the Molecular Signatures Database (http://www.broad.mit.edu/gsea)[40] on all coding probes expressed above background for *MYC* and KTN1-AS1 shRNA treated ST486 cells. In addition, we used Enricher (http://amp.pharm.mssm.edu/Enrichr/)[41, 42], to assess transcription factor binding site (ENCODE/CHEA TFs dataset) and pathway (KEGG, Reactome) enrichment.

2.9 Detailed Myc and Max binding site analyses

In order to select for well-annotated genes, only probes mapping to either RefSeq genes or annotated lincRNAs[2] were considered for Myc binding site analysis (52,728 and 57,480 probes on Agilent and Custom arrays, respectively). ChIP data were retrieved from Seitz et al[43] (Myc binding sites in BL cell lines) or ENCODE (CpG islands and Myc and Max binding sites in the GM12878 lymphoblastoid cell line). For all transcripts with a probe on the microarray, the region of the transcription start site (TSS) +/- 5kb was fetched and overlapped with binding site locations using Galaxy[44, 45].

2.10 RNA isolation and quantitative (q)RT-PCR

RNA isolation was performed using Phase Lock Gel Heavy tubes (5 Prime Inc., Hilden, Germany) in combination with the miRNeasy mini/micro kit (Qiagen) according to the manufacturer's instructions. RNA from FFPE tissue sections was isolated using the RNeasy FFPE kit (Qiagen). On-column DNAse digestion was performed for all samples. RNA concentration was measured with a NanoDropTM 1000 Spectrophotometer (Thermo Fisher Scientific Inc., Waltham, USA), RNA integrity was assessed on a 1% agarose gel. cDNA was synthesized using random primers, dNTP mix and the Superscript II Reverse Transcriptase Kit (Life Technologies Europe BV, Bleiswijk, NL) according to manufacturer's instructions. An input of 500ng RNA was used per sample in a total reaction volume of 20µL.

To detect transcript levels, SYBRgreen mix (Applied Biosystems) was used in a qPCR reaction volume of 10µL with 1ng cDNA and 150-300nM primers in on a Lightcycler 480 system (Roche, Penzberg, Germany). Primer sequences used in this study are listed in TABLE S1. Relative expression levels are calculated based on the expression of the housekeeping genes TBP or U6. TBP was generally used for normalization in BL cell line experiments. U6 was the most stable housekeeping gene (of 10 tested) in highly heterogeneous sample populations including the cell line panel and primary patient cases.

2.11 Nanostring

Expression levels in FFPE patient samples were assessed using the NanoString nCounter technology using a custom design including the Lymph2X COO gene panel[37], probes for MYC, KTN1-AS, MELLR-2 and 410 probes not relevant for this project. All reagents were purchased from NanoString Technologies (Seattle, MA, USA). Hybridization was performed according to the manufacturer's instructions, with minor modifications. In brief, 100ng RNA input was incubated with ½ the recommended concentration of capture and reporter probes for 16h at 65C. Samples were stored at 4C until further processing. The sample was loaded on a nCounter SPRINT Cartridge and processed on the nCounter SPRINT Profiler (NanoString Technologies). NanoString nSolver software (v3.0) was used for analysis. Counts passing the standard QC parameters were processed as follows: The raw counts were normalized against the geometric mean of internal positive controls to compensate for technical variability. Next, the data were normalized against the geometric mean of the housekeeping genes R3HDM1, WDR55, ISY1, UBXN4 and TRIM56. The probes used for NanoString analysis for these latter genes are listed in TABLE S1.

2.12 Statistics

For GFP competition assays, the significance of changes in GFP+ cell percentages in knockdown samples compared to controls was assessed as described previously[5]. Differences in GFP+ cell percentages upon DMSO or Q-VD treatment were determined using repeated measures ANOVA and Tukey's Multiple comparison test, comparing DMSO and Q-VD per shRNA construct. Significantly differential expressed microarray probes between knockdown and control conditions were determined using a moderated T test (p-value cutoff 0.05) and Benjamini-Hochberg multiple testing correction, followed by selection of probes with a >1,5-fold change in expression. For comparison, previously published microarray data from P493-6 cells[5] (GSE59480) were reanalyzed using the same settings and a >2-fold change in expression cutoff (conditions: MycOFF vs MycON all time points; MycOFF vs. MycON-4h). Enrichment of binding sites or CpG islands was evaluated using chi-square test. The percentage of probes with a specific feature within all probes deregulated upon Myc or KTN1-AS1 knockdown was compared to the percentage of probes with the same feature within all probes present on the

array. Significant expression differences for the qPCR data for the cell line panel and BL vs CLL patient cases were tested using standard T test with Welch's correction (naïve, memory, PMBL, HL, DLBCL or BL vs. GC B; BL vs CLL). Nanostring data were analyzed using Kruskal-Wallis and Dunn's multiple comparison tests.

3 Results

3.1 *MYC* depletion in BL cells to identify Myc-regulated transcripts

To confirm the dependency of BL cell lines on Myc we determined the effect of shRNA-mediated *MYC* inhibition on cell growth. Endogenous Myc levels were lowest in ST486, highest in CA46 and intermediate in DG75 cells, at both the mRNA and protein level (FIG. S1A). Two shRNAs against *MYC* caused a decrease of 60-80% at the protein level (FIG. S1B) in *MYC* shRNA treated samples compared to non-targeting control shRNA treated samples. More moderate decreases were observed at the mRNA level (data not shown). Three known Myc-induced genes (CAD, TFAM and PGK1)[46] showed a consistent decrease in expression upon shRNA treatment, further confirming effective knockdown of Myc (FIG. S1C). GFP competition assays showed a highly significant decrease in GFP+ cells for both shRNA constructs in all 3 BL cell lines ($p < 0.0001$), with the most prominent effect in ST486 cells (FIGURE 1A). These data confirm high expression of Myc in BL cell lines and their addiction to the Myc protein for survival.

To identify Myc-regulated lncRNAs, we assessed Myc-dependent gene expression changes using two shRNAs in ST486, being the most sensitive BL cell line. 10,670 mRNA and 5,040 lncRNA loci were expressed in ST486 cells. Significant enrichment of Myc target genes in control over knockdown samples was confirmed by gene set enrichment analysis (GSEA) showing enrichment of genes induced by Myc (FIG. S1D and TABLE S2). Enricher analysis using the MYC-induced genes showed the expected enrichment of the transcription factor Myc and also revealed enrichment of its binding partner Max (see also below) and pathways related to metabolism. Together, these data confirm effective Myc depletion in ST486 cells.

3.2 Identification of high confident Myc-regulated lncRNA loci in BL cells

A total of 1,048 mRNA and 324 lncRNA loci (1,152 and 377 probes, respectively) were significantly deregulated in Myc shRNA treated compared to non-targeting control shRNA ST486 cells (Fold change >1.5 fold). A similar number of Myc-induced (47%, 492 loci) and Myc-repressed (53%, 556 loci) mRNA transcripts were identified (FIGURE 1B). In contrast, fewer lncRNAs were Myc-induced (37%, 120 loci) than Myc-repressed (63%, 204 loci) (FIGURE 1C). Unsupervised clustering of the 1,048 Myc-regulated protein-coding genes

revealed a separate clustering of shMYC and shControl samples (FIGURE 1B). Unsupervised clustering of the 324 MYC-regulated lncRNA loci revealed a similar clustering with exception of the shMYC1-2 replicate, which clustered with the control samples (FIGURE 1C).

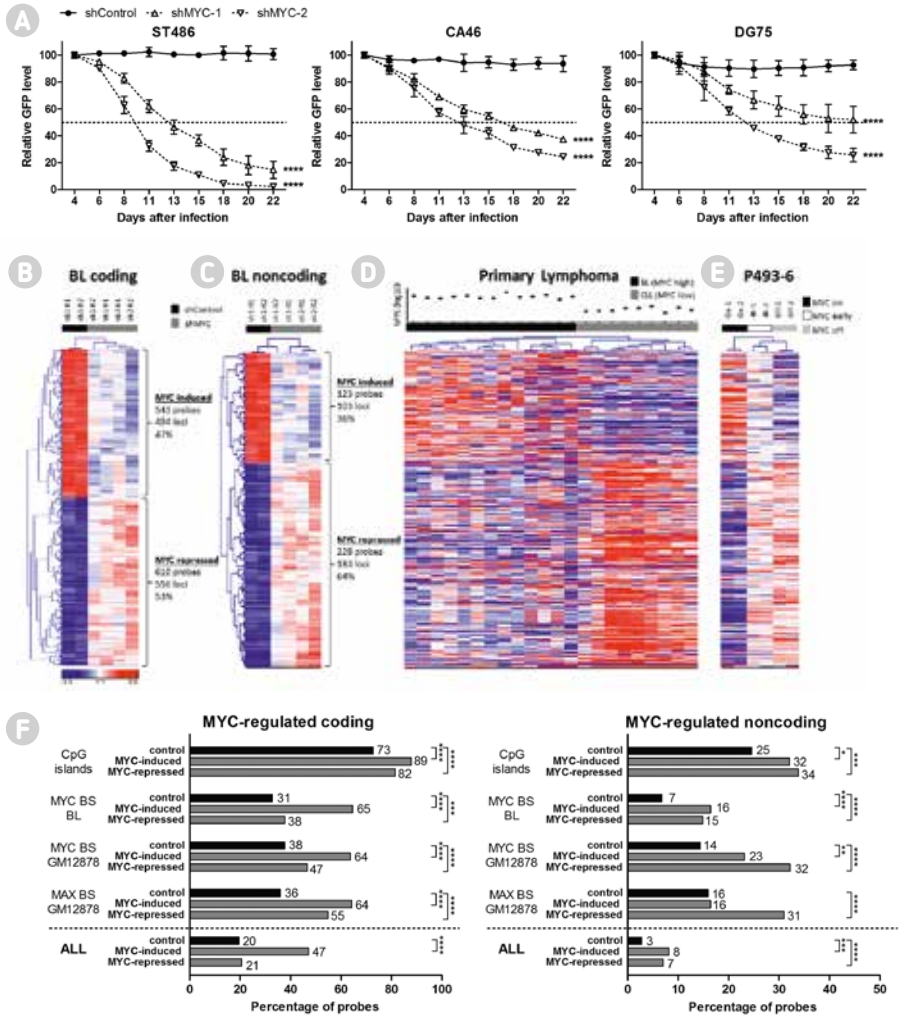

FIGURE 1 Myc-regulated coding and noncoding genes in Burkitt lymphoma. (A) Myc knockdown causes a significant growth decrease in all cell lines tested. Dashed lines indicates 50% reduction in GFP positive cells; n = 3, mean ± SEM. **(B)** Heatmap of differentially expressed coding genes upon Myc knockdown in ST486 generated by unsupervised hierarchical clustering using Pearson correlation. **(C)** Myc-responsive lncRNA probes in ST486 cells. **(D)** The same probes as in (C) are shown in P493-6 cells expressing high, medium (i.e. 4h after Myc-induction) or absent levels of Myc; and **(E)** in primary lymphoma cases with high (Burkitt lymphoma; BL) or low Myc expression (chronic lymphocytic leukemia; CLL). *MYC* mRNA expression in the individual cases is indicated on top of the heatmap. qRT-PCR normalized to U6. **(F)** Analysis of Myc occupancy at coding and noncoding loci responsive to Myc knockdown in ST486. The percentage of probes detecting genes with a CpG island, Myc or Max binding site within 5kb of the transcription start site are shown. The percentage of all probes on the array with the respective feature within 5kb is displayed as control. Chi-square test (Myc-induced or Myc-repressed vs control). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $P<0.0001$.

In order to further support the relevance of these lncRNAs for B-cell lymphoma, we assessed their expression in primary lymphoma samples with high and low *MYC* levels (i.e. BL and chronic lymphocytic leukemia; CLL, respectively)[5]. 23% of Myc-induced and 27% or Myc-repressed lncRNAs (24 and 49 loci) showed significant (>2-fold) expression changes in the expected direction. Unsupervised hierarchical clustering of all 377 Myc-regulated lncRNA probes identified in ST486 cells caused a clear separation of lymphoma samples according to their Myc status (**FIGURE 1D**). Further subgrouping within the BL and CLL clusters was not related to the level of *MYC* expression (**FIGURE 1D**, top). Next, we reanalyzed our previous data from the B cell P493-6 model[5], which carries a conditional, tetracycline-repressible *MYC* allele. 40% of the Myc-induced (41 of 103) and 42% of the Myc-repressed lncRNAs (76 of 183) showed a significant >2-fold change in expression between MycON and MycOFF conditions in P493-6. Unsupervised clustering of these samples and the total Myc-regulated gene set showed a separation that was based on their Myc status (**FIGURE 1E**).

To check the consistency of our data with earlier studies, we investigated the expression of known Myc-regulated lncRNAs based on presence of probes flagged as present on the microarray. This included 5 Myc-induced (H19, HOTAIR, MIF, MINCR, MYU) and one Myc-repressed (LPP-AS2) lncRNA. Despite not being detected as significantly differentially expressed in Myc-depleted ST486 cells, MIF, MYU and HOTAIR showed the expected expression changes in ST486 (**FIGURE S2A**), MINCR and MYU in P493-6 cells (**FIG. S2B**) and H19, MIF, MINCR and MYU in Myc high versus Myc low lymphoma samples (**FIG. S2C**). Myc-repressed LPP-AS2 does not appear as such in any of the three models. Thus, for three of the six known Myc-regulated lncRNAs we see the expected direction of Myc-regulation in at least two of the three *in vitro* models.

Next, we analyzed presence of CpG islands, Myc and Max binding sites at the affected coding and noncoding loci. CpG islands were observed more frequently near the TSS of coding genes as compared to noncoding genes (73% vs 25%). Both, Myc (31-38% vs 7-14%) and Max (36% vs 16%) binding occurred more often near the TSS of coding genes as compared to noncoding genes (**FIGURE 1F,** controls). This is in accordance with the lower number of lncRNAs responding to Myc knockdown compared to mRNAs. Both Myc-induced and Myc-repressed coding genes are significantly enriched for CpG islands, Myc and Max binding sites compared to control ($p < 0.001$) (**FIGURE 1F,** left). This enrichment was especially pronounced for Myc-induced mRNAs. Analysis of Myc-regulated noncoding loci revealed an enrichment of all three features in Myc-repressed lncRNAs, while Myc-induced lncRNAs showed enrichment of CpG islands and Myc binding, but not for Max binding (**FIGURE 1F,** right). Together, these data confirm the robustness of the identified Myc-regulated genes and reveal differences in the binding profiles of Myc and Max between coding and noncoding loci.

To identify high confidence Myc-regulated lncRNAs, we overlapped the lncRNAs up- and downregulated by Myc in ST486 cells with early response (within 4h after Myc

induction, 445 lncRNA loci) lncRNAs identified in P493-6 cells and noncoding gene loci with a proven Myc binding site in BL[43] (**FIG. S3A AND B**). 23% (22 of 103 loci) of Myc-induced and 7% (12 of 183 loci) of Myc-repressed loci defined in ST486 were among the early responders identified in P493-6 cells. Myc binding sites were identified near the TSS of 16 Myc-induced and 35 Myc-repressed lncRNAs. Overlap of all three datasets resulted in six high confidence Myc-regulated targets including two Myc-induced and four Myc-repressed lncRNAs.

3.3 Characteristics and validation of the high confident Myc-regulated lncRNA loci

The identified set of 6 lncRNAs (**TABLE 1** and **FIGURE 2**) included two transcripts that have a bidirectional orientation with a nearby protein coding gene (Myc-induced KTN1-AS1 and Myc-repressed PSMG3-AS1), one true antisense lncRNA (Myc-repressed TNFRS14F-AS1) and three intergenic transcripts hereafter termed MELLI-1, MELLR-1 and MELLR-2 (Myc Effectuated LncRNA Locus Induced / Repressed). To validate a Myc-regulated expression pattern, we inhibited *MYC* in ST486 and two additional BL cell lines. Knockdown efficacy was confirmed through assessment of MYC, CAD, TFAM and PGK1 transcript levels (**FIG. S4A**). All six lncRNAs showed the expected transcriptional changes in ST486 cells (**FIG. S4B-G**). In addition, four lncRNAs showed the same pattern in at least 1 additional cell line (KTN1-AS1, MELLR-2, TNFRSF14-AS1, PSMG3-AS1; **FIG. S4B, E-G**), while one lncRNA was not expressed in the other two cell lines (MELLR-1; **FIG. S4D**) and one showed an opposite expression pattern in the other two cell lines (MELLI-1; **FIG. S4C**). As previously observed in P493-6 cells[5] and in line with a putative *cis*-regulatory role, the nearby protein coding gene TNFRSF14 is co-repressed with TNFRSF14-AS1 (**FIG. S4F**), while PSMG3 located adjacent to MYC-repressed PSMG3-AS1 is increased in MYC-high conditions (**FIG. S4G**). KTN1 mRNA levels were increased upon *MYC* depletion only in DG75 cells and not in the two other BL cell lines.

TABLE 1 High confidence selection of Myc-regulated lncRNAs.

lncRNA locus	Genomic location (hg19)	Myc-regulation ST486#	Early Myc response P493-6#	Distance Myc BS to TSS (nt)$	Myc binding site*
KTN1-AS1	chr14:56024690-56046810	+3.6	+9.7	-82	BS-5042
MELLI1	chr22:42670646-42673855	+2.0	+7.0	-4.841	BS-1447
MELLR1	chr7:150130742-150145228	-4.1	-2.3	+1.957	BS-4900
MELLR2	chr11:331734-333645	-3.6	-5.5	-1.123	BS-3104
TNFRSF14-AS1	chr1:2481359-2488450	-3.1	-3.6	-1.227	BS-4547
PSMG3-AS1	chr7:1609709-1629261	-1.8	-5.5	-37	BS-4351

#Fold changes as measured by microarray; $Distance from the center of the binding site peak to the transcription start site. *Myc binding site region as published in[43]

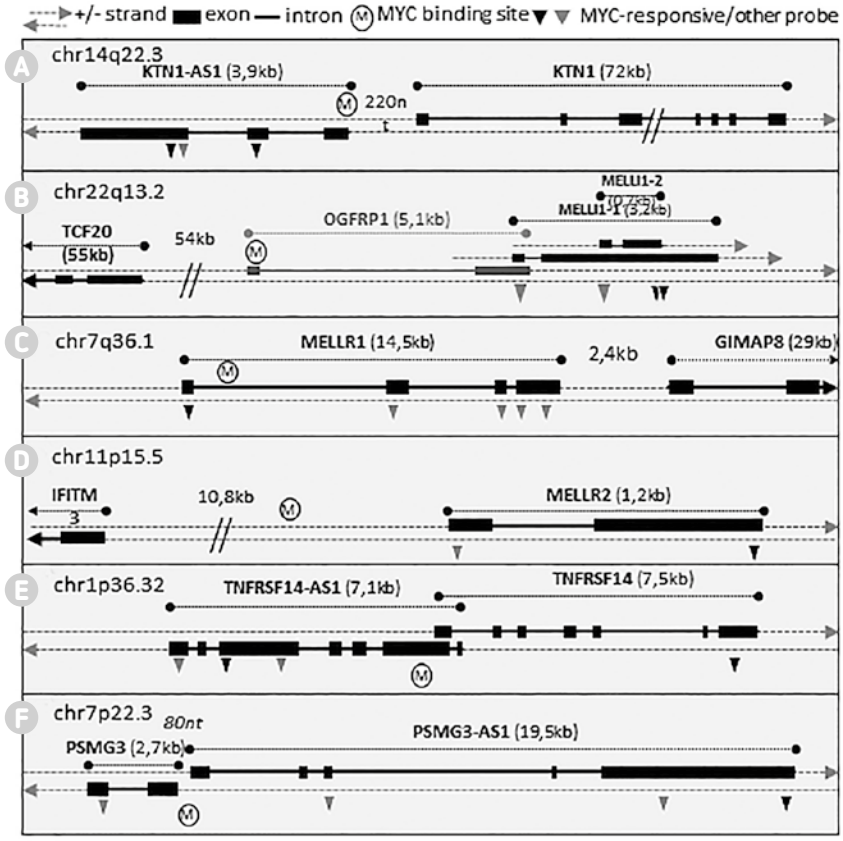

FIGURE 2 Schematic of Myc-regulated lncRNA loci in BL. Schematic representation of the lncRNA loci defined as high confidence Myc targets. Myc-induced loci **(A)** KTN1-AS1, and **(B)** MELLI-1, and MYC-repressed loci **(C)** MELLR-1, **(D)** MELLR-2, **(E)** TNFRSF14-AS1 and **(F)** PSMG3-AS1 are shown. Myc binding sites (M), differentially expressed probes (black triangles), as well as non-responsive probes (grey triangles) are indicated. Image is not to scale.

3.4 Expression of the top Myc-induced and Myc-repressed lncRNA candidate in B-cell lymphoma

To get an indication of the biological significance of our top candidates, i.e. Myc-induced KTN1-AS1 and Myc-repressed MELLR-1, we assessed their expression levels in a broad panel of lymphoma cell lines, primary cases and normal B-cell subsets. As a control, we also analyzed *MYC* expression levels in these samples. *MYC* levels (FIGURE 3A) were significantly increased in HL ($p = 0.0059$), DLBCL ($p = 0.0005$) and BL ($p = 0.048$) cell lines compared to GC-B cells. KTN1-AS1 expression levels (FIGURE 3A) were slightly increased in GC-B cells compared to naïve ($p = 0.046$), but not compared to memory cells ($p = 0.16$). Similar to *MYC*, KTN1-AS1 levels were increased in HL ($p = 0.034$), DLBCL ($p = 0.015$) and BL ($p = 0.005$). MELLR1 was expressed in memory, but not in naïve or

GC-B cells; and was expressed only in four of seven HL, one of nine DLBCL and two of seven BL cell lines. *MYC* ($p < 0.001$) and KTN1-AS1 ($p = 0.011$) levels as assessed by qRT-PCR were higher in frozen tissue samples of BL as compared to CLL, while mean expression of MELLR1 was slightly, but non-significantly increased in CLL compared to BL ($p = 0.3$ (FIGURE 3B).

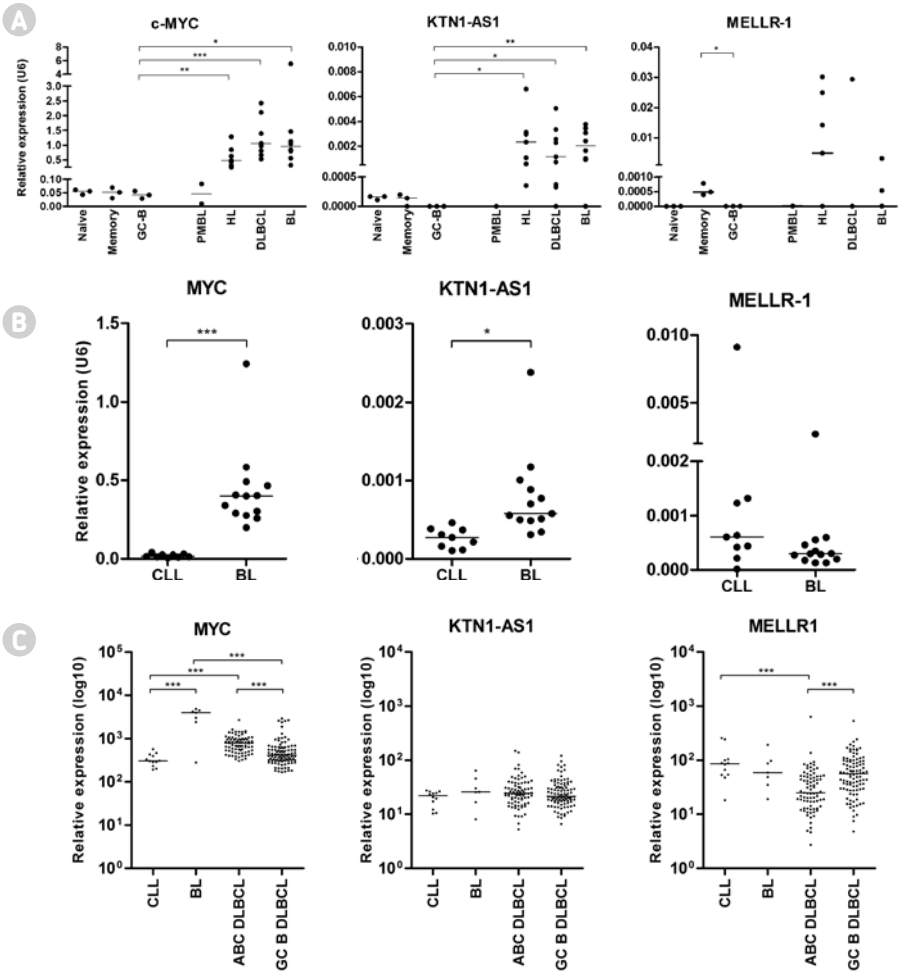

FIGURE 3 **Expression of top candidates, Myc-induced KTN1-AS1 and MYC-repressed MELLR-1, and MYC accross normal B cells, lymphoma cell lines and primary lymphoma cases. (A)** RT-qPCR expression analysis of MYC, KTN1-AS1 and MELLR-1 in sorted normal naïve, memory and germinal center (GC) B cell subsets (n = 3) and multiple lymphoma cell lines including primary mediastinal B-cell lymphoma (PMBL; n = 2), Hodgkin lymphoma (HL; n = 7), Diffuse Large B-cell Lymphoma (DLBCL; n = 9) and Burkitt lymphoma (BL; n = 7). qRT-PCR, mean ± error; t-test vs. GC-B, Welch's correction. **(B)** RT-qPCR expression analysis in a set of fresh frozen tissue samples consisting of primary chronic lymphocytic leukemia (CLL) characterized by low MYC levels and BL characterized by high MYC expression. qRT-PCR, mean ± error; t-test, Welch's correction. **(C)** Expression assessed using Nanostring in a set of FFPE tissue samples including CLL (n = 12), BL (n = 7) activated B cell like (ABC; n = 81) and germinal center B like (GC B; n = 99) DLBCL cases. Lines indicates median. Kruskal-Wallis test, Dunn's multiple comparison test. * p < 0.05, ** p< 0.01, *** p < 0.001, **** p < 0.0001.

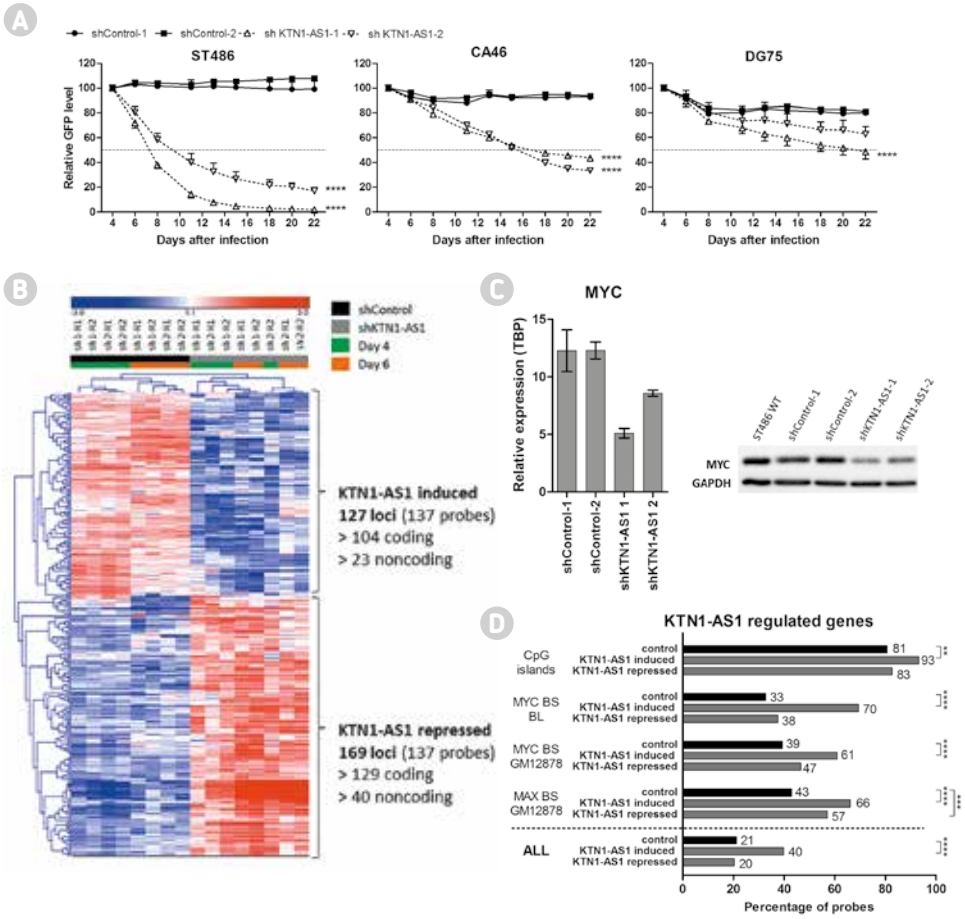

FIGURE 4 KTN1-AS1 depletion decreases cell growth mainly through downregulation of Myc. **(A)** KTN1-AS1 depletion causes a strong growth decrease in three BL cell lines. GFP competition assay; mean ± SD, n = 3, **** p < 0.0001. **(B)** Genome wide expression analysis after KTN1-AS1 depletion in ST486 cells reveals 296 differentially expressed loci. Heatmap of affected probes by unsupervised hierarchical clustering. **(C)** MYC mRNA and protein levels upon depletion of KTN1-AS1 (day 3 and day 6 after shRNA infection, respectively). Both performed in duplicate; qRT-PCR shows mean of two replicates; representative Western Blot. **(D)** Presence of CpG islands, Myc binding sites in BL, and Myc and Max binding sites in GM12878 lymphoblastoid cells near the KTN1-AS1 affected probes. Chi-square test; * p < 0.05, ** p < 0.01, *** p < 0.001, **** p < 0.0001.

Nanostring analysis on BL, CLL and DLBCL FFPE tissue samples revealed the most abundant *MYC* levels in BL, intermediate levels in DLBCL and low levels in CLL (**FIGURE 3C**). In line with the literature *MYC* levels were significantly higher in ABC compared to GCB DLBCL cases (*p < 0.0001*). KTN1-AS1 showed similar expression levels across all lymphoma cases. Levels of MELLR1 were the highest in CLL, intermediate in BL and GCB DLBCL. ABC DLBCL cases had the lowest MELLR1 levels (*p < 0.0001*, in comparison to GCB DLBCL). Overall, expression levels of KTN1-AS1 and MELLR1 are in line with KTN1-AS1 being Myc-induced and MELLR1 being Myc-repressed lncRNAs.

3.5 KTN1-AS1 depletion affects BL cell growth

To establish relevance of KTN1-AS1, we studied the effect of shRNA-mediated knockdown on growth of BL cell. shKTN1-AS1-1 was more efficient with a > 50% reduction in transcript abundance compared to a ~30% reduction observed for shKTN1-AS1-2 (FIGURE S6B). shKTN1-AS1-1 caused a significant ($p < 0.0001$) growth decrease in all and shKTN1-AS1-2 in two of three cell lines (FIGURE 4A). A 50% decrease in GFP positive cells was achieved at day 5, 15 and 20 after infection with shKTN1-AS1-1 in ST486, CA46 and DG75 cells, respectively. These data indicate a growth supportive role for KTN1-AS1 in BL.

To check if the observed growth decrease is due to apoptosis, we treated BL cells infected with non-targeting control, KTN1-AS1 or MYC shRNAs with the caspase inhibitor Q-VD. Q-VD treatment induced a minor rescue of the growth reduction phenotype in ST486 cells infected with both non-targeting control (8 and 11% difference Q-VD vs. DMSO) and both KTN1-AS1 shRNAs (20 and 14%). In CA46 cells treated with Q-VD we observed a minor rescue for shKTN1-AS1-2 (14%) and both *MYC* shRNAs (5 and 6%) (FIGURE S5A). Consistent with these modest effects, we only observed a minor increase in cleaved PARP in KTN1-AS1 depleted ST486 cells (FIGURE S5B). Thus, these data show that the phenotype observed upon KTN1-AS1 knockdown is not due to the induction of caspase-mediated apoptosis.

3.6 KTN1-AS1 transcripts show a nuclear localization

To obtain insights in the function of KTN1-AS1 we analyzed abundance of its transcripts in cytoplasmic, nucleoplasmic and chromatin fractions of three BL cell lines. To confirm purity of the fractions, we assessed enrichment of RNA transcripts with known subcellular localization: RPPH1 and DANCR as cytoplasmic, SNHG4 and ANRIL as nucleoplasmic and XIST (cell lines derived from female patients) or MIAT (for cell lines derived from male patients) as chromatin-associated transcripts[47,48]. These control RNAs were enriched in the expected fractions in all three cell lines (FIG. S6A). KTN1-AS1 showed strong enrichment in the nuclear fraction (85-98% of KTN1-AS1 RNA), with preferential localization in the chromatin fraction (50-90% of KTN1-AS1 RNA) in all three cell lines (FIG. S6A). The nuclear localization of KTN1-AS1 hints at a putative function in gene expression regulation.

3.7 Identification of KTN1-AS1 regulated gene expression changes

As KTN1-AS1 maps close to the protein coding gene kinectin 1 (KTN1), we checked whether depletion of KTN1-AS1 affects expression of the neighboring gene. The KTN1-AS1 shRNAs had no effect on the levels of KTN1 mRNA (FIG. S6B). Conversely, knockdown of KTN1 (~75% mRNA reduction with two shRNAs) did not affect the expression levels of KTN1-AS1 (FIG. S6C). These data suggest that KTN1 and KTN1-AS1 do not influence each other's expression.

Next, we assessed genome wide gene expression changes upon KTN1-AS1 depletion in ST486 cells, which showed the strongest phenotype on cell growth. 11,368 coding and 2,425 noncoding loci showed expression levels above background in either control or KTN1-AS1 shRNA samples. To define genes responsive to KTN1-AS1 depletion, we assessed significant expression changes in cells harvested at day 4 (212 deregulated probes) and day 6 (170 deregulated probes). Combination of the significantly altered genes resulted in a total of 219 coding and 82 noncoding loci (220 and 90 probes, respectively) responding significantly to KTN1-AS1 depletion. Unsupervised hierarchical clustering of all responsive genes resulted in a clear separation between control and knockdown samples (**FIGURE 4B**). Probes for 47% of the loci were increased in control vs. knockdown samples (137 probes, hereafter referred to as 'KTN1-AS induced'), while the remainder was decreased (175 probes; 'KTN1-AS1 repressed').

3.8 KTN1-AS1 depletion affect levels of Myc and a subset of its target genes

GSEA on KTN1-AS1 induced coding genes showed the strongest enrichment for a Myc target gene set that was also detected as being enriched in the Myc-induced coding genes (**FIG. S7A** and **TABLE S3**). Enricher analysis showed, similar to our observations for Myc knockdown, enrichment of the transcription factors Myc and Max and of pathways related to metabolism. Together, the enrichment analyses suggest that KTN1-AS1 may directly affect Myc levels.

To establish a potential direct effect of KTN1-AS1 on Myc, we analyzed Myc levels in KTN1-AS1 knockdown vs. control ST486 cells. A decrease in Myc levels was evident both at the mRNA and protein level and the reduction in Myc (**FIGURE 4C**) correlated with the efficiency of KTN1-AS1 knockdown (**FIG. S6B**). These data indicate that KTN1-AS1 enforces high Myc expression levels. This was also supported by the highly similar expression pattern observed in the heatmap for 259 of the 312 KTN1-AS1-regulated probes (i.e. the probes available on both array versions) in the shMYC-treated ST486 samples and the MycOn / MycOFF P493-6 cells (**FIGURE S7B**). Of the KTN1-AS1 affected genes, 21% and 55% showed significant expression changes to Myc modulation in ST486 (>1.5-fold) and P493-6 cells (>2-fold), respectively.

To further assess the potential collaborative effect of Myc and KTN1-AS1, we analyzed presence of CpG islands, Myc and Max binding sites nearby the TSS of the KTN1-AS1-regulated genes. As only 63 noncoding loci were affected, we restricted this analysis to the 233 coding loci. KTN1-AS1-induced coding genes (n = 104) showed a highly significant enrichment for CpG islands as well as Myc and Max binding sites (**FIGURE 4D**), very similar to what we observed for Myc-regulated coding genes (**FIGURE 1F**). KTN1-AS1-repressed coding genes (n = 129) also showed enrichment for Max binding, but not for Myc binding.

4 Discussion

As Myc is one of the most frequently overexpressed oncogenes but considered to be an undruggable target, the study of its downstream targets may lead to the identification of novel therapeutic targets. By intersecting lncRNAs responsive to Myc knockdown in BL cells with lncRNAs that are early responsive to Myc induction and map at genomic loci showing Myc occupancy within the TSS we identified six high-confidence Myc-regulated lncRNA loci. The relatively low number of lncRNA candidates identified is likely related to the high stringency of the analysis, requiring an over two-fold change in expression within 4h after Myc re-induction in P493-6 cells to qualify as 'early responder', as well as having a consistent Myc binding in the TSS of five individual BL cell lines[43].

In general, the Myc-responsive genes identified in ST486 show high consistency with our earlier approaches to identify Myc-regulated genes in P493-6 cells and in primary BL cases. Furthermore, over half of the Myc-responsive coding genes identified in ST486 cells also had a Myc binding site in other BL cell lines[43] and in EBV-transformed lymphoblastoid cells[49]. The Myc-dependent regulation of the protein coding genes was validated by gene-set and Myc binding site enrichment analyses, showing enrichment of Myc-related gene and ChIP sets. This high consistency between our expression analysis in ST486 cells and additional models shows that assessment of Myc-dependent expression changes even in a single cell line is a reliable approach to identify cell type specific Myc-targets.

The global analysis of features around the TSS of genes represented on the microarray shows that lncRNAs are less frequently associated with CpG islands (25%) compared to coding genes (73%). This observation is in accordance with lncRNA transcripts being evolutionary less conserved as well as highly tissue specific. The more often 'younger' age of lncRNAs as compared to protein coding genes, could explain the lower percentage of CpG islands, for which the generation is an evolutionary long process that depends on active transcription of the locus. Binding of transcription factors was also observed at a lower frequency in noncoding as compared to protein coding gene loci.

We found a strong enrichment of Myc binding sites for Myc-induced, but not for Myc-repressed coding genes, in line with our earlier observations in P493-6 cells and BL/CLL samples[5]. Max binding sites were enriched in both Myc-induced and Myc-repressed coding gene loci. The lack of Myc binding at repressed loci in combination with the more marked enrichment of Max binding may point towards a repressive mechanism involving the competition of Myc with other repressive Max-interacting factors, such as Mxi1[15]. For noncoding loci we observed strong enrichment of Myc binding sites in both

Myc-induced and Myc-repressed loci, whereas Max binding sites were enriched only in Myc-repressed noncoding gene loci. Our analysis is in line with a model in which Myc can activate and repress the expression of noncoding loci.

MELLR1 was highly expressed in 4 out of 7 HL cell lines, but only sporadically expressed across cell lines of other lymphoma subtypes. In tissue samples, MELLR1 was differentially expressed between ABC and GCB DLBCL subtypes. Lower expression of MELLR1 in the more aggressive ABC subtype warrants further study of this lncRNA as a predictor for survival or as an additional marker that could aid in the classification of DLBCL.

KTN1-AS1 expression was high in most of the cell lines and also in BL samples. Knockdown of KTN1-AS1 in BL cell lines had strong effects on growth without clearly affecting caspase-mediated apoptosis. In ST486 cells, the loss of GFP+ cells occurred faster upon KTN1-AS1 depletion (**FIGURE 3A**) than upon Myc knockdown. Genome wide expression analysis in KTN1-AS1 depleted cells pointed towards an active role in the Myc pathway. This was supported by the accompanying strong reduction of Myc observed at both the mRNA and protein level. The decrease in Myc correlated with the KTN1-AS1 knockdown efficiency. The effect on *MYC* combined with the chromatin-associated localization and the strong and early phenotype on growth upon KTN1-AS1 depletion suggests that KTN1-AS1 has a main role in regulating Myc-induced effects. This might be achieved by a direct effect on Myc expression and / or by serving as a co-factor for Myc-dependent transcriptional regulation for a subset of the Myc target genes. Our data suggests a positive feedback loop between Myc and KTN1-AS1, enforcing each other's expression. Intriguingly, a few other Myc-induced lncRNAs have already been described to promote *MYC* expression at the transcriptional (CCAT-1, PCGEM1)[20, 22] or posttranscriptional level (MIF, GHET1, XLOC_010588, PCAT-1, RoR)[26-28, 50, 51]. The existence of multiple Myc-regulating lncRNAs may be related to the tissue specific expression pattern of lncRNAs, requiring cell type specific lncRNAs that support or modulate Myc expression and Myc effects. The tissue specific expression of the Myc-regulating lncRNAs CCAT-1[20] and PCGEM1[22] in colorectal and prostate tissue respectively, is in line with this notion. It will be interesting to determine whether there are any similarities in sequence, secondary structure or binding partners of these lncRNAs that can support a similar function.

In conclusion, we identified six high-confidence Myc-regulated lncRNAs in BL. Further analysis of the most prominent Myc-induced lncRNA KTN1-AS1 revealed that this nuclear lncRNA is critical to the growth of BL cells by enhancing Myc levels and regulating a subset of the Myc targets. As such, KTN1-AS1 may represent a novel and potent target for alternative therapeutic approaches in lymphoma cells.

5 References

1 Derrien, T. *et al*. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* **22**, 1775-1789 (2012).

2 Cabili, M. N. *et al*. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* **25**, 1915-1927 (2011).

3 Leucci, E. *et al*. Melanoma addiction to the long non-coding RNA SAMMSON. *Nature* **531**, 518-522 (2016).

4 Schmitt, A. M. & Chang, H. Y. Long Noncoding RNAs in Cancer Pathways. Cancer. *Cell*. **29**, 452-463 (2016).

5 Winkle, M. *et al.* Long noncoding RNAs as a novel component of the Myc transcriptional network. *FASEB J.* **29**, 2338-2346 (2015).

6 Hart, J. R., Roberts, T. C., Weinberg, M. S., Morris, K. V. & Vogt, P. K. MYC regulates the non-coding transcriptome. *Oncotarget* **5**, 12543-12554 (2014).

7 Doose, G. *et al.* MINCR is a MYC-induced lncRNA able to modulate MYC's transcriptional network in Burkitt lymphoma cells. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E5261-70 (2015).

8 Molyneux, E. M. *et al.* Burkitt's lymphoma. Lancet 379, 1234-1244 (2012).

9 Karube, K. & Campo, E. MYC alterations in diffuse large B-cell lymphomas. Semin. *Hematol*. **52**, 97-106 (2015).

10 Petrich, A. M., Nabhan, C. & Smith, S. M. MYC-associated and double-hit lymphomas: a review of pathobiology, prognosis, and therapeutic approaches. *Cancer* **120**, 3884-3895 (2014).

11 Dang, C. V. MYC on the path to cancer. *Cell* **149**, 22-35 (2012).

12 Nie, Z. *et al*. c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells. *Cell* **151**, 68-79 (2012).

13 Lin, C. Y. *et al.* Transcriptional amplification in tumor cells with elevated c-Myc. *Cell* **151**, 56-67 (2012).

14 Walz, S. *et al.* Activation and repression by oncogenic MYC shape tumour-specific gene expression profiles. *Nature* **511**, 483-487 (2014).

15 Link, J. M. & Hurlin, P. J. The activities of MYC, MNT and the MAX-interactome in lymphocyte proliferation and oncogenesis. *Biochim. Biophys. Acta* **1849**, 554-562 (2015).

16 Mao, D. Y. *et al.* Analysis of Myc bound loci identified by CpG island arrays shows that Max is essential for Myc-dependent repression. *Curr. Biol.* **13**, 882-886 (2003).

17 Sabo, A. & Amati, B. Genome recognition by MYC. *Cold Spring Harb Perspect. Med*. **4**, 10.1101/cshperspect.a014191 (2014).

18 Tu, W. B. *et al*. Myc and its interactors take shape. *Biochim. Biophys. Acta* **1849**, 469-483 (2015).

19 Yang, F. *et al.* Long noncoding RNA CCAT1, which could be activated by c-Myc, promotes the progression of gastric carcinoma. *J. Cancer Res. Clin. Oncol.* **139**, 437-445 (2013).

20 Xiang, J. F. *et al.* Human colorectal cancer-specific CCAT1-L lncRNA regulates long-range chromatin interactions at the MYC locus. *Cell Res.* **24**, 513-531 (2014).

21 He, X. *et al.* C-Myc-activated long noncoding RNA CCAT1 promotes colon cancer cell proliferation and invasion. *Tumour Biol*. *35*, 12181-12188 (2014).

22 Hung, C. L. *et al*. A long noncoding RNA connects c-Myc to tumor metabolism. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 18697-18702 (2014).

23 Pu, H. *et al*. CUDR promotes liver cancer stem cell growth through upregulating TERT and C-Myc. *Oncotarget* **6**, 40775-40798 (2015).

24 Hu, G., Lou, Z. & Gupta, M. The long non-coding RNA GAS5 cooperates with the eukaryotic translation initiation factor 4E to regulate c-Myc translation. *PLoS One* **9**, e107016 (2014).

25 Sawaya, A. P. *et al*. Topical mevastatin promotes wound healing by inhibiting the transcription factor c-Myc via the glucocorticoid receptor and the long noncoding RNA Gas5. *J. Biol. Chem*. (2017).

26 Prensner, J. R. *et al*. The long non-coding RNA PCAT-1 promotes prostate cancer cell proliferation through cMyc. *Neoplasia* **16**, 900-908 (2014).

27 Huang, J. *et al.* Linc-RoR promotes c-Myc expression through hnRNP I and AUF1. *Nucleic Acids Res.* **44**, 3059-3069 (2016).

28 Zhang, P., Cao, L., Fan, P., Mei, Y. & Wu, M. LncRNA-MIF, a c-Myc-activated long non-coding RNA, suppresses glycolysis by promoting Fbxw7-mediated c-Myc degradation. *EMBO Rep*. **17**, 1204-1220 (2016).

29 Barsyte-Lovejoy, D. *et al*. The c-Myc oncogene directly induces the H19 noncoding RNA by allele-specific binding to potentiate tumorigenesis. *Cancer Res*. **66**, 5330-5337 (2006).

30 Ma, M. Z. *et al.* Long non-coding RNA HOTAIR, a c-Myc activated driver of malignancy, negatively regulates miRNA-130a in gallbladder cancer. *Mol. Cancer*. **13**, 156-4598-13-156 (2014).

31 Kawasaki, Y. *et al.* MYU, a Target lncRNA for Wnt/c-Myc Signaling, Mediates Induction of CDK6 to Promote Cell Cycle Progression. *Cell. Rep.* **16**, 2554-2564 (2016).

32 Hu, T. & Lu, Y. R. BCYRN1, a c-MYC-activated long non-coding RNA, regulates cell metastasis of non-small-cell lung cancer. Cancer. *Cell. Int.* **15**, 36-015-0183-3. eCollection 2015 (2015).

33 Kim, T. *et al.* Role of MYC-regulated long noncoding RNAs in cell cycle regulation and tumorigenesis. *J. Natl. Cancer Inst.* **107**, 10.1093/jnci/dju505. Print 2015 Apr (2015).

34 Kim, T. *et al.* MYC-repressed long noncoding RNAs antagonize MYC-induced cell proliferation and cell cycle progression. *Oncotarget* **6**, 18780-18789 (2015).

35 Uphoff, C. C. & Drexler, H. G. Detection of mycoplasma in leukemia-lymphoma cell lines using polymerase chain reaction. *Leukemia* **16**, 289-293 (2002).

36 Swerdlow, S. in *WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues* (IARC, 2008).

37 Xue, X., Zeng, N., Gao, Z. & Du, M. Q. Diffuse large B-cell lymphoma: sub-classification by massive parallel quantitative RT-PCR. *Lab. Invest.* **95**, 113-120 (2015).

38 Danko, C. G. *et al*. Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. *Mol. Cell* **50**, 212-222 (2013).

39 Sturn, A., Quackenbush, J. & Trajanoski, Z. Genesis: cluster analysis of microarray data. *Bioinformatics* **18**, 207-208 (2002).

40 Subramanian, A. *et al*. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 15545-15550 (2005).

41 Chen, E. Y. *et al*. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128-2105-14-128 (2013).

42 Kuleshov, M. V. *et al*. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*. **44**, W90-7 (2016).

43 Seitz, V. *et al*. Deep sequencing of MYC DNA-binding sites in Burkitt lymphoma. *PLoS One* **6**, e26837 (2011).

44 Goecks, J., Nekrutenko, A., Taylor, J. & Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol*. **11**, R86-2010-11-8-r86. Epub 2010 Aug 25 (2010).

45 Blankenberg, D. *et al*. Galaxy: a web-based genome analysis tool for experimentalists. *Curr. Protoc. Mol. Biol*. **Chapter 19**, Unit 19.10.1-21 (2010).

46 Li, C. *et al*. Copy number abnormalities, MYC activity, and the genetic fingerprint of normal B cells mechanistically define the microRNA profile of diffuse large B-cell lymphoma. *Blood* **113**, 6681-6690 (2009).

47 Clark, M. B. & Mattick, J. S. Long noncoding RNAs in cell biology. Semin. *Cell Dev. Biol*. **22**, 366-376 (2011).

48 Cabili, M. N. *et al*. Localization and abundance analysis of human lncRNAs at single-cell and single-molecule resolution. *Genome Biol*. **16**, 20-015-0586-4 (2015).

49 ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).

50 Yang, F. *et al*. Long non-coding RNA GHET1 promotes gastric carcinoma cell proliferation by increasing c-Myc mRNA stability. *FEBS J*. **281**, 802-813 (2014).

51 Liao, L. M. *et al*. Low expression of long noncoding XLOC_010588 indicates a poor prognosis and promotes proliferation through upregulation of c-Myc in cervical cancer. *Gynecol. Oncol*. **133**, 616-623 (2014).

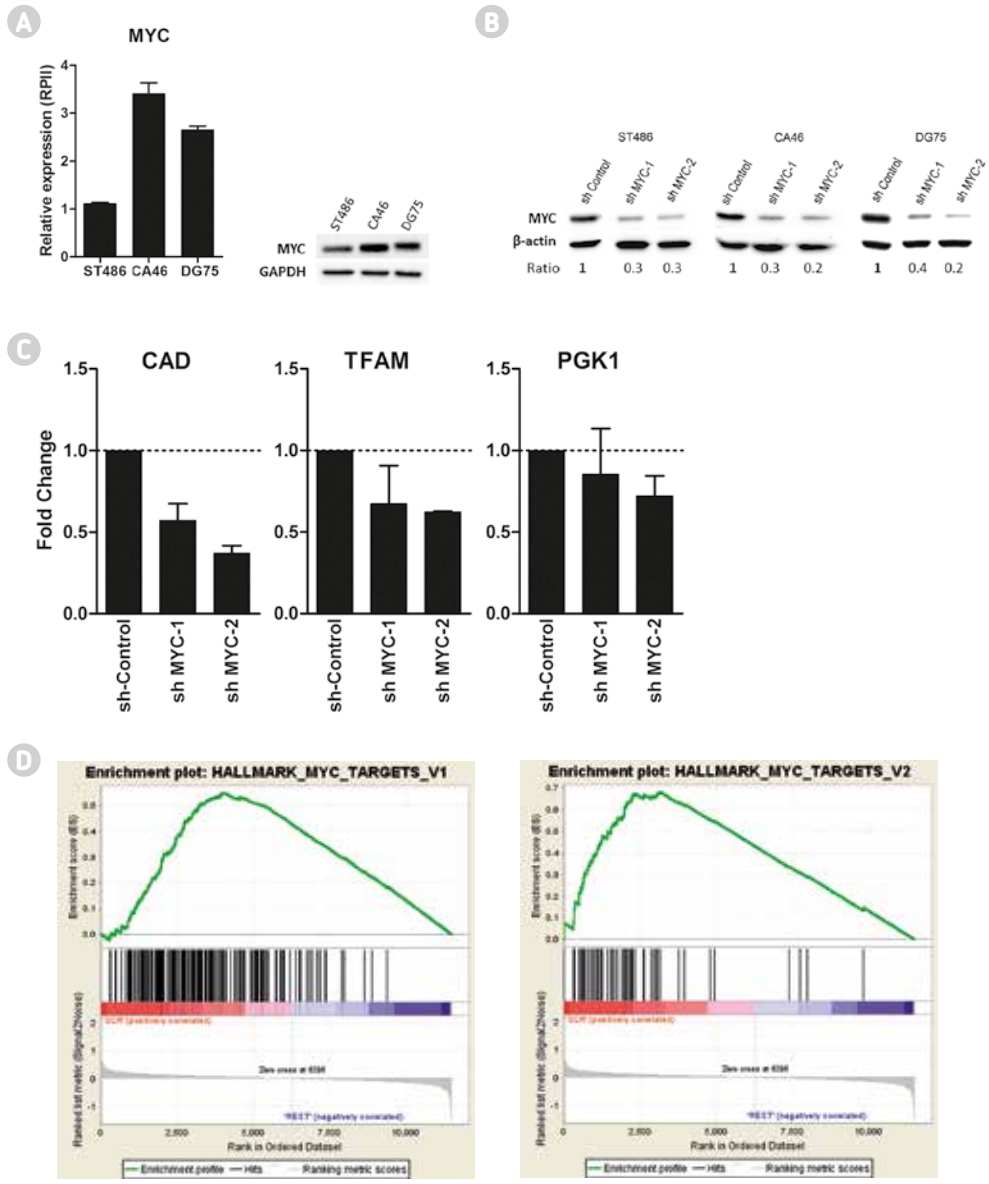

FIGURE S1 **Myc knockdown in Burkitt lymphoma cells reliably identifies Myc-regulated genes. (A)** Endogenous mRNA and protein Myc levels in three BL cell lines as determined by qPCR and Western blot, respectively. **(B)** A 60-80% reduction in Myc protein is achieved by infection of two individual anti-Myc shRNAs in three cell lines. A representative blot is shown. **(C)** Known Myc-induced target genes CAT, TFAM and PGK1 show the expected expression decrease upon Myc knockdown in ST486 cells. qRT-PCR; sh-control set to 1; mean ± error; n = 2. **(D)** Gene set enrichment analysis after microarray analysis shows strong enrichment of previously defined Myc-target hallmark gene sets. As expected, Myc-induced genes are enriched in the sh-Control (SCR) and depleted in the Myc knockdown cells (REST; shMYC-1 and shMYC-2). $p < 0.0001$, FDR < 0.0001.

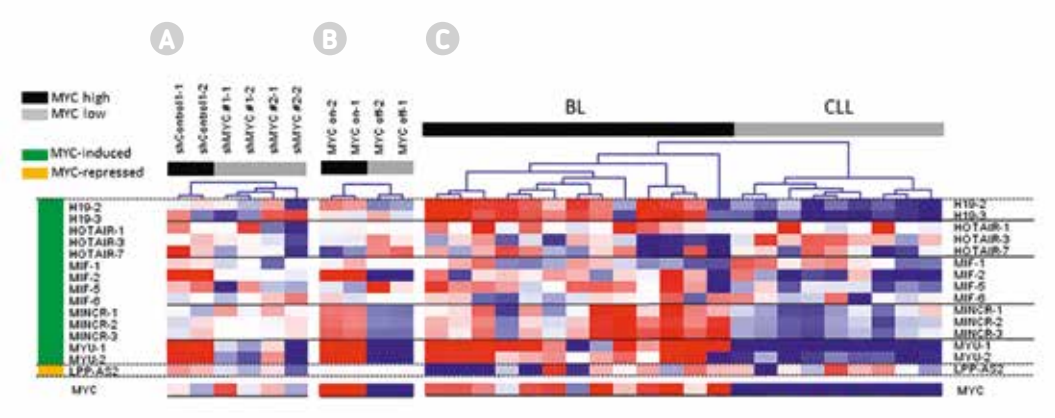

FIGURE S2 Expression of lncRNAs in the Myc pathway in B-cell lymphoma cell lines and patient samples with high or low Myc. Five Myc-induced and one Myc-repressed lncRNA are detected with one or multiple probes on the microarray. Expression differences in the following cell types are shown: **(A)** ST486 cells with shRNA-mediated MYC knockdown, **(B)** P493-6 cells untreated (MYC on) or tetracycline-repressed (MYC off), **(C)** BL cell lines and sorted normal GC B cells (I.e. cell of origin), (D) BL cases and CLL cases with high and low Myc levels, respectively. The expression level of MYC in BL and CLL patient cases is indicated below. For each lncRNA, all available probes expressed above background are shown.

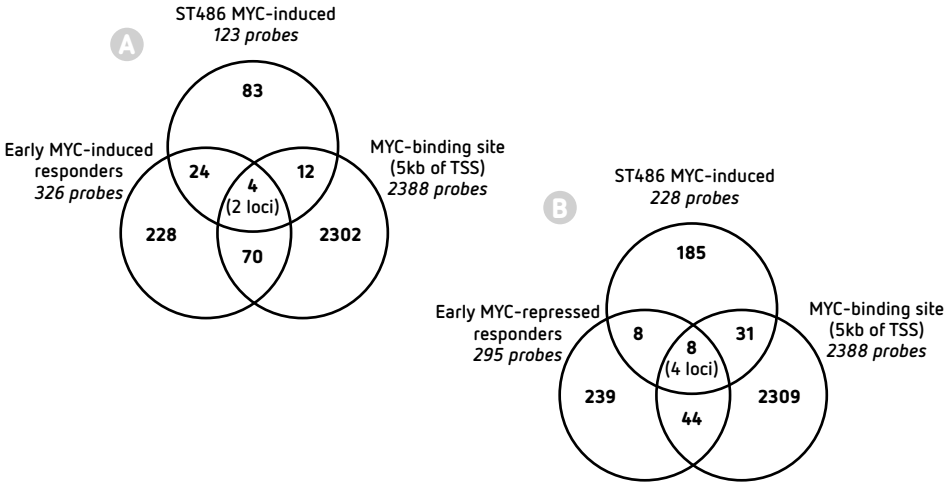

FIGURE S3 Candidate selection based on Myc-response kinetics and Myc occupancy. The intersection of three datasets is shown: (i) Myc-responsive lncRNA probes in ST486 cells, (ii) early responders (i.e. probes significantly deregulated within 4h after Myc induction in P493-6 cells), and (iii) lncRNA probes with a Myc binding site within 5 kb of the transcription start site (TSS) for **(A)** Myc-induced lncRNA targets and **(B)** Myc-repressed lncRNA targets.

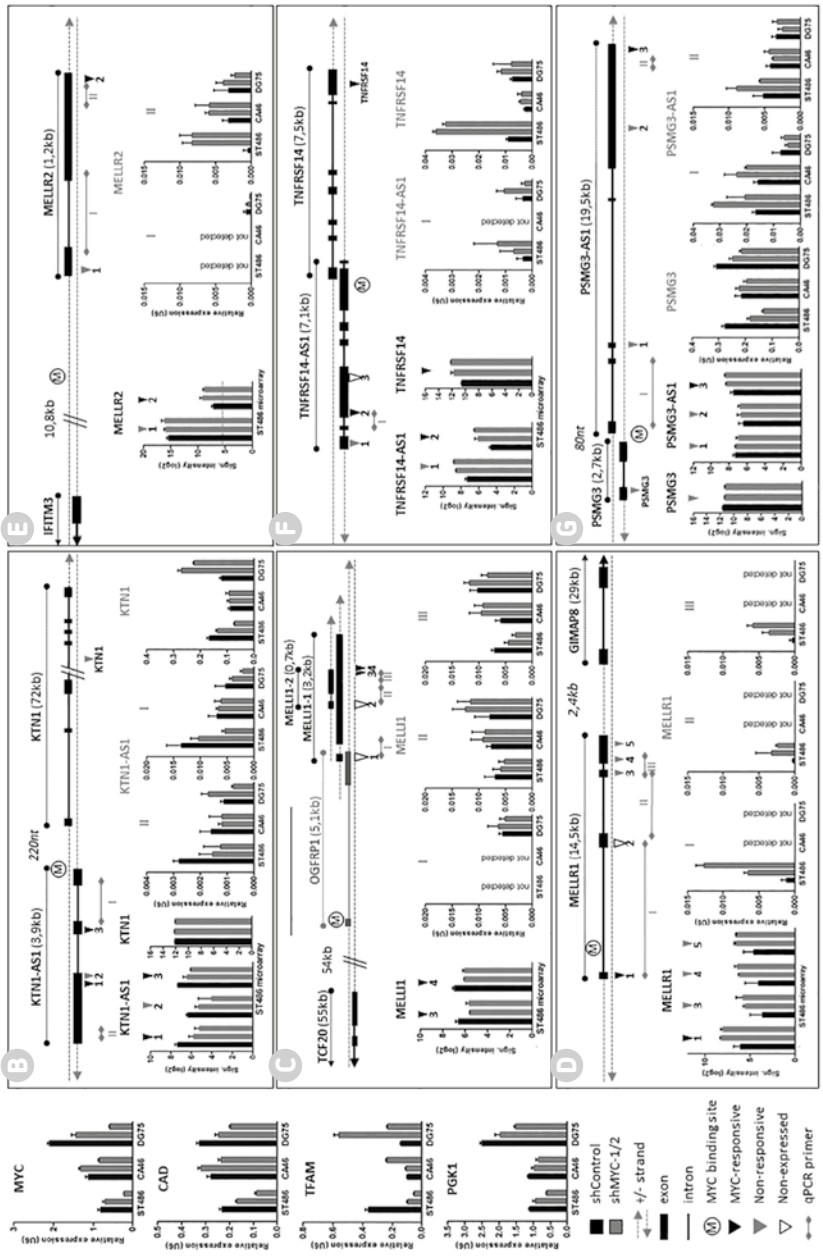

FIGURE S4 Validation of high confidence Myc-regulated lncRNAs in Myc knockdown experiments in three BL cell lines. (A) Expression decreases of *MYC* and at least two out of three Myc-induced target genes is confirmed at the mRNA level for all cell lines. qRT-PCR, mean±error. A schematic of the locus as well as array and qRT-PCR expression is shown for Myc-induced lncRNA candidates **(B)** KTN1-AS1 and **(C)** MELLI1, and Myc-repressed lncRNAs **(D)** MELLR1, **(E)** MELLR2, **(F)** TNFRSF14-AS1 and **(G)** PSMG3-AS1. Each schematic shows the closest coding gene, expression of neighboring genes within 10kb is assessed.

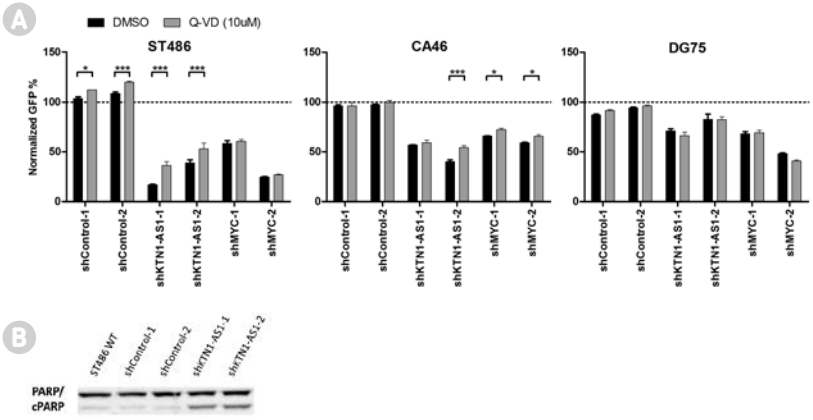

FIGURE S5 Caspase-mediated apoptosis is a minor contributor to the growth reduction phenotype observed upon KTN1-AS1 depletion. (A) ST486, CA46 and DG75 cells infected with non-targeting control or shRNAs targeting KTN1-AS1 or Myc and consequently treated with caspase ihbibitor Q-VD or DMSO. Differences between DMSO and Q-VD treated samples indicate a contribution of apoptosis to the observed growth decrease. Repeated measures ANOVA; Tuckey's Multiple comparison test; * $p < 0.05$; ** $p < 0.01$, *** $p < 0.001$. **(B)** An increased ratio of cleaved PARP over PARP indicates the presence of apoptosis in ST486 cells treated with shRNAs targeting KTN1-AS1. Western Blot, a representative blot is shown.

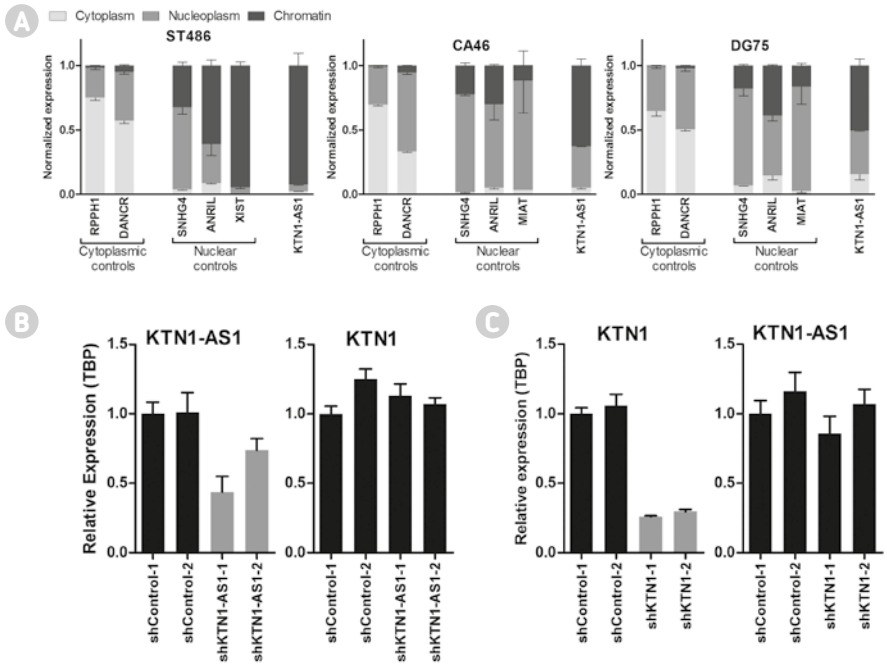

FIGURE S6 Despite a strong nuclear localization KTN1-AS1 depletion does not affect transcription of neighboring gene KTN1. (A) Subcellular fractionation was performed to study lncRNA localization in three BL cell lines. The indicated control lncRNAs confirm successful fractionation of cells being enriched in either the cytoplasm or nucleus as expected. KTN1-AS1 shows strong nuclear localization. **(B)** KTN1-AS1 knockdown does not affect mRNA levels of neighboring KTN1 or **(C)** *vice versa*. Knockdown efficacy (left) and expression of neighboring gene (right) in the same samples is indicated. qRT-PCR, mean ± error.

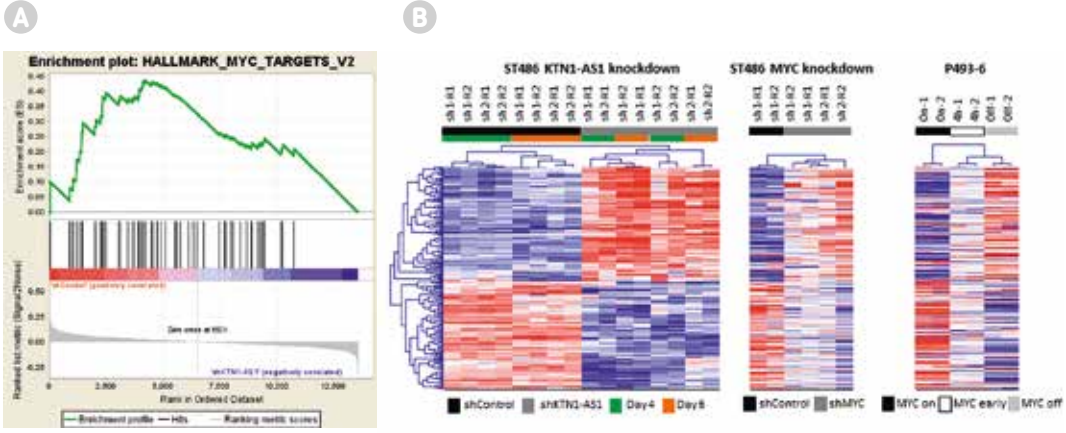

FIGURE S7 KTN1-AS1 depletion in ST486 cells affects Myc target genes. (A) Three Myc-target gene sets are found enriched in KTN1-AS1-regulated genes identified by lncRNA knockdown. GSEA. p = 0.002; FDR = 0.15. **(B)** Heatmap of KTN1-AS1 regulated probes also present on the previously used array version (259 of 312 KTN1-AS1 regulated probes) in ST486 cells with KTN1-AS1 depletion (*left*). The same genes are shown in Myc-depleted ST486 cells (*middle*) and in P493-6 cells in the presence and absence of Myc.

TABLE S1 Primer, probe and shRNA sequences.

Primer sequences	
U6	F: 5'-TGGAACGATACAGAGAAGATTAGCA-3' // R: 5'-AAAATATGGAACGCTTCACGAATT-3'
TBP	F: 5'GCCCGAAACGCCGAATAT--3' // R: 5'-CCGTGGTTCGTGGCTCTCT-3'
MYC	F: 5'-GCTCATTTCTGAAGAGGACTTGTTG-3' // R: 5'-TTACGCACAAGAGTTCCGTAGCT-3'
CAD	F: 5'-ACCTGTGGGGCTAATGACTG-3' // R: 5'-ACACCCAACACCACGTCAG-3'
TFAM	F: 5'-GTTTCTCCGAAGCATGTGG-3' // R: 5'-AGATGAAAACCACCTCGGTAAA-3'
PGK1	F: 5'-CAGCTGCTGGGTCTGTCAT-3' // R: 5'-GCTGGCTCGGCTTTAACC-3'
KTN1-AS1 I	F: 5'-GTTCAGTAGCCCCGCTCTC-3' // R: 5'-ACGCTCAGAAGATAGCCAGT-3'
KTN1-AS1 II	F: 5'-GCAAAGACACAAGGCTCACA-3' // R: 5'-ATGGTATTGGGGCACGTACA-3'
KTN1	F: 5'-CGCATCAAAGATTCCTGGCA-3' // R: 5'-TCCTGGTCATCGCTTCCATT-3'
MELLI1 I	F: 5'-TTCTCTTGGGTCTCCGGCAT-3' // R: 5'-GGTCCATCTCTGTCTTTCCAGT-3'
MELLI1 II	F: 5'-TGGGATTTGTCCTCAGGTCC-3' // R: 5'-CCAGAAGACAATCGCCTTGG-3'
MELLI1 III	F: 5'-CCCTGTGGCGGATTCATTAC-3' // R: 5'-CTCAGGACCCACTGACAGAC-3'
MELLR1 I	F: 5'-ACTCTCTGCCACATCGTTCG-3' // R: 5'-TCCTTCCGCTTCTTACGCTG-3'
MELLR1 II	F: 5'-CACGGCTTCATTCGCTTCA-3' // R: 5'-TTCTTCACCTTCTTAGCTGGC-3'
MELLR1 III	F: 5'-GGGCACTTTGTCTTACTTGGC-3' // R: 5'-AGGGGCATCCAAAGGTCTTG-3'
MELLR2 I	F: 5'-CTGTCAGGCCTCCGAGC-3' // R: 5'-GGAGAGTTGAGTGGCTTCCT-3'
MELLR2 II	F: 5'-TTCCTTCCTTTCGCCTCAGT-3' // R: 5'-TAGGCTGCGTAACAAATGGC-3'
TNFRSF14-AS1	F: 5'-GTGGCAACACAAAATGAACCG-3' // R: 5'-TCTTCAGACACAAGCACTCCC-3'
TNFRSF14	F: 5'-GCCTCGTCATCGTCATTGTTTG-3' // R: 5'-CACCTTCTGCCTCCTGTCTTT-3'
PSMG3	F: 5'-CACACGCCTCAAAGTCTTCC-3' // R: 5'-TGTGGACATTCAGCCAATGC-3'
PSMG3-AS1 I	F: 5'-TTAGAGCCGAACCCCGTTTT-3' // R: 5'-TTCACTGGTGTCCACTCACA-3'
PSMG3-AS1 II	F: 5'-TGTGCTCACCACAAAAGTCC-3' // R: 5'-GAGACACAAACGCTACCAGG-3'
tRNA Lys	F: 5'- CGGCTAGCTCAGTCGGTAGA -3' // R: 5'- CCAACGTGGGGCTCGAAC -3'
RPPH1	F: 5'-AGCTTGGAACAGACTCACGG-3' // R: 5'-AATGGGCGGAGGAGAGTAGT-3'
DANCR	F: 5'-CGTCTCTTACGTCTGCGGAA-3' // R: 5'-TGGCTTGTGCCTGTAGTTGT-3'
U3 SNORNA	F: 5'-AACCCCGAGGAAGAGAGGTA-3' // R: 5'-CACTCCCCAATACGGAGAGA-3'
ANRIL	F: 5'-AAGCCGCTCCGCTCCTCTTCT-3' // R: 5'-GCCGTGTCCAGATGTCGCGT-3'
MIAT	F: 5'-TGGAGGCATCTGTCCACCCATGT-3' // R: 5'-CCCTGTGATGCCGACGGGGT-3'
XIST	F: 5'-GTCCTTTCTTTTGACCCCAGAA-3' // R: 5'-GAGCCTGGCACTTTTTTTTCC-3'

Nanostring probes (selection)

R3HDM1	5'CCTGTGTTCCCAAGAGAATTACATTATTGACAAAAGACTCCAAGACGAGGATGCCAGTAGTACCCAGCA-GAGGCGCCAGATATTTAGAGTTAATAAAGAT-3'
WDR55	5'-CTACCTCTTCAATTGGAATGGCTTTGGGGCCACAAGTGACCGCTTTGCCCTGAGAGCTGAA TCTATCGACTGCATGGTTCCAGTCACCGAGAGTCTGCTG-3'
ISY1	5'-GGCAAAACATCAGTGTCTGTGGGTAGTTGGAATCTTCAGTTCCTGTGAGCGTCGGCGTCTT CTGGGCCTGTGGAGTTTCTTGGACAGGGGCCGCGGGGCT-3'
UBXN4	5'-CATCGCGACGGCCAAAAGGAGCGGCGCGGTCTTCGTGGTGTTCGTGGCAGGTGATGATGA ACAGTCTACACAGATGGCTGCAAGTTGGGAAGATGATAAA-3'
TRIM56	5'-GTGGAGGCCGAGGACATTTTCCTGAAGGGCAGGGGTTGGCAACTTTTCAACATGGAGTGC CAAACTGCTAACCCGTCTTCTAGTGTGTGAGAATAGGGAC-3'
KTN1-AS1	5'-AAGCAGAAGGAAGTTTCTACTTTGTTCATACAGTCTCAAGACATCCCTGGAGCCTTTCAA GGGAAATTTGGGCAGAAGTAGGTAGAGAGGTAAAGCTATT-3'
MELLR1	5'-CCGGTTCCTGCTGGTTCTGTGCCTTCTGGACTCATTTCCTTTCCAGGGTCCGCAACCGCTTC CGACACAATCACTATTTTGGCTGCGGCCACGGCTTCAT-3'

shRNA sequences

shMYC-1	S: 5'- GATCCGATGAGGAAGAAATCGATGTTCAAGAGACATCGATTTCTTCCTCATCTTTTTG-3' AS: 5'-AATTCAAAAAGATGAGGAAGAAATCGATGTCTCTTGAACATCGATTTCTTCCTCATCG-3'
shMYC-2	S: 5'- ATCCAACGACGAGAACAGTTGAAACATTCAAGAGATGTTTCAACTGTTCTCGTCGTTTTTTTG-3' AS: 5'- AATTCAAAAAAACGACGAGAACAGTTGAAACATCTCTTGAATGTTTCAACTGTTCTCGTCGTTG-3'
shKTN1-AS1-1	S: 5'-GATCCGCAGATCTAGAAGTCATTAGTTCAAGAGACTAATGACTTCTAGATCTGCTTTTTG-3' AS: 5'-AATTCAAAAAGCAGATCTAGAAGTCATTAGTCTCTTGAACTAATGACTTCTAGATCT GCG-3'
shKTN1-AS1-2	S: 5'-GATCCGAAGTAGGTAGAGAGGTAAATTCAAGAGATTTACCTCTCTACCTACTTCTTTTTG-3' AS: 5'-AATTCAAAAAGAAGTAGGTAGAGAGGTAAATCTCTTGAATTTACCTCTCTACCTACTTCG-3'
shKTN1-1	S: 5'-GATCCGAAAGTGACGCATCAAAGATTTCAAGAGAATCTTTGATGCGTCACTTTCTTTTTG-3' AS: 5'-AATTCAAAAAGAAAGTGACGCATCAAAGATTCTCTTGAAATCTTTGATGCGTCACTTTCG-3'
shKTN1-2	S: 5'-GATCCGTGGTAGATAAGAGAGAGGTTTCAAGAGAACCTCTCTCTTATCTACCACTTTTTG-3' AS: 5'-AATTCAAAAAGTGGTAGATAAGAGAGAGGTTCTCTTGAAACCTCTCTCTTATCTACCACG-3'

TABLE S2 **GSEA and Enrichr analyses of Myc-regulated coding genes.**

GSEA - Hallmarks

MYC-induced (p < 0.05; FDR < 0.25) (Top 5 of 10)	**P-value**	**FDR**
HALLMARK_E2F_TARGETS	<0.0001	<0.0001
HALLMARK_MYC_TARGETS_V2	<0.0001	<0.0001
HALLMARK_MYC_TARGETS_V1	<0.0001	<0.0001
HALLMARK_G2M_CHECKPOINT	<0.0001	0.0002
HALLMARK_MTORC1_SIGNALING	<0.0001	0.0550
MYC-repressed (p < 0.05; FDR < 0.25) (Top 5 of 11)	**P-value**	**FDR**
HALLMARK_INTERFERON_ALPHA_RESPONSE	<0.0001	<0.0001
HALLMARK_INTERFERON_GAMMA_RESPONSE	<0.0001	<0.0001
HALLMARK_TNFA_SIGNALING_VIA_NFKB	<0.0001	<0.0001
HALLMARK_IL6_JAK_STAT3_SIGNALING	<0.0001	0.0006
HALLMARK_KRAS_SIGNALING_UP	<0.0001	0.0007

Enrichr - Encode and ChEA consensus TFs

MYC induced (Top 5 of 17)	**Adj. p-value**
MYC_ENCODE	<0.0001
MAX_ENCODE	<0.0001
MYC_CHEA	<0.0001
E2F6_ENCODE	<0.0001
SIN3A_ENCODE	<0.0001
MYC repressed (Top 5 of 14)	**P-value**
SPI1_ENCODE	<0.0001
RELA_ENCODE	<0.0001
USF2_ENCODE	0.0001
USF1_ENCODE	0.0001
ZBTB7A_ENCODE	0.0007

Enrichr - KEGG pathways

MYC induced (Top 5 of 11)	**Adj. p-value**
Metabolic pathways_Homo sapiens_hsa01100	<0.0001
Biosynthesis of amino acids_Homo sapiens_hsa01230	0.0003
Arginine and proline metabolism_Homo sapiens_hsa00330	0.0025
Pyrimidine metabolism_Homo sapiens_hsa00240	0.0042
RNA transport_Homo sapiens_hsa03013	0.0051
MYC repressed	**P-value**
Lysosome_Homo sapiens_hsa04142	<0.0001
B cell receptor signaling pathway_Homo sapiens_hsa04662	0.0002
NF-kappa B signaling pathway_Homo sapiens_hsa04064	0.0003

Enrichr - Reactome pathways

MYC induced (Top 5 of 19)	***Adj. p-value***
Mitochondrial translation initiation_Homo sapiens_R-HSA-5368286	<0.0001
Metabolism_Homo sapiens_R-HSA-1430728	<0.0001
Mitochondrial translation_Homo sapiens_R-HSA-5368287	<0.0001
Gene Expression_Homo sapiens_R-HSA-74160	<0.0001
Mitochondrial translation termination_Homo sapiens_R-HSA-5419276	<0.0001
MYC repressed	**Adj. p-value**
N/A	N/A

TABLE S3 GSEA and Enrichr analyses of KTN1-AS1 knockdown affected coding genes.

GSEA - Hallmarks (p < 0.05; FDR < 0.25)

KTN1-AS1 induced	P-value	FDR
HALLMARK_MYC_TARGETS_V2	0.00184	0.1547
HALLMARK_WNT_BETA_CATENIN_SIGNALING	0.03162	0.1703
HALLMARK_UV_RESPONSE_UP	0.00559	0.2249
HALLMARK_ALLOGRAFT_REJECTION	0.01066	0.2136
KTN1-AS1 repressed	**P-value**	**FDR**
N/A	N/A	N/A

Enrichr - Encode and ChEA consensus TFs

KTN1-AS1 induced (Top 5 of 9)	Adj. p-value
MAX_ENCODE	0.00545
MYC_ENCODE	0.00545
UBTF_ENCODE	0.01035
NFYA_ENCODE	0.02225
BRCA1_ENCODE	0.02225
KTN1-AS repressed	
IRF8_CHEA	

Enrichr - KEGG pathways

KTN1-AS1 induced	Adj. p-value
Metabolic pathways_Homo sapiens_hsa01100	0.02094
Terpenoid backbone biosynthesis_Homo sapiens_hsa00900	0.01911
Steroid biosynthesis_Homo sapiens_hsa00100	0.01911
KTN1-AS1 repressed	**P-value**
Glycosphingolipid biosynthesis - ganglio series_Homo sapiens_hsa00604	0.01967

Enrichr - Reactome pathways

KTN1-AS1 induced	Adj. p-value
Cholesterol biosynthesis_Homo sapiens_R-HSA-191273	0.00010
Activation of gene expression by SREBF (SREBP)_Homo sapiens_R-HSA-2426168	0.02454
Metabolism_Homo sapiens_R-HSA-1430728	0.04688
KTN1-AS1 repressed	**Adj. p-value**
N/A	N/A