The MYCN oncoprotein and helicases DDX17 and DDX5 have opposite effects on the production of chimeric transcripts in neuroblastoma cells

Valentine Clerc, Jessica Valat, Xavier Grand, Nicolas Fontrodona, Matéo Bazire, Nicolas Rama, Didier Auboeuf, Benjamin Gibert, Franck Mortreux, Cyril Bourgeois

The MYCN oncoprotein and helicases DDX17 and DDX5 have opposite effects on the production of chimeric transcripts in neuroblastoma cells

Valentine Clerc[1], Jessica Valat[1], Xavier Grand[1,2], Nicolas Fontrodona[1], Matéo Bazire[1], Nicolas Rama[3], Didier Auboeuf[1], Benjamin Gibert[4], Franck Mortreux[1], Cyril F. Bourgeois[1]

[1.] Equipe labellisée "La Ligue", Laboratoire de Biologie et Modelisation de la Cellule, Ecole Normale Superieure de Lyon, CNRS, UMR 5239, Inserm, U1293, Universite Claude Bernard Lyon 1, 46 allee d'Italie F-69364 Lyon, France

[2] INSERM U1052- Cancer Research Center of Lyon (CRCL), F-69008 Lyon, France

[3] Apoptosis, Cancer and Development Laboratory- Equipe labellisée "La Ligue", LabEx DEVweCAN, Institut Convergence PLAsCAN, Centre de Recherche en Cancerologie de Lyon, INSERM U1052-CNRS 5286, Université Claude Bernard Lyon 1, F-69008 Lyon, France.

[4] Gastroenterology and technologies for health, Centre de Recherche en Cancerologie de Lyon, INSERM U1052-CNRS5286, Université Lyon 1, 69008, Lyon, France

Corresponding author: Cyril F. Bourgeois (cyril.bourgeois@inserm.fr).

ORCID number: 0000-0002-0756-5501

Abstract

DEAD box helicases DDX17 and DDX5 control transcription termination and the associated processing of the 3' end of transcripts. Here we show that the transcriptional readthrough induced by their depletion in neuroblastoma cells results in increased production of chimeric transcripts from tandemly orientated genes. Analysis of neuroblastoma tumours in which chimeric transcripts are abundant revealed that low expression of DDX17 and DDX5 genes is associated with high-risk tumours and poor overall patient survival, and inversely linked with *MYCN* oncogene amplification. We demonstrate that changes in MYCN expression do not affect the expression of either helicase, but alter transcription termination leading to the production of chimeric transcripts. MYCN acts on termination through its direct binding to the 3' region of genes and it interacts with DDX17, suggesting that it may inhibit the activity of the helicase. Collectively, our work reveals a novel function of MYCN in transcription termination and suggests that the deregulation of MYCN and DDX17/DDX5 expression in neuroblastoma may lead to the expression of non-canonical and potentially harmful RNA molecules.

Introduction

During transcription by RNA polymerase II (RNAPII) in eukaryotes, the nascent RNA undergoes different processing steps, including 5' capping, splicing, 3' end cleavage and addition of the poly-A tail. These steps are mostly co-transcriptional and are impacted not only by RNA binding proteins, but also by transcription and chromatin-associated factors (Neugebauer, 2019, Tellier et al., 2020). The tight connection that exists between transcription termination and 3' end RNA processing is archetypal of this coordination of events. The slow down and release of RNAPII after the transcription of the polyadenylation site (PAS) and the cleavage of the nascent transcript are mutually dependent on each other, in a complex interplay (Boreikaite and Passmore, 2023, Porrua and Libri, 2015, Proudfoot, 2016, Rodriguez-Molina et al., 2023). The regulation of this process involves a number of factors that act on chromatin, at the level of the RNAPII complex or on the RNA molecule, or which can have an effect at several levels.

An alteration in the expression or activity of factors involved in transcription termination and/or RNA cleavage classically results in transcriptional readthrough and in the extension of the 3' extremity of the transcript. Readthrough transcription has been observed in cells exposed to various stresses or viral infection, as well as in cancer (Cardiello et al., 2018, Grosso et al., 2015, Rutkowski et al., 2015, Heinz et al., 2018, Rosa-Mercado et al., 2021, Vilborg et al., 2015, Vilborg et al., 2017, Hennig et al., 2018, Bauer et al., 2018), as recently reviewed (Rosa-Mercado and Steitz, 2022, Morgan et al., 2022). The fate of these extended transcripts is uncertain, but it has been shown that stress-induced DoGs

(downstream of gene containing transcripts) lack a defined endpoint and are non-coding RNAs which remain associated to the chromatin in the nucleus (Vilborg et al., 2015, Vilborg et al., 2017).

In some cases, RNAPII that fails to terminate past the PAS can invade the downstream gene on the same DNA strand, generating *cis*-spliced transcripts containing exons from both genes. These RNA molecules have been called transcription-induced chimeras (Akiva et al., 2006, Parra et al., 2006), tandem RNA chimeras (Greger et al., 2014), products of conjoined genes (Prakash et al., 2010, Kim et al., 2012), readthrough fusions (Varley et al., 2014) or *cis*-splicing products of adjacent genes (Qin et al., 2016, Qin et al., 2015). Below we will simply define them as chimeric transcripts, keeping in mind that in this study we are not considering RNA molecules resulting from chromosomal rearrangements. The formation of chimeric transcripts can be associated to splicing defects in the adjacent invaded gene, highlighting the necessary coordination of events that ensures the correct and autonomous expression of neighbouring genes (Alpert et al., 2020, Hadar et al., 2022). Thanks to the increasing number and deeper analysis of RNA-seq experiments, chimeric transcripts have been identified in many types of cancers and associated with oncogenesis (Dorney et al., 2023, Sun and Li, 2022). However, they are also found in normal tissues (Babiceanu et al., 2016) and their production is considered as a possible mechanism to increase protein diversity (Parra et al., 2006). Indeed, chimeric transcripts can retain a functional open reading frame (ORF) allowing the production of chimeric or fusion proteins (Varley et al., 2014, Yun et al., 2014, Han et al., 2017, Egashira et al., 2019).

Several factors have been shown to affect the formation of chimeric transcripts in human cells, including splicing and/or polyadenylation factors (Chwalenia et al., 2019). Chemical splicing inhibition represses the formation of chimeric transcripts in neuroblastoma, a cancer in which most RNA fusions are intrachromosomal, likely resulting from transcriptional readthrough and *cis*-splicing (Shi et al., 2021). Chromatin-associated factors are also involved in chimeric RNA production, such as the histone methyltransferase SETD2 (Grosso et al., 2015) or the CCCTC-binding factor CTCF (Qin et al., 2016, Qin et al., 2015), which plays important roles in spatial genome organization and gene expression regulation via chromatin looping or insulation (Braccioli and de Wit, 2019). Recent reports further supported a function of CTCF in controlling PAS choice and transcription termination (Nanavaty et al., 2020, Terrone et al., 2022).

DEAD box ATP-dependent RNA helicase DDX17 and its paralog DDX5 have a variety of functions in gene expression, especially in transcription and RNA processing (Giraud et al., 2018, Xing et al., 2019). Thanks to their helicase activity, they act by modulating local RNA secondary structures or DNA/RNA structures such as R-loops and G-quartets, impacting promoter and exon choice (Camats et al., 2008, Kar et al., 2011, Lai et al., 2019, Mersaoui et al., 2019, Wu et al., 2019). Recently, we and others have shown that DDX17 and DDX5 control the 3' end processing of RNAPII transcripts, and that their knockdown leads to a decreased termination downstream of the expected PAS (Katahira et al., 2019,

Lai et al., 2019, Mersaoui et al., 2019, Terrone et al., 2022). We showed that this function is linked to the binding of CTCF near the PAS and to the specific 3D organization of their target genes (Terrone et al., 2022). We hypothesized that the helicases could control the dynamics of transcription and RNA processing across DNA and/or RNA structured regions, in line with other reports that described an impact on R-loops or 3'UTR RNA structure (Katahira et al., 2019, Lai et al., 2019, Mersaoui et al., 2019). We show hereafter that the transcription readthrough induced by *DDX17* and *DDX5* silencing in neuroblastoma cells results in the production of several hundreds of chimeric transcripts, a fraction of which is also found in neuroblastoma tumours. Interestingly, we found that high *DDX17* and *DDX5* expression correlates with good survival of neuroblastoma patients and that it is inversely correlated to the amplification of *MYCN*, a driving oncogene in this cancer (Otte et al., 2020). Like the closely related MYC protein, MYCN is a transcription factor which binds to most active promoters and regulates several aspect of RNAPII activity (Baluapuri et al., 2020). Here, we present evidence supporting a direct role of MYCN in controlling transcription termination through its binding near the PAS of DDX17/DDX5-targeted genes. This new function of MYCN suggests that the combined amplification of this oncogene and suboptimal expression of *DDX17* and *DDX5* could be determinant to explain the increased expression of chimeric transcripts in neuroblastomas.

Results

DDX17 and DDX5 depletion enhances the expression of chimeric transcripts

We recently described a genome-wide effect of DDX17/DDX5 depletion on transcription termination in neuroblastoma cells SH-SY5Y cells (Terrone et al., 2022). In our RNA-seq results, we observed numerous examples of transcriptional readthrough induced by DDX17/DDX5 depletion that also displayed splicing junction reads across adjacent genes. Typically, splicing skipped the terminal exon of the altered gene and linked its penultimate exon to the second exon of the following gene, generating a chimeric mRNA molecule (Fig. 1A, Supp. Fig. 1A). However, we also frequently observed complex alternative splicing patterns across exons from both genes (Supp. Fig. 1B). Interestingly, in some cases chimeric junctions linked more than 2 adjacent genes, suggesting that DDX17/DDX5 depletion could induce the formation of multimeric mRNA molecules (*NTRK1-PEAR1-LRRC71* and *NDUFA13-YJEFN2-CILP2*, Supp. Fig. 1B).

We next used the previously described Arriba algorithm (Uhrig et al., 2021) to determine more precisely the number of chimeric transcripts whose expression is modified upon DDX17/DDX5 depletion. We identified 282 chimeric mRNAs, 218 of which (77%) were upregulated and 64 downregulated (Fig. 1B, top panel, Supp. Table 1). Applying a more stringent selection threshold confirmed that the depletion of both helicases indeed mostly increased the expression of chimeric

transcripts (Fig. 1B, bottom panel). We found that a large proportion of the genes displaying an induced chimeric transcript had been previously identified as genes with readthrough induced by siDDX17/DDX5 (Fig. 1C) (Terrone et al., 2022). From these analyses, we defined a list of 953 genes whose termination is inhibited upon depletion of DDX17 and DDX5 (Supp. Table 1).

We next sought to experimentally validate the expression of a selection of chimeric transcripts. We first carried out RT-PCR experiments using a primer overlapping the RNA fusion between both genes, to specifically detect the chimeric transcript, that we systematically compared to the canonical form of the gene. We observed an increased amount of all chimeric transcripts tested upon DDX17/DDX5 depletion, which was not consistent with the expression of the corresponding canonical mRNA (Supp. Fig. 2A). We confirmed this result by RT-qPCR on a panel of chimeric transcripts, demonstrating that the normalized expression of all molecules was increased between 2 and 40-fold in absence of DDX17/DDX5, compared to the canonical transcript (Fig. 1D and Supp. Fig. 2B). Note that validated *CTSD-IFITM10*, *NDUFA13-YJEFN2* and *TRIM3-HPX* chimeric RNAs were not predicted as significantly induced by Arriba but were identified while browsing through our RNA-seq data (Supp. Fig. 1A-B). This suggested that the actual number of chimeric RNAs resulting from DDX17/DDX5 depletion is likely underestimated.

Chimeric transcripts can produce chimeric proteins

Some chimeric transcripts maintain a coding sequence that matches the reading frame of their respective canonical transcripts, potentially allowing them to escape degradation by the nonsense-mediated mRNA decay pathway. For example, the *CTSD-IFITM10* chimer lacks the last exon of *CTSD* and the first exon of *IFITM10*, but it is cleaved and polyadenylated at the canonical PAS of the *IFITM10* gene. This chimeric mRNA is predicted to produce a chimeric protein of 557 amino acids (60 kDa), which lacks 55 and 28 amino acids from the C-terminal part of CTSD and N-terminal part of IFITM10 proteins, respectively (Fig. 2A), and which was identified in breast cancer cells (Varley et al., 2014).

We carried out a western-blot analysis to detect this chimeric protein in neuroblastoma cells, using an antibody against CTSD. The *CTSD* gene normally encodes a unique precursor that is processed into two peptidic chains that compose Cathepsin D. As expected, we identified the two subunits between 25 and 45 kDa (Fig. 2B), but we also detected a discrete 60 kDa band that appeared only upon depletion of DDX17 and DDX5. All bands disappeared when cells were treated with a siRNA targeting the body of the *CTSD* transcript. Furthermore, treating cells with another siRNA specifically targeting the *CTSD-IFITM10* junction also led to a loss of the 60 kDa band. Note that this treatment also reduced the level of CTSD isoforms, likely because of a non-specific effect of the junction siRNA on canonical mRNAs. This result demonstrated that the siDDX17/DDX5-induced chimeric mRNAs can produce chimeric proteins, which could have a strong impact on the functions of their respective canonical proteins.

The expression of the *DDX17* gene is reduced in high-risk neuroblastomas

Recently, neuroblastoma tumours were shown to express large amounts of chimeric transcripts (Shi et al., 2021), which prompted us to look whether DDX17/DDX5-regulated chimeras can be found in these tumours. Of the 97 main chimeric transcripts from our analysis, 25 were also identified in neuroblastoma tumours (Fig. 3A, top panel). All these mRNAs belong to the pool of chimeric transcripts upregulated upon DDX17/DDX5 depletion, while none of the downregulated transcripts was identified (the top 10 transcripts are shown in Fig. 3B, all transcripts are shown in Supp. Table 1). Of note, only 1 of these 25 transcripts (*SLC29A1-HSP90AB1*) was also detected in a control cohort of 161 healthy adrenal gland samples, while 4 of them were detected in at least 10% of neuroblastomas (Shi et al., 2021) (Fig. 3A, bottom panel and Fig. 3B).

Our earlier work showed that DDX17 is involved in the early phases of retinoic acid-induced differentiation of neuroblastoma cells (Lambert et al., 2018), and recent findings have associated a high expression DDX5 with an unfavourable outcome in neuroblastoma patients (Zhao et al., 2020). To explore further the link between the two helicases and this cancer, we analysed the expression of *DDX17* and *DDX5* genes in the clinically annotated SEQC neuroblastoma cohort (Zhang et al., 2015). We found that a high expression of *DDX17* was significantly correlated with a better survival probability of patients (Fig. 3C, left panel). This was true also when the Kaplan-Meyer analysis was performed only on the most aggressive stage 4 patients (according to the International Neuroblastoma Staging System INSS) (Fig. 3C, right panel). Specifically, *DDX17* expression level is significantly lower in stage 4 tumours than in any other neuroblastoma stage (Fig. 3D), and is also significantly associated with high-risk tumours (Fig. 3E). Finally, we monitored *DDX17* expression according to the amplification of the *MYCN* oncogene, which is a well-established marker of poor prognostic in neuroblastoma (Otte et al., 2020), and we found a negative association between these two parameters (Fig. 3F). The same analysis showed a more contrasting result for the *DDX5* gene, but a low *DDX5* expression was also found to correlate with a lower survival probability and with *MYCN* amplification (Supp. Fig. 3).

In conclusion, those results established a link between the expression of *DDX17*, and to a lower extent of *DDX5*, and the severity of neuroblastoma. This raised the interesting possibility that a suboptimal expression of DDX17 and DDX5, which leads to transcription termination defects of hundreds of genes in cultured cells, could favour the abnormal expression of a subset of chimeric transcripts in aggressive neuroblastomas.

MYCN increases transcriptional readthrough and chimeric transcript formation

The negative link between *MYCN* amplification and a lower expression of both *DDX17* and *DDX5* genes suggested that MYCN could control the expression of both helicases, but also that chimeric transcript

expression could be increased upon MYCN overexpression. To test this double hypothesis, we used two different neuroblastoma cell lines with and without an amplification of *MYCN* to modulate the expression of the oncogene.

First, we transfected SH-SY5Y cells (no *MYCN* amplification) with increasing amounts of an HA-tagged MYCN-encoding plasmid. MYCN protein was barely detectable in control cells and strongly induced upon transfection, but this did not result in any significant change in DDX17 and DDX5 protein levels (Fig. 4A). Under these conditions, we observed an increased production of chimeric transcripts (Fig. 4B, Supp. Fig. 4). MYCN also induced transcription readthrough beyond the 3' end of genes that we had previously identified as regulated by DDX17 and DDX5 (Terrone et al., 2022) (Fig. 4B right panel, Supp. Fig. 4).

In parallel, we carried out the opposite experiment and downregulated MYCN expression in the *MYCN*-amplified SK-N-BE(2) cell line, using a mixture of specific siRNAs which reduced MYCN protein level by nearly 60%. Under these conditions, neither DDX17 nor DDX5 protein level was altered (Fig. 4C). However, we observed a significant reduction in the expression of previously tested chimeric transcripts, as well as reduced transcriptional readthrough (Fig. 4D, Supp. Fig. 4). Because steady-state levels of the corresponding canonical transcripts remained largely unaffected (Supp. Fig.4), our results suggested that variations in MYCN expression altered transcription termination and as a consequence, chimeric RNA production, in a manner that is distinct from its promoter-associated activity.

MYCN binds near the 3' end of DDX17/DDX5-regulated genes and interacts with DDX17

To test this hypothesis of a direct effect of MYCN on transcription termination, we re-analysed previously published ChIP-seq datasets performed in several neuroblastoma cell lines (Buchel et al., 2017, Upton et al., 2020, Zeid et al., 2018). We found that the 3' region of genes displaying transcriptional readthrough or chimeric RNA production upon DDX17/DDX5 knock-down or MYCN overexpression generally exhibited at least one MYCN peak, often detected across several cell lines (Fig. 5A and Supp. Fig. 5). A global analysis of all DDX17/DDX5-regulated genes confirmed this observation, as the chromatin region overlapping their terminal exon was significantly enriched in MYCN binding sites, compared to non-regulated genes (Fig. 5B). We then performed ChIP-qPCR in SK-N-BE(2) cells and validated the binding of MYCN near the 3' end of our model genes (Fig. 5C). Together with results of Figure 4 showing that MYCN does not affect DDX17 and DDX5 expression, these experiments strongly suggested that the effect of MYCN on transcription termination is direct.

We next hypothesized that through its binding to chromatin at the 3' end of DDX17/DDX5 target genes, MYCN could interfere with the function of helicases. Therefore, we tested whether MYCN could interact with DDX17 and DDX5 by carrying out co-immunoprecipitation (co-IP) assays in SK-N-BE(2) cells. As shown in Fig. 5D, endogenous DDX17 and MYCN efficiently co-immunoprecipitated, whatever

protein was targeted. The co-IP was even enhanced when extracts were treated with benzonase, excluding the contribution of nucleic acids and suggesting a direct interaction between the two proteins. In contrast, no co-IP was detected between MYCN and DDX5 (Fig. 5E), which underlines the specificity of the MYCN-DDX17 interaction.

Altogether these results indicate that MYCN impacts transcription termination in a direct manner, possibly by inhibiting the normal function of DDX17.

The forced recruitment of MYCN near the termination region is sufficient to induce the formation of chimeric transcripts.

We sought to provide a direct proof that MYCN binding near the termination region of a gene is sufficient to promote transcriptional readthrough and the subsequent production of a chimeric transcript. We designed an experiment in SH-SY5Y cells (in which endogenous MYCN level is very low) in which we could force the recruitment of MYCN downstream of the PAS of the *MAPK7* gene. For this, we co-transfected cells with two plasmids, one expressing MYCN protein fused to a catalytically inactive "dead" Cas9 protein (dCas9-MYCN) and the other expressing specific RNA guides to target dCas9-MYCN downstream of the *MAPK7* gene (Fig. 6A). We verified by ChIP-qPCR that dCas9-MYCN was indeed bound at its target site upon transfection of RNA guides (Fig. 6B). A plasmid expressing dCas9-GFP fusion protein was used as a control to rule out a possible non-specific action of the dCas9 fusion protein on transcription through steric hindrance.

We then quantified the expression of the *MAPK7* gene, which remained unaffected by the recruitment of either dCas9-MYCN or dCas9-GFP, as did the control *ADCK4* gene (Fig. 6B, left panel). In contrast, the chimeric *MAPK7-RNF112* chimer was induced only upon the binding of the dCas9-MYCN between the two genes while the *ADCK4-NUMBL* chimer remained unchanged (Fig. 6B, right panel).

This experiment demonstrated that the binding of MYCN near the termination site can inhibit the normal termination process and promote the formation of a chimeric transcript.

Discussion

In this study, we show that the transcription readthrough induced by depletion of DDX17 and DDX5 helicases also leads to the production of chimeric transcripts in neuroblastoma cells. Among the most significantly induced chimeric transcripts, 27% were previously found in neuroblastoma tumors, and not in non-tumoural adrenal gland tissue (Shi et al., 2021). This suggests that this subset of chimeric transcripts might be specific of neuroblastomas and could potentially be used as biomarkers for this cancer. Further studies will be required to determine their level of specificity, as chimeric transcripts have been identified in many types of cancers (Dorney et al., 2023, Sun and Li, 2022).

If production of chimeric transcript in healthy cells is considered as a possible mechanism to increase protein diversity (Parra et al., 2006), their abnormal expression in cancer cells can have several consequence. First, a fraction of chimeric transcripts can retain a functional ORF allowing the translation of chimeric or fusion proteins (Varley et al., 2014, Yun et al., 2014, Han et al., 2017, Egashira et al., 2019), and we validated the production of CTSD-IFITM10 protein upon DDX7/DDX5 depletion. Chimeric proteins can generate highly immunogenic neo-antigens, which may be unique to the tumor and could be targeted by the immune system (Weber et al., 2022). A recent study identified a novel chimeric transcript produced by gene fusion and demonstrated that it produces a neo-antigen that can specifically elicit a host cytotoxic T-cell response in metastatic head and neck cancer (Yang et al., 2019). Chimeric proteins produced from readthrough induced transcripts may be much less abundant than those produced from a gene fusion event, but depending on their expression level, they could alter the protein interaction network and activate cell growth pathways by acting as oncoproteins, or alter the cell response to antitumour treatments. Second, even if not translated, chimeric transcripts can act as functional long noncoding RNAs (lncRNA). For example, *SLC45A3-ELK4* chimeric transcripts act as lncRNAs and regulate cell growth in prostate cancer, and its higher expression level has been associated with disease progression and metastasis (Qin et al., 2017, Zhang et al., 2012). Future analyses will determine whether the chimeric transcripts we have identified have protumoral properties or can influence tumour response to treatment.

Our results suggest that the production of chimeric transcripts in neuroblastomas could be favoured by a reduced expression of *DDX17* and *DDX5*, especially in high-risk tumours. Indeed, a low expression of both helicases is significantly associated with higher risk and shorter patient survival. Note that our results differ from a previous report showing a link between high *DDX5* expression and poor survival of patients with neuroblastoma (Zhao et al., 2020). However, we used a much larger and more detailed cohort of tumours (n=498 versus n=42), which allowed us to stratify the expression of helicases across all neuroblastoma stages. Importantly, the fact that high expression of *DDX17* and *DDX5* is associated with better survival when looking at stage 4 patients only, reduces the possibility that it is directly and solely linked to *MYCN* amplification, which is a less predominant risk factor in this class of tumours. It will be important to confirm our results at the protein level, and to determine whether variations in the expression of both helicases contribute to the mechanism of oncogenesis in neuroblastoma.

Interestingly, neuroblastoma is only one of few examples of cancers in which a high expression of *DDX17* and/or *DDX5* is associated with good patient survival, among many studies describing their positive contribution to major cancer signaling pathways and cancer development (Fuller-Pace and Moore, 2011, Xu et al., 2022). This suggests that the two helicases play contrasted roles depending probably on the cancer type and on the context in which they are studied. Beside neuroblastoma, high *DDX17* expression and high *DDX5* expression are respectively associated with a good prognosis in

breast cancer (Wortham et al., 2009) and lung adenocarcinoma (Zhou et al., 2024), and with pancreatic ductal adenocarcinoma (Morimachi et al., 2021) and hepatocellular carcinoma (Zhang et al., 2016a, Zhang et al., 2019). In addition to their multiple roles in controlling gene expression and RNA processing, for example by repressing readthrough transcription, DDX5 and DDX17 are also involved in genomic stability and DNA repair (Bader et al., 2022, Cargill et al., 2021), which may also explain how their low expression can impact high-risk tumours.

We show that MYCN, a transcription factor with a major oncogenic effect in neuroblastoma, promotes transcriptional readthrough and the production of chimeric transcripts, without altering the expression of either helicase. Our results indicate that this new function of the oncoprotein is direct, yet the mechanism by which it inhibits termination is still unclear. MYCN co-immunoprecipitates with DDX17 and binds preferentially downstream of DDX17/DDX5-regulated genes, suggesting that it may interfere with the activity of the helicases. We showed previously that those genes present a looped 3D conformation and that their transcription termination of is also controlled by CTCF (Terrone et al., 2022), so it will be interesting to test whether MYCN overexpression impacts gene looping and CTCF binding. Supporting this idea, MYCN and CTCF bind to hundreds of overlapping intergenic sites (Lee et al., 2012, Buchel et al., 2017), and MYC was recently shown to affect CTCF binding (Wei et al., 2023). Another hypothesis is that MYCN could prevent termination at the end of the gene by activating RNAPII, as MYCN or MYC do to induce promoter escape and productive elongation (Baluapuri et al., 2019, Herold et al., 2019, Papadopoulos et al., 2022). Alternatively, MYCN interacts and recruits multiple chromatin remodeling complexes, whose ectopic recruitment to the 3' end of gene may alter the local chromatin state and affect RNAPII kinetics and interactome and disturb the chain of events leading to termination.

Future work will help to determine which of these non-mutually exclusive hypotheses is correct, or if another mechanism is involved. Nevertheless, our study is a new illustration of the diversity of actions that a transcription factor can have upon its binding to non-promoter sites. To summarize, our results show that MYCN overexpression or a suboptimal expression of either DDX17 or DDX5, two parameters associated to poor survival of neuroblastoma patients, induce transcription readthrough of a subset of common genes and promote the production of chimeric transcripts. Understanding the interplay between these factors is of great importance to appreciate their impact on tumorigenesis.

Materials and Methods

Plasmids and cell culture

Human SH-SY5Y and SK-N-BE(2) cells (ECACC) were grown and transfected essentially as described previously (Lambert et al., 2018). For standard silencing experiments, 20 nM of siRNA (sequences in Supp. Table 2) were mixed with Lipofectamine™ RNAiMax (ThermoFisher Scientific) following the manufacturer's instructions and cells were harvested 48 h after transfection. For MYCN overexpression, cells were plated in 6-well-plates to reach 70% confluence, and then transfected with 0.5 to 1 µg pCDNA3-HA-hMYCN (Addgene #74163) using JetPrime® (Polyplus Transfection) following the manufacturer's instructions. Cells were harvested 48 h after transfection.

For dCas9-MYCN experiments, the hMYCN cDNA was subcloned into the dCas9-empty-GFP plasmid (a gift from Reini F. Luco, Institut Curie, Orsay, France). This cloning step was done by RD Biotech. We cloned the sequences corresponding to the two RNA guides into the *Bsm*BI site of the CRIZI plasmid (provided by Philippe Mangeot, CIRI, Lyon, France). SH-SY5Y cells were plated in 6-well-plates to reach 70% confluence and transfected with 1 µg of dCas9-HA-MYCN plasmid and 1 µg of sgRNA-containing plasmid (500 ng of each guide, sequences in Supp. Table 2) using jetPRIME (PolyPlus Transfection).

Western Blot and co-immunoprecipitation

Protein extraction and western blotting were carried out as previously described (Dardenne et al., 2014). Primary antibodies used for western-blotting were: DDX5 (ab10261, Abcam), DDX17 (ab24601, Abcam), GAPDH (sc-32233, Santa Cruz Biotechnology), MYCN (sc-53993, Santa Cruz Biotechnology), CTSD (21327-1-AP, Proteintech). For co-immunoprecipitation, SH-SY5Y cells were harvested and gently lysed for 5 min on ice in a buffer containing 10 mM Tris-HCl pH 8.0, 140 mM NaCl, 1.5 mM $MgCL_2$, 10 mM EDTA, 0.5% NP40, completed with protease and phosphatase inhibitors (Roche #11697498001 and #5892970001), to isolate the nuclei from the cytoplasm. After centrifugation, nuclei were lysed in IP buffer (20 mM Tris-HCl pH 7.5, 150 mM NaCl, 2 mM EDTA, 1% NP40, 10% glycerol and protease/phosphatase inhibitors) for 30 min at 4°C under constant mixing. The nuclear lysate was centrifuged for 15 min to remove cell debris and soluble proteins were quantified by BCA (ThermoFisher). The lysate was pre-cleared with 30 µg of Dynabeads Protein G/A (ThermoFisher) for 30 min under rotating mixing, and then split in aliquots of 1.5 mg proteins for each assay. Each fraction received 5 µg of antibody and the incubation was left overnight at 4°C under rotating mixing. The following antibodies were used for IP: rabbit anti-DDX17 (19910-1-AP, Proteintech) or control rabbit IgG (ThermoFisher), goat anti-DDX5 (ab10261, Abcam) or control goat IgG (Santa Cruz Biotechnology), and mouse anti-MYCN (sc-53993, Santa Cruz Biotechnology) or control mouse IgG. The next day, the different lysate/antibody mixtures were divided, treated or not with 250 U/ml benzonase (Merck-

Millipore) for 30 min at 37°C, and then incubated with 50 µg Dynabeads Protein G/A (ThermoFisher) blocked with bovine serum albumin, for 4 h at 4°C under rotating mixing. Bead were then washed 5 times with IP buffer. Elution was performed by boiling for 5 min in SDS-PAGE loading buffer prior to analysis by western-blotting.

RNA extraction and real-time quantitative PCR

Total RNAs were isolated using TriPure Isolation Reagent (Roche). For reverse transcription, 2 µg of purified RNAs were treated with Dnase I (ThermoFisher) and retrotranscribed using Maxima reverse transcriptase (ThermoFisher), as recommended by the supplier. Potential genomic DNA contamination was systematically checked by performing negative RT controls in the absence of enzyme and by including controls with water instead of cDNA in qPCR assays. PCR reactions were carried out described previously (Lambert et al., 2018). For qPCR analyses, the specificity and linear efficiency of each primer set (sequences are in Supp. Table 2) was first verified by establishing a standard expression curve with various amounts of human genomic DNA or cDNA. qPCR reactions were carried out on 0.625 ng of cDNA using a LightCycler 480 System (Roche), with the SYBR® Premix Ex Taq (Tli RNaseH Plus, Takara), under conditions recommended by the manufacturer. Melting curves were controlled to rule out the existence of non-specific products. Relative DNA levels were calculated using the 2-ΔΔCt method (using the average Ct obtained from technical duplicates or triplicates) and were normalized to the expression of *GAPDH* RNA.

Chromatin immunoprecipitation

A total of 10^7 cells were crosslinked with 1% formaldehyde for 10 min at room temperature. Crosslinking was quenched by addition of 0.125 M glycin. Nuclei were isolated by sonication using a Covaris S220 (2 min, Peak Power: 75; Duty Factor: 2; Cycles/burst: 200), pelleted by centrifugation at 1000 g for 5 min at 4°C, washed once with FL buffer (5 mM HEPES pH 8.0, 85 mM KCl, 0.5% NP40) and resuspended in 1 ml shearing buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.1% SDS). Chromatin was sheared in order to obtain fragments ranging from 200 to 800 bp using Covaris S220 (20 min, Peak Power: 140; Duty Factor: 5; Cycles/burst: 200). Chromatin was next immunoprecipitated overnight at 4°C with 4 µg of mouse anti-MYCN II antibody (sc53993, Santa Cruz Biotechnology) or an equivalent amount of the corresponding IgG Isotype control (ThermoFisher) and 30 µl of Dynabeads® Protein A/G (ThermoFisher). Complexes were washed with 4 different buffers: Low salt buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% Triton X-100, 0.1% SDS, 2 mM EDTA), High salt buffer (20 mM Tris-HCl pH 8.0, 500 mM NaCl, 1% Triton X-100, 0.1% SDS, 2 mM EDTA), Low LiCl buffer (10 mM Tris-HCl pH 8.0, 0.5 M LiCl, 1% NaDoc, 1% NP40), Tris/EDTA (10 mM Tris-HCl pH 8.0, 1 mM EDTA), and were eluted in Elution buffer (200 mM NaCl, 1% SDS, 0.1 M $NaHCO_3$, 20 µg Proteinase K) overnight at 65°C. The

immunoprecipitated chromatin was purified by phenol-chloroform extraction and ethanol precipitation, and analysed by qPCR. Values were normalized to the signal obtained for the immunoprecipitation with control IgG.

In silico prediction of chimeric RNAs from RNA-seq data

The raw RNA-seq data were described previously (Terrone et al., 2022) and are accessible from the Gene Expression Omnibus repository (accession number GSE183205). Raw reads were pre-processed using fastp (v0.23.2) (Chen et al., 2018) and mapped to the human reference genome (hg19, GRCh37.87) using STAR (v2.7.8a) (Dobin et al., 2013). Mapped reads were filtered using samtools (v1.11) (Danecek et al., 2021). Gene fusions were detected using Arriba (v2.3.0) (Uhrig et al., 2021) and results were parsed using an R script. Split and discordant reads identified by Arriba were counted for each fusion and compared to control condition. Differentially fused genes (DFG) were analyzed using the DESeq2 package (v1.40.1) (Love et al., 2014). The complete pipeline is available in Nextflow (Di Tommaso et al., 2017) at https://gitbio.ens-lyon.fr/LBMC/regards/nextflow/-/blob/master/src/arriba_fusion.nf.

Statistical analysis of *DDX17* and *DDX5* expression in neuroblastoma tumours

To test the expression of *DDX17* and *DDX5* in neuroblastomas, we used the previously described cohort of 498 tumours (GEO accession number: GSE62564) (Zhang et al., 2016b), in which patients groups were defined according to the International Neuroblastoma Staging System (INSS). Survival curves were generated by the Kaplan-Meier method, and statistical analyses were carried out as described previously (Gibert et al., 2014, Jiang et al., 2021).

Analysis of MYCN ChIP-seq data

To analyse the relative proximity of DDX17/DDX5-regulated terminal exons to MYCN binding sites, we first generated a BED file containing a merged list of MYCN peaks identified in several neuroblastoma cell lines (Buchel et al., 2017, Upton et al., 2020, Zeid et al., 2018) (Supp. Table 3), re-analysed as follows. Raw reads were pre-processed using fastp (v0.23.2) (Chen et al., 2018) and mapped to the human reference genome (hg19, GRCh37.87) using bowtie2 (v2.5.2) (Langmead et al., 2019). Mapped reads were filtered using samtools (v1.11) (Danecek et al., 2021) and then formatted using deepTools2 (v3.5.1) (Ramirez et al., 2016). Peak calling analysis was carried out with MACS (v3.0.0a6) (Zhang et al., 2008). Nearest exons from peak summits were identified using to BEDtools (v2.25.0) (Quinlan and Hall, 2010). The complete ChIP-seq pipeline is available in Nextflow (Di Tommaso et al., 2017) at https://gitbio.enslyon.fr/xgrand/ChIPster.

We next calculated the genomic distance (negative or positive for upstream and downstream peaks, respectively) between each exon from the FasterDB database (Mallinjoud et al., 2014) and the summit of the closest MYCN peak. Only genes having at least one internal exon were considered. We performed a logistic regression analysis to test if the 3' terminal exons regulated by DDX5/DDX17 (n = 933) are closer to MYCN peaks than unregulated terminal exons. (n = 18,482). We modeled the proximity to a MYCN peak according to the two groups of exons using the glm function, with family = binomial ('logit') in R software. A MYCN peak was considered as close to an exon if its center is located within an exon or within an interval of 1 base to 2 kb upstream or downstream of the exon. To test the differences between the two groups of exons, a Tukey's test was used (with R, emmeans function from emmeans library).

Acknowledgments

We wish to thank Reini F. Luco (Institut Curie, Orsay, France) and Philippe Mangeot (CIRI, Lyon, France) for sharing reagents. We are grateful to LBMC members for fruitful discussions and advice. We acknowledge the CBPSMN (Centre Blaise Pascal de Simulation et de Modélisation Numérique) of the ENS-Lyon for computing resources. This work was supported by grants from Ligue contre le Cancer (Equipe labellisée), Fondation ARC, and Association Hubert Gouin "Enfance et Cancer". V.C. was supported by a doctoral fellowship from Fondation pour la Recherche Médicale (FRM). J.V. received doctoral fellowships from the French Ministry of Research and Education and from the Ligue contre le Cancer.

Author Contributions

All authors contributed to the study conception and design. V.C., J.V. and M.B. carried out experiments, collected and interpreted the data. X.G. and N.F. analysed RNA-seq and ChIP-seq datasets. N.R. and B.G. analysed the cohort of neuroblastoma tumours. The first draft of the manuscript was written by V.C. and C.F.B. and all authors commented, corrected and approved the manuscript.

Conflict of interest

The authors declare no competing interests.

References

Akiva P, Toporik A, Edelheit S, Peretz Y, Diber A, Shemesh R, Novik A & Sorek R 2006 Transcription-mediated gene fusion in the human genome. *Genome Res* 16:30-6

Alpert T, Straube K, Carrillo Oesterreich F, Herzel L & Neugebauer KM 2020 Widespread Transcriptional Readthrough Caused by Nab2 Depletion Leads to Chimeric Transcripts with Retained Introns. *Cell Rep* 33:108324

Babiceanu M, Qin F, Xie Z, Jia Y, Lopez K, Janus N, Facemire L, Kumar S, Pang Y, Qi Y, et al. 2016 Recurrent chimeric fusion RNAs in non-cancer tissues and cells. *Nucleic Acids Res* 44:2859-72

Bader AS, Luessing J, Hawley BR, Skalka GL, Lu WT, Lowndes NF & Bushell M 2022 DDX17 is required for efficient DSB repair at DNA:RNA hybrid deficient loci. *Nucleic Acids Res* 50:10487-10502

Baluapuri A, Hofstetter J, Dudvarski Stankovic N, Endres T, Bhandare P, Vos SM, Adhikari B, Schwarz JD, Narain A, Vogt M, et al. 2019 MYC Recruits SPT5 to RNA Polymerase II to Promote Processive Transcription Elongation. *Mol Cell* 74:674-687 e11

Baluapuri A, Wolf E & Eilers M 2020 Target gene-independent functions of MYC oncoproteins. *Nat Rev Mol Cell Biol* 21:255-267

Bauer DLV, Tellier M, Martinez-Alonso M, Nojima T, Proudfoot NJ, Murphy S & Fodor E 2018 Influenza Virus Mounts a Two-Pronged Attack on Host RNA Polymerase II Transcription. *Cell Rep* 23:2119-2129 e3

Boreikaite V & Passmore LA 2023 3'-End Processing of Eukaryotic mRNA: Machinery, Regulation, and Impact on Gene Expression. *Annu Rev Biochem* 92:199-225

Braccioli L & de Wit E 2019 CTCF: a Swiss-army knife for genome organization and transcription regulation. *Essays Biochem* 63:157-165

Buchel G, Carstensen A, Mak KY, Roeschert I, Leen E, Sumara O, Hofstetter J, Herold S, Kalb J, Baluapuri A, et al. 2017 Association with Aurora-A Controls N-MYC-Dependent Promoter Escape and Pause Release of RNA Polymerase II during the Cell Cycle. *Cell Rep* 21:3483-3497

Camats M, Guil S, Kokolo M & Bach-Elias M 2008 P68 RNA helicase (DDX5) alters activity of cis- and trans-acting factors of the alternative splicing of H-Ras. *PLoS One* 3:e2926

Cardiello JF, Goodrich JA & Kugel JF 2018 Heat Shock Causes a Reversible Increase in RNA Polymerase II Occupancy Downstream of mRNA Genes, Consistent with a Global Loss in Transcriptional Termination. *Mol Cell Biol* 38

Cargill M, Venkataraman R & Lee S 2021 DEAD-Box RNA Helicases and Genome Stability. *Genes (Basel)* 12

Chen S, Zhou Y, Chen Y & Gu J 2018 fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884-i890

Chwalenia K, Qin F, Singh S & Li H 2019 A cell-based splicing reporter system to identify regulators of cis-splicing between adjacent genes. *Nucleic Acids Res* 47:e24

Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, Whitwham A, Keane T, McCarthy SA, Davies RM, et al. 2021 Twelve years of SAMtools and BCFtools. *Gigascience* 10

Dardenne E, Polay Espinoza M, Fattet L, Germann S, Lambert MP, Neil H, Zonta E, Mortada H, Gratadou L, Deygas M, et al. 2014 RNA helicases DDX5 and DDX17 dynamically orchestrate transcription, miRNA, and splicing programs in cell differentiation. *Cell Rep* 7:1900-13

Di Tommaso P, Chatzou M, Floden EW, Barja PP, Palumbo E & Notredame C 2017 Nextflow enables reproducible computational workflows. *Nat Biotechnol* 35:316-319

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M & Gingeras TR 2013 STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15-21

Dorney R, Dhungel BP, Rasko JEJ, Hebbard L & Schmitz U 2023 Recent advances in cancer fusion transcript detection. *Brief Bioinform* 24

Egashira S, Jinnin M, Makino K, Ajino M, Shimozono N, Okamoto S, Tazaki Y, Hirano A, Ide M, Kajihara I, et al. 2019 Recurrent Fusion Gene ADCK4-NUMBL in Cutaneous Squamous Cell Carcinoma Mediates Cell Proliferation. *J Invest Dermatol* 139:954-957

Fuller-Pace FV & Moore HC 2011 RNA helicases p68 and p72: multifunctional proteins with important implications for cancer development. *Future Oncol* 7:239-51

Gibert B, Delloye-Bourgeois C, Gattolliat CH, Meurette O, Le Guernevel S, Fombonne J, Ducarouge B, Lavial F, Bouhallier F, Creveaux M, et al. 2014 Regulation by miR181 family of the dependence receptor CDON tumor suppressive activity in neuroblastoma. *J Natl Cancer Inst* 106

Giraud G, Terrone S & Bourgeois CF 2018 Functions of DEAD box RNA helicases DDX5 and DDX17 in chromatin organization and transcriptional regulation. *BMB Rep* 51**:**613-622

Greger L, Su J, Rung J, Ferreira PG, Geuvadis c, Lappalainen T, Dermitzakis ET & Brazma A 2014 Tandem RNA chimeras contribute to transcriptome diversity in human population and are associated with intronic genetic variants. *PLoS One* 9**:**e104567

Grosso AR, Leite AP, Carvalho S, Matos MR, Martins FB, Vitor AC, Desterro JM, Carmo-Fonseca M & de Almeida SF 2015 Pervasive transcription read-through promotes aberrant expression of oncogenes and RNA chimeras in renal carcinoma. *Elife* 4

Hadar S, Meller A, Saida N & Shalgi R 2022 Stress-induced transcriptional readthrough into neighboring genes is linked to intron retention. *iScience* 25**:**105543

Han P, Chen RH, Wang F, Zeng JY, Yu ST, Xu LH, Cai Q, Liang FY, Xia TL, Lin ZR, et al. 2017 Novel chimeric transcript RRM2-c2orf48 promotes metastasis in nasopharyngeal carcinoma. *Cell Death Dis* 8**:**e3047

Heinz S, Texari L, Hayes MGB, Urbanowski M, Chang MW, Givarkes N, Rialdi A, White KM, Albrecht RA, Pache L, et al. 2018 Transcription Elongation Can Affect Genome 3D Structure. *Cell* 174**:**1522-1536 e22

Hennig T, Michalski M, Rutkowski AJ, Djakovic L, Whisnant AW, Friedl MS, Jha BA, Baptista MAP, L'Hernault A, Erhard F, et al. 2018 HSV-1-induced disruption of transcription termination resembles a cellular stress response but selectively increases chromatin accessibility downstream of genes. *PLoS Pathog* 14**:**e1006954

Herold S, Kalb J, Buchel G, Ade CP, Baluapuri A, Xu J, Koster J, Solvie D, Carstensen A, Klotz C, et al. 2019 Recruitment of BRCA1 limits MYCN-driven accumulation of stalled RNA polymerase. *Nature* 567**:**545-549

Jiang S, Richaud M, Vieugue P, Rama N, Delcros JG, Siouda M, Sanada M, Redavid AR, Ducarouge B, Hervieu M, et al. 2021 Targeting netrin-3 in small cell lung cancer and neuroblastoma. *EMBO Mol Med* 13**:**e12878

Kar A, Fushimi K, Zhou X, Ray P, Shi C, Chen X, Liu Z, Chen S & Wu JY 2011 RNA helicase p68 (DDX5) regulates tau exon 10 splicing by modulating a stem-loop structure at the 5' splice site. *Mol Cell Biol* 31**:**1812-21

Katahira J, Ishikawa H, Tsujimura K, Kurono S & Hieda M 2019 Human THO coordinates transcription termination and subsequent transcript release from the HSP70 locus. *Genes Cells* 24**:**272-283

Kim RN, Kim A, Choi SH, Kim DS, Nam SH, Kim DW, Kim DW, Kang A, Kim MY, Park KH, et al. 2012 Novel mechanism of conjoined gene formation in the human genome. *Funct Integr Genomics* 12**:**45-61

Lai YH, Choudhary K, Cloutier SC, Xing Z, Aviran S & Tran EJ 2019 Genome-Wide Discovery of DEAD-Box RNA Helicase Targets Reveals RNA Structural Remodeling in Transcription Termination. *Genetics* 212**:**153-174

Lambert MP, Terrone S, Giraud G, Benoit-Pilven C, Cluet D, Combaret V, Mortreux F, Auboeuf D & Bourgeois CF 2018 The RNA helicase DDX17 controls the transcriptional activity of REST and the expression of proneural microRNAs in neuronal differentiation. *Nucleic Acids Res* 46**:**7686-7700

Langmead B, Wilks C, Antonescu V & Charles R 2019 Scaling read aligners to hundreds of threads on general-purpose processors. *Bioinformatics* 35**:**421-432

Lee BK, Bhinge AA, Battenhouse A, McDaniell RM, Liu Z, Song L, Ni Y, Birney E, Lieb JD, Furey TS, et al. 2012 Cell-type specific and combinatorial usage of diverse transcription factors revealed by genome-wide binding studies in multiple human cells. *Genome Res* 22**:**9-24

Love MI, Huber W & Anders S 2014 Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15**:**550

Mallinjoud P, Villemin JP, Mortada H, Polay Espinoza M, Desmet FO, Samaan S, Chautard E, Tranchevent LC & Auboeuf D 2014 Endothelial, epithelial, and fibroblast cells exhibit specific splicing programs independently of their tissue of origin. *Genome Res* 24**:**511-21

Mersaoui SY, Yu Z, Coulombe Y, Karam M, Busatto FF, Masson JY & Richard S 2019 Arginine methylation of the DDX5 helicase RGG/RG motif by PRMT5 regulates resolution of RNA:DNA hybrids. *EMBO J* 38**:**e100986

Morgan M, Shiekhattar R, Shilatifard A & Lauberth SM 2022 It's a DoG-eat-DoG world-altered transcriptional mechanisms drive downstream-of-gene (DoG) transcript production. *Mol Cell* 82:1981-1991

Morimachi M, Hirabayashi K, Takanashi Y, Kawanishi A, Saika T, Ueyama Y, Nakagohri T, Nakamura N, Suzuki H & Kagawa T 2021 Low expression of DDX5 is associated with poor prognosis in patients with pancreatic ductal adenocarcinoma. *J Clin Pathol* 74:741-745

Nanavaty V, Abrash EW, Hong C, Park S, Fink EE, Li Z, Sweet TJ, Bhasin JM, Singuri S, Lee BH, et al. 2020 DNA Methylation Regulates Alternative Polyadenylation via CTCF and the Cohesin Complex. *Mol Cell* 78:752-764 e6

Neugebauer KM 2019 Nascent RNA and the Coordination of Splicing with Transcription. *Cold Spring Harb Perspect Biol* 11:a032227

Otte J, Dyberg C, Pepich A & Johnsen JI 2020 MYCN Function in Neuroblastoma Development. *Front Oncol* 10:624079

Papadopoulos D, Solvie D, Baluapuri A, Endres T, Ha SA, Herold S, Kalb J, Giansanti C, Schulein-Volk C, Ade CP, et al. 2022 MYCN recruits the nuclear exosome complex to RNA polymerase II to prevent transcription-replication conflicts. *Mol Cell* 82:159-176 e12

Parra G, Reymond A, Dabbouseh N, Dermitzakis ET, Castelo R, Thomson TM, Antonarakis SE & Guigo R 2006 Tandem chimerism as a means to increase protein complexity in the human genome. *Genome Res* 16:37-44

Porrua O & Libri D 2015 Transcription termination and the control of the transcriptome: why, where and how to stop. *Nat Rev Mol Cell Biol* 16:190-202

Prakash T, Sharma VK, Adati N, Ozawa R, Kumar N, Nishida Y, Fujikake T, Takeda T & Taylor TD 2010 Expression of conjoined genes: another mechanism for gene regulation in eukaryotes. *PLoS One* 5:e13284

Proudfoot NJ 2016 Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science* 352:aad9926

Qin F, Song Y, Zhang Y, Facemire L, Frierson H & Li H 2016 Role of CTCF in Regulating SLC45A3-ELK4 Chimeric RNA. *PLoS One* 11:e0150382

Qin F, Song Z, Babiceanu M, Song Y, Facemire L, Singh R, Adli M & Li H 2015 Discovery of CTCF-sensitive Cis-spliced fusion RNAs between adjacent genes in human prostate cells. *PLoS Genet* 11:e1005001

Qin F, Zhang Y, Liu J & Li H 2017 SLC45A3-ELK4 functions as a long non-coding chimeric RNA. *Cancer Lett* 404:53-61

Quinlan AR & Hall IM 2010 BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841-2

Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F & Manke T 2016 deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 44:W160-5

Rodriguez-Molina JB, West S & Passmore LA 2023 Knowing when to stop: Transcription termination on protein-coding genes by eukaryotic RNAPII. *Mol Cell* 83:404-415

Rosa-Mercado NA & Steitz JA 2022 Who let the DoGs out? - biogenesis of stress-induced readthrough transcripts. *Trends Biochem Sci* 47:206-217

Rosa-Mercado NA, Zimmer JT, Apostolidi M, Rinehart J, Simon MD & Steitz JA 2021 Hyperosmotic stress alters the RNA polymerase II interactome and induces readthrough transcription despite widespread transcriptional repression. *Mol Cell* 81:502-513 e4

Rutkowski AJ, Erhard F, L'Hernault A, Bonfert T, Schilhabel M, Crump C, Rosenstiel P, Efstathiou S, Zimmer R, Friedel CC, et al. 2015 Widespread disruption of host transcription termination in HSV-1 infection. *Nat Commun* 6:7126

Shi Y, Yuan J, Rraklli V, Maxymovitz E, Cipullo M, Liu M, Li S, Westerlund I, Bedoya-Reina OC, Bullova P, et al. 2021 Aberrant splicing in neuroblastoma generates RNA-fusion transcripts and provides vulnerability to spliceosome inhibitors. *Nucleic Acids Res*

Sun Y & Li H 2022 Chimeric RNAs Discovered by RNA Sequencing and Their Roles in Cancer and Rare Genetic Diseases. *Genes (Basel)* 13

Tellier M, Maudlin I & Murphy S 2020 Transcription and splicing: A two-way street. *Wiley Interdiscip Rev RNA* 11:e1593

Terrone S, Valat J, Fontrodona N, Giraud G, Claude JB, Combe E, Lapendry A, Polveche H, Ameur LB, Duvermy A, et al. 2022 RNA helicase-dependent gene looping impacts messenger RNA processing. *Nucleic Acids Res* 50:9226-9246

Uhrig S, Ellermann J, Walther T, Burkhardt P, Frohlich M, Hutter B, Toprak UH, Neumann O, Stenzinger A, Scholl C, et al. 2021 Accurate and efficient detection of gene fusions from RNA sequencing data. *Genome Res* 31:448-460

Upton K, Modi A, Patel K, Kendsersky NM, Conkrite KL, Sussman RT, Way GP, Adams RN, Sacks GI, Fortina P, et al. 2020 Epigenomic profiling of neuroblastoma cell lines. *Sci Data* 7:116

Varley KE, Gertz J, Roberts BS, Davis NS, Bowling KM, Kirby MK, Nesmith AS, Oliver PG, Grizzle WE, Forero A, et al. 2014 Recurrent read-through fusion transcripts in breast cancer. *Breast Cancer Res Treat* 146:287-97

Vilborg A, Passarelli MC, Yario TA, Tycowski KT & Steitz JA 2015 Widespread Inducible Transcription Downstream of Human Genes. *Mol Cell* 59:449-61

Vilborg A, Sabath N, Wiesel Y, Nathans J, Levy-Adam F, Yario TA, Steitz JA & Shalgi R 2017 Comparative analysis reveals genomic features of stress-induced transcriptional readthrough. *Proc Natl Acad Sci U S A* 114:E8362-E8371

Weber D, Ibn-Salem J, Sorn P, Suchan M, Holtstrater C, Lahrmann U, Vogler I, Schmoldt K, Lang F, Schrors B, et al. 2022 Accurate detection of tumor-specific gene fusions reveals strongly immunogenic personal neo-antigens. *Nat Biotechnol* 40:1276-1284

Wei Z, Wang S, Xu Y, Wang W, Soares F, Ahmed M, Su P, Wang T, Orouji E, Xu X, et al. 2023 MYC reshapes CTCF-mediated chromatin architecture in prostate cancer. *Nat Commun* 14:1787

Wortham NC, Ahamed E, Nicol SM, Thomas RS, Periyasamy M, Jiang J, Ochocka AM, Shousha S, Huson L, Bray SE, et al. 2009 The DEAD-box protein p72 regulates ER alpha-/oestrogen-dependent transcription and cell growth, and is associated with improved survival in ER alpha-positive breast cancer. *Oncogene* 28:4053-4064

Wu G, Xing Z, Tran EJ & Yang D 2019 DDX5 helicase resolves G-quadruplex and is involved in MYC gene transcriptional activation. *Proc Natl Acad Sci U S A* 116:20453-20461

Xing Z, Ma WK & Tran EJ 2019 The DDX5/Dbp2 subfamily of DEAD-box RNA helicases. *Wiley Interdiscip Rev RNA* 10:e1519

Xu K, Sun S, Yan M, Cui J, Yang Y, Li W, Huang X, Dou L, Chen B, Tang W, et al. 2022 DDX5 and DDX17-multifaceted proteins in the regulation of tumorigenesis and tumor progression. *Front Oncol* 12:943032

Yang W, Lee KW, Srivastava RM, Kuo F, Krishna C, Chowell D, Makarov V, Hoen D, Dalin MG, Wexler L, et al. 2019 Immunogenic neoantigens derived from gene fusions stimulate T cell responses. *Nat Med* 25:767-775

Yun SM, Yoon K, Lee S, Kim E, Kong SH, Choe J, Kang JM, Han TS, Kim P, Choi Y, et al. 2014 PPP1R1B-STARD3 chimeric fusion transcript in human gastric cancer promotes tumorigenesis through activation of PI3K/AKT signaling. *Oncogene* 33:5341-7

Zeid R, Lawlor MA, Poon E, Reyes JM, Fulciniti M, Lopez MA, Scott TG, Nabet B, Erb MA, Winter GE, et al. 2018 Enhancer invasion shapes MYCN-dependent transcriptional amplification in neuroblastoma. *Nat Genet* 50:515-523

Zhang H, Xing Z, Mani SK, Bancel B, Durantel D, Zoulim F, Tran EJ, Merle P & Andrisani O 2016a RNA helicase DEAD box protein 5 regulates Polycomb repressive complex 2/Hox transcript antisense intergenic RNA function in hepatitis B virus infection and hepatocarcinogenesis. *Hepatology* 64:1033-48

Zhang H, Zhang Y, Zhu X, Chen C, Zhang C, Xia Y, Zhao Y, Andrisani O & Kong L 2019 DEAD Box Protein 5 Inhibits Liver Tumorigenesis by Stimulating Autophagy via Interaction with p62/SQSTM1. *Hepatology* 69:1046-1063

Zhang S, Wei JS, Li SQ, Badgett TC, Song YK, Agarwal S, Coarfa C, Tolman C, Hurd L, Liao H, et al. 2016b MYCN controls an alternative RNA splicing program in high-risk metastatic neuroblastoma. *Cancer Lett* 371**:**214-24

Zhang W, Yu Y, Hertwig F, Thierry-Mieg J, Zhang W, Thierry-Mieg D, Wang J, Furlanello C, Devanarayan V, Cheng J, et al. 2015 Comparison of RNA-seq and microarray-based models for clinical endpoint prediction. *Genome Biol* 16**:**133

Zhang Y, Gong M, Yuan H, Park HG, Frierson HF & Li H 2012 Chimeric transcript generated by cis-splicing of adjacent genes regulates prostate cancer cell proliferation. *Cancer Discov* 2**:**598-607

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. 2008 Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9**:**R137

Zhao X, Li D, Yang F, Lian H, Wang J, Wang X, Fang E, Song H, Hu A, Guo Y, et al. 2020 Long Noncoding RNA NHEG1 Drives beta-Catenin Transactivation and Neuroblastoma Progression through Interacting with DDX5. *Mol Ther* 28**:**946-962

Zhou X, Zhao M, Fan Y & Xu Y 2024 Identification of a necroptosis-related gene signature for making clinical predictions of the survival of patients with lung adenocarcinoma. *PeerJ* 12**:**e16616

Figure legends

Figure 1. DDX17 and DDX5 depletion enhances the expression of chimeric transcripts. **A.** Schematic representation of the production of readthrough induced chimeric transcripts. **B.** Diagram showing the number of up and downregulated chimeric trancripts displaying at least a fold change of 2 and the most significant transcripts (*P*-val<0.1). **C.** Venn diagram showing the overlap between genes producing chimeric transcripts and the genes previously identified as displaying transcriptional readthrough events. The list of 953 genes corresponds to the genes on which we carried out global analyses (Supp. Table 1). Note that among the 218 upregulated chimeric transcripts from panel B, there are 2 cases in which one exon from the gene of origin is spliced with exons from different target genes, generating 2 different chimeras. As these events can be counted only once in panel B, this explains the total of 216 upregulated chimeras. **D.** Quantification of chimeric transcripts. Expression of chimeric transcripts was monitored by RT-qPCR using primers spanning the chimeric junction and then normalized on the expression of the 5' parental gene (see also Supp. Fig. 2B). Two-tailed paired t-test (* *P*-val<0.05 ; ** *P*-val<0.01).

Figure 2. Chimeric transcripts can produce chimeric proteins. **A.** Schematic representation of *CTSD-IFITM10* chimeric transcripts and putative chimeric protein. The sites targeted by the siRNAs used to validate the specificity of the band corresponding to CTSD-IFITM10 protein are indicated. **B.** Western-blot showing the expression of the chimeric CTSD-IFITM10 protein and the canonical protein CTSD, as well as the validation of *DDX5* and *DDX17* silencing.

Figure 3. The expression of the *DDX17* gene is reduced in high-risk neuroblastomas. **A.** Venn diagrams showing the number of chimeric transcripts (defined as in Figure 1A) that are both up or down-regulated after DDX17/DDX5 depletion and identified in neuroblastoma tumors. **B.** Top 10 of transcripts identified both in neuroblastoma tumours and upregulated upon DDX17/DDX5 depletion. **C.** Kaplan-Meier curves showing the overall survival of neuroblastoma patients, separated in 2 groups of high and low *DDX17* expression (relative to the median value of the group). Left diagram: all patients (n=498). Right diagram: stage 4 patients (n=183). The number of patients in each group and for each time point is indicated below. **D.** Box plot of *DDX17* expression relative to tumor stage (INSS classification). ANOVA corrected for multiple comparisons with Tukey's tests. Only the comparison between Stage 4 and other stages is shown, the other comparisons were not significant. **E.** Box plot of *DDX17* expression relative to low/high risk classification. Two-tailed *t*-test. **F.** Box plot of *DDX17* expression relative to amplified (A) or non-amplified (NA) *MYCN* status. Two-tailed *t*-test.

Figure 4. MYCN overexpression does not impact the expression of DDX17 and DDX5 but it induces the formation of chimeric transcripts. **A.** Western-blot showing DDX17 and DDX5 proteins levels upon increasing overexpression of MYCN in SH-SY5Y cells. Quantification of the western-blot (normalized to vinculin) is shown on the right. **B.** Expression of chimeric and readthrough transcripts upon MYCN overexpression. Details are as in Figure 1D (see also Supp. Fig. 4). Ratio paired *t*-test (*P-val<0.05 ; **P-val<0.01 ; ***P-val<0.001). **C.** Western-blot showing DDX17 and DDX5 proteins levels upon MYCN depletion in SK-N-BE(2) cells. Quantification of the western-blot (normalized to GAPDH) is shown on the right. **D.** Expression of chimeric and readthrough transcripts upon MYCN depletion. Unpaired Mann-Whitney test (*P-val<0.05 ; **P-val<0.01).

Figure 5. MYCN binds to the 3' end of DDX17/DDX5-regulated genes and interacts with DDX17. **A.** Typical examples of DDX17/DDX5-regulated genes displaying MYCN binding near their 3' end. **B.** Genome wide analysis of MYCN binding near terminal exons of DDX17/DDX5-regulated genes and control genes in neuroblastoma cell lines. Tukey's test. **C.** ChIP-qPCR analysis of MYCN binding near the 3' end of DDX17/DDX5-regulated genes. Paired *t*-test (*P-val<0.05 ; **P-val<0.01). **D.** Co-immunoprecipitation of endogenous MYCN and DDX17 in SK-N-BE(2) cells, in the presence or absence of benzonase, as indicated. **E.** Co-immunoprecipitation of endogenous MYCN and DDX5 in SK-N-BE(2) cells.

Figure 6. The forced recruitment of MYCN near the termination region is sufficient to induce the formation of chimeric transcripts. **A.** Strategy for the recruitment of MYCN (fused to a catalytically inactive dCas9) at the 3' end of the *MAPK7* gene. **B.** ChIP-qPCR analysis of dCas9-MYCN and dCas9-GFP

binding near the 3’ end of *MAPK7*. Each ChIP signal was normalised to its respective control (ChIP with IgG). **C.** Quantification of canonical *MAPK7* transcript and chimeric *MAPK7-RNF112* transcript in cells transfected with dCas9-MYCN or dCas9-GFP, relative to transfection without specific gRNA. Paired *t*-test (*P-val<0.05 ; **P-val<0.01).

Figure 1

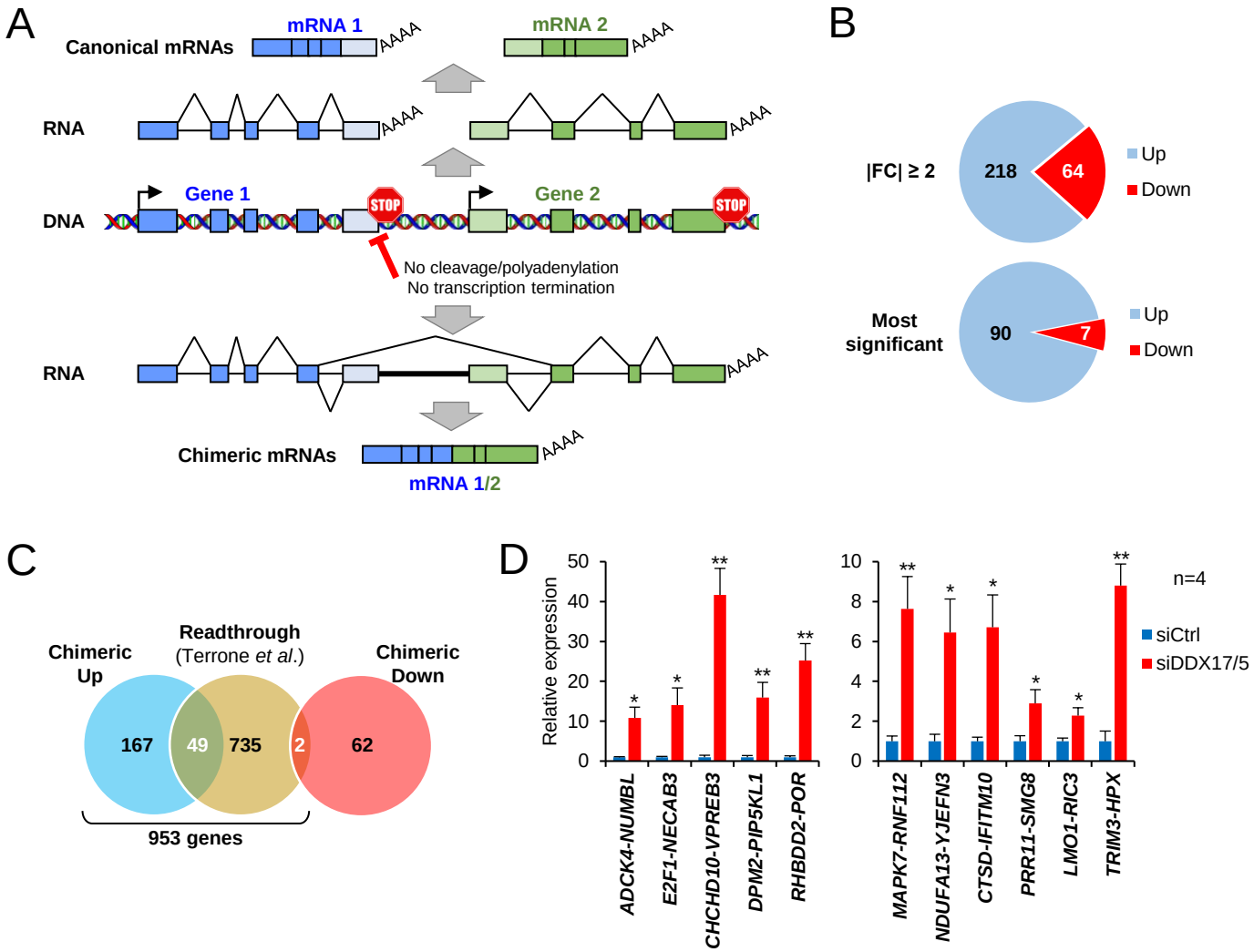

A
Canonical mRNAs
mRNA 1
mRNA 2
AAAA
RNA
DNA
Gene 1
Gene 2
STOP
No cleavage/polyadenylation
No transcription termination
RNA
Chimeric mRNAs
mRNA 1/2
B
|FC| ≥ 2
218
64
Up
Down
Most significant
90
7
Up
Down
C
Chimeric Up
Readthrough (Terrone et al.)
Chimeric Down
167
49
735
2
62
953 genes
D
Relative expression
n=4
siCtrl
siDDX17/5
ADCK4-NUMBL
E2F1-NECAB3
CHCHD10-VPREB3
DPM2-PIP5KL1
RHBDD2-POR
MAPK7-RNF112
NDUFA13-YJEFN3
CTSD-IFITM10
PRR11-SMG8
LMO1-RIC3
TRIM3-HPX

Figure 2

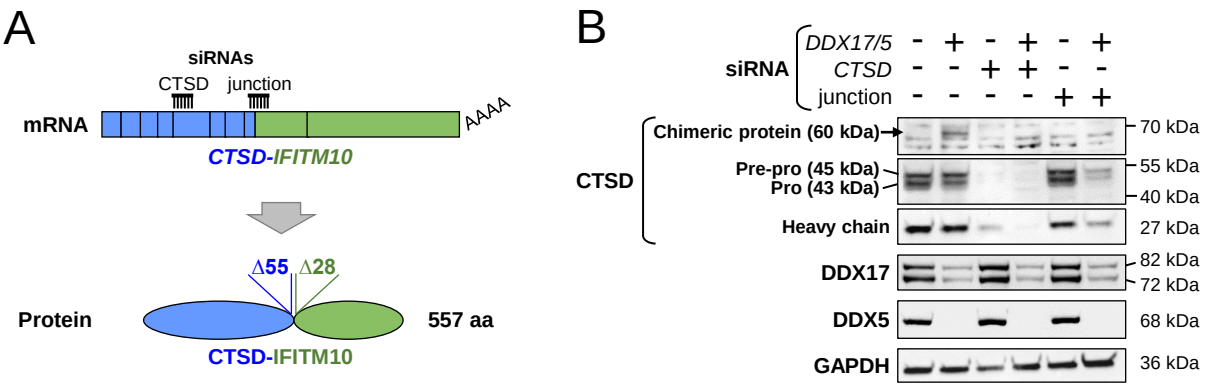
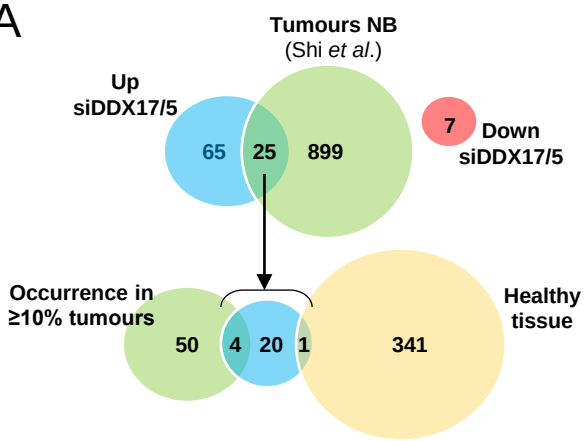

A
siRNAs
CTSD
junction
mRNA
AAAA
CTSD-IFITM10
Δ55
Δ28
Protein
557 aa
CTSD-IFITM10
B
siRNA
DDX17/5 - + - + - +
CTSD - - + + - -
junction - - - - + +
CTSD
Chimeric protein (60 kDa)
Pre-pro (45 kDa)
Pro (43 kDa)
Heavy chain
DDX17
DDX5
GAPDH
70 kDa
55 kDa
40 kDa
27 kDa
82 kDa
72 kDa
68 kDa
36 kDa

Figure 3

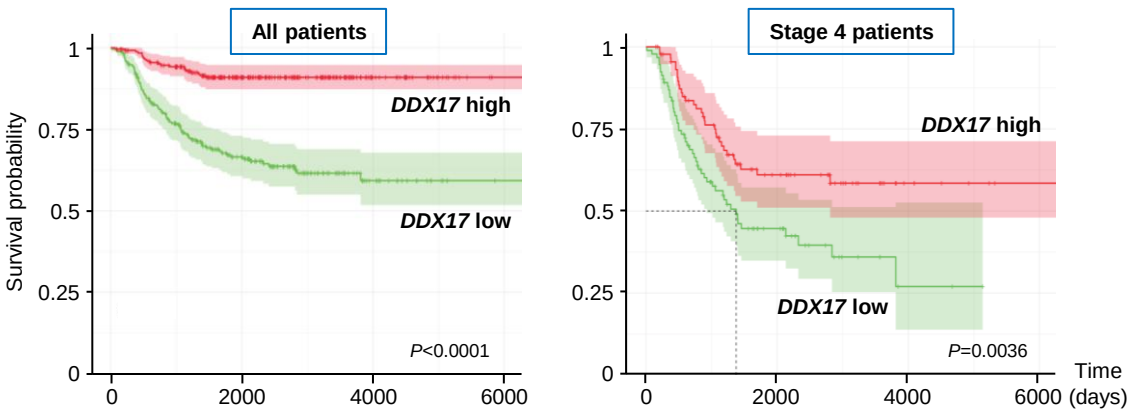
A



B

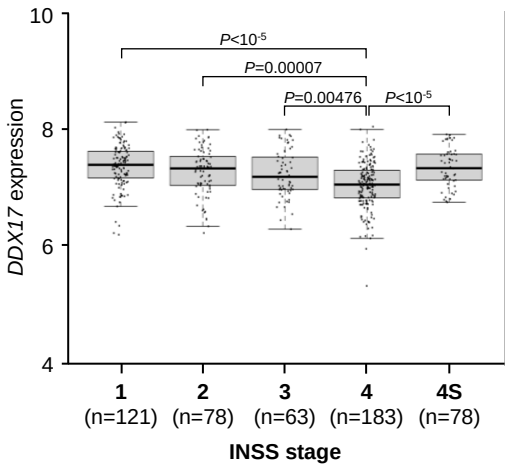
Chimeric transcript	Frequency in neuroblastoma	Fold change in siDDX17/DDX5
CHCHD10-VPREB3	27.33%	11.9
MAPK7-RNF112	18.60%	14.5
ZDHHC1-TPPP3	13.95%	6.6
PRDX5-CCDC88B	10.47%	14.6
E2F1-NECAB3	9.88%	11.5
PELI3-DPP3	4.07%	7.7
MAP3K11-KCNK7	3.49%	28.4
NTRK1-PEAR1	2.91%	17.6
PEAR1-LRRC71	2.91%	23.4
RUNDC1-RPL27	2.33%	12.1

C

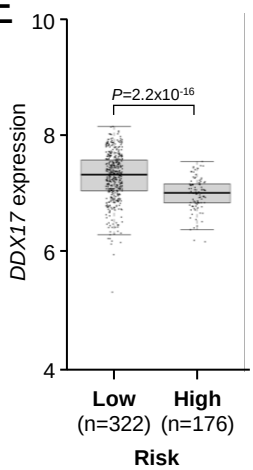


***DDX17* low**	249	107	21	1	92	24	2	0
***DDX17* high**	249	143	31	1	91	34	7	1

D



E



F

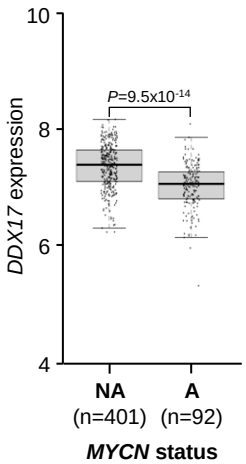


Figure 4

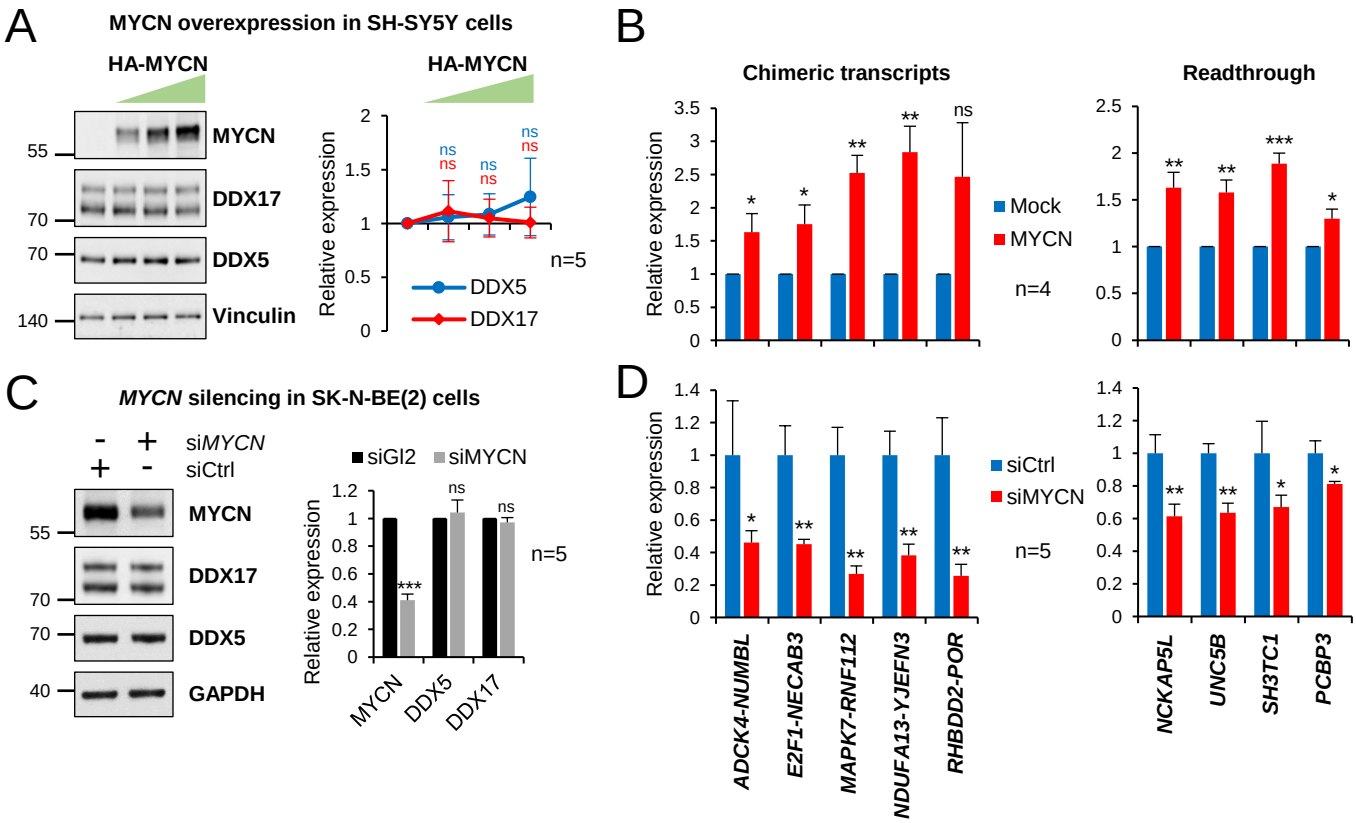
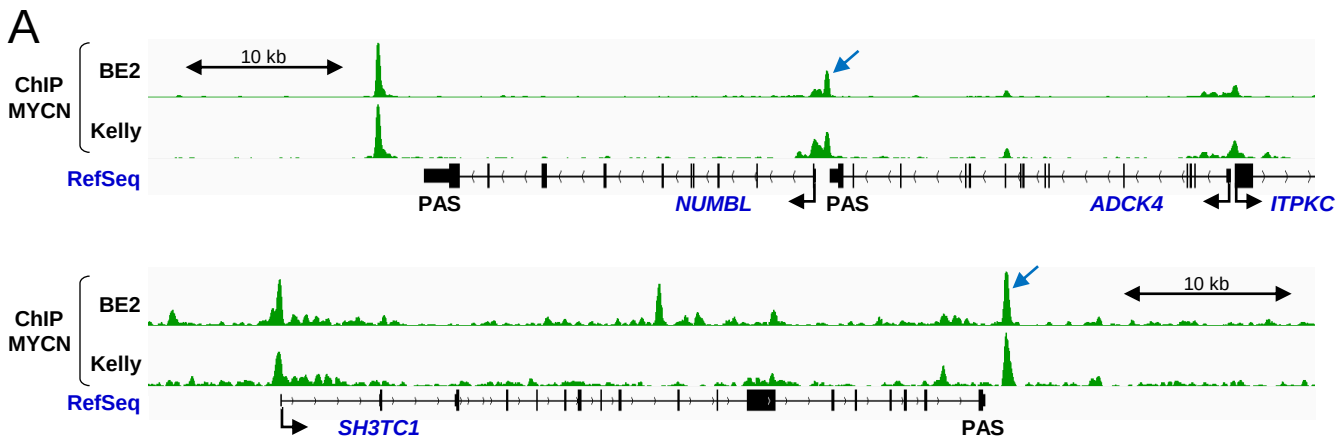
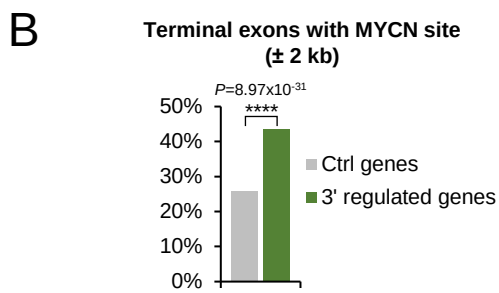
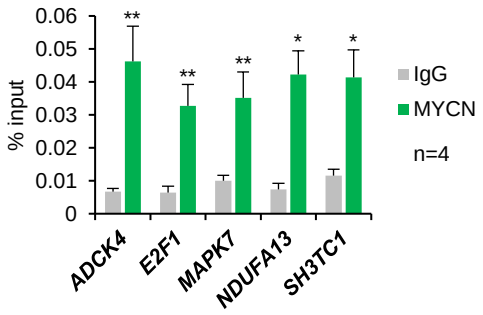
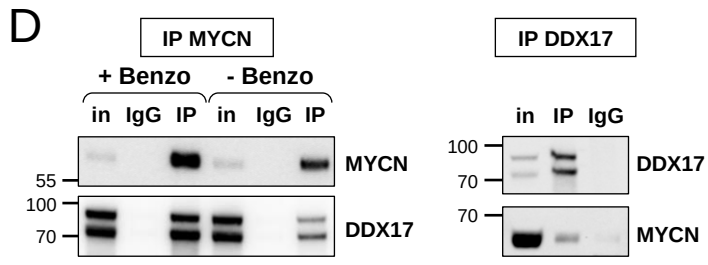

A
MYCN overexpression in SH-SY5Y cells
HA-MYCN
HA-MYCN
55
70
70
140
MYCN
DDX17
DDX5
Vinculin
Relative expression
ns
n=5
DDX5
DDX17
B
Chimeric transcripts
Readthrough
Relative expression
Mock
MYCN
n=4
C
MYCN silencing in SK-N-BE(2) cells
siMYCN
siCtrl
55
70
70
40
MYCN
DDX17
DDX5
GAPDH
siGl2
siMYCN
Relative expression
***
n=5
MYCN
DDX5
DDX17
D
Relative expression
siCtrl
siMYCN
n=5
ADCK4-NUMBL
E2F1-NECAB3
MAPK7-RNF112
NDUFA13-YJEFN3
RHBDD2-POR
NCKAP5L
UNC5B
SH3TC1
PCBP3

Figure 5


A
ChIP MYCN
BE2
Kelly
RefSeq
10 kb
PAS
NUMBL
PAS
ADCK4
ITPKC
ChIP MYCN
BE2
Kelly
RefSeq
10 kb
SH3TC1
PAS


B
Terminal exons with MYCN site
(± 2 kb)
P=8.97x10-31
****
50%
40%
30%
20%
10%
0%
Ctrl genes
3' regulated genes


C
% input
0.06
0.05
0.04
0.03
0.02
0.01
0
**
**
**
*
*
IgG
MYCN
n=4
ADCK4
E2F1
MAPK7
NDUFA13
SH3TC1


D
IP MYCN
+ Benzo
- Benzo
in IgG IP in IgG IP
55
100
70
MYCN
DDX17
IP DDX17
in IP IgG
100
70
70
DDX17
MYCN

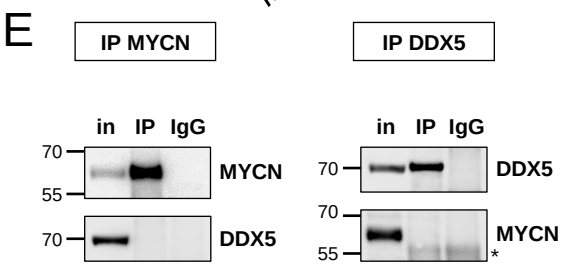
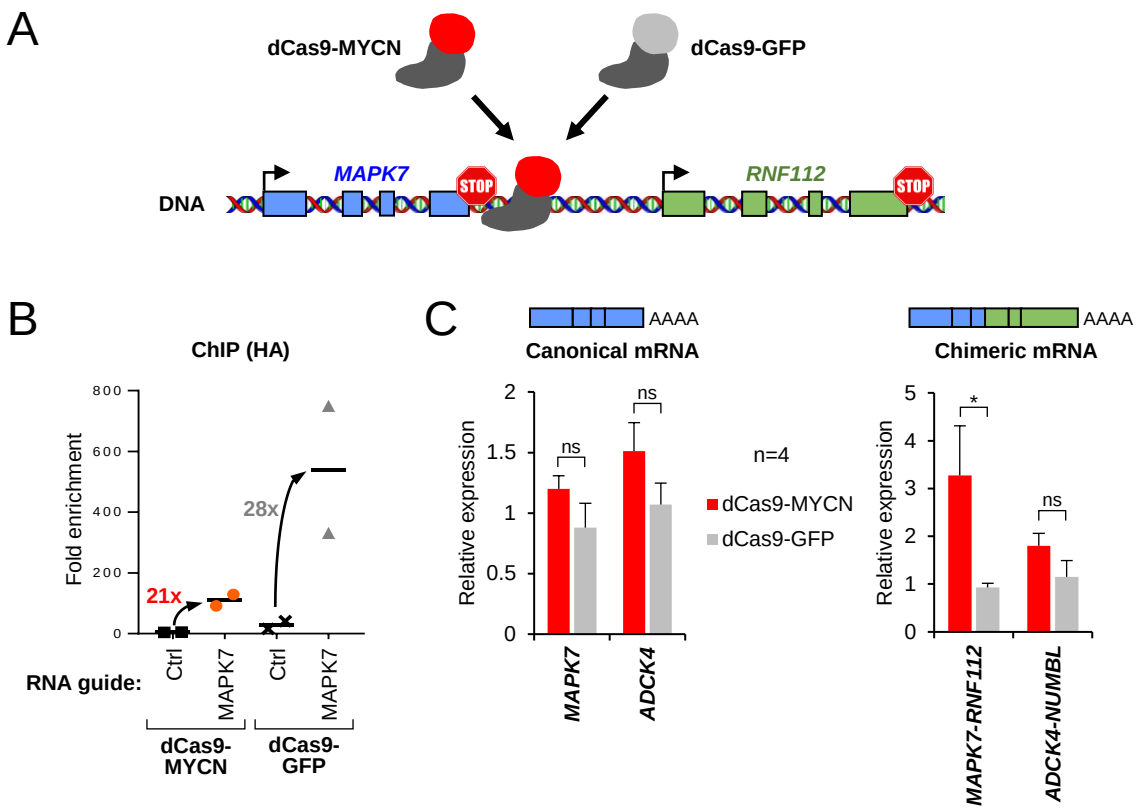

E
IP MYCN
in IP IgG
70
55
70
MYCN
DDX5
IP DDX5
in IP IgG
70
70
55
DDX5
MYCN
*

Figure 6


A
dCas9-MYCN
dCas9-GFP
DNA
MAPK7
STOP
RNF112
STOP
B
ChIP (HA)
Fold enrichment
800
600
400
200
0
21x
28x
RNA guide:
Ctrl
MAPK7
Ctrl
MAPK7
dCas9-MYCN
dCas9-GFP
C
AAAA
Canonical mRNA
Relative expression
2
1.5
1
0.5
0
ns
ns
n=4
dCas9-MYCN
dCas9-GFP
MAPK7
ADCK4
AAAA
Chimeric mRNA
Relative expression
5
4
3
2
1
0
*
ns
MAPK7-RNF112
ADCK4-NUMBL